

FOURTH EDITION

International Encyclopedia of Education

Globalization and the Shifting
Geopolitics of Education

VOLUME
1

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Fazal Rizvi
Jason Beech



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 1

**Globalization and the Shifting
Geopolitics of Education**

VOLUME 1 EDITORS

Fazal Rizvi and Jason Beech



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

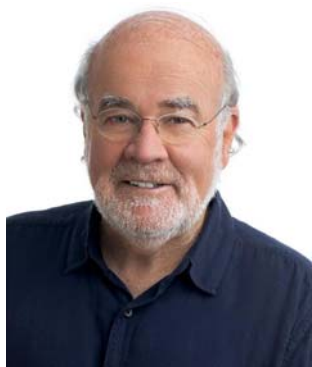
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



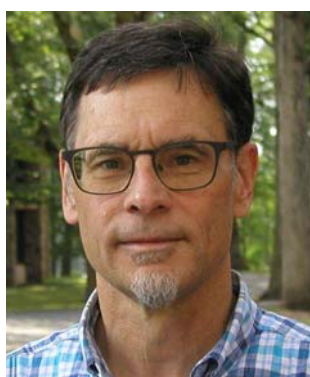
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

PREFACE

An encyclopedia has traditionally been defined as a compilation of academic work that either contains information on a particular branch of knowledge in the form of short entries or includes longer articles arranged by subjects or topics. It represents a body of knowledge useful to students, scholars, and researchers interested in an overview of key concepts and theories, and an introduction to the emerging issues and debates within a discipline. Although encyclopedias often seek to be comprehensive and objective, they seldom achieve this goal, not least because knowledge is invariably ordered from a particular perspective, against the backdrop of an understanding of the context in which it is produced. The access that the editors have to various epistemic traditions and scholarly communities is also not irrelevant. In this way, it needs to be acknowledged that articles contained in any encyclopedia represent a selective portrayal of a field of knowledge—diced and differentiated, ordered into segments and divisions, with some major topics often overlooked.

This was inevitably the case with the previous three editions of the *International Encyclopedia of Education* published by Elsevier. Looking back at their contents, the topics and themes included in those editions evidently reflected an understanding of the field of education that prevailed at the time. The articles drew on theoretical resources that were widely used then. The claims made in the articles were based on a range of assumptions about the nature of educational knowledge itself, and how it might be useful to scholars, policymakers, and educational practitioners. Among these was the assumption that the knowledge that is most valuable needed to be either general and universal or positioned within various national contexts. Little attention was paid to transnational crosscurrents within which educational knowledge is produced and utilized, despite the use of term ‘International’ in the title of the *Encyclopedia*. The idea of ‘International’ appeared more honorific than substantive, since much of the research reported emerged out of Anglo-American traditions, with the misleading assumption that this body of knowledge applied equally to all systems of education around the world.

In this, the fourth edition of the *Encyclopedia*, we have actively resisted this assumption and, in constructing the current volume, have, wherever possible, sought to problematize the meaning of the idea of ‘International’. We have recognized that the concept of ‘International’ is deployed in several different ways, has multiple meanings, and is highly contested. Furthermore, we have viewed internationalization as a process in knowledge-making, more like a journey than a destination. We have therefore regarded this edition of the *Encyclopedia* as a project in which an attempt is made to be mindful of the ways educational research, policy, and practice are shaped by both historical forces and emerging global conditions and are enacted in local contexts in a wide variety of different ways. We have acknowledged that the idea of internationalization itself has been used ideologically to promote various interests. In this way, the idea of internationalization has both empirical and normative meanings: it describes things as they supposedly are, but also prescribes how they ought to be.

These considerations have led us to approach the compilation of this edition of the encyclopedia from ‘an international perspective.’ What this means is that we have sought to ensure that the contents included in this edition reflect educational knowledge produced in all regions of the world. To so ‘deparochialize’ the information and analysis provided, as editors in chief, we have sought out the expertise of educational scholars drawn from every continent of the world, with knowledge of diverse cultural and epistemic traditions. We have thus tried to depart from the assumptions of both nation-centricity and universality, toward an ‘international’ perspective, not only by inviting exploration of issues, examples, and citations drawn from around the world but encouraging also the interrogation of the facts of transnational connectivity and exchange and how these are historically constituted and politically negotiated.

We do not believe, however, that this approach is based on a choice we have made but is rather demanded by the changing global conditions and their implications for education. The world of education has clearly been

transformed over the past two decades, driven by new developments in technology, resulting in: new modes of communication; the globalization of economy, politics, and culture; the changing nature of work and labor relations; the growing awareness of environmental crisis and climate change; the recognition of the ubiquity of cultural diversity and exchange and the legitimate claims of the colonized populations; an increasing level of distrust in traditional values and institutions and the rise of populist politics; and the growing interconnectedness of populations, practices, and policies through data collected from all aspects of life in artificial intelligence algorithms. These developments, among others, have transformed the space in which education now takes place, leading to the need to rethink the purposes of education and its governance.

The requirements of education knowledge and the ways of conducting and reporting research have been transformed by these significant shifts, highlighting the need to understand how educational experiences are shaped by locally and nationally specific conditions, and that the Anglo-American traditions of knowing can no longer be assumed to apply globally. Therefore, it is paramount to hear the voices of those who had previously been marginalized and to appreciate the local knowledges they produce, their priorities for education, the ways in which they wish to negotiate forces of globalization, and imagine their futures. We take these principles of inclusivity as fundamental to the normative concept of internationalization and have therefore sought to include, in this edition of the *Encyclopedia*, issues, themes, and debates that are global in their reach and significance, as interpreted by authors drawn from around the world.

Moreover, this edition of the *Encyclopedia* represents an effort to offer more than what might be derived from most Internet searches—in view especially of their likely unevenness or lack of trustworthiness in their treatment of topics. It is a compilation of chapters by section editors who in turn have mapped out subsections, topics, and authors to discuss key and timely issues being addressed in their fields of expertise. It is the section editors who approached the authors with request to represent the issues related to key topics and to do so in ways that animate the topic, especially going forward. Each of the chapters thus aims to provide background on important past developments largely as a way to call attention to contemporary issues that are being pursued or to represent new challenges and possibilities. The orientation to exploring issues as they are unfolding is a purposeful departure from more definitional approaches or from historical recantation. The emphasis upon internationalization begins to move us beyond what has been a form of western insularity. For aspiring scholars of education, we think the various chapters offer foundational deliberations intended to provide more lively overviews, and to not only generate understandings of where we have been but also where we are heading, and to do so in ways that spur questions and further inquiry.

This encyclopedia contains over 850 papers structured around 14 volumes on themes that we, as editors in chief, selected but which are thoughtfully developed, reviewed, and vetted by the content-specific editors for each volume, which in turn is divided in certain sections. The themes addressed within the volumes represent, as we have noted, a particular take on the currently existing body of knowledge in education. Inevitably some major themes are not included, and others only tangentially so. Nonetheless it is hoped that as a collection they introduce readers to a complex array of topics and issues, along with approaches to addressing them that have emerged over the past decade and remain poignant for thinking seriously and critically about education from an international perspective. In this way, our desire has been to complement and expand past considerations of developments in the field and to move forward beyond almost exclusively Anglo-centric representations that have predominated in past representations of the state of inquiry in the field, especially in annals in education such as highly cited journals, handbooks, and reviews.

We are pleased with the ways in which volume editors have helped us to move beyond an exclusive focus upon tropes within western scholarship. Their guidance has led to contributions from “northern” and “southern” educational scholars and to expanded discussions of developments in the field that extend to international examples and perspectives. So, for example, one of the volumes describes the shifting geopolitics in education and how they have posed new challenges for education. Another volume highlights issues of indigeneity and decolonization in education, addressing the reckoning between global and local understandings and issues of indigeneity, sovereignty, and domestication. Issues of how the curriculum is now contested across various cultural, political, and epistemic traditions is the theme of another volume.

In keeping with the expanded nature of studies of literacy and languages, a new comprehensive volume is included in this edition. The papers chronicle burgeoning forms and modes of languages and literacies as well as literature for young people, and the global contexts of their use. The myriad practices of language and the rich set of guiding theoretical and epistemological perspectives, from psychological to posthumanist, that guide inquiries are well represented here.

Just as many of the issues in language and literacy have become deeply political, the debates about science, technology, engineering, and mathematics (STEM) education, a focus on which reflects the shifting goals of education toward meeting the needs of the global economy. Another section involves a volume focused on the cultural politics of difference, and how this relates to issues of democracy and social justice. The politics of difference are also discussed in the volume on inclusive education and disability, which considers how issues of disability have traditionally been conceptualized and addressed in education, and why new approaches are now necessary. The focus here is on issues of access and participation in education through an interrogation of and call for inclusion—the mitigation or dismantling of barriers to access, participation, and success in education for all children and young people.

In the past two decades, issues of educational governance have been widely debated, especially with the growing embrace of the policies of marketization and privatization around the world. These have given rise to new techniques of educational administration and leadership. A volume in this encyclopedia throws critical light on these developments, exploring alternatives to the neoliberal imaginary. The issues of policy and governance are also central to a volume devoted to higher education and another focused upon teacher professionalism representing critical and global discussions on issues and trends related to teaching and teacher education that exposes the connections and disconnections with policy, and the impact of policy on teacher and teacher educator professionalism and practices.

Consistent with these aforementioned developments, the volume on cognition, learning, and human development, central aspects of all levels of education, has in recent years been researched from a wider variety of theoretical perspectives on learning, informed by sociocultural theories, as well as cognitive perspectives. It explores a broad array of topics (e.g., learning sciences, embodied cognition, social and emotional learning, learning analytics, and machine learning) covering themes of knowledge, epistemic cognition, cognitive load, executive function, curiosity and interest, self-efficacy, grit and resilience, academic self-concept, and emotion.

A discussion of the current developments in education is not possible without an examination of recent thinking in student assessment and research methodologies used in gaining insights about these developments. One of the volumes focuses on educational assessment, evaluation, and accountability. Since the last edition of the *Encyclopedia*, there has been an immense increase in the role of educational assessment in informing learning, policy, practice, and governance. Growth in expectations has been accompanied by expansion of methodologies used in educational assessment as well as challenges to its underpinnings. With these developments comes an awareness of a bidirectional relationship between assessment and social and cultural contexts, including its role in perpetuating social inequities.

And, finally, over the past decade, developments in use of computer technologies in daily life as well as education have transformed the ways it is possible to collect a vast amount of data. This has led to major innovations in research methods, both quantitative and qualitative. Two volumes in this encyclopedia explore these innovations, highlighting not only their technical aspects but also their efficacy and potential to assist attempts at educational reform. The volume on quantitative and measurement methodologies includes chapters focusing on foundational methodologies such as regression and reliability as well as machine learning and intelligent tutoring. The section on qualitative, multiple, and mixed methodologies includes philosophical and methodological aspects of these methodologies as was done in the previous edition. However, there are a significantly larger number of chapters on mixed and multiple methods.

This *Encyclopedia* is thus constituted by diversity in all its forms arising from recent developments, including the need to consider how our hegemonic approaches to educational thinking in the past had narrowed our perspectives, spaces, and voices and diminished the sovereignty and respect given nonmainstream groups. We hope the encyclopedia reinforces a shift to embracing transcultural matters in keeping with postcolonial matters and heralds the end of solely western ways of knowing in Education. We also hope the discussions advance a reconsideration of education and educational research in ways consistent with a fuller global embrace of the local. We believe that the global discussion of issues here reflects some of the wider sociocultural–political developments perspectives emerging in education.

Putting together this *Encyclopedia* has been an enormous task. It has taken us 3 years of hard work under the difficult conditions of COVID-19. Its completion would not have been possible without the collaborative support we have received from the volume editors, who volunteered to contribute not only because of their passion, commitment, and generosity, but also their sense of responsibility to the field of Education. And of course, we are enormously grateful to the rich complement of excellent authors, established and new voices in their respective fields, who have done their best to meet the deadlines and address editorial queries in a challenging time. We would like to offer a special note of thanks to the staff at Elsevier, and especially to

Paula Davies (Content Project Manager) who oversaw and ensured the integrity of the papers as well as managed the production processes for all of us. While there may be only a few brave souls who review the entire 4th edition of the *Encyclopedia*, we hope that the material will be valuable to a range of scholars in education, to students engaging with the many subfields, and to our various stakeholders. We offer this edition of the *Encyclopedia* as an exploration of issues represented in ways that transcend what we often encounter in the news or in social media, or even in our respective journals and other scholarly outlets, and to challenge conventional thinking in favor of vital ways of knowing and understanding our ever-changing landscape of Education.

Rob Tierney, Fazal Rizvi, and Kadriye Ercikan
Editors in Chief

Volume 1: Globalization and the Shifting Geopolitics of Education

Fazal Rizvi and Jason Beech

© 2023 Elsevier Ltd. All rights reserved.

Introduction

The papers in this volume of the 4th edition of the *International Encyclopedia of Education* are based on the premise that geopolitical shifts invariably affect the ways in which the purposes of education are conceptualized and its governance organized. This much is clear when we look at the major shifts that the world has experienced historically since formal educational systems were first established in the early 19th century. These systems were created then against the emerging demands of colonization, industrialization, and urbanization. They reflected the pursuit of modernity, with the use of industrial techniques to expand the populations that received formal education. While churches and other civic organizations had always promoted education for some sections of the community, from early 19th century the state began to play a dominant role in cultivating subjects that it felt the processes of modernization demanded.

The geopolitical shifts associated with colonialism deeply shaped the ways in which systems of education were organized, both in the colonizing and the colonized territories, and also in newly independent postcolonial states in the Americas. In this way, the spatial politics of colonialism largely created and fashioned modern systems of education, as well as the relationship between them in most parts of the world. The colonial powers, such as Britain and France, thus forged the character of education in the lands they occupied in ways that invariably served their interests, though in ways very different to each other and other colonial powers such as Spain and Portugal. Everywhere, however, education systems operated within the registers of the geopolitical imagination of the colonizers. Their curriculum and pedagogic approaches were designed mostly to produce subjects who were loyal to the colonial interests. Conversely, to ensure the legitimacy of colonial exploits, educational institutions in Europe had the task of providing their own local populations an understanding of the colonized world which portrayed the colonized people as inferior, who needed to be civilized.

As more colonies began to gain political independence after World War II, the geopolitics of the world shifted markedly. Yet most of the colonial arrangements persisted, despite attempts to cultivate new nationalist sensibilities (Fanon, 1967). This persistence was partly due to the failure of the decolonized states to imagine new ways of thinking about the nature of knowledge and the role that their own educational institutions could play in creating and transmitting a new appreciation of their place in the world. The lack of resources also led them to turn to the economically developed countries, their former colonizers, for development assistance to expand their systems of education, to create a knowledge and skills base they regarded as necessary to realize their nationalist aspirations.

The notion of development with its binary categorization into developed and underdeveloped country became a fundamental feature of the geopolitics of education and in the expansion of international organizations as key actors in promoting certain educational principles worldwide. The so-called developed nations often portrayed foreign aid as their moral responsibility. Strategically, however, such aid was always a way of extending their political influence and commercial interests, as the World Systems Theories (Wallerstein, 2004) demonstrated. This ideology of “developmentalism” (Escobar, 1995) also played an important role in the machinations of the Cold War, with education becoming aligned to the competing geopolitical interests. The Soviet Union and the United States competed for hearts and minds around the world in a cultural war that was fought through propaganda, cultural influence, and through education. The Soviet Union, for example, influenced the education systems of China and Cuba, while the United States had a strong impact on education in Brazil. The two superpowers also sought to extend their geopolitical influence through the scholarships they offered to students in the “developing countries,” supposedly to prepare them to meet the requirements of national economic development. Many scholarship students remained in these countries and now constitute a highly influential diasporic community.

Since the end of the Cold War, while this understanding of education and development persists to an extent, it is now tied to a view of geopolitics that has increasingly been shaped by modes of thinking associated with the ideologies of free markets and liberal democracy. This geopolitical shift is both driven by and associated with the discourses of globalization. The idea of globalization has of course been defined in a wide variety of ways (Held and McGrew, 2005), but its generalized definition highlights “the acceleration and compression of spatial and temporal relations” facilitated by the radically new developments in communication and transport technologies (Ray, 2007). It is expressed in and through the rapidly growing levels of mobility of people, money, ideas, and goods across nations and regions. This embedding of global flows has helped to cultivate active subjects whose interactions, intentions, and meaning now constitute global forms of sociality. It has created conditions that have reconstituted economic, political, and social institutions, including education, based on new policy settlements, articulated in various perceptions of global imperatives and opportunities.

Over the past three decades, these developments have encouraged various theoretical innovations underpinned by what has been referred to “the spatial turn” (Nieuwenhuis and Crouch, 2017), involving a new understanding of the relationship between geography and politics, encapsulated by the term “geopolitics.” While the idea of “geopolitics” has been widely used in the academic field of International Relations, a core argument of this volume is that it is also helpful in theorizing the ways in which

globalization has transformed the economic, political, and cultural institutions, relationalities, and practices, and how educational systems are now embedded within its evolving dynamics. It is possible therefore to speak of the geopolitical shifts in education, in attempts to understand how the various contours of education the ways are now invariably shaped by global forces, connections, and imaginations.

The concept of geopolitics, however, is a slippery and highly contested one: it refers both to a set of practices and institutional arrangements as well as various discursive constructions. At its core is the premise that geography plays an important role not only in determining the great power politics of International Relations but also in shaping the conduct of citizens, corporations, international bodies, social movements, governments, as well as institutions. The idea of geopolitics thus suggests that the connections between place, the state, and politics are affected by geographical arrangements, such as boundaries, coalitions, spatial networks, natural resources, and mobilities. These arrangements have the potential to redefine the ways in which political power is exercised, enforced, or undermined at global, regional, national, and local levels. In this way, an understanding of spatiality is helpful in making sense of global changes and in providing insights into the future behaviors of states, and their likely impact on individuals and institutions.

In the literature in International Relations, the concept of geopolitics has mostly been addressed through a realist lens (Dalby, 2013). The realist approach to geopolitics assumes that the relations between nation-states are largely anarchical since there is no world government capable of restricting their actions. Hence, self-interests often drive the exercise of power in International Relations, with nation-states viewed as primary actors. The core function of nation-states is to provide security and protect the domestic space and its citizens from the threat of the chaotic international sphere. In this way, realism presupposes a binary between the “inside” and the “outside” of the state. It also assumes the relationship between states to be inherently asymmetrical. It regards a stable global political space to be one in which chaos and anarchy can only be brought under a degree of control either through various forms of strategic agreements between nations or through the dominance of some nations over others.

The critics of this approach in the field of International Relations (for example, Sharp, 2009) argue that the realist understanding of geopolitics overstates the extent of conflict and competition in the world, and that the interstate system equally displays a capacity of both states and institutions to collaborate, negotiate international law, and work cooperatively through intergovernmental bodies such as the European Union. Without denying their importance, they insist moreover that nation-states are not the only actors in the configuration of geopolitics. Furthermore, the critical reading of geopolitics refuses to see the world as it supposedly is but highlights instead the need to examine the relationship between geography and politics as ideologically constructed, ‘imbued with social and cultural meaning’ (Dodds, 2019: 34). It regards the relationship between place and politics as always contingent, complex, and contextually determined. The critical approach thus insists “the everyday experiences of people and the strategies they have to adopt to in order to cope with the geopolitical and geoeconomic processes as fundamentally varied” (p. 36), subject to interpretations of the dynamic shifts in International Relations.

The papers in this volume are largely inspired by a critical reading of recent geopolitics shifts in education. They point to the ways in which the facts of global interconnectivity and interdependence have been debated and have reshaped our understanding of education, its purposes, as well as its modes of governance. They highlight how educational ideas and ideologies travel across national boundaries, broadening the scope of educational networks across various spatial configurations. This realization does not only give rise to new issues and challenges but also demands new ways of understanding educational policy processes and priorities, based on the assumption that the facts of global interconnectivity and their implications for education are never self-evident but can be imagined in a variety of different ways. Yet what has become abundantly clear is that in recent decades the idea of globalization has acquired a hegemonic meaning, linked to the ideological assumptions of neoliberalism. The neoliberal understanding of globalization involves a kind of political imagination that does not quite abandon the traditional norms of education but rearticulates them in commercial terms, commodifying educational outcomes.

This neoliberal imaginary highlights the benefits that can be derived from a global interconnected market economy, suggesting that education needs to create citizens who are well equipped to participate in economic activities that are increasingly knowledge based and globally extended. This understanding of economy has given rise to a wide variety of institutional practices, including organizational reforms such as globally stretched production, outsourcing, strategic alliances, clustering, and diversification and technological innovations especially in the areas of information, communication, and transport. It has promoted the formation of transnational networks to boost the production and distribution of goods and services, leading to the expansion of the movement of capital, goods, services, and people, and the uses of high information technologies, telecommunication networks, and intellectual capital to promote productivity in all fields of human endeavors. It has encouraged most aspects of education to be viewed in neoliberal terms (Rizvi and Lingard, 2010).

A range of explanations have been put forward to explain how market ideologies have become hegemonic. Some authors have pointed to the crisis of socialism on a global scale, rendering neoliberalism triumphant. At the same time, intergovernmental organizations (IGOs), global corporations, and other non-state actors have been highly influential. Since the early 1990s, the World Trade Organization, for example, has negotiated various rules to govern patterns of international trade, in goods and services alike to embed competition in most spheres of life, including education that has been recasted as a global commodity. The World Bank has become highly effective in persuading nation-states to adopt policies consistent with neoliberal precepts, while the OECD has created mechanisms to encourage global competition in education. Global corporations, such as Pearson and Microsoft, are no longer reluctant to steer national education policies in the direction of its preferred understanding of education. Other non-state actors such as foundations and think tanks are also major carriers in the global circulation of neoliberal sentiments. These shifts

have reconstituted the geopolitical space within which the nation-states now relate to each other, contributing to now only to massive changes in economic policy but also in setting educational priorities.

These developments have created conditions under which systems of education around the world have undergone major changes. The ways in which educational institutions are now governed have been radically transformed. The role of the state in funding education has diminished, with nonpublic sources, including student tuition, becoming dominant. The neoliberal sentiments have now begun to define approaches to teaching and learning, as well as approaches to assessment and accountability. Narratives that legitimate education have also shifted, as humanist ideals of human progress, salvation, equality, liberty, and access to truth have been marginalized at the same time as “the capacity of countries ... to compete in the global knowledge economy” (Schleicher, 2007: 9) is given precedence. At the same time, the number of students attending schools globally has grown rapidly and has more than doubled over the past three decades. This has transformed the demographic landscape of educational institutions, making diversity ubiquitous, giving rise to a new politics of difference that can no longer be ignored, and putting issues of inequalities at the center of educational debates.

These changes are historic and have seemingly become entrenched in our imagination regarding the purposes and governance of education. The papers in this volume attempt to understand this historical transformation, showing how educational changes have been promoted, managed, and contested and pointing to the major issues that systems of education now face, in various regions of the world. Of course, the issues and challenges examined in this volume do not claim to provide a complete and exhaustive account but are illustrative of how various geopolitical shifts are transforming educational landscape, such as those linked to the emergence of knowledge capitalism, privatization, the changing nature of work and labor relations, the mobility of students, and the like. Many issues that are equally if not more important are not included in this volume. This is so because some of these issues are addressed in other volumes of this encyclopedia, but also because some invitations were not taken up for a variety of perfectly understandable reasons, including the daunting effects of the COVID-19 pandemic.

However, beyond these practical limitations, it needs to be admitted that no account of the geopolitical shifts in education can ever be complete. The best that can be done is to highlight some of the key debates, albeit in an admittedly selective manner, to demonstrate how current developments in education take place within the context of the geopolitical shifts associated with a neoliberal imaginary of globalization. Within this context, the national systems of education are affected by the work of international organizations, both intergovernmental and nongovernmental, though in ways that vary and often contested. A theme common to many of the papers in this volume is that the contradictions and deeply harmful consequences of neoliberalism are now increasingly evident, and this provides a space in which the purposes and governance of education can be reimagined. The volume is structured in seven sections.

The first section shows how geopolitics shifts associated with globalization have challenged some of the traditional modes of theorizing educational purposes, priorities, and processes. As a result, new concepts, approaches, and methodological tools have been advanced in recent decades to apprehend the conditions emerging from the shifting geopolitics. The papers show how globalization has demanded new ways of thinking about comparative education, the world systems theory, and the notion of development in education and have investigated the potential of spatial theories, Southern Theory, as well as the idea of transnationalism. Also highlighted throughout these papers are the ways in which neoliberalism has transformed the social and ethical space of education, in ways that are both destructive and unsustainable.

The papers in the next section, “Global Norms and Challenges in Education,” discuss how recent geopolitical shifts have unsettled some of the traditional norms associated with education, such as equality, human rights and democracy, and peace education. The papers show how these issues now must be addressed in new registers, under conditions of global interconnectivity and interdependence. These conditions have also produced new economic, political, and economic challenges for education relating, for example, to migration and new diaspora formations. These changes have had profound impact on youth cultures around the world, demanding new approaches to curriculum and pedagogy.

Under the title “Policy Responses to the Shifting Landscape of Education,” the next set of papers considers how policy responses to these geopolitical shifts have been diverse, predicted on a contrasting range of assumptions relating to the nature of the good life and the ways in which education should be organized to promote it. Based on neoliberal assumptions, the dominant response has come in the form of what has been called the Global Educational Reform Movement. This has involved attempts to privatize education, with the corporatization of public education and the imposition of structural adjustment programs that restrict the investment of state funds on public education. The politics of foreign aid for education is increasingly aligned to the geopolitical interests of global markets and profit-making in and through education and can now be viewed as an exercise in public diplomacy. Global student mobility and the internationalization of education is often linked to the imperatives of the global knowledge economy and competition over skills. New communication technologies have also engendered in practices in online education, often as a way of expanding the reach of education, both through state-sponsored and commercially driven technology companies.

In the next section, “Policy Mobilities and the Global Politics in Education,” the discussion commences with the recognition of how recent changes in education display a remarkable drift toward what has been called “isomorphism” produced by patterns of policy transfer and borrowing across educational systems in the pluriverse. While this isomorphism does not imply global convergence of ideas and ideologies across the world in any absolute sense, it does indicate how the global forces are at play in their impacts on education, creating complex patterns of policy transfer, borrowing, and adaptation, creating possibilities of both competition and collaboration across national borders. Emerging also are new ways of organizing education at various levels of education.

Major differences in educational responses to the geopolitical shifts can be observed across regions and nations. The next section of this volume, “Global Imperatives and Regional Configurations,” includes a range of papers that explore the ways in which

globalization has impacted on educational policy and practice in various regions, including the Middle East and North Africa, sub-Saharan Africa, the various regions of Asia, and in the small island states in the Pacific and the Caribbean. What these papers demonstrate is the importance of examining how, in responding to the pressures of globalization, educational systems within the regions both collaborate and compete with other and are a major factor in shaping the geopolitics of education.

The next section in this volume examines the role of IGOs, such as the OECD, the World Bank, and UNESCO, in the formation and ongoing dynamics of the contemporary geopolitics of education. The papers in this section seek to explore of IGOs and influence national systems of education through the construction of assessment and accountability regimes, encouraging cooperation and competition, developing various conventions such as Bologna, and sharing information, as through various mechanisms of coercion, such as those involved in development assistance and demands of structural adjustment. Of course, these institutions do not promote the same ideological preferences for education, but often sponsor contrasting ideologies. UNESCO and UNICEF, for example, promote social democratic ideals, which the World Bank and the OECD view education in terms of human capital formation. Along with IGOs, nongovernmental organizations and for-profit corporate and philanthropic organizations have feature prominently in the processes of policy development and implementation in education.

The final section in this volume, "Emerging Conditions and Challenges in Education," points to the growing recognition of the complexity and contradictions of education spawned by the geopolitical shifts in education that have been witnessed over the past three decades. The papers point to the rapidly changing global conditions, relating, for example, to the greater recognition of the environmental crisis and climate change, inequalities, and the changing nature of work and communication; datafication and the emergence of new technologies of governance; the revival of nationalism; and anti-globalization sentiments (Rizvi et al., 2022). These developments have shaped educational politics itself, albeit in contradictory and conflicting ways. On the one hand, attempts to consolidate a neoliberal imaginary of education, through the relentless drive toward privatization, datafication of educational processes and the growing appeal of right-wing populism continue to hold power, while on the other hand, calls for the decolonization of knowledge and curriculum, and the reassertion of the educational significance of the values of equality, democracy and human rights strive for a different role of education in the construction of a more fair, sustainable and secure global order.

References

- Dalby, S., 2013. Realism and geopolitics. In: *The Ashgate Research Companion to Critical Geopolitics*, pp. 33–47.
- Dodds, K., 2019. *Geopolitics: A Very Short Introduction*. Oxford University Press, Oxford.
- Escobar, A., 1995. *Encountering Development: The Making and Unmaking of the Third World*. Princeton University Press, New Jersey.
- Fanon, F., 1967. *Black Skins, White Masks*. Grove Press, New York.
- Held, D., McGrew, A. (Eds.), 2005. *The Global Transformations Reader*, second ed. Policy.
- Nieuwenhuis, M., Crouch, D. (Eds.), 2017. *The Question of Space: Interrogating the Spatial Turn between Disciplines (Place, Memory, Affect)*. Rowman & Littlefield.
- Ray, L., 2007. *Globalization and Everyday Life*. Routledge.
- Rizvi, F., Lingard, 2010. *Globalizing Education Policy*. Routledge.
- Rizvi, F., Lingard, B., Rinne, R. (Eds.), 2022. *Reimagining Globalization and Education*. Routledge.
- Schleicher, A., 2007. "Foreword". In: Barber, M., Mourshed, M. (Eds.), *How the World's Best-Performing School Systems Come Out on Top*. London: McKinsey & Company.
- Sharp, J.P., 2009. *Critical Geopolitics*. In: *International Encyclopedia of Human Geography*. ScienceDirect, Oxford.
- Wallerstein, I., 2004. *World Systems Analysis: An Introduction*. Duke University Press.

CONTRIBUTORS TO VOLUME 1

Shirin Abdmolaei

*The University of Western Ontario, London, ON,
Canada*

Abdeljalil Akkari

University of Geneva, Genève, Switzerland

Richard Arnold

Loyola University Chicago, Chicago, IL, United States

Alejandro Artopoulos

*Universidad de San Andrés, Universidad de Buenos
Aires, Comisión de Investigaciones Científicas, Victoria,
Argentina*

Helen Avery

*Centre for Advanced Middle Eastern Studies, Lund
University, Lund, Sweden; and Department of
Languages, Linnaeus University, Växjö, Sweden*

Jim Baker

Education International, Brussels, Belgium

Stephen J Ball

*Sociology of Education, Institute of Education, University
College London, London, United Kingdom; Journal of
Education Policy, University of Copenhagen,
Copenhagen, Denmark; and University of Glasgow,
Glasgow, United Kingdom*

Annette Bamberger

*Hebrew University of Jerusalem,
Department of Sociology and Anthropology,
Jerusalem, Israel*

Kristyna Basna

*Institute of Sociology of the Czech Academy of Sciences,
Prague, Czech Republic*

Zvi Bekerman

*The Seymour Fox School of Education, The Hebrew
University of Jerusalem, Jerusalem, Israel*

Aaron Benavot

*School of Education, University at Albany-SUNY,
Albany, NY, United States*

Karen L Biraimah

*Department of Learning Sciences and Educational
Research, University of Central Florida, Orlando, FL,
United States*

Xavier Bonal

*Department of Sociology, Universitat Autònoma de
Barcelona, Catalonia, Spain*

Nigel OM Brissett

Clark University, Worcester, MA, United States

Birgit Brock-Utne

*University of Oslo, Oslo, Norway; University of Dar es
Salaam, Dar es Salaam, Tanzania; and EDCON, Oslo,
Norway*

Patricia Burch

*Center on Education, Policy, Equity and Governance,
Rossier School of Education, University of Southern
California (USC), Los Angeles, CA, United States*

Kiyong Byun

*Department of Education, Korea University, Seoul,
South Korea*

Yuzhuo Cai

*Faculty of Management and Business, Tampere
University, Tampere, Finland*

Niranjan Casinader

*Faculty of Education, Monash University, Melbourne,
VIC, Australia*

Zhe Chen

University of Rochester, Rochester, NY, United States

Dan Cohen

*Department of Geography and Planning, Queen's
University, Kingston, ON, Canada*

Robert Cowen

*University College London Institute of Education,
London, United Kingdom*

Aditi Desai
Center for International Teacher Education, Cape Peninsula University of Technology, Cape Town, South Africa

Victoria Desimoni
Loyola University Chicago, Chicago, IL, United States

Joyeeta Dey
University of Melbourne, Australia

Tenzin Dolma
O.P. Jindal Global University, Sonipat, Haryana, India

Caitlin Donnelley-Power
Loyola University Chicago, Chicago, IL, United States

Inés Dussel
DIE/Cinvestav, CDMX, Mexico

D Brent Edwards, Jr.
University of Hawai'i at Mānoa

Leona M English
St. Francis Xavier University, Antigonish, Nova Scotia, Canada

Óscar Espinoza
Facultad de Educación y Humanidades, Universidad de Tarapacá, Arica, Chile

Luca Favero
University of Turin (ESOMAS Department), Torino, Italy

Patricia Ferrante
FLACSO Argentina, Ciudad Autónoma de Buenos Aires, Argentina

Clara Fontdevila
School of Education, University of Glasgow, Glasgow, United Kingdom

Daniel Friedrich
Teachers College, Columbia University, New York, NY, United States

Byoung-gyu Gong
Sorenson Impact Center, David Eccles School of Business, University of Utah, Salt Lake City, UT, United States

Radhika Gorur
Deakin University, Australia

Andrew Greene
Teachers College, Columbia University, New York, NY, United States

Tom G Griffiths
Oslo Metropolitan University, Oslo, Norway

Jennifer Guevara
Dublin City University, Dublin, Ireland

Kalervo N Gulson
Sydney School of Education and Social Work, Faculty of Arts and Social Sciences, The University of Sydney, Camperdown, NSW, Australia

Eva Hartmann
Faculty of Education, University of Cambridge, Cambridge, United Kingdom

Mary Hayden
Mary Hayden Professor of International Education at the University of Bath, Bath, United Kingdom

Bawool Hong
Higher Education Policy Research Institute, Korea University, Seoul, South Korea

Martin Humburg
European Investment Bank, Luxembourg, Luxembourg

David Hursh
University of Rochester, Rochester, NY, United States

Jieyu Jiang
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, United States

Bronwen MA Jones
Institute of Education, University College London, London, United Kingdom

Agreement L Jotia
Department of Languages and Social Sciences, University of Botswana, Gaborone, Botswana

tavis d jules
Loyola University Chicago, Chicago, IL, United States

Sangeeta Kamat
International Education Program, College of Education, University of Massachusetts at Amherst, Amherst, MA, United States

Kerry J Kennedy
Department of Curriculum and Instruction, The Education University of Hong Kong, Tai Po, Hong Kong SAR, PR China The University of Johannesburg, Johannesburg, South Africa

Terri Kim
Comparative Higher Education, UEL, London, United Kingdom; UCL Institute of Education, London, United Kingdom; and St Antony's College, Oxford, United Kingdom

Sin Yee Koh
Institute of Asian Studies, Universiti Brunei Darussalam, Gadong, Bandar Seri Begawan, Brunei Darussalam

Ruthanne Kurth-Schai

Educational Studies, Macalester College, Saint Paul, MN, United States

Brianna A Kurtz

College of Education, Mary Baldwin University, Staunton, VA, United States

Jack T Lee

Moray House School of Education and Sport, University of Edinburgh, Edinburgh, United Kingdom

Hannu Lehti

University of Turku, INVEST Flagship Center, Faculty of Social Sciences, Turku, Finland

Steven Lewis

Senior Research Fellow/Senior Lecturer, Institute for Learning Sciences and Teacher Education, Australian Catholic University, Brisbane, QLD Australia; and REDI Centre, Deakin University, Burwood, VIC, Australia

Chris Lubieniski

School of Education, Indiana University – Bloomington, Bloomington, IN, United States

Josef Kuo-Hsun Ma

Department of Sociology, National Taipei University, New Taipei City, Taiwan

Suzanne Majhanovich

Faculty of Education, Western University, London, ON, Canada

Amar Mali

O.P. Jindal Global University, Sonipat, Haryana, India

Claire Maxwell

Department of Sociology, University of Copenhagen, Copenhagen, Denmark

Peter Mayo

University of Malta, Malta

Noel McGinn

School of Education, Harvard University, Swampscott, MA, United States

Simon McGrath

School of Education, University of Glasgow, Glasgow, United Kingdom

Marcia McKenzie

Melbourne Graduate School of Education, University of Melbourne, Melbourne, VIC, Australia

Juliana Mohok McLaughlin

Queensland University of Technology, Brisbane, QLD, Australia

Anja Mihr

Center on Governance Through Human Rights, Berlin, Germany

Marcella Milana

University of Verona, Department of Human Sciences, Verona, Italy

Greg William Misiaszek

Beijing Normal University, Faculty of Education, Institute of Educational Theories, China; and Paulo Freire Institute, UCLA, United States

Ricardo Montelongo

Sam Houston State University, Eleanor & Charles Garrett Teacher Education Center, Huntsville, TX, United States

Mauro C Moschetti

Autonomous University of Barcelona, Barcelona, Spain

Mousumi Mukherjee

O.P. Jindal Global University, Sonipat, Haryana, India

Diep Thi Bich Nguyen

School of Education, Deakin University, Burwood, VIC, Australia

Verity Norman-Tichawangana

International Education Program, College of Education, University of Massachusetts at Amherst, Amherst, MA, United States

Stephanie Oberheidt

ICF, Bruxelles, Belgium

Carlos Ornelas

Education and Communications at the Metropolitan Autonomous University, Mexico City, Mexico

Pushkarni Panchamukhi

R.V. University, Bengaluru, Karnataka, India

Helen Penn

UCL Institute of Education, London, United Kingdom; and University of East London, London, United Kingdom

Sahara Pradhan

International Education Program, College of Education, University of Massachusetts at Amherst, Amherst, MA, United States

Isaac Quist

Wesley College, Melbourne, VIC, Australia

Robyn Read

Independent Researcher, Toronto, ON, Canada

Goli M Rezai-Rashti

The University of Western Ontario, London, ON, Canada

Susan L Robertson
Faculty of University of Cambridge, Cambridge, United Kingdom

Emma Rowe
Faculty of Arts and Education, Deakin University, Burwood, VIC, Australia

Pasi Sahlberg
Southern Cross University, Lismore, NSW, Australia

Kenneth J Saltman
University of Illinois at Chicago, Chicago, IL, United States

Edda Sant
Manchester Metropolitan University, Manchester, United Kingdom

Yusuf Sayed
School of Education and Social Work, University of Sussex, Brighton, United Kingdom; and Center for International Teacher Education, Cape Peninsula University of Technology, Cape Town, South Africa

Sam Sellar
University of South Australia, Adelaide, SA, Australia

Gunjan Sharma
School of Education Studies, Dr. B. R. Ambedkar University Delhi, Delhi, New Delhi, India

Robin Shields
School of Education, University of Bristol, Bristol, United Kingdom

Iveta Silova
Mary Lou Fulton Teachers College, Arizona State University, Chandler, AZ, United States

Manuel Souto-Otero
School of Social Sciences, Cardiff University, Cardiff, United Kingdom

S Srinivasa Rao
Zakir Husain Center for Educational Studies (ZHCES), School of Social Sciences (SSS), Jawaharlal Nehru University (JNU), Delhi, India

Marcin Starnawski
University of Lower Silesia, Wrocław, Poland

Keita Takayama
Graduate School of Education, Kyoto University, Kyoto, Japan

Paul Tarc
Western University, London, ON, Canada

Ulrich Teichler
International Centre for Higher Education Research Kassel (INCHER-Kassel), University of Kassel, Kassel, Germany

Jeff Thompson
Jeff Thompson Emeritus Professor of Education at the University of Bath, Bath, United Kingdom

Matthew Thorpe
Manchester Metropolitan University, Manchester, United Kingdom

Jandhyala BG Tilak
Council for Social Development, New Delhi, Delhi, India

Ly Thi Tran
School of Education, Deakin University, Burwood, VIC, Australia

Elaine Unterhalter
Centre for Education and International Development (CEID), Institute of Education, University College London, London, United Kingdom

Mathias Urban
Dublin City University, Dublin, Ireland

Antoni Verger
Universitat Autònoma de Barcelona, Cerdanyola, Spain

Geoffrey Walford
Department of Education, University of Oxford, Oxford, United Kingdom; and Green Templeton College, University of Oxford, Oxford, United Kingdom

Anthony Welch
School of Education and Social Work, University of Sydney, Camperdown, NSW, Australia

Rui Yang
Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong

Miri Yemini
School of Education, Tel Aviv University, Tel Aviv, Israel

Ee-Seul Yoon
Faculty of Education, University of Manitoba, Winnipeg, MB, Canada

Adrián Zancajo
Department of Sociology, Universitat Autònoma de Barcelona, Catalonia, Spain; and University of Manchester, Manchester, United Kingdom

Kejin Zhu
Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong

CONTENTS OF VOLUME 1

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Preface</i>	<i>xix</i>
<i>Volume 1: Globalization and the Shifting Geopolitics of Education</i>	<i>xxiii</i>
<i>Contributors to Volume 1</i>	<i>xxvii</i>

THEORETICAL APPROACHES

Globalization and education <i>Steven Lewis</i>	1
Comparative education: then, now, and next <i>Robert Cowen</i>	11
World-systems analysis <i>Tom G Griffiths and Marcin Starnawski</i>	19
Spatial theories, methods and education policy <i>Kalervo N Gulson, Dan Cohen, Steven Lewis, Emma Rowe, Ee-Seul Yoon, and Chris Lubienski</i>	29
Southern theory as an educational project: Shinto, self-negation and comparative education <i>Keita Takayama</i>	37
Transnationalism <i>Niranjan Casinader</i>	45
Development and education <i>Simon McGrath</i>	52
The neoliberal dispositif: understanding the transformation of the social and ethical space of education <i>Bronwen MA Jones and Stephen J Ball</i>	60

GLOBAL NORMS AND CHALLENGES IN EDUCATION

Global human rights <i>Anja Mihr</i>	70
The light in the distance: global democracy and humanity's hope for the future <i>Ruthanne Kurth-Schai</i>	79
Peace education in conflict zones: general trends and Israeli particularities <i>Zvi Bekerman</i>	94
Enhancing a Global South-North partnership to promote global justice in K-12 classrooms: a Botswana case study <i>Agreement L Jotia, Karen L Biraimah, and Brianna A Kurtz</i>	104
Parental unemployment and children's educational outcomes – a Literature review <i>Hannu Lehti</i>	118
Diaspora and internationalization of higher education: a critical approach <i>Annette Bamberger</i>	128
Education of refugees <i>Helen Avery</i>	135
Online education <i>Ricardo Montelongo</i>	148
Digital learning divide <i>Josef Kuo-Hsun Ma</i>	157

POLICY RESPONSES TO THE SHIFTING LANDSCAPE OF EDUCATION

Global education reform movement revisited <i>Pasi Sahlberg</i>	165
The privatization of education. Drivers, social effects and regulatory challenges of private sector participation in schooling <i>Antoni Verger, Clara Fontdevila, and Adrián Zancajo</i>	174
Making the public private: the privatization of public education <i>David Hursh and Zhe Chen</i>	183
Shifting priorities of foreign aid to Pacific Island Countries: some considerations of new imperialism and geopolitics <i>Juliana Mohok McLaughlin</i>	193
Six decades of development: the impact of structural adjustment programs on education in developing countries <i>Óscar Espinoza and Noel McGinn</i>	205
Erasmus(+) student mobility: individual and institutional motivations and effects <i>Manuel Souto-Otero, Luca Favero, Kristyna Basna, Martin Humburg, and Stephanie Oberheidt</i>	218
Academic mobility: global trends and future prospects <i>Robin Shields and Jack T Lee</i>	230
Internationalization of higher education <i>Ulrich Teichler</i>	239

Global processes: global trade in international education and the international education-migration nexus	250
<i>Ly Thi Tran and Diep Thi Bich Nguyen</i>	

Teach for all and the visual regime of global teacher education reform	260
<i>Daniel Friedrich and Andrew Greene</i>	

POLICY MOBILITIES AND THE GLOBAL POLITICS IN EDUCATION

Policy isomorphism: a lens for understanding the influence of globalization on national education policymaking	271
<i>Yuzhuo Cai</i>	

Knowledge economy meets development imaginaries	280
<i>Alejandro Artopoulos</i>	

Redefining educational transfer and borrowing in the pluriverse	290
<i>Byoung-gyu Gong, Jieyu Jiang, and Iveta Silova</i>	

Global mobility and middle class families—parenting and education	302
<i>Claire Maxwell and Miri Yemini</i>	

Transnational academic collaboration: a critical analysis of the global and national contexts and constraints of co-authorship	308
<i>Terri Kim</i>	

South-South Cooperation in Education	317
<i>tavis d jules, Richard Arnold, and Victoria Desimoni</i>	

Global testing regimes: accountability, standardization, racial inequality and the rise of audit culture	327
<i>Goli M Rezai-Rashti and Shirin Abdmolaei</i>	

International schools: shifting sands and winds of change	336
<i>Mary Hayden and Jeff Thompson</i>	

International Baccalaureate: meanings, uses and tensions in a globalizing world	344
<i>Paul Tarc</i>	

Low-fee private schools	355
<i>Geoffrey Walford</i>	

False and distracting promises: early childhood education and care in low income countries	366
<i>Helen Penn</i>	

English as a medium of instruction and research in education: an international and comparative analysis	371
<i>Birgit Brock-Utne</i>	

GLOBAL IMPERATIVES AND REGIONAL CONFIGURATIONS

Impact of globalization on higher education in the Middle East and North Africa	380
<i>Abdeljalil Akkari</i>	

Challenges facing education in Sub-Saharan Africa	388
<i>Isaac Quist</i>	

Shifting policies of education in Latin América	398
<i>Carlos Ornelas</i>	

Asia-Pacific-realist and imaginary constructions <i>Kerry J Kennedy</i>	412
Education in South-East Asia. Much achieved, much still to do <i>Anthony Welch</i>	419
Globalization and political economy of education development in South Asia <i>Jandhyala BG Tilak and Pushkarni Panchamukhi</i>	425
Regionalization of higher education and geopolitics in Northeast Asia <i>Kiyong Byun and Bawool Hong</i>	447
Education in small island (and micro) states <i>tavis d jules, Victoria Desimoni, and Caitlin Donnelley-Power</i>	456

INTERNATIONAL ORGANIZATIONS

The work of international organizations: ten theoretical approaches and a way forward <i>D Brent Edwards, Jr. and Mauro C Moschetti</i>	465
The World Bank and education in a changing global arena <i>Xavier Bonal, Clara Fontdevila, and Adrián Zancajo</i>	480
UNICEF: self-transformation in a reconfigured global childhood space? <i>Jennifer Guevara and Mathias Urban</i>	488
European Union <i>Marcella Milana</i>	494
The World Trade Organization <i>Susan L Robertson</i>	503
UNESCO: promising education multilateralism? <i>Yusuf Sayed, Gunjan Sharma, and Aditi Desai</i>	510
Education for all and MDGs: global education policy translation in India <i>Mousumi Mukherjee, Amar Mali, and Tenzin Dolma</i>	526
The education Sustainable Development Goal 4: a critical appraisal <i>Nigel OM Brissett</i>	539
Labor markets, SDG4 and vocational education and training <i>Leona M English and Peter Mayo</i>	547
Global Education Monitoring Report <i>Robyn Read and Aaron Benavot</i>	553
NGOs & international development: ongoing trends and new architectures <i>Sahara Pradhan, Verity Norman-Tichawangana, and Sangeeta Kamat</i>	565
Education International and its global contribution to educational debates <i>Jim Baker</i>	574
Global education philanthropy <i>Kenneth J Saltman</i>	586
For profit corporations and education reform: global dimensions and impacts <i>Patricia Burch and S Srinivasa Rao</i>	591

EMERGING CHALLENGES IN EDUCATION

The geopolitics of climate change and education policy <i>Marcia McKenzie</i>	597
Ecopedagogy and ecopedagogical literacy <i>Greg William Misiaszek</i>	601
What works or what matters? Contested ideas of girls' education, gender equity and global connection <i>Elaine Unterhalter</i>	613
Global connective media: YouTube as an educational infrastructure <i>Inés Dussel and Patricia Ferrante</i>	622
Global competition over skills <i>Sin Yee Koh</i>	630
Public diplomacy and international higher education <i>Eva Hartmann</i>	640
Populism and education <i>Edda Sant</i>	648
The China challenge; China's challenges <i>Anthony Welch</i>	658
China's engagement with Southeast Asia in higher education <i>Kejin Zhu and Rui Yang</i>	666
Datafication <i>Matthew Thorpe and Sam Sellar</i>	673
Issues: neocolonialism and global languages <i>Suzanne Majhanovich</i>	682
The mangle of contemporary geopolitics <i>Radhika Gorur and Joyeeta Dey</i>	691

This page intentionally left blank

Globalization and education

Steven Lewis, Senior Research Fellow/Senior Lecturer, Institute for Learning Sciences and Teacher Education, Australian Catholic University, Brisbane, QLD Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Globalization as spaces and flows	2
Theorizing global spaces and relations	4
Scalar understandings of space	4
Topological understandings of space	5
Conclusion: the implications of globalization for education research	7
Acknowledgments	8
References	8

Introduction

In 2021, we find ourselves amid unprecedented uncertainty as COVID-19 continues to rage unchecked across the world. By its very definition, a global pandemic presents humanity with a crisis that is uniquely *global* in nature, regardless of one's geographic location or national allegiance. The extent of this shared experience is perhaps best expressed by the common responses countries have employed to limit virus transmission: the sudden halting of the mass circulation of people and products; the (re)erection of hard security borders along political boundaries to curtail freedom of movement; and the desperate scramble of national (and sometimes subnational) political leaders for ventilators and vaccines on the international market. Perhaps unlike any time in recent history, a significant majority of humanity is caught within an inescapable paradox, whereby global transportation, travel and trade bring both the risk of COVID-19 infection and, at the same time, the very means of overcoming the pandemic via the development and distribution of viable vaccines. Like Homer's Odysseus, we all now all caught somewhere between Scylla and Charybdis; that is, between a rock and a hard place, with neither option especially appealing.

Be it reliance on global supply chains or an aversion to international travel, a common – and necessary – concept to explain and hopefully resolve these diverse COVID-related issues is globalization. And yet, like many similarly nebulous terms (see here *neoliberalism*, *liberal*, *expertize*, *truth*, etc.), *globalization* has become something of an empty signifier, a catch-all term to diagnose and denote whatever feature of contemporary life is at hand, be it positive or negative or ambivalent. While ostensibly removed from the significance of global pandemics, I would note here that education is by no means immune to the processes and effects of globalization. Indeed, many of the more profound developments in education (and, for that matter, education research) have involved the global mobility of people, policies and products. Consider, for example, the now ubiquitous work of the Organization for Economic Cooperation and Development (OECD), including the numerous iterations of its Program for International Student Assessment (PISA) test; or the various local inflections of global discourses concerning standards-based reforms and standardized assessment in schooling systems; or the multiple instances where local schooling communities have sought to challenge and contest the imposition of solutions “from elsewhere”. In short, globalization is arguably alive and well in terms of informing how education is understood and practiced, even if a singularly agreed conceptual definition of *globalization* might well remain elusive.

It is clearly beyond the scope of any one chapter to adequately address every such possible definition, especially given that diverse epistemological and ontological possibilities will necessarily shape how one approaches globalization, both conceptually and empirically. That said, it is still worthwhile defining the contours of globalization theorizing, if only to acknowledge the complexities of globalization as both an empirical/theoretical construct and, relatedly, a site of study. To this end, [Axford \(2013\)](#) distinguishes between *globalization* as a process (i.e., how globalized connections are made and sustained) and *globalism* as a set of ideological frameworks that seek to endow globalization with certain values and meanings. For instance, [Steger \(2009\)](#) establishes three main forms of globalisms: (i) *market globalism*, associated with neoliberal notions of the free market; (ii) *justice globalism*, based on global solidarity and redistributive justice; and (iii) *religious globalism*, which serves to counter the forces of secularism and consumerism. Such diverse ideological underpinnings, comprising “patterned clusters of normatively imbued ideas and concepts” ([Steger and James, 2013, 19](#)), suggest, arguably, that the study of globalization *can never be* politically neutral, and a close reflexivity is therefore warranted when considering the theoretical legacies inherent in adopting different approaches, as well as how they then shape the (globalized) object of study.

These foundational definitions are further complicated if we also consider the related concept of *globality*; that is, how one practises and experiences the globalized world, in which “the totality of global flows, networks, interactions and connections (...) triggers a shift in the organization of human affairs and in ways of thinking about social relations and enacting them” ([Axford, 2014, n. p.](#)). Here, *globality as condition* transcends globalization as process and globalism as ideological assumption, and instead offers what [Axford \(2014, n. p.\)](#) describes as “a constitutive framework for action and a framework itself constituted through action and

consciousness". Put differently, globality seeks to undermine any *a priori* assumptions regarding a "global" entity, process or consequence, which also precludes essentializing any particular vantage point (geographical, social location, theoretical, ideological ...) from which to observe and study globalization. In a related manner, Connell's (2007) critique of "northern theories of globalization" notes the all-too-common absence of Southern and non-metropolitan social theory (and theorists), suggesting the future study of globalization needs to be far more global and inclusive, in both intent and substance. As Connell (2007, 379) pointedly remarks, there is very often "no one with a black face, no women, and no one from outside Europe", which might cause us all to reflect on how studies of globalization can ever be truly *global* under such circumstances.

Despite these caveats, and fully realizing the limitations of my efforts here, this chapter nonetheless seeks to explore and describe some of the major theoretical traditions that have informed thinking about globalization and, in turn, its implication for education policy and governance. I particularly situate an understanding of globalization within broader developments in the social sciences around spatial issues, as well as the various theoretical frameworks that have been mobilized to better understand *where* and *to whom* education policy is spatially and relationally located in a globalized setting. As Rizvi and Lingard (2010, 69) well note, "critical policy analysis in an era of globalization requires that we recognize the relationality and interconnectivity of policy developments". Put differently, a concern for globalization and education requires our attention to how different constellations of people, practices and places are brought together coherently; how policies, in a material and discursive sense, are mobilized to flow (or not flow) within and between these spaces; and, finally, the implications these movements have for how power relations are exercised. In short, globalization suggests that space needs to be employed "not simply as an object of concern but [also] as a conceptual tool for analysis" (Larsen and Beech, 2014, 201).

Deriving from this interest in spatial matters, I introduce here two theoretical approaches for studying globalization in education and the socio-spatial dynamics of policymaking and governance. These include (i) *scalar approaches*, or an understanding of space as nested multi-levelled scales, such as national, subnational and international (Papanastasiou, 2016; Fraser, 2010; Brenner 1999b, 2004; Jessop, 2002); and (ii) *topological approaches*, or an understanding of space as constituted through relations between social actors (Martin and Secor, 2014; Lury et al., 2012; Lewis et al., 2016; Decuyper and Simons, 2016). This is not to presume these frameworks are the only means by which to interrogate globalization and related spatial matters; however, they arguably provide two of the key (and most cited) approaches to apprehend global processes in education policy. I would also emphasize that both scalar and topological understandings of space foreground, importantly, the inherent relationship between space and power (Massey, 1993, 1999; Allen 2011, 2016; Allen and Cochrane, 2010). Attending in this way to the spatial dimensions of globalization helps bring to light the multiple complex processes at play in the constitution of "educational space", and the manner in which space is "deeply implicated in power, production and social relations" (Robertson, 2010, 15).

The structure of the chapter will proceed as follows. After first discussing globalization itself as an object of study, I then consider, in turn, scalar and topological theorizations of space, including their origin within the field of critical human geography and how each provides a distinct spatial lens through which to view globalized spaces and relations. I conclude the chapter with a discussion of how these scalar and topological approaches help us better apprehend the implications of globalization for undertaking education research, whereby space is not only – after Larsen and Beech (2014) – a matter for concern but also a central theoretical and analytical focus.

Globalization as spaces and flows

Although it has often been ambivalently defined within academic and popular literature, globalization is foremost concerned with how spaces and people beyond the traditional focus of the nation-state are situated in relation to one another. It is focused on "meanings of place and space ... and raised global connectivity" (Amin, 2002, 385), and how these have come to shape, and be recursively shaped by, a diverse array of social, economic and political exchanges. Different academic fields and traditions, drawing on different theoretical frameworks, will necessarily be driven to a variety of specific interests associated with globalization. Consider, for instance, how labor is spatially divided (Massey, 1993); how the world is imagined and the effects of this imagining (Appadurai, 1996); how the "Global North" (itself a problematic construction) has witnessed the hegemony of Americanized culture (Conrad, 2014); or how comparative testing regimes help constitute commensurate spaces of measurement and comparison (Meyer and Benavot, 2013; Lewis, 2020; Lingard, 2021a). But regardless of one's onto-epistemic orientation, a globalized world is connected and dynamic, in which world-level *organization* and *flows* bring forth competing sites and sources of political authority and social collectivity.

As noted previously, a concern for globalization requires our attention to how people, practices and places are brought together across countries and regions; how material and discursive flows are mobilized within and between these spaces; and how these movements shape the exercise of power relations. Although now well-known and oft-cited, the foundational definition of globalization by Held et al. (1999, 16) still warrants inclusion here:

[G]lobalization can be thought of as a process (or set of processes) which embodies a *transformation in the spatial organization* of social relations and transactions – assessed in terms of their extensity, intensity, velocity and impact – generating transcontinental or interregional *flows* and *networks* of activity, interaction, and the exercise of power. (emphasis added)

This highlights the processual nature of globalization as an ongoing and dynamic development that – while often sharing many features between “globalized spaces” – unfolds in highly contingent and context-specific ways. While there is no single experience of globalization that completely transcends individual cultures and histories, as these characteristics will always manifest in highly “vernacular” ways (Appadurai, 1996), the utility of such a definition is that the *spatial* and *relational* are explicitly emphasized. In short, it provides a lens for viewing a world of multi-lateral connectivity between spaces that are no longer reducible to the nation-state. Rather, we can see globalization making possible a “stretching and deepening of social relations and institutions” (Held, 1995, 20), whereby our activities are both shaped by and shape activities that might otherwise be considered geographically distant and distinct from one’s own.

Among the many implications of moving away from strict state-centrism is that all places, and their respective people and objects, are now “tied into at least thin networks of connections that stretch beyond each such place” (Sheller and Urry, 2006, 209), meaning no person or place can any longer be considered an island, entire of itself. Such wholesale changes to the spatial arrangement and connectivity of the world challenges, in turn, the primacy of political territory and geographic location as the sole determinant of “where” one is. To cite a well-worn example, so-called “global” cities – think here London, New York and Hong Kong – arguably exist in *multiple spatialities*. They are each at once part of broader national and subnational polities, and yet they also exist in direct relation to and alongside other similarly global metropolizes, seemingly disembodied from their broader territorial contexts through their advanced international transport, financial and cultural connections. It is not that one spatial configuration automatically needs to supersede another (for instance, London is not necessarily more “aligned” to New York than it is to the city of Manchester, or to the UK more broadly), but it does clearly introduce new complications to how place is spatially embedded. We might say that it occurs both topographically (i.e., *location*) and topologically (i.e., *relation*). Accompanying this shift in spatial belonging is the need to reimagine the ontological distinction between the local and elsewhere, in which we can no longer assume that “local happenings or geographies are ontologically separable from those ‘out there’” (Amin, 2002, 386).

However, it is equally important not to conflate efforts to complicate the ontology of global (“out there”) and local (“in here”) spaces with a view of globalization as somehow totalizing in nature, whereby the global overwhelms the local. Such a “determinist fallacy” (de Sousa Santos, 2006) wrongfully positions globalization as something spontaneous and irreversible, an *ex nihilo* process complete with its own inner logic that can be externally imposed “from above” upon local people and spaces. Rather, globalizing developments – such as Sahlberg’s (2011) *Global Educational Reform Movement* in education, or Ritzer’s (2013) *McDonaldization* in society more broadly – are always, to a greater or lesser degree, locally inflected by the historical and political specificities of a given locality or context. As de Sousa Santos (2006, 396) notes, “there is no originally global position; what we call globalization is always the successive globalization of a *particular localism*” (emphasis added). Such a position suggests there are no uniquely global spaces, processes or actors that exist entirely decoupled from their more local (be it subnational or national, as relevant) origins or contexts, as if they could be free-floating in some global ether. We can well observe many subsequently “global” policies as first being developed through either urban or regional initiatives. For instance, “Vancouverism”, as well as the conditional cash transfers (CCTs) of Mexico and participatory budgeting (PB) of Brazil, are clear examples of urban design policy taken up around the world *after* first originating in specific cities and regions (McCann and Ward, 2012; Peck and Theodore, 2015). In short, global policy does not need to solely emanate from notionally “global” intergovernmental organizations or agencies, like the OECD, to exert a global effect (see also McKenzie et al., 2021).

We can thus move away from a sense that globalization is somehow ontologically distinct from sub/national processes, and instead recognize the implicit relations that exist between the global and local. To take a specific education example, GERM may be regarded as a “globalized localism” (de Sousa Santos, 2006), or the progressive globalization of an approach to schooling accountability that first originated in the US and UK; that is, largely test-based, evidence-based, top-down and marketized (see Lingard and Lewis, 2016). The national origin of the policy advice proffered by the OECD, now arguably an influential policy actors in its own right, often explicitly references the policies and practices of PISA “poster children”, or schooling systems that have demonstrated high performance on the PISA survey, such as the much-fêted Singapore and Finland (Lewis, 2017). Moreover, the actors who inhabit national and subnational policy spaces are themselves frequently networked into whatever we might consider the global to be, with such “thought leaders” and policy entrepreneurs – or, in education, *edupreneurs* – often key nodes and facilitators in the development and dissemination of policy solutions (see also Lewis et al., 2020). Thus, even if global organizations are subsequently involved in promoting or revising initiatives that were originally local in origin (e.g., the World Bank promoting CCTs), these policy solutions cannot be considered to have simply dropped from the (global) sky.

Conversely, the manner that global(ized) policy is adopted by and informs local places (cities, states/provinces, countries) in material and discursive ways itself necessarily reflects the local and highly contextualized peculiarities of the local places in question. Such a process of mediation highlights what de Sousa Santos (2006) describes as a “localized globalism”, or the local inflection of a “globalism” that first emerged elsewhere (e.g., the Anglo-American approach to schooling accountability first emerging in the UK and US). Put differently, the local touching-down of “global” developments will always occur in “vernacular” (Appadurai, 1996) or “path-dependent” (Cohen, 2017; Tan and Yang, 2019; Takayama, 2012) ways that are mediated by local histories, politics and cultures. Given these more localized and context-driven enactments, it is more correct then, for instance, to recast the diverse expressions of Sahlberg’s (2011) GERM as *GERMs*, rather than otherwise suggest there is an essentialized version of global policy ensembles. And while politics and policy are never wholly local in nature, it is equally important to acknowledge that “neither do local spaces, nor local enactments of policy, ever entirely cease to retain their local-ness, regardless of the persistent (and increasing) presence and influence of global flows” (Lewis, 2021, 7).

We can thus see that “global” policy flows are still very much shaped by decidedly “local” conditions and contexts, which requires our attention to both dynamic movement and, simultaneously, the contextual embeddedness of so-called globalized policy, *vis-à-vis* the diverse people and places through which policy is developed, travels and is enacted. Rather than appearing from nowhere or else solely being imposed “from above”, there are considerably more nuanced processes at hand that risk being overlooked if globalization is instead reified as something always being done “here” from elsewhere.

Theorizing global spaces and relations

So far, we have defined globalization as something processual and contingent, rather than totalizing; and as deeply implicating local context through the development and importing of “global” policy solutions, rather than superimposed on unsuspecting recipients below. Taken together, it is foremost a concern with spaces and, importantly, how these spaces are held in relation to one another via the “stretching and deepening” of social connections. And yet, this seemingly simple focus is complicated not just by how one theorizes globalization as the constitution of relational spatiality, but also how one even more fundamentally understands what *social spaces* are. Despite the previous ontological primacy of the Westphalian nation-state, the acceleration of globalization and digital technologies in recent decades has problematized how the concept of space itself is understood, both ontologically (i.e., as spatial configuration) and epistemologically (i.e., as socially shaped category of practice). Critical geographers have long grappled with this problem, and the increased deployment of human geography within education research – or, the study of geographies of education (policy) (McKenzie et al., 2021) – provides an even greater impetus to explore how globalization in/of education is spatially enacted. There have been two main responses to these questions of spatial onto-epistemology, and how globalized social organization is arranged: (i) *scalar* approaches, and (ii) *topological* approaches. We will now consider each of them at length, respectively.

Scalar understandings of space

At its simplest, a scalar understanding of globalization recognizes different levels, or *scales*, of spatial organization and action, with subordinate scales (i.e., the local) nested within a series of progressively larger and notionally more complex spaces (i.e., the regional, national, international, transnational). The oft-quoted analogue of scalar space is the Russian *Matryoshka* doll, whereby carved dolls of decreasing sizes are placed within one another to represent the nation-state alongside other supranational and subnational political entities, which then aid or impede one another through the dynamics of inter-scalar contestation and/or cooperation (Jessop, 2002). Federal political systems provide a ready example of this inter-scalar jostling, especially in countries with a vertical fiscal imbalance between the revenue-raising powers of the federal government and the service-provision responsibilities of the states or provinces (e.g., Australia, Canada, Germany). We can also see such dynamics at play between the supranational European Union and its various member states and territories via the process of subsidiarity, which preserves responsibilities for certain policy areas for individual EU members and precludes “over-reach” by central authorities in Brussels.

Whichever scale is of particular focus, the common feature of all such scalar spaces is their definition via “geographically proximate links or as territorial units” (Amin, 2002, 386); that is, spatial organization is frequently determined by well-established political boundaries between or within countries that enclose continuous political units. This position does not assume that scalar spaces are somehow ontologically pregiven or stable, as territorial spaces and scales are always social constructions that reflect what Agnew (1999, 504) describes as the “historicity of spatiality”. Scalar thinking does, however, arguably foreground nation-states as fixed and mutually exclusive units of sovereign space, which are defined, more or less, by the territory within which they can exercise power (Agnew, 1994). As such, the nation-state effectively becomes a “container” of society that is centered upon the territory of the State, with smaller or larger organizational units arranged below or above such nationally defined territories. The presence of scalar logics, as well as the continued centring of the State (see Axford, 2014), is even implicit in the naming of fields such as “international relations” and “international development”, as these are defined by virtue of their activities occurring *beyond* the bounded scale and political territoriality of national space.

Attempts to view globalization through scalar lenses have been especially prominent within the field of political economy, especially regarding the increasing complexity of inter-territorial relationships and the resulting proliferation of spatial scales outside the strictures of the nation-state (see Jessop, 2002). One need only consider the rise of the “global city” (e.g., London, New York, Hong Kong), supranational political institutions (e.g., the European Union) or the increasing role of intergovernmental organizations in processes of global governance (e.g., the OECD, United Nations and World Bank) to observe how national territory has conceivably become but one spatial scale – and site of political authority – among many (for instance, see Brenner, 2004; Jessop, 2008; Sassen, 2006; Wallner et al., 2020). Importantly, the presence of multiple scales and globalized activity does not preclude the continuing relevance of the nation-state, nor the “simple replacement of national scales by a global scale of action” (Amin, 2002, 387). It does, however, acknowledge that the nation-state has been de-centred and can no longer be considered the “sealed container” of socio-economic activity, leading to what Brenner (1999b) has described as the “rescaling of territoriality”. In short, these attempts recognize the reconstitution of spatial scales and, in turn, the growing importance of subnational and supranational forms of political organization (Brenner, 1999a).

Such a scalar understanding of globalization also posits that “non-global” spaces (i.e., the national and subnational) are frequently the very sites where the global is practiced and where its effects are most apparent. For instance, Brenner (1999b) makes the useful distinction between, on one hand, the significant role of the *nation-state* as a site for the reterritorialization of global capital and, on the other, the reduced importance of the *national scale* as the means for delineating and organizing socio-economic relations. Put differently, the nation-state still clearly retains significance, but this role does not need to be solely exercised through nationally scaled institutions; that is, the organization and effects of the State are not automatically coterminous with the national scale of the State. The former ontological privileging of the national is thus overcome through scalar “relativization”, in which State power is rescaled downwards via devolution to subnational scales – turned “inside-out” – and, at the same time, rescaled upwards toward supranational institutions of regulation, or turned “outside-in” (Brenner, 1999b, 442). Relativization thus captures not only that the national scale is no longer preeminent, but that “no spatial scale is currently privileged” (Jessop, 1998, 90; emphasis added). This accommodates the so-called *re-scaling* of statehood (Brenner, 2004), whereby political authority and organization is relocated within new scalar dimensions.

More recent thinking around scalar space has sought to move further away from ontological understandings of scale, which has followed prior critiques of scale as a category of analysis (Robertson et al., 2002; Brenner, 2004). Such work has instead emphasized that scale should be understood more as an epistemological concept (see Papanastasiou, 2017; Savage et al., 2021), with scalar space seen as constituted through socially constructed relations. While this position aligns with earlier work that emphasized the historical nature of scalar spaces (see Agnew, 1999), it also reflects that “scale” arguably provides a useful lens through which to view and understand the world, both in respect of social science research concerned with globalization and, more mundanely, how individuals seek to imagine and construct their understanding of the world. Indeed, Papanastasiou (2016) has argued that scale is useful not because vertical scalar hierarchies are a pre-given feature of the globalized social world, but rather because they are (or, at least, have been) central to how many policy actors *imagine* a globalized social world to be one that comprises “local”, “national” and “international” scales. Attending to this active constitution of scale by policy actors – otherwise described as *scale-craft* (Papanastasiou, 2017) – is arguably useful to help reveal the political practices of policymaking, especially given that “scale is a central category used by policy actors to imagine and assemble political spaces and reforms” (Savage et al., 2021, 2; see also McKenzie and Aikens, 2021). A related consideration to this approach is concerned with recognizing the political and social processes of “territorialization”, where territory is seen as purposefully constructed to achieve certain ends (e.g., national policy consensus for federal systems), rather than construing territories as necessarily bounded and static (Lewis, 2021).

It should be noted, however, that this more epistemological notion of scale, even while accounting for the influence of spaces outside and within the nation, is still arguably overlaid upon a preconceived territorial spatiality and logic. Indeed, the scales on which globalization unfolds, even if only the imaginings of policy actors, “simultaneously bound social relations within determinate geographical arenas” (Brenner, 1999b, 447), effectively creating a series of hierarchized places and territories made coherent through complex inter-connections. More traditional notions of the nation-state, territory and borders are thereby retained, which arguably does little to move beyond a strict interior (“in here”) and exterior (“out there”) binary of territorialized socio-spatial relations. Amin (2002) uses the useful example of expatriate or diaspora populations to highlight the inability of scalar thinking to overcome this lingering local-global knot. Specifically, do these cosmopolitan populations belong to their new countries of work and residence, or are intimate connections still retained with the local spaces from whence they moved? And, relatedly, which scale(s) do they now occupy: local, or national, or transnational? And, if there are multiple scales, which one is most influential or present? While not negating the utility of a scalar approach, I would simply argue there are limits to its explanatory powers when considering “overlapping near-far relations and organizational connections that are not reducible to scalar spaces” (Amin, 2002, 386). As such, I will now consider another key effort to theorize spaces of globalization that attempts to address these more relational, non-territorial issues: *topological approaches*.

Topological understandings of space

Topology was initially envisaged as a strictly mathematical heuristic, with the first steps in its development taken by Leonhard Euler in 1736 as a means of solving the Seven Bridges of Königsberg (present-day Kaliningrad) problem; that is, whether a pedestrian could walk through the city and cross all seven bridges only once (Shields, 2012). The concept has since exceeded its earlier perambulatory confines to emerge, in recent decades, as an increasingly influential conceptual device in social and cultural theory, and across disciplines as diverse as political science, economics, human geography and, increasingly, education. Despite its broad application, the central principle of topology is the representation of space as emergent instead of absolute, moving beyond Euclidean and Cartesian geometries “to introduce a new spatial thinking that identifies lines of relation, rather than discontinuous points and lines” (Mezzadra and Neilson, 2012, 60). In short, a topological space is a space of flows that is defined by relations *between* points, instead of only considering the spatial location of points. Thus, topological understandings of globalization are post-Euclidean and post-territorial spatial theories (Martin and Secor, 2014), at least in contrast to more territorially-oriented scalar spaces (i.e., subnational, national, supranational ...).

Consider, for example, the space represented by the iconic map of the London Underground, in which different train stations (in topological parlance, *nodes*) are connected by a series of train lines, or *arcs* (see <https://tfl.gov.uk/maps/track/tube>). Upon inspection, it is readily apparent that the location of stations on the map *does not* correspond with the exact physical location of stations across the city of London, and neither are the distances or directions between stations on the map at scale with the

actual train lines. And yet, the map is still exceptionally useful for those wishing to travel around London using a transport system that is significantly underground and which, in turn, denies commuters the possible benefit of above-ground landmarks as a visual reference. The map is perhaps most useful because, topologically speaking, it eschews any precise alignment with physical location and focuses instead on the connections (i.e., train lines) *between* points (i.e., train stations). What we then have is a map whose purpose is not to accurately represent physical features of the space in question (e.g., location, elevation, distance), but which instead helps us apprehend how this space is made to relationally cohere through social practices (e.g., designing and using a functional public transport network). As such, the creation and use of such a map requires a decidedly topological approach to space, in which relational continuities are considered *in addition to* physical location alone (Lewis et al., 2016; Lury et al., 2012).

Topological imaginings of space have gained a particular traction, perhaps unsurprisingly, in studies of human geography, which is concerned primarily with the relational issues of how people organize, move and interact. Such thinking has sought to approach space *relationally*; that is, as “co-constituted, folded together, produced through practices, situated, multiple and mobile” (Amin, 2002, 389), and thus not reducible to prefigured scalar spaces. A key premise of topological understandings of globalization is that space is no longer a passive backdrop upon which actions unfold, but instead it *actively unfolds* and is constituted through actions and relations. Moreover, spaces once notionally distant and distinct from the local are now increasingly relevant and, perhaps more significantly, *present*. The current COVID-19 pandemic provides a clear example of how once-distant global spaces are now readily enfolded within and made near to the local by virtue of relational practices. Consider, for instance, how international travel and global supply chains, or the ubiquitous Zoom meetings that characterize working remotely, can technologically compress space-time (Harvey, 1989), or *timespace* (May and Thrift, 2003), and bring otherwise distant spaces together through relations. On the other hand, the initial onset of the pandemic from a specific location clearly reflects the ability of localities to shape more globally oriented processes and spatio-social dynamics, which can then affect other relationally connected local spaces. This suggests that the local very much retains its significance, rather than globalization leading to a wholesale effacing of difference into an amorphous and dematerialized global space of flows. It is not that space no longer matters in a globalized world, but rather that there are now so many spaces that *do matter*. Put simply, topological understandings of globalization signal the “rise of new spatio-temporalities affecting what goes on in place” (Amin, 2002, 392).

To consider globalization through the lens afforded by topology is then to apprehend spaces as emergent: as *a posteriori*, rather than *a priori* (Lash, 2012). These insights have particular importance for helping to overcome the ontological distinction between the local and global. Rather than territorial borders determining what is near and far via the mediating presence of the nation-state, it is the “mutable quality of relations” (Harvey, 2012, 78) that most determines whether something is proximate or distant. In other words, actors, organizations or places are rendered near if they can exert influence through their relations with other actors, organizations or places (Allen, 2011; Lata and Minca, 2016). This forces a rethinking of globalized space to not only involve scalar layering, but also how the folding of topological space can connect previously distant points to enable placement and proximity in multiple spaces. As Lury et al. (2012, 13) note, globalization creates a topological spatial continuum of the *in-between* that “not only enacts the scalar entities of the “local”, the “national” and the “global”, but also puts them in multiple relations to each other”. This enables an exploration of global and local places and processes, but without having to explicitly ground such endeavors within external territorial framings or “parametrization” (see Michael and Rosengarten, 2012).

To draw on a now commonplace example from education, the OECD’s PISA has contributed toward creating a globally commensurate space of testing and comparison (Lingard et al., 2016), whereby each participating (national or subnational) schooling system can be known and evaluated using the same PISA-derived metrics. Given this universal measure of “effectiveness”, high-performing PISA “poster children” (e.g., Singapore and Finland), as well as the OECD’s own educational policy work, can now be readily enfolded within more local schooling spaces, discursively shaping what counts as successful policies and practices (Grek, 2009; Sellar and Lingard, 2014; Lewis, 2020). Within such a topological space, it then matters less than these countries and schooling systems might be geographically distant from one’s own. What matters is that these schooling spaces can be rendered near in a relational sense via spatial “reach” (e.g., schooling systems comparing themselves with Singapore), thereby establishing a direct presence between actors, organizations and localities. The distance separating the local and global are then overcome by the relational enfolding of the global *into* the local; as well as by local people and places “reaching out” to global discourses and organizations, in order to inform and shape their own local politics and policymaking (Allen, 2016; Allen and Cochrane, 2010). By this reasoning, it can be argued that governing or steering “at a distance” (Kickert, 1995; Rose, 1999) is something of a misnomer, an artifact from a strictly Euclidean mindset in which power has defined coordinates of both origin and extension. Instead, the topological folding of space brings relations of influence into direct contact with those upon whom it acts, assembling new geographies of power and, in turn, possibilities for action.

Such an emphasis on topological relations should not, however, be entirely at the expense of considering physical location or territoriality. Although topology does admittedly draw attention to the numerous spaces, overlapping relations and organizational connections that shape the materiality of everyday life (Urry, 2007; Amin, 2002), and despite globalization being associated with a prevailing “ontology of movement” (Smith, 2003), this is not to say that topological space entirely discards connection to fixed places or locations. Indeed, topological spaces are invariably *woven into* Euclidean spaces to produce what Smith (2003, 569) describes as “a world of transitory hardenings and fluids”; or, put differently, a world of (territorial) frictions and (topological) flows (Lewis, 2021). As Hartong and Piattoeva (2019) suggest, topological space-making can perhaps not occur without an underpinning topographical frame of reference. Returning to the example of the London Underground, the

utility of the topological map is ultimately predicated on someone needing to get to an *actual* physical place. Similarly, the topological spaces of comparison enabled by PISA tests are meaningful only to the extent that sub/national territorial jurisdictions and political authorities can draw on such data to mobilize and justify their own local policy initiatives. Nevertheless, topological renderings of space would assert that the character of these jurisdictions and spaces, even if outwardly territorial in appearance, are no longer determined *solely* by recourse to territorial characteristics. Rather than seeing the local as exclusively embodying a politics of *place*, which only reflects local practices, topological accounts of globalization instead gesture toward a politics *in place* (see also [Amin, 2002](#)); that is, a localization of different spaces folded together through their relations and inflected through their various contexts.

Conclusion: the implications of globalization for education research

A key purpose of this chapter has been to explore how globalization has complicated social organization and relations, meaning that new spatial lenses (including scalar and topological approaches) are useful, and perhaps even necessary, if we are to apprehend the ongoing processes of relational space-making. In this concluding section, I would like to now briefly consider how these spatial frameworks are being productively applied to better understand the complexities of education policy and governance in a time of globalization. Put simply, education policies are now being realized in relational spaces not defined solely by the territorial boundaries of the nation-state ([Ball, 2016](#); [Lewis, 2020](#)); are developed by actors and organisations not necessarily situated in government bureaucracies and traditional sites of policy formation ([Gulson and Sellar, 2019](#); [Decuyper et al., 2021](#)); and, judging by their global reach, transfer and take-up, are in constant states of movement, mutation and local adaptation ([McKenzie and Aikens, 2021](#); [Lewis, 2021](#)). Not only has the ontological primacy of the nation-state been challenged through expanding sites and locations of political authority, but the socio-spatial processes of policy production have also brought new policy spaces into focus as sites of policy making and intervention. Accommodating these empirical changes, I would like to sketch an outline of how – and, perhaps most importantly, *where* – globalization and education research are now intersecting by addressing two points in particular: i) the danger of *theoretical churn*; and ii) the rise of *interdisciplinary approaches*.

As we have seen, the need to accommodate spaces, relations and processes associated with globalization is now increasingly apparent. These new frameworks of scalar and topological spatialities are both empirical and theoretical, insofar as contemporary social life is marked by the growth of globalized spaces and cultural practices that, in turn, provides fertile ground for deploying space as an analytical tool, both in everyday life and in social and cultural theory. Spatiality at once provides a new conceptual vocabulary (i.e., the analytical), while at the same time indicating new, or at least newly explored, empirical phenomenon (i.e., the descriptive). But despite the utility of scalar and topological lenses, or whatever else may subsequently follow, I would argue that it is equally important to not prematurely jettison theoretical models or frameworks in our haste to adopt new approaches. As former advocates of scale, [Jessop et al. \(2008\)](#) have since cautioned against what they describe as the “unflexive churning” of spatial turns, including the rush to privilege any particular dimension of socio-spatial relations (e.g., territory, scale, topology) as the essentialized feature of globalized space. To that end, I have sought here to present scalar and topological approaches as two separate but related attempts to reconcile the emergence of new globalized spaces. And while I personally find much to commend a topological spatiality, I have reserved casting some final judgment in this chapter, not to attempt an objective ambivalence but rather to allow researchers interested in these approaches to determine utility for themselves.

I would also hesitate to assume the current delineation of these approaches, or any such others, is the final statement or endpoint of their development. For instance, the rising relevance of spaces beyond the nation-state, especially following the end of the Cold War, led policy scholars to challenge the presumptive homology between the nation and the social, a situation that has been described pejoratively in research as “methodological nationalism” ([Beck, 2000](#)). While questioning the prerogative of national bureaucracy as the “natural” home of policymaking and political authority has undoubtedly been useful, there is also a clear danger of the pendulum swinging too far in the opposite direction; that is, where the (territorial) space of the nation is overlooked at the expense of attending to global spaces and flows. Indeed, [Clarke \(2019\)](#) and [Lingard \(2021b\)](#) argue we are now at risk of “methodological globalism”, or an inflexible attention to global(ized) socio-spatialities, organizations and actors that downplays more national and local policy processes. This recognition of the continuing relevance of the nation-state, in addition to the presence of new scalar and non-scalar (topological) spaces, should encourage a deep and considered engagement with these spatial lenses, and a commitment to observe where and how globalization is occurring, but without recourse to a presumed spatial coordination.

I would also like to say something here about the importance of interdisciplinary approaches to studies of education and globalization. There has been increasing scholarly awareness and attention in policy sociology and cognate fields, such as policy mobilities and geographies of education (policy), regarding how one should research education in the context of globalization if spatial matters are to be emphasized. Indeed, [Ozga's \(2021, 298\)](#) useful categorization of policy as “made and remade in process; as multi-sited, fluid and unstable; and as open to contestation” suggests that the spatial imagination of the policy researcher needs to similarly be attuned to the globalized and dynamic spaces of education policy processes. This has resulted in numerous theoretical and methodological approaches being adopted and adapted into education research to accommodate

these new spatial realities, including *policy mobilities* (Gulson and Symes, 2007; Lewis, 2021; McKenzie, 2017), *policy assemblage* (Savage, 2020; Thompson et al., 2021), *social topology* (Decuyper, 2021; Decuyper and Lewis, 2021; Lewis et al., 2016), *scalecraft* (Savage et al., 2021) and *digital data infrastructures* (Decuyper, 2019; Lewis and Hartong, 2021; Williamson, 2017). While remaining alert to the dangers of theoretical churn noted above, the complexity of contemporary policy processes has encouraged, and likely even requires, a further broadening of our conceptual and methodological horizons if we are to adequately foreground spatial issues in education research.

By way of closing, I return to the concise definition of globalization offered by Amin (2002, 392); that is, as the “rise of new spatio-temporalities affecting what goes on in place”. For all the complexities that globalization entails and for all the uncertainties it introduces, a central consideration for education research is the need to address these new globalized spaces, as well as the always-emerging relations and effects they make possible. I hope this discussion has contributed in some small way to encourage work with and the adoption of spatial approaches to understanding globalization, and thus fosters the further development of these ideas in education research.

Acknowledgments

This work was supported by the Australian Research Council [grant number DE190101141].

References

- Agnew, J., 1994. The territorial trap: the geographical assumptions of international relations theory. *Rev. Int. Polit. Econ.* 1 (1), 53–80.
- Agnew, J., 1999. Mapping political power beyond state boundaries: territory, identity and movement in world politics. *Millenn. J. Int. Stud.* 28 (3), 499–521. <https://doi.org/10.1177/03058298990280030701>.
- Allen, J., 2011. Topological twists: power's shifting geographies. *Dialogues Hum. Geogr.* 1 (3), 283–298. <https://doi.org/10.1177/2043820611421546>.
- Allen, J., 2016. *Topologies of Power: Beyond Territories and Networks*. Routledge, Oxon.
- Allen, J., Cochrane, A., 2010. Assemblages of state power: topological shifts in the organisation of government and politics. *Antipode* 42 (5), 1071–1089. <https://doi.org/10.1111/j.1467-8330.2010.00794.x>.
- Amin, A., 2002. Spatialities of globalisation. *Environ. Plann.* 34 (3), 385–399. <https://doi.org/10.1068/a3439>.
- Appadurai, A., 1996. *Modernity at Large: Cultural Dimensions of Globalisation*. University of Minnesota Press, Minneapolis.
- Axford, B., 2013. *Theories of Globalization*. Polity Press, Cambridge.
- Axford, B., 2014. Interview with professor barrie axford on “theories of globalization”. In: Global Studies Association. Retrieved from. <https://globalstudiesassoc.wordpress.com/2014/02/13/interview-with-professor-barrie-axford-on-theories-of-globalization/>. (Accessed 10 September 2021).
- Ball, S., 2016. Following policy: networks, network ethnography and education policy mobilities. *J. Educ. Pol.* 31 (5), 549–566. <https://doi.org/10.1080/02680939.2015.1122232>.
- Beck, U., 2000. The cosmopolitan perspective: sociology of the second age of modernity. *Br. J. Sociol.* 51 (1), 79–105. <https://doi.org/10.1080/000713100358444>.
- Brenner, N., 1999a. Beyond state-centrism? Space, territoriality, and geographical scale in globalization studies. *Theor. Soc.* 28 (1), 39–78. <https://doi.org/10.1023/A:1006996806674>.
- Brenner, N., 1999b. Globalisation as reterritorialisation: the re-scaling of urban governance in the European Union. *Urban Stud.* 36 (3), 431–451. <https://doi.org/10.1080/0042098993466>.
- Brenner, N., 2004. *New State Spaces: Urban Governance and the Rescaling of Statehood*. Oxford University Press, Oxford.
- Clarke, J., 2019. Foreword. In: Papanastasiou, N. (Ed.), *The Politics of Scale in Policy: Scalecraft and Education Governance*. Policy Press, Bristol, pp. v–xii.
- Cohen, D., 2017. Market mobilities/immobilities: mutation, path-dependency, and the spread of charter school policies in the United States. *Crit. Stud. Educ.* 58 (2), 168–186. <https://doi.org/10.1080/17508487.2016.1242507>.
- Connell, R., 2007. The northern theory of globalization. *Socio. Theor.* 25 (4), 368–385. <https://doi.org/10.1111/j.1467-9558.2007.00314.x>.
- Conrad, P., 2014. *How the World Was Won: The Americanization of Everywhere*. Thames Hudson, London.
- de Sousa Santos, B., 2006. Globalisations. *Theor. Cult. Soc.* 23 (2–3), 393–399. <https://doi.org/10.1177/026327640602300268>.
- Decuyper, M., 2019. Researching educational apps: ecologies, technologies, subjectivities and learning regimes. *Learn. Media Technol.* 1–16. <https://doi.org/10.1080/17439884.2019.1667824>.
- Decuyper, M., 2021. The topologies of data practices: a methodological introduction. *J. N. Approaches Educ. Res.* 9 (2), 67–84. <https://doi.org/10.7821/naer.2021.1.650>.
- Decuyper, M., Grimaldi, E., Landri, P., 2021. Introduction: critical studies of digital education platforms. *Crit. Stud. Educ.* 62 (1), 1–16. <https://doi.org/10.1080/17508487.2020.1866050>.
- Decuyper, M., Lewis, S., 2021. Topological genealogy: a methodology to research transnational digital governance in/through/as change. *J. Educ. Pol.*, 1–23. <https://doi.org/10.1080/02680939.2021.1995629>.
- Decuyper, M., Simons, M., 2016. Relational thinking in education: topology, sociomaterial studies, and figures. *Pedagog. Cult. Soc.* 24 (3), 371–386. <https://doi.org/10.1080/14681366.2016.1166150>.
- Fraser, A., 2010. The craft of scalar practices. *Environ. Plann.* 42 (2), 332–346. <https://doi.org/10.1068/a4299>.
- Grek, S., 2009. Governing by numbers: the PISA “effect” in Europe. *J. Educ. Pol.* 24 (1), 23–37. <https://doi.org/10.1080/02680930802412669>.
- Gulson, K., Sellar, S., 2019. Emerging data infrastructures and the new topologies of education policy. *Environ. Plann. Soc. Space* 37 (2), 350–366. <https://doi.org/10.1177/0263775818813144>.
- Gulson, K., Symes, C. (Eds.), 2007. *Spatial Theories of Education: Policy and Geography Matters*. Routledge, London.
- Hartong, S., Piattoeva, N., 2019. Contextualising the datafication of schooling: a comparative discussion of Germany and Russia. *Crit. Stud. Educ.* 1–16. <https://doi.org/10.1080/17508487.2019.1618887>.
- Harvey, D., 1989. *The Conditions of Postmodernity: An Enquiry into the Origins of Cultural Change*. Basil Blackwell, Oxford.
- Harvey, P., 2012. The topological quality of infrastructural relation: an ethnographic approach. *Theor. Cult. Soc.* 29 (4–5), 76–92. <https://doi.org/10.1177/0263276412448827>.
- Held, D., 1995. *Democracy and the Global Order*. Polity Press, Cambridge.
- Held, D., Anthony, M.G., Goldblatt, D., Perraton, J., 1999. *Global Transformations: Politics, Economics and Culture*. Polity Press, Oxford.

- Jessop, B., 1998. The enterprise of narrative and the narrative of enterprise: place marketing and the entrepreneurial city. In: Hall, T., Hubbard, P. (Eds.), *The Entrepreneurial City*. Wiley, Chichester, pp. 77–102.
- Jessop, B., 2002. *The Future of the Capitalist State*. Polity Press, Cambridge.
- Jessop, B., 2008. *State Power*. Polity Press, Cambridge.
- Jessop, B., Brenner, N., Jones, M., 2008. Theorising sociospatial relations. *Environ. Plann. Soc. Space* 26, 389–401. <https://doi.org/10.1068/d9107>.
- Kickert, W., 1995. Steering at a distance: a new paradigm of public governance in Dutch higher education. *Governance* 8 (1), 135–157. <https://doi.org/10.1111/j.1468-0491.1995.tb00202.x>.
- Larsen, M., Beech, J., 2014. Spatial theorising in comparative and international education research. *Comp. Educ. Rev.* 58 (2), 191–214. <https://doi.org/10.1086/675499>.
- Lash, S., 2012. Deforming the figure: topology and the social imaginary. *Theor. Cult. Soc.* 29 (4–5), 261–287. <https://doi.org/10.1177/0263276412448829>.
- Lata, I.B., Minca, C., 2016. The surface and the abyss/Rethinking topology. *Environ. Plann. Soc. Space* 34 (3), 438–455. <https://doi.org/10.1177/0263775815622086>.
- Lewis, S., 2017. Governing schooling through “what works”: the OECD’s PISA for Schools. *J. Educ. Pol.* 32 (3), 281–302. <https://doi.org/10.1080/02680939.2016.1252855>.
- Lewis, S., 2020. PISA, Policy and the OECD: Respatialising Global Educational Governance through PISA for Schools. *Springer Nature*, Singapore. <https://doi.org/10.1007/978-981-15-8285-1>.
- Lewis, S., 2021. The turn towards policy mobilities and the theoretical-methodological implications for policy sociology. *Crit. Stud. Educ.* 62 (3), 322–337. <https://doi.org/10.1080/17508487.2020.1808499>.
- Lewis, S., Hartong, S., 2021. New shadow professionals and infrastructures around the datafied school: topological thinking as an analytical device. *Eur. Educ. Res. J.* 1–15. <https://doi.org/10.1177/14749041211007496>.
- Lewis, S., Savage, G.C., Jessica, H., 2020. Standards without standardisation? Assembling standards-based reforms in Australian and US schooling. *J. Educ. Pol.* 35 (6), 737–764. <https://doi.org/10.1080/02680939.2019.1636140>.
- Lewis, S., Sellar, S., Lingard, B., 2016. “PISA for Schools”: topological rationality and new spaces of the OECD’s global educational governance. *Comp. Educ. Rev.* 60 (1), 27–57. <https://doi.org/10.1086/684458>.
- Lingard, B., 2021a. *Globalisation and Education*. Routledge, London.
- Lingard, B., 2021b. Multiple temporalities in critical policy sociology in education. *Crit. Stud. Educ.* 62 (3), 338–353. <https://doi.org/10.1080/17508487.2021.1895856>.
- Lingard, B., Lewis, S., 2016. Globalisation of the Anglo-American approach to top-down, test-based educational accountability. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York, pp. 387–403.
- Lingard, B., Martino, W., Rezai-Rashti, G., Sellar, S., 2016. *Globalising Educational Accountabilities*. Routledge, New York.
- Lury, C., Parisi, L., Terranova, T., 2012. Introduction: the becoming topological of culture. *Theor. Cult. Soc.* 29 (4–5), 3–35. <https://doi.org/10.1177/0263276412454552>.
- Martin, L., Secor, A.J., 2014. Towards a post-mathematical topology. *Prog. Hum. Geogr.* 38 (3), 420–438. <https://doi.org/10.1177/0309132513508209>.
- Massey, D., 1993. Power-geometry and a progressive sense of space. In: Bird, J., Curtis, B., Putnam, T., Tickner, L. (Eds.), *Mapping the Futures: Local Cultures, Global Change*. Routledge, London.
- Massey, D., 1999. Power-geometries and the Politics of Space and Time: Hettner Lecture, 1998. Department of Geography, University of Heidelberg, Heidelberg.
- May, J., Thrift, N. (Eds.), 2003. *Timespace: Geographies of Temporality*. Routledge, London.
- McCann, E., Ward, K., 2012. *Mobile Urbanism: Cities and Policymaking in the Global Age*. University of Minnesota Press, Minnesota.
- McKenzie, M., 2017. Affect theory and policy mobility: challenges and possibilities for critical policy research. *Crit. Stud. Educ.* 58 (2), 187–204. <https://doi.org/10.1080/17508487.2017.1308875>.
- McKenzie, M., Aikens, K., 2021. Global education policy mobilities and subnational policy practice. *Glob. Soc. Educ.* 19 (3), 311–325. <https://doi.org/10.1080/14767724.2020.1821612>.
- McKenzie, M., Lewis, S., Gulson, K., 2021. Response: matters of (im)mobility: beyond fast conceptual and methodological readings in policy sociology. *Crit. Stud. Educ.* 62 (3), 394–410. <https://doi.org/10.1080/17508487.2021.1942942>.
- Meyer, H.-D., Benavot, A. (Eds.), 2013. *PISA, Power and Policy: The Emergence of Global Educational Governance*. Symposium Books, Oxford.
- Mezzadra, S., Neilson, B., 2012. Between inclusion and exclusion: on the topology of global space and borders. *Theor. Cult. Soc.* 29 (4–5), 58–75. <https://doi.org/10.1177/0263276412443569>.
- Michael, M., Rosengarten, M., 2012. HIV, globalisation and topology: of prepositions and propositions. *Theor. Cult. Soc.* 29 (4–5), 93–115. <https://doi.org/10.1177/0263276412443567>.
- Ozga, J., 2021. Problematising policy: the development of (critical) policy sociology. *Crit. Stud. Educ.* 62 (3), 290–305. <https://doi.org/10.1080/17508487.2019.1697718>.
- Papanastasiou, N., 2016. How does scale mean? A critical approach to scale in the study of policy. *Crit. Pol. Stud.* 11 (1), 39–56. <https://doi.org/10.1080/19460171.2015.1119052>.
- Papanastasiou, N., 2017. The practice of scalecraft: scale, policy and the politics of the market in England’s academy schools. *Environ. Plann.* 49 (5), 1060–1079. <https://doi.org/10.1177/0308518x16687764>.
- Peck, J., Theodore, N., 2015. *Fast Policy: Experimental Statecraft at the Thresholds of Neoliberalism*. University of Minnesota Press, Minneapolis.
- Ritzer, G., 2013. *The McDonaldization of Society*. Sage, Los Angeles, CA.
- Rizvi, F., Lingard, B., 2010. *Globalising Education Policy*. Routledge, New York.
- Robertson, S.L., 2010. “Spatialising” the sociology of education: stand-points, entry-points and vantage-points. In: Apple, M.W., Ball, S., Gandin, L.A. (Eds.), *The Routledge International Handbook of the Sociology of Education*. Routledge, Oxon, pp. 15–26.
- Robertson, S.L., Bonal, X., Roger, D., 2002. GATS and the education service industry: the politics of scale and global reterritorialization. *Comp. Educ. Rev.* 46 (4), 472–496. <https://doi.org/10.1086/343122>.
- Rose, N., 1999. *Powers of Freedom: Reframing Political Thought*. Cambridge University Press, Cambridge.
- Sahlberg, P., 2011. *Finnish Lessons: What Can the World Learn from Educational Change in Finland?* The Series on School Reform. Teachers College Press, New York.
- Sassen, S., 2006. *Territory, Authority, Rights: From Medieval to Global Assemblages*. Princeton University Press, Princeton.
- Savage, G.C., 2020. What is policy assemblage? *Territ. Polit. Governance* 8 (3), 319–335. <https://doi.org/10.1080/21622671.2018.1559760>.
- Savage, G.C., Di Gregorio, E., Lingard, B., 2021. Practices of scalecraft and the reassembling of political boundaries: the contested nature of national schooling reform in the Australian federation. *Pol. Stud.* 1–22. <https://doi.org/10.1080/01442872.2021.1885640>.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. *Br. Educ. Res. J.* 40 (6), 917–936. <https://doi.org/10.1002/berj.3120>.
- Sheller, M., Urry, J., 2006. The new mobilities paradigm. *Environ. Plann.* 38 (2), 207–226. <https://doi.org/10.1068/a37268>.
- Shields, R., 2012. Cultural topology: the seven bridges of Königsberg, 1736. *Theor. Cult. Soc.* 29 (4–5), 43–57. <https://doi.org/10.1177/0263276412451161>.
- Smith, R.G., 2003. World city topologies. *Prog. Hum. Geogr.* 27 (5), 561–582. <https://doi.org/10.1191/0309132503ph447oa>.
- Steger, M.B., 2009. *Globalisms: The Great Ideological Struggle of the 21st Century*. Rowman & Littlefield, Lanham, MD.
- Steger, M.B., James, P., 2013. Levels of subjective globalization: ideologies, imaginaries, ontologies. *Perspect. Global Dev. Technol.* 12 (1–2), 17–40. <https://doi.org/10.1163/15691497-12341240>.

- Takayama, K., 2012. Exploring the interweaving of contrary currents: transnational policy enactment and path-dependent policy implementation in Australia and Japan. *Comp. Educ.* 48 (4), 505–523. <https://doi.org/10.1080/03050068.2012.721631>.
- Tan, C., Yang, J.A., 2019. Path-dependency or path-shaping? An analysis of the policy to target exam-orientation in South Korea. *Crit. Stud. Educ.* 1–16. <https://doi.org/10.1080/17508487.2019.1611612>.
- Thompson, G., Sellar, S., Buchanan, I., 2021. 1996: the OECD policy-making assemblage. *J. Educ. Pol.* 1–20. <https://doi.org/10.1080/02680939.2021.1912397>.
- Urry, J., 2007. *Mobilities*. Polity Press, Cambridge.
- Wallner, J., Savage, G.C., Hartong, S., Engel, L., 2020. Laboratories, coproducers and venues: roles played by subnational governments in standards-based reforms in four federations. *Comp. Educ. Rev.* 64 (2), 249–268. <https://doi.org/10.1086/708111>.
- Williamson, B., 2017. *Big Data in Education: The Digital Future of Learning, Policy and Practice*. Sage, Los Angeles.

Comparative education: then, now, and next

Robert Cowen, University College London Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	11
The invention of traditions	11
Traditions and tensions: science	12
Traditions and tensions: the practical	13
The end of “history” and the fragmentation	14
Globalization and governance	15
Conclusion	15
References	17

Introduction

Comparative education is an academic subject taught in universities and teacher education institutions, in many countries. It can also, more broadly, be thought of as a complex form of action on the world which creates policy knowledge through, for example, consultancies, funded research, and advisory work for governmental agencies (of whichever country) that stress the development of educational systems in other countries. Even more loosely, the phrase “comparative and international education” can be stretched to include international testing of children’s attainment (such as that undertaken by PISA) and the advice which international agencies (like the World Bank or OECD) give on educational reform. This article comments only briefly on such forms of “applied comparative education” and concentrates on trying to analyze comparative education as an academic subject. Overall, the argument of the article follows a “then, now, and next” sequence.

As might be expected, there are around the world considerable variations in how the academic subject has been defined in universities since what was arguably the first “taught course” in comparative education in 1899–1900 (Bereday, 1963). The bibliography that is available on comparative education grew rapidly during the 1960s and even more rapidly after (Altbach and Kelly, 1986; Cowen and Kazamias, 2009). The consequence is that, currently, in addition to the perennial problem of the potential hegemonies of the English language, there are considerable differences between, say, “comparative education” as this is understood in southern Europe (Palomba and Cappa, 2018) and how comparative education is thought about in what is normally referred to as CIES, the (American) Comparative and International Education Society (Cook et al., 2004). These variations, in how “comparative education” is defined by university academics, involve different assumptions about the relations of comparative education to other branches of educational studies, varying choices about which political and geographic spaces should be investigated, and divergent expectations about the “usefulness” of comparative education for sorting out educational or social problems at home, and overseas.

These patterns and differences have been discussed in greater and greater detail in the last 30 years (Bray et al., 2007a; Halls, 1990; Wolhuter et al., 2008). What has also been documented is, in the phrasing of Erwin Epstein (2008), the “professionalization” of the field of study in academic societies and specialist journals (Masemann et al., 2007; Cook et al., 2004; Cowen, 1990). All this adds up to a considerable body of literature and - not least as part of the invention of a field of study—there are the textbooks.

The invention of traditions

Some of these textbooks were written in the interwar period and soon after 1945—partly to construct an academic field of study but the books were also intended to be read by students. Some of the best-known early texts which suggest a scope and focus for “comparative education” include Hans (1949) and Kandel (1933) and in the late 1950s—at a lower level of intellectual difficulty, presumably because the texts were intended to be attractive to a wider student audience—books by Edmund King (1958) and Malinson (1957). Subsequent introductory textbooks became more self-consciously “theoretical” (e.g., Jones, 1971; Trethewey, 1976). By the 1990s and after, the concept of “textbook” included relatively complex constructions of the field of study, outlines of its new ideas, and explanations of its epistemic choices (Alexander et al., 1999; Arnone and Torres, 2007; Bray et al., 2007a; Burns and Welch, 1992; Kubow and Fossum, 2003; Phillips and Schweisfurth, 2008).

Thus, since the 1960s with its first major expansion of the literature, and certainly by the 1990s, it has been possible to use the expression “the history” of the field of study to mean scholarly literature, published in journals or in book form, which offers accounts of “comparative education” as this has been thought about in the last couple of hundred years. Some of the journal literature on that history, not least because of illustrations *en passant* of, say, Tolstoy’s interest in “education elsewhere”, remains fascinating (Brickman, 1960) but the raw materials of history—documents—were also beginning to be assembled (Fraser and Brickman, 1968) and there were succinct “histories” in books written for students doing courses in comparative education.

One of these early books is of major assistance in identifying some of the peculiarities of the conventional definition our “tradition”. The first part of Noah and Eckstein’s (1969)—with the title *Towards a science of comparative education*—offers a lengthy “history”, but it is a missionary history; teleological in its unfolding and destination, and also missionary in that it was designed to persuade. Its finishing point is similar to its starting point: comparative education should become a science. The “history” is in the first part of the book. The account is enthusiastic, well written, and informed by library research. Just in case the reader misses the strategic point, the remaining parts of the book and a companion volume (Eckstein and Noah, 1969) give examples of the proper architecture of good comparative studies. Of course, fifty years later, we have better histories—notably Maria Manzon’s (2011) book *Comparative Education: the construction of a field*—as well as some new problems with our “history”.

Nevertheless, the Noah and Eckstein text is useful as a tribal myth. It offers a neat over-simplification of a story—a history—that tends to become more detailed and more interesting each year (Epstein, 2020; Manzon, 2019; Phillips, 2020a; Takayama, 2018) and the Noah and Eckstein book is also good as a starting point because the text deliberately creates tensions about the identity of comparative education—a theme which is still being strenuously debated in 2020, albeit on very different terms.

Traditions and tensions: science

The invocation of the theme of “science” is (within comparative education) normally linked to the name of Marc-Antoine Jullien (Fraser, 1962). Jullien’s argument included looking at the cantons of Switzerland to see patterns of variation in educational provision and using this “comparative” perspective to escape from the whims and prejudices on which educational decisions had (allegedly) been based earlier. Jullien deliberately used the expression “science of education”. Given he was writing in 1817, his reason for using this vocabulary is normally linked to the Comtean positivist movement in French thought. Currently, as work on the history of comparative education grows more nuanced and refined, there is a question in the specialist literature about whether Jullien really was “the Father” of comparative education—but that is not the point here.

The question, here, is how did comparative education get positioned, ideologically, as a science? The short answer is “early” and by Jullien. A more difficult question to answer is: how does this positioning get re-enforced over time and what tensions does this create? Interestingly, the struggle is not at the end of the nineteenth century as the social sciences (e.g., economics and, gradually, sociology) begin to be so labeled in some university systems. Nor was the struggle part of the rebalancing in universities of the importance of “classics” (the study of the Greek and Latin languages and the civilizations of Greece and Rome) versus the natural or exact sciences. Because comparative education was part of educational studies—studies often not institutionalized at university level in the English-speaking world until the 1930s—recovery of the theme of a “science” of comparative education did not occur until the mid and late 1960s. But then the debate was quite sharp and at the same time, paradoxically, quite muddled.

The muddle was, and is, that the academics involved judged that their debate was about methodology, but the debate is more complicated. It was simultaneously a debate about the usefulness of comparative education for the purpose of shaping educational policy *and* about the earlier academic identities of the academics themselves *and* about modes of explanation. Thus, for example, George Bereday in Teachers College Columbia saw comparative education as an empathetic way of understanding the patterns of educational systems—including one’s own. Bereday was originally from Poland and his epistemic vision was that of a European encyclopedist: all forms of social science knowledge should be used to describe and understand patterns of education—with perhaps a slight preference for understanding via “political science” (Bereday, 1964). In contrast, Harold Noah, with an original training in London School of Economics and Political Science (and despite, later, attending the lectures of Nicholas Hans in King’s College in London) tended in his work with Max Eckstein in New York to prefer a perspective that emphasized “variables”, almost as if the ideal form of comparative education would be akin to econometrics. This was in complete contrast to the perspective of Edmund King who, after Hans, also worked in King’s College London. King (originally trained in classics) was perhaps the most extreme advocate of the mainstream British post-war knowledge-tradition in comparative education, which Bill Halls (1973) called a “culturalist” perspective, emphasizing languages, and an understanding of “the local” culture (in whichever “local” was being discussed). And worse: the Bereday and Noah and Eckstein and King positions contrasted with Brian Holmes’ firm view (Holmes had originally trained in physics) that both academic understanding and policy advice were dependent upon science. However—and this was the originality of Holmes’ work to the point where he was prepared to claim toward the end of his career that he had invented a new paradigm within comparative education (Holmes, 1984)—science was not a matter of following Roger Bacon or J.S. Mill and discovering causes, but of having theories which are able to predict, successfully. This, in comparative education, means being able to predict the consequences of “borrowing”; that is, the transfer of educational ideas and principles, institutions and practices, from one country to another.

The term “science” lingered on. For example, Barber (1973) offered a good article on science and comparative education. By the time we are in this century, the concept of science is centrally positioned by Gita Steiner-Khamsi in her opening chapter as a frame for the 2012 *World Yearbook of Education* (Steiner-Khamsi and Waldow, 2012). Her argument was that comparative education is a policy science. Juerger Schriewer (2000) discusses a “science of complexity”. Paradoxically the comparative educator most fully trained in the natural sciences, Joseph A. Lauwerys—who held degrees in both biology and chemistry and who began his career as a Lecturer in Science Education in the Institute of Education of the University of London—had tended to follow Nicholas Hans in the way in which he lectured and wrote about education. Lauwerys retained the word “science” mainly to confirm that his personal world-view was that of a “scientific humanist.” When Lauwerys reflected on comparative education and its uses—as he did in

writing but also in action as one of the core creators of the Comparative Education Society in Europe—he used the expression “educational statesmanship” (Cowen, 2020a).

Traditions and tensions: the practical

The theme of “educational statesmanship” (although it is left implicit and undefined) is in “the history” of Noah and Eckstein (1969). They note the work of some of the major administrators of education of the nineteenth century—persons such as Victor Cousin in France, James Kay-Shuttleworth in Britain, and Horace Mann in the United States. Historically, the point is much more complicated than merely a motif within comparative education—as the current crisis in Canada over the name of Egerton Ryerson demonstrates. Nevertheless, the original point made by Noah and Eckstein is clever: to include practical people in the history of an academic subject. The move makes “the history” more sequential and legitimates the theme of educational reform.

These administrators, in a series of practical actions, did indeed create “the educational system” in its initial nineteenth century models and also undertook “research”; that is, they collected information from other countries about the elementary school, its teaching practices, the training of teachers for it, and its finance and administration. Prussia was of major interest, but the search for information roamed widely across Europe and “Reports” became an early form of “comparison”. Thus, the sense of shock that “educational administrators” should be included in a history of academic comparative education is reduced. The administrators addressed two “unit ideas” of comparative education (Cowen, 2021): “transfer” and “the educational system” itself. They were also “investigating” foreign educational systems. Like so many doctoral theses currently, they did have research questions (though they were spared having to write interminable chapters on “the literature” and “the method”) and so an important way to think about the administrators is to note how they link information about “education elsewhere” with policy-and-practice. That is, they do an early version of “normal-puzzle” comparative education.

Indirectly, the administrators also help to construct current tensions over “normal-puzzle” comparative education and its political positioning. By the end of the 19th century, a famous question is asked by Sir Michael Sadler—himself a policy person and a senior administrator in the English university system (Phillips, 2020b). In what seems—from its title—to be a very simple lecture, Sadler asks, “How far may we learn anything of practical value from the study of foreign systems of education?” (Sadler, 1900). His answer is far from simple—he invokes the battles of Kosovo and Waterloo to illustrate how difficult it is to extricate educational principles and practices from the flows of history; but the question is a crucial one because the title of the talk provided a legitimization for decades of work in academic comparative education—and it still does. Sadler’s question and the work of the major administrators of national systems of the 19th century permit the routine assumption, particularly after 1945 and in the 1960s and 1970s debate about “methodology”, that the purpose of comparative education is to contribute to the (gradual) reform of educational systems and to advise on educational policy.

This tradition takes various forms. One is when Lauwerys and a small group of founding professors constructed the Comparative Education Society in Europe (whose meetings these days are much more like a lively and relaxed family conversation) as something like an “Academy of Letters”, in which educational statesmen, carefully elected by their peers, would be available to advise Ministers of Education. A similar “practical-policy” framing was also given to the (World) *Yearbooks of Education*—initially edited by Lauwerys and Hans in the Institute of Education in the University of London and later edited mainly by Lauwerys and George Bereday from the late 1940s until the beginning of the 1970s. The books were (or were intended to be) “useful”: for example, after volumes on post-war reconstruction, there were Yearbooks on teacher education systems; secondary schools; vocational-technical education; examinations; curriculum; and on universities and their reform. That pattern of doing “normal-puzzle” comparative education in the WYBs continued until about 2004 (Cowen, 2022).

The tradition was very strong. In the 1960s and 1970s, when the nominal theme of “methodology” splintered a generation of mid-career comparative educationists, almost the only point of agreement among them was that comparative education should continue to aim to be useful for policy purposes. The convention that the normal-puzzles of comparative education are sectors of the education system continues. For example, the comparative literature on higher education is very large indeed. There is also a copious comparative literature on teacher education, vocational-technical education, secondary schools, examinations, and curriculum in comparative perspective, and the relatively small comparative literature on early childhood education has just been revitalized (Sousa and Moss, 2022).

Overall, then, the long tradition of writing on policy themes which are often, de facto, the administrative boxes of schooling systems and reflecting on reform and “amelioration” seems to be continuous within comparative education. It is, but there are two interesting changes in the topography of the field of study. One redrawing of our topography is deliberate, recent and very visible: the investigation of new patterns in the “governance” of education. (That theme will be discussed later as part of the theme of “globalization”). The other theme is the change in assumptions about the importance of history as an approach to doing comparative education. This had major consequences for our topography. However, these consequences are not held together in the literature, not even in the brilliant “mapping” done by Rolland Paulston (2009). The consequences have not been seen as patterned; but they are.

The end of “history” and the fragmentation

The debate about “history” as if it were merely “a method” has already been mentioned. Kandel and Hans had worked within the perspective of history but the new embracement of a range of social sciences (e.g., by Bereday) and enthusiasms for something akin to a Noah and Eckstein vision of variables, rapidly affected the journals and the articles they were publishing. The debate about history seemed to have a simple beginning and end. Historians became Humpty Dumpties. More precisely, Andreas Kazamias, who has a sharp sense of the visual, chose when he was the Editor of *Comparative Education Review* to publish a drawing of Humpty Dumpty, fallen to the ground, with the dramatic consequences of a large fall made visible. This was a brilliant visual *aperçu* in a journal not given to using the visual. Later—the ironies of history being what they are—Kazamias was to write of persons he saw as historians-and-philosophers-and-liberal humanists as “forgotten men” (2009). This is in many ways true, but it also understates the complexity of the theme of history as a way of thinking in comparative education and the consequences of not worrying about history.

What at first glance seems a minor aberration, a touch of overconfidence by the new generation of comparative educationists in the 1960s—separating comparative education from history—marks the growth of a massive strategic problem for the identity of comparative education as a field of study. The problem was and is how to see the multiple shards of the fragmentation of the field of study as one “event”. It is not until this century that the literature has become sufficiently complex to see how the crisis “in methodology” had multiple effects that were far more serious. Hence the label used here: “The Fragmentation”: the burst of incoherence that came from separating what had—almost—been kept together.

The Fragmentation includes sudden limits on our thinking about (i) educated identity; (ii) time; (iii) space; (iv) political belief systems, including nationalism; (v) Empires; and (vi) confusions hidden within the re-naming of the field of study itself.

The flow of assertions that comparative education should address questions of policy and reform in “modern” educational systems meant that forms of “educated identity” in which different assumptions about society and power were compressed—for example, in the education of The Prince, in the education of the samurai, in the education of the courtier—were taken out of the imagination, despite the obvious originality of questions raised by conventional books on the history of education, but also by comparative analyses such as Max Weber writing on the Chinese literati (Gerth and Mills, 1970) or the new work of Margaret Archer (1979). More generally our professional sense of time was rebalanced: we became interested only in contemporary-time and future-time, in which to see the effects of policy advice. Meanwhile, the historians continued to develop a specialist comparative literature of their own to the point where we are now trying to catch up on the themes of transnational, comparative, and global history (Cowan, 2020b) and to rescue and redefine our sense of “uneven times and places”; a theme that was extremely well discussed by the McLeod, Sobe and Seddon *World Yearbook of Education* (2018).

However, the job is massive and some points are worth continued attention. For example, comparative educationists got themselves into a muddle over space, a rather interesting thing for comparative educationists to get muddled about. And they have not yet recovered. There used to be a broad but fairly clear definition—Bereday used to write of the “northern crescent”: Europe and Russia, and north Asia and China and then north America and back to Europe. This was the “space” that comparative educationists thought about, with the occasional excursion into, say, Chile. The point was paralleled in the Institute of Education in London where the Comparative Education Department was created in 1948. It also tended to study “the northern crescent”. However, as the economics of education as a set of theories about development became influential and the politics of the Cold War grew colder, this terminology collapsed. Societies were now being variously labeled: developed or underdeveloped; agricultural or industrial, or “newly-industrialized”; traditional and modern; and there was a Third World (and a first, second and fourth). The confusions seem retrospectively to be dramatic—but in one way, there were no confusions at all. At one point, it was sensible to think of Alex Inkeles or Talcott Parsons as the relevant theoretician and labeler of spaces; at another, maybe Theodore Schultz or C. Arnold Anderson; or the World Bank or OECD. The themes that Kandel and Hans had begun to explore (such as nationalism and fascism) were forgotten. The theme of Empire almost disappeared (although cf. Clignet and Foster, 1964; Madeira, 2008); but where is the comparative education work on, say, the Austro-Hungarian and Ottoman Empires?

Indeed, the theme of the British Empire—much analyzed by historians—came close to being an embarrassment within comparative education. The Institute of Education in London had established a Department for “Education in the Colonies” in 1928 (about 20 years before the Comparative Education Department was established under Lauwerys). Of course, post-1945, there was a rather embarrassed—and embarrassing—sequence of renamings, which included the remarkable title “Education in Tropical Areas” before “Education in Developing Countries” fitted comfortably into a changed *Zeitgeist*. What did pop up suddenly (given that Noah and Eckstein had issued their challenge—become a science—in the very late 1960s) was writing on “education as cultural imperialism” (Carnoy, 1974) and on “colonialism” (Altbach and Kelly, 1978). But the theme languished. The concept of Empire took less and less attention, and only now has Empire and the themes of “post-socialism”—not least in the superb analyses of Iveta Silova—and concern with the “post-colonial” and globalization marked early by Tikly (1999), become a major contemporary debate (Takayama et al., 2017) and brought everything back to life in a very serious way.

However, *entre-acte*, as it were, there was a fascinating debate that is far from finished: a renaming of the field. The vocabulary of “comparative and international education” permitted the absorption of the theme of development. Two academic societies changed their name—the American Comparative Education Society founded in 1956 became CIES and the British Society, initially a branch of the Comparative Education Society in Europe, finally became the British Association of International and Comparative Education (BAICE). David Wilson (1994) rather charmingly uses the expression “twins” which perhaps glides past the point of flows of international power captured in the question: who exports and who imports. The broader debate was considerable and it continues

(Carnoy, 2006; Crossley, 1999; Epstein, 2016) though this is slightly obscured, hidden by the muddle-headed absorption of the concept of “globalization” into comparative education.

Globalization and governance

“Globalization” became a major word for work in education and comparative education from, let us say, 1979. That date is a useful as a symbol, rather than an historical fact. The date however can be used as marker of the choice between a soft version of globalization which merely emphasizes mobile factors of production (capital, labor and sites of production) and a harder version which was well captured by Mrs. Thatcher’s first government and the insistence that “there is no alternative” (TINA). Despite some major efforts to stabilize the concept for comparative education, for example, in the writing of Tony Welch (2001), the soft version of the topic of globalization tended to edge into an even softer version of “internationalism” or be routinized and reduced to normal-puzzle solving comparative education and the reform of skills and curricula. However, there is hard-edged work in the literature, especially on the economic ideologies which began to inform policy thinking, notably by Rizvi and Lingard (2009) and Arnove (1980) working within and contributing to the Wallerstein tradition of thinking about the global. Fortunately, the theme of globalization rapidly overlapped with governance and the international agencies (Beech, 2009; Grek, 2009; Lawn and Grek, 2012; Li and Auld, 2020). A shift in one of the “unit ideas” was occurring: the nature of the State and loci of “governance” were changing and international economic relations were moving into a new form. Finally, this attracted the term “neo-liberalism”—attended by its major inequalities, about which historians such as Tony Judt had been grumbling for some time.

Amid these politics, a new kind of quasi-comparative education was finally taking very clear shape—as part of “governance” and implicitly as part of a hard-data “science” of governance, which will provide “solutions”. One of the major forms of “comparing-as-a-solution” is international testing (Cowen, 2014). An early version of this (part of the work of the International Association for the Evaluation of Educational Achievement—the IEA) was framed by the power politics of the Cold War and ranked the achievements by children of different nations (for example and famously, in mathematics). A contemporary version of this is PISA—whose testing is as exact as contemporary social science permits. Thus, the technical work of PISA might be termed “scientific” and PISA testing manifestly “compares”.

However, it would be more cautious to note that a thermometer is not medicine. PISA will give you a low or a high score (and the score may move you into PISA shock or Finland frenzy) but PISA tells you nothing useful about the social embeddedness of the Finnish education system, nothing useful about “transfer”, and it embraces specific notions of educated identity. PISA is politics (Auld and Morris, 2016; Auld et al., 2019; Tröhler, 2013). The OECD and the World Bank and UNESCO are part of the problem which comparative education, in its theory work, is analyzing: the flows and mix of international and domestic power as these are compressed into educational forms (Cowen, 2009).

And that, amazingly, brings us back to “science” and what academic comparative education has done, and should do next. The term “science” re-gathers its full strategic force for comparative education in Novoa’s (2018) paper. That brings us almost—but fortunately not quite—back to the starting point of this analysis.

Conclusion

The analysis began by trying to sketch “academic comparative education” and its “traditions”. The word “tradition” implies stability, and while it is possible to stress the continuities of “normal-puzzle” comparative education, tensions became obvious in academic comparative education by the 1960s. By the end of the 1970s, the discontinuities in academic comparative education had become dramatic, to the point where it becomes possible—even without examining seriously difficult concepts such as “context” or “inter-disciplinarity”—to discuss fragmentation.

Currently there are new questions giving an edge to contemporary discussion: for example, the theme of the politics of “epistemic space”, and its cultural and hegemonic significance. The “history” of comparative education sketched here is a history of Anglo-American and European work, which has influenced comparative education in Latin America, Australia, Hong Kong, India, Japan, Singapore, South Africa, South Korea (and so on). Maybe the strategic shift in vision will come from thinking about “Asia-as-method”, but continued progress on that theme is slow and is visible in the rest of Asia including East Asia, rather than in mainland China where comparative education is tending to take on a politically safe, “applied”, identity.

As an extension of this point, comparative educationists have not systematically addressed the theme of internal political space: the political, economic and cultural limits which shrink academic “comparative education” until its imagination and analytical reach is trivialized in universities by States which have become authoritarian—whether this authoritarianism is a consequence of populist parochialism, political philosophy, religious belief, or the infallibility of the Local Great Leader. This problem is not raised rhetorically: the demand by English politicians for “robust and relevant research” and later for “impact” has been corrosive (Cowen, 2000) altering both the external relations of the university—increasingly to be understood and measured in economic terms—and internal definitions of the “quality” of teaching and “good” research.

Such political questions are crucial for comparative education the moment it decides that it is not its job to advise the local Minister of Education. Unfortunately, except in the Cold War—or perhaps even because of the Cold War—it has not bothered much with its political positioning. It should, now, do that. The world recovering from the Second World War, a world energetically

expanding educational provision as part of a proper definition of citizenship, is the world of George Bereday and Joseph Lauwerys, a world of optimism. We have perhaps returned to a world in which the questions are more brutal: for which kind of political regimes should “comparative education” be useful these days?

Amid these serious (and rather esoteric) anxieties, the traditional normal-puzzle work of comparative education continues—and it should do so for perfectly proper reasons which have to do with a commitment to teachers, children, and education. Efforts to think about the improvement of education, with education normally thought of in its “boxes”—higher education, teacher education, vocational-technical schools, curriculum (including “teaching and learning”)—is our normal-puzzle work, but this is rooted in delicacies (and difficult theorization) about the nature of “educated identity”—our moral fulcrum and one of our core ethical questions. This normal-puzzle work emphasizes children and the “re-forming” of the educational service itself. It would be improper to ignore or denigrate normal puzzle work because it includes a powerful commitment to the importance of teachers (and, where this is possible, to initial teacher education). In addition, it is our consultancies and funded research (usually on normal-puzzle problems) which help comparative education to survive as an academic subject in some university systems. Right now, “comparative education” is a goose laying rather golden eggs.

That said, the reasons for looking beyond normal-puzzle work in comparative education are important. Comparative education has never quite sorted out its ethical issues (Cowen, 1973) as it tries to act on the world. We need to worry about that. There was no end to history, nor indeed to “our history”. We need to re-think that history and “The Fragmentation” and to offer fresh readings-of-the-global; not least by noting the global political economic and cultural relations of the contemporary period and redefining what should be studied—accepting fully that certain themes such as inequalities in human rights, and inequalities in education in region and race, gender and class, nations and ethnicities will continue to need major and unremitting attention. As this work proceeds (and as argued earlier), the international agencies—including even UNESCO—should be thought of as part of the problem, part of the problematic of new readings of the global; and not absentmindedly accepted as a benign.

Of course, among these pressures and problems—and perhaps partly because of them—the acquisition of a coherent sense of “tradition” seems proper. However, settling too comfortably into the self-identity given by “a tradition” is dangerous because the problem suddenly mutates. It becomes how to shake loose from tribal memories and to re-imagine. And we should re-imagine. For example, we have:

- forgotten war and revolution and the massive problems of understanding (comparatively) Empires—except through our discussions of colonialism and post-colonialism and post-socialism;
- until very recently, failed to wonder about the Gaia hypothesis’ and global warming;
- tended to treat the words “development” and “international” and “globalization” as if they were politically neutral and benign;
- until very recently, forgotten about the problem of nationalism—whether banal or virulent;
- often avoided questions about the political positioning of academic knowledge about societies, as if such knowledge is neutral (which it ceases to be once a choice of place, time, and topic is made);
- not linked our literature with new thinking about transnational and global history and the best work in “intercultural education”; and
- we have tended to assume “educational systems” will—with some fine-tuning—continue in their 20th century forms.

Overall, our literature tends to be based not only on “normal-puzzle” comparative education, but also on “normal-theory” comparative education which has a weakness for complexities that need very long words—post-structuralism, post-foundationalism, neo-institutionalism—and an urge to borrow from “savior saints” who have included Beck, Bourdieu, Derrida, Foucault, Giddens, Habermas and Luhmann. That is in a way reassuring; almost a bad tradition. Earlier generations of comparative educationists drew on racist versions of British anthropology, structural functionalism, human capital theory, Talcott Parsons, Walt Rostow, and so on. The mono-optical deduction of the future of comparative education from one perspective, from one theoretician or from one social science, has the virtue of simplicity, and is perhaps to be preferred to a scattered future deduced, *seriatim*, from all of the social sciences.

There is another way, fortunately. A risky but exciting step into the future might be to put more emphasis on inventing some theories of our own. The good news is that there is already grit in the oyster: there is comparative writing of major originality available in the literature by those who would not routinely label themselves as “comparative educationists” (for example, Margaret Archer, Stephen Ball, Roger Dale, Martin Lawn, Bob Lingard, Heinz-Dieter Meyer, Jenny Ozga, Tom Popkewitz, Terri Seddon, Theda Skocpol, Fazal Rizvi, Susan Robertson, Daniel Tröhler, Rupert Wilkinson).

Substantively, there are also a number of World Yearbooks of Education which break new theoretical ground. In addition to the 2018 WYB on “uneven time and spaces”, there is, depending on personal interest, quite a choice in terms of topics and new theorization. For example, the 2005 WYB was on globalization and nationalism; the 2006 volume addressed the politics of educational research; the 2008 volume reflected on “geometries of power” and changes in higher education; and the 2014 WYB analyzed “governing knowledge”. The latest volume (Tröhler et al., 2022) recovers a sense of history and takes a fresh look at nationalism and its “global universalization”.

Furthermore, there are currently three short but major papers within the specialist comparative literature which raise vital, strategic, theoretical, and political questions. The pattern of answers to those questions will re-orient the field: Nóvoa (2018), Nóvoa and Yariv-Mashal (2003), and Takayama (2018). Happily, the energy to answer those questions is there: some of our early and mid-career scholars are already creating disruptive and challenging theories. Their work has also begun to affect the “unit ideas” (including time and space); to reassert the relativisms of difference in educated identity while using that concept as an ethical

bedrock for comparative education; and to trace the flows of international and domestic power as these are compressed in educational systems. The earlier problem of “the fragmentation” of comparative education has not yet been fully and coherently stated, but ways to think about it and finally to move on, past it, are emerging.

References

- Alexander, R., Broadfoot, P., Phillips, D., 1999. *Learning From Comparing: New Directions in Comparative Educational Research*. Symposium Books, Oxford.
- Altbach, P.G., Kelly, G.P., 1978. *Education and Colonialism*. Longman, London.
- Altbach, P.G., Kelly, G.P., 1986. *New Approaches to Comparative Education*. University of Chicago Press, Chicago.
- Archer, M., 1979. *Social Origins of Educational Systems*. Cambridge University Press, Cambridge.
- Arno, R.F., 1980. Comparative education and world-systems analysis. *Comp. Educ. Rev.* 24, 48–62.
- Arno, R.F., Torres, C.A., 2007. *Comparative Education: The Dialectic of the Global and the Local*. Rowman and Littlefield, Lanham, MD.
- Auld, E., Morris, P., 2016. PISA, Policy and Persuasion: translating complex conditions into education “best practice”. *Comp. Educ.* 52 (2), 202–229.
- Auld, E., Rappleye, J., Morris, P., 2019. PISA for Development: how the OECD and World Bank shaped education governance post-2015. *Comp. Educ.* 55 (2), 197–219.
- Barber, B.H., 1973. Science, salience, and comparative education: some reflections on social scientific enquiry. In: Edwards, R., Holmes, B., van de Graaff, J. (Eds.), *Relevant Methods in Comparative Education*. UNESCO Institute for Education, Hamburg.
- Beech, J., 2009. Who is strolling through the global garden? International agencies and educational transfer. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.
- Bereday, G.Z.F., 1963. James Russell’s syllabus of the first academic course in comparative education. *Comp. Educ. Rev.* 7 (2), 189–196.
- Bereday, G.Z.F., 1964. *Comparative Method in Education*. Holt, Rinehart and Winston, Inc., New York.
- Bray, M., Adamson, B., Mason, M., 2007a. *Comparative Education Research: Approaches and Methods*. Comparative Education Research Centre, The University of Hong Kong, Hong Kong, and Springer, Dordrecht.
- Brickman, W.W., 1960. A historical introduction to comparative education. *Comp. Educ. Rev.* 3 (3), 6–13.
- Burns, R.J., Welch, A., 1992. *Contemporary Perspectives in Comparative Education*. Garland Publishers, New York.
- Carnoy, M., 1974. *Education as Cultural Imperialism*. Longman, New York.
- Carnoy, M., 2006. Rethinking the comparative—and the international. *Comp. Educ. Rev.* 50 (4), 551–570.
- Clignet, R., Foster, P., 1964. French and British colonial education in Africa. *Comp. Educ. Rev.* 8 (2), 191–198.
- Cook, B.J., Hite, S.J., Epstein, E.H., 2004. Discerning trends, contours, and boundaries in comparative education: a survey of comparativists and their literature. *Comp. Educ. Rev.* 48 (2), 123–149.
- Cowen, R., 1973. A query concerning developments within and the responsibilities of comparative education. *Can. Int. Educ.* 1 (1973), 15–29.
- Cowen, R., 1990. The national and international impact of comparative education infrastructures. In: Halls, W.D. (Ed.), *Comparative Education: Contemporary Issues and Trends*. UNESCO, Paris and Jessica Kingsley, London.
- Cowen, R., 2000. The market-framed university: the new ethics of the game. In: Cairns, J., Gardner, R., Lawton, D. (Eds.), *Values and the Curriculum*. Woburn, London.
- Cowen, R., 2009. Then and now: unit ideas and comparative education. In: Cowen, R., Kazamias, A. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.
- Cowen, R., 2014. Ways of knowing, outcomes and “comparative education”: be careful what you pray for. *Comp. Educ.* 50 (3), 282–301.
- Cowen, R., 2020a. Joseph Lauwers 1902–1981. In: Phillips, D. (Ed.), *British Scholars of Comparative Education: Examining the Work and Influence of Notable 19th and 20th Century Comparativists*. Routledge, London.
- Cowen, R., 2020b. Comparative, international, and transnational histories of education. In: Fitzgerald, T. (Ed.), *Handbook of Historical Studies in Education*. Springer, Singapore.
- Cowen, R., 2021. Educated identity: concepts, mobilities, and imperium. In: Klerides, E., Carney, S. (Eds.), *Identities and Education: Comparative Perspectives in Times of Crisis*. Bloomsbury Academic, London.
- Cowen, R., 2022. Nations in the world: interpreting the world Yearbooks. In: Tröhler, D., Piattoeva, N., Pinar, W. (Eds.), *World Yearbook of Education 2022: Education, Schooling and the Global Universalization of Nationalism*. Routledge, Taylor and Francis, London and New York.
- Cowen, R., Kazamias, A.M., 2009. *International Handbook of Comparative Education*. Springer, Dordrecht.
- Crossley, M., 1999. Reconceptualising comparative and international education. *Compare* 29 (3), 249–267.
- Eckstein, M., Noah, H., 1969. *Scientific Investigations in Comparative Education: An Anthology Illustrating the Strategy and Tactics of Comparative Education*. Collier-Macmillan, London.
- Epstein, E.H., 2008. Crucial benchmarks in the professionalization of comparative education. In: Wolhuter, C., Popov, N., Manzon, M., Leutwyler, B. (Eds.), *Comparative Education at Universities World Wide*. Bureau for Educational Services, Sofia.
- Epstein, E.H., 2016. Why comparative and international education? Reflections on the conflation of names. In: Kubow, P.K., Blosser, A.H. (Eds.), *Teaching Comparative Education: Trends and Issues Informing Practice*. Symposium Press, Oxford.
- Epstein, E.H., 2020. *North American Scholars of Comparative Education: Examining the Work and Influence of Notable 20th Century Comparativists*. Routledge, London.
- Fraser, S., 1962. *M.A. Jullien’s Plan for Comparative Education, 1816–1817*. George Peabody College, Nashville, Tennessee.
- Fraser, S., Brickman, W.W., 1968. *A History of International and Comparative Education: Nineteenth-Century Documents*. Scott, Foresman, Glenview, IL.
- Gerth, H., Wright Mills, C., 1970. *From Max Weber: Essays in Sociology*. London. Routledge and Kegan Paul, London.
- Grek, S., 2009. Governing by numbers: the PISA “effect” in Europe. *J. Educ. Pol.* 24 (1), 23–37.
- Halls, W.D., 1973. Culture and education: the culturalist approach to comparative education. In: Edwards, R., Holmes, B., Graff, V.D. (Eds.), *Relevant Methods in Comparative Education*. UNESCO Institute for Education, Hamburg.
- Halls, W.D., 1990. *Comparative Education: Contemporary Issues and Trends*. UNESCO, Paris and Jessica Kingsley, London.
- Hans, N., 1949. *Comparative Education: A Study of Educational Factors and Traditions*. Routledge, London.
- Holmes, B., 1984. Paradigm shifts in comparative education. *Comp. Educ. Rev.* 28 (4), 584–604.
- Jones, P.E., 1971. *Comparative Education: Purpose and Method*. University of Queensland Press, St Lucia.
- Kandel, I.L., 1933. *Studies in Comparative Education*. Houghton Mifflin, Boston.
- Kazamias, A.M., 2009. Forgotten men, forgotten themes: the historical-philosophical-cultural and liberal humanist motif in comparative education. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.
- King, E.J., 1958. *Other Schools and Ours*. Rinehart and Winston, New York.
- Kubow, P., Fossum, P.R., 2003. *Comparative Education: Exploring Issues in International Context*. Merrill/Prentice Hall, Upper Saddle River, NJ.
- Lawn, M., Grek, S., 2012. *Europeanizing Education: Governing a New Policy Space*. Symposium Books, Oxford.
- Li, X., Auld, E., 2020. A historical perspective on the OECD’s “humanitarian turn”: PISA for development and the learning framework 2030. *Comp. Educ.* 56 (4), 503–521.
- Madeira, A.I., 2008. Comparing colonial education discourses in the French and Portuguese African Empires: an essay on hybridization. In: Cowen, R., Kazamias, A. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.

- Mallinson, V., 1957. *An Introduction to the Study of Comparative Education*. Heinemann, London.
- Manzon, M., 2011. *Comparative Education: The Construction of a Field*. CERC the University of Hong Kong, Hong Kong, and Springer, Dordrecht.
- Manzon, M., 2019. *Origins and Traditions in Comparative Education*. Routledge, Taylor and Francis, Oxford.
- Masemann, V., Bray, M., Manzon, M., 2007. *Common Interests, Uncommon Goals: Histories of the World Council of Comparative Education Societies and Its Members*. Comparative Education Research Centre, The University of Hong Kong, Hong Kong, and Springer, Dordrecht.
- McLeod, J., Sobe, N.W., Seddon, T., 2018. *Uneven Space-Times in Education: Historical Sociologies of Methods and Practices*. *World Yearbook of Education 2018*. Routledge, Taylor and Francis, London and New York.
- Noah, H., Eckstein, M., 1969. *Toward a Science of Comparative Education*. Macmillan, New York.
- Nóvoa, A., Yariv-Mashal, T., 2003. Comparative research in education: a mode of governance or a historical journey? *Comp. Educ.* 39 (4), 423–438.
- Nóvoa, A., 2018. Comparing Southern Europe: the difference, the public, and the common. *Comp. Educ.* 54 (4), 548–561.
- Palomba, D., Cappa, C., 2018. Comparative studies in education in southern Europe. *Comp. Educ.* 54, 4.
- Paulston, R., 2009. Mapping comparative education after postmodernity. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.
- Phillips, D., 2020a. *British Scholars of Comparative Education: Examining the Work and Influence of Notable 19th and 20th Century Comparativists*. Routledge, Oxford.
- Phillips, D., 2020b. Michael Sadler 1861–1943. In: Phillips, D. (Ed.), *British Scholars of Comparative Education: Examining the Work and Influence of Notable 19th and 20th Century Comparativists*. Routledge, Oxford.
- Phillips, D., Schweisfurth, M., 2008. *Comparative and International Education: An Introduction to Theory, Method, and Practice*. Continuum, London.
- Rizvi, F., Lingard, B., 2009. The OECD and global shifts in education policy. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer, Dordrecht.
- Sadler, M., 1900. How far can we learn anything of practical value from the study of foreign systems of education? *Comp. Educ. Rev.* 7 (3), 307–314.
- Schriewer, J., 2000. Comparative education methodology in transition: towards a science of complexity? In: Schriewer, J. (Ed.), *Discourse Formation in Comparative Education*. P. Lang, Frankfurt am Main.
- Sousa, D., Moss, P., 2022. Comparative studies in early childhood education: past, present and future. *Comp. Educ.* 58, 3 (Special Issue).
- Steiner-Khamsi, G., Waldow, F., 2012. *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. Routledge, London.
- Takayama, K., 2018. Beyond comforting histories: the colonial/imperial entanglements of the International Institute, Paul Monroe, and Isaac L. Kandel at Teachers College, Columbia University. *Comp. Educ. Rev.* 62 (4), 459–481.
- Takayama, K., Sriprakash, A., Connell, R., 2017. Toward a postcolonial comparative and international education. *Comp. Educ. Rev.* 61 (1), 1–24.
- Tikly, L., 1999. Postcolonialism and comparative education. *Int. Rev. Educ.* 45 (5/6), 603–621.
- Trethewey, A.R., 1976. *Introducing Comparative Education*. Pergamon Press, Rushcutters Bay, Australia.
- Tröhler, D., 2013. The OECD and Cold War culture: thinking historically about PISA. In: Meyer, H.-D., Benavot, A. (Eds.), *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Symposium, Oxford.
- Tröhler, D., Piattoeva, N., Pinar, W., 2022. *World Yearbook of Education 2022: Education, Schooling and the Global Universalization of Nationalism*. Routledge, Taylor and Francis, London and New York.
- Welch, A., 2001. Globalisation, post-modernity and the state: comparative education facing the third millennium. *Comp. Educ.* 37 (4), 475–492.
- Wilson, D.N., 1994. Comparative and international education: Fraternal or Siamese twins? A preliminary genealogy of our twin fields. *Comp. Educ. Rev.* 38 (4), 449–486.
- Wolhuter, C., Popov, N., Manzon, M., Leutwyler, B., 2008. *Comparative Education at Universities World Wide*. Bureau for Educational Services, Sofia.

World-systems analysis

Tom G. Griffiths^a and Marcin Starnawski^b, ^a Oslo Metropolitan University, Oslo, Norway; and ^b University of Lower Silesia, Wrocław, Poland

© 2023 Elsevier Ltd. All rights reserved.

World-systems analysis: origins	19
World-systems analysis: major concepts	20
World-systems analysis in educational research	23
World-systems analysis and education for a just, utopistic, transition	25
References	26

World-systems analysis: origins

World-systems analysis (WSA) developed within a wider tendency of social science scholarship in the later part of the 20th century working to integrate research on contemporary realities with the study of history. Historical and comparative methods featured prominently early on in the late 19th and early 20th century works of political philosophers, sociologists, economic historians, and state theorists. Still, an integrated analysis of the past and the present, the material and the symbolic, structure and agency, etc., was long overshadowed by disciplinary particularisms and epistemological divergences leading to supposed gaps between “explanation” and “interpretation,” or nomothetic and idiographic perspectives generally reflecting divisions between humanities and “proper” social sciences. For example, mainstream Western sociology around the 1950s was exemplified by a turn toward, on the one hand, abstract models (“grand theory”) that were difficult to operationalize, and, on the other hand, empirical approaches that prioritized statistical explanations of the present phenomena over other accounts, including those based on knowledge of long-term processes and qualitative inquiry. It was largely against these tendencies that a more comprehensive approach was famously reclaimed by Mills (1959, p. 6) in the statement that the “sociological imagination enables us to grasp history and biography and the relations between the two within society.” The proponents of *historical sociology* in the later decades attempted to bridge diachronic and synchronic dimensions of social analysis, and it is this field that WSA is often associated with.

Wallerstein (2004a, pp. 7–15) pointed to a number of developments after World War II, both socio-political changes and vital scholarly debates, which formed major contexts that inspired the shaping of WSA. All these changes resulted in questioning of the structures of social sciences that had crystallized by the early decades of the 20th century, and were in fact the previous century’s ideological legacy that divided research and knowledge disciplines in several ways. One division was between disciplines studying the “West” and “the Rest” (the non-West), the former coming with a focus on the past (history) and on the present, further marked by split between the market (economy), the state (political science) and the society (sociology). Separate from this was study of the “non-West,” organized into the disciplines of anthropology, exploring cultures viewed as non-literate or “primitive,” and oriental studies examining texts and artifacts of, mostly Asian, regions dominated by the non-Christian world religions and former empires.

The shifting realities of the post-1945 world that challenged the established view of social sciences included, firstly, the hegemony of the United States and the growth of *area studies*, focused on the notion of development and modernization as a challenge and a major stake in the cold war rivalry between “the First World” (capitalist and liberal-democratic West) and “the Second World” (the state-socialist bureaucratic “people’s democracies” mostly dominated by the Soviet Union). Secondly, the need for new knowledge and new political perspectives was prompted by the rise of anti-colonial movements, the decline of Western colonialism, and the formation of “the Third World” (largely overlapping with the Non-Aligned Movement). In fact, it was Africa of the independence era, including countries such as Ghana and the Ivory Coast, that drew early scholarly interest of Wallerstein (Wallerstein, 1961, 1964, 1967, 1986; Rich and Wallerstein, 1972). The third factor was the expansion of the university system and the growth in the number of scholars who increasingly borrowed from the related disciplines and shifted interests either toward topics demanding cross-disciplinary expertise, or away from their traditional focus (e.g. anthropologists turning to their own countries).

The significant scholarly debates that set the scene for the emergence of world-systems analysis included discussions among Marxists on the role of transnational factors in the origins of capitalism, as well as economic histories of non-European empires. These debates manifested beyond Marxist orthodoxy and coincided with the growing popularity of the French school of historiography (associated with the *Annales* journal) that sought to analyze long-term structures and patterns (*longue durée*) beyond narrow empiricism of political history. Of particular importance for WSA were the works of Fernand Braudel of the *Annales* school with his use cross-disciplinary inquiry, focus on different categories of social time, problematization of the notions of “market” and “capitalism,” and the treatment of *world-economy* as the unit of analysis (Wallerstein, 2001a,b; Braudel, 1982, 1985). Finally, and perhaps most importantly, WSA drew on dependency theory developed in the 1960s that was grounded in the analyses undertaken in the late 1940s by experts from the United Nations Economic Commission for Latin America (Prebisch, 1963). The main tenets of the

theory identified patterns of economic domination of the stronger countries over the weaker ones, i.e. obtaining surplus value by “center” from “peripheries.” This was further developed into notions of the “development of underdevelopment” as a result of international capitalist dynamics (Frank, 1978, 2009) and unequal exchange (Emmanuel, 1972) as opposed to the classical concept of comparative advantage.

As Wallerstein emphasized, WSA is not a finished theory (“not a theory about the social world, or about part of it”; Wallerstein, 2001c, p. 237), but an analytical perspective and method to use in empirical studies of various sorts and, possibly, on which to build theory-oriented reflection. WSA can thus be seen as a heuristic framework drawing on historiographic accounts and diverse conceptual tools, and as such it constitutes a basis for a range of thematically specific research programs. World-systems analysis transgressed traditional disciplinary divisions, and rejected the supposed incongruity of the idiographic and the nomothetic epistemologies, seeking a more inclusive and “unidisciplinary” domain that can be called *historical social science* (Wallerstein, 2001c, 2004a). The emergence and growth of WSA represented an intellectual protest within social science

against the ways in which social scientific inquiry was structured for all of us at its inception in the middle of the nineteenth century. ... World-systems analysis maintains that this mode of social scientific inquiry, practiced worldwide, has had the effect of closing off rather than opening up many of the most important or the most interesting questions. ... World-systems analysis was born as moral, and in its broadest sense, political, protest. However, it is on the basis of scientific claims, that is, on the basis of claims related to the possibilities of systematic knowledge about social reality, that world-systems analysis challenges the prevailing mode of inquiry.

Wallerstein (2001c, p. 237).

There were several targets of this “protest.” Firstly, modernization theory and its universal concept of development were contested for their determinism, linearity, and ideological premises. The latter echoed earlier West-centric notions of “progress” (as in Enlightenment’s philosophies) and “evolution” (as in some variants of positivist sociology and anthropology). Secondly, WSA challenged the dominant methodology that equated “society” with “nation-state,” and debated whether these terms were adequate units of analysis for historical social research. These were replaced by the term of *historical system*, or *historical social system*. Thirdly, WSA called for overcoming disciplinary divisions among social sciences as well as the gap between sciences and humanities as “two cultures” (Snow, 2002), and it also rejected both ahistorical approaches and narrow views of history (i.e. history as a sequence of events, mainly political ones, and based exclusively or mostly on archived written documents testifying on such events). Fourthly, WSA criticized social science for its avoidance of discussions on moral and political dimensions and consequences of scholarly endeavors. This criticism was evident in challenging modernization theory and in turning to analyses, such as dependency theory, that had been linked to political efforts by experts and scholars outside the Global North. Two publications marked the beginning of WSA, both published by Wallerstein in the same year: an overall article on the world capitalist system and the first volume (out of four) of his *opus magnum*, devoted to the origins of capitalism in the 16th century (Wallerstein, 1974a,b).

Institutionally, WSA was promoted by the Fernand Braudel Center at the Binghamton University (<https://www.binghamton.edu/fbc/>), operating between 1976 and 2020 within a scheme of conferences, research groups and a publication program. The research working groups focused on world-systems structures and processes (e.g. structural trends, crisis, hegemonies, world labor, antisystemic movements, cultural forms), regional historical patterns (e.g. Atlantic slavery, economic development in Americas, East Asia) and issues of social science knowledge (methodology across the disciplines, structures of knowledge and categories of social knowledge as well as “utopistics”) (see Report, n.d.). While the center’s website provides a list of over forty books or thematic journal issues published by the late 1990s by scholars associated with it, the actual output of WSA-related works is much greater. The journal which functioned as a key forum for WSA scholarship was *Review*, published by the Fernand Braudel Center between 1977 and 2016 (thirty-nine volumes; for contents, subjects and abstracts see <https://www.binghamton.edu/fbc/review-journal/index.html>). Another major outlet is a biannual cross-disciplinary *Journal of World-Systems Research*, initiated in mid-1990s by Christopher Chase-Dunn and hosted by the University of Pittsburgh, and since 2008 and official journal of the Political Economy of the World-System section of the American Sociological Association. Twenty-seven volumes have been published by 2021, with a purpose of theory-building “to take the world-systems approach beyond the stage of a perspective” (<https://jwsr.pitt.edu/ojs/jwsr/about>).

World-systems analysis: major concepts

The central concept in WSA is that of *historical (social) system* (Wallerstein, 2001d, 2004a). The term is not only an abstract reconciliation “between the static and the dynamic, the synchronic and the diachronic” of “long-term entities,” but rests on defining characteristics. These include the relative functional autonomy of systems with unique internal processes, that is, “life within it is largely self-contained, and ... the dynamics of its development are largely internal” (Wallerstein, 1974b, p. 347), temporal boundaries (beginning and end), and spatial boundaries (changeable over time). Besides *mini-systems* that once constituted basic entities of such kind, the major type of historical (social) system is *world-system*, which includes

two main types: *world-empires* and *world-economies* (Wallerstein, 2001d, pp. 229–233). The distinction consists in the fact that a world-empire is a single political unit, that is, politically centralized, with cultural and ethnic plurality, and a world-economy that is characterized by political plurality of statehoods as much as with ethno-cultural plurality. The hyphenated form of the term indicates that

we are talking not about systems, economies, empires of the (whole) world, but about systems, economies, empires that are a world (but quite possibly, and indeed usually, not encompassing the entire globe). ... in “world-systems” we are dealing with a spatial/temporal zone which cuts across many political and cultural units, one that represents an integrated zone of activity and institutions which obey certain systemic rules.

Wallerstein (2004a, pp. 16–17).

The *modern world-system* has its beginnings in the “long 16th century,” and takes the form of *capitalist world-economy* (Wallerstein, 2001e, pp. 267–268). It is the first historical system to eventually encompass the entire globe in our times, although for much of its history there remained external zones that were successively incorporated into it. That this world-economy is *capitalist* is defined not by the market mechanism as such—which, following Braudel, Wallerstein treats as a form of economic organization predating capitalism—but on the principle or imperative of endless accumulation of capital, a factor that forced the system’s spatial expansion and secured its existence for the past 500 years. Another crucial characteristic of the capitalist world-economy is an axial division of labor across the multiple polities, the transfer of surplus value via unequal exchange between core and periphery, and the existence of semiperipheral zones with both core-type and periphery-type processes. This structural dynamic is manifest on a political plane in that such a positioning characterizes single states or groups of states, or regions within them. Nonetheless, while “centrality” and “peripherality,” and the relationship between them, are spatial, the term “core–periphery” should not be simply reduced to geographic distinctions. Furthermore, the capitalist world-economy is composed of institutions: markets, firms, households as well as hierarchies of social classes and status groups or identities (echoing Weber’s distinction of *Klassen* and *Stände*).

The scope of the capitalist world-economy overlaps with that of an *interstate system* of sovereign states, some of which are hegemonic. From the period between the French Revolution and the revolutions of 1848 onwards, the modern world-system has been also an arena of rivaling ideologies: conservatism, liberalism, and a range of radical (left-wing) ideologies. The latter produced anti-systemic movements that historically have simultaneously undermined and reinforced the system (Wallerstein, 2014). The rivalry of ideologies led to the shaping of what Wallerstein termed *geoculture*, a shared cultural framework of norms and discourses across diverse polities and ostensibly competing ideologies. The geoculture of the post-WWII period, labeled “centrist liberalism” (Wallerstein, 2011), reflected a consensus between conservative, liberals and the “old Left” about progressive (over time) change, progress, and development, via state power and legislation “to change the world.”

The geoculture rested on three pillars: (1) the inclusion of populations into citizenship, with civil and political rights and freedoms not always de facto granted but at least accepted in principle, as reflected in the nation-states involvement in the creation of the UN human rights system; (2) the expansion of socio-economic rights, again to varying degrees, but manifested either in the Western-welfare state or in state-socialist variants of limited egalitarianism; (3) nationalism, that is, the unification of citizens into a nationhood possessing or aspiring to a nation-state. The revolts of 1968 marked the emergence of the “new Left” against both US hegemony and Soviet-style socialism (and the established workers’ movements), the turn toward race, gender and sexuality as political issues, democratization movements in socialist states and later environmentalism. The 1968 “world revolution,” seen as a marker in the declining legitimacy of the dominant geoculture of liberalism, was also an international crisis that led to the neoliberal phase of capitalist expansion that further exacerbated structural tensions of the world-economy (Wallerstein, 1999a, 2003a, 2004a).

Finally, the system has dynamic features that are both short/mid-term (conjunctural cycles or “waves” of growth and contraction as conceptualized by Kondratieff (1984)) and long-term (secular trends). The existence of cyclical rhythms and secular trends creates a paradox: the capitalist world-economy, as any historical system, changes constantly, yet its fundamental features remain the same. Nonetheless, the secular trends—long-term increase in wages, taxation, and externalization of costs in production, coupled with the shrinking of non-proletarian, mostly rural, reservoir of labor force—signal the structural crisis of the capitalist world-economy. This structural crisis begs the question of possible directions that such a transition takes toward a new system to replace capitalism, and possible choices or “utopistics” (Wallerstein, 1998a) available to individuals and movements to achieve a more democratic, just, and more egalitarian world-system.

WSA has been criticized from different angles: positivist or neopositivist (for shortage of quantitative data as well as insufficient reduction of clear variables and weak verifiability of hypotheses), orthodox Marxist (for being circulationist, overlooking production in accumulation processes and not emphasizing class struggles within nation-states), state-autonomist (for economic reductionism and downplaying the role of political agencies on the state and interstate levels), and culturalist particularism (for economism and providing a totalizing and Eurocentric “grand narrative” that ignores plurality of cultural identities). Wallerstein (2004a, pp. 19–22) recounted these major criticisms and highlighted their weak points or even their misreading of WSA. Firstly, world-systems analysts are not against quantification but rather opt for methods and data adequate to the task of explaining long-term and broad-scale changes. Secondly, WSA sees wage (productive) labor among many labor forms in the capitalist world-economy (Wallerstein, 2001f), and points to the fact that class conflict should be considered within a wider framework of the world-system, for which core–periphery relations are a vital feature. Thirdly, WSA does not neglect the state but rather posits

that states form an interstate system that is an integral element in shaping the historical processes of capitalism. And lastly, drawing on Braudel and other proponents of “total” historiography, WSA includes culture and identity among the components of the capitalist world-economy.

That WSA constitutes a “grand narrative” is inevitable and typical for any product of knowledge-building. The fact that WSA tracks the history of capitalist expansion and domination that initially led to the formation of the “European world-system” (Wallerstein, 1974b) is not in itself Eurocentric but reflects knowledge on long-term and large-scale processes. Moreover, Wallerstein (2004a) noted that:

What these four critiques have in common is the sense that world-systems analysis lacks a central actor in its recounting of history. For nomothetic positivism, the actor is the individual, *homo rationalis*. For orthodox Marxism, the actor is the industrial proletariat. For the state-autonomists, it is political man. For cultural particularists, each of us (different from all the others) is an actor engaged in autonomous discourse with everyone else. For world-systems analysis, these actors, just like the long list of structures that one can enumerate, are the products of a process. They are not primordial atomic elements, but part of a systemic mix out of which they emerged and upon which they act (p. 21).

A critical feature of WSA is the relationship between the universal and the particular, whether in terms of the historical structures of knowledge in accordance with nomothetic and idiographic epistemologies (Wallerstein, 2004b), or the dynamics between hierarchical constructions of universal and particular cultures (e.g. Wallerstein, 1990) and competing particularist claims. WSA consistently identified and elaborated the phenomenon of European universalism (Wallerstein, 2006) including in terms of how the shared cultural framework of the capitalist world-economy was based on European scientific universalism, its associated conceptions of development and progress, such that these particular world-views were given universal status. Moreover, Wallerstein wrote repeatedly about the struggles against Eurocentrism and its manifestations in the social sciences, and how within the structures of the capitalist world-economy such struggles against Eurocentrism may inadvertently reinforce it (Wallerstein, 1997b).

Many insights of WSA have been taken by authors not identifying with this particular framework but analyzing how Eurocentric or North-dominated epistemologies misrepresented histories and experiences of regions, countries, peoples and nations located beyond the “core” zones (e.g. Santos, 2016; Mignolo, 2021). A moderate criticism has been raised against Wallerstein’s and especially Braudel’s West-centrism in explaining the origins of capitalism, and even the very use of the term, something they share with classical authors such as Marx and Weber (Goody, 2006). On a still more critical note, Connell (2020, pp. 67–68) places WSA among other strands of “metropolitan social science,” which represent “the northerness” in that “a system of categories is created by metropolitan intellectuals and read outwards to societies in the periphery, where the categories are filled in empirically. ... [Such approaches] treat the majority world as object. This closes off the possibility of social science working as a shared learning process, a dialog, at the level of theory” (Connell, 2020, pp. 66–68). Her critique of WSA consists of two major points. Firstly, she exposes “the short list of social actors ... bourgeoisies, proletariats, intermediate classes, state elites, and not much more” adding that “[t]he difficulty this framework has in dealing with issues of gender and race—constitutive structures of colonialism, on other reckonings—is familiar” (Connell, 2020, p. 67). This critique is concise and lacks elaboration, but reflects the culturalist charges mentioned above.

While WSA is not gender studies or a theory of race, it integrates race/racism and gender/sexism within its framework (Wallerstein, 2004a) as demonstrated by analyses of households (Smith and Wallerstein, 1992), a number of studies on non-metropolitan cultural and political realities including social movements (Wallerstein, 1999b, 2002a,b, 2003b) and a multi-thematic preoccupation of diverse research working groups at the Fernand Braudel Center with these issues and with the dynamics of knowledge production relevant to them. Wallerstein’s interest in Africa, cited earlier, and his particular interest in Fanon (e.g. Wallerstein, 2000b, 2009) are examples of WSA drawing on non-metropolitan knowledges. Wallerstein (2004a) placed Fanon’s (1963) *The Wretched of the Earth* in the “Bibliographical Guide” among “Relevant Works: Forerunners or Influential Writings of Other Large-Scale Analysts” (p. 104). Wallerstein (2009) concludes that Fanon “offers a brilliant delineation of our collective dilemmas. Without violence the wretched of the earth can accomplish nothing. But violence, however therapeutic and however effective, solves nothing. Without breaking from the domination of pan-European culture, it is impossible to move forward” (pp. 124–25). Moreover, being an analytical perspective and not a complete theory (a closed system of concepts and prepositions), WSA is not incompatible with empirical and theoretical developments in research on gender, sexuality, race and ethnicity, nationalities and other culturally-grounded categories or identities. To the contrary, WSA leaves plenty of room for detailed empirical analyses and novel theoretical constructs, often inspired by the world-systemic framework rather than reproducing its full conceptual toolkit.

Connell’s second critical point is even less clear. She pointed to a “powerful tendency to reification” in WSA scholars’ use of the concept of social system (actually, Wallerstein talks about *historical [social] systems*) and other related concepts such as “a boundary, internal differentiation of functions, subsystems, and so on,” a tendency that “led one current of world system research to devise models of the pulsations within a capitalist world-economy, and led another into a trawl through history for world-systems ... It led quite early to the idea that the future could be inferred from world-systems’ laws of motion. ... the once-fruitful research paradigm has turned into portentous speculation” (Connell, 2020, p. 68). The reification charge seems off point

not only because it would have to be directed against *any* concept in social science, but first and foremost because WSA is fundamentally a constructivist perspective (though usually not presented as such), which is demonstrated in its treatment of time and space:

it must be emphasized that for world-systems analysts, time and space—or rather that linked compound TimeSpace—are not unchanging external realities which are somehow just there, and within whose frames social reality exists. TimeSpaces are constantly evolving constructed realities whose construction is part and parcel of the social reality we are analyzing. The historical systems within which we live are indeed systemic, but they are historical as well. They remain the same over time yet are never the same from one minute to the next. This is a paradox, but not a contradiction. The ability to deal with this paradox, which we cannot circumvent, is the principal task of the historical social sciences. This is not a conundrum, but a challenge.

Wallerstein (2004a, p. 22).

Furthermore, Wallerstein's idea of *utopistics*, that is, a systematic study of possibilities and choices in the period of systemic crisis and transition, is not a speculation but a call for a social science that treats human dilemma seriously and does not avoid considering moral (and political) stakes within the activities of knowledge-building. And while it is but one of the many themes of the methodological orientation of WSA, it not only facilitates opening of the social sciences to various non-dominant and non-metropolitan modes of thinking, but also possesses a great educational potential in encouraging new directions in public imagination.

World-systems analysis in educational research

The usefulness of WSA in educational research has not been immediately obvious. The work of WSA scholars has primarily focused on historical trajectories and the contemporary crisis of the modern world-system, analyzing broad economic and political dynamics. Little attention has been given to, for example, the *longue durée* of educational ideas and institutions as part of the capitalist world-economy, or to the role of antisystemic movements in shaping and transforming educational spaces. Notable exceptions include the critical discussion on the formation and institutionalization of the social sciences and on the effects of the revolts around 1968 on Western university curricula, as well as some works on core-periphery patterns in academic knowledge production and circulation (Schott, 1998; Demeter, 2019; Warren et al., 2020). Nonetheless, in WSA scholars' reflections on social sciences, education is not mentioned among the major disciplines developed in the 19th and 20th centuries, despite its emergence by the early 1900s from philosophy, sociology and psychology, and its subsequent, however short-lived, international flourishing both as a scientific endeavor and a democratization and peace effort during the era of the New Education movement after World War I. As discussed below, the use of WSA has been greatest in the broad field of Comparative and International Education (see for example Arnove, 1980). Used in this way, WSA framework helps to see education as a cross-disciplinary research effort, a socially engaged practice (i.e. committed to changes on various levels, from the institutional system to human interactions in the learning process) and as a transnational field of ideas, much in line with the late 20th century calls for open and pluralistic unity in social and human sciences (cf. Wallerstein et al., 1996). Within labor sociology, worth mentioning is Silver's (2003) study partly inspired by WSA, which discusses protests in "the education industry" within a global and historical perspective. Her work takes a particular focus on the context of post-Fordist changes and the "knowledge-based economy," pointing to proletarian conditions of "education workers (teachers) [who] are central to processes of capital accumulation in the 21st century" (p. 114). An outlet occasionally publishing WSA-related educational research is the *Journal for Critical Education Policy Studies* (<http://www.jceps.com/>).

How is and can WSA, as a macro account of the historical development and trajectory of a single capitalist world-economy encompassing multiple polities across a hierarchical global division of labor, be applied to educational research? What insights can it offer to the study of educational systems, their organizational, curricular, and pedagogical forms? Like other macro-level theorizing such as that under the banner of globalization in which some locate world-systems analysis (e.g. Sklair, 1999), the task of bringing WSA to bear on analyses of educational phenomena, and operationalize its application, brings certain challenges.

WSA directly connects with efforts within comparative and international education, and across the social sciences, to overcome "methodological nationalism" (Dale, 2005, p. 120) manifest in research focused on single national-case studies. It also adds the historical dimension to the study of the present educational systems, models, and values as an effect of long-term societal patterns, not limited to most recent policy changes and reforms. As outlined above, part of WSA's protest against established concepts and methodologies in the social sciences includes this dimension, locating analyses of phenomena within nation-states in their inclusion and trajectory within the capitalist world-economy. The concept of globalization invites similar methodological approaches, incorporating the influence of trans or supra-national forces on national and local phenomena. Wallerstein (2000a) provocatively argued that the discourse of globalization studies was "in fact a gigantic misreading of current reality—a deception imposed upon us by powerful groups and an even worse one that we have imposed upon ourselves, often despairingly" (p. 250). The provocation here was that globalization studies misread phenomena like the growth and power of international institutions and of transnational corporations, for example, without reference to their longer historical development as part of the modern world-system. In so doing, it deflected attention from the actual systemic crisis of the capitalist world-economy, and the potentials for human agency to influence the shape and nature of its transition toward an uncertain alternative.

Leaving aside the question of systemic crisis and transition for one moment, the general point for educational research is that the study of national systems of education, including their core structures and purposes, is enhanced by reference to the capitalist world-economy in which the nation-state is historically and structurally located. It is worth noting that a similar approach had been employed before, without direct reference to WSA, by Hobsbawm (1987) who pointed to various dynamics in the expansion of education in the late 19th and early 20th centuries. These included its function of composing and recomposing class stratification, its role in nationalisms of both liberating nations and oppressive nation-states, grassroots education in workers' movements, its emancipatory role for women, and a general progress associated with literacy that still revealed core-periphery discrepancies between states, regions, and racial groups. While Hobsbawm's historiographic account does not operate within WSA conceptual framework, his take on capitalism parallels that proposed by Wallerstein who clearly drew on Hobsbawm's idea of the "long 19th century" stretching between 1789 and 1914 (cf. Hobsbawm, 1962, 1975, 1987; Wallerstein, 2011).

Within the field of comparative and international education, WSA-framed educational research is also a point of departure from the "neo-institutionalist" branch of research also referred to as "world culture theory," that explains identified convergence in educational policies across the world in terms of an institutionalized world culture of the modern nation-state, its forms and ideas of national progress, extending to models and contents of mass education systems (see for example Baker, 2014). The world culture perspective downplays or excludes both the economic imperatives of capitalism as a world-economy, and the power-dynamics associated with its effective functioning, discounting them as significant factors in accounting for the spread of mass education systems (see for example Griffiths and Armove, 2015).

In contrast, WSA offers a more comprehensive framework for contextualizing, and understanding, single and comparative case-studies. At an important level, WSA offers a casual account of phenomena like the global expansion of systems of mass education, tying this phenomenon to nation-states' operation within the capitalist world-economy, and the interstate system. Specifically, WSA highlights nation-states efforts to advance projects of national economic development, whether efforts to "catch up to the wealthy countries" (Wallerstein, 2000a, p. 263) by states located in the periphery or semi-periphery; or the ongoing programs of endless, linear, growth and development in states within the core working to maintain their relatively privileged position. For educational research, the immediate connection is to the well documented conceptualization of education by international policy makers and institutions, and their national and local counterparts, in terms of the human capital formation deemed as required for national economic growth and development. This application of human capital theory by policy makers aligns with orthodox modernization theorizing projecting universal paths of (capitalist) national economic development, narrowing the focus and scope of mass education to such instrumental demands (see for example Klees, 2008).

The substance of this sort of macro-account of the phenomenon of mass education, its character and official purposes, can be fleshed out in terms of WSA's articulation of a shared geoculture or cultural framework of "centrist liberalism" (Wallerstein, 2011) across diverse political systems during the past century of the modern world-system. This shared cultural framework included a "developmentalist" logic related to modernization theory. It extended to a shared belief across diverse and ostensibly competing political systems, in the ability of policy makers, through the exercise of state power, to make use of scientific and technological knowledge and advances in (modernizing) projects of national economic development. In ways not dissimilar to Rostow's (1959) universal claims, the shared assumptions of the geoculture presented national development in these terms as universally possible, and capable of delivering the promise of material abundance for all at some point in the not too distant future (e.g. Wallerstein, 2005). Analyses of efforts by institutions like the World Bank to present education in terms of these universal development promises, and measure the economic return (to the nation) on investments in education, logically follow.

Educational research from this perspective highlights and builds on the identified shared aspirations of historical socialist and capitalist states, that arguably continues post-1989, to further pursue this agenda. Despite failings, the geoculture of the world-economy maintains that all nations can achieve desired levels of national economic growth and development, toward "a more prosperous, more egalitarian world" (Wallerstein, 2000a, p. 262), toward a promised utopia of material well-being and abundance for all. The particular perspective of WSA is important here, in which Wallerstein identifies the collapse of historical socialism (as a mid-term process between 1956/1968 and 1989/1991) as a marker of the demise of centrist liberalism as the dominant geoculture of the modern world-system, and so a marker in the crisis and period of transformation of the capitalist world-economy. The "end of liberalism" argument highlights the failures of both capitalist and historical socialist states and systems to deliver the universalized promises of endless growth, of improved social and material security, health, well-being, of abundant consumption for all (see for example Griffiths, 2009); attributing these failures to the systemic operation and requirements of the world-system and hierarchical distribution of surplus value across participating states in peripheral, semiperipheral, and core zones of the world-economy (Wallerstein, 1995, 1999a, 2005).

Importantly, educational research inspired by this perspective brings to the examination of educational phenomena the simultaneous crisis and a loss of legitimacy of states due, in part, to their failure to deliver the liberal promises, despite more than half a century of UN "decades of development" since the 1960s. For example, the phenomenon of "credential inflation" undermines the promised upward social mobility for individuals, through the attainment of educational credentials, coupled with global phenomena of de-industrialization, automation, and ongoing neoliberal prescriptions that reduce the role of the state in the provision of public services and so as an employer, underpinning structural unemployment (e.g. Collins, 2013). As Warmington (2015) argues from the United States context, "The fact is, however, that not everyone can learn their way out of poverty ... The economic system that makes some affluent is the same system that locks others into poor work" (p. 274).

WSA foregrounds these constraints on a world-system level—how the capitalist world-economy, and its "normal" operation, impose on educational policy goals like learning one's way out of poverty as the counterpart of human capital logic. WSA requires

researchers to present such policy ambitions as a structural impossibility for many, including increasing numbers and groups located within wealthy, core countries, and disproportionately those located in relatively poor, semiperipheral and peripheral contexts. That the promise continues to be offered and reflected in educational policy, and that people across the world continue to pursue improved social and economic conditions in part through participating in formal education, are phenomena worth investigating. But the persistence of the phenomena does not change the structural impossibility of the liberal promise, without a structural change to the operating principles of the capitalist world-economy.

In earlier work analyzing the use of WSA in a major comparative education journal, we found its use in terms of supporting the shift away from “methodological nationalism” as raised above, and building on dependency theory, but with some critiques of WSA that we argued missed the mark (Griffiths and Knezevic, 2010). In particular, WSA was positioned by some as overly deterministic, and so as excluding opportunities for local agency over educational policy and practice. Accounts of its theorizing tended to position WSA as a now redundant point in the linear development of theorizing in the social sciences, overtaken by more contemporary approaches. Such accounts fail to acknowledge WSA’s consistent emphasis on the constant interplay between structure and agency within particular historical conditions. Moreover, they fail to engage with WSA’s argument that the period of systemic crisis and transition that we are living in provides heightened possibilities for human agency to influence the course of history (Wallerstein, 1997a). Crucially, some of the characterizations of WSA present it as a static “theory,” failing to understand it as a knowledge movement and protest against established structures of knowledge that work to support existing systemic injustices and inequalities.

Importantly for educational research, WSA directly connects with well-established traditions of critical education and critical pedagogy. Wallerstein, for example, frequently conceptualized WSA in terms of politically engaged academic work, citing three key “tasks” of such work: (1) the intellectual task of understanding social reality; (2) the “moral” task of imagining alternatives that are more egalitarian, just, equal, democratic; and (3) the “political” task of conceptualizing and taking action to change reality in such egalitarian directions (Wallerstein, 1998a,b). The alignment with Freire’s concepts of developing critical consciousness through the analysis of existing reality and its injustices and inequalities, as a basis for taking action in the world to change the world, are clear and consistent in WSA (see for example Griffiths, 2019). WSA sets out to construct a social science that can “speak to the worldwide social transformation through which we shall be living” (Wallerstein, 1999a, p. 201).

World-systems analysis brings explicit political ambitions to educational research, including explorations of the potential for educational work to contribute to and influence the transition from capitalism to an uncertain alternative. Cho’s (2013) volume, for example, provides a good account of critical pedagogy that draws on Wallerstein’s work, elaborating the potential of critical pedagogy to further fulfill its “language of possibility,” and for WSA to inform this work in terms of defining and constructing an alternative education, including its purposes and potential contributions to the creation of a more egalitarian future.

As we suggest below, it is in this area that WSA’s take up/application in educational research holds most promise, notwithstanding the difficulties of operationalizing such a project. This application is based on a number of premises about capitalist world-economy in a period of crisis and transition, linked to a set of irresolvable contradictions or tendencies reaching their absolute limit, and so bringing this analysis to bear on our thinking about and practice in educational settings. In concrete terms, the WSA tasks of critical interpreting existing social problems and inequalities and their systemic character, and of imagining alternative and more egalitarian world-systems, are a guide for educational curricular and pedagogical practice. These are general principles that have, more or less, characterized critical pedagogy for decades. What does WSA specifically bring that makes its application to educational research necessary?

World-systems analysis and education for a just, utopistic, transition

In this chapter we have emphasized how WSA’s identification of centrist liberalism as the dominant geoculture of the capitalist world-economy has included its overt commitment to and focus on “economic development linked to social amelioration,” these being seen as “obverse sides of the same coin” (Wallerstein, 2011, p. 107). Regardless of their location and highly unequal states of “development,” the promise and aspiration were aligned and promoted, grounded in an expressed belief in the universal potential for such development. As Wallerstein (1995) observed:

The possibility of the (economic) development of all countries came to be a universal faith, shared alike by conservatives, liberals, and Marxists. The formulas each put forward to achieve such development were fiercely debated, but the possibility itself was not. In this sense, the concept of development became a basic element of the geocultural underpinning of the world-system.

Wallerstein (1995, p. 163).

So while this insight is not new, the ongoing influence and application to educational policy and practice remains, albeit under pressure in various ways linked to the structural failings of the “universal faith.” WSA foregrounds the manifestation of this promise in education systems as another indication of the weakening legitimacy of the geoculture and the nation-states and their institutions that continue to promote it.

This critical position is necessary now more than ever, in the face of the current global climate crisis, the incapacity of mainstream conceptions of capitalist “development” based on linear and endless growth and capital accumulation to resolve the climate crisis, and so the urgent need for educational interventions to contribute to and advance feasible alternatives. In [Wallerstein’s \(1998a\)](#) utopistic terms, this involves

... the sober, rational, and realistic evaluation of human social systems, the constraints on what they can be, and the zones open to human creativity. Not the face of the perfect (and inevitable) future, but the face of an alternative, credibly better, and historically possible (but far from certain) future (pp. 1–2).

The climate emergency, and fledgling global efforts to constrain carbon emissions and increases in global temperature, highlight the tensions in tackling this crisis within the constraints of the economic system and its cultural assumptions that has, through the systemic requirement of endless growth and capital accumulation, produced the climate crisis. One of the secular trends contributing to the crisis and period of transition of the capitalist world-economy has been capital’s externalization of environmental costs of production to the state, to maintain rates of profit and capital accumulation ([Wallerstein, 2013](#)). This trend added to seemingly impossible pressure on the state to meet these and other costs, alongside pressure to reduce taxation, and so arguably to some shifting of the costs of global warming and remediation efforts back to capital that has exacerbated pressures on profits and capital accumulation. Clearly the global climate crisis exacerbates such tensions, and has the potential to build support for the utopistic work for system change.

WSA provides an important and relevant perspective for current analyses associated with the global climate emergency, and efforts to achieve some sort of just settlement across space and time. Work like [Hickel’s \(2020\)](#) volume, for example, seeks to make the case for a post-capitalist (in his case, degrowth) future. The crucial point is that the argument is based on the incompatibility of capitalism’s logic of endless growth and capital accumulation which, even with the growth of renewable energies and efforts at carbon capture, cannot reduce emissions sufficiently to avoid irreversible climatic tipping points being passed. We argue that this and other important work (e.g. [Smith, 2016](#)) constitutes the sort of politically engaged work that Wallerstein and WSA consistently advocated. It locates fundamental causes of the current climate problem in the systemic operation and structures of the capitalist world-economy; attempting to “imagine” and elaborate an alternative (non-capitalist) system; and strategies to advance or build such alternatives. For educators, the work provides material for use in classroom activities designed to engage students in these critical issues and debates.

At this macro-level, WSA continues to stand as a guiding framework for politically engaged educators and policy makers at all levels. The scope of WSA is and should be broad, working to avoid non-productive debates and concentrate attention on the systemic nature of existing inequalities, tied to the operation of the capitalist world-economy, and so to ambitions to create more egalitarian, just, and democratic alternatives. From a WSA perspective, these are the key characteristics of this approach that are applicable to educational research, whether empirical work explaining and critiquing current realities, or critical education advocacy and interventions (e.g. see [Misiaszek, 2020](#), on ecopedagogy).

World-systems analysis explores how capitalism, as a world-economy, came to incorporate the whole geographical world, and how it established and requires the highly unequal and unjust structure across nation-states, for its normal operation. This rests on capitalism’s historical expropriation of land, resources, and labor; on the unequal transfer of surplus value from the periphery to the core; on the incapacity of the system to maintain this mode of capital accumulation and deliver promised development universally at the same time; and crucially, on the fundamental contradiction between the requirement for endless compound growth, and the rates of resource use and consumption this demands, and the mitigation measures needed to contain global warming. One could argue that Green New Deal advocates, and even the United Nation’s Sustainable Development Goals’ limited efforts to respond to the climate crisis, and its call for differential growth rates across the globe linked to historical carbon emissions and associated levels of “development,” implicitly invoke some of these basic premises of WSA.

We conclude then by reasserting WSA’s contribution to our examination and understanding of the unequal and unjust dimensions of our current world, and particularly the urgent global climate emergency that confronts the world. WSA, and the work of Wallerstein in particular, is unequivocal about the capacity of social science to approximate the truth, and so develop “scholarly analyses that are more correct ... [and are] ... more socially useful” ([Wallerstein, 1997a](#), p. 1254), contributing to the construction of “alternative, credibly better, and historically possible” ([Wallerstein, 1998a](#), p. 2) futures. This understanding can be utilized by politically engaged educators, embedded in educational efforts to prepare students who might in turn contribute to a just and genuinely sustainable transition.

References

- Armove, R.F., 1980. Comparative education and world-systems analysis. *Comp. Educ. Rev.* 24, 48–62.
- Baker, D., 2014. *The Schooled Society: The Educational Transformation of Global Culture*. Stanford University Press, Stanford, California.
- Braudel, F., 1982. *On History*. Translated by Sarah Matthews. The University of Chicago Press, Chicago.
- Braudel, F., 1985. *La dynamique du capitalisme*. Les Éditions Arthaud, Paris.
- Cho, S., 2013. *Critical pedagogy and social change: critical analysis on the language of possibility*. Routledge, New York.

- Collins, R., 2013. The end of middle-class work: no more escapes. In: Wallerstein, I., Collins, R., Mann, M., Derlugian, G., Calhoun, C. (Eds.), *Does Capitalism Have a Future?* Oxford University Press, Oxford New York, pp. 37–69.
- Connell, R., 2020. *Southern Theory: The Global Dynamics of Knowledge in Social Science*. Routledge, New York.
- Dale, R., 2005. Globalisation, knowledge economy and comparative education. *Comp. Educ.* 41 (2), 117–149.
- Demeter, M., 2019. The world-systemic dynamics of knowledge production: the distribution of transnational academic capital in the social sciences. *J. World. Syst. Res.* 25 (1), 111–144.
- Emmanuel, A., 1972. *Unequal Exchange: A Study of the Imperialism of Trade*. Monthly Review Press, New York and London.
- Fanon, F., 1963. *The Wretched of the Earth*. Translated by C. Farrington. Grove Press, New York.
- Frank, A.G., 1978. *World Accumulation, 1492–1789*. Algora Publishing, New York.
- Frank, A.G., 2009. *Latin America: Underdevelopment or Revolution. Essays of the Development of Underdevelopment and the Immediate Enemy*. Monthly Review Press, New York and London.
- Goody, J., 2006. *The Theft of History*. Cambridge University Press, Cambridge.
- Griffiths, T.G., Armove, R.F., 2015. World culture in the capitalist world-system in transition. *Glob. Soc. Educ.* 13 (1), 88–108.
- Griffiths, T.G., Knezevic, L., 2010. Wallerstein's world-systems analysis in comparative education: a case study. *Prospects* 40 (4), 447–463.
- Griffiths, T.G., 2009. 50 Years of socialist education in revolutionary Cuba: a world-systems perspective. *J. Iber. Lat. Am. Res.* 15 (2), 45–64.
- Griffiths, T.G., 2019. *Critical world-systems education for progressive system change*. In: Peters, M.A. (Ed.), *Encyclopedia of Teacher Education*. Springer, Singapore.
- Hickel, J., 2020. *Less is More: How Degrowth Will Save the World*. William Heinemann, London.
- Hobbsbawm, E., 1962. *The Age of Revolution: 1789–1848*. Weidenfeld & Nicolson, London.
- Hobbsbawm, E., 1975. *The Age of Capital: 1848–1875*. Weidenfeld & Nicolson, London.
- Hobbsbawm, E., 1987. *The Age of Empire: 1875–1914*. Weidenfeld & Nicolson, London.
- Klees, S.J., 2008. A quarter century of neoliberal thinking in education: misleading analyses and failed policies. *Glob. Soc. Educ.* 6, 311–348.
- Kondratieff, N., 1984. *The Long Wave Cycle*. Translated by Guy Daniels. Richardson & Snyder, New York.
- Mignolo, W., 2021. *The Politics of Decolonial Investigations*. Duke University Press, Durham and London.
- Mills, C.W., 1959. *Sociological Imagination*. Oxford University Press, London/Oxford/New York.
- Misiasek, G.W., 2020. *Ecopedagogy: Critical Environmental Teaching for Planetary Justice and Global Sustainable Development*. Bloomsbury Academic, New York.
- Prebisch, R., 1963. *Towards a Dynamic Development Policy for Latin America*. United Nations, New York.
- Report on an Intellectual Project: The Fernand Braudel Center, 1976–1991. <https://www.binghamton.edu/fbc/about-fbc/intellectual-report.html>.
- Rich, E.J., Wallerstein, I., 1972. *Africa: Tradition and Change*. Random House, New York.
- Rostow, W.W., 1959. The stages of economic growth. *Econ. Hist. Rev.* 12 (1), 1–16.
- Santos, B.d.S., 2016. *Epistemologies of the South: Justice Against Epistemicide*. Routledge, London and New York.
- Schott, T., 1998. Ties between center and periphery in the scientific world-system: accumulation of rewards, dominance and self-reliance in the center. *J. World Syst. Res.* 4 (2), 112–144.
- Silver, B.J., 2003. *Forces of Labor: Workers' Movements and Globalization since 1870*. Cambridge University Press, Cambridge.
- Sklair, L., 1999. Competing conceptions of globalisation. *J. World Syst. Res.* 5 (2), 1–15.
- Smith, J., Wallerstein, I. (Eds.), 1992. *Creating and Transforming Households: The Constraints of the World-Economy*. Cambridge University, Cambridge.
- Smith, J., 2016. *Imperialism in the Twenty-First Century: Globalization, Super-Exploitation, and Capitalism's Final Crisis*. Monthly Review Press, New York.
- Snow, C.P., 2002. *The Two Cultures. With Introduction by Stefan Collini*. Cambridge University Press, Cambridge.
- Wallerstein, I., et al., 1996. *Open the Social Sciences: Report of the Gulbenkian Commission on the Restructuring of the Social Sciences*. Stanford University Press, Stanford.
- Wallerstein, I., 1961. *Africa: The Politics of Independence*. Vintage, New York.
- Wallerstein, I., 1964. *The Road to Independence: Ghana and the Ivory Coast*. Mouton, Paris.
- Wallerstein, I., 1967. *Africa: The Politics of Unity*. Random House, New York.
- Wallerstein, I., 1974a. The rise and future demise of the world capitalist system: concepts for comparative analysis. *Comp. Stud. Soc. Hist.* 16, 387–415.
- Wallerstein, I., 1974b. *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. Academic Press, New York.
- Wallerstein, I., 1986. *Africa and the Modern World*. Africa World Press, Trenton, New Jersey.
- Wallerstein, I., 1990. Culture as the ideological battleground of the modern world-system. *Theor. Cult. Soc.* 7, 31–55.
- Wallerstein, I., 1995. *After Liberalism*. The New Press, New York.
- Wallerstein, I., 1997a. Social science and the quest for a just society. *Am. J. Sociol.* 102 (5), 1241–1257.
- Wallerstein, I., 1997b. Eurocentrism and its avatars: the dilemmas of social science. *New Left Rev.* 1, 93–107.
- Wallerstein, I., 1998a. *Utopistics: or, Historical Choices of the Twenty-First Century*. The New Press, New York.
- Wallerstein, I., 1998b. The time of space and the space of time: the future of social science. *Polit. Geogr.* 17, 71–82.
- Wallerstein, I., 1999a. *The End of the World as We Know It: Social Science for the Twenty-First Century*. University of Minneapolis Press, Minneapolis and London.
- Wallerstein, I., 1999b. Islam and the West, and the world. *J. Islam. Stud.* X (2), 109–125.
- Wallerstein, I., 2000a. Globalization or the age of transition?: a long-term view of the trajectory of the world-system. *Int. Sociol.* 15 (2), 249–265.
- Wallerstein, I., 2000b. Fanon and the revolutionary class. In: Wallerstein, I. (Ed.), *The Essential Wallerstein*. The New Press, New York, pp. 14–32.
- Wallerstein, I., 2001a. The inventions of TimeSpace realities: towards an understanding of our historical systems. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 135–148.
- Wallerstein, I., 2001b. Marx and underdevelopment. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 207–217.
- Wallerstein, I., 2001c. Call for a debate about the paradigm. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 237–256.
- Wallerstein, I., 2001d. Historical systems as complex systems. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 229–236.
- Wallerstein, I., 2001e. World-systems analysis: the second phase. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 266–272.
- Wallerstein, I., 2001f. Marx and underdevelopment. In: Wallerstein, I. (Ed.), *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia, pp. 151–169.
- Wallerstein, I., 2002a. Left politics in an age of transition. *Mon. Rev.* LIII (8), 17–23. January.
- Wallerstein, I., 2002b. New revolts against the system. *N. Left Rev.* (18), 29–39. November–December.
- Wallerstein, I., 2003a. *The Decline of American Power: The U.S. in a Chaotic World*. The New Press, New York and London.
- Wallerstein, I., 2003b. Racism: our albatross. In: Wallerstein, I. (Ed.), *The Decline of American Power: The U.S. in a Chaotic World*. The New Press, New York and London, pp. 69–99.
- Wallerstein, I., 2004a. *World-Systems Analysis: An Introduction*. Duke University Press, Durham and London.
- Wallerstein, I., 2004b. *The Uncertainties of Knowledge*. Temple University Press, Philadelphia.

- Wallerstein, I., 2005. After developmentalism and globalization, what? *Soc. Forces* 83 (3), 321–336.
- Wallerstein, I., 2006. *European Universalism: The Rhetoric of Power*. The New Press, New York.
- Wallerstein, I., 2009. Reading Fanon in the 21st century. *N. Left Rev.* 57, 117–125.
- Wallerstein, I., 2011. *The Modern World-System IV: Centrist Liberalism Triumphant, 1789–1914*, IV vols. University of California Press, Berkeley.
- Wallerstein, I., 2013. Structural crisis, or why capitalists may no longer find capitalism rewarding. In: Wallerstein, I., Collins, R., Mann, M., Derlugian, G., Calhoun, C. (Eds.), *Does Capitalism Have a Future?* Oxford University Press, Oxford New York, pp. 9–36.
- Wallerstein, I., 2014. Antisystemic movements, yesterday and today. *J. World Syst. Res.* 20, 158–172.
- Warmington, P., 2015. Dystopian social theory and education. *Educ. Theor.* 65 (3), 265–281.
- Warren, S., Starnawski, M., Tsatsaroni, A., Vogopoulou, A., Zgaga, P., 2020. How does research performativity and selectivity impact on the non-core regions of Europe? The case for a new research agenda. *High. Educ.* 81, 607–622.

Spatial theories, methods and education policy

Kalervo N. Gulson^a, Dan Cohen^b, Steven Lewis^c, Emma Rowe^d, Ee-Seul Yoon^e, and Chris Lubienski^f, ^aSydney School of Education and Social Work, Faculty of Arts and Social Sciences, The University of Sydney, Camperdown, NSW, Australia; ^bDepartment of Geography and Planning, Queen's University, Kingston, ON, Canada; ^cREDI Centre, Deakin University, Burwood, VIC, Australia; ^dFaculty of Arts and Education, Deakin University, Burwood, VIC, Australia; ^eFaculty of Education, University of Manitoba, Winnipeg, MB, Canada; and ^fSchool of Education, Indiana University – Bloomington, Bloomington, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Spatial theories, methods and education policy	29
Marxist political economy and spatial dimensions of education	30
A (post)spatial turn	31
Power and topological space	31
Geospatial research development in education: a focus on method	32
Conclusion: asking spatial questions of education policy	33
References	34

Spatial theories, methods and education policy

While education policy can be understood as a temporal intervention into the world (Thompson and Cook, 2014), the increased emphasis in the last decade on the globalized aspects of policy has raised the importance of the geographical aspects of policy. As Rizvi and Lingard (2010) posit, “[t]he disposition for critical policy analysis in an era of globalization requires that we recognize the relationality and interconnectivity of policy developments” (p. 69). Part of this work has been to emphasize the ways that globalization requires a focus on power, boundaries, the global flows of policy ideas and people, and the ways policy is created, moved and modified in “vernacular” ways. Broadly, there has been an increased use of spatial language in understanding education policy.

In this article, we focus on the *where* of policy to understand how and why policy is made and enacted. We locate this article in what can be called the “critical geographies of education policy,” a subfield of the broader field of “critical geographies of education” that is concerned with the spatial dynamics of issues of power, inequality and difference (McCreary et al., 2013; Pini et al., 2017; Nguyen et al., 2017). In this article, the work we are outlining is part of what Thiem (2009) highlights as an “inward-looking” geography of education, where education policy is “the ultimate ends of investigation, with spatial variations in the provision, consumption, and outcomes of schooling considered to be phenomena in need of exploration” (p. 156). This article outlines what could be seen as the utility of spatial theories and methods in investigating education policy and politics, where we are primarily looking at the analytical outcomes of using spatial theories. Thiem (2009), however, also outlines the potential of “outward-looking” geographies of education, where the spatial dynamics of education are themselves a generative site of theory building.

As will also be clear as the article proceeds, spatial theory has traveled through interdisciplinary fields and in this travel, and this theoretical expansiveness, has at times lacked a clear road map. While geography is the logical home of spatial concepts, these concepts travel particularly through and between social theory and application in empirical work (Bourdieu, 1999), such as in education (Yoon, 2020). Gulson and Symes (2007b) note, “In many ways the spatial trajectories in education have mirrored those of other disciplines in borrowing, appropriating, and employing the spatial theories of geography” (p. 5). The increased spatial language, or tropes of space, and the use of imagery or metaphors of space in education policy studies can be described by what Said (1993) calls the conjuring of a “geographical imagination”. It is unviable, Said’s work suggests, to explore notions of power or knowledge without also considering territory, geography and space. This “geographical imagination” is conceptualized as the “spatial turn”, a growing attention to notions of space in the social sciences (Gulson and Symes, 2007b). Tropes of space, such as belonging or unbelonging, spaces of inclusion or spaces of engagement, point to the spatial turn. After all, “the language of exclusion is, by and large, spatial” (Gulson and Symes, 2007a). However, at times this spatial language is left as self-explanatory. There are numerous examples of spatial metaphors utilized in self-explanatory ways; for example, Lefebvre’s (1968) famous reference, “The Right to the City”, is commonly used in educational work, but often without reference to Lefebvre. A search for Lefebvre will produce only a handful of results within an education search engine, whereas “the right to the city” will produce thousands.

Therefore, in being cognizant of the movement of ideas, we have tried in this article to be precise about definitions and usage of spatial ideas. We accept Gregory’s contention that it is important that when using the language of space that we should “interrogate its common-sense understandings” (Gregory, 1994). A brief foray can illustrate the importance of interrogating common-sense understandings, namely the example of power and space. Spatial theories are characterized by their focus on power and how this is layered and textured. This is reflected in one of Massey’s (1993) signature concepts, “power-geometries”, and her attention to the spatial division of labor. Access to resources and means of production is embedded within space, and individuals retain unequal access depending on their class, race, gender and able-bodiedness. Power is embedded within space, or as Soja (1996) writes, “power is ontologically embedded”; whereas for Lefebvre (1976), “power is everywhere; it

is omnipresent, assigned to Being. It is everywhere *in space*" (p. 86, emphasis in original). An interpretation of spatial theory, then, will depend on one's epistemological and ontological positioning; that is, one's reading of knowledge and constructs of power. In this article, we outline, rather than synthesize, a range of work that spans across critical and quantitative studies.

This article outlines key ways in which space has been used in education policy work, providing overviews of concepts and examples of application in education. These are necessarily broad strokes, but we hope to provide some insights into the richness of spatial ideas and the range of ways that they can help us understand education policy. In the first part of the article much of what we cover can be broadly located within the premise that geography shapes social relations and practices, and this geographical connection between things and people can be broadly understood as "spatiality" (Keith and Pile, 1993). Spatial theory is concerned with the epistemological and ontological; space is fundamentally tied to place-bound expressions of power and representations of power. Yet, space is not bounded or closed; it is interminable and open. We move from Marxist political economy to a focus on postmodern theories and notions of relationality, including the extension of this into ideas of place, and topological approaches. The second part of the article concerns how method and geospatial research are re-focusing on place and education policy, while also bringing back quantitative methods into critical geographical scholarship that emphasize inequity and opportunity gaps. We conclude with some brief thoughts on the types of questions and issues that an explicit focus on space allows in education policy studies.

Marxist political economy and spatial dimensions of education

One strain of spatial theory that has been influential in education policy, particularly in the North American context, is the use of Marxian political economy associated with geographer David Harvey. Harvey's work, beginning with *Social Justice and the City* (1973) and *Limits to Capital* (1982) and spanning decades since, emphasizes the importance of the circulation of capital and commodities in shaping spatial relations. Like Lefebvre and Massey, power for Harvey imbues spatial relations, but this power is rooted in the social relations of the capitalist system. This is most clearly illustrated in Harvey's concept of the "spatial fix" and its manifestations in urban space, where the over-accumulation of capital requires the production of spaces that can absorb surpluses of capital and labor. Here, the booms and busts of urban real estate are understood as the result of how the search for high returns on investment leads capital to alternatively flow in and out of neighborhoods, cities, and countries. Related concepts such as uneven development and the "rent gap theory of gentrification" further highlight how the circulation of capital unevenly shapes urban spaces in ways that maximize capital accumulation (Smith, 1982).

This work, and the work of geographic political economists following in Harvey's footsteps, has been taken up by scholars of urban education in order to understand the relationship between schools, policy, and the production of urban space (Pedroni, 2011; Lipman, 2011; Buras, 2015). As put forward by these authors, we cannot understand changes to urban schooling policy without investigating their connection to the capitalist spatial relations that continually make and remake cities and their neighborhoods, notably in the production of disequilibrium between inner-cities and suburbs in metropolizes, and in the subsequent gentrification of inner-cities. Most commonly, this has taken the form of in-depth examinations of how school closures, the marketization of school systems, and other facets of urban education reform have facilitated processes of gentrification (Butler et al., 2013; Hankins and Martin, 2006; Nethercote, 2017; Lipman, 2011).

As Anyon (2005) describes, in this view, "urban schools are at the center of the maelstrom of constant crises that beset low-income neighborhoods" (p. 179), crises that are the consequences of suburbanization and resource redistribution. For Lipman in their investigations of Chicago, education reforms can therefore best be understood as part of projects that "remake public education to remake the city" (Lipman, 2011). Through detailed work and longstanding activist engagement, Lipman describes how marketized reform policies allow for class and race segregated schools to act as the "third leg" (2011, p. 93) of the gentrification process. They do so by allowing middle-class residents to avoid schools deemed as "failing" and thereby facilitate their movement into the city, thus unlocking windfalls for real estate developers. Linking similar processes in Detroit to Harvey's ideas, Pedroni (2011) argues that educational policy "supports and guides" the spatial fix of capital into urban real estate, and furthers this argument by arguing that this spatial fix is also a "racial fix" that intensifies urban inequalities. Cohen (2020), in cognate work on Detroit, focuses on the intersection of race and urban change through a framework of market logics.

Working in the Australian context, Nethercote (2017) documents how this connection between the circulation of capital and schooling may take different forms according to particular spatial conjunctures. Examining school policies in Melbourne, Nethercote highlights that the underfunding of public schools in that city has resulted in middle-class displacement from the urban core in ways that contrast with the American experience, but which nonetheless aid "the production of exclusive, privatized urban spaces fueled by and for global capital flows" (p. 462). This highlights how approaches to understanding education policy through studying the spatial dimensions of Marxist political economy require an attention to how capital, schools, and the production of (urban) spaces coalesce in different sites. Also in the Australian context, Rowe (2015) utilizes Harvey's theories of relational and absolute space in order to explore geographically-bound school choices of the gentrified middle-class, arguing that class-identity is constructed in relation to geographical (or residential) positioning. While space is fluid and flexible—always manipulated and in flux—space can be operationalized as fixed and constrained within the context of a highly competitive urban schooling market. Drawing on Harvey (2006), relative space speaks to the *relationships* between objects, whereas absolute space is conceptualized as fixed and immovable, a grid that is calculated by measurements, maps and physical geometry. Harvey (1973) argues absolute space is limited and too definitive, as it becomes a "thing in itself" (p. 13). A schooling space and community is relative space, in

that it is always moving and reshaping itself; but one could simultaneously argue that a school is located in absolute space and the class status that it becomes attached to becomes a “thing in itself”; while it is clearly subject to change, it is simultaneously essentialized, known and conceptualized as fixed and stable. Thereby, space is rooted in performativity of its relations.

A (post)spatial turn

Another strain of spatial thought has been connected to postmodernist spatial theories. The spatial turn is marked out by a shift away from positivism and essentialism, and instead a move toward postmodernism, postcolonial and feminist spatial theorists. It is interesting to see how ideas that seem located in one theoretical aspect can be taken up across areas. For example, the work of Henri Lefebvre is located in French Marxism. A spatial reading of capitalism in his famous book *La production l'espace* outlined in a wide ranging and at times frustratingly meandering manner,¹ that space is socially produced, engineered and constructed; and importantly, that social relations are made, or produced, relative to space (Lefebvre, 1991). The eclectic nature of Lefebvre's work has allowed it to be taken up by postmodern theorists such as Soja (1996), who argues for “thirdspace”, and the “trialectics of spatiality”, drawing on Lefebvre's three-way “perceived, lived and conceived spatiality”. Soja reworks Lefebvre's ideas into “spatial justice”, which was meant to create a series of emerging ideas held together by the contention that “everything that is social (justice included) is simultaneously and inherently spatial” (Soja, 2010).

Education research drawing on reworkings of Lefebvre has drawn out his analysis of the connections between social production, practices and capitalism, including Rowe's work on school choice, middle class practices and urban change (Rowe, 2017, 2020). Soja's work both on “thirdspace” and “spatial justice” has been the focus of work, interestingly, on education policy work on rural spaces and policy change (Green and Letts, 2007; Beach et al., 2018).

Another key thinker whose work developed from political geography through to social and cultural geography is Doreen Massey. Massey's work from a focus on the spatial dimensions of labor (Massey, 1979), on gender and space (Massey, 1994b), and connecting globalization, cities and space (Massey, 2007) emphasized that space was a key site of politics (Massey, 1992, 2009). Her work, particularly in reference to cities, broadened and challenged spatial understanding of globalization as a Western, colonial economized idea (Massey, 2005). The conceptual apparatus to do this work was relational space with three key components. There is no room to fully expand on these ideas but for now it is enough to identify them: (1) *Interrelations*, in that space is constituted by relations between people, and between people and the non-human; (2) *Multiplicity*, that rather than being one thing, space can be simultaneously multiple; and (3) *Openness*, that rather than being a non-political fixed arena, space is political (Massey, 2005). This doesn't mean space is never fixed, for as Thrift posits, “Spaces can be stabilized in such a way that they act like political utterances, guiding subjects to particular conclusions” (Thrift, 2003).

Space and place have tended to be conflated when taken up in education policy studies. Butler and Sinclair (2020) suggest that “while a growing body of theoretical and conceptual work has explored the role of place in education practice ..., place has been somewhat underconceptualized in education research” (p. 65). When place has been conceptualized in education policy research, not as a conflation but a discrete spatial category, it has tended to draw on work that has attempted to show how places (schools, neighborhoods) have interconnections that reconfigure the idea that place is local. This idea draws on relational notions of place and space and identity such as those advocated by Massey, who asserts the importance of the relationship of place and space, arguing that space is not more important than place, but rather that place “is best thought of as a particular part of, a particular moment in, the global network of those social relations and understandings” (Massey, 1994a). The key here that place is not static, closed and only local or proximal. This idea has been extended to focus on ideas of spatial encounters that comprise place as an event, with Amin and Thrift positing that places “are best thought of not so much as enduring sites but as moments of encounter, not so much as ‘presents’, fixed in space and time, but as variable events; twists and fluxes of interrelation” (Amin and Thrift, 2002).

Work on relational notions of space and place has spanned areas such as globalization, policy and the construction of the teacher (Maguire, 2002), and notions of personalization and de-schooling (Pykett, 2009). Other work has used Massey's ideas of the links between politics, race and openness and relational space to explore how space works to shape political possibilities in education (Gulson and Parkes, 2009). Additional work has used relationality as a way to connect the materiality of urban change, such as gentrification with education policy change, such as school choice (Smith and Lupton, 2008; Gulson, 2011). While Massey's work on globalization has been often linked to cities, educational scholars have used it to explore the globalized spaces of rural education, policy and marginalization, including in Australia and Chile (Oyarzún, 2020; Reid, 2017).

Power and topological space

The focus on relationality is a key part of topological conceptions of power, or “power-topologies” (Allen and Cochrane, 2010), has proved useful to theorize how power is exercised in contemporary spaces and relations associated with globalization (Amin, 2002). While Euclidean space provides the a priori backdrop against which power relations are situated and exerted, power-topologies

¹This wide-ranging nature of his writing may well have been connected to his writing habits, or indeed his series of secretaries and relationships (Merrifield, 2006, p. xxii).

instead constitute the very spaces in which their effects and influence are exerted, in which power relationships are “not so much positioned in space or extended across it, as they compose the spaces of which they are a part” (Allen, 2011). Power-topologies also express how policy actors are able to overcome physical distance to “make their presence felt in more or less powerful ways” (Allen, 2011). These ideas have been taken up in education policy in a focus on two main concentrations of education research. These are the movement of policy, as well as processes of datafication and infrastructuring, respectively.

Education policy sociology research concerned with the movement of policy has primarily drawn on topology to emphasize the dynamic movement (*flows*) of policy and, at the same time, the contextual embeddedness of its uptake, contestation and enactment (*frictions*) by people and places. Adopting a topological approach to policy sociology requires: (1) accounting for the new relational spaces of the policy cycle; (2) observing how these policies move within and across, and simultaneously reconfigure, these spaces; and (3) noting how these interactions and movements through space, as well as local processes of implementation, reshape the policies in question. Often, this framework has been translated into research via a rationale of “following the policy” (Peck and Theodore, 2015; McCann and Ward, 2012), where one follows (virtually or in-person) the people, places and processes through which policy is made and moves. Arguably, this is now central for policy sociologists who seek to address the networked nature of education policymaking, as well as for understanding how the enactment of “global” policy flows is still very much shaped by decidedly “local” conditions and contexts. More recently, there have been an increasing number of contributions that study the movement of education policy via a topological lens (Gulson et al., 2017; Lewis and Hardy, 2017; Gulson and Sellar, 2019).

At the same time, topology has also provided generative means for exploring how digital data are now increasingly central to the constitution of schooling systems as knowable and governable spaces. The concept of data infrastructures – or the process of *data infrastructuring* – refers here to socio-technical assemblages (of databases, software platforms, people, buildings and practices) that laboriously translate things into data and fabricate supposedly objective governing knowledge (Piattoeva and Saari, 2020). Topological modes of culture have emerged through processes of continuity and ordering, and significant advances in computational capacities to collect, calculate and compare such data have enabled new kinds of “ordered-ness” to emerge. For example, the generation of large data sets relating to the comparative performance of education systems (perhaps best typified by the OECD’s PISA) enables modes of reasoning focused on the relationships between policy settings, despite such policies being deployed in geographically distant and contextually distinct schools and schooling systems. There has similarly been a rising body of critical attention devoted to the ongoing datafication of schooling in recent year (Hartong, 2020; Lewis and Holloway, 2019; Takayama and Lingard, 2019).

Geospatial research development in education: a focus on method

This section discusses an emerging body of educational research that focuses on geospatial analytical approaches and their further development, as part of mixed-methods approaches that connect space (e.g., location, proximity, and distance) to place (e.g., sense of belonging, neighborhood boundaries, and lived experience) (Waitoller and Lubienski, 2019; Yoon et al., 2018a). Geospatial research is based primarily on the use, application, and extension of “geographic information systems” (or “science,” but “GIS” in either case), a fusion of georeferenced data (with locational details) and software that enables the analysis of such data. Harvey (1972), for instance, uses GIS-generated maps to illustrate his analysis of urban inequality. In education research, any kind of demographic or school data (e.g., achievement, absenteeism) can be transformed into geographic information when it is linked to longitudinal and latitudinal coordinates on the earth’s surface. Researchers thus emphasize the role of GIS, referring to geospatial research as GIS-based or GIS research in education, highlighting the indispensability of GIS to this type of research (Cobb, 2020).

GIS-based or geospatial research has been used in social sciences like sociology and economics, as well as in commercial endeavors such as market analytics, for some time. In education, GIS has been used by school boards since the 1980s for the purpose of school planning (Rob Brown, personal communication, November 20, 2020), while academic researchers started using this type of research in the late 1990s, especially in the United States (Cobb and Glass, 1999, 2009). This early group of educational researchers who adopted GIS leveraged it as a technology and an analytical tool to create and manipulate geographically referenced information, such as the impact of charter schools or school choice on ethnic and racial segregation and stratification (Cobb and Glass, 1999). Also, in the United Kingdom and Europe, researchers started using GIS to explore the impact of parental choice of schools on student flows across catchment boundaries or competition between schools in local areas (Parsons et al., 2000; Taylor, 2002), as well as ethnic segregation as a result of school choice (Karsten et al., 2003). This body of work builds on the notions of spatial segregation that are associated with social and ethnic exclusion and hierarchy.

During this early period of geospatial research, educational researchers applied GIS to their research to varying degrees. Some applied it at the core of their analysis (Taylor, 2007), while others used only the products of GIS, such as maps, merely to complement or illustrate their studies, without stating explicitly why or how they used it (Lipman, 2002). Also, it is worth noting that the rise of GIS research among educational researchers came in the wake of a similar trend among human geographers, particularly as computer software became more advanced. This trend, which arose in the 1990s, followed extensive debate between the computer scientists who develop GIS software to measure spatial relationships, and the human geographers who co-construct maps that reflect a variety of voices (Nyerges et al., 2011). It was after this debate that human geographers began to accept the use of GIS as part of their place-based research that focuses on participants’ meanings of communities, neighborhoods, and mobility. The rise of GIS research in education is thus linked to these developments in human geography and computer science.

So, why has GIS become so popular among educational researchers? As is notable in the publications by an early group of educational researchers, geospatial research has been applied to answer some of the long-standing questions that are pertinent to education and can be (re)examined using a spatial lens, with a particular attention paid to geographical clustering, proximity, distance, or accessibility. For instance, questions related to school outcomes, school catchment or attendance areas, and educational quality, including teachers, curriculum, and resources, have a spatial dimension because schools are located in particular neighborhoods. Since the appearance of schoolhouses in the 19th century, educational researchers have studied school outcomes and how they vary from school to school. By introducing GIS, researchers have been able to use maps to more effectively illustrate which schools have higher or lower outcomes, especially against the backdrop of neighborhood characteristics and development. More importantly, geospatial research has facilitated the theorization of how school outcomes vary according to the resources that are (un)available in the surrounding neighborhoods, including the level of affluence, community resources, libraries, or economic resources, which are now commonly referred to as the geography of opportunity (Gulosino and Yoon, 2020; Tate, 2012). It also allows for an examination of gerrymandering of school attendance zones (Richards, 2014), and curriculum variations (Gulosino and Miron, 2020). Overall, geospatial approaches have been able to help us see the geographical dimension of these perennial issues, which are universal across different time periods and countries (Mann and Saultz, 2019; Yoon et al., 2018b). Themes explored in this area of educational research resonate with the political economy of urban education as well as theories of educational reproduction, that are rooted in Marxism and critical theories.

GIS has become particularly invaluable for the study of school choice reforms. As Lubienski and Dougherty (2009) note, the issue of school choice reforms was at the forefront of geospatial research because “choice is inherently a spatial issue, with geography, location, and distance all primary concerns for parents choosing schools” (p. 487). Geospatial tools available in GIS software are particularly well suited for understanding the importance of school proximity and accessibility (Burgess et al., 2011; Marshall et al., 2010). The tools allow for the analysis of school segregation and stratification patterns by race and class in the growth of different types of school choice options, including charter schools (Butler et al., 2007; Saporito and Sohoni, 2006). Geospatial research is effective in illustrating the racialized manipulation of school admission policies and the gerrymandering of school attendance policies by schools that wish to maintain their competitive position by keeping only “desirable” students (Lubienski et al., 2009, 2013). Also, when linked to critical social theories, GIS can provide a geographical dimension to school choice and its reproductive function (Yoon et al., 2018b, 2020).

While initially used to analyze quantitative education data, recent decades have seen geospatial research increasingly come to complement qualitative research in education. Historically, ethnographies and case studies have focused on the places where schooling happens (Willis, 1981; Gewirtz et al., 1995; Anyon, 2005). Interview-based case studies and ethnographic studies dominated, for instance, place-based research on school segregation by race and class, and these studies were used to analyze historically-formed racialized neighborhood schools and the politics of school closures (Andre-Bechely, 2007; Lipman, 2011; Yoon, 2017). Further, critical race theory and disability theory have been applied to shed light on how different racial and ability groups perceive and co-construct places of educational policy and schools (Waitoller and Super, 2017; Armstrong, 2003). Yet, more recently, building on the epistemologies of critical, hybrid, and heterogeneous constructivism, a new group of researchers – in both human geography and the geography of education – has taken a much more mixed-methods approach that incorporates GIS with qualitative data, such as field notes, interviews, or sketch maps (Kwan, 2008; Yoon and Lubienski, 2018). These researchers use a mixed-methods design to layer analysis of macro-level patterns with micro-level practices and politics. These methods are robust and have been developed based on a wide range of critical, feminist, and postmodern social theories that scaffold spatial methods.

Connecting space to place, mixed-methods researchers use maps to show what is going on across a city, region, and even countries, and to tell the stories of maps using micro-level data. These maps, moreover, leverage individual meaning-making, social capital building, and politics, all of which shape macro-level patterns of education. Their approaches are particularly powerful in telling local stories that other maps can miss (Lubienski and Lee, 2017), including the views of historically racialized groups (Yoon and Lubienski, 2017); the desired versus actual choices of first-generation, low-income college students (Jabbar et al., 2017); and historically-formed social networks along class backgrounds that underpin higher education choices (Gamsu and Donnelly, 2020). Also, participatory GIS mapping has shown that students from marginalized communities can participate in map-making in ways that challenge poverty and deprivation in their neighborhoods, thereby calling for greater spatial justice (Ghose and Welcenbach, 2018). These emerging studies indicate that geospatial research in education has enormous untapped potential that can help generate reflexive scholarship, in which researchers challenge the notion of GIS as a neutral scientific tool while also serving as catalyst to bring about social and spatial justice (Crampton, 2010; Yoon et al., 2018a).

Conclusion: asking spatial questions of education policy

Educational policy is very much considered a temporal intervention into the education systems and societies that produce and maintain them. Indeed, when a new party is elected to form a government, they introduce a set of new policies to mark the “era” of their political regime. Nevertheless, in this article, we explored some major theories and methods as rooted in critical geographies (of education) scholarship, and how they can help us not to lose sight of *where* power, inequality and difference are (re) generated in times of change, especially in the current education systems that are increasingly privatized and digitized.

Over a decade ago, [Gulson and Symes \(2007b\)](#) argued that, “Drawing on theories of space contributes in significant and important ways to subtle and more sophisticated understandings of the competing rationalities underlying educational policy change, social inequity, and cultural practices” (p. 2). As the above aspects outline, this has only become more evident in the intervening years. Since [Rizvi and Lingard \(2010\)](#) argued that globalization requires an understanding of the interconnectivity of policy developments a decade ago, we have seen further connectivity through processes such as the datafication of education, but also reminders of the unevenness of geographic flows, as circulating policies, pedagogies, and even viruses articulate with the politics of place to shape educational practice.

While it is obvious that there is no theoretical or methodological fidelity to what we have covered, and we have indeed glossed over important epistemological and ontological debates within human geography, we do think that what has been provided can be an impetus to ask new questions about enduring and novel educational problems. That is, what educational policy questions could be asked, or asked differently, with spatial approaches? A critical researcher needs to consider such questions. Traditionally, the study of education policy has been rooted in science and positivist approaches to understanding policy change and reform. As this has evolved to a more critical approach and sociology of education, there is a significant need to consider the role of space in education reform. Studying education policy traditionally takes place in “flatness”, an inert and even space by disinterested researchers. This is the unoccupied and distant lens, one that is inattentive to the politics and power of space. To take up a spatial approach is to occupy space; that is, noticing how the researcher and participants occupy and *take up* space in ways that are performative and rooted in “power-geometry” ([Massey, 1993](#)).

In this regard, the approaches reviewed in this article highlight that spatial theories are not unidirectional but relational, and ask us to consider educational policy in ways that push against tendencies to take the “local”, “national”, or “global” for granted. Whether understanding the role of schooling in producing urban spaces for capital, the relational encounters that make up educational practice, or “following the policy” through “flows” and “frictions”, the geographies of education serve to unbound the study of education as a discrete space and, instead trace both inwards and outwards to examine how educational practice and policies shape (and themselves reshape) the networks with which they are connected ([Lewis, 2020](#)).

As such, we hope that future research continues to take up questions related to the spatial dimensions of policy, whether place-specific or mobile. We hope this future research examines, moreover, how policies unfold in informal or formal learning spaces, open or closed programs of choice, public or private schools, open or gated communities, and locally elected or independent school boards, in the current reforms sweeping across the globe. Spatial theories and methods can offer a new generative lens for critical studies of educational policy, and help illuminate how educational spaces are inherently tied to the (re)production of power and privilege in the past, the present and future.

References

- Allen, J., Cochrane, A., 2010. Assemblages of state power: topological shifts in the organization of government and politics. *Antipode* 42, 1071–1089.
- Allen, J., 2011. Topological twists: power's shifting geographies. *Dialog. Hum. Geogr.* 1, 283–298.
- Amin, A., Thrift, N., 2002. *Cities: Reimagining the Urban*. Polity Press, Cambridge.
- Amin, A., 2002. Spatialities of globalisation. *Environ. Plann.* 34, 385–399.
- Andre-Bechely, L., 2007. Finding space and managing distance: public school choice in an urban California district. *Urban Stud.* 44, 1355–1376.
- Anyon, J., 2005. *Radical Possibilities: Public Policy, Urban Education, and a New Social Movement*. Routledge, New York.
- Armstrong, F., 2003. *Spaced Out: Policy, Difference and the Challenge for Inclusive Education*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Beach, D., From, T., Johansson, M., Öhrn, E., 2018. Educational and spatial justice in rural and urban areas in three Nordic countries: a meta-ethnographic analysis. *Educ. Inq.* 9, 4–21.
- Bourdieu, P., 1999. *The Weight of the World*. Stanford University Press.
- Buras, K.L., 2015. *Charter Schools, Race, and Urban Space: Where the Market Meets Grassroots Resistance*. Routledge, New York.
- Burgess, S., Greaves, E., Vignoles, A., Wilson, D., 2011. Parental choice of primary school in England: what types of school do different types of family really have available to them? *Pol. Stud.* 32, 531–547.
- Butler, A., Sinclair, K.A., 2020. Place matters: a critical review of place inquiry and spatial methods in education research. *Rev. Res. Educ.* 44, 64–96.
- Butler, T., Hamnett, C., Ramsden, M., Webber, R., 2007. The best, the worst and the average: secondary school choice and education performance in East London. *J. Educ. Pol.* 22, 7–29.
- Butler, T., Hamnett, C., Ramsden, M.J., 2013. Gentrification, education and exclusionary displacement in East London. *Int. J. Urban Reg. Res.* 37, 556–575.
- Cobb, C.D., Glass, G.V., 1999. Ethnic segregation in Arizona charter schools. *Educ. Pol. Anal. Arch.* 7, 1.
- Cobb, C.D., Glass, G.V., 2009. School choice in a post-desegregation world. *Peabody J. Educ.* 84, 262–278.
- Cobb, C.D., 2020. Geospatial analysis: a new window into educational equity, access, and opportunity. *Rev. Res. Educ.* 44, 97–129.
- Cohen, D., 2020. “A marketplace of schools”: race, power, and education reform in the Detroit region. *Urban Geogr.* 1–25.
- Crampton, J.W., 2010. *Mapping: A Critical Introduction to Cartography and GIS*. Wiley-Blackwell, Oxford.
- Gamsu, S., Donnelly, M., 2020. Social network analysis methods and the geography of education: regional divides and elite circuits in the school to university transition in the UK. *Tijdschr. Econ. Soc. Geogr.* 112 (4), 370–386.
- Gewirtz, S., Ball, S.J., Bowe, R., 1995. *Markets, Choice and Equity in Education*. Open University Press, Buckingham.
- Ghose, R., Welcenbach, T., 2018. “Power to the people”: contesting urban poverty and power inequities through open GIS. *Can. Geogr.* 62, 67–80. *Le Géographe canadien*.
- Green, B., Letts, W., 2007. Space, equity, and rural education: a ‘trialectical’ account. In: Gulson, K.N., Symes, C. (Eds.), *Spatial Theories of Education: Policy and Geography Matters*. Routledge, New York.
- Gregory, D., 1994. *Geographical Imaginations*. Blackwell, Oxford.
- Gulosino, C., Miron, G., 2020. Examining curriculum differentiation of education management organizations (EMOs) in 41 States. *Peabody J. Educ.* 95, 272–299.
- Gulosino, C., Yoon, E.-S., 2020. School choice policies shaping neighborhoods, school locations, and destinies. *Peabody J. Educ.* 95, 203–210.
- Gulson, K.N., Parkes, R.J., 2009. In the shadows of the mission: education policy, urban space and the ‘colonial present’ in Sydney. *Race Ethn. Educ.* 12, 267–280.

- Gulson, K.N., Sellar, S., 2019. Emerging data infrastructures and the new topologies of education policy. *Environ. Plann. Soc. Space* 37, 350–366.
- Gulson, K.N., Symes, C., 2007a. Knowing one's place: space, theory, education. *Crit. Stud. Educ.* 48, 97–110.
- Gulson, K.N., Symes, C., 2007b. Knowing one's place: educational theory, policy, and the spatial turn. In: Gulson, K.N., Symes, C. (Eds.), *Spatial Theories of Education: Policy and Geography Matters*. Routledge, New York.
- Gulson, K.N., Lewis, S., Lingard, B., Lubienski, C., Takayama, K., Webb, P.T., 2017. Policy mobilities and methodology: a proposition for inventive methods in education policy studies. *Crit. Stud. Educ.* 58, 224–241.
- Gulson, K.N., 2011. *Education Policy, Space and the City: Markets and the (In)visibility of Race*. Routledge, New York.
- Hankins, K.B., Martin, D.G., 2006. Charter schools and urban regimes in neoliberal context: making workers and new spaces in metropolitan Atlanta. *Int. J. Urban Reg. Res.* 30, 528–547.
- Hartong, S., 2020. The power of relation-making: insights into the production and operation of digital school performance platforms in the US. *Crit. Stud. Educ.* 1–16.
- Harvey, D., 1972. *Society, the City and the Space-Economy of Urbanism*. Association of American Geographers.
- Harvey, D., 1973. *Social Justice and the City*. University of Georgia Press, Athens.
- Harvey, D., 1982. *The Limits to Capital*. University of Chicago Press, Chicago.
- Harvey, D., 2006. Space as a keyword. In: Castree, N., Gregory, D. (Eds.), *David Harvey: A Critical Reader*. Blackwell, Oxford.
- Jabbar, H., Sánchez, J., Epstein, E., 2017. Getting from here to there: the role of geography in community college students' transfer decisions. *Urban Rev.* 49, 746–776.
- Karsten, S., Ledoux, G., Roelvelde, J., Felix, C., Elshof, D., 2003. School choice and ethnic segregation. *Educ. Pol.* 17, 452–477.
- Keith, M., Pile, S., 1993. Introduction. Part 1: the politics of place. In: Keith, M., Pile, S. (Eds.), *Place and the Politics of Identity*. Routledge, London.
- Kwan, M., 2008. Critical GIS. In: Kemp, K.K. (Ed.), *Encyclopedia of Geographic Information Science*. Sage Publications Inc, Thousand Oaks.
- Lefebvre, H., 1968. *Le droit à la ville*. Collection Société et Urbanisme, Anthropos, Paris.
- Lefebvre, H., 1976. *The Survival of Capitalism: Reproduction of the Relations of Production*. Allison & Busby.
- Lefebvre, H., 1991. *The Production of Space*. Blackwell, Oxford.
- Lewis, S., Hardy, I., 2017. Tracking the topological: the effects of standardised data upon teachers' practice. *Br. J. Educ. Stud.* 65, 219–238.
- Lewis, S., Holloway, J., 2019. Datafying the teaching 'profession': remaking the professional teacher in the image of data. *Camb. J. Educ.* 49, 35–51.
- Lewis, S., 2020. The turn towards policy mobilities and the theoretical-methodological implications for policy sociology. *Crit. Stud. Educ.* 1–16.
- Lipman, P., 2002. Making the global city, making inequality: the political economy and cultural politics of Chicago school policy. *Am. Educ. Res. J.* 39, 379–419.
- Lipman, P., 2011. *The New Political Economy of Urban Education: Neoliberalism, Race and the Right to the City*. Routledge, New York.
- Lubienski, C., Dougherty, J., 2009. Mapping educational opportunity: spatial analysis and school choices. *Am. J. Educ.* 115, 485–491.
- Lubienski, C., Lee, J., 2017. Geo-spatial analyses in education research: the critical challenge and methodological possibilities. *Geogr. Res.* 55, 89–99.
- Lubienski, C., Gulosino, C., Weitzel, P.C., 2009. School choice and competitive incentives: mapping the distribution of educational opportunities across local education markets. *Am. J. Educ.* 115, 601–647.
- Lubienski, C., Lee, J., Gordon, L., 2013. Self-managing schools and access for disadvantaged students: organizational behaviour and school admissions. *N. Z. J. Educ. Stud.* 48, 82.
- Maguire, M., 2002. Globalisation, education policy and the teacher. *Int. Stud. Sociol. Educ.* 12, 261–276.
- Mann, B., Saultz, A., 2019. The role of place, geography, and geographic information systems in educational research. *AERA Open* 5, 2332858419869340.
- Marshall, J.D., Wilson, R.D., Meyer, K.L., Rajangam, S.K., McDonald, N.C., Wilson, E.J., 2010. Vehicle emissions during children's school commuting: impacts of education policy. *Environ. Sci. Technol.* 44, 1537–1543.
- Massey, D., 1979. *The Geography of Industrial Reorganisation: The Spatial Effects of the Restructuring of the Electrical Engineering Sector under the Industrial Reorganisation Corporation*. Pergamon Press, Oxford and New York.
- Massey, D., 1992. Politics and space/time. *N. Left Rev.* 196, 65–84.
- Massey, D., 1993. Power-geometry and a progressive sense of place. In: Bird, J., Curtis, B., Putnam, T., Robertson, G., Tickner, L. (Eds.), *Mapping the Futures: Local Cultures, Global Change*. Routledge, London.
- Massey, D., 1994a. Double articulation: a place in the world. In: Bammer, A. (Ed.), *Displacements: Cultural Identities in Question*. Indiana University Press, Bloomington.
- Massey, D., 1994b. *Space, Place and Gender*. University of Minnesota Press, Minneapolis.
- Massey, D., 2005. *For Space*. Sage Publications, London.
- Massey, D., 2007. *World City*. Cambridge Polity Press.
- Massey, D., 2009. Concepts of space and power in theory and in political practice. *Doc. Anal. Geogr.* 55, 15–26.
- Mccann, E., Ward, K., 2012. Assembling urbanism: following policies and 'studying through' the sites and situations of policy making. *Environ. Plann.* 44, 42–51.
- Mccreary, T., Basu, R., Godlewski, A., 2013. Critical geographies of education: introduction to the special issue. *Can. Geogr.* 57, 255–259. *Le Géographe canadien*.
- Merrifield, A., 2006. *Henri Lefebvre: a critical introduction*. Routledge, London.
- Nethercote, M., 2017. When social infrastructure deficits create displacement pressures: inner city schools and the suburbanization of families in Melbourne. *Int. J. Urban Reg. Res.* 41, 443–463.
- Nguyen, N., Cohen, D., Huff, A., 2017. Catching the bus: a call for critical geographies of education. *Geogr. Comp.* 11, e12323.
- Nyerges, T.L., McMaster, R., Couclelis, H., 2011. *Geographic Information Systems and society: a twenty year research perspective*. In: Nyerges, T.L., McMaster, R., Couclelis, H. (Eds.), *The SAGE Handbook of GIS and Society*. SAGE Publishers.
- Oyarzún, J.D.D., 2020. Tensions between education and development in rural territories in Chile: neglected places, absent policies. In: Papa, R. (Ed.), *Handbook on Promoting Social Justice in Education*. Springer Nature Switzerland.
- Parsons, E., Chalkley, B., Jones, A., 2000. School Catchments and Pupil Movements: a case study in parental choice. *Educ. Stud.* 26, 33–48.
- Peck, J., Theodore, N., 2015. *Fast Policy: Experimental Statecraft at the Thresholds of Neoliberalism*. University of Minnesota Press, Minneapolis.
- Pedroni, T.C., 2011. Urban shrinkage as a performance of whiteness: neoliberal urban restructuring, education, and racial containment in the post-industrial, global niche city. *Discourse* 32, 203–215.
- Piattoeva, N., Saari, A., 2020. Rubbing against data infrastructure(s): methodological explorations on working with(in) the impossibility of exteriority. *J. Educ. Pol.* 1–21.
- Pini, B., Gulson, K.N., Kraftl, P., Dufty-Jones, R., 2017. Critical geographies of education: an introduction. *Geogr. Res.* 55, 13–17.
- Pykett, J., 2009. Personalization and de-schooling: uncommon trajectories in contemporary education policy. *Crit. Soc. Pol.* 29, 374–397.
- Reid, J.-A., 2017. Rural education practice and policy in marginalised communities: teaching and learning on the edge. *Austr. Int. J. Rural Educ.* 27, 88.
- Richards, M.P., 2014. The gerrymandering of school attendance zones and the segregation of public schools: A geospatial analysis. *Am. Educ. Res. J.* 51, 1119–1157.
- Rizvi, F., Lingard, B., 2010. *Globalizing Educational Policy*. Routledge, London.
- Rowe, E., 2015. Theorising geo-identity and David Harvey's space: school choices of the geographically bound middle-class. *Crit. Stud. Educ.* 56, 285–300.
- Rowe, E., 2017. *Middle-class School Choice in Urban Spaces: The Economics of Public Schooling and Globalized Education Reform*. Routledge, New York.
- Rowe, E., 2020. Buying and selling the public school in the market: the politics of space and boundary crossings for urban school choosers. In: Thompson, G., Hogan, A. (Eds.), *Privatisation and Commercialisation in Public Education: How the Public Nature of Schooling Is Changing*. Routledge, New York and London.
- Said, E., 1993. *Culture and Imperialism*. Vintage Books, New York.
- Saporito, S., Sohoni, D., 2006. Coloring outside the lines: racial segregation in public schools and their attendance boundaries. *Sociol. Educ.* 79, 81–105.
- Smith, J.L., Lupton, R., 2008. *Mixed Communities: Challenges for Urban Education Policy*. Taylor & Francis.

- Smith, N., 1982. Gentrification and uneven development. *Econ. Geogr.* 58, 139–155.
- Soja, E.W., 1996. *Thirdspace: Journeys to Los Angeles and Other Real-And-Imagined Spaces*. Blackwell, Malden, Ma.
- Soja, E.W., 2010. *Seeking Spatial Justice*. University of Minnesota Press, Minneapolis.
- Takayama, K., Lingard, B., 2019. Datafication of schooling in Japan: an epistemic critique through the 'problem of Japanese education'. *J. Educ. Pol.* 34, 449–469.
- Tate, W.F., 2012. *Research on Schools, Neighborhoods, and Communities: Toward Civic Responsibility*. Rowman & Littlefield.
- Taylor, C., 2002. Geography of the 'new' Education Market: Secondary School Choice in England and Wales. Ashgate, Hampshire.
- Taylor, C., 2007. Geographical information systems (GIS) and school choice: the use of spatial research tools in studying educational policy. In: Gulson, K.N., Symes, C. (Eds.), *Spatial Theories of Education: Policy and Geography Matters*. Routledge, New York, NY.
- Thiem, C.H., 2009. Thinking through education: the geographies of contemporary educational restructuring. *Prog. Hum. Geogr.* 33, 154–173.
- Thompson, G., Cook, I., 2014. Education policy-making and time. *J. Educ. Pol.* 1–16.
- Thrift, N., 2003. Performance and *Environ. Plann.* 35, 2019–2024.
- Waitoller, F.R., Lubienski, C., 2019. Disability, race, and the geography of school choice: toward an intersectional analytical framework. *AERA Open* 5, 2332858418822505.
- Waitoller, F.R., Super, G., 2017. School choice or the politics of desperation? Black and Latinx parents of students with dis/abilities selecting charter schools in Chicago. *Educ. Pol. Anal. Arch.* 25, 1–46.
- Willis, P., 1981. *Learning to Labour: How Working Class Kids Get Working Class Jobs*. Columbia University Press.
- Yoon, E., Lubienski, C., 2017. How do marginalized families engage school choice in inequitable urban landscapes? A critical geographic approach. *Educ. Pol. Anal. Arch.* 25, 1–25.
- Yoon, E., Lubienski, C., 2018. Thinking critically in space: toward a mixed-methods geospatial approach to education policy analysis. *Educ. Res.* 47, 53–61.
- Yoon, E., Gulson, K.N., Lubienski, C., 2018a. A brief history of the geography of education policy: ongoing conversations and generative tensions. *AERA Open* 4, 2332858418820940.
- Yoon, E., Lubienski, C., Lee, J., 2018b. The geography of school choice in a city with growing inequality: the case of Vancouver. *J. Educ. Pol.* 33, 279–298.
- Yoon, E., Marmureanu, C., Brown, R.S., 2020. School choice and the polarization of public schools in a global city: a Bourdieusian GIS approach. *Peabody J. Educ.* 95, 229–247.
- Yoon, E., 2017. Neoliberalizing race? Diverse youths' lived experiences of race in school choice. *Res. Educ.* 97, 76–94.
- Yoon, E., 2020. School choice research and politics with Pierre Bourdieu: new possibilities. *Educ. Pol.* 34, 193–210.

Southern theory as an educational project: Shinto, self-negation and comparative education

Keita Takayama, Graduate School of Education, Kyoto University, Kyoto, Japan

© 2023 Elsevier Ltd. All rights reserved.

Doing southern theory	37
Ecofeminist and decolonial turns	38
Shinto's decolonial potentials	39
Politics of shinto	41
Salvaging shinto	42
Conclusion with self-negation	42
References	43

Doing southern theory

Raewyn Connell (2007) coined the term Southern Theory in her book *Southern Theory: The Global Dynamics of Knowledge in Social Science*. The term South is deployed to draw attention to the “periphery-center relations in the realm of knowledge” (Connell, 2007, p. viii). Her intension is not to “name a sharply bounded category of states or societies, but to emphasize relations—authority, exclusion and inclusion, hegemony, partnership, sponsorship, appropriation—between intellectuals and institutions in the metro-pole and those in the world periphery” (p. viii–ix). By recognizing the South as a source of critical theoretical insights, Connell demanded that social sciences address the differential global power relations in knowledge practice, which have left the existing theoretical tools parochial at best. Her work forces us to problematize the kind of relations often implicitly accepted in the process of academic knowledge production and invites us to do research differently, in a way that subverts the prevailing structure of knowledge. Drawing on her work as well as other cognate postcolonial and decolonial scholarship, I have, along with other colleagues, explored what it means to do Southern Theory in the field of comparative and international education (see Takayama, 2018, 2020a,b, 2015; Takayama et al., 2016; Takayama et al. 2017).

In my view, doing Southern Theory involves two critical moves. First, it requires us to explicitly recognize the “othered” lands and peoples as a source of intellectual work with global implications. Southern theory departs from the usual relations in social scientific knowledge generation, where the “othered” lands and peoples are treated as “data mine” or “informants”. It repositions them as a source of multicentric intellectual work and theoretical insights whose implications extend beyond the original places of generation. That is, it is a call for epistemic justice on the planetary scale. Second, it involves putting those insights and knowledges generated in the othered lands in critical dialogue with the Eurocentric knowledge widely circulated as “universal”. It involves provincializing, peculiarizing and then reconstituting the so called “foundational knowledge” or “theories” of a given discipline. It distances itself from postcolonial politics of resentment and nativist essentialism that simply ascribe to and solidify the colonial division of the world (Bhambra, 2014; Chen, 2010). Instead, it calls for a critical engagement with the existing disciplinary knowledge. That is, it recognizes both the indispensability and inadequacy of the existing disciplinary knowledge developed under the strong influence of Western modernity/coloniality (Chakrabarty, 2000). A project could be called Southern Theory, postcolonial or decolonial, depending on the specific intellectual lineages drawn upon, which inform the particular diagnosis of the challenges and the strategies proposed to address them (Alatas, 2006; Bhambra, 2014; Chen, 2010; Connell, 2007; Kurasawa, 2004; Mignolo, 2011).

In pursuing Southern Theory work over the last decade or so, I have attempted to operationalize Connell's notion as an educational project while grounding it in the field of comparative and international education. Firstly, it is an educational project because doing Southern Theory demands that we take up a role as a “teacher” in relation to fellow researchers both in and outside education. It involves inviting others to take the risk of venturing into the unfamiliar intellectual world that sits outside the academic centers of the “West” so as to broaden their epistemic horizons. Assuming the role of teachers means that those who advocate Southern Theory must demonstrate to others how it can be done and what implications, both intended and unintended, it might generate for one's research work. A self-reflective account on the very process of doing this work and its complexities, including this article, forms a valuable pedagogic resource.

Secondly, it is an educational project because doing Southern Theory necessarily requires a series of learning and unlearning, or “negations” (Takayama, 2020b). It demands that we question the taken for granted premises of disciplinary knowledge and reconstitute it through a set of insights that have been historically excluded from the discussion. But this epistemological shift is incomplete without an equally transformative ontological shift, that is, a shift in how we make sense of our being and relationality to others (Takayama, 2018, 2020b). Taking cues from Keiji Nishitani, Andrea English and Tim Ingold among others, I have elsewhere situated it at the core of the Southern Theory project to transcend the modern empiricist separation between knowing and being and to embrace the view that a shift in knowing necessarily results in a corresponding shift in being (see Takayama, 2020b).

The proposed ontological shift is what makes Southern Theory an indispensable part of the notion of comparative education as cultural critique, as recently articulated by Rappleye (2020). Drawing on his extended transnational experiences, Rappleye (2020) puts forward the notion of “immigracy of being” as the key disposition that invites “a deep experience of Otherness” to “completely shift the ontological ground” of the researcher self (p. 15). In his view, to compare necessarily involves a cultural critique of self and by extension self-negation. The same cross-cultural critique was at the heart of some of the founding scholars of sociology, including Marx, Weber and Durkheim (Kurasawa, 2004). Kurasawa (2004) coins the same methodological disposition “ethnological imagination,” which he defines as “a critical and cross-cultural hermeneutics of Euro-American modernity whereby engagement with other societies has been essential to the project of self-understanding and self-critique of their own times and places” (p. 4). Practicing ethnological imagination is deeply self-transformative, as “learning about others is not simply an act of cosmopolitan open-mindedness, but an integral part of learning about ourselves and even viewing ourselves as other” (p. 5). What is proposed here is not a call for postmodern relativism, but rather to let “the force of difference” release our imaginations (Geertz, 2000, p. 259). Doing Southern Theory sees disruption and self-negation as central to its educational project. It is necessarily disruptive, as it forces us to particularize and peculiarize the deep-seated assumptions shaping what and how we know and how we exist. It demands that we let go of a secured sense of self, acknowledge the partiality of our knowing and unlearn things that have been embraced as part of what makes us who we are (see Rappleye, 2020; Takayama, 2020b).

Building on the discussion of Southern Theory as a pedagogical project, I am going to deploy my own experience/reflection as a central pedagogical tool. The article documents how I have done Southern Theory over the years, with a hope of inviting others to pursue this line of knowledge practice. It is a story of learning through disruption and self-negation, or cross-cultural unlearning, whereby I came to recognize my “home” as the other in an unexpectedly way, and in the course of unlearning, struggled to let go of my secured sense of self. It demonstrates how I learned to embrace the discomfort of having my own subjectivity and “expertise” challenged as a point of departure for renewed understanding of self and others.

Ecofeminist and decolonial turns

My latest unlearning journey began when I was introduced to ecofeminist and decolonial literature through a special issue project in Shanghai (see Silova et al., 2020; Takayama, 2020a). In a nutshell, both ecofeminist and decolonial literatures open up space for different epistemologies, ontologies, and cosmologies that have been suppressed by the global spread of Western modernity-coloniality nexus and global capitalism. Ecofeminism recognizes the more-than-human worlds and stresses the need for the coexistence of humans with multispecies communities. Triggered by the human-induced ecological challenges of planetary scale and learning from the Indigenous communities of Australia, Rose (2005), for instance, rejects the human exceptionalism of the Anthropocene era and rearticulates humans as a member of multispecies communities. Rose (2005) draws on her anthropological work with the Australian Indigenous people in Northern Territories to unsettle the human centric premises of Western philosophy. By describing how the Indigenous concept of “country” operates upon the mutual entanglement of benefits that cuts across the human-nature dualism, she puts forward an alternative vision of philosophical ecology that decenters human agency, knowledge, and intentions.

In addition to anthropological studies with Indigenous people, historical studies of the world that once existed is another source of wisdom. To encourage us to remember what has led us to the current epistemological and ontological conditions of modernity and what has been done to those who dare to be different, Stengers (2012) asks us to remember “the smell of burning witches”. It is an evocative reminder that there existed social worlds where witches and fairies were embraced as part of the “real” world of humanity and where the distinctions between science and superstition were left opaque and constantly transgressed. The hegemony of scientific rationality and its underside, the eradication of “other” worlds, were not a natural consequence of human progress but of ontological violence. By smelling the burning of witches’ flesh, we could develop “our closeness with those who have already been destroyed in the name of rationality, objectivity, and the great divide between nature and culture” (Stengers, 2011, p. 58).

Much of this literature has been taken up in recent “common worlds” educational scholarship, in particular within early childhood education (see Blaise, 2015; Taylor, 2017). These scholars push us to imagine education where humans learn to decenter themselves and reposition themselves as part of multispecies common worlds, and to recognize the agencies of non-human others. They raise important questions about the current articulation of education for sustainability which preserves the central humanistic assumptions, including the logic of human exceptionalism. Taylor (2017), for instance, challenges the positioning of children as environmental stewards, and humans, in general, as the sole agents for solving the current ecological crisis. Furthermore, these post-humanist and ecofeminist studies call upon us to recognize how existing early childhood pedagogic practices already allow for transgressive space where children disrupt the nature-culture binary and where children interact with the more-than-human worlds (Taylor, 2017). Silova (2019) also suggests the tenacity of the “creature communities” today and how social science, including education scholarship, has been blinded to numerous students’ encounters with the more-than-human worlds during their schooling.

Decolonial scholarship, represented by Mignolo (2011) and other South American thinkers (e.g., Grosfoguel, 2011), raises a similar set of questions about the powerful genocidal effects of the global spread of Western modernity-coloniality nexus, which began in the late 15th century European colonization of America. The South American decolonial project historicizes with the intention of de-historicizing it, the current structure of knowledge and recognizes the centuries of Western colonial violence as the constitutive aspect of its “epistemicide”, the obliteration of other epistemologies and ontologies from the surface of the planet

(Grosfoguel, 2011). It problematizes the modernist, Cartesian premises of science and knowledge that set up the hierarchy of knowing and being and privileged for European countries as the only legitimate source of knowledge and science. On the basis of this historical critique, it calls for the resuscitation of marginalized epistemologies and ontologies in colonized parts of the world that have been pushed to the limits by the Eurocentric geopolitics of knowledge. It is a call toward the world of “pluriversity” where multiple epistemologies and ontologies coexist side by side (Mignolo, 2011).

In a number of places, Mignolo (2011) uses de-westernization and decoloniality to distinguish between different levels of critique of Western modernity/coloniality project. According to Mignolo, de-westernization is a shallower critique of the West in that it rejects the West while maintaining its allegiance to its fundamental logic of modernity. In economy, for instance, de-westernization refers to the rejection of US or World Bank interventions in South America, while still aspiring for a regionally or nationally based model of “development” and “progress” that accepts the same economic rationality and exploitative logic of modernity toward the planet. In the domain of knowledge, de-westernization refers to the rejection of knowledge generated in the West, while still accepting the fundamental philosophical foundations of Western knowledge, namely its epistemology and ontology. Drawing on Mignolo (2011), Silova et al. (2020) recognize that the most penetrating critiques of the modernity/decoloniality nexus are to be found in the decoloniality, not the de-westernization, project, particularly in the realm of spirituality and ontology.

Needless to say, the decolonial critique of Western modernity/coloniality resonates closely with ecofeminist scholarship. Most crucially, both identify the role of culture in constituting the violence of modernity/colonial nexus, more specifically in securing the hyperseparation between adults and children, humans and nature, men and women, civilized and barbaric, and science and myth (see Plumwood, 1993). The Western Enlightenment notion of reason provides the logic of hyperseparation with the former terms defined as the embodiment of reason and the latter terms as “everything that reason excludes” (Plumwood, 1993, p. 20). Both decolonial theory and ecofeminist theory encourage us to engage in cross-cultural critique and practice ethnological imagination: the particularization and peculiarization of now and here. They invite us to imagine different ways of being and knowing that have been suppressed, or even erased, by a set of cultural norms and practices of modernity and imposed through colonial violence.

The above exploration of the ecofeminist and decolonial scholarship led me to one interesting/uncomfortable observation; the critique of modernity/coloniality resonates with what I knew about Shinto, a Japanese animism/religion. Growing up in Japan, I intuitively knew that the nature-centered world view and the rejection of human exceptionalism echoes the core ideas of Shinto animism, although I grew up in a non-traditional household where Shinto was not particularly embraced. It was uncomfortable, because I was aware of the tainted history of Shinto, its close link to Japan’s imperial aggression during the past wars and its postwar association with the nativist politics. Indeed, my previous scholarship was scathingly critical of any attempt to invoke Shinto-informed spiritualities in Japanese education (see Takayama, 2010, 2008). I was not sure how I could possibly reconcile the idea of drawing on Shinto as a potential intellectual resource toward ecologically minded and decolonial education with its unfortunate past. It was with a deep sense of discomfort and uncertainty that I ventured to learn about Shinto.

Shinto’s decolonial potentials

As Jensen and Blok (2013) point out, “Japan is probably the only major industrialized country in which widespread discussion of animism is still a part of ordinary intellectual discourse” (p. 97). Indeed, the Shinto-inspired, animated worldview pervades all of Japanese society, underpinning mundane aspects of life as well as art forms and cultural practices (Carter, 2013; Nakayama, 2019). Shinto values and practices are so enmeshed in Japanese people’s everyday life that they have become their second nature (Kasuils, 2004). According to Nelson (2000), shrine Shinto is “one of the most long-lived of all Japan’s institutions, largely because (after nearly fourteen centuries) it continues to help form, orient, and empower a sense of local and ethnic identity” (p. 3).

True to much of animisms in the world, Shinto stresses the greatness of the universe and the relative insignificance of human presence in its entire history. It also recognizes the agencies of the more-than-human worlds and their spiritual impacts on humans. Shinto locates spirits in both humans and non-humans, including stones, rivers, trees, foxes, thunder, ancestors, rice and waterfalls, that is, radical personalization of the universe (Jensen and Blok, 2013). According to Shinto principles, “gods, men, animals, plants and inanimate objects are mutually permeable entities, appearing as a unified and dynamic field of existence, characterized by particular forms of immanence and vitalism” (Jensen and Blok 2013, p. 97).

The Shinto cosmology continues to influence the meaning of the term *shizen*, the Japanese translation of nature, the Western concept introduced to Japan in the late 19th century. As Nakayama (2019) explains, the Japanese concept of nature does not contain the hierarchical Christian idea, where the creator of the world God exists at the top, then man created in the image of God, with all the other creatures comprising nature at the bottom. Here, “God as a transcendent being does not exist within nature, nor are human beings a part of it” (p. 8). This dichotomy between human beings and nature was central to the emergence of the new science in the 17th century, further developed via Bacon and Descartes into the Cartesian worldview and subsequently the conceptualization of nature merely as an object to be controlled by humans (Plumwood, 1993; Hickel, 2020). In contrast, the Shinto animistic view of nature recognizes something sacred in all the creatures who are both physically and spiritually a part of nature. Hence, it defies the usual “opposition ... between human subjectivity and natural objectivity” (Nakayama, 2019, p. 9).

The same reverence toward the more-than-human worlds is central to the Japanese thought and philosophy, according to Robert Carter (2013). In explaining the central component of the Zen/Shinto philosophy, Carter quotes a Japanese landscape architect/Zen monk, Shunmyo Matsuo:

I wonder just what kind of spirit a certain stone has and how it would prefer to be set out. This is also true of plants and I always consider how I think the plants would like to be displayed. I always feel at one with the plants, when I am planting them, and with the stones, when I am arranging them. (p. 34)

Carter (2013) refers to Matsuo in his illustration of the key thinkers of the Kyoto School of Philosophy, Kitaro Nishida and Keiji Nishitani, who were likewise influenced immensely by the Zen and Buddhist thought (see Sevilla, 2016). Matsuo's approach to landscaping epitomizes the kind of empathic identification with non-animated objects that transcend the modernist culture-nature, subject-object dualism. Through self-cultivation, one learns to be one with an object (flowers, stones, and trees); it is a state of nothingness where one comes to know a thing not through reason and language but through intuitions developed through direct experience. Much of this Shinto-informed sensibilities toward nature permeate various Japanese arts, including traditional garden architecture, calendar pictures and poetry (see Kato, 2021).

Though Shinto is not explicitly taught in Japanese education due to its principle of secularity (Nakayama, 2019), research has shown that the Japanese curriculum, most notably in the Japanese language textbooks, implicitly teaches children emotional identification with non-human creatures and their worldviews (Gerbert, 1993; Ishihara, 2005). In comparing primary school readers in Japanese and American schools, Gerbert (1993) highlights the nature centered view of the world and the relative insignificance of humans in Japanese school readers. Her analysis shows that in many of the stories included in Japanese readers, "human protagonists drop out of the picture altogether" (p. 162). In one of the texts that she closely examines, "(t)he self identifies with and merges with nature. It never become a fully constituted "personality" as often seen in American readers (p. 164). Japanese children are encouraged to develop "a passive attitude toward nature" (p. 163), "be sensitive to small changes in the environment" (p. 163) and "quietly lose the self in the contemplation of nature" (p. 165). This is contrasted to the American primary school readers where an explicit emphasis is placed on formal reasoning, analytical thinking and a strongly anthropocentric view of nature.

More recently, the Ministry of Education introduced moral education supplementary texts, arguably to counterbalance the increasing (Western-inspired) emphasis on critical literacy (Ishihara, 2005). What follows is an excerpt from one of the moral education booklets, *Kokoro no nooto* (*Notebook for Heart*):

Humans are moved by beauty.
When faced with the magnificence of nature, we feel moved and hold our breath.
Calm, great ocean that spreads endlessly
Vast hills and fields, and towering mountain ranges.
It feels as though they mercifully embraced us, as though we melted into them.
But, would this be true?
When nature bares its fangs
it engulfs us with its overwhelming might.
Thundering noise of crashing waves, the volcanic smoke that shuts out the sky
and the massive earthquake that shatters the earth.
A feeling of awe and respect towards the existence of matters beyond human control springs up. (MEXT, 2002, p. 65)

Similar texts are also found in more recently published moral education readers.

To feel the greatness
Rainbows after rain
the bright red sun about to set
shining aurora
drifting ice masses approaching the seashore
and splashing waterfalls.
(Our) heart moves with the overwhelming
presence of natural phenomena and sceneries.
When faced with supernatural world beyond
humans, (we) experience the beauty and
greatness of nature.
They move (us), and (we) appreciate human
Heart (MEXT, 2014, p. 114).

In these passages, children are to leave behind the modern scientific form of learning where they are to understand, comprehend, and control the awesome. Instead, they are to cultivate "a feeling of awe" (*ikei no nen*) by simply standing "under it, feel themselves to be inherently part of it and it part of themselves" (Kasulis, 2004, p. 167). The MEXT moral education readers, including the ones cited above, are organized around the interconnectedness of self to the broader universe, the environment (nature), family, school, local community, nation and the global, or an interdependent sense of selfhood (Komatsu et al., 2021). As Japanese children progress with the readers, they learn to position themselves within the broader collectives of expanding scales, but they all stress somewhere in the texts the sense of awe toward the overwhelming presence of nature, insignificance of self, and the existence beyond humans.

According to Kasulis (2004), the author of *Shinto: The Way Home*, the feeling of awe is central to Shinto. He argues that Shinto, as a contemporary religion in a highly technological society, is striking in its insistence that awe is not to be understood, nor to be comprehended in any systematic way. The point of Shinto practice is "more to make one feel at home with awe rather than try

to understand or control it" (p. 167). Kasulis argues that in contemporary modern societies, we have lost this attitude toward the awesome. According to him, "one result of the predominance of scientific thinking is that today our initial response to the awesome is to try to understand it rather than to stand under it. Instead of filling us with a sense of humility before the unknown, awe has come to challenge us as only the not-yet-known. I don't know has become an ego-bruising admission of ignorance instead of a sign of wisdom" (p. 167). Following Shinto, Kasulis invites us to "accept the awesome as part of the world in which we live" (p. 12). Kasulis seems to suggest that the feeling of awe has much to do with letting go of our sense of ego and embrace ignorance as virtue.

It is this rejection of liberal principle of rationality and critical reason and the restricted sense of self that results from it that are crucial for Bowers' (1993, 1987) articulation of eco-justice pedagogy. In his view, liberal and critical educational thoughts, represented by Dewey, Skinner and Freire among others, are fundamentally antithetical to the generation of culture that is ecologically sustainable. This is because the liberal theories of education accord the absolute status to reason and rationality as practiced by individuals, and this undermines the authority of traditional patterns of knowledge. As he states, liberal educational thoughts are guided by the Enlightenment notion of rational, autonomous and self-directing self, and this conceptualization of self as "the epicenter of the social world" (p. 23) has the effect of "relativiz(ing) the communal foundations of a shared sense of moral authority, with the consequence that individual judgment reflects what is perceived as useful, fulfilling, and pleasurable" (Bowers, 1987, p. 25). With "the subjectivism of personal experience" as "the final refuge" (p.28), it breeds nihilism; it erodes "the sense of being interdependent with the large social and biotic community," which is defined as "an unwelcomed constraint on individual freedom" (Bowers, 1993, p. 27). Hence, the crucial shortcoming of liberal educational theories in light of their potentials for ecologically balanced world is their inability to offer the self-limiting principle, the very cultural condition for small and interdependent notion of selfhood. As a possible source of the alternatives, Bowers draws attention to indigenous, place-based and intergenerational knowledges (see also Hickel, 2020).

Most recent education scholarship for sustainable future reinforces Bowers' conclusion reached more than a few decades ago. For instance, Komatsu et al. (2021) undertook a series of quantitative analyses to identify the possible relationship between what they call ontological individualism, which is diametrically opposed to relational, interdependent selfhood, social unsustainability and student-centered teaching in schools. They establish the empirical foundation for the very assertion made by Bowers more than three decades earlier; student-centered teaching, underpinned by the liberal principle of rational and autonomous self, can accelerate social and ecological unsustainability. Then, they take up the suggestion made by Bowers (1993), to "study other cultures that have evolved in more ecologically balanced ways" (57), and discuss three examples of alternative pedagogic practices designed to promote interdependent sense of self, namely from Japan, Botswana and minority groups within USA. As an intergenerational body of local knowledge, Shinto can be one of such "othered" cultures that have not only survived the imposition of modernity but is actually thriving today. It can serve as much needed resources to promote the kind of fundamental cultural shift required toward ecological sustainable, post-growth society (Hickel, 2020).

Politics of shinto

Endorsing the Shinto animism and its cosmologies in Japanese schooling, however, is a complicated matter. This is because of the highly politicized context within which Shinto has been articulated within Japanese education throughout its postwar history (see Koyasu, 2004). The use of Shinto beliefs in schooling has been intensely contested since the end of the Asia Pacific War. During the wartime, Shinto was incorporated into the state apparatus and played a central role in interpellating the people into the wartime imperialist ideology. Public education was the key state mechanism through which the widely accepted Shinto beliefs in nature were mobilized for ideological indoctrination (Hardacre, 2017; Shimazono, 2003). The postwar US-led Occupation ordered the complete removal of Shinto religious elements from state apparatus and put in place constitutional mechanisms to ensure separation of church and state and freedom of religion (Koyasu, 2004). The Occupation blamed Shinto as the ideological source of Japan's wartime ultra-nationalism and imperialism, while turning blind eyes to the roles played by other religions (Hardacre, 2017). Its hunt for any remnants of the ostensibly Shinto-informed wartime ideology was so comprehensive that it even banned Japanese martial arts, calligraphies, and other benign cultural artifacts and practices (Shibata, 2004; Tanikawa, 2021).

The US-led Occupation's "demilitarization" of education planted the seed for the postwar political struggles in Japan. Since Japan's formal independence in 1952, those on the political right have consistently demanded the resurrection of the traditional Shinto cosmologies and patriotic values back within formal schooling (see Takayama, 2010, 2008). They call their political agenda the "normalization" of Japanese education, that is, normalizing the "unusual" situation left by the Occupation's imposition of liberal values, which has arguably detached the postwar generations from the nation's cultural "essence". By contrast, the Japanese liberal and left see such a move as a retrogressive desire to reinstate the imperial state and its ideology. They insist on protecting the liberal, democratic and pacifist principles and constitutional frameworks introduced during the US Occupation. Constitutional arguments were frequently made to prevent Shinto-informed values and any patriotic teaching from entering into education, with liberal critics arguing that they violate the constitutionally guaranteed separation of church and state and freedom of thought (Koyasu, 2004; Miyake, 2003; Nishihara, 2006). The wartime indoctrination through education and the people's general remorse for its consequences were the broader historical backdrop against which the concerns have been expressed by liberals and supported by the broader public (Takayama, 2008, 2010).

Situated within this highly politicized policy context, the seemingly benign Shinto-informed concepts such as "awe and respect toward nature", or "matters beyond humans", and "the insignificance of humans in the whole universe" assume highly contested

meanings. Liberals argue that teaching the Shinto-informed worldview reinforces students' passivity toward nature and by extension toward the authority and the state. It is the unquestioned devotion to nature that was then politically appropriated to generate people's allegiance toward the emperor/imperial state and the catastrophic "sacred" war for the imperial household. Instead of teaching children to be passive, hence, the liberal-left critics call for teaching the principles of rationality and critical reason that are central to democratic citizenship (see Irie, 2004; Iwakawa and Funabashi, 2004; Miyake, 2003). To those on the political left, the Shinto-inspired moral education readers, quoted earlier, are nothing but an expression of neo-conservative desire to render people obedient in time of neoliberal economic restructuring. This is exactly the argument that I made in my assessment of the 2006 revision to the Fundamental Law of Education (see Takayama, 2008). As Grosfoguel (2011) points out, nationalism, as a response to the Eurocentric colonial imposition, "reproduces an internal coloniality of power within each nation-state and reifies the nation-state as the privileged location of social change" (p. 23). When Shinto is so tainted with the discriminatory nativist politics of exclusion, could it be "conceptually freed from its historically over-determined connotations of political regression"? (Jensen and Blok, 2013, p. 94).

Salvaging shinto

Would it be possible to salvage "the multivalent ethico-pragmatic character of Shinto" from the nationalistic overtone that has dominated its domestic articulation? (Jensen and Blok, 2013, p. 108). In a sense, Shinto is a placeholder for multiple interests within Japan. The dominant, nationalistic discourse of Shinto, or what Kasulis (2004) calls the highly normative and prescriptive, "essentialist" Shinto spirituality, is certainly with us. But there is also the ad-hoc, flexible, and descriptive form of Shinto as a popular praxis, or what Kasulis (2004) calls the "existentialist" Shinto spirituality, which pervades much of place-based invocations of Shinto cosmologies in festivals, rituals, and mundane life moments. The localized, existentialist Shinto spirituality gives us a way to be radically different today that centers on immanent connectedness of humans and non-humans, hence with considerable implications for reimagining education for sustainable futures. This is the aspect of Shinto that Jensen and Blok (2013) also sees as of great value for science and technology studies, when they argue that Shinto embodies "an alternative politics of the polymorphous enchantments of non-human agency" (p. 108) that can help broaden the theoretical horizon on the entanglements among science, technology, ecology, and cosmos.

However, as Kasulis (2004) reminds us, the two forms of Shinto spiritualities—essentialist and existential—should not be treated as another form of binary. They are "not separate religious traditions but instead overlap in an internal relation with each other" (p. 153). The history of Shinto in Japan has been infused with the tension between these two forms of spirituality, which remains unresolved today (Kasulis, 2004, p. 6). To complicate the matter further, this distinction is not available in the Japanese vocabulary for talking about Shinto, according to Kasulis (2004). Partly because of the nationalist political dominance in the domestic discourse of Shinto, the language of Shinto necessarily implicates essentialist assumptions. This poses immense challenges in terms of "reclaiming" Shinto (Stengers, 2012), that is, teasing out and mobilizing the existential, localized form of Shinto without invoking the essentialist form of Shinto spirituality that has exclusionary effects on the domestic political front. For now, all I can do is sit with the enormity of the conceptual and political challenges and suggest that this conundrum is not necessarily particular to Shinto and Japan. This discussion has wider international implications for education for sustainable futures, given that the political appropriation of nature and the popularized form of animism has been central to the modern formation of nation-state and its physical, epistemic, and ontological violence.

Conclusion with self-negation

The purpose of this article is educational in that it aims to show what it all means to do Southern Theory in educational research. The narrative exposes the sense of vulnerability one might experience as a result of doing Southern Theory and embracing the disruption that it causes. In my case, it all started with engaging with decolonial and ecofeminist literature and realizing its close relevance to the discussion of Shinto animism. Attempting to embrace Shinto as an alternative source of knowledge, however, forced me to unlearn my well-established views, largely shaped by the postwar political discourse of education in Japan, and let go of the political sensibilities. The whole Japanese domestic debate over the place of Shinto cosmologies in education remains trapped within the discursive legacies of the Occupation and the Cold War. Rethinking the role of Shinto in education is part of what Chen calls the de-cold-war politics, which he argues is central to decolonizing imagination in East Asia. It is "to confront the legacies and continuing tensions of the Cold War" so as to "reopen the past for reflection in order to make moments of liberation possible in the future" (Chen, 2010, p. x).

While recognizing the need to rethink the usefulness of postwar binary politics, I remain deeply ambivalent about leaving behind altogether the liberal principles underpinning education. Here I am acutely reminded of the dangers in transferring purely philosophical arguments to the domain of politics that Davis (1998) highlights. In discussing the historical role of Japanese philosophers in promoting Japanese cultural nationalism, Davis makes the following point: "The collapse of subject and object, thought and action—long the aim of Japanese philosophers—may be innocent enough as epistemology or Buddhist soteriology—but it can have a devastating effect when applies to politics" (p. 183). Like Davis, I am worried about the political implications of the Shinto cosmologies in Japanese schools today, when Japanese scholars identify the rise of retrogressive nationalism as well as the

depoliticization of education as the major concerns and call for resuscitating the roles of education for democratic politics (see Hirata, 2015; Kodama, 2013).

Foremost scholar of Japanese traditional unlearning and *mushin* (no mind or nothingness), Nishihira (Nishihira and Rapple, 2021) makes the following observation about the risk of political appropriation of Japanese traditional concepts of learning for ideological indoctrination:

For those of us that write about Japanese thought and tradition, we still operate in the shadow of the war, to some extent. We worry that the element of trust and persistence embedded within this model may get co-opted, in a nationalist context, and transformed into a tool of indoctrination. It is a wariness of getting mistaken for nationalists or apologists for the prewar period, and lingering uncertainties about the possible dangerous potentials of self-negation, that has prevented many Japanese scholars, including myself, from more assertively discussing these models. Future work needs to remain vigilant of this. That said, this “just trust” call to faith is not unique to this “Japanese” model, but is a major issue for virtually all forms of tradition—religious, secular, and otherwise (p. 10).

While Nishihira does not refer to Shinto, the “just trust” call to faith embodies the same rejection of liberal principle of rationality, reason and individuality as expressed in Shinto’s call for awe. It is notable even some like Nishihira, the foremost expert on the Japanese traditional thoughts in education, express reservation about possible political appropriation of his work toward nationalistic agenda.

What Nishihira’s reflection illuminates is the predicament of reclaiming what is indigenous or traditional in a context where Western liberalism and its associated educational thoughts permeate the political and educational discourse and institutions. In the postwar politics of Japanese education, the rejection of liberal ideas, which is always coupled with the assertion of what is traditional and indigenous, has been so tainted with regressive politics of nativism. As a result, the discourse of Shinto in education has been thoroughly politicized, and its potential for sustainable future has remain largely unexplored. In the face of the planetary ecological crisis, however, there is an urgent need to reassess this mainstream discourse of Japanese postwar education scholarship, “where Japanese-ness is automatically equated with negative distinctness, prewar myths, and an escape from the responsibility of making Japan “fully modern”” (Rapple, 2018, p.17).

This is not at all an endorsement of the nationalist attribution of anything liberal to the postwar US “imposition” and of romanticizing of Japan’s imperial past. Instead, it is a call for a careful discerning of what is desirable and otherwise toward both ecologically balanced and politically engaged educational thoughts. The outcome of this discerning work is like to be a creative eclecticism of elements that are Indigenous and modern (Western), inviting students to live in multiple worlds simultaneously (Silova, 2019). In this regard, it is encouraging to know that the daily lives of the people in Japan are saturated with both advanced forms of technology and science and the penetrating presence of Shinto animism. It is a country where rocket scientists visit a local shrine for their successful space endeavors (Nelson, 2000, p. 1). And indeed, as amply demonstrated by Gerbert’s (1993) study, the strongly nature-centered view of the world has been quietly integrated into Japanese schooling.

It is important to note that Gerbert’s insights were only made available to us, as she was willing to look at Japanese primary school readers not through political lens but through cultural ones. This point takes me back to my earlier critique of the English-language comparative education scholarship of Japanese education and forces me to reflect how much I have shifted over time. About 10 years ago, I published a discursive analysis of the whole scholarship of Japanese schooling and identified the continuing Orientalist logic in this body of research (Takayama, 2011). I rejected its predominantly culturalist framing of Japanese schooling, whereby nation-state was unproblematically conceptualized as a cultural “container” and the assumed existence of “fundamental cultural differences” between Japan and the West (usually USA) remained unquestioned. I maintained that the whole scholarship was deeply apolitical and re-inscribed the nationalistic thesis of cultural homogeneity. The critique included some of the best-known scholarship of Japanese schools, including those by Lewis, Tobin, Cave and Tsuneyoshi. The argument in this article has forced me to reconsider my critical assessment of the scholarship. While I still consider valid much of the critique of the predominantly culturalist framing of Japanese education, I now acknowledge my own blind spot; my critique was strongly shaped by the postwar discourse of Japanese education scholarship whose preoccupation with politics trumps any consideration of culture. After all, I would not have been able to acknowledge the tremendous potentials of Shinto animism for eco-justice pedagogy, had these scholars not offered a culturalist reading of Japanese schools (e.g. Gerbert). Culturalist accounts are problematic, so is a categorical rejection of culture.

Dealing with all these tensions and complexities, involving both intra-national and inter-national politics of difference (Ge, 2015; see also Takayama 2020, 2018), is part of what doing Southern Theory entails. It demands constant self-critique and reflecting back on where we have come from and continual search for ways to release ourselves from our habitual modes of thinking and being. Doing all these is a constitutive part of what I consider as doing Southern Theory, and it is one way to make comparative and international education research transformative, meaningful and politically astute at the same time.

References

- Alatas, S.F., 2006. *Alternative Discourses in Asian Social Sciences: Responses to Eurocentrism*. Sage Publications, London.
- Bhabra, G.K., 2014. *Connected Sociologies*. Bloomsbury, New York.
- Blaise, M., 2015. Fabricated childhood: uncanny encounters with the more-than-human. *Discourse* 37 (5), 617–626.
- Bowers, C.A., 1993. *Education, Cultural Myths, and the Ecological Crisis: Toward Deep Changes*. SUNY Press, Albany, NY.

- Bowers, C.A., 1987. *Elements of a Post-liberal Theory of Education*. Teachers College Press, New York.
- Carter, R.E., 2013. *The Kyoto School: An Introduction*. SUNY Press, Albany, NY.
- Chen, K., 2010. *Asia as Method: Deimperialization*. Duke University Press, Durham.
- Chakrabarty, D., 2000. *Provincializing Europe: Postcolonial Thought and Historical Difference*. Princeton University Press, Princeton, NJ.
- Connell, R., 2007. *Southern Theory: The Global Dynamics of Knowledge in Social Science*. Allen and Unwin, Crows Nest.
- Davis, W., 1998. Religion and national identity in modern and postmodern Japan. In: Heelas, P. (Ed.), *Religion, Modernity and Postmodernity*. Blackwell, Oxford, pp. 169–185.
- Ge, S., 2015. *Ajia wo katarukoto no jirenma: chiteki kyoudou kuukan wo motomete* [Dilemma of speaking about Asia: Towards a shared intellectual space]. Iwanami shoten, Tokyo [in Japanese].
- Geertz, C., 2000. *Available Lights: Anthropological Reflections Philosophical Topics*. Princeton University Press, Princeton, NJ.
- Gerbert, E., 1993. Lessons from the "Kokugo" (National Language) Readers. *Comp. Educ. Rev.* 37 (2), 152–180.
- Grosfoguel, R., 2011. Decolonizing post-colonial studies and paradigms of political-economy: Transmodernity, decolonial thinking, and global coloniality. *Transmodernity J. Peripher. Cult. Prod. Luso-Hisp. World* 1 (1), 1–38.
- Hardacre, H., 2017. *Shinto: A History*. Oxford University Press, Oxford.
- Hickel, J., 2020. *Less is More: How Degrowth Will Save the World*. Windmill Books, London.
- Hirota, T., 2015. *Kyoudou wa nani o subekika: noryoku, shokugyou, shimin* [What education ought to do: Competency, occupation and citizenship]. Iwanami shoten, Tokyo [in Japanese].
- Irie, Y., 2004. *Kyokasho ga abunai: "kokoro no nouto" to koumin/rekishi* [Dangerous textbooks: "Notebook for Heart" and Civil/History]. Iwanami shoten, Tokyo [in Japanese].
- Ishihara, C., 2005. *Kokugo Kyokasho no shisou* [The philosophy of the national language textbook]. Chikuma shobou, Tokyo [in Japanese].
- Iwakawa, N., Funabashi, K. (Eds.), 2004. *"kokoro no noto" no ho e wa ikanai* (I won't be going towards "the notebook of heart"). Terakaya Shinsha, Tokyo [in Japanese].
- Jensen, C.B., Blok, A., 2013. Techno-animism in Japan: Shinto cosmograms, actor-network theory, and the enabling powers of non-human agencies. *Theory Cult. Soc.* 30, 84–115.
- Kasulis, T.P., 2004. *Shinto: The Way Home*. University of Hawaii Press, Honolulu.
- Kato, M., 2021. The educational function of Japanese arts: an approach to environmental philosophy. *Educ. Philos. Theor.* <https://doi.org/10.1080/00131857.2021.1904396>.
- Kodama, S., 2013. *Gakuryoku Gensou* [Illusion of Academic Achievement]. Chikumashobou, Tokyo [in Japanese].
- Komatsu, H., Rappleye, J., Silova, I., 2021. Student-centered learning and sustainability: solution or problem? *Comp. Educ. Rev.* 65 (1). <https://doi.org/10.1086/711829>.
- Koyasu, N., 2004. *Kokka to shisai: Kokka Shintou no genzai* [Nation-state and Shinto ceremonies: The current state of State Shintoism]. Tokyo Seidosha [in Japanese].
- Kurasawa, F., 2004. *The Ethnological Imagination: A Cross-Cultural Critique of Modernity*. University of Minnesota Press, Minneapolis MN.
- MEXT, 2014. *Watashitachi no dotoku, shogakko 5.6 nen* [Our morality, Year 5 and 6]. MEXT, Tokyo [in Japanese].
- MEXT, 2002. *Kokoro no no-to, chugakkou* [Notebook for Heart for Junior High School]. Akatsuki kyoudou tosho, Tokyo [in Japanese].
- Mignolo, W., 2011. *The Dark Side of Western Modernity: Global Futures, Decolonial Options*. Duke University Press.
- Miyake, A., 2003. "Kokoro no no-to" wo kangaeru. [A thought on "the Notebook of Heart"]. Iwanami shoten, Tokyo [in Japanese].
- Nakayama, O., 2019. New spirituality in Japan and its place in the teaching of Moral Education. *Religions* 10 (4), 278. <https://doi.org/10.3390/rel10040278>.
- Nelson, J.K., 2000. *Enduring Identities: The Guise of Shinto in Contemporary Japan*. University of Hawaii Press, Hawaii.
- Nishihara, H., 2006. *Ryoushin no jiyuu to kodomotachi* [Freedom of conscience and children]. Iwanamishoten, Tokyo [in Japanese].
- Nishihira, T., Rappleye, J., 2021. Unlearning as (Japanese) Learning. *Educational Philosophy and Theory*. <https://doi.org/10.1080/00131857.2021.1906644>.
- Plumwood, V., 1993. *Feminism and the Mastery of Nature*. Routledge, New York.
- Rappleye, J., 2018. In favor of Japanese-ness: future directions for educational research. *Educ. Stud. Japan* 12, 9–21.
- Rappleye, J., 2020. Comparative education as cultural critique. *Comp. Educ.* 56 (1), 39–56.
- Rose, D.B., 2005. An Indigenous philosophical ecology: Situating the human. *Aust. J. Anthropol.* 16 (3), 294–305.
- Sevilla, A., 2016. Education and empty relationality: thoughts on education and the Kyoto School of Philosophy. *J. Philos. Educ.* 30 (4), 639–654.
- Shibata, M., 2004. *Japan and Germany under the U.S. Occupation: A Comparative Analysis of the Post-War Education Reform*. Lexington Books, Lanham, MD.
- Shimazono, 2003. State Shinto and the religious structure of modern Japan. *J. Am. Acad. Relig.* 73 (4), 1077–1098.
- Silova, I., 2019. Toward a wonderland of comparative education. *Comp. Educ.* 55 (4), 444–472.
- Silova, I., Rappleye, J., You, Y., 2020. Beyond the Western horizon in educational research: toward a deeper dialogue about our interdependent futures. *ECNU Rev. Educ.* 3 (1), 46–65.
- Stengers, I., 2012. Reclaiming Animism. *E-Flux Journal* #36. <https://www.e-flux.com/journal/36/61245/reclaiming-animism/>.
- Stengers, I., 2011. Comparison as a matter of concern. *Common. Knowl.* 17 (1), 48–63.
- Takayama, K., 2020a. Engaging with the more-than-human and decolonial turns in the land of Shinto cosmologies: "Negative" comparative education in practice. *ECNU Rev. Educ.* 3 (1), 46–65.
- Takayama, K., 2020b. An invitation to "negative" comparative education. *Comp. Educ.* 56 (1), 79–95. <https://doi.org/10.1080/03050068.2019.1701250>.
- Takayama, K., 2018. Towards a new articulation of comparative educations: cross-culturalising research imaginations. *Comp. Educ.* 54 (1), 77–93.
- Takayama, K., 2015. Provincialising the world culture theory debate: critical insights from a margin. *Glob. Soc. Educ.* 13 (1), 34–57.
- Takayama, K., 2011. Other Japanese educations and Japanese education otherwise. *Asia Pac. J. Educ.* 31 (3), 345–349.
- Takayama, K., 2010. From the Rightist "coup" to the new beginning of progressive politics in Japanese education. In: Apple, M.W. (Ed.), *Global Crises, Social Justice, and Education*. Routledge, New York, pp. 61–111.
- Takayama, K., 2008. Japan's Ministry of Education "becoming the right": Neoliberal restructuring and the Ministry's struggles for political legitimacy. *Glob. Soc. Educ.* 6 (2), 131–146.
- Takayama, K., Sriprakash, A., Connell, R., 2017. Toward a postcolonial comparative and international education. *Comp. Educ. Rev.* 61 (S1), 1–24.
- Takayama, K., Heimans, S., Amazan, R., Maniam, V., 2016. Doing Southern theory: towards alternative knowledges and knowledge practices in/for education. *Postcolonial Dir. Educ.* 5 (1), 1–25.
- Tanikawa, T., 2021. *Baseball in Occupied Japan: US Postwar Cultural Policy*. Kyoto University Press, Kyoto.
- Taylor, A., 2017. Beyond stewardship: common world pedagogies for the Anthropocene. *Environ. Educ. Res.* 23 (10), 1448–1461.

Transnationalism

Niranjan Casinader, Faculty of Education, Monash University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

The problematics of transnationalism as a 21st century concept	45
Transculturalism and contemporary transnationalism	46
New formations of transculturalism	47
The education - transnationalism nexus	48
Transnationalism as an activator of rights in education	49
Acknowledgments	50
References	51

The problematics of transnationalism as a 21st century concept

In the first quarter of the 21st century, the contemporary vision of transnationalism is that it is primarily a transactional concept, involving the movement of phenomena across national borders. In particular, it is often applied in the context of global demographic movement, whether that be temporary and short-term in the undertaking of employment or recreation, or whether it be more permanent, long term and diasporic, as in cases associated with migration or refugee movement. The result is that studies of migration or locational shifts in international population tend to focus on the impact of that geographical shift on the people concerned, such as in the use of language (see, for example, [Kirsch and Gogonas, 2018](#); [Robert and Yu, 2018](#)).

At the most visible level, such an obvious interpretation of transnationalism is not unsurprising, given the extent to which current demographic movement has been linked to modern globalization, which in many ways is symbiotically linked to the notion of international movement across national boundaries. In its initial emergence during the 1990s, the modern phase of globalization was formulated primarily as an economic phenomenon, umbilically linked to an emerging neo-liberal context ([Burbules and Torres, 2000](#)) that emphasized the value of freedom of economic thought and action – the transaction of commerce – across global spaces ([Rizvi and Lingard, 2010](#)). In this, and in concert with its associations with demographic mobility, the contemporary phase of globalization is no different from earlier iterations, such as the Ancient Roman Empire, the naval ventures of 13th Century China or the Christian Crusades of medieval Europe ([McGrew, 2007](#)). However, one of the fundamental principles of the modern sovereign State that makes the present phenomenon different from earlier manifestations of globalization is that it has formal control over activity within its defined territory, a power that it establishes and maintains through the maintenance of its geographical areal location, as well as the physical and psychological integrity of the borders of its demarcated territory. The degree of control over those borders, and thus the effectiveness of a State's power, is further influenced by one of the key criteria for the continued viability of a sovereign State; that is, its existence and the location of its boundaries are accepted, agreed to and supported by the actions of other sovereign States. Conflicts between sovereign States are often expressed in terms of the lack of cogency and acceptance of the boundary positions in question.

However, another defining aspect of contemporary globalization in the first quarter of the 21st century is the frequency, scope and extent of transnational events across national borders, facilitated and fueled by the rapid expansion and accessibility of modern communications technology, ranging from transport to electronic media. As a result, the global human system has evolved into a form where national boundaries are increasingly amorphous and seemingly more porous. Whether economic, social, or cultural, human activity in the 21st Century is characterized by an enhancement in the extensive nature and degree of global connectivities ([Rizvi, 2009](#)). Consequently, it is arguable that the entity of the sovereign State has now lost some of its initial lifeforce, existing now only in terms of creating the international and political environment or skeleton around which geographical exchanges of human activity can take place, accompanied by movements in related necessities for a globalized society. One of these is a workforce that is more mobile, diverse and transitional than had been the case previously. As a result, transnational behavior and actions in the modern idiom are more accurately framed within the context of global corporate power, which possesses the capacity to transcend national boundaries independent of national governments. Ironically, such power has resulted from the freedom given to global corporations by the same sovereign States in the desire for greater and more stable national growth, either individually or through multilateral agreements, brokered through organizations such as the United Nation or World Trade Organization.

The second decade of the century has seen the gradual emergence of national resistance in reaction to the more negative, localized aspects of globalization. Economic restructuring into a global economy, in which national production specialization and global trade agreements has played a key role, has led to the decline and disappearance of many national traditional jobs and industries in various countries. While part of that resistance has been economic, with countries such the USA being more resistant to multilateral international agreements, much of the observable transnational impact has been largely selective and demographic. As a series of global and international socio-political conflicts have led to increases in numbers of refugees and asylum seekers migrating from their places of origin, governments in each inhabited continent have used them as a means of justifying and symbolizing their right to control their national territories. The fluidity and prevalence of contemporary demographic transnationalism has

been invoked by such forces in an attempt to cement the importance of international political boundaries and the maintenance of perceived nationalized cultural identities. Instead of so doing, these actions have only served to highlight the reverse; that is, the dominance that transnationalist behavior and activity continue to exert on global events, as well as the decreasing significance of national borders in the imaginary of those who seek to cross them.

Such realities reinforce the tensions of conceptual structure that are exacerbated in the case of transnationalism because, despite its embodiment of human activity that transcends national borders, its very name is founded on the concept of a defined, delineated space; that is, a sovereign State. Transnationalism only occurs when national borders are crossed, but more extensive and ongoing movement over these political boundaries only leads to a diminution of the significance of national borders in the life imaginaries of those involved. The prevailing notion of transnationalism is therefore a form of paradox. It assumes that there is an underlying constant; that is, the border or boundary to be crossed is permanent and unyielding, clear to all. However, the reality is that the effectiveness of national borders, regardless of any international legality, are dependent on people accepting the existence, reality and validity of a specific region or "place", the boundaries of which are considered to be highly managed and under the control of a governing authority. The strength of a national border is only as robust as its existence being accepted by the people affected by it.

This current paradox of transnationalism arises because of the persistence in people's conception of location and place as fixed entities. However, such a notion of "place" in a globalized society has long been challenged by geographers. Under the geographical interpretation of place, the fixity of boundaries that governments of nation-states desire – "... a site of an authenticity, as singular, fixed and unproblematic in its identity" (Massey, 1994) – is a false illusion, as it relies on the assumption that any space on the Earth's surface is, in fact, unchanging. In truth, any "place" is inherently transnational, the character or identity of which is being created and modified continually by forces within it and by its connections with phenomena or places beyond it; it is forever in a process of formation, never a completed and stable entity. The sovereign national boundaries on which the contemporary definition and utilization of transnationalism depend are, and have always been, themselves unfixed and permeable, evolving in location, nature and degree of porosity as a matter of course. Their geographical fixedness has always been temporary in the reality that their positions have changed over historical time in response to political change (often driven by military events). Further, while national boundaries may have had some impact in actual demographic movement, they have rarely been effective in excluding cross-border interaction based on more psychological or non-corporeal phenomena.

Such discontinuities are a strong argument for a reconfigured notion of transnationalism that is more reflective of a globalized and globalizing world, one which also reconstitutes the concept as a human condition in a way that is as much representative of the past as it is of the present and future. The starting point for this rebirth is the combination of ideas about transnationalism and its association with cultural identity(ies), or more specifically, its connections to more hybrid and multi-faceted substantiations of personal identity.

Transculturalism and contemporary transnationalism

To confine transnationalism to an imperfect perspective of the transactional is to deny its capacity and power to describe and explain patterns of human behavior and action in far more complex ways. The principal constraint on a wider and more reflective conceptualization of transnationalism is this intellectual umbilical cord to 21st Century globalization. In such a context, the rationale for the argument that transnationalism is a transactional process of movement, and not a phenomenon in itself, depends ultimately on the degree to which the notion of the sovereign State – itself only a conception that is no more than two hundred years old – is confirmed or compromised. Consequently, any intellectual shift that can expand the potential of the term to contribute to a wider understanding of human experience, depends ultimately on the negation, or at least a reduction, in the significance attributed to the notion of territorial "borders" or "boundaries".

The notion of transnationalism as being primarily founded on the physicality of geographical movement across national boundaries ignores the holistic nature of human "existence", in which the mind is vital as the corporeal. Instead, we need to both acknowledge and promote the view that transnationalism also incorporates a very distinct psychological, attitudinal belief in the power and the necessity of thinking beyond one's existing social/political/economic/national borders. In essence, this deeper conceptualization of transnationalism involves a subjective, universal consciousness, one that absorbs a disposition or state of mind in favor of international movement and change that is present across all periods of time and space. It is not just a spatial movement of people between locations, but an attitudinal drive and way of thinking that does not see national boundaries as unsolvable restraints.

Such a renewed configuration is not, as some might suggest, better identified as a "post-colonial cosmopolitanism", nor is it totally aligned with Arjun Appadurai's notion of transnationalism as being comprising "post-national" forces that counteract "territorial nationalism" (1996). Further, some might argue that his conception of global cultural flows as comprising a set of individually constructed "imaginary worlds" or scapes, which incorporate the "... way some transnational forces have come to be configured in the imagination of residents ..." (1996, p. 60), resembles the subjective consciousness outlined above. However, Appadurai's constructions still rely ultimately on the proposition that they result from movements of people and identities between places, even if the "de-territorialization of culture" has diminished the significance of the boundaries associated with sovereign or nation-states in an era of more rapid and frequent global mobility. Instead, in its posited reconfigured conception, transnationalism

is its fullest sense is more of a *pre-existing* sensibility that is *fundamentally* dis-associated from place, location and boundaries, one in which the notion of “transculturalism” is an essential aspect.

New formations of transculturalism

A conceptual reconfiguration of transnationalism is best served by a renewed focus on the meaning and significance of its prefix. To interpret the “trans-” in transnationalism as signifying a change of location across national boundaries is simplistic and unnecessarily instrumental. Instead, in the current state of an increasingly diverse world community at both the global and local scales, it is far more apt to view “trans” as embodying a more prominent emphasis on changes or shifts in personal identity or culture as a result of global mobility.

Similar arguments can be found in the work of, for example, Papastergiadis (2000) and Rizvi (2011); indeed, for Papastergiadis, the key feature of transnationalism is its “deterritorialization of culture”, or the

... the ways in which people now feel they belong to various communities despite the fact they do not share a common territory with all the other members. It also refers to the way the national or even the regional culture can no longer be conceived as reflecting a coherent and distinct identity (Papastergiadis, 2000).

Cultures have moved on from “... being formed in particular territorial relationships with carefully established borders, separating one from another ...” (Papastergiadis, 2000), founded on anthropological constructions that are based on a “... notion of culture as a visible set of manifestations of a way of life ...” (Casinader, 2014, p. 31) that is shared by its adherents. Reflections of the wider parameters that a more comprehensive conception of transnationalism embodies can be also found in the contention that transnational migration has placed more priority on the need for decolonization of knowledge.

In concert with Vertovec (2009), Rizvi (2011) and Alexander (2001), Papastergiadis has argued that the increased degree and spatial extent in the movement of people and ideas between territories demanded an image of culture that was not “... a fixed script which actors are bound to follow ...” (2000). Instead, culture needed to be viewed as a more dynamic entity, not located within the “... objective conditions of a society, or purely in the subjective consciousness of the individual, but [more in] the constantly mobile and transformative exchange between the two ...” (p. 109). Culture is a mind-centered, personal entity, in which identity is more dependent on the inner manifestations of attitudes toward values of life than outward indicators (Casinader, 2020).

Transnationalism, then, is more accurately and comprehensively viewed as personifying the dynamic instability of transformative transition that occurs when people and ideas from different parts of the world meet and cross-pollinate dimensions of thought. These perspectives are also reflected in Ong’s reminder that the increasingly rapid processes of cultural interconnectedness and mobility across space are inherently symbolic of transnationalism (Ong, 2004). Embedded in this term, she argues, are the concepts of trans denoting “... moving through space or across lines, as well as the changing nature of something ...”. In addition to suggesting “... new relations between nation-states and capital ...” the term also “... alludes to the transversal, the transactional, the translational, and the transgressive aspects of contemporary behavior and imagination that are incited, enabled, and regulated by the changing logics of states and capitalism.” (Ong, 1999).

One of the more interesting assumptions in the current dominance of the transactional view of transnationalism has been an almost universal conviction that the conceptualization and enactment of thinking skills is culturally neutral. In part, this has derived from, as previously outlined, a conception of culture as an immutable construction, a phenomenon that is seen as being clearly identifiable, despite the fact that the inevitable dynamism of cultures as entities over time and space has been part of the discourse for several decades (for example, Said, 1993; Bhabha, 1994; Alexander, 2001; Rizvi, 2011; Casinader, 2014). Instead of being a fixed representation of the stasis within a society at the moment in time, culture is a mobile and liquid entity that is continually evolving as result of the tensions between societal conditions and individual perceptions. Indeed, Papastergiadis argues that the commanding of multiculturalism as the basis of national identity by some nation-states shows a lack of “... appreciation of the dynamic flux mobilized in all cultural identities ...” (2000, p. 113).

The concomitant of this fixed notion of culture is that there can be no “transcultural” way of thinking or reasoning. If thinking processes are the same across all cultures, or are not influenced by cultural difference, there is only one way to “think”, and ipso facto, there can be no process of thinking that incorporates multiple cultural approaches to thinking because multiple variations do not exist. This conclusion about the relationship between culture and thinking has emerged as one of the consequences of “Western” educational systems and principles dominating global education over the last century. It is one of the key outcomes of contemporary globalization and its geographical spread, which has facilitated the diffusion of such educational philosophies, conceptions and processes at a rapid rate.

However, over the last decade or so, a number of researchers, such as Dahl (2010), Singh (2013) and Casinader (2014), have argued that the notions of reasoning that are embedded in the educational systems of more industrially developed societies have been almost exclusively derived from Euro-American or “Western” perspectives. They argue that different cultural contexts do utilize different approaches to thinking, with Casinader referring to these as “cultural dispositions of thinking” (2014, p. 148), in which different cultural groups view and enact skills such as critical thinking in contrasting ways.

In themselves, the linear and abstract forms of critical thinking that underpin Euro-American systems of thought do not encourage a more inclusive notion of reasoning; that is, one that derives from an acceptance that there are cultural variations in how thinking takes place. It is more a matter of whether or not a priority is placed on the value of contextual thinking and group problem solving, as opposed with a philosophy that prizes individual resolution above all else, which is one of the underpinning common values of “Western” societies. As a result, the concept of a “transcultural disposition of thinking” (Casinader, 2014) accepts the existence and natural validity of all cultural approaches to thinking, and generates the capacity for individuals to operate and function across multiple cultural contexts. The development of such transcultural dispositions is arguably more apparent those who have been more exposed to globalizing experiences (Casinader, 2014, 2020), which enable individuals to acquire the capacity to relate to and work in thinking approaches across cultural identities and spaces. In order to be effective in this transformation, the nature of these globalizing experiences must involve the individual experiencing some form of cultural displacement, or “... the learnings that derive from [being] physically and psychologically ... embedded in a culturally unfamiliar environment for a period of time” (Casinader, 2020, p. 149). Fundamental to these lessons that emerge from the acquisition of attitudes of transcultural capacity is the perspective that cultural variation is endemic in human society; it is uniformity that is more “unreal” than the “problem” or “anomaly” that cultural diversity is seen by some to represent. It is possible, therefore, to conceive of “trans”-nationalism as being more about the reflection of attitudes that transcend (rather than cross) national borders. In such a configuration, which is more representative of the cultural soup within contemporary global society, transnationalism can be viewed as embodying transcultural attitudes that are derived directly from the positive absorption of the cultural flux of globalization and are not bound territorially.

In this new framework, transnationalism becomes a far more sophisticated and long-reaching concept that extends beyond essentialized references to population movement; it is more inclusive of ideas of transferability that are able to incorporate notions of mutually inclusive cultural consciousnesses, which are themselves born out of globalized migrations of people, notions and perceptions. Consequently, as a phenomenon, rather than a process, transnationalism personifies the dynamic instability of transformative transition that occurs when people and ideas from different parts of the world meet, cross-pollinating dimensions of thought in the process. It is a function of the rejection of prior and existing assumptions that have been asserted as to the definition of culture. Whether demarcated in terms of anthropology or a more mind-centered conception based upon consciousness, cultures are not necessarily nationalistic and co-existent with political territorial borders. Instead, a more holistic notion of transnationalism exists, promoting “... that the movement of people and cultures contributes to a *Welbild* (conception of the world) in which borders and boundaries of nation, culture, race, and gender need to be reconceptualized, blurred, challenged, and, potentially, eliminated” (Friedman and Schulermandl, 2011, p. 6). This more encompassing umbrella of transnationalism has, at its heart, a high degree of embedded transcultural understanding, the nature of which is characterized by acknowledgment and acceptance of difference as the norm. Cultural diversity is better characterized as a collective of cultural dispositions of thinking, rather than being approached as a chess set of cultural pieces, existing side by side but in opposition to each other. The crucial feature here is the acceptance of cultural difference as the standard state of human society, one that needs to be confronted and mediated with the aim of co-existence, rather than seeing difference as a boil that needs to be lanced with policies of homogenization and assimilation. More significantly, the reality of globalization demands that this shift in the perspective toward transnationalism must take place, because it is “... [a paradox] of globalization ... that difference is becoming increasingly normative” (Suárez-Orozco and Baolian Qin-Hilliard, 2004).

The education - transnationalism nexus

A reconfiguration of transnationalism as an attitudinal and aspirational phenomenon, as opposed to being just a simple geographical transaction, provides a deeper understanding of its educational associations in both school and higher education. Another of the salient symbols of contemporary globalization has been a substantial growth in the number of international students from the industrially emerging economies attending schools and universities in Euro-American societies and economies, such as those in North America, Australasia and Europe. However, such interactions and the manner in which they have been perceived by relevant institutions in these regions have been overwhelmingly utilitarian, in which transnationalism has been seen primarily in the context of the international commodification of teaching and learning; the economic aspects of the globalizing process have been very much in focus.

Such a limited perception does not mirror the range of transnational priorities reflected in Vertovec’s six-part classification of the phenomenon (2009). In the main, the transactional nature of population movement across national borders has been overlain with the singularity of utilitarian individualism, focusing on either those who migrate to other countries, usually temporarily, for the sole purpose of obtaining qualifications, or on the delivery of education by an institution in a country outside that institution’s home base. The globalization of MBA degree studies to include regions outside the traditional cosmopolitan centers such as North America is an apt example of such an emphasis, illustrated by the Caribbean context, which links regional universities to Australia (Allahar and Sookram, 2018).

In many ways, the current vision of transnationalism as a commercial concept in higher and school education, in which people move temporarily (at least, initially) from one State to another for the purposes of acquiring education for a paid fee, is not unexpected. An interesting reflection of how different educational sectors are perceived by higher level policy makers can be observed in the general application of transnationalist principles to higher education more than school education, no better exemplified by the

definition of transnational education (TNE) adopted by the Council of Europe and UNESCO, in which transnationalism is reduced to simple definitions of student origin:

All types of higher education study programmes, or sets of courses of study, or educational services (including those of distance education) in which the learners are located in a country different from the one where the awarding institution is based. Such programmes may belong to the education system of a State different from the State in which it operates, or may operate independently of any national education system Transnational arrangements should be so elaborated, enforced and monitored as to and education widen the access to higher education studies ...

(Council of Europe, 2014)

Under a more expanded configuration as a disposition of the mind, the transnationalist concept offers a much deeper lens for interpretation of how global events can have an impact on educational demographic movements, both temporary and permanent, educational or otherwise. The wider limitations of such a constricted vision of transnational education was illustrated by one of the major educational impacts of the Coronavirus pandemic that began in 2020. The sudden loss of income that would have come from international students transferring from countries throughout Asia, Africa and South America to study in countries along the Euro-American axis had dramatic economic and intellectual impacts on the education sectors of those countries. For instance, In Australia, and in the face of minimal financial assistance from the Australian government for what were arguably ideological reasons, universities that had high proportions of international students had to retrench staff, contract their operations and reduce the extent and scope of their research activities considerably. This was in contrast to the policies of governments in countries such as the United Kingdom and Canada, which determined that the cost of supporting international students would have longer term benefits in terms of goodwill.

What the Australian Government and some institutions in the higher education sector were unable to fully comprehend was that transnationalism also embodies a desire for self- and life-improvement that transcends the mere utilitarian. International students are examples of people who possess the dispositional attitudes that accompany the willingness to follow educational opportunity, even if it means migrating internationally for a few years. The desire to improve the lives of oneself and that of the wider family overrides any feelings of separation, especially in a digital world where communication and contact are now only a click away.

The damage caused by the pandemic to international students, and the impact of the effective abandonment by the Australian government, was therefore not just confined to the inability to fulfill the transactional requirements for a qualification. It struck far more deeply as it tore into the long term effort to improve prospects for a better life, for which individuals – and often, their families in the place of origin – had worked and sacrificed much. Fundamentally, governmental attitudes such as these totally misunderstood the significance of the international student experience, in that it created the life-changing possibility of “... transforming themselves through mobility.” (Tan and Thao, 2018).

Transnationalism as an activator of rights in education

Contrary to popular imagination, this psychological nexus between education and transnationalism on an international skill is not a new phenomenon. The perspective is an established historical phenomenon, well-illustrated by examples from an earlier phase of globalization that peaked in the mid-twentieth century, which was aligned with the age of European colonialism. Of particular relevance are those that emerged from the British sphere of direct colonial influence, and the notion of education as an emancipator of individuals or groups who wished to actualize their own cultural priorities, rather than being empty vessels that succumbed to the colonizing power as the dominant form.

The principle that formal education is the primary and most effective means toward personal liberation and self-improvement is, arguably, one of the more enduring aspects of Euro-American colonial philosophy, especially in the context of the British Empire. It also provides one exemplar of the main contention of this article, that transnationalism is not only a more personal psychological orientation, but one that has been in existence throughout recorded history and is not just a feature of the contemporary globalized era. In that context, it is an intellectual precept that dates back to the Age of Enlightenment in Europe and John Locke’s assertion that, since all are born equal, the quality of the life that people will experience depends upon how the mind is molded by the quality of education that they receive. It is a principle that itself can be threaded back, in the “Western” context, to the Socratic emphasis on reason as the highest form of human achievement on the one hand, as well as to the prominences placed on the search for truth through the focus of mind and spirit that are inherent in Eastern philosophies, including those in Hinduism and Buddhism, on the other.

In this sense, the capacity of education to elevate individual lives can be presented as a universal truth, even though the form, organization and accessibility of that education has been, and remains, one of the debatable aspects of human existence. Until the nineteenth century, civilizations throughout the millennia have generally seen education as being reserved for those with economic and political power, or for whom there was a perceived need, in the mold of Plato’s philosopher kings. Of more universal relevance, however, is the capacity and readiness of individuals to harness the powers of education when they become accessible, as well as the relationship of education to the generation, existence and impact of transnationalism as a personal attribute, rather than a societal phenomenon. An example of this can be outlined in relation to colonial education during the period of the British Empire in the 19th and 20th centuries.

One of the more fundamental intellectual bases of postcolonial perspectives is that imperial education was a destructive force because that it denied the right of local people to actualize their own traditions of education and knowledge in ways that reflected their own way of life. However, such rationales lack some nuance in their construction and application, for they automatically deny self-agency to those who were colonized and do not allow for forms of pushback by that group in ways that benefited them directly. One example of this was how transnational disposition, in its posited sense, was harnessed by certain minority groups in colonial Sri Lanka during the 19th and 20th centuries and used to position themselves within Ceylonese society.

For ethnic groups such as the Ceylon Tamils and Burghers in British Ceylon, colonial education was the means through which a renewed cultural integrity within the community was able to be born out of the colonial experience, centered on their grasp of English and willingness to use the trappings of colonial administration for their own purposes and benefit (Casinader, 2017). The introduction of mass education helped to open up lines for improvement and lessen the socio-economic gap between the elite and the general population (Hobsbawm, 1989), providing the means by which these groups and individuals could activate their transnational drive for freedom from existing socio-political structures. It was the means by which they could choose to enforce their own rights to emancipation from socio-economic constraint, assisted by a capacity for transnationalism of the spirit and mind, as part of which personal, spiritual, social and economic advancement and security were more important than attachment to any particular location or cultural home.

Although the process and practice was initiated by missionary and religious groups in the early 1800s, the British colonial administration presided over – albeit in an ad hoc manner until the Colebrooke-Cameron Report of 1831 (Colebrooke, 1831) – the development of a mass education system that reached into rural communities, targeted both boys and girls, and saw education being conducted in the local vernacular languages, unlike in India, where English was the nominated language of administration and education. However, the schools that were established as English-language institutions to educate Ceylonese for what was to become the Ceylon Civil Service in the colonial administration were predominantly created, not by the colonial government, but by religious groups such as the churches and missionary societies. Regardless of sector, these schools were characterized by their egalitarian approach, open to able students across all ethnocultural groups, and being either low-fee or free. This pattern had been established largely because of the lack of available funds for education in the early years of the colonial administration. Governors such as Robert Brownrigg were nevertheless keen for education to be established as part of the colonial imperative and supported the ventures of missionary groups from the early 1800s.

As a result, marginalized minority ethno-cultural groups such as the Dutch Burghers and Ceylon Tamils, especially those who were Christian, were able to use the system of educational accessibility and equity within colonial Ceylon as an administrative framework to forge new directions of life and fulfill longstanding aspirations. For many, it freed up their ability to shift status, lifestyle and social position, activated in the process by a previously dormant sense of transnationalism that was later translated into a mind-centered disposition, a change guided by principles of self-improvement and security, and not necessarily an attachment to place, as assumed in the current interpretation of transnationalism (see Casinader, 2017, for further detail).

Overall, the example of transnational thinking and its connections to education from colonial Sri Lanka highlights the need and value for a more enlightened concept of transnationalism in current educational thinking, especially in the context of developing pedagogies and teacher education structures that are more compatible with the increasing cultural complexities in the mindsets of 21st Century students and communities. The growing recognition of such an educational imperative is perhaps best reflected in arguments that suggest that Spivak's notion of transnational literacy (1999) – a knowledge and understanding of how globalization is affecting the nature of education and the national implications of international developments – has now become a global imperative for educational leaders, and not just an optional extra. In one example, Brooks and Normore (2010) argued that a greater understanding of glocalization – the notion that globalization has re-energized the connections of individuals to their local place (Appadurai, 1996) – is becoming indispensable in the preparation and practice of educational leaders in the 21st century. Noting the scarcity of literature on connecting glocalization with educational leadership, they suggest nine literacies of glocalization (which can re-interpreted as a form of transnationalism) that are essential for leaders to promote if students are to be educated for a transnationalist era: political; economic; cultural; moral; pedagogical; information; organizational; spiritual and religious; and temporal. All of these are expressed within a framework of a definition of literacy that incorporates new information and communication technologies, critical thinking, active citizenship and linguistic and cultural diversity. Since educational leaders are in positions to provide direction and exercise influence, the educational resources that they make available to their students and school systems need to be more reflective of the realities of their students' futures, countering current deficiencies in student literacies across these areas. As we move toward the middle of the 21st century, it is not enough to educate people to live in a world where transnationalism merely involves movement across borders as a common occurrence. It is now necessary to acknowledge that transnationalism incorporates attitudes of self-improvement and awareness that are essential and integral parts of global mobility.

Acknowledgments

This article is based on material first published in Casinader N. (2017). *Transnationalism, Education and Empowerment: the latent legacies of empire*. Routledge. Chapters 1 and 2. Used by permission.

References

- Alexander, R., 2001. *Culture and Pedagogy: International Comparisons in Primary Education*. Blackwell Publishers, Oxford.
- Allahar, H., Sookram, R., 2018. Globalization of MBA studies and transnational educational institutions in the Caribbean. *Ind. High. Educ.* 32, 290–301.
- Appadurai, A., 1996. *Modernity at Large: Cultural Dimensions of Globalization*. University of Minnesota Press, Minneapolis.
- Bhabha, H., 1994. *The Location of Culture*. Routledge, London.
- Brooks, J.S., Normore, A.H., 2010. Educational leadership and globalization: literacy for a global perspective. *Educ. Pol.* 24, 52–82.
- Burbules, N.C., Torres, C.A., 2000. Globalization and education: an introduction. In: Burbules, N.C., Torres, C.A. (Eds.), *Globalization and Education: Critical Perspectives*. Routledge, London.
- Casinader, N., 2014. *Culture, transnational education and thinking: case studies in global schooling*. Routledge, Abingdon.
- Casinader, N., 2017. *Transnationalism, education and empowerment: the latent legacies of empire*. Routledge, Abingdon.
- Casinader, N., 2020. *Transnationalism and Teacher Capacity: Professional Readiness in the Globalised Age*. Routledge, Abingdon.
- Colebrooke, W.M.G., 1831. Commission of Eastern Enquiry: Report Upon the Administration of the Government of Ceylon. National Archives, United Kingdom, Series CO54, Folio 122.
- Council of Europe, 2014. Code of Good Practice in the Provision of Transnational Education [Online]. Available at: http://www.coe.int/t/dg4/highereducation/recognition/Code%20of%20good%20practice_EN.asp. (Accessed 19 April 2016).
- Dahl, M., 2010. *Failure to Thrive in Constructivism: A Cross-Cultural Malady*. Sense, Rotterdam.
- Friedman, M., Schultermund, S., 2011. Introduction. In: Friedman, M., Schultermund, S. (Eds.), *Growing up Transnational: Identity and Kinship in a Global Era*. University of Toronto Press, Toronto.
- Hobsbawm, E.J., 1989. *The Age of Empire 1875–1914*. Vintage Books, New York.
- Kirsch, C., Gogonas, N., 2018. Transnational experiences, language competences and worldviews: contrasting language policies in two recently migrated Greek families in Luxembourg. *Multilingua* 37, 153–175.
- Massey, D., 1994. *Space, Place, and Gender*. University of Minnesota Press, Minneapolis.
- McGrew, A., 2007. Organized violence in the making (and remaking) of globalization. In: Held, D., McGrew, A. (Eds.), *Globalization Theory: Approaches and Controversies*. Polity Press, Cambridge.
- Ong, A., 1999. *Flexible Citizenship: The Cultural Logics of Transnationality*. Duke University Press, Durham, Nc.
- Ong, A., 2004. Anthropological concepts for the study of nationalism. In: Nyiri, P., Breidenbach, J. (Eds.), *China inside Out : Contemporary Chinese Nationalism and Transnationalism*. Central European University Press, New York.
- Papastergiadis, N., 2000. *The Turbulence of Migration: Globalization, Deterritorialization and Hybridity*. Polity Press, Cambridge.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Rizvi, F., 2009. Towards cosmopolitan learning. *Discourse* 30, 253–268.
- Rizvi, F., 2011. Beyond the social imaginary of 'Clash of Civilizations'? *Educ. Philos. Theory* 43, 225–235.
- Robert, S.A., Yu, M., 2018. Intersectionality in transnational education policy research. *Rev. Res. Educ.* 42, 93–121.
- Said, E.W., 1993. *Culture and Imperialism*, First Vintage Books Edition. Vintage Books, New York.
- Singh, M., 2013. Educational practice in India and its foundations in Indian heritage: a synthesis of the East and West? *Comp. Educ.* 49, 88–106.
- Spivak, G.C., 1999. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Harvard University Press, Cambridge, Mass; London.
- Suárez-Orozco, M.M., Baolian Qin-Hilliard, D., 2004. Globalization: culture and education in the new millennium. In: Suárez-Orozco, M.M., Baolian Qin-Hilliard, D. (Eds.), *Globalization: Culture and Education in the New Millennium*. University of California Press in association with the Ross Institute, Berkeley and Los Angeles.
- Tran, L.T., Thao, T.P.V., 2018. 'Agency in mobility': towards a conceptualisation of international student agency in transnational mobility. *Educ. Rev.* 70, 167–187.
- Vertovec, S., 2009. *Transnationalism*. Routledge, Milton Park, Abingdon.

Development and education

Simon McGrath, School of Education, University of Glasgow, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	52
Education and economic development	53
Education and human development	54
Education and sustainable development	55
Education and human rights	56
Radical critiques of education and development	57
Conclusion	57
References	58

Introduction

The relationship between development and education is widely seen in policy circles and in popular understanding as education for development. That is to say, education is the servant of development. Indeed, most of the time the core understanding is of schooling for development. However, the intention here is to see both development and education as inherently contestable concepts both separately and when viewed in relationship with each other.

Although educational expenditure has long been an important part of development assistance, both official and non-official, and education featured in the world conferences of the 1990s and the resultant Millennium Development Goals, education has occupied a marginal role in development studies departments and journals.

As debate raged around 2012–2014 as to what should be included within the proposed new Sustainable Development Goals, it was seen as quite plausible that education would not “get its own SDG” for contradictory reasons. For some prominent economists, education issues were so easily solvable as to be unworthy of a focus in the SDGs. Sachs (2008: 301–2), for instance, declared that “Of all the MDGs, universal access to basic education is surely the easiest to achieve. The technology is the best understood and the most straightforward.” (cf. McGrath, 2010 and 2014). Burnett (2012), on the other hand as an economist with a far more detailed understanding of education, was concerned that it was the intractability of achieving desired education outcomes that would make it seem a poor focus of a target for the SDGs, as policymakers would inevitably look for easier wins. Yet, education did get included under SDG4, “ensure inclusive and equitable quality education and promote lifelong learning opportunities for all” (UN, 2015: 17). Moreover, educationalists stressed that education was integral to the wider SDG program, mapping the need for educational input to the achievement of all goals (UNESCO GEMR, 2016; McGrath, 2018; Tikly, 2020).

I have hinted at the power of economists in driving development thinking, and the orthodox position of development policy and studies is still that education is an investment in individuals that then has firm and economy level pay-offs. However, there are many other accounts of the education-development relationship. In a piece of this length, a selection of these can only be dealt with. Which then to choose? In a widely cited paper, the philosopher, Ingrid Robeyns, focuses on three main accounts: “rights, capabilities and human capital” (Robeyns, 2006). I follow her lead in my book-length exploration of the education-development relationship (McGrath, 2018), although noting a series of other positions there too. Economic development, human development and human rights, therefore, will be the foci of **Education and economic development**, **Education and human development** and **Education and human rights** sections of this article respectively. In returning to this issue some three years later, my relegation of sustainable development to the supporting cast seems unjustifiable in the light of the climate emergency. It is clear that the long-standing education for sustainable development (ESD) tradition needs more mainstream attention in light of the climate emergency, and this will form the focus of **Education and sustainable development** section. In my book, I note also more critical perspectives on the negative sides of both development and education, and I will summarize these debates in **Radical critiques of education and development** section.

Of course, other themes could be highlighted. For instance, SDG target 4.7 brings together environmental concerns with a range of others. To take one example, education for peace is enshrined at the start of the Constitution of UN’s specialized agency for education, UNESCO: “since wars begin in the minds of men [sic], it is in the minds of men that peace must be constructed” (http://portal.unesco.org/en/ev.php-URL_ID=15244&URL_DO=DO_TOPIC&URL_SECTION=201.html).

For space reasons, I am going to concentrate on contemporary debates. However, it is also important to remember that education has had a long and complex history of relationships with small “d” development long before big “D” development emerged in the aftermath of the Second World War (see McGrath, 2018, for a review of this history). And a final caveat: while I present these five accounts separately, reflecting their origins in distinct traditions, policy and theory reality is far messier, with considerable blending of approaches, a point I will reflect on in the conclusion.

Education and economic development

In the rise of both high and low-cost private schooling, in increasing expenditure on supplementary tuition for children, in the introduction of loans for higher education, North and South, there is a clear understanding that education is a good investment. It is a good investment for individuals and families, to be reflected in higher incomes, and for entrepreneurs, who can make good profits from private schooling. Even for the state, education can still be a good investment, particularly where it is acting against market failures (most notably, under investment by parents in girls' education).

For nearly two hundred years after Adam Smith's account of economic development, a notion of intangible human capital lay under the surface of economic theory. Only in the 1960s did this coalesce into a theory (Schultz, 1961; Becker, 1964). This human capital theory (HCT) argued that forms of human capital, most obviously schooling, contribute to economic productivity. Emerging at the time of rapid decolonization, and the heightened attention this gave to development assistance, it was quickly grasped by planners, North and South, that education could play a crucial role in development by enhancing national productive capacities. By 1980, this convergence of development theory and aid practice led to the World Bank replacing UNESCO as the leading global actor in education and development. Indeed, several of the key texts about education and economic development were written by Bank staff (McGrath, 2018).

The key concept that the Bank helped move into the education policy mainstream was rate of return analysis. If education is understood as an investment in the development of capital, then the decision to invest in education should be subject to the same calculation as other investments, that is, a calculation of the expected rate of financial return on the sums invested. For an investment in education to be justified, at either household or national level, its expected rate of return on education (RORE) should be higher than the alternative rate of return from other possible investments.

We have had nearly 50 years of rate of return to education analyses, generating thousands of publications. These tend to show relatively high rates of return on education as compared to other possible investments, and have been key to continued national and international investment in education, and particularly in the education of girls.

The analysis was quickly developed to suggest that all education did not matter equally. In developing a RORE methodology, Mincer (1974) had assumed constant returns to years of education. However, the newly-formed World Bank economics of education team soon focused on which level(s) of education provided the best returns (Heyneman, 2003; Psacharopoulos, 2006). Three World Bank research projects formed the core of an argument that retains its intellectual and policy influence 40 years later. First, Cochrane (1979) showed that an investment in primary schooling for girls had positive effects on fertility and infant mortality. Second, Lockheed et al. (1980) showed that even a few years of primary schooling brought benefits for farmer productivity. Third, and most influentially, Psacharopoulos (1981) showed that primary schooling had the best RORE of all forms of education. While these accounts have been subjected to widespread methodological, theoretical and philosophical critique, their core arguments remain hugely powerful and more recent work in the tradition is far more methodologically sophisticated (see McGrath, 2018, for a summary of the criticisms and of key developments in the approach).

A major concern has been that the use of years of schooling as a central variable is not robust enough given what we know from other research about variations in time on task and learning outcomes, both nationally and internationally. In response, Hanushek and Woessmann have pioneered attempts to measure educational quality rather than quantity for RORE. Using data from the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) as a superior proxy for human capital, they argued that this provides triple the explanatory power of the years of schooling approach (Hanushek and Woessmann, 2007). Indeed, they have subsequently argued that years of schooling are statistically insignificant once they are combined with measures of quality (Hanushek and Woessmann, 2012). Thus, we have seen a shift in policy focus from quantity to quality of schooling.

Although the argument that investments in primary education (especially for girls) are the most justified continues to be powerful, the past 40 years has seen research that also makes a case for other sectors of education being the best investment, such as early childhood development (Carneiro and Heckman, 2003) and higher education (McCowan and Schendel, 2015).

As noted briefly above, HCT has received considerable criticism from a number of angles. Many neoclassical economists of education argue that the social RORE is actually far greater than the private RORE at the heart of the Mincerian approach (McMahon, 1999). This has spawned the "wider benefits of learning" literature (Feinstein et al., 2008). Heterodox economists stress the role of education as part of more complex interactions between firms, state and market that drive innovation in economic development (Malerba, 1999; Lall, 1992; Von Tunzelmann, 2010; Kruss et al., 2015). Other economists, most famously Bennell (e.g., 1996), have pointed to very serious problems in using HCT in economies with high levels of public, informal and subsistence work, as is prevalent across the South. Similar weaknesses of course lie in the systematic underestimation of the economic value of care work, predominantly done by women (Donath, 2000). There have been criticisms too of the move toward measuring quality through PISA in particular, which is seen to be highly flawed (Meyer and Benavot, 2013). For Marxist economists, such as Bowles and Gintis (1976), the approach is flawed in its assumption that there is no conflict between capital and labor. Sociologists have continued this critique, and have focused on what they consider to be the false assumptions of HCT that mean that it is a highly partial account of how economic development takes place and what it means for individuals and society (Lauder, 2015).

While the education for economic development has well-defined tools and clear practice and policy implications, such critiques point to concerns about whether it has only the semblance of authority. Put simply, more than a half-century of this orthodoxy has not resulted in well-functioning education systems or economies around the globe. In many countries, and particularly for those most marginalized, the prospect of a smooth transition from quality education to decent work is a mirage.

Education and human development

Although the 1960s saw the emergence of the human capital orthodoxy, the decade also saw the emergence of a counter-narrative as it quickly became apparent that the expansion of schooling was not leading to the presumed economic take-off. This basic needs account (Streeten, 1979) laid the foundation of an approach within the United Nations Development Program that resulted in the Human Development Index and Report (McGrath, 2018). The driving force behind these two outputs, Ul Haq (1995), argued that the objective of development should be to create an enabling environment for people to enjoy long, healthy and creative lives. This work was developed in a new direction by his close collaborator, Amartya Sen, and by Martha Nussbaum, into the human development and capabilities approach (HDCA). Sen argues that economic development is simply a means to the ultimate goal of human flourishing. Thus, "Development can be seen ... as a process of expanding the real freedoms that people enjoy." (Sen, 1999: 3).

Sen places the building of individual capabilities at the heart of this. For him, capabilities mean "what a person is able to do or be ... [and represents] the opportunity to achieve valuable combinations of human functionings" (Sen, 2005: 153). Sen is concerned to stress that these vary between individuals and that we should respect this. Nussbaum (2000), in contrast, seeks to offer a set of capabilities that might be universally valuable that could then form the basis for legal and policy enactment.

Education is seen by both Sen and Nussbaum as crucial to the larger human development project. Underpinning the notion of capability for both is the ability of an individual to identify their "valued beings and doings" through reasoning. Education is explicitly seen as building this reasoning.

Yet, some later education and HDCA scholars, most notably Unterhalter (2005), have made it clear that education may actually diminish certain capabilities, such as where traveling to or attending school is unsafe, especially for girls. More positively, the HDCA approach to education allows two new questions to be asked about education. First, what are the multiple capabilities that learners value? Second, how are educational institutions successful, or not, in supporting capability achievement?

The first restates Sen's objection to a simple assertion that education is an investment with economic outcomes. Rather, it encourages us to consider why an individual has enrolled and persisted in education. Among the myriad reasons might include to be a better parent, to contribute more to one's community, or to be better able to take part in religious activities (Powell and McGrath, 2019).

The second follows on from this to develop a critique of evaluation of educational institutions' performance that emerges from new public management and HCT and which focuses on pass rates and employment destinations. Rather, the emphasis in HDCA-inspired evaluation is on measuring capability promotion (Walker and McLean, 2013; Powell and McGrath, 2014).

Like HCT, the approach was originally applied to schooling but has come to be used increasingly in other areas of education, such as higher education (Walker and McLean, 2013; Walker and Fongwa, 2016) and vocational education and training (VET) (Powell and McGrath, 2019; McGrath et al., 2020).

The HDCA approach to the development-education relationship has been critiqued in a number of ways. From within the tradition, there has been a strand of writing that has sought to make up for the perceived limitations of an approach that emanates from liberal economics and philosophy. The status of individualism in the approach is much debated (Robeyns, 2017) with some arguing for inclusion of a notion of collective capabilities (Evans, 2002; Ibrahim, 2006).

Recent work in education seeking to take a more post-colonial stance has stressed the need to think more relationally about capabilities (De Jaeghere, 2019; Tikly, 2020). Others within the HDCA tradition have also pushed for a stronger sense of the way that structure impacts on capability formation (Deneulin, 2011; Stewart, 2013). In her review of the nearly 40 years of capability writing, Robeyns (2017: 66) argues, "having an account of structural constraints is therefore non-optional: every capability theory has one; although sometimes this account will be very implicit." Within the education and development community, VET researchers have gone furthest in seeking to address this by recourse to combination of HDCA with sociological theory and political economy accounts (McGrath et al., 2020). There has also been criticism that the approach tends to privilege the interests of current generations at the expense of later ones, although here again more recent work has done much to address this and the links between human and sustainable development (Watene, 2013; Schultz et al., 2013).

From outside HDCA, there has been considerable sociological critique (Dean, 2009; Sayer, 2012) that focuses on HDCA's lack of an adequate account of institutions, society or power, although there are attempts to address this in some of the internal critique noted above. The approach is also questioned for its methodological limitations. These critiques focus largely on the works of Sen and Nussbaum, which are philosophical rather than empirical. The capability approach has been applied in methodologically pluralistic ways that make it ill-suited for the development of highly refined methodological tools as used in HCT. However, this pluralism also leads to a further challenge that HDCA work is not an orthodoxy in any discipline, and so subject to methodological critique from a wide range of directions.

The approach is also widely criticized for its limited policy value (cf. Robeyns, 2017). Clearly, Nussbaum's list is an attempt to talk to policy. However, it is not clear that the academic mainstream of the approach has strong enough theories of power and change to get much purchase at national or global policy levels. It is contended that an approach that stresses the position that we need to recognize individual's complex and multiple agendas rather than a simple set of measurable targets is very far from what policymakers and planners find useful. In the case of education, it is difficult for this liberal position to compete with the instrumental argument that we should prioritize those parts of education that provide the greatest economic returns and those elements that do the most to reduce income poverty.

Where HDCA does appear to have most policy purchase is through the Human Development Reports and Human Development Index at the global level, and their use to generate national reports in a number of states. However, there is a tension here within the tradition regarding whether HDCA is even a thing (Robeyns, 2017). Many academic philosophers, such as Robeyns, see HD and CA as separate approaches, though having common roots. In this view, policy work on human development is not work within the capabilities tradition. Others, most notably development economists, are much more comfortable with the notion of HDCA (Alkire and Denuelin, 2009) and so see the UNDP's work on human development as within the wider family.

Such work has been at the macro scale but there is also policy potential at the meso level. For instance, there does seem some potential that the evaluative perspective offered by HDCA can help educational institutional leaders to refocus their practices on expanding learner freedoms (Walker and McLean, 2013; Powell and McGrath, 2014, 2019).

Education and sustainable development

Just as the origins of the human development counter-narrative can be found in concerns about the problems of economic development in the South in the 1960s, so the origins of sustainable development can be seen in concerns about the costs of apparently successful economic development in the North at the same time (Carson, 1962; Boulding, 1966; Meadows et al., 1972). These concerns entered the global policy stage in 1972 through the Stockholm Conference on the Human Environment, quickly followed by the 1974 Cocoyac Conference. This called for the harmonization of development and environment concerns through "eco-development". The 1980 World Conservation Strategy saw the first international policy use of the term "sustainable development", popularized in 1987 by the Brundtland Commission report, *Our Common Future* (World Commission on Environment and Development, 1988).

The dominant position of these international reports is that environmental sustainability and economic development can be successfully harmonized, and this remains the global policy orthodoxy, as represented by the SDGs. However, many researchers and activists have been skeptical about this optimistic picture from the outset and such views have grown hugely as the extent of the environmental crisis has become clearer (Rockström et al., 2009; Jackson, 2017; Raworth, 2017). Belief in a tension between capitalism and the environment is put starkly by the journalist, Naomi Klein, "our economic system and our planetary system are now at war" (2014: 21). Attention has turned, therefore, to calls for transformative change, such as degrowth (D'Alisa et al., 2014) or just transitions (Swilling, 2020).

This tension between an ameliorative and a transformative vision of sustainable development plays out in accounts of education for sustainable development (ESD) (Vare and Scott, 2007). Vare and Scott characterize two broad trends within ESD. They see ESD1 as promoting changed attitudes, behaviors and skills, what they term "learning for sustainable development" (2007: 193). They see this as an expert-led and official approach to ESD, as reflected in UNESCO's UN Decade of Education for Sustainable Development, and this can be extended to much of the work around SDG 4.7. In contrast, ESD2 seeks to build a range of knowledge, skills, attitudes and values to support sustainable futures that need to be struggled for. It sees capitalism as being at the root of the environmental problem and calls for a radical transformation. It is critical too of formal education as being part of a wider system of extraction and exploitation.

Crucially, as Sterling (2019) notes, the patterns of learning embedded in conventional education have contributed to behaviors that exacerbate and accelerate the climate crisis, what Wals (2020: 825) calls a "hidden curriculum of unsustainability". Therefore, proponents of ESD2 call instead for transgressive learning and democratic pedagogies. Wals (2020) identifies three dimensions: criticality, emancipation and relationality. Criticality requires allowing learners the space to ask disruptive questions about why we are where we are today in terms of the climate crisis and how we can move toward sustainable futures, including what are the barriers to this. Emancipation requires empowering learners to find their own paths toward sustainability rather than directing them to one "correct" set of responses. Relationality is about supporting learners to connect to the environment at local levels while seeing the global connections, to other species in respectful ways, and to other people, especially across different backgrounds. What much of ESD of either mode 1 or 2 has is a place-based element that sees the centrality of the adage of "acting locally and thinking globally".

There is considerable attention in ESD to addressing key elements of formal educational systems: pedagogy, curriculum, institutional policies, etc. This has led to major initiatives to "green" schools, vocational colleges and universities (Sterling et al., 2013; Majumdar, 2011). However, there is also a very strong strand of ESD thinking that focuses on non-formal and informal learning sites as offering greater potential for transformative practices (Chaves et al., 2017; Lotz-Sisitka et al., 2017).

Tikly (2020: 57) offers a synthesis of post-colonial, sustainable development and human development perspectives on ESD, defining it as, "socially and environmentally just education that facilitates the capabilities of existing and future generations and of natural systems to flourish". He insists that an environmental education is not enough, but that ESD must be about environmental, economic and social justice.

In contrast with the other accounts of education and development above, ESD has generated a huge wealth of well-theorized critical reflection on educational practice. However, even ESD1 type approaches have struggled for purchase in formal education systems characterized by curriculum overload, resourcing constraints and performative pressures. The more radical approach that Vare and Scott characterize as ESD2 is necessarily threatening to powerful interests and dominant discourses of what matters in education and how it should be organized. The approach is far better at pointing to the need to change than offering ways of

overcoming resistance to change or providing a clear route map to an inevitably complex future. Nonetheless, its identification of many problems and priorities cannot be ignored.

Education and human rights

The three above accounts all explicitly talk about development, even if their visions of it radically differ. However, interacting with all of them is a powerful alternative account that does not seek to adjectivize development at all: the human rights account. Philosophical accounts of human rights are long-standing and well-known (e.g., [Paine, 1984\(1791\)](#); [Wollstonecraft, 2004\(1792\)](#)). However, it is with the 1948 Universal Declaration on Human Rights (UDHR) ([UN, 1948](#)) that the modern rights movement, centered on international legal conventions and in the work of NGOs and activists on the ground, begins.

There are three “generations” of rights: civil and political rights; economic, social and cultural rights, both of which are enshrined in the UDHR; and a more recent set of rights that focus on group and collective rights to self-determination, development and environmental sustainability ([de Schutter, 2010](#)).

Education is clearly an economic, social and cultural right and is reflected both in the majority of the core UN human rights treaties and in all regional human rights systems. The right to education is expressed in Article 26 of the UDHR. This states:

- (1) Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit.
- (2) Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. It shall promote understanding, tolerance and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace.
- (3) Parents have a prior right to choose the kind of education that shall be given to their children.

Since the UDHR, the right to education has been widely recognized and developed by a number of key rights documents from the United Nations. The two most important of these are the International Covenant on Economic, Social and Cultural Rights (ICESCR), 1966 and the Convention on the Rights of the Child (CRC), 1989. The right to education is also addressed in a series of documents from United Nations specialized agencies, most notably UNESCO and the International Labor Organization. This normative work goes significantly beyond the CRC in stressing a lifelong perspective on the right to education, but has less legal force.

To say that education is a human right means several things. It establishes that it is legally-enforceable and a right inherent to everyone without any discrimination. States are obliged to act on this and can be held accountable for any violations, at least in theory.

Although the language of rights is universal, it is evident that the ability to enjoy them is clearly unequal. This leads to a particular focus on the right to education of the most marginalized, such as cultural, ethnic and linguistic minorities, people living with disabilities, nomads, and migrants, refugees and internally displaced persons [UNESCO \(2008a,b\)](#).

The most comprehensive framework for thinking about the right to education is the “4A” model developed by the first UN Special Rapporteur on the Right to Education, Katarina [Tomaševski \(2001\)](#):

Availability—Education is free and there is adequate infrastructure and trained teachers able to support the delivery of education.

Accessibility—The education system is non-discriminatory and accessible to all, and positive steps are taken to include the most marginalized.

Acceptability—The content of education is relevant, non-discriminatory and culturally appropriate, and of quality; schools are safe, and teachers are professional.

Adaptability—Education evolves with the changing needs of society and challenges inequalities, such as gender discrimination; education adapts to suit locally specific needs and contexts.

The position of the rights perspective in education and development is disputed. [McGrath \(2018\)](#) argues that what made the international development commitments to education so powerful from the 1990 World Conference on Education for All to the SDGs is the conjunction of human capital and human rights arguments,

That primary education could be shown to be both a right and an investment created a previously unprecedented alliance between the lead UN specialized agencies, UNESCO and UNICEF, and the World Bank, as well as national governments, bilateral agencies and national and international NGOs.

[McGrath \(2018: 116\)](#).

This coming together of two powerful agendas was instrumental in the very significant progress on accessing schooling that has occurred since 1990. Right to education activists have pushed hard and successfully for the achievement of the joint commitments of the rights system and the international development targets. International civil society has been mobilized behind slogans such as

“no child left behind”. The increased policy focus on the right to education has encouraged the huge growth in research in this area. These are very real achievements.

Nonetheless, there are arguments that the rights agenda has been compromised by its accommodation with aid agendas and the education for economic development account. What has resulted is such an attenuated achievement of the right to education for many millions that has undermined that right and the ability to claim that it was not being realized (McGrath, 2015). At the same time, the combination of world conferences and international development goals has disempowering actors at the local and national levels who could be advocates for educational rights (Unterhalter, 2014; McGrath, 2015; McGrath, 2018).

From a lifelong perspective, there is a danger also that the right to education approach, with its emphasis on primary schooling as first priority, taken together with the problematic claims of the early World Bank economists of education, encourages a narrow emphasis on this single part of education systems (McGrath, 2018). Robeyns (2006) argues that there is a danger that rights provide a minimum floor for governments that allows them to underinvest in building real quality education. By using the notion of progressive realization, that states only have to meet the full set of rights as they can afford them, international human rights law can make it hard to determine whether a country is doing all it can, and what responsibility lies with the donors. What is clear from research on international education finance, such as presented in the Global (Education) Monitoring Report series is that funding is inadequate in spite of international promises. This appears to be one of the key areas of weakness of the right-based approach.

Finally, Robeyns argues that there is a major problem with thinking about quality from a legal rights perspective, and this is compounded by attempts at measurement that remain too focused on the tangible and more easily measurable inputs and outputs of education, and have a limited and overly formalistic sense of educational processes.

Radical critiques of education and development

While rights arguments are often combined with development accounts, of all three kinds outlined above, there are positions that offer far more radical critiques of both development and education.

In the case of development, Hettne (1995) highlights a series of alternative imaginings of development, which he collectively termed “Another Development”. He includes basic needs in this category, as well as self-reliance; eco-development; women and development; and ethno-development. In the subsequent quarter-century, we have seen further elaboration of the former strand around “indigenous” notions as *buen vivir* in Latin America, *ubuntu/ukama* in Africa and radical *swaraj* in India (e.g., Gudynas, 2011; Kothari, 2014; Swilling, 2020). Taken together, these accounts are more radically relational than the Northern development discourses, a relationality that extends to other persons and to nature. Each insists “on the social as well as the economic; the growth of the person over the growth of economy; and the location of that person as a member of a wider community with responsibilities as well as rights” (McGrath, 2018: 51).

Others have written from what have been called variously anti and post-development positions (Ferguson, 1990; Sachs, 1992; Escobar, 1995). Much of this work has been influenced by post-structuralism. This has led to a critique of the discourse and practice of development as constituting a new neocolonial phase of Western domination (Ziai, 2015). Ziai detects two broad trends within this literature, which he terms skeptical and neo-populist, although he suggests that much writing in the field is a confused mixture of the two. He argues that the skeptical strand contains a critical, dynamic, constructivist concept of culture that seeks to engage with development practice but in reflexive ways. However, he contends that neo-populism tends toward a romantic rejection of modernity and celebration of peasant societies that essentializes and ossifies culture. This may be contrasted with the neo-populist’s superior insights. As Ziai argues, this tends to a reactionary and, perversely, elitist dismissal of people’s desire for “development” as false consciousness, making post-developmentalists look very much like the “development experts’ they criticize so sharply” (2015: 837).

All of the accounts above take it as a given that education is a good thing, although I cited Unterhalter as noting that “actually existing education” can lead to capability diminution. Here, however, I want to briefly note some of the arguments that exist about education as not simply a passive bystander in violence but as being itself implicated in it. This in some way has echoes in Wals’ critique above of unsustainable education.

For critics in this area, education is understood a set of systems, institutions and actors that generates multiple forms of violence. Education is seen as playing a role in increasing levels of ethnic/racial, sexual and class-based violence (Bush and Salterelli, 2000; Dunne et al., 2006; Paulson and Rappleye, 2007). It also commits symbolic violence (Bourdieu, 1984) or epistemic injustice (Fricker, 2007) through valuing certain cultures, languages and knowledges over others, and more conventional violence that results for those societies and individuals marginalized and disempowered in its wake (Smith, 1999; Santos, 2014). These accounts are largely framed as critique but where there is a call to positive action, it lies in educational moves that are similar to those being advocated in ESD2.

Conclusion

Although there are fundamentalists who can only see merits in one of these five accounts, it is evident to me that each has strengths, and weaknesses. Education can bring economic benefits but is a fundamental human right regardless of this. Economic development and preparation for the work are important but need to be located within a wider framework of human flourishing. However,

it is increasingly evident that economic development threatens the future of our environment, of nature and of us. Indeed, development has come with very particular costs for many, linked to colonialism, extraction and the destruction of cultures. Education has been deeply complicit in this.

That the education and development relationship should be seen as both pluralistic and two-edged seems obvious to me, but it also feels that it need to be regularly repeated in the face of much education and development research and policy that acts as if there is only one possible understanding of education and development, and that this is entirely unproblematic. However, academic critique of this imperial project is not enough. Rather, what is needed from academics is that they become allies of those who are actively resisting a narrow education for development account and actively building educational alternatives that are pluralistic, democratic and relational, and which seek to build opportunities for flourishing for people and planet now and into the future.

References

- Alkire, S., Deneulin, S., 2009. Introducing the human development and capability approach. In: Deneulin, S., Shahani, L. (Eds.), *An Introduction to the Human Development and Capability Approach*. Earthscan, London.
- Becker, G., 1964. *Human Capital*. Columbia University Press, New York.
- Bennell, P., 1996. Using and abusing rates of return. *Int. J. Educ. Dev.* 16 (3), 235–248.
- Boulding, K., 1966. *The Economics of the Coming Spaceship Earth*. <http://dieoff.org/page160.htm>.
- Bourdieu, P., 1984. *Distinction*. Routledge, London.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America*. Basic Books, New York.
- Burnett, N., 2012. Are We Becoming Irrelevant? Results for Development Institute, Washington.
- Bush, K., Saltarelli, D., 2000. *The Two Faces of Education in Ethnic Conflict: Towards a Peacebuilding Education for Children*. Innocenti Research Centre, UNICEF, Florence.
- Carneiro, P., Heckman, J., 2003. *Human Capital Policy*. Working Paper 9495. National Bureau of Economic Research, Cambridge, MA.
- Carson, R., 1962. *Silent Spring*. Houghton Mifflin, Boston.
- Chaves, M., Macintyre, T., Verschoor, G., Wals, A., 2017. Towards transgressive learning through ontological politics. *Sustainability* 9, 21.
- Cochrane, S., 1979. *Fertility and Education*. World Bank, Washington.
- D'Alisa, G., Demaria, F., Kallis, G. (Eds.), 2014. *Degrowth*. Routledge, Abingdon.
- De Jaeghere, J., 2019. A capability pedagogy for excluded youth. *Educ. Citizen Soc. Justice*. <https://doi.org/10.1177/1746197919886859>.
- De Schutter, O., 2010. *International Human Rights Law*. Cambridge University Press, Cambridge.
- Dean, H., 2009. Critiquing capabilities. *Crit. Soc. Pol.* 29 (2), 261–278.
- Deneulin, S., 2011. Development and the limits of Amartya Sen's the idea of justice. *Third World Q.* 32 (4), 787–797.
- Donath, S., 2000. The other economy. *Fem. Econ.* 6 (1), 115–123.
- Dunne, D., Humphreys, S., Leach, F., 2006. Gender violence in schools in the developing world. *Gend. Educ.* 18 (1), 75–98.
- Escobar, A., 1995. *Encountering Development*. Princeton University Press, Princeton.
- Evans, P., 2002. Collective capabilities, culture, and Amartya Sen's development as freedom. *Stud. Comp. Int. Dev.* 37, 54–60.
- Feinstein, L., Budge, D., Vorhaus, J., Duckworth, K., 2008. *The Social and Personal Benefits of Learning*. Institute of Education, University of London.
- Ferguson, J., 1990. *The Anti-politics Machine*. University of Minnesota Press, London.
- Fricker, M., 2007. *Epistemic Injustice*. Oxford University Press, Oxford.
- Gudynas, E., 2011. Buen vivir. *Development* 54 (4), 441–447.
- Hanushek, E., Woessmann, L., 2007. The Role of Education Quality for Economic Growth. Policy Research Working Paper 4122. World Bank, Washington.
- Hanushek, E., Woessmann, L., 2012. Do better schools lead to more growth? *J. Econ. Growth* 17 (4), 267–321.
- Hettne, B., 1995. *Development Theory and the Three Worlds*, second ed. Longman, Harlow.
- Heyneman, S., 2003. The history and problems in the making of education policy at the World Bank, 1960–2000. *Int. J. Educ. Dev.* 23, 315–337.
- Ibrahim, S., 2006. From individual to collective capabilities. *J. Hum. Dev.* 7 (3), 397–416.
- Jackson, T., 2017. *Prosperity Without Growth*, second ed. Routledge, Abingdon.
- Klein, N., 2014. *This Changes Everything*. Simon and Schuster, London.
- Kothari, A., 2014. Radical ecological democracy: a path forward for India and beyond. *Development* 57 (1), 36–45.
- Kruss, G., McGrath, S., Petersen, I.H., Gastrow, M., 2015. Higher education and economic development. *Int. J. Educ. Dev.* 43, 22–31.
- Lall, S., 1992. Technological capabilities and industrialization. *World Dev.* 20 (2), 165–186.
- Lauder, H., 2015. Human capital theory, the power of transnational companies and a political response in relation to education and economic development. *Compare* 45 (3), 490–493.
- Lockheed, M., Jamison, D., Lau, L., 1980. Farmer education and farm efficiency. *Econ. Dev. Cult. Change* 29 (1), 37–76.
- Lotz-Sisitka, H., Mukute, M., Chikunda, C., Baloi, A., Pesanayi, T., 2017. Transgressing the norm. *Int. Rev. Educ.* 63 (6), 897–914.
- Magrath, B., 2015. The role of civil society in education for development. In: McGrath, S., Gu, Q. (Eds.), *Routledge Handbook of International Education and Development*. Routledge, Abingdon.
- Majumdar, S., 2011. Developing a Greening TVET Framework. https://unevoc.unesco.org/fileadmin/user_upload/docs/Greening_TVET_Framework-Bonn-Final_Draft.pdf.
- Malerba, F., 1999. Sectoral Systems of Innovation and Production. Paper Presented at the DRUID Conference on National Innovation Systems, Industrial Dynamics and Innovation Policy, Rebild, June.
- McCowan, T., Schendel, R., 2015. The impact of higher education on development. In: McGrath, S., Gu, Q. (Eds.), *Routledge Handbook of International Education and Development*. Routledge, Abingdon.
- McGrath, S., Powell, L., Alla-Mensah, J., Hilal, R., Suart, R., 2020. New VET theories for new times. *J. Vocat. Educ. Train.* <https://www.tandfonline.com/doi/ref/10.1080/13636820.2020.1786440>.
- McGrath, S., 2010. The role of education in development. *Comp. Educ.* 46 (2), 237–253.
- McGrath, S., 2014. The post-2015 debate and the place of education in development thinking. *Int. J. Educ. Dev.* 39, 4–11.
- McGrath, S., 2015. Education and the post-2015 development agenda. In: McGrath, S., Gu, Q. (Eds.), *Routledge Handbook of International Education and Development*. Routledge, Abingdon.
- McGrath, S., 2018. *Education and Development*. Routledge, Abingdon.
- McMahon, W., 1999. *Education and Development*. Oxford University Press, Oxford.
- Meadows, D.H., Meadows, D.L., Randers, J., Behrens, W., 1972. *The Limits to Growth*. Universe Books, New York.
- Meyer, H.D., Benavot, A., 2013. *PISA, Power, and Policy*. Symposium, Oxford.

- Mincer, J., 1974. Education, Experience and Earnings. National Bureau of Economic Research, New York.
- Nussbaum, M., 2000. Women and Human Development. Cambridge University Press, Cambridge.
- Paine, T., 1984(1791)[]. Rights of Man. Penguin, Harmondsworth.
- Paulson, J., Rappleye, J., 2007. Education and conflict: essay review. *Int. J. Educ. Dev.* 27, 340–347.
- Powell, L., McGrath, S., 2014. Exploring the value of the capability approach for vocational education and training evaluation. In: Carbonnier, G., Carton, M., King, K. (Eds.), *Education, Learning, Training*. Brill Nijhoff, Leiden.
- Powell, L., McGrath, S., 2019. Skills for Human Development. Routledge, Abingdon.
- Psacharopoulos, G., 1981. Returns to education. *Comp. Educ.* 17 (3), 321–341.
- Psacharopoulos, G., 2006. World Bank policy on education: a personal account. *Int. J. Educ. Dev.* 26 (3), 329–338.
- Raworth, K., 2017. Doughnut Economics. Penguin, Harmondsworth.
- Robeyns, I., 2006. Three models of education: rights, capabilities and human capital. *Theor. Res. Educ.* 4 (1), 69–84.
- Robeyns, I., 2017. Wellbeing, Freedom and Social Justice. Open Book, Cambridge.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F., Lambin, E., Lenton, T., Scheffer, M., Folke, C., Schellnhuber, H., Nykvist, B., De Wit, C., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R., Fabry, V., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J., 2009. Planetary boundaries. *Ecol. Soc.* 14 (2), 32.
- Sachs, W., 1992. The Development Dictionary. Zed, London.
- Sachs, J., 2008. Common Wealth. Penguin, Harmondsworth.
- Santos, B., 2014. Epistemologies of the South. Routledge, Abingdon.
- Sayer, A., 2012. Capabilities, contributive injustice and unequal divisions of labor. *J. Hum. Dev. Capab.* 13 (4), 580–596.
- Schultz, E., Christen, M., Voget-Kleschin, L., Burge, P., 2013. A sustainability-fitting interpretation of the capability approach. *J. Hum. Dev. Capab.* 14 (1), 115–133.
- Schultz, T., 1961. Investment in human capital. *Am. Econ. Rev.* 51 (1), 1–17.
- Sen, A., 1999. Development as Freedom. Oxford University Press, Oxford.
- Sen, A., 2005. Human rights and capabilities. *J. Hum. Dev.* 6 (2), 151–166.
- Smith, L., 1999. Decolonizing Methodologies. Zed Books, London.
- Sterling, S., Maxey, L., Luna, H., 2013. The Sustainable University. Routledge, Abingdon.
- Sterling, S., 2019. Becoming “Learner Drivers” for the Future. <https://www.taylorfrancis.com/sdgo/goal/QualityEducation/all/timcontent/f9a62148-b0f9-454c-ba7a-fa0a5942a551>.
- Stewart, F., 2013. Nussbaum on the capabilities approach. *J. Hum. Dev. Capab.* 14 (1), 156–160.
- Streeter, P., 1979. Basic needs. *J. Pol. Model.* 1, 136–146.
- Swilling, M., 2020. The Age of Sustainability. Routledge, Abingdon.
- Tikly, L., 2020. Education for Sustainable Development in the Post-colonial World. Routledge, Abingdon.
- Tomaševski, K., 2001. Human Rights Obligations. www.right-to-education.org/sites/right-to-education.org/files/resource-attachments/Tomasevski_Primer%203.pdf.
- Ul Haq, M., 1995. Reflections on Human Development. Oxford University Press, Oxford.
- UN, 1948. Universal Declaration on Human Rights. UN, Paris.
- UN, 2015. Transforming Our World. UN, New York.
- UNESCO GEMR, 2016. Global Education Monitoring Report: Education for People and Planet. UNESCO, Paris.
- UNESCO, 2008a. The Right to Education: Monitoring Standard-Setting Instruments of UNESCO. UNESCO, Paris.
- UNESCO, 2008b. The Right to Education. UNESCO, Paris.
- Unterhalter, E., 2005. Global inequality, capabilities, social justice. *Int. J. Educ. Dev.* 25 (2), 111–122.
- Unterhalter, E., 2014. Measuring education for the millennium development goals. *J. Hum. Dev. Capab.* 15 (2–3), 176–187.
- Vare, P., Scott, W., 2007. Learning for a change. *J. Educ. Sustain. Dev.* 1 (2), 191–198.
- Von Tunzelmann, N., 2010. Alignment, misalignment and dynamic network-based capabilities. In: Dyker, D. (Ed.), *Network Dynamics in Emerging Regions of Europe*. Imperial College Press, London.
- Walker, M., Fongwa, S., 2016. Universities, Employability and Human Development. Palgrave Macmillan, Basingstoke.
- Walker, M., McLean, M., 2013. Professional Education, Capabilities and Contributions to the Public Good. Routledge, London.
- Wals, A., 2020. Transgressing the hidden curriculum of unsustainability. *Educ. Philos. Theor.* 52 (8), 825–826.
- Watene, K., 2013. Nussbaum's capability approach and future generations. *J. Hum. Dev. Capab.* 14 (1), 21–39.
- Wollstonecraft, M., 2004(1792)[]. A Vindication of the Rights of Woman with Strictures on Political and Moral Subjects. Penguin, Harmondsworth.
- World Commission on Environment and Development, 1988. Our Common Future (Brundtland Report). Fontana, London.
- Ziai, A., 2015. Post-development. *Dev. Change* 46 (4), 833–854.

The neoliberal dispositif: understanding the transformation of the social and ethical space of education

Bronwen M.A. Jones^a and Stephen J. Ball^{b,c,d}, ^a Institute of Education, University College London, London, United Kingdom; ^b Sociology of Education, Institute of Education, University College London, London, United Kingdom; ^c Journal of Education Policy, University of Copenhagen, Copenhagen, Denmark; and ^d University of Glasgow, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	60
Neoliberalism as “almost” a doctrine	60
Neoliberalization	61
The dispositif	62
The neoliberalization of education	63
The education market: epistemological, ontological and ethical reorganizations	64
Choice and accountability—the prioritization of instrumental ethics	64
The commodification of education-choice	65
The commodification of students-accountability	65
The commodification of the self	66
The discourse of numbers—epistemological dominance	67
Discussion	67
References	68

Introduction

Neoliberalism is a term that has enabled insightful and meaningful analysis of many of the economic, political, social and cultural changes that have occurred across the globe since the 1970s. It is impossible to make sense of the wholesale and significant changes in the world of education over this period without coming to terms with the phenomena that is neoliberalism. However, there is considerable disagreement over how best to conceptualize neoliberalism.

In some respects, neoliberalism appears to have assumed the status of a political theory or ideology, without many of the requisite defining characteristics. There is no central text, no validated list of key founding principles against which to define an idea, value or practice as neoliberal. We might reference the Washington Consensus as a “quintessential expression” (Hilgers, 2012, p. 81 cited in Collier, 2012, p. 192), or Hayek’s *Road to Serfdom* (1944) as a seminal text but to call neoliberalism a “theory”, economic, philosophical or political, offers it a cohesion and unity it does not have. Furthermore, the various ways in which neoliberal ideas and principles have been enacted and modified have pushed and pulled at the limits of any general doctrinal framework that could be attached to neoliberalism. It is hard not to concur with commentators that as an analytic category, neoliberalism is a “rascal concept” (Peck, 2013, p. 133), a “rather overblown notion” (Dean, 2012, p. 70) with “a perplexing mix of overreach and under-specification” (Brenner et al., 2010, p. 183). Moreover, it is impossible to ignore that much discussion of neoliberalism has something of the end of days about it. Referenced as a “big Leviathan” (Collier, 2012), a “slouching beast” (Ball, 2016) it has arguably become “loose shorthand for a prevailing dystopian zeitgeist” (Venugopal, 2015, p. 168) and in certain quarters has evolved into “a generic term of deprecation” (Thorsen and Lie, 2006, p. 9). There is no shortage of rhetorical flourish in the academic debate about neoliberalism, it is personified as sly and deceitful, insinuating itself into societal structures and governance—a cannibalizing and colonizing force programming the population to embrace the market and its concomitant social relations. Neoliberalism is often anthropomorphized as the villain in a grand narrative lamenting the global shift away from the golden era of the Keynesian welfare states (Barnett, 2005). Compelling though this story is, caution is needed. We do not offer a fully comprehensive re-telling of the story of neoliberalism, rather we highlight some of the key aspects and transitions in the way it has been understood and employed. This includes an exploration of the concept and practice of neoliberalization and also addresses the evolving and problematic nature of “actually existing neoliberalism.” Prompted by some of the shortfalls in existing accounts of neoliberalization, we introduce Foucault’s concept of the dispositif as an analytic device. We then focus on education, examining first how the two main approaches to neoliberalization have been used to make sense of the impact of neoliberal reform on education. Finally, we explore the usefulness of the *dispositif* by employing it as a methodological and conceptual tool to examine the way neoliberal education reform has transformed the social, cultural and ethical space of education.

Neoliberalism as “almost” a doctrine

Most accounts of neoliberalism begin with its emergence as an economic “approach” in the aftermath of the Second World War. The founding of the Mont Pelerin society in 1947 serves as a convenient marker for the development of neoliberal thinking. Following

the catastrophic resurgence of nationalism in 1930s Germany and faced with the spread of communism, this society brought together economists, philosophers and historians to discuss the role of liberalism in securing the freedom and rights of the individual over and against the role and responsibilities of the state. The likes of Hayek in Europe and Friedman, Stigler, Becker, Schultz and like-minded economists of the Chicago school feted the potential of *laissez faire* market economics as an alternative to collectivist political philosophy and Keynesian economics (Peck and Tickell, 2002, p. 388). At this point, neoliberalism can be described as a theoretical and abstract undertaking, existing in a “proto” stage (Peck and Tickell, 2002, p. 384), which saw it evolve to exhibit a degree of intellectual coherence, refining what can be described as a “neoliberal” economic perspective:

a free market ideology based on individual liberty and limited government that connected human freedom to the actions of the rational, self-interested actor in the competitive marketplace.

Stedman Jones (2012, p. 2).

However, this coherence should not be over-emphasized since even at this stage of “abstract intellectualism” (Peck and Tickell, 2002, p. 388), a range of different convictions and principles informed the broad agreement on the value of market practices (Blomgren cited in Thorsen and Lie, 2006). This is unsurprising, given the different approaches and disciplines represented among its advocates—neoliberalism does not represent a unified and consistently justified set of principles or tenets. This is significant because it means that the enactment and practices of neoliberalism carry greater weight in helping to define and understand it. The difficulties scholars grapple with in terms of how to conceptualize and analyze neoliberalism do not really pertain to its diversity or lack of rigor and coherence at a theoretical stage. They come from what is understood to be the enactment of neoliberal ideas into “actually existing neoliberalism.”

The 1980s marked the transition of neoliberalism into an active political project, partly as a consequence of and response to the perceived failure of Keynesian economics, the Oil Crisis, stagflation in the 1970s. Neoliberal economics was positioned as an alternative to state planning and *ungovernability* (Jessop, 2016) and a route out of global crisis. In the US and the UK, this was pursued in combination with *neoconservatism* as what was termed the politics of the New Right. The elections of Margaret Thatcher and Ronald Reagan inaugurated an era in which neoliberal thinking came to dominate and frame political thinking and practice in parts of the West (Hall, 2003; Jessop, 2004). Indeed, neoliberalism became a global project exemplified by the Washington Consensus and policies of structural adjustment promoted by the International Monetary Fund and World Bank. Enactments of neoliberalism across the world in multiple, different ways intensified the debate about how best to understand it. From the 1980s onwards then neoliberalism began to reference something “actually existing” (Peck et al., 2018) with all the accompanying epistemological and ontological complications that “actually existing” brings. It became necessary for researchers and commentators to account for and understand certain continuities and coherences and differences and contradictions that characterize neoliberalism in practice. A key question was how to explain the rapid and extensive but uneven development and spread of neoliberal thinking—to understand how and why it is that “neoliberalism seems to be everywhere” (Peck and Tickell, 2002, p. 380). To capture this complexity, scholars turned their focus to the processes of *neoliberalization*. Studies of neoliberalization pay attention to context and change, to neoliberalism as an ongoing “process, not an end-state” (Peck and Tickell, 2002, p. 383), as practice rather than theory. This emphasis on empirical study is both a means of sharpening theoretical understandings of neoliberalism and capitulation/recognition that “There is no such thing as neoliberalism!” (Barnett, 2005, p. 7) since its contingent existences call into question its existence as anything other than “a pure archetype” (Castree, 2006, p. 4). Either way, these explorations of the ways neoliberal thinking has been adopted and adapted in practice raise difficult questions about the form and nature of neoliberalism itself. We present below an overview of the two main approaches that have emerged in respect of analyzing and understanding neoliberalization.

Neoliberalization

Heuristically speaking, accounts of neoliberalization have fallen into two camps. Firstly, those that identify neoliberalization at a macro structural level as a phenomenon exhibiting key characteristics and practices of the market form. This kind of macro approach attempts to convey and apprehend a “common thread” (Hall, 2003, p. 22) or “pervasive metalogic” (Peck and Tickell, 2002, p. 383) identified in a range of economic, political and social changes that have taken place in some manner or another across the globe. These are analyses of big N neoliberalism or “neoliberalism writ large” (Ong, 2007, p. 4). They can frame neoliberalism as a “Leviathan” (Wacquant, 2010, p. 197), “an economic tsunami that is gathering force across the planet” (Ong, 2007, p. 1), a monolithic radical free market project whose practices of deregulation and privatization and reduction or reformation of the state are recognizable to all. This approach helps us to apprehend commonalities in processes of neoliberalization. Often these macro structural accounts include within or alongside them a sense of neoliberalism as a hegemonic project. It is “the spider at the center of the hegemonic web that is world wide market rule” (Peck, 2013, p. 133).

By contrast, an alternative camp begins by critiquing macro structural approaches for portraying neoliberalism as a monolithic and hegemonic venture rolling out a template of neoliberalism across the world. It points to variations, contingencies and discrepancies in the way neoliberalization has proceeded. The “flux and diversity” of neoliberalism is highlighted and explored (Brenner et al., 2010, p. 183) and neoliberalization is observed to take place in “circumstances of contradictory cohabitation” (Peck, 2013,

p. 140) reflecting its input “as just one of many pulses” (Peck, 2013, p. 141) in a local situation. Ong is an exemplar of this approach with her analyses of the way neoliberal strategies have been taken up by East and Southeast Asian states as exceptions to their usual practices. Proposing an understanding of neoliberalism “not as a standardized universal apparatus” (Ong, 2007, p. 5) but a “migratory set of practices” (Ong, 2007, p. 4), Ong highlights the diversity and contingency of new manifestations of neoliberalism (Ong, 2007, p. 1). In this schema, neoliberalization occurs not as an imposition but as an interaction with pre-existing social, political and cultural conditions to give rise to varied and diverse manifestations. This mirrors the more general notion of *glocalization* (Robertson, 1995) as opposed to globalization. This “low flying” (Brenner et al., 2010, p. 199) case study approach works with an understanding of political power not as a “hegemonic, thoroughly structurant, state dwelling power” (Cotoi, 2011, p. 110) but as governance. Indeed, examining neoliberalization as a contingent, relational process inevitably problematizes the idea of neoliberalism as a hegemonic ideology. That neoliberalism may co-exist with alternative or even competing government rationalities necessitates an understanding of it as exhibiting a more “light touch”, adaptable model of power.

The second approach is associated with another distinction in the general field of neoliberal studies. Rather than an exclusive focus on structural and economic aspects of neoliberalism there is, to a greater or less extent, an attempt to explore the neoliberalization of social, ethical and personal experience—the way neoliberalism makes up people. This is tied to an understanding of neoliberalism as “as a hybrid form of governmentality” (Brenner et al., 2010, p. 183) with its inclination “to speak in terms of neoliberal modes of subject reformation and strategies of rule” (Brenner et al., 2010, p. 199). That said, putting together the understandings of neoliberalism as a flexible, contingent and diverse assemblage with “neoliberalism as governmentality” do present some problems. The latter tends to present neoliberalism as a dominant or ruling government rationale in a way similar to the overblown, dystopian accounts of economic change noted previously.

This cursory review belies much of the nuance that informs different conceptualisations of neoliberalization. Peck and Tickell acknowledge neoliberalism as “necessarily variegated and uneven” and stress that “analyses of neoliberalization must be sensitive to its contingent nature” (2002, p. 383). Similarly, Ong makes repeated mention of the mobility of neoliberal logic (2007). Thinking about the abstract and the concrete together is never easy but doing so allows us to consider the ways that neoliberalism is implicated and imbricated in multiple relations of power. These relations of power facilitate its spread as a discursive regime and at the same time enable it to take hold in specific and mundane places where one might not expect it—the personal, social and ethical spaces of human life. At the heart of these tensions is the struggle to understand and articulate how neoliberalism appears to transform the way we understand and relate to ourselves, others and wider society. The question is how best to address the suspicion that we are somehow unwittingly and unwillingly created and shaped as neoliberal subjects. It is this concern with restructuring of our personal and social life that underlies and provokes descriptions of neoliberalism’s pernicious and manipulative nature. Fundamentally, the anxiety is that neoliberalization represents an extension of government power, of neoliberal governance. Many accounts of neoliberalism and neoliberalization are limited in their capacity to explore this. Hence, we suggest Foucault’s concept of the dispositif as a way of tracing and exploring/connecting, though not necessarily resolving, these tensions and anxieties.

The dispositif

As discussed, macro structural approaches of neoliberalization often imply or work with notions of a neoliberal hegemonic ideology. However, the emphasis on social and cultural dominance is a blunt instrument with which to interrogate the apparently weasel-like, infiltrative ability of a global discourse—neoliberalism—to reframe social and personal life. Equally, the creation of individual subjectivities that embody and instantiate a government rationale seems to miss something of the wider diffuse atmosphere and milieu, the lived experience, produced by neoliberal practice. Foucault’s concept of the dispositif embraces these tensions and in doing so allows us to understand and unpack more subtly how neoliberalism and neoliberalization operate as modes of power.

Foucault is often referred to as a theorist of power and indeed his work proposes and traces multiple categorizations of power: disciplinary, sovereign, pastoral, bio power and governmentality. However, the dispositif is not a categorization of power but rather a way of conceptualizing and apprehending experience that allows relations of power to be foregrounded. It directs our attention to the bricks and mortar of neoliberalism and neoliberalization and enables us to identify how they might constitute, develop and connect different modalities of power.

Foucault describes the dispositif as a:

thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions—in short, the said as much as the unsaid (1980, p. 194).

This heterogeneous ensemble is cohered and connected by a particular web of meaning, a theme, a concept. To examine or propose a dispositif is to understand and analyze its existence as a formation that encompasses multiple forms. However, it is the relations between such diverse items that visibilize the dispositif and is critical to naming it. Analyzing a dispositif begins with an interrogation of concrete, quotidian practices and discourses. It involves tracing a varied and dynamic network of policies, practices, and discourses and establishing patterns of connection and coherence across material and discursive and macro and micro

levels of enactment. This is not a straightforward or finite process since the dispositif is based on or generated by constituents that are multivalent. They have their own genealogies and potential trajectories that may concur or contradict any overall strategic direction of a dispositif and in doing so alter it. It is the set of relations, a net of meaning that holds together a motley selection of discourses, practices, buildings, gestures, signs etc. that is the dispositif. "The apparatus [dispositif] itself is the system of relations that can be established between these elements" (Brenner et al., 2010, p. 194).

These relations involve a play of power and Foucault talks of and distinguishes between local tactics and overall strategies while reflecting on the mutuality of them both. They are conjoined at specific "points of contact." While individual parts may be connected by strategic objectives, they are also subject to "strategic elaboration" in ways that may be unforeseen, contradictory or both. The key in deploying the dispositif is to identify how these strategic affiliations, elaborations and relations cohere to generate, exert and manifest power.

The dispositif makes it possible to conceptualize the production, maintenance and evolution of a neoliberal social, cultural and ethical milieu. It is able to do this by capturing the uneven, diverse and evolving network of regimes of knowledge and practice that comprise our social reality. It looks to highlight the relations of power that lend coherence to this network and traces them across material and discursive and macro and micro levels of enactment. In taking seriously the materiality of thought, of power and knowledge, it is able to track the instantiation and reiteration of neoliberal truths as they cohere to produce a tangible and identifiable regime of truth. The materiality and contingency of its component parts and the conceptual relations that cohere them are integral to the structure of the dispositif. Without both and without an apprehension and experience of both, there is no dispositif.

We now want to turn our attention to the way in which neoliberal thinking and neoliberal practices have influenced and manifested themselves in the world of education. In other words, we want to examine those processes and developments that have been understood as the neoliberalization of education. Following this, we illustrate how the construct of the dispositif can develop our understanding of how the neoliberalization of education develops, is maintained and expands.

The neoliberalization of education

Education systems around the world have been subject to patterns of reform that speak to the influence of neoliberalism. Drawing on the wider theoretical debates concerning the nature of neoliberalism and neoliberalization can help in tracing the extent and evolution of educational reform, both nationally and globally.

Approaches that emphasize macro structural accounts of neoliberalization have been valuable in apprehending how the neoliberal reform of education has spread as a global practice. They aim to identify fundamental structural patterns that have been repeated across the world as a kind of "orthodoxy" that can be traced back to early "New Right" education reforms introduced in the United States and England and Wales (Fuller and Stevenson, 2019, p. 1). This is most clearly represented by Sahlberg's evocation of GERM-Global Education Reform Movement-in which he outlines five defining features of the global manifestation of neoliberal practice and approaches to education. Sahlberg's acronym-come -analogy also neatly conveys the reifying tendency of scholarship with regards to neoliberalism, conceptually constructing neoliberal reforms as "an epidemic that spreads and infects education systems through a virus" (Sahlberg, nd).

This level of analysis has also provided an understanding of the "staging" of the cumulative and extensive nature of the structural change many education systems have undergone. Stages of neoliberalization are identified in which the role of the state is progressively reconfigured from a "deregulation and dismantlement" of the Keynesian-welfare state to an "emergent phase of active state building and regulatory reform" as neoliberal principles are enacted more extensively and creatively (Peck and Tickell, 2002, p. 384). Peck and Tickell refer to this as roll back and roll out neoliberalism. Studies of the "ratchet" (Ball, 2007, p. 19) from endogenous to exogenous privatization in education in England and Wales illustrate this kind of theorizing as we set out below. England and Wales are not alone in exhibiting this pattern and are seen by some commentators as a testing ground for neoliberal reforms, which are then mimicked or adapted elsewhere (Grek and Ozga, 2010).

The UK's 1988 Education Reform Act initiated the deregulation and devolution of power and funding away from Local Education Authorities to individual schools and parents and concomitantly the assertion of central state control (Ball, 1990). This was a form of endogenous privatization through which the education system was restructured and re-cultured as a quasi-market—a market made and managed by the central state (Glennerster, 1991; Whitty, 1989). At the same time parents were given the opportunity to choose between competing providers, where competition was articulated by institutional performance, and funding followed choice on a per capita basis. These moves were of strategic importance in paving the way for, or making thinkable and therefore doable forms of exogenous privatization. That is, the introduction of private firms, companies, organizations and individuals as providers within the education marketplace. It is the intensification and extension of the latter that has been understood as a more extensive and ongoing phase of roll out neoliberalism that far from representing a shrinking of the state represents a reconfiguration of the state and intensification of governance. Examining this pattern of reconfiguration, scholars such as Ball and Junemann (2012) have drawn on the work of Jessop to argue that this extensive neoliberal reform has resulted in the reconstitution of the state as a form of network governance involving a complex inter relationship of private, public, voluntary organizations in the role of governance. This kind of analysis has implications for how education relates to wider forms of governance. The important point here is that neoliberalism may be set over and against the state in many respects but as a practical project it is also dependent on the state—in particular the facilitative role of the state as a "market maker."

In order to explore the nature and manner of that dependence, accounts of neoliberalization that emphasize the hybridity and contingency of its manifestations have proved particularly useful. These approaches have placed greater emphasis on variations in the manifestation of neoliberal influence on national education systems. They also point to a more complex process of policy borrowing and policy mobility that acknowledges the variations in the pace, extent and convolution of neoliberal reform (Ball, 2012a, 2017). Studies of school choice and accountability in particular have evidenced the diversity and adaptability of neoliberal practice to its contingent political and social environments. Chile is cited as the most extreme example of a neoliberal system of school choice with its adoption of a voucher scheme in the 1980s. England and Wales with policies of local management of schools and the more recent development of Academies and Free schools in England are equally seen as an outlier of neoliberal adoption (Gunter, 2011). The key is that while a global spread of choice policies is evident, increasingly research has identified the idiosyncrasies and specificities of the way they are taken up and enacted (Lingard, 2010; Wu, 2014). A good example of this is Tan's adoption of Ong's notion of neoliberalism "as exception" (Ong, 2006) to characterize Shanghai's neoliberal school choice and assessment policies as "neoliberalism with Chinese characteristics" (2019).

Most pertinent are those analyses that marry this emphasis on hybridity and contingency with an understanding of neoliberalism as a form of governmentality. Work in this area investigates how the neoliberalization of the structures and economics of education has changed the way in which education is understood and practiced and how this in turn positions those involved with education. It probes the social, cultural, ethical transformation of education as strategically fundamental to the promotion and production of particular kinds of neoliberal subjectivities. This emphasis maintains the importance of modalities of state power but focuses on the operation of these as the production of particular kinds of social and ethical subjects (Youdell, 2004). Ball discusses the manifold ways in which neoliberal education policy reworks us into neoliberal subjects (2003, 2012b, 2016). Further studies extend this kind of analysis to explore neoliberalization as a process that effects the formation of neoliberal subjectivities across education (Bailey, 2015; Bradbury, 2019; Spohrer et al., 2018). However, this approach acknowledges neoliberalization as a process of messy overlays and complex interpenetration with earlier and different forms of governance in which other forms of power are constantly evoked and interacted with (Allen, 2014; Gore, 1995; Hunter, 1994).

The neoliberalization of education undoubtedly entails structural change, but it also references a transformation of the social, cultural and ethical milieu and meaning of education—what it means to teach and learn, to be educated. Whether situated within the debate about the nature of neoliberalism or not, this transformation of our understanding and articulation of education is a recurring theme of much educational research over the past thirty years. Drawing on and bringing together some of this research, we develop the argument that neoliberalization is not just a process that occurs "out there" in structures and economic relations but also takes place "in there" as a social, cultural and ethical process, as experience. Neoliberal education rests upon, brings about, and flourishes not simply on the basis of a set of tenets or principles, or through a set of changes in systems of delivery, or in the enactment of new forms of social relations but in and through a complex arrangement and ensemble of practices, methods, ethics, interactions and subjectivities that alter, perhaps profoundly the epistemological basis of social and educational experience. It is in outlining and exploring the elements of this ensemble that the Foucauldian concept of the *dispositif* proves valuable.

In order to explore how it is that the social, cultural and economic spaces of education are constituted within and as part of a neoliberal dispositif, we have identified the epistemological, ontological and ethical coherences and affinities that span and frame relations between policies, practices and discourses. We interrogate two critical aspects of the neoliberalization in education: the generation and securement of neoliberal principles and practices as truthful and sensible and their evolution as an ongoing and contingent process. It is unwieldy to trace this from a global perspective and so we have focused on education reform in the "neoliberal heartlands or neoliberal 'home' spaces" (Peck and Tickell, 2002, p. 382) of England and Wales. This focus acknowledges the role of English education as "a social laboratory of policy ideas" that have been exported to and adapted by many other national systems (Exley and Ball, 2014).

The education market: epistemological, ontological and ethical reorganizations

We focus on school choice and accountability while recognizing the impossibility of separating these from each other or from concomitant and subsequent policy developments. Our aim is to consider how structural changes and particular discourses interact and connect. In particular we focus on the way these practices and discourses are cohered by and consequently instantiate, replicate and promote particular epistemological, ontological and ethical approaches. These approaches are fundamental to building and constituting a neoliberal social, cultural and ethical educational milieu.

Choice and accountability—the prioritization of instrumental ethics

Focusing on the UK as an exemplar of neoliberal education, the structural changes brought in by the 1988 ERA of devolving financial control to schools, per capita funding and open enrollment were the first steps in a continuing process that is reconfiguring almost all educational spaces as a competitive market in England and Wales (Ball, 1990; Bowe et al., 1992). This process of structural reorganization was accompanied by the development of a National Curriculum with standardized testing. This testing regime served as a system of accountability in a devolved market ensuring comparability across organizations and individuals. Systems of accountability are fundamental to the competitive environment of an education market in which schools, teachers, and children are constantly assessed, compared and selected (Ozga, 2009, 2013; Power, 1997) (see below). These changes to the education system in

England and Wales are seen as exemplifying a fundamental characteristic of neoliberalism and processes of neoliberalization. They evidence a devolvement of power and responsibility from government to a variety of providers and contractors while simultaneously instigating comprehensive national regimes of accountability that represent a modification, arguably an increase and extension, of government power. Over the past thirty years, processes of devolvement and systems of accountability have escalated and intensified and the school system in England and Wales is now a fragmented, competitive and diverse market simultaneously subject to centralized imperatives and interventions. The acceleration of the Academies program and the development of Free Schools have seen an influx of new providers that have been explored as various forms of heterarchical network governance (Ball and Junemann, 2012). Alongside competition and freedom and diversity and innovation the educational apparatus is animated and inundated by a bewildering and reactive form of “policy hyperactivity”—mostly aimed at raising performance—that is driven by ministerial enthusiasms and biases, international orthodoxies and often ill-informed and ill-thought out borrowings from other systems (Morris, 2012).

Policies and practices of choice and accountability as they are enacted in England and Wales are part of a seemingly ad hoc ensemble. We want to examine how they are cohered in relatively consistent way at the level of and by social and ethical relations. Below we identify processes of instrumentalization and commodification and the dominance of numerical knowledge as features characteristic of many of these relations.

The commodification of education-choice

Policies and practices of school choice reposition schools, parents and students as producers and consumers of education (Gewirtz et al., 1995; Olmedo and Wilkins, 2017; Rowe and Lubienski, 2017; Vincent and Ball, 2006; Wilkins, 2010). This reworks education as a product or commodity to be marketed as schools competed with one another for students (Gewirtz et al., 1995, chapter 5; Whitty and Power, 2000). The introduction of a National Curriculum and national testing to facilitate choice further enables the commodity education to be defined, standardized and measured so that parents are able to make informed choices. The academic performance of a school as recorded and published in league tables stands for and represents the value of the product education offered by each school (Ball, 2004). Both school choice and accountability entail and contribute to the commodification of education. Education is now positioned as a private rather than public good that is of value to the individual student. This prompts an instrumental appreciation of education and knowledge with an emphasis on what value it can accord them in the future.

The process of school choice is situated within a wider framework of how education is understood and valued. The notion of the knowledge economy, introduced by Drucker (1966), references the idea that knowledge can be treated as a product, valued in terms of its transactional leverage in the competitive job market. Whereas once workers had exchanged their labor for reward, now information and knowledge are the new “wealth creating assets” (Ball, 2017, p. 25). Again, this positions knowledge/education as a private good valued and chosen for the capacity to improve success in a competitive job market. This is an economic perspective on education that commodifies knowledge/education, instrumentalizing their value to produce a kind of educational *homo oeconomicus* (Foucault, 2010, p. 268–269). This wider policy context of the knowledge economy imports economic rationalism into education (Lingard et al., 2003) and in so doing, reiterates the commodification and instrumentalization of education that already characterize practices of school choice.

We can see two processes of commodification here, one embedded within the other. The wider market of the knowledge economy shapes the contours and orientation of the education market, and reiterates the relationship with education as one of self-interested instrumentality. This involves a layering and legitimization of an instrumental ethic and logic that runs through both and secures markets as sites of veridiction and value allocation (Foucault, 2010, p. 31).

The commodification of students-accountability

The commodification of education is also a product of and integral to systems of accountability. While school choice and accountability are integrated and overlapping practices, commodification is contextualized differently with different consequences. Systems of accountability are framed by and imbricated within wider policy discourses of efficiency, standards, “what works”, productivity and performance. These systems of accountability place emphasis on the performances of individuals and organizations as “measures of productivity or output, or displays of ‘quality’, or ‘moments’ of promotion or inspection” (Ball, 2003, p. 216). A performative culture uses measures and indicators to judge, compare, punish or reward and not only spotlight the individual or institution, but position them in a relation of competition with each other. It produces winners and losers, averages, norms and trends-numerical data- and this process has outcomes not only for the way those involved perceive themselves but also how they relate to others. Ball describes this process as performativity.

Research shows that due to the pressure for schools to “perform” and to “compete”, schools that are able to seek students who will perform well in public examinations with minimum input and at minimum cost (Ball, 2004; Ball and Gewirtz, 1997; Kenway and Bullen, 2001). In addition, the pressure for schools to produce good exam results has led to practices that can be understood as “gaming the system”, the construction of admissions strategies, exclusion policies, direction of teacher attention to “productive” students, putting children in for easier exams (Lingard and Sellar, 2013; Waslander et al., 2010). In the classroom, research has shown how teachers focused on students on grade boundaries in order to maximize percentage pass rates (Gillborn and Youdell, 2000; Youdell, 2004). These strategies make clear that children are valued, or not, as producers of academic performance, and

approached as commodities. The commodification of education that necessitated the marketization of education has led to the commodification of students:

Thus, schools and teachers are effectively encouraged to value students according to what these children can offer the school financially and in terms of exam performance and image. In this way, students have become objects of the education system, to be attracted, excluded, displayed and processed, according to their commercial and semiotic worth, rather than subjects with needs, desires and potentials

Gewirtz (2002, p. 124).

Ball refers to this as an economy of student worth:

The demands of competition, the “information” provided by League Tables, pressures from the state for performance improvement and target-achievement and per-capita funding, in a period of spending constraints, work together to create local “economies of student worth” (2004, p. 10).

The commodification of the child as a bearer of potential and realized academic performance transforms the relationship between the school, teacher and student. The academic performance of the child becomes highly significant to the teacher and school as an indicator of their own professional value, which in some systems is translated into performance related pay. It charges those relationships with an ethic of self-interested instrumentality; a built-in imperative to see and treat others as a means to an end, more specifically your own end. There is little doubt that this has brought about a radical transformation of the nature of the ethical relationship between teacher, student and school. Moreover, that this has been an issue of considerable concern. Academic research has identified anxiety about the impoverished relationship between students and teachers that result from neoliberal policy and practice (Ball, 2003; Codd, 2005; Cooper, 2004; Gewirtz, 1997; Gewirtz et al., 2009; Jeffrey and Woods, 1998). Teacher stress, generated by work overload resulting from the pressure to perform, has allowed little space for teachers to develop relationships with students or to care for them in a way that many teachers feel they wish to (Gewirtz, 2002). Such change has made it increasingly difficult to maintain personal relationships. Cooper summarizes what this looks like:

Empathic teachers exhaust themselves finding pockets of profound empathy for needy children in corridors and in the entrances and exits to lessons, but it is never enough (2004, p. 20).

All of this exemplifies the ways that structural changes associated with the market, school choice and accountability serve as a hinge that prompts the emergence of a particular ethical stance to others. The creation and maintenance of a quasi-market in education precipitates processes of commodification and instrumentalization that shape school practices and teacher student relationships. It fosters a particular kind of ethical practice. In such articulations the social and the economic blur or converge as more and more of the social, the educational, the psychological, and the interpersonal is opened up to the possibilities of calculation and competition in our relations to ourselves and to others.

The commodification of the self

Allied and integral to this process of commodification is an understanding of the self as a form of human capital (Becker, 1964). It is clear that the practices and structures of market competition in education position the child, and indeed schools and teachers, as sites of investment, projects demanding work and improvement in order to succeed. The positioning of education as a private good to be acquired for the benefit of the student means the student is prompted to understand him or herself as a site of investment, to approach their own development as a means to an end. It creates a divided self, a kind of self-commodification and a relationship of instrumentality toward yourself. This is a particular ontological organization of the self. Students are encouraged to “turn themselves into a project” in order to best compete. This echoes Foucault’s notion of the *homo oeconomicus*, the enterprising self—a character fleshed out in policies of entrepreneurialism and recently resilience (Peters, 2001). The enterprising self is characterized by autonomy, self-interest and a calculative logic in their decision-making (Olssen et al., 2004, p. 169–170). This represents another reiteration of instrumental ethics and alters profoundly the social relations of teaching. Both the teacher and the student are exhorted to be entrepreneurs of themselves and the moral implications of this are clear. “Every social transaction is conceptualized as entrepreneurial, to be carried out purely for personal gain” (Olssen et al., 2004, p. 137). Indeed, the reiteration of self-interested ethic of instrumentality demonstrates an ability to infiltrate even those areas of the curriculum that might have been intended to offer an authentic opportunity to engage students. Thomson and Gunter comment on the appropriation of “pupil voice”:

Within the education portfolio, there is a marked tendency for senior policy makers to bring “pupil voice” into the policy conversation as a means of achieving school improvement and higher standards of attainment, rather than as a matter of the UN convention, citizenship and rights (2006, p. 840).

This pervasive understanding of the self and others as a form of human capital reiterates a form of instrumental ethics that effectively acknowledges and secures the power of the market—a school or job market—to define and fix value. This process is facilitated and characterized by the privileging of certain types of knowledge and we want to consider this below.

The discourse of numbers—epistemological dominance

The discourse of the number and numerical “knowledges” are integral to practices of school choice and accountability and processes of commodification and instrumentalization. The privileging and disproportionate emphasis on and employment of numerical knowledge is fundamental to the restructuring of education as a market (Ball, 2017, p. 218; Ozga, 2008). The economic model of the market is rooted in scientific and numerical knowledge, as are the discourses of efficiency, productivity, and standards etc. that come from it. Such epistemological dominance extends beyond the league table and begins before it framing policy through a privileging of quantitative methodologies and policy agendas that are “research based” (Ozga, 2008). The education policy field has become dominated by and preoccupied with the production of numerical data (Bradbury and Roberts-Holmes, 2018). Numerically based knowledges hold an epistemological grip on the way that education is understood, articulated and practiced. It is important to understand the role of this discourse in transforming education.

In the context of education, numerical knowledge effects and represents the reduction of complex personal, social and moral realities for the purposes of assessment, comparison and competition; a process neatly described by Rose as “a rhetorical technique for black boxing” (1999, p. 208). In other words, the hegemony of the “number” through technical, mathematical, statistical, scientific discourses serves to obfuscate the complexity and fragility of the processes, relationships and commitments that lie behind and beyond the production of a number. This is an oversimplification, invisibilization and invalidation of categories of experience that many see as fundamental to the educative experience. Numerical discourse intensifies commodification by reducing and sidelining emotional, affective, social and moral relations. The numerical articulation represents an epistemological framework that facilitates commodification not only cohering multiple practices and discourses but by shaping them. This is a key way in which the ethical and social space of education is re-formed.

Moreover, the production of techno-scientific data at once constitutes that which it claims to represent while portraying itself as a neutral and purely representative form of knowledge. It hides not only the complexity of real-life processes but its own value-driven perspective (Ozga, 2008). It is a reductive account of education that presents itself as objective, impartial fact.

Numerical and economic and scientific discourse is a key strand of epistemological coherence that runs across and joins up elements of a neoliberal dispositif. It is fundamental to systems of accountability and performativity; it facilitates processes of commodification and instrumentalization. It is a language that articulates market logic and legitimates and renders neutral practices such as accountability while devaluing relational aspects of the educational experience. It creates and reiterates a frame of reference that bastardizes aspects of the educative process, the self and relationships with others.

Discussion

Outlined above are some of the ways that structural practices of neoliberal educational reform embody and facilitate the prioritization of particular epistemological, ontological and ethical standpoints. In tracing the key epistemological, ontological and ethical continuities and affinities that run throughout neoliberal discourses and practices, we are able to identify a critical strand cohering the neoliberal dispositif. This strand contributes to a powerful context in and through which “neoliberal education” takes place and is defined. It also helps to set discursive boundaries that delegitimize alternative competing discourses and practices that might encourage us to think differently about education. The dispositif then allows us to visibilize points of ethical coherence, neoliberal moral truths, that come together to create an environment that is, somewhat ironically, greater than the sum of its parts—a powerful milieu and marriage of calculation and judgment.

Policies of school choice and systems of accountability work hard on various fronts to standardize commodify and instrumentalize education, individuals and relationships. These processes overlap and are replayed from different perspectives and on multiple fronts. Education is instrumentalized by the prioritization of examination results but also by an economic understanding of education. There is a layering of this epistemological and ethical framework that serves to normalize and legitimate particular ways of understanding education, the self and others.

Equally, these processes are circumscribed and sometimes contradicted each by the others and by other discourses—they are bordered by multiple horizons. Supporting discourses and practices are manifold: the knowledge economy is a powerful frame within which the quasi market of education operates, performance culture lends a degree of urgency and commitment to both commodification and instrumentalization, numerical discourse provides a cloak of objectivity and facilitates easy practice.

In tracing their intersections and abutments, we can see the way that policies and practices of choice and accountability interact to open up spaces in which the prioritization of particular epistemological, ontological and ethical positions becomes “logical” and legitimate. Reiterated in a cumulative process of multiple, intersecting policies and practices, they constitute a powerful regime of truth. This goes beyond the logic and practices of the market and is about the assumptions and values built into that logic that makes it valid and tenable. They tether the market logic to wider, possibly more “substantial” discourses of truth concerned with the dominance of science and the number. This is a deeper level of epistemological and ontological and ethical coherence that locates neoliberalism’s intransigent existence not only in the nature and force of its own rationality but also in its kinship

and affinity with wider supporting discourses or regimes of truth. But where does this leave our understanding and conceptualization of neoliberalism?

Peck argues convincingly that the term neoliberalism remains necessary because it effectively, though not perfectly, marks out space that can be the focus of resistance. That seems right, but the question remains how best to mark out such “neoliberal” space. Neoliberalism needs to be conceptualized in a way that centers those aspects of its “actually existing” nature that are problematic. For many, this boils down to the way neoliberalism works as a form of governance—ineluctably referenced in terms of its persistent and pernicious global spread and adaptability and its surreptitious and sinister reformation of the soul. We have focused specifically on the extension of that governance from the “out there” of the structural reform of education to the “in there” of ethical constitution. We argue that the methodological construct of the dispositif helps us to “mark out” and understand the changes that have taken place in the social and ethical understanding and practice of education in three critical ways.

The dispositif captures the materiality of power. It is rooted in analysis of the concrete and quotidian practices and discourses that constitute part of the mundane educational environment in which education is defined and people are constituted as neoliberal educational subjects. Importantly, it recognizes the force and power that the reiteration and replaying of consistent themes and tropes lend to practice. The dispositif captures the dynamic creation of a milieu that circumscribes what constitutes reasonable and legitimate thought and action. This is the delineation of a neoliberal space within which relationships, meanings and subjectivities develop and are enacted while they in turn reiterate and extend that space.

Furthermore, the dispositif takes seriously the notion of neoliberalism as process. As a model, it places at its center the constant tension between fixedness and fluidity of the social reality of neoliberalism. As a methodological construct, it is not misled by the singular and intimidating moniker of “neoliberalism” that can direct us to essentialism and all the problems associated with that. The notion of a neoliberal dispositif is premised on and takes account of the ongoing and perpetual construction of social reality.

Lastly, and again because of its inherent epistemological and ontological foundations, the dispositif directs attention to coextensive and supporting discourses and practice at its edges and boundaries. This is important. Discourses of instrumental ethics and of the number exist beyond the remit of neoliberalism. They stand and fall on their own merits and limitations and can be tackled on this basis. It’s important to acknowledge when we are addressing neoliberalism and when we are addressing its support acts, even when they have been coopted. The resistance Peck refers to should be accurately directed because the process of unraveling does not always begin with the most obvious strand.

The dispositif is a methodological tool that adds process, fluidity and evolution as key characteristics to the modality of a noun. It is therefore founded on and expresses a model of social reality that allows us to capture how neoliberalism works as a form of power/governance in setting boundaries of thought and action. The process of constant thematic layering of the heterogeneous ensemble of practices, discourses, policies and technologies together mark out the “strategic envelope” of a neoliberal dispositif (Foucault, 1998, p. 100). Identifying and analyzing this thematic ordering as done here allows us to trace the transformation of the social and ethical space of education. It enables us to understand that transformation as the development of a neoliberal dispositif that directs and shapes the way we think about and practice education. In doing this, it visibilizes and explicates that transformation as an extension of neoliberal governance.

To return to Peck, if neoliberalism as a concept “does define a problem space and a zone of (possible) pertinence ... this represents the beginning of a process of analysis” (2013, p. 153). We suggest Foucault’s notion of the dispositif may prove a way forward in this process of analysis and one that encourages us to think differently about, and outside of neoliberalism.

References

- Allen, A., 2014. *Education in and Beyond the Age of Reason*. Palgrave Macmillan, Basingstoke.
- Bailey, P., 2015. Consultants of conduct: new actors, new knowledges and new “resilient” subjectivities in the governing of the teacher. *J. Educ. Adm. Hist.* 47 (3), 232–250.
- Ball, S.J., Gewirtz, S., 1997. Girls and the education market. *Gen. Educ.* 9 (2), 207–222.
- Ball, S., Junemann, C., 2012. *Networks, New Governance and Education*. Bristol University Press, Bristol.
- Ball, S.J., 1990. *Politics and Policy Making in Education*. Routledge, London.
- Ball, S.J., 2003. The teacher’s soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Ball, S.J., 2004. *Education for Sale! The Commodification of Everything*. Annual Education Lecture, Kings College, London.
- Ball, S.J., 2007. *Education Plc.: Understanding Private Sector Participation in Public Sector Education*. Routledge, London.
- Ball, S.J., 2012a. *Global Education Inc. New Policy Networks and the Neo-liberal Imaginary*. Oxon, Routledge.
- Ball, S.J., 2012b. Performativity, commodification and commitment: an I-Spy guide to the neoliberal university. *Br. J. Educ. Stud.* 60 (1), 17–28.
- Ball, S.J., 2016. Subjectivity as a site of struggle: refusing neoliberalism? *Br. J. Sociol. Educ.* 37 (8), 1129–1146.
- Ball, S.J., 2017. *The Education Debate*. The Policy Press, Bristol.
- Barnett, C., 2005. The consolation of neoliberalism. *Geoforum* 36 (1), 7–12.
- Becker, G., 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. University of Chicago Press, London.
- Bowe, R., Ball, S.J., Gold, A., 1992. *Reforming Education and Changing Schools: Case Studies in Policy Sociology*. Routledge, London.
- Bradbury, A., Roberts-Holmes, G., 2018. *The Datafication of Early Years and Primary Education: Playing with Numbers*. Routledge, Abingdon.
- Bradbury, A., 2019. Making little neo-liberals: the production of ideal child/learner subjectivities in primary school through choice, self-improvement and “growth mindsets”. *Power Educ.* 11 (3), 309–326.
- Brenner, N., Peck, J., Theodore, N., 2010. Variegated neoliberalization: geographies, modalities, pathways. *Glob. Netw.* 10 (2), 182–222.
- Castree, N., 2006. From neoliberalism to neoliberalization: consolations, confusions, and necessary illusions. *Environ. Plann.* 38 (1), 1–6.
- Codd, J., 2005. Teachers as “managed professionals” in the global education industry: the New Zealand experience. *Educ. Rev.* 57 (2), 193–206.
- Collier, S.J., 2012. Neoliberalism as big Leviathan, or...? A response to Wacquant and Hilgers. *Soc. Anthropol.* 20 (2), 186–195.
- Cooper, B., 2004. Empathy, interaction and caring: teachers’ roles in a constrained environment. *Pastor. Care Educ.* 22 (3), 12–21.

- Cotoi, C., 2011. Neoliberalism: a foucauldian perspective. *Int. Rev. Soc. Res.* 1 (2), 109–124.
- Dean, M., 2012. Free economy, strong state. In: Cahill, D., Edwards, L., Stilwell, F. (Eds.), *Neoliberalism: Beyond the Free Market*. Edward Elgar, Cheltenham.
- Drucker, P., 1966. *The Effective Executive*. Harper and Row, New York, NY.
- Exley, S., Ball, S.J., 2014. Neo-liberalism and English education. In: Turner, D., Yolcu, H. (Eds.), *Neo-liberal Educational Reforms: A Critical Analysis*. Routledge, London.
- Foucault, M., 1980. The confession of the flesh. In: Gordon, C., Foucault, M. (Eds.), *Power/Knowledge: Selected Interviews and Other Writings 1972–1977*. The Harvester Press, Sussex, pp. 194–228.
- Foucault, M., 1998. *The Will to Knowledge: The History of Sexuality*, vol. 1. Penguin, London.
- Foucault, M., 2010. The Birth of Biopolitics Lectures at the College de France 1978–1979 (G. Burchell. Trans.). Palgrave, London.
- Fuller, K., Stevenson, H., 2019. Global education reform: understanding the movement. *Educ. Rev.* 71 (1), 1–4.
- Gewirtz, S., Ball, S.J., Bowe, R., 1995. *Markets, Choice and Equity*. Open University Press, Milton Keynes.
- Gewirtz, S., Mahony, P., Hextall, I., Cribb, A. (Eds.), 2009. *Changing Teacher Professionalism: International Trends, Challenges and Ways Forward*. Routledge, London.
- Gewirtz, S., 1997. Post-welfarism and the reconstruction of teachers' working the UK. *J. Educ. Pol.* 12 (4), 217–231.
- Gewirtz, S., 2002. *The Managerial School: Post-welfarism and Social Justice in Education*. Routledge, New York.
- Gillborn, D., Youdell, D., 2000. *Rationing Education: Policy, Practice, Reform and Equity*. Open University Press, Buckingham.
- Glennerster, H., 1991. Quasi- markets for education? *Econ. J.* 101 (408), 1268–1276.
- Gore, J.M., 1995. On the continuity of power relations in pedagogy. *Int. Stud. Sociol. Educ.* 5 (2), 165–188.
- Grek, S., Ozga, J., 2010. Governing education through data: Scotland, England and the European education policy space. *Br. Educ. Res. J.* 36 (6), 937–952.
- Gunter, H.M., 2011. Introduction: contested educational reform. In: Gunter, H.M. (Ed.), *The State and Education Policy: The Academies Programme*. Continuum, London, pp. 1–18.
- Hall, S., 2003. New Labour's double shuffle. *Soundings* 24, 10–24.
- Hayek, F.A., 1944. *The Road to Serfdom*. Routledge, Abingdon.
- Hilgers, M., 2012. The historicity of the neoliberal state. *Soc. Anthropol.* 20, 80–94.
- Hunter, I., 1994. *Rethinking the School: Subjectivity, Bureaucracy, Criticism*. Allen and Unwin, St. Leonard's, NSW, Australia.
- Jeffrey, B., Woods, P., 1998. *Testing Teachers: The Effect of School Inspections on Primary Teachers*. Routledge, London.
- Jessop, B., 2004. New Labour's doppelte Kehrtwende: Anmerkungen zu Stuart Hall und eine alternative Perspektive zu New Labour. *Argument* 256, 494–504.
- Jessop, B., 2016. *The State: Past, Present, Future*. Polity Press, Cambridge.
- Kenway, J., Bullen, E., 2001. *Consuming Children: Education-Entertainment- Advertising*. Open University Press, Buckingham.
- Lingard, B., Sellar, S., 2013. "Catalyst data": perverse systemic effects of audit and accountability in Australian schooling. *J. Educ. Pol.* 28 (5), 634–656.
- Lingard, B., Ladwig, J., Luke, A., 2003. School effects in postmodern conditions. In: Slee, R., Weiner, G., Tomlinson, S. (Eds.), *School Effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements*. Falmer Press, London, pp. 84–100.
- Lingard, B., 2010. Policy borrowing, policy learning: testing times in Australian schooling. *Crit. Stud. Educ.* 51 (2), 129–147.
- Morris, P., 2012. Pick "n" mix, select and project; policy borrowing and the quest for "world class" schooling: an analysis of the 2010 schools White Paper. *J. Educ. Pol.* 27 (1), 89–107.
- Olmedo, A., Wilkins, A., 2017. Governing through parents: a genealogical enquiry of education policy and the construction of neoliberal subjectivities in England. *Discourse* 38 (4), 573–589.
- Olsen, M., Codd, J., O'Neill, A., 2004. *Education Policy: Globalization, Citizenship and Democracy*. Sage Publications, London.
- Ong, A., 2006. Neoliberalism as Exception: Mutations in Citizenship and Sovereignty. Duke University Press, Durham.
- Ong, A., 2007. Neoliberalism as a mobile technology. *Trans. Inst. Br. Geogr.* 32 (1), 3–8.
- Ozga, J., 2008. Governing knowledge: research steering and research quality. *Eur. Educ. Res. J.* 7 (3), 261–272.
- Ozga, J., 2009. Governing education through data in England: from self-regulation to self-evaluation. *J. Educ. Pol.* 24 (2), 149–162.
- Ozga, J., 2013. Accountability as a policy technology: accounting for education performance in Europe. *Int. Rev. Adm. Sci.* 79 (2), 292–309.
- Peck, J., Tickell, A.T., 2002. Neoliberalizing space. *Antipode* 34 (3), 380–404.
- Peck, J., Brenner, N., Theodore, N., 2018. Actually existing neoliberalism. In: Cahill, D., Cooper, M., Konings, M., Primrose, D. (Eds.), *The SAGE Handbook of Neoliberalism*. Sage Publications Ltd.
- Peck, J., 2013. Explaining (with) neoliberalism. *Territ. Politic. Gov.* 1 (2), 132–157.
- Peters, M., 2001. Education, enterprise culture and the entrepreneurial self: a Foucauldian perspective. *J. Educ. Enq.* 2 (2), 58–71.
- Power, M., 1997. *The Audit Society: Rituals of Verification*. Oxford University Press, Oxford.
- Robertson, R., 1995. Globalization: time-space and homogeneity-heterogeneity in global modernities. In: Featherstone, M., Lash, S., Robertson, R. (Eds.), *Global Modernities*. Sage Publications, London, pp. 25–44.
- Rose, N., 1999. *Powers of Freedom: Reframing Political Thought*. Cambridge University Press, Cambridge.
- Rowe, E., Lubienski, C., 2017. Shopping for schooling or shopping for peers: public schools and catchment area segregation. *J. Educ. Pol.* 32 (3), 340–356.
- Sahlberg, P., Global Educational Reform Movement is Here. Available at: <https://pasisahlberg.com/global-educational-reform-movement-is-here>.
- Spohrer, K., Stahl, G., Bowers-Brown, T., 2018. Constituting neoliberal subjects? "Aspiration" as technology of government in UK policy discourse. *J. Educ. Pol.* 33 (3), 327–342.
- Stedman-Jones, D., 2012. *Masters of the Universe*. Princeton University Press.
- Tan, C., 2019. Neoliberalism as exception: the new high-quality school project in Shanghai. *Discourse* 40 (4), 443–457.
- Thomson, P., Gunter, H., 2006. From "consulting pupils" to "pupils as researchers": a situated case narrative. *Br. Educ. Res. J.* 32 (6), 839–856.
- Thorsen, D.E., Lie, A., 2006. What is Neoliberalism? Department of Political Science: University of Oslo. Retrieved June, 2020 from: <https://jagiroadcollegelive.co.in/attendance/classnotes/files/1589998418.pdf>.
- Venugopal, R., 2015. Neoliberalism as concept. *Econ. Soc.* 44 (2), 165–187.
- Vincent, C., Ball, S., 2006. *Childcare, Choice and Classed Practices*. Routledge, London.
- Wacquant, L., 2010. Crafting the neoliberal state: workfare, prisonfare, and social insecurity. *Socio. Forum* 25 (2), 197–220.
- Waslander, S., Pater, C., van der Weide, M., 2010. *Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education*. OECD Education Working Paper No. 52.
- Whitty, G., Power, S., 2000. Marketization and privatization in mass education systems. *Int. J. Educ. Dev.* 20, 93–107.
- Whitty, G., 1989. The new right and the national curriculum: state control or market forces? *J. Educ. Pol.* 4 (4), 329–341.
- Wilkins, A., 2010. Citizens and/or consumers: mutations in the construction of concepts and practices of school choice. *J. Educ. Pol.* 25 (2), 171–189.
- Wu, X., 2014. *School Choice in China: A Different Tale?* Routledge, London.
- Youdell, D., 2004. Engineering school markets, constituting schools and subjectivating students: the bureaucratic, institutional and classroom dimensions of educational triage. *J. Educ. Pol.* 19 (4), 407–431.

Global human rights

Anja Mihr, Center on Governance Through Human Rights, Berlin, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	70
Global human rights	71
Global human rights education	71
Global programs for human rights dissemination	73
Global human rights regimes	74
The (r-)evolution of global human rights	74
Glocal human rights: from the global to the local	76
Conclusions	77
References	77
Relevant websites	78

Introduction

Global Human Rights is a composition of universally shared moral values, such as friendship, solidarity, fairness, respect, and trust, encompassing legal and political bindings instruments. One vital human rights instrument is the United Nations (UN) International Bill of Human Rights.¹ It enshrines ten key declarations and conventions that define and outline universal human rights norms and standards—most of them legally binding in customary law. They define the right to physical integrity (torture and ill-treatment), anti-racism and xenophobia (discrimination), women and girls rights, social, cultural, religious and economic rights, political and civil rights, the rights of migrants, the fight against disappearances, the rights of people with disability and the rights of children. These human rights are norms and standards, and for some, they are “tools” or “keys” to develop or “unlock” our capacities. Enhancing, respecting, and enforcing human rights also prevent us from harm in the form of ill-treatment and discrimination that could prevent us from prospering. One of these rights (Art 26 in the UDHR) is our entitlement to education which allows us to develop ourselves to the best of our capacities and live a dignified life. From that article, the concept of Human Rights Education (HRE) and, later, the 2011 UN Declaration on Human Rights and Learning derives.

Globally and locally, human rights can only be upheld and enforced by duty bearers and rights holders. State authorities are the most genuine duty bearers, but business and we as individuals bear the duty to respect these rights vis-à-vis others. These “negative human rights” protect us from violence, discrimination, and any form of abuse by third parties, overall corrupt state authorities, or armed groups. An example of a negative right is our human right to a fair and open trial in front of a court. The transparency of a trial and non-partisan judges aim to protect us from corrupt and arbitrary judicial systems. Other rights, such as freedom of religion and media, allow us to express and share our thoughts in a mutually beneficial way (Mihr, 2009). Nevertheless, human rights are a two-way road, always. If one expects that his/her physical integrity is respected and protected, that person needs to do anything in his/her capacity to do respect the rights of others, too.

Globally and locally, human rights are monitored, implemented, and enforced through legal and political mechanisms, such as human rights commissions, committees, courts, councils, human rights defenders or rapporteurs, and other monitoring bodies. Educational and training programs in schools, higher education institutions, think tanks, training of judges, and massive open online courses and video clips are other means to disseminate universal human rights norms.

International and regional courts, such as the Inter-American Court, African Court or European Court for Human Rights, OSCE and Council of Europe or African Union parliamentary commissions, and UN treaty-bodies are other legal and political binding means and ways of monitoring state obligations toward human rights fulfillment, as well as the human rights situation in general.²

Overall global human rights norms have become part of our awareness, daily judgments, and subsequently identity and conscience, habits, and behaviors (Fischlin and Nadorfy, 2007). We not only expect human rights standards to be fulfilled by others toward us, but we all have also become entrepreneurs, defenders, and agents of human rights promotion ourselves. That is what has made human rights truly glocal, hence global and local, today.

¹International Bill of Human Rights, ten core global human rights treaties: UN Office of the High Commissioner for Human Rights, International Bill of Human Rights, <https://www.ohchr.org/Documents/Publications/FactSheet2Rev.1en.pdf>.

²UN OHCHR <https://www.un.org/en/about-us/udhr/foundation-of-international-human-rights-law>.

Global human rights

Nevertheless, our globally shared thrive for justice and equality needs instruments and mechanisms that allow us the best personal and community-based development. Universal human rights norms and international human rights law (IHRL) are guidelines and standards for materializing the thrive people share. How we live and realize these values and norms are paraphrased in principles, laws, and rules in UN or other regional EU, CoE, OSCE, AU, OAS, ASEAN, and SCO treaties, declarations, and guidelines.³ For example, the human rights standards and law on the “right to a fair and open trial,” to a proper wage, to adequate housing, clean water, or our freedom to move and migrate, have to be formulated in such a way, that they are universally applicable and realistic.

Organizations institutions such as the UN monitoring treaty bodies and commissions, committees, and regional human rights courts, for example, aim to safeguard these global human rights norms. Nevertheless, they can have different regional priorities, often based on the history, memory, and legacy of war or conflicts in a country or region. Regional human rights regimes, such as the African Union (AU), the Organization for the American States (OAS), or the Council of Europe (CoE) and its monitoring committees, reporting systems, commissions, and courts, respond to domestic and local human rights matters.⁴ These instruments and mechanisms are complementary and mutually reinforcing toward the UN instruments and mechanisms, such as the Human Rights Council, the Treaty Body Regime, or the Special Procedures by the UN Office of the High Commissioner for Human Rights. Representatives of the different mechanisms, such as regional or domestic courts and human rights commissions, are in constant dialog and negotiation with decisions taken by international courts, claims, and reports by civil society organizations (CSOs) and human rights defenders (HRD). National and international governmental and non-governmental human rights organizations (NGOs), such as national human rights institutions (NHRI) and governmental bodies, complete the picture of the global human rights mechanism and regime (Christopher and Roberts, 2015).

Apart from the norm and standards-setting mechanism, for example, the UN treaty bodies and committees or commissions, CSOs or international human rights mechanism such as the European Court of Human Rights (ECtHR) and the Interamerican Court for Human Rights (IACtHR), that have often clarified, determined, and wrote human rights law and implemented them; the dissemination of human rights largely depends on the way human rights are put to practice, adhered to and enforced.

Human Rights Education (HRE) has thus become an indispensable tool and method to empower, lucite, teach, train, build capacities, illustrate and describe human rights, to adhere and fulfill them. By this, HRD, trainers, teachers, and academics empower individuals of all ages and backgrounds to understand the meaning and purpose of human rights and how to realize and materialize them in day-to-day life. Through HRE and empowerment, we understand how to implement equity, fairness, and mutual respect on a day-to-day basis (Mihir, 2010).

The fact that over the past decades, millions of young people have started their protest movements against social injustice, unfair trials, inequal payment, climate injustice, and unbearable living standards. The number of CSOs has dramatically risen in all parts of the world, and human rights have become an integral part of political rhetoric and business, resulting in HRE. The way the global human rights idea can be disseminated is defined in the United Nations Declaration on Human Rights Education and Training (2011).⁵

In the Declaration, HRE is defined as threefold, namely, (1) a lifelong process that fosters knowledge about human rights and acquiring skills to exercise them in daily life; Secondly (2) the shift of understanding and attitudes that determine our values and beliefs that uphold human rights; and Thirdly, (3) this learning process is about our behavior and taking action to defend and promote human rights.⁶

Global human rights education

In line with the 2011 Declaration, the UN Office of the High Commissioner for Human Rights describes HRE today as a tool.

promotes values, beliefs, and attitudes that encourage all individuals to uphold their rights and those of others. It develops an understanding of everyone's common responsibility to make human rights a reality in each community.

UN Declaration for Human Rights Education and Training (2011).

Consequently, Global Human Rights is a universally and internationally agreed set of human rights norms enshrined in international human rights law (IHRL) since the UN Charter of 1945 and the UDHR in 1948. Global Human Rights results from the worldwide mainstreamed dissemination process through CSOs, governments, Human Rights defenders, movies, media, pieces

³A general overview of Regional Human Rights systems, regimes, and mechanisms can be found under: “A Rough Guide to Regional Human Rights” <https://www.universal-rights.org>; The International Justice Resource Center <https://ijrcenter.org/regional/>; and the UN Office of High Commissioner for Human Rights, Regional Mechanism <https://www.ohchr.org/EN/Countries/NHRI/Pages/Links.aspx>.

⁴International Justice Resource Center: Regional Systems: [https://ijrcenter.org/regional/\(2021\)](https://ijrcenter.org/regional/(2021)).

⁵<https://www.ohchr.org/en/issues/education/training/pages/undhrededucationtraining.aspx>.

⁶UN Declaration on Human Rights Education and Training, GA, December 2011, <https://www.ohchr.org/en/issues/education/training/pages/undhrededucationtraining.aspx>.

of training, and online portals of which HRE is one aspect. HRE has become the focus of disseminating the idea of Global Human Rights since the World Conference for Human Rights in Vienna in 1993 and the subsequent decade for human rights education (1994–2004). Several World Programs followed up the decade in Human Rights Education over the past 30 years until the present.⁷ Global human rights norms and their implementation and enforcement tools, such as HRD, courts, changing of laws, and others, manifest themselves in a plethora of textbooks, educational video clips, and training programs in all languages, ideally free and easy to be download textbooks. Hence, HRE became global with the rise of Information and Communication Technologies (ICTs) and even with Artificial Intelligence (AI) which allows for more straightforward multiplication of information, data, and access to HRE. ICT and overall the Internet is seen as a double sword in human rights promotion because they also allow for false, fake, and manipulated information and images about human rights abuses. Because of this challenge of fake information today, HRE programs include Internet literacy in their pieces of training. It empowers people to know about human rights and abuse and detect false from accurate information. Furthermore, algorithms and intelligent AI programs can help scan and detect human rights violations both online and offline. Depending on how they are programmed, they can also discriminate based on skin color or age, exclude and censor Internet users based on their political ideas or sexual orientation, and hence increase the level of control and suppression in the hands of autocratic regimes.

Against this backdrop after the UN World Conference for Human Rights in Vienna in 1993, where all UN members states of the world agreed on the fundamental principles and norms of human rights and made human rights a truly global project, the first-ever UN Decade for HRE was launched and lasted until 2004. The idea of individual empowerment was universally endorsed—albeit not practiced. The more and better people are informed and trained in their human rights, the more they pose a threat to authoritarian regimes and organizations.

Albeit we notice a dramatic backslide of democracy and a rise of populist ethnic-nationalistic governments over the past two decades, at the same time the UN passed the HRE and Training Declaration in 2011, and in the 2015 Sustainable Development Goals (SDGs) a few years later.⁸ Both enshrine human rights standards as the essential tools to solve the world's critical problems and challenges, such as climate change, migration, and digitalization.

The recent fourth phase of the World Program on Human Rights Education was launched in 2020 and lasts until 2024, focusing on Youth and their needs for human rights empowerment in times of migration urbanization. At the same time, bearing in mind that a third of the world's population is young and frustrated about bad governance and lack of social mobility, the program comes in timely. The program responds to the continuous protesters and Youth movements ranging from Hong Kong, Bangkok, Moscow, Beirut, Tel Aviv, Khartoum, Almaty to Santiago de Chile. Today's Youth protests and demands result from HRE, yet they are demanding more human rights compliance within their social and political regimes, not less.

Furthermore, the number of online and offline materials used to teach and train citizens on their human rights has reached a record number, thus resulting in people's empowerment to stand for one own and the human rights of others. The first three decades mark an essential era for human rights education and pedagogy, and since 1993 millions of globally acting international, national and local CSOs and globally acting NGOs such as Human Rights Watch and Amnesty International, elevate human rights empowerment and education to one of their primary goals to unlock human rights. HRE is seen as the most effective prevention against human rights violations (Steven and Jensen, 2016).

Article 26 in the UDHR from 1948 and Articles 13 and 14 in the UN International Covenant on Social, Economic, and Cultural Rights from 1966 underline that everyone has then natural right to education and that the fulfillment of these rights should be free, at least in the elementary and fundamental stages. Education and lifelong learning should also be directed to fully developing the human personality and respect for human rights and fundamental freedoms. It has been specified in many conventions, treaties, declarations, working programs, and action plans.

Furthermore, in the 1999 UN General Comment, HRE is seen as a tool to promote global human rights, namely that "(...) education shall be directed to the human personality's sense of dignity, it shall 'enable' all persons to participate effectively in a free society, and it shall promote understanding among all 'ethnic' groups, as well as nations and racial and religious groups".⁹ Moreover, but philosophically, it is emphasized that HRE is a method that enables "(...) a well-educated, enlightened and active mind, (able) to wander freely and widely, is one of the joys and rewards of human existence".¹⁰

HRE has come a long way, from being a subject taught in school and an activist campaign to raise awareness to a pro-active tool to change our minds, attitudes, habits, and behavior. Nevertheless, teaching *about* human rights and knowing them remains its primary purpose and hence, is only half the story of global human rights. To know about human rights is not yet empowered to proact in the name of human rights-beyond the awareness-raising and attitudinal aim.

Nevertheless, HRE aims to guide our day-to-day decisions and options to materialize human rights. Today, millions of people are forced to leave their homes based on their religious or ethnic background, failed governmental policies combatting climate

⁷UN World Program for Human Rights Education and Training since 2004, <https://www.ohchr.org/en/issues/education/training/pages/programme.aspx>.

⁸The UN Human Rights Council, in its resolution 24/15 (October 8, 2013), decided to focus the World Program's third phase (2015–2019) on strengthening the implementation of the first two phases and promoting human rights training for media professionals and journalists. This resolution was adopted following the OHCHR consultation on the focus of the third phase, as presented in the High Commissioner's report A/HRC/24/24. OHCHR, in consultation with States, intergovernmental organizations, national human rights institutions, and civil society elaborated a plan of action for the third phase (2015–2019) of the World Program.

⁹UN OCHCHR, *General Comments* (1999): CESCR General Comment No. 13: The Right to Education (Art. 13) *Adopted at the Twenty-first Session of the Committee on Economic, Social and Cultural Rights, on 8 December 1999 (Contained in Document E/C.12/1999/10)*.

¹⁰UN OCHCHR, *General Comments* (1999).

change pollution or overpopulation, and war is one key challenge for human rights. These people need to be integrated, empowered, and resettled in different societies and cultural environments, asking each member of that society, migrants and citizens alike, to act according to IHRL and standards.

To get evidence of concentration and labor camps, humans being trafficked and tortured, people starving, disappearing, and getting abducted is putting a gloomy picture against the rise of human rights awareness. One of the reasons we wish to stop this injustice is our awareness of human rights and the dramatic rise of human rights activists, HRD and CSOs, trained attorneys, and judges that aim to put an end to it. Overall, human rights abuse is a sign of failed state authority to be guardians of human rights, the duty bearers. Instead, they are the key violators. The corrupt misconduct of civil officers, paramilitaries, and governmental authorities also allows non-state actors, warlords, and organized crime to operate within their state territories and violate fundamental rights. The challenge is that many warlords, paramilitaries, and organized criminals cannot easily be held accountable in front of courts since they operate within a deep-state, in the remote uncontrolled territory, in cyberspace, or under the umbrella of the government.

Moreover, the current challenges and developments in global human rights illustrate that even though we globally share the horror and aversiveness about these conducts, we feel individually incapable of resolving such issues. HRE helps to understand better the root causes of such violations of human rights and what to do about it.

Global programs for human rights dissemination

On December 10, 2004, the UN General Assembly of the United Nations proclaimed the first World Program for Human Rights Education to advance programs in all sectors (General Assembly's resolution 59/113, December 10, 2004). The World Programs are globally implemented and promoted, mainly through NGOs (Mahler et al., 2009). As enshrined in the International Bill of Human Rights and IHRL, global human rights norms have been an integral part of the Sustainable Development Goals (SDGs) action plans on national and local levels since 2015. Millions of capacity building training, information sessions, films, projects, programs, training activities, and clips on YouTube, Instagram, and other social media channels, seek to promote a common understanding of the three basic principles of HRE within the SDGs, namely (1) knowledge, (2) attitudes and (3) behavior. In order to show mutual respect and fairness and guarantee not only my rights but the rights of others, we need tools, methods, and a level of self-determination to act and decide in this matter.

If one lives and works in a society that is dominated by authoritarian rules—as is the case in over 50% of the world's countries, often rooted in the past, ethnical rivalries or ideologically disputes, the options to realize human rights and protect them are restricted to the private realm, such as family, school, neighborhood, work or city (Coysh, 2018). The critical question is, what can one do to safeguard human rights, and what is one capable of doing under the existing circumstances? Whether HRD, Youth organizations, CEOs of companies, policymakers, or city councilors, the various actors and stakeholders who commit to HRE provide a concrete framework for action and strengthen partnerships and cooperation from the international level down to the grassroots (Oberleitner, 2013).

The recent UN World Programs to foster global human rights awareness and HRE have been structured in consecutive phases to target specific sectors/issues national efforts. For example, some focus on primary and secondary school systems or exclusively on higher vocational and university education and human rights training programs for teachers and educators, civil servants, law enforcement officials, and military personnel. Human rights awareness-raising programs target media professionals and journalists: relevant resolutions, plans of action, reports, and other information.

In 2018, the UN Human Rights Council, in its resolution 39/3, makes Youth and civil society a priority. The council wants to strengthen those who are often the only forces to oppose corrupt security forces, judges, and authoritarian governments. In some countries, up to 40% of the population is below 30. The emphasis is on empowering these young women and men to understand better and use the tools striving for equality, development, and non-discrimination.

Likewise, in the Agenda for Sustainable Development (SDGs) and specifically with target 4.7, the UN OHCHR, in collaboration with national governments, elaborated a Plan of Action which puts Youth in the center of activities (World Program (A/HRC/42/23—adopted by the UN Human Rights Council through resolution 42/7 on September 26, 2019)).¹¹ The consequences were a dramatically rising global Youth Movements activism. Millions of activists, supporters, and groups network against racism, climate change-induced violations, and the lack of social mobility worldwide. The Youth and citizen movements are the strongest ever seen across the globe.

¹¹ Provisions on human rights education have been incorporated into many international instruments and documents, including the Universal Declaration of Human Rights (art. 26); the United Nations Educational, Scientific and Cultural Organization (UNESCO) Convention against Discrimination in Education (art. 5); the International Convention on the Elimination of All Forms of Racial Discrimination (art. 7); the International Covenant on Economic, Social and Cultural Rights (art. 13); the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (art. 10); the Convention on the Elimination of All Forms of Discrimination against Women (art. 10); the International Labor Organization Indigenous and Tribal Peoples Convention, 1989 (No. 169) (arts. 30 and 31); the Convention on the Rights of the Child (art. 29); the International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families (art. 33); the Convention on the Rights of Persons with Disabilities (arts. 4 and 8); the Vienna Declaration and Program of Action (Part I, paras. 33–34; Part II, paras. 78–82); the Program of Action of the International Conference on Population and Development (paras. 7.3 and 7.37); the Durban Declaration and Program of Action (Declaration, paras. 95–97; Program of Action, paras. 129–139) and the outcome document of the Durban Review Conference (paras. 22 and 107); and the 2005 World Summit Outcome (para. 131).

The UN OHCHR believes that HRE methods increase awareness changes attitudes, such as race, political diversity, religion, and gender, and hence aims to promote peaceful and participatory (democratic) solutions that lead to sustainable development and social justice. Still today, opinions vary on achieving social justice, democracy, the rule of law, and sustainable development through education. Even though studies have shown that HRE is fundamentally important in transitional justice processes and during the societal and political transition process, continuous learning about human rights and becoming self-reflective and critical is an ability that also helps consolidated democracies to stay a tune (Dalsgaard and Paulsen, 2009).

Global human rights regimes

Regional human rights regimes are vital mechanisms to uphold and enhance human rights realization, such as the AU, the CoE, the OAS, the European Union (EU), ASEAN or the Office for Human Rights, and Democratic Institutions of the OSCE in Warsaw (ODIHR). The monitoring, read the report, hold a hearing, write recommendations, and regional courts hold sessions on the country or individual cases that breach human rights. ASEAN has recently adopted the Bangkok Declaration of Human Rights. Other international organizations or summits have called global human rights norms integral to their statements, declarations, and policies (Sundrijo, 2020).

Human rights norms, principles, and values as laid out in the UN International Bill of Human Rights with its core treaties, including the UDHR and the UN Treaty Bodies (OHCHR Database), underpin international human rights mechanisms and instruments. With the establishment of the UN Human Rights Council in 2006, its legal and political instruments and mechanisms, and the reforms and rise of regional human rights regimes in the existing realm of the OAS, AU, CoE, EU, OSCE, AU, and ASEAN. They all incorporate, to some extent, elements of the Bill and foster human rights promotion of women's rights, minority rights, and elements of security (Tsutsui, 2018).

International human rights law (IHRL) and the concept "universal jurisdiction" that allows any domestic court to charge a person who commits crimes against humanity and HRE have grown into universal tools to unlock human rights in any aspect of life and any corner of the world. Universal jurisdiction refers, for example, to the practice of national courts. In Belgium, Germany, the US, or the Netherlands, for example, individuals have been prosecuted and charged for serious crimes against international law, such as crimes against humanity, war crimes, genocide, and torture, and thus crimes that are enshrined in the 1999 Rome Statute and practices by the International Criminal Court (ICC) in The Hague. This universal and thus global act of jurisdiction can also be applied by domestic courts in any part of the world and has thus contributed to globalizing human rights.

Universal jurisdiction and any human rights compliance are based on the principle of not harm others, in the light of IHRL. Hiding, escaping, fleeing to avoid jurisdiction has become more difficult for war criminals and torturers over the past year (Macedo, 2006).

International Law based on global norms and its jurisdiction is no longer of temporal nature. It can be applied anywhere at any time and is truly global. Other forms of justice, such as transitional justice (TJ), are temporal by referring to a particular past period, i.e., civil war, genocide, Apartheid-era, or autocratic regimes. TJ aims to bring the perpetrator to justice by illustrating and exemplifying the importance of human rights norms and the Rule of Law.

The (r-)evolution of global human rights

The earliest foundation of HRE and universal jurisdiction in modern times is embodied in the UN Charta of 1945 and the UDHR of 1948. Chapter 10 of the UN Charta empowered the Social and Economic Council (ECOSOC) to act on behalf of human rights and was later in 2006 replaced by the UN Human Rights Council in Geneva.

However, human rights never really became an integral part of UN activities but instead had to be laid down in additional treaties and a plethora of non-legally binding declarations, resolutions, and recommendations ever since the UDHR of 1948. The core human rights conventions, of which there are ten today that define human rights as global norms against discrimination of any kind, based on gender, faith or ethnicity, consciousness or race and promoting freedoms to self-determination and individual development in which our idea of dignity is based, came in the decades after. The international UN covenant on social-economic rights (ICESR), and the international UN covenant on civil and political rights (ICCPR) from 1966 (1976), are probably the two most significant legally binding agreements that have thus far determined our globally shared understanding of safety and freedoms. Moreover, article 26 of the UDHR guarantees the promotion, understanding, tolerance, and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace. (UN General Assembly *Universal Declaration for Human Rights. Resolution 217 A (III.)*, 1948).

The subsequent UN Treaty Bodies, emerging in the 1970s first on political and on women, children, economic and migrant rights, and later on torture, disappearances, and people with disability, have become standard setters and best practices for NHRI and regional organizations regimes and courts, such as the African Court and Commission for Human Rights (2006) or the EU Fundamental Rights Agency (FRA) since 2010. The foundation of regional human rights regimes started in post-WWII Europe (Strasbourg), followed by initiatives in the Americas (Costa Rica) (Pereira, 1997).

The Universal Period Review (UPR) of human rights was installed in 2006 by the UN OHCHR and the Human Rights Council. Up to the present day, it is the most universal, inclusive, and hence global procedure to monitor human rights performance and violation. The UPR procedures, led by the OHCHR, regularly list the human rights records of all 193 UN Member States equally, based on the International Bill of Human Rights. Governments are seen in this process as the duty bearer vis-à-vis billions of individual right-bearers worldwide. Governmental representatives at the UN level have to justify why certain human rights are not upheld and implemented or violated, and thanks to the ICT, today often as live-streams and webinars, in front of a world's public audience, at least in all vital world languages, Arabic, Chinese, English, Russian, Spanish and French.

When we ask how and who disseminated the principles, norms, and standards of human rights first in modern times? The answer is NGOs and CSOs, respectively. The story of global human rights dates back to the rise of NGOs, such as Amnesty International and Human Rights Watch, during the high times of the Cold War. A period also describes one of the battles of the minds between communist totalitarian regimes in the East and democratic liberal regimes in the West. The notion that human rights are a "Western" concept often dates back to when human rights activists and human rights courts could only be active in the West. In the 1960s, Amnesty International and other civil rights and non-governmental organizations started to fight for the rights of political prisoners. Later in 1977, the Helsinki Committees and Human Rights Watch emerged. Soon HRE, campaigning, advertising. Became an integral part of awareness-raising and behavior shifts these NGOs aimed at. With NGOs' rise of human rights campaigns in the 1980s, people became more aware of human rights and their abuses. The growing human rights awareness, triggered through peace education and human rights campaigns, evolved into a more elaborated concept of HRE, which became an educational concept of its own.

Human rights awareness campaigns, HRE and HRD, gained more legitimacy after NGOs and UN agencies such as UNESCO worked jointly in the 1990s. After the Cold War, human rights NGOs became an integral part of governmental advice and monitoring procedure. Human Rights Master programs, summer schools, and training have flourished at Academic Institutions since the 2000s. LL.Ms and MA programs focusing on human rights, peace, sustainability, governance are found in all major universities today. Business and transnational companies have human rights policies and hire "human rights officers." Any political statement or campaign by governmental leaders highlights—in one way or the other—the importance of human rights, knowing that sooner or later, they can be held to account for it. CSOs and NGOs are growing by the day and in numbers focusing their work on a large array of human rights matters.

For decades, UNESCO, for example, organized expert meetings based on Article 26 of the UDHR and called upon governments to introduce HRE, but without much success. HRE remained a rather elite and extra-curricula endeavor, only subject to voluntary teaching and training. After the 1993 Vienna Conference that led to the Global Action Plan for Human Rights, HRE, training and capacity building slowly became part of the international human rights agenda. Former ideological borders of the Cold War no longer blocked the dissemination of human rights. Many newborn states wanted to become members of the UN and the Council of Europe.

International governmental organizations (IGOs), governments, and CSOs work together more closely than ever to safeguard human rights. However, much of today's human rights protection and dissemination is left to CSOs, media, and IGOs, particularly in terms of schooling, in remote areas, and in protecting minorities and migrants, which poses a challenge to global human rights. Human rights have become more local than ever, and hence glocal, but at the same time, state authorities withdraw more and more from taking responsibilities as the vital duty bearer. That leaves human rights fulfillment often to HRD, CSOs, individual judges, and even business, whereas the state withdraws. Alongside various stakeholders, private, governmental, non-governmental, and public have embraced the concept of HRE as a means to strengthen civil society and democracy (Brysk, 2018).

Private companies, enterprises, startups, and development organizations make political and constitutional concessions within their business plans, including human rights compliance procedures, such as supply-chains, environmental and labor laws, human resource policies, and practice equal opportunity guidelines, etc.

Since the beginning of global human rights, outsourcing human rights protection from governmental duties has remained a crucial challenge. Millions of CSOs today are the job of the governments, but also IGOs such as the CoE, the EU, the AU, the OAS, and the UNESCO, or OSCE and the UN established offices, branches, and agencies in countries and regions supporting governments and NGOs. They conduct programs and projects to safeguard human rights.

Consequently, it was just a matter of time before Cold War established education concepts. Those working in the informal peace, tolerance, democracy, genocide, and inter-cultural education sectors would want to participate in HRE (Mihir, 2015). From then on, NGOs which traditionally worked in tolerance, peace, or development issues also started competing for money and resources for HRE with other agencies in this field. That is also why today, many of these agents and NGOs claim that their concepts, such as peace, tolerance, development are equally essential parts of HRE. Most educational concepts that deal with peace, conflict resolution, reconciliation, history, tolerance or democracy, and civic issues claim today that they are, in their essence, HRE programs. The problem is that they do so often without changing or adapting their pedagogical or didactical outlines to the objectives of HRE (Mahler et al., 2009).

Overall, the early 1990s constituted a window of opportunity for international organizations, NGOs, and progressive states alike to foster HRE and promote full-spectrum human rights education programming. This became most evident when the Council of Europe.

Glocal human rights: from the global to the local

Human rights have become glocal when global and universal norms and standards meet local realities and civil societies (Mihir, 2019). Think global-act locally: *glocal* practices are the norm, not the exceptions, and are the new normal today. Glocal human rights mean that global norms are individually and locally understood and lived by all of us on a day-to-day basis—or at least aimed to live by.

Glocalizing human rights have always been the UN's aim when inviting CSOs and public-private partnerships to make the UDHR a living document to begin with. It is written in any human rights treaty that human rights can only be achieved through and by people's will. Governments are one among many actors to realize human rights. However, other stakeholders such as CSOs, CEOs, NGOs, businesses, and HRD worldwide are crucial ones to make human rights a glocal reality.

Article 10, of the Convention of the Right of People with Disability from 2006, for example, states the glocality of human rights, when highlighting that it is national governments, that “reaffirm that every human being has the inherent right to life and shall take all necessary measures to ensure its effective enjoyment by persons with disabilities on an equal basis with others” (Mihir and Gibney, 2014).

In practice, this means that governments have to support, facilitate and undertake any possible measure to ensure full realization by and for the people, namely full inclusion of people with mental and physical disability and those with none, in school, sports, public administration, police, military. This, and all other core treaties, guarantee equal opportunities for all people.

Notwithstanding, realizing universal human rights norms is often a long way. It faces obstacles in political will, doability, and money. However, Article 6 of the ICCPR from 1966 was already paraphrased in Article 1 that

“All peoples have the right of self-determination. Under that right, they freely determine their political status and freely pursue their economic, social and cultural development (...) and that government shall promote the realization of the right of self-determination, and shall respect that right, in conformity with the provisions of the Charter of the United Nations (...) and hence for every person or citizen living in their territorial boundaries”.¹²

Conventions such as the ICCPR aim to support and empower people, make decisions, and resolve on a day-to-day basis what has always been the core problem of humanity, namely inequality (Tibbitts, 2020). Whether inequality is perceived or actual, when people are treated unequally and unfairly, they will strive for either revenge or justice. To regulate that strive in a fair, free and peaceful manner, this is what the Bill of Rights aims at. HRE is one way of explaining how that can be done. People-driven global movements for human rights and HRE have always faced challenges and obstacles, mostly politically and culturally driven (Moses et al., 2020).

Language obstacles, cultural diversity, political regime type, or so-called “traditional values” which are never closely defined have challenges and shaped HRE programs, as illustrated by Nancy Flowers, one of the founders of worldwide HRE initiatives since the 1970s. It was in fact, during the Cold War, “(...) radical teachers in the Global South who showed the world the power of HRE to further both civil-political and social-economic rights”; had always argued any exceptionalism when it came to global human rights, as she argues (Flowers, 2015, p. 2). As written in the Bill and hundreds of other regional or domestic action plans, conventions, and Declaration, human rights have always been aimed at all people equally, not at nationalities. Their purpose is not to privilege any ethnic or gender group, any age or religious group over others. The fact that privileges and exceptionalism are often the root cause for human rights violations is why HRE has mattered so much over the past decades.

Today, think global—act locally is the norm enshrined in the SDGs from 2015. The 75th UN Anniversary Working Groups in 2020 and 2021 have tirelessly reiterated and addressed local threats to global human rights, such as girls exclusion from schools, climate change-induced migration, access to clean water, and so on; all matters that increase inequality and hence lead to a breach of global human rights norms. Daily challenges and threats to equal development for all determine the framework of HRE in the present and the future. Against this backdrop, to dedicate the fourth UN World Program for HRE to Youth in 2020 has not been a coincidence. It targets those peers that are mostly affected by the challenges.

In this *glocalized* and interconnected world, HRD and governments must address three broad goals: how human rights awareness and knowledge firstly, (1) can first detect the root causes of human rights violations that fuel a fearful, weak, and socially dysfunctional society. Secondly, (2) the need for a legal and political framework to foster human rights and empowerment to take action, to emancipate and build confidence among people? Thirdly, (3) collaboration between duty-bearers and right-holders, governments, businesses, and CSOs, to effectively protect people from violation? Organizing and financing safe houses for women victims of domestic violence, promoting police training, and protecting people from genocidal violence can best be realized in collaboration among different stakeholders. HRE can properly equip scholars, learners, students, and people to tackle these issues and find sustainable solutions (Fischlin and Nadorfy, 2007).

¹²UN ICCPR (1966) International Covenant on Civil and Political Rights Adopted and opened for signature, ratification, and accession by General Assembly resolution 2200A (XXI) of December 16, 1966 entry into force March 23, 1976.

Conclusions

Think global, act locally—*be glocal*! Since the end of the Cold War in 1991 and the World Action Plan for Human Rights in 1993, human rights norms have undergone breathtaking dissemination and glocalization process. A multi-stakeholder governance approach, in which CSOs, NGOs, private enterprises, and governments have primarily cooperated to enhance human rights compliance, has led to a “natural alliance” between the global norms and local actions. Today, global human rights can be best seen in a certain level of glocality, namely implementing them in day-to-day life by HRD, local authorities, and CSOs while adhering to global principles of equal treatment, fair trials, free elections, inclusive schooling etc. Hence Glocalizing Human Rights is a new analytical framework to contextualize global human rights. Glocality, and the process of bringing global norms to local stakeholders, hence glocalization, thanks to ICT, volunteers, educators, and international mechanisms, give practical guidance to everyone, no matter where.

Globally, we see networks of human rights cities and human rights educator networks such as HREA, with over 20,000 educators involved in over 140 countries. Human Rights Campuses and programs are mushrooming, not decreasing. The Global Campus for Human Rights in Venice, Italy, and subsequent regional MA in Thailand for Southeast Asia, in Buenos Aires for Latin America, in Johannesburg for Africa, in Yerevan for the Caucasus, in Sarajevo for the Balkan, in Beirut for the Middle East and Venice for Europe; and the plethora of academic programs for human rights in law, social work and sciences, international relation and public policy, around the worlds. In addition, we count today millions of local, school projects, activities and programs, human rights days and defender days, awards, celebrations, anniversaries, commemorials, and others, that have mushroomed over the past decades—hundreds of Think Tanks working on global justice locally. Human Rights Research Centers, Schools, and Universities for Peace in Costa Rica, Tokyo, or Helsinki set academic and normative standards further to promote the understanding and implementation of human rights. Most of these initiatives can be traced back to the massive endeavors to promote HRE on higher and primary education in the formal and informal sectors after 1994 to the present times. It has since developed its dynamics, as shown in Joseph Zajda’s edition on Human Rights Education Globally (Joseph, 2020).

The fact and experience that the glocal realization of human rights takes time because our understanding best progress when we experience human rights both in a positive and also negative way, by experiencing exclusion and discrimination or violence, but also by being protected from it through equal opportunity acts, gender policies, or winning lawsuits over a property against corporations. Even in discomfort and as an experience of injustice and violation of our freedoms, we learn about human rights. This can only be achieved locally and in person, where concrete examples of violation, abuse, and discomfort can be discussed. Nevertheless, human rights-related rhetoric, movies, clips on YouTube and Instagram, campaigns, and actions raise awareness and offer solutions. Hence promoters of human rights have to be sensitive and reasonable in establishing trust and being compassionate. Indeed, we need to move toward a “strategic empathy,” which is an admittedly long and challenging task that needs the entire continuum and years of training and teaching. This is because the mere understanding of human rights is not enough, as emphasized by the UN Declaration for HRE and Learning in 2011. Scholars of human rights will become more susceptible to affective transformation when they enact compassionate action early on in their lives, starting with simple things such as learning to be more patient and tolerant with peers who do not grasp a “difficult” concept in language or mathematics, as Nancy Flowers, a teacher, schoolmaster and one of the earliest pioneers for HRE, has summarized (Flowers, 2015, p. 12).

References

- Brysk, A., 2018. *The Future of Human Rights*. Polity Press.
- Christopher, N., Roberts, J., 2015. *The Contentious History of the International Bill of Human Rights*. Cambridge University Press.
- Coysh, J., 2018. *Human Rights Education and the Politics of Knowledge*. Routledge Research in Human Rights Law.
- Dalsgaard, C., Paulsen, M.F., 2009. Transparency in cooperative online education. *Int. Rev. Res. Open Distance Learn.* 10 (3), 1–13. <https://doi.org/10.19173/irrodl.v10i3.671>.
- Fischlin, D., Nadorfy, M., 2007. *The Concise Guide to Global Human Rights*. Black Rose Books.
- Flowers, N., 2015. The global movement for human rights education. In: *Radical Teacher, Journal on Theory and Practice of Teaching*.
- Joseph, Z. (Ed.), 2020. *Human Rights Education Globally*. Springer Publisher.
- Macedo, S., 2006. *Universal Jurisdiction: National Courts and the Prosecution*. University of Pennsylvania Press.
- Mahler, C., Mihr, A., Toivanen, R. (Eds.), 2009. *The UN-Decade for Human Rights Education and the Inclusion of National Minorities*. Peter Lang Verlag, Frankfurt/Berlin.
- Mihr, A., Gibney, M. (Eds.), 2014. *Handbook of Human Rights*, vol. 2. SAGE Publication, London.
- Mihr, A., 2009. Human rights awareness, education and democratization: the challenge for the 21st century. *J. Hum. Right.* 8 (2), 177–189.
- Mihr, A., 2010. Human rights education. In: Denemark, R.E. (Ed.), *The International Studies Compendium Project*. Wiley & Blackwell Publishing, Oxford University Press, pp. 3439–3456.
- Mihr, A., 2015. Why holocaust education is not always human rights education. *J. Hum. Right.* 16 (4), 525–544.
- Mihr, A., 2019. “Glocal” public policy in times of global migration. In: Grimm, H.M. (Ed.), *Public Policy in the Global South*. Springer Publishing VS, pp. 43–66.
- Moses, D., Duranti, M., Burke, R. (Eds.), 2020. *Decolonization, Self-Determination, and the Rise of Global Human Rights Politics*. Cambridge University Press.
- Oberleitner, G., 2013. *Global Human Rights Institutions*. Wiley Publisher.
- Pereira, W., 1997. *Inhuman Rights: The Western System and Global Human Rights Abuse*. Other Indian Press.
- Steven, L., Jensen, B., 2016. *The Making of International Human Rights, the 1960s, Decolonization and the Reconstruction of Global Values*. Cambridge University Press.
- Sundrijo, D.A., 2020. *Regionalizing Global Human Rights Norms in Southeast Asia*. Springer International Publisher.
- Tibbitts, F., 2020. Deliberative democratic decision making, universal values, and cultural pluralism: a proposed contribution to the prevention of violent extremism through education. *Prospects* 48 (1), 79–94.
- Tsutsui, K., 2018. *Rights Make Might, Global Human Rights and Minority Social Movements in Japan*. Oxford University Press.

UN Declaration on Human Rights Education and Training, 2011. GA, December 2011. <https://www.ohchr.org/en/issues/education/training/pages/undhreducationtraining.aspx>.
UN OCHCHR, General Comments, 1999. CESCR General Comment No. 13: The Right to Education (Art. 13) Adopted at the Twenty-First Session of the Committee on Economic, Social and Cultural Rights, on 8 December 1999 (Contained in Document E/C.12/1999/10).

Relevant websites

Human Rights Education Associates (HREA), <https://hrea.org/>.
International Justice Resource Center: Regional Systems, <https://ijrcenter.org/regional/>.
OHCHR Database, <https://www.ohchr.org/EN/HRBodies/Pages/TreatyBodies.aspx>.
UN OHCHR, <https://www.un.org/en/about-us/udhr/foundation-of-international-human-rights-law>.
World Program (document A/HRC/27/28). Available at <http://www.ohchr.org/EN/Issues/Education/Training/WPHRE/ThirdPhase/Pages/ThirdPhaseIndex.aspx>.

The light in the distance: global democracy and humanity's hope for the future

Ruthanne Kurth-Schai, Educational Studies, Macalester College, Saint Paul, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

The peril and promise of global democracy	79
Divided in darkness: the crisis of neoliberal globalization	80
Neoliberalism as justificatory social vision	80
Consequences of compliance	81
Returning to the wellspring of democracy: politically engaged holism	82
Toward a transformative social philosophy	83
The path toward voluntary embrace	84
Evolving democracy: re-envisioning civic life	84
Cultivating the social-self: re-envisioning civic learning	85
Clarity	86
Communion	86
Creativity	86
Courage	87
Compassion	87
Inspiring principled action: re-envisioning civic pedagogy and practice	87
Pedagogies of enlightenment	88
Insight	88
Inspiration	88
Pedagogies of empowerment	89
Social inquiry	89
Social advocacy	90
Moving toward the light: our call to action	90
References	91

The peril and promise of global democracy

(W)e have taken democracy for granted ... it has to be enacted anew in every generation, in every year and day, in the living relations of person to person in all social forms and institutions.

Dewey (1937a, pp. 472–474).

The struggle for democracy has to be maintained on as many fronts as culture has aspects: political, economic, international, educational, scientific and artistic, religious.

Dewey (1939a, p. 187).

Throughout the 20th century, John Dewey heralded the peril and promise of democracy, calling each generation to accept a formidable task. He noted that democracy is always fragile. In every time, in every place, in every social and cultural context, it is necessary for all people to participate in processes of democratic reform and renewal. He further asserted that democratic renewal requires reimagining and revitalizing all aspects of culture, all forms of association, all public processes and institutions—most importantly mass state-supported education. In this exploratory essay, informed and inspired by diverse perspectives across time and cultures, I advocate response to Dewey's call, moving first to confront the challenges of our time and then to chart a path forward.

We live in dark times. Sustained exposure to greed, corruption, inequality, intolerance, violence and environmental degradation has taken its toll. Troubled by deepening divides within and among nations, I argue that the constellation of humanitarian and environmental crises we now face is compelled and intensified by the imposition of neoliberalism as the dominant worldview. A justificatory social vision emphasizing consumption, competition, fear and compliance; neoliberalism breeds oppression, division, delusion and despair. We retreat from the natural world and from each other, from unfamiliar ideas and experiences, from our hopes and our dreams. We seek certainty and control as if this will keep us safe. We abdicate agency—our capacity and responsibility to think, feel and act on our own and others' behalf. Global democracy is threatened; mass state-supported education is imperiled.

And yet there is hope. The challenges before us, though daunting, are not unprecedented. Inspired by philosophers, teachers, community and spiritual leaders representing diverse contexts and perspectives, I trace the development of a socially transformative

worldview that I will refer to as 'politically engaged holism'. In response to seemingly insurmountable social and environmental crises, reform advocates have worked tirelessly to compel collective experience of a unifying ethical vision. In sharp contrast to neoliberalism, politically engaged holism emphasizes clarity, compassion, creativity and courage; thereby promoting connection, collaboration, integrity and hope. Embracing this vision, we find guidance necessary to restore the promise of global democracy—to extend and evolve its processes and institutions, to reimagine and revitalize our aspirations for, and our approaches to, civic learning and life. This is our task. Herein lies humanity's hope for the future, our collective path from the darkness into the light.

Divided in darkness: the crisis of neoliberal globalization

Neoliberalism is a philosophy in which ... the operation of a market or market like structure is seen as an ethic in itself, capable of acting as a guide for all human action, and substituting for all previously existing ethical beliefs.

Treanor (2005).

The great wave of globalization sweeping contemporary society ... is a contrast of light and dark ... the negative aspects are war and conflict, rising economic disparities ... destruction of the global ecology. These dark shadows ... stir a vortex of malice and mistrust and provoke an identity crisis in the very depths of the human spirit.

Ikeda (2005, p. ix).

Grounded in classical liberalism and neo-classical economics, emanating from US based global institutions including the IMF and the World Bank, then embraced by power brokers across the globe; neoliberalism has become the transformative ideological paradigm of our time. This comprehensive worldview now dominates social, political, economic and educational landscapes, driving understanding of and response to systemic social and environmental crises (Comaroff and Comaroff, 2001; Dardot and Laval, 2013; Olssen et al., 2004; Shah, 2010; Treanor, 2005).

Pervasive acceptance of global capitalism permeates contemporary public discourse, policy and practice. Priorities centered on acquisition and security define and drive nearly all aspects of human identity and interaction. From commerce and entertainment, to civic discourse and interpersonal relationships, to governance and schooling; pressures toward increasing competition, consumption, privatization, compliance and inequality shape our lives (Gabbard, 2008; Giroux, 2004; Harvey, 2005; Stromquist, 2002). Justifications for public policy and practice are framed within a logic of inevitability. As noted by Michael Apple, neoliberal policies reflect an emerging global consensus, one that defies justification having acquired "something of a sacred aura, especially since we are repeatedly told there are *no* alternatives worth considering" (2006, p. 15).

Though framed as common sense or "just the way things are", neoliberalism entails powerful assumptions regarding pivotal philosophical questions: What is the nature of reality? What does it mean to be human? What is valued knowledge and how is it acquired? What does it mean to live an ethical life? Neoliberal responses to such questions reflect unacknowledged ideological commitments; self-perpetuating and mutually reinforcing assumptions that constrain social vision and restrict social action.

Neoliberalism as justificatory social vision

Neoliberalism as a social philosophy is characterized by faith in economic determinism, acquisitive individualism, entitlement ethics and meritocracy. Its central tenets are explicitly materialistic. Economic determinism, granted status similar to natural law, is accepted as the fundamental force shaping learning and life. To be human is to experience, to interpret, and to act within materially constrained social and physical worlds.

The individual is the central focus of human development, taking precedence over collective social entities. Humans are characterized as essentially and necessarily self-interested and competitive. Individuals act as autonomous agents exercising rational choice across transactional systems at all levels of experience from interpersonal relationships and family life, to public service provisions and consumer transactions, to national and global political and economic systems. Personal development is advanced through accumulation of material goods—understood as wealth—that is exchanged in free, impartial and competitive markets. Acquisitive individualism is required for survival and equated with personal virtue. Consequently, retreat from efforts aimed at advancing the "common good" is socially sanctioned.

Neoliberalism's emphasis on acquisitive individualism is reflected in its commitment to the ethics of entitlement. Ethical treatment is centered on protection of individual rights to acquire property and to seek opportunity defined largely in economic terms. Justice is defined as access to a "level playing field" and to fair and impartial "rules of the game." Worthiness and reward, framed as human capital, are distributed in response to personal merit determined in relation to social hierarchies mirroring pervasive economic structures.

To live an ethical life is to act as a conscientious decision-maker, a well-informed and skilled consumer. Core personal values include responsibility, industriousness, competitiveness, realism and efficiency. Personal responsibility entails being accountable for one's situation and reliable when relating to others. Industriousness combines personal ambition with a strong work ethic engaged in economic-political-social contexts that are fundamentally competitive. To be realistic means to conform to prevailing

materialistic assumptions and to subordinate all non-material aspects of human experience. To be efficient means to rely almost exclusively on economic prescriptions (e.g., “more for less”, “cost/benefits analysis”) in navigating a moral path through life.

Society is considered ethical to the extent that its institutions protect freedom of choice across individual and collective transactions. Free markets provide the mechanism for individual gain and social progress while governance institutions manage stable exchange systems with minimal interference. Order and security are maintained through adherence to the rule of law. In exchange for compliance with legal authority, citizens are entitled to protection (e.g., civil rights, limitations on government intrusion) and to public services not provided by the private sector.

Society is organized on the basis of what are accepted as natural and necessary hierarchies responsive to differences in individual merit. Entrepreneurial risk-taking is rewarded by monetary gain and widely regarded as a definitive indicator of social superiority. Wealth is asymmetrically distributed and correlated with power. Its directives and justifications shape not only the business sector, but also all levels of politics and governance, of public and private education, of high and popular culture including the arts, recreation and sports.

Political and economic elites define truth, determine priorities, impose standards and regulate public processes. Significant differences between people, groups and nations are assessed in relation to economic disparities. It is understood that inevitable inequalities in resource distribution will determine differences in quality of life, rather than barriers to access or opportunity associated with race, religion, gender or ethnicity. Gaps between the rich and the poor widen and deepen over time.

Engulfed in this pervasive, largely unacknowledged yet determinant social vision, the direct and often dire implications for human development and interaction—for social and environmental sustainability—are rarely articulated. Beyond the immediate political and economic consequences of neoliberalism are profound implications for what it means to survive, to live and to learn as fully engaged and evolving people; to participate in governance, civic life and mass state-supported education.

Consequences of compliance

Among the consequences of global compliance with neoliberal philosophy, policy and practice are increased energy consumption reliant on fossil fuels, large scale corporate agriculture/mining/deforestation, urbanization, industrialization, population growth and dislocation. Environmental systems begin to fail resulting in a cascade of dire consequences—global warming, ozone layer depletion, loss of biodiversity, pollution, lack of arable land and fresh water, intensification of natural disasters including global pandemics.

Social and cultural crises are exacerbated as competition for scarce, non-renewable resources intensifies. Expressions of white supremacy, neo-colonialism, political and religious intolerance; dramatic economic inequality and associated loss of political representation; human rights violations, violent conflict resolution, starvation, unmitigated disease and dislocation; deepen and expand.

Tragically, the imposition of neoliberal culture renders individuals, voluntary collectives and social institutions woefully unprepared to address systemic, mutually reinforcing humanitarian and environmental crises. Although actively promoted as a culture of unbridled opportunity, neoliberalism in reality is experienced as a culture of fear and constraint. Across all social domains, global capitalism diminishes human experience, development and potential by amplifying proclivities toward self-centeredness, greed, aggression, insecurity and demoralization. As opportunities for personal intellectual, social, emotional and ethical development are diminished, so too is the efficacy of primary social institutions—most significantly, systems of governance and mass state-supported education (Casey, 2016; Giroux, 2004, 2008).

Democracy as a form of governance, while always difficult to sustain, is uniquely threatened during times of deep material and moral challenge. When confronted with intertwined environmental and social crises, typical societal responses feature diminishing equity aspirations, seeking certainty and control, and persisting along familiar policy paths even though these may not have proved effective in the past. As political processes are corrupted via access to wealth and power, democracy is reduced to the illusion of representation masking underlying totalitarian tendencies. Both established and emerging democracies are threatened not only by authoritarian regimes beyond their borders, but also by extremist populist movements within. Relentless disinformation campaigns, voter suppression, civil rights violations, restrictions on free press and political protest fuel dangerous levels of distrust, aggression, political tribalism and legislative impasse (Applebaum, 2020; Diamond, 2008; Snyder, 2018; Ziblatt and Levitsky, 2018).

Dependent upon full participation of diverse, civically educated and engaged populations, democracies are further threatened by demonstrable declines in the quality of education provided for the masses. State-supported school systems are confronted with a convergent constellation of pressures emerging from increasingly centralized and bureaucratized government policies (e.g., restrictive ministry directives, NCLB) and global mandates (e.g., IMF/World Bank), aggressively marketed entrepreneurial initiatives (often associated with emerging educational technologies), powerful special interest lobbies (e.g., test preparation and curriculum development/textbook publishing) and persistent efforts to privatize public educational systems. Instructional strategies are crafted to ensure conceptual control and ease of replication. Curriculum is narrowed to support strong performance on standardized tests in addition to promoting proficiency in subjects contributing directly to competitive state economies (Apple et al., 2005; Apple, 2006; Burch, 2009; Casey, 2016; Hyslop-Margison and Sears, 2006; Lipman, 2004; Porfilio and Malott, 2008; Suarez-Orozco and Qin-Hilliard, 2004; Ward, 2012).

Of particular concern to democratic governments is the decline of civic education. Documented loss of time and attention is accompanied by efforts to propagate inaccurate and socially unjust historical representations that perpetuate state-sanctioned, though unacknowledged, systems of social oppression (e.g., colonialism, white supremacy, racism, patriarchy, heteronormativity).

Decline is further intensified by efforts to promote shallow representations democratic participation limited to acquiring basic understanding of the functions and structures of government accompanied by episodic voting. Capacities critical to sustaining vibrant democracy—social empathy, inquiry, critique and imagination—are neither recognized nor supported; in many cases, actively discouraged. Aspirations for governance and state-supported education are conceptually and strategically narrowed in response to changing demographics, financial crises, civil unrest, relentless natural disasters and deeply polarized politics. Retreat from deep, systemic, democratic reform is repeatedly reinforced (Allen and Reich, 2013; Banks, 2021; Barrow, 2017; Heybach and Sheffield, 2014; Ladson-Billings, 2004; Noddings, 2013; Tarozzi and Torres, 2016; West, 2004).

Overall, neoliberal culture constrains expression of human intelligence and integrity thereby threatening global democracy. Any social vision that discourages social inquiry and civic participation, thereby diminishing the overall quality of human life, can neither be ignored, nor unwittingly perpetuated. There must be another way to respond to deep crisis; another way to orient ourselves to our world and to each other. In response to the significant challenges posed by neoliberal globalization, careful and continuing enactment of a contrasting social philosophy, an alternative justificatory social vision, is required.

Returning to the wellspring of democracy: politically engaged holism

The spirit of democracy is not a mechanical thing to be adjusted by abolitions of forms. It requires a change of heart.

Mahatma Gandhi.

Democracy is a form of government only because it is a form of moral and spiritual association.

Dewey (1888, p. 18).

The challenges before us, though daunting, are not unprecedented. In response to severe social and environmental crises, in the face of seemingly insurmountable odds, reform advocates across time and cultures have worked tirelessly to create, and then to compel, collective experience of a unifying ethical vision. Of particular importance to prospects for global democracy are two streams of imagination and activism reflected in the words and actions of “spiritually engaged Progressives” and “socially engaged Buddhists”. Although both Progressivism and Buddhism encompass diverse yet interconnected perspectives, each originating and then developing within dramatically different social contexts, complimentary aspirations for and approaches to social and educational reform are evidenced in distinctive streams emergent within each tradition.

Informed by original works and contemporary interpretations, we find streams of Progressive thought signaling that spiritual development is necessary for full expression of human potential and therefore, education of the masses must acknowledge and attend to a spiritual dimension of learning and life. Drawing from both Western and Asian interpretations of classical Buddhist texts and lay Buddhist social movements, we find Buddhist reform traditions embracing the belief that spiritual development is promoted not through distancing oneself from the trials of human life, but instead by working to alleviate personal and collective suffering on a day by day, moment by moment, basis.

Glimpses of a spiritual dimension consistent with Progressive philosophy are expressed by proponents of *Pragmatism* (Dewey, 1934; Richardson, 2010; Garrison, 1997, 2010; Hickman, 2007; Kesson, 2003; Unger, 2007), *Critical Theory* (Freire, 2000; Greene, 1995; Purpel, 2004b; Shapiro, 2009), *Feminism* (Hooks, 2003), *Postmodernism* (Carlson, 2002; Oldenski and Carlson, 2002), and the *Black Radical Tradition* (Du Bois, 2019; McCluskey, 2014); in addition to teacher/scholars advocating *Indigenous/Anti-colonialist* (Cajete, 1994; Deloria, 1979; Sefa Dei, 2011), *Ecological Justice/Sustainability* (Riley-Taylor, 2002), *Montessori* (Standing, 1957) and *Waldorf* (Steiner, 2008) education. As noted by Thomas Oldenski and Dennis Carlson, “Progressive conceptions of spirituality are about the *transcendence* of human consciousness and culture ... the journey inward leads outward again, with a renewed sense of engagement with life and a sense of being reconnected to others (2002, p. 5).” The call to attend to the spiritual development of all learners is carried through the 20th century and on into present, gaining renewed attention in relation to systemic crises (over-consumption, violent conflict resolution, religious and political fundamentalism, environmental destruction) deemed irresolvable within the philosophical frameworks dominating contemporary schools and society.

Concurrently, social activism consistent with Buddhist reform traditions—including Nichiren, Zen and Tibetan streams of Mahayana and Thai Forest and Vipassana streams of Theravada—is expressed through teachings of clergy and lay practitioners throughout Asia. Included are Tsunesaburo Makiguchi (Bethel, 1973), Daisaku Ikeda (2010a, b), Thich Nhat Hanh (Ellsberg, 2001), Aung San Suu Kyi (2010), Somdech Prea Maha Ghosananda and A. T. Ariyaratne (King, 2006), Bhikkhu Buddhadasa (2005), Sulak Sivaraksa (2005), and His Holiness the 14th Dalai Lama, Tenzin Gyatso (1999). These teachings, responsive to contexts shaped by war, political oppression, mass dislocation and natural disaster, are enacted as social movements referred to as “engaged Buddhism”. In the words of the Dalai Lama, “the purpose of life is not to transcend the body, but instead to embody the transcendent.”

Emerging within tragic social contexts, spiritually engaged Progressives and socially engaged Buddhists have charted parallel paths to guide collective movement. Rather than submitting to imposed constraint and control; rather than accepting an impoverished view of humanity and the natural world; rather than lapsing into fear, anger, cynicism, apathy and despair; philosopher/scholar/teacher/activists representing both traditions have attempted to educate the masses to confront social crises with creativity

and courage. They have centered their hope in an expansive assessment of human potential, united in the belief that as humans are complexly unique and equally worthy. They have framed their parallel visions as a new “way of being”—an existential construct asserting dynamic integration of purpose and process, means and ends, path and destination; an inclusive, life affirming and sustainable global worldview; a social philosophy and political framework that I will refer to as “politically engaged holism”—the “change of heart” required to reclaim, to reimagine and revitalize global democracy.

Toward a transformative social philosophy

Looking beyond neoliberal assumptions regarding the primacy of materialism are visions of a multifaceted reality that holds the “material” and the “spiritual” in dynamic balance. Spirit is defined not as a distinct or separate dimension, but rather as a fundamental attribute that infuses and radiates from all entities; an exploratory, integrative, inherently idiosyncratic and unpredictable yet unifying force. Spirit is further defined not as a religious doctrine or prescribed set of practices, but instead as an orientation aimed at revealing ultimate meaning to be expressed as fidelity of high purpose (achievable aspiration) and comprehensive action (transformation). Though understandably controversial in societies embracing either secularism or religious fundamentalism, recognition and cultivation of a spiritual dimension provides a generative frame for guiding learning and life in a fiercely materialistic, competitive, cynical and divisive neoliberal age.

The nature of reality, and of humanity, is further perceived as inherently relational and profoundly social; everything and everyone are joined in webs of complex, ever evolving relationship. People exist, interact and learn as social selves—actively drawn to sustain relationships within (integration of multidimensional and integrative personal identities); relationships among (engagement on a person-to-person basis, building connections that build community); and relationships that move beyond (connections with broader, at times newly emergent, social and environmental systems and forces). Humans are difficult to limit or contain. They are necessarily curious, creative, experimental, altruistic; compelled to imagine and enact more expansive and more powerfully life-affirming ways of living and learning. Humans are empathetic and pragmatically collaborative. Although imperfect, still they possess unlimited potential to create and to express all forms goodness. Resources are received and distinctive contributions developed in light of insightful analyses of personal and communal needs and aspirations. In response to human efforts to initiate and sustain purposeful, diverse, multidimensional relationships; deep learning and deep change occur.

In sharp contrast to neoliberal assumptions asserting a profoundly instrumental appraisal of socially sanctioned knowledge and values, are aspirations for widespread participation in processes of social inquiry aimed at experiencing and evolving wisdom. Wisdom is relational—at once intensely personal and essentially social. Wisdom is multisensory, interdisciplinary, multi-paradigmatic—effectively connective across physical, temporal, emotional, intellectual, ethical, political and spiritual domains. Wisdom is exploratory—open, divergent, creative, prophetic, surprising. Wisdom is integrative—synthesizing efforts toward truth, beauty and justice thereby extending unrestricted and unlimited access to learning experiences of highest significance and deepest meaning. As such, wisdom is experienced and evolved as a gift, rather than as a commodity; valued not for short-term utility approached through narrowly rational means/ends strategies, but cherished instead for the promise of time-less possibility, for access to significant challenge and opportunity that yields transformative response on personal and societal levels.

Wisdom is acquired by means of fully relational, exploratory and integrative processes. Teaching/learning takes shape as an ascending spiral moving from active preparation (seeking, questioning, expanding receptivity while focusing intent), to inspiration (attending to experiences of intellectual, emotional, social, ethical and esthetic resonance), to meaning-making (reflecting, interpreting, grounding knowledge of inspirational quality in order to fully comprehend its real world significance), to sharing (translating, living, giving one's knowledge in a manner that enhances the quality of life, the quality of relationships among self and others—this, the very essence not only of teaching/learning, but also of principled social action).

A social ethic refusing to accept the ideological constraints of neoliberalism emphasizes mobilization of communal resources to protect personal freedom to be included, to create and to collaborate; over freedom to segregate, to compete and to acquire. Relationships among diverse individuals, within voluntary collectives, within social institutions, are characterized by expressions of mature interdependence. Diversity is necessarily maintained within any and all expressions of social unity. Social order and security, states reflective of enduring aspirations rather than paralyzing fear, are reliant on expressions of equity—collective creativity blending justice with compassion. Social prosperity is measured in terms of the extent to which public policy is experienced as responsibly empathetic and evaluated as demonstrably effective in enhancing collective well-being.

Released from the grips neoliberalism, a contrasting vision of personal ethics provides the foundation for movement toward new ways of living and learning in the world. Beyond an ethic assigning personal responsibility for pursuit of personal gain restricted only by principles of non-interference and market driven fairness/impartiality; emerge possibilities for an individual ethic aimed at “social-self realization” (Brameld, 2000) whereby capacities for self-cultivation and social contribution are continually expanded, deepened and integrated. Although a highly individualized process, moral development—the ascent of human spirit—is equally and necessarily a cultural and communal endeavor (Adorjan and Kelly, 2008; Doll and Gough, 2006; Eppert et al., 2015; Garrison et al., 2014; Hansen, 2007; Jackson, 2003; Kimmerer, 2013; King, 2006; Kurth-Schai, 1992; Kurth-Schai and Green, 1997; Maslow, 1971; Mason, 2008; Miller, 2019; Reagan, 1996; Rogers, 2009; Scott, 2000; Standish and Naoko, 2012).

The path toward voluntary embrace

Both spiritually engaged Progressives and socially engaged Buddhists understood that full realization of politically engaged holism as a pervasive justificatory social vision could only be achieved and sustained through persistently experiential, experimental and relational processes; through complex pedagogies grounded in creative response to the challenges of daily life, yet radiating outward to shape principled action in response to emerging crises on a global and planetary scale. Orchestration of such processes would only be possible within dynamic social contexts steadfastly engaged in reimagining and revitalizing democratic governance, civic participation and mass education.

Evolving democracy: re-envisioning civic life

Democracy is not a state. It is an act, and each generation must do its part to help build the Beloved Community, a nation and world society at peace with itself.

Lewis (2020).

Democracy is more than a form of government; it is primarily a mode of associated living (1916, p. 93) ... a way of life controlled by a working faith in the possibilities of human nature ... the task of democracy is forever the creation of a freer and more humane experience to which all contribute.

Dewey (1939b, pp. 341–343).

In search of a vision of democracy aligned with politically engaged holism, we turn to the writings of John Dewey, reinforced and extended by contemporary democratic theorists and practitioners (Ansell, 2011; Burch, 2000; Freire, 1998; Green, 2008b; Gutmann and Thompson, 2009; Hytten, 2009; Katz et al., 2009; Lummis, 1996; Palmer, 2011; Purpel, 2004a; Noddings, 2013; Wang, 2009; Westbrook, 2005).

A social philosopher in the American Pragmatist tradition, Dewey asserted that the fundamental purpose of philosophy is to afford all people opportunities to pursue life with dignity, affiliation and an ever-evolving sense of purpose and possibility. To this end, he re-envisioned philosophy as a moral process, to be tested in action, requiring continual reconstruction to remain responsive to changing contexts of time, place and culture (1919, 1920).

Dewey's philosophy was grounded in an evolutionary form of naturalism; a religious form of humanism that assumes an active, inquiring, esthetic, immediate, inspiring, transformative and "intimate relation with existence wherein our creative acts matter in the course of creating cosmos from chaos" (Garrison, 2008, p. 6); see also Garrison et al. (2014), Hickman (2007, 2008). Dewey described his spirituality as a common faith (1934); a metaphysics of human relationship grounded in lived experience with nature and with each other. This process-oriented metaphysics would become particularly important in times of deep crisis (Green, 2008a).

In 1916, Dewey identified *Education and Democracy* as the entwined perennial path toward enacting his philosophy across all dimensions of human activity. His conceptualization of and commitment to democracy as a creative, evolving approach to learning and living continued over the course of his lifetime, culminating with the writing of two pivotal essays, *Democracy is Radical* (1937b) and *Creative Democracy—The Task Before Us* (1939b).

Though aspirational, Dewey's vision, his faith, was not naive. He understood that the "end of democracy is a radical end" that could only be achieved through radical means (1937b, pp. 338–339). He knew that it was necessary to move beyond conventional enactments of democracy as a form of governance that limits personal freedom by promoting majority rule at the expense of minority perspectives, or restricts interaction for the sake of the public security. Instead, Dewey envisioned creative democracy as an expansive and experimental process through which the needs and aspirations of diverse individuals could be fulfilled within the context sustained communication and social contribution (1916, 1927, 1939b).

Dewey's path toward creative democracy was centered in collective experience of a radically social, experimental, exploratory, ethical and esthetic approach to interaction and inquiry—responsive to "belief in the ability of human experience to generate the aims and methods by which further experience will grow in ordered richness" (1939b, p. 343).

He argued that the experience of creative democracy is radically social in that it demonstrates the "power of voluntary action based upon public collective intelligence" (Dewey, 1937b, p. 339). Public collective intelligence is achieved through complexly inclusive and collaborative deliberation that accommodates free expression and full consideration of numerous and varied positions (1916).

The experience of creative democracy is radically experimental in that the multiplicity of perspectives on societal crises—their causes, consequences and possible solutions—are accepted as "hypothetical, demanding trial in terms of social action" (Dewey, 1928, p. 314). The experience becomes radically exploratory as the knowledge acquired through social experimentation is then subjected to human "need and desire—out of which grow purpose and direction of energy ... (that) continually open the way into the unexplored and unattained future" (1939b, p. 343).

The experience of creative democracy is radically ethical and esthetic as these critical dimensions of learning and life are engaged to ensure that collective action is undertaken with integrity (Dewey, 1927, 1932). Deeply felt personal and communal values play

a pivotal role in shaping public policy and practice. Rather than accepting moral or esthetic judgments derived from a transcendent reality, tradition, social convention, self- or special group interest; collective action is to be evaluated experimentally, assessing the harmonization of aims/means/ends and the capacity to generate what Dewey referred to as “contributory efficacies” (1922, p. 273). Esthetic and ethical expressions further serve as sources of inspiration. In response to profound challenges of human existence, artistic and ethical expressions affect and sustain public imagination and determination as needed to support continuing civic engagement.

Perhaps most important, Dewey understood that creative democracy is always vulnerable. Among the greatest threats is complacency; unreflective or resigned passivity that renders processes of communal decision-making and civic action irrelevant and ineffective. In order to remain vibrant, democracy must continually evolve. Each time/culture/location offers distinctive challenges that must be confronted with focused intent, wisdom and strategic response. Rather than depending on political institutions or legislative mandates, it is the continuing responsibility of all people, in all aspects of their lives, to participate in sustained, experimental and pluralistic interactions that support diverse expressions of openness, humility, cooperation and willingness to grapple with uncertainty and ambiguity.

In light of continuing threats to democracy held captive by the crippling effects of neoliberal globalization, drawing from Dewey and related contemporary democratic theorists/activists, I propose that collective experiences of democratic civic life be designed and assessed in relation to the following performance-based aspirations. Such experiences can be developed socially and evaluated experimentally, repeatedly striving to answer the question: In what manner, and to what extent, are the following criteria realized both individually and collectively?

- Shared Inspirational Purpose
 - clarifying the primary purposes to be advanced in relation to clearly articulated and widely embraced philosophical assumptions and values, in addition to ethically and esthetically compelling aspirations
- Free Expression
 - giving voice to all citizens, especially those whose access to power is limited
- Full Consideration
 - ensuring free and full participation of all participants, with special attention devoted to the deliberative barriers posed for historically disadvantaged populations
 - ensuring unbiased analysis across power differences and social divides
- Enriched Relationships
 - increasing the depth and variety of connections among diverse participants
- Multi-Dimensional Learning
 - engaging varied dimensions of human experience and understanding; intellectual, emotional, intuitive, ethical, esthetic, sensory, spiritual
- Creative Insight
 - challenging assumptions of certainty; inspiring new, often surprising, ideas
- Skilled Social Analysis & Imagination
 - encouraging movement beyond personal experience to deepen analysis and imagination regarding social systems (education, government, economy) and social forces (racism, colonialism, classism, patriarchy, political or religious intolerance, etc.)
- Authentic Consensus
 - facilitating group commitment to a common cause while sustaining respectful attention to diverse perspectives and experiences
- Enhanced Social Efficacy
 - translating collective goals into principled and effective social action
 - supporting sustained engagement in principled and effective social action

Cultivating the social-self: re-envisioning civic learning

The task of education must be to stimulate and unleash the wisdom that lies dormant in the lives of all young people ... This is not a forced process ... but rather drawing out the potential that exists within. I firmly believe that every young person has the power ... to change the world.

Ikeda (2010b, p. 151).

The *problem* of education ... is one with the *problem* of finding out what democracy means ... Democracy means ... a social order in which all the forces that make for friendship, beauty, and knowledge are cherished in order that each individual may become what he(/she), and he(/she) alone, is capable of becoming.

Dewey (1937c, pp. 416–417).

The call to initiate and sustain participation of large and diverse publics in genuinely collaborative, visionary, resourceful and effective personal and social transformation is daunting. Education of the masses, consistent with politically engaged holism, marks the path. Emergent within contexts of crisis, it is a path that seeks to prepare all people, not only to cope with challenges of personal and communal life, but also to contend with systemic change arising in response to large-scale social and environmental forces. It is a path that engages complex knowledge, skills and dispositions in a manner accessible to everyone; a path that must be experienced, refined, expanded and evolved over the course of a lifetime.

Understanding that civic learning toward creative democracy must be secured across all ages and social settings, those engaged in education reform guided by politically engaged holism have found it necessary to re-envision the very purpose of state-supported education. Rather than seeking to prepare all individuals to serve the political and economic needs of the nation-state; state-supported education is redirected toward preparing all people to find meaning, to act with integrity, and therefore to experience genuine hope, even in the face of seemingly insurmountable barriers. Wary of mass education that serves to perpetuate proclivities toward self-centeredness, greed, aggression, insecurity, and demoralization; spiritually engaged Progressives and socially engaged Buddhists have worked diligently to cultivate the distinctive potential of each person to act as a higher self—a socially conscious self—expressing clarity, communion, creativity, courage and compassion throughout all aspects of their lives. Collectively, it is asserted that:

Clarity

All people can be supported in learning to interpret the world with clarity. Clarity is the expression of integrative and illuminating potential. Clarity reveals a comprehensive interpretive frame characterized by synthesis of varied dimensions of truth, of truth and value, of multiple ways of knowing. Truth is understood as complex and profoundly consequential. Although some aspects of truth are perceived to be intrinsic—representative of essence, of found order inherent within the very nature of reality; other aspects are accepted as socially constructed and subject to continually changing conditions and contexts.

The civic function of clarity is to seek, to challenge and continually refine shared truth necessary to guide collective action. Neither universal nor relativistic aspects are to be passively accepted as abstractions, distant from the rigors of daily life. Instead, assertions of truth must be subjected to careful observation, experimentation and verification within a community of inquiry. Yet pursuit of truth alone is not sufficient to ensure clarity. Dimensions of truth must fully engage with dimensions of value, thereby entering the intricate dance of evaluation—a dance that draws from varied ways of knowing (sensory, cognitive, emotional, intuitive, kinesthetic, ethical, esthetic) to form judgments necessary to guide social action in a constantly changing world (Bethel, 1973, 1989; Dewey, 1922, 1932; Ellsberg, 2001; Garrison, 1997; Hickman, 2007; Ikeda, 2004, 2010a; Hanh, 2012).

We educate to cultivate clarity so that we might illuminate the complexities of human existence. For only through clarity are we freed from personal prejudice; from the allure of propaganda, deceit and disinformation; so that we might engage in social inquiry and civic contribution in an open, enlightened and efficacious manner.

Communion

All people can be supported in learning to establish and sustain communion. Communion is the expression of connective and harmonizing potential. Communion defines the quality of relationship necessary to enhance life across all dimensions of existence—interdisciplinary (across fields of inquiry), intrapersonal (within one's self), interpersonal (among self and others), and transpersonal (among self and broader social and natural systems, phenomenon and forces). Although each unique self is inherently expansive, growth can only occur within the context of reciprocal, trusting, generous and supportive yet challenging relationships.

The civic function of communion is to transform the experience of diversity as an enduring threat, to diversity as an integral and generative aspect of collective life. "Diversity within unity" cannot be sustained by abandoning or actively ignoring competing perspectives. Neither can it be sustained by blurring critical differences through forced compromise. Instead, through genuine dialog and empathetic deliberation, individual and special group interests are freely expressed and fully considered making it possible to more genuinely engage with each other, to learn from each other, to synthesize new patterns of response and to experience strength and solidarity when confronted with challenge and change (Bethel, 1989; Lama, 1999; Dewey, 1916, 1927; Garrison, 2010; Garrison et al., 2014; Hickman, 2007; Ikeda, 2004, 2010a,b).

We educate to cultivate communion so that in response to diverse perspectives we might seek common ground rather than resort to conflict; so that we might join together to "continuously recreate an inquiry-guiding, imagination-stirring 'big picture' or 'back-ground map'" of who we collectively are, and who we might become, beyond our differences (Green, 2008a, p. 43).

Creativity

All people can be supported in learning to respond to discord and uncertainty with creativity. Creativity is the expression of poetic and resonant potential. Creativity is a generative process engaging conceptual, material, emotional, ethical and esthetic dimensions. Jim Garrison suggests the essence of creativity is poesis, the act of calling into existence that which has never appeared before. He contends that for Dewey, the very act of inquiry—the pursuit of wisdom necessary to support creative democracy—lies in envisioning "ethereal things", for such imaginings move us to "accept life and experience in all its uncertainty, mystery, doubt ... and turns that experience upon itself to deepen and intensify its own qualities" (1997, quoting Dewey, p. 84). This unlimited poetic potential

exists within us all. As we cultivate it we are transformed, capable of recognizing and responding to each new situation in an innovative, unique and harmonious manner (Garrison, 2009).

The civic potential of a creative act is engaged as its power resonates beyond its source. Positive qualities are amplified and radiated outward through webs of interdependent relationship. Evoking the metaphor of “strong poet”, Ibrahim (2009) notes the depth of eloquence, vision, faith and audacity needed to engage, to understand, to walk through and effectively address conceptual impasses and emotion-laden controversies crossing personal to global domains.

We educate to cultivate creativity so that we might “intervene in the world to ameliorate suffering and discord. Such poetry recognizes the possibilities of experienced situations and overcomes obstacles by creatively transforming them to realize our ethical ideals” (Garrison et al., 2014, pp. 203–204).

Courage

All people can be supported in learning to confront crisis with courage. Courage is the expression of the exploratory and ethical potential required to sustain more deeply democratic approaches to learning and life. As noted by Ikeda (2021), “No matter how wonderful our dreams, how noble our ideals, or how high our hopes, ultimately we need courage to make them a reality.” Those engaged in comprehensive democratic reform will unavoidably face loss of confidence, hope and resolve. Advocates will always operate within contexts of scarcity and injustice. They will be called to push beyond constraints of past action and present resistance, to outperform those in positions of power, to accomplish more with less. Yet the future of global democracy is at risk. Systemic change is neither easy nor safe. It takes courage to enact bolder, more beautiful visions (Ellsberg, 2001; Freire, 1998; Ikeda, 2010a).

The civic function of courage is to engage and sustain social advocacy—principled risk-taking with and for those who are least well served by neoliberal social, political and economic arrangements. The path toward social justice requires mutual trust, shared aspiration and collective imagination; and so we challenge each other to dream. Yet dreaming is dangerous; aspiring to advance the greater good, we face possibilities for even larger loss. Failure is rightfully feared, for expectations—when heightened and then denied—can only deepen disillusionment among those who have already lost too much.

We educate to cultivate courage so that we might encounter civic risk-taking with integrity. To act with integrity is to act on the basis of principle; to hold fast to one’s ideals regardless of circumstance or consequence; to exhibit sensitivity, strength, patience and resilience in the face of adversity; to respond with sincerity, efficacy and honesty across all forms of interaction thereby diminishing dissonance between voiced intent and lived experience. The spirit of democracy is forged in the fires of civic courage and then liberated through persistent, inventive, impactful expressions of civic virtue (Kurth-Schai and Green, 2016).

Compassion

All people can be supported in learning to respond to life’s challenges and opportunities with compassion. Compassion is the expression of empathetic and emancipatory potential. While responsive to individual needs and aspirations, approaches to education aimed at nurturing compassion emphasize each person’s unique capacity and responsibility to contribute acts of courage, respect, caring and solidarity in support of others (Bethel, 1989; Lama, 1999; Hooks, 2003; Ikeda, 2010b; Noddings, 2005; Palmer, 2011). Genuine empathy expressed in action is not paternalistic, passive, unimaginative or self-sacrificial. Compassionate response instead marks the path toward liberation from personal and societal limitations. Responding with sensitivity and generosity promotes personal growth and fulfillment as, “self-realization is reciprocal ... Bestowing value, generating love, creating goodness enhances others while affirming and expanding ourselves” (Garrison, 1997, p. 40). As we practice compassion we acquire wisdom and therefore grow in our capacity for ethical response, not only to like-minded others, but also to those whose life experiences and commitments differ significantly from our own.

We educate to cultivate compassion so that we—as individuals, community members and citizens—might consistently express our highest ideals as intelligent, creative and benevolent response to whatever comes our way. For this is the basis of personal and social transformation—imaginative, substantive, sustained collective effort toward ever more enlightened, more principled, ways of knowing and being in the world.

Inspiring principled action: re-envisioning civic pedagogy and practice

We have to continue to learn. We have to be open. And we have to be ready to release our knowledge in order to come to a higher understanding of reality.

Thich Nhat Hahn.

Knowledge emerges only through invention, and re-invention, through the restless, impatient, continuing, hopeful inquiry human beings pursue in the world, with the world, and with each other.

Paulo Freire (1970, p. 58).

The task of re-envisioning civic learning and life in times of division, distrust and despair can only unfold along challenging paths. As educators and citizens our pressing responsibility, our critical opportunity, is twofold. We must cultivate our personal potentials for clarity, communion, creativity, courage and compassion while working to shape collective experience of these qualities across educational settings, inclusive of all people.

This is of course a daunting challenge considering the depth and extent of change required on individual, classroom, community and societal levels. It is in relation to this challenge that pedagogy and practice consistent with politically engaged holism holds promise. Compelled by ethical and practical imperatives to ameliorate pressing social and environmental crises, pedagogical approaches are developed, evaluated and evolved across complementary Progressive and Buddhist perspectives. Drawing from historical and contemporary crisis-driven education reform efforts, I propose “pedagogies of enlightenment” and “pedagogies of empowerment” as conceptual and experiential maps to guide our path forward.

Pedagogies of enlightenment cultivate personal and communal gifts of intrapersonal growth. Transformative experiences of insight and inspiration are engaged through skilled and persistent individual practice. Pedagogies of empowerment cultivate personal and communal gifts of interpersonal growth. Transformative experiences of social inquiry and social advocacy are engaged through skilled facilitation of diverse and inclusive collective participation. Pedagogies of enlightenment and empowerment represent categories of varied and complex approaches to teaching and learning. There are many options; each to be approached and implemented as a science, an art-form and social mission (Dewey, 1897); each holding important implications not only as teaching strategies but also as research methods. Based upon my own efforts to re-envision democratic civic education in times of crisis—in collaboration with students, teachers and community colleagues in the United States, Japan and Thailand—I offer examples of accessible points of entry as follows:

Pedagogies of enlightenment

Even in the darkest times, individuals can experience renewed meaning, fulfillment and growth through skilled and persistent expression of insight and inspiration.

Insight

Insight is the capacity to suspend habits of thought and emotion, thereby enhancing receptivity and heightening awareness, leading to deepened understanding and enriched relationships.

Surprising, expansive, integrative, resonant, uplifting, transformative—experiences of insight are described across time and cultures. Often framed as religious or spiritual experience, insight entails profound “awakening—a reorientation of the whole person” (Hickman, 2007, p. 199). Such moments of enlightenment are accessible to all, emergent within the rhythms of daily life (Garrison, 2010, 2019; Garrison et al., 2014; Ikeda, 2010a; Maslow, 1971; Hanh, 2012; Reagan, 1996).

Insight, as direct unmediated perception, cannot be planned or orchestrated. Insights can however be cultivated as gifts of disciplined openness; all people can learn to invite insight. The process of refining and expanding awareness often begins by honing skills in naturalistic observation (Bethel, 1989; Dewey, 1938; Hickman, 2008; Ikeda, 2010b). The process can be advanced through meditative practice.

Although there are many approaches to meditation—varied sitting and walking forms, tai chi, yoga, chanting, calligraphy, dance, uplifting anthems, creating mandalas—the path of inquiry remains the same. When practiced with discipline and persistence, the learner moves through experiential stages progressing from focused attention, to freedom from distractions, to enhanced receptivity, to heightened awareness whereby the self is not lost but instead expanded to “stand at the threshold of infinite” while maintaining identity and capacity necessary to act in the finite (Garrison, 2010, p. 3). It is in this state that insight may occur, and when it does, we are fundamentally changed. Rather than seeking to escape the pressures of the world to dwell in some utopian state, instead we are compelled to engage our daily lives with deeper clarity and compassion.

Meditation is a discipline, an art to be practiced and refined throughout the course of a lifetime. And yet there are accessible points of entry suitable for use in secular or mixed-faith settings. A simple 15 min meditation, composed of five 3 min components, is described as follows:

- Exploring Alignment: Sit or stand in an uplifted yet relaxed position.
- Exploring Focused Awareness: Attend to dominant physical sensations.
- Exploring Undifferentiated Awareness: No words, no images, no boundaries; concentrate on the sound of silence, a sense of space, the background of discrete experience.
- Exploring Heightened Awareness: Attend to subtleties not previously experienced; sustain a sense of expectant openness.
- Free write to ground the experience.

Inspiration

Inspiration is the capacity to find meaning and hope that motivate altruism and determination when confronted with social challenge and change.

Poetic representations—including storytelling and other esthetic linguistic forms—emphasize emotional, ethical, artistic and spiritual dimensions of civic life. As noted by Gregory Cajete, indigenous traditions of oral poetry were engaged to discourage conceptual analysis and suspend critical reflection in order to “‘enchant’ the hearers and draw them into ... a dance of meaning

in which complex images, symbols, and meanings are explored in direct and personal ways ... affecting and engaging individuals deeply and multi-dimensionally" (1994, p. 133). Beauty thereby awakens wisdom and deepens informed commitment.

Use of poetry as a source of inspiration pervades politically engaged holistic practice. For example, both Thich Nhat Hahn (1999) and Daisaku Ikeda (2014) have published multiple volumes of moving poetry designed to inspire in times of crisis. John Dewey, also a poet, was taken by the work of Walt Whitman who described democracy as a spiritual force (Garrison et al., 2014). But it was Tsunesaburo Makiguchi who pioneered the use of poetry as an approach to esthetic inquiry accessible to common people (Bethel, 1989; Goulah and Gebert, 2009).

While visiting Soka University in Tokyo, I heard stories of Makiguchi's use of a disciplined, exploratory approach to poetry. His approach was disciplined in that it was highly structured and repetitive—centered in Japanese non-rhyming forms of either haiku (3 lines: 5-7-5 syllables) or tanka (5 lines: 5-7-5-7-7 syllables)—both aimed at focusing intent, at moving to the heart of the matter. His approach was exploratory in that it was open-ended and creative; aimed at awakening, at motivating principled action. Because their lives were so harsh, Makiguchi encouraged young children to write a poem at the beginning of each school day describing how they might create, or do, or offer something of value. Teachers and older students were asked to write at the end of each day so that they might capture and enact the most important lessons learned.

Poetry can be used not only to inspire oneself, but also to encourage and sustain each other, particularly in response to traumatic social events. In classrooms and community settings, poetry as a socially redemptive process, can be engaged as follows:

- Using a structured format, each participant is asked to compose a poem reflecting their response to a specific social event or concern for the future.
- Individual poems are then shared in small groups, attending to both common themes and unique perspectives.
- Together group members then compose one or a series of poems describing actions they might take to address critical concerns raised (Kurth-Schai and Green, 2016).

Pedagogies of Enlightenment, experiences of insight and inspiration often entwined, expand beyond pedagogical applications (Cajete, 1994; Ikeda, 2004, 2010b; Kumar, 2013; Miller et al., 2005; Hanh, 2017; Rudd and Garrison, 2012; Seidel and Jardine, 2014), to reconceptualize and revitalize research. In research settings, pedagogies of enlightenment can be used to promote introspection and enhance clarity among researchers; this, essential at all stages of the research process in pursuit of complex, evolving, unbiased truth. Equally important, pedagogies of enlightenment can be employed to enhance personal expression and reflection among research participants, thereby enhancing interpretation of research findings and recommendations. Applications aimed at deepening qualitative and phenomenological research traditions include poetry, storytelling, portraiture, reflection/contemplation/meditation and efforts to center research ethics on developing authentic understandings of researcher positionality (Creely et al., 2020; Eisner, 1997; Feldman, 2004; Janesick, 2015; Lawrence-Lightfoot and Davis, 1997; Leavy, 2009; Stone, 2018; Vagle, 2018).

Pedagogies of empowerment

Even in the darkest times, we can join together to deepen democracy and promote life-affirming systemic social change through skilled and persistent facilitation of social inquiry and social advocacy.

Social inquiry

Social inquiry is the capacity to learn with, from and for diverse others; to construct knowledge in a respectful, consensual and purposeful manner.

Social inquiry, a fundamental act of empowerment, can be approached by engaging all people in learning to teach and teaching to learn; this, I believe to be a fundamental human right, responsibility and source of valued civic contribution. The science, the art, the power of teaching as a primary means of advancing both personal and social transformation is a resonant theme throughout politically engaged holistic practice; as is the call to support all people in taking on such tasks within the context of their daily lives. These pivotal themes are best captured in descriptions of the "Bodhisattva way".

I was first introduced to this metaphor while visiting a temple in northern Thailand. A young monk focused our attention on an ancient painting of the Buddha seated under a bodhi tree beside a lotus pond. He said it was at this moment that the Buddha, though enlightened, decided to forgo the opportunity to leave this world, choosing instead to remain as a teacher assisting others in moving along the path. He further explained that all beings, like each lotus, chart a distinctive course as they rise through the turbulent waters of life. Therefore the Bodhisattva, the effectively responsive teacher, must continually engage deeper wisdom and compassion so that they might grow in their capacity to tailor instruction to the varied needs and interests of diverse learners.

The works of politically engaged holists—from Freire and Dewey to Montessori and Makiguchi—suggest that we all have the potential, the opportunity and the responsibility to practice the Bodhisattva way (Bethel, 1973; Dewey, 1939b; Freire, 1994; Garrison, 2010; Goulah and Ito, 2012; Hooks, 2003; Ikeda, 2010a,b; Hanh, 2012; Standing, 1957). Instructional strategies framed as peer teaching, co-teaching, communities of inquiry, cooperative and community-based experiential learning are effectively utilized across ages, disciplines and educational settings around the world. We all can engage in complimentary processes of learning to teach and teaching to learn; within our schools, our families, our communities and beyond.

An accessible point of entry would be to routinely require all participants engaged in educational and civic settings—regardless of age, race, gender/sexuality or perceived academic ability etc.—to learn specific concepts or skills for the expressed purpose of

teaching these to others (e.g., peers, younger students, family or community members, civic leaders). Diligent, skilled facilitation must then be devoted to guiding teacher/learners in developing emancipatory learning goals, mastering content, understanding learner characteristics, exploring engaging methods, receiving feedback, reflecting on implementation and refining practice (Kurth-Schai, 2014).

Social advocacy

Social advocacy is the capacity to deepen understanding across domains of power and difference in order to effect inclusive, egalitarian, just and compassionate social change.

In his classic text, *Pedagogy of the Oppressed*, Paulo Freire (2000) charted what was to become a widely embraced philosophical and pedagogical path toward effective social advocacy. He began by asserting that education is inherently a political act. He said that teaching/learning for social stability is accomplished through the “banking approach” whereby strategically controlled knowledge is imposed by the powerful on the masses. He asserted that teaching/learning for social justice must begin with those whose power is limited, and then move to engage others across political spectrum. To this end, he developed “problem-posing education”—a cyclical process designed to grant “voice” to the oppressed, moving from expression of individual experience to a sense of shared solidarity; engage “critical consciousness”, moving from the personal to the political, demystifying systems of structural oppression; and then facilitate “praxis”, reflective participation in principled social action (Freire, 1994, 1998, 2000, 2005).

Committed to a parallel democratic social ethic, John Dewey envisioned and evolved “deliberation” as an inclusive, contextually responsive, imaginative and conciliatory approach to social advocacy. Rather than attempt to determine a course of collective action, or to resolve social disputes via debate—an approach designed to persuade by force of argument—Dewey worked to ensure free and full consideration of diverse, often conflicting perspectives as necessary to yield authentic consensus; acts of creative compromise through which multiple positions and aspirations could be included and represented with integrity. Such an approach would prove to be critical, not only in developing empathy and solidarity across dimensions of power and difference, but also in mounting effective response to both immediate/local and long-term/systemic social challenges (Dewey, 1916, 1927, 1932, Dewey, 1939b).

As a point of entry to social advocacy, skilled facilitation can be employed to ensure that all participants engage fully in each of the following steps adapted from Dewey's (1910) conceptualization of a “complete act of thought”:

- Identify the problematic situation, a complex dilemma that creates confusion because it cannot be understood or resolved based solely on past experience.
- Define and clarify one specific problem; identify and analyze every important dimension devoting special attention to issues of immediate collective concern.
- Develop a series of possible strategies for resolving the specific problem.
- Arrive at genuine consensus regarding one potentially promising solution path.
- Assess the solution path's validity by acting on it in a challenging community-based setting and then evaluating the social impact.
- Collectively evaluate participants' experience of the deliberative process; then join together in determining next steps.

Pedagogies of Empowerment, acts of inquiry and advocacy often entwined, expand beyond applications posed by politically engaged holists to include teaching/learning techniques drawn from *democratic* (Hyslop-Margison and Sears, 2006; Noddings, 2013; Parker, 2002, 2004; Stitzlein, 2012), *critical* (Freire, 1998, 2005; Giroux, 2020; Greene, 1995), *feminist* (Hernandez, 1997; Hooks, 2014), *eco-justice/sustainability* (Bowers, 1992; Kurth-Schai, 1992; Martusewicz et al., 2015; Nolet, 2015; Sulak, 2009), *culturally responsive* (Gay, 2018), *anti-racist* (Casey, 2016; Giroux, 2021; Ladson-Billings, 2003, 2004; May, 1999), *anti-colonialist* (Calderon, 2014; Sefa Dei and Kempf, 2006; Zainub, 2019), *indigenous* (Cajete, 1994; Deloria and Wildcat, 2001; McCoy et al., 2017; Sefa Dei, 2011), and other categories of liberatory pedagogy.

Related research methods—aimed at enhancing the relevance and utility of research findings by engaging both researchers and participants in social inquiry and advocacy—include critical participatory action research, youth participatory action research, cultural futures Delphi and ethnographic futures research (Caraballo et al., 2017; Cochran-Smith and Lytle, 2009; Fine, 2018; Kurth-Schai et al., 2000; Mirra et al., 2016; Textor, 1978).

Moving toward the light: our call to action

When the day comes, we step out of the shade,

aflame and unafraid,

the new dawn blooms as we free it,

for there is always light,

if only we're brave enough to see it,

if only we're brave enough to be it.

Gorman (2021).

Escaping delusion, division and despair requires sustained and mindful grappling with deeply troubling personal and collective concerns. Seeking the light, seeking wisdom, requires discipline, humility, persistence, creativity and compassion. We can afford neither to remain constrained, pessimistic and exclusionary; nor romantic, disorganized and naïve. Systemic, inclusive, egalitarian reform requires careful orchestration and pragmatic innovation. And so we are called to engage pedagogies of enlightenment and empowerment in all aspects of our lives—in our classrooms, boardrooms, voluntary associations, political institutions, public service bureaucracies, philanthropic organizations, and on into the streets as social movements. Herein lies the path toward creative, evolving democratic learning and life capable of bridging the gap between cherished hopes and challenging realities. Herein lies our collective path from the darkness into the light.

References

- Adorjan, M., Kelly, B., 2008. Pragmatism and "engaged" Buddhism: working toward peace and a philosophy of action. *Hum. Architect. J. Sociol. Self Knowl.* 6 (3), 37–49.
- Allen, D., Reich, R., 2013. *Education, Justice and Democracy*. University of Chicago Press, Chicago, IL.
- Ansell, C., 2011. *Pragmatist Democracy: Evolutionary Learning as Public Philosophy*. Oxford University Press, Oxford, UK.
- Apple, M., Kenway, J., Singh, M. (Eds.), 2005. *Globalizing Education: Policies, Pedagogies, & Politics*. Peter Lang, New York, NY.
- Apple, M., 2006. *Educating the Right Way*, second ed. Routledge, New York, NY.
- Applebaum, A., 2020. *Twilight of Democracy: The Seductive Lure of Authoritarianism*. Doubleday, New York, NY.
- Banks, J., 2021. Diversity, Transformative Knowledge, and Civic Education: Selected Essays. Teachers College Press, New York, NY.
- Barrow, E., 2017. No global citizenship: re-envisioning global citizenship education in times of growing nationalism. *High Sch. J.* 100 (3), 163–165.
- Bethel, D., 1973. *Makiguchi the Value Creator: Revolutionary Japanese Educator and Founder of Soka Gakkai*. Weatherhill, New York, NY.
- Bethel, D. (Ed.), 1989. *Education for Creative Living: Ideas and Proposals of Tsunesaburo Makiguchi*. Iowa State University Press, Ames, IA.
- Bowers, C., 1992. *Education, Cultural Myths and the Ecological Crisis: Toward Deep Changes*. State University of New York Press, Albany, NY.
- Brameld, T., 2000. *Education as power*. In: Paperback Republication of 1965 Classic. Caddo Gap Press, San Francisco, CA.
- Buddhadasa, B., 2005. *Handbook for Mankind*. Office of National Buddhism Press, Bangkok, Thailand.
- Burch, K., 2000. *Eros as the Educational Principle of Democracy*. Peter Lang, New York, NY.
- Burch, P., 2009. *Hidden Markets: The New Education Privatization*. Routledge, New York, NY.
- Cajete, G., 1994. *Look to the Mountain: An Ecology of Indigenous Education*. Kivaki Press, Asheville, NC.
- Calderon, D., 2014. Anticolonial methodologies in education: embodying land and indigeneity in Chicana feminisms. *J. Latino-Latin Am. Stud.* 6 (2), 81–96.
- Caraballo, L., Lozenski, B., Lyiscott, J., Morrell, E., 2017. YPAR and critical epistemologies: rethinking educational research. *Rev. Res. Educ.* 41, 311–336.
- Carlson, D., 2002. *Leaving Safe Harbors: Toward a New Progressivism in American Education and Public Life*. Routledge, New York, NY.
- Casey, Z., 2016. *A Pedagogy of Anticapitalist Antiracism: Whiteness, Neoliberalism and Resistance in Education*. State University of New York Press, New York, NY.
- Cochran-Smith, M., Lytle, S., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York, NY.
- Comaroff, J.L., Comaroff, J., 2001. *Millennial Capitalism and the Culture of Neoliberalism*. Duke University Press, Durham, NC.
- Creely, E., Southcott, J., Carabott, K., Lyons, D., 2020. *Phenomenological Inquiry in Education: Theories, Practices, Provocations and Directions*. Routledge, New York, NY.
- Dardot, P., Laval, C., 2013. *The New Way of the World: On Neoliberal Society*. Verso, London, UK.
- Deloria, V., Wildcat, D., 2001. *Power and Place: Indian Education in America*. Fulcrum, Golden, CO.
- Deloria, V., 1979. *The Metaphysics of Modern Existence*. Harper & Row, New York, NY.
- Dewey, J., 1888. *The Ethics of Democracy*. In: University of Michigan Philosophical Papers, Second Series, No. 1. Andrews & Company Publishers.
- Dewey, J., 1897. My pedagogic creed. *Sch. J.* 54, 77–80.
- Dewey, J., 1910. *How We Think*. D.C. Heath & Company, New York, NY.
- Dewey, J., 1916. *Democracy and Education*. Macmillan, New York, NY.
- Dewey, J., 1919/1998. *Philosophy and democracy*. In: Hickman, L., Alexander, T. (Eds.), *The Essential Dewey, Volume I: Pragmatism, Education, Democracy*. Indiana University Press, Bloomington, IN.
- Dewey, J., 1920/1980. *Reconstruction in philosophy*. In: Boydston, J. (Ed.), *John Dewey: The Middle Works, vol. 12*. Southern Illinois University Press, Carbondale, IL, pp. 1899–1924.
- Dewey, J., 1922/1998. *Valuation and experimental knowledge*. In: Hickman, L., Alexander, T. (Eds.), *The Essential Dewey, Volume II: Ethics, Logic, Psychology*. Indiana University Press, Bloomington, IN.
- Dewey, J., 1927/1984. *The public and its problems*. In: Boydston, J. (Ed.), *John Dewey: The Later Works, vol. 2*. Southern Illinois University Press, Carbondale, IL, pp. 1925–1953.
- Dewey, J., 1928/1998. *The inclusive philosophic idea*. In: Hickman, L., Alexander, T. (Eds.), *The Essential Dewey, Volume I: Pragmatism, Education, Democracy*. Indiana University Press, Bloomington, IN.
- Dewey, J., 1932/1985. *Ethics*. In: Boydston, J. (Ed.), *John Dewey: The Later Works, vol. 7*. Southern Illinois University Press, Carbondale, IL, pp. 1925–1953.
- Dewey, J., 1934. *A Common Faith*. Yale University Press, New Haven, CT.
- Dewey, J., 1937a. *Education and social change*. *Bull. Am. Assoc. Univ. Professors* 23 (6), 472–474.
- Dewey, J., 1937b/1998. *Democracy is radical*. In: Hickman, L., Alexander, T. (Eds.), *The Essential Dewey, Volume I: Pragmatism, Education, Democracy*. Indiana University Press, Bloomington, IN.

- Dewey, J., 1937c/1991. Education and social change. In: Boydston, J. (Ed.), *John Dewey: The Later Works*, vol. 7. Southern Illinois University Press, Carbondale, IL, pp. 1925–1953.
- Dewey, J., 1938. *Experience and Education*. Macmillan, New York, NY.
- Dewey, J., 1939a/1985. Freedom and culture. In: Boydston, J. (Ed.), *John Dewey: The Later Works*, vol. 13. Southern Illinois University Press, Carbondale, IL, pp. 1925–1953.
- Dewey, J., 1939b/1998. Creative democracy—the task before us. In: Hickman, L., Alexander, T. (Eds.), *The Essential Dewey, Volume I: Pragmatism, Education, Democracy*. Indiana University Press, Bloomington, IN.
- Diamond, L., 2008. *The Spirit of Democracy: The Struggle to Build Free Societies Throughout the World*. Times Books, New York, NY.
- Doll, W., Gough, N. (Eds.), 2006. *Curriculum Visions*, second ed. Peter Lang, New York, NY.
- Du Bois, W.E.B., 2019. *The Souls of Black Folk*, original ed. G&D Media.
- Eisner, E., 1997. *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*, second ed. Pearson, New York, NY.
- Ellsberg, R. (Ed.), 2001. *Thich Nhat Hanh: Essential Writings*. Orbis Books, Maryknoll, NY.
- Epert, C., Vokey, D., Nguyen, T., Bai, H., 2015. Intercultural philosophy and the nondual wisdom of “basic goodness”: implications for contemplative and transformative education. *J. Philos. Educ.* 49 (2), 274–293.
- Feldman, R., 2004. Poetic representation of qualitative research. *J. Crit. Inq. Curric. Instr.* 5 (2), 10–14.
- Fine, M., 2018. *Just Research in Contentious Times: Widening the Methodological Imagination*. Teachers College Press, New York, NY.
- Freire, P., 1994. *Pedagogy of Hope: Reliving Pedagogy of the Oppressed*. Continuum, New York, NY.
- Freire, P., 1998. *Pedagogy of Freedom: Ethics, Democracy, and Civic Courage*. Rowman & Littlefield, Lanham, MA.
- Freire, P., 2000. *Pedagogy of the Oppressed*, 30th anniversary ed. Continuum, New York, NY.
- Freire, P., 2005. *Education for Critical Consciousness*. Continuum, New York, NY.
- Gabbard, D., 2008. *Knowledge and Power in the Global Economy*, second ed. Lawrence Erlbaum, New York, NY.
- Garrison, J., Hickman, L., Ikeda, D., 2014. *Living as Learning: John Dewey in the 21st Century*. Dialogue Path Press, Cambridge, MA.
- Garrison, J., 1997. *Dewey and Eros: Wisdom and Desire in the Art of Teaching*. Teachers College Press, New York, NY.
- Garrison, J., 2008. A pragmatist approach to emotional expression and the construction of gender identity. In: Garrison, J. (Ed.), *Reconstructing Democracy, Recontextualizing Dewey*. State University of New York Press, New York, NY, pp. 157–184.
- Garrison, J., 2009. John Dewey and Daisaku Ikeda: A Religious Dialogue. Lecture at the Institute of Oriental Philosophy, Tokyo, Japan.
- Garrison, J., 2010. Compassionate, spiritual, and creative listening in teaching and learning. *Teach. Coll. Rec.* 112 (11), 2763–2776.
- Garrison, J., 2019. Nichiren Buddhism and Deweyan pragmatism: an eastern-western integrations of thought. *Educ. Stud.* 55 (1), 12–27.
- Gay, G., 2018. *Culturally Responsive Teaching: Theory, Research, and Practice*, third ed. Teachers College Press, New York, NY.
- Giroux, H., 2004. *The Terror of Neoliberalism*. Paradigm Press, Boulder, CO.
- Giroux, H., 2008. *Against the Terror of Neoliberalism: Politics Beyond the Age of Greed*. Paradigm Press, Boulder, CO.
- Giroux, H., 2020. *On Critical Pedagogy*, second ed. Bloomsbury, New York, NY.
- Giroux, H., 2021. *Race, Politics and Pandemic Pedagogy*. Bloomsbury, New York, NY.
- Gorman, A., 2021. *The Hill We Climb: An Inaugural Poem for the Country*. Viking Books, New York, NY.
- Goulah, J., Gebert, A., 2009. Tsunesaburo Makiguchi: introduction to the man, his ideas, and the special issue. *Educ. Stud.* 45 (2), 115–132.
- Goulah, J., Ito, T., 2012. Daisaku Ikeda's curriculum of Soka Education: creating value through dialogue, global citizenship, and “human revolution” in the mentor-disciple relationship. *Curric. Inq.* 42 (1), 56–79.
- Green, J., 2008a. Dr. Dewey's deeply democratic metaphysical therapeutic for the post- 9/11 American democratic disease: toward cultural revitalization and political rehabilitation. In: Garrison, J. (Ed.), *Reconstructing Democracy, Recontextualizing Dewey*. State University of New York Press, New York, NY, pp. 31–54.
- Green, J., 2008b. *Pragmatism and Social Hope: Deepening Democracy in Global Contexts*. Columbia University Press, New York, NY.
- Greene, M., 1995. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*. Jossey-Bass, San Francisco, CA.
- Gutmann, A., Thompson, D., 2009. *Why Deliberative Democracy?* Princeton University Press, Princeton, NJ.
- Hanh, N., 1999. *Call Me by My True Names: The Collected Poems of Thich Nhat Hahn*. Parallax Press, Berkeley, CA.
- Hanh, N., 2012. *Good Citizens: Creating Enlightened Society*. Parallax Press, Berkeley, CA.
- Hanh, N., 2017. *Happy Teachers Change the World*. Parallax Press, Berkeley, CA.
- Hansen, D. (Ed.), 2007. *Ethical Visions in Education: Philosophies in Practice*. Teachers College Press, New York, NY.
- Harvey, D., 2005. *A Brief History of Neoliberalism*. Oxford University Press, Oxford, UK.
- Hernandez, A., 1997. *Pedagogy, Democracy, and Feminism: Rethinking the Public Sphere*. State University of New York Press, Albany, NY.
- Heybach, J., Sheffield, E., 2014. Creating citizens in a capitalist democracy: a struggle for the soul of American citizenship education. In: Petrovic, J., Kuntz, A. (Eds.), *Citizenship Education Around the World: Local Context and Global Possibilities*. Routledge, New York, NY.
- Hickman, L., 2007. *Pragmatism as Post-modernism: Lessons From John Dewey*. Fordham University Press, New York, NY.
- Hickman, L., 2008. Evolutionary naturalism, logic, and lifelong learning: three keys to John Dewey's philosophy of education. In: Garrison, J. (Ed.), *Reconstructing Democracy, Recontextualizing Dewey*. State University of New York Press, New York, NY, pp. 119–135.
- Hooks, B., 2003. *Teaching Community: A Pedagogy of Hope*. Routledge, New York, NY.
- Hooks, B., 2014. *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, New York, NY.
- Hyslop-Margison, E., Sears, A., 2006. *Neo-liberalism, Globalization, and Human Capital Learning: Reclaiming Education for Democratic Citizenship*. Springer, Dordrecht, The Netherlands.
- Hyttén, K., 2009. Deweyan democracy in a globalized world. *Educ. Theor.* 59 (4), 395–408.
- Ibrahim, A., 2009. Learning to learn: Makiguchi as a “strong poet” of geography, courage and happiness. *Educ. Stud.* 45 (2), 221–226.
- Ikeda, D., 2004. *Light of Education: Thoughts and Vision for Humanistic Education*. Soka Gakki Malaysia, Kuala Lumpur, Malaysia.
- Ikeda, D., 2005. *Forward*. In: Noddings, N. (Ed.), *Educating Citizens for Global Awareness*. Teachers College Press, New York, NY.
- Ikeda, D., 2010a. *A New Humanism: The University Addresses of Daisaku Ikeda*. I.B. Tauris & Co. Ltd, New York, NY.
- Ikeda, D., 2010b. *Soka Education: For the Happiness of the Individual*. Middleway Press, Santa Monica, CA.
- Ikeda, D., 2014. *Journey of Life: Selected Poems of Daisaku Ikeda*. I.B. Tauris & Co. Ltd, New York, NY.
- Ikeda, D., 2021. Daisaku Ikeda: Buddhist Philosopher, Peacebuilder and Educator. Soka Gakki. <https://www.daisakuikeda.org/sub/quotations/theme/courage.html>.
- Jackson, P., 2003. *Buddhadasa: Theravada Buddhism and Modernist Reform in Thailand*. Silksworm Books, Chiang Mai, Thailand.
- Janesick, V., 2015. *Contemplative Qualitative Inquiry: Practicing the Zen of Research*. Routledge, New York, NY.
- Katz, M., Verducci, S., Biesta, G. (Eds.), 2009. *Education, Democracy, and Moral Life*. Springer, New York, NY.
- Kesson, K., 2003. Individuation as spiritual process: Jung's archetypal psychology and the development of teachers. *Encount. Educ. Mean. Soc. Justice* 16 (4), 2–31.
- Kimmerer, R., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge, and the Teaching of Plants*. Milkweed Editions, Minneapolis, MN.
- King, S., 2006. *Being Benevolence: Social Ethics of Engaged Buddhism*. University of Hawaii Press, Honolulu, HI.
- Kumar, A., 2013. *Curriculum as Meditative Inquiry*. Palgrave, New York, NY.
- Kurth-Schai, R., Green, C., 1997. *Schooling stories*. In: Schram, S. (Ed.), *Tales of the State: Narrative in Contemporary U.S. Politics and Public Policy*. Rowman & Littlefield, Lanham, MA.

- Kurth-Schai, R., Green, C., 2016. *Re-envisioning Education and Democracy*, second ed. Information Age Publishing, Charlotte, NC.
- Kurth-Schai, R., Poolpatarachewin, C., Pitiyanuwat, S., 2000. Using the Delphi cross-culturally: towards the development of policy. In: Cogan, J., Derricott, R. (Eds.), *Citizenship for the 21st Century*. Kogan Page, London, UK, pp. 93–108.
- Kurth-Schai, R., 1992. Ecology and Equity: toward the rational re-enchantment of schools and society. *Educ. Theor.* 42 (2), 147–163.
- Kurth-Schai, R., 2014. Fidelity in public education policy: reclaiming the Deweyan dream. *Educ. Stud.* 50 (5), 420–446.
- Ladson-Billings, G. (Ed.), 2003. *Critical Race Theory Perspectives on Social Studies: The Profession, Policies and Curriculum*. Information Age Press, Greenwich, CT.
- Ladson-Billings, G., 2004. Culture versus citizenship: the challenge of racialized citizenship in the United States. In: Banks, J. (Ed.), *Diversity and Citizenship Education: Global Perspectives*. Jossey-Bass, San Francisco, CA.
- Lama, D., 1999. *Ethics for the New Millennium*. Riverhead Books, New York, NY.
- Lawrence-Lightfoot, S., Davis, J., 1997. *The Art & Science of Portraiture*. Jossey-Bass, San Francisco, CA.
- Leavy, P., 2009. *Method Meets Art: Arts Based Research Practice*. Guilford Press, New York, NY.
- Lewis, J., 2020. Final Essay Published on the Opinion Page of the New York Times on the Day of the Funeral of the Honorable. John R. Lewis. <https://www.nytimes.com/2020/07/30/opinion/john-lewis-civil-rights-america.html>.
- Lipman, P., 2004. *High Stakes Education: Inequality, Globalization, and Urban School Reform*. Routledge, New York, NY.
- Lummis, C., 1996. *Radical Democracy*. Cornell University Press, Ithaca, NY.
- Martusewicz, R., Edmundson, J., Lupinacci, J. (Eds.), 2015. *Ecojustice Education: Toward Diverse, Democratic, and Sustainable Communities*, second ed. Routledge, New York, NY.
- Maslow, A., 1971. *The Farther Reaches of Human Mature*. Viking Press, New York, NY.
- Mason, M., 2008. *Complexity Theory and the Philosophy of Education*. Wiley-Blackwell, West Sussex, UK.
- May, S. (Ed.), 1999. *Critical Multiculturalism: Rethinking Multicultural and Antiracist Education*. Falmer Press, London, UK.
- McCluskey, A., 2014. *A Forgotten Sisterhood: Pioneering Black Women Educators and Activists in the Jim Crow South*. Rowman & Littlefield, Lanham, MD.
- McCoy, K., Tuck, E., McKenzie, M. (Eds.), 2017. *Land Education: Rethinking Pedagogies of Place From Indigenous, Postcolonial, and Decolonizing Perspectives*. Routledge, New York, NY.
- Miller, J., Karsten, S., Denton, D., Orr, D., Kates, I. (Eds.), 2005. *Holistic Learning and Spirituality in Education*. State University of New York Press, New York, NY.
- Miller, J., 2019. *The Holistic Curriculum*, third ed. University of Toronto Press, Toronto, Canada.
- Mirra, N., Garcia, A., Morrell, E., 2016. *Doing Youth Participatory Action Research: Transforming Inquiry With Researchers, Educators, and Students*. Routledge, New York, NY.
- Noddings, N., 2005. *The Challenge to Care in Schools*, second ed. Teachers College Press, New York, NY.
- Noddings, N., 2013. *Education and Democracy in the 21st Century*. Teachers College Press, New York, NY.
- Nolet, V., 2015. *Educating for Sustainability: Principles and Practices for Teachers*. Routledge, New York, NY.
- Oldenski, T., Carlson, D. (Eds.), 2002. *Educational Yearning: The Journey of Spirit and Democratic Education*. Peter Lang, New York, NY.
- Olsen, M., Codd, J., O'Neill, A., 2004. *Education Policy: Globalization, Citizenship, and Democracy*. Sage Publications, Thousand Oaks, CA.
- Palmer, P., 2011. *Healing the Heart of Democracy: The Courage to Create a Politics Worthy of the Human Spirit*. Jossey-Bass, San Francisco, CA.
- Parker, W., 2002. *Teaching Democracy: Unity and Diversity in Public Life*. Teachers College Press, New York, NY.
- Parker, W., 2004. Diversity, globalization, and democratic education: curriculum possibilities. In: Banks, J. (Ed.), *Diversity and Citizenship Education*. Jossey-Bass, San Francisco, CA.
- Porfilio, B., Malott, C., 2008. *The Destructive Path of Neoliberalism: An International Examination of Education*. Sense Publishers, Rotterdam, The Netherlands.
- Purpel, D., 2004a. *Critical Social Issues in American Education: Democracy and Meaning in a Globalizing World*. Routledge, New York, NY.
- Purpel, D., 2004b. *Reflections on the Moral and Spiritual Crisis in Education*. Peter Lang, New York, NY.
- Reagan, T., 1996. *Non-western Educational Traditions: Alternative Approaches to Educational Thought and Practice*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Richardson, R. (Ed.), 2010. *The Heart of William James*. Harvard University Press, Cambridge, MA.
- Riley-Taylor, E., 2002. *Ecology, Spirituality & Education: Curriculum for Relational Knowing*. Peter Lang, New York, NY.
- Rogers, M., 2009. *The Undiscovered Dewey: Religion, Morality, and the Ethos of Democracy*. Columbia University Press, New York.
- Rudd, A., Garrison, J., 2012. *Teaching With Reverence: Reviving an Ancient Virtue for Today's Schools*. Macmillan Palgrave, New York, NY.
- Scott, D., 2000. William James and Buddhism: American pragmatism and the orient. *Religion* 30, 1–20.
- Sefa Dei, G., Kempf, A. (Eds.), 2006. *Anti-colonialism and Education: The Politics of Resistance*. Sense Publishers, Rotterdam, The Netherlands.
- Sefa Dei, G. (Ed.), 2011. *Indigenous Philosophies and Critical Education: A Reader*. Peter Lang, New York, NY.
- Seidel, J., Jardine, D., 2014. *Ecological Pedagogy, Buddhist Pedagogy, Hermeneutic Pedagogy*. Peter Lang, New York, NY.
- Shah, A., 2010. A primer on neoliberalism. In: *Global Issues: Social, Political, Economic and Environmental Issues that Affect Us All*. <https://www.globalissues.org/article/39/a-primer-on-neoliberalism>.
- Shapiro, H., 2009. *Education and Hope in Troubled Times*. Routledge, New York, NY.
- Snyder, T., 2018. *The Road to Unfreedom: Russia, Europe, America*. Tim Duggan Books, New York, NY.
- Standing, E., 1957. *Maria Montessori: Her Life and Work*. Hollis & Carter, London, UK.
- Standish, P., Naoko, S. (Eds.), 2012. *Education and the Kyoto School of Philosophy: Pedagogy for Human Transformation*. Springer, London, UK.
- Steiner, R., 2008. *The Essential Rudolf Steiner*. Wilder Publications, Radford, VA.
- Stitzlein, S., 2012. *Teaching for Dissent: Citizenship Education and Political Activism*. Routledge, New York, NY.
- Stone, L., 2018. Research ethics and a case of positionality. In: Smeyers, P., Depaepe, M. (Eds.), *Educational Research: Ethics, Social Justice, and Funding Dynamics*, pp. 55–72.
- Stromquist, N., 2002. *Education in a Globalized World: The Connectivity of Economic Power, Technology, and Knowledge*. Rowman & Littlefield, Lanham, MD.
- Suarez-Orozco, M., Qin-Hilliard, D. (Eds.), 2004. *Globalization, Culture and Education in the New Millennium*. University of California Press, Berkeley.
- Sulak, S., 2005. *Conflict, Culture, Change: Engaged Buddhism in a Globalizing World*. Suksit Siam Publications, Bangkok, Thailand.
- Sulak, S., 2009. *The Wisdom of Sustainability: Buddhist Economics for the 21st Century*. Silkworm Books, Chiang Mai, Thailand.
- Suu Kyi, A.S., 2010. *Letters From Burma*. Penguin Books, London, UK.
- Tarozzi, M., Torres, C., 2016. *Citizenship Education and the Crises of Multiculturalism: Comparative Perspectives*. Bloomsbury, London, UK.
- Textor, R., 1978. Cultural futures for Thailand: an ethnographic inquiry. *Futures* 10 (5), 347–360.
- Treanor, P., 2005. Neoliberalism: Origins, Theory, Definition. <http://web.inter.nl.net/users/Paul.Treanor/neoliberalism.html>.
- Unger, R., 2007. *The Self Awakened: Pragmatism Unbound*. Harvard University Press, Cambridge, MA.
- Vagle, M., 2018. *Crafting Phenomenological Research*, second ed. Routledge, New York, NY.
- Wang, J., 2009. Reconstructing Deweyan democratic education for a globalized world. *Educ. Theor.* 59 (4), 409–425.
- Ward, S., 2012. *Neoliberalism and the Global Restructuring of Knowledge and Education*. Routledge, New York, NY.
- West, C., 2004. *Democracy Matters: Winning the Fight Against Imperialism*. Penguin Press, New York, NY.
- Westbrook, R., 2005. *Democratic Hope: Pragmatism and the Politics of Truth*. Cornell University Press, Ithaca, NY.
- Zainub, A., 2019. *Decolonization and Anti-colonial Praxis*. Brill, Leiden, the Netherlands.
- Ziblatt, D., Levitsky, S., 2018. *How Democracies Die*. Crown Publishers, New York, NY.

Peace education in conflict zones: general trends and Israeli particularities

Zvi Bekerman, The Seymour Fox School of Education, The Hebrew University of Jerusalem, Jerusalem, Israel

© 2023 Elsevier Ltd. All rights reserved.

Historical development	94
Theoretical premises	95
Peace education in Israel	96
The educational settings	96
Research findings	97
Critical perspectives and more	99
Adding educational concerns to critical perspectives	99
Adding methodological concerns to critical perspectives	100
Conclusion	100
References	101

The purpose of this paper is to offer readers a short introduction into the multifaceted field of Peace Education. In the first section, we offer some insights into the historical development of peace education, and point at the main intellectual perspectives and activities which have evolved since its inception. We then summarize three of the main theoretical frameworks which support peace educational work; the contact hypothesis, the socialization and social learning theories, and the social-cognitive developmental theory. In the following section, and with hope of helping readers envision the implementation of a variety of peace educational programs and their potential influence on conflicted societies, we introduce the case of Israel detailing some of the educational programs that have been developed in this country and presenting findings of research that investigated these programs. In the last section, we challenge present perspectives on peace education and research to (re) consider some larger educational and methodological concerns so as to become more effective in their path to contributing to peace and justice in complex global and local contexts.

Historical development

Presently, our world seems to be ailing from political and social instability and conflict. This is not really a new phenomenon. Technological developments have shortened distances and expanded our potential to witness conflicts afar adding to our growing sense of alarm and anxiety. Political and social instability and conflict are, for the most part, considered to be the product of group's territorial conflicts (and/or other material resources) set at the background of the group's differentiated cultures.

When considering possible solutions, to existing tensions, governments and international organizations point at education as one of the main arenas through which to manage and sooth existing conflicts. Educational institutions are expected to help overcome social divisions and thus contribute to social stability and peaceful development. Schools as central sites for socialization through carefully designed curriculum and pedagogies have in their power to improve interpersonal relationships and international understanding and thus help deal with a variety of intra and interstate conflicts. Educational initiatives of this sort go under the general title of Peace Education. In this compound, education points at the formal institutionalized efforts (mostly within state dominated educational systems) to transmit to younger generation knowledge and skills and a set of values and norms accepted in a given society; yet peace is a fussy concept one which evades clear definitions or which's definition suffers from ambiguity.

Peace has been conceptualized as being positive or negative; John Galtung (1969, 2008) one of the contemporary fathers of peace education, defined negative peace in terms of absence of war or direct or physical violence, whereas positive peace is defined as a process aimed at achieving the absence of indirect or structural violence. Neologisms such as "non-violence," "non-aggression" and "non-conflict," identify the negative peace approach while buzzwords such as cooperation, collaboration and harmony, are set as identifiers of positive peace approaches (see Fox, 2013). While the negative definition might allow, paradoxically, for the understanding of a state of absolutism as a situation contributing to peace, positive approaches link peace to justice not allowing for peace to be acknowledged as the result of violent imposition.

Historically, Peace education can be traced back to Comenius, the Czech educator who by the seventeenth century assumed that the road to peace could be facilitated by universally shared knowledge (Harris, 2010). In the 19th and 20th centuries, interest in the field of peace education dramatically increased following the horrors of the world wars and growing threats of nuclear confrontations. In 1946 The United Nations Educational, Scientific and Cultural Organization (UNESCO) was founded as an umbrella institution of the United Nations, and it was charged with planning, developing, and implementing general changes in education for it to align to the newly articulated international politics of peace and security (Lerch and Buckner, 2018).

Peace Education, however, is not an exact term but includes a variety of initiatives, programs, and studies conducted from numerous intellectual perspectives. These perspectives govern value criteria whose roots lie in the seminal works of scholars such as Galtung (1969, 1973), Brock-Utne (1985), and Reardon (1988). Galtung, as mentioned, distinguished between negative and positive peace. Brock-Utne, from a feminist perspective, identified different levels at which violence must be addressed distinguishing between the “organized” state level, and the “unorganized” level, pointing at violence that occurs in micro-structures, such as in families and communities. Reardon (1988) emphasized the connection between education as an intentional ethical and political activity, and the development of social value which can support peace.

More recently, there has been a rise in critical approaches to peace education (Bajaj, 2015; Brantmeier, 2011; Hantzopoulos and Bajaj, 2016; Bekerman, 2000, 2002, 2007, 2016). These approaches offer contextual and conceptual resources for understanding the structural impediments to advancing the promise of peace education in diverse locations; their main contribution rests on the analysis of both power dynamics and intersectionalities among race, class, gender, (dis)ability, sexual orientation, language, religion, geography, and other forms of stratification in conflictual settings (Bajaj, 2019).

Moreover, diverse theoretical approaches, underlying philosophies, and goals support the multiple issues on which peace education may focus. Issues ranging from bullying in schools to collaboration on areas of international security, from conflicts between the industrialized and the undeveloped world to the protection of the environment, and sustainable growth. Critics could argue that the field is too wide and that peace education lacks a unique theoretical framework, a firm methodology, and evaluation tools that will allow it to, properly, measure its outcomes.

The main approaches developed in Peace Education come from different appreciations of what is needed to deal with conflict (Minow, 2002), understood as the result of cultural heterogeneity on the local and national levels. Although heterogeneity is not new, much conflict is ascribed to the fact that technological developments and economic needs have made heterogeneity, in and among nation-states, the rule and not the exception. In spite of the recognition that diversity can enrich human experience, nation-states wrestle with tensions they see as following necessarily from diversity. In both the public and private spheres, equal distribution, citizenship, minority rights (including language), democratic participation, and cultural recognition are only some of the issues perceived to be a threat to peace in any of its forms (Opatow et al., 2005).

Other educational approaches share similar goals. Such are Conflict resolution, Multicultural education, and Intercultural education. Conflict resolution (for a review see Garrard and Lipsey, 2007) can be considered a retroactive approach trying to solve a conflict after it has already occurred; Multicultural education is based on a theory of cultural pluralism; ideals of social justice and the end of racism, sexism, and other forms of prejudice and discrimination. Multicultural education asserts the centrality of culture in the teaching and learning process; and as a prerequisite for educational equity and excellence (for a review see Nieto, 2017). Intercultural education, in its traditional approaches, focuses on supporting and celebrating diversity while more critical approaches support cultural diversity and social justice as well as the need to counter marginalization and discrimination in education and society (for a review see Zilliacus et al., 2017).

Peace education could be conceptualized as trying to provide a more comprehensive framework that includes the aims noted above and adds to them the need to address fears; to offer information regarding security systems; to support world disarmament; to understand violence and violent behavior; to provide for a future orientation; and to teach peace as a process (Harris, 2002). One of the great challenges of Peace Education as any other education geared toward the bettering of that which is human is its need to confront and change the ways in which education devalues diversity, and nurtures violent attitudes (Fredland, 2008; Singh, 2015) for education can prevent conflict but can also exacerbate it.

Theoretical premises

All peace educational initiatives base themselves in one way or another on a variety of theoretical developments, the most important of which we review in the following.

In general, we can think of peace education as guided by diverse perspectives on the type of societal unity needed to sustain a peaceful environment. These perspectives can be described as ranging from a rather strong sense of unity through integration/assimilation to a somewhat more relaxed sense of unity sustained through social cohesion. In both cases, education is supposed to be able, to contribute in this direction by helping groups overcome prejudices and reduce intergroup tensions.

The contact hypothesis, the socialization and social learning theories, and the social-cognitive developmental theory are the three main theoretical models through which to approach educational frameworks, which’s goals are the bettering of intergroup relations (Berger et al., 2016).

The “contact hypothesis” is the basis of most educational efforts toward integration. It can be traced far into the past, but it was given its contemporary form by Gordon W. Allport in *The Nature of Prejudice* (1954). The contact hypothesis suggests that promoting contact between members of different racial and/or ethnic groups will reduce tension, resulting in more tolerant and positive attitudes. The contact hypothesis lists a large number of variables which might influence the outcomes of intergroup contact. Contact will be more successful when it occurs under conditions of equality and interdependence that permit sustained interaction between participants as well as friendships in situations legitimized through institutional support (Allport, 1954; Amir, 1976; Pettigrew and Tropp, 2006).

Social Identity Theory (Tajfel and Turner, 1979), suggests that under certain conditions, people change from viewing the “self” and “other” as individuals to seeing them as representatives of broader social groups. Different social-cognitive processes operate

when such a shift takes place (Dovidio et al., 2003): individuals belonging to the in-group and who share salient characteristic, are chosen as interaction partners, while out-group members are negatively stereotyped. Thus, social settings in which social identity is salient lead to reduced intergroup contact, increased ethnic segregation, and increased intergroup prejudice (Gaertner et al., 2000; Rosenthal and Levy, 2010). Contact in this context can also become a beneficial option to remedy tensions but also other strategies are considered such as for example, individuation (Wilder, 1986) which renders members more distinctive and blurs the a prior categorization. Crisscrossing categories by forming new groups, which now include members from former groups, changes clarity as to who belongs to which group (Brewer et al., 1987; Vanbeselaere, 1996). Another possible strategy is re-categorization that in contrast to de-categorization is not designed to reduce salient categories but rather to redefine group affiliations in ways that reduce intergroup bias and conflict.

Social learning theories propose that attitudes toward outgroups are shaped by information and knowledge assembled from immediate social contexts as well as from multiple media channels. Both intercultural training (Stier, 2003) and anti-bias information (Bigler and Liben, 2006) are seen as capable of breaking down negative generalizations.

Social cognitive developmental theory establishes that children's attitudes are based on the developmental stage of their cognitive skills (Aboud et al., 2003); research in this field has argued against a unidirectional understanding of prejudice development anchoring it in complex variables as these relate to social contexts and relationships with others that make particular conflictual relationships between groups and group identities highly salient or place an emphasis on the universal application of moral principles of fairness and equality (Killen et al., 2016).

In addition, to the just mentioned, it is worth pointing at Bar-Tal (2013) who underlines the socio-psychological repertoire (i.e., collective memories and an ethos of conflict, and collective fear, hatred and anger orientations) which makes intractable conflicts especially difficult to resolve. These socio-psychological structures paradoxically work both to enable better adaptation to the conflict conditions, and also to help maintain and prolong the conflict. Thus, even when the parties to the conflict find a peaceful resolution to the existing conflict the socio-psychological repertoire does not change overnight. For it to change, Bar-Tal states, what is needed is a long process of peace building, which requires thoughtful planning and active efforts to overcome the narrow vision which have evolved and which exclude incongruent information and alternative approaches to the conflict.

Though, relatively less identified with the educational sphere, social cohesion and acculturation theories can also help substantiate the work of peace educational initiatives. Space limitations prevent us from expanding on them here but the interested reader will benefit from reviewing literature on social cohesion as it relates to social exclusion and inclusion (Sobhi Tawil and Harley, 2004; Tawil et al., 2003), integration (Green et al., 2011; Green et al., 2009), and critical perspectives (Green et al., 2009; Soudien, 2012). Issues related to acculturation can be pursued reading, among others, Berry (1997, 2005), Esser (2010), and Rudmin (2003).

Peace education in Israel

The educational settings

Intergroup encounters within the Israeli context have been implemented from the early 1950's. Yet, encounters with a specifically educational focus designed to overcome hostility and contribute to coexistence were initiated in the 1980s (Helman, 2002) in response to the publication of a survey (Zemach, 1980) that revealed anti-democratic attitudes among Jewish Israeli youth toward the Palestinian-Arab minority. This disclosure generated fears that Israeli society might reject its democratic character (Katz and Kahanov, 1990) and brought about the formation of what Rabinowitz (2000) defines as the "coexistence sector," which focused on the development of activities geared toward building coexistence (Maoz, 2009; Suleiman, 2004). During the 90s, following the Oslo accords, coexistence dialogs focusing on the conflictual relations among the national groups expanded notoriously under the hospices of dedicated NGOs (Bar et al., 1995).

The above initiatives which can be considered as belonging to alternative or informal educational efforts have been complemented since the 1980's with formal ones in the shape of integrated schools. The first integrated Israeli Jewish Palestinian school was created at Neve Shalom/Wahat Al-Salam in 1984. Toward the end of the 90's dedicated NGO's followed this new direction and today there are eight such schools serving a population of approximately 2480 students (Meshulam, 2019).

A rather recent development has been the creation, in 2012, by the Ministry of Education of the Shared Education program in mixed Israeli cities (there are five such cities in Israel - Haifa, Jaffa, Lod, Akko and Ramle - of which Ramle and Jerusalem are the main beneficiaries of the program). The shared education program offers Arab and Jewish neighboring schools in Israel the opportunity to share classes together on a variety of disciplinary subjects (Payes, 2017).

Lastly, a few programs emphasize active educational approaches - sports, the arts, and social activism - rather than cognitive ones. Since 2001 "Football 4 Peace" has used sport to deliver values based training which aims at endorsing respect, responsibility, inclusion, neutrality, and equality (Sugden, 2008). "Peace Child Israel" was founded in 1988 to teach coexistence using theater and the arts. Their programs bring together Palestinian and Jewish teens from Israeli schools to meet weekly for 8 months before creating original dramas about coexistence and its challenges (Ross, 2015). The Sadaka-Reut program was founded in 1983, with the aim of educating and empowering through uni-national and bi-national programs Jewish and Palestinian Israeli youth and university students to pursue social and political change through binational partnership (Ross, 2015). Seeds of Peace is a peacebuilding and leadership development organization founded in 1993 and headquartered in New York City. Seeds of Peace conducts yearly a summer camp program for Israelis and Palestinians (other groups in conflict are represented in smaller numbers e.g., Indians and Pakistanis) which necessitates the complete immersion of youth in an interactive exchange with youth from the other side of the

conflict, as well as with the predominantly American staff (Hammack, 2006). The Parents Circle - Families Forum (PCFF) established in 1995 is a joint Palestinian Israeli organization of over 600 families, all of whom have lost a close family member as a result of the prolonged conflict. The organizations conduct a variety of educational projects with both adult and younger populations geared toward building trust and supporting reconciliation efforts (Kuriansky, 2007). The Middle East Entrepreneurs of Tomorrow (MEET <http://meet.mit.edu/>) is a three-year program developed in partnership with the Massachusetts Institute of Technology which brings together young leaders through the common language of technology, entrepreneurship and leadership (Ramadi et al., 2019).

For whatever their organizational context all of these initiatives in one way or another consider their work as evolving from the theoretical developments described in the previous section. This is not to say that they developed their programs based on these theoretical perspectives nor on the results of research produced by academic circles. Yet these theoretical perspectives, in one-way or another, serve as reference points for the activities developed or, at least, to rhetorically justify them.

Research findings

As a rule research on the effects of educational initiatives toward coexistence, conflict resolution, and peace are limited and not definitive (Salomon, 2004; Smith, 2010) (for some more positive results see Goldenberg et al., 2017; Ross, 2019; Shani and Boehnke, 2017), Israel is no exception (Abu-Nimer, 2004; Bar-Tal and Rosen, 2014; Maoz, 2011). In this section, we first report on research findings on mid/short term intergroup encounters, and then we move to research findings on long term encounter initiatives (e.g. the integrated bilingual schools).

During the 80s, studies were conducted on a variety of short term intergroup encounter programs through a variety of quantitative and action research methods (Hertz-Lazarowitz et al., 1998). These studies replicated previous findings from the 70s on casual meetings between Palestinians and Jews in Israeli universities (Hertz-Lazarowitz et al., 1998) which indicated that while the groups involved expressed a high level of readiness for social contact, the Palestinian-Arab group's desire was consistently higher than that of the Jewish group.

Studies conducted by Bar-Natan (cited in Salomon, 2004) and Maoz (2000) on intergroup encounters in the post-Oslo era found that youth from both groups arrived at the encounters with a limited acquaintance of each other and holding mutually negative stereotypes. The qualitative analysis undertaken showed that after participating in the workshops, each group's perceptions of the other significantly improved according to dimensions such as "considerate" and "tolerant," indicating that transformative practices can be effective in spite of a harsh socio-political context. Yet, in Bar-Natan's study, while for the Jewish group the encounters experience seemed to have also contributed to the legitimization of the Palestinian group narrative, the same effect was not found for the Palestinian cohort. Biton and Salomon (2006) and Maoz and Ellis (2008) found that participants in intergroup encounter programs became more positive about compromise solutions when compared with individuals who did not take part in such programs. Rosen (2006) examined the encounter settings for the possible differential changes in central and peripheral beliefs, as well as the durability of the impact. The study showed that peace programs could efficiently influence youngsters' peripheral attitudes and beliefs (e.g. stereotypes, prejudices, and negative emotions).

Recent qualitative critical studies show less reason for optimism. Studies that tracked university students participating in intergroup encounter courses revealed the dominance of Jewish participants who focused on the interpersonal level of the dialog while avoiding confrontation with the Palestinian-Arabs on issues related to conflict and national identity (Halabi and Sonnenschein, 2004). Helman's (2002) interpretative analysis showed that when intergroup encounters were positioned in contexts of domination and structural inequality, the dialogs tended to reproduce monological discourses of culture and identity, turning the dialogs into tools that ultimately legitimized power differentials and structural inequality. Similarly, Bekerman (2002) examining the discursive resources used by participants in dialog encounters, observed that national majority-minority rhetoric was shaped within the context of the nation-state, which also seemed to guide the encounters' communicational exchange, casting the participants' ethnic/national identities in an essentialist framework. Maoz et al. (2004) illustrated the dynamics of "good enough" dialogs that fulfill the basic purpose of an intervention in contexts of conflict, i.e. one in which different and divergent voices actually can be heard. Hammack's (2010) study of a group of Israeli Jews, Palestinian-Arabs (Israeli citizens), and Palestinians from the Palestinian Authority showed mixed results. On the one hand, participants were able to construct new narratives, transcending the initial polarized identity discourse that had sustained and reproduced the conflict. On the other hand, identity accentuation was seen among the participating adolescents for as long as two years after the encounter.

A longitudinal study conducted by Litvak-Hirsh and Bar-On (2007) described the results of a four-year qualitative study of Jewish and Palestinian-Arab university students who took part in a year-long encounter group workshop based on a narrative/life story model. Interviewees in both groups agreed that the narrative/life story model contributed to their personal enrichment and to the creation of a positive and favorable communication on the interpersonal and group levels. The interviews also demonstrated a growing complexity in the way participants to the dialog perceived each other. Ross (2016, 2019) in a comparative study of the Peace Child and Sadaka Reut programs concluded that for intergroup encounters to provide encouragement for continued participation in social change activities they need offer opportunities for critically examining the role of one's in-group within conflict contexts.

Hager et al. (2011) studied a college Jewish Arab dialog model recognized as part of a bachelor degree requirement within the Department of Education. The program, which was successful in turning the campus into a place which advances equal opportunities and dialog between the national groups, encouraged reflexive identity study and the examination of existing power structures,

and directed participants toward social activism. David et al. (2017) using both qualitative and quantitative methods studied Jewish-Israeli undergraduate students who underwent a yearlong process learning about the Israeli–Palestinian conflict, the Palestinian narratives, and the ways in which the Palestinian “other” impacts on their own identity. The study showed that participants developed an increased capacity for acceptance of both Israeli and Palestinian collective narratives, and demonstrated a greater willingness toward reconciliation.

Golan and Shalhoub-Kevorkian (2014) studied regularly accredited action-learning courses taught by joint Palestinian–Jewish faculty teams, and in which both Palestinian and Jewish students participated in them. Their research reveals that the participants, though well aware of the political situation and social inequalities, were reluctant to engage in political discussion or activism. The authors suggest that when neither the faculty nor the students in these courses are willing to challenge the hegemonic silencing of political issues pertaining to the conflict in Israeli academic institutions much of the potential benefits of the mixed programs are lost.

We move now to research results related to long-term educational initiatives. Research on the integrated bilingual schools, though less extensive, has added educational considerations and perspectives into the research scene which has been traditionally dominated by psychological perspectives (Bekerman, 2007). Feuerverger (2001) qualitative study highlights the complexities with which integrated bilingual/binational education must contend in its attempt to respect differences, sustain dialog, and inspire a moral vision. The study also reveals how an ever-changing political reality constantly frustrates the participants’ efforts to reshape the boundaries that divide their existence; making political contexts central to an understanding of transformative educational initiatives.

Glazier’s (2003) ethnographic study focused on the ability of children in one of the Jewish/Palestinian schools to step back and forth between cultures, a skill she calls “cultural fluency.” Glazier’s study underlines that contact alone does not always promote cultural fluency; rather, curriculum and pedagogy needs to guide individuals to engage in ongoing, meaningful shared tasks across borders. Mark (2013), investigated classroom participation and interaction that might characterize and differentiate the work of Jewish and Palestinian children and teachers. Marc concluded that although distinctive classroom discourse patterns could be identified in the Jewish and Palestinian uni-national discussions (sessions in which the groups study separately), the patterns were associated not with different ethnicities but with different socioeconomic groups.

The influence of language on discourse patterns and conceptions of identity was explored by Schlam-Salman and Bekerman (2011) in their examination of how students in a Palestinian/Jewish integrated school defined their identities in an advanced English-learning group. The study showed that the students’ use of a third language enabled them to step outside of ideologies that are culturally embedded in Arabic and Hebrew, and that the discussion in English provided the students with resources that influenced the ways in which they constructed their identities.

Meshulam’s (2011) study extended the research lens beyond the Israeli context. In a comparative study between a Spanish-English bilingual school in Wisconsin and a Palestinian-Jewish school in Israel, Meshulam found that the Israeli school’s attention to the binational encounter constricts its critical perspective. He noted that criticism of oppression in the Israeli school was limited to lines of national affiliation and that discussions did not challenge racism and oppression. In Meshulam’s interviews with teachers in the Israeli school, only one teacher brought up the topic of racism. In the Wisconsin school, on the other hand, multiculturalism was explicitly connected to antiracism. There Meshulam found a critical approach “... which expands culture to denote any category of oppression and marginalization. Under this conception, culture also encompasses gender, sexuality, ability, and so on” (p. 262).

Bekerman has conducted research on Jewish – Palestinian integrated bilingual schools for over two decades. He examined the connection between power relations in Israeli society and the difficulties of creating a truly bilingual educational program for Jews and Palestinians in Israel (Bekerman, 2005), and, how the different status of Hebrew and Arabic in Israeli society influences each group’s motivation to acquire the language of the other (Bekerman, 2009a). The practical importance of Hebrew language acquisition is clear to Arab children and to their families. As a minority group, Palestinians need Hebrew to advance academically and professionally and they regularly require Hebrew to communicate on the street. On the other hand, Jewish parents generally hope that their children will learn Arabic, but there is no apparent price that the children will pay if they fail to acquire the language. Without a practical need for the language, the Jewish pupils’ level of Arabic is generally far below the Palestinians’ level of Hebrew, despite the great educational effort invested in the bilingual program. Amara (2007) presented similar findings regarding the place of and the challenge of Arabic language acquisition. Bekerman and Tatar (2009) demonstrated how the different social realities of Jews and Palestinians influence the families’ motivation to send their children to the Jewish-Palestinian integrated schools.

Bekerman (2003) has also examined how the multicultural goals of the schools shapes religious and national narratives and shown how parents and teachers see culture and religion as areas in which mutual understanding can help bridge the gaps that separate the populations in Israel. Parents emphasize getting to know and understand the others’ culture more deeply and believe that the schools are achieving this goal. Teachers emphasized similar goals and educational activities/celebrations around these issues appear to be conducted with ease and in fruitful collaboration. These celebrations carry a strong religious emphasis. In fact, it could be said that religious aspects are disproportionately emphasized given that the majority of the Jewish parents belong to secular sectors of Israeli society and the Muslim populations, though more traditional, are also mostly non-religious. While at times, Jewish parents expressed concerns and ambivalence about this religious emphasis, they also seemed to find solace in the religious underpinning of cultural activities given their (mostly unarticulated) fear that their children’s Jewish identity will be eroded as a result of participation in an integrated binational program.

The ethnographic data gathered also suggests that issues of national identity (Bekerman, 2009b) and historical narratives have become the ultimate educational challenge for parents and educational staff alike (Bekerman and Zembylas, 2012). National issues are compartmentalized into a rather discrete period in the school year corresponding, in the Jewish Israeli calendar, to Memorial

Day and Israel's Independence Day and, in the Palestinian calendar, to the Day of the Nakba. In accordance with the policy of the Ministry of Education, all schools hold a special ceremony for the Jewish cohort on Memorial Day, which the Palestinian cohort does not attend. Depending on the schools' (complex) relations with the surrounding community and the Ministry of Education's supervision, a separate ceremony is conducted for the Palestinians in commemoration of the Nakba. For the Palestinian group, tensions are apparent, particularly among the teachers, who see themselves at the forefront of the struggle to safeguard the Palestinian national narrative, which remains unrecognized by Israeli educational officialdom. For most liberal Jews, Israeli Palestinian cultural and religious expression in school is legitimate. However, national identification with the Palestinian Authority is not welcomed, and neither are perspectives that would, in any way, try to deny the right of Israel to be a Jewish state.

A recent study on graduates of the school (Bekerman, 2018) has shown that in line with recent research conducted at bilingual and integrated educational initiatives (Cameron and Rutland, 2008; Wright and Tropp, 2005) the Palestinian and Jewish graduates show decreased in-group favoritism and higher perceived similarities between the groups. Exposed, as they are, to high levels of intercultural education and interethnic contact within the school setting they show, if at all, low levels of stereotyping and discrimination supporting previous findings in multicultural and anti-bias research (Banks, 2006). The graduates, for the most part, have little romantic expectations from education; they realize its limitations and yet profoundly appreciate what the integrated school has afforded them. Palestinians value the pragmatic benefits of what an integrated school has to offer them relative to the benefits of participating in the regular Arab educational track in Israel and Jews, though complaining at times about the difficulties of participating in a setting which in a paradoxical sense (at least at the high school level) reverses the asymmetry which benefits them in the wider Israeli context, stay assured that their parents' choice has profited them significantly.

All seem to have come to realize that belonging to one's group does not necessarily imply the denial of the other group and that acknowledging the existence of multiple perspectives does not necessarily need to be interpreted as renouncing one's own, but might enrich one's repertoire and open a path to dialog.

Finally, the research shows school activities, at the intergroup level, to be working well. While knowing and clearly recognizing their own ethnic/religious/national affiliation, the children seem able to create and sustain social interactional spheres where identity is not necessarily addressed. This ability of children stands in sharp contrast to the adult stakeholders' tendency to adopt a purely categorized identity approach, based on the premise that strengthening ethnic and national identities is the path to achieving their aims. The study suggests that the adoption of a categorized approach needs to be critically considered and revised if the schools do not want to replicate the discourse of the reification of rigid identities, which are central to the present conflict (Bekerman, 2000, 2009).

As the research shows, the bilingual school initiative is comprised of multiple, overlapping facets that must be viewed in concert. Moreover, even perspectives that strive to be critical often overlook alternative explanations or crucial processes that might open doors to potentially successful educational strategies. The path toward reconciliation, tolerance, and recognition in conflict-ridden societies presents difficulties that cannot easily be overcome.

Critical perspectives and more

A variety of scholars have expressed skepticism regarding the possible influence of peace education initiatives in general (Bajaj, 2018; Hajir, 2019; McKeown and Dixon, 2017) and those conducted among Jews and Palestinians in particular (Bekerman, 2007; Maoz, 2011; Schimmel, 2009; Silberberg, 2019). Though, these initiatives might have some short-term effects on changing intergroup attitudes they seem to have failed to influence attitudes in society at large. Among the reasons mentioned are the short-term exposure that most encounters afford participants, the lack of follow up activities build into these programs, the fact that participation in them are usually self-selected (Schimmel, 2009), their essentializing and utopian tendencies (Silberberg, 2019), and their lack of criticality (Bajaj, 2015). It has also been argued that group encounters are for the most part detached from participants' daily lives and focused on rather essentialized perspectives of culture and a static understanding of identity while dealing almost exclusively on interpersonal issues. Such an approach, is stated, helps avoid any confrontation with deep structural and institutional asymmetries between groups thus possibly contributing to diminish the extent to which social injustice is acknowledged, rejected and challenged (Bekerman, 2007; Dixon et al., 2013; Saguy and Kteily, 2014).

The above limitation, important as they are, seem to focus narrowly on peace and coexistence educational initiatives and their practice; as if these critiques would be carefully accounted for, the said education could achieve its goals.

Adding educational concerns to critical perspectives

Bettering peace education in its multiple varieties might be indeed possible; but we believe this change, to be achieved, needs to take into account more foundational issues, which are usually not accounted for. The issues relate to wider educational concerns, more specifically to modern massive state education being guided by functional, psychologized, and idealistic perspectives (Bekerman and Zembylas, 2018). Thus, guided these educational initiatives echo modern white, western, totalizing conceptualizations that identify the individual mind as the locus of an illness (conflict) which needs to be treated. The treatment is offered to (universalized) individuals while ignoring their socio historical contexts. Doing so replicates premises that are constituted and constitutive of the modern Western world, under which many of the conflicts that coexistence education is expected to help smooth and ultimately overcome have flourished, bringing us to wonder if these premises are the right basis for peace education.

Within this paradigmatic perspective peace and reconciliation if properly managed and applied could eradicate violence and conflict from the face of the world. This approach bears remarkable similarities to the paradigmatic dichotomies set by Western epistemology (male/female, good/evil, us/them) and as such seems to be able only to replicate past outcomes. Inasmuch as there is a “true” way all need to follow, the understanding, recognition, and dialog with alterity becomes a difficult task (Biesta, 2004). In the Western tradition, differences need to first be pointed at, and then assimilated or destroyed; denied differences are the secret of “our” (good, right) existence; and what is more important, differences are set in the realm of meaning and not in the realm of power relations.

When peace education is regarded everywhere as an universal ideal it disregards the social arrangements which institutionalize inequality and injustice, avoiding the problem of questions such as who “we” are, what perceptions of justice we hold to, what dialog we want to sustain, and under which conditions. If peace education is not just a rhetoric of affirmation for the recognition, and rehabilitation of “alterity” it needs to critically approach the epistemological and metaphysical certainties of Western modernity (Bekerman, 2007). It needs to confront the Western construct of alterity which seems inevitably to demonize those who are not like us, those who do not comply with the hegemonic standards of white western males – the ones who have successfully replace the force of arms with the coercion of currency and consumerism.

Adding methodological concerns to critical perspectives

Any critical approach to peace education will also need to consider the main product of political modernity the nation state, which has radically changed past traditional socio-political orders by implementing new forms of territoriality and novel surveillance capabilities that monopolize effective control over social relations across time–space distances and over the means of violence. The nation state exercises its control over the population comprising it by establishing a culture that is at once homogenizing, universal, and anonymous (all members share the culture the state has canonized). Moreover, nation state erases all individual and group affiliations (church, family, tribe, etc.) not allowing them to stand between state and citizen. Given the above, peace education will have to acknowledge what Michael Mann (2004) has identified as the dark side of democracy – ethnic cleansing. Ethnic cleansing, product of the Western inclination to confound into one *ethnos* and *demos* the two concepts inherited from classical Greece as the pillars of its democratic states. Modernity has added “ethnos” (a group that shares a common heritage) to “demos” (the rule of the people) thus allowing for “the people” to rule democratically but also “tyrannically” over any minority in the midst of its nation state. Along similar lines, Louis Dumont (1966) has argued that racism is a correlate of liberal democracies for if, as its credo goes, “All men are created equal,” then the evidence of inequality requires the dehumanization of the many. Equality from this perspective is a quality of man’s “nature,” not of the context within which he evolves.

The elements mentioned carry implications for social sciences research in general and more specifically for educational research. Research in these disciplines needs to acknowledge their unfortunate contribution to the naturalization to the global regime of nation-states thus surrendering their analytical scope to what Wimmer and Schiller (2003) have christened as methodological nationalism. They will also need to review their epistemological colonialism (Poulter, 2012); so as for researchers to better understand how their secular rational underpinnings bias their judgment and appreciation of traditional cultural/religious phenomena which in no small part is the phenomena they are trying to understand. Confusing political or ideological concepts with analytical ones can never be a good path to serious research. States are not natural entities, societies are not necessarily countries organized as states, minorities/immigration are not the flow in or between nations, and identities and cultures are not traits of solipsistic individuals. Taking them as such blinds research to the profound influence of socio-political contexts and processes in shaping present conflictual realities. Focusing on (created) units of analysis might be justifiable when approaching the study of complex realities but when these analytical units blind us to the wider spheres needed to understand complex processes they lose their relevance. As the political context of the nation state is not bounded, the individual and his or her culture are not bounded either; when the social sciences uncritically surrender to conceptualizations such as identity and culture, without attending to the potential dangers of naturalizing folkloristic concepts while embedding them in analytical discourses, they risk hiding the phenomena they intend to uncover.

Conclusion

In the last 70 years or so, peace education has emerged as a field of practice and research global in scope. In the first sections of this rather short piece I have offered a review of the historical and conceptual development of peace education, anchoring it in the Israeli reality, hoping that doing so offers practical insights into the potential benefits and challenges peace education confronts. In the last section I hope to have added to traditional critical perspectives some foundational, epistemological and methodological, issues which if not considered might prevent peace education from developing more effective tools for research and practice in its way to achieving its vision for peace and justice.

Freeing the imagination to take new educational paths and/or research approaches might imply adopting the old Hypocritical adagio “cura te ipsum” (take care of yourself first) while struggling to confront our paradigmatic perspectives so as to expose and try to overcome the structures and practices that have established the present conflict. Even if this is done it would be good to remember that the long standing and bloody conflicts that peace/coexistence educational initiatives hope to remedy are grounded in and sustained by the very material unequal allocation of resources more than in the heads of troubled individuals.

References

- Aboud, F., Mendelson, M., Purdy, K., 2003. Cross-race peer relations and friendship quality. *Int. J. Behav. Dev.* 27 (2), 165–173.
- Abu-Nimer, M., 2004. Education for coexistence and Arab-Jewish encounters in Israel: potential and challenges. *J. Soc. Issues* 60 (2), 405–422.
- Allport, G.W., 1954. *The Nature of Prejudice*. Addison-Wesley, London.
- Amara, M., 2007. Teaching Hebrew to Palestinian pupils in Israel. *Curr. Issues Lang. Plann.* 8 (2), 243–254.
- Amir, Y., 1976. The role of intergroup contact in change of prejudice and ethnic relations. In: Katz, P.A. (Ed.), *Towards the Elimination of Racism*. Pergamon, New York, pp. 245–308.
- Bajaj, M., 2015. 'Pedagogies of resistance' and critical peace education praxis. *J. Peace Educ.* 12 (2), 154–166.
- Bajaj, M., 2018. Conceptualizing transformative agency in education for peace, human rights, and social justice. *Int. J. Hum. Rights Educ.* 2 (1), 13.
- Bajaj, M., 2019. Conceptualising critical peace education for conflict settings. *Educ. Conf. Rev.* 2, 65–69.
- Banks, J.A., 2006. *Cultural Diversity and Education: Foundations, Curriculum, and Teaching*, fifth ed. Pearson/Allyn and Bacon, Boston.
- Bar, H., Bargal, D., Asaqla, J., 1995. *Living With Conflict: Encounters Between Jewish and Palestinian Israeli Youth*. Jerusalem Institute for Israeli Studies, Jerusalem.
- Bar-Tal, D., Rosen, Y., 2014. Conditions for the development of peace education in societies involved in intractable conflict. In: Berliner, D.C., Kupermintz, H. (Eds.), *Fostering Change in Institutions, Environments, and People: A Festschrift in Honor of Gavriel Salomon*. Routledge, New York, NY, pp. 231–250.
- Bar-Tal, D., 2013. *Intractable Conflicts: Socio-Psychological Foundations and Dynamics*. Cambridge University Press.
- Bekerman, Z., Tatar, M., 2009. Parents' perceptions of multiculturalism, bilingualism, and peace education: the case of the Israeli Palestinian-Jewish bilingual primary schools. *Eur. Early Child. Educ. Res. J.* 17 (2), 171–185.
- Bekerman, Z., Zembylas, M., 2012. *Teaching Contested Narratives: Identity, Memory and Reconciliation in Peace Education and beyond*. Cambridge University Press.
- Bekerman, Z., Zembylas, M., 2018. Psychologized Language in Education: Denaturalizing a Regime of Truth. Palgrave Macmillan, New York, NY.
- Bekerman, Z., 2000. Dialogic directions: conflicts in Israeli/Palestinian education for peace. *Intercult. Educ.* 11 (1), 41–51.
- Bekerman, Z., 2002. The discourse of nation and culture: its impact on Palestinian—Jewish encounters in Israel. *Int. J. Intercult. Relat.* 26 (4), 409–427.
- Bekerman, Z., 2003. Reshaping conflict through school ceremonial events in Israeli Palestinian—Jewish coeducation. *Anthropol. Educ. Q.* 34 (2), 205–224.
- Bekerman, Z., 2005. Complex contexts and ideologies: bilingual education in conflict-ridden areas. *J. Lang. Ident. Educ.* 4 (1), 1–20.
- Bekerman, Z., 2007. Rethinking intergroup encounters: rescuing praxis from theory, activity from education, and peace/co-existence from identity and culture. *Peace Educ.* 4 (1), 29–41.
- Bekerman, Z., 2009. Identity work in Palestinian—Jewish intergroup encounters: a cultural rhetorical analysis. *J. Multicult. Discourses* 4 (2), 205–219.
- Bekerman, Z., 2009a. On their way somewhere: integrated bilingual Palestinian Jewish education in Israel. In: Woods, P., Woods, G. (Eds.), *Learning from Alternative Educations: Priorities for Promoting Democracy and Holistic Human Growth in Contemporary Education Systems*. Palgrave, New York, NY.
- Bekerman, Z., 2009b. Social justice, identity politics, and integration in conflict ridden societies: challenges and opportunities in integrated Palestinian-Jewish education in Israel. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, New York, NY, pp. 138–151.
- Bekerman, Z., 2016. *The Promise of Integrated Multicultural and Bilingual Education: Inclusive Palestinian-Arab and Jewish Schools in Israel*. Oxford University Press, New York.
- Bekerman, Z., 2018. The graduate (s): the harvests of Israel's integrated multicultural bilingual education. *Race Ethn. Educ.* 21 (3), 335–352.
- Berger, R., Benatov, J., Abu-Raiya, H., Tadmor, C.T., 2016. Reducing prejudice and promoting positive intergroup attitudes among elementary-school children in the context of the Israeli—Palestinian conflict. *J. Sch. Psychol.* 57, 53–72.
- Berry, J.W., 1997. Immigration, acculturation, and adaptation. *Appl. Psychol.* 46, 5–68.
- Berry, J.W., 2005. Acculturation: living successfully in two cultures. *Int. J. Intercult. Relat.* 29 (6), 697–712.
- Biesta, G., 2004. The community of those who have nothing in common: education and the language of responsibility. *Interchange* 35 (3), 307–324.
- Bigler, R.S., Liben, L.S., 2006. *A Developmental Intergroup Theory of Social Stereotypes and Prejudice*.
- Biton, Y., Salomon, G., 2006. Peace in the eyes of Israeli and Palestinian youths: effects of collective narratives and peace education program. *J. Peace Res.* 43 (2), 167–180. <https://doi.org/10.1177/0022343306061888>.
- Brantmeier, E.J., 2011. Toward Mainstreaming Critical Peace Education in US Teacher Education. *Critical Pedagogy in the 21st Century: A New Generation of Scholars*, pp. 430–445.
- Brewer, M.B., Ho, H.-K., Lee, J.-Y., Miller, N., 1987. Social identity and social distance among Hong Kong schoolchildren. *Pers. Soc. Psychol. Bull.* 13 (2), 156–165.
- Brock-Utne, B., 1985. *Educating for Peace: A Feminist Perspective*, vol. 1. Pergamon Press, New York.
- Cameron, L., Rutland, A., 2008. An integrative approach to changing children's intergroup attitudes. In: Levy, S., Killen, M. (Eds.), *Intergroup Attitudes and Relations in Childhood Through Adulthood*, pp. 191–203.
- David, Y.B., Hameiri, B., Benheim, S., Leshem, B., Sarid, A., Sternberg, M., Sagy, S., 2017. Exploring Ourselves Within Inter-group Conflict: The Role of Intra-group Dialogue in Promoting Acceptance of Collective Narratives and Willingness Towards Reconciliation.
- Dixon, J., Durrheim, K., Kerr, P., Thoma, M., 2013. 'What's So Funny 'Bout Peace, Love and Understanding?' Further Reflections on the Limits of Prejudice Reduction as a Model of Social Change.
- Dovidio, J.F., Gaertner, S.L., Kawakami, K., 2003. Special issue on intergroup contact. *Group Proc. Intergr. Behav.* 6 (1).
- Dumont, L., 1966. Caste, racism and 'stratification': reflections of a social anthropologist. *Trans.* In: Sainsbury, M. (Ed.), *Homo Hierarchicus*. University of Chicago Press, Chicago.
- Esser, H., 2010. Assimilation, ethnic stratification, or selective acculturation? Recent theories of the integration of immigrants and the model of intergenerational integration. *Sociologica* 4 (1), 0–0.
- Fox, M.A., 2013. *Understanding Peace: A Comprehensive Introduction*. Routledge.
- Fredland, N.M., 2008. Nurturing hostile environments: the problem of school violence. *Fam. Comm. Health* 31, S32–S41.
- Gaertner, S.L., Dovidio, J.F., Samuel, G., 2000. Reducing Intergroup Bias: The Common Ingroup Identity Model. Psychology Press.
- Galtung, J., 1969. Violence, peace and peace research. *J. Peace Res.* 6 (3), 167–191.
- Galtung, J., 1973. Education for and with peace: is it possible? In: Galtung, J. (Ed.), *Essays in Peace Research*. Christian Ejlert, Copenhagen, pp. 317–333.
- Galtung, J., 2008. Form and content of peace education. In: *Encyclopedia of Peace Education*, pp. 49–58.
- Garrard, W.M., Lipsey, M.W., 2007. Conflict resolution education and antisocial behavior in US schools: a meta-analysis. *Conflict Resolut. Q.* 25 (1), 9–38.
- Glazier, J.A., 2003. Developing cultural fluency: Arab and Jewish students engaging in one another's company. *Harv. Educ. Rev.* 73 (2), 141–163.
- Golan, D., Shalhoub-Kevorkian, N., 2014. Community-engaged courses in a conflict zone: a case study of the Israeli academic corpus. *J. Peace Educ.* 11 (2), 181–207.
- Goldenberg, A., Endevelt, K., Ran, S., Dweck, C.S., Gross, J.J., Halperin, E., 2017. Making intergroup contact more fruitful: enhancing cooperation between Palestinian and Jewish-Israeli adolescents by fostering beliefs about group malleability. *Soc. Psychol. Personal. Sci.* 8 (1), 3–10.
- Green, A., Janmaat, J.G., Han, C., 2009. Regimes of social cohesion. In: *LLAKES Centre Research Paper Series: Centre for Learning and Life Chances in Knowledge Economies and Societies*.
- Green, A., Janmaat, G., Cheng, H., 2011. Social cohesion: converging and diverging trends. *Natl. Inst. Econ. Rev.* 215 (1), R6–R22.
- Hager, T., Saba, T., Shay, N., 2011. Jewish Arab activism through dialogical encounters: changing an Israeli campus. *J. Peace Educ.* 8 (2), 193–211. <https://doi.org/10.1080/17400201.2011.589255>.

- Hajir, B., 2019. Between Idealism and Realism: Critical Peace Education in Divided Post-Conflict Contexts.
- Halabi, R., Sonnenschein, N., 2004. The Jewish-Palestinian encounter in a time of crisis. *J. Soc. Issues* 60 (2), 373–387.
- Hammack, P.L., 2006. Identity, conflict, and coexistence: life stories of Israeli and Palestinian adolescents. *J. Adolesc. Res.* 21 (4), 323–369.
- Hammack, P.L., 2010. Narrating hyphenated selves: intergroup contact and configurations of identity among young Palestinian citizens of Israel. *Int. J. Intercult. Relat.* 34 (4), 368–385.
- Hantzopoulos, M., Bajaj, M., 2016. Conclusion: critical directions for peace education. In: *Peace Education: International Perspectives*, pp. 233–238.
- Harris, I., 2002. Conceptual underpinnings of peace education. In: Teoksessa, Salomon, G., Nevo, B. (Eds.), *Peace Education. The Concept, Principles and Practices Around the World*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Harris, I., 2010. History of Peace Education. *Handbook on Peace Education*, pp. 11–22.
- Helman, S., 2002. Monologic results of dialogue: Jewish-Palestinian encounter groups as sites of essentialization. *Identities* 9, 327–354.
- Hertz-Lazarowitz, R., Kupermintz, H., Lang, J., 1998. Arab-Jewish student encounter: beit Hagefen coexistence program. In: Weiner, E. (Ed.), *The Handbook of International Coexistence*. The Continuum Publishing House, New York, NY, pp. 565–584.
- Katz, I., Kahanov, M., 1990. A survey of dilemmas in moderating Jewish-Arab encounter groups in Israel. *Megamot* 33, 29–47 (in Hebrew).
- Killen, M., Elenbaas, L., Rutland, A., 2016. Balancing the fair treatment of others while preserving group identity and autonomy. *Hum. Dev.* 58 (4–5), 253–272.
- Kuriansky, J., 2007. Beyond Bullets and Bombs: Grassroots Peacebuilding Between Israelis and Palestinians. Greenwood Publishing Group.
- Lerch, J.C., Buckner, E., 2018. From education for peace to education in conflict: changes in UNESCO discourse, 1945–2015. *Glob. Soc. Educ.* 16 (1), 27–48.
- Litvak-Hirsh, T., Bar-on, D., 2007. Encounters in the looking glass of time: longitudinal contribution of a life story workshop course to the dialogue between Jewish and Arab young adults in Israel. *Peace Conflict Stud.* 14 (1), 19–42.
- Mann, M., 2004. *The Dark-Side of Democracy*. Cambridge University Press, Cambridge.
- Maaz, I., Ellis, D.G., 2008. Intergroup communication as a predictor of Jewish-Israeli agreement with integrative solutions to the Israeli–Palestinian conflict: the mediating effects of out-group trust and guilt. *J. Commun.* 58 (3), 490–507.
- Maaz, I., Bar-On, D., Bekerman, Z., Jaber-Massarwa, S., 2004. Learning about ‘good enough’ through ‘bad enough’: a story of a planned dialogue between Israeli Jews and Palestinians. *Hum. Relat.* 57 (9), 1075.
- Maaz, I., 2000. Multiple conflicts and competing agendas: a framework for conceptualizing structured encounters between groups in conflict - the case of a coexistence project of Jews and Palestinians in Israel. *J. Peace Psychol.* 6 (2), 135–156.
- Maaz, I., 2009. Educating for peace through planned encounters between Arabs and Jews in Israel: a reappraisal of effectiveness. In: Salomon, G., Cairns, E. (Eds.), *Handbook of Peace Education*. Psychology Press, New Jersey.
- Maaz, I., 2011. Does contact work in protracted asymmetrical conflict? Appraising 20 years of reconciliation-aimed encounters between Israeli Jews and Palestinians. *J. Peace Res.* 48 (1), 115–125. <https://doi.org/10.1177/0022343310389506>.
- Mark, B., 2013. Patterns of Students Engagement and Teacher Evaluation. Doctoral Dissertation. School of Education. Hebrew University of Jerusalem, Jerusalem, Israel.
- McKeown, S., Dixon, J., 2017. The “contact hypothesis”: critical reflections and future directions. *Soc. Personal. Psychol. Comp.* 11 (1), e12295.
- Meshulam, A., 2011. What Kind of Alternative?: Bilingual Multicultural Schools as Counterhegemonic Alternatives Educating for Democracy. University of Wisconsin–Madison.
- Meshulam, A., 2019. Palestinian-Jewish bilingual schools in Israel: unravelling the educational model. *Int. J. Educ. Dev.* 70, 102092.
- Minow, M., 2002. Education for co-existence. *Ariz. Law Rev.* 44 (1), 1–29.
- Nieto, S., 2017. Re-imagining multicultural education: new visions, new possibilities. *Multicul. Educ. Rev.* 9 (1), 1–10. <https://doi.org/10.1080/2005615X.2016.1276671>.
- Opatow, S., Gerson, J., Woodside, S., 2005. From moral exclusion to moral inclusion: theory for teaching peace. *Theory Into Pract.* 44 (4), 303–318.
- Payes, S., 2017. Education across the divide: shared learning of separate Jewish and Arab schools in a mixed city in Israel. *Educ. Citizen Soc. Justice* 13 (1), 19–35. <https://doi.org/10.1177/1746197917698489>.
- Pettigrew, T.F., Tropp, L.R., 2006. A meta-analytic test of intergroup contact theory. *J. Pers. Soc. Psychol.* 90 (5), 751–783.
- Poulter, S., 2012. From secular epistemology towards the pedagogy of ‘inter-knowledges’. *Muslim Educ. Q.* 25 (3/4), 19–30.
- Rabinowitz, D., 2000. Postnational Palestine/Israel? Globalization, diaspora, transnationalism, and the Israeli-Palestinian conflict. *Crit. Inq.* 26, 757–772.
- Ramadi, K.B., Srinivasan, S., Atun, R., 2019. Health diplomacy through health entrepreneurship: using hackathons to address Palestinian-Israeli health concerns. *BMJ Global Health* 4 (4), e001548.
- Reardon, B., 1988. *Comprehensive Peace Education*. Teachers College Press, New York.
- Rosenthal, L., Levy, S.R., 2010. The colorblind, multicultural, and polycultural ideological approaches to improving intergroup attitudes and relations. *Soc. Iss. Pol. Rev.* 4 (1), 215–246.
- Ross, K., 2015. Quality as critique: promoting critical reflection among youth in structured encounter programs. *J. Peace Educ.* 12 (2), 117–137.
- Ross, K., 2016. Peacebuilding through inter-group encounters: rethinking the contributions of ‘mainstream’ and ‘politicised’ approaches. *Peacebuilding* 1–14.
- Ross, K., 2019. Becoming activists: Jewish–Palestinian encounters and the mechanisms of social change engagement. *Peace Change* 44 (1), 33–67.
- Rudmin, F.W., 2003. Critical history of the acculturation psychology of assimilation, Separation, Integration, and Marginalization. *Rev. Gener. Psychol.* 7 (1), 3–37.
- Saguy, T., Kteily, N., 2014. Power, negotiations, and the anticipation of intergroup encounters. *Eur. Rev. Soc. Psychol.* 25 (1), 107–141.
- Salomon, G., 2004. Does peace education make a difference in the context of an intractable conflict? *J. Peace Psychol.* 10, 257–274.
- Schimmel, N., 2009. Towards a sustainable and holistic model of peace education: a critique of conventional modes of peace education through dialogue in Israel. *J. Peace Educ.* 6 (1), 51–68.
- Schlam Salman, J., Bekerman, Z., 2011. Emancipatory discourse? An ethnographic case study of English language teaching in an Arabic-Hebrew bilingual school. In: Nasser, I., Wong, S., Berlin, L. (Eds.), *Examining Education, Media, and Dialogue Under Occupation: The Case of Palestine and Israel*. Multilingual Matters, UK, pp. 89–106.
- Shani, M., Boehnke, K., 2017. The effect of Jewish–Palestinian mixed-model encounters on readiness for contact and policy support. *Peace Conflict* 23 (3), 219.
- Silberberg, R., 2019. The philosophical challenges of critical peace education in the Palestinian-Israeli context. *Ethics Educ.* 14 (2), 198–212.
- Singh, K., 2015. Concerns of school in nurturing peace education. *Int. J. Educ. Technol.* 5 (1 and 2), 9–13.
- Smith, A., 2010. The Influence of Education on Conflict and Peace Building, Background Paper Prepared for the Education for All Global Monitoring Report 2011 the Hidden Crisis: Armed Conflict and Education. UNESCO, Paris.
- Soudien, C., 2012. The modern seduction of race: whither social constructionism? *Transformation* 79 (1), 18–38.
- Stier, J., 2003. Internationalisation, ethnic diversity and the acquisition of intercultural competencies. *Intercult. Educ.* 14 (1), 77–91.
- Sugden, J., 2008. Anyone for Football for Peace? The challenges of using sport in the service of co-existence in Israel 1. *Soccer Soc.* 9 (3), 405–415.
- Suleiman, R., 2004. Planned encounters between Jewish and Palestinian Israelis: a social-psychological perspective. *J. Soc. Issues* 60 (2), 323–337.
- Tajfel, H., Turner, J., 1979. An integrative theory of intergroup conflict. In: Austin, W.G., Worchel, S. (Eds.), *In the Social Psychology of Intergroup Relations*. Brooks/Cole, Monterey, CA, pp. 33–47.
- Tawil, S., Harley, A., 2004. Education and Identity-Based Conflict: Assessing Curriculum Policy for Social and Civic Reconstruction. *Education, Conflict and Social Cohesion*, p. 9.
- Tawil, S., Harley, A., Porteous, L., 2003. Curriculum Change and Social Cohesion in Conflict-Affected Societies. Retrieved from Geneva.
- Vanbeselaere, N., 1996. The impact of differentially valued overlapping categorizations upon the differentiation between positively, negatively, and neutrally evaluated social groups. *Eur. J. Soc. Psychol.* 26 (1), 75–96.

- Wilder, D.A., 1986. Social Categorization: Implications for Creation and Reduction of Intergroup Bias. In: *Advances in Experimental Social Psychology*, vol. 19. Elsevier, pp. 291–355.
- Wimmer, A., Schiller, N.G., 2003. Methodological nationalism, the social sciences, and the study of migration: an essay in historical epistemology. *Int. Migrat. Rev.* 37 (3), 576–610.
- Wright, S.C., Tropp, L.R., 2005. Language and intergroup contact: investigating the impact of bilingual instruction on children's intergroup Attitudes. *Group Process. Intergr. Relat.* 8 (3), 309–328. <https://doi.org/10.1177/1368430205053945>.
- Zemach, M., 1980. Attitudes of the Jewish Majority in Israel towards the Arab Minority (Retrieved from Jerusalem).
- Zilliacus, H., Holm, G., Sahlström, F., 2017. Taking steps towards institutionalising multicultural education—The national curriculum of Finland. *Multicul. Educ. Rev.* 9 (4), 231–248.

Enhancing a Global South-North partnership to promote global justice in K-12 classrooms: a Botswana case study

Agreement L. Jotia^a, Karen L. Biraimah^b, and Brianna A. Kurtz^c, ^a Department of Languages and Social Sciences, University of Botswana, Gaborone, Botswana; ^b Department of Learning Sciences and Educational Research, University of Central Florida, Orlando, FL, United States; and ^c College of Education, Mary Baldwin University, Staunton, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: development of equitable Global South-North partnerships to support quality study abroad programs	104
Reflecting on two decades of research: the transformation of study abroad programs rewards and challenges	105
The need for intercultural competencies and pedagogy	105
Issues related to the Global South and North	105
Notions of decoloniality and “othering”	106
Developing two-way learning through sustainable partnerships	106
Reflecting on two decades of research: the need for sustainable and equitable Global South-North partnerships to enhance and support study abroad programs	106
The University of Botswana (UB) and the University of Central Florida (UCF) partnership	107
The challenge of over-committed faculty from the Global South	107
The challenge of linking collaborative partnerships to quality study abroad programs	107
Financial hurdles	108
Studying abroad while staying at home: a new partnership perspective	108
Theoretical perspectives on the impact of globalization, study abroad, and the global partnerships which sustain these initiatives	109
Internationalization and global competence: impact on culturally responsive pedagogy	109
Internationalization and global competency as perceived by the Global North and South	109
Colonialism, decoloniality and their impact on globalization	110
Critical cosmopolitan and social justice theories	110
Global citizenship	111
Moving from theory to practice	111
Research methodology	112
Participant demographics	112
Results	113
Basic knowledge acquisition	113
Participants’ most rewarding experiences	113
Participants’ most challenging experiences	114
Professional growth and development of educators through study abroad programs in emerging nations	114
Impact on participants’ professional development	115
Concluding remarks	115
References	116

Introduction: development of equitable Global South-North partnerships to support quality study abroad programs

The geo-political space and the diversity of cultures that exist between the Global South and Global North should be embraced in pursuance of education for co-existence of humanity. Institutions of higher learning therefore have a duty to train learners holistically beyond the boundaries of socio-economic and political disparities as well as beyond the limitations of the geo-political space. We have based this article on the concept that effective global teaching and service-learning experiences cannot come to fruition through unilateral plans that exclude input from and collaboration with host institutions. Thus, we focus initially on the absolute need (and challenges) of developing a durable and equitable partnership between universities in the Global South and Global North. To this end, we include a brief synopsis of “lessons learned” over a decade of mutual engagement to develop a truly equitable partnership between the University of Botswana (UB) and the University of Central Florida (UCF). In doing so, we explore the challenges embedded within such partnerships, as well as a path to develop durable, quality study abroad programs that can positively influence both the partner universities, and the host community. Our approach is that education should be advanced as a unifying factor between the Global South and Global North so that learners can have an appreciation of diverse cultures. Consequently, they can play instrumental roles at their professional as well as community levels by ensuring that appropriate lenses are applied when appreciating global multicultural diversity.

Reflecting on two decades of research: the transformation of study abroad programs rewards and challenges

Historically, study abroad programs have depended upon educational experiences designed to enable participants to focus on global issues and their responsibilities to contribute to the greater good (Kommers and Bista, 2021). For example, a decade ago scholarship (Biraimah and Jotia, 2012; Foronda and Belknap, 2012; Schroeder et al., 2009; Woolf, 2006) emphasized the need for study abroad programs to include effective communication skills, community engagement, and meaningful service-learning projects for all participants. Onyenekwu et al. (2017) also argued that previous promotional material for study abroad programs included depictions of host countries and cultures as “othered” (Caton and Santos, 2009), as well as the use of consumerism to envision programs through a tourism lens (Bishop, 2013; Woolf, 2006). This negativity was also represented within Western mass media where representations of the Global South often relied on poverty porn (Glick, 2015) and the White savior complex (Cole, 2012; Brown, 2013) which often prompted emotional responses from Global North participants that enhanced misconceptions of the developing world.

More recently, these foci, whether constructive or problematic, have begun to be replaced by new factors, such as the commercialization of study abroad, which seems to be reshaping program goals and outcomes, as well as impacting issues of equity and student mobility (Kommers and Bista, 2021). With an increase in the outsourcing of university study abroad programs to third-party vendors these scholars suggest that tourism, rather than educational purpose, is being promoted and learning goals are being replaced by marketing ploys that these programs will now provide participants with a competitive edge when entering the job market. Unfortunately, regardless of their end results, study abroad programs remain problematic for students from lower socioeconomic backgrounds, thus perpetuating and reproducing societal inequalities (Lorz et al., 2016).

Given this evolution of study abroad programs over the past few decades, the purpose of this study was to examine key factors, such as the need for intercultural competence, as well as the challenges presented when participants from the Global North descend upon their host countries and communities in the Global South; an experience which can either reflect continuing challenges to move beyond tourist perspectives and “othering,” or provide opportunities for the development of two-way learning and equitable partnerships. Finally, this study will examine the degree of professional growth and intercultural competence and attitudes that occurred during a 2017 study abroad program to Botswana.

The need for intercultural competencies and pedagogy

Recent research (Baecher, 2021; Pfingsthorn and Czura, 2017; Tripp et al., 2021) reflects an ongoing need and “compelling evidence for the creation of teacher study abroad programming that is intentionally conscious and critical of structural inequities and their connection to schooling” (Baecher, 2021, p. xvi). To this end, preservice and in-service teachers who participant in these programs enhance their intercultural competence while expanding critical perspectives and pedagogy informed by their global experiences and knowledge (Pfingsthorn and Czura, 2017). For Baecher (2021), there also exists an enhanced potential for international learning that promotes teachers’ global and critical perspectives regarding the potential impact of their roles as educators with increasingly culturally diverse classrooms. Baecher’s (2021) research documents the “transformative learning experiences which impact the way teachers think about learning, teaching, and identity” (p. xiv). Reinforcing these perspectives, Tripp et al. (2021) suggest that “the study abroad experience adds value in terms of cultural and language experiences, thus resulting in greater awareness of the diverse needs of students and their backgrounds” (p. 131). They continue by underscoring the need to provide more opportunities for preservice teachers to “develop greater awareness of and sensitivity to those whose backgrounds differ from their own [and they also posit a] critical need for data to inform international program developers about specific academic and cultural benefits and potential challenges associated with study abroad experiences” (p. 132). Recent research has clearly underscored the need for more culturally sensitive study abroad programs.

The authors analysis of a study abroad program in the emerging, Global South nation of Botswana provides insights and informs the design of sustainable study abroad programs within teacher education that supports the development of intercultural competence, while preparing teachers to provide quality instruction to their increasingly diverse students.

Issues related to the Global South and North

A few decades ago, traditional study abroad programs usually involved students and host countries within the Global North (such as American students studying in Italy). However, with expansion of globalization, programs increasingly provide opportunities for students from the Global North to participate in programs within the Global South (such as our study abroad program that provides American preservice teachers with cross-cultural experiences in the emerging nation of Botswana). With this broadening of opportunities also comes a need to develop sustainable programs that are mutually beneficial to both participant and host communities; programs where simply “helping” has been replaced by mutually beneficial partnerships between visitors and their host communities. This notion of helping, according to Tripp et al. (2021), is often linked to the “White Savior Complex” which calls upon participants to ask if this desire to “help” also meets the needs of their host community. (p. 136).

With this enhancement of study abroad programs to the Global South also comes the responsibility to move beyond tourist excursions to exotic destinations. However, program designers cannot simply replace these shallow excursions with assumptions that participants from the Global North have arrived to solve their host nation’s educational challenges. Unfortunately, even if study abroad programs in emerging nations avoid inappropriate or superficial philosophies, numerous and potentially negative

consequences remain. Jotia and Biraimah (2016) suggested that “Not only is there the problem of participants flaunting their ‘first world’ wealth and dominance over local communities, but there are real threats to a host nation’s environment, economy and cultures” (p. 94). For example, their presence could further deteriorate diminishing clean water supplies, or their obvious use of advanced technology could inflate consumerism within a rural, and relatively poor population.

Notions of decoloniality and “othering”

As the popularity of study abroad programs to the Global South increases, there are also added concerns regarding issues of decoloniality (the need to move beyond colonial perspectives), and the “othering” of host nationals by Global North participants. When programs extend beyond the usual Eurocentric destinations to include emerging nations, it also becomes essential to consider the unique challenges study abroad programs might have upon fragile social and economic environments, and the need to avoid “a kind of voyeurism in which privileged young Americans go to observe relative poverty in a developing country” (Woolf, 2006, p. 136). In these instances, it is necessary for participants to transform themselves from a helping/savior mentality to one that views their study abroad program as a learning experience that can benefit all involved. These new perspectives can enhance participants’ intercultural awareness, knowledge and competence while simultaneously enabling them to better understand and respect diverse cultural groups.

Tripp et al. (2021) have noted and built upon the scholarship of Von Wendorff (2013), that suggests when participants from the Global North first arrive in their Global South communities, there is often “A sense of separation and ‘othering’ [that] can create barriers to successful communication and understanding of the social climate in international settings” (Tripp et al., 2021, p. 131). This concept of “othering,” as explained by Powell and Menden (2018) is “a set of dynamics, processes and structures that engender marginality and persistent inequality across any of the full range of human differences based on group dynamics” (p. 17). In the case of study abroad programs, participants from the Global North often view their Global South counterparts as “others,” who do not necessarily share similar attitudes, aptitudes, knowledge or experiences. While the theories explaining “othering” will be explained in more detail below, it is of interest to note that recent research also suggests that study abroad participants from the Global North may also see themselves as “others,” at least at the beginning of their study abroad program. For example, Tripp et al. (2021) found that “Participants themselves felt like ‘the other,’ because of a lack of knowledge of social norms, as well as linguistic miscommunication” (p. 136). They also reported that preservice teacher participants in their Malawi study abroad program often saw themselves as the “outsider” when initially immersed within the culture and people of Malawi.

Developing two-way learning through sustainable partnerships

Tripp et al. (2021) have noted that study abroad participants from the Global North often initially reflect a “‘one-way’ attitude of coming to help the poor, African people,” instead of being more receptive to learning from their host communities (p. 140). To remedy this shortcoming, they underscore the need for their hosts to contribute to the program’s learning exchange, which they later observed led to participants shifting their perspectives to more of a two-way learning pattern while in Malawi.

While the need to establish solid, sustainable partnerships that support quality study abroad programs will now be addressed, it will also be suggested that we might need to expand our understanding of these partnerships, and the medium for these study abroad programs. As suggested by Joseph and Johnson (2021), during periods of travel restrictions, such as the current pandemic, educational leaders need to envision new “study abroad” experiences that omit traditional travel to distant locations, yet include virtual, sustained and more inclusive intercultural learning opportunities, at no additional cost.

Reflecting on two decades of research: the need for sustainable and equitable Global South-North partnerships to enhance and support study abroad programs

Moving beyond the importance of sustainable and transformative study abroad programs, we will now address the need for sustainable and transformative partnerships capable of supporting these programs, particularly when programs include Global South-North partnerships. We will also examine the impact of study abroad programs that are supported by these Global South-North partnerships. As the goal of developing sustainable, quality programs that positively impact both partner universities is often made more difficult due to varying expectations and miscommunications, these are also frequently exacerbated by socioeconomic disparities. As previously mentioned, partnerships that support vibrant study abroad programs may take multiple forms, but regardless of the medium, factors that undergird a successful, equitable, and sustainable partnership remain. While new forces and factors may occur over time, some issues raised nearly two decades ago often remain applicable to current challenges when building sustainable partnerships capable of ensuring quality study abroad programs.

For example, when analyzing the challenges of establishing an equitable Global South-North partnership, Wohlgemuth and Olsson (2003) suggested that the usual domination by the Global North be replaced by dialog, along with the encouragement of shared values. Moreover, while “capacity building” is often seen as an essential outcome of international partnerships, “power relationships (based on inequitable wealth) may become a significant barrier to communication and immersion,” and hinder anticipated positive study abroad program outcomes (Woolf, 2006, p. 142). Thus, engaging in such partnerships can only become productive when partner institutions respect one another within a common vision for advancing education within

cultural differences. These challenges remain concerns within current scholarship, as demonstrated by [Rose and Cooper-Duffy's \(2021\)](#) scholarship which reflects upon the critical role played by partnerships enabling a Fulbright-Hays program for Global North special educators and speech-language pathologists in Botswana. They emphasized that their successful Global South-North partnerships (between North Carolina and Botswana's Ministry of Basic Education) had to have a "bi-directional relationship where educators from both counties have participated in study abroad experiences to learn about teaching students with disabilities" (p. 149). They continued to explain the level of cooperation within the partnership that made their program successful:

The instructors cannot overstate the importance of the partnership created with educators and administrators in the host country and the support provided by the [Fulbright-Hays] grant. Many of the participants would not have had the opportunity for an international experience without grant funding. Partners in the host country took the lead in preparing for school visits, accommodations, transportation, meetings, and cultural activities. Additionally, several of our host country partners accompanied the group throughout the trip and assisted with navigating challenges that arose. This close partnership facilitated participant growth in cultural competence (p. 163).

It should also be noted that this partnership was truly "two-way," with US participants in Botswana during 2016, followed by reciprocal trips to North Carolina by a Botswana delegation of teachers and administrators in 2017. We now extend our examination of a continuing partnership through an in-depth look at the decade long challenges faced by two universities in the Global South and North.

The University of Botswana (UB) and the University of Central Florida (UCF) partnership

To review various theoretical perspectives regarding the development of equitable Global South and North partnerships, which in turn support viable Global South-North study abroad programs, we provide an in-depth look at one partnership between an institution in the Global South, the University of Botswana (UB), and one institution in the Global North, the University of Central Florida (UCF). To this end, we examine various challenges to maintaining a viable and equitable partnership including those emerging from (a) over-committed faculty in the Global South; (b) the need for a truly collaborative partnership; and (c) multiple financial barriers.

The challenge of over-committed faculty from the Global South

The strain on emerging universities due to heavy faculty teaching loads can negatively impact the quality of their research and service, including participation in Global South-North study abroad programs. While not a perfect solution, program directors scheduled their study abroad programs in Botswana during their long winter vacation (which coincided with Global North summer vacations). This one factor positively impacted UB participation by providing more time to devote to these programs, as well as concomitant research, publications, teaching practice supervision and future grant proposal submissions. Regarding study abroad programs for US students in Botswana, not only were they scheduled during months when most UB and UCF students and faculty were on break between semesters, but the partnership was designed to provide numerous opportunities for scholarly activities, which to date, have resulted in multiple joint publications and conference paper presentations ([Biraimah and Jotia, 2012, 2014, 2016, 2021](#); [Jotia and Biraimah, 2016, 2017, 2018](#); [Jotia et al., 2020, 2021](#)).

The challenge of linking collaborative partnerships to quality study abroad programs

Even when logistical challenges mentioned above are overcome, numerous problems related to establishing an equitable and sustainable partnership between institutions in the Global South and North remain. For instance, even perspectives regarding timelines and program details can challenge a partnership. As an example, logistical issues involving itineraries and time spent in rural communities and schools are often addressed from different levels of urgency by partners in the Global South and North. Program coordinators from UCF who were responsible for submitting final itineraries and budgets to their institution several months prior to the actual departure date may be driven by internal deadlines that are of less significance to their UB partners. More precisely, when developing a final budget, it is critical that actual travel locations and schools are identified, as these directly impact transportation and other related costs. However, when viewed from UB's perspective, these urgent demands for final itinerary and school destinations seemed unrealistic. For example, before our UB coordinator could identify with certainty which schools would participate in the study abroad program, multiple communication and travel challenges to various remote localities had to be overcome. As many rural schools had only unreliable cell phone connections, and usually no wi-fi, rapid communication was replaced by hours of driving over dusty, and often unpaved gravel roads to make personal contact with school heads to obtain the necessary permission. While all this was playing out in rural Botswana, the UCF counterparts waited impatiently for their phone calls or emails to be returned. In this case, the disparity in the socio-economic factors comes into play as a partner in the US can usually obtain necessary information rather quickly, while their counterpart in Botswana may have to wait for days to (hopefully) obtain internet access. Consequently, frustrations come to the fore as academics in the Global North may assume that their counterparts in the Global South are lax and less motivated, while in fact they are doing their very best, given travel and communication challenges.

This same erroneous assumption that Global North practices and expectations will automatically be reflected in Botswana often becomes quite clear the moment their students arrived in Botswana. It has been our experience that often the first comment or request uttered by UCF study abroad participants arriving in Botswana is “what is the wi-fi-password?” However, as program directors hoped, they soon began to grasp the shocking realities of global diversity and inequities when they are informed that wi-fi has not been available for days, is very slow, or that due to electrical outages they may have to wait hours, if not days, to catch-up with their online world.

Unfortunately, these unrealistic perspectives and expectations, as seen from the Global South, are often perceived as menacing and unrealistic, as well as displays of power and control from the Global North, with little regard for mitigating socio-cultural issues. Regrettably, this naivety on the part of Americans may be viewed as power plays by our Botswana colleagues, and can easily derail, if not eliminate future study abroad programs if left unattended. What remains crucial, however, is that the partnership successfully navigates these challenges effectively and from a shared perspective. However, these goals are often quite difficult to achieve, given that many learners from the Global North may not have been properly prepared for the realities in Botswana, and therefore may be perceived as entitled and privileged by those in the Global South. Thus, not having wi-fi, for example, may be viewed as an incomprehensible scenario, while those in the Global South view the same predicament as normal, remaining hopeful that at some point in time the electricity and the slow internet will return.

Financial hurdles

Programs funded through US government grants, such as the U.S. Department of State’s international grants and the U.S. Department of Education’s Fulbright-Hays programs, support a clearly uneven playing field as the US recipient is placed in a dominant financial role, though this inequity can also occur in smaller student-funded university study abroad programs. In these instances, the UB/UCF partnership has endured financial inequalities where UB may be allowed to manage a small portion of the budget, but the overall fiduciary responsibilities remain with UCF. This leaves the Global South partner in more of a dependent and/or observer role, as opposed to a partnership where both entities wield equal programmatic and financial decision-making. While in the above instances UB was allocated limited funds through sub-contracts (covering expenses such as lecturer fees, dormitory rentals, and program coordinator/student assistant stipends), these payments were usually based on a reimbursement model, which remains a challenge for emerging university budgets that may not have sufficient fiscal reserves necessary to manage lengthy repayment cycles. This lack of budgetary involvement also makes it more difficult for Global South institutions such as UB to take financial advantage of a grant’s lucrative overhead revenues, which are generally calculated based on the amount of actual spending by each institution. As UB had extremely limited spending authority regarding programs funded by federal grants, their ability to receive overhead revenues was also curtailed.

Beyond these unequitable funding issues, other factors may also keep a Global South-North partnership on unequal terms. These include complex and competing accounting procedures (Jotia and Biraimah, 2016), and differing perspectives regarding “due diligence” linked to receipts, currency conversions and auditors’ expectations. As Bradley (2008) posited, these asymmetrical relationships between Global South-North partners are often perceived as major obstacles to productive and collaborative research. This disproportionality influences Global South-North partnerships that include project administration, budget management, and the development of research agendas.

Studying abroad while staying at home: a new partnership perspective

Given the economic challenges regarding Global South-North partnerships and study abroad programs, along with current travel restrictions due to the pandemic, colleagues (Joseph and Johnson, 2021) in Australia have initiated a new form of study abroad while staying at home. This approach also addresses the issue of economic inequality linked to program participation, as lower socio-economic status (SES) students are often unable to participate in expensive study abroad programs. For those concerned about the current cycle of global warming, this option also provides a way to reduce carbon emissions due to airline travel.

Built on collaborative partnerships with universities in Australia, Spain, South Africa, Sri Lanka and New Zealand, Joseph and Johnson (2021) have developed the Internationalization at Home (IaH) program which provides “an alternative option of IaH at our respective workplaces [at] no cost to the university or to the student. We see this option as a fair, equitable and accessible opportunity for our students” (p. 141). While their concept may first appear to be an oxymoron, this “alternative to study abroad allowed for authentic learning experiences to take place through a blended environment (face-to-face and through technology) where our students were able to reflect on their intercultural understandings and learn from it across the semester” (p. 151). Moreover, this sustained experience, rather than the usual brief international visit that may seem more like a tourist excursion than an academic endeavor, provides for more intercultural experiences for students, at no additional cost. In the words of Joseph and Johnson (2021), “... our alternative option of internationalization at home proved to be a rich cultural and linguistic learning experience where the process of integrating an international and intercultural dimension was intentional” (p. 151). While those of us working with study abroad programs may not have considered this option, it certainly has merit and deserves more careful analysis.

Theoretical perspectives on the impact of globalization, study abroad, and the global partnerships which sustain these initiatives

Moving beyond the significant challenges to successful Global South-North partnerships, and the study abroad programs they sustain, it is advantageous to deconstruct these experiences to better understand the forces and factors which impact these programs. In past decades, researchers (Alfaro and Quezada, 2010; Merryfield, 1995; Phillion et al., 2009) have stated the need for educators to prepare their students to productively engage in a globalized world, using study abroad experiences as one tool to achieve this goal. More recent research (Engel, 2017) reiterates the positive impact of study abroad programs on post-secondary academic achievement and retention of US students, based on studies conducted by state university systems in Georgia, Florida and Texas. Farrugia and Sanger's (2017) research also found that university study abroad programs enhanced the career prospects of their participants, regardless of major, when compared to colleagues who did not engage in an international experience. These scholars also posited that study abroad programs might also effectively aid teachers in addressing stereotypical prejudices within their classrooms, while simultaneously enhancing their students' knowledge regarding a more globalized world. To this end, we will now briefly discuss select theoretical perspectives which are often at the core of international study abroad programs and partnerships. These select theoretical perspectives include (a) internationalization and global competence, (b) critical cosmopolitan theory, (c) colonialism and decoloniality, (d) social justice, and (e) global citizenship.

Internationalization and global competence: impact on culturally responsive pedagogy

During the past two decades we have seen a continued emphasis on the need for culturally responsive pedagogy as initially defined by Gay (2010) as "using cultural knowledge, prior experiences, frames of reference, and performance styles of ethnically diverse students to make learning encounters more relevant to and effective for them" (p. 31). This conceptual framework has been expanded over the years, suggesting that culturally responsive pedagogy is an important outcome of study abroad engagement. For example, Gay (2018) points to an increasing cultural divide between teachers and students, suggesting that pre-existing cultural biases and stereotypes continue to erode the quality of education provided to marginalized populations.

For Lopez and Morales (2021) globally competent teachers must not only acquire academic knowledge with an international dimension, but must also apply numerous pedagogical skills for their students to value multiple points of view and to work toward becoming more globally responsible citizens. More specifically, these scholars believe that global competency may include both dispositions and skills that underscore and sustain cultural interconnectedness, while affirming diversity and contributing to a fuller understanding of their students, families and communities. This infusion of a more critical analysis of inequities within the world requires that globally competent teachers help their students "identify local issues and connect them to global trends and employ a range of critical pedagogies to guide students in examining root causes of issues and facilitate opportunities for students to take action" (Kopish, 2017, p. 21).

Internationalization and global competency as perceived by the Global North and South

Over a decade ago, Zhao (2010) perceived global competence as being the skills and knowledge necessary "to be aware of the global nature of societal issues, to care about people in distant places ... to appreciate the interconnectedness and interdependence of peoples, to respect and protect cultural diversity, [and] to fight for social justice for all" (p. 426). Relating this theme more directly to teacher education, Byker and Putman (2019) suggested that internationalization with a Global North perspective means "integrating international experiences and intercultural dimensions into teaching and learning" (p. 90). To achieve these results, they suggested the need to foster greater global awareness among preservice teachers through enriched coursework and study abroad programs focusing on multiple geographic regions, global issues, greater cultural understanding, and intensified language learning.

However, Byker and Putman (2019) also warned that perceptions of global competency might be seen quite differently from different locals, suggesting that for the Global South, the goals of internationalization and global competency may seem well intentioned, but could also reflect tacit hegemony and the further reproductions of inequities. For example, participation in study abroad programs from the Global North are often skewed toward the urbanized wealthy, leaving behind students marginalized by poverty, location and race or ethnicity. This perspective is expanded by Lopez and Morales (2021) with research that suggests that contrary to previous results for participants from the Global North, students from the Global South do not necessarily reflect their experiences as being personally transformational. Explaining that the Global South often relies on the Global North for significant levels of development and is aware of the power dynamics between students and academics in the Global North and South, participants in study abroad programs initiated from the Global South may define global competency quite differently. For example, globally competent teachers from the Global South, unlike their Global North counterparts, may not see themselves as the center of the universe, as they are quite aware that they must also view global issues from a position of less power when compared to the US. According to Lopez and Morales (2021) their Chilean teacher candidates come from a previously colonized country with a unique history and indigenous cultures which, for the most part, were replaced by external hegemonic cultures. Thus, the way in which they perceive their role as teachers may be quite different from their counterparts in the Global North as they do not necessarily see themselves as superior to others, including ideologies that reflect the need to help other countries.

As we have seen, the conceptualization of globalization and global competencies remain fluid, and at times contested, depending on perspectives regarding the role and value of study abroad programs, and the partnerships which often struggle to bring them to fruition. To extend this conversation, we now consider the issues of how colonization and decoloniality may impact not only our current understanding of globalization, but how it might be improved.

Colonialism, decoloniality and their impact on globalization

Present day participants in study abroad programs, particularly initiated by the Global North, may be quite unaware of the continuing impact of colonization, naively assuming it simply disappeared with the independence of former colonies during the mid-twentieth century. However, its negative impact did not simply disappear with the decline of official colonial systems, as nations and individuals continue to exert power over one another, complete with systems of rules and laws that determine the rights of peoples and provide the most reward for those who dominate. The reality of the 21st century is that the inequality among countries, as well as the impact of powerful countries from the Global North, continues to influence the lives and livelihoods of citizens in the Global South.

Today, there continues to be a need to confront the impact of colonization through both education and action (which can include more equitable study abroad programs and partnerships). However, there is also a need to understand perspectives from the Global South which underscore their omission from essential and equitable roles within globalization. Clearly, global perspectives, attitudes and programs are often skewed toward the Global North, which reflect Global North positions of privilege through the continuing lens of colonialism. From the South African perspective [Christie \(2020\)](#) advances the notion that the concept of decoloniality, which moves us beyond continuing cycles of colonialism and unequal power among nations and people, must first separate knowledge production from its former Eurocentric focus, while leaving behind the obsolete assumptions regarding the superiority and universality of Western knowledge and culture, particularly regarding education and curricula decisions. For [Lopez and Morales \(2021\)](#), there is a need to recognize the historical (and often continuing) relationship between the colonizer and the colonized that can shape individual identities while impacting and influencing knowledge production with the inclusion of colonial ideologies and mythical assumption, often without the input or consent of those colonized. To this end, both educational institutions, and the study abroad programs they support, must resist practices of domination and oppression while working toward goals of equity and inclusion, particularly for historically subjugated peoples.

When applying theories of colonialism and decoloniality to current study abroad programs, participants and partnerships alike must challenge continuing reflections of privilege and positioning supported by the lens of colonialism. This might begin with an analysis of Global North study abroad participants' motives, which often include personal growth, travel and adventure. While not necessarily invalid, they do reflect "one-way" benefits that favor the Global North while failing to challenge colonialism. For example, often absent from the motivations for participating in programs destined for host nations in the Global South were the goals of social justice and solidarity. Program directors and partnerships need to examine the true ethics reflected within their study abroad programs, while ensuring the inclusion of a more critical reflection that will help move us all beyond the legacy of colonialism in global education. For example, participants might question the role of a colonizer's language within a Global South nation, as well as if dichotomies such as native/non-native speaker remain valid.

Unfortunately, colonialism is not limited to the political and economic realms, but also pervades the perspectives and mindsets of those involved in international study abroad and partnership activities. For example, [Sharpe \(2015\)](#) mentioned a study abroad program in Cuba where participants from the Global North reflected colonialist tendencies by maintaining comfortable American style pedagogy and daily routines, as if they were tourists, while further reinforcing "othering" identifications between themselves and their host community. Throughout their study abroad program, the American students continued to view their Cuban counterparts as "others," while maintaining colonialist perspectives such as consumerism, personal growth, and economic advancement.

Clearly, study abroad programs and the partnerships which support them must not only decolonize their goals and curriculum, but must also move beyond simplistic and colonialist views of "othering" those who do not share their home cultures. Building upon these theoretical notions is a clear need to move toward a more critical cosmopolitan citizen theory that underscores themes of social justice. The following discussion will briefly describe the development and infusion of critical cosmopolitan theory, as well as the need to work toward greater social justice globally while strengthening a more comprehensive and impactful vision of global citizenship.

Critical cosmopolitan and social justice theories

For [Byker and Xu \(2019\)](#), the goal of Critical Cosmopolitan Theory (CCT) is to develop global citizens that continue to develop an increased awareness of a wider, and more interrelated world. Moreover, CCT works toward creating more compassionate and inclusive global citizens who continue to work for social justice. It should be noted that this theory moves us beyond neoliberal perceptions of global citizenship that were focused on issues related to the global economy. With a call for global citizens to engage in both local and global dialogs aimed at challenging oppression while working collaboratively to overcome social injustices, CCT calls upon individuals, including participants in international study abroad programs and partnerships, to denounce social injustices while promoting an individual's human dignity and future contributions. Moving beyond rhetoric, however, Byker and Xu also view CCT as another means for teachers and students to positively impact future history through social action, instead of simply reading history as written by others.

Building upon these concepts, [Newton et al. \(2020\)](#) recommend that study abroad directors design programs that not only focus on intercultural sensitivity and perspectives such as global competence and culturally responsive pedagogy, but must also provide preservice participants with opportunities to critically reflect upon issues of social justice. These critical reflections should include all aspects of study abroad programs, including context, curriculum, and co-curricular activities, as well as essential pedagogical strategies. To this end, effective study abroad programs will need to transform preservice teachers' understanding of how their social class, gender, and race/ethnicity impact their perceptions of the world.

Within teacher education program study abroad field experiences should also provide a way for preservice teachers to begin linking theory to practice through programs that offer service-learning opportunities. For [Newton et al. \(2020\)](#), short-term study abroad programs offer potential opportunities to prepare teachers for their future classroom which will undoubtedly include increasingly diverse cultures. Programs that incorporate critical approaches to issues such as global social justice enhance preservice teachers' experiences, abilities, and propensity to promote concepts of globalization, while enhancing mutually beneficial partnerships with their host communities.

This goal, of course, requires significant transformations that move study abroad participants beyond tourist experiences and limited preconceived notions and expectations regarding their host communities. To this end, program leaders need to encourage their participants to move beyond viewing their host communities as "others," to perspectives that promote empathy and partnership. To achieve this goal, study abroad programs must include critical multicultural coursework that helps to develop participants' intercultural competencies. These assignments and activities should also be designed to counter a "white visitor" complex. This hypervisibility within their host communities, where they are viewed as outsiders due to their lack of knowledge regarding the local language and cultures, will hopefully diminish through extensive orientation and in-country programs and activities designed and managed through an equitable and sustainable Global South-North partnership.

Unfortunately, these goals will remain challenging until we move beyond the current population of preservice teacher candidates and study abroad participants that continue to represent a majority of white, middle-class, female participants. This ongoing underrepresentation of marginalized and economically poorer student populations remains a challenge that must also be addressed to promote programs that promote cultural diversity and social justice.

Global citizenship

For [Byker and Putman \(2019\)](#), the concept of global citizenship itself remains contested and continues in many instances to reflect Western ideals, often driven by neoliberal perspectives which are more focused on the development of economies than on social justice. Their research also suggests that our Global South colleagues may have a different approach to global citizenship which focuses more on the issues of power and inequality. Thus, when viewing global citizenship from a Global South perspective, there needs to be more emphasis on participatory citizenship which addresses systematic and historical issues of powerlessness; views which are challenged by Global North domination and colonization.

While infusing issues of global citizenship within international study abroad programs and partnerships remains a priority, this process needs to be initiated within teacher education curricula. For example, professional organizations such as the [National Council for the Social Studies \(2016\)](#) continue to highlight the importance of including global perspectives within K-12 classrooms to facilitate a better understanding of the benefits from an increasingly interdependent global network of cultures, economies, and political systems. During the past few decades researchers ([Gaudelli, 2016](#); [Hadis, 2005](#); [Stachowski and Sparks, 2007](#)) have continued to suggest that one of the most effective means for preparing citizens for a more globalized world is by immersing themselves in another's culture. However, to make the most of these opportunities, the concept of global citizenship, along with other perspectives such as social justice, are best introduced early in a student's schooling.

Moving from theory to practice

While the previous theoretical perspectives provide us with a better understanding of the forces and factors that impact international study abroad programs and partnerships, it is critical to apply these perspectives to actual study abroad programs. Moreover, to better understand the dynamics such as global competency and decoloniality that influence Global South-North partnerships, and the study abroad programs they support, we will now examine the outcomes of a study abroad program which took place in Botswana during the summers of 2015, 2017, 2018, and 2019. This program, and its' participant outcomes, are part of a decade-long partnership between UIB and UCF bringing students to Botswana for educational lectures, cultural engagement, and a two-week stay in Remote Area Dwelling (RAD) schools where they immersed in authentic classroom experiences at primary and junior secondary schools. Our research focused on whether study abroad programs can effectively prepare educators to engage in inclusive pedagogy, while building upon both cultural and contextual knowledge to help meet the needs and challenges of a culturally and linguistically diverse K-12 student community.

In this article we analyze the outcomes of one university-based study abroad program, whose overarching focus was on understanding that culture is a major variable in society and schooling, and as such must be valued by educators to promote educational equity and equality among all students. Moreover, international programs need to be designed such that both home and host institutions and nations benefit from the exchange and interaction of participants, and that participants are enabled to interact with their

host country and local schools to enhance meaningful communication. To this end, the authors suggest that study abroad programs, when carefully designed and supported by an equitable global partnership, will enhance an understanding of the roles and impact of diversity and plurality within and beyond educational institutions.

Research methodology

In order to ascertain participants' knowledge growth as well as their qualitative perspectives on the program itself, a mixed-methods methodology was employed using a multi-level model of triangulation design in accordance with Tashakkori and Teddlie (1998). Through this method, program participants engaged in two pre-departure knowledge assessments as well as a pre-departure reflection, which were completed five months prior to departure and before any synchronous orientation meetings. The pre-departure knowledge assessments fell into the categories of "Cultural Perspectives and Curriculum Resources" and "Background and Setswana." With the "Cultural Perspectives and Curriculum Resources" survey, students were asked to gauge their own feelings and experiences on interactions with different cultures, comfort in the classroom, and related teaching-specific qualities. For the "Background and Setswana" assessment, a forty-question quiz on background knowledge of the country, culture, and Setswana language was administered. On the qualitative side, the pre-departure reflection served as an open-ended survey asking participants the following questions:

1. Has your impression of this project changed from the time of your initial interview? First describe what this impression was, and then explain why it has, or has not, changed.
2. Have your expectations regarding Botswana changed from the time of your initial interview? First describe what these expectations were, and then explain why they have, or why they have not changed.
3. What do you expect will be the most rewarding experience(s) during your stay in Botswana?
4. What do you expect to be the most difficult experience(s) during your stay in Botswana?
5. At this point, how do you think this experience will affect you as a future educator and/or professional?
6. What is the one, most important thing you expect to bring back from Botswana?
7. Please share with us any other thoughts or impressions that you may have with us at this time.

As a follow-up, one month after returning from the study abroad experience, participants were asked to complete two post-program knowledge assessments in the same areas mentioned before for quantitative analysis. A post-program reflection was also submitted by students with the past tense versions of the seven-question survey of the pre-program reflection. This pattern was maintained for each of the four study abroad year cohorts analyzed: 2015, 2017, 2018, and 2019.

Participant demographics

Over the course of the four program cycles, there were a total of forty-two participants in the study abroad program to Botswana, with ten participants in 2015, eleven in 2017, eight in 2018, and thirteen in 2019. In total, four participants (9.5%) identified as male, while thirty-eight (90.5%) identified as female. Participants came from a wide variety of racial backgrounds, as seen in Fig. 1 below:

Regarding academics, the vast majority of program participants were in the last two years of their undergraduate careers, with twelve (28.6%) being rising juniors and twenty-two (52.4%) being rising seniors at the time of program participation. Graduate students only comprised seven (16.7%) of program participants, with the final participant having been a rising sophomore. As

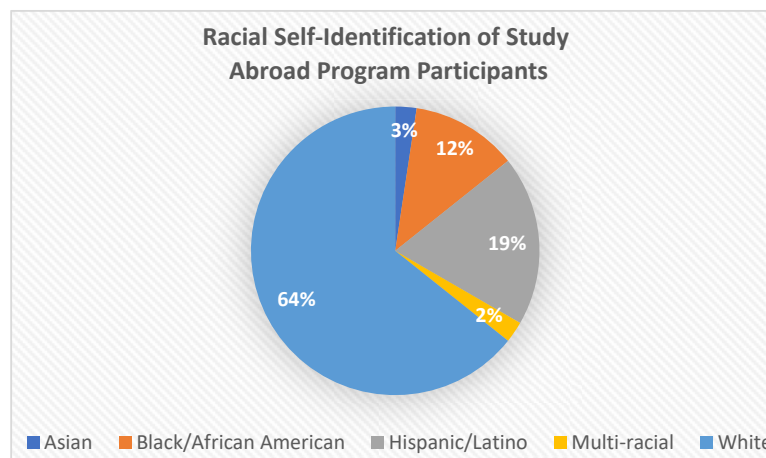


Fig. 1 Racial self-identification of program participants.

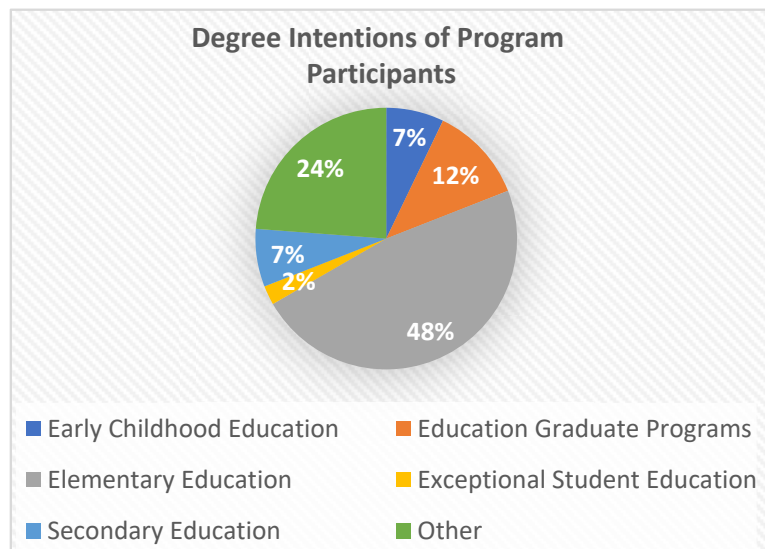


Fig. 2 Degree intentions of program participants.

the program was managed through the education colleges at the respective universities, most of the student participants' intended degree ran in accordance with this, with a breakdown of majors seen in Fig. 2 above. Majors in the "Other" category included such options as English, French, Health and Exercise Sciences, and Hospitality and Tourism Management.

Finally, one interesting point regarding the participant demographics was that 15 of the 42 (35.7%) self-identified as the first members of their family to attend college.

Results

In this section, we discuss the quantitative and qualitative results of data collected from the four cohort years. All data collection and informational storage procedures were conducted with the express permission of UCF's and UB's Institutional Review Boards, and student consent was also obtained.

Basic knowledge acquisition

In order to analyze participant's acquisition of knowledge regarding Botswana, a dependent samples t-test was performed on the pre-departure and post-program assessments for "Background and Setswana" on both the aggregated and individual data from the four cohorts. All recorded positive gains. Using all 42 participants' results, it was found that there was a statistically significant increase in basic knowledge acquisition in terms of country, culture, and language knowledge from pre-departure to post-program ($t = 11.291, p < 0.001, df = 41$). These statistically significant increases were identified regardless of a participant's academic year at the time of the program, racial identification, or first-generation college student status. As the data set only included four male participants, it was determined that there was not enough information to make a sound conclusion about gender results and comparison in basic knowledge acquisition.

Participants' most rewarding experiences

Participants repeatedly and overwhelmingly cited their in-school encounters as the most rewarding experiences they had in the study abroad program. One White Female participant from 2019 noted, "I loved teaching the students, but also learning from them (they taught me about their country, culture, and language)." Even as some of the participants faced teacher absenteeism in some of the schools, this in itself was also cited at times as a rewarding experience. One participant from 2018 noted:

My most rewarding experience was definitely being in the classroom with the kids. For the first week, they had no teacher, so I essentially had to be the teacher. This was so challenging, but beyond rewarding because I was able to see the issues that each child faced first-hand.

White Female, 2018

Another participant from 2017, who went on post-program to join Teach for America, agreed with the schooling experience being the most rewarding but specifically discussed the rural school environment and tied it into her own experience as a first-generation college student:

What drew me most to this project was the fact that we would be teaching in rural schools. Being part of She's the First, education is very important to me and I'm passionate about working to provide education to students who normally wouldn't attend school, or would attend a poor-performing school.
White Female, 2017

This participant then continued to cite the relationship she developed with one student and the positive impact it had upon her include overall experience in the program. Although many study abroad programs cultural experiences, visits to historical sites, and lectures by host country faculty, it appears that the in-school experiences made the UB/UCF program compare quite favorably to those from other institutions.

Participants' most challenging experiences

In terms of challenging experiences, a common theme among participants across the years was the cultural differences in the classroom environment, including the witnessing of corporal punishment. As one participant stated:

The most difficult part of this trip was how the students were treated in the rural schools. Most teachers I spoke to seemed indifferent to the students. I witnessed a teacher whip his students because they didn't bring in their assignments.
Hispanic/Latina Female, 2019

Although corporal punishment and its prevalence in a variety of nations in the Global South was discussed during the pre-departure orientation sessions, in readings and during in-country lectures prior to the school experience, there was a disconnect for multiple students between what they read, and what they observed within the classroom. This theme emerged within participant reflections.

Yet another challenge cited by multiple participants was a feeling of homesickness that emerged from being in a new environment and away from one's family and friends for an extended period of time, as was the case with these four years of programs. One student, a White female from the 2019 cohort, stated she did not anticipate "how hard it would be to detach from what was going on at home." Another White female from the same cohort stated that she had times when she felt "overwhelmed" but also grateful that she "was willing to go outside my comfort zone and sign up for this trip to travel 7000 miles away with a group of strangers." Creature comfort adjustments were also mentioned as challenging both in their classrooms, and in their homestays, and included varying amounts of hot water and internet access. Nonetheless, multiple participants did state that their previous notions of what a Global South nation looked like were shattered due to unexpected levels of urbanization and development within Botswana's capitol city, Gaborone, and its second-largest city, Francistown. In addition to living in rural communities during their school experiences, participants also visited and lived in these two urban areas during their program.

Professional growth and development of educators through study abroad programs in emerging nations

Many participants reported an increased sense of classroom confidence as well as preparation for teaching in challenging environments upon return to the United States. One student, a Hispanic/Latina female from the 2019 cohort, reported that after her experience, she feels that she "can teach in a Title 1 school" a sentiment echoed by many. Another 2019 participant, a White female, stated that she "look(s) forward to using the skills [she] learned in Botswana to become a better educator in the United States." This sentiment was echoed on the behavioral side by an African-American/Black female participant from 2018, who said she has come to the realization that, "everything doesn't need a reaction – be patient, be kind."

Other participants cited this experience as one that enhanced their desire for more international classroom experiences. Another participant from 2018, a White Female, stated that the experience, "gave me a little glimpse into what it might be like to work for the Peace Corps and made me realize that this is one hundred percent something I want to pursue." Similarly, a 2017 White female participant stated the "huge impact" of the program on her has made her consider entering Teachers Without Borders upon graduation.

Yet another aspect of professional growth reported was participants' newfound respect for language diversity in a classroom setting, and how this impacts equal access to quality education. Many students shared feelings that this experience paired nicely with what they had learned in previous courses, including the need to support second language learners and linguistic diversity.

One particular participant also spoke about realizations regarding the needs of students with disabilities and wanting to make sure appropriate accommodations were taken in their own classrooms once they returned. She stated:

I met two brothers, both suffering from speech disabilities in a school with no accommodations. I can't imagine how helpless and alone they must feel. I want to work to avoid any child feeling that way and do what I can to always provide proper accommodations for students.

White Female, 2019

The recognition of how best to support learners of varying needs was imbedded throughout all programs, regardless of participant gender, race, and academic standing. With this systematic theme carrying through, participants came to recognize the professional development opportunities within the study abroad program they had just experiences.

Impact on participants' professional development

Participants reported anticipating using the skills they learned in Botswana, in multiple ways. Referring back to the corporal punishment theme discussed in a previous section, a White female from 2019 program stated that she now speaks about the experience in her current education class and, "how to make [students] listen without hurting them."

Moreover, the theme of increased appreciation and understanding for the need to have culturally responsive and sensitive curriculum within the classroom also emerged from participants from all four cohorts. As one participant mentioned:

This experience will most definitely have a huge effect on my future professional career. This experience has made me so much more open-minded, aware of prejudices everywhere, conscious and respectful towards other cultures and beliefs, aware of racial differences and segregation in life and schools, and so much more ... I will also be able to add a lot more culture and social issues into my classroom, which in turn will make my students more open minded and accepting of others.

White Female, 2017

Concluding remarks

Educational programs should promote social justice, nurture equitable relationships, and unite people regardless of their global location. Quite clearly, the positive outcomes of this UB/UCF Global South-North collaborative partnership outweighed the negatives, and other study abroad programs may wish to include these outcomes in their partnership design. The multicultural exposure that both Botswana and American academics and students experienced prepared them to be more confident when functioning in an increasingly culturally diverse environment. The intent of these programs was not to assimilate, nor to oppress "the other," but rather to learn how to coexist amid cultural differences. As Mapp (2012) noted, the overall impact of study abroad programs is positive as they help participants to become more open in their ways of thinking, to develop a greater awareness of the culture and values of others, and to enhance their sense of appreciation for differences, especially regarding social justice matters.

That said, the results of this study suggest that while participants' knowledge base regarding their host country increased, though much of this could have been obtained without leaving home. However, as it is often said that "experience is the best teacher," it was through the qualitative and affective domains that the most intriguing results occurred. Participants' perceptions of their own professional development, cultural awareness, pedagogy and curricular material selection generally indicated a sustained positive growth over time. Their exposure to complex and diverse communities has enriched their cross-cultural conceptualizations and we anticipate that their ability to reach and teach 21st Century students will be enhanced. Recent research (Baecher, 2021; Pfingsthorn and Czura, 2017; Tripp et al., 2021) suggests similar outcomes that participation in study abroad programs enhance professional development opportunities for globalizing teacher education. It also suggested that ongoing globalization now demands that educators engage in programs that promote understanding of global issues and interventions necessary to address transnational problems. Both the professional knowledge and social/cultural perspectives of study abroad participants were positively influenced through their in-country experiences in Botswana.

Our discussion also supports the conclusion that equitable and durable institutional partnerships are crucial to the development of transformational study abroad programs. The findings of our study also provide recommendations for developing meaningful programs in emerging nations by establishing a more equitable partnership between universities in the Global South and North. However, to ensure sustainable and collaborative partnerships, there must first be transparent and mutually agreed upon "terms of engagement" regarding the development, management, and evaluation of projects. Above all, partners must exercise collegiality and respect for cultural diversity. These approaches will hopefully curtail situations in which institutions from the Global South are cast in dependent roles, while their Global North partners maintain control of "joint" projects. If ignored, inequities such as these can compromise the goal of mutual capacity building, the key to successful Global South-North partnerships.

Moving beyond the necessity of equitable Global South-North partnerships, we have also underscored the potential positive impact of study abroad programs in emerging nations on professional educators. Reflecting upon our partnership and study abroad programs, cultural immersion appears to play a significant role in developing transformational learning environments.

Undoubtedly, well-structured global experiences are crucial for teachers who must continue to address multiple issues linked to cultural diversity and global justice within their classrooms, while educating the next generation of globally competent K-12 students. We have hopefully provided a “wake up call” for our next generation of educators who will experience increasingly diverse K-12 classrooms.

Finally, as education has clearly become a global enterprise, we cannot afford to prepare teachers who do not have adequate exposure to an increasingly global education system and understanding of the role of global justice, both within their classrooms and worldwide. Study abroad initiatives, such as our programs in Botswana, provide students with global experiential learning opportunities. Study abroad participants supported by the UB/UCF Global South-North partnership have demonstrated how these international experiences have contributed to their professional growth, with some moving on to positions with national programs such as the Peace Corps, NGOs such as FHI360, or numerous multilateral agencies such as UNESCO. Although scholars will continue to debate the benefits of study abroad programs, the UCF/UB experience remains both distinctive and positive.

References

- Alfaro, C., Quezada, R.L., 2010. International teacher professional development: teacher reflections of authentic teaching and learning experiences. *Teach. Educat.* 21 (1), 47–59.
- Biraimah, K., Jotia, A., 2012. The longitudinal effects of study abroad programs on teachers' content knowledge and perspectives: Fulbright-Hays Group Projects Abroad in Botswana and Southeast Asia. *J. Stud. Int. Educ.* 17 (4), 433–454.
- Biraimah, K., Jotia, A., 2014. Developing Transformative Study Abroad Programs in Botswana: A Precarious Partnership Fraught With Angst and Hope Between North and South [Paper Presentation]. Comparative and International Education Society Annual Conference, Toronto, Canada.
- Biraimah, K., Jotia, A., 2016. Developing Equitable Capacity-Building Partnerships Between the Global South and North: Whose Agenda Dominates and Whose Capacity is Truly Enhanced? [Paper Presentation]. South Africa Educational Research Association 2016 Conference, Cape Town, South Africa.
- Biraimah, K., Jotia, A., 2021. Developing an Equitable Global South-North Partnership: A Study Abroad Program in Botswana [Paper Presentation]. Comparative and International Education Society Annual Conference, Seattle, Washington, U.S. <https://cies2021.org>.
- Baecher, L. (Ed.), 2021. *Study Abroad for Pre- and In-Service Teachers: Transformative Learning on a Global Scale*. Routledge. <https://doi.org/10.4324/9781003004387>.
- Bishop, S.C., 2013. The rhetoric of study abroad: perpetuating expectations and results through technological enframings. *J. Stud. Int. Educ.* 17 (4), 398–413.
- Bradley, M., 2008. North-south research partnerships: lessons from the literature. *NORRAG News*. <http://www.norrag.org>.
- Brown, A., 2013. Waiting for superwoman: white female teachers and the construction of the neoliberal savior in a New York City Public School. *J. Crit. Educ. Pol. Stud.* 11 (2), 123–164.
- Byker, E.J., Putman, S.M., 2019. Catalyzing cultural and global competencies: engaging preservice teachers in study abroad to expand the agency of citizenship. *J. Stud. Int. Educ.* 23 (1), 84–105.
- Byker, E., Xu, T., 2019. Developing global competencies through international teaching: using critical cosmopolitan theory to compare case studies of two study abroad programs. *Frontiers* 31 (2), 105–120.
- Caton, K., Santos, C.A., 2009. Images of the other: selling study abroad in a postcolonial world. *J. Trav. Res.* 48 (2), 191–204.
- Christie, P., 2020. *Decolonising Schools in South Africa: The Impossible Dream*. Routledge.
- Cole, T., 2012. The white-savior industrial complex. *Atlantic*. <http://www.theatlantic.com/international/archive/2012/03/the-white-savior-industrialcomplex/254843>.
- Engel, L.C., 2017. Underrepresented Students in US Study Abroad: Investigating Impacts. Institute of International Education (IIE). Research and Policy Brief Series, Issue Number 1.
- Farrugia, C., Sanger, J., 2021. Gaining an Employment Edge: The Impact of Study Abroad on 21st Century Skills and Career Prospects in the United States [Center for Academic Mobility and Impact]. International Institute of Education, Washington, DC.
- Foronda, C., Belknap, R.A., 2012. Transformative learning through study abroad in low income countries. *Nurse Educat.* 37 (4), 157–161.
- Gaudelli, W., 2016. *Global Citizenship Education: Everyday Transcendence*. Routledge.
- Gay, G., 2010. *Culturally Responsive Teaching*, second ed. Teachers College Press.
- Gay, G., 2018. *Culturally Responsive Teaching: Theory, Research, and Practice*, third ed. Teachers College Press.
- Glick, H., 2015. Bold New Photo Project Tackles Appalachian Poverty Porn. *Global Citizen*. Retrieved from. <https://www.globalcitizen.org/en/content/bold-new-photo-project-tacklesappalachian-poverty/>.
- Hadis, B.F., 2005. Why are they better students when they come back? Determinants of academic focusing gains in the study abroad experience. *Frontiers* 11, 57–70.
- Joseph, D., Johnson, R., 2021. Studying abroad while staying at home in Australia: diverse intercultural learning and teaching through partnerships. In: Kommers, S., Bista, K. (Eds.), *Inequalities in Study Abroad and Student Mobility: Navigating Challenges and Future Directions*. Routledge, pp. 140–155.
- Jotia, A., Biraimah, K., 2016. Developing research partnerships in emerging nations: bridging the north-south divide. *Msingi J.* 1 (2), 87–115.
- Jotia, A., Biraimah, K., 2017. Decolonizing a Global South-North Partnership: is an Equitable and Transformative Research Agenda Possible Between the University of Botswana and the University of Central Florida, USA? [Paper Presentation]. South African Educational Research Association Conference, Port Elizabeth, South Africa.
- Jotia, A., Biraimah, K., 2018. Is Decolonization and Equitable Transformative Research Possible Within a Global South-North Partnership? A Case Study of Botswana and the US [Paper Presentation]. Comparative and International Education Annual Conference, Mexico City, Mexico. <https://cies2018.org>.
- Jotia, A., Biraimah, K., Kurtz, B., 2020. Transformative study abroad programs in emerging nations: moving toward equitable global north-south partnerships. *Theor. Pract.* 59 (3), 279–288. <https://doi.org/10.1080/00405841.2020.1751527>.
- Jotia, A., Biraimah, K., Kurtz, B., 2021. Transformative Study Abroad Programs in Emerging Nations: Moving Toward Equitable Global North-South Partnerships [Paper Presentation; Invited]. American Association of Colleges of Teacher Education 73rd Annual Meeting, Seattle, Washington, U.S.
- Kommers, S., Bista, K., 2021. *Inequalities in Study Abroad and Student Mobility: Navigating Challenges and Future Directions*. Routledge.
- Kopish, M., 2017. Global citizenship education and the development of globally competent teacher candidates. *J. Int. Soc. Stud.* 7 (2), 20–59.
- Lopez, M.M., Morales, P.R.L., 2021. From global south to global north: lessons from a short-term study abroad program for Chilean teacher candidates in English pedagogy. *J. Res. Child. Educ.* 35 (2), 248–267.
- Lorz, M., Netz, N., Quast, H., 2016. Why do students from underprivileged families less often intend to study abroad? *High Educ.* 2, 153–174. <https://doi.org/10.1007/s10734-015-9943-1>.
- Mapp, S.C., 2012. Effect of short term study abroad programs on students' cultural adaptability. *J. Soc. Work. Educ.* 48 (4), 727–737.
- Merryfield, M.M., 1995. Institutionalizing cross-cultural experiences and international expertise in teacher education: the development and potential of a global education PDS network. *J. Teach. Educ.* 46 (1), 19–27.
- National Council for the Social Studies, 2016. Global and International Education in Social Studies. [NCSS Position Statements]. <https://www.socialstudies.org/position-statements/global-and-international-education-social-studies>.

- Newton, J., Oudghiri, S., Obenchain, K., Phillion, J., 2020. Preservice teachers' understandings of social justice within the context of study abroad programs. *Theor. Pract.* 59 (3), 259–268.
- Onyenekwu, I., Angeli, J., Pinto, R., Douglas, T., 2017. Misrepresentation among U.S. study abroad programs traveling to the African continent. A critical content analysis of a Teach Abroad program. *Frontiers* 29 (1), 68–84.
- Pfingsthorn, J., Czura, A., 2017. Student teachers' intrinsic motivation during a short-term teacher training course abroad. *Lang. Cult. Curric.* 30 (2), 107–128.
- Phillion, J., Malewski, E.L., Sharma, S., Wang, Y., 2009. Reimagining the curriculum: future teachers and study abroad. *Frontiers* 18, 323–339.
- Powell, J., Menden, S., 2018. The problem of othering: towards inclusiveness and belonging. In: Grant-Thomas, A., Galloway-Popotas, R., Menden, S., Omi, M. (Eds.), *Othering and Belonging Expanding the Circle of Human Concern*. Haas Institute for a Fair and Incisive Society, pp. 14–39.
- Rose, A.J., Cooper-Duffy, K., 2021. Implementation of an interprofessional study abroad experience in Botswana for special educators and speech-language pathologists. In: Baecher, L. (Ed.), *Study Abroad for Pre- and in-Service Teachers: Transformative Learning on a Global Scale*. Routledge, pp. 146–165. <https://doi.org/10.4324/9781003004387>.
- Schroeder, K., Wood, C., Galiardi, S., Koehn, J., 2009. First do no harm. Ideas for mitigating negative community impacts of short-term study abroad. *J. Geogr.* 108 (3), 141–147.
- Sharpe, E.K., 2015. Colonialist tendencies in education abroad. *Int. J. Teach. Learn. Higher Educ.* 27 (2), 227–234. <https://files.eric.ed.gov/fulltext/EJ1082881.pdf>.
- Stachowski, L.L., Sparks, T., 2007. Thirty years and 2000 student teachers later: an overseas student teaching project that is popular, successful, and replicable. *Teach. Educ. Q.* 34 (1), 115–132.
- Tashakkori, A., Teddlie, C., 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Sage, Thousand Oaks, CA.
- Tripp, L.O., Love, A., Barry, N., Thomas, C.M., Russell, J., 2021. A qualitative analysis of teacher candidates' study abroad experience in Malawi. In: Baecher, L. (Ed.), *Study Abroad for Pre- and in-Service Teachers: Transformative Learning on a Global Scale*. Routledge, pp. 129–145. <https://doi.org/10.4324/9781003004387>.
- Von Wendorff, F., 2013. Top 5 Challenges of Studying Abroad. Retrieved from: www.topuniversities.com/blog/top-5-challenges-studying-abroad.
- Wohlgemuth, L., Olsson, J., 2003. Dialogue in pursuit of development. In: Expert Group on Development Issues, vol. 2, pp. 58–71.
- Woolf, M., 2006. Come and see the poor people: the pursuit of exotica. *Frontiers* 13, 135–146.
- Zhao, Y., 2010. Preparing globally competent teachers: a new imperative for teacher education. *J. Teach. Educ.* 61, 422–431. <https://doi.org/10.1177/0022487110375802>.

Parental unemployment and children's educational outcomes – a Literature review

Hannu Lehti, University of Turku, INVEST Flagship Center, Faculty of Social Sciences, Turku, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	118
Results of the reviewed studies	119
Educational performance	119
Grade repetition	121
School dropout at secondary level	121
Upper secondary enrollment and completion	122
Tertiary education enrollment and completion	123
Mechanisms and institutional context	124
The reduction of family income	124
Noneconomic mechanisms	125
Conclusions	126
References	126

Introduction

Unemployment is a negative experience that produces economic deprivation, stress, and other wellbeing problems for unemployed individuals. The negative consequences of unemployment may not only be restricted to an unemployed individual but also to other people who are dependent on the unemployed individuals' resources, namely their children. Parental unemployment has been shown to be negatively associated with children's self-esteem, psychological wellbeing (Christoffersen, 1994), attitudes toward work (Payne, 1987), educational ambitions (Andersen, 2013), occupational attainment (Karhula et al., 2017) own unemployment in adulthood (Ekhaugen, 2009), and educational attainment (Ström, 2003).

In this article, I review studies that measure the effect of parental unemployment on children's educational outcomes, such as grade point averages (GPAs), grade repetition, educational enrollment and completion, and educational attainment. Because no previous articles have reviewed the effect of parental unemployment on children's education globally, all the studies that have been conducted in developed and developing economies are included. However, only one of the reviewed studies was conducted in a developing economy; thus, the results consider mainly developed countries.

The aim of this article is to observe how parental unemployment influences children's education in different countries and institutions globally. Further, the main mechanism between parental unemployment and children's educational outcomes, as suggested in the reviewed literature, are examined.

This literature review concentrates only on articles published after 2004. Earlier published studies can be found in a previous review on parental unemployment and different children's outcomes (Ström, 2003). I have only reviewed English articles that are published in peer-reviewed journals; thus, I have excluded working papers in the review. Google Scholar was used to find articles using search words such as "parental job loss/unemployment," "father's/mother's job loss/unemployment," and "children's educational attainment/achievement/outcomes/academic achievement." These restrictions and search words produced 20 peer-reviewed articles that studied parental unemployment and children's educational outcomes in 11 different countries, and one article that analyzed 21 countries. Thus, this review consists of 21 articles that have investigated the effects of parental unemployment and job loss on children's education globally. Eight of the reviewed articles were retrieved from sociological journals, one from statistical, and twelve from economic journals. Because some of the articles studied many outcomes, such as GPA and secondary and tertiary attainment, this review reflects on 28 studies in total.

Parental unemployment can be involuntary or voluntary. I have only focused on studies that have investigated involuntary parental unemployment or job loss. Previous literature has indicated that unemployment and job loss are not precisely the same concepts. Job loss is usually referred to as a discrete event, whereas unemployment refers to a state that includes more heterogeneity according to duration. In these studies, job loss or displacement have been described as an exogenous shock, that has enabled better estimates of the effects of parental unemployment on children's educational outcomes (Brand, 2015). Although job loss and unemployment can be understood as separate concepts, I have considered both in this review. Usually, job loss does not always precede unemployment because an individual can be unemployed for a different reason other than losing a job, for example, because an individual is dismissed. Usually, economists have favored the concept of parental job loss, while sociologists have focused on parental unemployment.

Many of the reviewed articles were published after 2010. This indicates that parental unemployment and children's educational outcomes have recently been a topic of interest in the field of unemployment studies. Fig. 1 illustrates a growing publication trend.

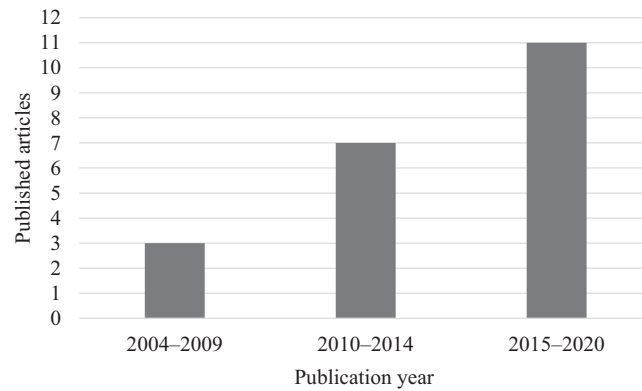


Fig. 1 Number of published articles by 5 years periods.

Only three articles were published between 2004 and 2009, whereas between 2015 and 2019, 11 articles were published on this topic. The Great Recession that started in the US in 2007, and the growing insecurity of the labor markets in many countries thereafter, may have influenced scholars' interest in studying not only how unemployment is associated with individuals, but also how unemployment influences families, particularly the children's prospects regarding education.

Next, I have reviewed the results of the 28 studies in 21 articles, after which I have explored the main mechanisms of the reviewed studies and how the institutional context moderates the effect of unemployment on children's educational outcomes.

Results of the reviewed studies

Table 1 illustrates the basic information regarding all reviewed articles. Five of the 21 articles studied the effect of parental unemployment on children's GPA. Grade repetition as an outcome was studied in two articles, tertiary enrollment or attainment in eleven articles, school dropout at secondary level in four articles, educational attainment, in years, in two articles, and secondary attainment or enrollment in four articles. In total, 21 articles included 28 different studies, because some articles described many outcomes. Overall, almost all studies indicated a negative effect on children's educational outcomes. However, the size of the estimated effect varied, and one study found a positive effect. This indicates that institutional and country contexts influence on the effect of parental unemployment. On average, the effects of parental unemployment are moderate, or marginal for each outcome. The articles incorporate studies conducted in the following countries: USA (6), the Netherlands (1), Germany (2), Finland (2), Hungary (2), Spain (1), Norway (1), Sweden (1), Britain (2), Palestine (1), and Canada (1). Next, I have reviewed the basic results of these studies more precisely, based on the outcome.

Educational performance

Educational performance was studied in five articles. Four articles studied GPA and one examined GCSE point scores (in Britain). Studies were conducted in Sweden, Finland, Britain, Spain, and Norway. All the studies found a negative effect of parental unemployment or job loss on children's GPA, although the effect size varied between countries. The average negative effect of parental unemployment on children's performance measured by GPA or GCSE point scores according to the five studies, is 10.1% of the standard deviation. Thus, the negative effect on children's school performance can be considered to be moderate in effect size.

[Rege et al. \(2011\)](#) studied the effect of parental job loss using Norwegian register data. In their study, they used plant closure for their method, because a plant closure can be assumed to be an exogenous shock, thus removing the problem of selection that is related to unemployment. They found that the father's, but not mother's, job loss had a significant negative effect on their children's school performance, reducing their GPA by 6% of the standard deviation.

In Finland, [Lehti et al. \(2019\)](#) studied the effects of parental unemployment (father or mother) on children's GPA by applying sibling fixed-effect methods. Thus, they could compare siblings within a family, where only some siblings were exposed to parental unemployment before they acquired their final grades from compulsory school. They found that, on average, parental unemployment reduced GPA by 15% of the standard deviation. However, paternal unemployment had a stronger effect on children's GPA (21% on average) than maternal, although the mother's unemployment had a statistically significant negative effect (15% on average).

In addition, [Ruiz-Valenzuela \(2020\)](#) found that parental job loss decreased the GPA by 15% of the standard deviation during the Great Recession in Spain. In the study, she used the individual fixed effects (FEs) model and panel data; thus, he could follow the children and their parents before and after the recession. According to the study, only paternal, not maternal, job loss had a negative influence on school performance.

Table 1 Detailed description of the review articles.

<i>Author(s), year, journal</i>	<i>Country</i>	<i>Data type</i>	<i>Sample size</i>	<i>Outcome variable</i>	<i>Method</i>	<i>Effect size</i>
Mooi-Reci et al. (2019), European Sociological Review	The Netherlands	Longitudinal survey and register data	812	Educational attainment	Weighted least square regression	Edu attainment = −0.51 education years
Lindemann and Gangl (2020), Social Forces	21 countries	Five different longitudinal surveys	13,769	Postsecondary attainment	Multilevel regression (LPM)	The effect depends on the country
Lindemann and Gangl (2019), Research in Social Stratification and Mobility	Germany	Longitudinal survey	1027, (mothers) 1321 (fathers)	Postsecondary attainment	Propensity score matching	Tertiary attainment = −13.7%-points (only father)
Lehti et al. (2019), Research in Social Stratification and Mobility	Finland	Longitudinal register data	113,100, 2508	GPA, upper secondary and tertiary enrollment	Sibling random and fixed effect regression	GPA = −13 or −17% of std. (depends on age), general sec. enrollment = −10%-points, tertiary = −12%-points
Kalil and Wightman (2011), Social Science Quarterly	USA	Longitudinal survey	1374	Postsecondary attainment	Probit regression	College attendance = −10%-points
Kallio et al. (2016), European Sociological Review	Finland	Longitudinal register data	157,135	Secondary educational attainment (school dropout)	Random effects regression	Secondary attainment = −5%-points
Kalil and Ziol-Guest (2008), Social Science Research	USA	Longitudinal survey	4476	grade repetition	Logistic regression	Grade repetition = 2 odds ratio (father only)
Kertesi and Kezdi, (2008), The BE Journal of Economic Analysis & Policy	Hungary	Longitudinal survey	991	School dropout (no secondary education)	Probit regression	School dropout = 5.8%-points per age year
Brand and Thomas (2014), American sociological Review	USA	Longitudinal survey	4,412, 3,993, 2817	High school completion, College attendance, College completion	Propensity score matching	High school = −5.2, College attendance = −9.9, College completion = −4.3%-points (single mothers)
Stevens and Schaller (2011), Economics of Education Review	USA	Longitudinal survey	53,268	Grade repetition	Individual fixed effect regression	15% increase grade repetition
Ruiz-Valenzuela (2020), SERIES	Spain	Longitudinal survey	300	GPA	Individual fixed effect regression	GPA = −15% of std. (only father)
Rege et al. (2011), Review of Economic Studies	Norway	Longitudinal register data	16,164	GPA	Plant closure (LPM)	GPA = −6% of std.
Pan and Ost (2014), Economics of Education Review	USA	Longitudinal survey	695	College enrollment	Gotshalk (LPM)	College enrollment = 10% −points
Mörk et al. (2020), Labor Economics	Sweden	Longitudinal register data	332,505–419,300	GPA & high school completion	Workplace closure and propensity score matching	GPA = −2.3% −points (only mother)
Hilger (2016), American Economic Journal: Applied Economics	USA	Longitudinal register data	Over 7 million fathers' layoffs	college enrollment	DID regression	College enrollment = −1% −points (only father)
Hajdu et al. (2019), Acta Oeconomica	Hungary	Longitudinal survey	4765	Secondary education attainment	Linear regression (LPM)	Secondary attainment = −4% −points
Gregg et al. (2012), Fiscal Studies	Britain	Survey	10,028	GCSE point score (GPA)	Linear regression	GCSE = −14% of std.
Di Maio and Nisticò (2019), Journal of Development Economics	Palestine	Longitudinal survey	9539	School dropout	Instrumental variable 2SLS regression	dropout = 9.2% −points
Coelli (2011), Labor Economics	Canada	Longitudinal survey	2403	Enrollment in university and community college (postsecondary enrollment)	Linear regression (LPM)	Any postsecondary −10.5, University −10.1% −points

Table 1 Detailed description of the review articles.—cont'd

<i>Author(s), year, journal</i>	<i>Country</i>	<i>Data type</i>	<i>Sample size</i>	<i>Outcome variable</i>	<i>Method</i>	<i>Effect size</i>
Müller et al. (2017), Oxford Economic Papers	Germany	Longitudinal survey	2,200, 1800	Completion and enrollment of upper secondary education, college attendance, educational attainment	Sibling fixed effect and Gottschalk	Daughters' upper sec attendance 9.6, upper sec. degree: 33.3, college attendance: 18.4%-points, years education: 0.540% of std.
Ermisch et al. (2004), Journal of the Royal Statistical Society: Series A	Britain	Longitudinal survey	1787, 1183	A-level (postsecondary attainment)	Sibling fixed effect and logit regression	A-level = -5.1%-points

In Britain, Gregg et al. (2012) studied the effect of fathers' job loss on children's GCSE grades during the recession of 1980–83, using the British Cohort Study, and observed the effect of the father's job displacement from major industry. They found that paternal job loss reduced children's GCSE grade points to 14% of the standard deviation lower than their counterparts who did not experience paternal job loss during the recession. In this study, maternal job displacement was not observed.

Mörk et al. (2020) studied the effect of parental job loss in Sweden. They used large and reliable Swedish register data on displaced and non-displaced workers, and propensity score matching. Their findings contradicted other studies that observed parental job loss effects on children's GPA. They reported that the effect was limited in Sweden and was associated only with maternal, but not paternal, job loss. They argued that Swedish welfare institutions successfully support families and provide protection against negative shocks in family. A dual-earner norm and strong incentives for female labor may have contributed to the negative effects of maternal job loss.

Grade repetition

Two of the articles studied the effect of parental unemployment on children's grade repetition. Both studies were conducted in the US, and both found a positive effect on grade repetition. Applying US survey data and logistic regression, Kalil and Ziol-Guest (2008) found an association between the father's job loss and the children's grade repetition and school suspension by two odds compared to those children who did not experience paternal unemployment. The effect was only related to paternal unemployment and was not statistically significant for maternal unemployment. In their subgroup analysis, they found that the association was only significant in households where mothers earned more than fathers.

Stevens and Schaller (2011) conducted a longitudinal survey from 1996 to 2004 in the US, and used an individual fixed effects estimation method. They found that parental job loss increased the probability of children's grade retention by 0.8 percentage points, or 15%.

The results of the two studies were aligned regarding the negative effect on children's school attainment; however, Stevens and Schaller (2011) found a lesser effect than Kalil and Ziol-Guest (2008), although the studies used the same Income and Program Participation survey. This may have been due to the different methods and sample features used in the studies. Kalil and Ziol-Guest (2008) used logistic regression and only one wave that was collected at 1999. Logistic regression can indicate only correlations between variables. In contrast, Steven and Schaller used longitudinal characteristics of the survey with a more accurate individual fixed effects estimation method to indicate causal negative effects of parental job loss.

School dropout at secondary level

Four papers analyzed dropout from secondary school. Studies were conducted in Finland (1), Hungary (2), and Palestine (1). All the studies found a negative effect, and the average effect of parental unemployment on children's secondary school attainment according to all four studies was -6.1 percentage points.

Kallio et al. (2016) used a longitudinal Finnish register dataset and random effects multilevel regression models. They observed children's secondary attainment until the age of 22 years and did not separate maternal and paternal unemployment. They found that parental unemployment, on average, decreased children's secondary attainment by five percentage points.

Kertesi and Kezdi (2008) investigated whether children's secondary school dropout was more likely at the age of 15–20 than if they were younger when they experienced parental job loss. They estimated the effect of parental job loss using Hungarian data at the time of the post-communist transition. They argued that the transition from a communist economy, where virtually everybody in the working-age population had a stable job, to post-communist, where many people lost their jobs, could be considered an exogenous effect. They used the probit regression model and a pooled longitudinal survey. Their sample included only two-parent families that were likely to have a lower skill distribution. They found a substantially stronger effect for younger children

to drop out from secondary education. The dropout rate increased by 5.8 percentage points for each year that the child was younger than 15 when they experienced parental job loss. They did not find any differences between maternal and paternal unemployment.

Hajdu et al. (2019) studied the effect of parental job loss on children's secondary school dropout rate in Hungary when they were 14–21 years of age. This study used linear probability models and a longitudinal panel survey. On average, the authors found that parental unemployment increased dropout from secondary education by four percentage points. However, they also indicated that the average negative effect was concentrated among families that did not provide a stimulus for their children at home. In these families, parental unemployment led to a decrease of nine percentage points in the school completion rate. They did not find any negative effects among families that provided a cognitively stimulating home environment. They also suggested that the home environment was more influential than family income in moderating the negative effect of parental unemployment.

Di Maio and Nisticò (2019) studied the parental job loss in a developing country, namely households that were living in the occupied Palestinian territories. They utilized quarterly variation in conflict intensity across districts in the occupied territories as an instrument for parental job loss in Israel. Their results indicated that job loss of the household head increased the probability of a child's school dropout by nine percentage points. They found that the negative effect varied with household characteristics, and that children with lower-educated parents were more strongly affected. According to their results, the negative effect stemmed from the income reduction in unemployed families.

These studies indicated that country context did not influence school dropout in developed economies, because the estimated effect sizes were similar in every country. The negative effect of parental unemployment was rather marginal in every developed country. However, the negative effect of parental unemployment was approximately twofold in Palestine, indicating that in developing countries, the negative effect of parental unemployment on school dropout could be more detrimental.

Upper secondary enrollment and completion

Four articles analyzed upper secondary school or high school completion and enrollment. The studies were conducted in Finland (1), Sweden (1), the US (1), and Germany (1). Two studies found a negative effect, and one did not find any effect. In addition, one study found a positive effect on daughters. According to these studies, the effect of parental unemployment on children's upper secondary attainment/enrollment was 5 percentage points on average. However, when one study that found a positive effect with a large error term was removed, the effect was negative at 5 percentage points.

Brand and Thomas (2014) studied how single mothers' job loss influenced children's high school completion by age 19 in the US. The study was conducted using propensity score matching and 30 years of nationally representative panel samples from the National Longitudinal Survey of Youth (NLSY). According to the study, single maternal unemployment lowered college attendance by about 5.2 percentage points. However, the negative effect was only observed at age 12–17. Because the children whose mothers had a lower likelihood of job loss had a larger negative effect on children's high school completion, the authors interpreted that the distress and social stigma associated with unemployment explained the results.

Lehti et al. (2019) studied the effects of parental unemployment (father or mother) on children's upper secondary enrollment at age 16 in Finland. They applied sibling fixed effects regression models. With the sibling FE method, they could compare siblings within a family where only some of the siblings were exposed to parental unemployment before they went to upper secondary school. They found that, on average, siblings who were exposed to parental unemployment had an average of 10 percentage points lower probability of tertiary enrollment. Similar to Brand and Thomas (2014), found that parental unemployment had a negative effect only in adolescents age 14–15 during the educational transition period. Parental unemployment had a negative effect only among children of lower-educated parents. Highly-educated parents could compensate for the negative effect of parental unemployment.

Mörk et al. (2020) found no effect of parental unemployment on children's high school completion. This study was conducted in Sweden, and the authors used Swedish register data on displaced and non-displaced workers and propensity score matching. Again, this indicated that Swedish welfare institutions could compensate for the negative effect of parental unemployment on high school completion.

Only the study by Müller et al. (2017) found a positive effect of parental unemployment on children's high school enrollment and completion (or any educational outcomes reviewed here). The study was conducted in Germany, using longitudinal German socioeconomic panel data and sibling fixed effects regression, as well as the Gottschalk method. They studied only the effects of paternal unemployment and found different patterns for sons and daughters. Using sibling FE models, paternal unemployment did not have any effect on sons' high schooling outcomes. Further, paternal unemployment did not influence daughters' enrollment, but it increased the likelihood for daughters to complete upper secondary education. Daughters who experienced paternal unemployment had, on average, 30 percentage points higher probability of completing upper secondary education compared to the sibling who did not. This could be considered a significant positive effect. However, this study was conducted with rather limited data, and many interactions were performed; thus, the error term of the study was significant (and non-reliable). The confidence intervals varied from 8.9 percentage points to 57 percentage points. The authors explained that daughters who experienced their fathers being unemployed were motivated to seek a partner with low unemployment risk and thus had a greater likelihood of completing upper secondary school. However, the test of the marriage market hypothesis was very vague in the study, and authors could only speculate that it was a mechanism behind the positive effect of parental unemployment on daughters' upper secondary attainment.

Tertiary education enrollment and completion

Nine studies of the reviewed articles analyzed the effects of parental unemployment and children's tertiary education enrollment and completion. In addition, two of the articles studied parental unemployment and children's educational attainment in years. On average, the studies indicated that parental unemployment during childhood could have a crucial impact on tertiary educational enrollment and completion. The reviewed studies revealed that children who experienced parental unemployment had, on average, 13 percentage points lower probability of achieving tertiary education compared to children who had not experienced parental unemployment. These studies were conducted in the US (4), Canada (1), Germany (2), Finland (1), and Britain (1).

In the reviewed articles, only two studies examined the effects of parental unemployment on children's educational attainment. In these studies, educational attainment was measured by in years. The studies were conducted in Germany (1) and the Netherlands (1).

Using longitudinal data from Canada, [Coelli \(2011\)](#) found that when children were at high school age (16–17), parental job loss affected post-secondary education enrollment. On average, children whose parents had experienced job loss had 10 percentage points lower enrollment in post-secondary education in general, and also enrollment in university education. He attributed this negative effect to unemployed parents' loss of income. Family disruption, namely parental divorce and household residential relocation, did not explain the effect of parental unemployment. According to the study, parental unemployment may have influenced the financial constraints of families, and that would have affected tertiary education enrollment in Canada, where tertiary education has tuition fees. This study did not distinguish between paternal and maternal unemployment.

[Lindemann and Gangl \(2019\)](#) studied the effect of parental unemployment on entering tertiary education in Germany. They used the longitudinal German Socio-Economic Panel and applied propensity score matching. They found that children exposed to parental unemployment had, on average, 13 percentage points lower likelihood of entering tertiary education, and that only paternal unemployment had this effect, while maternal unemployment had no effect. Contrary to Coelli's study in Canada, they did not find that family income mediated the negative effect of parental unemployment. However, subjective expectations and parental stress levels partly explained the negative effects of paternal unemployment. They also found cumulative effects of parental unemployment because the effect was more disadvantageous when the duration of paternal unemployment was longer.

In Finland, [Lehti et al. \(2019\)](#) studied the effects of parental unemployment (father or mother) on children's tertiary enrollment by applying the sibling fixed effects regression method. Thus, they could compare siblings within a family where only some of the siblings were exposed to parental unemployment before they enrolled in tertiary education. They found that, on average, siblings who were exposed to parental unemployment had a 13 percentage points lower likelihood of tertiary education enrollment. They found that the negative effect was only observed during the educational transition periods and among children whose parents were highly educated. According to their study, family income did not mediate the negative effect of parental unemployment on tertiary education enrollment. Paternal unemployment was, on average, more disadvantageous for children's tertiary enrollment than maternal unemployment, although the differences were not statistically significant.

[Kalil and Wightman \(2011\)](#) studied the association of parental unemployment and postsecondary education among children in Black and White middle-class households, at 21 years of age. They used US data from the Panel Study of Income Dynamics (PSIDs) and conducted multivariate regression analyses. They found that among White children, parental unemployment lowered college attendance by 9 percentage points. However, the negative association of parental unemployment and postsecondary education among Black children was three times higher (25 percentage points lower) compared to children who had not experienced parental unemployment. A large share of these different associations between parental job loss and children's college attendance could be explained by household wealth, family income, and parental experiences of long-term unemployment.

[Brand and Thomas \(2014\)](#) studied how single mothers' job loss affected children's college attendance by age 21, and completion by age 25, in the US. The study was conducted using propensity score matching and 30 years of nationally representative panel samples. According to this study, single maternal unemployment lowered college attendance by about 10 percentage points, and college completion, on average, by 4.3 percentage points.

Using the longitudinal Panel Study of Income Dynamics (PSIDs) from the US, [Pan and Ost \(2014\)](#) studied how the age when exposed to parental unemployment influenced children's college enrollment. They used Gottschalk and linear probability models. With this method, they could compare children who experienced parental unemployment at different ages. Further, in some analyses, they applied sibling fixed effects models to obtain causal estimates. They find that parental layoff was disadvantageous for the college enrollment at age 15–17 but not in the younger age groups. Parental layoff decreased tertiary enrollment by, on average, 10 percentage points in this age group, compared to children aged 21–23. Further, the authors studied heterogeneity by investigating the association between parental layoff and homeownership status, parental education, and local state tuition fees. They found that the interaction between unemployment and homeownership status was statistically significant. There was a stronger negative effect of parental unemployment on children whose parents did not own home compared to those whose parents were homeowners. Although the coefficient of the interaction between state tuition fees and unemployment predicted that the effect of unemployment was stronger in the states with higher tuition fees, the effect was not statistically significant. In addition, the interaction between parental education and unemployment was not statistically significant. The authors stated that the interaction estimates were only suggestive, because their sample was very small and the interaction should be interpreted with caution.

Hilger (2016) also studied the effect of parental layoff on children's college enrollment in the US, using longitudinal administrative data. His data included seven million layoffs of fathers of children age 12–29, and he used the difference-in-difference regression method. Although he found that unemployment greatly reduced family income, paternal layoff had only a one-percentage point negative effect on children's college enrollment.

Ermisch et al. (2004) studied how parental unemployment was associated with children's A-levels. They used sibling FE modeling and data from the British Household Panel Survey (BHPS). Using sibling FE models, they found that parental joblessness in early childhood reduced the likelihood of children achieving A-levels in Britain by five percentage points, on average. The effect did not vary significantly between daughters and sons.

Müller et al. (2017) studied how paternal unemployment influenced college attendance of children in Germany using longitudinal German socioeconomic panel data and sibling fixed effects regression models, as well as the Gottschalk method. Surprisingly, they found that daughters who experienced a father's unemployment had an 18 percentage points higher probability of attaining college education compared to their sister who did not experience paternal unemployment. This was the only study that found a positive effect of parental job loss on children's educational outcomes. However, they found a negative effect on sons (8.6 percentage points), but the effect was not statistically significant. They also studied educational attainment and found a positive effect on daughters, but no effect on sons. Paternal unemployment increased daughters' educational attainment by 0.5 years compared to their sisters who were not exposed to paternal unemployment.

Mooi-Reci et al. (2019) studied the association between parental unemployment and children's educational attainment in the Netherlands. They used data from the first three waves of OSA, which is a nationally representative longitudinal sample of over 2000 households and weighted least squares regression as a method. They found that fathers', but not mothers' unemployment decreased children's educational attainment by about half a year. This was attributed to parental views about work becoming more pessimistic during their unemployment.

One of the reviewed articles studied the relationship between parental unemployment and children's tertiary education in multiple countries. Lindemann and Gangl (2020) examined how parental unemployment was associated with children's tertiary education in a different institutional context. They used data from five longitudinal surveys and studied associations in 21 countries with multilevel regression models. The datasets used in this study were EU-SILC for 18 European countries, the SIPP panel for the US, the German Socioeconomic Panel (SOEP) for Germany, BHPS, as well as understanding society, and the UK Household Longitudinal Study (UKHLS) for Britain. Their results indicated that the countries that provided insurance against unemployment, namely social transfers, recorded lesser effects of parental unemployment. Further, in countries that provided more equalizing policies in terms of financial support to students, and where there was minimal private expenditure required for tertiary education, the effect of parental unemployment was less than in countries that did not provide these benefits. They observed strong effects in East and South European countries, such as Hungary, Bulgaria, Slovakia, Portugal, and Greece, as well as in the US. However, in the Nordic and some Continental European countries such as Sweden, Belgium, Austria, and Finland, the effects were remarkably less.

Mechanisms and institutional context

Two main mechanisms are suggested to be associated with parental unemployment and children's educational outcomes: (1) Economic mechanisms directly related to parental income reduction due to unemployment. (2) Non-economic mechanisms that initiated by parental unemployment, such as stress and other psychological factors that may influence the behavior of the parents and, subsequently, children's educational performance and choices.

It has been suggested that the mechanisms vary in different countries and institutional contexts (Lindemann and Gangl, 2020). For example, in liberal regimes such as the US, Canada, and Britain, where there are costly tuition fees for higher education, household income may be the primary mechanism that explains the negative effect of parental unemployment on higher education enrollment and completion. In contrast, in Nordic countries such as Finland, Sweden, and Norway, that have been described as social-democratic regimes, where social security is generous and higher education is free of charge, household income may not explain the effect of unemployment. However, parental loss of status and stress may have an effect.

Income loss and stress due to unemployment may also lead to other disadvantages among families, namely parental separation and residential mobility. These disadvantageous life-course events may increase children's stress and disrupt their educational attainment. Children within families that are forced to move for a parent's new job may suffer from stress and a lack of social networks.

The reduction of family income

One of the most obvious result of parental unemployment is the reduction of a family's economic resources, particularly consumption. The effects of economic resources on education are usually explained by parents' potential to invest monetary resources in their children and the material endowments available for the children's use (Becker and Tomes, 1979).

Some empirical evidence supports the assumption that the negative intergenerational effects of unemployment are at least partially related to a family's reduced household income. In Canada, Coelli (2011) found that parental job loss at high school age (16–17) reduced children's postsecondary education enrollment. He attributed this result to the income loss of unemployed parents. Kalil and Wightman (2011) found differences in the effect of fathers' job loss on White and Black children's tertiary

attainment that was mediated by household income and wealth. Household income reduction may have a particularly strong mediating effect in developing countries. In Palestine, as revealed by [Di Maio and Nisticò \(2019\)](#), parental unemployment was associated with income reduction and children's dropout from secondary education. However, this was the only reviewed study from a developing country. Further studies are needed to draw broader conclusions.

Although some studies have found that parental unemployment is associated with household income reduction, many studies have indicated that the negative effect of parental unemployment is not explained by household income reduction. Surprisingly, in the reviewed studies that directly investigated whether parental income mediated the effects of unemployment, only 3 studies out of 12 found a mediating effect.

One explanation for why many studies did not find the mediating effect of household income between parental job loss and children's educational outcomes is that the institutional context appear to influence this relationship. For example, in Finland, [Lehti et al. \(2019\)](#) studied the effect of parental unemployment after controlling family income, and the mediation effect of income loss was negligible. Also, [Rege et al. \(2011\)](#) found a similar result in Norway; the negative effect of paternal job loss was unrelated to the family income. In Sweden, paternal unemployment had no effect on children's educational attainment, and maternal unemployment had only a marginal negative effect ([Mörk et al., 2020](#)). In addition, study by [Lindemann and Gangl \(2020\)](#), using the different institutional contexts of 21 countries, indicated that higher social benefits and lower private educational charges in tertiary education reduced the effect of parental unemployment in general, and may have also influenced the mediation effect of parental unemployment.

Overall the reviewed studies clearly elucidated that institutional context influences the magnitude of the effect of parental unemployment on children's educational outcomes. Further, the mediation effect of parental income could only be found in countries where higher education required tuition fees.

Noneconomic mechanisms

A nonmonetary mechanism related to parental stress and psychological factors could also explain the negative effects of parental unemployment on children's educational outcomes. In general, it is difficult to estimate the extent to which parental stress mediates the negative effect of parental unemployment, because the reviewed studies did not measure stress levels directly. Many studies concluded that distress may be an explanatory factor behind parental unemployment. Other mechanisms, such as residential moves ([Stevens and Schaller, 2011](#)), parental dissolution ([Rege et al., 2011](#)), time spent with children ([Stevens and Schaller, 2011](#)), and income ([Kalil and Ziol-Guest, 2008](#); [Lehti et al., 2019](#); [Lindemann and Gangl, 2019](#); [Mörk et al., 2020](#)) did not explain the negative effects on children's educational outcomes. In addition, some studies found that parental unemployment was only harmful for children's educational outcomes among lower-status and lesser-educated families ([Brand and Thomas, 2014](#); [Lehti et al., 2019](#)). This could have indicated that unemployment generated more stress in families with lower socioeconomic resources. For example, [Brand and Thomas \(2014\)](#) studied the effect of single mothers' unemployment on children's college education, and found that unemployment was negatively associated with college enrollment and completion only in families that had a lower unemployment propensity. They concluded that the distress experienced by unemployed single mothers mediated the negative effect of unemployment on children's college attainment.

Further, according to the reviewed studies, paternal unemployment is more disadvantageous for children's educational outcomes than maternal unemployment. All but one study, that compared the effects of paternal and maternal unemployment, found that paternal unemployment had a stronger negative effect on children's educational outcomes ([Kalil and Ziol-Guest, 2008](#); [Lehti et al., 2019](#); [Lindemann and Gangl, 2019](#); [Mörk et al., 2020](#); [Rege et al., 2011](#)). Because previous studies have shown that men are mentally more distressed by unemployment ([Paul and Moser, 2009](#)), the results indicated that the stress that unemployment generated for men can have an impact on their children's educational outcomes.

One explanation for why unemployment affects men more than women is status loss ([Andersen, 2013](#)). Because men are usually considered the family breadwinners and men's earnings are on average higher compared to women, unemployment may have more detrimental effects on paternal rather than maternal status loss, creating stress that influences children and family dynamics in general. Indeed, some of the reviewed studies found that paternal status loss related to unemployment is associated with children's educational outcomes. For example, [Lindemann and Gangl \(2019\)](#) found that paternal unemployment was associated with children's educational aspirations.

Similar to [Lindemann and Gangl \(2019\)](#), [Lehti et al. \(2019\)](#) found that paternal unemployment is associated with less risky educational choices. Thus, the mediation effect of paternal unemployment is not only related to stress, but it can also trigger a sense of insecurity within families that can lead to the adoption of educational trajectories that are more secure, have more predictability, and thus involve fewer risks for the children. It was also found that paternal status loss due to unemployment was associated with children's school efforts ([Andersen, 2013](#)).

Overall, the psychological consequences are a more important factor explaining the disadvantageous effect of parental unemployment on children's educational outcomes than economic factors, particularly in the institutional context where social security is at a high level and education does not include any tuition fees. Although parental income does not mediate the negative effect of parental unemployment, the psychosocial consequences of unemployment remain and can harm children's educational performance and attainment.

Table 2 The average estimates of the reviewed studies by outcome variable.

Reviewed studies	GPA	Secondary/dropout	Upper secondary	Tertiary	Education attainment	Total
All the studies	−0,10	−0,06	0,04	−0,13	0,02	26
One study removed ^a	−0,125	−0,67	−0,05	−0,157	−0,51	21
All observations	5	4	4	11	2	26

Note. Because of different modeling strategies, the average effect cannot be calculated for studies that measured grade repetition.

^aA study with the most positive effect removed.

Conclusions

In this review article, I have reviewed 21 articles on the effects of parental unemployment on children's educational outcomes. The 21 articles included studies that examined outcomes such as GPA, grade repetition, school dropout at secondary level, secondary and tertiary education, and educational attainment. Articles included five studies on GPA, two on grade repetition, four on school dropout, four on upper secondary enrollment and completion, eleven on tertiary enrollment and completion, and two on educational attainment, totaling 28 studies.

The results of the reviewed articles indicated that parental unemployment or job loss had, on average, negative effects on children's educational outcomes. However, the effect varied according to country and institutional context, and was greater in countries with tuition fees and low social security.

Table 2 illustrates the average effects of the outcomes that were included in the review articles. According to the studies, children who experienced parental unemployment had, on average, 10% of the standard deviation lower GPA than children who did not experience parental unemployment. Parental unemployment increased the likelihood of dropout from secondary education by an average of six percentage points. Because one study found a significant positive effect for daughters regarding upper secondary and educational attainment, **Table 2** shows the average estimates without this study. The results of the study cannot be considered reliable because of the low number of cases and the method used. Thus, it can be suggested that parental unemployment reduces the likelihood of upper secondary enrollment and completion, on average, by 6 percentage points, and tertiary education, on average, by 13 percentage points. Parental unemployment decreases children's educational attainment by an average of five percent of the standard deviation.

Overall, the effects of parental unemployment can be considered small because the estimated effect sizes are lower than 10 percentage points. However, parental unemployment has the largest negative effect on children's tertiary education that can be considered a moderate effect size. A favorable institutional context equalizes the negative effect of parental unemployment on children's educational outcomes, and the average estimates (in **Table 2**) do not take into account that the negative effect varies between countries. However, future research should investigate, in more detail, how different institutional settings can amplify or mitigate the negative effects of parental unemployment, and the deeper mechanisms that can influence children's education when parents face unemployment.

Only one article analyzed in this study was conducted in a developing country. It is evident that more research is needed to draw conclusions on how parental unemployment influences children's educational outcomes in developing economies when there is no suitable social security system.

All the articles included in this review found at least a marginal effect of parental unemployment (one found a positive effect) on children's educational outcomes. However, this could also reflect publication bias, meaning that articles indicating negative effects had a greater likelihood of being published than those that did not find any effect.

References

- Andersen, S.H., 2013. Common genes or exogenous shock? Disentangling the causal effect of paternal unemployment on children's schooling efforts. *Eur. Socio Rev.* 29, 477–488. <https://doi.org/10.1093/esr/jcr088>.
- Becker, G.S., Tomes, N., 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *J. Polit. Econ.* 87, 1153–1189.
- Brand, J.E., 2015. The far-reaching impact of job loss and unemployment. *Annu. Rev. Sociol.* 41, 359–375.
- Brand, J.E., Thomas, J.S., 2014. Job displacement among single mothers: effects on children's outcomes in young adulthood. *Am. J. Sociol.* 119, 955–1001.
- Christoffersen, M.N., 1994. A follow-up study of longterm effects of unemployment on children: loss of self-esteem and self-destructive behavior among adolescents. *Childhood* 2, 212–220.
- Coelli, M.B., 2011. Parental job loss and the education enrollment of youth. *Lab. Econ.* 18, 25–35.
- Di Maio, M., Nisticò, R., 2019. The effect of parental job loss on child school dropout: evidence from the occupied Palestinian territories. *J. Dev. Econ.* 141, 102375.
- Ekhaugen, T., 2009. Extracting the causal component from the intergenerational correlation in unemployment. *J. Popul. Econ.* 22, 97–113.
- Ermisch, J., Francesconi, M., Pevalin, D.J., 2004. Parental partnership and joblessness in childhood and their influence on young people's outcomes. *J. Roy. Stat. Soc.* 167, 69–101.
- Gregg, P., Macmillan, L., Nasim, B., 2012. The impact of fathers' job loss during the recession of the 1980s on their children's educational attainment and labour market outcomes. *Fisc. Stud.* 33, 237–264.

- Hajdu, T., Kertesi, G., Kézdi, G., 2019. Parental job loss, secondary school completion and home environment. *Acta Oecon.* 69, 393–423.
- Hilger, N.G., 2016. Parental job loss and children's long-term outcomes: evidence from 7 million fathers' layoffs. *Am. Econ. J. Appl. Econ.* 8, 247–283.
- Kalil, A., Wightman, P., 2011. Parental job loss and children's educational attainment in black and white middle-class families. *Soc. Sci. Q.* 92, 57–78.
- Kalil, A., Ziol-Guest, K.M., 2008. Parental employment circumstances and children's academic progress. *Soc. Sci. Res.* 37, 500–515. <https://doi.org/10.1016/j.sresresearch.2007.08.007>.
- Kallio, J.M., Kauppinen, T.M., Erola, J., 2016. Cumulative socio-economic disadvantage and secondary education in Finland. *Eur. Socio Rev.* 32, 649–661.
- Karhula, A., Lehti, H., Erola, J., 2017. Intergenerational scars? The long-term effects of parental unemployment during a depression. *Res. Finn. Soc.* 10, 87–99.
- Kertesi, G., Kézdi, G., 2008. Children of the post-communist transition: age at the time of the parents' job loss and dropping out of secondary school. *B E J. Econ. Anal. Pol.* 7.
- Lehti, H., Erola, J., Karhula, A., 2019. The heterogeneous effects of parental unemployment on siblings' educational outcomes. *Res. Soc. Stratif. Mobil.* 64, 100439.
- Lindemann, K., Gangl, M., 2020. Parental unemployment and the transition into tertiary education: can institutions moderate the adverse effects? *Soc. Forces* 99, 616–647.
- Lindemann, K., Gangl, M., 2019. The intergenerational effects of unemployment: how parental unemployment affects educational transitions in Germany. *Res. Soc. Stratif. Mobil.* 62, 100410.
- Mooi-Reci, I., Bakker, B., Curry, M., Wooden, M., 2019. Why parental unemployment matters for children's educational attainment: empirical evidence from The Netherlands. *Eur. Socio Rev.* 35, 394–408.
- Mörk, E., Sjögren, A., Svaleryd, H., 2020. Consequences of parental job loss on the family environment and on human capital formation-Evidence from workplace closures. *Lab. Econ.* 67, 101911.
- Müller, S., Riphahn, R.T., Schwientek, C., 2017. Paternal unemployment during childhood: causal effects on youth worklessness and educational attainment. *Oxf. Econ. Pap.* 69, 213–238.
- Pan, W., Ost, B., 2014. The impact of parental layoff on higher education investment. *Econ. Educ. Rev.* 42, 53–63.
- Paul, K.I., Moser, K., 2009. Unemployment impairs mental health: meta-analyses. *J. Vocat. Behav.* 74, 264–282.
- Payne, J., 1987. Does unemployment run in families? Some findings from the General Household survey. *Sociology* 21, 199–214. <https://doi.org/10.1177/0038038587021002003>.
- Rege, M., Telle, K., Votruba, M., 2011. Parental job loss and children's school performance. *Rev. Econ. Stud.* 78, 1462–1489.
- Ruiz-Valenzuela, J., 2020. Job loss at home: children's school performance during the great recession. *SERIEs* 11, 1–44.
- Stevens, A.H., Schaller, J., 2011. Short-run effects of parental job loss on children's academic achievement. *Econ. Educ. Rev.* 30, 289–299.
- Ström, S., 2003. Unemployment and families: a review of research. *Soc. Serv. Rev.* 77, 399–430.

Diaspora and internationalization of higher education: a critical approach

Annette Bamberger, Hebrew University of Jerusalem, Department of Sociology and Anthropology, Jerusalem, Israel

© 2023 Elsevier Ltd. All rights reserved.

Introduction	128
Internationalization of higher education	129
Constructions of and approaches to diaspora in HE scholarship	129
Diaspora as human capital and state possession	129
Diaspora as marginalized other and social construct	130
A critical approach	131
Future directions	132
Conclusion	133
Acknowledgments	133
References	133

Introduction

The concept of diaspora, long associated with forced dispersion, exile from a “homeland” and oppression in “hostlands”, has undergone a significant shift in recent years. In the period of contemporary globalization, with its emphasis on global markets and connectivity, a greater demand emerged for “intercultural skills”, global competencies and networks. The advantageous nature of such global networks and skills, for both individual and national competitive advantage, saw a shift occur in which diaspora shed much of its pejorative connotations and became associated with a desirable form of cosmopolitanism which emphasized both dispersion (no longer exclusively traumatic) and connectivity (to a homeland, but more broadly, communities and people around the world, and circuits of migration and mobility) (Bamberger et al., 2021a; Rizvi, 2021). Although the COVID-19 pandemic has spurred increasing racism and nationalism, this positive view of diaspora largely prevails.

The intensive processes of contemporary globalization characterized by interconnectedness and interdependence have changed the conditions of diaspora formation and perpetuation, and its perceived value for individuals, communities, states, regional and supranational organizations. Individuals with the “right” forms of intercultural skills and global connections, have experienced a boon in their opportunities as markets globalized. States began to see the usefulness of diasporas in pursuing their own foreign relations and enhancing their global economic competitiveness. Countries as diverse as China, South Korea, Ireland, Israel, India, Australia have created diaspora schemes, part of what Gamlen (2020) traces as the explosion of diaspora institutions. Likewise, international and regional organizations such as the OECD, UNESCO, World Bank and EU have highlighted the importance and potential of diaspora in spurring development and entrepreneurship and crafted programs, policies, working reports, etc. to this effect. Higher education within these portrayals of development in which “skills transfer,” “knowledge diaspora,” “innovation” and the like are invoked, are closely linked with the oft-touted “diaspora option” and evolving discourses of “brain drain to circulation.” These growing trends connect national, regional, and supranational diaspora strategies, with the aspirations and identities of mobile people – particularly academics and students.

The increasing relevance and changing conditions of diaspora formations in the age of contemporary globalization has triggered a surge in diaspora scholarship and theorization across the humanities and social sciences, particularly in the fields of Migration, Human Geography, Political Science, Cultural Studies and increasingly in Education (e.g. Gholami, 2017; Rizvi, 2021; Shirazi, 2019). Such scholarship, combined with institutional, national, regional and supranational engagement, has broadened the meaning of and approaches to diaspora (Bamberger, 2021; Bamberger et al., 2021a). To the traditional idea of dispersion due to trauma, has been added a more generalized conception of “dispersion” and “exile” (Brubaker, 2017); binary views of homeland/host land are increasingly challenged, and networks, circuits and flows are more often invoked (Appadurai, 1990); homelands are no longer viewed as merely static, physical places, but rather, as symbols of identity, and important cultural discourses (Gilroy, 1993, 1997). These views indicate the different perspectives on what diaspora is and how it can be studied. Broadly, the proliferation of diaspora scholarship has reflected the expanding use of post-modern perspectives and theories, with much diaspora scholarship since the 1990s representing a move from diaspora as a static and essentialized notion of dispersed groups, maintaining distinctive identities to “an indicator of an identity in flux” (Delano and Gamlen, 2014, p. 44), hybridity and third space (Bhabha, 1994).

The relevance and appeal of diaspora, thus, appears widespread. However, it is not always clear how diaspora is being used and how it might represent a more critical and profound concept around which to understand key processes in higher education research, particularly around internationalization. Building on my recent research (Bamberger, 2020a, b, 2021; Kim and Bamberger, 2021; Bamberger et al., 2019, 2021a) I argue that diaspora is becoming a key concept in higher education research, closely connected to internationalization in particular; and that increasing diaspora theorization provides potential tools/lenses for understanding internationalization. I further argue that this potential is not being met, and I suggest a critical approach to analyze diaspora, outlining some possible avenues for future research.

Internationalization of higher education

Internationalization is one of the key features of contemporary higher education. Scholarship abounds on different forms of internationalization (e.g. curriculum, research collaborations), however, [Buckner and Stein \(2020\)](#) argue that the major focus has been on mobility – particularly of students and academic staff. [Rizvi \(2011\)](#) has shown how international mobility of students has, in recent years, been increasingly framed within the precepts of neoliberalism and market rationality. [Bamberger \(2020a\)](#) argues that this mobility is often framed as the rational pursuit of “cosmopolitan capital”, a form of competitive and positional advantage that is associated with being accustomed to travel and foreign cultures, having international social networks, and possessing prestigious credentials ([Bühlmann et al., 2013](#); [Igarashi and Saito, 2014](#)). For academic staff, this may extend to the cultivation of international academic networks and research funding, boosting individual advantage and positional worth in the increasingly stratified global higher education system ([Kim, 2017](#)).

The literature tends to assume that higher education credentials from the Global North are universally desired (for their largely economic returns on investment) and that economic and competitive considerations are a priori driving factors in international student/faculty mobility and migration. This approach does little to explain different trajectories of mobility (e.g. regional, horizontal mobility; North-South trajectories) ([Teichler, 2015](#)) and motivations for mobility. The literature often promotes an ahistorical and apolitical stance, in that it does little to engage with the myriad roots of contemporary mobility, often ignoring local (Global South) histories, cultures, legacies and enduring forms of global domination and coercion (e.g. colonialism), and undervalues individual considerations in mobility. This approach also tends to reflect a view of internationalization as an unconditional good within broadly humanitarian, progressive purposes; its more undesirable affects, such as a focus on meritocracy as opposed to equity, undervaluing “non-elite” forms of cosmopolitan capital, perpetuating inequalities in the Global South, etc. are often under investigated ([Bamberger et al., 2019](#)). Moreover, the role of the state in internationalization research is often relegated to that of a Western, democratic and “small” neoliberal state.

These initiatives, likewise, cast renewed interest in the role of the state in internationalization. A recent special issue of *British Journal of Educational Studies* ([Bamberger et al., 2021a](#)) reveals the intensive role of authoritarian states in diaspora policies and initiatives (e.g. [Brooks and Waters, 2021](#); [Han and Tong, 2021](#); [Rensimer, 2021](#)). The role of the state, however, in internationalization literature, [Bamberger et al. \(2019\)](#) argue, has been largely relegated to that of a market regulator. Thus, diaspora policies and initiatives draw our attention to the myriad roles which different states may take in their approaches to internationalization.

While there is renewed interest in the role of the state, there is also a dearth of critical research on the role of global institutions and international organizations exercising “global governance” on internationalization – beyond the limited scope of global university rankings. Little research has investigated the roles of the OECD, World Bank, UNESCO and other influential international organizations and regional blocs (outside of the EU) in internationalization. Moreover, little research has explored the role of transnational diaspora philanthropies and institutions in promoting and shaping internationalization.

Recognizing the limitations of the widespread economic approach, higher education scholars are searching for more nuanced and sophisticated theoretical lenses to analyze internationalization. In recent years, a research agenda which views mobile academics and students in processes of “becoming,” connecting HE with complex processes of self-formation, that entail multiple, interwoven intentions and identities that include but go beyond rational economic concerns has emerged (e.g. [Marginson, 2014](#); [Tran, 2016](#)). Within this emerging literature, scholarship has revealed the existence of international student mobility along diaspora trajectories (e.g. Israel ([Bamberger, 2020a](#)), South Korea ([Kim, 2011](#)); China ([Jian, 2017](#))). Diaspora has also been shown to play a role in creating international research collaborations, in which feelings of shared culture, language, religion and politics, have spurred connectivity ([Bamberger et al., 2019](#); [Fernando and Cohen, 2016](#)). Thus, an interest in rethinking (long-distance) nationalism and cosmopolitanism within globalization and agency is emerging.

Overall, diaspora has become a topical area of research in internationalization research. “Diaspora” provides a potentially powerful frame for interpreting the main themes of internationalization, moving beyond a focus on “mobility” (e.g. of people, ideas, programs) to focuses on movement, connectivity, identities, agency and control. Thus, I argue that diaspora is both a topical issue and important theoretical lens in contemporary internationalization studies in (higher) education in general, and in internationalization research in particular.

Constructions of and approaches to diaspora in HE scholarship

[Bamberger’s \(2021\)](#) systematic review of the use of diaspora in higher education research, revealed two widespread approaches: diaspora as human capital and state possession; and diaspora as marginalized other and social construct. Below I elaborate on these findings in detail.

Diaspora as human capital and state possession

Articles which invoked diaspora as human capital and state possession, defined people by their dispersion and degree of belonging to the state. Studies were framed around discourses of globalization, national development and competition. The growing importance of HE in driving innovation and national development was highlighted and the increasing competition between states for the

highly skilled. She finds that these studies employed literature, theories and concepts which were based in a “brain drain/gain/circulation” framework and took for granted the “pull” of global north countries and universities for “their” diaspora. States were engaged in attempts to reverse this trend, or at least to entice “their” diaspora to aid in national development from abroad. Against this backdrop, [Bamberger \(2021\)](#) finds that these articles primarily focused on describing and evaluating (state/university) diaspora policies and programs that either acquiesced to these global forces and aimed to create/maintain connections to its diaspora, who would presumably stay in the global north; or to combat these forces with both carrots and sticks, to alternatively entice and penalize scholars/students to encourage homeland engagement or repatriation. She finds that articles were overwhelmingly descriptive, prescriptive, and evaluative, focusing more on outcomes as opposed to processes.

So, what emerges in these articles is that “diaspora” is viewed through a human capital/neo-liberal frame. [Bamberger \(2021\)](#) argues that this produces a view of diaspora as an exogenous source of knowledge capital to be harnessed to the state’s desire for national development and to succeed in the global knowledge economy. There was little (if any) discussion of heterogeneity *within* this group. There was often unquestioned, affective allegiance toward the “homeland” or “motherland” on the part of “its” knowledge diaspora. It was the state (or universities, serving as a proxy for the state) that were leading diaspora initiatives and forging diaspora cooperation in HE. Diaspora members were assigned considerable goodwill and volunteerism on behalf of their “motherland” and were spurred to engage largely through their strong national identities. [Bamberger \(2021\)](#) argues that this rather simplistic approach likewise under analyzes the actual affective connections scholars had with “homelands” and the connections which exist.

[Bamberger \(2021\)](#) reveals that while there were some critical perspectives, these were neither aimed at universities nor states in the Global North, nor the competitive, neoliberal logics at play; rather, critical statements focused primarily on chiding states for not doing enough to harness the potential of their (supposedly eager) knowledge diaspora, thus missing an opportunity to improve their own national conditions. Alternatively, blame was laid on knowledge diasporas for not doing enough to forge connections, give back, and spur national development in their homelands. She finds that these views elicited (policy) criticism and recommendations from scholars, indicating that if states provided the right opportunities, engagement, knowledge and development would follow.

[Bamberger \(2021\)](#) shows that a large concentration of these studies addressed states with authoritarian governments, widespread corruption, and backgrounds of domestic strife and conflict. Despite this, she notes that there was surprisingly little discussion of “dispersion” (e.g. as a result of conflict, political oppression). More common, [Bamberger \(2021\)](#) argues, was an economic framing (often the push/pull approach) in which dispersion was portrayed as rational choice for scholars and students moving toward high-status cities in the global north – similar to the widespread framing of mobility in internationalization literature. This framing was supplemented by considerably underplaying political issues of mobility – and the desirability of return or engagement with the “motherland.” She argues that these ahistorical and de-political tendencies in the literature, disconnected diaspora from its “tragic” roots, indicating that diaspora has come to be a prized “possession” and an extension of the state.

[Bamberger \(2021\)](#) calls attention to which states use – and do not use – “diaspora” – and with what intention. With few exceptions, scholars hailing from the global north appear not to refer to repatriation/diaspora strategies in their own states – neither does it seem that states in the global north are in need of the “diaspora option.” Rather [Bamberger \(2021\)](#) argues that it may be that these Global North states are instead using different words like “highly talented” policies and do not refer to “their” “diasporas” as such (e.g. expatriates, international or global scholars) or to the role they play in creating other state diasporas. She argues that more attention to the types of states, their governments, discursive representations, and so on would greatly help to shed light on this phenomenon and its politics.

Overall, while the state centered approach may be useful for practitioners and/or policy-makers, intellectually, this approach perhaps sheds more light on the field of HE Studies, and normative views of many of its authors, than on diaspora. Criticality is limited and studies exude an essentializing tendency, with limited critical consideration of scholar/student agency. Its lack of reliance on scholarship within international relations, likewise does little to investigate the states themselves (e.g. mechanisms and strategies for creating diaspora; geopolitical positions of states; legitimacy; conflict) thus does little to shed light on diaspora and its connections with HE. In the absence of such analysis, the result is a partial narrative based on global capitalism and competition, which is certainly a part of the story, but cannot be all of it. There is likewise considerable underplaying of the implications these national/state narratives of diaspora have on scholars, students and institutions.

Diaspora as marginalized other and social construct

Articles which invoked diaspora as marginalized other and social construct, linked (implicitly and explicitly) diaspora with a social identity. Articles adopting this view, portrayed people primarily by their (externally imposed and/or self-identified) belonging to an ethnic, racial or religious group. However, there were several studies which went beyond these categories, and employed “diaspora” to understand identity formation processes and experiences of academics and students in HE that were defined along other social categories (e.g. heteronormative culture; first generation), representing a considerable expansion of the term diaspora, to include those with similar social characteristics, or even intellectual interests (e.g. [Jöns et al. \(2015\)](#) “elective” knowledge diaspora).

[Bamberger \(2021\)](#) finds that articles that employed this approach were more concerned with understanding, exploring and critiquing, than with description, prescription and evaluation. Often critical theories were employed, particularly Feminist, Post-Colonial and Pan-African. Literature framed diaspora as a way to explore individual and collective identities, experiences and practices – and the role of HE in this process. Articles emphasized heterogeneity in “diaspora” and several articles aimed to

“unpack” diaspora and its homogenizing effects. These articles often called for greater equality between diasporic views in and de-centering of dominant/majority diaspora views.

Bamberger (2021) reveals that diaspora identities and experiences were alternatively expressed as sources of social alienation, empowerment and support. HE was viewed as an institution to foster, reshape and grapple with a social identity (e.g. through heritage study abroad programs, courses on heritage languages and histories); or as an institution with inherent structural biases and inequalities which should be reformed – but that in the meantime, should be managed through workarounds by students/scholars. Thus, Bamberger (2021) calls attention to the extent of much of the criticality in these articles. She argues that while it is essential for students and scholars to find and build support networks to thrive within discriminatory systems, a lack of attention to social structures themselves, at times led to a discourse that was similar to rationalist (albeit not economic) discourses. She argues that this lack of attention to social structures risks leading to over glorification of hybridity and agency without significant focus on oppressive and discriminatory social structures. Bamberger (2021) further finds that social attributes and categories were largely unquestioned and taken for granted (e.g. few articles focused on how students/staff were “diasporized” – even if racial, ethnic, religious and other social categories were critiqued and problematized). Rather the “fact” that they were part of a diaspora was often taken for granted, and the analysis focused on how individuals used their agency to shape these identities in and through HE.

Overall, considering the portrayals, foci and discourse of the two types of diaspora scholarship, it is evident that they broadly reflect well-rehearsed criticisms of their respective approaches to diaspora. Essentialist/outcomes-based approaches are critiqued for reflecting homogenizing discourses of “the” diaspora (e.g. around nation, state, ethnicity) and for its static view of social reality. Constructivist/post-modern approaches are critiqued for its loose conceptualization of “diaspora,” its focus on agency and neglect of social structures which can at times echo rationalist discourses (see Vertovec, 1997); and for its inattentiveness to history (Alexander, 2017).

In sum, Bamberger’s (2021) review of the field indicates that although there is growing scholarship on diaspora and higher education, the literature is fragmented between those concerned with “knowledge diaspora” that focus on national and institutional diaspora policy and practice (e.g. evaluation, implementation and improvement); and those that focus on the identity processes and experiences of diaspora individuals. Moreover, “diaspora” – is loosely connected to internationalization in higher education research, despite the clear connections with internationalization at the system, institution, and individual levels. Bamberger posits that this is likely because “internationalization” is associated with humanitarianism, cosmopolitanism, global citizenship and the de-centering of states, nations, ethnicities, religions, and particularistic identities (Bamberger et al., 2019). Diaspora is often connected with these latter categories and thus, is shunned as nationalist, essentialist, and so on. However, Bamberger (2021) also argues that the blending of (often conflictual) ideas, intentions, values and identities is precisely what diaspora engages in. And thus, research approaches which aim to connect these areas, and bridge their respective weaknesses are sorely lacking.

A critical approach

I propose a critical approach to bridge the chasm in contemporary diaspora research and to guide a future research agenda. Such an approach would be based on seven precepts: eschewing (national) exceptionalism; developing alternative views of universality; operating from diverse historical bases; employing theoretical diversity and flexibility; pursuing methodological pluralism; focusing on a multiplicity of units of analysis; and developing new bridges and understanding of individual agency and social structure. These tenets and the critical approach I put forward, resonate with Acharya’s (2014, 2016) “Global International Relations” approach, however, also differs in important ways. Below I elaborate.

Eschewing (national) exceptionalism: As revealed in the previous section, there is a distinct strand of literature which views nation-states, institutions or individuals from the perspective of exceptionalism. While the climate of global competition has spurred many nation-states toward narratives of national exceptionalism in HE, a discourse particularly reflected in studies on national branding campaigns to attract international students (e.g. Lomer et al., 2018; Stein, 2018), however, it is imperative for scholarship to maintain a critical distance from such national, promotional narratives. Likewise marketing tropes which address individuals such as “talents” should be critically considered. Thus, an approach which consciously aims to critically interrogate such claims and tropes of exceptionalism – instead of reproducing them - will shed greater light on issues of diaspora and internationalization.

Alternative views of universality: There is a tendency to view a particular reality (often Anglo-European) as universal and to project this into the Global. However, such projections may clash with and/or provide limited explanatory value or hold limited purchase in other societies. An approach which is conscious of such tendencies, and that reflects diverse social and ontological views, and lived realities; greater self-awareness of situated knowledge; caution when applying “borrowed” terms, and recognition of their development and connotation in different contexts, would be beneficial. Such a view would also critically interrogate “Western” views of the “diaspora option” (i.e. as widely positive for combating brain drain, often caused singularly by their own actions; universally appealing to those on the “losing” end of the brain drain equation).

Diverse historical bases: As opposed to primarily Western historical bases for the development of HE and internationalization, a critical approach would recognize a plurality of historical bases from which to view antecedents and developments in these areas. It would also recognize how diverse aspects of world history, often viewed through the lens of the Anglo-European tradition, have different impacts and implications across the world.

Theoretical diversity and flexibility: Scholars must pay attention to the wider theoretical literature and advances in other fields. This should not be a difficult task; HE Studies is already an interdisciplinary field, which borrows widely from other disciplines

(Tight, 2018). However, it is important to consider the extent to which empirical research informs theoretical choices and explanations. Some critical internationalization scholars adopt a priori theoretical and political stances, espousing an agenda for social justice and advocating for radical change. While it is sensible to view scholarship as a tool to better society, greater understanding is the first step to social change, and thus a priori theoretical approaches – which may be led by researcher interest, instead of grounded in empirical data – should be avoided. Such an approach, based in deep empirical investigation avoids many pitfalls, particularly applying theoretical frameworks which are inappropriate to discover and explore other possible influences and interpretive frameworks which could hold explanatory weight in particular contexts. Thus, this approach takes a position that is critical of projecting a priori theories and explanations (particularly in contexts outside of the West); and emphasizes the need for theoretical plurality and flexibility to study and promote understanding of diaspora and internationalization of higher education, and its different articulations across contexts.

Methodological pluralism: It is valuable to bridge the respective gaps in the different approaches in the diaspora and HE literature. For example, many studies focus on the outcomes of policy analysis, while ignoring scholar/student agency; and likewise studies which take the social construction/postmodern approach, tend to underplay/analyze social structures, widespread discourses of state policies, and historical antecedents. To nuance the literature, a shift toward a more critical perspective which focuses on processes and mechanisms (as opposed to solely outcomes), is an important step. Such a step would be aided by methodological plurality. More varied research samples, which would include a wider range of participants (e.g. those who may not espouse “volunteerism” to the motherland); diverse research methodologies, specifically purposeful comparative studies, may have much to offer in understanding the conditions and mechanisms of diaspora engagement; and wider forms of research “data” (e.g. moving beyond policies, surveys and interviews, to include analyses of curricular components, academic networks, historical/archival materials, media analyses) among others.

Multiplicity of units of analysis: While methodological nationalism has been a widespread critique of comparative education, and prominent scholars, notably Cowen (2000) have encouraged scholars to “read the global,” there is still a considerable emphasis on the role of the state. While the state is an important unit, it should not be taken as a single unit, or the only unit. More emphasis on diversity within “the” state and on regions, supranational/international organizations, and local, global and transnational actors would shed considerable light on this area.

Agency and structure: Agency has tended to be largely placed with nation-states or individuals, however, there is a need to recognize multiple sources of agency, as well as continued consideration of how different forms of agency interact with social structures. Specifically, such considerations would avoid the volunteerism narrative in the state possession/human capital framing of diaspora, and also avoid glorifying individual agency while diminishing critiques of discriminatory/oppressive social structures.

Future directions

Employing the precepts outlined in the previous section, a research agenda around diaspora and higher education could be taken in many different directions. Below, I outline a few examples, which indicate how such an approach could be applied to reconsider/explore higher education institutions; the role of the state; and supranational/international and regional organizations.

Critical analysis of national diaspora policies, programs, institutions and actors – which eschews national exceptionalism and evaluative/descriptive/prescriptive tendencies – has considerable potential to shed light on the roles of (authoritarian) states in internationalizing higher education – both at home and abroad (e.g. China’s encouragement of students to study abroad). Such scholarship would greatly benefit from combining critical approaches with diverse disciplinary theories, particularly those from international relations (e.g. when/how states “deploy” diaspora; geopolitical conditions for nation-building outside state borders; state securitization and legitimization). It would likewise analyze which states are involved in “diaspora”; while recent work has uncovered the strong connection between authoritarian states and national diaspora policies and programs (e.g. Brooks and Waters, 2021; Han and Tong, 2021), more needs to be done to understand this. Likewise, Koinova and Tsourapas (2018), argue that “the state” is often used as a broad, unitary category. Thus, while state actions, discourses, and strategies should be examined, more attention to the plurality of the state is important (e.g. political parties; comparative analyses of different state ministries and departments) could likewise shed light on the plurality of the state, and diverse voices in diaspora “deployment” and relations with the state through HE. Such approaches may reveal the mechanisms, strategies and conditions for state-diaspora engagement and have wide reaching implications for HE. It likewise has the potential to reveal divergent forms of internationalization, shedding light on its multiple purposes and uses around the world. This approach could possibly reveal new discourses around internationalization which are starkly in contrast with much of the dominant internationalization discourses organized around the “unconditional good” or internationalization as a progressive humanitarian/neoliberal construct (Bamberger et al., 2019). This approach could likewise shed greater light on a (re)emerging and increasingly common trope of “soft power” in the role of higher education in international relations, in claiming “hearts and minds” in what is shaping up to be a new global battle for superpower status. Applying a critical approach, Bamberger et al. (2021) alight on a form of internationalization practiced by China and Israel toward Hong Kong and East Jerusalem respectively; it aims to bolster national forms of identity and extend state control over “troublesome” minorities within the nation state. This form of internationalization operates within a broader program of “internal colonialization” that is neither well developed in the literature nor explained by prominent typologies of internationalization.

Institutional analysis of diaspora policies and practices tend to be limited to evaluative/descriptive studies, often in relation to the “execution” of national policies and programs, within universities/HEIs, often focusing on the mobility (“return”) of diaspora

scholars. This leaves much to be explored. There are many different ways in which institutions can craft, lead, and respond to diaspora initiatives from diverse actors. Critical analysis of institutional policies, practices and actors – and their relationship with the state, diaspora individuals, donors, and communities – may shed light on issues of governance and agency which would be insightful. Moreover, beyond HEIs, philanthropic institutions, or NGOs may have diaspora missions which impact on HE, however, may operate outside of the HE system, or direct state control, and may be led by diaspora individuals/donors/communities. Increasing focus on the role of these types of institutions would likewise be insightful.

Supranational/international and regional organizations policy and program analysis is lacking. There has been little critical policy analysis of international organizations role in creating and perpetuating discourses around diaspora, and the implications for higher education. My intuition is that the numerous research papers, policies and programs emanating from such organizations as the UN, OECD, and World Bank promote concepts of diaspora that may be particularly influential in understanding how the “diaspora option” is being portrayed and deployed as part of the global governance of these organizations – and its reliance on the highly skilled and higher education institutions as bases for recruitment and implementation. Bamberger’s (2021) review indicates, and Brooks and Waters’ (2021) study suggests, that target states for such policies may not have fully embraced the simplistic “brain circulation” and “diaspora option” narrative propagated by these organizations. Thus, a critical analysis of these initiatives and their negation and implementation across contexts may shed light on resistance to supranational and regional governance, how HE is implicated therein and the implications which then arise for HEIs, scholars and students.

Conclusion

Over the past decades, a shift has taken place in which diaspora, once a pejorative term associated with a marginalized minority group, in the era of contemporary globalization became associated with a desirable form of cosmopolitanism. This shift in perceptions dovetailed with a significant expansion of diaspora theorization across the humanities and social sciences. I argued that diaspora is becoming a key concept in higher education research, closely connected to internationalization in particular; and that increasing diaspora theorization provides tools/lenses for investigating key areas of internationalization in HE.

The relevance and appeal of diaspora, thus, appears widespread. However, it is not always clear how diaspora is being used and how it might present a more critical and profound concept around which to understand key processes in higher education research, particularly around internationalization. Drawing on a systematic review of the literature (Bamberger, 2021), I revealed and critiqued dominant portrayals and approaches to diaspora within HE research. I demonstrated that although there is growing scholarship on diaspora and higher education, the literature is fragmented between those concerned with “knowledge diaspora” that focus on national and institutional diaspora policy and practice (e.g. evaluation, implementation and improvement); and those that focus on the identity processes and experiences of diaspora individuals. This fragmentation is largely due to differences in approach and notably, results in limited criticality. Moreover, this research largely ignores diaspora scholarship as part of internationalization, despite its clear connections.

The potential for expansion of a diaspora research agenda in higher education is timely, however, I argued that alternatives and a fresh approach to thinking and researching diaspora in (higher) education is needed. I suggested a critical framework to analyze diaspora, which I propose will shed light on this emerging phenomenon. This critical approach would be based on seven precepts: eschewing (national) exceptionalism; developing alternative views of universality; operating from diverse historical bases; employing theoretical diversity and flexibility; pursuing methodological pluralism; focusing on a multiplicity of units of analysis; and developing new bridges and understanding of individual agency and social structure. As diaspora institutions and initiatives expand, and global mobility and connectivity continue, fresh ways of thinking about diaspora and internationalization in higher education offer consider scholarly potential. This potential is increased in light of rising nationalism and bordering practices during the COVID-19 pandemic.

Acknowledgments

This research was supported by the Golda Meir Post-Doctoral Fellowship Fund.

References

- Acharya, A., 2014. Global international relations (IR) and regional WorldsA new agenda for international studies. *Int. Stud. Q.* 58 (4), 647–659.
- Acharya, A., 2016. Advancing global IR: challenges, contentions, and contributions. *Int. Stud. Rev.* 18 (1), 4–15.
- Alexander, C., 2017. Beyond the “The ‘diaspora’ diaspora”: a response to Rogers Brubaker. *Ethn. Racial Stud.* 40 (9), 1544–1555.
- Appadurai, A., 1990. Difference and disjuncture in the global cultural economy. *Theor. Cult. Soc.* 7 (2), 295–310.
- Bamberger, A., Morris, P., Yemini, M., 2019. Neoliberalism, internationalisation and higher education: connections, contradictions and alternatives. *Discourse* 40 (2), 203–216.
- Bamberger, A., Kim, T., Morris, P., Rizvi, F., 2021a. Diaspora, internationalization and higher education. *Br. J. Educ. Stud.* 69 (5), 501–511.
- Bamberger, A., Yan, F., Morris, P., 2021b. Adapting ‘internationalization’ to integrate ‘troublesome’ minorities: higher education policies towards Hong Kong and East Jerusalem. *J. Educ. Policy*, 1–23.
- Bamberger, A., 2020a. Accumulating cosmopolitan and ethnic identity capital through international student mobility. *Stud. High Educ.* 45 (7), 1367–1379.

- Bamberger, A., 2020b. *Diaspora, State and University: An Analysis of Internationalisation of Higher Education in Israel*. Doctoral dissertation. UCL (University College London).
- Bamberger, A., 2021. From human capital to marginalized other: a systematic review of diaspora and internationalization in higher education. *Br. J. Educ. Stud.* 1–23.
- Bhabha, H., 1994. *The Location of Culture*. Routledge, London.
- Brooks, R., Waters, J., 2021. International students and alternative visions of diaspora. *Br. J. Educ. Stud.* 1–22.
- Brubaker, R., 2017. Revisiting “the ‘diaspora’ diaspora”. *Ethn. Racial Stud.* 40 (9), 1556–1561.
- Buckner, E., Stein, S., 2020. What counts as internationalization? Deconstructing the internationalization imperative. *J. Stud. Int. Educ.* 24 (2), 151–166.
- Bühlmann, F., David, T., Mach, A., 2013. Cosmopolitan capital and the internationalization of the field of business elites: evidence from the Swiss Case. *Cult. Sociol.* 7 (2), 211–229.
- Cowen, R., 2000. Comparing futures or comparing pasts? *Comp. Educ.* 36 (3), 333–342.
- Delano, A., Gamlen, A., 2014. Comparing and theorizing state–diaspora relations. *Polit. Geogr.* 41, 43–53.
- Fernando, W.D.A., Cohen, L., 2016. Exploring career advantages of highly skilled migrants: a study of Indian academics in the UK. *Int. J. Hum. Resour. Manag.* 27 (12), 1277–1298.
- Gamlen, A., 2020. *Human Geopolitics: States, Emigrants, and the Rise of Diaspora Institutions*. Oxford University Press, Oxford.
- Gholami, R., 2017. The ‘sweet spot’ between submission and subversion: diaspora, education and the cosmopolitan project. In: *Diaspora as Cultures of Cooperation*. Palgrave Macmillan, Cham, pp. 49–68.
- Gilroy, P., 1993. *The Black Atlantic: Modernity and Double Consciousness*. Harvard University Press, Cambridge.
- Gilroy, P., 1997. Diaspora and the detours of identity. In: Woodward, K. (Ed.), *Identity and Difference*. Sage, pp. 299–346.
- Han, C., Tong, Y., 2021. Students at the nexus between the Chinese diaspora and internationalisation of higher education: the role of overseas students in China’s strategy of soft power. *Br. J. Educ. Stud.* 1–20.
- Igarashi, H., Saito, H., 2014. Cosmopolitanism as cultural capital: exploring the intersection of globalization, education and stratification. *Cult. Sociol.* 8 (3), 222e39.
- Jiani, M.A., 2017. Why and how international students choose mainland China as a higher education study abroad destination. *High Educ.* 74 (4), 563–579.
- Jöns, H., Mavroudi, E., Heffernan, M., 2015. Mobilising the elective diaspora: US–German academic exchanges since 1945. *Trans. Inst. Br. Geogr.* 40 (1), 113–127.
- Kim, J., 2011. Aspiration for global cultural capital in the stratified realm of global higher education: why do Korean students go to US graduate schools? *Br. J. Sociol. Educ.* 32 (1), 109–126.
- Kim, T., 2017. Academic mobility, transnational identity capital, and stratification under conditions of academic capitalism. *High. Educ.* 73 (6), 981–997. <https://doi.org/10.1007/s10734-017-0118-0>.
- Kim, T., Bamberger, A., 2021. Diaspora, ethnic internationalism and higher education internationalization: the Korean and Jewish cases as stateless nations in the early 20th century. *Br. J. Educ. Stud.* 1–23.
- Koinova, M., Tsourapas, G., 2018. How do countries of origin engage migrants and diasporas? Multiple actors and comparative perspectives. *Int. Polit. Sci. Rev.* 39 (3), 311–321.
- Lomer, S., Papatsiba, V., Naidoo, R., 2018. Constructing a national higher education brand for the UK: positional competition and promised capitals. *Stud. High Educ.* 43 (1), 134–153.
- Marginson, S., 2014. Student self-formation in international education. *J. Stud. Int. Educ.* 18 (1), 6–22.
- Rensimer, L., 2021. Negotiating educational choices in uncertain transnational space: South Asian diaspora in the United Arab Emirates. *Br. J. Educ. Stud.* 1–22.
- Rizvi, F., 2011. Theorizing student mobility in an era of globalization. *Teach Teach: Theory Pract* 17 (6). <https://doi.org/10.1080/13540602.2011.625145>.
- Rizvi, F., 2021. International higher education and the formation of business diasporas. *Br. J. Educ. Stud.* 1–20.
- Shirazi, R., 2019. “Somewhere we can breathe”: diasporic counterspaces of education as sites of epistemological possibility. *Comp. Educ. Rev.* 63 (4), 480–501.
- Stein, S., 2018. National exceptionalism in the ‘EduCanada’ brand: unpacking the ethics of internationalization marketing in Canada. *Discourse* 39 (3), 461–477.
- Teichler, U., 2015. Academic mobility and migration: what we know and what we do not know. *Eur. Rev.* 23 (S1), S6–S37. <https://doi.org/10.1017/S1062798714000787>. Mobility and Migration in Science, May 2015.
- Tight, M., 2018. *Higher Education Research: The Developing Field*. Bloomsbury Publishing.
- Tran, L.T., 2016. Mobility as ‘becoming’: a Bourdieuan analysis of the factors shaping international student mobility. *Br. J. Sociol. Educ.* 37 (8), 1268–1289.
- Vertovec, S., 1997. Three meanings of “diaspora,” exemplified among South Asian religions. *Diaspora A J. Transnatl. Stud.* 6 (3), 277–299.

Education of refugees

Helen Avery^{a,b}, ^a Centre for Advanced Middle Eastern Studies, Lund University, Lund, Sweden; and ^b Department of Languages, Linnaeus University, Växjö, Sweden

© 2023 Elsevier Ltd. All rights reserved.

The position of refugee education in a globalized world	136
Who is a refugee learner? Numbers, global distribution and the issue of definitions	137
Refugee learners and the right to education	137
Particular challenges in refugee education	138
Geographical distribution and focus of research on refugee education	138
Pedagogical and psychosocial challenges in international education	139
Challenges in refugee education compared to other international and transnational education	139
Language policies in refugee education and in the global educational landscape	140
Curricular content and pedagogical orientation in refugee education	141
Brain drain, repatriation of refugees and post-conflict recovery	141
Conclusions and further research	142
References	143
Further reading	146
Relevant websites	147

Glossary

Country of origin A country of nationality or of former habitual residence of a person or group of persons who have migrated abroad, irrespective of whether they migrate regularly or irregularly (IOM, n.d.)

Diaspora A population scattered outside a country, region or location, and who retain connections to this location, as well as to other individuals and communities with a similar origin

Displacement The movement of persons who have been forced or obliged to flee or to leave their homes or places of habitual residence, in particular as a result of or in order to avoid the effects of armed conflict, situations of generalized violence, violations of human rights or natural or human-made disasters (IOM, n.d.)

Global education (GE) Aims at developing awareness and appreciation of cultural difference, recognition of planetary interdependence, social injustice and environmental degradation, as well as commitment to social engagement on these issues (Gaudelli, 2007)

HIC High-income countries—country classification based on economy as used by the UN World Economic Situation and Prospects (WESP), 2021

Internally Displaced Persons (IDPs) Persons or groups of persons who have been forced or obliged to flee or to leave their homes or places of habitual residence (see Displacement) but “who have not crossed an internationally recognized State border” (IOM, n.d.)

International education A broad range of practices, that range from promoting international understanding and peace (UNESCO, 1974), to marketized international schools or internationalization in higher education

LMIC Low- and middle-income countries—country classification based on economy as used by the UN World Economic Situation and Prospects (WESP), 2021

Mobility regimes The extent to which processes of globalization are also concerned with the prevention of movement and blocking of access (Shamir, 2005)

Open Education (OE) A philosophy and practices that aim to give access to education for all, including formal and non-formal education, as well as informal learning communities

Refugee Two main definitions: (1) “Under international law and UNHCR’s mandate, refugees are persons outside their countries of origin who are in need of international protection because of feared persecution, or a serious threat to their life, physical integrity or freedom in their country of origin as a result of persecution, armed conflict, violence or serious public disorder.” (UNHCR, n.d.) and (2) “A person who, owing to a well-founded fear of persecution for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country; or who, not having a nationality and being outside the country of his former habitual residence as a result of such events, is unable or, owing to such fear, is unwilling to return to it.” (1951 Convention, IOM, n.d.)

Refugee education Formal and non-formal education provided for refugee learners, whether in educational settings dedicated to refugees, integrated in national education systems or provided through international education

Refugee learner Any learners whose educational opportunities and career prospects are affected by their own or their families' experience of forced displacement

Stateless person A person who is not considered as a national by any State under the operation of its law (IOM, n.d.)

Transnational education (TNE) All types of higher education study programs, or sets of courses of study, or educational services (including those of distance education) in which the learners are located in a country different from the one where the awarding institution is based (UNESCO/Council of Europe, 2001).

When not otherwise specified, terminology on forced displacement and education follows IOM and UNESCO definitions.

The position of refugee education in a globalized world

Refugee education takes place in a highly politicized context. Various geopolitical and economic interests lie behind armed conflicts across the globe and motivate attitudes of third countries to parties to these conflicts, as well as to the neighboring countries that receive refugee flows (Chimni, 1998; Hyndman, 2013; Avery and Said, 2017; Micinski, 2018; Morris, 2021; Abdelaaty, 2021). Authoritarian regimes are supported or condemned, depending on economic interests in the natural resources or strategic location of the country. Xenophobic and nationalist discourses are deliberately used by politicians to rally support for military agendas, or to avoid responsibility for economic or social problems by instead diverting public attention to the convenient scapegoat of the “foreigner” or other minority groups (Gulmez, 2019; Rossell Hayes and Dudek, 2020; Wang, 2020; Wasem, 2020; Winn, 2021).

Such political and geopolitical agendas shape conditions for teaching and learning, and thus cannot be neglected when considering purely “pedagogical” issues in education. Profound insecurity—and frequently also hostility from the population of host countries—are among the most clearly defining features of refugee education, and which distinguish it from other kinds of education or other educational situations.

For refugee learners, the overall political agendas within which refugee education is embedded translate in the best of cases into a “culture of welcome” (Gereke and Nijhawan, 2019), supported by state policies. For many, however, their studies instead take place in a context of exclusion and bullying (Guo et al., 2019). The hostility of host populations may be aggravated if there is a perception that refugees are “stealing jobs”, or that they benefit from state or humanitarian support that is not available to the country’s own disadvantaged populations. Often welcoming and excluding policies or drivers exist side-by-side in the same national context (Bellino and Dryden-Peterson, 2019; Garvik and Valenta, 2021). Education and support measures may thus be offered with a humanitarian or integration rationale, at the same time that other elements of national policy maintain permanent insecurity regarding settled status, or close pathways to further education and future employment (see Dryden-Peterson et al., 2019; Lee et al., 2020; Baker et al., 2021).

Depending on the context, other geopolitical issues can have significant impacts on the lives and future of refugee youth. This notably includes recruitment as child soldiers by non-state armed groups (Darden, 2019), or by armed groups used by states against neighboring countries. Various forms of “refugee militarization” are used to justify policies of “securitization” (Nanopoulos et al., 2018; Ansorg, 2020), which in turn contribute to deteriorating conditions in which refugee youth live and study, rendering them increasingly vulnerable to exploitation. Host countries may also use refugee populations as money of exchange (refugee “rent-seeking behavior”, see Tsourapas, 2019) to pressure agreements with the European Union or other high-income regions that wish to prevent the entry of refugees. Viewing refugees as a commodity can thus constitute an incentive to perpetuate a state of crisis, as well as to keep refugees in camps where their numbers are countable and their plight is visible. Although corruption and the monetization of refugees in host countries is a well-known phenomenon, donors and international agencies are reluctant to support refugee-led initiatives (Alio et al., 2020; Pincock et al., 2021), leaving a relatively small number of independent NGOs financed by diasporas that are committed to education by refugees, for refugees.

A considerable amount of research has been devoted to the structural ways globalization affects education (see e.g., Stromquist, 2002; Stromquist and Monkman, 2014; Zajda and Rust, 2021). This includes phenomena such as marketization of education, standardization, national curricular content geared toward the needs of transnational corporations, or brain drain of competence toward these corporations and countries of the global North (Marsh and Oyelere, 2018; UNESCO, 2018a). Much of academic research in the field highlights negative impacts on countries of the global South, but may also discuss more general effects of mobility and interconnectedness on identities, life trajectories and worldviews (Rizvi, 2019), or concern strategies adopted by education institutions—particularly in higher education—in order to remain competitive on global education markets. Another body of literature is dedicated to policies at national levels, concerning for instance migration, mobility, and strategies to secure qualified labor, while international organizations tend to frame the issues in terms of development, crisis response, humanitarian aid or rights to protection. While very little of the literature on globalization and education explicitly focuses on refugees, many of the structural phenomena that determine refugees’ education opportunities—such as the power balances of the world order that contribute to armed conflicts, inequity within and between nations, policies and value systems, aspirations of young people, as well as xenophobic backlash or securitization of refugees—are all strongly influenced by globalization.

Who is a refugee learner? Numbers, global distribution and the issue of definitions

There is no simple definition of who shall be considered a refugee (Sandelind, 2020; Cole, 2021), and many learners with a refugee experience or background will not be identified as such in statistics or administrative systems. The term “refugee” is used by researchers and authorities to cover a wide range of circumstances beyond the definition of the Geneva Convention (United Nations, 1951), and a multitude of related designations are used to indicate specific categories and situations. The UN Refugee Agency (UNHCR) uses the term “forcibly displaced people”, to cover the various situations where people have fled their homes “as a result of persecution, conflict, violence, human rights violations and events seriously disturbing public order”, estimating their number to be 82.4 million in 2021 (UNHCR, 2021a). Of these, the majority are internally displaced people (48 million, representing 58%), while only a small proportion are asylum seekers (4.1 million, representing 5%). Roughly 15% of forcibly displaced people are children of school age. In several parts of Africa and Asia, the proportion of children can exceed 50%. The vast majority of forcibly displaced are hosted in low- and middle-income countries (LMIC), mostly neighboring the country they fled from (NRC, 2021). By comparison, relatively small number of refugees are able to seek asylum in high income countries (HIC). For instance, in May 2021, 8549 asylum seekers were granted refugee status across European countries, while an estimated 792,800 cases were pending (EASO, 2021). Obstacles in access to education for refugees can be compounded by statelessness (Mahruf and Shohel, 2022). Situations of flight can lead to statelessness, but estimating numbers is even more difficult than for other refugees. Thus, while the UNHCR in 2021 estimated that 4.2 million people were stateless worldwide, based on numbers reported by 79 countries), in 2018 Filippo Grandi of the UNHCR suggested that the true figure might be three times greater (UN, 2018).

All statistics concerning refugees must be seen as rough approximations, and estimating numbers is notoriously difficult (Blitz et al., 2019). This is in part because UNHCR statistics aim to assess needs. Refugees who have been granted settled status in high income countries are therefore removed from statistics after 10 years, while refugees in developing countries are likely to remain in need of protection and humanitarian aid and are therefore mostly included as long as the situation of forced displacement continues (NRC, 2021); in some cases, this extends for generations (Dryden-Peterson, 2017). Refugees granted residence on humanitarian grounds and resettlement refugees are also excluded from UNHCR statistics (NRC, 2021). Many refugees are unable or unwilling to register. Thus, refugees living informally in other countries, those who have traveled to countries with no visa requirements, as well as those who have been able to migrate through family reunification, or with work or study permits, are not included in the numbers (Blitz et al., 2019; Cole, 2021; NRC, 2021).

Refugee learners and the right to education

Refugees are protected by international law, through the Geneva Convention of July 28, 1951 (United Nations, 1951), the Protocol of 1967, 1969 OAU Convention, and the 1984 Cartagena Declaration (Freier et al., 2022), as well as other instruments which are based on these, such as the African Charter on the Rights and Welfare of the Child, or the EU Charter of Fundamental Rights. Importantly, the principle of “non-refoulement” ensures that people who have fled cannot be sent back as long as threats persist, regardless of the status they are granted, and regardless of whether the host country is party to the Geneva Convention. Article 22 of the Convention stipulates that refugees should enjoy the same right to elementary education as nationals, while for education other than elementary, they should at least have the same access as non-nationals, including the recognition of prior qualifications. Children’s right to education and protection are established in numerous other international documents, such as the Universal Declaration of Human Rights, the United Nations Convention on the Rights of the Child, and Sustainable Development Goal 4 of Agenda 2030 (UNESCO, 2018b). The Geneva Convention does not include mechanisms of compliance, however (Willemms and Vernimmen, 2018), and in practice a limited proportion of refugees enjoy the rights stipulated in the Convention, including the right to education (see also Freier et al., 2022). Less than two-thirds of refugees have access to primary education, one-fourth get a secondary education, and approximately 3% enroll in tertiary education (UNESCO, 2018a). Statistics often focus on children and young people, but also adults are in need of retraining and further education, since former qualifications are seldom recognized and new skills are needed in the new location.

International instruments which offer protection and rights to refugee learners, such as the Geneva Convention, resulted from specific historical developments in the post WWII period, including the Cold War dynamics and processes of decolonization that significantly affected the world order reflected in UN bodies (Chimni, 1998), as well as ensuring their funding. Since then, the world order is in a process of renegotiation, and the factors that initially motivated wide support and adherence to UN principles have weakened. UN agencies are experiencing a financial crisis (Mir, 2020), while their autonomy, credibility and scope of action is reduced (Thakur, 2020). At the same time, humanitarian crises have been accelerating worldwide, and this trend is likely to increase in the coming years. UN yearly efforts to secure national commitments and to mobilize humanitarian funding from donor countries are thus set to fail.

It is against this background that the UNHCR shifted strategy in 2012 from supporting teaching of the curriculum of the country of origin, to instead aiming at integration in the educational systems of the host country (Dryden-Peterson, 2017; Bellino and Dryden-Peterson, 2019). Similarly, the Global Compact on Refugees corresponds to a need to find viable long-term funding for

refugee education, and entails a renegotiation of responsibilities aiming not to place a disproportionate burden on the countries that host refugees—mainly neighboring countries situated in the global South (Sharpe, 2018; NRC, 2021). There are, however, a number of serious issues with this strategy. Chimni (2018) underlines that the Global Compact lacks mechanisms of enforcement and will therefore, like the Geneva Convention, rely on the goodwill of states. Furthermore, it does not address the pushback policies (see e.g., Parliamentary Assembly, 2019; Auer, 2021) and other obstacles put in place to prevent refugee protection by high-income states (Morris, 2021).

Particular challenges in refugee education

It is seldom possible for local authorities in host countries or international agencies organizing education in refugee camps to assess how long a conflict will last, or what the probable outcome will be. Such uncertainties make it difficult to take appropriate decisions concerning whether to set up temporary arrangements based on the language used for teaching and learning and curricula of the country of origin, or to instead aim at rapid integration in the education system of the host country (Dryden-Peterson, 2017, 2021). Further obstacles are, on the one hand, that curricula in both host country and country of origin are unlikely to be oriented toward the challenges that post-conflict countries typically face for recovery and peacebuilding. This, in turn, reduces the likelihood that circumstances in the country of origin will eventually permit the safe return of refugees. On the other hand, since refugees and returnees face an increased risk of renewed flight or onwards migration, their education should equip them both to continue their education and eventually work in specific national contexts, and to be able to use qualifications and knowledge in other unspecified contexts about which they cannot be sure in advance.

Although legal status and access to protection differ depending on context and individual factors, from the perspective of refugee education, all learners who have been obliged to flee their country of prior residence will share many of the same needs. Their past experience and future expectations will be marked by the trauma of displacement (cf. Sandelind, 2020) and loss of social, economic and cultural capital—a loss that can be more difficult to deal with for learners with high aspirations, who had extensive opportunities before their flight. Refugee learners will face administrative and pedagogical obstacles, and often be exposed to discrimination, social exclusion, financial hardship and a multitude of uncertainties (UNESCO, 2018a). School fees can be a challenge (Aquila, 2022). While refugee learners are likely to face challenges due to various aspects of their background and situation, there is also a likelihood that education they are offered will not be well matched to their needs and aspirations (Molla, 2021).

Furthermore, experiences that characterize refugee learners do not only apply to the first generation. Impacts of forced displacement persist over generations (see e.g., Sajdi et al., 2021), including historical and intergenerational trauma. As with other learners with a migrant background, refugee learners are likely to have friends, relatives and other members of their communities scattered over many countries, leading to a transnational identity and understanding of the world (see Bloch and Hirsch, 2018; Rizvi, 2019). Furthermore, as other learners from disadvantaged backgrounds, refugee learners need a greater degree of self-efficacy and community support to manage negative images they meet in education as well as in society at large. As many learners from disadvantaged backgrounds also, they will face expectations to provide assistance to relatives and members of their community, which increases the pressure to succeed and obtain qualifications that allow social or geographical mobility. Others drop out early, to work and contribute to livelihoods (Erden, 2019), while particularly girls may marry young (El Arab and Sagbakken, 2019; Dagane and Aden, 2021). This can create tensions with educational systems that see learners as individuals detached from their social context.

Geographical distribution and focus of research on refugee education

A search made on August 12, 2021 of peer-reviewed articles in English published in Scopus-listed journals using the search term “refugee education” in title-abstract-keywords yielded 1641 publications in the subject area social sciences. Of these, 21% were published between the years 1958 and 2009, and 79% were published after 2010. The US, Australia, UK and Canada represented 62% of the total publications for the period 1958–2021, while Sweden, Turkey, Germany, the Netherlands, Israel and Norway represented a further 16 %. The US alone produced 544 publications (33% of total publications). According to the UNHCR (2021a), the country hosted a total of 1 million asylum-seekers and 300,000 refugees in 2020, corresponding to 0.4% of its population. Among the other major host countries, Sweden produced 66 publications (4%), hosting 267,000 refugees in 2020 (2.6% of its population), Turkey had 62 publications (3.8%, with 3,976,000 refugees (4.7% of population), and Germany 57 publications (3.5%), with 1,454,000 registered refugees (1.7% of population) (NRC, 2021).

Although several host countries, and above all the US, thus contribute substantially to research on refugee education, other host countries are noticeably absent in the academic literature. Excluding Turkey, Germany and the US, the ten main countries that refugees had fled to were Jordan, Palestine, Colombia, Pakistan, Uganda, Lebanon, Peru, Sudan, Bangladesh and Ethiopia, with a total of 15.1 million refugees (NRC, 2021, including UNRWA figures), but together only representing 5.1% of Scopus-listed publications. The largest number of these (28 publications) came from Lebanon, where refugees make up 19.5% of the population. By comparison, Turkey, Germany and the US jointly host 6.8 million refugees, but represent 39.5% of publications, with the vast majority coming from the US. Literature from countries of origin is almost absent, with a total of 13 publications (0.8%) from the ten

main countries of origin. Where funding sponsor was indicated, a total 121 publications were funded by the EU, various national governments, foundations, academic institutions or humanitarian organizations from high-income countries; 17 were sponsored by UN agencies, and 18 were sponsored by academic institutions in LMICs.

The 1641 publications found in Scopus covered a very wide range of topics. However, only less than one-third (511 publications) had keywords that specifically referred to education (e.g., teaching, learning, students, education policy). The most frequent keywords among the 511 articles focusing education were: higher education (68), migration (45), education policy (34), United States (30), female (26), ethnology (25), language (25), economics (24), employment (24), and young population (24).

Many of the types of challenges faced by refugee learners are similar to other categories of learners, and strategies to address these challenges can therefore be found in literature that does not explicitly focus refugees, such as regarding international (IE) and transnational education (TNE), education for newcomers, migrants and learners with an immigrant background, or literature on social justice, anti-racism and inclusion more generally. Other strands of literature relevant to the large proportion of refugee learners hosted in the global South can be found under designations such as “emergency” education, “education for development”, and education in fragile, humanitarian, or low- and middle-income country (LMIC) settings.

Although numerous publications relevant to refugee education can certainly be found using other search terms or in other languages, in journals not listed by Scopus and in the gray literature, the Scopus overview nevertheless suggests certain trends. As in other fields of research, English speaking countries dominate international publications in English. From around 2000, the number of publications appears to increase exponentially, and a certain correlation can be discerned in the geographical distribution between publications and number of refugees hosted. However, there is a sharp asymmetry between publications from HICs and LMICs, and US publications are disproportionately represented. The dominance of US research in the field is accentuated when looking at OA publications, which are likely to have greater impact due to ease of access. Here, of 121 OA publications, 102 (84%) were from the US.

Pedagogical and psychosocial challenges in international education

Refugee education shares many characteristics with what can be broadly termed “international” or “transnational” education (TNE), including internationalization efforts and measures to encourage mobility in education, such as the European Erasmus+ program. Thus, from a purely pedagogical point of view, we are looking at an educational situation where language of teaching and learning and official curriculum contents differ from the curriculum and prior language of the learner (cf. [Waters and Leung, 2017](#); [Dryden-Peterson, 2021](#)). The learner will have to adapt to a different societal and educational culture and may experience social isolation due to the distance from their previous social networks. The needs of refugee learners and other types of learners in international education are also similar, to the extent that there is an assumption or anticipation of transnational life and career trajectories, characterized by mobility.

In the case of transnational and international education, these issues are well known, and are addressed in constructive ways. Teachers and educational institutions are aware of language issues and may have special teacher training to better provide for their international students. The educational institutions may have special support staff or offices, designed to assist students with academic writing. Differences in students’ prior knowledge and educational culture are understood to be normal and a consequence of different educational systems and curricula. Such challenges in transitioning between educational systems are therefore not interpreted as an indication of lower academic capacity or value of the students. Rather, differences in background are one of the arguments for internationalization, since international students enrich the learning environment with their prior experiences, culture, perspectives, and knowledge. Support services for international students ([Arthur, 2017](#); [Ammigan and Jones, 2018](#); [Madden-Dent et al., 2019](#)) may involve mentoring, or specially arranged social and cultural events, to facilitate friendships and interaction, so that international students can fully benefit from their stay, while students at the host institution can learn from their presence ([Arkoudis et al., 2013](#)). Educational counselors are familiar with the mental health issues that international students may have due to isolation and navigating in an unfamiliar environment, and host institutions may also provide support regarding practical challenges, such as accommodation or transport.

Challenges in refugee education compared to other international and transnational education

The most fundamental difference between refugee education and other forms of international education is that, in the latter, the international students are seen as desirable ([Lo, 2019](#); [Serpa et al., 2020](#)), and the different experiences that they bring are seen as an asset to other students, as well as to the standing and reputation of the host institution ([Delgado-Márquez et al., 2013](#); [de Wit, 2019](#)). From a policy perspective, internationalization and transnational mobility are considered means to support the economy, by preparing the future workforce to be better adapted to mobile working conditions, cooperate in multicultural teams, work in international corporations or organizations, support trade relationships, tourism, diplomatic or geopolitical interests, as well as to strengthen future collaborations in research and development (see e.g., [Rose and McKinley, 2018](#); [Waters and Leung, 2017](#); [Wihlborg and Robson, 2018](#)). As well as serving economic interests, educational mobility may be seen as strengthening regional identity. For example, within the European Union mobility is seen as a way of fostering a “European identity”, to ensure

the stability, cohesion and future development of the union (Van Mol, 2018). Similar motivations lie behind the expansion of the Bologna process, or the educational collaboration among countries of the British Commonwealth, as well as in other regions that wish to consolidate their internal cohesion or gain new allies (Ziguras, 2018; Ploner and Nada, 2019; Khan et al., 2020).

In the case of international schools and programmes that receive paying students (see Kim and Mobrand, 2019), as well as in the case of higher education institutions receiving international students, higher fees demanded from international students are often vital to the finances of the educational institution (de Wit, 2019). Such educational institutions are therefore highly motivated to attract the international students, and to make their educational experience positive and academically successful. At the same time, with paying students, financial resources are also available for the various support measures.

By contrast, in refugee education, students can rarely afford high tuition fees. To the extent that host countries do not see refugee populations as an asset, public funding will be limited. Given the high percentage of refugees located in so-called developing countries, which may themselves have difficulty in supporting education systems for their existing populations, much education provision will devolve to international humanitarian organizations and NGOs. However, these have insufficient funding to address growing challenges worldwide and can therefore only offer limited support. In a great number of educational settings, financial resources available to fund necessary support measures are therefore lacking.

Not only do host countries rarely provide adequate funding for support measures, but in many cases there are hostile policies aimed at controlling refugee populations, deterring them from arrival or facilitating rapid deportation. These may lead to legislation and regulations that multiply formal and administrative obstacles to education; for example asylum seekers and refugees may be required to reside in particular locations or they may be moved from one location to the other at short notice. Grüttner et al. (2021) speak of an “administrative jungle” confronting refugee students. Educational institutions may also not be allowed to show flexibility concerning ordinary requirements of documentation. Translation of prior qualifications, documentation concerning identity, parental custody, health certificates, proof of residence and status, may additionally entail considerable costs and administrative delays. Original records may have been lost (Gilliland, 2017; UNESCO, 2018a) or destroyed due to armed conflict, and contacts with authorities representing the country of origin may expose the refugee or asylum seeker to danger. Even when documents are available, prior qualifications tend to be devalued (see Eggenhofer-Rehart et al., 2018). The administrative formalities connected to enrollments and assessment of prior qualifications can thus become major obstacles to educational access.

Although refugee students are less likely to pursue higher education (UNESCO, 2018a) and also more likely to drop out from pre-study programmes, in their study of a German context, Grüttner et al. (2021) found that this difference no longer appeared when gender, age, social origins and other relevant factors were accounted for. Financial problems were identified as major obstacles, while feelings of social exclusion also increased the risk of drop out. Nevertheless, refugee students were more likely to be resilient in the face of adversity than international students. Experts consulted in Grüttner, Schröder and Berg’s study therefore believed that refugee students would benefit from similar support as other international students.

Language policies in refugee education and in the global educational landscape

The curriculum and language used for teaching and learning in refugee education can be oriented toward the education system of the country of origin, in view of an assumed later return or be oriented toward integration in the national education system of the host country (Bellino and Dryden-Peterson, 2019; Koehler and Schneider, 2019; Reddick and Dryden-Peterson, 2021). Countries that wish to keep the refugees in the long term, will generally encourage learning the host country language, and may provide special support to this effect.

The choice of language of instruction is crucial for refugee learners’ future educational pathways and opportunities but learning any language to a level that allows future studies and careers entails an investment of many years. From the perspective of refugee learners, such choices are particularly difficult in the general context of uncertainty that characterizes refugee education. Inevitably, any choice of language will open certain future pathways, while creating obstacles for others. Some countries and educational institutions have a policy of supporting plurilingualism and high-quality bilingual education—Canada and Bolivia, or the European school of Luxembourg are, for instance, notable examples. However, this is not a systematic practice in settings of refugee education and objectives of national plurilingualism policies are likely to diverge from those relevant to refugee learners. Often also, refugee learners will be reluctant to invest time in developing academic competence in smaller non-dominant languages that are only used locally, and will instead prefer dominant “international” languages such as English, particularly if they hope to move onwards (Karam et al., 2017; Zsófia, 2018; Yilmaz, 2022). For refugee education thus, learning the language(s) of the host country, maintaining language(s) of country of origin or prior education, and acquiring international languages in view of future mobility, all play significant roles in preventing loss of skills and cultural capital, as well as enabling choice of future trajectories. As in the case of other international and transnational education, for elite learners acquiring additional languages is seen by the states or organizations that fund education as a valid investment and advantage, while for refugee learners the issue is framed as a problem and a “burden” that institutions are reluctant to take responsibility for (cf. Hult and Hornberger, 2016).

Refugee learners can be forced into renewed displacement, or may themselves wish to move onwards in search of better opportunities or to live close to family settled elsewhere. However, the landscape of languages that enable mobility is structured around powerful interests, due to the role that language in education plays in diplomacy and geopolitical strategies (Khan et al., 2020). Thus, language, content, curricular structure and education pathways, recognition of qualifications and transnational links between

educational institutions can serve to strengthen the economic and geopolitical position of a country, region or group of allied countries. Internationalization and educational collaboration are therefore often coupled to a strategy of promoting certain languages, thereby also consolidating the role of these languages in research and trade (Diaz, 2018; Erfurt, 2018; De Costa et al., 2021). Student mobility is encouraged in this type of internationalization strategies, since it is anticipated that some international students will remain in the country they received their qualifications in, and thereby contribute to the pool of highly qualified professionals. Others will return to their countries of origin or pursue careers internationally, but still retain their ties and networks to the country where they received their education, as well as an understanding of the perspectives and interests of that country. They will be able to communicate in the language of education and can be expected to show sympathies for that country's interests in future positions of authority and power that they may occupy. For a small group of refugee learners who receive scholarships, similar considerations may play a role, and in some cases global powers involved in the conflicts that led to the flight may wish to support certain groups of refugee learners. Languages used in refugee education in LMICs that is funded by humanitarian organizations is also not neutral to the geopolitical dynamics of donor countries.

From a pedagogical perspective, among approaches of particular interest for refugee education are those involving creativity and artistic expression (McLeod et al., 2020) since these are less dependent on language proficiency and allow refugee learners to be active and visible in the community, express experiences of resilience and process adversity.

Curricular content and pedagogical orientation in refugee education

Since prior knowledge is difficult to transfer (Hajian, 2019; Dryden-Peterson, 2021) self-directed and inquiry-based approaches (Almasri et al., 2019) are suggested to enable learners to independently reorganize and supplement their knowledge or skills, as well as defining their own learning agendas. However, learners with trauma can find it difficult to navigate learning environments that lack clear structure and where instructions are not sufficiently explicit. Similar difficulties may be encountered by refugee learners moving from teacher-led environments to more fluid learner-centered contexts, such as situations experienced during emergency distance learning during Covid (Fujii et al., 2020). Nevertheless, once they have cracked the code of school culture and teacher expectations, refugee learners often have considerable resourcefulness and skills in solving complex tasks, as well as tenacity in dealing with adversity (Sleijpen et al., 2017; Ryu and Tuvilla, 2018).

In refugee camps and settings of temporary residence, educational content may consist of ad hoc assemblages of whatever volunteers, NGOs or donor countries are able and willing to offer (Karim and Hussain, 2019), including curricula modeled on the donor country or informed by the ideological or religious identity of the NGO. Language and curriculum differences in education systems, combined with lack of recognition of prior qualifications mean that moving between systems entails losing years of study, as well as other resources invested in education. To the extent that refugee trajectories have a likelihood to involve onward migration or flight, and from the perspective of enabling refugees to lead meaningful lives, there is therefore a strong argument for offering high status curricula such as the international baccalaureate program, that are recognized and can be pursued internationally (cf. Streitwieser, 2019).

Refugee education is insufficiently funded, however, leaving little resources for quality education (UNHCR, 2011; Mir, 2020), while even high-quality education offers limited protection against discrimination. Also, like other measures to enhance mobility, channeling refugee learners toward elite, transnationally-recognized programmes could aggravate brain drain (UNESCO, 2018a), standardization, or loss of locally relevant curricular content (Stromquist, 2002; Stromquist and Monkman, 2014). As an alternative, Dryden-Peterson et al. suggest including content and using teaching approaches in refugee education that build transferable skills, which will be useful regardless of how the future unfolds (Dryden-Peterson et al., 2019; Dryden-Peterson, 2021). Such approaches are particularly valuable in the context of short-term non-formal programmes directed toward refugees, but to the extent that the strategy only partly addresses issues of recognition of qualifications, it does little to open pathways out of the parallel universe of constrained education and career opportunities that refugees are frequently assigned to.

Drawing on evidence concerning long term outcomes of European education systems for learners with immigrant background, Koehler and Schneider (2019) conclude that early-tracking aggravates educational disadvantage, compared to education systems aiming at "education for all" and with life-long learning perspectives. Policy that does not permanently penalize learners for disadvantage at a particular age is therefore important. Pedagogical approaches and institutions developed in the open education movement (Weller, 2020) are in line with such aims. Distance learning addresses the mobility constraints refugee learners are subjected to, while open educational resources (OER) can be a valuable asset in resource-constrained settings. However, OER do not necessarily correspond to the curricular content, languages or educational needs of different refugee education settings, while digitally mediated instruction comes with a number of other challenges (see e.g., Joynt and James, 2018; Bock et al., 2020).

Brain drain, repatriation of refugees and post-conflict recovery

In the context of refugee education, the issue of brain drain is frequently a focus, since it affects a relatively small number of countries at a massive scale, where the challenges of post-conflict recovery for these countries are aggravated by the interrupted or limited education of populations who remained in the country of origin (see e.g., Avery and Said, 2017; UNESCO, 2018a). Mechanisms of

exclusion directed toward refugees in host countries also mean that the latter only benefit to a limited extent from the influx of educated professionals and youth in education. Although less visible, authoritarian regimes that drive brain drain through flows of refugees suffering individual persecution, also drive a steady exodus of youth and educated professionals who seek better and safer life conditions elsewhere. However, despite the real impacts of brain drain, it is simplistic to frame refugee repatriation as a solution, since this does not address the underlying drivers to flight (Walter, 2011; Bara et al., 2021).

Even in the best of cases, conditions to benefit from an inflow of qualified returnees are unlikely to be favorable in many post-conflict contexts (see e.g., Williams, 2020). In cases where diasporas do not, or only intermittently return, grievances may push them to oppose new governments, while in other cases they play active roles in recovery from afar (Cochrane, 2015). Diaspora remittances can be important, as well as social, cultural and economic initiatives based on the international networks and competences diasporas have acquired (Amazon, 2014; Dryden-Peterson and Reddick, 2019; Williams, 2020; Halilovich and Efendi, 2021; San-dož et al., 2021). For refugee education, a significant issue affecting the ability of diasporas and returnees to contribute to post-conflict recovery are not only the institutional, security and economic conditions of the country of origin, but to which extent their skills and qualifications match needs in the country of origin, and whether knowledge and experience are transferable (Dryden-Peterson, 2021). Curricular content, medium of instruction and qualification structures of national education systems are not geared to meet such needs, while transnational education tends to be oriented toward producing mobile elites at the service of large corporations or national diplomacy and civil service.

The objective of repatriation to countries of origin outlined in the Global Compact is therefore problematic (Chimni, 2018; Tegenbos and Vlassenroot, 2018; Zetter, 2021). While many refugees do hope to return and ensuring the right to return is an important issue, conditions for return are often highly precarious and refugee return is frequently implemented as a deportation (Lemberg-Pedersen, 2021; Naimou, 2021; Lindberg and Edward, 2021; Garvik and Valenta, 2021). Peace agreements seldom settle underlying root causes, and the risk for future conflicts and continued violence is high (Walter, 2011; Bara et al., 2021). In yet other cases, such as Bosnia and Lebanon, the power sharing arrangements of the peace agreements lead to long-term political stalemates that block development and produce a state of protracted crisis. Also, groups having previously experienced persecution remain exposed, as illustrated by recent examples for Rohingya, Syrians, and Afghans. In the best of cases, for refugee learners, repatriation will still represent a disruption of their studies, with efforts to adjust to a place they may never have known, and rebuilding a new life from scratch, rather than resuming “life as it was”. For many, shifting to a new medium of instruction that they may not be proficient in will also represent a significant educational challenge (Reddick and Dryden-Peterson, 2021), since their academic skills will be encoded in another language. Second and third generation refugee returnees may not even have basic proficiency in the languages of the country of origin of their families.

Conclusions and further research

Globalization involves a regime that enables mobility of capital, profits and certain goods and services, coupled with free movement of elites (Bauman, 1998; Solimano, 2020), but reduced freedom of mobility for other groups (Shamir, 2005; Liu-Farrer et al., 2021). This regime makes it possible for strong global actors to externalize negative environmental, social and economic impacts, aggravating global disparities and concentrating wealth. Globalization dynamics thereby contribute to contradictory discourses on mobility and migration, belonging and entitlements, that determine fundamental conditions for refugee education. Travel and mobility can be profoundly enriching, when they are freely chosen. By contrast, refugee lives and the educational opportunities of refugee learners are characterized by forced mobility, and often also by forced immobility, containment or detention (UNESCO, 2018a; Yilmaz, 2022). The largest numbers, and those that are generally focused as a matter of international concern, have been forced to flee as a consequence of armed conflict, while ethnic, political or religious persecution also play significant roles.

The educational trajectories of refugees are transnational, as a consequence both of past displacement and future forced or voluntary onwards migration and uncertainties (Dryden-Peterson et al., 2019; Dryden-Peterson, 2021). Their friends, relatives and sense of belonging are distributed across multiple countries, as are the educators, policymakers and authorities that fund, plan, record, validate and implement their education. International agencies, international law (Willems and Vernimmen, 2018) and global mobility regimes (Schapendonk et al., 2020) play a major part in these processes. Refugee education can thus be understood as a phenomenon embedded in a globalized educational landscape, that is strongly impacted by geopolitical interests and global economic forces. At the same time, the circumstances under which refugees lead their lives place them at odds with national educational systems organized to serve citizens, and which tend to presuppose an uninterrupted educational trajectory within that country. Nevertheless, it is important to bear in mind that although refugee learners share certain experiences and challenges with respect to the structures of educational systems globally, they can by no means be viewed as a homogenous group. The trajectory, circumstances and aspirations of each learner will always be unique. Also, refugees learners’ situation varies significantly—great differences exist in mobility and opportunities between refugees who have obtained citizenship in a HIC, and those who are undocumented, unrecognized, or living in camps in LMICs.

Refugee education is constrained by global policy, citizenship and mobility regimes that combine to place refugee learners in a position of disadvantage, but also make refugee populations into a “burden” for host countries and the international community. Among the main issues that currently directly impact refugee education globally, renegotiations of responsibility for funding education and livelihoods of refugee learners included in the Global Compact can be noted, coupled with the risk of forced repatriations

(Chimni, 2018; Tegenbos and Vlassenroot, 2018; Zetter, 2021). Another important issue is whether education is respectively seen as a process of integration into the national education systems of host countries in view of permanent settlement, or as temporary arrangement in view of future return (Dryden-Peterson, 2017).

A number of factors combine to place refugees in a position of disadvantage with respect to the intersection of both national and transnational regimes. From this perspective, education systems can be understood as organized globally in a three-tier hierarchy. At the top, are foreign degrees and transnational education, designed to optimize mobility and status for global elites (see e.g., Gardner-McTaggart, 2018; Tuxen, 2019), and for selected groups of professionals that serve their interests. In the middle, national education systems are designed to enable economic activity within countries. National systems can also be stratified internally, while there is a stratification between countries, corresponding to the current world order. At the bottom of the hierarchy, refugee education at best provides limited access to education within national systems, at the same time that it does not address the extended needs of refugee learners for transnational mobility and access to transnational resources. Consequences of reduced freedom of mobility for refugees and refugee learners are particularly dire, since their mobility needs are not only motivated by an ambition to improve their situation, but by survival (Streitwieser, 2019). Also, refugee learners benefit from limited protection by states (Sandelind, 2020), and therefore need to mobilize greater resources of their own, through networks, communities and education (see Dryden-Peterson et al., 2017), in order to compensate for the position of systemic disadvantage in which they are placed. There is thus a fundamental mismatch between the needs of refugee learners, on the one hand, and the education provisions that are offered, on the other (see also Molla, 2021).

Considering limited funding to tailor programmes designed specifically to support refugee learners, global citizenship education (UNESCO, 2014) and global education (Gaudelli, 2007; Nordén and Avery, 2021) have the advantage that they are aimed at enabling societies to collectively address major global challenges. Importantly, these approaches are not specifically aimed at refugee learners, but correspond to general commitments by states to contribute to sustainable development. From the perspective of refugee education, they offer a framework where transnational identities and experiences are highly valued, within a narrative of collective efforts and solidarity, rather than the narrative of transnational elites.

According to the UNHCR (2021b), in 2020 more than two-thirds (68%) of refugees worldwide came from just five countries of origin (Syria, Venezuela, Afghanistan, South Sudan, Myanmar), while 39% of all refugees were hosted in the top five countries (Turkey, Colombia, Pakistan, Uganda, Germany). The vast majority of refugees (86%) were 2020 hosted in developing countries (27% in the Least Developed Countries). The dominance of a relatively small number of HICs in the research, coupled with a low proportion of research from LMIC host countries, thus constitutes a structural bias in the way our understanding of issues in refugee education is shaped, as well as with respect to the perspectives and agendas that are represented in proposed solutions. Other types of bias in the research come from difficulties in access to refugees, finding refugees who are hiding, or not formally identified as such (see Cole, 2021), as well as from the fact that very little research is driven or carried out by refugees themselves. Among topics that would require additional research in the context of refugee education are, for instance, the specific challenges of IDPs, who make up the majority of the forcibly displaced, education issues concerning refugee returnees, and the educational situation of refugees who are not recognized or identified. Disability in refugee education is another neglected area. Just as for other research on, for or with any vulnerable individuals, research on refugee education entails considerable ethical challenges (Clark-Kazak, 2017), and with increasingly stringent requirements for ethical review, future research will face challenges.

References

- Abdelaaty, L.E., 2021. *Discrimination and Delegation: Explaining State Responses to Refugees*. Oxford University Press.
- Alio, M., Alrihawi, S., Milner, J., Noor, A., Wazefadost, N., Zigashane, P., 2020. By refugees, for refugees: refugee leadership during Covid-19, and beyond. *Forced Migr. Rev.* 64, 76–78.
- Almasri, N., Tahat, L., Skaf, S., Masri, A.A., 2019. A digital platform for supervised self-directed learning in emergencies: the case of the Syrian crisis. *Technol. Pedagog. Educ.* 28 (1), 91–113. <https://doi.org/10.1080/1475939X.2019.1568293>.
- Amazan, R.C., 2014. When the Diaspora returns: analysis of Ethiopian returnees and the need for highly skilled labour in Ethiopia. In: Streitwieser, B. (Ed.), *Internationalisation of Higher Education and Global Mobility*. Symposium Books Ltd, Oxford, pp. 169–186.
- Ammigan, R., Jones, E., 2018. Improving the student experience: learning from a comparative study of international student satisfaction. *J. Stud. Int. Educ.* 22 (4), 283–301. <https://doi.org/10.1177/1028315318773137>.
- Ansorg, N., 2020. Securitisation strategies to prevent conflict diffusion in Tanzania and former Zaire. *J. Contemp. Afr. Stud.* 38 (4), 579–593. <https://doi.org/10.1080/02589001.2020.1811211>.
- Aquila, L., 2022. *Child Labour, Education and Commodification in South Sudan*. Rift Valley Institute, South Sudan.
- Arkoudis, S., Watty, K., Baik, C., Yu, X., Borland, H., Chang, S., Lang, I., Lang, J., Pearce, A., 2013. Finding common ground: enhancing interaction between domestic and international students in higher education. *Teach. High. Educ.* 18 (3), 222–235. <https://doi.org/10.1080/13562517.2012.719156>.
- Arthur, N., 2017. Supporting international students through strengthening their social resources. *Stud. High. Educ.* 42 (5), 887–894. <https://doi.org/10.1080/03075079.2017.1293876>.
- Auer, M., 2021. Witness: "If You Scream, They Will Beat You More". Migrants Face Human Rights Abuses at EU Borders; Impunity Persists. Human Rights Watch. Witness: "If You Scream, They Will Beat You More" | Human Rights Watch. [hrw.org](https://www.hrw.org).
- Avery, H., Said, S., 2017. Higher education for refugees: the case of Syria. *Policy Pract.* (24), 104–125.
- Baker, S., Due, C., Rose, M., 2021. Transitions from education to employment for culturally and linguistically diverse migrants and refugees in settlement contexts: what do we know? *Stud. Cont. Educ.* 43 (1), 1–15. <https://doi.org/10.1080/0158037X.2019.1683533>.
- Bara, C., Deglow, A., van Baalen, S., 2021. Civil war recurrence and postwar violence: toward an integrated research agenda. *Eur. J. Int. Relat.* 27 (3), 913–935. <https://doi.org/10.1177/13540661211006443>.
- Bauman, Z., 1998. *Globalization: The Human Consequences*. Columbia University Press, New York.

- Bellino, M.J., Dryden-Peterson, S., 2019. Inclusion and exclusion within a policy of national integration: refugee education in Kenya's Kakuma Refugee Camp. *Br. J. Sociol. Educ.* 40 (2), 222–238. <https://doi.org/10.1080/01425692.2018.1523707>.
- Blitz, B.K., D'Angelo, A., Kofman, E., 2019. Counting the population in need of international protection globally. In: Evans, J., Ruane, S., Southall, H. (Eds.), *Data in Society: Challenging Statistics in an Age of Globalisation*. Policy Press, Bristol, UK, pp. 91–102. <https://doi.org/10.2307/j.ctvmd84wn>.
- Bloch, A., Hirsch, S., 2018. Inter-generational transnationalism: the impact of refugee backgrounds on second generation. *Comp. Migration Stud.* 6 (30), 1–18. <https://doi.org/10.1186/s40878-018-0096-0>.
- Bock, J.G., Haque, Z., McMahon, K.A., 2020. Displaced and dismayed: how ICTs are helping refugees and migrants, and how we can do better. *Inf. Technol. Dev.* 26 (4), 670–691. <https://doi.org/10.1080/02681102.2020.1727827>.
- Chimni, B.S., 1998. The geopolitics of refugee studies: a view from the South. *J. Refug. Stud.* 11 (4), 350–374.
- Chimni, B.S., 2018. Global compact on refugees: one step forward, two steps back. *Int. J. Refug. Law* 30 (4), 630–634. <https://doi.org/10.1093/jr/ryy067>.
- Clark-Kazak, C., 2017. Ethical considerations: research with people in situations of forced migration. *Refuge* 33 (2), 11–17. <https://doi.org/10.7202/1043059ar>.
- Cochrane, F., 2015. *Migration and Security in the Global Age: Diaspora Communities and Conflict*. Routledge, London. <https://doi.org/10.4324/9781315881768FTA>.
- Cole, G., 2021. Pluralising geographies of refuge. *Prog. Hum. Geogr.* 45 (1), 88–110. <https://doi.org/10.1177/0309132519900925>.
- Dagane, A.S., Aden, A.Y., 2021. The Causes of Female Student Dropout in a Secondary School in the Dadaab Refugee Camps of Kenya. *Local Engagement Refugee Research Network Paper No. 12*. Carleton University, Canada.
- Darden, J.T., 2019. *Tackling Terrorists' Exploitation of Youth*. American Enterprise Institute, Washington, DC.
- De Costa, P.I., Green-Eneix, C., Li, W., 2021. Problematizing language policy and practice in EMI and transnational higher education: challenges and possibilities. *Aust. Rev. Appl. Ling.* 44 (2), 115–128. <https://doi.org/10.1075/aral.00036.edi>.
- de Wit, H., 2019. Internationalization in higher education, a critical review. *SFU Educ. Rev.* 12 (3), 9–17. <https://doi.org/10.21810/sfuer.v12i3.1036>.
- Delgado-Márquez, B.L., Escudero-Torres, M., Hurtado-Torres, N.E., 2013. Being highly internationalised strengthens your reputation: an empirical investigation of top higher education institutions. *High Educ.* 66 (5), 619–633. <https://doi.org/10.1007/s10734-013-9626-8>.
- Díaz, A., 2018. Challenging dominant epistemologies in higher education: the role of language in the geopolitics of knowledge (re) production. In: Liyanage, I. (Ed.), *Multilingual Education Yearbook 2018: Internationalization, Stakeholders & Multilingual Education Contexts*. Springer, Cham, pp. 21–36.
- Dryden-Peterson, S., 2017. Refugee education: education for an unknowable future. *Curric. Inq.* 47 (1), 14–24. <https://doi.org/10.1080/03626784.2016.1255935>.
- Dryden-Peterson, S., 2021. Toward cognitive and temporal mobility: language considerations in refugee education. *Mod. Lang. J.* 105, 590–595. <https://doi.org/10.1111/modl.12728>.
- Dryden-Peterson, S., Reddick, C., 2019. "What I believe can rescue that nation": diaspora working to transform education in fragility and conflict. *Comp. Educ. Rev.* 63 (2), 213–235.
- Dryden-Peterson, S., Dahya, N., Adelman, E., 2017. Pathways to educational success among refugees: connecting locally and globally situated resources. *Am. Educ. Res. J.* 54 (6), 1011–1047. <https://doi.org/10.3102/0002831217714321>.
- Dryden-Peterson, S., Adelman, E., Bellino, M.J., Chopra, V., 2019. The purposes of refugee education: policy and practice of including refugees in national education systems. *Sociol. Educ.* 92 (4), 346–366. <https://doi.org/10.1177/0038040719863054>.
- EASO, 2021. Latest Asylum Trends—May 2021. European Asylum Support Office. Latest Asylum Trends | European Asylum Support Office. europea.eu.
- Eggenhofer-Rehart, P.M., Latzke, M., Pernkopf, K., Zellhofer, D., Mayrhofer, W., Steyrer, J., 2018. Refugees' career capital welcome? Afghan and Syrian refugee job seekers in Austria. *J. Vocat. Behav.* 105, 31–45. <https://doi.org/10.1016/j.jvb.2018.01.004>.
- El Arab, R., Sagbakken, M., 2019. Child marriage of female Syrian refugees in Jordan and Lebanon: a literature review. *Glob. Health Action* 12 (1), 1–12. <https://doi.org/10.1080/16549716.2019.1585709>.
- Erden, O., 2019. Child Labour and School Attendance in Turkey. <https://www.fmreview.org/>. (Accessed 5 May 2022).
- Erfurt, J., 2018. Languages and spaces: La Francophonie and other "phonies". In: Middell, M. (Ed.), *The Routledge Handbook of Transregional Studies*. Routledge, Oxford and New York, pp. 159–168.
- Freier, L.F., Berganza, I., Blouin, C., 2022. The Cartagena refugee definition and Venezuelan displacement in Latin America 1. *Int. Migrat.* 60 (1), 18–36. <https://doi.org/10.1111/imig.12791>.
- Fujii, M.S., Hüttmann, J., Kutscher, N., Friedrichs-Liesenkötter, H., 2020. Participation?! Educational challenges for young refugees in times of the Covid-19 pandemic. *Media Educ.* 11 (2), 37–47. <https://doi.org/10.36253/me-9605>.
- Gardner-McTaggart, A., 2018. The promise of advantage: englishness in IB international schools. *Perspectives* 22 (4), 109–114. <https://doi.org/10.1080/13603108.2018.1456491>.
- Garvik, M., Valenta, M., 2021. Seeking asylum in Scandinavia: a comparative analysis of recent restrictive policy responses towards unaccompanied Afghan minors in Denmark, Sweden and Norway. *Comp. Migration Stud.* 9 (15), 1–22. <https://doi.org/10.1186/s40878-020-00221-1>.
- Gaudelli, W., 2007. Global education. In: Mathison, S., Ross, E.W. (Eds.), *Battleground: Schools*. Greenwood Publishing Group, Westport, pp. 280–287.
- Gereke, M., Nijhawan, S., 2019. "Start ins deutsche"—students teach German to refugees at Goethe University Frankfurt. In: Sengupta, E., Blessinger, P. (Eds.), *Language, Teaching, and Pedagogy for Refugee Education*. Emerald Publishing Limited, Bingley, pp. 91–105.
- Gilliland, A.J., 2017. A matter of life and death: a critical examination of the role of official records and archives in supporting the agency of the forcibly displaced. *J. Crit. Library Inf. Stud.* 1 (2), 1–24. <https://doi.org/10.24242/jclis.v1i2.36>.
- Grüttner, M., Schröder, S., Berg, J., 2021. University applicants from refugee backgrounds and the intention to drop out from pre-study programs: a mixed-methods study. *Soc. Incl.* 9 (3), 130–141. <https://doi.org/10.17645/si.v9i3.4126>.
- Gulmez, R., 2019. The securitization of the Syrian refugee crisis through political party discourses. *J. Int. Migrat. Integrat.* 20 (3), 887–906. <https://doi.org/10.1007/s12134-018-0637-2>.
- Guo, Y., Maitra, S., Guo, S., 2019. "I belong to nowhere": Syrian refugee children's perspectives on school integration. *J. Contemp. Issues Educ.* 14 (1), 89–105. <https://doi.org/10.20355/jcie29362>.
- Hajian, S., 2019. Transfer of learning and teaching: a review of transfer theories and effective instructional practices. *IAFOR J. Educ.* 7 (1), 93–111.
- Halilović, H., Efendić, N., 2021. From refugees to trans-local entrepreneurs: crossing the borders between formal institutions and informal practices in Bosnia and Herzegovina. *J. Refug. Stud.* 34 (1), 663–680. <https://doi.org/10.1093/jrs/ryy066>.
- Hult, F.M., Hornberger, N.H., 2016. Revisiting orientations in language planning: problem, right, and resource as an analytical heuristic. *Bilingual Rev.* 33 (3), 30–49. Retrieved from https://repository.upenn.edu/gse_pubs/476.
- Hyndman, J., 2013. A refugee camp conundrum: geopolitics, liberal democracy, and protracted refugee situations. *Refuge* 28 (2), 7–15. <https://doi.org/10.25071/1920-7336.36472>.
- IOM (n.d.). Key Migration Terms. IOM UN Migration. Key Migration Terms | International Organization for Migration (iom.int).
- Joynes, D., James, Z., 2018. An Overview of ICT for Education of Refugees and IDPs. K4D Helpdesk Report. UK Department for International Development.
- Karam, F.J., Kibler, A.K., Yoder, P.J., 2017. "Because even us, Arabs, now speak English": Syrian refugee teachers' investment in English as a foreign language. *Int. J. Intercult. Relat.* 60, 169–182. <https://doi.org/10.1016/j.ijintrel.2017.04.006>.
- Karim, A., Hussain, F., 2019. When will you start teaching the REAL curriculum? In: Nielsen, P., Kimaro, H.C. (Eds.), *Information and Communication Technologies for Development. Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D*. Springer Nature, Cham, pp. 675–686. https://doi.org/10.1007/978-3-030-18400-1_55.
- Khan, M.M., Ahmad, R., Fernald, L.W., 2020. Diplomacy and education: a systematic review of literature. *Global Social Sci. Rev.* 5 (2), 1–9. [https://doi.org/10.31703/gssr.2020\(V-III\).01](https://doi.org/10.31703/gssr.2020(V-III).01).

- Kim, H., Mobrand, E., 2019. Stealth marketisation: how international school policy is quietly challenging education systems in Asia. *Glob. Soc. Educ.* 17 (3), 310–323. <https://doi.org/10.1080/14767724.2019.1571405>.
- Koehler, C., Schneider, J., 2019. Young refugees in education: the particular challenges of school systems in Europe. *Comp. Migration Stud.* 7 (1), 1–20. <https://doi.org/10.1186/s40878-019-0129-3>.
- Lee, E.S., Szkudlarek, B., Nguyen, D.C., Nardon, L., 2020. Unveiling the canvas ceiling: a multidisciplinary literature review of refugee employment and workforce integration. *Int. J. Manag. Rev.* 22 (2), 193–216. <https://doi.org/10.1111/ijmr.12222>.
- Lemberg-Pedersen, M., 2021. The humanitarianization of child deportation politics. *J. Borderl. Stud.* 36 (2), 239–258. <https://doi.org/10.1080/08865655.2020.1835524>.
- Lindberg, A., Edward, S., 2021. Contested dreams, stolen futures: struggles over hope in the European deportation regime. In: Bhatia, M., Canning, V. (Eds.), *Stealing Time: Migration, Contested Temporalities and State Violence*. Palgrave Macmillan, Cham, pp. 83–104. https://doi.org/10.1007/978-3-030-69897-3_2.
- Liu-Farrer, G., Yeoh, B.S., Baas, M., 2021. Social construction of skill: an analytical approach toward the question of skill in cross-border labour mobilities. *J. Ethnic Migrat. Stud.* 47 (10), 2237–2251. <https://doi.org/10.1080/1369183X.2020.1731983>.
- Lo, W.Y.W., 2019. Beyond competition: a comparative review of conceptual approaches to international student mobility. *Glob. Soc. Educ.* 17 (3), 261–273. <https://doi.org/10.1080/14767724.2018.1525283>.
- Madden-Dent, T., Wood, D., Roskina, K., 2019. An inventory of international student services at 200 US universities and colleges: descriptive data of pre-departure and post-arrival supports. *J. Int. Stud.* 9 (4), 993–1008. <https://doi.org/10.32674/jis.v9i4.346>.
- Mahruf, M., Shohel, C., 2022. Education in emergencies: challenges of providing education for Rohingya children living in refugee camps in Bangladesh. *Educ. Inq.* 13 (1), 104–126. <https://doi.org/10.1080/20004508.2020.1823121>.
- Marsh, R.R., Oyelere, R.U., 2018. Global Migration of Talent: Drain, Gain, and Transnational Impacts. In: Dassin, J., Marsh, R., Mawer, M. (Eds.), *International Scholarships in Higher Education*. Palgrave Macmillan, Cham, pp. 209–234. https://doi.org/10.1007/978-3-319-62734-2_11.
- McLeod, H., Lewis, L.B., Li, X., 2020. Resilience and hope: exploring immigrant and refugee youth experiences through community-based arts practice. *Engaged Scholar Journal: Community-Engaged Research. Teach. Learn.* 6 (2), 88–104. <https://doi.org/10.15402/esj.v6i2.70765>.
- Micinski, N.R., 2018. Refugee policy as foreign policy: Iraqi and Afghan refugee resettlements to the United States. *Refug. Surv. Q.* 37 (3), 253–278. <https://doi.org/10.1093/rsq/hdy007>.
- Mir, W., 2020. Financing the United Nations Secretariat: Resolving the UN's Liquidity Crisis. Issue Brief. International Peace Institute, New York, Vienna and Manama.
- Molla, T., 2021. Refugee education: Homogenized policy provisions and overlooked factors of disadvantage. *Int. Stud. Sociol. Educ.* 1–22. <https://doi.org/10.1080/09620214.2021.1948892>.
- Morris, J., 2021. The value of refugees: UNHCR and the growth of the global refugee industry. *J. Refug. Stud.* 34 (3), 2676–2698. <https://doi.org/10.1093/jrs/feaa135>.
- Naimou, A., 2021. Refuge and deportation: the future as property in the border regime. In: Moraña, M. (Ed.), *Liquid Borders: Migration as Resistance*. Routledge, London, pp. 57–68. <https://doi.org/10.4324/9781003142911>.
- Nanopoulos, E., Guild, E., Weatherhead, K., 2018. Securitisation of borders and the UN's global compact on safe. In: *Orderly and Regular Migration*. Queen Mary School of Law Legal Studies Research Paper (270), pp. 1–16.
- Nordén, B., Avery, H., 2021. Global learning for sustainable development: a historical review. *Sustainability* 13 (6), 3451. <https://doi.org/10.3390/su13063451>.
- NRC, 2021. Global Displacement Figures 2021. Norwegian Refugee Council. Global figures | NRC.
- Parliamentary Assembly, 2019. Resolution 2299 (2019). Pushback Policies and Practice in Council of Europe Member States. Text Adopted by the Assembly on 28 June 2019 (27th Sitting). Council of Europe.
- Pincock, K., Betts, A., Easton-Calabria, E., 2021. The rhetoric and reality of localisation: refugee-led organisations in humanitarian governance. *J. Dev. Stud.* 57 (5), 719–734. <https://doi.org/10.1080/00220388.2020.1802010>.
- Ploner, J., Nada, C., 2019. International student migration and the postcolonial heritage of European higher education: perspectives from Portugal and the UK. *High Educ.* 80 (2), 373–389. <https://doi.org/10.1007/s10734-019-00485-2>.
- Reddick, C., Dryden-Peterson, S., 2021. Refugee education and medium of instruction: tensions in theory, policy, and practice. In: Benson, C., Kosonen, K. (Eds.), *Language Issues in Comparative Education II*. Brill Sense, Leiden and Boston, pp. 208–233.
- Rizvi, F., 2019. Global mobility, transnationalism and challenges for education. In: Drinkwater, M., Rizvi, F., Edge, K. (Eds.), *Transnational Perspectives on Democracy, Citizenship, Human Rights and Peace Education*. Bloomsbury Publishing, London, London, pp. 27–47.
- Rose, H., McKinley, J., 2018. Japan's English-medium instruction initiatives and the globalization of higher education. *High Educ.* 75 (1), 111–129. <https://doi.org/10.1007/s10734-017-01251-1>.
- Rossell Hayes, A., Dudek, C.M., 2020. How radical right-wing populism has shaped recent migration policy in Austria and Germany. *J. Immigr. Refug. Stud.* 18 (2), 133–150. <https://doi.org/10.1080/15562948.2019.1587130>.
- Ryu, M., Tuivilla, M.R.S., 2018. Resettled refugee youths' stories of migration, schooling, and future: challenging dominant narratives about refugees. *Urban Rev.* 50 (4), 539–558. <https://doi.org/10.1007/s11256-018-0455-z>.
- Sajdi, J., Essaid, A.A., Taleb, H.A., Azzam, M.A., 2021. "Why should I stay in the classroom?": drivers of school dropout among stateless Palestinian adolescents in Jordan. In: Jones, N., Pincock, K., Abu Hamad, B. (Eds.), *Adolescents in Humanitarian Crisis*. Routledge, Abingdon, Oxon and New York, NY, pp. 140–159. <https://doi.org/10.4324/9781003167013-7>.
- Sandelind, C., 2020. Refugees, displacement and territorial stability. *J. Global Ethics* 16 (2), 162–181. <https://doi.org/10.1080/17449626.2020.1784250>.
- Sandoz, L., Mittmasser, C., Riaño, Y., Pigué, E., 2021. A review of transnational migrant entrepreneurship: perspectives on unequal spatialities. *Z. für Wirtschaftsgeogr.* <https://doi.org/10.1515/zfw-2021-0004>.
- Schapendonk, J., van Lierp, I., Schwarz, I., Steel, G., 2020. Re-routing migration geographies: migrants, trajectories and mobility regimes. *Geoforum* 116, 211–216. <https://doi.org/10.1016/j.geoforum.2018.06.007>.
- Serpa, S., Caldeira, S.N., Serpa, M.S.D., Gonçalves, R.L., Montenegro, H.M., Rego, I.E., 2020. Mobility in the internationalisation of higher education institutions. *Int. J. High. Educ.* 9 (4), 46–60. <https://doi.org/10.5430/ijhe.v9n4p46>.
- Shamir, R., 2005. Without borders? Notes on globalization as a mobility regime. *Socio. Theor.* 23 (2), 197–217.
- Sharpe, M., 2018. The global compact on refugees and conflict prevention in Africa: "root causes" and yet another divide. *Int. J. Refug. Law* 30 (4), 707–710. <https://doi.org/10.1093/ijrl/eyy052>.
- Sleijpen, M., Mooren, T., Kleber, R.J., Boeije, H.R., 2017. Lives on hold: a qualitative study of young refugees' resilience strategies. *Childhood* 24 (3), 348–365. <https://doi.org/10.1177/0907568217690031>.
- Solimano, A., 2020. Global Mobility of the Wealthy, Their Assets and Inequality. Working Paper 45. CIGLOB International Center for Globalization and Development, Santiago, Chile.
- Streitwieser, B., 2019. International education for enlightenment, for opportunity and for survival: where students, migrants and refugees diverge. *J. Comp. Int. Higher Educ.* 11, 4–9. <https://doi.org/10.32674/jcihe.v11iFall.1357>.
- Stromquist, N.P., 2002. Education in a Globalized World: The Connectivity of Economic Power, Technology, and Knowledge. Rowman & Littlefield, Lanham, MD.
- Stromquist, N.P., Monkman, K. (Eds.), 2014. *Globalization and Education: Integration and Contestation Across Cultures*, second ed. Rowman & Littlefield, Lanham, MD.
- Tegenbos, J., Vlassenroot, K., 2018. Going Home? A Systematic Review of the Literature on Displacement, Return and Cycles of Violence. Politics of return working paper, (1). Conflict Research Programme. London School of Economics and Political Science, London, UK.
- Thakur, R., 2020. The United Nations and the North-South partnership: connecting the past to the future. *Ethics Int. Aff.* 34 (3), 305–317. <https://doi.org/10.1017/S0892679420000507>.

- Tsourapas, G., 2019. The Syrian refugee crisis and foreign policy decision-making in Jordan, Lebanon, and Turkey. *J. Global Secur. Stud.* 4 (4), 464–481. <https://doi.org/10.1093/jogss/ogz016>.
- Tuxen, N., 2019. Seeking “The Foreign Stamp”: International Education and the (Re) Production of Class Status in Mumbai, India. Doctoral dissertation. The Australian National University.
- UNESCO, November 19, 1974. Recommendation Concerning Education for International Understanding, Co-operation and Peace and Education Relating to Human Rights and Fundamental Freedoms. UNESCO, Paris.
- UNESCO, 2014. Global Citizenship Education: Preparing Learners for the Challenges of the 21st Century. UNESCO, Paris.
- UNESCO, 2018a. Global Education Monitoring Report 2019: Migration, Displacement and Education—Building Bridges, Not Walls. UNESCO, Paris.
- UNESCO, 2018b. The Status of the Right to Education of Migrants: International Legal Framework, Remaining Barriers at National Level and Good Examples of States’ Implementation. UNESCO, Paris.
- UNESCO/Council of Europe, 2001. Code of Good Practice in the Provision of Transnational Education. UNESCO/Council of Europe, Riga.
- UNHCR, 2011. Refugee Education: A Global Review. UNHCR, Policy Development and Evaluation Service, Geneva.
- UNHCR, 2021a. Global Trends—Forced Displacement in 2020. UNHCR UNHCR Global Trends—Forced displacement in 2020.
- UNHCR, 2021b. Refugee Data Finder UNHCR—Refugee Statistics.
- UNHCR (n.d) Master Glossary of Terms. UNHCR. UNHCR Master Glossary.
- United Nations, 1951. Convention Relating to the Status of Refugees. Adopted on 28 July 1951 by the United Nations Conference of Plenipotentiaries on the Status of Refugees and Stateless Persons Convened Under General Assembly Resolution 429 (V) of 14 December 1950.
- United Nations, 2018, 12 November. “12 Million” Stateless People Globally, Warns UNHCR Chief in Call to States for Decisive Action. UN News. <https://news.un.org/en/story/2018/11/1025561>.
- Van Mol, C., 2018. Becoming Europeans: the relationship between student exchanges in higher education, European citizenship and a sense of European identity. *Innovat. Eur. J. Soc. Sci. Res.* 31 (4), 449–463. <https://doi.org/10.1080/13511610.2018.1495064>.
- Walter, B.F., 2011. Conflict Relapse and the Sustainability of Post-conflict Peace. World Bank, Washington, DC.
- Wang, J., 2020. Limits of multicultural imagination and the anti-refugee controversy in contemporary China. *J. Contemp. Eastern Asia* 19 (2), 125–147. <https://doi.org/10.17477/jcea.2020.19.2.125>.
- Wasem, R.E., 2020. More than a wall: the rise and fall of US asylum and refugee policy. *J. Migration Human Secur.* 8 (3), 246–265. <https://doi.org/10.1177/2331502420948847>.
- Waters, J.L., Leung, M., 2017. Trans-knowledge? Geography, mobility, and knowledge in transnational education. In: Jöns, H., Meusbürger, P., Heffernan, M. (Eds.), *Mobilities of Knowledge*. Springer, Cham, pp. 269–285.
- Weller, M., 2020. Open and free access to education for all. In: Burgos, D. (Ed.), *Radical Solutions and Open Science. Lecture Notes in Educational Technology*. Springer, Singapore, pp. 1–15. https://doi.org/10.1007/978-981-15-4276-3_1.
- Wihlborg, M., Robson, S., 2018. Internationalisation of higher education: drivers, rationales, priorities, values and impacts. *Eur. J. High Educ.* 8 (1), 8–18. <https://doi.org/10.1080/21568235.2017.1376696>.
- Willems, K., Vernimmen, J., 2018. The fundamental human right to education for refugees: some legal remarks. *Eur. Educ. Res. J.* 17 (2), 219–232. <https://doi.org/10.1177/1474904117709386>.
- Williams, N., 2020. Moving beyond financial remittances: the evolution of diaspora policy in post-conflict economies. *Int. Small Bus. J.* 38 (1), 41–62. <https://doi.org/10.1177/0266242619878064>.
- Winn, M., 2021. The far-right and asylum outcomes: assessing the impact of far-right politics on asylum decisions in Europe. *Eur. Union Polit.* 22 (1), 70–93. <https://doi.org/10.1177/1465116520977005>.
- Yilmaz, B., 2022. Language and humanitarian governmentality in a refugee camp on Lesbos Island. In: De Fina, A., Mazzaferro, G. (Eds.), *Exploring (Im) Mobilities: Language Practices, Discourses and Imaginaries. Multilingual Matters*, Bristol, pp. 207–226.
- Zajda, J., Rust, V.D., 2021. Globalisation and comparative education: changing paradigms and issues. In: Zajda, J., Rust, V.D. (Eds.), *Globalisation and Comparative Education*. Springer, Dordrecht, pp. 139–157. https://doi.org/10.1007/978-94-024-2054-8_9.
- Zetter, R., 2021. Refugees and their return home: unsettling matters. *J. Refug. Stud.* 34 (1), 7–22. <https://doi.org/10.1093/jrs/feab005>.
- Ziguras, C., 2018. Soft power projection: the political return on investment in international higher education. In: Cantwell, B., Coate, H., King, R. (Eds.), *Handbook on the Politics of Higher Education*. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA, pp. 167–186.
- Zsófia, N., 2018. From camps to schools: the participation of refugee children in Greek public education. *J. Mod. Greek Stud.* 36 (2), 373–399. <https://doi.org/10.1353/mgs.2018.0028>.

Further reading

- Becker, J., 2021. Refugees, migration and the tightening borders in the Middle East. Perspective from biographical research on the re-figuration of spaces and cross-cultural comparison. *Forum* 22 (2), 184–212. <https://doi.org/10.17169/fgs-22.2.3598>.
- Bešić, E., Paleczek, L., Barbara Gasteiger-Klicpera, B., 2020. Don't forget about us: attitudes towards the inclusion of refugee children with(out) disabilities. *Int. J. Incl. Educ.* 24 (2), 202–217. <https://doi.org/10.1080/13603116.2018.1455113>.
- Bonizzoni, P., 2020. The border(s) within: formal and informal processes of status production, negotiation and contestation in a migratory context. In: Ambrosini, M., Cinalli, M., Jacobson, D. (Eds.), *Migration, Borders and Citizenship: Between Policy and Public Spheres*. Palgrave Macmillan, Cham, pp. 217–235.
- Marley, C., Mauki, B., 2019. Resilience and protective factors among refugee children post-migration to high-income countries: a systematic review. *Eur. J. Publ. Health* 29 (4), 706–713. <https://doi.org/10.1093/eurpub/cky232>.
- Richardson, E., MacEwen, L., Naylor, R., 2018. Teachers of Refugees: A Review of the Literature. IIEP-UNESCO and Education Development Trust, Reading Berkshire and Paris, ISBN 978-1-909437-95-1.
- Streitwieser, B., Loo, B., Ohorodnik, M., Jeong, J., 2019. Access for refugees into higher education: a review of interventions in North America and Europe. *J. Stud. Int. Educ.* 23 (4), 473–496. <https://doi.org/10.1177/1028315318813201>.
- Summers, K., Crist, J., Streitwieser, B., 2022. Education as an opportunity for integration: assessing Colombia, Peru, and Chile's educational responses to the Venezuelan migration crisis. *J. Migration Human Secur.* 1–18. <https://doi.org/10.1177/23315024221085189>.
- Wiseman, A.W., Galegher, E., 2019. Teacher preparation, classroom pedagogy, and the refugee crisis in national education systems. In: Wiseman, A.W., Damaschke-Deitrick, L., Galegher, E., Park, M.F. (Eds.), *Comparative Perspectives on Refugee Youth Education: Dreams and Realities in Educational Systems Worldwide*. Routledge, New York, pp. 75–101.
- Unangst, L., Ergin, H., Khajarian, A., DeLaquil, T., de Wit, H. (Eds.), 2020. Refugees and Higher Education: Trans-National Perspectives on Access, Equity, and Internationalization. Brill Sense, Leiden and Boston.

Relevant websites

Academic Network on Global Education & Learning Academic Network on Global Education and Learning, angel-network.net.
UNHCR (n.d.) Master Glossary of Terms. UNHCR Master Glossary. **This page intentionally left blank**

Online education

Ricardo Montelongo, Sam Houston State University, Eleanor & Charles Garrett Teacher Education Center, Huntsville, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Online education defined	149
The development of online education	149
Online education as a response to the COVID-19 pandemic	150
Online education frameworks	151
Culturally relevant online education	152
Student engagement in online education	152
Video introductions	153
Weekly overviews	153
Video feedback	153
Synchronous live meetings	153
Opportunities for interaction	153
Immersive video technologies	153
Finance issues associated with online education	154
Summary	154
References	155
Additional Readings	156

Glossary

Critical Digital Pedagogy A pedagogical approach to online teaching and learning that begins with awareness of critical perspectives (feminist, queer, critical race, etc.) and then examines digital tools, learning management systems, and other digital resources to ensure a relational and human-centered pedagogy is established

Digital Pedagogy Pedagogical skills and practices that considers the relationships existing between course content, teaching activities, and technology knowledge to formulate course instruction effective in virtual online classroom spaces

High-Context Communication Communication that emphasizes relationship-building actions and behaviors between individuals. These relations typically are measured by the level of nonverbal communication that is occurring between individuals, places, and events

Interface Interaction In response to the growth of online education, this form of interaction describes the behaviors and actions students and faculty express when interacting with mobile and digital technology as part of being engaged in online courses, especially through course requirements and experiences

Point-of-View Action Camera Small, mobile, and high-resolution video camera recorders that provide viewers wide-angled perspectives and high-definition image quality. Considered an innovative video technology, these cameras have been used in online education to provide immersive video experiences as part of online course lessons and presentations

Online delivery of courses has a significant presence in the educational experiences of students at all grade levels and post-secondary learning. Since the early-2000s, the expanded use of technology in student learning has influenced the rapid transformation of education to move beyond traditional brick-and-mortar classroom settings (Davila and Montelongo, 2020; Galuszka, 2005; Gourlay, 2021; Zimmerman, 2012). Worldwide, the growing popularity of online education has been enhanced by governmental initiatives to provide digital literacy (e.g., India's *Digital India*) and resources to offer open access online education (e.g., Australia's *Open Universities Australia*) (Palvia et al., 2018). In fall 2018, over six million post-secondary students in the US alone took at least one online course (National Center for Educational Statistics, 2019). This trend reflects the steady growth of online enrollments, despite the decline of overall higher education enrollments since 2012 (Seaman et al., 2018). Online education has been seen for its ability to increase course enrollments and lower infrastructure costs for universities (Casement, 2013; Deming et al., 2015; Meyer, 2010). Universities that offered online courses and degrees also provided effective outreach to non-traditional learners and full-time working students who desired flexibility with course scheduling (Seaman et al., 2018; Tallent-Runnels et al., 2006).

Despite online education's overall impact in responding to the needs of a more technologically confident citizenry, this mode of instruction still faces challenges on perceptions of becoming an impersonal, anonymous, and less effective learning experience compared to in-person course instruction (McKeown, 2012; Sun and Chen, 2016). Online education also faces criticism with regards to how educational administrators and public policy leaders use such course delivery primarily as a cost-cutting strategy to increase profitability and student enrollments without first understanding the pedagogical issues connected to online teaching and

learning (Meyer, 2010; Montelongo, 2019). While online education continues to increase its presence in educational institutions worldwide, there needs to be further understanding on its evolution, development, and impact on student academic engagement and outcomes which are necessary for its further expansion and enhancement.

Online education defined

Online education is an umbrella term used for any learning and instruction that is presented on a computer or digital device usually virtually using the internet. Online education is also commonly referred to as distance education, online learning, web-based teaching, and electronic learning (e-learning). Online education is also associated with the instructional approaches of blended learning or hybrid learning, where courses provide face-to-face class sessions in combination with lessons which are completed online at the student's pace. Online education is typically linked to the educational trend of flipped classrooms due to its approach of blending face-to-face classroom interaction with independent study at home, which usually makes use of technology and multimedia.

Online education has several common characteristics that is shared among its variety of terms and approaches: Technology use, geographic separation, asynchronous nature, and digital mobility. Online education has reflected the shifts in using technology for student instruction and learning. Since the early-2000s, the use of technology has been evolving to utilize innovative and current software and devices (Davila and Montelongo, 2020; Galuszka, 2005). According to Seaman et al. (2018), technologies used in online education include:

Internet, one-way and two-way transmissions through open broadcasts, closed circuit, cable, microwave, broadband lines, fiber optics, satellite or wireless communication devices; audio conferencing; and video cassette. DVDs, and CD-ROMS, if the cassette, DVDs, and CD-ROMS are used in a course in conjunction with the technologies listed above (pg.5).

Technologies used in online education are dynamic and constantly being updated and introduced to consumers. Thus, some items listed above can be seen as already or quickly becoming outdated (e.g., video cassettes, DVDs). Currently, online education already makes use of application-based social media platforms, video conferencing, video and audio sharing platforms, and immersive video technologies such as point-of-view cameras (Montelongo, 2020; Montelongo and Eaton, 2019b; Sun and Chen, 2016).

Another common characteristic of online education is that the learner is separated physically from the course instructor (Seaman et al., 2018). This physical separation can be perceived as both an asset and disadvantage to online education. By being separated by the course instructor, the earliest critiques of online education characterized this mode of instruction as being "no more than a hodgepodge of ideas and practices taken from traditional classroom settings and imposed on learners" (McIsaac and Gunawardena, 1996; Gunawardena and McIsaac, 2013, p.5). The current forms of online education address this separation by emphasizing the importance of online student engagement and transferring high-impact teaching practices into virtual environments (Bailey et al., 2018; Fink, 2016; Montelongo, 2019). Current online education pedagogies have been developed to encourage online student engagement through intentional use of videos for relationship building (Costa, 2020; Montelongo and Eaton, 2019b), developing virtual community engagement activities (Montelongo and Eaton, 2018) and creating social justice opportunities in online learning (Guthrie and McCracken, 2010).

Asynchronous learning is a key component of the online education definition. By providing self-paced online course lessons, students complete course requirements and content according to their own schedules. While online education does not necessarily free students from stated deadlines and lesson tasks, the asynchronous nature of online education allows students educational opportunities that might have been challenging due to access, financial, and schedule challenges (Sun and Chen, 2016). Online education is seen to serve a large, adult learner population that tends to be older than traditional students (Tallent-Runnels et al., 2006). The asynchronous nature of online education provides these students learning spaces where they hold various roles outside the course environment and supports their goal-oriented drive for online education (Sun and Chen, 2016).

The last characteristic of online education is reflective of the rapid evolvement of digital technologies used in education. Online education is described by its ability to be available anytime and anywhere for students due to the emergence of mobile devices such as laptops, tablets, and smartphones. Online education instruction is no longer bound to desktop computers and much like Graham's (2012) assessment that learning spaces built in the 1960s were unlikely to meet the needs of students in the 21st century, online education resources prior to the late 2010s could already be obsolete for today's online learning spaces. Zimmerman (2012) initiated awareness of how students interact with online course content using mobile technological devices and educators need to keep pace with how the newest generation of learners use mobile technology in their academic learning and behaviors.

The development of online education

The exact origin of online education is debatable among educational scholars, but discussions on its evolution tend to start with how postal, radio, and television outlets have been used for instruction in the United States. Some of the earliest forms of online education can be traced back three centuries ago when shorthand was taught to students in rural areas of New England using letters

exchanged in the postal system (Larreamendy-Joerns and Leinhardt, 2006; Palvia et al., 2018). The idea of creating and enhancing correspondence courses reflected a growing idea in the late 1800's of expanding educational opportunities to women in Boston and seeing this form of learning to further democratize higher education (Larreamendy-Joerns and Leinhardt, 2006). Online education, even in its earliest forms in the United States, appeared to reach and recruit students that may not had the means or access to higher learning.

Since its earliest forms using letters sent through the mail, online education has made use of emergent technologies being developed within societies. Radio or television at the turn of the 20th century were being explored as possible modes of providing instruction by higher education institutions. As radio grew to become the primary communication medium during World War I, colleges and schools saw its possibility for use in course delivery (Sun and Chen, 2016). One of the earliest uses of radio for instruction can be found in Wisconsin with its use of a federally license radio station at the University of Wisconsin and the development of the School of the Air, which the Madison Public School district worked in collaboration with the university to deliver programs aimed at primary and secondary schools around the state (Palvia et al., 2018; Sun and Chen, 2016). The growth of television in the 1950's followed which allowed institutions to provide visual elements to distance learning. Visual instruction via television now allowed students to interact with instructors who were not in physically in the same space. Televised course instruction expanded the ability of universities and colleges in the United States to offer courses for college credit.

The emergence of the internet and the World Wide Web provided another wave of technological innovation for educators to consider for course delivery. The early 1990's was seen as a milestone for the forward progress of online education and the eventual study and growth of online teaching and learning (Sun and Chen, 2016). The rise of the internet not only expanded the number of course offerings that colleges and universities could offer to students, but also led to the advent of fully online degree programs and online campuses such as the University of Phoenix. At the K-12 level, school districts and charter schools US started to expand their use of the internet to provide instruction to students and to use it as a strategy to counter teacher shortages in some areas of study such as science and math.

Internationally, the evolution of online education is largely benefitted from governmental support of digital literacy and responding to demands for continuing education by working professionals and jobseekers (Pavia et al., 2018). Online education does face challenges with regards to cultural differences in learning styles compared to western models of education. Online education tends to be critiqued on its lack of providing high-context communication which is valued in global cultures found for example in Latin America, Asia, and Africa (Westbrook, 2014). High-context communication focuses on the visual and auditory facets of individual speaking such as non-verbal cues, hand gestures, vocal tone, and emotion. The asynchronous nature of online learning usually relies primarily on written communication, which misses these high-context elements of interaction. Despite this challenge, the evolving field of digital pedagogy is addressing these concerns and providing advice and strategies on how to provide culturally relevant online course instruction (Costa, 2020; Montelongo, 2019; Montelongo and Eaton, 2019b).

Online education as a response to the COVID-19 pandemic

According to UNESCO, an estimated 850 million individuals transitioned to some alternative form of teaching and learning worldwide as a response to the COVID-19 pandemic (UNESCO, 2020). In March of 2020, as the COVID-19 virus quickly spread throughout the world, educators and students found themselves wondering how their learning would continue in environments where lockdowns, avoidance of crowds in any space, and social distancing were quickly becoming the norm. Educational systems globally were seeking ways to keep students in the classroom to not disrupt their learning. As global societies were dealing with protecting citizens, "decisions to cancel, postpone, or move in-person classes online came within a matter of days in most countries" (Johnson et al., 2020, p.7).

Shifting to online learning in times of disruption such as natural disasters has been a strategy used previously by educational systems. In the United States, especially along the Gulf Coast where tropical weather events are common, transitioning courses to online environments "on the fly" has been a strategy used to assist students with their education when storms impact this area (Lorenzo, 2008, p.5). However, the COVID-19 pandemic was not a localized or regional event. Educational systems nationwide and globally had to respond quickly to transition to online learning environments. As of this writing, the COVID-19 pandemic continues to have global impact with variants of the virus (i.e., Delta variant) rapidly spreading across areas of the US, India, and other parts of the world. Despite this variant spread, schools and campuses are planning or considering a return to in-person instruction. The experiences of both students and teachers/faculty with online education during the pandemic continues to be researched and published in the literature. This chapter provides a brief summary of this emergent field of educational research.

Faculty and administrators' experiences with dealing with the COVID-19 educational response was studied by Johnson et al. (2020). In their study, they surveyed 897 higher education faculty and administrators in the US Four-year and two-year colleges and universities were represented and over one-third came from large public institutions student enrollments of over 20,000. Using their study as reflective of the experiences faced by this group of educators, Johnson and her associates found that faculty with no experience in online teaching reported applying new teaching methods in online courses. Key to this quick transition to online teaching and learning was the use of an institution's Learning Management System (LMS) to distribute course content and materials. Also notable for this study was the use of synchronous video (i.e., Zoom, GoToMeeting, Google Hangouts, etc.) as a new instructional tool for faculty and administrators. The rise of synchronous video platforms is one of the most memorable developments of

the pandemic response in education. All levels of education, from kindergarten to doctoral-level seminars, experienced the use of synchronous video class meetings as part of this response. Faculty and administrators also noticed the concern for providing student services and support to provide access to digital resources and tools during this transition (Gourlay, 2021). Care and concern for students was stated as primary and necessary reasons for the changes that faculty made to their courses during the pandemic (Johnson et al., 2020).

Student perspectives on the COVID-19 pandemic transition to online education is assisted by previous research investigating their perceptions of online courses compared to face-to-face instruction. Students tend to perceive online education as being more flexible and friendly to their work and life schedules, but not as effective to learning compared to face-to-face instruction (Platt et al., 2014). A national survey of over 1000 US undergraduates who experienced the transition from in-person classes to remote instruction was performed in spring 2020 (Means and Neisler, 2021). In this study, student satisfaction levels with courses took a sharp decline after transitioning to remote instruction. Students however provided helpful advice and recommendations to make online learning and teaching more effective and satisfying. Teaching practices that made the online education environment more personable and helpful included personal messages providing up-to-date information on how they are doing and course activities that asked students for reflections on their learning (Means and Neisler, 2021). Students in this study responded that their online education would be enhanced if “instructors provided multiple ways for students to engage in their online courses and demonstrate their learning” (Means and Neisler, 2021, p. 23).

Educators will continue to learn from the lessons provided by the transition to online education as a response to the pandemic crisis of 2020. As research is further developed on students, faculty, and administrators with their online education experience, the literature already highlights the need for institutions to enhance and improve their LMS's to match the demands of faculty and students who now have immersed themselves in online teaching and learning environments. Faculty and teachers are now aware of digital resources and technologies to enhance their teaching skills and administrators need to provide additional professional development for faculty and teachers to boost their skills in using educational technology and digital learning tools. Students' perceptions of online education will likely be similar to those prior to the pandemic, but there is now a new knowledge on how to improve online learning through understanding effective digital pedagogies and intentional strategies to break down the anonymous nature of virtual learning spaces by using tools such as video and social media.

Online education frameworks

The rapid emergence of technology in education has created an urgent need for educators to understand how course content, teaching practices, and student engagement intersect within digital spaces. These relationships impact the online learning experience of students and how teachers convey their knowledge of content into online environments. Two frameworks provide conceptualizations in which to understand the complex interactions occurring between digital technology and pedagogical practice: Technological Pedagogical Content Knowledge (Mishra and Koehler, 2006) and Critical Digital Pedagogy (Morris and Stommel, 2018).

Mishra and Koehler's (2006) Technological Pedagogical Content Knowledge, also known as the TPACK model, challenge faculty who teach with technological resources to think beyond content knowledge and pedagogical strategy, arguing that educators have undertheorized the relationship between technology, pedagogy, and content. With their model, Mishra and Koehler remind educators when introducing technology into the educational environment, they must also carefully consider how this technology is used and its relationship with content knowledge and pedagogical practices. As Mishra and Koehler (2006) state, “knowledge about content, pedagogy, and technology is central for developing good teaching” (p. 1025). All three are in relation.

Digital technologies come into the educational arena with new possibilities and constraints. The dual nature of technology constraining familiar and opening new pedagogical practices challenge educators to think innovatively and differently with their courses. For online educators, what might be done easily in face-to-face classrooms can sometimes become challenging or impossible to do within online learning environments. The TPACK Model underscores the need for educators, especially those teaching online, to have awareness, knowledge, skills, and understanding of technology use in teaching and learning. Online educators must examine how implementation or engagement with digital technological tools, spaces, and places enhance or detract pedagogical opportunities related to a course or module content.

Critical digital pedagogy is an emerging field of study which looks at how students and educators look at the tools used in online teaching and learning and how these create community, collaboration, engagement, and empowerment in digital spaces (Morris and Stommel, 2018). Online education is viewed as a social environment and digital pedagogy is about human relationships that exist within these spaces. To establish and work on these relationships, online educators consider how technology enhances or weakens this goal within the digital environment. Critical digital pedagogy views technology as neither neutral or value free and that online environments such as those found in the LMS must do more than just be repositories for course content (Montelongo and Eaton, 2019b; Morris and Stommel, 2018). Technological tools are part of this relationship building according to this framework. In using critical digital pedagogy in an online course, educators pay attention to the relationships between teaching and student; between student and student; between user and technological tool; and the social identities held by those within the online course environment (Montelongo and Eaton, 2019a).

Culturally relevant online education

Ideas such as critical digital pedagogy allows online education to expand its purposes to allow educators to reimagine and reconsider what it means to teach and build relationships with students. Online education, thus, can continue efforts to foster cultural diversity and social justice in digital spaces.

Online education can be viewed as a learning space that is not anonymous, impersonal, and value-free (Morris and Stommel, 2018). In fact, online education can become a powerful strategy for educators to strengthen relationships with students, provide agency to improve support, and motivate students to actively engage students in course content.

One area of research that has been addressing how diversity and social justice is presented in digital space is on educators who teach fully online diversity courses. As institutions continue to increase the number of online course offerings and fully online degree programs are established, students have the possibility of enrolling in a required diversity course for their degrees. Online education's growth includes the growth of teaching multicultural education courses that address issues of racism, discrimination, privileges, and other issues involved in equity and inclusion discussions. The complexity apparent in these topics are challenging for educators to address in traditional face-to-face classroom environments, even more so in online environments (Akintunde, 2006). However, the emerging research on online diversity courses allows online educators knowledge and understanding on concepts which should be examined for all online course instruction. Two concepts to consider for effective online teaching from a diversity and social justice lens are culturally relevant pedagogy (Ladson-Billings, 1995) and high-context learning (Westbrook, 2014).

Culturally relevant pedagogy is described by Ladson-Billings (1995) as a series of pedagogical practices that can be placed along a continuum. This series of practices have educators reflect on how they conceive of themselves and others; how they build relationships with their students; and where they go to find and process knowledge. This form of pedagogy focuses on eliminating the deficit-minded thinking usually placed on students of color with regards to their academic difficulties. Ladson-Billings pedagogical stance focuses more on how educators can develop skills and practices that values student experiences and their communities. Key to advancing these pedagogical ideas is the notion of providing care and support to students in the classroom, along with valuing linguistic and cultural differences to help students navigate a sometimes "inequitable and undemocratic" world (Ladson-Billings, 1995, p. 474). Advancing this in online education requires digital classrooms to become spaces where students are heard and appreciate, not just seen, through the LMS. High-context communication and relationship building within online course modules are initial strategies to promote this environment.

Recognition of low-context and high-context cultures (Westbrook, 2014) provide awareness of cross-cultural communication between individuals within classroom settings. High-context cultures emphasize the relationship-building that can occur when individuals communicate with each other and messages received through nonverbals such as body gestures, hand movements, and facial expressions. Low-context cultures rely primarily on communication received through written and spoken word with less emphasis on establishment of relationships and use or presence of nonverbal cues. In online education, awareness of these cultural contexts apparent in communication is important in creating effective learning experiences that not only engages online students, but also connects to how diverse students' cultural norms are represented in the digital learning space. When considering high-context communication, attention is placed on communication, visual cues, and relationship building.

Montelongo (2019) outlined the use of high-context teaching in online graduate student instruction. In his recommendations to enhance this form of culturally relevant teaching, he encourages online educators to make extensive use of videos to provide instructor presence within online course modules. The use of videos provides nonverbals such as facial expressions, hand gestures, and emotional reactions. The addition of one or more scheduled synchronous live class meetings in online courses also promotes community building among classmates, where they are joined live together to see and hear each other and to converse on course topics. An understanding of pedagogical strategies to use in online teaching to foster cultural validation also provides opportunities for increase student engagement.

Student engagement in online education

Online education usually faces a level uncertainty and reluctance from both students and faculty due to unfamiliarity and negative experiences within this teaching environment (Bailey et al., 2018; McKeown, 2012). Ideas on the effective pedagogy typically uses the lens of traditional college classroom environments and spaces, which rely on brick-and-mortar physical classroom settings. Despite this bias, a growing body of knowledge is developing strategies to apply in online education.

High-impact teaching practices (Fink, 2016) are being applied by online instructors to strengthen engagement with online course content. Fink created a list of teaching practices highlighting the need to help students learn more about learning, designing courses that are learner-centered, creating academic community engagement activities, and increasing classroom interaction with students. Montelongo and Eaton (2019a) describe how they utilize high-context learning and high-impact teaching in their online courses, especially with their online graduate-level diversity courses for their fully online master's program. Their colleagues in the same program also found that creating learning environments that required active critical thinking and problem solving tend to improve student engagement in online courses (Holzweiss et al., 2014). Strategies to build an online environment for active student engagement include video introductions, weekly overviews, video feedback, synchronous live meetings, opportunities for interaction, and immersive video technologies.

Video introductions

In order to provide an initial faculty presence within the online course environment, video introductions offer students information on the instructor's background and interest in the course topic. These videos, which are usually no more than 5 min in length, are usually placed within the course LMS space at a location easily found outside any course lesson module. The introduction video offers students information on their instructor, including any personal reflections on why they teach the course. Establishing faculty video presence in the online course was found to be an important characteristic of a high-impact, high-context online course ([Montelongo and Eaton, 2019a,b](#)). Students who see their instructor's face and voice in the digital space eliminates the anonymous nature usually attached to online learning. Students can also develop video introductions of themselves to share to classmates within their online course. These student introductions can include information about who they are, their interest in the class, and their goals for the course and semester. The video introductions are a possible strategy to reduce the anxiety and uncertainty students may come with due to ideas that online learning is impersonal and distant.

Weekly overviews

Weekly overviews are videos produced by the online instructor which are placed at the start of each online course module. In these videos, which again are usually 5 min or less, the instructor provides their initial thoughts on the module topics for that week and offers any personal experiences with the topics at hand. If appropriate, these weekly overview videos can also connect current, real-time news events to the introduction to further add context to that week's online course module. Weekly overviews support student engagement and helps manage the asynchronous characteristic by enticing online learners to check into their course modules early.

Video feedback

Video feedback is a strategy used to provide students high-context communication on graded assignments where students see and hear comments directly from the instructor via video. In using this strategy, online instructors can produce short videos, usually 3 min in length, with responses and reactions to students' coursework. Rather than just relying on the low-context feedback offered by written comments or grading rubrics embedded in the LMS, students can receive high-context feedback where the instructors voice tone, facial expressions, and body language are observed and heard as part of the grade. Video feedback also offers high-context cues to highlight areas of improvement, quality, and excellence.

Synchronous live meetings

As mentioned earlier, one memorable outcome of education's response to the pandemic was extensive use of video conferencing platforms for synchronous class meetings. While fully online courses are desired largely for their asynchronous nature, including 1–2 synchronous meetings in online courses provide opportunities for students to meet each other, ask questions for the class, and the possibility to include real-time discussions and presentations on course material. To make the synchronous live meetings as effective as possible, instructors are recommended to make student attendance mandatory. Advanced scheduling is necessary for this strategy to be employed. Real-time conversations through synchronous class meetings are effective and simple ways to create high-impact, high-context online education.

Opportunities for interaction

After several semesters of online teaching, [Montelongo and Eaton \(2018\)](#) found that students often enter online courses with a level of uncertainty about online learning, especially with class dynamics. In preparation for online teaching, these instructors in their course planning included opportunities for students to interact with faculty. An effective part of this strategy, especially in teaching online diversity courses, was the inclusion of one-to-one meetings to discuss course ideas. In the diversity course, this was seen in a course activity title "Multicultural Chats." The chats were usually scheduled in the latter part of the semester and offered a chance for both student and instructor to share conversations on personal development and social identity. While this strategy requires time and effort to meet individually, doing so provided a concluding experience where both student and instructor learned more about specific connections, experiences, and personal development on the course topic. In other courses, different approaches to creating this interaction are possible using assignments (e.g., podcasts, interviews, case studies, etc.) as opportunities to connect.

Immersive video technologies

Placing videos with lesson introductions, interviews, and interactions with campus settings provide high-context communication within online lesson modules ([Montelongo and Eaton, 2019b](#)). One distinct method for this visual communication is use of point-of-view action cameras (e.g., GoPro) in producing online course lessons. Innovative video technologies like that found in point-of-view action cameras (e.g., GoPro) are being tried and tested in online instruction. Point-of-view camera recordings provide immersive experiences in viewing by providing wide-angled perspectives and high-definition high-resolution recordings. Immersive video technologies were described in their use in teaching in a fully online master's education program ([Montelongo, 2020](#)). The

immersive video captures rich contextual information that replaces static presentation slides with dynamic video filled with rich detail. The video technology fosters a culturally relevant teaching approach to online education in the course that goes beyond relying on written text to describe course content and ideas.

Finance issues associated with online education

The increased presence of online education brings attention to the finance issues facing such instruction. As the growth of online education continues to be used as a cost-reduction strategy, caveats are presented to say such strategies diminish the quality of education as an outcome of the supposed cost benefits (Deming et al., 2015). The cost-effectiveness of online education is evident in the literature, but also are its supposed shortcomings. For example, in a comparison of grade-based outcomes between online and face-to-face courses, one study found that struggling university students in the US, especially those with lower grade point averages, struggle with online courses and fare worse in them compared to face-to-face courses (Cavanagh and Jacqueman, 2015). While these researchers focused more on student outcomes and not course experiences, their findings raise the issue on whether online learning experiences provide similar classroom learning experiences compared to brick-and-mortar classrooms. A frequently mentioned outcome in discussions on online education is increased revenue while reducing operating costs (Bailey et al., 2018).

To transform online education into high-impact practice, concentrated focus on effort and resources needed by instructors to develop online courses with these qualities is necessary (Tallent-Runnels et al., 2006). Online education research findings often feed notions of what this form of learning cannot do, rather than what it can do, especially with regards to its possibilities of providing high-impact teaching. Online education provides beneficial outcomes for faculty, students, and administration when viewed for its potential outcomes. Under the lens of educational finance, awareness of faculty effort, organizational support and professional development, and equitable student support should be considered in fiscal discussions.

Financial considerations need to take account the support necessary to not only provide online delivery of courses, but also the professional training and development of online instructors and faculty. Despite the increasing prevalence of online education, most organizational structures in the US and globally are organized according to brick and mortar, face-to-face models. Organizational support needs to understand what it means to instruct and learn in digital space. Institutions need to provide funding for faculty professional development so that they can navigate a landscape of continually proliferating technological changes. The importance of online educators receiving regular training and professional development in technological knowledge is critical for the improvement and effectiveness of online educational programs (Major, 2015). Online educators need to be kept informed on the rapid innovations apparent in technology and be comfortable with utilizing a variety of technological tools used to enhance student learning and engagement. Educational institutions now support and sponsor a variety of educational and professional development seminars and conferences on digital pedagogy and learning. For example, the Digital Pedagogy Lab based at the University of Mary Washington in Virginia is an international professional development institute specializing in training on critical digital pedagogy and online learning (Morris and Stommel, 2018). Chief academic officers need to understand what this entails and promote professional development to ensure that technological competence is held by online instructors (Montelongo, 2019).

Delivery of online course content starts with a learning management system (LMS) that is functional, reliable, and secure. To move online education toward high-impact levels, there also needs to be institutional support for tools outside the prescribed LMS for course instruction such as video production services, instructional content designers, and additional applications and programs to enhance the LMS' online instruction. Higher education organizational structures tend to reflect the traditional brick-and-mortar, face-to-face model. As we realized in 2020, institutions need significant investments to enhance and improve their online education support offices and staff. This investment should include a "central team of instructional designers, web designers, multimedia personnel, data analysts, quality assurance experts, and student support services staff" (Bailey et al., 2018, p. 29).

In a report six colleges and universities in the US who state having high-quality online learning organizations, all were found in the report to have strong teams of instructional support professionals, instructional designers, web designers, and multimedia personnel, just to name a few (Bailey et al., 2018). The report found that these US institutions provide an investment of between \$2 to \$14 per student credit hour to support these online education support units (Bailey et al., 2018). While the report did mention the economic benefits of institutions reduce their need to provide and maintain physical campus learning spaces, the return on investment included increased retention, graduation rates, and other positive outcomes for institutions in the report. Interestingly, faculty in the report stated that by having the strong institutional online education support, they were able to create online courses that were more rigorous and challenging compared to face-to-face courses (Bailey et al., 2018). Organizational support and a well-developed strategic plan that includes professionals to assist faculty and students in their online experience appears to have positive institutional outcomes, while also providing a reduction of costs in delivering online education. This is in large part in viewing online education more for its potential, rather than its shortfalls (Bailey et al., 2018).

Summary

Since its first presence in education in the late-1800's as correspondence courses using mail, teaching and learning outside traditional brick and mortar classrooms continues to "shake up pedagogical routines" (Larreamendy-Joerns and Leinhardt, 2006, p. 597). Today's online education is constantly evolving to meet the dynamic needs of today's learners and to respond to the rapidly

changing technologies which can be applied to our transforming educational pedagogies. Online education is seen globally as a strategy to apply when disruptions in student learning occur from natural disasters, political upheavals, or health crisis. The strengths and weaknesses of online education have been experienced 2020 academic year by multitudes of students and instructors, from K-12 to all levels of higher education, when classes transitioned to online spaces in a matter of weeks due to the COVID-19 global pandemic. Research is being conducted on what we learned from this unprecedented global crisis in education. The experiences from students, instructors, and administrators continue to be investigated and the future literature will offer an abundance of information to use further enhance and improve online education. With this emergent research, theories and frameworks to provide guidance and foundations to online teaching and learning are assured to be reviewed and explored to understand what creates the most effective and high quality online educational experience. Montelongo (2019) writes that in order “to promote online education in the ‘more than’ lens that reflects high impact teaching and learning, institutions need to invest in the effort provided by faculty in developing and implementing rigorous and impactful online learning” (p. 72). As educational institutions worldwide continue to increase the number of online courses, educators need to learn which pedagogical practices are effective in course delivery and which financial issues need to be addressed to effectively deliver online courses. Online education already has significant presence in the educational experiences of students and instructors in all parts of the world. As more students include online elements in their educational careers and more instructors are asked to teach in online learning spaces, the future of education is impacted by current and future trends in technology and growing innovation in digital pedagogy. Online education is constantly dynamic and while most educational organization structures center their instructional models toward a face-to-face classroom spaces, teaching and learning is shifting to include elements of digital pedagogy and understanding strategies to create high-impact online learning environments. The future of online education provides opportunities to reimagine critical thinking, problem-solving, and discovery in digital spaces.

References

- Akintunde, O., 2006. Diversity.com: teaching an online course on white racism and multiculturalism. *Multicult. Perspect.* 8 (2), 35–45.
- Bailey, J., Vaduganathan, N., Henry, T., Laverdiere, R., Pugliese, N., 2018. Making Digital Learning Work. The Boston Consulting Group, Inc. Retrieved from: <https://edplus.asu.edu/what-we-do/making-digital-learning-work>.
- Casement, W., 2013. Will online learning lower the price of college? *J. Coll. Admiss.* 220, 14–18.
- Cavanaugh, J.K., Jacquemin, S.J., 2015. A large sample comparison of grade based student learning outcomes in online vs. face-to-face courses. *Online Learn.* 19 (2), n2.
- Costa, K., 2020. 99 Tips for Creating Simple and Sustainable Educational Videos: A Guide for Online Teachers and Flipped Classes. Stylus Publishing, LLC.
- Davila, V., Montelongo, R., 2020. Considering digital technology and innovative learning spaces as “structures for serving” at hispanic serving institutions. In: Garcia, G.A. (Ed.), *Hispanic Serving Institutions (HSIs) in Practice: Defining “Servingness” at HSIs*. Information Age Publishing, pp. 117–134.
- Deming, D.J., Goldin, C., Katz, L.F., Yuchtman, N., 2015. Can online learning bend the higher education cost curve? *Am. Econ. Rev.* 105 (5), 496–501. Available at: <https://doi.org/10.3386/w21080>.
- Fink, L.D., 2016. Five high impact teaching practices: a list of possibilities. In: *Collected Essays on Learning and Teaching*, vol. 9, pp. 3–18.
- Galuszka, P., 2005. Technology's latest wave, Black issues High. *Educ.* 22 (2), 24–28.
- Graham, C., 2012. Transforming spaces and identities: the contributions of professional staff to learning spaces in higher education. *J. High Educ. Pol. Manag.* 34 (4), 437–452.
- Gourlay, L., 2021. There is no “virtual learning”: the materiality of digital education. *J. New Approaches Educ. Res.* 10 (1), 57–66.
- Gunawardena, C.N., McIsaac, M.S., 2013. Distance education. *Handbook of Research on Educational Communications and Technology*. Routledge, pp. 361–401.
- Guthrie, K.L., McCracken, H., 2010. Teaching and learning social justice through online service-learning courses. *Int. Rev. Res. Open Dist. Learn.* 11 (3), 311–323.
- Holzweiss, P.C., Joyner, S.A., Fuller, M.B., Henderson, S., Young, R., 2014. Online graduate students' perceptions of best learning experiences. *Dist. Educ.* 35 (3), 311–323. <https://doi.org/10.1080/01587919.2015.955262>.
- Johnson, N., Veletsianos, G., Seaman, J., 2020. U.S. faculty and administrators' experiences and approaches in the early weeks of the COVID-19 pandemic. *Online Learn.* 24 (2), 6–21.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491. <https://doi.org/10.3102/00028312032003465>.
- Larreamendy-Joens, J., Leinhardt, G., 2006. Going the distance with online education. *Rev. Educ. Res.* 76 (4), 567–605.
- Lorenzo, G., 2008. The Sloan semester. *J. Async. Learn. Network* 12 (2), 5–40. <https://files.eric.ed.gov/fulltext/EJ837474.pdf>.
- Major, C.H., 2015. *Teaching Online: A Guide to Theory, Research, and Practice*. JHU Press.
- McIsaac, M.S., Gunawardena, C.N., 1996. Distance education. *Handbook of Research for Educational Communications and Technology* 403–437.
- McKeown, K.D., 2012. Can Online Learning Reproduce the Full College Experience? Center for Policy Innovation. Discussion Paper. Number 3. Heritage Foundation.
- Means, B., Neisler, J., 2021. Teaching and learning in the time of COVID: the student perspective. *Online Learn.* 25 (1), 8–27.
- Meyer, K.A., 2010. If higher education is a right, and distance education is the answer, then who will pay? *J. Async. Learn. Network* 14 (1), 45–68.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach. Coll. Rec.* 108 (6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>.
- Montelongo, R., Eaton, P.W., 2018. Beyond geographical space: online learning and community engagement initiatives. In: Evans, H. (Ed.), *Community Engagement Best Practices across the Disciplines: Applying Course Content to Community Needs*. Rowman & Littlefield Publishing Group, Lanham, MD, pp. 89–102.
- Montelongo, R., Eaton, P.W., 2019a. Online learning for social justice and inclusion: the role of technological tools in graduate student learning. *Int. J. Inf. Learn. Technol.* 37 (1–2), 33–45. <https://doi.org/10.1108/IJILT-11-2018-0135>.
- Montelongo, R., Eaton, P.W., 2019b. Strategies and reflections on teaching diversity in digital space(s). In: Kyei-Blankson, L., Blankson, J., Ntuli, E. (Eds.), *Care and Culturally Responsive Pedagogy in Online Settings*. IGI Global, Hershey, PA, pp. 41–62.
- Montelongo, R., 2019. Less than/more than: issues associated with high-impact online teaching and learning. *Adm. Issues J.* 9 (1), 68–79. <https://doi.org/10.5929/9.1.5>.
- Montelongo, R., 2020. Using immersive video camera technology to enhance online graduate preparation courses. *J. Coll. Student Dev.* 61 (5), 654–657. <https://muse.jhu.edu/article/766020>.
- Morris, S.M., Stommel, J., 2018. An Urgency of Teachers: The Work of Critical Digital Pedagogy. Hybrid Pedagogy, Inc. available at: <https://urgencyofteachers.com/>. (Accessed 15 June 2021)
- National Center for Education Statistics, United States, Office of Educational Research, Improvement, Center for Education Statistics and Institute of Education Sciences (US), 2019. Digest of Education Statistics, 46. US Department of Health, Education, and Welfare, Education Division, National Center for Education Statistics.

- Palvia, S., Aeron, P., Gupta, P., Mahapatra, D., Parida, R., Rosner, R., Sindhi, S., 2018. Online education: worldwide status, challenges, trends, and implications. *J. Global Inf. Technol. Manag.* 20 (4), 233–241. <https://doi.org/10.1080/1097198X.2018.1542262>.
- Platt, C.A., Raile, A.N.W., Yu, N., 2014. Virtually the same? Student perceptions of the equivalence of online classes to face-to-face classes. *MERLOT J. Online Learn. Teach.* 10 (3), 489–503.
- Seaman, J.E., Allen, I.E., Seaman, J., 2018. Grade Increase. Tracking Distance Education in the United States. Babson Survey Research Group. Available at: <http://www.onlinelearningsurvey.com/highered.html>.
- Sun, A., Chen, X., 2016. Online education and its effective practice: a research review. *J. Inf. Technol. Educ.* 15.
- Tallent-Runnels, M.K., Thomas, J.A., Lan, W.Y., Cooper, S., Ahern, T.C., Shaw, S.M., Liu, X., 2006. Teaching courses online: a review of the research. *Rev. Educ. Res.* 76 (1), 93–135.
- UNESCO, March 19, 2020. Half of world's student population not attending school: UNESCO launches global coalition to accelerate deployment of remote learning solutions. <https://en.unesco.org/news/half-worlds-student-population-not-attending-school-unesco-launches-global-coalition-accelerate>.
- Westbrook, T.P., 2014. Global contexts for learning: exploring the relationship between low-context online learning and high-context learners. *Christ. High Educ.* 13 (4), 281–294.
- Zimmerman, T., 2012. Exploring learner to content interaction as a success factor in online courses. *Int. Rev. Res. Open Dist. Learn.* 13 (4), 152–165. <https://doi.org/10.19173/irrodl.v13i4.1302>.

Additional Readings

- Bain, K., 2004. What the Best Teachers Do. Harvard University Press, Cambridge, MA.
- Crawford-Ferre, H.G., Wiest, L.R., 2012. Effective online instruction in higher education. *Q. Rev. Dist. Educ.* 13 (1), 11–14.
- Fedynich, L., Bradley, K.S., Bradley, J., 2015. Graduate students' perceptions of online learning. *Res. High. Educ. J.* 27. Retrieved from. <https://files.eric.ed.gov/fulltext/EJ1056187.pdf>.
- Herie, M., 2005. Theoretical perspectives in online pedagogy. *J. Technol. Hum. Serv.* 23 (1–2), 29–52. https://doi.org/10.1300/J017v23n01_03.
- Koc, S., Liu, X., 2016. An investigation of graduate students' help-seeking experiences, preferences and attitudes in online learning. *Turkish Online J. Educ. Technol.* 15 (3), 27–38.
- Kyei-Blankson, L., Blankson, J., and Ntuli, E. (eds.), Care and Culturally Responsive Pedagogy in Online Settings (pp. 41-62). Hershey, PA: IGI Global.
- McGinley, V., Osgood, J., Kenney, J., 2012. Exploring graduate students' perceptual differences of face-to-face and online learning. *Q. Rev. Dist. Educ.* 13 (3), 177–182.
- Stommel, J., Friend, C., Morris, S.M. (Eds.), 2020. Critical Digital Pedagogy: A Collection. Hybrid Pedagogy Incorporated.
- Zembylas, M., 2008. Engaging with issues of cultural diversity and discrimination through critical emotional reflexivity in online learning. *Adult Educ. Q.* 59 (1).

Digital learning divide

Josef Kuo-Hsun Ma, Department of Sociology, National Taipei University, New Taipei City, Taiwan

© 2023 Elsevier Ltd. All rights reserved.

A glance at the facts	157
Inequality of digital access	157
Inequality of digital use and online skills	158
Theories of the digital divide	158
Different types of digital divide	158
Digital cultural capital	159
Highlights from recent empirical research	160
Issues associated with ICT use at home	160
Problems associated with ICT use at school	161
The role of education in bridging the digital divide	162
References	162

Information and communication technology (ICT) has dramatically changed the educational landscape in recent decades, creating a new form of inequality—the digital divide. In the wake of this inequality, a number of scholars have well documented the causes and consequences of the digital divide. Educators and parents seek to understand how differentiated access and use of ICT influences students' daily lives. At the national and international levels, a number of efforts have been made with respect to minding the digital gap. Although digital inequality has become a worldwide phenomenon that affects every student around the globe, there is substantial cross-national variation in the way ICT is integrated into education as well as how the use of ICT generates educational inequality. In this review work, I integrate findings based on recent international reports, empirical research, review papers, and theoretical articles to present a renewed look at the evolving digital divide. Taking a cross-national comparative approach, I also explain how ICT generates a new form of educational inequality that has become a common problem around the globe.

A glance at the facts

Inequality of digital access

Globally, there remains substantial variation in the digital access divide. According to a report by the International Telecommunication Union (ITU, 2018), in 2018 more than 80% of households among economically developed countries had access to a computer and the internet at home. The number went down to less than 40% among developing countries. In contrast, less than 10% of people living in the least economically developed regions of the world had a computer or an internet connection in their homes. Since the coronavirus pandemic in 2020, the inequality between digital haves and have-nots has apparently worsened. And this form of inequality also has dramatic impact in highly developed and affluent countries. For example, in Korea, more than 200,000 students reported they did not have proper digital devices or internet access which were required to participate in online/remote education when schools were locked down (Bicker, 2020). The same problem has occurred among other affluent countries such as the United States, where many underprivileged or working class children faced obstacles to online education (Vogels et al., 2020). The situation was even worse among poor or less-developed countries, where their ICT infrastructure was not matured and educational technology was not prevalent, which made remote education or e-learning almost impossible when schools were closed due to the surge of coronavirus cases (Vegas, 2020). In summary, the intersection between the digital divide and the inequality of educational opportunities has worsened since the COVID-19 outbreak.

The engagement of digital learning at home has profound impact on students' reading achievement (Dynarski et al., 2007; Kuhfeld and Tarasawa, 2020). Students with limited to no digital access at home have few opportunities in digital learning; the more available ICT resources at home, the more likely that they can apply them in the learning process. On one hand, there is a substantial "global digital learning (or e-reading) divide" that appears between countries. A report based on the 2016 Progress in International Reading Literacy Study (PIRLS) data suggests that more than 40% of fourth grade students living in Scandinavian countries (Norway 58%, Finland 53%, Denmark 49%, and Sweden 42%) had high-level digital access at home; specifically, this means that these students had seven or more different ICT devices in the household that can be used for reading and academic learning. On the contrary, less than 5% of fourth graders in several less-developed countries had high-level digital access in the household. Among these nations, at least a third of pupils reported that they did not have a computer or internet access at all (Morocco 47%, Azerbaijan 34%, South Africa 33%, Egypt 30%, and Iran 23%; IEA, 2017). On the other hand, the digital learning divide is also prevalent within countries. The 2018 Program for International Student Assessment (PISA), which focuses on 15-year-old pupils, reported that, nowadays, more than nine out of 10 students are able to use the internet at home. And from 2012 to 2018, their hours on the internet outside of school lessons increased from 18 h per week to 27 h per week (OECD, 2019b). However, it is important to

note that there are large variations in the speed of internet connection and the kinds of digital devices or ICT appliances that they mainly use when surfing the internet. Until recently, however, more than three out of 10 students from socio-economically disadvantaged households did not have a computer that can be properly used for schoolwork in their households (OECD, 2019a). This indicates that they might need rely on mobile phones or other cheap devices to browse the internet and/or to complete school assignments.

Within countries, the inequality of ICT access is also pronounced, and it has become a critical issue in tertiary education, since college students oftentimes rely heavily on ICT to work on various research projects and out-of-class assignments. In the United States, for example, students who are socioeconomically underprivileged or identified as racial minorities experience hardships as they rely on poorly functioning laptops for schoolwork; they often experience technology problems such as unreliable internet connections, data limits, and broken hardware (Gonzales et al., 2020).

At the school site, the availability of ICT access in the classroom has improved a lot over the past decades (UNESCO, 2015), but there remain sizable cross-national variations in the quantity and quality of ICT provision and the availability e-learning resources at school (Ma, 2021). In 2017, for example, a great majority of primary students in New Zealand (93%), Denmark (92%), Netherlands (87%), and Sweden (83%) were able to use computers in reading lessons. In contrast, only a small proportion of students in Trinidad and Tobago (12%), South Africa (8%), the France area of Belgium (7%), and Morocco (6%) used computers in a reading course. Among most countries around the world (except for a few such as Georgia, 60%, and Denmark, 38%), no more than 30% of primary schoolers were able to attend schools where one-computer-per-student is universally available in classrooms (IEA, 2017).

Inequality of digital use and online skills

For most countries of the world, there is still a lot of room for growth in the development of school curriculums that aim to train students in digital skills and online literacy (ITU, 2018; van Deursen and van Dijk, 2014a). Most teachers rarely or only occasionally require students to engage in computer-based activities in core subject courses such as reading and mathematics. Taking primary education as an example, less than a quarter of students across the globe are engaged in ICT-based activities in reading courses on a weekly basis. The exceptions are Australia, Denmark, Georgia, Israel, and New Zealand, where a larger proportion of students are able to engage in activities such as reading digital texts, conducting research, and looking up information throughout the internet; also, their teachers are more likely to teach strategies for digital reading (IEA, 2017). It is also noteworthy that a growing number of students living in developed societies are mainly engaged in digital learning *outside* of schools (van Deursen and van Dijk, 2014a)—either because schools have not provided adequate ICT training courses or because there have been more e-learning opportunities at home.

ICT can be used for a variety of purposes. And it is widely believed that ICT can enrich lives. But without possessing proper ICT skills and having related online usage experiences, it will be difficult for students to perceive the benefits of ICT. It is also apparent that when online activities get more complicated, fewer people are able to participate in these activities. Within countries, educational status is positively associated with the level of digital skill (Hargittai and Hinnant, 2008). Cross-nationally, people living in more-developed countries possess more basic and advanced digital-related skills than those from less-developed countries (ITU, 2018).

For children and adolescents, the level of the digital divide in ICT-related skills is substantial, which appears both within- and between-countries (ITU, 2018; Ma et al., 2019). Even among countries with similar economic and social backgrounds, the gap in ICT skills exists. Taking eight European countries (France, Germany, Italy, Netherlands, Poland, Spain, Sweden, and United Kingdom) as an example, Dutch children have higher levels of online operational skills but lower levels of online social skills compared to their peers from the other seven nations. Both children from Spain and Italy have relatively low levels of online skills with regard to info-navigation activities and content creation activities (ITU, 2018). A study based on 55 countries suggested that increasing national income is associated with an increase of the internet literacy among teenage students. But more interestingly, national income predicts the size of the digital divide; that is: The inequality in internet literacy between upper-class and lower-class students is greater in its magnitude in poor countries than rich countries (Ma et al., 2019).

In summary, it is equally important to study the digital divide in the access, use, and skills of ICT within countries and between countries, particularly focusing on the differences between more- and less-developed countries. But due to the lack of data collection, it is quite difficult to estimate the level of inequality in ICT skills among people living in the poor and least developed regions of the world (UNESCO-UIS, 2016).

Theories of the digital divide

Different types of digital divide

For at least two decades, a burgeoning literature has well documented the explanations and patterns of the digital divide. In the early 2000s, scholars were more concerned with the inequality of digital access, namely the first-level digital divide (Davison and Cotten, 2003). They focused on the disparities in access to computers and the internet that are associated with economic, social, cultural, and geographical conditions (for instance, see DiMaggio et al., 2004). The main objectives to bridge this form of digital divide are to institutionalize educational policies and initiatives to expand the coverage and access to ICT infrastructure within school systems

(Attewell, 2001; Warschauer et al., 2004). In all, this group of research focuses on digital inequality between haves and have-nots, with an implicit assumption that the more connected to the digital world people are, the better their lives will be (Beuermann et al., 2013).

According to scholarship on the second-level digital divide, it is more important to examine differentiated digital use and ICT-related skills, even if the equality of digital access has been achieved (Hargittai, 2002). The basic thesis under this line of research is that even if ICT devices such as computers and digital tablets become equally accessible to all groups of people, regardless their socio-demographic characteristics, there remains persistent inequality in terms of how people use technology and how much ICT-related skill they already possess. Among developed and high internet-penetration countries, research suggests that those who are more educated are more likely to use the internet in more beneficial ways, which help them accumulate human and social capitals (Büchi et al., 2015; Hargittai and Hinnant, 2008). On the other hand, the less-educated are more likely to participate in online social media activities and gaming that are time-consuming in nature (van Deursen and van Dijk, 2014b). These accounts align with the knowledge gap hypothesis (Tichenor et al., 1970), which states that those who are more educated derive more knowledge from the use of mass media such as radio, television, newspapers and the internet. Finally, it is possible that the social (re) production of digital inequality is a sequential process. This means that those who are not able to engage in a particular way online are less likely to be involved in other ways, and those lacking a certain type of digital skill are also short of another type of skill (van Deursen et al., 2017).

More recently, scholars have highlighted the need to examine both the first-level and the second-level digital divides (van Deursen and van Dijk, 2019; Gonzales, 2015). They claim that even among economically developed and high internet-penetration countries, new ICT devices and appliances keep emerging every day. People who are socioeconomically underprivileged are lagging behind in the access and use of the newest models of digital devices/appliances compared to those who are privileged. Van Deursen and van Dijk (2019) recommended focusing on differences in material access, such as computers (e.g., desktops, tablets, and smart TVs), software (e.g., online apps and web subscriptions), and peripheral equipment (e.g., scanners, printers, additional monitors, and external hard drives). They suggested examining material access inequality in three aspects: (1) differences in device opportunities, including the availability of using other devices with different functions and capacities as replacements, (2) differences in the diversity of ICT devices and peripherals in the household; and (3) variation in the maintenance costs of devices and peripherals as well as the monthly cost of subscribing to high-speed broadband internet service. Even though internet access is more available even to poor and working class people, they tend to rely on less sophisticated or low-end ICT devices to surf the internet (Gonzales et al., 2020; Rideout and Katz, 2016).

Moreover, the use of smartphones may not reduce but aggravate the existing problem of digital inequality. Smartphone use with unlimited data plan increases the likelihood that one uses the internet autonomously and engages in different online activities; it is also associated with greater communication competence in terms of networking, instrumental, and creative skills (Lee et al., 2015). However, we should note that smartphones may not be beneficial or productive for everyone; instead, only those who are able to access wired or high-speed internet from highly functional devices (e.g., desktops and laptops)—when they are at home or in the workplace—are able to accrue advantages from smartphone use.

A review work by Marler (2018) concluded that even in developed countries such as the US, those who are socially, economically, and racially underprivileged are more likely to rely solely on smartphones to access the internet. Due to the limited functionality of smartphones, they may experience barriers while accessing particular websites and online services that are not suitable for mobile phone users (e.g., job application websites). This problem becomes more severe in the developing world, such as Africa, where there is underdeveloped internet infrastructure, leading to the outcome that a large proportion of Africans can only access the internet via mobile phones (Donner et al., 2011). Taken together, because digital technology is changing and developing rapidly, the digital divide as evidenced by differentiated digital use and unequal material access in ICT devices remains crucial across the globe, even among affluent countries that have high internet penetration rates.

Moreover, recent scholarship has called attention to the third-level digital divide that is concerned with variation in tangible benefits of digital/online usage (van Deursen and van Dijk, 2019). In education, this suggests that the outcomes of digital learning may differ across students with various sociodemographic profiles, thereby reinforcing existing educational inequalities. The following section(s) will share insights into understanding the process of how this form of inequality is produced.

Digital cultural capital

Bourdieu (1984) explained how elite and upper-class parents pass on their class privilege to children through “capital.” The capital is presented as different forms of *assets*—such as economic, social, cultural, and technical assets—that can be stored, accumulated, and transmitted within fields, thereby making social differentiation and inequality possible to exist/persist. As Halford and Savage (2010, p. 945) explained, this analytical approach:

allow us to explore how technical skills might allow the accumulation of other capacities—for instance, in terms of reputation among peers, and also the potential to convert these skills into other areas—for instance, in [educational fields], the labor market, or political participation.

Building upon Bourdieu's conceptualization of capital, Ollier-Malaterre et al. (2019) introduced the concept of digital cultural capital. This concept can help us gain new insight into how digital technology (re)generates inequality in the educational landscape due to the following reasons. First, digital cultural capital is about how "individuals perform to gain control over technology and its associated social norms to align their use of technology with their values and goals" (Ollier-Malaterre et al., 2019, p. 432), which is called as technology management. Note that a combination of awareness, motivation, and skill is essential to the performance of technology management. Therefore, it is important that children and adolescents are aware of how to properly use technology, but not let technology control them. They need to know the consequences if they immerse themselves into the digital world—either to enrich their lives or to bring more troubles. They also need both motivation and skills to monitor their online activities, so that they are able to extend their learning from school to home while connecting to the internet, discerning the quality and reliability of web content, and knowing from whom to seek help (e.g., teachers, peers, family members) when they are in trouble.

Researchers like Ollier-Malaterre et al. (2019) argued that the awareness, motivation, and skill that are necessary for technology management generate class inequality. When parents transmit digital cultural capital to the next generation, it is likely that children of privileged social class backgrounds end up developing a "digital habitus" that gives them clear advantages over others. For example, parents who are highly educated and affluent know how to help children to sign up for ICT-related activities that can enhance their learning experiences and educational performance; in addition, they also know how to teach their children to avoid certain online usage behavior or patterns that may distract them from learning or put themselves in danger. Therefore, it is important for families to provide an atmosphere to help children learn how to utilize ICT devices in a safe and smart way, and assist them to avoid negative impact from improper digital/internet use (Forkosh-Baruch and Erstad, 2018).

A growing body of empirical studies have applied the concept of cultural capital into their research on digital inequality. Inspired by parental mediation theory, Yuen et al. (2018) referred to cultural capital as the variation in parenting styles, values, and practices pertinent to digital technology. They examined different types of parents' strategies in how to regulate children's online behavior as well as their overall ICT use at home. Highly educated and affluent parents were more likely to employ the kind of parenting strategy to help accumulate ICT skill that can be transformed into digital cultural capital, leading their children to receive better educational and occupational outcomes in the future. Yuen et al. explained that when parents are able to transmit digital cultural capital to their offspring through successful parenting styles and regulations, their children tend to become the "celebrating" users, indicating those who find ICT enjoyable and are able to use digital technology in an effective and productive way. In contrast, students without cultural capital at home tend to become the "copying" users or the "struggling" users; the former refers to those who make conscious efforts to use ICT for academic or productive purposes but fail to do so, whereas the latter indicates those who keep having negative experience of ICT use because of basic problems including poor digital access and lack of basic ICT-related competence. To understand how ICT interacts with students and teachers in school settings, Paino and Renzulli (2013) referred to digital cultural capital as certain ICT-related attributes of students that contribute to teachers' expectations. That is, teachers have greater educational expectations for students with advanced computer competency, despite their actual levels of academic performance in the classroom. This suggests that ICT ownership and its related skill become a signal that can be used by teachers to evaluate or judge students.

Highlights from recent empirical research

Issues associated with ICT use at home

In less-developed societies, public spaces such as schools, libraries, and community centers become the main places where students learn basic ICT competencies such as operational and formal skills. This is quite different from developed societies, where students tend to receive these skills at home through self-learning or by receiving help from family members (van Deursen and van Dijk, 2014a). This prompts the question of what home environment is the best to enhance children's e-learning process.

Socioeconomic status (SES) seems to be one of the most influential factors in determining the quantity and quality of ICT use. Earlier studies found a disadvantage of ICT access at home among students of racial minority backgrounds and those from single-parent households (DiMaggio et al., 2004). But when accounting for the impact of SES, the digital access gap by race and family type becomes relatively smaller in size. This suggests that parents' educational, income, and occupational statuses play a great part in shaping students' experience with ICT at home (Scheerder et al., 2019). Based on the extant research, the digital divide along the socioeconomic line persists even among developed countries [e.g., Australia (Smith et al., 2013); the Netherlands (Goedhart et al., 2019); the United Kingdom (Helsper and Reisdorf, 2016); and the United States (Gonzales, 2015)]. This literature suggests that students of higher SES have greater incentives and are more motivated to engage in online learning activities; in contrast, students of low SES are more likely to use technology with respect to social networking and entertainment. Until recently, a great proportion of low-income students in developed societies lived in an under-connected household because of slow or mobile-online internet access, intermittent disconnection, and poor hardware (Gonzales et al., 2020).

The rise of globalization and the knowledge economy can prompt feelings of increased competition among students. In elite and middle-class families, parents are able to mobilize resources to ensure their children always stay a few steps ahead in the competition for excellence in educational performance (e.g., earning a high grade in class, receiving academic rewards, getting admission to an elite university, and enhancing international learning experience) (Brown, 2013). To help their children succeed in the future, these parents may arrange and pay for additional out-of-school activities both offline and online. In Korea, for example, Park et al. (2016) found that there is a growing demand for online private tutoring services. Teachers with great teaching skills are hired; more

than hundreds of e-courses are offered, and each class is tailored to meet students' needs with an aim to help them top the national college entrance exam. These online learning services may further lead to the growth of educational inequality insofar as only the elite and the middle-class are able to afford this e-form of private education (Choe, 2009).

There is no doubt that all children and adolescents want to succeed in education. As e-learning becomes more important, they all know that ICT can—more or less—help facilitate their learning process. But a question arises as to why the use of ICT is more beneficial to students of higher-SES backgrounds than their lower-SES peers? A qualitative study of Dutch families conducted by Scheerder et al. (2019) gives an insightful glimpse into understanding the differences along the socioeconomic line: First, students of higher-SES families began using new technology in an early age. They described a strong interest in learning ICT as they often had opportunities to participate in computer clubs and see how their parents worked at home through online networked computers. Second, not only did parents in higher-SES families take an active role in helping children learn how to use ICT, but children who were believed to be “digitally savvy” also assisted their parents. It was common that members in these families mutually helped each other to solve problems they encountered with technology. Comparatively, due to the lack of time and know-how, lower-SES parents rarely helped their kids. Third, when buying new ICT devices, members of higher-SES families were more concerned with the functionalities of the device like processor speed and computer memory. They intended to integrate different devices to add value, save time, and make their life more convenient. Also, they developed habits of using ICT for professional and educational purposes. In contrast, lower-SES families were only concerned with the price and cost when they were buying new ICT devices. In most occasions, they used ICT for maintaining social contacts. Similarly, a research on Brazilian young people demonstrates that parents who are more educated are more likely to apply active technical mediation as a way to coach their children, which is effective in promoting children's digital skills (Cabello-Hutt et al., 2018).

In sum, socioeconomically advantaged people are more likely to explore potential benefits accrued from the internet. Additionally, they are also more conscious with regard to the validity and reliability of the information received from the web. In his qualitative study of British adolescents, Davies (2018) found that working class students relied more on Google to seek quick answers. They were less likely to verify the sources of the information behind the web, which is quite different from other upper class students who have greater motivations to think analytically and stay alert to the information provided on the web.

Problems associated with ICT use at school

As noted above, some would argue that schools in developed societies are not the primary place to accumulate a wide range of computer-related skills (e.g., van Deursen and van Dijk, 2014a). Nevertheless, how schools generate digital learning inequality remains a notable issue even in the most affluent regions of the world.

There are at least three main explanations that are helpful to address the digital divide at the school-level. First, school resource explanations contend that different access to material resources (e.g., classroom equipment and school funds) and non-material resources (e.g., teacher credentials and support from professional staff) are associated with differentiated e-learning opportunities. In the US, for example, racial minority and poor students are more likely to attend economically disadvantaged and underperforming schools that are unable to provide enough computers or digital literacy courses (Robinson, 2014). Note that these schools encounter basic problems that are fundamental to keep themselves functioning, such as the lack of eligible teachers, staff shortages, and dilapidated classrooms. A great proportion of students are underperforming in basic subjects such as reading and math. It is likely that the average ICT capacity of teachers is low (Warschauer, 2016). Therefore, to put a massive amount of resources into equipping ICT-related devices (e.g., desktops, screens, and projectors) may result in the misallocation of educational resources, because the cost of ICT is drained from other parts of the budget that are vital to run the school (Natriello, 2001).

The second explanation highlights how the cultural and institutional context of schools influences the attitudes and behavior of individuals, even if variations in school resources have been controlled. This generates cultural distinctions along the line of social class. Motivated by social reproduction theory, this research examines how teachers treat working-class students differently from other middle-class peers (Bowles and Gintis, 2011). In schools with a great majority of poor and racial minority kids, teachers are more likely to undervalue students' competencies and have lower educational expectations for them. In contrast, teachers in elite and high-SES schools are more likely to help students to learn in an interactive and creative way and encourage them to engage in self-learning in classrooms (Jack, 2016). This may influence how students value the role of ICT in learning. In his recent qualitative work of American high-schools, Rafalow (2018) suggested that teachers in resource-rich and elite schools have greater ICT competence. More importantly, teachers perceive e-learning to be a valuable part of education and encourage students to “pursue their interest-driven digital play during class” (p. 1446). This is quite different from other schools with a concentration of poor and minority students, where teachers see digital play to be irrelevant to learning or a threat to their authority. Taken together, the socioeconomic composition at the school-level influences the opportunities for digital learning at school (Ma, 2021). In line with this view, Leu et al. (2015) suggested that when schools are located in affluent neighborhoods, the average of students' online research and comprehension skill is at least 1 year ahead of other students living in mid- or low-income neighborhoods.

Third, it is not enough to merely consider the binary division between digital haves and have-nots. Rather, Halford and Savage (2010) argued that we should focus on how digital technologies “are associated with the crystallization of social relations of different kinds” (p. 950), thereby creating social-technical networks that are durable, enduring, and consisted of both human and non-human actors. According to the actor network theory (ANT), our living social world is inseparable from the material world (Latour, 1990). Law and Singleton (2000, p. 774) contended that “technologies, knowledges and working may be understood as the effects of materially, socially, and conceptually hybrid performances. In these performances different elements assemble together

and act in certain ways to produce specific consequences.” Using the notion of “digital pedagogics,” Alirezabeigi et al. (2020) explained that how schools are operated and classes are arranged (for instance, existing school hierarchies, school policies, and the adoption of materials in classrooms) have been entirely rearranged and reconstructed due to the emergence of digital pedagogics. That is, the integration of ICT in education may reshape “all aspects of school and not only moments when actors are engaged with learning” (Alirezabeigi et al., 2020, p. 203). It also involves “a complex configuration of clouds, software and interfaces, algorithms and patterns, standards, protocols and negotiations” (p. 193), and creating networks between human actors (e.g., school staff, IT service providers, school principals, teachers, and students) and non-human actors. To better understand how this new form of digital pedagogics (or e-learning in school settings) is used—which becomes opportunities to some but obstacles to others—Fluckiger et al. (2017) argued for the need to adopt a “subjectivist” approach that can better account for teachers’ experiences, strategies, attitudes, and skills pertaining to their adoption of ICT technologies in classrooms, such as interactive whiteboards.

Taken together, the aforementioned accounts have clearly pointed out that the mission to bridge the digital learning divide in schools can be a complicated task, which is much more than just improving the digital infrastructure or computer/internet access in schools. Past efforts that simply focus on “bring your own device” (BYOD; Alirezabeigi et al., 2020) or “one laptop per child” (OLPC; Ames, 2019) is likely to result in a failure to bridge the digital divide.

The role of education in bridging the digital divide

New technology is changing the way in which students live. Since the COVID-19 pandemic, it is also changing how students are educated when they are asked to navigate online educational material, discuss homework with others on social media sites, search for information, consult professionals, and interact with teachers throughout the internet. This raises several questions: Can school play a crucial role in the promotion of e-learning opportunities? What is the role of education in bridging the digital learning gap?

Despite this ubiquity of ICT in students’ everyday lives, there seems to be a missing link between sociological research on stratification and inequality and research on the digital divide; that is, to date we still know little about how individuals’ technical capacities, affordances, and their experiences with the use of ICT have profound impact on their current and future life outcomes (Halford and Savage, 2010).

As noted earlier, in developed countries, students may learn more digital skills at home compared to what they learn from schools (van Deursen and van Dijk, 2014a; Robinson, 2014). But this does not mean that schools should not take precedence over other places in the provision of ICT training and e-learning opportunities. Van Deursen and van Dijk (2014a) contended that formal education remains necessary for learning appropriate ICT-related skills. First, the digital divide between students with different socioeconomic backgrounds remains significant. This shows that lower-income and minority students have fewer digital learning resources/opportunities outside of schools; compared to others, they also receive fewer benefits from the use of ICT at home. Therefore, schools may become the only place for them to gain basic ICT competence. Second, some types of ICT skills (e.g., operational skill and formal skill) are basic and can be learned from home by trial and error or practicing. But other forms of ICT skills (e.g., online information skill and strategic online skill) are not suitable to learn from home, because they are more complicated, require a number of steps, and need a considerable amount of time to learn. Without a clear guidance and help from teachers, students may easily feel frustrated and distressed when they are engaged in difficult tasks on the internet.

Noteworthy, findings based on Eurostat’s Community survey on ICT usage in households and by individuals suggest that formal education provides insufficient support as a means of learning and acquiring ICT-related skills. In 2011, less than one-third of young people and adults learned ICT-related skills via formal education. Instead, most of them relied on self-studying or sought assistance from proximate people like friends and family members (van Deursen and van Dijk, 2014a). Aside from Europe, it would be reasonable to argue that this finding can be applied to other parts of the world. This clearly indicates that most schools nowadays have not put enough effort in the investment of digital technology as well as the promotion of e-learning opportunities in classrooms.

Even before the onset of COVID-19, digital access had already become a critical issue for the global south as well as the developing world (Henwood and Wyatt, 2019). In Paraguay, for example, teachers were more likely to face a great deal of technical difficulties that were associated with ICT use in classrooms, such as compute breakage, missing software, operating difficulties, limited teacher agency, and student disinterest (Ames, 2019). Because digital inequality is more severe in poor and developing countries (Ma et al., 2019), there is an urgent need to provide appropriate ICT learning opportunities in schools and to incorporate e-learning into the curriculum. But in the former section, I have suggested that how to better address the problem of (or bridge) the digital divide in school environment seems to be more complicated than what people used to think (see **Problems associated with ICT use at school** section). Last but not least, we should not overlook the gender digital divide, to the extent that gender oppressions and inequalities are reproduced or even become exacerbated in digital spaces, even when men and women have equal access to digital technologies (Henwood and Wyatt, 2019).

References

- Alirezabeigi, S., et al., 2020. Investigating digital doings through breakdowns: a sociomaterial ethnography of a Bring Your Own Device school. *Learn. Media Technol.* 45 (2), 193–207.
- Ames, M.G., 2019. *The Charisma Machine: The Life, Death, and Legacy of One Laptop Per Child*. The MIT Press, Cambridge, MA.

- Attewell, P., 2001. The first and second digital divides. *Sociol. Educ.* 74 (3), 252–259.
- Beuermann, D., et al., 2013. Home Computers and Child Outcomes: Short-Term Impacts From a Randomized Experiment in Peru. NBER Working Paper Series. Available from: <http://www.nber.org/papers/w18818>.
- Bicker, L., 2020. Coronavirus: How South Korea is Teaching Empty Classrooms. BBC News, 10 April. Available from: <https://www.bbc.com/news/world-asia-52230371>. (Accessed 24 October 2020).
- Bourdieu, P., 1984. *Distinction: A Social Critique of the Judgement in Taste*. Harvard University Press, Cambridge, MA.
- Bowles, S., Gintis, H., 2011. Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life, reprint ed. Haymarket Books, Chicago, IL.
- Brown, P., 2013. Education, opportunity and the prospects for social mobility. *Br. J. Sociol. Educ.* 34 (5–6), 678–700.
- Büchi, M., et al., 2015. Modeling the second-level digital divide: a five-country study of social differences in Internet use. *New Media Soc.* 18 (11), 2703–2722.
- Cabello-Hutt, T., et al., 2018. Online opportunities and risks for children and adolescents: the role of digital skills, age, gender and parental mediation in Brazil. *New Media Soc.* 20 (7), 2411–2431.
- Choe, S.-H., 2009. Tech company helps South Korean students ace entrance tests. *N. Y. Times*, 1 June. Available from: https://www.nytimes.com/2009/06/02/business/global/02cram.html?_r=0.
- Davies, H.C., 2018. Learning to Google: understanding classed and gendered practices when young people use the Internet for research. *New Media Soc.* 20 (8), 2764–2780.
- Davison, E., Cotten, S., 2003. Connection discrepancies: unmasking further layers of the digital divide. *First Monday* 8 (3).
- DiMaggio, P., et al., 2004. Digital inequality: from unequal access to differentiated use. In: Neckerman, K. (Ed.), *Social Inequality*. Russell Sage Foundation, New York, pp. 355–400.
- Donner, J., et al., 2011. Exploring mobile-only internet use: results of a training study in urban South Africa. *Int. J. Commun.* 5, 574–597.
- Dynarski, M., et al., 2007. Effectiveness of Reading and Mathematics Software Products: Findings From the First Student Cohort. US Department of Education, Institute of Education Sciences. Available from: <https://files.eric.ed.gov/fulltext/ED496015.pdf>. (Accessed 15 June 2021).
- Fluckiger, C., et al., 2017. What role does the textbook play in the era of the Interactive Whiteboard? A didactic approach to teachers' professional writings. *J. Subject Didactics* 1 (2), 101–117.
- Forkosh Baruch, A., Erstad, O., 2018. Upbringing in a digital world: opportunities and possibilities. *Technol. Knowl. Learn.* 23 (3), 377–390.
- Goedhart, N.S., et al., 2019. "Just having a computer doesn't make sense": the digital divide from the perspective of mothers with a low socio-economic position. *New Media Soc.* 21 (11–12), 2347–2365.
- Gonzales, A.L., et al., 2020. Technology problems and student achievement gaps: a validation and extension of the technology maintenance construct. *Commun. Res.* 47 (5), 750–770.
- Gonzales, A., 2015. The contemporary US digital divide: from initial access to technology maintenance. *Inf. Commun. Soc.* 19 (2), 234–248.
- Halford, S., Savage, M., 2010. Reconceptualizing digital social inequality. *Inf. Commun. Soc.* 13 (7), 937–955.
- Hargittai, E., Hinnant, A., 2008. Digital inequality: differences in young adults' use of the Internet. *Commun. Res.* 35 (5), 602–621.
- Hargittai, E., 2002. Second-level digital divide: differences in people's online skills. *First Monday* 7 (4). <https://doi.org/10.5210/fm.v7i4.942>.
- Helsper, E.J., Reisdorf, B.C., 2016. The emergence of a "digital underclass" in Great Britain and Sweden: changing reasons for digital exclusion. *New Media Soc.* <https://doi.org/10.1177/1461444816634676>.
- Henwood, F., Wyatt, S., 2019. Technology and in/equality, questioning the information society: (almost) 20 years later. *Digit. Cult. Soc.* 5 (1), 183–194.
- IEA, 2017. PIRLS 2016: International Results in Reading. TIMSS & PIRLS International Study Center, Chestnut Hill, MA. <https://doi.org/10.1787/9789264250246-en>. Available from.
- ITU, 2018. *Measuring the Information Society*. ITU, Geneva.
- Jack, A.A., 2016. (No) harm in asking class, acquired cultural capital, and academic engagement at an elite university. *Sociol. Educ.* 89 (1), 1–19.
- Kuhfeld, M., Tarasawa, B., 2020. The COVID-19 Slide: What Summer Learning Loss Can Tell Us About the Potential Impact of School Closures on Student Academic Achievement. NWEA. Available from: <https://eric.ed.gov/?id=ED609141>. (Accessed 15 June 2021).
- Latour, B., 1990. Technology is society made durable. *Socio. Rev.* 38, 103–131.
- Law, J., Singleton, V., 2000. Performing technology's stories: on social constructivism, performance, and performativity. *Technol. Cult.* 41 (4), 765–775.
- Lee, H., et al., 2015. A new dimension of the digital divide: exploring the relationship between broadband connection, smartphone use and communication competence. *Telematics Inf.* 32 (1), 45–56.
- Leu, D.J., et al., 2015. The new literacies of online research and comprehension: rethinking the reading achievement gap. *Read. Res. Q.* 50 (1), 37–59.
- Ma, J.K.-H., et al., 2019. National income, political freedom, and investments in R&D and education: a comparative analysis of the second digital divide among 15-year-old students. *Soc. Indic. Res.* 144 (1), 133–166.
- Ma, J.K.-H., 2021. The digital divide at school and at home: a comparison between schools by socioeconomic level across 47 countries. *Int. J. Comp. Sociol.* 62 (2), 115–140.
- Marler, W., 2018. Mobile phones and inequality: findings, trends, and future directions. *New Media Soc.* 20 (9).
- Natriello, G., 2001. Bridging the second digital divide: what can sociologists of education contribute? *Sociol. Educ.* 74 (3), 260–265.
- OECD, 2019a. PISA 2018 Results (Volume II): Where All Students Can Succeed. PISA, OECD Publishing, Paris. <https://doi.org/10.1787/b5fd1b8f-en>. Available from.
- OECD, 2019b. PISA 2018 Results (Volume III): What School Life Means for Students' Lives. PISA, OECD Publishing, Paris. <https://doi.org/10.1787/acd78851-en>. Available from.
- Ollier-Malaterre, A., et al., 2019. Technology, work, and family: digital cultural capital and boundary management. *Annu. Rev. Sociol.* 45 (1), 425–447.
- Paino, M., Renzulli, L.A., 2013. Digital dimension of cultural capital: the (in)visible advantages for students who exhibit computer skills. *Sociol. Educ.* 86 (2), 124–138.
- Park, H., et al., 2016. Learning beyond the school walls: trends and implications. *Annu. Rev. Sociol.* 42, 231–252.
- Rafalow, M.H., 2018. Disciplining play: digital youth culture as capital at school. *Am. J. Sociol.* 123 (5), 1416–1452.
- Rideout, V.J., Katz, V.S., 2016. Opportunity for All? Technology and Learning in Lower-Income Families. A Report of the Families and Media Project. The Joan Ganz Cooney Center at Sesame Workshop, New York, NY. Available from: <https://kramden.org/technology-learning-lower-income/>.
- Robinson, L., 2014. Freeways, detours, and dead ends: search journeys among disadvantaged youth. *New Media Soc.* 16 (2), 234–251.
- Scheerder, A.J., et al., 2019. Internet use in the home: digital inequality from a domestication perspective. *New Media Soc.* 21 (10), 2099–2118.
- Smith, J., et al., 2013. Beneath the "digital native" myth: understanding young Australians' online time use. *J. Sociol.* 49 (1), 97–118.
- Tichenor, P.J., et al., 1970. Mass media flow and differential growth in knowledge. *Publ. Opin. Q.* 34 (2), 159–170.
- UNESCO, 2015. Education for All 2000–2015: Achievements and Challenges. UNESCO, Paris. Available from: <https://en.unesco.org/gem-report/report/2015/education-all-2000-2015-achievements-and-challenges#sthash.bfMFAV.dpbs>.
- UNESCO-UIS, 2016. Paper Commissioned for the Global Education Monitoring Report 2016, Education for People and Planet: Creating Sustainable Futures for All. Available from: <http://unesdoc.unesco.org/images/0024/002455/245572e.pdf>.
- van Deursen, A.J., van Dijk, J.A., 2014a. Digital Skills: Unlocking the Information Society, 2014th ed. Palgrave Macmillan, New York, NY.
- van Deursen, A.J., van Dijk, J.A., 2014b. The digital divide shifts to differences in usage. *New Media Soc.* 16 (3), 507–526.
- van Deursen, A.J., van Dijk, J.A., 2019. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media Soc.* 21 (2), 354–375.
- van Deursen, A.J., et al., 2017. The Compoundness and sequentiality of digital inequality. *Int. J. Commun.* 11, 452–473.

- Vegas, E., 2020. School Closures, Government Responses, and Learning Inequality Around the World during COVID-19. The Brookings Institution, 14 April. Available from. <https://www.brookings.edu/research/school-closures-government-responses-and-learning-inequality-around-the-world-during-covid-19/>. (Accessed 20 September 2020).
- Vogels, E.A., et al., 2020. 53% of Americans Say the Internet Has Been Essential During the COVID-19 Outbreak. Pew Research Center, 30 April. Available from. <https://www.pewresearch.org/internet/2020/04/30/53-of-americans-say-the-internet-has-been-essential-during-the-covid-19-outbreak/>. (Accessed 23 October 2020).
- Warschauer, M., et al., 2004. Technology and equity in schooling: deconstructing the digital divide. *Educ. Pol.* 18 (4), 562–588.
- Warschauer, M., 2016. Addressing the social envelope: education and the digital divide. In: Greenhow, C., et al. (Eds.), *Education and Social Media: Toward a Digital Future*. MIT Press, pp. 29–48.
- Yuen, A.H., et al., 2018. The significance of cultural capital and parental mediation for digital inequity. *New Media Soc.* 20 (2), 599–617.

Global education reform movement revisited

Pasi Sahlberg, Southern Cross University, Lismore, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Road to the global education race	165
Tentative typology of 20th century whole system school reforms	166
The arrival of the GERM as an alternative typology	168
Increasing competition between schools over enrollments	169
Standardization of teaching and learning in schools	169
Increasing priority to basic literacy and numeracy	169
Test-based accountability	170
Privatization of public education	170
Antidotes to GERM and conclusion	171
References	172

Road to the global education race

If this chapter was written 20 years ago, it would have a totally different tone. Not only because of the ongoing global pandemic that has changed so many things in education, but also because of the much better understanding today of key education policy drivers for whole system success. Earlier education policies relied mostly on indicators describing what education systems do, or inputs into education systems. Today we know much more about how education systems work, how different parts of the system interact, how important some key educational outcomes are and how this all varies over time. This chapter looks back at some earlier efforts to conceptualize large-scale education reforms and then explains how the global education reform movement as a second-generation typology has evolved over the past decade from the earlier typologies. The aim of this chapter is to contribute to the ongoing debate about wrong versus right drivers for education system improvement.

Expectations for global education development were set high in the early 1990s. There was a strong common understanding among the coalition of rich countries and intergovernmental organizations, including the World Bank, OECD, and UNESCO, that fixing the world's educational challenges was primarily about finding enough money to do so rather than knowing how to do it. Consequently, increasing international development aid through "Education for All," "Millennium Development Goals," and other global campaigns promised to provide high quality and equitable education for more children around the world. Most wealthy countries saw their education budgets grow while middle-income countries were borrowing money from international banks to upgrade their education systems, and more students than ever before continued their learning in secondary and higher education. The emerging new millennium was expected to be an era of digital technologies and innovation that would transform pedagogy in schools and unleash professional potential in education systems. The future of education looked optimistic.

What was missing, however, was a good understanding of how to change whole education systems. By the year 2000, research on school effectiveness and school improvement that was essential for whole-system reforms had established a knowledge base that made turning individual ineffective schools into better performing schools. But to change all the schools in an education system was much more difficult. One challenge was that large-scale education reforms that aim to transform whole education systems are complicated and expensive to research. Another challenge was the absence of more advanced metrics that would provide comparable and comprehensive information about an education system's performance. Systematic data from genuine whole-system reforms were also rare. Examples of large-scale education reforms that have been studied as natural systemic experiments include: the school choice and voucher reforms in Chile in the 1980s and in Sweden in the 1990s (Adamson et al., 2016), the Education Reform Act 1988 in England (Ball, 2008), and the introduction of standards-based education in many parts of the United States in the 1990s and in Australia in the 2000s. These education reforms share a common assumption that if education could only just act more like a marketplace, excellence will prevail. These system-wide reforms have been extensively reviewed in education policy and research literature and they were influential in many national education reforms around the world in the 1990s.

A significant milestone in the history of international whole system education reforms was the launch of the Organization for Economic Co-operation and Development's (OECD) new standardized student assessment in 2000 called PISA, or the Program for International Student Assessment. This triennial international large-scale assessment (ILSA) is coordinated and administrated by the OECD and repeated in all its member states as well as a large number of other countries every three years testing 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges (OECD, 2001). At the time of launching PISA, radically new whole-system education reforms in Chile, Sweden, England, and in many states of the US, for example, had created high expectations among policymakers and the education expert community about whole system success in these countries. In this light the significance of PISA was three-fold.

First, the OECD required that all its member countries and new candidate countries participated in PISA in each three-year cycle. This made PISA a common mechanism in the OECD's regular economic cooperation meetings and thereby elevated the importance

of PISA in regard to national economic and social policies and strategies. Other international large-scale assessments that were coordinated by the International Association for the Evaluation of Educational Achievement (IEA) in four-year cycles were primarily used for researchers to analyze national education systems, having less influence in policy formation or national education reforms (Fischman et al., 2019). Since the beginning, the OECD has played an active role in national level policy advice and inter-governmental policy transfer using data from PISA.

Second, PISA was the first time that the entire OECD member community (in year 2000 it had 28 member states, including then 15 European Union member states) would employ the same learning metric to test the performance of their school systems. In other words, PISA was also the first standardized test to justify the impact of whole-system education reforms mentioned earlier. It also put education systems with very different structures and policies, like those in Germany, France and in the Nordic countries, to the test that would allow more detailed comparative analysis of the success or failure of different models of school education.

Third, PISA collected a broader range of background information about schools, their teachers and students that allowed better analysis of relationships between educational performance, students' family backgrounds and other social outcomes in participating countries. These new data broadened the scope of "education performance" when comparing different education systems to include different aspects like educational equality, inclusion, and fairness, in other words equity of education.

The first results of PISA in 2000 were a litmus test for the previous large-scale education reforms. In the absence of more systematic data about education systems' performance, there were education system leaders who believed they had found the key to success in transforming whole systems. Then there were those who thought that their education systems were world-class without needing any radical reforms. As the OECD's Andreas Schleicher, who has worked on PISA since the beginning recalled from a mid-1990s meeting in Paris where all OECD member-government representatives had gathered to discuss this proposed new learning metric (2018, 18): "some of them were boasting that they had the world's best school system – perhaps because it was the one they knew best." Although no one expected PISA to be the ultimate yardstick for education systems excellence, many expected that it would prove that the education policies and reform designs adopted earlier in their own countries were keeping their promise. The results of the first PISA study in 2000 launched new era in international comparative education. Sam Sellar and colleagues (2017) argue that PISA also created a new global education race and paved the way to the Global Education Reform Movement.

Tentative typology of 20th century whole system school reforms

There are various ways to conceptualize the evolution of systemwide education reforms that date back to the mid-20th century. Some scholars see the 1990s an era of comprehensive systemic education reforms around the world that happened for different reasons but with similar aims to change whole school education systems (Fullan, 2011; Hargreaves and Shirley, 2009; Hopkins et al., 2014). The collapse of socialist governments in Eastern Europe and economic globalization in the rest of world offered fresh opportunities to test and export new education policy ideas in the countries that needed rapid transformation of their outdated education systems. South Africa, the Middle East and Latin America had similar needs to modernize their education systems. Modernization often meant updating curricula and qualifications, raising the quality of student learning outcomes, narrowing the achievement gap between different social groups, integrating new digital media and technologies to serve both teaching and learning in schools, and improving efficiency of education systems as they grew bigger and became more expensive (Zajda and Rust, 2021). Many government authorities and educational reform architects were looking for places where systemwide educational change experiments were underway hoping to find inspiration to fix their own system challenges.

It was well-known during that time that education reforms rarely succeed even when significant resources were invested in implementing them. In his 1990 book "The Predictable Failure of Educational Reform" Seymour Sarason (1990) describes how educational reforms, one after another, turn into failures primarily because policymakers and system leaders are unable to learn from the past change efforts. In the 1990s many scholars shared Sarason's skepticism and concluded that the more we try to change schools, the more they seem to remain the same. School effectiveness research that was born in the 1970s in North America, Europe, and Australia improved understanding of the key characteristics of successful schools (Hopkins et al., 2014). At the same time, school improvement research began to show how to make low-performing schools more effective (Joyce & Showers, 1995). However, how to change whole education system remained an unanswered question.

One of the most influential large-scale education reforms was started by the issuance of the Education Reform Act (ERA) in 1988 in England. This new legislation set the direction for education in England for the next decades and influenced governments around the world to follow that reform model so much so that Levin and Fullan (2008) call it "a watershed event in the international educational reform movement." Neoliberal public sector policies that formed an ideological foundation for the New Public Sector Management constituted a particular approach to educational reform inspired by education as a marketplace. Free market principles had been experimented with in the education sector before (e.g., Chile), but England soon became a laboratory that caught wide international attention due to its systemwide nature (Carnoy, 1999; Levin and Fullan, 2008; Adamson et al., 2016). The Education Reform Act 1988 and the policies and practices it launched in the UK happened at the right time politically and economically in an environment that was fertile for scaling up bold new ideas to reform the entire education system.

The theory of change that was built in the ERA was simple and easy to understand by politicians and other non-educators. As a neoliberal design it surprisingly survived a decade of Tony Blair's Labor Government rule from 1997 to 2007. The following four key principles were the foundation of ERA's modus operandi (Levin and Fullan, 2008):

- The more schools compete against other schools for student enrollment, the more aligned to student needs they will have to be, thereby improving student outcomes for all as well as overall quality of education in general.
- For proper competition schools need administrative and pedagogical autonomy to operate, and parents need freedom to choose a school that they think fits best for their child.
- A common national curriculum is needed to set the same educational performance standards for all schools.
- Publicly available comparable data for parents and the public about student outcomes and performance of schools should form the basis of school accountability.

These operational premises that resonate closely with marketplace logic promised efficiency gains in many education systems just like competition does in the world of commerce. International development organizations, global consulting firms, many philanthropic organizations and education think tanks around the world adopted the ERA's reform logic as part of the evolving New Public Management paradigm (Hood and Jackson, 1991). Education reforms in many parts of the Anglo-Saxon world followed this reform often adding their own peculiarities in implementation. School competition and parental choice, aligning teaching and learning to common standards, standardized tests in literacy and numeracy, holding schools accountable for these test results to authorities, and new forms of public-private partnerships soon spread around the world, including countries in developing parts of the world. This new way of improving education systems' performance that was included in policy recommendations and education reform designs around the world soon became a "gold standard" model for national education reforms and laid a foundation for educational change in the 21st century.

In the 1990s the most influential actors in the global education arena included United States, United Kingdom, the Netherlands, and the Nordic countries as bilateral donors, and the World Bank (WB), the OECD, the European Union (EU), and the United Nations Education, Scientific and Cultural Organization (UNESCO). Their influence in national education policymaking and reform designs varied from soft advice and coordination (EU and UNESCO) to government-level policy recommendation (OECD) and conditional funding of reform programmes (WB) that required governments to enact new legislation and related reform policies. A common theory of change was to identify international best practices from more successful education systems and include them in national education policies. Some education systems in Eastern Europe, Northern Africa and the Middle East shifted focus from static knowledge to flexible skills including technological literacy and on new labor market needs whereas many other education systems were looking for more efficient ways of system management, alternative ways of financing education, and new information management systems. Although the overall goal in most education reforms was similar, how they were implemented varied greatly from one system to another.

The challenge for both borrowers of education reform ideas and lenders of solutions to improve education systems at that time was an absence of a commonly accepted framework to judge how effective different reforms have been. Instead, education reforms were categorized by looking at their aims, purposes, and orientations (Carnoy, 1999; Fullan, 2011; Hopkins et al., 2014; Sahlberg, 2004). The four commonly identified education reform categories in the 1990s were those that focused on a system's restructuring, economic efficiency, standardization of education, and enhancing equity. These are summarized in Table 1.

Table 1 Tentative typology of system-wide education reforms in the end of the 20th century.

<i>Restructuring education systems</i>	<i>Enhancing economic efficiency</i>	<i>Standardization of education process</i>	<i>Strengthening equity in education</i>
The key assumption was that high performing education systems share similar structures, purposes, and operational principles, and that global "best practices" should guide education reforms in all other countries. The key indicators of education systems productivity included pupil-teacher ratios, class-sizes, school-sizes, time-allocation per subject in different levels of schooling, education expenditure per capita, participation and completion rates, and grade repetition.	The main aim was to reduce the cost of education to taxpayers by encouraging new forms of between school competition and introducing user-fees to pay for education. These reforms became prevalent during austerity actions in early 1990s and again following the global financial downturn after 2008 fiscal crisis. Since education is a significant proportion of public sector spending, reducing public spending inevitably means also shrinking education budgets financed from public taxes. This, in turn, forces governments to seek new funding sources or to reduce the unit costs within education sector.	The starting point was a belief that an effective way to improve the quality of education is to set higher standards for teaching and learning in all schools. This was particularly relevant at the time of decentralized governance and increasing school autonomy. The most common reform strategy includes (1) a closely scripted curriculum with predetermined learning standards; (2) aligned standardized tests that measure how these standards have been achieved; (3) tougher accountability to control teachers' and schools' performance vis-à-vis the set standards; and (4) performance-related incentives and sanctions to teachers and schools.	The dual intention of was to promote social equity and increase economic opportunity simultaneously. Education is seen as an important means in closing the gap between the socio-economic groups in society. These reforms emphasize strengthening the political role of education in building democracy, social mobility and equal opportunities for all citizens. In general, these reforms focused on: shifting public spending from higher to lower levels of education, addressing special needs education, gender issues, schooling of minorities, and broadening moving toward a more integrated curriculum and inclusive organization of schooling.

National education system improvement efforts often included aspects of several of the types of reforms described in [Table 1](#). The first three types of education reforms (restructuring, economic efficiency, and standardization) were common templates in advice and technical assistance offered by international development organizations and many bilateral donors in the 1990s. Equity-oriented education policies and reforms that were common in the Nordic countries remained less popular often due to a prejudice that equity and equality of opportunity, by default, jeopardizes the system-wide efforts to produce excellent learning outcomes. Indeed, many education pundits and private school advocates pointed to the Nordic countries as valid proof that being more educationally equal happens at the expense of educational excellence.

Optimism regarding large-scale educational reforms' power to change education systems was high in the end of the 1990s as part of the Third Way political and economic movement that tried to close the gap between old divisions political left and right. Anthony Giddens who is often recognized as the father of the Third Way idea argued that governments, especially those in the UK, the US, and Germany should not rely on reform programmes that represent the interests of any one political block, but that political parties should build consensual support from the middle – the Third Way – particularly in advocating a synthesis of center-right economic reforms combined with center-left social policies. As the Third Way promised to strengthen public services, it strived to do that with a spirit of free market logic. Within this logic education is seen as a marketplace where parents have the freedom to choose their children's education. This required more competition between schools so that they better serve families' increasingly diverse expectations and needs. In practice this gave a boost to new types of schools governed in partnership between public authorities and private operators. Charter schools in the US, Academies in England, and Free Schools in Sweden are examples of new forms of publicly funded "private" schools.

Andy Hargreaves and Dennis Shirley have argued that the Third Way of educational change is pursuit by focusing performance through partnerships ([Hargreaves and Shirley, 2009](#), p. 44) that "Third Way advocates sought to address the dawning realization that education in the 21st century must move beyond the control of self-serving professionals under free-wheeling progressivism and beyond the dark thicket of prescription and standardization that limits capacity and stifles initiative." The Third Way was a form of governance built on a synchrony to serve private economic interests and enhance public good. The Third Way directions in education reforms around the world have served as birthplace for what we now know as the Global Education Reform Movement.

The arrival of the GERM as an alternative typology

It is often claimed that the launch of the OECD's PISA in 2000 is the most significant single event in global education. PISA shifted focus from assessing what students have been taught in school more toward what they can do with knowledge and skills they have learned in school. At the dawn of the 21st century, OECD countries, in general, were increasingly interested to see how their education systems met the needs of the knowledge economy, innovation and the unknown future. PISA was expected to shed more light on how successful different education systems could raise the quality of education so that it includes all children. Indeed, PISA data also allowed a new way of comparing education systems in terms of their fairness and inclusiveness to provide equal education opportunities to different students across a wide range of social groups.

After the first three PISA cycles, it had become clear that the market-based education policies adopted by many education systems as key drivers of improvement had turned out to be disappointments. What accumulative triennial data that was available by then revealed was that the market-based reform models, i.e., competition between schools, frequent standardized student assessments, tougher accountability with associated rewards and sanctions, and a growing range of public-private partnerships to govern schools were not associated with higher education system performance ([Fullan, 2011](#); [Hargreaves and Shirley, 2009](#); [OECD, 2007](#); [2021](#); [Sahlberg, 2016](#)). Researchers and analysts around the world were especially interested in using the OECD's comparable data to test how some of the whole system education reforms of the 1990s were, however, keeping their promises of whole system success.

Analysis and early research became extremely important in answering two essential questions: One, what are the common features of those education systems that were systematically high performing in the early PISA cycles in the 2000s? Two, what are common characteristics in education reforms associated with lower-than-expected educational performance? One of the big disappointments was Sweden, a country that introduced a radical voucher reform to bring more competition, choice, and innovation in its traditional public education system in 1992. By 2006 Sweden became the lowest performing Nordic country in PISA ([OECD, 2007](#)). The conclusion was, by and large, that the neoliberal education reforms of the 1990s were a wrong way for Sweden ([Dahlstedt and Fejes, 2019](#); [Adamson et al., 2016](#); [Wiborg, 2015](#)). Another failing neoliberal education reform was the ERA in 1988 in England. Again, England's educational performance in the 2000s was disappointing to many of those who had believed in competition, school choice, standardization and student testing as key drivers of educational success. The ERA, which is one of the most researched and investigated national education reforms, has largely been judged as a failure although like in Sweden some positive gains have also been accomplished ([Strain, 2009](#); [Levin and Fullan, 2008](#); [Ball, 2018](#)). A similar verdict has been made after two decades of No Child Left Behind reforms in the US ([Ravitch, 2013](#); [Berliner, 2011](#); [Abrams, 2016](#)). Research from other countries that had adopted similar neoliberal education reforms, including Australia, New Zealand, The Netherlands, and Denmark have added to that growing evidence-base of unsuccessful education reforms for whole system improvement. In brief, the assumption that market-like competition between schools would be the driver for improving quality and equity of education was not vindicated by the evidence. This gave rise to describing such ill reforms as germs, or what is known as Global Education Reform Movement (GERM).

It is worth noting that the Global Education Reform Movement is not any official policy or program but rather informs a “new educational orthodoxy”, as Andy Hargreaves has named it. GERM was born out of the realization that there seemed to be a set of “wrong drivers” in national education reforms that have evolved over time from the growing exchange of education policies and “best practices” globally. GERM is an organic, evolving construction that relies on a particular set of assumptions, some of them mentioned above, to improve education systems (Adamson et al., 2016; Fullan, 2011, 2020; Hargreaves and Shirley, 2009; Sahlberg, 2021). In international educational change literature GERM first appeared in 2006 (Sahlberg, 2006) but Andy Hargreaves had earlier used term Global Education Reform Agenda (Hargreaves et al., 2001) that is a narrower and less detailed notion of the same thing. Other scholars like Michael Fullan wrote about “Wrong Drivers” and “The Right Drivers” of whole system reform” (Fullan, 2011, 2021), and others labeled these “Neoliberal Education Reforms” (Teng et al., 2020; Apple, 2006; Ross and Gibson, 2007) that resonates closely to the key assumptions of GERM.

Five distinct features have been dominant in the global education reform movement as it is defined in this chapter.

Increasing competition between schools over enrollments

Perhaps the most fundamental feature is increased parental choice that is associated with growing *competition* between schools for student enrollment. Most education systems have introduced alternative forms of establishing and running school to offer parents with more *choice* of schools for their children (Abrams, 2016; Heyneman, 2009; OECD 2018). But there are significant differences in the way school choice is enacted in practice. The voucher system in Chile in the 1980s, free schools in Sweden in the 1990s, charter schools in the United States in the 2000s, and secondary academies in England in the 2010s are examples of faith in pure competition as a main engine of educational progress. Consequently, the proportion of more advantaged students studying in private schools or independent schools has grown. In Australia, for example, where the federal government gives attractive financial incentives for non-government schools, more than every third primary and secondary school student studies in non-governmental schools. In Canada, for comparison, where the federal government has no role in provincial and territorial education (except education for First Nations children), most provinces and territories have had strong policies to protect and advance public education as the main form of schooling in the country (one of every fifteen students attend private school). In many of the countries where schools compete over students, school rankings based on their performance in national standardized assessments are annually published in media that further increases between-school competition. Recent data shows that, according to school principals across all OECD countries more than three-quarters of the students attend schools that compete with at least one other school for enrollment (OECD, 2019a, b). Finally, students, especially in many Asian countries, experience stronger pressure to perform better against their peers due to the tough race to be accepted to the best high schools and universities.

Standardization of teaching and learning in schools

Shifting the focus from education inputs to outcomes in the 1990s had raised the popularity of standards-based education policies, especially in the Anglo-Saxon world. Standardization-oriented education reforms initially aimed to have a stronger emphasis on learning outcomes and schools’ educational performance rather than just syllabi, curriculum, instruction time, and various structures of schooling. Presence of externally set high expectations or standards for schools, teachers, and students has been a commonly accepted precondition to improved learning and overall higher performance of schools. Standardization draws from an assumption that all students should be educated to the same (externally) agreed learning targets. In practice, this has led to a prevalence of government prescribed curricula and even homogenization of curriculum policies worldwide in the quest of universal education standards.

It goes without saying that some standardization is necessary in education, including in curriculum, teaching, and learning. But what necessity is in different contexts is where the problems start to emerge. Too restrictive externally mandated standardization can have inconvenient side effects to the teaching profession. First, it narrows the professional autonomy and flexibility in schools and classrooms that may be harmful to creative teaching and personalized learning. Second, it prevents teachers from exploration and experimentation, reduces the use of alternative pedagogical approaches, limits risk-taking in schools and thereby reduces chances of innovation. Research on educational systems that have adopted policies emphasizing steering education through externally mandated and controlled standards on core subjects, suggests that teaching and learning are narrower, and teachers focus on “proven methods” and “guaranteed content” to best prepare their students for the high-stakes tests (Abu-Alhija, 2007; Berliner, 2011; Emler et al., 2019). The consequence is that the more schools are expected to conform to predetermined teaching and learning standards in schools, the lower the degree of freedom and risk-taking in classroom teaching and learning there will be. When teachers and students start to avoid taking risks, they stop trying new ideas in school that, in turn, will limit creativity and innovation.

Increasing priority to basic literacy and numeracy

Learning to read, write and do mathematics well at school is important, but so are many other things in today’s schooling. Those who have championed “back to basics” in school education often do that at the expense of social studies, arts, music, physical education, and play. The dominance of literacy and numeracy means that they also become the prime criteria of improvement in national education reforms. According to the OECD’s own conclusion (Schleicher, 2018) and comparative education research

(Ball, 2012; Sellar et al., 2017; Fischman et al., 2019) on national education policies in several countries, national education policies are increasingly influenced by international student assessments, especially PISA. Since PISA primarily measures literacy and numeracy in schools, those subjects have had growing attentions in many countries during the past two decades.

Literacy and numeracy strategies that increased instructional time for the core subjects in England, Ireland, and Ontario (Canada), for example, are concrete consequences of the global educational reform movement in the teaching profession. *No Child Left Behind* legislation in the United States has led most school districts in the country to shift teaching time from other subjects, especially from social studies, arts and music, to teaching reading, mathematics and science so that schools are better prepared for tests that measure student performance and hold schools accountable in these subjects (Jennings and Stark Rentner, 2006). Australia imported its National Assessment Program – Numeracy and Literacy (NAPLAN) that started to measure literacy and numeracy annually in school years 3, 5, 7 and 9 from the US as a response to the nation's less-than-satisfactory performance in global education rankings. Moreover, poorer than expected performance in the OECD's PISA has led Sweden, Norway, and Denmark to revise their curricula and related school policies to give more attention to the basic academic domains and standardized measurements of student achievement in these domains. Worryingly, strategic privilege given to literacy and numeracy comes at the cost of eliminating recess and free playtime at school. When literacy and numeracy time is prioritized above all else, other necessary aspects of school and child development are easily ignored.

Test-based accountability

Increased school autonomy and flexibility in teaching normally mean that accountability begins to play clearer role in administration and management. Accountability is a term that has become fashionable in education, illustrating education's closer partnership with the business world. In education systems that have adopted corporate management models, teachers and schools are held accountable for students' achievement using data from standardized assessments and opinion surveys. School performance—especially raising students' achievement in academic subjects—is intimately tied to the processes of assessing, inspecting, appraising, and rewarding or punishing schools and teachers. Performance-based pay, data management systems, and school rankings are some examples of new accountability mechanisms. Growing infatuation with the performance-based pay scale and test-based accountability relies upon two erroneous ideals. First, teachers will work harder for extra pay. Second, test scores are the largest factor in determining effective teachers. Using this logic, if a teacher is boasting stellar student test scores, she must be working diligently, teaching successfully, and worthy of higher pay. It is this thinking that contributes to the undermining of the teaching profession in certain countries across the globe.

The problem with test-based accountability is not that students, teachers, and schools are held accountable per se, but rather the way accountability mechanisms are designed and how they affect teachers' work and students' studying in school. A solid empirical finding for 50 years (e.g., Alexander and Morgan, 2016; Haertel, 2013) is that students' family background is more powerful factor associated to academic achievement in school than people think. If most of the variation in students' test scores in school is associated to factors outside the school gate, schools cannot be held entirely accountable for these test scores. Moreover, recent research on the social side of education reform has revealed that social capital, in other words teacher collaboration and teamwork in teaching have positive effect on student learning (Quintero, 2017). This means that measured student achievement is not always a result of any one single teacher but rather a collective outcome of the school. A conclusion that the OECD (2021) made in its cross-country analysis is that test-based accountability is not linked to academic achievement, nor has a notable impact on educational inequity. More specifically, the OECD (2021, 25) concluded that "the intensive competition across schools that test-based accountability promotes could be disruptive in some educational contexts, producing unintended consequences in school communities." This raises important questions about correctness of these policies as they often do more harm than make education better.

Privatization of public education

Parental choice that is the key to expanding private education emerged from Milton Friedman's economic theories in the 1950s. Friedman maintained that parents must be given the freedom to choose their children's education and thereby encouraged healthy competition among schools so that they could better serve families' diverse individual needs. This spirit of school choice led to charter schools and free schools in the US and Sweden in early 1990s, Secondary Academies in England, and expanding independent school sectors and private actors in education in Australia, Denmark, and New Zealand. Typically, school choice manifests itself through the emergence of private schools where parents pay part or all of the tuition for their children's education. Today, there are scores of various types of alternative schools other than fee-based private schools to expand choice in education markets. Privately funded schooling ideology maintains that parents should have a right to use the public funds set aside for their children's education to choose the schools—public or private—that work best for them. Privatization of public education also includes other forms of private actors engaged in the education sector (Adamson et al., 2016). This may have other consequences in regard to the ways education systems are able to serve their purpose for the common good. Thurpp et al. (2020) conclude that "in many ways the New Zealand case illustrates how a public education system can lose capacity and become dependent on private actors."

The five features of the global education reform movement are not the complete list of failed efforts to enhance education system performance, especially equity and excellence. There are some other common but less consistent aspects of educational change that have emerged through borrowing "best practices" rather than deeper policy learning. Whole system reforms by technology (digital devices to all students), de-professionalization of the teaching profession (fast-track pathways to teaching and leadership), and

complex information management systems (multiple levels of data from schools) are examples of other efforts that have been launched with high hopes but ended up less than expected outcomes.

Antidotes to GERM and conclusion

If GERM is not the right way to improve quality and equity of an education system, what is the alternative? Critiques of GERM should not state that all earlier mentioned features of education reform are wrong and should be avoided. What is important to understand is that those features are wrong priorities for system wide education reform, as Michael Fullan (2011, 2021) says, but they may have a role in some well-thought ways serving other policy drivers. The answer to the question of what the alternative directions to reform education systems could be, or an antidote to GERM if you wish, requires finding out the key policy drivers in more successful education systems. This is where comparable data from the OECD member states' education systems is useful.

Comparative analysis of OECD's data since 2000 revealed a surprising finding: successful education systems in terms of quality and equity of outcomes, participation and completion rates, and systems efficiency were quite different than what most expected. Finland, Canada, The Netherlands, South Korea, and Japan were among the highest performing education systems at the end of 2000s (OECD, 2007; Sahlberg, 2010). These are very different education systems in many ways but, interestingly, they have more things in common as far as the key drivers of educational system improvement are concerned. Table 2 illustrates some of the main features in high performing education systems compared to the GERM.

It is difficult to prove the success or failure of any education reform. Education systems are complex social systems and causal associations between events or elements are difficult, if not impossible to identify reliably. When education leaders make decisions regarding their policies and reforms, they can rely on (1) available evidence about reported possible success factors of high performing education systems in Table 2; and (2) current research and analysis of those large-scale reforms that were based on one or more policy drivers in Table 1. It is beyond the scope of this chapter to make a comprehensive review of research evidence regarding either of these options. Instead, it is worthwhile to mention some of the most significant conclusions made by the OECD to inform policy-makers about wrong and right drivers for whole system success.

In 2012 the OECD (2012) published its first comprehensive analysis of quality and equity in education. The review included PISA data since 2000 from all OECD countries and provided a detailed comparison of different education policy contexts from the point of view of quality and equity of educational outcomes. The OECD argued that greater equity in education can contribute to economic competitiveness and social cohesion. Another finding was that fairer and more inclusive education can serve as a lever out of financial crisis that most OECD countries found themselves in at the time of this analysis. The overall evidence was conclusive

Table 2 Key education policy and reform features in high performing education systems.

<i>Global education reform movement</i>	<i>High performing education systems</i>
Increasing competition between schools over enrollments	Collaboration and networking The basic assumption is that education is fundamentally a collaborative process and therefore cooperation, networking, and sharing ideas among schools benefits everyone in raising the quality of education. Collaboration is encouraged in education policies and school leadership focuses on schools and teachers helping one another.
Standardization of teaching and learning in schools	Creative customization of curriculum Schools have a key role in designing the best possible teaching and learning for all students guided by flexible national framework or curriculum. Schools are supported to come up with new ideas to curriculum and to experiment different methods of teaching. Individual learning plans are commonly used to accommodate needs of different learners.
Increasing priority to basic literacy and numeracy	Whole-child approach Teaching and learning focus on deep, meaningful learning, giving equal value to all aspects of the growth of an individual's personality, moral character, creativity, knowledge, ethics, and skills. Student wellbeing is closely integrated in the process of schooling.
Test-based accountability	Trust-based professionalism Teachers and schools as communities of professionals are trusted to do what is best for students. Accountability is viewed as reciprocal professional link between school and authorities. Student assessments are primarily designed on sample-based testing methods and high-quality research is used to inform both schools and policymaking about improvement.
Privatization of public education	Social equity and adequacy in education Basic premise is that all children should have equal prospects for educational success in school and public education is the best way to guarantee that right to everyone. Because school learning is strongly influenced by children's family background and associated factors, social equity and adequacy in education outcomes are the basic policy and reform principles. School choice is made possible within public education system as much as possible.

- investing in educational equity pays off. The OECD (2012, p. 14) made a bold conclusion: “The highest performing education systems across OECD countries are those that combine high quality and equity.” Six years later the OECD released its second analysis about the central role that equity in education plays in economic and social development. It stated that “Disparities in performance related to socio-economic status develop early and widen throughout students’ lives” (OECD, 2018, p. 30), and in most OECD countries the association between socio-economic status and academic performance did not change over the past decade leaving inequity in education one of the key challenges in most countries. These findings conflict with earlier beliefs that educational excellence comes with a cost of inequity. Studies by Parker and colleagues and Pfeffer (Parker et al., 2018; Pfeffer, 2015) that used OECD data to test this assumption confirmed that “a negative relationship exists between average academic excellence and inequity” (Parker et al., 2018, p. 855). In other words, education system can have high education outcomes and be equitable regarding these outcomes simultaneously.

The OECD has not found support to the claim that between-school competition is associated with improved student learning outcomes (OECD, 2019c). Empirical evidence on whether student achievement is higher when there is greater competition between schools is mixed. Instead, the OECD (2012, p. 64) warns that

School choice advocates often argue that the introduction of market mechanisms in education allows equal access to high quality schooling for all. Expanding school choice opportunities, it is said, would allow all students – including disadvantaged ones and the ones attending low performing schools – to opt for higher quality schools, as the introduction of choice in education can foster efficiency, spur innovation and raise quality overall. However, evidence does not support these perceptions, as choice and associated market mechanisms can enhance segregation.

Rather than encouraging governments to have more parental choice and school-to-school competition, the OECD (2019, p. 12) advises them to “provide the checks and balances that prevent choice from leading to more segregation”, and “ensure that all parents can exercise their right to choose the school of their preference.”

Further antidote against GERM can be found in the OECD’s assessment of whether and which accountability practices affect equity and education performance in student academic achievement in school. By using data available from consecutive PISA cycles (2006–2015) the OECD analyzed how accountability practices affect equity and quality of education outcomes in different countries. The conclusion was that there is no conclusive evidence of accountability positively affecting education systems performance (OECD, 2021).

The *Educational Review* journal dedicated an entire special issue in 2019 to the analysis of GERM and its manifestations in international education. In the editorial Kay Fuller and Howard Stevenson (2019) summarize the problematic consequences of GERM in different countries (e.g., Chile, England, and China). Articles in that special issue describe awareness of GERM has provided students and educators with a new vocabulary to resist wrong-headed reforms and speak about education systems that are visibly “cracked”.

The OECD has had dual influence on the emergence of the global education reform movement (Heyneman and Lee, 2014; Lingard and Sellar, 2016; Sellar and Lingard, 2013). On one hand it has promoted many aspects of GERM, especially standardization and homogenization of education by promoting standardized assessments in measuring education systems’ success. OECD’s active role in facilitating international policy dialogs has often led to simplistic transfer of “best practices” and even adopting education policies from other countries. On the other hand, it has been instrumental in producing more data and analyzing some of the most controversial topics in education – school choice, accountability, and equity in education. Conclusions that more competition and choice or tougher accountability don’t guarantee sustainable educational improvements, and that investing teacher professionalism and equity in education are smarter policies have been fundamental reasons why GERM was realized two decades ago. By acting this dual role the OECD has created antidotes to the infections that it has been spreading globally.

Despite frequent school reforms around the world, educational performance overall has not improved during the past decade. Indeed, it has been in decline compared benchmark levels of academic results around 2010. The global statistics reveal system-wide inequities that are boosted by the present global pandemic. As this chapter makes it clear, the inspiration for systemwide education reforms have often been imported from the US and England. Yet, the evidence to support many of these large-scale policy changes is weak or non-existent. For instance, research also referenced in this chapter has shown that school choice, test-based accountability, or privatisation are wrong drivers for whole-system educational success. So, what to do instead? Success in fighting the global pandemic is a result of relying on best available science and expert knowledge to maximise the effectiveness of treatments while minimising their side effects. We should follow that same principle that have led success in public health in many parts of the world in education, too. Herein lies silver lining amid global disaster. No doubt there will be new variants of GERM that will affect education systems in the future. If we have learnt anything since 2020, it is that we need to act in protecting and improving education more like we act in defending public health systems in keeping each of us safe.

References

- Abrams, S., 2016. Education and the commercial mindset. Harvard University Press, Cambridge, MA.
- Abu-Alhija, F.N., 2007. Large-scale testing: benefits and pitfalls. Stud. Educ. Eval. 33, 50–68.
- Adamson, F., Åstrand, B., Darling-Hammond, L. (Eds.), 2016. Global Education Reform. How Privatization and Public Investment Influence Education Outcomes. Routledge, New York, NY.

- Alexander, K., Morgan, S., 2016. The Coleman Report at Fifty: Its Legacy and Implications for Future Research on Equality of Opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences*. <https://doi.org/10.7758/rsf.2016.2.5.01>. Retrieved from: https://socweb.soc.jhu.edu/faculty/morgan/papers/Alexander_and_Morgan_2016.pdf.
- Apple, M.W., 2006. Understanding and interrupting neoliberalism and neoconservatism in education, pedagogies. *Int. J.* 1 (1), 21–26. https://doi.org/10.1207/s15544818ped0101_4.
- Ball, S., 2008. The legacy of ERA, privatization and the policy Ratchet. *Educ. Manag. Adm. Leader* 36 (2), 185–199.
- Ball, S., 2012. *Global Education Inc. New Policy Networks and the Neo-Liberal Imaginary*. Routledge, Abingdon, UK.
- Ball, S., 2018. The tragedy of state education in England: reluctance, compromise and muddle—a system in disarray. *J. Br. Acad.* 6, 207–238. <https://doi.org/10.5871/jba/006.207>.
- Berliner, D.C., 2011. Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Camb. J. Educ.* 41, 287–302.
- Carnoy, M., 1999. *Globalization and Educational Reform: What Planners Need to Know?* Unesco/IEP, Paris, France.
- Dahlstedt, M., Fejes, A. (Eds.), 2019. *Neoliberalism and Market Forces in Education: Lessons from Sweden*. Routledge, London.
- Emler, T.E., Zhao, Y., Deng, J., Yin, D., Wang, Y., 2019. Side effects of large-scale assessments in education. *ECNU Rev. Educ.* 2 (3), 279–296. <https://doi.org/10.1177/2096531119878964>.
- Fischman, G., Sahlberg, P., Silova, I., Marcetti Topper, A., 2019. International large-scale student assessments and their impact on national school reforms. In: Larry Suter, L., Smith, E., Denman, B. (Eds.), *The Handbook of Comparative Studies in Education*. SAGE Publishing, New York.
- Fullan, M., 2011. *Choosing the Wrong Drivers for System Wide Reform* (Paper No. 204). Centre for Strategic Education Seminar Series, Melbourne.
- Fullan, M., 2021. *The Right Drivers for Whole System Success* (CSE Leading Education Series No. 1). Centre for Strategic Education Seminar Series, Melbourne.
- Fuller, K., Stevenson, H., 2019. Global education reform: understanding the movement. *Educ. Rev.* 71 (1), 1–4.
- Haertel, E.H., 2013. *Reliability and Validity of Inferences about Teachers Based on Student Test Scores*. Educational Testing Service, Princeton, NJ.
- Hargreaves, A., Shirley, D., 2009. *The Fourth Way. The Inspiring Future of Educational Change*. Corwin, Thousand Oaks.
- Hargreaves, A., Earl, L., Moore, S., Manning, S., 2001. *Learning to Change. Teaching beyond Subjects and Standards*. Jossey-Bass, San Francisco, CA.
- Heyneman, S., Lee, B., 2014. Impact of international studies of academic achievement on policy and research. In: Rutkowski, L., von Davies, M., Rutkowski, D. (Eds.), *Handbook of International Large Scale Assessments*. Taylor and Francis, Boca Raton, pp. 37–75.
- Heyneman, S., 2009. International perspectives on school choice. In: Berends, M., Springer, M.G., Ballou, D., Walberg, H.J. (Eds.), *Handbook of School Choice*. Routledge, New York.
- Hood, C.C., Jackson, M., 1991. *Administrative Argument*. Gower, London.
- Hopkins, D., Stringfield, S., Harris, A., Stoll, L., Mackay, T., 2014. School and system improvement: a narrative state-of-the-art review. *Sch. Effect. Sch. Improv.* 25 (2), 257–281. <https://doi.org/10.1080/09243453.2014.885452>.
- Jennings, J., Stark Rentner, D., 2006. *Ten Big Effects of the No Child Left behind Act on Public Schools*. Center on Education Policy.
- Joyce, B., Showers, B., 1995. *Student achievement through staff development: fundamentals of school renewal*, 2nd ed. Longman, White Plains, NY.
- Levin, B., Fullan, M., 2008. Learning about system renewal. *Educ. Manag. Adm. Leader* 36 (2), 289–303. <https://doi.org/10.1177/1741143207087778>.
- Lingard, B., Sellar, S., 2016. The changing organizational and global significance of the OECD's education work. In: Mundy, K., Green, A., Lingard, R., Verger, A. (Eds.), *The Handbook of Global Policy and Policymaking in Education*. Wiley-Blackwell, New York, pp. 357–373.
- OECD, 2001. *Knowledge and Skills for Life: First Results from PISA 2000*. OECD Publishing, Paris.
- OECD, 2007. *PISA 2006. Science Competencies for Tomorrow's World*, vol. I. OECD Publishing, Paris.
- OECD, 2012. *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. OECD Publishing, Paris.
- OECD, 2018. *Equity in Education: Breaking Down Barriers to Social Mobility*. OECD Publishing, Paris.
- OECD, 2019a. *PISA 2018 Results (Volume I): What Students Know and Can Do*. OECD Publishing, Paris.
- OECD, 2019b. *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OECD Publishing, Paris.
- OECD, 2019c. *Balancing School Choice and Equity. An international perspective based on PISA*. OECD, Paris.
- OECD, 2021. Does test-based school accountability have an impact on student achievement and equity in education? A panel approach using PISA. *OECD Education Working Papers* No. 250. OECD, Paris. <https://doi.org/10.1787/0798600f-en>.
- Parker, P., Marsh, H., Jerrim, J., Guo, J., Dicke, T., 2018. Inequity and excellence in academic performance: evidence from 27 countries. *Am. Educ. Res. J.* 55 (4), 836–858. <https://doi.org/10.3102/0002831218760213>.
- Pfeffer, F., 2015. Equality and quality in education. A comparative study of 19 countries. *Soc. Sci. Res.* 51, 350–368. <https://doi.org/10.1016/j.ssresearch.2014.09.004>.
- Quintero, E. (Ed.), 2017. *Teaching in Context: The Social Side of Education Reform*. Harvard Education Publishing, Cambridge, MA.
- Ravitch, D., 2013. *Reign of Error. The Hoax of the Privatization Movement and the Danger to America's Public Schools*. Alfred A. Knopf, New York, NY.
- Ross, E.W., Gibson, R. (Eds.), 2007. *Neoliberalism and Education Reform*. Hampton Press, Cresskill, NJ.
- Sahlberg, P., 2004. Teaching and globalization. *Int. Res. J. Manag. Glob. Trans.* 2 (1), 65–83.
- Sahlberg, P., 2006. Education reform for raising economic competitiveness. *J. Educ. Change* 7 (4), 259–287.
- Sahlberg, P., 2010. Educational change in Finland. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*. Springer, New York, pp. 323–348.
- Sahlberg, P., 2016. Global educational reform movement and its impact on teaching. In: Mundy, K., Green, A., Lingard, R., Verger, A. (Eds.), *The Handbook of Global Policy and Policymaking in Education*. Wiley-Blackwell, New York, pp. 128–144.
- Sahlberg, P., 2021. *Finnish Lessons 3.0: What Can the World Learn from Educational Change in Finland?* Teachers College Press, New York, NY.
- Sarason, S., 1990. *The Predictable Failure of Education Reform. Can We Change the Course before Its Too Late?* Jossey Bass, San Francisco.
- Schleicher, A., 2018. *World Class. How to Build a 21st-Century School System*. OECD Publishing.
- Sellar, S., Lingard, B., 2013. The OECD and the expansion of PISA: new global modes of governance in education. *Br. Educ. Res. J.* 40 (6), 917–936.
- Sellar, S., Thompson, G., Rutkowski, D., 2017. *The Global Education Race: Taking the Measure of PISA and International Testing*. Brush Education.
- Strain, M., 2009. The Education reform act 1988: success or failure? A short report of a BELMAS Discussion Forum. *Management in Education* 23 (4), 151–154. <https://doi.org/10.1177/0892020609344036>.
- Teng, S.S., Abu Bakar, M., Layne, H., 2020. Education reforms within neoliberal paradigms: a comparative look at the Singaporean and Finnish education systems. *Asia Pac. J. Educ.* 40 (4), 458–471. <https://doi.org/10.1080/02188791.2020.1838884>.
- Thrupp, M., O'Neill, J., Powell, D., Butler, P., 2020. Private actors in New Zealand schooling: the path to saturation. *J. Educ. Adm. Hist.* 52 (4), 385–402. <https://doi.org/10.1080/00220620.2020.1726882>.
- Wiborg, S., 2015. Privatising education: free school policy in Sweden and England. *Comp. Educ. Rev.* 59 (3), 473–497.
- Zajda, J., Rust, V., 2021. *Globalisation and Comparative Education. Changing paradims*. Springer, Dordrecht.

The privatization of education. Drivers, social effects and regulatory challenges of private sector participation in schooling

Antoni Verger^a, Clara Fontdevila^b, and Adrián Zancajo^c, ^a Universitat Autònoma de Barcelona, Cerdanyola, Spain; ^b University of Glasgow, Glasgow, United Kingdom; and ^c University of Manchester, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Definition and scope	174
Conceptual demarcation and relevant distinctions	174
World trends	175
Main drivers	175
Privatization by design versus <i>de facto</i> privatization	177
Beyond the neoliberal creed: privatization as a globalizing policy	177
The effects of privatization: emerging consensus and persisting discrepancies	178
Regulation and governance of private provision	179
The regulation of PPPs in education	179
Regulating independent schools: the case of LFPS	180
To conclude	180
References	181

The privatization of education is a global phenomenon taking place at all educational levels and growing worldwide. In this entry we define educational privatization and examine its growth in different regions, discussing why it emerged, and how it has developed. We also reflect on the main effects of education privatization policies, with an emphasis on the challenges these policies pose for educational equity and the public regulation of education. We focus on basic education, including primary and secondary education.

Definition and scope

Conceptual demarcation and relevant distinctions

Privatization is a process through which the private sector increasingly participates in economic and social activities that, in most countries, used to be the remit of the State. In the last decades, privatization has crystallized in public services that were previously de-commodified and in relation to which the State has had an important historical presence over the 20th century, as tends to be the case of education. Nonetheless, privatization processes can occur in various guises and to different degrees. To start with, we can distinguish between drastic and gradual modes of privatization. Drastic privatization implies that both service ownership and management responsibilities are transferred from public to private hands at a large scale. Although this form of privatization has been more common in gas, water or telecommunications sectors, the expansion of fully private schools, which are exclusively funded by families' fees and/or philanthropic resources, would be also considered as part of a drastic privatization trend. In contrast, gradual privatization entails a progressive or partial transfer of service ownership and/or management and delivery responsibilities from the public to the private sector, with the State retaining full or an important part of the ownership. In most contexts, the privatization of education has tended to follow a gradual privatization path rather than a drastic one. In fact, the privatization of education tends to crystallize at the level of provision (via the outsourcing of service provision to private management organizations) and/or funding (via co-payment policies). Conversely, it rarely entails changes at the ownership level- with the rapid *academization* of schools in England being more an exception than the rule (Verger, 2020).

The concept of Public-Private Partnerships (PPPs) has been broadly used in educational literature to capture the predominantly gradual character of privatization processes in education, as well as the hybrid nature that educational provision is acquiring in numerous school systems. PPPs in education encompass a broad range of policies and programs that entail a contract between the State and the private sector in which public sources fund an education service, and a private actor takes responsibility for its delivery. In the context of PPPs, both sectors, the private and the public, are expected to share risks, ideas, and other resources in delivering services (Hodge et al., 2010). In education, it is possible to distinguish at least three main policy categories that fall under the PPP umbrella (Zancajo et al., 2021a):

1. **Charter schools.** Schools are owned by the State but managed by private entities. Frequently, these schools enjoy higher levels of autonomy than regular public schools.
2. **Vouchers.** Private schools receive public funding depending on the number of students enrolled. In some cases, private schools can combine voucher and self-funding students.
3. **Contracted schools.** Private schools that the government directly contracts to provide the educational service in exchange for public funding covering part or all the costs they require to operate. Under this scheme, and in contrast with vouchers, schools are not necessarily publicly funded on a per-capita basis.

Another relevant distinction is that advanced by Ball and Youdell (2007), who have famously distinguished between exogenous and endogenous forms of educational privatization. Exogenous privatization involves the opening up of public services to private sector participation, whereas endogenous privatization involves the importing of ideas and management techniques from the private to the public sector. Endogenous privatization can occur independently from exogenous privatization, in the sense that it does not necessarily imply the privatization of service provision, ownership, or financing—but the adoption of a private management styles and techniques within public sector, including outcomes-based management, competition between providers, job flexibility, performance-based incentive schemes, etc.

In Table 1, we show the educational policies associated with each type of privatization. The table makes it clear that certain policies can promote both exogenous and endogenous privatization. Nonetheless, the privatization concept is far more frequently used to refer to exogenous than endogenous privatization.

It should finally be noted that, while privatization and marketization are not synonymous terms, they tend to be inextricably linked. Most privatization processes, both exogenous and endogenous, lead to the generation of market dynamics in educational systems (including the promotion of school choice and school competition to attract users). By default, a greater presence of private providers in a given system creates an environment that is conducive to market logics and competitive practices. In the same way that privatization processes tend to progress in parallel with marketization processes, privatization tends to correlate to liberalization, with the latter understood as the loosening and/or dismantlement of government regulations to favor private sector participation. Nonetheless, the relationship between privatization, liberalization and deregulation is not always linear. As we show in this article, the privatization of education has generated numerous social and political challenges, and many governments around the world are responding to these challenges with additional layers of public regulation. Thus, although education privatization processes eventually require of some form of deregulation, over time, these processes tend to generate more regulatory than deregulatory dynamics.

World trends

The privatization of education can be considered a global phenomenon, but it has intensified further in some regions and countries than in others. Fig. 1 shows that Latin America is the region with the highest percentage of countries with a substantive proportion of private provision in their educational systems, followed by Middle East and North Africa, and Sub-Saharan Africa. The figure also shows that, in Latin America, the percentage of countries with more than 20% of private enrollments in primary education expanded the fastest, whereas in Europe and Central Asia the growth concentrated among countries with 10%–20% of private enrollment.

The pace and historical evolution of such developments also vary across regions. In the Middle East and North Africa, the expansion of private enrollment has gradually increased since 1995. Conversely, in South Asia, private enrollment expanded rapidly during the 2000s after a period of relative stability and predominance of public provision. Finally, the figure also indicates that world regions are also becoming internally more diverse, in the sense that the share of private enrollment varies greatly across the countries constituting each region.

Main drivers

Not only are countries becoming more diverse in how much space private provision is acquiring in educational systems, but the reasons and causes driving the rise of private schooling are also multiple and subject to change. It is possible to discern at least six different paths that lead toward the privatization of education. In Table 2, we describe each of these paths and identify the country cases that better correspond to each of them. These paths differ not only in the policy outputs into which privatization

Table 1 Types of educational privatization and main policy programs.

<i>Type of privatization</i>		<i>Aim</i>		<i>Education policies</i>	
Exogenous	Promoting the emergence and expansion of private providers in the education sector	• Liberalization of the education sector • Tax incentives to private schools and/or private schooling consumption • Public subsidies to private schools • Contracting out private providers	• Vouchers and similar competitive formula in which financing follows the demand • Charter schools' programs aiming to expand school choice options • Other freedom of school choice policies (such as the elimination or enlargement of catchment areas)		
Endogenous	Introducing management norms, rules and logics of the private sector within public education	• Performance-related pay for schools and/or teachers • Disaggregation of units in the educational system; school-based management • Standardized evaluation and school rankings			

Authors based on Ball and Youdell (2007).



Fig. 1 Percentage of countries by their share of private enrollment in primary education, 1975–2015. Authors with data from UNESCO Institute for Statistics and World Bank Data Bank.

Table 2 Paths toward privatization.

Path	Privatization process	Paradigmatic cases
Reshaping the role of the State in education	Drastic advancement as part of a structural State reform adopted by neoliberal governments and consolidated by succeeding center-left administrations. Resulting in a quasi-market system where the State focuses on regulation and monitoring.	Chile, United Kingdom, New Zealand
Education privatization in social-democratic welfare states	Pro-market reforms framed as part of a necessary modernization of the welfare state in a globalized context. Proactive leading or collaboration of social-democratic forces, usually less inclined to promote competition but to respond to middle-class demand for diversification.	Sweden, Denmark, Norway, Finland
Scaling-up privatization	Uneven but progressive alteration of the system through the authorization and encouragement of new forms of private provision and management. School choice as the guiding principle and with the key contribution of non-state actors forming advocacy coalitions.	United States of America, Canada, Colombia, Brazil
De facto privatization in low-income countries	Expansion of a low-fee private school sector, originally set by local <i>edupreneurs</i> responding to a growing education demand. Increasingly promoted by the international development community and supported by national governments as part of PPP schemes.	Ghana, India, Kenya, Malawi, Nigeria, Pakistan, Peru, Dominican Republic, Jamaica
Historical PPPs	Historical role and importance of private faith-based schools, incorporated in the State network during the expansion of basic in the 20th century. Process mainly oriented to achieve a State-Church compromise rather than on ideological grounds.	The Netherlands, Belgium, Spain, Argentina
Privatization by way of catastrophes	Advancement catalyzed by natural disasters or violent conflicts, which are framed by privatization advocates as an opportunity to reconstruct the education system. Developed in contexts where the democratic deliberative process is undermined because of the sense of urgency.	New Orleans, El Salvador, Nicaragua, Haiti, Guatemala, Honduras, Iraq

Source: Adapted from Verger et al. (2016) and Verger et al. (2018).

crystallizes (e.g., vouchers, charter schools, etc.) but also in the drivers that explain the nature and intensity of these transformations—that is, in the causes, rationales and contextual factors triggering and mediating in educational privatization trends.

Privatization by design versus *de facto* privatization

Privatization can happen by design or *de facto*, depending on the role of national governments in the expansion of private provision. Privatization by design implies public authorities proactively promoting pro-private sector policy measures – often motivated by ideological reasons. In the late 1970s and the beginning of the 1980s, for instance, the first wave of neoliberal governments, among which the Chilean and UK governments stand out, privatized public assets to generate state revenue and encourage market competition. These reforms were strongly inspired by public choice theory and the then popular economic ideas of the Chicago School. Therefore, the education privatization measures adopted in these contexts frequently combined free choice and competitive funding formulas aimed at empowering families and making educational providers more responsive to clients' demands. The privatization agenda that emerged in the 1980s was strongly grounded in the belief that the government's role in all economic and social domains had to be curtailed, or otherwise, the public sector would lack the necessary incentives to provide services efficiently and responsively. According to neoliberal advocates, the public sector is easily captured by vested interests such as self-serving bureaucracies and labor unions, which is why they conceive the privatization agenda as an opportunity to undermine the power and density of trade unions in the public sector.

Nonetheless, privatization policies have not necessarily followed such an ideological logic in all contexts. In countries such as the Netherlands, Argentina or Belgium, for example, the State has contracted out the provision of education to private – usually religious – schools, due to the historical role of these institutions in the educational system, the political and social significance of faith-based organizations and the centrality of freedom of education values. In fact, the historical establishment of PPPs between the State and the private sector adopted in these countries significantly precedes in time the emergence of neoliberalism.

De facto privatization, in contrast to privatization by design, advances for reasons that are apparently unrelated to public action and policy intentions. *De facto* privatization often unfolds in contexts where the State does not react to citizens' demands for new educational services and these citizens turn to private provision to satisfy their needs. Lack of state administrative capacity for education planning, insufficient resources to invest in public education and the massification of classrooms in government schools are some of the reasons why the phenomenon of Low-Fee Private Schools (LFPSS) has boomed in many low-income countries (Härmä, 2021).

Beyond the neoliberal creed: privatization as a globalizing policy

In the 1990s, after the collapse of the Soviet Union, the privatization agenda penetrated different regions of the world. During this period, international financial institutions started to give economic support to transitioning economies and low- and middle-income countries, generally under the form of aid programs conditional on substantive budget cuts in public services that led to pro-private sector reforms. These structural adjustments particularly affected Latin America and Sub-Saharan Africa regions, as well as Eastern Europe, where many countries experienced economic crises in this period and public debt had reached high levels.

In Europe, the Maastricht Treaty, signed in 1992, also created a fertile ground for privatization due to the fiscal austerity it imposed on European Union members. This restriction on public expenditure created an appetite for the private provision of public services which translated into greater reliance on private entities as providers of public services. In continental Europe, privatization was more pragmatic and moderate than in Anglo-Saxon countries, and did not penetrate into public services such as education so substantially. This is because of the strong welfare state tradition and a culture of social dialog between government, unions and civil society that prevails in most European countries held back more drastic or structural reforms (Vere, 2007).

It was also during the 1990s that social democratic governments started to experiment with the engagement of private entities in the delivery of public services. This represented a turning point since, historically, left-leaning parties have been reluctant to adopt education privatization measures. However, with the advent of the Third Way in the 1990s, fully public and homogenous services were no longer perceived as a possible or a desirable horizon among social democratic forces. This led to the neoliberal discourse about the "inherent superiority of the private sector" being replaced with the belief that certain pro-market measures could help modernize public education and promote greater innovation and diversification. In the 1990s, for instance, the Labor Party in the UK, the Social Democratic Party in Sweden and the Democrats in the US were swept up by this change, which translated into a rapid increase in private education provision. It should be noted that, as Hicks (2015) points out, left-wing or social-democratic parties tend to be more open to favoring markets and the private sector in education systems when social inequality is low, as they tend to be aware that, in contexts of high socio-economic inequality, education markets contribute to amplifying social and educational inequalities.

While political variables (and the government ideology in particular) appear to be poor predictors of privatization reforms, economic drivers are more influential in understanding privatization trends in different sectors. A more liberalized global economy motivates governments to outsource all types of public services, including education, particularly during periods of economic recession (Obinger et al., 2016). Economic factors also predispose families to invest privately in educational services to position their children more favorably in increasingly competitive and polarized labor markets.

The effects of privatization: emerging consensus and persisting discrepancies

In academic and policy circles alike, advocates of education privatization have long emphasized the potential of the private sector to improve educational provision. In industrialized countries, privatization policies have been promoted as a strategy to enhance education quality, equity and effectiveness, as well as to incentivize pedagogical innovation and the diversification of school supply. This line of reasoning is informed by the work of a number of influential pro-market scholars such as Chubb and Moe (1990), Friedman (1994) or Peterson (1990). Within development circles, in turn, the popularity of private education delivery owes much to the assumption that such schooling modality will widen access to education while leading to a better distribution of educational opportunities. This expectation is informed by the empirical and advocacy work of a group of experts and practitioners operating at the interstices of academia and international agencies (see for instance, Patrinos et al., 2009; Tooley, 1996; or LaRocque and Lee, 2011).

However, while privatization proponents have relied heavily on studies emphasizing the benefits of private provision, the broader academic conversation appears to be much more complex and nuanced, with the bulk of investigations bringing mixed or inconclusive results (Verger et al., 2019). Indeed, the growing volume of research published on the topic reaches disparate and even conflicting conclusions regarding the impact of private education delivery. Reviews of evidence published over the last decade have concluded that research on the impact of private education provision is insufficient and often yields contradictory results – see for instance the work of Waslander et al. (2010), Day Ashley et al. (2014); Aslam et al. (2017); Verger et al. (2020); or Gruijters et al. (2021). Remarkably, this applies to research examining the effects of publicly subsidized private education under the form of PPPs, but also to research focusing on LFPs which are not necessarily benefitting from government subsidies or other forms of support.

These studies find that evidence on PPP's potential to improve the quality, effectiveness and diversity of education systems is mixed at best. Hence, whereas some widely-circulated studies identify a positive effect in the improvement of learning outcomes (e.g., Hoxby, 2003 or Woessmann, 2005), others find that PPP's impact in this area is virtually nil (see Musset, 2012; Belfield and Levin, 2002, or Rouse and Barrow, 2009). In the case of developing countries, a systematic literature review conducted by Aslam et al. (2017) observed that evidence on the aggregate impact of these interventions on learning outcomes is still scarce and often inconclusive; whereas the review conducted by Day Ashley et al. (2014) concluded similarly that the size of LFPs's academic advantage was small. The same is true for OECD countries, as suggested by the review conducted by Waslander, Pater, and van der Weide (2010). The authors conclude that available evidence on the impact of market mechanisms is limited and frequently undermined by methodological limitations – and that, even when positive effects are found, they are generally modest.

Likewise, evidence on the cost-efficiency advantages associated with private education provision is relatively scarce and frequently contested – with some authors drawing attention to the fact that any gains in this area are likely to be neutralized by transaction costs and the administrative burden associated with the supervision of PPPs. This is the case of the literature review conducted by Languille (2017), who argues that available evidence demonstrates that in the case of education (but also health) PPPs, the monitoring and management costs assumed by the governments challenge the alleged superiority of these programs in terms of cost-efficiency. While research on such questions is relatively scarce in the area of education, Languille's findings are consistent with the argument advanced by Bartlett and Le Grand (1993) in relation to the provision of welfare and social services, according to which high transaction costs are likely to reduce the efficiency of quasi-markets compared to social services allocated through bureaucratic, State-led mechanisms.

Claims on the superiority of private provision as an innovation and diversification strategy also remain an open question. As argued extensively by Lubienski (2003, 2006, 2009), there is evidence that privatization arrangements do not necessarily promote pedagogical innovation, and that market pressures are likely to operate as standardizing forces. This would be a consequence of the fact that, in certain contexts, competitive dynamics lead schools to adopt teacher-centered pedagogies and back-to-basics curricula as a way to raise or preserve their reputation and respond to parental preferences for traditional instruction practices (Glatter et al., 1997; Hastings 2005). Experimentation is thus confined to management practices or marketing strategies (Lubienski, 2006; Lake, 2008).

The impact of PPPs in relation to educational equity appears one of the areas in which existing evidence is more consistent in its conclusions. A wide range of studies conducted in different contexts show that education systems where private provision and market mechanisms (i.e., school choice and competition) play a relevant role tend to exhibit higher levels of inequality (see for instance, Alegre and Ferrer, 2010; Elacqua, 2012, OECD, 2012, or Dumay and Dupriez, 2014). According to these studies, this occurs since the engagement of private providers tends to exacerbate dynamics of school segregation and social stratification among schools, and frequently leads to a rise in student performance inequalities. The high levels of school segregation and social stratification associated with education privatization are, in turn, explained by the fact that market pressures incentivize schools to develop practices of school selection, facilitate the exodus of better-off families from local schools, and are conducive to greater levels of social closure (Zancajo and Bonal, 2021). In addition, market mechanisms tend to translate into an unequal distribution of economic and human resources across schools, which also contributes to amplify social and educational inequalities (Zancajo et al., 2021a).

Once again, the negative impact of private education provision in terms of equity appears to apply to different contexts and privatization modalities. Although private provision has contributed to the expansion of primary and secondary education in some low-income countries (Pedró et al., 2015), evidence shows that this has been at the cost of increasing educational inequalities. From a comparative perspective, a study analyzing the effect of educational PPPs on equity in countries of the Global North and

South found that PPP schools reinforce social disparities in all types of contexts (Baum, 2018). Also, a recent comparative study focusing on European countries concludes that educational systems with a large share of privately managed schools tend to have higher levels of academic segregation – an effect that intensifies in systems with low public investment in education (Eurydice, 2020). In relation to LFPS in developing countries, different researchers have found that the emergence of these schools is likely to reinforce and deepen school segregation (see Macpherson (2014) for an overview of these debates).

Regulation and governance of private provision

The increasing evidence about the negative impact of privatization on educational equity has fostered numerous debates. The most relevant international organizations in the education policy field have actively engaged in such debates, and most of them increasingly problematize the relationship between pro-privatization policies and educational inequalities. UNESCO has warned on different occasions about the risks that privatization and marketization represent for equity.

The reproduction and possible exacerbation of inequalities of learning opportunities resulting from privatization in all its forms raises important questions about the notion of education as a public good and about the role of the State in ensuring the right to education. (UNESCO, 2015, p. 75; see also UNESCO, 2017 or Moummé and Saudemont, 2015).

The OECD has also produced extensive evidence on the relationship between private provision, pro-market policies (e.g., school choice or competition between schools) and educational inequalities (OECD, 2011, 2012, 2014, 2019). Even the World Bank, one of the most active international promoters of education privatization, has recently recognized that the rapid increase of private provision, can have unintended negative consequences in terms of equity, particularly in the Global South (World Bank, 2018). This does not necessarily mean that these organizations advocate for reversing privatization and marketization policies, but to rethink them in order to make their supposed benefits (such as greater diversification and more freedom of school choice) compatible with certain equity standards. International organizations, but also civil society initiatives such as the Abidjan Principles, have pointed out the need to adopt effective regulations to address the equity challenges possessed by privatization and market mechanisms in education. It is expected that improving the regulatory frameworks of private provision and school choice will allow reconciling the presence of private school offer with education systems' equity objectives (OECD, 2017).

Despite this being a relatively emerging debate, it is possible to identify two main positions regarding the need to regulate the participation of non-state actors in the provision of basic education. On the one hand, a regulation strategy in which the State plays a central role in setting standards, enforcing them and sanctioning those providers that do not meet these requirements. This is, for instance, the position of the UN Special Rapporteur on the right to education, UNESCO or the Abidjan Principles (United Nations, 2015; Moummé and Saudemont, 2015; The Abidjan Principles, 2019). To these organizations, states should assume a more active role in ensuring that private providers comply with a series of regulatory standards, including equity standards, and respond to policy frameworks defined at the national level. On the other hand, the World Bank or the OECD are more aligned with market-based or post-bureaucratic regulatory strategies, in which the State is supposed to be mainly responsible for establishing learning goals and accountability policies, while school autonomy and information systems should be reinforced to ensure that private providers face the necessary incentives to provide quality education in an equitable way (Lewis and Patrinos, 2011; Baum et al., 2014; OECD, 2017).

Nonetheless, pro-equity policy recommendations and regulatory reforms differ importantly according to the policy framework in which private providers operate. Here, we distinguish between policy initiatives that focus on the regulation of PPPs (i.e., vouchers, charter schools, contracted schools) and those that focus on the regulation of private independent schools (including LFPSs).

The regulation of PPPs in education

In the case of PPPs, regulatory debates and initiatives have focused on three main dimensions: 1) authorization of private providers, 2) funding, and 3) school choice and admissions.

The regulation of the *authorization of new providers* is expected to provide the State with more control over the quality and the quantity of the private education supply. This type of control can minimize situations of oversupply, the unequal geographic distribution of publicly funded private providers, or the over-representation of private schools with certain religious, linguistic or philosophical orientations (Zancajo et al., 2021a). Whilst in most PPPs social demand has been traditionally the main criterion to authorize a new private provider for receiving public funding, the role of educational planning is becoming increasingly relevant. In countries such as Sweden, Chile or Belgium, recent reforms have given priority to educational planning criteria in the authorization of new private subsidized providers (Zancajo et al., 2021b). Similarly, different Southern countries such as Liberia, Colombia or Pakistan have established bidding processes and eligibility criteria to manage the access of private providers to PPP schemes (Cameron, 2020; Edwards et al., 2017; LaRocque and Sipahimalani-Rao, 2019; Bano, 2008).

The *funding of private providers* has also emerged as one of PPPs' core regulatory dimensions to address education inequalities (Elacqua et al., 2018; Levin, 2002; OECD, 2019). The calibration of public funding of private providers in PPP frameworks can contribute to better distributing educational opportunities across the education system, and disincentivize selective and

discriminatory practices among private schools. There is an international regulatory trend toward increasing the resources available by those schools that enroll higher levels of low SES students or that stop charging add-on tuition fees to families. Chile, Sweden or Uganda are examples of countries that have prohibited private subsidized schools from charging add-on tuition fees in exchange for additional public funding (Lachance, 2020; Treviño, 2018; Baum 2018). Whereas education systems such as the Flemish and the French-speaking communities of Belgium, England or Chile have adopted formula funding mechanisms or targeted funding schemes that provide higher levels of public funding to schools depending on the number of disadvantaged students they enroll (De Witte et al., 2019; Friant, 2016; Foster and Long, 2020; Elacqua and Santos, 2013).

School choice and admissions is the third core regulatory dimension within the debate about educational PPPs. Several international organizations have pointed out that the regulation of school choice is essential to guarantee that policies that provide families with the possibility to choose school do not harm equity, particularly in contexts that combine public and private subsidized provision (OECD, 2019; Eurydice, 2020). The rationale here is to ensure that school choice systems are equivalent for private and public schools, as well as to establish common priority criteria of admissions. In the past few years, the French-speaking community of Belgium or Chile have adopted controlled school choice systems (Cantillon, 2013; Muñoz and Weinstein, 2019), England has regulated the admission criteria for public and private subsidized schools (West et al. 2011) and the Flemish community of Belgium, India or Colombia have established admissions criteria that give priority to socially disadvantaged students (Cantillon, 2011; Morgan, 2017; Edwards et al., 2017).

To date, the evidence on the impact of PPPs' regulatory reforms is relatively scarce. Nonetheless, in general terms, it seems that these reforms have had little or no effect in improving equity. Available evidence shows that regulatory reforms focusing on more equitable school choice or funding have not significantly reduced the segregation and stratification dynamics that characterize most educational PPPs. Most impact evaluations of these reforms show for instance, that school segregation has remained stable or has only decreased slightly after the adoption of these reforms. However, as Boeskens (2016) and Verger et al. (2020) noted, PPP arrangements not allowing schools to turn a profit, select students or charge tuition add-ons to families tend to be less harmful in terms of equity. So, future pro-equity regulatory changes may be more effective if they pay attention to these aspects, and adopt a more systemic and multi-dimensional approach to the problem of educational inequality within PPPs.

Regulating independent schools: the case of LFPS

The important expansion of private independent schools in some low- and middle-income countries has also fostered the debate about the need to regulate this sector to avoid potential negative effects on equity. In the particular case of LFPSS, one of the main challenges for national governments is to guarantee that these schools are registered and certificated (Härma, 2021), assuming that this is a first and necessary step that will allow governments to control the growth of private provision and to guarantee that this sector contributes to providing equitable access to education to all students (Baum et al., 2018). However, available evidence shows that the efforts made by some governments to regulate LFPSS have not necessarily achieved the pursued objectives. For instance, Day Ashley et al. (2014) observe that, in many cases, governments fail to enforce regulations because of technical or resource constraints. The lack of effectiveness of LFPSS regulation is also explained by the fact that strict regulatory frameworks tend to incentivize LFPSS to remain — or even go — unregistered (Härmä and Adefisayo, 2013; Baum et al., 2018; Härma, 2021). Finally, corruption has been also reported as a factor that affects the capacity of governments to regulate LFPSS. In some contexts, LFPSS can easily overcome regulatory requirements by bribing government officials (Härma, 2019, 2021). As in the case of PPPs, the evidence available shows that despite the fact that some international donors pushed for higher integration of LFPSS in national education systems and national governments have made an effort to regulate the private sector in developing countries, strict regulations have not necessarily improved equity (Härma, 2019).

To conclude

Education privatization is a global phenomenon that manifests in multiple and changing forms. Understanding why and how education privatization happens, and how it evolves, requires acknowledging the simultaneously global and idiosyncratic character of privatization — avoiding sweeping generalizations, while transcending the mere sum of particularized accounts. It also requires expanding research efforts beyond the few country cases in which empirical investigation has tended to focus. Scholarship on the topic of privatization has privileged a relatively small number of country settings — with many countries and even entire regions remaining largely overlooked. This is all the more relevant considering that in some of these under-researched areas, such as the Middle East and North Africa or Eastern Europe, private enrollment has soared over the last years.

Education privatization features as one of the most contentious and polarizing issues in education research. The available evidence on the impact of privatization on learning outcomes yields mixed and inconclusive results, and is subject to multiple interpretations. However, there is growing consensus that market policies and the engagement of private providers in public education systems affect equity negatively. As it has been reported repeatedly, privatization tends to exacerbate school segregation, social stratification among schools and the marginalization of the most socioeconomically disadvantaged students. Private schooling presents thus a major challenge for policymakers and education planners across the globe. Gaining a richer understanding of these trends is particularly important given their social implications.

There is a growing global awareness about the need of addressing the challenges that privatization possesses on equity. A number of national governments are engaging in reform processes that aim to strengthen the regulation and oversight of private provision within their educational systems. Nonetheless, while public regulation and oversight are being portrayed as solutions to long-standing inequalities within educational systems with high percentages of private provision, such efforts might not work in all circumstances. Regulatory initiatives need to be multi-dimensional, tailored to the context and to advance together with the strengthening and dignification of the public sector, especially in low- and middle-income country settings, if they are to be effective and sustainable over time.

References

- Alegre, M.A., Ferrer, G., 2010. School regimes and education equity: some insights based on PISA 2006. *Br. Educ. Res. J.* 36 (3), 433–461.
- Aslam, M., Rawal, S., Saeed, S., 2017. Public-private Partnerships in Education in Developing Countries: A Rigorous Review of the Evidence. Ark's Education Partnerships Group, London.
- Ball, S.J., Youdell, D., 2007. Hidden Privatisation in Public Education. Education International, Brussels.
- Bano, M., 2008. Public Private Partnerships (PPPs) as 'anchor' of educational reforms: lessons from Pakistan. In: Background Paper Prepared for the Education for All Global Monitoring Report 2009. Retrieved from: <https://www.semanticscholar.org/paper/Public-Private-Partnerships-%28PPPs%29-as-%27anchor%27-of-Bano/c3b80e5fca86127eebaa7a9a913ecd5202a60721?p2df>.
- Bartlett, W., Le Grand, J., 1993. The theory of quasi-markets. In: Le Grand, J., Bartlett, W. (Eds.), *Quasi-markets and Social Policy*. Palgrave Macmillan, London, pp. 13–34.
- Baum, D.R., 2018. The effectiveness and equity of public-private partnerships in education: a quasi-experimental evaluation of 17 countries. *Educ. Pol. Anal. Arch.* 26.
- Baum, D.R., Cooper, R., Lusk-Stover, O., 2018. Regulating market entry of low-cost private schools in Sub-Saharan Africa: towards a theory of private education regulation. *Int. J. Educ. Dev.* 60, 100–112.
- Baum, D., Lewis, L., Lusk-Stover, O., Patrinos, H., 2014. What Matters Most for Engaging the Private Sector in Education: A Framework Paper (SABER. Working Paper Series, number 8). World Bank Group.
- Belfield, C.R., Levin, H.M., 2002. The effects of competition between schools on educational outcomes: a review for the United States. *Rev. Educ. Res.* 72 (2), 279–341.
- Boeskens, L., 2016. Regulating Publicly Funded Private Schools: A Literature Review on Equity and Effectiveness (OECD Education Working Papers, No. 147). OECD Publishing, Paris.
- Cameron, E., 2020. The sweeping incrementalism of Partnership Schools for Liberia. *J. Educ. Pol.* 35 (6), 856–870.
- Cantillon, E., 2011. School Choice Regulation in Practice: Lessons from Antwerp, Brussels and Ghent (Unpublished manuscript). Retrieved from: <https://rethinkingbelgium.eu/wp-content/uploads/2019/08/Cantillon.pdf>.
- Cantillon, E., 2013. Mixité sociale: le rôle des procédures d'inscription scolaire. In: Maystadt, P., Cantillon, E., Denayer, L., Pestieau, P., Van der Linden, B., Cattelain, M. (Eds.), *Le Modèle Social Belge: Quel Avenir? Éditions de l'Université Ouverte, Charleroi, Belgium*, pp. 847–864.
- Chubb, J.E., Moe, T.M., 1990. *Politics, Markets and America's Schools*. The Brookings Institution, Washington, DC.
- Day Ashley, L., McLoughlin, C., Aslam, M., Engel, J., Wales, J., Rawal, S., Batley, R., Kingdon, G., Nicolai, S., Rose, P., 2014. The Role and Impact of Private Schools in Developing Countries. Department for International Development/Evidence for Policy and Practice Information and Co-ordinating Centre, London.
- De Witte, K., Vitezslav, T., Holz, O., Smet, M., 2019. Financing Quality Education for All. The Funding Methods of Compulsory and Special Needs Education. Leuven University Press, Leuven, Belgium.
- Dumay, X., Dupriez, V., 2014. Educational quasi-markets, school effectiveness and social inequalities. *J. Educ. Pol.* 29 (4), 510–531.
- Edwards Jr., D.B., DeMatthews, D., Hartley, H., 2017. Public-private Partnerships, Accountability, and Competition: Theory versus Reality in the Charter Schools of Bogotá, Colombia, vol. 25. Education Policy Analysis Archives.
- Elacqua, G., 2012. The impact of school choice and public policy on segregation: evidence from Chile. *Int. J. Educ. Dev.* 32 (3), 444–453.
- Elacqua, G., Iribarren, M.L., Santos, H., 2018. Private Schooling in Latin America. Trends and Public Policies. Education Division Social Sector Inter-American Development Bank (No IDB-TN-01555). IDB Working Paper Series.
- Elacqua, G., Santos, H., 2013. Preferencias reveladas de los proveedores de educación privada en Chile: el caso de la Ley de Subvención Escolar Preferencial. *Gest. Polit. Publica* 22 (1), 85–129.
- Eurydice, 2020. Equity in School Education in Europe: Structures, Policies and Student Performance. Luxembourg: Publications Office of the European Union.
- Friant, N., 2016. NESET Country Report: Belgium (Unpublished manuscript). Retrieved from: <https://halshs.archives-ouvertes.fr/halshs-01392996>.
- Foster, D., Long, R., 2020. The Pupil Premium (Briefing Paper, No. 6700). House of Commons Library, London, UK.
- Friedman, M., 1994. The case for choice. In: Billingsley, K.L. (Ed.), *Voices on Choice: The Education Reform Debate*. Pacific Research Institute for Public Policy, San Francisco, CA, pp. 91–101.
- Glatzer, R., Woods, P.A., Bagley, C., 1997. Diversity, differentiation and hierarchy: school choice and parental preferences. In: Glatzer, R., Woods, P.A., Bagley, C. (Eds.), *Choice and Diversity in Schooling: Perspectives and Prospects*. Routledge, London, pp. 7–28.
- Gruijters, R.J., Alcott, B., Rose, P., 2021. Do private schools improve learning outcomes? Evidence from within-household comparisons in East Africa and South Asia. *Comp. Educ. Rev.* 65 (4), 640–666.
- Härmä, J., 2019. Ensuring quality education? Low-fee private schools and government regulation in three sub-Saharan African capitals. *Int. J. Educ. Dev.* 66, 139–146.
- Härmä, J., 2021. Low-fee Private Schooling and Poverty in Developing Countries. Bloomsbury Publishing, London.
- Härmä, J., Adefisayo, F., 2013. Scaling up: challenges facing low-fee private schools in the slums of Lagos, Nigeria. In: Srivastava, P. (Ed.), *Low-Fee Private Schooling: Aggravating Equity or Mitigating Disadvantage?* Symposium Books, Oxford, pp. 129–152.
- Hastings, J.S., Kane, T.J., Staiger, D.O., 2005. Parental Preferences and School Competition: Evidence from a Public School Choice Program (NBER Working Papers, No. 11805). National Bureau of Economic Research, Cambridge, MA.
- Hicks, T., 2015. Inequality, marketisation and the left: schools policy in England and Sweden. *Eur. J. Polit. Res.* 54 (2), 326–342.
- Hodge, G.A., Greve, C., Boardman, A.E. (Eds.), 2010. *International Handbook on Public-Private Partnerships*. Edward Elgar, Cheltenham, UK.
- Hoxby, C.M., 2003. School choice and school competition: evidence from the United States. *Swed. Econ. Pol. Rev.* 10, 9–65.
- Lake, R.J., 2008. In the eye of the beholder: charter schools and innovation. *J. Sch. Choice* 2 (2), 115–127.
- Lachance, A., 2020. Feedback effects and coalition politics in education reforms: a comparison of school voucher programs in Sweden and Wisconsin. *J. Comp. Pol. Anal.* 22 (5), 405–421.
- Languille, S., 2017. Public private partnerships in education and health in the global South: a literature review. *J. Int. Comp. Soc. Policy* 33 (2), 142–165.
- LaRocque, N., Lee, S., 2011. Non-state Providers and Public-Private Partnerships in Education for the Poor. UNICEF/Asian Development Bank, Bangkok.
- LaRocque, N., Sipahimalani-Rao, V., 2019. Education Management Organizations Program in Sindh, Pakistan: Public-Private Partnership Profile (ADB Briefs, No. 107). Asian Development Bank. Retrieved from: <https://www.adb.org/sites/default/files/publication/492646/adb-brief-107-sindh-pakistan-education-management-ppp-profile.pdf>.

- Levin, H.M., 2002. A comprehensive framework for evaluating educational vouchers. *Educ. Eval. Pol. Anal.* 24 (3), 159–174.
- Lewis, L., Patrinos, H., 2011. Framework for Engaging the Private Sector in Education. System Assessment and Benchmarking Education for Results. World Bank, Washington, DC.
- Lubienski, C., 2003. Innovation in education markets: theory and evidence on the impact of competition and choice in charter schools. *Am. Educ. Res. J.* 40 (2), 395–443.
- Lubienski, C., 2006. School diversification in second-best education markets. *Educ. Pol.* 20 (2), 323–344.
- Lubienski, C., 2009. Do Quasi-Markets Foster Innovation in Education? A Comparative Perspective (OECD Education Working Papers, No. 25). OECD Publishing, Paris.
- Macpherson, I., Robertson, S., Walford, G. (Eds.), 2014. Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia. Symposium Books, Oxford.
- Morgan, C., 2017. Public private partnerships and corporate social responsibility: needs for and impacts on education in India and Indonesia. In: Hasan, S. (Ed.), *Corporate Social Responsibility and the Three Sectors in Asia*. Springer, New York, pp. 199–224.
- Moumné, R., Saudemont, C., 2015. Overview of the Role of Private Providers in Education in Light of the Existing International Legal Framework. *Investments in Private Education: Undermining or Contributing to the Full Development of the Human Right to Education?* (Working Papers on Education Policy, 01) UNESCO, Paris.
- Muñoz, G., Weinstein, J., 2019. The difficult process in Chile: redefining the rules of the game for subsidized private education. In: Ornelas, C. (Ed.), *Politics of Education in Latin America. Reforms, Resistance and Persistence*. Brill, Leiden, the Netherlands, pp. 72–100.
- Musset, P., 2012. School Choice and Equity: Current Policies in OECD Countries and a Literature Review (OECD Education Working Papers, No. 66). OECD Publishing, Paris.
- Obinger, H., Schmitt, C., Traub, S., 2016. The Political Economy of Privatization in Rich Democracies. Oxford University Press, Oxford.
- OECD, 2011. Private Schools: Who Benefits? (PISA in Focus 7) OECD Publishing, Paris.
- OECD, 2012. Public and Private Schools: How Management and Funding Relate to Their Socio-Economic Profile. OECD Publishing, Paris.
- OECD, 2014. When Is Competition between Schools Beneficial? (PISA in Focus 42) OECD Publishing, Paris.
- OECD, 2017. School Choice and School Vouchers: An OECD Perspective. OECD Publishing, Paris.
- OECD, 2019. Balancing School Choice and Equity: An International Perspective Based on PISA. OECD Publishing, Paris.
- Patrinos, H.A., Barrera-Orsorio, F., Guáqueta, J., 2009. The Role and Impact of Public- Private Partnerships in Education. World Bank, Washington, DC.
- Pedró, F., Leroux, G., Watanabe, M., 2015. The privatization of education in developing countries: evidence and policy implications. UNESCO Working Paper on Education Policy number 2.
- Peterson, P.E., 1990. Monopoly and competition in American education. In: Clune, W.H., Witte, J.F. (Eds.), *Choice and Control in American Education, Volume 1: The Theory of Choice and Control in Education*. The Falmer Press, London, pp. 47–78.
- Rouse, C.E., Barrow, L., 2009. School vouchers and student achievement: recent evidence and remaining questions. *Annual Review of Economics* 1 (1), 17–42.
- The Abidjan Principles, 2019. Guiding Principles on the Human Rights Obligations of States to Provide Public Education and to Regulate Private Involvement in Education. Retrieved from: https://static1.squarespace.com/static/5c2d081daf2096648cc801da/t/5dc414bb9f409d285dc9abf2/1573131454068/Online+version_A4_WEB_COUV%2BTEXTE+THE-ABIDJAN-PRINCIPLES_Nov_2019.pdf.
- Tooley, J., 1996. Education without the State. IEA, London.
- Treviño, E., 2018. Diagnóstico del sistema escolar: las reformas educativas 2014–2017. In: Sánchez, I. (Ed.), *Ideas en Educación II. Definiciones en tiempos de cambio*. Ediciones Universidad Católica de Chile, Santiago, Chile, pp. 135–165.
- UNESCO, 2015. Rethinking Education. Towards a Global Common Good? UNESCO, Paris.
- UNESCO, 2017. Global Education Monitoring Report 2017/18. Accountability in Education: Meeting Our Commitments. UNESCO, Paris.
- United Nations, 2015. Report of the Special Rapporteur on the Right to Education (A/70/342). Retrieved from: <https://undocs.org/A/70/342>.
- Vere, A., 2007. Social Dialogue in the Education Sector: An Overview (ILO Working Papers, No. 256). ILO, Geneva.
- Verger, A., 2020. Privatization. In: *The Blackwell Encyclopedia of Sociology*. <https://doi.org/10.1002/9781405165518.wbeosp099.pub2>.
- Verger, A., Fontdevila, C., Zancajo, A., 2016. The Privatization of Education: A Political Economy of Global Education Reform. Teachers College Press, New York.
- Verger, A., Fontdevila, C., Rogan, R., Gurney, T., 2019. Manufacturing an illusory consensus? A bibliometric analysis of the international debate on education privatisation. *Int. J. Educ. Dev.* 64, 81–95.
- Verger, A., Moschetti, M., Fontdevila, C., 2018. The expansion of private schooling in Latin America: multiple manifestations and trajectories of a global education reform movement. In: Saltman, K., Means, A. (Eds.), *The Wiley Handbook of Global Educational Reform*. Wiley Blackwell, Hoboken, NJ, pp. 131–155.
- Verger, A., Moschetti, M.C., Fontdevila, C., 2020. How and why policy design matters: understanding the diverging effects of public-private partnerships in education. *Comp. Educ.* 56 (2), 278–303.
- Waslander, S., Pater, C., van der Weide, M., 2010. Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education (OECD Education Working Papers, No. 52). OECD Publishing, Paris.
- West, A., Barham, E., Hind, A., 2011. Secondary school admissions in England 2001 to 2008: changing legislation, policy and practice. *Oxf. Rev. Educ.* 37 (1), 1–20.
- Woessmann, L., 2005. Public-private Partnerships in Schooling: Cross-Country Evidence on Their Effectiveness in Providing Cognitive Skills (Program on Education Policy and Governance, Research Paper 05-09). Available at: <https://www.innovations.harvard.edu/sites/default/files/PEPG-05-09woessmann.pdf>.
- World Bank, 2018. World Development Report 2018: Learning to Realize Education Promise. International Bank for Reconstruction and Development/The World Bank, Washington, DC.
- Zancajo, A., Bonal, X., 2021. Education markets and school segregation: a mechanism-based explanation. Compare. <https://doi.org/10.1080/03057925.2020.1858272>.
- Zancajo, A., Fontdevila, C., Verger, A., Bonal, X., 2021. Regulating Public-Private Partnerships, Governing Non-state Schools: An Equity Perspective (Background paper prepared for the 2021 Global Education Monitoring Report). Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000380092.locale=en>.
- Zancajo, A., Verger, A., Fontdevila, C., 2021b. The instrumentation of public subsidies for private schools: different regulatory models with concurrent equity implications. *Eur. Educ. Res. J.* <https://doi.org/10.1177/14749041211023339>.

Making the public private: the privatization of public education

David Hursh and Zhe Chen, University of Rochester, Rochester, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Charter schools, vouchers, and the privatization of education	183
The rise of testing and the demise of the social good	184
From privatization of education to privatization in education	187
Parents opting children out of the tests; teachers going on strike	188
Develop a strong base using social media	189
The rise of the unthinkable: teacher strikes	190
Concluding ideas	190
References	191

Across the globe, K-12 schools are increasingly privatized as for and non-profit corporations play a larger role in teaching and assessing, or take over whole schools (Williams and Hogan, 2021; Verger et al., 2016). Here, we describe some of the central means through which privatization is occurring. We focus more on the US than on other countries because the US has been at the forefront in implementing neoliberal reforms and a leader in what Finish educator Pasi Sahlberg (2016) labels the Global Education Reform Movement or GERM.

In researching education privatization, we examine two distinct aspects of public schooling. First, some privatizers are interested in exogenous privatization, which “promotes the emergence and expansion of private providers in the education sector” (Verger et al., 2016, p. 9). The two major examples of this are, first, the privatization of whole schools—such as through charter schools—and the privatization of whole school districts—such as the dissolution of most of the New Orleans school district after Hurricane Katrina (Buras, 2014). The second form of privatization is endogenous, or the privatization of what occurs inside schools, including standards, curriculum, and assessments (Verger et al., 2016) as exemplified by the Pearson corporation’s efforts to integrate teacher education, teaching, assessment, and professional development into one seamless experience throughout educators’ careers and marginalizing teachers, parents, students, and community members.

Because privatization occurs in a myriad of ways, we will begin by examining one of the two dominant ways in which privatization occurs across the globe: the privatization of whole schools or school districts—whether they are called charter schools in the US, academies in England Eyles et al. (2017), or low-fee schools in developing countries.

However, because there is disagreement over what counts as a public or private school and why we conclude that charter schools are private, even though some charter schools identify themselves as a “public charters schools,” we begin by providing our criteria for what counts as public. Then, we will place the current debates within the context of the history of education in the US. In particular, we suggest that beginning in the late 1800s, two major conflicting education paradigms emerged. Historian Herb Kliebard (2004) described one paradigm as building on “scientific efficiency,” which were modeled after the factory assembly lines that were central to the new industrial age. Scientific efficiency emphasized the quantification of everything, uniformity, and competition. Those promoting a scientific or social efficiency approach to education incorporated standards as determined by “experts” (Snedden, 1924), intelligence testing, and student tracking (Yerkes, 1919).

The second paradigm, which Kliebard described as “social reconstructionist” and for which John Dewey was the most well-known proponent, valued democracy over efficiency. Rather than asking corporations what they want students to learn, schools should aim to develop democratic citizens and require corporations to adapt to citizens’ demands for a critical education (1915). The goal of democratic schooling requires focusing on the whole child and creating interdisciplinary curriculum and authentic assessments.

We will then turn to our two examples of exogenous privatization. The first example examines exogenous privatization with the corporate takeover of whole schools or school districts and the second exogenous privatization with the corporate takeover of curriculum, teaching and assessment.

Charter schools, vouchers, and the privatization of education

Here we need to be clear regarding what makes a school public versus what makes a school private. This is a distinction often debated when discussing charter schools, which are schools chartered by the state, often administered by private corporations, and primarily funded by local, state, and federal governments along with philanthropies supportive of charter schools, such as the Bill and Melinda Gates Foundation and the Walton Family Foundations. Charter school proponents often argue that they are public schools because they are primarily government funded.

However, unlike public schools, which are governed by an elected school board, charter schools are typically governed by an appointed board. Further, some charter schools do not welcome visitors in the schools. Moreover, but it is not an issue we can

go into depth here, the lack of transparency may contribute to the high percentage of instances of charter schools that have had funding or staffing irregularities (Lecker, 2013).

Forty-five states and the District of Columbia have laws permitting charter schools. However, states may be promoting charter schools not because they are pedagogically sound but because in these financially challenging times, Race to the Top (RTTT) legislation offers additional federal funding to states that legislate the creation of charter school.

Many charter school proponents claim that charter schools are more effective than traditional public schools, but that evidence is questionable. First, the most reliable and probably most often cited studies, conducted by the Center for Research on Education Outcomes (CREDO), states that based on scores on standardized tests about three-fourths of charter schools are the same or worse than public schools. Second, we should not only consider student achievement, but also the negative effect charter schools have on efforts to desegregate schools. Because charter schools are not required to enroll all applicants, including students with disabilities, English language learners, and students with behavioral problems, they tend to enroll fewer students with learning disabilities and students living in poverty than in the typical school in the school district in which the charter school is located. Consequently, given that test scores are closely related to family income, those that score better on the standardized exams are more likely to congregate in a few schools.

Given the information questioning the effectiveness of charter school and that it exacerbates racial and class segregation, one might wonder why the number of charter schools and students enrolled continues to increase. However, charter school parents may not be aware of the many ways that test scores are misleading (Hursh, 2015) and how charter schools exacerbate segregation. Rather, they are in search of that one good classroom seat for their child.

Voucher programs are similar to charter schools in that both receive most of their funding directly from the government based on their enrollment. Voucher programs differ from charter schools in that families are provided with a voucher that can be used to pay for some or all the tuition to a private or parochial school. Consequently, depending on a family's wealth, the cost of tuition, and the value of the voucher, voucher programs may benefit from a few to most of the students enrolled in the private school. In the US, charter schools continue to increase in number while voucher programs remain limited. According to the National Conference of State Legislatures fourteen states and the District of Columbia have passed school voucher laws and none of them are open to all students. Four of them are only up to students with a disability, two states only provide vouchers to students from low-income families, and two states restrict eligibility to students who live in either the Cleveland or Milwaukee metropolitan areas (Pedroni, 2007).

Some countries have the equivalent of charter schools (UK academies). Others, particularly in developing countries Williams and Hogan (2021), have low-fee private schools, which are attractive to entrepreneurs because of the possibility of making a profit.

A second form of privatization—endogenous—is the privatization of what occurs inside of schools. This kind of privatization is facilitated by the increasing use of technology in schools—whether the use of mass administered tests beginning with Robert Yerkes' 1919 World War IQ tests and quickly picked up by educators and continuing up to the present with efforts, such as the Common Core standards and exams to digitize and place on computers not only the standards, curriculum, and assessments, but also surveil everything that teachers and students do. Therefore, it is useful to provide a quick summary of how the increasing use of classroom technology is enabling corporations to take over all aspects of teaching and learning.

In this article, we examine the increasing privatization of education and suggest that it is part of the long struggle between two conceptions of education: education for the advancement of the economically productive individual or for the whole child within the common good.

By turning those processes over to private interests, it shifts control over what occurs in schools away from teachers, students, parents, and the local community and toward corporations, and for-profit and nonprofit organizations. Teachers would increasingly be forfeiting control over what occurs in schools to the corporate employees developing the standards, curriculum and assessments.

Bill Gates and the Pearson Corporation both supported the Common Core standards, curriculum, and assessments with the hope that states would adopt the Common Core curriculum and assessments. This would facilitate the corporate takeover of what occurs in schools (Schneider, 2015; Strauss, 2014a,b). Stephen Ball, 2012, in *Global Education Inc.: New Policy Network and the Neoliberal Imaginary* briefly describes the conflict over which societal vision would become the dominant social imaginary (Rizvi and Lingard, 2009) in each era: social efficiency in the period after World War I, social democracy from the 1930 to 1970, or neoliberalism from the 1970s to the present. Then, we will describe how some teachers, students, and parents are resisting the neoliberal agenda in education of privatizing by resisting the increased standardization of schools through high-stake testing and funding chart schools.

The rise of testing and the demise of the social good

The current agenda for privatizing teaching and learning within schools complements the agenda promoting social efficiency (Kliebard, 2004). Social efficiency in education arose as a movement in the early 1900s as proponents sought to make economic efficiency central to in determining the goals and process of education. Schools were modeled after the new more efficient factory assembly lines to "eliminate waste" in education (Bobbitt, 1912; Snedden, 1924). Standardized tests would inform teachers whether and what students were learning. Therefore, in the next section we return to the conflict described between the two opposing education paradigms of efficiency and social reconstruction, to suggest that the current emphasis on privatization supports the agenda to privatized education.

Many of the characteristics of social efficiency were adopted by schools and continue today—frequent measurement and quantitative data, sorting students into different vocational tracks (manual, skilled, and professional) like on factory assembly lines and “working children up as raw material” to a finished product (Bobbitt, 1912, p. 269).

Some of the teachers and students who opposed the social efficiency approach explicitly built on John Dewey’s ideas (1938/1997) were not interested in educating children to meet the needs of industry but rather to prepare students to be critical democratic citizens. Industry should adjust to the democratic needs of students rather than have students adjust to industry. Dewey wrote that the primary purpose of all social institutions, including workplaces, is to facilitate personal growth. Dewey disagreed with the idea that students were to meet the needs of businesses and stated that he was not interested in adapting students to the “existing industrial regime,” but, instead, wanted the existing industrial regime to adapt to democratic citizens, including students (Dewey, 1915, p. 42).

Although Dewey’s ideas regarding pedagogy seem more honored in their breach than in their practice, the idea that public schools should serve all children and educate the whole child persists and forms the basis for many of the progressive policies of the 20th century.

Progressives like John Dewey or social democrats like Franklin Delano Roosevelt supported expanding public education and other public programs, such as unemployment insurance, social security, health care, and higher education. To distinguish this view as opposed to less progressive members of the Democratic party, we refer to the former as social democrats and the latter as Democrats. However, while social democrats favored expanding social welfare programs, neoliberals—an opposing group of social theorists—began to push back, arguing that rather than expanding government, everyone would benefit from downsizing, if not eliminating, government. Rather than having the government make policy decisions, neoliberals argued that all decisions should be based on individual market choices which would be more “objective.”

Beginning in the 1930s and 40s, neoliberals established several organizations to investigate and promote their ideas, including The Center for Public Choice at the University of Virginia, the Mont Pelerin Society in Vienna, and the economics department at the University of Chicago. Since then, as in much of the world, neoliberalism has become the accepted and dominant way to conceptualize society, so much so, that most people do not even recognize that it has become the default social theory guiding societal decisions. Many people have come to agree with the former UK Prime Minister Margaret Thatcher (1979–1990) that There Is No Alternative (TINA) to neoliberalism (Leitner et al., 2007). Neoliberalism has come to dominate both society as a whole and education in particular, which Finnish educator Pasi Sahlberg has labeled the Global Education Reform Movement or GERM (year).

The current move toward privatizing public schools as a whole or in part comes after a long struggle over the aims and organization of schools which social democrats and other progressives seemed to be winning until the 1970s. Until then, public schools, in the United States, were becoming more inclusive as many schools and school districts were becoming (albeit, with some resistance) more integrated and educating students with a wider range of abilities. Similarly, our understanding of how children learn was giving more significance to students’ interests and abilities. Critical educators, such as Apple (1982), Freire (1970) began to challenge the dominant textbook-based curriculum as inadequate, and looked to Gilborn (2008), and others began to argue for developing curricula that responded to the needs and interests of adults and children at their specific time and place. In addition, the science of teaching and learning, reflected in the National Research Council’s *How people learn: Brain, mind, experience, and school* (2001), criticized traditional scientific efficiency approaches to teaching and learning and replaced them with more Deweyan approaches.

Similarly, in the two decades after World War II, social democrats, such as presidents Roosevelt, Kennedy, and Johnson, successfully expanded the government’s role to include providing social services such as social security, health care, and education from preschool through university. Consequently, by the mid 1960s, it was becoming accepted that all children, regardless of race, (*Brown vs. Board of Education*) deserved a quality education.

During this time period, many of the social democrats built on John Dewey’s notion that schools should be democratic institutions composed of communities of learners preparing students to live in a democratic society for the benefit of the common good. For example, in New York State, in the late 1980s, some two dozen public schools came together to form the Performance Assessment Consortium Schools where graduation from secondary school required not passing standardized tests but demonstrating proficiency in 13 areas. The state department of education granted the consortium schools waivers so that their students did not have to take the statewide Regents exams, giving the schools flexibility in how they would achieve the standards.

However, just as proponents of social democratic policies and practices gained dominance in the middle of the 20th century, neoliberals who opposed social democratic policies began to replace social democratic policies with neoliberal policies. In the United States, Milton Friedman, an economics professor at the University of Chicago, did the most to spread neoliberal ideas not only in the academy but also in the popular media, such as his best-selling book *Free to Choose: A Personal Statement* (1980) where he called for the end of public or “government schools,” and privatizing public services so that they would have to compete for clients (or in the case of schools: students), and providing “objective assessments” either in monetary terms, ratings, or, in education, test scores.

Globally, neoliberal education policies have focused on privatizing public schools as charter schools, funding private schools through governmental vouchers, developing education standards and assessing students, teachers, and schools through standardized tests, and handing over to private corporations much of the decision making regarding what is to be taught and how (Verger et al., 2016).

Friedman was able to put his ideas into practice when in Chile, in 1973, General Pinochet led a military coup overthrowing the democratically elected President Salvador Allende. Pinochet enthusiastically adopted neoliberal policies not only in education but throughout society, such as in transportation and healthcare. With the guidance of Friedman and his students at the University of Chicago (dubbed the “Chicago boys”) Pinochet drastically reduced public funding for education at all levels. Funding to schools was cut so much that some teachers (who came to be known as “taxi teachers”) had to teach up to fifteen classes a day in three different schools.

Collins and Lear (1995) write regarding Chile that “per-pupil school vouchers, the encouragement of for-profit schools ... all decreed in the name of choice—have greatly magnified the differences in educational opportunities and results” (pp. 8–9).

Later, Ronald Reagan, as president of the United States (1981–1989), and Margaret Thatcher, Prime Minister in the United Kingdom (1979–1990), allied with one another with the goal of shrinking government and cutting its funding so that it would have a diminished role in society. They both worked to reduce the power of unions and to privatize whatever could be privatized. Thatcher’s anti-government views were reflected in her 1987 statement that “there is no such thing as society, only people and their families” (Thatcher, 1995, pp. 626–627), while Reagan (1986) opined that “the most terrifying words in the English language are ‘I’m from the government and I’m here to help.’”

Over the last four decades neoliberalism has become in the United States and in many other countries the predominant view of how society works, what Rizvi and Lingard (2009) refer to as the “social imaginary.” Neoliberalism has become so dominant that neoliberals need not refer to themselves as neoliberals, preferring instead to describe themselves as free market advocates who insist that there is no alternative to the neoliberalism of education.

The basic tenets of neoliberalism are reflected in the quotes above from Thatcher and Reagan: they argued that neoliberal policy makers could make society more efficient by reducing the size or eliminating through privatization social services such as education, social security, transportation, and health care. As private organizations, social services, such as schools, should compete with one another for clients. Individuals would be responsible only for themselves or their families and be held accountable through objective criteria.

These tenets form the rationale for the last almost four decades of education policies, whether from Republicans or Democrats, beginning with *A Nation at Risk* 1993, and continued with *RttT* (2015). *RttT* required states to amend their constitutions to allow for charter schools. These acts blamed public education for the failures caused by economic inequality and instituted regulations that favor privatization.

These acts demonstrate the ways local, state and federal governments contribute to the disasters by placing private above public interests. For example, because inadequate funding of the infrastructure contributed to the temporary closure of most of the public schools in New Orleans, the state government, demonstrating how capitalism often uses so called “natural disasters,” to privatize the public, dissolved all the public schools in New Orleans, fired all the teachers, and replaced the public schools with charter schools (Buras, 2014).

Charter schools were originally proposed by Albert Shanker, long-time (1974–1997) president of the American Federation of Teachers, one of the two teachers’ unions in the United States and the union more likely to represent teachers in urban school districts (Kahlenberg and Potter, 2015). Shanker’s original proposal was for teacher-led semi-autonomous schools composed of six or more teachers, which would need to be approved by panels “run jointly by districts and the local teacher union affiliates” (Ravitch, 2020, p. 131). These schools would serve as a source of innovation that would improve all the schools, much like New York’s “performance assessment schools” (<http://www.performanceassessment.org>) that formed in 1998 as teacher-run small schools.

The first charter school in the United States opened 1992 in Minneapolis, Minnesota but differed from what had been originally proposed. Rather than encouraging small groups of teachers to propose small teacher-run schools, the new regulations encouraged entrepreneurs to start charter schools, approved by the state, not local communities, and that the rules favored the involvement of private companies rather than community organizations. Consequently, in the end, charter schools are more like private than public schools.

The number of charter schools in the United States has been increasing, although at a lower rate than five years ago (Ravitch, 2020). Charter schools, vouchers, standards, standardized testing, and top-down often computer-generated curriculum are central features of neoliberal policies. A major reason for the formation and increasing number of these neoliberal policies is the state and federal support for them as reflected in *No Child Left Behind*, *Success for All*, among others. Success for All requires states, if they want to be eligible for federal funding, to adopt laws promoting the establishment of charter schools (Ravitch, 2020). Currently, all but four states have laws authorizing charter schools and charter schools enroll about five percent of the 50.7 million students in K-12 education (Ravitch, 2021).

Under Arne Duncan, Obama’s Secretary of education, the federal government aimed to incorporate neoliberal principles into the governance of education by incorporating neoliberal education reformers into education policy making.

One came directly from the Gates Foundation, and several came from the Gates funded New School Venture Fund, including the Under Secretary of Education, Ted Mitchell (Schneider, 2014). Besides privatizing schools through charter schools and vouchers (exogenously), Obama appointees aimed to take control of assessments and curriculum (endogenously). Arne Duncan’s chief of staff, Joanne Weiss, clearly laid out the neoliberal strategy here:

The development of common standards and shared assessments radically alters the market for innovation in curriculum development, professional development, and formative assessments. Previously, the markets operated on a state-by-state basis, and often on a district-by-district basis. But the adoption of common standards and shared assessments means that education entrepreneurs will enjoy national markets where the best products can be taken to scale.

Weiss (2011), cited in Ravitch (2020, p. 24).

Both exogenous and endogenous reforms are being resisted, as parents, students, educators and community members have pushed back against neoliberal reforms.

While the number of charter schools continues to rise, beginning in 2015–16, the rate of increase began slowing (Ravitch, 2020), for which there may be several explanations, including a decline in the reputation of charter schools (Miron et al., 2021). Charter schools initially claimed that their students outperformed students in the public schools. But, numerous studies have shown that the difference can be attributed to differences in student enrollment. Charter schools enroll a much smaller percentage of students with learning disabilities (Stern et al., 2015) or who are English Language Learners (Finnigan et al., 2004). Further, while the public schools must enroll all the students who show up at the schoolhouse door, charter schools can be selective in who they admit (California Teachers Association, 2019) and can expel those whom they deem no longer fit (Frankenberg et al., 2010). In fact, when demographics are held constant, charter schools perform the same or worse than comparable public schools (Chen, 2020) and this doesn't recognize that charter schools typically receive more funding per student than the public schools in the district where the charter school is located, nor that charter schools often receive additional funding from conservative organizations such as the Walton Family Fund, which between 2016 and December 2019, channeled more than \$185 million to over 100 charter schools to improve the charter schools' real estate holdings effort (Harden, 2020).

In addition, a high percentage of charter schools are plagued by scandals including the manipulation of test scores (Hursh, 2015) and misuse of funds. Finally, parents and the public are becoming aware that charter schools contribute to growing inequality, either by contributing to segregation (Wells, 2019) or by taking funds originally intended for public schools (Ravitch, 2020).

Not only do charter schools perform on average worse than public schools, they are increasingly privatized as charter schools are either begun by or taken over by private education management organizations (EMOs). This shift undermines the claims of charter school proponents "that charter schools would be locally run, innovative, autonomous, and highly accountable" (Miron et al., 2021, p. 1). Profiles of EMOs reveal that more than half of the nation's charter school students are now enrolled in privately operated charter schools (Miron et al., 2021).

From privatization of education to privatization in education

While states and municipalities vary in whether they permit charter schools or vouchers, and if they do, the percentage of students enrolled in either program, almost all the states and their educators and students are affected by the privatization of what occurs in schools or *endogenously*, such as private corporations taking over the creation of standards, standardized tests, and the curriculum. Private corporations have played a major role in this development, with companies like Pearson taking control of developing standards, assessment, and curriculum. Williams and Hogan (2021) write:

What we have observed is a huge effort to create multisector coalitions of public-private partnerships and networks in which commercial actors play a key part, and we have seen the multilateral policy influencers, wealthy tech philanthropies and government departments to global technology corporations, edu-business, venture capital investors, and public institutions (p. 66).

The most recent wave of neoliberal reforms been typically justified as the best way to improve student learning in schools and to close the "so-called" achievement gap (Ladson-Billings, 2006). In the United States, beginning with A Nation at Risk in 1983, neoliberal reformers have argued that schools and students have failed because schools, teachers, and students have not been held accountable and that teachers have failed to hold students to high standards. Consequently, argue the authors of A Nation at Risk, we need standards and standardized tests so that we have "objective" measures of student learning with serious consequences for failing to achieve them.

With the passage of the No Child Left Behind Act in 2001, real consequences were added to the testing requirements. Students, teachers, and schools were evaluated based partially on their students' aggregated test scores and these scores would be made public. Schools would have to make Adequate Yearly Progress (AYP), which would be a rising minimum aggregated passing rate for students on a standardized exam in math and English/language arts. Every state determined their own passing rate, but they all had to achieve a 100% passing rate by 2014, which was patently impossible. Schools failing to achieve AYP could be forced to reorganize, close, or privatize as charter schools (Ravitch, 2013).

Given that the NCLB Act was signed by President George W. Bush, in 2001, he knew that by 2014 he would be out of office and, therefore could not be blamed for the inevitable high but misleading "failure" rate of schools. Instead, Obama was president and not wanting to preside over and be blamed for a massive "failure" of the schools, he made it possible for states to receive waivers

from the passing requirements if they would institute further “reforms,” which included more privatization, and the Common Core curriculum and standardized tests.

However, there is no evidence that the neoliberal reforms were achieving their goals; instead, the evidence strongly suggests that the reforms not only did not improve teaching and learning but damaged it (Ravitch, 2013; Winerip, 2003, 2011). Consequently, they, like many educators, parents, and students, began to call for limits on the number of charter schools, the end of voucher programs, and asserted that teachers should be treated as professionals. Here we will focus on New York State, where the state and the federal government together have required increased testing tied to standards and curriculum developed initially by Pearson, but subsequently by other corporations.

The first increase in the testing requirements occurred in the late 1990s when the state imposed the new graduation requirement of passing five Regents exams: one each in math, science, and English, and two in history. Because Commissioner Mills wanted to appear to have high standards, but also not significantly harm the graduation rate, tests and test scores were manipulated to achieve the right balance—not resulting in so many students failing that it would look like the schools were getting worse—not so few that it looked like the reforms were not rigorous. As reported elsewhere—Winerip (2003) and Ravitch (2013)—this resulted in the commissioner setting the cut score low in the introductory high-school courses (remember, students needed to only pass ONE exam in each subject, so the ninth grade tests could be scored easier).

At the federal level, the first test results of the Common Core tests came in 2013. Proponents of standardized testing hoped to develop unified Common Core standards, curriculum, and tests delivered to classrooms via Microsoft technology. Ideally, for Gates, was the possibility of creating national standards for math and language arts. Gates funded over 180 organizations to develop the Common Core standards. The New York State Commissioner of Education, John King, stated that because of where they placed the cut score—the score needed to pass—only about 5% of the students in the Rochester City School District and only 30% statewide “achieved proficiency” or passed the standardized tests (Cody, 2014; Strauss, 2014a,b). Not surprisingly, since King set the cut score that students needed to reach to achieve proficiency, that percentage achieving proficiency was what the Commissioner desired. When asked what was the value of the test when the Commissioner pre-determined the percentage passing, King responded that at least now they knew which districts needed intervention, as if they did not know that before the tests were given (Hursh, 2015).

Given that King and others gained power during the Obama administration, it may not be surprising that King is one of two central advisors to President Biden who recently pushed the president to renege on his campaign promise not to require states to give high-stakes tests in the middle of a pandemic. Again, when asked what was the value of standardized testing in the middle of a pandemic when simply opening the schools was a challenge, King claimed we would at least know where the learning gaps were.

Parents opting children out of the tests; teachers going on strike

Over the past few years, educators, parents, students, and community members have become more active in pushing back against the neoliberal reforms and working for the common good. As we will describe, both groups have had success and can serve as models for others. We will begin by identifying some of the ways they are similar and then turn to describing both in more detail. We will end? By pulling together the different arguments and the challenge of promoting the public good over the private gain. In describing the opt-out movement, we will focus on the movement in New York. While there have been parallel movements in 12 other states, New York’s opt out movement has been the most successful (Hursh et al., 2021). In describing teacher strikes, we will describe the rise of teacher resistance in West Virginia, Arizona, and Los Angeles.

Our research shows that the neoliberal reforms have been resisted by parent-led movements such as New York State Allies for Public Education (NYSAPE) and Long Island Opt Out (LIOO) and by teacher organizations, only some of which are unions. Both movements are similar in that they were willing, as Lisa Rudley, the head of NYSAPE stated, to “throw a wrench in the system” (interview by Hursh) by withholding their children from the test, and in the case of the teacher strikes, teachers withholding their labor.

Both movements have similar goals emphasizing fighting for the common good rather than just looking out for their own interests. Even though, over the last decade teachers’ salaries have barely increased, teachers have a broader set of demands, including better working conditions not just for themselves but for other employees, such as school bus drivers, teachers’ aides, and cafeteria workers (Blanc, 2019a,b). They are also demanding smaller class sizes, more school nurses, and more resources. Likewise, while parents are understandably looking out first for their own children, they are demanding not only the elimination of high-stakes testing but also promoting the welfare of the whole child through interdisciplinary teaching, providing meals for students, and recreation facilities.

The two movements also share strategies and tactics in that they are organizing both horizontally and vertically to develop widespread support so that they can influence the powerful. In the case of the opt-out movement, they have organized horizontally building coalitions that include over fifty local organizations initially asserting their political power at the local level by supporting candidates running for local school boards. In the case of teacher unions or organizations they worked to develop complementary relationships with parents and community members.

The opt-out movement and teachers’ unions have both used similar strategies using inexpensive or free social media to build a strong base. Central to their efforts are websites and Facebook pages. They also work at keeping the media informed of the issues. Having strengthened their base they can then work to elect people to higher positions in the state legislature, and the governorship.

We will expand on this later when we examine how teachers, often along with other employees—such as bus drivers and cafeteria workers—went on strike demanding not only higher salaries but also more classroom resources, better facilities and, in some cities and states, contesting the neoliberal education reforms have been contested by educators, political activists, parents, and students. Here, we offer the United States and New York, in particular, as sites where the resistance has been most successful. We begin by describing recent efforts by parents to opt their children out of the Common Core math and literacy exams in grades three through eight. The New York State education department made the exams high stakes by using them to evaluate teachers, students, schools, and school districts and to determine the amount of funding the schools would receive from the state. While some parents would be happy to only reduce the amount of time devoted to testing, many desire to completely eliminate the tests.

Initially, efforts to eliminate the Common Core consisted of parents, educators, students, and community members holding literally hundreds of hearings across the state, usually with the Commissioner or Chancellor as a participant, with the hope that those in power might be persuaded to reduce the amount of testing by revealing that the tests did not lead to the positive results they claimed. However, those in power typically feigned paying attention to parents and teachers, ultimately ignoring their pleas and continuing on the same path.

After several years of not being heard, parents began to conclude that they needed to force the issue by withholding their child from the exams. After all, proponents reasoned, even if only 10% of students failed to take the exams, and if the percentage varied between classrooms and schools, the exams would be invalidated and could not be a significant part of the metric for evaluating teachers, students, or schools, or so they reasoned. The pandemic has led to the tests being canceled, or being administered but with promises that the scores would not be used, which raised questions of why administer them?

While some parents were optimistic about achieving a high opt-out rate, doing so would be difficult. Some school administrators refused to inform parents that they could opt their children out. Other administrators let students opt out but instituted a “sit and stare policy” requiring students sit quietly at their desks while the tests were administered for several—up to 6 hours a day over several days and do nothing other than meditate. Furthermore, the federal department of education threatened school districts that if more than five percent of their students failed to take the exams, their federal funding would be withheld. This was particularly pernicious for urban high-poverty schools in the United States because their lack of resources often made it difficult for these students to get to school and because urban schools are more likely to receive most of their funding from the federal government, while typically wealthier suburban districts receive most of their funding from local property taxes. Consequently, the opt-out rate for New York’s big cities were lower than other parts of the state. But this did not necessarily mean that they were less interested in opting out. In many communities, parents and their children would have to remind teachers and administrators that they have the right to opt their children out.

Develop a strong base using social media

However, the opt-out rate exceeded most people’s expectations. In the three most recent years (2016–2019), around 20% of the students statewide did not participate in the tests, with much higher opt-out rates in the suburbs (NYSAP website). For example, the two counties east of New York City on Long Island have had an opt-out rate of 50.7% on the English/language arts exam over the last three years. Each year the opt-out rates for schools and school districts were reported and debated, which focused attention on the tests and the resistance.

In addition, LIOO and NYSAP dropped the strategy of reducing or eliminating testing by appealing to those in power, focusing, instead, on removing those in power. They began by a strategy of first developing a strong base of local supporters—working horizontally—who could then organize to work vertically to influence those in power (McAlevey, 2016). They began by developing liaisons to local districts who would work with teachers, administrators and school board members to replace high-stakes testing with more whole-child approaches to teaching and learning. Then, as they gained allies at the local level, they worked vertically to elect state legislators who would, in turn, support progressive education policies and name progressive educators to the Board of Regents, which is responsible for all aspects of education, including schooling from infancy through graduate school, and from museums to libraries. It is also the Regents who choose one member to chair the board as chancellor. They also appoint the commissioner of education, who implements the policies by directing the department of education. At this moment in time the strategy has resulted in replacing the chancellor, commissioner, and regents with representatives who are more likely to endorse progressive policies. However, it is the federal government where the Biden (Ujifusa, 2021) administration continues the Obama administration’s policy of high-stakes testing (Ravitch, 2013).

At the same time that the parent-led opt-out movement was achieving impressive results, including defeating the Bill Gates’ led effort to install in schools the student data gathering technology InBloom (Bulger et al., 2017), teachers were beginning to also push back against many of the same things as the opt out movement. Both are pushing back against high-stakes standardized testing for the way in which they shift control over education away from teachers and students to corporate employees and private foundations. Indeed, under the false claim that sitting children in front of a personal computer enabled students to experience “personalized learning,” (Boninger et al., 2019; Hursh et al., 2021) what students were increasingly receiving was curriculum delivered based on computer algorithms.

The rise of the unthinkable: teacher strikes

In the same way that it took withholding children from school to force those in power to acknowledge the concerns of parents, students, teachers, and others, teachers have learned that they can withhold their contribution to the educational enterprise by striking. Here we will show that teachers, students, and parents were unable to affect educational policy through the usual means of lobbying and demanding hearings. Instead, in parallel to parents withholding children from the standardized exams, teachers began to withhold not their children from the exams, but their labor. They began to go on strike ([Blanc, 2019a,b](#)).

Blanc, whose book *Red State Revolt: The Teachers' Strike Wave and Working-Class Politics* (2019), is the most extensive study of the teachers' strikes in West Virginia, Oklahoma, and Arizona. He reports that while in most states teachers' pay was already low and falling even further behind other professions, teachers were more likely to leave the profession not over pay but because of decreasing autonomy as standards and curriculum were being developed by other than classroom teachers.

Until the recent wave of strikes, teachers rarely went on strike, partly because in many states it was illegal to strike. But, a combination of low pay and the fear of increasing privatization of education through charter schools and corporatization of testing and curriculum led desperate teachers to risk jail or other punishments by striking. In addition, teachers formed alliances with other school employees and parents. By allying with teacher aides, bus drivers, cafeteria workers, and others, and by arguing for more classroom resources, they revealed how dire the situation was for all and how they were stronger together.

Once teachers observed other teachers successfully striking, such as in West Virginia, where the teachers dared the state to arrest them for striking, knowing that the public was likely to be disturbed by seeing teachers jailed, teachers in other states began to do the same. As each strike resulted in improved working conditions for all, the teachers and other employees began to realize that there is power in numbers. As they gained power, they realized they had the support of the community and the ability to demand changes.

By devoting years to organizing and then calling for strikes, they demonstrated that "ordinary people have two major forms of power—our numbers—there are more of us than them—and our ability to withhold our participation in a system that denies us fairness, dignity and health" ([Schirmer, 2021](#), p. 250). Where teachers have been successful, writes [Schirmer \(2021\)](#), they emphasized organizing, which went beyond advocacy and mobilizing by recruiting leaders out of the rank and file and involving large numbers of people. Jane [McAlevy \(2020\)](#), a labor organizer who was central to the success of the Los Angeles teachers' 2019 strike, wrote that organizing, especially organizing with an eye toward building strikes, activates both forms of power. Successful labor and social movements often use a combination of advocacy, mobilizing and organizing.

At the same time that the parent led Opt-Out Movement in New York was growing and achieving previously unthinkable results, teachers were starting to push back against many of the same things as the opt out movement. Both were pushing back against high-stakes standardized testing for the way in which they shifted control over education away from teachers and students in the classroom toward corporate employees and private foundations. Indeed, under the false claims of "personalized learning," what is taught and how is increasingly determined neither by the teacher nor the student but by nameless computer programmers ([Hursh et al., 2021](#), p. 75).

Lastly, as states have been paying for unknown others to construct and grade the tests, most public schools, as a result of both fiscal cutbacks over the last two decades, are receiving less funding per pupil than in the past. Consequently, teachers' salaries have not kept pace with other professions.

In the same way that it took withholding children from school to force those in power to acknowledge the concerns of parents, students, teachers, and others, teachers have learned that they can withhold their contribution to the educational enterprise by striking. Teachers, students, and parents were less likely to affect educational policy through the usual means of lobbying and holding forums. Instead, it took teachers going on strike and withholding their labor to reassert their interests.

Until the recent wave of strikes, teachers rarely went on strike. But once teachers observed other teachers successfully striking first in West Virginia, they began to do the same. As each strike resulted in improved working conditions for all, teachers began to realize that as they had the support of the community, they had the power to demand changes. As community organizer Eleni Schirmer writes, "Ordinary people have two major forms of power—our numbers—there are more of us than them—and our ability to withhold our participation in a system that denies us fairness, dignity and health" ([Schirmer, 2021](#), p. 250).

Where teachers have been successful, writes [Schirmer \(2021\)](#), they have gone beyond advocacy and mobilizing to organizing. Organizing is crucial because, as opposed to advocacy and mobilizing, it aims to organize power structures to favor constituents and diminish the power of their opposition, and recruit and develop leaders from the rank and file.

Concluding ideas

Here we have focused on some of the principal ways education is being privatized. For-profit corporations are increasingly taking over public education by privatizing schools as charter schools or using vouchers to provide public funding to private secular and religious schools. Further, corporations are privatizing what occurs inside schools by taking over the process of setting standards, creating curriculum, and administering assessments.

But over the last few years the public—parents, students, educators, and others—have been pushing back and reasserting the public in public education. Indeed, we have shown how as much as corporate neoliberal reformers have the monetary and often the legal advantages over the general population, those arguing for schools that support the common good continue the fight.

References

- Apple, M., 1982. *Education and Power*. Routledge and Kegan Paul, Boston, 1982w.
- Blanc, E., 2019a. *Red State Revolt: The Teachers' Strike Wave and Working-Class Politics*. Verso, New York, NY.
- Blanc, E., 2019b. Ten Lessons From the Teacher Strikes. *The Democratic Left*. Retrieved from: <https://www.dsasua.org/democratic-left/ten-lessons-from-the-teacher-strikes/>.
- Bobbitt, F., 1912. The elimination of waste in education. *Elem. Sch. Teach.* 12 (6), 259–271.
- Boninger, F., Molnar, A., Saldaña, C.M., 2019. Personalized Learning and the Digital Privatization of Curriculum and Teaching. National Educational Policy Center. Retrieved from: <https://nepc.colorado.edu/publication/personalized-learning>.
- Bulger, M., McCormick, P., Pitcan, M., 2017. The Legacy of InBloom. Working paper. https://datasociety.net/pubs/eci/InBloom_feb_2017.pdf.
- Buras, K.L., 2014. *Charter Schools, Race, and Urban Space: Where the Market Meets Grassroots Resistance*. Routledge, New York.
- California Teachers Association, 2019. State of Denial: California Charter Schools and Special Education Students. The report is available at: <https://www.utla.net/news/new-study-reveals-privately-run-charter-schools-under-enroll-students-disabilities>.
- Chen, G., 2020. Charter schools vs. traditional schools: which one is underperforming? *Public Sch. Rev.* Available at www.publicschoolreview.com/blog/charter-schools-vs-traditional-public-schools-which-one-is-under-performing.
- Cody, A., 2014. Common core-aligned tests and the new Pearson GED: failure by design? *Educ. Week*. Available at: <https://www.edweek.org/teaching-learning/opinion-common-core-aligned-tests-and-the-new-pearson-ged-failure-by-design/2014/06>.
- Collins, C., Lear, C., 1995. *Chile's Free Market Miracle. A Second Look*.
- Dewey, J., 1915. Industrial education—the wrong kind. *New Republ.* 3 (42), 178–180.
- Dewey, J., 1997. *Experience and Education* [1938]. Touchstone, New York, NY.
- Eyles, A., Machin, W., McNally, S., 2017. Unexpected school reform: academization of primary schools in England. *J. Public Econ.* 155, 108–121.
- Finnigan, K., Adelman, N., Anderson, L., Cotton, L., Donnelly, M.B., Price, T., 2004. Evaluation of Charter Schools Program: 2004 Final Report. U.S. Department of Education, Policy and Programs Study Service. Available at: <https://www2.ed.gov/rschstat/eval/choice/pcsp-final/finalreport.pdf>.
- Frankenberg, E., Siegel-Hawley, G., Wang, J., 2010. *Choice Without Equity: Charter School Segregation and the Need for Civil Rights Standards*. The Civil Rights Project/Proyecto Derechos Civiles at UCLA, Los Angeles, CA.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Seabury, New York.
- Friedman, M., Friedman, R., 1980. *Free to Choose: A Personal Statement*. Harcourt, Brace, and Company, New York, NY.
- Gilborn, D., 2008. *Racism and Education: Coincidence or Conspiracy*. Routledge, New York.
- Harden, N., 2020. Walton Foundation Taps Wall Street to Fund Charter Schools. *Real Clear Education*. Available at: https://www.realcleareducation.com/2020/01/31/walton_foundation_taps_wall_street_to_fund_charter_schools_48156.html.
- Hursh, D., Deutermann, J., Rudley, L., Chen, Z., McGinnis, S., 2021. *Opting Out: The Story of the Parents' Grassroots Movement to Achieve Whole-Child Public Schools*. Myers Education Press, Gorham, ME.
- Hursh, D.W., 2015. *The End of Public Schools: The Corporate Reform Agenda to Privatize Education*. Routledge, New York, NY.
- Kahlenberg, R.D., Potter, H., 2015. Restoring Shanker's vision for charter schools. *Am. Educat.* 38 (4), 4–14.
- Kliebard, H., 2004. *The Struggle for the American Curriculum: 1893–1958*, third ed. Routledge, New York, NY.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in U.S. schools. *Educ. Res.* 35 (7), 3–12.
- Lecker, W., 2013. The hidden costs of charter schools. *Stamford Adv.* 1. <https://www.stamfordadvocate.com/news/article/Wendy-Lecker-The-hidden-costs-of-charter-schools-4635437.php>.
- Leitner, H., Sheppard, E., Sziarto, K., Maringanti, A., 2007. Contesting urban futures: decentering neoliberalism. In: Leitner, H., Peck, J., Sheppard, E. (Eds.), *Contesting Neoliberalism: Urban Frontiers*. Guilford Press, New York, NY.
- McAlevy, J., 2016. *No Shortcuts: Organizing for Power in the New Gilded Age*. Oxford University Press, New York, NY.
- McAlevy, J., 2020. *A Collective Bargain: Unions, Organizing, and the Fight for Democracy*. HarperCollins, New York, NY.
- Miron, G., Gulosino, C., Shank, C., Elgeberi, N., Davidson, C., Hernandez De Alvarez, F., Jurdzy, B., Larsen, J., Pham, D., Ruder, K., Urdapilleta, L., Urschel, J., 2021. Profiles of For-Profit and Nonprofit Education Management Organizations Fifteenth Edition. National Education Policy Center, Boulder, CO. Available at: <http://nepc.colorado.edu/publication/emo-profiles-fifteenth-ed>.
- Pedroni, T.C., 2007. *Market Movements: African American Involvement in School Voucher Reform*. Routledge, New York.
- Ravitch, D., 2013. *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*. Vintage, New York, NY.
- Ravitch, D., 2020. *Slaying Goliath: The Passionate Resistance to Privatization and the Fight to Save America's Public Schools*. Knopf, New York, NY.
- Ravitch, D., 2021. Who Signed the Pro-testing Appeals? Diane Ravitch's Blog. Retrieved from: <https://dianeravitch.net/2021/03/05/who-signed-the-pro-testing-appeals/>.
- Reagan, R., 1986. Ronald Reagan Presidential Foundation and Institute. Website is available at: <https://www.reaganfoundation.org/ronald-reagan/reagan-quotes-speeches/news-conference-1/>.
- Rizvi, F., Lingard, B., 2009. *Globalizing Education Policy*. Routledge, New York, NY.
- Sahlberg, P., 2016. The global education reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. John Wiley and Sons, Malden, MA, pp. 128–144.
- Schirmer, E., 2021. Three theories of change: organizing, mobilizing, and advocacy. In: Charney, M., Hagopian, J., Peterson, B. (Eds.), *Teacher Unions and Social Justice: Organizing for the Schools and Communities Our Students Deserve*. Rethinking Schools, Milwaukee, WI.
- Schneider, M.K., 2014. Obama's USDOE Appointed to Privatize. Period. Mercedes Schneider's Blog. Available at: <https://deutsch29.wordpress.com/2014/11/23/obamas-usdoe-appointed-to-privatize-period/>.
- Schneider, M.K., 2015. *Common Core Dilemma: Who Owns Our Schools?* Teachers College Press, New York, NY.
- Snedden, D., 1924. Education for a world of team players and team workers. *Sch. Soc.* 20 (514), 552–557.
- Stern, M., Clonan, S., Jaffee, L., Lee, A., 2015. The normative limits of choice: charter schools, disability studies, and questions of inclusion. *Educ. Pol.* 29 (3), 448–477.
- Strauss, V., 2014a. Everything you need to know about the common core. *Wash. Post*. <https://www.washingtonpost.com/news/answer-sheet/wp/2014/01/18/everything-you-need-to-know-about-common-core-ravitch/>.
- Strauss, V., 2014b. The scary way common core "cut scores" are selected. *Wash. Post*. Available at: <https://www.washingtonpost.com/news/answer-sheet/wp/2014/04/29/the-scary-way-common-core-test-cut-scores-are-selected/>.
- Thatcher, M., 1995. *The Downing Street years 1995*. Harper Collins, New York.
- Ujifusa, A., 2021. States still must give standardized tests this year, Biden administration announces. *Educ. Week*. Available at: <https://www.edweek.org/teaching-learning/states-still-must-give-standardized-tests-this-year-biden-administration-announces/2021/02>.
- Verger, A., Fontdevila, C., Zancajo, A., 2016. *The Privatization of Education: A Political Economy of Global Education Reform*. Teachers College Press, New York, NY.
- Weiss, J., 2011. The innovation mismatch: "smart capital" and education innovation. *Harv. Bus. Rev.* Available at <https://hbr.org/2011/03/the-innovation-mismatch-smart.html>.
- Wells, A.S., 2019. An Inconvenient Truth About the New Jim Crow of Education, 2019 Presidential Address, Presented at the American Educational Research Association Annual Conference, Toronto, Canada.

Williams, B., Hogan, A., 2021. Pandemic Privatization in Higher Education: Edtech and University Reform. Education International Research, Brussels, Belgium.

Winerip, M., 2003. On education; when a passing grade defies laws of physics. N. Y. Times. Available at. <https://www.nytimes.com/2003/03/12/nyregion/on-education-when-a-passing-grade-defies-laws-of-physics.html>.

Winerip, M., 2011. 10 years of assessing students with scientific exactitude. N. Y. Times. Available at. <https://www.nytimes.com/2011/12/19/education/new-york-city-student-testing-over-the-past-decade.html>.

Yerkes, R., 1919. The Mental Rating of School Children. National School Service.

Shifting priorities of foreign aid to Pacific Island Countries: some considerations of new imperialism and geopolitics

Juliana Mohok McLaughlin, Queensland University of Technology, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	193
Foreign development aid and Pacific education systems	194
Framing foreign aid to Pacific Islands' development and education	195
A conceptual framing of new imperialisms—architecture of foreign aid	196
Globalization	197
New imperialism	197
The global/local nexus—geopolitics and hegemonic power	198
Global imperatives, new forms of foreign aid and local agendas	199
Global connections and new forms of foreign aid	199
Opportunity, threat or special responsibility: Pacific step-up	200
Politics of education reform—global imperatives national implementation	201
Summary and implications for education	202
References	203
Websites	204

Introduction

Research exploring foreign aid provision concludes that aid is motivated more by the economic priorities of donor governments and multilateral agencies than the economic needs of developing countries. Foreign aid had historically been negotiated through bilateral agreements in which donors worked collaboratively with developing countries to address agreed challenges (Fagerlind and Saha, 1989). In contrast, over the last three decades, foreign aid has a more explicit political agenda because it is unashamedly motivated to fulfill regional political agendas than addressing local needs of recipient countries (Coxon and Munce, 2008; Luteru and Teasedale, 1993; Richards, 1977). Recent experiences of aid agreements in the Pacific suggest a complex relationship driven by economic priorities, national security agenda and global events, inviting scholars to examine how new imperialisms (Tikly, 2004; Richards, 1977) underpin foreign aid in postcolonial contexts.

This paper examines the impact of globalization and shifting geopolitics on foreign aid for development and education by focusing on some experiences of Pacific Island Countries. Drawing from recent experiences of aid projects, this paper argues that the language of foreign financial arrangements remains confined to colonial and deficit discourses of developed and underdevelopment and does not take into account an intensified transnational order that directs foreign policies and educational agenda. An eclectic conceptual framework consisting of postcolonial reconsideration of neo-colonialism and dependency, globalization and new imperialisms in the context of geopolitics is proposed to allow for a deeper examination of contemporary architecture of foreign aid. Despite economic fatigue, foreign aid is channeled and powered on through new forms of imperialism, an understanding of which is imperative for educational development research.

Australia's role as the major aid donor for Pacific Island Countries is the focus of this discussion. Some examples of global imperatives through which new forms of aid are expended are also discussed, including the manifestations of tensions between superpowers as Pacific Island Countries embark on participating in regional economic dialog. The second narrative draws on Australia's perspective of its neighbors, considering how it sees the Pacific as an opportunity, threat or special responsibility (Hawksley, 2009) while national security remains as the key platform. Given the intensified geopolitical movements in the region, a new approach is being adopted through the Pacific Step-Up initiative, with the Australian government investing in corporate capital and expertise to counteract hostile powers (Kehoe, 2021). I conclude by returning our focus to the implications for education, specifically on the politics of education reform. This brings forward an examination of multi-lateral organizations, such as the United Nations and global goals for educational development. This final narrative confirms that educational priorities set by external agencies are noble (Tikly, 2004) and attract foreign financial assistance for implementation. However, local conditions and resourcing risk appropriate implementation and sustainability of these projects. An appreciation of the Pacific Island Countries' context is useful for this analysis.

Foreign development aid and Pacific education systems

In the context of geopolitics and foreign aid, the Pacific Island Countries in Oceania are particularly dependent, vulnerable and implicated (Dorman and Pryke, 2017; Funaki, 2016; Gani, 2008; Hughes, 2003). Vulnerability and high dependency on foreign aid, to a large extent, is caused by their geographical features and demography, particularly with small populations and remote locations within the vast Pacific Ocean (Pavlov and Sugden, 2006). A contextual discussion of the Pacific Islands is necessary, providing a background to foreign aid imperatives and its significance to the region. While Australian aid to the countries of the Pacific will be the focus of this examination, major aid donors and lenders include Governments of New Zealand, Japan, China, the European Union, the World Bank and the Asian Development Bank.

The Pacific region incorporates the Earth's largest ocean, comprising 35% of its surface (Morgan, 2020; Thomas, 1993) and includes both sovereign nations and colonial territories, in free association with colonial powers. The term "Pacific Island Countries" (PICs) specifically refers to fourteen nation states consisting of the Cook Islands, Fiji, Kiribati, Nauru, Niue, Palau, Papua New Guinea, Republic of Marshall Islands, Samoa, Solomon Islands, Tokelau, Tonga, Tuvalu and Vanuatu. Other island countries including New Caledonia, French Polynesia, and Federated States of Micronesia are French and United States territories respectively. These island countries are collectively grouped into three regional groups consisting of Melanesia, Micronesia and Polynesia.

The Pacific, although home to approximately twelve million people (Pacific Community, 2020), is extremely culturally and linguistically diverse, thus collectively, the region holds the world record for the largest number of languages spoken in one region, with the larger Melanesian countries of Papua New Guinea, Solomon Islands and Fiji containing the densest populations. As a consequence of colonial history, either English or French are often the languages used for government and business communication. Likewise, these are the languages of instruction in schools, although creole languages for example, *Tok Pisin* in Papua New Guinea and Solomon Islands, and *Bislama* in Vanuatu, are widely spoken in informal contexts. Fisheries, agriculture, and mining, particularly in the larger islands, are the primary economic resources. Growing service industries, such as tourism are emerging which provide employment to citizens, most of which were severely impacted by the COVID-19 pandemic from 2019. Importantly, issues concerning sustainability of resources (land and sea), land denigration, and rising sea levels caused by global warming make Pacific countries vulnerable to climate change and especially dependent on foreign aid (Dorman and Pryke, 2017; Johansson-Fua, 2016; Funaki, 2016; Morgan, 2020).

While PICs contribute the least to greenhouse gas emissions, they experience the negative impacts of climate change due to frequent king tides caused by raising sea levels and extreme weather patterns. These often culminate in heat waves, droughts, landslides, and floods which can have an impact on livelihoods and development progress. For Pacific Island communities, livelihoods and wellbeing depend on geography and natural resources including the ocean being a major source of sustainable access for healthy living and economic growth. There is a growing dissenting voice by Pacific peoples on the impact of global warming and consequent rising sea levels on low-lying atolls as seen at the recent United Nations Climate Change Summit in Glasgow in 2021. However, these voices are often drowned out by developed nations and agencies who contribute most to global warming via greenhouse gas emissions. Ultimately, developed countries with financial power and investments in developing alternative and environmentally friendly energy sources determine the nature and delivery of solutions. Contextual solutions drawn from indigenous ontologies and realities are silenced, begging for strategies that could inspire local innovation for sustainability through education.

As such, any critique of foreign aid to education invites an appreciation of its historical foundations, while particularly important, recognizing the important role education can play in shifting colonial discourses embedded within foreign aid relationships. Formal education in the Pacific is attributed to the Christian missionaries, during the so-called Pacification Era as the precursor for European colonial annexation of Pacific Islands. Colonial educationalists ignored traditional ways of knowing, value systems and vernaculars, and to various extents, suppressed indigenous languages and banned cultural practices. These were replaced with a variant of European education delivered through English or French, with the Bible often used as a foundational text. Such education aimed to "civilize", enlighten and eventually train a colonized population to support the work of the colonial administration. Thomas (1993) asserts that formal colonial education was alien to indigenous educational structures, methods, personnel and content, legacies of which are carried over in post-colonial times. As Thaman (2013) argues, no one asked how Pacific people conceptualized wisdom, knowledge, learning and teaching, or values that informed and sustained these communities (McLaughlin, 2018).

Education systems in these developing countries remain largely a colonial inheritance, evident in learning structures, educational policies and institutional architecture. Implicated in the colonial inheritance are the roles of donor agencies, nature and aims of technical assistance on the one hand and local political and educational priorities on the other, pointing to an ambivalence of post-colonial conditions characterized by simultaneous reluctant resistance to and compliance with new forms of power embedded in geopolitical agendas.

Decades after declaration of political independence, Pacific Island Countries continue to emulate systems of education inherited from the colonial past (McLaughlin and Hickling-Hudson, 2005; Nabobo-Baba, 2013; Sanga, 2006; Thaman, 2005, 2013). These colonial legacies include a stratified system of schooling, culturally inappropriate syllabuses, high stakes examinations which generate large numbers of relatively illiterate "drop-outs" and culturally alienated youth. This is compounded by insufficient employment opportunities, a scenario that is experienced in Papua New Guinea with high numbers of unemployed educated university graduates in 2022. Ironically, the end of classic colonialism witnessed the emergence of new forms of hegemonic powers that continue to control educational structures and processes. Importantly, reliance on foreign aid for educational change and

innovation is complicated by respective PICs and donor political agendas rather than the real needs of the Pacific communities (Coxon and Munce, 2008; Sanga, 2016). For example, a recent announcement in December 2021 of A\$35 million from the Australian sector budget to support Papua New Guinea tuition-free primary education in 2022 exemplifies new forms of financial support, in addition to A\$340 million in COVID-19 related assistance to assist the country's response to the pandemic (The National PNG, December 14, 2021).

A growing social divide between nation's elite and those of lower socio-economic status in rural, remote and urban poor exists in developing countries, and the PICs are of no exception. Tackling contemporary educational and consequent social divide involves reconciliation of modern forms of socialization with the traditional indigenous ways of knowing and value systems. The challenge of leading contemporary educational change then demands new ways of interrogating perennial questions (see Thomas, 1993, p. 240) such as the purpose of formal education, ownership and decision making, teacher education and certification, curriculum, and pedagogical models. In recent decades, the generation of possible responses to these challenges has been exacerbated by global imperatives and provisions of foreign aid for educational development.

Apart from colonial legacies, the postcolonial education systems are influenced by global agendas. Over the last three decades, Pacific Island Countries have been signatories to global imperatives, such as the United Nations (UN) Declaration on Education for All (EFA) in the 1990s, the Millennium Development Goals and the UN Sustainable Development Goals set for the 2030 Agenda for Sustainable Government (UNESCO, 2021). These global declarations are honorable, containing appropriate goals targeting serious concerns such as poverty alleviation, increasing access to education, gender equality and women's empowerment, and environmental sustainability. However, given the scale of these educational imperatives, often beyond national education budgets, implementation of global educational reforms depends on foreign governments, donors, and lenders such as the World Bank, the Asian Development Bank, and multiple non-government organizations.

While it is argued that the Pacific Island Countries are small, isolated and vulnerable to aid dependency (Dorman and Pryke, 2017; Gani, 2008; Pavlov and Sugden, 2006), their vast marine and natural resources attract competition from foreign governments and global conglomerates. In addition, the geographical location triggers competition between superpowers bordering the Pacific Ocean, specifically between the United States of America and China, making the Pacific Island Countries a pawn in their geostrategic power rivalry (see Hawksley, 2009). Accordingly, this geo-political rivalry implicates all countries in the region, including Australia and New Zealand, who maintain "special relationships" with Pacific Island Countries.

Given this scenario, foreign aid to the Pacific Island Countries and specifically educational development proceeds beyond bilateral agreements based on former colonial relationships but are increasingly linked to geopolitics by allies to superpowers in the region. A discussion of development aid through an eclectic conceptual framework for examining aid to education in the Pacific Islands is essential.

Framing foreign aid to Pacific Islands' development and education

There has been much controversy over foreign aid in both the recipient and donor countries over the last three decades. Hawksley (2009) contends that the Pacific region is characterized by asymmetries of power involving both former and current colonial powers, and between independent Island states themselves (p. 115). Although global comparisons would categorize Australia as middle power, it is considered a superpower within the Pacific as it is the largest and wealthiest state in the region. Australia and New Zealand are the two major donors of foreign aid to the Pacific Islands (Hawksley, 2009), and continue to instantly respond to natural disasters and emergencies such as those caused by cyclones, volcanic eruptions and consequent tsunamis, landslides and long periods of drought. Instant responses by both countries to the volcanic eruption in Tonga in January 2022 illustrates their humanitarian and special responsibilities to their Pacific neighbors. In cases of civil unrests and constitutional crisis, both countries led regional security forces to attend to threats to national stability for example the Bougainville conflict (Dorney, 1998), the Fiji military coup in 2006 (Schultz, 2014) and the civil unrests in the Solomon Islands in 2010 and 2021. Both Australia and New Zealand responded to the COVID-19 pandemic providing much needed vaccinations to Pacific countries. Such attention demonstrates an illusion of neighborly connectedness and responsibility, with readily available financial assistance budgeted through their foreign aid allocations.

Apart from bilateral agreements, both Australia and New Zealand financially support regional agencies, such as the Pacific Islands Forum which was established as a key site for Pacific diplomacy and as a means of gaining control of regional diplomatic agenda (Morgan, 2020). The Pacific Islands Forum succeeded the South Pacific Commission (SPC), established as a regional institution after the second world war to promote corporation among colonial powers. SPC was limited in its remit as it constrained discussions around decolonization and nuclear weapons testing, thus impeding Pacific leaders' voices and decision-making powers. The Pacific Islands Forum is a key site for Pacific diplomacy (Morgan, 2020) and acts as a regional power block for the interest of Pacific Islanders.

Schultz (2014), in his detailed analysis of theorizing Australia-Pacific Island relations, contended Australia's relationship with its Pacific Island neighbors was characterized by incoherency and inconsistency of policy making set to achieve its commercial, strategic and humanitarian goals. He argues Australia engages in generous development and humanitarian assistance which coexists within policies indifferent to Pacific Island interests and particularities, thus attracting negative feedback of the "overbearing and bullying Australia" (p. 549). Underpinning Australia's objectives with its Island neighbors are three distinct ideas that inform its relationships, based on the perception of the Pacific "as a threat, an opportunity and a special responsibility" (Schultz, 2014, p. 549). A

key element here is evident in the way Pacific Islands can be thought of as a steppingstone for hostile powers to reach Australia. Thus, soft power through foreign aid as a special responsibility can maintain friendly ties to ensure rival powers are excluded from setting up strongholds in the neighboring Island countries. The recognition of wealth to be extracted in the region, with current extraction of natural resources including commercial operations, mining including under the sea, forestry, trading and service industries are significant in respect to Australia's neoliberal economy. Thus, tying foreign aid to the purchase of Australian equipment, expertise, contractors, and agencies to negotiate agreements with Pacific Islands that may be seen as mutually beneficial for both donor and recipient countries.

Australian foreign aid policies have shifted from its original responsibility to former colonial territories. [Hawksley \(2009\)](#) postulates this difference in policies are more in style rather than substance of Australia's aid relationship. Comparing Prime Minister Rudd's new direction in 2007 with the Howard Government's position was a change of the former government's pushy and bullying Australia, and despite rhetoric of greater engagement with the Pacific Islands, the emphasis remained on economics of foreign aid ([Hawksley, 2009](#)). It can be argued that neo-liberalism continues to inform foreign aid policy and relationships with PICs. Australia's own colonial history of dispossession of Aboriginal lands, racialized ideologies and policies, and contemporary relationship with its First Nations people is postcolonizing ([Moreton-Robinson, 2003](#)) rather than decolonizing its policies for Aboriginal and Torres Strait Islander peoples.

In the absence of theorization of Australia's aid relationships ([Schultz, 2014](#)), the provision of foreign aid has been founded on theories of development and modernization and definitions of under-development. Economic models have been adopted to project aid provisions, such as the Migration, Remittance, Aid and Bureaucracy (MIRAB) model ([Dorman and Pryke, 2017](#)). Thus, such platforms for development aid to education suggest this practice is instituted on human capital theories, modernization discourses and colonial binaries view of the world ([McLaughlin, 2002](#)). [Luteru and Teasedale \(1993\)](#) highlight the mismatch between donor and recipient development priorities, and while control is in the hands of donors, exacerbates suggestions of neo-colonial dependency and un-equal power relationship. [Adams and Solomons \(1991\)](#) foretold that foreign aid may be inherently debilitating rather than strengthening to the recipient countries, an argument that remains relevant in the 21st century. In a postcolonial analysis of the Australian sponsored Papua New Guinea Secondary School Students' Project from 1988 to 2001, [McLaughlin and Hickling-Hudson \(2005\)](#) argued for theorization beyond dependency theories. A complex conceptual framework is therefore required to critique the intricate interfaces of foreign aid.

A conceptual framing of new imperialisms—architecture of foreign aid

The end of classic colonization and the post-colonial era sparked anti-colonial movements toward achieving self-determination by formal colonized countries and territories. Liberation from western colonial power sparked anti-colonial movements toward decolonizing institutional power structures. However, the colonial architecture of power remains deeply embedded within the post-colonial context, translating through capitalism, patriarchy and racism, a few lingering examples of oppression and legacies of colonialism. As [Yamahata \(2015\)](#) argues, the unequal colonial power relations are translational and intersectional, which perpetuates inequalities in division of labor, marginalization of women and reinforces patriarchal systems in contemporary society. As [Satre \(2001\)](#) reminds us, colonization is a system, its legacies implanted in modern institutions and forms of socialization. Consequently, education systems around the world which have their origins from the colonial experience, as [Tikly \(2009\)](#) contends, are artifacts of this legacy.

Decolonization is both a political process and as well as an epistemic project ([Yamahata, 2015](#)). Thus, any authentic analysis of former colonized states requires analysis of their contextual social, political, ideological, and cultural environments. Postcolonial theory offers a theoretical lens which incorporates a sophisticated critique of how colonial influences and values are embedded in the bureaucracy and education systems of independent nations ([Crossley and Tikly, 2004](#); [Hickling-Hudson, 2009](#)). Postcolonial theory extends a critical lens for understanding colonial legacies, and how colonial ideologies continue to influence political, economic, social, and cultural practices. Moreover, postcolonial theory provides a framework for analysis which explores and reveals systems of domination and hegemonic power relationships, ambiguity and contradictions in educational policy formation and practice. Implicated in postcolonial educational analysis is the provision of foreign aid projects for educational development, and their imposition of foreign education practices that filter into the systems of the post-colonial contexts.

Current systems of education are often perceived as elitist, lacking relevance to local realities, and at variance with indigenous cultures, values and belief systems ([Crossley and Tikly, 2004](#), p. 149; [Tikly and Bond, 2013](#)). Postcolonial perspectives are useful to critique colonial discourses ambiguities, contradictions, and ambivalences of educational policies, strategies and the everyday pedagogical relationships. Indeed, postcolonial theory offers a strategy to explore power structures and knowledge systems ([Fox, 2016](#)) as well as a catalyst that forces a deeper consciousness in generating an intellectual energy to combat injustice in the form of neo-colonialism and imperialism ([McLaughlin, 2018](#)).

In consideration of remarkable global change since the establishment of foreign aid after the Second World War, postcolonial analysis is appropriate for examining new forms of international relationships and agents involved in its operations. Interrogating power relationships between and within national borders, as those offered by colonizer—colonized relationships, as well as internal power structures, require critique beyond classic and new forms of colonization. Indeed, colonization has not ended at the point of political independence for former colonized countries, current global movements represent salient configurations of power,

transforming global economic transactions, cultural and social institutions. A discussion of globalization and its impact on educational processes and foreign aid provisions is therefore necessary.

Globalization

The last four decades have witnessed the onslaught of global economic, political and cultural transformation through globalization, and has thus inspired a plethora of literature on globalization and education (see for example, Henry, 1999; Hickling-Hudson, 1999; Rizvi and Lingard, 2010; Scholte, 2000; Spring, 1998; Taylor et al., 1997; Zadjia, 2015). Scholte (2000) postulates that globalization is a transformation of social geography shaped by the growth of supraterritorial spaces, thus, both territoriality and supraterritoriality coexist in complex interrelations (p. 8).

Globalization involves supra-national ideas and process which are borderless, over and beyond the nation state (Hickling-Hudson, 1999, p. 82), resulting in an intensified global distribution of a particular economic system based on capital, prompted by rationalism, capitalism, technological innovation, and regulation (Scholte, 2000). The rise of supraterritoriality was a result of the combination of the emergence global consciousness as a product of rationalist knowledge, development trends of capitalism, innovative communication and data processing technologies, and the construction of regulatory frameworks through states and supranational institutions (Scholte, 2000, p. 3). Founded on the basis of a bias regulated global economy, Brecher and Costello (1994) warn of the competitiveness associated with countries' propositions to attract corporate investments that would result in a reduction of labor, social and environmental costs. Impacts of globalization are realized as nation states relinquish control of resources to multinational conglomerates, who control the operations and benefit most from profits of their investments. Although the West stood to benefit most from capital investments, globalization facilitated the growth of the markets in other regions, witnessing the growth of the Asian Tigers. Neoliberal policies, on the other hand, prompted many negative consequences which transpired and led to increased ecological degradation, persistent poverty, deteriorated employment conditions, cultural violence, widening inequalities and deepened democratic deficits (Scholte, 2000, p. 9). As Rizvi (2009b) argues, neoliberal states celebrated the emergence of global markets and "free trade" in the production and consumption driven by the view of "globalization from above", yet little attention was paid to the realities, "globalization from below" (p. 287). The notion of the equal economic opportunities offered through globalization remains a myth, as it was never an equal playing field in the first place.

Moreover, globalization facilitated an unprecedented flow of capital and opportunities to participate in the global economy, it also created an unprecedented flow of ideas, migration, refugees, travelers and cultural movements (Hickling-Hudson, 1999). Rizvi (2009b) details the complexities, multi-dimensional and nuanced triggers of international mobility that has transformed global spaces, drawing our attention to how people movements consequently have constructed transnational identities and forms of socialization. Human security (and in-security) and identities are then transnational with cross border connections and interconnections. These have implications for educational research as a geographically bounded field of study is no longer sustained, instead there is a need to pay attention to new non-linear, complex and evolving relationalities (Rizvi, 2009a,b, p. 287).

As globalization facilitated new forms of communication, economic and cultural opportunities, it created winners and losers, and can be argued that it represents manifestations of nineteenth century colonialism (see Rizvi, 2007; Tikly, 2004; Tikly and Bond, 2013). Powered by information and communication technologies, globalization massively impacted education policies, institutional architecture, curricula, assessment, and pedagogies. Globalization involves the diffusion of uniform education goals and systems, complicating the processing of educational planning and provision. Rizvi (2007) argues for educational research to acknowledge, understand and respond to powerful forces impacting educational processes beyond the jurisdiction of nation states.

The reality of globalization means that decisions made in Europe, the United States and Australia influence how Pacific Island Countries communicate, conduct business, use technology and educate their citizens. It is also important to recognize the role of international organizations such as the World Bank and the Organization for Economic Co-operation and Development (OECD) (Rizvi, 2009a,b), as these multilateral development agencies provide the language, set the goals and underlying rationale for educational policy and practice (Tikly, 2004, p. 174). Some clarification and definitions of new forms of power, specifically, new imperialism is required.

New imperialism

Imperialism is commonly accepted "as the practice, theory and the attitudes of a dominating metropolitan center ruling in a distant territory" (Said, 1993, p. 8), with early definitions framed according to the imperial relations between the ruling nations and their colonies (Sen and Marcuzzo, 2018, p. 1). Imperialism differs from colonialism in terms of implementing settlements on foreign lands, however, colonialism facilitated the foundations on which imperial interest can transpire (Tikly, 2004; also see Ferguson, 2003; Ferro, 1997). According to Tikly (2009), new imperialism differs from neo-colonial critique in two ways. First, when Kwame Nkrumah (1965) coined the term neo-colonial, new imperialism was not fully conceptualized particularly with the changing context of Western dominance (p. 26). The second difference is the way new imperialism is understood and analyzed, as it coincides with poststructuralist and culturalist thought in the social sciences. Notably, when colonialism was legitimated on race and biological racism which originated in the eighteenth-century eugenics movement, complex cultural differences (a by-product of colonization) has emerged as an element of conflict in western countries (Tikly, 2004). Cultural diversity and social change have indeed impacted on colonial binaries and definitions.

Marcuzzo and Sunanda, 2018 contends that imperialism was an arrangement of exploitation and expropriation, sustained over time and space, through various forms to achieve its set aims and depending on a network of power structure. It can be argued that imperialism prevails from its early domination through the colonies—that maintains empire power relationships to contemporary capitalism, based on neoliberalism, globalization and free market ideology (Sen and Marcuzzo, 2018). On the other hand, the decolonization era triggered new configurations of empire—colony relationships, which then necessitated a re-working of overseas development assistance, tying aid to priorities of powerful nations. Trade becomes the new mantra, unfortunately exacerbating the gap between the rich and economically poor nations, thus witnessing imperialism in its new incarnation (Sen and Marcuzzo, 2018, p. 2). Consequently, dominating agendas and interests are driven by economic, political, and military concerns which legitimate powerful nations imposing their democratic ideologies on the economically poor countries around the globe.

Apart from the economic aspects of imperialism (Marcuzzo and Sunanda, 2018), the discursive nature of imperialism remains an important angle for analysis. Tikly (2009) is concerned with the impacts and implications of globalization and cultural change and how these have been experienced through social transformative agents, including feminist groups, indigenous groups and social movements. Foucault's (1991) work is useful with this analysis in ways that comprise people as subjects, thus manifested through authority and exclusion, surveillance and governmentality (Tikly, 2009; O'Farrell, 2005). New imperialism as transpired through foreign aid and loans to lower income countries illustrates the exercise of power, surveillance and the governing architecture, providing a shrewd critique to new forms of domination and hegemony.

The impact of new imperialism on education has been meticulously delineated by Tikly (2004, 2009) who contends that education is a central feature of new imperialism, used by the multilateral agencies and the World Bank as they envision the nature of education as an aspect of Western discourse around development. For example, most United Nations agencies operate under the belief that public intervention is necessary to ensure basic needs and human rights, a platform that is necessary and resonates with development priorities of partner countries. On the contrary, low-income countries are trapped between policy imperatives set by global organizations in contradictory ways (Tikly, 2009, p. 26), particularly when projects designed to respond to human rights and basic human needs are directed and controlled by donors and funding agencies, often without much consideration of local contexts. It is, therefore, crucial that Pacific Island Countries understand the genesis and processes of new imperialism in a way that allows society, through education, to counteract ongoing imperial politics and mechanisms of domination and hegemonic control.

The global/local nexus—geopolitics and hegemonic power

The examination of key theoretical positions offered through an analysis of discourses of development and foreign aid, postcolonialism, globalization and new imperialism allows the interrogation of contemporary architecture of foreign aid policies and global interactions. New forms of foreign aid are considered as humanitarian, poverty alleviation, and development imperatives of foreign aid arrangements shift focus to global security and geopolitics. The Pacific region provides some manifestations of this.

It is not possible to offer a single definition of geopolitics in the context of how we understand, critique and analyze world politics, however, geopolitics is briefly described as "the struggles over the control of spaces and places and focus of how power operates across scales" (Nyuyen, 2020, p. 1). The complexity around defining geopolitics is simply because, as Flint (2006) postulates, the word invokes images and ideas of war, empire and diplomacy therefore, propping up the perspectives of geopolitics as the practice of states controlling and competing for territories (p. 13). In this sense, we also consider how countries are competing for not simply territories, but also for the extraction of resources contained therein. "Power" is classified by sizes of countries, the racial character of its population, its military and economic capacity, and more recently, on capabilities of industrial strength, educational levels and military might (Flint, 2006, p. 28). At a theoretical level, geopolitics creates images by providing language, practice and classifications of territories and masses of people, consequently, geopolitics becomes a way of seeing the world (Flint, 2006, p. 13). Further, geopolitics is not simply understood as an exercise of territorial competition, and/or that geopolitics is the preservation of states, but invokes "situated knowledge" that may include racial conflicts, women's movements, diplomacy over greenhouse gas emissions. Geopolitics involves multiple practices and representations by various agents and territories and presents a new form of imperialism.

Countering dominant ways of thinking about world affairs, critical geopolitics encourages agents, including groups of citizens, feminist groups, individuals, indigenous peoples and marginalized others to resist hegemonic control imposed by states and powerful institutions. Feminist critique of geopolitics considers how patriarchy (male-dominated elite) define and exercise power relations through what is perceived as common sense and normal. Adopting Antonio Gramsci's (1971) theory of hegemonic power, the use of exert force for compliance is not required. On the contrary, the lower class or minority groups follow willingly the ideology that is presumed natural and unpolitical, which remains to serve and benefit the interest of the powerful (Flint, 2006). For instance, policies for economic development designed by the rich powerful nations are adopted willingly by poor nations under the disguise of "progress and development". Representation of geopolitics as a manifestation of power (O'Tuathail, 1996), necessitates critical understanding of how such ideas and practices are justified as normal while belittling alternative forms of knowing and being (Flint, 2006, p. 28). Although nation states are key geopolitical agents, seen through the policies and practices of both domestic and international competition, we consider the roles of non-state organizations and multi-state organizations, such as the North Atlantic Treaty Organization (NATO) and the United Nations connections to the geopolitical structure, forming a dynamic and hierarchical architecture of power relations. Thus, critical understandings of the connections between geopolitical agents can allow space for other groups and alternative ways of viewing the world.

Global imperatives, new forms of foreign aid and local agendas

A conceptual framework consisting of neo-colonialism, globalization, new imperialism and geopolitics provide powerful perspectives in the examination of foreign aid to education globally and specifically to Pacific Island Countries. Indeed, global imperatives for education are noble and address the essence of humanity, incorporating human rights, poverty alleviation, gender relations, and sustainable development and environmental protection. With foreign governments and donor agencies primarily financing global educational agendas, it has generated contestations around local implementation, issues of sustainable outcomes and the capacity of lower income countries to maintain momentum once projects are completed. This next section of the chapter illustrates some of the experiences of foreign funding projects driven by development ideologies and practices beginning with the enticement to participate in globalization processes.

Global connections and new forms of foreign aid

The Papua New Guinea experience provides a case in point for critique of foreign aid to education and national development. Papua New Guinea gained its political independence in 1975, after Australia coached the territories of Papua and New Guinea toward a sovereign nation and relinquished its colonial administrative regime. Over the last 47 years, PNG remained a major beneficiary of Australia's overseas aid program, organized mostly via bilateral agreements and remains strategically important to Australia's national security. Apart from Papua New Guinea's membership of and leadership within the Pacific Islands Forum as an important agency for regional solidarity, it became a formal member of the Asia-Pacific Economic Cooperation (APEC) in 1993, allowing it to participate in the processes of globalization, making Papua New Guinea the only Pacific Island Country to enter this cooperation. APEC is an economic cooperation forum in the Asia-Pacific region, incorporating Asian countries, Australia and New Zealand, and countries bordering the Pacific Ocean, including Canada, Japan, Russia, United States of America, Mexico, Peru and Chile, with a membership of twenty-one countries in total. As Plummer (1997) predicted, Papua New Guinea stood out to economically gain from this membership, but also cautioned that it needs to avoid seeing APEC as another vehicle for special and differential treatment for foreign aid. Instead, Plummer (1997) called for Papua New Guinea to reform its economic policies and programs consistent with World Trade Organizations (WTO) obligations and align with important trading partners in APEC.

In 2018, Papua New Guinea hosted the 26th APEC Summit in Port Moresby, with the host government aiming to promote its economic development prospects, establish trading links with member states and increase foreign investments. Papua New Guinea invited Pacific leaders to a special meeting, the *APEC Leaders' Dialogue with Pacific Island Leaders*. With APEC's position as a powerful economic group, the 2018 Summit had global media coverage through the Australian Broadcasting Commission (ABC), the British Broadcasting Commission (BBC), The Guardian, Reuters and the Washington Post, among others (Huang, 2020).

Conversely, Papua New Guinea is the least economically developed among the APEC member countries (ABC, 2018), so hosting the Summit was a gamble, with its 2018 national budget projecting to record a deficit of two billion PNG Kina (\$800 million). Hosting the APEC Summit was a major financial commitment for the country which prompted outrage at a time when many citizens were living below the poverty line as stated by the United Nations Development Program (ABC, 2018; The Guardian, 2018). With six million citizens living outside urban centers with health and educational services severely exhausted and poor infrastructure, Papua New Guinea welcomed superpowers into Port Moresby. With the cost of hosting the Summit projected to over one billion dollars (including the purchase of luxury Maseratis and Bentley Spurs), the Papua New Guinea Government sought financial assistance from other donors and lenders resulting in China financing the building of the APEC House (International Conventional Center) and other infrastructure projects. Papua New Guinea had signed up to China's One Belt One Road initiative (ABC, 2018; Morgan, 2020), consisting of billions of dollars for infrastructure investment primarily to expand the land and sea links between Asia, Africa and Europe (Wen et al., 2018). Consequently, a showcase of Chinese aid and investment was on display with the Chinese President Xi Jinping opening a Beijing funded boulevard in Port Moresby and met with Pacific leaders to promote the Road and Belt initiative, in addition to discussions on respective bilateral financial agreements particularly with fragile and distressed economies, agitating concerns around increasing Chinese global influence particularly in the Pacific region (Hayward-Jones, 2013).

Clear tensions between China and the United States of America prevented a Summit consensus. The United States Vice President Mike Pence stated that the underlying issues preceded economic concerns to freedom of navigation and human rights, and cautioned Pacific Island leaders to avoid accepting debt that compromises their sovereignty. Competition between the two world superpowers over the Pacific region swiftly became the focus, with the United States and its allies including Australia, New Zealand and Japan counteracting by committing to deliver reliable electricity and high-speed internet connection to Papua New Guinea (Wen et al., 2018).

The APEC Summit Chairman's Final Report contained reference to education, framed along "skilled human resources", with the purpose of education firmly aligned with on human capital theories (APEC Ministerial Report, 2018). In true Papua New Guinea's political style, a change of government eventuated in 2019, with the resignation of the then Prime Minister. The receipt of anticipated billions of dollars of foreign investment as a consequence of the 2018 APEC Summit was uncertain. Papua New Guinea's APEC experience, however, illustrates the imbalance and inequality of participation in international economics as facilitated by globalization processes, the geopolitical competition by superpowers, and the unequal and hegemonic relationships between powerful nations and emerging economies. The APEC discussions herein are useful for a renewed focus of aid relationship between Australia and the Pacific Islands. As Hayward-Jones (2013, p. 1) states, the "center of global economic gravity has shifted to the Asia-Pacific region".

Opportunity, threat or special responsibility: Pacific step-up

For almost half of a century, Australia enjoyed an exclusive superpower status in the region and recognized as such by its Pacific Island neighbors. The last two decades, however, saw Australia's relationship with the Pacific Island Countries defined by its own national security, manifested by the introduction of the Pacific Solution formally launched in 2001 (see Fox, 2015; Fraenkel, 2016; Grewcock, 2018; Larcertosa, 2014; Taylor, 2005). The Pacific Solution was established when Australia's government reformed aspects of its immigration system which involved a third country immigration processing scheme, "wherein asylum seekers who arrived in Australian territory by boats are detained by authorities and shipped to Pacific Island nations to have their refugee claims processed and reviewed under the receiving nations' laws" (Larcertosa, 2014, p. 322). Although it can be argued that "sovereign borders" policy directives are principled in combating and stopping people smuggling schemes and saving lives from drowning in oceans, the Pacific Solution demonstrates how national security and domestic politics can influence Australia's relationship with neighboring countries in the region. In negotiation with the Pacific Island governments, detention or refugee processing centers were established on Nauru and Manus Island, Papua New Guinea. International refugee and asylum seeker rights under International Conventions aside, the Australian government negotiated and re-negotiated the two-phases of the Pacific Solutions (Phase 1 2001–2008; Phase 2 2012–2006), agreed through Memorandum of Understandings with the two countries. As Nauru's Phase 1 negotiations illustrate, Australia pledged ongoing financial assistance in the form of tied aid in exchange for continued operations of the detention center (Grewcock, 2018, p. 363) which was estimated at the cost of AU\$1 billion, while Phase 2 including both operational expenditures on Nauru and Manus Island blew out to AU\$3.9 billion between 2012 and 2016 (Fraenkel, 2016, p. 282). In 2016, however, the supreme court found that the Manus Island detention center violated Papua New Guinea's constitution and ruled its immediate closure. The benefit and impact of the Pacific Solution on Manus Island and Nauru may require further analysis from the perspectives of the islanders themselves. This experience, however, illustrates geopolitical arrangement and requires critical understandings of how foreign aid can be used as an "unmitigated bribe" (Grewcock, 2018, p. 363). Similarly, avoiding a "debt trap diplomacy" is crucial to Pacific Island Countries (Wall, 2020).

The recent "Pacific Step-Up" initiative has renewed Australia's focus and relationship building through engagement with the Pacific Islands while strategically targeting a secure and economically stable region. Launched in 2016, the Pacific Step-Up captures Australia's intent for renewed partnership that draws on historical, cultural and personal ties with its Pacific neighbors. Economic prosperity, regional security, people to people connections, gender equality, strengthening resilience to climate change and disaster, and supporting regional organizations are key priorities in this initiative. Commitment to fisheries industry is also a highlight of the Pacific Step-Up, a recognition of ocean as a resource in terms of food security for Island communities' and global sustainability. Similarly, these priorities are translated in the *Aid Investment Plan Pacific Regional 2015 – 2016 to 2017 – 2018* (Government of Australia, 2018) which compliments separate bilateral agreements. Papua New Guinea, for example, continues to be the major recipient of Australia's ODA program with a total of AU\$607.5 million in 2020 (Australian Government, 2020). The *Papua New Guinea-Australia Comprehensive Strategic and Economic Partnership* is defined by six pillars, including strong democracies for stable future, close friends/enduring ties, economic partnership for prosperity, strategic cooperation for security and stability, social and human development, and neighbors, and global partners (Australian Government, 2020).

A number of key developments have been completed since the 2018 APEC Summit by the Australian government through its Development Assistance Program. Among these was the completion and operation of the Coral Sea Cable System, a fiber optic submarine cable that links Port Moresby and Honiara (Solomon Islands), providing both countries with reliable and affordable telecommunications access and next generation capabilities. Further, Australia provided a short-term loan of US\$300 million to meet a financial shortfall to support economic reform as requested by the Papua New Guinean government (Australian Government, 2020, p. 2). In 2021, Telstra (a telecommunication network in Australia) partnered with the Australian government to acquire Digicel Pacific which is the biggest mobile operator in the South Pacific region, including Nauru, Solomon Islands, Tonga, Vanuatu, Papua New Guinea and Fiji. At the cost of US\$1.33 billion, the Australian government committed to the financial package through the Export Finance Australia to support Telstra's acquisition, reflecting the Pacific Step-Up initiative. Conversely, observers and media reporters concur that this deal is an effort to counter China's influence in the region (Faa et al., 2021; Baird and Kehoe, 2021; Probyn, 2021). As Kehoe (2021) bluntly argues, injecting AU\$1.8 billion of taxpayers' money to "help Telstra buy Digicel Pacific is a totemic shift for a Western government to weaponize corporate power and fight back against the creeping influence of China's Belt and Road Initiative". This deal combines foreign aid, strategic and national security objectives to counter, after years of observing, Chinese state-owned enterprises buy and build strategic infrastructures such as telecommunication systems and ports in the region to boost its strategic influence through soft power diplomacy and espionage powers (Kehoe, 2021).

These recent developments of Australia's engagement with Pacific Island neighbors reflects its foreign objectives through regional and bilateral agreements and initiatives. However, given the intensified geopolitical movements in the region, foreign aid can be used as a strategy to maximize economic opportunities, maintain special relationships with its neighbors as well as response to threats by perceived hostile powers. Conversely, national security emerges as a powerful feature in the provision of foreign aid. The Pacific, a region that has been relatively neglected and economically improvised is attracting attention of superpowers in the region (Hayward-Jones, 2013). It is therefore, necessary to understand the way new forms of imperialism as manifested through globalization and geopolitics impacts on aid to education and how education can cultivate critical understandings to respond to global challenges.

Politics of education reform—global imperatives national implementation

The preceding discussions illustrated how new imperialisms exemplified through globalization and geopolitics impact on development aid. Education systems in postcolonial contexts have grappled with developing contextually and culturally appropriate systems of schooling comparable to global standards of quality. Indeed, global goals for education are noble, designed to address real issues of poverty alleviation and liberation, however the contextual complexities that impede successful translation of global imperatives are deeply entrenched with ongoing uncritical acceptance of educational priorities and models, dependency on foreign finance and conceptualization of educational change. Such practice illustrates neo-colonial practice and imperialism, social and ongoing cultural politics involved in desiring quality education capable of ensuring participation in the competitive global marketplace while simultaneously cultivating a national identity embracing ethnic cultures and wisdom. Again, the Papua New Guinea narrative illustrates the gamut of global imperatives for education in postcolonial contexts, as it demonstrates the complex dialectics of implementing global agendas into educational settings.

From the early 1990s, Papua New Guinea's National Department of Education executed extensive nationwide educational reform. This reform was largely inspired by global developments in education; in particular, the Declaration for Education for All (EFA) movement as a global commitment to provide basic education for children, young people and adults in 1990 and affirmed in 2000 (Cassity, 2008, 2010). An Education Sector Review was commissioned in 1991, tasked to identify and develop strategies to rectify endemic problems in the education system. This review confirmed high attrition rates at the primary level, low transition rates following grades 6 and 10, an irrelevant curriculum, and weak management and administration. It also identified declining resources allocation, high unit costs, and a severe imbalance in the allocation of funding to higher education. A significant change was recommended which led to the national structural education reform in 1993, tasked to respond to poverty alleviation and to increase access to formal education. The main targets included access to nine years of relevant basic education; formal school learning at age six in students' first language; strengthening all curriculum areas targeting improvements in standards and relevance, increased emphasis on relevant practical skills for life, and expanded access to secondary and vocational education (NDOE, 2000; PASTEP, 2002).

Central to this reform were the development of elementary education, the promotion of bilingual education, restructuring of primary and secondary schooling and curriculum, and the development of elementary teacher education programs. This translated in an intensive curriculum reform adopting the Outcomes Based Education (OBE) model as the curricula and pedagogical framework, and elementary education with vernacular as language of instruction, in a country of approximately 860 languages. The Papua New Guinean education landscape was ablaze with reform activities, attracting funding from international donor agencies, including the Australian, New Zealand, Chinese, and Japanese governments, European Union, the Asian Development Bank and the World Bank (McLaughlin, 2011). After two decades of implementation, the OBE model was replaced by standards-based curriculum.

Given the global imperatives on educational reform, elementary education with the vernacular as the language of instruction attracted international aid donors. AusAID provided technical and funding assistance toward elementary teacher education program through the Elementary Teacher Education Support Project (ETESP). An apprenticeship model was adopted by training elementary teacher trainers at the Education Institute in Port Moresby while self-instructional learning packages were designed as home study, supported by trainer-directed training in residential workshops, and supervised teaching in schools (Hahambu, 2011, p. 9). This model commissioned sixteen thousand teachers under 10 years (NDOE, 2008). Notwithstanding, Papua New Guinea's social and cultural contexts are diverse, complex, dynamically changing, and demand qualified teachers with knowledge of the local context, subject disciplines and wider global community. If teacher education and preparation programs are important factors in quality teaching and learning, the process of elementary teacher education in Papua New Guinea remained an incomplete project (McLaughlin, 2018).

The absence of appropriate teacher education in early childhood development and bilingual education resulted in incompetent teaching, misappropriation of cultural and nationally prescribed knowledge content, and whole scale missed opportunities to reform a relevant and appropriate education system. Consequently, the misconception of bilingual education processes, and low academic performance generated tension between elementary and primary teachers, primary and secondary teachers, communities, and the National Department of Education (Hahambu, 2011, p. 9). Further tensions emerged as secondary teachers and community members accused the use of vernacular in elementary schools for students' lower academic performance and raised implications for inequity and quality education (Kombra, 2012; Kombra et al., 2011).

Consequently, the Ministry of Education, issued a policy directive to phase out vernacular as the language of instruction in all elementary schools and instead reinstate English as the medium of instruction in elementary and primary schools (PNG Ministerial Policy Statement, 2013). In 2021, the Minister of Education called for the phasing out of elementary schools and integrating early childhood classes into the primary schools (Patjole, 2018). Such policy reversals raise questions around the mismatched education reform policies between local and global priorities, ownership of educational reforms, maintenance of indigenous languages as an artifact of systems of knowing and being. Serious questions remain around the conceptualization of quality teacher education with the goal of graduating teachers capable of cultivating young minds to negotiate traditional wisdom of their people and navigate challenges of global economics, geopolitical powers and forms of socialization.

In the meantime, Pacific Island Countries signed up to the United Nations Millennium Development Goals (MDGs) containing eight universal aims to reduce extreme poverty, achieve universal primary education, promote gender equality and women empowerment, reduce child mortality, improve maternal health, combat HIV/AIDS, malaria and other diseases, ensure environment

sustainability and global partnerships for development by 2015. Again, Pacific Countries are signatories to the 2016 United Nations Sustainable Development Goals (SDGs) with seventeen ambitious targets to eliminate social ills including eradication of poverty and hunger, fight inequality, promote good health and well-being, quality education, gender equality and tackle climate change by 2030. Conceding that not all MDGs were met globally depending on regions and state of country's development, the SDGs seventeen goals have extended the MDGs, however, a concentrated focus on environment sustainability and climate change is notable. For those Pacific Island Countries whose key interests remain on protection of their islands, land and sea resources, these SDGs including climate change, life below water, life on land, responsible consumption and production, clean water and sanitation, affordable and clean energy, sustainable cities and communities are noteworthy. Targets including partnerships for the SDGs; peace, justice and strong institutions; decent work and economic growth; industry, innovation and infrastructure are particularly useful and invite critical understandings within the context of globalization and geopolitics. However, with SDGs described as "our shared vision of humanity and a social contract between world's leaders and the people", remind us of the authority of superpowers who decide the way these targets are designed and can be accomplished through the provision of foreign finance. It is, therefore, critical that we engage with understandings of how new imperialisms are manifested by global power structures and monitor how these can impact the outcomes of the SDGs by 2030.

Summary and implications for education

This paper mapped out some recent key features that impact on the conceptualization and delivery of foreign aid for development since its inception. With a focus on the Pacific, a region that is vulnerable to global warming and is aid dependent, the narratives herein illustrate a postcolonial perspective illustrating global shifts in the purpose and understandings of the provision and implementation of foreign aid initiatives. These shifts are exemplified through global capitalism and regional events instigated by powers external to the Pacific Island Countries, and as [Hayward-Jones \(2013\)](#) maintains, the center of global economic gravity has indeed shifted to the Asia-Pacific region, thus consequently providing a platform for aid. Nevertheless, in contrast to the post-colonial era, time immediately after political independence, sovereign nations are also implicated in the negotiation and implementation of foreign aid with donor governments, financially powerful corporations and multi-lateral agencies. National and local priorities are often overlooked by the enticement of quick and unsustainable solutions. The implications for educational development and as a process for cultivating mindsets capable to critically engage with geopolitical tensions and new imperialism demands further investigation. New theoretical perspectives beyond neo-colonial and dependency theories are crucial as these are insufficient to illuminate contemporary global complex dynamics.

Within the current geopolitical climate, the Pacific region remains an opportunity and a threat to global superpowers, including Australia ([Schultz, 2014](#)). The extraction of natural resources, with revenues in billions of dollars generated through harvesting resources, including oil, metals and minerals, seafood, and wood products while causing environmental devastation by poisoning of rivers and forests and degradation of food security, create serious challenges for sustainable livelihoods and economic development for Pacific Islanders ([The Guardian, 2021](#)). Combined with the exploitation of resources, Pacific governments and land-owners, given that land remains largely under customary ownership, are short-changed with billions of dollars returning to foreign hands. Communities are left to suffer the consequences of environment degradation and at the frontline of global warming through raising sea levels. Global warming and the threats around submerging of low-laying atolls and communities, the loss of their lands and cultural heritage is the substantial concern of Pacific peoples and needs to be prioritized in negotiations with foreign governments, donors and corporations. Pacific perspectives on climate change are deeply cultural, entrenched in identity, belonging to place (islands) and value systems. The strengthening of regional voices through established regional institutions may provide a stronger bargaining power from collective ontologies of Pacific peoples ([Morgan, 2020](#)).

The role of education in cultivating critical understandings of the processes of and engagement with globalization and new imperialism in this geopolitical climate is imperative. For this, two possible propositions are made. First, these developments on foreign aid relationships pose implications for educational research. Sophisticated conceptual frameworks are required which can illuminate new non-linear, complex and evolving relationalities and between jurisdictions of nation states ([Rizvi, 2009a,b](#)). Second, the shift around aid development discourses directs us to how aid recipients can maintain the longevity of educational project objectives. The Rethinking Pacific Education by Pacific Peoples project (RPEPP), sponsored by New Zealand aid, celebrated its 20th anniversary in 2021 ([Nabobo-Baba, 2013](#); [Sanga, 2006, 2011](#); [Thaman, 2013](#)). This initiative draws on Pacific systems of knowing and being, centering indigenous pedagogies and value systems based on relationality and reciprocity, and nurtures educational leaders and researchers ([Toumua et al., 2021](#); [McLaughlin, 2018](#)).

Contrary to suggestions of "aid fatigue", I argue that aid will continue in the foreseeable future, in formats different from budget support and humanitarian objective. The Pacific region remains resource rich offering future economic opportunities while threats to donor governments' national security will direct aid flows. It is, however, that way global systems and agendas driving foreign financial assistance are understood, negotiated, and articulated, that presents a new challenge for Pacific Island Countries in the 21st century.

References

- Adams, P., Solomon, L., 1991. In the Name of Progress - the Underside of Foreign Aid. Earthscan, London.
- APEC Ministerial Chair, 2018. 2018 APEC Ministerial Meeting Chair's Statement. Government of PNG, Port Moresby.
- Australian Government, 2020. 2019–2020 Papua New Guinea Development Program Progress Report. Government of Australia, Canberra.
- Baird, L., Kehoe, J., 2021. Taxpayers bankroll Telstra's \$2.1b Digicel buy. Aust. Financ. Rev. 25 October. <https://www.afr.com/companies/telecommunications/telstra-snaps-up-digicel-for-2-1b-with-government-backing-20211025-p592s7>.
- Brecher, J., Costello, T., 1994. Taking on the multinationals. *Nation* 259 (21), 757–760.
- Cassidy, E., 2008. More effective aid policy? AusAID and the global development agenda. *Int. Educ. J.* 9 (2), 2–17.
- Cassidy, E., 2010. New partnerships and education policy in Asia and the Pacific. *Int. J. Educ. Dev.* 30, 508–517.
- Coxon, E., Munce, K., 2008. The global agenda and the delivery of aid to Pacific education. *Comp. Educ.* 44 (2), 147–165.
- Crossley, M., Tikly, L., 2004. Postcolonial perspectives and comparative and international research in education: a critical introduction. *Comp. Educ.* 2, 147–156.
- Dorman, M., Pryke, J., 2017. Foreign aid to the Pacific: trends and development in the twenty-first century. *Asia Pac. Stud.* 4 (Policy 3), 386–404.
- Dorney, S., 1998. The Sandline Affair. ABC Books, Sydney.
- Faa, M., Dziedzic, S., Burgess, A., 2021. Telstra to Buy Pacific Arm of Telecommunications Gaint Digicel with Canberra's Support Amid China's Rising Influence. 25 October.
- Fagerlind, I., Saha, L., 1989. Education and National Development: A Comparative Perspective, second ed. Pergamon Press, Oxford.
- Ferguson, N., 2003. Empire: How Britain Made the Modern World. Penguin, London.
- Ferro, M., 1997. Colonisation: A Global History. Routledge, London.
- Flint, C., 2006. Introduction to Geopolitics. Routledge, London.
- Foucault, M., 1991. Governmentality. In: Burchell, G., Gordon, C., Miller, P. (Eds.), The Foucault Effect: Studies in Governmentality. University of Chicago Press, Chicago, pp. 87–104.
- Fox, P., 2015. International Asylum and boat People: the Tampa affair and Australia's "Pacific solution". *Brook. J. Int'l L.* 40 (2), 356–373.
- Fox, C., 2016. Who is my neighbour? Unleashing our postcolonial consciousness. *Int. Educ. J.* 15 (3), 57–76.
- Fraenkel, J., 2016. Australia's detention centres on Manus Island and Nauru: an end of constructive Pacific engagement? *J. Pac. Hist.* 51 (3), 278–285.
- Funaki, K., 2016. New visions for international aid: perspectives from the Pacific islands. *Jpn. Soc. Pac. Island Stud.* 4, 39–62.
- Gani, A., 2008. Governance and foreign aid in Pacific island countries. *J. Int. Dev.* 21, 112–125.
- Gramsci, A., 1971. Selections From the Prison Notebooks of Antonio Gramsci. Q. Hoari & G. Smith, Trans. Lawrence & Wishart, London.
- Grewcock, M., 2018. Systems of exclusion: the Rudd government and the "end" of the Pacific solution. *Curr. Issues Crim. Justice* 19 (3), 363–368.
- Hahambu, C., 2011. Perspectives of Teaching in Papua New Guinea: A Case Study of Elementary Teacher Trainers and their Mentors. QUT Unpublished Thesis, Brisbane.
- Hawksley, C., 2009. Australia's aid diplomacy and the Pacific Islands: change and continuity in middle power foreign policy. *Global Change Peace Secur.* 21 (1), 115–130.
- Hayward-Jones, J., 2013. Big Enough for All of Us: Geo-Strategic Competition in the Pacific Islands. Lowy Institute, Sydney.
- Henry, M., 1999. Working with/against globalisation in education. *J. Educ. Policy* 14 (1), 85–97.
- Hickling-Hudson, A., 1999. Globalisation, postcolonialism and educational change. In: Meadmore, D., Burnett, B., O'Brien, P. (Eds.), Understanding Education: Contexts and Agendas for the New Millennium. Prentice Hall, Sydney, pp. 82–90.
- Hickling-Hudson, A., 2009. Southern theory and its dynamics for postcolonial education. In: Coloma, R.S. (Ed.), Postcolonial Challenges in Education. Peter Lang, New York, pp. 365–376.
- Huang, X., 2020. An analysis of the APEC news in Washington Post from the perspective of engagement system based on appraisal theory. *Theor. Pract. Lang. Stud.* 10 (9), 1054–1059.
- Hughes, H., 2003. Aid Has Failed the Pacific. Issue Analysis No. 33. Centre for Independent Studies, Sydney.
- Johansson-Fua, S., 2016. The oceanic researcher and the search for a space in comparative and international education. *Int. Educ. J.* 15 (3), 30–41.
- Kehoe, J., 2021. Digicel deal a totemic shift for a Western government. Aust. Financ. Rev. 25 October. <https://www.afr.com/policy/foreign-affairs/digicel-deal-a-totemic-shift-for-a-western-government-20211025-p592tb#:~:text=injecting%20%241.8%20billion%20of%20taxpayer,China's%20Belt%20and%20Road%20Initiative>.
- Komra, U., 2012. Front Page/Back Page Schools: A Case Study of Factors Perceived to Affect Academic Performance in Five Papua New Guinea Secondary Schools. Unpublished PhD Thesis. QUT, Brisbane.
- Komra, U., Hickling-Hudson, A., McLaughlin, J., Pillay, H., 2011. The Effects of Social Relations on Academic Performance Within Selected Secondary Schools in Papua New Guinea: A Study of Teacher/Student Perceptions. Paper presented at The UKFIET Conference. Oxford: Oxford University.
- Larcertosa, J.L., 2014. Unfriendly shores: an examination of Australia's "Pacific Solution" under international law. *Brooklyn J. Int. Law* 40 (1), 321–356.
- Luteru, P.H., Teasedale, G., 1993. Aid and education in the South Pacific. *Comp. Educ.* 29, 293–305.
- Marcuzzo, M.C., Sunanda, S. (Eds.), 2018. The Changing Face of Imperialism: Colonialism to Contemporary Capitalism. Routledge, London.
- McLaughlin, J.M., 2002. The Outcomes of the Australian/Papua New Guinean Secondary School Students' Project: An Analysis From a Postcolonial Perspective. Unpublished PhD Thesis. QUT, Brisbane.
- McLaughlin, J.M., 2011. Lost in translation: partnerships for authentic education in Papua New Guinea. *Int. J. Educ.* 10 (2), 86–98.
- McLaughlin, J.M., 2018. Shifting boundaries and education development discourses: implications for comparative and international education in Oceania. *Annu. Rev. Comp. Int. Educ.* 34, 205–230.
- McLaughlin, J.M., Hickling-Hudson, A., 2005. Beyond dependency theory: a postcolonial analysis of educating Papua New Guinea secondary school students in Australian schools. *Asia Pac. J. Educ.* 36 (s1), 118–130.
- Moreton-Robinson, A., 2003. I still call Australia home: indigenous belonging and place in a postcolonising society. In: Ahmed, S., Fortier, A.M., Sheller, M., Castaneda, C. (Eds.), Uprootings/Regroundings: Questions of Home and Migration. Berg Publications, Oxford, UK.
- Morgan, W., 2020. Oceanic Hydro-politics: Exploring Contemporary Geopolitics in the Blue Pacific. Regional Outlook Paper, 64. Griffith Asia Institute, Brisbane.
- Nabobo-Baba, U., 2013. Transformations from within: rethinking Pacific Education Initiative. The development of a movement for social justice and equity. *Int. Educ. J.* 12 (1), 82–98.
- National Department of Education (NDOE), 2000. The State of Education in Papua New Guinea. National Department of Education, Port Moresby.
- National Department of Education, 2008. State of Education Report. NDOE, Port Moresby.
- Nkrumah, K., 1965. Neo-colonialism. The Last Stage of Imperialism. Thomas Nelson & Sons, Ltd., London.
- Nyuyen, N., 2020. On geopolitics and education: intervention, possibilities and future directions. *Geogr. Compass* 14, 1–11.
- Ó Tuathail, G., 1996. Critical Geopolitics. University of Minnesota Press, Minneapolis, MN.
- O'Farrell, C., 2005. Michel Foucault. SAGE, London.
- Pacific Community, 2020. Pacific Islands populations. Statistics for Development Division, SPC, Noumea.
- Papua New Guinea Ministerial Statement, 2013. Ministerial Policy Statement no 01/2013. NDOE, Port Moresby.
- PASTEP, 2002. Primary and Secondary Teacher Education Project. NDOE and AusAID, Port Moresby.
- Patjole, C., 2018. Elementary Schools to be Phased Out by 2021. Loop News PNG, 30 May.
- Pavlov, V., Sugden, C., 2006. Aid and growth in the Pacific islands. In: Asia Pacific Economic Literature. Australian National University & Blackwell Publishing, Canberra, pp. 38–55.

- Probyn, A., 2021. Australian Taxpayers Could Help Buy Pacific of Telecommunications Giant Digicel to Stop China from Nabbing it First. Australian Broadcasting Commission News Analysis. 20 July.
- Richards, L., 1977. The context of foreign aid: modern imperialism. *Rev. Radic. Polit. Econ.* 9 (4), 43–75.
- Rizvi, F., 2007. Debating globalisation and education after september 11. *Comp. Educ.* 40 (2), 157–171.
- Rizvi, F., 2009a. Postcolonialism and globalisation in education. In: Coloma, R.S. (Ed.), *Postcolonial Challenges in Education*. Peter Lang, New York, pp. 23–45.
- Rizvi, F., 2009b. Global mobility and the challenges of educational research and policy. *Teach. Coll. Rec.* 111 (14), 268–289.
- Rizvi, F., Lingard, B., 2010. *Globalising Education Policy*. Routledge, London.
- Said, E., 1993. *Culture and Imperialism*. Vintage Books, London.
- Sanga, K., 2006. Brief Summary of the Rethinking Pacific Education Initiative For and By Pacific Peoples (RPEIPP). Report to AZAID, Wellington: New Zealand.
- Sanga, K., 2011. Re-framing Pacific regional service delivery: opportunity spaces for together and apart. *Int. Educ. J.* 10 (2), 22.
- Sanga, K., 2016. What if form was the aid? Possible opportunities for renewed neighbourliness in aid relationships. *Int. Educ. J.* 15 (3), 7–15.
- Satre, J.-P., 2001. *Colonialism and Neocolonialism*. Translated by A. Haddour, S. Brewer & T. McWilliams. Routledge, London.
- Scholte, J.A., 2000. *Globalisation: A Critical Introduction*. Palgrave, New York.
- Schultz, J., 2014. Theorising Australia–Pacific islands relations. *Aust. J. Int. Aff.* 68 (5), 548–568.
- Sen, S., Marcuzzo, M., 2018. Introduction. In: Marcuzzo, M. (Ed.), *The Changing Face of Imperialism: Colonialism to Contemporary Capitalism*. Routledge, London, pp. 1–9.
- Spring, J., 1998. *Education and the Rise of the Global Economy*. Lawrence Erlbaum Associates, New Jersey.
- Taylor, S., 2005. The Pacific Solution or a Pacific nightmare? The difference between burden shifting and responsibility sharing. *Asian Pac. Law Pol. J.* 6, 1–43.
- Taylor, S., Rizvi, F., Lingard, B., Henry, M., 1997. *Educational Policy and the Politics of Change*. Routledge, London.
- Thaman, K., 2005. The Future of the Past: Teachers as leaders in advancing cultural and cognitive democracy in education in Oceania. WIPCE Conference Keynote Presentation.
- Thaman, K., 2013. Quality teachers for indigenous students: an imperative for the twenty-first century. *Int. Educ. J.* 12 (1), 98–118.
- The Guardian, 2018. Not a Magic Pill: Readers in PNG Respond to the APEC Summit, 12 November.
- Thomas, R.M., 1993. Education in the South Pacific: the context for development. *Comp. Educ.* 29 (3), 233–248.
- Tikly, L., 2004. Education and new imperialism. *Comp. Educ.* 40 (2), 173–198.
- Tikly, L., 2009. Education and new imperialism. In: Coloma, R.S. (Ed.), *Postcolonial Challenges in Education*. Peter Lang, New York, pp. 23–45.
- Tikly, L., Bond, T., 2013. Towards a postcolonial research ethics in comparative and international education. *Compare: A Journal of Comparative and International Education*. 43 (4), 422–442.
- Toumu'a, R., Sanga, K., Johansson, F. (Eds.), 2021. *Weaving Education Theory and Practice in Oceania*. University of South Pacific, Nuku'alofa, Tonga.
- UNESCO, 2021. 2021/2022 Global Education Monitoring Report. Non-state Actors in Education—Who Chooses? Who Loses? UNESCO, Paris.
- Wall, J., 2020. China's Debt-Trap Policy is About to Challenge Papua New Guinea—and Australia. Australian Strategic Policy Institute – The Strategist, 8 September.
- Wen, P., Greenfield, C., Barrett, J., Westbrook, T., 2018. APEC Fails to Reach Consensus as U.S.—China Divide Deepens. Reuters. November, 18.
- Yamahata, Y., 2015. Decolonising World Politics: Anti-colonial Movements Beyond the Nation-State. *E-International Relations*. <https://www.e-ir.info/2019/12/15/decolonising-world-politics-anti-colonial-movements-beyond-the-nation-state/>.
- Zadja, J. (Ed.), 2015. *Second International Handbook on Globalisation, Education and Policy Research*. Springer, Dordrecht, Netherlands.

Websites

- 2018 Australian Broadcasting Commission. APEC 2018. (Accessed 02 February 2022).
- 2018 Stepping-Up Australia's Engagement with our Pacific Family. (Accessed 02 February 2022).
- 2021 Telstra to buy Pacific arm of Telecommunications Giant Digicel with Canberra's Support amid China's Rising Influence. (Accessed 04 February 2022).
- 2021 Australian Taxpayers could Help Buy Pacific Arm of Telecommunications Giant Digicel to Stop China from Nabbing it First. (Accessed 05 February 2022).
- 2021 The Pacific Plunder: the Extraction of Valuable Resources from Across the Pacific has made Companies Billions But Comes at a Sometimes Catastrophic Cost. (Accessed 01 December 2021).

Six decades of development: the impact of structural adjustment programs on education in developing countries

Óscar Espinoza^a and Noel McGinn^b, ^aFacultad de Educación y Humanidades, Universidad de Tarapacá, Arica, Chile; and ^bSchool of Education, Harvard University, Swampscott, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	205
The first two development decades 1960–79: the introduction of structural adjustment	206
The third development decade 1980–1989: rethinking structural adjustment loans	207
Mexico	208
Chile	208
Ghana	208
Zambia	209
A broad overview	209
The fourth decade: structural loans and programs during the 1990s	210
Latin America	210
Tanzania	210
Former Soviet republics	210
Summary	210
2000—development decades five and six: structural adjustment in new clothes?	211
Partnerships	211
Poverty reduction strategy papers	212
The knowledge society	212
Results-based lending	213
Summary: twenty-first century observations on structural adjustment in education	214
References	214

Introduction

National education systems originated in the pursuit of a common language and culture, thought to be essential for the political and economic growth of a nation-state. To reach a large portion of their population, governments designed schools like the factories just beginning to appear in European economies. Mass production of loyal citizens could be achieved by imposing a common curriculum, training of teachers, schedules and supervision (McGinn, 1987). Thus, even though each new nation-state had its own culture, their national educational systems shared some elements of structure and practice. Trade, emigration and tourism provided opportunities to learn about other practices. These similarities found across national systems were the result of relatively balanced processes of lending and borrowing (Cowen, 2012; Steiner-Khamsi, 2004, 2012).

In many other countries, however, school systems were developed as the result of an imposition by their colonial masters. In these cases, schooling originally was designed to acculturate only a small portion of the population, in values and skills that would serve the colonial masters. Colonial schools educated only those few chosen to represent the empire's interests.¹

Over time, the variety of forms of educational systems was reduced, both by lending and borrowing and by colonization. One result has by change in the original objectives of schooling. Some systems now focus more on teaching of work-related knowledge and skills and give little importance to citizen formation and cultural traditions. One analyst has proposed that today there are six major types of national systems, those of Germany, France, the United Kingdom, Japan, the United States, and Russia (Cummings, 1999, 2003). Each of these types of education systems has shaped the systems of other countries. As they themselves have borrowed from others, their variety is more like the broad spectrum of a rainbow rather than few distinct colors.

In recent years, some have come to believe that education systems should (and are) converging toward a common mission and structure. This conviction is fostered by observations of the apparent growing similarity of economic, political, and social institutions and practices. The ideational forces that led to the formation of the United Nations, emphasizing common goals and universal rights, are one reflection of the pursuit (by some) of a generic definition of progress or development. Some belief that human progress can best be achieved by following a unilinear path toward the ideal society. Movement along this path would have five stages,

¹In 1835, Lord Macaulay, advising the Corporation ruling India for the English Crown, proposed that the Indian education system be designed to produce: "a class who may be interpreters between us and the millions whom we govern; a class of persons, Indian in blood and color, but English in taste, in opinions, in morals and in intellect".

the highest of which would be high mass consumption (Rostow, 1960).² To measure progress along the path, a variety of indicators have been developed (United Nations, 1971) some aggregated to a single expression, such as Gross National Product (GNP).

As levels of educational attainment are higher in wealthier countries than in poor, education was considered essential for development (Harbison and Myers, 1964). The results of education were defined as a form of capital, which (capitalist) economic theory posited as the dominant factor in production and economic growth. Education is the major producer of human capital; countries seeking to develop should therefore invest in education (Schultz, 1961; Becker, 1962).

The identification of relationships between social and economic characteristics of a country and its level of development encouraged the elaboration of a theory that argued that with increased contact the economies (nations, societies, cultures) of the world were becoming or would over time become more alike. This progress was driven by the assimilation of the values, practices, skills of the more modern, developed nations (Inkeles, 1966; Kahl, 1968; Boli et al., 1985; Reich, 1992; Meyer et al., 1997).

Convergence in educational policies and practices, it was argued, would accelerate a country's progress along that path. Modernization required replacing the institutional structures found in less developed countries with those found in the more advanced countries. As one enthusiast put it:

Clearly educational systems must adjust to the new global environment ... the appropriate response is a restructuring of the policy changes [that had been] dictated by local needs and responses ... policies and practices formed around a logic of local schooling for local needs are no longer viable. Neither is a system of centrally controlled national schooling. What is needed is a system of schools which serve specific populations with specific resources and students. Their future workplace is the global market place (Ilon, 1994, pp. 100 and 104).

The horrors of World War II and the formation of the United Nations prepared the ground for a more concerted effort to build a world culture. "Modernization" of education and other social institutions moves countries along that path. To that end, the Western allies in the UN agreed to provide the financial resources necessary to assist developing countries. The administration of funds subscribed by the participating countries was to be mainly the responsibility of two organizations, the International Monetary Fund and the World Bank.³ In addition to other activities, these two organizations made interest-bearing loans to developing countries. The Bank also provided technical assistance, often in the form of expatriate resident advisors paid through the loan.

At the same time, a number of countries created their own development agencies. These provided grants, rather than loans to developing countries. These included USAID, DFID of the United Kingdom, IDRC of Canada, SIDA of Sweden, JICA of Japan and others. Their policies and practices were not necessarily aligned to those of the IMF and the World Bank or those of other countries.

The next sections of this article describe the origin, intent, content and implementation of structural adjustment programs affecting loans for education. The UN's six Development Decades frame the changes that occurred in IMF and WB loans to national education systems. Not all developing countries received IMF and WB loans, and relatively few reports of their outcomes have been published. Where available, country-level detail is provided.

The first two development decades 1960–79: the introduction of structural adjustment

The UN's First Development Decade began in 1960. The UN agencies placed major emphasis on achieving economic growth in poorer nations, to be stimulated by transfers of capital. Increased national revenues in recipient countries, it was believed, would prompt increased public spending on education and health. The lenders (IMF and World Bank) anticipated that significant advances would be achieved by the end of the decade. And in fact, many of the low-income countries had by 1970 achieved economic growth rates of 5%. Excited by what seemed to have been the first step toward a Golden Age, the United Nations General Assembly, announced a Second Development Decade.⁴

For various reasons, the Second Decade was less successful. Inflation in the United States associated with the costs of the war in Vietnam and pursuit of the Great Society, led to US withdrawal from the Bretton Woods Agreement that had controlled currency exchange rates (Bordo, 2017). This contributed to price instabilities around the world. Between 1973 and 1975 the economies of Western Europe and the United States slowed significantly. The developed countries reduced their assistance to developing countries. Economic growth in newly industrializing countries created instabilities in the economies of the more developed nations. Oil producing countries reduced their production quotas resulting in a rapid rise in gasoline prices. Economists opposed to the use of economic planning (by the state) called for a market system (Toye, 1987). Confidence in state-level economic planning declined; economic growth was characterized principally by expansion of the private sector and the development of a market economy with competition and competitive pricing. International trade increased, as did cooperation between countries. National government planning and economic control diminished (Cable, 1995).

²It may be remembered that one of the popular indicators of a country's level of development was its consumption of coal and oil.

³The regional development banks (Inter American Development Bank, African Development Bank, Asian Development Bank), which followed, were included in this coordination. For a description of their linkages see Group of Twenty (2017). The OECD, originally called the Organization for European Economic Cooperation, was created by the United States and the European nations to implement the 1948 Marshall Plan.

⁴Dramatic changes in education, health and general prosperity, however, were not yet visible by 1970 (Koehler, 2015).

In order to continue their efforts to expand their economies, more developing countries then turned to the International Monetary Fund for loans.⁵ The collapse of the Bretton Woods agreement on exchange rates required the IMF to monitor its loan agreements more carefully. The creditor countries that supplied the majority of the funds to the IMF found it in their interest to cooperate with this surveillance. Borrower countries were obliged to acquiesce. Beginning in 1979 some developing country governments changed financial policies in order to receive IMF loans (Boughton, 2000).

The World Bank, also a creature of the United Nations, works closely with the IMF (Driscoll, 1995). While the IMF was assigned responsibility for the international monetary system, the World Bank's mission was "to reduce poverty and improve living standards by promoting sustainable growth and investment in people" (World Bank, 1998). The Bank's principal tool for the realization of its mission was loans to the poorer countries. During its early years, Bank loans had been limited primarily to basic infrastructure such as roads, dams, irrigation systems, electricity production. By 1970, its attention had expanded to include poverty reduction, education, and health.

In 1968, a World Bank Discussion Paper argued that as the world economy was becoming more knowledge-intensive, all countries had to modernize their education systems. To do this,

Low-income countries ... need to develop the scientific personnel who will understand fully the latest technological advances coming out of the industrialized countries and be able to adapt and apply them for local production of goods and services.

(Haddad et al., 1990, p. 72).

In exchange for financial aid, governments were required to accept the Bank's economic development strategies; these emphasized greater participation in the global economy. In the Philippines, for example, the government was required to abandon its emphasis on import substitution in order to redistribute income, in favor of an export promotion strategy that would increase economic growth (Ascher, 2009).

Most of the economic growth of the 1970s was financed by loans and credits. The amount of debt which poor countries carried sometimes exceeded their GNP. The 1979 regime change in Iran, which interrupted the global flow of oil, led to a sharp increase in world prices and increased inflation. Developing countries that had funded economic growth with loans faced increased interest rates. This reduced government spending.

In Latin America, non-democratic governments were more likely to reduce government spending on social services and especially education. Democratically-elected governments most often reduced other forms of spending and sometimes increased spending on education (Reimers, 1990).

Some countries were no longer able to meet scheduled repayments of loans from commercial banks located in the more developed nations. They turned to the IMF for assistance. In 1982 the Mexican government announced its inability to repay its international debts, exacerbating the world economic situation (FDIC, 1997).

Especially after 1979, the IMF and the World Bank began to insist on Structural Adjustment Programs (SAPs) as the basis for loans made to countries. The SAPs specified the conditions that countries must meet prior to receiving funds from the loan. The principles were laid out in what came to be known as the Washington Consensus. That label was eventually applied to a set of 10 policy actions described as reflecting what many economists in the IMF and the World Bank thought were necessary for modernization (Espinoza, 2002; Williamson, 2004).

Although conceived with Latin America in mind, the underlying principles of the Consensus were considered relevant for all countries. The recommended policies included avoidance of budget deficits, more public spending on sectors with high economic yields, lower taxes, privatization of public enterprises and secure property rights. They represented a turn away from import substitution and a dynamic public sector, toward liberalization, a market system and a smaller state apparatus. They came to define neo-liberal economics.

Given the evident success of the industrialized capitalist nations, neo-liberal economists presumed that convergence toward free markets would result in maximum efficiency in the use of resources. A free market in education requires, it was believed, less direct government control and involvement, and support of a private market by direct grants or vouchers to consumers. Holding educational workers "accountable" for student learning would insure higher productivity.

The third development decade 1980–1989: rethinking structural adjustment loans

National governments reacted in different ways to the imposition of conditions on structural adjustment loans. Several examples illustrate the variety of reactions of recipient countries and the extent to which outcomes matched the expectations of the Washington Consensus. Our focus is on the effect of structural adjustment loans and programs on economic growth, education and social equality.

⁵The IMF at the time had about 40 member countries; the republics of the Soviet Union remained outside.

Mexico

The first Structural Adjustment Program was proposed by the IMF in 1983. Overwhelmed with debt, in 1982 Mexico had threatened to stop payment on its loans. In response, the IMF insisted that Mexico take four major actions: decrease spending by the government; devalue the national currency; promote exports rather than substitution; and encourage or at least permit foreign investment in Mexican enterprises (Hellman, 1997).

Implementation of the SAP had unexpected results. Real wages dropped to levels of 20 years earlier (Latapí Escobar and González de la Rocha, 1995). Foreign-owned factories introduced automatized production technologies, lowering the demand for more highly educated labor (Alarcón-González and McKinley, 1999). The Mexican government reduced its budget by spending less money on health and on secondary and higher education. Looking back, one analyst likened the experience to a “perfect storm” in which a series of unexpected events combined producing an undesired outcome (Kafferstan, 2017).

Chile

The military junta that seized power in Chile in 1973 implemented economic and social policies consistent with the logic of the IMF and World Bank’s structural adjustment loans (Espinoza, 2008).⁶ They included reduction of government spending, openness to trade, privatization of public enterprises and free markets. In higher education, the government sought to have students or their families share in the costs of teaching, and established a competitive market among universities. Costs of higher education soared. In 1981 the Ministry of Education initiated a policy of vouchers for both public and private schools based on attendance. The voucher amount was fixed, independent of family income. Private school enrollment in primary and secondary schools increased dramatically, public school enrollment increased only in higher income neighborhoods. Faced by rapidly mounting inflation, the government imposed a fixed exchange rate for the national currency. This worsened the balance of payments with other countries, increasing inflation. Government spending on education declined by 20%. By 1983 Chile was in a recession (Ritter, 1990).

In 1985, with a new Minister of Finance and structural adjustment loans from the IMF and the World Bank, Chile implemented a new set of macroeconomic policies consistent with the principles of the Washington Consensus. Economic growth and employment generation followed (Ritter, 1990). By 1990 the budget was in balance, inflation was declining, and recently de-nationalized industries were expanding their production.

The effects on education were mixed: enrollment at all levels increased in number but disproportionately across income groups. Private school enrollment (which was primarily drawn from upper income groups) continued to increase as a proportion of total enrollment. As the government reduced its subsidies to public universities, the universities introduced special fees and became increasingly selective in admissions. Students unable to afford four years of university enrolled in 1- or 2-year technical training centers. Less selective (and even more expensive) private universities increased their enrollments. Many lower-income students financed their university studies with loans from commercial banks (Espinoza, 2008). Structural reform in Chile had enabled the country to move from a system of “elite” higher education to one of “mass” higher education (Trow, 1973), but the mass system dominated by private universities reinforced its elite character. By 1990, 72% of students in private universities were from families in the top 20% income bracket, compared to 4.2% whose families were in the lower 40% of the household income distribution (Larrañaga, 1992).

In 1989, after a national plebiscite, the military government was replaced by a coalition of left and center parties. Macroeconomic policies were left intact with only minor changes, but new social policies reversed the earlier orientation toward reduced government involvement in education and health.

Ghana

A SAP in Ghana in 1984 enabled the country to recover from a recession that had persisted from 1977. The inflation rate dropped from over 100% in 1983 to below 40% in 1984 and continued to decline. Economic growth, which had been negative, averaged over 5.6% for the next 6 years. The success of Ghana was such that it was considered a model country. USAID, the World Bank and the IMF named Ghana as one of its outstanding successes.

Ranked on the basis of growth, savings, exports and investments, the five most successful adjusting countries were, in order: South Korea, Mauritius (dubbed the Korea of Africa), Morocco, Ghana and Thailand.

(World Bank 1991, p. 3).

Economic stabilization had been made possible by establishment of a powerful state apparatus willing to use violence to insure compliance with its policies (Kraus, 1991). The military leader seizing power in Ghana in 1981 tolerated no criticism of the policies of economic recovery recommended by the IMF and World Bank. His refusal to re-negotiate or to soften the recommended policies provoked considerable resentment among the people of Ghana. For example,

⁶The military’s advisors were economists trained at the University of Chicago influenced by Milton Friedman, whose ideas also were absorbed by others espousing the premise of a global economy.

The Trade Union Congress (TUC) ... vigorously protested SAP policies, students the rising school fees and university costs, peasants new borehole and well fees, and consumers the end of subsidized goods and rising prices. Businessmen ... protested restraints on credit and 30% interest rates, industrial managers the liberalization of trade which undercuts the sale of Ghanaian goods, and some state corporations their sale to private interests or liquidation. Intellectuals, with little access to the public, ...criticized SAP policies by pointing out their inadequacies, inequities, and social costs. (Kraus, 1991, pp. 19–20).

The government survived continuous opposition during the 1980s, in part because of the political actions taken to contain protest, and in part because of cultural values of the Ghanaian people and perhaps sub-Saharan Africans in general (Haynes, 1991). In distinction with the United Nations Declaration of Human Rights, or the rights listed in the constitutions of Europe or the United States, Africans may give more importance to social harmony rather than to individual freedom. The “people” are entitled to be treated with the same dignity as the “person”.⁷ A comparative review of theories of human rights noted that from this perspective, collective rights should be honored first, then economic and social rights. Only then are political and civil rights respected (Vincent, 1981).

During the 1980s public spending on education in Ghana was lower than it had been during the 1970s. The government instituted fees for books and residence in secondary and higher education institutions. Schools were left lacking fundamental services. Only beginning in 1989 did the Ghanaian government began to restore drastic cuts made in health services and education (Kraus, 1991). By then the government had organized supporters into “community watch” groups that helped control unrest. The government’s success and that of Ghana could be attributed to factors not anticipated by the IMF or the World Bank.

Zambia

The implementation of structural adjustments during the last years of Zambia’s Second Republic were marked by high levels of public opposition. The program begun in 1983 called for a currency devaluation, a cap on wage increases, a reduction in government employment, and a currency exchange auction system which benefitted foreign commercial firms and contributed to increased prices of commodities. At the same time public expenditure on education was reduced, affecting the purchasing power of teachers and access to schooling. Education’s share of the national budget was between 1984 and 1988 reduced by 8%, and spending per student fell by 60%. The secondary school gross enrollment ratio dropped from 30 to 20% of the age group (Simutanyi, 1996).

As in Ghana, Zambia was governed by one party. And, as in Ghana, there was considerable opposition to the effects of the structural reform agreements, from the business sector as well as from the labor movement. Also as in Ghana, policies were imposed by the president without agreement from those outside his party. Unlike Ghana, the opposition in Zambia was able to generate enough resistance to frustrate the government’s efforts (Simutanyi, 1996).

The effects of structural adjustment in Zambia can be attributed to a different set of contingencies than those that operated in Ghana. As one analyst has written, in order to succeed a government must be determined to impose its will but also know how to insulate itself from pressures by those in opposition (Simutanyi, 1996). Zambia had prior experience with structural adjustment loans, but the one-party government then in power (1973–91) had not been able to carry through policies promoted by the World Bank and IMF. The movement to restore democracy was motivated (in part) by opposition to structural reforms. Although they had competing objectives, organized labor cooperated with business interests to force out the government. They then negotiated a series of compromises that led to significant reform.

A broad overview

The effect of SALs and SAPs between 1980 and 1989 has been analyzed by comparing government spending on education during 1980–1989 in the countries receiving loans with that in 50 other countries not receiving loans listed in the (World Bank’s, 1992) report on lending (Reimers, 1994). During the period the Bank initiated structural adjustment loans or programs in 57 countries around the world (World Bank, 1992). Some 27 countries experiencing economic difficulties received two loans during that period, another 30 received at least one loan by 1990.

Using World Bank data, Reimers (1994, 1997) first examined whether government expenditures on education changed (in proportion to GNP) between 1980 and 1990. He assumed that a reduction in expenditures would indicate that the government considered other sectors to be of more importance than education. In general, countries receiving loans reduced their spending on education in proportion to GNP. This response was more marked among the sub-Saharan African countries, but also true in Latin America. Overall, Asian countries receiving loans increased their spending on education in relation to their GNP. Within each category of countries (e.g., region, number of loans) there is considerable variation in GNP. There is no association between 1980 level of GNP and amount of change between 1980 and 1990.

The next comparison was with respect to changes in government expenditures on education. The effect of the austerity induced by the loan conditions was an overall reduction of government spending. On average, although total budget expenditure was reduced, the proportion allocated to education increased. This was true for the Latin American and African recipients, but not

⁷Paraphrasing the Banjul Charter on Human and People’s Rights, passed unanimously in June 1981 at the 18th assembly of heads of state and government of the Organization of African Unity (Vincent, 1981, p 39).

for the Asian countries, who reduced their (already higher than in the other regions) spending on education (but increasing in relation to GNP).

The increase in expenditures in Africa and Latin America more heavily favored higher education although those regions had lower gross enrollment ratios in primary and secondary education than did Asia. Overall, by 1990 African countries that had received two loans were spending 77 times as much per higher education students compared to primary students, compared to a ratio of 38 to 1 in countries not receiving loans. The ratio had been 70 to 1 in 1980. In Latin America and Asia higher level education students cost only 7 times as much as primary students. The analysis concluded that:

The politics of adjustment favoured the groups with greater lobbying power, namely urban dwellers and those from higher income groups. This explains why the disparities in allocation between education levels remained, or in some cases worsened, in spite of adjustment.

(Reimers, 1997, p. 16).

The fourth decade: structural loans and programs during the 1990s

Latin America

Additional structural adjustment loans and programs were made during the 1990s. Some were seen as highly successful, others the target of strong criticisms. The policy actions of SAPs were judged as inequitable, as they affected some groups in society more than others. For example, a reduced federal budget in Mexico lowered food subsidies, on which many unemployed single mothers depended. In six Latin American countries, the effect of World Bank policies between 1990 and 1999 was reduced economic growth and great income inequality (Espinoza, 2002, 2008; Bonal, 2004).

Tanzania

Noting that “inter/national structural adjustment policies rarely attend to local settings”, an ethnographer described their impact on women and children in a rural village in Tanzania (Vavrus, 2005). Under the SAP, the government agreed to commercialize agriculture, and removed subsidies for fertilizer. At the same time, it reduced duties on food imports which resulted in competition with products from other countries. While food prices dropped so did sales of local production. Foreign investors introduced new technologies and workers from other countries created high unemployment. The GDP increased but so too did foreign debt and government spending on education. Youth unemployment increased, affecting women more than men (Buchmann, 1996).

Former Soviet republics

After the end of the Soviet Union in 1989, the IMF and World Bank offered structural adjustment loans and advice to former Soviet republics. All three of the Baltic states, Estonia, Latvia, and Lithuania had, with the collapse of the Soviet economy, suffered a severe economic shock, the worst in their recorded history. While in the Soviet Union they had enjoyed a relatively high standard of living, with high levels of employment, free education and access to health services. By 1992 unemployment had increased to 17%, and the life span of men reduced by 3 years. Once Lithuania joined the European Union, many young people migrated to the United Kingdom seeking to learn English. One commentator noted that “The more severe damage seems to have been inflicted on people’s sense of security” (Harris, 1995).

Reports of the withdrawal of government funding from social services in recipient countries was, in 1997, lamented by the Director of the IMF and again in 2001 (reported in executive board minutes 95/55 and 00/121). But structural adjustment in the former Soviet nations had far less effect on education than in it did in Africa and Latin America. Under Soviet rule education had received given high priority; governments spent proportionately more on public education than any except those of the most advanced capitalist nations. Given their well-educated populations the former Soviet nations recovered from their economic shocks much more quickly than elsewhere.

Summary

Structural adjustment in the 1980s and 1990s worked well in some countries but had little effect in others. Reduced spending on education had disastrous effects in some countries. Reviewing 20 years of experience with structural adjustment in Africa, analysts concluded that by not paying attention to the effects of reduced government spending on the most vulnerable members of society it had been “an inadvertent enemy of human development” (Geo-Jaja and Mangum, 2001). Another suggested that repeated loans to some countries (30 in Argentina, 20 to Cote d’Ivoire and Ghana between 1980 and 1999) indicated the ineffectiveness of structural adjustment conditions (Easterly, 2005). A review of loans to 100 countries in the period 1970–2005 concluded that enrollments in primary education are increased by higher per capita aid. Level of government spending *per se* does not relate to expanded enrollments. Nor do estimates of institutional quality (Dreher et al., 2008). A review of IMF and World Bank loans in all sectors concluded that aid does work, sometimes, but that it often fails to work. The reviewer noted that:

... For the IMF and ... World Bank, ownership is understood as the process whereby recipient countries come round to accepting ... [how they should change] the respective institution’s programmes, policies, and approaches to development and poverty reduction.

(Riddell, 2008, pp. 240–241).

Critics of structural adjustment (and the World Bank's interventions) focused on the Bank's failure to involve national and local actors in decisions about education. Several critiques noted the disjunction between those making decisions and those responsible for carrying them out. They complained that while the Bank acknowledges that conditions vary across countries and should be taken into account and calls for participatory planning, it then prescribes policies conceived in Washington that emphasize compliance with standards and the importance of surveillance. The Bank makes evidence-based decisions, but its data are its own or commissioned by it. The decisions made by Bank economists are different, in their assumptions and their data, than those of educators, especially educators in other cultures (Lauglo, 1996).

Although highly supportive of the IMF's support of developing country economies and a firm believer in the "colossal productive forces of capitalism", the then Deputy Secretary of Treasury of the United States recommended against transferring lessons learned in one country to another. Referring to the failures of structural adjustment, he noted

... These are failures of extrapolation ... if most now agree that macroeconomic reforms took precedence over microeconomic in the earlier steps of reform, and reducing the size of government took precedence over improving its quality, then it is fair to say that education and other basic social services were especially ill-served by these biases.

(Summers, 1998, p. 3).

In 1998 the World Bank published its analysis of 220 programs involving adjustment loans (Dollar and Svensson, 1998). The conclusion reached by the Bank was that in 75% of all cases success or failure could have been predicted using a few political economy variables. The critical factors controlled by the Bank, including conditions, had no relationship with the outcome. What mattered was whether the country, its leaders and people, were committed to the reform. To improve its "success with adjustment lending, the World Bank must become more selective and do a better job of understanding which environments are promising for reform and which are not". The most promising environments were those that demonstrated a strong commitment to reform (Dollar and Svensson, 1998).

A former Bank staff member later wrote that the Bank had tended to be captured by single methodologies and an insistence on use of Bank-directed research to justify policy decisions. Policy decisions often were made before doing on-site research. "Global" solutions were applied until they were proven to work only in some settings. The former staff member concluded that more attention should be paid to local priorities (Heyneman, 2003). Another analyst concluded that the IMF should not impose conditions with its loans. Not only was imposition of conditions not included in the IMF statement of mission, but also it was counterproductive (Dreher, 2009).

2000—development decades five and six: structural adjustment in new clothes?

During this period SAPs were replaced with new procedures. Critics observed similarities between the arguments supporting the old and the new. This section discusses several of the new forms of intervention by the World Bank and the IMF, accompanied by the OECD.

Partnerships

In 1996 the Organization for Economic Cooperation and Development (OECD) had proposed that aid should be limited to poverty reduction and human development objectives. Their call for a partnership of donors and countries was seconded by several countries and eventually led to the formulation of the UN's Millennium Development Goals (Kayizzi-Mugerwa, 1998). The short-term effect was a reduction in aid from OECD countries. The amounts of grants from donor countries declined from 5.5 billion dollars in 1990 to 1.8 billion in 1997 (Banya and Elu, 2001). This increased developing country dependence on the IMF and the Bank.

Anticipating the new millennium, the IMF and Bank began to describe themselves not as single actors but as partners with other international agencies committed to assisting the nations of the world. In 2002 the Bank became a member of the Global Partnership for Education (GPE), a non-governmental organization to carry out the United Nations' Education for All Fast Action strategy (<https://www.globalpartnership.org/>). By explicitly committing to the self-government (sovereignty) of states receiving aid, the Bank could, it was argued, increase its influence not by domination or imposition, but by encouraging participants to accept self-discipline (Abrahamsen, 2004). The Bank eventually provided about 75% of the funds distributed through the GPE. Technical assistance and grants were made to needy countries on the basis of recommendations by external consultants who did Education Sector Plan appraisals. The IMF also joined, asserting that its lending to a country encouraged donors to increase their support for social spending in that country (Clements et al., 2013).

The recommendations made to recipient countries were similar to those that appeared in structural adjustment programs. The construct of partnership was not well-defined, prompting several questions. The proposal called for contracts between (lending and receiving) partners: critics asked whether contractual agreements were not the same as conditions (Maxwell and Riddell, 1998). After all, countries might define their own priorities but they had to be acceptable to the lending partners. What would be the effect of unequal relationships of power among partners (Bailey and Dolan, 2011)? Other critics observed that the World Bank would control the dominant portion of all resources in the partnerships proposed (Menashy, 2018; Menashy and Shields, 2017).

The formation of the GPE had little effect on relationships with recipient countries. One commentator observed that the World Bank was operating with the same ideology expressed in [Rostow's \(1960\)](#) theory of stages of development. By classifying countries according to their stage of development, the Bank located countries on a common path of progress or modernization. In order to develop, they must follow in the footsteps of the more "advanced" capitalist countries ([Klees, 2002](#)).

Poverty reduction strategy papers

To highlight their focus on poverty reduction, the IMF and World Bank in 1999 initiated a process of preparing Policy Reduction Strategy Papers. The process was meant to engage representatives of the civil society in dialog about policy formulation, in order to "help poor countries and their development partners strengthen the impact of their common efforts on poverty reduction" ([World Bank, 2000](#)). Although this was seen as a radical shift of objectives, it maintained a relationship of imposition of the Bank's conditions ([Pender, 2001](#)). The process to achieve globalization by transforming informal into formal institutions would, it was argued, displace social norms and rules of conduct with legally binding rules ([Craig and Porter, 2002](#)). The IMF and the Bank claimed the poverty reduction approach was new, others saw it as nothing more than window dressing for the same old structural adjustment approach that had pushed so many countries further into poverty ([Joy et al., 2003](#)).

Between 2000 (the first Poverty Reduction Strategy Paper) and 2006 the specification of how to achieve development went through four different iterations. Initially, the prescribed actions conformed to the original Washington Consensus which called for deregulation and privatization. A second approach, a post-Washington Consensus, put more emphasis on building quality institutions and safety nets. The New York Consensus emphasized economic growth, aid and governance. The fourth approach, the Social Protection Agenda (SPA), aimed specifically at reduction of poverty and vulnerability. Emphasis on the SPA increased as a result of the 2008 crisis. The Washington Consensus was no longer the Bank's development agenda ([Elkins et al., 2018](#), p. 4).

Viewed almost twenty years later, the emphasis on poverty reduction appeared to have been successful. Using longitudinal data from 123 countries, researchers found a positive rate of economic growth, a significant reduction in poverty and in infant mortality, and positive progress toward realization of universal primary education and gender parity ([Elkins et al., 2018](#)).

The knowledge society

As another shift in emphasis, at the turn of the century in 2002 the World Bank announced its dedication to building "knowledge societies". This would be accomplished by providing increased support for tertiary education, which the Bank identified as a major factor in the production and dissemination of knowledge. Global partnerships would provide access to knowledge about "innovation systems" that would move countries along the path to development ([World Bank, 2002](#)).

The announcement was cheered by many, as a radical shift from earlier policies that had, by emphasizing primary education, wreaked drastic harm on higher education in developing countries ([Obamba, 2013](#)). Between 1960 and 2000 the Bank twice reversed its course. Initially the Bank had espoused human capital theory which assigned more value to higher education. When research showed that economic rates of return to primary schooling were higher than those to higher ([Psacharopoulos, 1973, 1980](#)), support for universities was reduced. Drastic consequences for hapless recipient countries followed from these changes ([Samoff and Carrol, 2004; Obamba, 2013](#)).

By 1998 new "knowledge" prompted another reversal, re-defining universities as critical actors in the economic development process. Following Bank recommendations, higher education spending sub-Saharan Africa was set at 22% of government expenditure on education, even though only 2% of the student population would benefit ([Banya and Elu, 2001](#)). The Bank called for construction of systems of learning, including lifelong learning, with universities as the major source of knowledge ([World Bank, 2002](#)). It announced that only by producing, adopting, adapting, disseminating and commercializing knowledge could a country compete economically and improve the general welfare ([Obamba, 2013](#), p. 97). The university would become the primary means by which Africa (and other regions) could take their place in the global economy. Meanwhile, Africa had not yet achieved universal primary education and less than half its adult population was literate ([Watkins, 2013](#)). Some viewed the premises and promises of higher education development initiatives with suspicion ([Molla and Cuthbert, 2018](#)).

The term "knowledge" was not defined in the Bank's report, nor did the document provide illustrative examples. A distinction was made between technical knowledge and knowledge about attributes, but issues of validity and reliability of knowledge were not discussed. Referring to itself as a "knowledge bank" the Bank's argument drew (principally) on its own research and analysis of "knowledge economies", to highlight how (institutional) science and technology contribute to development ([World Bank, 2002](#)).

By 2015 the Bank was one of the world's leading producers of education research. Most of the research cited to justify Bank proposals and conditions was produced by Bank employees or contractors, or cited other authors linked to the Bank ([Zapp, 2016](#)). The IMF's and Bank's "authority" in negotiations with loan-seeking countries also rested on their command of the technologies used to assemble and present financial and educational data. Both institutions were skilled in use of accounting procedures that encouraged loan recipients to exercise "self-discipline" in conforming to loan conditions ([Neu et al., 2006, 2010](#)).

The report stated that support for tertiary institutions was necessary to assist them to adapt to changes in internal and international market forces. Low rates of economic growth in recipient countries were attributed (in part) to their lack of knowledge.⁸ If a country (re)built its own knowledge-producing institutions, international trade regimes and foreign investments would generate employment (Mehta, 2001).

In this new perspective, universities were seen as key agents in the construction of the knowledge society, and hence drivers in the (global) knowledge economy. Success in the fulfillment of the mission of the university, at one time a disinterested search for truth, would be assessed in terms of its contribution to research and economic productivity, and scores on quality assurance measures. Both universities and the economy would benefit, it was argued, from closer links with industry and commerce (Olssen and Peters, 2005).

Results-based lending

The shift in orientation of the World Bank was accompanied by an increased emphasis on the importance of measurement of the outcomes of national education systems (Baker and LeTendre, 2005; Kamens and McNeely, 2010). Interest in comparing the learning outcomes of different countries may have been a product of the Cold War. The Soviet Union's success with Sputnik had focused the Western Alliance's attention on education and its relationship with innovations in technology. National assessment of education systems quickly followed (Purves, 1987; Tröhler, 2013).

Originally, the International Association for the Evaluation of Educational Achievement (IEA) was a collaborative project to identify common principles that explain the performance of the educational systems in economically advanced countries (Husén and Postlethwaite, 1996). The test they developed measured students' knowledge of the national curriculum.⁹ What started as a scholarly enterprise, however, soon took on a different aspect. National comparisons on academic achievement were associated with measures of economic performance. Politicians and citizens in lower-scoring countries sought to improve their economies by imitating practices in those wealthier countries scoring more highly in academic achievement.

Among the tests that followed was that produced as part of the Program for International Student Achievement (PISA). This test was developed by the Organization for Economic Cooperation and Development (OECD), which works closely with the IMF and the World Bank. Originally PISA was offered as a service to OECD member countries seeking to evaluate student achievement. Currently it is used, together with an instrument that assesses the effects of teaching practices (TALIS), to assess teacher and system accountability.

Unlike the tests of the IEA, however, PISA is not a measure of school-based knowledge. Instead, it purports to measure the life skills and knowledge of youth who are 15 and 16 years old. OECD assertions about the validity of the PISA and the TALIS instruments are based on the association between their scores and measures of national development, such as GNP. Neither of the tests meets conventional standards of either construct validity (does the test measure what it claims to measure?) or predictive validity (do scores correlate with some independent event or measure?) (Araujo et al., 2017; Berliner, 2020; Hanberger, 2014; Komatsu and Rappleye, 2017; Pepper, 2020; Zhao, 2020).

By associating PISA scores with levels of "development" and the global economy, OECD presented PISA as capable of valid measurement of the quality of an education system, arguing that by comparison of one country's students' "knowledge and real-world problem-solving skills" it is possible to determine the system's effectiveness in contributing to economic growth. Teaching should be assessed in terms of knowledge (test scores) rather than learning ability or creativity (Smith, 2014).

Critical discourse analysis of published documents demonstrates how OECD uses test-based "knowledge" to shape the education policies of participating countries.

... OECD (a) uses a discourse of fear to market teacher quality in light of global changes, implicitly framing [low-scoring] teachers as "bad teachers"; (b) advocates reliance on the organization as a protector and (c) promises a remedy by regulating teachers in the name of effectiveness and the knowledge economy.

(Berkovich and Benoliel, 2020).

PISA was applied initially in high income countries, but then came to be used by the World Bank to assess loan applications from poor countries. The "effect" of PISA on education in recipient countries has been described as one of "governing by numbers" (Grek, 2009). As more countries were invited to join OECD, PISA has come to be described as a "global governing complex" for education (Ydesen, 2019).

With the competition eliminated, by 2013 the OECD had constructed a system of "global educational governance". PISA was central to imposing an "audit culture" which served to decontextualize schooling, bringing participants into the world culture (Meyer and Benavot, 2013). Although not experimentally verified, OECD recommendations of effective practices based on use of PISA are given high credence.

⁸In part, this "lack" was the result of prior investment in higher education. University graduates in Africa migrated to employment in more advanced economies.

⁹That is, the IEA tests, like those that followed, measured a student's current knowledge, not necessarily learned in school.

Observers see a “Global Education Industry” working in all stages of policymaking, emphasizing the “best global practices”. In effect, the generation of policy options is today frequently outsourced to the private sector (Mohamed and Morris, 2021). In some instances, the policy agenda of OECD supplants national priorities of recipient countries (Auld et al., 2018).

It does seem that PISA has contributed to persuading countries to align their education systems with those of the more prosperous countries. Once a relatively small organization (less than 40 countries), OECD now has 100 members with more asking to join and willing to conform to OECD’s educational policy mandates. Structural adjustment, guided by the IMF, World Bank, OECD and others continues today (Addey and Gorur, 2020; Meyer and Benavot, 2013; Pons, 2017; Ydesen, 2019).

Summary: twenty-first century observations on structural adjustment in education

The concept of a “knowledge economy” appears to have displaced “world culture” as the justificatory framework for “recommendations” to recipient countries by the IMF, World Bank and OECD. Much more emphasis now is placed on research to justify selection among policy choices based on evidence or results. In most instances, however, the evidence provided is not from the country receiving the loan but rather from another in which a given policy has proved to be effective. To the extent that “recommendations” function as conditions, the framework is normative.

The IMF and World Bank, after 50 years of immersion in national economies, believe that its fund of economic knowledge should be applied uniformly to all. The OECD with 100 or more members whose economies account for perhaps 90% of the global GNP (but a smaller proportion of the world’s population), shares this confidence. The poor countries of the world stand outside the board room; inside others plan policies that will involve radical changes in the values and cultures of those left out. Given past experience, the policies proposed by the OECD, World Bank and IMF are likely to fail to reach their objectives.

World culture theory initially acknowledged differences across cultures, but the path taken by globalization is believed by some to have removed those differences as an obstacle. By the late 90s some came to believe that development had reached the peak which is “the end of history” (Fukuyama, 1992). Other claimed that differences between cultures and countries have disappeared to the point that global understanding flows easily across the globe; the world is now “flat” (Friedman, 2006).

Today, however, those pronouncements seem less credible. Our political and economic world appears not flat, but rather “spiky” (Feiock et al., 2008). Whether in consideration of geography, climate, economy, culture, population, language or religion, the regions, countries and communities of the world seek and find much on what to differ. Nations vary in the objectives knowledge, skills, beliefs and values of those who live in them, as well as in the policies that can be implemented, and their beliefs as to which will be effective. Over time they may have become similar in some ways, but at the same time they have developed unique and often contradictory features. The acquisition of knowledge has become a source of continuing change in beliefs, values, skills, actions.

The apparent convergence of national education systems assumes acceptance of uniform global policies, but the reality of countries is highly diverse. As a consequence, one reason for the failure of structural adjustment programs is they are not implemented as intended. In Mongolia, following the Soviet collapse, a structural adjustment loan included a condition of decentralization of education. Over a 10-year period the Ministry of Education shifted back and forth, decentralizing then recentralizing. After close up study researchers concluded that the cultural legacy of socialism collided with the culture of the international lenders (Steiner-Khamsi and Stolpe, 2004, 2006). Loan implementation would be more probable if conditions were determined after conducting an analysis of conditions and priorities in the recipient country, rather than relying on normative “best practices” (Steiner-Khamsi, 2016).

Some may attribute non-implementation to incompetence, but more likely failure results from political resistance. A study of SALs and SAPs in 90 countries between 1980 and 2005 concluded that the most common explanation for non-implementation was refusal by democratically elected officials to comply with imposed conditions. Resistance was explained either by their own conclusions about likely results, or because of political opposition (Henisz and Mansfield, 2019).

Lack of attention to the political economy of lending has also threatened the ability of the Global Partners to agree on strategies and specific policies. The effectiveness of the GPE depends in part on its ability to appear before recipient countries as a cohesive, consensual partnership of countries representing diverse positions. Internally, however, the GPE is plagued with disagreements among partners about key principles. The partners themselves hold different political philosophies and cultural values, with the consequence that the GPE itself is “a venue for serious contestation” (Knutsson and Lindberg, 2017, 2019, 2020).

From this perspective, the normative theory of “world culture” and a global economy is destined for failure in its quest to bring all countries to a comparable level of development. Taking the political into account will require a different understanding of what can be achieved through globalization. Rather than thinking of a global society, we should work toward a global system of societies in which the participation of unique participants generates synergy. Cooperating societies, each contributing unique knowledge derived from unique experiences, can develop education systems that not only expand knowledge, but also our awareness of its unlimited possibilities.

References

- Abrahamsen, R., 2004. The power of partnerships in global governance. *Third World Q.* 25 (8), 1453–1467. <https://doi.org/10.1080/0143659042000308465>.
 Addey, C., Gorur, R., 2020. Translating PISA, translating the world. *Comp. Educ.* 56 (4), 547–564. <https://doi.org/10.1080/03050068.2020.1771873>.
 Alarcón-González, D., McKinley, T., 1999. The adverse effects of structural adjustment on working women in Mexico. *Lat. Am. Perspect.* 26 (3), 103–117. <https://doi.org/10.1177/0094582X9902600306>.

- Araujo, L., Saltelli, A., Schnepf, S., 2017. Do PISA data justify PISA-based education policy? *Int. J. Comp. Educ. Dev.* 19 (1), 1–16. <https://doi.org/10.1108/IJCED-12-2016-0023>.
- Ascher, W., 2009. Robert McNamara's Second Vietnam. <https://www.tni.org/my/node/9961>.
- Auld, E., Rappleye, J., Morris, P., 2018. PISA for development: how the OECD and World Bank shaped education governance post-2015. *Comp. Educ.* 55 (2), 187–219. <https://doi.org/10.1080/03050068.2018.1538635>.
- Bailey, F., Dolan, A., 2011. The meaning of partnership in development: lessons in development education. *Pol. Pract.* 0 (13), 30–48.
- Baker, D.P., LeTendre, G.K., 2005. National Differences, Global Similarities: World Culture and the Future of Schooling. Stanford University Press.
- Banya, K., Elu, J., 2001. The World Bank and financing higher education in Sub-Saharan Africa. *High Educ.* 42, 1–34.
- Becker, G.S., 1962. Investment in human capital: a theoretical analysis. *J. Polit. Econ.* 70 (5, Part 2), 9–49. <https://doi.org/10.1086/258724>.
- Berkovich, I., Benoliel, P., 2020. Marketing teacher quality: critical discourse analysis of OECD documents on effective teaching and TALIS. *Crit. Stud. Educ.* 61 (4), 496–511. <https://doi.org/10.1080/17508487.2018.1521338>.
- Berliner, D.C., 2020. The implications of understanding that PISA is simply another standardized test. In: Fan, G., Popkewitz, T. (Eds.), *Handbook of Education Policy Studies*. Springer, pp. 239–258.
- Boli, J., Ramirez, F.O., Meyer, J.W., 1985. Explaining the origins and expansion of mass education. *Comp. Educ. Rev.* 29 (2), 145–170.
- Bonal, X., 2004. Is the World Bank education policy adequate for fighting poverty? Some evidence from Latin America. *Int. J. Educ. Dev.* 24 (6), 649–666. <https://doi.org/10.1016/j.ijedudev.2004.03.002>.
- Bordo, M., 2017. The operation and demise of the Bretton Woods system: 1958 to 1971. In: *Bretton Woods Agreements*. Yale University Press, pp. 217–235. <https://voxeu.org/article/operation-and-demise-bretton-woods-system>.
- Boughton, J.M., 2000. Silent Revolution: The International Monetary Fund, 1979–89. International Monetary Fund. <https://www.imf.org/external/pubs/ft/silent/>.
- Buchmann, C., 1996. The debt crisis, structural adjustment and women's education. *Int. J. Comp. Sociol.* 37 (1–2), 5–30. <https://doi.org/10.1163/002071596X00208>.
- Cable, V., 1995. The diminished nation-state: a study in the loss of economic power. *Daedalus* 124 (2), 25–53.
- Clements, B., Gupta, S., Nozaki, M., 2013. What happens to social spending in IMF-supported programmes? *Appl. Econ.* 45 (28), 4022–4033. <https://doi.org/10.1080/00036846.2012.744136>.
- Cowen, R., 2012. World yearbook of education 2012: policy borrowing and lending in education. In: Steiner-Khamsi, G., Waldow, F. (Eds.), *International Studies in Sociology of Education*, vol. 22. Routledge. <https://doi.org/10.1080/09620214.2012.737694>, 3.
- Craig, D., Porter, D., 2002. Poverty reduction strategy papers: a new convergence. *World Dev.* 31 (1), 53–69. www.elsevier.com/locate/worlddev.
- Cummings, W., 1999. The Institutions of education: compare, compare, compare! *Comp. Educ. Rev.* 43 (4), 413–437.
- Cummings, W., 2003. The Institutions of education. In: *A Comparative Study of Educational Development in the Six Core Nations*. Symposium Books, London, UK.
- Dollar, D., Svensson, J., 1998. What Explains the Success or Failure of Structural Adjustment Programs. Policy Research Working Paper 1938.
- Dreher, A., 2009. IMF conditionality: theory and evidence. *Publ. Choice* 141 (1/2), 233–267. <https://doi.org/10.1007/S11127-009-9486-Z>.
- Dreher, A., Nunnenkamp, P., Thiele, R., 2008. Does aid for education educate children? Evidence from panel data. *World Bank Econ. Rev.* 22 (2), 291–314. <https://doi.org/10.1093/wber/lnh003>.
- Driscoll, D.D., 1995. The IMF and the World Bank: how do they differ?. In: *The IMF and the World Bank: How Do They Differ?* <https://doi.org/10.5089/9781557754059.038.a001>.
- Easterly, W., 2005. What did structural adjustment adjust? The association of policies and growth with repeated IMF and World Bank adjustment loans. *J. Dev. Econ.* 76 (1), 1–22. <https://doi.org/10.1016/j.jdeveco.2003.11.005>.
- Elkins, M., Feeny, S., Prentice, D., 2018. Are poverty reduction strategy papers associated with reductions in poverty and improvements in well-being? *J. Dev. Strat.* 54 (2), 377–393.
- Espinoza, Ó., 2002. The Global and National Rhetoric of Educational Reform and the Practice of in (Equity) in the Chilean Higher Education System (1981–1998). School of Education, University of Pittsburgh. Doctoral dissertation.
- Espinoza, O., 2008. Creating (in) equalities in access to higher education in the context of structural adjustment and post-adjustment policies: the case of Chile. *High Educ.* 55 (3), 269–284. <https://doi.org/10.1007/s10734-007-9054-8>.
- FDIC, 1997. An Examination of the Banking Crises of the 1980s and Early 1990s. Federal Deposit Insurance Corporation.
- Feiock, R.C., Moon, M.J., Park, H.J., 2008. Is the world “flat” or “spiky”? Rethinking the governance implications of globalization for economic development. *Publ. Adm. Rev.* 68 (1), 24–35. https://doi.org/10.1111/j.1540-6210.2007.00832_2.x.
- Friedman, T.L., 2006. The world is flat: a brief history of the twenty-first century. *Int. J.* 61 (Issue 3). <https://doi.org/10.2307/40204208>.
- Fukuyama, F., 1992. The end of history and the last man. Penguin.
- Geo-Jaja, M.A., Mangum, G., 2001. Structural adjustment as an inadvertent enemy of human development in Africa. *J. Black Stud.* 32 (1), 30–49. <https://doi.org/10.1177/002193470103200102>.
- Grek, S., 2009. Governing by numbers: the PISA “effect” in Europe. *J. Educ. Pol.* 24 (1), 23–37. <https://doi.org/10.1080/02680930802412669>.
- Group of Twenty, 2017. Coordination Between the International Monetary Fund and Multilateral Development Banks on Policy-Based Lending: Update on the Implementation of the G20 Principles. Group of Twenty, Washington DC.
- Haddad, W.D., Carnoy, M., Rinaldi, R., Regel, O., 1990. Education and development: evidence for new priorities. In: *World Bank Discussion Papers* (No. 95; Development Discussion Papers, Vol. 95).
- Hanberger, A., 2014. What PISA intends to and can possibly achieve: a critical programme theory analysis. *Eur. Educ. Res. J.* 13 (2), 167–180. <https://doi.org/10.2304/eej.2014.13.2.167>.
- Harbison, F., Myers, C.A., 1964. Education, Manpower and Economic Growth: Strategies of Human Resource Development. McGraw-Hill.
- Harris, N., 1995. Structural adjustment on the Baltic. *Econ. Polit. Wkly.* 30 (48), 3062–3065.
- Haynes, J., 1991. Human rights and democracy in Ghana: the record of the rawlings' regime. *Afr. Aff.* 90 (360), 407–525.
- Hellman, J.A., 1997. Structural Adjustment in Mexico and the Dog that Didn't Bark. CERLAC Working Paper Series.
- Henisz, W.J., Mansfield, E.D., 2019. The political economy of financial reform: De Jure liberalization vs. De Facto implementation. *Int. Stud. Q.* 63 (3), 589–602. <https://doi.org/10.1093/isq/sqz035>.
- Heyneman, S.P., 2003. The history and problems in the making of education policy at the World Bank 1960–2000. *Int. J. Educ. Dev.* 23, 315–337. [https://doi.org/10.1016/S0738-0593\(02\)00053-6](https://doi.org/10.1016/S0738-0593(02)00053-6).
- Husén, T., Postlethwaite, T.N., 1996. A brief history of the international association for the advancement of educational assessment (IEA). *Assess. Educ.* 3 (2), 129–147.
- Ilon, L., 1994. Structural adjustment and education: adapting to a growing global market. *Int. J. Educ. Dev.* 14 (2), 95–108. [https://doi.org/10.1016/0738-0593\(94\)90015-9](https://doi.org/10.1016/0738-0593(94)90015-9).
- Inkeles, A., 1966. *Becoming Modern*. Harvard University Press.
- Joy, J., Malaluan, C., Guttal, S., 2003. Poverty reduction strategy papers: a poor package for poverty reduction. In: *Economic Policy* (Issue January 2002).
- Kafferstan, S., 2017. The Perfect Storm: Lasting Impacts of Structural Adjustment Programs and Pressures of Climate Change in Latin America and Ghana's Africa.
- Kahl, J.W., 1968. Convergence theories. In: *Encyclopedia.com*, p. 18. <https://www.encyclopedia.com/social-sciences/encyclopedias-almanacs-transcripts-and-maps/convergence-theories>.
- Kamens, D.H., McNeely, C.L., 2010. Globalization and the growth of international educational testing and national assessment. *Comp. Educ. Rev.* 54 (1), 5–25. <https://doi.org/10.1086/648471>.
- Kayizzi-Mugerwa, S., 1998. Africa and the donor community: from conditionality to partnership. *J. Int. Dev.* 10 (2), 219–228.
- Klees, S.J., 2002. World Bank education policy: new rhetoric, old ideology. *Int. J. Educ. Dev.* 22, 451–474. www.elsevier.com/locate/ijedudev.

- Knutsson, B., Lindberg, J., 2017. Studying "the political" in international aid to education: methodological considerations. *Comp. Educ. Rev.* 61 (4), 701–725. <https://doi.org/10.1086/693924>.
- Knutsson, B., Lindberg, J., 2019. On the absent ground of transnational partnerships in education: a post-foundational intervention. *Glob. Soc. Educ.* 17 (4), 432–444. <https://doi.org/10.1080/14767724.2019.1583090>.
- Knutsson, B., Lindberg, J., 2020. Depoliticisation and dissensus in the global partnership for education: rethinking the post-political condition. *J. Int. Relat. Dev.* 23 (2), 436–461. <https://doi.org/10.1057/s41268-018-0141-5>.
- Koehler, G., 2015. Seven decades of "development", and now what? *J. Int. Dev.* 27 (6), 733–751. <https://doi.org/10.1002/jid.3108>.
- Komatsu, H., Rappleye, J., 2017. A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth. *Comp. Educ.* 53 (2), 166–191. <https://doi.org/10.1080/03050068.2017.1300008>.
- Kraus, J., 1991. The struggle over structural adjustment in Ghana. *Afr. Today* 38 (4), 19–37.
- Larrañaga, O., 1992. Financiamiento universitario y equidad. Chile 1990. Serie Investigación I-45 Junio.
- Latapi Escobar, A., González de la Rocha, M., 1995. Crisis, restructuring and urban poverty in Mexico. *Env. Urban.* 7 (1), 57–76.
- Lauglo, J., 1996. Banking on education and the uses of research. A critique of: World Bank priorities and strategies for education. *Int. J. Educ. Dev.* 16 (3), 221–233.
- Maxwell, S., Riddell, R., 1998. Conditionality or contract: perspectives on partnership for development. *J. Int. Dev.* 10 (2), 257–268. [https://doi.org/10.1002/\(SICI\)1099-1328\(199803/04\)10:2<257::AID-JID527>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-1328(199803/04)10:2<257::AID-JID527>3.0.CO;2-S).
- McGinn, N., 1987. The creation of an education operating system: the bureaucratic organization of the French national system of education. In: Thomas, R.M., Kobayashi, V. (Eds.), *Educational Technology: Its Creation, Development and Cross-Cultural Transfer*. Pergamon Press, pp. 209–232.
- Mehra, L., 2001. Commentary: the World Bank and its emerging knowledge empire. *Hum. Organ.* 60 (2), 189–196. <https://doi.org/10.17730/humo.60.2.cl4dw5yv0gv7867c>.
- Menashy, F., 2018. Multi-stakeholder aid to education: power in the context of partnership. *Glob. Soc. Educ.* 16 (1), 13–26. <https://doi.org/10.1080/14767724.2017.1356702>.
- Menashy, F., Shields, P., 2017. Unequal partners? Networks, centrality, and aid to international education. *Comp. Educ.* 53 (4), 495–517. <https://doi.org/10.1080/03050068.2017.1323822>.
- Meyer, H.D., Benavot, A., 2013. In: Meyer, H.D., Benavot, A. (Eds.), *PISA, Power and Policy: The Emergence of Global Educational Governance*. Symposium Books.
- Meyer, J.W., Boli, J., Thomas, G.M., Ramirez, F.O., 1997. World society and the nation-state. *Am. J. Sociol.* 103 (1), 144–181. <https://doi.org/10.1086/231174>.
- Mohamed, M., Morris, P., 2021. Buying, selling and outsourcing educational reform: the Global Education Industry and 'policy borrowing' in the Gulf. *Compare* 51 (2), 181–201. <https://doi.org/10.1080/030507925.2019.1607255>.
- Molla, T., Cuthbert, D., 2018. Re-imagining Africa as a knowledge economy. *J. Asian Afr. Stud.* 53 (2), 250–267.
- Nations, U., 1971. *World Economic Survey, 1969–1970: The Developing Countries in the 1960s: The Problem of Appraising Progress*.
- Neu, D., Ocampo Gomez, E., Graham, C., Heincke, M., 2006. "Informing" technologies and the World Bank. *Account. Org. Soc.* 31, 635–662. <https://doi.org/10.1016/j.aos.2005.07.002>.
- Neu, D., Rahaman, A.S., Everett, J., Akindayomi, A., 2010. The sign value of accounting: IMF structural adjustment programs and African banking reform. *Crit. Perspect. Account.* 21 (5), 402–419. <https://doi.org/10.1016/j.cpa.2008.06.002>.
- Obamba, M.O., 2013. Uncommon knowledge: World Bank policy and the unmaking of the knowledge economy in Africa. *High Educ. Pol.* 26 (1), 83–108. <https://doi.org/10.1057/hep.2012.20>.
- Olssen, M., Peters, M.A., 2005. Neoliberalism, higher education and the knowledge economy. *J. Educ. Pol.* 20 (3), 313–345.
- Pender, J., 2001. From "structural adjustment" to "comprehensive development framework": conditionality transformed? *Third World Q.* 22 (3), 397–411. <https://doi.org/10.1080/01436590120061679>.
- Pepper, D., 2020. When assessment validation neglects any strand of validity evidence: an instructive example from PISA. *Educ. Meas.* 39 (4), 8–20.
- Pons, X., 2017. Fifteen years of research on PISA effects on education governance: a critical review. *Eur. J. Educ.* 52 (2), 131–144. <https://doi.org/10.1111/ejed.12213>.
- Psacharopoulos, G., 1973. Returns to education: an international comparison. In: *The Economics of Education*. Elsevier. <https://doi.org/10.1016/b978-0-12-815391-8.00005-7>.
- Psacharopoulos, G., 1980. Higher education in developing countries: a cost benefit analysis. In: *Staff Working Paper No.440*. World Bank, Washington, D.C.
- Purves, A.C., 1987. The evolution of the IEA: a memoir. *Comp. Educ. Rev.* 31 (1), 10–28. <https://doi.org/10.1086/446653>.
- Reich, R.B., 1992. *The Work of Nations: Preparing Ourselves for 21st Century Capitalism*. Vintage.
- Reimers, F., 1990. *Deuda Externa y Financiamiento de la Educación. Su Impacto en Latinoamérica*. UNESCO.
- Reimers, F., 1994. Education and structural adjustment in Latin America and Sub-Saharan Africa. *Int. J. Educ. Dev.* 14 (2), 119–129. [https://doi.org/10.1016/0738-0593\(94\)90017-5](https://doi.org/10.1016/0738-0593(94)90017-5).
- Reimers, F., 1997. Education and structural adjustment: unmet needs and missed opportunities. In: Lynch, J., Modgil, C., Modgil, S. (Eds.), *Education and Development. Tradition and Innovation. Volume One. Concepts, Approaches and Assumptions*. Casells, pp. 1–18.
- Riddell, R., 2008. Does foreign aid really work?. In: *Choice Reviews Online*, vol. 45. Oxford University Press. <https://doi.org/10.5860/choice.45-3312>, 6.
- Ritter, A.R.M., 1990. Development strategy and structural adjustment in Chile, 1973–1990. *Can. J. Lat. Am. Caribb. Stud.* 15 (30), 159–195.
- Rostow, W.W., 1960. *Tow, the Stages of Economic Growth: A Non-communist Manifesto*. Cambridge University Press.
- Samoff, J., Carrol, B., 2004. Partnership and dependence: external support higher Carrol Joel Samoff and Bidemi. *Afr. Stud. Rev.* 47 (1), 67–199.
- Schultz, T.W., 1961. Investment in human capital. *Am. Econ. Rev.* 1 (2), 1–17.
- Simutanyi, L., 1996. The politics of structural adjustment in Zambia. *Third World Q.* 17 (4), 825–839.
- Smith, W.C., 2014. The global transformation toward testing for accountability. *Educ. Pol. Anal. Arch.* 22 (116), 55. <https://doi.org/10.14507/epaa.v22.1571>.
- Steiner-Khamsi, G., 2004. *The Global Politics of Educational Borrowing and Lending*. Teachers' College Press.
- Steiner-Khamsi, G., 2012. Understanding policy borrowing and lending: building comparative policy studies. In: Steiner-Khamsi, G., Waldow, F. (Eds.), *World Yearbook of Education 2012*, pp. 23–38. <https://doi.org/10.4324/9780203137628-8>.
- Steiner-Khamsi, G., 2016. New directions in policy borrowing research. *Asia Pac. Educ. Rev.* 17 (3), 381–390. <https://doi.org/10.1007/s12564-016-9442-9>.
- Steiner-Khamsi, G., Stolpe, I., 2004. Decentralization and recentralization reform in Mongolia: tracing the swing of the pendulum. *Comp. Educ. Rev.* 40 (1), 29–53.
- Steiner-Khamsi, G., Stolpe, I., 2006. Educational import: local encounters with global forces in Mongolia. In: *Educational Import: Local Encounters with Global Forces in Mongolia*. Palgrave Macmillan. <https://doi.org/10.1057/9781403982971>.
- Summers, L.H., 1998. "Equity In A Global Economy", Remarks by Lawrence H. Summers, Deputy Secretary Of the Treasury, International Monetary Fund. Treasury News, Washington, D.C., June 8, 1998. <https://fraser.stlouisfed.org/title/6111/item/587214/toc/551097>.
- Toye, J., 1987. *Dilemmas of Development*. Basil Blackwell.
- Tröhler, D., 2013. The OECD and Cold War Culture: Thinking Historically About Pisa. In: *PISA, Power, and Policy. The Emergence of Global Educational Governance*, vol. 23, pp. 141–161. Symposium Books. <https://doi.org/10.1007/978-1-4039-8297-1>.
- Trow, M., 1973. *Problems in the Transition from Elite to Mass Education*. Carnegie Commission on Higher Education.
- Vavrus, F., 2005. Adjusting inequality: education and structural adjustment policies in Tanzania. *Harv. Educ. Rev.* 75 (2), 174–201. <https://doi.org/10.17763/haer.75.2.565v0213145413t5>.
- Vincent, R., 1981. *Human Rights and International Relations*. Cambridge University Press.
- Watkins, K., 2013. *Narrowing Africa's Education Deficit*. www.brookings.edu/foresight/africa.
- Williamson, J., 2004. The strange history of the Washington Consensus. *J. Post Keynes. Econ.* 27 (2), 195–206.
- World Bank, 1991. *World Development Report 1991: The Challenge of Development* (Washington, D.C.).

- World Bank, 1992. Adjustment Lending And Mobilization of Public and Private Resources for Growth.
- World Bank, 1998. Assessing aid: what works, what doesnt and why. Washington, DC.
- World Bank, 2000. World Development Report 2000/2001: Attacking Poverty. World Bank, Washington, DC.
- World Bank, 2002. Constructing Knowledge Societies: New Challenges for Tertiary Education. www.worldbank.org.
- Ydesen, C., 2019. In: Ydesen, C. (Ed.), The OECD's Historical Rise in Education: The Formation of a Global Governing Complex. Palgrave Macmillan.
- Zapp, M., 2016. The World Bank and education: governing (through) knowledge. *Int. J. Educ. Know.* 53, 1–11. <https://doi.org/10.1016/j.ijedudev.2016.11.007>.
- Zhao, Y., 2020. Two decades of havoc: a synthesis of criticism against PISA. *J. Educ. Change* 21 (2), 245–266. <https://doi.org/10.1007/s10833-019-09367-x>.

Erasmus(+) student mobility: individual and institutional motivations and effects

Manuel Souto-Otero^a, Luca Favero^b, Kristyna Basna^c, Martin Humburg^d, and Stephanie Oberheid^e, ^a School of Social Sciences, Cardiff University, Cardiff, United Kingdom; ^b University of Turin (ESOMAS Department), Torino, Italy; ^c Institute of Sociology of the Czech Academy of Sciences, Prague, Czech Republic; ^d European Investment Bank, Luxembourg, Luxembourg; and ^e ICF, Bruxelles, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	218
Literature review	219
Erasmus(+) students: motivations and effects	219
Motivations	219
Impact on students	220
Erasmus(+) and higher education institutions: motivations and effects	221
Motivations	221
Impact	222
Data: the Erasmus+ impact study survey	222
Why do higher education institutions take part in international student mobility? Strategies, motivations and effects	223
Contextualizing international student mobility motivations: ISM and HEIs' internationalization strategies	223
Why is it important? Institutional motivations for international student mobility	224
Institutional effects of Erasmus participation	224
Internationalization of the student population	224
Curriculum and learning outcomes	225
Organization and governance: capacity building	226
Conclusions	227
Acknowledgment	228
References	228

Introduction

The Erasmus program was established in 1987, after almost ten years of development of pilot student exchange schemes, underpinned by the view that “a strong educational dimension” was “absolutely vital to the construction of an open and democratic Europe” (Jones 2016). In its more than three decades of existence, Erasmus has become the largest international student mobility program in the world and is seen by Europeans as one of the most positive results achieved by the European Union (EU) (Kantar, 2018). As noted by de Wit et al. (2013:22) “Much more than in other parts of the world, student mobility has been promoted in Europe as an intrinsically positive and desirable development, and has become at many levels a policy goal in itself”.

The Erasmus program has experienced a number of changes since it was set up. The number of participants has expanded significantly from the original 11 participating countries and 3244 higher education students traveling abroad for their studies in 1987. This expansion has derived from various enlargements of the EU itself and from the progressive opening of the program to participants from the rest of the world (non-EU countries within Europe as well as countries from outside Europe). Changes have also entailed the expansion of the program budget and of the range of activities it supports, to also include traineeships abroad for higher education (HE) students, staff mobility, cooperation projects (focusing mainly on innovation, exchange of good practices, network creation) and support for policy development. Finally, Erasmus has been integrated within increasingly broader “umbrella programs” that encompass a wider range of activities in education, training, youth and sports, beyond higher education. The name and scope of these umbrella programs have changed approximately every 6–7 years since the setting up of the original Erasmus, as “program periods” for EU education programs matured: the Socrates program started in 1994, Socrates II in 2000, Lifelong Learning Program in 2007 and most recently, “Erasmus+”, which has been the label used since 2013 and will continue to be used at least until 2027.

Something that has remained constant since the launch of the original Erasmus is that international mobility continues to be the “flagship activity” in Erasmus+ (European Commission, 2021a:4). For the current programming period (2021–27), the Erasmus+ program will have a budget in excess of 26 billion Euros. 70% of this budget is allocated to supporting international mobility opportunities across the life cycle and types of education: schools, vocational education and training, higher education and adult education (European Commission, 2021a). The mobility of higher education students represents the bulk of all mobility supported by the Erasmus+ program; around half of the mobility during the 2013–20 programming period was undertaken by this group

(European Commission, 2021b). Millions of higher education students have by now participated in Erasmus(+)¹ international student mobility. Erasmus+ projects contracted in 2019 alone funded over 500,000 higher education student and staff mobilities in Europe and the rest of the world (European Commission, 2020a). The large majority of the mobilities funded take place from and to EU countries (European Commission, 2021c).

International student mobility can take many different forms. The program defines international student mobility as an opportunity for learners “to undertake learning and/or professional experience in another country” (European Commission, 2021c:14), which is normally referred to as the “host” country for the mobility -whereas the country where the student normally follows their program of study is referred to as their “home country”. Erasmus+ supports students in any field and cycle of HE (short cycle, bachelor, master and doctoral levels). The mobility abroad is most frequently undertaken at a higher education institution, where the student will typically study for a period between 3 and 12 months following an agreed course of studies. About 70% of all student mobilities for higher education students in the 2018 Erasmus+ call was “study mobility” (European Commission, 2020a:32). Students pay no additional fees at the host university, receive a grant from the EU (sometimes complemented by other national or regional funding schemes) which aims to cover the additional costs generated by international mobility, and the credits and grades obtained at the host university are recognized by the home institution as part of the students’ degree. The remainder mobilities funded by the program were traineeships, which can take place at various types of organizations, such as “an enterprise, a research institute, a laboratory, an organisation or any other relevant workplace abroad” (European Commission, 2021c:41).

In the next section we review the literature on the motivations of students and institutions to take part in the Erasmus(+) program -with a focus on study mobility- and the effects of international student mobility, while also discussing the linkages between Erasmus(+) higher education (HE) and internationalization in higher education more broadly. This review concludes that there is much more information on the motivations and program effects on students than at the institutional level, where information is very scarce and often covers only a small sample of institutions. In order to address this gap, we present results from a large survey of Higher Education Institutions (HEIs): the Erasmus+ HE impact study survey. Section three introduces these data and section four presents the survey results on: the role of international student mobility in HEIs international strategies, the motivations of HEIs to take part in international student mobility and the effects of Erasmus(+) participation at the institutional level, in a range of areas. Section five presents our conclusions.

Literature review

Erasmus(+) students: motivations and effects

Motivations

The literature investigating motivations for student mobility has identified a large set of factors that explain students’ motivations to spend part of their studies abroad, as done within Erasmus(+) HE. The large majority of studies refers to four main types of motivation: Academic learning, other skills development, employability enhancement and personal development/life experiences.

Academic learning

Academic learning has been identified as an important motivation to study abroad. Students often wish to gain academic learning experience in another country and to have access to types of learning and knowledge not offered at their home institution (Murphy-Lejeune, 2002). Some students also wish to attend world class institutions that may be perceived to provide high quality learning experiences during their study abroad (Findlay et al., 2010; Rodríguez González et al., 2011).

Other skills development

Other skills development is a further important motivator for students. This is particularly the case of language learning, which is frequently mentioned by a large share of mobile students across different surveys (Rodríguez González et al., 2011; Souto-Otero et al., 2019). Soft-skills development is also an important reason for studying abroad (Lesjak et al., 2015; Teichler, 2004). Students aim to improve soft skills such as adaptability, taking initiative or proactivity (Souto-Otero et al., 2019) as a result of mobility. Moreover, students also often want to increase their cross-culture awareness (Roy et al., 2019).

Employability

Employability improvement is an important motivation to study abroad. Students often see mobility experiences abroad and the knowledge and skills acquisition and personal development facilitated by it, to be associated with the improvement of career chances and employability (Souto-Otero et al., 2019). Students view mobility experiences as a way to improve their professional skills (Lesjak et al., 2015; Murphy-Lejeune, 2002) and their career prospects, by differentiating them from other applicants (Deakin, 2014). Some students expect mobility to be a first step toward specific career pathways, in particular an international career

¹In this chapter we sometimes use “Erasmus(+)” to refer to Erasmus+ and predecessor programs.

(Findlay et al., 2010): in the views of these students international mobility can increase their chances of employment in a multinational company or in a different country thanks to the language, professional and soft-skills developed during their stay abroad (Beadle et al., 2015).

Personal development and life experiences

Personal development and life experiences are central motivations for study abroad: living abroad and meeting new people are among the most frequent reasons for Erasmus+ students to go abroad (Brandenburg et al., 2014; Souto-Otero et al., 2019; Ulicna et al., 2017). This has been the case for decades: Maiworm and Teichler (2002) reported that almost nine in ten Erasmus mobile students were motivated by personal development. Other studies have highlighted the importance of personality maturation (Fom-bona et al., 2013), traveling and adventure/having a break from usual surroundings (Findlay et al., 2010; Rodríguez González et al., 2011), fun and excitement (Stronkhorst, 2005). Simply the wish of spending some time in a different country and experiencing life abroad is one of the strongest reasons for mobility (Keogh and Russel-Roberts, 2009; Lesjak et al., 2015). This motivation was mentioned by 70% of students in the Erasmus+ HE impact study (Souto-Otero et al., 2019). Finally, students are also motivated by the possibilities of expanding their social network (Souto-Otero et al., 2019), meeting new people and making new friends (Brandenburg et al., 2014; Ulicna et al., 2017).

Impact on students

The literature identifies four main types of impact of Erasmus(+) on participants, most of which are related to students' motivations to become mobile. These are: academic learning, other skills development, labor market outcomes and personality and identity. These impacts are likely to be correlated with each other -for example, increased employability is likely to be a result of progress made with regard to the other types of impact. It is also important to bear in mind that while most studies on individual effects are based on empirical analysis, subjective measures of impact (such as perceived improvement on competences, academic development or employability, reported changes in identity) are more often employed than more objective measures (such as salaries, probability of being employed or working abroad, measured changes in identity or behavior). In this sense, the literature has a tendency to explore experienced/perceived effects.

Academic learning

The literature has identified an impact of mobility on academic performance. Teichler (2012) looks at a self-reported measure of academic performance, and found that mobile graduates perceived to have made better academic progress abroad than what they would have made at home during the same time. Other studies use less subjective measures of performance, e.g., graduation marks, time of graduation. d'Hombres and Schnepf (2021), for example, find evidence that mobility leads to increases in the take-up of further studies in the case of Italian mobile bachelor degree holders, who are 8 percentage points more likely to take up further studies -but report no effect for mobile students from the UK. For some students, study abroad may also entail positive "vertical mobility", defined in this context as students moving to institutions or countries with higher academic reputation than the home institution or country.

Other skills development

Study abroad could be expected to have strong effects in terms of language learning. Erasmus alumni surveyed in Teichler and Janson (2007) were found to be three times more confident in foreign language proficiency than their non-mobile peers. Similarly, using a survey of Italian students, Sorrenti (2017) shows how the program improves self reported hard skills such as the command of foreign languages.

The 2014 Erasmus Impact Study (EIS) (Brandenburg et al., 2014) and Erasmus Impact Study+ (EIS+) (Souto-Otero et al., 2019) provide evidence that mobility helps participants build up their soft skills and develop their personality -see also below. In these studies, personality development is measured through the memo© psychometric tool. This tool records personality, attitudes, and behavioral traits that are correlated with career success and employability. Respondents in the EIS+ study were tracked over time and answers from the same individuals were compared before and after their mobility experience. Mobile graduates displayed higher values of memo© scores prior to mobility, which shows that they are a selective group compared to the general student population. However, and despite starting from higher scores, they still recorded remarkable gains in self-confidence compared to students who remained at home -who experienced close to no development over the same time period.

Other effects on personality development concern the ability to interact with foreign cultures and adapt to new situations. According to the EIS+, 9 in 10 mobile graduates perceived an improvement in terms of adaptability, interaction with people from foreign cultures, communication skills and intercultural competences. These findings are confirmed by the literature measuring self-reported competences before and after mobility (Onorati et al., 2017; Stronkhorst, 2005) or interrogating students about the importance of these competences in their jobs (Teichler and Janson, 2007).

Labor market outcomes

Assessing the effects of Erasmus(+) participation in employability and earnings is challenging due to endogeneity problems: participants are a selective group and it is difficult to assess whether differences between participants and non-participants are due to participation in the program or pre-existing differences (for example in terms of entrepreneurship, motivation, etc.) between mobile students and the wider student population. The majority of Erasmus(+) alumni surveyed believe that their time abroad helped

them secure their first jobs and their careers (Souto-Otero et al., 2019; Teichler and Janson, 2007), notably through the emphasis placed in recruitment on academic achievement and personality, and the increasing importance of foreign language skills and international experience. Moreover, they also report higher levels of happiness with their occupation and score higher on job quality measures such as job security and career prospects.

Several econometrics studies find that international mobility increases the probability of being employed (Di Pietro, 2015 for Italy; d'Hombres and Schnepf, 2021 for Italy and the UK; Iriondo, 2020 for Spain) or of working overseas (Parey and Waldinger, 2011 for Germany; Di Pietro, 2012 for Italy; Pinto, 2020 for Spain). The evidence on the impact of Erasmus(+) mobility on earnings is mixed -see Di Pietro (2021) and Netz and Cordua (2021) for recent reviews of the literature on the wage premium of international mobility more generally. Teichler and Janson (2007) report how Erasmus students appear to be skeptical about the impact of stay abroad periods on income. However, some studies that analyze graduates' salaries at different times after graduation point to positive wage premia for Erasmus(+) alumni. There is no consensus over the exact magnitude of such premium, ranging from 3% (Messer and Wolter, 2007 for Switzerland; Rodrigues, 2016 for EU countries) to about 7%–14% (Jahr and Teichler, 2002 for EU countries; Cammelli et al., 2008 for Italy; Iriondo, 2020 for Spain). A significant share of the wage premium is likely to stem from a better command of foreign language skills and the job opportunities this opens. Sorrenti (2017), for example, suggests that foreign language skills are rewarded in the labor market, with a positive wage premium of 6%.

Some studies investigate the role of "vertical mobility" on labor market outcomes. Iriondo (2020) analyzed the impact of vertical mobility on employment and salaries for Spanish students. His results show significant premia for Spanish Erasmus graduates going to Germany, France, Nordic countries and the UK while failing to find any effect for other host countries such as Italy and Portugal. The importance of vertical mobility is consistent with the results from the EIS+, which found that mobile graduates from low GDP countries, Eastern and Southern Europe enjoyed the largest benefits from mobility in terms of finding their first job. Respondents from Eastern and Southern Europe were more likely to consider their mobility periods beneficial or highly beneficial (74% for Southern Europe and 73% for Eastern Europe) than individuals from Western (67%) and Northern Europe (67%). Van Mol et al. find that, after controlling for selection into ISM, differences in career outcomes between mobile and non mobile students disappear in the Netherlands, which they explain with reference to the characteristics of the Dutch education system and labour market, where restricted possibilities for upward vertical mobility limit returns to ISM in the local labour market.

Personality and identity

The previous discussion on skills development notes a range of personality changes derived from Erasmus(+) mobility. In addition, the EIS+ study found that 95% of the Erasmus+ HE participants surveyed reported having learned how to better get along in different cultures and 93% felt that they had improved their ability to take cultural factors into consideration.

One particular aspect of interest in the Erasmus(+) related literature has been the effect of international mobility on the European identity of participants. Some studies reported little or no effect (see for example Sigalas, 2010), whereas others report a positive effect on the development of EU identity. Evidence in the EIS+ shows that mobile students feel more European than the average student before the Erasmus+ experience, but also that their attachment to the EU grows during their study abroad. Observed changes in European identity are larger among students traveling to non-neighboring countries and can be surmised to having been exposed to greater cultural differences.

Among mobile students, those who identified themselves as Europeans to a lower extent before the Erasmus(+) program experienced the largest development in their European identity. This is confirmed by regression evidence in the EIS+. This result is also in line with earlier findings in Kuhn (2012:999) who, using survey data, claims that the largest gains of mobility are to be expected among those who are less likely to be "the winners of European integration who are already likely to be convinced of its benefits and who are already prone to feeling European".

Erasmus(+) and higher education institutions: motivations and effects

While an increasing body of research aims to explore the motivations and effects of international student mobility in general and Erasmus(+) students in particular, the institutional level has received much less attention in the literature. In this section, we review extant literature, before we present additional findings on these topics, derived from the HE leadership and HE staff EIS+ surveys.

Motivations

The large majority of European higher education institutions participate in the Erasmus+ program. By 2019 over 5700 higher education institutions held the Erasmus University Charter (European Commission, 2020a), which contains the fundamental principles that they need to observe when participating in the program. The large majority of institutions that hold the Charter are active in sending or receiving students and staff (European Commission, 2015). Given widespread participation in the program, institutional participation seems to have become internalized and assumed as "the norm" by researchers and European HE institutions alike. As a result, the motivations for participation in the program is a little researched phenomenon.

There is a dearth of studies on institutional motivations to participate in Erasmus(+) and the importance of international student mobility for the internationalization of HEIs, but it is possible to look at motivations for internationalization to gain an insight into motivations for mobility. The meaning of internationalization in HE is contested, but a popular definition is provided by Knight (2008:6) as "the process of integrating an international, intercultural, or global dimension into the purpose, functions or delivery of higher education". While we still know little about how internationalization takes place in practice (Seeber et al., 2020) it

is clear that at its core is “border-crossing” (Teichler, 2017). Such border crossing could be from home to abroad or from abroad to home: as such, internationalization activities have been noted to have “at home” and “abroad” components (Knight, 2012). Given this, it is not surprising that work on the conceptualization and classification of internationalization (Knight, 2004; Horn et al., 2007) often highlights international student and staff mobility as one of its central components. In the case of international student mobility this entails the hosting of international students (at home dimension) or the sending of students (abroad dimension). In fact, international student mobility became the most prominent theme in internationalization debates in Europe in the 1980s and 1990s, and the absolute number of foreign or international mobile students the most frequently used indicator for internationalization (Teichler, 2017).

When exploring motivations for internationalization it is possible to discern a number of core themes (de Wit, 2001; Altbach and Knight, 2007; Tadaki and Tremewan, 2013):

- Socio-cultural and academic factors (e.g. extension of the academic horizon, profile/status, updating curricula with international content, knowledge acquisition -including foreign language learning- and quality enhancement)
- Economic factors (e.g. economic growth and competitiveness, employability, financial incentives for institutions and governments), and
- Political factors (e.g. foreign policy, security, national and regional identity)

However, some of the motivations outlined above do not apply easily to institutional motivations for mobility. For example, political motivations around foreign policy, security or national/regional identity may not go a long way in explaining institutional appetite for internationalization. By contrast, other motivations not captured in the above list may exist; for example, HEIs may see an intrinsic value in international network development. As such, we propose four sets of motivations for international student mobility, which we employ below in this chapter:

- Academic related (competence development, which students would be unlikely to acquire otherwise),
- Economic (increase the attractiveness of and demand for study programs²),
- Social (increase the diversity of the student population) and
- Profiling (increase in institutional visibility, network development, increase national international prestige/quality assessments/ranking or reputation).

These different motivations are sometimes considered as contrasting or in tension. As Castro et al. (2016:419) put it: “to polarise and simplify the issue two idealised discourses of internationalisation can be identified. On the one hand, there is the neo-liberal instrumental, economic agenda. On the other hand, there is the educational agenda”. But as we will see, these motivations can also co-exist in the rationalization of student mobility practices in HEIs.

Impact

The institutional impact of participation in the Erasmus(+) program identified in the literature (Vossensteyn et al., 2008; Souto-Otero et al., 2019) tends to refer to four main areas: teaching and learning (e.g. improving the quality in teaching and learning, improving the development and recognition of student learning outcomes, internationalizing teaching and learning - e.g. internationalization of the curriculum -, internationalizing the student population), research (e.g. internationalizing research, increasing cooperation with industry), organization and governance (e.g. enhancing institutional governance - establishment or development of internationalization strategies, etc. -, capacity building for information and support services for students -in particular in counseling and in support structures for international student mobility, such as setting up international offices -) and organizational profiling (e.g. increasing institutional visibility, network development, benchmarking, reputation management and learning from quality standards in other institutions).

In general, however, there is scarce research on the extent to which these different effects have been associated and continue to be associated with Erasmus(+) participation. Moreover, not all these types of effects could be expected to result from the international student mobility elements of the Erasmus(+) program, and in equal measure. Those elements more closely related to student mobility are in the areas of teaching and learning (internationalization of the student population; internationalization of the curriculum and development and recognition of learning outcomes achieved abroad) and organization and governance (internationalization strategies that recognize international student mobility, capacity building). These are the areas that we cover in our analysis.

Data: the Erasmus+ impact study survey

The data that we use in this chapter was gathered as part of the “Erasmus+ Higher Education Impact study” (Souto-Otero et al., 2019). The study collected data through surveys of higher education students, graduates, academic and non-academic staff and leadership between the Spring of 2017 and the summer of 2018, making use of a census database of Erasmus+ participants. In total, 76,893 valid responses to the survey were received -making it one of the largest Erasmus(+) surveys to date. In this chapter

²We recognize that at least some HEIs may also want to increase the attractiveness of their programs to be more selective and increase the quality of the student population, rather than increase their recruitment, however we expect that for many institutions economic viability will be the primary concern when thinking about increasing attractiveness.

we make use of the results from the survey of HE leadership and Erasmus(+) mobile staff. 708 responses were obtained in the survey of higher education institutions, which was directed to management in HEIs (Rector, Vice-Rector, International Relations Officer, other members of the management team of the institution) and which gathered information on the internationalization strategy of the institution, student and staff mobility, and their impact on the institution as well as the characteristics of the institution. We also make use of over 3000 responses to a EIS + survey of mobile teaching staff. The survey collected information on their demographic characteristics, motivations for international mobility and the impact of their mobility at the personal and institutional level.

Why do higher education institutions take part in international student mobility? Strategies, motivations and effects

Given the lack of studies on the institutional dimension of international student mobility, compared to the individual dimension, its ramifications are explored in this section making use of data collected through the EIS+ surveys of leadership and HE staff (Souto-Otero et al., 2019). We first contextualize the motivations for international student mobility, then explore those motivations, and conclude exploring its effects at the institutional level. The data used to explore the first two questions goes beyond the Erasmus+ program, to expand to international student mobility in general, although it should be noted that Erasmus+ is the main instrument for international student mobility for European institutions (d'Hombres and Schnepf, 2021) and the data was collected in the context of surveys on Erasmus+. The exploration of effects refers, specifically, to Erasmus(+) HE international student mobility.

Contextualizing international student mobility motivations: ISM and HEIs' internationalization strategies

Internationalization is high on the agenda of the leadership of HEIs, and student mobility is considered the most important aspect of internationalization. Indeed, almost 9 in 10 institutions responding to the EIS+ leadership survey (87%) reported to have an internationalization strategy. Of these, around 95% of reported "recognition of student mobility as part of study programmes" to be an important or very important element in their strategy -see Fig. 1, making it the most important aspect of international strategies, above internationalization of the curriculum, of research, staff mobility or internationalization of staff. Internationalization of research cooperation was especially important among HEIs in Eastern European countries. The internationalization of the student population, which is related to student mobility, also came high in the responses. This is consistent with suggestions made by previous research, which highlights the importance of international student mobility as part of internationalization strategies, but confirms this finding in a large sample of HE leadership and explicitly comparing its importance with other aspects of HEIs' internationalization strategies.

This finding is consistent with Seeber et al. (2020), who analyzed survey data on the international activities portfolio of 431 institutions in 33 European countries collected in 2013, and found that mobility opportunities had a privileged place in these portfolios: mobility was the most frequent internationalization activity of HEIs: 96% of HEIs offered outgoing mobility for students (study abroad, international internships, etc.) and 93% for staff.

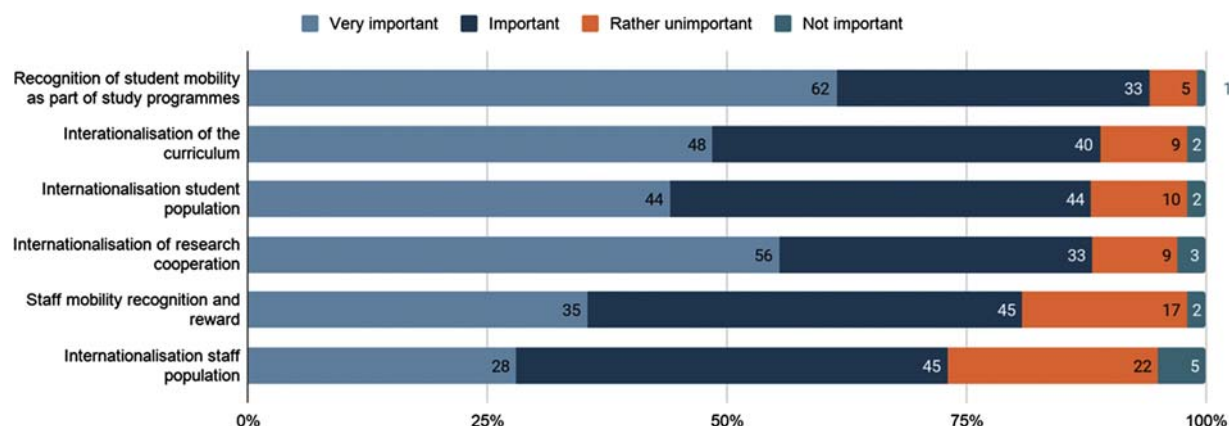


Fig. 1 Importance of selected aspects internationalisation strategy. Source: Erasmus+ impact study (survey of HE leadership) (Souto-Otero et al., 2019). N = 561. "What is the relative importance of the following aspects in the internationalization strategy of your institution?"

Why is it important? Institutional motivations for international student mobility

Why is international student mobility so important for HEIs? We approached this question by asking HE leadership about their institutions' objectives for international student mobility. In the **Motivations** section four types of motivations were identified: academic related, economic, social and profiling. Academic-related considerations were reported as the most important by HE leadership: international mobility provided institutions with the possibility to develop competences in their students that they would otherwise be unlikely to develop. This contrasts with a reportedly popular view of study abroad as tourism and a purely consumerist act. It also contrasts with a focus on international student mobility as a source of income for HEIs competing in global markets for students (Guruz, 2011) and more general views about the economic rationale of internationalization whereby: "Although still present in the rhetoric of international education, traditional values such as cooperation, peace and mutual understanding, human capital development, and solidarity, have been moved to the sidelines as universities strive for competition, revenue, and reputation/branding" (de Wit and Altbach's, 2021:35).

To be sure, economic considerations (program attractiveness) also featured highly, but at a long distance from competence development -see Fig. 2. Some aspects of profiling (network development) also featured highly in the response, but the importance of others (visibility, reputation/ranking) was much more modest, in spite of the purported association between excellence, rankings and mobility (see Souto-Otero and Enders, 2017). This may be explained by the emphasis of Erasmus+ on "horizontal" rather than "vertical" mobility (Van Mol et al., 2020), whereby the value of mobility is based on reciprocity and mutual exchange rather than being motivated by prestige hierarchies. There are, nevertheless, striking differences by region, with HEIs in Partner and Eastern European countries putting greater emphasis on prestige and visibility. In fact, 50% of the HEIs in Eastern European countries wish to increase their visibility abroad through participation in student mobility activities.

Social motivations around diversity featured low, which is surprising given that the diversity brought about by international student mobility is often considered as closely linked to internationalization (Castro et al., 2016:425) -although it should be noted that respondents were asked to choose the three most important objectives of students mobility for their institution only. The objective of increasing diversity of the student population was more frequently mentioned by HEIs in Western and Northern countries.

Institutional effects of Erasmus participation

In this section we report on the views of HE leadership on the effects of participation in the Erasmus(+) in three of the aspects noted in the **Impact** section, which were covered by the EIS+ survey: the internationalization of the student population, the curriculum and recognition of learning outcomes obtained abroad, and organization and governance -where we focus on institutional capacity building for mobility.

Internationalization of the student population

Erasmus(+), like any other mobility programs, aims to internationalize the student population. Fig. 3 shows that almost half of the HE leaders surveyed reported that the program had had a large impact in this respect at their institution. A further 33% reported a moderate impact. Erasmus+ is not the only tool that HEIs have to internationalize their student population. However, it provides a critical mass of mobile students from a wide range of countries (and arguably in a more balanced way than other tools such as full program mobility) and remains a key tool for the internationalization of the student population in European HEIs.

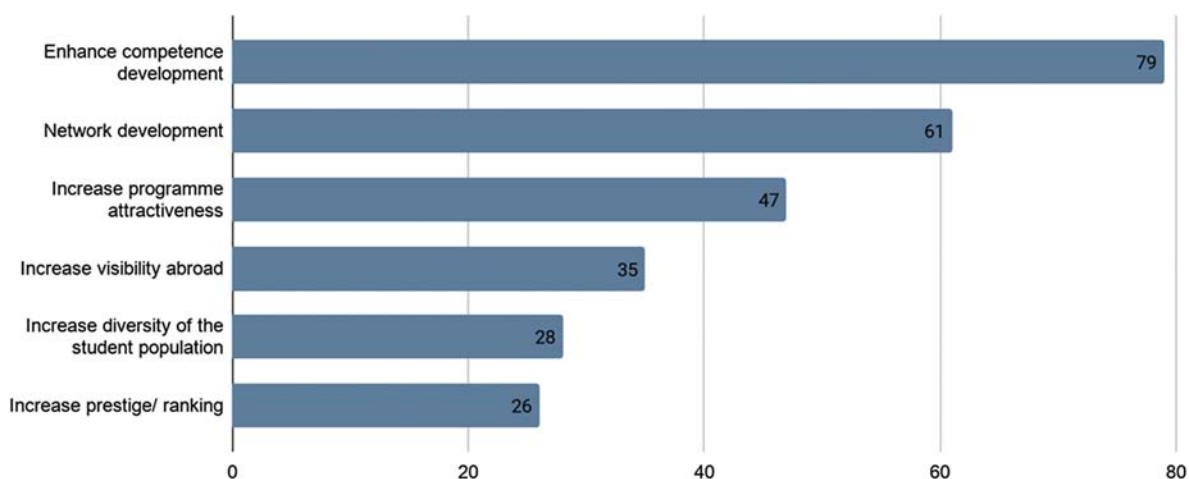


Fig. 2 Institutional objectives for international student mobility. Source: Erasmus+ impact study (survey of HE leadership) (Souto-Otero et al., 2019). N = 619. "What are the main objectives of student mobility for your institution?"

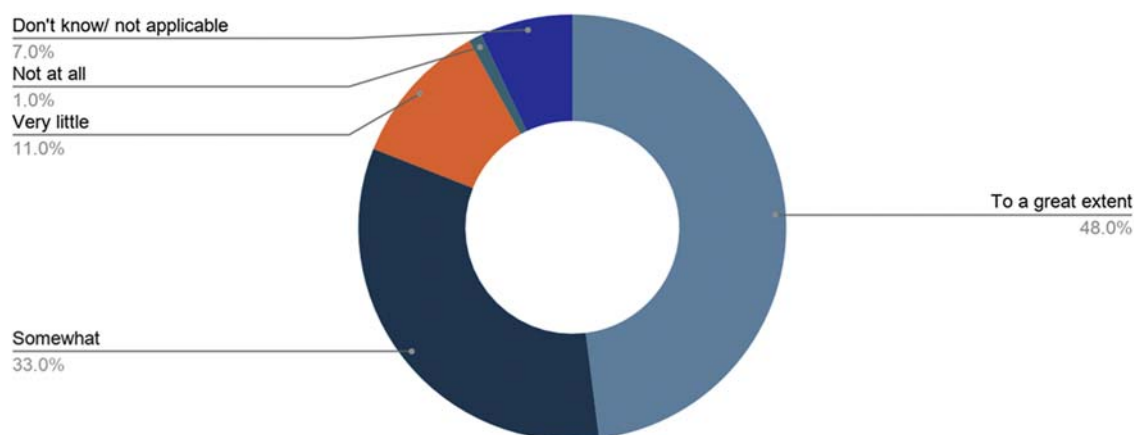


Fig. 3 Impact of Erasmus(+) on the internationalization to the student population over the last decade (2009–18). Source: Erasmus+ impact study (survey of HE leadership) (Souto-Otero et al., 2019). N = 450. “To what extent has participation in Erasmus+ or predecessor programs contributed to the internationalization of the student population in your institution in the past 10 years?”

Curriculum and learning outcomes

Erasmus(+) student mobility has various effects on the curriculum, and on the recognition of learning outcomes. The effects on the curriculum can come from Erasmus(+) cooperation projects too, but international student mobility may also stimulate the systematic integration of mobility into study programs. Fig. 4 shows the impact of participation in Erasmus(+) staff mobility on a range of curricula-related aspects in the home department of mobile teaching staff, as reported by staff. The results suggest that participation in Erasmus(+) has had widespread effects on the curricula, including adding more international perspectives into the curriculum or the modification of the curricula to integrate mobility into study programs, thus facilitating the recognition of learning outcomes acquired through international mobility experiences.

HE leadership were asked the extent to which participation in Erasmus(+) had also facilitated the recognition of mobility outside of those programs where mobility may be mandatory or is explicitly integrated into the curriculum -what we call embedded mobility. Over 40% of HE leaders reported that Erasmus(+) had helped the recognition of mobility outside of embedded mobility in their institution very much, and a further 42% that it had done so to a considerable extent -see Fig. 5. Whereas HEIs from Partner countries reported the highest level of impact in this area - more than 50% selected the highest impact category -, HEIs from Northern European countries reported much lower impact - only around 30% selected that category.

The next section provides information on the impact of Erasmus(+) on capacity building, including information on the impact on the recognition of learning occurred during mobility periods -regardless of whether this took place within or outside the context of embedded mobility.

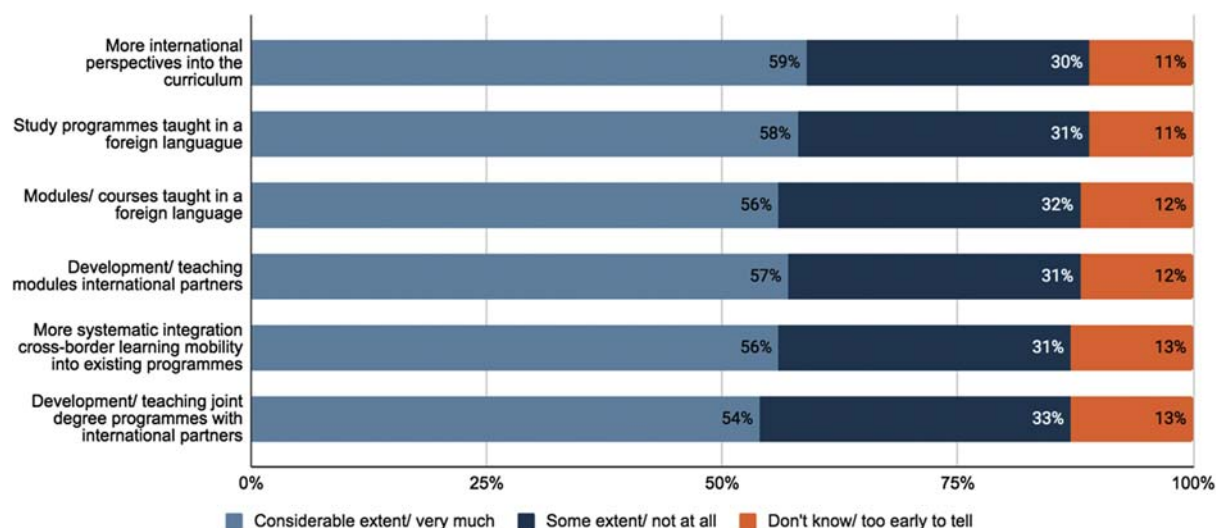


Fig. 4 Impact of Erasmus(+) staff mobility on curricula in the home department. Source: Erasmus+ impact study (survey of mobile staff) (Souto-Otero et al., 2019). N = 2512–3408 depending on the item. “Have the knowledge, competences or contacts acquired by you or other staff during mobility supported by Erasmus+ or predecessor programs led to improvements in the development and implementation of the following activities in your department?”

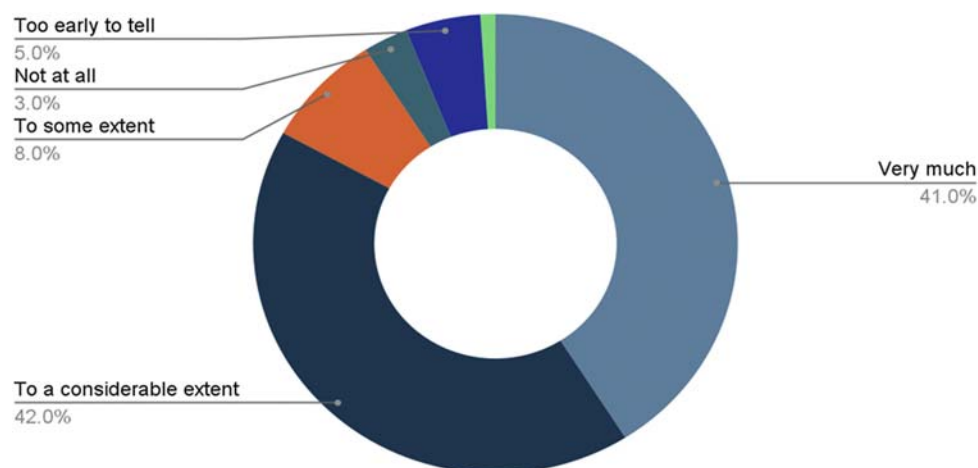


Fig. 5 Impact of Erasmus(+) in the recognition of mobility outside of embedded mobility. Source: Erasmus+ impact study (survey of HE leadership) (Souto-Otero et al., 2019). N = 459. "Has participation in the Erasmus+ and predecessor programs led to improvements in the recognition of mobility (outside of embedded mobility) at your institution?"

Organization and governance: capacity building

Lack of students' (and staff) preparation for mobility is often seen as a challenge (Castro et al., 2016), and there is, no doubt, room for development in these areas. Yet, it is widely acknowledged that Erasmus(+) has also had widespread impact in relation to organizational and capacity building aspects associated with student mobility. The EIS+ survey asked about seven capacity building aspects that support students in their mobility experiences. The impact of Erasmus(+) on these elements varied quite markedly, but was generally high to very high -see Fig. 6. The program impact was reported as particularly high in increasing information about available options for international mobility (in open days, information days, preparatory sessions, etc.) and in improving the management of student mobility more generally. Impact was lower in relation to the identification of additional sources of financing (showing an important degree of reliance in the program as a funding source for mobility) and the organization of language preparation pre-departure, although a still sizable share of institutions reported impact on these areas. The results also show that over 60% of respondents considered Erasmus(+) to have improved the recognition of learning occurred during mobility periods to a considerable or a large extent. HEIs from Eastern and Southern European countries reported higher impact of Erasmus(+) on student mobility and support systems.

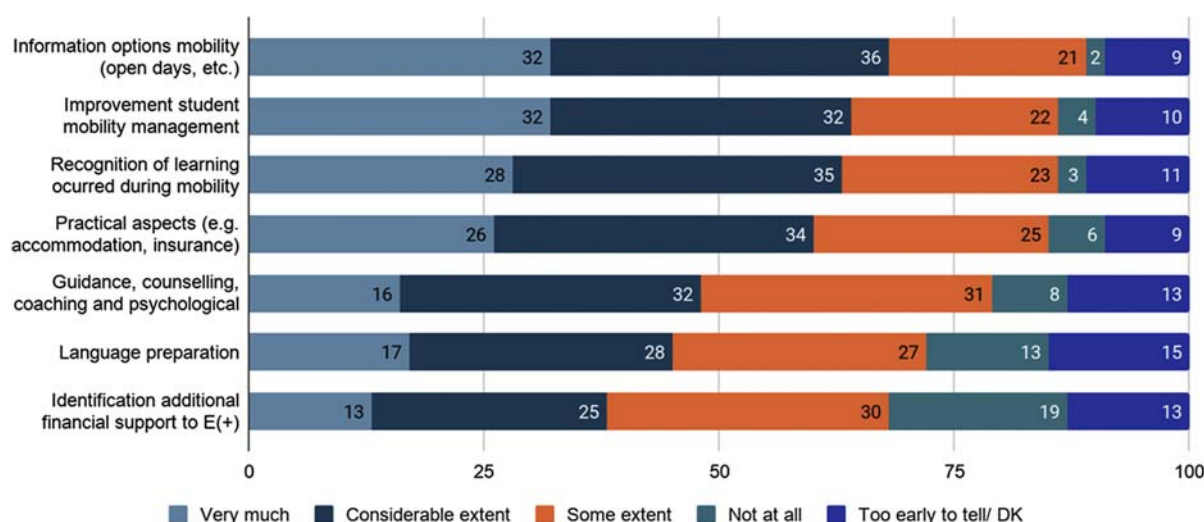


Fig. 6 Impact of Erasmus(+) on student mobility management and support. Source: Erasmus+ impact study (survey of HE leadership) (Souto-Otero et al., 2019). N = 495. "Have the knowledge, competences or contact acquired by staff during participation in Erasmus+ or predecessor programs led to improvements in the following aspects of student support available to students at your institution?"

Conclusions

This chapter has provided an overview of the motivations for and effects of individual and institutional participation in the Erasmus(+), the largest international student mobility program in the world. In this concluding section we offer some reflections on the past, present and future of international student mobility.

The review of the literature on motivations and effects of student mobility at the individual level showed an alignment between the expectations and effects of student mobility through Erasmus(+). The literature identifies four main factors that motivate students to take part in mobility: academic learning, other skills development, employability enhancement and personal development/life experiences. Among them, language learning and life experiences such as traveling and getting to know a new country appeared to be particularly important motivators. The literature has explored the effects of Erasmus(+) in, primarily, four areas, which are closely related to the motivations previously identified: academic learning, other skills development, labor market outcomes and personality and identity. Available evidence points toward the existence of Erasmus(+) effects across these areas, with effects being connected to each other and often mutually supportive, implying that improvement in one area is often connected to advancement in other areas as well.

The review of the literature has also highlighted that the majority of studies that examine the effects on students focus on single countries and use different methodologies, hindering the possibility of drawing international comparisons. Results from cross-country analysis, however, seem to point to heterogeneous effects with stronger impacts for Southern (d'Hombres and Schnepf, 2021) and Eastern European countries (Rodrigues, 2016). There is also some evidence that the country of origin and country of destination can matter, with students moving vertically to countries with more prestigious educational systems and stronger labor markets benefiting more from their mobility (Iriondo, 2020). In the absence of such vertical mobility benefits are more uncertain (Van Mol et al., 2020).

Information on motivations and effects at the institutional level is harder to come by. To address this gap, the chapter used data from two Erasmus(+) impact study surveys, covering over 700 HE leaders and over 3000 mobile teaching staff. The analysis revealed the central role that international student mobility has in the internationalization strategies of HEIs in Europe. The importance of such student mobility for institutions is primarily associated with its role in the improvement of the academic experience, and the development of competences that would otherwise be difficult to develop in students. Economic considerations are also relevant, as institutions are aware of the popularity of study abroad options among prospective students. Mobility also helps institutions to project specific profiles through network development and, somewhat less importantly for leadership, by increasing the visibility and reputation of their institutions. The social elements of international student mobility (increasing the diversity of the student population) seem to be less of a priority for institutional leadership in comparison to internationalization and economic motivations. This may be because HEIs may consider themselves to be already diverse institutions, because they have other ways to increase the diversity of their student population and because Erasmus(+) has traditionally not been particularly designed as a program that aimed to promote diversity in the student population, beyond national diversity. The inclusion theme within mobility itself (making mobility more inclusive along socio-demographic characteristics) has been in the Erasmus(+) agenda for some time (Souto-Otero et al., 2013; Souto-Otero, 2008; Bunesco et al., 2020) and it is being given greater attention in the new 2021–27 Erasmus+ programming period.

In terms of effects, the results show rather widespread and large effects of the program when it comes to the internationalization of the student population and the curriculum, the recognition of learning outcomes obtained abroad, and organization and governance. At this point, the impact on the internationalization of the student population and recognition of learning outcomes abroad seems somewhat higher than the impact on organization and governance. Organization and governance are the aspects where institutions may have at this point already established systems and may rely less on the program. Akin to the case of students, effects on institutions have been also found to vary by region (Souto-Otero et al., 2019).

Where next for mobility? After more than three decades of existence, the program continues to grow, and has ambitious participation targets for the future. Internationalization continues to have a prominent role in the mission of HEIs, which seems to have resisted contrary movements in recent years in politics and other areas of social life in several European countries and in the USA. Physical international student mobility within the program may change in the future: patterns of mobility will be affected by the UK exit from the EU (Brexit) as well as other political moves, for as Teichler (2017:210) noted “international mobility and cooperation is certainly one of the most “political” themes in higher education” -see also (França et al., 2018). de Wit and Altbach (2021) note that it is too early to tell the effects that the rise of nationalist-populist movements, bans on migration and anti-integration trends can have on HE internationalization and student mobility. International student mobility is part of a broader political and economic EU project, and has both pragmatic and ideological rationales. While these are in constant evolution (Dvir and Yemini, 2017; Shields, 2016; Sigalas, 2010), student mobility (like the development of university networks and partnerships) has a dual role as a tool to promote economic competitiveness and in forging support for the European Union project, and is likely to continue to be a central element of European education policy. But mobility is also being affected by other factors. These include public health issues such as the COVID-19 pandemic and concerns with HE sustainable practices and “the green university”, which are being scrutinized in more detail than in the past, although existing analyses have given international student mobility a small role in carbon emissions compared to other factors, such as staff mobility (ETH Zurich, 2020).

The types of student mobility supported by the program have evolved taking some of the above considerations in mind. In particular environmental concerns and “virtual mobility” have become more central. Since 2018 Erasmus+ Virtual Exchange enables participants in Europe and the Southern Mediterranean to “engage in meaningful intercultural experiences online as

part of their formal or non-formal education" (...) that "can form part of a higher education degree or a youth project" through "online-facilitated activities, Interactive Open Online Courses (IOOCs) and online debates" and has had a significant take-up (European Commission, 2020a:24 and 69–70). Blended mobility, combining physical and virtual mobility has also been on the rise (European Commission, 2020a:56). The Commission communication on Achieving the European Education Area promised an Erasmus+ program that "will be greener and more digital. Virtual and blended mobility could complement physical mobility" (European Commission, 2020b:18).

Some stakeholders, however, have raised concerns with such conceptualizations of international student mobility. The student-led Erasmus Student Network, for example, has been emphatic in its critical assessment of the cited communication and of the idea of virtual mobility, which it contends: "is the wrong approach", "is in our view not a form of mobility, even though the term has unfortunately entered into the public discourse" and "cannot be a substitute for a mobility period abroad (as...) an easy way to reach higher numbers of mobility participants" compromising funding for future physical mobilities (ESN, 2020:6–7). How these debates resolve the definition, shape and size of student mobility within the program, in turn, will affect the motivations of stakeholders to engage in it, and the effects Erasmus+ mobility will continue to produce on students and institutions in the future.

Acknowledgment

The EIS+ survey was financed by the DG EAC of the European Commission. The opinions and analysis presented in this chapter are those from the authors only. They do not purport to reflect the view of the European Commission, who financed the Erasmus+ impact study survey, or their organizations.

References

- Altbach, P.G., Knight, J., 2007. The internationalization of higher education: motivations and realities. *J. Stud. Int. Educ.* 11 (3–4), 290–305.
- Beadle, S., Humburg, M., Smith, R., Vale, P., 2015. Study on Foreign Language Proficiency and Employability. Final Report from ICF, Brussels.
- Brandenburg, U., Berghoff, S., Taboadela, O., 2014. The Erasmus Impact Study: Effects of Mobility on the Skills and Employability of Students and the Internationalisation of Higher Education Institutions. European Commission, Brussels.
- Bunescu, L., Davies, J., Gaebel, M., 2020. Erasmus+ International Credit Mobility -a Study of the Mobility of Disadvantaged Students from Partner Countries. SPHERE.
- Cammelli, A., Ghiselli, S., Mignoli, G.P., 2008. Study experience abroad: Italian graduate characteristics and employment outcomes. In: Byram, M., Dervin, F. (Eds.), *Students, Staff and Academic Mobility in Higher Education*. Cambridge Scholars Publishing, Newcastle, pp. 217–236.
- Castro, P., Woodin, J., Lundgren, U., Byram, M., 2016. Student mobility and internationalisation in higher education: perspectives from practitioners. *Lang. Intercult. Commun.* 16 (3), 418–436.
- de Wit, H., Altbach, P.G., 2021. Internationalization in higher education: global trends and recommendations for its future. *Pol. Rev. Higher Educ.* 5 (1), 28–46.
- de Wit, H., Ferencz, I., Rumbley, L.E., 2013. International student mobility: European and US perspectives. *Perspectives* 17 (1), 17–23.
- de Wit, H., 2001. Internationalisation of Higher Education in the United States of America and Europe. Greenwood Press, London.
- Deakin, H., 2014. The drivers to Erasmus work placement mobility for UK students. *Child Geogr.* 12 (1), 25–39.
- d'Hombres, B., Schnepf, S.V., 2021. International Mobility of Students in Italy and the UK: Does It Pay off and for Whom? forthcoming in *Higher Education*.
- Di Pietro, G., 2012. Does studying abroad cause international labor mobility? Evidence from Italy. *Econ. Lett.* 117 (3), 632–635.
- Di Pietro, G., 2015. Do study abroad programs enhance the employability of graduates? *Educ. Finan. Pol.* 10 (2), 223–243.
- Di Pietro, G., 2021. Studying abroad and earnings: a meta-analysis. *J. Econ. Surv. Online First*.
- Dvir, Y., Yemini, M., 2017. Mobility as a continuum: European commission mobility policies for schools and higher education. *J. Educ. Pol.* 32 (2), 198–210.
- ESN, 2020. The Future of the European Education Area. European Students Network. <https://esn.org/news/future-european-education-area-achieving-high-quality-education-future-young-europeans>. (Accessed 30 March 2021).
- ETH Zurich, 2020. Stay Grounded, Keep Connected. <https://ethz.ch/services/en/organisation/executive-board/vice-president-infrastructure/mobilitaetsplattform/air-travel.html>.
- European Commission, 2015. Erasmus: Facts, Figures and Trends. https://ec.europa.eu/assets/eac/education/library/statistics/erasmus-plus-facts-figures_en.pdf. (Accessed 20 March 2021).
- European Commission, 2020a. Erasmus+ Annual Report 2019. Publications Office of the European Union, Luxembourg.
- European Commission, 2020b. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Achieving the European Education Area by 2025. COM, Brussels (2020) 625 final. (Accessed 30 September 2020).
- European Commission, 2021a. Erasmus+ 2021–2027: Enriching Lives, Opening Minds. Publications Office of the European Union, Luxembourg.
- European Commission, 2021b. Erasmus+ Key Figures. https://ec.europa.eu/programmes/erasmus-plus/about/key-figures_en. (Accessed 19 March 2021).
- European Commission, 2021c. Erasmus+ Programme Guide. https://ec.europa.eu/programmes/erasmus-plus/programme-guide/part-b/key-action-1/mobility-higher-education-students-staff_en.
- Findlay, A.M., King, R., Geddes, A., Smith, F., Stam, A., Dunne, M., Skeldon, R., Ahrens, J., 2010. Motivations and Experiences of UK Students Studying Abroad. Project Report. Department for Business, Innovation and Skills, London.
- Fombona, J., Rodríguez, C., Pascual Sevillano, M.A., 2013. The motivational factor of Erasmus students at the university. *Int. Educ. Stud.* 6 (4), 1–9.
- França, T., Alves, E., Padilla, B., 2018. Portuguese policies fostering international student mobility: a colonial legacy or a new strategy? *Glob. Soc. Educ.* 16 (3), 325–338.
- Guruz, K., 2011. Higher Education and International Student Mobility in the Global Knowledge Economy: Revised and Updated Second Edition. State University of New York Press, New York.
- Horn, A.S., Hendel, D.D., Fry, G.W., 2007. Ranking the international dimension of top research universities in the United States. *J. Stud. Int. Educ.* 11 (3–4), 330–358.
- Iriondo, I., 2020. Evaluation of the impact of Erasmus study mobility on salaries and employment of recent graduates in Spain. *Stud. High Educ.* 45 (4), 925–943.
- Jahr, V., Teichler, U., 2002. Employment and work of former mobile students. In: Teichler, U. (Ed.), *ERASMUS in the SOCRATES Programme: Findings of an Evaluation Study*. Lemmens, Bonn, pp. 117–135.
- Jones, H.C., 2016. An Insider's History of How Erasmus Developed. <https://fedtrust.co.uk/an-insiders-history-of-how-erasmus-developed/>. (Accessed 25 March 2021).
- Kantar, 2018. Standard Eurobarometer 89. Spring 2018. Publications Office of the European Union, Luxembourg.
- Keogh, J., Russel-Roberts, E., 2009. Exchange programmes and student mobility: meeting student's expectations or an expensive holiday? *Nurse Educ. Today* 29 (1), 108–116.
- Knight, J., 2004. Internationalization remodeled: definition, approaches and rationales. *J. Stud. Int. Educ.* 8 (1), 5–31.

- Knight, J., 2008. The internationalization of higher education: are we on the right track? *Acad. Matters* 5–9 (Oct–Nov), 2008.
- Knight, J., 2012. Concepts, rationales, and interpretive frameworks in the internationalization of higher education. In: *The SAGE Handbook of International Higher Education*. SAGE, London, pp. 27–42.
- Kuhn, T., 2012. Why educational exchange programmes miss their mark: cross-border mobility, education and European identity. *J. Common. Mark. Stud.* 50 (6), 994–1010.
- Lesjak, M., Juvan, E., Ineson, E.M., Yap, M.H.T., Podovschnik Axelsson, E., 2015. Erasmus student motivation: why and where to go? *High Educ.* 70 (5), 845–865.
- Maiworm, F., Teichler, U., 2002. The students' experience. In: Teicher, U. (Ed.), *Erasmus in the Socrates Programme. Findings of an Evaluation Study*. Lemmens Verlag, pp. 83–115.
- Messer, D., Wolter, S.C., 2007. Are student exchange programs worth it? *High Educ.* 54, 647–663.
- Murphy-Lejeune, E., 2002. *Student Mobility and Narrative in Europe: The New Strangers*. Routledge, London.
- Netz, N., Cordua, F., 2021. Does studying abroad influence graduates' wages? A literature review. *J. Int. Stud.* 11 (4), 768–789.
- Onorati, M.G., d'Ovidio, F.D., Antonucci, L., 2017. Cultural displacement as a lever to global-ready student profiles: results from a longitudinal study on International Lifelong Learning Programs (LLP). *Qual. Quantity* 51 (2), 545–563.
- Parey, M., Waldinger, F., 2011. Studying abroad and the effect on international labour market mobility: evidence from the introduction of ERASMUS. *Econ. J.* 121 (551), 194–222.
- Pinto, F., 2020. The effect of university graduates' international mobility on labour outcomes in Spain. forthcoming in *Stud. High Educ.*. <https://doi.org/10.1080/03075079.2020.1725877>.
- Rodrigues, M., 2016. Does Student Mobility during Higher Education Pay? Evidence from 16 European Countries. Publications Office of the European Union, Luxembourg.
- Rodríguez González, C., Bustillo Mesanza, R., Mariel, P., 2011. The determinants of international student mobility flows: an empirical study on the Erasmus programme. *High Educ.* 62 (4), 413–430.
- Roy, A., Newman, A., Ellenberger, T., Pyman, A., 2019. Outcomes of international student mobility programs: a systematic review and agenda for future research. *Stud. High Educ.* 44 (9), 1630–1644.
- Seeber, M., Meoli, M., Cattaneo, M., 2020. How do European higher education institutions internationalize? *Stud. High Educ.* 45 (1), 145–162.
- Shields, R., 2016. Reconsidering regionalisation in global higher education: student mobility spaces of the European Higher Education Area. *Compare* 46 (1), 5–23.
- Sigalas, E., 2010. Cross-border mobility and European identity: the effectiveness of intergroup contact during the ERASMUS year abroad. *Eur. Union Polit.* 11 (2), 241–265.
- Sorrenti, G., 2017. The Spanish or the German apartment? Study abroad and the acquisition of permanent skills. *Econ. Educ. Rev.* 60, 142–158.
- Souto-Otero, M., Enders, J., 2017. International students' and employers' use of rankings: a cross-national analysis. *Stud. High Educ.* 42 (4), 783–810.
- Souto-Otero, M., Huisman, J., Beerkens, M., De Wit, H., Vujić, S., 2013. Barriers to international student mobility: evidence from the Erasmus program. *Educ. Res.* 42 (2), 70–77.
- Souto-Otero, M., Gehlke, A., Basna, K., Doka, A., Endrodi, G., Favero, L., Humburg, M., Jantos, M., Key, O., Oberheidt, S., Stiburek, S., 2019. Erasmus+ Higher Education Impact Study. Final report to the European Commission. European Commission, Brussels.
- Souto-Otero, M., 2008. The socio-economic background of Erasmus students: a trend towards wider inclusion? *Int. Rev. Educ.* 54 (2), 135–154.
- Stronkhorst, R., 2005. Learning outcomes of international mobility at two dutch institutions of higher education. *J. Stud. Int. Educ.* 9 (4), 292–315 (d).
- Tadaki, M., Tremewan, C., 2013. Reimagining internationalization in higher education: international consortia as a transformative space. *Stud. High Educ.* 38 (3), 367–387.
- Teichler, U., Janson, K., 2007. The professional value of temporary study in another European country: employment and work of former ERASMUS students. *J. Stud. Int. Educ.* 11 (3–4), 486–495.
- Teichler, U., 2004. Temporary study abroad: the life of ERASMUS students. *Eur. J. Educ.* 39 (4), 395–408.
- Teichler, U., 2012. International student mobility and the Bologna process. *Res. Comp. Int. Educ.* 7 (1), 34–49.
- Teichler, U., 2017. Internationalisation trends in higher education and the changing role of international student mobility. *J. Int. Mobil.* (1), 177–216.
- Ulicna, D., Oberheidt, S., Mathis, J., Milio, S., Basna, K., Endrodi, G., 2017. Combined Evaluation of Erasmus+ and Predecessor Programmes: Main Evaluation Report. Publications Office of the European Union, Luxembourg.
- Van Mol, C., Caarls, K., Souto-Otero, M., 2020. International Student Mobility and Labour Market Outcomes: An Investigation of the Role of Level of Study, Type of Mobility, and International Prestige Hierarchies. *Higher Education*, pp. 1–27.
- Vossensteyn, J.J., Beerkens-Soo, M., Cremonini, L., Antonowitsch, D., Epping, E., Lanzendorf, U., Teichler, U., Buerger, S., Tubail, A., Souto-Otero, M., McCoshan, A., Vega, S., Allen, K., Fernandez, J., Solorzano, B., Torrecillas, C., 2008. The Impact of ERASMUS on European Higher Education: Quality, Openness and Internationalisation. Final Case Studies Report from a Consortium of CHEPS, INCHER-Kassel and ECOTEC.

Academic mobility: global trends and future prospects

Robin Shields^a and Jack T. Lee^b, ^a School of Education, University of Bristol, Bristol, United Kingdom; and ^b Moray House School of Education and Sport, University of Edinburgh, Edinburgh, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	230
International student mobility	230
Faculty mobility	232
Academic mobility: current issues and future challenges	233
Barriers in policy and practice	233
Brain drain and brain circulation	233
Geopolitical currents	234
Climate change	235
Futures of mobility	235
References	236

Introduction

Mobility of individuals is a key driver of globalization and a hallmark of international education. Globally, the mobility of people across international borders has risen rapidly since the mid-twentieth century. For example, the UN World Tourism Organization recorded 2.28 billion crossings in 2018, which means that 29.4% of the global population crossed a border that year, a significant increase from the 1.03 billion when data was first collected in 1995 (18.0% percent of the global population that year, [World Bank, 2021a](#)). With this rise, mobility has become increasingly complex; long-term migration (including forced migration and refugees) is now an integral aspect of many societies around the world and an underpinning feature of global labor markets.

The rise of international mobility is particularly apparent in higher education, where the movement of faculty, researchers and students across international borders has reshaped universities and systems of higher education. However, universities' commitment to internationalism has a long history and considerable geographic scope, with records of significant mobility of faculty and students in medieval universities in Europe and East Asia ([Kim, 2009](#); [Rizva and Teichler, 2007](#)). Recent changes in higher education have accelerated mobility considerably: neoliberal policies in several countries have created immense pressures for universities to recruit international students in order to compensate for the loss of public funding, while shifts in the academic labor market and increasing international collaboration have spurred faculty mobility ([Shin et al., 1999](#); [Royal Society, 2011](#)).

This chapter discusses the drivers and impacts of academic mobility in international higher education. It considers mobility broadly as an integral part of international education and contemporary higher education, including mobility of students, faculty, and researchers. We argue that although mobility has been a longstanding and rapidly growing dimension of international higher education, a re-thinking of both policies and practices is paramount given the increasing inequalities, changing geopolitics, and growing environmental impact of academic mobility. We begin by discussing international student and faculty mobility, including current trends and drivers. We then turn to recent developments and their potential impacts on academic mobility, including shifting geopolitical powers and relationships, the rise of populism and associated racist and xenophobic movements, and the existential challenges of the climate emergency. We conclude by reflecting on the benefits of physical mobility in a world that is increasingly driven by virtual interactions.

International student mobility

The mobility of students across international borders has a long history. [Rizva and Teichler \(2007\)](#) estimate that 10% of students in medieval universities were international, while [Kim \(2009\)](#) recounts international mobility between Korea and China during the Tang Dynasty (618–907). Academic mobility also fueled European colonial projects, with faculty in Imperial powers undertaking a significant share of mobility to colonized societies as part of a purported “civilizing” mission ([Jöns, 2016](#); [Pietsch, 2010](#)). Following World War II, the geopolitics of the Cold War also provided a strong impetus for mobility, as the First (Western Democratic) and Second (Soviet Communist) Worlds vied for global influence. For example, scholarships to students from non-aligned states (“Third World”) was a powerful form of cultural diplomacy that asserted the intellectual strengths of host countries, which also gained influence when their alumni attained elite positions around the world ([Altbach, 2004](#)).

However, international student mobility has reached unprecedented levels of growth in the 21st century. According to the UNESCO Institute for Statistics, there were 2.1 million students who went abroad to complete a full degree at the turn of the century, and by 2018 this number had grown to 5.7 million ([Fig. 1](#)). Literature commonly differentiates between two common types of international mobility ([de Wit, 2008](#)):

Degree Mobility: Refers to students who cross an international border to undertake a full-degree. They are enrolled in a university outside their country of ordinary residence.

Credit Mobility: Refers to students who cross an international border to undertake credits as part of a degree program, but are not enrolled in a university abroad. This mobility is often informally referred to as “study abroad.”

In addition to these common types of mobility as identified in the literature, other significant modes of international student mobility exist. For example, students may complete an international trip as part of a course in which they are enrolled in their home country (e.g., a field visit or international service learning) or may undertake short-term, non-credit bearing mobility for educational purposes during holiday periods in their studies. Taken together, these various modes of mobility point to the vast scope and range of international student mobility in recent years. For example, in addition to the 5.7 million degree mobile students identified by the [UNESCO Institute for Statistics \(2021\)](#), 505,000 students took part in credit-mobility through the European Commission’s Erasmus+ program (2020), while the [Institute of International Education \(2020\)](#) reported 347,099 students from the US participated in some form of study abroad for credit (including short-term programs).

Much literature understands this explosive growth in the context of the global knowledge economy, referring to the growing demand for highly-skilled labor that often requires advanced levels of education ([Gürüz, 2011](#); [Kaushan and Lanati, 2019](#)). In this framework, international education becomes an investment in human capital (i.e., economically productive skills), which yields a return in the form of higher salaries paid because of the demand for these skills in the labor market ([Kratz and Netz, 2018](#)).

However, many others take a much more expanded view of the drivers and benefits of international study, pointing out that it can be part of longer term migration strategies and access to job markets ([Kahanec and Králiková, 2011](#)), development of intercultural competencies often broadly described as global citizenship ([Tarrant, 2010](#)), or the acquisition of a set of broader cosmopolitan identities and dispositions ([Tran, 2016](#); [Rizvi, 2011](#)). Literature also contains much evidence on the benefits from participation in mobility, including increased employability ([Crossman and Clarke, 2010](#); [Mohajeri Norris and Gillespie, 2009](#)) and confidence and language abilities ([Roy et al., 2019](#)), and engagement in global issues ([Paige et al., 2009](#)).

However, a smaller but growing body of literature takes a more critical approach to the topic, highlighting the role of international mobility options in social class reproduction, through which students access elite institutions to acquire advantages in the labor market that would otherwise not be available to them ([Findlay et al., 2012](#); [Shaw and Thomas, 2006](#)). This critique is supported by the fact that financial barriers and home caring responsibilities are significant barriers to participation in mobility ([Souto-Otero et al., 2013](#)). Other forms of social discrimination also appear to prevent international student mobility. Literature shows that racialized minority groups are less likely to participate in international student mobility ([Stroud, 2010](#)), and students with LGBTQ+ identities may experience hardships due to intolerance in host countries ([Michl et al., 2019](#)). While women tend to be over-represented in study abroad participation in many Western countries ([Böttcher, 2016](#)), this trend is reversed among students from the Middle East and East Asia, which [Findlay \(2011\)](#) attributes to gender discrimination in countries of origin. Thus, in terms of individual participation, international student mobility can reproduce or increase existing inequalities rather than reduce them.

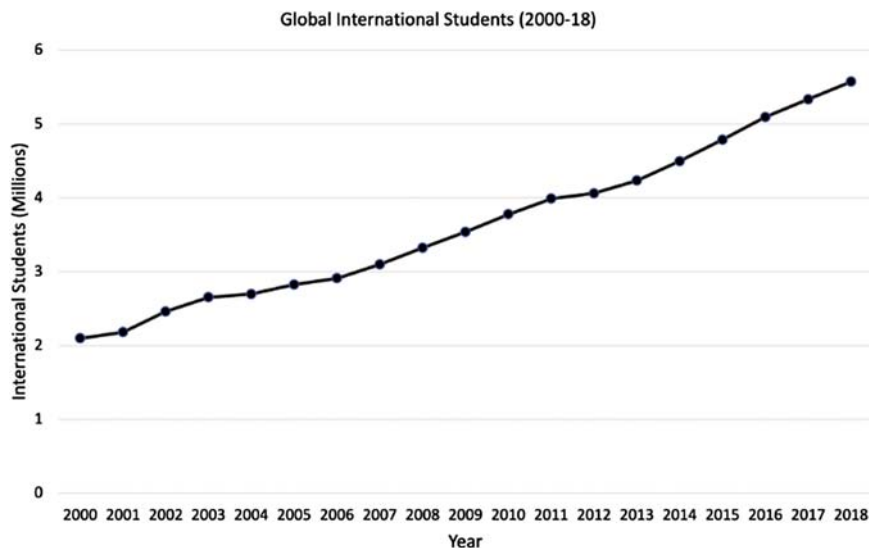


Fig. 1 Global numbers of degree mobile students between 2000 and 2018 ([UNESCO Institute of Statistics, 2021](#)).

On the global level, relationships of inequality and groupings are also pertinent to international student mobility. For example, Shields (2013) uses social network analysis to examine global flows of international students and finds that these flows are becoming increasingly polarized and unequal, and that they closely follow other relationships in the global political economy. In particular, English-speaking host countries account for a disproportionately large share of incoming students (Shields and Edwards, 2010). However, other changes may counteract this trend; a growing share of mobility also occurs at the regional level (i.e., between groups of countries in the same global region), supported by initiatives to facilitate mobility through the recognition of credits and qualifications (Chan, 2012; Shields, 2016). For example, mobility within Asia (intra-regional) is rising rapidly as the region's higher education systems and economies develop rapidly (Collins and Ho, 2018; Le Ha and Fry, 2021).

The COVID-19 pandemic disrupted travel on a global scale with serious effects on international student mobility. As national governments implemented new border protections, many students had difficulty returning to their home countries when universities closed halls of residence and moved teaching online in 2020 (Dickerson, 2020). Demonstrating the importance of international student fee income to universities' financial models, several universities commissioned charter flights to relocate students to campuses in 2021, as commercial flights were drastically cut (Tang, 2020). Even so, quarantines and lockdowns meant that international students spent little time in classrooms and were largely confined to halls of residence (Mueller, 2020). It remains to be seen whether this disruption has a long-term influence on either the numbers of international students or the patterns of mobility.

Faculty mobility

While student mobility dominates both the policies and literature of international higher education, the transnational flow of academics reveals many critical junctures in contemporary higher education systems worldwide. The peripatetic life of a scholar is not a new phenomenon. However, globalization has fundamentally changed the nature and scale of mobility among academics. Today faculty mobility ranges from brief sabbaticals and guest lectures to full-time employment and institutional leadership. For example, according to data from the Institute of International Education (2021), the number of mobile scholars migrating to the US rose from 106,000 in 2008 to 136,000 in 2018. The United States and United Kingdom have reaped tremendous benefits over many generations from scholars flocking to their universities (Fleming and Baily, 1969; Coser, 1984). This migratory pattern was particularly acute following World War II as scientists fled war-torn Europe. Even if mobility did not lead to outright immigration, foreign scholars have contributed immensely to the growth of American and British higher education through research.

Although the migration from the Global South to the Global North continues to persist and exacerbate epistemic inequality, significant changes in higher education systems worldwide are challenging this historical pattern. Rapidly developing higher education systems and prominent institutions in Asia and the Middle East are increasingly attracting Western academics and members of its diaspora. For example, the King Abdullah University of Science and Technology in Saudi Arabia has recruited several leading American and European scientists as well as proven university presidents from Asia to provide management expertise. Likewise, Chinese and Singaporean universities actively recruit scientists through the provision of well-equipped laboratories, generous research funding and competitive salaries (Lee, 2014; Han, 2021). Some of these scientists continue to hold appointments and lead research teams at their home institutions in the Global North while commuting regularly to Asia and the Middle East. Among high-profile academics, institutional affiliations among different countries is increasingly the norm. Beyond the institutional level, numerous countries have also launched prominent policy initiatives to attract global talent to their higher education systems. These schemes demand mobility even if the commitment is a short visiting professorship. While some schemes are genuinely open calls for talent, other schemes explicitly target the diaspora. For example, both China and Colombia have a long history in courting their diasporas to return since the 1990s. China's recent Thousand Talents¹ program contains a strong emphasis on the repatriation of Chinese scientists. Migration for Development in Africa, a scheme that serves 11 countries, promotes the repatriation of talented Africans. Brazil's well funded São Paulo Research Foundation (FAPESP) has also attracted numerous diasporic members to return even though its initiative did not explicitly target Brazilians. Some of these recruitment schemes are certainly well intentioned for capacity building and social aspirations; however, the policy discourse sometimes belies an undercurrent of ethnic nationalism. In other words, despite the claims of meritocracy in policy rhetoric to attract global talent, policy implementation favors a dominant ethnic group to entrench or expand power well beyond academia.

Another key driver behind faculty mobility is the shifting landscape of higher education in many countries due to changes in demography and political economy. The declining birth rate in many industrialized societies naturally results in a shrinking cohort of domestic students entering higher education (Grawe, 2018; Lau, 2021). The onslaught of neoliberalism steadily erodes public funding for higher education. Consequently, program closures, budget cuts and institutional insolvency reduce the number of academic posts available. Nevertheless, the doctoral pipeline continues to produce a steady stream of new academics. The oversupply of doctorates in many of these societies persuade a growing number of early career researchers to consider employment overseas especially in countries with thriving higher education systems (Huisman et al., 2002; Chou, 2020). At the other end of the career spectrum, mandatory retirement age in some higher education systems motivates some senior academics to extend their careers or provide consultancy abroad to support new programs and institutions (Czaika and Toma, 2017). These contractions

¹This program was created in 2008 and renamed in 2019 to the "National High-End Foreign Experts Recruitment Plan."

and expansions among higher education systems worldwide mediate individual agency and ultimately affect faculty mobility on a global scale.

Academic mobility: current issues and future challenges

The international mobility of students and faculty share many commonalities: both are generally assumed to be positive and transformative for individuals and institutions (Le Ha and Fry, 2021); both are largely driven by the value of certified knowledge and talent in the contemporary global economy and its labor markets; both are underpinned by the discourses of cosmopolitanism, modernity and diversity as enshrined in the mission statements of many universities. To be sure, there are also notable differences: student mobility is largely driven by individual choice and aspirations (Tran, 2016), often in relation to long-term migration plans (Kahanec and Králiková, 2011). However, faculty mobility is not necessarily due to individual agency when constraining factors such as poor academic job markets, family responsibilities, and expectations of international experience as a rite of passage in academia force scholars to move (Ackers, 2008; Lee and Kuzhabekova, 2018; Sautier, 2021; Luczaj and Holy-Luczaj, 2022).

The parameters of faculty mobility are also frequently uncertain given the nature of career development and family relocations compared to student mobility with reasonably well defined time frames and outcomes. The acceptance of faculty mobility also varies widely among institutions and disciplines unlike the overall positive support for student mobility. Specifically, the rhetoric of “international mobility” as a critical part of contemporary academic work does not reflect actual practices at many higher education institutions when hiring committees view overseas experience with skepticism or lack the contextual knowledge to properly assess performance at a foreign institution. Discriminations against returnees and academics with a very mobile track record raise serious questions about the misalignment between policy and practice (Ackers, 2008). Due to their many similarities, student and faculty mobility share a common set of ongoing trends and future challenges. In this section, we discuss several issues that affect academic mobility: barriers in policy and practice, the debate on brain drain and brain circulation, geopolitical currents, and the existential challenges of climate change. These challenges demand a fundamental re-thinking of mobility as a pillar of internationalization when restricted access and destructive outcomes contravene internationalism.

Barriers in policy and practice

Mobility demands numerous conditions which are often unspoken in the discourse of international higher education. The vast majority of higher education students and faculty worldwide are not mobile despite the tireless efforts among universities and governments in promoting academic mobility. Many individuals cannot access mobility due to structural barriers and constraints. For example, a student who must manage part-time employment to fund his/her studies is not likely to participate in mobility due to financial constraints. An employer may be reluctant to grant the student extended leave from work to study abroad, and the student would also need to find an alternative income while overseas. Astronomical tuition fees for international students also place degree mobility beyond the reach of many students worldwide. Individuals with physical disabilities are largely excluded from academic mobility (Matthews et al., 1998; Johnstone and Edwards, 2020). Other marginalized individuals include those who face social restrictions on mobility: women in Middle Eastern countries who are unable to go abroad without a male guardian (Aldossari, 2015), and LGBTQ+ minorities who may encounter hostile attitudes and violence in foreign countries (Hubain, 2017; Adeyemo, 2020). For faculty mobility, studies show that participants are predominantly male in their early career (Selmer and Luring 2010; Bauder 2012), while female and married academics are far less mobile (Leemann, 2010; Morley et al., 2018). Therefore, without policies to widen access, mobility is an exercise in privilege by a selected few rather than an opportunity that is available to all.

On the policy front, some governments and institutions have also enacted explicit measures to regulate mobility and penalize immobility. It is not uncommon to find scholarships for studying abroad with restrictions on the choice of institutions. For example, both Chile and Kazakhstan generously fund overseas studies on the condition that the host institution must rank in the top 50 and 200, respectively, in the world (Perna et al., 2015). This gate-keeping in effect legitimizes the metrics of global rankings, overlooks disciplinary excellence at lower ranked institutions, and restricts individual agency. Such policies on mobility celebrate the templates of higher education in Anglophone countries and reinforce the inequality of global higher education. Furthermore, virtually all international scholarships require recipients to travel and live abroad because mobility is conceived as the essence of an international education. Online programs and hybrid programs are therefore dismissed as unworthy or assumed to be low quality. These regulatory policies effectively reify mobility as a signifier of excellence in higher education with a disregard for new innovations in higher education delivery in the last few decades.

Brain drain and brain circulation

In the past, academic mobility has followed predominantly South-to-North patterns, with universities in the Global North (high-income, industrialized countries) attracting students and faculty because of perceived opportunities and resources, in many cases coupled with the widespread use of English as a *lingua franca* (Shields, 2013; van der Wende, 2015). These persistent patterns have elicited criticisms about “brain drain,” a term that emerged in the mid-twentieth century to describe a persistent extraction of intellectual resources from institutions and societies in the Global South (Cervantes and Guellec, 2002). As posed in much critical

literature, brain drain creates a vicious cycle in which countries in the Global South must continually replenish their intelligentsia, ultimately leading to a chronic “under development” of education systems and skilled professions (Straubhaar, 2000).

However, recent literature takes a more critical view of “brain drain” as a concept, pointing out that it is underpinned by a narrow and archaic view of mobility and immobility. For example, Rizvi (2005) argues that academic mobility is driven not only by the demand for skilled labor, but also by networks within diasporic communities and by individuals’ own identities and aspirations. Rizvi (2005) also points out how the realities of globalization also mean that social networks and technology are better positioned to enable diasporic communities to make social contributions across borders, including but not limited to their country of origin.

By definition, the concept of brain drain assumes that contributions to a society require in-situ residence. Therefore, the diasporic members living outside the homeland cannot offer any meaningful contributions to national development. This view also represents an antiquated conception of knowledge production by assuming that an academic can only contribute to the development of one country – in other words, a zero-sum accounting of knowledge that refuses to consider concurrent benefits to multiple locales. Critics of brain drain frequently elevate nation-building as a collective moral responsibility among the intelligentsia. This perspective places an additional burden and responsibility on academics from the Global South by chastising those who make mobility decisions based on academic opportunities and personal aspirations (Siekierski, 2018). A more productive approach would be to invest resources in improving work conditions at local universities rather than to censure indigenous talents who are pursuing dreams overseas.

For these reasons, “brain circulation” is now often a preferred concept to describe how patterns of academic mobility affect the capacity and development of institutions and societies (Postiglione, 2013). Unlike “brain drain,” the concept of “brain circulation” recognizes the potential benefits of mobility to *both* the country of origin and destination as well as mobility that may be temporary or that may involve “North to North” or “South to South” migration. Nevertheless, brain circulation can increase inequality between countries, even for those in the Global North, as institutions and governments compete to attract highly-educated workers (van der Wende, 2015). This competition is part of a larger geopolitical context that has become increasingly acerbic and complex.

Geopolitical currents

Academic mobility is deeply implicated in wider geopolitical relationships in the age of globalization. The exchange of ideas and individuals is fraught with political ramifications and instrumental motives. For example, in 1946 the US Fulbright Hays Act was passed, which provided significant support for academic mobility in order to promote American interests and values abroad in the context of the emerging Cold War (Johnson, 2018). The liberal world order that crystalized at the end of the twentieth century offered a context in which academic mobility thrived: growing economic liberalization and integration reduced barriers to international mobility and enabled individuals to seek opportunities abroad. The Fulbright program continues to operate today to generate benefits for the American higher education system by facilitating the bilateral exchange of academics.

However, the early 21st century has witnessed significant cracks appearing in the political norms and geopolitical equilibrium that have provided a stable base for rising levels of academic mobility. For example, the increasing power of populist, nationalist, and xenophobic political movements in Western democracies – most apparent in the election of Donald J. Trump as president and the UK’s vote to leave the European Union in 2016 – signaled an insular turn in countries that had previously been active in international fora. Rising incidents of racism toward international students in the US, UK, and Australia have alarmed prospective students and sending countries (Yao, 2019; Menchin, 2021; Zewolde, 2021). The ongoing COVID-19 pandemic has further intensified racism toward international students of Asian descent as scapegoats of a global crisis. For the first time since data were recorded, the number of international students in the United States decreased (Laws and Ammigan, 2020). Altbach and de Wit (2017) argue that this change will lead to a longstanding shift in the ways these countries are perceived by international students, with accompanying shifts in mobility patterns. Statements from host country governments and universities may condemn acts of discrimination and reaffirm cultural openness, but action rather than rhetoric is critical in ensuring that mobility is conducive to meaningful experiences in learning, teaching and research.

Another important geopolitical shift has been the rise of new global powers and their adoption of academic mobility as part of a wider geopolitical strategy. In particular, the higher education systems of Brazil, Russia, India, China, and South Africa (collectively known as BRICS) are increasingly implicated in national aspirations and ambitious foreign policies. Of all the BRICS countries, China has been the most prominent in leveraging higher education as a part of its foreign policy through its network of Confucius Institutes worldwide, its recruitment of international students, and its revival of the historical Silk Road (Lo and Pan, 2016; Lee et al., 2021). Confucius Institutes are funded by China and established within existing foreign universities to promote the study of Mandarin and Chinese culture. Started in 2004, there were more than 500 Confucius Institutes in 134 countries by the end of 2017 (Ye and Edwards, 2018). Ongoing operation of these institutes entails significant outbound mobility of over 440,000 Chinese teachers to teach abroad (Ye and Edwards, 2018). Students and business professionals worldwide also travel to China after gaining some language and cultural competencies at Confucius Institutes. Due to its close connection to the Chinese state, Confucius Institutes have attracted significant controversy and even the label “Chinese trojan horse” (Paradise, 2009, p. 659). As concerns about academic freedom and intellectual property grow, the US Government moved to restrict federal funds to universities with Confucius Institutes, and many universities closed their Confucius Institutes (Hall-Martin, 2020). However, others have argued that the use of Confucius Institutes as a form of “soft power” in global politics is not different from the ubiquitous presence of the British Council or Alliance Française (Yang, 2010; Liu, 2019). In addition to these institutes, China has sponsored scholars from other countries and regions in which it has a strategic interest, most notably sub-Saharan Africa (Mulvey, 2021). While these strategies are not

conceptually different from those taken by other countries in the past, China's determination, coupled with its resources, can dramatically shift the dynamics of international mobility in higher education.

In other ways, geopolitical realities are sparking and restricting academic mobility. In this respect, the growing global population of refugees is a key trend to which higher education institutions must adapt and respond (Arar et al., 2020). The global refugee population is at a historic high and is growing rapidly, rising from 15.9 million in 2007 to 26.3 million in 2020 (World Bank, 2021b). Additionally, many of these individuals experience increasingly long periods of displacement, with the duration of protracted displacement exceeding 21 years on average (Devictor and Do, 2017). This growth means that many international students at universities are also refugees. However, in many policies and practices, the narrow framing of international student mobility as an endeavor couched in individual choice, opportunities, aspirations and returns on investment fails to address the needs of refugee students (Earnest et al., 2010). Organizations such as Scholars At Risk (US) and the Council for At-Risk Academics (UK) have been instrumental in relocating and protecting thousands of students and academics fleeing repressive regimes and conflict areas in the world. On another front, sudden diplomatic crises can also infringe academic mobility and effectively turn international students into pawns in a game of realpolitik. For example, in 2018, the Canadian Government criticized Saudi Arabia's record on gender rights only to face threats by the Saudi Government to recall all its students from Canadian universities (McKie, 2018). For some Canadian institutions that host a large number of Saudi students, this exodus posed a serious financial risk. Similarly, in the same year, confrontations between the right-leaning Hungarian Government and the progressive Central European University resulted in the entire university relocating to Vienna and dramatically upending academic mobility to this regional center of higher education.

Climate change

As described above, levels of academic mobility have accelerated greatly in recent decades: with over 5 million mobile students and significant levels of academic mobility for migration, research collaboration, and international conferences, the number of individuals participating in academic mobility approaches the population of a megacity. However, the air travel required for this mobility has a significant environmental cost, as aviation is one of the most carbon-intensive forms of travel (Rockström et al., 2017), and growing concentrations of carbon-dioxide and other greenhouse gases are responsible for creating disruptive changes in earth's climate that pose serious threats for all of humanity (UN, 2018). To highlight the extent of these costs, Shields (2019) estimates that international student mobility results in the release of 14–34 megatons of carbon dioxide into the atmosphere, which is equivalent to the entire annual national emissions of a small country such as Jamaica or Croatia.

While the benefits of academic mobility have been the subject of much research, serious consideration of its environmental costs is only just beginning. Such concerns must be raised from two perspectives: first, in terms of the issues of equity and justice involved, and second in terms of the contradictions they present for universities. In relation to equity, pollution from air travel highlights immense inequalities. For instance, Shields (2019) also shows that the carbon emissions from a one-way flight from London to Shanghai are greater than the average annual emissions of an individual in many countries. This disparity raises questions of whether the benefits of mobility, as great as they may be, can be compared with the needs of another individual over an entire year of life. Such concerns also affect faculty mobility, especially given evidence that air travel has little impact on research productivity (Wynes et al., 2019).

Second, a growing number of universities have issued statements about the seriousness of climate change and the need to take urgent action (Latter and Capstick, 2021), yet many also rely on large numbers of international students as a core revenue source in their financial models (Gill, 2020). This places universities in a state of contradiction, in which their own organizational behavior undermines their declarations based on scientific evidence (O'Neill and Sinden, 2021). Both the issue of equity and of contradiction suggest that a fundamental transformation of universities may be required in order for higher education to contribute meaningfully to sustainable development. This sentiment was expressed in the Peoples' Sustainability Treaty on Higher Education (2012) from the Rio+20 conference, which declared

Before higher education can genuinely contribute to sustainable development, it must transform itself. The dominant education paradigm is centered on values and priorities that threaten sustainable development.

By extension, for international mobility to remain a core aspect of higher education in the future, fundamental changes are likely required. The COVID-19 pandemic has forced students and researchers to learn and work together without being able to meet physically, whether they are in the same city or separated by great distances. Large academic conferences have been moved online with seemingly little impact on the quality or volume of research.

Futures of mobility

This chapter has discussed academic mobility as a significant force that shapes contemporary higher education. A dramatic rise of international students and faculty has reshaped universities and systems of higher education to the extent that many could not exist

as we know them today without pursuing high levels of mobility. At the same time, there are increasing questions about whether this unrelenting rise will continue, or even whether current levels of mobility can be sustained. On the one hand, the drivers for both student and faculty mobility continues unabated; the need for skilled labor is high in many fields; opportunities for faculty are opening in rapidly growing higher education systems; universities are providing high levels of support for international students. On the other hand, it is hard to ignore the clouds on the horizon: xenophobic politics, tensions between new global powers, and the existential threats of climate change all pose very serious challenges to mobility as we know it today.

The COVID-19 pandemic is unequivocally a tragedy in terms of loss of life and health, but it also presents an opportunity to rethink academic mobility and international education (Oleksiyenko et al., 2021). As the spread of the pandemic prevented physical mobility, the academic community quickly found new ways to work together online, as large international conferences and university degree programmes moved online (Silova et al., 2020). These adaptations can be seen as part of a more general trend to replace or complement large-scale mobility of individuals by new forms of provision. For example, twinning programs and international branch campuses focus on the mobility of programmes and institutions rather than individuals (Knight, 2016). Online platforms also have many opportunities for sustained intercultural exchanges among students if embedded within courses and program curricula rather than organized as single events.

The direction that academic mobility takes in the future may help us better understand its drivers and underlying causes. If academic mobility is driven by human capital (the demand to acquire economically productive skills and to source skilled labor) then new forms of online learning and collaboration may replace physical movements across borders. In contrast, if a large part of mobility is actually driven by the demand for experience, of realizing new identities and cosmopolitan dispositions, then it will be hard for online engagement to replace physical mobility. Wider cultural assumptions, which tend to view immobility as negative and associate prestige with residence in the Global North, may also be hard to dispel (Chou, 2020). Thus, through the continued study of academic mobility, one gains insights not only into an important aspect of contemporary higher education, but also into larger, underlying changes in societies around the world.

References

- Ackers, L., 2008. Internationalisation, mobility and metrics: a new form of indirect discrimination? *Minerva* 46 (4), 411–435. <https://doi.org/10.1007/s11024-008-9110-2>.
- Adeyemo, K.S., 2020. Filipino university students' attitude toward sexual minorities. *J. Int. Stud.* 10 (1), 203–219. <https://doi.org/10.32674/JIS.V10I1.1030>.
- Aldossari, A., 2015. Experience, Struggle and Change: Saudi Women Studying in a Canadian Context, under the King Abdullah Scholarship Program. Unpublished Doctoral Dissertation. University of Toronto, Canada.
- Altbach, P.G., 2004. Globalisation and the university: myths and realities in an unequal world. *Tert. Educ. Manag.* 10 (1), 3–25.
- Altbach, P.G., de Wit, H., 2017. Trump and the coming revolution in higher education internationalization. *Int. High. Educ.* (89), 3–5.
- Arar, K., Kondakci, Y., Streiwieser, B., 2020. Higher education for forcibly displaced migrants, refugees and asylum seekers. *High Educ. Pol.* 33 (2), 195–202. <https://doi.org/10.1057/s41307-020-00184-z>.
- Bauder, H., 2012. The international mobility of academics: a labour market perspective. *Int. Migrat.* 53 (1), 83–96.
- Böttcher, L., Araújo, N.A., Nagler, J., Mendes, J.F., Helbing, D., Hermann, H.J., 2016. Gender gap in the ERASMUS mobility program. *PLoS One* 11 (2), e0149514.
- Cervantes, M., Guellec, D., 2002. The brain drain: old myths, new realities. *OECD Obs.* 230, 40–42.
- Chan, S.J., 2012. Shifting patterns of student mobility in Asia. *High Educ. Pol.* 25 (2), 207–224.
- Chou, M.-H., 2020. Sticky and slippery destinations for academic mobility: the case of Singapore. *High Educ.* 1–16. <https://doi.org/10.1007/S10734-020-00574-7>.
- Collins, F.L., Ho, K.C., 2018. Discrepant knowledge and interAsian mobilities: unlikely movements, uncertain futures. *Discourse* 39 (5), 679–693.
- Coser, L.A., 1984. *Refugee Scholars in America: Their Impact and Their Experiences*. Yale University Press, New Haven, CT.
- Crossman, J.E., Clarke, M., 2010. International experience and graduate employability: stakeholder perceptions on the connection. *High Educ.* 59 (5), 599–613.
- Czaika, M., Toma, S., 2017. International academic mobility across space and time: the case of Indian academics. *Popul. Space Place* 23 (8), e2069. <https://doi.org/10.1002/PSP.2069>.
- Devictor, X., Do, Q.T., 2017. How many years have refugees been in exile? *Popul. Dev. Rev.* 355–369.
- De Wit, H., 2008. Changing dynamics in international student circulation: meanings, push and pull factors, trends, and data. In: de Wit, H., Agarwal, P., Said, M.E., Sehoole, M.T., Sirozi, M. (Eds.), *The dynamics of international student circulation in a global context*. Brill, Leiden, pp. 15–45.
- Dickerson, C., 2020. "My World Is Shattering": Foreign Students Stranded by Coronavirus. Accessed at: <https://www.nytimes.com/2020/04/25/us/coronavirus-international-foreign-students-universities.html>. (Accessed 15 July 2021).
- Earnest, J., Joyce, A., De Mori, G., Silvagni, G., 2010. Are universities responding to the needs of students from refugee backgrounds? *Aust. J. Educ.* 54 (2), 155–174.
- European Commission, 2020. Erasmus+ and European Solidarity Corps: Investing for Europe's Future Generations. Press Release, 16 December 2020. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2422. (Accessed 5 July 2021).
- Findlay, A.M., 2011. An assessment of supply and demand-side theorizations of international student mobility. *Int. Migrat.* 49 (2), 162–190.
- Findlay, A.M., King, R., Smith, F.M., Geddes, A., Skeldon, R., 2012. World class? An investigation of globalisation, difference and international student mobility. *Trans. Inst. Br. Geogr.* 37 (1), 118–131.
- Fleming, D., Bailyn, B., 1969. *The Intellectual Migration: Europe and America, 1930–1960*. Harvard University Press, Cambridge.
- Gill, M., 2020. High flying business schools: working together to address the impact of management education and research on climate change. *J. Manag. Stud.* 58 (2), 554–561.
- Grave, N.D., 2018. *Demographics and the Demand for Higher Education*. Johns Hopkins University Press, Baltimore, US.
- Gürüz, K., 2011. *Higher Education and International Student Mobility in the Global Knowledge Economy: Revised and Updated Second Edition*. SUNY Press, Albany.
- Hall-Martin, M., 2020. Institutional responses to environmental pressures: Confucius institute closures in the United States. *J. Comp. Int. High. Educ.* 12 (6S1).
- Han, S., 2021. Empowered or disempowered by mobility? Experience of international academics in China. *Stud. High. Educ.* <https://doi.org/10.1080/03075079.2021.1876649>.
- Hubain, B.S., 2017. *The Spaces between Us: A Queer ↔ Intersectional Analysis of the Narratives of Black Gay International Students*. Unpublished doctoral dissertation, University of Denver, USA.
- Huisman, J., de Weert, E., Bartelse, J., 2002. Academic careers from a European perspective: the declining desirability of the faculty position. *J. High. Educ.* 72 (1), 141–160.
- Institute of International Education, 2020. United States Hosts Over 1 Million International Students for the Fifth Consecutive Year.
- Institute of International Education, 2021. Open Doors Data. <http://www.iie.org/opendoors>. (Accessed 15 July 2021).

- Johnson, L.R., 2018. The program and the philosophy and Fulbright geography of US exchange programs since World War II. In: Tournès, L., Scott-Smith, G. (Eds.), *Global Exchanges: Scholarships and Transnational Circulations in the Modern World*. Berghahn Books, New York, pp. 173–185.
- Johnstone, C., Edwards, P., 2020. Accommodations, accessibility, and culture: increasing access to study abroad for students with disabilities. *J. Stud. Int. Educ.* 24 (4), 424–439.
- Jöns, H., 2016. The University of Cambridge, academic expertise and the British empire, 1885–1962. *Environ. Plann.* 48 (1), 94–114.
- Kahanec, M., Králíková, R., 2011. Pulls of International Student Mobility. Discussion Paper 6233. Institute for the Study of Labor, Bonn.
- Kaushal, N., Lanati, M., 2019. International Student Mobility: Growth and Dispersion. Working Paper No. 25921. National Bureau of Economic Research.
- Kim, T., 2009. Shifting patterns of transnational academic mobility: a comparative and historical approach. *Comp. Educ.* 45 (3), 387–403.
- Knight, J., 2016. Transnational education remodeled: toward a common TNE framework and definitions. *J. Stud. Int. Educ.* 20 (1), 34–47.
- Kratz, F., Netz, N., 2018. Which mechanisms explain monetary returns to international student mobility? *Stud. High. Educ.* 43 (2), 375–400.
- Latter, B., Capstick, S., 2021. Climate emergency: UK universities' declarations and their role in responding to climate change. *Front. Sustain.* 2, 660596. <https://doi.org/10.3389/frsus.2021.660596>.
- Lau, J., 2021. Demographic dive leaves South Korea struggling to fill campuses. *Times Higher Education*. Retrieved from: <https://www.timeshighereducation.com/news/demographic-dive-leaves-south-korea-struggling-fill-campuses>. (Accessed 20 July 2021).
- Laws, K., Ammigan, R., 2020. International students in the Trump era: a narrative view. *J. Int. Stud.* 10 (3), xviii–xxii.
- Lee, J.T., 2014. Education hubs and talent development: policymaking and implementation challenges. *High Educ.* 68 (6), 807–823. <https://doi.org/10.1007/s10734-014-9745-x>.
- Lee, J.T., Kuzhabekova, A., 2018. Reverse flow in academic mobility from core to periphery: motivations of international faculty working in Kazakhstan. *High Educ.* 76 (2), 369–386. <https://doi.org/10.1007/s10734-017-0213-2>.
- Lee, J.T., Lo, W.Y.W., Abdrasheva, D., 2021. Institutional logic meets global imagining: Kazakhstan's engagement with China's Belt and Road Initiative. *High Educ.* 82 (2), 237–253. <https://doi.org/10.1007/s10734-020-00634-y>.
- Le Ha, P., Fry, G.W., 2021. Editorial: international educational mobilities and new developments in Asia's higher education: putting transformations at the centre of inquiries. *Res. Comp. Int. Educ.* 16 (3), 199–208.
- Leemann, R.J., 2010. Gender inequalities in transnational academic mobility and the ideal type of academic entrepreneur. *Discourse* 31 (5), 605–625.
- Liu, X., 2019. So similar, so different, so Chinese: analytical comparisons of the Confucius Institute with its western counterparts. *Asian Stud. Rev.* 43 (2), 256–275.
- Lo, J.T., Pan, S., 2016. Confucius Institutes and China's soft power: practices and paradoxes. *Compare* 46 (4), 512–532. <https://doi.org/10.1080/03057925.2014.916185>.
- Luczaj, K., Holy-Luczaj, M., 2022. International academics in the peripheries: a qualitative meta-analysis across fifteen countries. *Int. Stud. Sociol. Educ.* 1–26. <https://doi.org/10.1080/09620214.2021.2023322>.
- Matthews, P.R., Hameister, B.G., Hosley, N.S., 1998. Attitudes of college students toward study abroad: implications for disability service providers. *J. Postsecond. Educ. Disabil.* 13, 67–77.
- McKie, A., 2018. Saudi removal of students from Canada is 'warning shot' to US and UK. *High Educ.* <https://www.timeshighereducation.com/news/saudi-removal-students-canada-warning-shot-us-and-uk>. (Accessed 27 July 2021).
- Menchin, J., 2021. Address "Deep-rooted" Racism to Support Int'l Students, Australian Gov't Told. *PIE News*. <https://thepienews.com/news/address-deep-rooted-racism-to-support-intl-students-australian-govt-told/>. (Accessed 26 July 2021).
- Michl, T., Pegg, K., Kracen, A., 2019. Gender X culture: a pilot project exploring the study abroad experiences of trans and gender expansive Students. *Frontiers* 31 (2), 32–50.
- Mohajeri Norris, E., Gillespie, J., 2009. How study abroad shapes global careers: evidence from the United States. *J. Stud. Int. Educ.* 13 (3), 382–397.
- Morley, L., Alexiadou, N., Garaz, S., González-Monteaudo, J., Taba, M., 2018. Internationalisation and migrant academics: the hidden narratives of mobility. *High. Educ.* 76 (3), 537–554.
- Mueller, B., 2020. 'It Really Was Abandonment': Virus Crisis Grips British Universities. <https://www.nytimes.com/2020/10/06/world/europe/virus-UK-universities.html>. (Accessed 15 July 2021).
- Mulvey, B., 2021. "Decentering" international student mobility: the case of African student migrants in China. *Popul. Space Place* 27 (3), e2393.
- Oleksienko, A., Blanco, G., Hayhoe, R., Jackson, L., Lee, J., Metcalfe, A., Sivasubramanian, M., Zha, Q., 2021. Comparative and international higher education in a new key? Thoughts on the post-pandemic prospects of scholarship. *Compare* 51 (4), 612–628. <https://doi.org/10.1080/03057925.2020.1838121>.
- O'Neill, K., Sinden, C., 2021. Universities, sustainability, and neoliberalism: contradictions of the climate emergency declarations. *Polit. Govern.* 9 (2), 29–40.
- Paige, R.M., Fry, G.W., Stallman, E.M., Josić, J., Jon, J.E., 2009. Study abroad for global engagement: the long-term impact of mobility experiences. *Intercult. Educ.* 20 (Suppl. 1), S29–S44.
- Paradise, J.F., 2009. China and international harmony: the role of Confucius Institutes in bolstering Beijing's soft power. *Asian Surv.* 49 (4), 647–669.
- Perna, L.W., Orosz, K., Jumakulov, Z., 2015. Understanding the human capital benefits of a government-funded international scholarship program: an exploration of Kazakhstan's Bolashak program. *Int. J. Educ. Dev.* 40, 85–97. <https://doi.org/10.1016/j.ijedudev.2014.12.003>.
- Pietsch, T., 2010. Wandering scholars? Academic mobility and the British world, 1850–1940. *J. Hist. Geogr.* 36 (4), 377–387.
- Postiglione, G.A., 2013. Anchoring globalization in Hong Kong's research universities: network agents, institutional arrangements, and brain circulation. *Stud. High Educ.* 38 (3), 345–366. <https://doi.org/10.1080/03075079.2013.773605>.
- Rivza, B., Teichler, U., 2007. The changing role of student mobility. *High Educ. Pol.* 20 (4), 457–475.
- Rizvi, F., 2005. Rethinking "brain drain" in the era of globalisation. *Asia Pac. J. Educ.* 25 (2), 175–192.
- Rizvi, F., 2011. Theorizing student mobility in an era of globalization. *Teach. Teach.* 17 (6), 693–701.
- Rockström, J., Gaffney, O., Rogelj, J., Meinshausen, M., Nakicenovic, N., Schellnhuber, H.J., 2017. A roadmap for rapid decarbonization. *Science* 355 (6331), 1269–1271.
- Roy, A., Newman, A., Ellenberger, T., Pym, A., 2019. Outcomes of international student mobility programs: a systematic review and agenda for future research. *Stud. High Educ.* 44 (9), 1630–1644.
- Royal Society, 2011. *Knowledge, Networks, and Nations: Global Scientific Collaboration in the 21st Century*. The Royal Society, London.
- Sautier, M., 2021. Move or perish? Sticky mobilities in the Swiss academic context. *High Educ.* 82 (4), 799–822. <https://doi.org/10.1007/s10734-021-00722-7>.
- Selmer, J., Laurant, J., 2010. Self-initiated academic expatriates: inherent demographics and reasons to expatriate. *Eur. Manag. Rev.* 7 (3), 169–179.
- Shaw, S., Thomas, C., 2006. Discussion note: social and cultural dimensions of air travel demand: hyper-mobility in the UK? *J. Sustain. Tourism* 14 (2), 209–215.
- Shields, R., 2013. Globalization and international student mobility: a network analysis. *Comp. Educ. Rev.* 57 (4), 609–636.
- Shields, R., 2016. Reconsidering regionalisation in global higher education: student mobility spaces of the European Higher Education Area. *Compare* 46 (1), 5–23.
- Shields, R., 2019. The sustainability of international higher education: student mobility and global climate change. *J. Clean. Prod.* 217, 594–602.
- Shields, R., Edwards, R.M., 2010. Student mobility and emerging hubs in global higher education. In: Portnoi, L.M., Rust, V.D., Bagley, S.S. (Eds.), *Higher Education, Policy, and the Global Competition Phenomenon*. Palgrave Macmillan, New York, pp. 235–248.
- Shinn, C., Welch, A., Bagnall, N., 1999. Culture of competition? Comparing international student policy in the United States and Australia. *J. Furth. High. Educ.* 23 (1), 81–99.
- Siekierski, P., Lima, M.C., Borini, F.M., 2018. International mobility of academics: brain drain and brain gain. *Eur. Manag. Rev.* 15 (3), 329–339.
- Silova, I., Millei, Z., Goebel, J., Manion, C., Read, R., 2020. Beyond the human: rethinking education and academic conferencing during the times of climate crisis. *Comp. Educ. Rev.* 64 (4), 749–752.
- Souto-Otero, M., Huisman, J., Beerens, M., de Wit, H., Vujić, S., 2013. Barriers to international student mobility: evidence from the Erasmus program. *Educ. Res.* 42 (2), 70–77.
- Straubhaar, T., 2000. *International Mobility of the Highly Skilled: Brain Gain, Brain Drain or Brain Exchange*. HWWA Discussion paper, No. 88. Hamburg Institute of International Economics (HWWA), Hamburg.
- Stroud, A.H., 2010. Who plans (not) to study abroad? An examination of US student intent. *J. Stud. Int. Educ.* 14 (5), 491–507.

- Tang, D., 2020. British Universities Charter Jets to Fly in Chinese Students. <https://www.thetimes.co.uk/article/british-universities-charter-jets-to-fly-in-chinese-students-0p3775rr2>. (Accessed 15 July 2021).
- Tarrant, M.A., 2010. A conceptual framework for exploring the role of studies abroad in nurturing global citizenship. *J. Stud. Int. Educ.* 14 (5), 433–451.
- Tran, L.T., 2016. Mobility as 'becoming': a Bourdieuan analysis of the factors shaping international student mobility. *Br. J. Sociol. Educ.* 37 (8), 1268–1289.
- UNESCO Institute for Statistics, 2021. UIS.Stat: Inbound Internationally Mobile Students by Continent of Origin. <http://data.uis.unesco.org/>. (Accessed 5 July 2021).
- United Nations, 2018. Climate Change: An 'existential Threat' to Humanity, UN Chief Warns Global Summit. <https://news.un.org/en/story/2018/05/1009782>. (Accessed 12 July 2021).
- Van Der Wende, M., 2015. International academic mobility: towards a concentration of the minds in Europe. *Eur. Rev.* 23 (S1), S70–S88.
- World Bank, 2021a. International Tourism, Number of Arrivals. <https://data.worldbank.org/indicator/ST.INT.ARV.L>. (Accessed 26 July 2021).
- World Bank, 2021b. Refugee Population by Country or Territory of Asylum. <https://data.worldbank.org/indicator/SM.POP.REFG>. (Accessed 26 July 2021).
- Wynes, S., Donner, S.D., Tannason, S., Nabors, N., 2019. Academic air travel has a limited influence on professional success. *J. Clean. Prod.* 226, 959–967.
- Yang, R., 2010. Soft power and higher education: an examination of China's Confucius Institutes. *Glob. Soc. Educ.* 8 (2), 235–245. <https://doi.org/10.1080/14767721003779746>.
- Yao, C.W., Mwangi, C.A.G., Brown, V.K.M., 2019. Exploring the intersection of transnationalism and critical race theory: a critical race analysis of international student experiences in the United States. *Race Ethn. Educ.* 22 (1), 38–58. <https://doi.org/10.1080/13613324.2018.1497968>.
- Ye, W., Edwards, V., 2018. Confucius institute teachers in the UK: motivation, challenges, and transformative learning. *Race Ethn. Educ.* 21 (6), 843–857.
- Zewolde, S., 2021. Racism and Othering in International Higher Education: Experiences of Black Africans in England. Global Centre for Higher Education. Working Paper no. 62. <https://www.researchcghe.org/perch/resources/publications/working-paper-62.pdf>. (Accessed 26 July 2021).

Internationalization of higher education

Ulrich Teichler, International Centre for Higher Education Research Kassel (INCHER-Kassel), University of Kassel, Kassel, Germany

© 2023 Elsevier Ltd. All rights reserved.

The exceptional term and its undercurrents	239
Definitions and conceptualizations	240
International student mobility	241
Direction of mobility	241
Temporary mobility versus mobility for a complete study program	242
“Vertical” versus “horizontal” mobility and various objectives of mobility	242
Varied functions	242
Statistics and indicators	243
The value of student mobility	243
International mobility and international interaction of academics	244
Academic mobility—a major feature of internationalization	244
Irritating statistics	244
Beyond mobility: international approaches of academic work and international collaboration	245
Varied modes of knowledge transfer	245
Transnational higher education	245
MOOCs	246
Internationalization of the substance of curricula, teaching and learning	246
Future perspectives	247
References	247

The exceptional term and its undercurrents

“Internationalization of higher education” became a widely used term in the public discourse as well as in research on higher education since about the 1980s. It has been even more often on the agenda since the beginning of the 21st century. Obviously, the frequent border-crossing of knowledge and persons in higher education is viewed nowadays as a phenomenon deserving utmost attention.

The term differs from other key terms employed with reference to higher education through the ending “-ization”. In contrast to terms such as governance, teaching and learning or even “knowledge society”, it implies that the subject of analysis inevitably is in a process of growth or at least is always desired or expected to grow. When the trend of growing managerial power in higher education institutions is addressed, nobody sees the need to talk about “managerialization”, but proposals to use other terms in this area, such as “internationality” instead (see Teichler, 2012) did not gain popularity. This, of course, raises the question whether there is a stronger growth trend of border-crossing phenomena than of other key issues of higher education, whether there is a stronger normative consensus that growth should happen in this area, or whether this is just a terminological coincidence without substantial meaning.

“Internationalization”, however, is not the only term employed, when border-crossing features of higher education are reflected. There seem to be more parallel or competing terms than other key topics discussed in the public discourse on higher education.

First, “international higher education”—or occasionally “international education”—is preferred in many instances (see various articles in Deardorff et al., 2012): Either as an umbrella term without reference to specific features of internationalization, or specifically for addressing substantive thrusts and modes in learning or various ways of thinking and value judgments across countries (Marginson, 2014). We have to bear in mind, though, that the term “international higher education” often is used as well much wider: Not limited to border-crossing phenomena, but rather as “higher education in other countries”, i.e. higher education features of any kind outside the observers’ country.

Second, “globalization” has become a popular term since about the beginning of the 21st century. Occasionally, it seems to be understood as synonymous to internationalization. In contrast, specific notions are presented frequently: Either claiming that higher education all over the world is exposed to common powerful political thrusts or societal trends, or assuming that higher education itself converges rapidly all over the world (see for example various notions in OECD, 2010). Very often, we note a less polarized utilization of the two terms: While “internationalization” addresses border-crossing phenomena with the assumption that differences between countries and higher education systems remain salient, “globalization” suggests that frequent border-crossing is accompanied by or even supports the blurring of differences between countries (Teichler, 2004).

Third, in contrast, some terms underscore that cross-border phenomena in higher education are by no means world-wide similar. “Regionalization” and more specifically terms such as “Europeanization of higher education” (see for example Huisman and van der Wende, 2005) signal closer linkages between neighbor countries than between distant countries. Actually, international

cooperation and exchange in higher education is emphasized most strongly in Europe (see for example [Wächter, 2006](#)), but also plays an increasing role in other parts of the world (see for example [Yonezawa et al., 2014](#)). There are other country priorities of border-crossing activities as well, e.g. post-colonial or possibly neo-imperialist ties, targeted development aid policies, ties based on a common languages, etc.

Definitions and conceptualizations

Efforts have been made frequently to define “internationalization” and to explain the respective features of higher education—in publications comprising the views of a broad range of authors (e.g. [OECD, 2010](#); [Gaebel et al., 2011](#); [Deardorff et al., 2012](#); [de Wit et al., 2015](#); [Ogden et al., 2020](#)), in monographs of individual experts ([de Wit, 1997](#); [Knight, 2008](#)), in syntheses of the state of research on internationalization of higher education ([Teichler, 2004, 2019](#); [Kehm and Teichler, 2007](#); [Bedenlier et al., 2018](#)), and in critical accounts of such concepts and their underlying ideologies (e.g. [Buckner and Stein, 2020](#)).

In 1990, the US scholar and university manager Clark Kerr published an article titled “The internationalization of learning and the nationalization of the purposes of higher education: Two ‘Laws in Motion’ in conflict”. He pointed out that notions of universal knowledge, search of new knowledge all over the world, international knowledge and learning, and international physical mobility as widespread means of knowledge transfer are highly appreciated and can look back on long traditions. In contrast, the regulatory systems, the institutional settings, and many organizational features in higher education have been all the time under national regimes—[Kerr \(1990\)](#) even observed “nationalization” in those respects. Even though this article has not been referred so frequently, it might be viewed as a major early voice pointing out that “internationalization” in terms of knowledge is embedded into a logic different from that in terms of organization, cooperation, etc. This raises the question, of course, whether we could observe eventually also an organizational and substantial internationalization of the institutional setting of higher education in recent years.

Most overviews on internationalization of higher education name between half a dozen and a dozen thematic areas. A frequently quoted overview on respective research, for example, named seven major areas: (a) International mobility of students and academic staff, (b) mutual influences of higher education systems on each other, (c) internationalization of the substance of teaching, learning and research, (d) institutional strategies of internationalization, (e) knowledge transfer, (f) cooperation and competition, (g) national and supranational policies regarding the international dimension of higher education ([Kehm and Teichler, 2007](#)).

Based on a review of the state of respective research, six “key meanings” of internationalization of higher education had been presented by the author of this article as “most widely spread”:

- Worldwide/border-crossing knowledge transfer (books, media, etc.),
- Physical mobility across countries (students, academic staff, administrative staff, etc.),
- International cooperation and communication (between countries, institutions of higher education, individual scholars, etc.),
- International education and research (comparative approaches, intercultural learning, socialization for international understanding, etc.),
- International similarity (convergence, globalization, Europeanization, etc.),
- International reputation (“world-class universities”, “international quality”, etc.) ([Teichler, 2017](#)).

This list of “key meanings” does not only refer to border-crossing of knowledge and border-crossing of persons and institutional cooperation, but also to the perceptions of the world scenery of higher education: Do national higher education system remain clearly distinct? Do they become more similar? Does increasing international interaction in higher education call the “nationality” of higher education into question? Or do we even observe a re-nationalization in terms of a fierce competition to put one’s own national system strongly on the international map?

It is not possible to present a list of the major areas of activities in the domain of “internationalization of higher education”, which could be accepted as appropriate by more or less all the key actors and key experts. Rather, we can easily present a multitude of diverse lists.

First, even a minimum consensus is lacking as regards the seemingly most obvious phenomena, as will be discussed below: For example, who is included and what is excluded, when we talk about international student mobility.

Second, internationalization policies and activities traditionally differ between countries and institutions according to a widely assumed pecking order of academic quality and reputation. For example, some academically highly reputed countries try to attract students, doctoral candidates and young scholars to go there and eventually the brightest of them to stay there. What often is called “brain gain” for the hosting countries, obviously is viewed as “brain drain” for the countries of origin, even though the latter countries eventually might also have some benefits in terms of academic cooperation, late returns, etc.—“brain circulation” (see the overview of this discourse in [Wächter, 2006](#)).

Third, we observe dramatic differences of internationalization policies between individual countries and regions of the world, which cannot be explained as consequence of the above named pecking order of quality and reputation. In an analysis of dominant policies in European countries undertaken in the early years of the 20th century, [Huisman and van der Wende \(2004\)](#) noted a spirit of “cooperation” prevailing in some countries, of “competition” dominant in some other countries, and a mixture of both in others.

Fourth, internationalization approaches change over time. Many observers noted a recent increase of economic and utilitarian thrusts of internationalization policies in various countries, for example vividly characterized as “From the pursuit of knowledge to the pursuit of revenue” (Reisberg and Rumbley, 2014): While some actors and experts believe that this growing competitive spirit is beneficial for academia and the knowledge system in many respects, others fear a loss of academic creativity (see the discussion in Krücken et al., 2018). Similarly, some observers note a growing nationalistic undercurrent—one expert even coined such an attitude in the most influential countries of the academic world “hegemonic internationalization” (Scott, 2015): Efforts to shape international cooperation in higher education in such a way that one’s own country gains financially, economically and politically at the expense of other countries. In contrast, “international understanding” and “global citizenship” are high on the agenda in other countries. Finally, some observers argue that there are changes of the political Zeitgeist which affect internationalization policies and activities almost everywhere; for example, a mood of “anti-internationalization” (Rhoades, 2017) seemed to spread in the second decade of the 21st century—a mood which on the one hand seems to endanger the growth of international activities and on the other hand to stimulate efforts of strengthening activities aimed at reinforcing “international understanding”.

Fifth, we note quite a dynamic development of actual activities: While most internationally mobile students aimed to spend the whole study program in another country initially, short-term mobility gained popularity after a while. While individual scholars were the obvious key actors of international cooperation for many years, higher education managers increasingly opted for institutional arrangements of cooperation in recent years supported by a growing number of professional staff in international offices, and many universities all over the world turned to explicit formulations of internationalization strategies (Sursock, 2016). While physical mobility—of students and academics—had been in the limelight for a long time, efforts to promote international learning without physical mobility spread in recent years—for example called “transnational higher education”, “MOOCs” and “internationalization at home”.

A close view at the available literature also shows that hundreds of definitions of internationalization of higher education have been put forward over the years. Given the multitude of phenomena under consideration as well as the diversity of policies, strategies and value judgments (see for example Altbach and Knight, 2007; Seeber et al., 2016), it does not come as a surprise to note that not a single definition has been accepted more or less generally and the academics and students have varied notions of the terms frequently employed (Cotton et al., 2019). The authors of a recent overview claim: “The most commonly accepted definition of internationalization is ‘the process of integrating an international, intercultural, or global dimension into the purpose, functions and delivery of post-secondary education’ (Knight, 2008)” (de Wit et al., 2015, p. 38). Obviously, however, this definition—putting emphasis on the teaching function of higher education—does not cover the actual breadth of notions.

The multitude of policies, strategies and value judgments of the actors involved and the multitude of activities actually undertaken cannot be mirrored completely in an encyclopedia article. Therefore, information will be provided primarily on international student mobility, international mobility and cooperation of academics, and various recent approaches to strengthen international dimensions of the substance of curricula, teaching and learning.

International student mobility

The use of the terms “international mobility in higher education”, obviously addressing all groups of mobile persons, further “international academic mobility”, either referring to the mobility of the academic profession (for example Cavalli and Teichler, 2015) or jointly to mobility of academics, students and administrators (see for example Blumenthal et al., 1996), and finally—most frequently—of “international student mobility” (see for example Byran and Dervin, 2008; van Mol, 2014; Ogden et al., 2020) underscores that border-crossing physical movement—notably of students, doctoral candidates and academic staff—is most frequently addressed in the public discourse about international features of higher education. Impressive figures are named, substantial funds are allocated to it, and institutions of higher education are visibly active to make mobility a success story. Thereby, international student mobility is most often in the limelight of public policies, support schemes, institutional strategies, administrative support as well as of assessment of the respective processes and outcomes—not the least, because statistics report even seven-digit figures.

Based on a superficial first-glance view one might draw the conclusion that international student mobility is a relatively homogeneous feature. For example, many institutions of higher education have a single international office taking care for all strands of student mobility, and international educational statistics present aggregates of all mobile students in higher or in tertiary education. A careful look, however, reveals an enormous diversity of goals, expectations, functions and results.

Direction of mobility

All moves of students from one to another country can be viewed from the side of the country of origin and from the side of the country of destination. Terms are employed such as “outwards” and “inwards” or “outbound” and “inbound” mobility, “outgoing” and “incoming” students, but also “foreign students” and “study abroad”—the latter with reference to the students’ nationality rather than directly to the location of the movement (see the overview in Teichler, 2017).

In some countries, strong efforts are made primarily to attract students from other countries, and a high proportion of foreign or incoming students often is viewed as indicating the academic quality in the hosting country or of the hosting university: Some

“rankings” of “world-class universities” take figures of inward mobile or foreign students as well as of foreign academic staff as indicators of academic excellence (Kehm and Stensaker, 2009). Notably the United States, the United Kingdom, Australia and New Zealand strive for and succeed in hosting more than ten times as many foreign students, who as a rule have to pay enormous tuition fees, as in encouraging students of their own country successfully to opt for study abroad: This is based on the widespread assumption in these countries that knowledge generated and disseminated at home is worth to be exported rather than being internationalized through cooperation and inflow of persons from other societies and cultures (see for example the critique voiced by Scott, 2021, p. 55). Other countries, for example France and Germany, are active in attracting students from many countries, in stimulating reciprocal flows, e.g. “exchange” with partner institutions abroad, for example through regular cooperation with partner institutions in other countries, and in encouraging their own students to collect international experience. Similarly, a reciprocal exchange with neighbor countries was encouraged in most European countries cooperating in the so-called “Bologna Process” (see Diener, 2019). Finally, outwards mobility dominates in many cases: About two-thirds students studying abroad—according to international educational education statistics—come from developing countries, where outwards mobility as a rule is much higher than inwards mobility (OECD, 2021). Many observers fascinated by absolute figures instead of paying attention to the proportion among all students, name China and India in these contexts, i.e. the countries with the largest overall population.

In contrast to the attention paid to the frequency of foreign or inwards mobile students, the ministers of the European countries cooperating in the “Bologna Process” even concluded in 2009, as will be discussed below, that increasing the proportion of one’s own student students studying or having other study-related experiences abroad is the single most important measurable target of internationalization policy (see details of this policy in Vögtle, 2019). Accordingly, international knowledge and international understanding on the part of own students became the most important goals to be strived for.

Temporary mobility versus mobility for a complete study program

The majority of international mobile students spend the whole study program up to the degree in the hosting country, for example 3, 4 or more years up to bachelor degree. A substantial proportion, however, studies in another country only for a year, half a year or an even shorter period. The former students are characterized, for example, as “degree mobile” or “diploma mobile” students, the latter as “temporarily mobile”, “short-term mobile” or “credit-mobile” students. Temporary student mobility notably got in the limelight of public discourse, since the European Commission established the ERASMUS program of the European Commission in the late 1980s and eventually provided funds for hundred thousands of students aimed at covering the supplementary costs of studying at a partner institution of another European country for at most 1 year. Research undertaken on student mobility suggests that a large proportion of students mobile for a whole study program often have chosen a university and a country considered to be academically prestigious, and that they aim to adapt more or less completely to the host institution’s and the host country’s environment. In contrast, temporary mobility is often preferred to collect other experience to that at home at an institution of more or less the same quality as at home: Variety of experience and learning from contrast are strived for to widen the horizon.

“Vertical” versus “horizontal” mobility and various objectives of mobility

The option to study in another country, the option to study there for a short period or all the study program as well as the choice of the host country and of the host institution are influenced by varied expectations as regards the conditions, processes and results of learning in a country different from that of origin. The author of this article has argued that two different types of mobility can be established along those lines: On the one hand “vertical mobility”, when students are not primarily interested in acquiring knowledge being distinctive from that at home, but rather in experiencing an academically superior level of teaching and learning to that at home, and on the other hand “horizontal mobility”, when a similar level of academic level is expected to help students to get acquainted with contrasting knowledge and a contrasting academic environment (Teichler, 2004). Enhancing academic quality versus broadening understanding of the variety of academic and cultural ways of thinking are the respective key goals.

Varied functions

A close look reveals that the functions of international student mobility are enormously varied. But it seems to be appropriate to name three major directions. First, students aim to learn differently in another countries in domains, in which they could learn more or less the same at home as well: For example a student of physics—a subject shaped by universal knowledge—is likely to opt for study in another country either for getting specialized in sub-areas not well provided at home or for getting exposed to areas of knowledge also provided at home, but abroad on a higher academic level. Second, students opt for study in another country in order to become specialists on other countries, cultures or languages, e.g. for Chinese studies or American studies, to become interpreters, foreign language teachers, etc., or to become experts of international phenomena, for example of international law. Third, students are internationally mobile, because they want to become versatile in doing comparative analyses as well as understanding different cultures and societies. This “learning from contrast” and this awareness of the diversity of options in most areas of academia and life are not bound necessarily to specific academic disciplines and professions, but often viewed as valuable in many life and work situations.

Statistics and indicators

Statistical figures tend to be presented to underscore the prominent role of international student mobility in the process of internationalization of higher education. We often read that student mobility increased impressively from about 200,000 or 300,000 about five decades ago to more than six millions nowadays (see for example [Cummins, 1991](#); [Banks and Bhandari, 2012](#); [Wächter and Ferencz, 2012](#); [Gérard and Sanna, 2017](#); [OECD, 2021](#)). The growth rate, however, looks quite moderate, if the increase of the overall number of students worldwide is taken into consideration: Accordingly, only about 2% of the students were enrolled in another country over a long period, and this figure has risen to at most 3% in recent years.

As a rule, annual statistics collected and published jointly by UNESCO, OECD and Eurostat (the statistical agency of the European Commission) are referred to. Reports published by the European Parliament ([Lanzendorf and Teichler, 2005](#)) and by the Academic Cooperation Association, the federation of national agencies for the promotion of international mobility in academia in European countries ([Kelo et al., 2006](#); [Teichler et al., 2011](#); [Wächter and Ferencz, 2012](#)), however, showed “deplorable” weaknesses of these statistics. Notably, the statistical agencies of the individual countries were asked by the international agencies to deliver figures of “foreign students”, even though more detailed figures available in some European countries indicated that only about three-quarters of the foreign students in these countries are “foreign mobile”, while about one-quarter are “foreign non-mobile”, i.e. do not change the country for the purpose of studying there, but rather have lived and learned already there prior to embarking on higher education; additionally, possibly about one-tenth of mobile students can be named “non-foreign mobile”: They have lived and learned in a country different from that of their citizenship prior to study and have turned to the country of their citizenship for the purpose of study, or they have lived in another country prior to study and have taken over the citizenship of host country of study while studying there. Moreover, the international agencies ask the national agencies not to include temporarily mobile students into the statistics, defined at times as students studying abroad for less than 1 year and at times as those studying abroad for up to 1 year—i.e. the type of mobile students viewed to be most important one in many countries and possibly comprising about a quarter of all internationally mobile students in the world.

The international agencies responded to this critique recently in encouraging the individual countries to distinguish foreign students, who turned to host country for the purpose of study, and foreign students who lived and/or learned there already prior to study; the former are called “international students” in recent statistical publications. But temporarily mobile students are aimed to be excluded in international statistics up to the present.

As mentioned above, most European countries involved in the Bologna Process have departed in 2009 from the strong emphasis placed on foreign or inwards mobile students. First, they pay primarily attention now and are primarily interested to raise the proportion of outgoing students: How many of our students gain international experience? Second, they suggested to measure international student mobility in a way that “degree-mobile” and temporarily mobile students are taken into consideration alongside; this is achieved by setting target for, what might be called the “event of student mobility in the course of study” ([Teichler, 2013](#)): Actually, the target was set in 2009 that 20% of all students from European countries ought to have acquired international experienced through study or study-related experience—temporarily or for the whole study period—in another country by the year 2020. New means of statistical data collection had to be promoted. Available data suggested at that time that the mean rate across countries was already above 10%; subsequent data collections showed a growth almost in tune with that target ([Education Audiovisual and Culture Executive Agency, 2018](#)). One could ask, however, what such a target for a whole continent really might mean, because the rate seems to have differed by country at the time, when the goals was formulated, from less than 5% to more than 20%.

The value of student mobility

An analysis of articles on internationalization of higher education published in reviewed journals ([Kosmützky and Putty, 2016](#)) suggests that “impact” and “outcomes” of student mobility are neglected area of research. In contrast, detailed substantive reviews ([Kehm and Teichler, 2007](#); [de Wit, 2009](#); [Deardorff and van Gaalen, 2012](#); [van Mol, 2014](#); [Egron-Polak et al., 2015](#)), drawing from a broad range of sources, showed a wealth of studies. Actually, most studies addressing the results of student mobility have been undertaken by administrators and practitioners and present information on a single or a few higher education institutions. Large comparative studies such as the “Study Abroad Evaluation Project” in the 1980s ([Oppen et al., 1990](#)) and those reviewing the ERASMUS program (notably [Teichler and Maiworm, 1997](#); [Janson et al., 2009](#); [European Commission, 2014](#))—all focusing on temporary mobile students—remained exceptional.

Many analyses address issues of study success. Obviously, many students mobile “vertically” for the whole study program face difficulties in coping with the academic expectations and the social settings abroad: According to some studies, rates of degree-mobile students’ prolongation and drop-out are often more than twice as high as those of host country students. As regards temporary mobility, attention often is was paid to the question, whether students’ study achievements abroad are completely or only selectively recognized upon return. Although the studies undertaken could not clearly determine the extent to which limited recognition was due to lower achievement, to different views about the compatibility of the substance of course abroad and at home, or to unfair procedures and judgments, this reinforced policies in favor of more complete recognition—most prominently the Lisbon Declaration 1997 of various international organizations (see [Teichler, 2003](#)).

Other analyses aim at establishing whether international mobility enhances the students’ international orientation. There are indications that “global awareness”, “international understanding”, etc. grow, but some studies suggest not to overestimate such

impact. For example, the above named comparative study undertaken in the 1980s showed that students opting for temporary mobility were already more internationally oriented prior to the sojourn than non-mobile students and hardly changed their attitude later; also students seem to have mixed experiences during the study period abroad and on average do not have a more positive view of the host country afterward. Further, quite a number of studies looks at the career of formerly mobile graduates; in most cases, slightly higher success as compared to formerly non-mobile graduates is reported, but the strongest impact is not “vertical” (easy transition to employment, higher status and income, etc.), but rather “horizontal”: A more frequent turn to international occupations and work task (temporary work assignments abroad, frequent contact with foreign countries and people, regular use of foreign languages, etc.). Finally, research in this domain suggests that the motives for studying in another country are enormously varied; the majority state later that they appreciate the experience and results, but according to a broader range of criteria than those in the limelight of the dominant internationalization policies and strategies.

International mobility and international interaction of academics

Academic mobility—a major feature of internationalization

International mobility of academic staff is not as frequently referred to in the public discourse as on internationalization as student mobility, but might be viewed as more influential for teaching, learning and research than any other means. A look at the diverse modes and functions of academic mobility supports such a view (see [Teichler, 2015](#)).

A survey undertaken in 2012 in 33 European countries showed that the majority of researchers in Europe had been abroad for some period during their academic career ([IDEA Consult, 2013](#)). Some academics spend a “sabbatical” abroad to refresh their academic horizon; others cooperate with scholars from other countries and visit them for a while. A survey undertaken in 2007–2008 in almost 20 countries worldwide indicated that almost one-fifth had been in another country during the last 3 years for teaching purposes ([Jung et al., 2014](#)). Support schemes have been established in many countries to facilitate short-term mobility of scholars, and some supra-national provisions are worth named as well: ERASMUS also comprised support continuously over the years for short-term teaching mobility; impressive programs of the European Union for short-term junior academic staff mobility were eventually made more visible in the late 1990s through the choice of the name “Marie Curie-Program”.

Available statistics suggests that about one-fifth of persons being awarded a doctoral degree are foreign, i.e. not citizens of the country in which the awarding-institution is located. On average of economically advanced countries, more than one-tenth of academics employed at universities are foreigners—in some countries even more than one-third.

Scholars involved in the above named survey in almost 20 countries developed typologies of academic mobility ([Goastellec and Pekari, 2014](#); [Huang, 2014](#))—without claiming, though, that the typologies cover the variety of moves completely. On the one hand, many scholars professionally active in the country of their citizenships have been internationally mobile in early stages of their career, e.g. “study mobile academics”, “Ph.D. mobile academics”, “early career mobile academics”, etc. On the other hand, many scholars having a citizenship different from that of the country of work or having had a different citizenship but eventually having turned to the citizenship of the country of work differ according to the stage of move to the current country of work: “Early immigrants”—already prior to study, “study mobile immigrants”—formerly mobile students who remained in the host country, “Ph.D. mobile immigrants”, and “professional immigrants”, i.e. those changing the country of employment at later career stages. The terms indicate that the study period, the doctoral phase and the early career are crucial mobility stages of academics.

Irritating statistics

Statistical overviews about the frequency of academics’ mobility are even more irritating than those on student mobility. The educational divisions of the international organizations collecting and publishing education statistics provide information on “teaching staff”, while the research divisions of the same international organizations collect and publish completely different statistics on “scientists”; only few reports draw from both sources. All these statistics inform about the number of foreigners and not about mobility for the purpose of academic work, about foreigners employed and not about those being in the country academically active for short periods without an employment status, and finally about those currently employed in the respective country and not those who had been mobile in earlier stages of live ([Teichler et al., 2011](#)).

Many observers conclude that a high proportion of foreign and other internationally experienced scholars is beneficial for the quality of teaching and research. As already mentioned with respect to student mobility, the rate of foreign academic staff is taken as an indicator in some ranking studies (see the overviews in [Kehm and Stensaker, 2009](#)). In some countries, deliberate efforts are made to enhance the academic capacity through the recruitment of talented foreign scholars. As already discussed above, such “brain gain” policies and the corresponding “brain drain” situation in other countries indicate most directly the often conflicting scene of internationalization of higher education ([Wächter, 2006](#)). Further drawbacks are reported: Substantial costs of mobility and costs of knowledge transfer ([Wagner, 2006](#); [Kwiek, 2018](#)); costs for advanced countries for helping catch-up developments in medium-income and low income countries; last not least problems of foreign and mobile academics to adapt to the administrative and cultural milieu of the host country and the host institution.

Beyond mobility: international approaches of academic work and international collaboration

Altogether, mobility of academics certainly is a key driver to the internationalization of knowledge. But we do not know as a rule, what role factors such as physical mobility actually play. Some surveys at least provide information on international activities of academics and on the role international elements play for their academic work. According to the above named survey undertaken in almost 20 countries, about 60% of academics on average across countries stated that they emphasize international perspectives in their teaching, more than half characterized their research as international in scope or orientation, and about 40% reported that they collaborate with colleagues abroad in research. The figures, however, vary substantially by country, as it might be illustrated for economically advanced countries. An international approach in research was stated by 75% of academics in Italy, 55% in Germany and only 41% in the United States. Collaboration in research with colleagues abroad was reported by more than 60% of academics in the Netherlands and the United Kingdom, 50% in Germany, and only 33% in the United States and 24% in Japan (Teichler et al., 2013; Huang et al., 2014; Rostan and Ceravolo, 2015). Internationality is widely spread in some countries across all these various measures. For example, few Japanese scholars are internationally mobile and collaborate with foreign scholars, but many emphasize internationality in the substance of their academic work. Finally, a substantial proportion of scholars in the United States is not international according to the measures named, while internationality of higher education in this country seems to be reinforced notably through the “import” of doctoral candidates and junior academics.

Varied modes of knowledge transfer

Even though physical mobility of students and academics has been an impressively growing component as well as a major target of policies and strategies to internationalize higher education in recent decades, other modes of knowledge transfer across borders never were viewed as marginal and have been more strongly emphasized in recent decades. Notably, the conditions for spreading information across borders improved through ICT, and efforts intensified to internationalize the views and expertise also of the non-mobile students.

Cooperation based on informal links or on formal contracts spread enormously between small units, departments or whole institutions of higher education of two or more countries (see for example Seeber et al., 2016). Many universities have a list of more than a hundred partnerships. Approaches and arrangements of such partnerships are so manifold that hardly any generalizable statement can be made about concepts, operations and impact. Evaluations of internationalization strategies of universities undertaken in many countries, however, show that this is a typical tension: If a department or a university sets a priority of cooperating with certain partners based on certain goals, some internationally active academics of that university and department fear that the broad range of other individual initiatives based on the variety of expertise within this institutional setting might be disadvantaged.

Some moves to shape international cooperation newly became overarching movements. “Transnational higher education”, “MOOCs” as well as various approaches to internationalize the substance of study programs were the most visible ones in the first two decades of the 21st century.

Transnational higher education

In recent decades activities increased to provide students in selected universities or departments the opportunity to benefit from the substance and the mode of teaching and learning prevailing in partner countries without being internationally mobile. For example, Knight (2012) identified a generational shift from “people mobility” to “program and provider mobility” and eventually “educational hubs”. Terms such as “borderless”, “cross-border”, “off-shore” higher education and “foreign-backed” universities spread, but eventually “transnational higher education” became the most widely accepted term (Amaral, 2016; Knight and McNamara, 2017; Cracium and Oroc, 2018; see also Kosmützky and Putty, 2016).

Two types of transnational higher education developed: On the one hand, what also often had been called “off-shore higher education” or also the “satellite model” of transnational higher education (Knight, 2015): Universities with a high academic standard and often with a strong commercial interest—frequently universities in the US, the UK and Australia—established regular study programs in academically less favored countries. A “branch campus” or a regular university might have been founded there—possibly with a substantial proportion of mobile academic staff, whereby the curricula and the degree award are controlled by the mother institution abroad. On the other hand, “cooperative higher education” is preferred by governments and higher education institutions in Germany, France and some other European countries and their partners. Departments or universities of a bi-national character are established—often characterized as “joint” or “bi-national” institutions: The academic thrust of the partner institution abroad shapes the substance of teaching, learning and research to a considerable extent, mobile teacher play a major role at least in the initial years, and degrees of the host partner might be awarded initially, but administrative matters are decided upon jointly from the outset, and the institution is expected to be fully recognized in the country of location and to grow gradually into its national system in administrative terms while keeping key elements of border-crossing substance.

Transnational higher education cannot be arranged easily. Some “branch campuses” operate well, but there have been cases criticized as “degree mills” or “rogue” arrangements; Therefore, international organizations—notably UNESCO and OECD—cooperated in establishing quality criteria (see [Rosa et al., 2016](#)). Governments, managers and academics of the two countries involved in each case often turned out to have conflicting views about the extent of influence of the host side and the host academic characteristics versus its adaption to the home country’s thrusts (for example [Smith, 2010](#); [Qin, 2021](#)).

MOOCs

Around the turn of the century the idea gained momentum—initially in the United States and subsequently in other countries—to make individual courses or even complete study programs of universities available online all over the world. Various distance universities decided to offer their courses to their students online, and various highly reputed universities decided to make individual courses publicly accessible everywhere. Within a few years, virtual knowledge transfer expanded at an enormous pace. In this framework “Massive Open Online Courses”, abbreviated “MOOCs”, became a popular concept. Some advocates of this movement even believed that most students all over the world eventually would listen to presentations of a few famous scholars and take whole study programs from famous universities far-away, while local academics would shift to the role of tutors. The title of a book, which took stock of this movement, i.e. “From Books to MOOCs?” ([de Corte et al., 2016a](#)), underscores the impressive shift as regards the media of knowledge transfer, and the book itself points out as well the possible organizational and financial changes in higher education.

After a few years, the MOOCs movement could be called a “hype” in the public discourse on innovative developments in higher education (see for example [Kim, 2014](#)). It advocates argued that MOOCs promote democratization of higher education by providing free open access to knowledge and teaching for everybody everywhere. Moreover, acquiring knowledge online in such a way was expected to be embedded quickly by many didactic innovations, e.g. creative modes of group learning and creative mixes of online-learning and face-to-face teaching and learning arrangements. High hopes, however, soon gave way to criticisms that MOOCs involve the risk of just distributing conventional standardized modes of teaching and curricula, e.g. heavy reliance merely on lectures and multiple choice exams. Practices of offering online courses freely, but charging substantial fees for the assessment of students’ achievements and the certification of courses were viewed as even increasing inequality in higher education. Some research projects came to the conclusion that drop-outs rates in MOOCs arrangements are higher than in traditional provisions ([de Corte et al., 2016b](#)). Yet, activities spread to safeguard the quality, and MOOCs and other “open learning” arrangements continue to be expected to expand in the future.

Internationalization of the substance of curricula, teaching and learning

Already at times, when international mobility was paid utmost attention as a means for internationalization of higher education, actors and experts never lost sight of the view that the substance of learning, teaching and research is crucial. While mobility is a means for opening the opportunity of enhancing international knowledge and international values—often in ways of individual “experiential learning”, more direct ways of shaping the acquisition of international knowledge and stimulating shifts of attitudes and values toward increasing international open-mindedness and—more broadly—toward “intercultural competences” (see [Gregersen-Hermans, 2015](#)) frequently were strived for—for example with reference to the term “international higher education”.

When efforts to increase international student mobility were high on the agenda in the 1980s and 1990s, awareness grew as well than even the most optimistic targets discussed assumed that only a minority of students will be mobile. Eventually the idea gained popularity since the 1990s that non-mobile students should be prepared similarly to live in and to contributing to an internationalizing world. The concept of “internationalization at home”, first formulated in Europe in the 1990s, gained popularity in the 21st century (see [Beelen and Leask, 2011](#); [Robson et al., 2018](#))—defined as “the purposeful integration of international and intercultural dimension into the formal and informal curriculum for all students within domestic learning environments” ([Beelen and Jones, 2011](#), p. 76). Many institutions of higher education actually intensified efforts to offer chances of knowledge acquisition and personality development for non-mobile students similar to those of mobile students.

Further concepts surface to increase internationality of the substance of teaching and learning comprehensively. “Internationalization of the curriculum” was advocated—defined as the “incorporation of international, intercultural and/or global dimensions into the content of the curriculum as well as the learning outcomes, assessment tasks, teaching methods and support services of a program of study” ([Leask, 2015](#)). Also the term “comprehensive internationalization” is used to describe efforts to integrate strategically international, intercultural and global features into the values, processes and outcomes of higher education ([Hudzik, 2015](#)). Some concepts underscore both the importance of formal and informal processes of international learning in general, while others emphasize specific targets of specific activities, for example education students for “global citizenship” ([Horey et al., 2018](#)).

Efforts to internationalize the substance have been most frequently voiced with respect to curricula, teaching and learning. It is widely assumed as well that the substance of research becomes more international along the rising opportunities for knowledge transfer and along the increasing international cooperation of academics, as discussed above. Notably, scholars are increasingly expected to reflect the internationality of their knowledge transmission activities and to contribute in a targeted way to their students’ international competences.

Future perspectives

Notions about likely future perspectives often are shaped by the recent past. Most reflections regarding the domain under consideration here still adhere to the ending “-ization” in assuming that cross-border issues of higher education will be addressed increasingly in the future. The range of activities are likely to spread even further, and we might move toward a more “intelligent internationalization” (Goodwin and de Wit, 2020).

Moreover, many experts believe that most innovative movements in higher education will internationalize. For example, more direct ways of shaping society through higher education, as often called “third mission of higher education”, have led to concepts such as “internationalization of higher education for society” (Brandenburg et al., 2020; see also Leask and de Gayardon, 2021). Or growing research priorities, such as those emphasizing “sustainable development”, gradually will become themes in efforts to internationalize the curriculum (Gegersen-Hermanns, 2021).

Some thoughts about possible future development, however, are stimulated by ambivalent or by problematic developments in recent years. For example, experts observing the declining international mobility and the increasing digitalization of international communication in the course of the Corona pandemic (see for example Reimers and Marmolejo, 2021) ponder the relative strengths and weaknesses of future virtual communication versus face-to-face communication in higher education. And many actors and observers fear that moves toward a re-nationalization of politics and toward increasing “international misunderstanding” visible in the second decade of the 21st century might intensify in the future and might undermine the mutual trust required for the continuous internationalization of higher education.

References

- Altbach, P.G., Knight, J., 2007. The internationalization of higher education: motivations and realities. *J. Stud. Int. Educ.* 11 (3–4), 290–305.
- Amaral, A., 2016. Cross-border higher education: a new business? In: Rosa, M.J., Sarrico, C.S., Tavares, O., Amaral, A. (Eds.), *Cross-Border Higher Education and Quality Assurance*. Palgrave Macmillan, London, pp. 1–24.
- Banks, M., Bhandari, R., 2012. Global student mobility. In: Deardorff, D.K., de Wit, H., Heyl, J.D., Adams, T. (Eds.), *The Sage Handbook of International Higher Education*. Sage, Los Angeles, CA, pp. 379–397.
- Bedenlier, S., Kondakci, Y., Zawacki-Richter, O., 2018. Two decades of research on internationalization of higher education. *J. Stud. Int. Educ.* 22 (2), 108–135.
- Beelen, J., Jones, E., 2011. Redefining internationalisation at home. In: Curaj, A., Matei, L., Pricopie, R., Salmi, J., Scott, P. (Eds.), *The European Higher Education Area: Between Critical Reflections and Future Policies*. Springer Open, Cham, pp. 67–80.
- Beelen, J., Leask, B., 2011. Internationalisation at home on the move. In: Gaebel, M., Purser, L., Wächter, B., Wilson, L. (Eds.), *Internationalisation of European Higher Education: An EUA/ACA Handbook*. Raabe, Berlin.
- Blumenthal, P., Goodwin, C., Smith, A., Teichler, U. (Eds.), 1996. *Academic Mobility in a Changing World*. Jessica Kingsley Publishers, London and Bristol, PA.
- Brandenburg, U., de Wit, H., Jones, E., Leask, B., Drobner, A., 2020. *Internationalisation in Higher Education for Society (IHES)*. DAAD, Bonn.
- Buckner, E., Stein, S., 2020. What counts as internationalization? Deconstructing the internationalisation imperative. *J. Stud. Int. Educ.* 24 (2), 151–166.
- Byran, M., Dervin, F. (Eds.), 2008. *Students, Staff and Academic Mobility in Higher Education*. Cambridge Scholars Publishing, Newcastle.
- Mobility and migration in science (special issue). In: Cavalli, A., Teichler, U. (Eds.), *Eur. Rev.* 23 (Suppl. 1).
- Cotton, D.R.E., Morrison, D., Magne, P., Payne, S., Heffeman, T., 2019. Global citizenship and cross-cultural competency: student and expert understanding of internationalization terminology. *J. Stud. Int. Educ.* 23 (3), 346–364.
- Cracium, D., Orosz, K., 2018. *Benefits and Costs of Transnational Collaborative Partnerships in Higher Education*. Publications Office of the European Union, Luxembourg.
- Cummings, W.K., 1991. Foreign students. In: Altbach, P.G. (Ed.), *International Higher Education: An Encyclopedia*. Garland, New York, pp. 107–125.
- De Corte, E., Engwall, L., Teichler, U. (Eds.), 2016a. *From Books to MOOCs? Emerging Models of Learning and Teaching in Higher Education*. Portland, London.
- De Corte, E., Engwall, L., Teichler, U., 2016b. The hype of MOOCs. In: De Corte, E., Engwall, L., Teichler, U. (Eds.), *From Books to MOOCs? Emerging Models of Learning and Teaching in Higher Education*. Portland, London, pp. xv–xxv.
- De Wit, H., Hunter, F., Howard, L., Egron-Polak, E. (Eds.), 2015. *Internationalisation of Higher Education*. European Parliament, Brussels.
- De Wit, H., 1997. *Studies in international education*. *J. Stud. Int. Educ.* 1, 1–8.
- De Wit, H., 2009. *Measuring the Success of Internationalisation of Higher Education*. European Association for International Education, Amsterdam (EAIE Occasional Paper, 22).
- Deardorff, D.K., van Gaalen, A., 2012. Outcome assessment in the internationalization of higher education. In: Deardorff, D.K., de Wit, H., Heyl, J.D., Adams, T. (Eds.), *The SAGE Handbook of International Higher Education*. Sage, Los Angeles, CA, pp. 167–189.
- Deardorff, D.K., de Wit, H., Heyl, J.D., Adams, T. (Eds.), 2012. *The SAGE Handbook of International Higher Education*. Sage, Los Angeles, CA.
- 20th Anniversary of the Bologna Process: Europeanization through soft governance. In: Diener, C. (Ed.), *Innovat. Eur. J. Soc. Sci. Res.* 32 (4).
- Education, Audiovisual and Culture Executive Agency, Eurydice, 2018. *The European Higher Education Area in 2018: Bologna Process Implementation Report*. Publications Office of the European Union, Luxembourg.
- Egron-Polak, E., Hudson, R., Sandstrom, A.-M., 2015. Quantifying internationalisation—empirical evidence of internationalisation of higher education in Europe. In: De Wit, H., Hunter, F., Howard, L., Egron-Polak, E. (Eds.), *Internationalisation of Higher Education*. European Parliament, Brussels, pp. 59–76.
- European Commission, 2014. *The ERAMUS Impact Study: Effects of Mobility on the Skills and Employability of Students and the Internationalisation of Higher Education Institutions*. Publications Office of the European Union, Luxembourg.
- Gaebel, M., Purser, L., Wächter, B., Wilson, L. (Eds.), 2011. *Internationalisation of European Higher Education: An EUA/ACA Handbook*. Raabe, Berlin.
- Gérard, M., Sanna, A., 2017. Students' mobility at a glance: efficiency and fairness when brain drain and brain gain are at stake. *J. Int. Mobil.* 5, 43–74.
- Goastellec, G., Pekari, N., 2014. Gender differences and inequalities in academia: findings in Europe. In: Teichler, U., Hoehle, E.A. (Eds.), *The Work Situation of the Academic Profession in Europe: Findings of a Survey in Twelve Countries*. Springer, Dordrecht, pp. 55–78.
- Goodwin, K.A., de Wit, H. (Eds.), 2020. *Intelligent Internationalisation: The Shape of Things to Come*. Sense-Brill, Rotterdam.
- Gegersen-Hermanns, J., 2015. The impact of exposure to diversity in the international university environment and the development of intercultural competence in students. In: Curaj, A., Matei, L., Pricopie, R., Salmi, J., Scott, P. (Eds.), *The European Higher Education Area: Between Critical Reflections and Future Policies*. Springer Open, Cham, pp. 81–102.

- Gregersen-Hermans, J., 2021. Towards a curriculum for the future: synthesizing education for sustainable development and internationalization of the curriculum. *J. Stud. Int. Educ.* 25 (4), 461–481.
- Horey, D., Fortune, T., Nicolacopoulos, T., Kashima, E., Mathisen, B., 2018. Global citizenship and higher education: a scoping review of the empirical evidence. *J. Stud. Int. Educ.* 22 (5), 472–492.
- Huang, F., Finkelstein, M.J., Rostan, M. (Eds.), 2014. *The Internationalisation of the Academy: Changes, Realities and Prospects*. Springer, Dordrecht.
- Huang, F., 2014. The internationalization of the academic profession. In: Huang, F., Finkelstein, M.J., Rostan, M. (Eds.), *The Internationalization of the Academy: Changes, Realities and Prospects*. Springer, Dordrecht, pp. 1–21.
- Hudzik, J.K., 2015. *Comprehensive Internationalisation: Institutional Pathways to Success*. Routledge, New York.
- Huisman, J., van der Wende, M. (Eds.), 2004. *On Cooperation and Competition: National and European Policies for the Internationalisation of Higher Education*. Lemmens, Bonn.
- Huisman, J., van der Wende, M. (Eds.), 2005. *On Cooperation and Competition II: Institutional Responses to Internationalisation, Europeanisation and Globalisation*. Lemmens, Bonn.
- IDEA Consult, 2013. *Support for Continued Data Collection and Analysis Concerning Mobility Patterns and Career Paths of Researchers: Deliverable 8—Final Report More2*. European Commission, Research Directorate-General, Brussels.
- Janson, K., Schomburg, H., Teichler, U., 2009. *The Professional Value of ERASMUS Mobility*. Lemmens, Bonn.
- Jung, J., Kooij, R., Teichler, U., 2014. Internationalization and the new generation of academics. In: Huang, F., Finkelstein, M.J., Rostan, M. (Eds.), *The Internationalization of the Academy: Changes, Realities and Prospects*. Springer, Dordrecht, pp. 207–236.
- Kehm, B.M., Stensaker, B. (Eds.), 2009. *University Rankings, Diversity and the Landscape of Higher Education*. Sense Publishers, Rotterdam and Taipei.
- Kehm, B.M., Teichler, U., 2007. Research on internationalisation in higher education. *J. Stud. Int. Educ.* 11 (3–4), 260–273.
- Kelo, M., Teichler, U., Wächter, B. (Eds.), 2006. *EURODATA: Student Mobility in European Higher Education*. Lemmens, Bonn.
- Kerr, C., 1990. The internationalisation of learning and the nationalisation of the purposes of higher education: two “Laws in Motion” in conflict. *Eur. J. Educ.* 25 (1), 5–22.
- Kim, P. (Ed.), 2014. *Massive Open Online Courses: The MOOC Revolution*. Routledge, London.
- Knight, J., McNamara, J., 2017. *Transnational Education: A Classification Framework and Data Collection Guidelines for International Program and Provider Mobility (IPPM)*. British Council, London.
- Knight, J., 2008. *Higher Education in Turmoil: The Changing World of Internationalization*. Sense Publishers, Rotterdam and Taipei.
- Knight, J., 2012. *Internationalization: Three Generations of Cross-Border Higher Education*. India International Center, New Delhi.
- Knight, J., 2015. International universities: misunderstandings and emerging models? *J. Stud. Int. Educ.* 19 (2), 107–121.
- Kosmützky, A., Putty, R., 2016. Transcending borders and traversing boundaries: a systematic review on transnational, off-shore, cross-border, and borderless higher education. *J. Stud. Int. Educ.* 20 (1), 8–33.
- Krücken, G., Engwall, L., de Corte, E., 2018. Introduction to the special issue on “university governance and creativity”. *Eur. Rev.* 26 (Suppl. 1), 1–5.
- Kwiek, M., 2018. International collaboration and international research orientation: comparative findings about European academics. *J. Stud. Int. Educ.* 22 (2), 136–160.
- Lanzendorf, U., Teichler, U., 2005. *Student Mobility. European Parliament, Directorate General Internal Policies of the Union 2005*, Brussels (IP/B/CULT/ST/2004_010, PE 361.212, EN).
- Leask, B., de Gayardon, A., 2021. Special issue: reimagining internationalisation for society. *J. Stud. Int. Educ.* 25 (4).
- Leask, B., 2015. *Internationalising the Curriculum*. Routledge, London.
- Marginson, S., 2014. Student self-formation in international education. *J. Stud. Int. Educ.* 18 (1), 6–22.
- OECD (Ed.), 2010. *Higher Education to 2030. Globalisation*, vol. 2. OECD, Paris.
- OECD, 2021. *Education at a Glance: OECD Education Indicators 2021*. OECD, Paris.
- Ogden, A.C., Streitwieser, B., van Mol, C. (Eds.), 2020. *Education Abroad: Bridging Scholarship and Practice*. Routledge, London.
- Opfer, S., Teichler, U., Carlson, J., 1990. *Impacts of Study Abroad Program on Students and Graduates*. Jessica Kingsley Publishers, London.
- Qin, Y., 2021. *Building Cross-Border Joint Universities in China: A Study of Organizational Dilemma*. Routledge, London and New York.
- Reimers, F.M., Marmolejo, F.J. (Eds.), 2021. *University and School Collaborations during a Pandemic*. Springer, Cham.
- Reisberg, L., Rumbley, L.E., 2014. Redefining academic mobility: from the pursuit of scholarship to the pursuit of revenue. In: Maldonado-Maldonado, A., Basset, R.M. (Eds.), *The Forefront of International Higher Education: A Festschrift in Honor of Philip G. Altbach*. Springer, Dordrecht, pp. 115–126.
- Rhoades, G., 2017. Backlash against “others”. *Int. High. Educ.* 89 (2–3).
- Robson, S., Almeida, J., Scharfner, A., 2018. Internationalization at home: time for review and development. *Eur. J. High. Educ.* 8 (1), 19–35.
- Rosa, M.J., Sarrico, C.S., Tavares, O., Amaral, A. (Eds.), 2016. *Cross-Border Higher Education and Quality Assurance*. Palgrave Macmillan, London.
- Rostan, M., Ceravolo, F.A., 2015. The internationalization of the academy: convergence and divergence across disciplines. *Eur. Rev.* 23 (Suppl. 1), 38–54.
- Scott, P., 2015. Dynamics of academic mobility: hegemonic internationalisation or fluid globalisation. *Eur. Rev.* 23 (Suppl. 1), 55–69.
- Scott, P., 2021. *Retreat or Resolution? Tackling the Crisis of Mass Higher Education*. Bristol University Press, Bristol.
- Seeber, M., Cattaneo, M., Huisman, J., Paleari, S., 2016. Why do higher education institutions internationalize? *High. Educ.* 40 (2), 238–253.
- Smith, K., 2010. Assuring quality in transnational higher education: a matter of collaboration or control? *Stud. High. Educ.* 35 (7), 793–806.
- Sursock, A., 2016. The shift to strategic internationalisation approaches. In: Rosa, M.J., Sarrico, C.S., Tavares, O., Amaral, A. (Eds.), *Cross-Border Higher Education and Quality Assurance*. Palgrave Macmillan, London, pp. 73–91.
- Teichler, U., Maiworm, F., 1997. *The ERASMUS Experience: Major Findings of the ERASMUS Evaluation Research Project*. Office of Official Publications of the European Communities, Luxembourg.
- Teichler, U., Ferencz, I., Wächter, B. (Eds.), 2011. *Mapping Mobility in Higher Education in Europe. Volume I: Overview and Trends*. Deutscher Akademischer Austausch Dienst, Bonn.
- Teichler, U., Arimoto, A., Cummings, W.K. (Eds.), 2013. *The Changing Academic Profession: Major Findings of a Comparative Survey*. Springer, Dordrecht.
- Teichler, U., 2003. Mutual recognition and credit transfer in Europe: experiences and problems. *J. Stud. Int. Educ.* 7 (4), 312–341.
- Teichler, U., 2004. The changing debate on internationalisation of higher education. *High. Educ.* 48 (1), 5–26.
- Teichler, U., 2012. Excellence and internationality of higher education. In: Wächter, B., Lam, Q.K.H., Ferencz, I. (Eds.), *Tying It All Together: Excellence, Mobility, Funding and the Social Dimension in Higher Education*. Lemmens Medien, Bonn, pp. 24–56.
- Teichler, U., 2013. The event of international mobility in the course of study: the European policy objective. In: Zgaga, P., Teichler, U., Brennan, J. (Eds.), *The Globalisation Challenge for European Higher Education*. Peter Lang, Frankfurt a. M., pp. 55–78.
- Teichler, U., 2015. Academic mobility and migration: what we know and what we do not know. *Eur. Rev.* 23 (Suppl. 1), 6–37.
- Teichler, U., 2017. Internationalisation trends in higher education and the changing role of student mobility. *J. Int. Mobil.* 5, 177–216.
- Teichler, U., 2019. Bologna and student mobility: a fuzzy relationship. *Innovat. Eur. J. Soc. Sci. Res.* 32 (4), 429–449.
- Van Mol, C., 2014. *Intra-European Student Mobility in International Higher Education Circuits: Europe on the Move*. Palgrave Macmillan, Cheltenham.
- Vögle, E.M., 2019. 20 years of Bologna—a story of success, a story of failure. *Innovat. Eur. J. Soc. Sci. Res.* 32 (4), 406–428.

- Wächter, B., Ferencz, I., 2012. Student mobility in Europe: recent trends and implications of data collection. In: Curaj, A., Scott, P., Vlasceanu, L., Wilson, L. (Eds.), *European Higher Education at the Crossroads: Between the Bologna Process and National Reforms*. Springer, Dordrecht, pp. 387–411.
- Wächter, B., 2006. Brain drain: what we know and what we do not know. In: Teichler, U. (Ed.), *The Formative Years of Scholars*. Portland Press, London, pp. 51–66.
- Wagner, C.S., 2006. International cooperation in science and technology: promises and pitfalls. In: Box, L., Engelhard, R. (Eds.), *Science and Technology Policy for Development: Dialogues at the Interface*. Anthem Press, London, pp. 165–176.
- Yonezawa, A., Kitamura, Y., Meerman, A., Kuroda, K. (Eds.), 2014. *Emerging International Dimensions in East Asian Higher Education*. Springer, Dordrecht.

Global processes: global trade in international education and the international education-migration nexus

Ly Thi Tran and Diep Thi Bich Nguyen, School of Education, Deakin University, Burwood, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	250
Forces underpinning global trade in international education	250
Historical and geopolitical processes	251
Economic processes	251
Commercialization of international education in different countries and impacts	252
International education and skilled migration	254
Conclusion and implications for future research	256
References	256

Introduction

International education has been driven by a range of rationales and aspirations, related to human and institutional capacity building, transnational knowledge and cultural exchange, geopolitical and commercial gains. The degree of global trade in international education has been significantly expanded over the last three decades. In major international education providing countries, currently international education is largely shaped by commercial orientations and often measured in economic terms (Rizvi, 2020; Tran, 2020). Global trade in international education is driven by a web of interrelated historical, geopolitical, and economic processes. As the trade orientation in international education heightens, universities have shifted from merely public entities to corporate enterprises that function in business models (Blackmore and Sawers, 2015; Connell, 2019). The COVID-19 pandemic has exposed the fragility of the transactional higher education model, characterized by increased reliance on revenues from international students to substantially subsidize institutional operations, especially in major receiving countries of international students such as Australia, United States, United Kingdom, Canada and New Zealand.

Key education destinations have actively capitalized on international education as a mechanism to support their nation's economy through its connection with the competition over skills or "global race for talent" (Shachar and Hirschl, 2013). The international education and migration nexus is pushed through by the view that the post-study work and migration prospects are both critical to host countries' destination attraction and responsive to their need to tap into a qualified pool of workforce to fill out areas of skills shortages and support economic growth. The importance of post-study work and migration opportunities in international students' decision about study destination is reinforced by the literature which shows that the tightening of immigration policy in the aftermath of events such as September the 11th in the US and the UK's 2012 decision to remove post study work rights (Kaplan, 2019; Johnson, 2018; Ilieva, 2017) resulted in the decline of international students in periods shortly after the above events. However, there has been a mismatch between policy goals and the actual processes and practices associated with the link between international education and migration, leading to a range of consequences. This situation has resulted in the modifications of international education and skilled migration policies across key destination countries, including the introduction of the post-study work visas that allow international students to remain in the host countries to gain work experiences post-graduation. The past decade has seen the evolution of the international education-migration nexus to the international education-work-migration nexus.

This article provides a critical review of the ways in which international education has been shaped and re-shaped as global trade. First, it outlines the complex historical, geopolitical, and economic processes underpinning international education as global trade. Second, it reviews policies and programs driven by the commercialization of international education in different countries and the impacts. Third, it provides insights into the development of the international education and migration link and the emergence of the international education-work-migration nexus. Fourth, it discusses emergent research agendas, including how international education as trade is manifested in traditional sending countries of international students and how international education has been re-shaped in response to COVID-19.

Forces underpinning global trade in international education

While international education is not a new phenomenon, it is only in the last three decades that it has become part of global trade and seen as a tradable commodity or service across borders (Knight, 2004). Global trade in international education is not driven by one solitary rationale but rather, by a set of complex historical, geopolitical, and economic processes.

Historical and geopolitical processes

Rationales for international education have evolved over different historical periods, from the pursuit of academic and cultural exchange, to include development aid, international cooperation, public diplomacy and commercial motivations (de Wit, 2013; Knight, 2015; Tran and Bui, 2021; Tran et al., 2021). As Guruz (2008) and Welch and Denman (1997) stated, traces of international education dated back to medieval times when scholars “wandered” across the world in search of knowledge and to conduct enquiry-based work. During colonial time, cross-border education was fueled by the idea of enlightenment whereby colonizers used education as a tool to equip the local elite of the colonized nations with Western knowledge, skills and attitudes to meet the economic and geopolitical interests of colonial powers (Rizvi, Forthcoming).

In the post-independence period until early 1980s, international education served the developmental and aid purposes as developed countries provided technical support and knowledge transfer to assist with capacity building for newly independent nations (de Wit, 2013; Rizvi, 2011). Geopolitical drivers also underpinned international education program during this period both in the West and among the socialist countries, which could now be seen as a mechanism of soft power development. Examples of education aid include the Colombo Plan by Australia, Fulbright Program by the US and development programs for socialist countries by the Soviet Union. While these programs took the form of developmental undertaking, they were at the same time believed to serve the political purposes of aid-providing countries as international education aid was used as a means of soft power and public diplomacy for aid providers to influence aid recipients. For example, the Australian Colombo Plan and the US Fulbright Plan were designed to keep developing countries from falling into the communist bloc (Rizvi, 2011). Therefore, development needs were taken into account but so were the benefits of serving geopolitical purposes and building relationships between states, many of which were former colonial powers, and elites in postcolonial states. While these international education as aid programs assisted with human capital building for aid-receiving countries and fostered long-lasting people-to-people connections and country-to-country ties, international education during this period was also seen to represent a new form of colonial practice that reinforced power inequalities across the globe (Escobar, 1995). Further to that, the development intent underpinning these forms of education aid was not fully achieved when a proportion of international students from aid-receiving countries decided not to return home, causing brain drain for these countries (Rizvi, 2005). This was accompanied by financial pressures experienced by aid providers who no longer had the capacity to support poorer countries (Rizvi, 2011; Lenn, 2000). Subsequently, by the mid-1980s, the developmentalist ideology of international education lost its grounds, giving way to new rationales of international education. The shift was from aid to academic exchange and curriculum development in continental Europe, as compared to the move from aid to trade in countries such as the UK and Australia (de Wit, 2013) that will be discussed further in later parts of this chapter.

Economic processes

In addition to the historical and geopolitical drivers, the shift of focus to trade in international education has been shaped and strengthened by market forces, liberalization of trade and neoliberal ideologies.

Market forces

International trade in education services is driven by market forces on the demand-supply basis (Ziguras, 2005). With regard to program mobility, institutions in the Global North provide education services and in return gain financial profits and international branding from transnational education. Meanwhile, foreign partners, predominantly from the Global South receive international education, largely to build the capacity of their education systems and national economies (Altbach and Knight, 2007). This reflects the push-pull model borrowed from international trade (Choudaha and de Wit, 2014; Maringe and Carter, 2007). In turn, international student mobility can be seen as an outcome of supply and commercialization orientation from providers in the Global North and a response to students’ demand and growth of middle-class families from the Global South countries (Bamberger et al., 2019; Tran, 2020). That is, students tend to study in countries other than their own where the supply is perceived to be greater and their demand could be met. The key change leading to the commercialization of international education was that students were able to travel more easily, that more elites in postcolonial states could afford to study in high-income countries, and that many developed countries realized this growing demand and decided to capitalize on it by allowing their institutions to recruit massive number of international students and charge fees.

Liberalization of international trade

Commercialization and commodification of international education have been facilitated by liberalization of trade in education services that was primarily catalyzed by free trade agreements. The General Agreement on Trade in Services (GATS) of the World Trade Organization coming into force in 1995 has lowered barriers to movement, enhanced the transparency and measures affecting trade and therefore created global markets for goods and labor (Bamberger et al., 2019; Ziguras, 2005). Accordingly, under GATS, education—traditionally seen as a “public good” and a “social responsibility” (Knight, 2008)—has been identified as an international commodity or service to be traded on commercial basis and governed by international trade rules (Knight, 2015). Through GATS, common educational standards, mutual recognition and liberalization of the processes by which professionals are permitted to practice have been developed. A notable feature of GATS is that this agreement seeks to create more freedom for private and foreign providers and equal treatment between foreign and domestic providers (Knight, 2004, 2015; Ziguras, 2005). That is, once a foreign supplier has been allowed to supply a service, education included, there should be no discrimination in treatment from domestic providers in that country.

GATS has paved the way for other trade agreements at the regional, plurilateral and bilateral levels within and outside the scope of the World Trade Organization. Regional trade agreements include the Trans-Pacific Partnership (TPP) involving 12 countries in the Pacific Region dominated by the US; the EU-Canada Comprehensive Economic and Trade Agreement (CETA) between Canada and Europe; the Transatlantic Trade and Investment Partnership (TTIP) involving Europe and the US, and the Trade in Services Agreement (TISA) largely between OECD countries. Plurilateral agreements were in force between, for example, ASEAN, Australia and New Zealand. Examples of bilateral free trade agreements are between Singapore and the US, between Europe and South Korea (Robertson, 2017). Australia signed the first bilateral agreements with New Zealand in 1983, and most recently Indonesia in July 2020 (DFAT, 2021). Similar to GATS, these regional, plurilateral and bilateral free trade agreements facilitate international trade in education services, and at the same time expose the education sector to privatization and commercialization, which introduce opportunities and at the same time pressures for both developed and developing countries.

Global, regional and bilateral trade agreements are based on principles of non-discriminatory treatment for service providers. On the one hand, these create new opportunities for cross-border mobility in education and benefits for education institutions, including diversity in sources of funding including cost-sharing with students, and free entry of foreign capital and investment in education (Gupta, 2015). On the other hand, many fear that free trade agreements introduce risks to the idea of education as a public good, the capacity of governments to shape the cultural and social dimensions of education, and the role of public education providers (Ziguras, 2005). Specifically, the concern is that free trade measures will weaken or remove the power of the state as a key provider of public education and its role in determining national education policies. In the same line, functions of education as nation building and civic development might be hindered by free trade. Also, increased reliance on private and foreign providers will threaten the existence of public institutions when they lose local students and staff, creating a two-tier education system in developing countries.

Neoliberal ideologies

More broadly, increased trade in international education is associated with the larger ideological shift to neo-liberalism (Yang, 2006; Rizvi, Forthcoming). As Bamberger et al. (2019) state, cross-border mobility is portrayed by the desire of individuals, institutions, countries to “maximize their gains through employment, global rankings or a global competition” (p. 208). In particular, countries and institutions providing international education view it as a source of revenue generation for national economy and education sector. Within higher education, both government and universities see income from international education as a substitute source of revenue in place of public funding cut. Such economic motivations trigger nations and universities to boost international student recruitment and therefore make efforts to build their international branding and reputation (Lomer et al., 2018). Likewise, faculty mobility is not purely for academic or humanistic purposes but it also aims at seeking international research funding, boosting transnational research partnerships and performance (Woldegiyorgis et al., 2018), and increasing the number of co-authored publications, which are among key indicators in international university rankings (Wan et al., 2017). With these motivations, countries and institutions now take a more competitive approach to international education rather than a cooperative one (de Wit, 2013; Knight, 2015). For individual students, mobility is closely linked to the quest for global skills, competencies, and social networks (Bamberger et al., 2019). These are then transferred into increased economic gains for individuals, for example, in relation to enhanced employability and migration opportunities.

Critics point out a number of tensions and risks facing countries, institutions and students as a result of a neoliberal approach to international education. First, international students are positioned as “cash cows” when too much emphasis is placed on revenue generation from international student recruitment (De Vita and Case, 2003; Lomer, 2014). As such, universities are transformed into international enterprises operating under a transactional model (Sidhu, 2009; Wu and Naidoo, 2016). Second, sending countries of international students face the risk of brain drain when “talented” students choose to stay in host countries instead of returning home after study. Third, there might be an issue with transnational education with regard to the substandard quality of branch campuses and delivery of overseas education (Altbach, 2010; Lieven and Martin, 2006).

Commercialization of international education in different countries and impacts

The volume and value of international trade in education have increased significantly over the recent decades. The top providing countries include the USA, UK and Australia, among which the USA currently ranks first while UK and Australia take the second and third positions respectively. In 2019, in higher education sector alone, there were 851,957 international student enrollments in the US (Open Doors, 2020), as compared to 485,645 international enrollments in the UK (HESA, 2020), and 442,219 international students in Australia (DESE, 2019). In these countries, the traditional form of international education is delivered to onshore international students. Recently, the emerging form of international education is transnational education or also known as offshore education (Ziguras, 2005), in which students are not required to leave their countries. Education providers might also be based in their own countries when delivering education overseas with the assistance of digital technology.

The US is currently the largest provider of international education with respect to the inflow of international students and their offshore activities. Particularly, the US ranks first in terms of receiving international students pursuing a full degree (de Wit, 2013). Over its long history, the US education system was driven and characterized by an inward-looking orientation, leading to isolationism among American institutions. This is because “Americans saw themselves as ‘the best’ and took pride in the belief that everyone else really yearned to follow the American example” (Hudzik, 2011, p. 14). The post-World War II period saw the US move away

from simplistic isolationism as a geopolitical strategy (Hudzik, 2011). However, the US still saw itself as being the best and took comfort in inwardly derived strengths. Therefore, international education in the US is not a two-way exchange insofar as the US would involve itself in teaching, but not so much in learning. The period from the mid-1970s to the mid-1980s marked the beginning of large-scale overseas recruitment efforts by many US institutions, and the sector experienced considerable growth in the 1980s and 1990s (Dunnet, 2013). Despite a sharp decrease in international recruitment after the 9/11 event, international enrollments at many institutions have increased substantially until now. As part of the effort to expand international enrollments, there has been an increase in the use of commission-based third-party recruitment agents by US institutions. Dunnet (2013) has pointed out that this has led to issues such as some unethical recruitment practices, incidences of gross misrepresentation of American education overseas, and in some cases violations of US immigrations laws.

In the UK, commercialization of international education was marked by the introduction of full fees for international students in 1980 under the Thatcher government (de Wit, 2013; de Wit and Altbach, 2020). From that time onward, international education has been treated as an exporting industry, with the main focus being recruitment of full fee paying international students. Another major feature of education export, which has been expanded rapidly, is transnational education. Both international student recruitment and transnational education have been identified to be the key dimensions of international education export in the UK due to the economic values, as stated in a government policy document in 2013:

Education exports are defined as those activities where money comes to the UK from an overseas source, either for an education-related activity taking place in the UK (e.g. international students studying at a UK HEI) or from an education-related activity occurring overseas (e.g. TNE).

BIS (2013, p. 21)

Worton (2012) even projected that UK transnational education would grow more significantly and rapidly than onshore provision. Accordingly, economic values generated from both forms of education exports include not only tuition fees but also living costs, language costs and related ancillary services.

Following the UK, Australia experienced the transition from “aid” to “trade” in the late 1980s, driven partly by concerns about the effectiveness of the sponsored overseas student program but more particularly due to the recognition of the commercial potential in providing international education (Harman, 2005; Rizvi, 2011). The “trade” phase was characterized by the substantial growth in international students across all sectors following the introduction of the “Overseas Student Policy” in 1985, which introduced the implementation of full-fee programs, allowing Australian universities to charge international students full fees (de Wit and Adams, 2010; Marginson, 1995). This was coupled with the Commonwealth government policy reforms (often referred to as the Dawkins reforms) and the *Higher Education: A Policy Statement 1988*, which marked the decrease in direct government funding from the Commonwealth government for HE (Tran, 2020). These policy changes have paved the way for the business model of Australian universities, driven by market-based rationales (Rizvi, 2020; Tran, 2020).

The current Australian international education is still deeply shaped by a commercial rationale. Prior to the COVID-19 pandemic, international education was the third largest export sector of the country that contributed nearly AU\$40 billion to the Australian economy in 2019 (ABS, 2020). Major source countries providing international students for Australia prior to COVID-19 include China (accounting for 28% of the total international students), India (15%), followed by Nepal (7%), Brazil (4%), Vietnam (3%), Malaysia (3%), South Korea (3%), Colombia (3%), Indonesia (2%), and Thailand (2%) (IEAA, 2020). Nowadays, Australian universities heavily depend on full-fee-paying international students as a major source of revenue to substantially subsidize research and institutional operations, which has placed the Australian HE sector in a biggest financial crisis during COVID-19 (Blackmore, 2020; Rizvi, 2020).

Over the recent decades, New Zealand and Canada have also moved toward a market approach to international education (de Wit and Altbach, 2020). Likewise, although continental Europe has maintained a cooperative model of international education during the past 25 years, they have switched to commercialization of international student recruitment (de Wit and Altbach, 2020).

The shift to commercial approach of international education in major providing countries indicates their growing dependence of education institutions, particularly universities, on international students as a significant financial source. This has led to the higher level of global competition in the space of branding and marketing in search for international students (Marginson and Rhoades, 2002; Ramírez, 2018). In the context of such heightened competition, universities are no longer considered as merely public entities but multinational corporations that function in business models (Blackmore and Sawers, 2015; Connell, 2019). That is, entrepreneurial goals seem to be prioritized over academic and intellectual foci. Further, this entails risks for the sustainability of the education sector and national economy of the providing countries more broadly. Particularly, COVID-19 has revealed the vulnerability of many major providers of international education, such as the US, UK and Australia when losing a large proportion of international students due to border closing and the fear among students and their families about the uncertainties in host countries. Whereas, supplying countries have identified new opportunities to expand international education at home to cater for the needs of their domestic students (Phan and Phung, 2020).

From humanitarian perspectives, the focus on profit-making over international student fees has contributed to inequality in access to international education. The cut of education aid through scholarship and financial assistance schemes, and high fees applicable to international students mean that international education is more accessible for a small proportion of privileged

students. This has raised ethical concerns about the inclusivity of international education and the increased gap between students who can gain access and those who do not (de Wit, 2016). Further to this, increased trade has reinforced commodification of international education in which education services are framed as products and international students are defined as consumers. As such, international students are valued in economic terms rather than according to their other contributions to host countries' universities, culture and society (Tran, 2020). In this corporate model, academic staff are expected to play a significant role in increasing institutional reputation and ranking, which has put heightened level of pressure on their shoulder (Blackmore and Sawers, 2015).

While Anglophone and European countries are known as main exporters of international education services, middle-income and low-income countries from the developing world are positioned as typical importers (Altbach and Knight, 2007). These countries are major sources of international students, importers of twinning and franchised programs, and also borrowers of the curriculum from more advanced systems to deliver in their at-home internationalized programs. Although these dimensions of international education contribute to enhancing the quality of education in importing countries, concerns are raised in relation to brain drain as a result of outbound student mobility, as well as Westernization and neo-colonialization due to curriculum importation with little adaptation to local contexts (Phan, 2017; Tran and Marginson, 2018; Tran et al., 2018).

International education as global trade is increasingly politicalized. The politicalization of international education over the past decade has been shaped by ideologies and movements such as populism, national protectionism or inward-looking nationalism, *anti-immigration*, *anti-globalization* and *anti-multiculturalism* (Altbach and de Wit, 2017; Hsieh, 2020). Brexit, the Trump isolationism, and the rise of far-right parties in Europe have substantially influenced international education programs and strategies, including inbound and outbound mobilities, international student recruitments, transnational research collaboration and transnational education partnerships and internationalization at home in many countries. International education has also been used as a mechanism to boost soft power influences by a range of countries. China, for example, is very active in increasing its soft power and its position regionally and globally through aggressive strategies to attract international students to China, open up offshore campuses and promote the country as a regional education hub (Cheng, 2009). Lifting China's destination attraction and soft power as part of China's political agenda has been accelerated through the Silk Road Scholarship programs under the Belt and Road Initiative (Flint and Zhu, 2019). In addition to China, for the last decade, many traditional importing countries or regions such as Hong Kong, Malaysia, Singapore, Turkey, and Mexico have emerged as new players in the international education market and become active competitors with Western systems (Dunnet, 2013; Wen and Hu, 2019). These countries have begun to promote themselves as regional and highly affordable hubs for world class education, thereby attracting an increasing population of international students. The development of education hubs by these countries is underpinned by nation states' desire to strengthen leadership position in international education so as to exert their geopolitical influence (Lee, 2015).

International education and skilled migration

Student mobilities as one of the prominent dimensions of international education are often linked to the development of human capital and transnational knowledge, ideas and skills mobilities. The link between international education and skilled migration has been seen as a win-win policy for many host nations to support both the trade orientation of international education and the attraction of a qualified pool of workforce to respond to areas of skills shortages and boost economic growth.

The emergence of the international education and migration nexus

The nexus between international education, skills development and skilled migration is shaped by the desires and goals at individual, institutional and government levels. Research evidence suggests that students' motivations to pursue international education can include securing skilled migration to the host country and developing skills and knowledge that can enhance their employment prospects, future life, and/or personal and intercultural transformation (Tran and Nyland, 2011; Tran, 2016; Tran and Nguyen, 2016; Tran et al., 2022). At the institutional level, student mobilities are situated within the broader internationalization agenda which aims to generate revenues, strengthen international partnerships and internationalize the curriculum, teaching and learning to enrich the student experience and enhance graduate employability. At the government level, many host nations see student mobilities as a source of export income, a vehicle for public diplomacy and a potential supply of future skilled migrants (Tran and Vu, 2018; Ziguras and Law, 2006). Viewed from the lens of many host nations, international students are situated at the intersection of skills production and skills attraction (Mosneaga, 2014). On the one hand, they are seen as students who aspire to accumulate skills, knowledge and experiences through an international education. On the other hand, they are positioned as a potential pool of skilled migrants that host countries are seeking to tap into (Riaño et al., 2018; Tremblay, 2005). The recent message from Canada's Minister of Immigration, Refugees and Citizenship demonstrates this perspective: "These new policies will help those with a temporary status to plan their future in Canada, play a key role in our economic recovery and help us build back better. Our message to them [international students] is simple: your status may be temporary, but your contributions are lasting—and we want you to stay" (Government of Canada, 2021).

Key destination countries often revise their international education and skilled migration policies. These regular modifications respond to changes in the broader context of international education and the social, economic and geopolitical circumstances of the host countries as well as to the intended and unintended consequences of the international education-migration nexus. Nevertheless, they cause complexities and uncertainties for current and prospective international students. In Australia, the direct link of international education and migration was facilitated through the introduction of the government policy in July 2001 which

encouraged international students with skills in demand who completed their course at an Australian institution, and met general eligibility requirements, to apply for permanent residency on shore (Spinks, 2016). This skilled migration pipeline from international education was supported through the Skilled-Independent category of the General Skilled Migration (GSM) program. In May 2005 the number of trades occupations listed on the Migration Occupations in Demand List (MODL), including cooking and hospitality, was expanded (Spinks, 2016), leading to the increase in international student enrollments in vocational education courses.

However, the direct link between international education and migration caused a range of consequences. Some colleges in the private vocational education and training (VET) in Australia were charged with promoting VET courses to international students as a pathway to permanent residency (PR) rather than an education opportunity and were producing poorly trained graduates (Mariginson et al., 2010; Perkins, 2009; Smith, 2010). In response to this situation, the Australian government modified the General Skilled Migration Scheme leading to the restriction in graduates' ability to secure permanent residency and the requirement of independent skills testing of graduates (Tran and Nyland, 2011).

Existing literature suggests that the popular generalization of all international VET students as "PR hunters" who have little or no interest in pursuing a high-quality education have had adverse impacts on both international students and the VET sector (Tran and Nyland, 2011). This popularization of international VET students as those who are only interested in migration largely ignores their aspirations and efforts put in learning, development of skills, knowledge and experiences, and enhancement of cultural, personal and professional attributes. The discourse promoting the "international VET students simply want to migrate" perspective has led to destructive attitudes toward this cohort and negatively impacted their learning, wellbeing and connectedness with institutional community, the workplace and the broader Australian community (Tran and Vu, 2016). The VET sector as a whole has encountered "imposed and/or self-imposed" consequences of this stereotype since "without the protectionism provided by the migration-education link," coupled with the destructive stereotype of international students as mere "PR hunters" and many VET courses as "PR factories", the sector struggles to "market courses that have been devalued and stereotyped as low quality" (Tran and Nyland, 2011, p. 28).

Decoupling the direct international education and migration link and the emergence of the post-study work rights policy

Post-graduate work policy is considered as a strategic mechanism to attract international students in key destination countries, including Australia, Canada, New Zealand and European countries such as Germany, Ireland, Sweden, the Netherlands and the UK. In an effort to decouple the direct international education and migration link while still maintaining Australia's destination attractiveness, the Australian government introduced the revised post-study work rights policy and the Temporary Graduate visa (subclass 485) in 2013. This move followed the recommendations from the 2011 Knight Review which evaluated the quality, integrity and competitiveness of Australia's international student visa program (Tran et al., 2019). The Knight Review highlighted that the introduction of an expanded post-study work visa was crucial to enhance Australia's attractiveness as a study destination. The 485 Temporary Graduate visa comprises two streams: the Graduate Work stream and the Post-Study Work stream. Under the Graduate Work stream, international graduate visa holders are allowed to remain in Australia for 18 months while the Post-Study Work stream allows holders with a Bachelor, Masters or Doctorate degree to live and work in the country between 2 and 4 years, after the completion of their studies. In addition to the difference in terms of the visa duration, applicants for the Graduate Work visas need to nominate an occupation on the Medium and Long-term Strategic Skills List and have their skills assessed by the assessing body while this is not required for international graduates applying for the Post-Study Work stream (Australian Government, 2019a,b). In the UK, the post-study work rights policy was originally introduced in 2004 but removed by then Home Secretary Theresa May in 2012. As a move to regain its destination competitiveness, in October 2019, the UK government announced the re-introduction of a 2-year post-graduation work visa policy (UKCISA, 2019).

Research on the impacts of the post-study work policy in Australia shows that up to 76% of 1150 international graduates in a survey indicated that access to post-study work rights was an important factor influencing their decision to choose Australia as a study destination (Tran et al., 2019). In addition, the top five citizenship countries of temporary graduate visa holders in Australia (India, China, Nepal, Pakistan, Vietnam) have mirrored the top five source countries (China, India, Nepal, Pakistan, Vietnam) of international enrollments in Masters by coursework programs since 2013 (Tran et al., 2019). However, in reality, there seems to be a mismatch between international students' intent to stay in the host country post-study and the actual uptake (Berquist et al., 2019). While around 60–80% of international students surveyed express their desire to remain in the host country after graduation, statistics show that only around 25% actually stay for a long term in the OECD countries (Berquist et al., 2019; OECD, 2011).

Despite the introduction of the post-study work rights policy being welcome by education providers and its effectiveness in strengthening host countries' attraction as a study destination, the work experiences of international graduates on these temporary graduate visas do not seem satisfactory. A survey with 1156 international graduates on temporary graduate visas in Australia shows that only 36% of current and past holders of the visas who stayed in Australia secured full-time employment in their field of study (Tran et al., 2019). International graduates on the temporary post-study work visas often face significant challenges and structural barriers in gaining a foothold in the host labor market. Employers in Australia (Blackmore et al., 2014; Tran et al., 2020a,b), Canada (Flynn and Arthur, 2013) and the US (Monahan, 2018; Sangganjanavanich et al., 2011) reportedly hesitate to recruit international graduates due to the temporary status of their visas, discrimination and prejudices about their cultural skills and ability to adapt to the workplace, their level of familiarity with the host environment, their English proficiency and their ethnicity. In particular, a significant proportion of international graduates who seek post-study work rights opportunities are on the 2-year visas, which is seen as being too short to provide employers a sense of security and confidence in recruiting this cohort (Tran et al., 2019). The recruitment practice in some host countries is heavily shaped by the "best fit" principle that often disadvantages international

students and graduates as this group are considered as “outsiders” who do not possess the “needed” abilities and qualities to fit in “our” local workplace (Blackmore and Rahimi, 2019; Tran and Bui, 2019). In addition, there is a lack of understanding among employers in Australia about the purpose and nature of the temporary graduate visas (Tran et al., 2020a,b).

While permanent residency is a significant factor enhancing international graduates’ access to the Australian labor market, the temporary graduate visa does not provide them with a substantial competitive advantage. Instead, the temporary graduate visas give them extra time to remain in the host country to develop their social and professional networks, enhance their English language proficiency, gain some work experience and get some return on investment in their international study (Tran et al., 2019, 2020a,b). However as international graduates struggle to get foot in the door, especially in their field of study, they are facing a risk of exploitation, deskilling, precarity, financial stress and vulnerability. Many desperately need work in the host country to pay for living expenses, pay back their loans, achieve their career goals, and acquire permanent residency (Tran et al., 2020a,b).

International graduates who remain in the host country on post-graduation visas are important actors who help further the commercialization of education in the context of global trade but are subject to various injustices. In Australia, while international graduates on temporary post-study work visas constitute around 0.7% of Australia’s labor force and contribute taxes to the Australian economy, they are not entitled to subsidized government services due to their non-citizenship status (Tran and Bui, 2019). In rhetoric, key destination countries such as Australia, Canada, New Zealand and the UK position the post-study work rights policy as providing local businesses and industry with an opportunity to access a pool of highly qualified human resource who are educated in the host countries and possess multilingual capabilities, global skills, international experiences and transnational networks (Tran and Bui, 2019). Nevertheless, in reality, international students and graduates are marginalized and disadvantaged in the host labor market due to structural factors including existing prejudices toward these cohorts, the length and conditions of their visas and the lack of understanding of these cohorts as well as the temporality of the visa they are on.

Conclusion and implications for future research

Over the past three decades, international education has been shaped and re-shaped into a form of global trade, driven by historical, economic and geopolitical rationales. This re-positioning of international education as global trade has been pursued through policies and programs across a range of dimensions, including inbound and outbound student mobilities, international student recruitment, internationalization of the curriculum, international research collaborations and transnational education and offshore programs. The commercialization of international education has also been increasingly politicalized and associated with “a scalar geopolitics linking the education of the individual to global political relations” (Brooks, 2015, p. 65).

Research and debates in international education have raised concerns about the international education-migration nexus. Designed as a win-win policy for some key destination countries, this nexus has resulted in some unintended consequences, leading to the introduction of the international education-work-migration nexus. Within this emergent nexus, the post-study work opportunities have been a drawcard for Australian universities to attract international students (Tran et al., 2020a,b). However, there are a range of tensions and paradoxes related to the reality of how international students and graduates are positioned within this nexus. Due to structural barriers, including visa regimes and prejudices in the host labor market, destination countries like Australia have yet to tap on this pool of young and highly qualified workers to deliver benefits to its economy (Chew, 2019; Tran et al., 2019; Tran and Bui, 2019).

While there is considerable research describing the ways in which international education has become commercialized and the international education-migration link, the following topics related to international education in the global trade are less understood:

- Whether and how traditional sending countries have approached international education as trade and how international education as trade is manifested in these countries
- The impacts of global trade in international education on pedagogies and curriculum and student experiences
- How students can be engaged to capitalize on the academic and humanistic values of international education while minimizing the negative impacts of the commercial orientation of international education
- The role of students as partners and consumers in (re)defining international education services
- The nature and characteristics of the international education-work-migration
- The effectiveness of post-study work rights policies across different destination countries
- The ways in which the relationship between geopolitics and international education been re-shaped during and post-COVID-19
- The impacts of escalated geopolitical turbulences during and post COVID-19 on different aspects of international education and the landscape of international education marketplace.

References

- ABS, 2020. International Trade in Goods and Services. Australian Bureau of Statistics. <https://www.abs.gov.au/statistics/economy/international-trade/international-trade-goods-and-services-australia/latest-release>. (Accessed 5 March 2021).
- Altbach, P., 2010. Why branch campuses may be unsustainable. *Int. High. Educ.* 58, 2–3. <https://doi.org/10.6017/ihe.2010.58.8467>.
- Altbach, P.G., de Wit, H., 2017. Trump and the coming revolution in higher education internationalization. *Int. High. Educ.* (89), 3–5.

- Altbach, P.G., Knight, J., 2007. The internationalization of higher education: motivations and realities. *J. Stud. Int. Educ.* 11 (3/4), 290–305. <https://doi.org/10.1177/1028315307303542>.
- Australian Government, 2019a. Temporary Graduate Visa (Subclass 485): Post-study Work Stream. Department of Home Affairs. Retrieved from. <https://immi.homeaffairs.gov.au/visas/getting-a-visa/visa-listing/temporary-graduate-485/post-study-work#About>. Department of Home Affairs. 2019a.
- Australian Government, 2019b. Temporary Graduate Visa (Subclass 485): Graduate Work Stream. Department of Home Affairs, Canberra. Retrieved from. <https://immi.homeaffairs.gov.au/visas/getting-a-visa/visa-listing/temporary-graduate-485/graduate-work#Eligibility>.
- Bamberger, A., Morris, P., Yemini, M., 2019. Neoliberalism, internationalization and higher education: connections, contradictions and alternatives. *Discourse* 40 (2), 203–216. <https://doi.org/10.1080/01596306.2019.1569879>.
- Berquist, B., Hall, R., Morris-Lange, S., Shields, H., Stern, V., Tran, L.T., 2019. Global Perspectives on International Student Employability. International Education Association of Australia (IEAA). Retrieved from. <https://www.ieaa.org.au/documents/item/1586>.
- BIS, 2013. International Education: Global Growth and Prosperity. Department for Business, Innovation and Skill, London.
- Blackmore, J., Rahimi, M., 2019. How "best fit" excludes international graduates from employment in Australia: a Bourdeusian perspective. *J. Educ. Work* 32 (5), 436–448.
- Blackmore, J., Sawers, N., 2015. Executive power and scaled-up gender subtexts in Australian entrepreneurial universities. *Gend. Educ.* 27 (3), 320–337. <https://doi.org/10.1080/09540253.2015.1027670>.
- Blackmore, J., 2020. The carelessness of entrepreneurial universities in a world risk society: a feminist reflection on the impact of Covid-19 in Australia. *High. Educ. Res. Dev.* 39 (7), 1332–1336. <https://doi.org/10.1080/07294360.2020.1825348>.
- Blackmore, J., Gribble, C., Farrell, L., Rahimi, M., Arber, R., Devlin, M., 2014. Australian International Graduates and the Transition to Employment: Final Report. Deakin University, Melbourne.
- Brooks, C., 2015. The apostle of internationalism: Stephen Duggan and the geopolitics of international education. *Polit. Geogr.* 49, 64–73.
- Cheng, X., 2009. Education: the intellectual base of China's soft power. In: Li, M. (Ed.), *Soft Power: China's Emerging Strategy in International Politics*. Lexington Books. ProQuest Ebook Central.
- Chew, J., 2019. Economic Opportunities and Outcomes of Post-study Work Rights in Australia. Retrieved from. www.ieaa.org.au.
- Choudaha, R., de Wit, H., 2014. Challenges and opportunities for global student mobility in the future: a comparative and critical analysis. In: Steitwieser, B (Ed.), *Internationalisation of Higher Education and Global Mobility*. Symposium Books Ltd, Oxford, UK, pp. 19–33.
- Connell, R.W., 2019. *The Good University: What Universities Actually Do and Why It's Time for Radical Change*. ZED Books, London.
- de Wit, H., 2016. Internationalisation should be ethical and for all. *University World News*. <https://www.universityworldnews.com/post.php?story=20160712085821857>. (Accessed 15 February 2020).
- de Wit, H., Adams, T., 2010. Global competition in higher education: a comparative study of policies, rationales, and practices in Australia and Europe. In: Portnoi, L.M., Rust, V.D., Bagley, S.S. (Eds.), *Higher Education, Policy, and the Global Competition Phenomenon*. Palgrave Macmillan, New York, pp. 219–233.
- de Wit, H., Altbach, P.G., 2020. Internationalization in higher education: global trends and recommendations for its future. *Pol. Rev. High. Educ.* 2020, 1–19. <https://doi.org/10.1080/23322969.2020.1820898>.
- De Vita, G., Case, P., 2003. Rethinking the internationalisation agenda in UK higher education. *J. Furth. High. Educ.* 27 (4), 383–398. <https://doi.org/10.1080/0309877032000128082>.
- de Wit, H., 2013. Internationalization of higher education, an introduction on the why, how and what. In: de Wit, H. (Ed.), *An Introduction to Higher Education Internationalization*. Center for higher education internationalization, Università Cattolica Del Sacro Cuore, Milan, pp. 13–46.
- DESE, 2019. International Student Monthly Data December 2019. Retrieved from. <https://internationaleducation.gov.au/research/International-Student-Data/Documents/MONTHLY%20SUMMARIES/2019/Dec%202019%20MonthlyInfographic.pdf>.
- DFAT, 2021. Free trade agreements in force. Department of Foreign Affairs and Trade. <https://www.dfat.gov.au/trade/agreements/in-force/free-trade-agreements-in-force>. (Accessed 1 October 2021).
- Dunnet, S.C., 2013. A brief history of international recruitment in U.S. higher education. In: de Wit, H. (Ed.), *An Introduction to Higher Education Internationalization*. V&P Vita E Pensiero, Milan.
- Escobar, A., 1995. *Encountering Development*. Princeton University Press, Princeton NJ.
- Flint, C., Zhu, C., 2019. The geopolitics of connectivity, cooperation, and hegemonic competition: the belt and road initiative. *Geoforum* 99, 95–101. <https://doi.org/10.1016/j.geoforum.2018.12.008>.
- Flynn, S., Arthur, N., 2013. International students' views of transition to employment and immigration. *Can. J. Career Dev.* 12 (1), 28–37.
- Government of Canada, 2021. New Pathway to Permanent Residency for Over 90,000 Essential Temporary Workers and International Graduates. Media Release. Retrieved from. <https://www.canada.ca/en/immigration-refugees-citizenship/news/2021/04/new-pathway-to-permanent-residency-for-over-90000-essential-temporary-workers-and-international-graduates.html>.
- Gupta, T., 2015. The Perils of Commercialization of Higher Education. *Int. J. Res. Eng. Technol. Sci.* 5 (10), 1–8.
- Guruz, K., 2008. *Higher Education and International Student Mobility in the Global Knowledge Economy*. University of New York Press, Albany NY.
- Harman, G., 2005. Internationalization of Australian higher education: a critical review of literature and research. In: *Internationalizing Higher Education*. Springer, pp. 119–140.
- HESA, 2020. Higher Education Student Statistics: UK, 2018/19. Retrieved from. <https://www.hesa.ac.uk/news/16-01-2020/sb255-higher-education-student-statistics>.
- Hsieh, C.C., 2020. Internationalization of higher education in the crucible: linking national identity and policy in the age of globalization. *Int. J. Educ. Dev.* 78, 102245.
- Hudzik, J.K., 2011. Comprehensive Internationalization: From Concept to Action. NAFSA: Association of International Educators, Washington, D.C.
- IEAA, 2020. International Students by Numbers. Retrieved from. <https://www.ieaa.org.au/documents/item/1151>.
- Ilieva, J., 2017. Do Political Events in a Host Country Affect International Education Engagement? Paper Presented at the AIEC Conference. Hobart: Australia. Retrieved from. <https://aiec.idp.com/aiec2017-past-janet-ilieva>.
- Johnson, K.A., 2018. 9/11 and international student visa issuance. *J. Stud. Int. Educ.* 22 (5), 393–413.
- Kaplan, 2019. The UK's Tax Revenues From International Students Post-graduation. London. <https://www.hepi.ac.uk/wp-content/uploads/2019/03/The-UK-tax-revenues-from-international-students.pdf>.
- Knight, J., 2004. Cross-border education as trade: issues for consultation, policy review and research. *J. High. Educ. Afr.* 2 (3), 55–81.
- Knight, J., 2008. The role of crossborder education in the debate on education as a public good and private commodity. *J. Asian Public Policy* 1 (2), 174–188.
- Knight, J., 2015. Meaning, rationales and tensions in the internationalization of higher education. In: McGrath, S., Gu, Q. (Eds.), *Routledge Handbook of International Education and Development*. Routledge, pp. 325–339.
- Lee, J.T., 2015. The regional dimension of education hubs: leading and brokering geopolitics. *High. Educ. Pol.* 28 (1), 69–89.
- Lenn, P.M., 2000. Higher education and the global marketplace: a practical guide to sustaining quality. *Horizon* 8 (5), 7–10. <https://doi.org/10.1108/10748120010803555>.
- Lieven, M., Martin, G., 2006. Higher education in a global market: the case of British overseas provision in Israel. *High. Educ.* 52 (1), 41–68. <https://doi.org/10.1007/s10734-004-6374-9>.
- Lomer, S., 2014. Economic objects: how policy discourse in the United Kingdom represents international students. *Policy Futures Educ.* 12 (2), 273–285. <https://doi.org/10.2304/pfie.2014.12.2.273>.
- Lomer, S., Papatsiba, V., Naidoo, R., 2018. Constructing a national higher education brand for the UK: positional competition and promised capitals. *Stud. High. Educ.* 43 (1), 134–153. <https://doi.org/10.1080/03075079.2016.1157859>.

- Marginson, S., 1995. Markets in higher education: Australia. In: *Academic work, the changing labour process in higher education*. Open University Press and The Society of Research into Higher Education, Bristol, pp. 17–39.
- Marginson, S., Nyland, C., Sawir, E., Forbes-Mewett, H., 2010. *International Student Security*. Cambridge University Press, Melbourne.
- Marginson, S., Rhoades, G., 2002. Beyond national states, markets, and systems of higher education: a glonacal agency heuristic. *High. Educ.* 43 (3), 281–309. <https://doi.org/10.1023/a:1014699605875>.
- Maringe, F., Carter, S., 2007. International students' motivations for studying in UK HE: insights into the choice and decision making of African students. *Int. J. Educ. Manag.* 21 (6), 459–475. <https://doi.org/10.1108/09513540710780000>.
- Monahan, 2018. *Starting a Career in the U.S.: International Student Perceptions of the Optional Practical Training Experience*. PhD thesis. University of Missouri-Columbia.
- Mosneaga, A., 2014. Student migration at the global trijunction of higher education, competition for talent and migration management. In: Tejada, G., Bhattacharya, U., Khadria, B., et al. (Eds.), *Indian Skilled Migration and Development: To Europe and Back*. Springer India, New Delhi, pp. 87–111.
- OECD, 2011. *Education at a Glance 2011: OECD Indicators*. OECD Publishing, Paris.
- Open Doors, 2020. U.S. Study Abroad for Academic Credit Trends. Retrieved from. <https://opendoorsdata.org/data/us-study-abroad/u-s-study-abroad-for-academic-credit-trends/>.
- Perkins, M., 2009. Students offered false hope of residency, report warns. In: *The Age*.
- Phan, L.H., Phung, T., 2020. COVID-19 Opportunities for Internationalization at Home. *University World News*. Retrieved from. <https://www.universityworldnews.com/post.php?story=20200828113510793>.
- Phan, L.H., 2017. *Transnational Education Crossing "Asia" and "The West": Adjusted Desire, Transformative Mediocrity and Neo-colonial Disguise*. Routledge, New York.
- Ramírez, G.B., 2018. Introduction: mapping the global brandscape of higher education. In: Papadimitriou, A. (Ed.), *Competition in Higher Education Branding and Marketing: National and Global Perspectives*. Palgrave Macmillan, pp. 1–12.
- Riño, Y., Van Moi, C., Raghuram, P., 2018. New directions in studying policies of international student mobility and migration. *Glob. Soc. Educ.* 16 (3), 283–294. <https://doi.org/10.1080/14767724.2018.1478721>.
- Rizvi, F., 2005. Rethinking "brain drain" in the era of globalisation. *Asia Pac. J. Educ.* 25 (2), 175–192. <https://doi.org/10.1080/02188790500337965>.
- Rizvi, F., 2011. Student mobility and the shifting dynamics of internationalization. In: Davis, D., Mackintosh, B. (Eds.), *Making a Difference: Australian International Education*. UNSW, Sydney.
- Rizvi, F., (Forthcoming). *Rise of Asia, the Shifting Geopolitics and Higher Education*.
- Rizvi, F., 2020. Reimagining recovery for a more robust internationalization. *High. Educ. Res. Dev.* 39 (7), 1313–1316. <https://doi.org/10.1080/07294360.2020.1823325>.
- Robertson, S.L., 2017. Making education markets through global trade agreements. *Glob. Soc. Educ.* 15 (3), 296–308. <https://doi.org/10.1080/14767724.2017.1345408>.
- Sangganjanavanich, V.F., Lenz, A.S., Cavazos Jr., J., 2011. International students' employment search in the United States: a Phenomenological Study. *J. Employ. Counsel.* 48 (1), 17–26.
- Shachar, A., Hirschl, R., 2013. Recruiting super talent: the new world of selective migration regimes. *Indiana J. Glob. Leg. Stud.* 20 (1), 71–107. <https://doi.org/10.2979/indjglolegstu.20.1.71>.
- Sidhu, R., 2009. The "brand name" research university goes global. *High. Educ.* 57 (2), 125–140. <https://doi.org/10.1007/s10734-008-9136-2>.
- Smith, E., 2010. *Pedagogy Not Political Pointscoring: How Training Providers Teach International Students*. Paper Presented at the 13th AVETRA Conference, Gold Coast.
- Spinks, H., 2016. *Overseas Students: Immigration Policy Changes 1997–2015: Research Paper Series 2015–16*. Parliamentary Library, Canberra. https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1516/OverseasStudents. (Accessed 26 July 2021).
- Tran, L.T., Bui, H.T.N., 2019. International graduates: navigating the host and home labour markets. *International Education Association of Australia (IEAA)*. <https://www.ieaa.org.au/documents/item/1743>.
- Tran, L.T., Bui, H., 2021. Public diplomacy and social impact of Australian student mobility to the Indo-Pacific: host countries' perspectives on hosting New Colombo Plan students. *J. Stud. Int. Educ.* 25 (4), 1028315320984833.
- Tran, L.T., Bui, H., Nguyen, M.N., 2021. Youth agency in public diplomacy: Australian youth participation in outbound mobility and connection building between Australia and the Indo-Pacific region. *Int. Stud. Sociol. Educ.* 1–25. <https://doi.org/10.1080/09620214.2021.2017784>.
- Tran, L.T., Bui, H., Tan, G., Rahimi, M., 2022. "It is not OK to think that you are good just because you have graduated from overseas": Agency and contextual factors affecting Vietnamese returning graduates. *Int. Migr.* <https://doi.org/10.1111/imig.12954>.
- Tran, L.T., Marginson, S., 2018. Internationalisation of Vietnamese higher education: an overview. In: Tran, L.T., Marginson, S. (Eds.), *Internationalisation of Vietnamese higher education*, 51. Springer, Cham.
- Tran, L.T., Nguyen, N.T., 2016. The cosmopolitan sojourners: understanding International Korean, Vietnamese and Chinese students' motivations to study overseas. In: *Handbook on Comparative and International Studies in Education*, pp. 561–584.
- Tran, L.T., Nyland, C., 2011. International vocational education and training—the migration and learning mix. *Aust. J. Adult Learn.* 51 (1), 8–31.
- Tran, L.T., Vu, T.T.P., 2016. "I'm not like that, why treat me the same way?" the impact of stereotyping international students on their learning, employability and connectedness with the workplace. *Aust. Educ. Res.* 43 (2), 203–220.
- Tran, L.T., Vu, T.T.P., 2018. Beyond the "normal" to the "new possibles": Australian students' experiences in Asia and their roles in making connections with the region via the New Colombo Plan. *High. Educ. Q.* 72 (3), 194–207.
- Tran, L.T., Phan, H.T.T., Marginson, S., 2018. The "advanced programmes" in Vietnam: internationalising the curriculum or importing the "best curriculum" of the West? In: Tran, L.T., Marginson, S. (Eds.), *Internationalisation in Vietnamese Higher Education*. Springer, Dordrecht, pp. 55–76.
- Tran, L., Rahimi, M., Tan, G., 2019. *Temporary Graduatification: Impacts of Post-study Work Rights Policy in Australia*. Final Report. Deakin University, Victoria, Australia. Retrieved from. https://www.deakin.edu.au/_data/assets/pdf_file/0004/1964317/Post-study-work-rights_Final-report_20092019.pdf.
- Tran, L.T., Phan, H., Tan, G., Rahimi, T., 2020a. "I changed my strategy and looked for jobs on Gumtree": the ecological circumstances and international graduates' agency and strategies to navigate the Australian labour market. *Compare*. <https://doi.org/10.1080/03057925.2020.1837613>.
- Tran, L.T., Rahimi, M., Tan, G., Dang, X.T., Le, N., 2020b. Post-study work for international graduates in Australia: opportunity to enhance employability, get a return on investment or secure migration? *Glob. Soc. Educ.* 18 (5), 495–510.
- Tran, L.T., 2016. "Mobility as becoming": a Bourdieuan analysis of the factors shaping international student mobility. *Br. J. Sociol. Educ.* 37 (8), 1268–1289.
- Tran, L., 2020. Understanding the Full Value of International Students. *University World News*. Retrieved from. <https://www.universityworldnews.com/post.php?story=20200820103708349>.
- Tremblay, K., 2005. Academic mobility and immigration. *J. Stud. Int. Educ.* 9 (3), 196–228. <https://doi.org/10.1177/1028315305277618>.
- UK Council for International Student Affairs (UKCISA), 2019. *Return of Post-study Work*. Retrieved from. <https://www.ukcisa.org.uk/studentnews/1387/Return-of-Post-Study-Work>.
- Wan, C.D., Chapman, D., Hutcheson, S., Lee, M., Austin, A., Zain, A.N.M., 2017. Changing higher education practice in Malaysia: the conundrum of incentives. *Stud. High. Educ.* 42 (11), 2134–2152. <https://doi.org/10.1080/03075079.2015.1134475>.
- Welch, A., Denman, B., 1997. Internationalization of higher education: retrospect and prospect. *Forum Educ.* 52 (1), 14–29.
- Wen, W., Hu, D., 2019. The emergence of a regional education hub: rationales of international students' choice of China as the study destination. *J. Stud. Int. Educ.* 23 (3), 303–325. <https://doi.org/10.1177/1028315318797154>.
- Woldegiorgis, A., Proctor, D., de Wit, H., 2018. Internationalization of research: key considerations and concerns. *J. Stud. Int. Educ.* 22 (2), 161–176. <https://doi.org/10.1177/1028315318762804>.
- Worton, M., 2012. UK higher education today and the place of internationalization. *Reperes Campus France* 15, 1–6.

- Wu, T., Naidoo, V., 2016. International marketing of higher education. Palgrave Macmillan, New York.
- Yang, R., 2006. The commodification of education and its effects on developing countries: a focus on China. *J. fur Entwicklungspolitik* 22 (4), 52–69. <https://doi.org/10.20446/JEP-2414-3197-22-4-52>.
- Ziguras, C., 2005. Chapter 6: International trade in education services: governing the liberalization and regulation of private enterprise. *Counterpoints* 280, 93–112.
- Ziguras, C., Law, S.-F., 2006. Recruiting international students as skilled migrants: the global “skills race” as viewed from Australia and Malaysia. *Glob. Soc. Educ.* 4 (1), 59–76.

Teach for all and the visual regime of global teacher education reform

Daniel Friedrich and Andrew Greene, Teachers College, Columbia University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	260
Themes in the visual analysis	261
Youth	261
Hands	262
Gaze	264
Permanence/change	265
Infographics	266
On the visual regime	267
Concluding thoughts	269
References	269

Introduction

Institutionalized education has had global entanglements since its modern foundation in the 19th century, with school models and ideas traveling, being adapted, and transformed across national borders. Yet the idea of global education reform is a relatively new phenomenon. Through what Stephen Ball (2012) terms Transnational Advocacy Networks, concerted efforts to bring about specific visions of education across the world –including funding models, public-private partnerships, student assessment metrics and rankings, and as is the case for this article, teacher recruiting and training models–have had a growing influence on the localized thinking and practice of educators and policy-makers. Narratives woven as success stories are packaged and sold (be it for monetary or other kinds of resources), as policy entrepreneurs (Ball, 2012) make the case for their global adoption.

One such narrative is that of Teach For America, the program created by Wendy Kopp in 1990 as a way of addressing what she saw as the main problem in US education: low quality teachers. According to Kopp's vision, if high-achievers chose to go into teaching instead of higher-paying professions, they would bring with them their innate brilliance and values to America's neediest students. Thus, Teach For America was set up as a highly selective and prestigious program that would recruit college graduates with high-GPAs from top-ranked institutions to teach for two years in disadvantaged schools. After participating in a five-week training program, "corps members" would be deployed as full-time teachers to receive "real-life" experience in the classroom, so that they would then return to their careers and professions with an eye to enact education reform.

The program grew rapidly and in 2001, Teach First sought to replicate its "success" in the UK. By the mid-2000s, there was a demand for more programs, and in 2007, at the Clinton Global Initiative's annual meeting, Teach For All was founded with Kopp as its CEO. As of July 2021, "Teach For All is a network of 60 independent, locally led and governed partner organizations and a global organization that works to accelerate the network's progress ... Each network partner recruits and develops promising future leaders to teach in their nations' under-resourced schools and communities and, with this foundation, to work with others, inside and outside of education, to ensure all children are able to fulfill their potential" (<https://teachforall.org/about>).

While there is much still to be written, both about Teach For All as a global network, and about the individual affiliate organizations, in the last decade and a half a growing body of literature has emerged that engages with different aspects of this global reform movement.

Inspired by Ball and others' work on policy networks, Saura (2016); Subramanian (2018); Olmedo et al. (2013), and La Londe et al. (2015) seek to illuminate the ways in which the funding and ideological networks behind these programs carry with them a neoliberal understanding of the role of education in furthering market rationality, privatizing the public sector, and undermining public education. The edited volume by Thomas et al. (2020) combines articles dedicated to a study of the network with close-up analyses of specific cases. Friedrich (2014, 2016); Straubhaar (2020); Caetano and de Oliveira Costa (2018), and Yin et al. (2019) go more in depth into one or two programs in order to gain a better understanding about how global discourses are adapted, changed, and/or resisted in specific contexts. Mirroring the movement of former Teach for America corps members to produce counter-narratives to the organization's view of its own success (see e.g. Schonfeld, 2013), the most recent set of academic literature brings to light reflections on the experiences of individuals recruited by different affiliates from across the globe, as compiled in Brewer et al. (2020).¹

For the purpose of this article, the most relevant writing is that of Adhikary et al. (2018); Ellis et al. (2016); Blumenreich and Gupta (2015), and Ahmann (2015), who center their attention on the global discourses that provide meaning to either the network as a whole or specific programs, tracing their emergence and the ways in which these initiatives are used to frame problems and

¹For a more in-depth literature review on Teach For All, see Crawford-Garrett and Thomas (2018).

solutions in the field of education. Southern (2018) aims her work directly at the media representations of Teach First found in brochures, newspaper articles, and documentaries, with the lens on textual language. More specifically, the analysis that follows is a continuation of the project initiated by Friedrich et al. (2015), which looked at the ways in which data and *dataspeak* are mobilized through social media to produce a common sense about what is wrong with education, and therefore what is to be done, positioning Teach For All as the ultimate pragmatic solution to the issues plaguing schools worldwide.

The present article engages in an analysis of visual advertisements posted by Teach For All programs on their respective Facebook pages from January 2017 through July 2020. The goal is to provide a novel entry point into the conversation about the global production of narratives about education, its problems, and solutions. By comparing the visual language being utilized by these programs to appeal to prospective corps members,² donors, and the general public across the globe, we aim at gaining a better understanding of the different ways in which this specific transnational education reform movement imagines and builds its own community. For this purpose, we collected over 1300 images from 52 Facebook pages. While these pages usually had over 1000 images each, we curated our archive by only capturing images that contained the logo of the program (as opposed to photographs documenting everyday activities or images celebrating national holidays or achievements). In our analysis, we did not pay attention to the words –whether we understood the language or not– instead focusing solely on the visual language the images displayed. In doing so, this work can be seen as part of the “visual turn” (Rose, 2001). When discussing this turn in the history of education, Dussel (2012, p. 223) explains: “The visual is probably the dominant mode of representation in the popular historical imagination nowadays ... Pictures do not circulate in isolation, but are part of visual régimes that organize the way we relate to them (Mirzoeff, 2005; Poole, 1997). Researchers, at the same time spectators and producers of contemporary visualities, are part of that régime. Failing to acknowledge this can result in poor scholarship.” Dussel continues on to argue that in order for the visual turn to not be another attention-seeking fad in the history of education (and we would argue, in curriculum studies and educational studies writ large), images cannot be taken as mere surfaces that depict “reality,” but that we must read them as part of those complex visual regimes. Our analysis, then, seeks to better comprehend the visual regime that makes the global education reform movement Teach For All possible and legible.

The more than 1300 images selected were coded into themes based on their visual language. These themes emerged organically as the authors collected images and were then refined as they returned to the full data set. The two authors worked on coding independently, and then compared their work to fine-tune criteria for the purpose of reliability. Our analysis identified five themes in the advertisements: youth, hands, gaze, permanence/change, and infographics. After a review of these themes, we will piece them together to reveal the visual regime that undergirds Teach for All’s narrative strategy, and will conclude by providing tools to place this visual regime within a wider context of global education reform.

Themes in the visual analysis

Youth

(Fig. 1) Likely the most ubiquitous theme presented by the Teach for All network’s partner organizations is that of youth. This is no surprise among the student populations represented in images, but youthfulness and its associated virtues are central to the idealized version of the teacher presented through a broad array of depictions. It is noteworthy that we did not find a single image, across all programs, showing teachers that appeared older than 40.³ Corps members, however, are not only identified as youthful because they look young in age, but also through their high energy, their physical abilities, the compositional style of images, their attire, and their attractiveness.

In order to understand the primary function of these images of youthfulness, it is important to note that the central narrative disseminated by Teach for All is that the current state of education is a problem, and that it can only be solved through radical innovation and “disruption” (Gautreaux and Delgado, 2016). Through such a lens, images of youthfulness in a variety of forms can be read as evidence of Teach for All’s successful approach to change and reform within what have become stagnant and dysfunctional educational systems that are failing students. Representations of youthfulness thus differentiate Teach for All’s approach and its corps members from the traditional image of the teacher. One type of portrayal that accomplishes this goal is the depiction of corps members in energetic, passionate states that are often reciprocated by students. This conveys a vision of teachers committed to and able to connect with students, ones who are driven by their social conscience and dedication to change. This type of portrayal is exemplified by images from Thailand, Bulgaria, and Nigeria, where teachers are shown smiling or with mouth agape, touching or in close proximity to students who appear equally enthusiastic. Such visual constructions of youthful teachers expressing their dedication, coupled with a strong positive response from students, suggests that the path forward in educational reform is based on the ability of talented and hard-working young people to act as saviors for marginalized children.

These images of enthusiasm and ability to relate to students work closely with depictions of youthful physical ability and action. In some ways these portrayals are more subtle than the previous examples, but they convey something that makes young people the logical and only choice for educational reform. Whether in the case of a teacher actively working with building blocks on the floor of

²We utilize the term “corps members” to refer to program participants following the language used by Teach For America. Each program in the network has its own terminology.

³Taking Sweden as an example, this is striking given that 40.5% of secondary public school teachers are over 50 years old, but only 7.4% are under 30 (<https://data.oecd.org/teachers/teachers-by-age.htm>).

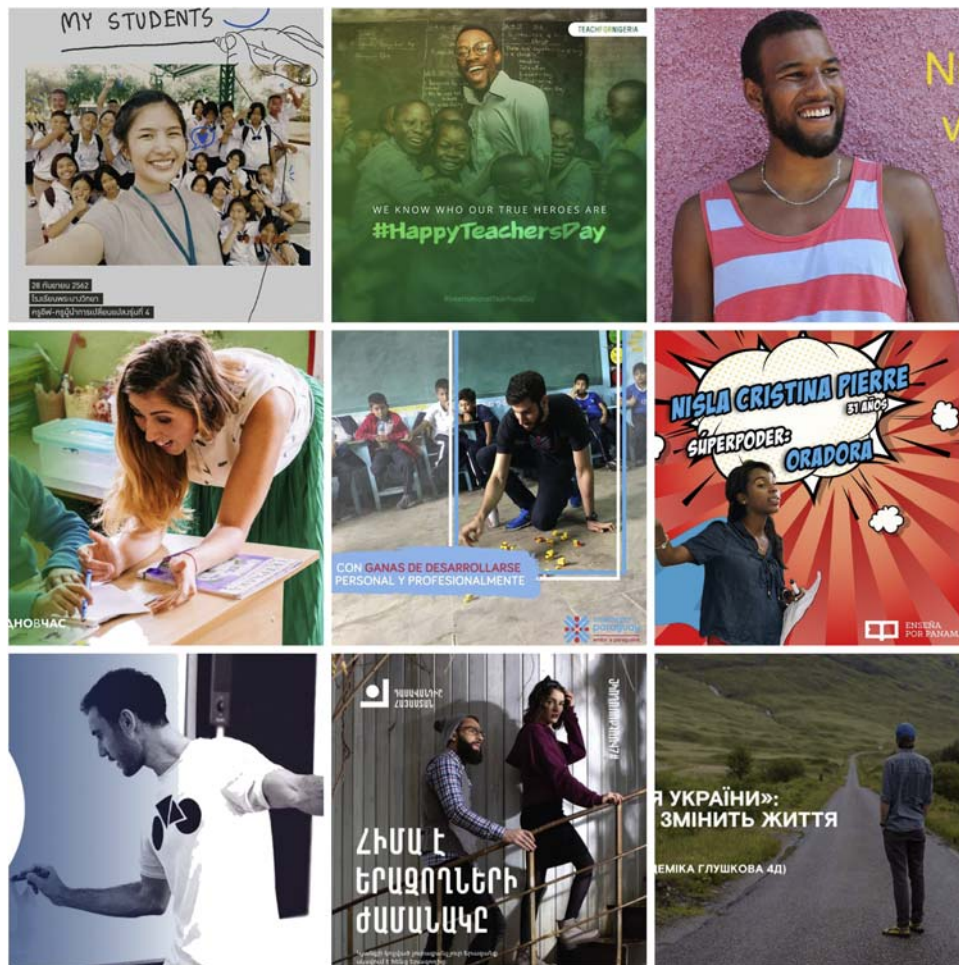


Fig. 1 Images from left to right and top to bottom extracted from programs in: Thailand, Nigeria, Sweden, Bulgaria, Paraguay, Panama, Denmark, Armenia, and Ukraine.

a classroom (Paraguay), a lone individual looking down a long open road in a natural landscape (Ukraine), or a brawny male teacher with an abstracted image of a bicycle on his shirt (Denmark), there is an implication that traditional conceptions of who teachers are and what they are capable of makes them ill-suited to the rigors of the long, metaphorical road to “good” education. Additionally, this sort of image fits neatly within the individualistic, meritocratic narratives that Teach for All relies on more broadly, which focus on the individual’s will to change and innate talent as the keys to reform (Friedrich, 2014).

Many images also gesture toward youthfulness by adopting compositional and stylistic forms that are associated with popular and youth culture. There are various approaches that adopt youthful culture, ranging from the comic book style associated with youth (and superheroes, another instance of individual ability and the savior mentality) seen in materials from Panama, the “selfie” composition equated with young people’s usage of social media seen in an image from Thailand, and even the hip, avant-garde style of the image from Armenia that looks as though it was pulled from the pages of a fashion magazine. These choices again distinguish corps members from the images of teachers that tend to circulate in the collective imaginary, and seek to make teaching within the Teach for All model seem cool.

Each of these three varieties of image perform one component of a recruiting strategy focused on attracting young people to the network. The first and second are appeals to young people’s social consciousness and unique abilities, the third is an alignment with youth culture, and a fourth and final form is aspirational. This last element is less a specific category of image, and more a common quality of many of these depictions; In most cases, corps members are shown wearing fashionable or casual clothing and, overwhelmingly, are physically attractive. This strategy relies on potential applicants’ desires to be like and be associated with the people shown in the images. Could this be a sign of the “instagramification” of teacher recruitment?

Hands

(Fig. 2) Images of children and young people raising their hands appear in 49 out of the 53 programs’ online images. The images call our attention toward engaged students and teachers, to their schooled bodies, and their willingness to succeed. An array of wrist



Fig. 2 Images from left to right and from top to bottom extracted from programs in Italy, Bangladesh, Malaysia, Argentina, Colombia, Germany, Bulgaria, Haiti, and Pakistan.

and hand gestures propping up one, two, or five fingers, denote cultural differences in distinct locations, yet the seemingly universal sign of classroom participation remains easily recognizable. The camera tends to focus either on an individual or a small group of students –capturing their faces either straight ahead or from a slight angle– or on the class as a whole, depicting them from behind with the teacher in the front of the room (yet the figure farthest away from the lens). In the first type of image, the students' faces are trained beyond the camera, likely at the instructor who commands their attention. This focus indicates that they seek to demonstrate their knowledge and be validated as knowing beings. This specific gesture has historical connections to the idea of altitude, height, and elevation being associated with proximity to the divine (for example, teachers' desks sitting on an elevated platform) (Álvarez-Uría and Varela, 1991). The raised hand is the attempt to reach higher than the seated body, and thus to be closer to the knowledge of the teachers and their authority. In these images, it is not only the hand that is raised; the whole body seems poised to jump in anticipation of the teacher's calling, almost unable to be contained under the desk when one is available. This also suggests the unique capacity of the corps member as teacher to inspire students to break the restraints of "bad" education. In a sense it is as though the physical structures of the classroom are not real confinements for students, instead they are held back by the lack of personal will that can only be cultivated by the "good" teacher.

In the second set of images, the focus is divided between the raised arms, heads and necks seen from behind, and the teacher commanding attention at the front of the class. The instructor is the vanishing point of an image that is simultaneously about a group and an individual, with the viewer becoming an observer into a successful pedagogical moment. For those of us who have supervised teachers, these images are the kind that produce nods of acceptance and post-observation compliments: the act of teaching is achieving the goal of engagement.

There is a third, much smaller, subset of images. In these (represented in the collage above by the image from the German program) it is not children or youth raising their hands, but young adults, seemingly corps members. Their faces are blurred, but their gestures are clear. If raising hands in classrooms would appear to be an almost universal symbol of participation, that same participation is what these organizations are seeking from young adults. To be an agent of change requires participating and leading in that change. To take part in the successful reform driven by the Teach For All network, prospective corps members

need to be noticed. They need to channel the energy of their will into becoming a corps member, thus they raise their hands and hope that the organizations will call on them. One can almost imagine the teacher/organization asking: "Who knows what the solution is for educational inequity?" only to be met by students/corps members raising their hands to say: "I do, it is you!"

Gaze

(Fig. 3) Our visual analysis of the ads posted by Teach For All affiliates also reveals the high frequency of images focusing on specific types of gaze. In general terms, the images manifest two forms of gaze: one directed toward the viewer, and one between teachers and students.

Perhaps the most common kind of image in this collection could be characterized as the "team member portrait." In these photographs, young adult men and women pose looking straight at the camera, with either just their faces or the top half of their bodies visible to the viewer. Open smiles and proud expressions accompany eyes that engage the spectator, inviting them to join in and imagine themselves as the future faces of the organization. As we stated above, there are no images of older people, and the youthful gaze seems to be interested solely in meeting other youthful eyes.

Perhaps the only bodies more youthful than those of the corps members are the ones of students, depicted also in a series of portraits. The gaze plays a different, though connected role here. Both the portraits of young teachers and those of children appeal to a sense of participation in social change on the part of the audience for these ads. However, while the former seeks self-identification and a view of futurity (tomorrow *I could be* such an agent of change), the latter seeks to be an interpellation to the viewer as the reason to enact change. What is at stake in these images of children looking directly at the prospective corps member and/or donor is not his or her future as an individual, but their fate as educated subjects. These images recall campaigns by UNESCO and similar organizations using individual children in close-up to stand for the future not only of children, but of the nation (and the world) as a whole. For the viewer, it becomes difficult to ignore the call of a young pair of eyes directed toward one's own will to

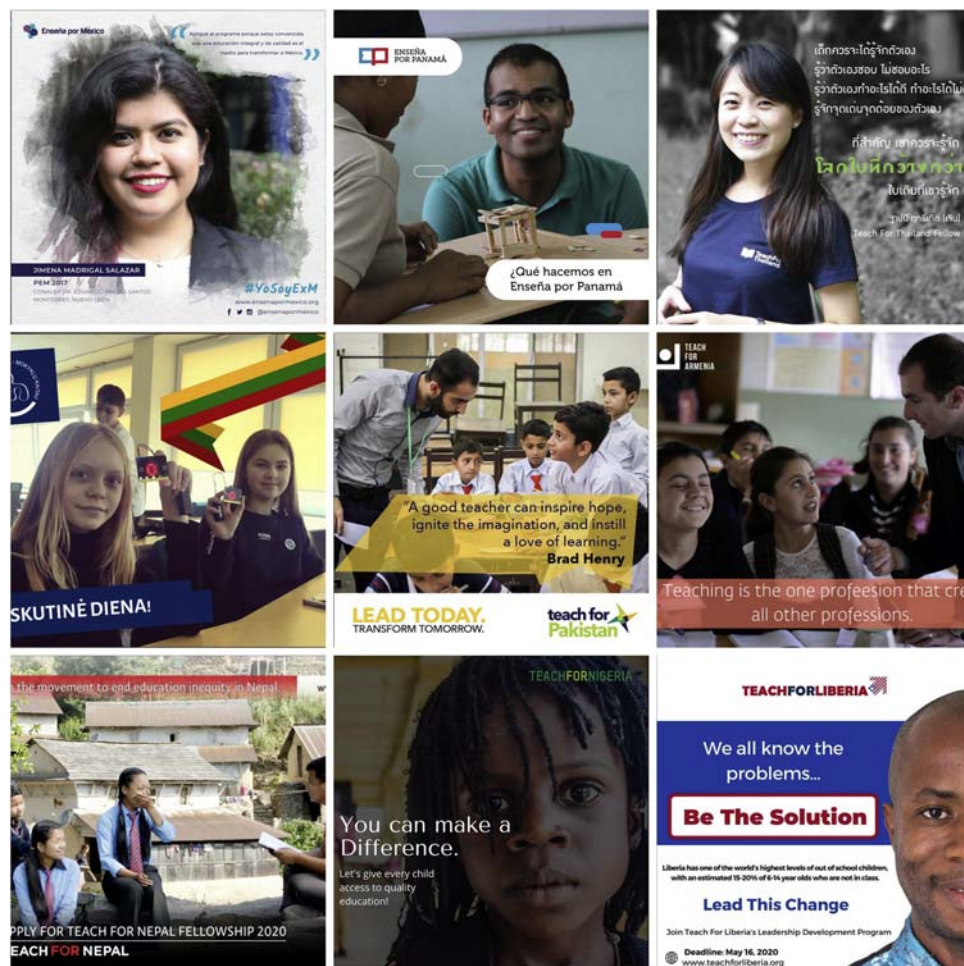


Fig. 3 Images from left to right and from top to bottom extracted from programs in Mexico, Panama, Thailand, Lithuania, Pakistan, Armenia, Nepal, Nigeria, and Liberia.

reform, and when that gaze is accompanied by what appears as a simple solution to a complex problem, the ad fulfills its function of presenting Teach for All's model as the only remedy.

If portraits of corps members are the "who" and portraits of children are the "why," then images of pedagogical interactions (in or out of the classroom), in which there is a clear connection between teacher and student, are the "how." The teachers in these images may or may not be the same ones depicted in the portraits, but they are always young, appear healthy, and look directly at the eyes of a specific student. The student returns the gaze with admiration and awe, as if hungry for someone to *finally* listen and engage. Together with the images of raising hands discussed above, the gaze shared by teacher and student represents the essence of pedagogy and the practical way to effect change in the world. The intimacy of that gaze represents the everyday evidence of the intervention's efficacy, the much-lauded *glimmer in the child's eye*. It also serves as a counter to the enormity of the task. If addressing educational inequity seems like an impossible goal vexed by structural, political, and historical issues, connecting with one child is a much more concrete and achievable way of contributing to social change. It is important to note that in these images the focus is on the two people gazing at each other, while the context is blurred. The spectator cannot see the conditions of the school building; identify the materials present in the classroom; sense how crowded, hot, or cold the rooms are; or experience the smells, sounds, and textures of these educational spaces. What is clearly visible (and therefore the only relevant feature of educational quality) is the satisfaction on the two individuals' faces at a moment of meaningful connection.

Permanence/change

(Fig. 4) It is not enough to read the visual program of Teach for All network partners simply as a critique of stagnant educational systems around the world. Instead, it is essential to understand how Teach for All's problematization of current education and reform is conceptualized as an issue of human capital, one that can be solved effectively by recruiting different (and better) teachers to access classrooms through alternative pathways rather than traditional certification routes (Kopp, 2003). A theme that recurs



Fig. 4 Images from left to right and from top to bottom extracted from programs in Spain, Italy, Uganda, Slovakia, Portugal, Latvia, Japan, Argentina, and India.

through analysis of network partners' social media images is the dichotomy between what is permanent to education, and what change is necessary for successful reform. When looking at images produced in countries across regions, traditions, and societies of all types, it is striking that schools are always recognizable and that certain visual signifiers consistently represent what "has always been" schooling. Yet, this acceptance of permanence in schooling demands that the viewer adopts Teach for All's conception of what is open to reform and what that reform requires.

Chalkboards, neat rows of desks, pencils, paper, and books point to the materiality of permanence in the classroom that seems not to have changed much since the 19th century. Math and literacy appear in white traces on the chalkboard, and the (young) adult stands in front of the class or by the students imparting knowledge. This sense of permanence across time and space is what assures that any viewer will recognize these as images of schooling. However, there are also divergences from what most spectators know schooling to be; something has changed. In these images all children are engaged, working diligently, and show signs of being happy. Teachers are always attentive and invariably exhibit the traits of youthfulness described above; they are energetic, socially conscious, physically able, and have the capacity to connect deeply and directly with children. Schooling in these images seems effective and joyful, in other words, ideal. To most people who have gone through the experience of schooling, this is change; this is not what schools are, but what they should be.

Notably, if the classroom and its material features are accepted as permanent, then the change is in who teaches, how they teach, and the reaction of those who learn, turning those aspects into the target of reform. In other words, the solution to the problem of education is not a rethinking of what schooling looks like and the material conditions that sustain it (or not), but rather, it is a matter of getting smarter, more able, and more committed teachers to do a better job of teaching in order to bring schools closer to their ideal form. This fits neatly within the discourses promoted by Teach for All with its emphasis on the importance of the meritorious individual, unburdened by the inefficiencies of bureaucratic teacher education and certification (Crawford-Garrett and Thomas, 2018), and the idea that what underserved populations need, more than anything, is a different kind of teacher (Friedrich, 2014). It also provides a convenient and cost-effective resolution to the failings of schooling. There is no need to waste resources on laptops, smartboards, other costly technologies, or even infrastructure change in schools. Moreover, these images confirm that the modern schooling conceptualized in the 19th century works (or at least can work) with the right people in charge. All that the students of the world really need are smarter teachers who have survived competitive selection processes and can use their natural capabilities to figure out how to teach better on their own.

Part of what makes Teach for All attractive to a specific segment of the population across the globe is that it appeals to widespread belief in meritocracy and individualism, coupled with a call to be part of significant social change. Teach for All has found purchase in its appeal to young people's sense of social responsibility as well as the future career opportunities its prestige and selectiveness create. For funders, usually those who have reaped the benefits of late capitalism, these narratives often appeal to their views on government (de)regulation and talent. Ultimately, these images, that contrast the permanence of the physical elements of school with a new vision of the teacher, serve as a microcosm of the Teach for All model and its simplicity. If the problem is to be located in bad teachers who do not care enough and have been in the system *forever*, then change is about all students being placed with young teachers who do care.

Infographics

(Fig. 5) The last set of imagery produced by members of the Teach for All partner network that our analysis focuses on is infographics. These representations seek to convey information and data in ways that make them quickly comprehensible in glossy and novel forms. Infographics from partner organizations frequently incorporate easily legible bar and pie charts, maps, or flow charts with illustrations. These graphics are often accompanied by or represented within other highly recognizable symbols that signify something about the data they communicate. These visual choices not only make data the focal point of the stories Teach for All partners tell, but they also convey certain narratives about what is wrong with education and how the particular reform approach of Teach for All is the solution (Friedrich et al., 2015).

As has been established previously, the Teach for All network employs narratives that make the "use of data" (Friedrich et al., 2015) a central technique of problematizing and solving education internationally. Within this context, infographics represent a logical and widely understandable means of disseminating key data to prospective corps members, funders, and the wider public in order to justify Teach for All's approach to educational reform. But it is not only the use of infographics to convey information that is relevant to the Teach for All narrative, it is also the specific symbolism and style that can be read as a strategic program.

One recurring feature is the use of abstracted bodies to represent data. In some cases these bodies are partially filled in, a group of bodies may be colored to indicate quantity or percentage, a number of bodies may expand or decrease over time, and they may be abstracted as male and female symbols. These representations perform two functions: first, they are a reminder that Teach for All's approach seeks to impact people, but, second, the approach does not work on recognizable individuals, rather the program seeks to make change on a scale representable only in numerical data. This variety of infographic, then, is reminiscent of images produced by other large international NGO's that focus on human rights. Additionally, this form of infographic works together with the presentation of youthful bodies visible in some of the images highlighted in previous sections. If images of energetic, able, and desirable bodies are a metaphor for the unique capacity of Teach for All partners and corps member to effect change, then the symbols of these infographics represent the abstracted, but dysfunctional bodies of young people around the world in need of Teach for All's intervention.

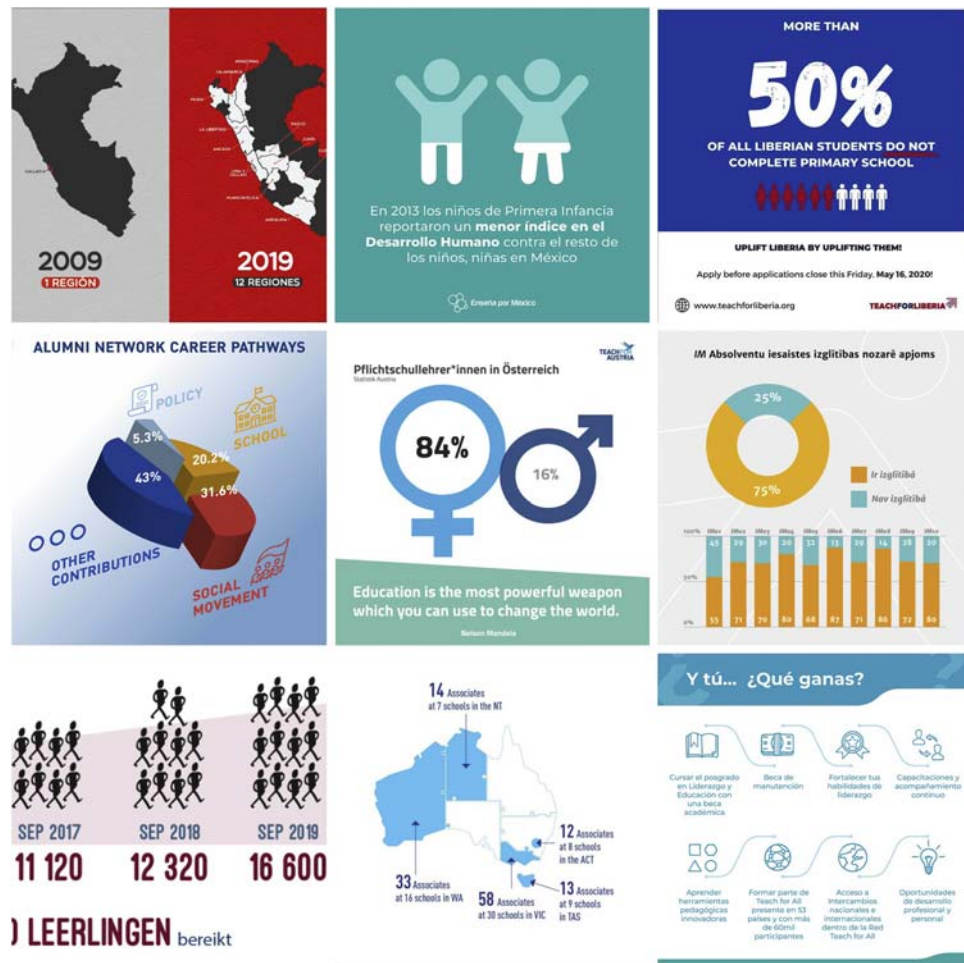


Fig. 5 Images from left to right and from top to bottom extracted from programs in Peru, Mexico, Liberia, Thailand, Austria, Latvia, Belgium, Australia, and Mexico.

A second category of infographic is constructed around maps of countries, usually with state or provincial boundaries included. These variously include color fill or numbers to indicate where the network partner is operating or some other piece of information, usually about the inefficacy of current education in the country represented. The former are more common, and sometimes include multiple maps to indicate change over time, as in the case of Peru above. These images of maps also work in two ways: they both represent the impact that the network organization has on the country and they signal that the network partner is working on behalf of the country. This localizes the Teach for All approach to the national context and demonstrates the success the network partner has already had in reforming nationwide education.

A final category of infographics that recurs across partner organizations is made up of the sort of generic data and flow charts that address the impacts that the Teach for All experience has on its corps members. Here the data being represented do not address problems in education or the effectiveness of Teach for All's approach, but rather indicate the specific value proposition that joining a network partner offers for corps members. These infographics make use of symbols that indicate particular industries or professions and either the percentage of former corps members now working in those areas, or flow charts that suggest a certain path which prospective applicants can look forward to on their way to a desirable career. In either case, the message is clear: if you give two years of your life to teaching through our program, we will offer you a lifelong career of your choosing. Gone here is the play on the individual sense of social responsibility and activism seen in other images, replaced instead by the promise of a personal reward for doing good, a promise backed up by numerical and visual data.

On the visual regime

Taken together, the five themes –youth, hands, gaze, permanence/change, and infographics– are part of an intricate visual regime that, while not exclusive to Teach For All, allows the organization to enter a visual conversation with specific values and beliefs. The

role of social media as not only the medium, but also an active participant in shaping these images cannot be understated. Jurgenson (2019) explains:

Traditional analyses of photography fixate on the photo *object*. This is the thingness of the photograph, as a discrete something with borders. This was central to film photography and continues to animate contemporary discussions about digital photography. The thingness of the social media image is undoubtedly still interesting, but its status as an object is not as central; the what and how of a social photo is less important than the why. As van Dijck puts it, young people “take less interest in sharing photographs as *objects* than as sharing them as *experiences*.” There is a fast-evolving literacy in the circulation of images as communication. Images within the social stream evoke more than they explain; they transmit a general alertness to experience rather than facts. This is what happens when photography is oriented more toward the normal than the exceptional and becomes woven through the contours of everyday life ... Every image has a border. It is a singular document, a record, and a piece of information. But as part of a stream, as an everyday lightweight practice, what emerges is a more nuanced visual literacy. As a visual discourse, social photos are a means to express feelings, ideas, and experiences in the moment, a means sometimes more important than the specific ends of a particular image. (pp. 15, 18)

If we consider a visual regime to refer to “what is seen and how it is seen” (Rose, 2001, p. 6), then Teach For All member organizations have provided us in their advertisements with some guidelines as to how that visual conversation is shaped, as well as the feeling, ideas, and experiences it seeks to evoke. What is seen in these images, although they are static, is purposeful action. The action is the consequence of a will to change by young actors, who participate in a social movement, get people (students and other young adults) engaged, connect at the individual level with those in need, replace previous -older- teachers who do not represent that will to change, and rely on data to leverage their power. By looking across programs, there is a sense that while languages might vary, the visual conversation crosses boundaries. Following the work of Friedrich et al. (2015) on the uses of data and data speak on twitter, our analysis reveals that the production of a global language of education, its problems, and its solutions is as much visual as it is numerical. Without knowing where each image comes from it would be almost impossible to point to specificities and local idiosyncrasies. Invariably the visual regime presents attractive young people, teaching younger generations, in settings commonly understood to be schools, doing what traditional teachers seem to be unable to do, backed by numbers representing both the problems at hand and the impact of this particular solution. What is seen, then, is the power of young entrepreneurship.

A focus on how this regime is seen reveals a set of tensions that undergird it. A first tension is that between the individual –and their will and desires– and the larger system. Images of singular young corps members are presented side by side with infographics that point to the failures of the system to provide an equitable education. The gaze that connects a teacher and a student lives online next to an image in which the percentage of children who do not finish primary school takes up the entire screen.

This first tension between the individual and the system interacts with a second tension between the local and the global aspects of education reform. The images depicting students and young adults cannot but bring to the fore idiosyncratic aspects of localized living: hijabs on girls’ heads, specific colors on pupils’ uniforms, logos drawn using national flags as inspiration. The very name of each one of the member organizations includes the country where it is located, and the portraits of corps members gazing at prospective teachers appeals to a sense of commitment to the nation and its future. Yet, the “all” in Teach for All aims higher. Looking across images produced by these programs it becomes clear that the visual language being is essentially universal. A teacher is a teacher is a teacher; a student is a student is a student. Teach For All’s CEO, Wendy Kopp, expressed: “If our problems are the same, then it gives me hope that the solutions must be shareable” (cited in Murlidhar, 2013), however as Friedrich et al. (2015) make clear, the making of the problems and solutions of education as global phenomena is never smooth or straightforward. If, as some comparative infographics (and the shared aesthetic of the graphs), the involvement of global partners as represented in their logos, and the idea of an “all” that transcends specificities imply, the problems of education are the same and need to be tackled by global policy entrepreneurs (Ball, 2012), then what is the role of an individual’s commitment to the nation? Where are the boundaries between the generic call for energetic young teachers, and the specificities of the self that heeds that call?

The third tension that inhabits the visual regime that governs and is constituted by these images is that between present and future. The images of children looking directly at the lens work together with the colorful graphs to highlight the present crisis, the catastrophe that needs to be addressed now. They are the clarion call for intervention and reform. The images of students raising their hands, and the portraits of corps members have a more ambiguous temporality. While they do depict a moment in time of successful teaching and people who already chose to be part of the organizations, they also have a foot in the future, presented as what needs to happen (more) in order for the first images to disappear. The images of the possible tomorrow (happy and engaged students, successful teachers) are only legible within a theory of change that relies on individual action, interacting with the first tension described above. Absent here is any consideration of the past, including past efforts at addressing inequities. The present in crisis hints at the inefficacy of past reforms, and the focus on youth places all hopes in the future, yet the specifics of what is different this time, what will make the jump from those so-far-unbridgeable gaps and the bright future of effective and meaningful schooling is left in the empty space between images.

The visual regime we are describing exists, then, supported by these three tensions and their interactions, as thousands of images that serve as advertisements for the Teach For All member organizations work to contribute to a common sense about the global education arena, their actors and structures, and their problems and solutions.

Concluding thoughts

As a Transnational Advocacy Network (Ball, 2012), Teach for All has demonstrated a remarkably stable visual narrative among its 54 network partners. The key themes of the visual regime presented on the Facebook pages of partner organizations (youth, hands, gaze, permanence/change, and infographics) work as a language to support and persuade viewers of the universal efficacy of Teach for All's model, and though specific network partners' visual programs are not intended to be read side by side, the clarity of their messages work in unison. Whether these images are read by potential corps members, funders, or a wider audience, they employ a common discourse that proves attractive both ideologically and aesthetically.

But what this common visual regime seems to obfuscate are the problems of education globally, and in particular, how they vary by locality. For any viewer of these images, it is easy to assume that the similarity of network partners' Facebook pages means that the educational problems and solutions for each of their contexts are the same. However, further reflection challenges the notion that the diverse cultural, political, and economic environments of countries such as Sweden, Nepal, and Nigeria (and the differences within those contexts) could possibly produce the same barriers to educational quality, and thereby be easily solved by the Teach for All approach. As Friedrich (2016) has illustrated previously, even the purchase that Teach for All network partners have gained in different nations varies widely. For example, in the free market-oriented Chilean political context, Enseña Chile has spread quickly through the publicly funded schools system, while in neighboring Argentina, Enseñá por Argentina has been barred from placing its corps members in state schools, remanding them instead to a small number of private, largely Catholic schools. Yet, the uniformity of network partners' visual programs glosses over these not so minor details in presenting one common language for understanding the promise and successes of Teach for All.

Finally, it is important to acknowledge that at the center of the visual narrative presented by these images is an absence: the bad teacher. While the theory of change promoted by Teach for All, one focused on replacing bad teachers with new, young, talented ones, is explicit in most other outlets, these images have no way of presenting the problem of bad teachers in a direct way. Though there are a number of infographics that highlight data and statistics about the poor quality of education, overwhelmingly, network partners present common images of what good education looks like. Viewers, then, are left to imagine the bad teacher in the negative space left behind after defining the good teacher as conceived by Teach for All. This seems to be the result of both the challenges of presenting a negative image in advertisements and the potential for misreading what is being shown. For example, while Teach for All would like to convey that older teachers are the problem, were they included in these representations, some viewers may associate their age with experience, which seems like an asset. As such, it makes sense then that the images we do see on these Facebook pages are oriented in the way they are, with a fixation on youth, physical and cognitive ability, the capacity to engage and inspire, and the willingness to base decisions on rational data.

Teach for All has made it clear that if the solution to the problems of education is fundamentally a human capital one, and that the central mechanism for enacting change is recruiting talented young individuals and getting them into a classroom as quickly as possible, it follows that answer is to elevate the ideals of social consciousness and entrepreneurship. As globalizing narratives about the value of competition and market forces grow ever stronger, Teach for All continues to lean into the universality of its approach and the visual regime of its partner networks' Facebook pages play a central role in its public messaging.

References

- Adhikary, R.W., Lingard, B., Hardy, I., 2018. A critical examination of Teach for Bangladesh's Facebook page: 'social-mediatisation' of global education reforms in the 'post-truth' era. *J. Educ. Pol.* 33 (5), 632–661.
- Ahmann, C., 2015. Teach for all: storytelling "shared solutions" and scaling global reform. *Educ. Pol. Anal. Arch.* 23 (45).
- Álvarez-Uría, F., Varela, J., 1991. *Arqueología de la escuela*. Ediciones de la piqueta, Madrid.
- Ball, S.J., 2012. *Global Education Inc.: New Policy Networks and the Neo-Liberal Imaginary*. Routledge, London, UK.
- Blumenreich, M., Gupta, A., 2015. The globalization of Teach for America: an analysis of the institutional discourses of Teach for America and Teach for India within local contexts. *Teach. Teach. Educ.* 48, 87–96.
- Brewer, T.J., deMarrais, K., McFaden, K.L. (Eds.), 2020. *Teach for All Counter-narratives: International Perspectives on a Global Reform Movement*. Peter Lang, New York, NY.
- Caetano, M.R., de Oliveira Costa, M., 2018. Gerencialismo, Internacionalização da educação e o papel da Teach for All no Brasil. *Revista Educação Em Questão* 56 (48), 249–275.
- Crawford-Garrett, K., Thomas, M.A., 2018. Teacher education and the global impact of Teach for All. In: *Oxford Research Encyclopedia of Education*. Oxford University Press, Oxford, UK.
- Dussel, I., 2012. The visual turn in educational history: just another fad, or a serious challenge to historians? *Int. J. Historiogr. Educ.* 12 (2), 221–227.
- Ellis, V., Maguire, M., Trippestad, T.A., Liu, Y., Yang, X., Zeichner, K., 2016. Teaching other people's children, elsewhere, for a while: the rhetoric of a travelling educational reform. *J. Educ. Pol.* 31 (1), 60–80.
- Friedrich, D., Walter, M., Colmenares, E., 2015. Making all children count: teach for all and the universalizing appeal of data. *Educ. Pol. Anal. Arch.* 23 (48).
- Friedrich, D., 2014. Global microlending in education reform: enseñá por Argentina and the neoliberalization of the grassroots. *Comp. Educ. Rev.* 58 (2), 296–321.
- Friedrich, D., 2016. Teach for All, public-private partnerships, and the erosion of the public in education. In: Verger, A. (Ed.), *World Yearbook of Education 2016: The Global Education Industry*. Routledge, London, UK, pp. 180–194.
- Gautreaux, M., Delgado, S., 2016. Portrait of a Teach for All (TFA) teacher: media narratives of the universal TFA teacher in 12 countries. *Educ. Pol. Anal. Arch.* 24 (110).
- Jurgenson, N., 2019. *The Social Photo: On Photography and Social Media*. Verso Books, London, UK.
- Kopp, W., 2003. *One Day, All Children: The Unlikely Triumph of Teach for America and what I Learned along the Way*. PublicAffairs, New York, NY.
- La Londe, P.G., Brewer, T.J., Lubienski, C.A., 2015. Teach for America and Teach for All: creating an intermediary organization network for global education reform. *Educ. Pol. Anal. Arch.* 23 (47).
- Mirzoeff, N., 2005. *Watching Babylon: the war in Iraq and global visual culture*. Routledge, United Kingdom.
- Murldhar, C., 2013. Looking Back on the Global Conference. Retrieved from <http://teachforall.org/en/news/looking-back-global-conference>.

- Olmedo, A., Bailey, P.L., Ball, S.J., 2013. To infinity and beyond...: heterarchical governance, the Teach for All network in Europe and the making of profits and minds. *Eur. Educ. Res. J.* 12 (4), 492–512.
- Poole, D., 1997. *Vision, race and modernity: a visual economy of the Andean image world*. Princeton University Press, New Jersey.
- Rose, G., 2001. *Visual Methodologies: An Introduction to Researching with Visual Materials*. Sage, London, UK.
- Saura, G., 2016. Saving the world through neoliberalism: philanthropic policy networks in the context of Spanish education. *Crit. Stud. Educ.* 59 (3), 279–296.
- Schonfeld, Z., 2013. Meet the Teach for America resistance movement that's growing from within. *Atlantic*. Retrieved from. <https://www.theatlantic.com/national/archive/2013/07/meet-teach-america-resistance-movement-s-growing-within/313358/>.
- Southern, A., 2018. Disrupting the habitus? Media representations and participant experience of Teach First: an exploratory case study in Wales. *Teach. Teach.* 24 (5), 584–597.
- Straubhaar, R., 2020. Teaching for America across two hemispheres: comparing the ideological appeal of the Teach for All teacher education model in the United States and Brazil. *J. Teach. Educ.* 71 (3), 307–318.
- Subramanian, V.K., 2018. From government to governance: teach for India and new networks of reform in school education. *Contemp. Educ. Dial.* 15 (1), 21–50.
- Thomas, M., Rauschenberger, E., Crawford-Garrett, K., 2020. *Examining Teach for All: International Perspectives on a Growing Global Network*. Routledge, United Kingdom.
- Yin, Y.M., Dooley, K., Mu, G.M., 2019. Why do graduates from prestigious universities choose to teach in disadvantaged schools? Lessons from an alternative teacher preparation program in China. *Teach. Teach. Educ.* 77, 378–387.

Policy isomorphism: a lens for understanding the influence of globalization on national education policymaking

Yuzhuo Cai, Faculty of Management and Business, Tampere University, Tampere, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	271
Conceptualization of policy isomorphism	271
Policy isomorphism as one of the perspectives for understanding national education policymaking in the context of globalization	273
Policy isomorphism in a broad context of social sciences	273
Boundary conditions for applying the analytical framework of policy isomorphism	274
Isomorphic mechanisms in education policymaking in a global education system	274
Controversial views concerning policy isomorphism	275
Confirmative views	275
Criticisms	275
Alternative approaches	276
Conclusions	277
References	277

Introduction

It has been argued that “the globalization of economy is accompanied by the globalization of policy making” (Moutsios, 2009, p. 467). As education is increasingly embedded in the process of globalization, national education development and policymaking are unavoidably subject to global pressures (Carney et al., 2012; Schofer and Meyer, 2005; Rizvi and Lingard, 2009; Yang, 2003). The influence of globalization on national policymaking can be understood through the lens of “policy isomorphism.” In the literature on public policy, the term “policy isomorphism” generally refers to the phenomenon of similar policies across nations (Villadsen, 2011) or “the convergence and similarities in structures or policy outputs of different governing bodies, including localities and nations” (Zhu and Meng, 2018, p. 1). The theoretical root of policy isomorphisms is based on DiMaggio and Powell’s (1983) notion of institutional isomorphism, which best captures the process of the homogenization of organizations in an institutional environment. Policy isomorphism concerns the application of institutional isomorphism in policy analysis, specifically concerning the process of policymaking.

Education studies applying institutional isomorphism have been used mainly for organizational analysis, whereas those dealing with policy issues mostly address national contexts (Cai and Mehari, 2015; Baker and Wiseman, 2006). Nevertheless, issues related to policy isomorphism in a global context are not rare in the literature on education research. However, these studies often use the concepts of policy convergence, policy diffusion, policy borrowing, policy learning, and policy transfer (Verger et al., 2012; Green, 1999) and apply approaches such as “globalizing education policy analysis” (Rizvi and Lingard, 2009) and “mechanisms of external effects on national policies” (Dale, 1999). Policy isomorphism, grounded in institutional isomorphism, could offer an effective lens via which to analyze the phenomenon of policy convergence, especially concerning the influence of globalization on national education policymaking. Policy isomorphism also has the potential to help enhance and build synergies among existing approaches to policy convergence in the context of globalization. However, there is a lack of systematic elaboration of the concept of policy isomorphism, as well as the use of the concept in education policy studies when globalization is concerned.

This article elaborates on the concept of policy isomorphism as a lens for analyzing national education policymaking in the context of globalization by integrating the insights of institutional theory, public policy studies, and education research on policymaking in a global context. It highlights the mechanisms of policy isomorphism but also reminds the reader of the theoretical limits of policy isomorphism. The remainder of the paper is structured as follows. First, it begins with conceptualizing policy isomorphism and introducing its underlying mechanisms based on the theory of institutional isomorphism. Then, it discusses how policy isomorphism can be used as an analytical framework to understand the processes of national education policymaking in the context of globalization, in comparison with some alternative approaches. This is followed by a discussion of controversial views regarding policy isomorphisms reflected in the literature. This paper concludes with suggestions for future research on applying policy isomorphism in education policy research.

Conceptualization of policy isomorphism

The conceptualization of policy isomorphism cannot be done without understanding the notion of isomorphism and, more specifically, institutional isomorphism (DiMaggio and Powell, 1983). Drawing on Hawley (1968), DiMaggio and Powell (1983) define

isomorphism as “a constraining process that forces one unit in a population to resemble other units that face the same set of environmental conditions” (p. 149). While Hawley (1968) refers to the social, economic, political, and biophysical aspects of the environment in his conceptualization of isomorphism, DiMaggio and Powell (1983) distinguish between the technical and institutional environments because the mechanisms of isomorphism in these two contexts are different.

Demands driven by technical/market environments include changes in consumers’ preferences, competitive conditions, and other characteristics of the task environment; the institutional environment refers to the rules, norms, understandings, beliefs, and taken-for-granted assumptions regarding what constitute appropriate or acceptable organizational forms and behaviors (Meyer and Scott, 1983). In the technical environment (e.g., a marketplace), competitive isomorphism is used to understand the phenomenon of homogeneity, in which increasing similarities between organizations are caused by competitive pressures (Hannan and Freeman, 1977). The notion of institutional isomorphism stresses the importance of legitimacy and the logic of appropriateness for processes of homogenization (DiMaggio and Powell, 1983).

DiMaggio and Powell (1983) identify three processes of institutional isomorphism, namely coercive, mimetic, and normative. Coercive isomorphism stems from political influence and the need for legitimacy. In other words, coercive pressures are derived from other organizations in that an organization is dependent on the expectations of the social surroundings in which that organization is embedded. The chief coercive forces include political power, the legal environment, and governmental mandates. The extent of the structural impact on organizations depends on the resources received from powerful central institutions. Such a logic is not far from the resource dependency theory (Pfeffer and Salancik, 1978), which predicts that a more plural resource base should lead to higher organizational autonomy and, hence, potentially greater organizational diversity. However, institutional theory and resource dependency theory focus on the institutional environment and technical environment, respectively.

Mimetic isomorphism occurs when actors face uncertainty and attempt to emulate successful organizations as a solution. Mimetic isomorphism is based on imitation, which is characterized by copying structural elements of organizational patterns that are believed to be successful and legitimate. Normally, an organization tends to imitate well-established organizations when it faces uncertainty. Uncertainty refers to “the degree to which future states of the world cannot be anticipated and accurately predicted” (Pfeffer and Salancik, 1978, p. 67). According to DiMaggio and Powell (1983), the extent of mimetic isomorphic effects on an organization depends on the degree of uncertainty and the ambiguity of goals.

Normative isomorphism arises primarily from professionalization, which refers to “the collective struggle of members of an occupation to define the conditions and methods of their work” (DiMaggio and Powell, 1983, p. 152). Professionalization involves two aspects: one is the homogenizing influence of established norms, and the other is the growth and elaboration of professional networks (DiMaggio and Powell, 1983). As explained by Levy (2006, p. 145), “Mimetic isomorphism occurs when actors are otherwise unclear on what to do and therefore copy successful organizations, whereas normative isomorphism arises where professions or others feel capable of mapping their own policy but do so based on their socialization of dominant norms.” According to DiMaggio and Powell (1983), the greater the reliance on using professional credentials to choose staff, the greater similar to other organizations will be. Also, the greater the participation of members in professional organizations, the more alike organizations will be.

Associated with the notion of institutional isomorphism is the concept of organizational field, which is defined by DiMaggio and Powell (1983, p. 148) as “those organizations that, in the aggregate, constitute a recognized area of institutional life: key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar services or products.” They also note, “Organizational fields provide a context in which individual efforts to deal rationally with uncertainty and constraint often lead, in the aggregate, to homogeneity in structure, culture, and output” (DiMaggio and Powell, 1983, p. 147). The notions of organizational field and institutional isomorphism largely enhance neo-institutional theory, which was originated in Meyer and Rowan (1977) and Meyer (1977).

When conducting empirical research on institutional isomorphism, two components are essential: organization and organizational field—the environment in which isomorphism takes place (DiMaggio and Powell, 1983). It should be noted that DiMaggio and Powell (1983) mainly used the mechanisms of isomorphism to explain “what makes organizations so similar” (p. 147), not specifically for analyzing policy similarities across nations. When analyzing policy isomorphism in a global context, the environment is the world or global system of education, and organizations are national governments and other actors involved in national policymaking (Cai, 2010; Drori, 2008). Then, the three mechanisms of institutional isomorphisms can be used to analyze policy isomorphism.

While the mechanisms of institutional isomorphism provide a theoretical account of how organizations are subject to institutional pressures in the process of homogenization, Villadsen (2011) contends that the mechanisms are not sufficient in explaining “how institutional pressures are conveyed and under which circumstances they are translated into decisions and action in organizations” (p. 574). Thus, he proposes integrating a social network perspective into institutional isomorphism analysis. This is, indeed, in line with the fundamental theoretical assumption of institutional theory, which primarily deals with a “network of organizations” (Greenwood and Hinings, 1996, p. 1026). Owen-Smith and Powell (2008) assert that networks are “the channels through which institutional effects flow” and “networks of like-minded individuals are central reference groups that promote widely emulated practices” (p. 599).

Villadsen (2011) particularly suggests that institutional expectations/pressures, together with the associated information, are conveyed through the social networks of political executives (i.e., mayors in their research setting). Therefore, public organizations with their executives centrally positioned in networks (or organizational fields) are likely to be the agents of policy isomorphism. The social network of policymakers across nations is an important source of policy isomorphism in a global context because

conveying norms, information, knowledge, and expectations relies heavily on international networks, sometimes through their connections to international organizations (Cai, 2010). Not only policymakers but also participants in education policy processes, such as teachers, can also be agents of isomorphism through their international mobility (Brown and Stevick, 2014).

In short, the conceptualization of policy isomorphism can be summarized as follows. First, from the perspective of institutional isomorphisms, policy isomorphism can be defined as a constraining process that forces policymakers, in the form of the governing bodies of regions and nations, in a global environment to resemble one another in terms of policymaking, especially when addressing similar policy problems. Second, the core elements of policy isomorphism analysis are national governments (as organizations) and the world system, e.g., the global system of education (as the organizational field or institutional environment). Third, there are three types of mechanisms of policy isomorphism, namely coercive, mimetic, and normative mechanisms.

Policy isomorphism as one of the perspectives for understanding national education policymaking in the context of globalization

While the concept of policy isomorphism is to be further elaborated on here to build an analytical framework with which to understand national policymaking in a global educational system, it should be noted that the applicability of the policy isomorphism perspective depends on certain conditions. The power and limits of policy isomorphism as an analytical tool are also related to theoretical assumptions about human behaviors.

Policy isomorphism in a broad context of social sciences

Although the power of institutional isomorphism in national education policymaking has been widely acknowledged in the literature (e.g., Jarvis, 2014; Joo and Halx, 2012; Cai, 2010; Rakic, 2001), this same literature has warned that not all policymaking processes, especially those perceived as institutional change, can be explained by institutional isomorphism (Joo and Halx, 2012; Levy, 2006; Beckett, 2010). To understand the limits of institutional isomorphism, one must also understand two competing views in the social sciences on understanding the behavior of individuals. These are the cultural anthropology and rational choice theory, which are located at the two extremes of a spectrum.

Anthropologists believe that human behavior is governed by culturally transmitted norms, and that such norms contain accumulated wisdom, which allows people to behave sensibly even though they do not understand what they do and why they do. Rational choice [on the contrary] consider that people make behavioral choices [based] on their rational calculations, in which the central element involves a cost benefit analysis.

Cai and Yan (2011, p. 56).

Institutional isomorphism, as a core aspect of new institutional theory, is largely shared with anthropologists' thinking. It has become popular in policy analysis because of the increasingly strong "world culture" (Schofer and Meyer, 2005) that affects national education policies (Kim and Boyle, 2012; Carney et al., 2012). Meanwhile, sovereign states are supposed to design and implement their policies independently (Joo and Halx, 2012). This can be explained by the rational theory of decision making, according to which policymakers' choices are influenced by their preferences and projections of maximized policy outcomes (Howlett et al., 2009).

Recent developments in institutional theory have integrated the word-culture perspective and the role of human agency (Cai and Mehari, 2015). Within such a perspective, individuals' behaviors and choices are directly driven by their interests and preferences, but the ways in which they pursue their interests are bound by institutional rules, especially in legitimated institutional settings (Ingram and Clay, 2000; Scharpf, 1997). Because national policymakers are embedded in intricate global networks, their interests and preferences are shaped by global institutional constraints (Cai, 2010).

In addition, policymakers are also influenced by rules and routines from the past—as explained by the path-dependency theory (Greener, 2002). Path dependency is tightly associated with historical institutionalism (Suddaby et al., 2013), as well as other schools of neo-institutionalism. As noted by Raadschelders (1998), "path dependency not only connects the past to the present but highlights the fact that the past limits the range of choices in the present" (p. 570). Kay (2005, p. 562) further articulates, "It occurs only once in a certain context and at a certain time," and the condition of path dependency is "based on the fact that events in nature are repeatable and recurrent," often in a national context. However, we are now in the age of innovation, in which the world is changing quickly and increasingly uncertain thanks to the rapid advancement of globalization and technology. Nations are not simply embedded in their own institutional environments rooted in the past but in a complex global system.

When the global system can be treated as an organizational field, institutional isomorphism can be a more suitable analytical framework than the path-dependency theory for understanding national policymaking processes. However, if the global system consists of multiple or even competing institutional elements, the explanatory power of institutional isomorphism (and neo-institutionalism in general) may be constrained. In this regard, the institutional logics perspective has its advantages. The concept of institutional logics was introduced by Alford and Friedland (1985) when describing how contradictory practices and beliefs inherent in modern Western societies shape individuals' actions in the political arena. One central pervasive argument of the institutional logics perspective is that multiple and contending logics provide the dynamic for potential change in both organizations

and societies (Thornton et al., 2012). It has been shown that there are multiple and conflicting institutional logics in the global system of education, particularly higher education (Cai and Mountford, 2021).

Thus, policy isomorphism is only one of the perspectives for understanding national education policymaking in a global context. When isomorphism is applied to policy analysis in a specific research setting, one must be aware of its theoretical limits and alternative perspectives.

Boundary conditions for applying the analytical framework of policy isomorphism

When using policy isomorphisms as an analytical tool to explain the influence of globalization on national education policymaking, three conditions must be met. First, policy problems are concerned with institutional pressures. Policy problems exist in social contexts, including both technical and institutional dimensions (Meyer and Scott, 1983). While in a technical environment, “organizations compete ... for resources and customers,” in an institutional environment, organizations compete “for political power and institutional legitimacy, for social as well as economic fitness” (DiMaggio and Powell, 1983, p. 150). Thus, policy isomorphism has its most explanatory power when the dynamic of policymaking is primarily concerned with the institutional environment.

Second, in a policy isomorphism analysis, the global environment in which public organizations and policymakers are embedded can largely be seen as an organizational field, instead of an institutional system. Cai and Liu (2020) identify three differences between an organizational field and an institutional system, both of which are used to understand the organizational environment. First, an institutional system may cut across several organizational fields (Thornton et al., 2012). Second, an organizational field is characterized by a structuration that results in less diversity (DiMaggio and Powell, 1983), while an institutional system is comprised of mingling and conflicting institutional logics (Thornton et al., 2012). Third, the concept of an institutional system emphasizes the active roles of both organizations and individuals for institutional changes (Thornton et al., 2012), while the organizational field mainly deals with relationships in which similar types of organizations share (or are constrained by) similar norms.

The third condition is that the policymakers are embedded in international social networks concerning education development. Typical examples of national policymakers involving the networks occur through joining intergovernmental organizations or being members of international nongovernmental organizations. The global networks are key carriers of “world culture” (Schofer and Meyer, 2005), which explains global isomorphism in education—a phenomenon first noted by John Meyer in his observation of the global expansion of mass schooling (Meyer, 1971, 1977). The global norms, expectations, and knowledge conveyed by international networks influence national policymaking through the transferring the global ideologies, as a source of legitimacy, into local contexts. This corroborates the central theoretical claim of world culture (Carney et al., 2012): educational policies are not particularly responsive to the political, economic, and social characteristics of individual nation-states but often result from the “characteristics of the contemporary world system” (Meyer et al., 1977, p. 242).

It should be acknowledged that any national policymaking is influenced by multiple factors, from both technical and institutional environments, as well as global and local pressures. Policy isomorphism can be used as a useful lens in policy analysis when national policymaking is primarily affected by institutional factors and global networks or the research focus is on that.

Isomorphic mechanisms in education policymaking in a global education system

Policy isomorphism is represented by the three mechanisms of institutional isomorphisms. Drawing on Chen (2008), Cai (2010) labels the mechanisms of isomorphisms in education policymaking as (1) mechanisms of international regulation, (2) mechanisms of mimetic learning, and (3) mechanisms of consultancy involvement when analyzing the influences of global pressures on national higher education policymaking.

Under the mechanism of international regulation, nations develop similar policy agendas due to the requirements of/suggestions from international organizations or intergovernmental organizations. The most influential international organizations in education policy are, for example, the United Nations Educational, Scientific, and Cultural Organization (UNESCO); World Bank; World Trade Organization (WTO); International Monetary Fund (IMF); and Organization for Economic Cooperation and Development (OECD) (Beech, 2009; Moutsios, 2009). Moutsios (2009) claims that these organizations’ decisions “are decisively shaping current directions and developments in national education systems” (p. 467). This reflects the coercive mechanism. In Moutsios’s (2009) words, “the mode of decision making characterizing transnational political structures is network-like or ‘netocratic’, rather than democratic” (p. 469). For instance, national policymakers’ decisions, influenced by “human capital and neoliberal ideologies,” are made “through largely asymmetric, non-democratic and opaque procedures” (Moutsios, 2009, p. 476).

Under the mechanism of mimetic learning, nations attempt to imitate successful nations when they face uncertainties concerning future development or ambiguous development goals. If there is one word that could capture contemporary societal development, including the domain of education, it is uncertainty (Altbach and Reisberg, 2018). The uncertainties in education have been widely reported in the literature in terms of their various aspects, such as learning (e.g., Susnea et al., 2014), education-to-work transition (e.g., Lechner et al., 2016), educational leadership (e.g., Jameson, 2012), the financing of education (e.g., Maman and Rosenhek, 2020), and education policy (Sirat, 2010). Following DiMaggio and Powell (1983), uncertainty invites processes of imitation and copying—the search for apparently legitimate and successful solutions elsewhere. For instance, many aspects of higher education policy development in Central and Eastern Europe (Dobbins and Knill, 2009), Russia (Gounko and Smale, 2007), and China (Yang, 2019; Cai, 2012) have, to a large extent, been driven by memetic learning. It should be noted that “policy

mimesis" is "the imitation or reproduction of a policy in another context, rather than a simple transposition across geographical and political boundaries" (Massey, 2009, p. 383).

Under the mechanism of consultancy involvement, nations formulate similar policies when the policymakers and/or their consultants are influenced by globally shared norms within a profession. This can be understood as the normative pressures induced by professionalization. In the process of education policymaking, not only do politicians make decisions but other professions are also involved as policy consultants, including international organizations, e.g., the World Bank, think tanks operating on a global scale (Spring, 2005), as well as individual educational professionals (Gunter and Mills, 2017a) and researchers (Gunter and Mills, 2017b; Ion et al., 2019). All these policy consultants may convey global norms to national contexts but in different ways. While the role of international consultancies' influences on the national policy agenda is apparent, educational researchers and professionals bring global ideas to local contexts due to their international experience and mobility (Brown and Stevick, 2014). When think tanks are engaged with policymaking in different countries, they help transfer experiences and best practices among these countries.

The three isomorphic mechanisms take place simultaneously and reinforce one another (Tuttle and Dillard, 2007). Although, in some cases, one mechanism may more strongly influence the homogenization process than others (Boxenbaum and Jonsson, 2017), the three mechanisms are often intricately interlinked in both the practice of policymaking and empirical analysis (Hallonsten and Hugander, 2014).

Controversial views concerning policy isomorphism

After conceptualizing policy isomorphism from a theoretical perspective, this section examines how policy isomorphism is reflected in empirical research concerning the influence of globalization on national education policies. Based on reviewing the relevant literature, three controversial views on policy isomorphism are identified: (1) those confirming the explanatory power of policy isomorphism, (2) those criticizing policy isomorphism, and (3) those proposing alternative approaches to the phenomenon of policy convergence. These controversial views are further explained as follows.

Confirmative views

Most studies about the influence of globalization on national education policymaking confirm the trends of policy isomorphism. The empirical cases in this camp lie between two extremes. On the one end are European countries. The European Union intends to streamline higher education policies among European member states, and thus, policy isomorphism is expected (Dobbins and Knill, 2009; Rakic, 2001). The cases on the other end are concerned with statist countries, such as China. There is a perception that the Chinese government makes its own decisions in formulating national education reform policies, while global ideologies play a minor role (Cai, 2010). In between the two extreme ends, there are countries in diverse contexts, such as Korea (Joo and Hah, 2012) and Australia (Croucher and Woelert, 2016). Policy isomorphism is mainly reflected in aspects such as the governance of higher education (Cai, 2010), quality assurance systems (Seyfried et al., 2019), and research funding structure (Hallonsten and Hugander, 2014).

While a large amount of literature finds policy convergence, sometimes, in the same report, both convergence and diversity are observed in terms of different aspects of education policy. For instance, Rakic (2001) found, in his analysis of the Netherlands, Belgium/Flanders, and Germany, that "a number of important indications of convergence are present in national policy arrangements in the field of higher education (in particular student mobility and quality assurance), but less in the structure of higher education systems" (p. 225). Although not directly discussing policymaking in a global context, Cai and Yan (2011) explain why the convergence and diversity of education development may co-exist. They argue that there is no simple tendency toward the convergence or divergence of policy design in an organizational field; what aspects are isomorphic or divergent depends on whether decision-makers want to pursue legitimacy in a (global) institutional environment or economic fit in a local technical environment.

Criticisms

While current education research shows that policy isomorphism is becoming prevalent in the global education field, some education researchers are very concerned about its negative effects. As LeTendre et al. (2001) put it, "it is dangerous to imply that a global dynamic culture pushes everything everywhere to be the same at all times. Empirically this is not true and conceptually it leads to a theoretical dead end where unique national and local change is impossible" (p. 13). Certainly, there are other factors besides world culture at work. Brint et al. (2009) point out that decision-makers often face the need to balance pressures from multiple interest groups instead of simply gaining legitimacy by conforming to "rationalized myths" (Meyer and Rowan, 1977) in society. In the same vein, Wiseman et al. (2014) criticize two misconceptions of isomorphisms or neo-institutionalism: "(1) the belief that the 'world culture' strand is the only version of neo-institutional theory applicable to comparative education research; and (2) the assumption that the global homogenization of society, culture and schooling is a goal of researchers applying neo-institutional theory to comparative education phenomena" (p. 687).

Indeed, policy isomorphism, or institutional isomorphism, has its theoretical limits. It focuses too heavily on structuration in an institutional environment while ignoring the technical (or task) environment, as well as institutional complexity. In their literature analysis of the use of institutional theory in higher education research, Cai and Mehari (2015) remind us that the new

institutionalism, in which institutional isomorphism is rooted, “mainly provides a good account of the similarity or stability of organizational arrangements in a given organizational field” but is inefficient in explaining institutional changes due to its ignorance of “power, self-interest, efficiency and human agency” (p. 11). According to Boxenbaum and Jonsson (2017), current research applying institutional isomorphism has not even utilized the full power of neo-institutionalism. For instance, “decoupling,” another core concept of neo-institutionalism, is often absent in studies analyzing institutional isomorphism. As “organizations decouple their formal structure from their production activities when institutional and task environments are in conflict, or when there are conflicting institutional pressures” (Boxenbaum and Jonsson, 2017, p. 78), bringing both concepts of institutional isomorphism and decoupling into the analysis may provide more comprehensive views of the dynamics involved in organizational changes and policymaking.

The criticism to institutional isomorphism is associated with the “paradox of embedded agency” (Seo and Creed, 2002): if the actions of organizational actors are constrained by taken-for-granted institutions, how and why can the actors induce institutional changes (Horton and Wanderley, 2018)? Institutional theory has been further developed, largely for a better understanding of the paradox of embedded agency. For instance, the concepts of both institutional entrepreneurship (DiMaggio, 1988; Battilana et al., 2009) and institutional work (Lawrence and Suddaby, 2006) attempt to address the purposive and proactive action of individuals and organizations to create, maintain and disrupt institutions. The institutional logics perspective further explains how institutions both enable and constrain actions in a multi-institutional environment (Thornton et al., 2012).

Alternative approaches

As a response to the theoretical limits of policy isomorphism, some alternative approaches to analyzing education policymaking in the context of globalization have been proposed. One commonly used strategy is to integrate institutional isomorphism theory with other supplementary theoretical perspectives. For instance, in a literature analysis, Cai and Mehari (2015) identify several weaknesses of neo-institutionalism and list a number of other theories applied to supplement neo-institutionalism in empirical higher education studies (Table 1).

Without moving outside the framework of institutional theory, the combination of both institutional isomorphism and the institutional logics perspective may offer the most promising approach to gain a comprehensive and deep understanding of the dynamics of education policymaking in the processes of globalization. This is because the insights of institutional logics have the advantage in terms of explaining how an environment with multiple and conflicting institutional pressures enables and constrains action via the macro-structure, local culture, and human agency (Thornton et al., 2012). This makes up for the weakness

Table 1 Supplementary theories to neo-institutionalism.

<i>Weaknesses of neo-institutionalism</i>	<i>Supplementary theories</i>
Lack of attention to the technical environment	Resource dependency theory Population ecology theory Cultural and socioeconomic factors Globalization perspective (highlighting the economic and technical impacts of globalization) Customer choice theory World system theory and dependency theory Administrative lattice Monopsony
Little attention paid to the role of human agency in institutional changes	Sensemaking Rational system theory Principal-agency model Actor's perspective Negotiated order theory Bourdieu's theory of practice Organizational change theory Political, teleological, and cultural models Professional theory Self-referential theory Transformative approach Social theories Dialectical theory Instrumental perspective
Less sensitive to geographical distance	Network theory
Over-emphasis on one organizational field	Allomorphism (organizations are embedded in multiple organizational fields)

Adapted from Cai and Mehari (2015).

of institutional isomorphism. Such a combination has proven useful in an analysis of the evolution of university strategies in the organizational field of Swedish higher education institutions (Holmberg and Hallonsten, 2015). Such a combined approach is likely to be relevant for studies of education policies in a global context.

There are also efforts to transform the analytical framework of policy isomorphism. Zapp and Ramirez (2019) suggest using the concept of a global higher education regime to understand the processes of global higher educational integration with regard to the “global education system” (Meyer, 2006). The regime concept focuses on the integrative mechanisms underlying the interactions of diverse actors across nations that are involved in the construction of a global higher education regime, “instead of conceptualizing integration as a process leading to a single ‘system’ or ‘society’ of higher education systems” (p. 474), as the emphasis of institutional isomorphism. Zapp and Ramirez (2019) theorize the global higher education regime concept by specifying three mechanisms of regime construction, namely discursive, normative, and regulative mechanisms, which, in combination, “transcend local-global and inter-state distinctions, thus facilitating integrated ‘oneness’ beyond growing inter-national ‘similarity’” (p. 474).

Conclusions

As a global education system is becoming more evident, policy isomorphism may serve as a useful lens with which to analyze national education policymaking in such a context. It is especially effective in explaining why education policies across countries tend to become similar. However, when applying policy isomorphism in empirical analysis, one must be aware of its theoretical limits. For instance, the theory of isomorphism, as part of neo-institutionalism, is not designed to analyze institutional change, conflicting institutional elements, and the role of human agency. The analytical framework of policy isomorphism is suitable for analyzing the influence of globalization on national education policy when three conditions are met: (1) the policy problem is primarily concerned with institutional pressures, (2) the global environment of policymaking is largely dominated by a single set of norms, and (3) policymakers/consultants are involved in global networks. However, both the global and national systems of education are becoming increasingly complicated. Instead of one set of predominating norms, there are multiple or even conflicting institutional logics in these systems (Lepori, 2016; Cai and Mountford, 2021). Also, education policymaking processes are not merely influenced by internationally shared norms but also by national interests, as well as local economic conditions (Kogan, 2018).

This does not mean that policy isomorphism, as an analytical lens, is out of date. Regardless of changes and complexities in society, there are always forces driving policies and organizations to become similar. My suggestion is that, when applying policy isomorphism in policy analysis, researchers should keep broader institutionalist perspectives in consideration and particularly bear in mind the following issues. First, the policy strategies and actual processes of implementation can be loosely coupled or decoupling, which is another core concept of neo-institutionalism. Second, the recent development of institutional theory has offered many useful insights to supplement neo-institutionalism. In a policymaking context in which there are mingling institutional logics and actors’ roles are crucial, researchers should consider alternative institutional theory approaches (e.g., institutional logics and institutional entrepreneurship) or integrate these theoretical insights with policy isomorphisms. Third, the lens of policy isomorphism helps to not only explain the convergence of education policies across nations but also observe those aspects that are not isomorphic. Eventually, in the practices of education policymaking, we must “work with and against the pressures of globalization in education” (Henry et al., 1999, p. 85).

References

- Alford, R., Friedland, R., 1985. *Powers of Theory: Capitalism, the State, and Democracy*. Cambridge University Press, Cambridge.
- Altbach, P.G., Reisberg, L., 2018. Global trends and future uncertainties. *Change* 50, 63–67.
- Baker, D.P., Wiseman, A.W., 2006. *The Impact of Comparative Education Research on Institutional Theory*. Emerald Group Publishing Limited, Oxford.
- Battilana, J., Leca, B., Boxenbaum, E., 2009. How actors change institutions: towards a theory of institutional entrepreneurship. *Acad. Manag. Ann.* 3, 65–107.
- Beckert, J., 2010. Institutional isomorphism revisited: convergence and divergence in institutional change. *Socio. Theor.* 28, 150–166.
- Beech, J., 2009. Who is strolling through the global garden? International agencies and educational transfer. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer Netherlands, Dordrecht.
- Boxenbaum, E., Jonsson, S., 2017. Isomorphism, diffusion and decoupling: concept evolution and theoretical challenges. In: Greenwood, R., Oliver, C., Lawrence, T., Meyer, R.E. (Eds.), *SAGE Handbook of Organizational Institutionalism*, second ed. SAGE, London.
- Brint, S., Proctor, K., Murphy, S.P., Turk-Bicakci, L., Hanneman, R.A., 2009. General education models: continuity and change in the U.S. Undergraduate Curriculum, 1975–2000. *J. High. Educ.* 80, 605–642.
- Brown, K.D., Stevick, D., 2014. The globalizing labor market in education: teachers as cultural ambassadors or agents of institutional isomorphism? *Peabody J. Educ.* 89, 4–16.
- Cai, Y., Liu, C., 2020. The role of university as institutional entrepreneur in regional innovation system: towards an analytical framework. In: Preto, M.T., Daniel, A., Teixeira, A. (Eds.), *Examining the Role of Entrepreneurial Universities in Regional Development*. IGI Global, Hershey, PA.
- Cai, Y., Mehari, J., 2015. The use of institutional theory in higher education research. In: Huisman, J., Tight, M. (Eds.), *Theory and Method in Higher Education Research III*. Emerald, Bingley.
- Cai, Y., Mountford, N., 2021. Institutional logics analysis in higher education research. *Stud. High Educ.* 1–25.
- Cai, Y., Yan, F., 2011. Organizational diversity in Chinese private higher education-an institutional perspective. In: Teixeira, P., Dill, D. (Eds.), *Public Vices, Private Virtues? Assessing the Effects of Marketization in Higher Education*. Sense Publisher, Rotterdam.
- Cai, Y., 2010. Global isomorphism and governance reform in Chinese higher education. *Tert. Educ. Manag.* 16, 229–241.
- Cai, Y., 2012. Traditional reform philosophy and challenges of higher education reforms in China. *Int. J. Humanit. Soc. Sci.* 2, 60–69.

- Carney, S., Rappleye, J., Silova, I., 2012. Between faith and science: world culture theory and comparative education. *Comp. Educ. Rev.* 56, 366–393.
- Chien, S.S., 2008. The isomorphism of local development policy: a case study of the formation and transformation of national development zones in post-Mao Jiangsu, China. *Urban Stud.* 45, 273–294.
- Croucher, G., Woelert, P., 2016. Institutional isomorphism and the creation of the unified national system of higher education in Australia: an empirical analysis. *High Educ.* 71, 439–453.
- Dale, R., 1999. Specifying globalization effects on national policy: a focus on the mechanisms. *J. Educ. Pol.* 14, 1–17.
- Dimaggio, P.J., Powell, W., 1983. The iron cage revisited: institutional isomorphism and collective rationality. *Am. Socio. Rev.* 42, 147–160.
- Dimaggio, P.J., 1988. Interest and agency in institutional theory. In: Zucker, L.G. (Ed.), *Institutional Patterns and Organizations: Culture and Environment*. Ballinger, Cambridge, MA.
- Dobbins, M., Knill, C., 2009. Higher education policies in central and eastern Europe: convergence toward a common model? *Governance* 22, 397–430.
- Drori, G.S., 2008. Institutionalism and globalization studies. In: Greenwood, R. (Ed.), *The SAGE Handbook of Organizational Institutionalism*. SAGE, London.
- Green, A., 1999. Education and globalization in Europe and East Asia: convergent and divergent trends. *J. Educ. Pol.* 14, 55–71.
- Greener, I., 2002. Theorizing path-dependency: how does history come to matter in organizations? *Manag. Decis.* 40, 614–619.
- Greenwood, R., Hinings, C.R., 1996. Understanding radical organizational change: bring together the old and new institutionalism. *Acad. Manag. Rev.* 21, 1022–1054.
- Gunter, H.M., Mills, C., 2017a. Professionals as consultants in and outside of companies. In: Gunter, H.M., Mills, C. (Eds.), *Consultants and Consultancy: The Case of Education*. Springer, Cham.
- Gunter, H.M., Mills, C., 2017b. Researchers as consultants in and outside of universities. *Educational G.* In: Gunter, H.M., Mills, C. (Eds.), *Consultants and Consultancy: The Case of Education*. Springer, Cham.
- Hallonsten, O., Hugander, O., 2014. Supporting “future research leaders” in Sweden: institutional isomorphism and inadvertent funding agglomeration. *Res. Eval.* 23, 249–260.
- Hannan, M.T., Freeman, J., 1977. The population ecology of organizations. *Am. J. Sociol.* 82, 929–964.
- Hawley, A.H., 1968. Human ecology. In: Sills, D.L. (Ed.), *International Encyclopedia of the Social Sciences*. Macmillan, New York.
- Henry, M., Lingard, B., Rizvi, F., Taylor, S., 1999. Working with/against globalization in education. *J. Educ. Pol.* 14, 85–97.
- Holmberg, D., Hallonsten, O., 2015. Policy reform and academic drift: research mission and institutional legitimacy in the development of the Swedish higher education system 1977–2012. *Eur. J. High Educ.* 5, 181–196.
- Horton, K.E., Wanderley, C.D.A., 2018. Identity conflict and the paradox of embedded agency in the management accounting profession: adding a new piece to the theoretical jigsaw. *Manag. Account. Res.* 38, 39–50.
- Howlett, M., Ramesh, M., Perl, A., 2009. *Studying Public Policy: Policy Cycles & Policy Subsystems*. Oxford University Press, Oxford.
- Ingram, P., Clay, K., 2000. The choice-within-constraints new institutionalism and implications for sociology. *Annu. Rev. Sociol.* 26, 525–546.
- Ion, G., Marin, E., Proteasa, C., 2019. How does the context of research influence the use of educational research in policy-making and practice? *Educ. Res. Pol. Pract.* 18, 119–139.
- Jameson, J., 2012. Leadership values, trust and negative capability: managing the uncertainties of future English higher education. *High Educ. Q.* 66, 391–414.
- Jarvis, D.S.L., 2014. Policy transfer, neo-liberalism or coercive institutional isomorphism? Explaining the emergence of a regulatory regime for quality assurance in the Hong Kong higher education sector. *Pol. Soc.* 33, 237–252.
- Joo, Y.H., Halx, M.D., 2012. The power of institutional isomorphism: an analysis of the institutionalization of performance-based pay systems in Korean National Universities. *Asia Pac. Educ. Rev.* 13, 281–297.
- Kay, A., 2005. A critique of the use of path dependency in policy studies. *Publ. Adm.* 83, 553–571.
- Kim, M., Boyle, E.H., 2012. Neoliberalism, transnational education norms, and education spending in the developing world, 1983–2004. *Law Soc. Inq.* 37, 367–394.
- Kogan, M., 2018. *Educational Policy-Making: A Study of Interest Groups and Parliament*. Routledge, London.
- Lawrence, T.B., Suddaby, R., 2006. Institutions and institutional work. In: Clegg, S. (Ed.), *The Sage Handbook of Organization Studies*, second ed. SAGE, London.
- Lechner, C.M., Tomasik, M.J., Silbereisen, R.K., 2016. Preparing for uncertain careers: how youth deal with growing occupational uncertainties before the education-to-work transition. *J. Vocat. Behav.* 95–96, 90–101.
- Lepori, B., 2016. Universities as hybrids: applications of institutional logics theory to higher education. In: Huisman, J., Tight, M. (Eds.), *Theory and Method in Higher Education Research*. Emerald, Bingley.
- Letendre, G.K., Baker, D.P., Akiba, M., Goesling, B., Wiseman, A., 2001. Teachers’ work: institutional isomorphism and cultural variation in the U.S., Germany, and Japan. *Educ. Res.* 30, 3–15.
- Levy, D.C., 2006. How private higher education’s growth challenges the new institutionalism. In: Meyer, H.D., Rowan, B. (Eds.), *The New Institutionalism in Education*. State University of New York Press, Albany.
- Maman, D., Rosenhek, Z., 2020. Facing future uncertainties and risks through personal finance: conventions in financial education. *J. Cult. Econ.* 13, 303–317.
- Massey, A., 2009. Policy mimesis in the context of global governance. *Pol. Stud.* 30, 383–395.
- Meyer, J.W., Rowan, B., 1977. Institutionalized organizations: formal structure as myths and ceremony. *Am. J. Sociol.* 83, 340–363.
- Meyer, J.W., Scott, W.R., 1983. *Organizational Environments: Ritual and Rationality*. Sage, Beverly Hills, London.
- Meyer, J.W., Ramirez, F.O., Robinson, R., Boli-Bennett, J., 1977. The world educational revolution, 1950–1970. *Sociol. Educ.* 50, 242–258.
- Meyer, J.W., 1971. Economic and political effects on national educational enrollment patterns. *Comp. Educ. Rev.* 15, 28–43.
- Meyer, J.W., 1977. The effects of education as an institution. *Am. J. Sociol.* 83, 53–77.
- Meyer, J.W., 2006. Foreword. In: Baker, D., Wiseman, A. (Eds.), *The Impact of Comparative Education Research on Institutional Theory*. JAI Press/Elsevier, Oxford.
- Moutsios, S., 2009. International organizations and transnational education policy. *Compare* 39, 469–481.
- Owen-Smith, J., Powell, W.W., 2008. Networks and institutions. In: Greenwood, R., Oliver, C., Sahlin, K., Suddaby, R. (Eds.), *The SAGE Handbook of Organizational Institutionalism*. SAGE, Los Angeles, London.
- Pfeffer, J., Salancik, G.R., 1978. *The External Control of Organizations*. Harper & Row, New York.
- Raadschelders, J.C.N., 1998. Evolution, institutional analysis and path dependency: an administrative-history perspective on fashionable approaches and concepts. *Int. Rev. Adm. Sci.* 64, 565–582.
- Rakic, V., 2001. Converge or not converge: the European Union and higher education policies in The Netherlands, Belgium/Flanders and Germany. *High Educ. Pol.* 14, 225–240.
- Rizvi, F., Lingard, B., 2009. *Globalizing Education Policy*. Routledge, London.
- Scharpf, F.W., 1997. *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*. Westview Press, Boulder, Colo.
- Schofer, E., Meyer, J.W., 2005. The worldwide expansion of higher education in the twentieth century. *Am. Socio. Rev.* 70, 898–920.
- Seo, M., Creed, D.W.E., 2002. Institutional contradictions, praxis and institutional change: a dialectical perspective. *Acad. Manag. Rev.* 27, 222–247.
- Seyfried, M., Ansmann, M., Pohlentz, P., 2019. Institutional isomorphism, entrepreneurship and effectiveness: the adoption and implementation of quality management in teaching and learning in Germany. *Tert. Educ. Manag.* 25, 115–129.
- Sirat, M.B., 2010. Strategic planning directions of Malaysia’s higher education: university autonomy in the midst of political uncertainties. *High Educ.* 59, 461–473.
- Spring, J.H., 2005. *Political Agendas for Education: From the Religious Right to the Green Party*. L. Erlbaum Associates.
- Suddaby, R., Foster, W., Mills, A.J., 2013. Historical institutionalism. In: Buchell, M., Wadhvani, D. (Eds.), *Organizations in Time: History, Theory, Methods*. Oxford University Press, Oxford.
- Susnea, I., Pecheanu, E., Tudorie, C., Cocu, A., 2014. The Education for Creativity—The Only Student’s Tool for Coping with the Uncertainties of the Future Prague. MAC Etel 2014—International Conference on Education, Teaching and e-Learning.

- Thornton, P.H., Ocasio, W., Lounsbury, M., 2012. *The Institutional Logics Perspective: A New Approach to Culture, Structure and Process*. Oxford University Press, Oxford.
- Tuttle, B., Dillard, J., 2007. Beyond competition: institutional isomorphism in U.S. accounting research. *Account. Horiz.* 21, 387–409.
- Verger, A., Novelli, M., Altinyelken, H.K., 2012. Global education policy and international development: an introductory framework. In: Verger, A., Novelli, M., Altinyelken, H.K. (Eds.), *Global Education Policy and International Development: New Agendas, Issues and Policies*. Continuum, New York.
- Villadsen, A.R., 2011. Structural embeddedness of political top executives as explanation of policy isomorphism. *J. Publ. Adm. Res. Theor.* 21, 573–599.
- Wiseman, A.W., Astiz, M.F., Baker, D.P., 2014. Comparative education research framed by neo-institutional theory: a review of diverse approaches and conflicting assumptions. *Compare* 44, 688–709.
- Yang, R., 2003. Globalization and higher education development: a critical analysis. *Int. Rev. Educ.* 49, 269–291.
- Yang, R., 2019. Internationalization of Chinese higher education. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Singapore, Singapore.
- Zapp, M., Ramirez, F.O., 2019. Beyond internationalization and isomorphism—the construction of a global higher education regime. *Comp. Educ.* 55, 473–493.
- Zhu, X., Meng, T., 2018. Geographical leadership mobility and policy isomorphism: narrowing the regional inequality of social spending in China. *Pol. Stud. J.* 48, 806–832.

Knowledge economy meets development imaginaries

Alejandro Artopoulos, Universidad de San Andrés, Universidad de Buenos Aires, Comisión de Investigaciones Científicas, Victoria, Argentina

© 2023 Elsevier Ltd. All rights reserved.

The “rational” knowledge economy	281
The bubbling knowledge economy	282
Missing links of knowledge-based development (KBD)	283
The labyrinths of educational digital inclusion in the Global South	284
The translating power of “Creative Britain”	286
Conclusions	286
Uncited references	287
References	287

The term “Knowledge Economy” (KE) holds relevance in current discussions about human and economic development. Since 1996, when the OECD introduced it, it has acted as an umbrella concept that triggered various lines of public policy at the global and regional levels (Godin, 2006).

The concept emerged as an explanatory device for processes of structural change from industrial economies to those based on the use, creation and distribution of knowledge; and was intended to explain and justify a variety of public policies, ranging from economic to educational reforms.

Alongside public policies, forms of collective action have emerged based on the appropriation of such explanatory device. Social actors from civil society, industry, and academia have built a public-private agenda that supports their interests around projects for educational change based on the concept of knowledge economy.

The initial debates around the notion of KE assumed that the transition to KE was to be completed in advanced countries; and only then other countries could consider it as a reachable destination in their pursuit for development. Moreover, its outcome -the “knowledge society”—was considered an “objective” and “universal” fact—only a set of rational steps to be decoded by less advanced countries to initiate the needed transition.

Although the idea of KE suffered significant setbacks, such as the dotcom bubble in 2000–2002 or the subprime mortgage crisis in 2007–2008, escalations in structural changes did not always follow a market logic, as evidenced in the free software or open science movements. Moreover, the belief in the assumed benefits of a transition to KE maintained a steady presence in the policy agenda of a variety of countries around the world, showing that the concept still holds relevance.

Due to the persistent resilience of the concept, it is necessary to further develop non-linear explanations of the post-industrial transition to KE. The relative absence of alternative critical approaches to the notion of KE and its implications in economic and educational terms, even in advanced countries, contributes to the reproduction of unequal conditions of access to knowledge and its application to individual and collective life projects.

One of the biggest challenges in apprehending the idea of a KE is that the concept has become a moving target. Meanings shift as sociotechnical transitions become more complex, both when digital infrastructure becomes a private-public endeavor and when globalization blurs the national state as a valid unit of analysis of the symbolic sphere of the development process (Bhatt and MacKenzie, 2019).

In this essay, we will examine the goal of constructing a KE as a national endeavor that comprises symbolic and socio-material dimensions, analyzing discontinuities in specific structural and sociotechnical transitions. We propose the hypothesis that there is no single, universal KE as a desirable development destination per se. We assume, after the observation of case studies, that each nation-state puts its stamp or name on the construction site of its KE.

In their eagerness to promote public development policies, international organizations have oversimplified complex and unstable processes and made development imaginaries invisible. It is necessary to open the black box of KE to link it to symbolic work, so that each country can find its development path, avoiding the manifest dangers that platforms and algorithms present to the political and economic institutions necessary for sustainability. Society has reached a point where the technologies that build the KE also give body to the same sociotechnical material that lays the foundations of ignorance as a social practice (Latour, 2018).

I attempt a critical approach to KE that incorporates the multidirectionality of the transition to KE and the interrelation between economic-based infrastructural changes and socio-cultural appropriation strategies at the nation-state level. The individuation process during the late capitalism has modified the premises of development theories but not the diverse identity that supports them (Hui, 2020).

In the first section, we will examine the concept of KE in the context of its emergence in the second half of the 20th century, both in applied research by International Organizations such as the OECD and in their theoretical foundations.

In the second section, I will trace UNESCO’s humanistic response, contrasting it with the perspectives of the OECD’s monodisciplinary approach. The shift to the idea of a Knowledge Society (KS) in the first decade of the 21st century promoted an expansion

of the conceptual space of KE to establish frameworks that integrate economic and human development. UNESCO developed an interactive perspective of the economy-society relationship rather than the unidirectional one proposed by economic theory.

In the last two sections, from a critical perspective of linear approaches to the transition to KE, I develop the analytical framework on the missing cultural link between economics and education. I apply the concept of Sociotechnical Imaginaries to the interpretation of the Finnish and South Korean cases in the third section. In the final section, I evaluate educational reforms in various countries and analyze how they related with institutional development.

The “rational” knowledge economy

The notion of the knowledge economy, introduced by OECD (1996) as knowledge-based economies (KBE), has tried to describe economies whose core activities to develop and sustain long-term growth are characterized by knowledge. Research pursued by the organization to coin the term drew data from the most advanced economies and the transition processes of certain “exemplar” countries that have caught up, including Finland (Dahlman et al., 2006) and South Korea (Bank et al., 2001; Marginson, 2010). It tried to make sense of transformations in economic structure and innovation processes due to technological change triggered by the diffusion of microelectronics and the personal computer 20 years earlier.

KBE was defined as “economies which are directly based on the production, distribution and use of knowledge and information. This is reflected in the trend in OECD economies toward growth in high-technology investments, high-technology industries, more highly-skilled labor and associated productivity gains (...). In this view, investments in research and development, education and training and new managerial work structures are key” (OECD, 1996, p. 7). By focusing on the KBE, the OECD established a new growth strategy characterized by the use, creation and distribution of knowledge. The introduction of the concept also meant a significant shift in the epistemic communities of development politics, both in cases’ interpretation and policy recommendations. The core of advanced economies was no longer to be based on processing physical or material inputs dependent exclusively on traditional resources such as natural resources (e.g., land), labor, or capital (Haas, 1992).

Peter Drucker, management consultant and founder of the management field, served as a source of inspiration for OECD. In 1959 he coined the term “knowledge worker” to distinguish a new type of labor productivity. A new management, a new economy, and therefore a new society, the knowledge society, would emerge (Drucker, 1959). Although the influential thinking of Drucker was decisive in crafting the narrative of KE in the public sphere, his texts were based on primitive forms of KE prior to the micro-computing revolution and the spread of the internet. Thus, he did not go beyond prospective arguments.

Among the foundations of the OECD, we can also find Machlup’s estimates on the growth of knowledge industries, Gibbons et al.’s proposal on the new mode of knowledge production (Gibbons, 1997), the theory of National Innovation Systems (Freeman et al., 1987), Lundvall’s *Learning Economy* (2010), and Romer’s endogenous growth theory (Romer, 1994). The first theorists of KE such as Mashlup or Bell turned their attention to information processing, knowledge production, and the creation of intangible assets. Nevertheless, intangible assets such as patents and other products of scientific and technological research and development in the form of explicit knowledge, do not account for tacit knowledge accumulated in knowledge-based organizations and communities of practice.

Still, early stages in the process of building the concept of the KE also enhanced human talent through the massive training of highly specialized manpower at informal education systems and in continuous training circuits where the knowledge of the workforce is updated. Furthermore, International organizations stimulated and followed the study of the evolution of National Innovation Systems (NIS) proposed by Christopher Freeman and Bengt-Åke Lundvall. They described the innovation process as an interactive relationship between private and public organizations and actors that evolved in time to a dense network of institutions “located within or rooted inside the borders of a nation state” (OECD, 1997; Lundvall, 2010).

Focused on the diffusion and modification of new technologies, the production of applied knowledge combined with theoretical and empirical production in the early 1990s formed a variety of epistemic cultures in IOs (Haas, 1992; Kallo, 2021). Nevertheless, these failed to factor into the equation what we know today as the creative industries or creative economy. Such a new line of inquiry opened from 2002 with the works of Hesmondhalgh and Florida (Florida, 2003; Hesmondhalgh, 2002). Indeed, the 1990s OECD’s “rational” approach to the knowledge economy encountered a blind spot in creative activities and startups when identifying how knowledge is distributed and circulated, only considering technology drivers.

Furthermore, networked information technologies combined with knowledge management and new economic theories triggered a “rational and control” mindset about development. Knowledge was now to be seen as input, output, and capital, even if imperfectly accounted for or understood. Public agencies were convinced that knowledge could be managed in sophisticated, rational ways, where diffusion of information technology was essential. Sociotechnical transition, in particular the organizational transformation after the first wave of internet use in the production system, operated as (partial) evidence of structural change. Ideas about Knowledge Management as a new and powerful lever for development gave, beyond the hype, a tinge of rationality and control to development policies.

The framework was completed at the macrosocial level with the role of the “information society” in distributing knowledge through computer and communication networks. Daniel Bell had observed in the late 1960s through the theory of the Postindustrial Society that the axis of advanced societies was shifting toward service activities (Bell, 1976). Bell claimed information rather than energy was the transforming resource. Nevertheless, Bell did not visualize the interactive relationships between economy and culture. Rather than considering the emergence of new economic actors such as startups, entrepreneurs, and creative industries,

Bell's theory turned out to be a dead-end theory of problems and contradictions of post-industrial capitalism. Still, and in the context of conceptualizing KEs, Bell's writings reinforced the assumption that endogenous abstract knowledge growth processes could determine KEs.

The bubbling knowledge economy

The idea of a KE model gained recognition as the momentum of the growth of IT companies as expressed in the value of their shares. Between 1995 and its peak in March 2000, the Nasdaq Composite stock market index grew by 400%.

The early stages of the popularization of Internet use fueled unrealistic expectations around explosive economic growth. The launch of Mosaic in 1993, the first web browser, and the spread of its successor web browsers during the following years gave computer users access to the World Wide Web, the first version of a service-based digital economy.

Momentum met its end, or a long pause, with the dotcom bubble, a stock market bubble caused by excessive speculation of Internet-related companies. Between March 2000 and October 2002, Nasdaq fell 78%, giving up all its gains during the bubble, and causing a drastic change in the optimistic expectations about the KE.

Drastic change in expectations raised critical voices that pointed to the importance of integrated more fully the economy-society relationship from approaches to technological change that account for the political, cultural, and social aspects that affect these seemingly exclusively economic development processes (Florida, 2003; Peters, 2003; Godin, 2004, 2006; Beech, 2005a,b; Dale, 2005; Olssen and Peters, 2005; Kenway et al., 2006; Sörlin and Vessuri, 2007).

The topic's relevance stimulated a "second wave of knowledge economy studies." Theoretical and empirical research production in different fields of the social sciences overflowed the boundaries and reformulated the strictly economic approaches of early applied research (Choong and Leung, 2021, p. 18).

In this new context, UNESCO (2005) published its strategy to transition to Knowledge Societies by placing the Knowledge Economy in a macro-social framework. UNESCO narrative introduced a sociological twist in the references to the Knowledge Society theories of Nico Stehr (Stehr, 1994), Manuel Castells' Informational Capitalism (1996), and Amartya Sen's capability theory approach to development (2000).

The UNESCO document attempts an interdisciplinary and integrative approach to human and institutional development aspects. Identifying a plural destiny for "Knowledge Societies," the document mentions that the knowledge existing in all societies may or may not be mobilized for application by the knowledge economy model. It asks, what are the limits to the commodification of knowledge?

UNESCO (2005) stated that societies must be based on four pillars: freedom of expression, universal access to information and knowledge, respect for cultural and linguistic diversity, and quality education for all. It is the first significant report by an international body to define the digital divide as a form of multidimensional inequality on the global public agenda.

The document incorporates the idea developed by Warschauer (2003) that there is not one digital divide but several: material, geographical, age, gender, linguistic, educational-cultural, accessibility. The UNESCO proposal set out an agenda that has become even more urgent in times of the Covid 19 pandemic (Van Dijk, 2020).

Inequality in knowledge societies is not exclusively technological in origin, but the material basis of digital disconnection is combined with cognitive and socio-cultural aspects of digital inequality. This cognitive divide includes the capacity to interpret information on the internet, what we today call applied computational thinking, and educational and communicational aspects of inequality, such as the availability of digital content in the local language appropriate to the cultural context (Denning, 2017).

Influenced by the trilogy *The Information Age* published in 1996 by Manuel Castells, UNESCO cited the notion of Network Society to serve as a theoretical anchor for the applied narrative of the transition from the Information Society to Knowledge Societies. Castells goes a step further in explaining the triangular interdependence between economy-technology-culture.

For UNESCO, unlike the OECD, the new social infrastructure of the KE is the techno-social networks and not only the circulation of information. The KE is contained in the Network Society. The production of goods and services in the KE requires information and knowledge to be processed by the body of the workforce connected by the internet and, nowadays, by platforms (UNESCO, 2005, p. 49).

What characterizes the sociotechnical nature of the KE is "the application of such knowledge and information to knowledge generation and information processing/communication devices, in a cumulative feedback loop between innovation and the uses of innovation" (Castells, 1996, p. 31). Therefore diffusion of technology amplifies their power as it becomes appropriated and redefined by its users. Computers, information networks and algorithms are not simply tools to be handled, but digital infrastructure to be developed. Creating and manipulating symbols (culture) are strongly related to the capacity to produce and distribute goods and services (economy).

Castells based his theory of knowledge accumulation on Nonaka and Takeuchi's (1995) work on knowledge management, that introduced the analysis of the circulation of tacit dimension in the production of knowledge through the creation of new products and services. Nonaka and Takeuchi described how new innovative products were conceived in an iterative process of researching tacit knowledge as a primary source of value. They emphasized the growing importance of personalization and socialization of knowledge, to the detriment of knowledge codified in objects, text, or patents (Nonaka and Takeuchi, 1995).

Knowledge accumulation in KE occurs primarily through knowledge embodied in practices, capabilities, and skills. Disciplinary or scientific knowledge, while an input, is just the codified part of the source of value. The primary source of value is the tacit

knowledge embodied in intellectual practices applied to the production and distribution of goods and services assisted by online information.

Although Information and communication technologies (ICTs) act as a driver of change since they provide the material basis for human activities, KE involves an economy based on the manipulation of signs. That is why the main factor to be considered is the human intellectual activity around manipulating signs that require a reconfiguration of labor under a new technical division of creative activities. Thus informational capitalism relies as much on ICTs as on human creativity, tacit knowledge, and creative talent (Castells, 1996, p. 93; Lash and Urry, 1998; Lash, 2002).

Castells (1996) developed a model to interpret the transition from an industrial to informational capitalism, which identified common patterns of change observable in different types of societies, both central and peripheral. Informationalization describes a common process of structural change that, regardless of the level of development of each country, occurs and is expressed in each national cultural sphere according to its identities and trajectories.

Informationalism according to Castells “cannot be equated to the emergence of the service economy” (Castells, 1996, p. 100). For the author, the socio-technical transition between the industrial economy and the knowledge economy crosses sectoral activities of intellectual work, management and design, supported by the power of information technologies, resulting in informational agriculture, informational manufacturing, and informational services.

By connecting economics and culture, the Theory of Informational Capitalism brings into focus a third element of development: institutions. In addition to economic and human development, the construction of each state is based on institutional development. In the next section, we try to address the missing cultural link between economics and education institutions by applying the concept of Sociotechnical Imaginaries to selected cases.

Missing links of knowledge-based development (KBD)

Despite UNESCO’s efforts to reconcile economic development with human development, the third wall of development, the institutional one, persists in an unbridgeable divide. How to address the institutional dimension of development divides?

Knowledge discourses in policy documents of advanced countries such as Canada, the UK, and Australia have suggested that Institutions must promote generic skills acquisition in an individualist and universalist model of education. Universities were under increasing pressure from companies to respond to training and professional updating demands. The vision shared by IOs and companies is that the ideal type of a knowledge worker is a “technopreneur” that grasps innovation discovery (Al-Fehaid and Shaili, 2021; Fenwick, 2010).

KBD strategies for reassuring the development of developed countries rely on techno-determinism derived from a Schumpeterian approach to innovation as preceding and “driving” the economy—a hard agenda on institution building where educational goals are vaguely directed at upskilling, emphasizing control and measurement (Powell and Snellman, 2004).

By placing the techno-entrepreneur at the center as the determining factor of the KE the narrative of KBD underestimate or fail to consider dense aspects of innovation culture, such as the maker movement and open source communities of practice, where knowledge assets are collectively owned and driven by the individual will. The institutionalist agenda of development policies overlooked the interactive relationship between economy and culture (Stiglitz, 1999; Peters, 2003).

Ironically, large technology companies such as IBM or Microsoft understood sooner or later that they had to integrate such communities of technological practice and their cultures into their business strategies. This led to the development of management research on open innovation (Chesbrough, 2007).

In order to explain cultural preconditions for the development of economies based on knowledge, Peters presented the concept of Knowledge Cultures as the hubs of meaning that bind collective action in the transitions to KE, “shared practices of epistemic communities” that embody culturally preferred ways of doing things at a supranational level (Peters, 2003, p. 375).

Knowledge production requires a grounded set of shared ideas that depend upon certain cultural conditions, including trust, reciprocal rights, and responsibilities between different knowledge partners, institutional regimes, and strategies. However, it is unclear at what social levels communities of practice share knowledge and create their “ways of doing things.” While Knowledge Cultures assume supranational spaces, Innovation Systems, on the other hand, assume national or even regional spaces.

In that sense, the revision of the theory of National Innovation Systems adds to the interactive institutional processes the learning capacities to connect diversity. KE depends on a delicate balance of trust, social distance, and feelings. Attempts to reduce this complex interaction to “exchange of information” do not capture the most critical dimensions (...) lead to normative conclusions and recommended arrangements that differ from those reached on the basis of standard economics analysis” (Lundvall, 2010, p. 20).

The case of Finland’s transition to KE was often taken to illustrate a virtuous process. The case’s attraction came from a leapfrog from an industrial economy with little or no competitiveness to a KE. Nevertheless, while the IOs present the case as exemplary, other less linear analyses show dynamic and interactive aspects and even shadowed areas of development “models” (Himanen and Castells, 2004; Dahlman et al., 2006; Schienstock, 2007).

Himanen and Castells distinguished between egalitarian national strategies to KE (informationalism) and individualistic and exclusionary forms. In the egalitarian strategy, the Finnish welfare state built-up public goods leading to accelerated material wealth creation while integrating and equalizing opportunities for its population.

The second form (USA) maintained its informational leadership without human development, leading to the decline of its middle class, as can be observable in the “Rust Belt,” and a scarce presence of public goods that can integrate poor immigrant workers, often also illegal (Himanen and Castells, 2004; Schienstock, 2007; Saxenian, 2014).

A comparison between Finland and the United States reveals the shape of the ideal transition but also hides the more viscous side of transitions. When comparing Finland with other paralleled, less extreme cases, such as South Korea, we can observe a pattern describing the construction of a national institutional superstructure based on a sociotechnical imaginary.

Sociotechnical Imaginaries, according to Jasanoff and Kim (2013) are “collectively imagined forms of social life and social order reflected in the design and fulfillment of nation-specific scientific and/or technological projects” (Jasanoff and Kim, 2013, p. 120). While research has largely applied this theoretical framework to energy policy formulation processes, we have found cases of policy formulation processes to promote the knowledge economy. These cases are examined in the following paragraphs.

The national narrative of Finland’s development project was based on the “information society” as a national destination. A hacker vision of building an egalitarian post-Soviet “open” society. Himanen (2014), one of the curators of this narrative, turned it into theory in a well-known essay “The Hacker Ethic, and the Spirit of the Information Age,” in which he attempts to re-seed the Weberian formula for the 21st century. A narrative rooted in material and real achievements such as the Linux Operating System financed by university scholarships put the country at the center of an epistemic community that became a game-changer for the global IT industry (Castells and Himanen, 2002).

Himanen’s efforts as a curator of the national narrative later evolved into an attempt to construct a global proposal based on globalist readings of the hacker spirit and the free software movement. This proposition was shared by Peters (2003) when he pointed to the free software movement as one of the components neglected by the KE formulators. Proposals that ignored the national roots of the construction of socio-technical imaginaries (Himanen and Committee for the Future in Parliament of Finland, 2004; Peters, 2009; Sahlberg, 2011).

South Korea, on the other hand, presented dissociation and discontinuities. It developed its economy in two distinct stages: advanced industrial capitalism through a long path of improving industrial labor productivity concentrated in its industrial conglomerates (Chaebols), mainly under a dictatorial regime (1953–1979). A second stage after the 1997 crisis described a transition to a KE promoted by the state under the imaginary of the “Smart Nation.”

South Korea’s “last mile” knowledge economy development strategy consisted of reforms to institutional incentives for creativity-oriented smart education investment and cultural openness, with enhanced knowledge exchanges between universities, firms, local governments, and research institutes. Between 2008 and 2013, following president Lee Myung-bak’s cabinet reorganization, the Ministry of Trade, Industry, and Energy was named the Ministry of Knowledge Economy (Biggart and Guillén, 1999; Bank et al., 2001; Marginson, 2010; OECD, 2015).

Finland and South Korea, despite their differences, are cases of sustained economic growth and an accelerated transition to KE that transformed their innovation and educational systems based on a common national narrative. Sociotechnical imaginaries supported both political consensus and a set of national symbolic elements that built a cohesive identity that contributed to the wide appropriation of social change. While both are KE, each has its way of building its own culture of knowledge.

The cases of Finland and South Korea describe streamlined narratives with the power to produce development. But what happens when sociotechnical imaginaries do not achieve the expected socio-economic outcomes? What can we find in failed educational reforms about the missing cultural link between economics and education?

The labyrinths of educational digital inclusion in the Global South

During the wave of neoliberal education reforms in the 1990s, when decentralization and choice through privatization were promoted, linear approaches of the transition to KE promoted economic policies without considering appropriation processes in Brazil and Argentina. Centralized bureaucracies were dismantled from educational organizations without ensuring capacities for autonomy and flexibility (Beech, 2011).

The education reforms implemented in several Latin American countries in the 1990s were based on the idea that education systems had to adapt to the “knowledge society”. Under the influence of views mainly imported through international agencies, national institutions suffered an increasing division of labor in the top-down implementation process. While experts defined policy goals and curriculum guidelines, teachers were downgraded to mere technicians, and local and national educational institutions were reduced to plain transmission chains (Beech, 2005a,b).

Paradoxically in the quest to produce the transition to KE, the IOs tailored policy implementation in such a way, that the result was a weakening of existing institutions by reducing the autonomy and flexibility of teachers, principals, and education officials. Thus, they reduced the capacity to manage the diversity of the institutions that were supposed to expand them. In the absence of teachers’ and principals’ autonomy, with narratives that empower their agency, the transition to national education systems ready for KE could result in an Orwellian nightmare (Wilson and Segal, 2005; Zuboff, 2019).

Cases such as Brazil and Argentina in the 1990s offered evidence of dangerous outcomes of structural change when nation-states choose symbolic passivity. Because passivity did not lead to a narrative vacuum but to the almost mechanical reproduction of the imaginaries that gave rise to the notion of KE. This is the enthronement of the figure of the techno-entrepreneur together with the promotion of the “innovative culture” whose main reference is located in the Silicon Valley.

As was required of the rest of society, teachers also had to become entrepreneurs, to be innovative. They failed to achieve appropriation of economic or educational reforms, but they opened the door to tearing apart the previous institutions that sustained the social relations of knowledge.

A decade later, weak capacities eroded new digital inclusion plans when countries such as Argentina, Peru, and Nigeria, tried to bridge the digital divide and modernize their educational systems by deploying 1:1 programs. The exception was the case of Plan Ceibal in Uruguay, probably the most successful implementation of Negroponte's OLPC project (One Laptop per Child) (OECD, 2010; Rivoir, 2012; Warschauer et al., 2014; Artopoulos, 2020).

The OLPC Project, born in 2005, promoted the idea of one-to-one computing in education institutions in developing countries. The independent attempt triggered the quest for a stabilized classroom device that finished with Chromebook teachers' acceptance in 2015 in leading OCDE countries, adding a computational layer to school sociomateriality for good (Ames, 2019).

One-to-one computing in education (sometimes abbreviated as "1:1") allowed new teaching practices (ex. Flipped Teaching) based on student's activities on a computational device both in class and out promoting more autonomous learning based in teaching platforms before the COVID Pandemic (Artopoulos et al., 2020).

Although Latin American countries such as Chile, Uruguay, and Costa Rica reached considerable leaps in public goods, educational reforms did not achieve enough sustainability and/or quality standards. Informationization of the region advanced in the primary sectors (mining, agriculture, construction) without producing significant changes in the labor force or national innovation systems (Castells, 2003; Calderón, 2015; Artopoulos, 2020).

Decentralization of state apparatuses facilitated conditions for private investments in education. The absence of a soft institutional agenda and adequate regulations caused a legitimacy crisis in the expansion of education systems, as was the case of Chile with the so-called "Penguin Revolution." The social movement gave rise to a new political coalition that changed the structure of the Chilean political system. The social conflicts of the knowledge society seem to provoke strong political earthquakes (Calderon and Castells, 2014).

Civil society mobilization around the big tech industry was the other side of "laissez-faire" approaches to education reform based on the techno-entrepreneur imaginary. ONGs tried to set up an agenda of educational reform that fits Big-Tech industry interests in alliance with certain parts of academia. Two relevant cases describe how currents of opinion born in the USA become social movements of an international scale that seek reforms in basic education in consecutive moments of technological change. They are still relevant today because they significantly impacted dominant narratives of educational change.

The first wave of "digital culture" and 1 to 1 model between 2002 and 2013 was encouraged by the educational use of new web services such as Wikipedia and participatory digital culture through blogs. At the beginning of the "convergence culture" in 2002, the Partnership for 21st Century Skills (P21) was founded, advocating for integrating generic skills such as critical thinking, problem-solving, and communication into teaching core academic subjects. It represented a mixture of creative industries, media, and information technology multinational companies such as AOL Time Warner, Apple Computer, Cisco Systems, Dell Computer Corporation, Microsoft Corporation, and SAP (Jenkins, 2006).

The second wave, "1 h of code," promoted coding for all and introduced computational thinking as fundamental knowledge in the curriculum. Ten years later, in 2013, Code.org and "Hour of Code Challenge" platformed the social media scene and heated the debate around coding at the basic education level. Non-profit organizations started their activity by promoting learning computer science through a website with free coding lessons and tutorials. It also has targeted schools to encourage them to include more computer science classes in the curriculum. In contrast, the code.org initiative seemed to be more grassroots but with a more tech-driven agenda and less sensitive to curriculum and pedagogical issues.

As the new century progressed, other attempts to transition to KE offered slightly more sophisticated cases of failure. In "Imagining a Modern Rwanda: Sociotechnological Imaginaries, Information Technology, and the Postgenocide State," Bowman (2015) illustrated how a Rwanda government in need of KE models found a mirror to look at with images that sound attainable.

Rwanda was a case of importing KE models without an IOs intervention or abstract models. Relying on the term "sociotechnical imaginaries," Bowman explained how the new democratic government of Rwanda announced its ambition to be an "African Singapore." Rwanda's government unsuccessfully tried to build a KE by borrowing the socio-technological Imaginary of an autocratic Asian nation. The KE Singaporean narrative of the "smart nation" was abandoned without solid local roots (Jasanoff and Kim, 2015).

When addressing digital inclusion It is necessary to consider the historicity of KE in terms of changes in the socio-materiality of digital learning environments and work, commerce, or entertainment. Since the OECD defined it in 1996, techno material foundations of education have evolved dramatically. In 2005 1.02 billion people were connected to the internet, only 15.76%. In 2021, 16 years later, those connected amounted to 4.9 billion, exceeding 62.5% of the world's population (ITU, 2021).

During the COVID-19 pandemic, many education systems have been able to continue providing lessons and teachers assistance to students, not only because schools and educators have shown remarkable resilience in trying to compensate for the crisis, but also because techno material conditions of learning had changed over the years (Vincent-Lancrin et al., 2022).

The Pandemic Crisis showed the value of smartphones, introduced in 2006 when the iPhone was launched, maintaining student-teacher contact even in precarious conditions. Paradoxically at the same time, it revealed new layers of digital divides that were not contemplated 16 years earlier in UNESCO's proposal (Warschauer, 2003; Toyama, 2015; van Dijk, 2020).

Algorithms and intelligent technologies turn literacy into multiliteracies and the meaning of digital progress. Platforms are increasing uncertainty and risk instead of providing rationality and control. We need to address how new attempts to build more egalitarian KE are translated into policies (Beck, 1998; OECD, 2016, 2021; Smicek, 2017, 2021).

The translating power of “Creative Britain”

Advanced countries with a degree of bureaucratic capacities, such as Canada, Australia, or the United Kingdom, transition to KE allowed renewed airs of change in education systems. In the presence of adequate autonomous capacities, both teachers and management teams could trigger virtuous processes of bottom-up school quality improvement from communities of practice (Fullan, 2002; Hargreaves, 2003).

As we will see in the present section, in these softer processes of transition toward KE, the construction of the institutional superstructure played a fundamental role in the convergence of change in innovation systems and education systems. Nevertheless, KS could become less egalitarian when top-down reform without the same capacities disregarded such institutional arrangements.

Each culture has its manner of understanding knowledge and learning. It is impossible to bridge the divide between the economic and the educational sides of the knowledge society by bypassing the cultural nature of the connection. In 2022 the concept of KE is still meaningful. Sustaining semantic weight is less a result of a consensus on an abstract model of KE than instead supported by the evidence of the above-mentioned national experiences such as South Korea, Finland, Canada, or Australia.

Perhaps the case of Creative Education under the “Creative Britain” policy sharply illustrates how complex forms of educational policies merge into development policies. How they intertwine and hybridize transversally with economic and cultural policies through national strategies of constructing their own “sociotechnical imaginaries” of KE.

Throughout the New Labor administration, between 1997 and 2007, Chris Smith, Secretary of State, promoted the KE in the form of creative industries and economic contribution of culture. Creativity was one of the *leitmotifs* of the political agenda of prime minister Tony Blair (Smith, 1998; Hewison, 2011; Mangabeira Unger et al., 2019).

Smith, head of the Department for Culture, Media and Sport, in 1998 set up Nesta, a decentralized innovation agency’s national endowment for science, technology, and the arts. Nesta was originally funded by a £250 million endowment from the UK National Lottery. New Labor’s institutional reform led to a substantial investment in the cultural infrastructure. The visual and performing arts, museums, and galleries were persistently promoted as a stimulus to national economic revival.

In this context, a national education commission chaired by Ken Robinson, then professor of arts education at the University of Warwick, issued the “All Our Futures: Creativity, Culture and Education” Report that called to create a national strategy focusing on creative and cultural education. It argued that the national debate on education in Britain had been expressed as a series of failed dichotomies—“as a choice between the arts or the sciences; the core curriculum or the broad curriculum; between academic standards or creativity; freedom or authority in teaching methods” (1999).

“Creative Education Trust” was an example of the implementation of transversal creative education/economy policies based on a national sociotechnical imaginary (“Creative Britain”) NESTA introduced a new model of organizations in charge of policy implementations. The trust worked in England’s post-industrial cities and coastal towns: areas of economic disadvantage and with a history of academic underachievement. Its methodology involved integrating a knowledge-rich curriculum with skills and creativity (CET, 2020).

There were also international effects of these creative education policies. Sir Ken Robinson, also known for having one of the most viewed TEDTalks, “Do schools kill creativity?” With 72 million views, became one of the first educational ‘influencers’ of a global epistemic community. A new figure converged with other educational change thinkers such as Nicholas Negroponte or Michel Fullan in the international arena.

Later, others followed Robinson, such as Salman Khan, Jeanette Wing, and Sugata Mitra. Another interesting case was Manuel Castells appointment as Minister of Universities of the Kingdom of Spain, which began a reform of upper education in 2020 that is still in progress.

Conclusions

In this chapter, we have examined KE building first as a narrative that justified the policies proposals of IO’s both of OECD and UNESCO, and then as a national endeavor that comprises symbolic and socio-material dimensions.

We proposed the hypothesis that there is no single, universal KE as a desirable development destination per se. Although much more empirical research is needed after examining case studies, it is plausible to assert that attempts at symbolic construction of each KE can be found in both paradigmatic and regular cases. We can even find them in the transition to KE of central advanced countries.

We found that exemplary types of development such as Finland and South Korea did not respond to the KE narrative of the OECD’s monodisciplinary approach. Instead we observed that the construction of a national institutional superstructure based on a sociotechnical imaginary requires a state leadership effort.

When public-private leadership deserts the crafting of a socio-technical imaginary of development, the vacant voids are filled by global narratives such as the techno-entrepreneur figure and the culture of innovation. Countless countries have tried to build their own Silicon Valley by turning their backs on their educational institutional contexts. They have failed by enmeshing social actors who could have been recruited into cohesive national enterprises.

In their eagerness to promote public development policies, international organizations have oversimplified complex and unstable processes making the symbolic construction of sociotechnical imaginaries invisible or at least irrelevant. When no symbolic work was pursued, and only mimicked operating models of other “successful” countries, institutions built in the previous industrial period ceased to be effective in their functions.

Although the KE narrative has managed to survive setbacks and transformations, it is doubtful that its power of persuasion can continue to inspire sustainable development processes without reflection on its blind spots. Cases such as “Theranos” and the trial of its founder Elizabeth Holmes or the bankruptcy of “WeWork” point out the crescent exhaustion of the production of meaning of the generic KE narratives.

One blind spot was the idea that entrepreneurs are in charge of producing new knowledge when they only take the last link in a chain of public goods investments. Another blind spot was the overestimation of knowledge as a marketable good. The circulation of knowledge is highly dependent on the tacit knowledge circulating in communities of practice around and within educational institutions.

If we have learned anything about KE definitions, it is that they are all transitory. Due to current sociotechnical infrastructure conditions, the Knowledge Society is a moving target, a black box that is constantly resetting itself. It is advisable to constantly exercise systematic reflexivity in the social sciences dealing with the development of KE. Perhaps the closest we came to a definition of KE was with an economy based on the manipulation of signs that requires human creative activity exercised in a given institutional space and cultural (national) context through platform technologies.

Twenty-six years after the arrival of the term “Knowledge Economy,” the public agenda today is not very different from early linear, globalist ideas based on reliance on entrepreneurial individualism, even though they have been modified by the evolution of the experience of sociotechnical change.

The humanist reaction to KE did not question the very concept in their attempt to bring KE into dialog with social and cultural diversity and address inequalities in access to up-to-date education and the tools of thought. UNESCO’s attempts to put KE in the context of human development while putting the digital divide and cultural diversities on the agenda met with stiff resistance. But it attempted to do so from a weak globalist and identity-defensive position that dismissed the importance of the symbolic national construction of the KE.

The case with the 1:1 (OLPC) model of educational digital inclusion was an example of the fragility of attempts at well-intentioned globalist interventions. Since the free or open knowledge movements flourished in the Global North civil society, weaknesses of the nation-state in developing countries have not been able to take advantage of their projects. Then the 1:1 proposals that were born open source became a battering ram of the more closed platform.

The sustained demand for digital talent in the new millennium, combined with the reliance on Edtech technology solutions, has pressured education systems with experiments such as MOOCs and Microdegrees. Without conclusive results, secondary and higher education moves post-pandemic toward hybrid or online scenarios. The Negroponte dichotomy between atoms and bits is diluted in post-digital hybrid fluids.

Among the difficulties in thinking about the interactive relationship between economy and society in the datification era, the most difficult challenge is to overcome a narrative of the single abstract model of KE. Educational reforms of the future need inspiration within the framework of development policies around a national-specific endeavor that comprise symbolic and sociomaterial dimensions, capable of justified new endeavors of digital educational infrastructures.

Few efforts have been made to build an enriched body of theory that integrates educational, institutional, and cultural change around the uniqueness of each nation-state. It is necessary to encourage a multi-scalar explanatory framework that aims to connect micro, meso, and macro levels of educational change concerning Sociotechnical Imaginaries.

There is no longer an automatic translation between the digitization of the world and the advancement of KE. Digital transformation can also lead to the permanence of collective ignorance. Advanced digital technologies are no longer necessarily synonymous with development based on the personal dignity of dignified life projects.

Uncited references

Castells et al., 2014; Castells and Himanen, 2014; Department for Business Innovation and Skills. UK, 2016; Van Dijk et al., 2018; Frantsi, 2008; Fussey and Roth, 2020; Heeks, 2002; Lauder et al., 2012; Lo and Tang, 2020; Lupton, 2014; Luzardo and Gasca, 2018; Mansell, 2001; Melo, 2001; Moore, 2019; Morin, 1999; National Research Council, 2002; OCDE, 2010; Pasquale, 2018; Pérez, 1996; Selwyn, 2019.

References

- Al-Fehaid, H.Y.N., Shaili, V., 2021. Knowledge economy and its implications in the Kingdom of Saudi Arabia. SSRN Electron. J. 1–38. <https://doi.org/10.2139/ssrn.3846918>.
- Ames, M., 2019. The charisma machine: The life, death, and legacy of One Laptop per Child. MIT Press.
- Artopoulos, A., 2020. Orígenes del subdesarrollo informacional De la industria del software al extractivismo de talento en Argentina (2002–2019). In: Rivoir, A. (Ed.), *Tecnologías digitales y transformaciones sociales Desigualdades y los desafíos en el contexto latinoamericano actual*. CLACSO—ObservaTIC.
- Artopoulos, A., Huarte, J., Rivoir, A., 2020. Plataformas de simulación y aprendizaje. *Propuesta Educativa* 1, 1–16.
- Bank, T.W., Thomas, A., Carl, D., 2001. Korea and the Knowledge-Based Economy Making the Transition: Making the Transition. OECD Publishing (Information society).
- Beck, U., 1998. La sociedad del riesgo: hacia una nueva modernidad. Paidós Ibérica.
- Beech, J., 2005a. International Agencies, Educational Discourse, and Teacher Education Reform in Argentina and Brazil (1985–2002). Available at. <http://eprints.ioe.ac.uk/6657/>.
- Beech, J., 2005b. Sociedad del conocimiento y política educativa en Latinoamérica: invirtiendo los términos de la relación. *Quaderns Digital* (38), 1–13.
- Beech, J., 2011. Global Panaceas, Local Realities: International Agencies and the Future of Education. Peter Lang (Comparative studies series).

- Bell, D., 1976. The coming of the post-industrial society. Taylor & Francis.
- Bhatt, I., MacKenzie, A., 2019. Just Google it! Digital literacy and the epistemology of ignorance. *Teach. High. Educ.* 24 (3), 302–317. <https://doi.org/10.1080/13562517.2018.1547276>.
- Biggart, N.W., Guillén, M., 1999. Developing difference: social organization and the rise of the auto industries of South Korea, Taiwan, Spain, and Argentina. *Am. Socio. Rev.* 64 (5), 722–747.
- Bowman, W., 2015. Imagining a modern Rwanda: sociotechnological imaginaries, information technology, and the postgenocide state. In: Jasanoff, S., Kim, S.-H. (Eds.), *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press, Chicago.
- Calderón, F., 2015. Navegar contra el viento ... O las perspectivas de América Latina en la era de la información 1. *Revista de Sociología* 30, 11–29.
- Calderon, G., Castells, M., 2014. Chapter 7. Development, democracy, and social change in Chile. In: *The Information Society and the Welfare State: The Finnish Model*. Oxford University Press, Oxford.
- Castells, M., 1996. La era de la información. Economía, Sociedad, y cultura. Alianza Editorial.
- Castells, M., 2003. Panorama de la Era de la Información en América Latina ¿es sostenible la globalización? In: Calderón, F. (Ed.), *¿Es sostenible la globalización en América latina?* FCE, Santiago de Chile.
- Castells, M., Himanen, P., 2002. *The Information Society and the Welfare State: The Finnish Model*. Oxford University Press, NY, USA.
- Castells, M., Himanen, P., 2014. Reconceptualizing Development in the Global Information Age. Oxford University Press, Oxford.
- Castells, M., Caraga, J., Cardoso, G., 2014. Aftermath: The Cultures of the Economic Crisis. OUP, Oxford. Available at: <https://books.google.com.ar/books?id=bXwmnwEACAAJ>.
- Chesbrough, H., 2007. Why Companies Should Have Open Business Models. *MIT Sloan Management Review*, 48208.
- Choong, K., Leung, P., 2021. A critical review of the precursors of the knowledge economy and their contemporary research: implications for the computerized new economy. *J. Knowl. Econ.* 13, 1573–1610.
- Dahlman, C.J., Routti, J., Ylä-Anttila, P., 2006. In: Dahlman, C.J., Routti, J., Ylä-Anttila, P. (Eds.), *Finland as a Knowledge Economy. Elements of Success and Lessons Learned*, Knowledge for Development Program. World Bank.
- Dale, R., 2005. Globalisation, knowledge economy and comparative education. *Comp. Educ.* 41 (2), 117–149. <https://doi.org/10.1080/03050060500150906>.
- Denning, P.J., 2017. Remaining trouble spots with computational thinking. *Commun. ACM* 60 (6), 33–39. <https://doi.org/10.1145/2998438>.
- Department for Business Innovation & Skills UK, 2016. Success as a Knowledge Economy: Teaching Excellence, Social Mobility and Student Choice. Gov.Uk. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/523546/bis-16-265-success-as-a-knowledge-economy-web.pdf.
- Drucker, P., 1959. *The Landmarks of Tomorrow*. Heineman, Canada.
- Finwick, T., 2010. Policies for the Knowledge Economy: Knowledge Discourses at Play. In: Malloch, Margaret, Cairns, Len, Evans, Karen, Bridget, N O'Connor (Eds.), *The SAGE Handbook of Workplace Learning*. SAGE.
- Florida, R., 2003. *Rise of the Creative Class*. Turtleback.
- Frantsi, H., 2008. What Makes Finnish Kids So Smart ?, pp. 1–5.
- Freeman, C., Dosi, G., Freeman, C., Nelson, R., Silverberg, G., Soete (comps.), G. y L., 1987. Changes in the national system of innovation. *Technical Change and Economic Theory*. Pinter.
- Fullan, M., 2002. El significado del cambio educativo: un cuarto de siglo de aprendizaje. *Revista de currículum y formación del profesorado* 6 (1998), 1–14. Available at: <http://www.ugr.es/~recfpro/rev61ART1.pdf>.
- Fussey, P., Roth, S., 2020. Digitizing sociology: continuity and change in the internet era. *Sociology*. <https://doi.org/10.1177/0038038520918562>.
- Gibbons, M., 1997. La Nueva producción del conocimiento: la dinámica de la ciencia y la investigación en las sociedades contemporáneas. Pomares-Corredor.
- Godin, B., 2004. The New Economy: what the concept owes to the OECD. *Res. Pol.* 33 (5), 679–690. <https://doi.org/10.1016/j.respol.2003.10.006>.
- Godin, B., 2006. The knowledge-based economy: conceptual framework or buzzword? *J. Technol. Tran.* 31 (1), 17–30. <https://doi.org/10.1007/s10961-005-5010-x>.
- Haas, P.M., 1992. Introduction: Epistemic Communities and International Policy Coordination. International Organization. <https://doi.org/10.1017/S0020818300001442>.
- Hargreaves, A., 2003. Enseñar más allá de la sociedad del conocimiento: del valor del dinero a los valores permanentes. *Enseñar en la sociedad del conocimiento* 49–88.
- Heeks, R., 2002. Information Systems and Developing Countries: Failure, Success, and Local Improvisations, pp. 101–112.
- Hesmondhalgh, D., 2002. *The Cultural Industries*. SAGE.
- Hewison, R., 2011. "Creative Britain": myth or monument? *Cult. Trends* 20 (3–4), 235–242. <https://doi.org/10.1080/09548963.2011.589703>.
- Himanen, P., 2014. Reconstructing the Finnish model 2.0: building a sustainable development model? In: Castells, M., Himanen, P. (Eds.), *Reconceptualizing Development in the Global Information Age*. Oxford University Press, Oxford.
- Himanen, P., Castells, M., 2004. In: Castells, M. (Ed.), *Institutional Models of the Network Society: Silicon Valley and Finland. The Network Society. A Cross-cultural Perspective*, p. 5. <https://doi.org/10.2307/778114>.
- Himanen, P., Committee for the Future in Parliament of Finland, 2004. Challenges of the Global Information Society. Helsinki.
- Hui, Y., 2020. Fragmentar el futuro: ensayos sobre tecnodiversidad. Caja Negra, Buenos Aires.
- Jasanoff, S., Kim, S.H., 2013. Sociotechnical imaginaries and national energy policies. *Sci. Cult.* 22 (2), 189–196.
- Jasanoff, S., Kim, S.-H., 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*.
- Jenkins, H., 2006. *Convergence Culture*. NYUP, New York, NY, USA.
- Kallo, J., 2021. The epistemic culture of the OECD and its agenda for higher education. *J. Educ. Pol.* 36 (6), 779–800. <https://doi.org/10.1080/02680939.2020.1745897>.
- Kenway, J., et al., 2006. *Haunting the Knowledge Economy*. Taylor & Francis (International Library of Sociology).
- Lash, S., 2002. *Critique of Information, Crítica de la información*. Amorrotu.
- Lash, S., Urry, J., 1998. Economías de signos y espacios: Sobre el capitalismo de la posorganización. Amorrotu (Biblioteca de sociología).
- Latour, B., 2018. *Down to Earth: Politics in the New Climatic Regime*. Wiley.
- Lauder, et al. (Eds.), 2012. *Educating for the Knowledge Economy? Critical Perspectives*. Taylor & Francis Group, Routledge.
- Lo, W.Y.W., Tang, H.H.H., 2020. Chasing Phantoms? Innovation policy, higher education and the pursuit of a knowledge economy in Hong Kong. *J. High. Educ. Pol. Manag.* 42 (2), 178–193. <https://doi.org/10.1080/1360080X.2019.1687268>.
- Lundvall, B.-Å., 2010. National Systems of Innovation: Toward a Theory of Innovation and Interactive Learning. Anthem Press (Anthem Other Canon Economics Series). Available at: <https://books.google.com.ar/books?id=iDXGwacw-4oC>.
- Lupton, D., 2014. *Digital Sociology*. Taylor & Francis.
- Luzardo, A., Gasca, L., 2018. *Launching an Orange Future*.
- Mangabeira Unger, R., et al., 2019. *Imagination Unleashed Democratizing the Knowledge Economy*. London. Available at: www.nesta.org.uk.
- Mansell, R., 2001. Digital opportunities and the missing link for developing countries. *Oxf. Rev. Econ. Pol.* 17 (No. 2), 17(2).
- Marginson, S., 2010. Higher education in the global knowledge economy. *Procedia Soc. Behav. Sci.* 2 (5), 6962–6980. <https://doi.org/10.1016/j.sbspro.2010.05.049>.
- Melo, A., 2001. The innovation systems of Latin America and the Caribbean. In: Research Department Working Paper Series, 460. Washington: IADB.
- Moore, R., 2019. Rethinking US education policy: paradigms of the knowledge economy. *J. Educ. Pol.* 34 (3), 446–447. <https://doi.org/10.1080/02680939.2018.1492223>.
- Morin, E., 1999. Los siete saberes necesarios para la educación del futuro. *Revista De Innovación E Investigación Educativa* fdg.
- National Research Council, 2002. *The Knowledge Economy and Postsecondary Education. The Knowledge Economy and Postsecondary Education*, Washington. <https://doi.org/10.17226/10239>.
- Nonaka, I., Takeuchi, H., 1995. *The knowledge-creating company: how Japanese companies create the dynamics of innovation*. Oxford university press.

- OCDE, 2010. 1:1 en Educación Prácticas actuales, evidencias del estudio comparativo internacional e implicaciones en políticas, pp. 1–23, 44.
- OECD, 1996. *The Knowledge-Based Economy*. Paris.
- OECD, 1997. *National Innovation Systems*. Paris.
- OECD, 2015. *Students, Computers and Learning: Making the Connection*. OECD Publishing, OECD Publishing, PISA. <https://doi.org/10.1787/9789264239555-en>.
- OECD, 2016. *Skills for a Digital World*.
- OECD, 2021. *Pushing the Frontiers With AI, Blockchain, and Robots*. Available at. https://www.oecd-ilibrary.org/education/oecd-digital-education-outlook-2021_589b283f-en.
- Olsen, M., Peters, M.A., 2005. Neoliberalism, higher education and the knowledge economy: from the free market to knowledge capitalism. *J. Educ. Pol.* 20 (3), 313–345. <https://doi.org/10.1080/02680930500108718>.
- Pasquale, F., 2018. *Digital Capitalism: How to Tame the Platform Juggernauts*. Available at. <http://library.fes.de/pdf-files/wiso/14444.pdf>.
- Pérez, C., 1996. La modernización industrial en América Latina y la herencia de la sustitución de importaciones. *Comercio exterior* 46 (5), 347–363.
- Peters, M.A., 2003. Education policy in the age of knowledge capitalism. *Pol. Futures Educ.* 1 (2), 361–380. <https://doi.org/10.2304/pfie.2003.1.2.12>.
- Peters, M., 2009. National education policy constructions of the "knowledge economy": towards a critique. *J. Educ. Enq.* 2 (1), 1–22.
- Powell, W.W., Snellman, K., 2004. The Knowledge Economy, *Annual Review of Sociology*, pp. 199–220. <https://doi.org/10.1146/annurev.soc.29.010202.100037>.
- Rivoir, A., 2012. Estrategias Nacionales para la Sociedad de la Información y el Conocimiento en América.
- Romer, P., 1994. The origins of endogenous growth. *J. Econ. Perspect.* 8 (1), 3–22.
- Sahlberg, P., 2011. *Finnish Lessons*. Teachers College Press, NY.
- Saxenian, A., 2014. The Silicon Valley model: economic dynamism, social exclusion. In: Castells, M., Himanen, P. (Eds.), *Reconceptualizing Development in the Global Information Age*. Oxford University Press, Oxford.
- Schlenker, G., 2007. From path dependency to path creation: Finland on its way to the knowledge-based economy. *Curr. Sociol.* 55 (1), 92–109. <https://doi.org/10.1177/0011392107070136>.
- Selwyn, N., 2019. *What is Digital Sociology*. Wiley (What is Sociology?).
- Sen, A., 2000. Desarrollo como Libertad. *Gaceta Ecológica* (55), 14–20.
- Smith, C., 1998. *Creative Britain*. Faber & Faber.
- Sörlin, S., Vessuri, H., 2007. Knowledge Society vs. Knowledge Economy: Knowledge, Power, and Politics. Palgrave Macmillan US (Issues in Higher Education).
- Smicek, N., 2017. *Platform Capitalism*. Wiley (Theory Redux).
- Stehr, N., 1994. *Knowledge Societies*. SAGE.
- Stiglitz, J.E., 1999. *Public Policy for a Knowledge Economy*.
- Toyama, K., 2015. *Geek Heresy: Rescuing Social Change From the Cult of Technology*. PublicAffairs, New York, NY, USA.
- UNESCO, 2005. *Towards Knowledge Societies*.
- Van Dijk, J., 2020. *The Digital Divide*. Wiley.
- Van Dijk, J., Poell, T., de Waal, M., 2018. *The Platform Society: Public Values in a Connective World*. Oxford University Press.
- Vincent-Lancrin, S.C., Román, C., Reimers, F. (Eds.), 2022. *How Learning Continued During the COVID—19 Pandemic*.
- Warschauer, M., 2003. Demystifying the digital divide. *Sci. Am.* 289 (2), 42–47. Available at. <http://www.jstor.org/stable/26060401>.
- Warschauer, M., et al., 2014. Balancing the one-to-one equation: equity and access in three laptop programs. *Equity Excell. Educ.* 47 (1), 46–62. <https://doi.org/10.1080/10665684.2014.866871>.
- Wilson, E.J., Segal, A., 2005. Trends in China's transition toward a knowledge economy. *Asian Surv.* 45 (6), 886–906. <https://doi.org/10.1525/as.2005.45.6.886>.
- Zuboff, S., 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Redefining educational transfer and borrowing in the pluriverse

Byoung-gyu Gong^a, Jieyu Jiang^b, and Iveta Silova^c, ^a Sorenson Impact Center, David Eccles School of Business, University of Utah, Salt Lake City, UT, United States; ^b Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, United States; and ^c Mary Lou Fulton Teachers College, Arizona State University, Chandler, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	290
Educational transfer in the western transcendental worldview	291
The logic of universality	292
“Best practice” as the ideal norm	292
Theory over context	293
Imagining educational transfer beyond western transcendence: toward pluriversality	294
No comparison option: interrupting education transfer	294
Post-colonial option: counter-balancing the inequity and injustice in comparative education research	295
Decolonial option: changing reference points and terms of knowledge production and transfer	295
Spiritual/ontological option: rebuilding the pluriverse through education comparison and transfer	296
Implications for future research	297
References	299

Introduction

The study of educational transfer has become increasingly prominent in the context of globalization, signaling a rapidly intensifying circulation of education policies and practices worldwide. Similarly to mainstream research on globalization, western scholars have generally approached the study of educational transfer as a rational, linear, and predictable process, seeking a *universal* logic behind the global phenomenon. While some research on educational transfer has focused on knowledge exchange across western or northern countries, most research has examined a unidirectional knowledge transfer from West to East, from North to South, or from global to local—but rarely in reverse or multiple directions—overlooking the problems of uneven and unequal knowledge exchange. Furthermore, many international agencies (e.g., Organization of Economic Cooperation and Development, the World Bank, United Nations, etc.) have used educational transfer as a central mechanism to maintain the geopolitical and economic status quo, assuming a universal trajectory toward “progress” through a trickle-down effects of mass schooling and infinite economic growth. Based on the assumption that “best” practices can be “lent” and “borrowed,” much research and practice of educational transfer has been historically approached from the perspective of “global linear thinking” as a deliberate convergence toward Westernization (Mignolo, 2021). With some exceptions, such research has actively tapped into and relied on the hierarchical global diffusion networks of educational knowledge sustained since the colonial period (for critique see Carney et al., 2012).

While there has been a growing critique of mainstream approaches to research and practice of educational transfer, such critique is too often based within the dominant western modern(ist) paradigm (Phillips and Ochs, 2003; Steiner-Khamsi, 2004; Phillips, 2004; Steiner-Khamsi and Waldow, 2012; Anderson-Levitt, 2003). In particular, it assumes that educational transfer can help—if properly implemented through the right mix of policies and practices or careful attention to contextual details—to make progress toward a “better” world for all. Even when research is not driven by ameliorative motives, it tends to use the “global” as a taken-for-granted starting point, focusing on the diffusion of dominant education paradigms and tracing the political, economic, or cultural dynamics associated with their adoption in local contexts. Fixated on using the “global” as the main reference point and drawing on universality as its epistemological framework, such research produces “theoretical effects of closure,” making it impossible to imagine what does not fit its hegemonic definition and what may thus lie “outside of the global” (Staheli, 2003, p. 2).

The challenge is how to redefine the notion of educational transfer in ways that would bring into focus—and into policy conversation—multiple education knowledge traditions, including those that have been systematically ignored, outcast, and sometimes even erased by the dominant frameworks that insist on *universality* as an epistemological foundation and starting point. Such redefinition entails moving beyond “changes in the content” by simply adding “missing” knowledge traditions into the framework of universality to instead changing “the terms of the conversation” itself by grounding it within the logic of *pluriversality* (Mignolo, 2021). Taking pluriverse (rather than universe) as a starting point means acknowledging that reality is constituted not only by many worlds but “by many kinds of worlds, many ontologies, many ways of being in the world, many ways of knowing reality” (Querejazu, 2016, p. 3). The pluriverse thus implies a nonhierarchical coexistence of and exchange between different knowledge traditions, including a multiplicity of Indigenous cosmologies and non-western knowledge traditions, as well as a wide array of western ontologies and epistemologies. From this perspective, the pluriverse acts as “a strategic response to the violence of universalism,” advocating for pluriversal ethics, cosmopolitics, and world-making practices (Mercier, 2019, p. 1). It constitutes “the re-existence of ways and modes of knowing and living” that disobey the hegemonic and homogenizing expectations of western and North Atlantic scholarship (Mignolo, 2021, p. 721).

In this chapter, we engage in the process of changing the terms of the conversation by writing from “the receiving end of the epistemic colonial difference”—i.e., ontological and epistemic differences produced as a consequence of colonial and imperial designs—in order to bring into focus “ways of knowing and knowledge that do not bend to western European and North Atlantic epistemic regulations” (Mignolo, 2021, p. 721). In particular, we aim to lay the foundation for redefining the concept of educational transfer in ways that would enable multidirectional and non-hierarchical knowledge exchange in a pluriverse. We will start by critically reviewing the colonial and imperial western-centric frameworks of the existing educational transfer research paradigms grounded in the western transcendental worldview, particularly focusing on the “theoretical effects of closure” that stem from the blind faith in universality, the belief in “best practice” as an ideal education norm, and prioritization of theory over context. We will then draw on non-western perspectives to articulate alternative theoretical frameworks beyond the western transcendental view, opening the space to approach educational transfer as a non-hierarchical interaction between and among multiple knowledge systems, perspectives, and ideas.

Educational transfer in the western transcendental worldview

Educational transfer has been one of the key instruments of the modernity project in the West, deeply rooted in the western notion of transcendence that has positioned western thought (the so-called scientific knowledge) not only as unique and univocal, but also as universal.

To avoid arbitrarily using the concept of transcendence in understanding other alternative frameworks, Hall and Ames (1998) explain that in western philosophy, transcendence indicates “an asymmetrical relationship” between the two categories, where A is independent from B and affects B, while A will not be affected by B in turn, reflecting the predominant discursive power and control in two-dimensional, linear interactions (p. 191).¹ That is to say, the transcendent category is independent from or outside of the ordered system, serving as a prominent model to be imitated by others, such as the God to the created world in the theological tradition (p. 191) or the western “expert” or “best practice” to the so-called “developing world.” Furthermore, such “strict transcendence” of the one over the other precludes any substantive interaction between the two, while reinforcing binary logic that divides mind and body, individual self and society, knowledge and practice, or theory and context.

According to Hall and Ames (1998), the appeal to “strict” transcendence in western societies seems to have been grounded in attempts “to meet the challenge of the pluralism of beliefs and practices by recourse to objective, unassailable norms” through transcendence toward universality (p. 189). In this context, universalist discourses of transcendence assume that the world is one, that it is knowable across different contexts within single modes of thought, and is thus manageable and governable in those terms (Conway and Singh, 2011, p. 701). This logic of transcendence is clearly reflected in the mainstream frameworks of educational transfer research and practice, as illustrated in Fig. 1 below. In particular, the majority of the existing theories exploring educational transfer are generally based on the transcendental assumption that universal knowledge exists at the “global” level and diffuses across the “local” and non-western contexts with less or more variation of its original form. Assuming a linear, unidirectional

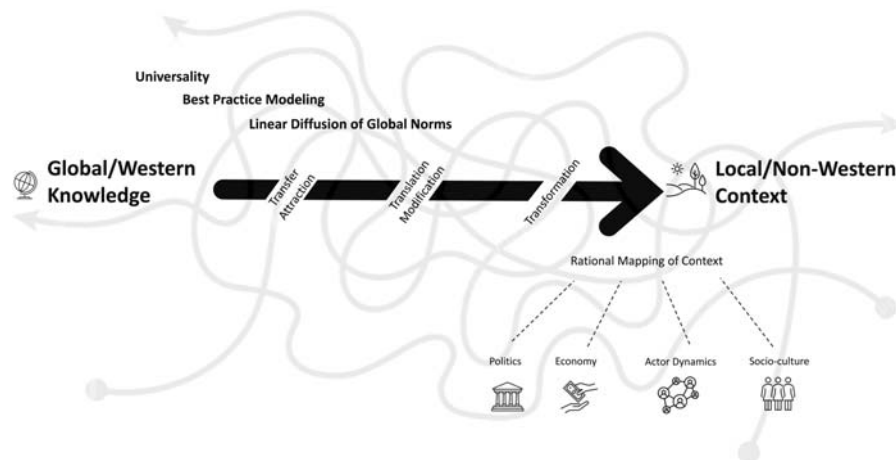


Fig. 1 A visual summary of the dominant approaches to the study of education transfer. *Note:* The Fig. 1 visually summarizes dominant approaches to the study of education transfer, drawing on Dolowitz (2000), Phillips and Ochs (2004), and Rogers (2003). The dark arrow represents the process of education transfer as envisioned by the (western) transcendental worldview. The light gray (thin) lines show the multiple and multidirectional knowledge flows in the pluriverse.

¹In *Thinking from the Han*, Hall and Ames (1998) define the Western notion of transcendence as “A is transcendent with respect to B if the existence, meaning, or import of B cannot be fully accounted for without recourse to A, but the reverse is not true” (p. 190).

flow of knowledge transfer, such approaches define specific stages or temporal phases of the transfer process, which may include (1) policy attraction, (2) interpretation or translation, (3) modification, adaptation, or reception, and (4) internalization or indigenization, among others (e.g., see Dolowitz, 2000; Rogers, 2003; Phillips and Ochs, 2004). Importantly, any existing deviations from the universal logic of educational transfer are usually explained by elaborating on the hidden politics, complex dynamics, multiple actors, and intricate processes. Not surprisingly, such conceptualization of education transfer has resulted in research that is grounded in the logic of universality, the belief in “best practice” as an ideal or universal education norm, as well as prioritization of theory over context, thus introducing—and reproducing—asymmetrical relationships in educational transfer research.

The logic of universality

The idea of universality has been deeply embedded in the foundations of the western modern(ist) thought. From western bureaucrats in colonial administrations to like-minded academics and international development practitioners, the belief in universality has set policymakers on the quest for efficient education policy solutions to “move” countries toward a future imagined in the image of the West. This policymaking logic has intensified after the end of the Cold War in 1989, reaching a point where education policy studies came to be almost entirely focused on identifying and offering “solutions” to any problems or obstacles hindering the promise of “Progress” (Silova et al., 2020). The most obvious examples today are found within OECD and World Bank led projects seeking to gather immense amounts of data and apply the most advanced statistical techniques in analysis to identify “what works” in raising students’ achievement and then translate these findings into “best practices” to be shared globally (Auld and Morris, 2016). International initiatives such as Education for All (EFA), the Millennium Development Goals (MGDs), and more recently the Sustainable Development Goals (SDGs) reveal a similar logic of universality.

In comparative education research, the logic of universality is reflected most clearly in the world culture theory research, which aims to depict the typical and prevalent (universal) images, ideas, and messages “irrespective of history, identity, or national context” (Carney et al., 2012, p. 372). In other words, instead of describing diverse kinds and forms of world cultures and theories, world culture theorists “have been increasingly involved in selectively identifying and advocating for the global diffusion of particular education models which reflect particular western and, especially, North American ideals” as a global common good (p. 368). While critics have challenged these assumptions by focusing on local manifestations of world-level tendencies (e.g., researching location adaptations of global reforms), their critique has been comfortably accommodated within world culture theory. In fact, some critique may have inadvertently contributed to the expansion of the logic and hierarchies of western universal (and universalizing) tendencies by taking for granted the global (heavenly) core as a starting point and adding more comparative cases of global diffusion across the world—offering thick descriptions of education policy and practice variations in different local contexts—without critically questioning its underlying logic.

Relying on the logic of universality inevitably incapacitates any attempts to account for contextual diversity, difference, and distinctiveness, which become instead subsumed into universal—and universalizing—frameworks. In the context of global economic competitiveness driven by international student achievement tests, standardized evaluations, and metrics, the West (along with the policies emanating from there) appears as a model of “progress” for others to follow. In this process, binary categories are used to set apart the West from the “rest”, the so-called “developed” countries from the “developing”, or the high-ranking countries or students from the low-ranking ones. The differences are often explained in terms of “lagging temporality,” a notion which not only redeploys the teleological construction of progress from East to West (or South to North), but also embeds the teleology itself, focusing attention once again on the future and the West (Hörschmann and Stenning, 2008, pp. 320–321). As a result, the diversity, depth and scale of the world’s particular histories, cultures, and geographies are erased as they become (just like) the West (Hörschmann and Stenning, 2008, p. 321).

More problematic, however, is that the logic of universality continues to be taken for granted as “real” and inevitable in much of education scholarship, inadvertently solidifying and promoting the *assumption* of global “sameness” by suppressing non-western realities as sites or phenomena to be discovered and explained, while failing to recognize them as self-sufficient in their own right. In this process, it blocks from our view “a much richer picture of *varied* motivations and levels of explanatory power for a *combination* of theoretical perspectives—functionalism, micro-realism, cultural/historical institutionalism, world systems theory, feminism, post-structuralism, post-colonialism and others” (Silova and Rapple, 2015, p. 2). This points out two problems with the logic of universality in comparative education research. First, the concept of universality assumes that the “correct” answers and “best” solutions do indeed exist, with countless studies being produced to illustrate this assumption empirically. Second, the research cases are often presented as effective manifestations of “best” practice, offering concrete examples for others to follow in order to become “successful” and “advanced” if those examples are followed step by step. In other words, universality appears to provide both the diagnosis of particular problems and ready-made answers to these problems by offering the methods and procedures of achieving the predetermined goals.

“Best practice” as the ideal norm

Historically, educational borrowing has been conceptualized as one of the key tools or mechanisms of pursuing “progress” and development. Whether speaking from the contexts of the 19th or 21st centuries, scholars and policymakers have frequently approached educational borrowing from a very pragmatic perspective, seeking to employ ideas taken from the experience of one educational system or context to another for the purposes of improving society or the individual. In the post-World War II context,

educational borrowing became perceived as a tool for achieving broader ends, directly “relating education to economic growth, social amelioration, and political development” (Noah and Eckstein, 1969, p. 116). What educational borrowing offered comparative and international education was an “instrumental value” for correcting the course of history:

No longer were the horizons for comparative education limited to simple cross-national borrowing ... by understanding forces and factors that molded education and society, men might be able to chart the course they were taking, and, if they did not like either direction or speed, conceivably they could hope to modify them.

Noah and Eckstein (1969, p. 41).

We see here the idea of “the ideal norm,” which first appeared in Jullien’s work in the early 19th century (Jullien, 1817; also see Beech and Rizvi, 2017) beginning to gain dominance. This policy-oriented approach to educational borrowing intensified during the period of the Cold War. Following the launch of Sputnik by the Soviet government in 1957, the study of educational borrowing re-positioned comparative education in the United States at the frontlines of the Cold War in order to “keep the United States ahead of the Soviet Union through education” (Noah, 2006, p. 10). With the funding available from the National Defense Education Act (NDEA) and Ford Foundation, comparative educators in the United States eagerly engaged in the study of “best practices” to ensure the educational competitiveness of the US globally, while at the same time pursuing other strategic interests of national importance—frequently expressed in the “concern for the plight of less fortunate people” (Noah and Eckstein, 1969, p. 38)—in non-aligned countries. Finally, international development opened worldwide opportunities for comparative educators to engage in technical assistance, marking the “development turn” in comparative education (Steiner-Khamsi, 2006) and thus further codifying the connection between normative educational borrowing and the ideal of “progress.”

Common to such policy-oriented approaches to the study of educational borrowing is the belief that education is future-directed with the purpose of improving the world, society, and the individual. Beech (2006) observed that this trajectory generally includes three stages, including (1) identification of a local problem, (2) identification of solutions in foreign educational systems, and (3) finally an adaptation of these “solutions” to the local contexts (p. 2). By infusing scholarship with western democratic ideals and liberal-humanist values about the nation-state and the individual—“liberty, equality, and fraternity”—it reflects “an abiding faith in the idea of progress through education” (Kazamias, 2009, pp. 41–42). The expectation is that “lessons drawn from comparative study could be used to aid in the development of the new nations” (Noah and Eckstein, 1969, p. 117). In more recent and increasingly complex iterations, a diverse range of international organizations, think tanks, and policy entrepreneurs have used cross-national achievement tests, such as the OECD’s Program for International Student Assessment (PISA), to construct and promote knowledge of transferrable “best practices,” while putting pressure on national governments to realign their education goals and curriculum according to “global” standards (Auld and Morris, 2016). In this context, “best practice” functions as one of the central tools of constructing and moving toward the ideal norm based on the model of western “progress.”

Theory over context

The binary and dualistic assumption of the transcendental worldview in western philosophy has caused a divide between theory and context. The transcendental worldview weighs more on “what” (in relation to theory), than “how,” i.e., the pragmatic concern to approach the context with the so-called “what-priority attitude” that places greater value on the truth of knowledge rather than the knowledge of how (Hall and Ames, 1998, p. 104). From this perspective, context is viewed as a barrier to actualize or prove theory, because it is creating variation, uncertainty, and mutation. The variant context is seen as threatening to destabilize knowledge of truth, while theory is assumed to be based on consistent, complete, and discoverable conditions that can produce an overarching truth claim. Such strict assumptions on truth tend to separate theory from its “sphere of praxis” (Hall and Ames, 1998, p. 128). As Takayama (2011) points out, such theory over context attitude is deeply rooted in the western transcendence worldview, creating a division of the comparative education research into western theory and non-western context. It is often assumed that theory is produced by western “experts,” while it is tested and applied by local education practitioners (for critique see Silova et al., 2017). In a similar manner, the context has been often approached as a puzzle or a black box full of complexity and uncertainty, while most comparative research has continued to focus on the easily identifiable and transferable “global” knowledge (Rapple, 2006). Even when recognizing the importance of context, many studies have focused on producing detailed descriptions, elaborations, and lists of actors, mechanisms, and processes in relation to the dominant trajectory of traveling “global” policies and practices in various contexts, as if nothing else exists outside of its scope.

Further reinforcing the dominance of theory over context is the increasing popularity of numbers in comparative research, as evidenced by the rise of international large scale assessments such as OECD’s PISA or IEA’s TIMSS studies as tools to facilitate comparisons and educational transfer across diverse contexts. As Piattoeva (2021) points out, “numbers work by stripping away the contexts of their production and the granular and ambiguous detail of the phenomena they claim to represent” (p. 511). In particular, both producers and users of numbers—whether large-scale student learning assessments or administrative statistical data—tend to articulate measuring or being measured in a decontextualized manner to claim legitimacy of numbers as objective

representations of reality or impartial tools of governance and decision-making (Piattoeva, 2021). In addition to “standardizing the context” by “translating diverse contexts into a single commensurate space,” numbers are also used to establish and maintain the standardized frameworks, benchmarks, and processes required for these large scale education comparisons (Gorur et al., 2019, p. 302). From this perspective, numbers effectively serve to legitimize the universality of theory (and western knowledge) over context, while simultaneously attempting to subsume multiple and diverse contexts into universalizing frameworks.

Departing from this dominant approach are various alternatives within both western and non-western scholarship. For example, comparative education scholars working in the area of Science and Technology Studies (STS) see context as “relationally emergent,” arguing that any standardization of context simultaneously entails the contextualization of global standards, i.e., “standards have to be adapted and made more flexible to accommodate new contexts” (Gorur et al., 2019, p. 304). Similarly, researchers may think about context with the sensibilities of Actor-Network Theory toward the relationality and flatness of actor-networks, suggesting that researchers themselves participate in context-making and thus act upon the world through research and policy practices (Piattoeva et al., 2018). Such perspectives contribute to destabilizing the notion of contexts as “neatly packaged matter-of-fact cubes,” gesturing toward imagining contexts as a temporarily realized “confluence of practices and objects coming together and never permanently stabilizing” (see Sobe and Kowalczyk, 2013). In non-western scholarship, the Chinese intellectual tradition refers to the so-called “art of contextualization,” i.e., an approach that places less emphasis on theory building or a single coherent notion of truth, instead drawing on the analogical reflection to suggest how the seemingly fixed notion of truth can permanently change across different contexts. In the Eastern philosophy more broadly, knowledge (Zhi) is not separable from context, i.e., “the realizer is not independent of the realized circumstances, but rather is a constituent element in the creative enterprise of making a world” (Hall and Ames, 1998, p. 30). By avoiding metaphysical questions such as what is truth, what is eternal, what is essence, the Chinese thinkers have been creating “site specific” and “provincial” knowledge that allows for a myriad of unique details and values to coexist.

Imagining educational transfer beyond western transcendence: toward pluriversality

Writing against the dominant understanding of educational transfer in the framework of universality, this chapter attempts to move beyond the limitations of the western-centric worldview and instead establish a set of theoretical frames acknowledging the multiple facets of the phenomenon as well as the multi-directional and always relational knowledge flows. Building on recent research inspired by the decolonial thought (see Silova et al., 2020a,b; Takayama et al., 2017a; Tlostanova and Mignolo, 2012, etc.), we will explore alternative ways of understanding educational knowledge production and circulation with a commitment to pluriversality (rather than universality) as guiding principles, while expanding an education space where various views and worldviews can coexist on a non-hierarchical basis, without undermining each other’s ontological and epistemological grounds. Once we make a commitment to moving beyond the idea of universality as a single vision of comparative research on educational transfer, what are our options? Of multiple alternatives, we will briefly highlight only some options that align with the principle of pluriversality, acknowledging that many other options coexist along these ones. Keeping in mind that these options are not fully divergent and may overlap in various ways, we will briefly discuss each option separately with a focus on educational policy transfer research, including (1) no comparison and/or transfer option, (2) the post-colonial option, (3) the decolonial option, and (4) the spiritual/ontological option.

No comparison option: interrupting education transfer

The “no comparison” option is the most radical approach against the business-as-usual in education comparison and transfer research, as it means a complete withdrawal from any comparative practice whatsoever. This withdrawal implies a refusal to contribute to and perpetuate the hierarchical knowledge production system of international and comparative education research and practice. The open letter debate among Andreas Schleicher, a director of the OECD, and a group of critical academics surrounding the devastating effects of PISA is one such example (Andrews et al., 2014). Signed by nearly 100 academics from various academic institutions worldwide, the letter expresses deep concern that “measuring a great diversity of educational traditions and cultures using a single, narrow, biased yardstick could, in the end, do irreparable harm to our schools and our students” (Andrews et al., 2014). While the letter calls OECD to “slow down the testing juggernaut” in order to make room for the full range of relevant constituents and to incorporate collective learning (rather than a top-down and unidirectional approach), some governments have either withdrawn or refused to participate in the PISA or other international large-scale assessment tests all together.² For these governments, a refusal to participate in international large-scale assessment tests may have signified an attempt to recover a sense of existence without the dominance of one education vision over another or the “arrogance of centrism” in a world of global economic competition (Radhakrishnan, 2009, p. 461).

A withdrawal from the comparison can be a first step to bring into focus and re-assert previously subjugated, dismissed, and even erased voices in the educational policy-making processes that tend to prioritize the pursuit of global competitiveness or modernization of education at the expense of other education goals. It is related to abandoning top-down policy processes that impose

²For example, India participated in the 2009 cycle of PISA, with low results, and withdrew from PISA 2012, only to re-join again in 2019. Kyrgyz Republic also opted to withdraw from PISA 2012 after its low performance in the 2006 and 2009 cycles, but recently announced to rejoin the study in 2024. For a discussion of different motivations behind India’s decision to withdraw and then re-enter PISA, please see Gorur et al. (forthcoming).

foreign education models through externalization and transfer. Many PISA-related education reforms have been introduced without considering national priorities, paralyzing local level policy-decision making processes. Furthermore, a refusal to participate in PISA may signal an attempt to reinvigorate various national efforts of education innovation against the preconceived notions that local (or non-western) educators are incapable of producing credible knowledge or articulate meaningful education priorities. The “no comparison” approach seems radical, but it makes sense considering that the western countries rarely learn from the non-western countries to reflect on and reform their own education policies and practices. Reversely, what is problematic is that innovation in the west is seen as a self-invention claiming originality, while innovation in the non-west is perceived as a reinvention of the borrowed features from elsewhere (i.e., usually from the west). Hence, sometimes the withdrawal from education comparison and transfer can cultivate diversity and reinvigorate comparative education research in the long term by rebalancing global power structures and dynamics, while returning centrality to diverse cultures in the non-west.

Post-colonial option: counter-balancing the inequity and injustice in comparative education research

Another alternative is the so-called post-colonial option that may extend more broadly in the area of critical education studies. This approach defines educational transfer itself as problematic because of the unidirectionality of transfer and diffusion. Most studies in comparative education examine the educational transfer from North to South, from West to non-West, or from South to South at best. The educational transfer in other directions—from South to North or from non-West to West—is rarely an object of research or source of knowledge. It is not easy to find such transfer cases in the real world. From the perspective of world culture theory and best practice modeling, the educational transfer is for a survival of the fittest, supposedly a value-neutral selection process to sort out the best in quality that is scientifically proven. In contrast, post-colonial perspective considers the transfer as a value-laden political process.

The post-colonial option entails a systematic effort to provincialize the West/North as the core of global knowledge production and counterbalance the inequity inherent in comparative education studies. For example, a growing number of scholars have challenged the “new” policy regime—driven by “numbers” and advanced by OECD and other international development agencies—for imposing western-centric lens to understand education across different contexts and converting complexity into (western) “best practices” (Auld and Morris, 2016; Beech, 2011; Dale and Robertson, 2012; Grek, 2012; Takayama, 2011, 2015; You, 2020a), for introducing market principles into public education settings (Verger and Moschetti, 2017; Verger, 2012; Robertson et al., 2012; Brehm, 2021), for deepening socioeconomic and gender inequities (Stromquist, 2016a,b; Unterhalter, 2017), or for overlooking other policy alternatives (Edwards and Loucel, 2016; Silova, 2010; Silova et al., 2020). For example, You (2020a) problematizes the process of educational transfer in the context of the recent “rise” of East Asia in the international PISA rankings, arguing that England—and more broadly Anglo-American societies—has represented high-performing East Asian societies as both an inspiration for education reforms and a threat to the domestic economy. Taking the OECD and McKinsey as two illustrative examples, she convincingly concludes that “the dominant ways of perceiving, representing and referencing East Asian education and the embedded East–West power relation are largely framed in a manner that continues the legacy of Orientalism” (You, 2020a, p. 742). Many others join forces in order to draw urgent attention to the problematic geopolitics of knowledge production that enables persisting racism and ethnocentrism in comparative education research, including studies on educational transfer (for critique, see a special issue of *Comparative Education Review* edited by Takayama et al., 2017a; Tikly, 1999, 2004).

In addition to provincializing the dominant (western/northern) paradigms and problematizing the knowledge hierarchy in comparative education research, the post-colonial option opens an opportunity to bring into focus and dialog knowledge(s) that have been previously disregarded or dismissed, including knowledge(s) seeking alternatives to a single vision of neoliberal education futures. In the comparative study of education policy transfer, some research has focused on the benefits of South–South cooperation and grassroots mobilization as possible ways to build more symmetrical relationships between the lenders and borrowers of education policies and practices (see Chisholm and Steiner-Khamsi, 2008; Jules and Sá e Silva, 2008; Sá e Silva, 2009). In particular, scholars have approached the study of the South–South transfer from multiple perspectives, ranging from regional cooperation initiatives e.g., Hickling-Hudson’s (2004) research on Cuba’s international educational assistance to schools in Jamaica) to community-driven forms of political mobilization e.g., Luschei’s (2004) study on Escuela Nueva in Colombia, Tarlau (2012) and McCowan’s (2003) work on the Landless Workers’ Movement in Brazil) to transnational social movements (e.g., Mundy and Murphy (2001) on Education International and other organizations). Ultimately, these studies gesture toward the multiplicity and multidirectionality of educational transfer phenomena by offering numerous cases of alternative educational transfer practices, while simultaneously deconstructing western dominance and counter-balancing systemic inequity in the global knowledge production and diffusion system.

Decolonial option: changing reference points and terms of knowledge production and transfer

The decolonial option represents a distinct break with the abstract (western) idea of universality, refusing to accept a single vision of what constitutes a “good” life or “best” practice. One example of an epistemological delinking from universality is Kuan-Hsing Chen’s *Asia as Method* (2010), which offers a decolonial, de-imperial, and de-Cold War analytical framework that moves research beyond western-centric interpretations of history and enables scholars to imagine historical experiences in Asia as “an alternative horizon, perspective, and method for posing a different set of questions about world history” (p. xv). Chen (2010) has analytically approached “Asia” as both a geographic region and a constructed cultural-political space with complex and contested—yet deeply

interconnected—historical relations within the region and with the “West.” By reorienting the conventional reference points away from the “West” and instead focusing on knowledge “inter-referencing” within the Asian region, [Chen \(2010\)](#) has re-centered Asia “as the source of a multiplicity of new [knowledge] flows” (p. 8), thus effectively interrupting the hegemony of western knowledge and offering a new view on global history, as well as global knowledge production. By de-centering the west as a single reference point for all comparisons and instead focusing on “inter-referencing” within Asia, the “Asia as method” approach opens ways to bring to the foreground multiple histories, while revealing the relationality and interdependence of different global spaces.

In comparative and international education, [Zhang et al. \(2015\)](#) elaborated on [Chen’s \(2010\)](#) ideas in an edited book *Asia as Method in Education Studies*, offering examples of what it means to rethink and reexamine education in Asia beyond both the western imperialist eye and the post-colonial “politics of resentment.” In response to [Chen’s \(2010\)](#) challenge, the authors engage in constructing new research imaginations for a more meaningful East-West dialog and knowledge co-creation. Similar attempts to redefine education knowledge production and transfer have also been proposed by other scholars, including [Takayama \(2016\)](#) and [Yonezawa et al. \(2018\)](#) reflecting on Japanese education research communities, [Lee \(2019\)](#) on South Korean education, [Kuswandono et al. \(2015\)](#) on Indonesian education, [Park \(2016\)](#) on education in Hong Kong and the broader Asian region, and others. Furthermore, multiple articulations of the “Asia as method” approach in education suggest that this decolonial project goes beyond a question of geographic focus and also entails a change of analytical approach in education research (see, for example, [Burman’s \(2019\)](#) discussion of “child as method” as a resource for interrogating models of development in childhood and education).

In addition to research inspired by “Asia as method,” important decolonial options have been developed by scholars working with Southern Theory ([Connell, 2007](#); [Takayama et al., 2017a,b](#)), epistemologies of the South ([Santos, 2007, 2018](#); [Esteva and Babones, 2013](#)), postsocialist/decolonial studies ([Silova et al., 2017, 2018](#); [Tlostanova and Mignolo, 2012](#)), indigenous decolonial practice ([Tuck and Yang, 2012](#); [Smith et al., 2018](#)), pluriversal rights education ([Williams and Bermeo, 2020](#)), among many others. While diverse in their approaches and foci, these various projects have some things in common. First and foremost, they neither claim universality nor attempt to replace one epistemology with another or others. Rather, these decolonial projects strive to contribute to creating a space where many different worlds and worldviews could coexist on a non-hierarchical basis ([Tlostanova, 2012](#); [Tlostanova et al., 2016](#); [Silova et al., 2020a,b](#)). Furthermore, these decolonial projects demonstrate that decoloniality of knowledge production requires scholars to also engage in decolonizing (and rebuilding) their subjectivity, which entails reclaiming their positions as epistemic subjects who have both the legitimacy and the capacity to interpret the world from their own origins and lived realities, while simultaneously looking outwards “to alternative and multiple forms of identification through the practice of ‘becoming others’” ([Zhang et al., 2015](#), p. 26). Collectively, they contribute to the new geopolitics of knowledge and being, unsettling modernity’s logic of universality and embracing a global viewpoint that reflects pluriversality.

Spiritual/ontological option: rebuilding the pluriverse through education comparison and transfer

The spiritual/ontological option signals a move beyond materialism, gesturing toward the importance of ontological concern in comparative education research. This is an elaboration of [Mignolo’s \(2011\)](#) brief description of the “spiritual option,” which challenges two assumptions of western science and capitalism: first, that a material world of ontologically discrete objects can be known “objectively,” and second, that it is a derivative of a particular religious worldview (Christianity). As [Silova et al. \(2020\)](#) have elaborated previously, the spiritual/ontological option opens up “new ontological possibilities, or if some prefer, metaphysical universes, ones that refuse a secular, materialist worldview as a starting and end point of research” (p. 18). It partially overlaps with the decolonial option discussed above and recent engagement with the post-secular turn ([Wu and Wenning, 2016](#)), the post-foundational, post-humanist, and new materialism turns ([Zhao, 2019a, 2020](#); [Murriss, 2021](#); [Bozalek, 2021](#); [Corson and Ress, 2021](#); [Petersen, 2018](#); [Epstein, 2019](#); [Carney and Madsen, 2021](#)), the ontological turn ([Jensen, 2017](#); [Holbraad and Pederson, 2017](#)), and bears close connections to work re-centering spirituality in the academy ([Shahjahan, 2004, 2005](#); [Edwards, 2016, 2020](#)).

To further illustrate how the spiritual/ontological option is being articulated in comparative education research, we offer three examples that challenge some of the foundational assumptions of dominant (western) approaches to education transfer studies and offer new alternatives that are built on non-hierarchical and non-binary relationality and deep interdependence. Comparing western understandings of the knowledge of truth to the Confucian one, [Hall and Ames \(1995, 1998\)](#) explain that in Confucianism, how individuals are defined, evaluated, and described depends on their relational networks (not confined to human beings relations), such as the relationships with individual self, parents, elders, friends, fellows, and objects. Hence, it is difficult to strictly divide the individual from the surroundings and tell relative superiority or inferiority of one over another through comparison. Similarly, the knowledge of truth is not static, abstract, or separate; rather, it is constructed as a contextually dynamic, concrete, and participatory ([Hall and Ames, 1995](#), p. 104). Challenging the transcendental status of “the only one truth and knowledge,” such an approach redefines the relationship between knowledge and human beings, as well as knowledge and reality, i.e., knowledge and truth are not single, sacred, and inviolable above individuals, and knowledge is not simply mirroring the reality. Instead, they are conditional, questionable, and variable because they grow out of daily interactions which are relational, dynamic, complex, and diverse. Drawing on this perspective, we can begin redefining modern and linear approaches to comparison and the education transfer. For example, [Tan \(2019\)](#) uses a Confucian conception of competence to critically reexamine a response to the competencies agenda that underpins international large-scale assessments such as PISA. Challenging the global competencies agenda based on the principles of technical rationality, discrete skills, and instrumental worth, she draws upon a relational model of competence that views competence as essentially communal, situated within social practices, and manifested through tacit achievement (e.g.,

a Confucian notion of competence premised on the virtue of *ren* or humanity as well as the social, cultural, and ethical dimensions of competence).

Similarly, the notion of *tianrenheyi* in Daoism suggests a trajectory against and beyond the notion of (western) transcendence in terms of reexamining the binary of fixed relationships of human/*tian*, human/environment, and differentiations/hierarchies (Zhao, 2019a). In particular, Daoism assumes that humans are a part of nature and the universe, as Zhuang Tzu famously said: “There is a man, there is Tian” and “all things and me are one.” To explore this traditional Chinese concept further in academic research, Zhao (2019a) illustrates the concept of *tianrenheyi* by explaining a Confucian person-making education and a “correlative cosmology” (p. 6) in terms of conceptualizing human as being relational, thus convincingly challenges the anthropocentric logic. By introducing ancient Chinese cosmology of *qi*, Zhao argues that the correlative cosmology and *tianrenheyi* transforms individual persons into relational roles within their situations, and posits individuals in their communities temporally as well as spatially necessary to achieve the ecological consonant state of co-creating and co-existing among beings, natural and social worlds (p. 6). This correlative cosmology echoes what Hall and Ames (1995) refer to as “Chinese acosmotic thinkers” (p. 183), meaning that “they do not depend on the majority of their speculations upon either the notion that the totality of things has a radical beginning, or that these things constitute a single-ordered world” (pp. 183–184). *Tianrenheyi* and the correlative cosmology thus extend the notion of comparison to a new level: recognizing *tian* and other surroundings beyond human beings and learning from and with them. This correlative and acosmotic education bridges the divergences between human and non-human beings, thus opening up to multiple cosmos. It further challenges the prevailing ideas of “best” practices and universality, thus broadening the approaches and possibilities for comparison and educational transfer beyond the one world/one worldview framework (see also You (2020b) for a discussion of learning/pedagogy from a Confucian and Daoist perspective).

Finally, alternative articulations of comparison and knowledge transfer have also been proposed within western scholarship aiming to challenge the principles of universality. For example, Donna Haraway’s (2016) idea of sympoiesis as a practice of “making-with” resonates deeply with some of the non-western thought, referring to the complexity and diversity of many co-existing worlds both from the ontological and epistemological perspective. In addition to acknowledging the presence of all other beings, there is a recognition of deep interconnectedness between all beings in the process of becoming-with each other and (re) making worlds together. Haraway (2016) explains that sympoiesis is a simple word that means “making-with:” “Nothing makes itself; nothing is really autopoietic or self-organizing ... earthlings are never alone Sympoiesis is a word proper to complex, dynamic, responsive, situated, historical systems. It is a word for worlding-with ...” (p. 58). Introducing the principles of sympoiesis in comparative and international education means resituating and reconnecting human (and more-than-human) beings within this inclusive and relational space which accepts those alternatives that have been marginalized in the hierarchical and modernized comparative education system. However, sympoiesis implies neither a neglect of difference (which may result in reverting back to the principle of universality in policy transfer) nor elevation of particularity to a transcendental space and knowledge (which may render comparison intangible and non-transferrable in a global context). As Stengers (2012) explains, the idea is not to transcend the particularity but “to think with this particularity, to induce the capacity to imagine the possibility that it can be regenerated” without being universalized:

It means thinking with its own specific, dangerous and never innocent ways of weaving relations. It means thinking with the resources—imaginative, scientific and political—that it may be able to activate in order to enable us, perhaps, to think with other peoples and natures. (p. 156)

In the field of education, the Common Worlds Research Collective (2020) has been translating many of these ideas into “common world” pedagogies and research, calling for a complete paradigm shift—“from learning about the world in order to act upon it, to learning to become *with* the world around us” (p. 2, emphasis added). In comparative and international education, collective efforts of reimagining knowledge production and transfer in more relational, more-than-human ways include research that draws on African cosmology of Ubuntu (e.g., Assié-Lumumba, 2017; LeGrange, 2018; Cossa, 2020), contemplative research and practice (e.g., Lin et al., 2019), negative education or unlearning (e.g., Takayama, 2020; Nishihira and Rapple, 2021; Zhao, 2019b), and more. What is common to these different approaches is that they engage with questions of knowledge production and transfer in relational ways, while refusing to see individuals, beings, objects, and institutions in isolation. From this perspective, comparison and educational transfer cannot be unidirectional or hierarchical; rather, they enable us to see and engage with similarities and differences in all layers of complexity. Furthermore, relationality in the knowledge of truth accepts neither the assumption of universality nor the belief in “best” practices, instead taking difference and diversity as a starting point and encouraging the mobility of truth and multiplicity in ways of knowing. Because the knowledge of truth is concrete, participatory, and about “know-how” in each unique context, the act of comparison and policy transfer cannot be generalizable and copyable regardless of different contexts, thus making the assumption of universality irrelevant and the mechanism of learning from (western) “best” practice obsolete.

Implications for future research

Given the centrality of educational transfer research in transforming education policy and practice worldwide, it is important to critically reexamine the underlying assumptions of the dominant research paradigm, as well as its implications for education

futures. In this chapter, we have highlighted how the existing research on educational policy transfer is generally rooted in the logic of modernity/coloniality, reflecting the established hierarchies in the global education knowledge production system and thus perpetuating colonial relationships. Despite a growing criticism of the logic of coloniality in comparative education research, a review of existing literature indicates that policy transfer research continues to reflect the western transcendental worldview, including its blind faith in universality, the belief in “best practice” as an ideal education norm, and prioritization of theory over context. While producing “unalienable gains” (e.g., human rights, liberty, democracy) and bringing attention to the global diffusion of dominant education norms (always based on a unidirectional knowledge flow from North to South or from West to East), such an approach has created theoretical, methodological, and existential effects of closure. In addition to standardizing comparative education theory and practice through scientific modeling and “best practice” approaches, it has contributed to displacing, marginalizing, and sometimes completely erasing education knowledge(s) and experiences that cannot be readily subsumed into the dominant globalization frameworks. It has also contributed to reproducing hierarchy and rank across distinctive educational systems and cultures, resulting in “unacceptable losses” and leading to the oppression, domination, marginalization, and even elimination of the Other and their ways of knowing and being (Zhao, 2009, p. 391; see also Wagner, 2002).

Moving beyond the western transcendental view, we aim to open the space for articulating research on education transfer as a non-hierarchical interaction between and among multiple knowledge systems, perspectives, and ideas. This requires changing the terms of the conversation rather than its content only, thus redefining the concept of educational transfer in ways that enable multidirectional and non-hierarchical knowledge production and exchange. This process is about taking the pluriverse (rather than universe) as an ontological starting point, thus acknowledging a nonhierarchical coexistence of different worlds and worldviews, where everyone and everything—both human and more-than-human—are interconnected. Pluriverse makes possible—and desirable—the co-existence of more than two options, of spaces in-between, and of more than one ontological reality (Silova et al., 2020). It enables us to move beyond the critiques of colonialism and imperialism in order to fundamentally rethink the adequacy of our vocabularies, theories, methods, practices, movements, and ways of knowing and being in order to articulate alternatives.

Then, what are the specific ways to set pluriverse as a foundation for future education transfer research? First and foremost, we can redefine educational transfer beyond the western transcendental worldview by bringing into focus different concepts of relationality, i.e., those that do not presume superiority of (western) modern worldmaking over alternatives in the pluriverse. For example, the notions of relationality in Confucianism, Daoism, or Ubuntu, along with many other examples from Indigenous and ecofeminist thought, all gesture toward breaking the consistent hierarchy of knowledge and being in comparative and international education by offering alternatives that are built on the principles of deep interconnectedness of all forms of life—human and more-than-human. These forms of relationality offer an alternative to modern epistemology and its many dichotomies, including those that divide science/nature, self/other, human/non-human, and more. Instead of relying on dichotomies, the pluriversality asks us to engage with difference and diversity by weaving mutual relationships. These non-western philosophies can be our reference point to building up a new intellectual foundation for educational transfer studies in the future.

Second, research approaches highlighting context (over theory) can contribute to a paradigm shift in educational transfer studies, opening up pathways toward the pluriverse. Unlike the transcendental worldview that aims to isolate “truth” from its surrounding contexts, many other thought traditions—from Confucianism and Daoism to ecofeminism and Indigenous knowledges—approach context as central to understanding, interpreting, and engaging with the world. For example, Haraway’s (1988) concept of “situated knowledges” reminds that knowledge production is always particular, partial, and historically contingent, that is, all knowledge claims are inevitably “views from *somewhere*” as opposed “universal truths” generated by disembodied and disembedded scientists who claim to observe “everything from nowhere”—a so-called “god-trick” (p. 590). In Confucianism, it is widely understood that pure “truth” without context is impossible, which explains the emphasis on the situatedness of all beings and knowledge (Hall and Ames, 1998). In particular, Hall and Ames (1998) call attention to “the art of contextualization,” which enables individuals to ally themselves with different contexts that they constitute and that in turn constitute them (p. 40). In a similar vein, educational transfer studies need to pay more attention to the context. We can regain a sense of context not by considering it as a final missing piece to complete a preconceived theory, but by bringing it to the fore from the very conception of research and practice. It may require more time and patience as we engage with the unequivocal concepts and unclear notions.

Third, embracing a pluriversal worldview fundamentally changes the notion of comparison itself. As Stengers (2011) explains, comparison is only meaningful if all those engaged in a comparison have the opportunity to present their own version of what the comparison is about, while avoiding the imposition of irrelevant criteria and categories on others. By redefining the principle of comparison in a non-hierarchical way, we can ensure that education transfer research does not contribute to differentiating, ranking, or homogenizing, but instead serves as a “connective tissue” that brings different worlds and worldviews into copresence with each other (Silova, 2020). It thus offers a powerful alternative to the currently dominant processes of globalization and international development as a single vision of our global futures.

Finally, we cannot underestimate the role of language itself in perpetuating hierarchical knowledge structure. Daoism considers that language both fragments and fixes the reality which is fluid (Hall and Ames, 1998). The adverse effect of language intensifies even further when ideas in one language are translated into another. Translation has been always an issue for comparativists because untranslated notions have been considered non-existing—thus, untranslatability often means uncertainty, invisibility, and untruth. The translation defines the boundary of modernity, setting borders between visibility and erasure (Vázquez, 2011). Translation is often seen as an essential step to minimize difference, while maximizing sameness and enabling comparison, classification, and exclusion. It can thus serve as an “imposition of an economy of truth” (Vázquez, 2011). Therefore, establishing new translation and communication practices and norms in academic work seems critical in order to make our world more pluriversal. We need

more research that can loosen the existing standards and norms so that untranslatable things (and phenomena) can find space in our work. We need to learn more about how to “stay with the trouble” (Haraway, 2016)—the uncertainty, untranslatability, spontaneity, and imprecise concepts. In such a context, comparative research is no longer about ranking, bordering, and classifying, but rather about “mirroring self and world, [recognizing] both the uniqueness of each perspective and parity among them” (Hall and Ames, 1998, p. 80). Such an approach to comparative research has the potential to redefine unidirectional, linear concepts of education transfer rooted in (western) universality, while opening up spaces for rebuilding the pluriverse.

References

- Anderson-Levitt, K. (Ed.), 2003. *Local Meanings, Global Schooling: Anthropology and World Culture Theory*. Springer.
- Andrews, P., Atkinson, L., Ball, S.J., Barber, M., Beckett, L., Berardi, J., et al., 2014. OECD and PISA tests are damaging education worldwide—academics. *Guardian*. Retrieved from: <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>.
- Assié-Lumumba, N.D.T., 2017. The Ubuntu paradigm and comparative and international education: epistemological challenges and opportunities in our field. *Comp. Educ. Rev.* 61 (1).
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education “best practice”. *Comp. Educ.* 52 (2), 202–229.
- Beech, J., Rizvi, F., 2017. Revisiting Jullien in an era of globalisation. *Compare* 47 (3), 374–387.
- Beech, J., 2006. The theme of educational transfer in comparative education: a view over time. *Res. Comp. Int. Educ.* 1 (1), 2–13.
- Beech, J., 2011. *Global Panaceas, Local Realities*. International Agencies and the Future of Education. Lang, Frankfurt.
- Bozalek, V.G., 2021. Doing Academia Differently: Creative Reading/Writing-With Posthuman Philosophers. *Qualitative Inquiry*, 10778004211064939.
- Brehm, W., 2021. *Cambodia for Sale: Everyday Privatization in Education and Beyond*. Routledge.
- Burman, E., 2019. Child as method: implications for decolonising educational research. *Int. Stud. Sociol. Educ.* 28 (1), 4–26.
- Carney, S., Madsen, U.A., 2021. *Education in Radical Uncertainty: Transgression in Theory and Method*. Bloomsbury Publishing.
- Carney, S., Rappleye, J., Silova, I., 2012. Between faith and science: world culture theory and comparative education. *Comp. Educ. Rev.* 56 (3), 366–393.
- Chen, K.H., 2010. *Asia as Method*. Duke University Press.
- Chisholm, L., Steiner-Khamsi, G., 2008. *South-South Transfer: Cooperation and Unequal Development in Education*. Teachers College Press, New York.
- Common Worlds Research Collective, 2020. *Learning to Become With the World: Education for Future Survival*.
- Connell, R., 2007. *Southern Theory: Social Science and the Global Dynamics of Knowledge*. Polity.
- Conway, J., Singh, J., 2011. Radical democracy in global perspective: notes from the pluriverse. *Third World Q.* 32 (4), 689–706. Retrieved June 13, 2021, from: <http://www.jstor.org/stable/41300342>.
- Corson, J., Rens, S., 2021. Post-foundational approaches in comparative and international education: uncertain moves toward unknown horizons. In: Jules, T.D., Shields, R., Thomas, M.A. (Eds.), *The Bloomsbury Handbook of Theory in Comparative and International Education*. Bloomsbury Publishing.
- Cossa, J., 2020. Cosmo-Ubuntu: toward a new theorizing for justice in education and beyond. In: *Critical Theorizations of Education*. Brill Sense, pp. 31–43.
- Dale, R., Robertson, S.L., 2012. Towards a critical grammar of education policy movements. In: *World Yearbook of Education 2012*. Routledge, pp. 41–60.
- Dolowitz, D.P., 2000. *Policy Transfer and British Social Policy*. Open University Press, Buckingham and Philadelphia.
- Edwards Jr., D.B., Loucel, C., 2016. The EDUCO Program, impact evaluations, and the political economy of global education reform. *Educ. Pol. Anal. Arch.* 24 (49), 1–50.
- Edwards, S., 2016. Seeking collective wisdom: a spiritual-dialogic approach. In: Lin, J., Oxford, R.L., Culham, T. (Eds.), *Towards a Spiritual Research Paradigm: Exploring New Ways of Knowing, Researching, and Being*. Information Age Publishing, Charlotte, pp. 257–273.
- Edwards, S., 2020. Allowing ourselves to reimagine ecologically responsible futures for education research and practice globally: critiquing the limitations imposed by Christian hegemony. *East China Normal Rev. Educ.* 3 (1), 164–168.
- Epstein, I., 2019. *Affect Theory and Comparative Education Discourse: Essays on Fear and Loathing in Response to Global Educational Policy and Practice*. Bloomsbury Academic.
- Esteva, G., Babones, S.J., 2013. *The Future of Development: A Radical Manifesto*. Policy Press.
- Gorur, R., Maddox, B., Sørensen, E., 2019. Standardizing the context and contextualizing the standard: translating PISA into PISA-D. In: Prutsch, M.J. (Ed.), *Science, Numbers and Politics*. Palgrave.
- Gorur, R., Dey, J., Arnold, B., (forthcoming). *From Adapting to Context to Engineering the Context: A Case Study of Pisa in India*. Paris: Cnesco-Cnam.
- Grek, S., 2012. Learning from meetings and comparison: a critical examination of the policy tools of transnationals. In: *World Yearbook of Education 2012*. Routledge, pp. 61–81.
- Hall, D.L., Ames, R.T., 1995. *Anticipating China: Thinking Through the Narratives of Chinese and Western Culture*. SUNY Press.
- Hall, D.L., Ames, R.T., 1998. *Thinking From the Han: Self, Truth, and Transcendence in Chinese and Western Culture*. State University of New York Press, Albany, New York.
- Haraway, D., 1988. Situated knowledges: the science question in feminism and the privilege of partial perspective. *Fem. Stud.* 14 (3), 575–599.
- Haraway, D., 2016. *Staying With the Trouble: Making Kin in the Chthulucene*. Duke University Press, Durham and London.
- Hickling-Hudson, A., 2004. South–south collaboration: Cuban teachers in Jamaica and Namibia. *Comp. Educ.* 40 (2), 289–311.
- Holbraad, M., Pedersen, M.A., 2017. *The Ontological Turn: An Anthropological Exposition*. Cambridge University Press, Cambridge.
- Hörschelmann, K., Stenning, A., 2008. Ethnographies of postsocialist change. *Prog. Hum. Geogr.* 32 (3), 339–361.
- Jensen, C.B., 2017. New ontologies? Reflections on some recent “turns” in STS, anthropology and philosophy. *Soc. Anthropol.* 25 (4), 525–545.
- Jules, T.D., Sá e Silva, M.M.D., 2008. How different disciplines have approached South-South cooperation and transfer. *Soc. Int. Educ. J.* 5 (1), 45–64.
- Jullien, M.A., 1817. *Esquisse et vues préliminaires d'un ouvrage sur l'éducation comparée* [Outline and Preliminary Views for a Work on Comparative Education]. L. Colas, Paris.
- Kazamias, A.M., 2009. Forgotten men, forgotten themes: the historical-philosophical-cultural and liberal humanist motif in comparative education. In: *International Handbook of Comparative Education*. Springer, Dordrecht, pp. 37–58.
- Kuswandono, P., Gandana, I., Rohani, S., 2015. Towards deimperialization of instructional strategies: cases of task-based learning and reflective practice implementation in the Indonesian context. In: *Asia as Method in Education Studies*. Routledge, pp. 97–110.
- Le Grange, L., 2018. The notion of Ubuntu and the (post) humanist condition. In: *Indigenous Philosophies of Education Around the World*. Routledge, pp. 40–60.
- Lee, Y., 2019. A critical dialogue with “Asia as method”: a response from Korean education. *Educ. Philos. Theor.* 51 (9), 958–969.
- Lin, J., Culham, T., Edwards, S. (Eds.), 2019. *Contemplative Pedagogies for Transformative Teaching, Learning, and Being*. Information Age, Charlotte, NC.
- Luschei, T.F., 2004. Timing is everything: the intersection of borrowing and lending in Brazil’s adoption of Escuela Nueva. In: *The Global Politics of Educational Borrowing and Lending*, pp. 154–167.
- McCowan, T., 2003. Participation and education in the landless people’s movement of Brazil. *J. Crit. Educ. Pol. Stud.* 1 (1), 1–18.
- Mercier, T.C., 2019. Uses of “the pluriverse”: cosmos, interrupted-or the others of humanities/uses of “the pluriverse”: cosmos, interrupted-or the others of humanities. *Ostium* (2).
- Mignolo, W., 2011. *The Darker Side of Western Modernity: Global Futures, Decolonial Options*. Duke University Press, Durham, NC.
- Mignolo, W.D., 2021. Coloniality and globalization: a decolonial take. *Globalizations* 18 (5), 720–737.
- Mundy, K., Murphy, L., 2001. Transnational advocacy, global civil society? Emerging evidence from the field of education. *Comp. Educ. Rev.* 45 (1), 85–126.

- Morris, K., 2021. Navigating the Postqualitative, New Materialist and Critical Posthumanist Terrain Across Disciplines: An Introductory Guide. Routledge, New York.
- Nishihira, T., Rappleye, J., 2021. Unlearning as (Japanese) learning. *Educ. Philos. Theor.* 1–14.
- Noah, H.J., Eckstein, M.A., 1969. *Towards a Science of Comparative Education*. Macmillan, New York.
- Noah, H., 2006. U.S. Social and Educational Research during the cold war: an interview with Harold J. Noah by Gita Steiner-Khamsi. *Eur. Educ.* 38 (3), 9–18.
- Park, J., 2016. Asian education and Asia as method. In: Lam, C.M., Park, J. (Eds.), *Sociological and Philosophical Perspectives on Education in the Asia-Pacific Region*. Springer, Singapore, Heidelberg, New York, Dordrecht, London, pp. 205–225.
- Petersen, E.B., 2018. "Data found us": a critique of some new materialist tropes in educational research. *Res. Educ.* 101 (1), 5–16.
- Phillips, D., Ochs, K., 2003. Processes of policy borrowing in education: some explanatory and analytical devices. *Comp. Educ.* 39 (4), 451–461.
- Phillips, D., Ochs, K. (Eds.), 2004. *Educational Policy Borrowing: Historical Perspectives*. Symposium Books Ltd.
- Phillips, D., 2004. Toward a theory of policy attraction in education. In: *The Global Politics of Educational Borrowing and Lending*, pp. 54–67.
- Piattoeva, N., Klutas, A., Suominen, O., 2018. Making and mobilizing contexts in policy and research. In: *World Yearbook of Education 2019*. Routledge, pp. 202–218.
- Piattoeva, N., 2021. Numbers and their contexts: how quantified actors narrate numbers and decontextualization. *Educ. Assess. Eval. Account.* 33, 511–533.
- Querejazu, A., 2016. Encountering the pluriverse: looking for alternatives in other worlds. *Rev. Bras. Pol. Int.* 59.
- Radhakrishnan, R., 2009. Why compare? *New Lit. Hist.* 40 (3), 453–471. Retrieved July 2, 2021, from <http://www.jstor.org/stable/27760271>.
- Rappleye, J., 2006. Theorizing educational transfer: toward a conceptual map of the context of cross-national attraction. *Res. Comp. Int. Educ.* (3), 223–240.
- Robertson, S.L., Mundy, K., Verger, A., Menashy, F., 2012. An introduction to public private partnerships and education governance. In: *Public Private Partnerships in Education*. Edward Elgar Publishing.
- Rogers, E.M., 2003. *Diffusion of Innovations*, fifth ed. The Free Press, New York.
- Sá e Silva, M.M.D., 2009. South-South Cooperation: Past and Present Conceptualization and Practice. *South-South Cooperation in Education and Development*, pp. 39–61.
- Santos, B.D.S., 2007. *Another Knowledge is Possible: Beyond Northern Epistemologies*. Verso.
- Santos, B.D.S., 2018. *The End of the Cognitive Empire. The Coming of Age of Epistemologies of the South*. Duke University Press.
- Shahjahan, R., 2005. Spirituality in the academy: reclaiming from the margins and evoking a transformative way of knowing the world. *Int. J. Qual. Stud. Educ.* 18 (6), 685–711.
- Shahjahan, R., 2004. Centering spirituality in the academy: towards a transformative way of teaching and learning. *J. Transformative Educ.* 2 (4), 294–312.
- Silova, I., Rappleye, J., 2015. Beyond the world culture debate in comparative education: critiques, alternatives and a noisy conversation. *Glob. Soc. Educ.* 13 (1), 1–7.
- Silova, I., Millei, Z., Piattoeva, N., 2017. Interrupting the coloniality of knowledge production in comparative education: postsocialist and postcolonial dialogues after the Cold War. *Comp. Educ. Rev.* 61 (S1), S74–S102.
- Silova, I., Piattoeva, N., Millei, Z., 2018. *Childhood and Schooling in (Post) Socialist Societies. Memories of Everyday Life*. Houndmill/Basingstoke.
- Silova, I., Rappleye, J., Auld, E., 2020a. Beyond the Western Horizon: Rethinking Education, Values, and Policy Transfer. *The International Handbook of Education Policy Studies*. Springer, Dordrecht.
- Silova, I., Rappleye, J., You, Y., 2020b. Beyond the Western horizon in educational research: toward a deeper dialogue about our interdependent futures. *ECNU Rev. Educ.* 3 (1), 3–19.
- Silova, I., 2010. Private tutoring in eastern Europe and central Asia: policy choices and implications. *Compare* 40 (3), 327–344.
- Silova, I., 2020. Anticipating other worlds, animating our selves: an invitation to comparative education. *ECNU Rev. Educ.* 3 (1), 138–159.
- Smith, T.L., Tuck, E., Yang, K.W., 2018. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, New York.
- Sobe, N.W., Kowalczyk, J.A., 2013. Exploding the cube: revisioning "context" in the field of comparative education. *Curr. Issues Comp. Educ.* 16 (1), 6–12.
- Staheli, U., 2003. The outside of the global. *CR N. Centen. Rev.* 3 (2), 1–22.
- Steiner-Khamsi, G., Waldow, F. (Eds.), 2012. *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. Routledge.
- Steiner-Khamsi, G. (Ed.), 2004. *The Global Politics of Educational Borrowing and Lending*. Teachers College Press.
- Steiner-Khamsi, G., 2006. The development turn in comparative education. *Eur. Educ.* 38 (3), 19–47.
- Stengers, I., 2011. Comparison as a matter of concern. *Common. Knowl.* 17 (1), 48–63.
- Stengers, I., 2012. Reclaiming animism. *E-Flux J.* 36, 1–10.
- Stromquist, N.P., 2016a. *Feminist Organizations and Social Transformation in Latin America*. Routledge.
- Stromquist, N.P., 2016b. Using regression analysis to predict countries' economic growth: illusion and fact in education policy. *Real-World Econ. Rev.* 76, 65–74.
- Special issue of Comparative Education Review. In: Takayama, K., Sriprakash, A., Connell, R. (Eds.), 2017a. *Contesting Coloniality: Rethinking Knowledge Production and Circulation in Comparative and International Education*, vol. 61. S1.
- Takayama, K., Sriprakash, A., Connell, R., 2017b. Toward a postcolonial comparative and international education. *Comp. Educ. Rev.* 61 (S1), S1–S24.
- Takayama, K., 2011. A comparativist's predicaments of writing about "other" education: a self-reflective, critical review of studies of Japanese education. *Comp. Educ.* 47 (4), 449–470. <https://doi.org/10.1080/03050068.2011.561542>.
- Takayama, K., 2015. Provincialising the world culture theory debate: critical insights from a margin. *Glob. Soc. Educ.* 13 (1), 34–57.
- Takayama, K., 2016. Beyond "the West as method": repositioning the Japanese education research communities in/against the global structure of academic knowledge. *Educ. Stud. Japan* 10, 19–31.
- Takayama, K., 2020. An invitation to negative comparative education. *Comp. Educ. Rev.* 56 (1), 79–95.
- Tan, C., 2019. Beyond the competencies agenda in large-scale international assessments: a confucian alternative. *Philos. Inq. Educ.* 26 (1), 20–32.
- Tarlau, R., 2012. Soviets in the countryside: the MST's remaking of socialist educational practices in Brazil. In: *Logics of Socialist Education*. Springer, Dordrecht, pp. 53–72.
- Tikly, L., 1999. Postcolonialism and comparative education. *Int. Rev. Educ.* 45 (5/6), 603–621.
- Tikly, L., 2004. Education and the new imperialism. *Comp. Educ.* 40 (2), 173–198.
- Tlostanova, M., Mignolo, W., 2012. *Learning to Unlearn: Decolonial Reflections From Eurasia and the Americas*. The Ohio State University Press, Columbus.
- Tlostanova, M., Thapar-Björkert, S., Koobak, R., 2016. Border thinking and disidentification: postcolonial and postsocialist feminist dialogues. *Fem. Theor.* 17 (2), 211–228.
- Tlostanova, M., 2012. Postsocialist ≠ postcolonial? On post-Soviet imaginary and global coloniality. *J. Postcolonial Writ.* 48 (2), 130–142.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Unterhalter, E., 2017. A review of public private partnerships around girls' education in developing countries: flicking gender equality on and off. *J. Int. Comp. Soc. Policy* 33 (2), 181–199.
- Vázquez, R., 2011. Translation as erasure: thoughts on modernity's epistemic violence. *J. Hist. Sociol.* 24 (1), 27–44.
- Verger, A., Moschetti, M., 2017. *Public-Private Partnerships as an Education Policy Approach: Multiple Meanings, Risks and Challenges*. Education Research and Foresight Series. No. 19.
- Verger, A., 2012. Framing and selling global education policy: the promotion of public-private partnerships for education in low-income contexts. *J. Educ. Pol.* 27 (1), 109–130.
- Wagner, P., 2002. *A Sociology of Modernity: Liberty and Discipline*. Routledge.
- Williams, H.M.A., Bermeo, M.J., 2020. A decolonial imperative: pluriversal rights education. *Int. J. Human Rights Educ.* 4 (1), 1.
- Wu, J., Wenning, M., 2016. The postsecular turn in education: lessons from the mindfulness movement and the revival of confucian academies. *Stud. Philos. Educ.* 35 (6), 551–571.
- Yonezawa, A., Kitamura, Y., Yamamoto, B., Tokunaga, T. (Eds.), 2018. *Japanese Education in a Global Age: Sociological Reflections and Future Directions*, vol. 46. Springer.
- You, Y., 2020a. The "new Orientalism": education policy borrowing and representations of East Asia. *Compare* 50 (5), 742–763. <https://doi.org/10.1080/03057925.2018.1559038>.

- You, Y., 2020b. Learning experience: an alternative understanding inspired by thinking through confucius. *ECNU Rev. Educ.* 3 (1), 66–87.
- Zhang, H., Chan, P.W.K., Kenway, J. (Eds.), 2015. *Asia as Method in Education Studies: A Defiant Research Imagination*. Routledge.
- Zhao, G., 2009. Two notions of transcendence: confucian man and modern subject. *J. Chin. Philos.* 36 (3), 391–407.
- Zhao, W., 2019a. Historicizing tianrenheyi as correlative cosmology for rethinking education in modern China and beyond. *Educ. Philos. Theor.* 51 (11), 1106–1116.
- Zhao, W., 2019b. Daoist onto-un-learning as a radical form of study: reimagining study and learning from an Eastern perspective. *Stud. Philos. Educ.* 38 (3), 261–273.
- Zhao, W., 2020. Eastern ethico-onto-epistemologies as a diffracting return: implications for post-qualitative pedagogy and research. In: *Navigating the Postqualitative, New Materialist and Critical Posthumanist Terrain Across Disciplines: An Introductory Guide*. Routledge, New York, pp. 85–98.

Global mobility and middle class families—parenting and education

Claire Maxwell^a and Miri Yemini^b, ^a Department of Sociology, University of Copenhagen, Copenhagen, Denmark; and ^b School of Education, Tel Aviv University, Tel Aviv, Israel

© 2023 Elsevier Ltd. All rights reserved.

Living mobility	303
A sense of belonging	304
Strategies for social reproduction	304
Conclusions	305
References	306

Migration is a phenomenon with a very long history and it figures centrally as a key social, economic and political question in today's world (Burrell, 2010; De Haas, 2010; Kunz, 2016). Research into the experiences of migrant families has focused on a range of themes—the reasons for migration (McGhee et al., 2017); parental strategies around engaging with and attempting to benefit from the host education systems (Antony-Newman, 2020; Byrne and De Tona, 2012; Lopez Rodriguez, 2010); the ways in which specific migrant groups are “managed” by governments, integrated into new education systems (Faas, 2016) and how various groups of migrants understand their autonomy in relation to the services they encounter (Kunz, 2020). One further domain of research examines the perceptions of migrant families in relation to their homeland (Faas, 2016). This line of study tends to deal with identity formation and its negotiation between where people have come from, their journeys to their current location, and experiences in their new homes. Their relationships with and within their new places of residence are drawn on to develop self-definitions and articulate self-distinctions (Ryan, 2011). Traditionally, research on migration has focused on immigrants from less developed countries who moved to Western Europe and North America (Massey et al., 1993; Portes and Rumbaut, 2006). More recently, researchers have also begun to explore the mobility practices of other groups, including various economic elites, but also highly-skilled professionals who are very mobile due to their professional responsibilities and aspirations (e.g., Koh and Wissink, 2018; Kunz, 2016; Scott, 2006).

With the “deterritorialization of capital” (Embong, 2000: 991), it has been argued, comes the emergence of a transnational capitalist class (TCC) (Sklair, 2002) and an accompanying global middle class (GMC) (Ball, 2010). It is claimed that the GMC are constituted as a group of globally mobile professionals who have become essential in facilitating the exponential growth of multinational corporations. These professionals, do not own the capital and therefore are not understood as an elite, but are defined as providing the expert knowledge and skills needed for the operation of such business entities by facilitating the companies’ global networks of production, consumption, and bureaucracy. Other terms sometimes deployed to describe them are expatriates or, more recently, “international talent”. They are highly skilled workers of diverse national origins who circulate the globe—mostly between key global cities such as New York, London, Berlin, Sydney, San Francisco, and Hong Kong—and serve as high-tech, financial, and legal specialists; middle managers; engineers; and other professions (Ball, 2010; Beaverstock, 2005; Embong, 2000; Sassen, 1999, 2000; Sklair, 2001; Yeoh and Willis, 2005). This new group of globally mobile professionals are crucial for the transfer of specialized knowledge internationally and the consolidation of global networks. They are conceptualized as a service class to the transnational capitalist class, and as such understood to be a social class of their own, different from elites but also distinct from non-mobile middle classes in their home nations. Thus, as a group that are untethered from the nation state and always on the move, their social position is not, theoretically, determined by a national field of social relations. In earlier writings on modern societies and the cosmopolitanization of space (Beck, 2012), this group are presented as essentially nomads roaming a transnational space. It is this nation-less existence that is argued to bind them as a group—“the global middle class”.

This is a fundamental empirical and theoretical question for sociology today—is this collection of globally mobile professionals, who are differentiated from others by their expected and relatively frequent mobility for work, a new fraction of the middle classes? Do their orientations to the self, others around them, questions of belonging, capital resources that are valorized, and their desired imagined futures bind them as a “well-formed class with its own distinct social identity” (Lockwood, 1995: 3)? Before we seek to explore what we know about this group of globally mobile professionals in terms of their social class description, it is important to highlight that there are various definitions of the “global middle class”. The first is the one we offer above—professionals, who are highly educated, who are mobile for work and take up positions that facilitate the work of multinational organizations and corporations (Ball, 2010). Originally, they were largely from the Global North, and often known as expatriates. Increasingly now though, these professionals come from all over the world and mobility trajectories are not limited to the large cities of the Global North, but also to metropolises across the Global South (Friedberg, 2007). This makes the study of these globally mobile professionals even more interesting in terms of seeking to define them and ascertain whether that are norms and practices that bind them, but also more challenging as their heterogeneity increases.

However, the term “global middle class” has also been used to describe the significant growth of middle class segments within countries across the Global South. Researchers have been tracking the distribution of populations by income, and noting significant growth, at least in certain regions, of people who can now be considered ‘middle class’ (Kochhar, 2020). This definition of the global

middle class understands them as members of this particular group on purely economic terms (Das, 2009), so for instance in this report: “as all those living in households with daily per capita incomes of between USD10 and USD100 in PPP terms” (Kharas, 2010: 6). Thurow (1987) argued that all people with an income in the interval from 75% to 125% of median income of a particular nation should be considered part of the global middle class. Koo (2016), meanwhile, offers a third definition for the global middle class, which seeks to integrate a focus on class-making practices: “the affluent and globally oriented segments of the middle classes in developing countries” (Koo, 2016: 440; see also Derné, 2005). Koo’s focus on Korea allows him to illustrate how these more globally-oriented members of a society leads to internal divisions within the Korean middle classes, as practices of distinction in relation to education and mobility come into play. Koo (2016) argues that the global middle class he studied were a: “globally oriented, globally connected, and globally mobile segment of the middle class” (p. 449). We found similar distinctions in orientations in our mobile and “immobile” Israeli middle class participants (Yemini et al., 2019), but argue here that regular global mobility for work is an important distinction between those members of a middle class who are globally oriented and perhaps mobile for education, work trips and leisure travel but remain reasonably strongly embedded within their “home” nation, and those who become more transnational as a regular feature of their lives. Koo (2016) also points to this differentiation, but it is here that the scholarly work is not developed enough—is there is social class fraction distinction between a more “locally-moored” middle class and one that is transnational, and what degree of transnationality is needed to push a person into this second grouping?

To ensure we are consistent in our definition of the global middle class—we focus here on globally mobile professionals—who are relatively and continuously mobile for work—that is to say—they have moved for work purposes more than twice and continue to be open to the prospect of future mobility. In what follows we offer a summary of the relevant research done to date on globally mobile professionals and their family lives—with a particular emphasis on education and other parenting strategies. In this way we are able to highlight what appears to connect this group of people, in terms of family practices, future aspirations and how children’s needs intersect with professional desires and demands to drive this group of individuals’ practices; while also seeking to illuminate where and why they might diverge. Our overview is organized into three sections: (1) what it means to lead lives that are mobile; (2) their sense of national belonging; (3) and their strategies for social reproduction, when detached from a nation state.

Living mobility

One of the most distinguishing features of this group of global professionals is their pervasive mobility. Mobility means moving to live in one or more places outside their “home” country, having to travel regularly for work and liaise with professionals based around the world, while also, necessarily, maintaining social and family networks virtually and through travel. As Favell (2008) found in his study of young professionals moving across Europe for work, mobility was articulated as “liberating” them from some of the national—social and cultural—structures they perceived as oppressive back “home.” Yet, many of his participants also anticipated returning “home” when they planned on starting a family. What about those professionals and their families who keep moving? Questions around what motivates this engagement with mobility, how it shapes family practices and educational strategies, whether it constructs a particular orientation to the world, how transnational ties are maintained and the ways they affect the societies they enter are all pertinent. Examining these issues also allows us to identify ways these globally mobile professionals are similar or different from others who also migrate (but come from differently resourced backgrounds), middle class peers from the societies they come from and those they enter, and how their relative social class position is affected by the hierarchical constitution of the countries and cities they relocate to.

One of the most fundamental questions research should engage with is how being constantly mobile is conceptualized by those who make a choice to relocate themselves and their families for work. What anticipated benefits and challenges do they see, and how do they work to smooth the transition from place to place? For some, relocation is necessitated/required of them by employers, for others it is an active choice—to leave behind a social and political system they have become disengaged from (Yemini and Maxwell, 2018; Adams and Agbenyega, 2019), or it might be a desire to facilitate aspirations for their children’s futures to be located anywhere in the world (Yemini et al., 2019), as they seek for their children to become “the best of the best” and open new possibilities and horizons for them.

But mobility is unsettling—as breaks with “home” and “what we know” occur, and they are uncertainties about how to thrive in new fields of social relations and education systems. In our work we have drawn on the concept of boundary objects (Akkerman and Bakker, 2011) to help theorize how families work to maintain some familiarity during a relocation with “what they know”, while simultaneously drawing on already-acquired resources (such as language or cosmopolitan orientations to difference) to facilitate the movement into a new geographical and social space (Yemini et al., 2020). But, as the practice of relocation becomes more habitual, the motivations for mobility and the ways it is managed will necessarily change (as demonstrated in Beech et al., 2021, where school choice strategies were found to change over time and number of relocations). Another facet of living mobility is the work needed to maintain an ever disparate social network of family and friends, located all around the world, alongside the continued cultivation of a sense of belonging and identity (see the next sub-section).

As mobility becomes more or less habitual, tracing how people variously draw on their resources to secure their positions in a new space is an important focus for research. In particular, focusing on how globally mobile professional parents seek to secure their children’s future advantages will facilitate an examination of what might constitute middle class practices in transnational space—a hitherto unresolved theoretical quandary. From the scarce research that exists, it can be suggested that practicing mobility does distinguish the kinds of values and family practices identified within global professional families, when compared with

migrant (i.e., mobile once) and non-mobile families who are similarly well-educated and have professional jobs (Yemini et al., 2019). However, a much larger empirical base is needed to study this question.

Some scholars have drawn on Bourdieu to examine whether capitals from one social field can be transferred to another—whether their value is recognized and can be exchanged (Soong et al., 2018; Waddling et al., 2019). This approach could be employed to examine whether particular middle class capitals can be transnational (perhaps a university degree from an internationally-renowned university); and/or to examine how different (nationally-located) social fields differ from one another; and how middle class people moving into a new space can compete with those emplaced social groups who have been struggling over the articulation of specific capital resources. Previous research has argued that the “geopolitical relations” (Lillie, 2021a: 91) or positionings of a person’s “home” country will affect how transferable privilege from one setting to another is (see also Lillie, 2021b; Soong et al., 2018).

A sense of belonging

If a feature of being globally mobile is being untethered from the nation state, how do globally mobile professionals and their families articulate a sense of identity, as this is so often linked to “the nation”—either as “being from somewhere” or as now belonging “somewhere new”. Meanwhile, a sense of identity and/or belonging may be more connected to ethnicity and/or religion, and so does not necessarily need the nation state as an object for recognition. But work on globally mobile professionals opens up a new line of enquiry—does a sense of belonging emerge that connects these diverse individuals (diverse in terms of their citizenship, ethnicity, religion etc) because of their mobility or “being global”? Here the concept of “frames of reference” (Savage et al., 2005) might help in examining this further. Do globally mobile professionals draw on local, national or global frames of reference that are in some way connected to geographical space? Or perhaps they engage with a differently constituted idea of spatial relations which connects them to others “like me”.

A paper on the school choices made by globally mobile families, for instance, demonstrates that Chinese ethnicity (Chinese–Malaysian, Chinese–Singaporean etc) is critical to the anchoring work (Grzymala-Kazłowska, 2018) these families do in creating a sense of identity (Beech et al., 2021). Or, as Agbaria (2019) argues, religious-belonging might be a central connecting feature that overrides ethnicity or nation. Meanwhile, for other globally mobile families, maybe the “global” becomes their frame of reference as any other spatial configurations are too disparate to embed themselves in. A global frame of reference invites the suggestion of a cosmopolitan or global citizenship orientation (Goren and Yemini, 2016, 2017); namely, of acceptance, interest, and comfort in engaging with the “Other” but also potentially about being the “Other” in a particular context. While some of the limited research on this group identifies a commitment to cosmopolitanism in these families’ practices (Maxwell and Yemini, 2019), other more concrete forms of elective belonging (Savage et al., 2005) also emerge (Yemini et al., 2020). This might be connected to the psychic need to have belonging tethered to something much more secure, or may be a practice of distinction that is seen to complement the families’ cosmopolitan status (Howard and Maxwell, 2021).

A second articulation of “belonging” or “identity”, might be understood as the relations of sociability globally mobile professional families forge in their new countries of residence. Favell’s (2008) research on the “Eurostars”, found that his participants tended to congregate with “people like us” when they moved to live in new countries. Meanwhile, Andreotti et al. (2015) found that those mobile professionals in the European cities they studied, still retained strong links to family “back home” and engaged only in a “selective rootedness” (p. 180) when they moved somewhere new. Whether families “integrate” in their new communities, often depends on the school choices they make for their children, which residential areas they move to, what friends and family they may already know who are also located there, the leisure activities they take up, and how critical knowing the country’s main language is for forming friendships (Yemini and Maxwell, 2018). If globally mobile professionals congregate in real or imagined gated communities and attend a school that follows their home country curriculum—ties to the home nation stay firmly intact (Toh, 2020).

A third concept, taking from migration research, which could be insightful is that of “incorporation” (Beaverstock, 2002; Dubucs et al., 2017; Jaskulowski, 2018; Plög and Becker, 2015; Scott, 2006). The focus of “incorporation” can be within the companies that employ these globally mobile professionals, but also within the local communities they join (residential areas, shared communities such as those that form around a school, shared interests, nationality, ethnicity and/or religion).

Strategies for social reproduction

There is a relatively large body of scholarly work highlighting the “concerted cultivation” strategies that middle-class parents worldwide engage in Irwin and Elley (2011), Lareau (2003), Nogueira (2010), Van Zanten (2009), Vincent et al. (2012). Whether directly drawing on a Bourdieusian framework or not, middle-class parents are argued to be ambitious in seeking to secure and extend their children’s advantages, expose their sons and daughters to a broad range of extra-curricular opportunities, help them to identify their talents, but also develop them as a whole rounded person (Stefansen and Aarseth, 2011). Such practices are argued to maintain current relations of inequality, where the middle classes continue to secure and extend their advantages over their working class peers (Weis et al., 2014). As Van Zanten (2015) argues, middle-class parents can draw on their own knowledge and personal experiences but also on professional experience of how to maneuver through national systems of education to their children’s

advantage—conceptualized as a form of cultural capital. Critically, it is also the economic resources that are available to middle-class families that facilitate strategies of concerted cultivation—paid extra-curricular activities, private academic tutoring, (usually) mothers able to work part-time or become a full-time stay-at-home-parent, trips abroad for leisure but also language tuition, and so forth (Van Zanten, 2015). What is usually highlighted in social reproduction strategies practiced by the middle classes is their dependence on an understanding of, and familiarity with, the state's structures and institutions, so that the capitals they have acquired over time have value in that particular field, and can therefore be activated to secure advantage and therefore social reproduction.

Globally mobile professional families are also “middle class” in terms of their education, income and professional status. But a central question for this field of study is how being constantly globally mobile affects parenting and education strategies. As Breidenstein et al. (2018) demonstrated in their study of globally mobile professionals interacting on a website discussion forum about school choices in Berlin, some parents want to replicate what they know in a new space, while others seek to embrace the new opportunities that mobility affords them and their children. Another study of globally mobile middle-class families in London (Yemini and Maxwell, 2018) found that while parental aspirations are being negotiated within a global frame of reference in terms of future employment and anticipated mobility of their children, significant resources are also invested in order to secure the best opportunities for their children in their current locations (via the area of residence chosen, the school attended, the friendship networks promoted, extra-curricular activities engaged in, how family and other developments “back home” are kept abreast of etc.). Knowledge of the local system they have entered is accumulated through networks and connections. This raises the critical issue, highlighted by Lareau et al. (2016), that in order to make “choices” or draw on the necessary capitals to shape practices to their advantage, parents must understand “the rules of the game,” as Bourdieu (1984) would argue. While Lareau et al. (2016)'s study is focused on a non-mobile population in one part of the US seeking to gain access to specific early-childhood institutions, their argument is even more critical for mobile families as they settle into a new and usually unfamiliar cultural and organizational spaces (Waddling et al., 2019). Thus, research is needed to understand how GMC families decide upon entering a new education system how best to successfully maneuver their way through. One paper tackles this question directly (Beech et al., 2021), arguing that as the number of moves increases among globally mobile professional families, their school choice strategies change too, as mobility becomes more habitual and more deeply engrained in their values and anticipated futures.

Conclusions

We know too little about this growing group to make many definitive statements about them. As their number and diversity grow (diversity in terms of where they “come from” and where they are moving to), the research that is being done is not keeping pace. However, the theoretical contributions studies to date have offered (see a summary in Maxwell et al., 2019) give us some tools with which to think further about how global mobility is affecting both the lives of globally mobile professional families, but also the host societies they leave and enter. This latter potential effect is not insignificant we would argue, just as other migrant communities alter the composition of cities and neighborhoods, globally mobile professionals have the resources to demand more and change local understandings of what is desirable in terms of education offers, travel expectations and future aspirations (see also Higginson et al., 2019; Yemini and Maxwell, 2021, with an important challenge to this argument by Windle and Maire, 2019).

Despite the lack of comprehensive data sets to examine the question—are globally mobile professionals a “well-formed class with its own distinct social identity” (Lockwood, 1995: 3)—we would suggest there are likely different fractions that make up a loosely-formed group bound in some way by actual or anticipated frequent mobility. The range of professional backgrounds (working for the private or public sector, working across different industries, in a range of professions—Scott, 2006) suggests that other distinctions within the middle classes—cultural and economic fractions (Ball et al., 2004; Aarseth, 2018)—will also shape the family practices of those professionals who are globally mobile. This differentiation (e.g., among corporate professionals, UN civil servants and diplomats, university academics) has not yet been comparatively studied. Furthermore, Urry's (1999) emphasis on the range of mobilities is also important to consider further—differences between families that relocate frequently, those who move back and forth between “home” and a posting by their employer, those that do not necessarily relocate for work continuously, but have a “mobile” orientation in terms of travel, social networks, anticipated futures.

The two arguments we make above—about globally mobile professionals having a broader impact on societies, and that important differentiations are likely to exist within this group, fits well with Kaufmann et al. (2004) understanding of mobility as a form of capital. They argue mobility is not just about spatial movement, but rather about how such spatial movement is made possible or is perceived as being possible through social position. Thus, according to Kaufmann et al. (2004), physical mobility not only affects an individual's or group's positioning within the social structure, but will also itself affect social structures more broadly. Kaufmann et al. (2004) therefore re-conceptualize mobility as “motility” in order to emphasize its association with a form of capital, and define motility as examining “how entities access and appropriate the capacity for socio-spatial mobility” (Kaufmann et al., 2004: 750). Critically, Kaufmann et al. focus on both the potential and actual capacity for mobility—alerting us to the importance of aspirations in shoring up practices (Kaufmann et al., 2018).

This article has sought to discuss how we might conceptualize and understand a middle class group of professionals who, with their families, regularly relocate for work. Drawing on recent empirical studies, we have been able to summarize some ways in which this group is distinctive, and also how in other ways there are important differences to be found within. Whether or not there are enough similarities that bind them for them to be constituted as a “global middle class”, it is too early to say. However, the ways mobility shapes understandings of self, the future, and family practices, means these middle class families are continuously

negotiating local, national and transnational spaces of capital accumulation and activation, as well as a sense of belonging that straddles the glocal (Marginson and Rhoades, 2002). In this way, we argue they are an important, growing social grouping who affect societies they leave and those they relocate to. The research we have summarized focuses on educational and parenting strategies, and highlights that despite the uncertainty mobility can engender, these families are imagining futures on a global scale, seeking to create cosmopolitan subjects that nonetheless have roots in various communities of belonging.

References

- Aarseth, H., 2018. Fear of falling – Fear of fading: the emotional dynamics of positional and personalised individualism. *Sociology* 52 (5), 1087–1102.
- Adams, M., Agbenyega, J., 2019. Futurescapes: school choice of internationally mobile global middle class families temporarily residing in Malaysia. *Discourse* 40 (5), 647–665.
- Agbaria, A.K., 2019. Religion and the global middle class: towards a new research agenda. *Discourse* 40 (5), 734–742.
- Akkerman, S.F., Bakker, A., 2011. Boundary crossing and boundary objects. *Rev. Educ. Res.* 81 (2), 132–169.
- Andreotti, A., Le Galès, P., Moreno-Fuentes, F.J., 2015. Globalized Minds, Roots in the City: Urban Upper-Middle Classes in Europe. John Wiley & Sons.
- Antony-Newman, M., 2020. Parental involvement of Eastern European immigrant parents in Canada: whose involvement has capital? *Br. J. Sociol. Educ.* 41 (1), 111–126.
- Ball, S.J., Vincent, C., Kemp, S., Pietikainen, S., 2004. Middle class fractions, childcare and the 'relational' and 'normative' aspects of class practices. *Sociol. Rev.* 54 (4), 478–502.
- Ball, S., 2010. Is there a global middle class? The beginnings of a cosmopolitan sociology of education: a review. *J. Comp. Educ.* 69 (1), 137–161.
- Beaverstock, J., 2002. Transnational elites in global cities: British expatriates in Singapore's financial district. *Geoforum* 33 (4), 525–538.
- Beaverstock, J.V., 2005. Transnational elites in the city: British highly-skilled inter-company transferees in New York City's financial district. *J. Ethn. Migr. Stud.* 31, 245–268.
- Beck, U., 2012. Redefining the sociological project: the cosmopolitan challenge. *Sociology* 46 (1), 7–12.
- Beech, J., Koh, A., Maxwell, C., Yemini, M., Tucker, K., Barrenechea, I., 2021. "Cosmopolitan start-up" capital: mobility and school choices of global middle class parents. *Camb. J. Educ.* 51 (5), 1–15.
- Bourdieu, P., 1984. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press.
- Breidenstein, G., Forsey, M., La Gro, F., Krüger, J.O., Roch, A., 2018. Choosing international: a case study of globally mobile parents. In: *Elite Education and Internationalisation*. Palgrave Macmillan, Cham, pp. 161–179.
- Burrell, K., 2010. Staying, returning, working and living: key themes in current academic research undertaken in the UK on migration movements from Eastern Europe. *Soc. Ident.* 16 (3), 297–308.
- Byrne, B., De Tona, C., 2012. Trying to find the extra choices': migrant parents and secondary school choice in Greater Manchester. *Br. J. Sociol. Educ.* 33 (1), 21–39.
- Das, D.K., 2009. Globalization and an emerging global middle class. *Econ. Aff.* 29 (3), 89–92.
- De Haas, H., 2010. The internal dynamics of migration processes: a theoretical inquiry. *J. Ethnic Migrat. Stud.* 36 (10), 1587–1617.
- Derné, S., 2005. Globalization and the making of a transnational middle class: implications for class analysis. In: McLaren, P., Jaramillo, N.E. (Eds.), *Alternative Globalizations: Critical Globalization Studies*. Routledge, London, pp. 177–186.
- Dubucs, H., Pfirsich, T., Recchi, E., Schmoll, C., 2017. Je suis un Italien de Paris: Italian migrants' incorporation in a European capital city. *J. Ethnic Migrat. Stud.* 43 (4), 578–595.
- Embong, A.R., 2000. Globalization and transnational class relations: some problems of conceptualization. *Third World Q.* 21 (6), 989–1000.
- Faas, D., 2016. *Negotiating Political Identities: Multiethnic Schools and Youth in Europe*. Routledge.
- Favell, A., 2008. *Eurostars and Eurocities: Free Movement and Mobility in an Integrating Europe*. Wiley-Blackwell, Oxford.
- Friedberg, R., 2007. The economic impact of knowledge workers from India and China. In: Tambar, U. (Ed.), *Movement of Global Talent: The Impact of High Skill Labor Flows from India and China*. Policy Research Institute for the Region, Woodrow Wilson School of Public and International Affairs, Princeton University, pp. 35–53.
- Goren, H., Yemini, M., 2016. Global citizenship education in context: teacher perceptions at an international school and a local Israeli school. *Compare* 46 (5), 832–853.
- Goren, H., Yemini, M., 2017. Global citizenship education redefined—a systematic review of empirical studies on global citizenship education. *Int. J. Educ. Res.* 82, 170–183.
- Grzymala-Kazłowska, A., 2018. From connecting to social anchoring: adaptation and "settlement" of Polish migrants in the UK. *J. Ethnic Migrat. Stud.* 44 (2), 252–269.
- Higginson, J., McLeod, J., Rizvi, F., 2019. Globally mobile middle class lives in government secondary schools. *Discourse* 40 (5), 633–646.
- Howard, A., Maxwell, C., 2021. Conferred cosmopolitanism: class-making strategies of elite schools across the world. *Br. J. Sociol. Educ.* 40 (2), 164–178.
- Irvin, S., Elley, S., 2011. Concerted cultivation? Parenting values, education and class diversity. *Sociology* 45 (3), 480–495.
- Jaskulowski, K., 2018. A study of local incorporation of migrant professionals in Wrocław. *Int. Migrat.* 56 (5), 123–136.
- Kaufmann, V., Bergman, M.M., Joye, D., 2004. Motility: mobility as capital. *Int. J. Urban Reg. Res.* 28 (4), 745–756.
- Kaufmann, V., Dubois, Y., Ravalet, E., 2018. Measuring and typifying mobility using motility. *Appl. Mobil.* 3 (2), 198–213.
- Kharas, H., 2010. *The Emerging Middle Class in Developing Countries*. OECD Development Centre Working Papers, No. 285. OECD Publishing, Paris. <https://doi.org/10.1787/5kmmpp8lncrns-en>.
- Kochhar, R., 2020. A global middle class is more promise than reality. In: Suter, C., Madheswaran, S., Vani, B.P. (Eds.), *The Middle Class in World Society: Negotiations, Diversities and Lived Experiences*. Routledge, India, pp. 15–48.
- Koh, S.Y., Wissink, B., 2018. Enabling, structuring and creating elite transnational lifestyles: intermediaries of the super-rich and the elite mobilities industry. *J. Ethnic Migrat. Stud.* 44 (4), 592–609.
- Koo, H., 2016. The global middle class: how is it made, what does it represent? *Globalizations* 13 (4), 440–453.
- Kunz, S., 2016. Privileged mobilities: locating the expatriate in migration scholarship. *Geogr. Compass* 10 (3), 89–101.
- Kunz, S., 2020. Expatriate, migrant? The social life of migration categories and the polyvalent mobility of race. *J. Ethnic Migrat. Stud.* 46 (11), 2145–2162.
- Lareau, A., Adia Evans, S., Yee, A., 2016. The rules of the game and the uncertain transmission of advantage: middle-class parents' search for an urban kindergarten. *Sociol. Educ.* 89 (4), 279–299.
- Lareau, A., 2003. *Unequal Childhoods*. University of California Press.
- Lillie, K., 2021a. Multi-sited understandings: complicating the role of elite schools in transnational class formation. *Br. J. Sociol. Educ.* 42 (1), 82–96.
- Lillie, K., 2021b. Mobile and elite: diaspora as a strategy for status maintenance in transitions to higher education. *Br. J. Educ. Stud.* <https://doi.org/10.1080/00071005.2021.1948965>.
- Lockwood, D., 1995. *Marking Out the Middle Class(es)*. Social Change and Middle Classes. UCL Press, London, pp. 1–12.
- Lopez Rodriguez, M., 2010. Migration and a quest for "normalcy": Polish migrant mothers and the capitalization of meritocratic opportunities in the UK. *Soc. Ident.* 16 (3), 339–358.
- Marginson, S., Rhoades, G., 2002. Beyond national states, markets, and systems of higher education: a glocal agency heuristic. *High Educ.* 43 (3), 281–309.
- Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., Taylor, J.E., 1993. Theories of international migration: a review and appraisal. *Popul. Dev. Rev.* 431–466.
- Maxwell, C., Yemini, M., 2019. Modalities of cosmopolitanism and mobility: parental education strategies of global, immigrant and local middle-class Israelis. *Discourse* 40 (5), 616–632.
- Maxwell, C., Yemini, M., Koh, A., Agbaria, A., 2019. The plurality of the global middle class(es) and their school choices—moving the "field" forward empirically and theoretically. *Discourse* 40 (5), 609–615.

- McGhee, D., Moreh, C., Vlachantoni, A., 2017. An “undeliberate determinacy”? The changing migration strategies of Polish migrants in the UK in times of Brexit. *J. Ethnic Migrat. Stud.* 43 (13), 2109–2130.
- Nogueira, M.A., 2010. A revisited theme—middle classes and the school. In: Apple, M.W., Ball, S.J., Gandin, L.A. (Eds.), *The Routledge International Handbook of the Sociology of Education*. Routledge, pp. 271–281.
- Plöger, J., Becker, A., 2015. Social networks and local incorporation—grounding high-skilled migrants in two German cities. *J. Ethnic Migrat. Stud.* 41 (10), 1517–1535.
- Portes, A., Rumbaut, R.G., 2006. *Immigrant America: A Portrait*. University of California Press.
- Ryan, L., 2011. Transnational relations: family migration among recent Polish migrants in London. *Int. Migrat.* 49 (2), 80–103.
- Sassen, S., 1999. Whose city is it. *Sustainable cities in the 21st century*, pp. 145–163.
- Sassen, S., 2000. New frontiers facing urban sociology at the Millennium. *Br. J. Sociol.* 51, 143–159.
- Savage, M., Warde, A., Devine, F., 2005. Capitals, assets, and resources: some critical issues 1. *Br. J. Sociol.* 56 (1), 31–47.
- Scott, S., 2006. The social morphology of skilled migration: the case of the British middle class in Paris. *J. Ethnic Migrat. Stud.* 32 (7), 1105–1129.
- Sklair, L., 2001. The transnational capitalist class. Blackwell, Oxford.
- Sklair, L., 2002. Democracy and the transnational capitalist class. *Ann. Am. Acad. Polit. Soc. Sci.* 581 (1), 144–157.
- Soong, H., Stahl, G., Shan, H., 2018. Transnational mobility through education: a Bourdieusian insight on life as middle transnationals in Australia and Canada. *Glob. Soc. Educ.* 16 (2), 241–253.
- Stefansen, K., Aarseth, H., 2011. Enriching intimacy: the role of the emotional in the “resourcing” of middle-class children. *Br. J. Sociol. Educ.* 32 (3), 389–405.
- Thurow, L., 1987. A surge in inequality. *Scientific American* 256, 30–37.
- Toh, G., 2020. Schooling (for) Japanese children in cosmopolitan Singapore: building bridges and erecting barriers. *Int. Stud. Sociol. Educ.* 29 (3), 274–292.
- Urry, J., 1999. Globalization and citizenship. *J. World Syst. Res.* 310–324.
- Van Zanten, A., 2009. The sociology of elite education. *Routledge Int. Handb. Sociol. Educ.* 329–339.
- Van Zanten, A., 2015. Promoting equality and reproducing privilege in elite educational tracks in France. In: Maxwell, C., Aggleton, P. (Eds.), *Elite Education: International Perspectives*. Routledge, pp. 114–125.
- Vincent, C., Rollock, N., Ball, S., Gillborn, D., 2012. Being strategic, being watchful, being determined: black middle-class parents and schooling. *Br. J. Sociol. Educ.* 33 (3), 337–354.
- Wadding, J., Bertilsson, E., Palme, M., 2019. Struggling with capital: a Bourdieusian analysis of educational strategies among internationally mobile middle class families in Sweden. *Discourse* 40 (5), 697–716.
- Weis, L., Cipollone, K., Jenkins, H., 2014. *Class Warfare: Class, Race, and College Admissions in Top-Tier Secondary Schools*. University of Chicago Press.
- Windle, J., Maire, Q., 2019. Beyond the global city: a comparative analysis of cosmopolitanism in middle-class educational strategies in Australia and Brazil. *Discourse* 40 (5), 717–733.
- Yemini, M., Maxwell, C., 2018. De-coupling or remaining closely coupled to “home”: educational strategies around identity-making and advantage of Israeli global middle-class families in London. *Br. J. Sociol. Educ.* 39 (7), 1030–1044.
- Yemini, M., Maxwell, C., 2021. Mobilities of policy and mobile parents-creating a new dynamic in policy borrowing within state schooling. *Glob. Soc. Educ.* 19 (1), 70–80.
- Yemini, M., Maxwell, C., Mizrahi, M.A., 2019. How does mobility shape parental strategies—a case of the Israeli global middle class and their “immobile” peers in Tel Aviv. *Glob. Soc. Educ.* 17 (3), 324–338.
- Yemini, M., Maxwell, C., Koh, A., Tucker, K., Barrenechea, I., Beech, J., 2020. Mobile nationalism: parenting and articulations of belonging among globally mobile professionals. *Sociology* 54 (6), 1212–1229.
- Yeoh, B., Willis, K., 2005. Singaporean and British transmigrants in China and the cultural politics of ‘contact zones’. *J. Ethn. Migr. Stud.* 31 (2), 269–285.

Transnational academic collaboration: a critical analysis of the global and national contexts and constraints of co-authorship

Terri Kim^{a,b,c}, ^aComparative Higher Education, UEL, London, United Kingdom; ^bUCL Institute of Education, London, United Kingdom; and ^cSt Antony's College, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	308
The macro context: politics of scientometrics and the datafication process	309
Micro context: organisational culture and social psychology of transnational academic collaboration	311
Ethical implications for transnational academic collaboration in co-authorship	312
Discussion and conclusion	314
References	315

Introduction

Academic collaboration is a common, integral feature of academic life - understood as academics working jointly especially in an intellectual endeavor. It takes on many forms and many levels of relationships in both formal and informal contexts. This article considers transnational academic research collaboration in the form of co-authorship in the humanities and social sciences (HSS).

Nurturing academic collaboration has been high on research policy agendas not just at national levels but also at transnational levels (Hoekman et al., 2009; Yegros-Yegros et al., 2021) such as the European Commission's Horizon 2020 programmes and the European Research Council's Synergy Grant schemes. The HSS field is no exception - e.g. Humanities in the European Research Area (HERA) - a partnership between 26 Humanities Research Councils across Europe and the European Commission - committed to strengthening the European platform for the humanities research. HERA (<https://heranet.info/about-us/governance-and-structure/network-of-members/>)'s Joint Research Program (JRP) aims to enable large transnational Collaborative Research Projects (CRPs) in the Humanities.

However, the policy emphasis on academic research collaboration reflects the trend of promoting a certain type of research designed for measurable socio-economic impacts that match the interests of both funders and disciplinary gatekeepers. For instance, Wellcome Trust also funds research in HSS "to form the bedrock of more focused and instrumental research projects that lead to improvements in health". It supports "collaborations between UK research networks and international partners to maximize the impact of their work" (<https://wellcome.org/what-we-do/our-work/research-humanities-and-social-sciences>).

Such funding schemes indicate that productivity in academic research is correlating with the high level of collaboration. According to the jobs.ac.uk survey (2012) (<https://www.jobs.ac.uk/media/pdf/careers/resources/making-academic-collaboration-work.pdf>), 91% of researchers agreed that collaboration increases research impact; 84% of researchers think that building personal, on-the-ground relationships (informal collaborations) are as important to research excellence as strategic initiatives (formal collaborations) in their institution; and 79% of researchers said that collaboration makes them more productive.

Given these known benefits of academic collaboration, there is also an increasing emphasis on funding collaboration. Many supranational and national funders seek to stimulate transnational academic research collaboration through a mix of grant schemes and grant criteria (e.g. the European Science Foundation (ESF) programmes, DAAD, NordForsk, JSPS, ESRC, AHRC, etc.). The recent UKRI call for applications: "Addressing COVID-19 challenges with Japanese researchers" is exclusively open to the HSS fields jointly funded by the ESRC and AHRC in the UK and the JSPS in Japan: (<https://www.ukri.org/opportunity/addressing-covid-19-challenges-with-japanese-researchers/>). Such measures are designed to fund collaborative research based on particular models of collaboration; yet, much of the disciplinary heterogeneity in research practices reside in the differing patterns of collaboration, including forms of publication authorship (Lewis et al., 2012).

This paper considers both the macro and micro contexts of transnational academic research collaboration and co-authorship in HSS. It is argued that academic collaboration in the form of "co-authorship" in the fields of HSS may be intrinsically an oxymoron given the nature of knowledge creation in HSS. Nevertheless, transnational academic collaboration in co-authorship is increasingly promoted in HSS as well, valorized and rewarded in the contemporary rules of the game in academia. International scholarly impacts are measured in various academic research evaluation schemes in many countries. In other words, the rapid surge of transnational academic collaboration is not only driven by the type of funded academic research *per se* but also closely entwined with the university performativity, participating in the neoliberal rent-seeking economy.

Accordingly, the paper analyses both the structural and agentic attributes to transnational academic collaboration and the embedded power relationships within. The paper examines, first of all, the macro-context, politics of scientometrics and the datafication process and its implications for transnational academic collaboration; secondly, the micro-context, organisational culture and social psychology of transnational academic collaboration; and thirdly, ethical implications of transnational

academic collaboration in co-authorship. The conclusion of the paper returns to its argument about the oxymoronic nature of transnational academic collaboration in co-authorship in the fields of HSS and discusses the structural dissonance in the paradoxical processes of knowledge creation and production in HSS. The paper ends with a macro level consideration of contemporary geopolitics, its perils and impacts on transnational academic collaboration in the shadow of the Covid-19 pandemic as an aperçu.

The macro context: politics of scientometrics and the datafication process

The number of internationally co-authored papers has continued to increase exponentially over the last two decades. According to [Marginson \(2021a; 2021b\)](#), the growth in total bibliometric output loosely correlates to the expanded national system capacity, as shown in government funding of research in universities and public institutes, and the researcher workforce, from which we can apprehend the contemporary globalization of transnational academic collaboration.

At the same time, we can also surmise that this trend indicates the intensification of international comparisons and competitions driven by the datafication process. Datafication is the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis; it is about taking previously invisible process/activity and turning it into data, which then can be monitored, tracked, analyzed and optimized ([Mayer-Schoenberger and Cukier, 2013](#)). [Sadowski \(2019\)](#) illustrates the perpetual cycle of data accumulation that corresponds to flows of power and profit relying upon a world in which everything is made of data: “the alchemy of datafication promises to produce infinite reserves of both. At the same time, the rhetoric of universality reframes everything as within the domain of surveillance/platform/digital capitalism.” (p. 3) Datafication is also integrated into the university systems of academic evaluation in many countries around the world.

Networked global science has been expanding at such speed in the last two decades that global research collaboration has been considered an imperative of knowledge innovation. [Kwiek \(2020\)](#)’s research on international research collaboration (IRC) in Europe during the period 2009–18 in terms of co-authorship and citation distribution of globally indexed publications reveals that the growth of European science has been driven solely by internationally co-authored papers. Ironically, however, the existing metrics of global science is still focusing on inter-national comparisons and competitions. The value of academic research outputs is measured to rank one country against another, pointing to the “arms race in innovation” ([Marginson, 2020](#)) and the “sportification of science” ([Kaldewey, 2018](#)).

For instance, between 2010 and 2019, China has seen 62.5% increase of internationally co-authored publications, which accounted for 14.4% of its total publications in 2010 and 23.4% in 2019. Since 2018, China’s research output has grown by 13.2%, overtaking the US as the top country by publication output in the world ([Universities UK, 2020](#)). China has the world’s largest market for academic publishing, partly because many professions require publishing research papers as a job performance indicator ([Lew, 2021](#)). Fang Shimin, a US-based scholar and commentator who has been exposing pseudoscience and fraud for 20 years, confirms that there has been no fundamental change in China regarding research ethics. In 2020, Dutch microbiologist and science integrity advocate Elisabeth Bik and other experts found more than 400 published scientific papers with potentially fabricated images that were suspected to have been produced by one paper mill in China ([Lew, 2021](#)).

In terms of the proportion of co-authored publications in the overall research outputs, France (58%) was leading, followed by Australia (57.5%) and the UK (57.2%) in 2019 ([Universities UK, 2020](#), p. 24). The UK saw the increase of international collaboration compared from 25.7% in 1999 to 57.2% of the overall research outputs in 2019 to rank the third largest producer of internationally co-authored publications in the world after China and the United States. However, given the size of its population, the UK’s internationally co-authored research publications ratio per academic is the highest in the world. Furthermore, the UK’s field-weighted citation impact has ranked first every year since 2007 – which is a testament to the global impact of its research output ([Universities UK, 2020](#), p. 23).

Scientometricians have devised a multitude of metrics to help in these rankings. Such inter-national comparisons and competitions have a trickle-down effect on the intra-national systems of academic evaluation in many countries, to allocate public funds between universities or measure individual academics’ research outputs.

In the rise of scientometrics (as integral part of datafication in the emerging digital capitalism), there is increasing interdependency between the higher education sector and the academic journal publishing industry, whose oligopolistic system has been increasingly criticized. Although non-profit publishing houses once dominated the academic publishing industry, now about 50% of global academic journals are part of commercial for-profit publishing houses ([Gaille, 2018](#)). [Larivière et al. \(2015\)](#) looked at all scientific articles published in the Web of Science database between 1973 and 2013, and found that five companies have published more than half of them since 2006: Reed-Elsevier, Taylor & Francis, Wiley-Blackwell, Springer and Sage. Almost 70% of journal articles published in psychology and social sciences are owned by these big publishers.

Academic publishing is an extremely lucrative business with total global revenues of more than £19bn despite the narrow audience ([Buranyi, 2017](#)). In 2010 Elsevier scientific publications had 36% profit margins – higher than Apple, Google, or Amazon posted that year. The profit margin has been increasing each year and by 2017 it was 40% ([Buranyi, 2017](#)).

Publishers get copyrights of the work produced by academics for free. They do not even pay for quality control as it is done by academics themselves for free in the form of peer review. Once the academic papers are published, publishers sell them back to the academic communities at a monopolistic price set by the publishers, with subscription contracts, to be read by academics who, in a collective sense, created the product in the first place.

For instance, Harvard is currently paying \$3.75 million each year to maintain their current journal subscriptions. This cost represents 10% of the total cost of everything that the library acquires. Some journals cost the university up to \$40,000 every year, with the two top publishers increasing the price of content 145% over the last six years (Gaille, 2018). A memo from Harvard's Faculty Advisory Council to its Library in 2012 states that it "reached this conclusion: major periodical subscriptions, especially to electronic journals published by historically key providers, cannot be sustained: continuing these subscriptions on their current footing is financially untenable" (Wagstaff, 2012).

Publishers also control bibliometrics that are used to evaluate research impacts which affect academic career advancement. This oligopolistic system is prevalent worldwide, with which all research institutions are compliant. A logical question raised then out of this situation is: why and what exactly are we paying these big publishers for?

The exploitations of academic work – not only knowledge production *per se* but quality control by the collaborative work of editorial board and peer reviews for free – have become a universal system applied worldwide. Peer reviews play an essential role in ensuring the quality and validity of academic papers. Publishing papers in academic journals is a main way of achieving professional recognition, although scholars in the fields of HSS would consider publishing their most significant work in book form rather than in journal articles. Nevertheless, journals still play an important role for HSS as a lot of articles are developed in the context of a book in progress; thus, peer reviewers for HSS journals play an important role in moving a potential book to eventual publication.

Publications in top-tier journals are often more valued than publications in lower ranked ones. Given the convention of traditional journal publishing, switching to a different model may be considered difficult. It is expected that a major publisher acquiring a journal will have the effect of increasing the journal's visibility and impacts. However, as Larivière et al. (2015)'s research indicates, there has been no clear increase in terms of citations after a journal ownership changes from a small company to a larger major publisher. As Harzing and Mijnhardt (2014) suggest, citation-based performance metrics can be more democratic as their "verdict" is based on the reception of the paper by the academic community as a whole and self-promotion of individual academics, whereas acceptance in a high-impact journal is dependent on only a handful of "gatekeepers" (i.e. the editor and reviewers).

There have been academic protests to recover the control and ownership of academic research outputs in order to make them available and useable to anyone and to any institution for free for the sake of the advancement of knowledge. For instance, "The Cost of Knowledge" (<http://thecostofknowledge.com>) protest initiated by a group of prominent academics points out that the cost of journal publishing has gone down because the cost of typesetting has been shifted from publishers to authors and the cost of publishing and distribution is significantly lower than it used to be. Then why do academics contribute all this volunteer labor, and their employers pay all this money, for a service whose value no longer justifies its cost?

Multiple groups of academics and organisations (as listed in http://openscience.ens.fr/ABOUT_OPEN_ACCESS/ARTICLES/) have insisted for more than a decade on developing and promoting new open access publishing platforms, called "intellectual commons", publicly owned and freely used by academics, companies and citizens (Eisen, 2003; Farge, 2017). They argue that academic papers should be available for free in open access platforms and their content should be useable to anyone and to any institution. Ideas are not of the same nature as material objects. When ideas are shared, they are not lost. Ideas are only fruitful if they are communicated for discussion, verification, improvement, and education. Ideas are not merchandise but intellectual commons. Accordingly, they suggest that journals should be owned by their editorial board while editors and referees continue to work for free; and the publishing companies could continue their business only "as service providers" to the publishing platforms (Farge, 2014, 2017).

Such open access publishing platforms in fact already exist, albeit not in the main stream yet – e.g. "scielo.org" in Brazil, "revue.s.org" in France, and "PLOS ONE" in the United States – a peer-reviewed open access scientific journal published by the Public Library of Science (PLOS) under Creative Commons licenses (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>) since 2006. More recently, in Germany, the so-called "Alliance Initiative" (a task force of all German research institutions) has been assigned with negotiating collective, nationwide open access agreements with the three largest commercial publishers of scholarly journals, namely Elsevier, Springer Nature, and Wiley, on behalf of all German academic institutions, including universities, research institutes, and their libraries (Haucap et al., 2021).

In the field of scientometrics, there are also crucial contested issues about co-authorship and types of contribution (theoretical, empirical, methodological) and the context of the citation (praising or criticizing, central or peripheral) matters (Nosek, 2012). For instance, h-index (an author-level metric that measures both the productivity and citation impact of the publications of a scientist or scholar) cannot be used to compare academics that work in different disciplines or are at different career stages. Harzing et al. (2014) devised the hI_{annual} (or hI_a for short) that represents the average annual increase in the individual h-index, and showed that the hI_a index attenuates h-index differences that are purely attributable to (disciplinary) co-authorship practices and career lengths.

Overall, knowledge is the *raison d'être* and *sine quo non* of academic work. However, as illustrated above, there is an increasing separation of knowledge production value (academic capital) from the producer (academics) in the rent-seeking economy of academia, concomitantly subsumed in the datafication process by the oligopolistic academic publishing industry.

The notion of academic alienation is far more than the issue of copyright ownership. Academic alienation is a structural problem caused by the current academic system, which exacerbates the paradox embedded in academic collaboration in HSS. I argued that the very notion of academic collaboration in the form of co-authorship in HSS may be an oxymoron in nature; hence it requires a closer investigation in line with a better understanding of the micro contexts of knowledge production and innovation.

Micro context: organisational culture and social psychology of transnational academic collaboration

The contemporary academic management structure is heavily reliant on the quantification and datafication of academic work (especially research, publication, and impact). The research evaluation systems established in many countries (e.g. Australia, Belgium, France, Italy, Korea, New Zealand, the UK, etc.) are purported to increase academic performance and productivity and efficiency – according to the principles of New Public Management (NPM), with its performance contracts and performance-based institutional funding (Sandström and Besselaar, 2018).

However, there is an acknowledged decline in the conditions of academic work, and enterprising academic subjects live an existence of calculation for performativity (Ball, 2012; Winn, 2015). Ball (2012:17) narrates his own lived experience of changes in his academic subjectivity, having succumbed to the contemporary “market-framed university” imperative (Cowen, 2000) to make himself “calculable rather than memorable”.

Overall, the university is organized around the neoliberal market principles (both real and imagined) that are rationalized and enter into the deep structure (*doxa*) – the agreement between all participants in the field (Bourdieu, 1984; Kim, 2017: 995) – to achieve “the universalization of market-based social relations, with the corresponding penetration in almost every single aspect of our lives” (Shamir, 2008: 3; Ball, 2012: 18). Performative accountability regimes have had impact at the macro-level on policies and systems, and at the meso- and micro-levels of basic units and individuals (Scott, 2006; Xu et al., 2021). This speaks to the condition of “governmentality” in Foucault’s term, the way that individuals come to govern themselves (McKinlay et al., 2012), and academic work is routinely monitored and evaluated accordingly. As Cowen (2000) forewarned more than two decades ago, smaller and smaller pieces of information are being collected, more and more surveillance are being devised to intensify the “factorization of performance”, by which academic research outputs are evaluated on a piece-rate basis (Cowen, 2000: 102).

Overall, the highly skewed processes of academic knowledge production and extremely competitive up-or-out nature of academic career have become more tied to neoliberal performativity regimes as *doxa*, which discourage long-term research without funding, while encouraging more tactics and short-term strategies to perform and play by the rules of the game to increase research outputs and impacts. The current sociology of academic performativity mandate speaks to micromanaging organization, measuring everything for quantification and datafication.

The contemporary notion of metric fixation is entwined with the mercantilisation of knowledge. Lyotard (1979/1984) forewarned four decades ago that knowledge ceases to be an end in itself and the goal of knowledge production is to be “exchange”. However, the more exchange value the academic creates, the more her/his own value decreases as the “self-exploiting entrepreneur” who is forced to sell herself ‘piecemeal’ in order to remain employable (Hall, 2018: 2).

Rent-seeking activities in neoliberal academia are differentiated from “value-producing” labor or “profit-making” entrepreneurialism (Reitz, 2017; Welsh, 2020). Rent-seeking activities in universities are directly related to performativity, which operates by system optimization, i.e. playing the calculation of input and outputs by the rules of the game – e.g. the Research Excellence Framework (REF), research grants and various league tables with strings attached. In some countries (e.g. China, South Africa, South Korea), there are direct incentives for international publications (Cowen, 2020; MacFarlane, 2017; Xu et al., 2021). The publication incentive mechanism in China, for instance, is symptomatic of performativity regimes, in which Chinese academics in HSS proactively align their academic pursuits with national interests (Zha and Shen, 2018), which are underpinned by publishing incentives (Xu et al., 2021: 5).

For Lyotard (1979/1984), “performativity” is defined by what constitutes knowledge, what knowledge is of worth, and whose knowledge has legitimacy (Clapham, 2013), which is increasingly conditioned by scientometrics – an integral part of the datafication process in the “computerization of society” (Lyotard, 1984: 7). Overall, the proliferation of rent-seeking behaviors in academic life indicates the systemic change of organizational culture inside academia that is compliant with the national regulatory regimes as well as the transnational tacit rules of competition, which altogether reflect the transformed grounds for the legitimacy of knowledge.

Against the backdrop, transnational academic collaboration in the form of co-authorship has become an important currency of academia, vital for career progression, funding, and success in research assessments. Many academics go into transnational academic collaboration for co-authorship with a mixture of ambition and opportunism for an “international reputation” with reference to the ranking system used in the research assessments. Collaboration in the form of co-authorship is regarded as a solution for getting more papers published, more quickly. This trend is particularly strong for “early-career researchers” trying to launch their academic careers (Taylor and Francis, 2017).

According to Taylor and Francis (2017: 6)’s white paper, “Co-authorship in the Humanities and Social Sciences: A Global View”, a typical paper published in HSS is now more likely than not to have multiple authors: two or three authors per paper in HSS (61%); single author (27%); four or five (11%); more than six (1%). The survey also confirms that the notable growth in co-authorship in HSS contemporaneously is attributed to increasing competition and the performance-based pressures in academic life. The next most highly scoring motivation was the growth of opportunities to collaborate internationally and the growth of multi-disciplinary work. Scholars are benefitting from research networks sharing multi-disciplinary perspectives and funding to support collaboration (Taylor and Francis, 2017: 6–8).

Overall, the trend of co-authorship in academic publications has led to a sharp decrease in the proportion of single authored publications in HSS. According to Ossenblok et al. (2014)’s data analysis using the bibliographic information between 2000 and 2010, in social sciences, majority of publications (four-fifth of all publications) were co-authored, whereas in the humanities, more than one-third were co-authored publications.

As research collaborations proliferate, authorship credit has increasingly become a complicated and often contested issue in conflict. The contemporary obsession with measuring everything (including “impacts”) to quantify knowledge and the influence it has on us (datafication) itself have further complicated the terms and conditions of transnational academic collaboration. Especially the issues around co-authorship are the most common concerns in publishing ethics under scrutiny (MacFarlane, 2017).

Ethical implications for transnational academic collaboration in co-authorship

The commonly occurring problems in academic co-authorship are related to authorship credit – i.e. if the order of named author reflects the level of contribution fairly. In what order should the authors’ names appear on a published paper? According to Taylor and Francis (2017), HSS researchers do not have a settled view on how to come to an agreement, and yet the majority of HSS researchers have not received enough guidance on how to do collaborative writing.

There is a general assumption or conventional belief that senior academics are often over-credited. Taylor and Francis (2017) reports that a researcher in HSS complained about “supervisors insisting their name goes on as a co-author when they have basically done their job, nothing more” (Taylor and Francis, 2017: 12). However, there exist the opposite cases that junior researchers are over-credited in comparison to the other authors (Taylor and Francis, 2017); and yet such cases are often less scrutinized or even tend to be condoned.

For example, a senior academic in HSS has encountered a calculous behavior of an early-career researcher in the process of co-authoring a paper. Her episode is narrated here, with her permission, as a case study:

A senior academic (S) invited a doctoral student (J) who is close to the completion of her Ph.D. based in another country to write a joint paper for a Special Issue. The joint paper was built on S’s ideas and theoretical framework. S already shared her 2500-word draft conceptualization and outline of the paper when inviting J to collaborate. However, even before making any contribution, more or less immediately after agreeing to join the writing project, J insisted that she needs to be the first author. Her given reason is that in her country, only the first author counts; hence she wants to be the first author for the sake of her career progression. Further to that, J also assured S that her contribution to the joint paper would be equal. S then hesitantly agreed, under the condition that J would make substantial contribution in writing the joint paper. However, J’s contribution turned out to be a minimum.

To produce the joint paper, S did the vast majority of the conceptualization, research and writing. Eventually when S suggested changing the authorship credits to reflect this, J accused S of abusing her power. From the perspective of those not well informed, the situation may be easily portrayed in the public domain as a senior academic exercising her power over a junior researcher. However, S’s experience was that J took advantage, using her position of structural vulnerability (as an early-career researcher) to her advantage in the authorship dispute – instead of delivering her expected contribution to the joint paper professionally.

In the authorship dispute, the senior academic (S) explains the situation:

“In my view, authorship should be determined by the level of contribution as simple as that. It is not about seniority but intellectual honesty ... J keeps saying she wrote the ‘theoretical’ part. But it was not. She wrote just one section in the literature review within the conceptual framework that I had outlined in the paper. Furthermore, J’s draft literature review section didn’t meet the publishable standards as it was mostly descriptive, and thus I had to rewrite it. Her allegation that I merely paraphrased what she drafted is totally untrue. I rewrote the section substantially, drawing on more relevant literature from my own fresh research as well as from my book in line with my argument and conceptual framework that I sketched in the Introduction. ...

... All in all, I have never experienced this level of aggression and grotesque ambition from any of my colleagues – let alone any of early-career researchers – in my academic career so far over the last few decades. I have helped J whenever I could on various occasions whenever she requested – e.g. reading and providing feedback on her other draft papers and her research proposal. I also agreed to be one of her referees in her research grant application when she requested ...

... However, when it comes to intellectual and academic work to produce a joint paper for publication, it is important to be honest and transparent to acknowledge the level of contributions made by each of us correctly, regardless of the ranks and seniority of the contributors.”

According to the Research Integrity policy and guidelines from UCL (<https://www.ucl.ac.uk/research/integrity/policies-and-guidelines/publication-and-authorship>):

- Assigning appropriate authorship is fundamental to guarantee professional conducts and ethical practices when publishing academic and research papers.
- Major contributors sign first.
- Alphabetical order is applied only when authors have contributed equally.
- It is also worth noting that the balance of contribution can alter over time, and so if this occurs researchers may find it beneficial to revisit authorship prior to publication.

- Given the variances and the importance of accuracy not only for the publication but in recognizing contributions, it is a matter that should be openly addressed.
- Incorrect authorship goes beyond raising disputes amongst research teams to matters of research integrity and good research conduct.

The RCUK Policy and Guidelines on the Governance of Good Research Conduct (<https://www.ukri.org/about-us/policies-and-standards/research-integrity/>) states that unacceptable research conduct includes 'Misrepresentation of involvement, such as inappropriate claims to authorship and/or attribution of work where there has been no significant contribution, or the denial of authorship where an author has made a significant contribution'."

Following the authorship dispute, however, J took advantage of her invited position as one of the guest editors of a Special Issue (SI) which involves S and two other senior academics. Whilst not making proper contribution to the joint paper led by S, J was writing a solo paper without telling them. Once she finished drafting her paper, she asked the other guest SI editors (including S) to provide feedback on it. Once the paper was revised for submission, J insisted that her single-authored paper should be published by the same journal first ahead of the SI and that the authors of SI papers should cite her paper. J made that happen.

Her solo paper was published extremely fast within less than three months of submission. Without consulting the other guest editors of the SI, J succeeded in getting some of the authors of the SI papers cite her forthcoming paper through placing this a condition of revision requirements for publication in the SI. In that way J could achieve her ambition to increase citations of her new publication instantly.

When two or more academics from very different national cultural backgrounds (including race/ethnicity, language, and religion) decide to collaborate, they are likely to rely on some common grounds and networks that help their collaboration – e.g. having doctorates from the same university and/or working in the same field of studies, which speaks to the first episode about the co-authorship. It is, therefore, important to understand how the co-authorship in dispute was intersected (and compounded) by the heterogeneous personal attributes and behaviors of S and J, and how structural forces (the rules of the game) within academic publishing and promotion toxify the collaboration.

Another possible scenario is presented here: when academics in two or more countries decide to work in collaboration on a research/writing project, they may actually come from the same country of origin and the same ethnic cultural backgrounds especially if their academic collaboration is derived from their ethnonational diasporic networks and the research project relies on their social and cultural capital – e.g. overseas Chinese academics collaborating with Chinese academics in mainland China to write a joint English paper on China-related topics for international publication. Such a case is narrated in the second episode:

For instance, when a UK-based East Asian academic (C) joined her alumni network-based group research project, she realized that the research collaboration was far from intellectual endeavors. The research group consists of 10 female academics who graduated from the same university in East Asia. The project leader, a US-based alumna (Y) determined the nature of collaboration in the group research project. Y is senior as eight years older than C, but in terms of academic rank, Y is Associate Professor and C is full Professor. C was surprised to find that Y imposed a hierarchical order in research collaboration. The group communication was almost like militaristic top-down order rather than open and equal among co-investigators, even though the project team was made up of female alumni and the research was on gender issues, concerning women with doctorates and their professional career progression reflecting the specific East Asian cultural contexts.

C expected multidisciplinary approaches for intellectual discussions and wanted to bring interdisciplinary interpretations and comparative perspectives in the process of collaborative research. However, none of C's comments and suggestions were welcomed, nor accepted by Y in drafting a joint research paper.

C was equally shocked by the attitudes and roles of the other co-investigators who are based in their native country in East Asia. Regardless of their advanced academic career and research experience, they all seem to be familiar with and complacent about Y's top-down order in research collaboration, limiting their roles as mere functionary in the process of data collection and coding. The deference to authority by doing what they are told to do in the hierarchical structure seems to be what these East Asian academics based in their East Asian home country consider as good 'collaboration'.

The project leader's dictatorial leadership did not end there. Y decided the name order of co-authors without communicating with the project team members. When the co-authors' bio notes were collated and shared among the project team members to get ready for submission, C discovered her author bio note was drastically shortened, even though her original bio note was written within the word limit. Y (the project leader and the first author) edited it, without consulting C, to omit the information about key achievements from C's bio note.

Since only C and Y were the overseas based academics in the project team, C felt Y's sense of competition and insecurity in relation to her age/academic rank, and more importantly, the quality of research paper led by Y's autocratic leadership was nowhere near the international standards of excellence. In the end, Y also admitted that the paper was not good enough to be submitted to a high ranked SSCI journal - despite the original plan. Publication in high-ranked SSCI journals was, in fact, the main goal and the main reason of this transnational diasporic academic collaboration (especially for those based in the East Asia country) as there are strong incentives for international publications in their academic evaluation system. Such a trend is longstanding and intensifying. Even 25 years ago when C was offered a university faculty position in her home country in East Asia, a condition of the job offer (albeit informally suggested) was to write joint papers with a male colleague in the Department on his subject area (Philosophy) in English as he needed international publications.

All in all, these episodes narrated so far as case studies speak to the fact that relationships between and among academics in transnational collaboration are shaped by different assumptions they bring in, which segue into the nature of collaboration. Such assumptions are both directly and indirectly influenced by embodied knowledge (not only epistemic, disciplinary knowledge capital but also social, cultural, symbolic and identity capital) in the increasingly stratified academic systems in both global and national contexts (Kim, 2017).

Discussion and conclusion

This paper has illustrated that transnational academic collaboration in the form of co-authorship in HSS is shaped in perverse ways by the existing structures of the fields to highlight the oxymoronic nature of collaboration and the paradoxical processes of knowledge creation and production in HSS. The process of knowledge creation in HSS is more often than not based on a heroic model of producing narrative knowledge with explanatory power (internalities), while the funds of knowledge is dedicated to optimization of performativity (externalities). As warned by Latour and Callon (1997), “externalities would constantly perturbate, put at risk, internalities. Calculation would then become infinite; private appropriation would turn out to be impossible; profit would always be the object of endless disputes.” (Latour and Callon, 1997: 7). It is important to recognize a radical differentiation (incommensurability) between narrative knowledge in HSS produced for our understanding (*Verstehen*) and its storage and capitalization for scientification (*Verwissenschaftlichung*) and datafication. Such dissonance between ‘use-value’ and ‘exchange-value’ of knowledge has become more pronounced than ever in the current system of data-manufacturing, digital capitalism (Sadowski, 2019), which exposes the omnipresence of calculation underlying the exponential increase of transnational academic collaboration in the form of co-authorship.

As forewarned by Lyotard (1979; 1984) more than four decades ago,

Knowledge is and will be produced in order to be sold, it is and will be consumed in order to be valorised in a new production: in both cases, the goal is “exchange”. Knowledge ceases to be an end in itself, it loses its “use-value”. (Lyotard, 1984: 4–5).

In fact, however, not just in HSS but also in many other areas of scientific research, ideas are the currency for innovation. Popper (1934) empathized the importance of ideas in the process of inventing a theory (the context of “discovery”), which is not logical but creative – e.g. Freud’s theory of the Ego, the Super-ego, and the Id, taken as an example by Popper, originated from Greek myths for which no substantially stronger claim to scientific status can be made than for Homer’s collected stories from Olympus. However, Popper illuminates,

Historically speaking almost all scientific theories originate from myths ... a myth may contain important anticipations of scientific theories ... If a theory is non-scientific or metaphysical, it is not thereby found to be unimportant or nonsensical ... Newton’s theory of gravity, and especially the lunar theory of the tides, was historically speaking an offspring of astrological lore ... But to get feedback from Nature, the metaphysical frameworks must ultimately provide scientific – falsifiable – theories (Popper, 1963: 50).

As one of the interviewees in my research on the process of new knowledge creation in HSS expressed, “I am an artist. Creativity is needed – not only in terms of the original idea but also in the implementation of the solution – to see the whole picture in writing.” Latour and Callon (1997) assert that the categorical imperative “you should not calculate!” is applicable to all forms of mobilization of ideas, goods and people, no matter how paradoxical this might seem at first. I endorse their views, believing that innovation, especially in the fields of HSS, stem from individual artistic dispositions enabling creative imagination communicated in writing. Accordingly, when it comes to the business of co-authorship, I argue, it is difficult to claim true collaboration in absolute terms.

However, to understand the infrastructure and process of scientification and datafication of knowledge and its impacts on transnational academic collaboration, it is useful to comprehend Latour (1987)’s Actor-Network Theory (ANT), which suggests scientific knowledge production is built upon a “cycle of accumulation” of resources - through a systematic mobilization and circulatory movements of human and nonhuman resources to and between different “centers of calculation” (Jöns, 2011). For Latour, network can be understood as a distributed sociocultural cognitive system (a mind machine), and application of science to the world is a matter of extending the network from the centers to the peripheries. One of the ways of extending networks is through “metrology” – the creation of standards.

From this notion of “metrology”, the current system of scientometrics is inferred as a gigantic enterprise that helps to extend metrological chains globally. Once “a cycle of accumulation” of “immutable mobiles” (such as writings, inscriptions, documents and illustrations) is established, cascading of resources accumulated in the “centers of calculation” starts, which makes dominance at a distance feasible and generates new cognitive power and the production of new scientific knowledge. For Latour, “theory” refers to the order of inscriptions that allow the “centers of calculation” to mobilize, manipulate, combine, rewrite and tie together all the traces obtained through the ever-extending networks (Latour, 1987: 241–242). The higher order translates the lower order, and in that way, it increases the mobility of the lower order inscriptions across a wider terrain, across multiple domains. Overall, ANT

analyses the infrastructure of actor-network, albeit without attending to power imbalances within the actor-network. Latour (1999)'s ANT concentrates attention on a movement – a *circulating* entity: one of the very phenomena of the social order not being made of agency and structure.

In the ongoing global pandemic, the infiltration of geopolitics in transnational academic collaboration has been exposed. Most recently, at the US GOP Medical Witnesses: COVID-19 meeting, held on June 20, 2021, Richard Muller, Professor Emeritus of physics at the University of California, Berkeley, who has worked on scientific efforts that have won Nobel Prizes, testified that virus, which came out of China, carried with it genetic information about its origins. Muller also shared an anecdote that occurred with a colleague of his — a story he says is “as horrifying and more frightening than almost anything else in my life.” In the early days of the pandemic, he called on an expert virologist friend to help him review literature suggesting there may have been a lab leak. The friend said no, so he asked if someone in his laboratory could do it. But the answer was no again. Muller pressed him on the refusal, to which he responded:

If anyone in my laboratory is discovered to be working on a laboratory leak hypothesis, China will label us enemies of China and the laboratory will be blacklisted and we will no longer be able to collaborate. We collaborate all the time with China. Nobody will take that risk.

The idea that China has managed to interfere, to break United States' freedom of expression, freedom of investigation, freedom of thought through this collaboration is really scary. Mercola (2021).

Overall, academic knowledge production is an integral part of diffused network-based career building process – both epistemic and institutional networks, in which scholarship is increasingly a means for business in the rent-seeking economy rather than a vocation (invoking Weber). In that regard, love of knowledge, passion for pursuing truth and courage to speak truth to power would no longer be *a priori* reasons in pursuit of scholarship. As evidenced in the testimony quoted above, transnational academic collaboration in times of crises has many implications – both geopolitical and epistemic – for the *longue durée* as witnessed in history. The fields of humanities and social sciences will have to bear witness continually.

References

- Ball, S., 2012. Performativity, commodification and commitment: an I-spy guide to the neoliberal university. *Br. J. Educ. Stud.* 60 (1), 17–28. <https://doi.org/10.1080/00071005.2011.650940>.
- Bourdieu, P., 1984. *Distinction: A Social Critique of the Judgement of Taste* (Richard Nice, Trans. and Tony Bennett, Intro.). Routledge, London and New York.
- Buranyi, S., 2017. Is the staggeringly profitable business of scientific publishing bad for science? *Guardian*. <https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>.
- Clapham, A., 2013. Performativity, fabrication and trust: exploring computer-mediated moderation. *Ethnogr. Educ.* 8 (3), 371–387. <https://doi.org/10.1080/17457823.2013.792676>.
- Cowen, R., 2020. Changing comparative education: interpreting foreigners, routinising research, and writing history. *Comp. Educ.* 56 (1), 127–141. <https://doi.org/10.1080/03050068.2020.1714332>.
- Cowen, R., 2000. The market-framed university: the new ethics of the game. In: Cairns, J., Gardener, R., Lawton, D. (Eds.), *Values and the Curriculum*. Routledge, London and New York.
- Eisen, M., 2003. Publish and be praised. *Guardian*. <https://www.theguardian.com/education/2003/oct/09/research.highereducation>.
- Farge, M., 2014. The future of scientific publishing. *Open World Forum*.
- Farge, M., 2017. Scholarly Publishing and Peer-Reviewing in Open Access in European Commission. Europe's Future: Open Science, Open Innovation, and Open to the World, Reflections of the Research, Innovation and Science Policy Experts (RISE) High Level Group, Publications Office of the European Union, Chapter 2; 2.4 73–82. <https://www.rd-alliance.org/sites/default/files/attachment/ki0217113enn.pdf>.
- Gaille, B., 2018. 17 Academic Publishing Industry Statistics and Trends. <https://brandongaille.com/17-academic-publishing-industry-statistics-and-trends/>.
- Hall, R., 2018. *The Alienated Academic: The Struggle for Autonomy inside the University*. Palgrave Macmillan, Gewerbestrasse, Switzerland.
- Harzing, A., Mijndhardt, W., 2014. Proof over promise: moving citation metric systems beyond journal impact towards a career impact approach. *LSE Blog*, 17 September 2014. <https://blogs.lse.ac.uk/impactofsocialsciences/2014/09/17/proof-over-promise-career-impact-approach-harzing/>.
- Harzing, A., Alakangas, S., Adams, D., 2014. hla: An Individual Annual H-index to Accommodate Disciplinary and Career Length Differences.
- Haucap, J., Moshgbar, N., Schmal, W.-B., 2021. The Impact of the German “DEAL” on Competition in the Academic Publishing Market. Heinrich-Heine-University Düsseldorf, Düsseldorf. March, 2021. http://openscience.ens.fr/ABOUT_OPEN_ACCESS/ARTICLES/2021_03_15_University_of_Dusseldorf_on_the_negative_impact_of_DEAL_in_Germany.pdf.
- Hoekman, J., Frenken, K., Oort, F., 2009. The geography of collaborative knowledge production in Europe. *Ann. Reg. Sci.* 43, 721–738. <https://doi.org/10.1007/s00168-008-0252-9>.
- Jobs.ac.uk Survey, 2012. Making Academic Collaboration Work. <https://www.jobs.ac.uk/media/pdf/careers/resources/making-academic-collaboration-work.pdf>.
- Jöns, H., 2011. Centre of calculation. In: Agnew, J.A., Livingstone, D.N. (Eds.), *The SAGE Handbook of Geographical Knowledge*. Sage, pp. 158–170.
- Kaldewey, D., 2018. The grand challenges discourse: transforming identity work in science and science policy. *Minerva* 56, 161–182. <https://doi.org/10.1007/s11024-017-9332-2>.
- Kim, T., 2017. Academic mobility, transnational identity capital, and stratification under conditions of academic capitalism. *High Educ.* 73 (6), 981–997.
- Kwiek, M., 2020. What large-scale publication and citation data tell us about international research collaboration in Europe: changing national patterns in global contexts. *Stud. High Educ.* <https://doi.org/10.1080/03075079.2020.1749254>.
- Larivière, V., Haustein, S., Mongeon, P., 2015. The oligopoly of academic publishers in the digital era. *PLoS One* 10 (6), e0127502. <https://doi.org/10.1371/journal.pone.0127502>.
- Latour, B., 1999. On recalling ANT. *Socio. Rev.* 47 (S1), 15–25. <https://doi.org/10.1111/j.1467-954X.1999.tb03480.x>.

- Latour, B., Callon, M., 1997. "Thou shall not calculate!" or How to Symmetrize Gift and Capital [« Tu ne calculeras pas ! »- ou comment symétriser le don et le capital]. La revue du MAUSS 9, 45–70. Comment peut-on être anticapitaliste?, La Découverte, edited by Alain Caillé. <http://www.bruno-latour.fr/sites/default/files/downloads/P-71%20CAPITALISME-MAUSS-GB.pdf>.
- Latour, B., 1987. *Science in Action*. Harvard University Press, Cambridge, Mass.
- Lew, L., 2021. Behind China's "mind power" egg scandal, where academic ghost writers quote US\$400 to get bogus science published. South China Morning Post, 15 May 2021. https://www.scmp.com/news/china/politics/article/3133548/behind-chinas-mind-power-egg-scandal-where-academic?module=lead_hero_story_1&pgtype=homepage.
- Lewis, J., Ross, S., Holden, T., 2012. The how and why of academic collaboration: disciplinary differences and policy implications. High Educ. 64, 693–708. <https://doi.org/10.1007/s10734-012-9521-8>.
- Liotard, J.-F., 1979/1984. *The Postmodern Condition: A Report on Knowledge*, Trans. Geoff Bennington and Brian Massumi. Manchester University Press, Manchester. Trans. of La Condition postmoderne: rapport sur le savoir (Paris: Minuit, 1979).
- MacFarlane, B., 2017. The ethics of multiple authorship: power, performativity and the gift economy. Stud. High Educ. 42 (7), 1194–1210. <https://doi.org/10.1080/03075079.2015.1085009>.
- Marginson, S., 2020. Knowledge Network or Hegemonic Hierarchy Dynamism and Power in Global Science, 27 October 2020. <https://www.researchcghe.org/events/cghe-seminar/knowledge-network-or-hegemonic-hierarchy-dynamism-and-power-in-global-science/>.
- Marginson, S., 2021a. Heterogeneous Systems and Common Objects: The Relation between Global and National Science. CGHE Special Report. April 2021. <https://www.researchcghe.org/perch/resources/publications/cghe-special-report-april-2021.pdf>.
- Marginson, S., 2021b. Knowledge Network or Hegemonic Hierarchy Dynamism and Power in Global Science. <https://www.researchcghe.org/perch/resources/events/cghe-webinar-27-oct-2020.pdf>.
- Mayer-Schoenberger, V., Cukier, K., 2013. *Big Data. A Revolution that Will Transform How We Live, Work, and Think*. John Murray Publishers, London.
- McKinlay, A., Carter, C., Pezet, E., 2012. Governmentality, power and organization. Manag. Organ. Hist. 7 (1), 3–15. <https://doi.org/10.1177/1744935911429414>.
- Mercola, J., 2021. Expert Scientists Testify on Virus Origin. <https://articles.mercola.com/sites/articles/archive/2021/07/17/expert-scientists-testify-on-virus-origin.aspx>.
- Nosek, B., 2012. Publish or Perish? Grade Yourself and Persist. Association for Psychological Science (APS), 28 March 2012: <https://www.psychologicalscience.org/observer/publish-or-perish>.
- Ossenblok, T., Verleysen, F., Engels, T., 2014. Coauthorship of journal articles and book chapters in the social sciences and humanities (2000–2010). J. Assoc. Inf. Sci. Technol. 65 (5), 882–897. <https://doi.org/10.1002/asi.23015>.
- Popper, K., 1934. *The Logic of Scientific Discovery*.
- Popper, K., 1963. *Conjectures and Refutations*. Routledge and Keagan Paul, London.
- Reitz, T., 2017. Academic hierarchies in neo-feudal capitalism: how status competition processes trust and facilitates the appropriation of knowledge. High Educ. 73 (6), 871–886.
- Sadowski, J., 2019. When data is capital: datafication, accumulation, and extraction. Big Data Soc. 1–12. <https://doi.org/10.1177/2053951718820549>. January–June 2019.
- Sandström, U., Besselaar, P., 2018. Funding, evaluation, and the performance of national research systems. J. Informetr. 12 (1), 365–384.
- Shamir, R., 2008. The age of responsibilization: on market-embedded morality. Econ. Soc. 37 (1), 1–19.
- Scott, P., 2006. The academic profession in a knowledge society. In: Teichler, U. (Ed.), *The Formative Years of Scholars*. Portland Press, London, pp. 19–30. https://portlandpress.com/pages/volume_83_the_formative_years_of_scholars.
- Taylor, Francis, 2017. Co-authorship in the Humanities and Social Sciences A Global View: A White Paper from Taylor & Francis. <https://authorservices.taylorandfrancis.com/wp-content/uploads/2017/09/Coauthorship-white-paper.pdf>.
- Universities UK, 2020. International Facts and Figures 2020. <https://www.universitiesuk.ac.uk/universities-uk-international/insights-and-publications/uuki-publications/international-facts-and-figures-2020>.
- Wagstaff, K., 2012. If harvard can't afford academic journal subscriptions, maybe it's time for an open access model. Time, 26 April 2012. <https://techland.time.com/2012/04/26/if-harvard-cant-afford-academic-journal-subscriptions-maybe-its-time-for-an-open-access-model/>.
- Welsh, J., 2020. Struggling beyond the paradigm of Neoliberalism. Thesis Eleven 158 (1), 58–80. <https://doi.org/10.1177/0725513619888667>.
- Winn, J., 2015. Writing about academic labour, Workplace. A Journal for Academic Labor 25, 1–5. ISSN 1715-0094.
- Xu, X., Oancia, A., Rose, H., 2021. The impacts of incentives for international publications on research cultures in Chinese humanities and social sciences. Minerva. <https://doi.org/10.1007/s11024-021-09441-w>.
- Yegros-Yegros, A., Capponi, G., Frenken, K., 2021. A spatial-institutional analysis of researchers with multiple affiliations. PLoS One 16 (6), e0253462. <https://doi.org/10.1371/journal.pone.0253462>.
- Zha, Q., Shen, W., 2018. The paradox of academic freedom in the Chinese context. Hist. Educ. Q. 58, 447–452. <https://doi.org/10.1017/heq.2018.22>.

South-South Cooperation in Education

tavis d. jules, Richard Arnold, and Victoria Desimoni, Loyola University Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	317
Toward a Definition of South-South Cooperation	318
Historizing the Origins of South-South Cooperation	319
The Three Historical Phases in the Development of SSC	320
The First Phase	320
The Second Phase	321
The Third Phase	321
South-South Cooperation and Education Policy	322
South-South Cooperation, Transfer, and the Sustainable Development Goals	323
Conclusions and Further Research	325
References	325

Introduction

Since the end of the Cold War, the concept of a Global North and Global South has arisen in the attempt to abandon former pejorative conceptualizations, such as First, Second, and Third Worlds and developed/underdeveloped countries (Altinbas, 2013; Shneider, 2017). In the Three-Worlds conceptualization, the First World referred to those countries with liberal, democratic, and capitalist governments and a high level of industrialization, the Second World was often used to represent communist countries with ties to the former Soviet Union and Russia, and the Third World was viewed as being composed of developing or underdeveloped states that are not aligned with the first or second worlds (Altinbas, 2013). This categorization fitted nicely into the era of the Cold War, as the primary antagonistic blocs of the West or First World and the East or Second World vied for hegemonic power, control, and influence over the non-organized Third World. This conceptualization lost relevance after the Cold War ended.

The global development architecture has metamorphosed, and today, the division of North/South is preferred terminology because it is less polemical, focusing on economic development levels rather than political ideology. While development aid flows have shifted and Southern countries are engendering novel cooperation and collaboration models, this newest categorization is far from flawless. As Jules and Moraes de Sá e Silva (2008) point out, the terms North and South are themselves Western constructs. Shneider (2017) argues that the concept of Global South still has similar problems as earlier terms such as of Third World or underdeveloped countries, the main problem being a persistent geographical understanding of the term and the oversimplification that comes from trying to fit the whole world in two categories. Through similar arguments, Chisholm and Steiner-Khamsi (2009) offer two clarifications to explain what it means to belong to the Global South. The first one is that belonging to the Global South does not necessarily imply being located in the South hemisphere. The second clarification is that the South is a relational concept used in opposition to the North. The term refers to a relationship of inequality, and “the notions of North and South have become a proxy or metaphor for rich and poor, developed and underdeveloped, First and Third World, givers and recipients of aid” (Chisholm and Steiner-Khamsi, 2009, p. 3). They further argue that the problem with this definition is its generality and oversimplification of how relationships between and within countries work. In Moraes de Sa e Silva’s (2009) words, “the problem is not so much with the terminology (South, developing countries, periphery), but rather with the assumption that the label implies homogeneity” (p. 54). To summarize, these conceptual difficulties mirror a more profound problem of power relationships between the Global North and the Global South, including their geopolitical and economic boundaries. These difficulties have led to South-South cooperation (SSC), a newer form of collaboration among developing countries that is based on equality, solidarity, respect, and complementarity. In this paper, we use SSC and South-South transfer interchangeably.

In what follows, we first discuss the difficulty in defining SSC. We show that SSC can be compared to triangular cooperation or North-South-South cooperation. Next, we review the literature that discusses the emergence of SSC and its three modalities, or phases, as described by Moraes de Sá e Silva (2009) (“self-reliance and political strengthening,” “demobilization,” and “best-practice transfer”). Third, we examine the role of SSC in education policy by looking specifically at how triangular cooperation has affected education systems in the Global South. With this objective in mind, we present a summary of the role of SSC in education policy across the Global South and discuss the role of the Sustainable Development Goals (SDGs), specifically Goal 4: on quality of education, in informing and facilitating educational transfer across Southern nations. While UNESCO and the SDGs are not direct tools of SSC, the alignment of goals that they provide as benchmarks promote and often facilitate educational transfer across the Global South. UNESCO encourages intervention in educational policy across borders in this way. This chapter, especially the third section, aims to present how SSC has impacted education in the Global South.

Toward a Definition of South-South Cooperation

The emergence of the Global South concept was in response to the steady rise of economies in the so-called Third World countries that began in the 1970s (Shneider, 2017). This rise meant that the previous categorization that placed Southern countries under the subordination of Northern countries was no longer useful. However, this concept of subordination has come under scrutiny in recent years as

developing countries and civil society have repeatedly criticized the way aid is often used as a neo-colonial tool by developed countries—imposing policy conditionalities on developing countries and trying to aid commercial, political and military interests of donors.

Reality of Aid Management Committee (2010, p. 2)

Given this, the concept of the Third World was not adequate to represent the new joint efforts of previously colonized nations to lift themselves up and stand against Western hegemony.

It has been and continues to be incredibly hard for developing countries to lift themselves from poverty and colonial dependence. This incapacity has been explained through a number of related theories, including Dependency Theory, World-System Analysis (WSA), and the Theory of Unequal Exchange. Immanuel Wallerstein (1930–2019) has been a critical figure in the development of WSA, a more comprehensive theory that is based on Dependency Theory. In Griffiths' (2021) words,

like dependency theory, WSA argues that the economic structures of so-called underdeveloped countries should not be understood as being in the early stages of a linear and universal transition to industrialization and development, but rather as “the result of being involved in the world-economy as peripheral, raw material producing area.” (p. 91)

Similarly, regarding the Theory of Unequal Exchange, Dargin (2014) states that “unequal exchange occurs when developing countries trade low-priced products (e.g., bananas, cocoa, and oil) for expensive manufactured goods (e.g., cars and computers)” (p. xxi). This theory upholds that the global market is, in reality, an exploitive trade system, as all exchanges between the peripheral and core countries in this market result in wealth transfer from the periphery to the core (where the peripheral countries are seen as the developing and under-developed countries, and the core the developed countries). In other words, “the proponents of the Theory of Unequal Exchange concluded that, in the face of uneven economic development, and for a variety of other reasons, free trade, or even trade per se, is inherently iniquitous” (Dargin, 2013, p. xxii). In the twentieth-century, two Argentinian economists, Raúl Prebisch (1901–1986) and Hans Signer (1910–2006), broadened this theory and suggested that primary-product-producing countries were going to decay more and more because primary products were condemned to a long-term and non-stop price degradation (Dargin, 2013).

This realization and the subsequent efforts of the Global South to revert their unprivileged standing in global trading did not by itself lead to the “Rise of the Rest.” The actions of the Global North were equally significant in this process. The USA, for instance, believed that enhancing the economic development of its allies was important, especially as a way to prevent communist ideologies from expanding. American policies also boosted the expansion of China; in their intent to fight the Soviet Union, “the highest levels of US policy making decided to open up China in order to contain the Soviet Union” (Dargin, 2013, p. xxiii). A third example is the movement of manufacturing centers from the USA to Southern countries, which occurred because of the increase of work stoppages in American factories. In order to make more profits in relation to production costs, big corporations decided to move their production to countries with cheaper labor (Dargin, 2013). In this context SSC, and particularly triangular cooperation which involves three actors—two from the South and one from the North or an international organization—became the main instrument for the Global South's economic and political flourishing. After the concept of the Third World was phased out by the term South or Global South, every sort of collaboration among countries in the South began to be called SSC (Jules and Morais de Sá e Silva, 2008).

SSC is “any cooperative initiative between two or more developing countries; it may be carried out by governmental institutions, non-governmental organizations, universities, independent professionals, scholars and researchers” (Morais de Silva, 2009, p. 2). UNESCO (2019) defines SSC as an “exchange of knowledge and resources between governments, organizations and individuals in developing countries, or those from what is known as the Global South” (para. 1). Triangular cooperation or North-South-South cooperation is a form of collaboration within SSC. Abdenur (2009) defines triangular cooperation as

the initiative of one or more southern countries that wish to cooperate with one another. In order to maximize their financial, logistic and technical resources, such countries can ask for the support of a Northern donor as a third partner. Alternatively, a donor can partner with a developing country willing to provide technical cooperation to other Southern partners and whose initiative will make triangular cooperation the Northern donor's priorities and interests. The Northern donor would then offer to support South-South cooperation through a triangular approach by providing financial and/or technical support. (pp. 158–159)

Other forms of cooperation include bilateral SSC (which is government-to-government cooperation between developing countries), triangular SSC (where the financial supporter, the donor, and the recipient are developing countries), triangular multilateral SSC (where a multilateral organization works as the financial supporter for two developing countries), and regional SSC (which is the cooperation between two regional organizations formed by developing countries) (SEGIB, 2008, 2009, 2010). In this chapter, however, we treat SSC and triangular cooperation as one, given that donors from Northern countries, including developed countries themselves, non-governmental organizations (NGOs), and individual philanthropists, play an essential role in enhancing and functioning cooperation among and within developing countries, especially in education matters. In fact, many collaborations and policy transfers among countries in the Global South stem from international agreements and are regulated by the donors' benchmarks. Steiner-Khamsi (2009) argues that "South-South transfer is currently central to the operations of international donors" (p. 256). A clear example is the Education for All-Fast Track Initiative (EFA-FTI), launched in 2002 as a partnership between the donor and developing countries, connecting donors with low-income countries (named "partner countries") to help them achieve the education goals established in the United Nations' Millennium Development Goals (MDGs) agenda. Through the EFA-FTI, donors and the World Bank gave financial support to those low-income countries for whom donors' help was not enough to achieve the needed improvements. To become a partner country and matching with a donor, countries were required to have a poverty reduction strategy and a sound education sector plan that included borrowing educational practices from comparable educational systems.

SSC has received a lot more attention in the past decades since some developing countries became crucial actors in the world economy. In 2010, for instance, "the so-called BRICs economies—Brazil, Russia, India, and China—with 40% of the world's population spread out over three continents, already account for 25% of the global gross domestic product (GDP)" (*The Reality of Aid Management Committee*, 2010, p. 1). However, as Bilal (2012) explains, the rise of the Global South did not happen evenly across Southern countries. Most of the growth of GDP in the Global South is due to the development of the BRICs countries, especially Asian countries. Asian developing countries represented around 10% of the world economy in 1990, and in 2016 they reached approximately 30%. In contrast, Latin America and the Caribbean decreased their GDP representation from 9.5% in 1990 to 8.5% in 2012 (Bilal, 2012). Additionally, there are also significant disparities within countries in each continent. As Bilal (2012) points out, "optimism about the South, its economic development and the growing importance of South-South relations should not blind us to this strongly uneven process between and within countries, and the serious poverty challenges that the South is still confronted with" (p. 11).

Nevertheless, economic development is not the only goal of cooperation and partnership among the Global South. Political strategies are as important as economic growth. The multidimensional character of SSC, which includes economic, political, and social aspects in its cooperation agreements, is a crucial attribute of this type of relationship between countries. As Bilal (2012) states, "the focus on mutual learning and exchanges on good practices, without promoting a certain model of development and imposing a set of policy recommendations, has been particularly appreciated by developing countries" (p. 22). This is the reason why SSC has become a fundamental feature in the rise of the Global South. As Gray and Gills (2016) explain,

development is a concept that attempts to encompass a vast complexity of processes of social transformation. It conveys meanings of great promise and hope to billions of human beings concerning human betterment, and refers to a long-term historical project of the liberation of peoples and nations from the vestiges of colonialism, poverty, oppression and underdevelopment. South-South cooperation (SSC) has been a key organizing concept and a set of practices in pursuit of these historical changes through the vision of mutual benefit and solidarity among the disadvantaged of the world system. It conveys the hope that development may be achieved by the poor themselves through their mutual assistance to one another. (p. 557)

Jules and Morais de Sá e Silva (2008) also highlight the importance of SSC in postcolonial studies, arguing that SSC and transfer are an essential part of the decolonization process helping "developing countries decolonize their educational systems and break their dependency from educators, researchers and scientists from the North" (p. 57).

Historizing the Origins of South-South Cooperation

Given the status of SSC as a consequence of the Cold War, its origins are generally thought of in that era. While SSC existed before the 1950s in more informal and non-institutionalized ways, it was the 1955 Conference in Bandung, Indonesia, that is undoubtedly considered a landmark in SSC development and its transformation into a global political movement (Gray and Gills, 2016). Caruso (2009) offers an example of SSC as far back as the end of the 1800s, when different Spanish colonies across Latin America started to transfer the Bell-Lancaster-System educational model that emerged in Cadiz, Spain. Steiner-Khamsi (2009) also believes SSC dates back to the colonial period. Despite this, Morais de Sá e Silva (2009) places the emergence of SSC around the 1950s, during the Cold War period, and differentiates three historical phases in the development of SSC, discussed in depth below.

As an institutional entity, SSC is realized through principles of cooperation and interconnectedness. Bilal (2012) states that SSC is guided by certain principles that all countries in the Global South share

South-South cooperation, its agenda, and development objectives should be set by countries of the South. It should not be driven by charity and power/dependency relationships. On the contrary, it claims to be based on the principles of equality, solidarity, the respect of national independence and ownership, mutual benefits (promoting win-win outcomes) and complementarity. The principle of non-interference in domestic affairs is key. (p. 22)

Non-interference as a principle is one of the cornerstones of SSC, as “South-South cooperation on development aims to observe the principle of non-interference in internal affairs, equality among developing partners and respect for their independence, national sovereignty, cultural diversity and identity and local content” (*The Reality of Aid Management Committee, 2010*, p. 2). Though this description of SSC, as one of mutual cooperation and respect, seems appealing, the reality is more complex, and there are also multiple disadvantages. For example, the Bretton Woods institutions (specifically, the International Monetary Fund (IMF) and the World Bank) launched Structural Adjustment Programs (SAPs) in the 1980s with global prescriptions aimed at improving international competitiveness, adjusting countries’ economic structures, and restoring their balance of payments. These measures would push developing countries toward exploring deeper forms of Southern cooperation. Countries failing to enact SAPs, whose terms and financial institutions determined conditions, faced fiscal discipline. This would lead to the erosion of national control of the economy and open it up to interference from foreign markets. For instance, in 1988, the World Bank noted that its SAPs placed nations at “risk of ‘ceding national sovereignty’ to transnational firms whose main interests are the perpetuation of unequal distribution of value added from the fruits of export manufacturing” (Mills, 1989, p. 28). Therefore, despite a lack of direct interference, inequality was nevertheless further perpetuated. This is because “SSC as an organizational medium rest on two pillars: political and economic. Any conceptual separation of these pillars is difficult because the two separate arenas are actually interconnected” (Altinbas, 2013, p. 34). With this understanding, “non-interference” during SSC is a paradox; there will always be an imbalance.

The Three Historical Phases in the Development of SSC

The First Phase

In the aftermath of the Cold War, countries in the Global South were predisposed to avoid conflict by any means; they likely understood they could not withstand the same economic devastation that had befallen Europe. Many were also newly decolonized and looking to attain financial independence as well. Much of SSC begins with the anti-colonial movement, “fueled by strong political motivations and practiced by autonomous militants or officially provided by socialist countries” (Morais de Silva, 2009, p. 2). The first phase of South-South transfer began with the solidifying of the concept of “underdevelopment” by the American President Harry Truman in a 1949 speech, wherein he introduced the Four Point Plan, “a bold new program for making the benefits of our scientific advances and industrial progress available for the improvement and growth of underdeveloped areas” (Truman, 1949). Truman’s Four-Point Plan would solidify the concept of “underdeveloped countries” as a homogenous group and create an identity that promoted SSC, given that “the creation of a Third World identity and the image of an in-between block created the conditions for the establishment of an alliance between countries that were—and still are—quite heterogeneous” (Morais De Silva, 2009, p. 8).

While SSC has been practiced informally for centuries, the defining entry point to institutionalized SSC is the 1955 Bandung Conference, which “brought together 29 countries from Asia and Africa to promote economic and cultural cooperation in the Asian-African region ‘on the basis of mutual interest and respect for national sovereignty’” (*The Reality of Aid Management Committee, 2010*, p. 2). After the Bandung conference, “developing countries proposed to build a new international partnership for development on the basis of respect for national sovereignty, equality and mutual benefit” (Cui, 2016, p. 2). This led to the Non-Aligned Movement (NAM) in 1961, the United Nations Development Program in 1965, the Group of 77 in 1967, and the 1978 conference of the Global South held in Buenos Aires, which resulted in the Buenos Aires Plan of Action. The Buenos Aires Plan of Action saw delegates from 138 states adopting the Plan of Action for promoting and implementing Technical Cooperation among Developing Countries (TCDC), which was seen as a blueprint for major changes in approaches to development assistance and called on parties to establish a new international economic order based on the pooling of knowledge.

From the outset, SSC was self-deterministic and anti-interventionist. As Gray and Gills (2016) point out, “the ‘Bandung Spirit’ came to encapsulate policies of non-interference and non-alignment, with the Non-Aligned Movement (NAM) further developing this solidarity to challenge the deepening global inequality while lessening the Third World’s economic and political dependence on the Global North” (p. 558). The “Bandung Spirit” was “an ideology that tried to merge Asian and African nationalism, religion, and humanity” (Widyatmadja, 2005). Realizing their strength in numbers in voting blocs, countries started to consolidate for multilateral negotiations, with leaders such as India’s Jawaharlal Nehru, Tanzania’s Julius Kambarage Nyerere, and China’s Mao Zedong jockeying for leadership roles within these collectives; for example, Mao tried to unite this “Trade Union of the Poor” under his “Three World Theory ... [with] China in the center of the South from where the anti-imperialist struggles against the richer nations were to be orchestrated” (Morais De Silva, 2009, p. 10). Even though little was achieved in terms of creating the desired New International Economic Order, economic integration, military cooperation, humanitarian aid, and technical assistance soared during this time, and since then, “South-South cooperation has been practiced in numerous ways ranging from economic integration,

the formation of negotiating blocs within multilateral institutions, military alliances, [to] cultural exchanges ... these efforts have made important contributions to strengthening the conditions for social and economic development" (*The Reality of Aid Management Committee*, 2010, p. 2). In short, the 1960s and 1970s saw leaders in developing countries looking for alternatives to the existing economic and political order of a bipolar Cold War era. As such, they embarked upon technical cooperation among themselves.

The Second Phase

From 1980 to 1998, there was a demobilization period due to debt crises and SAPs. While Global South countries had benefited from Bretton Woods loans, soaring interest rates during Reaganomics created an economic crisis in Latin American countries, with high recession and inflation leading to the 'lost decade.' Given this situation, "the idea of different states engaging in the generation of a New International Economic Order made no sense vis-à-vis the preached argument that the market was the one to tell what kind of economic order was to prevail" (Morais De Silva, 2009, p. 38) and as states focused inwardly on domestic affairs, "with the outbreak of the debt crisis, the poverty related aid gradually reduced and more and more assistance began to be used for privatization and structural adjustment" (Cui, 2016, p. 2).

Much of the demobilization was in response to the rise of the ten market prescriptions of the Washington Consensus, which obligated neoliberal reforms as part of the SAPs. Intergovernmental organizations "no longer counted on the political mobilization of governments of the South, nor on the advocacy of specialists who defended the view that South-South trade was the way forward" (Morais De Silva, 2009, p. 16) since individual states were now focusing on integrating into the global economy rather than attracting regional South-South investment. This led to the renewed influence of the "First World" countries; for example, the World Bank noted during this time that "African policymakers perceive the staff of international organizations as promoting the view that right-minded policy is a matter of 'macho-political will'" (Mills, 1989, p. 2). The Washington Consensus had indirectly opened up labor markets in Southern countries thanks to the removal of trade barriers, and it was soon noted that "there is a need to analyze the advantages and the disadvantages of structural adjustment policies from the point of view of gainers and losers the Bank has a tendency ... to see more clearly the positive (economic) aspects of reform and to minimize the social and political risks" (Mills, 1989, p. 12).

The Third Phase

With the creation of a public international organization—the Global Development Network (GDN), which was created in 1999² and aims at using research to promote better lives in developing and transitioning countries—SSC has entered a new phase characterized by a reversal in the trade liberalization of the previous decades. Between 2006 and 2016, GDN's focus was on enhancing cooperation with regional networks globally. From 2017 to 2022, its current strategy highlights three areas: (i) strengthening research in low-capacity environments; (ii) joining hands for global excellence; and (iii) putting development research to better use as part of its focus. In the modern era, SSC is shifting in how it is approached, as the appeal is no longer limited to developing countries but also to international entities such as knowledge banks who seek the promotion of "best practices," hence, this can be referred to as the "Best Practices phase." This concept is the cornerstone of the GDN, which was created to foster the transfer of "best practices" between underdeveloped countries. As it became clear that SAPs were only leading to rising costs and high unemployment, Global South countries mobilized economically, first through the World Trade Organization (WTO) and later at the G8³/G20⁴/GX⁵ conferences. This has led to "a rapidly expanding South-South trade, investment, and cooperation in a growing variety of fields—including regional economic integration, national security, health and the environment—rest primarily on the increasing competitiveness of Southern countries in world production" (Arrighi and Zhang, 2011, p. 33).

The new century led to an increase in up-and-coming economies that were relying on promoting SSC and which would encompass new features, including "a number of new development finance institutions led by developing countries" (Cui, 2016, p. 2). This is a significant reversal from the era of the SAPs, and now the "Washington Consensus ... has backfired, creating conditions of a reversal of power relations between the global North and South that may well be reshaping world politics as well as the theory and practice of national development" (Arrighi and Zhang, 2011, p. 1). In a way, it has come full circle to the antagonism of the first phase, as leftist governments (especially in Latin America) returned to people-centric policies.

²The agreement was signed by Colombia, Egypt, India, Italy, Senegal, and Sri Lanka.

³Group of Eight (G8) refers to the group of eight highly industrialized nations—France, Germany, Italy, the United Kingdom, Japan, the United States, Canada, and Russia—who meet yearly at an intergovernmental forum from 1997 until 2014 to discuss international issues. Today it is the G7 as Russia was suspended. At times we speak of the G8+5 where the five leading emerging markets—Brazil, China, India, Mexico, and South Africa—are invited to participate in G7 meetings.

⁴Built as the premier forum for international cooperation, since 1999, the G20 brings together the world's major economies whose members account for more than 80% of world's Gross Domestic Product (GDP), 75% of global trade, and 60% of the world's population. Its participants are Argentina, Australia, Brazil, Canada, China, France, Germany, Japan, India, Indonesia, Italy, Mexico, Russia, South Africa, Saudi Arabia, South Korea, Turkey, the United Kingdom, the United States, and the European Union. Spain is invited as a permanent guest.

⁵The Group X (or 10), not to be confused with the Groups of 7, 8, 20, or 24, consists of a group of 11 countries with similar economic interests who consult each other and cooperate on international financial matters. They are Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, Switzerland, the United Kingdom, and the United States, with Switzerland playing a minor role.

the Southern countries' hostility is not against globalization or capitalism, because they use each of these processes and systems to their own advantages. Their reaction is against a capitalist system that somehow designates them as perennial losers and against global trade policies that deprive them of influence. Their goal is not to change the system completely but to make the system more inclusive.

Altınbas (2013, p. 59).

International agencies now act as conduits for the transfer of South-South practices, since "the real significance of South-South development cooperation (SSDC) lies not so much in the magnitude of official development assistance (ODA) or financial resources flowing between developing countries but rather in the character of the relationship expressed by these exchanges" (The Reality of Aid Management Committee, 2010, p. 1).

South-South Cooperation and Education Policy

SSC is about technical cooperation, knowledge and resources transfer, and innovative knowledge exchange between developing countries in areas to share knowledge. When development became more than just economic growth and started to encompass other aspects such as freedom, empowerment, and citizens' well-being, education became a key area of development strategies. Therefore, education became central in any SSC initiative (Morais de Sá e Silva, 2009). SSC in education has not always involved the mediation of third parties, such as international organizations or Northern countries. This means triangular cooperation was not, and it is not the only form of SSC in education. The forms of cooperation that do not include mediators that serve as financial supporters are bilateral South-South, triangular south-south-south, and regional South-South cooperation (SEGIB, 2008, 2009, 2010). An example of a non-mediated SSC in education is the case of the *¡Yo, Sí Puedo!* (Yes, I can) literacy method further explained below. This educational initiative emerged from the South (Cuba) and was transferred to numerous other Southern countries as a successful way to fight analphabetism.

However, since the 1990 Jomtien Education for All (EFA) conference, SSC in education has become more and more mediated by international organizations and powerful "donor" states. In this conference, countries worldwide agreed that education is critical for development and poverty reduction (Samoff, 2009). However, education also started to be seen as a unique area for foreign aid as it enables donors' control over knowledge. Samoff (2009) defines foreign aid as a "tool" or "set of tools," as "managing transfers and exchanges means not only supporting schools and learning but also managing the flow of ideas and the accepted procedures for validating knowledge" (p. 125). He continues by saying that "understood in this way, the aid relationship clearly has the potential to extend rather than eliminate poverty and to entrench and institutionalize rather than reduce global inequalities" (Samoff, 2009, p. 125).

The EFA strategy designed in the Jomtien conference significantly exposes the huge expectations in the role of SSC or triangular cooperation in helping reach educational benchmarks across the Global South. Morais de Sá e Silva (2009) explains how SSC for EFA promotion has been specially designed and executed by the Group of E-9 countries,⁶ which represents some of the largest education systems in the world and over half of the world's population:

E-9 countries have not only been recognized as crucial places for the promotion of universal quality education; they have also demonstrated internal capacity to design efficient EFA programs. They are, therefore, seen as important "reference societies" for other developing countries and, by means of South-South cooperation, their nationally engineered programs can be spread to other places counting on UNESCO's facilitation. (pp. 51–52)

As suggested here, the concept of SSC in education today mainly entails policy transfer of international "best practices" that can lead to educational development as defined by the Global North. This intention behind SSC has changed the nature of collaboration across the Global South. For example, during the first phase of SSC, the foremost advocates of cooperation among the South were the same Southern countries. Today, in contrast, "South-South cooperation has, to some extent, escaped the 'hands of the South'" (Morais de Sá e Silva, 2009, p. 53), meaning that it is in the interest of Northern governments and international organizations, especially international knowledge banks. Another critical change is the motivations. Even though the original aims of the Global South to gain autonomy and international power are still important, SSC has been focusing much of its energy on expanding "best practices" and meeting global development benchmarks. A last change to highlight is how increasingly blurry the concepts of North and South have become, given the growing inequalities and differentiation among countries in the so-called Global South (Morais de Sá e Silva, 2009).

In some of these "best policy" transfers, the principles of SSC mentioned above can be seen as being facilitated by international knowledge banks. UNESCO has aligned many of its policy goals with the principles of SSC. For example, its Malala Fund for Girls' Right to Education has the core goals of "expand[ing] access to education for girls and women, especially those hardest to reach and affected by conflict and disaster" (UNESCO, 2020, para 1), which aligns with the principles of equality and solidarity, while implementing programs that train teachers directly, without attempting to sway internal policy. In terms of non-interference, the

⁶E9 countries are Bangladesh, Brazil, China, Egypt, India, Indonesia, Mexico, Nigeria, and Pakistan.

motivation of SSC to expand best practices across the Global South has darkened this principle, as benchmarking is creating indirect governance by Northern countries, international institutions, and international knowledge banks.

Considering that donors from the North and international knowledge banks regulate SSC in education development, several scholars wonder whether SSC is really “a way out from the dependency track in educational development” (Steiner-Khamsi, 2009, p. 257). This begs the question, are there any educational policies really emerging from the Global South and regulated by it? Or is SSC in education just a way to hinder the persistence of Western hegemony over educational practices worldwide? As Steiner-Khamsi (2009) states, “South-South cooperation can be seen as a vehicle to accelerate the accomplishment of development targets established by the North” (p. 257).

In addition, countries in the North have become involved in triangular cooperation to play a role in international development cooperation or official development assistance (ODA). They participate actively in this role, as appearing as a great global donor increases countries’ legitimacy as powerful and influential. Mochizuki (2009) presents the case of Japan, which became one of the largest donor countries, aiding developing nations in both Asia and Africa as “a new aid scheme that would allow it to achieve its broader goals in foreign relations, including attaining the UN Security Council permanent membership” (p. 81). Getting involved in SSC or, more specifically, triangular cooperation, became a strategy used by Japan to “establish itself as a ‘first class’ developed country in the international community” (Mochizuki, 2009, p. 81).

South-South Cooperation, Transfer, and the Sustainable Development Goals

This section looks at educational SSC in the context of the SDGs, which provides an external series of benchmarks that helps align SSC. With its use of political dialogue and financial cooperation, SSC has contributed significantly to shaping the 2030 Agenda for Sustainable Development. Education plays an essential role in achieving many SDGs, and SDG 4 aims to promote lifelong learning opportunities for all while ensuring inclusive and equitable, quality education. Global education is the pivot point for it. SDG 4 is comprehensive; it expands from early childhood education through basic education to vocational and secondary education and lifelong learning. The idea of “lifelong learning for all” is a crucial element of SDG 4. SDG 4 has 10 targets⁷ which are measured by 11 indicators. These 10 targets are considered the backbone of Global Campaign for Education (GCE)’s policy and advocacy work since 2015.

UNESCO has a crucial role in promoting triangular cooperation in educational matters by strengthening cooperation between developing countries and attracting donors (developed countries or NGOs) who affect SSC. As part of its mission statement, UNESCO (n.d.) calls for “cooperat[ion] with Member states, civil society, and the academic world [to] promotes public-private partnership in education” (para. 4). Some examples of educational projects run by UNESCO that promote SSC are: (i) the Capacity Development for Education Program (CapED) created in 2003 and modified in 2016, and is aligned with the SDG 4—Education 2030 agenda to offer quality education opportunities for all (ii) the Enhancing Teacher Education for Bridging the Education Quality Gap in Africa project, which is funded by the Chinese government and uses information communication technologies to enhance teacher training; (iii) UNESCO’s Malala Fund for Girls’ Right to Education, which was established in 2012 to increase girls’ access to quality and gender-responsive education and ensure safe learning environments; (iv) Youth Employment in the Mediterranean Project (YEM), which is funded by the European Union and implemented between 2018 and 2020 and addressed issues of youth employment; (v) Best Practices in Mobile Learning, launched in 2016, as a 5-year program, with support from the Fazheng Group, seeks to create equitable learning environments through school-wide planning and use of mobile learning. There are many more programs that UNESCO runs under the banners of capacity development, teacher education, girls and women’s education, technical and vocational education and training, and ICT in education.

Although all 10 SDG targets are important to consider when talking about education in relation to SSC, four targets are more relatable in this context:

- Goal 4.3: By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university.
- Goal 4.4: By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills.
- Goal 4.5: By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations.
- Goal 4.6: By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy.

⁷These 10 targets are broken down into seven outcome targets and three means of implementation. The seven outcome targets are (i) universal primary and secondary education; (ii) early childhood development and universal pre-primary education; (iii) equal success to technical/vocational and higher education; (iv) relevant skills for decent work; (v) gender equality and inclusion; (vi) universal youth literacy; and (vii) education for sustainable development and global citizenship. The three means of implementation are: (i) effective learning environments; (iii) scholarships, and (iii) teachers and educators.

According to Groenez et al. (2007) and Blossfeld et al. (2014), a correlation exists between education participation rates and governmental investment in industries, innovation, and infrastructures. This justifies the idea of education being used as a powerful “tool” to develop a more inclusive and equal society.

With the introduction of the SDGs, we can see SSC educational transfer aligning with the SDGs; in particular, Goal 4. Goal 4 of the SDGs integrates within UNESCO’s earlier EFA project, which had six specific goals, including “achieve[ing] a 50% improvement in adult literacy by 2015, especially for women, and equitable access to basic and continuing education for all adults” (World Bank, 2014, para 5) which matches Goal 4.6, and “eliminate gender disparities in primary and secondary education by 2005 and achieve gender equality in education” (World Bank, 2014, para 6) which matches Goal 4.5. With overlaps such as these occurring in congruent SSC education goals, a stronger picture emerges of the direction between triangular cooperation donor countries want to align their efforts.

We can see various SSC initiatives aligning with SDG 4, both directly and indirectly. China has committed to aligning its Belt and Road Initiative (BRI), which focuses on infrastructure development predominantly in the Global South, with the SDGs, including Goal 4. For instance, the “BRI conceptualizes an education community as a multicultural group that transcends geographical, cultural and linguistic boundaries, while considering common interests and common responsibilities as the core foundation on which equality and respect are built” (Peters et al., 2020, pp. 1044–1045) and “both are committed to promoting inclusive and sustainable economic growth and social development; the BRI’s vision to realize diversified, independent, balanced, and sustainable development in the countries along the BRI echoes the sustainable development goals set out in the 2030 Agenda” (Yin, 2019, p. 8). While not referencing this SDG specifically, the United Arab Emirates has also been aligning its SSC funding toward cohesion with the SDGs; for example “in 2016, 76% of the UAE foreign assistance supported the eight SDGs highlighted in the UAE Foreign Assistance Policy” (Khalid, 2018, p. 456), while benchmark 4.4 neatly matches the UAE’s goal to “create and strengthen existing technological capacities ... in order to improve the effectiveness with which such capacities are used and to improve the capacity of developing countries to absorb and adapt technology and skills to meet their specific developmental needs” (Khalid, 2018, p. 454) in their development aid. Meanwhile, in Japan, education funding matches 4.3 when “the 1999 Medium-Term Policy on ODA also stated that Japan would focus on the ‘quality,’ rather than quantity, of aid activities in the promotion of its ODA” (Mochizuki, 2009, p. 10) to developing countries, as well as the promotion of basic literacy

a dramatic shift of focus from higher education and technical and vocational education in Asia to basic education in Africa was an integral part of Japan’s new ODA strategies aimed at enabling Japan to play a more central role in international development cooperation.

Mochizuki (2009, p. 11).

It is clear that in broad terms, at least, SDG 4 is guiding SSC development.

Through its own donor-funded arms, the UN coordinates much of the alignment of SSC with its own SDGs. For example, in 2015, it is the Addis Ababa Action Agenda, which “brought together over 1.5 trillion for education and skills development infrastructure by consolidating funds from AIIB, the Global Infrastructure Hub, the BRICS NDB, the Asia Pacific Project Preparation Facility, the World Bank Group’s Global Infrastructure Facility” (UNOSSC, n.d., p. 139). The UN’s Science, Technology, and Innovation Strategy for Africa (STISA) “have already decided to support capacity-building efforts in education and research across Africa through the creation of scientific centers and networks of excellence and the upgrading of both research and teaching at universities” (UNDP, p. 18). Moreover, in Bangladesh, “BRAC recruited a village woman with eight or more years of education to serve as para-professional teachers for 3–4 h per day for a small stipend” while “a School Management Committee (SMC) selected the times and days the school would meet according to the needs of the parents” (Chabbot, 2009, p. 3). By using its donor apparatuses to engage in triangular cooperation, the UN is positioning the SDGs as a driving force across SSC development.

The case of the ¡Yo, Sí Puedo! ([YSP] *Yes, I can*) literacy method, designed by ALBA-TCP (Bolivarian Alliance for the Peoples of Our America—Peoples’ Trade Treaty), is an example of where we can see SDG 4 in a specific SSC initiative. The ¡Yo, Sí Puedo! program was designed in the 2000s by the Cuban educator Leonela Relys Diaz, to expand basic literacy among people of all ages. From 2001 onwards, YSP started to be employed in numerous countries in the Global South, gaining a honorable mention in UNESCO’s international Literacy Prize (Boughton and Durnan, 2014; Muhr, 2015). Boughton and Durnan (2014) define YSP as “a model for mounting low-cost mass adult literacy campaign across regions and countries, involving extensive coordination and mobilization of all the relevant government and non-government agencies and actors and the population as a whole” (p. 327). The objectives of this program, therefore, are very much in line with the goals of the SDG 4, quality of education, in that it has enhanced literacy among millions of adults (in line with goal 4.6), promoting long-life learning opportunities (one of the critical components of goal 4). Especially the ¡Yo, Sí Puedo Seguir! (*Sure, I can continue*) focuses on ensuring literacy for all throughout life. As Muhr (2015) explains

by late 2014, 3,815,092 people of all ages (meaning that no particular group age is targeted) are officially stated to have acquired basic literacy through ¡Yo, Sí Puedo! associated literacy campaigns in the alba-tcp member territories, and 1,174,312 have completed post-literacy non-formal elementary education ... by late 2014, 8,203,324 people had acquired literacy through ¡Yo, Sí Puedo! worldwide, and over one million have benefited from ¡Yo, Sí Puedo Seguir! (p. 128)

In the last 15 years, the program has reached 29 countries, and it has made basic literacy accessible to more than six million people, mainly in the Global South. However, as [Muhr \(2015\)](#) and [Boughton and Durnan \(2014\)](#) highlight, the program has not been significantly noticed and studied in the Global North. In addition, it works by bringing together multiple actors in collaboration. In fact, this literacy campaign that emerged in the Global South was only able to function thanks to the joint efforts of two of ALBA-TCP's founding members, Cuba, and Venezuela, guided by the SSC principles of complementarity, solidarity, and cooperation.

[Muhr \(2015\)](#) argues that “the case of ¡Yo, Sí Puedo! should not be regarded as ‘best practice transfer’ among developing countries, but as integral to South-South cooperation as a collective counter-hegemonic process of Third World liberation and emancipation for structural transformation toward a socially just and democratic world order” (p. 126). Counter-hegemony is part of SSC's motivations and ideas, representing the aims of the Global South to gain economic independence from the Global North and overcome their “under-development.” The counter-hegemonic process emerged at the same time as SSC, in the 1955 Bandung conference, and acquire strength with the foundation of the Non-aligned Movement (NAM), the foundation of the G-77 Group (now composed by 134 developing countries), and the 1978 Buenos Aires Plan of Action for Promoting and Implementing Technical Cooperation among Developing Countries (BAPA), signed by 138 UN members. The BAPA is unique in that it was the first framework to incorporate this type of cooperation and called for respect for sovereignty, non-interference in internal affairs, and equality of rights, among others. The YSP literacy method promoted by ALBA-TCP, which became global, is critical because it represents the counter-hegemonic efforts that started to emerge from the SSC and the role of education in this process. In [Muhr's \(2015\)](#) words, YSP

does not limit itself to learning to read and to write only. It goes much further. It pursues a conscientising and transformatory literacy, besides schooling ... It has as its fundamental objective the active insertion of the participants in the social, economic and political activities of the community and the countries in which they live. (p. 130)

In this way, YSP contests the “new global governance of education” ([Muhr, 2015](#), p. 132), as it does not focus on transferring “best practices” nor involves any type of conditionalities or coercion from tertiary parties. It is a program based on voluntary participation and, therefore, a real example of pure SSC that is engendered within and strengthens the Global South.

Conclusions and Further Research

This chapter has presented SSC in the arena of education. After providing the theoretical lens of SSC, it has charted the history of SSC, from its decolonial origins through its three phases of development since the Cold War. It then provided examples of how SSC, through the guidance of the UN, aligned with SDG 4. Finally, it presented a practical model of educational SSC transfer that is likely informed by SDG 4. While North-South cooperation is the orthodoxy, SSC is used among countries that share similar historical realities and similar challenges. SSC has contributed to many knowledge exchanges through programs, projects, and initiatives that have helped solve challenges faced by Southern countries.

While SSC can spur educational development in countries, the educational arena it operates in is still highly postcolonial and Western-oriented. SSC is affecting the hegemony of developed countries in the educational policy arena by realigning policies with Western notions of “development.” While “on the one hand, SSC is a multiregional initiative designed to reduce the South countries' economic dependency on markets in the North, ... on the other hand, SSC must necessarily involve changes in political dependencies” ([Altinbas, 2013](#), p. 58). As long as the UN is a significant donor arm in SSC coordination, there will be a political dependency on Western nations. While “SSC's ultimate goal is to close the socio-economic gap between the developed and developing countries, in any logic scheme, SSC should reduce the South's vulnerability to arbitrary or unilateral pressures from industrialized nations” ([Altinbas, 2013](#), p. 34). While SSC emphasizes equality and basic functional skills in the education arena, it is the international donors who define and delineate these concepts.

Future research can be informed by locating what is needed to revise the agenda that was set when EFA was introduced. We have moved into a post-knowledge economy where the agenda should focus on lifelong learning. It is time for the international community to convene again around the topic of SDG 4 and try to build consensus around what the next policy agenda framework would look like. Before the SDGs can become gospel, “triangular cooperation must overcome the challenges of high transaction costs, limited funds for projects, lack of trust between OECD DAC countries and their counterparts among emerging economies, and the possibility of adding to the already strained bureaucracies in the Global South” ([UNOSSC, n.d.](#), p. 136). The control of infrastructure development will set the stage for the next decade of educational reform in the Global South.

References

- Abdenur, A., 2009. The strategic triad: forum and content in Brazil's triangular cooperation practices. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 157–170.

- Altınbaş, D., 2013. South-South cooperation: a counter-hegemonic movement? In: Dargin, J. (Ed.), *The Rise of the Global South: Philosophical, Geopolitical and Economic Trends of the 21st Century*. World Scientific Publishing, Singapore, pp. 29–65.
- Arrighi, G., Zhang, L., 2011. Beyond the Washington consensus? A new Bandung. In: *Globalization and Beyond: New Examinations of Global Power and Its Alternatives*, pp. 25–57.
- Bilal, S., 2012. What is the rise of south-south relations about? In: *Development, Not Aid*. MO papers, Brussel.
- Blossfeld, H.-P., Kilpi-Jakonen, E., Vono De Vilhena, D., Buchholz, S. (Eds.), 2014. *Adult Learning in Modern Societies: An International Comparison from a Life-Course Perspective*. Edward Elgar Publishing, Cheltenham.
- Boughton, B., Duman, D., 2014. Cuba's Cuba's Yo, sí puedo. A global literacy movement? *Postcolonial Direct. Educ.* 3 (2), 325–359.
- Caruso, M., 2009. Imperial connections, entangled peripheries: Cadiz and the Latin American monitorial schools. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 17–38.
- Chisholm, L., Steiner-Khamsi, G. (Eds.), 2009. *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town.
- Chabbot, C., 2009. BRAC goes global. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press, New York, NY, pp. 192–209.
- Cui, W., 2016. Comparison between north-south aid and south-south cooperation: based on the analysis of the new development finance institutions. *J. Shanghai Jiaotong Univ.* 21 (1), 25–32.
- Dargin, J. (Ed.), 2013. *The Rise of the Global South: Philosophical, Geopolitical and Economic Trends of the 21st Century*. World Scientific Publishing, Singapore.
- Gray, K., Gills, K., 2016. South-South cooperation and the rise of the global south. *Third World Quarterly* 37 (4), 557–574.
- Griffiths, T.G., 2021. Dependency theory and world-systems analysis in Comparative and International Education. *Critical accounts of education and development*. In: Jules, T.D., Shields, R., Thomas, M.A.M. (Eds.), *The Bloomsbury Handbook of Theory in Comparative and International Education*. Bloomsbury Publishing, Great Britain, pp. 87–108.
- Groenez, S., Desmedt, E., Nicaise, I., 2007. Participation in Lifelong Learning in the EU-15: The Role of Macro-level Determinants. Paper Presented at the European Conference for Education Research (ECER), Held 19–21 September 2007 in Ghent, Belgium.
- Jules, T.D., Morais de Sá e Silva, M., 2008. How different disciplines have approached South-South cooperation and transfer. *Soc. Int. Educ. J.* 5 (1), 45–64.
- Khalid, A., 2018. Implementing global strategy in the UAE foreign aid: from Arab solidarity to South-South Cooperation. *Вестник Российского университета дружбы народов. Серия: Международные отношения* 18 (3).
- Mills, C.A., 1989. *Structural Adjustment in Sub-Saharan Africa*. World Bank, Washington, DC.
- Mochizuki, Y., 2009. Japan's official development assistance strategies and UN conceptualizations of South-South cooperation. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 63–86.
- Morais de Sá e Silva, M., 2009. South-South cooperation: past and present conceptualization and practice. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 39–59.
- Muhr, T., 2015. South-south cooperation in education and development: the ¡Yo, sí puedo! Literacy method. *Int. J. Educ. Dev.* 43, 126–133.
- Peters, M., Oladele, O., Green, B., Samilo, A., Lv, H., Amina, L., Wang, Y., Chunxiao, M., Chunga, J., Rulin, X., Ianina, T., Hollings, S., Yousef, M., Jandrić, P., Sturm, S., Li, J., Xue, E., Jackson, L., Tesar, M., 2020. Education in and for the belt and road initiative. *Educ. Philos. Theory* 52 (10), 1040–1063.
- Samoff, J., 2009. Foreign aid to education: managing global transfers and exchanges. In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 123–156.
- SEGIB, 2008. *II Informe de la Cooperación Sur-Sur en Iberoamérica*. SEGIB, Madrid.
- SEGIB, 2009. *III Informe de la Cooperación Sur-Sur en Iberoamérica*. SEGIB, Madrid.
- SEGIB, 2010. *IV Informe de la Cooperación Sur-Sur en Iberoamérica*. SEGIB, Madrid.
- Shneider, N., 2017. Between promise and skepticism: the Global South and our role as engaged intellectuals. *Global South* 11 (2), 18–38.
- Steiner-Khamsi, G., 2009. Conclusion: a way out from the dependency trap in educational development? In: Chisholm, L., Steiner-Khamsi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press and HSRC Press, New York and Cape Town, pp. 241–258.
- The Reality of Aid Management Committee, 2010. *South-South cooperation: a challenge to the aid system?*. In: *Special Report on South-South Cooperation 2010*. IBON Books, Philippines.
- Truman, H., 1949. Second Inaugural Address. Harry S. Truman. Library. Museum. <https://www.trumanlibrary.gov/education/presidential-inquiries/challenge-international-aid>.
- UNESCO, 2019. *South-South Cooperation in Education*. UNESCO. <https://en.unesco.org/themes/education/south-south-cooperation>.
- UNESCO, 2020. *UNESCO Malala Fund for Girls' Right to Education*. UNESCO. <https://en.unesco.org/themes/education-and-gender-equality/malala-fund>.
- United Nations Office of South-South Cooperation (UNOSSC), (n.d.). *Cooperation Beyond Convention*. South-South and Triangular Cooperation in a Changing Global Landscape.
- Widyatmadja, J., 2005. *The Spirit of Bandung*. Retrieved from. <https://yaleglobal.yale.edu/content/spirit-bandung>.
- World Bank, 2014. *Education for All*. Retrieved from. <https://www.worldbank.org/en/topic/education/brief/education-for-all>.
- Yin, W., 2019. Integrating sustainable development goals into the Belt and road initiative: would it be a new model for green and sustainable investment? *Sustainability* 11 (24), 6991.

Global testing regimes: accountability, standardization, racial inequality and the rise of audit culture

Goli M. Rezai-Rashti and Shirin Abdmolaei, The University of Western Ontario, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	327
International organizations and comparative data	327
OECD's PISA, high performing and high equity	328
Neoliberalism, racialized students, testing and accountability	329
Mediatization and standardized testing	332
Conclusion	333
Acknowledgment	334
References	334

Introduction

While testing and using data to monitor student performance is not a new phenomenon, its current manifestation emanating from the Organization for Economic Co-operation and Development (OECD) and various national education systems corresponds with changes in the global education policy discourse that have placed greater emphasis on numbers, metrics and the promotion of an audit culture (Lewis, 2020; Lingard et al., 2016; Lingard, 2011; Rizvi and Lingard, 2010; Moreno-Salto and Robertson, 2021). To draw on the work of Shore and Wright (2015a,b), we are increasingly governed by numbers, indicators, and audits, and “to be audited and inspected is now regarded as an axiomatic part of personhood” (p. 23). This article foregrounds this policy shift in the significance of data and standardization and argues that this shift does not impact all students equally but has serious consequences for racialized students and students from low-income families. Moreover, it leads to further segregation of racialized students and, coupled with the mediatization of test data and ranking of schools based on test scores, testing has resulted in stigmatization, lower enrollment rates and closure of low-performing schools in certain racialized and poor neighborhoods.

This paper begins with an overview of the global context of testing and a discussion of the role international organizations such as the OECD play in steering education policy globally and nationally. This paper then proceeds to discuss the impact of neoliberal governance in driving national education systems toward datafication, accountability and comparisons with consequences for racialized students and students from low-income families.

International organizations and comparative data

Since the 1990s, international organizations such as the OECD have been important players in the development of educational indicators (Henry et al., 2001). The connection between comparative international performance data and global economic competition has resulted in the OECD taking on an increasingly significant role as an education policy actor on the global stage, performing an important task by globalizing educational accountabilities across nations (Rizvi and Lingard, 2010). In fact, Leuze et al. (2007) attribute educational changes around the globe to two main trends: (1) the increasing involvement of international organizations and, (2) the growing marketization of education. According to the authors, these two changes not only resulted in international organizations such as the OECD becoming more involved in policymaking, but education has turned into a tradable commodity, subsequently creating competition among countries and between students. Current research regarding the role of the OECD in education positions it as a significant “institutional network that evolved into a ‘global political superstructure’” (Sørensen et al., 2021, p. 101).

Grek (2013) suggests that the rise of international comparative assessment began in the early 1990s with the comparative assessment of international adult literacy in several countries. This, Grek argues, has offered the OECD visibility and exposure which resulted in placing international tests in a comparative dimension. Discussing the historical trajectory of the OECD in education, Grek (2013, p. 706) argues that the “OECD’s capacity has not only been its ability to ‘move’ experts around the world and bring them around the same table, but also to then effectively steer and direct toward its own prespecified agenda.” In his analysis of the role of international organizations in education such as the OECD, Ball (2003, p. 216) discusses “new” policy technologies and how the metamorphosing of these international organizations is creating greater emphasis in “monitoring systems and the production of information.” Ball (2003, p. 216) suggests that “Performativity is a technology, a culture and mode of regulation that employs judgments, comparisons and displays as means of incentive, control, attrition and change based on rewards and sanctions (both material and symbolic).” As Jakobi and Martens (2010, p. 176) argue, “The OECD today not only define the problem but offers the solution With the new generation of indicators, the Organization has therefore gained an important status in several stages

of national policy-making from agenda setting to policy formulation and implementation.” In other words, the OECD, as a major international organization, has garnered the capacity to become a key policymaker and policy player, shifting educational policy discourses and driving global educational competition.

OECD's PISA, high performing and high equity

In promoting the important role of the OECD in the collection of comparative data, the OECD refers to the significant role of political leaders and politicians and the values they place on education, which results in achieving equitable outcomes regardless of cultural and economic diversity:

Starting from very different levels, a number of countries and regions have succeeded over the last few years in raising their students' performance substantially. They display some important common features. Their politicians and social leaders share with parents, teachers and students a strong belief in the value of education ... The best systems deliver strong and equitable learning outcomes across widely varying cultural and economic contexts.

OECD (2010) cited in Lingard et al. (2016, p. 95).

Following this connection between equity and achievement, the OECD (2010), through the Program for the International Student Assessment (PISA), an international assessment that measures the performance of reading, mathematics and science among 15-year-old students, identified Canada as a country whose students are high achieving. Using the high-performance rates of select immigrant populations on PISA, this was taken as indication for Canada's apparently highly equitable education system. However, the implications of this data is that it fails to account for the reality of Canadian-born racialized students—students who have low educational success rates and are much more likely to drop-out of school. According to a 2006 report by the Toronto District School Board (TDSB), the largest board of education in Canada with a high percentage of racialized students, of the 25% of students who do not graduate secondary schools, most of them are from Indigenous, Black, Hispanic and Middle Eastern backgrounds. The TDSB (2006) also mentions that these students have among the lowest rates of attendance and highest rates of suspension. They also have the lowest test score rates on standardized tests, such as the Education Quality and Accountability Office (EQAO), which tests the reading, writing, and math skills of all students across the province of Ontario in Grades 3, 6, 9, and 10.

Interestingly, this distinction between immigrants and non-immigrant racialized Canadians is ignored and overlooked in the OECD's report of Canada's performance on PISA, which has implications for racialized students; an important issue that we will delve into later in this paper. But for now, it is important to note that Canada's immigrants are admitted through the Point System and are thus likely to be highly educated. Using the OECD's data, Beach et al. (2006, p. 3) suggest that under the Point System, immigration has shifted from family reunification toward a class of economically privileged, skilled workers and a greater emphasis on accepting immigrants who are educated, young, and have capital. Indeed, the influx of more financially privileged, young, educated immigrants directly correlates with the success of immigrant students in Canadian schools which the OECD celebrates. It is ironic, then, that the OECD's (2006) earlier statement on immigrant success contradicts their own data that “on average, first generation Canadian students had parents with as many or more years of education as native-born parents” (OECD, 2006 cited in Mehta and Schwartz, 2011, p. 150).

Martino and Rezai-Rashti (2013, p. 607) research on equity, the OECD's PISA and the gender achievement gap provides a cautionary note for education policymakers to “not rely too heavily on PISA measures as basis for decision-making regarding questions pertaining to equity.” The authors also critique the OECD's PISA in their use of comparative data that often endorses school-based interventions while giving less attention to significant structural dimensions of inequality regarding race and social class—two significant factors that shape educational outcomes. Despite this, however, there is an undeniable role that international organizations such as the OECD play in not only producing new knowledge, but steering education policy and disseminating performance data globally, which has implications for education and educational equity that we should not take lightly. As stated by Grek (2013, p. 697), “international comparative testing appears then as much more than simply a statistical project; it has become part of a consistent effort to restore legitimacy and trust between populations and their governments.” This political use of PISA is also discussed by Moreno-Salto and Robertson (2021) in their research on the role of OECD's PISA in Mexico. They show how politicians have used PISA's comparative performance data to justify their own political objectives and how policymakers “draw upon much earlier renditions of PISA so that it now enters into Mexican education policymaking and shapes practice through a metaphoric back door” (Moreno-Salto and Robertson, 2021, p. 213).

It is also important to recognize that the significant role of international organizations does not reduce the position of nation-states in steering policies relevant to their own contextual factors (politics, culture, and history) while also considering the influence of global educational accountability policies. Novoa and Yariv-Marshall (2003), for example, have written about the “global eye” and the “national eye” and how they govern together through complementary international and national testing. However, it needs to be acknowledged that the OECD plays a significant role in driving education policies in order to facilitate an exercise of global governance which has broader implications for education policy and practice (Sharman, 2012).

Neoliberalism, racialized students, testing and accountability

Recent focus on the use of quantitative data and performance-based assessment have been critiqued by critical policy scholars as a technology of neoliberal governance in the education field (Ball, 2003; Grek, 2013; Hursh, 2013; Lipman, 2013; Lingard et al., 2013; Ozga, 2009). It is argued that the focus on comparative performance data of various countries with distinct political, economic, racial and cultural structures makes it quite problematic and “results in a ‘governance by comparison’ approach that generates truth claims and a particular way of seeing the world in which the material effects of race and poverty remain unmarked and are expunged from political and policy consciousness” (Lingard et al., 2016, p. 92).

Nichols et al. (2006), in their study of the relationship between pressures for high-stakes testing and student achievement, suggest that although the US has had a long history of standardized testing, its current emphasis on high-stakes testing could be traced to the 1983 publication, “A Nation at Risk”. According to the authors, “it was believed that if the public education system did not receive an overhaul, our economic security would be severely compromised” (Nichols et al., 2006, p. 3). The authors believe that this view of education and economic competitiveness was internalized in American culture. Through this logic, the neoliberal point of view considers students as human capital and future workers who “must be given the requisite skills and dispositions to compete efficiently and effectively” (Apple, 2005, p. 214). But as will be discussed in this section, we need to think about which students are benefiting from high-stakes and standardized testing and which students are not, and what purposes testing, under the guise of objectivity, meritocracy and self-responsibility, serves in sustaining a racist, capitalist system in the age of neoliberalism.

One of the most significant criticisms of testing has been its connection to datafication and a neoliberal mode of governance in which metrics, data and accountability play an important role in the production and enactment of policies that justifies and individualizes the performance of students regardless of structural and institutional challenges. Hursh (2013) situates high-stakes testing to the rise of neoliberalism in the US and its particular ideology of reducing the size of government while establishing the market as the model for all economic and social activities, including creating self-responsibilizing individuals who need to rely on themselves. This market model of efficiency has focused on standardized and high-stakes testing as an objective assessment tool to make teachers and students accountable. As Hursh (2013) contends, current reform measures have been justified and rationalized by creating more efficiency, accountability, equity and objective assessment of student performance and competition in the global job market. In this context, individual ability and performance is highlighted while underlying systemic and structural inequalities remain invisible. For example, in relation to race and racial inequality, it has been argued that the move toward comparative performance data and neoliberal governance in education in countries such as Canada and the United States has legitimized racial inequality by using the discourse of meritocracy. Goldberg (2009, p. 339), focusing on race and neoliberal modes of governance, demonstrates how the shift from a welfare/pastoral state to a neoliberal one has resulted in the acceptance of the ideology of self-responsibilizing individuals and “individualized merit and ability.” According to Goldberg (2009, p. 339), such ideologies have contributed to the discourse of racelessness under neoliberalism:

In diluting, if not erasing, race in all public affairs of the state, neoliberal proponents nevertheless seek to privatize racisms alongside most everything else Categories of race disappear as much from keeping account of discrimination itself, thus leaving the condition it is supposed to articulate, to mark and express as well as identify and assess, as untouchable as it tends now to be untouched.

In the US, there has been increasing critical investigation into the role of accountability and standardization in sustaining racial and class inequities, an argument rooted in the historical trajectories of standardized testing. A legacy mired in racism and the notion of White intellectual superiority, justified and institutionalized through the eugenics and IQ testing movements (Au, 2009, 2013), Au (2016) contends that testing began as a “racial project”. Eventually working its way into education, testing was instituted as an objective, fair, value-laden instrument to measure student performance and competence. The implications of this, however, was that such “scientific” testing ultimately declared the poor, racialized, women and immigrants, as intellectually subordinate while serving White supremacy (Au, 2013). As Knoester and Au (2017, p. 7) write:

Standardized tests, in conjunction with the ideology of meritocracy, thus have operated as a tool of White supremacy because they make racist outcomes of the tests appear as a by-product of the way the world works objectively and naturally- they “scientifically” justify the existing racial order, and they do so within a false promise of measuring everyone equally, accurately, and fairly.

The presumed objectivity of standardized tests has left no room for structural considerations of racism and class privilege; negating and denying how race, socioeconomics and other social locations may differentially impact students and their performance on tests. But existing scholarship on this proves otherwise. For example, in their study of standardized testing in the state of Virginia, Brunn-Bevel and Byrd (2015) draw on the historical and structural dimensions of education, detailing how historically, Black and White students were segregated by race, which relegated Black students to schools which were poorly resourced while White students received an education which matched their economic, social and material privileges. This history continues to inform the educational and social outcomes of students in the US, including the persistence of segregated schools, which provides important context to understanding the outcomes of Black and White students on tests. According to the authors, structural

disparities help to explain why, in almost every subject, Black students in their study were more likely to fail standardized tests in comparison to their White counterparts. Utilizing data to ascertain school district size, the ratio of students to teachers, school funding, school segregation and income levels to investigate testing results, Brunn-Bevel and Byrd (2015, p. 444) note that discrepancies in education are “exasperated by accountability systems based on standardized testing that discount the reality of segregated and resource-depleted school environments” that Black students are more likely to encounter than White students. Indeed, structural circumstances cannot be disregarded in conversations about testing, given that, as Marmol (2016, p. 8) writes:

[s]tudents cannot be expected to achieve comparable results if presented with vastly unequal material resources and social conditions, which are then compounded when their histories, cultures and ways of learning are devalued and deemed superfluous. Poverty, unequal funding of schools and distribution of resources need to be remedied and a newfound respect for cultural diversity and equality, or indeed equity, need to be emphasized.

The presumption that standardized tests are neutral, objective and unbiased is ultimately a fallacy when varying lived experiences and access to opportunity, funding and resources are absent from the discussion. Such an absence fails to recognize how testing has ultimately contributed to worsening racial, class and educational inequities (Marmol, 2016). As Thompson and Allen (2012) study of the disastrous impact of the Bush administration’s education policy “No Child Left Behind” demonstrates, the over-emphasis on high-stakes testing in the US has worked at the expense of Black American students who have become largely apathetic and disengaged as teachers have lost autonomy over what and how they teach due to increasing pressures on teachers to ensure that students perform well on tests. Creative, culturally relevant lessons are increasingly being replaced with more rigid and structured curricula as a “narcissistic education system” (Thompson and Allen, 2012, p. 218), where administrators and school boards are overly preoccupied with achieving an image of success based purely on test scores alone, becomes the norm.

The fears and concerns of racialized and low-income students who are subject to standardized testing has also been the focus of scholarly attention. In Kearns (2016) study of students who failed the high-stakes Ontario Secondary School Literacy Test (OSSLT) (a test that all students must pass to obtain their high school diplomas), the author shows that standardized testing works in favor of sorting and ranking students while constructing those who do not pass as “deficient” and illiterate. In the urban setting in which this study was conducted, many of those who did not pass the test were racialized, low-income and did not have English as their first language. Already marginalized based on their varying social identities, failing the OSSLT worked to further highlight their presumed difference; Kearns (2016, p. 135) writing that “the marginalized are confronted with their ‘otherness’ through the literacy test.” The findings of Spencer’s (2012, p. 132) research also suggests that in schools with high numbers of students who are visible minorities or English language learners, “the effects of standardized testing include a range of practices that reinforce inequity and increase social disparity.” These practices include narrowed expectations and intensive curricula that privilege dominant linguistic and cultural norms.

In a separate study, Kearns (2011) found that students who attended school in a community where the poverty rate was nearly 50% and the population consisted largely of first-generation Canadians, experienced shame, humiliation and stress for having failed the test. The youth in this study perceived the OSSLT as inequitable and counterproductive to their learning and well-being, finding the test to be less useful than their high school English classrooms. Indeed, an implication of standardized testing as found in Kearns (2011, 2016) studies is that such tests “alter youths’ perceptions of themselves: they question themselves, their abilities, experience themselves as inferior to others, or consciously oppose an inferior label given to them by a large testing agency” (Kearns, 2011, p. 123).

Such feelings attributed to fears of failing tests, such as stress and angst experienced by Kearns’ high school participants, begins at a much younger age. Eizadirad (2020) found that among racialized children in Grade 3 who had to take the EQAO standardized test preparation in Ontario, the test produced extreme anxiety among young children who feared failure and being considered “dumb” as a result. Examining the harmful impact of testing on racialized and low-income students, Eizadirad (2020, p. 288) argues that “externally administered standardized tests such as the EQAO function as a tool for stereotyping racialized identities” and outlined how stereotyping in this context functions on three levels: (a) constructing racialized students as low-achievers early in childhood contributes to deficit thinking among children and teachers; (b) perpetuating the stereotype that racialized students have poor intellectual abilities and placing them in separate behavioral and special education classes at an early age; and (c) challenging the stereotype that racialized students are low-achievers “by producing positive achievement results” (Eizadirad, 2020, p. 290). Regarding the latter, Eizadirad (2020, p. 290) explains that:

Although this might seem positive at the surface, the constant effort of having to defend one’s intellectual abilities while navigating predominately White elite spaces embedded with hierarchal power relations saturated with stereotypical assumptions about one’s race, culture, ethnicity, and/or socio-economic status is exhausting and socio-emotionally draining leading to subsequent poor performance, feelings of exclusion and not belonging, and/or triggering identity issues and crisis that can contribute to dropping out of school.

Such fears rooted in the stereotype threat that often accompanies academic failure thus works to harm racialized and low-income students early in their educational lives, “marginalizing and oppressing racialized identities via the interpretations and socially constructed labels created from standardized test results” (Eizadirad, 2020, p. 292). Likewise, in Wasserberg (2017) study of the

stereotype threat among high-achieving African American students, students expressed an active need to not confirm fixed preconceptions that considered Black students as failures; a stereotype that was applied to students given their school's low test scores.

In her ethnographic study of accountability driven reform in Chicago, [Lipman \(2011\)](#) provides significant evidence of the devastating impact of such reforms for African American and Latinx students in urban schools, which has resulted in the closing of schools, displacement and disenfranchisement. [McNeil et al. \(2008\)](#), in their research on a test-based accountability policy in Texas, expose drastic consequences of high dropout rates among students who are poor, English language learners, and come from African American and Latinx backgrounds. Similarly, [Fontaine \(2016, p. 8\)](#) discusses fundamental problems with the narrow framework of accountability and datafication, arguing that the:

Education marketplace disproportionately advantages those with the time, social capital, and institutional knowledge to navigate the system. They benefit the relatively privileged within all racial and social class groups and function to keep middle and upper class families invested in public schools in gentrified areas (Ball et al., 1995). As a result, accountability programs may backfire and have the unintended consequences of reinforcing school segregation as parents eschew neighborhood schools in favor of higher performing schools elsewhere.

As addressed, the implications of standardized testing for racialized and low-income students has been subject to much scholarly critique in Canada while a growing body of literature is also exposing the negative impact of testing and accountability regimes on students' experiences of schooling and teachers' pedagogical practices ([Kearns, 2011, 2016](#); [Lubiensky and Yoon, 2017](#); [Rezai-Rashti and Lingard, 2020](#); [Spencer, 2012](#)). [Lubiensky and Yoon \(2017, p. 6\)](#), researching test-based accountability policies in two Canadian provinces (Ontario and British Columbia), conclude that "school choice has become a middle-class privilege, or is practiced largely by parents who place high value on education as means to further advantage their children" (p. 6). [Rezai-Rashti and Lingard \(2020\)](#), drawing on the perspectives of racialized students in their comparative study of test-based accountability in Canada and Australia, challenged the notion that test-based accountability benefits all students equally. The authors further conclude that test-driven accountability has resulted in further segregation of racialized and poor students in both countries while benefitting those families with social and cultural capital who are able to access the choices offered in the education marketplace.

Analyzing the impact of neoliberal reforms on teachers in Ontario's secondary schools, [Rezai-Rashti \(2009\)](#) found that standardization resulted in a prescriptive curriculum and reduction in teacher autonomy as a result of mandated appraisals of teacher performance. Increasing standardization in Ontario's education system manifested itself through a "results-based curriculum focusing on what students are able to do at the end of the program; and standard discipline oriented (subject) curriculum based on measurable items" ([Majhanovich, 2002, p. 165](#)). Furthermore, [Majhanovich \(2002\)](#) argues that although educational reform, hallmarked by standardized curriculum and high-stakes testing, has been promoted under the banner of excellence and quality enhancement, it has had dire consequences on the de-skilling of teachers, leaving little room for creativity, innovation or autonomy. Similarly, speaking of the use of large-scale assessments, [Duncan \(2011\)](#) detailed the unintended consequences of this for teachers, including constraints on instructional time, higher order thinking skills, ability to cover the curriculum and teacher professionalism. It is not uncommon to hear teachers lament that "more time is spent preparing for tests or testing than is actually spent teaching and that seems counter intuitive" ([Duncan, 2011, p. 83](#)).

In the context of school leadership, [Spencer \(2012, p. 132\)](#) investigated how the provincially mandated system of accountability has "constructed social practices and relations, and how it constituted agents in schools as the subjects of reform." Using governmentality as her approach to empirical data, [Spencer \(2012\)](#) found that high-stakes tests and accountability policies have resulted in significant changes in the practices of school administrators, writing that "administrators have moved into new management roles as, increasingly, there time is devoted to tasks for monitoring, accounting for, and reporting on the administration of policies concerned with performance and outcomes, such as standardized testing" ([Spencer, 2012, p. 132](#)). [Rezai-Rashti and Segeren \(2020\)](#) also argued that the pressure for test-based accountability has resulted in urban school leaders working to game the system and engage in strategies of manipulating data to show school improvement. Their research on the experiences of urban secondary school leaders with the Ontario Secondary School Literacy Test (OSSLT) (a test that all students must pass to obtain their high school diplomas) revealed how they articulated the use of various methods or tactics for manipulating or fabricating test result data to artificially raise a school's reported performance in an effort to be seen as complying with accountability expectations. Participants revealed that it was not an uncommon practice to ask certain students to stay home on the day of the test so that only those students who were likely to pass the test would write it. A literacy coordinator in their research study, who worked directly with schools to improve test results, explained that there are two methods for reporting the results of standardized tests. Method one reports on the results of all eligible students and how many passed. So, if a student was absent or was deferred, this would be reported as a failure, impacting the school's overall score. Method two only reports on students who wrote the test. The school leader further explains that:

there are schools in the city that defer all of the applied level students and so 90 something percent are passing, but that gap between the two scores is large. So, method 2 looks good on paper. So, newspapers report on method 2. Schools report on method 2 because the scores are higher in comparison to method 1 that included all eligible students.

This is similar to the findings of [Ball et al. \(2012\)](#) research which showed a clear distinction between policy implementation and enactment. The authors suggest that it is important to focus attention on how schools deal with “contradictory policy demands, and the diverse ways that they creatively work to fabricate and forge practices out of policy texts and policy ideas in the light of their situational realities” ([Ball et al., 2012](#), p. 142).

As will be detailed further in the next section, the effects of public perception in an age of data-driven decision-making and engineered crises of failing schools and failing students is that the general public is being conditioned to believe that “the implementation of standards-based programs signals students, parents, and society at large that teachers are not to be trusted or respected and that technical/managerial control is what is needed to fix the problems that teachers helped create” ([Duncan, 2011](#), p. 84). But as the existing literature reveals, the overemphasis on accountability through testing has concerning outcomes for racialized and poor students and teacher efforts to effectively reach the educational needs of such students. The implications of this for educational equity must be seriously acknowledged.

Mediatization and standardized testing

One of the most significant but perhaps unintended consequence of standardized testing has been the mediatization of test results and the ranking of schools. Addressing the mediatization of educational performance and the way in which various forms of media play an important role in shaping public perceptions, [Rawolle and Lingard \(2011\)](#) detail the impact that journalism and journalists have in placing issues on the education agenda. Indeed, journalists identify problems and pose solutions, which serve to deeply shape and inform public perceptions of the education system. This was well documented in [Parker \(2017\)](#) historical critical policy analysis, where she investigates how neoliberal education policies were “sold” to the public. Through educational reforms, intended to resolve the educational crisis that Ontario students were falling behind their international counterparts, Parker demonstrates how the news media highlighted the need for educational policy changes that favored neoliberal reforms. Documenting the publicization of EQAO test results, [Parker \(2017, p. 53\)](#) found that the “news media saturate discussions of whether education policy needs an overhaul based on the latest test scores.”

Research conducted by [Rezai-Rashti and Segeren \(2020\)](#) shows how urban school leaders in Ontario and British Columbia understand and critique the role of the media in reporting and ranking schools, which often results in stigmatization of racialized and low-income students, their schools, as well as their communities. For example, one of the school principals in their research implicated the Fraser Institute, a conservative think tank that reports on the performance of all schools across four provinces in Canada. The principle noted the impact of the institute’s annual report on test scores, stating: “when these Fraser reports and media reports come out and bash certain schools, it does stigmatize places. And the reputation of buildings become affected. Those stigmatizing reports don’t take the pulse of the whole school.”

Indeed, this public shaming has contributed to stigmatizing schools in low-income, racially diverse neighborhoods where test results are often lower. [Rezai-Rashti and Segeren \(2020\)](#) suggest that the publication of comparative performance data has enabled middle-class families with above-average levels of capital to opt out of their local, under-subscribed schools and choose higher performing ones instead, a trend identified in two provincial contexts. Another school principal in British Columbia explained that media reports have impacted the reputation of his school, noting that it is a contributing factor to low enrollment rates and the flight of middle-class families to other schools in the area. [Martino and Rezai-Rashti \(2013\)](#), in their analysis of the media and the discourse of the gender achievement gap and failing boys, which were based on PISA test scores, provide important insight into the significant role of the media within the network of policymaking and practice. The authors write that the media “constitute a particular policy habitus, particularly in relation to the endorsement of a gender achievement gap and the mobilization of discourses about failing boys in the Canadian context” ([Martino and Rezai-Rashti, 2013](#), p. 604).

The publication of student performance data and public shaming of low performing schools has unintended consequences for schools that are spending increasing amounts of time and effort engaged in entrepreneur-like activities. Reflecting trends noted in the literature associated with the new public management of schools, principals and guidance counselors reported trying to “sell” or “market” their schools to attract students and families. This trend is especially evident in urban schools or low performing schools at risk of closure, and school leaders are the ones tasked with recruiting new students, as [Rezai-Rashti and Segeren \(2020\)](#) found. Interestingly, school leaders are now submerged in business related activities, such as marketing their school by “selling” special “features” of their schools through extra-curricular activities or specialized academic programming. As the head of the guidance department in an urban school explained: “I feel a lot of pressure to go out to the feeder schools and to market our school and to promote our school, because we’re all competing for the same students downtown”. Part of her marketing campaign involved crafting a more robust portrait of her school to bolster its public reputation:

I’m always on Twitter and Facebook reposting positive news stories and posting pictures of all the great things happening here. There is pressure to be in control of your own story as a school, because parents and the community are using technology. We’re trying to update our website all the time, make it more student friendly with lots of photos and things like that, just to be relevant and recent.

Rezai-Rashti and Segeren (2020).

While school choice appears to be desirable and increasingly available, research shows that those exercising this choice are predominantly middle-class parents. School choice, however, works to the disadvantage of students and families from racialized and low-income backgrounds who lack the social and cultural capital of their middle-class counterparts to take advantage of a wide array of schooling options (Rezai-Rashti and Lingard, 2020; Yoon et al., 2017).

It is important to mention that the negative role of the media has been acknowledged by academics and more recently by national and provincial governments. For example, in 2017, the Ontario Ministry of Education appointed two advisors to guide and transform assessment and reporting practices by acknowledging the significance of equity and Indigenous education. The report predominantly focused on misuse of performance data and ranking of schools. Their report listed 18 recommendations to transform the culture of assessment in Ontario in order to achieve the goal of equity, well-being and enhancement of public confidence in Ontario's publicly funded education. Together, these recommendations sought to connect assessment to curriculum, identify and respond to inappropriate misuse of provincial large-scale assessment data (EQAO) and oppose the public ranking of schools. The advisory committee raised concerns over the current nature of provincial assessments in the context of Ontario's commitment to equity and cultural diversity, especially the Grade 10 Literacy Test (OSSLT), which the committee agreed should not be linked to the high school graduation requirement. The recommendations of the advisory group included reducing the number of tests by eliminating standardized tests in several grades as well as creating a mechanism to prevent "inappropriate use and misuse of provincial large-scale assessment data by other organizations/individuals, including opposing public ranking of schools" (Campbell et al., 2018, p. 70).

The literature on high-stakes testing and accountability in several other countries such as the UK and Australia also shows the impact of national testing and its consequences for the promotion of comparison data and an audit culture. Lingard and Sellar (2013, p. 652) discuss how collecting performance data and reporting through the "My School" Website is causing reputational damage to schools, arguing that the reporting of Australia's National Assessment Program-Literacy and Numeracy (NAPLAN):

has become high stakes for systems through reputational damage caused by perception of poor performance ... the pressure and nervousness felt by policy makers and politicians in response to NAPLAN outcomes are more immediately motivated by concerns to improve or maintain the reputation of schools and systems and to secure funding.

Gillborn (2008, p. 65) work on the "gap talk" in the UK also addresses how the government is using performance data as justification for educational improvements, suggesting that "closing and/or 'narrowing' gaps operates as a discursive strategy whereby statistical data deployed to construct the view that things are improving, and the system is moving in the right direction."

In contrast to many of the countries discussed in this review, Finland has an exceptional trajectory of not engaging with neoliberal accountability while being one of the top performers in the international PISA test. Writing about Finnish distinctiveness, Simola et al. (2013, p. 623) discuss the rise of Finland as an exemplary model through incremental policy rather than through neoliberal declaration, writing that "Finland did not follow the Anglo-Saxon accountability movement in education" as teachers and schools are the ones who are tasked with evaluating students. In Finland, with a long history of no national testing (with the exception of a high school matriculation examination), politicians and educational officials have been consistently rejecting the notion of school ranking based on performance indicators. In 2005, the Finish Education Evaluation Council formally stated their opposition to ranking and mediatization of individual school results:

In publicising evaluation results schools will not be ranked, nor will schools or teachers will be labeled as of high or low standard on the basis of one-sided evidence. When reporting upon an analysis based on a nationwide sample, no data identifying individual schools will be given
cited in Simola et al. (2013, p. 623).

In nations where neoliberalism is increasingly shaping the push toward the marketization of education, it remains to be seen to what extent the mediatization of test scores and rankings will shift, and perhaps even undermine, the democratic foundations of public education. As trends mentioned in this section reveal, the implications for educational equity are vast.

Conclusion

The role of performance-based accountability cannot be understood without carefully investigating national testing regimes, key actors, power dynamics, as well as the inflecting role of international organizations such as the OECD in global governance. In this paper, we emphasized the influence of international organizations in the collection of comparative data and their significant role in steering education policy around the globe. We believe such comparative performance data is decontextualized with limited focus on local and national contexts of schooling.

We also focused on a select number of studies which have explored the impact of test-based accountability on racialized students and students from low-income families. Our main objectives have been to show how change to test-based accountability nationally and globally have had serious consequences for racialized and poor students. Situated within the context of neoliberal rationality

and relying on discourses of meritocracy and self-responsibilizing individuals, we believe that performance-based accountability has resulted in focusing too much on individual students and teachers while leaving broader and more significant structural and institutional inequities invisible. Our research raises important questions regarding the winners and losers of neoliberal educational reforms and whether these forms of top-down accountability measures and assessments are the right approach in the first place. We believe, as argued by Shore and Wright (2015a, p. 27), that “the construction of easy-to-read, decontextualized numbers, however faulty, also makes them popular and a useful tool of management and governance.” However, these rankings and numbers do not impact all individuals equally and “audit culture is based on a zero-sum game: to succeed in this game, someone has to lose. This is the basis of the most corrosive effects of this form of governance” (Shore and Wright, 2015b, p. 432).

Acknowledgment

This research was financially supported by the Social Sciences and Humanities Research Council of Canada, award no. 435-2014-1014.

References

- Apple, M.W., 2005. Are markets in education democratic? Neoliberal globalism, vouchers, and the politics of choice. In: Singh, M., Kenway, J., Apple, M.W. (Eds.), *Globalizing Education: Policies, Pedagogies, & Politics*. Peter Lang Publishing, New York, pp. 209–230.
- Au, W., 2009. *Unequal by Design: High-Stakes Testing and the Standardization of Inequality*. Routledge, New York.
- Au, W., 2013. Hiding behind high-stakes testing: meritocracy, objectivity and inequality in U.S. education. *Int. J.* 12, 7–19.
- Au, W., 2016. Meritocracy 2.0: high stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educ. Pol.* 30, 39–62.
- Ball, S., Maguire, M., Braun, A., 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. Routledge, London, New York.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18, 215–228.
- Beach, C., Green, A., Worwick, C., 2006. Impact of Point System and Immigration Policy Levers on Skill Characteristics of Canadian Immigrants. Queen's University Economics Department, Kingston.
- Brunn-Bevel, R.J., Byrd, W.C., 2015. The foundation of racial disparities in the standardized testing era: the impact of school segregation and the assault on public education in Virginia. *Humanit. Soc.* 39, 419–448.
- Campbell, K., Fullan, M., Hargreaves, A., James, C., Longboat, K.D., 2018. *Ontario: A Learning Province, Findings and Recommendations From the Independent Review of Assessment and Reporting*. Ontario Ministry of Education.
- Duncan, C., 2011. Re/Viewing student successes in an era of accountability. *J. Unsch. Altern. Learn.* 59, 58–99.
- Eizadirad, A., 2020. External assessment of stereotyping: experiences of racialized grade 3 children, parents and educators with standardized testing in elementary schools. *Rev. Educ. Pedagog. Cult. Stud.* 42, 277–295.
- Fontaine, C., 2016. The myth of accountability: how data (mis)use is reinforcing the problem of public education. *Data Soc.* 1–13. www.datasociety.net.
- Gillborn, D., 2008. *Racism and Education: Coincidence or Conspiracy?* Routledge, London.
- Goldberg, D., 2009. *The Threat of Race: Reflections on Racial Neoliberalism*. Wiley Blackwell, Malden.
- Greks, S., 2013. Expert moves: international comparative testing and the rise of expertocracy. *J. Educ. Pol.* 28, 695–709.
- Henry, M., Lingard, B., Rizvi, F., Taylor, S., 2001. *The OECD, Globalization and Education Policy*. Pergamon, Oxford.
- Hursh, D., 2013. Raising the stakes: high-stakes testing and the attack on public education in New York. *J. Educ. Pol.* 28, 574–588.
- Jakobi, A., Martens, K., 2010. Expanding and intensifying governance: the OECD in education policy. In: Martens, K., Jakobi, A.P. (Eds.), *Mechanisms of OECD Governance: International Incentives for National Policy-Making?* Oxford University Press, Oxford, pp. 163–176.
- Kearns, L.L., 2011. High-stakes standardized testing and marginalized youth: an examination of the impact of those who fail. *Can. J. Educ.* 34, 112–130.
- Kearns, L.L., 2016. The construction of “illiterate” and “literate” youth: the effects of high-stakes standardized testing. *Race Ethn. Educ.* 19, 121–140.
- Knoester, M., Au, W., 2017. Standardized testing and school segregation: like tinder for fire? *Race Ethn. Educ.* 20, 1–14.
- Leuze, K., Martens, K., Rusconi, A., 2007. New arenas of education governance: the impact of international organizations and markets on education policy making. In: Martens, K., Rusconi, A., Leuze, K. (Eds.), *New Arenas of Education Governance—The Impact of International Organizations and Markets on Education Policy Making*. Palgrave Macmillan, New York, pp. 3–15.
- Lewis, S., 2020. *PISA, Policy and the OECD: Respatializing Global Educational Governance Through PISA for Schools*. Springer, Singapore.
- Lingard, B., Sellar, S., 2013. “Catalyst data”: perverse systemic effects of audit and de-skilling of education in Australian schooling. *J. Educ. Pol.* 28, 634–656.
- Lingard, B., Martino, W., Rezai-Rashti, G., 2013. Testing regimes, accountabilities and education policy: commensurate global and national developments. *J. Educ. Pol.* 28, 539–556.
- Lingard, B., Martino, W., Rezai-Rashti, G., Sellar, S., 2016. *Globalizing Educational Accountabilities*. Routledge, New York, London.
- Lingard, B., 2011. Policy as numbers: ac/counting for educational research. *Aust. Educ. Res.* 38, 355–382.
- Lipman, P., 2013. Economic crisis, accountability, and the state's coercive assault on public education. *J. Educ. Pol.* 20, 557–573.
- Lipman, P., 2011. *The new political economy of urban education: neoliberalism, race, and the right to the city*. Routledge.
- Lubiensky, C., Yoon, E., 2017. Introduction to special issue: studying school choice in Canada. *Educ. Pol. Anal. Arch.* 25, 1–10.
- Majhanovich, S., 2002. Conflicting visions, competing expectations: control and de-skilling of education—a perspective from Ontario. *McGill J. Educ.* 37, 159–176.
- Marmol, E., 2016. The undemocratic effects and underlying racism of standardized testing in the United States. *Crit. Intersect. Educ.* 4, 1–9.
- Martino, W., Rezai-Rashti, G., 2013. “Gap talk” and the global rescaling of educational accountability in Canada. *J. Educ. Pol.* 28, 589–611.
- McNeil, L.M., Coppola, E., Radigan, J., Heilig, J.V., 2008. Avoidable losses: high-stakes accountability and the dropout crisis. *Educ. Policy Anal. Arch.* 16 (3), 1–45.
- Mehta, J.D., Schwartz, R.B., 2011. Canada looks a lot like us but gets much better results. In: Tucker, M.S. (Ed.), *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems*. Harvard Education Press, Cambridge, pp. 141–165.
- Moreno-Salto, I., Robertson, S., 2021. On the “life of numbers” in governing Mexico's education system: a multi-scalar account of OECD's PISA. *Glob. Soc. Educ.* 19, 213–227.
- Nicholas, S., Glass, J., Berliner, D., 2006. High-stakes testing and student achievement: does accountability pressure increase student learning? *Educ. Pol. Anal. Arch.* 14, 1–56.
- Novoa, A., Yariv-Marshall, T., 2003. Comparative research in education: a mode of governance or historical journey? *Comp. Educ.* 39, 423–438.
- OECD, 2010. *Programme for International Student Assessment (PISA): Strong Performers and Successful Reformers in Education*. www.oecd.org.
- Ozga, J., 2009. Governing education through data in England: from regulation to self-evaluation. *J. Educ. Pol.* 24, 149–162.
- Parker, L., 2017. Creating crisis: selling neoliberal policy through the rebranding of education. *Can. J. Educ. Adm. Pol.* 183, 44–60.
- Rawolle, S., Lingard, B., 2011. New scalar politics: implications for education policy. *Comp. Educ.* 47, 489–502.

- Rezai-Rashti, G., Lingard, B., 2020. Test-based accountability, standardized testing and racialized students' perspectives in urban schools in Canada and Australia. *Discourse* 1–16.
- Rezai-Rashti, G., Segeren, A., 2020. The game of accountability: the perspectives of urban school leaders in Ontario and British Columbia, Canada. *Int. J. Leader. Educ.* 1–18.
- Rezai-Rashti, G., 2009. The assault of neo-liberalism on education reform, restructuring, and teaching in Ontario secondary schools. In: Levine-Rasky, C. (Ed.), *Canadian Perspectives on the Sociology of Education*. Oxford University Press, Toronto, pp. 307–321.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Sharman, J.C., 2012. Seeing like OECD on tax. *New Polit. Econ.* 17, 17–33.
- Shore, C., Wright, S., 2015a. Governing by numbers: audit culture, ranking and the new world order. *Soc. Anthropol.* 23, 22–28.
- Shore, C., Wright, S., 2015b. Audit culture revisited: rankings, ratings, and the reassembling of society. *Curr. Anthropol.* 56, 421–431.
- Simola, H., Rinne, R., Varjo, J., Kauko, J., 2013. The paradox of the education race: how to win the ranking game by sailing the headwind. *J. Educ. Pol.* 28, 612–633.
- Sorensen, B., Ydesen, C., Robertson, S.L., 2021. Re-reading the OECD and education: the emergence of a global governing complex-an introduction. *Glob. Soc. Educ.* 19, 99–107.
- Spencer, B.L., 2012. Counting in, counting out, and counting for: administrative practices of standardized literacy testing. In: Spencer, B.L., Gariepy, K.D., Dehli, K., Ryan, J. (Eds.), *Canadian Education: Governing Practices and Producing Subjects*. Sense Publishers, Rotterdam, pp. 131–144.
- Thompson, G.L., Allen, T.G., 2012. Four effects of the high-stakes testing movement on African American K-12 students. *J. Negro Educ.* 81, 218–227.
- Toronto District School Board, 2006. *Research Report: The TDSB Grade 9 Cohort Study: A Five-Year Analysis, 2000–2005*. Toronto District School Board, Toronto.
- Wasserberg, M.J., 2017. High-achieving African-American elementary students' perspectives on standardized testing and stereotypes. *J. Negro Educ.* 86, 40–51.
- Yoon, E., Lubiensky, C., Lee, J., 2017. The geography of school choice in a city with growing inequality: the case of Vancouver. *J. Educ. Pol.* 33, 279–298.

International schools: shifting sands and winds of change

Mary Hayden^a and Jeff Thompson^b, ^a Mary Hayden Professor of International Education at the University of Bath, Bath, United Kingdom; and ^b Jeff Thompson Emeritus Professor of Education at the University of Bath, Bath, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	336
Early days	336
International education	337
International curriculum	338
International schools: developments and contemporary issues	338
Terminology and definitions	340
Nature of education	340
Going forward	341
References	342

Introduction

It is undoubtedly the case that education in its many forms has undergone notable change since the first edition of this encyclopedia was published in the 1980s. There can have been few changes more marked, however, than those experienced within the group of schools known as 'international schools'. In this chapter, we begin by setting the scene with a short overview of the development of this area from its earliest days, before showing how an initially steady increase in numbers was superseded by a massive unprecedented growth in the closing years of the 20th century. The remainder of the chapter will focus on a number of issues and challenges arising in what will be described as the international school sector, as it has not only grown substantially in size but has also markedly changed in nature as the environment in which it is located, and the expectations made of it, have themselves undergone major changes.

Early days

One feature bedevilling international schools (a term used here to encompass all schools worldwide that might self-identify or be considered as such) and challenging those who write about, study, work in, or research them is the lack of a clear and agreed lexicon. Frustratingly, as the sector has grown and broadened in scope, the terminology has if anything become more rather than less complicated, as new developments have introduced new terminology – in some cases to describe concepts very similar to those that already exist, though with different labels for what are essentially variants on an underlying theme. Discussion about such variations has been described by Marshall as 'the big terminology debate' (2007: 38), and since Marshall's observation there has been no reduction in the complexity or lack of clarity which, if anything, has continued to increase.

In the absence of a universally accepted definition of the term 'international school', it could be argued that the origins of the concept of schooling outwith national borders can be traced back to the nineteenth century. Maseru English Medium Primary School in the then British Crown colony of Basutoland (since 1966 the Kingdom of Lesotho), for instance, was established in 1890 to cater for the children of English-speaking missionaries, traders and officials of the British administration (MEMPS, 1990), while the International College at Spring Grove, West London, UK, was founded in 1866 by such luminaries as Thomas Huxley, Richard Cobden, John Tyndall and the novelist Charles Dickens – all advocates of free trade, whose vision of a number of international schools in Europe was based on students from different countries being educated 'among school fellows of all nations ... the method of study being precisely the same in each international school' (Dickens, 1864 in Sylvester, 2002: 8), with the students thus becoming 'international ambassadors' (Stewart, 1972: 5). Sylvester (2002) provides a more detailed analysis of developments leading to the establishment of this, arguably, 'first' international school. Twin rationales for what might be thought of as international schools were therefore evident from the earliest days: the *pragmatic* (offering education to children displaced from their home education system due to a parent's global mobility) and the more *ideological* (intentionally bringing together students from different national/cultural contexts with a view to them learning within a broader context than the national borders within which education would more usually be prescribed).

More commonly described as the earliest international schools are the International School of Geneva, Switzerland and Yokohama International School, Japan, both established in 1924 and with the former indeed claiming to be the 'first international school' (Walker, 1996). Though with different provenances from MEMPS or the International College at Spring Grove, clear in both schools is a pragmatic rationale for their establishment: to cater for the children of globally mobile or otherwise temporarily displaced professional parents who wished their children to accompany them rather than to 'board' in the home country, and for whom other schooling available locally was deemed inappropriate – while overlapping with a post-Great War ideological vision of

encouraging the broadening of horizons, development of intercultural understanding and preparation of adults equipped for effective and peaceful interaction with those from other backgrounds than themselves (see Knight, 1999; Stanworth, 1998). With increasing global mobility emerged growing numbers of international schools worldwide, often established by parents themselves: as, for instance, in the case of The Alice Smith School in Kuala Lumpur, Malaysia which began in 1946 as the eponymous Alice Smith taught her own young daughter, and subsequently growing numbers of children of friends and colleagues (Alice Smith School, 1996). Based on a different rationale, post-World War 2 concerns that such horrors should never be experienced again were central to the establishment, inspired by the visionary Kurt Hahn, of Atlantic College in south Wales (Peterson, 1987), founded in 1962 with the express purpose of bringing together students of 16–18 years of age from around the world, on a scholarship basis, to study and live together. Having something in common then with Spring Grove in being founded expressly with an ideological purpose, Atlantic College was the progenitor of the worldwide United World College (UWC) movement which, at the time of writing, has 18 member schools and colleges worldwide, all espousing the philosophy that ‘UWC makes education a force to unite people, nations and cultures for peace and a sustainable future’ (UWC, 2021) and embodying a response to the question asked in his 1957 Nobel Prize acceptance speech by Lester B Pearson, then Prime Minister of Canada, and after whom the UWC in Vancouver is named: ‘How can there be peace without people understanding each other, and how can this be if they don’t know each other?’ (Lester B Pearson College of the Pacific, 1982: 9).

Together with a growth in numbers of United World Colleges in the following half century, until relatively late in the twentieth century a steady increase was seen in numbers of individually established international schools established principally to cater for growing numbers of globally mobile professionals in a particular location. Though in the absence of a universal definition of the term, exact numbers of international schools cannot be cited, growth has nevertheless been evident over a number of years. (See, for instance, Hayden, 2011). While having an essentially pragmatic *raison d’être*, such schools would also very often promote a more ideological form of education in preparing the adults of tomorrow for an increasingly globalised world. In arguably the first piece of research-based academic writing with a focus on international schools, in 1995 Hayden and Thompson observed that, at that point, the international school sector could be characterised as a ‘conglomeration of individual institutions which may or may not share an underlying educational philosophy’ (1995). Reflecting later on what they described as the ‘globalist’ and ‘internationalist’ approaches to education, where the former relates to serving a market that requires the global certification of educational qualifications with portability between schools and transferability between systems, while the latter is underpinned by an existential, experiential philosophy of education that values the moral development of the individual and the development of a sense of responsible citizenship, Cambridge and Thompson (2004) point out that the form of education practised in international schools is (or was at that time) the ‘reconciliation of these contrasting approaches’, with each reconciliation being unique to the individual institution. We shall return to more recent developments in the sector after exploring first a number of related concepts.

International education

If the terminology surrounding international schools is unclear, that relating to the concept of international education is no clearer. In the first edition of this encyclopedia, Husén suggested that international education can be considered as including ‘all educative efforts that aim at fostering an international orientation in knowledge and attitudes’ (1985: 2660). In the same year, the Harvard Educational Review’s Special Issue on international education argued in its editorial preface that ‘International, global, cross-cultural and comparative education are different terms used to describe education which attempts ... to come to terms with the increasing interdependence that we face and to consider its relationship to learning’ (Fasheh, 1985). More recently, as interest has grown in education across national borders, the term ‘international education’ has come to be used in a range of contexts. It may be encountered, for instance, in debate relating to the higher education (university level) sector, where the internationalisation of higher education is the focus of increasing research interest as, in some if not all parts of the world, universities accept increasing proportions of ‘international’ or ‘overseas’ students, appoint growing numbers of academic staff from around the world, engage in research across national boundaries and establish ‘outposts’ in other parts of the world than their home country. The research base for this form of international education, or ‘transnational education’, is steadily broadening (see, for instance, Clifford and Montgomery, 2017).

At school level, in addition to the international schools on which this chapter largely focuses, a recent phenomenon in some countries has been increasing interest in internationalisation of the school experience. In some contexts this may take the form of children from one country attending a school in the national system of another country (in residential/boarding schools in England, for instance). In other contexts, it may arise from growing recognition that education for any child, whatever their background, needs to prepare them for a future that will not be entirely nationally-focused. Whether or not a child growing up in, say, the UK aspires to adulthood as a diplomat, or a career with a multinational company, it is likely that their future life – even if they never travel beyond their home borders – will be heavily influenced by events beyond those borders. While much of the food they eat, goods they purchase and entertainment in which they engage is likely to emanate from other parts of the world, it is also the case that global events and geopolitics will undoubtedly impact in some way upon their future lives. Recognition of the importance of preparing them, as far as possible, for that future has led to initiatives including, in England for instance, schools supplementing the statutory national curriculum with Oxfam materials (2015) that provide a guide for teachers on global citizenship, and projects including the UK’s Global Learning Programme (2021).

The area of education where, perhaps, the most widely recognised work has been done on internationalising the educational experience at school level is that relating to the formal curriculum, where major strides have been made in the past half century that are of direct relevance to international schools.

International curriculum

As the international school sector has experienced rapid growth and become increasingly recognised in recent years, so too has the concept of a formal curriculum that goes beyond curricula offered in national contexts that essentially prepare young people (notwithstanding more recent developments such as those described above) for a future in that national context. The first curriculum developed specifically to cater for the needs of international schools is generally accepted to be the International Baccalaureate (IB) Diploma Programme, first examined in 1970 and developed by a small number of schools with input from the International Schools Association and from experts in national contexts including the USA, UK and France following developments spearheaded in the first instance by the International School of Geneva (Peterson, 1987). Prompted by the pragmatic need of schools preparing 16–19 year old students who aspired to university entrance in a range of different national systems for a curriculum that could be studied by all and would be recognised worldwide, the IB Diploma has since been joined by its sibling Primary Years Programme (PYP), Middle Years Programme (MYP) and Career-related Programme (CP) (IB, 2021a), thus offering a suite of programmes for which schools can apply for authorisation. Though IB programmes may also be offered within national systems of education, they are found in many international schools worldwide, facilitating as they do the transition of globally mobile students between international schools and countries, while at the same time offering a recognised form of education that not only satisfies pragmatic needs but also supports an ideological approach to education in encouraging students to develop international mindedness (Barratt Hacking et al., 2018) as well as a number of attributes including being open-minded, principled and caring, as summarised in the IB Learner Profile (IB, 2021b).

Not all international schools, however, offer all or any of the IB programmes. While the IB is undoubtedly a major actor in the international school sector, so too is the widely recognised Cambridge Assessment International Education. Cambridge's International General Certificate of Secondary Education (IGCSE), for instance, is offered in large numbers of international schools worldwide, with other programmes also offered across the age-range (Cambridge Assessment International Education, 2021). More recently joining the international school scene are programmes developed by Fieldwork Education, whose International Primary Curriculum (IPC) was launched in 2000 and has since been joined by the International Middle Years Curriculum (IMYC) and International Early Years Curriculum (IEYC) (Fieldwork Education, 2021). It should be noted, however, that not all international schools elect to offer what might be described as an international curriculum. Depending on the nature of the international school (about which more below), it may choose rather to offer the curriculum of a 'home country' or, indeed, its own school-based curriculum. It is not uncommon, for instance, for British-type international schools to offer the national curriculum of England, across part or all of the age range, in conjunction at pre-university level with either A levels from England or, in some cases, the IB Diploma. International schools with other national affiliations may similarly offer curricula from the home country, possibly in combination with an international curriculum for part of the age range.

Thompson's (1998) categorisation of international curricula offered in international schools proposed four categories which continue to be helpful today, as follows:

- Exportation: the marketing abroad of existing national curricula and examinations, with little if any adjustment to take account of the different context and a value system 'unapologetically that of the country from which it is exported'
- Adaptation: where existing national curricula and examinations are adapted for the national context, with the 'inherent value system' not likely to change at all and the risk of, as Thompson puts it, an 'unwitting process of educational imperialism'
- Integration: where 'best practices' from a range of 'successful' curricula are brought together into one curriculum for operation across a number of systems or countries (with attendant challenges potentially to be faced from the different values and ideological positions in question)
- Creation: the development of a programme 'from first principles'

Issues relating to curriculum in the international school context will be returned to later.

International schools: developments and contemporary issues

In the quarter century and more since Hayden and Thompson described international schools as a 'conglomeration of individual institutions which may or may not share an underlying educational philosophy' (1995), much has changed. Massive geopolitical changes in the global environment and increasing forces of globalisation manifested in a 'growing magnitude or intensity of global flows such that states and societies become increasingly enmeshed in worldwide systems and networks of interaction' (Held et al., 2000:3), have been majorly affected by, and have influenced, international schools. Coulby and Zambeta argue that 'Educational institutions themselves are part of the process of globalisation because of their central role in the development of the knowledge economy' (2005: 1), and there can be little doubt that this assertion applies to international schools to at least the same extent as it does to other educational institutions. From their earliest days as largely pragmatic responses to a practical need in specific locations

where globally mobile expatriate families were found, which indeed describes the context in 1995 as noted by Hayden and Thompson, international school numbers have massively increased and, with that growth, has grown the diversity that characterises the sector today. A quarter of a century ago, few children attending international schools were not globally mobile expatriates. Indeed it was not uncommon at that time for 'host country national' students to be prohibited by law from attending such schools, as was the case in, for instance, China, Malaysia, Singapore and Thailand. Since that time, the perceived pragmatic benefits of an international school education – offering a curriculum and end-of-school qualification recognised by universities worldwide, and taught through the medium of English, to fluency in which many non-first language English speakers aspire for their children – have persuaded affluent and aspiring middle class parents in some countries of the competitive edge that such an education will provide for their children. Some, though not all, countries where host country nationals were previously forbidden from attending such schools have since deregulated. Thailand is a case in point where, following relaxation of the law prohibiting Thai students from attending international schools, numbers of international schools saw a steady growth with many new schools catering for the 'home' rather than expatriate market (MacDonald, 2006).

Of particular interest in this respect, given the implications for growth, is the case of China, which has retained its earlier prohibition of Chinese nationals (other than those of dual or multiple nationality) attending previously existing international schools (or 'schools for the children of foreign workers'), and has more recently begun the development of a new type of international school intended for Chinese nationals rather than expatriates, which show every sign of experiencing large growth in the coming years: what Poole (2021) describes as 'Chinese internationalised schools'. In addition to growth across Asia (Machin, 2017), other areas of notable growth include the United Arab Emirates (see, for instance, Azzam, 2021) – again, where international schools are catering not only for the globally mobile but also for the local population. As just some examples of growth in the international school sector, it is clear that, while the numbers of what might be described as 'traditional' international schools, catering for globally mobile expatriates, may have seen some growth, the more marked growth has been in those schools established principally to cater for the 'host country' national families who aspire to their children attending university, and possibly developing a career, in another (often western, anglophone) country. Indeed such has been the shift in the balance of these two types of school in the international school sector over recent years that it has been estimated that, between 1983 and 2013, the ratio of 80% expatriate: 20% host country national students attending international schools worldwide completely reversed to 20%: 80%.

Issues of terminology continue to complicate the international school debate, with no one body having the authority to determine whether a school may or may not be described as an international school. It is thus, in the absence of a universally-accepted definition, impossible to chart definitive growth of international schools over the decades. In approximate terms, however, it has been estimated that there were around 50 international schools worldwide in 1964 (Bereday and Lauwerys, 1964) and 1000 by 1989 (Matthews, 1989). More recently, ISC Research – which curates a global database that monitors activity in the international school sector and uses a very specific definition of international schools, for their own market intelligence purposes, of schools delivering a curriculum at any age range wholly or partly in English outside an English-speaking country, or that offer an English-medium curriculum other than the country's national curriculum and are international in their orientation, if based in a country where English is one of the official languages (ISC Research, 2021a) – record 12,373 international schools as at July 2021, from a total of 7655 in July 2011 (ISC Research, 2021b).

Many attempts have been made over the years to bring order to the large number of international schools in terms of types and categories, by authors including, *inter alia*, Matthews (1989) and Sylvester (1998), all of them developed before the more recent shift toward international schools catering for host country nationals, and many of them challenged by the fact that their proposals began to date rapidly almost as soon as they were disseminated. (See Hayden, 2006 for details of other categorisations). Our own more recent categorisation (Hayden and Thompson, 2013) proposed that the rapidly-changing sector could be thought of as including three main 'types' of international school, described as follows:

- 'Type A' 'traditional' international schools: established principally to cater for globally mobile expatriate families for whom the local education system is not considered appropriate
- 'Type B' 'ideological' international schools: established principally on an ideological basis, bringing together young people from different parts of the world to be educated together with a view to promoting global peace and understanding
- 'Type C' 'non-traditional' international schools: established principally to cater for 'host country nationals': the socio-economically advantaged elite of the host country who seek for their children a form of education different from, and perceived to be of higher quality than, that available in the national education system

It has, however, been clear for some time that this typology – relatively recent though it is – is in need of updating. While Types A and B may still reasonably represent schools today, Type C has – since the term was coined – become more complex as the part of the international school sector catering principally for host country national students has grown and diversified. Recent marked growth in this area, for instance, includes what may be considered exports of elite private schools from England including, among others, Dulwich, Harrow and Shrewsbury (see Bunnell, 2008) together with the Chinese 'internationalised' schools referred to above. Unlike the not-for-profit nature of many Type A schools, where surplus income is ploughed back into developing the school rather than into profits for owners, shareholders and/or investors, notable among the characteristics of many, though not all, schools that could be included in 'Type C' is that they are for-profit and have been established on a more commercial basis than traditionally has been the case for international schools. Notable too as a major change in the international school sector in the last quarter century has been the shift from the 'conglomeration of individual institutions', observed by Hayden and Thompson (1995: 332) as noted above, to the creation of *groups* of international schools established on a for-profit basis. Previously unusual,

other than in the case of a small number of groupings such as the European Schools (Gray, 2003) or Shell Schools (see Hayden, 2006), groups of international schools may now be found, often catering essentially for the local market, within national contexts including China and Vietnam, but also cutting across national borders. Some well-established, though still relatively recent, groupings of international schools worldwide that may cater for both local students and the more globally mobile include, *inter alia*, the GEMS group, Nord Anglia and Cognita. And while there may be a lack of clarity with respect to the meaning of the term 'international school', means of judging the 'quality' of particular international schools have grown up in the form of accrediting agencies, including among others the Council of International Schools (CIS), which accredits schools where evidence is found of achievement of pre-specified standards, thus providing quality assurance in the sector.

In the context of the current growth in the international school sector, and the likelihood of that growth continuing, a number of issues and challenges arise, as discussed in the following sections.

Terminology and definitions

If terminology in this context was bedevilled by lack of clarity when Marshall highlighted the issue in 2007, how much more is that now the case as the sector grows and diversifies, with schools having different reasons for including 'international' in their name. Undoubtedly in some cases a marketing strategy, in other cases what might to a casual observer appear to be an international school may choose not to be described as such. Others make clear their affiliation with a particular national context through a title such as 'British International School of X' or 'American School of Y', while others eschew any national affiliation and offer an international curriculum throughout the school, to students of many different nationalities, and cultural and linguistic backgrounds. Whether a lack of consistency or clarity in terminology is a cause for concern may be debatable. Some would argue that it is (see Azzam, 2021); we share a concern that, at the very least, the lack of consistency should be recognised and acknowledged.

Nature of education

In the earliest days of educating away from their home country young people who were expected to return in adulthood to that context, the form of education offered might not have necessitated much in the way of debate. If the purpose of education was to prepare young people for adult life in their 'home' context, whether their childhood was spent there or elsewhere, then the traditional concept of curriculum – as exemplified for instance in Lawton's concept of curriculum as a selection from the culture of a society (1989) – would apply equally to those temporarily displaced from that society. American expatriate students for instance could follow a US-style curriculum in a school outside the US, in preparation for their return. Complexity emerged with the growth of international schools catering for students from multiple national and cultural backgrounds, whose teachers were faced with preparing students at the pre-university level for return to a 'home' university and for the qualifications which would demonstrate their preparedness to embark upon university education in that context. Challenges presented by teaching the subject of history, as one notable example, to a class of students needing to prepare for national qualifications with different curricula and examinations were, as noted earlier and discussed in more detail elsewhere (see, eg. Peterson, 1987) instrumental in leading to the creation and development of the International Baccalaureate (IB) Diploma as a curriculum that could be studied by all students regardless of destination, in the anticipation of acceptance as a university entrance qualification worldwide.

As the demand for international programmes is growing, however, so too are questions as to what it means for a form of education to claim to be 'international'. Is it sufficient that a curriculum is effectively a form of educational 'home away from home'? And how should the 'ideological' theme running through many such programmes (claiming to develop in young people attributes such as international mindedness, global competence, and respect for others different from themselves) relate to the more 'pragmatic' theme of preparing young people for life in an increasingly globalised world, where attributes and skills including multilingualism and intercultural literacy (Heyward, 2002) will be at a premium in the global knowledge economy (Lauder et al., 2012) of the 21st century?

As noted earlier, many different programmes are now offered in international schools worldwide. And as Thompson has highlighted, they may have been designed intentionally to be international, or be exported or adapted from a national context while, as Thompson puts it, risking an 'unwitting process of educational imperialism' (1998). Those that claim explicitly to be 'international' rather than arising from the context of any one national education system have in so doing laid themselves open to challenge. For some time there have been claims that the IB programmes, for instance, are 'Eurocentric' (not least because of the Diploma's three main working languages of English, French and Spanish). More recently, questions have been raised about the meaning of 'international' in this context, with suggestions that such programmes (others that claim to be 'international', as well as those of the IB) might in fact be more accurately described as western or western-liberal in values and focus, as discussed by, *inter alia*, Drake (2004), Van Oord (2007) and indeed by Walker (2010), a previous Director General of the IB. Such questions seem likely to persist, with proposals that international programmes should acknowledge other cultural perspectives than simply that of a western liberal value system arguably being in tension with the fact that the recent rapid growth in international schools is highlighting the latent demand from growing middle classes in non-western contexts for the form of international (western) education expected to provide their children with a competitive edge through the acquisition of English-medium internationally recognised qualifications that offer access to prestigious western universities, and thus to the western privileged lifestyle to which many such families aspire.

Debate will no doubt continue as to what it means, or should mean, to offer an international form of education. Having emerged as a concept through the IB Diploma Programme, principally as a pragmatic means of catering for globally mobile (largely

western) expatriates, and at a time (1960s) when the term 'globalisation' was only just beginning to emerge (Held and McGrew, 2000: 1), expectations have been raised in the intervening years about what might be expected of a form of education with such aspirations. Thompson's question about the risk of an 'unwitting process of educational imperialism' is echoed in more general questions of 'post-colonialism' raised by Andreotti, for instance (2011). Perceived exclusivity of programmes such as the IB Diploma Programme that are relatively expensive to offer compared with their nationally-based counterparts and thus beyond the reach of students attending most non-fee-paying schools is no less an issue, though with some exceptions in a small number of national contexts including, for instance, Ecuador where 72% of its 254 IB schools are state-funded (Hill, 2022). Recent analysis by Tate (2022) of the place of international schools in current debate arising from Goodhart's identification of 'Anywheres' ('a large minority group of the highly educated and mobile') and 'Somewheres' ('a larger and less influential group ... who are more rooted and less well educated') (2017: vi) raises related questions as to the potential of international schools, however ideologically-focused, to exacerbate differences in our rapidly changing world.

Going forward

This chapter has so far provided an overview of the international school sector from its earliest days, and in doing so has highlighted a number of challenges facing the sector as its *raison d'être* has changed in response to the changing global environment. From individual schools established principally to meet local needs of globally mobile expatriate families, an international school 'sector' has developed of some 12,300 schools (as at July 2021: ISCR, 2021b) which either claim to be international or might be considered as such on the basis of their characteristics, while acknowledging that the number could be somewhat larger if similar schools are included that do not meet ISCR's English-medium criterion. While to what extent international schools do actually form a 'sector' is debatable, given the diversity of the institutions and the absence of any overarching body with authority to determine whether a particular school may describe itself as such, major differences evident between the present day and the earliest years include the emergence of *groups* of international schools such as elite schools from a national context (for instance the UK's Dulwich and Harrow) and other groups including Nord Anglia and GEMS. And while the vast majority of international schools are fee-paying (with some notable exceptions such as the Netherlands-based Dutch International Primary Schools and Dutch International Secondary Schools, and those already noted as offering the IB Diploma in Ecuador), and 'not-for-profits' feed back surplus funds into development of the school (the basis of many of the earlier international schools), those for which surplus funds are channelled as profit to owners, shareholders and investors have become a clear trend in very recent years, with the majority of new international schools established on a commercial basis. The changes in nature of international school student populations is clearly an associated phenomenon, with the vast majority of international school students no longer the globally mobile expatriates for whom such a school facilitated relocation with a parent's professional mobility, but rather the elite and aspirational middle classes in developing countries for whom an international school education and internationally recognised pre-university qualification provide a stepping stone to membership of what has been described as the 'transnational capitalist class' (Sklair, 2001).

The rapid changes in even very recent years, with notable growth in particular contexts including but not limited to China, suggests that the international school 'sector' is likely to continue to change and develop in coming years. As the 'sector', and associated curriculum organisations, grows, its interest as a focus for academic research will undoubtedly also grow. A focus on international schools has already been noted as one of the main areas of research in international education (Dolby and Rahman, 2008) and seems set to be the basis of increasing scrutiny and critique by those with interests both in this specific context and in related areas. Questions increasingly being raised in the growing international school research base include, *inter alia*, those related to the nature of international schools (see, for instance, Bunnell et al., 2017), leadership of international schools (Keller, 2015) and teachers in international schools (Bailey, 2015).

Likely also to be of increasing interest in the international school context are areas that, while not unique to this context, may have particular relevance where mission statements and ideologies claim the development of attributes in students including respect for others, international mindedness and global citizenship. So, for instance, the Black Lives Matter movement's highlighting of issues relating to equality and diversity, and the focus on decolonisation of the curriculum in many contexts (see, for instance, Aman, 2017) will almost certainly be increasingly prevalent in the coming years – as will the expectation of a proactive role from international schools in ensuring that global topics including the Climate Emergency will continue to have high visibility in the curriculum.

Anticipated growth in use of technology in education, though again not unique to the international school context, may be expected to be associated with notable changes in the teaching and learning process. The global pandemic's impact on schooling has opened the eyes of many educators and students to the benefits of technology, with earlier speculation about a possible model for the future of international schools being one that relies less on buildings and more on remote interaction (Hayden and Thompson, 2013) becoming, as a result of COVID-19, a reality far more quickly than could ever have been anticipated. And not only increased use of technology, but also advances already being made in Artificial Intelligence, Virtual Reality and Augmented Reality are clearly likely to influence developments in the coming years (see, eg, MacDonald, 2022). Though not discussed here in any detail, two other issues may be expected to be the focus of continuing debate. One very pragmatic question relates to those who teach in international schools. In earlier days, almost exclusively expatriate and largely from a limited number of western, English-speaking countries (except for specialist teachers of local language and culture), as the numbers of schools have grown so has the diversity of teachers – though many schools still struggle with pressure from parents to appoint only first language English speakers and,

some would argue, underlying discrimination may be found in at least some contexts in favor of white, western, western-educated applicants. Already a hot topic too in relation to teachers in international schools is the question of just how the rapidly-growing number of schools will find the increasing numbers of teachers needed to sustain them. And if these growing numbers of teachers are to be recruited not only from the Global North, challenging issues arise about the ethics of relatively wealthy schools attracting from the home context in which they were educated teachers who leave behind shortages in a local, less well-resourced, education system. While, for many expatriate teachers, a short time spent teaching in international schools, or even a career developing from an initial foray into the international school world can be rewarding, exciting and stimulating, the precarity of the context (Bunnell, 2015), with short term contracts and little of the security that might be expected by teachers in national contexts, needs also to be acknowledged.

There can be few forms of education that have followed quite such a rapidly changing trajectory in the past half century as have international schools. From the origins of individual institutions established to cater for a local need, they are emerging as major influencers in an increasingly globalised world, arguably contributing to growing societal divides and facilitating the growing gap between 'Somewheres' and 'Anywheres' (Goodhart, 2017; Tate, 2022). Although the phrase 'international school sector' is widely used, there has arguably never been a 'sector' as such; nor does it seem that the existence of such a thing is closer now than it ever was. It may well be time for the concept of an international school *sector* to be abandoned, given the impossibility of making meaningful generalisations that apply equally to schools catering for expatriates relocating temporarily abroad, as to schools established principally to cater for aspirational local families in developing countries. Indeed, as noted by MacDonald (2022: 222), 'International schools catering primarily to host country nationals are often filling a void in the local independent school market ... [and] ... have become *de facto* independent school networks', citing the United Arab Emirates and Thailand as two examples. Though times have changed since Hayden and Thompson's (1995) description of international schools as a 'conglomeration of individual institutions', it is still the case that it would be misleading to assume that international schools in general share specific characteristics common to all.

Some would argue that the earlier forms of international schools are closer to the concept of a 'true' international school than those that have developed more recently. We would argue, however, that it does not need to be the case that one type of school can be assumed to be superior to another. Particularly given the lack of clarity with respect to terminology, and that programmes such as those of the IB, Cambridge and Fieldwork are offered across different types of schools, it cannot be assumed, for instance, that a for-profit school is by its nature somehow deficient in the international nature of the education offered compared with that of a longer-standing more traditional not-for-profit institution. Were it the case that a global organisation had the power to determine whether a particular school could be described as an international school, this is clearly a time when directives as to terminology would be beneficial to all concerned. In the absence of such authority, a lack of clarity is likely to continue. This does not, however, prevent international schools from being an increasingly high profile and influential contributor to education worldwide, and an important focus for research by those with interests in international education and its associated areas.

References

- Alice Smith School, 1996. The Alice Smith School, Kuala Lumpur: A Celebration of 50 Years, 1946-1996. The Alice Smith School, Kuala Lumpur.
- Aman, R., 2017. Decolonising Intercultural Education: Colonial Differences, the Geopolitics of Knowledge, and Inter-epistemic Dialogue. Routledge, London.
- Andreotti, V., 2011. Actionable Post-colonial Theory in Education. Palgrave Macmillan, New York.
- Azzam, Z., 2021. International Schools: whose definition is it anyway? International School. Summer 10-13.
- Bailey, L., 2015. Reskilled and 'Running Ahead': teachers in an international school talk about their work. J. Res. Int. Educ. 14 (1), 3-15.
- Barratt Hacking, E., Blackmore, C., Bullock, K., Bunnell, T., Donnelly, M., Martin, S., 2018. International mindedness in practice: the evidence from international baccalaureate schools. J. Res. Int. Educ. 17 (1), 3-16.
- Bereday, G.Z.F., Lauwerys, J.A. (Eds.), 1964. The Yearbook of Education 1964: Education and International Life. Evans Brothers, London.
- Bunnell, T., 2008. The exporting and franchising of elite English private schools: the emerging 'second wave'. Asia Pac. J. Educ. 28 (4), 383-393.
- Bunnell, T., 2015. Teachers in International Schools: a global educational 'precariat'. Glob. Soc. Educ. 14 (4), 543-559.
- Bunnell, T., Fertig, M., James, C., 2017. Establishing the legitimacy of a school's claim to be 'international': the provision of an international curriculum as the institutional primary task. Educ. Rev. 69 (3), 303-317.
- Cambridge Assessment International Education, 2021. www.cambridgeinternational.org.
- Cambridge, J., Thompson, J.J., 2004. Internationalism and globalisation as contexts for international education. Compare 34 (2), 161-175.
- Clifford, V., Montgomery, C., 2017. Designing an internationalised curriculum for higher education: embracing the local and the global citizen. High Educ. Res. Dev. 36 (6), 1138-1151.
- Coulby, D., Zambeta, E., 2005. In: World Yearbook of Education 2005: Globalization and Nationalism in Education. Routledge, London.
- Dickens, C., 1864. International education. All the year round: a weekly journal, XII (281), 106-108.
- Dolby, N., Rahman, A., 2008. Research in international education. Rev. Educ. Res. 78 (3), 676-726.
- Drake, B., 2004. International education and IB programmes: worldwide expansion and potential cultural dissonance. J. Res. Int. Educ. 3 (2), 189-205.
- Fasheh, M., 1985. Talking about what to Cook for Dinner when Our House is on Fire: the poverty of existing forms of international education. Harv. Educ. Rev. 55, 121-126.
- Fieldwork Education, 2021. www.fieldworkeducation.com.
- Global Learning Programme, 2021. Available at: <https://globaldimension.org.uk>.
- Goodhart, D., 2017. The Road to Somewhere. The New Tribes Shaping British Politics. Penguin Books, London.
- Gray, M.R., 2003. Tense conjugations: translating political values into an educational model: the European schools. J. Res. Int. Educ. 2 (3), 315-330.
- Hayden, M.C., Thompson, J.J., 1995. International schools and international education: a relationship reviewed. Oxf. Rev. Educ. 21 (3), 327-345.
- Hayden, M.C., 2006. Introduction to International Education: International Schools and Their Communities. SAGE, London.
- Hayden, M.C., 2011. Transnational Spaces of Education: the growth of the international school sector. Glob. Soc. Educ. 9 (2), 211-224.

- Hayden, M.C., Thompson, J.J., 2013. International Schools: antecedents, current issues and metaphors for the future. In: Pearce, R. (Ed.), *International Education and Schools: Moving beyond the First 40 Years*. Bloomsbury Academic, London.
- Held, D., McGrew, A., Goldblatt, D., Perraton, J., 2000. Global transformations: politics, economics and culture. In: Pierson, C., Tormey, S. (Eds.), *Politics at the Edge*. Springer, New York, pp. 14–28.
- Heyward, M., 2002. From international to intercultural: redefining the international school for a globalized world. *J. Res. Int. Educ.* 1 (1), 9–32.
- Hill, I., 2022. The democratisation of international education. In: Hayden, M.C. (Ed.), *Interpreting International Education: In Honour of Professor Jeff Thompson*. Routledge, London.
- Husén, T., 1985. International education. In: Husén, T., Postlethwaite, T.N. (Eds.), *The International Encyclopaedia of Education: Research and Studies*. Pergamon, Oxford.
- IB, 2021a. www.ibo.org.
- IB, 2021b. The IB Learner Profile. Available from: www.ibo.org/benefits/learner-profile.
- ISC Research, 2021a. ISC Research. www.iscresearch.com.
- ISC Research, 2021b. Data on International Schools. www.iscresearch.com/data/.
- Keller, D., 2015. Leadership of International Schools: understanding and managing dualities. *Educ. Manag. Adm. Leader* 43 (6), 900–917.
- Knight, M., 1999. *Ecolint: A Portrait of the International School of Geneva 1924–1999*. International School of Geneva, Geneva.
- Lauder, H., Young, M., Daniels, H., Balarin, M., Lowe, J., 2012. *Educating for the Knowledge Economy? Critical Perspectives*. Routledge, London.
- Lawton, D., 1989. *Education, Culture and the National Curriculum*. Hodder and Stoughton, London.
- Lester B Pearson College of the Pacific, 1982. *Adventures in High Endeavour: People Understanding People. The Story of Lester B Pearson College of the Pacific*, Toronto: Brownstone.
- MacDonald, J., 2006. The international school industry. *J. Res. Int. Educ.* 5 (2), 191–213.
- MacDonald, J., 2022. Where to next? What does the future hold for international education? In: Hayden, M.C. (Ed.), *Interpreting International Education: In Honour of Professor Jeff Thompson*. Routledge, London.
- Machin, D., 2017. The great Asian international school gold rush: an economic analysis. *J. Res. Int. Educ.* 16 (2), 131–146.
- Marshall, H., 2007. The global education terminology debate: exploring some of the issues. In: Hayden, M.C., Levy, J., Thompson, J.J. (Eds.), *SAGE Handbook of Research in International Education*. SAGE, London.
- Matthews, M., 1989. The scale of international education. *Int. Sch. J.* 17, 7–17.
- Maseru English Medium Primary School, 1990. *Maseru Prep 1890 – 1990: The School in the Sky*. Maseru English Medium Preparatory School, Maseru.
- Oxfam, 2015. *Global Citizenship in the Classroom: A Guide for Teachers*. Oxfam GB, Oxford.
- Peterson, A.D.C., 1987. Schools across frontiers: the story of the international baccalaureate and the 'united world colleges'. La Salle II: Open Court.
- Poole, A., 2021. *International Teachers' Lived Experiences: Examining Internationalised Schooling in Shanghai*. Palgrave Macmillan, New York.
- Sklair, L., 2001. *The Transnational Capitalist Class*. Blackwell, Oxford.
- Stanworth, D., 1998. *An International Education at Yokohama International School: From Theory to Practice*. unpublished MA Education dissertation, University of Bath, Bath.
- Stewart, W.A.C., 1972. *Progressives and Radicals, 1750–1970*. Macmillan, London.
- Sylvester, R., 1998. Through the lens of diversity: inclusive and encapsulated school missions. In: M.C. and. In: Thompson, J.J. (Ed.), *International Education: Principles and Practice*. Kogan Page, London.
- Sylvester, R., 2002. The 'first' international school. In: Hayden, M.C., Thompson, J.J., Walker, G.R. (Eds.), *International Education in Practice: Dimensions for National and International Schools*. Kogan Page, London, pp. 3–17.
- Tate, N., 2022. Academies for Anywheres: international schools, the nation state and identity. In: Hayden, M.C. (Ed.), *Interpreting International Education: In Honour of Professor Jeff Thompson*. Routledge, London.
- Thompson, J.J., 1998. Towards a model for international education. In: Hayden, M.C., Thompson, J.J. (Eds.), *International Education: From Principles to Practice*. Kogan Page, London.
- United World Colleges, 2021. www.uwc.org.
- Van Oord, L., 2007. To Westernize the Nations? An analysis of the International Baccalaureate's philosophy of education. *Camb. J. Educ.* 37 (3), 37–390.
- Walker, G.R., 1996. *Keeping the Balance*, vol. 11. IB World, pp. 19–20 (April).
- Walker, G.R., 2010. East Is East and West Is West. International Baccalaureate Organization, Geneva. Available from: <https://www.ibo.org/globalassets/publications/east-is-east-and-west-is-west-en.pdf>.

International Baccalaureate: meanings, uses and tensions in a globalizing world

Paul Tarc, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	344
The phenomenon of IB	345
Research on IB	346
IB and the shifting geopolitics of education under globalization	347
IB as a window on (Anglo-Western) international education	347
IB as a window on school internationalization	351
Conclusion	352
References	353
Relevant website	354

Introduction

Much could be written about the multiple dimensions of the current phenomenon of the *International Baccalaureate* (IB)—its presence, growth and impact in the world, as well as its uptake in the academic literature. Therefore, I want to make explicit my approach to constructing this entry. First, to fully appreciate what IB represents today, it is vital to consider its history. And it is critical to understand not only how IB emerged in the field of a small set of multilateral international schools in the 1960s (Peterson, 1972; Mayer, 1968), but to differentiate the conditions of this past historical moment to those of today (Scott, 2004; Tarc, 2009, 2011). There are instances when policy expressions from the period of IB's creation have been read through the filter of the (21st-century) present, misconstruing what was meant by those expressions and, by consequence, misrepresenting the past's mark on the present.

Moreover, with the current mainstreaming of terms like “global citizenship” sloganized across multiple organizations, it is easy to forget that it is only relatively recently that nation-states have been more sympathetic to such terms as “global citizenship education” and to forms of internationalizing education, given the traditionally tight grip “sovereign” nations have held upon their (idealizations of) state schooling (Heater, 1980; Tarc, 2009). Indeed this shift, beginning in the early 1990s, represents a core feature of the “shifting geopolitics of education” under globalization. Accordingly, to understand the character and development of IB requires an examination of the larger conditions that have shaped its concrete manifestations and evolving policy rhetoric on its purposes, achievements, modifications and plans. On the current (2020) “about-the-ib” webpage, viewers confront IB's long-standing aspirational vision, “Now in our 52nd year, we're more dedicated than ever to developing international education that creates a better world” (ibo.org). This aspirational goal, of making a better world through a progressive education for “international understanding,” has endured since IB's inception. However, how these aspirational visions are expressed and manifested in practice are enabled and constrained by institutional and regional pressures/agendas and larger temporal conditions (Tarc, 2009).

Second, the official stories that IB leadership tells of IB represent only one part of the reality of IB. The other part is what happens on the ground, why and how schools, universities, governments and families open to and (potentially) adopt or use IB. These parts reflexively inform one another, but they also produce discord, contradictions and tensions. IB, then, is constituted by both top-down governance and policy as well as bottom-up engagements, above and below the IB organization's core function of providing its four educational programs in schools. For this reason, I employ in my title the more performative terms, “meanings and uses” of IB, consistent with a pragmatist lens (Rizvi, 2014). From this perspective, IB is not some essential “thing,” but has flexible meanings and tangible uses and intended and unintended effects across the diverse contexts in which it is adopted and engaged. These meanings and uses (and tensions) are mediated by a confluence of factors, such as the following: larger conditions of neoliberalization (exogenous and endogenous to nation-states), IB's policy rhetoric and governing practices, state and university admissions policies, school/curricular practices and the (cosmopolitan) perspectives and (global) class-making strategies of IB users' families.

Finally, the section on the historical development and evolving tensions of IB, that situates international education under the globalization processes of recent decades, is mainly derived from a periodization of IB presented in my book, *Global Dreams, Enduring Tensions: International Baccalaureate in a Changing World* (2009). For greater explication on the historical development of IB and how the three structuring tensions (citizenship, curricular aims and operational function) find altered dynamics across time, readers can turn to this 2009 publication. Given the date of this publication, I have particularly considered the literature on IB published in the last decade; further, I have reviewed more recent IB policy statements to extend the analytic trajectory of the *Global Dreams* text. For additional historical analyses of IB, see Bagnall (1994), Bunnell (2008), Fox (1985), Hahn (2003), Hill (2002a,b) and Peterson (1972, 1987).

The phenomenon of IB

As of July 2020, the IB Organization reports that there are 7002 IB programs offered in 5284 schools in 158 countries; about 52% of these schools are state-funded schools (ibo.org/programmes/find-an-ib-school/, accessed, Nov 25, 2020). From 2015 to 2019, the numbers of IB programs adopted increased by a significant 37.9%. Adoption of the IB programs by schools is geographically uneven particularly with IB schools in public sectors. Just over half of the IB schools are in “The Americas” (predominantly in the US and Canada). “Africa, Europe and the Middle East” have about 21% of the schools (despite recent growth, Africa accounts for only 2.3% worldwide). And the “Asia-Pacific” region accounts for about 27% (ibo.org/programmes/find-an-ib-school/, accessed, Nov 25, 2020). 2018 statistics analyzed by Bunnell (2020) are illuminative:

Put simply, in 2018 there were at least 53 nations where there existed authorized “IB World Schools” yet zero public schooling activity, whilst four nations (Australia, Canada, Ecuador, and the United States) accounted for 81% ... of the IB’s overall body of public schools.... In most parts of the world, the IB (still) operates out of a traditional, private and relatively elite schooling mode of activity. (p 60)

Thus, apart from a handful of unique arrangements with governments and IB (such as in the cases of Ecuador and Sweden), the majority of IB publicly funded schools are located in the Anglo-West; while, in “developing country” contexts, most IB schools are private institutions, primarily serving mobile and national elites.

The IB is run by a non-profit foundation registered in Switzerland. In the most recently published Annual Review (2018–2019), the Director General, Siva Kumari reiterates IB’s “three business areas:”

Working closely with our passionate community of educators (over 5000 schools in more than 150 countries), our mission inspires us to continual improvement in all aspects of our work in our three business areas: curriculum development, working with schools, and assessment. (ibo.org/about-the-ib/facts-and-figures/ib-annual-review/year-in-review-2018-2019/a-message-from-dr-siva-kumari-to-the-ib-community/, accessed on November 25, 2020).

Centering IB’s “work” is the provision of four preK-12 educational programs; the 5000 plus *IB World Schools* offer at least one of these programs to their students. The *IB Diploma Program* (IBDP), the longest-standing and most popular program (offered by more than 3500 schools), is provided for students aged 16–19 years; it officially began in 1968. The *Middle Years Program* (MYP) for ages 11–16 began in 1994. The *Primary Years Program* (PYP) started in 1997 and is for children aged 3–11 years. More recently, in 2012, IB launched the *Career-related Program* (CP) for 16-19-year-olds that leads to “further/higher education apprenticeships or employment” (ibo.org/programmes). Currently there are 274 schools offering this new program. The IB’s website (ibo.org) is a well-updated site hosting materials and comprehensive details on its mission, philosophy, governance structure, finances, operations, history, curricular programs, geographic spread and growth, annual review statements, research summaries on IB, IB events and initiatives (some showcased in the “IB World” magazine), etc. This entry will not provide a description of the various elements of IB; the website is a good source for accessing these details.

Beyond the increasing numbers of authorized IB World schools and users of IB programs and IB courses in the K-12 private and state-funded sectors, IB has found notoriety in additional arenas. On the one hand, IB has a growing presence in national and transnational educational policy spheres (Tarc, 2009; Tarc and Beatty, 2012); on the other hand, the IB has entered new domains to spread its influence (Tarc, 2009). For example, across the last two decades, IB has partnered with multilateral policy actors and philanthropic foundations on non-IB educational projects. More recently, the IB has partnered with a select number of universities’ faculties of Education. In concert with the IB organization, these faculties now offer *International Baccalaureate Education Certificates* (IBEC) in or alongside their preservice teacher education or graduate education programming (see ibo.org/contentassets/f23b082dbc184e379a5bec2d42009e73/ibec-2020-university-directory.pdf). Additionally, the 2018–19 *Annual Review* highlights new partnerships and projects with the governments of United Arab Emirates, Japan and South Korea, and the launching of a Master of Education program with the University of the People to offer “a tuition-free online university degree to benefit teachers worldwide” (ibo.org/about-the-ib/facts-and-figures/ib-annual-review/year-in-review-2018-2019/impact/ accessed on July 22, 2020). Such examples illustrate the IB Organization’s commitments to “service” and having “impact” beyond its core mandate of providing its four educational programs.

Anecdotally, a good number of colleagues and acquaintances over the years have incidentally mentioned IB; I am always intrigued to know what they mean by it. As suggested above, the IB has multiple meanings and uses; and, as I have argued (2009), this flexibility of IB has been instrumental to its widespread adoption and financial sustainability. Is IB a cosmopolitan social movement (as implied in DG Siva’s invocation of a “passionate community” above and as showcased/advanced in the IB World magazine and various IB networks, blogs and groups)? An education for global citizenship (Dvir et al., 2018) or international mindedness (Hacking et al., 2018)? A private school education (within a publicly funded school) for only the price of examination fees (Tarc, 2007)? An inquiry-based pedagogical model (Twigg, 2010)? A “gold standard” (of quality) for well-established international schools (Lauder, 2007) and/or for the fast growing, for-profit sector of Anglo-Western-inspired international schools (Waterson, 2016)? A liberal—but not political—model of international education (Tarc, 2011) acceptable to more authoritarian nation-states? An UN-inspired infringement on state schooling in the US (Bunnell, 2012)? A model of gifted education (Kyburg

et al., 2007; Poelzer and Feldhusen, 1997)? An “international passport” to elite universities in the West (Lee and Wright, 2016)? An academically challenging program for US Tier One/low SES schools (Mayer, 2008)? Or a school choice option for (upper-) middle class families (Doherty, 2009)?

As reported in the literature, IB is understood and used in each of these ways depending on the context and stakeholders involved. This malleability has proven useful for the IB’s viability across contexts and across time, but it also has produced concerns and tensions that require labor by the IB organization to assert and re-assert its authorship over the IB brand/ideals. I will continue to address this malleable feature of IB as it (in)forms the research literature and the terrain upon which IB policy is (re)formulated and the ongoing cultural production of the IB brand.

Research on IB

As a “learning organization” (Tarc, 2009), the IB is invested in research in terms of better understanding and improving upon its programs and operation, as well as to leverage findings to build its reputation and deepen and broaden the IBDP’s acceptability to university admissions offices (and state educational ministries) globally. As Resnik (2019) asserts: “IB research is one of the main nonhuman actors that encourages DP recognition ...” by universities (p. 347). Thus, the research IB does on IBDP students’ readiness for, or success in, university programs becomes part of the way that IB inserts itself into the national; positive research findings thus act as a “non-human actor” in a larger “assemblage” of actors influencing university admission policies on the IBDP. Such evidenced-based research also allows IB to more legitimately participate in the performative cultures of transnational policy making spheres alongside more prominent agents as OECD or UNESCO.

IB has also increasingly come on the radar of academic researchers, including graduate students. The malleability and multidimensionality of the IB is well reflected in the growing number of research studies engaging IB. Research on IB crosses a broad spectrum from more insider or practice-based studies, conducted or commissioned by the IBO, that investigate (some dimension) of IB with a focus on evaluation or improvement, to more outsider or academic research that takes IB as an exemplar of a particular form, or proxy, of education (such as gifted education) or as a window onto a larger phenomenon (as school choice for class making). In these latter approaches to research, the aim is to illuminate the form of education or the larger phenomenon more than features intrinsic to IB. However, there are also insider studies that engage larger questions of educational aims or methods and use IB as the example, and academic research that does more intrinsically study the IB. The IBO hosts a research page (ibo.org/research/) which profiles IB in-house and IB commissioned research categorized as either “outcomes research,” “curriculum research” or “policy research.” They also have commissioned and posted annual annotated bibliographies inclusive of academic research conducted on IB for the period 2010–19. These bibliographies, conducted by university academics, cite and provide abstracts of academic journal articles, theses and dissertations, book articles, reports and conferences (see ibo.org/research/research-resources/).

A review of these sources as well as cited sources found through educational database searches surface common strands of research on the IB. A number of studies take the IB programs as a prominent exemplar of international education (for example, Hill, 2007, 2012); some studies more particularly engage the tension between the idealist and instrumental agendas or visions of international education in a context of globalization (Cambridge and Thompson, 2004; Gardner-McTaggart, 2016; Hill, 2006); relatedly some studies engage the (philosophical) mission or cultural affinities/translations of IB (Drake, 2004; Hayden and Wong, 1997; Lineham, 2013; Rizvi et al., 2020; van Oord, 2007; Wells, 2011). A number of studies examine the trends and prospects of/for IB in specific geographic areas, such as Australia (Kidson et al., 2019) and China (Wright and Lee, 2014). One of the more developed strands of research employs sociological analysis to illuminate IB’s use as a choice option with neoliberal school reform and how IB offers advantage or distinction for (global) class making (Doherty, 2009, 2012, 2013; Doherty et al., 2009) along the schooling to university trajectory (Wright and Lee, 2019). These strands make evident the multidimensionality of IB and the roles that IB plays in education and in educational markets worldwide, as well as the ongoing salience of IB as an object of scholarly research.

Most compelling, perhaps, is the IB organization’s enduring viability and strong reputation as a non-state provider of progressive curricular programs for “international mindedness” and its attendant teacher professional development and examination/oversight regimes operating for more than half a century (Tarc, 2021). Also striking is the character of IB’s global geographic dispersion and significant entry into, and ability to work within or alongside, state-funded systems. How has the IB navigated such a complicated terrain, and for so long? How (well) does it hold to its “global dreams” (Tarc, 2009) of making a better world through education? How does it respond to the dynamic tensions that arise as the “dream” enters the practical realities across different geopolitical and cultural contexts? As IB expands, how does it ensure quality of its programs (Charleson, 2010) as well as remain distinctive, and thereby desirable, in light of competition (Doherty, 2013)? What are the current trends and prospects for IB (both functionally and aspirationally) in a still hyper-connected, uneven world, now in further global crises? (How) will/might IB (continue to) be a global agent, as well as a reflection, of educational reforms in the shifting geopolitics of education? These questions are very salient for current and future scholarly research on IB. Some of the historical and analytic material to support these prospective inquiries is offered in this entry.

The following subsections focus more specifically on two research strands most relevant to this volume’s focus on the shifting geopolitics of education. The first strand takes IB as an exemplar of international education under globalization and the second strand takes IB as constitutive of internationalization processes of K-12 schooling. My approach is to draw a distinction between international education and its variants as a long-standing set of educational ideals, practices and initiatives (Elvin, 1960; Good,

2020; Heater, 1980; Méras, 1932) and the internationalization of education as a more recent trend emerging from the 1990s under processes of neoliberalization (Tarc, 2019; Tarc et al., 2012). This distinction is useful to differentiate older and newer modalities of international education with their potentially different objectives or animating visions. It also parallels the distinction between the *literal* definition of international education as educational activities crossing or connecting across political borders and the *ideal* of an outward looking education for international understanding (Tarc, 2019). Internationalization of education as a recent trend entails both the literal and aspirational definitions of international education, but the larger neoliberalizing conditions that drive internationalization agendas from above, favor the literal and “instrumental” definitions over the “ideological” or “educational” (Stier, 2004). Consequently, IB’s adoption may have more to do with “international education” than “internationalization” or vice versa; but the point here is that these empirical and normative differences matter. Considering IB as an exemplar of international education (as in educating for international understanding) leads to a different set of questions (and critiques) than considering it as an exemplar of the internationalization of K-12 schooling (as supra-national presencing in state schooling). Of course, linkages need to be made as part of the analytic, but under-thought confluences risk clouding research aims and findings.

IB and the shifting geopolitics of education under globalization

This section presents two specific analytic strands most relevant to this volume’s theme of “globalization and the shifting geopolitics of education.” The first strand considers IB as an exemplar of international education and the trends and prospects for the 20th century dream of international education under the unfolding 21st century conditions. The second strand centers on the internationalization of K-12 schooling, where “IB is [taken as] an emblematic case of educational globalization” in terms of “de-nationalizing” state schooling (Resnik, 2012, p. 249) or “school internationalization” (Engel et al., 2019). For this second strand, I read Resnik (2012) article, *The denationalization of education and the expansion of the International Baccalaureate*, with and against Bunnell (2020) recent article, *The “internationalization of public schooling” in practice: A “skeptical reality” approach*.

IB as a window on (Anglo-Western) international education

While the idea of an international baccalaureate was not new in the 1960s when IB came to life (Hill, 2002b), there was, at this time, sufficient practical demand for an internationally-recognized secondary school leaving diploma, to facilitate expatriate families’ access to home-country universities in the West (Peterson, 1972). The practical necessity and logistics of developing an internationally recognized diploma for multilateral international schools was foundational to the development of the IBDP. However, equally foundational, were the “global dreams” of IB (Tarc, 2019)—the progressive educational and cosmopolitan visions of the creators and supporters of IB to develop an innovatory educational program for international understanding. For most of the 20th century “international understanding” was the dominant signifier of the pedagogical goal of international education (Heater, 1980; Tarc, 2009). More than representatives of their own national systems, founding Director General (DG) Alec Peterson and his collaborators were largely educational reformers, critical of encyclopedic (and nationalist) approaches to schooling (Mayer, 1968). They envisioned an education for international understanding as a humanist “education of the whole person” (Peterson, 1972; Renaud, 1974), where students would study across the humanities, arts and sciences, engage a second language and experience social service and aesthetic activities. IB had a lineage to the Kurt Hahn-inspired service/outdoors movement, as well as to the relatively independent English private schools’ movement. While a regime of centralized examinations would be IB’s method for ensuring a level of standards for university acceptability, the program was aimed at deepening students’ understanding of the world through disciplinary and interdisciplinary study which included the cultivation of the moral and the aesthetic (Peterson, 1972). In the founding period of IB, a classical progressive education in the internationalist milieu of multilateral international schools and their communities was the means to international understanding (Tarc, 2019). For this non-state actor, a multimodal examination regime would allow for the steering of a curricula for international understanding and be the accountability mechanism needed to gain acceptability from university admission offices (Tarc, 2009).

The following six books are particularly illuminative of the *time-space milieu* of IB’s creation and experiment in the mid-1960s to early 1970s: Leach (1969) *International Schools and their Role in the Field of International Education*, Mayer (1968) *Diploma: International Schools and University Entrance*, Malinowski and Zorn (1973) *The United Nations International School: Its History and Development*, Renaud’s (1974) *Experimental Period of the International Baccalaureate: Objectives and Results*, Peterson (1972) *The International Baccalaureate: An Experiment in International Education*, and Peterson (1987) *Schools Across Frontiers: The Story of the International Baccalaureate and the United World Colleges*. These books provide a glimpse of the ethos, motivating ideals, practical realities and concrete problems and logistics that represent the contextual features of the community of IB creators and supporters (many of whom were teachers) that brought the IB to life in a historical moment that is quite distinct from our current one.

The confluence of the idealist educational/cosmopolitan visions and the practical demands and logistics produced a set of tensions that have endured from the period of IB’s creation and experiment (1962–73) to the present day (Tarc, 2009). In the historical moment when IB emerges, three core tensions constitute the “international” of IB:

The structuring tensions of IB emerge from the interplay of the dream of international understanding and the functional operation of an international diploma at work in the world. The term "international understanding," on its own, is under tension as an educational aim in a historical period when a dominant purpose of schooling was to produce loyal national subjects. The educational ideal of IB as a progressive education of "the whole person" is in tension with the need for IB to have internationally acceptable standards for university entry. And the ideal of IB representing a modern, forward-looking model of schooling, oriented to making a more peaceful and humane world in a historical period of democratization movements, becomes strained where IB was effectively used by a social elite. These three examples signal the core tensions of the "international" of IB in the founding moment.

Tarc (2009, p. 23).

The first tension of "citizenship," thus, centers on the IB's mission of developing international understanding when schooling is to foster national understanding and loyalties. Although international understanding was not contentious within the communities of the participating multilateral international schools, the IBO had to temper its internationalist sentiments in seeking recognition and funding from national governments and institutions. In its policy statements, IBO's predominant focus is on the forms and aims of the IBDP education and assessment (Tarc, 2009). Where international understanding is discussed explicitly, the IBO emphasizes that students first must identify with their national identity and later develop an openness to other nations and cultures (Peterson, 1972). Additionally, the IBDP is consistently described as a "complement to," or as a potential "laboratory" for, national schooling (Peterson, 1972).

The second and more consequential "curricular tension" is explicitly stated by Peterson (1987):

One of the problems which from the start face the IBO in developing an international curriculum was the tension between the academic requirements of university entrance procedures and these personal requirements of the whole human being growing up in an interdependent world. (p. 199)

Thus, as with progressive education more generally, the innovatory and progressive elements are constrained by demands for standards and accountabilities. In the case of the IBDP, deputy DG Gerard Renaud admitted that some university stakeholders working with the IBO had "dictated the content of some programs ... and sometimes imposed a greater degree of conservatism than the promoters of the experiment desired" (IBO, 1972, p. 27).

Still, in these early years, the IBDP curricular structure and assessment regime were innovatory and potentially enabling of more progressive and internationalist pedagogies. In the first place, with the IBDP, senior secondary students in international schools were no longer required to be split up to study for national entrance examinations (Leach, 1969). Second, the examination system was nuanced and multimodal to mitigate the back-wash effect of teaching to the test and the use of rote learning approaches (Peterson, 1972). Moreover, the curricular design of the IBDP compelled students to study a range of subjects and take an innovatory "core" that included the Theory of Knowledge course, a student-initiated "extended essay" and the "creative, aesthetic and social service experience" (CASS). Additionally, there was room for individual schools to create a school-based syllabus (SBS) to address local interests. One of the core international schools participating in the creation of IB, Atlantic College, designed and offered a *Peace and Conflict Studies* course, which represents an example of the (still under-used) innovatory possibilities of the IBDP (Tarc, 2009).

The third "operational" tension refers to the disconnect between the larger internationalist-egalitarian vision of IB (emerging in a time of political decolonization and the democratizing and massifying of secondary and postsecondary schooling) and the not-so international character of the organization and program and with the elite social class backgrounds of the users of the IBDP. While the idea that international education could be massified beyond elites and that IBDP should be open for the "academically-able," the schools that offered the IBDP in the experimental period served socially elite families (Tarc, 2009). Limited access to the IBDP thus represented one pillar of the operational tension.

In terms of its inter-national make-up or representativeness, the IBO and the IBDP curricula reflected specifically the position-alities and perspectives of individuals from a small set of wealthy Western nations. These were the voices of consequence in terms of university partners and of the development of the program, curricula and assessment operation. On the one hand, the internationalism of IB signals the inclusion of perspectives from a multiplicity of nations; but, on the other hand, the international refers to a "chain of equivalencies: West = democratically advanced = modern = international" (Tarc, 2009, p. 42). "Representation" thus represents the second pillar of the operational tension:

For the IBO, the structuring tension of representation in the founding period is produced out of the desire to include national perspectives and voices within a hierarchy of [assumed] relevance and expertise. University entrance requirements in England, Switzerland, Germany and France, needs of Anglo-American international schools for the mobile elite and recommendations of funders and other liberal-minded enthusiasts magnify the influence of particular perspectives and voices over others. (p. 42–43)

Given the enormity of the task to secure funding and support to create and pilot the IBDP, the IBO seemed positioned to be able to respond to these tensions only with aspirational commitments to enlarge access and internationalize representativeness over time.

Analysts of IB in contemporary times will recognize that these tensions have endured, sometimes presented as critiques of IB. For example, lack of access to the IB programs continues to be a point of critique (Dickson et al., 2017). In response, the IB organization, continues to be actively and strategically engaged in broadening access to the IB, as I will detail below. Nevertheless, as the IB has moved through its phases of “creation and experiment” to “growth and sustainability” to “diffusion and diversification” to “branding and impact” over the last fifty years, dynamics of these tensions have altered (Tarc, 2009). Most obvious is that, for the most part, international understanding as an aim of education is no longer contentious; indeed, many governments advocate for it as a component of human capital development in a globalizing world (Green, 1997). For a more detailed analysis of the shifting dynamics of the tensions and how the IBO navigates them in a changing world, refer to the Global Dreams text (Tarc, 2009). In the remaining part of this **Introduction** section, I will outline the most significant continuities and discontinuities of these tensions precipitated by larger 21st century transformations and the attendant responses of the IB Organization.

First, amidst the ascendancy of neoliberal economic globalization in the 1990s with the breakup of the Soviet Union, international education begins to move from a potentially politically contentious and marginal activity to an expedient (Tarc, 2009, 2013). For example, governments want globally savvy/mobile citizens who can contribute to the national economy, universities seek out international students as a new generation stream under declining public funding, businesses want interculturally competent employees who can exploit niche markets globally and students want to build their resumes with international certificates and experience (Tarc, 2013). Often entangled with these pragmatic agendas of this neoliberal internationalization movement come the more idealist/aspirational agendas privileging the potential educational, cosmopolitan and ethical potentialities of international education (Tarc, 2019). In this sense the “citizenship tension” of IB has largely abated. For only a fringe right, admittedly energized under the recent rise in strongman populism (Geiselberger, 2017), does international education remain contentious (for the US context, see Bunnell, 2012). In terms of IB’s diffusion and acceptance into state-funded schooling, IB’s mission of developing international mindedness or global citizenship is either inconsequential or seen as an asset by schools and ministries of education also open to the internationalization trend. What this change means is that rather than trying to minimize its internationalist vocabulary, IB’s international, becomes a “value added,” a marker of distinction (Tarc, 2009); for upwardly mobile middle-class parents, cosmopolitan capital is increasingly recognized and pursued as a form of cultural capital (Forsey, 2017; Weenink, 2008).

Today, neither the vocabulary nor the liberal-humanist pedagogy of international education is contentious. However, the larger political tension of international education’s aims and uses remains. The expediency of international education is tied into nationalist agendas and strategic capital accumulation of mobile elites (Ong, 1999; Tarc, 2013). Nationalist internationalisms have long been critiqued (Leach, 1969) and remain dramatically present and problematically at odds with the ideals of equity, reciprocity and dialog founding ethical internationalist engagements. In this sense, the citizenship tension has merged with the operational tension (access and representativeness). For example, where IB is used to further social advantage by elites or where IB curricula remain Eurocentric and complicit with hierarchizing societal and human value, the aspirational (world) citizenship goal of (massifying) international understanding remains stunted.

The curricular tension has endured but with changing dynamics. First, the centrally examined IBDP remains a college preparatory degree and thus the tension remains between the development of the whole person through a general education and the standards or accountability mechanisms¹ necessary to facilitate access to top universities. However, through the development of the younger-years programs (MYP, PYP), IB has moved forward in realizing its progressive educational visions. These programs are less academically content-rigid and therefore tend to better support the kinds of inquiry-based, thematic, interdisciplinary, progressive and innovative approaches to which the IB brand aspires. Given that IB still must ensure quality standards over the younger years program, there still exist constraints on innovation. However, it remains arguable whether the IB programs or the (national) school and community contexts in which IB is enacted represent the “bottle neck” to realizing more progressive, internationally minded or innovatory pedagogies.

As for the curricular tension within the IBDP, some studies have shown that many IBDP students find the IBDP program to be a very intensive and stressful experience (Hertberg-Davis and Callahan, 2008); my colleague and I (Tarc and Beatty, 2012) found in one IB World school in Ontario, that some IBDP students had to limit or eliminate their extra-curricular activities in sports, arts and service in the school and community, to fulfill the academic requirements of the IBDP. Obviously, this sole focus on academics contradicts the goal of developing the whole person through a general education.

IB leadership is aware of this curricular tension in the IBDP and periodically have discussed alternatives to its high stakes culminating examination regime, but the role of the examination regime in assuring IB’s “high quality” standards has much inertia (Tarc, 2009). Most recently, the IBDP’s 2020 Spring examinations were canceled due to the Covid-19 pandemic. In a recent interview, DG Kumari discussed the need to learn from the pandemic and IB’s responses to it and specifically of “plans to shift focus away from end-of-program exams.” She is reported stating:

Before Covid, we were already designing our strategy for the next 10 years and as part of that we have been having these conversations about the end program ... where this heavy-duty summative experience does not reflect the real world anymore. (<https://www.tes.com/news/international-baccalaureate-siva-kumari-exams-future-education-coronavirus>, accessed December 8, 2020)

¹Albeit, as one anonymous reviewer noted, IBDP’s examinations maybe less “high stakes” than the local jurisdiction such as with many Australian systems that use “norm-based” external assessment and moderate internal assessment with the external scores. In contrast IB employs “criterion-based” assessment and does not adjust internal grade components.

It is possible that the pandemic pushes “these conversations” forward and catalyzes a change in the status quo. Perhaps the IB brand has a solid enough reputation to maintain its high regard with respect to the quality and oversight of the programs without the IBDP high stakes testing regime. Time will tell.

Another key shift over time has been the development of IB’s focus on internationalizing IB curricula—supporting schools to be “internationally minded,” beyond providing international schools with an internationally-recognized diploma. Initially the IBDP enabled national groups to study together on a single less nationalistic curriculum. International understanding was implicit to the milieu of multilateral international schools. In the mid-1980s as IB found financial sustainability via the diffusion of the IBDP into state schools in the United States in Canada, IBO realized that an international student body could no longer be assumed (IBO, 1988). From the early 1990s the IB took a more conscious focus on the goal of international understanding (and later international mindedness) and its integration into its curricula. In its current “branding and impact” phase in the 21st century, fostering international mindedness is an overarching goal. Although still tied to a set of progressive learner dispositions of “open mindedness” and “inquiry”² the IB organization has become more explicit about its aim of fostering international mindedness and what that entails (IBO, 2017). In this important mission document with just over 6 pages of content, a full page is afforded to international mindedness. Significant also are recently conducted IB-sponsored studies, specifically focused on educating for international mindedness (Hacking et al., 2018; Singh and Qi, 2013; Sriprakash et al., 2014). Thus, IB now projects a more “positive” definition of international education (Tarc, 2009), that *any* school can work at.

The operational tension also remains significantly on the radar in IB’s rhetoric and reform policy (Tarc, 2009). As would be expected national/cultural diversity has expanded greatly in the make-up of IB policy actors and employees. New working languages have been added to the original languages of English and French. Curricular modifications have incrementally opened-up the Euro-centric beginnings of IB programs; for example, in 2015 Indigenous ways of knowing was formally added to the “knowledge areas” in the IBDP Theory of Knowledge course. Also, of note, is a recent push from IB commissioned research toward “intellectual equality” and “multilingualism” (Sriprakash et al., 2014) to inform IB’s conception of international mindedness, as its overarching pedagogical ideal. The Western-centeredness of IB, which can be attributed to its foundations and from the still hegemonic global status of Anglo-Western education, must be assumed. However, my speculation, to repeat, is that the IB curricula itself (especially the more open PYP and MYP) is probably more open to epistemic diversity than are the (national) contexts of IB school classrooms and teachers. Thus, we might say that the tension of the inter-national representativeness of IBO and of IB curricula have at least diminished somewhat.

In terms of broadening access, the IB has expanded beyond the more elite private international schools of its foundational period and entered state schools in the West; it has developed more accessible younger years programs, entered inner-city schools in the United States and the United Kingdom and entered partnerships with state schools in the global South (Ecuador). The more recent provision of the “Career-related Program” also represents an important access-broadening initiative. Still, criticisms remain of IB as elitist, as inaccessible, and as (unintentionally) furthering educational inequality given its use as a choice option within neoliberal school reform. For its part, the IBO began to prioritize its commitments to broadening access to IB programs in its 21st century policy discourse (IBO, 2006; Tarc, 2009). Given the 21st century zeitgeist of inclusivity and equity and given IB’s mission to massify international education, concerns around access and equity remain key pressure points for the IBO today (see the “E2” initiative, IBO, 2018).

As they made explicit in their Growth to Access document: “Today, over 1/2 million students from all continents have graduated from our Diploma Program but it is obvious that our goal requires millions of people worldwide to benefit from an international education (IBO, 2006, p. 2). In this key policy document, IBO makes it clear that despite a variety of their broadening access initiatives over the years, there remains very uneven access to IB, both geographically and in terms of social class. The policy document also includes a set of strategies and plans to mitigate this uneven access; while the past decade has seen incremental and singular instances of broadened access, in the larger picture, limited and uneven access remains a core challenge for IB. For example, in Dickson et al. (2017) recent examination of the Australian context, they find that “whether private or public, IB schools in Australia are overwhelmingly located in higher-SES areas and enroll students from higher-SES backgrounds” (p. 75).

In his DG report of 1972, Alec Peterson directly responded to IBDP student charges that IB is elitist, Alec Peterson suggested that IB could foster “intellectual elitism” over “social elitism” (IBO, p. 17). Here he was signaling how one’s education and academic performance might overtake social status and familial wealth in a meritocratic society. Perhaps, at that historical moment, the feasibility of dis-entangling educational elitism from social elitism seemed more credible. Despite schooling reforms for equality, catalyzed by the “new sociology of education” of the 1970s, social class and educational attainment have remained tightly correlated; we now witness countless ways in which familial cultural capital and resources are applied to facilitate academic success for middle and upper class families. As an education for distinction (particularly the IBDP) built on a “user-pays” model (Tarc, 2009), IB can little extract itself from these larger conditions of schooling and social class stratification. Consequently, the very neoliberalizing conditions allowing new models, as IB programs, into state educational systems, also steer the uses of IB education (as cultural capital) for social class advantage; and, given IB’s user-/institution-pays model that produces uneven access to its product, it is thus difficult to imagine how the operational tensions can be resolved.

²The “IB Learner Profile” sets out the following 10 “attributes” that “IB learners ... strive to be,” as follows: inquirers, knowledgeable, thinkers, communicators, principled, open-minded, caring, risk takers, balanced and reflective (IBO, 2017, p. vii). The profile, originally developed in the context of the PYP, is a key instrument of IB’s branding.

Historicizing the structuring tensions of IB, as outlined above, illustrates how IB articulates within and alongside state educational systems and private independent schools embedded in national settings. It also shows how the IB has been somewhat proactive in the internationalization of education as well as reactive to its flows and pressures. These tensions of IB, invoked by the encounter of its aspirational dreams with practical realities (structured by larger forces) are also resonant with internationalizing schools and universities, navigating the limits and possibilities for (citizenship) education in an interdependent and asymmetric world. The next section turns to IB's relations to K-12 school internationalization.

IB as a window on school internationalization

IB's diffusion in state educational systems can also be studied as an instance of the internationalization of K-12 schooling. In her 2012 article, scholar Julia Resnik offers a most explicit analysis of how the IB's insertion and expansion in state schools can be interpreted as a process of de-nationalizing education. From its foundation, the IB was designed and adapted to articulate with state systems for educational and qualificatory alignments and recognition. For example, as the IBDP expanded in monolingual areas of North America, a beginning-level language course was added to its previously more demanding second language course requirement (Tarc, 2009). However, state educational systems have also been changing under globalization and not only as passive victims to exogenous global forces. Drawing on Sassen's scholarship (2000, 2003) uncovering how globalization is advanced within national spaces and by state actors, Resnik asserts that "the theoretical significance of IB schools ... is that they embody the denationalization of educational systems" (p. 249). Consequently, "certain national contexts and educational traditions encourage IB schools, while others hinder their propagation" (p. 249). I would add that since these contexts and traditions are also in flux, IB's propagation and prospective sustainability (see Beech and Guevara, 2020) within countries or educational jurisdictions also shift across time.

The "global" forces or actors interact with national systems, institutions and processes in what Sassen (2000) calls "frontier zones" as distinct "spatialities" embedded in the national (Resnik, 2012, p. 251). Resnik offers a kind of spatial typology that is useful in breaking apart the different levels of global-local interactions that can be considered part of k-12 school internationalization. She lists them, in ascending order of *thickness of the global*: (1) the IB international brand, (2) the International Baccalaureate Organization (IBO), (3) IB regional offices, (4) international schools for mobile families, (5) private schools that recruit local children, and (6) public schools that recruit local children. (my emphasis, p. 251)

IB as brand is the most global, symbolic and least tethered to a national physical space, materialized with the IB logo stamped on all things IB. The IBO is a transnational entity, made up of individuals of different nations and working with multiple states; yet, it has a central headquarters in geographical space. The least global is the state school serving "local" children of which the "global" of IB has very little presence in relation to the totality of the school operations and processes. In terms of focusing on school internationalization within state systems over the "global" of IB, it is helpful to invert Resnik's ordering. Under this inversion the thickest form of school internationalization is represented by the presence of IB programs in the "local" public school, where students' learning and subject formation is (at least subtly) being shaped by a supranational entity. And the thinnest and most symbolic manifestation, but much more widespread, is the IB brand recognized by wider national publics without tangible experience with the IB programs *per se*. The recent instituting of IBEC partnerships with state teacher education programs, mentioned above, also represents a new and somewhat thick form of school internationalization in the de-nationalized frontier zone of the global (IB) embedded in the national (here, teacher education). Thus, internationalization of K-12 schooling, as de-nationalization, is thickest where students receive a compulsory, state-funded education in the local school provided by a supranational entity. However, given the very small percentage of IB schools and IB students (and educators) within state educational systems, the breadth of this thickest level of de-nationalizing seems extremely narrow. On the other end, the thinner forms reach a wider audience and influence/represent internationalization (as de-nationalization) in more symbolic ways that are admittedly playing a role but also more difficult to trace.

A recent article by Tristan Bunnell takes a more "skeptical" view of IB's role and influence in "the internationalization of public schooling in practice" (p. 56), emphasizing just how narrow this thickest form may be. In his article, he "seek[s] to show that in practice the extent of contact between the IB and public schooling is relatively scarce, small scale and minimally funded and prioritized" (p. 57). First, Bunnell emphasizes the uneven distribution, where public schooling activity is predominantly located in a minority of countries and in clusters of only "urban settings" within these countries; the government-funded growth in Ecuador is a case in point (p. 60–61). Further, while there are private international schools where a majority of students take the IBDP, in many publicly funded schools in the other top three countries with public IB schools (United States, Canada and Australia), a minority of students are enrolled in the IBDP; for example, sometimes there exists a small cohort of 20–25 students taking the IBDP within a school of one thousand students (Tarc, 2009). In the case of Ecuador, its 270 IBDP public schools had an average of 22 examination candidates per school (Bunnell, p. 62). Thus, the IB brand(ing), with its "5000+ IB World schools in more than 150 countries," belies the relatively small number of IB students and teachers in public schools and the very small numbers of IB schools in many of these 150 countries.

Bunnell concludes his skeptical framing by considering state funding and IB's reliance on "political champions." There are a few countries, as Ecuador, Japan and the United States "led by an IB-government assemblage" (p. 63), indicative of a distinctive spacialization of de-nationalization (Resnik, 2012). However, Bunnell (2020) continues, "very few governments directly support the IB, and it tends to involve relatively small grants of money" (p. 63). In the countries of the UK and the US, which have supported the adoption of IB programs through government funding, "funding is usually merely to cover the basic costs of applying for the

accreditation process, i.e., there is no long term funding available" (p. 63). Finally, IB's diffusion has continued to rely on individual influential contacts (Tarc, 2009) including, more recently, political champions (UK, City of Chicago, Ecuador). Where support is contingent on individuals and political outcomes, such support may not be sustainable.

Thus, Resnik (2012) illustrates dynamics of the denationalization of education via the propagation of IB in state schooling, across its thinner and thicker types; whereas, Bunnell (2020) provides cautionary statistics on the actual depth, scope and accessibility of the thickest type. Taken together we can see that the IB is indeed emblematic of transnationalizing processes, but the depth of these processes is questionable. My sense is that it is useful to examine all the types/levels in Resnik's typology and particularly their complex interactions, as they work together to produce denationalizing zones that constitute school internationalization. The IB branding works dynamically with the IBO's regional offices and the concrete manifestations of IB programs, to heighten the prominence and effect of IB. And while actual numbers of IBDP exam candidates, IB students and IB teachers remain low in relation to the "buzz" of IB, these actors have relations across multiple domains. For example, anecdotally, IB teachers talk about how teaching IB has positively affected their teaching in non-IB classes.

Both Resnik (2012) and Bunnell (2020) articulate such positive potential secondary effects. Resnik discusses "percolation" as the influence of aspects of the DP program ... on curricula and programs that are not related to the "IB" (p. 263). She provides examples from two state-sponsored schools in London, England that adapt the IBDP's innovatory core elements to provide critical thinking courses, extended inquiry projects and volunteering to the larger majority of non-IB students in the schools. These percolatory effects are particularly significant given IB's "horizontal networks of governance," where "teachers are trained by other teachers ... and a range of mechanisms for promoting the exchange of know-how and experiences among practitioners have been fostered" (Beech and Guevara, 2020, p. 104). This horizontality of governance and teachers' professional development explains why pockets of IB schools within certain geographies can be found, as in the country of Ecuador and in the city of Chicago, and why the IBO seeks to develop these close clusters of schools. It also explains how *private* IB schools are actors in the assemblage of school internationalization in the *public* sector. Additionally, within the public sector, this horizontality and teacher collaboration may percolate into non-IB schools under board level initiatives, including non-IB professional development and educator networks.

Bunnell also elaborates on this "ripple" effect by discussing research on IB public schools in Spain, Ecuador and Japan. In summary these (potential) effects included: energizing learning and extra-curricular activities and positive school culture for non-IB students, showcasing high academic approaches, model pedagogies and internationally minded educational approaches for other programs, schools and system-wide reforms (p. 64–65). However, not all ripple or secondary effects are positive. Researchers have also raised concerns of the negative secondary effects of bringing IB into state educational systems. These potential negative effects, which connect to the stratifying uses of IB under neoliberal privatizing reform already discussed, include: the potential negative effects of "IB-choosing" students exiting from local schools (Lauder, 2007); the funneling of resources to the already privileged mobile (Doherty, 2013; Tarc, 2009) and urban (Bunnell, 2020) middle classes; and the siloing of IB and non-IB student groupings (Culross and Tarver, 2007).

On the one hand, from a state schooling perspective, IB is implicated in the de-nationalizing of education that, in recent decades, have been challenging idealizations and practices of schooling as a territorially bounded activity of the sovereign nation-state. At the very least, the propagation and growth of IB into state educational systems represents a window on these de-nationalizing processes constituted by the mix of transnational (educational) policy forces and flows, national/ministerial sovereignties and (flexible) citizenships inside, and stretching across, borders. In some respects, IB is itself a transnational force in school internationalization, albeit with its program provision having limited scope, depth and accessibility (Bunnell, 2020).

From a more multilateral perspective, on the other hand, the IB is a salient exemplar of international education and, analytically, offers a window on how ideals and manifestations of an education for international understanding (toward a more peaceful, egalitarian world) have evolved across the last fifty years. Such analysis provides insights into the trends and prospects for international education and its variants under contemporary conditions. Both perspectives reveal how IB as implicated in the globalizing geopolitics of education.

Conclusion

To conclude, "IB" is a longstanding, multi-faceted and study-worthy phenomenon as a particular manifestation of international/progressive education, as a window on the shifting meanings and uses of (international) education in global times and as an enabler and manifestation of school internationalization. IB is a compelling exemplar of the heightened expediency of international education in the 21st century (Tarc, 2009, 2019). For many students and families, IB has proven to be an enriched and value-added educational program (Tarc et al., 2019). The IB organization has navigated relatively successfully across decades of global transformations, beyond its own making, to govern its core mandate of providing an IB diploma on a school-by-school basis. It has also expanded its activities to include younger-years and career-related programs and to seek out other initiatives and partnerships to "create a better world" through education. How IB has navigated, points to the larger conditions of globalization that have also animated the trend of internationalization of K-12 schooling. Concurrently, IB is also an agent in school internationalization across its different levels of embeddedness in national schooling.

IB's continued success raises several new and old questions that will require re-thinking and negotiation as societies and institutions respond to the regressive and potentially progressive forces catalyzed by the Covid-19 pandemic and its still uncertain aftermath. On the one hand, there will be pressures for IB to "up" its (discursive) commitments to social justice/anti-oppressive and

environmentalist pedagogy in the wake of the Black Lives Matter movement and given global ecological crises. A turn to anti-racist pedagogies, for example, would represent a shift given IB's emergence from, and niche uses in, the traditionally "color-blind" elite international schools culture (for example, David, 2020). On the other hand, the threat to liberal internationalism by the rise of populist nationalisms, border closings and trade-wars might mean that IB's viability finds more traction in the accrual of academic distinction and capital, over its promise of a humanist international mindedness or even the facilitation of global mobility. As with other educational providers, IB will have to respond to a confluence of social justice desires, the pedagogical needs of the learner (21st century learning) and the continued dominance of neoliberal performativity (Tarc, 2015). My sense is that, programmatically, it will continue to stick with its core foundation as a provider of a learner-centered liberal-humanist education, while, adapting rhetorically, to the wider conditions and cultural politics in a globalizing (and de-globalizing) world.

References

- Bagnall, N., 1994. *The International Baccalaureate in Australia and Canada: 1980–1993*. University of Melbourne, Melbourne.
- Beech, J., Guevara, J., 2020. Multiple internationalizations: the idiosyncratic enactment of the International Baccalaureate in State Schools in Costa Rica, Peru and Buenos Aires. In: Engel, L.C., Maxwell, C., Yemini, M. (Eds.), *The Machinery of School Internationalization in Action: Beyond the Established Boundaries*. Routledge, New York, pp. 103–118.
- Bunnell, T., 2008. The global growth of the International Baccalaureate Diploma Program over the first 40 years: a critical assessment. *Comp. Educ.* 44 (4), 409–424.
- Bunnell, T., 2012. *Global Education Under Attack: International Baccalaureate in America*. Peter Lang, Frankfurt.
- Bunnell, 2020. The "internationalization of public schooling" in practice: a "skeptical reality" approach. In: Engel, L.C., Maxwell, C., Yemini, M. (Eds.), *The Machinery of School Internationalization in Action: Beyond the Established Boundaries*. Routledge, New York, pp. 56–69.
- Cambridge, J., Thompson, J., 2004. Internationalism and globalization as contexts for international education. *Compare* 34 (2), 161–175.
- Charleson, C.T., 2010. Book review: global dreams, enduring tensions: International Baccalaureate in a changing world. *J. Res. Int. Educ.* 9 (3), 330–334.
- Culross, R.R., Tarver, E.T., 2007. Teacher and student perceptions of the International Baccalaureate program: a first year perspective. *J. Sch. Choice* 1 (4), 53–62.
- David, X., 2020. Decolonize IB: How International School Alumni Are Mobilizing to Diversify the Expat Curriculum. Retrieved August 05, 2020, from: <https://medium.com/@xoidavid/decolonise-ib-how-international-school-alumni-are-mobilising-to-diversify-the-ecat-curriculum-cf3471816fa6>.
- Dickson, A., Perry, L.B., Ledger, S., 2017. How accessible is IB schooling? Evidence from Australia. *J. Res. Int. Educ.* 16 (1), 65–79. <https://doi.org/10.1177/1475240917696037>.
- Doherty, C., Mu, L., Shield, P., 2009. Planning mobile futures: the border artistry of International Baccalaureate Diploma choosers. *Br. J. Sociol. Educ.* 30 (6), 757–771.
- Doherty, C., 2009. The appeal of the International Baccalaureate in Australia's educational market: a curriculum of choice for mobile futures. *Discourse* 30 (1), 73–89.
- Doherty, C., 2012. Optimizing meritocratic advantage with the International Baccalaureate Diploma in Australian Schools. *Crit. Stud. Educ.* 53 (2), 183–196.
- Doherty, C., 2013. Making a point of difference: the glocalized ecology of the International Baccalaureate Diploma in Australian Schools. *Glob. Soc. Educ.* 11 (3), 379–397.
- Drake, B., 2004. International education and IB programs: worldwide expansion and potential cultural dissonance. *J. Res. Int. Educ.* 3 (2), 189–205.
- Dvir, Y., Shields, R., Yemini, M., 2018. Three faces of global citizenship education: IB schools' self-representations in four local contexts. *Br. J. Educ. Stud.* 66 (4), 455–475.
- Elvin, H.L., 1960. Nationalism and internationalism in education. *Sch. Rev.* 68 (1), 1–22.
- Engel, L.C., Maxwell, C., Yemini, M. (Eds.), 2019. *The Machinery of School Internationalization in Action: Beyond the Established Boundaries*. Routledge.
- Forsey, M., 2017. Education in a mobile modernity. *Geogr. Res.* 55 (1), 58–69.
- Fox, E., 1985. International schools and the International Baccalaureate. *Harv. Educ. Rev.* 55 (1), 53–69.
- Gardner-McTaggart, A., 2016. International elite, or global citizens? Equity, distinction and power: the International Baccalaureate and the rise of the South. *Glob. Soc. Educ.* 14 (1), 1–29.
- Geiselberger, H. (Ed.), 2017. *The Great Regression*. Polity Press, Cambridge.
- Good, K.D., 2020. *Bring the World to the Child: Technologies of Global Citizenship in American Education*. MIT Press, Cambridge, MA.
- Green, A., 1997. *Education, Globalization, and the Nation State*. St. Martin's Press, New York.
- Hacking, E.B., Blackmore, C., Bullock, K., Bunnell, T., Donnelly, M., Martin, S., 2018. International mindedness in practice: the evidence from International Baccalaureate schools. *J. Res. Int. Educ.* 17 (1), 3–16.
- Hahn, A.M., 2003. *Intersection of Language, Power and International Education: A Critical Discourse Analysis of the International Baccalaureate Organization*. Unpublished Ph.D. Columbia University, New York.
- Hayden, M.C., Wong, C.S., 1997. The International Baccalaureate: international education and cultural preservation. *Educ. Stud.* 23 (3), 349–361.
- Heater, D.B., 1980. *World Studies: Education for International Understanding in Britain*. Harrap, London.
- Hertberg-Davis, H., Callahan, C.M., 2008. A narrow escape: gifted students' perceptions of Advanced Placement and International Baccalaureate programs. *Gift. Child. Q.* 52 (3), 199–216.
- Hill, I., 2002a. The International Baccalaureate: policy process in education. *J. Res. Int. Educ.* 1 (2), 183–211.
- Hill, I., 2002b. The history of international education: an International Baccalaureate perspective. In: Hayden, M., Thompson, J., Walker, G. (Eds.), *International Education in Practice: Dimensions for National & International Schools*. Kogan Page, London, pp. 18–35.
- Hill, I., 2006. Do International Baccalaureate programs internationalize or globalize. *Int. Educ. J.* 7 (1), 98–108.
- Hill, I., 2007. International education as developed by the International Baccalaureate Organization. In: Hayden, M., Levy, J., Thompson, J.J. (Eds.), *The Sage Handbook of Research in International Education*. Sage, London, pp. 25–37.
- Hill, I., 2012. Evolution of education for international mindedness. *J. Res. Int. Educ.* 11 (3), 245–261.
- International Baccalaureate Organization, 1972. *IBO Annual Bulletin No. 8*. IBO, Geneva.
- International Baccalaureate Organization, 1988. *IBO Annual Bulletin*. IBO, Geneva.
- International Baccalaureate Organization, 2006. *From Growth to Access: Developing an IBO Access Strategy*. Retrieved May 24, 2007, from: ibo.org/mission/strategy/documents/FromGrowthtoAccess.pdf.
- International Baccalaureate Organization, 2017. *What is an IB Education?* International Baccalaureate Organization, Cardiff, Wales. Retrieved September 3, 2020, from: <https://ibo.org/globalassets/what-is-an-ib-education-2017-en.pdf>.
- International Baccalaureate Organization, 2018. *Excellence and Equity Initiative*. Retrieved December 15, 2020, from: ibo.org/excellence-and-equity.
- Kidson, P., Odhiambo, G., Wilson, R., 2019. The International Baccalaureate in Australia: trends and issues, compare. *J. Comp. Int. Educ.* 49 (3), 393–412. <https://doi.org/10.1080/03057925.2017.1415751>.
- Kyburg, R.M., Hertberg-Davis, H., Callahan, C.M., 2007. Advanced Placement and International Baccalaureate programs: optimal learning environments for talented minorities? *J. Adv. Acad.* 18 (2), 172–215.
- Lauder, H., 2007. International schools, education and globalization: towards a research agenda. In: Hayden, M., Levy, J., Thompson, J.J. (Eds.), *The Sage Handbook of Research in International Education*. Sage, London, pp. 441–449.

- Leach, R.J., 1969. *International Schools and Their Role in the Field of International Education*. Pergamon Press, Oxford.
- Lee, M., Wright, E., 2016. Moving from elite international schools to the world's elite universities. *Int. J. Comp. Educ. Dev.* 18 (2), 120–136.
- Lineham, R., 2013. Is the International Baccalaureate diploma program effective at delivering the International Baccalaureate mission statement? *J. Res. Int. Educ.* 12 (3), 259–282.
- Malinowski, H.W., Zorn, V., 1973. *The United Nations International School: Its History and Development*. United Nations International School, New York.
- Mayer, M., 1968. *Diploma: International Schools and University Entrance*. Twentieth Century Fund, New York.
- Mayer, A.P., 2008. Expanding opportunities for high academic achievement: an International Baccalaureate Diploma Program in an Urban High School. *J. Adv. Acad.* 19 (2), 202–235.
- Méras, E.A., 1932. World-mindedness. *J. High Educ.* 3 (5), 246–252.
- Ong, A., 1999. *Flexible Citizenship: The Cultural Logics of Transnationality*. Duke University Press, New York.
- Peterson, A.D.C., 1972. *The International Baccalaureate: An Experiment in International Education*. G.G. Harrap, London.
- Peterson, A.D.C., 1987. *Schools Across Frontiers: The Story of the International Baccalaureate and the United World Colleges*. Open Court, La Salle, IL.
- Poelzer, G.H., Feldhusen, J.F., 1997. The International Baccalaureate: a program for gifted secondary students. *Roeper Rev.* 19 (3), 168–171.
- Renaud, G., 1974. *Experimental Period of the International Baccalaureate: Objectives and Results*. UNESCO, Paris.
- Resnik, J., 2012. The denationalization of education and the expansion of the International Baccalaureate. *Comp. Educ. Rev.* 56 (2), 248–269.
- Resnik, J., 2019. Struggling for recognition: access to higher education through the International Baccalaureate. *Crit. Stud. Educ.* 60 (3), 340–357.
- Rizvi, F., Savage, G.C., Quay, J., Acquaro, D., Sallis, R.J., Sobhani, N., 2020. Transnationalism and the International Baccalaureate learner profile. *Prospects* 48 (3), 157–174.
- Rizvi, F., 2014. *Encountering Education in the Global: The Selected Works of Fazal Rizvi*. Routledge, New York.
- Sassen, S., 2000. Spatialities and temporalities of the global: elements for a theorization. *Publ. Cult.* 12 (1), 215–232.
- Sassen, S., 2003. Globalization or denationalization? *Rev. Int. Polit. Econ.* 10, 1–22.
- Scott, D., 2004. *Conscripts of Modernity: The Tragedy of Colonial Enlightenment*. Duke University Press, Durham, NC.
- Singh, M., Qi, J., 2013. 21st century international mindedness: an exploratory study of its conceptualization and assessment. Centre for Educational Research School of Education, Australia, University of Western Sydney.
- Sriprakash, A., Singh, M., Jing, Q., 2014. A Comparative Study of International Mindedness in the IB Diploma Program in Australia, China and India. International Baccalaureate Organization, Cardiff, Wales. Retrieved from: <http://www.ibo.org/globalassets/publications/ib-research/dp/international-mindedness-final-report.pdf>.
- Stier, 2004. Taking a critical stance toward internationalization ideologies in higher education: idealism, instrumentalism and educationalism. *Glob. Soc. Educ.* 2 (1), 1–28.
- Tarc, P., Beatty, L., 2012. The emergence of the International Baccalaureate Diploma in Ontario: diffusion, pilot study and prospective research. *Can. J. Educ.* 35 (4), 241–275.
- Tarc, P., Mishra Tarc, A., Ng-A-Fook, N., Triolokekar, R.D., 2012. Re-conceiving international education: theorizing limits and possibilities for transcultural learning. *Can. Int. Educ. J.* 41 (3), 6.
- Tarc, P., Mishra Tarc, A., Wu, X., 2019. Anglo-Western international school teachers as global middle class: portraits of three families. *Discourse* 40 (5), 666–681. <https://doi.org/10.1080/01596306.2019.1570635>.
- Tarc, P., 2007. What is the “international” of the International Baccalaureate? Towards a Periodization of IB in the World. ProQuest.
- Tarc, P., 2009. *Global Dreams, Enduring Tensions: International Baccalaureate in a Changing World*. Peter Lang, New York.
- Tarc, P., 2011. How does “global citizenship education” construct its present? The crisis of international education. In: Andreotti, V., Menezes de Souza, L.M. (Eds.), *Postcolonial Perspectives on Global Citizenship Education*. Routledge, London, pp. 105–123.
- Tarc, P., 2013. The rise of international education: expanded opportunities, new complications. In: Tarc, P. (Ed.), *International Education in Global Times: Engaging the Pedagogic*. Peter Lang, New York, pp. 1–18.
- Tarc, P., 2015. What is the “active” in 21st century calls to develop “active global citizens”? Justice-oriented desires, active learning, neoliberal times. In: Harshman, J., Augustine, T., Merryfield, M. (Eds.), *Research in Global Citizenship Education*. Information Age Publishing, Charlotte, NC, pp. 35–58.
- Tarc, P., 2019. Internationalization of education as an emerging field of study? A conceptualization of international education for cross-domain analyses. *Pol. Futures Educ. Internet* 17 (6), 732–744.
- Tarc, P., 2021. Transnational Governing for the Pedagogical Ideals of K-12 International Education: contrasting PISA and IB. *Educ. Rev.*
- Twigg, V.V., 2010. Teachers' practices, values and beliefs for successful inquiry-based teaching in the International Baccalaureate Primary Years Program. *J. Res. Int. Educ.* 9 (1), 40–65.
- van Oord, L., 2007. To westernize the nations? an analysis of the International Baccalaureate's philosophy of education. *Camb. J. Educ.* 37 (3), 375–390.
- Waterson, M., 2016. The corporatization of international schooling. In: Hayden, M., Thompson, J. (Eds.), *International Schools: current Issues and Future Prospects*. Symposium Books, Oxford, pp. 185–213.
- Weenink, D., 2008. Cosmopolitanism as a form of capital: parents preparing their children for a globalizing world. *Sociology* 42 (6), 1089–1106.
- Wells, J., 2011. International education, values and attitudes: a critical analysis of the International Baccalaureate (IB) Learner Profile. *J. Res. Int. Educ.* 10 (2), 174–188.
- Wright, E., Lee, M., 2014. Developing skills for youth in the 21st century: the role of elite International Baccalaureate Diploma Program schools in China. *Int. Rev. Educ.* 60 (2), 199–216.
- Wright, E., Lee, M., 2019. Re/producing the global middle class: International Baccalaureate alumni at “world-class” universities in Hong Kong. *Discourse* 40 (5), 682–696.

Relevant website

ibo.org.

Low-fee private schools

Geoffrey Walford^{a,b}, ^a Department of Education, University of Oxford, Oxford, United Kingdom; and ^b Green Templeton College, University of Oxford, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	355
Definitions and the complexity of reality	355
The development of low-fee private schools	356
Factors leading to low-fee private schools	358
Over-supply of teachers	358
Overall poor performance of the public sector	358
Hidden costs and private tuition costs	360
Language of instruction	360
Access and affordability	360
Low-cost private school chains and international investment	361
Conclusion	363
Acknowledgments	363
References	363

Introduction

The private sector of schooling has always been highly diverse. While much of the academic research and public discussion about private schools has long been linked to issues related to privilege and elitism ([Green and Kynaston 2019](#); [Peel, 2015](#); [Walford, 1990](#)), in practice, this focus has always been misleading. In most countries the private sector is very varied - while the best known schools may well be highly selective, expensive, and likely to lead to high-status universities, there have always been many other private schools that are far more modest.

It is worth remembering that historically all schools were originally private schools. Schools owned, funded, managed, and regulated by the state are relative newcomers in most countries. Generally, just two centuries ago it would have seemed a needlessly expensive and potentially dangerous activity to provide schooling for the poor, but now free universal basic education is perceived as a human right ([Aubry and Dorsi, 2016](#)).

Over the last few decades diversity within the private sector has been growing, and there has been an unexpected increase in private schooling in developing countries. Countries as varied as China, Mongolia, India, Nigeria, Uganda, and Vietnam have all seen a dramatic growth in the private sector of schooling - particularly at the primary level, but also at secondary. The schools themselves are very wide-ranging in nature. Some are very expensive and elitist schools for the wealthy or for ex-patriots living within these countries, but a major part of the growth of private schools has been those designed for local, often poor, families. The schools with the lowest fees may charge only a few US dollars per month, and have come to be known as “low-fee” or “low-cost” private schools. These schools are aimed at a target market of some of the poorest families in each of the countries, providing schooling for children whose families may be earning a few dollars per day. Where there has been a demand for private schooling a range of individuals, small groups, charities, and non-governmental organizations have stepped in to provide schooling for those who cannot afford the high fees of the elite schools, but are able and prepared to spend a significant proportion of their income on lower-fee private schools.

Definitions and the complexity of reality

There is no internationally accepted definition for “low-fee private schools.” This is not unexpected as there are no internationally accepted definitions for any of the four words that make up that phrase. Although many, even within the worlds of academia and of international development, use these terms as if they are internationally agreed, all four words represent simplifications of the reality that is to be found in the complex and ever changing world in which we live.

The concept of a “school” usually goes without question, but various countries have different ideas about what is to count as a school. There are obvious questions about how many children are involved, how old, for how long, for what proportion of the day or year, what subjects are taught, who does the teaching, and so on. Now that learning from the internet is ubiquitous, the idea that students need be gathered together in a physical space called a “school” comes under examination, and homeschooling has become a reality for many. When does private tutoring become a school? In terms of recognition and regulation, different countries, and sometimes regions within countries, have made different decisions about what should count as a school - which are totally ignored in most studies in comparative education.

But the problems of definition become far more complex when considering the term “private.” Private might refer to ownership, funding, management, or regulation of the schools. Religious organizations have been centrally involved in private provision in many countries for centuries. In many more-economically-developed countries they were the backbone on which schooling developed, and have become so much a part of the educational landscape that they receive substantial financial support from the state. In the UK, for example, the buildings of many religious schools remain in ownership of religious organizations, but most religious schools receive full per-pupil funding from the government, and are highly regulated by the state. In the Netherlands the majority of schools are officially private, but they receive full funding and even their buildings are provided by the state. This means that the distinction between private and state schools is often blurred, and whether a school is defined as being “private” or “government” depends on the history and politics of each country.

Similar problems can be seen in the words “low” and “fee.” What may be considered low in terms of international comparisons, may not be considered low by the families of the children attending these schools where they represent a considerable proportion of their income. For them, the fees are all they can or wish to afford. “Low” is operationalized differently in various academic research reports. Thus, for example, [Srivastava \(2006: 498\)](#) defined it as a fee “not exceeding about one day’s earnings of a daily wage laborer at the primary and junior levels, and two days’ earnings at the high school and intermediate levels”. Other studies have defined “low” in relation to state-enforced minimum wages, or a “poverty line” of US\$1.25 or US\$2.00 per person per day ([Tooley and Longfield, 2016](#)), or as a proportion of the country’s per-capita income, or per-pupil expenditure in government schools ([Kingdon, 2020: 13](#)). Any definition is ultimately arbitrary, for the schools are simply at the lower end of the fee scales for private schools in any particular country. Government statistics rarely take account of any such definitions and there is no evidence that parents, teachers, or owners of schools do so either.

What is to be included in the term “fee” is also contentious and some writers have used the term low-cost rather than low-fee to try to deal with this. Some use the word “cost” to indicate that families have to pay more than just the stated school’s fee for their children to attend (e.g. [Tooley and Longfield, 2016](#)). They may have to pay for books, school uniforms, examination entrance fees, transport, and other extras. For older children there may also be the additional cost of forgone earnings, as children often make major contributions to family’s income. The use of the term “cost” also allows metaphorical costs, such as the safety costs of lengthy travel, to be more easily included. For this type of low-cost, the schools are usually geographically near to the home, so may be considered to be safer for children (especially for girls) than a school at a greater distance. Other academics (for example, [Lewin, 2017: 82](#)) use the term “cost” to indicate a totally different consideration, where the cost is the per-pupil cost of providing schooling by the owner rather than the larger fee (or price) paid by the parents. Thus the profit made by the owner is emphasized, along with the reasons why schools are able to charge low fees.

In his edited study of India, Nigeria and Uganda for the Commonwealth Secretariat, [Phillipson \(2008: 1\)](#) adopts a very narrow definition of low-fee private schools that is described by what he does not consider to be encompassed by the category as much as what is to be included. He states:

It is not a school run by a non-governmental organization for charitable or developmental purposes. It is not a school run by a religious organization for the furtherance of a particular set of moral values and beliefs. It is not a school offering an educational advantage to its pupils and charging a high price for the privilege of gaining access to it. Finally, it is not a school set up by the local community until the government agrees to take over ownership. In contrast to these distinctions, the low cost private school is a school that has been set up and is owned by an individual or individuals for the purpose of making a profit.

This narrow definition is unhelpful, and it is not even consistently applied within the book itself. Not only are these distinctions not clear, in that there may well be multiple reasons for starting and continuing to run a school for the poor, the exclusion of non-governmental organizations (especially small, local NGOs), religious organizations, and those who might eventually wish to obtain some state funding, omits from consideration a huge part of the growth in low-fee private schooling. It also restricts our understanding of why such schools might be started and how the schools themselves, and the motivations for their continued existence, may change in nature over a period of time. Even the idea that they must “make a profit” collapses within the complexity of individuals and groups paying themselves salaries, or establishing schools so that they might gain employment. The vast majority of these new private schools were not (at least initially) the result of shareholders investing money in schools because they saw them as the way to obtain the highest financial return. The reasons for starting and continuing with the schools are much more complex, and there is thus the need to consider the whole range of non-government sector schools with low fees that are designed to serve some of the poorest families in each society.

Any generalizations about low-fee private schools, even within a single country, should thus be treated with great caution (including within this article). Readers of research on low-fee private schools need to carefully check the definitions being used and the sources of information cited. There are many cases in the literature where researchers do not adhere to their own definitions, and where the data sources are inadequate to support the claims made.

The development of low-fee private schools

Within the vast majority of countries there has been a range of private schools with fees from very expensive to relatively low fees for many centuries. A good number of major private schools in England were originally established in the 15th and 16th centuries as

schools for “poor scholars” in the local area. While there is some doubt about who was defined as a “poor scholar” (Orme, 2008), these schools certainly provided some students with free schooling. Meanwhile most children from poorer families had to make do with dame schools and other local fee-paying provision. Many children had little schooling, but church and other private organizations provided a basic level of schooling to many others. It was only as late as 1870 that the British government began to build a national system of schools and, even then, its initial purpose was to fill the gaps in private and charitable provision rather than to provide schools for all.

Tooley (2009: 212–234) argues that in the early nineteenth century India had a somewhat similar system of local private village schools which it was estimated in 1826 provided for about 25% of boys. With the addition of home schooling, about a third of boys were being schooled and, most interestingly, many of these boys were from lower caste families paying very low fees. These schools were later replaced by European-style schools which, in 1931, Mahatma Ghandi argued were too expensive for the poor, leading to the result that “today India is more illiterate than it was fifty or a hundred years ago” (quoted by Tooley, 2009: 212). Ghandi called for a revival of this type of low-fee private school in India, and Tooley and others found what might be seen as their reincarnation in the low-fee private schools in both rural and urban areas in India and in other countries in the late 1990s and 2000s.

Although the growth in low-fee private schools probably started in the 1980s, it was in the late 1990s and early 2010s that academics and others gradually became aware of the rapid increase in school designed to provide places for poor students in the less-economically-developed world. A range of reports, papers, and books such as those by Majumdar and Vaidyanathan (1995), Kitaev (1999), The PROBE Team (1999), and Tooley (2001), and those involved with international development policy used a variety of terms to describe these schools such as “budget schools” (Tooley and Dixon, 2005), “schools for the poor,” “new private schools,” and so on. The term “low-fee private schools” was first used to describe these schools by Srivastava (2006). The phrase has been widely used since but, as discussed above, has not gone unchallenged. As the role of these schools in educational provision for the poor is highly contested (discussed below) the terminology used by the various stakeholders varies so as to emphasize particular aspects of the nature of the schools.

It is important to recognize that the growth of low-fee private schools took place in parallel with the development and some progress toward the Millennium Development Goals for education where free basic education for all children was seen as a Human Right. Their growth might be seen as a reaction to the failure of various countries to provide schooling of an appropriate quality and effectiveness, and various authors saw them as vital part of reaching Millennium Goals. Now the focus is on Sustainable Development Goals and the extent to which low-fee schools might be promoted and supported by international aid, other funders, and by the governments of these less economically developed countries to achieve those Goals. This is a highly politically contentious area, and most researchers in the field are probably either advocates or critics of the involvement of non-state actors in education, especially if those actors wish to make a profit.

The study of such schools became particularly important when it was argued that low-fee private schools were playing an important part in achieving *Education for All* targets and the Millennium Development Goals. For example, Tooley and Dixon (2007), Tooley et al. (2007) suggested that the many low-fee private schools that were practically unknown to the governments of, for example, Ghana, India, Nigeria, and Pakistan, were making a substantial contribution toward meeting those targets. In contrast, others such as Lewin (2007), argued that the contribution that private schools made to the achievement of *Education for All* was small. But the work of such critics often relied on government statistics and interviews with high ranking administrators neither of which, as Tooley (2009) and others described, usually knew very much about the existence or nature of low-fee private schools in their own countries. One of the themes throughout Tooley’s book is the way that government officials, university academics, development advisors, and educationalists continually denied the existence of such schools in their countries in the early 2000s yet, by going into the slums and poor districts of so many developing countries, researchers were able to locate very many such schools. Many parents had rejected the schools provided by the government (where they were provided) and decided to devote a large proportion of their incomes (but a small actual sum) to the fees for local private schools. Most of these schools were entrepreneurial schools in the sense that they had to make enough money for the owner (who who might also be the Principal) to survive, yet there was usually more to the establishment of such schools than simple profit-making or earning a living. Tooley (2009) gives many examples where wishing to help the local community was as much a part of the motivation for starting a school as increasing personal income, and this is shown by the free or subsidized places that are sometimes given to orphans or families without any means to pay. He also describes the way that an initial plan to run a kindergarten gradually expands into a full primary school as parents encourage the owner to extend the age coverage year by year as their children get older. Owners respond not only because they see the chance to increase their income, but because they see a need in their local community that they feel they can meet. This same sort of mixture of motives was found by Srivastava (2007a, 2008a) in her study of low-fee private schools in Uttar Pradesh. The role of local philanthropy should not be excluded.

There has now been considerable research on low-fee private schools in the developing world, and the existence of these schools has become more recognized. However, even recent studies have shown that there is a need to systematically search the streets in some areas to locate schools: the administrative authorities simply do not have any records (e.g. Härmä (2016b) in rural Nigeria). There has been a particular concentration of research on India where many separate studies have been conducted. This includes work by Chudgar and Creed (2016), Härmä (2009, 2011), Kingdon (2007, 2008, 2020), Muralidharan and Kremer (2009), Shukla and Joshi (2008), Srivastava (2007 a,b), Tooley and Dixon (2006), Tooley (2007), and Tooley et al. (2007). But there has also been much work conducted in other developing countries such as Argentina (Moschetti and Snaider, 2019), Colombia (AlWindi (2015), Ghana (Akyeampong, 2009), Kenya (Edwards et al., 2017), Liberia (Dixon and Humble, 2019), Nepal (Bhatta, 2014), Nigeria (Umar, 2008; Rose and Adelabu, 2007), Malawi (Chimombo, 2009), Mozambique (Härmä, 2016a), Uganda (Kisira, 2008),

Pakistan (Fennell, 2013), and Peru (Balarin et al., 2019). Obviously, the details of low-fee private schooling differ according to the local context, but what is clear from all of these studies is the recent expansion of low-fee private schools and their growing importance to education, even in war-torn countries such as Liberia.

Factors leading to low-fee private schools

In 2008 Phillipson identified some of the most important factors leading to the expansion of low-fee private schools as: an over-supply of teachers, poor performance of the public sector, high hidden costs of government schooling and private tuition costs, and language of instruction. These factors still appear to be central. While each may not apply in all cases, most can be seen to be present in developing countries where an expansion of low-fee private schools is occurring.

Over-supply of teachers

Initially, it might seem strange that various developing countries have an over-supply of teachers. Yet, in countries such as India, teaching is seen as a well-paid and secure job which is also one that is “respectable” for women. Indeed, Kingdon (2020:16–18) shows in India government elementary schools teachers earn, on average, around 8 times the country’s per capita income. This ratio is very much higher than for teachers “in China, Pakistan, Indonesia, Bangladesh etc. where the ratio is typically between 1 and 2” (pp. 16–17). Over-supply is also a matter of definition, for in many countries private school teachers do not have to be qualified by the state. In many parts of India, the over-supply is not in terms of trained teachers, but of graduates who are deemed to be suitable to teach simply because they have a degree. In low-fee schools many teachers do not even have a degree or sometimes even any schooling after primary. In other parts, and in some other countries, there are many more trained teachers than places available due to poor planning and the desire of many to teach.

The oversupply of people qualified, or believing themselves qualified, to teach leads directly to new entrepreneurial schools being opened. It also means that new and existing low-fee private schools are able to employ teachers at very low cost, often at a fraction of the salaries that trained teachers can obtain in the government schools. Data on private school salaries is difficult to research for negotiations are often conducted at an individual level, and vary considerably by District and State. Kingdon (2020: 18) shows that the ratio of average private school compared to government school salaries seems to have declined over the last 20 years, partly due to the overnight almost doubling of state salaries in 2009, such that the worst ratio was found in the early 2010s in two districts of the Punjab where private school teachers were earning only 3% of that of Government teachers. Perhaps one-tenth is a reasonable guess overall – a figure which is still roughly equivalent to per capita income. Those teachers who are trained are prepared to tolerate these low salaries for a short period while adding to their experience of teaching and waiting for a job in a government school. Muralidharan and Kremer (2009), for example, estimate that (before the hike in government salaries) teachers in rural low-fee private schools in India earned at most one-fifth of the salaries of government school teachers. They were on average 10 years younger and twice as likely to come from the same village where the school is situated. They explained that, not only does this poorly paid teaching give them further teaching experience, it also gives them time for further study via distance education or local colleges so that they became better placed in the labor market for a future job in teaching or elsewhere.

Low salaries for teachers is the central reason why low-fee private schools are able to operate with low-fees. Härmä (2015) found that in Lagos the mean private school salary was only 60% of the legal minimum wage and only 42% of the starting salary for teachers in government schools. Even where extra funding is obtained from religious or philanthropic organizations, low-salaries are usually crucial in balancing the budget.

Overall poor performance of the public sector

The failure of the government sector can be simply that schools are not available for all children, or not within what the parents regard as a reasonable distance from their homes. This is the case in many rural parts of India, Nigeria, Uganda, and so on. It can also be that the schools available do not offer the type of schooling that parents desire, in particular with regard to religion or nature of the curriculum. Interestingly, this can apply as much in developed countries as in developing ones where schools for religious minorities such as Muslims or evangelical Christians are often not available and these groups feel forced to provide their own (often low-fee) private schools (Walford, 1991).

But the most significant way in which government schools are thought to have failed is in terms of academic success and factors linked to this. The reduced salaries paid to teachers in low-fee private schools allow them to employ proportionally more teachers than in government schools. In Muralidharan and Kremer’s (2009) study, for example, they found that the pupil-teacher ratio for the private schools of 19.2 was less than half the ratio of 43.4 in government schools. But, perhaps most importantly, teachers in private schools are often perceived as having a greater commitment to the school and to the pupils than teachers in government schools. This is seen in terms of the number of absences by teachers (which many studies have shown is often far lower in low-fee private schools, e.g. Härmä, 2015: 184), the amount of time spent teaching by teachers when they are present, and by the fact that teachers in low-fee are more likely to be local to the school and thus “understand” the children better, speak the local language, and be more accountable to parents. Private schools, of course, have a greater ability

to sack teachers who do not meet expectations than do far more bureaucratic, and unionized, government schools. In contrast, as we have seen, teachers in low-fee schools are less likely to be trained teachers and to have a lower level of education than those in government schools.

It has long been known that there are problems with many schooling systems in the developing world. For example, an important *Public Report on Basic Education* (The PROBE Team, 1999) in four northern Indian States, showed many problems associated with the quality of schooling. When researchers called unannounced on a large random sample of government schools only half were engaged in any teaching activity. In a third, the Principal was absent. Examples were given of teachers being drunk, sleeping on the job, getting children to do their domestic chores for them, and teachers keeping schools closed for weeks at a time. The report concluded that, generally, teaching activity in these government schools had been reduced to a minimum, in terms of both time and effort. More importantly, they claimed that “this pattern is not confined to a minority of irresponsible teachers – it has become a way of life in the profession.” Similar conclusions have been drawn by many research studies.

There is something of an irony in that, while there are such clear problems in the government sector, for many educationists it is the quality of the schooling provided in the private sector that is a key concern. But these are reasonable concerns given that the fees of such schools are low mainly because the teachers are not paid salaries anywhere near those of government school teachers, and teachers in these low-fee schools are often not trained or qualified teachers. It is important that parents are not being exploited and part of their very limited incomes being wasted. Yet, Tooley et al. (2007) found, in an extensive census and survey of private and government schools in the notified slums of Hyderabad, that on a variety of measures (including pupil-teacher ratio, teaching activity, teacher absenteeism, and classroom facilities) private unaided schools (including the unrecognized ones) were actually superior to the comparable government schools. In a later studies of academic performance in India, Nigeria and Ghana, Tooley and his colleagues (Tooley, 2009, ch 9) gave standardized tests to thousands of children and showed that the children in private schools in general scored higher on these tests in key curriculum areas than children in government schools. This was true even when the results were controlled for several background variables to try to account for the differences between the children’s backgrounds. Such studies are always open to criticism for not controlling for sufficient variables.

In the mid 2010s there were two major reviews of published studies on low-fee schools conducted by Day Ashley et al. (2014) and Wales et al. (2015). These reports drew together a great deal of data and are a useful guide to relevant publications, but are less useful for their overall findings. The basic problem is that any review, by its very nature, cannot take into consideration the social and historical contexts in which schools and education are embedded. Countries vary greatly in the regulatory and financial context in which both state and non-state schools function. To try to make generalizations about the nature and impact of non-state schools is fraught with problems – a fuller understanding can only be achieved a country at a time and, even then, there is considerable variation in the quality, nature, and effectiveness of individual low-fee private schools.

However, it is worth briefly examining some of the findings of these reports on academic advantage. Where there are nearby government schools, one of the most important questions that researchers have attempted to answer is whether low-fee private schools actually offer any academic advantage to children over that available in government schools. Day Ashley et al. (2014) examined research published in the previous five years focusing on 11 countries that had been prioritized by the UK Department for International Development – Bangladesh, Ghana, India, Jamaica, Kenya, Malawi, Nepal, Nigeria, Pakistan, South Africa, and Tanzania. This review considered private schools that were dependent on fees for all or part of their of their operational and development costs, and were owned and/or founded independently of the state. It did not include schools run by NGOs, charities, or religious organizations. Fifty-nine studies were judged to be of appropriate quality and included in the review. As such review studies are unable to take account of each country’s differing histories and contexts, while not including elite private schools, some of the research considered may include moderate-rather than low-fee schools. Overall, there was only one finding where they were able to present strong evidence – that teaching is better in private schools than in state schools as measured by levels of teacher presence and teaching activity as well as teaching approaches that are more likely to lead to improved learning outcomes. Simply stated, the teachers are more likely to turn up and teach their classes. The report found moderate evidence that parents perceived that private schools were of better quality than government schools in terms of teaching, teacher attendance, school performance, small class size and discipline. The study found moderate evidence that private school students actually achieved better learning outcomes, but found ambiguity about the size of the school effect and they noted that “many children may not be achieving basic competences even in private schools” (Day Ashley et al., 2014: 1). The report was heavily criticized by Tooley and Longfield (2014) and an active blog ran on the UKFIET website for a while (2015), with some of the criticism being accepted by the original authors of the report.

A more recent review by Akmal et al. (2019) examined 33 empirical papers published between 2014 and 2019 specifically looking at the relative academic success of children in government and low-fee private schools. Of 13 papers judged of appropriate quality, they found eight showing a positive effect on learning outcomes for private schools (including Rolleston and Moore (2018) on Andhra Pradesh and Alcott and Rose (2016) on Kenya and Uganda) and five studies showing no or very little difference between government and low-fee private schools (including Dixon et al. (2019) on Delhi, Zuilkowski et al. (2020) on Kenya). Akmal et al. (2019) found that the evidence had not changed since the original DFID Report, and gave only weak support for private schools advantage. A study by Gruijters, Adcott, and Rose (2020) of schools in India, Kenya, Pakistan and Uganda found that once they had controlled for family background, there were (at most) small private school learning advantages in all cases – ranging from 0.06 to 0.25 standard deviations. This is non-negligible, but it is from a very low baseline. One thing that is clear is that if low-fee private schools do give better outcomes than government schools, in practically all cases in the developing world the differences are small and the average level of achievement attained is low.

The second major review linked to that of (Day Ashley et al., 2014) but covering philanthropic and religiously-based schools (Wales et al., 2015) has had less impact, despite of the fact that this type of school plays a large part in provision in certain countries. BRAC, for example, has 32,000 schools mainly in Bangladesh, while madrasa schools play a large part in Pakistan and Bangladesh. In the Democratic Republic of Congo Catholic schools dominate this form of provision. A major problem is that there is limited high quality research on whether children are more academically successful in these schools than in government schools, but the review found moderate evidence that some philanthropic schools gave an academic advantage while limited evidence suggest that religious schools were similar to government schools.

In summary, while parents clearly believe that low-fee private schools are likely to produce academic results that are better than the corresponding local government schools, there are mixed research results on this. In some areas of some countries there is fairly strong overall positive support, but the differences are small, and the overall level of performance is low. In other areas of the same or different countries there is little or no evidence for their success. In the end, as in the wider private sector it depends on the quality of each individual private school and the quality of the competing government school. Härmä (2019) has shown in her case studies of Abuja, Accra and Kampala that regulation is weak for both types of school. In such a situation variation in the quality of individual schools will be high, such that broad generalizations and averages are only partially informative.

Hidden costs and private tuition costs

Even in most developing countries, at least primary schooling is nominally free. In practice, there are still costs to be met for school uniforms, books, stationary, “voluntary” gifts to teachers, and to the school for specific activities. But if government schools are not performing effectively, many parents also pay for supplementary private tutoring – sometimes given by the very same teachers who should have been teaching their children during the school day. The decision that parents make is not one of free schooling for their children versus paying fees for private schools, but one of how much more a low-fee school would cost.

Private supplementary tutoring has boomed during recent decades (Aslam and Atherton, 2014; Brehm and Silova, 2014; Bray et al., 2018). Mark Bray (2006, 2009) has documented some of this growth in a wide range of countries from Canada and France, to Cambodia, Bangladesh, Kenya and Egypt. Gathering together results from several surveys he shows, for example, that about two-thirds of Kenyan Grade 6 pupils, about a third of students from Bangladesh and Namibia receive supplementary tutoring. Many of these families using private tuition will be more affluent than those using the low-fee private schools but, if the child is attending a “free” government school, it might be cheaper to move the child to a private school that actually teaches well instead. Low-fee private schools can be seen to be partly in competition with what parents can obtain by employing private tutors. Indeed, as in some countries children in low-fee private schools are sometimes enrolled in government schools in addition to the low-fee school (to obtain free lunches or books, or to ensure that they can enter national examinations), the low-fee schools might be seen as a developed version of supplementary tuition.

Language of instruction

The last of Phillipson’s (2008) four factors that are liable to lead to a growth in low-fee private schools is the common demand among the world’s poor for a particular language of instruction. Sometimes it is minorities who wish their children to be taught in the language of the home rather than an official language but, in contrast, it may be that parents believe their child will have an advantage in life English is learned from a young age. English is seen as an international language that will provide access to future jobs. In Nepal, for example, many low-fee private schools advertise themselves as “English Boarding” or “EB” schools. In this case the word “boarding” has been transformed into a signifier of prestige, linguistically associating the schools with the major private boarding schools of Britain. The word “English” reinforces this high status and also indicates that the language of instruction is supposedly English for the main subjects. Whether or not this is correct is a matter of individual investigation. It is also highly debatable whether using English as the language of instruction at an early age actually helps children learn either English or the other subjects.

Access and affordability

An important area for research and debate has been whether low-fee schools are accessible and affordable for very poor families. There are many local studies particularly in rural areas which find that low-fee schools are both inaccessible and unaffordable (e.g. Härmä and Rose, 2012; Härmä, 2016a,b; Riep, 2014). The review by Day Ashley et al. (2014) found weak and inconclusive evidence about the accessibility of private schools for poor children. There were differences between urban and rural areas in the provision of private schools, and the spread varied in different countries with generally there being greater access in urban areas. There was inconsistent evidence about whether more boys than girls were attending private schools, with this being more common in India and Pakistan. In contrast, Wales et al. (2015) found strong evidence that philanthropic and NGO schools often purposefully located themselves to reach the poor and marginalized in rural areas and urban slums. They also often targeted girls.

The DFID Report (Day Ashley et al., 2014) also found moderate evidence on whether the poor are able to pay for private schools. Most was neutral, some negative and the report found no positive evidence that any but a few children from lower economic quintiles are able to pay the necessary fees. This was one of the areas strongly challenged by Tooley and Longfield (2014). The report by

Wales et al. (2015) identified questions of affordability as one of the significant gaps in research, yet many philanthropic and religious schools are targeted at the poor.

One of the problems with surveys of research relating to a wide range of different countries and circumstances is that they are likely to give weak evidence overall. One has to examine access and affordability of low-fee schools at a local level and within the historic, economic and social context of these low-fee schools within the spectrum of private schools and government schools available. It is obvious, for example, that urban areas are likely to have greater choice between private schools and that rural areas may only have one or no private schools that is possible for them to use. There is a spread of private schools in terms of geographical availability as well fee level.

Kingdon (2020) has used a variety of government data to examine the extent of private schooling in India, focusing on recognized and unrecognized unaided schools, while knowing that this may under-represent the number of unrecognized schools. First she charts the rapid increase in private schools and the corresponding decrease in government schools, which she describes as a “parental abandonment of government schools.” Next she uses data from a national household survey to find how much families pay in school fees. She presents both mean and median fees for States for both urban and rural families. She benchmarks these fees against various criteria and finds that, for example, median fees compared to annual minimum earnings of a daily-wage worker can be in one state as low as 4% and are about 10% overall. She also finds that, on average, about 26% of rural private pupils’ monthly fees are below their State’s daily minimum wage. She argues that in many States more than one-third of private schools are serving poor families. In other words, using these definitions, a good proportion of private schools are accessible to the poor.

This finding from Kingdon is in contrast to much of the literature, but an important paper by Tooley and Longfield (2016) examines the seeming conundrum that there is good evidenced that some of the very poorest in various nations are using low-fee schools yet various studies indicate that they are unaffordable to the poor. The problem is the very simple one that most researchers use average fees in their calculations and fail to consider that this means that very many schools charge less than the average. They are actually asking a question of “Can the poorest families afford the average fees that are paid by families who send their children to private schools?” But the appropriate question is whether there are schools available that the poorest can afford. Just as with the rest of the private sector, there is a range of prices even for those schools specifically aimed at the poor. Further, the additional costs such as for food, stationary, uniform, and transport are often treated as if they are fixed, yet the poorest will spend less than the average poor in each of these areas.

Even the stated fee in low-fee schools is more flexible than it may seem. Schools may advertise that there are reductions for second and subsequent children, but may also be prepared to offer lower fees in particular circumstances. In a similar way to major private schools around the world, some offer scholarships to able children because these children are likely to be successful in external examinations and become good advertisements for the school. Other schools may be genuinely philanthropic and allow lower fees to orphans or where families are having unexpected financial difficulties. While some schools may have owners who can choose to be personally philanthropic, other schools may be forced into philanthropy by the need to maintain student numbers. The incomes of many of the parents of children in low-fee schools are often erratic and unpredictable, which leads to similar problems for the schools (Rolleston and Adefeso-Olateju, 2014; Edwards et al., 2017). Srivastava (2008b: 454) reports “fee-bargaining” and “fee-jumping” as parents either bargain to reduce their payments or simply change schools leaving a debt. A child paying a smaller fee for tuition is better than a vacant place, so in some circumstances parents have negotiating power. We do not know the extent of such flexibility in fees actually collected but, even without any flexibility, in any discussion of affordability it needs to be remembered that a median has half below and half above that average value, and the very poorest will try to find a school in the bottom half.

This question of affordability is key to for those who believe that low-fee private schools have an important role in providing schooling for the poor (e.g. Dixon, 2013; Tooley, 2009), but the fact that there may always be some very poor families that cannot afford even the lowest fees does not mean that low-fee private schools cannot play a part in schooling the poor. Those low-fee private schools that are religious or philanthropic are usually designed to reach the poor.

Low-cost private school chains and international investment

While many of the early low-fee schools were provided by individual entrepreneurs who either wished simply to be employed or to make a profit, low-fee private schools soon became the target of investment. Between the first and second editions of James Tooley’s book on *The Global Education Industry* (1999, 2001: 13–14) a new introduction was written in which rather scathing comments on schools “charging very low fees, and offering education of variable quality” and “basic education with no frills” in the first edition had been regretted and new information added on the ubiquity of such schools in India with positive comments made about the schools and the “Federation of Private Schools’ Management” in Hyderabad. It is noteworthy that both books were so early, and that Tooley was beginning to think about brands of schools with might improve quality and management. By 2007 Tooley was suggesting that investors might help local entrepreneurs through loans to expand their schools into branded chains which would help parents make choices about the quality of schools. Alternatively, investors might establish joint ventures with local entrepreneurs to form their own school chains. Tooley (2007: 41) envisaged initial research and development involving technology, curriculum, pedagogy and teacher-training to generate a replicable model for education for the poor within a franchise model.

Since that time, and in part due to Tooley’s own involvement, in addition to being a way of enhancing Education for All and Global Development Goals, low-cost private schools (in his terminology) have become investment opportunities for some very

big players in the “education industry.” Ball (2012, ch. 3, 4) has traced Tooley’s involvement in various policy networks and his role as a policy entrepreneur at an international level urging investment in for-profit low-cost schools. In April 2007 he joined Orient Global, a Singaporean private investment company, as President of its newly created US\$100 million Education Fund. That fund has since given grants to several groups with which he previously conducted research: the Kenya Independent Schools Association, working in the slums of Nairobi, The Association of Formidable Schools in Nigeria, and Joy’s Schools in Zimbabwe. Organizations and chains of low-cost schools have many advantages over individual and isolated entrepreneurial schools. They enable the sharing of good management techniques and the development of curriculum ideas. They act as a brand name or assumed badge of quality. Franchised chains of schools go further and allow a centralized curriculum to be developed such that undereducated teachers can be trained to present standard lessons of a consistent quality. In Tooley’s (and other advocates’) view, low-cost schools provided through profit-making chains of schools can extend the reach and quality of basic education for the world’s poor.

Omega School Franchise is one such chain which was co-founded in 2008 by James Tooley and Ken Donkoh, a Ghanaian entrepreneur who had previously worked for Oxfam and USAID. The model is that of a “school in a box” which usually involves the construction of a 12-classroom building and all the initial materials and resources needed for 500 students. Each school is said to be immediately self-sustainable as parents pay on a “pay-as-you-learn” daily fee system. The Omega Schools Franchise was studied by Riep (2014) who, in 2013, was able to visit six Omega Schools and conduct interviews, observations and a survey of students. At that time the chain had grown to 20 schools with about 11,000 students and had more than 38 schools after four years. The initial growth was helped by an investment in the chain by Pearson Affordable Learning Fund established in 2012 by Pearson Education – the world’s largest multinational education corporation – with Michael Barber as its first Chairman. Growth seems to have slowed after this period and the chain was sold to the Rising Academy Network in 2020.

Riep’s (2014) interviews with Donkoh indicated that Pearson saw their involvement in low-cost schools as an extension of their global focus on education. There was huge wealth at the “bottom of the pyramid” (Pralhad, 2005) from which profit eventually could be made. Riep (2014: 266–271) describes the franchise in terms of the McDonaldization of education, where the emphasis is on efficiency, standardization of product, brand reliability, and consumerist pay-as-you-learn. He does not mean this to be a compliment, but as if the schools could actually offer these, it would be better than many other schools in most developing countries. The problem is price. While the price is low because the schools employ mainly untrained people to teach and then train them to follow set lesson plans and curriculum, it is not low enough for the very poor. Riep (2014: 272) shows that the lowest 7% of Ghana’s Central region would have to pay, on average, 43% of their family income to pay for one child, and even someone earning the average income for Ghana would have to pay 40% of income. While these are low-fee schools in comparison with other local private schools, they are not low for many families. Further, he found that in a survey of 437 students in four schools only one had not been in another school before, including other private schools, thus the schools are not providing new places.

It is worth noting that the idea of low-fee private schools employing non-trained people as teachers and then teaching them to follow a pre-set curriculum is far from new. Accelerated Christian Education (a US profit-making corporation) designed such a curriculum and teaching program in the 1970s, and there were some 5000 US ACE schools in the 1980s (Rose, 1988; Walford, 1991; Scaramanga and Reiss, 2017), with a further 800 elsewhere including the UK. In 2020 there were around 7000 ACE schools worldwide including schools in USA, England, Rumania, and Ethiopia. In 2020 fees at the Vine Christian School, Reading, UK were under £2000 per year for senior students. This is low compared with the vast majority of private schools in the UK, including the roughly £3000 charged at the Independent Grammar School, Durham, UK, which advertises as “Independent, affordable education” and of which James Tooley is Chair of the School Board and a primary investor (but which does not use a pre-packaged curriculum). This school is intended to be the first of a chain of “affordable” schools in the UK. However, with median earnings for full-time employees in the UK being just over £30,000 per year in 2019, the question of affordability for which groups is once again raised. Even Tooley (2018: 15) himself admits that families in the lowest quartile will not be able to afford the school, and to pay for two children will mean the family has to be in the middle quartile for discretionary income. Forgetting his warning about using averages (discussed above), he claims that this compares well with the average private school where only the top quartile can afford to send even one child. His use of the term “low-cost” has become very flexible.

The last decade has seen a remarkable growth in chains of private schools supported by business as an investment, by philanthropy, and through various Public-Private Partnerships. From being an educational sideline, low-fee private schools and private schools in general have moved center-stage in ideas about the provision of school places in the less developed world. The World Bank, various governments’ aid agencies, policy agencies, investment organizations, and others now see supporting private schools as a positive activity (Verger et al., 2018). Such chains as Bridge International Academies were first funded by Deutsche Bank America’s Foundation, Gray Ghost Ventures, and the Kellogg Foundation through the Clinton Global Initiative, then had further investment from the Omidyar Network (Santori et al., 2006). Low-fee private schools have become seen as a strategy for achieving Sustainable Development Goals in education by governments, the aid industry, and big business as they are seen as offering more efficient access to education with less government spending. What used to be the UK’s Department for International Development DFID was one of the most active in promoting the expansion of low-fee private schools in countries such as Kenya, Nigeria and Pakistan and have been widely criticized for such support (Junemann et al., 2016; Verger et al., 2018). One interesting aspect of this is the blurring of the boundaries between low-fee and other private schools. The International Finance Corporation, the International and Private Education Forum, the Qatar Foundation’s World Innovation Summit for Education, and others organizes regular international conferences on private education where investors, consultants, aid agencies, and education entrepreneurs are brought together, and low-fee private schools are simply seen as part of the continuum of schools in which investment might be made.

Srivastava (2016) sees this as the Second-Wave of low-fee private schools and as a part of a broader trend of privatization in education, and is cautious about the extent of activity reported and whether the chains are actually affordable by the poor. She also questions the involvement of governments with PPP schemes which have caused considerable costs even in developed countries. Indeed, low-fee private schools now seem to be being treated as an extension to the existing spectrum of private schools rather than something distinctly different. For example, in a very unusual and government legislated PPP, the 1999 Right to Education Act for India demanded that all private schools – from the most elite private schools to the poorest of the low-fee – offer 25% of their places free to children of specific poor groups with only minimum funding from the state. Theoretically the Act also introduced tougher criteria for the recognition of all private schools so that minimum standards were achieved, but what is unusual is the incorporation of the entire private sector within this policy.

Conclusion

We have seen that there is no clear definition of low-fee private schools, and no clear dividing line between these so-called low-fee schools and other private schools in each country. While there are cases in academic discussion where the research questions make it helpful to draw an arbitrary line between these schools and other slightly more expensive private schools, families are unlikely to be doing so. They will be looking for a school that they can afford and seems to meet their needs. These schools are best understood and analyzed as a part of the spectrum of private schools within each country providing for a range of incomes. As with other private schools, generalizations across countries and even within countries are often misleading. Again parents will not be concerned with averages, but with comparisons between their local government and private schools bearing in mind how much they are willing and able to pay and what each school offers.

The belief that there was huge wealth at the “bottom of the pyramid” (Prahalad, 2005) from which profit eventually could be made has driven some of the boom in private schools. However, it is often forgotten that Prahalad did not include education in his study and, even though the sub-heading of his book is “Eradicating poverty through profits” the two of twelve extended examples that are nearest to education are on eye care and prosthetic feet for the poor. Both of these examples are actually non-profit organizations, and several of the other examples rely on some elements of philanthropic activity. In fact most of the examples could equally be non-profit and do not depend on the profit motive, just cheaper and more efficient ways of doing things, which could be run by non-profit just as well as for-profit organizations.

There is a variety of reasons why low-fee private schools have been established in developing countries, but the main reason is failure of the government to provide enough schools or schools of an acceptable quality or nature. Emphasizing the profit-making of entrepreneurs as a solution to the problem of inadequate schools leads to particular proposals, such as aiming to support and extend such provision. Brands of low-fee schools may provide schools of a more consistent quality, but we have seen that such chains of schools do not necessarily serve the poorest, and within the private sector as a whole there has been a tendency to try to move up the income and fee scales, thus gradually excluding the poorest families. This is less true with religious and philanthropic schools which have a clear mission.

The fundamental problem of private schooling applies to low-fee private schools as much as to any other private schools (Walford, 1990). Those parents who are able and willing to afford a private school, because they believe it will give advantages to their children, are taken out of the government educational system. These are the very parents who are most likely to be able and willing to challenge politicians to improve the government system. Politically this can be disastrous for the needs of the poor and voiceless as governments respond to pressure.

Treating the phenomenon of low-fee private schools as separate from the rest of the private sector, and from the government sector, limits our understanding of how these schools have been established and how they fit within their local educational systems. There is now a globalized desire among poor parents to have their children attend school (Walford, 2015). Families have put their faith and dreams in schools as a way out of poverty for their children. There is a great deal of evidence that parents would prefer the government to provide such schools free at the point of delivery, thus sharing the costs of educating all young people between the whole population – few would pay fees if good schools were provided.

Acknowledgments

This paper draws heavily on previous papers (Walford, 2011, 2015). The work on the paper was partially supported through my work for the Open Society Institute, but the statements made and conclusions drawn are those of the author alone.

References

- Akmal, M., Crawford, L., Hares, S., 2019. Low-cost Private Schools: What Have We Learned in the Five Years since the DFID Rigorous Review? Center for Global Education, Washington, DC. www.cgdev.org/blog/low-cost-private-schools-what-have-we-learned-five-years-dfid-rigorous-review. (Accessed 22 September 2020).
- Alcott, B., Rose, P., 2016. Does private schooling narrow wealth inequalities in learning outcomes? Evidence from East Africa. *Oxf. Rev. Educ.* 42 (5), 495–510.
- Akyeampong, K., 2009. Public-private partnerships in the provision of basic education in Ghana: challenges and choices. *Compare* 39 (2), 135–150.

- AlWindi, J., 2015. Low-cost private schools in the slums of Colombia. In: Dixon, P., Humble, S., Counihan, C. (Eds.), *Handbook of International Development and Education*. Edward Elgar, Cheltenham, pp. 200–214.
- Aslam, M., Atherton, P., 2014. The shadow education sector in India and Pakistan: opening Pandora's Box. In: MacPherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 137–158.
- Aubry, S., Dorsi, D., 2016. Towards a human rights framework to advance the debate on the role of private actors in education. *Oxf. Rev. Educ.* 42 (5), 612–628.
- Balarin, M., Fontdevila, C., Marius, P., Rodriguez, M.F., 2019. Educating on a budget: the subsistence model of low-fee private schooling in Peru. *Educ. Pol. Anal. Arch.* 27 (132), 1–39. <https://doi.org/10.14507/epaa.27.4328>.
- Ball, S.J., 2012. *Global Education Inc. New Policy Networks and the Neo-Liberal Imaginary*. Routledge, London.
- Bhatta, P., 2014. Public desire for private schooling. In: Macpherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 67–87.
- Bray, M., 2006. Private supplementary tutoring: comparative perspectives on patterns and implications. *Compare* 36 (4), 515–530.
- Bray, M., 2009. *Confronting the Shadow Education System. What Government Policies for what Private Tutoring?* UNESCO, Paris.
- Bray, M., Kwo, O., Jokić, B., 2018. *Researching Private Supplementary Tutoring: Methodological Lessons from Diverse Cultures (CERC Studies in Comparative Education)*. Springer, Hong Kong.
- Brehm, W.C., Silova, I., 2014. Ethical dilemmas in the education marketplace: shadow education, political philosophy and social (in)justice in Cambodia. In: MacPherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 159–178.
- Chimombo, J., 2009. Expanding post-primary education in Malawi: are private schools the answer? *Compare* 39 (2), 167–184.
- Chudgar, A., Creed, B., 2016. How are private school enrolment patterns changing across Indian districts with a growth in private school availability? *Oxf. Rev. Educ.* 42 (5), 543–560.
- Day Ashley, L., McLoughlin, C., Aslam, M., Engel, J., Wales, J., Rawal, S., Batley, R., Kingdon, G., Nicolai, S., Rose, P., 2014. *The Role and Impact of Private Schools in Developing Countries: A Rigorous Review of the Evidence*. Department of International Development, London.
- Dixon, P., 2013. *International Aid and Private Schools for the Poor. Smiles, Miracles and Markets*. Edward Elgar, Cheltenham.
- Dixon, P., Egalite, A.J., Humble, S., Wolf, P.J., 2019. Experimental results from a four-year targeted education voucher program in the slums of Delhi, India. *World Dev.* 124. <https://doi.org/10.1016/j.worlddev.2019.104644>.
- Dixon, P., Humble, S., 2019. **School choice in Liberia**. In: Dixon, P., Humble, S. (Eds.), *School Choice Around the world...and the Lessons We Can Learn*. Institute of Economic Affairs, London, pp. 117–130.
- Edwards Jr., B., Klees, S.J., Wildish, J.L., 2017. Dynamics of low-fee private schools in Kenya: governmental legitimacy, school-community dependence, and resource uncertainty. *Teach. Coll. Rec.* 119 (7), 1–42.
- Fennell, S., 2013. Low-fee private schools in Pakistan: a blessing or a bane? In: Srivastava, P. (Ed.), *Low-fee Private Schooling: Aggravating Equity or Mitigating Disadvantage?* Symposium, Oxford, pp. 65–82.
- Green, F., Kynaston, D., 2019. *Engines of Privilege. Britain's Private School Problem*. Bloomsbury, London.
- Gruiters, R.J., Alcott, B., Rose, P., 2020. *The Effect of Private Schooling on Learning Outcomes in South Asia and East Africa: A Within-Family Approach*. REAL, University of Cambridge, Cambridge, UK.
- Härmä, J., 2009. Can choice promote Education for All? Evidence from growth in private primary schooling in India. *Compare* 39 (2), 151–165.
- Härmä, J., 2015. Private schooling and development: an overview. In: Dixon, P., Humble, S., Counihan, C. (Eds.), *Handbook of International Development and Education*. Edward Elgar, Cheltenham, pp. 171–199.
- Härmä, J., 2011. Low cost private schooling in India: is it pro poor and equitable? *Int. J. Educ. Dev.* 31, 350–356.
- Härmä, J., 2016a. Is there a private school market in poor neighbourhoods in Maputo, Mozambique? Exploring the role of the non-state education sector. *Oxf. Rev. Educ.* 42 (5), 511–527.
- Härmä, J., 2016b. School choice in rural Nigeria? The limits of low-fee private schooling in Kwara State. *Comp. Educ.* 52 (2), 246–266.
- Härmä, J., 2019. Ensuring quality education? Low-fee private schools and government regulation in three sub-Saharan African capitals. *Int. J. Educ. Dev.* 66, 139–146.
- Harma, J., Rose, P., 2012. Is low-fee private primary schooling affordable for the poor? Evidence from rural India. In: Robertson, S.L., Mundy, K., Verger, A., Menashy, F. (Eds.), *Public Private Partnerships in Education*. Edward Elgar, Cheltenham, pp. 243–258.
- Junemann, C., Ball, S.J., Santori, D., 2016. Joined-up policy: Network connectivity and global education governance. In: Mundy, K., Green, A., Lingard, R., Verger, A. (Eds.), *Handbook of Global Education Policy*. Wiley-Blackwell, Chichester, pp. 535–553.
- Kingdon, G., 2007. The progress of school education in India. *Oxf. Rev. Econ. Pol.* 23, 168–195.
- Kingdon, G., 2008. School-sector effects on student achievement in India. In: Chakrabarti, R., Peterson, P.E. (Eds.), *School Choice International*. MIT Press, Cambridge, MA, pp. 111–139.
- Kingdon, G.G., 2020. The private school phenomenon in India: a review. *J. Dev. Stud.* 56 (10), 1795–1817.
- Kisira, S., 2008. Uganda. In: Phillipson, B. (Ed.), *Low-cost Private Education. Impacts on Achieving Universal Primary Education*. Commonwealth Secretariat, London, pp. 131–171.
- Kitaev, I., 1999. *Private Education in Sub-Saharan Africa: A Re-examination of the Theories and Concepts Related to its Development and Finance*. UNESCO-IIEP, Paris.
- Lewin, K.M., 2007. The limits to growth of non-government private schooling in Sub-Saharan Africa. In: Srivastava, P., Walford, G. (Eds.), *Private Schooling in Less Economically Developed Countries: Asian and African Perspectives*. Symposium Books, Oxford, pp. 41–65.
- Lewin, K.M., 2017. Making rights realities. Does privatizing educational services for the poor make sense? In: Ndlimande, B.S., Lubienski, C. (Eds.), *Privatization and the Education of Marginalized Children. Policies, Impacts, and Global Lessons*. Routledge, New York, pp. 81–102.
- Majumdar, M., Vaidyanathan, A., 1995. *The Role of Private Education in India: Current Trends and New Priorities*. Centre for Development Studies, Trivandrum, Kerala.
- Moschetti, M., Snaider, C., 2019. Speaking cooperation, acting competition: supply-side subsidies and private schools in socioeconomically disadvantaged contexts in Buenos Aires. *Educ. Pol. Anal. Arch.* 27 (131), 1–26. <https://doi.org/10.14507/epaa.27.4330>.
- Muralidharab, K., Kremer, M., 2009. Public-private schools in rural India. In: Chakrabarti, R., Peterson, P.E. (Eds.), *School Choice International*. MIT Press, Cambridge, MA.
- Orme, N., 2008. For richer, for poorer? Free education in England. *J. Hist. Child. Youth* 1 (2), 169–187.
- Peel, M., 2015. *The New Meritocracy: A History of UK Independent Schools 1979–2014*. Elliott and Thompson, London.
- Phillipson, B., 2008. In: *Low-cost Private Education. Impacts on Achieving Universal Primary Education*. Commonwealth Secretariat, London.
- Prahalad, C.K., 2005. *The Fortune at the Bottom of the Pyramid: Eradicating Poverty through Profit*. Wharton School Publishing, Upper Saddle River.
- Riep, C.B., 2014. Omega Schools Franchise in Ghana: affordable private education for the poor or for-profitteering? In: MacPherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 259–277.
- Rolleston, C., Adefeso-Olateju, M., 2014. De facto privatisation of basic education in Africa: a marked response to government failure? A comparative study of the cases of Ghana and Nigeria. In: MacPherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies from Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 25–44.
- Rolleston, C., Moore, R., 2018. *Young Lives School Survey, 2016–7*. University of Oxford, Department of International Development, Oxford.
- Rose, S.D., 1988. *Keeping Them Out of the Hands of Satan. Evangelical Schooling in America*. Routledge and Kegan Paul, London.
- Rose, P., Adelabu, M., 2007. Private sector contributions to education for all in Nigeria. In: Srivastava, P., Walford, G. (Eds.), *Private Schooling in Less Economically Developed Countries: Asian and African Perspectives*. Symposium Books, Oxford, pp. 67–87.

- Santori, D., Ball, S.J., Junemann, C., 2016. Financial markets and investment in education. In: Verger, A., Lubienski, C., Steiner-Khamsi, G. (Eds.), *The Global Education Industry*. Routledge, London, pp. 193–210.
- Scaramanga, J., Reiss, M.J., 2017. Accelerated Christian education: a case study of the use of race in voucher-funded private Christian schools. *J. Curric. Stud.* 50 (3), 333–351.
- Shukla, S., Joshi, P., 2008. India. In: Phillipson, B. (Ed.), *Low-cost Private Education. Impacts on Achieving Universal Primary Education*. Commonwealth Secretariat, London.
- Srivastava, P., 2006. Private schooling and mental models about girls' schooling in India. *Compare* 36 (4), 497–514.
- Srivastava, P., 2007a. For philanthropy or profit? The management and operation of low-fee private schools. In: Srivastava, S., Walford, G. (Eds.), *Private Schooling in Less Economically Developed Countries: Asian and African Perspectives*. Symposium Books, Oxford, pp. 153–186.
- Srivastava, P., 2007b. Low-fee private schooling: challenging an era of education for all and quality provision? In: Verma, G., Bagley, C., Jha, M.M. (Eds.), *International Perspectives on Educational Diversity and Inclusion: Studies from America, Europe and India*. Routledge, New York.
- Srivastava, P., 2008a. School choice in India: disadvantaged groups and low-fee private schools. In: Forsey, M., Davies, S., Walford, G. (Eds.), *The Globalization of School Choice*. Symposium Books, Oxford, pp. 185–208.
- Srivastava, P., 2008b. The shadow educational framework: towards a new institutional understanding of an emerging private sector in India. *Res. Pap. Educ.* 23, 451–474.
- Srivastava, P., 2016. Questioning the global scaling up of low-fee private schooling. The nexus between business, philanthropy, and PPPs. In: Verger, A., Lubienski, C., Steiner-Khamsi, G. (Eds.), *The Global Education Industry*. Routledge, London, pp. 248–263.
- The PROBE Team, 1999. *Public Report on Basic Education in India*. Oxford University Press, New Delhi.
- Tooley, J., 1999. *The Global Education Industry. Lessons from Private Education in Developing Countries*. Institute for Economic Affairs, London.
- Tooley, J., 2001. *The Global Education Industry. Lessons from Private Education in Developing Countries, second ed.* Institute for Economic Affairs, London.
- Tooley, J., 2007. Educating Amarech: private schools for the poor and the new frontier for investors. *Econ. Aff.* 27 (3), 37–43.
- Tooley, J., 2009. *The Beautiful Tree, A Personal Journey into How the World's Poorest People Are Educating Themselves*. Penguin, New Delhi.
- Tooley, J., 2018. A chain of low-cost private schools for England. *Policy* 34 (2), 13–17.
- Tooley, J., Dixon, P., 2005. An inspector calls: the regulation of 'budget' private schools in Hyderabad, Andhra Pradesh, India. *Int. J. Educ. Dev.* 25, 269–285.
- Tooley, J., Dixon, P., 2006. "De facto" privatisation of education and the poor: implications of a study from sub-Saharan Africa and India. *Compare* 36 (4), 443–462.
- Tooley, J., Dixon, P., 2007. Private education for low-income families: results from a global research project. In: Srivastava, P., Walford, G. (Eds.), *Private Schooling in Less Economically Developed Countries: Asian and African Perspectives*. Symposium Books, Oxford, pp. 15–39.
- Tooley, J., Dixon, P., Gomathi, S.V., 2007. Private schools and the millennium development goal of universal primary education: a census and comparative survey in Hyderabad, India. *Oxf. Rev. Educ.* 33 (5), 539–560.
- Tooley, J., Longfield, D., 2014. The Role and Impact of Private Schools in Developing Countries: A Response to DfID's "Rigorous Literature Review". E.G.West Centre, University of Newcastle, Newcastle.
- Tooley, J., Longfield, D., 2016. Affordability of private schools: exploration of a conundrum and towards a definition of 'low-cost'. *Oxf. Rev. Educ.* 42 (4), 444–459.
- UKFIET (2015). www.ukfiет.org/2015/what-we-know-and-dont-know-about-the-impact-of-private-schooling-in-developing-countries/ (Accessed 22 September 20).
- Umar, A., 2008. Nigeria. In: Phillipson, B. (Ed.), *Low-cost Private Education. Impacts on Achieving Universal Primary Education*. Commonwealth Secretariat, London, pp. 91–129.
- Verger, A., Fontdevila, C., Zancajo, A., 2018. Constructing low-fee schools as an educational model for the Global South: from local origins to transnational dynamics. In: Verger, A., Novelli, M., Altinyelken, H. (Eds.), *Global Education Policy and International Development. New Agendas, Issues and Policies, second ed.* Bloomsbury, London, pp. 255–276.
- Vine Christian School (2020). <https://www.vinechristianschool.org.uk/> (Accessed 10 October 20).
- Wales, J., Aslam, M., Hine, S., Rawal, S., Wild, L., 2015. *The Role and Impact of Philanthropic and Religious Schools in Developing Countries*. DFID, London.
- Walford, G., 1990. *Privatization and Privilege in Education*. Routledge, London.
- Walford, G., 1991. The reluctant private sector: of small schools, people and politics. In: Walford, G. (Ed.), *Private Schooling: Tradition, Change and Diversity*. Paul Chapman, London, pp. 115–132.
- Walford, G., 2011. Low-fee private schools in England and in less economically developed countries. *Compare* 41 (3), 401–413.
- Walford, G., 2015. The globalisation of low-fee private schools? In: Zajda, J. (Ed.), *Second International Handbook on Globalisation, Education and Policy Research*. Springer, Amsterdam, pp. 309–320.
- Zuilkowski, S.S., Piper, B., Ong'ele, S., 2020. Are low-cost private schools worth the investment? Evidence on literacy and mathematics gains in Nairobi primary schools. *Teach. Coll. Rec.* 122 (1), 1030.

False and distracting promises: early childhood education and care in low income countries

Helen Penn^{a,b}, ^a UCL Institute of Education, London, United Kingdom; and ^b University of East London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	366
Early child development	366
Early education	368
Conclusion	369
References	369

Introduction

Economic inequality between high-income and low-income countries is so pronounced that one estimate suggests that for every \$1 given in aid, \$24 is extracted in terms of resources, patents, trade deals and money manipulations (Kar and Schjelderup, 2016). Within countries too, inequalities between rich and poor, dominate daily life. Inequality, driven by the movement of global capital (but not the movement of peoples), distorts life and expectations for billions of people (Deaton, 2013; Milanovic, 2016; Piketty, 2020; Stiglitz, 2011; Hickel 2020; Woodhead et al., 2009; and many others). At a time when millions are dying of Covid, very wealthy billionaires, like Jeff Bezos, Richard Branson and Elon Musk, oblivious to the needs of their fellow citizens, are spending billions on personal space travel (Nolan, 2021). Social mobility for most people is a fantasy, such are the structural inequalities that shape lives.

Yet the myth of social mobility continues to be powerful, and is embedded in the rhetoric of education. It may be true that for a very few, educational opportunities are fruitful and offer a route out of poverty, but very generally education systems tend to perpetuate inequality and confirm poor children in their poverty and in their position at the bottom of a hierarchy rather than offering them a way out. Literacy and numeracy are highly valued in modern societies, and it is undoubtedly better for survival to have some schooling rather than none, but state education systems are rarely redistributive. The elite almost always choose private, segregated education if it is available to them (Reay, 2017).

Early child development

Nowhere is this myth of social mobility more pervasive than in the education of very young children. This is partly due to the work of James Heckman, who won a Nobel prize for his work on econometrics, and his exploration of the predictive power of statistics to inform social policies. Heckman argued that early intervention programs provide higher rates of return compared to remediation programs targeted at older children and young adults. It is more economically efficient to provide early education and care to young children than for any other age group, in terms of long-term outcomes-higher wage earners, less call on social services and support. It is cheaper in the long term to provide education/stimulation for young children than to deal with results of no or poor education.

This conclusion about the relevance of early intervention has been reinforced expanded by various international bodies, who have linked it to work on neuroscience. The argument put forward is that the brain grows most rapidly in the first three years; neurones expand and form dense networks, but in order to do this, the brain needs both adequate mental stimulation and adequate nutrition. The relationship between neurone growth and “stimulation” is tenuous (Penn, 2019a,b). The evidence on the importance of nutrients for growth, including brain growth, is stronger, but the World Food Program estimates that billions of children are hungry. Food poverty is very widespread, and getting worse. Programmes which dish out nutritional supplements to very young children do so in a context where elder children and adults often go hungry, and the nutrients may have to be distributed in such a way that no other family members can access them! Whatever the practical and cultural drawbacks in running programmes for very young children, they are seen to offer a magic solution to global problems.

What is one of the best ways a country can boost shared prosperity, promote inclusive economic growth, expand equitable opportunity, and end extreme poverty? The answer is simple: Invest in early childhood development. Investing in early childhood development is good for everyone – governments, businesses, communities, parents and caregivers, and most of all, babies and young children. It is also the right thing to do, helping every child realize the right to survive and thrive. And investing in ECD is cost effective: For every \$1 spent on early childhood development interventions, the return on investment can be as high as \$13. Early childhood development is also key to upholding the right of every child to survive and thrive.

(WHO's Nurturing Care foreword 2017).

Healthy development in the early years (particularly birth to three) provides the building blocks for educational achievement, economic productivity, responsible citizenship, lifelong health, strong communities, and successful parenting of the next generation. What can we do during this incredibly important period to ensure that children have a strong foundation for future development?

(Harvard Center for the Developing Child, 2020).

International aid programmes, especially those, like the Children's Investment Foundation Fund (one of the biggest donors-who also benefit from tax breaks through charitable giving) have latched onto it as an opportunity to entrepreneurially engage in an area previously neglected in international aid portfolios, and which appears to promise good investment returns. The approach to early childhood generated by Heckman is referenced by almost every international organization – including major bi-lateral agencies like the World Bank, UNICEF, UNESCO and WHO— as well as by almost all leading child charities and think tanks. It has become an international development commonplace. Many policy documents, as evidenced above, argue that early interventions in childhood, through providing high quality early education for targeted populations, can be an effective strategy to reduce the prevalence of later adult 'problems' such as poverty, unemployment, criminal offending and intergenerational disadvantage.

WHO in particular has promoted an early intervention approach which targets parents (inevitably mothers) by using home visiting programmes. Using tenuous arguments about brain development to bolster their interpretation of the Heckman data (Penn, 2019a,b) they argue that the first 1000 days are the most important for laying down the foundations of later life. WHO claims that interventions which enable parents and children to learn more about optimal child development strategies, will improve children's brains and thereby their life-chances in an economically efficient way and save money on subsequent remedial services, as Heckman predicted. There are many claims for these home visiting programmes. One of the most cited is a Jamaican program which claims that long-term home-visiting interventions produce higher income earners over a lifespan (World Bank, 2021). The program, targeted at 170 children in a total population of 2.3 million inhabitants has in no way impacted on Jamaica's considerable wealth inequalities, nor has it altered its wealth as a nation vis a vis other countries. Yet these programmes are touted as a long-term opportunity to substantially improve children's future, and in doing so, make countries more prosperous and better able to compete in the world market. This is such a staggering and preposterous claim for so small an impact, that one wonders what is at stake.

Heckman's work draws on an underlying theory of skills formation. Key to this theory is an assumption that directly rejects the impact of structural inequality:

disadvantaged families do not invest sufficiently in their children because of information problems rather than limited economic resources or capital constraints.

(Heckman et al., 2013).

This approach focuses on individual success in world of inequalities. It focuses on what contributes to individual success and takes the unevenness of the world, and the reluctance to pay for state services other than for a needy minority, for granted. It seeks only to justify small interventions which may make some children able to cope better than they would otherwise have done, at a relatively small cost.

This research is invariably couched in hyper-rigorous statistical terms, which excludes any wider conceptual discussion outside of the statistical formulas employed. Its frame of reference is explicitly narrow and it deliberately does not take structural inequality into account. This is increasingly problematic, since Heckman's original data, on which he constructed his predictive statistical curve, was drawn from a small pool of USA studies which only use randomized controlled trials, some of which were initiated over 50 years ago and date back to the 1960's and 1970's. The initial assumptions and procedures of those studies, especially those concerning race, do not bear re-examination. Although Heckman has since been involved in longitudinal studies in a number of countries, including Jamaica and Italy, the highly contextual nature of the assumptions, and the limitations on its generalization to other countries, is rarely conceded (Rea and Burton, 2020)

One of the underlying assumptions in the WHO approach is that child development tenets, like medical knowledge, are scientifically based and universal (Phillips et al., 2017). Just as poor nutrition in early childhood leads to stunted bodies, poor stimulation in early childhood leads to stunted brains. Leaving aside the strength or otherwise of the neuroscientific and brain growth arguments, what is "poor stimulation"? Robert Serpell, who has long convened a group of African psychologists, argues that ideas about child development are predominantly and explicitly Westernized and Americanized, heavily emphasizing individualism, cognitive and linguistic precocity and personal striving at the expense of – for example in Africa - of co-operation and deferential helpfulness. Many other psychologists and anthropologists have commented on the cultural biases of typical child development texts from the global North (Burman, 2018; LeVine and New, 2008; Gottlieb, 2004). Tobin et al. (2011) gives a bleak description of the "shopping mall" mentality encouraged by American caregivers of young children, continually urging them to "choose" between a limitless range of commercial goods, and enabling them to vigorously defend and argue for their choices even when they inconvenience others. Perhaps the most trenchant comment is that of Jerome Brunauer. *"Perhaps even more than with most cultural matters, childrearing practices and beliefs reflect local conceptions of how the world is and how the child should be readied for living it."*

Many of those concerned with promoting "child development" are often singularly uninterested in how parents view the world, and what, as parents, they consider children should be learning. The statistical methods used in randomized controlled trials deliberately exclude qualitative data, and there is no account in any of the Jamaican material for example, of how mothers view the

program. Their views are an irrelevance in calculating efficacy. The home visiting programmes also assumes the availability of mothers, and ignore any paid work in which they may be engaged, an assumption which UN women (2017), for example, has severely criticized.

It is partly that -relatively speaking - early child interventions are such a cheap option for improving the world, and do not disturb the status quo, that makes it such a popular idea among international policy makers. Teaching mothers to look after their children better, is a typical Westernized welfare solution that neatly puts the burden and the blame for poverty on mothers themselves, or at the very least, refuses to take other explanations into account. The poor can be taught how to do their job of child-rearing better. From a historical perspective, it is what the blinkered rich have been saying to poor mothers for centuries, it's your fault, not ours, that you are poor, and you should improve yourself and do more for your children and feed them properly (Cunningham, 2020). As in earlier times, there is a high moral tone to the advice. A series of articles in the Lancet on the topic of *Nurturing Care- Promoting Early Child Development*, suggests the lives of millions of children could be saved. *New estimates, based on proxy measures of stunting and poverty, indicate that 250 million children (43%) younger than 5 years in low-income and middle-income countries are at risk of not reaching their developmental potential. There is therefore an urgent need to increase multisectoral coverage of quality programming that incorporates health, nutrition, security and safety, responsive caregiving, and early learning. Equitable early childhood policies and programmes are crucial for meeting Sustainable Development Goals, and for children to develop the intellectual skills, creativity, and well-being required to become healthy and productive adults* (Black et al., 2017).

Reducing human suffering is a wholly admirable goal, but home-visiting programmes for children under two, driven by a statistically sound but extraordinarily narrow econometric theory, is likely to be no more than a distraction and an excuse for wider inaction, given the geo-politics of inequality. The logical conclusion of this reasoning is that if an estimated 43% of a population has stunted bodies and stunted brains, their contribution to international debate cannot be taken too seriously, and those from high-income countries who are seeking to redress the situation are the best informed to know what to do, and likely to remain so.

Early education

Another variant of early intervention is that promoted by UNICEF and other agencies, and incorporated into the Sustainable Development Goals. Target 4.2 is by 2030 ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education. Drawing on the same Heckman data, but applying it in a different context, the assumption is that pre-school experiences will enable better performance at school, although very often the caveat "high-quality" is attached. But more often than not, pre-school, even less than mainstream school, is not of high quality, but provided by a plethora of private operators, many of them of poor quality. Pre-school programmes have been monitored for "learning outcomes" and there are various, mainly North American schedules to monitor the quality of these interventions. But often the same problems arise, the lack of competent, trained caregivers in the private sector - even in countries like the UK (Christie and Co, 2019). Ex-communist countries, which have a history of early childhood education and care tend to have more rigid provision, but the standards also tend to be higher in terms of infrastructure and training (Penn, 2019a,b).

As well as incorporating very particular, Westernized, views of child development, and about what constitutes a high quality early education, many of the lobbyists for early intervention also hold traditional views about the organization of family life. Mamdani (2017), in a famous study, used court records, to show how patriarchal colonial assumptions in Kenya have shaped the law, so that ownership of land was invariably registered with male heads of households, rather than as traditional dual or female ownership. This colonial process of diminishing the rights of women and transfer of ownership of rural land has in the long-term had disastrous consequences for women, and contributed to the migration of women from rural to urban or peri-urban areas. Razavi (2011) has shown that the phenomenon of female single-headed households in poor urban areas has become a norm, throughout the world, exacerbated by both the legacy of patriarchal colonial policies and by current fiscal policies. Women in these urban and peri-urban circumstances mostly work in the informal sector, as market traders, domestic servants, or in the sex industry. Poor women tend to work for long hours in precarious and largely undocumented jobs. As a result, there is a major childcare problem. Millions of very young children are often left unattended or in the care of barely older siblings (Sammans et al., 2016). In the cities of subSaharan Africa, unlicensed private entrepreneurs have profited from childcare enterprises, cramming very young children into unsuitable premises with minimum supervision, while their mothers are at work (Hughes et al., 2021).

This contradiction, between a prosperous, idealized family life, and everyday reality in low-income countries, is at last beginning to percolate through to international agencies. The World Bank, following on from a recent report by UN Women, has now emphasized the importance of gender issues, and suggests childcare, rather than other forms of early childhood interventions, should be a public priority for all countries. This has long been a position held by the OECD, whose documents on the issue date back to 1972. In 2006 OECD published a landmark document, *Starting Strong II* in 2006 and continues to follow this up with annual figures on its Family database, and updated versions of *Starting Strong*. Similarly, the EU has made childcare and education an integral plank in its education services, publishing a series of reports and analyses in its Eurydice program (Eurydice 2019). OECD/EU recognize that to provide comprehensive ECEC services requires substantial investment and competent governance rather than fragmented private initiatives, and also publish figures on the level of this investment and its administration across member states.

The World Bank *Better Jobs and Brighter Futures* document on childcare (Dervicelli and Beaton-Day 2020) is rather late on the scene, and still very much USA focused in its examples. It outlines some of the key regulatory areas for childcare, according to criteria enumerated by (mainly) American psychologists and educationalists, as noted above. The quality of a childcare or ECE program

depends on several different elements, which are usually categorized as either “structural” or “process”. Structural quality, typically the easiest aspect of quality to be defined, measured, and regulated, includes staff/child ratio, group size, and physical infrastructure. Process quality includes the program or curriculum; the workforce; the interactions that take place between adults and children; and the wider system which ensures these aspects of quality are in place. All five categories should be considered as part of an integrated approach, with each category supporting the others. However, the document is much vaguer as regards financing these initiatives. Educating and caring for young children is labor intensive, and requires a relatively safe and an environment which protects against hazards. The 2006 OECD document suggests 1% GDP as a target for funding an integrated childcare and early education system, and some few countries have achieved this target. But currently, in the World Bank literature, the assumption is that funding services can be achieved by a mixture of private enterprise, workplace nurseries (surely a fantasy in an informal economy), co-operatives, donations, and a very little state aid in the form of training and regulatory activities. The usual way of describing this unsatisfactory and tenuous mixture of financial incentives in which no organization carries ultimate responsibility is “a partnership”.

So although there is more recent recognition of the needs of women and children for childcare alongside education in low-income countries and elsewhere, the question of what it might cost and how it might be delivered, and to whom, is left largely unaddressed and undocumented.

Conclusion

This chapter has focused on the efforts of international agencies to improve the educational chances of young children in low-income countries. There are of course dialogs within countries and among groups within countries about education and what sustains it, which may or may not correspond to Western ideas. But many international agencies, heavily, if not totally influenced by Western ideas, argue that investment in early childhood education and care can bring about transformational change, in the lives of young children and in the prospects for the countries they live in. This article has argued that poverty and inequality are so entrenched and ignored, and the actions of international agencies, although indubitably well-meaning, are so clouded by a Westernized interpretation of childhood, that attempts to improve the situation are unlikely to succeed. It is unlikely that the education of millions of young children will be improved by current efforts of international agencies. But the illusion of success deters any substantial effort to get to grips with the problems that really exist.

References

- Black, M.M., Walker, S.P., Fernald, L.C.H., et al., 2017. Early childhood development coming of age: science through the life course. *Lancet* 389 (10064), 77–90. [https://doi.org/10.1016/S0140-6736\(16\)31389-7](https://doi.org/10.1016/S0140-6736(16)31389-7).
- Burman, E., 2018. *Fanon, Education, Action: Child as Method*. Routledge, London.
- Christie and Co, 2019. Early Childhood Education and Care: Workforce Trends and Associated Factors. Christie and Co. [https://www.christie.com/getattachment/News-Resources/Publications/September-2019-\(1\)/Early-Childhood,-Education-Care-Workforce-Trends-A/C-Co-ECEC-Report-\(Pages\)-\(1\).pdf.aspx?lang=en-GB&ext=.pdf](https://www.christie.com/getattachment/News-Resources/Publications/September-2019-(1)/Early-Childhood,-Education-Care-Workforce-Trends-A/C-Co-ECEC-Report-(Pages)-(1).pdf.aspx?lang=en-GB&ext=.pdf).
- Cunningham, H., 2020. *Children and Childhood in Western Society Since 1500 (Studies In Modern History)*, third ed. Routledge.
- Deaton, A., 2013. *The Great Escape: Health, Wealth and the Origins of Inequality*. Princeton University Press, Princeton.
- Eurydice/EU, 2019. Key Data on Early Childhood Education and Care in Europe – 2019 Edition. https://eacea.ec.europa.eu/national-policies/eurydice/content/key-data-early-childhood-education-and-care-europe-%E2%80%93-2019-edition_en.
- Gottlieb, A., 2004. *The Afterlife Is where We Come from*. Chicago University Press.
- Harvard Center on the Developing Child. (2020). <https://developingchild.harvard.edu/>.
- Heckman, J.J., Pinto, R., Savelyev, P., 2013. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *Am. Econ. Rev.* 103, 2052–2086. <https://doi.org/10.1257/aer.103.6.2052>.
- Hickel, J., 2020. *Less Is More: How Degrowth Will Save the World*. Penguin, London.
- Hughes, R.C., Kitsao-Wekulo, P., Muendo, R., Bhopal, S.S., Kimani-Murage, E., Hill, Z., Kirkwood, B.R., 2021. Who actually cares for children in slums? Why we need to think, and do, more about paid childcare in urbanizing sub-Saharan Africa. *Phil. Trans. R. Soc. B.* 376, 20200430. <https://doi.org/10.1098/rstb.2020.0430>.
- LeVine, R.A., New, R.S. (Eds.), 2008. *Anthropology and Child Development: A Cross-Cultural Reader*. Malden, MA: Blackwell Publishers. New York, NY: Wiley-Blackwell.
- Mamdani, M., 2017. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism (Princeton Studies in Culture/Power/History)*. Princeton University Press, Princeton, NJ.
- Milanovic, B., 2016. *Global Inequality*. Harvard University Press, Cambridge, MA.
- Nolan, H., 2021. Why Is Bezos Flying to Space? Because Billionaires Think Earth Is a Sinking Ship. <https://www.theguardian.com/commentisfree/2021/jul/20/jeff-bezos-space-flight-billionaires-earth-sinking-ship>.
- OECD, 2006. *Starting Strong II*. OECD, Paris.
- Penn, H., 2019a. Reflexiones acerca del desvío de la atención sobre la pobreza en el área de desarrollo infantil temprano. In: Lipina, S., Segretin, M. (Eds.), *Exploraciones neurocientíficas de la Pobreza. Mind Brain and Education Symposium*, Erice, Sicily.
- Penn, H., 2019b. Uzbekistan: implementing ECEC services in authoritarian regimes. In: Kjørholt, A., Penn, H. (Eds.), *Early Childhood and Development Work*. Palgrave Macmillan, New York, pp. 153–164.
- Piketty, T., 2020. *Capital and Ideology*. Harvard University Press.
- Phillips, D.A., Lipsey, M.W., Dodge, K.A., Haskins, R., Bassok, D., Burchinal, M.R., et al., 2017. Puzzling it Out: The Current State of Scientific Knowledge on Pre-kindergarten Effects A Consensus Statement. Brookings Institution, Washington, DC. https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf.
- Rea, D., Burton, T., 2020. New evidence on the Heckman curve. *J. Econ. Surv.* 34 (2). <https://doi.org/10.1111/joes.12353>.
- Reay, D., 2017. *Miseducation: Inequality, Education and the Working Classes*. Policy Press, Bristol.
- Seen, heard and counted: rethinking care in a development context. In: Razavi, S. (Ed.), *Dev. Change* 42 (4), 873–904.

- Sammans, E., Presler-Marshall, E., Jones, N., 2016. Women's Work— Mothers, Children and the Global Childcare Crisis. ODI (Overseas Development Institute), London. Available from: <https://www.odi.org/publications/10349-women-s-work-mothers-children-and-global-childcare-crisis>.
- Stiglitz, J., 2011. The Price of Inequality. Penguin, London.
- Tobin, J., Hsueh, Y., Karasawa, M., 2011. Preschool in Three Cultures Revisited: China, Japan, and the United States. University of Chicago Press.
- Woodhead, M., Ames, P., Vennam, U., Abebe, W., Streuli, N., 2009. Early Child Development. Equity and Quality? Working Papers in Early Child Development. Van Leer Foundation. <https://www.younglives.org.uk/sites/www.younglives.org.uk/files/BvLF-ECD-WP55-Woodhead-Equity%26Quality.pdf>.
- World Bank, 2021. 30 Years Later, Does an Early Childhood Program Still Have Impact? <https://www.worldbank.org/en/programs/sief-trust-fund/brief/30-years-later-does-an-early-childhood-program-still-have-impact>.

English as a medium of instruction and research in education: an international and comparative analysis

Birgit Brock-Utne^{a,b,c}, ^a University of Oslo, Oslo, Norway; ^b University of Dar es Salaam, Dar es Salaam, Tanzania; and ^c EDCON, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	371
The LOITASA project	372
The introduction of private primary schools	373
Strengthening of a government primary school	373
The rhetoric of English and development	373
The situation in India	374
From one colonial language to the other in Rwanda	375
The situation in the Nordic countries	375
Academic publishing—in whose language?	376
Higher rewards for academic publications in English	376
The 2008 White paper on Norwegian language policy	376
Norwegian researchers defending our language as a language of research	377
Where there is a political will—an example from Tanzania	377
Prestige Planning	377
References	378

Introduction

In the years 1987–1992 I was a professor at the University of Dar es Salaam in Tanzania. The meeting with students and colleagues there developed my interest in the language issue in Africa, which later became the main focus of my research (Brock-Utne, 2000, 2012, 2014, 2021; Brock-Utne and Skattum, 2009). When I first came to Tanzania I knew little about the country. I got hold of some books about Tanzania. I read in several of them that there were 120 languages spoken in the country. I asked people in the NORAD office in Dar es Salaam if they had learnt to speak the national language, Kiswahili. They answered that they had learnt some greetings. That was really all that was needed because the people they had meetings with and negotiated with all spoke English well. “And you will be teaching in English and all of the meetings you attend will be in English”, they said to me. “And besides” they said, “your colleagues will come from all over the country and will speak very different languages. It is only a waste of time to try to become a fluent Kiswahili speaker, but knowing some greetings may be useful.”

The first day I came to the Department of Education, there was a tea—break and all my prospective colleagues sat in chairs in a quadrangle, sipping very sweet tea. They greeted me in English but very quickly switched to speaking Kiswahili with each other. I said: “I had heard that you would be speaking different Tanzanian languages. Don’t you all come from different parts of the country?”. They answered that they did. They could speak many different local languages, but they all also could speak Kiswahili as a lingua franca, a language that united them. I decided already then that I would learn the language, so that the people sitting next to me in the tea break would not need to switch into English. When I started working full-time at the university and met my master students, I noticed that they all came into the class-room talking with each other in Kiswahili.

Teaching undergraduate classes with several hundred students was a great challenge. To make the classes a bit more interactive, I just lectured for about a quarter of an hour and then wrote some questions on the blackboard, which I asked the students to discuss in small groups for about a quarter of an hour. I had written the questions in English. Already the first time I did this, I started wandering in the auditorium, listening to the group discussions. I soon noticed that the students had quickly translated the questions. They all discussed them in Kiswahili. The discussions were lively (Brock-Utne, 2000/2008).

I shared office with my only female colleague at the time. We both had students coming to us for tutorials. I noticed that she gave several of her tutorials in Kiswahili or she code-switched from English to Kiswahili and back. I discussed this practice with her. She said that she was not supposed to give tutorials in Kiswahili, but she communicated much better with the students in that language than in English, a foreign language to them all. I told her that I was now bent on learning Kiswahili and she volunteered to help me. She taught me several proverbs and we built an ever-lasting friendship, doing research together and living together in the same room while conducting our research. As part of my duties as a professor of education at the UDSM I was also required to participate for 6 weeks each year in teacher practice supervision. This meant listening to our education students teach in secondary school, evaluate them and help them improve their teaching practice. I very quickly discovered the great difference in activity level among the

secondary school students depending on the language of instruction used. I noticed the lively atmosphere in the classes taught in Kiswahili (at that time not only the subject Kiswahili but also the subject *siasa*¹ was taught in Kiswahili) and the passivity of the students and aggressiveness of the teachers when the teaching was conducted in English.

The LOITASA project

The abbreviation LOITASA stands for Language of Instruction in Tanzania and South Africa. This was a research project which started in 2001 and lasted officially for ten years but in reality at least five more years. The last student who took her doctorate under the project did so in 2016 (Bakahwemama, 2016). In my years at the UDSM I had, through my yearly teaching practice supervision, experienced the problems students in secondary education had mastering the foreign language, English, when it was used as a language of instruction. I liked initiating discussions on the language of instruction issue with my colleagues. These discussions could become quite heated. Many colleagues thought, like I did, that it would be best for Tanzanian secondary school students to switch to Kiswahili as the language of Instruction (LOI). This was the language they normally communicated in, and were familiar with. It was the language of instruction for all of the seven years of primary school. Others claimed that this was impossible to do because English was such an important language; it was the language of science and technology. It was an international language.

One of my colleagues, Dr. Sulema Sumra, and I discussed the need for a research project on the language of instruction in Tanzania. I promised to look for funding for such a project.

There was at the time a board between universities in Norway, NUFU,² that could, on application, give money to Norwegian universities that would cooperate with and build capacity at universities in the South. NUFU preferred South-South-North cooperation. The Dean of the Faculty of Education at the University of Western Cape at the time, Prof. Harold Herman (2008:10) writes:

In 1999, Prof. Birgit Brock-Utne and I met during a Comparative and International Education Society (CIES) conference in the USA to consider the possibility of a North-South cooperation on language of instruction policies in Africa. In our ensuing years of dialogue and presentations at international conferences it was clear to me that Birgit, although Norwegian, was *an African scholar at heart*, a person with a flair and commitment to confront the multiple dilemmas facing the African continent. She had discussed the possibility of launching a research project on the language of instruction in Tanzania with colleagues at the University of Dar es Salaam. Together, Birgit and I, conceptualized a language in education project between the University of Dar-es-Salaam, the University of Oslo and the University of the Western Cape.

Later in the same year, I, gave a key-note address at the Oxford Conference (Brock-Utne, 2001a). It was called “Education for All - in whose language?” There I met Zubeida Desai from the language department at the University of Western Cape, who also presented a paper on the language issue in Africa - exemplified by the situation in the Western Cape. The two of us continued the discussion that had started between Harold Herman and me about a research cooperation around the Language of Instruction in Tanzania and South Africa. Zubeida and I decided that I should apply to NUFU for seed money to plan a project with partners from UDSM and UWC as part of a research project. The research project involved both a policy analysis and an experimental phase in both countries. In Tanzania, where the language of instruction shifts to English in Form I in secondary school, we decided that we would let some classes in Form I be taught science and geography in Kiswahili (we translated both the learning material and the tests into Kiswahili) for some months and compare the students with students taught in English only. In South Africa we looked into the situation for black isiXhosa speaking children living in a non-formal settlement near the airport of Cape Town. Both the teachers, the parents and the children spoke isiXhosa. That was the language of instruction only for the first three years. Then the language of instruction was supposed to switch into English, a language foreign to them all, a language they never communicated in outside of school. Here our intervention happened in grade 4 and 5. We translated learning material into isiXhosa. In both countries we found that children learnt much better when they were taught in a familiar language. Both in South Africa (Holmarsdottir, 2005; Nomlomo, 2007) and in Tanzania (Mwinsheikhe, 2007; Vuzo, 2007) the students said they learnt better and they also got better grades when concepts were explained to them in Kiswahili in Tanzania, and in isiXhosa in the informal settlement in the Western Cape in South Africa, languages they were familiar with.

Two of my Ph.D. students (Mwinsheikhe, 2007; Vuzo, 2007) and I (Brock-Utne, 2007) noticed that when the teachers were teaching in English, they frequently punished the students, mostly by having them stand all through the lesson. They never punished the students when the teaching was in Kiswahili. The teachers themselves were not aware of this fact, before we pointed it out for them.

¹«siasa» - a social science subject closely connected to the ideology of CCM—Nyerere’s party—was taught in Kiswahili all through the secondary school from 1969 to 1992 when the subject—under the influence of the reintroduction of a multiparty system—was renamed “civics” for the first 4 years and “general studies” for the next 2 years and taught in English (Mkwizu, 2002).

²In Norwegian: Norsk utvalg for Utviklingsrelatert Forskning og Utvikling (NUFU). Though it was started in 1991, it really took off only from 1996. The Fund, which originally came from NORAD, was administered by academics interested in developing countries from all of the universities in Norway (see Brock-Utne, 2019). It was located at the University of Bergen.

The introduction of private primary schools

It was important for the educator and President of Tanzania, Julius Nyerere, to create a good primary school for all children. In the policy directive, called "Education for Self-Reliance" Nyerere (1967) notes that for the education which independent Tanzania wants to build, "the purpose is not to provide an inferior education to that given at present. The purpose is to provide a different education" (p. 63). He wanted the educational system of Tanzania to emphasize co-operative endeavor, not individual advancement, and to stress concepts of equality and responsibility (p. 52).

In my book "Whose Education for All? The Recolonization of the African Mind" (Brock-Utne, 2000/2008) I show how imported assessment systems from the West along with the destruction of indigenous publishing and curriculum development have strengthened the neo-liberal policies outlined by the World Bank and have undermined the policies professed by a newly independent Tanzania. Nyerere saw it as important to have Kiswahili be the language of instruction all through the seven years of schooling. He was much against opening up for private schooling. He saw such a strategy as one which would undermine the equal society he wanted to create.

In an analysis of the formulation of educational policies in Tanzania in the nineties, the Danish researcher Buchert (1997a) found that the Tanzania Social Sector Review was worked out by the Bank (World Bank, 1995). When it comes to the "Education and Training Policy" (URT, 1995) paper, Buchert (1997a:52) reports that many of the government officials as well as bilateral aid agencies and people from the academic environment saw "a determined World Bank hand behind it." One of Buchert's interviewees said: "It has been stuffed down the throat of the Government by the IMF and the World Bank" (Buchert, 1997a: 52). Buchert mentions that during 1994 and 1995 the World Bank held a number of education seminars in Washington DC on Educational Policies in Africa and the Coordination of Aid. The seminars were held for key Tanzanian educators and had both a direct and an indirect impact on government thinking on education.

Work on the Education and Training Policy started in February 1993. According to Buchert (1997b) it was initiated due both to an internally felt need for an official policy that reflected the state of the art of education in the 1990s and to agency pressure for a policy framework to guide education assistance. The point of departure for the policy paper was the task force report The Tanzania Education System for the 21st Century—completed in 1992 and published in 1993 (URT, 1993). The Report analyzes educational needs in light of a future Tanzanian society able to cope in an increasingly globalized world. The Education and Training Policy (URT, 1995) opened up for private primary schooling. The language of instruction in these private primary schools is English. The private schools have mushroomed since the government opened up for them. Parents pay part of the costs of these schools. In the best schools parents have to pay very high school-fees, in the not so good schools parents pay less. The language of instruction also in these low-paying schools is still English or rather; bad and broken English. Parents send their children to private schools not so much because the language of instruction is English, but more often because these schools are better resourced. The pupils there have textbooks, there are fewer pupils in class, the teachers get their skills up-dated through in-service courses, are better paid and are not so often absent.

Strengthening of a government primary school

In the second phase—the last 5 years—of the LOITASA project the Tanzanian group decided to strengthen a government Kiswahili speaking primary school with textbooks for all the pupils in grade 6 and 7 and give the teachers in-service teacher training through two up-dating seminars. We gave this government school the fictitious name "Mweshipandeka". These two small interventions were enough to make the pupils in Mweshipandeka get better grades and be happier at school than the pupils in the English speaking private primary school with the highest school fees in the district. The teachers at Mweshipandeka were convinced that the better scores their pupils got had to do not only with the fact that they were provided with textbooks but also with the fact that they understood what the teacher was saying. They were taught in a familiar language (Bakahwemama, 2009, 2010).

But in the village where the experiment took place we heard several people say to each other, "Have you heard that Mweshipandeka" has become English medium?" The school had continued teaching in Kiswahili, but the village people thought that the fact that the school did so well, even better than the fee-paying private schools, meant that the school had switched to using English as the language of instruction. Nothing could be further from the truth.

The rhetoric of English and development

The growing craze for English all over the world is associated with the rhetoric of English and development, permeating into popular perception of the significance of English, often without any critical scrutiny. Broadly, English is projected as a global language (Graddol, 2000) or a language needed for maintaining a competitive edge in a globalized world. Yet, as Coleman (2011:104) notes:

Globalisation and competitiveness are associated with a need for English and then with a need to use English as a medium of instruction, although the logical relationships between these concepts remain unclear.

There is actually no logical relationship. The best way for a pupil living in a non-English speaking country to learn English is not have it as a language of instruction, but rather learn it as a subject, as a foreign language from teachers who are good at teaching foreign languages. Many parents are not aware of this fact.

The situation in India

Mohanty (2017) drawing on his work in respect of English in Indian society and education, discusses the processes through which English in India gets situated in a position of dominance, disadvantaging the other language communities. He analyses the role of English in Indian education in perpetuating social discrimination. He shows that while some groups benefit from English, most do not (Mohanty, 2017:266).

The first Prime Minister of India, Jawaharlal Nehru, was very concerned about the language question in India. He saw that the choice of English as the language of instruction created a class society. As leader of the socialist part of the Indian Congress party he had fought against the caste system of India. He saw that a new caste was being developed: an English speaking caste, separated from the rest of the people. Only 5% of the Indian population are fluent in oral and written English.

Mohanty (2017:275) notes that even though Prime Minister Nehru feared a development of an “English knowing caste” in India, he was not able to stop this development. In the struggle for Indian independence Mahatma Gandhi warned that English represented cultural alienation. Prime Minister Nehru (educated at Cambridge University and imprisoned by the British before independence) expressed in a letter that he was “convinced that real progress in India can only be made through our own languages and not through a foreign language. I am anxious to prevent a new caste system being perpetuated in India—an English knowing caste separated from the mass of our public” (cited in Gopal, 1980:25). Philipson (2003:6) comments:

In fact, an English-using caste has emerged, because the management of multilingualism in India has largely been left to market forces. These strengthen the position of users of English, here as elsewhere. Roughly 30 million Indians are fluent users of English, but they account for under 5 per cent of the population.

Philipson (2003) mentions that there are elites in Africa and India who speak exclusively English to their children. Mazrui (1978) called them the Afro-Saxons. It is not uncommon for Indian grandparents, who do not speak English, to have no language in common with their grandchildren. Philipson (2003:75) further laments: “The young upwardly mobile, internationally oriented generation of Indians and Africans have more in common with ‘global’ culture than with the mass of inhabitants of India and Africa.”

Mohanty writes that there are now English knowing sub-castes in India, differentiated on the basis of the level of competence in English. Some pupils have superb English and have been schooled in very expensive private schools with excellent teachers. Another sub-caste of pupils have only average English, even though they too have gone to English-medium private schools, but less expensive ones. The lowest sub-caste of pupils are those with poor English from the low-cost private English medium schools. He questions the practice of low-fee private schools, which claim to be providing English-medium education, but which in reality fail to teach English and fail to teach the subjects which are supposedly being delivered through English. Mohanty (2017:275) describing the English knowing sub-castes asks: “Whose development does English promote?” As my example from Tanzania shows, the same question could be asked there as well as in the other so-called “anglophone” countries in Africa, where only about five percent of the population are fluent speakers of English.

Referring to the role of “the superimposed international languages” in the African context, Heugh (1999, 306) points out that “these languages serve only the interests of the elites”. Thus, any claim of a positive role of English in development cannot be taken to be a universal phenomenon. English is not a culturally neutral medium that puts everyone on the same footing; it empowers some and disempowers many.

In their chapter on Language Policy and Education in the Indian Subcontinent Mohanty and Panda (2017) write about the “double divide”—the hierarchical relationship between English and the major national/regional languages and the other divide between the major languages and the Indigenous tribal minority ones (ITM). On the Indian subcontinent English is promoted along with the major national regional languages, while the ITM languages are grossly neglected. Languages in education reflect the linguistic double divide; private schools use English as the medium of teaching-learning whereas public schools are in the medium of the dominant regional languages. The affluent send their children to the English medium private schools. This situation is rather similar to the one we find on the African continent, where there is a double divide between the former colonial languages English, French and Portuguese and the larger regional languages like Kiswahili, Hausa, Oromo, Amharic on the one hand and the regional languages and community languages on the other. In Tanzania Kiswahili is the language of instruction for all children in government primary school. Though the language is a national language, most newspapers printed in it and is frequently used on television, there are rural areas where the language is not so frequently spoken. Children in these areas start school in a language they do not speak (Brock-Utne and Qorro, 2015). They normally become quite fluent in Kiswahili through primary school. But in secondary school they have to cope with another language as the language of instruction. And this

language is a foreign language with no similarity to Bantu languages. English becomes the language of instruction in secondary school. At this stage a third of the students drop out of school, most likely because they do not understand what the teacher is saying.

From one colonial language to the other in Rwanda

The new language policy of Rwanda is likely to have disastrous effects on the learning of school subjects for thousands of children. Since independence from Belgium, Rwanda had retained French as the official language and so was termed a “Francophone” country, although the whole population, Hutus and Tutsis alike, speak Kinyarwanda and many of them also speak Kiswahili. In Parliament, in administration at the national level and in the Supreme Court, Kinyarwanda is the language predominantly used. Before the genocide Kinyarwanda was used as the language of instruction in the first three grades, then a switch to French occurred. [Schweisfurth \(2006\)](#) mentions that the Government of Rwanda after the genocide insisted on a trilingual education policy (Kinyarwanda, French and English) to secure greater equity between groups who favored one or the other language. A trilingual policy might have been good for Rwanda, provided that Kinyarwanda—a language that is spoken by 99.4% of the population, according to the Government—had been the language of instruction with French and English learnt as foreign languages (as subjects). In 2008, the international “development” partners and the new elite in Rwanda decided to do away with French and to introduce English from the first grade of schooling. English is the language of Rwanda’s new elite—especially the Rwandan Patriotic Front under the leadership of Paul Kagame and other Tutsi returnees from so-called Anglophone countries. This did not work and in 2011 Kinyarwanda was reintroduced as the language of instruction, but only for the three first grades and then a switch to English as the language of instruction occurs. In many schools, especially in the capital Kigali, English is the language of instruction from the very first grade of primary school. This policy was implemented in violation of recommendations by UNESCO and the African Union ([Brock-Utne, 2017](#)).

The situation in the Nordic countries

The intention of the European Union with its Bologna Declaration³ was to streamline educational standards in Europe. The streamlining also had the consequence of strengthening English as the language of instruction. According to Luc Soete, the Rector of Maastricht University: “National languages were perceived as a hindrance for student mobility akin to customs barriers, so the creation of an open market in English is another way for them to sell their educational products” (referred to by [Thomas and Breidlid, 2015: 350–351](#)). Some years back I was involved in organizing a European master degree program involving three universities, one in Belgium (Leuven, where the languages of instruction are Flemish and French), one in Finland (Oulu, where the language of instruction is Finnish) and one at the University of Oslo (where the language of instruction is Norwegian). The language of instruction of the whole master program was to be English, whatever university the students visited.

When I taught at the University of Dar es Salaam (1988–1992) many of my students said they wanted to come to Norway and continue their Ph.D. studies at my university. I had to tell them that if they wanted to do so, they had to learn Norwegian. At that time we did not have any Ph.D. courses in English. Neither did we have any master courses. But this has changed as witnessed in the table below:

Studies in English at Norwegian universities and colleges

Year	2007	2012	2016
Number of studies in English, all places of study	2379	4543	5798
Percentage of all studies offered in English, all places of study	8.9	15.7	19.6

Source: [Schwach and Elken \(2018:61\)](#).

In 1997 I took the initiative of organizing a master of Comparative and International Education at the Institute for Educational Research at the University of Oslo. The master started in 1998 and I was the leader of it most of the years until 2008. It was a study mostly for our students from Africa, Asia and Latin-America but some few Norwegian students also enrolled. I noticed that the Norwegian students actually commanded the written English language better than e.g., the Tanzanian students. For the Norwegian students it was the first time they were exposed to the use of English as the language of instruction while the Tanzanian students had used that language for at least nine years. The Norwegian students had had Norwegian as the language of instruction but learnt English well as a foreign language, taught by teachers who were experts in teaching English as a foreign language.

³The Bologna declaration (in full, Joint Declaration of the European Ministers of Education convened in Bologna on 19 June 1999) is the main guiding document of the Bologna process.

Academic publishing—in whose language?

The Swedish language activist Lindblom (2009) notes that during the period 1960 to 1979 90% of Ph.D. theses delivered at the University of Copenhagen were written in Danish and 10% in English. From 2000 to 2004 100% were written in English! (Lindblom, 2009). Schwach and Elken (2018:51) note that in 2017 90.8% of the Ph.D. thesis examined at Norwegian universities were written in English. Only 8.5% were written in Norwegian. Of these 7.9% were written in the urban variety of Norwegian⁴ and only 0.6 in the more rural variety of Norwegian.⁵

At the level of master theses there has been a clear increase in the number of theses delivered in English over the last 20 years.

Higher rewards for academic publications in English

In 1991 Norwegian state institutions were given the possibility of introducing “performance salary” as a part of local salary negotiations. Before that all associate professors had the same salary and so did all professors. The whole reward system fits well with the commercialization of higher education, which has also hit European universities (Brock-Utne, 2002).

In 2004 the Norwegian Association of Higher Education Institutions published a dossier called: *Vekt på forskning*⁶ (UHR, 2004). This publication institutionalized a reward system dividing journals and publishing companies into three levels, level zero (no reward given to the institution or researcher—most publishing companies in developing countries belong to this category—even if they publish in English), level 1 (reward given), level 2 (higher reward given—normally three times higher as level 1).

On the internet one can find a list of 486 ranked publishing companies. Of these 55 companies are ranked at level 2, while 431 companies are ranked at level 1. No Norwegian publishing company is ranked at level 2, not even the University Publishing Company. More than 80% of the publishing companies ranked at level 2 are based in the US.⁷

When it comes to academic journals, a list of 1758 ranked journals are given, among which a 10th are ranked at level two and the rest at level one. Only two of the many peer-reviewed academic journals published in Norwegian have been ranked at level two. Within the field of Educational Research no academic journal, where any of the articles is written in another language than English, has been ranked at level two (Brock-Utne, 2001b, 2009).

The engagement of many Norwegian academics in the defense of Norwegian as an academic language can be contrasted with the attitude of many Dutch academics. In 1989 Prof. Ritzen was appointed the Minister of Education in Holland. Minister Ritzen, who has a doctoral degree both in economics and physics, and had studied in the United States, had as a professor of economics in Holland felt frustrated because of the use of Dutch in the academia. As a Minister he now proposed that English should be the sole medium of instruction in all Dutch universities. His proposal met with overwhelming support from the academia. His proposal met, however, with harsh critique when it was presented in Parliament. Parliament insisted on regulating the language issue because it neither trusted the Minister,⁸ nor the academics. Therefore Parliament passed an amendment to the university law now saying that no courses can be offered in another language if it is not offered in Dutch. This was actually seen as a step backwards for those professors who wanted more English language instruction in Dutch higher education. There has, however, still been a steady growth of Master courses taught in English within Dutch higher education.

Some years back I felt that there was an acute need for a textbook in comparative and multicultural education written in Norwegian. I took the initiative to edit such a book, which appeared in 2006 (Brock-Utne and Bøyese, 2006). With two exceptions all the authors were native Norwegians. We all did most of our academic writing in English. Almost all the authors, including myself (Brock-Utne, 2006) had problems finding academic terms in Norwegian, describing phenomena we normally wrote about in English. Not long after the book was published my institute decided that the course in comparative and international education at the Bachelor level, which had been taught in Norwegian, should hereafter be taught in English. That may be one of the reasons why the book has not reached the sale figures we had hoped for.

The 2008 White paper on Norwegian language policy

In the summer of 2008 the Norwegian Ministry of Culture and Church presented a White paper to Parliament on Norwegian Language policy.⁹ I found that the paper was disappointing on two accounts and wrote an article in the largest Norwegian newspaper, *Aftenposten*, about this (Brock-Utne, 2008). This article led to a radio debate between the Minister of Culture, Trond Giske from the Social-Democratic party, and me on the July 21, 2008. The Norwegian government in 2008—the so-called red-green government—was a coalition consisting of the following three parties: The Social Democrats, the Center party and the Socialist Left party. These three parties together with the Christian People's party in 1995 had Parliament adopt the paragraph: “The language

⁴This variety is in Norwegian called bokmål and is developed from Danish.

⁵This variety is in Norwegian called nynorsk and is built on dialects spoken all over Norway.

⁶In English: Emphasis on Research.

⁷This is the web-site dealing with the ranking of publications: <https://dbh.nsd.uib.no/publiseringskanaler/Forside>.

⁸Former Minister Ritzen later worked in the World Bank in Washington DC.

⁹The White paper was called *Mål og mening. Ein heilskapleg norsk språkpolitikk* (Language and meaning. A holistic Norwegian language policy).

of instruction in Norwegian universities and colleges shall as a rule be Norwegian” as part of the Law of Universities and Colleges. But the bourgeois coalition government in the beginning of the new millennium—consisting of the Conservative party, the Liberals and the Christian party saw to it that this paragraph was taken out of the University law of 2005, even though Norwegian is more threatened as an academic language to-day than it was in 1995.

One would have expected that the red-green government would argue in their White Paper on Norwegian language policy for reinserting this paragraph in the Law of Universities and Colleges. The White Paper “Mål og mening” mentions that the Norwegian Language Council in 2006 had suggested to reinsert this paragraph. The White Paper argues against the Language Council as the Minister of Culture did against me in the radio debate in July 2008, maintaining that deletion of this paragraph was a consequence of the “internationalization of higher education.” But it is exactly the pressure from the internationalization of higher education which requires a legal protection of the Norwegian language in places of higher learning. In 2001 the Legal Court of the EU denied Iceland the right to have lower taxation of Icelandic literature than of literature written in foreign languages—a practice Iceland had embarked on to protect their own language. Knowing about this fact I asked our Minister of Culture whether he was afraid of sanctions from the EU if Norway wanted further promotion of the Norwegian language through legal measures. The Minister chose not to answer that question.

My other caveat when it comes to the White paper of the Norwegian Language policy of 2008 has to do with the fact that the paper goes against the proposal from one of our law professors, Professor Ola Mestad, of giving the Norwegian language legal protection in our Constitution like France has done for the French language. The indigenous minority language, the Saami language, is given legal protection in our Constitution—the majority language, Norwegian, is not.

Norwegian researchers defending our language as a language of research

Norwegian academics seem to be more prepared to defend the Norwegian language as an academic language than colleagues in some other European countries, like especially in the Netherlands.

On the fifth of May 2006 a petition in defense of Norwegian as an academic language signed by 223 well-known Norwegian professors from the humanities and the social sciences was published in our largest newspaper *Aftenposten*. This is the newspaper which also had been the leading news channel for the debate. The Norwegian case shows how a smaller European language, like Norwegian, is threatened as an academic language. When Norwegian academics are discouraged from publishing in Norwegian, it means that academic Norwegian will deteriorate and vocabulary will not be further developed. We shall reach a situation, which African academics are in, when they have difficulties discussing academic matters in African languages because the academic concepts have not been developed in their languages.

All languages develop through use and they also fail to develop or stagnate through disuse.

Where there is a political will—an example from Tanzania

At the University of Dar es Salaam there is one department that uses Kiswahili as the language of instruction and one research institute that uses Kiswahili as the language of communication. The department is the department of Kiswahili and the research institute is the Institute of Kiswahili Research. Also this department and research institute used to be run in English. There were big discussions before the language was switched to Kiswahili. Some of the lecturers and researchers claimed that it would not be possible to teach and write in Kiswahili because they lacked terms for “guttural sounds”, “phonemes”, even “terminology”. Others said: “we shall just invent these terms as we go along.” That became the political decision and now all the linguistic terms have been developed in Kiswahili. The terminology (istalahi) has been developed. This means that there is no problem in holding a linguistic conference in Kiswahili in East Africa.

Prestige Planning

Kamwangamalu (2015) suggests that language-education planning needs a paradigm shift. Rather than critiquing inherited colonial policies, he proposes what he calls *Prestige Planning* for African languages. According to Kamwangamalu (2015:14) Prestige Planning:

... entails associating African languages with an economic value on the labour market and requiring academic skills in these languages as one criteria for access to employment

The focus is on the reception (and not on the production) of language planning. Kamwangamalu (2015) argues that until African languages are associated with a market value, English will continue to dominate the educational systems of African Commonwealth countries, much as it did in the colonial era. But how can African languages be assigned a market value to

make them instrumentally competitive with English at least in the local labor market? The call for Prestige Planning for African Languages is made against the background of theoretical developments in language economics (Grin, 1996; Vallancourt and Grin, 2000). Kamwangamalu proposes Prestige Planning for African languages if these languages are to become, like English, an instrument of upward social mobility.

While the translanguaging approach being used especially among the black learners in South Africa (Makalela, 2014) or the code-switching approach used in many other African countries are strategies used on the micro-level, Prestige Planning is a macro-strategy which requires political willingness to change the power relations in a country. It is a strategy that would lead to the development of the majority population of Africa and thereby the African continent.

In this paper I have shown the hegemonic dominance of English in education and its consequences for the majority population in the so-called Anglophone countries in the developing world. The Anglo globalization of education has also hit non-English speaking countries in Europe, undermining the national languages as languages of academic publishing and research.

References

- Bakahwemama, J.B., 2009. What is the Difference in Achievement of Learners in Selected Kiswahili and English Primary Schools in Tanzania? MA thesis Institute for Educational Research, University of Oslo.
- Bakahwemama, J.B., 2010. What is the difference in achievement of learners in selected Kiswahili and English primary schools in Tanzania? In: Desai, Z., Qorro, M., Brock-Utne, B. (Eds.), *Educational Challenges in Multilingual Societies*. African Minds, Cape Town, pp. 204–229.
- Bakahwemama, J.B., 2016. Change of Language of Instruction for Science and Mathematics in Grade Five in Zanzibar: What Was the Reasoning Behind the Policy Change? PhD thesis Institute for Educational Research, University of Oslo.
- Brock-Utne, B., 2002. Language, Democracy and Education in Africa. Discussion Paper No. 14. The Nordic Africa Institute, Uppsala.
- Brock-Utne, B., Bøyesen, L. (Eds.), 2006. Å greie seg i utdanningssystemet i nord og sør: Innføring i flerkulturell og komparativ pedagogikk, utdanning og utvikling [How to Survive in the Educational System in the North and in the South: Introduction to Multicultural and Comparative Education, Education and Development]. Fagbokforlaget, Bergen.
- Brock-Utne, B., Qorro, M., 2015. Multilingualism and language in education in Tanzania. In: Yakoumetti, A. (Ed.), *Multilingualism and Language in Education: Sociolinguistic and Pedagogical Perspectives From Commonwealth Countries*. The Cambridge Education Research Series. Cambridge University Press, Cambridge, pp. 19–31. ISBN 978-1-107-57431-1.
- Brock-Utne, B., Skattum, I. (Eds.), 2009. *Languages and Education in Africa: A Comparative and Transdisciplinary Analysis*. Symposium Books, Oxford.
- Brock-Utne, B., 2000/2008. Whose Education for All? Recolonization of the African Mind. Falmer Press, New York, p. 340. Reprinted in 2008 in Seoul: Homi Publishers.
- Brock-Utne, B., 2001a. Education for all—in whose language? *Oxf. Rev. Educ.* 27 (1), 115–134.
- Brock-Utne, B., 2001b. The growth of English for academic communication in the Nordic countries. *Int. Rev. Educ.* 47 (3–4), 221–233.
- Brock-Utne, B., 2006. Innviklet utdanning for utvikling (Difficult education for development). In: Brock-Utne, B., Bøyesen, L. (Eds.), *Å greie seg i utdanningssystemet i nord og sør: Innføring i flerkulturell og komparativ pedagogikk, utdanning og utvikling* [How to Survive in the Educational System in the North and in the South: Introduction to Multicultural and Comparative Education, Education and Development]. Fagbokforlaget, Bergen, pp. 221–236.
- Brock-Utne, B., 2007. Learning through a familiar language versus learning through a foreign language: a look into some secondary school classrooms in Tanzania. *Int. J. Educ. Dev.* 27 (5), 487–498.
- Brock-Utne, B., 2008. En sprikende språkmelding (An Ambivalent White Paper on Language Policy). *Aftenposten*.
- Brock-Utne, B., 2009. Is Norwegian threatened as an academic language? In: Brock-Utne, B., Garbo, G. (Eds.), *Language is Power. The Implications of Language for Peace and Development*. Mkuki na Nyota, Oxford, Dar es Salaam. African Books Collective. East Lansing: Michigan State University Press.
- Brock-Utne, B., 2012. Language and inequality: global challenges to education. *Compare* 42 (5).
- Brock-Utne, B., 2014. Language of instruction in Africa—the most important and least appreciated issue. *Int. J. Educ. Dev. Africa* 1 (1), 4–19. ISSN 2312-3540.
- Brock-Utne, B., 2017. Multilingualism in Africa, marginalisation and empowerment. In: Coleman, H. (Ed.), *Multilingualism and Development*. British Council, London, pp. 64–80 (Chapter 4).
- Brock-Utne, B., 2019. Models of cooperation between a university in Norway and two universities in Africa: an autoethnographic report. In: Halvorsen, T., Orgeret, K.S., Krøvel, R. (Eds.), *Sharing Knowledge, Transforming Societies: The NORHED Programme 2015–2018*. African Minds, Cape Town, pp. 379–403. ISBN: 978-1-928502-00-5, e-Book: 978-1-928502-01-2 e-Pub: 978-1-928502-02-9. Pages in book: 612. Ch. 16.
- Brock-Utne, B., 2022. The language of instruction in Asia and Africa—some issues. *Sri Lankan J. Educ.* 1 (1).
- Buchert, L., 1997a. Education Policy Formulation in Tanzania: Coordination Between the Government and International Aid Agencies. International Institute for Educational Planning, Paris.
- Buchert, L., 1997b. Coordination of Aid to Education at the Country Level: Some Experiences and Lessons From Tanzania in the 1990s. Paper Presented at the Oxford International Conference on Education and Development, “Education and Geopolitical Change,” 11–15 September.
- Coleman, H., 2011. Allocating resources for English: the case of Indonesia’s English medium International Standard Schools. In: Coleman, H. (Ed.), *Dreams and Realities: Developing Countries and the English Language*. British Council, London, pp. 89–113. <http://www.teachingenglish.org.uk/transform/books/dreams-realities-developing-countries-english-language>.
- Gopal, S. (Ed.), 1980. Jawaharlal Nehru. An anthology. Oxford University Press, Delhi.
- Graddol, D., 2000. *The Future of English*. British Council, London.
- Grin, F., 1996. The economics of language: survey, assessment and prospects. *Int. J. Sociol. Lang.* 121, 17–44.
- Herman, H.D., 2008. The start and progress of a language of instruction project in Africa—the spirit of Bagamoyo. In: Qorro, M., Desai, Z., Brock-Utne, B. (Eds.), *LOITASA: Reflecting on Phase I and Entering Phase II*. E&D Vision Publishing, Dar es Salaam, pp. 10–26.
- Heugh, K., 1999. Language, development and reconstructing education in South Africa. *Int. J. Educ. Dev.* 19, 301–313.
- Holmarsdottir, H., 2005. From Policy to Practice: A Study of the Implementation of the Language-In-Education Policy in Three South African Primary Schools. PhD dissertation. University of Oslo.
- Kamwangamalu, N.M., 2015. The sociolinguistic and language education landscapes of African Commonwealth countries. In: Yakoumetti, A. (Ed.), *Multilingualism and Language in Education: Current Sociolinguistic and Pedagogical Perspectives From Commonwealth Countries*. Cambridge University Press, Cambridge, pp. 1–18.
- Lindblom, P.Å., 2009. Are the Nordic Languages threatened as academic languages? In: Brock-Utne, B., Garbo, G. (Eds.), *Language and Power. The Implications of Language for Peace and Development*. Mkuki na Nyota, Dar es Salaam, pp. 283–289.
- Makalela, L., 2014. Teaching indigenous African languages: the effects of translanguaging for multilingual development. In: Hibbert, L., van der Walt, C. (Eds.), *Multilingual Universities in South Africa: Reflecting Society in Higher Education*. Multilingual Matters, Bristol, pp. 88–104.
- Mazrui, A., 1978. *Political Values and the Educated Class in Africa*. University of California Press, Berkeley.

- Mkwizu, M., 2002. The Pedagogical Implications of Using English in Teaching Civics in Tanzanian Secondary Schools. Master thesis. Institute for Educational Research. Unit for Comparative and International Education, University of Oslo, p. 120.
- Mohanty, A.K., Panda, M., 2017. Language policy and education in the Indian subcontinent. In: McCarty, T.L., May, S. (Eds.), *Language Policy and Political Issues in Education*, third ed. Springer International Publishing, pp. 507–519.
- Mohanty, A., 2017. Multilingualism, education, English and development: whose development? In: Coleman, H. (Ed.), *Multilingualism and Development*. British Council, London, pp. 261–281.
- Mwinsheikhe, H.M., 2007. Overcoming the Language Barrier: An In-Depth Study of Tanzanian Secondary School Science Teachers' and Students' Strategies in Coping With the English/Kiswahili Dilemma in the Teaching/learning Process. PhD thesis. University of Oslo: Faculty of Education.
- Nomlomo, V.S., 2007. Science Teaching and Learning Through the Medium of English and isiXhosa: A Comparative Study of Two Primary Schools in the Western Cape. Unpublished PhD dissertation. University of the Western Cape, Cape Town.
- Nyerere, J., 1967. Education for Self Reliance. Policy booklet published in March 1967. Reprinted in: Nyerere, Julius, 1968: *Ujamaa: Essays on Socialism*. Oxford University Press, Dar es Salaam, pp. 44–76.
- Phillipson, R., 2003. *English-Only Europe? Challenging Language Policy*. Routledge, London & New York.
- Schwach, V., Elken, M., 2018. Å snakke fag på et språk andre forstår. Norsk fagspråk i høyere utdanning og arbeidsliv (Speaking a Language People Understand. Norwegian Academic Language in Higher Education and in Employment). NIFU-STEP, Oslo. RAPPORT 20/2018.
- Schweisfurth, M., 2006. Global and cross-national influences on education in post-genocide Rwanda. *Oxf. Rev. Educ.* 32 (5), 697–709.
- Thomas, P., Breidlid, A., 2015. In the shadow of "Anglobalization", national tests in English in Norway and the making of a new English underclass. *J. Multicult. Discourses* 10 (3), 349–368. <https://doi.org/10.1080/17447143.2015.1041963>.
- UHR (Universitets og Høgskolerådet – the Norwegian Association of Higher Education Institutions), 2004. Vekt på forskning . Nytt system for dokumentasjon av vitenskapelig publisering (Emphasis on Research. A new system for documentation of scientific publishing).
- URT (United Republic of Tanzania), 1993. The Tanzania Education System for the 21st Century: Report of the Task Force, 1993. United Republic of Tanzania, Ministry of Education and Culture and Ministry of Science, Technology, and Higher Education, Dar es Salaam (Printed by the Print Service, a division of University of Leeds Media Services).
- URT (United Republic of Tanzania), 1995. Education and Training Policy. Ministry of Education and Culture.
- Vallancourt, F., Grin, F., 2000. The choice of a language of instruction: the economic aspects. In: Distance Learning Course on Language Instruction in Basic Education. World Bank Institute, Washington, DC.
- Vuzo, M.S., 2007. Revisiting the Language of Instruction Policy in Tanzania: A Comparative Study of Geography Classes Taught in Kiswahili and English. PhD dissertation. University of Oslo.
- World Bank, 1995. Tanzania Social Sector Review. World Bank, Washington, DC.

Impact of globalization on higher education in the Middle East and North Africa

Abdeljalil Akkari, University of Geneva, Genève, Switzerland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	380
From nomadic scholars to colonial dominance	380
University as a national development project	381
International cooperation and privatization	382
Rethinking university in MENA uncertain times	384
Missed links between education and development	385
References	387

Introduction

Contemporary higher education is submerged by rapid social, political, economic and cultural transformations, which affect most countries in various ways and have an impact on its functioning. Changes in higher education in Middle East and North Africa (MENA) countries are linked to globalization, understood as a multidimensional economic, political and cultural process, which dynamic have direct and indirect repercussions on universities. In the context of globalization, economic, political and cultural activities do not stop at the boundaries of national societies, but tend to transcend territorial boundaries. Overall, MENA region lags behind most of the rest of the world on many key indicators of higher educational development. Although several countries in this region have profited greatly from the considerable wealth derived from large oil and gas revenues over the past half century, universities are not performing high in international comparisons.

The term globalization within the university context typically refers to the increasing willingness of universities to be open to international influence: exchange programs, staffs and students, expand partnerships between institutions based in various countries, and to harmonize academic programs. Such globalization may be highly desirable. Yet, globalization affects actors in a highly disproportionate way. Under the guise of a universalist or international narratives, it increases inequality, reinforces a competitive approach of learning, and removes national specificities and interests. Because their activities take place in an economic, historic, social, and geographic context that obviously cannot be conveyed by few indicators or standards, universities are affected in various and different ways by globalization (Strassel, 2018).

Figge and Martens (2014) categorized globalization according to four different trends: (1) internationalization, that is, transactions across country borders; (2) liberalization, in reference to the policies or efforts for neoliberal economic globalization; (3) universalization or Westernization, which is the spread of certain values, objects and experiences; and (4) de-territorialization, in reference to international connections with a significant autonomy from territorial locations and national borders. These 4 trends apply with different intensity in the MENA region and radically transform the landscape of higher education.

Globalization affects state education policies, with specific consideration of the balance of transnational/international forces and national resistances, capacities and responses. This relationship between the global and the local in education is a core preoccupation and relates directly to the possibilities for a politics of education (Ozga and Lingard, 2007). Internationally oriented public universities are caught between their internationalizing student bases and research aspirations on the one hand, and the priorities of their national and sub-national governments as public funders on the other (Knight et al., 2021).

In particular, in many countries across the MENA, globalization increased the weight of international organizations and the privatization and commodification of higher education. As suggested by Robertson et al. (2007), a key element of globalization is the thickening of multilateral interactions and interconnections in the global governance landscape. The traditional International Organizations (IOs), such as the UN institutions (World Bank, UNESCO and ECLAC) and the OECD, are now joined by an expanding array of non-governmental international organizations and transnational firms. Another effect of globalization is that higher education sector, having modeled its goals and strategies on the market-oriented and entrepreneurial business model, is compelled to embrace the corporate ethos of the efficiency, accountability and profit-driven managerialism (Zajda, 2020).

From nomadic scholars to colonial dominance

Universities in the Arab-Muslim world began to emerge between the ninth and the 10th centuries. These include Ez-Zeituna (737, Tunisia), Al Quaraouiyine University (859, Morocco), al-Azhar University (970, Egypt) and Nizamiyyah University in Baghdad (Iraq, 1062). Some historians limit the concept of university to the Euro-Christian world. However, we can say that medieval Islamic universities saw the emergence of many scholars in diverse fields of knowledge ranging from mathematics to philosophy. Major ground-breaking advances in mathematics, astronomy, optics, navigation, architecture, irrigation and agriculture, medicine, and the development of hospitals, commerce and trading, and the willing acquisition of transcultural knowledge were special features of the Islamic Golden Age (Hillman and Baydoun, 2018; Lai et al., 2016).

Two characteristics deserve to be highlighted in this golden age of Arab-Muslim universities: the multiculturalism of ancient universities in the MENA region and the links with other civilizations, in particular, Europe. On the one hand, the crossing of ethnic, religious and continental borders of thinkers. Many thinkers in Arab-Muslim universities came from regions far from the center of Arab-Muslim civilization. They can come from India, Central Asia, Andalusia and sub-Saharan Africa. They can also be of Jewish faith. These thinkers circulate in the Arab-Muslim space. For example, al-Farâbî circulated between Kazakhstan and Syria. Ibn Khaldoun navigated between North Africa, Spain and Egypt. They were truly nomadic thinkers. Hence, multiculturalism and multilingualism characterize medieval thinkers in the MENA region. On the other hand, there were many links and exchanges of ideas with other civilizations. Ties have been intense with Europe, Africa, India and China. In Europe, what is attested is the contribution of Arab scholars to knowledge by translations from Arabic to Latin of philosophers and thinkers of Arab-Muslim culture: al-Kindi (802–866), al-Farâbî (850–950), Ibn Sinâ (Avicennes) (980–1037), Ibn Rushd (Averroes) (1126–1198). The density of intellectual exchanges between the so-called Arab-Muslim civilization and the rest of the world is a reality that is difficult to deny (Freely, 2010).

However, progressively Guardians of Islamic law (Sharia) developed an uncaring attitude toward secular disciplines, critical thinking and excluded them from the curriculum of the madrasa system. This action deprived the Islamic education system from intellectual challenge and stimulation. Learning by rote, a process that encourages instruction rather than original and critical thinking impoverished intellectual creativity in the region.

First modern universities were founded in the MENA region during the political turmoil of the late 19th century, which brought together the Arab awakening, decline of the Ottoman Empire, and European colonialism and imperialism. Education and higher education were associated with nationalist and geopolitical issues in the region. While Arab nationalist seek to take control of higher educational institutions, Western powers were engaged in a severe competition in higher education. The creation in 1865 of the Syrian Protestant College (future American University of Beirut), followed ten years later by the foundation of Saint Joseph University by a French Jesuit mission exhibit the intense cultural transfer between Western World and MENA region during this period. Competition of colonial powers and nationalist struggles combined here to confer significant symbolic power to higher education, in Egypt, Lebanon and Palestine in particular (Romani, 2012).

University as a national development project

After World War II, the newly independent Arab States progressively promoted an equalizing nationalist ideology and staked their legitimacy in their ability to foster economic development and provide social welfare to their citizens. Education was a central component of the states' nation-building projects. Higher education became an important means to garnering political support from the upwardly mobile middle classes, while also centralizing state power (Mazawi, 2005). A new dynamic contributes to the politicization of the question of higher education in the Arab world after the Second World War and the newly acquired independence. New national governments started ambitious development programs, which focus on education, technology and science (Siino, 2003). In this context of the "developing state", both in the version linked to the western bloc or in the other version linked to Soviet Union, many universities of research institutes were created in the region, often with the government guarantee of employment for any university graduate. From Iraq to Morocco, thousands of students accessed higher education until then the prerogative of narrow social elite (Siino, 2003).

It is possible to point out both massification of higher education in MENA and durability of social inequalities by the end of 20th century. In Egypt, higher education increased from just above 10% among those born in 1955 to almost 30% by the 1985 cohort. In Jordan, higher education amplified from around 25% for the 1955 cohort to almost 40% by the 1985 cohort. Tunisia's education system expanded later than Egypt or Jordan. Higher education was below 10% for cohorts born into the 1970s, but higher education expanded rapidly in recent decades, approaching 30% by the 1985 cohort (Krafft and Alawode, 2018).

Despite this rapid growth in higher education enrollment, there is substantial inequality of opportunity in Egypt, Jordan, and Tunisia. Tunisia and Egypt had the most important inequality of opportunity, whereas Jordan was considerably less unequal. The two primary drivers of inequality in Egypt are mother's education (27.2% of inequality) and father's education (42.3% of inequality). There are also smaller contributions from father's work sector and region. Jordan has less inequality of opportunity. Jordan would need to redistribute 18.7% of opportunities to be equality of opportunity. The inequality in Jordan is primarily related to mother's education (35.8%) and father's education (43.0%). Father's job sector as well as number of siblings also contribute to inequality. Inequality is high in Tunisia, where 37.3% of opportunities to attain higher education would have to be redistributed for equality of opportunity to prevail. This inequality is driven primarily by father's education (42.9%), followed by mother's education (22.1%) and urban/rural disparities (15.8%). There are small contributions from father's work sector, region, and number of siblings. Across countries, sex contributes less than 3% to inequality (Krafft and Alawode, 2018). An important finding of Krafft and Alawode (2018) is that, despite policies centered on free public education, equality of opportunity in higher education does not prevail in MENA. Free higher education is a regressive policy that primarily benefits the rich. The two countries that guarantee free public higher education, Egypt and Tunisia, have the highest inequality of opportunity.

There were about ten universities in the Arab world in the 1940s. The number is in the hundreds today. The share of the rich Gulf countries in this quasi-exponential growth is itself growing. A particular feature of this university boom is the overlap of privatization and claimed globalization. Two-thirds (or about 70) of the new universities founded in the region since 1993 are private, and are mostly branches of US universities. Even in Saudi Arabia, the government, which operates 8 public universities, agreed in 2008 to establish two private universities and numerous post-secondary education institutes (Romani, 2012).

Table 1 Gross enrollment ratio for tertiary education, both sexes (%).

Country	2013	2014	2015	2016	2017	2018
Algeria	33.89	34.48	36.78	42.62	47.64	51.36
Bahrain	39.54	41.32	43.27	46.56	47.14	50.48
Egypt	30.11	31.06	35.02	33.85	35.16	–
Jordan	–	–	36.63	35.58	31.14	34.41
Kuwait	–	51.02	55.14	57.28	55.36	54.36
Mauritania	5.41	–	5.61	5.35	4.99	–
Mauritius	40.04	39.12	37.41	38.84	40.59	–
Morocco	22.64	25.33	28.40	31.98	33.78	35.93
Oman	29.93	34.55	39.23	44.05	38.12	38.036
Palestine	45.85	44.76	45.28	43.92	43.67	44.25
Qatar	11.99	13.58	14.70	15.64	16.62	17.86
Saudi Arabia	52.23	58.29	61.05	67.33	69.69	68.03
Sudan	17.67	16.90	16.91	–	–	–
Syrian Arab Republic	32.84	41.58	42.67	40.05	–	–
Tunisia	34.95	35.33	35.19	32.81	32.14	31.74
Western Asia and Northern Africa	36.71	38.97	41.67	43.87	45.78	46.26
Arab states	28.33	29.60	31.17	31.98	32.90	33.36

Prepared by author using data from UIS-UNESCO (2020).

Table 1 below compare ratio for tertiary education in several MENA countries. We observe both important growth between 2013 and 2018 and differences between countries in the region. Saudi Arabia has the highest enrollment rate (68%) and Mauritania the lowest (5%).

According to [Calderon \(2018\)](#), gross enrollment ratio of Arab States in 2016 (32.0%) which stood at the same level of Latin America and the Caribbean in 2006 or the world's average in 2012. This improvement in participation is a result of increased diversification and focus on part of states to develop higher education, based on the US-style liberal arts institution model in the Middle-east and on the French public model in the Maghreb. Higher education reform in the region is confronted with a myriad of challenges that spread over demographic, political and economic considerations, and these are unlikely to overcome in the next years. States in this region with the largest number of enrollments are Egypt (2.8 million in 2016), followed by Saudi Arabia (1.6 million) and Algeria (1.4 million).

International cooperation and privatization

Globalization of higher education is slow because of the following realities and constraints. First, educational systems all over the world have to preserve political and cultural identity and serve political and ideological objectives. Second, universities and other higher education institutions worldwide retain their autonomous, individualistic, institutional and elitist characteristics. Even today, universities and other tertiary institutions are sealed into national contexts and most especially in developing countries, are state institutions. Third, it is a fallacy to talk about the globalization of higher education where social and economic inequalities are extensive ([Ishengoma, 2003](#)).

While national context is still the most important factor shaping university, we observe worldwide a growing impact of globalization on Higher education in the 21st century. In MENA region, we consider that international organizations and privatization are two key tenets affecting higher education and its links with globalization. For a decade or two after gaining independence, national higher education systems in MENA countries were structured and functioned primarily within the boundaries of individual countries. However, from the mid-1970s, a series of new phenomena, directly and indirectly linked to higher education, appeared and considerably modified its modus operandi to extend its presence and action beyond national borders ([Martins, 2019](#)).

As a global player in international cooperation, the World Bank is involved in almost all MENA countries in projects to reform the governance and to improve the quality of education of higher education. This involvement comprises both funding but also diagnostic and policy recommendations ([Samoff and Carrol, 2003](#); [Shahjahan, 2012](#)). The [World Bank \(2008\)](#) called MENA countries for flexibility and additional public funding that could be associated with outcomes and innovation, thus ensuring accountability for performance. Professionalization of staff was also considered as a top priority. The European Union within the framework of EUROMED policy ([Fouchet and Moustier, 2010](#)) adopted in the region the same line of conduct than the World Bank. In addition, bilateral cooperation particularly in the Maghreb shaped higher education reforms ([Mullin, 2017](#)).

Privatization and blurring of the lines between public and private characterize the financing of higher education in the MENA region. While privatization is high in many countries, it presents contrasting management models illustrating diverse degrees of government control over higher education. In Qatar, funding is mainly public through the Qatar Foundation. Since 2003, the campus called Education city has hosted at least 9 foreign universities (6 US, one French, one British) in addition to a Qatari university and a dozen research centers. The state covers construction and investment costs there but the universities remain private and

relatively autonomous. In the United Arab Emirates, the International Academic City was established in 2007 in Dubai, where it is part of the "Knowledge Village", built in a free zone and hosting 27 universities, mostly foreign. Funding is more symmetrical there because these universities have to cover their own expenses, in what is referred to as a co-investment operation. Finally, in Saudi Arabia, government control remains the main norm. King Abdullah University of Sciences was founded in 2009, with a sovereign wealth fund of \$10 billion, making it the richest university in the world even before it started operating (Romani, 2012).

The privatization of universities in the MENA region is unfolding mainly through four mechanisms found in other parts of the world. First, privatization can be linked to historical arrangements that have seen the emergence of private universities linked to religious minorities. This is typically the case in Lebanon, Syria and Palestine where the birth of private universities at the end of the 19th century was linked to Christian religious congregations.

Second, we find privatization linked to public policies aimed at increasing the proportion of students in private universities. This can be accelerated by regulatory and fiscal incentives from the state or by public-private partnership (PPP). We may consider offshore campuses as a form of PPP, particularly in the rich gulf states where public funding is used in building a network of foreign universities.

Third, privatization can also emerge through the degradation of the quality of public education and the dynamism of private entrepreneurs, finding in higher education an opportunity for economic gain.

Fourth, privatization is also an ideology/trend present in public universities by the promotion of good governance, accountability and entrepreneurship.

While these four mechanisms of privatization are present in most regions of the Global South, it is important to observe two specificities in the MENA region. The first is the key role played by the state in promoting and accelerating privatization. By injecting a lot of public funding used by private universities, the state is a main player of privatization. The second specificity is the historical mission played by private universities in promoting foreign models of higher education, mostly the Anglo-Saxon model, but also the French model and to a less extent and more recently the German, Japanese or Turkish models.

In Lebanon, students in private higher education come from the middle-class families, i.e., government officials and employees of the public and private sectors benefiting from collective agreements, granting a financial support for their children education. Taking advantage of this principle of freedom of choice of school, the overwhelming majority of these civil servants and employees choose private education, including at the university level, in search of better quality, but also of the symbolic prestige attached to elitist private universities. Thus, we observe the existence of a subtle mechanism that allows private higher education to be partially financed with public funds (Kabbaj, 2012). This situation illustrates the blurring borders between private and public sectors.

Alliances, collaboration, and partnerships forged between higher education institutions in the North and South are one-sided and often benefit one side only. In a situation where countries and multinational corporations in the North are dominating the process of globalization, to talk about globalization of higher education is synonymous to neocolonialism or academic imperialism in higher education. South countries have little to offer in the process of globalization other than receiving whatever accrues from the process. Educational globalization may be not different from such earlier concepts as neocolonialism, dependency, or center-periphery (Ishengoma, 2003).

It is also important to observe that privatization is not following the same trend in the sub-regions across MENA: The Maghreb, the "Middle East," and the Gulf (Oil) States. Compared to the Middle East where privatization of higher education is old, North Africa seems to be more reluctant to massive privatization: « The nations' strong orientation toward the French model, and its commitment to a free public system were specifically cited as the reason for the decade long delay in the implementation of Morocco's privatization reforms and Tunisia's continuing hesitance to support private higher education. » (Buckner, 2018, p. 1302). However, the growth of private higher education both in quantity and quality have been observed in Morocco, Tunisia and more recently in Algeria in the last two decades. In the Gulf States, higher education development is recent and privatization is a policy endorsed and supported by the state. Offshore campuses and foreign university represent a significant share of higher education enrollment (Powell, 2014).

After several decades of international cooperation and privatization, the results of higher education in the MENA region are weak. Young people and their family invest in their higher education with hope to get better job and greater opportunities. However, higher education has become a source of widespread frustration because it is not delivering the skills and competencies needed by young people today. Universities in the region are still centered on credentials rather than competencies (El-Kogali, 2018). Rather than leading to qualified jobs and greater well-being, a university degree in the MENA region is more likely to lead to frustration, and a dead end of unemployment. More curious is the situation of the rich countries of the Gulf who have invested heavily to build public and private universities of good quality. Yet, these countries are experiencing a high unemployment rate of national graduates while importing numerous foreign workers with a university level. The achievements of Arab higher education systems do not seem to be on par with the economic weight and the long scholarly tradition of this region. Compared to emerging economies in Asia and Latin America, Arab higher education systems are way behind in terms of quality and research outcomes. They suffer from relatively high levels of graduate unemployment, and are characterized by inadequate governance arrangements (Salmi, 2015).

Furthermore, Arab universities are in low position in the world university rankings. In the top 500 places of the Shanghai ranking, since its launch in 2003, only 5 of the 550 Arab universities appear, compared to 7 of the 9 Israeli universities (El Amine, 2016). Higher education sector in the MENA region (and particularly in the gulf) is thriving, with billions of dollars spent on impressive-looking infrastructure over the past three to four decades, the rapid growth in the number of young people (including women) enrolling at universities and tertiary colleges, and the recruitment of thousands of expatriate academics to work in regional

universities. However, this is a misleading impression and even with the huge sums of money spent on building infrastructure, it is evident that there are very few excellent universities in the MENA region.

According to Forster (2018), the reasons leading to this situation are diverse. First, the lack of investment in research and development at universities across the MENA region. Countries in the MENA region collectively spend a comparatively low amount of capital on R&D, and on both basic and applied scientific research in universities. In 2013, the 57 member countries of the Organization of the Islamic Conference invested just 0.81% of their annual gross domestic product (GDP) on R&D. The Muslim world as a whole spends less than half of 1% of their cumulative GDP on R&D, compared to an average of about 5% in the countries affiliated to the Organization for Economic Development (OECD). Second, the lasting absence of intellectual freedom and the constraints and limitations imposed on free expression of ideas and academic freedom in universities in the MENA region. While a lack of investment in high-quality research at MENA universities is one plausible reason for their low ranking, it does not provide a complete explanation for this. For example, several scholars have asked if states of the Gulf and broader MENA region will ever encourage the academic standards and intellectual freedom that are necessary to generate the “world-class research” they often claim to support in their public speeches and in their national economic development plans. Third, the failure of universities in the MENA region to encourage, monitor and reward high-quality and high-impact research across all academic disciplines. While the volume of research generated in the MENA region has increased over the last 15 years, there are no indications that research quality and impact have also been improved. For example, the overall impact rating of scientific research publications from all universities in the MENA region was 0.33 in 2016. This was well below a global average of 1.00, and significantly lower than the impact level of publications from the world’s top 500 universities (Forster, 2018).

The anti-scientific mind-sets of some ruling political elites, of Islamic theologians and of those who are responsible for the oversight and administration of universities in the MENA region provides some clarification of the low productivity of higher education. Many scholars have suggested that Islam, and its pervasive influence in primary and secondary education throughout the entire MENA region, is one of the principal reasons why this part of the world is still characterized by so many endemic and deep-seated problems—including what appears to be an almost instinctively anti-scientific worldview. Islam’s most influential theologians all promote the belief that spiritual and secular authority are inseparable, and the latter must—in all circumstances and at all times—be subservient to the former (Forster, 2018). However, we have observed a process of innovation and critical academic thinking emerging in universities in countries such as the United Arab Emirates, Lebanon, Jordan or Tunisia.

In addition, MENA universities are disadvantaged in respect of academic publishing. Such publishing is typically in the form of articles in scholarly journals, in book form, or as theses, usually with peer-review and/or editorial oversight. Few universities worldwide have their own publishing houses. There are no Arab journals in the sciences and humanities that are universally regarded as being high impact academically at a time when the global literature is becoming dominated by the use of English, western journals, and citation systems developed in the USA and Europe. Financially weak and poorly supported, Arab academic societies are unable to redress the disadvantage (Hillman and Baydoun, 2018).

Universities, especially the newer ones, are running out of steam in their attempt to identify with the dominant models in hegemonic countries from their initial position as dependent universities, which relegates them to a peripheral role. In this context, such creeping standardization risks being reflected not only in a race to the bottom in terms of university education and training, but also in a shrinking of the university public space and, in its wake, of the education as a public good, with the continued regression of the place and role of the public university (Kabbaj, 2012).

Rethinking university in MENA uncertain times

Higher education in the MENA region is at a crossroads and it is urgently needed to rethink it, particularly in terms of governance and the university’s mission in national development.

Some observers call for a revision of models of higher education governance in MENA region. El Amine (2016) analyzed governance models in Arab public and private universities. In public universities, he distinguishes 4 models: (1). The electoral model (Tunisia), (2). The bureaucratic model (Egypt), (3). The politicized model (Lebanon) and (4). The controlled model (Jordan). The governance models in private universities are: (1) The business model, (2) The community model and (3) The autonomous model. In comparison with international models of university governance, Arab public universities are very far from the academic type that had been historically established in Europe. They are also far removed from the market-oriented (neoliberal) type prevalent in the United States or in private universities around the world. Rather, they can be classified in the state-centered type but with some differences between countries. We can say that their main points of divergence from the European state-centered type are their submission to government and the involvement of political bodies from top to bottom, or both ways. The inter-Arab divergences refer to social structures, the relationship with the State, and the union movement, etc. As for the non-public universities, they are polarized either toward the market or by a community organization, or both at the same time. The product of this bipolarity between politics and the market is the loss of autonomy, the weakness of academic freedom and the deterioration of overall quality (El Amine, 2016).

There are a number of important differences in states’ reform strategies in MENA region. Bucknera (2011) classifies states’ reform strategies into three major groups: Neoliberal, Quality Assurance, and Imported Internationalization. Neoliberal reforms aim to expand access to higher education while offsetting costs to consumers and the private sector. Egypt and Tunisia both initiated neoliberal reforms over a decade ago. However, they have not been as successful as Jordan. Today the percentage of youth in higher education in both countries is approximately 30%, but the percentage in private universities is only less than 5% of total enrollment

in higher education. In contrast to the neoliberal model, North African states are emphasizing the importance of the state in providing higher education while pursuing strategies to strengthen the internal and external efficiency of tertiary education. In addition, in all three Maghreb countries of Tunisia, Algeria, and Morocco, the Ministries of Higher Education have recently pursued large-scale quality assurance programs inspired by the Bologna process by restructuring degree requirements to accord to a Bachelors- Masters-PhD system. This approach attempts to align the higher education curricula with European models and permit the mobility of qualifications and labor cross-nationally. The third model of higher education policies pursued in the MENA Region is that of Americanization, whereby the desire to “modernize” higher education systems has not only taken the form of extensively privatizing the provision of higher education, like many neoliberal states, but also establishing extensive international partnerships with American and British universities. A number of scholars have examined how these institutions are changing tertiary education in the UAE, Qatar, Kuwait and Saudi by linking it to global trends.

Zewail (2011) suggested a fourfold strategy to reform higher education in MENA region: (a) Achieve the full emancipation of Arabic women and ensure their full participation in civic, economic and political life; (b) Make reforms to national constitutions to allow freedom of thought, decreasing of bureaucracy, development of merit-based systems, and the creation of credible and enforceable legal codes (c) Build centers of excellence in science and technology in each Muslim country to show it can be done, to show that Muslims can indeed compete in today’s globalized economy, and to instill in their youth the desire for learning and (d) Develop more partnerships between leading Western universities and those in the region in order to improve their research capabilities.

The other dimension to be clarified is the mission of university in the region and its relationships with socioeconomic development. The Arab world represents 5.8% percent of the world population and produces 4.5% percent of the planet’s GDP, but its universities account for only 0.08% percent of the top 500 institutions in the Shanghai ranking. The poor results of a large country like Egypt—the 15th most populated nation of the world—are striking in contrast to the impressive performance of a small country such as the Netherlands, which places four universities in the top 100 of the Shanghai ranking. The tiny territory of Hong Kong has more universities in the Shanghai ranking as all Arab countries considered together (Salmi, 2015).

While participating in economic globalization and the internationalization of higher education, the countries of the MENA region have weak capacities to orient these processes in the direction of their own interests and national development. As stated by Ishengoma (2003), many countries in the global South are in danger of being largely excluded from the whole process of globalization, this exclusion will also apply to higher education.

Missed links between education and development

It is important not to limit the analysis of university situation in the Arab countries to the education sector alone. Indeed, education is part of a socio-political and economic ecosystem favorable or not to the production of knowledge. The apparent paradox is that the Arab countries have not reaped the fruits expected from massive investments in education. Obviously, “schooling” and “socio-economic development” have no mechanical ties stipulating that investing in schooling would automatically allow economic to leverage. Harber (2014) suggests rightly for the “need to be cautious about the assumed automatic benefits of formal education for development” (p. 18).

In addition, the fragile legitimacy of political regimes in the Arab countries does not allow higher education to contribute to better participation of young people in civil society and political life. Education systems produce unemployed graduates whose voices are rarely heard in the political arena. While the Arab Spring movements made possible more youth participation in some Arab countries, a great disillusionment of youth persists everywhere due to the absence of economic prosperity.

Several researchers have identified low rates of return to education in the Arab world as a fundamental cause of the current socio-political unrest and upheaval (Fargues, 2011; Shafiq and Vignoles, 2015). Using years of schooling, Patrinos and Montenegro (2014) estimated that the rate of return to schooling in the MENA region is 5.6%, which is considerably behind the rest of the world. Using empirical data, Kingsbury (2018) tested several hypotheses to explain the low return to schooling in Arab countries. The first hypothesis stipulated that religious orthodoxy overwhelm the return on investment of education, including mathematics, science and foreign languages. The second hypothesis developed by Kingsbury (2018) suggested that poorly performing schools in the region had a negative impact on quality. The third hypothesis indicated that the region’s dependence on natural resource exports limits returns to education. The fourth hypothesis, more difficult to corroborate with empirical data, suggested that corruption, nepotism and non-meritocratic government policy (coupled with the absence of the implementation of law) reduce returns to education. However, it is necessary to discuss certain hypotheses of Kingsbury (2018). The over-representation of religious content is not present in all the countries of the region. Additionally, education systems in the region are very open to foreign languages, especially in the Maghreb and Middle East countries.

In a recent empirical study, Blackeburg and Tholen (2018) pointed out that the share of young people suffering from social exclusion is the highest in Tunisia, followed by Algeria, Egypt, and Lebanon. Youth exclusion does not affect only the young generation but society as a whole. Social exclusion as a function of the respondent’s employment status varies significantly across countries. In contrast to Lebanon, where unemployed young people feel especially excluded, young people in Algeria, Egypt, and Tunisia are more likely to be assigned the status socially excluded if they are employed. A finding pointing toward the dominance of the informal sector.

During the 2000s, Devarajan and Ianchovichina (2018) suggested that inequality in Arab countries was low or moderate and, in many cases, on its way of declining. Different measures of wealth inequality were also lower than elsewhere in the world. Yet, there were revolutions in four countries and protests in several others. Devarajan and Ianchovichina (2018) explained this so-called

“inequality puzzle” by first noting that, despite favorable income inequality measures, subjective well-being measures in Arab countries were relatively low and falling sharply. This fall was especially true for the middle class and in the countries where the uprisings were most intense. The increasing of disappointment, reflected in perceptions of deteriorating standards of living, was associated with dissatisfaction with the quality of public services, the shortage of formal-sector jobs, and widespread corruption. The transition to the labor Market is characterized by unemployment, long waits and informality (Dhillon et al., 2009).

These sources of dissatisfaction suggest that the old social contract, where the government provided jobs, free education and healthcare, and subsidized food and fuel, in return for the subdued voice of the population, was broken. Thus, it became clear that this social contract was not sustainable. In particular, the fiscal deficits associated with public-sector employment and high subsidies were becoming unsustainable. Governments slowed down, and in some cases, stopped hiring workers. Unfortunately, the private sector did not create jobs fast enough to absorb the large number of young people entering the labor force. MENA has the highest unemployment rate in the developing world, with the rate for young people and women about double the world average (International Labor Office, 2020). This problem is worsening due to the emergence of COVID-19 pandemic; that forces governments of all countries to close their operational sectors, as a result, this affects the economy of most countries, and Arab countries are not an exception, many lost their jobs in addition to those who have not got any job yet. A dislocation exists between policy discourse, which emphasizes the centrality of education for work in the process of economic development, and the inability of Arab State to penetrate society effectively (Mazawi, 2007).

The development of higher education should in theory contribute to the reduction of social inequalities, improve access to employment and make the participation of citizens more tangible in political life. The paradox is that the results of the MENA region are mixed on these three levels despite advances in schooling. The progress in education in the Arab region over the past decades should be seen in the light of patent inequalities in the level and quality of education received. These inequalities are illustrated through a differing schooling experiences depending on the social background of the students: Children from privileged socioeconomic backgrounds versus children from disadvantaged socioeconomic backgrounds. Most often, the first attend private school and benefit from quality foreign language teaching—Children living in urban (coastal) areas versus children living in rural (interior) areas. Most often, the latter endures poor conditions of schooling, including insufficient and untrained teachers. Schools are overcrowded and under-equipped, and working in two-shifts, especially after the Syrian crisis, many countries received flux of refugees, like Jordan, Egypt, and Lebanon, the governments accepted students at the age of primary stage to be enrolled at public schools UNRWA schools. Since the capacities of these schools are already limited, having extra number of refugees added more pressure on the educational systems, and thereby affecting the quality of education.

In an ethnographic study, Boutieri (2016), analyzes the arrangement of the different education streams in secondary education in Morocco and dissects the inequalities based in particular on language skills. The most popular options and those most likely to lead to employment after schooling are the scientific and technical streams. In this context, the humanities streams are abandoned to the poorest students from disadvantaged social backgrounds. The bifurcation between school subdivisions takes place at the end of the third year of secondary school. Boutieri (2016) emphasizes that to gain a place in the scientific fields, one must not only have obtained good grades in science and mathematics, but also in French (with a view to continuing studies at university level). Thus, even if chemistry, physics or mathematics courses are given in Arabic at the secondary level, enrollment in a scientific field and subsequent success in these studies paradoxically depend, largely, on mastery of a foreign language, French or English. Arab education systems still have the footprints of colonial encounter in that they remain elitist, lack pertinence to local contexts and are often disconnected from local knowledge and wisdom. Arab society and education systems remain engaged with a Western-driven project of modernity and a consumerist dependence on foreign markets (Mazawi, 2006).

Based on that, social inequality has emerged and it is partially associated with the limited and low-quality skills that many young people acquire in the Arab world because of their low-quality schooling, as they mainly focus on rote learning rather than high-order thinking skills. Likewise, the authorities of the Arab region could not afford the youth sufficient number of jobs and opportunities to take their places as responsible and productive members in their societies. In this vein, it is undeniable that the current social, economic and educational systems could not help the youth to make their dreams come true and become productive members in their communities. On the one hand, there is the story of poor economic performance, characterized by volatility between high and low growth rates, mainly linked to the instability of oil prices. Although, there is a steady growth of education in terms of quantitative measures, learners show inability of knowledge production or acquiring 21st century skills. It is important to underline the absence of strong relations between economic development and education in Arab countries and the need to rethink of both directions simultaneously. As pointed out by the International Labor Office (2020), youth unemployment is the highest in Northern Africa and in the Arab States, at around 2.2 and 1.7 times the global rate, respectively. In these two sub-regions, youth unemployment rates have been considerably higher than those in the rest of the world since at least 1991, suggesting that there are structural barriers stopping young people from engaging in the labor market.

Finally, we need to remember that the education of the population has two major objectives. The first one is instrumental, which consists of having a more productive and educated active generation. The second is related to the exercise of political rights and citizenry participation. However, younger generations in the Arab region do not have access to the status of an educated and socio-economic workforce, nor to the status of citizens enjoying political freedom and dignity. Hence, revolts broke out in certain Arab countries demanding that these two objectives should be achieved. Malik and Awadallah (2013) point out that young people are not only unemployed, but also unemployable. This problem persists despite the fact that public spending on education in the Arab world is higher than or comparable to that in other regions of the world.

It is unlikely to address any aspect of education and educational policy in the Arab region without due attention to the fact that the crisis of education is also in many ways a crisis of the development and the state, of which the utter instability—political and economic—are the major two manifestations. There is hardly any Arab society across the region in which the state is not—in many ways—on the defensive, fragile if not already collapsed (Mazawi, 1999).

References

- Backeberg, L., Tholen, J., 2018. The frustrated generation youth exclusion in Arab Mediterranean societies. *J. Youth Stud.* 21 (4), 513–532.
- Boutier, C., 2016. *Public Cultures of the Middle East and North Africa*. Indiana University Press, Bloomington.
- Buckner, E., 2011. The role of higher education in the Arab State and society. *J. Comp. Int. Higher Educ.* 3, 21–26.
- Buckner, E., 2018. The growth of private higher education in North Africa: a comparative analysis of Morocco and Tunisia. *Stud. High Educ.* 43 (7), 1295–1306.
- Calderon, A., 2018. *Massification of Higher Education Revisited*. RMIT University, Melbourne.
- Devarajan, S., Ianchovichina, E., 2018. A broken social contract, not high inequality, led to the Arab spring. *Rev. Income Wealth* 64, S5–S25.
- Dhillon, N., Salehi-Isfahani, D., Dyer, P., Yousef, T., Fahmy, A., Kraetsch, M., 2009. *Missed by the Boom, Hurt by the Bust: Making Markets Work for Young People in the Middle East*. MEYI, Washington, DC.
- El Amine, A., 2016. Les modèles de gouvernance des universités arabes. *Lebanese Nat. Defense Rev.* 96, 35–36.
- El-Kogali, S., 2018. *Expectations and Aspirations: A New Framework for Education in the Middle East and North Africa: Overview* (English). World Bank Group, Washington, D.C. <http://documents.worldbank.org/curated/en/527931542039352771/Overview>.
- Fargues, P., 2011. International migration and the demographic transition: a two-way interaction. *Int. Migrat. Rev.* 45 (3), 588–614.
- Figge, L., Martens, P., 2014. Globalization continues: the Maastricht globalization index revisited and updated. *Globalizations* 11 (6), 875–893. <https://doi.org/10.1080/14747731.2014.887389>.
- Forster, N., 2018. Why are there so few world-class universities in the Middle East and North Africa? *J. Furth. High. Educ.* 42 (8), 1025–1039.
- Fouchet, R., Moustier, E., 2010. *L'enseignement supérieur et la recherche en Méditerranée: enjeux et défis de demain en sciences sociales*. EUROMED, 10 Papers for Barcelona 2010, Paper 4.
- Freely, J., 2010. *Aladdin's Lamp: How Greek Science Came to Europe Through the Islamic World*. Vintage.
- Harber, C., 2014. *Education and International Development: Theory, Practice and Issues*. Symposium Books Ltd.
- Hillman, J.R., Baydoun, E., 2018. The future of universities in the Arab region: a review. In: Badran, A., Baydoun, E., Hillman, J. (Eds.), *Universities in Arab Countries: An Urgent Need for Change*. Springer, Cham. https://doi.org/10.1007/978-3-319-73111-7_1.
- International Labour Office, 2020. *Global Employment Trends for Youth 2020: Technology and the Future of Job*. ILO, Geneva.
- Ishengoma, J.M., 2003. The myths and realities of higher education globalization: a view from the southern hemisphere. In: *Focus Journal, Open Forum*. Retrieved Month, day, year, from: <http://www.escotet.org/infocus/forum/ishengoma.htm>.
- Kabbajji, J., 2012. Heurs et malheurs du système universitaire libanais à l'heure de l'homogénéisation et de la marchandisation de l'enseignement supérieur. *Rev. Mondes Musulmans Méditerranée* (131), 127–145.
- Kingsbury, I., 2018. Making sense of low private returns in MENA: a human capital approach. *Int. J. Educ. Dev.* 61, 173–183.
- Knight, E., Jones, A., Gertler, M.S., 2021. The public university and the retreat from globalization: an economic geography perspective on managing local-global tensions in international higher education. *Environ. Plann.* 53 (1), 210–218.
- Kraft, C., Alawode, H., 2018. Inequality of opportunity in higher education in the Middle East and North Africa. *Int. J. Educ. Dev.* 62, 234–244.
- Lai, Y.M., Ahmad, A.R., Da Wan, C. (Eds.), 2016. *Higher Education in the Middle East and North Africa: Exploring Regional and Country Specific Potentials*. Springer.
- Malik, A., Awadallah, B., 2013. The economics of the Arab spring. *World Dev.* 45, 296–313.
- Martins, C.B., 2019. L'enseignement supérieur à l'heure de la mondialisation. *Socio. New Soc. Sci. J.* (12), 205–227.
- Mazawi, A.E., 1999. The contested terrains of education in the Arab states: an appraisal of major research trends. *Comp. Educ. Rev.* 43 (3), 332–352.
- Mazawi, A., 2005. Contrasting perspectives on higher education in the Arab states. In: Smart, J. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XX. Springer, London, pp. 133–189.
- Mazawi, A.E., 2006. Globalization, development, and policies of knowledge and learning in the Arab states. In: *Concepts of Knowledge and Learning: The Learning Society in Europe and Beyond*. Peter Lang, New York, pp. 335–383.
- Mazawi, A.E., 2007. "Knowledge society" or work as "spectacle"? Education for work and the prospects of social transformation in Arab societies. In: *World Yearbook of Education 2007*. Routledge, pp. 269–285.
- Mullin, C., 2017. L'enseignement supérieur en Tunisie, lieu de pouvoir (néo) colonial et de lutte décoloniale. *Tumultes* 1, 185–205.
- Ozga, J., Lingard, B., 2007. Globalization, education policy and politics. In: *The Routledge Falmer Reader in Education Policy and Politics*. Routledge, pp. 75–92.
- Patrinos, H., Montenegro, C.E., 2014. *Comparable Estimates of Returns to Schooling Around the World*. World Bank Policy Research. Working Paper, 7020.
- Powell, J.J., 2014. University roots and branches between "glocalization" and "mondialisation": Qatar's (inter) national universities. In: *Education for a Knowledge Society in Arabian Gulf Countries*. Emerald Group Publishing Limited.
- Robertson, S., Novelli, M., Dale, R., Tikly, L., Dachi, H., Alphonse, N., 2007. *Globalization, Education and Development: Ideas, Actors and Dynamics*. DfID, London.
- Romani, V., 2012. Internationalisation des politiques universitaires et contournement de leurs publics? *Rev. Mondes Musulmans Méditerranée* (131), 13–21.
- Salmi, J., 2015. What Future for Higher Ed in the Arab World? <https://www.wise-qatar.org/what-future-higher-ed-arab-world-jamil-salmi/>.
- Samoff, J., Carrol, B., 2003. *From Manpower Planning to the Knowledge Era: World Bank Policies on Higher Education in Africa*. UNESCO Forum on Higher Education, Research and Knowledge.
- Shafiq, M.N., Vignoles, A., 2015. *Beyond the Arab Spring: Education, Earnings, and Protest Participation*. Paper Prepared for the AALIMS—Princeton Conference on Islam and Human Capital, October 2015.
- Shahjahan, R.A., 2012. The roles of International Organizations (IOs) in globalizing higher education policy. In: Smart, J., Paulsen, M. (Eds.), *Higher Education: Handbook of Theory and Research*. Springer, Dordrecht, pp. 369–407. https://doi.org/10.1007/978-94-007-2950-6_8.
- Siino, F., 2003. Sciences, savoirs modernes et pouvoirs dans le monde musulman contemporain. *Rev. Mondes Musulmans Méditerranée* 101–102, 9–264.
- Strassel, C., 2018. Les enjeux géopolitiques de la mondialisation universitaire. *Herodote* (1), 9–38.
- World Bank, 2008. *The Road Not Traveled. Education Reform in the Middle East and North Africa*. The World Bank, Washington, DC.
- Zajda, J., 2020. *Globalization, Ideology and Neo-liberal Higher Education Reforms*. Springer, New York.
- Zewail, A., 2011. A scientific revolution: the Arab spring puts the Middle East in a position to become a scientific powerhouse, but it needs help from the west. *New Sci.* 8.

Challenges facing education in Sub-Saharan Africa

Isaac Quist, Wesley College, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	388
Colonialism and its ambivalent legacies	389
The role of education in national development	390
Contrasting goals of the World Bank and UNESCO	391
Emergence of private corporate and philanthropic actors in education	392
Current challenges	394
Conclusion	396
References	396

Introduction

As we head into the second quarter of the 21st century, Sub-Saharan Africa continues to encounter a myriad of issues relating to its development and the future of its peoples. Of enduring significance are matters relating to the provision of formal education in the region, particularly those concerning access, inclusion, relevance and quality. These issues are of course not new. A close look at the history of the development and consolidation of formal education as an endeavor of significance should reveal that its provision in Sub-Saharan Africa has always been associated with questions about who provides, who participates, at what cost, for whose benefit, of what relevance and of what quality. Fundamental, as they are, these questions trigger further questions about the modes of thinking that have informed the strategic choices made and approaches adopted to the provision of formal education in Sub-Saharan Africa from its early beginnings in the Colonial era. It is of interest that much the same questions are being asked now that Sub-Saharan Africa is comprised of modern independent nation states as when it was a collection of colonial territories owned and overseen by the major European powers.

Tracing the journey over time, this paper aims to identify and review the key actors of educational activity in Sub-Saharan Africa. It also seeks to understand their overriding objectives, primary considerations and overall aspirations, and the role(s) these play in how the enactment of formal education in the region has been both approached and prosecuted.

The paper assumes however that though the questions asked about educational provision in Sub-Saharan Africa may not have changed, the national development needs of the politically independent countries that now make up the region should constitute the guiding intelligence that informs the provision, enactment, and ongoing reform of education in the region. This paper will assess the soundness of this assumption. In addition, it will seek to reflect the objectives, considerations and aspirations of educational endeavor and its reforms over time through a lens that sees education as critical to addressing the workforce, citizenship and leadership needs of the relatively young independent countries that make up the region.

Given what appears to be the proverbial chicken and egg relationship that exists between the quality of education (measured by literacy rates, student achievements etc.) and the level of a country's development, it is important to acknowledge at the outset that in Sub-Saharan Africa we are dealing with countries that are routinely classified as low and lower-middle income economies and, therefore, developing if not under-developed. For a provision that is generally acknowledged as a fundamental human right, education is not cheap. This fact constitutes an ongoing challenge for countries in the region. The experts associated with a recent Global Education Monitoring Report point out the following:

Children are more at risk, especially in sub-Saharan Africa, where extreme poverty affects 49% of children, accounting for 52% of extremely poor children globally Education is an opportunity with the potential to transform lives. Yet an estimated 258 million children, adolescents and youth, or 17% of the global total, are not in school. The number out of school in sub-Saharan Africa has passed that of Central and Southern Asia and is growing. The share of sub-Saharan Africa in the global total increased from 24% in 2000 to 38% in 2018.

Global Education Monitoring Report Summary (2020, pp. 6–7).

Against this sobering context, it is important to note moreover the reality that: "Large differences in education spending as a share of GDP between low-income and high-income countries are not due to differences in the priority accorded to education in government budgets but due to differences in the share of overall government spending in GDP" (Al-Samarrai et al., 2021, p. 7). Essentially, it seems, actual spending on education in Sub-Saharan Africa is dictated more by the size of GDP than by the value placed on it as a developmental need. With a capacity to provide so compromised, it is important to ask what the real forces are that drive educational activity in Sub-Saharan Africa, the dominant motivations at play, and the prospects and challenges that arise for countries in the region as a result. This paper, aims to engage with and, hopefully, shed some light on these questions.

Still, and before proceeding any further, there is no intention, as might appear to be the case so far, to characterize Sub-Saharan Africa as a homogenous undifferentiated space. Indeed, to highlight its diversity and the many ways by which often undifferentiated educational approaches and policies manifest at local, community and national levels, the paper will routinely illustrate its claims by pressing into service examples from individual countries as relevant, including, Ghana, Nigeria, Tanzania, South Africa and Rwanda. It will moreover focus predominantly on the K-12 sector in Africa. Still, the issues highlighted may well also be applicable, to various extents, to the story of higher education in Sub-Saharan Africa.

Colonialism and its ambivalent legacies

A common historical experience shared by the countries of the vast geographical territory, now labeled Sub-Saharan Africa, is that of colonialism. With the exception of Ethiopia, much of Sub-Saharan Africa was colonized by the powers of 19th Century Europe, notably Belgium, France, Germany, Great Britain and Portugal. Still, it is to the Christian churches and their missionaries that introduction of European style education in the region can be credited. As Becker (2021: n.p.) notes, “Mission societies often expanded into territories before colonial powers did”; further, they “continued to be the main providers of education even after colonial powers established control during the so-called Scramble for Africa which occurred between 1884 and 1914”. The schools set up by the missionaries facilitated their work by training interpreters who supported their evangelism and helped gain new converts through various networks of schools. The education provided through these mission schools also “spread Eurocentric norms” many of which, for good or ill, remain visible even today in the lived culture of many Sub-Saharan countries. This however is only the story of *European style* education in Sub-Saharan Africa.

Often forgotten, or neglected, is the fact that as far back as the 11th century the Arab-Berber merchants had also brought *Arab-Islamic education* to West Africa. It was “subsequently spread by [Islamic] religious brotherhoods in the 19th century” (d’Aiglepiere et al., 2018: n.p.). The European colonial authorities saw the Koranic schools of the Islamic tradition as a force to contend with and created their own schools to train competent civil servants’ (d’Aiglepiere et al., 2018), familiar with western norms and Christian ethos. The Islamic madrasas however have continued to exist throughout the Sub-Saharan region. It is important to recognize this alternative source of formal education in Sub-Saharan Africa and its present-day manifestations if we are to have a fuller understanding of how the contemporary educational landscape in Sub-Saharan Africa is constituted. As well, an acknowledgment of Arab-Islamic education in the region has implications for how we understand the complexities of access and inclusion in the region. This is a matter of relevance to the aspirations of the *education for all* agenda. d’Aiglepiere et al. rightly note that:

Arab-Islamic education in general, and Koranic schools in particular, are largely excluded from programs advocating education for all in Africa. Yet this education concerns a large number of children, many of whom are considered as “out of school” by the most public authorities in Africa, as well as the various international development agencies. According to d’Aiglepiere et al. (2018), however, recognizing their existence and importance, as well as their diversity, is a prerequisite for building a dialog between all stakeholders about the challenges of education reform in the Sub-Saharan Africa.

Additionally, this corrective is useful as it takes us beyond some of the more common clichés that, typically, diminish the presence and role of Arab-Islamic education in the region, including rather popular misperceptions about it as: a recent phenomenon in Sub-Saharan Africa; a negligible phenomenon in Sub-Saharan Africa; being only for boys and for the poor; and concerned only with memorizing the Koran. While Islam, Christianity and their religious pursuits can be said to account for the advent of formal¹ education in Sub-Saharan Africa, there is no denying the role played by colonial governments following the Scramble for Africa. Much has been made of the role of education in colonialism’s “civilizing mission” in Sub-Saharan Africa and of the duty that the “superior races” have to “civilize the less fortunate, ‘inferior races’” (Serequeberhan cited in Mart, 2011, p. 1). Neither time nor space permit detailed elaboration of the whys and wherefores of these claims. Still, the question must be asked as to whose interests were served by the advent and consolidation of formal education in pre-colonial and colonial Sub-Saharan Africa.

Boiled to its essence, the historical evidence points to the complicity of religious missionaries in Europe’s project of colonization in Sub-Saharan Africa. The introduction of Eurocentric norms and values clearly aided the colonial project by nurturing an inadequate sense of self among the African populations who came to believe in their own inferiority to their European counterparts (Fanon, 1967). Furthermore, the colonial governments used education as a tool to train Africans for colonialism’s own purposes, namely, administration of colonial territories in the interests of the colonizers, as was clear with Lord Lugard’s ‘policy of indirect rule’ (Scott, 2013, p. 2). Such policies promoted the “modernization for colonized people” but served the purpose of creating a pool of skilled but cheap labor. It led to cultural alienations that entrenched notions of racial inferiority among Africans while also disrupting existing social structures and establishing new hierarchies in African societies.

Across present day Sub-Saharan Africa, the ambivalent legacies of the colonial purposes and impact of formal education remain very much in evidence—these often manifest in issues that challenge the internal cohesion of countries in the region as modern nation states. The most readily visible legacy, in this context, are languages of colonial domination now elevated to the status of

¹Some commentators present Arabic-Islam education as informal. This would seem to be because it is often considered to be purely religious in focus—a contestable perception.

official languages of Sub-Saharan Africa. English, French, Portuguese continue to function as languages of instruction in the educational systems of the region and play a critical role in the social mobility of its citizens (Okolo, 2005). It is widely argued (Wolff, 2018) that these languages, owing to their prime position in administration of the modern independent nations, are best placed to provide a shared language of communication across different clans, tribes and kingdoms of each country and internationally. Such an approach to the language of instruction is viewed as essential to nation building in a context of different languages, dialects and cultures each vying for prominence. The sober reality is that in the elevation of Europe's languages, the African languages are effectively silenced and rendered inferior. Consequently, those without access to English, French, Portuguese are also excluded from opportunity and social mobility in the context of the modern state.

The role of education in national development

Assuming it to be a powerful force for economic and social development, it is not surprising that all nations—rich and poor alike—commit to investing significant proportions of their revenues to the education of their people. This is consistent with the other widely held view that any nation's most important resource is its people, a notion made even more popular by Barack Obama in his Yangon University speech of 2012. While Sub-Saharan Africa is consistently described as a region rich in natural resources such as oil, coltan, lithium and a wide variety of other valuable minerals both renewable and non-renewable, it is equally maintained that harnessing of these resources for national development requires a knowledgeable and skilled workforce. There is no overstating the important role education plays in the development of such a workforce. This is even more so, given the relatively recent arrival of Sub-Saharan African countries to the status of independent nationhood with the presumed capacity to determine their own directions for development and to manage their own affairs.

As already noted, however, formal education in colonial Sub-Saharan Africa seldom served the interests of the colonized. More than anything the training and orientations it provided to local populations simply facilitated the carting off of the region's significant natural wealth to serve the material enrichment of Europe. That this exploitative arrangement worked to the detriment of the region's own development is evident in the generally inadequate, if not altogether non-existent, infrastructure for socio-economic development in the several countries of the region at their respective attainments of political independence. This includes the poor levels of human resource development. On independence, it might be noted that countries such as Ghana, Tanzania and Kenya inherited very low literacy rates, together with a low number of trained doctors, engineers, lawyers, teachers, nurses, tradesmen. Not surprisingly therefore, commentators such as Austin (2010) have noted that political independence for these countries did not come with economic independence. Significantly, with hardly any exceptions, the erstwhile colonial powers, especially the United Kingdom and France, maintained their controlling involvement in the management of the region's natural wealth. In a more nuanced analysis than this cursory survey can give justice to, Austin (2010) acknowledges severally that "the French government remained closely involved with its former colonies after their independence, not least through the franc zone." Citing Uche (2008), Austin further acknowledges that "... the British government's attitude to the Biafran secession was influenced by the interests of British oil companies." As a result, these resource rich countries were reduced to mere raw material producers with little influence in determining the value of their own wealth. In more recent decades, this situation has been made worse with the emergence of a globalized world economy that allows economic power to be retained by capital and its owners. This is a context in which enables China, now a major economic player, to extract Africa's raw materials to feed its own industrial manufacturing interests with little change to the region's status as a generally powerless raw material source.

The upshot is that despite all their natural wealth, the preponderance of Sub-Saharan African countries remains poor and unable to afford the cost of educating their own peoples to provide the workforce, citizenship and leadership skills that will advance their own national development interests. This vicious cycle of poverty diminishes the political agency of the Sub-Saharan countries, leaving their development aspirations at the mercy of external funding or donor agencies whose interests are not necessarily convergent, if not altogether at odds, with theirs. It is a poverty that generates dissatisfactions and disillusionments with the promise of independence. It is not unheard of for sections of the disappointed populations in these countries to express a desire to return to the relative 'comforts' of colonial rule (Moro, 1989 cited in Ayittey, 2006). The unfulfilled aspiration among the various sectors of the populace results, ultimately, in the political instability of coups d'états for which Sub-Saharan Africa has something of a reputation.

The revolving door of political leadership further undercuts the prospects of economic development for countries in the region as they are taken hostage by successive groups of military adventurists and their ilk, mostly ill-suited to national leadership and without the skills to match their ambition. Examples from the DRC, the Gambia and Burkina Faso indicate extreme versions of this phenomenon (Suleiman and Onapajo, 2022). Their inevitable setting back of the economic development aspirations, contributes in no small way to exacerbating the challenges of providing the education necessary for the economic uplift of the nation and the upward mobility of its citizenry.

Clearly, for the several countries of Sub-Saharan Africa, the historically long arm of colonialism continues to cast a grave impoverishing shadow which weakens their very viability as independent nations that possess the agency to determine and manage their own affairs. Sub-Saharan African nations have the most need to educate their own peoples for the purposes of harnessing their considerable natural resources for national development. And yet, they find themselves, unwittingly and helplessly, assisting the economic expansion of the very countries whose wealth constitute the resource base of the international finance institutions to which they must go begging for aid, such as the International Monetary Fund and the World Bank. A lamentable irony.

The aid sought and received from international organizations (IOs) routinely comes with ideological strings that shape the approach to education reform in directions that are not always of the choosing nor necessarily in the interests of the Sub-Saharan African countries themselves. The donor organizations and countries have their own interests and perspectives which define their understanding of the educational reforms needed in the countries of the Sub-Saharan Africa. Invariably IOs seek to marry the locally articulated requirements with their own preferences. This can be illustrated by examining the contrasting agendas of educational reform promoted by UNESCO and the World Bank.

Contrasting goals of the World Bank and UNESCO

Most countries in Sub-Saharan Africa gained independence from colonial rule during the 1960s–70s. Political independence, for these countries, was the first rung on the ladder to nation building and national prosperity. Eager to improve life and living for their respective peoples, the governments of these countries embarked on a broad range of national development initiatives to achieve economic self-sufficiency and affirmation of national identity. Typically, these initiatives were ambitious in scale, span of focus and speed, especially for countries whose populations were relatively small, their institutions in the formative stages of development and their sense of nationhood only incipient. Still those heady days of development saw significant investments in building the infrastructure for national development, including education. The example of Tanzania, where “public spending on education rose during most of the 1970s, reaching a high of about 6% of GDP in 1979” (Gerard et al., 1998) cited in [Johnson \(2002\)](#) is a case in point.

Unfortunately, such expanded investment in nation building through education came to an untimely end in the 1980s. The collective bite of the OPEC oil crises of the late 1970s, followed by the global recession of the 1980s and the financial toll of an ambitious national development drive ushered in a period of declining investments. Of Tanzania, Gerard et al. (1998, cited in [Johnson \(2002:n.p.\)](#), note that, “The share of budgetary contributions to education fell just over 4% in the early 1980s, and then dropped to below 4% by the early 1990s”, along with “other spheres of social endeavors were also dramatically affected under the impact of fiscal retrenchment.”

Across the region, governments sought the support of the World Bank and IMF to revitalize economies suffering a similar fate as Tanzania’s. Their response, the Structural Adjustment Program ([Heidhues and Obare, 2011](#)) of the 1980s, deserves comment. In order to ensure debt repayment and efficient restructuring of economies in distress, the banks imposed “conditionalities” in return for financial assistance. A key characteristic of these “conditionalities” included the requirement for governments to reduce public expenditure on the very activities that would advance economic growth and raise the standards of living for their peoples—these included:

- cuts to government expenditure on essential nation building activity like education, health and development—effectively this requirement placed stifling limits on state participation in activity to grow their economies
- encouragement of privatization in the provision of development infrastructure including, critically, education and health
- removal of protections for domestic industries so as to attract foreign investors
- increased exports of raw materials to support debt repayments and stabilize local currencies

The negative impact of these policies on the economies of Sub-Saharan African countries, and the extreme hardships they created for daily living have been well-documented ([Konadu-Agyemang, 2002](#)). In education, the reductions in government expenditure had dire consequences for provision of educational infrastructure (schools, textbooks and other learning resources), teacher development and the opportunities to expand access—a perfect combination of factors that ultimately served to divest education in Sub-Saharan Africa of any semblance of quality, a situation that continues to endure in many of the countries affected. As well, in countries like Tanzania, the government was compelled to reintroduce school fees (cost sharing)—earlier abolished to stimulate broader participation in the immediate post-independence years—with predictable consequences ([Brock-Utne, 2013](#)).

Private sector participation, intended to make up the shortfall in government provision, generally operated on a commercial basis. The costs involved in participation added to household costs in a time of financial hardship and parents/guardians had to make choices. This led to the exclusion of families without the means; it also tended to exclude girls, orphans and other children of school going age whose participation in education parents felt to be unnecessary or not worth the investment. [Pamba \(2012\)](#) notes that in countries like Burkina Faso, Burundi, Angola and Kenya and elsewhere, the policy of cost sharing had an adverse impact on enrollments with the number of children out of school increasing markedly in the period between 1980 and 1997.

In a world where education is widely held to be a fundamental human right ([Task Force on Higher Education and Society, 2000](#)), one can only wonder at the human sensitivity of policies that overtly, and some might say gratuitously, undermine this right for the poorest of the world. Article 26.1 of the Universal Declaration of Human Rights states without equivocation that “Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages” ([United Nations, 1948](#)). Did the whole world turn a wilful blind eye to these statements while the International Financial Institutions (IFIs) brazenly put a cost on education via their SAPs and, thereby, commodified it without warning, to the exclusion of large swathes of children from Sub-Saharan Africa?

It is important to note that with their insistence on SAPs and attendant conditionalities, the IFIs, including the IMF and the World Bank, arrogated to themselves decision making authority for shaping macro-economic policy in those countries unfortunate enough to request their financial assistance. As [Pamba \(2012\)](#) argues, “It was not long before the World Bank begun playing an

active role in guiding macroeconomic development in recipient and even applicant countries in the direction it considered appropriate” (pp. 19–20). Of significance, the World Development Report (World Bank, 2000) indicated that the World Bank committed some USD1.8 billion, to the education programmes in the world’s poorest countries between 1991 and 1999. This constituted 8.2% of its overall lending for the period, almost double the 4.8% it allocated to education in the 1980s. However, most of the increased lending for education was committed to recurrent expenditure—salaries, teaching and learning supplies and the like—at the expense of development expenditure. In, Kenya, for example, recurrent expenditure accounted for 95% of total education expenditure with development expenditure making up only 5% of same over the 2002/3 to 2004/5 school years (Pamba, 2012). The curtailed investment in development expenditure had negative consequences for access, teacher recruitment and training, and infrastructure developments.

The absence of a concern for social values in the conception and implementation of the neoliberal SAP is perhaps best underscored by the extent of physical violence and death its resistance led to, as students and teachers protested the cuts to educational funding. In country after country these protests, in the form of demonstrations and teacher strikes resulted in actual loss of human life. From Benin, through Niger to Nigeria, Uganda and the Democratic Republic of Congo, cuts to government funding for education in the 1990s led variously to the attenuation of educational programs, withdrawal of study grants for students and teachers alike, salary reductions and redundancies, deterioration of school facilities and learning equipment.

The ensuing protests, especially by students and lecturers in the tertiary sector resulted in expulsions, arrests and incarcerations, violent deaths, long school boycotts and closures. In her research on the subject, Pamba (2012) details the armed assaults unleashed on students of the Ahmadu Bello University (ABU) in Zaria, Nigeria who “were massacred by security forces after staging peaceful protests over impending introduction of SAPs” in 1986. In an equally violent manner, the Ugandan police “fired into a crowd of protesting students” of Makerere University—at least two students were killed on that occasion in December 1990.

Placed against the interventions of the IFIs, UNESCO’s aspirations, motivations, and approaches to education in Sub-Saharan Africa in contrast could not be more different. With a clear goal for *all* children to be participating in school by 2030, UNESCO’s *Education For All* agenda and Sustainable Development Goals explicitly have prioritized equity, access and inclusion. They have sought to grow the capacity of Sub-Saharan African countries to optimize opportunities for equipping their citizens with the competences for effective participation in nation building in a time of global volatility, uncertainty, complexity and ambiguity. Upfront and clear-eyed about the magnitude of the need, UNESCO acknowledges that:

Of all regions, sub-Saharan Africa has the highest rates of education exclusion. Over one-fifth of children between the ages of about 6 and 11 are out of school, followed by one-third of youth between the ages of about 12 and 14. According to UIS data, almost 60% of youth between the ages of about 15 and 17 are not in school. Without urgent action, the situation will likely get worse as the region faces a rising demand for education due to a still-growing school-age population.

UIS Factsheet No. 56 September (2019, p. 7).

In direct contrast to the IFIs, UNESCO seems focused on addressing, not creating or exacerbating, the problems of access traceable, in part at least, to the impact of SAPs in the Sub-Saharan Africa region. Whereas the IFIs are driven by economic imperatives rooted dispassionately in the workings of market forces, UNESCO is unequivocal in declaring that “Education in Africa is major priority for UNESCO and the UIS.” In pursuit of this priority, its Institute of Statistics (UIS) “Develops indicators to help governments, donors and UN partners better address the challenges.” These indicators focus, for example, on school’s access to basics, including access to electricity and potable water, classroom conditions—from the availability of textbooks to average class sizes. They also focus on data relating to teacher training, recruitment and working conditions (UNESCO-UIS, 2019). Of particular interest to UNESCO is Girls’ education—an aspect of education in Sub-Saharan Africa that was severely hampered by the regime of SAPs.

UNESCO’s approach to education in Sub-Saharan Africa is rooted in a commitment to inclusion as a moral imperative for a rapidly changing world that “faces constant major challenges—from technological disruption to climate change, conflict, the forced movement of people, intolerance and hate—which further widen inequalities and exert an impact for decades to come.” (UNESCO, 2020) Unapologetically, UNESCO embraces education as a fundamental human right for all, including, especially, the world’s most vulnerable and disadvantaged in Sub-Saharan Africa. As its Director General, asserts in her foreword to the *Global Education Monitoring Report (2020)* with the telling theme “Inclusion and Education: All means all”, “More than ever, we have a collective responsibility to support the most vulnerable and disadvantaged, helping to reduce long-lasting societal breaches that threaten our shared humanity.”

Emergence of private corporate and philanthropic actors in education

Attempts to reform education in the countries of Sub-Saharan Africa thus take place within this contradictory space, with the international financial institutions, UNESCO and other United Nations agencies with a responsibility for education, such as UNICEF, pulling them in conflicting directions. These international organizations are however not the only players. The region’s very conditions of disadvantage and inadequacy as regards educational infrastructure, facilities, resources, and the qualified personnel, to

support access and high-quality outcomes for all, appear to have made sub-Saharan education an object interest for a variety of non-governmental players.

Noted already has been the roles played in education by the Christian church and Islam in pre-colonial and colonial times. Both have maintained their interest and presence in the post-colonial era, too, and continue to support promote their own distinctive approach to educational reform in present day Sub-Saharan Africa. Schools belonging to or closely associated with the various denominations—Catholic, Methodist, Anglican, Presbyterians, Baptist, T I Ahmadiyya—remain very visible on the sub-Saharan educational landscape.

Alongside the religious groups, philanthropic organizations such as the Aga Khan Development Network (AKDN) and SOS Kinderdorf International have also had a long involvement in bringing education of various countries within the region. The AKDN has a strong presence especially in East Africa (Kenya, Madagascar, Mozambique, Rwanda, Tanzania, Uganda) and West Africa (Burkina Faso, Cote d'Ivoire, Mali, Senegal) where it supports the education of children, the youth and adults (AKDN; n.d.). SOS Kinderdorf International operates in 46 African countries (SOS Children's Villages; n.d.). Like the AKDN, it also supports the education and training of disadvantaged children and youth by providing access to traditional K-12 education as well as vocational training, apprenticeships and work experience.

As philanthropic organizations both of these organizations seek to augment the work of governments in making fee-free education accessible to as many as their own resources and resourcefulness makes possible. They can be said to be partners of governments within the region in this regard as early adopters and implementors of the UN's SDG 4 on quality education (see <https://www.akdn.org/gallery/sustainable-development-goals-partnerships-goals>). Notable recent arrivals on the terrain of philanthropic educational endeavor in Sub-Saharan Africa include the African Leadership Academy (ALA), which currently located in South Africa, Mauritius and Rwanda, is dedicated to the cause of developing the next generation of African leaders through programs that emphasize entrepreneurship. Another active a philanthropic organization is Still I Rise, which seeking to make a high-quality education available to children who live in Mathare, a slum in Nairobi, Kenya. An additional, and miscellaneous, group of actors can also be found in the several so-called "International Schools" often run through partnerships between host countries and others to support education for the children of expatriates, diplomats and relatively affluent local elite. For example, the United World College International now has a school in Swaziland and another in Tanzania.

Over the past two decades, though, some new non-governmental players have emerged from the global corporate world alongside a motley crew of small local actors, who work individually or in partnership with others across the region, to take advantage of the opportunities for entrepreneurship presented by the deficiencies of education within the region. Some of the names and organizations of interest in this regard would be Pearson (and its local franchisees in Sierra Leone and Ghana), the Enko Group which describes itself as "a fast-growing network of African international schools, increasing access to the world's best universities for learners across Africa"—supported by several international financing institutions, including Proparco, I&P, Oiko Credit, Liquid Africa and BIO, Enko Education currently operates in Senegal, Côte d'Ivoire, Cameroon, Mali, Burkina Faso, Mozambique, Botswana, Zambia and South Africa, and aims to open 50 more schools over time.

This group of non-governmental educational players operate a for profit fee-paying model of education across the income spectrum with the global corporate, Pearson, for example, targets the lower end, the "bottom of the pyramid" "among the poorest the fastest growing market" with "collective 'untapped buying power' ... by providing services to the poorest, businesses are not only making profit, but at the same time doing good! Indeed, this is the argument of 'philanthrocapitalism'" (Prahalad, 2004) cited in Lingard (2015).

Working in partnerships, Pearson claims a commitment to supporting the educational needs of the underprivileged in Sub-Saharan Africa. According to Kate James, Chief Corporate Affairs and Global Marketing Officer at Pearson:

In large parts of Africa, upwards of 40% of children are living in poverty and not receiving an education or gaining the crucial skills needed to get a job. Together with Camfed we're determined to play a role in helping improve the life prospects and career opportunities for young people in Africa – with a particular focus on empowering young girls.

Pearson (2017).

Camfed CEO, Lucy Lake says that, "Our partnership with Pearson has opened up opportunities among some of the most marginalized young people around the world ..."

While Pearson's declared intent is laudable, it has drawn significant criticism owing to questions about its actual mode of operation. Pearson's critics complain about education of a questionable quality as the very low level of remunerations it offers means that it uses unqualified Learner Guides, in place of qualified teachers, to deliver pre-packaged mass-produced curriculum. In an era of increasing sophistication in personalized learning and differentiated instruction techniques, such a strategy is incapable of catering to the reality of diversity as regards student learning needs. Commenting on the contradictions of Pearson's edu-business activities globally, Lingard (2015, n.p.) notes:

Pearson proffer strong support for the 'quality teachers' agenda in the nations of the Global North (see, for example, Pearson's series of publications, *Open Ideas* and in particular John Hattie's two 2015 pieces), but their support for low-fee private schools that will generate profit is economically dependent upon the employment of un- and under-qualified and very lowly paid, non-union organised teachers often using scripted pedagogies. There

is a stark moral contradiction here. It becomes apparent that the bottom-line is profit, rather than social good.

In addition to these contradictions relating to teacher quality, pedagogy and remunerations, Lingard (2015, n.p.) also identifies contradictions relating to the “low-fees” charged by Pearson and its partners. In his view, in contexts of predominantly low incomes, school fees, however low, are still a significant expenditure for families and leads to exclusion of the culturally disadvantaged. In this constituency would belong girls, the disabled, orphans, and the like, for whom formal education is often seen by parents and guardians as an investment of limited benefit. Lingard also notes that Pearson’s concept of “low fee” schooling both challenges and undermines “the notion that ‘free’, high quality public education for all is central to democracy and a socially just society and concerningly there is little evidence to support the effectiveness of low-fee, for-profit schools” (Lingard, 2015, n.p.).

Current challenges

The contemporary challenges facing education in Sub-Saharan Africa are arguably located within a history of failed attempts driven by ideological interests that often undermined Africa’s distinctive cultural and political interests. The histories formal European (Western) and Arab (Islamic) persist in the region but are now rooted in a new set of conditions emanating from philanthropic and corporate organizations that espouse a language of reform but often pursue their own interests. Within the context of the reality of limited government funding relative to need, the challenges which at educational systems in sub-Saharan Africa continue to face relate to access and participation, especially of girls and of children disadvantaged by poverty and various forms of disability; inadequacy of basic infrastructure for education in the region and the limited availability of essential resources, facilities and equipment—these cover the range from physical classrooms, stationery, textbooks, laboratory equipment to teacher supply and teacher/leadership training; and the nature (i.e., focus and content) of education as enacted, its relevance and responsiveness to the needs of students and to the needs of the communities (local, national, regional, global) in which they will likely operate.

It is instructive to acknowledge that these challenges persist despite evidence of recent increases in the proportion of expenditure on education, relative to other sectors of the economy, by governments across the region. From a relative low of 15.88% in 2016, government expenditure on education (current, capital, and transfers) ... expressed as a percentage of total general government expenditure on all sectors (including health, education, social services, etc.) has seen a steady increase resulting in 17.88% in 2018 (World Bank Macrotrends, 2022). It is important to note however that these increases, however modest, are occurring against a backdrop of persistent balance of payments deficits for the majority of countries in the region, and mark clear commitment by their governments to address the unfortunate phenomenon that some commentators label “education poverty”.

That access and participation remain a seemingly intractable challenge is underscored by the data available from the UNESCO Institute for Statistics (UIS) from which we learn that:

Of all regions, sub-Saharan Africa has the highest rates of education exclusion. Over one-fifth of children between the ages of about 6 and 11 are out of school, followed by one-third of youth between the ages of about 12 and 14. According to UIS data, almost 60% of youth between the ages of about 15 and 17 are not in school (UIS UNESCO, 2019, p. 1).

Scrutinizing these sobering figures closely, Mogoatlhe (2020, n.p.) reminds us that the situation is “particularly grim for girls. 9 million girls in Sub-Saharan Africa, aged between six and 11, will never set foot in school.” She also notes that “23% of girls are out of primary school, compared to 19% of boys.” And yet, countries in the region have and continue to initiate and implement strategies, by themselves and in collaboration with other countries and/or agencies to work toward the ideal of education for all.

Noting the difficulties posed to access by school fees, previously discussed, a country like Ghana instituted the Free Compulsory Universal Basic Education (FCUBE) program aimed at ensuring equal and free access to basic education for all Ghanaian children. This program operated between 1995 and 2005. Not surprisingly, the expanded access to basic education—covering 2 years of Kindergarten, 6 years of Primary and 3 years of Junior Secondary—created a need for similar expansion at the Senior Secondary and Tertiary levels. The government of Ghana responded, in 2017, with a fee free education for students in Senior Secondary. Ambitious and well-intentioned as such initiatives are, they are not problem-free. Again, the Ghanaian example is instructive as the expansion in student numbers is not always met with commensurable expansions in learning spaces (physical school blocks/classrooms), teacher supply, teaching and learning resources.

Of interest, the recent move to a fee-free education for students in Ghana’s Senior Secondary system has had to be accompanied by the introduction of a double-track² system of schooling to cope with the significant numbers of students taking up Senior Secondary education—a surge from 306,000 students in 2016 to 430,000 in 2018 (Kerr, 2020). This arrangement has had implications for continuity and quality of learning. Briefly, while the access agenda is progressively being met, that of quality, if the critics are to be believed, is progressively deteriorating.

The Ghanaian example makes clear that access is dependent on adequate infrastructure—a fee-free policy without the requisite number of learning spaces, well-qualified teachers and learning materials to cope with the increased numbers of students hampers

²The double track is a shift system of schooling that uses the same facilities to provide for two different student cohorts. In the Ghanaian system, the cohorts alternated their occupancy of the school space such that while one was in session, the other was on a break. This resulted in adjustments to the school year with each cohort spending a significantly shorter period of time at school when it was in session.

both access and quality. Abolishing school fees is not enough—this is a point worthy of emphasis as the Ghanaian example (dubbed Free SHS) took a very comprehensive approach to “free” which extended beyond school fees to uniforms, textbooks, feeding and all costs relating to boarding and lodging for students admitted to boarding schools. It is remarkable, therefore, that anecdotes abound, in Ghana, of parents asking to be allowed to pay fees if that would ensure that their children/wards did not have to participate in the double track system.

According to [Kerr \(2020\)](#), “Currently, 17 governments, representing 37% of Sub-Saharan African (SSA) countries, offer some form of fee-free lower and/or upper secondary education.” While these policies vary in their conception and implementation across countries it is clear that governments across the region are reckoning seriously with the impediment that a regime of school fees poses to participation in education especially by the disadvantaged. As the report notes, in many cases there are hidden costs to what is held up as “free education”, as parents/guardians have to finance the costs of textbooks, uniforms, feeding and the like.

Even in Ghana, where the government pays for all formal costs associated with the Free SHS, ambitious parents are quick to point out that they are spending more of their own money than previously to provide private tuition for their children—this is a complaint that goes to the heart of the quality debate. These hidden costs are no less an impediment to participation especially for girls and children from disadvantaged backgrounds. Enhancing access and participation for girls continues to be a matter of complexity as factors like the lack of access to basic toilet facilities and adolescent hygiene products at school combine with cultural expectations of early marriage, childbirth, farm labor in rural settings to keep girls away from school. Fortunately, the latter are on the decline.

The disruptions wrought on the world of education by the Covid-19 pandemic have thrown into sharp relief the inadequacies of school infrastructure across the region. In many countries of the region, education effectively came to a halt for students as the infrastructure of digital technology did not exist to support the shifts to online learning that were initiated and consolidated in other parts of the world. Even where some form of online learning was made available neither teachers nor students had the knowledge and facilities necessary for effective participation in teaching and learning. In some parts of Kenya, some lessons were made available to students via radio networks. Children from families without a radio could not participate.

As Human Rights Watch reports, “Many children received no education after schools closed across the continent in March 2020” ([Human Rights Watch, 2020](#); n.p.). In the presentation to their *35th Ordinary Session of the African Committee of Experts on the Rights and Welfare of the Child* they demonstrated in very vivid terms the ways in which “school closures caused by the pandemic exacerbated previously existing inequalities, and that children who were already most at risk of being excluded from a quality education have been most affected.”

Covid-19 aside, however, the issue of educational infrastructure, or lack thereof, has underlined a long-standing challenge in the region. Reference has already been made to the adoption of a double track system in Ghana as a strategy to cope with the surge in student numbers at the introduction of the Free SHS. The region does not have enough schools or classrooms for its students. UNESCO’s Institute for statistics identifies “poor infrastructure” as a critical barrier to education ([UIS UNESCO, 2016](#), Slide 4). It is not unusual for children in urban centers to have overcrowded classrooms relative to the physical size of the class or to attend schools that lack the basics such as a board for each class, or toilet facilities or clean water or electricity or adequate textbooks and stationery. Similarly, it is not unusual for children in country towns or rural villages to have to walk long hazardous journeys, daily, to a school or attend classes in dilapidated buildings or under trees.

If a school’s most important resource is its teachers, then this is an area of significant challenge in the region owing to problems of supply. As government policies and actions lead to expansion of access to education for all children in the region, more teachers will continue to be needed. However, as a relatively poorly paid profession in the region, teaching does not attract many takers or stayers and there is, thus, a high attrition rate. Against such low take up and high attrition—from retirements, resignations, emigrations, mortality—the demand for teachers, has always outstripped supply. Historically, countries in the region have made do with significant numbers of “pupil teachers” and university graduates without teaching qualifications but this was before the policies and programs to make education free and accessible to all children of school going age. For example, “In Lesotho a 2007 study found that 40% of primary teachers were unqualified (Education International, 2007). In the same year there were 40,000 qualified teachers unemployed in Kenya, and a further 15,000 unemployed in Zambia (Education International, 2007)” ([UNESCO, 2010](#), p. 1).

These figures suggest a general dissatisfaction within the region with teaching as a worthwhile profession. Challenges of remuneration, retention, ongoing professional learning, career progress and fulfillment loom large alongside those of training and recruitment. Any attempt to narrow the gap between demand and supply of teachers must seek to address these challenges as a matter, also, of addressing the quality of teaching and learning. In Ghana, for example, the number of teacher training institutions, known as Colleges of Education (CoE) is on the rise—Isaac [Buabeng et al. \(2020\)](#) note that, from 38 CoEs in 2014, the number now stands at 48 CoEs. The increase is a result of deliberate government intervention which resulted in the absorption of “10 private CoEs with the aim of expanding the infrastructural facilities in the colleges” (p. 87) Typically, the CoEs offer three-year Diploma in Basic Education (DBE) programmes. Alongside the CoEs, the University of Cape Coast and the University of Education, Winneba, well-known as traditional teacher education institutions run programmes leading to academic degrees in Education ([Buabeng et al., 2020](#)). Beyond initial training and certification, however, there seems to be very little effort on the part of governments in the region to tackle the needs of ongoing teacher professional development beyond massive subject focused teacher workshops to introduce new or revised curriculum as necessary.

The challenges highlighted in this discussion, while important in themselves, are also of significance because of the adverse impact they stand to have on the overall quality of education in the region. Even as literacy rates increase with the efforts to expand

educational access, it is of concern that the quality of education on offer does not result in outcomes that meet minimum standards globally. Mogoatlhe (2020) tells us that “90% of children aged between 6 and 14 ... won’t reach the minimum reading requirements even after they complete secondary education.” This telling statistic underscores Mogoatlhe’s assertion that “it’s also not enough to just be in school—the quality of the education also has to be good for a child to thrive.”

However, I would argue that the quality of education for a region as distinctive as Sub-Saharan Africa, should be determined by a more comprehensive set of performance markers than are currently available from reliance on student achievement on standardized testing. Proceeding by questioning in the concluding segment of this paper, it is important to note the need to approach the issue of quality from the standpoints of both curriculum content and pedagogy, together with the recognition that Sub-Saharan African countries are among the poorest of the world (Mogoatlhe, 2020) and lack the resources to invest heavily in education. Against this recognition, a significant regional and national objective for education must not only be a focus on educational outcomes but also the alleviation of poverty.

Conclusion

The discussion in this chapter has highlighted the continuing challenges that systems of education face in Sub-Saharan education. The challenges are not only a product of historical legacies but also the current pressures, economic, political and cultural. Various international organizations have attempted to help but have not been able to align their interests with those of Sub-Saharan Africa. May reform attempts have failed, while others lack conditions of sustainability and suitability. The criteria by which the quality of education in the region should be assessed continue to be contested. Ultimately however these criteria need to consider the following questions more seriously and critically:

- To what extent does the content of curriculum in the region align with the needs of national development for its various countries?
- What importance is given to the languages and cultures of the region in the education of its children?
- Are the ways of knowing and being in the region acknowledged and valorized in the curriculum?
- To what extent does teaching and learning routinely connect students to the wealth of their own cultural contexts as resources for learning?
- To what extent does the local community participate in the teaching and learning process?
- How does teaching and learning provide opportunities for schools and students to collaborate with community (local, national, or regional) to solve/address problems of interest?
- How does teaching and learning harness technology to provide experiences that support the learning needs of each student and, thereby, enhance the development and consolidation of desired competences?
- To what extent and how effectively do teachers actively collaborate with their peers, students and their parents in the pursuit of effective praxis as lifelong learners?

These questions should, hopefully, provoke reflection on the nature of education appropriate to the needs of the region. In many ways they reflect standards and practices to be aspired toward if the quality of education in Sub-Saharan Africa is to improve in the ways that address the needs of national development and poverty alleviation while also affirming the cultural identity of students and supporting their holistic development. They also point to the need for the curriculum and its teaching to be decolonized through engagement in culturally relevant, responsive and sustaining practices. As well, they nudge into consciousness the levels of professionalism and professional practice that teachers in the region should be aspiring toward.

Taken together, the questions return us to the purposes of education for the Sub-Saharan African region. They also offer a perspective on what the experience of learning should be for students, teachers, schools and the communities that host and, in many ways, sponsor them. Implicit is the consideration of whether an education rooted in the resources (and daily challenges) of the community might prove to be more accessible and engaging to all, regardless of financial standing? Might it also engender modes of problem solving that serve to create new knowledge for sharing with others? But perhaps, we are getting ahead of ourselves. There is yet much work for the region’s governments and stakeholders in education to do to realize the possibilities of quality education envisaged here. Sadly, too much of the leadership of educational reform is currently ceded to the region’s development partners. The governments of the region need to step fully into this role as a matter of urgency.

References

- Al-Samarrai, S., Cerdan-Infantes, P., Bigarinova, A., Bodmer, J., Vital, M.J.A., Antoninis, M., Barakat, B.F., Murakami, Y., 2021. Education Finance Watch 2021. World Bank Group, Washington, DC. <http://documents.worldbank.org/curated/en/226481614027788096/Education-Finance-Watch-2021>. (Accessed March 2022).
- Austin, G., 2010. African economic development and colonial legacies. Open Edition Journals. <https://journals.openedition.org/poldev/78>. (Accessed March 2022).
- Ayittey, G.B.N., 2006. The Postcolonial Destruction and Betrayal of Africa, Ch.11. Indigenous African Institutions, second ed. Brill | Nijhoff. <https://doi.org/10.1163/ej.9781571053374.i-586.63>.
- Becker, B., 2021. Mission education left an uneven legacy: an analysis of 26 African countries. The Conversation. <https://theconversation.com/mission-education-left-an-uneven-legacy-an-analysis-of-26-african-countries-166315#:~:text=Mission%20societies%20often%20expanded%20into,which%20occurred%20between%201884%201914>. (Accessed October 2021).

- Brock-Utne, B., 2013. Globalisation and Inequality: Challenges to Education in Africa". World Studies in Education. James Nicholas Publishers. <https://doi.org/10.7459/wse/14.1.03>.
- Buabeng, I., et al., 2020. Teacher education in Ghana: policies and practices. *J. Curric. Teach.* 9 (1), 86–95. <https://eric.ed.gov/?id=EJ1248146>.
- d'Aiglepiere, R., et al., 2018. Arab-Islamic education in Sub-Saharan Africa: going beyond clichés to build the future. *The Conversation*. <https://theconversation.com/arab-islamic-education-in-sub-saharan-africa-going-beyond-cliches-to-build-the-future-93442>. (Accessed October 2021).
- Fanon, F., 1967. *The Wretched of the Earth*. Penguin Classics.
- Global Education Monitoring Report, 2020. Inclusion and Education: All Means All. UNESCO. https://reliefweb.int/report/world/global-education-monitoring-report-2020-inclusion-and-education-all-means-all?gclid=EAlaQobChMI7MWpkYXM9glVRIJLBR3cog56EAAAYASAAEgKs3_D_BwE. (Accessed September 2021).
- Heidhues, F., Obare, G., 2011. Lessons from structural adjustment programmes and their effects in Africa. *Q. J. Int. Agric.* 50 (1), 55–64. https://ageconsearch.umn.edu/files-4_HeidhuesPDF.
- Human Rights Watch, 2020. Impact of Covid-19 on Children's Education in Africa. <https://www.hrw.org/news/2020/08/26/impact-covid-19-childrens-education-africa>. (Accessed November 2020).
- Johnson, T., 2002. Educational reforms in sub-Saharan Africa: 1980s–1990s. Perspective. <https://www.theperspective.org/educationalreforms.html>. (Accessed October 2021).
- Kerr, K., 2020. Secondary Education Isn't Always Accessible for Youth Across Africa. <https://www.theafricareport.com/26442/secondary-education-isnt-always-accessible-for-youth-across-africa/>.
- Konadu-Agyemang, K. (Ed.), 2002. IMF and World Bank Sponsored Structural Adjustment Programs in Africa: Ghana's Experience, 1983–1999. Routledge. <https://www.routledge.com/IMF-and-World-Bank-Sponsored-Structural-Adjustment-Programs-in-Africa-Ghanas-Konadu-Agyemang/p/book/9780367666385#>.
- Lingard, B., 2015. Complementarities and Contradiction in the Pearson Agenda. *Education International*. <https://www.ei-ie.org/en/item/21300:complementarities-and-contradiction-in-the-pearson-agenda>.
- Mart, C.T., 2011. British colonial education policy in Africa. *Academic Journals*. https://academicjournals.org/article/article1379428447_Mart.pdf. (Accessed October 2021).
- Mogoathe, L., 2020. What Does Universal Quality Education Really Mean for Sub-Saharan Africa? *Global Citizen*. <https://www.globalcitizen.org/en/content/universal-quality-education-in-africa/>.
- Okolo, M.S.C., 2005. Reassessing the impact of colonial languages on the African identity for African development. *Afr. Media Rev.* 13 (2), 85–103. https://codesria.org/IMG/pdf/05_Okolo.pdf. (Accessed March 2022).
- Pamba, E.S., 2012. An Analysis of the Impact of Structural Adjustment Programmes on Education in Sub-Saharan Africa: A Case Study of Kenya (1980–2010). University of Nairobi. http://erepository.uonbi.ac.ke/bitstream/handle/11295/8564/pamba_An%20Analysis%20of%20the%20Impact%20of%20Structural%20Adjustment%20Programmes%20on%20Education%2C%20in%20Sub-saharan%20Africa%20a%20Case%20Study%20of.pdf?sequence=1&isAllowed=y. (Accessed November 2021).
- Pearson, 2017. Pearson Extends Partnership With Camfed to Improve Girls' Education in Africa. <https://plc.pearson.com/en-US/news/pearson-extends-partnership-camfed-improve-girls-education-africa>. (Accessed March 2022).
- Scott, R., 2013. Indirect rule in Tropical Africa: up close and personal. *J. Hist. Biography*. https://espace.library.uq.edu.au/data/UQ_322054/UQ322054_OA.pdf?Expires=1647487626&Key-Pair-Id=APKAJKNBJ4MJBJNC6NLQ&Signature=Q4UQz3nyW7D1j6hiv6zDQg8abjsySZIYvPdEwO22RbG0ej7Fe2FTdoGLL5LC5SWgQIPIMljo0Mt7HPuXi~mzEcNjg2hMTQXayZaamGGEI6eQFURTRK-sG8kXGSNNHsSE2kcrTOPuYOWJkJ8TO-BAMngYaHezNs66KyCH1xclapYnTpkN6GxGdqsI-hZk9OcOvwHgHph2MYfxLIFNadZ9TVbelmcNjX2te4lppCmAxymsmbwYE5gEtZVFXId3UJiaqOi5~xNdcH04ZM2HYVSPU5AF0y9qBDqZaG0ea6svrVkl4hq8mrzP~xyIZDalwjliUPyHTraR~ligyKIQxM~g__. (Accessed March 2022).
- Suleiman, M., Onapajo, H., 2022. Why west Africa has had so many coups and how to prevent more. *The Conversation*. <https://theconversation.com/why-west-africa-has-had-so-many-coups-and-how-to-prevent-more-176577>. (Accessed March 2022).
- Task Force on Higher Education and Society. 2000. Higher Education in Developing Countries Peril and Promise. World Bank Group. <https://elibrary.worldbank.org/doi/epdf/10.1596/0-8213-4630-X>.
- UIS, 2019. Education in Africa. UNESCO. <http://uis.unesco.org/en/topic/education-africa>. (Accessed October 2021).
- UIS UNESCO, 2016. School Resources and Learning Environment in Africa. <http://uis.unesco.org/sites/default/files/school-resources-and-learning-environment-in-africa-2016-en/school-resources-and-learning-environment-in-africa-2016-en.pdf>. (Accessed October 2021).
- UIS UNESCO, 2019. Fact Sheet No. 56 New Methodology Shows that 258 Million Children, Adolescents and Youth Are Out of School. <http://uis.unesco.org/sites/default/files/documents/new-methodology-shows-258-million-children-adolescents-and-youth-are-out-school.pdf>. (Accessed October 2021).
- UNESCO, 2010. Teacher Attrition in Sub-Saharan Africa: The Neglected Dimension of the Teacher Supply Challenge; A Review of Literature. <https://unesdoc.unesco.org/ark:/48223/pf0000188197>. (Accessed November 2020).
- United Nations, 1948. Universal Declaration of Human Rights. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>. (Accessed March 2022).
- Wolff, E., 2018. Africa Failing to Address Linguistic Imperialism. *University World News*. <https://www.universityworldnews.com/post.php?story=20180222092733807>. (Accessed March 2022).
- World Bank Macrotrends, 2000–2022. Sub-Saharan Africa Education Spending 2000–2022. www.macrotrends.net. (Accessed 19 May 2022).
- World Bank, 2001. World Development Report 2000–2001: Attacking Poverty. World Development Report. Oxford University Press. © World Bank. <https://openknowledge.worldbank.org/handle/10986/11856> License: CC BY 3.0 IGO. (Accessed March 2022). <https://openknowledge.worldbank.org/handle/10986/11856>.

Shifting policies of education in Latin América

Carlos Ornelas, Education and Communications at the Metropolitan Autonomous University, Mexico City, Mexico

© 2023 Elsevier Ltd. All rights reserved.

Introduction	398
Contextual record and basic concepts	398
Education and the consolidation of the national state	399
The expiration of the 20th-century milieu	401
Decentralization: a frail course of action	401
Lifelong learning: an upright target	402
The curricula: broken bridges between law and practice	402
Assessment to cope with low quality of education	405
Teacher professionalization, identity, and collective representation	406
Covid-19 and the right to education	407
Ending	409
References	409

Introduction

Since the 1980s, the educational systems of Latin America underwent analogous shifts despite significant disparities between countries. Education became an artifact to join globalization from being a lever for the nation-state's consolidation from independence to the early twentieth century. Besides emphasizing growth on enrollments, educational policies move to inclusion, equity, quality, and labor training. Nonetheless, the public education enterprise has been historically poorly founded. Intergovernmental organizations and national governments pushed for shifts in politics and education policies under the umbrella of a universal educational reform project. Still, it does not eradicate Latin America's distinctive feature and its educational systems prevalence of vast inequalities due to ethnicity, social class, gender, religious beliefs, or sexual orientation fermented in the 19th and 20th centuries. Inequalities that worsened in 2020 because of the Covid-19 contagion.

Although educational reform processes have been complex, diverse, and with innumerable design and implementation problems, to facilitate the analysis, they can be grouped into two main models. One of the initiatives of such reforms to cope with quality and equity is privatizing education systems for economic and ideological reasons. A neoliberal call's dictums favor market mechanisms that emphasize training for productive work, free choice, and school autonomy. However, although without the economic power and propaganda means, democratic tendencies oppose that vision. These conceive education as a public good and propose reforms that stress the human being, respect for cultural differences, and substantive equality.

By the 1980s, most quantitative education indicators showed improvement even though the population grew due to progress in preventive health, but the general perception was that schooling was of low quality. Thus, in the 1990s, intergovernmental institutions nurtured the tendency to decentralize school systems and strengthen adult education and lifelong learning. But at the same time, under central government control, encouraged shifts in the pedagogical orientations from the content and teacher-based curriculum with an emphasis on general teaching to competence-based or constructivist approaches that focus on students' learning. And governments of most countries created specialized institutions to measure such education. The human capital pattern to train for work dominates, but diverse nature experiences flourish to educate for responsible citizenship and freedom. In the reform processes, all participants agreed that the professionalization of teachers was essential. Yet, teacher unions resisted the push for neoliberal reforms.

This article focuses on educational politics and policy. It starts picturing trends after independence in the 1800s and emphasizes the shifts in the 1980s until the present day. It covers from initial education to K12.

Contextual record and basic concepts

Latin America is a geographical area of great contrasts, with modern economies that mix with subsistence or extreme scarcity and vast social disparities and regimes that postulate humanism, democracy, and social justice, but which political practices result in authoritarianism and patronage. Despite so, with few exceptions, the governing groups proclaim the purpose of reforming school systems to improve the quality of education, advance social inclusion, and further equity. Even with all the impetus and expenses, overall, the expected results have not been produced, although there is some progress in enrollment growth, a higher transition between school levels, and vocational education provision.

Although the concepts are not antagonistic, there is a query for education between quality and equity. Both conceptions coincide in the proposals for change. The search for quality education is closer to technocratic approaches under human capital theory, while

the quest for equity pursues humanistic ends. The first stresses measure student performance in literacy and numeracy, the second in the personality's harmonious development.

The word inequality refers to economic or any other magnitudes that are different and feasible to measure. The term inequity and others with the same connotation pass value judgments on those magnitudes. The notion of equity does not mean treating everyone the same way; the concept connects with distributive justice: the fundamental postulate is that the search for equity implies equal (normative) concern for all members of a society (all are equal before the law) and dissimilarity in consideration of those with social disadvantages (Sen, 2009). However, there is a global trend to transform basic education, to which Latin America is no alien. Although it does not disregard the objectives of equity, it superimposes the urgency of raising the quality of teaching, emphasizing the acquisition of skills for the information age: the universal educational reform model.

The universal model of education reform for the information age, mainly promoted by UNESCO, the World Bank, and the OECD, depicts five features: (1) Decentralization / school autonomy. (2) Lifelong learning. (3) Centralized curriculum based on competencies. (4) Central evaluation systems. And (5) Professionalization of teachers (Beech, 2008). Some Latin American governments accepted and adopted such features in the 1990s, however, amid social conflicts and political changes that affected the governance of education systems because, in each case, there is a dialectic between the national and the local (Arnové, 2013). Governments have to deal with internal political forces, face social groups' contentions, and pressure from teachers' unions.

Education and the consolidation of the national state

Schwartzman synthesizes colonization and exploitation history that the Spanish and Portuguese empires subjected to native peoples of what is now Latin America. The conquerors arrived to exploit the new world's treasures. "Many of them did not bring their families, and, even if they did, mingled with the local population and the slaves, creating large, stratified mixed-blood populations that also combined the cultural, religious, and linguistic traits of their different origins." (Schwartzman, 2015) Still, each of the future independent nations developed distinct cultural characteristics. There are countries with a large native population such as Bolivia, Mexico, Peru, and part of Central America, and others where people of European descent are the majority, such as Argentina, Chile, and Uruguay. Immigrants from other regions, Palestine, China, or Japan, for example, enriched the miscegenation and enlarged the cultural complexity of the area.

In the Spanish colonies and Brazil, a Portuguese settlement, the dominant groups were of European ascendance (criollos), and the original population constituted the oppressed castes. The Europeans conquered the native peoples and forced them to work in mining and farming to extract raw materials and grow products for the world market. They also imported millions of Africans to work as slaves on large plantations; the conquerors enforced Catholicism as the official religion and Spanish or Portuguese as the overriding languages.

The wars of independence of Spanish America led to the establishment of republics in the early nineteenth century. Still, most of them suffered civil wars, military dictatorships, and the division of territories. For decades there was nothing, urban and regional centers that struggled to maintain an external representation of a national daydream (Córdova, 1977). Brazil became the Portuguese empire's headquarters until 1889 when it shifted to an independent nation and a republican regime.

Although enlightenment and political liberalism had intellectual influence in Latin American independence struggles, the nascent republics did not forge a liberal democratic state. It was an oligarchic state whose economic structure was inherited from the colonial period: the self-sufficient hacienda (latifundio). It focused on the indigenous population's forced labor, poor mestizos, and descendants of slaves (Medina Echaverría, 1964). In that state, the military caste took charge of the bureaucracy, and the Church of the few education centers. Furthermore, given the collection of tithes and mortgages, the Church became an economic power.

Governments were unstable and bureaucratic, where corruption and compadres abounded. The military and ecclesiastical institutions were the most important. Despite the republican regime, social, economic, and political inequality persisted in the region's countries. Throughout the first century of independence, authorities reestablished political ties with Spain, and Catholicism kept being the official religion; the Catholic Church was part of those states, and even it was the hegemonic power until the late 1860s (Halperin Donghi, 1969).

Although the law established that education would be free in many countries, both due to civil wars and fiscal hardship, education systems were inexistent. Schooling was for a minority, to prepare clergymen and public servants in seminaries and universities. Countries like Mexico, Brazil, Colombia, Argentina, Uruguay, and Ecuador started Lancaster and Bell schools from the 1820s onwards. It was oriented to serve the poorest classes with a student-monitor method. Parishes, humble houses, and public buildings hosted the groups. In Mexico, those schools became part of official instruction in the 1830s (García Benavente, 2015).

The dominant groups, with the landowners at the top, tried to strengthen the state; they maintained their concern for egalitarianism in the field of law, not in practical life. The governing classes wanted the masses to identify with the nation. Strengthening national identity became a mission for education. Against the Church's requests and conservative forces, liberal governments sustained public instruction institutions. The patriotic and civic teachings supported the construction of legitimacy and the crystallization of ferments of collective identity. The concept of homeland became identified with the nation. The first curricula introduced a heroic sense of the recent past constituted by the Independence revolutions. Likewise, the oligarchic groups used

bordering conflicts with other republics as elements to promote the nationalist spirit. Patriotism became an actual national project (Ossenbach Sauter, 1993).

However, school systems' institution was slow and charged with disputes between religious and secular conceptions; nonspiritual thinking was gaining ground, especially in the more populous countries. The most radical expression was in Mexico that after a social revolution, it decreed secular education in the Constitution of 1917, and socialist education in the amendment of 1933 (Ornelas, 1995). During the first decades of the 20th century, Latin American countries struggled to create education systems; first, to increase literacy and to instruct the mass of workers and, second, to force indigenous peoples to speak Spanish or Portuguese and incorporate them into the national state. According to Bergés (2009), literacy percentages grew in the first half of the 20th century, but unevenly. For instance, in 1900, Argentina and Uruguay had rates of literate persons of 51% and 59%. By 1950 it grew to 88% in the first country and 86% in the second. Others, like Mexico and Peru, barely exceeded 20% at the beginning of the century, and, by 1950, they reached 61% and 51%.

As shown in Table 1, except for Bolivia, Brazil, Peru, and the Central American countries, all with high percentages of the indigenous population, illiteracy was less than 20%. In the Dominican Republic and Brazil, there is also a high percentage of people of African descent. Except for Cuba, there are more illiterate women than men. Nevertheless, the gross enrollment rates in primary school were already over 70%, and those in secondary education stood growing. Similarly, the average number of students per teacher in primary school seemed acceptable, but in Colombia, Ecuador, Mexico, Peru, and Central America. The lag in the Dominican Republic was considerable. Funding for public education, however, was (and continues to date) low. Only Costa Rica and Ecuador allocated more than 5% of their GNP to education. According to the Economic Commission for Latin America and the Caribbean (ECLAC), over 90% of spending went to teacher and administrator salaries across the region (ECLAC, 2001).

Although illiteracy among adults was still high in the poorest nations, universal attendance to primary education was a fact in almost all the region's countries. There were progress and some social mobility; schooling systems incorporated large popular sectors. However, the expansion of enrollments did not reduce social inequality; the growing urban middle classes were the primary beneficiaries of the evolution of schooling, while the low quality of education became a matter of concern.

Once national states were consolidated, and the borders between countries varied little, toward the end of the 1960s, the human capital theory began to penetrate educational policy discussions. In contrast, intellectuals started to debate the concept of equal opportunities in the 1970s, and by the 1980s, the emergence of global trends marked the educational politics and policies of Latin America.

Table 1 Basic indicators of education in Latin America (1980).

Country	Illiteracy 15 years + %			Gross enrolmen rates		Gross enrolmen rates		Public spending	
				Primary schooling		Secondary schooling			
	M	F		Ages	Rate	Ages	Rate	Student/teacher primary schooling	%
1 Argentina	6.0	5.7	6.4	6–12	111	13–17	56	19	1.9
2 Bolivia	30.9	19.9	41.3	6–13	76	14–17	37	18	3.8
3 Brazil	25.4	23.7	27.2	7–14	98	15–17	34	26	0.7
4 Chile	8.5	7.9	9.1	6–13	113	14–17	61		4.1
5 Colombia	15.6	14.7	16.4	6–10	107	11–16	34	31	2.5
6 Costa Rica	8.3	8.1	8.4	6–11	107	12–16	70	20	6.2
7 Cuba	7.9	8.1	7.8	6–11	106	12–17	81	17	–
8 Ecuador	18.1	14.4	21.8	6–11	117	12–17	53	35	5.3
9 El Salvador	33.8	29.1	38.4	7–15	74	16–18	25	48	3.4
10 Guatemala	46.2	38.1	54.3	7–12	73	13–18	17	34	1.8
11 Honduras	39.0	37.2	40.8	7–12	98	13–17	30	37	3.0
12 Mexico	17.0	13.8	20.2	6–11	120	12–17	49	39	3.1
13 Nicaragua	41.8	41.5	42.1	7–12	98	13–17	43	35	3.5
14 Panama	14.3	13.7	15.1	6–11	106	12–17	61	27	4.5
15 Paraguay	14.1	10.6	17.5	7–12	105	13–18	27	27	1.3
16 Peru	20.2	11.7	28.8	6–11	114	12–16	64	37	2.5
17 Dominican Republic	26.2	25.2	27.3	7–12	107	13–18	42	56	2.1
18 Uruguay	5.3	5.7	4.8	6–11	107	12–17	59	21	1.9
19 Venezuela	15.1	13.3	16.9	7–15	110	16–18	40	27	4.1

Source: Own elaboration based on ECLAC (2001).

The expiration of the 20th-century milieu

Through the 20th century, the oligarchic state archetype moved to various forms of corporatism and military dictatorships; democracy was weak. The idea of education at the homeland service continued, but with increases in enrollment and escalation to secondary and higher education. Also, many governments, like in Argentina, Brazil, Chile, Colombia, Mexico, and Peru, established vocational training to prepare workers for the industry. Still, growth did not mean proper quality instruction; thus, by the 1980s, the leading groups started education reforms aiming at accomplishing better-quality teaching and equity in education. Nevertheless, social exclusion continued to be the landmark locus in Latin America (Treviño, 2005).

The influence of intergovernmental organizations grew; in many countries, they even played as local political actors. In the 1980s, those organizations began to advocate for educational reforms to improve quality; enrollment growth, they argued, while it should continue, was insufficient. Most countries in the region accepted the Jomtien declaration of Education For All (EFA), which UNESCO called to continue with the expansion of enrollment and offer enhanced education (UNESCO, 1993). The era of globalization manifested itself in the strain between studying for work—as proclaimed by organizations such as the World Bank and the OECD—and education as a fundamental human right, championed by UNESCO.

Most of the region's governments accepted the EFA slogans but were unable to meet the goals for the year 2000. However, there were impulses to equity in education due to the propagation of new actors interested in learning. Intellectuals and organizations, such as the Economic Commission for Latin America and the Caribbean, reviewed the slogan of equal opportunities and began to speak of educational equity and social inclusion (Treviño, 2005). In addition to access, it was necessary to guarantee permanence and graduation from schooling levels. Furthermore, the gaps in the quality of education offered to boys and girls, middle and poor classes, urban and rural populations, had to be reduced. The most vulnerable segments of the populaces should be favored. Like Pablo Latapí in Mexico, intellectuals introduced the concept of equity to public debate and proposed compensatory programs to support the disadvantaged (Latapí, 1993).

Throughout the 20th century, governments in the region, churches of various faiths, civil organizations, service clubs, and people in business launched non-formal education projects to alleviate some of the systems' deficiencies. Such as illiteracy and problems derived from social inequity or race or gender issues. UNESCO and other international donors did the same. Unfortunately, by the end of the twentieth century, except for a few short-term or literacy projects that had some success, the record included more failures (Cortina and Stromquist, 2000; Post, 2002; Reimers, 2000; Rivero, 1999). Social exclusion continues to be one of the negative badges of Latin American societies.

Despite these deficiencies and many governments' shortcomings, the hope for better quality and equitable education remained. Dispersed sets of educational experiences that met—not all, but many—of the population's expectations maintained the flame. The ideas of Paulo Freire on popular education took off in the 1970s. Those were edifying exercises whose primary purposes and their center of action were quality and equity: good educational practices (Ornelas, 2005). Still, global institutions pushed for the international model of education reform.

Decentralization: a frail course of action

The World Bank began to gain prominence with proposals for more financing and better administration; educational decentralization, it argued, was the precondition for optimizing the use of scarce resources. Furthermore, bringing the authorities closer to schools and parents would allow governments to understand the schools' problems better and efficiently channel energies to solve them (Winkler, 1989).

Perhaps for internal reasons, but also under the influence of the World Bank, national and local governments took the slogan of educational decentralization. Senior government officials replicated it in their discourse, there were academic forums, and the Bank allocated resources to finance research to strengthen local autonomy and management capacity. The drive covered the entire continent, from Chile and Argentina to Central America and Mexico (Carnoy et al., 2005; Edwards, 2019; Ornelas, 1995). Brazil was a case apart. Its federal system has been in operation since the 19th century, but the states transferred control and primary education financing to the municipalities. The reforms of the 1990s planned to strengthen the counties (Guimarães de Castro, 2015). Cuba and Uruguay decided to keep its school systems under central control.

The policy of educational decentralization had radical tones. In Chile, the extreme case during the Augusto Pinochet government, which transferred public education to the municipalities, established voucher programs that promoted privatization so that it would function according to the logic of the market (Ávalos and Bellei, 2019; Bellei and Vanni, 2015; McEwan and Carnoy, 2000; Muñoz Stuardo and Weinstein Cayuela, 2019). There were community control experiences of schools and school autonomy in areas emerging from civil wars such as EDUCO in El Salvador (Edwards, 2020), later replicated in Guatemala and Honduras (Ganimian, 2016). The ideology of efficiency governed the decentralization efforts in Latin America's education. However, the process in countries that proclaimed educational decentralization transferred degrees of authority and control to sub-national entities—and considerable financial resources in Mexico's case—but centralized the curricular and political decisions in the central ministries (Kubal, 2003).

Although it tried to change discourse and incorporate other actors into the educational policy, the decentralization move did not work overall. There were some exceptions, but a good part of sub-national governments did not go beyond a reformist discourse; in their daily practice, however, they showed disinterest, cooperation with the previous uses and corruptions of the centralized

structure, fear of conflict, or they replicated the speechmaking but did not act on a consequence. By the end of the twentieth century, there was no proof that through the decentralization policies, the region's school systems enhanced the quality of education (Carney, 2012).

Teacher unions and traditional bureaucracy opposed educational decentralization. Some for good reasons, because if there were not enough administrative capacities at the national scale, even less at the sub-national or local levels. Union leaders knew that with decentralization, they would lose power and mediation with the central government (Schneider, 2021). Parents, business associations, or civil society organizations interested in educational tasks without reaching concrete and binding levels of participation, nor institutional and permanent mechanisms (Kubal, 2003).

Lifelong learning: an upright target

Although lifelong learning is part of the international reform model, it is also an ingredient of sustainability projects, with even more history and contributions. The first type of prioritizes economic development and labor productivity skills to increase participation in global value chains and promote equity and inclusion (OECD, 2017). For the second, lifelong learning has a critical role in eradicating poverty, protecting the planet, securing human rights, building peaceful, inclusive, and equal societies, and promoting social, economic, cultural, and technological progress based on the values of the United Nations (UNESCO Institute for Lifelong Education, 2012).

The generic concept in practice until a few decades ago was that of adult education. Lifelong education has two main aspects, literacy (includes digital literacy) and training for productive work. It is a global ideal even before the term globalization was known. Since its creation, UNESCO has promoted six International Conferences on Adult Education CONFITEA (from the French, Conférence International en pratique sur l'Education des Adultes).

A proposal from Mexican diplomats gave way to the Regional Center for Fundamental Education for Latin America (CREFAL). It started in 1951 with the Mexican government's sponsorship and the UNESCO and the Organization of American States' support. Its purpose was to support Latin American governments to meet urgent needs for fundamental education: the training of adult education teachers and the production of teaching materials. To promote a broad multidisciplinary outlook, CREFAL ensured specialized personnel provided by other United Nations Organizations. The International Labor Organization (offered a specialist in arts and crafts); the World Health Organization (experts in health education, childcare, rational nutrition, and individual and social hygiene); the Food and Agriculture Organization professionals in crop improvement methods and conservation of natural resources, home economics, family industries and recreational leisure, organization of rural cooperatives (Torres Bodet, 2017).

Throughout the 1970s, CREFAL became one of Paulo Freire's main disseminators and practitioners' methods and ideas on resistance, hope, and critical consciousness. Freire's best-known texts, *Pedagogy of the oppressed* and *Education as a practice of freedom*, were compulsory readings in seminars and congresses. This institution published books and manuals on literacy and adult education; it even tried to formalize specific Freire's non-formal education proposals (Rodrigues Brandao, 1977). Together with other Mexican and Latin American institutions, CREFAL promoted distance education, job training, strengthening of local cultures and traditions, and international citizenship at its Center for Sustainable Community Development (Beltrán Morales, 2015). Since October 1990, it is named Regional Cooperation Center for Adult Education in Latin America and the Caribbean. Eleven countries are part of it.

All Latin American countries agreed to accomplish the UNESCO's Sustainable Development goal 4, primarily targets 3. Equal success to technical/vocational and higher education; 4. Relevant skills for decent work; and 6. Universal youth literacy (United Nations Organization, 2016). However, due to the World Bank's influence, and because governments detected the need for a competent workforce for industry and services, grew the supply of technical vocational education (ETP) in secondary school. The ECLAC documents that all countries have promoted it, as shown in Table 2 (Sevilla, 2017). The theory of human capital reigns in that territory. However, many social sectors consider it a second-rank training—an offer for the poor classes' offspring, which refutes the motto of social equity.

Although lifelong learning is a shared concept, the query between quality and equity is underlying. It is also in the pedagogical field, but openly between traditional teaching and competence-based curricula, amid contending teaching approaches and the emphasis on transmitting information to learning and critical thinking claims.

The curricula: broken bridges between law and practice

Although diminished by incorporating more subjects in the curriculum, patriotism continues in-school programs in almost all countries. The traditional method, centered on teaching, may still be dominant; teachers' and institutions' cultural persistence is steady in Latin America. However, the competency-based approach makes its way in legislation, agreements between actors and programs. Intergovernmental agencies advocate a pedagogy that emphasizes communication skills, creativity, flexibility, learning to learn, working in groups, and problem-solving (Beech, 2008). Broadly, those programs of change propose a national core curriculum under central control. On a smaller scale, popular education offers similar changes, emphasizing critical awareness, and decolonization of knowledge under the umbrella of intercultural understanding or interculturality (Cortina, 2017).

Table 2 Latin America (ETP programs in school systems with plans for segmented secondary studies).

Country	<i>Propedeutic ETP in the upper secondary</i>		<i>ETP provision variants</i>		
	<i>Years of study</i>	<i>Distinctive title granted</i>	<i>ETP</i>	<i>ETP</i>	<i>Optional ETP deepening</i>
			<i>Prevocational</i>	<i>Professionalizing</i>	
Argentina	5-6 ^{a,b}	Middle technician	Integrated to high ETP	–	–
Brasil	1–3	Middle technician	–	–	–
Chile	2	Middle technician	–	–	–
Colombia	2	Middle technician	–	–	–
Costa Rica	3 ^a	Middle technician	–	–	–
Cuba	3–4	Middle technician	–	Skilled worker	–
Ecuador	3 ^d	Technical bachelor	–	–	Productive technical bachelor
Guatemala	2–3	Bachelor specific application	–	–	Proficient
Honduras	3 ^c	Technical or proficient	Prevocational common	–	–
Mexico	3	Middle technician	Technical junior secondary	Secondary for workers - technical professional	–
Nicaragua	2–3	Middle technician	–	General technician	Specialized technician
Panama	3	Middle technician	–	Second industrial cycle	–
Paraguay	3	Middle technician	Agricultural professional initiation	Middle vocational training	–
Dominican Republic	3	Vocational technician	–	Basic technician	–
El Salvador	3 ^c	Middle technician	–	–	–
Uruguay	3	Middle technician	Technological basic cycle	Basic professional training. Qualified operator	–
Venezuela (Bolivarian República of)	6 ^{a,c}	Middle technician	Integrated to high ETP	–	–

^aTotal duration low and high ETP.

^bConsiders the changes included in the reform in progress in that country that extends the ETP from 2 to 3 years and incorporates the basic technical 2-year duration.

^cDuration of ETP equal to one additional year to the academic study plan.

^dDuration of ETP equal to the academic study plan but with a higher hourly load.

Source: Own elaboration based on Sevilla (2017).

Educational reforms increase in the region from the down of the 21st century. Despite significant challenges, Latin America's reformers propose inclusion, respect for human rights, gender equality, interculturality, and recognition of the different. Besides, a few improvements in the governance of educational systems (Saavedra and Gutierrez, 2020). The states carried out reforms; yet, with many of its instruments of power diminished. The weakening of the state's power was either by expanding democracy or following neoliberal policies or authoritarian regimes.

Despite the exciting project of education reform under intergovernmental bodies' guidance, there is no single route for each nation. Several countries in the region undertook ambitious reforms, with different results. In almost all the region's countries, the Constitution or the law establishes that the primary purpose of education is the full development of the human personality, as stated by UNESCO's Education for All. Besides, school systems must promote freedom, peace, social justice, solidarity, cooperation, integration of peoples, and respect for nature. They establish democratic rights and principles, commit to each country's cultural identity or region, intellectual, artistic, physical, intercultural, moral, and civic education. They ban social discrimination, either in the curricula or the classrooms. Still, it seems quite challenging to put into practice such proposals with a centralized curriculum. An overview will illustrate the point (Posner, 2017b; Schwartzman, 2015).

In 2006 the Argentine Congress approved the National Education Law and two critical programs to promote inclusion: Universal Child Allowance, a massive conditional cash transfer program for families in poverty, and Connect Equality (Beech, 2019). In Chile, during Bachelet's second government (2014–18), three laws were enacted after intense debates: Inclusion, New Public Education, and Teacher Professional Development, with them the state tried to correct the adverse effects on the educational system of neoliberal ideology inherited from the dictatorial period, from 1973 to 1990 (Ávalos and Bellei, 2019; Muñoz Stuardo and Weinstein Cayuela, 2019). After some progress since the 1990s, President Evo Morales dismantled democratic aspects of Bolivian education that had guaranteed a diverse and decentralized administration as a necessary means for a multicultural Bolivia, and projected a national curriculum, although maintaining multiculturalism and intercultural approaches (Andrade Ramos and Blanco Cossío, 2015). While in Brazil, the Common Bases of the National Curriculum (2017) promotes the inclusion of marginalized groups, encouraging an egalitarian society that takes plurality in terms of gender, ethnicity, race, language, and others (Ivenicki, 2019). But the Bolsonaro government tries to dismantle its few achievements.

Uruguay and Costa Rica have enviable political systems in Latin America: democratic governments, low corruption, stability, and their population and territories are small; they enjoy vigorous institutions. In 2010, the Uruguayan government designed the National Strategy for Children and Adolescents 2010–1030, including flexible hours in the classroom (Vaillant, 2015). Costa Rica abolished the army in 1948. According to the OECD, Costa Rica is among the countries in the region with the best educational performance; it has made great strides in early childhood education, elementary schooling, and a more significant transition to high school. However, the same OECD recommends diversifying its curriculum to emphasize critical thinking instead of rote memorization (Román Vega, 2019).

In Peru, in the government of Alejandro Toledo, the Congress approved the General Law of Education (2003). It emphasized democratization and decentralization of the education system and proposed a Sectorial Emergency Plan to improve the quality of education. However, the changes are in the air due to political instability (Balarin and Rodríguez, 2019). Likewise, in Ecuador, when the government of Rafael Correa pushed in 2008 for a new Constitution and the National Development Plan: “Good Life” (Buen Vivir) 2009–13, which articulated a popular alternative (Baxter, 2019). While Juan Manuel Santos, the President of Colombia, crowned his effort in 2018, he applied for his country’s entry to the OECD in 2011. One of the requirements involved an evaluation of educational policy, especially to reduce inequality. The 2014–18 National Development Plan expressed 19 strategic lines (Guevara Ramírez and Téllez Rico, 2019).

In Mexico, President Peña Nieto launched an educational reform in 2012. The government and allied parties achieved constitutional amendments, changes in the General Education Law, and two new laws. One granted autonomy to the National Institute for the Evaluation of Education, and the General Law Professional Teaching Service, which instituted a meritocratic system for teachers and administrators. In 2017 the Public Education Secretary launched the “Educational model for compulsory education: educating for freedom and creativity.” However, in 2018, Andrés Manuel López Obrador’s new government buried the reform (Ornelas, 2019).

Other countries tried changes, but institutional inertia or political conditions did not allow them, or the governments did not seek them, but in rhetoric. In Paraguay, after the overthrow of the Alfredo Stroessner dictatorship in 1989, the different governments promote education. The educational system shows a significant increase in enrollment, but poverty, especially among indigenous people and children with disabilities, imposes barriers to the right to education (Brizuela, 2015). In Venezuela, after Hugo Chávez’ rise to power, a process began to institutionalize a parallel educational system to build a communal state: the educational missions, whose main criterion is inclusion, measured by the number of students enrolled. The 2009 Constitution of the Bolivarian Republic of Venezuela draws a path for constructing socialism in the 21st century (Mundó, 2015). Panama is a rare case of sustained economic growth and poor performance of its educational system, according to the Latin American Laboratory for Assessment of the Quality of Education (LLECE) and OECD evaluations due to the low investment in education and the lack of certified teachers (Guardia Wald, 2017). In Nicaragua, during the second Sandinista period — that started in 2006 — despite proposing significant actions in the National Human Development Plan, there are no substantive advances; the Ministry of Education recentralized the educational system (Flóres and Elvir, 2017). Still, in Guatemala, the situation is worse. According to Posner (2017a), there are ghost schools; others abandoned, an inadequate number of teachers with low qualifications. Poor peasants, indigenous of Mayan descent, are the most disadvantaged. The Dominican Republic ranks last the LLECE tests. It frustrates the mood for modernization, despite the government’s efforts to introduce competency-based curricular reform to focus on student learning, increasing instructional time, extending the school day, improving teaching training, and bolstering early childhood education (Hamm and Martínez, 2017).

Other countries have highs and lows in reform and institutionalization of educational projects; yet, there is some progress. Although modestly, El Salvador and Honduras show achievements in improving their school systems. There are a few advances in social inclusion due to full-time schools and subsidies to the poor, especially in rural areas, with textbooks, uniforms, and other supplies. In El Salvador, those improvements are not entirely due to the Ministry of Education, but to social inclusion programs like Education with Community Participation [EDUCO] (Edwards, 2019). Thanks to the support and dissemination of the World Bank, EDUCO became famous as a successful innovation. The government turned it off in 2011; nevertheless, the Bank bets to replicate it in other countries, including Guatemala and Honduras (Ganimian, 2016). In Honduras, despite poverty, violence, and mass migration, reforms are underway, attached to UNESCO’s Education for All. The Secretary of State for Education prepared the Educational Reform Program underway (Luque, 2017).

However, putting prescribed law into practice is a complicated matter. The panorama of the reforms is of prominences and gorges, except in Cuba.

According to international observations, in Cuba, education of high-quality is a universal right. An exceptional state, in many ways. Cuba has a dynamic and egalitarian education system. Since the triumph of the Revolution, in 1959, the government undertook an educational policy called the four revolutions to improve education (Ginsburg and García Batista, 2019). According to Carnoy (2007), the Cuban school’s success is due to the social context that is highly supportive of academic achievement. Besides, the uninterrupted political commitment of national leaders, the strict monitoring of the national curriculum, the supervision and support for teachers, and continued funding, despite the scarcity of public resources.

A conceivable motive for the low educational performance in Latin America may have less to do with embracing the universal reform wrap. It would have more to do with rougher governance issues since several decades of neoliberal reforms or dictatorships undermined states’ ability to successfully implement political changes. It is a matter of political allegiance. Cuba’s academic advantage is due to two leading causes, according to Carnoy (2007): (1) Teachers stress a child-centered and problem-solving approach, which aims the students to develop the ability to work in groups and be creative. (2) The state strictly enforces the curriculum’s

implementation through a chain of command that begins with the Minister of Education and ends with schools' principals supervising teachers in their classrooms. Neither free choice nor school autonomy.

Nevertheless, other of the universal project pillars, centralized learning assessment, is underway in most countries of the region.

Assessment to cope with low quality of education

The discourse on the low quality of education in Latin American countries took supremacy in the globalization era. At the turn of the century, the assessment of learning became the new mantra to raise educational quality. That was a global tool advocated by inter-governmental organizations. Most states established a standardized assessment of learning as a new regulator of instructional practices, which acted on students, schools, and teachers (Rivas, 2015:19). The most enthusiastic promoters were the Latin American Laboratory for Assessment of the Quality of Education and the Organization for Economic Cooperation and Development.

The UNESCO documents that LLECE emerged in the 1990s, where many countries in the region were promoting education reforms with insufficient data for their design. They also lacked a critical mass to measure educational quality. In the 1990s, only a few countries conducted national educational assessments, which were generally not released to the public. In 1997 the Laboratory applied the First Regional Comparative and Explanatory Study (PERCE) among students in third and fourth grades of primary education, reading, and mathematics. LLECE implemented the second comparative study (SERCE) in 2007 and the third (TERCE) in 2013. The tools and results of these last two allowed comparisons over time in students' learning achievements in the third and sixth grades of primary education in the areas of reading and mathematics, as shown in Table 3 (LLECE, 2014).

The natural sciences test was only applied in sixth-grade primary and were voluntary in SERCE. As a result, both studies for comparison purposes are only available for seven countries and the Mexican state of Nuevo León. Although Cuban students participated in the three assessments showed a higher performance than the rest of their peers from the 16 participating countries. However, their results are not comparable, according to LLECE.

Nonetheless, even though only nine countries in the region applied it, the OECD's Program for International Student Assessment (PISA) stood as the superlative measure. Only Brazil and Mexico have participated in PISA since 2000. Argentina did not participate in 2003, and the OECD did not publish its 2015 outcomes. Chile excluded its involvement in 2003. Colombia has partaken the tests since 2006; the Dominican Republic is associated with PISA in 2015 and 2018. Panama joined the assessment in 2009, but in 2012. Peru maintained its students out in 2003 y 2006, and Uruguay joined since 2003. In 2018, the nine countries participated. Table 4 summarizes the historical trend.

PISA has two contradictory drifts, a switch on the traditional modes of communication of achievement, not only enrollment, infrastructure, and curricula, but also what the OECD assumes it is worth learning. On the one hand, it promotes the project of universal reform that tries to ensure that all countries tend to be homogeneous (Beech, 2008) and that "Skills for life," which

Table 3 Student's learning achievements.

SERCE	Mathematics		Reading		Science	TERCE	Mathematics		Reading		Science
	3rd grade	6th grade	3rd grade	6th grade	6th grade		3rd grade	6th grade	3rd grade	6th grade	6th grade
Argentina	505	513	510	506	489	Argentina	533	530	512	509	501
Brazil	505	499	504	520		Brazil	540	520	519	524	
Chile	529	517	562	546		Chile	582	581	571	557	
Colombia	499	493	511	515	504	Colombia	519	515	519	526	527
Costa Rica	538	549	563	563		Costa Rica	558	535	543	546	
Dominican Rep.	396	416	395	421	426	Dominican Rep.	448	437	454	456	444
Ecuador	473	460	452	447		Ecuador	524	513	508	491	
Guatemala	457	456	447	451		Guatemala	501	488	495	489	
Honduras						Honduras	508	480	497	479	
Mexico	532	542	530	530		Mexico	549	566	519	529	
Nicaragua	473	458	470	473		Nicaragua	485	462	478	479	
Panama	463	452	467	472	473	Panama	494	461	490	483	475
Paraguay	486	468	469	455	469	Paraguay	488	456	481	469	455
Peru	474	490	474	476	465	Peru	533	527	521	505	501
Uruguay	539	578	523	542	533	Uruguay	551	567	524	532	517
Average countries	491	492	491.21	494	480	Average countries	522	511	510	507	488
			557.80								
Nuevo León	563	554	558	542	511	Nuevo León	561	587	528	548	539

Note: Data for Honduras is presented, but no comparison with its performance in SERCE is possible as the country did not take part in the second study. Its scores and performance levels in TERCE will be reported using the SERCE scale.

Source: Own Elaboration based in LLECE (2014).

Table 4 PISA scores for Latin American countries.

	2000				2003				2006				2009				2012				2015				2018			
	R	M	S	A	R	M	S	A	R	M	S	A	R	M	S	A	R	M	S	A	R	M	S	A	R	M	S	A
Argentina	418	388	396	401					374	381	391	382	398	388	401	396	396	388	406	397					402	379	404	395
Brasil	396	334	375	368	403	356	390	383	393	370	390	384	412	386	405	401	410	389	405	401	407	377	401	395	413	384	404	400
Chile	410	384	415	403					442	411	438	430	449	421	447	439	441	423	445	436	459	423	447	443.0	452	417	444	438
Colombia									385	370	388	381	413	381	402	399	403	376	399	393	425	390	416	410.3	412	391	413	405
Dominican Republic																					332	358	328	339.3	342	325	336	334
Mexico	422	387	423	411	400	385	405	397	410	406	410	409	425	419	416	420	424	413	415	417	423	408	416	415.7	420	409	419	416
Panama													371	360	376	369					332	358	328	339.3	377	353	365	365
Peru	327	292	333	317									370	365	369	368	384	368	373	375	398	387	397	394.0	401	400	404	402
Uruguay					434	422	438	431	413	427	428	423	426	427	427	427	411	409	416	412	437	418	435	430.0	427	418	426	424
Average	500	500	500		494	500	500		492	498	500		493	496	501		496	494	501		493	490	493		487	489	489	

R = Reading; M = Mathematics; S = Science; A = Average.

Source: Own elaboration with the help of Diego Amado De Leon and Eduardo López Guadarrama: OECD, (2019). PISA Databases 2000, 2006, 2009, 2012, 2015, 2018: <https://www.google.com/search?q=pisa+dataset&oq=PISA+data&aqs=chrome>.

includes the teaching of language, mathematics, and science, become the cornerstones of the education through K12. On the other hand, it hierarchizes the countries when preparing rankings according to their students' results in those tests' performance.

As in the rest of the world, PISA sparks debates; it has advocates beyond governments and the OECD itself and critics who point out its irrelevance to education because it focusses on a few subjects. Ravela (2011) sees two advantages. First, PISA forces to leave the endogenous reference to analyze Latin American countries' educational situation in a broader context. Also, it helps to understand the overall situation better. Second, participation in international studies represents an opportunity for learning and accumulating knowledge for national evaluation institutions.

At the other extreme, as summarized by Antonio Bolívar (2011), it produces a feeling of failure: "The population is disconnected or dissatisfied with appearing in the successive PISA Reports, which help to determine their respective policies as 'losers' in comparison with other countries." And, what is critical in terms of social perceptions, is that although PISA contains much more than test results, the "[...] news media prioritizes the results of the rankings, reinforcing the dichotomy of 'winners' and 'losers,' between countries that offer a good education and others that provide a mediocre or poor one."

The OECD sees itself as a promoter of equity and good governance in education (Melchor, 2008). In comparison, PISA critics note that in addition to offering its technical expertise in measuring learning and schooling performance, the OECD also became an internal political actor (Lingard et al., 2016). It also wields power and pushes the technocratic approach to rule by numbers (Grek, 2009). Certainly, PISA provokes new institutional arrangements in many countries.

Participation in PISA caused discomfort for the results, but it also generated protests from academics and journalists to governments for the low quality of education; SERCE and TERCE did not make as much debate. Nonetheless, the tendency to evaluate education had institutional consequences. In much of Latin America, governments established national institutes to measure the quality of education and students and teachers (Bruno et al., 2019). Standardized international and domestic assessments led to rejection by the teachers' unions. They alleged that high ranking functionaries and politicians blamed educators for low test scores of all kinds and other ills in education.

As in other parts of the world, in Latin America, certain segments of the governing groups criticized teachers for schools' deficiencies and, at the same time, pinned hopes for substantive change on teachers. UNESCO, the World Bank, the OECD, and other intergovernmental organizations proposed reforms emphasizing effective schools, principal's leadership, student achievement, and dynamic forms of governance to raise the quality of education. To achieve any of these purposes, the professionalization of teachers became a required subject.

Teacher professionalization, identity, and collective representation

Although the nation-states were not yet fully consolidated, liberal governments began to build schools. In Chile, Argentina, Mexico, Peru, and Colombia, they looked to France for a model to follow: a system under state control. However, governments lacked assets and personnel to teach. As in the schools of the Lancaster-Bell system, they used upper-grade students to help the few teachers and governesses; these students became instructors by empirical means and were among the first teachers in normal schools to train preceptors from the 1820s.

The modern normal schools began their institution from the 1850s (Weinberg, 1995), molded in the French style. Being graduated from a teacher college was the beginning of the normalista identity. Identity includes the notion of equality, permanence, continuity, durability, a fixed human attribute with which one is born, or has as a destiny (Navarrete Cazales, 2016); it is collective and individual. The collective identity stresses the idea of equality, the individual, of merit; the collective generates shared beliefs,

symbols, and rituals that give a group a particular distinction. The normalista identity is one of the most deeply rooted characteristics in the educational systems of Latin America.

These community attributes combined well with the formation of school systems under the state's tutelage and, therefore, facilitated unions' institution to represent the working benefits of teachers first and then their professional interests. In its origins, as early as the 20th century, teachers' unions distinguished themselves by defending the rights of their members: fighting for better wages, job security, recognition for seniority, and fringe benefits such as health and housing. They also advocated improving their schools and social acceptance of their work (Tedesco and Fanfani, 2006). However, they had fractures regarding their professional identity.

Throughout the 20th century, there were debates to characterize teachers' work; authors disputed whether their labor was a trade (like that of a master craftsman), a profession, or a salaried job (Tenti Fanfani, 2006). A widespread belief among public school teachers is that teaching was a state profession, ergo, that they were civil servants (but not bureaucrats). Some thought that having a bachelor's degree was the attribute of the trade. And the teachers' unions reinforced that collective idea. Following the international educational reform project, the changes of recent years refute such a trend (Ornelas, 2018). But teachers' unions are more than representative union bodies; they are also interest groups, like in Chile, Brazil, and Argentina, or robust political machines, like in Mexico and Ecuador (Schneider, 2021), and represent the opposition to neoliberal reforms (Loyo, 2019). Other authors point to the positive impact that unions have on instruction and a teachers' sense of purpose (Gindin and Finger, 2013).

Almost all the authors agree that teachers' unions play a significant role in education policymaking in Latin America. That is because unionized public sector workers can process information that affects the policy arena under their influence and can monitor government activities. Through mandatory fees from their members, unions often have access to vast financial resources. Federations with large membership can also use strikes and protests politically in the streets to advance their demands (Bruns and Luque, 2015). The extreme case of influence is that of the National Union of Education Workers (SNTE) of Mexico, with around 1,400,000 members. Through their affiliation to the Institutional Revolutionary Party, since the founding of the union in 1943, and later in an alliance with the National Action Party (2000–12), the leaders controlled teachers' labor trajectory from entering the schools to their retirement (Chambers-Ju and Finger, 2016). The politics of patronage reached the extreme that teachers who retired could inherit their jobs or, if they did not have teacher heirs, sell them to the highest bidder (Ornelas, 2018). Corruption reached the point that there were ghost schools and teachers and more than 100,000 deceased and retired people in the payroll (H.T., 2014).

There are many teacher unions in Latin America (Table 5) that, although have differences between them, defend the traditional identities attached to rote-memory, collective action for recruitment, and certification by the state. The modernizing reformers thought that such attributes were a cause—not that it was exclusive—of the low quality of education. For that reason, the 21st century reforms incorporated critical legislative changes to professionalize teachers (which implies that they were not professionals).

Elacqua et al. (2018) discuss how the laws of the Teacher Professional Development System in Chile (2016), the Statute of Teacher Professionalization in Colombia (2002), the Organic Law of Intercultural Education in Ecuador (2011), the General Law of the Professional Teacher Service in Mexico (2013), and Law of Teacher Reform in Peru (2012) seek to make the teaching profession more attractive to talented young people. Also, to improve future teachers' training, that implies reforming the traditional schools—the crucial aspect: selecting the best candidates for the profession, which requires rigorous examinations. And, in return, supporting new teachers with tutoring. The current Mexican government repealed the teachers' law in 2019 and, although it does not mean a return to the past, it substantially recovers the normalista spirit; the SNTE recovered a good part of the power that the Peña Nieto government (2012–18) wanted to take from it (Ornelas, 2019).

The core of the teaching profession's reform purposes is to accentuate individual identity, put merit as the primary attribute, knowledge, and practices appropriate for the present and the future. That includes using more information and communication technologies. Changes in recruitment methods, decision-making skills (personal initiative and open mind), and a new type of work ethic, as proposed by UNESCO, the World Bank, and the OECD.

Covid-19 and the right to education

In Mexico, there is a saying. "It rains over wet." It means that if something is wrong, an external event aggravates it. And that happens to educational systems with Covid-19. The Latin American Campaign for the Right to Education (CLADE, by its initials in Spanish) expressed its concern that the pandemic deepens historical inequalities. In Latin America and the Caribbean, 74% of all enrolled students were affected by schools' closure (Croso, 2020). For millions of children, adolescents, and young people, especially those in situations of greater vulnerability, the shutting of schools implies the temporary loss of studies and the stoppage of dialogs between teachers and students. Furthermore, it injures a fundamental safety net, including nutrition, health, emotional support, and protection against domestic violence and sexual abuse.

This article comes to an end in August 2020. There is still not enough academic work to substantiate the consequences of Covid-19 on education. Still, the press and media report that the disease aggravated pre-existing problems and increased economic and social inequality, femicide and sexual assault, and discrimination based on ethnic origin, gender, or belief.

Many governments responded to the emergency of Covid-19 with promptness through distance education, as in Argentina, Chile, Colombia, Mexico, Peru, and Uruguay. They even made efforts to ensure that the programs reached the most significant number of inhabitants. Still, children from low-income families—the majority indigenous or Afro-descendants—do not have, or

Table 5 Teachers' unions in Latin America, unity, and fragmentation. Selected countries by the availability of data, as of 2010.

Country	Teachers union(s)	Unity density	Fragmentation
Argentina	Confederación de Trabajadores de la Educación de la República Argentina - Confederation of Education Workers of the Argentine Republic (CTERA)	50.8% (234 000)	High fragmentation (effective coordination through CTERA)
Brazil	Confederação Nacional dos Trabalhadores em Educação National Confederation of Education Workers (CNTE)	44.2% (925 229)	High fragmentation (partial coordination through CNTE)
Chile	Colegio de Profesores - Collegiate of Teachers	53.3% (71 982)	Monopoly of representation
Colombia	Federación Colombiana de Educadores - Colombian Federation of Educators	81.6 (as of 2007)	Monopoly of representation
Costa Rica	Asociación Nacional de Educadores - National Association of Educators (ANDE); Asociación de Profesores de Segunda Enseñanza - Second School Teachers Association (APSE); Colegio de Licenciados y Profesores en Letras, Filosofía, Ciencias y Artes - Collegiate of Graduates and Professors in Letters, Philosophy, Sciences and Arts (COLYPRO); Sindicato de Trabajadores de la Educación Costarricense - Union of Costa Rican Education Workers (SEC)	100% (membership estimated 140,000–200,000)	Fragmentation by education level
Ecuador	Unión Nacional de Educadores - National Union of Educators (UNE); Frente Unionista de los Trabajadores de Educación del Ecuador - Unionist Front of Education Workers of Ecuador (FUTE)	79%–90% for UNE during 1990's, 79% first decade of 20th century	Monopoly of representation
Honduras	Colegio Profesional para la Superación Magisterial de Honduras - Professional Collegiate for Teacher Improvement of Honduras (COLPROSUMAH); Primer Colegio Profesional Hondureño de Maestros - First Honduran Professional Collegiate of Teachers (PRICPHMA); Colegio de Profesores de Educación Media de Honduras - Collegiate of Teachers of Secondary Education of Honduras (COPEMH)	Mandatory affiliation	Fragmented
México	Sindicato Nacional de Trabajadores de la Educación - National Union of Education Workers (SNTE)	100% (between 1200 000 and 1, 500,000 members)	Monopoly of representation
Paraguay	Federación de Educadores de Paraguay - Federation of Educators of Paraguay (FEP), Organización de Trabajadores de la Educación del Paraguay - Organization of Education Workers of Paraguay (OTEP); Unión Nacional de Educadores - National Union of Educators (UNE); Agrupación de Docentes y Funcionarios de la Educación Paraguaya - Association of Paraguayan Education Officials and Teachers (ADOFEF); and Federación de Educadores de Capital e Interior Federation of Educators of Capital and Interior (FECI).	55% (15 000; 14 000; 15 000; information on The membership for ADOFEF and FECI is not available)	Fragmented
Peru	Sindicato Único de Trabajadores de la Educación del Perú - Unique Union of Education Workers of Peru (SUTEP); Sindicato de Docentes de Educación Superior del Perú - Union of Higher Education Professors of Peru (SIDESP)	41.9% between 145,000 and 200,000)	Monopoly of representation
Uruguay	Federación Uruguaya de Magisterio – Trabajadores de la Enseñanza - Uruguayan Federation of Teaching - Teaching Workers (FUM-TEP) for primary education; Federación Nacional de Profesores de Enseñanza Secundaria - National Federation of Secondary Education Teachers (FENAPES)	77% for primary schools	Fragmented by education level

Source: Fernández M.A., 2012. The Political Challenges in Pursuing an Agenda for Quality of Education in Latin America. Background paper for Latin America Regional Study on Teacher Quality, World Bank, Washington, DC.

is scarce, access to computers and the Internet; they do not even have electricity in many places. The contagion widened the digital divide. CLADE underlines that it highlighted the dramatic consequences of the precariousness of public services that so many governments maintain in Latin America and the Caribbean: "There is no doubt that the pandemic and the closure of schools will have an impact on the realization of the right to education in the short, medium, and long term."

Ending

Covid-19 contained, at least for a few more years, the wave of educational reforms that reached Latin America as an imported good of the globalization. At the end of the 1980s, global intergovernmental organizations inserted themselves into the political life of nations. They demanded educational reforms, which many governments accepted willingly, and others forced by the context of dependency. Reform drives have removed specific attributes of the educational systems formed in the 19th century, but—maybe—not their fundamental matrix.

Although organizations such as UNESCO, the World Bank, and the OECD promote the educational reform's universal project, it is not homogeneous. UNESCO is faithful to its humanistic origin. It encourages changes under the motto that education is a fundamental human right, that its primary purpose is to develop the faculties of the human being harmoniously. Native educational streams of Latin America, such as popular and intercultural education, or the Southern epistemologies promote hundreds of small-scale projects that pursue similar purposes. Since the Jomtien Conference on Education for All, that mandate is now in bylaws and even constitutions.

Furthermore, all the region's countries signed the Millennium Goals and endorsed the Inchon Pact on September 25, 2015. The area leaders adopted a set of global goals to eradicate poverty, protect the planet, and ensure prosperity for all as part of a new sustainable development agenda. It assumes that democracy is the political basis for sustainability. Each objective has goals that must be achieved by 2030. Goal 4 proposes that governments should establish quality, inclusive, and equitable education.

The language of the World Bank and the OECD also includes those words as purposes. However, the approach is different; the emphasis is on preparing a mass of laborers to acquire productive work skills in a global world. The concept of long-life education, for example, puts skills for life as the critical tool, with productivism or neoliberal approaches. At the same time, humanists argue that it should serve to forge the character of people.

Those are two different bets so that, through education, countries achieve progress in social inclusion, democratization, and sustainable development. However, together with political authoritarianism, inequality and social segregation are firmly established in Latin America's reality. It is a historical heritage that is difficult to eradicate and, perhaps, globalization will not solve such legacy, but rather exacerbate it.

Yet, it is an obligation of teachers, scholars, and democratic organizations of civil society to continue the struggle at the local and national scales for a quality, humanistic education that prepares the largest possible number of children and adolescents for a world increasingly complex.

References

- Andrade Ramos, F.d., Blanco Cossío, M., 2015. In: Schwartzman, S. (Ed.), *Bolivia: Perspectives and Challenges for Multicultural Education*, Education in South America. Bloomsbury, London (Kindle).
- Arnone, R.F., 2013. Introduction: reframing comparative education: the dialectic of the global and the local. In: Arnone, R.F., Torres, C.A., Franz, S. (Eds.), *Comparative Education: The Dialectic of the Global and the Local*, fourth ed. Roman & Littlefield Publishers, Lanham, MA, pp. 1–24.
- Ávalos, B., Bellei, C., 2019. Recent education reforms in Chile: how much of a departure from market and new public management systems. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 43–71.
- Balarin, M., Rodríguez, M.F., 2019. Endurance and absences in Peru's reform: the challenge of second-order reforms in the core of educational practice. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 116–133.
- Baxter, J., 2019. The paradox of power in Ecuador: governance and education reform (2007–2015). In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 134–165.
- Beech, J., 2008. The institutionalization of education in Latin America: loci of attraction and mechanisms of diffusion. In: Baker, D.P., Wiseman, A. (Eds.), *The Impact of Comparative Education Research on Institutional Theory*. Emerald, Wagon Lane, pp. 281–303.
- Beech, J., 2019. The long and winding road to inclusion: educational policies in Argentina (2003–2015). In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 18–42.
- Bellei, C., Vanni, X., 2015. In: Schwartzman, S. (Ed.), *Chile: The Evolution of Educational Policy*, Education in South America. Bloomsbury, London, pp. 1980–2014 (Kindle).
- Beltrán Morales, V.M., 2015. Modelo de sistematización de experiencias de aprendizaje del CEDECOSU, del CREFAL, Decisio: Saberes para la Acción en Educación de Adultos, pp. 59–61, 40, January–April 2015.
- Bergés, A., 2009. Vertical and Horizontal Inequalities in Human Capital: Education Attainment and Literacy in Latin America During the Twentieth Century. Mini-Conference: A Comparative Approach to Inequality and Development: Latin America and Europe. Fundación Ramón Areces and Instituto Figuerola, Madrid. May 8–9.
- Bolívar, A., 2011. The dissatisfaction of the losers: PISA public discourse in Ibero-American countries. In: Pereyra, M.A., Kotthoff, H.-G., Cowen, R. (Eds.), *PISA under Examination: Changing Knowledge, Changing Tests, and Changing Schools*. Sense Publishers, Boston, pp. 61–74.
- Brizuela, C., 2015. In: Schwartzman, S. (Ed.), *Paraguay: An Overview*, Education in South America. Bloomsbury (Kindle).
- Bruns, B., Luque, J., 2015. Great Teachers: How to Raise Student Learning in Latin America and the Caribbean. The World Bank, Washington, DC.
- Bruns, B., Akmal, M., Birdsall, N., 2019. The Political Economy of Testing in Latin America and Sub-Saharan Africa. Center for Global Development, Washington, DC.
- Carnoy, M., Cosse, G., Cox, C. (Eds.), 2005. Las reformas educativas en la década de los 90: Un estudio comparado de Argentina, Chile y Uruguay. Ministerio de Educación, Ciencia y Tecnología, Buenos Aires.

- Carnoy, M., 2007. Cuba's Academic Advantage: Why Students in Cuba Do Better in School. Stanford University Press, Stanford, CA.
- Carnoy, M., 2012. Does More and Better Education Equalize Opportunity and Income in the Knowledge Economy? Cátedra Globalización y Democracia, Universidad Diego Portales, Santiago.
- Chambers-Ju, C., Finger, L., 2016. Teachers unions in Mexico: the politics of patronage. In: Moe, T. (Ed.), *Teachers Unions and Education Systems across the World*. Cambridge University Press, Cambridge, UK, pp. 216–239.
- Córdova, A., 1977. Los orígenes del Estado en América Latina. *Cuad. Polit.* 14, 23–43.
- Cortina, R., Stromquist, N. (Eds.), 2000. *Distant Alliances: Promoting Education of Girls and Women in Latin America*. Routledge Falmer, Nueva York.
- Cortina, R. (Ed.), 2017. *Indigenous Education Policy, Equity, and Intercultural Understanding in Latin America*. Palgrave MacMillan, New York.
- Croso, C., 2020. COVID-19 in Latin America and the Caribbean. Goba Campaign for Education. August 1: <https://www.campaignforeducation.org/en/2020/05/19/covid-19>.
- ECLAC, 2001. *Statistical Yearbook for Latin America and the Caribbean*. Economic Commission for Latin America and the Caribbean, Santiago.
- Edwards Jr., D.B., 2019. Education in El Salvador: recent political economy of education policy. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 190–214.
- Edwards Jr., D.B., 2020. False advertising and first principles: misrepresentation of research evidence from El Salvador and the need to change the paradigm of community-based management. In: Nishimura, M. (Ed.), *Community Participation Withschools in Developing Countries:towards Equitable and Inclusive Basiceducation for All*. Routledge, London, pp. 200–212.
- Elacqua, G., Hincapié, D., Vegas, E., Alfonso, M., 2018. *Profesión profesor en América Latina: ¿Por qué se perdió el prestigio docente y cómo recuperarlo?* BID, New York.
- Flóres, C.A., Elvir, P., 2017. Nicaragua: education for sustainable human development. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 215–237.
- Ganimian, A.J., 2016. Why do some school-based management reforms survive while others are reversed? The cases of Honduras and Guatemala. *Int. J. Educ. Dev.* 47, 33–46.
- García Benavente, J.F., 2015. La escuela Lancasteriana en México y en América Latina como solución del estado liberal ante el vacío dejado por la Iglesia. *Boletín Redipe* 4 (7), 48–66.
- Gindin, J., Finger, L., 2013. *Promoting Education Quality: The Role of Teachers' Unions in Latin America* (Paris).
- Ginsburg, M.G., García Batista, G., 2019. Education and teacher education in Cuba. revolución y perfeccionamiento. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 215–243.
- Grek, S., 2009. Governing by numbers: the PISA 'effect' in Europe. *J. Educ. Pol.* 24 (1), 23–37.
- Guardia Wald, R.d., 2017. Panama: an overview. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 238–262.
- Guevara Ramírez, R., Téllez Rico, S.M., 2019. From the indicative to the imperative: Colombia, the most educated in 2025? In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 166–189.
- Guimarães de Castro, M.H., 2015. In: Schwartzman, S. (Ed.), *Brazil: The Role of States and Municipalities in the Implementation of Education Policies*, Education in South America. Bloomsbury, London (Kindle).
- H.T., 2014. Education in Mexico: Phantom Teachers. *The Economist*. April 7 2014.
- Halperin Donghi, T., 1969. *Historia contemporánea de América Latina*. Alianza Editorial, Madrid.
- Hamm, M., Martínez, R., 2017. Dominican republic: an overview. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 283–305.
- Ivenicki, A., 2019. Education reform in Brazil: multicultural reflections. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 101–115.
- Kubal, M.R., 2003. *The Politics of Education Decentralization in Latin America: Rhetoric and Reality in Chile, Mexico, Argentina, and Nicaragua*, 2003 Meeting of the Latin American Studies Association. Dallas, TX, March 27–29.
- Latapi, P., 1993. Reflexiones sobre la justicia en educación. *Rev. Latinoam. Estud. Educ.* XXIII (2), 9–41.
- Lingard, B., Martino, W., Rezai-Rashti, G., Sellar, S., 2016. *Globalizing Educational Accountabilities*. Routledge, New York.
- LLECE, 2014. *Serce and Terce 2006-2013*. Latin American Laboratory for Assessment of the Quality of Education-UNESCO Regional Bureau for Education in Latin America and the Caribbean, Santiago.
- Loyo, A., 2019. Teacher unions and education reforms: the 90s in Argentina, Brazil, and Mexico. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 272–292.
- Luque, J., 2017. Honduras: contemporary education reform. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 191–214.
- McEwan, P.J., Carnoy, M., 2000. The effectiveness and efficiency of private schools in Chile's voucher system. *Educ. Eval. Pol. Anal.* 22 (3), 213–239.
- Medina Echaverría, J., 1964. *Consideraciones sociológicas sobre el desarrollo económico de América Latina*. Editorial Solar, Buenos Aires.
- Melchor, O.H., 2008. *Managing Change in OECD Governments: An Introductory Framework*, vol. 12. OECD Publishing, Paris.
- Mundó, M., 2015. In: Schwartzman, S. (Ed.), *Venezuela: Political and Institutional Transition, 1999–2014*, Education in South America. Bloomsbury, London (Kindle).
- Muñoz Stuardo, G., Weinstein Cayuela, J., 2019. The difficult process in Chile: redefining the rules of the game for subsidized private education. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 272–279.
- Navarrete Cazales, Z., 2016. Investigación sobre formación e identidad profesional de los docentes en México. *Revista de Educación Superior del Sur Global* 1 (1), 5–26.
- OECD, 2017. *OECD Skills Outlook 2017*. OECD, Paris.
- OECD, 2019. PISA databases. OECD, Paris. <https://www.google.com/search?q=pisa+dataset&oq=PISA+data&aqs=chrome>.
- Ornelas, C., 1995. *El sistema educativo mexicano: la transición de fin de siglo*. Fondo de Cultura Económica, Mexico City.
- Ornelas, C. (Ed.), 2005. *Buenas prácticas de educación básica en América Latina*, 2 vols. Consejo Empresario de América Latina – Instituto Latinoamericano de la Comunicación Educativa, México.
- Ornelas, C., 2018. *La contienda por la educación: globalización, neocorporativismo y democracia*. México: Fondo de Cultura Económica.
- Ornelas, C., 2019. Hills and valleys of the Mexican reform: politics and contentions. In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Boston, pp. 245–271.
- Ossenbach Sauter, G., 1993. Estado y educación en América Latina a partir de su independencia (siglos XIX y XX). *Revista Iberoamericana de Educación* 1 (1). Available online: <https://rieoei.org/historico/oeivirt/rie01a04.htm>.
- Posner, C.M., 2017a. Guatemala: a failing system. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 167–190.
- Posner, C.M., 2017b. Introduction: regional overview. In: Posner, C.M., Martin, C., Elvir, P. (Eds.), *Education in Mexico, Central America and the Caribbean*. Bloomsbury, London, pp. 1–25.
- Post, D., 2002. *Children's Work: Schooling and Welfare in Latin America*. Westview Press, Boulder.
- Ravela, P., 2011. *¿Qué hacer con los resultados de PISA en América Latina?* PREAL, Santiago.
- Reimers, F. (Ed.), 2000. *Unequal Schools, Unequal Chances: The Challenge of Equal Opportunity in the Americas*. The David Rockefeller Center Series on Latin American Studies/Harvard University, Cambridge, MA.
- Rivas, A., 2015. *América Latina después de PISA: lecciones aprendidas de la educación en siete países (2000-2015)*. Fundación CIPPEC, Buenos Aires.

- Rivero, J., 1999. Educación y exclusión en América Latina: reformas en tiempos de globalización. Miño y Dávila Editores, Buenos Aires.
- Rodríguez Brandao, C., 1977. El método Paulo Freire para la alfabetización de adultos Pátzcuaro. CREFAL, MX.
- Román Vega, I., 2019. Estado de la educación en Costa Rica: principales hallazgos. Colopro, San José.
- Saavedra, J., Gutierrez, M., 2020. Peru: a wholesale reform fueled by an obsession with learning and equity. In: Reimers, F.M. (Ed.), *Audacious Education Purposes: How Governments Transform the Goals of Education Systems*. Springer International Publishing, Cham, pp. 153–180.
- Schneider, R.B., 2021. Teacher unions, political machines, and the thorny politics of education reform in Latin America. *Polit. Soc.* <https://doi.org/10.1177/00323292211002788>.
- Schwartzman, S., 2015. Introduction: regional overview. In: Schwartzman, S. (Ed.), *Education in South America* (London: Bloomsbury, Kindle Edition).
- Sen, A., 2009. La idea de la justicia. Santillana, Mexico City.
- Sevilla, B., 2017. Panorama de la educación técnica profesional en América Latina y el Caribe. CEPAL- Ministerio de Asuntos Exteriores de Noruega, Santiago de Chile.
- Tedesco, J.C., Fanfani, E.T., 2006. Nuevos tiempos y nuevos docentes. In: Sindicato Nacional de Trabajadores de la Educación. IV Congreso nacional de educación: documentos de discusión. SNTE, Mexico City, pp. 57–81.
- Tenti Fanfani, E. (Ed.), 2006. El oficio de docente. Siglo XXI Editores Argentina, Buenos Aires.
- Torres Bodet, J., 2017. *Memorias II: el desierto internacional/la tierra prometida/Equinoccio*. Fondo de Cultura Económica, Mexico City.
- Treviño, E., 2005. Estado del arte. desigualdad y discriminación en educación: una mirada para América Latina. In: Hevia, R. (Ed.), *Políticas educativas de atención a la diversidad cultural: Brasil, Chile, Colombia, México y Perú*. Oficina Regional de Educación de la UNESCO para América Latina y el Caribe, Santiago de Chile, pp. 23–61.
- UNESCO Institute for Lifelong Education, 2012. Observatory website for Adult and Young Education for Latin America and the Caribbean. UNESCO Institute for Lifelong Education, Mexico City.
- UNESCO, 1993. *Education for All: Status and Trends*. UNESCO, Paris.
- United Nations Organization, 2016. *The Sustainable Development Goals Report 2016*. UNO, New York.
- Vaillant, D., 2015. In: Schwartzman, S. (Ed.), *Uruguay: The Teachers' Policies Black Box*, Education in South America. Bloomsbury, London (Kindle).
- Weinberg, G., 1995. *Modelos educativos en América Latina*. UNESCO-CEPAL-PNUD- A Z Editora, Buenos Aires.
- Winkler, D.R., 1989. *Decentralization in Education: An Economic Perspective*. World Bank Group, Washington, DC.

Asia-Pacific-realist and imaginary constructions

Kerry J. Kennedy, Department of Curriculum and Instruction, The Education University of Hong Kong, Tai Po, Hong Kong SAR, PR China
The University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Multiple ways of understanding the Asia-Pacific	412
Asia-Pacific – a realist perspective	412
Asia-Pacific – alternative constructions	412
Constructions of education within the multiple geographies of the Asia-Pacific	414
International assessments and the education divide	414
Civic and social education	415
Asia-Pacific geographies as pathways for student mobility	416
Conclusion	417
References	417

There are multiple ways to consider the Asia-Pacific, including the fact that for some it is no longer relevant. It would be easier to adopt a realist approach to definition, but as shown below such an approach underestimates not only the complexity of geographies, social systems and politics but also fails to interrogate the idea that underlies the construct of Asia Pacific. As the Australian historian, [Spate \(1977\)](#) commented, with reference to the Pacific element of the construct, “the outsiders named the Ocean, gave it bounds, and in a not trivial sense ‘the Pacific’ as a concept is a Euroamerican artifact. But it was not just ‘made in Europe’; the local materials were not inert.” (p. 211). Apart from providing a history of how what was called by some the “South Sea” became the “Pacific”, [Spate \(1977\)](#) also pointed out that the name change was linked to concerns of European traders over centuries (p. 210). Naming, therefore, is not necessarily, or even primarily, a simple realist action. Rather, naming can be symbolic of other meanings and intentions, some might say an act of imagination. In what follows three broad themes will be explored:

- Multiple ways of understanding the Asia-Pacific;
- Constructions of education within these multiple geographies; and
- Multiple geographies as pathways for student mobility

Multiple ways of understanding the Asia-Pacific

Asia-Pacific – a realist perspective

The United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP) provides a good example of a realist perspective on the Asia-Pacific. It unambiguously identifies five major sub-regions making up the Asia Pacific: East and North Asia, North and Central Asia, South and South East Asia, South and South West Asia and the Pacific ([UNESCAP, 2020](#)). This is largely an administrative construction of the region that complements other regional groupings within the United Nations (UN) system. Such groupings, for example, are used to report progress on the UN’s Sustainable Development Goals (SDG). The information provided is limited, but not unimportant, as shown below.

In terms of SDG 4 related to Quality Education the 2020 report showed that three of the sub-regions were making reasonable progress (East and North Asia, North and Central Asia, South and South East Asia), one was making moderate progress (South and South West Asia) while one was making little progress at all (the Pacific). There was a different pattern on SDG 5 related to Gender Equality. Two were making moderate progress (the Pacific and South and South East Asia) while the remainder were making very little progress. In relation to SDG 10, Reduced Inequalities, only North and Central Asia was making reasonable progress, South and South East Asia was actually regressing while the remainder were making moderate progress. More detailed and fine-grained assessments are also provided, but they provide limited insights. Nevertheless, what stands out is the diversity across the region, but with no explanation of why such diversity exists or how it might be remedied. This is an important issue in the context of a realist Asia-Pacific. Yet that is not the only way to consider the region and the education and social developments embedded within it.

Asia-Pacific – alternative constructions

In contrast to the realist perspective, [Dirlik \(1992\)](#) argued that:

In a fundamentalist sense, there is no Pacific region that is an “objective” given, but only a competing set of ideational constructs that project upon a certain location on a globe the imperatives of interest, power or vision of these historically produced relationships (p. 56).

Dirlik's (1992) subsequent narrative argued that the Asia-Pacific region would be better named "the Euro-American Pacific" (p. 59). This is a reflection that the region, no matter how it is described geographically, has little to do with the people who actually inhabit it, their needs, wishes or priorities. Rather, it is about the way those outside the "region" see its value in trading and commercial terms. At one time both Asia and the Pacific were the focus of European imperialism when naked force was used to exert the will of the invaders. In post-colonial times the will for domination has not disappeared it has simply re-emerged in a different form. Even a country like the United States, itself born out of colonial conflict, emerged in the nineteenth century with its own colonialist agenda driven by what was seen as economic needs and imperatives. As Dirlik's (1992) pointed out "this structure expressed the demands of the nascent capitalist order that informed the new vision, rather than expressing needs intrinsic to the region" (p. 66). This broader issue has been highlighted recently by Chua (2021) who, in writing about "the monstrosities of maritime capitalism" made the point that "unique among other spaces of capitalist production, the sea is both passageway and precondition for capitalism to conquer new frontiers."

Both Dirlik (1992) and Spate (1977) stressed the historical contexts that led to the construction of a region driven by external forces rather than by those who actually lived in it. Toward the end of the twentieth century, however, there were new attempts to reconstruct the region initially for economic reasons but subsequently for security concerns.

The first of these was a forum for Asia Pacific Economic Cooperation (APEC), an exclusivist organization set up at the instigation of Australian politicians to take advantage of the growing economies of East Asia. It reflected late twentieth century neo-liberalism and a continuing colonial mentality toward both Asian and Pacific Ocean participants. Its purpose was "to preserve and extend a process of market integration among rapidly growing economies and their major regional partners" (Garnaut, 1999, pp. 2–3). Initial membership included Korea, Japan and ASEAN members Australia and New Zealand, the United States and Canada. Within a few years Hong Kong, Taiwan and China were added. The criterion for inclusion was that "these were economies ... were committed to internationally-oriented growth and conducted a high proportion of their trade with other Asia Pacific economies" (Garnaut, 1999, pp. 4–5). A number of other countries joined at a later stage (Vietnam, Russia, Chile, Peru). The most notable exclusion was the Pacific Islands that were seen to lack capacity with Papua New Guinea being the exception. Thus "the Pacific" in the context of APEC was Australia, North America and the three Latin American countries with the entire area of the Pacific Ocean excluded. This was neoliberalism's ideology for a twentieth century "Asia-Pacific" grounded firmly in the way developed economies constructed their needs and excluding all those who were judged as incapable of contributing.

Even though there were exclusions from APEC in order to create a set of countries of equal economic capacity and potential this did not mean it was an organization of equals. As Ravenhill (1999) pointed out, APEC's agenda was transformed in part by pressure from North American and Oceanic countries which saw APEC as another forum for exerting leverage to open up East Asian markets (p. 232). Societies in East Asia such as South Korea, Taiwan and Hong Kong became known as of the four "Asian tigers" (p. 33), the fourth one being Singapore (Castells, 1992). They managed to "maintain ... the highest rate of GNP growth in the world over 25 years, and won substantial shares of the world market economy" (Castells, 1992, p. 34). Countries such as the United States, Australia and New Zealand, former colonies themselves, adopted what seemed like a very colonial attitude to East Asia (Sidaway, 2002):

In the 1980s and 1990s, a sense of rising Asian confidence and development produced a desire for engagement/response within Australia, New Zealand, Canada and the USA. In the 1980s, in the context of a discourse of Asian miracles and dynamism, these "old" colonial-settler states became anxious about their prospect (p. 205).

Thus, APEC's construction of the "Asia-Pacific" can be seen as an attempt by these settler societies to keep the "Asian tigers" in check, or at the very least to share in their expansion and growth. This neoliberal idea continued, but it was not the only idea that sought to use the region for its own purposes.

If the late twentieth century saw the Asia-Pacific constructed in neoliberal terms, the 21st century's focus has been on security. The geography of the new security entity has been referred to as the Indo-Pacific, bringing together the two oceans with ASEAN sitting almost at the fulcrum of this giant entity. Medcalf (2015) argued that the idea of this two oceans strategy had its origins in Australian defense planning. Yet Choong (2019) traced the idea to exchanges between Japan and India as far back as 2007 when Japanese Prime Minister Shinzo Abe made an impassioned call for a "dynamic coupling" of the Indian and Pacific Oceans as "seas of freedom and of prosperity." Soon after the Quadrilateral Security Dialog, subsequently referred to as 'the QUAD', especially by the Biden administration, consisted of the United States, Australia, Japan and India, engaged in a rare meeting. Tellis (2020) traced US thinking about the Indo-Pacific to earlier foreign policy experience in the region. Thus when in 2019 the US announced its Indo-Pacific Strategy (IPS) Report, it was little more than a restatement of these earlier attempts to focus on these parts of the world. Yet there was a new element. Wanandi (2019) pointed out that, "Quad 2 can be seen as the implementation of the US' national strategy for containing China" (p. 55). Denisov et al. (2021), on the other hand, provided a more nuanced analysis suggesting that entities like ASEAN and countries like Japan were more multilateral in their orientation and did not necessarily endorse the anti-China elements of the US position. This view was supported by (Tan, 2020; Choong, 2019).

The US' strategic opposition to China was multi-pronged but its Belt and Road Initiative (BRI) raised particular issues (Ching and Wu, 2018a,b). The BRI was global in nature and, from the Chinese perspective, operated within the region providing infrastructure support and funding opportunities for a range of countries not only in Asia but also as far afield as Africa and even in

Europe and Australia (Rolland, 2017). Others, however, saw it as creating debt traps (Hurley et al., 2019), creating trade imbalances (Cheng, 2016) and causing environmental problems (Ascensão et al., 2018). The IPS, as a security concept, therefore, was an attempt to curtail what was seen as Chinese expansionism and promote a Western oriented alternative, albeit with an Asian flavor provided by Japan and India. An expanded alternative was proposed by President Biden in 2021 called “Build Back Better World,” (B3W) specifically to counter BRI. Both B3W and BRI were strategic global initiatives representing the changed balance of power in the Asia-Pacific in the 21st century.

Finally, people of the Pacific Ocean came up with their own idea of how to deal with the region. Referred to as *The 2050 Strategy for the Blue Pacific Continent* (Pacific Islands Forum, n.d.), The Strategy identified multiple challenges for Pacific Island countries: climate change, aid dependency, poor economic growth, poverty and inequalities distance from markets, small economic base, transport costs with just one brief reference to security (Council Pacific Affairs, 2021) although there was a recognition that external factors influenced the region. Yet these were not of central concern to those who lived in the region. Nevertheless, at a discussion of the region’s strategic issues that subsequently took place in Australia almost the entire focus was on security highlighted particularly by those in Australia’s security establishment (Wallis and Batley, 2020). At one point in the discussion one Pacific island participant responded, “the Pacific doesn’t need national security; it needs wellbeing”.

For reasons that are difficult to comprehend, Australia security officials continued to argue that China’s presence in the Pacific was a security threat to Australia and therefore this should be a concern for Pacific Island countries. Yet the “Blue Pacific” was an idea of local people concerned about the future and their wellbeing; the Australian response was to highlight the security goals of the Indo-Pacific, subduing local needs and the people who actually lived in the Pacific rather than on its rim. In reality this represented how the Pacific side of the Asia-Pacific had always been treated. The indigenous strategy, the Blue Pacific, was an important fight back against external and peripheral constructions of the Asia-Pacific.

The multiple ideas of Asia-Pacific as discussed above show that such ideas have changed over time, even in terms of the geographies that define the region. The Asia-Pacific has been constructed to suit assumed economic and security needs usually with little or no participation of the people who actually inhabit the region. While there are realist conceptions of the region adopted by organizations such as UNESCO, they are little more than lines drawn on a map for administrative convenience. In what follows, reference will be made to educational issues within a broadly defined Asia-Pacific but with the understanding that the idea of the region can change at any time under multiple external influences that over time have constantly reconstructed it.

Constructions of education within the multiple geographies of the Asia-Pacific

Chen (2005) has argued that research should “cease ... to look at Asia as object of analysis” (p. 141) but rather as a point of reference, and importantly as a means of “inter-referencing.” Koh (2018) pointed out that this can be taken to mean that Western ideas, values or countries do not have to be reference for any discussion. As Chen (2005) pointed out:

Mediating through the horizon of “Asia” as an imaginary anchoring point, societies in Asia could begin to mutually see the existence of one another and become one another’s reference points (p. 140).

This “inter-referencing approach” will be taken in this section that seeks to view the Asia-Pacific through the lenses provided by various aspects of education.

International assessments and the education divide

PISA, 2018 (OECD, 2019) tested students from 77 countries in the areas of mathematics, science and reading. Putting aside technical issues about the resulting measurements and the social and political implications that arise from international testing regimes, one key issue rarely given any attention is a significant regional education divide across all three areas. East Asian societies (Mainland China, Macao Hong Kong, Korea and Taiwan) and Singapore all perform above the OECD average while South-East Asian societies (Malaysia, Brunei Darussalam, Thailand, Indonesia and the Philippines) all score below the OECD average. While the comparison of scale scores is not always helpful, the differences between the top East Asian society (China in each case) and the top South East Asian society (Malaysia in each case) is close to 150 score points representing 1.5 standard deviations.

There is no explanation for this divide in the report, although there are some general attempts to assess the impact of structural variables (e.g., resources available to education) on reading achievement (p. 65). Most high performing East Asian societies (the exception is Taiwan) expend more than the OECD average on education (although Mainland China does not seem to be included in the analysis) and most of the low performing countries (the exception is Brunei Darussalam) spend below the OECD average (p. 65). The association is not causal, and it is not perfect. Taiwan scored above the average OECD but spends below the OECD average while Brunei Darussalam spends above the average and scores below. In any event, understanding that the availability of resources maybe associated with learning outcomes is a very small step in understanding this regional education divide. In addition, this divide has also been revealed in another study with a quite different focus.

Chow and Kennedy (2014), using data from the 2009 International Civic and Citizenship Education Study (ICCS) (Schulz et al., 2010) examined students' intentions for future civic participation in Taiwan, Korea, Hong Kong, Indonesia and Thailand. They were asked specifically whether they would engage in various forms of protest (legal and illegal), in elections, in other forms of civic participation and in informal kinds of participation. They identified at one extreme those they called "Active Participators" (meaning participants in this group scored highly on each of the different forms of participation), with 10% of students from East Asian societies falling into this category and 33% from South East Asia. At the other end, they identified "Minimal Participators" (meaning participants in this group scored at the low end on each of the different forms of participation) 35% of whom were from East Asia and 10% from South East Asia (p. 481). The study suggested that these contrasts should be followed up in subsequent research as no explanation was attempted (p. 480). Yet, taken together with the PISA results reported above, it does seem that intra-regional variations, across broad subject areas, are part of the make-up of the region. To date there has been very little explanation for these variations although there have been some scattered attempts to understand better the educational dynamics across the region.

Kennedy (2012) compared Hong Kong and Thai students' civic values and their relationship to civic knowledge. Several results stand out, although they perhaps raise more questions than provide answers. Thai students' attitudes to religion were negatively related to the level of civic knowledge. The latter was relatively low in terms of the ICCS results thus establishing a negative association between religious values and civic knowledge. For Hong Kong students the reverse was true – their civic knowledge score was relatively high and the association with their attitudes to religion was negative. In relation to democratic values while Hong Kong students scored relatively highly on these values, they did not have any effect on their level of civic knowledge. The reverse was true for Thai students with their democratic values showing a positive association with their level of civic knowledge. In neither case are these associations causal and since ICCS is a cross sectional survey it does not indicate that the measurements are repeatable. Nevertheless, the differences revealed here are important and they did lead to further work of a qualitative nature seeking in particular to understand the values and attitudes of Thai students.

Kennedy (2014) reported the results of interviews with three groups of secondary students in Northern Thailand. Religion was clearly important to students in each of these groups but so too were families and friends. At the heart of each of these relationships was a concern for morality and for the development of "good people" and "good deeds and actions". Family members were trusted but, they needed to be "good people" and it was the same for politicians, in whom there was not a great deal of confidence. The requirement for a "good politicians" was that they needed to be a "good people". The moral concerns of these students is likely the product of their religious engagement and family values. They understand the importance of the civic in the sense that they had ideas about how political projects could achieve important things for the nation. Yet it seemed that their focus was on the morality of "good works" rather than the politics necessary to achieve them.

These analyses only start to paint a picture of complex social, political and cultural values. It takes us further than the somewhat bald results of ICCS, 2009 (Schulz et al., 2010), yet it is still incomplete. Much more work like this is needed to unfold the intricacies that underpin the civic values of students in South East Asia and therefore help us to understand better the regional education divide between East and South East Asia.

The form such work should take is important to consider. There is little doubt that surveys similar to PISA and ICCS can help to highlight issues and point to new ideas. Their quantitative methodology ensures large samples, good designs, relevant questions and easily interpretable results. Yet these results are limited in providing deeper explanations for the complex phenomena referred to above. There needs to be a new program of qualitative research that will enable countries, and groups of countries, to explore significant issues related to the education of young people. Whether it is the impact of resources, the influence of cultural values, the impact of poverty, gender stereotypes or racism, these issues need to be explored in depth with respect to their influence on education. Western explanatory models are not always relevant for the region and models developed for East Asian students may not be relevant for students in South East Asia. Research needs to be grounded in the local, as only qualitative research can be. Such research needs to be a priority for the future if we are to understand better the intricacies of a region that affects the lives of millions of young people every day.

Civic and social education

While the international testing regime has focused largely on mathematics, science, reading and technology (with ICCS being the exception), at the regional level there has been a broader interest in social education (Kennedy 2021a,b; Kennedy et al., 2014; Grossman and Lo, 2008; Kennedy et al., 2010; Grossman et al., 2008; Lee et al., 2004). The motivation for this extensive collection of work was to explore the distinctiveness of regional approaches to social education, identify commonalities and differences and reach a better understanding of regional constructions of social education. Attempts were made to be inclusive but not every country in the region was included in this work. Several articles focus on the Pacific Islands but the bulk of the work highlights issues in East and South East Asia, although there are some notable exceptions including Laos, Cambodia, Nepal and more. Nevertheless, this collection took seriously the notion of the Asia Pacific and attempted to understand how it provided social education for its young people. Several issues stand out.

Across the region there are school systems many of which owe their origins to the colonists of earlier periods. At the same time traditional values that preceded colonialism still survive (Kennedy, 2021a) meaning that present day nation states continue to engage with the issues of modernity often fueled by value systems that are thousands of years old. These values are often given expression in the religions of the region whether it is Buddhism in Thailand and Myanmar, Islam in Indonesia and Pakistan or

Hinduism in India. In addition there are major philosophical systems such as Confucianism that is an important influence in many East Asian countries. The religious values often find themselves in schools, unlike the secular education so often found in Western contexts, and if these values not reflected in schools, students will encounter them in their homes and in the community. Thus young people in the region are likely to carry with them these traditional values as well as the values of their modern globalized societies.

As students these young people will confront in school a curriculum structured along lines similar to those in Western countries – Mother tongue language, Mathematics, Science, Physical Education, Creative Arts and Social Studies. The latter in secondary schools is likely to be discipline based rather than integrated (although there may be integrated forms of Social Studies in primary education (Kennedy, 2021a,b)). The focus of social studies education across the region is likely to be nationalistic in nature. It will highlight the importance of the nation state, especially in the post-colonial environment that characterizes so many countries in the region. To this end there is also likely to be an emphasis on what is often referred to as 21st century skills (Kennedy, 2021a) since these nation states not only need to assert their post-colonial identity, they also need to engage in the challenges posed by economic and social conditions of the 21st century. Social studies education in different contexts is designed to help students with these future challenges.

Accompanying programmes of social studies education across the region there is likely also to be a subject referred to as Civic Education or Citizenship Education or Moral Education. As has been shown on multiple occasions (Kennedy et al., 2014; Grossman and Lo, 2008; Kennedy et al., 2010; Grossman et al., 2008; Lee et al., 2004) the focus of these programmes is likely to be more moral than civic with a focus on preparing the “good citizen”. While similar subject names are used in Western contexts, regional variants are more likely to follow traditional values that seek strong versions of social cohesion around a set of moral values. This is one of the key characteristics of school based citizenship preparation programmes in the region (Kennedy and Fairbrother, 2004).

Despite the generalizations outlined above, the diversity in providing social and civic education should not be underestimated. Kennedy et al. (2014), for example, provided the opportunity for Mainland Chinese scholars to discuss citizenship education in their unique context. Most often citizenship education is considered in terms of democratic political systems, but China provides an example of how single party systems seek to prepare future citizens. The examples provided leave little doubt that the Chinese state is as serious about citizenship preparation as any democratic society. But China is not preparing democratic citizens – it is preparing citizens who will support socialism with Chinese characteristics. This disrupts any notion of social cohesion across the region and reflects part the region’s diversity.

On the issue of diversity and citizenship, the Pacific Islands provides a different example from the one above. In the 21st century, countries like Fiji had a British colonial legacy on one side and the pressure of foreign aid donors on the other (Koya Vakauta, 2011). Citizenship education was a relatively new requirement for Fiji schools and the form it was meant to take under the influence of Australian and New Zealand aid donors was the traditional civic literacy approached based on political institutions and values. Yet scholars such as Thaman (2004) pointed out that citizenship education in the Fiji context was more than civic literacy. Koya Vakauta (2011) used terms like “being Fijian” and “belonging to Fiji” (p. 77) to indicate that there are indigenous values that underpin life in Fiji and these are as important as the democratic frameworks neo-colonialist nations wished to impose. This view was not about rejecting democracy, but rather about recognizing local values as the basis for citizenship and consequently citizenship education. The concept of understanding cultures and meaning in the Pacific Island context was a theme running through Thaman’s work over a long period of time (Thaman, 1997, 2004, 2008, 2009). She highlighted the importance of decolonizing thinking about the Pacific and highlighting indigenous values as part of the development of modernity in Pacific Island countries.

The emphasis on values in Pacific Island thinking about citizenship and citizenship is consistent with views in East and South and South East Asia. Epistemologically, Pacific Island values, as proposed by people such as Thaman, are different from those elsewhere. Christianity, brought to the Islands by the different colonial powers, has exerted a significant influence claiming he majority of adherents across the sub-region. Yet, as Dickie (2005) pointed out, “for many of those indigenous to the Pacific Islands region, seeing nature as sacred is still very prevalent. They continue to maintain great pride in their language, culture, and land, and they still hold strongly to the traditional myths and practices” (p. 1). Just as in the other parts of the region, older value systems remained even when overlaid by colonized values and so-called modernization processes.

Kennedy and Fairbrother (2004) pointed out that while globalization has been a major influence across the region the societies within it “are resilient and their values are deeply held. Citizenship education has been an important means of sustaining these values and ensuring that future citizens are aware of them” (p. 300). On account of its political system, China has been able to do this more successfully than the Pacific Islands. Western liberalism has been challenged in this way different parts of the region and this is often reflected in school programmes of social and civic education.

Asia-Pacific geographies as pathways for student mobility

According to the UNESCO Institute for Statistics, 2021, Australia currently acts as host to 444,514 international students representing 8% of the total number of mobile students in the world. The top 12 country/society sources of these students are in Asia (China, India, Nepal, Vietnam, Malaysia, Indonesia, Pakistan, Hong Kong, Sri Lanka, Republic of South Korea, Singapore and Thailand) representing around 75 % of the total. A small number, representing 0.09% of the total, come from the Pacific Islands (Fiji, Solomon Islands, Papua New Guinea and Tonga). In return, there are 13, 319 Australian outbound students with the top five destinations being the United States, New Zealand, the United Kingdom, Germany and Canada taking 78% of the total. Only 4.6% of these outbound students go to universities in the region (Japan, Republic of South Korea and Malaysia). These internationalization

processes, therefore, are quite unequal. Not only is there very little reciprocity between Australia and Asia-Pacific countries in terms of student exchanges, but Australian outbound students make up just 3% of incoming students. This is an issue that has both policy and theoretical dimensions.

From a policy perspective it is clear that for the Australian government and its universities the economic aspect of international education most likely outweighs the intercultural benefits of student mobility. Rizvi (2011) pointed to the influence of neoliberalism on Australian international student policy. In this context, the internationalization of higher education becomes business tool to attract full fee-paying students than for intercultural development. In 2018, the top three countries from which students came (China, India, Nepal) contributed \$18.4 billion to the Australian economy in what were referred to as "education related travel services exports" (Department of Foreign Affairs and Trade, 2019, p. 11). There is little doubt that Australian universities provide intercultural opportunities for Australians and students from Asia and the Pacific. Yet the internationalization of higher education is a business in a neoliberal economic environment and the country's fourth largest export industry (Brumby, 2021). Across the region, however, there is an attempt to create multiple destinations for the Asia-Pacific's mobile students.

Kuroda et al. (2018) referred to the "Asianization of Asia" by which they meant the intra-regional mobility of students. Two-way student flows between Japan, China, Korea and ASEAN countries account for an increasing number of students remaining within the region (p. 27). The numbers by no means match the student flows to Australia (or to areas outside the region such as the United States and the United Kingdom), yet they do show a developing reliance on the regional rather than the international. Structures such as University Mobility in Asia and the Pacific (UMAP) and The Asian Universities Alliance (AUA) show the determination of regional players to focus on intra-regional mobility to make the region a beneficiary of student mobility rather than entities outside the region.

Whether this focus on redirecting regional student flow within the region rather than outside of it is motivated by any more noble reasons than the pure business motives of Australian universities and their government is an issue that has not been widely discussed. Overall, the general literature tends to suggest that the main reason for seeking education outside of the home country is for human capital purposes (Kaushal and Lanati, 2019). There is at times an interest in migration, but there is very seldom mention of intercultural sensitivity. It is also of interest to note that the chief proponents of intra-regional student mobility are the economically developed nations of East Asia (Japan, South Korea, China) whose concerns are as much about their knowledge economies and neo-liberal policies as their equivalents in the West. One exception to this is Malaysia where in bound student flows from countries like Indonesia and the Middle East based on religious compatibility is a chief pull factor (Luo, 2017). This new area is neglected by Western researchers and further work is necessary to understand the policy and theoretical implications of students moving within the Asia-Pacific rather than outside it.

Conclusion

The Asia-Pacific as a construct has changed overtime largely because of the influence of those external or at least peripheral to the region. No matter how conceived, it is diversity rather than uniformity that characterizes the region whether the focus is on religion, social systems, political systems or educational achievement. There is some commonality in the social aspects of the school curriculum and citizenship education. Yet unlike in the West, however, the focus of the latter is not so much on institutions and structures as in producing "good citizens".

In the 21st century structures have emerged to try to mold a regional identity not so much to keep outsiders out in these globalized times but to ensure that those in the region can play a proactive role in rapidly changing economic times. What is needed in the future is a research agenda that will help to develop greater understanding of the region, its values and its priorities. The marginalization of Pacific Islanders and their values in the light of neo-colonial aid donors is an issue of considerable importance to issues of equity and equality in the region. The same issue applies to those countries outside the East and South East Asian orbit, most of which have not even been mentioned in this review article. In its most recent guise as the Indo-Pacific security region, there remains much that is unknown outside of the so-called security priorities. There is much to learn about the people who live, work and develop their lives across the region's breadth. In the end it is such people who represent the region rather than the constructions placed on it by external forces.

References

- Ascensão, F., Fahrig, L., Clevenger, A., Richard, T., Corlett, R., Jochen, A., Jaeger, A., Laurance, W., Pereira, H., 2018. Ironmental challenges for the Belt and Road initiative. *Nat. Sust.* 1, 206–209. <https://doi.org/10.1038/s41893-018-0059-3>.
- Brumby, J., February 19, 2021. Foreign students are our fourth largest export, even with the pandemic. *Syd. Morning Her.* Retrieved on 17 May 2021 from: <https://www.smh.com.au/national/foreign-students-are-our-fourth-largest-export-even-with-the-pandemic-20210218-p573pz.html>.
- Castells, M., 1992. Four Asian tigers with a dragon head: a comparative analysis of the state, economy, and society in the Asian Pacific Rim. In: Appelbaum, R., Henderson, J. (Eds.), *States and Development in the Asia Pacific Rim*. Sage Publications, Inc, Newbury Park, CA, pp. 33–70.
- Chen, K.S., 2005. Asia as method. *Taiwan A Radic. Q. Soc. Stud.* 57, 139–218.
- Cheng, L., 2016. Three questions on China's "Belt and Road initiative". *China Econ. Rev.* 40, 309–313.
- Chong, A., Wu, S.S., 2018a. Indo-Pacific Versus Asia-Pacific: Contending Visions. *RSIS Commentaries*, No. 034. Nanyang Technological University, Singapore. Retrieved on 6 ay 2021 from: <https://www.rsis.edu.sg/rsis-publication/cms/co18034-indo-pacific-vs-asia-pacific-contending-visions/#.YJQDF7UzbyE>.

- Chong, A., Wu, S.S., 2018b. Indo-Pacific Versus Asia-Pacific: Contending Visions? Retrieved on 28 July 2021 from 'Indo-Pacific' vs 'Asia-Pacific': Contending Visions? - AsiaViews.
- Choong, W., 2019. Indonesia, ASEAN and the return of the Indo-Pacific strategy. *Aust. Outlook*. Retrieved on 28 July from. <https://www.internationalaffairs.org.au/australianoutlook/indonesia-asean-return-indo-pacific-strategy/>.
- Chow, K.F., Kennedy, K., 2014. Secondary analysis of large-scale assessment data: an alternative to variable-centred analysis. *Educ. Res. Eval.* 20 (6), 469–493.
- Chua, C., 2021. The ever given and the monstrosity of maritime capitalism. *Boston Rev.* 4 May, Retrieved on 4 May from. <https://bostonreview.net/class-inequality-politics/charmaine-chua-ever-given-and-monstrosity-maritime-capitalism>.
- Council Pacific Affairs, 2021. The 2050 strategy for the Blue Pacific Continent. Retrieved on 4 April 2022 from. <http://www.councilpacificaffairs.org/ressources/the-2050-strategy-for-the-blue-pacific-continent/?msclid=91d73b80b3bd11ecafb069aa2611a307>.
- Denisov, I., Paramonov, O., Arapov, E., Safranchuk, I., 2021. Russia, China, and the concept of Indo-Pacific. *J. Eur. Stud.* 12 (1), 72–85.
- Department of Foreign Affairs and Trade, 2019. Trade in Services Australia 2018. Retrieved on 16 May 2021 from. <https://www.dfat.gov.au/sites/default/files/trade-in-services-australia-2018.pdf>.
- Dickie, R., 2005. Indigenous traditions and sacred ecology in the Pacific Islands. *Univ. Wiscon. Lacr. J. Undergr. Res.* VIII. Retrieved on 14 May 2021. from. <https://www.uwlax.edu/globalassets/offices-services/urc/jur-online/pdf/2005/dickie.pdf>.
- Dirlik, A., 1992. The Asia-Pacific idea: reality and representation in the invention of a regional structure. *J. World Hist.* 3 (1), 55–79.
- Garnaut, R., 1999. APEC ideas and reality: history and prospects. In: Paper Presented at the 2th Pacific Trade and Development Council, Osaka, 16–18 July. Retrieved on 6 May 2021 from. <https://cpb-ap-se2.wpmucdn.com/blogs.unimelb.edu.au/dist/a/142/files/2016/01/APEC-IDEAS-AND-REALITY-PAFTAD-1999-2g0tcdk.pdf>.
- Grossman, D., Lo, T.Y., 2008. *Social Education in Asia: Critical Issues and Multiple Perspectives*. Information Age Publishing, Charlotte, North Carolina.
- Grossman, D., Lee, W.O., Kennedy, K. (Eds.), 2008. *Citizenship Curriculum in the Asia Pacific*. Centre for Comparative Education Research and Springer, Hong Kong and Dordrecht.
- Hurley, J., Morris, S., Portelance, G., 2019. Examining the debt implications of the Belt and Road initiative from a policy perspective. *J. Infrastr. Pol. Dev.* 3 (1). <https://doi.org/10.24294/jipd.v3i1.1123>.
- Kaushal, N., Lanati, M., 2019. International Student Mobility: Growth and Dispersion. National Bureau of Economic Research (NBER). Working Paper 25921. Retrieved on 17 May 2021 from. <http://www.nber.org/papers/w25921>.
- Kennedy, K., 2012. Asian students' citizenship values and their relationship to civic understanding: an exploratory study comparing Thai and Hong Kong students. *Res. Comp. Int. Educ.* 7 (2), 248–259.
- Kennedy, K., 2014. Thai students talk about citizenship responsibilities: what the ICCS 2009 survey didn't tell us. In: Paper Presented at the Conducting Cross-Cultural and Cross National Research in International Setting Conference. Chulalongkorn University, Bangkok, 15–17 January.
- Kennedy, K. (Ed.), 2021a. *Social Studies Education in South and South East Asia*. Routledge.
- Kennedy, K. (Ed.), 2021b. *Social Studies Education in East Asia Contexts*. Routledge, New York & London.
- Kennedy, K., Fairbrother, G., 2004. Asian perspectives on citizenship education in review: postcolonial constructions or precolonial values? In: Lee, W.O., Grossman, D., Kennedy, K., Fairbrother, G. (Eds.), *Asia-Pacific Perspectives on Citizenship Education: Concepts and Issues*. Centre for Research in Comparative Education and Kluwer Academic Publishers, Hong Kong, pp. 289–302.
- Kennedy, K., Lee, W.O., Grossman, D. (Eds.), 2010. *Citizenship Pedagogies in Asia and the Pacific*. Centre for Comparative Education Research and Springer, Hong Kong and Dordrecht.
- Kennedy, K., Fairbrother, G., Zhao, Z.Z., 2014. *Citizenship Education in China: Preparing Citizens for the "Chinese Century"*. Routledge, New York and London.
- Koh, A., 2018. "Inter-referencing" as methodology: the "emotional contagion" of PISA and the discursive formation of an emergent East Asian policy field. In: Kennedy, K., Lee, J.C.K. (Eds.), *Routledge International Handbook of Schools and Schooling in Asia*. Routledge, New York, NY, pp. 40–57.
- Koya Vakauta, C., 2011. A reflection on 'being' 'Fijian' and 'belonging' to Fiji: conceptualizing a value-theory approach to citizenship education. *Pac. Asian Educ.* 23 (2), 77–88.
- Kuroda, K., Sugimura, M., Kitamura, Y., Asada, S., 2018. Internationalization of higher education and student mobility in Japan and Asia. In: Background Paper Prepared for 2019 Global Monitoring Report. Retrieved on 17 May 2021 from. https://www.jica.go.jp/jica-ri/publication/other/175nbg000010mg5u-att/Background_Kuroda.pdf.
- Lee, W.O., Grossman, D., Kennedy, K., Fairbrother, G. (Eds.), 2004. *Citizenship Education in Asia and the Pacific – Concepts and Issues*. Centre for Comparative Education Research and Kluwer Academic Publishers, Hong Kong.
- Luo, N., 2017. Japan, Malaysia, Taiwan, and South Korea remap regional student flows. *World Education News and Reviews*. Retrieved on 27 August 2021, from. <https://wenr.wes.org/2017/08/global-mobility-japan-malaysia-taiwan-and-south-korea-look-to-shift-student-mobility-by-2025>.
- Medcalf, R., 2015. Reimagining Asia: from Asia-Pacific to Indo-Pacific. *Asian Forum*, 26 June. Retrieved on 7 May from. <http://www.theasianforum.org/reimagining-asia-from-asia-pacific-to-indo-pacific/>.
- OECD, 2019. *PISA 2018 Results*. In: *What Students Know and Can Do*, vol. 1. OECD Publishing, Paris.
- Pacific Islands Forum, n.d. The 2050 Strategy For the Blue Pacific Continent. Retrieved on 9 May 2021 from. <https://www.forumsec.org/pacific-regionalism/>.
- Ravenhill, J., 1999. APEC and the WTO: which way forward for trade liberalization? *Contemp. S. Asia* 21 (2), 220–237.
- Rizvi, F., 2011. Theorizing student mobility in an era of globalization. *Teach. Teach.* 17 (6), 693–701.
- Rolland, N., 2017. China's "Belt and Road initiative": underwhelming or game-changer? *Wash. Q.* 40 (1), 127–142.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., Losito, B., 2010. ICCS 2009 International Report: Civic Knowledge, Attitudes, and Engagement Among Lower-Secondary School Students in 38 Countries. International Association for the Evaluation of Educational Achievement (IEA), Amsterdam, the Netherlands.
- Sidaway, J., 2002. Pacific dreaming, APEC, ASEAN and their geographies: reflections on Poon (2001). *Roy. Geogr. Soc.* 34 (2), 204–214.
- Spate, O.H.K., 1977. 'South Sea' to 'Pacific Ocean'. *J. Pac. Hist.* 12 (4), 205–211.
- Tan, S.S., 2020. Consigned to hedge: South-east Asia and America's 'free and open Indo-Pacific' strategy. *Int. Aff.* 96 (1), 131–148.
- Tellis, A., 2020. Waylaid by contradictions: evaluating trump's Indo-Pacific strategy. *Wash. Q.* 43 (4), 123–154.
- Thaman, K., 1997. Reclaiming a place: towards a Pacific concept of education for cultural development. *J. Polyn. Soc.* 106 (2), 119–130.
- Thaman, K., 2004. Whose values and what responsibilities? Cultural and cognitive democracy in education. In: Keynote Address, Pacific Circle Consortium Conference, Hong Kong, April 21 – 23.
- Thaman, K., 2008. Culture, Teaching and Learning With Specific References to Oceania. Keynote address, Shimane-Yamaguchi global seminar, Hiroshima, August 7.
- Thaman, K., 2009. Towards cultural democracy in teaching and learning with specific references to Pacific Island Nations (PINs). *Int. J. Scholarsh. Teach. Learn.* 3 (2), 1–11.
- UNESCAP, 2020. *Asia and the Pacific SDG Progress Report 2020*. Retrieved on 3 May 2021 from. https://www.unescap.org/sites/default/files/publications/ESCAP_Asia_and_the_Pacific_SDG_Progress_Report_2020.pdf#page=41.
- UNESCO Institute of Statistics, 2021. Global flow of tertiary level students. Retrieved on 16 May 2021 from. <http://uis.unesco.org/en/uis-student-flow>.
- Wallis, J., Batley, J., 2020. How does the 'Pacific' fit into the 'Indo-Pacific'? The changing geopolitics of the Pacific islands. *Sec. Challenges* 16 (1), 2–10.
- Wanandi, J., 2019. East Asian cooperation: retrospect and prospect. In: Tay, S., Armstrong, P., Drysdale, P., Intal, P. (Eds.), *Collective Leadership, ASEAN Centrality, and Strengthening the ASEAN Institutional Ecosystem*. ERIA, Jakarta, pp. 49–55.

Education in South-East Asia. Much achieved, much still to do

Anthony Welch, School of Education and Social Work, University of Sydney, Camperdown, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Regional background	419
Diversity	419
National identity	420
Inequality	420
Colonial legacy	420
Demographic challenge	420
Dropout: forms and problems	421
Education and decentralization	421
Financing of education	421
Corruption	421
Public private balance	422
Regionalism	422
Teachers and university faculty: training and qualifications	423
Quality assurance	423
References	424

Regional background

The diverse pattern of education in South-East Asia is composed of a rich mix of colonial legacies, post-colonial ambitions, growth and development, governance, and its growing importance as a world region. Embracing the world's most populous Muslim-majority nation (Indonesia) with a total population of 270 million, and some of the smallest (Brunei, population 441,000), the region is an opulent blend of religious traditions (Islam, Buddhism, Hinduism, Christianity and local spiritualism and pantheism), some of which spawned important educational initiatives. The Khmer empire developed important centers of Hindu and (Theravada) Buddhist learning at Angkor Wat's famed temples, built in the first half of the 12th century and replete with substantial libraries (Hayhoe, 2019), while the grand Buddhist temple of Borobudur, built in the 8th to 9th centuries by the Shailendra dynasty of Java, featured an architectural design based on a cosmology offering an ascending path to enlightenment. Catholic missionaries who pushed into current Philippines and Việt Nam, developed schools and universities, while in Burma (now Myanmar) the Baptist Judson College became part of the new Rangoon University (now Yangon University, ရန်ကုန် တက္ကသိုလ်) in 1920. The Confucian system of ethics, including the core principle regarding the value of knowledge and self-cultivation, was the basis of higher learning and governance in Việt Nam for something like a thousand years, and its heritage is still tangible in today's Singapore (Welch, 2019).

Religious influences remain substantial: in both Brunei and Malaysia, for example, Islam is the state religion, while in neighboring Indonesia, a separate Ministry of Religious Affairs (MORA) continues to administer Islamic schools and higher education institutions (HEIs). In Myanmar, monastic education based on Theravada Buddhism survives, if mostly for the poorest. The relative independence of monastic education, however, allowed it to introduce innovations that were, at times, subsequently taken up in the public sector. In this instance, this meant that "... children from the most disadvantaged background ultimately had access to better teaching practices than those at government schools, as a result of the work pioneered in (monastic schools)." (Lall, 2020: 101).

Diversity

The multi-faceted jewel of South-East Asia displays a wide range of development levels, political systems, languages and ethnicities. Overall population in the region totaled 660 million in 2020, with regional GDP reaching US\$3.2 trillion; collectively, this made it the 10th largest economy in the world (ASEAN, 2020). While, on average, services, including education, comprise around half of the regional economy, with both industrial and agricultural sectors declining in proportional terms, the share of the national economy that the service sector occupies varies widely from 39% (Việt Nam) to 65% (Singapore). (ASEAN, 2020).

Political systems vary substantially, from monarchies (Thailand, Brunei), to various forms of military regime (Thailand, and Myanmar, previously called Burma), strong-man authoritarian regimes (Philippines, Cambodia), single-party socialist market economies (Việt Nam, Laos), and guided democracies (Singapore). Education is considered important throughout the region, and families may often make considerable sacrifices to send their children to the best education they can afford. This may often involve significant additional payments for private extra-school classes, after hours.

National identity

National identities remain strong, and are promoted within the education system, for example in Indonesia's *Pancasila* (Five Principles). But national identities are also challenged by the strength and persistence of local affiliations, languages, and ethnic ties that at times transcend national borders. Linguistic diversity is a feature of several systems, with Indonesia alone, for example, featuring 300 languages (including the national language Bahasa, spoken by 200 million as their second language). At the same time, efforts to promote a trans-national regional identity based around the 11-member Association of South-East Asian Nations (ASEAN), founded in 1967, are ongoing, including via regional mobility programmes such as ASEAN International Mobility for Students, and the EU's SHARE scholarship program that offers scholarships to study in an HEI within ASEAN. A further pan-ASEAN feature is the adoption of English as the official language of ASEAN. In practice most countries permit non-dominant languages as a medium of instruction.

The politics of ethnicity can shape educational outcomes. In Malaysia, Chinese-Malaysian families overwhelmingly send their children to Chinese medium schools, while at the higher education level, ethnic quotas, formally abandoned in 2003, still mean that, in practice, Chinese Malaysians and Indian Malaysians have limited options: private higher education, or a university overseas. Students from Tamil or Chinese medium schools who cannot demonstrate sufficient proficiency in the national language must enroll in a one year transition class, since *Bahasa Melayu* is the official medium of instruction at the secondary level (UNESCO, 2014b: 30–31).

Inequality

Participation rates, particularly at secondary level, have risen in recent decades but vary appreciably. Secondary school net enrollment ratios (NER) vary from 35% in Laos to 99% in Singapore. No ASEAN member state has an NER of less than 90% at the primary level. (ASEAN, 2020: 15) Substantial levels of inequality in education have been exacerbated by neo-liberal forms of globalization in recent decades, and deepened by successive shocks such as the regional currency crisis of the late 1990s, the Global Financial Crisis a decade later, and the COVID pandemic of 2020–21. Gini co-efficients, (a widely used measure that charts the degree of income inequality in a society) that were already high (mostly around 0.40 or so in 2019), worsened in response to COVID. In many systems, the effects pushed millions of families, especially from the informal economy, back into poverty, with severe consequences for school participation. UNESCO data on the effects of the COVID-19 pandemic revealed widespread school closures, and the number of out of school children worldwide rising to around 9% of total learners, or 157 million individuals worldwide. The consequences within South-East Asia were profound with estimates that 68 million children were affected in Indonesia alone, with a further 15 million in Thailand and 10.5 million in Myanmar (UNESCO, 2021a).

Colonial legacy

A kaleidoscope of colonialism has marked the region, including in education. Dutch colonialism shaped education in Indonesia, British colonialism helped structure education in Malaysia, Myanmar, Brunei and Singapore, while French colonialism molded education in Cambodia, Laos and Việt Nam (Symaco, 2013). The post WWII era saw progressive de-colonization across the region, with Indonesia declaring independence in 1945 (although sovereignty was not formally transferred from the Dutch until 1949), Laos and Cambodia in 1953, and Malaysia in 1957. But in some cases, the independence struggle took much longer. While Việt Nam's Ho Chi Minh declared independence from the French immediately after the Japanese surrender in 1945, a united Communist state was not achieved for another 30 years. Thailand was the only country in the region not to have been colonized, although still subject to significant external influences.

But the colonial legacy was not easily overcome, and the challenge of building national education systems was compounded by the post-independence departure of many colonial administrators and educators, depriving the local system of much-needed experience and specialist skills. New curricula and textbooks were needed, as well as many new teachers. Moreover, education ministries tended to be reluctant to cede much control to either institutions or to trans-national initiatives. While pilot projects were cautiously introduced in Indonesia, for example, to accord more autonomy to select universities, to this day in Myanmar, "Everything is very centralized and the universities have hardly any autonomy." Lall (2020: 134).

Demographic challenge

An added challenge was demographic: a burgeoning youth population profile that remains common to the region, imposed substantial demands on nascent national systems of education. A current population cohort of 213 million youth (15–34 years) in ASEAN countries, is the largest ever, and is expected to peak at over 220 million by the late 2030s (ASEAN, 2017). The high proportion of the overall population aged under 15, that ranges between a quarter and a third for almost all ASEAN member states, ensures major ongoing demand for schooling, while growing levels of wealth and aspiration underpin rising demand for higher education, at times of uneven quality (ASEAN, 2020; World Bank, 2020a). Rising aspirations for education, and in several cases a swelling middle class, add to aggregate demand for education, in both quantitative and qualitative terms.

Dropout: forms and problems

But equity issues persist. Dropout rates continue to be a problem in Laos, Việt Nam, Cambodia and Myanmar, especially at the secondary stage, and particularly among the poor and rural youth. In Myanmar, over half of the wealthiest 20% of the population had either completed secondary or gained higher qualifications and only 7.5% had never attended school. But for the poorest 20%, the ratio with secondary or higher qualifications was less than 4%, while 30% had never attended school (Republic of Myanmar, 2014). In Cambodia, 84% of children complete primary school, but only 20% complete the secondary stage, with a dropout rate from lower secondary to upper secondary of 27.5% (SEAMEO, 2017: 43). In Myanmar, rates of attendance fall from 80% at age 11 to 36% at age 16, and one in six children who should be attending primary school were not enrolled (Republic of Myanmar, 2014). Regional variations are also substantial with the rate of those who never attended school varying from 1.7% in the national capital region (Naypidaw) to 21.9% in Shin, 10.1% in Kayin and (a claimed) 5.5% in Rakhine (Republic of Myanmar, 2014, see also Lall, 2020: 29). Across the region gross enrollment ratios at secondary level range widely from 42 to 108% with an average of 77%. By contrast, gross tertiary enrollment ratios range from 11 to 21%, with the regional average being 18% (ASEAN, 2017: 27).

Education and decentralization

The most common age at which schooling begins in the region is 6. Most ASEAN member states feature 6 years of primary schooling, often with 3 + 3 at the secondary level. The systems in both Laos and Việt Nam systems operate on the basis of 5 years of primary schooling with 4 + 3 at the secondary level (UNESCO, 2014b: 12). With the exception of Myanmar and Laos, where only 5 years of schooling is free and compulsory, all other ASEAN systems mandate at least 6 years of compulsory and free schooling. (UNESCO, 2014b) Governance and administration tend to remain rather centralized, especially regarding primary source of funds, quality control and standards and teacher administration, despite varying degrees of decentralization instituted across the region in recent decades, via laws or regulations. Standard setting tends to be centralized, but in Việt Nam, for example, teacher recruitment may be devolved to the Province or District level, while in Indonesia, which introduced a substantial degree of decentralization in 1998, most teachers are selected and paid by local government. At the local levels, various issues such as underfunding, lack of capacity, and accountability are problems in systems such as Laos, Việt Nam and Myanmar.

Financing of education

The proportion of public spending on education as a percent of total government spending varies widely, from less than 5% in Myanmar, to over 22% in Thailand. For the years 2007–2010, public expenditure on education as a proportion of Gross Domestic Product (GDP) ranged from less than 1% in Myanmar, to almost 6% in Malaysia, although the latter figures is somewhat misleading, since much of the funding is concentrated on educational institutions dominated by *Bumiputra* (ethnic Malays). (UNESCO, 2014b: 9) For some states, the financial burden of catering to the growing demand for education (among numerous other competing demands on the state purse) is challenging, and together with varying traditional levels of state provision, leads to rising levels of private educational supply (UNESCO, 2021b; Welch 2021). As well, state educational institutions may be under pressure to respond to rising demand under conditions of constrained state budgets, which leads to varying degrees of privatization of public sector educational institutions. State capacity, too, varies widely from wealthy, sophisticated Singapore, to Myanmar, which is still rebuilding an education system that was neglected for decades. Systems that spend higher proportions of their budget on education as a proportion of Gross Domestic Product (GDP) also tend to have greater capacity and stronger quality assurance systems, as well as more robust teacher training systems and better teacher remuneration.

Corruption

Corruption is a substantial issue throughout the region and is linked to both governance, state capacity and finance. Afflicting most parts of the economy, the education sector is by no means exempt, and the phenomenon reduces available funding and undermines both quality and equity of schools and universities (Fig. 1).

Transparency International's Corruption Perceptions Index (CPI) ranks and rates 180 countries on a 1–100 scale, with 1 representing wholly corrupt and 100 being wholly free of the taint of corruption. With the notable exception of Singapore, the data below reveals significant levels of corruption in several ASEAN member states. Specific forms of corruption in education range from cronyism in appointments and promotions, to the sale of positions, bribery, sexual abuse of students (for admission, or higher grades), fake qualifications, and research problems such as plagiarism and fake results. (Welch, 2020).

Country	Country Rank	Corruption Score
Singapore	3	85
Malaysia	57	51
Indonesia	102	37
Thailand	104	36
Philippines	115	34
Việt Nam	104	36
Myanmar	137	28

Fig. 1 Levels of corruption, selected South East Asia. Source: [Transparency International \(2020\)](#).

Public private balance

Limited state capacity and tensions between rising demand and limited state finances has resulted in a varying balance between public and private sectors in South-East Asian education systems. Some systems have long had private sector establishments, both schools and higher education institutions. In many cases, their origins relate to missionary activities: the Philippines' Santo Tomas university, founded in 1611 by the Archbishop of Manila, has a genealogy longer than Harvard's, and there were also Catholic schools and universities in the south of Việt Nam, at least until re-unification in 1975. (Welch, 2011; 2021) Driven by increased urbanization, the decline of the extended family, limited state capacity, demand for English language instruction, the growth of the middle class, and increased interest in an international education, current estimates are that spending on private education in South-East Asia totals around US\$60 billion. (EY nd).

But the growth of the private sector has not always been well regulated. In a number of systems, private sector enrollment growth has outstripped public sector growth rates. The proportions within ASEAN education systems vary significantly, however; private schools account for 42% of the total in Indonesia, but only 9% in neighboring Malaysia, 25% in the Philippines, and 2% in Cambodia. Rates of private provision also vary by level, with the share often higher at higher levels (UNESCO, 2014b: 17, World Bank, 2020b; EY Parthenon, 2018). In Laos, the private sector is an important component of educational provision at all levels from pre-primary to higher education (SEAMEO, 2017: 92).

The share of total enrollments held by private sector higher education institutions (HEIs) also varies widely across the region: from around 15% in Việt Nam to around 66% in neighboring Cambodia. In Indonesia, the region's largest system with more than 8 million higher education enrollments, the private sector share is 62% (UNESCO, 2014a) But because private HEIs are much smaller on average, some with as little as 100 enrollments, and many with 500 or so, their share of institutional totals is exaggerated (Welch, 2021). In general, quality is higher in the public sector, although in Indonesia and the Philippines a small number of private HEIs vie for equivalent leading status with leading public HEIs. Rates of private spending as a proportion of total expenditure on education also vary significantly, by country and level, with private expenditure at the tertiary level accounting for more than half the total in Indonesia and two-thirds in the Philippines (UNESCO, 2014a: 18).

But the relatively faster growth of the private sector in a number of systems, at least in higher education, is only one side of the Janus face of privatization. The fact that growing pressure on demand for schooling and higher education has often not been matched by equivalent increases in state funding has put additional strains on public sector educational institutions. Governments remain keen to have more and more young people in education; but are unwilling or unable to match such increases with equivalent increases in financial support. The increasing gap between enrollments and finance places extra strain on public sector educational institutional finances. In order to shore up their bottom line, some HEIs in the region resort to dubious practices, such as lowering admission standards, enrolling more students than their allocated quota, accepting bribes to gain a place or for a higher grade, turning a blind eye to fake qualifications, illicitly granting degrees without the requisite work, and mounting similar courses for high fees in parallel private institutions, at times with uncertain financial transparency. The combination of rising fees at public HEIs, and significant fees at private HEIs, often of poorer quality, narrows options for some parts of the populace, and can mean that poor families may ultimately pay more for an inferior education.

Regionalism

While the national administration of education systems remains strong in practice, some decentralization has occurred, for example in Indonesia after 1998. Perhaps in part a response to a history of colonialism, countries within the region tend to be reluctant to cede much authority to regionalism. Nonetheless, programmes to foster a greater sense of regional identity, at both school and university level, are part of the ASEAN educational agenda (ASEAN, 2017; 2018). Schemes at the higher education level tend to be better developed and include the ASEAN University Network (AUN) of 30 HEIs across the region, the EU-ASEAN Support to Higher Education in the ASEAN Region (SHARE) scheme, and the South-East Asian Ministers of Education (SEAMEO) ASEAN International Mobility for Students (AIMS) program, administered by SEAMEO's Regional Institute for Higher Education (RIHED). For systems such as Myanmar, poor and underdeveloped after decades of neglect, such programmes have been of particular importance; by 2018, Myanmar had helped select 148 students for study in other ASEAN systems, while also hosting 17 in Myanmar.

But there are also wider links that arguably extend the notion of region. For many of the countries in the region, China is their major trading partner. Several, notably Singapore and Malaysia, have substantial ethnic Chinese minorities, which across the region

exercise a disproportionate economic influence. A second link derives from the fact that China borders several South-East Asian states, notably Myanmar, Laos and Việt Nam. Ethnic groups such as the Dai, Wa and Kachin straddle China's border, and universities on China's Southern border regions, for example, have long had connections to Việt Nam. Overall, education links between China and ASEAN have also grown significantly, including via more than 30 Confucius Institutes located within ASEAN, each of which is in the form of a bi-lateral agreement between a Chinese and an ASEAN university. But other networks have been expanded, too: subsequent to a bi-lateral *Memorandum of Understanding on an Academic Cooperation and Exchange* being signed in 2007, the AUN network now extends to several Chinese universities. The agreement included 13 Chinese HEIs and an equivalent number of ASEAN HEIs, as well as national representatives from each side. Numerous Confucius Classrooms also exist in the region, with Thailand hosting 20, for example, while China awards thousands of international scholarships to enable ASEAN students to enroll at its universities. Another extra-territorial link consists of links with Australia, where tens of thousands of students from ASEAN have long studied. Research collaboration between countries such as Singapore, and Malaysia, with Australia, also continues to grow. (Finnane, 2021; Welch, 2014a,b).

Teachers and university faculty: training and qualifications

Qualification levels for admission to teacher training vary somewhat across the region: most systems require Year 12 graduation, but Brunei Darussalam, Laos and Myanmar accept students with either Year 11, or Year 10 qualifications. Teacher training programmes vary in length, but higher levels of qualification are usually needed for secondary teachers. While teacher recruitment has mostly been devolved to provincial, or district level, Singapore, Malaya, Cambodia and Myanmar adhere to centralized recruitment schemes. Programmes are often of two years duration for primary school teachers and three to four years for secondary. In rural areas, teachers tend to be less qualified, and some are not registered or accredited. Efforts are ongoing in Malaysia and elsewhere to upgrade teacher qualifications, especially for those untrained, or with minimal education levels. Continuing professional development is a feature of several systems. Cambodian community teachers in the provinces are entitled to 16 days of professional development, for example, while Việt Nam requires teachers to engage in 30 days of in-service training per year. (UNESCO, 2014b: 20).

As with other public sector workers, teachers and academics in several ASEAN systems are poorly remunerated. Average annual public secondary school teacher salaries expressed as a proportion of GDP per capita range widely across the region, with Indonesia, Cambodia and Laos featuring the lowest proportion, and Malaysia and Thailand offering the highest. (UNESCO, 2014b: 39) In several systems, low salaries lead to varying degrees of moonlighting, in an effort to supplement overall earnings. For teachers, this often involves private tutoring (classes) after school hours, for fees. For school pupils, it is often not possible to pass the exam without participating in such out-of-hours classes, which means students from poor families who cannot afford the additional fees, are excluded and condemned to fail. For academic staff, moonlighting at private HEIs boosts their income considerably, but at the cost of reduced quality and energy at both public and private HEIs. Wealthy, highly-developed Singapore is an outlier, and awards salary increments based on performance and meeting professional standards. At the higher education level, bi-lateral agreements to recognize degrees as equivalent, such as that between Malaysia and China, aid in staff and student mobility.

Quality assurance

Quality assurance (QA) is a priority for ASEAN member states, but the effectiveness of regulation differs from system to system, depending on state capacity, and levels of development. In some systems different agencies address different levels, as with Indonesia's National Board of School Accreditation (BAN) and National Accreditation Board for Higher Education (BAN-PT). In Laos, the Educational Standards and Quality Assurance Center is responsible for QA at all levels, as is Việt Nam's General Department for Educational Testing and Accreditation (GDETA). In practice, corruption undermines the robustness and transparency of the process to varying degrees in several systems, with bribes, favors, cronyism and other nefarious practices not uncommon. Such problems weaken quality across the education system, reduce levels of available resources, and undermine public confidence in both the education and wider political system (Welch, 2011, 2020).

ASEAN's rich diversity of peoples, languages, cultures and religions is intersected by a common tradition of respect for learning, teachers and education, as well as the ambition to improve the quality of education and encourage wider participation. But sharp differences in wealth, state capacity, and quality of personnel lead to very different outcomes. The common aspiration to have at least one so-called "World Class University" provides an example. According to all the major global higher education ranking schemes, wealthy, highly developed Singapore has two, Malaysia has one, but no other ASEAN system has a university ranked among the leading 500 HEIs worldwide.

There is therefore a mix of achievement, aspiration and challenge that has differentially shaped the trajectory of individual ASEAN education systems. The rising importance of the region, the mix of persistent national identity and quest for a pan-ASEAN regionalism, cautious moves to decentralize, and ongoing problems of corruption in education will continue to shape both the regional and national form of education in decades to come.

References

- Association of South-East Asian Nations (ASEAN), 2017. First ASEAN Youth Development Index. https://asean.org/storage/2017/10/ASEAN-UNFPA_report_web-final-05sep.pdf.
- Association of South-East Asian Nations (ASEAN), 2018. ASEAN Empowers Students to Enhance Regional Identity. <https://asean.org/asean-empowers-students-enhance-regional-identity/>.
- Association of South-East Asian Nations (ASEAN), 2020. ASEAN Key Figures 2020. https://www.aseanstats.org/wp-content/uploads/2020/11/ASEAN_Key_Figures_2020.pdf.
- EY Parthenon Education Practice, 2018. Education in South-East Asia. Opportunities for Investors and Operators. https://cdn.ey.com/echannel/parthenon/pdf/perspectives/EY-Parthenon_education-in-SE-Asia_Paper_final_web_032018.pdf.
- Finnane, A. (Ed.), 2021. Australia Asian Research Collaboration in the Humanities: Mapping the Present Planning the Future. Australian Academy of the Humanities Report. <https://www.humanities.org.au/wp-content/uploads/2021/06/LASP-ASIA-LAYOUT-for-ARC.pdf>.
- Hayhoe, R., 2019. The gift of Indian higher learning traditions to the global research university. *Asia Pac. J. Educ.* 39 (2), 177–189.
- Lall, M., 2020. Myanmar's Educational Reforms. A Pathway to Social Justice? Press, London UCL.
- Republic of the Union of Myanmar and UNFPA, 2014. Policy Brief on Education. https://myanmar.unfpa.org/sites/default/files/pub-pdf/policy%20brief%20and%20infographics_Education.pdf.
- South-East Asian Ministers of Education Organization (SEAMEO), 2017. Guidebook to Education Systems and Reforms in Southeast Asia and China. SEAMEO, Bangkok.
- Symaco, L., 2013. Introduction. Education and issues in South-East Asia: a regional overview. In: Symaco, L. (Ed.), *Education in South-East Asia*. Bloomsbury, London, pp. xiii–xviii.
- Transparency International, 2020. Corruption Perceptions Index 2020. <https://transparency.org.au/cpi-2020/>.
- UNESCO, 2014a. Higher Education in Asia. Expanding Out, Expanding Up..... <http://uis.unesco.org/sites/default/files/documents/higher-education-in-asia-expanding-out-expanding-up-2014-en.pdf>.
- UNESCO, 2014b. Education Systems in ASEAN+6 Countries: A Comparative Analysis of Selected Educational Issues. UNESCO, Bangkok.
- UNESCO, 2021a. Global Monitoring of School Closures..... <https://en.unesco.org/covid19/educationresponse>.
- UNESCO, 2021b. Non-State Actors in Education. Who Chooses, Who Loses? https://en.unesco.org/gem-report/non-state_actors.
- Welch, A., 2011. Higher Education in South-East Asia. Blurring Borders Changing Balance. Routledge, London.
- Welch, A., 2014a. China-ASEAN relations in higher education: an analytical framework. In: Yonezawa, Kitamura, Y., Meerman, A., Kuroda, K. (Eds.), *Emerging International Dimensions in East Asian Higher Education*. Springer, Dordrecht, pp. 103–120.
- Welch, A., 2014b. Richer relations? Four decades of ASEAN-Australia relations in higher education. In: Wood, S., He, B. (Eds.), *The Australia-ASEAN Dialogue: Tracing 40 Years of Partnership*. Palgrave Macmillan, New York, pp. 145–166.
- Welch, A., 2019. Higher education in Asia. In: Rury, J., Tamura, E. (Eds.), *The Oxford Handbook of the History of Education*. Oxford University Press, Oxford. <https://doi.org/10.1093/oxfordhob/9780199340033.013.17>.
- Welch, A., 2020. Governance and corruption in east and South-East Asian higher education. In: Capano, G., Jarvis, D. (Eds.), *Convergence and Diversity in the Governance of Higher Education*. Cambridge University Press, Cambridge, pp. 355–395.
- Welch, A., 2021. Private higher education in East and South-East Asia: growth, challenges, implications. In: UNESCO Global Monitoring of Education Report 2021. Non-State Actors in Education. UNESCO, Paris.
- World Bank, 2020a. World Bank Indicators. Population 0–14 (% of Population)..... <https://data.worldbank.org/indicator/SP.POP.0014.TO.ZS>.
- World Bank, 2020b. School Enrolment, Secondary (Private), % of Total Secondary. <https://data.worldbank.org/indicator/SE.SEC.PRIV.ZS>.

Globalization and political economy of education development in South Asia

Jandhyala B.G. Tilak^a and Pushkarni Panchamukhi^b, ^a Council for Social Development, New Delhi, Delhi, India; and ^b R.V. University, Bengaluru, Karnataka, India

© 2023 Elsevier Ltd. All rights reserved.

Introduction	425
Human development, literacy and education development	427
Universalization of elementary education	429
Secondary education	431
Higher education in South Asia	431
Growth of higher education	432
Inequalities in higher education	434
Policy dynamics: grappling with globalization	436
Privatization of education	436
Globalization and medium of instruction	438
Internationalization and ranking of institutions	439
Globalization and public funding of education	441
External aid for education	443
Concluding observations	443
References	445

Introduction

Given that South Asia is the most populous and most densely populated region in the world, comprising about one-fourth of the world population, it is important to analyze developments that have taken place in there as they are influenced by, and in turn, influence developments in the other parts of the world. The South Asian region, also known as Indian sub-continent, comprising Islamic Republic of Afghanistan, People's Republic of Bangladesh, Kingdom of Bhutan, Republic of India, Republic of Maldives, Federal Democratic Republic of Nepal, Islamic Republic of Pakistan and Democratic Socialist Republic of Sri Lanka, which is a vast area of the size of Europe, exhibits some important similarities and significant differences. While striking dissimilarities exist across these countries, for instance, in terms of population-- the second most populated country in the world exists alongside the world's least populated one-- geographical area, social and political contexts, economic levels of development and the size of education systems, there are major similarities too. The region shares a common heritage, culture and history, starting from the ancient period; the several countries in the region are connected not only geographically sharing national boundaries and colonial legacies, but also through similar cultural practices, and religious and linguistic diversity. India has been the central core of the region both geographically and more importantly historically and culturally. The other countries, some of which were part of India earlier, form the peripheral region, and have been influenced for centuries, historically and politically by India. The region is characterized by an immense degree of diversity within a very broad contour of unity. In terms of a multitude of features in the economic, social, and political history, South Asia has been a "mini world," suggesting "about the early 'modern' world economy" (Washbrook, 1988).

The countries also face comparable challenges and problems in many spheres, including in education. Any discourse on how the landscape of education has evolved in the countries of South Asia over the past several centuries clearly speaks of a decline of a highly evolved knowledge society and a robust educational system, one of the most ancient civilizations in the world that used to be rich not only in its treasure of scriptural knowledge and practices, but also one that nurtured and inspired indigenous knowledge in the fields of science, mathematics, technology, engineering and medicine into a highly disintegrated irrelevant system. Knowledge and knowledge systems typically evolve over time implying that there is a gradual progress in not only in ways that knowledge is produced and disseminated, but also in the ontological meaning of the term, knowledge is redefined. Speaking of knowledge systems and education in the context of South Asia, one cannot but note that there has been a radical displacement of what defines knowledge and education by dominant Western systems. The Western systems have refused for long to acknowledge the contributions of scientists and mathematicians of the ancient India like those of Bouddhyana (for the discovery of several concepts in mathematics, including calculating the value of phi), Aryabhatta (known for his contribution of zero, as well as his discovery of the curvature of the earth and the fact that it rotates on its axis around the sun, the solar and lunar eclipses, the duration of a day as well as the distance between the celestial bodies of the earth and moon, etc.), Pingala (for using binary values), Acharya Kanad (for atomic energy), Bhaskaracharya (for his work on differential calculus), Varahamihira (for his contributions to astrology, hydrology, geology, ecology, including earthquakes and the concept of gravity), Brahmagupta (for the concept of gravity), Nagarjuna (for the doctrine of emptiness), Charaka (for Ayurveda – Indian medicine), Shusruta (for surgery, including plastic surgery), and Patanjali (for yoga). There were also many great authors of monumental literary and spiritual works of immense value. Ancient India laid the foundations in many areas of knowledge; particularly in mathematics, science, logic, medicine, psychology,

spiritualism and ethics, besides social and cultural studies, law, and language and linguistics. A closely related area refers to the great Indian engineering knowledge demonstrated in many constructions of the past 1,000–2,000 years in south Asia, which is not acknowledged in modern knowledge systems either (Tilak, 2022). The region once known as *Visvha Guru* in knowledge development and dissemination, has become the most ignorant and illiterate society. This is attributed to several invasions and the long colonial rule that most countries in South Asia endured, which uprooted “the beautiful tree” (Dharampal, 1983) and transformed an advanced intermediate society of India into an illiterate society, besides converting it into a raw material appendage on the economic front. The result of sustained colonial oppression was that the region sunk into abysmally low levels of literacy with high levels of human deprivation, shattered social and cultural structures and a ruined economy. As a result of continuous efforts made by the countries since the fifties of the twentieth century when they became independent, the several countries have come back gradually to the path of rebuilding their nations, even though they continue to face serious hurdles in their endeavor to achieve progress. Education is one such area where one finds paradoxically impressive growth along with mounting anomalies. As is the case in most other countries of the world, globalization, the present wave that started in the early 1980s, has dramatically impacted the landscape of education in South Asia.

Even before the present wave of globalization commenced, South Asian countries and their education systems have been seriously inflicted after independence, by conflicts based on differences in ideologies, religion and ethnicity. Some of them still continue to suffer. The ethnic war in Sri Lanka between the majority Sinhalese and the minority Tamils lasted for 26 years till 2009 resulting in a loss of countless lives and stress for the economy. The impact of the dispute was seen in different aspects of education beginning with school education, text books, higher education and the quality of teachers. During the colonial time of “divide and rule”, British favored the Tamils who venerated education and did not have much resistance to studying in English medium Catholic schools. Consequentially, a lot of them made it to higher education in good numbers. Several studies (e.g., Perera, 2001) provide detailed accounts of how the Tamils had progressively occupied majority of the positions in administrative services, railways, postal services, armed forces, law, medicine etc. Threatened by this, the Sinhalese government soon introduced education policies and schemes that majorly benefitted the Sinhalese, such as making Sinhalese the only official language, highlighting Sinhalese history, culture etc. in text books and favoring Sinhalese students for admissions into higher education. The Sinhalese hegemony created further problems, leading to a long civil war thereby significantly affecting especially the lives of children and the economy. A recent study on the impact of the war on children (Lakshman et al., 2020) showed how displacement and trauma caused by the war resulted in high dropouts and loss of interest in studying among children in the Northern Tamil region; impact that had both immediate and long term effects.

The religious conflict between India and Pakistan that led to their historic partition in 1948 post-independence has also severely damaged education in Pakistan and Kashmir in India. Pre-occupied with war against India over Bangladesh initially and later over Kashmir, Pakistan’s educational system suffered with very poor infrastructure and access. On the other hand, education in Kashmir remained disrupted due to constant tension and conflicts. Bangladesh, after its independence from Pakistan in 1971 began as an impoverished country for decades and only in 2015 gained the status of a lower-middle income economy. The resilient Bangladesh economy surpassed Pakistan on many fronts and its performance in education was also impressive since human resource development was a key policy thrust of the government of Bangladesh (ADB, 2008).

Nepal had strict dictatorial rule for more than a century from 1846 and education was unavailable to the masses. Faced with severe political turmoil, a Board of Education was established as late as in 1951 after which organized educational planning was undertaken with substantial foreign aid and access to education was provided to all. Owing to severe political instability with frequent Maoist insurgencies, Nepal has not been able to make significant progress in educational indicators.

Internal conflicts due to political and ethnical differences have had huge impact on the educational trajectories that countries in South Asia have taken. The combination of factors such as crisis of governance, failure of the democratic system, decentralization of power, imbalanced intra-state and inter-state developments and increased ethnic conflicts jeopardized the South Asian system (Lahiry, 2013). Because of its vast geographical size and being economically and politically more powerful, India has dominated the region in general and in education progress in particular.

Using most recent quantitative and qualitative information on some key parameters that define educational development, such as literacy, enrollment ratios in primary, secondary and higher levels of education, out-of-school children, inequalities, and public funding, and foreign aid, the article analyses the achievements and under-achievements in education in South Asian countries and factors that have contributed to them, in a broad political economy framework, acknowledging the importance of not merely the economic factors but also of social, cultural, historical, and political factors, the roles played by divergent actors within and outside the government—politicians, civil society organizations, community members, private investors, businessmen, philanthropists—national and international, working at macro level as well as at grassroots in shaping education policies and development. After all, education is not situated in isolation; it is influenced by a broad macro environment and the prejudices and interests of various players. In turn, education influences the broader environment in a variety of ways. The article highlights the trajectory of developments in both school and higher levels of education in South Asian countries in the wake of globalization. It also discusses the major guiding principles and factors behind educational policies and strategies introduced in these countries and some paradigm shifts that have taken place in their approach toward education. As an immediate and significant fallout, the general trends on how policies have facilitated the growth of certain types of educational institutions, apart from discussing the global dynamics and connections that have distorted national priorities, policies and practices. In this context, the article refers to some landmark global declarations in the field of education that have shaped national policies, which caused shifts in national priorities and changes in principles of educational planning in these countries. It may be noted that only a few select key dimensions of education development have been analyzed as in a short article of this size a comprehensive analysis of the vast area of education cannot be made.

Human development, literacy and education development

Education was always placed on a high pedestal in the history of South Asia. Education has always been valued for its own sake, recognized for the value it has added to the overall dignity of life. Numerous invasions and colonial rule devastated not only existing education systems, but also economic and political structures. Hitherto wealthy countries were transformed into less or least developed countries. The high-standing that education had enjoyed could not be restored even after independence. After Independence, the countries of South Asia struggled to revive their economies and re-build their educational systems. Economically, one can argue that all the countries in this region are still “developing”, though some like India have experienced high rates of economic growth over the last two-three decades, and have been aspiring to enter the group of “upper middle-income” countries soon, and eventually the group of “high-income countries”. At the same time, whatever has been the economic growth and modernization, they got confined to a few urban areas in the countries, showing a very high degree of rural-urban inequalities within the countries.

Countries in this region have had more than their share of strife. Political turmoil, wars with neighbors, internal civil war and unrest and coups have not been unknown. These are in addition to ethno-religious and sectarian type conflicts, some of which have cross-border implications. Some of these are actually legacies of the colonial rule. Today the region is also one of the tensest regions in the world with two nuclear power nations at loggerheads all the time. Politically the region is marked by federal and unitary structures; multi-party political democracies, martial rule and also constitutional monarchies. It is also characterized by a high degree of religious diversity--the region has the largest population of Hindus, Muslims, Sikhs, Jains and Zoroastrians in the world. Apart from accounting for more than 90% of the world's Hindus and Sikhs, the region is home to 31% of the world's Muslims, 35 million Christians and 25 million Buddhists. The region is also one of the most linguistically diverse areas of the world containing four major language families and more than 650 individual languages, though only a few dozens of languages remain “functional”. With respect to political administration, decentralization has become an important development strategy not only in large countries such as India and Pakistan, but also in relatively smaller ones like Nepal.

The poverty ratios are high, in many countries of the sub-continent, with many countries following divergent strategies to reduce penury, with varying degrees of success. On the whole, South Asia has some of the worst human development outcomes of the poorest population quintile in the world. Except Sri Lanka and the Maldives, all the countries of this region score badly in the overall Human Development Index, though in the total list of 189 countries in the world, they are described as “medium human development countries” (Table 1). A special mention may be made of Bhutan, a tiny Himalayan kingdom, which is best famous internationally for its development philosophy of “Gross National Happiness” initiated in early 1970s.

Though, as has been stated earlier, South Asians have had held education in high esteem from the ancient period onwards, that regard was lost till it was partially revived by the famous “human investment revolution” propounded by Schultz (1961) in the 1960s that led many countries, both developing and developed, to re-recognize the undisputed role education has in the socio-economic development of a nation, and accordingly to invest in education, and to build or revisit their education policies with the objectives of improving access to better quality and more equitable education. It has been clearly recognized that human capital produced through investment in education (and training) is an important factor in the economic growth of South Asia countries, along with physical capital and technology. Estimates of growth accounting between 1981 and 2010, suggest that human capital has contributed directly as a productive input to about 22% of the annual gross domestic product (GDP) per worker in India. During the same period, it has contributed around 21% in Bangladesh and 16% in Sri Lanka (Asian Development Bank, 2017). Even though the momentum that was created by “the human investment revolution” in economic thought suffered a break in the 1970s, the following decades witnessed the emergence of “new growth theories” (in the 1980s) that have stressed again the role of human capital in economic growth. As a result, public interest in education increased significantly, along with a rise in the commitment to universal education. This was also partly attributable to the various global initiatives

Table 1 Human development index of the countries in South Asia.

	1990	2000	2010	2019	Rank in 2019
High Human Development					
Sri Lanka	0.789	0.691	0.754	0.782	72
Maldives	0.555	0.622	0.685	0.74	95
Medium Human Development					
Bhutan	0.236	0.520 ^a	0.574	0.654	129
India	0.439	0.495	0.579	0.645	131
Bangladesh	0.318	0.478	0.557	0.632	133
Nepal	0.273	0.453	0.537	0.602	142
Pakistan	0.423	0.447	0.512	0.557	154

^a2005.

Source: UNDP (various years).

of the United Nations and others and the new thinking that has crystallized in the name of human development. Education's instrumental role as a means to socio-economic development, and its constitutive role as a measure in itself of development have been recognized (Sen, 1999). All this has led to a colossal explosion of education in the South Asian countries as also in a few other developing nations.

Beginning with literacy rates, all major indicators of education development point toward progress in all the South Asian countries after independence. It may have to be underlined that progress was made and the structures were built mostly on the firm base established by colonial rule. Indigenous systems have been completely lost or are at best existing as fringe elements in certain parts. Modern education having replaced the traditional ones, began to grow fast in all the countries of the region. South Asian countries truly witnessed in general an "education explosion". Adult literacy in the region almost doubled from 36% in 1970 to nearly 73% by 2020 (Fig. 1). Enrollments in every level of education have grown exponentially over the years. However, the growth has not been uniform, as shown in Table 2. Forty percent of adult population in Pakistan are still illiterate. Enrollments increased by 2–5 times between 1981 and 2018 in all countries except in Sri Lanka, where the growth has been modest, having large numbers in the base year itself.

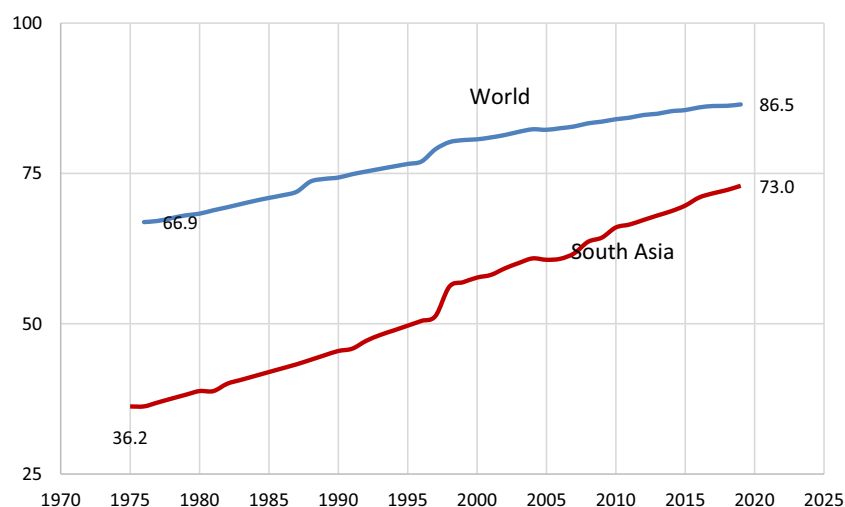


Fig. 1 Growth in adult literacy rates (%): world and South Asia. Source: World Bank (2021).

Table 2 Growth of education in South Asia.

	Adult Literacy Rate (%)		Enrollments (in millions)							
			Primary		Secondary		Tertiary		Total	
	1981	2018/19	1981	2019/20	1981	2018/19	1981	2018	1981	2018
Afghanistan	18 ^b	31 ^e	1.02	6.77	0.033	3.06	NA	0.015 ^f	1.05	9.83
Bangladesh	35	75	8.29	17.6	2.35	15.71	0.244	3.695	10.88	37.01
Bhutan	53 ^d	67 ^h	0.035	0.094	0.004	0.076	NA	0.013 ⁱ	0.039	0.183
India	41	74	73.87	121.7	32.75	130.93	3.545	35.148	110.17	287.88
Maldives	92 ^c	98 ^g	0.034	0.049	NA	0.021	NA	0.012 ^d	NA	0.082
Nepal	21	68	1.39	3.97	0.32	3.46	0.051	0.438	1.761	7.868
Pakistan	26	58	5.47	23.59	2.17	14.19	0.157	1.878	8.00	39.30
Sri Lanka	87	92	2.13	1.73	1.29	2.73	0.044	0.326	3.46	4.79
South Asia	69	74	92.24	173.72	39.27	170.27	0.244	41.525	131.75	343.99

^a1971

^b1979

^c1985

^d2005

^e2011

^f2014

^g2016

^h2017

ⁱ2020

Source: World Bank (2021) and national sources.

Universalization of elementary education

Post-World War II, education was envisaged as a universal fundamental human right according to the *Universal Declaration of Human Rights* (1948). Article 26 of the Declaration made it clear that elementary education was to be free and compulsory and education, including higher education, was a human right. Among the South Asian countries, India, Afghanistan and Pakistan (that included the present Bangladesh) were signatories to this Declaration, along with 45 other countries. The 1990 Convention on the Rights of the Child went a step further and imposed a legal obligation on this right. Colonial India, which then included Pakistan and Bangladesh, attempted as far back as in 1910 to make primary education free and compulsory throughout the country. But this assurance could only be implemented after Independence when it was incorporated in the Constitution of India in 1950 as a Constitutional Directive which, after a constitutional amendment, was made a Fundamental Right in 2002. This right was incorporated into an Act in 2009 under the name of the Right to Free and Compulsory Education, briefly known as the Right to Education. Today, all of the South Asian countries recognize education as a fundamental right: Sri Lanka declared education as a fundamental right in 1945, years before Independence; Bangladesh in 1993; Nepal declared it in its Constitution of 1990 and via the seventh amendment in 2001; India did so in 2002 and Pakistan in 2010. The Directive Principles incorporated in the Constitutions of Bangladesh and Pakistan, like the Constitution of India (1950), aim at free elementary education but it does not carry any legal compulsions or implications. Other global initiatives, such as the World Declaration on Education for All (EFA), a UNESCO-led movement in 1990, followed by the Millennium Development Goals (MDGs), and Sustainable Development Goals (SDGs) have had an important effect on the growth of elementary education all over the world, including in the countries of South Asia.

Historically, Sri Lanka was way ahead of other countries in the education sector. As a result of early initiatives, it has achieved impressive enrollment levels in primary education. The most striking aspect seen in Sri Lanka, which like Kerala in India, is, despite a low per capita income, it could universalize basic education, ensure equity in education and better its performance with respect to indicators in human development (Tilak, 2001). The goal of education in Sri Lanka has been to promote national and social cohesion. India has also been able to raise its net enrollment ratio in primary education to above 90%. Measures of affirmative action implemented by governments, for instance, programmes like the mid-day meal scheme and the Integrated Child Development Scheme in India have proved strong incentives for children to come to school, helping them improve their academic performance. After the launch of the Education for All movement in 1990, between 1990 and 2000, “countries in the region have made good and also rapid progress on many fronts in quantitative terms—in terms of literacy, increase in enrollments, enrollment ratios (gross and also net), retention rates, gender parity, teachers and public expenditures” (Tilak, 2000; see also UNESCO, 2009). Further, as the *South Asia Sub-Region Synthesis Report* (Tilak, 2000) stated, overall the South Asia has shown remarkable progress in achieving the millennium development goal of universal primary education; and much of the credit for this goes to India as, enrollments rose in South Asia from 72% to 89% over the period 1990–2010, largely as a result of progress India made. Currently, enrollment figures in primary education in South Asia have been impressive with an above 90% net enrollment ratio, except in Pakistan, where it has been below 70% in 2018 (Table 3). In case of Afghanistan no information is available on net enrollment rate for recent years; but with 107% gross enrollment ratio in 2019, one can expect a reasonably high net enrollment rate. All South Asian countries, with the exception of Pakistan, certainly seem to be moving toward universal primary education.

Teacher has always been regarded as the pivot of an education system. Globally, school systems employ trained and qualified teachers. But over the years, particularly since the beginning of the 1990s, when most countries adopted structural adjustment policies, constrained by budgetary resources and unavailability of qualified teachers, almost all South Asian countries resorted to recruiting teachers who had been under qualified and under trained. Such a practice was also influenced by managerial advantages of recruiting such teachers. Though over the last couple of decades, efforts have been made to improve the situation resulting in a slight increase in the proportion of qualified teachers available in primary schools, the improvement still remains far below the needed requirement, except in Bhutan (Table 4). Sizable numbers of teachers in countries like Bangladesh remain under-qualified. On the

Table 3 Net and gross enrollment ratio (%) in primary education in South Asian countries.

	Net enrollment ratio		Gross enrollment ratio	
	Year	Ratio	Year	Ratio
Afghanistan	1993	27	2019	107
Bangladesh	2010	91	2018	116
Bhutan	2018	88	2018	100
India	2013	92	2019	97
Maldives	2017	95	2019	98
Nepal	2017	96	2019	142
Pakistan	2018	68	2019	95
Sri Lanka	2018	99	2019	100
South Asia	2018	88	2019	107

Source: World Bank (2021).

Table 4 Proportion of teachers with the minimum required qualifications in primary education (%).

	2014	2015	2016	2017	2018	2019	2020
Bangladesh	–	47.6	50.4	50.4	–	–	–
Bhutan	–	100	100	100	100	–	100
India	–	–	69.5	69.8	–	73.1	–
Maldives	86.1	82.8	85.7	90.1	91.9	88.8	–
Nepal	93.6	94.4	97.0	97.3	–	97.3	–
Pakistan	84.0	82.5	–	–	78.0	76.9	–
Sri Lanka	–	86.3	85.0	85.5	83.1	–	–

Source: UNESCO Institute of Statistics (2021).

whole, there is a huge shortage of teachers in South Asian countries. According to the estimates of UNESCO, there is a requirement of 15 million teachers in the region over and above teacher replacement by 2030.

While the improvement in literacy levels, enrollment figures and accessibility in primary education is certainly no mean achievement for South Asia, countries in this region have responded differently to meet the challenge of universal primary education. For most there have been challenges on the social, cultural and political front resulting in a lacuna in the progress in the case of a few key indicators. Maldives, Sri Lanka and Bhutan have shown exceptional progress in most indicators; India, Bangladesh and Nepal have shown moderate progress while enrollment and literacy levels in Pakistan have been alarmingly low. Further, research points to glaring failures in the efficiency and quality of educational progress; one still notices continuing high numbers of out-of-school children, high drop-out rates, low completion rates and poor age-specific educational outcomes. The outcomes and progress of the educational sector as a whole and specifically primary education in South Asia leave a lot to be desired.

Out-of-school children is a phenomenon that plagues countries in this region acutely with approximately 22% of the world's out-of-school children living here. These children, in the age group 6 to 14, have never even been enrolled in elementary education, or have been absent from school without prior intimation for 45 days or more in a year. Out of the 12 million out-of-school children in South Asia, ten million are in India and Pakistan. "Perhaps most strikingly, Pakistan has the highest number of out-of-school children worldwide after Nigeria" (Hunter, 2020). Post-2000 there has been about a 56% decline in such school children in South Asia. There has been a decline in primary school age children, from around 43% in the 1970s to seven per cent in 2019. This decline has been pronounced beginning 2000 when the ratio was 21% and declined to eight per cent in 2007. In 2020, in India and Pakistan, as many as 17% girls in the age-group 10–17 were not in school (Béteille et al., 2020).

The primary completion rate, another standard indicator, which is defined as the number of students in the last grade of primary school minus the number of repeaters in that grade, divided by the total number of children who are of an official graduation age, is widely considered as a measure of efficiency of primary education. Interestingly, the ratios are high in the region, except in Bangladesh and Pakistan, and are higher in the case of females except in Pakistan (Table 5).

While many indicators show that there has been reasonably good to impressive progress in terms of accessibility to education, increased schooling has not translated into any significant improvement in learning. Learning outcomes are hugely unsatisfactory with more than 50% of pupils in Bangladesh, Pakistan and India recording below minimum reading proficiency levels even after completing primary school education, as the annual reports of Pratham Foundation in India and citizen-led assessments in South Asian (and other) countries reveal. The citizen-led assessments indicates that only 46% of children in Pakistan's public schools and 48% in India's public schools can read a Grade II level text by the time they reach Grade V (ASER Pakistan, 2015; ASER India, 2016). In India, Pakistan, Afghanistan, Bangladesh, and Nepal, less than half the number of students who leave primary school are equipped with the basic foundational skills that one would expect them to acquire by the end of primary level. Summing up available evidence, the Asian Development Bank (2017) has stated that a large share of primary school graduates in South Asia lacks basic

Table 5 Completion rate in primary education (%).

Country	Year	All	Males	Females
Afghanistan	2019	84	102	66
Bangladesh	2010	68	64	72
Bhutan	2020	82	76	88
India	2020	95	93	96
Maldives	2019	92	90	93
Nepal	2019	120	118	123
Pakistan	2019	73	78	68
Sri Lanka	2017	102	100	99
South Asia	2019	92	91	93

Source: World Bank (2021).

numeracy and literacy skills; and the [World Bank \(2018\)](#) has estimated that about 120 million youth (age 15–24 years) in South Asia lack foundational learning skills. Children who do not acquire the basic foundational skills of literacy and numeracy by the age of ten, are feared to fall into ‘learning poverty’ ([World Bank, 2020a](#)), a pre-cursor to other forms of poverty. In short, with a chronic quality deficit, South Asia is in the midst of a learning crisis: with over one-third of its students having no grade-level competencies which is a serious challenge. Recruitment of quality teachers in required numbers is essential to break the “low-quality education equilibrium” that the region is characterized with ([Richards et al., 2022](#)).

Secondary education

With the rapid growth in enrollments in primary education and improved efficiency rates, secondary education can be expected to grow fast. So obviously there is an impressive growth in enrollment ratios in secondary education in the countries of the South Asian region. The gross enrollment ratio in Sri Lanka was above 100% in 2019. At the regional level, the corresponding ratio is lower when it is compared with the world average of 76%, with Pakistan having the lowest enrollment ratio at 44%. But other countries are catching up with the world average and are even performing better ([Fig. 2](#)).

The net attendance rate is a better indicator of access to education as it measures the total number of students in the given official age group for a given level of education who are attending the given level of education, expressed as a percentage of the corresponding age-group population. Statistics on secondary education are available separately for lower secondary (generally grades IX and X) higher/upper secondary level (grades XI and XII). Net enrollment rates in the lower secondary level are higher than in the upper (higher) secondary level in all countries of the region, as shown in [Fig. 3](#). Both Nepal and India are ahead of Pakistan. A drop in net enrollment ratio is evident as we climb from the lower secondary to the higher secondary level in India, Nepal and Pakistan. In addition, [Fig. 3](#) highlights gender differences which are anyhow, not very high, except again in Pakistan.

Recognizing that secondary education has failed to provide any job-relevant skills, and has served only as a transitory phase toward higher education, it is not considered a viable terminal level of education in these nations ([Tilak, 2020](#)), and many countries have planned for significant expansion of vocational education streams at the secondary education level. While secondary education has expanded somewhat impressively in South Asia, the attempts to expand vocational/technical education at the secondary level have not been successful, with only 3.5% of upper secondary school students of South Asia being enrolled in vocational programmes as against the 21.7% world average. Vocational and technical education has failed to attract students in good numbers as policy plans have projected vocational and technical education programmes as cheap alternatives essentially meant for poor and academically less meritorious students. In fact, meager investments have been made on vocational secondary education in South Asian countries ([Tilak, 1988a, 2003b](#)). Some countries like India have started vocational educational programmes as add-ons in higher education.

Higher education in South Asia

For a long time, the discourse on strong linkages between education and economic development, reflected in the estimates of rates of return to education, has poignantly been argued, for raising priority for primary education especially in developing countries, relegating secondary and higher education to the background. As a result, secondary and higher education have severely suffered for several decades although ironically primary education too did not receive adequate funding. It is only recently that some changes have been observed. The importance of human capital has been recognized as an essential element in developing a knowledge

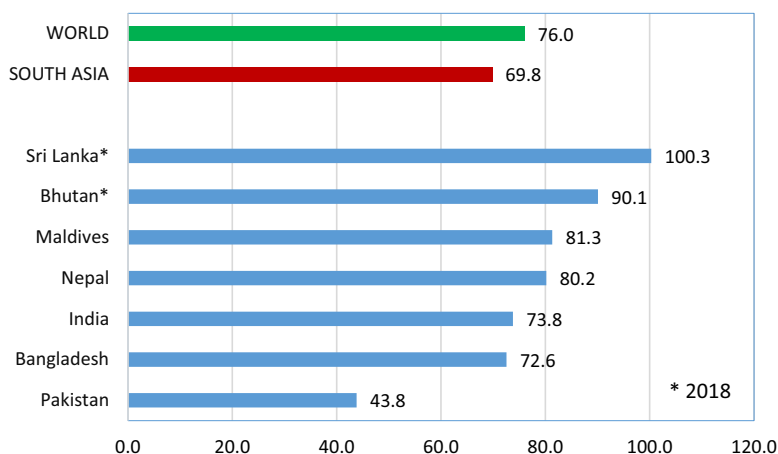


Fig. 2 Gross enrollment ratio in secondary education in South Asia (2019). Source: UIS (2021).

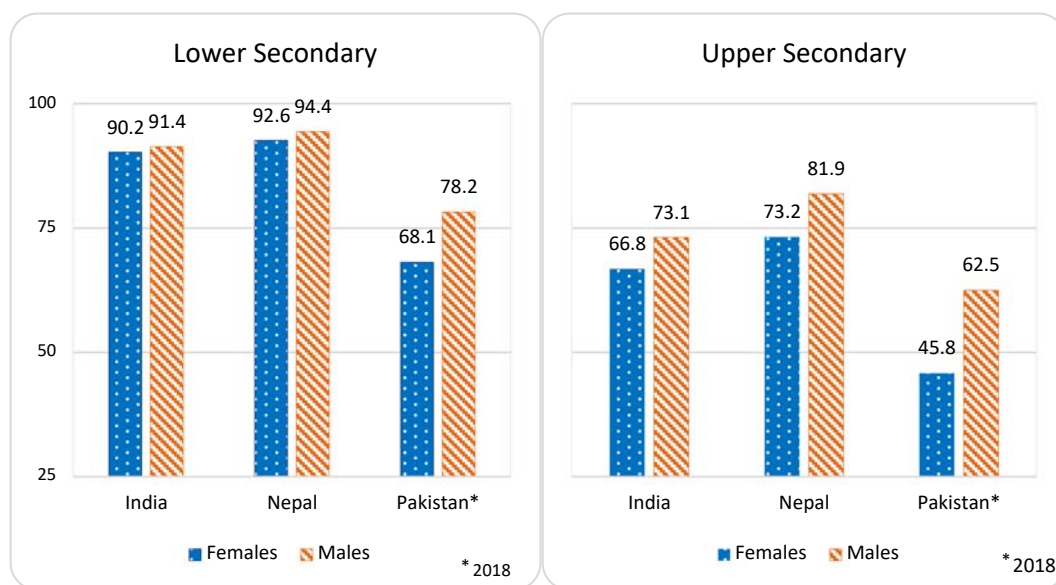


Fig. 3 Net attendance rate in secondary education in India, Nepal and Pakistan by gender (2016) (%). *Source: UIS (2021).*

economy which relies on the creation and use of intellectual capital, a specialized human capital, which in turn depends on higher education, including technical and professional higher education. This form of human capital is believed to be a significant determinant of economic output and growth, arresting diminishing returns to scale. As a result of this new understanding, since the late 1990s, higher education has grown substantially in South Asian countries as it has been recognized as essential for any country to be successful with globalization policies.

The higher education landscape in South Asia has been marked by a number of varying trends, perhaps under pressure from globalization. Tilak (2006b) has identified six global trends in higher education: (i) the rapid expansion of higher education; (ii) a decline in public subsidies for higher education; (iii) an increase in cost-recovery, particularly through student fees and student loans, and the generation of funds from the corporate sector and other non-State segments of the society; (iv) a neglect of the liberal arts and humanities, sciences and social sciences, (v) the large-scale adoption of new market modes of higher education including privatization, commoditization and commercialization; and (vi) internationalization of a new type. South Asian countries have also experienced similar trends in varying degrees.

Growth of higher education

There has been a significant rise in absolute numbers in South Asia especially with regard to enrollments, the increased number of colleges, universities, and in teachers available for higher education. As Bêteille et al. (2020) have estimated in South Asia, enrollments in higher education have grown at a staggering 387% over the past 15 years—from 11 million students in the early 2000s to 42 million students in 2019/20. The average annual growth rate over the period was 14%. The Maldives has experienced the highest rate of growth (35%), while Sri Lanka and Bhutan, the lowest (6%). The growth rate was steepest in the first decade of the 21st century, but even today higher education is growing rapidly and trends indicate that it is unlikely to slow down. The gross enrollment ratio in higher education in the region as a whole has increased remarkably by six times from four per cent in 1970 to 25% in 2019. In 2019, the corresponding enrollment ratio in the world is 39%. While the ratio is above 20% in India, Bangladesh and Sri Lanka, Pakistan, with an enrollment ratio of 8.9%, is almost at the bottom.

However, these figures given in Fig. 4 relating to South Asia are much lower than the corresponding figures of other Asian countries like China (50%) or South Korea (95%). India, it can be argued, has been a champion in this area, having developed the second largest higher educational system in the world, with nearly a 1000 universities, 40,000 colleges and 37 million students; the gross enrollment ratio in 2019 was 29% (Fig. 5). Bangladesh and Pakistan stand at a distant second and third positions with regard to the size of the system and enrollments. Student numbers are very small in Sri Lanka and Nepal. Bhutan has only just two universities, one of which is a medical one and the Maldives has none. As a result, hardly 0.4% of the Maldives' adult population has completed higher education. The percentage of population aged 25+ years with tertiary education varies between 9.5% in Sri Lanka, and 6.1% in India, on the one side and 2.3% in Nepal and 0.4% in the Maldives on the other end as shown in Fig. 6. On the whole, the ratio is very low in South Asia as compared to other advanced regions of the world.

The expansion of higher education in South Asian economies however, has been more in the area of general higher education. Bangladesh, for instance, in 2019 has 88% of its graduates in fields other than science, technology, engineering or mathematics.

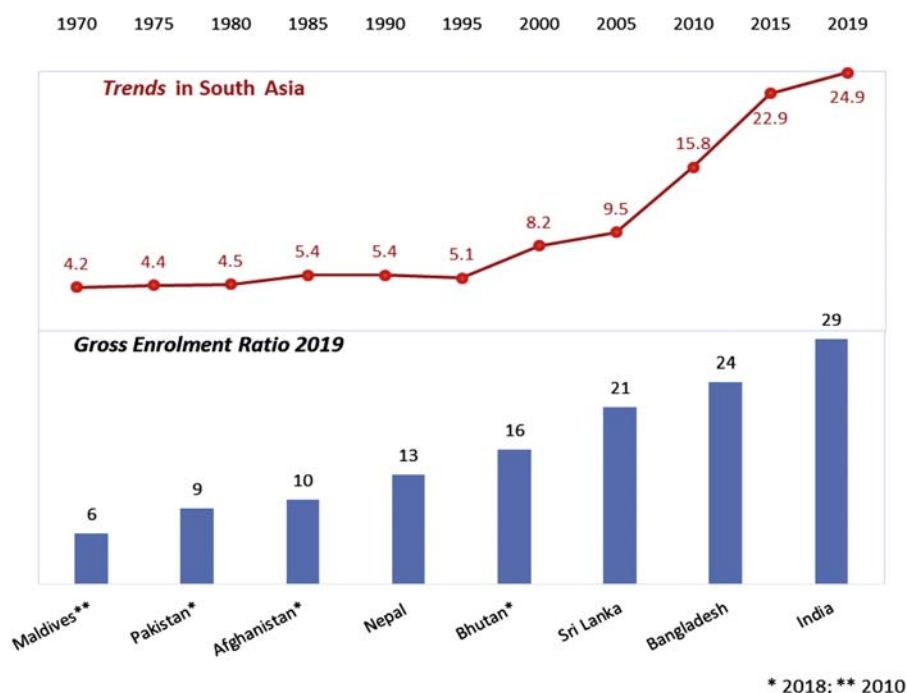


Fig. 4 Growth in gross enrollment ratio in higher education in South Asia. *Source:* UIS (2021) and World Bank (2021).

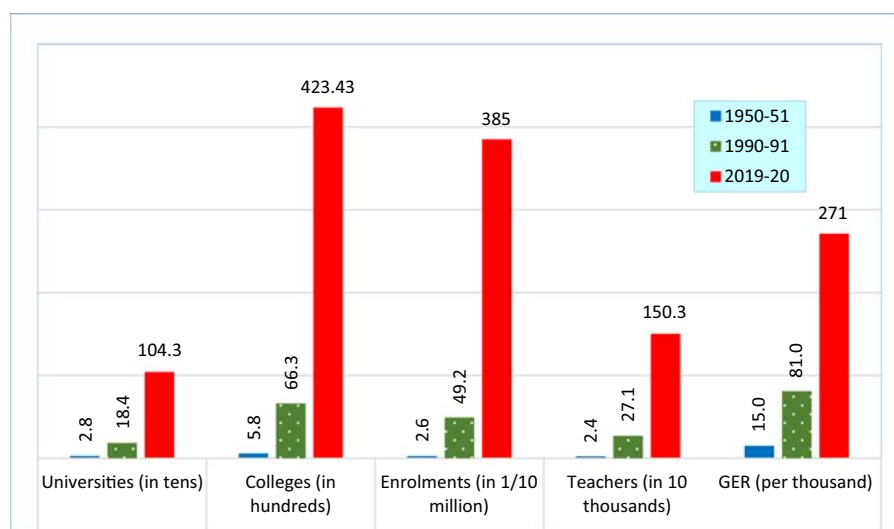


Fig. 5 Growth of higher education in India. *Source:* Education in India 1950–51, Statistics on Higher and Technical Education in India (1990–91) and All India Survey on Higher Education, 2019–20 (Ministry of Human Resource Development, Government of India, New Delhi).

The figure for India is 68% and for the Maldives, 95%. This does not however imply that a higher value is placed on streams of education such as humanities or the liberal arts; the reality is that the exorbitant cost of acquiring professional education means that only certain privileged sections of society are able to access it and many from lower strata end up in less-expensive general higher education.

Globalization, which ushered a revolution in information and communication technologies (ICTs) as its prodigal offspring, has successfully packed the markets with educational courses offered in different forms and modes, with or without instructors and physical campuses. Internet-based study programmes and distance education have further helped the massification of higher education with students competing to acquire degrees that they hope will get them some jobs or others, irrespective of considerations like graduates' interest or merit. Most of South Asia is also facing a demographic dividend also known as the demographic bonus which refers to the likely economic growth that results from a significant increase in the skilled working-age

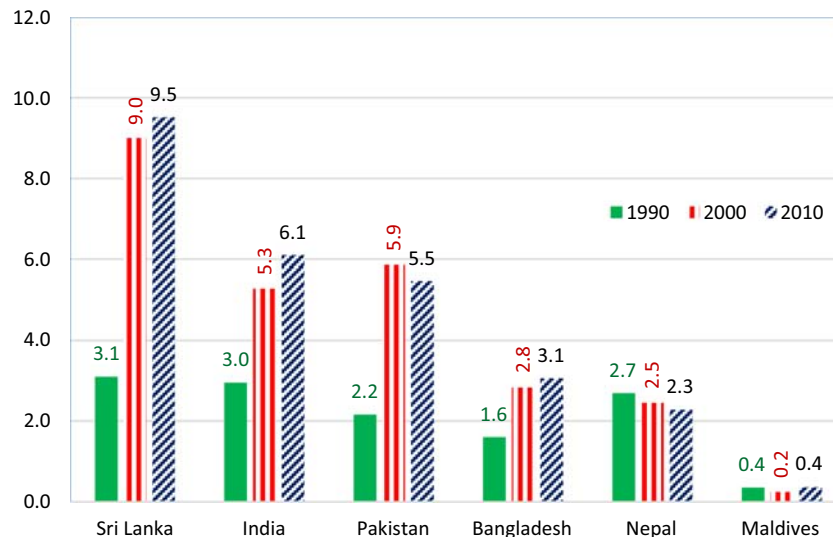


Fig. 6 Rate of higher education attainment in South Asia. (Percentage of population aged 25+ who have completed tertiary education). *Source:* World Bank (2021).

population (age 15–64). For example, 30% of Indians are under the age of 19 while Pakistan reports that 64% of its population is below the age of 30, and 29% between the age of 15 and 29; Bangladesh, Pakistan and Nepal have all the indications of a “youth bulge” (Ahmad, 2018; Trines, 2018b). In the face of a severe resource crunch and demographic pressures to meet the rising demand for higher education, governments have resorted to open and distance learning modes. Massive open online courses (MOOCs) have not only allowed access to higher education being provided by universities within the country at a cheaper cost, but have also enabled students to access universities abroad from any part of the country, without the hassle of travel and associated expenses, as long as their location, however remote, has internet access. Around 16,000 MOOCs have been launched by 950 universities all around the world (Shah, 2020), barring China. Distance education accounts for 11.1% of the total enrollments in higher education in India in 2019–20 (GoI, 2021). India is having the world’s largest open university, the Indira Gandhi National Open University with a student population of about ten million that far exceeds the population of some European countries. In addition, there are 15 state open universities and 110 dual mode universities. The questionable quality of the courses, however, implies that dreams of better prospects do not necessarily materialize into decent employment options for the marginalized. The disadvantaged often fail to catch up with the privileged who enjoy a better standard of education and have a natural advantage in terms of their social capital. So there is a crisis of graduate unemployment in South Asia. According to World Bank data, around 15% of the total labor force in India with advanced education is unemployed (2019); in Bhutan the figure is 18% (2013) and in Bangladesh, 11% (2017).

Inequalities in higher education

The higher education system in South Asia is also characterized by a high degree of inequality. As Tilak (2003a, p. 2274) notes, “the regional, social and economic inequalities that are a glaring feature in the societies of South Asia, are reflected in the educational systems, with the poor and socially backward areas suffering a severe degree of exclusion from education.” The enrollment and attendance figures in higher education reveal high levels of inequality across gender, income groups and rural and urban backgrounds. For South Asia as a whole the gross enrollment ratio for females was 3.5% in 1990 while for males it was 7.2%. The gender gap has narrowed over the years in favor of females with a gross enrollment ratio of 27% as against 26% for males in 2020. But the disparity against females continues and is pronounced in countries like Bangladesh where the gross enrollment ratio for males is 26% and for females is 20% in 2020. Pakistan also shows low and unequal enrollment figures in higher education in stark contrast with the Maldives which shows the highest gross enrollment ratio in South Asia with 67% among females, much higher than that of males at 18%. Indeed, the progress of the Maldives on certain socio-economic indicators have been impressive (World Bank, 2011). This tiny island nation has achieved five out of the eight Millennium Development Goals and today has achieved a MDG+ status. However, the *Education Sector Analysis Report* (Government of Maldives, 2019), shows that the apparent gender parity level achieved in higher education could be misleading: a high number of females have enrolled in hairdressing and beauty courses introduced in polytechnic colleges and in online nursing programmes. Further, the overall enrollment has increased due to the introduction of multiple modes of delivery initiated by the private sector. On the whole, it must be stated that women fare better than men in South Asia, except in Afghanistan and Bangladesh; and in all the countries of the region there has been significant improvement in gender inequality in higher education (Table 6).

Table 6 Gross enrolment ratio and gender inequality in higher education in South Asia, 2020.

	Gross enrollment ratio (%)						Gender inequality index ^a		
	1981			2020			2020	1981	Improvement
	Male	Female	All	Male	Female	All			
Afghanistan	3.2	1.2	1.7	15	6	11	0.818	1.176	0.358
Bangladesh	5.2	0.9	3.1	26	20	23	0.261	1.387	1.126
Bhutan	1.3	0.4	0.7	16	17	16	-0.063	1.286	1.348
India	7.4	2.9	5.2	28	31	29	-0.103	0.865	0.969
Maldives	NA	NA	NA	18	67	34	-1.441	NA	NA
Nepal	4.3 ^b	1.3 ^b	3.5 ^b	13	14	13	-0.077	0.857 ^b	0.934
Pakistan	3.1	1.2	2.2	12	12	12	0.000	0.864	0.864
Sri Lanka	3.2	2.4	2.8	17	27	22	-0.455	0.286	0.740
South Asia	6.7	2.3	4.7	25	27	26	-0.077	0.936	1.013

Note: ^aenrollment ratio of male minus ratio of female, as a ratio of 'All'.

^b1983

Source: World Bank (2021).

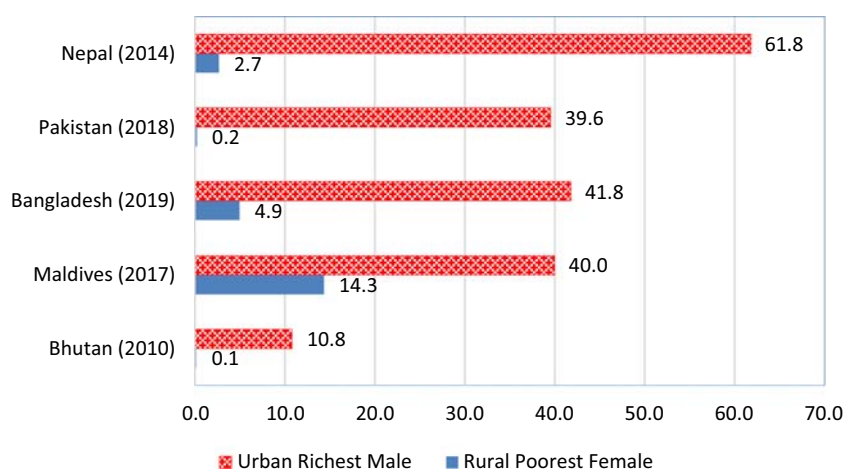


Fig. 7 Inequality in gross attendance ratio in higher education in South Asian countries (Ratio among urban richest male vs. rural poorest female). Source: World Bank (2021).

But when one examines gender-wise ratio by economic classes and by rural and urban regions, inequalities sharpen. For example, in all countries of the region the enrollment ratio is highest for urban males belonging to the richest quintile and lowest among rural females belonging to the poorest quintile. The difference between these two groups is as high as 198 times in Pakistan and 108 times in Bhutan, while it is only 8.5 times in Bangladesh as depicted in Fig. 7.

Such a disparity reveals a gross neglect of higher education as a means to achieve equitable development levels. Examining the problem of unequal access to higher education in South Asia and Sub-Saharan Africa, Ilie and Rose (2016, pp. 446–447) have observed:

... While gender inequality continues to be a cause for concern, the magnitude of gender gaps is smaller than that of wealth-driven gaps. However, given that countries with higher rates of higher education participation also tend to have wider gender gaps amongst both the rich and the poor, there is a danger that as higher education systems expand in countries with lower levels of access, so could the gender gaps. This suggests the need to put in place strategies to address gender and wealth inequalities early on. As such higher education attendance rates are overall very small, the focus should be on measured expansion that is pro-poor and tackles any gender imbalance from the very beginning.

Despite high rates of return (Table 7) to higher education, in comparison to the other levels, there has been an insufficient expansion of higher education, specifically seen in the low participation of females and disadvantaged sections of the society, like the rural and poor, in higher education.

Table 7 Average returns to education in South Asia, by level and gender.

	Primary	Secondary	Higher
All	6.0	5.0	17.3
Males	4.7	3.9	16.6
Females	4.8	6.2	23.3

Source: Montenegro and Patrinos (2014).

Policy dynamics: grappling with globalization

The current wave of globalization which began in the 1980s, which includes economic globalization, cultural globalization and political globalization, saw an important phase in the reform of education in South Asia. Here countries have been compelled to grapple with forces at the political, economic and cultural levels that have pushed them to “adjust”, “readjust”, “adapt” and “innovate” in the policy space. Globalization and associated market forces have raised a few basic questions such as: is education a public good or a private good? Who should take the responsibility of providing education—the state or the market? Why should education be provided free? What is wrong with profit-making in education? How much should students be charged for their education? At what level can education be largely driven by market forces and at what level is it necessary for the government to intervene? While some of these questions have never been asked for centuries, if not millennia, some issues have been settled, but they have re-emerged again. Generally, as no clear answers could be found, governments have tended to adopt *laissez faire* in education which has been marked by non-intervention and an absence of public policy (Tilak, 2005).

One can identify four major trends that have either originated or been accentuated with the adoption of the policies of globalization. They are: privatization of education; emergence of English as a medium of instruction; internationalization of higher education and emergence of ranking systems; and changes in the financing of education, including the entry of foreign aid into the education sector. These are briefly discussed here.

Privatization of education

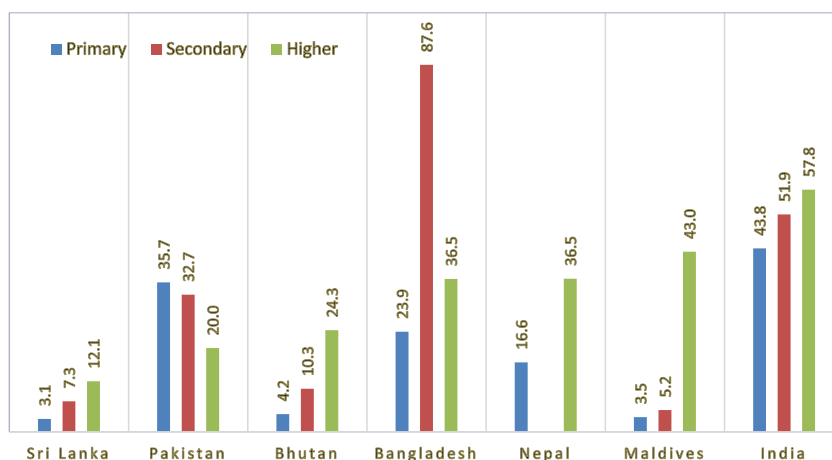
Globalization has meant a large-scale entry of the private sector into the area of education which hitherto had been almost a state monopoly, or had insignificant levels of a private presence. Privatization, which originated as a charitable initiative in the field of education, gradually became a natural consequence of globalization dovetailing with simultaneous cuts in the education budget by governments. The private sector has projected itself as a rescuer of the education system from this resource crunch. The deficit in public expenditure on education, attributable to a lack of resources in general and more importantly a lack of political will, on the one hand, and a rising private demand for education on the other hand, have provided enough space for the private sector to enter and exploit the situation to the maximum. Soon a huge number of private players entered every level of education in almost all countries, more so in developing countries; and the nature of private involvement changed dramatically. Philanthropic principles were being replaced by purely principles of profit (Tilak, 2006a). These new “principles” did help education to grow enormously, albeit in different forms.

The private sector is a major player in every level of education in many countries including specifically in India, Bangladesh and Pakistan, as shown in Fig. 8. According to World Bank estimates, private schools account for 44% of enrollments in primary schools in India, 36% in Pakistan, and 24% in Bangladesh. In Bangladesh growth of several “private or community, NGO and non-mainstream religious schools” has been encouraged by the state (Hossain et al., 2002). In a study undertaken on the growth of private education in India, Kingdon (2017) shows that almost 50% of primary school-going children in urban India and more than 20% in rural India go to private schools. Public private partnership has also emerged as a strong presence in India when the Right to Education Act (2009) made it mandatory for all private schools to give a minimum of 25% of admissions to disadvantaged children against a compensation of fees by the state to the schools, which is seen by many more as a pro-private policy than as an equity-promoting initiative.

Secondary education is heavily dominated by the private sector in Bangladesh with 88% of secondary school students going to private schools. The corresponding ratio is 58% in India. Though relative numbers tend to decline in Bangladesh at every level of education, they still remain high (Table 8). Further, the share of private enrollments are highest in tertiary education in the region (except for Bangladesh where secondary education has the highest private enrollments) and the trends continue upwards.

In the area of higher education, most South Asian countries, except for Sri Lanka, have shifted a substantial part of the responsibility from the state to the private sector. During the first two decades of the 21st century, the percentage of enrollments in private higher education has grown from seven per cent in Pakistan (in 2003) to 35%; in Nepal from 26% to 37%; in Sri Lanka it has doubled from six per cent to 12% in five years between 2014 and 2019. It is only in Bangladesh that one notices the proportion significantly declining from 63% to 37% between 2000 and 2019. In India, it has been high at around 58%. Some of these zig-zag trends and patterns can be noticed in Fig. 9.

The “laissez-faire” policy of the governments, or policy absence, and/or pro-private policies explain the growth of private education in the region. Taking advantage of the absence of or the presence of weak regulatory systems, private players have opened so far as many as 408 private universities in India a kind of which were completely absent until the end of the 1980s. In 2019, 78% of



Note: Some figures refer to different years. Maldives: Primary (2017); Pakistan: Primary (2018); Sri Lanka: Tertiary (2013); Bhutan: all levels (2018)

Fig. 8 Percentage of enrollments in private institutions by education levels (2019). *Source:* World Bank (2021).

Table 8 Growth in enrollments in private education in Bangladesh.

	Primary	Secondary	Higher
2000		95.7	62.7
2005	41.7	96.0	46.8
2009	41.0	95.3	43.0
2016	21.4	94.0	35.5
2019	23.9 ^a	87.6	36.5

^a2018.

Source: World Bank (2021).

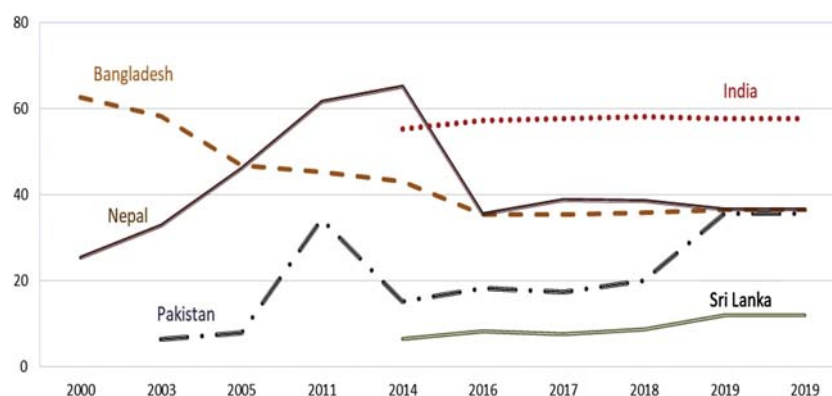


Fig. 9 Percentage of enrollments in private institutions of tertiary education. *Source:* World Bank (2021).

colleges were in the private sector. In Sri Lanka there was widespread popular resistance to the establishment of private universities and it was only legitimized as late as 2013. Ever since the trend has been changing. Currently along with the recognized private higher education providers there is a sector of unrecognized providers that is flourishing. Affiliated with foreign universities, as permitted by Sri Lankan law, they offer degree programmes jointly with foreign providers.

On the whole, South Asia has suffered from a lack of well-defined government policy in these dimensions. In the absence of a clear government vision and directives, the private sector has expanded recklessly and created a system that is infested with a dubious quality of education, exorbitant fees, rogue institutions providing fake degrees and a neglect of certain streams in the liberal arts and the need for developing “holistic” global citizens. All this weakens the linkages with the employment market. Yet they survive and even flourish by adopting cheap and dubious marketing strategies. These private institutions almost

exclusively rely on student tuition fees not only for their operating costs, but also for their development needs. The almost extortionist fees charged by these institutions result in an unequal access to education. For instance, in Pakistan the highest fees charged for a semester by private higher educational institutions can be as high as US \$3106 as against the US \$310 to US \$585 charged by public institutions (Hunter, 2020). Similarly, in India the cost per student varies between US \$4520 to US \$90,400 for a 2–4 year, “self-financing” professional degree course in public universities that are comparable to the fees charged in private institutions (ADB, 2012). For comparable programmes, public institutions charge a fraction of these figures. The high levels of fees charged by the private institutions drive students away from higher education, or force them to go abroad seeking admission in relatively less known public universities in small countries in Europe or Central Asia where they also struggle to adjust to difficult and strange cultures and conditions. Nepal’s higher education system had until 1986 one state university, Tribhuvan University funded by the government. Universities that have been established subsequently in Nepal are also owned by the government, but the major difference is that the new universities are required to charge high fees to cover their operating expenses; thus they function *de facto* as private institutions. Nepal also has an unusual system of public universities with privately owned campuses or affiliated colleges. In 2015–16, they numbered 777 as compared to community campuses (532) which receive some grants from the government or constituent campuses (98) that are entirely government funded.

The private institutions are also characterized by a lack of diversity in not only student population, as they cater to the needs of relatively better off sections of the society, but also by a lack of diversity in academic disciplines. Concentrating on revenue-generating disciplines like commerce, business management engineering, information technology, etc., they ignore liberal arts, humanities, social sciences, and even life sciences and physical sciences, which are important for development of a balanced higher education system on the one hand and of a humane society on the other.

An important dimension of private higher education institutions, apart from the fact that they treat education as a private good like any other commodity in the markets and charge exorbitant fees, is the questionable quality of the education they impart. Many private institutions don’t comply with state rules and regulations in terms of quality assurance, transparent admissions and the fair recruitment of faculty. Since making profits is the primary goal of private providers, and in the absence of any proper strong regulatory mechanisms, the foremost goals of equity and quality are relegated to the background and all other educational considerations including most importantly the public good characteristics are completely overlooked.

Globalization and medium of instruction

Globalization has also meant setting global uniform standards for education and moving away from local and regional knowledge systems. For instance, English as a medium of instruction in schools has become extremely popular and there has been a sharp and steady rise in private schools in South Asian countries offering English medium education, particularly in India, Nepal and Pakistan. Of course while the roots of ascribing a higher value to the English language can be seen in the colonial regime, these tendencies have become more entrenched with increased international interaction through growing globalization and liberalization. English language skills have been viewed as a powerful tool of globalization, growth and prosperity by individuals as well as collectively, by the nations. It is assumed that language not only increases the “cultural capital” (Bourdieu, 1986) of a person but also reduces social inequalities and poverty by increasing the overall employability, productivity and efficiency in a globalized labor market. English is ideologically perceived as a means of acquiring “linguistic capital,” which is often believed to be providing wide access to the global labor market. Modernization, employment in white collar jobs, and high wages are regarded as correlates of knowing English language. As a result, instruction in English in schools and colleges has become a dominant cultural paradigm (Sah, 2022) in South Asian societies. The discourse of a neoliberal language policy has led common people to perceive that an English medium education as the ideal one they aspire for their children in order to help them attain fast upward socio-economic mobility (Bhattacharya and Mukherjee, 2013; Erling and Seargeant, 2013).

Post-independence, most countries of the region adopted their local languages/mother tongues as medium of instruction and English as a subject in schools, which may still be pedagogically the best option. However, countries also got engulfed in controversies. Though there are five languages in Pakistan, Urdu was recognized as the official language, which has also been the medium of instruction in most of the schools. Sinhalese and Tamils in Sri Lanka insisted their own language to be the national official language and medium of instruction in schools. The debates on the medium of instruction are much more pronounced in Bangladesh as the country was formed with language identity as a key issue. A larger percentage of children studies in Bengali medium schools, and only a smaller percentage in English medium schools which are mostly private. However, the government policy allows, if not encourages, well-resourced schools to introduce English version of the national curriculum and to do away with the Bengali medium (Sharma, 2020). The problem is not so severe in other countries of the region, except in India, where initially a three-language formula was proposed which was never seriously implemented, and which was again proposed in the National Education Policy 2020 (GoI, 2020). On the whole, given the diversity in languages, “India has been particularly susceptible to the forces promoting [linguistic] fragmentalism” (Emerson, 1971, p. 181). Countries like India, Pakistan and Nepal have actually facilitated teaching of English soon after Independence, adopting a multi-lingual or a bilingual educational policy. In India the medium of instruction in both languages, mostly in local language, is offered in schools, while in colleges and universities the medium used to be English only for a long time, until the local language was introduced as an option. In most private schools, English is the medium of instruction while government schools use the local language. However, with globalization, many state governments have felt pressurized to offer English as a medium of instruction, along with the local language as an option to the students at all levels from pre-primary to higher in government schools and colleges. Though UNESCO’s and UNDP’s policy

programmes like the Education For all, and the Sustainable Development Goals (SDG 4.5.2) and constitutional provisions in countries like India, e.g., the Right to Education Act, clearly state that the local language/mother tongue/home language should be the medium of instruction at the primary level, many countries have ignored this noble vision, as English won the “language wars” in many South Asian countries (Chowdhury and Kabir, 2014). Though the non-availability of trained and qualified teachers in English is an important constraint many countries in the region face, they still seem to prefer a system with poor quality English teachers to one which has better qualified teachers proficient in local languages. It has been widely acknowledged that private schools thrive essentially because the medium of instruction is English. Sri Lanka introduced English as a medium of instruction only in 2001, viewing it as a means to prepare the younger generation to face “globalization” effectively (Wijayatunga, 2018). Similarly, Nepal has also seen an increasing trend in the growth of English medium schools, but in the absence of qualified teachers there is increasing criticism that students have not been able to develop sufficient skills either in their home language or in English, or in the subject (Sah, 2017). The emphasis given to English language is at the cost of local languages as an increasing number of students do not feel the need to learn local language or any other language. As Mehendale, 2020, p. 397) observes,

These shifts and the ensuing challenges need to be seen within contrasting forces of globalization and nationalism where policies aim at accommodating diverse imperatives while simultaneously aiming to develop a national identity through use of a common language such as English. Given the linguistic diversity and the pressures of globalization, maintaining a multilingual balance challenges the national education systems.

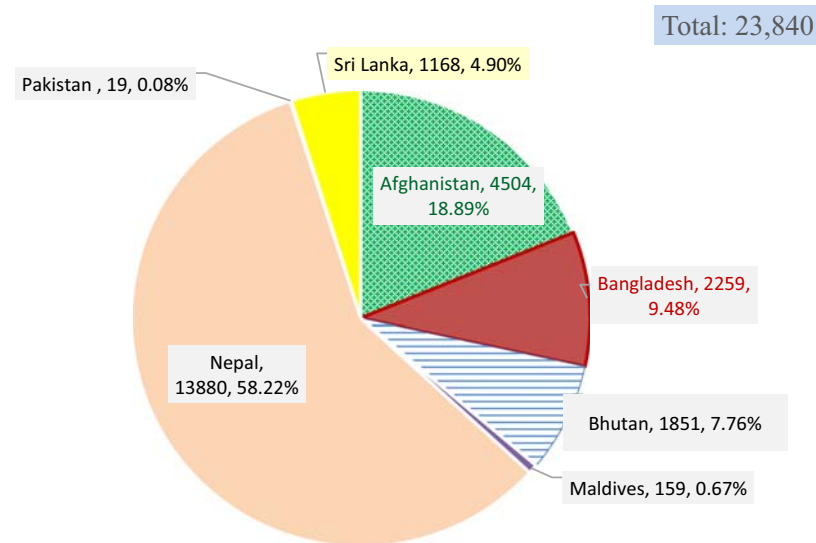
While in the post-colonial societies, the adoption of English has emerged as a form of neo-colonialism, in other countries, the adoption of English has become a neo-liberalist market response under globalization. As some rightly argue, “English medium schooling is a recipe for disaster for the learning of overwhelming majority of the disadvantaged, ... it will neither provide them basic communication skills in English, nor provide them education and social mobility, or higher order cognitive skills that will equip them to deal with challenges of the 21st Century ... [and this conclusion] is broadly applicable to the experience of many other former British colonies in Asia and Africa” (Kurrien, 2021).

Internationalization and ranking of institutions

A natural consequence of globalization has been the internationalization of higher education. Collaborations with universities worldwide and the exchange of students and faculty for teaching and research have been in practice for a long time essentially with a view to enhance the quality of the learning environment in university campuses. The same model, along with a few new ones like the setting up of campuses in foreign countries under new trade agreements, have made the internationalization of higher education a game changer in the global arena of education. The transformed models have altogether different objectives.

The inclusion of education in the General Agreement on Trade in Services (GATS) has opened up new avenues for better inter-connections among universities world-wide, by trading education. For developing countries of South Asia, however globalization has meant a strong influx of things and ideas from the global West; internationalization of universities has been more about emulating western standards, and grappling to survive in the global education market. Taking a “peripheral” role, universities in South Asia have been importers of education and the mobility of students has been predominantly one-way, from South Asia to the rest of the world.

Trade in education as defined by GATS occurs in four modes (Tilak, 2011): in the first mode, cross-border supply of education takes place without any movement of consumers or providers which includes all “e-learning” courses. As mentioned earlier, India has the second largest market for distance education after the United States. With increasing access to the internet, other countries like Nepal and Bangladesh have also sought refuge in foreign distance learning to increase access to higher education with limited resources. Private providers from the US, such as Coursera and edX, have colossal markets in these countries and local providers in South Asian countries are still way behind in competition. With respect to the second mode, viz., student mobility, one notes that the diaspora of South Asian students is one of the largest with the outbound mobility ratio, defined as the number of students studying abroad as a percentage of the total tertiary enrollment, which is 1.5% against the inbound ratio of 0.12% for the whole of South Asia. The highest outbound ratio in South Asia is 38% in the case of Bhutan, followed by Nepal, 20% and Sri Lanka, eight per cent. However, in absolute numbers, India is the second largest country after China with more than one million students studying abroad in 2021 as reported by the Ministry of External Affairs, India. This is projected to increase to 1.8 million by 2024. At the same time India is a popular destination for other South Asian students, and a sort of a leader in providing higher education, accounting for above half of the entire inbound student population from Asia (ADB, 2017; AIU 2017). As one can note from Fig. 10, the inbound students from South Asia number nearly 24 thousand in 2019–20; and they constitute about half of the all foreign students in India. Among the students from South Asia, who doubled in a decade from about 10 thousand in 2010–11, students from Nepal and Afghanistan account for more than three fourths (2019–20). Apart from proximity and culture, English medium education provided at comparatively lower cost has made India attractive to students from South Asia and also from the Caribbean and the Gulf countries which have a huge Indian diaspora. On the whole, with about 50,000 students from all countries in 2019–20, more than half of whom are from South Asia, India is still only a “neighbourhood’s education hub” (Xavier et al., 2020).



Note: Legends include: name of the country, number of students and their percentage.

Fig. 10 Country-wise distribution of inbound students from other South Asian countries to India, 2019–20. Source: Based on Gol (2021).

The third mode enables the commercial presence of universities in foreign countries and is by far the most contested one. On the one hand, it might limit the number of students traveling abroad and the expenses incurred therein, but on the other hand it may wipe out domestic universities which will find it hard to compete with international brands. In the context of India, Sahni and Kale (2004) argue that “the opening up of the domestic market for such branches however, would intensify the level of competition in India as an internal level of ‘brain drain’ of academicians may also occur in this field.” Finally, the last mode focusses on the temporary cross-border movement of researchers and teachers as service providers. Most developing countries, including India, have been lobbying for the removal of barriers in this mode owing to the comparative advantage that they possess in terms of abundant and specialized service providers in the areas of information technology, medicine, engineering, entertainment and hospitality to name a few (Sahni and Kale, 2004). On the whole, putting together all the four modes, the performance of South Asia in internationalization of higher education has been poor.

An important outcome of globalization is global competition along with internationalization. Every country wants to compete with others in attracting students, foreign students in particular. Global university rankings, which is also an offshoot and also a strengthening feature of globalization, have further accentuated competition, and every country wants to have its universities appear in the top positions in the global rankings. They are also keen on developing world class universities. For example, India at some point had planned to establish high quality research and innovation universities which might come close, in nature, to the world class universities. To this end, the government has planned to identify universities of eminence and strengthen them further to become world class universities. India also developed its own national university ranking systems to improve competition.

As per the current situation, South Asian universities rank poorly in international rankings. Except for a handful of institutions in the region, no higher education institutions in South Asia figures in any of the global rankings of the universities. Only eight South Asian universities, of which six are in India, one in Pakistan and one in Sri Lanka, have featured in the top 500 world universities in 2020 as ranked by the Times Higher Education (THE) World University Rankings. In the 2021 rankings, only four from South Asia figured among the top 500 (Table 9). Even in the QS Asia University Rankings (2021), only seven Indian institutions and the National University of Science and Technology, Pakistan appear in the top 100. None from South Asia figures in the top 30. As per the latest World University Rankings of the QS (2022), the performance of South Asia is marginally better, with 11 institutions — 8 from India and 3 from Pakistan — appearing among the top 500, with better ranks. Of course, not even one figures in the top 100.

The poor position of universities in global rankings reflects the poor quality of higher education in South Asia. The quality of education is a highly contested area covering a wide array of parameters ranging from graduate out-turn, graduate attributes, employability of graduates, and so on. These are clearly related to the quality of teaching programmes the universities offer, the research practices in the institutions and the performance of the institutions in social relevance and service. Both in terms of inputs, such as the quality of teachers, infrastructure, and outputs in terms of quality of graduates and research publications, many universities in South Asia fare very poorly.

Though there are certain limitations with such rankings, they are useful at least for certain limited purposes as they highlight the gaps in some key indicators on research performance and also on performance in teaching, highlighting wide gaps that exist

Table 9 South Asian institutions of higher education in top 500 institutions in the Times Higher Education Ranking (2020 and 2021) and QS World University Rankings (2022)

<i>Rank</i>	<i>Name of the Institution</i>
Times Higher Education Rankings	
2021	
301-350	Indian Institute of Science (India)
351-400	Indian Institute of Technology Ropar (India)
401-500	Indian Institute of Technology Indore (India)
401-500	University of Peradeniya (Sri Lanka)
2020	
301-350	Indian Institute of Science (India)
301-350	Indian Institute of Technology Ropar (India)
351-400	Indian Institute of Technology Indore (India)
401-500	Indian Institute of Technology Bombay (India)
401-500	Indian Institute of Technology Delhi (India)
401-500	Indian Institute of Technology Kharagpur (India)
401-500	University of Peradeniya (Sri Lanka)
401-500	Quaid-i-Azam University (Pakistan)
QS World University Rankings	
2022	
177	Indian Institute of Technology Bombay (India)
185	Indian Institute of Technology Delhi (India)
186	Indian Institute of Science (India)
255	Indian Institute of Technology Madras (India)
277	Indian Institute of Technology Kanpur (India)
280	Indian Institute of Technology Kharagpur (India)
358	National University of Sciences and Technology (Pakistan)
378	Quaid-i-Azam University (Pakistan)
395	Indian Institute of Technology Guwahati (India)
398	Pakistan Institute of Engineering and Applied Sciences (Pakistan)
400	Indian Institute of Technology Roorkee (India)

Source: The World University Rankings: https://www.timeshighereducation.com/world-university-rankings/2020/world-ranking#!/page/0/length/25/sort_by/rank/sort_order/asc/cols/stats; <https://www.timeshighereducation.com/world-university-rankings/2021/world-ranking> [and] QS World University Rankings: <https://www.topuniversities.com/university-rankings/world-university-rankings/2022>

between different countries, apart from many other aspects, which may or may not necessarily be considered relevant by all countries (Tilak, 2016). Such information the ranking systems provide may be helpful to the governments and universities in taking appropriate policy initiatives and interventions.

On the whole within the South Asian region, higher education and research tend to privilege India, Pakistan, and Bangladesh, and among the three, India is more privileged. For example, in Table 9 we note that three out of four in top 500 universities in 2020, and six out of eight in 2021 in the Times Higher Education rankings, are Indian institutions. One university from Sri Lanka figures in the list both in 2020 and 2021; in 2020 one university from Pakistan was also in this category. India's advantage owes to its recognition of higher education and research in nation building and accordingly its laying strong foundations for higher education and research during the first three five year plans after independence. Moreover, India could absorb and benefit relatively better than other countries of the region, from the treasure of ancient Indian (*Vedic*) knowledge, as well as from the modern knowledge of the west. Sri Lanka has laid strong foundations on school education, but restricted the growth of higher education and research. Bangladesh is also laying emphasis on school education, but not much on higher education. So are the Maldives and Bhutan. India has also a scale advantage of having a large mass of higher educated graduates and a large number of higher education and research institutions.

Globalization and public funding of education

Among many aspects of globalization, the most noticeable one is the funding of education. Public funding reflects the priority that the state accords to education. The proportion of GDP allocated to education is an important standard indicator of the priority a nation gives to education. On an average, South Asia spent around 2.5% of its GDP on education as a whole as against a world average of 4.5% in 2019, in contrast to the recommendation of many including the UNESCO (2015) that governments should allocate at least 4–6% of GDP and/or at least 15–20% of total public expenditure to education. With quite a few ups and downs, including a peak of 3.4% in 2010, a dip of 2.4% in 2012, and a peak again of 3.4% in 2014, the GDP on education in this region has finally plateaued at 2.5% in 2016, that is marginally less than the level spent in 2000 (2.6%). The zig-zag trends of expenditure on education as a proportion of the GDP or of the total public expenditure shown in Fig. 11 reflect the wavering

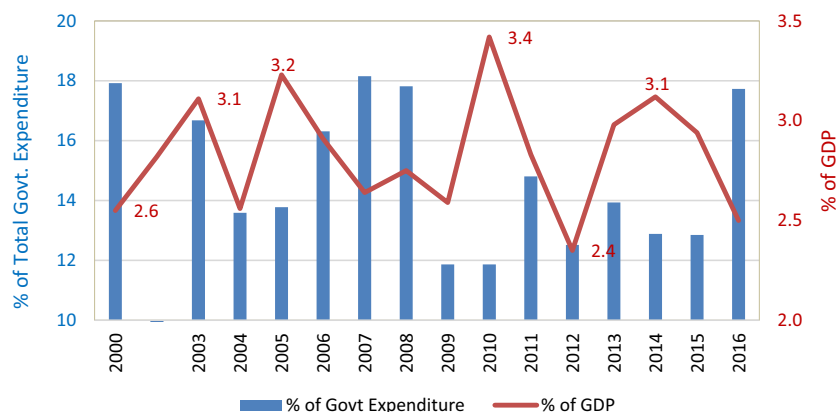


Fig. 11 Public expenditure on education as a percentage of total public expenditure and GDP in South Asia. Source: World Bank (2021).

priority governments have accorded to education. The regional average in 2016 was about 17%, which is about the same in 2000 and 2007. These fluctuating trends also reveal an underlying instability, disparity, lack of vision for the future and an absence of clarity in the state's commitment to education.

Among all the South Asian countries, Bhutan accords the highest priority to education by allocating 7.2% of its GDP in 2017. This is also one of the highest figures in the world. The corresponding figure in Bangladesh is the lowest at 1.5%. According to 2017 data, Pakistan and Sri Lanka apparently devoted similar percentages of their GDP to education at 2.9% and 2.8% respectively. However, one needs to note that Sri Lanka has a historical advantage in terms of public spending on education and a well-established robust system; and these figure reflect mostly current expenses. On the other hand, Pakistan is still groping to set up its educational system; and its current levels of spending are the second lowest in the region (after Bangladesh).

Between elementary, secondary and higher levels, the share of higher education in GDP is the lowest in South Asian economies (Table 10). Except in the case of Nepal, intra-sectoral allocations do not significantly favor elementary education, as elementary education, an apparent priority sub-sector in education, does not account for even half the total allocation to education, though this is a stated goal in countries like India.

During the past three or four decades, the education sector has been subjected to severe cuts in public budgets in many countries, including in South Asia, which is also partly the result of adopting the adjustment policies that require the downsizing of budgets in all social sectors. These policies also required the introduction of user charges and cost recovery methods which are heavily resorted to in education, particularly in secondary, and more clearly, in higher education. According to the national Constitutions and international commitments, primary education is to be provided free. Though some nominal fees are charged in some countries, and high charges in private primary schools, public education at the primary level is not much subjected to cost recovery measures.

Most public universities in South Asia that have functioned hitherto with government support have also had to resort to "alternative" means of financing themselves in the face of severe cuts in state funding. Apart from raising student fees significantly, they have started offering "self-financing" courses in many subjects and diplomas or "executive" programmes, sometimes of dubious quality, provided at exorbitant fees (Asian Development Bank, 2012). User charges have been introduced for a variety of items and services which earlier were provided either free or at nominal charges. Student loan schemes have been introduced or restructured in India, Pakistan, Bangladesh, Nepal, and Sri Lanka, though they appeal only to a small proportion of students. Because of the increase in fees and associated expenditure in public and private institutions, household spending on education has become very sizable. As UNESCO (2015) has concluded, the proportion of a private (household) spending on education as compared to total spending on education in South Asia is among the highest in the world and continues to rise further. These trends have led to "a weakening of the state as the primary provider of education and [have] contribute[d] to the emergence of a dynamic political economy landscape where non-state actors have gained prominence and inequality of opportunities have gone unaddressed for generations" (Chudgar and

Table 10 Expenditure on education (by levels) as a percent of GDP.

	Bangladesh		Bhutan		India		Maldives		Nepal		Pakistan		Sri Lanka	
	2000	2020	2009	2014	2000	2016	2010	2016	2010	2016	2013	2015	2010	2019
All	2.1	1.5	4.9	6.0	4.3	3.8	4.1	4.1	3.6	3.7	2.5	2.7	1.7	2.1
Elementary	1.0	NA	1.1	1.9	1.6	1.0	1.8	1.5	2.0	2.0	0.8	0.8	0.4	0.6
Secondary	0.8	0.7	2.7	3.4	1.7	1.2	1.0	0.9	1.1	1.2	0.6	1.0	0.9	1.0
Higher	0.2	0.4	NA	0.6	0.9	1.1	0.3	0.8	0.5	0.4	0.8	0.6	0.3	0.4

Source: World Bank (2021).

Aslam, 2020, p. 1053). The dominance of the private sector and high levels of private expenditure have created problems of inequitable access, poor quality and an uneven growth of higher education. High levels of private/household spending represent high levels of inequalities in education. Theoretically, private expenditure may complement public expenditure raising the overall level of expenditure on education, but as Tilak (2001) points out, the assumption of a complementary relationship between public and private expenditure on higher education can be erroneous, especially in the case of South Asia where they have been substituting each other. It is also important to note that there is a very strong correlation between educational attainment and government expenditure on education (Siddiqui, 2006; Tilak, 1999a), suggesting the need to protect government expenditure on education from frequent budget cuts.

The large-scale growth of the private sector and the associated growth of private expenditure on education are occurring because essentially education, particularly, but not just higher education, is increasingly being viewed as a private good, and the cost of losing the public good nature of education is not understood (Tilak, 2008b).

External aid for education

One particular aspect that perhaps needs to be mentioned in the present context of funding is foreign aid for education. Programmes like Education for All, the Millennium Development Goals and the Sustainable Development Goals that aim to achieve universal primary education serve as the rationale for foreign aid in the area of education in developing countries, including South Asia. Initially, the presence of the international community in the context of domestic education was more in the nature of aid in the form of grants and technical support. During the phase of adjustment, however, aid has taken on the form of loans. Realizing that “aid to education is unlike aid to other sectors because education is most susceptible to charges of interventionism, dependence and cultural domination, etc” (Tilak, 1988b), many countries have not been receptive to the idea of aid, particularly in the form of loans for education. But soon almost all countries in the region, India, Pakistan, Bangladesh, Bhutan, Sri Lanka and Nepal have resorted to the new mode of educational aid from multilateral organizations like the World Bank and the International Monetary Fund. India’s experience with foreign aid for education presents an interesting story. Initially India was averse to take any loan for primary education, while the World Bank was keen on entering into loan business in primary education in the late 1980s. Due to political economy factors including pressurized by the World Bank on one side and deteriorating economic situation in the country on the other, India gave in, initially agreeing for partial funding of primary education in 10 districts in the state of Uttar Pradesh in north India; but soon the project known as District Primary Education Project expanded covering nearly the entire country (Tilak, 1999b). Subsequently India has drastically reduced its dependence on aid for education. On the whole, most countries in the region have received external aid from multilateral organizations like the World Bank and Asian Development Bank and many bilateral sources. India has not only received aid from several sources, but also provided assistance to some countries in the region like Nepal, Bhutan, Maldives, and Afghanistan and also a few countries in Africa.

Overall, overseas development assistance (ODA) has been high in case of education in South Asian region. It increased by 2.5 times from US\$866 million in 1997–2000 to US\$2,185 million in 2019. In 2019 basic education accounted for 40% of the total, 22% by secondary education and post-secondary education 35% (Table 11). In case of India, the direction of assistance changed over the years significantly from primary education to post-secondary education, the latter accounting for 82% of total education aid it received in 2019, compared to less than 15% in 1997–2000. In Bhutan and the Maldives also we find post-secondary education receiving higher proportions of aid, but the absolute numbers are small. In Sri Lanka secondary education received 41.8% of the total education aid. In all other countries the priority of the international development organizations is in favor of basic education, as these countries are still farther from reaching the SDG goals relating to primary education. It is rather hard to make definite conclusions about the effectiveness of such aid in the education sector. A lot of research has analyzed the impact of foreign aid on education. Foreign aid has improved quantitative indicators of educational development that were in alignment with the goals of Education for All and the Millennium Development Goals for achieving universal primary education by 2015. While its impact on the physical expansion of education, in terms of enrollment or provision of educational infrastructure has been positive, the impact of aid on sustainable development in education in terms of providing quality education, in improving the quality of teachers, by providing high quality training and by bettering the learning outcomes of children has been rather disappointing. Aid has produced several desirable and not so desirable effects on educational policies in developing countries. The latter have been significant in some South Asian countries: some effects that aid has produced include a distortion in priorities and policies, a substitution of public funds and creating a culture of donor dependency (Tilak, 1997, 2008a). The perceived benefits of greater international interactions, an exchange of ideas and extension of financial aid have not all been achieved in South Asia. With a staggering 11.3 million children in the primary school-going age and 20.6 million children in the lower secondary school-going age in South Asia not going to school (UNICEF, 2018), there is enough reason to question the paths treaded by countries of the region to borrow for primary education. Therefore one can conclude that there is much scope for improvement in the aid mechanism (Riddell, 2012).

Concluding observations

Most of the South Asian countries gained independence from colonial rule in the middle of the 20th century. Nationalism in colonial period was essentially anti-colonial; but nationalism in independent nations is different. Nationalism and national values differ from country to country. The ancient Indian philosophy behind “*Vasudhaiva kutumbakam*” underlined the broad approach of India;

Table 11 Overseas development assistance to education, by recipient country in South Asia.

	1997–2000 (average annual average)				2019			
	Post-			Total	Post-			Total
	Basic	Secondary	secondary		Basic	Secondary	secondary	
US\$ millions								
Bangladesh	91	44	9	149	401	206	113	720
India	331	14	75	522	45	36	365	446
Pakistan	22	1	14	37	158	72	116	392
Afghanistan	2	1	5	8	180	82	91	353
Nepal	55	4	6	66	85	58	49	191
Sri Lanka	5	53	2	60	18	28	21	67
Bhutan	1	2	2	6	5	4	9	13
Maldives		12	5	18	1	0	2	3
% Distribution								
Bangladesh	61.1	29.5	6.0	100.0	55.7	28.6	15.7	100.0
India	63.4	2.7	14.4	100.0	10.1	8.1	81.8	100.0
Pakistan	59.5	2.7	37.8	100.0	40.3	18.4	29.6	100.0
Afghanistan	25.0	12.5	62.5	100.0	51.0	23.2	25.8	100.0
Nepal	83.3	6.1	9.1	100.0	44.5	30.4	25.7	100.0
Sri Lanka	8.3	88.3	3.3	100.0	26.9	41.8	31.3	100.0
Bhutan	16.7	33.3	33.3	100.0	38.5	30.8	69.2	100.0
Maldives	0.0	66.7	27.8	100.0	33.3	0.0	66.7	100.0

Source: UNESCO (2010, 2021).

and it is secularism that determined the Indian nationalism after independence. But nationalism meant Muslim nationalism in Islamic countries like Pakistan, Bangladesh, and the Maldives (Emerson, 1971); Sri Lanka and Bhutan are influenced by Buddhism, and Nepal by Hinduism. Obviously schools are used to propagate the nationalistic, social, cultural, and political values based on which the political systems evolved in South Asian countries, like everywhere else. They are also used to propagate, working in their own specific cultural contexts, values of religion that the country is wedded to. For example, Vedic schools in India focused on imparting Vedic values, Sikh schools in India and Pakistan on Sikh values, Christian schools on Christian values, while madrasas in Islamic countries – Pakistan, Bangladesh, the Maldives – and also in India focused on Islamic values, though all are attempting to modernize themselves. The curriculum in many of the schools also includes basic sciences, mathematics and languages. The mainstream schools in India emphasize secular and universal human values. While broad values remained unaltered, over the period they have undergone marginal to significant changes in intensity and direction; and these changes necessitated somewhat frequent curricular reforms, as political ideologies and overall development paradigms also changed drastically. For example, given the alleged links with militancy, madrasas in India and also in other countries are subject to curricular reforms and modernization (Bano, 2010). Many other reforms are also introduced in the education systems in South Asia in the wake of globalization and growth of neo-liberalism.

The introduction of new policies and policy reforms in education are certainly influenced by political economy factors in almost every society. In fact, no policy reform takes place in a vacuum. Severely affected by several invasions and a long period of colonial rule, many countries in South Asia have been struggling to reform their education systems. Though they have not been able to develop their own indigenous systems replacing the colonial system of education, a phenomenal expansion has taken place in education mostly on the lines of the colonial model. Alongside, it has been crippled with a plethora of problems, partly accentuated by globalization.

Education development may be evaluated in terms of the quantity, quality and equity dimensions that form a triangle, an “elusive triangle” (Naik, 1979). In terms of all the three dimensions, South Asian countries have fared poorly in comparison with the more advanced regions of the world and they have a long way to catch up. Even within South Asian countries, there exists wide variations in education development.

To sum up the overall education development and prospects for future in terms of human capital index (World Bank 2020), South Asia fares very poorly compared to many other regions of the world. Among the 174 countries for which such an index has been developed, all the countries of the region, except Sri Lanka, figure in the second 100 only, as shown in Table 12. The value of the index ranged between 0.29 (Central African Republic) and 0.88 (Singapore), and it ranged between 0.4 and 0.6 among the countries of the South Asian region. The index captures “the amount of human capital a child born today could expect to attain by age 18, given the risks ... that prevail in the country where [the child] lives.” If so, as the index shows, the educational future of the South Asian countries should indeed be a matter of serious concern of all.

Table 12 Human Capital Index in South Asia, 2020

	<i>Index</i>	<i>Rank</i>
Afghanistan	0.40	148
Bangladesh	0.46	123
Bhutan	0.48	121
India	0.49	116
Nepal	0.50	109
Pakistan	0.41	144
Sri Lanka	0.60	71
Maldives	NA	NA

Source: World Bank (2020b).

Education in many South Asian countries suffers from a policy crisis which can also be described as policy paralysis (Tilak, 2015). A coherent education policy, formulated based on a long-term vision for the development of society, is an important pre-requisite for building India has announced a visionary policy in 2020. a strong education system that will help in the social transformation, economic progress and nation-building. Of all the necessary requirements, a nation's political will is an important determinant of the nature, kind and pace of growth in a country's education framework and the public investments made in it. Unfortunately, most South Asian countries, with the exception of Sri Lanka, have shown a lack of strong public commitment toward education (Haq and Haq, 1998; Jahan, 2021). The complex challenges facing South Asia require exploring creative policies in education, by learning from each other's present and past, including ancient experiences and from other countries of Asia and the rest of the world.

References

- ADB: Asian Development Bank, 2008. Education Sector in Bangladesh: What Worked Well and Why under the Sector-Wide Approach? <https://www.oecd.org/countries/bangladesh/42228468.pdf>.
- ADB: Asian Development Bank, 2012. Counting the Cost: Financing Asian Higher Education for Inclusive Growth: Higher Education Dynamics in Asia (Manila).
- ADB, 2017. Human Capital Development in South Asia: Achievements, Prospects, and Policy Challenges. Manila.
- Ahmad, S., 2018. Unleashing the Potential of a Young Pakistan, Human Development Reports (24 July 2018). <http://hdr.undp.org/en/content/unleashing-potential-young-pakistan>.
- AIU: Association of Indian Universities, 2017. Internationalisation of Higher Education in India. New Delhi.
- ASER India, 2016. Annual Status of Education Report (Rural) 2016. ASER Centre, Pratham Foundation, New Delhi.
- ASER Pakistan, 2015. Annual Status of Education Report ASER-Pakistan 2015. South Asia Forum for Education Development, Lahore.
- Bano, M., 2010. Madrasas as partners in education provision: the South Asian experience. *Dev. Pract.* 20 (4/5), 554–566.
- Béteille, T., Tognatta, N., Riboud, M., Nomura, S., Ghorpade, Y., 2020. Ready to Learn Before School, in School, and Beyond School in South Asia. World Bank, Washington DC.
- Bhattacharya, S., Mukherjee, A., 2013. Strategic information revelation when experts compete to influence. *Rand J. Econ.* 44 (3), 522–544.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood, Westport, CT, pp. 241–258.
- Chudgar, A., Aslam, M., 2020. Political Economy of Education in South Asia: An Introduction. In: Sarangapani, P., Pappu, R. (Eds.), *Handbook of Education Systems in South Asia*. Global Education Systems. Springer, Singapore, pp. 1049–1058.
- Chowdhury, R., Kabir, A.H., 2014. Language wars: English education policy and practice in Bangladesh. *Multil. Educ.* 4 (1). <https://doi.org/10.1186/s13616-014-0021-2>.
- Dharampal, 1983. The Beautiful Tree: Indigenous Indian Education in the Eighteenth Century. Biblia Impex, New Delhi.
- Emerson, R., 1971. Post-independence nationalism in South and Southeast Asia: a reconsideration. *Pac. Aff.* 44 (2), 173–192 (Summer).
- Erling, E.J., Seargeant, P. (Eds.), 2013. *English and Development: Policy, Pedagogy and Globalization*. Multilingual Matters, Bristol.
- Gol: Government of India, 2020. National Education Policy 2020. Government of India, New Delhi.
- Gol: Government of India, 2021. All-India Survey of Higher Education 2019–20. Government of India, Ministry of Education, New Delhi. https://www.education.gov.in/sites/upload_files/mhrd/files/statistics-new/aishe_eng.pdf.
- Government of Maldives, 2019. Education Sector Analysis Maldives. Male. <https://www.globalpartnership.org/sites/default/files/2019-05-maldives-education-sector-analysis.pdf>.
- Haq, M.U., Haq, K., 1998. *Human Development in South Asia*, 1998. University Press, Dhaka.
- Hossain, N., Subrahmanyam, R., Kabeer, N., 2002. The Politics of Educational Expansion in Bangladesh. Working Paper 167. Institute of Development Studies, Sussex.
- Hunter, R., 2020. Education in Pakistan. *World Education News + Reviews*. Available at: <https://wenr.wes.org/2020/02/education-in-pakistan>. (Accessed 18 April 2021).
- Ilie, S., Rose, P., 2016. Is equal access to higher education in South Asia and sub-Saharan Africa achievable by 2030? *High Educ.* 72 (4), 435–455.
- Jahan, R., 2021. Universalisation of secondary education: questions for discussion and debate. In: Tilak, J.B.G. (Ed.), *Universal Secondary Education in India*. Springer, Singapore, pp. 419–427.
- Kingdon, G.G., 2017. The private schooling phenomenon in India: a review. *J. Dev. Stud.* 56 (10), 1795–1817.
- Kurrien, J., 2021. Delusions and traps for the unwary: english medium schooling for the poor in India. *Academia Letters*. <https://doi.org/10.20935/AL2893>. Article 2893.
- Lahiry, S., 2013. India's relationship with its South Asian neighbours: contemporary trends. *World Aff.* 17 (4), 70–79.
- Lakshman, I., Rajeshkannan, R., Schubert, M., 2020. Continuing education during times of war: experiences of children in northern Sri Lanka. *Eur. Scient. J.* 16 (35), 1–18. <https://doi.org/10.19044/esj.2020.v16n35p1>.
- Mehendale, A., 2020. School education systems and policies in South Asia: an introduction. In: Sarangapani, P., Pappu, R. (Eds.), *Handbook of Education Systems in South Asia*. Springer, Singapore, pp. 393–410.
- Montenegro, C.E., Patrinos, H.A., 2014. Comparable Estimates of Returns to Schooling Around the World. Policy Research Working Paper no. 7020. World Bank, Washington DC.
- Naik, J.P., 1979. Equality, quality and quantity: the elusive triangle in Indian education. *Int. Rv. Educ.* 25 (2–3), 167–185.
- Perera, S., 2001. The Ethnic Conflict in Sri Lanka: A Historical and Sociopolitical Outline. Background Paper prepared for the World Bank. <https://documents1.worldbank.org/curated/en/727811468302711738/pdf/677060WP00PUBLOio0political00Outline.pdf>.
- Richards, J., Ahmed, M., Islam S. 2022. *The Political Economy of Education in South Asia: Fighting Poverty, Inequality and Exclusion*. Toronto, University of Toronto Press.

- Riddell, P., 2012. Crisis management. *Polit. Insight* 3 (1), 20–22.
- Sah, P.K., 2017. Using the first language (L1) as a resource in EFL classrooms: Nepalese university teachers' and students' perspectives. *J. NELTA* 22 (1–2), 26–38.
- Sah, P.K., 2022. English medium instruction in South Asian's multilingual schools: unpacking the dynamics of ideological orientations, policy/practices, and democratic questions. *Int. J. Biling. Educ. Biling.* 25 (2), 742–55.
- Sahni, R., Kale, S., 2004. GATS and higher education: some reflections. *Econ. Polit. Wkly.* 39 (2), 2174–2180.
- Schultz, T.W., 1961. Investment in human capital. *Am. Econ. Rev.* 51 (1), 1–17.
- Sen, A., 1999. *Development as Freedom*. Clarendon, Oxford.
- Shah, D., 2020. By the numbers: MOOCs in 2020. Report by Class Central. <https://www.classcentral.com/report/mooc-stats-2020/>.
- Sharma, G., 2020. Conceptions of school curriculum in South Asia: a study of India, Pakistan, and Bangladesh. In: Sarangapani, P., Pappu R. (Eds.), *Handbook of Education Systems in South Asia*. Springer, Singapore, pp. 1704–1728.
- Siddiqui, A., 2006. Macroeconomic returns to education in South Asia. *J. Econ. Coop. among Islam. Ctries.* 27 (4), 25–44 (January).
- Tilak, J.B.G., 1988a. Vocational education in South Asia: problems and prospects. *Int. Rev. Educ.* 34 (2), 244–257.
- Tilak, J.B.G., 1988b. Foreign aid for education. *Int. Rev. Educ.* 34 (3), 313–335.
- Tilak, J.B.G., 1997. Effects of adjustment on education: a review of Asian experience. *Prospects* 27 (1), 85–107.
- Tilak, J.B.G., 1999a. Investment in human capital in India: an inter-state analysis of stock and flow of human capital. *J. Ind. Sch. Polit. Econ.* 11 (1), 39–75.
- Tilak, J.B.G., 1999b. Development assistance to primary education in India: transformation of enthusiastic donors and reluctant recipients. In: King, K., Buchert, L. (Eds.), *Changing International Aid to Education: Global Patterns and National Contexts*. UNESCO in cooperation with NORRAG, Paris, pp. 307–317.
- Tilak, J.B.G., 2000. Education for All in South and West Asia: A Decade After Jomtien: An Assessment. A Synthesis Report. UNESCO, New Delhi.
- Tilak, J.B.G., 2001. Education and development: lessons from the Asian experience. *Indian Soc. Sci. Rev.* 3 (2), 219–266.
- Tilak, J.B.G., 2003a. Education in South Asia. In: Guthrie, J.W. (Ed.), *The Encyclopedia of Education*, second ed. Macmillan Reference, New York, pp. 2272–2277.
- Tilak, J.B.G., 2003b. Vocational education and training in Asia. In: Kleeves, J.P., Watanabe, R. (Eds.), *The Handbook on Educational Research in the Asia Pacific Region*. Kluwer Academic Publishers, Dordrecht, pp. 673–686.
- Tilak, J.B.G., 2005. Are we marching towards laissez-faireism in higher education development? *J. Int. Cooper. Educ.* 8 (1), 153–165.
- Tilak, J.B.G., 2006a. Private higher education: philanthropy to profits. In: *Higher Education in the World: The Financing of Universities*. Global University Network for Innovation, Barcelona and Palgrave Macmillan, pp. 113–125.
- Tilak, J.B.G., 2006b. Global trends in funding higher education. *Int. High. Educ.* 42, 5–6.
- Tilak, J.B.G., 2008a. Political economy of external aid for education in India. *J. Asian Pub. Pol.* 1 (1), 1–20.
- Tilak, J.B.G., 2008b. Higher education: a public good or a commodity for trade? Commitment to higher education or commitment of higher education to trade. *Prospects* 38 (4), 449–466.
- Tilak, J.B.G., 2011. Trade in Higher Education. The Role of the General Agreement of Trade in Services (GATS). *Fundamentals of Educational Planning* No. 95. UNESCO International Institute for Educational Planning, Paris.
- Tilak, J.B.G., 2015. Higher education in South Asia: crisis and challenge. *Soc. Scient.* 43 (1–2), 43–59 (January–Feb).
- Tilak, J.B.G., 2016. Global rankings, world-class universities and dilemma in higher education policy in India. *Higher Educ. Future* 3 (2), 1–18 (July).
- Tilak, J.B.G., 2020. On dilemmas in reforming higher education in India. *Higher Educ. Future* 7 (1), 54–66.
- Tilak, J.B.G., 2022. Book review of 'An alternative perspective on education (Atul Kohari)'. *Int. Rev. Educ.* 68 (1), 155–159 (Feb).
- Trines, S., 2018a. Education in India. *World Education News + Reviews*. Available at: <https://wenr.wes.org/2018/09/education-in-india>. (Accessed 27 March 2021).
- Trines, S., 2018b. Educating the Masses: The Rise of Online Education in Sub-Saharan Africa and South Asia. *World Education News + Reviews*. Available at: <https://wenr.wes.org/2018/08/educating-the-masses-the-rise-of-online-education>.
- UIS (UNESCO Institute of Statistics), 2021. Data for Sustainable Development Goals: Education and Literacy. uis.unesco.org, 2021. (online database). Available at: <http://uis.unesco.org/>.
- UNDP, Human Development Report, (various years), New York.
- UNESCO, 2009. Asia and the Pacific Education for All (EFA): Mid-Decade Assessment: South Asia Sub-region Synthesis Report. UNESCO Bangkok, Bangkok.
- UNESCO, 2015. Education 2030: Incheon Declaration and Framework for Action—Towards Inclusive and Equitable Quality Education and Lifelong Learning for All. UNESCO, Paris.
- UNESCO, 2010. EFA Global Monitoring Report 2010: Reaching the Marginalized. UNESCO, Paris.
- UNESCO, 2021. Global Education Monitoring Report 2021/2: Non-state actors in education – Who chooses? Who loses? UNESCO, Paris.
- UNICEF, 2018. Improving Education Quality in South Asia: A Review of UNICEF's Efforts. UNICEF Regional Office for South Asia, Kathmandu.
- Washbrook, D.A., 1988. Progress and problems: South Asian economic and social history c.1720–1860. *Mod. Asian Stud.* 22 (1), 57–96.
- Wijayatunga, A., 2018. English as a medium of instruction in secondary schools in Sri Lanka. In: *Proceedings of the Fourth International Conference of Education*, vol. 4, pp. 151–161. <https://www.researchgate.net/deref/https%3A%2F%2Fdoi.org%2F10.17501%2Ficedu.2018.4117>.
- World Bank, 2011. *Human Capital for a Knowledge Society, Higher Education in the Maldives: An Evolving Seascape*. Washington DC.
- World Bank, 2018. *World Development Report 2018: LEARNING to Realize Education's Promise*. Washington DC.
- World Bank, 2020a. *Ending Learning Poverty: What will it Take?* Washington DC.
- World Bank, 2020b. Human Capital Project. Washington DC. <https://www.worldbank.org/en/publication/human-capital>.
- World Bank, 2021. Databank. [Databank.worldbank.org](https://databank.worldbank.org), 2021. Education | Data | DataBank. (online database). Available at: <https://databank.worldbank.org/databases/education>.
- Xavier, C., Chaba, A., Dang, G., 2020. Is India Still the Neighbourhood's Education Hub? *Policy Brief* Brookings Institution India Center, New Delhi. <https://www.brookings.edu/wp-content/uploads/2020/03/s-India-still-the-neighbourhood%E2%80%99s-education-hub-1.pdf>.

Regionalization of higher education and geopolitics in Northeast Asia¹

Kiyong Byun^a and Bawool Hong^b, ^a Department of Education, Korea University, Seoul, South Korea; and ^b Higher Education Policy Research Institute, Korea University, Seoul, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Introduction	447
Intraregional dependence among Northeast Asian countries	448
Economic interdependence in Northeast Asia	448
Growth & diversity of cross-border higher education in Northeast Asia	448
Regionalization of higher education in Northeast Asia	450
Emergence of regional higher education cooperation framework in Northeast Asia: a historical overview	450
CAMPUS ASIA program (China, Japan & South Korea)	452
Challenges for regionalization of higher education in Northeast Asia	453
Geopolitical tensions in Northeast Asia	453
Regional leadership	453
US influence on Northeast Asian higher education	454
Conclusion	454
References	454
Relevant websites	455

Introduction

The geographical notion of East Asia or Northeast Asia has been defined in many different ways. Following the usage of East Asia and Northeast Asia in previous regional integration literature (i.e., [Cheong, 2005](#); [Li, 2007](#); [Kuroda and Passarelli, 2009](#)), in this article, Northeast Asia is referred to as the three major Northeast economies of China, Japan, and South Korea, with East Asia covering Northeast Asia, as well as Southeast Asia (comprising the ten ASEAN countries), unless otherwise specified. The East Asian region, however, is extremely diverse and complex (e.g., including Hong Kong and Taiwan), and does not easily constitute a single political, economic, or cultural entity. Within East Asia, the Southeast Asian region was, in many aspects, more advanced in multi-lateral higher education cooperation efforts than their Northeastern counterparts, and made much progress over the years through many programs (i.e., the ASEAN International Mobility for Students: AIMS program) administered by the Southeast Asian Ministers of Education Organization Regional Institute for Higher Education and Development (SEAMO RIHED hereafter) and ASEAN University Network (AUN hereafter). However, certain changes are being detected in this situation. The most notable change is the “CAMPUS ASIA” program, a multi-lateral student mobility program among the three countries, launched with a purpose to develop further into a more comprehensive higher education cooperation framework within Northeast Asia and beyond.

Nevertheless, due to the bitter history shrouding Northeast Asia, interregional higher education cooperation and one of its major projects in the form of the CAMPUS ASIA program is inevitably affected by the ever-changing geopolitics in the region. One of the major conflicts in this region occurred in 2017 when the U.S. deployed Terminal High Altitude Area Defense (THAAD) missile system on the Korean Peninsula, which caused heavy backlash from the Chinese government against South Korea, in the form of business sanctions, boycotts and tourist restrictions. Subsequently, the Korean government announced its New Southern Policy (NSP) that signaled its strategic shift toward Southeast Asia, in response to escalating tensions with China. To make matters worse, a 2019 Korean Supreme Court ruling on “comfort women” rekindled old colonial tensions between Japan and South Korea and ensued a trade dispute. Furthermore, Chinese suppression of democratic protests in Hong Kong and attempts to extend geopolitical influence through international education (e.g., Confucius Institutes), continuing missile tests by North Korea and the ongoing US-China trade war has pushed Japan to lean on its alliance with the US more than ever before, furthering its divide with China, while South Korea has been left in a precarious position of balancing its geopolitical position between the US-Japan alliance and China. These geopolitical conditions have cast an unpredictable outlook for the future of regionalization of higher education in Northeast Asia and threaten to undermine the achievements of the CAMPUS ASIA program so far.

Against this backdrop, the purpose of this article is to (i) overview developments in the regionalization of higher education in Northeast Asia, particularly focusing on China, Japan, and South Korea, and then to (ii) investigate the geopolitics surrounding higher education in the region, and the impact it will have on the outlook for the region.

¹This article is based on [Byun and Um \(2012\)](#). However, many parts of this article were updated and rewritten mainly by Bawool Hong in order to reflect the developments since 2012, when the manuscript was originally published.

Intraregional dependence among Northeast Asian countries

Economic interdependence in Northeast Asia

Kuroda et al. (2010) observed that “behind the concept of the East Asian Community lies a situation where the weight of this region in the world economy is expanding ... a comparatively more independent economic system that does not rely on the Western economy is forming (5).” To provide some empirical data in line with this observation, we examined the relative importance of the three major Northeast Asian economies in relation to the world, and also investigated how they depended on each other. **Table 1** presents the economic profiles of China, Japan, and South Korea in terms of their share of the world's total in population, GDP(PPP), trade, and FDI (inflows) in 2019.

It is evident from the data below that China, Japan, and South Korea are key players in the global economy. These three countries account for over twenty percent of the world's population (20.49%) and GDP (22.95%), respectively. The Northeast Asian economy has grown rapidly so that it is now larger in economic size than that of North America and even greater than the EU in terms of GDP (PPP). In addition, the Northeast Asian region represented by China, Japan, and South Korea occupies 16.83% of the world's traded goods and services, the second largest among the regional blocs in the world economy, although it still lags behind the EU and North America in terms of Foreign Direct Investment (FDI) inflows.

Statistics have indicated that economic inter-dependency among the Northeast Asian countries has increased rapidly since the start of the 21st century, albeit having slightly slowed down in the past decade. **Table 2** below shows the changes in trade volumes in terms of both exports and imports among China, Japan, South Korea, and the US from 2001 to 2019. Upon close examination of **Table 2**, in general, the economic inter-dependence among these three countries has been increasing, while trade dependence between the three Northeast Asian countries with the US in terms of both exports and imports, has gradually decreased over time.

With China's ever-growing presence in the world economy, the share of trade in both Japan and South Korea with China has drastically increased in terms of exports and imports. For instance, South Korea's imports from China rose from US\$ 13.3 billion or 9.4% (of South Korea's total imports) in 2001 to US\$ 107.2 billion or 21.3% in 2019. The share of South Korea's exports to China also increased to US\$ 136.2 billion or 25.1% in 2019, up from US\$ 18.1 billion or 12.1% in 2001. Similar patterns can also be observed between Japan and China. For China, the share of trade with Japan and South Korea also substantially increased in terms of total trade, while, percentage-wise, slightly decreased for Korea (exports: 4.7 => 4.4%; imports: 9.6 => 8.3%) and decreased significantly for Japan (exports 16.9 => 5.7%; imports 17.6 => 8.2%).

Furthermore, the trade volume for all 3 Northeast Asia countries in term of exports to and imports from the US, at least percentage-wise, has sharply decreased during the same period (China 20.4 => 16.7%, Japan 30.4 => 19.8%, Korea 20.8 => 13.5% on the export side, and China 10.8 => 5.9%, Japan 18.3 => 11.2%, Korea 15.9 => 12.3% on the import side). These data clearly indicate that, despite the fact that Northeast Asian countries still heavily rely on exports to countries outside the region (i.e., the US), intra-regional trade has played an increasingly important role for these countries' economies. Thus, China is currently one of the most important trade partners of Japan and South Korea in terms of both exports and imports.

Growth & diversity of cross-border higher education in Northeast Asia

As was illustrated in the previous section, the interdependence of the Northeast regional economy has grown substantially over the past two decades. Moreover, when closely examining recent developments of cross-border higher education activities in the region, similar trends can also be identified in the area of higher education as well. This section describes some of the recent developments in cross-border higher education among the three Northeast Asian countries, especially in terms of intraregional student mobility.

Table 3 below shows the international student flow in Northeast Asia over a span of nearly 2 decades (2000–2019). The data indicate that during this period all 3 Northeast Asian countries (China, Japan, South Korea) experienced exponential growth in recruiting international students within their territories, which demonstrates growing student mobility in this region. For instance, the number of international students in South Korea increased by almost 48-fold from 3373 in 2000 to 160,165 in 2019. The

Table 1 Economic profile of the major Northeast Asian countries: share of the world's total.

	(Unit: %)			
	Population (2019)	GDP (PPP) (2019)	Traded goods & services (2019)	FDI (inflows) (2019)
China	18.18	17.29	10.57	9.17
Japan	1.64	4.02	3.64	0.95
S. Korea	0.67	1.64	2.62	0.69
C-J-K	20.49	22.95	16.83	10.81
EU	6.67	17.73	34.82	29.02
N. America	4.77	17.22	13.74	19.26
World	100	100	100	100

Source: PRB World Population Data Sheet (www.prb.org); World Bank World Development Indicators (www.worldbank.org); United Nations Conference on Trade and Development (www.unctad.org); WTO World Trade Statistical Review (www.wto.org).

Table 2 Intra-regional trade among China, Japan, and South Korea: 2001–2019.

(Unit: billions of USD, % in parentheses)											
Country	Year	Exports to					Imports from				
		CH	JP	KR	US	All	CH	JP	KR	US	All
China (CH)	2001	–	44.9 (16.9)	12.5 (4.7)	54.4 (20.4)	266.1 (100)	–	42.8 (17.6)	23.4 (9.6)	26.2 (10.8)	243.6 (100)
	2010	–	121.0 (7.7)	68.8 (4.4)	283.8 (18.0)	1578 (100)	–	176.7 (12.7)	138.3 (9.9)	102.7 (7.4)	1396 (100)
	2019	–	143.2 (5.7)	110.9 (4.4)	418.5 (16.7)	2498.5 (100)	–	171.5 (8.2)	173.5 (8.3)	123.2 (5.9)	2068.9 (100)
Japan (JP)	2001	31.0 (7.7)	–	25.3 (6.3)	122.5 (30.4)	403.4 (100)	57.8 (16.6)	–	17.2 (4.9)	63.8 (18.3)	349.3 (100)
	2010	149.5 (19.4)	–	62.4 (8.1)	120.5 (15.6)	769.8 (100)	153.2 (22.1)	–	28.6 (4.1)	69.1 (10.0)	692.6 (100)
	2019	134.6 (19.0)	–	29.6 (4.1)	140.4 (19.8)	705.6 (100)	169.2 (23.4)	–	29.6 (4.1)	81.2 (11.2)	720.8 (100)
S. Korea (KR)	2001	18.1 (12.1)	16.5 (11.0)	–	31.4 (20.8)	150.4 (100)	13.3 (9.4)	26.6 (18.9)	–	22.4 (15.9)	141.1 (100)
	2010	117.2 (25.1)	28.3 (6.0)	–	50.1 (10.7)	467.7 (100)	71.5 (16.8)	64.3 (15.1)	–	40.6 (9.6)	425.1 (100)
	2019	136.2 (25.1)	28.4 (5.2)	–	73.5 (13.5)	542.1 (100)	107.2 (21.3)	47.5 (9.4)	–	62.1 (12.3)	503.2 (100)

Note: Figures in parentheses are exports to and imports from foreign countries as the percentage shares of the total exports and imports of each country concerned. All figures for China exclude the Special Administrative Regions (SAR) of Hong Kong and Macau.

Source: United Nations Commodity Trade Statistics Database (comtrade.un.org).

Table 3 International student flow in Northeast Asia: 2000–2019².

(Unit: number of students)								
Country	Inbound				Outbound			
	2000(A)	2009(B)	2019(C)	C/A	2000(A)	2009(B)	2019(C)	C/A
China	52,150	238,184	492,185	9.43	140,501	511,763	662,100	4.71
Japan	59,691	131,599	312,214	5.23	59,294	45,130	115,146	1.94
S. Korea	3,373	50,030	160,165	47.48	70,991	125,725	213,000	3.00
C-J-K	115,214	419,813	965,564	8.38	270,786	682,618	990,246	3.6
USA	514,723	671,616	1,095,299	2.12	143,590	260,327	347,099	2.41

Note: Statistics include all forms of international students (e.g., study abroad, language programs etc.).

Source: The Educational Statistics Yearbook of China, Ministry of Education of the People's Republic of China (en.moe.gov.cn), Japan Student Services Organization (jasso.go.jp), Korean Educational Development Institute (kedi.re.kr), Korean Ministry of Education (moe.go.kr), UNESCO Institute of Statistics (uis.unesco.org), Open Doors (opendoorsdata.org).

growth in international student enrollments at China's HEIs was even more astounding in terms of sheer quantitative mass, whereby over 440,000 additional international students were admitted at Chinese HEIs within the past 20 years. As for Japan, although not as exponentially as its other Northeast Asian counterparts; the number of international students still increased at a relatively rapid rate, especially when compared to the US.

In addition, similar trends were observed for outbound student mobility as well. China sent 662,100 students abroad in 2019, representing almost a five-fold increase from 2000, while Korea sent 142,009 more students abroad than it did in 2000. Japan on the other hand experienced a slight decrease in outbound students between 2000 and 2009, before experiencing a sharp increase in the past decade. With that in mind, it can be empirically assumed that international student mobility over the past 2 decades has grown faster, in terms of growth rates and absolute numbers, in the Northeast Asian countries, and even more so in relative terms to the traditional center of international student mobility: the US.

Furthermore, such rapid increases in inbound international students to these 3 Northeast Asia countries, have also changed the makeup of international students in each respective country, as shown in [Table 4](#) below. In 2000, the Northeast Asian countries were the top 2 feeder countries for each other. In the case of China, South Korea accounted for 32.2% of all international students while Japan came in second at 26.5%. And for Japan and South Korea, Chinese students constituted the majority of international students at 47% and 35% respectively, while the two countries (Japan & South Korea) were each other's second highest feeder country of international students. By 2019, the Northeast Asian countries were still each other's top feeding country and increased their sending numbers quantitatively, but also expanded their makeup to include a Southeast Asian country as their second largest feeder country (Thailand rose to be the second largest feeder country for China, while Vietnam emerged as the second largest feeder country for both Japan and South Korea).

This showed that not only did cross border higher education in Northeast Asia increase quantitatively, but also became increasingly more diversified. China being the prime example, branched out to even include more students from South Asia (Pakistan and

²It is noted that number of student flow by itself is not the best indicator for quality of mobility. Information on how long (on average) students spent abroad was not included due to limited statistics and data on student mobility in Northeast Asia.

Table 4 Makeup of international students in Northeast Asia.

<i>(Unit: number of students, % in parentheses)</i>			
<i>Country</i>	<i>Makeup (%)</i>		
	<i>2000</i>	<i>2009</i>	<i>2019</i>
China	1. S. Korea—16,787 (32.2%) 2. Japan—13,806 (26.5%) 3. USA—4280 (8.2%) 4. Indonesia—1947 (3.7%) 5. Germany—1270 (2.4%)	1. S. Korea—64,232 (27.0%) 2. USA—18,650 (7.8%) 3. Japan—15,409 (6.5%) 4. Vietnam—12,247 (5.2%) 5. Thailand—11,379 (4.8%)	1. S. Korea—50,600 (10.3%) 2. Thailand—28,608 (5.8%) 3. Pakistan—28,023 (5.7%) 4. India—23,198 (4.7%) 5. USA—20,996 (4.3%)
Japan	1. China—28,076 (47.0%) 2. S. Korea—18,237 (30.6%) 3. Malaysia—1956 (3.3%) 4. Indonesia—1413 (1.9%) 5. USA—1077 (1.8%)	1. China—79,394 (60.3%) 2. S. Korea—24,850 (18.9%) 3. Vietnam—2895 (2.2%) 4. Thailand—2193 (1.7%) 5. Malaysia—2147 (1.6%)	1. China—124,436 (39.9%) 2. Vietnam—73,389 (23.5%) 3. Nepal—26,308 (8.4%) 4. S. Korea—18,338 (5.9%) 5. Taiwan—9584 (3.1%)
S. Korea	1. China—1182 (35.0%) 2. Japan—613 (18.2%) 3. USA—195 (5.8%) 4. Russia—77 (2.3%) 5. Vietnam—62 (1.8%)	1. China—39,309 (78.6%) 2. Mongolia—1621 (3.2%) 3. Vietnam—1456 (2.9%) 4. Japan—989 (2.0%) 5. USA—758 (1.5%)	1. China—71,067 (44.4%) 2. Vietnam—37,426 (23.4%) 3. Uzbekistan—7492 (4.7%) 4. Mongolia—7381 (4.6%) 5. Japan—4392 (2.7%)
USA	1. China—54,466 (10.6%) 2. Japan—46,872 (9.1%) 3. India—42,337 (8.2%) 4. S. Korea—41,191 (8.0%) 5. Taiwan—29,234 (5.7%)	1. India—103,260 (15.4%) 2. China—98,235 (14.6%) 3. S. Korea—75,065 (11.2%) 4. Canada—29,697 (4.4%) 5. Japan—29,264 (4.3%)	1. China—369,548 (33.7%) 2. India—202,014 (18.4%) 3. S. Korea—52,250 (4.8%) 4. S. Arabia—37,080 (3.4%) 5. Canada—26,122 (2.4%)

Note: Statistics include all forms of international students (e.g., study abroad, language programs etc.).

Source: The Educational Statistics Yearbook of China, Ministry of Education of the People's Republic of China (en.moe.gov.cn), Japan Student Services Organization (jasso.go.jp), Korean Educational Development Institute (kedi.re.kr), UNESCO Institute of Statistics (uis.unesco.org), Open Doors (opendoorsdata.org).

India as third and fourth largest feeder countries respectively), while Japan too expanded its reach to South Asia (Nepal as third largest feeder country). South Korea would take a slightly different route by attracting more students from Central Asia (Uzbekistan and Mongolia as third and fourth largest feeder countries respectively). This increasing diversity of international student makeup in Northeast Asia, lies in sharp contrast to the US, whereby the international student body over the past 2 decades has increasingly become more concentrated on China and India.

To summarize, the data presented in this section suggest that, since the early 2000s, cross-border higher education activities in Northeast Asia have drastically increased and has led to form a certain degree of de facto integration of higher education in this region. This growing inter-dependence of higher education systems in Northeast Asia has also resulted in the development of a more concrete regional higher education cooperation framework in Northeast Asia. The next section further explores this issue.

Regionalization of higher education in Northeast Asia

Emergence of regional higher education cooperation framework in Northeast Asia: a historical overview

Contrary to Southeast Asia, where ASEAN has played a pivotal role in developing a regional identity and systematic higher education cooperation frameworks over the last couple of decades; there is no regional body comparable to ASEAN in Northeast Asia, that enables a multilateral approach. This is probably due to several intertwined factors having shaped the complex geopolitical situation in Northeast Asia: (i) the political tension among countries in Northeast Asia during and even after the Cold War era, (ii) rivalry between Japan and an emerging China, and (iii) the complicated historical legacies between China, South Korea, and Japan, in particular, of the first half of the 20th century (Chapman, 2010; OECD, 2004). For these reasons, in Northeast Asia, discussions on regional higher education integration started only since the early 2000s. In particular, to avoid this delicate geopolitical situation in Northeast Asia, at the beginning, these discussions usually took place in a much broader context, such as ASEAN + 3 or East Asian Summit (EAS) where ASEAN in most cases played an important mediating role.

The first critical momentum to facilitate close cooperation among these three countries was brought on by the East Asian financial crisis in the late-1990s. The crisis provided East Asian economies with a new perspective on regional cooperation, which actually introduced several significant government-led regional integration processes in East Asia. For instance, in response to the crisis, the leaders of China, Japan, and South Korea were unofficially invited to the ASEAN's 30th anniversary, which paved a way for forming the ASEAN + 3 mechanism. Through this newly created mechanism, all three major Northeast Asian countries participated in the discussions, for the first time, on intra-regional higher education cooperation in a much broader regional context of East Asia. The importance of higher education cooperation in East Asian countries has since then been actively discussed at ASEAN + 3 meetings, and

Table 5 Major recommendations & declarations emphasizing the regional integration of East Asian higher education involving China, Japan, and South Korea.

<i>Regional process</i>	<i>Attention to higher education</i>
East Asia vision group (EAVG) prospect report in 2001	To work together with cultural and educational institutions to promote a strong sense of identity and an East Asian consciousness, and to promote East Asian studies in the region through cooperative programs, teaching or languages, establishment of networks; expanding the ASEAN University Network (AUN) to the rest of East Asia, and profiting by existing bilateral initiatives between ASEAN and China, Japan, and South Korea (EAVG recommendations, section Conclusion “Social, Cultural and Education Cooperation”, points 98 and 100)
ASEAN + 3 leaders’ declaration at the 7th ASEAN + 3 summit in 2003	To promote lifelong learning programs; credit transfer systems; scholarships and exchange programs for students, faculty, staff; research and development cooperation; “centers of excellence,” including e-learning; and curricular development as bases for common regional qualification standards among interested centers/institutions (recommendations adopted in the area of education)
Kuala Lumpur declaration at the 1st East Asian summit in 2005	To enhance people-to-people exchange aimed at developing a “we feeling”; to encourage the sharing of ideas through greater interactions between students, academicians, researchers, artists, media, and youths among countries in East Asia; to conduct regular exchange of intellectuals, members of think tanks, religious personalities, and scholars, which will benefit from East Asia and the world through deeper knowledge and understandings so as to fight intolerance and improve understanding among cultures and civilizations (Articles 6, 7, 8)

Note: Modified from Yepes (2007).

Source: Yepes (2007), Kuroda and Passarelli (2009).

sometimes in an even broader context, such as the EAS that started in 2005, which expanded its membership further to Australia, New Zealand, and India. Some of the most important recommendations and declarations emphasizing the role of East Asian higher education cooperation from various regional processes involving China, Japan, and South Korea, are summarized in Table 5 above.

However, the most dramatic change in the developments of higher education cooperation framework in Northeast Asia was brought about by the establishment of the trilateral summit meeting among the leaders of China, Japan, and South Korea. Not surprisingly, as was the case with the initiation of discussions on establishing an East Asian community at the ASEAN + 3 Summit; the development of a leaders’ network among the three Northeast Asian countries was again the most important factor in accelerating recent talks on the regionalization of higher education in Northeast Asia. The trilateral summit meeting, involving exclusively China, Japan, and South Korea, first took place in 2008. Since then, these three core countries in Northeast Asia have met annually under the formal institutional framework of the China-Japan-Korea Trilateral Summit (the Trilateral Summit hereafter). This shows that these three countries consider, on top of the greater ASEAN + 3 option, the Northeast Asian option as one viable regional scope for more fruitful regional cooperation.

Seen from European experiences, however, the success of regional integration is dependent on the development of a true regional identity. This point was well stressed by then Japanese Prime Minister Hatoyama’s proposal of establishing an “East Asian Community” at the second Trilateral Summit meeting held at Beijing in October 2009, and which was again reflected as a key item in “the Joint Statement on the 10th Anniversary of Trilateral Cooperation among the People’s Republic of China, Japan, and the Republic of Korea”, adopted by the leaders of the three countries (Yonezawa and Meerman, 2010). The Joint Statement said that “we have agree [d] ... [to] continue to conduct exchanges among all sectors of the three countries, particularly friendly youth exchanges and exchanges among universities[to] consider establishing a long-term mechanism for youth and media exchanges, encourage academic institutions and local authorities, and promote closer trilateral cooperation in areas such as ... education”

In the area of higher education, this agreement had an immediate impact on the ongoing movement toward establishing a common regional framework to encourage student exchanges among the Northeast Asian countries. To implement this agreement, the governments of the three countries set up a “Joint Expert Committee for Promoting Exchange and Cooperation (Joint Expert Committee hereafter)” to discuss and develop guidelines to support exchange programs among universities in China, Japan, and South Korea. On May 29, 2010 in Jeju, South Korea, the leaders of China, Japan, and South Korea convened again at the third Trilateral Summit Meeting and agreed on the early realization of the CAMPUS ASIA program, the Northeast Asian version of the European ERASMUS program.

At this Summit, the leaders of the three countries also adopted a “Trilateral Cooperation VISION 2020”, articulating “[we] share the common recognition that by presenting specific goals and visions to be achieved through cooperation by 2020, our future-oriented comprehensive cooperative partnership will be more solid.” These two historic documents at the Trilateral Summits have so far been the most high-profile and comprehensive agreements on higher education cooperation among the three Northeast Asian countries, which finally resulted in the CAMPUS ASIA program. It was the first and most concrete multilateral student exchange initiative taken by the Northeast Asian countries, which has high potential to be further developed into a more comprehensive regional higher education cooperation framework in this region.

Table 6 List of CAMPUS ASIA consortia.

No.	Korea	Japan	China	Project theme
1	Korea University	Kobe University	Fudan University	Risk Management on East Asia
2	Dongseo University	Ritsumeikan University	GDUFS	East Asian Humanities
3	Pusan National University	Kyushu University	SJTU	Energy, Environmental Science & Technology
4	Seoul National University	Hitotsubashi University	Peking University	Asia Business Leaders Program
5	Seoul National University	University of Tokyo	Peking University	BESETO DDMP International & Public Policy Studies
6	Sungkyunkwan University	Okayama University	Jilin University	Achievement of Common Good & Reevaluation of Classical Culture in E. Asia
7	Sungkyunkwan University Seoul National University	Nagoya University	Renmin University Tsinghua University SJTU	Law and Political Science
8	KAIST	Tokyo Institute of Technology	Tsinghua University	Science & Technology
9	Korea University	Waseda University	Peking University	E. Asia Global Program for Multi-layered Conflict & Social Innovation
10	Pusan National University	Kyushu University	Tongji University	Cooperative Educational Program for Fostering HR
11	Seoul National University	University of Tokyo	Peking University	The BeST Regional Experts Education Project
12	Seoul National University of Education	Tokyo Gakugei University	Beijing Normal University	International Graduate Program for Teacher Education in E. Asia
13	Sungkyunkwan University	Nagasaki University	Shandong University	China–Japan–Korea Intercollegiate Cooperative Project
14	Yonsei University	Osaka University	Peking University Tsinghua University SJTU Tianjin University of Traditional Chinese Medicine	Medical Science and Public Health Program
15	Yonsei University	Chiba University	Tsinghua University Zhejiang University	Plant & Environment Innovation Program
16	KNU of Arts	Tokyo University of Arts	Communication University of China	Japan–China–Korea International Animation Co-work Curriculum
17	Korea Maritime & Ocean University	Tokyo University of Marine Science & Tech	Shanghai Ocean University	Collaborative Education Program in Marine Science & Technology

Source: <http://www.campusasiaprogram.kr/index.do?lang=eng>.

CAMPUS ASIA program (China, Japan & South Korea)

The CAMPUS ASIA program implemented its small-scale pilot program from 2012 to 2015. A joint call for pilot program proposals was announced in May 2011, and ten consortiums were selected through a two-stage evaluation process (one by each country and the other by a joint trilateral evaluation). Then in January 2016, the first South Korea–Japan–China Education Minister Meeting was held, whereby all three countries agreed to expand the pilot program into a full-fledged program that would include 17 consortia. As a result, the official CAMPUS ASIA program was launched in December 2016, along with the first Trilateral Rectors' Forum. And as of 2019, the CAMPUS ASIA program has expanded to include 17 consortia consisting of 40 universities. Table 6 above, shows the list of CAMPUS ASIA consortia and participating universities from South Korea, Japan, and China, along with their areas of cooperation.

Unfortunately, however, throughout the life of the CAMPUS Asia program, the Northeast Asian region has gone through a complicated geopolitical rollercoaster, with tensions between Japan–Korea constantly being rekindled over historical (i.e., comfort women) and territorial sovereignty (i.e., Dokdo/Takeshima Islands) issues, strained relations between all three states due to military issues (e.g., North Korean military threat, deployment of American THAAD missile system in Korea, Japan's increasing military presence), while the looming US–China trade war has further pushed a geopolitical wedge in the region.

As a result, the progress of the CAMPUS ASIA program has come to a stall in recent years. According to the plan agreed upon by the three countries' government in 2016, the three countries participating in CAMPUS ASIA program have been seeking to expand the scope of the program in the second stage to link it to universities in the ASEAN region, but discussions have not progressed at all due to the recent political tensions among the three countries. Therefore, in addition to the CAMPUS ASIA program among the three countries; South Korea (2016) and Japan (2013) have also participated in the AIMS program, a government-backed official student exchange platform where 7 ASEAN partner countries (Malaysia, Indonesia, Singapore, Thailand, Vietnam, Philippines, and Brunei) and 2 Northeast Asian Countries (Japan and South Korea) are involved.

Challenges for regionalization of higher education in Northeast Asia

Substantiated policy frameworks so far aimed at the integration of higher education systems in East Asia, were initiated first at the sub-region level on the assumption that (i) higher education cooperation frameworks at the sub-regional level would create better opportunities for countries in the sub-region to take full advantage of geographical proximity, and more intimately shared cultural and educational heritage of neighboring countries, and that (ii) once they matured, these sub-regional cooperation frameworks would then ultimately contribute to the formation of a Pan-East Asian higher education community. Initially it seemed as if this approach would be very efficient and somewhat inevitable at least for a while, given the huge diversity and developmental gaps existing across various sub-regions in East Asia. From this viewpoint, the regionalization of higher education in Northeast Asia was interpreted as a first step or a building block toward forming a greater East Asian higher education area.

From a more practical perspective, significant complementarities exist between China, Japan, and South Korea. For instance, in the case of South Korea and Japan; they have a strong domestic capacity to provide higher education, as domestic enrollment rates in these countries have continued decreasing in recent years due to the decline of college age students, while China still remains as the top feeder for both countries. Thus, closer regional higher education cooperation is inevitable in the near future, as it will provide both resources and markets to each other.

However, the engine that once accelerated the movement toward regionalization of higher education in Northeast Asia, has been losing steam as of late. The governments of the Northeast Asian countries are facing many complicated issues at hand, such as escalating geopolitical tensions, leadership issue over the region, the US influence on Northeast Asian higher education and most recently, the outbreak of COVID-19.

Geopolitical tensions in Northeast Asia

The geopolitical situation in Northeast Asia is much less conducive to the regionalization of higher education than the situation in Europe or in Southeast Asia. Territorial and political conflicts remain unsettled, and regional factors driving regional integration seem far more complicated in Northeast Asia (Seliger, 2011).

One important momentum triggering heavy geopolitical tensions in Northeast Asia was initiated by the deployment of the Terminal High Altitude Area Defense (THAAD) missile system (in South Korea) by the US in April of 2017, generating deep discontent in China. The Chinese government responded by embarking on business sanctions and pulling back on Chinese tourists, which hurt the sales performance of major Korean companies such as Lotte, Hyundai and Kia in China, as well as businesses in Korea's Jeju Island, whereby nearly 90% of tourists come from mainland China (South China Morning Post, 2017). Despite various efforts, it was difficult for both sides to come to a mutual diplomatic agreement, as there was a gap in understanding on the intentions of each side. This placed Korea in quite a precarious position, as it was forced to find a "middle way" between its military partnership with the US-Japan alliance, while protecting its economic partnership with China (The Korea Herald, 2017). It was during this period where Korean President Moon Jae-In declared Korea's New Southern Policy (NSP), signaling the country's strategic pivot toward the ASEAN region so as to diversify away from Korea's heavy economic reliance on China. It would also reprioritize Korea's diplomacy toward the so-called "Big Four" (US, China, Japan, Russia) in the region (The Korea Times, 2020).

Geopolitical tensions in the region were further escalated when in 2018, the South Korean Supreme Court ordered Japanese firms to compensate Koreans whom they had used as forced labor. Two Japanese companies had their assets seized in Korea, while Mitsubishi Heavy, one of the firms involved, refused to comply with the court order (BBC News, 2019). In addition, the Japanese government tightened control on exports of three chemicals essential for making semiconductors and flat panel screens used in smartphones and TVs, and by doing so threatened Korea's high-tech economy. The Abe administration then turned up the heat in August 2019, by formally taking Korea off its "whitelist", which meant Japanese companies now needed government approval to export sensitive materials to Korea (Nikkei Asia, 2019). This sparked a nation-wide boycott of Japanese goods in Korea, with anti-Japanese sentiment being spread through social networking services and other major media outlets. Hence, it ignited a mini trade war between the two countries, which not only troubled the region, but also the global supply chain system that was already reeling from a US-China trade war.

Regional leadership

With the above-mentioned geopolitical situation in mind, will the Northeast Asian region become a more coherent regional bloc of higher education integration in the near future? If so, what implications does this have for the establishment of a broader East Asian higher education area? Many scholars in the past (e.g., Kuroda and Passarelli, 2009; Mok, 2011) have argued that, as opposed to what has happened and is happening in Europe; the developments of regional higher education cooperation and its related institutions in East Asia can only be called nascent, as the majority of these agreements and institutional frameworks are either very shallow or rather sub-regional. For instance, Evans (2005, cited in Mok, 2011:20) argued that regional agreements for higher education cooperation in East Asia, such as the Kuala Lumpur Declaration, may comprise just "little more than talking and becoming familiar with one another (211)." And when considering the current state of affairs in the region, this view seems to still hold true even in the present moment.

There is not anything like the European Union or the Bologna process in East Asia. In addition, the government-level efforts in East Asia have taken place at a sub-regional level through agreements between neighboring states, rather than at a pan-regional level as in the case of Europe. For instance, ASEAN countries reached a consensus in 2003 to establish an ASEAN community, and South-east Asian Ministers of Education Organization—Regional Center for Higher Education and Development (SEAMEO RIHED) discussed an even more ambitious idea of creating a Southeast Asian Higher Education Area (SEA-HEA) within its own region (Aphijanyatham, 2010; Kuroda et al., 2010). In parallel, Northeast Asia also initiated the CAMPUS Asia program within its own regional boundary as well, however these efforts are still lagging behind in terms of becoming a truly pan-regional approach.

Why is this the case in East Asia? The underlying reasons behind the present situation need to be understood. First of all, East Asia is complex and diverse. There is less common cultural, linguistic, and religious ground in the East Asian region than in Europe. This extreme diversity poses various challenges for regional integration in East Asia, that would be difficult to settle in a short time period. Second, contrary to Europe, there is no powerful supranational body that is able to accelerate a multilateral approach. ASEAN, EAS, and Asia-Pacific Economic Cooperation (APEC) do not have enough institutional capacity to forge a strong consensus that cuts across national agendas. This is largely in part due to the emerging intra-regional rivalry between China and Japan, as a rapidly rising China aims to expand its sphere of influence in the region, while Japan leans heavier on its US-led alliances in order to control China's growing influence.

US influence on Northeast Asian higher education

Another salient feature characterizing Northeast Asia is the strong influence of the US in this region. The US is "an indirect political and economic factor in Northeast Asian integration (Seliger, 2011:4)." The area of higher education is not an exception for this US influence. As shown in Table 4 of section **Intraregional dependence among Northeast Asian countries**, Chinese and South Korean students still make up a large percentage of international students that choose to study in the US as a favored study destination, due to the ever-growing role of English as well as its better quality of higher education. These patterns and trends were threatened with the rise of American isolationism during the Trump administration years, and it seemed as if Northeast Asian countries would be coerced away from US influence in the area of higher education. However, the Biden Administration that tends to favor more multi-lateral approaches to government policies, may influence the status quo into a new direction once more.

Although in the past decade the rate of growth of cross-border activities between Northeast Asian countries and the US has slowed down; at the end of the day, the US will still inevitably influence the speed and shape of future regionalization of higher education in Northeast Asia. Yet, until now, no political consensus has yet emerged as to how this strong interdependence between Northeast Asian countries and the US can properly be taken into account in achieving a higher education integration in Northeast Asia.

Conclusion

The outbreak of COVID-19 has posed an unprecedented threat to Northeast Asia across all facets of society. In 2020 alone, the total number of confirmed cases was over 60,000 in South Korea, 87,000 in China, and 230,000 in Japan (Worldometer, 2020; World Health Organization, n.d.). The region faced similar difficulties during the 1997 Asian Financial Crisis and 2008 Global Recession, however, never has there been such a global biological threat, that has either legally prevented international students from studying abroad or caused them to voluntarily choose not to go due to health concerns. This has only further stalled the regionalization of higher education in Northeast Asia, as students, faculty and government officials from each country are simply not able to meet one another in person like in the past.

Although there have been recent talks between governments on possible hybrid forms of student mobility (i.e., combining online and offline courses/educational activities); COVID-19 has cast a cloud over an already, bleak outlook for regionalization of higher education in Northeast Asia. Nevertheless, after the full-scale outbreak of COVID-19 in early 2020, the number of online courses offered has been rapidly expanding and its potential has been newly recognized by higher education policymakers and practitioners around the world. Given the situation we are now facing, geographical proximity may not necessarily be a key consideration of exchanges in the future, unlike before, due to the nature of online courses. However, it remains to be seen whether these new challenges brought on by COVID-19 will prove to be an opportunity or a threat to higher education cooperation in the Northeast Asian region.

In this sense, future developments for the regionalization of higher education in Northeast Asia are yet rife with uncertainty. The only thing that can be said for certain is that politicians of the three countries must change their approach toward historical and territorial disputes, if they truly wish to achieve regionalization in Northeast Asia, including the domain of higher education. Collaborative efforts to ease age-old tensions rather than aggravate them will be crucial.

References

- Aphijanyatham, R., 2010. East Asian Internationalization of Higher Education: A Key to Regional Integration. A Program Report by Visiting Foreign Research Fellows. No. 25. Center for Research on International Cooperation in Educational Development, University of Tsukuba.
- Byun, K., Um, S., 2012. The regionalization of higher education in Northeast Asia. *Korean J. Comp. Educ.* 22 (3), 1–30.

- Chapman, D.W., Cummings, W.K., Postiglione, G.A., 2010. Transformations in higher education: crossing borders and bridging minds. In: Chapman, D.W., Cummings, W.K., Postiglione, G.A. (Eds.), *Crossing Borders in East Asian Higher Education*. CERC Studies in Comparative Education 27. Comparative Education Research Center: The University of Hong Kong. Springer, pp. 1–24.
- Cheong, I., 2005. East Asian Economic Integration: Korea's Perspective. Available from. <http://www.rieb.kobe-u.ac.jp/academic/ra/seminar/2004/seminar-details/cheong.pdf>. (Accessed 8 July 2011).
- Evans, P., 2005. Between regionalism and regionalization: policy networks and the nascent East Asian Institutional Identity. In: Pempel, T.J. (Ed.), *Remapping East Asia: The Construction of a Region*. Cornell University Press, Ithaca.
- Kuroda, K., Passarelli, D., 2009. Modeling TNE directions in Asia. In: *The Observatory on Borderless Higher Education*. International Strategic Information Service.
- Kuroda, K., Yuki, T., Kang, K., 2010. Cross-Border Higher Education for Regional Integration: Analysis of the JICA-RI Survey on Leading Universities in East Asia. JICA-RI Working Paper No. 26. JICA Research Institute.
- Li, X., 2007. East Asian Regional Integration: From Japan-Led "Flying-Geese" to China-Centred "Bamboo Capitalism". CCIS Research Series Working Paper No. 3, Center for Comparative Integration Studies (CCIS) Aalborg University Denmark. ISSN: 1901-9718.
- Mok, K.H., 2011. Regional Cooperation or Competition? The Rise of Transnational Higher Education and the Emergence of Regulatory Regionalism in Asia. Paper Prepared for the 2011 Senior Seminar Co-hosted by the East-West Center, UNESCO Bangkok, and Hong Kong Institute for Education. May 4–6, 2011.
- Organisation for Economic Co-operation and Development, 2004. *Internationalization and Trade in Higher Education: Opportunities and Challenges*. Organisation for Economic Co-operation and Development, Paris.
- Seliger, B., 2011. Economic Integration in Northeast Asia: Preconditions and Possible Trajectories. Available from. http://biblioteca.universia.net/html_bura/ficha/params/title/economic-integration-in-northeast-asia-preconditions-and-possible-trajectories/id/54395257.html. (Accessed 8 July 2011).
- Yepes, C.P., 2007. Regionalization of higher education services in Europe and East Asia and potential for global change. *Asia Eur. J.* 5 (83). <https://doi.org/10.1007/s10308-006-0069-z>.
- Yonezawa, A., Meerman, A., 2010. Multi-polar Initiatives in Asian Higher Education: The Repositioning of Japanese Higher Education in Times of Budgetary Constraint. Paper Presented at the 2010 General Conference Organized by the OECD Institutional Management in Higher Education, 13–15 September 2010, Paris France.

Relevant websites

- BBC News, 2019. South Korea and Japan's Feud Explained. Available at. <https://www.bbc.com/news/world-asia-49330531>. (Accessed 1 October 2020).
- Campusasia, kr, 2020. CAMPUS Asia Korea-Japan-China. Available at. <http://www.campusasiaprogram.kr/index.do?lang=eng>. (Accessed 4 October 2020).
- Koreaherald, 2017. Gap Exists Between Seoul, Beijing Over THAAD. Available at. <http://www.koreaherald.com/view.php?ud=20171124000809>. (Accessed 30 September 2020).
- Lee, S., 2020. New Southern Policy. *Koreatimes*. Available at. https://www.koreatimes.co.kr/www/opinion/2020/09/674_287872.html. (Accessed 18 September 2020).
- Mitsuru, O., Kim, J., 2019. Inside the Lose-Lose Trade Fight Between Japan and South Korea. *Nikkei Asia*. Available at. <https://asia.nikkei.com/Spotlight/The-Big-Story/Inside-the-lose-lose-trade-fight-between-Japan-and-South-Korea>. (Accessed 26 September 2020).
- Tselichtchev, I., 2017. The Real Reason Why China is Angry With South Korea. *South China Morning Post*. Available at. <https://www.scmp.com/week-asia/politics/article/2114232/beyond-thaad-real-reason-why-china-angry-south-korea>. (Accessed 1 October 2020).
- World Health Organization, 2021. WHO Coronavirus Disease (COVID-19) Dashboard. Available at. <https://covid19.who.int/table>. (Accessed 19 January 2021).
- Worldometer, 2020. Coronavirus Cases. Available at. <https://www.worldometers.info/coronavirus>. (Accessed 19 January 2021).
- Youtubecom, 2019. Japan Marks S. Korea's Dokdo Islands as Japanese Territory in Annual Defense White Paper. Available at. <https://www.youtube.com/watch?v=YW7loCaZ7pM&list=PLzLYx9A-gW8a7q15s0OWtE6eOzpRBhNn6&index=177>. (Accessed 26 September 2020).
- Youtubecom, 2020. Japan Asserts Senkaku Islands Claim in Dispute With China. Available at. <https://www.youtube.com/watch?v=9TFgEU97EIE>. (Accessed 19 September 2020).

Education in small island (and micro) states

tavis d. jules, Victoria Desimoni, and Caitlin Donnelley-Power, Loyola University Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	456
Toward a definition of small states	457
Historiography of small states research: deconstructing the different generations of small states studies	459
An ecology of educational policies and priorities in small states	461
Conclusion	463
References	463

Introduction

It is complicated to say whether small (and micro) states are more or less effective than large ones, not least because of the complexities associated with what is meant by a small state. Following small states scholars, we frequently use different terms interchangeably: small states, small vulnerable economies (SVE), microstates, small open economies, middle power, small islands developing states (SIDS),¹ to talk about small states (Armstrong et al., 1998; Commonwealth Advisory Group, 1997, 1985; Commonwealth Secretariat and World Bank, 2000; Read, 2004). In what follows, we use the term small states as a generic term to encapsulate all of the meanings above.

Many small states are considered SIDS and microstates, given their territorial and demographic economies of scale. While scale is a healthy prism to analyze the micro-climates of small states, it should not be the only way to examine the social realities of small states. Small states, which are primarily young post-colonial nations, have historically been singled out as cases for special attention, given that many of them have not achieved high levels of industrialization or economic diversification. The high degree of openness in their economies, stemming from high ratios of imports, also poses challenges to these states. While small states have sought to overcome their constraints, they are still saddled with high public expenditure on education and other social services. Moreover, many post-colonial small states are still aid-dependent, thus exacerbating the core/periphery challenge. On average, most small states receive 15% of their gross national income or USD 201 per capita from the Organisation for Economic Co-operation and Development's (OECD) Official Development Aid (ODA) program. As such, educational provision and attainment of global benchmarks in small states vary from Jamaica to Singapore. Small states differ significantly in their adoption and implementation of education reforms, with some achieving high rankings on global matrices such as the human development index, while others are perceived as vulnerable to "existential threats" (Girvan, 2010). As such, attention has been focused on strengthening institutional capacity, reducing vulnerability, adjusting to the international trade regimes, peripherality, aid dependence, and coping with the effects of globalization.

As Table 1 shows, many conceptualizations of small states are based on population size with an upper limit of 5 million. Regardless of their population size, the educational systems in small states face tremendous challenges in a changing global context. In the immediate post-independence period, the educational challenges of small states were centered on access and quality. Today, their challenges include meeting international commitments to the educational agendas of multilateral and financial institutions while implementing incremental educational reforms. While globalization has allowed small states to become less isolationary, it has also brought about unprecedented challenges for small states ranging from climate change to deregulated financial markets to changing global migration patterns to weak public sector capacity. Despite these challenges and the rise of the educational intelligent economy, many small states have managed to being closer to achieve universal access to basic education and gender parity in primary and secondary education as they have been the first to shift their educational priorities toward more pressing global commitments and issues of achieving quality, equity, and inclusion. Moreover, small states have remained committed to the educational priorities outlined in the Education for All for All (EFA) agenda, the Millennium Development Goals (MDGs), and the recent Sustainable Development Goals (SDGs).

Much of the recent literature on small states has avoided broad generalizations based on size and instead emphasizes how small states adapt to the changing nature of globalization and the onset of the fourth industrial revolution (Crossley and Sprague, 2017; Jules and Ressler, 2016; Salajan and Jules, 2019). Educational provisions in small states remain multipurpose and flexible, with schools having to accommodate the needs of diverse learners.

This chapter provides a literature review of education in small states. First, we discuss the difficulty of clearly defining small states. We present an overview of the different theories describing small states' smallness and vulnerability. We then offer a historiographical genealogy of small states research, focusing on the scholarship categorizing the two small state research generations. Third, we

¹The United Nations Department of Economic and Social Affairs (UNDESA) recognises 58 SIDS who share similar physical and structural challenges.

Table 1 Small States across the globe.

Region	Population < 1.5 million	Population, 1.5–5 million
Africa	Cape Verde; Equatorial Guinea; Gabon; São Tomé & Príncipe; Swaziland	Botswana; Central African Republic; Congo (Republic of); Eritrea; The Gambia; Guinea Bissau; Lesotho; Liberia; Namibia
Americas	French Guiana (FRORD); Suriname	Costa Rica; Panama; Uruguay
Arab States	Bahrain; Djibouti; Qatar	Lebanon; Mauritania; Oman; United Arab Emirates; West Bank and Gaza
Atlantic	Bermuda (BROT); Falkland Islands (BROT); Faroe Islands (DENSG); Greenland (DENSG); Iceland ; St Helena (BROT); St Pierre & Miquelon (FRTC)	
Asia	Bhutan; Brunei Darussalam; Macao-China (SAR); Timor Leste	Georgia; Mongolia; Singapore
Caribbean	Anguilla (BROT); Antigua & Barbuda; Aruba (NETHFA) The Bahamas; Barbados; Belize; British Virgin Islands (BROT); Cayman Islands (BROT); Dominica; Grenada; Guadeloupe (FRORD); Guyana; Martinique (FRORD); Montserrat (BROT); Netherlands Antilles (NETHFA); St Barthelemy (FROC); St Kitts & Nevis; St Lucia; St Martin (FROC); St Vincent & the Grenadines; Trinidad & Tobago; Turks & Caicos (BROT); US Virgin Islands (UST)	Jamaica; Puerto Rico (SGUT)
Europe	Andorra; Cyprus; Estonia; Gibraltar (BROT); Guernsey (UKCD); Isle of Man (UKCD); Jersey (UKCD); Liechtenstein; Luxembourg; Malta; Monaco; Montenegro; San Marino; The Vatican	Albania; Armenia; Bosnia & Herzegovina; Croatia; Ireland; Latvia; Lithuania; Macedonia FYR; Moldova; Norway; Slovenia
Indian Ocean	Christmas Island (AUST); Cocos Islands (AUST); Comoros; Mayotte (FROC); Maldives; Mauritius; Réunion (FRORD); Seychelles	
Pacific	American Samoa (UST); Cook Islands (SGNZ); Federated States of Micronesia; Fiji Islands; French Polynesia; Guam (SGUT); Kiribati; Marshall Islands; Nauru; New Caledonia (FRORD); Niue (SGNZ); Norfolk Island (AUST); Northern Marianas (SGCUS); Palau; Samoa; Solomon Islands; Tokelau (NZSAT); Tonga; Tuvalu; Vanuatu; Wallis & Futuna (FROC)	New Zealand

Countries in **bold** are members of the United Nations. Countries underlined are members of the Commonwealth. Data refer to 2008. *Abbreviations:* AUST – Australian Territory Administered from Canberra; BROT – British Overseas Territory; DENSG – Self-governing Overseas Administrative Division of Denmark; FROC – French Overseas Collectivity; FRORD – French Overseas Regions and Departments; NETHFA – Part of the Kingdom of The Netherlands with Full Autonomy in Internal Affairs; NZSAT – New Zealand Administering Territory; SAR – Special Administrative Region of the People's Republic of China; SGCUS – Commonwealth in Political Union with the USA; SGNZ – Self Governing in Association with New Zealand; SGUT – Self-Governing Unincorporated Territory of the USA; UKCD – United Kingdom Crown Dependency; UST – Unincorporated territory administered by USA Office of Insular Affairs.

Based on Crossley and Sprague (2012, p. 28).

explore education policies in small states. Finally, we offer some new directions in studying small states in the face of the fourth industrial revolution and the educational intelligent economy.

Toward a definition of small states

Despite the various attempts to define small states, there is still no single agreed-upon definition (Baldacchino, 2012a; Bacchus, 2008; Crowards, 2002; Crossley and Sprague, 2012; Jules and Ressler, 2017a,b). Most definitions of small states have created a “fundamental definitional ambiguity” (Long, 2017, p. 144) and are embedded in the colonial heritage of statehood. The colonial roots of current understandings of small states appear in the history of the Commonwealth Parliamentary Association (CPA), which was formed in 1911 to coordinate Britain's colonies and small protectorates and further centralize imperial power in the hands of the UK (CPA, n.d.). In 1981, the CPA recognized small states as having distinctive concerns that need to be addressed, and many formerly colonized small states have become members of the CPA. As such, current understandings of small states tend to begin with the definition derived “for” and “by” the Commonwealth States' London Declaration and are thus embedded in colonial hierarchies (Commonwealth Secretariate, 1949).

Early definitions of small states differentiated *small societies* from *small territories* and suggested that small societies exist in large states (Benedict, 1967). In this way, being small has been regarded as an impediment to state action and survival, and it is tied to “a specific spatio-temporal context” (Thorhallsson and Wivel, 2006, p. 654). However, small states have come to be defined by some common characteristics, which ranges from population size (1.5 million or less, though there are a few exceptions), geography (two-thirds of small island states are small states), land area, gross national product, vulnerability to natural disasters and climate change, their debt burden (which is usually higher for these countries than other developing countries), to how their characteristics make it difficult for them to fit the standard Western-centric post-colonial development model. However, as Table 1 shows, there are

challenges with these measurement characteristics, particularly around population size. For example, Kuznets (1960) sets an upper population limit of 10 million for small states; today, 134 countries could be considered small by this gauge. By the current measurement of between 1.5 and 5 million people (Armstrong et al., 1998; Bacchus, 2008; Commonwealth Secretariat, 2009; World Bank Group, 2016), a few large countries (e.g., Jamaica, Lesotho, Namibia, and Papua New Guinea) are also categorized as small since they share several of the same characteristics of smallness – including small Gross Domestic Products (GDP) (Commonwealth Secretariat/World Bank Task Force, 2000).

Additionally, the literature highlighting vulnerabilities associated with smallness typically makes assumptions that “small states share a high degree of conformity” (Jules and Ressler, 2017a,b, p. 21). However, changing global economies and definitions of scale mean that formal demographic characteristics defining small states may need to be re-examined. Jules and Ressler (2017a,b) argue that researchers should focus on the advantageous aspects of smallness rather than constraints in the absence of one clear definition of small states. Such thinking allows us to question categories such as “smallness” or even “state.” For example, some literature describes sub-national areas and jurisdictions within a large urban metropolis, such as the favelas of Rio de Janeiro, as exhibiting, facing, and responding to some of the characteristics that define small states (Straubhaar, 2017).

The discourse around small states frequently focuses on vulnerability narratives such as geography, demography, and economy to changing global markets, natural disasters, and limited productive and infrastructural capacity (Baldacchino, 2011; Jules, 2012). This narrative has led to a pervasive “deficit discourse” that centers on four perceived characteristics defining small states, including “exaggerated personalism, limited resources, inadequate service delivery and donor dependence” (Sutton, 2006, p. 13). The consequences of these vulnerability narratives have left small states frequently understood to be “passive recipients” (Jules and Ressler, 2017a,b) of uncritical educational transfer of global policy discourse. While all small states are unique, their smallness places the “idiosyncrasy of the state at the center of the stage” (Brock, 1987, p. 12) in their educational priorities.

While small states experience many of the same complex problems as big states, they are not smaller versions of large states but rather entities with distinct challenges, needs, and responses. Despite their distinct needs, educational policies in small states are often generalized copies of development trajectories that originated in states with very different contexts (Baldacchino, 2011; Brock, 1987; Brock and Crossley, 2013). The external push toward the educational policy transfer process that does not critically examine small states’ particular contexts is a challenge in and of itself. Baldacchino (2011) discusses the implications of this trend in education, where he argues that the transfer of policy used in large states has little bearing on the realities of small states and notes that “perhaps that is also one of the destinies, and definitions of a small state: not to be equipped with formal instruments of teaching and learning that speak directly and meaningfully to its social, economic, political and other narratives” (p. 461). Although small states are often viewed as takers of global policies, recently, we have begun to witness that as political and economic political and economic contexts become increasingly important in policy discussions, and the practice of uncritical policy educational transfer has declined (Jules and Ressler, 2016).

Small states face particular vulnerabilities that correspond to inherited colonial histories, population limits, and sensitivities to external markets; however, smallness can also operate as an essential strength in a rapidly changing global economic and political landscape. Many existing definitions of small states cite economic dimensions such as “limited institutional capacity and aid dependence” (Jules and Ressler, 2016, p. 30) as evidence of small states’ vulnerability. Some small states’ reliance on a single export for income has led to these nations’ characterization as being particularly vulnerable to external market fluctuations and shocks (Bacchus, 2008; Brock and Crossley, 2013). Moreover, small population size may prohibit small states from participating in economies of scale (Brock and Crossley, 2013). Baldacchino (2011) describes “small scale syndrome” as a defining characteristic of small states and small economies, in that it gives rise to the dynamics of “intimacy, monopoly, and totality, with the only realistic exit option being emigration” (p. 461). It has been argued that this particular narrative originated and is primarily perpetuated outside of small states themselves and tends to rely on stereotypical imaginaries of small states (Baldacchino, 2011).

Several essential considerations emerge in literature around small states’ strengths in responding to geopolitical and economic challenges: adaptability to change, the geostrategic projection of smallness, and the emergence of transnational coordination (Crossley and Sprague, 2017; Jules and Ressler, 2016). Frequently, small states, microstates, and particularly SIDS, are the first to experience the effects of climate change. Crossley and Sprague (2017) argue that “this ‘sharp end’ experience has positioned small island states as forerunners in the development of new ideas, approaches, and technologies to combat these challenges” (p. 44). Indeed, the “sharp end” experiences of small states have resulted in an increasing number of small states engaging with cross-sectoral approaches to sustainable development, including incorporating UNESCO’s Education for Sustainable Development (ESD) goals (Crossley and Sprague, 2014). The “Islands of Inspiration” report commissioned by the Commonwealth focuses on contextualizing good practices in small state education systems that seek to incorporate ESD goals. A similar argument is discussed by Jules and Ressler (2016) when they point to typical small-state behaviors like adaptability, strategic flexibility, and informality being invoked during times of crisis. Many small states have exercised their relative systemic flexibility to adapt to emerging crises better than larger states (Jules and Ressler, 2016). Additionally, Jules & Ressler (2016) argue that it is crucial to investigate how small states define themselves. Often, small states project smallness only when it is advantageous for them and do not necessarily identify as “small” otherwise (Jules and Ressler, 2016). This projection has been used in instances of cooperation where several small states share knowledge and resources to develop regional solutions to problems (Jules and Ressler, 2016). Several studies argue that much can be learned from reframing discussion of these challenges from chronic vulnerability toward investigating how small states respond to changes and threats (Brock and Crossley, 2013; Crossley and Sprague, 2017; Jules and Ressler, 2016).

It is clear from recent literature on small states that a new conceptualization of “smallness” that departs from the orthodoxy of the vulnerability or deficit discourse is necessary. A new model of analyzing small states in policy research could recognize the particular challenges that small states face, indicate their strengths, contribute to international policy discourse, and avoid over-generalizing the similarities between small states.

Historiography of small states research: deconstructing the different generations of small states studies

Given the lack of a clear definition of what constitutes small states, Jules and Ressler (2017a,b) argue that small states’ research can be broken down into “first generation studies,” and “second generation studies,” which are not chronological. *First generation* or “orthodoxy studies” are trapped in the vulnerability thesis, which exhibits the notion that small states are “special cases” for investigation. Spearheaded by “the international community that developed an interest in the explosion of smallish sovereign units onto the world stage after the 1940s” (Baldacchino, 2011, p. 458), small state research began with the September 1957 conference of the International Economics Association. Under this thesis, colonialization has created a high degree of conformity and distinctively open economies, and thus, until the formation of the World Trade Organization in 1995, small states had been able to utilize the networks of their former colonizers to secure preferential market treatment for their goods. However, “while this may be true, for example, within the Commonwealth grouping, it does by far not apply to all states that can be considered small based on certain formal criteria like population size” (Jules and Ressler, 2017a,b, p. 21). This line of inquiry often emphasizes how post-colonial societies of the South are poor, fragile, and vulnerable, and victims of an oppressive global economic system. In essence, these small states’ studies centered on population size, ecology, vulnerability to external shocks, limited human and natural resources, nature of their economies, cost per capita of services, and dependence on trade or tourism, and established that smallness was vulnerable and fragile (Briguglio, 1995; Bune, 1987; Jules, 2012; Holmes, 1976).

The work of the 1985 Pan-Commonwealth meeting resulted in the influential publication of *The Challenges of Scale* (Bacchus and Brock, 1987). In education, the work of Bacchus and Brock (1987) was the first to look at the “behavioral characteristics” of smallness by focusing on the economic impediments of development in areas like South-South migration (Bartlett, 2012); the politics of education in small states (Brock, 1988; Grant, 1993); size and survival (Lewis, 1976); implications of smallness for educational development (Bray and Fergus, 1986; Atwood and Bray, 1989); the role of indigenous knowledge development and values upon the policy process (Holmes and Crossley, 2004); expansion, improvement, and equality in education (Bird, 1984; Grant, 1987; Payne, 1980); the impact of isolation and dependence (Brock, 1984); research, institutional and evaluation capacity building (Crossley, 2008; Crossley and Holmes, 1999; Louisy, 2001); financial and human capital limitations (Jules, 1994); the impact of donor aid on local decision-making (Jules, 2006, 2010; Louisy, 2001); adult education in small states (Jules, 2006); higher education financing (Baldacchino and Farrugia 2002; Nkrumah-Young et al., 2008); cooperation and collaboration (Jules, 2008, 2012), and post-socialist transformation (Jules, 2014). These works shaped and perpetuated the vulnerability narrative while recognizing only a few strengths given that they tend to focus on “political and formalistic definitions of ‘capacity’ and ‘capacity building’” (Baldacchino, 2010, p. 65).

The first-generation studies used economic vulnerability as the critical measurement of “small size, insularity and remoteness, proneness to natural disaster, environmental factors, and other characteristics, such as demography and external dependency” (Jules and Ressler, 2016, p. 24) and thus characteristic of small states. In the words of Baldacchino (1995), illustrative concepts of smallness syndrome have

included vulnerability, scarcity, combustibility, non-viability, indivisibility and intimacy; while the metaphors of an overcrowded barracoon; a ship-wrecked community; a skeleton crew; a phoney terrestrial paradise; the flotsam and jetsam of empire; forlorn peripheries always meant to be part of something larger; boats set adrift; or envious paragons of beauty and democratic consensus, are poignant and evocative. (p. 264)

While many first generation studies are rooted in the perception of the weakness inherent in small states as they sought to respond to both endogenous and exogenous influences, a few scholars have focused on the ecology of smallness by highlighting that “insularity, smallness of scale and remoteness of location combine in different degrees ... to further promote locale-specific educational situations, the problems associated with them being less amenable to resolution through theories and methods developed from the experience of larger industrialized nations” (Brock, 1990, p. 3). Such locale-specific educational problems stem from the fact that “the style of educational development ... [in small states] is too frequently modeled on what is appropriate and fashionable in large states” (Commonwealth Secretariat, 1986, pp. 5–6).

Part of the reason for this limited scope of first-generation studies is the conceit that small states are often perceived as having to overcome smallness, territorial compactness, monocrops, and monoservice economies. First-generation studies did not identify educational advantages that small states possess, such as territorial compactness and economic diversification in executing educational policy reforms. Rather, these advantages were often downplayed, and the focus became on scale, isolation (spatial, political,

social, and cultural), dependency (economic, cultural, and educational), and remoteness (Brock, 1987). Yet, these characteristics have allowed small states to combine cooperation and small scale in education to establish and achieve the provision of tertiary education through entities such as the University of the West Indies (established in 1948 with 17² contributing members) and the University of the South Pacific (established in 1968 and jointly owned by the governments of 12³ member countries). In first generation research, such accomplishments, which are unique as these are the only two universities in the world to be jointly owned and simultaneously operated by several governments, are not marked as small state exceptionalism. Similarly, small states in the Caribbean have used cooperation to set up regional assessment and examination bodies such as the Caribbean Examination Council (CXC) and regional planning projects and programs such as the Pacific Regional Initiatives for the Delivery of Basic Education (PRIDE). First generation studies often overlooked the development of indigenous or regional models as their focus was on showing that small systems are prone to “incoherence in policy making” (Rodhouse, 1987, p. 30), thus leaving them exposed. Moreover, first generation scholars highlighted that limited capacity is exacerbated by power asymmetry and how smallness reduces competition and gives rise to monopolistic or oligopolistic arrangements in the marketplace rather than stressing the benefits derived from informality, intense personalization less hierarchical societies (Baldacchino and Wive, 2020).

In the second generation studies, small states’ definition is developed beyond measures such as population, landmass size, and vulnerability. Attention is placed on resilience and flexibility rather than vulnerability, since vulnerability implies that a state has little or no control. For instance, what constitutes vulnerability is frequently inadequate, and in certain respects, some small states are not more vulnerable than many big states. Moreover, vulnerability is a somewhat subjective concept that is often ideologically charged and deeply embedded in prevailing notions of (Western) progress. In this way, second generation studies move from *priori definitions* to *posterior conceptualizations* of smallness by problematizing concepts such as size and scale and offering evidence of a “small scale syndrome” (Baldacchino, 1997), which is based on a “syndrome of behavioral issues” (Baldacchino, 2012b). Instead, second generation studies concentrate on the role of globalized nation-states and their trading status – both intraregional and inter-regional – to emphasize GDP and resources spent in trade. In other words, second generation studies move away from the vulnerability and fragility thesis and place attention on the strengths of smallness “while at the same time recognizing the consequences of the fragilities and vulnerabilities that many small states possess as part of their inherited material baggage” (Jules and Ressler, 2017a,b, p. 24). Second generation scholars recognize that

globalization has proven to be a double-edged sword for small states allowing most of them to boost trade and avoid poverty, but at the same time increasing vulnerabilities and dependency for many as a consequence of unconventional security risks stemming from mass migration, terrorism, money laundering and environmental degradation.

Baldacchino and Wive (2020, p. 2).

These studies draw attention to how small states use forums such as South-South cooperation and transfer and mechanisms such as soft power to project a stature of influence to garner diplomatic leverage (Baldacchino, 2000). Scholars under this paradigm recognize that in an era of hyperconnectivity, the “deficit discourse” (Baldacchino, 2012a) and the “economic vulnerability index”⁴ – which measures small size, insularity and remoteness, proneness to natural disaster, environmental factors, and other characteristics, such as demography and external dependency – (Briguglio, 1995) are obsolete since large populations are no longer confined to particular spaces, topographies, and ecologies (Baldacchino, 2012b). They aim to highlight how small states use “strategic flexibility,” which asserts that many small states develop nimble responses to both threats and opportunities in a rapidly changing global environment (Baldacchino, 2012a,b; Briguglio, 1995). Thus, they argue that it is quite possible that “small states thrive in the modern global order” (Baldacchino and Bertram, 2009, p. 142) as they can in providing “niche-filling export strategy, flexible specialization, enhanced entrepreneurship and economic deregulation” (Briguglio, 1995, p. 1624).

Given that most small states exist as part of a geographical typology, second generation studies draw attention to how new hemispherical and regional players, such as customs unions, regional trade agreements, and “trans-regional regimes” (Jules, 2012) and the shift from “inter-regionalism” – the relationship between two distinct, separate regions – to “trans-regionalism” – that is, common “spaces” between and across regions that constituent agents shape (e.g., individuals, communities, and organizations (Dent, 2003)) – is reorganizing the ways in which we think about small states. In some cases, small states eschew globalization trends to pursue multilateralism at the regional level. Here we can see regional states amplifying their own educational initiatives. This multilateralism can be referred to as educational exceptionalism, wherein small states come together to create a “policy space in which educational policies, performances, and practices are inhabited at the regional level” (Jules, 2015, p. 2). In the second generation, scholars “accept that there is no unified definition of small states in an interconnected world where large populations are no longer confined to particular spaces, topographies, and ecologies” (Jules and Ressler, 2017a,b, p. 24).

²Contributing countries are Anguilla, Antigua, Bahamas, Barbados, Belize, Bermuda, British Virgin Islands, Cayman Islands, Dominica, Grenada, Jamaica, Montserrat, St. Kitts & Nevis, St. Lucia, St. Vincent, Trinidad and Tobago, Turks and Caicos. There are five campuses - Mona, Jamaica in 1948; St. Augustine Trinidad in 1960; Cave Hill, Barbados in 1963; Open Campus in 2008 and the Five island Campus in 2019.

³These are the Cook Islands, Fiji, Kiribati, Marshall Islands, Nauru, Niue, Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu and Samoa. The main campuses are Laucala, is in Fiji, the Alafua Campus in Samoa the Emalus Campus in Vanuatu.

⁴The Economic Vulnerability Index is composed of trade openness, export concentration, dependence on strategic imports and peripherality.

The political ecologies of small states have been shaped by existential threats such as structural adjustment programs (SAPs), which saw the imposition of externally driven modules of development as a way to achieve images of Western modernity. Small states are indeed vulnerable to external shocks, but how they respond makes them a special case for studying. As such, second generation scholars have sought to move the pendulum away from fragility and vulnerability (e.g., economically or ecologically) and toward flexibility and strategic alliances; there have been new conceptualizations of what it means to be viewed as a small state. While smallness is often based upon institutional constraints in that

[A] small state is a state that either believes it is small, and/or else is seen to be one, and is expected to behave accordingly; also because of its historical unfolding and resource availability. '[Q]uite convincingly, it can be argued that a state is 'small' when it feels and acts small – implying that it could become smaller or less small at different points in its history.

Baldacchino (2012a, p. 31).

Jules and Ressler (2017a,b) advance that small states use “geostrategic leveraging” to rely on networks to constrain their efforts on the best possible solutions. Educational geostrategic leveraging refers to “the use of soft power grounded in strategic-level bargain and cooperation at the national level to achieve regional consensus” (Jules and Ressler, 2016, p. 28). This implies using soft power grounded in strategic-level bargaining and cooperation to project the perceived notion of smallness and fragility when it is strategically advantageous. In this way, size is viewed as providing states with opportunities since “decision-making and adaptation may be easier where there is a small population and shorter lines of communication” (Keating, 2021, p. 224) and “small state citizens may need to seek out and grasp every opportunity for purposeful economic activity” (Baldacchino, 2011, p. 457).

There is also a growing recognition that while small states may not have military strength or research capability, they may have opportunities that their larger neighbors may not. For example, they may be able to innovate and respond flexibly and rapidly to shifts in the external environment. Moreover, small states, especially islands, are the first to experience the impact of climate change. In this area, there is a lot that researchers can learn from small states and their strategies to cope with climate challenges. Thus, small post-colonial states are now global players.

An ecology of educational policies and priorities in small states

Small states face challenges that have essential impacts on these states’ education policies and provide valuable insights into potential responses to these challenges. The “social ecology” of these societies is such that they are “closely knit, integrated but open communities with highly personalized relationships” (Farrugia and Attard, 1989, p. 17). While small states are more responsive to educational reforms given their size, educational specialists in small states need different skills from their counterparts in larger states. Although various factors shape small states’ priorities in education, policymakers and planners may become dependent on external actors. The primary educational challenge for small states is providing quality education for a small number of students. Small states tend to exhibit a higher unit cost in their public expenditure devoted to education. They also lack local research capacity building. Several scholars have extensively discussed how resource dispersal issues, lack of capacity, institutional coordination, absence of professional collaboration, local research and evaluation, tensions between local educational content and international expectations, and barriers to information and communication technologies have challenged educational provision in small states (Baldacchino, 2010).

A significant criticism of education difficulties in small states is that these challenges stem from a system developed through pervasive colonial penetration and a peculiar colonial administrative legacy. Educational developments in small states are a product of Western colonization that was metropolitan orientated, and “governments have criticized the alien educational system they inherited with political independence: ‘Like Ruth amidst the alien corn’” (as cited in Baldacchino, 1995, p. 272). Bacchus (2008) points to the small population size limiting career opportunities and demand for workers with high education levels. While the number of trained personnel is limited, most external educational projects rely on foreign technical experts who do not know small states’ social and cultural roots. The “blocked mobility” (Bacchus, 2008) that characterizes small states’ career landscape causes highly educated people to pursue jobs in larger or more developed states. This trend extends into the teacher workforce that small states have available to them (Edwards-Gums, 2014). Trained teachers in small states may be recruited by large states with higher economic capacity and mobility, which severely limits small states’ ability to retain trained teachers (Edwards-Gums, 2014). In Anguilla, for example, much of their teacher workforce is made up of high school graduates who view teaching as a stepping stone for other career paths making teacher turnover and finding trained teachers a major issue (Edwards-Gums, 2014). First generation studies have maintained that relationship among individuals in small states tends to be communal instead of associational (Bacchus, 2008). This has consequences for education as well. First generation scholars view education as a means of improving cultural and material wellbeing as well as instrumental in endangering the social and economic mobility of the poor. Education in small states is therefore viewed as multipurpose.

The challenges associated with limited population are compounded by the population distribution trends in some SIDS, mainly where populations are scattered across large internal distances or several different islands, which affect educational provision. This distribution makes educational resources and infrastructure challenging to reach all populations (Bacchus, 2008). Other problems that come from the limited population include, but are not limited to, restricted internal markets for the products of small states,

difficulty in attracting investments or receiving loans (which comes with their propensity to remain poor and inability to provide the quality of education needed), uneven spending of public funds in urban areas in contrast with less populated areas (which creates difficulty to provide quality education in suburban and rural areas) (Bacchus, 2008). Given that small states have sometimes been characterized as economically vulnerable because of limited internal markets, high levels of income inequality, and reliance on tourism or a single major export which may be devastatingly disrupted during periods of global crisis or climate change, all have implications for these states' ability to fund and sustain their education systems (Bacchus, 2008).

The tendency of international discourse and literature to emphasize small states' vulnerabilities has led to a problematic view of these entities as passive recipients of global "best practice" and uncritical policy transfer (Baldacchino, 2011; Brock and Crossley, 2013; Jules and Ressler, 2017a,b). Prince (2012) argues that policy transfer frequently acts as a normative project where external, private sector consultants are increasingly tasked with making links between hegemonic projects and other nation-states. This kind of policy transfer has resulted in a lack of educational policies that are relevant and sensitive to the particular contexts and landscapes of small states (Baldacchino, 2011; Brock and Crossley, 2013). Education systems are essential for their economic growth, and "most small states have to depend on their human resources as the main factor in their production efforts and therefore the quality of these resources is key to their economic growth. A major consideration in preparing high quality resources for their labor force is their educational systems" (Bacchus, 2008, p. 132). Indeed, small states often achieve lofty global education development goals like accessibility and gender parity earlier than their larger counterparts (Crossley and Sprague, 2014). As such, reframing education policy discussion in small states is an essential step in moving away from a deficit discourse and exploring how the international community can learn from these countries (Brock and Crossley, 2013). Here we suggest that small states have responded to the challenges associated with their size and geographic realities in education policy by embracing regionalism, new mutualism, cooperative educational transfer, and educational geostrategic leveraging.

Recent literature on policy responses of small states suggests increasing participation in networks engaging with "new regionalism," or 'open regionalism' centered on open membership, regional and global trade liberalization, trade facilitation, and the inclusion of regional level into the global market" (Jules, 2012). New regionalism also departs from more traditional modes of cooperation between regions and moves toward cooperation within regions (Jules and Ressler, 2016). This new regionalism has implications for education policy in small states as well, as education governance moves away from states as the sole legitimate education authority and toward regional bodies. Regional educational bodies provide opportunities for the "mitigation of transactional costs, sharing of policy best practices and techniques, and dissemination of information" (Jules and Ressler, 2016, p. 32). The kind of regionalism documented among small states is characterized by new mutualism operationalized through cooperative educational transfer. New mutualism is the idea that the pooled resources of states forming a regional body will facilitate "regional solutions to national problems in an era of enhanced regionalism" (Jules, 2012, p. 270). Central to new mutualism is the process of cooperative transfer, which allows for national policy responses and best practices to be shared at a regional level and subsequently disseminated back to the national level through regional directives or benchmarks (Jules, 2012). In the Caribbean, the Caribbean Community (CARICOM) visibly engaged in new mutualism and cooperative transfer in their education policy response to the HIV/AIDs epidemic as multisectoral regional mandates shaped national reforms. Small states have adopted a similar model in the South Pacific through their Educational Quality and Assessment Program, which seeks to align educational approaches, practices, and goals across different member states (Educational Quality and Assessment Program, 2019). New mutualism offers a framework that explains several education policy responses in small states and aligns with increasingly important regional education bodies and goals.

It has been documented that many small states engage in geostrategic leveraging as a part of new regionalism and cooperative transfer (Jules and Ressler, 2016). Thus, small states use their comparatively flexible institutional and bureaucratic systems to strategically adapt to global changes in education areas. In many cases, smallness makes it easier for these states to reach specific educational benchmarks since "the small scale of their education systems has enabled many small states to move ahead relatively rapidly in terms of many of the international targets established for basic education" (Brock and Crossley, 2013, p. 399). These strengths position small states to attract investments, diversifying single-export economies that have typically been considered a significant vulnerability. In these ways, educational geostrategic leveraging demonstrates that the strategic thinking that small states engage in is often ignored in small state literature and indicates how small states are equipped to respond to external forces.

Much of the recent literature on education policy in small states argues that the international community can learn how small states respond to challenges through policy (Brock and Crossley, 2013; Crossley and Sprague, 2014; Jules and Ressler, 2016). Small state research that participates exclusively in the deficit discourse has historically surrounded small states ignores how smallness has been integral in many states' ability to respond to changing economic, political, and environmental landscapes. Additionally, the emphasis on vulnerability and corresponding tendency toward uncritical policy transfer limits broader international discussion about how policy, benchmarks, and success metrics are determined and implemented. Brock and Crossley (2013) argue that the distinct "small state experience can help to caution against the similar replication of a new set of fixed, universal and inflexible educational goals and targets" (p. 399) that have been typical of global educational policy discourse. Educational policy research surrounding small states must rigorously engage with new forms of collaborative governance, attention to contextual nuance, and grapple with changing definitions of scale.

In response to the multipurpose nature of education in small states and the developing smallness syndrome, regional cooperation in education or "educational regionalism," particularly in the small states of the Caribbean, has emerged as there has been a movement from an idea of national development and nation building, toward the notion that modernization comes through

regionalism and economic coordination. Today, small states realized that they must band together to fend off existential threats, and economic regionalization in functional areas, such as health, education, and transportation, has emerged as a way to define small state exceptionalism. Participation in regional projects is seen as engendering educational cooperation as educational developments in small states are now seen not as nationalistic instruments but as regional projects. Regional cooperation in education acts as a buffer against economic exploitation.

Conclusion

It is important to remember that small states are not just “scaled-down varieties” of large states. This chapter discusses the historical legacy and explanatory factors influencing small states, which tend to occupy a marginal position in global affairs and are often only heard once they band together. This chapter ultimately suggests that there are two generations of research on small states. The first generation focuses on almost exclusively on the vulnerability narrative, and the second generation pays attention to the strengths of small states. It then considered education in small states by paying attention to the challenges that small states face and how they have sought to overcome these, given that many small states suffer from economic vulnerability and are prone to external economic shocks. While the structures and constraints faced by small states, which are often ex-colonial states, are the same as those faced by larger states, what differentiates the two is how small states can circumvent these challenges. As Bacchus (2008) reminds us, the movement toward deeper economic globalization driven by the information age and knowledge society has only dismantled trade barriers and preferential systems that small states relied upon. The rise of the so-called fourth industrial revolution – which blends the physical with the cyber-physical and the “educational intelligent economy,” which is viewed as “the exponential production of digital data to measure, analyze and predict educational performance in comparative perspective” (Salajan and Jules, 2019, p. 1) – has had drastic consequences for educational developments in small states. No longer can small states choose to be insular or play second fiddle to other states. Instead, they must exert pressure where necessary and maximize their human development capacities. While the level of education is often higher in small states, and some small states have found success in educating human infrastructure for the global service economy through expanding access to basic education, they continue to struggle with producing quality education.

While several small states have met the poverty alleviation agendas and EFA and MDG objectives, many have a ways to go while working on SDG goals. Most small states have been marginalized from shaping international educational agreements in the last decade. For countries where data exists, many small states have net primary enrollment rates of over 90% but still face challenges of access and equity. With the advent of the knowledge economy and the transition to the educational intelligent economy, there has been a demand to expand education in small states. Today, small states must walk a fine line between promoting national values at home to engender development and promoting cosmopolitan values abroad (Baldacchino and Wive, 2020). With the rise of the educational intelligent economy, educational developments in small states will rely on their human resource development. As Bacchus (2008) reminds us, “small states will only be able to survive economically and culturally if their populations receive an education that helps them to perceive new opportunities and provide them with the competencies and dispositions that would equip them to adapt readily to the changing social and economic world environment” (p. 142). In this way, small states today need to have a higher degree of flexibility since they cannot change their size and disadvantages, given that they are on the global periphery and frequently rely on international aid. Finally, this chapter argues that small states are unique and require a distinctive category of analysis.

References

- Armstrong, H., Jouan de Kervenoael, R., Li, X., Read, R., 1998. A comparison of the economic performance of different micro-states, and between micro-states and larger countries. *World Dev.* 26 (4), 639–656.
- Atwood, C., Bray, M., 1989. Multiple Shift Schooling: Design and Operation for Cost-Effectiveness. Commonwealth Secretariat, London.
- Bacchus, M., Brock, C., 1987. The Challenge of Scale: Educational Development in the Small States of the Commonwealth. Commonwealth Secretariat, London.
- Bacchus, M.K., 2008. The education challenges facing small nation states in the increasingly competitive global economy of the twenty-first century. *Comp. Educ.* 44 (2), 127–145.
- Baldacchino, G., Wivel, A., 2020. Small states: concepts and theories. In: Baldacchino, G., Wivel, A. (Eds.), *Handbook on the Politics of Small States*. Edward Elgar Publishing Limited, U.K., pp. 2–19.
- Baldacchino, G., 1995. Labour formation in small developing states: a conceptual review. *Compare* 25 (3), 263–278.
- Baldacchino, G., 1997. *Global Tourism and Informal Labour Relations: The Small Scale Syndrome at Work*. Mansell, London.
- Baldacchino, G., 2000. The challenge of hypothermia: a six-proposition manifesto for small island territories. *Round Table* 89 (353), 65–79.
- Baldacchino, G., 2010. Entrepreneurship in smaller jurisdictions: appraising a glocal elite. In: Mayo, P. (Ed.), *Education in Small States: Global Imperatives Regional Initiatives and Local Dilemmas*. Routledge, London, pp. 64–78.
- Baldacchino, G., 2011. The small state in higher education: a conceptual assessment. *Teach. High. Educ.* 16 (4), 457–467.
- Baldacchino, G., 2012a. Meeting the tests of time: small states in the 21st century. *Curr. Issues Comp. Educ.* 15 (1), 14–25.
- Baldacchino, G., 2012b. Governmentality is all the rage: the strategy games of small jurisdictions. *Round Table* 101 (3), 235–251.
- Baldacchino, G., Farrugia, C., 2002. *Educational Planning and Management in Small States: Concepts and Experiences*. Commonwealth Secretariat, London.
- Baldacchino, G., Bertram, G., 2009. The beak of the finch: insights into the economic development of small economies. *Commonwealth J. Int. Aff.* 98 (401), 141–160. <https://doi.org/10.1080/00358530902757867>.
- Bartlett, L., 2012. South-south migration and education: the case of people of Haitian descent born in the Dominican Republic. *Compare J. Compar. Int. Educ.* 42 (3), 393–414.
- Benedict, B., 1967. Sociological aspects of smallness. In: Benedict, B. (Ed.), *Problems of Smaller Territories*. University of London/Athlone Press, London, pp. 45–55.

- Bird, E., 1984. The university of the West Indies in the 1980s. *Prospects* 14 (3), 413–421.
- Bray, M., Fergus, H., 1986. The implications of size for educational development in small countries: Montserrat, a Caribbean case study. *Compare* 16 (1), 91–102.
- Briguglio, L., 1995. Small island developing states and their economic vulnerabilities. *World Dev.* 23 (9), 1615–1632.
- Brock, C., 1987. Problems of policy and development. In: Bacchus, K., Brock, C. (Eds.), *The Challenge of Scale*. Commonwealth Secretariat, London, pp. 11–25.
- Brock, C., 1984. Scale, Isolation, and Dependence: Educational Development in Island, Developing and Other Specially Disadvantaged States. Commonwealth Secretariat, London.
- Brock, C., 1988. Education and national scale: the world of small states. *Prospects* 28 (3), 302–314.
- Brock, C., Crossley, M., 2013. Revisiting scale, comparative research and education in small states. *Comp. Educ.* 49 (3), 388–403.
- Bune, P., 1987. Vulnerability of Small States: The Case of the South Pacific Region and Fiji, pp. 85–87. *Courier* 104.
- Commonwealth Advisory Group, 1997. *A Future for Small States: Overcoming Vulnerability*. Commonwealth Secretariat, London.
- Commonwealth Consultative Group, 1985. *Vulnerability: Small States in the Global Society*. Commonwealth Secretariat, London.
- Commonwealth Parliamentary Association (n.d.). *Our History*. London: Commonwealth Parliamentary Association. [online] Available at: <https://www.cpahq.org/about-us/our-history/>.
- Commonwealth Secretariat, 1986. *Vulnerability: Small States in the Global Society*. Commonwealth Secretariat, London.
- Commonwealth Secretariat, 2009. *Small States: Economic Review and Basic Statistics- Volume 13*. Commonwealth Secretariat, London.
- Commonwealth Secretariat, 1949. *The London Declaration*. Commonwealth Secretariat, London.
- Commonwealth Secretariat and World Bank Task Force, 2000. *Small States: Meeting Challenges in the Global Economy*. Commonwealth Secretariat, London.
- Crossley, M., 2008. The advancement of educational research in small states. *Comp. Educ.* 44 (2), 247–254.
- Crossley, M., Holmes, K., 1999. *Educational Development in the Small States of the Commonwealth: Retrospect and Prospect*. Commonwealth Secretariat, London.
- Crossley, M., Sprague, T., 2012. Learning from small states for post-2015 educational and international development. *Curr. Issues Comp. Educ.* 15 (1), 26–40.
- Crossley, M., Sprague, T., 2014. Education for sustainable development: implications for small island developing states (SIDS). *Int. J. Educ. Dev.* 35, 86–95.
- Crossley, M., Sprague, T., 2017. Developing a bigger picture: re-theorising, applying, and extending the education in small states literature. In: Jules, T.D., Ressler, P. (Eds.), *Re-reading Education Policy and Practice in Small States: Issues of Size and Scale in the Emerging 'Intelligent Society and economy.'*. Peter Lang, Germany, pp. 37–51.
- Crowards, T., 2002. Defining the category of small states. *J. Int. Dev.* 14, 143–179.
- Dent, C., 2003. From inter-regionalism to trans-regionalism? Future challenges for ASEM. *Asia Eur. J.* 1 (2), 223–235.
- Education Quality and Assessment Programme, 2019. *EQAP Business Plan: 2019-2022*. EQAP, Pacific Community, ACER, New Zealand Foreign Affairs & Trade, Australian Aid. Available at: <https://eqap.spc.int/sites/default/files/resources/2021-06/EQAP%20Business%20Plan%202019-2022.pdf>.
- Edwards-Gums, O., 2014. Anguilla: the challenges of teacher recruitment. In: Thomas, E. (Ed.), *Education in the Commonwealth Caribbean and Netherlands Antilles*. Bloomsbury Academic, London.
- Farrugia, C.J., Attard, P.A., 1989. *The Multi-Functional Administrator: Educational Development in the Small States of the Commonwealth*. Commonwealth Secretariat, London.
- Girvan, N., 2010. Are Caribbean countries facing existential threats? *Latin Am. Movem.* Available at: <https://www.alainet.org/en/active/42028>.
- Grant, R., 1987. The politics of higher education within a regional context: the case of the Caribbean Community and Common Market (CARICOM). *J. Commonwealth Comp. Polit.* 25 (2), 180–195.
- Grant, R., 1993. Education in small jurisdictions, with reference to the Commonwealth Caribbean. *Compare* 23 (1), 25–35.
- Holmes, F., 1976. Development problems of small countries. In: Castle, L., Holmes, F. (Eds.), *Co-Operation and Development in the Asia Pacific Region: Relations between Large and Small Countries*. Japan Economic Research Centre, Tokyo, pp. 43–66.
- Holmes, K., Crossley, M., 2004. Whose knowledge, whose values? The contribution of local knowledge to education policy processes: a case study of research development initiatives in the small state of Saint Lucia. *Compare* 34 (2), 197–214.
- Jules, D., 1994. Adult education policy in micro-states: the case of the Caribbean. *Rev. Pol. Res.* 13 (3–4), 415–432.
- Jules, D., 2006. Power and educational development: small states and the labors of sisyphus. In: Afolayan, M., Browne, D., Jules, D. (Eds.), *Current Discourse on Education in Developing Nations: Essays in Honor of B. Robert Tanachnick and Robert Koehl*. Nova Science Publishers, New York, pp. 17–29.
- Jules, T.D., 2008. *Re/Thinking Harmonization in the Commonwealth Caribbean: Audiences, Actors, Interests, and Educational Policy Formation*. PhD Diss. Columbia University, New York: Teachers College.
- Jules, T.D., 2010. Beyond post-socialist conversions: functional cooperation and transregional regimes in the Global South. In: Silova, I. (Ed.), *Post-socialism Is Not Dead: Rereading the Global in Comparative Education*. Emerald Publishing, Bingley, pp. 401–426.
- Jules, T.D., 2015. Educational exceptionalism in small (and micro) states: cooperative educational transfer and TVET. *Res. Comp. Int. Educ.* 1–21.
- Jules, T.D., Ressler, P., 2016. What can be learned from small (and micro) states? Educational geostrategic leveraging' and the mechanisms of the fourth industrial revolution – the internet of things and disruptive innovation. *Island Stud. Indian Ocean* 3 (1), 27–35.
- Jules, T.D., Ressler, P., 2017a. Is 'small' always small and 'big' always big? In: Jules, T.D., Ressler, P. (Eds.), *Re-reading Education Policy and Practice in Small (and Micro) States in the Emerging 'Intelligent Society and economy.'*. Peter Lang, Germany, pp. 17–36.
- Jules, T.D., Ressler, P. (Eds.), 2017b. *Re-reading Education Policy and Practice in Small States: Issues of Size and Scale in the Emerging 'Intelligent Society and Economy.'*. Peter Lang, Germany.
- Jules, T.D., 2012. Re-reading the anamorphosis of educational fragility, vulnerability, and strength in small states. *Curr. Issues Comp. Educ.* 15 (1), 5–13.
- Jules, T.D., 2014. Trans-regional regimes and globalization in education: constructing the Neo-Caribbean citizen. In: Silova, I., Hobson, D. (Eds.), *Globalizing Minds: Rhetoric and Realities in International Schools*. Information Age Publishing, Chapel Hill, pp. 247–276.
- Keating, N., 2021. Small states in a big world. In: Hepburn, E., Keating, M., McEwen, N. (Eds.), *Scotland's New Choice Independence after Brexit U.K. Centre on Constitutional Change*. St. John's Land, pp. 223–230.
- Kuznets, S., 1960. Economic growth of small nations. In: Robinson, E.A.G. (Ed.), *Economic Consequences of the Size of Nations*. Macmillan, London, pp. 14–32.
- Lewis, V. (Ed.), 1976. *Size and Self-Determination and International Relations: The Caribbean*. Institute for Social and Economic Research, University of the West Indies, Kingston.
- Long, T., 2017. It's not the size, it's the relationship: from 'small states' to asymmetry. *Int. Polit.* 54 (2), 144–160.
- Louisy, P., 2001. Globalisation and comparative education: a Caribbean perspective. *Comp. Educ.* 37 (4), 425–438.
- Nkrumah-Young, K., Huisman, J., Powell, P., 2008. The impact of funding policies on higher education in Jamaica. *Comp. Educ.* 44 (2), 215–227.
- Payne, A., 1980. One University, many governments: regional integration, politics and the University of the West Indies. *Minerva* 18 (3), 474–498.
- Prince, R., 2012. Policy transfer, consultants and the geographies of governance. *Prog. Hum. Geogr.* 36, 188–203.
- Read, R., 2004. The implications of increasing globalization and regionalism for the economic growth of small island states. *World Dev.* 32 (2), 365–378.
- Rodhouse, J., 1987. Problems of policy and development. In: Bacchus, K., Brock, C. (Eds.), *The Challenge of Scale*. Commonwealth Secretariat, London, pp. 26–34.
- Salajan, F.D., Jules, T.D., 2019. Introduction: the educational intelligent economy, educational intelligence, and big data. In: Jules, T.D., Salajan, F.D. (Eds.), *The Educational Intelligent Economy: Big Data, Artificial Intelligence, Machine Learning and the Internet of Things in Education*. Emerald Publishing, Bingley, pp. 1–32.
- Straubhaar, R., 2017. Expanding the definition of small states: how Rio de Janeiro's favelas are a small state that 'talks big'. In: Jules, T.D., Ressler, P. (Eds.), *Re-reading Education Policy and Practice in Small States: Issues of Size and Scale in the Emerging 'Intelligent Society and Economy.'*. Peter Lang, Germany, pp. 56–70.
- Sutton, P., 2006. *Modernizing the State: Public Sector Reform in the Commonwealth Caribbean*. Ian Randle Publishers, Kingston.
- Thorhallsson, B., Wivel, A., 2006. Small states in the European Union: what do we know and what would we like to know? *Camb. Rev. Int. Aff.* 19 (4), 651–668.
- World Bank Group, 2016. *Approach Paper Clustered Country Program Evaluation for Small States*. Independent Evaluation Group.

The work of international organizations: ten theoretical approaches and a way forward

D. Brent Edwards, Jr.^a and Mauro C. Moschetti^b, ^aUniversity of Hawai'i at Mānoa; and ^bAutonomous University of Barcelona, Barcelona, Spain

© 2023 Elsevier Ltd. All rights reserved.

Theorizing the work of international organizations: ten approaches	466
World Culture Theory	466
International political economy	468
Cultural political economy and the strategic-relational approach	469
International relations and global governance	469
Historical and sociological institutionalism	470
Principal-agent theory	471
Institutional entrepreneurship	471
Policy streams and policy entrepreneurs	472
Organizational legitimacy	472
Science and technology studies	474
A way forward: circuits of power	474
Conclusions: strengths and possibilities	476
References	477

International organizations are central to the field of education, and particularly to the field of global education policy. Among other things, they are sources of policy and curricular ideas, provide leadership at various levels related to global goals and educational reform, engage in monitoring and evaluation of educational programs, offer financial and technical support for education projects, facilitate networks of education professionals, and partner with universities and schools to implement initiatives. Of course, international organizations can take many forms, and have many different missions. They can be multilateral in nature, as in the case of UN organizations, as well as bilateral, as in the case of international development agencies. They can also be a range of non-governmental or corporate organizations, such as think tanks, consulting and technology companies, philanthropies, or non-profit organizations with a social mission (e.g., Save the Children).

But regardless of the type of organization in question, it has been uncommon, or at least insufficiently common, for scholars in the field of education to bring theory to bear to explain the way these international organizations work. When we say “explain how they work,” we mean scholarship that theorizes what drives their activity, that is, theorizes what they do, their behavior, how they operate, and why. Because the work of international organizations focuses on policy in one way or another—whether making and implementing their own or contributing to the development and enactment of others’ policy—a discussion of international organizations’ nature and operation necessarily connects directly with this issue. However, it needs to be emphasized that we are not concerned here with explaining policy change, policy transfer, or policy borrowing at the country level or between international organizations. Rather, as noted, the discussion here is centered around characterizing theoretical approaches that shed light on the work of international organizations, though, of course, this work (of organizations) may have implications for policy movement more broadly. It should also be acknowledged that there are some theories that have been applied to both phenomena—that is, policy change within and beyond international organizations. In these cases, we will characterize the former (policy change within/inside the international organization) to the extent it is helpful to a discussion of understanding how the theory is applied to the latter (policy change beyond/outside international organizations).

Some scholarship does exist wherein theory has been applied, and a number of examples will be discussed in the sections that follow. Despite these examples, the point of departure for this entry is that there is a mismatch between the centrality and importance of international organizations, on one hand, and the extent to which theory is brought to bear in order to make clear how these entities are understood when it comes to their work, on the other. In other words, this entry takes as its focus those theories that attempt to explain international organizations because, it is argued, the idea of international organizations is routinely invoked but without making clear the theoretical assumptions or perspectives upon which authors are drawing when analyzing them. As one brief example, consider that the recent discussion of the “international architecture for education” by prominent scholars is characterized by a lack of engagement with theory that explains organizational behavior (e.g., Burnett, 2019; Heyneman and Lee, 2016; see also Benavot, 2011). This is a striking omission given that the entire purpose of this debate is, first, to diagnose what is wrong with how international organizations have carried out their work related to meeting international development goals and, second, to offer recommendations for how to improve their efforts going forward. However, while the omission is striking, is it not surprising, for it is a continuation of a larger trend where international organizations are discussed as actors involved in the politics

and provision of education but without a conceptualization of what it means to be—and what are the implications of being—an *organization*, nor any theorization of what guides their behavior.¹

Following from the above, the purpose of this entry is, first, to characterize a range of theoretical options for understanding the work of—and policy change within—international organizations, and, second, to argue for the benefits of a particular approach in that regard—known as the circuits of power model (Clegg, 1989). In addressing the first goal, the next section will characterize 10 approaches. The discussion of each approach will be relatively brief, given that this is an encyclopedia entry. The approaches over-viewed are intended to show the range of options available. They are indicative, not exhaustive. Moreover, they have been used by scholars of comparative and international education. The characterization of each approach is followed by examples of their application. As will be seen, examples focus on multilateral organizations such as the European Union; the Organization for Economic Cooperation and Development (OECD); the United Nations Education, Science, and Culture Organization (UNESCO); the United Nations Children's Fund (UNICEF), and the World Bank, since these are the organizations on which research has been conducted that provides insight into the issues of interest here. It is hoped that this entry can serve as a resource for those interested in international organizations, a resource that fills a gap in the existing literature by clarifying and bringing together in one place multiple options for theoretically approaching the work of international organizations.

In addressing the second purpose, this entry draws on organizational theorist Stewart Clegg, who developed the circuits of power approach in order to combine and to go beyond the core characteristics of many approaches to organizational change (Clegg, 1989). We see this approach as having certain benefits, particularly in the way that it is able to address the interaction of such things as structure, agency, and discourse. As will be more fully discussed, this framework, in explaining the work of organizations, is attentive to the interplay of agentic action, social relations of meaning, organizational structure, and exogenous environmental conditions. In focusing on this approach, the goal is to encourage more nuanced scholarship that overcomes historical divisions related, for example, to a focus on either political-economic structures, shared cultural values, discourse circulation, or the initiatives of individuals acting alone, among other tensions discussed in what follows. In other words, we see Clegg's (1989) approach as advantageous in that it brings together and extends the benefits of these other approaches discussed in the next section.

Theorizing the work of international organizations: ten approaches

Organizational theorists note that organizations are created in order to enable people to work together to accomplish tasks that would not be possible without coordinated activity. In their seminal paper, originally published in 1963, Blau and Scott (2003) write that “if the accomplishment of an objective requires collective effort, men [sic] set up an organization designed to coordinate the activities of many persona and to furnish incentives for other to join them for this purpose” (p. 175). Theoretically, however, there are divergent, though not necessarily mutually exclusive, perspectives on how to understand the behavior of these organizations, as discussed below. When it comes specifically to international organizations, four of the most prominent approaches, at least in the field of education, are World Culture Theory (WCT), international political economy (IPE), cultural political economy (CPE), and international relations/global governance. In addition to these, the next section discusses: historical and sociological institutionalism, principal-agent theory, institutional entrepreneurship, policy streams, organizational legitimacy, and science and technology studies (STS). These are then followed by a characterization of Clegg's approach in a separate section.

As the names of the approaches listed above imply, most of them are not rooted in organizational theory or do not emanate from the field of organizational studies. This is not seen as a weakness as the purpose of this entry is simply to map out approaches that have been used in the field of education to understand the work of international organizations. In accordance with this focus, we do not engage with deeper debates on how foundational sociological paradigms approach organizational analysis (for that, see, e.g., Burrell and Morgan, 1979; Clegg et al., 2006).²

World Culture Theory

In the field of education, WCT is commonly discussed in terms of how the spread of ideas has affected the ways that education systems are organized and in terms of what is taught. We will characterize WCT in relation to this area of application first and will then offer additional commentary on how organizations are theorized, before finally highlighting a few examples of the kinds of studies that look within international organizations.

¹Likewise, it is uncommon to theorize the way that international organizations engage with nation-states and other actors when trying to influence their policies. To this end, a recent systematic review of literature on the World Bank's influence on policymaking in education found that only 54% of the 77 studies analyzed clearly discussed their theoretical orientation (Edwards et al., 2021a).

²It should also be acknowledged that this discussion is related to but different from the discussion offered in a companion chapter entitled, “Globalization and education: Understanding and problematizing approaches to global governance and the global education policy field” (Edwards, 2021). The focus there was on characterizing diverse approaches for assessing the relationship between globalization and education in the field of global education policy. As such, that discussion, in addition to not having an exclusive focus on international organizations, was also centered around approaches that are broader in nature, in the sense that the kind of policy change of interest was not restricted to the context of international organizations. Thus, although three of the first four sub-sections below address approaches featured there, the discussion of these approaches here has been tailored to focus on their relevance and application to international organizations specifically.

Scholars writing within the WCT tradition see international organizations such as UN agencies and the World Bank as carriers of ideas that are freely taken up by nations around the world (Chabbot, 1998; McNeely, 1995; McNeely and Cha, 1994; Meyer et al., 1997; Ramirez et al., 2016). That is, from the perspective of World Culture Theory, through such avenues as international conferences and engagement with international organizations and their publications, policymakers at the national level are exposed to—and then adopt, as argued, often out of a desire to be seen as legitimate—prevailing ideas around what education systems should look like (Ramirez and Boli, 1987). This happens in the context of a “world society” composed of nation-states that have, especially in the post-WWII period, become increasingly dependent on seemingly rational, bureaucratic institutions, with the implication being that education is one among many sectors affected by this trend, a trend which, at foundation, is geared toward achieving “progress” and “modernization” through the spread and inculcation of liberal legal and cultural norms (e.g., primacy of the individual, consumption, faith in science, representative democracy, etc.) (Boli et al., 1985). When it comes to empirical studies of international organizations, World Culture scholars have tended to examine the extent to which educational policies, curricula, and textbooks reflect the ideas embedded in the discourse that is circulated in, through and by those international organizations (Chabbot, 1998; Kamens and McNeely, 2010). Differences between international ideas and globally circulated models, on one hand, and the forms they take at the national level, on the other, are attributed to “loose coupling” (Meyer and Rowan, 1977; Meyer et al., 1997).

The same assumptions of organizational imitation (i.e., “isomorphism”) for purposes of legitimacy are applied to international organizations as well. Writing on organizations generally—and not on national education systems specifically—Meyer and Rowan (1977), in their famous article, wrote that

organizations are driven to incorporate the practices and procedures defined by prevailing rationalized concepts³ of organizational work and institutionalized in society. Organizations that do so increase their legitimacy and their survival prospects, independent of the immediate efficacy of the acquired practices and procedures (p. 340).

This perspective on organizational behavior contrasts with the (functionalist) assumption that organizations “function according to their formal blueprints: coordination is routine, rules and procedures are followed, and actual activities conform to the prescriptions of formal structure. But much of the empirical research on organizations casts doubt on this assumption” (Meyer and Rowan, p. 342). In practice, and in contrast to the depictions of organizational charts, WCT stresses that “formal organizations are often loosely coupled,” meaning that their

structural elements are only loosely linked to each other and to activities, rules are often violated, decisions are often un-implemented, or if implemented have uncertain consequences, technologies are of problematic efficiency, and evaluation and inspection systems are subverted or rendered so vague as to provide little coordination.

Meyer and Rowan (1977, p. 343).

Whether or not organizations actually follow prevailing norms and expectations, they need to be perceived as doing so, for the sake of legitimacy, as noted.

More recently, Bromley and Powell (2012) have written on a related phenomenon, means-end decoupling. Even where internal systems, processes, and actions are tightly linked together, it may be the case that the activities of an organization (i.e., means) are misaligned with their stated goals (i.e., ends). It is suggested that this form of decoupling is growing in relevance in “an increasingly managerial world that emphasizes evaluation, standardization, and benchmarking, [whereas] the policy–practice form of decoupling may become less common” (Bromley and Powell, 2012, p. 3).

When it comes to research on international organizations, the WCT approach has been adapted through what Milana (2015) calls a global polity perspective in order to study, for example, the processes through which UNESCO “mobilizes” different education policies (p. 203). Milana (2015) highlights

three processes that cross geopolitical borders and professional interests: “landmarking”, by which a shared sense of a common past is created; “brokerage”, which helps shape a common future direction; and “framing”, which is used to convert ideational landscapes into material government-led actions (p. 203).

In these ways, UNESCO would be a source of “rationalized myths,” that is, narratives and expectations for education systems more broadly: “its work tends toward the homogenization of beliefs, guiding norms, values and ideas” (Milana, 2015, p. 210).

³Rationalized myths are accepted narratives “depicting various formal structures as rational means to the attainment of desirable ends” (Meyer and Rowan, 1977, p. 345).

Other studies have looked at decoupling within international organizations. Zapp (2019), for example, reviewed documents from 16 different international organizations during 2000–2015, revealing the “awkward entrenchment in the inconsistent goals of world culture: the formal celebration of cultural diversity, on the one hand, and the delegitimizing of such culture wherever it conflicts with the goals of equality and universalism, on the other hand” (p. 322). Zapp (2019) ultimately argues that international organizations are decoupled in that “their ceremonial commitments to culture and contextual adaptation of reforms are disconnected from their actual practices of promoting progressive cultural ideology and universalistic planning based on the premises of the social sciences” (p. 322). Elsewhere, Mundy and Menashy (2014) have drawn on the work of Bromley and Powell (2012) around decoupling to highlight “organizational hypocrisy” within the World Bank, where it is common to promote private schooling in rhetoric but uncommon to support related reforms through the day-to-day operations of the organization.

As a final point here, it is ironic that the studies that focus on the role of international organizations in creating or spreading norms that supposedly become or are globally-accepted, those same studies implicitly evince the importance of agency and material resources that a focus on cultural and rationalized myths historically sought to avoid. This is a tension to which the circuits of power approach responds, as discussed later.

International political economy

WCT has been critiqued for its lack of attention to the way that material power shapes organizational behavior (Carney et al., 2012; Downey et al., 2020). The contrasting perspective to WCT—known as IPE—takes this kind of power as its focus. However, just as with WCT, the use of IPE in the field of education has tended not to look *within* organizations. Rather, it has been more common to think about how organizational activity is connected to, or a consequence of, the larger political-economic structures within which it is embedded. As with the previous discussion on WCT, we characterize this broader, and more common, way of approaching international organizations from the perspective of IPE, before then drawing out the implications of this perspective for the internal work of such organizations.

For scholars operating within an IPE paradigm, the focus is on structure, power, and the nation-state as a site of negotiation among competing political and (capitalist) economic interests. Dale (2000), for example, in writing about what he labels the “globally structured approach to education” (GSAE), sees the role of the World Bank, through its loans and the attached conditions, as central to the settlement of national-level decisions related to education policy. When it comes to education, the important thing to note from the framework of GSAE, is that it (i.e., education) is at the heart of the process by which the government legitimates and generates consent to capitalist modes of production (e.g., through promises of education-based meritocracy and through the provision of the skills necessary to acquire jobs that will provide a decent quality of life). The role of the World Bank (and other institutions with similar ideology) is thus to promote (and to lobby for the acceptance of) policies that hold out the promise of improved education and certain outcomes (like enhanced productivity, increased employment, national competitiveness, etc.) without challenging the capitalist system of which the education sector is a part (Bonai, 2002). Scholars grounded in the general paradigm of IPE tend to examine the power-dynamics, processes, and structures that characterize the interaction of the World Bank and its government counterparts, as well as the pathways through which the World Bank is able to exert its influence (see, e.g., Edwards et al., 2021a; Moutsios, 2009, 2010; Samoff, 2004, 2009).

When shifting to look within international organizations from an IPE perspective, emphasis is still placed on understanding how larger-political economic structures affect behavior (Broad, 2006). Studies here question, for example, the extent to which international organizations should be seen as a “vector of the interests of their member states” (Wade, 1996, p. 5). For example, insightful studies have spoken to the extent to which World Bank policies, publications, knowledge production are shaped by whether or not they align with the accepted perspectives and paradigms for discussing international development, that is, the extent to which they align with the preferences of powerful actors (such as the US Treasury) who benefit from the reproduction of dominant approaches to development (Broad, 2006; Downey et al., 2020; Wade, 1996).

In the field of education, IPE’s preoccupation with political-economic structures has similarly led to studies that focus on how international organizations are able to use the material resources available to them to promote their ideologies (Edwards, 2018b; Klees, 2012; Mundy, 1998; Steiner-Khamsi, 2006). The World Bank has again been the subject of study from this perspective, perhaps unsurprisingly, given its prominent position in the field of global education policy and the fact that it is the largest international development organization—both in terms of its staff and in terms of the financial-technical resources available to it (Gilbert and Vines, 2006). Studies in this area have highlighted the way that the World Bank has been able to use its resources not only to produce research and other publications on an incredible scale (Klees and Edwards, 2014; Samoff, 1993) but also how it has been able to use its resources to leverage those studies globally (Edwards, 2018a; Edwards and Urquilla Loucel, 2016; Edwards et al., 2020) and to structure processes for participation by other international organizations (Bermingham, 2011; Verger et al., 2014).⁴ Importantly, this focus on how the production of knowledge and discourse affects subjective perceptions has been (more explicitly) taken up by a variation of IPE, discussed next.

⁴Interestingly, IPE’s focus on material resources overlaps here with Foucault’s emphasis on how the production of discourse leads to acceptance of certain ways of seeing the world (Clegg et al., 2006; Foucault, 1980).

Cultural political economy and the strategic-relational approach

In the past 20 years or so, scholars of political economy have been attentive to the absence of attention to the cultural dimension, that is, to how meaning is made and how commonly-shared meanings affect behavior. Increased attention to this area has led to a cultural turn in this field under the label of CPE. Jones (2008) writes that CPE is an extension of the Marxist tradition that explores “relations between the cultural (the world of *discourse and practice*), the political (actors and institutions) and the economic” (p. 1, emphasis added). Some scholars advancing CPE focus primarily on the production of discourse and meaning (semiotics)—within larger political-economic structures—and the way that actors interpret that discourse (Jessop, 2004). To that end, CPE introduces the issue of strategic selectivity, or what Jessop (2001a) labels the strategic-relational approach.

Per Jessop (2001a), individual policy actors, informed by discursive trends, participate in reform processes, in spaces where decisions are made, and in agenda-setting while engaging in structurally-oriented strategic calculations in that they analyze and take into account their institutional and political constraints and opportunities and subsequently “orient their strategies and tactics in light of their understanding of the current conjuncture and their ‘feel for the game’” (Jessop, 2001a, p. 1224). The point, then, is that, in extending the focus on ideas mentioned in relation to IPE (Hay, 2002), CPE makes an attempt to be less deterministic than IPE by emphasizing the role of individual agency, though that agency is still circumscribed by political-economic structures, by the vested interests of organizations, and the discursive context in which they are located (Jessop, 2010). The goal is to go beyond an analysis of policies as subject to political and economic constraints (already addressed by the IPE approach) to address as well the ways that policies, in practice, are brought to life in cultural contexts (including within organizations) that are imbued with their own “operational logics or instrumental rationalities” (Jessop, 2001b, p. 290). As will be further discussed later, these assumptions are also shared by the circuits of power approach (Clegg, 1989).

The CPE approach has been applied to the organizational context of the European Union. Through this approach, Jones’ (2010) study, for example, sought to accomplish the following:

to account for and explain the questions of why and how particular kinds of cultural, political, and economic processes develop at the [European Union] scale, the role that actors, institutions, and their mobilization and contestation of discourse play in these processes, and the extent to which the cultural, the political, and the economic are mutually constitutive in conditioning what happens within multilevel governance. In taking seriously the importance of actors and institutions, their discourses and practices, [cultural political economy] would clearly seek to place them in relation to the structured complexity of bigger economic, political, and social processes within the [European Union] (p. 34).

As can be seen from these goals, the CPE approach is demanding in analytic terms. However, where successful, its application can lead to nuanced insights. In the case of Jones (2010), CPE’s value was in that it offered:

a way to think about the extent to which: (1) the EU is becoming a site at which the governance of education is achieved; (2) the political context of the activities, including the discursive, which are pursued at this level; and (3) the importance of both the discourses and the realities of the economic context within which education reform can be pursued at the EU level (p. 38).

By shedding light on these dimensions, CPE manages to capture the whole policy translation process, including the emergence of different policy options (variation), which options are privileged (selection), and which are ultimately taken up for implementation (retention; Jessop, 2010).

International relations and global governance

A fourth group of scholars approaches international organizations through the lens of international relations and global governance (Barnet and Finnemore, 2005; Jakobi, 2009; Jones and Coleman, 2005; Jones, 2007a,b; Mundy, 2007). While there are multiple camps within international relations theory, the most basic two are idealist and realist in nature, with the former focusing on (or assuming) the inclination of nation-states to work together through organs of global governance (e.g., the United Nations) to make progress on liberal priorities (e.g., democracy, safety, health, universal education, etc.), while the latter has presumed independence and inherent conflict among states and, within the realm of global relations, “the pursuit of national interests in contexts of very uneven distributions,” including through multilateral organizations such as the World Bank (Jones and Coleman, 2005, p. 9). A third international relations perspective is neo-Marxist in nature and draws on world systems theory (Wallerstein, 1984). These scholars depart from a preoccupation with the capitalist world system and the way that nation-states occupy either core or peripheral positions within that system. From their privileged positions in this system, core nations use international organizations to their benefit and to extend the capitalist economic system (Mundy, 2007).⁵

However, since the 1990s, following the end of the Cold War, there has been a shift in the discussion by international relations scholars toward global governance (Mundy, 2007). The point is that, independent of nation-states, there is “an evolving set of processes and interactions ... that by definition involves heterogeneous private and public actors at multiple levels or scales of

⁵As can be seen, the world systems approach has much in common with—and indeed can be considered a form of—IPE analysis.

action: local, national, international, and transnational" (Mundy, 2007, p. 343). Within this space—which has been conceived as a global policy field in education (Lingard et al., 2005)—“multilateral organizations are seen as world actors in their own right, behaving as distinct components of global power relations and not merely as functional extensions of the systems that gave rise to them” (Jones and Coleman, 2005, p. 4; see also Barnett and Finnemore, 2005). A seminal study by Jones and Coleman (2005) has applied this perspective to explain the behavior of UNESCO, UNICEF, the United Nations Development Program, and the World Bank. Such research focuses on how “forms of international authority [grounded in norms and rational, bureaucratic organization] are socially constructed and historically contingent, rather than materially or historically fixed” (Mundy, 2007, p. 351). A number of studies beyond Jones and Coleman (2005) have also looked at the World Bank in order to explain its inner workings over time as well as changes to its position vis-à-vis other multilateral organizations (Jones, 1992, 2007a,b; Mundy, 2002).

Some recent scholarship that falls within the global governance umbrella shows that there is a point of connection between this approach, IPE, and CPE. For example, in drawing on critical constructivism (with its assumption that portrayals of reality reinforce actual power structures) and CPE (which is interested in how economic and political imaginaries are translated and institutionalized in structures), Verger (2014) emphasizes that such actors as the World Bank are key in global governance because of their ability to project visions of the world, to define problems that need to be solved, and then to mobilize ideas around which models are appropriate for solving those problems. This approach has been utilized to explain how the World Bank has framed and sold the idea of public-private partnerships in education (Verger, 2012; see also Samoff, 2012). We will discuss further below how the global governance perspective also connects with theorizing related to institutional entrepreneurs.

Historical and sociological institutionalism

Scholars of international relations, political science, and sociology have advanced different forms of institutionalism (Hall and Taylor, 1996). This section focuses on two variants: historical and sociological institutionalism. The first—historical institutionalism—is characterized by “attention to large questions with an explicit temporal scope that concern the creation, reproduction, development, and structure of institutions over time” (Fioretos, 2011, p. 372). Here, institutions are understood as

the formal or informal procedures, routines, norms and conventions embedded in the organizational structure of the polity or political economy. They can range from the rules of a constitutional order or the standard operating procedures of a bureaucracy to the conventions governing trade union behaviour or bank-firm relations.

Hall and Taylor (1996, p. 938).

More specifically, while taking into account pressures from the organizational context (e.g., developments in the global education policy field, such as the Sustainable Development Goals), historical institutionalism investigates how “periods of continuity [are] punctuated by ‘critical junctures’, i.e., moments when substantial institutional change takes place thereby creating a ‘branching point’ from which historical development moves onto a new path” (Hall and Taylor, 1996, p. 942). Finally, historical institutionalism is “attentive to the relationship between institutions and ideas or beliefs,” meaning the way that organizational arrangements reinforce certain ideas (e.g., in the case of the OECD, the way that its organizational arrangement has come to reinforce the Program on International Student Achievement, PISA) (Hall and Taylor, 1996, p. 942). As an additional example, historical institutionalism could be used to study how and why the OECD and the World Bank have changed their practices, rules, and even structures to work together in the 2010s to adapt PISA for low- and middle-income countries.

Studies that draw on historical institutionalism engage case study methods (interviews and archives) to produce “theory-guided narratives” (Capoccia and Keleman, 2007). These narratives center, first, on explaining critical junctures and “the main actors, their goals, preferences, decisions, and the events that directly influenced them” (Capoccia and Keleman, 2007, p. 357). Second, these narratives “reconstruct the consequences of the decisions that were taken” (p. 357). Importantly, these narratives explain how, during periods of “institutional fluidity,” influential actors “steer outcomes toward a new equilibrium ... in order to enhance their own position” (p. 354). As these quotes suggest, and in accordance with the strategic-relational approach discussed previously, historical institutionalism assumes that actors (i.e., individuals and groups within organizations) are knowledgeable, agentic, and reflexive, and that they “routinely (often intuitively) monitor the consequences of their action” (Hay and Marsh, 1999, p. 956). Moreover, in so doing, they “assess both the immediate and unfolding impact of their prior strategies in relation to earlier intentions and anticipated outcomes in the light of strategic assessments of the conduct of others” (p. 956).

In the field of global education policy, Menashy and Manion (2016) have combined historical institutionalism with sociological institutionalism. In their approach, sociological institutionalism is defined as “an examination of the dynamics within organizations by studying bureaucratic rules and associated behaviors” (Menashy and Manion, 2016, p. 321). More specifically, sociological institutionalism:

Draws ... attention to the cultural norms within organizations that shape how the bureaucracy operates and responds to external pressures (Barkin, 2006; Schmidt, 2008). Under this approach, policy-making processes are argued to be governed primarily by formal rules and procedures, shaping which ideas are considered and selected and thereby what actions are taken by an organization (Campbell, 1998, 2002).

Menashy and Manion (2016, p. 321).

Interestingly, the approach taken by [Menashy and Manion \(2016\)](#) not only demonstrates the plurality of approaches to institutionalism, but also, in this case, shows that sociological institutionalism can be taken to have contrary meaning by different authors. Whereas [Menashy and Manion \(2016\)](#) assume that behavior in organizations is driven by formal rules and procedures, other authors have used sociological institutionalism as a synonym for WCT ([Hall and Taylor, 1996](#)), which, as already explained, sees organizational behavior as driven by rationalized myths and as characterized by a policy-practice disconnect. In any event, [Menashy and Manion \(2016\)](#) use their approach to historical and sociological institutionalism in order to inform their examination of how “the organizational structures of UNESCO and UNICEF influenced the ways in which each organization interprets their educational mandates” (p. 320).

Principal-agent theory

The principal-agent approach draws on field of economics to examine the relationship between international organizations (agents) and the states to which they are accountable (principals). This area of scholarship asks such questions as: why do states delegate certain responsibilities to international organizations and, once they do, how do states attempt to control them? Within this relationship, international organizations are conceptualized as semi-independent actors with significant leeway in terms of how they do what they do, that is, how they “implement policy decisions and pursue their own interests strategically” ([Hawkins et al., 2006](#), p. 5). Just as with the global governance perspective discussed above, scholars concerned with this dimension of the principal-agent approach ask questions about how international organizations take advantage of the room to maneuver that is afforded to them by their principals—and how principals respond in order exert control.

To be specific, principals are defined as actors that have the ability “both to grant authority and rescind it” ([Hawkins et al., 2006](#), p. 7), while agents are those who “receive conditional grants of authority from a principal” (p. 8). The semi-independent nature of international organizations is a function of discretion, that is, the fact that agents are created to accomplish certain goals (the goals of the principal[s]) but they have latitude in the actions taken to work toward those goals. In this respect, the principal-agent approach connects with WCT theory, in that discretion can lead to decoupling.

Despite the focus on discretion, agency tends to be conceptualized at the level of the organization, rather than at the level of the individual operating within them. In other words, studies tend to discuss entire organizations (e.g., UNICEF, the World Bank, etc.) as the agents who principals seek to regulate (see, e.g., chapters in [Hawkins et al., 2006](#)). This focus is distinct from that of the strategic-relational approach, discussed above, or that of institutional entrepreneurs, discussed in the next section.

The principal-agent approach has not been common in research in the field of education. Scholars in this field have more commonly sought to understand the politics of how policies get made within international organizations (e.g., [Verger et al., 2014](#)). Where a principal-agent lens has been brought to bear, it has been extended, as in the case of [Mundy and Verger \(2015\)](#), who, first, take into account not only the relationship between the World Bank and the states that serve as its formal principals, but also the relationship between this organization and the civil society organizations who put pressure on it and to whom the World Bank must respond in order to manage perceptions of legitimacy. Second, they look within the World Bank to see how its leadership and staff use the organizational slack available to them to act in autonomous ways, thus going beyond the unitary focus mentioned above. Third, these authors are attentive to the ways that the World Bank is responsive to the exigencies and expectations of borrowing countries, who must be willing to borrow if the World Bank is to survive.

Institutional entrepreneurship

Beyond [Jessop's \(2001a\)](#) strategic-relational approach, discussed above, there are two additional, and closely related, approaches that the focus on the ability of individuals within organizations to act as change agents. One is institutional entrepreneurship; the other is the idea of policy entrepreneurs, associated with the policy streams perspective.

When it comes to individual agency, the first perspective aligns with historical institutionalism ([Hay and Marsh, 1999](#)) and reacts to sociological institutionalism ([Menashy and Manion, 2016](#)) and WCT ([Meyer and Rowan, 1977](#)). This is so in that institutional entrepreneurship zeroes in on the “activities of actors who have an interest in particular institutional arrangements and who leverage resources to create new institutions or to transform existing ones” ([Maguire et al., 2004](#), p. 657). The project of institutional entrepreneurship has been to reintroduce “agency, interests and power into institutional analyses of organizations” ([Garud et al., 2007](#), p. 957). Here, and as with historical institutionalism, the focus on “institutional analyses” in the previous sentence is meant to draw attention to the ways that actors, through their entrepreneurship, contribute to “deviations from some norm” within organizations ([Garud et al., 2007](#), p. 960). More fully, “to qualify as institutional entrepreneurs, individuals must break with existing rules and practices associated with the dominant institutional logic(s) and institutionalize the alternative rules, practices or logics that they are championing” ([Garud et al., 2007](#), p. 962).

Interest in understanding how institutional entrepreneurs can succeed in institutionalizing changes has brought attention not only to the material resources available to individuals. It has also generated a focus on the “linguistic and symbolic aspects of power where the focus is on the meanings that humans attribute to a situation which, in turn, influences how they act in relation to it” ([Garud et al., 2007](#), p. 962). In this way, institutional entrepreneurship connects with some of the core tenets of CPE and certain strands of IPE and global governance scholarship. As one example of scholarship related to this area, [Verger \(2012\)](#) examines the strategies of entrepreneurs in the context of different international organizations as they frame and promote the idea of public-private partnerships as beneficial for education sectors in low- and middle-income countries. Ultimately, this research “reveals

the complex way in which programmatic ideas, political actors, institutions, and material factors interact to strategically put forward policy alternatives in developing countries" (Verger, 2012, p. 110). An additional example can be found in the work of Auld and Morris (2016), who likewise inquire into the strategies of entrepreneurs from different international organizations, albeit with a focus on how they use data from international achievement tests to "construct and promote knowledge of transferable 'best practices'" (p. 202).⁶

Policy streams and policy entrepreneurs

Notably, both of the examples discussed just above focus not on change within international organizations but rather how actors located in them advocate for policy change in other entities, such as ministries of education.⁷ This is distinct from the work of Sumida (2017), who draws on Kingdon's (1984) policy streams approach to conceptualize agenda-setting processes within international organizations. Here, the term policy entrepreneur, as opposed to institutional entrepreneur, is employed. While the notion of the policy entrepreneur mirrors that of the institutional entrepreneur, the former is conceptualized as situated within an explicitly political process within which different organizations participate and compete as they advocate for different policy outcomes. This is different from assuming, as is often the case, that institutional entrepreneurs operate in the context of a given organization as they dedicate themselves to effecting change there.⁸ However, as Sumida (2017) has shown, the policy streams approach of Kingdon can be adapted to study the evolution of policy change within an organization. But what, then, is the policy streams approach, and how can it be applied to an organizational context?

In short, Kingdon's (1984) approach has three different "streams," which, when they come together, open a policy window. The three streams are the problem stream, the politics stream, and the policy stream. The politics stream refers to the political context wherein different political parties, governmental representatives, and other organizations vie for position and influence. The problem stream, which can be affected by the politics stream, refers to which issues are defined as needing to be addressed—usually through governmental action, since Kingdon's (1984) approach was developed with the US government in mind. The policy stream, then, refers to the many ideas, proposals, programs, and solutions promoted by a variety of actors, including researchers, academics, and interest groups. A policy window is said to open temporarily when these three streams align. However, in order for a given policy response to be taken up and to move forward as *the* solution, it typically must be championed by one or multiple policy entrepreneurs, with these individuals defined as those who are "willing to spend their resources, such as time, energy, reputation, and money, to promote their position [(i.e., their solution)] in return for future gain" (Sumida, 2017, p. 383).

Although, as indicated, Kingdon's (1984) approach has tended to be employed in analyzing domestic politics, Sumida (2017) adapted it to the context of global governance. More specifically, she applied it to understand how the initiative known as the Decade of Education for Sustainable Development came to be taken up as a formal policy initiative of the United Nations during 2005–2014. To do so, she examined how certain policy entrepreneurs were not only able to sell this initiative to decisionmakers within the United Nations, but also how they were able to take advantage of the predictable policy windows that periodically open within this organization, and, in so doing, how they positioned this initiative as a response to the problems of environmental degradation and unsustainable development.

Organizational legitimacy

Literature on organizational legitimacy has tended to emanate from the field of business and examines, for example, how companies respond to crises such as oil spills or factory fires (see, e.g., Desai, 2011; Massey, 2001). Although international relations scholars have begun to discuss legitimacy in the context of international organizations, the literature has focused on legitimacy as a function of the democratic management of these organizations (see, e.g., Buchanan and Keohane, 2006; Zürn, 2004).⁹ Put differently, international relations scholars have looked for the source of legitimacy within such organizations. However, given the socio-political nature of the field of global education policy (Lingard et al., 2015), wherein international organizations collaborate and compete for influence, it is important to look outside of organizations, to the ways that they interact and perceive one another.

Drawing on the theoretical literature on organizational legitimacy, a framework for understanding organizational legitimacy in this field has been developed by Edwards et al. (2018) by drawing on literature from the fields of organizational management, sociology, and communications (e.g., Bitektine, 2011; Dowling and Pfeffer, 1975; Massey, 2001). Although scholarship on multilateral institutions has suggested that legitimacy results from a "generalized perception ... that the actions of an entity are desirable or

⁶Note that the literature on education has not tended to engage with the concept of institutional entrepreneurs, at least not when it comes to the global education policy field. This characterization is also true of Verger (2012) and Auld and Morris (2016), who both use the term policy entrepreneur. Despite this difference in terminology, they are included here as examples of institutional entrepreneurship precisely because of the lack of examples in the field of global education policy, and because of the conceptual similarity in the notions of institutional and policy entrepreneurs. The difference is in their area of application—inside versus across organizations—not in the substance of what they are thought to do.

⁷Attention to the examples just above still fits with the focus set out at the beginning of this entry because, in these cases, the work of international actors was to change others' policies.

⁸This way of approaching institutional entrepreneurs is more common outside of education.

⁹In this sense, UNESCO is among the most legitimate multilateral organizations, due to its democratic structure, where each member country has one vote.

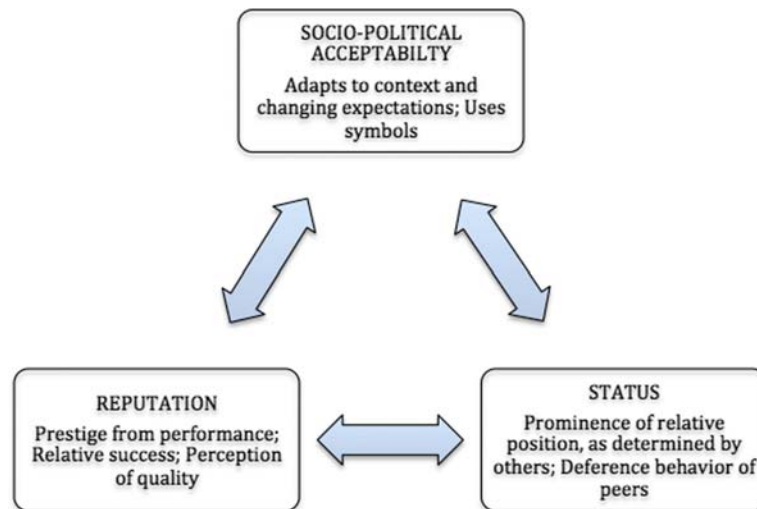


Fig. 1 Components of organizational legitimacy. Source: Edwards et al. (2018).

appropriate" (Suchman, 1995, p. 574), the framework put forth goes beyond the notion of "generalized perception" and to delineate a conception of legitimacy based on three interrelated components—socio-political acceptability, reputation, and status. These components are depicted in Fig. 1. As can be seen, an organization can dedicate itself to achieving socio-political acceptability by (a) adapting to its context, (b) responding to changing expectations, and (c) invoking or affiliating itself with symbols (or other organizations) that possess legitimacy. The other two components—reputation and status—are beyond the organization's direct control.¹⁰

Reputation has a specific meaning. It is defined as (a) the "prestige accorded" to the organization "on the basis of how they have performed particular activities in the past" (Jensen and Roy, 2008, p. 497); (b) the organization's "relative success in fulfilling the expectations of multiple stakeholders" (Fombrun and Shanley, 1990, p. 235); and (c) the perception of the quality of the work of organization (Rhee and Haunschild, 2006, p. 101). Reputation is therefore the counterpart to an organization's striving. That is, while an organization may strive for legitimacy, changes to reputation depend on how those efforts are perceived by others.

Socio-political acceptability and reputation are then linked to status. Though there are no guarantees, an organization can dedicate itself to improving the perception of its performance, fulfilling the expectations of its stakeholders, and ensuring that its work is perceived as high quality—in hopes of raising its status among peer institutions, where status is defined as the "prominence of an actor's relative position within a population of actors" (Wejnert, 2002, p. 304), and where increased prominence is associated with increased "deference behavior" (Huberman et al., 2004, p. 103). Of course, as status is based on perception, it is socially constructed and depends on agreement among the relevant actors in a social system (Washington and Zajac, 2005). Thus, while it may be sensible for organizations to strive to improve their status, it should be noted that "legitimacy is not something that can be claimed by organizations, but is instead something that is given by stakeholders" (Massey, 2001, p. 156).

Something that should not be lost here is the relative nature of reputation and status. The former is based on "relative success" while the latter is connected with "relative position." The implication is that actors in the global education policy field do not only judge an organization against the expectations which have been placed on it (socio-political acceptability), but also judge the performance of this organization in comparison to its peers (reputation), with that valuation affecting its overall position in the field (status).

Edwards et al. (2018) apply this framework to UNESCO. Having faced a crisis of legitimacy for decades, UNESCO sought, starting in the 1990s and early 2000s, to use both its coordination of the Education For All initiative and the production of its Global Education Monitoring Report to raise its status in the field of global education policy. The study by Edwards et al. (2018) provides insights into the successes and challenges faced by UNESCO in this effort. In using this framework, the authors were able to show how, due to UNESCO's efforts, this organization was able to raise its reputation by responding to the changing socio-political context, though it still could not elevate its status to match other organizations, such as the World Bank, which continued to pursue its own initiatives that helped to maintain its dominant position in the global education policy field.

¹⁰While they are sub-components of legitimacy in this framework, status and reputation have their own literatures. See Washington and Zajac (2005) for their excellent review of the theory on status and Fombrun and Stanley (1990) for their treatment of the literature on reputation. Bitektine (2011) provides a very thorough review of different definitions of legitimacy as well as the relationships among legitimacy, status, and reputation.

Science and technology studies

Research from the STS approach does not conceptualize directly the work of organizations. It is an area of inquiry that looks at the “sociology of knowledge processes” from a particular perspective (Gorur et al., 2019, p. 7). In drawing on Camic et al. (2011, p. 6), Gorur et al. (2019) emphasize that STS is concerned with “the day-to-day actions and processes through which the producers of social knowledge actually go about the on-the-ground work of making, evaluating, and disseminating the kinds of social knowledge that they are involved in producing” (p. 7). Of course, these processes do tend to play out within organizational contexts. However, rather than apply a ready-made theory of organizational behavior that separates structure from agency, the STS tradition looks at how assemblages of human and non-human elements are formed as they come into contact with one another (Latour, 2005). The idea is that these assemblages function as temporary structures that are co-constituted by the operation and intersection of human and non-human agency. In this way, STS encourages an approach to power that is “both relational and dynamic, rather than structural and monolithic” (Gorur et al., 2019, p. 4).

The avoidance of a priori theory means that STS depart from the perspectives of those who are embedded in the contexts and processes—or assemblages—of interest. One must explain phenomena from their perspective, rather than imposing conceptual categories or frameworks. In this respect, one can see how STS connects with approaches to global governance that are rooted in constructivism. A key difference, as noted above, is the focus on non-human agency. The core insight here is that “politics and power relations are distributed through ‘things’—materials and technologies, numbers and standards—things that often are underestimated or overlooked as agents in educational research” (Gorur et al., 2019, p. 4; Latour, 2005). The value of such an approach is intuitive when one considers that humans are constantly interacting with and, indeed, reacting to material and discursive elements as they go about their business (Koyama and Varenne, 2012).

Importantly, one goal of research thus oriented is to understand how temporary constellations of actors and objects allow for common understanding to develop and take hold among actors in a network. Drawing again on Gorur et al. (2019), STS scholars STS research, then, attempts to shed light on how categories for understanding come to be.

explore the mundane and everyday practices through which, eventually, ideas and fact-making are stabilized and systems are established. They attempt to open up the black boxes of the taken-for-granted and of common sense to reclaim the tentativeness and contingency by tracing their routes to naturalization.

(p. 4; see also Savage, 2019).

In the field of global education policy, the STS approach has historically received scant attention, though this is changing, as the special issue edited by Gorur et al. (2019) indicates. That said, there are still few studies that look explicitly at—or that connect with—the work of international organizations. Among the few exceptions, perhaps the most relevant to mention is the study by Addey and Gorur (2020), entitled “translating PISA, translating the world.” These authors drawn on the STS approach to argue that the adaptation of the OECD’s PISA test to low- and middle-income country contexts was not only a technical exercise but was also a political one, with both being necessary in order for country participation to make sense, and in order for the PISA results of these countries to be interpretable, that is, to have agreed upon meanings shared by those who come into contact with them.

A way forward: circuits of power

Each of the previously discussed perspectives brings a distinct approach to the issue of organizational work. While each perspective has its strength, we argue here that the approach of Clegg (1989), elaborated in his book, *Frameworks of power*, brings together and extends the benefits of these other approaches. However, before discussing further the ways that Clegg’s (1989) approach can be seen as an improvement, we first turn to explaining his approach.

In essence, there are two dimensions that work together. One is labeled the “sociology of translation,” while the other is “circuits of power.” The former is temporal in nature and has to do with the phases of activity that characterize (policy) change within an organization. The latter—circuits of power—must be taken into account during each phase, for its components affect the ability of change to take place over time. The combination of these two approaches enables one to engage in an analysis of policy change—or, really, any kind of substantive change that is made in terms of how things are done—within an organizational setting that considers such factors as individual and non-human agency, social relations, and organizational structures, as well as factors exogenous to the organization, among others.

With the above in mind, we note that Clegg (1989) understands organizations as a “locus of decision and action” (p. 197). The question becomes, within this understanding of organizations, how international organizations—but particularly those engaged in policy work—come to support certain reforms and not others. In other words, the analytic puzzle has to do with explaining the decisions and actions that lead to organizational change.

According to the approach of the “sociology of translation,” organizational change happens through four moments. In the first—problematisation—agents of an organization posit “the indispensability of their solutions for (their definition of) the others’ problems” (Clegg 1989, p. 204). For the second—*intéressement*—agents engage in “the process of ‘interesting’ or ‘enrolling’ another agent to one’s own agency” (p. 205). It is important to highlight that, here, the agent seeks to “lock in” what is meant or, we suggest,

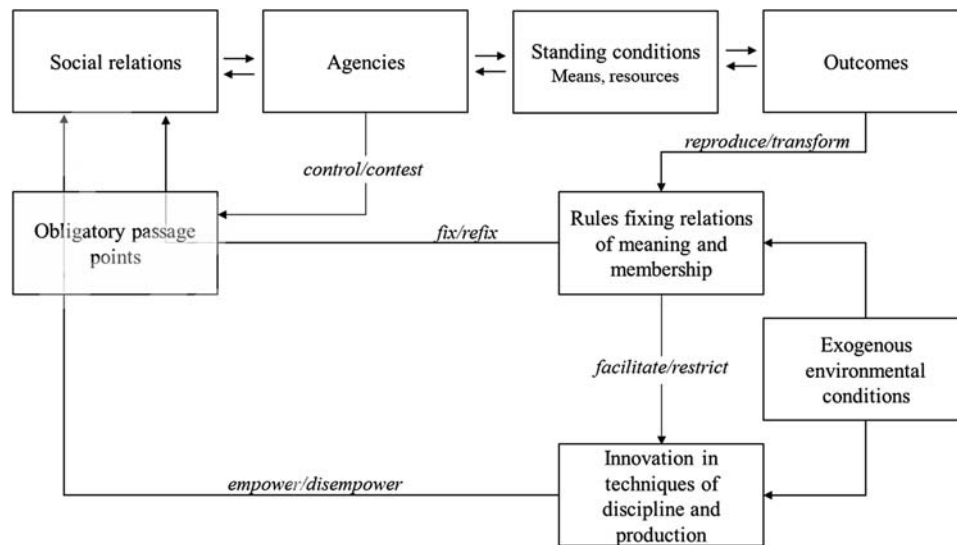


Fig. 2 Organizational Circuits of power. Source: Edwards and Moschetti (2021), adapted from Clegg et al. (2006).

envisioned by the policy being advanced. The third moment—enrollment—then endeavors to transform the interest of other agents into “alliances and coalitions” (p. 205) around shared meanings. Conceptually, moments two and three are distinct, though they may difficult to separate empirically, since a given action can seek both to generate interest in a policy and to create a coalition—that may be more or less formal—to support it. The last moment is mobilization; it refers to “the set of methods that agencies [(i.e., actions, not ‘international agencies’)] use to ensure that the representations of interest ... are in fact themselves fixed” (p. 205). Put differently, mobilization is simply the actions (or “agencies” in Clegg’s (1989) terminology) executed by an organization and its agents for the purpose of ensuring that the policy or reform of interest is defined and understood in a certain way.

Thus, overall, the sociology of translation describes a rather straight-forward political process, and one that has echoes of other approaches, such as Kingdon’s (1984) policy streams framework. However, Clegg’s (1989) approach has the benefit, first, that it focuses explicitly on the *internal* dynamics of organizations (as opposed to being concerned with more general political systems and policy arenas) and, second, that brings a particular approach to conceptualizing individual agency. We turn to this latter issue next.

Clegg (1989) offers a way of understanding agency that is usefully nuanced. For Clegg (1989), organizational change is not explained by the internalization of institutional “myths” (as in WCT) or by relations of economic dependence (as with certain IPE strands and neo-Marxist international relations). Rather, organizational change is the result of the way that “episodic agency” operates in a matrix of “social relations” and “standing conditions,” with these dynamics then dialectically intersecting with two circuits of power—namely, the “circuit of social integration” and the “circuit of system integration” (see Fig. 1). To be clear, episodic agency is the term used by Clegg (1989) to refer to autonomous actors’ free-thinking, whereas standing conditions refer to resources and the means to which actors have access for the purpose of advancing their preferred outcomes. The circuit of social integration, for its part, is “concerned with fixing or refixing relations of meaning” (p. 224), while the circuit of system integration is focused on changes to strategies for discipline and production. Here, “production” can be understood as the daily business of the organization in question. Furthermore, as agents engage in actions to create shared understandings and to gain approval for new ways of doing things, these new meanings and techniques of production (or discipline) must successfully pass through required points of organizational regulation or approval, labeled by Clegg (1989) as “obligatory passage points.” Once these new meanings and practices are approved, they shift the social relations and standing conditions within which actors maneuver. Finally, as Fig. 2 depicts, the impetus for pursuing organizational change can result from factors that are both endogenous (i.e., coming from within an organization) and exogenous (i.e., stemming from external pressures and opportunities).

It may be helpful to restate the language of this framework such that what it envisions is easily grasped. There are four key points. First, individuals—in our case, organizationally-positioned policy entrepreneurs—act within a network of relations that are not only social (i.e., human) and non-human¹¹ but also based on their position vis-à-vis standing conditions (i.e., based on the extent to which they can access and can utilize resources). Second, these actions seek to produce certain kinds of intermediate outcomes (understood as actions or milestones) that contribute to a transformation of the existing “relations of meaning” (i.e., which policies are seen as desirable) and the “techniques of production” (i.e., how the organization does what it does; in the case of the World

¹¹We note that Clegg’s (1989) framework connects with STS in that it accommodates non-human activity as a form of agency. Clegg (1989) was clearly attuned to this form of agency, as when he wrote that “one has to withstand any ontological impulse to restrict the category of agency only to human beings” (p. 187), or, as an additional example, that “not all agents are human actors. Agency must be vested in non-human entities as diverse as machines, germs, animals and natural disasters” (p. 188).

Bank, for example, this can mean the reforms and models that it funds and attempts to implement through its loans and technical assistance). Third, for these transformations to occur, the proposed change has to successfully move through “obligatory passage points,” that is, mechanisms of control within the organization that are defined by the rules of how the organization operates. These passage points are often individuals in positions of authority who serve as what we would call “guarantors of the status quo.” Taking the World Bank again as an example, this could mean the loan approval process, with all its safeguards and legal requirements. Finally, if these transformations succeed, they will alter, in a dialectic fashion, the network of social relations within which an actor operates as well as their position within the organizational standing conditions. This makes sense: As an actor contributes to (positively perceived) change within an organization, their social standing in that organization will naturally rise.

Because organizational agents are constrained by their human/non-human relations and standing conditions, not to mention the circuits of social and system integration as well as obligatory passage points, they are not free to pursue any change they may wish. That is, they cannot freely pursue any desired change. As Clegg (1989) writes, “the interests of actors in organizations and the decisions that they make are necessarily contingent on various forms of organizational calculation” (p. 197). It may be the case that certain agents in international organizations have “high discretionary strategic agency” (p. 199)—and that such discretion, as principal-agent theory suggests, may be “organizationally creative, productive, [or] reproductive” (p. 200)—yet it is also true that “power will always be inscribed within contextual ‘rules of the game’” and that “these rules may be taken to be the underlying rationale of those calculations which agencies routinely make in organizational contexts” (200). In other words, the “interests” upon which agents act are informed by the representations (i.e., “discourse”) they encounter of what is important to the organization, with these representations, of course, being in part a consequence of the structural characteristics of the organization in which one is embedded. The implication is that this approach is not concerned with the “true interests” of organizational agents, but rather only in the actions they take to work toward certain outcomes within the relevant structures of opportunity (e.g., career advancement) and discipline (e.g., dismissal).

In combining the sociology of translation with the circuits of organizational power approach, the task for scholars is to explain how organizational agents move through the four moments of translation while navigating the dimensions of the circuits of power framework. An example is provided by Edwards and Moschetti (2021), who examine policy change inside the World Bank. They look at how advocates within the World Bank fought for the acceptance of the policy of community-based management of education. More specifically, they look at how this policy was adapted from the context of El Salvador, promoted within the World Bank, and then finally fully embraced as one of this organization’s flagship policies during the 1990s and 2000s.

Conclusions: strengths and possibilities

The previous section began by asserting that Clegg’s (1989) approach brings together and extends the strengths of other approaches. We return to that point here. As the discussion above has shown, the circuits of power approach:

- (a) Accommodates the dissemination of values and discourse (as in WCT) but without sacrificing strategic individual agency (as promoted by the strategic-relational approach, institutional entrepreneurs, and the policy streams approach);
- (b) Takes the discretion of the principal-agent approach but opens it up to avoid the assumption of unitary organizational actors (similar to the global governance approach but with much more theoretical specificity when it comes to understanding organizational dynamics);
- (c) Reflects the key tenet of STS that non-human agency must also be taken into account, but without the need to relinquish a focus on intentionality and actors’ interests, and also without the need to eschew a general, guiding theory of organizational change dynamics;
- (d) Maintains emphasis on structures, for example, those larger (or exogenous) structures within which the organization is based (as with IPE, CPE, and Kingdon’s (1984) politics stream) but also those structures internal (or endogenous) to the organization (as with the strategic-relational approach and institutional entrepreneurship literature)—though without assuming that either individual or organizational behavior is entirely determined by them;
- (e) Is able to explain change to internal organizational structures (as with historical institutionalism) but with more analytic guidance (beyond claiming attention to organizational context and the role of ideas or beliefs);
- (f) Draws attention to “bureaucratic rules and associated behaviors” within organizations (as in the case of sociological institutionalism; Menashy and Manion, 2016, p. 321) but reframes the rules-social relations intersection as mediated and influenced by additional factors; and
- (g) Attends to contextual pressures related to organizational legitimacy but usefully combines attention to this with a nuanced notion of what propels and constraints behavior inside the organization, as individuals react to external pressures.

With the above in mind, we conclude by returning to the goals set out at the beginning. Ultimately, in providing this overview of different approaches, this entry has sought, first, to make clear the options available for guiding research on international organizations, second, to emphasize an approach that is seen as particularly insightful, and, third, to encourage more explicit and theoretically-informed discussions of the nature of international organizations. This last goal is, we suggest, important given the centrality of international organizations and their networks to the field of education generally and global education policy specifically.

On this last point, about networks, it should be stated that Clegg's (1989) approach—and the broader discussion offered here—can indeed be extended further. Although Clegg (1989) has advantages over other approaches, his framework, too, can, and should, be further developed by bringing it into conversation with networks. His focus in this area is perhaps underdeveloped because his emphasis was on what happens *inside* organizations. On the other hand, Clegg (1989) might respond that his attention to “social relations” captures or subsumes a focus on networks. In any event, while there is not a substantive focus on this issue in Clegg's (1989) approach, networks and the phenomenon of network governance have garnered significant attention since the publication of his book (Ball, 2012; Edwards et al., 2021b; Olmedo, 2017). Thus, one promising area for future work—among others, such as bringing more critical attention to how his framework accommodates, or doesn't, a focus on culture within organizations (Anderson-Levitt, 2012)—would be to flesh out how these two areas of scholarship complement one another.

References

- Addey, C., Gorur, R., 2020. Translating PISA, translating the world. *Comp. Educ.* 56 (4), 547–564.
- Andersen-Levitt, K., 2012. Complicating the concept of culture. *Comp. Educ.* 48 (4), 441–454.
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education “best practice”. *Comp. Educ.* 52 (2), 202–229. <https://doi.org/10.1080/03050068.2016.1143278>.
- Ball, S.J., 2012. *Global Education Inc.: New Policy Networks and the Neo-liberal Imaginary*. Routledge, London.
- Barkin, S., 2006. *International Organization: Theories and Institutions*. Palgrave Macmillan, New York.
- Barnett, M., Finnemore, M., 2005. The power of liberal international organizations. In: Barnett, M., Duvall, R. (Eds.), *Power in Global Governance*, pp. 161–184. Cambridge: Cambridge.
- Benavot, A., 2011. Imagining a transformed UNESCO with learning at its core. *Int. J. Educ. Dev.* 31, 558–561.
- Bermingham, D., 2011. The politics of global education policy: the formation of the Education for All–Fast Track Initiative (FTI). *J. Educ. Pol.* 26 (4), 557–569.
- Bitektine, A., 2011. Toward a theory of social judgments of organizations: the case of legitimacy, reputation, and status. *Acad. Manag. Rev.* 36 (1), 151–179.
- Blau, P.M., Scott, W., 2003. *Formal Organizations: A Comparative Approach*. Stanford University Press.
- Boli, J., Ramirez, F., Meyer, J., 1985. Explaining the origins and expansion of mass education. *Comp. Educ. Rev.* 29 (2), 145–170.
- Bonal, X., 2002. Plus ça change ... the World Bank global education policy and the post-Washington consensus. *Int. Stud. Sociol. Educ.* 12 (1), 3–22.
- Broad, R., 2006. Research, knowledge, and the art of paradigm maintenance: the World Bank's Development Economics Vice-Presidency (DEC). *Rev. Int. Polit. Econ.* 13 (3), 387–419.
- Bromley, P., Powell, W.W., 2012. From smoke and mirrors to walking the talk: decoupling in the contemporary world. *Acad. Manag. Ann.* <https://doi.org/10.1080/19416520.2012.684462>. *The Academy of Management Annals*.
- Buchanan, A., Keohane, R., 2006. The legitimacy of global governance institutions. *Ethics Int. Aff.* 20 (4), 405–437.
- Burnett, N., 2019. Invited essay: it's past time to fix the broken international architecture for education. *Int. J. Educ. Dev.* 68, 15–19. <https://doi.org/10.1016/j.ijedudev.2019.04.005>.
- Burrell, G., Morgan, G., 1979. *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*. Routledge.
- Camic, C., Gross, N., Lamont, M. (Eds.), 2011. *Social Knowledge in the Making*. University of Chicago Press, Chicago.
- Campbell, J.L., 1998. Institutional analysis and the role of ideas in political economy. *Theory Soc.* 27, 377–409.
- Campbell, J.L., 2002. Ideas, politics, and public policy. *Annu. Rev. Sociol.* 28, 21–38.
- Capoccia, G., Keleman, D., 2007. The study of critical junctures. *World Polit.* 59 (3), 341–369.
- Carney, S., Rappleye, J., Silova, I., 2012. Between faith and science: world culture theory and comparative education. *Comp. Educ. Rev.* 56 (3), 366–393.
- Chabbot, C., 1998. Constructing educational consensus: international development professionals and the world conference on education for all. *Int. J. Educ. Dev.* 18 (3), 207–218.
- Clegg, S.R., Courpasson, D., Phillips, N., 2006. *Power and Organizations*. SAGE Publications Ltd, London.
- Clegg, S.R., 1989. *Frameworks of Power*. SAGE Publications Ltd, London.
- Dale, R., 2000. Globalization and education: demonstrating a “common world educational culture” or locating a “globally structured educational agenda”? *Educ. Theor.* 50 (4), 427–448.
- Desai, V., 2011. Mass media and massive failures: determining organizational efforts to defend field legitimacy following crises. *Acad. Manag. J.* 54 (2), 263–278.
- Dowling, J., Pfeffer, J., 1975. Organizational legitimacy: social values and organizational behavior. *Pac. Sociol. Rev.* 18 (1), 122–136.
- Downey, L., Lawrence, E., Pyles, M., Lee, D., 2020. Power, hegemony, and world society theory: a critical evaluation. *Socius* 6, 1–22. <https://doi.org/10.1177/2378023120920059>.
- Edwards Jr., D.B. (2021). Globalization and education: Understanding and problematizing approaches to global governance and the global education policy field. In H. Steyn & C. Wolhuter (Eds.), *Critical issues in education systems: Comparative-international perspectives*. (pp. 247–278). Axiom.
- Edwards Jr., D.B., Moschetti, M., 2021. The sociology of policy change within international organizations: beyond coercive and normative perspectives—towards circuits of power. *Glob. Soc. Educ.* 19 (1), 55–69.
- Edwards Jr., D.B., Urquilla Loucel, C., 2016. The EDUO program, impact evaluations, and the political economy of global education reform. *Educ. Pol. Anal. Arch.* 24 (92), 1–46. <https://doi.org/10.14507/epaa.24.2019>.
- Edwards Jr., D.B., Okitsu, T., da Costa, R., Kitamura, Y., 2018. Organizational legitimacy in the global education policy field: learning from UNESCO and the Global Monitoring Report. *Comp. Educ. Rev.* 61 (1), 31–63.
- Edwards Jr., D.B., Morrison, J., Hall, S., 2020. The suspect statistics of best practices: a triple critique of knowledge production and mobilization in the global education policy field. *Glob. Soc. Educ.* 18 (2), 125–148.
- Edwards Jr., D.B., Rappoport, A., Sperdutti, V., Caravaca, A., 2021a. The Influence of the World Bank on Policy Formation, Policy Implementation, and Private Education: A Systematic Review of the Literature, Background Paper for the Global Education Monitoring Report. UNESCO. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000380064>.
- Edwards Jr., D.B., Moschetti, M., Caravaca, A., 2021b. Network governance and new philanthropy in Latin America and the Caribbean: the reconfiguration of the State in the Dominican Republic and beyond. *Br. J. Sociol. Educ.* 42 (8), 1210–1226.
- Edwards Jr., D.B., 2018a. *Global Education Policy, Impact Evaluations, and Alternatives: The Political Economy of Knowledge Production*. Palgrave MacMillan, New York.
- Edwards Jr., D.B., 2018b. *The Trajectory of Global Education Policy: Community-Based Management in El Salvador and the Global Reform Agenda*. Palgrave MacMillan, New York.
- Fioretos, O., 2011. Historical institutionalism in international relations. *Int. Organ.* 65, 367–399.
- Fombrun, C., Shanley, M., 1990. What's in a name? Reputation building and corporate strategy. *Acad. Manag. J.* 33 (2), 233–258.
- Foucault, M., 1980. In: Gordon, C. (Ed.), *Power/Knowledge*. Pantheon, New York.

- Garud, R., Hardy, C., Maguire, S., 2007. Institutional entrepreneurship as embedded agency: an introduction to the special issue. *Organ. Stud.* 28 (7), 957–969. <https://doi.org/10.1177/0170840607078958>.
- Gilbert, C.L., Vines, D., 2006. *The World Bank: Structure and Policies*. Cambridge University Press.
- Gorur, R., Hamilton, M., Lundahl, C., Sjödin, E.S., 2019. Politics by other means? STS and research in education. *Discourse* 40 (1), 1–15. <https://doi.org/10.1080/01596306.2018.1549700>.
- Hall, P.A., Taylor, R.C.R., 1996. Political science and the three new institutionalisms. *Polit. Stud.* 44 (5), 936–957. <https://doi.org/10.1111/j.1467-9248.1996.tb00343.x>.
- Hawkins, D., Lake, D., Nielson, D., Tierney, M. (Eds.), 2006. *Delegation and Agency in International Organizations*. Cambridge University Press, Cambridge.
- Hay, C., Marsh, D., 1999. Introduction: towards a new (international) political economy? *New Polit. Econ.* 4 (1), 5–22.
- Hay, C., 2002. Globalisation as a problem of political analysis: restoring agents to a “process without a subject” and politics to a logic of economic compulsion. *Camb. Rev. Int. Aff.* 15 (3), 379–392.
- Heyneman, S.P., Lee, B., 2016. International organizations and the future of education assistance. *Int. J. Educ. Dev.* 48, 9–22. <https://doi.org/10.1016/j.ijedudev.2015.11.009>.
- Huberman, B.A., Loch, C.H., Onculer, A., 2004. Status as a valued resource. *Soc. Psychol. Q.* 67 (1), 103–114.
- Jensen, M., Roy, A., 2008. Staging exchange partner choices: when do status and reputation matter? *Acad. Manag. Ann.* 51 (3), 495–516.
- Jakobi, A.P., 2009. Global education policy in the making: international organisations and lifelong learning. *Glob. Soc. Educ.* 7 (4), 473–487.
- Jessop, B., 2001a. Institutional re(turn)s and the strategic-relational approach. *Environ. Plann.* 33, 1213–1235.
- Jessop, B., 2001b. What follows Fordism? On the periodization of capitalism and its regulation. In: Albritton, R., Itoh, M. (Eds.), *Phases of Capitalist Development: Booms, Crises, and Globalization*. Palgrave, pp. 282–299.
- Jessop, B., 2004. Critical semiotic analysis and cultural political economy. *Crit. Discourse Stud.* 1 (2), 159–174.
- Jessop, B., 2010. Cultural political economy and critical policy studies. *Crit. Pol. Stud.* 3 (3–4), 336–356.
- Jones, P.W., Coleman, D., 2005. *The United Nations and Education: Multilateralism, Development and Globalisation*. RoutledgeFalmer, New York, NY.
- Jones, P.W., 1992. *World Bank Financing of Education: Lending, Learning, and Development*. Routledge, Abingdon, UK.
- Jones, P., 2007a. Education and world order. *Comp. Educ.* 43 (3), 325–337.
- Jones, P.W., 2007b. Global governance, social policy and multilateral education. *Comp. Educ.* 43 (3), 321–323.
- Jones, P., 2008. Cultural political economy and the international governance of education: theoretical framework. Paper Presented at the Workshop for “International Governance in Education: Theoretical Perspectives, Themes and Areas,”. University of Tübingen. October 2–5, 2008.
- Jones, P.D., 2010. Cultural political economy and the international governance of education: a theoretical framework. In: Amos, S.K. (Ed.), *International Perspectives on Education and Society*. Emerald, pp. 19–56.
- Kamens, D., McNeely, C., 2010. Globalization and the growth of international educational testing and national assessment. *Comp. Educ. Rev.* 54 (1), 5–25.
- Kingdon, J., 1984. *Agendas, Alternative, and Public Policies*. Little Brown, Boston.
- Klees, S., Edwards Jr., D.B., 2014. Knowledge production and technologies of governance. In: Fenwick, T., Mangez, E., Ozga, J. (Eds.), *World Yearbook of Education 2014: Governing Knowledge: Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education*. Routledge, New York, pp. 31–43.
- Klees, S., 2012. World Bank and education: ideological premises and ideological conclusions. In: Klees, S., Samoff, J., Stromquist, N. (Eds.), *The World Bank and Education: Critiques and Alternatives*. Sense, Rotterdam, pp. 49–68.
- Koyama, J.P., Varenne, H., 2012. Assembling and disassembling. *Educ. Res.* 41 (5), 157–162. <https://doi.org/10.3102/0013189x12442799>.
- Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford University Press, Oxford.
- Lingard, B., Rawolle, S., Taylor, S., 2005. Globalizing policy sociology in education: working with Bourdieu. *J. Educ. Pol.* 20 (6), 759–777.
- Lingard, B., Sellar, S., Baroutsis, A., 2015. Researching the habitus of global policy actors in education. *Camb. J. Educ.* 45 (1), 25–42.
- Maguire, S., Hardy, C., Lawrence, T.B., 2004. Institutional entrepreneurship in emerging fields: HIV/AIDS treatment advocacy in Canada. *Acad. Manag. J.* 47, 657–679.
- Massey, J.E., 2001. Managing organizational legitimacy: communication strategies for organizations in crisis. *J. Bus. Commun.* 38 (2), 153–183.
- McNeely, C., Cha, Y., 1994. Worldwide educational convergence through international organizations: avenues for research. *Educ. Pol. Anal. Arch.* 2 (14), 1–11. Retrieved from <http://epaa.asu.edu/ojs/article/view/677>.
- McNeely, C.L., 1995. Prescribing national education policies: the role of international organizations. *Comp. Educ. Rev.* 39 (4), 483–507.
- Menashy, F., Manion, C., 2016. The historical evolution and current challenges of the United Nations and global education policy-making. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*, first ed. John Wiley & Sons, Oxford, pp. 319–334.
- Meyer, J.W., Rowan, B., 1977. Institutionalized organizations: formal structure as myth and ceremony. *Am. J. Sociol.* 83 (2), 340–363.
- Meyer, J., Boli, J., Thomas, G., Ramirez, F., 1997. World society and the nation-state. *Am. J. Sociol.* 103 (1), 144–181.
- Milana, M., 2015. Global policy in adult education and UNESCO: landmarking, brokering and framing policy. *Glob. Soc. Educ.* 14 (2), 203–226. <https://doi.org/10.1080/14767724.2015.1010437>.
- Moutios, S., 2009. International organisations and transnational education policy. *Compare* 39 (4), 467–478.
- Moutios, S., 2010. Power, politics and transnational policy-making in education. *Glob. Soc. Educ.* 8 (1), 121–141.
- Mundy, K., Menashy, F., 2014. The World Bank and private provision of schooling: a look through the lens of sociological theories of organization hypocrisy. *Comp. Educ. Rev.* 58 (3), 401–427.
- Mundy, K., 1998. Educational multilateralism and world (dis)order. *Comp. Educ. Rev.* 42 (4), 448–478.
- Mundy, K., 2002. Retrospect and prospect: education in a reforming World Bank. *Int. J. Educ. Dev.* 22, 483–508.
- Mundy, K., 2007. Global Governance, educational change. *Comp. Educ.* 43 (3), 339–357.
- Mundy, K., Verger, A., 2015. The World Bank and the global governance of education in a changing world order. *Int. J. Educ. Dev.* 40, 9–18.
- Olmedo, A., 2017. Something old, not much new, and a lot borrowed: philanthropy, business, and the changing roles of government in global education policy networks. *Oxf. Rev. Educ.* 43 (1), 69–87.
- Ramirez, F., Boli, J., 1987. The political construction of mass schooling: European origins and worldwide institutionalization. *Sociol. Educ.* 60 (1), 2–17.
- Ramirez, F., Meyer, J., Lerch, J., 2016. World society and the globalization of education policy. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley, Oxford, UK, pp. 43–63.
- Rhee, M., Haunschild, P.R., 2006. The liability of good reputation: a study of product recalls in the U.S. automobile industry. *Organ. Sci.* 17 (1), 101–117.
- Samoff, J., 1993. The reconstruction of schooling in Africa. *Comp. Educ. Rev.* 37 (2), 181–222.
- Samoff, J., 2004. From funding projects to supporting sectors? Observation on the aid relationship in Burkina Faso. *Int. J. Educ. Dev.* 24, 397–427.
- Samoff, J., 2009. Foreign aid to education: managing global transfers and exchanges. In: Chishom, L., Steiner-Khamisi, G. (Eds.), *South-South Cooperation in Education and Development*. Teachers College Press, New York, pp. 123–156.
- Samoff, J., 2012. “Research shows that ...”: creating the knowledge environment for learning for all. In: Klees, S., Samoff, J., Stromquist, N. (Eds.), *The World Bank and Education: Critiques and Alternatives*. Sense, Rotterdam, pp. 141–156.
- Savage, G.C., 2019. What is policy assemblage? *Territ. Politics Gov.* 8 (3), 319–335. <https://doi.org/10.1080/21622671.2018.1559760>.
- Schmidt, V., 2008. Discursive institutionalism: the explanatory power of ideas and discourse. *Annu. Rev. Polit. Sci.* 11, 303–326.
- Steiner-Khamisi, G., 2006. The economics of policy borrowing and lending: a study of late adopters. *Oxf. Rev. Educ.* 32 (5), 665–678.
- Suchman, M.C., 1995. Managing legitimacy: strategic and institutional approaches. *Acad. Manag. Rev.* 20 (3), 571–610.

- Sumida, S., 2017. Agenda setting in multilateral contexts: the example of the decade of education for sustainable development. *Glob. Soc. Educ.* 15 (4), 381–396. <https://doi.org/10.1080/14767724.2017.1356706>.
- Verger, A., Edwards Jr., D.B., Kosar-Altinyelken, H., 2014. Learning from all? The World Bank, aid agencies and the construction of hegemony in education for development. *Comp. Educ.* 50 (4), 1–19.
- Verger, A., 2012. Framing and selling global education policy: the promotion of public-private partnerships for education in low-income contexts. *J. Educ. Pol.* 27 (1), 109–130.
- Verger, A., 2014. Why do policy-makers adopt global education policies? Toward a research framework on the varying role of ideas in education reform. *Curr. Issues Comp. Educ.* 16 (2), 14–29.
- Wade, R., 1996. Japan, the World Bank, and the art of paradigm maintenance: the east asian miracle in political perspective. *N. Left Rev.* 1 (217), 1–16.
- Wallerstein, I., 1984. *The Politics of the World Economy: The States, the Movements and the Civilisations*. Cambridge University Press, Cambridge, UK.
- Washington, M., Zajac, E.J., 2005. Status evolution and competition: theory and evidence. *Acad. Manag. J.* 48 (2), 282–296.
- Wejnert, B., 2002. Integrating models of diffusion of innovations: a conceptual framework. *Annu. Rev. Sociol.* 28 (1), 297–326.
- Zapp, M., 2019. International organizations, educational development, and conflict: a (world) cultural perspective. *Prospects* 47 (4), 321–337. <https://doi.org/10.1007/s11125-019-09442-4>.
- Zürn, M., 2004. Global governance and legitimacy problems. *Gov. Oppos.* 39 (2), 260–287.

The World Bank and education in a changing global arena

Xavier Bonal^a, Clara Fontdevila^b, and Adrián Zancajo^c, ^a Department of Sociology, Universitat Autònoma de Barcelona, Catalonia, Spain; ^b School of Education, University of Glasgow, Glasgow, United Kingdom; and ^c Department of Sociology, Universitat Autònoma de Barcelona, Catalonia, Spain

© 2023 Elsevier Ltd. All rights reserved.

Introduction	480
The emergence of the WB as a prominent actor in education for economic development	480
The World Bank and the Post-Washington Consensus	482
The World Bank after 2008: building a partial heterodoxy	484
Future directions: which role of the World Bank in the global education arena?	485
References	486

Introduction

The influence of the World Bank in shaping global education policies can hardly be disputed. Backed by the rationale of human capital theory, the quantitative and qualitative expansion of the World Bank's educational activities allowed this supranational organization to hold for many years a quasi-monopoly in the area of international aid for educational development. Interestingly enough, an institution that never had an explicit education mandate among its goals became the most prominent actor in the field of educational development. Changes in the economic world order, shifts in educational multilateralism and internal crises in institutions such as UNESCO resulted in the dominance of institutions such as the IMF or the World Bank in shaping the international agenda for development, in particular after the financial crisis of the early 1980s (Mundy, 1999). Even if the World Bank's influence appears to be dwindling and the organization is the object of mounting criticism, its power and ascendancy within the global education field remain unmatched.

For almost 60 years, the Bank¹ has exhibited the financial and technical capacity to become a prominent actor in education and development. However, its role has changed over time, from a more prescriptive and strict imposition of specific policies to a more versatile and flexible institution with a higher disposition to adapt to the needs of the borrowing countries. The debate about the nature of this change is open. A critical controversy is whether the organization has abandoned a neoliberal paradigm of development or has never done so, introducing only apparent or superficial changes that keep its ideology unaltered.

Nonetheless, and beyond this debate, there are visible changes in the policy orientation and programmatic priorities of the strategic sectoral papers that the Bank has produced in the last decades. The fluctuations in the World Bank's strategies have to be understood in a context of a changing landscape of multilateral aid. The legitimacy crisis of the Bank's neoliberal agenda, new international actors gaining higher presence in the field of international education or the globalization of targets and priorities in the field of economic and social development may explain why the Bank has shifted its priorities and strategies in education.

This chapter explores three main phases of the World Bank approach to education. The first one refers to the irruption and consolidation of the organization as a key actor in educational aid, especially through mechanisms of loan conditionality during the 1980s. A second period starts with the crisis of Structural Adjustment Programs (SAPs) and the opening of the Bank toward the Post-Washington Consensus in the 1990s. A change toward a less prescriptive and more comprehensive role was effected during the James Wolfensohn's mandate and the launch of the Poverty Reduction Strategy Papers. Finally, the third phase refers to the aftermath of the 2008 crisis and the importance given to learning in the international scene of educational development. New global agendas for education, new actors and changes within the organization that changed the balance of power explain a higher level of heterodoxy of the World Bank education policy, particularly visible in the 2018 World Development Report.

Although the three periods identified are not homogeneous in length, each one captures important shifts in the priorities and strategies in education set by the Bank. For each period, we describe the key features of the WB strategy and identify the main factors to understand its priorities and the reasons for the policy shifts. A final section concludes and reflects on the future directions that the organization may take in the current global scenario of the Sustainable Development Goals (SDGs).

The emergence of the WB as a prominent actor in education for economic development

Education as a capital investment sector was not part of the World Bank portfolio when the Bank was first established in 1944. Nevertheless, the formulation of Human Capital Theory by Theodore Schultz and Gary Becker in 1960 justified and opened the

¹The terms *the World Bank*, *the Bank* and *the WB* are used interchangeably in this entry.

door to the possibility of lending to education. Key works on the economic value of education gave a rationale for understanding the contribution of education to economic growth and valuing the concept of human capital (Becker, 1974; Psacharopoulos and Woodhall, 1985). In 1960 the International Development Association (IDA) was founded as a Bank affiliate to promote economic and social progress in the poorest countries. Education was considered a key sector for social progress and the improvement of living standards. For the WB, education was mainly understood as a key sector for social and economic development. From this perspective, investing in education became strategic for governments, since education was supposed provided returns to investment at both social and private levels.

The first WB education project was launched in 1962. Since then, the field of educational development cannot be understood without examining the role of this international organization. Despite education not being the most prominent sector of the Bank's capital investment, lending in education has accounted for proportions between 5% and 9% of the Bank's portfolio (Mundy and Verger, 2015). The arrival of Robert McNamara to the WB presidency in 1968 reinforced the role of the Bank as a social policy entrepreneur, and focused its role on the objective of poverty reduction and the satisfaction of basic needs. As Stephen Heyneman argues, by the end of the 1960s and during the 1970s, investment in education was seen as a key aspect for ensuring the necessary technical capacities for development (Heyneman, 2003). The focus on agriculture, industry and manufacturing increased the importance of recurrent expenditures -unlike the exclusive concentration on capital investment that characterized the first years of the World Bank- and put capacity building as one of the main goals for social development. Human capital needs were assessed through manpower planning forecast systems, which provided the rationales for lending in education.

However, and unquestionably, it was during the 1980s that the WB consolidated its position as the most influential and powerful agency in the area of educational development. The difficulties of many developing countries in facing their external debt in the context of an enormous increase in interest rates induced a change in the lending system of the Bretton Woods institutions. Their hegemonic position within the international market of financial assistance for development and the extreme dependency of highly indebted countries on international loans, mostly to refund external debt interests, completely shifted the relationship between the Bretton Woods institutions and borrowing countries. Both the IMF and the WB established a new system of loan conditionality based on the implementation of specific macroeconomic measures and institutional reforms. Loan conditionality thus allowed for the so-called Washington Consensus to become a blueprint for economic development. Through programs like Structural Adjustment Facilities, Extended Fund Facilities or Structural Adjustment Loans, both the IMF and the WB were able to impose the neoliberal paradigm of development to borrowing countries, regardless of the previous social and economic conditions. Countries receiving assistance needed to signal a strong commitment to fiscal adjustment, tax reforms, deregulation, trade liberalization, attraction of foreign investment and privatization in their plans in order to continue to have access to international funding (Cornia et al., 1997).

The translation of this loan conditionality in education was named by Stephen Heyneman as the "short education policy menu", visible in the 1980 *Education sector policy paper* (World Bank, 1980). Emerging from the rate of return analysis (Bennell, 1996), education policies were supposed "(i) to shift public expenditures away from vocational and higher education toward academic and basic education; (ii) to increase the private cost for attending universities; and (iii) to install loan schemes to set off the financial burden on individuals who now must face high tuition fees for higher education" (Heyneman, 2003, p. 325). In addition, the short education policy menu included cost-recovery strategies (including fees in basic education) and decentralization strategies to improve the efficiency of education systems. Interestingly, the priority for basic education bolstered the credibility of the Bank because it relied on apparently irrefutable arguments legitimated by the dominant paradigm of economics of education. It also allowed the Bank to appear as an institution concerned with the struggle against poverty and the expansion of educational opportunities. At the same time, the rates-of-return rationale legitimated other arguments favoring more user participation in the financing of post-basic education and the increase of the private sector in educational supply.

Structural adjustment programs and the adoption of the short education policy menu entailed a change in the political rationality of recipient countries (Morrow and Torres, 2000). Their higher dependency on multilateral funding changed the priorities and strategies for education and limited the number of options available. In this way, the new lending system did not only reduce national autonomy to set educational strategies, but significantly restricted the available policy options. Recipient countries were aware that they had to adopt the recommended policies to keep receiving funding from the WB.

A number of studies about the impact of SAPs on education (Reimers and Tiburcio, 1993; Carnoy and Torres, 1994) evidenced several negative effects on education indicators, such as the fall in educational expenditure as a percentage of the GDP, the reduction of per capita education expenditure, the increase in educational absenteeism and school failure or a contraction in access to primary and secondary education (Bonai, 2002). It is not clear whether these negative effects were the result of the direct effects of the short education policy menu or whether they were rather caused by the indirect effects of severe macro-economic adjustment measures dictated by international financial institutions, with clear implications on educational budgets (Dale, 1999). The rise of opportunity costs of educational investment for the poor (traditionally undermined by the World Bank estimations), the need to plan long-term investments to overcome poverty, the higher costs of education linked to cuts in public budgets and increasing privatization, or the opting-out of the middle classes from less protected and low-quality public services all are factors that can explain why SAPs had significant negative effects on public education in most developing countries.

Paradoxically, the failure of neoliberal policies in education did not induce immediate changes in the WB agenda. The creation of the World Bank's Development Economics Vice-Presidency played a key role in the "art of the paradigm maintenance" (Wade, 1996). The research department of this vice presidency played a critical role in legitimizing the neoliberal "free-trade"

paradigm for more than two decades. By means of hiring “adequate professionals,” creating incentives for internal promotion, discouraging dissonant discourses, selecting the presentation of data or projecting “rigorous and sound” research, the Bank reinforced its image as a knowledge Bank within the paradigm of neoliberal economics (Broad, 2007).

Nevertheless, and despite this art of paradigm maintenance, the reverse effects of SAPs on poverty explain the progressive abandonment of the strict neoliberal orthodoxy and the subsequent adoption of a development agenda that includes strategies complementary to a market-led model of economic growth. Thus, in the late 1980s, the Washington Consensus was experiencing difficulties in the face of a number of criticisms. The rise of poverty and subsequent complaints about structural adjustment effects led to demands for an adjustment with a human face (Ilon, 1996). Simultaneously, the success of East Asian economies opened claims for a different role of the State in development (Fine, 2001), which challenged the neoliberal assumptions of the Washington Consensus and claimed for context-based national strategies of industrialization to ensure the integration into the global economy (Gore, 2000).

Against such background, the Bank’s decision to join the 1990 World Conference on Education for All (WCEFA) celebrated in Jomtien can be mainly understood as an attempt to regain some of the organization’s lost legitimacy. Unlike UNICEF or UNESCO, the World Bank was more of a late and hesitant sponsor of the WCEFA (Chabbott, 2003). Nonetheless, the conference eventually came to be perceived within the WB’s education circles as an opportunity to regain a position of political leadership against a much-debilitated UNESCO, and to persuade other education agencies that the WB policy prescriptions were in line with the new education agenda then in the making (Jones, 2007a). However, efforts at the restitution of the World Bank’s credibility were not confined to the education sector—they also took place at an organizational level. As we will see in the next section, the opening of a debate about development that would eventually create the conditions for a Post-Washington Consensus owes much to the work of Joseph Stiglitz as Chief Economist, and to the changes introduced during the presidency of James Wolfensohn in the 1990s.

The World Bank and the Post-Washington Consensus

In the light of the mounting evidence on the negative impact of SAPs in education, by the mid-1990s, the “short education policy menu” started to show signs of exhaustion—outside but also within the WB education circles. While the commitment to the WCEFA allowed the Bank to regain a sense of alignment with the global education community, it could not entirely neutralize the internal and external legitimacy damage suffered by the organization. The publication of the *Strategies and Priorities for Education* policy document² in 1995 (World Bank, 1995) was a critical episode that laid bare the growing levels of internal dissent and the appetite for a change in the World Bank’s policy approach. Thus, the strategic document was prepared amid a climate of tension that pitted regional managers and operational staff against the Bank’s education staff based in the headquarters. According to Heyneman (2003), regional staff feared that the emphasis on the short policy menu would put at risk their relationship with borrowers. Their demands for a less orthodox approach were, however, largely ignored by the team in charge of drafting the new policy document. Thus, and quite predictably, the publication of the 1995 paper faced a considerable backlash within education policy circles internal and external to the Bank. The emphasis on the decentralization and privatization policies and the ratification of the short education policy menu did not bode well with the expansionist and aspirational spirit of the Education for All (EFA) agenda (Jones and Coleman, 2005), and the paper became the object of a number of replies and critical pieces (e.g., Lauglo, 1996; Samoff, 1996).

The poor reception of the 1995 paper gave rise to a period of soul-searching that would ultimately result in a reconfiguration of the World Bank’s discourse on education policy. As Jones and Coleman (2005) affirm, “much subsequent bank thinking and pronouncing about education [*post 1995*] needs to be seen partly as an attempt to project a softer, more inclusive and less strident view of education and its benefits” (p. 122). Thus, over the years to come, the World Bank education sector launched different initiatives oriented at appeasing its internal and external critics. However, and as discussed below, these changes left largely untouched some fundamentals of the short policy menu.

This relative appetite for institutional change within the Bank’s education sector needs to be understood in relation to broader changes affecting the organization. Thus, by the mid-1990s, the World Bank was close to a crisis of legitimacy. Different civil society campaigns had put the Bank under the spotlight, drawing attention to the regressive effects of SAPs (Vet-terlein, 2007). Prominent figures within the World Bank were also emphatic about the need for structural change—most notably, Joseph Stiglitz, appointed Chief Economist in 1997, and who would become a vocal critic of the Washington Consensus orthodoxy and market fundamentalism (cf. Stiglitz, 1998, 2003). In addition, the World Bank’s centrality as a development partner was being eclipsed by the emergence and consolidation of other development partners, including the Inter-American Development Bank and the Asian Development Bank. The market share of the WB was thus dwindling, and the Bank’s dogmatic approach proved increasingly problematic in responding to the demands of an increasingly heterogeneous pool of borrowers (Mundy and Verger, 2015).

²Note that although the paper was formally devised as a policy review, it had a distinctly prescriptive nature.

Thus, under the presidency of James Wolfensohn (1995–2005), the World Bank embarked on a process to boost its own credibility and reputation. Wolfensohn announced a revitalization of the Bank's focus on poverty alleviation as well as a willingness for the Bank to become more "client-centered"—which meant more responsive to the demands of its borrowers (Mundy and Verger, 2015; Mundy, 2002). Also, and as noted in the prior section, the Bank started to transition toward the so-called Post-Washington Consensus (Pereira, 2016; Gore, 2000). According to Fine (2001), the Post-Washington Consensus approach is openly critical of the Washington Consensus but rejects the State versus markets dualism. In this sense, and as observed by Mundy (2002), the new approach should not be understood as a radical departure from the belief in the power of the market or open trade, but rather as "an effort to understand the failures of various assumptions and expectations within the Washington Consensus package, based on the idea that package should be seen as a flexible set of policy options" (p. 496).

As part of Wolfensohn's attempts at institutional transformation, Poverty Reduction Strategy Papers (PRSPs) were promoted as a replacement for structural adjustment lending. PRSPs were national planning frameworks required for countries wishing to qualify for external assistance or debt relief. They combined elements from the Washington Consensus with a greater concern for human and social development and governance issues, thus bringing poverty-alleviation concerns into sharper focus and being one of the clearest manifestations of the Post-Washington Consensus (Mundy, 2006; Tarabini and Jacovkis, 2012). Another central component of Wolfensohn's new vision was the Comprehensive Development Framework (CDF), which promoted a multidimensional approach to development. The rise of the CDF entailed a crucial procedural change, for it required the inclusion of (or a dialogue with) the civil society organizations, NGOs and beneficiaries of the World Bank's projects in the very design of such interventions—as a means to guarantee greater country ownership and context-sensitivity (Jones, 2007a; Bonal and Tarabini, 2009).

However, and as noted by different authors, the combination of PRSPs and CDF did not clearly translate into a change in the WB's operational practices and policy approach. In the area of education, Klees (2002) noted that, in practice, many PRSPs showed little adherence to the principles of participation and continued to pay limited attention to social protection policies. It was thus soon apparent that the PRSP and the PWC did not entail a radical departure from the neoliberal principles that had long dominated much of the WB's activity. As noted by Bonal (2002), by the early 2000s, it was far from clear that the PWC had succeeded in truly altering the principles underpinning the Bank's education policy. The author thus observed that "basic features of WB education policy, like the rates of return rationale for educational investments, the importance of private education, and the marginalization of vocational education and training remain unaltered" (p. 13).

In fact, the 1995–2000 period can be characterized as one of internal inconsistency in which the gulf between the WB's policy positions and lending practices grew considerably (Jones, 2007a). On the one hand, the education sector continued to advance a neoliberal agenda and simply tweaked its language in response to the rhetoric of responsiveness and flexibility promoted by the WB leadership. On the other hand, regional managers and operational staff continued to ignore the centrally-formulated policy designs as a means to avoid alienating their borrowers.

Importantly, this is a period in which the WB's dominant position in the education for development area (both as a donor and as an agenda-setter) suffered a rapid decline. Particularly during the late 1990s, the quality of the World Bank's analytic work decayed (Jones, 2007a). This progressive weakening of the WB's technical capacity in education cannot be dissociated from the staff shortages that affected the educational sector during the late 1990s because of organizational reforms and budget cuts (Mundy, 2002). In parallel, education lending activity slowed down because of the combination of the World Bank's inflexibility regarding priority areas and lending instruments, and increased competition on the part of multilateral development partners (Jones and Coleman, 2005).

Matters were compounded by the fact that the publication of a new *Education Sector Strategy* in 1999 only succeeded partially in giving the World Bank's education sector a sense of direction. Thus, while the 1999 strategy signaled the abandonment of the short education policy menu and continued to put a premium on flexibility and diversity as *procedural* values, no new policy priorities were clearly identified, thus giving way to something of a "policy vacuum" (Mundy, 2002; Heyneman, 2003). On the plus side, this policy vacuum and the collapse of the short education policy menu allowed for greater heterogeneity in terms of policy responses and policy research—as noted by Jones and Coleman (2005), this is a period in which the range of research questions was significantly expanded.

The early 2000s were characterized by a certain return to the coordination between operational and policy work, while leading to a greater definition of the World Bank's education policy agenda (Jones, 2007a). Between 2000 and 2005, the WB defined new positions in a number of key areas, including higher education and the private financing of education. Thus, the Bank became an explicit supporter of the abolition of school fees in primary education (Mundy and Menashy, 2014) and recognized the need for greater public spending on higher education after years of ambivalence—a shift justified by the centrality of university education in the context of the knowledge economy (Mundy and Verger, 2015), patent in the higher rates of return of tertiary education (Verger and Bonal, 2012). The new position was summarized in the *Education Sector Strategy Update: Achieving Education for All, Broadening our Perspective, Maximizing our Effectiveness*, framed as an update of the 1999 Education Sector Strategy (World Bank, 2005). Remarkably, much of the new position embraced by the new strategy followed the approach already favored by the regional division and lending operations.

One of the most distinctive features of the new strategy was the relative loss of centrality of the rates of return analysis (Verger and Bonal, 2012). The document thus marked the abandonment of the obligation for borrowers to prioritize certain education levels and the advance to a more expansive and flexible approach to education spending. While this represented a clear departure from the prescriptive approach that had prevailed in the past, the new strategy also featured elements of continuity—most notably, the

insistence on a narrow set of policy options closely connected to the neoliberal agenda and the dominance of neoclassic economic orthodoxies, such as privatization and decentralization (Mundy and Verger, 2015). Indeed, the principles of choice and competition only acquired greater prominence during the mid-2000s.

The World Bank after 2008: building a partial heterodoxy

In the aftermath of the 2008 financial global crisis, the World Bank experienced important transformations that influenced its discourse and strategical approach to the education sector. According to Mundy and Verger (2015), two main geopolitical and political factors have affected the Bank in this most recent period. First, the increasing role of emerging economies, such as China, India and Brazil, as development actors and donors. These countries recently moved from being borrowers of the World Bank to influential international donors, particularly among low-income countries in Africa. The increasing relevance acquired by these countries in the international arena coincided with a period in which the historically influential members of the World Bank were still struggling with the consequences of the global economic crisis (i.e., low growth rates and increasing national debt). The second factor that has influenced this period was the presidency of Jim Yong Kim, who promoted a new institutional strategy (see World Bank, 2013) and internal organizational reform. Kim reinforced the “knowledge turn” (Zapp, 2017) embraced by the Bank in the mid-1990s, promoted the role of the Bank in the dissemination of best practices, and undertook a significant internal organization restructuring to better respond to this role.

Regarding the World Bank approach to the education sector, the most relevant policy document of this period is the *Education Strategy 2020* (ES2020), released in 2011 (World Bank, 2011). As different authors have pointed out, while this new strategy involved discursive changes in relation to the previous strategical documents, the policy proposals continued to be aligned with the most traditional paradigms of the World Bank (Verger and Bonal, 2012; De Siqueira, 2012). According to Edwards and Storen (2017), most of the policy ideas advanced by the ES2020, such as standardization, national learning assessments or accountability mechanisms, have been promoted by the Bank since the 1990s.

Despite the continuity in terms of policy options promoted by the World Bank, the ES2020 contained some relevant discursive transformations, one of the most relevant being the focus and centrality of learning. Although the World Bank recognized that significant achievements had been made in the past two decades in expanding access to education, the new strategic document considered that the main challenge faced by national governments, particularly in low and middle-income countries, was ensuring “learning for all.” In this way, the Bank recognized that access to education was not enough to ensure that students would learn at school, and highlighted the need to ensure students maximize the level of learning achieved during schooling. Based on the revised human capital approaches to the relationship between education and economic growth (see, for instance, Hanushek and Woessmann, 2007), the new emphasis on learning over access was justified by the World Bank because “growth, development, and poverty reduction depend on the knowledge and skills that people acquire, not the number of years that they sit in a classroom” (World Bank, 2011, p. 3). For the World Bank, the focus on learning as the main objective of education policies required to overcome the input-driven reforms and promote policy options to set the necessary incentives to boost learning. However, as Verger and Bonal (2012) point out, despite the emphasis on market solutions and demand-side interventions, the new strategy is ambiguous regarding the specific policy options that promote learning.

In line with the 2013 World Bank Group strategy, the ES2020 also entailed a significant change in the role that the World Bank was supposed to play in the definition of education policies. The new strategy emphasized and reinforced the role of the Bank as a provider of knowledge and technical expertise rather than as a prescriber of specific policy options (Verger and Bonal, 2012). According to this, one of the main initiatives included in the ES2020 was the Systems Approach for Better Education Results (SABER), which became an evolved and upgraded version of the previous Education Knowledge Management System (EKMS) program included in the 1999 Education Strategy. This new World Bank’s project, launched in 2011, had as its main objective “to collect and disseminate comparative data and knowledge on education policies, to help countries systematically evaluate and strengthen their education systems” (Rogers and Demas, 2013, p. 10). This knowledge initiative covers so far 13 different domains, from early childhood education to tertiary education. According to authors such as Zapp (2017) or De Siqueira (2012), SABER pretended to reinforce the “knowledge turn” of the Bank, becoming both a diagnostic and benchmark tool for policies and educational institutions.

Another relevant characteristic of the new education strategy is the renewed emphasis on non-state actors and market solutions. Although the World Bank has been a historical advocate and active promoter of the involvement of private actors and market mechanisms in the education sector, the ES2020 can be considered a step forward in this regard (Mundy and Menashy, 2014; Verger and Bonal, 2012). In contrast with the initial approach to non-state actors’ involvement developed during the 1990s, in which private provision was considered as a mechanism to alleviate the pressures faced by the public sector in addressing the expansion of educational demand, the new strategy considered that private providers play an essential role in allowing the access of poor populations, “especially in areas that governments do not reach” (World Bank, 2011, p. 20). Indeed, in the past years, the Bank has emphasized the need to establish collaboration schemes between national governments and private actors, mainly through Public-Private Partnerships (PPPs) (Verger, 2012).

To characterize the most recent WB education sector approach, the 2018 *World Development Report* (WDR) is worth mentioning. This edition of the WDR has a main area of focus in education, and it was entitled *Learning to Realize Education’s Promise* (World Bank, 2018). Although in this document, the World Bank continued to stress the centrality of learning and, even, considered

that there was a “learning crisis,” the report supposed a significant discursive switch regarding education for different reasons. First, the report framed education as a basic human right that “is central to unblocking human capabilities” (World Bank, 2018, p. 38). Both the recognition of education as a human right and the reference to human capabilities represent an important departure from the traditional approaches to education characterized by an economic-instrumentalist perspective (Menashy, 2013). Second, for the first time, the World Bank acknowledges the existence of a solid body of evidence warning about the potential negative effects of private provision. Indeed, the report recognizes that, in many contexts, the superiority of the private schooling sector can be mainly explained by the fact that private schools skim off the most academically or socially advantaged students and exclude those from the most socially disadvantaged backgrounds. The 2018 WDR even affirms “there is no consistent evidence that private schools deliver better learning outcomes than public schools, or the opposite” (World Bank, 2018: 176). The World Bank also recognizes that the dramatic expansion of private provision in some national contexts “can undermine the political constituency for effective public schooling in the longer term” (World Bank, 2018, p. 177).

Future directions: which role of the World Bank in the global education arena?

Although the role and influence of the World Bank in the education sector have evolved in the past decades, the current influence of the World Bank in the definition of the global education agenda cannot be overstated. Not only does it remain the largest funder of education in the developing world, but it also holds an unmatched power in the definition of education problems and adequate policy solutions. The role of the World Bank in global goal-setting is illustrative of the ideational and normative dominance this organization enjoys. Thus, while the direct participation of the World Bank in goal-setting processes has been somewhat erratic, global education goals have been invariably influenced by the Bank’s analytic work. Likewise, the ultimate impact of global education goals is often determined by the Bank’s interpretation of such agendas.

An example of this hegemonic position is the role of the Bank in the lead-up to the WCEFA celebrated in Jomtien in 1990. Despite the World Bank’s initial reservations about the project, the organization eventually managed to repurpose the conference and EFA as an opportunity to persuade the international community of the need to prioritize primary education—in other words, as a means to vindicate basic education as a development priority for both donors and borrowing countries (Heyneman, 2009; Jones, 2007b). Moreover, while other agencies were less enthused by this approach, and the EFA agenda was eventually settled in favor of a more expansive vision, the central role of the Bank in the implementation of EFA meant that, in practice, access to primary education became the overriding EFA priority.

More recently, the negotiation of Sustainable Development Goal 4 (SDG4) was directly shaped by the “learning turn” promoted by the World Bank in the mid-2000s. The decoupling of schooling and learning emphasized in a range of World Bank’s publications ended up permeating the priorities of the major development agencies over the years to come. The prominence of the learning outcomes in the SDG4 debates cannot thus be understood without reference to the normative, research and dissemination labor led by the Bank during the previous years (Moriarty, 2019; Fontdevila, 2021). And while, as a result of civil society advocacy, the final list of targets does not feature learning as the main priority of the SDG4, over the last years, the Bank has made an effort to reinstate learning outcomes as the centerpiece of the global education agenda. For instance, in 2019, the World Bank launched the Learning Target, which aimed to “reduce by at least half the share of 10-year-olds who cannot read:” by 2030 (World Bank, 2019, p. 21). While the Learning Target was introduced as aligned with the SDG agenda, it was also framed as a more feasible goal. The new targets thus appear to be reinterpreting SDG4 by prioritizing a subset of objectives and nudging countries toward focusing their efforts on a specific goal. This is why, even if the Learning Target was framed as an effort to complement SDG4, the initiative has raised suspicion within the global education community, being perceived in certain quarters as an instance of counter-targeting (e.g., Archer, 2019).

The indirect influence of the Bank in the configuration of the SDG debate is indicative of the considerable power it retains in terms of agenda-setting and issue-definition. Yet the role of the Bank in the years to come is more of a question mark. The aid landscape has experienced a rapid transformation over the last decade—a process described by Babb and Chorev (2016) as “the loosening of the development regime,” according to which the emergence of new actors has given rise to a multipolar development arena while the material and ideational power of the World Bank and the IMF has been declining. More recently, the COVID-19 pandemic has only magnified the overall sense of volatility and uncertainty about the future of the Bretton Wood Institutions (Babb and Kentikelenis, 2021; Verbeek, 2021). In the global education arena, in particular, a crucial shift has been the emergence of new pooled funds (most notably, the Global Partnership for Education (GPE)³ and the Education Cannot Wait Fund—Menashy, 2019), which have effectively disputed the centrality of the Bank not only as a donor but also as a producer of education policy ideas. Coupled with the rise in power of new bilateral and private donors, this means that the Bank operates today in an increasingly crowded marketplace of ideas—one in which a growing number of organizations vie for the visibility and influence of their own policy proposals.

In this multipolar environment, the WB can no longer afford to operate as a *siló*—a sharply-bound, self-referential entity. Nowadays, the World Bank has a stronger incentive to cooperate and engage with other multilateral and bilateral agencies perceived as

³While the World Bank enjoyed a prominent position in the management and administration of the Fast-Track Initiative (GPE’s precursor), efforts have been made since 2010 efforts demarcate clearer boundaries between the GPE and the World Bank (Menashy, 2017).

strategically aligned or ideological affine—not only in the delivery of development projects, but also in its analytic or ideational labor. Over the last years, the WB has launched a wide range of projects developed in collaboration with different development partners—recent examples include the Global Education Finance Platform, which is oriented at the mobilization of resources for education and established in consultation with the OECD, the GPE, the IMF, UNESCO and the Education Commission; or the Global Education Evidence Advisory Panel, aimed at synthesizing evidence on effective education interventions. Such collaborative initiatives arguably reinforce the WB's legitimacy as an organization committed to cooperation and to preventing operational duplicities. Yet the ultimate impact of such partnerships on the leadership and authority of the Bank remains to be seen. As noted by *Kranke (2017)*, growing expectations for regularized collaboration between international organizations are difficult to reconcile with the need of these organizations to differentiate themselves from other agencies (i.e., preserving a distinct area of expertise or a singular policy and research agenda).

Another transformation that renders unpredictable the future of the World Bank is the growing level of heterodoxy of its policy menu and the diversification of its education ideas. The discursive shifts advanced by the 2018 WDR have only contributed to further blurring the World Bank's agenda. While the times of “short education policy menu” are definitely over, and the legacy of the Post-Washington Consensus is at best uncertain, they have not been substituted by a coherent paradigm and clear policy programs. The flexibilization and diversification of the World Bank policy agenda have been largely welcome within development circles, but it has also taken a toll in terms of institutional visibility and distinctiveness. This paradox of greater flexibility and lesser orthodoxy can potentially make the Bank equally present in educational development projects and lending but less influential in setting the direction of the global education policy agendas.

Understanding how these developments affect the World Bank's position as a “global governor” represents an important empirical challenge. This is so as the loss of the Bank's hegemony has been paralleled by a deceleration of the academic debate on this organization. Thus, the paucity of literature on the World Bank's last decade contrasts with the vibrant scholarship this institution generated during the 90s and early 2000s. While recent episodes like the 2018 WDR have been the object of some comment, this literature has tended to adopt a normative or evaluative approach. Conversely, there is more limited scholarship on the inner workings of the institution, or on the changes in the bureaucratic culture that might explain this discursive evolution. Making sense of the World Bank in a rapidly-evolving global arena requires advancing toward more fine-grained accounts of the micro-politics of this organization, and of the institutional conditions in which it operates.

References

- Archer, D., 2019. The World Bank Shouldn't Be Setting Global Goals for Education. Devex. Retrieved from: <https://www.devex.com/news/opinion-the-world-bank-shouldn-t-be-setting-global-goals-for-education-95898>.
- Babb, S., Chorev, N., 2016. International organizations: loose and tight coupling in the development regime. *Stud. Comp. Int. Dev.* 51 (1), 81–102.
- Babb, S., Kentikelenis, A., 2021. Markets everywhere: the Washington Consensus and the sociology of global institutional change. *Annu. Rev. Sociol.* 47, 521–541.
- Becker, G.S., 1974. Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education. National Bureau of Economic Research, New York.
- Bennell, P., 1996. Using and abusing rates of return. *Int. J. Educ. Dev.* 16, 235–248.
- Bonal, X., Tarabini, A., 2009. Global solutions for global poverty? The World Bank education policy and the anti-poverty agenda. In: Simons, M., Olsen, M., Peters, M.A. (Eds.), *Re-reading Education Policies. A Handbook Studying the Policy Agenda of the 21st Century*. Sense Publishers, Rotterdam, pp. 96–111.
- Bonal, X., 2002. Plus ça change ... the World Bank global education policy and the post-Washington Consensus. *Int. Stud. Sociol. Educ.* 12 (1), 3–22.
- Broad, R., 2007. “Knowledge management”: a case study of the World Bank's research department. *Dev. Pract.* 17 (4–5), 700–708.
- Carnoy, M., Torres, C.A., 1994. Educational change and structural adjustment: a case study of Costa Rica. In: Samoff, J. (Ed.), *Coping With Crisis. Austerity, Adjustment and Human Resources*. Cassell, London.
- Chabbot, C., 2003. *Constructing Education for Development. International Organizations and Education for All*. Routledge Falmer, London/New York.
- Cornia, G., Jolly, R., Stewart, F., 1997. *Adjustment With a Human Face*, vol. 1. Clarendon Press, Oxford.
- Dale, R., 1999. Specifying globalization effects on national policy: a focus on the mechanisms. *J. Educ. Pol.* 14 (1), 1–17.
- De Siqueira, A.C., 2012. The 2020 World Bank education strategy: nothing new, or the same old gospel. In: Klees, S.J., Samoff, J., Stromquist, N. (Eds.), *The World Bank and Education*. Sense Publishers, Rotterdam, pp. 67–81.
- Edwards, D., Storen, I., 2017. The World Bank and Educational Assistance. *Oxford Research Encyclopedia of Education*. Retrieved from: <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-43>.
- Fine, B., 2001. *Social Capital Versus Social Theory*. Routledge, London.
- Fontdevila, C., 2021. Global Governance as Promise-Making. Negotiating and Monitoring Learning Goals in the Time of SDGs. Doctoral Dissertation, Universitat Autònoma de Barcelona. Retrieved from: <https://www.tesisenred.net/handle/10803/672579#page=1>.
- Gore, C., 2000. The rise and fall of the Washington Consensus as a paradigm for developing countries. *World Dev.* 28, 789–804.
- Hanushek, E.A., Woessmann, L., 2007. The Role of Education Quality for Economic Growth. Working Paper; No. 4122. World Bank, Washington, D.C.
- Heyneman, S.P., 2003. The history and problems in the making of education policy at the World Bank 1960–2000. *Int. J. Educ. Dev.* 23 (3), 315–337.
- Heyneman, S.P., 2009. The failure of education for all as political strategy. *Prospects* 39 (1), 5–10.
- Ilon, L., 1996. The changing role of the World Bank: education policy as global welfare. *Pol. Polit.* 24, 413–424.
- Jones, P.W., Coleman, D., 2005. *The United Nations and Education. Multilateralism, Development and Globalisation*. Routledge Falmer, London/New York.
- Jones, P.W., 2007a. *World Bank Financing of Education. Lending, Learning and Development*, second ed. Routledge, London, New York.
- Jones, P.W., 2007b. WCEFA: a moment in the history of multilateral education. *Int. Perspect. Educ. Soc.* 8, 521–538.
- Klees, S.J., 2002. World Bank education policy: new rhetoric, old ideology. *Int. J. Educ. Dev.* 22 (5), 451–474.
- Kranke, M., 2017. The Politics of Collaborative Global Governance: Organisational Positioning in IMF-World Bank Collaboration. Doctoral Thesis, University of Warwick. Retrieved from: <http://wrap.warwick.ac.uk/101269/>.
- Lauglo, J., 1996. Banking on education and the uses of research. A critique of World Bank priorities and strategies for education. *Int. J. Educ. Dev.* 16 (3), 221–233.
- Menashy, F., 2013. Interrogating an omission: the absence of a rights-based approach to education in World Bank policy discourse. *Discourse* 34 (5), 749–764.
- Menashy, F., 2017. The limits of multistakeholder governance: the case of the global partnership for education and private schooling. *Comp. Educ. Rev.* 61 (2), 240–268.

- Menashy, F., 2019. *International Aid to Education. Power Dynamics in the Era of Partnership*. Teachers College Press, New York.
- Moriarty, K., 2019. *Developing a Transformative Vision of Global Education? Unpacking Education Quality and Learning in the Policy Formulation and Content of Sustainable Development Goal 4*. Doctoral Dissertation, University of Sussex. Retrieved from. <http://sro.sussex.ac.uk/id/eprint/85462/>.
- Morrow, R.A., Torres, C.A., 2000. The State, globalization and educational policy. In: Burbules, N.C., Torres, C.A. (Eds.), *Globalization and Education. Critical Perspectives*. Routledge, London.
- Mundy, K., Menashy, F., 2014. The World Bank and private provision of schooling: a look through the lens of sociological theories of organizational hypocrisy. *Comp. Educ. Rev.* 58 (3), 401–427.
- Mundy, K., Verger, A., 2015. The World Bank and the global governance of education in a changing world order. *Int. J. Educ. Dev.* 40, 9–18.
- Mundy, K., 1999. Educational multilateralism in a changing world order: UNESCO and the limits of the possible. *Int. J. Educ. Dev.* 19 (1), 27–52.
- Mundy, K., 2002. Retrospect and prospect: education in a reforming World Bank. *Int. J. Educ. Dev.* 22 (5), 483–508.
- Mundy, K., 2006. Education for all and the new development compact. *Rev. Educ.* 53, 23–48.
- Pereira, J.M.M., 2016. Recycling and expansion: an analysis of the World Bank agenda (1989–2014). *Third World Q.* 37 (5), 818–839.
- Psacharopoulos, G., Woodhall, M., 1985. *Education for Development: An Analysis of Investment Choices*. The World Bank, Washington, D.C.
- Reimers, F., Tiburcio, L., 1993. *Education, Adjustment and Reconstruction: Options for Change*. UNESCO, Paris.
- Rogers, H., Demas, A., 2013. The What, Why, and How of the Systems Approach for Better Education Results (SABER). World Bank Group, Washington, D.C.. <http://documents.worldbank.org/curated/en/867151468180272110/The-what-why-and-how-of-the-Systems-Approach-for-Better-Education-Results-SABER>
- Samoff, J., 1996. Which priorities and strategies for education? *Int. J. Educ. Dev.* 16, 1–27.
- Stiglitz, J., 1998. More Instruments and Broader Goals: Moving Toward the Post-Washington Consensus. WIDER Annual Lecture. WIDER, Helsinki.
- Stiglitz, J., 2003. Challenging the Washington Consensus. *Brown J. World Aff.* 9 (2), 33–40.
- Tarabini, A., Jacovkis, J., 2012. The poverty reduction strategy papers: an analysis of a hegemonic link between education and poverty. *Int. J. Educ. Dev.* 32 (4), 507–516.
- Verbeek, N., 2021. The future of IMF and the World Bank: new dynamics in times of COVID-19 crisis, 2020. In: Hosli, M.O., Garrett, T., Niedecken, S., Verbeek, N. (Eds.), *The Future of Multilateralism: Global Cooperation and International Organizations*. Rowman & Littlefield, London, pp. 73–91.
- Verger, A., 2012. Framing and selling global education policy: the promotion of Public–Private Partnerships for education in low-income contexts. *J. Educ. Policy* 27 (1), 109–130.
- Verger, A., Bonal, X., 2012. “All things being equal”? Policy options, shortfalls and absences in the World Bank education sector strategy 2020. In: Klees, S.J., Samoff, J., Stromquist, N.P. (Eds.), *The World Bank and Education: Critiques and Alternatives*. Springer Science & Business Media.
- Vetterlein, A., 2007. Economic growth, poverty reduction, and the role of social policies: the evolution of the World Bank’s social development approach. *Glob. Gov.* 13 (4), 513–533.
- Wade, R., 1996. Japan, the World Bank and the art of paradigm maintenance: the East Asian miracle in political perspective. *N. Left Rev.* 217, 3–36.
- World Bank, 1980. *Education Sector Policy Paper*. The World Bank, Washington D.C.
- World Bank, 1995. *Priorities and Strategies for Education. A World Bank Review*. The World Bank, Washington, D.C.
- World Bank, 1999. *Education Sector Strategy*. The World Bank, Washington, D.C.
- World Bank, 2005. *Education Sector Strategy Update: Achieving Education for All, Broadening Our Perspective, Maximizing Our Effectiveness*. The World Bank, Washington, D.C.
- World Bank, 2011. *Learning for All: Investing in People’s Knowledge and Skills to Promote Development*. The World Bank, Washington, D.C.
- World Bank, 2013. *World Bank Group Strategy*. The World Bank, Washington, D.C.
- World Bank, 2018. *World Development Report 2018: Learning to Realize Education Promise*. The World Bank, Washington, DC.
- World Bank, 2019. *Learning Poverty: What Will It Take?* The World Bank, Washington, D.C.
- Zapp, M., 2017. The World Bank and education: governing (through) knowledge. *Int. J. Educ. Dev.* 53, 1–11.

UNICEF: self-transformation in a reconfigured global childhood space?

Jennifer Guevara and Mathias Urban, Dublin City University, Dublin, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Education for the new world order	488
The construction of a global children's policy agency	489
Organisational configurations in times of change	490
The construction of a global (early) childhood education space	490
UNICEF and the politics of global childhood	491
Conclusion	491
References	492

UNICEF—the United Nations Children's Fund—is the United Nations' agency tasked with humanitarian and developmental aid to children worldwide. It was established in 1946, in the aftermath of World War II as United Nations International Children's Emergency Fund. Since 1953, UNICEF has been part of the United Nations System consisting of the UN's principal organs,¹ specialised agencies² and related organisations.³ Today UNICEF is one of the most prominent global welfare organisations, with a presence in 192 countries and territories. UNICEF's activities include providing immunizations and disease prevention, administering treatment for children and mothers with HIV, enhancing childhood and maternal nutrition, improving sanitation, promoting education, and providing emergency relief in response to disasters.

The United Nations succeeded a previous intergovernmental body dedicated to global coordination and the pursuit of world peace in the aftermath of World War I—the League of Nations (1920–46). Together with its extended system of organisations and agencies, the UN was conceived as central part of a multilateral system to create and maintain a 'new world order', promote social progress and advance human rights. UNICEF is an indispensable part of this multilateral system and 'new world order'. Seen in this context, UNICEF and its activities have been affected by the current crisis of multilateralism as the 'new world order' of the second half of the 20th century and its legitimacy have become questionable. We discuss some of the implications, opportunities and challenges for UNICEF's mandate in the concluding section of this article.

Education for the new world order

There are four UN agencies that have historically been more committed to the development of international educational policy-making—the United Nations Educational, Scientific and Cultural Organisation (UNESCO), the World Bank, the United Nations Children's Fund (UNICEF) and the United Nations Development Program (UNDP). While UNESCO was the first to start supporting education in new nation states, by the 1960s all UN bodies were involved with educational development in nation states to some extent (Jones, 2006).

Today, three of these bodies are considered key players in the education field: UNESCO, UNICEF and the World Bank (Mahon, 2016; Menashy and Manion, 2016; Mundy et al., 2016). Among them, UNESCO and the World Bank received detailed attention in the literature (Edwards et al., 2018; Jones, 1990; Klees et al., 2012; Mundy and Verger, 2015). Although less explored in the academic world, UNICEF is arguably the best-known agency in the UN system amid the general public. Born as a humanitarian organisation, UNICEF transformed into a development agency (Beigbeder, 2001b). Due to its practical operations, in-country presence and its ability to respond to emergencies, UNICEF enjoys an overwhelmingly positive image in the public perception, especially when compared to the other UN bodies (Jones, 2006). It also occupies a prominent place across education multilateral organisations, where it challenged more economic approaches to education policy promoted, for example, by the World Bank.

Although UNICEF's main contribution has been 'on the ground' (Mahon, 2016), in recent years the agency has also played a major role as a global development agency and, more recently, a children's rights advocate. This chapter analyses the role of UNICEF as a global children's policy agency and its role in the construction of a global (early) childhood education space. We review the

¹General Assembly, Security Council, Economic and Social Council (ECOSOC), Trusteeship Council, International Court of Justice (ICJ), and the UN Secretariat.

²Food and Agriculture Organization (FAO), International Civil Aviation Organization (ICAO), International Fund for Agricultural Development (IFAD), International Labor Organization (ILO), International Maritime Organization (IMO), International Monetary Fund (IMF), International Telecommunication Union (ITU), United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Industrial Development Organization (UNIDO), Universal Postal Union (UPU), World Bank Group (WBG), World Health Organization (WHO), World Intellectual Property Organization (WIPO), World Meteorological Organization (WMO), World Tourism Organization (UNWTO), International Refugee Organization (IRO); ceased to exist in 1952.

³International Organization for Migration (IOM), Comprehensive Nuclear-Test-Ban Treaty Organization Preparatory Commission (CTBT PrepCom), International Atomic Energy Agency (IAEA), Organisation for the Prohibition of Chemical Weapons (OPCW), World Trade Organization (WTO).

process by which UNICEF evolves from a temporary humanitarian relief body to a children's rights advocate and a global player in (early) childhood education in the post war 'new world order'. We conclude with some thoughts around UNICEF's role in the emerging multi-polar 'new' world order of the 21st century.

The construction of a global children's policy agency

Initially a temporary agency of the UN focused on survival of marginalised children, UNICEF transitioned to an established and large organisation with a strong mandate 'to advocate for the protection of children's rights, to help meet their basic needs and to expand their opportunities to reach their full potential'. In the turn of the 1970s the agency started moving away from humanitarian relief and transforming into a 'fully-fledged member' of the global development agencies club (Black, 1996). According to Shusterman (2021) UNICEF did not lose focus on rapid relief, but rather shifted its footing on the relief–development continuum. UNICEF redefined itself in its role, where 'emergency action for children and women [became] a limited but significant part of its overall mandate' (UNICEF, 1996, p. 1). In the agency's identity and discourse, this dual mandate is presented as a comparative advantage, where UNICEF mobilises its presence not only during but also before and after emergencies and crises.

In the late 1990s, in a third key transition, UNICEF transformed itself into children's rights advocacy organisation. Following the adoption of the Convention on the Rights of the Child (CRC) in 1990, UNICEF started to firmly bring an emphasis on human rights as a program driver in education. However, the declaration was the achievement of a movement of non-governmental organisations (NGOs), mostly based in Geneva, in which the agency had played little part. To UNICEF, Human rights were a controversial field where the organisation's involvement could jeopardise its relationships—mainly its governmental network, since cooperation with national children programmes had become its main activity (Black, 1996). Yet, the General Assembly 'endorsed UNICEF as a practical vehicle for implementing the declaration's principles, thereby implicitly expanding UNICEF's role beyond relief and general child' (Beigbeder, 2001a: 22).

UNICEF was not explicitly included in the Declaration, but anonymously referred to in the preamble as part of the 'specialized agencies and international organizations concerned with the welfare of children whose statutes include special safeguards for children' (Assembly, 1989, p. Preamble). Other specialised agencies were not interested in making human rights the centrepiece of their agenda. The World Bank, for example, historically resisted that effort (Klees and Qargha, 2014).

While the responsibility for realising the CRC lies with states, UNICEF is to encourage its effective implementation and promote international cooperation related to the Convention (Beigbeder, 2001b). The agency's role has been to pursue the CRC's universal ratification, publicise and advocate for its implementation. Considering that the CRC is an aspirational and ambiguous document the organisation also plays a role in the interpretation of what it means to protect, provide, and promote children's rights in each local context.

The consolidation of UNICEF's role as a children's rights advocate did not mean losing its focus on either humanitarian relief or development. As Coleman and Jones (2004, p. 245) argue, UNICEF's 'key transitions – from a humanitarian focus to being a development agency and then on to human rights advocacy – have seen more of a mingling of these than the serial displacement of one by another'. UNICEF's trifold identity maybe ambiguous, or even contradictory, but it certainly functions as a strength for the organisation, helping it to keep a large base of support among the general public.

There has been, however, some vocal opposition to UNICEF's adaptation of a broad children's rights agenda. To give just one example, in 2004, the right-wing Oakland (US)-based think tank The Independent Institute (declared mission: '[...] to boldly advance peaceful, prosperous, and free societies grounded in a commitment to human worth and dignity.') published strongly worded criticism of what it terms 'UNICEF's implementation of its "children's rights" vision', especially in relation to UNICEF's focus on the rights of girls and women in family and society: 'Indeed, UNICEF's Medium Term Strategic Plan is more of a blueprint for social engineering along radical feminist lines' (<https://www.independent.org/news/article.asp?id=1443>). This is a clearly tendentious statement, aiming to undermine international advocacy for children's rights, considering it is coming from a right-wing organisation in the only country that has persistently denied ratification of the Convention on the Rights of the Child. It is questionable, too, as a meta-evaluation of UNICEF program evaluations by the OECD's Development Assistance Committee found that 'only one-third of the evaluations were able to demonstrate that UNICEF-supported programs effectively address gender equality' (OECD, 2013, p. 62). The authors conclude that 'UNICEF's performance with respect to gender equality is a serious concern' (ibid).

Some argue that, in the last decade, UNICEF is undergoing a fourth transition from a human rights approach into an equity focus, based on the argument that 'human rights cannot be realized so long as inequality persists' (UNICEF 2010, p. 19 cited in Klees and Qargha, 2014). This 'refocus' on equity can also be seen in the rationale of the Global Partnership for Education (GPE) and Education Cannot Wait (ECW). Founded in 2002, GPE is the largest multistakeholder agency in the sector; it is chaired by the World Bank and coordinated by UNICEF at the national level (Silva and Oliveira, 2021). ECW is a more recent multistakeholder global pooled fund that UNICEF has hosted since its incubation (Menashy, 2019). Some have questioned UNICEF's role in these partnerships, but specially in ECW, since the organisation may be both a funder and a potential fund-seeker on the ground (Menashy, 2019). In any case, this illustrates UNICEF has taken up a role in the relatively new trend of multistakeholder partnerships in international development.

Organisational configurations in times of change

With its particular history and tri-faceted role, UNICEF's organisational structure fundamentally differs from the rest of the UN bodies. While the World Bank and UNESCO are considered autonomous 'specialised agencies', UNICEF is accountable to the UN General Assembly (Beigbeder, 2001b; Menashy and Manion, 2016). Also, unlike other UN bodies, UNICEF's budget is met in full through voluntary contributions from different actors, including public national and supranational organisations (governments, European Commission, inter-organizational arrangements, global program partnerships and international financial institutions) (Mundy et al., 2016). A significant amount of the organisation's revenue comes from contributions of member states, with the United States of America, Germany, and the United Kingdom being the two largest contributors, which increases their decision-making power and influence (Mundy et al., 2016).

Private organisations and the general public are UNICEF's other contributors. This funding model has profound implications for the way the organisation operates. UNICEF is a donor-driven organisation that continuously needs to demonstrate and advertise its work not only to inform donors and the public about its work but, more importantly, 'sending public messages on its policies in an effort to influence national and international decision makers and obtain the support of public opinion for policies' (Beigbeder, 2001b, p. 145). Associated with this is fundraising, a central characteristic of UNICEF's identity, which brings it closer to how non-profit organisations operate.

As Mundy et al. (2016) argues, these features contribute to an organisational structure that is able to elude the criticism that has fallen on other UN bodies, such as UNESCO and the World Bank. Hence, compared to other UN agencies, UNICEF has a stronger legitimacy to advocate for children's rights. UNICEF's advocacy has consistently corresponded with a strong presence 'on the ground' and its presence in shaping the global (early) childhood agenda.

The construction of a global (early) childhood education space

UNICEF's initial agenda did not include education, except from maternal education programmes (Beigbeder, 2001a). The organisation struggled for decades to find its role and own positioning in the education field. In the new century UNICEF has been able to solidify its role as a major global player (Jones, 2006). In a rapidly expanding and diversifying network of development programs, the agency saw education as an area where it could position itself (Coleman and Jones, 2004). Its education advocacy started with an emphasis on informal education, rather than formal schooling. UNICEF's interests expanded progressively to most areas of education, from primary education and the education of girls to early childhood education and care (ECEC).

The concept of early childhood has entered UNICEF and other global players' agendas broadly since the beginning of the 21st century. That means it is not long ago that many international organisations, as well as national governments assumed young children to be primarily a domestic matter, with only very limited public or state responsibility for young children's care and education. Rather than providing education for all young children from a young age as a common good, early childhood services were often seen as necessary intervention in 'dysfunctional' families.

In many cases, state governments remain reluctant to enter the early childhood space for purposes other than general legislation and regulation, and services are often provided by non- and semi-state actors (e.g., charitable, community, or welfare organisations) and to a large extent by private-for-profit providers. A broad acceptance of the concept of ECEC helped promote the education and care of young children become to a global issue. This acknowledgment of ECEC and its expansion can be traced back to the rise of recent knowledge claims around young children (Mahon, 2016), including: the neuroscientific development of young children's brains; their economic potential as human capital and the return of investment on young children; the investment in ECEC as the 'greatest equaliser' of socioeconomic and gender inequalities; and the rights discourse, among other traveling ideas (Morabito et al., 2013). As in most international agencies, UNICEF's international webpage on early childhood reads

Children's brains are built, moment by moment, as they interact with their environments. In the first few years of life, more than one million neural connections are formed each second – a pace never repeated again. The quality of a child's early experiences makes a critical difference as their brains develop, providing either strong or weak foundations for learning, health and behaviour throughout life (<https://www.unicef.org/early-childhood-development>)

All these lines of argument are highly contested and are subject to interrogation, especially by scholars grounded in critical theory, critical pedagogy, anti-colonialist scholarship and indigenous epistemologies (e.g., Apple, 2003; Bloch et al., 2018; Brayboy, 2005; Chiara, 2020; Gupta, 2006; Nsamenang, 2004; Pence and Marfo, 2008; Steiner-Khamisi et al., 2018). Nonetheless, they have provided powerful arguments to increase attention and investment in early childhood education and care.

The discourse—and its critique—condense in a central concept that continues to drive international and national early childhood policy debate: *quality*. In itself a highly contested term, the quest for *quality* has become the most influential driver of early childhood policy globally. Its efficacy has been compared to 'the irresistibility of a steamroller' by German author Burkhard Müller in an early analysis: steamrollers are powerful, effective—but tend to flatten everything in their way (Müller, 1996, p. 8, our translation). It has led early childhood advocates into a dilemma that appears impossible to resolve:

Indeed, for those of us who value ECE but dislike the story of quality and high returns, the awkward fact is that the contemporary expansion of early childhood education would not have happened without this discourse gaining dominance.

(Moss, 2017, p. 14).

The core of the critique of the discourse of quality is that it renders a value-bound and hence contested and profoundly political social practice—the education of young children—as a practice of applied technology, certainty and, if done right, one of (pre)determinable outcomes. In the international discourse, *quality* has assumed a hegemonic position, has become unquestionable and, as such, amounts to early childhood education policy's very own regime—its *general politics*—of truth (Foucault and Rabinow, 1991; Urban, 2022 in preparation).

UNICEF and the politics of global childhood

With its global recognition and its programmes bridging the global south and north, it is not surprising that UNICEF activities and strategies are affected by—and contributing to—the dilemma of the construction of a global imaginary of children, childhood and early childhood education as captured in the debate about *quality* and other dominant paradigms. This becomes apparent, for instance, in anti-colonialist critiques of UNICEF and other global development operators' depiction of children in the global south in order to raise funds from donors in the global north. A recent analysis of UNICEF by the activist website karmacolonialism.org found 1390 references to the 'needy child' on UNICEF webpages, with the term 'twice as likely to be used on a page for Africa' than on pages referring to global north countries, and the 'needy' child in question overwhelmingly a girl (Alyson, 2021).

However, UNICEF has become an increasingly vocal advocate in the area of national early childhood policy development, policy monitoring, and accountability. Arguably, UNICEF's influence on national policies can be seen on a level similar only to the OECD's early childhood 'Starting Strong' policy initiatives (OECD, 2001, 2006, 2009, 2010, 2012; 2015a; 2015b, 2017). UNICEF's exercises its ECEC policy advocacy mainly through the UNICEF Innocenti Research Center, based in Florence, Italy. The Innocenti Center (officially: UNICEF Office of Research – Innocenti) was established in 1988 to 'strengthen the research capability of the United Nations Children's Fund and to support its advocacy for children worldwide'. According to its mission statement the center 'generates knowledge and analysis to support policy formulation and advocacy in favor of children; acts as a convener and catalyst for knowledge exchange and strategic reflections on children's concerns; and supports program development and capacity-building' (www.unicef-irc.org). One of the center's most recognised tools for policy advocacy are its 'report cards'—short, focused publications that measure countries' progress against declared early childhood policy aims. Unlike UNICEF's appeals for support of its programmes in the global south, these documents have targeted 'economically advanced' and 'rich' countries. Examples include the widely cited *Report Cards 8* and *16*: 'The child care transition. A league table of early childhood education and care in economically advanced countries' and 'Worlds of Influence: Understanding What Shapes Child Well-being in Rich Countries' (UNICEF Innocenti Research Center, 2008, 2020).

The OECD has been criticised for exerting significant but unaccountable 'soft power' on national education systems (Auld and Morris, 2016; Auld et al., 2019; Kaurel, 2020; Morris, 2016). UNICEF, arguably, has entered the very same space, albeit from a decisively children's rights-based perspective. With its global reach and visibility, UNICEF's approach is distinct from that of another influential 'global player' in the field of early childhood policy: the World Bank and its regional arms. Compared to UNICEF they have considerable 'hard power' at their disposal to shape countries' education policies through funding and loan arrangements (Klees et al., 2012; Steiner-Khamsi, 2013; Steiner-Khamsi et al., 2018).

Conclusion

In only six decades, UNICEF has evolved from a UN temporary humanitarian relief body to a global development agency, a children's rights advocate, and a driver of multistakeholder partnerships. As shown above, each of UNICEF's transitions did not imply a complete transformation, but rather a new dimension of the organisation's identity being developed. UNICEF continues to play a role in crises, while also implementing a global agenda on children's policies, all under the umbrella of children's rights. Its multifaceted role, we argued, is one of the keys to its success in being able to take part in the international organisation's 'club' but maintain an overwhelming positive image among the general public.

The involvement of UNICEF in education is quite recent. As argued in this chapter, the organisation struggled for years to find its place in the global education field. In fact, its approach to education did not solidify until the organisation undertook the mandate of promoting the UNCRC. Today, UNICEF is one of the main partners in the largest multistakeholder partnerships that promote education, and young children's rights. We have argued that UNICEF, with its different arms and strategies, exercises a significant but unaccountable amount of 'soft power', which may be rendered invisible since it comes from a decisively children's rights-based perspective.

The crucial question for UNICEF will be how and to what extent it will be able to use its 'soft power' in the dramatically changed context of the 21st century. Today, the 'new world order' that brought about UNICEF and its sister agencies in the UN space in the second half of the 20th century is fractured (at best). Its legitimacy has become questionable not least by the failure of UN

institutions to provide effective global governance, representing the interests of all of humanity as it faces a convergence of existential crises: climate, biodiversity, pandemic, democracy, equity. The apparent failure of global governance stands in contrast to UNICEF's efficacy 'on the ground', e.g., in providing concrete support for children through its programmes.

The litmus test, we would argue, will be whether UNICEF will be able to reposition itself as the boundaries between 'developed' and 'developing' countries, between global south and north, are increasingly blurred. This is particularly relevant in relation to young children, who as UNICEF warns are about to become the 'silent victims' of the globally entangled crises scenario (UNICEF, 2020). Where one UNICEF could safely focus on marginalised children in the global south, it is becoming apparent that, as Santos and other anti-colonialist scholars argue, the global south is no longer a geographical location (Chiara, 2020; Santos, 2018). Locating the new 'global south' has become a largely geo-political question, as the *terra firma* of 'us' and 'them' turns into a global archipelago of islands of deprivation, marginalisation and exclusion in a rising sea of inequality.

UNICEF underwent most of its transitions in the context of a post war 'new world order'. Its ground operations and its rights-based approach have helped the agency construct itself as a legitimate advocate for children. At a crossroads of multiple global crises, and with the world shifting toward multi-polar 'new' world (dis)order of the 21st century, will UNICEF grow into a possible new role as one of the legitimate agents of much needed global multi-lateral coordination?

References

- Allyson, S., 2021. UNICEF Needs the Needy. Retrieved from: <https://karmacolonialism.org/unicef-needs-the-needy/>.
- Apple, M., 2003. Is the new technology part of the solution or part of the problem in education? In: Darder, A., Baltodano, M., Torres, R.D. (Eds.), *The Critical Pedagogy Reader*. Routledge, New York, pp. 440–457.
- Assembly, U.G., 1989. Convention on the rights of the child. United Nations, Treaty Series 1577 (3), 1–23.
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education 'best practice'. *Comp. Educ.* 52 (2), 202–229.
- Auld, E., Rappleye, J., Morris, P., 2019. PISA for Development: how the OECD and World Bank shaped education governance post-2015. *Comp. Educ.* 55 (2), 197–219. <https://doi.org/10.1080/03050068.2018.1538635>.
- Beigbeder, Y., 2001a. The creation of UNICEF. In: *New Challenges for UNICEF*. Springer, pp. 1–17.
- Beigbeder, Y., 2001b. *New Challenges for UNICEF: Children, Women and Human Rights*. Springer.
- Black, M., 1996. *Children First*. Macmillan, Australasia.
- Bloch, M., Swadener, B.B., Cannella, G.S. (Eds.), 2018. *Reconceptualizing Early Childhood Education and Care: A Reader*. Critical Questions, New Imaginaries & Social Activism, second ed. Peter Lang Publishing, New York.
- Brayboy, B.M.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37 (5), 425–446. <https://doi.org/10.1007/s11256-005-0018-y>.
- Chiara, D., 2020. Global south childhoods. In: Cook, D.T. (Ed.), *The SAGE Encyclopedia of Children and Childhood Studies*. SAGE Publications, Inc, Thousand Oaks, California, pp. 867–870.
- Coleman, D., Jones, P.W., 2004. *The United Nations and Education: Multilateralism, Development and Globalisation*. Taylor & Francis Group, Abingdon, Oxon, UNITED STATES.
- Edwards Jr., D.B., Okitsu, T., da Costa, R., Kitamura, Y., 2018. Organizational legitimacy in the global education policy field: learning from UNESCO and the Global Monitoring Report. *Comp. Educ. Rev.* 62 (1), 31–63.
- Foucault, M., Rabinow, P., 1991. *The Foucault Reader*. Penguin, London.
- Gupta, A., 2006. Early Childhood Education, Postcolonial Theory, and Teaching Practices in India: Balancing Vygotsky and the Veda, first ed. Palgrave Macmillan, Basingstoke.
- Jones, P.W., 1990. UNESCO and the politics of global literacy. *Comp. Educ. Rev.* 34 (1), 41–60.
- Jones, P.W., 2006. Elusive mandate: UNICEF and educational development. *Int. J. Educ. Dev.* 26 (6), 591–604.
- Kaurel, J., 2020. From Inspiration to Uniformity? 20 Years of OECD in the Field of Early Childhood Education and Care, 2/2020. Union of Education Norway, Oslo.
- Klees, S.J., Qargha, O., 2014. Equity in education: the case of UNICEF and the need for participative debate. *Prospects* 44 (3), 321–333. <https://doi.org/10.1007/s11125-014-9295-0>.
- Klees, S.J., Smoff, J., Stromquist, N.P. (Eds.), 2012. *The World Bank and Education. Critiques and Alternatives*. Sense Publishers, Rotterdam, Boston, Taipei.
- Mahon, R., 2016. Early childhood education and care in global discourses. *Handbook Glob. Educ. Policy* 224–240.
- Menashy, F., 2019. *International Aid to Education: Power Dynamics in an Era of Partnership*. Teachers College Press.
- Menashy, F., Manion, C., 2016. The historical evolution and current challenges of the United Nations and global education policy-making. *Handbook Glob. Educ. Policy* 319–334.
- Morabito, C., Vandenbroeck, M., Roose, R., 2013. 'The Greatest of Equalisers': a critical review of international organisations' views on early childhood care and education. *J. Soc. Pol.* 42 (3), 451–467.
- Morris, P., 2016. *Education Policy, Cross-National Tests of Pupil Achievement, and the Pursuit of World-Class Schooling*. UCL Institute of Education Press, London.
- Moss, P., 2017. Power and resistance in early childhood education: from dominant discourse to democratic experimentalism. *J. Pedagogy* 8 (1), 11. <https://doi.org/10.1515/jped-2017-0001>.
- Müller, B., 1996. *Qualitätsprodukt Jugendhilfe. Kritische Thesen und praktische Vorschläge*. Lambertus, Freiburg im Breisgau.
- Mundy, K., Green, A., Lingard, B., Verger, A., 2016. Introduction: The Globalization of Education Policy-Key Approaches and Debates.
- Mundy, K., Verger, A., 2015. The World Bank and the global governance of education in a changing world order. *Int. J. Educ. Dev.* 40, 9–18.
- Nsamenang, B.A., 2004. *Cultures of Human Development and Education: Challenge to Growing up African*. Nova Science Publishers, New York.
- OECD, 2001. *Starting Strong*. Early Childhood Education and Care. OECD, Paris.
- OECD, 2006. *Starting Strong II*. Early Childhood Education and Care. OECD, Paris.
- OECD, 2009. *Doing Better for Children*. OECD, Paris.
- OECD, 2010. *Enhancing Quality in Early Childhood Education and Care (ECEC)*. Retrieved from: http://www.oecd.org/document/0/0,3746,en_2649_39263231_45149440_1_1_1_1,00.html.
- OECD, 2012. *Starting Strong III - A Quality Toolbox for Early Childhood Education and Care*. OECD, Paris.
- OECD, 2013. *Review of UNICEF's Development Effectiveness: Final Report 2009 – 2011*. Retrieved from Paris. <https://www.oecd.org/dac/evaluation/UNICEF%20DE%20Review%20Report%20final%20May%2015.pdf>.
- OECD, 2015a. *Starting Strong IV*.
- OECD, 2015b. *Starting Strong IV. Monitoring Quality in Early Childhood Education and Care*. OECD, Paris.
- OECD, 2017. *Starting Strong V*.
- Pence, A.R., Marfo, K., 2008. Early childhood development in Africa: interrogating constraints of prevailing knowledge bases. *Int. J. Psychol.* 43 (2), 78–87.
- Santos, B.d.S., 2018. *The End of the Cognitive Empire: The Coming of Age of Epistemologies of the South*. Duke University Press, Durham.

- Shusterman, J., 2021. Gap or prehistoric monster? A history of the humanitarian–development nexus at UNICEF. *Disasters* 45 (2), 355–377. <https://doi.org/10.1111/disa.12427>.
- Silva, R.d., Oliveira, J., 2021. Global education policy in african fragile and conflict-affected states: examining the global partnership for education. *Glob. Soc. Educ.* 1–15.
- Steiner-Khamsi, G., 2013. What is wrong with the 'what-went-right' approach in educational policy? *Eur. Educ. Res. J.* 12 (1), 20–33. <https://doi.org/10.2304/eeerj.2013.12.1.20>.
- Steiner-Khamsi, G., Appleton, M., Vellani, S., 2018. Understanding business interests in international large-scale student assessments: a media analysis of *The Economist*, *Financial Times*, and *Wall Street Journal*. *Oxf. Rev. Educ.* 44 (2), 190–203. <https://doi.org/10.1080/03054985.2017.1379383>.
- UNICEF, 1996. *Children and Women in Emergencies: Strategic Priorities and Operational Concerns for UNICEF*. Retrieved from, New York.
- UNICEF, 2020. Don't Let Children Be the Hidden Victims of COVID-19 Pandemic. Retrieved from. <https://www.unicef.org/press-releases/dont-let-children-be-hidden-victims-covid-19-pandemic>.
- UNICEF Innocenti Research Centre, 2008. Report Card 8. *The Child Care Transition. A League Table of Early Childhood Education and Care in Economically Advanced Countries* (Retrieved from Florence).
- UNICEF Innocenti Research Centre, 2020. Report Card 16. *Worlds of Influence: Understanding what Shapes Child Well-Being in Rich Countries*. UNICEF Office of Research - Innocenti, Florence.
- Urban, M., 2022. *Scholarship in Times of Crises: Towards a Trans-discipline of Early Childhood*. Comparative Education.

European Union

Marcella Milana, University of Verona, Department of Human Sciences, Verona, Italy

© 2023 Elsevier Ltd. All rights reserved.

Introduction	494
Supranational integration, Europeanization, and education policy	494
The EU political system and education	496
European governance, policy coordination, and education	497
Expert groups and the network governance of (European) education	498
Policy coordination under the European Semester	500
Conclusion	501
References	501

Introduction

This article outlines the nature of the European Union (EU) as a *sui generis* political system with policy capacity in support of supranational integration and Europeanization and identifies the implications of this nature for European education policy developments. In so doing, it privileges the perspective of European governance and that which occurs at the EU level rather than in domestic education systems across Europe.

Since the creation of the EU with its particular, hybrid nature, scholars have mobilized theories, concepts, and methods from across disciplines (international relations, political science, and sociology) and subfields (comparative politics, policy sociology) to account for supranational integration within and beyond the European region (Wiener et al., 2019), the institutionalization of the EU (Georgakakis, 2011), and the way it functions as a political system (Costa and Brack, 2019), but also to examine the EU's influence on member states, and vice-versa, as well as the horizontal (across member states) diffusion of discourses, identities, political structures, and public policies (Radaelli, 2003; Graziano and Vink, 2008). Alongside these studies, a rich cluster of sociological and governance-focused frameworks have been adopted to understand broader EU dynamics (Rowell and Mangenot, 2016). Recognizing that neither EU institutions nor member states are monolithic actors (Kohler-Koch and Eising, 1999), these frameworks place actors at the center of analyses. Accordingly, some sociologists have been inspired by critical and post-positivist approaches to make the EU political arena their focus (Georgakakis and Weisben, 2010), while others have considered the multi-level, non-hierarchical configuration of the EU's socioeconomic and political system (Peters and Pierre, 1998) and examined interactions among actors (individuals and corporate; public and private), their resources, strategies, interests, and constraints (Rhodes, 1997; Rhodes et al., 1996). With few exceptions (cf. Bevir and Phillips, 2019), much of this literature disregards education as an object of study; nonetheless, scholars concerned with education policy across disciplines (education, sociology, and political science) often draw on these studies either explicitly or implicitly to assess the implications of supranational integration and Europeanization for education policy and the governance of education (including, among others: Dale and Robertson, 2009; Lawn and Grek, 2012; Lawn and Normand, 2015; Vatrella and Milana, 2019; Sørensen et al., 2021).

This article begins by providing an overview of supranational integration, Europeanization, and the EU political system so as to appreciate the position education has come to hold in this system and its implications for policy developments. The article then assumes the perspective of European governance and the instruments that substantiate it, particularly expert groups. This perspective acknowledges that both EU institutions and various policy actors matter for the formation of European education policy and the educationalization of public policy in the EU. In this vein, attention is also paid to the European Semester, a novel architecture for EU socio-economic policy coordination that is reconfiguring the policy capacity of a plurality of actors, as an aspect deserving of further attention by education policy researchers.

Supranational integration, Europeanization, and education policy

Officially formed as the EU in 1993 (when the Maastricht Treaty entered into force), this entity's foundation dates to the 1950s when a regional organization, the European Economic Community (1957), was established to bring about economic integration by setting up a common market and customs union among six countries. These countries had come together after World War II to integrate the coal and steel industry by forming the European Coal and Steel Community (1951) and thereby averting the mobilization of a new armed force in the region. Today, the EU is considered an important actor in the global economy, security, health, and international affairs, and plays a growing role in education. It represents a unique economic and political coalition of 27

countries (following the withdrawal of the United Kingdom's membership in 2019) located in geographical proximity to one another and which together cover an area of approximately 4 million km², a large part of the corresponding continent, and host a population of almost 447 million people as of 2020. The EU's uniqueness lies in its pooled sovereignty through which member states remain independent sovereign nations but delegate part of their decision-making powers to the institutions of the EU. Although these institutions act within the limits of the powers conferred on them by the founding Treaty, member states can opt to revise the Treaty. The Lisbon Treaty, enforced in 2009, purposely adjusted the EU decision-making process to increase the power of national parliaments and give more voice to European citizens (Costa and Brack, 2019). Unlike federations of states like the United States of America or international organizations based on co-operation between governments like the United Nations (UN) or Organization for Economic Cooperation and Development (OECD), therefore, member states in the EU not only coordinate their policies through intergovernmental cooperation but also engage in joint decision-making. This leads to supranational integration (CVCE, 2016) or "the formation of economic and/or political linkages among countries that are geographically near to each other" (Graziano and Vink, 2008, pp. 7–8) and brings about transformations in the institutions, policies, ideas, etc. that constitute and explain Europeanization (Graziano and Vink, 2013).

Since the creation of the European Communities, scholars have used various approaches to comprehend supranational integration, Europeanization, and the very nature of the EU as an entity that, as neither a federal state nor a typical international organization, represents a *sui generis* political system with its own institutional architecture (Hix and Hoyland, 2011; Costa and Brack, 2019). Originally, international relations scholars developed grand theories such as neo-functionalism or intergovernmentalism to explain supranational integration as a functional necessity (Haas, 1958) or the factors behind the construction of "an international regime for policy coordination" (Moravcsik, 1993, p. 480). Over time, both of these grand theories have come under critique. Neo-functionalism specifically has been called into question for viewing the expansion of EU institutions' authority from the economic to the political domain as a simple spillover over effect and underplaying the significance of domestic political processes and structures; intergovernmentalism has been criticized for giving too much weight to national leaders' rational choice to better manage their interests through supranational integration as the reason for their support of Europeanization (Costa and Brack, 2019). Although these theories might be seen as complementary rather than opposed (cf. Wiener et al., 2019), some scholars, drawing on comparative politics, reject both of these normative perspectives to assert an analytical recognition of the EU as a specific form of federalism with a factual decision-making center (e.g. the EU institutional architecture) in which member states' interests and willingness play a vital role (Börzel and Hosli, 2003).

New grand theories were thus developed following the Maastrich and Lisbon treaties. For instance, post-functionalism attempts to better capture the coupling of national and EU politics by granting attention to interest groups well outside the economic arena (Hooghe and Marks, 2001) whereas new intergovernmentalism recognizes the centrality EU institutions have acquired in important policy domains (e.g. economic governance) as well as the weakened relative role EU institutions play in determining the direction and nature of supranational integration (Bickerton et al., 2015). Alongside these developments, multiple middle-theories and conceptual approaches have also been formulated drawing on political science (particularly comparative politics) and sociology. For instance, different "new institutionalisms" (Hall and Taylor, 1996) largely "agree on the importance of institutions (*institutions matter*) and believe that they influence the political process and the EU's public policy and consequently confirm their impact on the European integration process in the long term" (Costa and Brack, 2019, p. 56, emphasis in the original). These approaches have led to studies examining the EU as a political system on its own terms (Hix and Hoyland, 2011).

Also in the wake of the Maastrich Treaty, scholars have employed various perspectives to investigate Europeanization, that is, the effects of supranational integration. As Costa and Brack (2019) note, these encompass: (1) a top-down or "downloading" perspective that casts member states as passive recipients of policy and visions consolidated at the EU level; (2) a bottom-up or "uploading" perspective that considers how national governments and other private and public institutions influence EU policy; (3) a "horizontal" perspective emphasizing the transfer of politics, policies and policy-making between member states; and (4) a perspective that, combining all of the above, conceptualizes Europeanization as the outcome of EU member states' constant interactions as well as horizontal diffusion processes. However, Europeanization as a research object also remains a matter of methodological debate as it is not always clear the extent to which it captures something which is substantially different from the effects of non-EU specific phenomena (e.g. globalization) (cf. among others Radaelli, 2002; Graziano and Vink, 2008). There thus remains a great deal of confusion as to the causes and effects of Europeanization (Costa and Brack, 2019). Having acknowledged all these points, a comprehensive definition of Europeanization points at:

a process of (a) construction, (b) diffusion, and (c) institutionalization of formal and informal rules, procedures, policy paradigms, styles, "ways of doing things", and shared beliefs and norms which are first defined and consolidated in the making of EU public policy and politics and then incorporated in the logic of domestic discourse, identities, political structures, and public policies.

Radaelli (2003, p. 30).

Accordingly, the making of EU education policy is not independent from the EU political system and the position education has come to hold within this system.

The EU political system and education

Although it is hard to establish a clear-cut separation between executive and political powers among EU institutions, the EU political system can be described as a quadripartite structure with the European Council and European Commission holding executive power while the European Parliament and Council of the European Union hold legislative power (Costa and Brack, 2019). The European Council, conceptualized as “a collective Head of State” (Costa and Brack, 2019, p. 76), is composed of the heads of state or government of all member states, the President of the Council, and the President of the European Commission. It decides on the overall political direction of the EU and sets its policy agenda and priorities. These often encompass education-related actions, goals, and long-term priorities. The European Commission (hereafter the Commission) has been compared to a government by virtue of its central administration and is comprised of one commissioner per member state as well as over 32,000 policy officers, researchers, lawyers, and translators. It represents the interests of the EU, implements its decisions, manages its budget and programs, and is the only institution empowered to initiate EU legislation and represent the EU outside of Europe. Much attention has thus been paid to tensions deriving from the nature of the Commission, with its combination of administrative functions and political leadership (Christiansen, 2006), and many concur that it has acquired increased political influence over time (Cisotta, 2013) on many matters, including education (Eeva, 2018; Gornitzka, 2009; Krick and Gornitzka, 2019). The European Parliament (hereafter the Parliament), and Council of the European Union (hereafter the Council) have been likened to “a kind of bicameral parliament” (Costa and Brack, 2019, p. 116). On the one hand, the Parliament, made up of representatives elected directly by European citizens, is a co-legislator on equal footing with the Council and co-responsible for adopting EU legislation; it also holds budgetary and supervisory powers. Each member of the Parliament may join both informal groups and formal groups based on political affiliation and specialized committees focused on education-related matters (e.g. Culture and Education—CULT, Employment and Social Affairs—EMPL Committee). Consequently, the Parliament’s political influence is seen to depend on its members’ ability to organize deliberations (Costa and Brack, 2019). On the other hand, the Council, an equal co-legislator with the Parliament, represents member states through their ministries and is organized into ten policy areas (e.g. Education, Youth, Culture and Sport—EYCS; Employment, Social Policy, Health and Consumer Affairs—EPSCO). The Council thus coordinates member states in relation to economic and fiscal policies, education policies (associated with culture, youth, and sport), and employment policies.

In accordance with the subsidiarity principle, the EU political system has exclusive competence over certain policy areas (e.g. the internal market, customs union, monetary policy for euro countries) and shared or supporting competence over others (e.g. social policy, health). Since the Maastricht Treaty, education has been among the EU’s supporting competences. Accordingly, a number of historical accounts employing diverse theoretical perspectives have revealed how complex intergovernmental cooperation in the educational sphere is in the EU, demonstrating that “actual cooperation [in education] proceeded only slowly, as a by-product of the gradual development of the common market” (Rasmussen, 2014, p. 19) but has intensified at an extraordinary pace following the Lisbon Agenda (2000) through the growing use of standards (Lawn and Grek, 2012) and interaction among actors (i.e. networks) (Gornitzka, 2009).

Although archival research shows that there have been discussions around the role of communitarian education initiatives in fostering the European project since the 1950s, no mention of education was made in the Treaty (1957) setting up the European Community (St. John, 2021). Vocational training has been a sensitive issue since the 1950s (Milana, 2008) and economic support for training activities has been available since 1951 via the European Social Fund; nevertheless, the idea of a common vocational training policy, for instance, initially had little impact. In 1975, however, this idea led to the creation of the European Center for the Development of Vocational Training (CEDEFOP), to support the harmonization of communitarian policies (Pépin, 2006), and its twin agency the European Training Foundation (ETF), in 1990, to support reforms in non-EU countries. Similarly, an action program for European cooperation in education launched in the mid-1970s through a Council resolution was met with resistance from the ministries of a few member states (Rasmussen, 2014) but nonetheless paved the way for more Community action in education and training from the 1980s onwards (e.g. Erasmus program) (St. John, 2021). Intergovernmental cooperation in education has gradually expanded from that point on (Rasmussen, 2014), leading to the introduction of the subsidiarity principle in response to member states’ fear that the EU would otherwise centralize power over education as well (Amaral et al., 2009). However, the subsidiary principle is open to diverse interpretations (Føllesdal, 1998) and does not necessarily stem EU “encroachment” on national education decision-making. For instance, from the mid-1990s onward lifelong learning became the object of EU policy changes aimed at tackling wide-ranging economic and social issues under the banner of economic policy (Volles, 2016). In contrast, it was the education ministries of 29 countries—all EU members (15) or candidate members (10) at the time, plus Iceland, Liechtenstein, Norway, and Switzerland—that first agreed on greater intergovernmental cooperation in higher education, thereby giving rise to the Bologna process (1999) with no direct involvement by EU institutions at that time. This development demonstrates that member states “were apparently prepared to loosen their embrace of the principle of subsidiarity” (Amaral et al., 2009, p. 7), and may explain why, when the European Council agreed on the Lisbon Agenda in 2000, member states not only opted for a EU frame to modernize education through intergovernmental cooperation (aka the Education and Training work program) but also set aside their original resistance and allowed EU institutions to become involved in the Bologna process.

Stipulated alongside the Lisbon Agenda (2000) was the Open Method of Coordination, a mechanism through which the European Council agrees on broad policy goals and guidelines that member states then translate into national and regional policies, as well as monitoring and evaluation benchmarks and indicators overseen by the EU political system (Eggermont, 2012). Many observers thus agree that, through the Open Method of Coordination, member states have given the EU political system a mandate to develop a regional approach and common objectives in education as well (see, among others: Hingel, 2001; Ertl, 2006; Nóvoa

and de Jong-Lambert, 2003; Milana, 2014). This is evidenced by the advancement of the Bologna process, now a complex mechanism supporting cooperation in higher education among 48 governments—representing almost the totality of the European continent—and the European Commission. Its current architecture, based on extensive standardization devices and monitoring practices operating under the Open Method of Coordination, has come to represent “a regulative technology designed to govern at a distance [that] reaches deep into the ontological matter of everyday working life in higher education organizations” (Brøgger, 2019, p. 4) across EU member states and beyond. While the Bologna process was initiated as a form of intergovernmental cooperation supporting horizontal integration, therefore, the Lisbon Agenda—a supranational process supporting vertical integration in multiple arenas including education—represented a turning point in terms of European integration efforts increasingly affecting education through European governance.

European governance, policy coordination, and education

European governance is the regional regime underlying public action in the EU (Bevir, 2019), action that is based on multi-level relations in that it creates connections among EU institutions, national governments, and other regional and local authorities. It supports supranational integration and Europeanization through a mix of “hard” and “soft” mechanisms (Trubek et al., 2005), with the former operating “through rules that arise from treaties, directives and regulations” while the latter create “non-binding rules that are nevertheless expected to produce effects in practice” (Maggetti, 2015, p. 252) through flexible structures and arrangements (i.e. committees, task forces, working groups, expert groups, etc.). As seen in the previous sections, hard governance manifests in the EU political system in charge of defining and guaranteeing the operation of EU institutions: this is stronger in the areas of exclusive EU competence (e.g. the internal market) and softer in areas, like education, in which the EU has only supporting competences. In addition, in areas of shared competences (e.g. economic, and social policies), the EU, like member states, sometimes passes laws and regulations that may have (intended and unintended) effects on education as well. There is thus a mix of hard and soft mechanisms operating in diverse policy domains that can influence European education policy, including the most deeply-studied mechanism of the Open Method of Coordination (see, among others: Brøgger, 2019; Chatzopoulou, 2015; Ravinet, 2008; Tholoniati, 2010; Walters and Haahr, 2005). In order for policies agreed on at the EU level to prompt action by national governments and other regional and local authorities in areas of supporting competence, however, European governance requires coordinated action. Both the Education and Training work program (2000-onwards) and the European agenda on adult learning (2006-onwards) illustrate this point.

In the wake of the Lisbon Agenda, with the agreement on *The concrete future objectives of education systems* (European Commission, 2001) a comprehensive approach to education was first laid out at the EU level with the aim of increasing the quality and effectiveness of national education and training systems, facilitating better access for all, and opening up European education and training systems to the wider world. This led to the *Education and Training 2010* work program as the agreement’s operational arm. In keeping with its implementation, the Council set up a Strategic Framework for European Cooperation in education and training: *Education and Training 2020* (Council of the European Union, 2009). This strategic framework, “spanning education and training systems as a whole in a lifelong learning perspective” (Council of the European Union, 2009, p. 2), served as a policy umbrella for several parallel processes in higher education, vocational education and training, etc. Alongside the progression of the Education and Training work programs, a dedicated agenda on adult learning was also established at the EU level, the developments of which were examined recently (Milana and Klatt, 2019). Initially, naming 1996 as the European year of lifelong learning and the Council’s *Resolution on lifelong learning* (June 2002) laid the groundwork for adult learning to emerge as a distinctive policy domain at the EU level. At this stage, the adult learning side of lifelong learning was teased out in dialog between EU institutions and went to bolster ties between education, training and employment policies. EU institutions then concentrated on existing statistical data gaps at the microlevel (learner-centered) and more robust knowledge exchange and collaboration across member states and with other international organizations with an interest in adult and further education (such as the OECD). With the 2006 *Communication of the European Commission Adult Learning: it is never too late to learn* (European Commission, 2006) and Council’s Conclusions on adult learning (Council of the European Union, 2008), adult learning was established as a clearly defined policy domain. However, the outbreak of the 2009 global financial crisis also impinged on efforts to fine-tune a dedicated agenda on adult learning, so in 2011 the Council approved a *Resolution on a renewed European agenda for adult learning* that reinforced the links at EU level between adult learning and other policy domains (e.g. education and training, employment, macroeconomic policies).

Both *Education and Training 2020* and the *Renewed European agenda for adult learning* rely on a variety of mechanisms and policy instruments to ease policy coordination efforts on the part of EU institutions (and the Commission particularly) and bring about changes in domestic education systems. However, policy influence does not necessarily take place from the top down (from the EU to member states); rather, it works as a loop in which a plurality of policy actors partake. There are several mechanisms and policy instruments that can be taken as illustrating this loop, including standard-setting, capacity-building, elite-learning, and financial redistribution. Various characterized in the literature on supranational governance (cf. Dale, 1999; Lawn, 2011; Martens and Jakobi, 2010; Ozga et al., 2011; Woodward, 2009), these mechanisms have been examined in recent research on the European governance of lifelong learning (Milana et al., 2020a). Each of these mutually dependent mechanisms relies on a set of policy instruments. Specifically, standard-setting (a mechanism) involves taking normative actions and setting common goals that converge to establish a single, clear-cut European model in a specific policy domain (higher education, adult learning, etc.). This mechanism operates through implicit and explicit benchmarks (an instrument), namely the EU-established standards by which member state

performance is measured and compared and their level of quality judged. However, EU benchmarks are themselves the result of complex negotiations among heads of states and governments and their ministerial complements at EU institutions. Such negotiations thus imply capacity-building (a mechanism) to orient heads of states and governments and their ministerial complements in the direction of identifying policy solutions to what are perceived as common problems (e.g. high rates of adults with low levels of education and literacy). Capacity-building is facilitated by various policy instruments such as data generation, mutual- and peer-learning arrangements, and expert groups assembled by the Commission under the Open Method of Coordination framework to tackle specific issues (e.g. digital skills) for a certain amount of time. These instruments work more fruitfully in interaction with one another than they do in isolation. Specially, data generation encompasses the work of gathering quantitative and/or qualitative data, the method used to generate such data, and the procedure through which data is organized. And yet one of the instruments through which generated data circulates, and more data is generated, is the set of mutual- and peer-learning arrangements catering to member state representatives together with Commission staff members (among others) to help them identify and learn about ongoing initiatives in a certain area (i.e. the upskilling of low-skilled adults). Increasingly, one of the main collective actors to take advantage of these types of arrangements are expert groups.

Table 1 illustrates how the above-mentioned instruments are concretized in practice, thus contributing to the influence loop among the plurality of policy actors involved in European education policy (cf. [Milana et al., 2020a](#)). Against this backdrop, expert groups merit particular attention.

Expert groups and the network governance of (European) education

Expert groups are consultative bodies set up by the Commission or its departments for advisory purposes, but which under the principles of the Open Method of Coordination are engaged in fostering mutual learning, the exchange of good practice, and socialization among their members. These experts include representatives from member states, candidate countries, European Free Trade Association countries, and relevant EU bodies or agencies (e.g. CEDEFOP, ETF) as well as education and training associations and European social partners (e.g. the European Trade Union Confederation). As mentioned above, the Open Method of Coordination has been used extensively since the Lisbon Agenda to promote supranational integration and Europeanization in education, and expert groups have become essential for coordinating education policy at EU level (cf. [Gornitzka, 2009](#); [Krick and Gornitzka, 2019](#)) through a form of network governance “which is substantially alternative to both centralized government at the community level and to national government in single member states” ([Milana et al., 2000b](#), p. 32).

Current education policy research employs different methodological approaches for analyzing networks, and recognizes these approaches to be valuable ([Menashy and Verger, 2019](#)). One of these is Social Network Analysis ([Borgatti et al., 2009](#)), an approach recently used to delve into the structural configuration of the networks, and relationships among actors, formed as part of the activity of the expert groups active under the Education and Training work program, and in the adult learning domain (see **Table 1**) ([Milana et al., 2020b](#)). By applying Social Network Analysis, this study was able to conduct a structural analysis of relational phenomena (i.e. the emergent effect of the bonds between institutional actors) generated and/or reinforced by the establishment of expert groups in each of the policy domains under consideration (education and training vs. adult learning). To clarify, the forms of network governance in these two domains are not independent but rather connected through a specific expert group (the ET2020 Working Group on Adult Learning, cf. **Table 1**) whose actors are thus involved in both domains. The analysis brought attention to the mutual constitutiveness of a social network and its members as well as the potential power within and of a network to contribute to developments in European education policy. For instance, striking similarities were found across education and training vs. adult learning considered as distinct policy domains, suggesting that two national agencies—the Flemish Department of Education and Training (Belgium) and Finland’s Ministry of Education and Culture—stand out as institutional actors displaying a high degree of cohesiveness in both forms of network governance (education and training vs. adult learning). Meanwhile, remarkable differences between the two policy domains attest to the centrality achieved by organizations lobbying on behalf of workers and trades at the EU level. For instance, while the three major unions (i.e. European Trade Union Committee for Education, European Trade Union Confederation, and European Federation of Education Employers) were among the main “brokers” in the education and training network, these organizations have comparatively less connectivity and brokerage capacity in the adult learning network. Theoretically, greater centrality in a network is thought to entail better access to the information, material resources and/or relationships that represent important factors in decision-making. Moreover, high rates of integration and brokerage connectivity in a network are understood to unlock more possibilities for influencing others. Both aspects are important for more efficient policy coordination in areas of EU supporting competence like education where EU policies are not prescribed but “formed” through soft governance mechanisms and policy instruments such as expert groups. At the same time, however, “the practical effects of a high level of centrality are also dependent on individual capacities and organizational backing” ([Milana et al., 2020b](#), p. 45), two aspects that the authors did not consider in their study and which merit future attention.

Complicating the picture of European governance, a novel socio-economic architecture for EU policy coordination known as the European Semester has been developing over the last decade, affecting many domains including education and is thus worth examining.

Table 1 Examples of policy instruments in practice.

<i>Education and training 2020</i>	<i>Policy instruments</i>	<i>The European agendas for adult learning</i>
ET2020 working group on vocational education and training ET2020 working group on digital skills and competences ET2020 working group on modernization of higher education ET2020 working group on promoting citizenship and the common values of freedom, tolerance and non-discrimination through education ET2020 working group on schools ET2020 working group on adult learning In-depth country workshops—focus on policy development and practice in selected member states with the aim of identifying key factors for policy success. Peer learning activities—involve national experts in learning together based on evidence and experience and sharing both positive and negative policy experiences. Peer-counseling—a voluntary tool that brings together “professional peers from a small number of national administrations to provide external advice to a country in the process of a significant policy development” (European commission, 2015, p. 1) Other—ad hoc peer learning activities, thematic events, peer reviews, and policy learning exchanges. Education and training monitor—includes quantitative, comparative analyses and country-specific recommendations based on eurostat and OECD data. Country reports—annual individual country reports identify where each country stands in relation to the ET 2020 benchmarks and other indicators. Studies and reports—(e.g. <i>Promoting adult learning in the workplace</i>. Final report of the ET, 2020 working group, 2016–2018 on adult learning, see: WGAL 2018) Guidelines—(e.g. 20 guiding principles for high-performance apprenticeships and work-based learning in the 2014–2015 period, see: WGVET, 2017). Policy conclusions—(e.g. Higher education institutions as centers of regional development and innovation, see: WGMHE, 2016). Other—recommendations, background papers, flash reports, etc. At least 95% of children should participate in early childhood education Fewer than 15% of 15-year-olds should be under-skilled in reading, mathematics and science The rate of early leavers from education and training aged 18–24 should be below 10% At least 40% of people aged 30–34 should have completed some form of higher education At least 20% of higher education graduates and 6% of 18–34 year-olds with an initial vocational qualification should have spent some time studying or training abroad The share of employed graduates (aged 20–34 with at least upper secondary education attainment and having left education 1–3 years ago) should be at least 82%	Coordinated working groups/networks Mutual- and peer-learning arrangements Data generation Benchmarks	Working group on the implementation of the action plan on adult learning Thematic working group on quality assurance in adult learning Thematic working group on financing adult learning National coordinators implementing the european agenda on adult learning Regional meetings organized by the working group on the implementation of the action plan on adult learning (2008–2010). Workshops with country experts to hold in-depth discussions of topics (e.g. improving quality in adult learning, effective policies for increasing adults’ participation in basic skills provision). Conferences sponsored by the european commission (e.g. Adult skills conference: Empowering people, 6–7 December 2016) or in collaboration with other international organizations (e.g. Equipping adults for the 21st century: Joining forces for action on skills and competences, 9–10 December 2013, with the UNESCO institute for lifelong learning). Literature reviews (e.g. <i>Improving basic skills in adulthood: Participation and Motivation</i>). Studies and reports (e.g. <i>An in-depth analysis of adult learning policies and their effectiveness in Europe</i>). “Best” or “good practices” compendia (e.g. <i>ERASMUS+: Good practices in the implementation of the European Agenda for Adult Learning, 2012–2016</i>) State of the art reports (e.g. <i>State of the Art Report on the Implementation of the European Agenda for Adult Learning</i>, by the european association for the education of adults). Design and management of multi-country surveys (e.g. the labor force survey and the european adult education survey managed by EUROSTAT or the education and training monitor) Final support to multi-country household surveys (e.g. the program for international assessment of adult competences managed by the OECD).

(Continued)

Table 1 Examples of policy instruments in practice.—cont'd

<i>Education and training 2020</i>	<i>Policy instruments</i>	<i>The European agendas for adult learning</i>
At least 15% of adults should participate in lifelong learning	Funding schemes	European structural funds (ESF) (e.g. member states' participation in the program for international assessment of adult competencies was supported through these funds). Europe program for employment and social innovation (EaSI) (e.g. 2017 call for proposals dedicated to <i>Awareness-raising activities on "Upskilling Pathways: New Opportunities for Adults"</i>) Erasmus+ (e.g. Activities of national coordinators engaged in implementing the European agenda on adult learning in their countries have been partially financed through Erasmus+, under key 3 action: Support for policy reforms)

Policy coordination under the European Semester

The European Semester is a cycle of economic and fiscal policy coordination proposed by the European Commission, approved by the Council configuration for Economic and Financial Affairs (ECOFIN), and launched in January 2011. Framed as "a new framework for policy-coordination" at both horizontal and vertical levels, the European Semester "provide[s] a new socioeconomic governance architecture to co-ordinate national policies without transferring full sovereignty to the EU level" (Verdun and Zeitlin, 2018, p. 137). In extreme synthesis, the yearly cycle of the European Semester involves a preparatory phase (November-January), a phase of EU-level policy guidance (January-April), a country-specific phase (April-July), and an implementation phase (July-October). During the preparatory phase, EU institutions analyze the current situation of member states' budgetary and structural policies, as reported by the Commission. In the next phase (*policy guidance at EU level*), the Council studies the produced document and the Commission's recommendations, the Parliament expresses own opinions and, on these grounds, the European Council reaches an agreement on what policy directions should be pursued for the following year. During the country-specific phase, member states set their own objectives, policies, and plans, and these form the basis on which the Commission drafts Country-Specific Recommendations that are then subject to agreement, endorsement, and final adoption by the EU institutions. In 2021, this yearly cycle was adapted to allow member states to submit, in one integrated document, their national reform programs (as foreseen by the European Semester) together with their recovery and resilience plans to mitigate the economic and social impact of the COVID-19 pandemic.

Even before the outbreak of the COVID-19 pandemic and consequent EU mitigation strategy, scholars across multiple disciplines had begun to scrutinize the inherently supranational character of the European Semester through examinations of the political power and pivotal roles gained or lost by each EU institution (i.e. the Parliament, Council, and Commission) vs. member states, either in isolation or jointly (i.e. the European Council) (see Crum, 2018; Halleberg, et al., 2018; Maricut and Puetter, 2018; van der Veer and Haverland, 2018; Savage and Verdun, 2016; among others). The Commission emerged as the "core political driver" of the EU, with stronger powers vs. member states (Cisotta, 2013, p. 2–4, emphasis in original). A strand of the literature has focused on the (re)balancing of economic and social policy objectives within the EU (cf. Costamagna, 2013), questioning whether and to what extent the European Semester affected pre-existing relations between economic and social policies, including education and lifelong learning (cf. Stevenson, 2019a,b). Although empirical claims diverge, some point to social policy objectives being subordinated to economic ones (cf. Crespy and Menz, 2015a,b), with EU economic policy actors (the ECOFIN Council, among others) predominating in decision-making procedures (cf. de la Porte and Heins, 2014, 2015; Degryse et al., 2014) and education and other social reforms thus subordinated to economic imperatives (Stevenson et al., 2019). Alternative arguments, however, stress "a partial but progressive 'socialization' of the Semester" (Zeitlin and Vanhercke, 2018, p. 152). The European Semester, they suggest, is undergoing a progressive re-orientation to better balance economic and social policy objectives (including in education) and adopt a less hierarchical and more interactive decision-making process in which EU social and employment policy actors learn "to gain traction in 'evidence-based' deliberation with their economic policy counterparts" (Zeitlin and Vanhercke, 2018, p. 154), including on educational matters (cf. Stevenson et al., 2019). In short, many concur that the European Semester has produced a "shift of EU competences from the *soft* coordination of economic policies to a more binding and intrusive agenda-setting procedure" (Cisotta, 2013, p. 6, emphasis in original), "bringing together fiscal and macroeconomic coordination instruments with social and employment ones with the same cycle" (Louvais Fasois, 2017, p. 2). Under the European Semester's composite policy coordination process, therefore, both horizontal relations among EU institutions and EU-member states' vertical relations have been strengthened, including in education (Stevenson et al., 2019).

Conclusion

This article has outlined the EU as a *sui generis* political system with policy capacity in support of supranational integration and Europeanization and considered the implications of this for European education policy developments. Drawing on different theoretical perspectives, research on European education policy over the last two decades has acknowledged the supranational process supporting vertical integration triggered by the Lisbon Agenda and, lately, also the contribution to Europeanization made by the European Semester with its compound process of horizontal and vertical policy coordination. These findings have underlined that education (secondary, further, and higher education, as well as adult and vocational education) is not a stand-alone policy domain in the EU but part of “a complex interweaving” (St. John, 2021, p. 6) of policy domains increasingly built on the educationalization of social problems (Depaepe and Smeyers, 2008; Tröhler, 2016), a far more complex phenomenon if not an explicit strategy at the level of the EU (Sørensen et al., 2021; St. John and Murphy, 2019; Valiente et al., 2020). All this calls for stronger policy coordination across a large variety of actors with vested interests in education policy. This article has thus paid special attention to the need and capacity for EU policy coordination from a European governance perspective, thereby recognizing that both EU institutions and various policy actors matter for the formation of European education policy and the educationalization of public policy in the EU. Fittingly, major developments in European governance, encompassing the European Semester, address education (Milana et al., 2020a; Sørensen et al., 2021; Stevenson et al., 2020; Traianou and Jones, 2019), thus encouraging more research attention.

References

- Amaral, A., Neave, G., Musselin, C., Maassen, P. (Eds.), 2009. *European Integration and the Governance of Higher Education and Research*. Springer, Dordrecht.
- Bevir, M., Phillips, R. (Eds.), 2019. *Decentering European Governance*. Routledge, London.
- Bevir, M., 2019. Governance. In: *Encyclopædia Britannica*. <https://academic-eb-com.ezproxy.nottingham.ac.uk/levels/collegiate/article/governance/600930>.
- Bickerton, C.J., Hodson, D., Puetter, U., 2015. *The New Intergovernmentalism: States and Supranational Actors in the Post-maastricht Era*. Oxford University Press, Oxford.
- Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G., 2009. Network analysis in the social sciences. *Science* 323, 892–985.
- Börzel, T.A., Hosli, M.O., 2003. Brussels between Bern and Berlin: comparative federalism meets the European Union. *Governance* 16 (2), 179–202.
- Brøgger, K., 2019. *Governing Through Standards: The Faceless Masters of Higher Education—the Bologna Process, the EU and the Open Method of Coordination*. Springer, Cham.
- Chatzopoulou, S., 2015. Unpacking the mechanisms of the EU “throughput” governance legitimacy: the case of EFSA. *Eur. Politics Soc.* 16 (2), 159–177.
- Christiansen, T., 2006. The European Commission: the European executive between continuity and change. In: Richardson, J.J. (Ed.), *European Union: Power and Policy-Making*. Routledge, New York, pp. 99–119.
- Cisotta, R., 2013. What Role for the European Commission in the New Governance of the Economic and Monetary Union? *Istituto affari internazionali*, Roma.
- Costa, O., Brack, N., 2019. *How the EU Really Works*. Routledge, London, New York.
- Costamagna, F., 2013. *The European Semester in Action: Strengthening Economic Policy Coordination While Weakening the Social Dimension?* (Working Paper-Laboratorio di Politica Comparata e Filosofia Pubblica, 5) Centro Einaudi, Torino.
- Council of the European Union, 2008. Council conclusions of 22 May 2008 on adult learning. *Off. J. Eur. Union C* 140, 10–13.
- Council of the European Union, 2009. Council Conclusions of 12 May 2009 on a Strategic Framework for European Cooperation in Education and Training (“ET 2020”). Retrieved from: http://www.cedefop.europa.eu/files/education_benchmarks_2020.pdf.
- Crespy, A., Menz, G. (Eds.), 2015a. *Social Policy and the Euro Crisis: Quo Vadis Social Europe*. Palgrave Macmillan, Houndmills.
- Crespy, A., Menz, G., 2015b. Commission entrepreneurship and the debasing of social Europe before and after the Eurocrisis. *J. Common. Mark. Stud.* 53 (4), 753–768.
- Crum, B., 2018. Parliamentary accountability in multilevel governance: what role for parliaments in post-crisis EU economic governance? *J. Eur. Publ. Pol.* 25 (2), 268–286.
- CVCE, 2016. *The European Communities*. Retrieved 26 November 2020.
- Dale, R., Robertson, S.L., 2009. *Globalization and Europeanization in Education*. Symposium Books, Oxford.
- Dale, R., 1999. Specifying globalization effects on National Policy: a focus on the mechanisms. *J. Educ. Pol.* 14 (1), 1–18.
- de la Porte, C., Heins, E., 2014. Game change in EU social policy: towards more European integration. In: Rodrigues, M.J., Xiarchogiannopoulou, E. (Eds.), *The Euro Crisis and the Transformation of EU Governance*. Ashgate, Farnham, pp. 157–170.
- de la Porte, C., Heins, E., 2015. A new era of European integration? Governance of labor market and social policy since the sovereign debt crisis. *Comp. Eur. Polit.* 13 (1), 8–28.
- Degryse, C., Jepsen, M., Pochet, P., 2014. A worrying prospect: towards a more imbalanced European social governance? In: Natali, D. (Ed.), *Social Developments in the European Union 2013*. European Trade Union Institute/European Social Observatory, Brussels, pp. 57–83.
- Depaepe, M., Smeyers, P., 2008. Educationalization as an ongoing modernization process. *Educ. Theor.* 58 (4), 379–389.
- Eeva, K., 2018. *The EU's European Semester: Soft Power and Knowledge in the Governing of Education*. Doctoral dissertation. University of Oxford.
- Eggermont, F., 2012. *The Changing Role of the European Council in the Institutional Framework of the European Union*. Intersentia, Cambridge, MA.
- Ertl, H., 2006. European Union policies in education and training: the Lisbon Strategy as a turning point? *Comp. Educ.* 42 (1), 5–27.
- European Commission, 2001. *The Concrete Future Objectives of Education Systems*. Report from the Commission. COM (2001) 59 Final, 31 January 2001.
- European Commission, 2006. *Communication from the Commission—Adult Learning: It is Never Too Late to Learn*. COM(2006) 0614 Final. The European Commission, Brussels.
- European Commission, Peer Counselling, 2015. *A New Element in the ET 2020 Toolbox*. Retrieved from: https://ec.europa.eu/education/sites/education/files/document-library-docs/et-2020-peer-counselling-brochure_en.pdf.
- Follesdal, A., 1998. Subsidiarity. *J. Polit. Philos.* 6 (2), 231–259.
- Georgakakis, D., 2011. Don't throw out the “Brussels Bubble” with the bathwater: from EU institutions to the field of Eurocracy. *Int. Political Sociol.* 5, 331–334.
- Georgakakis, D., Weisben, J., 2010. From above and from below: a political sociology of European actors. *Comp. Eur. Polit.* 8 (1), 93–109.
- Gornitzka, A., 2009. *Networking administration in Areas of National Sensitivity: the commission and European higher education*. In: Amaral, A., Neave, G., Musselin, C., Maassen, P. (Eds.), *European Integration and the Governance of Higher Education and Research*. Springer, Dordrecht, pp. 109–132.
- Graziano, P., Vink, M.P., 2008. *Europeanization. New Research Agendas*. Palgrave Macmillan, Basingstoke.
- Graziano, P., Vink, M.P., 2013. *Europeanization: concept, theory and methods*. In: Bulmer, S., Lequesne, C. (Eds.), *The Member States of the European Union*. Oxford University Press, Oxford, pp. 31–54.
- Haas, E.B., 1958. *The Uniting of Europe: Political, Social and Economic Forces, 1950–1957*. Stevens & Sons, London.
- Hall, P.A., Taylor, R.C.R., 1996. Political Science and the three new institutionalisms. *Polit. Stud.* 44 (5), 936–957.
- Hallerberg, M., Marzotto, B., Wolff, G.B., 2018. Explaining the evolving role of national parliaments under the European Semester. *J. Eur. Publ. Pol.* 25 (2), 250–267.
- Hingel, A.J., 2001. *Education Policies and European Governance. Contribution to the Interservice Groups on European Governance*. DG EAC/A/1 (Development of Educational Policies). European Commission, Directorate-General for Education and Culture, Brussels.

- Hix, S., Hoyland, B., 2011. *The Political System of the European Union*, third ed. Palgrave Macmillan, Basingstoke.
- Hooghe, L., Marks, G., 2001. *Multi-level Governance and European Integration*. Rowman and Littlefield, Oxford.
- St John, S., Murphy, M. (Eds.), 2019. *Education and Public Policy in the European Union—Crossing Boundaries*. Palgrave Macmillan, Cham.
- St John, S., 2021. *Education and Solidarity in the European Union: Europe's Lost Spirit*. Palgrave Macmillan, Cham.
- Kohler-Koch, B., Eising, R., 1999. *The Transformation of Governance in the European Union*. Routledge, London.
- Krick, E., Gornitzka, A., 2019. The governance of expertise production in the EU Commission's "high level groups": tracing expertization tendencies in the expert group system. In: Bevir, M., Ryan, P. (Eds.), *Decentering European Governance*. Routledge, London, pp. 102–120.
- Lawn, M., Grek, S., 2012. *Europeanizing Education. Governing a New Policy Space*. Symposium Books, Oxford.
- Lawn, M., Normand, R., 2015. *Shaping of European Education: Interdisciplinary Perspectives*. Routledge, Abingdon, Oxon; New York, NY.
- Lawn, M., 2011. Standardizing the European Education Policy space. *Eur. Educ. Res. J.* 10 (2), 259–272.
- Louvais Fasois, C., 2017. The Subtle Mechanisms of Europeanization: The European Semester and Its Effects on the Social and Employment Policies of Belgium. Paper Presented at the Fifteenth Biennial Conference of the European Union Studies Association (EUSA) in Miami, Florida, May 4–6 2017.
- Maggetti, M., 2015. Hard and soft governance. In: Lynggaard, K., Manners, I., Löfgren, K. (Eds.), *Research Methods in European Union Studies*. Palgrave Macmillan, London, pp. 252–265.
- Maricut, A., Puetter, U., 2018. Deciding on the European Semester: the European Council, the Council and the enduring asymmetry between economic and social policy issues. *J. Eur. Publ. Pol.* 25 (2), 193–211.
- Martens, K., Jakobi, A.P., 2010. *Mechanisms of OECD Governance: International Incentives for National Policy-Making?* Oxford University Press, Oxford.
- Menashy, F., Verger, A., 2019. The value of network analysis for the study of global education policy—key concepts and methods. In: Gorur, R., Sellar, S., Steiner-Khamisi, G. (Eds.), *Comparative Methodology in the Era of Big Data and Global Networks (World Yearbook of Education)*. Routledge, London/New York, pp. 117–131.
- Milana, M., Klatt, G., 2019. Governing adult education policy development in Europe. In: McGrath, S., Mulder, M., Papier, J., Suart, R. (Eds.), *Handbook of Vocational Education and Training*. Springer, Cham, pp. 789–812.
- Milana, M., Klatt, G., Vatrella, S. (Eds.), 2000a. *Europe's Lifelong Learning Markets, Governance and Policy—Using an Instruments Approach*. Palgrave Macmillan, Cham, Switzerland.
- Milana, M., Klatt, G., Tronca, L., 2020b. Towards a network governance of European lifelong learning: a structural analysis of commission expert groups. *Int. J. Lifelong Educ.* 39 (1), 31–47.
- Milana, M., 2008. Is European active citizenship fostering inclusion within the Union? A critical review. *Eur. J. Educ.* 43 (2), 207–216.
- Milana, M., 2014. Europeanization and the changing nature of the (European) state: implications for studying adult and lifelong education. In: Milana, M., Holford, J. (Eds.), *Adult Education Policy and the European Union: Theoretical and Methodological Perspectives*. Sense Publishers, Rotterdam, pp. 73–90.
- Moravcsik, A.M., 1993. Preferences and power in the European Community: a liberal intergovernmentalist approach. *J. Common. Mark. Stud.* 314, 473–524.
- Nóvoa, A., de Jong-Lambert, W., 2003. The education of Europe: apprehending EU educational policies. In: Phillips, D., Ertl, H. (Eds.), *Implementing European Union Education and Training Policy. A Comparative Study of Issues in Four Member States*. Kluwer, Dordrecht, pp. 41–72.
- Ozga, J., Dahler-Larsen, P., Segerholm, C., Simola, H. (Eds.), 2011. *Fabricating Quality in Education—Data and Governance in Europe*. Routledge, London.
- Pépin, L., 2006. *The History of European Cooperation in Education and Training*. EU Publications Office, Brussels.
- Peters, B.G., Pierre, J., 1998. Governance without government? Rethinking public administration. *J. Publ. Adm. Res. Theor.* 8 (2), 223–243.
- Radaelli, C.M., 2002. The domestic impact of European Union public policy: notes on concepts, methods, and the challenge of empirical research. *Polit. Eur.* 1 (5), 105–136.
- Radaelli, C.M., 2003. The Europeanization of public policy. In: Featherstone, K., Radaelli, C.M. (Eds.), *The Politics of Europeanization*. Oxford University, Oxford, pp. 27–56.
- Rasmussen, P., 2014. Adult learning policy in the European Commission—development and status. In: Milana, M., Holford, J. (Eds.), *Adult Education Policy and the European Union*. Sense, Rotterdam, pp. 17–34.
- Ravinet, P., 2008. From voluntary participation to monitored coordination: why European countries feel increasingly bound by their commitment to the Bologna process. *Eur. J. Educ.* 43 (3), 353–367.
- Rhodes, R.A.W., Bache, I., George, S., 1996. Policy networks and policy making in the European Union: a critical appraisal. In: Hooghe, L. (Ed.), *Cohesion Policy and European Integration: Building Multilevel Governance*. Clarendon Press, Oxford, pp. 367–387.
- Rhodes, R.A.W., 1997. *Understanding Governance. Policy Networks, Governance, Reflexivity and Accountability*. Open University Press, Buckingham.
- Rowell, J., Manganot, M., 2016. *Political Sociology of the European Union*. Manchester University Press, Manchester.
- Savage, J.D., Verdun, A., 2016. Strengthening the European Commission's budgetary and economic surveillance capacity since Greece and the Euro area crisis: a study of five Directorates-General. *J. Eur. Public Policy* 23 (1), 101–118.
- Sørensen, T.B., Grimaldi, E., Gajderowicz, T. (Eds.), 2021. *Rhetoric or Game Changer: Social Dialogue and Industrial Relations in Education Midst EU Governance and Privatization in Europe*. ETUCE, Brussels.
- Stevenson, H., Milner, A., Winchip, E., Hagger-Vaughan, L., 2019. Education policy and the European Semester: challenging soft power in hard times. In: Tett, L., Hamilton, M. (Eds.), *Resisting Neoliberalism in Education: Local, National and Transnational Perspectives*. Polity Press, Cambridge, pp. 211–224.
- Stevenson, H., Hagger-Vaughan, L., Milner, A., Vanhercke, B., Belletti, C., Pond, R., 2020. *The European Semester – the challenges for public service trade unions - Final report*. Retrieved from: <https://www.epsu.org/sites/default/files/article/files/FINAL%20REPORT.pdf>.
- Stevenson, H., 2019a. *The European Semester: A Literature Review*. European Public Service Union, Brussels.
- Stevenson, H., 2019b. *The European Semester: An Analysis of the 2017–2018 Cycle*. European Public Service Union, Brussels.
- Tholoniati, L., 2010. The career of the open method of coordination: lessons from a "soft" EU instrument. *W. Eur. Polit.* 33 (1), 93–117.
- Traianou, A., Jones, K., 2019. Austerity and the remaking of education policy in Europe since 2008. In: Traianou, A., Jones, K. (Eds.), *Austerity and the Remaking of European Education*. Bloomsbury, New York.
- Tröhler, D., 2016. Educationalization of social problems and the educationalization of the modern world. In: Peters, M. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore.
- Trubek, D., Cottrell, M., Nance, M., 2005. "Soft Law", "Hard Law", and European Integration: Toward a Theory of Hybridity. University of Wisconsin Legal Studies. Research Paper No. 1002.
- Valiente, O., Capsada-Munsech, Q., de Otero, J.P.G., 2020. Educationalization of youth unemployment through lifelong learning policies in Europe. *Eur. Educ. Res. J.* 19 (6), 525–543.
- van der Veer, R., Haverland, M., 2018. Bread and butter or bread and circuses? Politicization and the European Commission in the European Semester. *Eur. Union Polit.* 19 (3), 524–545.
- Vatrella, S., Milana, M., 2019. European Governance and Educational Policy. Youth guarantee as a "trail" between labor and learning. *Sc. Democr.* 4, 189–205.
- Verdun, A., Zeitlin, J., 2018. Introduction: the European Semester as a new architecture of EU socioeconomic governance in theory and practice. *J. Eur. Publ. Pol.* 25 (2), 137–148.
- Volles, N., 2016. Lifelong learning in the EU: changing conceptualizations, actors, and policies. *Stud. High Educ.* 41 (2), 343–363.
- Walters, W., Haahr, J.H., 2005. *Governing Europe: Discourse, Governmentality and European Integration*. Routledge, New York.
- Wiener, A., Börzel, T.A., Risse, T., 2019. *European Integration Theory*. Oxford University Press, Oxford.
- Woodward, R., 2009. *The Organization for Economic Co-operation and Development*. Routledge, Milton Park, Abingdon, Oxon and New York, NY.
- Zeitlin, J., Vanhercke, B., 2018. Socializing the European Semester: EU social and economic policy co-ordination in crisis and beyond. *J. Eur. Publ. Pol.* 25 (2), 149–174.

The World Trade Organization

Susan L. Robertson, Faculty of University of Cambridge, Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	503
Early history of the WTO	503
What does GATS involve?	504
Education—from public services to a tradeable services sector	505
Forum shifting—Regional Trade Agreements	506
The Trade in Services Agreement (TISA)	507
Trade related aspects of intellectual property rights	507
Conclusions	508
References	509

Introduction

National education systems and the World Trade Organization (WTO) might appear an odd pairing of institutions. After all, we teach and learn in one, and we regulate the global trading system in the other. However, this chapter will make the case that their relationship has emerged out of a growing view since the 1980s that education should be regarded as a service, that education systems can be construed as belonging to a services sector, that knowledge can be treated as a commodity, and that education services can be regulated through rules agreed to as part of the global trading system.

Understanding the history behind the World Trade Organization, the agreements that are invoked to bring education into the global regulatory system, and the mechanisms to manage disputes is important. This is because across many of the WTO member states, education services are now reported in international trade figures, while education activity is itself framed in very specific ways to articulate with the language of global trade. I explore these modalities, the politics of contestation that have accompanied efforts to bring education into trading relations and their regulations, along with new initiatives that have tried to side-step the challenges this contestation has presented with regards to the functioning of the WTO.

Early history of the WTO

The World Trade Organization (WTO) is a multilateral institution with 164 members as of 2021. It is also a relatively newer institution. The WTO was launched in 1995 and emerged out of the Uruguay Round of multilateral trade negotiations spanning the years 1986–1994. Up until this point it was the General Agreement on Tariffs and Trade (GATT) that coordinated global trade in the post WWII period. However, beginning in the 1980s there was a growing view particularly in the United States that services—financial, tourism, education, and health—should be included in the global trading rules. But while financial services and tourism have been commodified in systems of exchange, sectors like education and health in many parts of the world was considered a public service and not a consumer good.

The interest in transforming education into the services sector is because services make up a growing share of the economies of the developed world and is valued in the trillions of dollars. As the WTO notes:

While services currently account for over two-thirds of global production and employment, they represent no more than 25% of total trade, when measured on a balance-of-payments basis. This—seemingly modest—share should not be underestimated, however. Indeed, balance-of-payments statistics do not capture one of the modes of service supply defined in the GATS, which is the supply through commercial presence in another country (mode 3). Furthermore, even though services are increasingly traded in their own right, they also serve as crucial inputs into the production of goods and, consequently, when assessed in value-added terms, services account for about 50 percent of world trade

WTO website (2021a).

The GATT remained part of the WTO and was now joined by two new agreements: the General Agreement on Trade in Services (GATS), and the Trade Related Intellectual Property Rights (TRIPS). Both have important implications for education. In GATS all levels of education, from early years learning to adult and higher education, were specified as up for negotiation to be included in education services sector negotiations with Member States. Where knowledge took a proprietary form, it was increasingly regarded as intellectual property and included in TRIPS. This refers to music, books, research papers in journals, and so on.

The WTO's mandate was to "... formalize, deepen and widen an international system of trade regulation. It was also to bring greater coherence in global policymaking by drawing together the work of the WTO with that of the IMF, the WB as well as to

develop relations with other bodies” (Wilkinson, 2002: 129). Yet there were problems in defining services using the language of trade in goods. As Winham points out:

... services are processes, defining them is difficult, unless a strict functional definition is employed. ... The tasks for the negotiators at the Uruguay Round were to incorporate GATT principles of transparency, national treatment and reciprocity, as well as newer principles such as market access, into areas of trade that were conceptually dissimilar from trade in goods

Winham (2005: 101).

The result was a services agreement that was “... not yet complete, not terribly user friendly, with a complex geometry and *ala carte* obligations set against the backdrop of near universal coverage and sovereign immunity in liberalization matters” (Sauve, 2002: 3).

What does GATS involve?

Members who joined the WTO were required to opt into a series of legally-binding rules and a built-in agenda to engage in on-going negotiations leading to progressive liberalization. And while all sectors of education were included in GATS, it was then up to the negotiators as to which modes and sectors might be listed for agreement (Verger and Robertson, 2012). This is called a “positive list” approach.

GATS has three components. The first component is a framework of rules that lays down the general obligations governing trade in services, much as it does with the GATT. It lays out the basic elements such as: (i) *most-favored nation*—treating other nations equally so that one nation is not favored over another though some exemptions will apply; (ii) *national treatment*—treating foreigners and locals equally especially in the context of those operating within the boundaries of the country; (iii) *market access*—eliminating or reducing various kinds of tariffs that limit the possibility for some entities to get access to the market; and (iv) *transparency* of all domestic regulations at all levels of government and including also non-governmental bodies exercising governmental power or authority. This last element is important to make visible the various kinds of regulations which might act as a barrier to a provider of services regarding entry to the market (Robertson et al., 2002).

In a second component there are very detailed annexes on specific services sectors like education. A third component consists of “schedules of commitments” detailing the liberalization commitments of each WTO member.

The GATS distinguishes between “four modes of supply” through which services can be traded (see Box 1). These are: Mode 1: cross border supply; Mode 2: consumption abroad; Mode 3: commercial presence; and Mode 4: movement of natural persons. These are odd sounding categories from the point of view of education—largely as they borrow from the language of the GATT. Briefly, when applied to education, cross border supply refers to services that are delivered by a supplier in another country so that it moves across a national border—such as distance education. Consumption abroad refers to students from one country paying fees to study in another country. Commercial presence refers to an education provider, such as a university based in one country, establishing itself in another country and supplying education—such as a branch campus. Movement of natural persons refers to academics and researchers who are providing a service when they are working in another country on a temporary basis.

An exemption is only granted to a services sector under GATS Article 1.3; when it is supplied “... in the exercise of governmental authority” (Art 1.3(b)), and is “supplied neither on a commercial basis nor in competition with one or more service providers” (Art. 1.3 (c)). Governmental authority is interpreted in a narrow sense (WTO, 2021; see also Krajewski, 2016: 2–3). However, the idea of

Box 1 General Agreement on Trade in Services -modes of supply, definition and example

Mode of supply	Definition	Example
Mode 1 Cross border supply	Supply of a service from one MS territory to another	Distance education Education software
Mode 2 Consumption abroad	Supply of a service in the territory of one MS to the service consumer	Fee paying international student studying abroad
Mode 3 Commercial presence	Supply of a service by a service supplier in the territory of another MS	Branch campus Language training company
Mode 4 Movement of natural persons	Supply of a service by a service supplier of one MS through presence of natural persons	Professor, teachers, researchers working abroad on a temporary basis

governmental authority is only applicable to core sovereign functions and not applicable where there is any remuneration, or where there is evidence of competition, meaning there are one or more service suppliers. Were a commitment to be made, a reversal (such as renationalizing) would be difficult in that other WTO Members would have the right to refuse such a reversal, demand a similar commitment in another sector, demand compensation for the reversal, or ultimately to reject the reversal of the commitment. This has happened where a country wanted to make a reversal to health commitments made by a previous government that had liberalized and privatized their economy (see [Robertson et al., 2017](#)).

Education—from public services to a tradeable services sector

Public services in many countries around the globe have been encouraged to implement market mechanisms, such as competition and choice (quasi-markets), or because they are existing markets made up of buyers (families/institutions/countries) and of sellers (firms/countries) ([Verger et al., 2016](#)). Most education systems are mixed systems, with a private sector playing a role (even if their education offer—such as the type of degree—differs) ([Robertson et al., 2002](#)). Furthermore, fees exist even in public schools. Education systems were therefore unlikely to have qualified for the status of “governmental authority” flagged under GATS Article 1.3 at the time that GATS was launched in 1995. This is even less likely today given the extent to which education systems—as public services—have been unraveled and now include many non-state actors supplying a wide range of education activities ([Robertson et al., 2012](#); [Verger et al., 2016](#)).

[Kelsey \(2008: 234\)](#) argues that including education as a public service or public utility was off the radar during the Uruguay Round, largely because education was widely understood by governments and citizens to be a public good. However, this norm was challenged in the 1990s when (i) trade departments saw the potential value of the sector, (ii) entrepreneurs saw new opportunities opening up as education became more and more central to developing competitive economies, and (iii) families were under growing pressure to buy additional education services and other opportunities to ensure social mobility ([Brown et al., 2011](#)).

Despite efforts to list education commitments across the WTO members beginning in 1995, the GATS negotiations proved to be hugely controversial. Each annual meeting of the WTO was marked by protest. The iconic “Battle in Seattle” in 1999 was hugely resonant of the times. Between meetings, campaigners organized and promoted education as a human right, and not a commodity, though some countries did make education commitments. [Verger \(2010\)](#) shows that the extent of commitments was quite limited, and the density of these commitments quite varied. Those opposing the inclusion of education in GATS enrolled a competing narrative; of education was a human right and entitlement, as recognized in international instruments. The most prominent here was the International Declaration of Human Rights (1948), and the Convention of the Human Rights of the Child (1959).

In a specially commissioned report on GATS, the Special Rapporteur for Education, Katerina Tomasevski, pointed out that the “... rapid development of international trade law necessitated a decisive reaffirmation of education as a human right” (2001: 5). This led her to observe:

... the liberalization of trade in services, without adequate government regulation and proper assessment of its affects, can have undesirable effects. Different service sectors require different policies and time frames for liberalization and some areas are better left under governmental authority (p. 20) ... While the WTO Agreements provide a legal framework for the economic aspects of the liberalisation of trade, they focus on commercial objectives. The norms and standards of human rights provide the means of providing a legal framework for the social dimensions of human rights ... A human rights approach to trade liberalisation emphasizes the role of the State, not only as negotiator of trade rules and setter of trade policy, but also as duty bearer for human rights

[Tomasevski \(2001: 10\).](#)

In her Report, the Rapporteur pointed to ways in which GATS would exaggerate social inequalities in education. For example, in Mode 1 (cross border supply) those advantaged by virtual suppliers were those who had the necessary infrastructures, such as the Internet, to access education. Mode 2 (consumption abroad) could lead to the introduction of a dual market of fees and would exaggerate inequalities if governments did not have a way of ensuring cross subsidization. Under Mode 3 (commercial presence), if user fees were introduced, then education would likely become more expensive, and it would not always be easy to regulate foreign suppliers. Finally, under Mode 4 (presence of natural persons), while it might enable knowledge transfer and remittances to be returned to the sending country, it could also lead to brain drain. Locking countries into their schedule of commitments, and penalizing them for seeking reversals, revealed the tension between commercial interests and societal interests. In making this point the Rapporteur observed:

From a commercial perspective, holding countries to their commitments to liberalise is important to ensure transparency and predictability in international trade and the payment of compensation is a legitimate commercial response to the settlement of disputes. From a human rights perspective, however, the focus is less on predictability and more on the need for flexibility to modify or withdraw commitments to liberalise services where experience demonstrates that a commitment constrains or limits the enjoyment of human rights. The need for flexibility is particularly relevant for developing countries given that they are in a dynamic process of building infrastructures. ... Moreover, while compensation to affected parties might be appropriate in some cases upon withdrawal of commitments, a human rights approach would question whether states should be sanctioned for taking action to protect human rights

[Tomasevski \(2001: 28\).](#)

Other influential commentators, including academic activist Jane Kelsey (2003, 2008), argued that GATS would have important consequences for the marginalized and poorer sectors of societies, including indigenous peoples, women, and girls, as GATS rules did not provide for any recognition of indigenous rights or cultural identity. And Frederiksen (2004) pointed out, the irony was that if a child does not get the education he/she is entitled to according to the Convention of the Human Rights of the Child (para 28 and 29), there is no way to force that government to meet its commitments. However, if a company trading in education services loses its right to trade in a particular country, that country where the company is based will have, according to the WTO rules, the right to compensation. Rules concerning trade seem to be much stronger in international law than rules concerning human rights. Similar concerns were aired, and statements issued, by the higher education sector. These concerns, along with growing disquiet among many developed and developing countries as to what free trade really meant for them, resulted in WTO negotiations grinding to a halt by 2005 (Kelsey, 2008).

In sum the inclusion of education as a services sector in the WTO's General Agreement on Trade in Services in 1996 marked a radical shift in understanding as to the purpose of education in society. From being a means of socializing individuals into a society as both economic and political citizens, GATS reframed education as a global services sector to be governed through global rules of trade. And while some countries had committed aspects of education sectors, well organized global protests around including the services sectors more generally resulted in limited progress by the GATS negotiators. This, if course, should not be read off as a success for the protestors. What emerged out of this impasse was that trade negotiators took a bilateral rather than a multilateral route.

Forum shifting—Regional Trade Agreements

Following the challenges facing the GATS negotiations, free trade advocates and corporations shifted their efforts to new forums to keep advancing the “trade” agenda. Forum shifting took a number of forms. Using the UK as a lens (and mindful that the UK was still in the European Union and the trade negotiators were in the European Commission's DG for Trade and Industry), we can see multiple movements and scales of activity.

At the national level, interested education services providers approached nationally located regulatory bodies to make claims to enter the sector (Robertson and Komljenovic, 2016). At the regional level, agreements like European Union's Directive on Services were advanced and finally agreed to in 2006. At the bilateral level, a significant number of Bilateral and Preferential Trading Agreements were advanced (Horn et al., 2010). In combination, these different forums to advance liberalization moved the trade agenda forward.

Here I will focus specifically on what are called Preferential or Regional Trade Agreements (RTAs). Horn et al. (2010) show that RTAs grew significantly in number when the World Trade Organization was launched—from 50 in 1995 to around 349 in 2021 (WTO, 2021b). A large part of this expansion involved agreements where either the US or the EC are at the center of the partnership. This has led some to argue that though the EC and US account for no more than 40% of world GDP and world trade, they account for 80% of the rules that regulate the functioning of world markets, and thus can be regarded as “regulators of the world” (Horn, 2010).

However, these newer RTAs are different from those formed prior to 1995. The earlier group focused on trade in goods and the reduction of tariff rates and quotas, while the post-1995 RTAs now included the regulatory aspects of services, and trade-related intellectual property (Sauvé, 2014). Commentators argue what is notable about this new generation of RTAs is that they tended to go further in their coverage of regulatory issues than under the WTO, and now include provisions and mechanisms such as: “... investment protection, competition policy, labor standards and protection of the environment” (see Horn et al., 2010: 1566). In a 2014 Trade and Development Report, the United Nations Conference on Trade and Development (UNCTAD) signaled concern over the steady erosion of national “policy space” as a result of RTAs “... some of which contain provisions that are more stringent than those covered by the multilateral trade regime, or they include additional provisions that go beyond those of the current multilateral trade agreements” (UNCTAD, 2014a: 19). They go on:

Provisions in RTA have become ever more comprehensive and many of them include rules that limit the options available in the design and implementation of comprehensive national development strategies. Even though these agreements remain the product of (often protracted) negotiations and bargaining between sovereign States, there is a growing sense that, due to the larger number of economic and social issues they cover, the discussions often lack transparency and the coordination—including amongst all potentially interested government ministries

UNCTAD (2014a: 19).

UNCTAD (2014b) also shows that there has been a steady rise in investor disputes cases; by 2021 UNCTAD (2020) reported this figure had gone over 1000 cases. UNCTAD also notes that the ad-hoc tribunals established to adjudicate disputes tend to display a pro-investor bias, and that there is a lack of transparency around the arbitration process. Research by van Harten (2012: 6) confirms this. Drawing on empirical evidence, he shows that there is systematic bias in the arbitration resolutions in favor of the major Western capital-exporting states. The secretive nature of the arbitration process, and the lack of any requirement to consider precedent, allows plenty of scope for creative adjudications.

The sharp rise in contentious arbitrations is interpreted as the result of companies learning to exploit ISDS clauses, going as far as buying firms in jurisdictions where they apply simply to gain access to them. Shopping around for firms in jurisdictions where they can make a claim is not just a case of “forum shopping”—a term used to describe firms who look around for agreements that will give them most “return” regarding potential investor pay-outs—but also of “place hopping”; that is, looking for those places, or jurisdictions, that enable them to launch a more lucrative claim—much as tax havens and states with a low taxation floors have been used. This uneven terrain is the result of different degrees of economic and political liberalization in countries. Most importantly, ISDS gives exclusive rights to foreign investors—rights not available to domestic investors or citizens. Foreign investors are not even obliged to exhaust the domestic legal system before turning to ISDS to make a claim.

The Trade in Services Agreement (TISA)

The origins of the Trade in Services Agreement (TISA) come from frustrations by a selective number of WTO members in the lack of progress by the WTO in negotiations including education. The TISA is thus an ad hoc coalition of WTO member countries willing to advance services negotiations outside of the WTO formal structures. These negotiations were initiated by the US and Australia (Peng, 2013: 614); both countries who have a huge stake in the global education services market. Calling themselves the “Really Good Friends of Services,” the parties to TISA are negotiating services independently of the GATS/WTO though they argue that they hope to inform them (Peng, 2013: 613). More broadly, the “Really Good Friends of Services” are predominantly OECD countries. Historically, the inclusion of services in the WTO was a concession for developing countries (Gould, 2014: 9) and Gould argues that the “... TISA negotiations are essentially a replay of the negotiations that produced the GATS, but this time without the delegations [developing countries] present in the room that might have pushed back against the more extreme demands of the transnational services lobby” (Gould, 2014: 9).

The negotiations formally began in March 2013, and a basic text was agreed by September of that year indicating which service markets participants were prepared to open, and to what extent. By August 2016 there were 23 negotiating parties: Australia, Canada, Chile, Chinese Taipei (Taiwan), Colombia, the EU (representing its 28 member states), Hong Kong China, Japan, South Korea, Mexico, New Zealand, Norway, Pakistan, Switzerland, Turkey, the United State. Costa Rica, Israel, Iceland, Panama, Peru and Mauritius. By November 2016, there had been 21 rounds of negotiation (EC, 2022). Sauve (2014: 7) notes that TISA was being advanced without the formal ascent of the broader WTO membership, negotiations were taking place in Geneva but outside of the perimeter of the WTO, it was at arm’s length to the WTO Secretariat, and the RGFS do not allow observers in, including the WTO.

The ultimate outcome in Sauve’s (2014: 5) view, is that TISA is likely to face the very same fate of the services negotiations under the Uruguay and Doha Rounds; a stand-off between the developed and developing world. Given the secrecy surrounding the negotiation of the TISA, the final outcome regarding how education will be included remains uncertain. The “negative list” principle for national treatment adopted in TISA means in principle, no service is excluded. As under the WTO/GATS, only those services that are services supplied in the exercise of “governmental authority”—see Article 1.3(b) (WTO, 1999)—are exempted.

The exclusion of such countries was broadly political. There was a widely shared view that the EU and US were attempting to create a set of rules for the global economy that work in their specific interests (cf. Bhagwati et al., 2015) regarding the share of global trade. Yet it is clear China (outside the TISA group) has become a major force in the global economy—though not a major exporter of education services. The USA continues to dominate global trade figures, while Germany has moved from second to third place as the powerhouse for the EU (WTO, 2021). Overall, however, TISA can be viewed as largely a failed project. Since 2016 there has been very little activity suggesting that like the WTO negotiations this project has stalled.

Trade related aspects of intellectual property rights

TRIPS was the second major agreement launched at the time of the WTO’s inception alongside GATS. TRIPS has important implications for education systems as producers and users of knowledge, and particularly so following the extent to which knowledge is being digitized.

A key claim of contemporary knowledge-economy policy agendas is that knowledge is Intellectual Property (IP) protected through law (e.g., patent, trademark, and copyright law). New digitally mediated innovations have enabled the creation of new products and services, with data and information now a form of currency (Curtis, 2012). Information is an important building block for new products and services around the creation, storage, analysis, sharing and innovative recombining of data and information. For instance, large student satisfaction surveys about their university experience are carried out by companies like i-graduate or the Princeton Review; these firms gather data and sell information to students and universities regarding where best to study. They also sell information back to the sector to enable institutions to see how their competitors are doing.

Curtis (2012) argues there have been rapid changes in intellectual property and policy due to the intersection between international trade and the numerous international trade agreements that have been bought into force more recently. Driving this dynamic are the developed countries who export intellectual property rights (music, writing, patents, trademarks, and so on) and who pressure other countries to change or create intellectual property rights–related policies and regulatory instruments to accommodate the exporter’s interests. As a result, have been major changes in intellectual property (IP), law and policy over the past 20 years. Realizing the terms of conditions for “rents” depends on legal infrastructures that bind producers and users into a contractual relationship.

Yet as [Curtis \(2012: 8–9\)](#) points out, there is a tension between the rights of *creators* and the rights of *users*. Lawyers are largely concerned with the rights of creators, turning those rights into intellectual property rights. Economic policymakers are more interested in determining the cost-benefit ratio between the rights of creators and the rights of users in terms of innovation, economic growth and productivity. This is because in particular circumstances, limiting access to products and services through putting a price on them, in turn slows down innovation. Similarly, copyright protections aimed at creators can limit creativity in that it makes it difficult for some texts to be easily accessed in sufficient volume by users. How closed and how open systems of regulation are to ensure that the creators of things (goods, services, artifacts, performances) get some financial return on their labor, is pitched against the costs or not for users.

Since the mid-1990s, global trade and intellectual property rules have come together with a view to regulating them through the World Trade Organization's Trade-Related Aspects of Intellectual Property Rights (TRIPS) (1996). TRIPS was later followed by the inclusion of IP standards and obligations in many regional and bilateral Free Trade Agreements trade agreements beginning in 2000.

Though many countries have been developing their own infrastructures, the United States dominates the digital economy and digital trade; that is, commerce in products and services delivered using the Internet in cross border information flows ([Aaronson, 2016](#)). The United States is also behind efforts to develop trade rules to govern cross border flows. Ranked by market capitalization, the United States is home to eleven of the fifteen largest internet-based businesses, including Apple, Google, Facebook, Amazon, LinkedIn and Twitter. China is home to four; Alibaba, Tencent, Baidu and JD.com ([Aaronson, 2016: 7](#)). No companies from Brazil, Canada, the EU28, India, Japan, or Korea are listed among the top fifteen digital corporations.

The stakes of who regulates the Internet has increased since 2008 following the global financial crisis, with governments around the world viewing the Internet and digital technologies as an economic savior. The global pandemic as a result of Covid19 in 2020 will have escalated this given the movement of education delivery to online learning. However, countries also worry about the dominance of the United States in issues of competition, human rights, and infrastructural policies. Education sectors include individuals and organizations (e.g., administrators, teachers, researchers, students, test and textbook producers, performance artists) who are both creators and users of intellectual property. As creators, their outputs may or may not be owned by the institutions in which they work. They may well sign away their rights as creators to publishers for a fee, or a share of the royalties.

More broadly, the global education publishing market is massive, and accounts for 15% of global media and entertainment ([Wischenbart, 2016: 4](#)) estimated to be worth over US\$1 trillion. The total world market value of publishing education in 2016 was estimated to around €95.6 billion and growing ([Wischenbart, 2016: 2](#)), with markets in the USA and China taking the major share. Figures reveal that the Chinese publishing industry is expected to grow further as a result of domestic consumption within China, and as a result of aspirations for a better education among the middle classes in urban regions ([Wischenbart, 2016: 3](#)). Recently e-books have revived a shaky book market, particularly in the US and the UK, both with strong export profiles. The global pandemic in 2020/21 has also created a major boost for electronic books given that access to physical books was made difficult if not impossible in some countries.

As users of digitally based intellectual property, many education providers have invested in virtual learning platforms to manage their teaching and learning interfaces. This applies to physically located institutions, as well as purely on-line providers. Massive Open Online Courses (MOOCs), which emerged around 2010, now sit on platforms enabling learners from around the world to connect to bespoke courses offered by universities, sometimes with a small fee collected by the MOOC platform owners from the students who enroll in exchange for a certificate of participation. Start-up companies like LinkedIn have been purchased by the technology giants like Microsoft, not only connect data on the learning and employment profiles of individuals but use this information to create new services around recruitment ([Komljenovic, 2019](#)). Many education institutions also use Google's Gmail system, as well as other services, such as DropBox. The Cloud is also routinely used to store education-related data from libraries of papers to large datasets. Facebook is used by education institutions to bring groups of learners and other study groups together, as well as being a means of keeping in touch with alumni. However, these global companies do not reach into each country; China for instance places limits on the operation of digital firms inside its borders—such as Facebook and YouTube.

Conclusions

I began by suggesting that educators are likely to look askance at the pairing of education systems with the World Trade Organization in that what is implied is that education might be part of the global trading system. Yet it is evident that there has been an ongoing project since the 1980s to bring services into the regulatory systems that operate globally, and that education should be treated as sector to be included in the ongoing negotiations. The WTO's two agreements, GATS and TRIPS are particularly relevant to education. It is also clear that WTO Members have been rather unwilling to negotiate over education, and the level of commitment has been quite low. In essence what the developed countries were seeking were ways of entering another country's education sectors either through providing education services, ensuring education professionals could operate on a temporary basis in another country, or it can provide services—such as on-line learning—across national borders. The critical thing was that foreign providers would be treated the same as local's which could also include eligibility for local provisions, such as access to state funded student loans.

The WTO has nevertheless not managed to get a high level of buy-in regarding commitments to open up. This is not to say that of liberalization of education has been stopped in its tracks. Far from it. Other means of liberalizing education system to open them to

a wide array of non-state actors headquartered in another country have been pursued, such as the Preferential or Regional Trade Agreements, which commentators argue have been even more aggressive than the WTO disciplines, or simply letting the logic of the market shape the dynamics of the sector. Where the WTO rules will matter if education has been committed is how to wrest education back from the logics of profit and market expansion were other politics to emerge which challenged this framing.

References

- Aaronson, S.A., 2016. The Digital Trade Imbalance and its Implications for Internet Governance, Paper Series No. 25. Global Commission on Internet governance, Waterloo.
- Bhagwati, J., Krishna, P., Panagariya, A., 2015. Where is the world trade system heading? *Adephi Rev.* 17–38. <https://doi.org/10.1080/19445571.2014.1019714>.
- Brown, P., Lauder, H., Ashton, P., 2011. *The Global Auction*. Oxford University Press, Oxford.
- Curtis, J.M., 2012. Intellectual Property Rights and International Trade: An Overview. Paper Series No. 3. Global Commission on Internet Governance, Waterloo.
- EC, 2022. [https://www.europarl.europa.eu/legislative-train/theme-a-balanced-and-progressive-trade-policy-to-harness-globalisation/file-trade-in-services-agreement-\(tisa\)](https://www.europarl.europa.eu/legislative-train/theme-a-balanced-and-progressive-trade-policy-to-harness-globalisation/file-trade-in-services-agreement-(tisa)).
- Fredrikssen, U., 2004. GATS, education and teacher union policies. *Eur. J. Educ. Res.* 3 (2), 415–441.
- Gould, E., 2014. *The Really Good Friends of Transnational Corporations Agreement*. Public Services International, Geneva.
- Horn, H., Mavroidis, P., Sapir, A., 2010. Beyond the WTO? An anatomy of EU and US preferential trade agreements. *World Econ.* 33 (11), 1565–1588.
- Kelsey, J., 2003. Legal fetishism and the contradictions of GATS. *Glob. Soc. Educ.* 1 (3), 267–280.
- Kelsey, J., 2008. *Serving Whose Interests? The Political Economy of Trade in Services Agreements*. Routledge-Cavendish, Abingdon, UK and New York.
- Komljenovic, J., 2019. LinkedIn, platforming labour, and the new employability mandate for universities. *Glob. Soc. Educ.* 17 (1), 28–43.
- Krajewski, M., 2016. Model Clauses for the Exclusion of Public Services From Trade Agreements. *European Public Service Unions with Gerechtigkeit Muss Sein*.
- Peng, S.Y., 2013. Is the trade in services agreement (TISA): a stepping stone for the next version of GATS? *HKLJ* 43 (2), 611–632.
- Robertson, S., Bonal, X., Dale, R., 2002. GATS and the education service industry. *Comp. Educ. Rev.* 46 (4), 472–497.
- Robertson, S., Verger, A., Mundy, K., Menashey, F. (Eds.), 2012. *Public Private Partnerships in Education*. Edward Elgar, Cheltenham.
- Robertson, S., Tidy, J., Arcas, S., 2017. What Educators Need to Know About Global Trade Deals. Education International, Brussels.
- Robertson, S.L., Komljenovic, J., 2016. Non-state actors and the advance of 'frontier' higher education markets in the global south. *Oxf. Rev. Educ.* 42 (5), 594–611.
- Sauve, P., 2002. Trade, Education and the GATS: What's In, What's Out, What's All the Fuss About? OECD, Paris.
- Sauve, P., 2014. *A Plurilateral Agenda for Services? Assessing the Case for a Trade in Services Agreement (TISA)*. NCCR Trade Regulation, Geneva.
- Tomasevski, K., 2001. Annual Report of the Special Rapporteur on the Right to Education, Submitted to the Commission on Human Rights. Economic and Social Council, New York.
- UNCTAD, 2014a. Trade and Development Report, 2014, Overview. UNCTAD, Geneva.
- UNCTAD, 2014b. Investor-State Dispute Settlement: UNCTAD Series on Issues in International Investor Agreements II. UNCTAD, Geneva.
- UNCTAD, 2020. Investor State Dispute Settlement Cases Pass the 1,000 Mark: Cases and Outcomes in 2019. Accessed at. <https://unctad.org/system/files/official-document/diaepcbinf2020d6.pdf>.
- Van Harten, G., 2012. Arbitrator Behaviour in Asymmetrical Adjudication: An Empirical Study of Investment Treaty Arbitration, Osgoode Hall Law School. Research Paper 41/2012. New York University, New York.
- Verger, A., Robertson, S., 2012. In: Robertson, S., Verger, A., Mundy, K., Menashey, F. (Eds.), *GATS and Education, Public Private Partnerships in Education*. Edward Elgar, Cheltenham.
- Verger, A., Lubienski, C., Steiner-Khamis, G. (Eds.), 2016. *The Global Education Industry, World Yearbook 2016*. Routledge, London and New York.
- Verger, A., 2010. *WTO/GATS and the Global Politics of Higher Education*. Routledge, New York and London.
- Wilkinson, W., 2002. The World Trade Organization. *New Polit. Econ.* 7 (1), 129–141.
- Winham, G., 2005. The evolution of the global trading regime. In: Ravenhill, J. (Ed.), *Global Political Economy*. Oxford University Press, Oxford.
- Wischenbart, R., 2016. IPA Global Publishing Statistics. <http://www.internationalpublishers.org/images/data-statistics/2012/ipa-global-publishing-statistics-2.pdf>.
- WTO, 1999. World Trade Organization Statistical Review 2019. Accessed at. https://www.wto.org/english/res_e/statis_e/wts2019_e/wts2019_e.pdf.
- WTO, 2021a. The General Agreement on Trade in Services (GATS): Objectives, Coverage and Disciplines. https://www.wto.org/english/tratop_e/serv_e/gatsqa_e.htm.
- WTO, 2021b. Regional Trade Agreements. https://www.wto.org/english/tratop_e/region_e.htm.

UNESCO: promising education multilateralism?

Yusuf Sayed^{a,c}, Gunjan Sharma^b, and Aditi Desai^c, ^a School of Education and Social Work, University of Sussex, Brighton, United Kingdom; ^b School of Education Studies, Dr. B. R. Ambedkar University Delhi, Delhi, New Delhi, India; and ^c Center for International Teacher Education, Cape Peninsula University of Technology, Cape Town, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	510
Positioning UNESCO's educational agenda: an overview	511
UNESCO's intellectual trajectories	513
Education for democratic redevelopment—the “education fundamental” approach	513
The global education goal setting: from Jomtien to the twin global agendas of the EFA and the MDGs	514
The Dakar EFA and MDG phase	515
The quality turn, the return of lifelong learning and the Covid-19 pandemic	516
The assertion of UNESCO's global education discourse: from the Delors Report to the Futures Report	517
Setting the problem	517
Knowledge and learning	518
Realizing the education vision	519
International cooperation and UNESCO's role	520
Negotiating the politics of multilateralism and global education governance	521
Conclusion	523
References	524

Introduction

The United Nations Educational, Scientific and Cultural Organization (UNESCO), is an agency of the United Nations (UN). As a key UN agency and an International Organization (IO) for education, science and culture, UNESCO has worked in a wide variety of areas including protecting heritage, strengthening biodiversity (especially ocean sciences), fostering freedom of expression, preventing violent extremism, and sustainable development.

UNESCO's stated mandate is to establish the “intellectual and moral solidarity” of humanity to prevent war and build peace.¹ One of the key activities that UNESCO aligns with this mandate is to “foster the free flow of ideas and knowledge sharing”.² UNESCO realizes this vision through three sets of interrelated activities:

- i. Knowledge generation and dissemination through data collection, reports, development of curricula, handbooks, guides and manuals. This function is not different from other IOs, such as the World Bank (WB), which positions itself as a knowledge bank.
- ii. Coordination, evaluation and monitoring of the various international commitments or goals. This is often referred to as the accountability function in education.
- iii. Advocacy through conferences, meetings and policy briefs, referred to by the staff as its “convening power”.³

It has its headquarters, nine education institutions, national and regional offices, networks and institutes across the globe that support these functions in varied ways, as discussed below.

UNESCO's website⁴ identifies its two key priorities which are Africa and gender equality. UNESCO has a presence in all 54 countries of Africa. Its Africa programming concentrates on six flagship areas that include: (1) education systems for sustainable development; (2) science for sustainable resource management and disaster risk reduction; (3) building on culture for sustainable development and peace; (4) freedom of expression and media development; (5) peace and non-violence; and (6) science, technology and innovation for sustainable development.⁵ With regards to gender equality, UNESCO has adopted a three-pronged approach of: (1) integrating or mainstreaming gender across its programming; (2) developing specific programmes on gender; and (3) monitoring and evaluating the progress on gender equality goals (UNESCO, 2019). UNESCO's Gender Equality Action Plan 2014–2021 (revised in 2019) defines gender equality as a human right and lays down a time-bound plan of action for achieving its targets. Much of the organization's programming and publications, especially in its first three decades, have been driven by these priorities.

¹<https://www.unesco.org/en/introducing-unesco> (Accessed on 02.11.2021).

²<https://www.unesco.org/en/introducing-unesco> (Accessed on 02.11.2021).

³Personal communication with UNESCO staff by authors.

⁴<https://www.unesco.org/en/introducing-unesco> (Accessed on 02.11.2021).

⁵<https://en.unesco.org/priorityafrica/flagshipprogrammes> (Accessed on 02.11.2021).

UNESCO is located within a range of diverse IOs comprising various donors/bilateral agencies, specialized agencies of the UN, such as UNICEF and UNDP, and Bretton Woods institutions such as the World Bank (WB), the International Monetary Fund (IMF), international non-government organizations (INGOs), alliances and networks, such as Global Alliance for Vaccine and Immunization (GAVI), the Global Campaign for Education (GCE), and philanthropic foundations such as the Bill and Melinda Gates Foundation. In this complex and diverse set of institutions, UNESCO, as this paper argues, has sought to ensure its hegemony over the global education agenda competing for space with, for example, the WB and the UN Children's Fund (UNICEF) whose education vision and orientations differ from those of UNESCO. What sets UNESCO apart from most other IOs and UN agencies is that it is governed by its member states. This means that each country that is a member has an equal say, a rare example of democratic participation in policy making in the UN system.

In this context, this paper examines the functioning of UNESCO as part of the international global architecture for development and as part of the field of global policy, in general, and education policy, in particular. We pay particular attention in this paper to the intellectual evolution and trajectory of UNESCO's education ideas and how it constitutes its space in the multilateral global field and the tension and contradictions that result from this. In particular, we seek to provide an ideational genesis of the education policy and programmes that populate UNESCO's work showing how its vision shifts and alters shape within specific conjunctural and spatial moments and periods as a result of its interactions with its own members and in relation to other IOs.

This paper draws on a general review of the UNESCO Digital Library⁶ over nearly eight decades of its existence. This library contains a large volume of publications about education. It has directly produced 127,381 publications and has sponsored 2659 publications, most in the English language. These publications include many types of documents predominantly program and meeting documents (77,185), books (12,821) and articles (17,988). The Digital Library classifies conference material (47,305) as the most frequently published type of UNESCO and UNESCO-sponsored publications. These publications are cataloged under a large variety of topics with the most prominent during 2000–2021 being higher education (8556). Geographic regions covered by these publications include Latin America (5506) and Africa (5016) as the two regions with the highest number of publications. Navigating the Digital Library is challenging because of the volume of publications and also because the topics that are cataloged are not mutually exclusive. Analysis indicates that these publications could be more coherent if they were classified by tracking UNESCO's intellectual history as we do below using particular publications to explain UNESCO's vision of education over time.

The paper is organized in five sections including this first introductory section. The second section positions UNESCO's educational agenda in the larger context of its emergence, its centrality in UNESCO's programming, and vis-à-vis other IOs. The third section traces the intellectual trajectory of UNESCO's educational agenda around certain key phases in its history. The fourth section builds on this tracking, and engages with two major intellectual reports of UNESCO to show how significant shifts in these reports are reflected in the practice of education. The fifth section juxtaposes the changes in the educational discourse with the flash points in UNESCO's multilateral politics. It also investigates how UNESCO's educational agenda is positioned as a UN agency and among the many IOs in the sector.

Positioning UNESCO's educational agenda: an overview

UNESCO originated from the League of Nations and the Committee on Intellectual Co-operation in 1922. It emerged from inter-governmental cooperation for the rebuilding of education, science, and culture in a post-World War I world. In particular, UNESCO was an outcome of the Conference of Ministers of Education of the Allied Governments and the French National Committee (CAME) that was first convened in London on November 16, 1942, and continued to meet until 1945. The first meeting was organized by the Board of Education of the United Kingdom (UK) and the British Council, with the Ministers of Education of Belgium, Czechoslovakia, France, Greece, the Netherlands, Norway, Poland, and Yugoslavia. The intent was to initiate dialog on the concerns collectively experienced by these countries, particularly around education, both during wartime and afterward (De Capello, 1970). Based on the deliberations at CAME, the governments of France and the UK, called for a United Nations conference for the establishment of an International Organization for Education and Culture, during November 1–16, 1945, in London. This conference resulted in the drafting and adoption of UNESCO's Constitution. The preamble of the Constitution states, "since wars begin in the minds of men [sic], it is in the minds of men [sic] that the defenses of peace must be constructed".⁷ UNESCO's mission is "to build peace through international cooperation in Education, the Sciences and Culture".⁸ Niebuhr (1950:4) divides UNESCO's initial functions into three categories: (a) bridging the gaps in the free flow of ideas in the world community; (b) bringing the assistance of the developed nations to bear upon the educational problems of the developing nations; and (c) overcoming cultural and inter-governmental misunderstandings.

Given the emphasis on education at its pre-conceptualization stage, UNESCO's vision, discourse and programming have prioritized education. This makes UNESCO a unique agency of the UN in the education sector compared to other specialized agencies, for instance, the WB Group and the IMF that added education as an addendum to their primary agenda of global economic growth. It is also different from UN programmes and funds that have overlaps with the education sector, such as the UN Development

⁶<https://unesdoc.unesco.org/library> (Accessed on 25.11.2021).

⁷<https://www.unesco.org/en/introducing-unesco> (Accessed on 24.11.2021).

⁸<https://www.unesco.org/en/introducing-unesco> (Accessed on 24.11.2021).

Program (UNDP) and the UN International Children's Emergency Fund (UNICEF), in terms of the centrality, aims and scope of its educational agenda.

UNESCO's institutional architecture has eight Category 1 institutes or specialized departments that focus on education, as compared to one each in science and statistics. These eight education institutes include:

- The International Bureau of Education (IBE)
- The International Institute for Education Planning (IIEP)
- The UNESCO Institute for Lifelong Learning (UIL)
- The Institute for Information Technologies in Education (IITE)
- The International Institute for Higher Education in Latin American and the Caribbean (IESALC)
- The International Institute for Capacity Building in Africa (IICBA)
- The Mahatma Gandhi Institute of Education for Peace and Sustainable Development (MGIEP)
- The UNESCO Institute for Statistics (UIS)

Out of the program implementation arm of UNESCO's 81 Category 2 Centers, 15 have a dedicated focus on education. According to UNESCO (2021b:8), education has the highest percentage of staff (45%) in the program sectors, followed by the natural sciences (24%) and culture (16%). It currently has 400 staff members of which 120 work at the headquarters in Paris, France, while the rest work in education programmes in its 53 field offices. UNESCO is also the only UN agency that has a global network of 199 National Commissions set up by the member states to support cooperation on different areas of work, including education. UNESCO also launched the UNTWIN/UNESCO Chairs Program in 1992 to develop an international network for knowledge generation and sharing by creating Chairs in 850 institutions in 117 countries. The Chairs are seen by UNESCO as bridges between policy development, research and global inter-cultural understanding. UNESCO is financed by compulsory contributions of the member states with Japan, China and Germany as the top three contributors.⁹

Compared to other IOs that work in the education sector, there are also other distinctive structural and organizational elements of UNESCO that stem from its foundational aim of rebuilding the post-war world. Historically, as an intergovernmental multilateral organization, UNESCO's focus has involved conventions, agreements, recommendations, and declarations, at an intergovernmental level. It was not conceived as a funding organization or as a forum for delivering the agenda of alliances of developed countries. In the context of its multilateral origins, UNESCO's democratic foundational structure has three principal organs: (a) the General Conference which is the principal policy decision making forum of UNESCO with representatives from all the member countries; (b) the Executive Board elected by the General Conference for supervising the agreed upon mandates which, until 1993, comprised elected individuals rather than nominees from states; and (c) the Secretariat headed by the Director General. The decision making in the General Conference and the Executive Board is based on discussion and debate over each new program proposal where each of the member country representatives has the right to express concerns or objections. Each member country of UNESCO has the right to vote in the General Conference and in the election of the Executive Board that has representation from different electoral groups (Hite and Bray, 2010:688–693). This is in contrast to some other UN agencies and IOs,¹⁰ for example, the WB, where the voting power privileges the developed nations based on their financial strength.

The democratic structure of UNESCO not only shapes its education policy, but also frames its politics and functional challenges. From 20 members in 1946, currently, UNESCO has 193 member countries and 11 associate members.¹¹ The membership fluctuates as an outcome of multilateral politics between the member nations (as discussed below), significant events in the world history, and international relations (such as the collapse of the Soviet Union). As more states join UNESCO, particularly the developing countries, its focus has shifted toward education in the Global South. This shift has caused North-South tensions that have directed its future.

As a UN agency, since 2001, UNESCO has participated in the observance of UN International Decades that are relevant to its institutional mandate and has assumed the role of a coordinating and monitoring agency or a "convening body". The UN decades that UNESCO has observed are as follows:

- 2001–2010: International Decade for a Culture of Peace and Non-Violence for the Children of the World
- 2003–2012: United Nations Literacy Decade
- 2005–2014: Second International Decade of the World's Indigenous People
- 2005–2015: International Decade for Action, "Water for Life"
- 2010–2020: United Nations Decade for Deserts and the Fight against Desertification
- 2011–2020: United Nations Decade on Biodiversity
- 2013–2022: International Decade for the Rapprochement of Cultures
- 2014–2024: United Nations Decade of Sustainable Energy for All
- 2015–2030: United Nations Decade of Education for Sustainable Development
- 2015–2024: International Decade for People of African Descent
- 2021–2030: Decade of Ocean Science for Sustainable Development

⁹<https://en.unesco.org/themes/education/about-us> (Accessed on 24.11.2021).

¹⁰Personal communication with official to one of the authors.

¹¹<https://en.unesco.org/countries> (Accessed on 25.11.2021).

- 2022–2032: Indigenous Languages Decade

These UN International Decades have shaped UNESCO's education sector even though they have not always reflected its intellectual interests in areas, such as higher education and teacher education, which are key to its policy, programming, and knowledge functions.

UNESCO's intellectual trajectories

This section describes the UNESCO's educational history by tracking significant shifts over time (see section [Negotiating the politics of multilateralism and global education governance](#)). Particular attention is paid to how UNESCO's multilateral politics have impacted its knowledge production and its position vis-a-vis other IOs in an increasingly competitive arena of global education governance.

UNESCO is the only IO that has simultaneously worked in non-formal as well as formal education that includes adult education through literacy campaigns, school education, tertiary or higher education, teacher education and vocational education. Over time, the program priorities have shifted toward primary education, even though higher education is still included in its agenda ([Mundy and Madden, 2009](#)). This section discusses UNESCO's engagement with universal and non-formal education including higher education.¹²

UNESCO's vision of education includes a human rights perspective. At its founding, UNESCO intended its educational programming to contribute to global peace building and development even though its work in conflict and conflict affected contexts emerges much later in its policy, programming and knowledge production. UNESCO's education vision is based on four phases.¹³ The first phase is a focus on fundamental education orientated toward literacy. The second phase focuses on the period from the Jomtien declaration to the global education agendas of Education for All (EFA) and Millennium Development Goals (MDGs.) The third is the Dakar EFA and MDG phase that includes the dual millennium global education goals and the fourth is a discussion of the SDG, COVID-19 and lifelong learning.

Education for democratic redevelopment—the “education fundamental” approach

The first phase of UNESCO's intellectual agenda,¹⁴ from its inception to the 1970s, prioritized education as a fundamental human right and placed “education for all” centrally in its constitution ([UNESCO, 1945:5](#)). It launched the “Fundamental Education” program soon after its inception in 1947 that, while including all levels of education, emphasized adult education that was included in Article 26 of the UN's 1948 Universal Declaration of Human Rights. UNESCO's First International Conference on Adult Education (CONFITEA I) in 1949, defined adult education as “the true spirit of democracy and the true spirit of humanity” ([Elfert, 2019:539](#)). This included eliminating racial and totalitarian propaganda through re-education ([Elfert, 2019](#)).

UNESCO's initial approach to literacy has been critiqued in the literature from a post-colonial perspective as it adopted Western notions of education. For instance, [Wagner \(2013:21\)](#) states that

... UNESCO often found itself adopting the exhortatory approach of illiteracy ‘eradication’ (treating illiteracy as a disease), the ‘battle’ against illiteracy, and the comparison of literacy to a ‘light bulb’, in that a person is in the dark until liberated into the ‘light’ of literacy.

Despite this, UNESCO's published documents consistently demonstrate a commitment to a human rights approach in which adult literacy is a core founding priority.

By the 1950s, as UNESCO's membership grew, so did new multilateral priorities that included commitment to “functional literacy” ([Jones, 1988](#); [Mundy, 1999](#); [Limage, 2007](#); [Elfert, 2019](#)) that focused on the formal education system and redefined its conception of literacy. This shift is evident in the Experimental World Literacy Program (EWLP) (co-funded by the WB and the UNDP) from 1966 until the mid-70s which tested and demonstrated “the economic and social returns of literacy” ([UNESCO, 1966:2](#)):

Rather than an end in itself, literacy should be regarded as a way of preparing man [sic] for a social, civic and economic role that goes far beyond the limits of rudimentary literacy training consisting merely in the teaching of reading and writing ... reading and writing should lead not only to elementary general knowledge but to training for work, increased productivity, a greater participation in civil life and a better understanding of the surrounding world, and should ultimately open the way to a basic human culture.

¹²We are engaged currently in further research about the genesis and trajectory of UNESCO's interest and involvement in other education priorities such as teacher and teacher education.

¹³The four phase periodization partly draws on and extends [Mundy's \(1999\)](#) work on UNESCO which has periodized the work of UNESCO into three phases. It is important to recognize that this intellectual periodization is not strictly chronological—instead it uses a broad historical period as a basis to sketch the intellectual idea and thoughts that animate the education policy and programming of UNESCO.

¹⁴Based on a review of secondary and primary literature relating to UNESCO from its very inception until the first three decades of its history (founding late 1980s).

While critiqued for its failure to achieve its goals,¹⁵ this approach suggested a shift in UNESCO's education vision to one which privileged the economic utilitarian purposes of education, reflecting the WB's rates of returns and human capital approach to education.¹⁶

With the formal ending of colonization in the 1960s for many countries in Africa, UNESCO paid attention to higher education,¹⁷ in particular, to the inclusion of higher education in national policies, connecting higher education with nation-building and development. During the late 1940s until the 1970s, UNESCO established the International Social Science Council (1948), the International Association of Universities (1950), and the Tokyo based UN University (1975) (Lebeau and Sall, 2011:130). The regional conferences organized by UNESCO on higher education show that, unlike the case of non-formal education, the aim of higher education was the potential of economic returns. This reflected the priorities of the donors as well as the members. However, this donor interest in higher education had declined by the 1970s (Lebeau and Sall, 2011; Mundy, 1999; Limage, 1999).

The publications of UNESCO during this phase suggested a strong focus on a more expansive view of education. Two documents, in particular, signaled this approach: (1) *Learning to be: The world of education today and tomorrow* (Faure Report) (UNESCO, 1972); and (2) the *Declaration of Persepolis* (UNESCO, 1975). These documents, especially the Faure Report, asserted UNESCO's educational agenda that committed to a broad notion of "lifelong learning". Both documents allude to the role of education in nurturing global citizenship while acknowledging the economic aims of education. The Faure report states that

[t]he explicitly humanistic and democratic approach to education that permeates *Learning to be* is hardly recognisable in contemporary discussions, in which the right to permanent education ... seems to have been replaced by the duty to continue to learn throughout one's life ...

Biesta (2021:para 3).

Lee and Friedrich (2011) maintain that the Faure Report has been drawn from the radical ideas of Paulo Freire, Ivan Illich and democratic liberalism that signify UNESCO's assertion of its intellectual autonomy and multilateral character.

The global education goal setting: from Jomtien to the twin global agendas of the EFA and the MDGs

UNESCO's second intellectual phase can be traced to the oil crises of the 1970s and the global shift to the right ushering in a period of new public management politics and economics which reshaped social sector provision and social policy¹⁸ that caused UNESCO to rebrand itself (Mundy, 1999; Elfert, 2019; Edward et al., 2018). This is reflected in the biennial World Education Reports of UNESCO (1991, 1993, 1995, 1999) that presented the state of education in different countries, while avoiding controversies (for instance, see Spaulding, 1992).

First, UNESCO remained committed to an expansive and universal education as represented in its International Literacy Year programmes, *Learning: A treasure within*, also known as the Delors Report 1996 (UNESCO, 1996), and the hosting of CONFENTIA V 1997. The Delors Report sought to establish UNESCO's hegemonic intellectual positioning in education and was a riposte to The World Bank's (1995) report, *Priorities and strategies for education*, which was, according to Elfert (2015b), narrow and reductionist. The Delors Report identified four pillars for social progress: learning to be, learning to do, learning to become and learning to live together, which reasserted UNESCO's earlier position on education and challenged the human capital orthodoxy of development thinking. Biesta (2021:8) notes:

Learning: The treasure within ... can be read as a response to a rapidly emerging alternative policy discourse that was strongly characterised by an economic rationale and a focus on lifelong learning as the development of human capital ... the Delors report made a case for shifting the attention from social cohesion to democratic participation and from economic growth to human development.

Second, UNESCO organized the World Conference of Higher Education 1998, that resulted in the World Declaration on Higher Education (UNESCO, 1998) based on UNESCO's initial core values of ensuring equal access to higher education to all, improving the quality of the provisions, and identifying higher education as a public good. However, UNESCO increasingly turned its focus toward EFA as discussed below.

Third, this intellectual phase marks the beginning of a global movement toward global education goal setting. This has caused a disagreement between IOs about the leadership of this global agenda, especially UNESCO and the WB whose visions and values have frequently competed. This focus on collective global goal setting in education represents a significant departure from the

¹⁵See Limage, 1999.

¹⁶See Benell and Furlong (1998) and Tilak (2005) for an overview and critique of this approach.

¹⁷See also Mundy (1999).

¹⁸The oil crises brought with it the long period of the displacement of Keynesian economics in favor of liberal/neo-liberal economics beginning the 1980s onwards, intensifying in the 1990s, and then reaching the global goals at the turn of the millennium.

consultation of a global education policy and demonstrates how a global framing of education became the driver of education change.¹⁹

This focus on global goal setting occurs in a context in which neoliberal frameworks were shaped by globalization that significantly influenced education policies across the world (Rizvi and Lingard, 2010). The 1990s saw a return of the Structural Adjustment program of the 70s in many countries in the Global South in which state responsibility, support and expenditure on social sector spending was curtailed in favor of the neo-liberal market, privatization approaches to education delivery that included a results-based management, a focus on aid and global goal setting. This is discussed in *Negotiating the politics of multilateralism and global education governance* section of this paper.

The global goal setting approach to development manifested in The World Education Forum held in Jomtien in 1990 (World Education Forum [WEF], 1990) that represented an uneasy and fractious global alliance between different stakeholders, including competing IOs, such as UNESCO and WB, national governments, and Civil Society Organizations (CSOs). The adoption of the Jomtien Declaration and the corresponding *Framework for Action to Meet Basic Learning Needs* by UNESCO, steered its education policy and programming for the next 10 years or more laying the groundwork for its Education for All movement in 2000 (UNESCO, 1990). These processes were in addition to UNESCO's initiatives, like the Delors Report, that focused on the broader aims of education. Thus, within UNESCO, two approaches to educational agenda—one in which education was seen in a broad framework of human rights, and the other that restricted the aims of education to human capital development—were beginning to consolidate.

Mundy (2006) argues that this period of global education was marked by a lack of coordination, decentralization and disorganization between the IOs that officially cooperated on educational programmes, with multiple views of the aims and goals of education. The period could be seen as a “fractious epistemic community” with “very loose coupling between rhetorical commitments and practical activities—creating, in effect, a smorgasbord of priorities and approaches from which donor countries might choose according to their own geopolitical and economic interests” (Mundy, 2006:27–28). What was emerging, albeit in an inchoate and opaque manner, was the discursive constitution of a global hegemony of what is valued and is to be measured in education. This intensified competition between the WB and UNESCO in education policy over the Global South.

The Dakar EFA and MDG phase

The period between the World Education Forum in Dakar in 2000 and 2015 could be identified as the third phase in UNESCO's intellectual history as the custodian of the global education agenda. It sought to bridge the focus on an expansive right based education agenda with a narrower technocratic- and instrumentalist view of education.

The Dakar Conference, led by UNESCO in April 2000, and the adoption of the Dakar Framework at the WEF (2000) provided the first formalized set of six global education goals that established a global education policy agenda with measurable strategic actions and targets.

The six EFA goals are:

(i) expand and improve comprehensive early childhood care and education, especially for the most vulnerable and disadvantaged children; (ii) ensure that, by 2015, all children, particularly girls, those in difficult circumstances, and those belonging to ethnic minorities, have access to and complete, free and compulsory primary education of good quality; (iii) ensure that the learning needs of all young people and adults are met through equitable access to appropriate learning and life-skills programmes; (iv) achieve a 50% improvement in adult literacy by 2015, especially for women, and equitable access to basic and continuing education for all adults; (v) eliminate gender disparities in primary and secondary education by 2005, and achieve gender equality in education by 2015, with a focus on ensuring girls' full and equal access to and achievement in basic education of good quality; and (vi) improve all aspects of the quality of education and ensure the excellence of all so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills.

WEF (2000).

Following the Dakar WEF, the UN General Assembly in September 2000 (UN, 2000) announced the eight MDGs out of which only one directly related to education: “to achieve universal primary education”. The other goal indirectly related to education was to promote gender equality and empower women with a corresponding target to eliminate “gender disparity” in primary and secondary education. The MDG education targets were considerably narrower than the Dakar EFA targets and clearly omitted early childhood education, adult education, and quality education. The MDGs also narrowed down the EFA concept of gender equality to gender disparity. In so doing, the adoption of the MDGs in September 2015 set in motion a twin track global education agenda. This put UNESCO, as the custodian of the six EFA goals, in opposition to other IOs, in particular the WB that was supported by multiple donors and certain CSOs who saw it as the driver of the MDG agenda and a strong advocate of a narrowly defined understanding of education as mere physical access to schools.

Hulme and Scott (2010) trace how the MDGs emerged from a complex negotiated process that involved not just the UN but also the OECD, the World Bank, and the IMF. They elucidate how the MDGs were aligned with OECD International Development goals (IDGs) so that there would be only one set of goals to follow (Hulme and Scott, 2010:297). However, this alignment with the IDGs

¹⁹There were global goals in education before Jomtien. However, Jomtien represents the first formal process by which multilateral agencies, donors, Bretton Woods institutions, national governments and CSOs came together to develop a hegemonic view of what is globally desirable in education.

caused a disagreement between the MDGs and the Dakar EFA goals that resulted in a dual educational architecture with rival planning processes and organizational commitments which exerted parallel and dual pressures on national governments (Sayed et al., 2012:193). These developments, rather than drawing attention to the need for more robust education agenda, increasingly replaced the EFA goals with narrowing priorities in the global education policy.

During this period, UNESCO was given the role of coordinating and monitoring the progress toward these goals with a specific mandate to produce annual monitoring reports. The implications of this for UNESCO are discussed in the next section. However, this role increasingly narrowed the scope for UNESCO to define its own independent intellectual trajectory and reduced its agenda to producing Global Monitoring Reports of global education policy goals. Also, the execution of its monitoring role reveals the narrowing of the education vision evident in the two rival global education agenda, for example, in the use of monitoring indices as monitoring tools as pointed out by Tilak (2005:75)²⁰ who states,

As a whole, the EFA goals refer to six specific dimensions, while the EDI refers to four. By contrast, the Millennium Development Goals require countries to focus in the area of education only on primary completion rates and elimination of gender disparities All of this may create confusion, if not chaos, among policy-makers and education planners in developing countries in their approach to EFA. Economic and finance ministries interested more in the Millennium Development Goals and associated aid may, in turn, pressure education ministries to concentrate on these alone. In fact, the set of six EFA goals can be viewed as a minimum set. Indeed, the EDI and the Millennium Development goals clearly reveal a phenomenon of 'minimization of minimum goals of development'!

As the 15 years of MDGs neared completion, UNESCO published a report titled *"Rethinking education: Toward a global common good"* (UNESCO, 2015) partly to position how it influenced the next generation of global educational goals. This was conceptually at variance from the Faure's and Delors' Reports in terms of excluding lifelong education as a concern, and in recognizing alternate visions of development (Elfert, 2015a). However, it does reaffirm a humanistic approach to education that goes beyond its economic aims and the principle of education as a public and common good (Tawil and Locatelli, 2015). It is difficult to say whether the report has served intellectual purposes as UNESCO has transited to adopting and achieving SDGs.

In the context of higher education, given the internal push from some member countries, UNESCO has re-instated its interest in the sector partnering with the WB. This collaboration resulted in the formation of the UNESCO-WB Task Force on HE that did not last for long. Its report, *Higher education in developing countries: Peril and promise* (UNESCO-WB Task Force on HE, 2000) pointed to the emergence of the twin track approaches in higher education goals. Halvorsen (2016:79) believes this report signifies the "silent compromise" between "UNESCO values" and "World Bank values".

There were several other reports that UNESCO produced during this time in HE such as *Higher Education in a Globalized Society: UNESCO Education Position Paper* (UNESCO, 2004), *UNESCO-OECD guidelines for quality provision in cross-border higher education* (UNESCO-OECD, 2005), *Towards Knowledge Societies Report* (UNESCO, 2005), and *Ranking and accountability in HE* (Marope et al., 2013). These reports were primarily concerned with policies, practices and debates on quality assurance that have emerged from a neoliberal approach to higher education. There is a lack of agreement in the literature on whether or not these works represent a "real" commitment to UNESCO's core principles (for example, Blackmur, 2007; Halvorsen, 2016; Lebeau and Sall, 2011).

This phase in UNESCO's intellectual history in the field of global education sought to position its prominence in education, while other IOs developed rival visions pointing to a tension between a broad-based human-rights approach to education and a narrowed vision of "basic education" that was interpreted as basic primary schooling or "learning competencies" by some countries and implementing agencies (Muller, 2000; Sayed and Moriarty, 2020). This can be seen, for instance, in the 1997 OECD report, *Lifelong learning for all* (OECD, 1997) that interpreted learning as acquiring competencies to meet the needs of the global knowledge economy (Rubenson, 2006; Bagnall, 2000; Biesta, 2021). At the same time, as UNESCO took on the custodian role of monitoring the global goals, it began to frame what gets measured in education. The question that emerged in this phase is whether, in reporting of the attainment of the global goals, UNESCO was measuring that which is valuable and valued in education.

The quality turn, the return of lifelong learning and the Covid-19 pandemic²¹

The adoption of the Incheon Declaration on SDG4 in May 2015 at the World Education Forum may be seen as marking the beginning of the fourth phase for UNESCO's educational agenda. This phase traverses the period of the declaration of the SDGs in 2015 as well as UNESCO's response to the COVID-19 pandemic. During this phase, UNESCO asserted itself more firmly in the global education policy space notwithstanding its tensions and contradictions with other IOs. The SDGs aspire to address an ambitious range of economic, social and environmental issues globally by 2030. The SDG4 makes a commitment to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". While this may signal a return of adult education in global policy as well as UNESCO agenda, critical policy literature on SDGs (Sayed and Moriarty, 2020; Elfert, 2019) delves into the political processes that shape these SDG4 concepts and their targets. Sayed and Moriarty (2020:212) document the complex struggles

²⁰In more recent times, the Global Education Monitoring Report, renamed from its original Global Monitoring Report, dropped the use of the EDI arguably due to the lack of utility and reductionism inherent in such an approach to monitoring.

²¹This we classify as the current phase encompassing both the SDG agenda agreed by the UNGA in September 2015 and the response to the Covid-19 pandemic.

during 2013 and 2015 with multiple intergovernmental meetings, such as in Dakar (2013), Muscat (2014) and Incheon (2015), between different alliances of stakeholders (mainly the MDGs and EFA alliances) negotiating their claims on the meaning of “quality”. The study has identified how SDG4 and its targets are based on a “pseudo-technicism” that omits “the rights-based understandings of quality, including inclusive and child-centered pedagogy, interactive teaching and learning processes in the classroom, a curriculum that encourages critical thinking, and respect for and understanding of human rights”.

Similarly, the inclusion of the concept of “lifelong learning” opportunities in the SDG4 sparks hopes for the return of the adult education agenda in UNESCO’s programming and global education policy. However, these are all speculations as the interpretations of lifelong learning could vary. In this context [Elfert \(2019\)](#) argues:

... two reports, both coming out of UNESCO, reflect divergent interpretations: The 3rd Global Report on Adult Learning and Education (GRALE) emphasises that the SDGs place a stronger focus on adult learning and education than did the MDGs ([UIL, 2016:15](#)), while UNESCO’s Global Education Monitoring Report 2016 observes that the 10 targets [of SDG 4] fail to deal explicitly with adult education. [UNESCO \(2016:222; see also p. 236\)](#).

The meaning of lifelong learning also needs to be seen in the post-2015 development context that is marked by a shift in investment attention of major donors toward secondary and higher levels of education. This shift is guided by the financial slump of 2008, that encouraged donors, such as the WB and the OECD, to focus lending programmes on higher levels of education that guarantee a better return on investment ([Mundy and Verger, 2015; Elfert, 2019](#)). In this context, the SDG4 framework of lifelong learning opportunities is an attempt to signal a focus on education beyond primary education especially tertiary education but not necessarily adult education ([Elfert, 2019](#)). This is evident from the recent exaggerated focus of UNESCO on higher, vocational and work skills-based education.²²

This phase is also marked by UNESCO establishing itself in the field of education amid the increased dominance and specialized programming of competing organizations, such as the WB, OECD, UNICEF, and INGOs (that are further discussed in [Negotiating the politics of multilateralism and global education governance](#) section). In this attempt, UNESCO has often collaborated with competing organizations in this congested terrain with a growing consensus on education for development or, what [Mundy \(2006\)](#) refers to as the “new development compact”. These collaborations have created a contradiction within UNESCO’s own programmes.

It is yet to be seen how Covid-19 pandemic will impact the SDG4 programming of UNESCO. UNESCO’s publication focus has understandably shifted to the Covid-19 pandemic education in the developing world since 2020. The focus of this has been on school closures, learning losses, digital divides, inequities in education, curricular and pedagogic challenges, socio-emotional learning, and psycho-social support. With Covid-19 triggering major humanitarian, social and economic crises in the developing as well as developed countries, the democratic goals of education also become as important as they were in the post-World War II context. UNESCO’s report, *Reimagining our futures together: A new social contract for education* (2021a), looks beyond the Education 2030 Agenda as discussed in detail below but a concerted effort to engage with the more complex education questions and contested issues is yet to be seen in UNESCO’s publications.

The assertion of UNESCO’s global education discourse: from the Delors Report to the Futures Report

As a producer of regimes of truth, UNESCO has developed global education through various reports. Four such significant reports have been produced by UNESCO since 1946. These are: (a) *Learning to be* (the Faure Report) ([UNESCO, 1972](#)) (b) *Learning: The treasure within* (the Delors Report) ([UNESCO, 1996](#)); (c) *Rethinking education: Toward a global common good* ([UNESCO, 2015](#)); and (d) *Reimagining our futures together: A new social contract for education* (henceforth the Futures Report) ([UNESCO, 2021a](#)).

The Delors Report (1996) was published at a critical juncture for UNESCO when the parallel educational discourses and compulsions were emerging in its work. The Futures Report, the most recent of its reports, was published at a time when aligning with the neoliberal frameworks and colluding with a narrowing economic view of education have impacted UNESCO’s educational agenda. In order to map this shift in the organization’s work, the discussions have been organized under four subsections.

Setting the problem

The Delors Report responds to a world in which education is “one of the principal means available to foster a deeper and more harmonious form of human development and thereby to reduce poverty, exclusion, ignorance, oppression and war” ([UNESCO, 1996:13](#)). Mindful of the global/local dialectic, it suggests that “people need to gradually become world citizens without losing their roots and while continuing to play an active part in the life of their nation and their local community” ([UNESCO, 1996:17](#)).

²² A large percentage of UNESCO’s publications cataloged in its Digital Library on higher education, skill development and technical and vocational education, were published 2015 in contrast to more recent publications on adult education and literacy. Thus, it can be argued that, despite its rights-based and conceptually broader policy choices, UNESCO’s work in relation to the SDG4 appears to delegitimize this as a priority. However, as SDGs are still being implemented, it is somewhat premature to draw a firm conclusion.

This report identifies globalization and its tensions, such as that between tradition and modernity, the equality of opportunity, and technology, as key socio-cultural challenges while noting the need for an individual humanistic education vision. In this report, the mission of education is “to enable each of us, without exception, to develop all our talents to the full and to realize our creative potential, including responsibility for our own lives and achievement of our personal aims” (UNESCO, 1996:19).

Coming on the back of the liberalization of economies in the 1990s after the fall of the Soviet bloc and changes in international world order, the report notes that:

[t]here exists today a global arena in which, whether we like it or not, the destiny of every individual is to some extent played out. World-wide economic, scientific, cultural and political interdependence – dictated by the opening up, under the pressure of free-trade theories of economic and financial frontiers, reinforced by the break-up of the Soviet bloc and finding an instrument in the new information technologies – is becoming ever more securely established ... What is more, the widespread realization that international relations are now globalized is itself a dimension of the phenomenon. Despite its latent promise, the emergence of this new world, difficult to apprehend and even more difficult to predict, is creating a climate of uncertainty, not to say apprehension, which renders the search for a truly global approach to problems even more half-hearted.

UNESCO (1996:39).

In this “new World”, the report advocates for an education which addresses inequity and plays a special role in the building of a mutually supportive world engendering a new humanism characterized by ethical principles. It also understands and respects different civilizations as a much-needed counterweight to a globalization that would otherwise be only in economic or technological terms (UNESCO, 1996:50).

The report also supports the strengthening of the nation state to ensure that it is not eroded and where democracy is a fundamental principle that nurtures “freedom, peace, genuine pluralism and social justice ...” (UNESCO, 1996:56).

The report does not view education as solely focused on economic growth but as “a broader perspective of human development” (UNESCO, 1996:69). It calls for establishing

new links between education policy and development policy, with a view to strengthening the bases of knowledge and skills in the countries concerned, encouragement of initiative, teamwork, realistic synergies taking local resources into account, self-employment and the spirit of enterprise.

UNESCO (1996:81).

Almost 25 years later, the Futures Report saw a global threat to the planet and humanity. Using similar language of interconnectedness and interdependencies, it advocates for a social contract arguing that the world is at a “turning point” and that development needs to recognize how the “quest for growth and development” has “overwhelmed our natural environment, threatening our own existence ...” (UNESCO, 2021a:1). In its analysis of the global context, it places a hopeful (and perhaps naive) belief in reason and knowledge “in creating better futures” (UNESCO, 2021a:1). The underlying idea and principle of a social contract invokes Rousseau’s ideals of “an implicit agreement among members of a society to cooperate for shared benefit” and as a starting point for “shared vision of the public purposes of education” (UNESCO, 2021a:2). Like the Delors Report, this report identifies multiple and interlocking crises and challenges including inequality, climate change, biodiversity loss, resource use that exceeds planetary boundaries, “backsliding in democratic governance” and disruptive technological automation (UNESCO, 2021a:3). Therefore, both reports draw on similar discourses even though the Futures Report does so in more subtle and non-controversial ways.

Knowledge and learning

The Delors Report identified four pillars of education with the aim of equipping each individual “to seize learning opportunities throughout life, both to broaden her or his knowledge, skills and attitudes, and to adapt to a changing, complex and interdependent world” (UNESCO, 1996:85). The Learning to know, Learning to be, Learning to do and Learning to live together framework, the commission believes, provides the basis for “meeting the challenges of the coming century” (UNESCO, 1996:85). Such an expansive view of education is designed

to enable each individual to discover, unearth and enrich his or her creative potential, to reveal the treasure within each of us. This means going beyond the instrumental view of education, as a process one submits to in order to achieve specific aims (in terms of skills, capacities or economic potential), to one that emphasizes the development of the complete person, in short, *learning to be*.

UNESCO (1996:86).

Unlike the Delors Report, the Futures Report pays more attention to education process in relation to curriculum, pedagogy and assessment as the new social contract proposed “needs to allow us to think differently about learning and the relationships between students, teachers, knowledge, and the world” (UNESCO, 2021a:3). Underlying the educational philosophy of this report is the following framework of principles:

- i. Organizing pedagogy around principles of cooperation, collaboration, and solidarity, to enable fostering of “the intellectual, social, and moral capacities of students to work together and transform the world with empathy and compassion” (UNESCO, 2021a:4). This pedagogy would also have to enable the unlearning of biases, prejudice and divisiveness and assessments would have to reflect these dimensions of pedagogy;
- ii. Curricula which emphasizes ecological, intercultural and interdisciplinary learning;
- iii. Professionalization of teaching, where collaboration and teamwork characterize the work of teachers, and wherein teachers are seen as “knowledge producers” and support is provided for their “autonomy and freedom” and their participation is “fully in public debate and dialog on the futures of education” (UNESCO, 2021a:4);
- iv. Schools as sites to foster inclusion, equity and individual and collective well-being and digital technologies. It highlights the need to support and not replace schools and that the schools need to also become “exemplars of sustainability and carbon neutrality” (UNESCO, 2021a:4);
- v. Learning, as not only taking place throughout life, but also in different “cultural and social spaces”, thus highlighting the need for education to be lifelong and encompass the right to “information, culture, science and connectivity” (UNESCO, 2021a:4).

This report takes a broad intersectional and decolonial approach to education arguing that:

[p]edagogies of solidarity must also recognize and redress the systematic exclusions and erasures imposed by racism, sexism, colonialism, and authoritarian regimes around the world ... Pedagogies of respect, inclusion, belonging, peacebuilding, and conflict transformation go beyond merely acknowledging or tolerating difference ... The world is rich in multicultural and multi-ethnic societies and education should promote intercultural citizenship ... The basis for intercultural citizenship is the affirmation of one's cultural identities. Knowing who you are is the starting point for respecting others. Intercultural education should not be used as a tool for the assimilation of cultural minorities, indigenous peoples, or other marginalized groups to the dominant society, but rather to promote more balanced and democratic power relations within our societies ...

UNESCO (2021a:52–54).

Based on the above vision for learning/education, the report sees knowledge as a pedagogical relationship formed by teachers and students in which teaching and learning “are both nourished by, and contribute to, the knowledge commons” and in which “... teachers and students need to form a community of knowledge-seekers and builders nourished by and contributing to humanity's knowledge commons” (UNESCO, 2021a:51).

While the Futures Report acknowledges the need to promote democratic power relations, it fails to articulate the basis on which socio-cultural difference is constituted. In an attempt to project a positive and hopeful view of the social contract, the report elides and simplifies difference and does not see how it is conjuncturally and politically shaped as a driver of conflict. It fails to provide a sound basis for dealing with the dialectic between sameness and difference. Similarly, while the report sees educational processes and knowledge generation as essentially participatory, without a corresponding understanding of difference that undercuts such action, it does not sufficiently support envisaging the systems where such vision translates to practice.

Realizing the education vision

The Delors Report recognizes that, to realize its vision for education, political choices/decisions need to be made by the “public authorities” in both developed and developing countries:

... the pressure of these demands, which fall largely upon the public authorities, necessitates organizational choices which are often in reality socio-political ones. Policy-makers are faced with conflicting interests: industry demands more and more skills and aptitudes; science wants funds for research and the kind of top-flight higher education that produces young researchers; the arts and academia want the wherewithal for improving school enrolments and general education; and parents' groups militate for more high-quality education ... a better supply of good teachers. The dilemma becomes especially acute when it becomes impossible to satisfy all these demands ... each demand is based on legitimate expectations that all correspond to one or other of education's basic functions.

UNESCO (1996:156–157).

The report emphasizes the role of democratic public debate “not only on the resources to be made available to education but also on its ultimate goals” (UNESCO, 1996:157). It advocates decentralized decision making, involving all stakeholders, encourages autonomy of schools in various decision-making aspects and in the presence of a strong central administration with a strong role for the state as representative of the whole community, in a pluralistic and partnership-based society where education is a lifetime affair.

Drawing from this, the report highlights that investing in education is more than a social investment; it is also an economic and political investment—in building the nation, citizenry and preparing the citizens for the world of work. In light of this, it lays emphasis on the role of public funds, as against a private funding/market, as one of education's central concern is not to “heighten social inequalities” so “large scale resources must be channeled toward the most disadvantaged ...” (UNESCO, 1996:167).

The Futures Report views a “massive commitment to social dialog, to thinking and acting together” as crucial to forge the new social contract for education which would “overcome discrimination, marginalization, and exclusion” (UNESCO, 2021a:5). Specifically, the report, while emphasizing the need for viewing education as a global public good and placing the central responsibility

on public authority, also suggests “more just and equitable cooperation among state and non-state actors. Beyond North-South flows of aid to education, the generation of knowledge and evidence through South-South and triangular cooperation must be strengthened” (UNESCO, 2021a:5). The international community is called upon to enable states and non-state actors to realize a common/shared vision of the new social contract. In this, it argues for a key leadership role for universities that are creative, innovative and committed to strengthening education as a common good for the future of education and in building the social contract.

International cooperation and UNESCO's role

The Delors Report argues for the need for global responses toward

... building a better-or less bad world [that] has become as never before everyone's business ... Co-operation in education should thus be viewed in the broader context of what the international community must do to stimulate awareness of all the problems requiring solution and achieve consensus on the issues needing concerted action; this action will involve international and intergovernmental organizations, governments, non-governmental organizations, business and industry, professional and labour organizations and ... those actively participating in the education system and the intellectual community, all working together as partners.

UNESCO (1996:177).

It highlights the need to radically rethink international cooperation through not only aid policies but also partnerships which foster exchange and cooperation within regional groupings, promote and encourage the education of girls and women through strong international cooperation for policy, encourage debt swaps to offset adverse effects of adjustment policies and strengthen national policies through international cooperation, among other aspects (UNESCO, 2021a:189).

Within this strengthened and broadened international cooperation, the report sees UNESCO as especially suited to exercise its moral authority in the international sphere:

Although UNESCO is neither a funding agency nor a simple research institution, it has always had the task of developing human potential, in co-operation with its member states and its many partners and associates on the international scene ... The distinctive character of UNESCO lies in the scope of its responsibilities – not only education, but also culture, research and science, and communication – which make it an intellectual organization in the broad sense, less subordinate than others to a purely economic view of the issues ... As a moral authority and a setter of international standards, it pays as much heed to human development as to purely material progress. In the field of education, all these attributes predispose it to undertaking action on several fronts: helping Member States to build and reform their education systems and obtain the most benefit from the scientific and technological revolution, but also making the right to education a reality for all people throughout the world and promoting everywhere the idea of peace and the principles of justice and tolerance.

UNESCO (2021a:187 emphasis added).

The Futures Report calls for global solidarity and international cooperation and for a “renewed commitment to global collaboration in support of education as a common good, premised on more just and equitable cooperation among state and non-state actors at local, national and international levels” (UNESCO, 2021a:132).

The report describes the current atmosphere for international education cooperation as “increasingly precarious” with the “notion of a world society anchored in common universal values profoundly eroded” and global fora, such as the UN system, “responsible for establishing common goals and organizing global collective action” facing “harsh critique and fiscal constraint” (UNESCO, 2021a:132). It also points out the restrictions imposed on the work of non-state actors and civil society in an increasingly fragmented world which restricts their ability to build strong coalitions and also the difficulties of functioning in a post-pandemic world where finances are dwindling. This report advocates for the inclusion of diverse non-state actors in global governance through partnerships—a movement away from top-down toward multi-centric action—and new forms of regional cooperation, especially South-South and triangular cooperation (UNESCO, 2021a:133).

The report also looks critically at UNESCO's role over the years and highlights the complex context within which it works. However, it still highlights the advantages of UNESCO to forge the new social contract for education as follows:

UNESCO has faced many challenges over the past 25 years. While it has retained the formal responsibility for coordinating global dialogue and standard setting in education, and for ensuring achievement of SDG4 for education, it has struggled to meet these obligations effectively, and has faced severe criticism. It is sobering to note that despite the breadth of its mandate in education, science and culture, UNESCO's entire budget is smaller than that of many European universities. To play an effective role in our vision for sustainable educational futures, UNESCO will need to rethink its approach to educational development. Building on the principle of subsidiarity, it should see itself first as a partner whose job it is to strengthen regional and national institutions and processes. Secondly it is an evidence broker and an advocate for strengthened data and accountability to citizens at all levels of educational systems. While maintaining its unique role in fostering global dialogue for a new social contract for education, UNESCO must focus the bulk of its financial and human resources on regions where the right to education is most threatened – and particularly on Africa, where the vast majority of the world's youth will live and learn by 2050. UNESCO will need a clearer sense of its comparative advantage within the complex ecosystem of global and regional actors involved in educational norm setting, financing and knowledge mobilization. It should work with UN partners to find innovative solutions to ensure the right to education of involuntary migrants and displaced populations whose numbers are expected to multiply over the course of our uncertain century UNESCO has a unique capacity to convene and mobilize people and institutions around the world

to shape our shared educational futures. Herein lies its great strength. And it is precisely this strength that is needed to build a new, internationally agreed social contract for education and, more importantly, a new deal for implementing it.

UNESCO (2021a:138).

From the above, it is evident that the broad discourse, inclusive of the linguistic choices and their organizing principles, underlying the two reports presents a more nuanced view of education that goes beyond considering only the utility of education for economy. Both the reports conform to the humanistic principles of education and its centrality for civilizations and citizenship. Both the reports call for plurality in perspectives in understanding difference. However, the approaches of these arguments allude to a shade of difference in framing the educational agenda. While both the reports understand the role of education from a human rights perspective, the Delors Report, more assertively than the Futures Report, underlines education as a fundamental right thereby calling for greater interest in and resources for it. Similarly, the Delors Report, in critical and direct terms, calls for a global response to the neoliberal paradigm of education and development, whereas the Futures Report hints of this need by referring to inequalities, marginalities and exclusion. However, there is a clear noticeable shift in how UNESCO's own role has been articulated in the two documents – from an “intellectual organization” with a “moral authority” as “a setter of international standards” to a “partner”, “evidence broker” and “advocate”. This shift in the visualization of its own role is also evident in the difference in the authoritative voice of the Delors Report, to the propositional-reflective tone of the Futures Report.

Negotiating the politics of multilateralism and global education governance

The intellectual evolution of UNESCO serves as a case in point for understanding contestations about hegemony among multilateral agencies and IOs in the international global system. As stated above, the shifts in the intellectual history of UNESCO are inherently tied to the significant political junctures in the multilateral nature of UNESCO and its positioning vis-a-vis other IOs as well as internal member state contestation within UNESCO, in particular, the role of the countries representing the Global North (USA, UK). These negotiations have significantly impacted UNESCO's claim to global leadership in education. In this part, we discuss some of the critical points of how UNESCO's multilateralism has evolved and been shaped over its history.

UNESCO was established to formalize multilateral cooperation between nations for building political order/stability globally in a post-World War II context. Its primary focus was establishing peace by promoting human rights and strengthening bonds between nations through educational, scientific and cultural development. From being an IO that emanated from the European experience of the war, over the years, UNESCO has increasingly come to focus on the economically underdeveloped countries or the Global South. This focus on the Global South is a reflection of the fact that, as a member state UN agency, UNESCO needs to respond to its member states, particularly those who, in the 1960s, saw the formal, albeit not *de facto* ending of colonization. This unique form of multilateralism underlies the distinct identity and role of UNESCO vis-à-vis the other IOs as it creates a space for the representation of the voices of the Global South thereby opening possibilities for counterhegemonic politics and a relatively even balance of power in global policy negotiations. In this context, critical international relations theorists have argued that such post-war multilateralism of UNESCO and other institutions has carved “a distinctive political space, through which an ‘imagined’ world order was negotiated, constructed, and (at times) contested” (Mundy, 1998:452).

Through the close to eight-decade history of UNESCO, the unique multilateralism that forms the foundation of the organization has been debated and contested with regards to its policy, programming and its modus operandi. In this section, we recount three forms of this contestation and conflict in the history of UNESCO's work in education impacting its sustainability and the execution of its work program (Elfert, 2019).

The first is the use of funding as a mechanism to question the intellectual, policy and programming standing and legitimacy of the work of UNESCO through the withdrawal of membership. The USA (as well as the UK) has used funding as a key tool in negotiating its relationship with UNESCO. For example, USAID funding was withdrawn in 1965 for the literacy programmes of UNESCO—stating that Fundamental Education was “contrary to American ideals and traditions” (Dorn and Ghodsee, 2012:375). This incident is one of the early markers in how funding is used to steer the policy and programming directions of UNESCO.

UNESCO's literacy agenda was critiqued by the Global North, particularly the USA, as promoting communism in the Cold War. This withdrawal of funding for UNESCO's literacy agenda led to questions about the organization's status as a lead agency for education. This tension continued until the mid-70s, with USA-UNESCO relationships taking an unprecedented turn when the USA froze its funding in 1974 as the organization granted recognition to the Palestine Liberation Organization. Again, in the 1980s, the USA, followed by the UK and Singapore, withdrew membership and funding due to UNESCO's adoption of the MacBride Report, “*One world many voices*” (UNESCO, 1980). The report challenged the commercialization and monopolization of information and communication media by the powerful, mostly Global North nations, especially the USA, while calling for the liberalization and democratization of media by inclusion of countries in the Global South. The US re-joined UNESCO in

2003²³ until 2011 when it again froze its funding (without formally leaving UNESCO) on the matter of Palestine being granted full membership—finally quitting again in 2019.²⁴

This constant withdrawal of membership and funding (and re-joining) reflects the broad geo-political interests of the global governance in the multilateralist approach of UNESCO. In their work, *Hope and folly: The United States and UNESCO* (1989) Preston, Herman and Schiller identify the withdrawal of the US in 1984 as a significant turning point in the tense relationship between UNESCO as representing all countries and being a full membership agency and being an agency that privileged the voice of certain countries in the Global North. In their analysis, Preston et al. (1989:xiv) note that, following these developments, the criticism of UNESCO by the USA media was so obvious and powerful that it triggered the “anti-UNESCO” campaign and posed an unprecedented challenge to the phenomenon of multilateralism and international cooperation, “indeed, the criticism of the report, and of UNESCO in general, deviated so sharply from reality that the notion of a hidden agenda was hard to avoid” an observer noted and added that the crisis at UNESCO is “part of a challenge to US hegemony by most of the rest of the world” (Preston et al., 1989:xiv).

Similarly, the issue of membership of the PLO and Palestine reflect how UNESCO, as member state multilateral agency, reflects the tension and contestation between speaking for all, in particular, the Global South and speaking for the powerful state of the Global North. This tension is reflective of the broader development in the field of international education and development and, in particular, the efforts of the Global South to seek a voice and representation within the global architecture.

Second, the politics of multilateralism in the international aid architecture and UNESCO, in particular, is reflected in the relationship between IOs. In particular, in the evolution of UNESCO, a key issue is about the leadership of the education agenda and the vision that drives the global education agencies (as discussed in the previous section). This relationship has been particularly significant as, unlike other IOs, UNESCO has, since its inception, been dependent on external funding sources and is not mandated to be a funding agency. Generally speaking, the social, economic and cultural rights (including education) that are focal areas of UNESCO’s work, have been much contested within these agencies (Morsink, 2000; Elfert, 2019; Dorn and Ghodsee, 2012). On the one hand, given the post-war aftermath, social, economic and cultural rights (including education) were seen by the different IOs as being strongly linked with building democracy. On the other hand, until 1962, IOs, including the World Bank, “did not lend a single dollar” to educational programming and rebuilding efforts as there was no guarantee of return on investment (Dorn and Ghodsee, 2012:373). As the WB and the OECD began to prioritize education in driving economic growth, the leadership of UNESCO over the global education agenda was called into question. By the 1980s, the WB’s combined disbursements in education were nearly double those of the UN and its agencies (Torres and Schugurensky, 2002:429). The WB also created its own extensive technical expertise in a wide range of specializations in education and was providing expertise to the Global South for education policy programming. This strategic development and deployment of expertise has played a central role in increasing WB’s influence in global governance and global education policy. The enormous funding and extensive technical capacity of the WB irreversibly impacted the work of UNESCO and ultimately did lead to, as noted above, the dual track global education agenda.

As rapprochement between UNESCO and other IOs emerges, particularly around the new SDG 2015 agenda, UNESCO has begun to focus on becoming a monitoring agency to rebuild its credibility. UNESCO thus currently places great store on taking a monitoring role and, as such, its larger and more expansive education agenda is being challenged. This reflects a broader shift in development thinking about result-based management and “evidence-based policy”. Such a policy truth regime was spearheaded by international agencies, such as the WB, DfID, Education Development Trust, and others, that legitimized their role as international knowledge brokers (Verger et al., 2017). This regime, as evident in the MDG and SDG frameworks and monitoring metrics, limited the scope and autonomy of states by de-territorializing local frameworks of action and establishing accountability frameworks of global education policy (Sayed and Moriarty, 2020). This arguably also resulted in a shift in which countries of the Global South are becoming part of the performativity regime of international development within IOs in general, including UNESCO. Mundy (1999) and Edward et al. (2018) argue that the signs of this narrowing role of UNESCO can be traced back to its attempts in the 1990s to rebrand itself from “lead agency in education globally” to “global intellectual forum”.

Third, the refashioning of its role as global education forum, reflecting its idea of convening power, is evident in the hosting of the International Task Force for Teachers in Education (ITFTE) by UNESCO. The ITFTE managed by UNESCO is “a voluntary global alliance including national governments, intergovernmental organizations, non-government organizations, international development agencies, CSOs, private sector organizations, and UN agencies”.²⁵ The ITFTE in UNESCO reveals two key issues about the multilateral role of UNESCO and its relationship to other IOs. In an “increasingly crowded policy space” for the work of teachers and teaching, and with larger organizations, such as the WB and OECD, developing similar “workstreams and activities”, the challenge for such entities hosted by UNESCO is to become a voluntary global education forum and to “have a clear view of the type of value it provides in comparison to the work conducted by other peers and incumbents” (ITFTE, 2021:12). Additionally, it is evident that such initiatives, including entities such as the Global Partnership for Education (GPE), raise questions about their relationship with their host and how the member state of UNESCO relates to such parallel hosting initiatives. Specially, in the case of the ITFTE, the question an evaluation of the ITFTE raises is of the “integration of the Task Force Secretariat in the UNESCO Section for Teacher Development” and its link to UNESCO’s policy and programming work. As such, initiatives like ITFTE, as is the case of the GPE

²³Announced in 2002. <https://2001-2009.state.gov/p/io/rls/fs/2003/24189.htm#:~:text=September%2022%2C%202003%20The%20United%20States%20Rejoins%20UNESCO,will%20participate%20fully%20in%20its%20mission%20to%20advance>.

²⁴<https://www.aljazeera.com/news/2019/1/1/us-and-israel-formally-quit-unesco> (Accessed on 10.11.2021).

²⁵<http://www.teachersforefa.unesco.org/v2/index.php/en/about-us>.

hosted by the WB, raise serious questions about the individual interests of IOs and efforts to work and collaborate in voluntary ways, and in ways in which funding is provided.

Fourth, while UNESCO happens to be the most critiqued and studied of all UN agencies, its organizational aspects and their broader influence on UNESCO's identity and multilateralism are much understudied, as is the case with many other IOs. In this respect, Jones' (1988) work focuses on how the internal organizational nature and dynamics, particularly related to its administrative and technical staffing, have critically shaped UNESCO's functions. As an intergovernmental organization, UNESCO's primary expertise is drawn from international civil servants who identify as experts in cognate areas of work. Who these expert civil servants are, who they represent, their positionality and their ideologies become crucial in shaping the multilateralism of UNESCO and its relationship to IOs. This is also the case with other IOs, such as the OECD and the WB, yet the gaze has not turned to them. This should therefore be a priority of research and scholarly attention.

This section highlighted how intricately linked UNESCO's work and functioning are to the geopolitical context which has repercussions for both its internal administrative/institutional functioning and its ability to negotiate with the larger global policy field to retain its intellectual and technical focus. It highlighted through four instances, how these impact the different dimensions of its functioning—as intellectual moral authority, as a global education agenda setter and as a convening body enabling wider global policy dialog and coherence. These, it is believed, would need to be studied through a critical interrogation of the working of the entire UNESCO machinery, to be able to shed light on the nuances of the role played by UNESCO in the fast-changing and increasingly uncertain global social, economic and political context within which education policy agendas are set, monitored and pursued.

Conclusion

The analysis in this paper has highlighted the shifts in UNESCO's intellectual trajectory unpacking its evolution in two respects—including its overall vision of education and its distinct positioning and identity as an IO. The discussions have tracked how UNESCO's initial conception of education as a human right and allegiance to a broad-based philosophically grounded connotation of education have come to coexist peacefully with the emergent global consensus on education for development compact. In the process of this evolution, UNESCO has both asserted its voice while resisting being a part of the new consensus, and colluded while finding its own space with competing IOs. It has eventually come to find more settled grounds in seeking to make itself visible in a crowded space, filled with financially stronger and arguably not equally democratically orientated IOs. This settlement has displaced it from the position of a distinct IO in the global education policy governance and has repositioned it to perform the relatively marginal role of a collaborator with strong Global South networks. This role includes being a repository of knowledge and expertise for capacity building but also being a convening and monitoring agency.

The analysis has suggested that this intellectual evolution of UNESCO is far from straightforward as it has been shaped by complex, multilateral politics and tensions not only with other IOs but within itself as well. Hegemonic negotiations between developed and developing member states of UNESCO have often resulted in sustainability, credibility and utility questions being raised on UNESCO's educational programming. Anti-UNESCO campaigns, arguably contrived by the hegemonic powers, provide examples of how establishing "intellectual and moral solidarity" is essentially a politico-economic project and not just an ideational or conceptual one. Both the internal and external tensions indicate how a politics of accords and compromises has determined UNESCO's intellectual ascendancy in the increasingly complicated North-South politics whereby the hierarchies of lending-borrowing as well as knowledge generation determine the weight of the voices of nations in the global economy. The analysis thus indicates how establishing a democratic multilateralism as a moral axis of a post-war world, a goal for which UNESCO was constituted, has remained a utopian dream.

These developments have resulted in a shift in UNESCO's approach toward its democratic vision of the post-war world, to increasingly skirting and eliding critical socio-cultural-economic questions of inequities and global tensions that shape the educational realities of our times. In certain instances, it has emerged that UNESCO's self-image has also seen a corresponding reshaping by losing its confident approach as a moral authority.

In this context, our analysis points to the need for future research on IOs in general and UNESCO in particular. Comparatively mapping the changes in UNESCO's discourse with those in other organizations around the critical time periods in global politico-economic history, in our view, will reveal much beyond the global governance politics, and would deepen the understanding about UNESCO. As a part of our ongoing research, we identify the need for a nuanced and fine-grained analysis of the day to day working and program level interventions of UNESCO at global and national levels. It also requires, in the case of UNESCO, a better understanding of its internal configuration including the relationships in work programmes between the headquarters, regional offices, institutes and country offices. This will partly assist in understanding how UNESCO manages itself as an intellectual/moral leader. A beginning point in this research agenda could be a rigorous study of UNESCO's Digital Archives to develop a detailed picture of its discourse. This proposal is not just to map broad trajectories in educational agenda but to study each of its focal areas in depth that would provide a basis for understanding how similar tensions are reflected in different areas of its work and its discourses. A study of UNESCO from the perspective of the international civil servants and technical staff, through collecting their institutional memory and oral history about the role that expertise has played and how it has shaped the evolution of UNESCO as an IO will, in our opinion, shed light on aspects that a global policy perspective may not reveal.

The intellectual trajectories of UNESCO reveal the complexity of an IO operating within a crowded international space, negotiating contested multilateralism as it seeks to be and become a repository of knowledge and expertise for capacity building, a convening and monitoring agency and a global leader in education. In so doing, it navigates the ways in which “political” and “technical” decisions are made. The analysis of UNESCO points to the need for a better understanding of how global education policy is shaped and reshaped across and between actors within the international development field and how positioning for hegemonic power discursively constitutes the formulation, implementation and monitoring of the global education agenda.

References

- Bagnall, R.G., 2000. Lifelong learning and the limitations of economic determinism. *Int. J. Lifelong Educ.* 19 (1), 20–35.
- Benell, P., Furlong, D., 1998. Has Jomtien made any difference? Trends in donor funding for education and basic education since the late 1980s. *World Dev.* 26 (1), 45–59.
- Biesta, G., 2021. Reclaiming a future that has not yet been: the Faure report, UNESCO’s humanism and the need for the emancipation of education. *Int. Rev. Educ.* 67 (5). <https://doi.org/10.1007/s11159-021-09921-x>.
- Blackmur, D., 2007. A critical analysis of the UNESCO/OECD guidelines for quality provision of cross-border higher education. *Qual. High Educ.* 13 (2), 117–130.
- De Capello, H., 1970. The creation of the United Nations educational, scientific and cultural organization. *Int. Organ.* 24 (1), 1–30.
- Dorn, C., Ghodsee, K., 2012. The Cold War politicization of literacy: communism, UNESCO and the World Bank. *Dipl. Hist.* 36 (2), 373–398.
- Edward, B., Okitsu, T., da Costa, R., Kitamura, Y., 2018. Organizational legitimacy in the global education policy field: learning from UNESCO and the Global Monitoring Report. *Comp. Educ. Rev.* 62 (1), 31–63.
- Elfert, M., 2015a. Rethinking Education: Towards a Global Common Good? UNESCO’s New Humanistic Manifesto? <https://www.norrag.org/rethinking-education-towards-a-global-common-good-unescos-new-humanistic-manifesto/>.
- Elfert, M., 2015b. UNESCO, the Faure Report, the Delors Report, and the political utopia of lifelong learning. *Eur. J. Educ.* 50 (1), 88–100. <https://doi.org/10.1111/ejed.12104>.
- Elfert, M., 2019. Lifelong learning in Sustainable Development Goal 4: what does it mean for UNESCO’s rights-based approach to adult learning and education? *Int. Rev. Educ.* 65, 537–556. <https://doi.org/10.1007/s11159-019-09788-z>.
- Halvorsen, T., 2016. Higher education in developing countries: Peril and promise, a decade and a half later: development lost? *Int. J. African Higher Educ.* 3 (1). <https://doi.org/10.6017/ijah.v3i1.9638>.
- Hite, S.J., Bray, M., 2010. UNESCO’s International Institute for Educational Planning. In: Peterson, P.L., Baker, E., McGaw, B. (Eds.), *International Encyclopaedia of Education*. Elsevier, Amsterdam, pp. 688–693.
- Hulme, D., Scott, J., 2010. The political economy of the MDGs: retrospect and prospect for the world’s biggest promise. *New Polit. Econ.* 15 (2), 293–306.
- ITFTE, 2021. External Evaluation of the International Task Force on Teachers for Education 2030 for Its 2018–2021 Strategic Plan, 2021. <https://teachertaskforce.org/knowledge-hub/external-evaluation-international-task-force-teachers-education-2030-its-2018-2021>.
- Jones, P.W., 1988. *International Policies for Third World Education: UNESCO, Literacy and Development*. Routledge, London.
- Lebeau, Y., Sall, E., 2011. Global Institutions, Higher Education and Development. In: King, R., Marginson, S., Naidoo, R. (Eds.), *Handbook on Globalisation and Higher Education*. Edward Elgar, Cheltenham, UK, pp. 129–147.
- Lee, M., Friedrich, T., 2011. Continuously reaffirmed, subtly accommodated, obviously missing and fallaciously critiqued: ideologies in UNESCO’s lifelong learning policy. *Int. J. Lifelong Educ.* 30 (2), 151–169.
- Limage, L.J., 1999. Literacy practices and literacy policies: where has UNESCO been and where might it be going? *Int. J. Educ. Dev.* 19 (1), 75–89.
- Limage, L.J., 2007. Organizational challenges to international cooperation for literacy in UNESCO. *Comp. Educ.* 43 (3), 451–468.
- Marope, P.T.M., Wells, P.J., Hazelkorn, E., 2013. *Rankings and Accountability in Higher Education: Uses and Misuses*. Unesco, Paris.
- Morsink, J., 2000. *The Universal Declaration of Human Rights: Origins, Drafting, and Intent*. Pennsylvania Studies in Human Rights Series. University of Pennsylvania Press, Philadelphia, PA.
- Muller, J., 2000. From Jomtien to Dakar: Meeting Basic Learning Needs of Whom? *Adult Education and Development*, 55/2000. <https://www.dv-international.de/en/adult-education-and-development/editions/aed-552000/dakar-education-for-all-from-jomtien-to-dakar-meeting-basic-learning-needs-ndash-of-whom>.
- Mundy, K., 1998. Educational multilateralism and world (dis)order. *Comp. Educ. Rev.* 42 (4), 448–478.
- Mundy, K., 1999. UNESCO and the limits of the possible. *Int. J. Educ. Dev.* 19 (1), 27–52.
- Mundy, K., 2006. Education for all and the new development compact. *Int. Rev. Educ.* 52 (1/2), 23–48. <http://www.jstor.org/stable/29737065>.
- Mundy, K., Madden, M., 2009. UNESCO and higher education: opportunity or impasse? In: Bassett, R.M., Maldonado-Maldonado, A. (Eds.), *International Organizations and Higher Education Policy: Thinking Globally, Acting Locally?* Routledge, New York, pp. 46–63. <https://doi.org/10.4324/9780203876664>.
- Mundy, K., Verger, A., 2015. The World Bank and the global governance of education in a changing world order. *Int. J. Educ. Dev.* 40, 9–18.
- Niebuhr, R., 1950. The theory and practice of UNESCO. *Int. Organ.* 4 (1), 3–11. <https://doi.org/10.1017/S0020818300028617>.
- OECD, 1997. *Lifelong Learning for All—Report of the Meeting of the Education Committee at Ministerial Level*, 16–17 January 1996. OECD, Paris. <https://www.voced.edu.au/content/ngv%3A25305>.
- Preston, W., Herman, E., Schiller, H., 1989. *Hope and Folly: The United States and UNESCO, 1945–1985*. University of Minnesota Press, Minneapolis, MA.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Rubenson, K., 2006. Constructing the lifelong learning paradigm: competing visions from the OECD and UNESCO. In: Ellers, S. (Ed.), *Milestones in Adult Education*. Danish University Press, Copenhagen, pp. 63–78.
- Sayed, Y., Ahmed, R., Mogloacci, R., 2012. The 2030 global education agenda and the SDGs: process, policy and prospects. In: Verger, A., Novelli, M., Altinyelken, K. (Eds.), *Global Education Policy and International Development: New Agendas, Issues and Policies*. Bloomsbury Academic, London, pp. 185–208.
- Sayed, Y., Moriarty, K., 2020. SDG4 and the “education quality turn”: prospects, possibilities, and problems. In: Wulff, A. (Ed.), *Grading Goal 4: Tensions, Threats and Opportunities in the Sustainable Development Goal on Education Quality*. Brill, New York, pp. 194–213.
- Spaulding, S., 1992. World education report, 1991: UNESCO. *Comp. Educ. Rev.* 36 (4), 547–550.
- Tawil, S., Locatelli, R., 2015. Rethinking Education: Towards a Global Common Good? UNESCO Education Research and Foresight. <https://www.norrag.org/rethinking-education-towards-a-global-common-good/>.
- The World Bank, 1995. *Priorities and Strategies for Education*. The World Bank, Washington, DC.
- Tilak, J.B.G., 2005. Gender and education for all: the leap to equality—EFA global monitoring report 2003/4. *Int. Rev. Educ.* 51 (1), 73–82.
- Torres, C.A., Schugurensky, D., 2002. The political economy of higher education in the era of neoliberal globalization: Latin America in comparative perspective. *High Educ.* 43, 429–455. <https://doi.org/10.1023/A:1015292413037>.
- UN, 2000. Millennium Declaration: Resolution 55/2. <https://undocs.org/A/RES/55/2>.
- UNESCO, 1945. Constitution of the United Nations Educational, Scientific and Cultural Organisation. UNESCO, Paris.
- UNESCO, 1966a. Final Report of the World Congress of Ministers of Education on the Eradication of Illiteracy. UNESCO, Paris.

- UNESCO, 1966b. General Conference: Thirteenth Session, 13 c/PRG/4. UNESCO, Paris.
- UNESCO, 1972. Learning to Be: The World of Education Today and Tomorrow. UNESCO, Paris.
- UNESCO, 1975. Declaration of Persepolis, 1975: International Symposium for Literacy. UNESCO, Paris.
- UNESCO, 1980. One World Many Voices. UNESCO, Paris.
- UNESCO, 1990. Framework for Action to Meet Basic Learning Needs. UNESCO, Paris.
- UNESCO, 1991. World Education Report. UNESCO, Paris.
- UNESCO, 1993. World Education Report. UNESCO, Paris.
- UNESCO, 1995. World Education Report. UNESCO, Paris.
- UNESCO, 1996. Learning: The Treasure Within. Report to UNESCO of the International Commission on Education for the Twenty-First Century. UNESCO, Paris.
- UNESCO, 1998. World Declaration on Higher Education for the Twenty-First Century: Vision and Action and Framework for Priority Action for Change and Development in Higher Education. UNESCO, Paris.
- UNESCO, 1999. World Education Report. UNESCO, Paris.
- UNESCO, 2004. Higher Education in a Globalized Society. UNESCO Education Position Paper. UNESCO, Paris.
- UNESCO, 2005. Towards Knowledge Societies Report. UNESCO, Paris.
- UNESCO, 2015. Rethinking Education in a Changing World. UNESCO, Paris.
- UNESCO, 2016. Education for People and Planet: Creating Sustainable Futures for All: Global Monitoring Education Report 2016. UNESCO, Paris.
- UNESCO, 2019. UNESCO Priority Gender Equality Action Plan: 2014–2021, 2019 Revision. UNESCO, Paris.
- UNESCO, 2021a. Reimagining Our Futures Together: A New Social Contract. UNESCO, Paris.
- UNESCO, June, 2021b. Key Data on UNESCO Staff. <https://en.unesco.org/careers/keydata>.
- UNESCO Institute for Lifelong Learning (UIL), 2016. 3rd Global Report on Adult Learning and Education (GRALE III). UIL, Hamburg.
- UNESCO-OECD, 2005. Guidelines for Quality Provision in Cross-Border Higher Education. UNESCO, Paris.
- UNESCO-WB Task Force on HE, 2000. Higher Education in Developing Countries: Perils and Promises. The International Bank for Reconstruction and Development, Washington, DC.
- Verger, A., Novelli, M., Altinyelken, K., 2012. Global Education Policy and International Development: New Agendas, Issues and Policies. Bloomsbury Academic, London.
- Verger, A., Fontdevila, C., Rogan, R., Gurney, T., 2017. Evidence-based policy and the education privatization debate: analysing the politics of knowledge production and mobilization through bibliographic coupling. *Int. J. Educ. Dev.* <https://ssrn.com/abstract=2898808>.
- Wagner, D., 2013. Literacy and UNESCO: conceptual and historical perspectives. *N. Dir. Adult Cont. Educ.* 2013 (138), 19–27. <https://doi.org/10.1002/ace.20050>.
- World Education Forum (WEF), 1990. World Declaration on Education for All: Meeting Basic Learning Needs. UNESCO, Paris.
- World Education Forum (WEF), 2000. The Dakar Framework for Action—Education for All: Meeting Our Collective Commitments. UNESCO, Paris.

Education for all and MDGs: global education policy translation in India

Mousumi Mukherjee, Amar Mali, and Tenzin Dolma, O.P. Jindal Global University, Sonipat, Haryana, India

© 2023 Elsevier Ltd. All rights reserved.

Introduction	527
Education for all	528
Background	528
The EFA declaration and the framework of action	528
Review of EFA goals	529
Indian education scenario	530
Lok Jumbish, Rajasthan	530
District Primary Education Program (DPEP)	531
Sarv Shiksha Abhiyan (SSA)	531
Achievement of Sarv Shiksha Abhiyan	532
Role of external aid	532
The Millennium Development Goals: Dakar framework of action	533
Local response to MDGs in India: reducing gender disparity	533
Early childhood care and education (ECCE)	534
Balwadis and Anganwadis	535
Midday meal scheme (MDM)	535
Implementation challenges	536
Conclusion	536
References	536
Further reading	537

Glossary

AWW Anganwadi worker
DFID United Kingdom's Department for International Development
DPEP District Primary Education Program
EFA Education for All
EC European Commission
ECCE Early Childhood Care and Education
ECE Early Childhood Education
ESCAP Economic and Social Commission for Asia and the Pacific
EU European Union
ICDS Integrated Child Development Services
KGBV Kasturba Gandhi Balika Vidyalaya
NCERT National Council of Educational Research and Training
MDG Millennium Development Goals
MDM Midday Meal Scheme
NEP National Education Policy
NGO Non-Governmental Organization
NIEPA National Institute of Educational Planning and Research
NPEGEL National Program for Education of Girls at Elementary Education
NP-NSPE National Program of Nutritional Support to Primary Education
OECD Organization for Economic Cooperation and Development
PRI Panchayati Raj Institutions
SSA Sarv Shiksha Abhiyan
TLM Teaching Learning Materials
UN United Nations
UNDP United Nations Development Program

UNESCO United Nations Education, Scientific and Cultural Organization
 UNICEF United Nations Children's Fund
 WB World Bank
 WCEFA World Conference on Education for All

Introduction

Global Education Policy as a field emerged in the post-war era with the establishment of UNESCO and the Universal Declaration of Human Rights in 1948 (Mundy et al., 2016). Though nation-states continued to enjoy sovereignty with regards to the framing of national public policies, for the first-time education was placed on the post-war agenda of multilateralism. The UNESCO and UDHR focused on ensuring shared principles and values across countries. Education was declared as a basic human right. It was also considered that education "shall promote understanding, tolerance and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace" (United Nations n.d.).

Since the 1950s in the liberal democratic countries of the West, policy studies emerged as a field of study to study the policy-making process and its implementation. Governments also sought to draw on the rich resources of knowledge in the social sciences to inform policymaking, rather than an ad hoc and intuitive traditional approach to public policymaking. There was a prevailing positivist view of social sciences and a rationalist approach to make evidence-based policies with the help of social scientists and empirical research conducted by them.

Several approaches to policy studies have also emerged within academia, mostly developed in the context of the Nation-states located in the Global North. All these approaches, and theories are bounded by a methodological nationalism since policymaking has been mostly understood as a process within the domain of the Nation-States. In recent years, this methodological nationalism employed in policy analysis has been critiqued by scholars, such as Ball (2012), Rizvi and Lingard (2010), Silova et al. (2020). These scholars argue that it is necessary to consider the strong advocacy role played by the powerful intergovernmental organizations and transnational actors, such as the European Union (EU), the United Nations Education, Scientific and Cultural Organization (UNESCO), the Organization for Economic Cooperation and Development (OECD), multinational and transnational corporations, international financial institutions (such as the World Bank and International Monetary Fund), as well as a network of private sector coalitions (such as including Hewlett and Gates Foundations) in assembling policies around the world in this age of neoliberal globalization.

Scholars further argue that it is also important to consider the historic experiences of the postcolonial nations of the Global South seriously because of the process of policy transfer from the colonial metropole of the Global North to the colonies in the Global South since the colonial times. Silova et al. (2020) argues, "All modern education systems arose relationally, be it through purposeful learning from systems elsewhere or by forceful implant by colonizing powers" (p. 3). Education policies and administration within these postcolonial nation-states have been historically shaped by congeries of local and global forces (Mukherjee and Suresh, 2018). Drawing on David Easton's (1953) definition of public policy as "authoritative allocation of values" and insights from a range of theoretical and methodological post-positivist approaches, Rizvi and Lingard (2010) attempted "to show how education policies represent a particular configuration of values whose authority is allocated at the intersection of global, national and local processes".

Moreover, policy translation is complex process that involves discursive practices, various national and transnational political actors, and institutional change. Much has been written in recent years about translation as a policy process analytical framework (Lendvai and Stubbs, 2007, Stone, 2012). They argue:

By reconsidering our understanding of the policy transfer process from the point of view of translation we would argue instead that the policy transfer process should be seen as one of continuous transformation, negotiation, and enactment, on the one hand, and as a politically infused process of dislocation and displacement ("unfit to fit"), on the other hand.

Lendvai and Stubbs (2007, p. 180).

Hence, in this article we will discuss the translation of two major global education policy mandates since the 1990s- Education for All and the Millennium Development Goals within the local context of postcolonial India, as a case study. We will discuss the continuous process of translation, transformation, negotiation, and enactment of these policy mandates according to the values embedded in the postcolonial Indian national constitution, existing policies, and priorities of the local Indian societal context. Thereby, we suggest that the EFA and MDGs might have been also interpreted and translated differently in other systems of education based on their local histories and political traditions. In doing so, we also highlight the challenges of implementing the SDGs in the decades ahead of us.

Education for all

Through the initiative of four international organizations, World Conference on Education for All (WCEFA) was organized from 5 to March 9, 1990 at Jomtien, hosted by the Royal Government of Thailand. These four organizations were United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Children's Fund (UNICEF), the United Nations Developmental Program (UNDP) and the World Bank. At this meeting over 1500 people representing 155 governments, 33 intergovernmental bodies and 125 non-governmental organizations, institutes and foundations signed the declaration titled "World Declaration on Education for all: Meeting Basic Learning Needs" and decided to guarantee quality basic education for all. Organized during the International Literacy Year, the conference set in motion renewed worldwide initiative to meet the basic learning needs of all children, youth and adults and adopted two documents: the World Declaration on Education for All and the Framework of Action to Meet Basic Learning Needs ([Inter-Agency Commission, 1990](#)).

Background

There have been several international conferences prior to the Jomtien conference focusing on the universal literacy and schooling and timebound approach of setting the targets. Some of them are UNESCO's Regional Conferences on Free and Compulsory Education in Bombay (1952), Cairo (1954) and Lima (1956); UNESCO Meeting of Representatives of Asian Member States on Primary and Compulsory Education, Karachi ("Karachi Plan, 1960"); UNESCO Conference of African States on the Development of Education in Africa, Addis Ababa ("Addis Ababa Plan, 1961"). What distinguishes WCEFA from the rest is the scale of participation and the commitments to enlarge the global resource pool to finance education.

This was also the result of a yearlong world-wide process of consultations prior to, during and after the conference at Jomtien. The three principal objectives of WCEFA were:

1. To highlight the importance and impact of basic education, and renew commitment to make it available to all
2. To forge a global consensus on a framework for action to meet the basic learning needs of children, youth and adults;
3. To provide a forum for sharing experiences and research results to invigorate ongoing and planned programs ([Inter-Agency Commission, 1990](#)).

WCEFA was happening following the decade of global economic recession of the 1980s. Most of the least developed countries and many developing countries were facing the challenges of enormous external debts, social and political instability, rapid population growth and environmental degradation. This concern was highlighted in the World Conference debate ([Inter-Agency Commission, 1990](#)). As a result of economic factors and partly due to the negative impact of structural adjustment programs, the state financing on the social infrastructure shrunk. Although total world expenditure on formal education increased to around 5.5% of world GNP in 1980s, material conditions of education in least developed countries deteriorated. There are two reasons for this decline: first, only about 12.5% of the total expenditure was incurred in developing countries. However, these countries have had three-quarter of the total world enrollment in formal education. Second, in many countries of the sub-Saharan Africa, Latin America and the Caribbean, there was a decline in growth of GNP and negative growth in some countries resulting into cutbacks in public expenditure on Education ([UNESCO, 1991](#)).

The EFA declaration and the framework of action

The EFA declaration begins by presenting a broad educational scenario highlighting the seriousness and need for adopting this resolution. It stated that, that there were more than 100 million children with no access to primary schooling in 1990, out of them more than 60% were girls, adult illiterate people were above 960 million out of them at-least two-third were women. More than one-third of the world's adults had no access to printed knowledge, new skills and technologies, and many more adults could not complete basic education programs. It is important to note here that the figures reported for literate and illiterate population are based on UNESCO's 1958 definition, based on the ability or inability to read and write short and simple sentences on everyday life. These data was gathered by the techniques involving self-reporting or based on number of years of schooling. However, if adhered to UNESCO's 1978 definition of "functional literacy" which focus on the ability to use literacy skills in self & community's development then estimated figures of illiterate population will substantially increase ([UNESCO, 1991](#)).

Therefore, at the outset the EFA declaration calls for synergetic efforts by combining exponentially increased wealth of information with the power of modern communication to make the goal of education for all achievable. This declaration recalls that education is a fundamental right of all people with an understanding that education will lead to make the world more safe, healthy and prosperous. It accepted that knowledge is an indispensable key for personal and social improvement with a recognition that traditional knowledge and indigenous heritage must be accepted as righteous.

The focus of the WCEFA was on revisiting and broadening the concept of basic education. Going beyond the rudimentary measures such as basic numeracy & literacy skills or number of years of schooling, the declaration emphasized on the expanded vision of basic education through ten articles:

1. Meeting basic learning needs to benefit every person of all ages across the world to be able to benefit from: basic education designed to cater their differential needs in-terms of learning content, pedagogy and empower them at the same to enable them to enrich cultural and moral values.
2. An expanded vision of basic education that encompasses person from all age groups, from different knowledge systems, education which is not restricted to the ambit of formal schooling institutions, and which promotes the values of life-long learning.
3. Universalizing access and promoting equity by rendering special attention to girls and women, disabled and the disadvantaged.
4. Focusing on learning rather than numerical expansion in enrollments to achieve and maintain an acceptable level of learning, to improve systems of assessing and allowing learners to achieve their fullest potential.
5. Broadening the means and scope of basic education by recognizing the significance of early childhood care and education, universalizing quality primary education for all person within and without formal schooling institutions, education that is not restricted to certain age groups or a single type of knowledge.
6. Enhancing the environment for learning by improving nutritional status, health care, physical and emotional support for the learners.
7. Strengthening partnerships between education department with other departments, at different levels of governing structures within the country and involving non-governmental organizations at local, national, and international levels in the process of financing, planning, implementing, and evaluating basic education programs.

To achieve the above vision, the WCEFA conference decided a set of strategies, which are listed in following three articles.

1. Developing a supportive policy context across the social, cultural, and economic sectors through strong political will, fiscal measures, institutional strengthening, appropriate policies in sectors such as trade, labor, health, and employment. Improving the quality of higher education and scientific technological research.
2. Mobilizing financial and human resources through enlarged public and private sector engagement.
3. Strengthening international solidarity by establishing fair economic relations within and between the nations to redress existing economic disparities.

Along with these strategies and requirements, the EFA declaration set the goals and targets for the decade of 1990s as following (EFA 1990):

1. Expansion of early childhood care and developmental activities, including family and community interventions, especially for poor, disadvantaged, and disabled children.
2. Universal access to, and completion of, primary education (or whatever higher level of education is considered as "basic") by the year 2000.
3. Improvement in learning achievement such that an agreed percentage of an appropriate age cohort (e.g. 80% of 14-year-olds) attains or surpasses a defined level of necessary learning achievement.
4. Reduction of the adult illiteracy rate (the appropriate age group to be determined in each country) to, say, one-half its 1990 level by the year 2000, with sufficient emphasis on female literacy to significantly reduce the disparity between male and female illiteracy rates.
5. Expansion of provisions of basic education and training in other essential skills required by youth and adults, with program effectiveness assessed in terms of behavioral changes and impacts on health, employment, and productivity.
6. Increased acquisition by individuals and families of the knowledge, skills and values required for better living and sound and sustainable development, made available through all education channels including the mass media, other forms of modern and traditional communication, and social action, with effectiveness assessed in terms of behavioral change.
7. With the signing of the world declaration, there was a major thrust globally through revitalizing existing policies and structures, new resolutions and legislations, and enhanced allocation of funds to align literacy and educational programs with that of EFA's framework.

Review of EFA goals

In the second meeting of international consultative forum on Education for All held at Delhi, India in 1993 presented progress made in 2 years. Within 1st years, 13 out of 155 countries expanded their budget on basic education substantially, more than 100 countries defined the goals and plan of action to implement EFA and 56 countries initiated mechanism to promote and monitor the progress of EFA goals. International organizations too improved their funding substantially, for example, the World Bank doubled its loans for basic education, UNDP's allocation for basic education increased to two-thirds of its programs for 1992–1996 and UNICEF committed to allocate 25% of its budget to basic education by the year 2000 (Torres, 1999, p. 7).

Mid-decade assessment presented in FEA Forum in 1996 showed that "school enrollments in developing countries grew by 50 million pupils (from 496 million to 545 million), double the pace of growth in the 1980s. Regions with the highest growth rates were South Asia and sub-Saharan Africa (33 million), followed by the Arab States (5 million), Latin America and the Caribbean (6 million) and East Asia (5 million). The number of out-of-school children fell from 128 million in 1990 to 109 million in 1995" (Torres, 1999, 5).

However, it is apparent from several reviews in the following years that there are several trajectories of the progress made within each country. Some of the main challenges that explain these shortcomings were presented in EFA forum in 1996. It includes lack of determined action; misperception on the part of developed countries that EFA not being their priority; misunderstanding that EFA is indeed “schooling for all”, thus prioritizing creating new schools and increasing enrollments while neglecting qualitative improvements in learning goals—especially those envisioned in “expanded vision of basic education”, paid inadequate attention toward learning needs of adults and out of school youth (Torres, 1999, p. 6). The Amman Affirmation conceded to the fact that despite giving utmost priority to close the gender gap in education, the progress toward this goal is “excruciatingly slow”. It was observed in this assessment that in many underdeveloped countries, less emphasis was given to early childhood education, adult education and out-of-school literacy programs. This assessment also noted the tendency of missing the significance of essential link that connects basic education with secondary and higher education, teacher training and development of technical and vocational skills (UNESCO, 1996a,b).

In a comprehensive review of Education for All (EFA) in a decade following the Jomtien conference in 1990, Torres (1999) argues that this conference redefined the scope of basic education but the original vision, design and targets of the EFA did not materialize and instead the goals have been interpreted selectively and their scope had “shrunk”. The policy programs designed on the principles of EFA in most countries failed to implement the ideas of rethinking and transforming of education. This was mainly caused by increased pressure from the international organizations to meet the goals of EFA in an insufficient timeframe, thus resulting into minimalistic approach and short-term solutions. Thus, globally there was a quantitative expansion in terms of provision of schooling infrastructure, increased enrollment numbers and training programs but the qualitative achievements were far from satisfactory.

Original EFA declaration was for all countries, but developed countries considered their role as funders and did not find it necessary to meet the targets of EFA owing to their developed status and considered the program applied for developing countries. Onus of implementation was shifted on individual countries whereas in the declaration, it was proposed that EFA and basic education was a shared responsibility of governments, civil society and international community. The overall trajectory indicates that in the process of the implementation of EFA strategies, most developing and underdeveloped countries followed the quantitative measures such as increasing number of schools, number of years of schooling and enrollments at primary levels and targeting the children in most vulnerable section of the population. The priorities were given to improve physical environment in schools rather than creating learning environment for all (Torres, 1999, p. 18).

Indian education scenario

In the colonial India, formal schooling institutions were established and expanded but its access was largely limited to wealthy class and upper caste communities, the literacy rate at the time of independence was as low as 16% and about two-third of Indian children had no access to schooling. Postcolonial Indian Constitution adopted the goal of free and compulsory education for all children till the age of fourteen as its directive principle of state policy (Article 45). Initial target of 10 years from the commencement of the constitution in 1950 to achieve this goal was soon realized to be ambitious and unattainable.

In the following decades, a priority was accorded primarily to quantitative expansion of schools to make them accessible with a target of improving enrollments and foundational literacy.

The goal of universal education has been reiterated through all important policy documents, schemes, and programs since independence, such as National Education Policy (NEP, 1968), NEP (1986) and updated policy of 1992 also known as Program of Action (PoA), and recently NEP (2020). Apart from these important policy documents, there have been several interventions such as the scheme of Non-Formal Education (1979), Operation Blackboard (1987), District Institutes of Education and Training (1987), Mahila Samakhya program (1988). To target its vulnerable population groups, state specific Basic Education Projects were launched like the Minimum Levels of Learning (1978), Andhra Pradesh Primary Education Project (APPEP, 1984), Bihar Education Project, Uttar Pradesh Basic Education Project, Lok Jumbish in Rajasthan; District Primary Education Program (DPEP, 1994). India has also implemented schemes to cover entire population such as National Program of Nutritional Support to Primary Education (1995); Sarv Shiksha Abhiyan (2000) and Education Grantee Scheme (2004–05) (Shirname, 2007, p. 11).

Lok Jumbish, Rajasthan

Launched in 1992 with the external aid from SIDA, Lok Jumbish was an important intervention toward achieving the goals of basic education in the state of Rajasthan. The main theme of this program was community mobilization and transforming mainstream education system to ensure that every child has access to basic education. The delivery of this program was channeled through community demand for more schools, a greater number of teachers and for improved school environment.

This program focused on empowerment of women and adolescent girl children belonging to vulnerable communities by expanding the access to formal and non-formal education. This project covered 58 blocks and 9755 villages of the state. Under this initiative, more than 2800 groups of women were mobilized, more than 800 schools were established.

Despite the immensely positive outcomes achieved under this initiative, this program was closed after 7 years of implementation. The main challenge this program faced was the resistance from dominant communities and political leaders from Rajasthan. It is important to take into consideration how the socio-cultural aspects such as caste hierarchy, patriarchy play important role in the success and sustainability of any educational intervention (Ramachandran, 2001).

District Primary Education Program (DPEP)

Launched in the year 1994, this program was the first one wherein in Government of India took support from the World Bank in the area of primary education. Apart from the loan provided by the World Bank, there were other funding agencies such as the Official Developmental Assistance from UK, European Union, the Government of Netherland and UNICEF.

The guidelines issued by the Government of India in 1995 highlights that the motivation for this initiative lies in the efforts made since independence toward the goal of achieving universal elementary education and limitations it faced (MHRD, 1995). This document recognizes the limitations of quantitative expansion of schools. It recognizes that it is not the sufficient condition to achieve the goals of basic education. The objectives in the DPEP program focuses on “a holistic planning and management approach”, inclusive approach to cover schools in remote areas and catering to vulnerable communities, improving the teaching-learning competencies and materials, reducing gender gap in educational outcomes and convergence with programs on early childhood education and health. DPEP strategy was also derived from the experiences of other externally funded interventions such as Bihar Education Project, Uttar Pradesh basic education program and Andhra Pradesh Primary Education Project (MHRD, 1995). The strategy was also based on the empirical data based on six studies conducted by the team of scholars from NCERT and NIEPA on education finance, learning achievements, teacher policy and training needs, gender issues, educational problem of tribal children and textbooks.

This program was launched on a mission mode in 42 districts of 7 states which was expanded to 150 districts, and in 15 districts in a later phase. The main objectives of this program were as following (MHRD, 1995, p. 6):

1. To reduce difference in enrollment, dropout and learning achievement among gender and social groups to less than 5%.
2. To reduce overall primary dropout rates for all students to less than 10%.
3. To provide access for all children to formal or non-formal education.
4. To strengthen the capacity of national, state and district institutions.

In a mid-term assessment conducted in 1997, there was a modest increase in enrollment varying from 2.5 to 6.2%. However, three states witnessed decline in enrollments namely, Assam, Bihar and Maharashtra. Most important gain was made in the growth in enrollments of girls, especially in the districts where there was low female literacy rate (Aggarwal, 2000).

A review of the DPEP conducted for The World Bank by Sipahimalani-Eao and Clarke (2003) highlights similar trends. There has been a substantial progress made in improving the enrollment rates, and most districts achieved the targets of reducing gender disparities in enrollment. Similarly, there have been reductions in gender disparities in education. Levels of learnings have been also improved, especially in grades I and II. Other dimensions such as community mobilization, in-service teacher trainings, revision of textbooks to do away with gender stereotyping also indicate positive gains.

The limitations and challenges experienced in the implementation of DPEP are also important to understand the overall impact. Despite targeting educationally poor performing districts, the results shows that the enrollment rates of children from scheduled castes and tribal communities are still less. The goal to reduce the dropout rates to less than 10%, and targets of improving learning levels seemed to be too ambitious to achieve (Sipahimalani-Rao and Clarke, 2003, p. 10).

Sarv Shiksha Abhiyan (SSA)

Following the global UN declaration of the Education for All in 1990, Sarv Shiksha Abhiyan (SSA), literally meaning the Education for All Mission was launched in India in 2001–2002 as a flagship program to attain universal elementary education for all children in the age group of 6–14 by 2010. This program was designed to be an umbrella program to cover existing programs such as District Primary Education Program, Lok Jumbish and Operation Blackboard.

Despite multiple programs that have been launched since independence, there was still a great concern over the quality and coverage to cater to the educational needs of the diverse Indian population. Following the “bottom-up” process of planning, SSA was designed for active involvement from the community and in partnership with Local Self Governments and State Governments.

The Objectives of SSA were as follows:

- All children in school, Education Grantee Center, Alternative School, or “Back to School” camp by 2003.
- Bridge social, regional, gender gaps at primary stage by 2007 and at elementary education level by 2010.
- All children to complete 5 years of primary schooling by 2007
- All children to complete 8 years of elementary schooling by 2010.
- Emphasis on education for life.

- Universal retention by 2010.

The SSA focused on access, equity and equality through the development of schooling infrastructure, human resource development and qualitative improvements in teaching and learning outcomes. It paid special attention to bridge the gender gaps in education and children from minority communities, disadvantaged groups, who drop out before completing primary schooling. The SSA strategy involved bringing holistic changes and bringing institutional reforms through community-based ownership, management and capacity building of all stakeholders.

As a major step toward making elementary education universal and as a fundamental right, in 2002, an article 21-A was inserted in the Indian Constitution through Eighty-Sixth amendment Act. It reads:

The state shall provide free and compulsory education to all children of the age six to fourteen years in such manner as the State, may by law, determine.

The same act also substituted article 45 of the Constitution to provide early childhood care and education for all children until they complete the age of 6 years. Subsequently in the year 2009, “The Right of the Children to Free and Compulsory Education Act” was passed in the parliament and came into effect from 2010 (MHRD, 2011).

Achievement of Sarv Shiksha Abhiyan

The 2010 evaluation report on SSA by the Planning Commission of the Government of India highlights the achievements in key areas under this program as following (Planning Commission, 2010, p. xiv):

- Improved access—more than 98% of the rural habitations had access to elementary schools within 3 km whereas in urban areas 93% of the children in slums had access to schools within 1 km.
- Enrollments—Overall gross enrollment ratio was increased from 89% in 2003 to 93% in 2007; Enrollment ratio of girls significantly rose leading to gender parity ratio of 0.89 in rural schools and 0.82 in urban schools; Share of differently-abled children in total enrollment increased from 0.43% in 2003 to 1.17% in 2007; share of socially disadvantaged groups was 32% in rural and 30% in urban areas.
- Attendance—Average attendance improved, 62% of the rural schools reported average attendance of more than 75%.
- Out of school children/dropouts—Approximately 7% of the rural households and 20% of the households from urban slums had out of school children or dropouts.
- Infrastructure—88% schools had pucca (all weather buildings); 90% schools had drinking water facilities; 82% of the schools had common toilets but only 50% of the schools had separate toilets.
- Electricity and computer—in rural areas, 60% of the schools did not have electricity, whereas 82% schools in urban slums had electricity; only 11% of the schools in rural areas had computer aided learning, whereas in urban areas the coverage was 62%.
- Teaching Learning Materials—In rural areas, 75% of the schools were equipped with adequate TLMs, whereas in urban areas 93% schools had the TLMs.
- Free Textbooks—nearly 98% of the students in urban areas received the books whereas 84% of the students in rural areas received books at the beginning of the academic year.
- Pupil-Teacher Ratio—60% of the schools in rural areas and 57% of the schools in urban areas had adequate PTR.

Considering the vast population of the country, the achievements made under SSA are tremendous. Despite this achievement, the goal of universal elementary education has still not been achieved due to the diversity of its population and geographic conditions, as well as socio-economic disparities.

Some of the major limitations that SSA has faced are high drop-out rates and quality issues; problems of vacancies of teachers; untrained staff and absenteeism; inadequate teaching learning material; need for including local community in monitoring school performance; weak linkages between monitoring and supervision and last but not the least, the shortages in number of classrooms, separate toilets for boys and girls, and drinking water facilities.

Role of external aid

Initially the SSA was launched without any external support. Since 2004 Government of India has been taking contribution from three development partners namely, United Kingdom’s Department for International Development (DFID), the European Commission (EC) and the World Bank to support SSA. From 2003 to 2007, approximately 30% of the total expenditure on SSA has been disbursed by the three external aid partners. This aid money has been integrated into the central government’s strategy to implement the SSA. It has also helped in supporting the strengthening of technical abilities, assessment models and helped intensifying monitoring and supervision of the program (Ward, 2011).

In their paper, Colclough and De (2010) argue that international aid had not altered the policy priorities of educational expansion but, at the same time had a great impact on management practice, financial reporting, accounting procedures and monitoring

arrangements and provided resources for qualitative improvements. The programs such as DPEP and SSA had special role in improving the access to primary education nearing the goal of universalization of primary education.

The global policy mandate of EFA and international aid agencies can reasonably claim the credit in this achievement along with the government of India. However, critical scholars have argued against external aid in the education sector. They have argued with evidence that the structural adjustment policy connected to external aid has rather increased the pre-existing inequities within the education sector in India (Upendranadh, 1993; Tilak, 2008).

The Millennium Development Goals: Dakar framework of action

After a decade of efforts to meet the goals of EFA outlined during the Jomtien conference, participant member states of the international organizations met at Dakar, Senegal in the year 2000 and reaffirmed their commitment for achieving the target of education for all by 2015. Extensive review was conducted by all participant countries of the Jomtien conference and presented their findings in six regional conferences organized during the year of 1999–2000. Despite significant progress made globally throughout the decade, concerns were raised regarding 113 million children, who still did not have access to primary education.

The Dakar framework highlighted that, 880 million adults were illiterate, gender disparity in education continued to be a daunting fact, and principles of basic education were not being met, wherein adequate skills and knowledge necessary for gainful employment and full participation in society, was not being imparted.

Hence, the Dakar framework set six goals:

1. Expanding and improving early childhood education and care with a special focus on the children in vulnerable conditions.
2. By 2015, all children to have access to and complete primary education of good quality and for free of cost. Priority accorded to girls and children in difficult circumstances and ethnic minorities.
3. Learning needs of all young people and adults should be met through like-skills programs and equitable access to appropriate learning.
4. Adult literacy—50% improvement by 2015, especially for women.
5. Eliminating gender disparities in primary and secondary education by 2005 and, to target gender equity in education by 2015.
6. Improving all aspects of the quality of education and ensuring excellence for all, so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy, and essential life skills.

In order to achieve these goals, the framework pledged the following strategies:

1. To enhance investment in basic education by developing national action plans.
2. EFA policies to be linked with poverty elimination and development strategies.
3. Participation of civil society in the formulation, implementation, and monitoring.
4. Developing accountable systems of educational governance and management.
5. To implement strategies for gender equality in education.
6. To enhance the status, morale and professionalism of teachers.
7. To harness information communication technologies in education effectively.

Local response to MDGs in India: reducing gender disparity

One of the most enduring challenges to attain a goal of universal elementary education in India is of gender equality. Since independence all the important policy documents and programmatic interventions have set the target of achieving equity and equality in educational outcomes for boys and girls. Many efforts have been put in through initiatives such as Mahila Samakhya Program, District Primary Education Program, Shiksha Karmi project of Rajasthan and Kasturba Gandhi Balika Vidyalaya (KGBV), National Program for Education of Girls at Elementary Education (NPEGEL)—to name a few. Reducing gender disparity in education has been the national priority, especially post 1990s EFA targets and the Millennium Development Goals (MDGs). The achievements made during last few decades were, although impressive, still far from the ideals of equality. There are socio-cultural aspects which are deeply embedded which perpetuates heterogeneous gendered realities and economical-political aspects, which creates multiple disadvantages for women and girl child to complete their schooling, higher education and actively participate in democratic process and contribute to the economic development.

There are some critical aspects, such as systemic constraints, pedagogic issues and social factors, that poses multiple challenges for women in the educational processes. Culture plays important role in educational access. Traditionally the Indian women were prohibited from any form of formal or religious education. Learning domestic chores and nurturing the children are seen as the education for girls. Traditionally, her position is vulnerable and powerless due to patrilocal residence, patrilineal descent and exclusion from inheritance of the property. In case of economic constraints, education of the boy was preferred over girl child, as the girl will get married and leave the home, whereas boys will take care of parents in old age. Child marriages has also been detrimental for excluding girls from educational process. But these practices are not uniformly applicable, different communities in different states indicates differential proportion in education and participation in the workforce.

Some of the important initiatives that the Indian government has launched to improve educational access for girls are:

- To ensure universal access, community mobilization has been important strategy in all programs/schemes/initiatives to improve enrollment rates.
- Reservations of women in village education committees, School Management committees and Panchayati Raj Institutions (PRIs) have been provided.
- Efforts have been made to improve number of female teachers, and large number of non-formal and special schools (for example, Kasturba Gandhi Balika Vidyalayas KGBVs¹) have been established for out of school girls.
- Different bridge courses, vocational courses have been included in the school curriculum.
- Through different incentives such as free uniforms, textbooks, stationary and through different awards and scholarships girls are encouraged.

It has been observed that if the schools are easily accessible and functioning well then participation of girls improves, as demonstrated in National Program for Education of Girls at Elementary Education (NPEGEL) wherein providing girls with cycles resulted into reducing drop-outs, improved attendance and levels of learning (Ramachandran, 2009). Similarly, making educational content and pedagogic practices gender sensitive, reducing stereotypes in training of teachers have led to improved learning outcomes for girls.

Girls' enrollment started improving at a greater pace in the decade of 1990s and this growth is directly associated with the establishment of new schools under various flagship programs, centrally sponsored schemes and through various NGOs. However, participation of girls from Scheduled Caste, Scheduled Tribe, Other Backward Classes and minority communities such as Muslims is much lower. The transition rates from primary to upper primary and upper primary to high school levels reveals even more worrisome situation. 28.57% girl's dropout before completing primary education, 52.32% dropout before elementary level and 62.69% of the girls drop-out before secondary education. The drop-out rate for the girls from SC and ST communities are approximately 10%–20% more than the national average respectively (Ramachandran, 2009). Through concentrated efforts of linking completion of primary education with successful grade specific achievements in upper primary stages and access to secondary schools, vocational courses providing the opportunity of self-employment, drop-outs could be reduced.

Early childhood care and education (ECCE)

This section will provide an overview of more flagship initiatives of the Indian government, i.e., Mid-Day Meals, Balwadis and Anganwadis, in response to the MDGs mandate on- Early Childhood Care and Education. Early Childhood Care and Education (ECCE) is a significant area of importance, as increasing empirical evidence points to the value that providing pre-school experience to children not only improves their readiness for school but also meets their basic growth and developmental needs.

The field of neuroscience has provided substantial evidence that "experience-based brain development in the early years sets neurological and biological pathways that affect health, learning and behavior throughout life" (Young and Richardson, 2007). Along with physical development, other important capabilities such as cognitive, social, and psychomotor competencies develop in the critical periods in the early years of life. With the stimulating and enriching physical and psycho-social environment, the chances of a child's brain development to its full potential are considerably higher leading to later success in life. Contrarily, if such a nourishing and stimulating environment is lacking in the early phases, there are irreversible impacts on the development of the child. As a result, many children from families with poorer economic conditions, by the time they reach school age, are at significantly disadvantageous position compared to other children in terms of social and cognitive ability (Kaul and Sankar, 2009).

"Expanding and improving comprehensive early childhood care and education, especially for the most vulnerable and disadvantaged" was the first goal stated in the Dakar Framework for Action UNESCO (2000). ECCE can be classified into the sub-stages of:

- (a) Prenatal to two and a half to three years.
- (b) 3–4 years.
- (c) 4 to 5/6 years.

In the first sub-stage of ECCE, which is from prenatal to 3 years, the most significant aspect is to provide adequate provision for health and nutritional wellbeing. This stage of life is the most vulnerable to proper physical growth and brain development. In this stage, it is important to pay adequate attention to parent counseling in nutrition and health education and in "early psycho-social stimulation" by following the home targeting approach.

In the next substage which spans from 3 to 4 years, the important aspect to cater to is early learning in an organized manner. For this purpose, the Early Childhood Education and Care (ECCE) program recommends the learning facilitated through a center-based

¹The KGBV scheme was launched in the year 2004 (and brought under the umbrella of SSA from 2007) for girls belonging to SC, ST, OBC or minority communities and girls from families below poverty line. KGBVs are residential schools till grade VIII, but some states have introduced higher grades also. As on 2015, there are over 3600 KGBVs in 460 districts from 27 states, which have low female literacy rates (NITI Aayog, 2015). The context and location specific nature of this scheme has been instrumental in reaching out to the girls situated where day schools could not.

approach using the play method. In the third substage from 4 to 6 years, school readiness elements are introduced in a structured approach (Kaul and Sankar, 2009).

The earliest formal documentation of preschool/early childhood education, the first government initiative in India was Central Social Welfare Board in 1953, which started as a grant-in-aid scheme for voluntary agencies (Kaul and Sankar, 2009). Gijubhai Badheka and (later joined by) Tarabai Modak the pioneers of ECCE in India; they established the first Montessori school in India in 1920 at Bhavnagar in Gujarat and together prepared material for teachers and started a training college for preschool teachers (Early Childhood Care, n.d.).

In 1975, one of the world's largest and unique program was launched i.e., the Integrated Child Development Services (ICDS) for early childhood care and development. Objectives of this scheme were to reduce malnutrition, morbidity, reduced learning capacity and mortality of children in the age group of 0–6 years, pregnant women, and lactating mothers (Ministry of Women and Child Development, 2009). Another significant landmark was the 86th Constitution Amendment Act (2002) which committed to the provision of “free early childhood education and compulsory education for children up to 14 years of age”.

Some of the important policy initiatives in India for strengthening ECCE are following: National Nutrition Policy (1993) identified children below 6 years old as high risk category; National Policy on Empowerment of Women (2001) made the provision for childcare facilities including creches at work place; National Plan of Action for Children (2005) set the target of universalization of ECCE and National Curriculum Framework (2005) emphasized 2 years of pre-schooling as essential aspect for the holistic development of children (Kaul and Sankar, 2009).

In three decades from its inception, ICDS has been expanded to cover entire nation and the budgetary allocations (both state and external) have been increased significantly. According to the latest data available, in 2015 there were 1,346,186 Anganwadis (early childhood care center) in India providing preschool education to 36,543,996 children (Ministry of Women and Child Development, 2015).

However, there are some persisting challenges in the structure of implementation and the coverage of beneficiaries. One of the critical concerns is regarding the number of tasks that has been assigned to Anganwadi workers. It includes keeping the track of new births, of health indicators of infants and children till age six; to provide supplementary nutrition feeding for children, expectant and nursing mothers; to provide health and nutrition education to mothers, counseling to newly married couples on birth control/family planning; coordination with different institutions such as primary health care centers, schools and to assist in the implementation of different government schemes. The Anganwadi worker is also supposed to conduct pre-school activities for the children in the age group of three to six and educating parents on their role in child's growth and development (Ministry of Women and Child Development, n.d.). So, in practice, the ICDS services have been focusing on the “feeding” aspect rather than behavior change in the child caring practices. There are also limitations in the skill sets and educational levels that AWWs have, to be able to effectively implement ECCE and behavioral aspects at the grassroot level (Kaul and Sankar, 2009).

Balwadis and Anganwadis

The first Balwadi was established in 1945 by Tarabai Modak as a pioneer effort for tribal people's education and non-formal educational activity center for pre-school education at Bordi village in Maharashtra (Shukla, 2004). Since then, Balwadis and Anganwadis are accepted concept and approach to early childhood education. Modak's two types of Balwadis are (Shukla, 2004):

1. Central Balwadi: functions during regular school hours and are centrally located.
2. Angan Balwadi or Anganwadi: functions at the hours convenient for the children in their own localities.

The purpose of Balwadis is to provide basic and fundamental child facilities for physical and mental growth at home and outside (Bhattacharya, 2008). Thousands of Balwadis have been set up all over India by government and non-government agencies, for instance, many Balwadis are set up by voluntary organization like Kasturba Gandhi National Memorial Trust as well as developed as a part of the government of India's poverty alleviation program by universalizing education. Hence, today most of the pre-primary sector is largely covered by the Balwadis or Anganwadis, and the independent nursery schools or nursery schools attached to regular schools are very few (Abraham, 2011).

The limited impact of early childhood care in India is because “the staff running the ECCE program is the single most important factor in determining the quality of the program yet, it is the most neglected aspect of the program” (Abraham, 2011). The teachers at this level are not merely imparting education but they are responsible for holistic care of a child, and yet most of the nursery teachers employed in Balwadis or Anganwadis are trained for a very short duration and the training has been mostly inadequate. But a ray of hope is in the National Education Policy (2020) where the government has proposed to train the current Anganwadi workers into high-quality ECCE teachers (Kulkarni, 2020).

Midday meal scheme (MDM)

Provision of school lunches is an old concept in India. Even before its independence, Madras Municipal Corporation was giving meals to disadvantaged children. In post-independence India, by the 1990s most of the states were providing school lunches or food grains to the children in primary grades through their own exchequer. The National Program of Nutritional Support to Primary

Education (NP-NSPE) was launched in 1995 as a centrally sponsored scheme with twin objectives—to enhance enrollment, retention and attendance; and to enhance nutritional standards of children in primary classes, through the provision of free food grains to the children in primary grades. In subsequent years, this program was expanded to all the block of the country and the coverage was expanded to include the children till grade VII (Ministry of Education, n.d.). This is the largest school feeding program in the world reaching out to over 120 million children in the country.

Implementation challenges

Although this scheme promised cooked meal, some states failed in this arrangement and instead, provided students with monthly dry ration based on their attendance in school until 2001 when the supreme court issued an interim order directing states to introduce mid-day meal in all public schools and to provide a minimum content of 300 calories and 8–12 g of protein each day of school for a minimum of 200 days (Sahai, 2014).

This was an important step toward promoting universalization of primary education in India. This scheme specifically targeted children belonging to economically poorer and socially vulnerable groups. The positive co-relation between provision of school feeding programs and enrollment rates, is evident in many developing countries (Jayaraman and Simroth, 2015). In their large-scale study on the impact of the mid-day meal scheme on primary school enrollment in India, Jayaraman & Simroth found that school enrollments increased by 13% between the years 2002 and 2004. This corresponds to additional 2.4 million children in school. Although, this increase is largely driven by enrollments in grade one, the study indicated that mid-day meal resulted into 7% reduction in out of school children in the age group of 6–10. Mid-day meal scheme has also provided safety net to raise the nutritional intakes among children, especially in the draught-stricken areas (Singh et al., 2014).

Despite its significant achievements, the effectiveness of this scheme has challenges: sometimes in the transportation to the rural areas, the food grains get adulterated and pilfered in the delivery, and some schools do not have the facility or kitchen/storeroom/clean water to cook (Sahai, 2014). Some field studies have also reported that there have been instances of caste-based discrimination in the provision of mid-day meal with regards to food cooked by *Dalit* (Scheduled caste) cooks (Khera, 2006). Another important concern is regarding the quality of the food and maintaining nutritional standards as prescribed in the guidelines (Shukla, 2014).

Conclusion

India has reduced poverty and hunger by half (MD Goal 1), achieved gender parity in primary school enrollment (MD Goal 3), reduced maternal mortality by three-quarters (MD Goal 5), however there is still need of improvement in achieving universal primary school enrollment and completion and, achieving universal youth literacy (MD Goal 2) according to the UN India report by Economic and Social Commission for Asia and the Pacific (UN India, 2015). India's progress in achieving the MDGs has been significant considering it being a developing country with 1.3 billion population. Yet a lot more needs to be done to attain the global goal of education for all as a human right. From the above discussion, it becomes clear that existing local histories-inequalities and discriminatory attitudes can cause major barriers for policy translation and implementation, just as structural adjustment policies connected to external aid.

The Sustainable Development Goal (4) has further strengthened the educational agenda by making governments pledge to “ensure equitable quality education and promote lifelong learning opportunities for all”. According to UNDP India, although India made substantial progress in achieving the EFA and MDGs' universal primary education through national initiatives such as Right of Children to Free and Compulsory Education Act (2009), the quality of education, out of school children and school drop-outs are major challenges, particularly among the children in rural areas and belonging to minority communities (UNDP, 2015). The integration of SDG 4 in the New Education Policy 2020 aims for positive impact on achieving inclusive and equitable quality education in India. SDG 4 claims to do the “unfinished business’ of the EFA agenda and the education-related MDGs” (UNESCO, 2016, p. 7). Strong political will, well-targeted interventions and resources are required to fulfill this “unfinished business”. Scholars have already raised alarm about the domination of “development” thinking during the era of MDGs through growth paradigm that has increased inequalities and environmental exploitation. It has been argued that it is necessary to spread awareness about development paradoxes and encourage alternative methods to educate for promoting sustainable development- SDGs through indigenous learning, ecopedagogy, and empowerment education (Kopnina, 2020). Hence, the success of Sustainable Development Goals (SDGs) and specifically, SDG 4 (pledge to achieve quality education for all) will once again depend on how this global policy mandate is translated and implemented within local contexts.

References

- Abraham, J., 2011. Teacher education in early childhood care and education: issues and concerns. In: The Primary Teacher. January and April 2011 PDF. <https://nroer.gov.in/home/file/readDoc/5983591116b51cb4b3e86e64/teacher-education-in-early-childhood-care-and-education-issues-and-concerns.pdf>.
- Aggarwal, Y., 2000. Monitoring and Evaluation Under DPEP: Measurement of Social Impact. NIEPA, New Delhi.
- Ball, S., 2012. New Policy Networks and the Neo-liberal Imaginary. Routledge.

- Bhattacharya, S., 2008. *Social Work: Interventions and Management*. Deep & Deep Publications, ISBN 978-81-8450-073-8.
- Colclough, C., De, A., 2010. The impact of aid on education policy in India. *Int. J. Educ. Dev.* 30 (5), 497–507.
- Early Childhood Care, (n.d.) Education and Development. 1. H10EC14 -Sri Aurobindo Ghosh, Gijubhai Badheka, Tarabai Modak. PDF access http://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/home_science/10_early_childhood_care_education_and_development/14_aurobindo_gijubhai_badheka_tarabai_modak/et/6716_et_et.pdf.
- Inter-Agency Commission, 1990. *Final Report World Conference on Education for All: Meeting Basic Learning Needs*. Jomtien, Thailand.
- Jayaraman, R., Simroth, D., 2015. The impact of school lunches on primary school enrollment: evidence from India's midday meal scheme. *Scand. J. Econ.* 117 (4), 1176–1203.
- Kaul, V., Sankar, D., 2009. *Early Childhood Care and Education in India. Education for All Mid-Decade Assessment*. National University of Educational Planning and Administration (NUEPA), New Delhi, India.
- Khera, R., 2006. Mid-day meals in primary schools: achievements and challenges. *Econ. Polit. Wkly.* 4742–4750.
- Kopina, H., 2020. Education for the future? Critical evaluation of education for sustainable development goals. *J. Environ. Educ.* <https://doi.org/10.1080/00958964.2019.1710444>.
- Kulkarni, S., 2020. Anganwadi workers key to ECCE implementation under NEP; training soon. In: *The Deccan Herald*. UPDATED: AUG 02 2020, 19:06 IST. <https://www.deccanherald.com/national/anganwadi-workers-key-to-ecce-implementation-under-nep-training-soon-868732.html>.
- Lendvai, N., Stubbs, P., 2007. Policies as translation: situating transnational social policies. In: Hodgson, S., Irving, Z. (Eds.), *Policy Reconsidered: Meanings, Politics and Practices*. Policy Press, Bristol, pp. 173–189.
- MHRD, 1995. *DPEP Guidelines*. MHRD, New Delhi.
- MHRD, 2011. *Sarva Shiksha Abhiyan: Framework for Implementation Based on the Right of Children to Free and Compulsory Education Act, 2009*.
- Ministry of Education, (n.d.). Mid Day Meal Scheme. Government of India. www.mdm.nic.in/mdm_website/.
- Ministry of Women and Child Development, 2009. *Integrated Child Development (ICDS) Scheme*. Government of India. <https://icds-wcd.nic.in/icds.aspx>.
- Ministry of Women and Child Development, 2015. All India Status of ICDS Scheme as on 31.3.2015. Government of India. <https://icds-wcd.nic.in/Qpr0314forwebsite23092014/currentstatus.htm>.
- Ministry of Women and Child Development, (n.d.). Roles and Responsibilities of AWWs. Government of India. <https://icds-wcd.nic.in/icdsimg/RolerresponseAWWs.pdf>.
- Mukherjee, M., Suresh, K., 2018. Local and Global Congeries Shaping Educational Administration in India. *Asia Pacific Education: Leadership. Governance and Administration*, p. 173.
- Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), 2016. *The Handbook of Global Education Policy*. Wiley, Oxford.
- NITI Aayog, 2015. *Evaluation Study on Kasturba Gandhi Balika Vidyalaya (KGBV)*. National Institution for Transforming India Aayog, Government of India, India: New Delhi.
- Planning Commission, 2010. *Evaluation Report on Sarva Shiksha Abhiyan*. Retrieved January, 8, 2014.
- Ramachandran, V., 2001. Community participation in primary education: innovations in Rajasthan. *Econ. Polit. Wkly.* 2244–2250.
- Ramachandran, V., 2009. *Education for All—Mid Decade Assessment. Towards Gender Equality in Education*. National University of Educational Planning and Administration, New Delhi.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge.
- Sahai, C.S., 2014. Mid-day meal scheme: achievements and challenges. *Int. J. Humanit. Soc. Sci. Invent.* 2319–7722. Volume 3 Issue 10, October. 2014, PP.06-09. www.ijhssi.org.
- Shirame, S., 2007. Education for All. India: Historical Development, Especially in the Light of Gender Equality and Impact on the Present-Day Situation, Advancing Social Well-Being and Gender Equity in a Globalizing World. *Dialogue between India and Finland*, Helsinki, pp. 27–28.
- Shukla, R.P., 2004. *Early Childhood Care and Education*. Sarup & Sons, ISBN 978-81-7625-474-8, p. 106.
- Shukla, S., 2014. Mid-day meal: nutrition on paper, poor food on the plate. *Econ. Polit. Wkly.* 51–57.
- Silova, I., Rappleye, J., Auld, E., 2020. Beyond the Western horizon: rethinking education, values, and policy transfer. In: *Handbook of Education Policy Studies*. Springer, Singapore, pp. 3–29.
- Singh, A., Park, A., Dercon, S., 2014. School meals as a safety net: an evaluation of the midday meal scheme in India. *Econ. Dev. Cult. Change* 62 (2), 275–306.
- Sipahimalani-Eao, V., Clarke, P., 2003. A Review of Educational Progress and Reform in the District Primary Education Program (Phases i and ii).
- Stone, D., 2012. Transfer and translation of policy. *Pol. Stud.* 33 (6), 483–499. <https://doi.org/10.1080/01442872.2012.695933>.
- Tilak, J.B.G., 2008. Political Economy of external aid for education in India. *J. Asian Publ. Pol.* 1 (1), 32–51. <https://doi.org/10.1080/17516230701835898>.
- Torres, R.M., 1999. One Decade of Education for All: The Challenge Ahead. IPE.
- UN India, 2015. India and the MDGs: Towards a Sustainable Future for All. PDF. https://www.unescap.org/sites/default/files/India_and_the_MDGs_0.pdf.
- UNDP, 2015. UNDP in India. In: *Millennium Development Goals*. Available at: <https://www.in.undp.org/content/india/en/home/post-2015/mdgoverview.html>.
- UNESCO, 1991. *World Education Report*. Paris.
- UNESCO, 1996a. *Education for All: Achieving the Goal; The Amman Affirmation*.
- UNESCO, 1996b. *EFA Forum—Education for All: From Delhi to Bali, Report From the Ministerial Review Meeting of the Nine High Population Countries (Jakarta, Indonesia, Sep. 1995)*. UNESCO, Paris.
- UNESCO, E., 2000. *The Dakar Framework for Action: Education for All: Meeting Our Collective Commitments*. Dakar Senegal, pp. 26–28.
- UNESCO, 2016. *Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4: Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning*. UNESCO Digital Library.
- Upendranadh, C., 1993. Structural adjustment and education: issues related to equity. *Econ. Polit. Wkly.* 28 (44), 2415–2419.
- Ward, M., 2011. Aid to education: the case of Sarva Shiksha Abhiyan in India and the role of development partners. *J. Educ. Pol.* 26 (4), 543–556.
- Young, M.E., Richardson, L.L. (Eds.), 2007. *Early Child Development From Measurement to Action: A Priority for Growth and Equity*. World Bank Publications.

Further reading

- Ball, S.J., 1998. Big policies/small world: an introduction to international perspectives in education policy. *Comp. Educ.* 34 (2), 119–130.
- Ball, S.J., 2012. *Politics and Policy Making in Education: Explorations in Sociology*. Routledge.
- Birdsall, N., Vaishnav, M., 2005. Education and the MDGs: realizing the millennium compact. *J. Int. Aff.* 257–264.
- Buchert, L., 1995. The concept of Education for All: what has happened after Jomtien? *Int. Rev. Educ.* 41 (6), 537–549.
- Gabriel, J.M., 2017. David Easton's "Authoritative Value Allocation"—Activating the Definition's Potential. Available at SSRN 2909910.
- Kainth, G.S., 2006. A mission approach to Sarva Shiksha Abhiyan. *Econ. Polit. Wkly.* 3288–3291.
- Kumar, K., Priyam, M., Saxena, S., 2001. Looking beyond the smokescreen: DPEP and primary education in India. *Econ. Polit. Wkly.* 560–568.
- Lacey, C., Cooper, B., Torrance, H., 1993. Evaluating the Andhra Pradesh primary education project: problems of design and analysis. *Br. Educ. Res. J.* 19 (5), 535–554.
- Ministry of Human Resource and Development, 2004. *A Program for Universal Elementary Education-Manuel for Planning and Appraisal*. Department of elementary education and literacy.

Prajapati, R., et al., 2017. First Quarter 2017. Significance of Life Skills Education Contemporary Issues in Education Research, vol. 10, 1.

Ramachandran, V., 2008. External Aid in Elementary Education: A Double-Edged Sword (No. id: 1673).

Ravi, J., 2014. Matthal Centre for Educational Innovation. Indian Institute of Management Ahmedabad. Mahila Samakhya 2014-A National Review. Ahmedabad. Retrieved from. http://mhrd.gov.in/sites/upload_files/mhrd/files/upload_document/Report-MSP.pdf.

United Nations, (n.d.). Universal Declaration of Human Rights. Available at <http://www.un.org/en/universal-declaration-human-rights/>.

The education Sustainable Development Goal 4: a critical appraisal

Nigel O.M. Brissett, Clark University, Worcester, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	539
Background to SDGs and SDG4	539
The emergence and structure of SDG 4	540
Optimism and virtues of SDG 4	542
The limits and liabilities of SDG 4	543
Conclusion	545
References	546

Introduction

The fate of our futures has been increasingly tied to the prospects of educational change and advancement. From its entrance into the post-World War II era's incarnation of modernity that envisioned a more unified world, albeit one that is western defined and led, education has gained prominence as a human right. Subsequent education-specific initiatives—including Education for All (EFA), the Dakar Framework, the Millennium Development Goals (MDGs) 2 and 3, and Decade of Education for Sustainable Development 2005–2014—have cemented education as a legitimate stakeholder in development discourses (Wulff, 2020). The new Sustainable Development Goals (SDGs), the global successors to the MDGs, have roped-in education more tightly for a seemingly more integrative role in the United Nations' "plan of action for people, planet and prosperity" and "to strengthen universal peace in larger freedom" (United Nation, 2015, p. 1). In fact, the Incheon Declaration, a declaration on education adopted at the World Education Forum in 2015, proclaims Sustainable Development Goal 4 (SDG 4 henceforth), *Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*, the "unfinished business" of the EFA agenda and the education-related MDGs" (UNESCO, 2016, p. 7). The vision that is promoted through SDG 4 "is to transform lives through education, recognizing the important role of education as a main driver of development and in achieving the other proposed SDGs" (UNESCO, 2016, p. 7).

This article presents the core principles contained in SDG 4 and critically interrogates the values embedded in them, as well as assesses practical issues of relevance and implementation. The paper unfolds in the following manner: the next section briefly examines the emergence of SDG 4 out of MDG 2 and 3, as well as the basic structure and content of SDG 4. The section after this discusses the positive aspirations of SDG 4, noting how education is framed to advance its key Target by itself and in support of the overall SDG framework. The final section addresses some of the deep detractions of SDG 4, in particular its embedded values and how they may hinder some of the potentially liberatory goals of education. The paper ends with a brief conclusion that summarizes some of the key points.

Background to SDGs and SDG4

The Sustainable Development Goals (SDGs) succeeded the Millennium Development Goals (MDGs) which consisted of 8 Goals that were established for the period 2000–2015 geared toward addressing development problems of the Global South. The MDGs contained two education-related goals: Goal number 2 "To achieve universal primary education," and Goal number 3 "To promote gender equality and empower women." A common view is that while the MDGs did help create a global focus on development issues, there were significant flaws in their initial construction, focus, implementation and achievements (Hulme, 2009; Fehling et al., 2013). Among the more common detractions of the MDGs, they were described as "being created by only a few stakeholders without adequate involvement of developing countries and overlooking development objectives previously agreed upon" and being "unachievable and simplistic, not adapted to national needs, [did] not specify accountable parties and reinforce[d] vertical interventions" (Fehling et al., 2013). Yet others point to some important achievements. Ban Ki-Moon, then UN Secretary General, said the MDGs "produced the most successful anti-poverty movement in history," and "helped to lift more than one billion people out of extreme poverty, to make inroads against hunger, to enable more girls to attend school than ever before and to protect our planet" (p. 1109). He added that the MDGs "generated new and innovative partnerships, galvanized public opinion and showed the immense value of setting ambitious goals" (United Nations, 2015b). Specifically related to the education goals, MDG Goals 2 and 3, their achievements are cited to include:

- Enrollment in primary education in developing regions reached 91% in 2015, up from 83% in 2000.
- Among youth aged 15–24, the literacy rate has improved globally from 83% to 91% between 1990 and 2015
- The gap between girls/women and boys/men has narrowed (United Nations, 2015c).

However, the MDGs' educational failures were also seen as consequential; for instance the [United Nations \(2015b\)](#) points out that there continue to be "large disparities in primary school enrollment, and the poorest and most disadvantaged children bear the heaviest burden" (p. 26). Increased enrollment did not necessarily translate to learning (an issue of quality), and while the gender gap had narrowed under the MDG 3 as evidenced by improved gender parity, equal treatment and learning outcomes were still problem areas. Importantly, too, the MDGs were silent on several marginalized groups ([Wulff, 2020](#)). Additionally, the combination of over-crowded, underequipped classrooms and unqualified teachers also reproduced patterns of inequality and exclusion, thus limiting the progress of the MDG Education goals ([Wulff, 2020](#)). The SDG 4 was supposed to address these MDG failures and tackle other challenges. In fact, the *The Millennium Development Goals Report 2015* framed the task of the SDGs in this way:

Drawing on lessons learned from the MDGs, interventions will have to be tailored to the needs of specific groups of children—particularly girls, children belonging to minorities and nomadic communities, children engaged in child labor and children living with disabilities, in conflict situations or in urban slums. Investing in the quality of education and ensuring a sustainable source of funding are also essential.

[United Nations \(2015b, p. 27\)](#).

The SDGs emerged "out of the much wider agenda [than] the MDGs and evolved out of a hugely interactive process of meetings, lobbying and advocacy at every level, national, regional and international" ([King, 2016](#), p. 18). These new Goals were more global than the MDGs which were largely designed with a focus on developing countries. The SDGs promised to eradicate global poverty, fight inequalities and tackle climate change through the multipronged approach of achieving 17 integrated goals and 169 targets covering social, economic, and ecological issues the world faces today. There was a markedly different tone and process of designing the SDGs in important ways compared to the MDGs, which were developed by a small group of Westerners ([Hulme, 2009](#)). Thus, a fundamental difference between the SDGs and the MDGs is the adoption of a more "participatory" approach to developing the Global Goals. One of the initiatives that preceded the SDGs was "The Future We Want," a declaration that specified the need for greater inclusion of civil society and marginalized populations universally in the creation of global development policies. The UN also produced "A Million Voices," a document that reflected input from governments, think tanks, NGOs, civil society, and academics regarding their aspirations for the post-2015 development framework. Additionally, the UN launched the "My World Survey" which allowed individuals across the globe to vote online for the top six issues they wished to see in the SDGs. These apparently more participatory initiatives, drawing from multiples sources, are different from the policy formulation process for the MDGs, which was left solely in the hands of a small group of UN experts ([Hulme, 2009](#)). Yet, even with this seemingly more open system of contributing to and developing the SDGs, some, like [King \(2016\)](#), suggest it is important to ask questions about who had final control over how the texts were edited, and who took decisions on the education targets. [McGrath and Nolan \(2016\)](#) also suggest that the highly technocratic process of indicator setting silences the voices of non-"experts" (p. 122). These issues are explored later in the paper.

Rhetorically, the SDG's participatory emphasis is evident in the language used in the initiative's framework, including references to various documents developed to shape the SDGs, such as "A Million Voices" and the "My World Survey". [United Nations \(2015\)](#) noted that "The Goals and targets were the result of over two years of intensive public consultation and engagement with civil society and other stakeholders around the world, which paid particular attention to the voices of the poorest and most vulnerable" (p. 3). Similarly, the Framework's emphasis on working "in a spirit of global solidarity" (p. 10), "embark[ing] on a collective journey" (p. 1) to ensure that "no one will be left behind" (p. 1 and 3), and adopting a "people-centered approach" (p. 3 and 8) makes clear the SDGs' effort to appear participatory and reflective of the needs of all populations, particularly "the poorest and most vulnerable" (p. 3). The SDG framework utilizes unifying rhetoric, stressing the importance of all countries playing their part to "free the human race from the tyranny of poverty and want to heal and secure our planet" (p. 1), and claims that "Never before have world leaders pledged common action and endeavor across such a broad and universal policy agenda" ([United Nations 2015](#), p. 18). On the surface then, by "involving" marginalized voices in the policy formation stages, by using unifying rhetoric, and by recognizing the role of governments in tailoring the SDGs to meet their own national contexts and realities, the SDGs appear highly participatory. This is the overall SDGs framework on which the SDG 4 is built.

The emergence and structure of SDG 4

[Wulff \(2020\)](#) notes that the inclusion of a stand-alone Education Goal was not a foregone conclusion when negotiations for a new development agenda began; instead, the SDG 4 Education Goal came about due to mobilization of the education community, allied with education's first level ranking on the "My World Survey." SDG 4 was designed with the intention of incorporating but also moving well beyond MDG 2 and 3 to take up important educational themes that were missing from these previous Development Goals. Thus, SDG 4 represents "a stronger and broader commitment to quality and equity in education" ([Wulff, 2020](#), p. 1) with individual targets that are much broader and covering areas that were absent from the MDG 2 and 3. Included in the SDG 4 are equitable access to primary and secondary education, early childhood development and pre-primary education, post-secondary/tertiary education of various forms, job-relevant skills, increased access for people with disabilities and other traditionally marginalized groups, adult literacy education, education for sustainable development, building of physical school facilities, increased scholarship abroad for higher education for students in low income countries, and increased teacher supply (see [Table 1](#)). Overall,

Table 1 Sustainable development goal no. 4, targets and indicators.

Outcome targets (4.1–4.7)	
Target 4.1	By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes
4.1.1	Proportion of children and young people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex
4.1.2	Completion rate (primary education, lower secondary education, upper secondary education)
4.1.3	Gross intake ratio to the last grade (primary education, lower secondary education)
4.1.4	Out-of-school rate (1 year before primary, primary education, lower secondary education, upper secondary education)
4.1.5	Percentage of children over-age for grade (primary education, lower secondary education)
4.1.6	Administration of a nationally representative learning assessment (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education
4.1.7	Number of years of (a) free and (b) compulsory primary and secondary education guaranteed in legal frameworks
Target 4.2	By 2030, ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education
4.2.1	Proportion of children aged 24–59 months who are developmentally on track in health, learning and psychosocial well-being, by sex
4.2.2	Participation rate in organized learning (one year before the official primary entry age), by sex
4.2.3	Percentage of children under 5 years experiencing positive and stimulating home learning environments
4.2.4	Gross early childhood education enrollment ratio in (a) pre-primary education and (b) early childhood educational development
4.2.5	Number of years of (a) free and (b) compulsory pre-primary education guaranteed in legal frameworks
Target 4.3	By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university
4.3.1	Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex
4.3.2	Gross enrollment ratio for tertiary education by sex
4.3.3	Participation rate in technical-vocational programs (15–24-year-olds) by sex
Target 4.4	By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship
4.4.1	Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill
4.4.2	Percentage of youth/adults who have achieved at least a minimum level of proficiency in digital literacy skills
4.4.3	Youth/adult educational attainment rates by age group and level of education
Target 4.5	By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations
4.5.1	Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated
4.5.2	Percentage of students in (a) early grades, (b) at the end of primary, and (c) at the end of lower secondary education who have their first or home language as language of instruction
4.5.3	Existence of funding mechanisms to reallocate education resources to disadvantage populations
4.5.4	Education expenditure per student by level of education and source of funding
4.5.5	Percentage of total aid to education allocated to least developed countries
Target 4.6	By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy
4.6.1	Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex
4.6.2	Youth/adult literacy rate
4.6.3	Participation rate of illiterate youth/adults in literacy programs
Target 4.7	By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and nonviolence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development
4.7.1	Extent to which (i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies, (b) curricula, (c) teacher education and (d) student assessment
4.7.2	Percentage of schools that provide life skills-based HIV and sexuality education
4.7.3	Extent to which the framework on the World Program on Human Rights Education is implemented nationally (as per the UNGA resolution 59/113)
4.7.4	Percentage of students in lower secondary education showing adequate understanding of issues relating to global citizenship and sustainability
4.7.5	Percentage of students in lower secondary showing proficiency in knowledge of environmental science and geoscience
4.7.6	Extent to which national education policies and education sector plans recognize a breadth of skills that needs to be enhanced in national education systems

(Continued)

Table 1 Sustainable development goal no. 4, targets and indicators.—cont'd

Means of implementation targets (4.a to 4.c.7)	
Target 4.a	Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all
4.a.1	Proportion of schools offering basic services, by type of service
4.a.2	Percentage of students experiencing bullying in the last 12 months in (a) primary, and (b) lower secondary education
4.a.3	Number of attacks on students, personnel and institutions
Target 4.b	By 2020, substantially expand globally the number of scholarships available to developing countries, in particular least developed countries, small island developing states and African countries, for enrollment in higher education, including vocational training and information and communications technology, technical, engineering and scientific programmes, in developed countries and other developing countries
4.b.1	Volume of official development assistance flows for scholarships by sector and type of study
Target 4.c	By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing states
4.c.1	Proportion of teachers with the minimum required qualifications, by education level
4.c.2	Pupil-trained teacher ratio by education level
4.c.3	Percentage of teachers qualified according to national standards by education level and type of institution
4.c.4	Pupil-qualified teacher ratio by education level
4.c.5	Average teacher salary relative to other professions requiring a comparable level of qualification
4.c.6	Teacher attrition rate by education level
4.c.7	Percentage of teachers who received in-service training in the last 12 months by type of training

Modified from [UN Institute for Statistics \(2021\)](#).

improved quality and equitable access under a rights-based umbrella serve as underpinning principles of the SDG 4. Further, the notion of education for sustainable development attempts to make education a more cross-cutting theme relevant to the overall goal of sustainability the SDGs seek to convey, and tries to make education seem more integrative of the Goals (Wulff, 2020). SDG 4 has 10 targets, that is, specific, measurable objectives which are expected to contribute to achieving the overall goal (Outcome Targets range from 4.1 through 4.7, and Means of Implementation Targets are 4.a to 4.c.; see Table 1). There are also 11 global indicators and 32 thematic indicators, which are markers of change or continuity that are expected to enable the measurement of progress toward meeting each target set to achieve the broader education Goal. The indicators are spread across four monitoring levels—global, thematic, regional and national. UNESCO-UIS (2018) notes that global level monitoring provides a global overview of progress toward each target, while thematic monitoring compares indicators within a specific sector (e.g., education, environment, energy, health) or cross-cutting themes (e.g., gender) and serves as a framework to track progress on a cross-nationally comparable basis. For regional monitoring, countries are expected to develop and agree to a set of indicators, and consider priorities and issues of common and shared interests of a particular region. National level monitoring is linked to the needs of national and sub-national governments in developing education sector plans and informing policies.

Optimism and virtues of SDG 4

Like the Education MDGs, the Education SDG, even more so, “is a valuable resource for lobbying and advocacy by civil society” (King, 2016, p. 22; also see Dost, 2020). With a more expanded set of objectives, targets and principles, SDG 4 lays out a wider array of education-related aspirations to advance development and justice in the modern era. Compared to the MDG Education Goals, SDG 4 targets “are transformative in scope and ambition, and ... they have the potential to challenge some of the pervasive forces in education” and are intended to be implemented under a rights-based approach (Wulff, 2020, p. 3). The broader, more transformative nature of the SDGs has direct implications for education too, due to “their universality and pledge to overcome silos within sustainable development” (Wulff, 2020, p. 11) which translate into the prospect of a global effort to address educational needs and for education to be considered in relation to all other goals and targets in the overall SDG agenda. Also, SDG 4 is significant in its endorsement of education for sustainable development. Target 4.7 is specific in its aspiration that “all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles ...”. Thus a notion of the “public good is implicit in the general SDG vision in which education is mentioned in targets for food security, women’s rights, decent work, health and equalities” (Allais et al., 2020, p. 135) and sustainable development. Additionally, because the SDGs overall are framed as Global in nature, the educational goals are more inclusive of marginalized people outside of the Global South compared the MDGs which were specifically designed for developing countries. By the principles of the SDGs being applicable to wealthy industrialized nations, as well, those pockets of educational inequities in “developed” countries should become subject to greater scrutiny and policy change under a global mechanism that has universal relevance.

The notion of inclusive education in the SDGs also takes more explicit forms under the umbrella of universal rights, for example, by “focusing on specific subgroups as target populations” such as “persons with disabilities, indigenous peoples and children in

vulnerable situations" (Johnstone et al., 2020, p. 97) as described in target 4.5 of Goal 4. Additionally, Goal 4 names other groups in various targets, including girls, rural children, students from developing countries, and small island developing states (SIDS) as educational rights bearers. Critically, indicators (as listed in 4.5.1) are attached to these targets that require member states to collect and provide data, which potentially creates an environment of verifiability and accountability and therefore accords rights of inclusion to various historically marginalized groups. As Johnstone et al. (2020) note, such rights are plural in the sense that they are "concerned with the presence, participation, and achievement of all children and particularly with educational needs of populations that have been historically marginalized" (p. 98). In real world context, Do et al. (2020), drawing on case study and evidence from Vietnam, conclude that these key rights-based elements in SDG 4 "can satisfactorily ensure the legal rights of ethnic minorities" (p. 15). Their study further note that by "... targeting substantive equality, these [aspects of SDG 4] resolve long-existing conflicts in equality law, while still focusing on human rights for underrepresented groups" (Do et al., 2020, p. 3), by offering concrete and relevant interventions for vulnerable groups to take place.

Another important aspect of SDG 4 is its movement of the education discourse beyond mere enrollment to one of learning outcomes and educational achievement. With this "quality turn," as Sayed and Moriarty (2020) refer to it, SDG 4 has brought about a focus on the quality of educational experiences compared to previous educational initiatives. The implication of SDG4's more qualitative focus is an "education in which one concentrates on each and every child as a whole, development of each student as social, emotion, mental, physical and cognitive irrespective of sex, gender, race, caste, religion, status and geographic location" (Rajput et al., 2020, p. 104). And because SDG 4 focuses on "learning as opposed to access to education" (212) and "not just as an overarching goal but embedded in the target" (Sayed and Moriarty, 2020, p. 195), notions of educational relevance, resource support, teacher quality, financing, among other educational issues, are forced to be parts of broader educational policy and planning. In this context, the promotion of lifelong learning also emerges as a key principle in SDG 4, going farther than the focus solely on universal primary and secondary education as was the case with the Education MDGs. With SDG 4, educational policy and practice agenda is extended from formal basic education to tertiary-higher education and continued learning and retooling for continued relevance of skills for the 21st century economy. The inclusion of higher education in SDG 4 is seen as a progressive development, keeping in mind that this level of education has traditionally been absent from previous education development initiatives, including the MDGs (Allais et al., 2020). It is noteworthy, too, that the often-limited educational capacities at the tertiary educational level in developing countries is recognized, and are support by international cooperation between developed and developing countries, and international scholarships as described in Target 4.b and Target 4.c. Here, there is a specific target to "substantially expand globally the number of scholarships available to developing countries" (4.b) and "substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries" (4.c).

The limits and liabilities of SDG 4

Like the SDGs overall, SDG 4 mirrors some of the very challenges, tensions and contradictions of Western modernist development and global public policy/governance. As it relates to the former, modernist development is subject to critique of being a Western led universalizing agenda with notions of imperialism where powerful industrialized countries set and dominate the international development agenda (Klak, 1998; Nisbet, 1980). In such a system, as many critical scholars will point out, there are power imbalances that define historical and existing relations globally and locally, which impinge on who shapes the dominant development discourse and impacts what rights and responsibilities emerge (Dost, 2020). Specifically with regard to SDG 4, Wulff (2020) suggests that because its contents were "agreed to by those in power in 2015, the agenda was bound to reproduce the power relations and imbalances of that time" (p. 2) (see also Dost, 2020; King, 2016). Likewise, differences in "level of development," regional, national and local priorities, as well as culture, religious beliefs, among other issues, create tensions and potential divergences from Globally set educational targets. With regard to the issue of global public policy and governance, global development goals and initiatives, like the MDGs and the SDGs, must contend with the complexities and tensions of pursuing international cooperation in the context of national sovereignty and multiple development actors with different interests and varying levels of power (Stone, 2008; Barnett and Duvall, 2018; Marshall, 2013; Dost, 2020). Though the framers of SDG 4—including governments, international organizations, non-profits organization, and civil society—claimed to have been deliberate in their efforts to make the Goals seem participatory in its formulation, content, implementation and monitoring and evaluation, different stakeholders have highlighted crucial challenges in the development of SDG 4.

One such challenge is that even though the SDG framework recognizes "different national realities, capacities and levels of development" and respects "national policies and priorities" (United Nations, 2015a, p. 21), the Western universalist nature of the Goals creates contradictions, tensions and inconsistencies between global and national (and regional) educational policy priorities (Ali, 2016). For example, when countries already have educational and regional plans that have been negotiated at every level but which do not consistently align with SDG 4, problems and tensions of various kinds can emerge, including a pressure to change national priorities through various forms of coercive global governance mechanisms including international aid (Ali, 2016; King, 2016). As Ali (2016) highlights by drawing on the context of the Sindh Province of Pakistan where the long negotiated Education Sector Plan is "being undermined by the new exigencies of SDG-4 Targets," educational planning becomes very difficult by the donor community who "continue to come with their own priority areas ... with money attached to these priorities" (Ali, 2016, p. 58). Where national priorities do take precedence, the notion of Global goals might not be so global after all, which may render the cost and effort of monitoring and evaluation SDG 4 at the global level for universality futile. Also, when there are such differences

between national and global educational goals, monitoring and evaluation become challenging and the efficacy of the SDG 4 cannot be truly assessed. This problem also reflects three broader methodological challenges of measurement and comparison of goals at the global level as cited by UNESCO Institute of Statistics. The first is that “national, regional and global data on basic competencies in literacy and numeracy are frequently collected but cannot be used in an integrated manner to create a global picture of learning” (UNESCO-UIS, 2018, p. 23). What this means is that different countries (or regions) use different assessment mechanisms rendering cross comparison of countries difficult or meaningless. Secondly, there is no shared technical standards to ensure data are of similar quality at national and regional levels. And third, is the challenge of taking into account the multiple viewpoints related to “what can and should be measured globally,” thinking of “how national and regional data can help inform global measurement,” and “striking an appropriate balance between global competencies and the role of local influences” and education goals (UNESCO-UIS, 2018, p. 23).

However, the most vociferous and consistent detraction of SDG 4 are the values that are embedded within the educational paradigm that the Goal endorses and is built on, which reflects some of the core Western visions of the world. A fundamental critique is that SDG 4 conforms to the same paradigm as that which has existed within the dominant modernist development agenda, which itself is built on colonial notions of coercion, hierarchies and universalism that favor Western ways and ideas (Sterling, 2016). For instance, SDG 4 is interpreted as fundamentally instrumentalist to the Western economic model, and committed to economic growth and therefore all other social, political, cultural and environmental imperatives are rendered subservient to the economic imaginary (Brissett and Mitter, 2017; Sterling, 2016). This adherence to the traditional development paradigm creates internal contradictions within SDG 4, because while it promises to liberate the world’s dispossessed and marginalized peoples, its theory of change seems to be based on the very educational systems that have supported capitalist, growth-driven consumerism that have created inequities, and ruined societies and their natural environments in the first place (Komatsu and Rapple, 2018; Dost, 2020). Rather than connecting various forms of human existence, education in the SDGs is primarily framed as education for employment, which reduces much of education’s potential liberating and transformational effects.

With regard to the environment specifically, the unquestioning of the current education paradigm results in the continued prioritization of economic growth over the environment (Komatsu et al., 2020). For example, there is no obvious recognition of the fact that countries having “better” education also have more detrimental impacts on climate change, or the fact that the world is the “most educated it has ever been and yet the nearest to environmental breakdown” (Komatsu and Rapple, 2018, p. 7), which suggests that there is a deep and fundamental flaw in the values embedded in existing educational model that SDG 4 is attempting to expand. Yet there is little attempt to interrogate and critically reconsider the existing model of education, and there is only anemic attempts to explore whether all education consistently support sustainable development (Komatsu and Rapple, 2018). As Sterling (2015) observes, while Goal 4 emphasizes education in terms of its potential socio-economic benefits, there is no recognition that education through awareness raising, training and capacity-building can help protect environmental quality and lead to wiser resource use. In fact, SDG 4 is seen by some as paying mere lip service to education for sustainable development by lumping it together with so many other topics in Target 4.7 that the focus on education for the environment is diluted (Sayed and Moriarty, 2020, p. 2121). Therefore, while SDG 4 was supposed to be a core clarion call for education to play crucial role in environmental preservation, its advent has not changed education’s “business as usual” approach and is therefore still in need of radical transformation itself for greater impact on sustainability.

As explored earlier, it is welcomed that SDG 4 is more ambitious than the education MDGs by expanding its focus on educational levels and framing quality education beyond mere access to schooling. However, even with these expansions to the educational agenda, “the concept of quality education and learning is narrowed and reduced in instrumental ways that reduce its potential reach and impact” (Sayed and Moriarty, 2020, p. 212). For example, definitions of indicators of quality seem more technicist and narrow and not sufficiently clear on what constitutes quality learning, particularly in defining how education should change behavior (that is, education’s affective component). There seems to be some acknowledgment of this concern, as in 2019, UNESCO (2019) conducted a study that “examined whether, and to what extent, three learning dimensions,” that is the cognitive, social and emotional and behavioral, “are prioritized in commitments to Education for Sustainable Development (ESD) and Global Citizenship Education (GCED) learning in pre-primary (PPE), primary (PE), lower secondary (LSE) and upper secondary (USE) education in a selection of countries from UNESCO’s five regions of the world” (p. 7). The study found that “the relative emphasis on each of the three interrelated learning dimensions varies across education levels and across countries, reflecting specific contexts, systems, and needs,” leaving UNESCO to encourage stakeholders to “pay more attention to [these] three learning dimensions” (p. 9).

Critics also add that there is a distinct quantitative bias of SDG 4 indicators and this advantages those characteristics of educational quality that are readily amenable to quantitative measurement; qualitative indicators that are more reflective of changed behaviors, perceptions, and people’s experiences and emotive responses to education are suppressed. Wulff (2020) also highlights the “narrow implementation of the Broad Commitment of Quality” (p. 21) by the disproportionate focus on measuring learning outcomes primarily in the areas of literacy and numeracy while other subjects and dimensions of quality are marginalized. Further, the narrow universal focus on quality as measurement of numeracy and literacy, along with its attendant narrow set of data points, pushes education “toward global convergence and denies the importance of contextually relevant education” (Wulff, 2020, p. 22). As a result, what is considered valuable knowledge becomes standardized regardless of context, and learning is narrowed and reduced to market-productive human capital foci. Some observers have also raised SDG 4’s non-engagement with the issue of language of instruction as a problematic education quality issue that should have been addressed by the Goal 4 (Milligan et al., 2020). Milligan et al. (2020), for example, cite voluminous research showing that language of instruction is central to learning

and therefore a key component of educational equity, given that “so many children around the world learn in a dominant language, often which is not their own for example colonial language or non-indigenous language” (p. 120).

An important component of SDG 4, like all the other Goals, is verification of their progress and effects through monitoring and evaluation and, as such, the Goals are heavily dependent on data collection. Some characterize this process as “standardization” and “datafication” of education that have now found their way into SDG 4 based on the assumptions that processes of teaching and learning can be standardized and used to measure the performance of students, teachers and overall education systems, regardless of context (Wulff, 2020; Sellar and Hogan, 2019). In addition to this reductive use of data, data collection can be expensive and this has increasingly become a critical challenge, especially for developing regions with limited data gathering infrastructure and tradition (Johnstone et al., 2020). This has the potential to pull already limited resources from important educational areas in order to satisfy data gathering requirement of the SDG 4, which may not necessarily advance key educational and equity goals. In fact, UNESCO (2020) in its *Global Education Monitoring Report 2020* cites problems of data collection and note that many countries still struggle to collect meaningful data, especially for inclusion of educationally disadvantaged populations. SDG 4’s preoccupation with certain types of standardized data also invites a gendered critique. Durrani and Halai (2020), for instance, note that SDG 4’s reliance on and preoccupation with measurement and international ranking “often misses the most important aspects of gender equality because those aspects cannot be quantified” (p. 94). Thus, the Education SDG is said to “reproduce colonial gender hierarchies, including gender binaries” where the most marginalized are often the most invisible social groups due to the nature of standardization and datafication of education (Durrani and Halai, 2020, p. 93).

Another consequential component of SDG 4, also related to monitoring and evaluation and accountability, is how the Goal will be financed. One irony that some have cited is that even though SDG 4 is much broader and more ambitious than the Education MDGs, commitment to resourcing the former has gone down from the recommended 6% of GDP to a more flexible range of 4–6% and from 20% of national budget to 15–20% (Archer and Muntasim, 2020; Wulff, 2020). However, it is estimated that at least \$39 billion of aid was needed annually to achieve just some of the main SDG 4 targets (King, 2016). The disappointing outcome of the Financing for Development (FfD) conference in Addis Ababa in 2015 that should have supported the development agenda, including SDG 4, is significant (Brissett, 2018; Archer and Muntasim, 2020). With the continued and even further diminishing budgets of developing countries, the SDG 4 framework has promoted public-private partnerships (PPP) as a solution to fill the wide financial gap of the education targets. However, scholars and activists are wary of the profit motive and market logic, as well as the reductive and instrumental view of education, that emerge when private interests shape education, and they worry about the privatization forces on equity and equality (Rizvi and Lingard, 2010; Wulff, 2020; Archer and Muntasim, 2020). There is genuine concern that as the market logic influences SDG 4, through such public private partnerships, private interests might “capture public funds to support commercial interests, whose long-term effects could be deeply damaging to quality, equality and the public good” (Draxler, 2020, p. 168).

Conclusion

The SDGs emerged out of a much wider agenda than the MDGs and were more participatory in their development. Education has taken an important place in the SDGs after initial doubts that it would have a designated place in the new global development agenda. Replacing MDG 2 and 3, SDG 4 expanded on these Goals to include equitable access to primary and secondary education, early childhood development and pre-primary education, post-secondary/tertiary education of various forms, job-relevant skills, increased access for people with disabilities and other traditionally marginalized groups, adult literacy education, education for sustainable development, building of physical school facilities, increased scholarship abroad for higher education for students in low income countries, and increased teacher supply. The prospects of SDG 4 rest on its claim of a rights-based approach, which underpins its more inclusive tone compared to the Education MDGs. SDG 4, therefore includes its widened capacity as a resource for lobbying and advocacy for education, and the adoption of a more elongated interpretation of education with a focus on lifelong learning to tertiary educational levels. But particularly important is the greater focus on quality education beyond mere access and enrollment, as well as its heightened, more deliberate focus on historically marginalized groups, including people with disabilities, indigenous peoples and children in vulnerable situation, among others.

However, SDG 4’s rights-based approach is juxtaposed with an instrumentalist approach where education, fundamentally, is built to support a particular economic model. As such, SDG 4 suffers from an internal struggle between a broader rights-based conceptualization and the instrumentalist approach to education. This latter approach generates limitations of SDG 4 that are similar to problems associated with the Western modernist development model of universalism and inequitable power relations, as well the challenge of global public policy making in a context of sovereign nations and multiple interests. Among the challenges emerging from SDG 4 in this context are the tensions between the SDG 4 Targets and those in countries and regions. However, most problematic are the fundamental values that some see embedded in SDG 4, particularly its commitment to the existing model of education that has historically supported what many believe to be a flawed economic system that has exploited people, societies and the environment. Thus, another central critique of SDG 4 is the lip service paid to education for the environment; many observers have noted the limited systemic integration of types of knowledge and learning that truly raise the natural world above its existing position of a consumable resource. And this critique is also tied to the degree to which some see SDG 4 as adhering to its rhetoric of quality education; here, questions are raised about how quality education is assessed due to the Goal’s bias toward standardization and datafication which limit the emergence of educational issues that are qualitatively important but not amenable to easy

quantification. These concerns are also extended to the extensive monitoring and evaluation necessary for SDG 4 and the associated costs. These above challenges are capstoned by the lack of financing for the SDG 4.

However, many laud these new global development Goals. And even with their challenges, there is still hope and, some would say, potential for the Education SDG 4 to serve and advance Global Citizenship and transform our world.

References

- Ali, S., 2016. Pakistan's Educational Planning and Alignment With the Education SDG. NORRAG NEWS, Agenda 2030—SDG4 Education 2030—One Year On: Challenges and Opportunities 54.
- Allais, S., Unterhalter, E., Molebatsi, P., Posholi, L., Howell, C., 2020. Universities, the public good, and the SDG 4 vision. In: *Grading Goal Four*. Brill Sense, pp. 135–155.
- Archer, D., Muntasim, T., 2020. Financing SDG 4: context, challenges, and solutions. In: *Grading Goal Four*. Brill Sense, pp. 170–193.
- Barnett, M., Duvall, R., 2018. International organizations and the diffusion of power. In: Weiss, T., Wilkinson, R. (Eds.), *International Organization and Global Governance*. Routledge, pp. 51–62.
- Brissett, N., Mitter, R., 2017. For function or transformation? A critical discourse analysis of education under the Sustainable Development Goals. *J. Crit. Educ. Pol. Stud.* 15.
- Brissett, N.O.M., 2018. Sustainable Development Goals (SDGs) and the Caribbean: unrealizable promises? *Prog. Dev. Stud.* 18, 18–35. <https://doi.org/10.1177/1464993417734440>.
- Do, D.-N.-M., Hoang, L.-K., Le, C.-M., Tran, T., 2020. A human rights-based approach in implementing sustainable development goal 4 (quality education) for ethnic minorities in Vietnam. *Sustainability* 12, 4179. <https://doi.org/10.3390/su12104179>.
- Dost, A.C., 2020. The goals concealed in the goal revealed: taking forward global education from SDG 4. *Econ. G20* 185–235. https://doi.org/10.1142/9789811214776_0008.
- Draxler, A., 2020. Education for all open for business? Public goods vs. Private profits. In: *Grading Goal Four*. Brill Sense, pp. 156–169.
- Durrani, N., Halai, A., 2020. Gender equality, education, and development: tensions between global, national, and local policy discourses in postcolonial contexts. In: *Grading Goal Four*. Brill Sense, pp. 65–95.
- Fehling, M., Nelson, B.D., Venkatapuram, S., 2013. Limitations of the Millennium development goals: a literature review. *Global Publ. Health* 8, 1109–1122. <https://doi.org/10.1080/17441692.2013.845676>.
- Hulme, D., 2009. The Millennium Development Goals (MDGs): a short history of the world's biggest promise. *SSRN Electr. J.* 1–55. BWPI Working Paper No. 100.
- Johnstone, C.J., Schuelka, M.J., Swadek, G., 2020. Quality education for all? The promises and limitations of the SDG framework for inclusive education and students with disabilities. In: *Grading Goal Four*. Brill Sense, pp. 96–115.
- King, K., 2016. The Sustainable Development Goals (SDGs) One Year On; The Case of Education—SDG4. NORRAG NEWS, Agenda 2030—SDG4 Education 2030—One Year On: Challenges and Opportunities 54.
- Klak, T., 1998. Thirteen theses on globalization and neoliberalism. In: Klak, T. (Ed.), *Globalization and Neoliberalism: The Caribbean Context*. Rowman & Littlefield, pp. 3–23.
- Komatsu, H., Rappleye, J., 2018. Will SDG4 Achieve Environmental Sustainability? Center for Advanced Studies in Global Education (CASGE) (Working Paper# 4). Arizona State University, Phoenix. Retrieved June 5, 2019.
- Komatsu, H., Rappleye, J., Silova, I., 2020. Will education post-2015 move us toward environmental sustainability?. In: *Grading Goal Four*. Brill Sense, pp. 297–321.
- Marshall, K., 2013. Global development governance. In: Weiss, T., Wilkinson, R. (Eds.), *International Organization and Global Governance*. Routledge, pp. 590–605.
- McGrath, S., Nolan, A., 2016. SDG 4 and the child's right to education. *Norrag News* 54, 122.
- Milligan, L.O., Desai, Z., Benson, C., 2020. A critical exploration of how language-of-instruction choices affect educational equity. In: *Grading Goal Four*. Brill Sense, pp. 116–134.
- Nisbet, R.A., 1980. *History of the Idea of Progress*. Transaction Publishers.
- Rajput, N., Gupta, M.K., Kumar, M.H., 2020. The impact and implication of SDG4 in India. *J. Compos. Theory* XIII, 101–114.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London; New York, NY.
- Sayed, Y., Moriarty, K., 2020. SDG 4 and the "education quality turn": prospects, possibilities, and problems. In: *Grading Goal Four*. Brill Sense, pp. 194–213.
- Sellar, S., Hogan, A., 2019. *Pearson 2025: Transforming Teaching and Privatizing Education Data*. Education International, Brussels.
- Sterling, S., 2015. Ensuring inclusive and equitable quality education and promote life-long learning opportunities for all. In: ICSU, I. (Ed.), *Review of the Sustainable Development Goals: The Science Perspective*. International Council for Science (ICSU), Paris.
- Sterling, S., 2016. A commentary on education and sustainable development goals. *J. Educ. Sustain. Dev.* 10, 208–213.
- Stone, D., 2008. Global public policy, transnational policy communities, and their networks. *Pol. Stud. J.* 36, 19–38.
- UN Institute for Statistics, 2021. Official List of SDG4 Indicators.
- UNESCO, 2016. *Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4: Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning*. UNESCO Digital Library.
- UNESCO, 2019. *Educational Content Up Close. Examining the Learning Dimensions of Education for Sustainable Development and Global Citizenship Education*. Organización de las Naciones Unidas para la Educación, la Ciencia y la
- UNESCO, 2020. *Inclusion and Education All Means All*.
- UNESCO-UIS, 2018. *Quick Guide to Education Indicator for SDG4*. UNESCO Institute for Statistics Montreal.
- United Nations, 2015a. *Transforming Our World: The 2030 Agenda for Sustainable Development*. Department of Economic and Social Affairs, United Nations, New York.
- United Nations, 2015b. *The Millennium Development Goals Report 2015*. United Nations, New York.
- United Nations, 2015c. *United Nations Millennium Development Goals* [WWW Document]. <https://www.un.org/millenniumgoals/education.shtml>. (Accessed 6 February 2021).
- Wulff, A., 2020. Introduction: bringing out the tensions, challenges, and opportunities within sustainable development goal 4. In: Wulff, A. (Ed.), *Grading Goal Four*. Brill | Sense, pp. 1–27. https://doi.org/10.1163/9789004430365_001.

Labor markets, SDG4 and vocational education and training

Leona M. English^a and Peter Mayo^b, ^a St. Francis Xavier University, Antigonish, Nova Scotia, Canada; and ^b University of Malta, Malta

© 2023 Elsevier Ltd. All rights reserved.

Adult education (goal 4) as a response	548
Language acquisition	548
Recognizing prior learning	549
Discussion	550
Final note	551
References	552

The Sustainable Development Goals scheduled for completion by 2030 stand as a beacon to the international community of a collective hope for quality life on the planet. They are especially important as we face enormous issues like climate change, migration and unbridle capitalism and globalization. Addressing each of the 17 goals is a challenge that has replaced the once equally bold, MDGs (UN, 2015) which ended unsatisfactorily in 2015. The 17 goals expand greatly the dreams of the 6 MDGs and show our collective energy and will to meet Agenda 2030. In Goal 4 and its ten targets, the UN (UNESCO, 2016) named its hope that it would “Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all”. This education or lifelong education goal became the focus of UN agencies like UNESCO which have heightened capacity in this area. However, they have not forgotten that the accomplishment of Goal 4 is deeply connected to all the goals, especially decent work (Goal 8), gender equality (Goal 5), poverty reduction (Goal 1), and health and well-being (Goal 3) (English and Mayo, 2019).

This chapter focuses primarily on Goal 4 and its targets as the way forward to reaching sustainability in 2030. A key piece of Goal 4 is that it is, unlike the MDGs which focused on children in school, inclusive of adult education. Despite the optimism of a lifelong learning approach, researcher Elfert (2019) offers a caution to the hope for adult education found in the SDGs. Her careful appraisal draws attention to the overall global prioritization of the education of children, the general lack of funding for adult learning and education, and the negligible interest in educating marginalized groups. Her analysis is clear and convincing, and supported by robust data. Yet, Goal 4 and the other 16 SDGs are the planet’s best means of fulfilling its global commitment to sustainable development. Education bears a major responsibility and is arguably, due to its reach and its centrality in the global knowledge economy, the main goal.

Goal 4 on education is the key goal in addressing many world crises, especially migration and the flows of people from country to country and state to state. The IOM reports that in mid-2019, the latest statistics, there were 271.6 million migrants (IOM, 2020). Yet, nations have failed to furnish comprehensive policies to welcome them and, more radically, change the unequal neocolonial relationship between North and South, the only way to really stem the tide of people leaving their homes to seek pastures new elsewhere. Education can constitute one of the contributory variables to change.

International pronouncements aside, even continents like Europe do not have a universal policy for immigration, let alone an educational one, for all the EU talk of LLL for employability and active citizenship. Each individual state is expected to create its own policies and responses, in part because the hope of coming to consensus is low. Within countries in Europe, for instance, the response to immigration has been mixed, especially in countries that faced situations of their empire striking back, not in Stuart Hall’s (Hall, 2017) sense of harkening to a sense of Empire but in the sense of victims of imperial politics seeking recourse to the former empire, and which have been witnessing the rise of Right wing, including Far Right, movements/parties (e.g. Lega Lombarda in Italy, UKIP and the BNP in Britain, the Vlaams Belang Party in the Flemish region of Belgium and the Front National in France) appealing to swathes of the “autochthonous” and even settled immigrant populations, the irony apparently lost on the latter. This mixed reception stands in sharp contrast to the EU’s hope for a Social Europe or the invocation of the long history of progressive social, economic and educational policy that is often generally attributed to Europe (Mayo, 2017).

European counties in the Mediterranean like Malta struggle to accommodate migrants, not only those who came in the waves of migrants in 2016 but those who continue to arrive, escaping inhospitable conditions in their countries of origin. Migrants are attracted, for the most part, toward the Eldorado, the image of the Western world as implanted by different means of colonial education. Even after the massive waves of migration in 2016–17, migration within and between countries and continents remains a crisis or, as in settler countries like Canada, a way of life. This country was colonized by the English and the French who came to live on Indigenous lands and it continues to welcome newcomers. Indeed, Canada continues to be a settler nation as about 20% of the population is foreign born, with the numbers much higher in cities like Toronto and Vancouver where the percentage is almost 50% (Statistics Canada, 2017). This country has a complex ethnic makeup and defines itself as multicultural, so immigration, migrants and refugees have been part of its history and its makeup, though it has a strict immigration policy that priorities skills and employability. This is also true of Australia (Webb et al., 2017). This system has its weaknesses, since many professional migrants are unemployed and underemployed. Canada struggles with social integration and employability, especially with the shift to the right in American governance.

It bears saying here that migration is not always occasioned by a crisis. Countries in the west often open their doors to migrant workers, who enter the country often as candidates in programs like the Seasonal Agricultural Workers Program (SAWP) and Live-in Caregiver Program (LCP) in Canada, designed to fill jobs that Canadians will not or cannot do. These workers become an integral part of our society and yet are denied full citizenship rights (Perry, 2019; Sharma, 2012; Walia, 2010). They are often tied to their employer, left with low wages and little choice, becoming unwittingly part of a global economy that capitalizes on their labor but offers little in return. Indeed, these migrant foreign workers are the mainstay of fish plant operations in Prince Edward Island, the heart and soul of fruit harvesting in Ontario and Nova Scotia. Filipinos, Jamaicans, and others are welcomed to do jobs that other Canadians will not do, yet they often remain in limbo without adequate access to citizenship or rights and benefits such as health and social services (Barber, 2008). As well, while the migrants live and work in the new country, they send remittances back to their country of origin and maintain close family connections, disrupting the notion of what it means to live in the new country. The United States has similar issues.

None of these nations has a foolproof strategy for immigration. Australia and Canada struggle with integrating non-European migrants, even highly skilled ones, and have been unable to make migration as smooth as it could be (Webb et al., 2017). Clearly, there are complicated relationships among and between states and citizens. While one would hope that universal sustainable goals such as quality education (Goal 4), gender equity (Goal 5), and decent work (Goal 8) and health (Goal 3) would be paramount, it seems that for now they are being sidelined. We look closer at the specific issues that confront them might be integrated in a fulsome immigration approach that integrates learning. Indeed, we see them as a map to be followed in creating a response.

Adult education (goal 4) as a response

As citizens move, the world is watching and trying to respond to the challenges that are inherent in the move: ability to find decent and meaningful work (Goal 8) and the ability to access education (Goal 4) or have credentials recognized. Here we follow the revised UNESCO Recommendation on Adult Learning and Education (RALE; UNESCO, 2016) which sees adult education as crucial in working with those who are in a precarious state. It calls for special attention to migrants and displaced peoples. This 2016 revision of the 1976 recommendation on adult education by UNESCO, was brought forward because of the international CONFITEA process in Belem, and is continuously being monitored as UNESCO realizes the value of global consensus on such an important issue. This understanding of the role of adult education is in sync with and supports the need to assist those who are migrating as they have been “affected by conflict or disaster, [or are] refugees, stateless or displaced persons” (UNESCO, 2016). Attention to the adult in this case is a means of addressing family needs such as seeking employment, language skills and inclusion in a new environment. In addressing education, receiving countries are better able to help the whole family or extended family. This serves as an important document in drawing our attention to the mandate of these countries, as part of their adult education agenda, to look to these asylum seekers, to provide them with help for full integration.

Language acquisition

Related to the attention to adults is second language acquisition which may be needed, as migrants may or may not have the language or indeed the literacy skills necessary to succeed in the new environment or state. In most states, language proficiency is required for participation in the labor market and all too often this proficiency is determined by test taking that may or not correspond to the job being sought or the society in which one will live. Gibb (2015) notes that in this “global knowledge economy, the new work order involves a “new word order”” (p. 251); language can be used to keep people out—it can be a way of strengthening national identity and building borders of exclusion for those found deficient. Rather than assisting with immigration, stringent language policies and lack of instruction support the continued dominance of western colonial powers and the economic and social stratification of the newcomer. Gibb points to the need for those language assessors and tests for assessment to be held to high standards of training and review, else they further reduce the complexity of language to a written test and ultimately fail the newcomer.

Second language acquisition can be a troubling area for women, in particular (see Webb, 2015). To begin with, learning another language does not guarantee employment or integration in receiving countries. Ennser-Kananen and Pettitt (2017) speak to the complexity for women engaging in second language learning in Austria and the US, noting the ways in which these women are positioned as deficits and as passively accepting banking education, while the providers are situated as saviors of the lost. The migrant may also be the target of an attempt to erase the culture and societal practices that were part of the migrant’s past. The assimilation techniques used in many language classes may have the potential of further stereotyping and limiting them. The authors call for the development of programs that develop sustainable agency. This approach works toward the goal of gender equality (Goal 5).

A clear caution on our language and other requirements is that they need to have flexibility and to account for the complexity of tasks, people, social and cultural situations. One of the things that we most forget when we get involved in work with migrants—is the need to incorporate the element of difference: societies that are integrative, not assimilative, work better for migrants (Shan, 2015). Fejes (2019) offers a particular critique of Sweden where welcoming programs like *Swedish from Day 1* work in many ways to homogenize immigrants, to valorize Swedish culture and language, and to ignore the ways that Sweden has already changed by its inclusion of asylum seekers. To present the culture as monochromatic and unchanging is to deny both Sweden and the immigrant.

Language is a key element in the adult education of migrant workers in most countries. In receiving countries like Malta there had been calls to transform the Maltese open centers into LLL centers involving the teaching of skills for either inclusion or relocation, with support from the potential receiving country in the latter case (Mayo, 2017). This seems to be coming to pass, as some of the open centers in Malta have witnessed projects by NGOs in teaching literacy in both Maltese and also English. These programmes are held either inside or outside open centers, in latter case at NGO premises, the Workers' Memorial Building (the main premises of the country's largest trade union, the General Workers' Union) or the Blue Door Institute, the latter managed and financed by St Andrew's Scots Church situated across the street from the trade union building in Valletta, the capital city (Medbalt, n.d., p. 12). The issue of relocation is a key one in recent European policy, with a strategy devised by the EU for this purpose. Few EUMS, however, responded to the strategy, with only Malta and Finland on course to fulfilling their obligations in this regard. Hungary, Austria and Poland, are refusing to participate in the scheme, while the Czech Republic, Bulgaria, Croatia and Slovakia have been critical and unsupportive to date (EC, 2017).

As for actual AE provision in Malta, migrants, regarded as Third Country Nationals (TCNs), enjoy the right of access to educational and vocational training as Maltese citizens, although there are differences with regard to costs incurred. While Maltese and EU citizens have the opportunity to attend MCAST and basic courses offered by the Directorate of Lifelong Learning, in the Ministry of Education and Employment (MEDE), free of charge while they pay for level 1 courses at a subsidized rate, migrants have to pay for the basic courses, and have to pay a higher fee to attend the level 1 courses. Refugees and people under subsidiary protection are often exempt from these charges. An effective LLL strategy intended to include migrants from third countries in Maltese society should address the course price issue, perhaps based on means testing. (Medbalt, n.d, p. 11).

One initiative addressing some of the needs of Third Country Nationals was organized by the state-funded corporation Employment & Training Corporation (ETC) as part of a EU co-financed project (IF 03-2012 Integrating TCNs in Maltese society). The program, "Living and Working in Malta, consisted of the provision of information regarding documents governing work in Malta and basic cultural information, all intended to abet the integration process. This kind of education is the staple of adult education for migrants characterized by what Paulo Freire would call "assistencialismo" (welfarism) (Freire, 1994), a notion of knowledge transmission rendering the learner "object" rather than "subject" (Medbalt, n.d., p. 11).

Recognizing prior learning

Goal 8 on RVA (Recognition, Validation and Accreditation), which has a cache in many spheres, as it purports to provide credentialing and recognition of acquired skills and abilities, with the intent of integration into the labor market, regardless of where they were trained or credentialed or where they previously worked (Singh, 2015). The term RVA is used by UNESCO (2012) to refer to the recognition, validation, and accreditation of the outcomes of nonformal and informal learning. The intent is to have a means of recognizing and credentialing learning that one has accumulated through everyday or otherwise developed purposes. In documenting best practices around the world, in developed and so-called developing contexts, Singh observes that some countries have been more effective than others in implementing the systems of RVA or prior learning and assessment and recognition (PLAR), as it is variously known (Andersson and Guo, 2009).

Despite western European attempts to provide equalization in areas such as university credits (the Bologna Process), there is little evidence these have had the results they expected. In the USA, the system of recognition is located mainly in the higher education sector where mechanisms such as portfolio assessment exist for learners to get advanced standing for degrees and diplomas (Singh, 2015). These systems can provide credit for experiences attained outside formal institutions of learning. In the secondary sector, the US and Canada have systems in place for adult learners to challenge the General Education Development (GED) program in order to attain a high school diploma. Australia, in contrast, has stronger RVA mechanisms in its vocation, education and technological (VET) sector. South Africa has a comprehensive approach that allows for assessment to help in gaining credits, to allow access and to facilitate advancement. The study of RVA makes clear that to decrease barriers to entry to education, the workplace, and social networks, each country needs mechanisms that provide equivalencies and ease access. They are essential to achievement of the SDGs.

One of the key issues raised by the RVA system in a world of knowledge migration and labor market needs is that many countries, including Canada, have sophisticated sets of language proficiency tests, the Canadian Language Benchmarks (CLB) being one. As problematic as the CLB is, it does exist and it is under necessary monitoring and scrutiny (Gibb, 2015). Attention has been drawn to the CLB's competency based framework and its inattention to language as a social practice. Meanwhile, there is no national or adequate RVA mechanism to assess skills and abilities in a similar way, for all occupational groups. Those systems that do exist are ad hoc and need to be more robust (Kahanec et al., 2013) and seen as trustworthy if they are to maintain their usability and practicality. Clearly, RVA is complex and multifaceted and it affects movement of migrants between provinces and countries.

Goal 8 for decent work is key to those migrating steadily across borders, carried and embodied by adults who are on the move. The recognition of diploma and qualifications, according to the Malta Qualifications Network levels, in synch with the European Qualification Framework for Lifelong Learning, falls under the National Commission for Further and Higher Education (NFCHE) that is also responsible for the validation of continuing vocational education and training (CVET). TCNs interviewed for the Medbalt Report by the NGO, Kopin (Medbalt, n.d, p. 11) identified the main problem as being the slowness of the bureaucratic process involved.

Many parts of Europe have been witnessing a situation where people are gaining more qualifications than their parents but cannot enjoy their standard of living, often working in precarious conditions (English and Mayo, 2012). Highly qualified graduates

from Europe emigrate from their country as “employability” does not translate into employment, as Ettore Gelpi once warned (Gelpi, 2002). The translation of skills and abilities across borders is required, as deskilling is the product of an inability to translate. While this problem negatively affects male immigrants, it is especially problematic for immigrant women. Drawing on a research project she participated in (Mirchandani et al., 2010), Slade (2015) notes that women in the study were immigrants with professional degrees in their home countries, but they were able only to find precarious work in Canada, work that in no way approached the level of skill and ability they had attained in their home country.

In Canada, the Federal Skilled Worker Program allows foreign trained workers to become credentialed (Government of Canada). Yet, anyone who has worked with these programs to be credentialed will realize that they are slow, flawed and not always accommodating. In short, as Maitra and Guo note (2019), “curricular and pedagogical approaches associated with lifelong learning can also have colonial undertones”. Guo sees this often as a racial issue (see also Webb et al., 2017). He points out that “It is the “color” of the skill associated with immigrants’ skin color rather than the skill itself which causes the deskilling and devaluation” (cited in Maitra and Guo, 2019). In Maitra’s interviews with racialized immigrants, this came through in the ways in which migrants were told to dress, talk and act like Canadians in order to get a job. This raises the issue of who is Canadian and whether conforming to desired behaviors is justified. Eurocentrism serves as a normalizing discourse that limits the migrant other.

Even in cases where migrants are able to access PLAR mechanisms, there are challenges. Andersson and Guo (2009) note that these mechanisms can be both controlling and devaluing of migrants’ experiences, despite the claims of the providers and supporters. The result is underutilization of skills and the devaluing of people, especially women. Looking especially at the Canadian system, they point to the lack of a centralized agency to assess qualifications and skills and to the existence of a diffuse array of bodies and organizations that serve to accredit immigrants. A similar view is provided, on the basis of research, by the Migrant Women Association of Malta whose sample contained women having significant levels of education. 93.3% of the surveyed were literate, having attended school. Over half completed secondary education, 19.8% completed university studies and 14.8% pursued vocational training. Those who pursued post-secondary education studied such areas as Law, Nursing, Economics and Business Administration. These were significantly represented (34.7%) (Merino, 2017, p. 7). Discrimination was flagged by all national groups represented in the survey. This discrimination took different forms with age being primarily reported by Syrians (81%), while Somalis cited factors, concerning race and gender, were reported solely by Somalis. Somalis and Libyans also reported religious discrimination and some Syrians cited “wearing hijab” as “an additional basis for discrimination” (Merino, 2017, p. 9).

Discussion

Nation states are struggling with migration and with the basics of providing support for newcomers. The conundrum is how to continue leading a nation of citizens while integrating new citizens in meaningful ways so that the quality of life for all is advanced. Governments are deeply immersed in the need to achieve the SDGs, not only for their own citizens but for the global community. From an adult education and lifelong learning perspective, the task is daunting, especially in light of the immediacy of the 17 SDGs slated to be achieved by 2030. We have a challenge like never before but we also have the tools and the knowhow to address it.

There is hope in knowing that migrants themselves are active in their own cause and that they both acquiesce and resist the sometimes oppressive systems they come into. In some cases, this is plain and simply bad behavior. In others, it is finding more and more black market sources of employment, and in others playing along with the system. As Mirchandani et al. (2010) remind us, migrants have always found ways to negotiate troubling issues. Given the unprecedented waves of migration in the 21st century, both invited, as in the Canadian foreign worker programs (Sharma, 2012), and uninvited as in the case of the refugees who literally washed ashore in the Mediterranean in 2016, the global community is facing issues of enormous import for the cohesion of societies and the strengthening of the economic and social systems in which they are embedded. While there has been a coordinated effort internationally to identify the SDGs and to articulate the world we want, there has not been a coordinated international response, in terms of education, health and work. The timely accomplishment of the SDGs demands that their needs be addressed. Agenda 2030 is impossible without active government and citizen support and advocacy for migrants.

Those concerned with adult education are in a key position to address the SDGs not only through assisting migrants to become integrated through language and housing efforts, but also through addressing government policies that affect the health and well-being of peoples (Goal 3). Operating on three levels—local and national, as well as international—those concerned for lifelong learning can help disturb the status quo of neoliberal regimes.

The challenge becomes deeper if we are to take the issue of democracy and social inclusion seriously. This raises important epistemological questions regarding the nature of knowledge. People who migrate have multiple subjectivities beyond that of working for the system, if at all. They represent the phenomenon of “portability of cultures” wisdoms and learning traditions across the globe. This presents the challenge for adult educators, as all educators, for that matter, to tackle the issue of “cognitive justice,” in Boaventura de Sousa Santos’ terms (de Sousa Santos, 2017) and be open to what the same Portuguese sociologist calls “epistemologies of the South” (de Sousa Santos). This takes us back to the earlier question of whose knowledge do we valorize or take on board and whether much of the credit for the origins of this knowledge should lie with, and therefore be restored to, the peoples from among whom many migrants emerge. Much of the knowledge labeled and patented as western is southern in origin. This applies to knowledge adopted in the academic or economic fields.

The best welcome one can provide to people new to the country is recognition of their contribution to the knowledge that has been appropriated and stolen from them. This would mark the restoration of cognitive justice. The other related task is to

learn from migrants' practical and holistic skills, many of which were discarded, in official colonial circles, as too Indigenous and unscientific, judged from a narrow western science standpoint. Some of this knowledge and its underlying philosophy can help teach western societies and their economies a few profound lessons regarding how one can produce in harmony with the rest of nature, using positive natural energy in lieu of that which pollutes, exacerbates global warming and threatens planetary existence, imposing rhythms beyond natural capacities. As [Shiva \(2016\)](#) puts it: soil not oil. In this way, migrants are not simply serving as objects of a top down approach to work but are potential agents of change in work, raising questions about not only how we produce but most importantly what do we produce, in whose interests and to the detriment of whom. All is connected in this world. Unfortunately, this is obscured through the reproduction of a North and South economic imbalance and a myopically perceived separation between humans and nature, the latter appearing to be there for conquest and exploitation. The real challenge for learning, living and producing is to recognize that we do not stand apart from nature but are rooted in it, in a global eco-system. Without romanticizing the cultures of migrants from continents with their own traditions and philosophies, or any culture for that matter, this cosmic vision lies at the heart of some of the traditions in which they have been raised. Think of the First Nations of the United States and their concept of the web of life ([Berry, 1999](#); [O'Sullivan, 1999](#)). Migrants from these backgrounds have much to teach the West with their "portability of cultures" as they cross frontiers and settle in foreign lands. Authentic dialog in the great Freire tradition can help facilitate this process of educational exchange which, in the long run, can help transform the nature of economic practices and the world in which we live.

More immediate, with regard to our present existential situation, without losing sight of the above long term vision, we conclude by offering some key points to consider with regard to policy making, curriculum development and face to face practice in adult education with sustainable development in mind. Education as it is viewed with Goal 4 speaks to the needs of migrants as it addresses their disposability and their vulnerability in a world that is troubled by racism, economic inequality and a constant search for sustainable livelihoods. We see exciting initiatives to address this through bodies such as UNESCO who are working with Member States to find effective policies and strategies to ensure the SDGs are accomplished, from the perspective of Goal 4.

Lifelong learning, stretching across the lifespan, has a key role in integrating all of the sustainable development goals. In its broadest sense, lifelong learning incorporates the informal, nonformal and formal dimensions of learning and infiltrates all the goals and their development.

It is able to help in creative ways, most predictably in preparing citizens for work. Nation states need to work with existing educational and credentialing bodies to make this a reality. Countries like Canada can do more to integrate and credential immigrants who are often highly qualified and skilled. An official central agency such as [Andersson and Guo \(2009\)](#) call for would help in this regard. It would also decrease the drive to westernize Canadianize immigrants through uniform schooling and higher education, as well as PLAR mechanisms that are obsolete and racist.

Adult Education has a direct role to play in maintaining and creating a viable economy. Education in all senses is important here. Sustainability is an ongoing process in a country like Canada which is not a resource rich and yet struggles to find ways to maintain our environment and provide income and decent work (Goal 8) to peoples, especially Indigenous peoples. ESL and EAL can help in creating opportunities for learning and engaging others. This, however, needs to be teaching by qualified educators in these fields, and this without taking anything away from the sterling work provided by volunteers engaged with NGOs. Qualifications matter and so all ESL instruction requires qualified educators ([Smythe, 2018](#)).

Geographic regions containing migrants need to have educational centers for social inclusion or resettlement. These need to be anti-racist and inclusive of efforts that engage productive discussions and practices of the underlying question of power: who dialogs with whom and from which position of power? This, to our mind, constitutes a key critical pedagogical question. It entails turning around some of our more limited efforts. Models of inclusion exist in cities like Toronto which are immigrant rich and multicultural. Community education programs in the context of adult education would foster the spirit of inter-ethnic dialogue.

Final note

Active citizenship requires us to invest in lifelong learning. Migrants need not just skills but also engagement in the democratic participation that is not limited to work. To have gender equality (goal 5) acknowledged is one key part of this work. Workers' education, in the context of adult education, can provide an understanding of the nature of workers' solidarity in this day and age, reaffirming the strong connection between class and ethnicity, indicating that categories as working and middle classes are international in scope and not national, thus helping to avoid misplaced alliances against the competition. LLL can also foster appreciation of the role of immigrant labor in the marketplace, in the homeplace and in the very fabric of our societies in the West. Their labor is value added to our collective wellbeing. The vestiges of colonialism must be dismantled.

Perhaps one of the key ways that lifelong learning can help in the achievement of the SDGs is its role in critical thinking and critique ([English and Mayo, 2021](#)). It can help in de-coding colonial language and policies, and in helping to understand global issues such as migration. It can help in tackling the issues once they are understood. Broader than schooling, lifelong learning can envelope the whole way of life and being, and help move us forward in addressing the predatory situation that wrecks havoc on the environment and people's lives, especially southern people's lives.

References

- Andersson, P., Guo, S., 2009. Governing through non/recognition: the missing "R" in the PLAR for immigrant professionals in Canada and Sweden. *Int. J. Lifelong Educ.* 28 (4), 423–437. <https://doi.org/10.1080/02601370903031264>.
- Barber, P.G., 2008. The ideal immigrant? Gendered class subjects in Philippine–Canada migration. *Third World Q.* 29 (7), 1265–1285.
- Berry, T., 1999. Foreword. In: O'Sullivan, E. (Ed.), *Transformative Learning. Educational Vision for the 21 Century*. University of Toronto Press, pp. xi–xv.
- de Sousa Santos, B., 2017. *Decolonising the University: The Challenge of Deep Cognitive Justice*. Cambridge Scholars Publishing.
- EC, 2017. Commission Calls for Renewed Efforts in Implementing Solidarity Measures under the European Agenda on Migration Press Release 2 March. http://europa.eu/rapid/press-release_IP-17-348_en.htm. (Accessed 13 March 2017).
- Elfert, M., 2019. Lifelong learning in Sustainable Development Goal 4: what does it mean for UNESCO's rights-based approach to adult learning and education? *Int. Rev. Educ.* 65, 537–556. <https://doi.org/10.1007/s11159-019-09788-z>.
- English, L., Mayo, P., 2012. *Learning With Adults. A Critical Pedagogical Introduction*. Sense Publishers, Rotterdam, Boston & Taipei.
- English, L.M., Mayo, P., 2019. Lifelong learning challenges: responding to migration and the sustainable development goals. *THEME Issue on Lifelong Learning Int. Rev. Educ.* 65 (2), 213–231. <https://doi.org/10.1007/s11159-018-9757-3>.
- English, L.M., Mayo, P., 2021. Lifelong Learning and the Neoliberal Scourge. *Sustainable Development, Global Social Justice and Social Solidarity*. Palgrave-Macmillan.
- Enns-Kananen, J., Pettitt, N., 2017. "I want to speak like the other people": second language learning as a virtuous spiral for migrant women? *Int. Rev. Educ.* 63, 583–604. <https://doi.org/10.1007/s11159-017-9653-2>.
- Fejes, A., 2019. Adult education and the fostering of asylum seekers as "full" citizens. *Int. Rev. Educ.* 65, 233–250. <https://doi.org/10.1007/s11159-019-09769-2>.
- Freire, P., 1994. *Pedagogy of Hope: Reliving Pedagogy of the Oppressed*. Continuum, New York & London.
- Gelpi, E., 2002. *Lavoro Futuro. La Formazione Professionale come Progetto Politico. Future Work. Vocational Education as a Political Project*. Guerini e associati, Milan (in Italian).
- Gibb, T., 2015. Regimes of language skill and competency assessment in an age of migration: the in/visibility of social relations and practices. *Stud. Cont. Educ.* 37 (3), 251–266. <https://doi.org/10.1080/0158037X.2015.1043986>.
- Government of Canada, 2020. Eligibility to Apply as a Federal Skilled Worker (Express Entry). <https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada/express-entry/eligibility/federal-skilled-workers.html>. Website).
- Guo, S., 2013. Lifelong learning in the age of transnational migration: emerging trends and challenges. In: Guo, S. (Ed.), *Global Issues and Perspectives*. Routledge, London, pp. 6–21.
- Guo, S., Maitra, S., 2019. Decolonising lifelong learning in the age of transnational migration. *Int. J. Lifelong Educ.* 38 (1), 1–4. <https://doi.org/10.1080/02601370.2018.1561534>.
- Hall, S., 2017. *The Empire Strikes Back in Selected Political Writings: The Great Moving Right Show and Other Essays*. Duke University Press, pp. 200–206.
- IOM, 2020. IOM Data Portal. International Organization for Migration, Geneva. https://migrationdataportal.org/?i=stock_abs_&t=2019.
- Kahanec, M., Kim, A.M.-H., Zimmermann, K.F., 2013. Pitfalls of immigrant inclusion into the European welfare state. *Int. J. Manpow.* 34 (1), 39–55. <https://doi.org/10.1108/01437721311319647>.
- Maitra, S., Guo, S., 2019. Theorising decolonisation in the context of lifelong learning and transnational migration: anti-colonial and anti-racist perspectives. *Int. J. Lifelong Educ.* 38 (1), 5–19. <https://doi.org/10.1080/02601370.2018.1561533>.
- Mayo, P., 2017. Storming "Fortress Europe" in Search of a Social World. *Counterpunch*. <http://www.counterpunch.org/2017/03/01/storming-fortress-europe-in-search-of-a-social-world/>. (Accessed 5 March 2017).
- Merino, A.C., 2017. *Stepping Up. An Investigation of Female Asylum Seekers and Employment in Malta*. Migrant Women Association of Malta, Malta.
- Mirchandani, K., Ng, R., Coloma-Moya, N., Maitra, S., Rawlings, T., Siddiqui, K., Shan, H., Slade, B., 2010. Transitioning into contingent work: immigrants' learning and resistance. In: Sawchuk, P., Taylor, A. (Eds.), *Challenging Transitions in Learning and Work: Reflections on Policy and Practice*. Sense Publishers, Rotterdam.
- O'Sullivan, E., 1999. *Transformative Learning. Educational Vision for the 21 Century*. University of Toronto Press.
- Perry, J.A., 2019. Images of work, images of defiance: engaging migrant farm worker voice through community-based arts. *Agric. Hum. Val.* 36, 627–640. <https://doi.org/10.1007/s10460-018-9861-9>.
- Shan, H., 2015. Settlement services in the training and education of immigrants toward a participatory mode of governance. Theme issue: Transnational migration, social inclusion, and adult education N. Dir. *Adult Cont. Educ.* 146, 19–28. <https://doi.org/10.1002/ace.20128>.
- Sharma, N., 2012. The "difference" that borders make: "Temporary foreign workers" and the social organization of unfreedom in Canada. In: Lenard, P.T., Straehle, C. (Eds.), *Legislated Inequality: Temporary Labour Migration in Canada*. McGill-Queen's University Press, Montreal, pp. 26–47.
- Shiva, V., 2016. *Soil, Not Oil. Climate Change, Peak Oil and Food Insecurity*. Zed Books.
- Singh, M., 2015. *Global Perspectives on Recognising Non-formal and Informal Learning*. Springer, New York.
- Slade, B.L., 2015. Migrating professional knowledge: progressions, regression, and dislocations. Theme issue: Transnational migration, social inclusion, and adult education N. Dir. *Adult Cont. Educ.* 146, 65–74. <https://doi.org/10.1002/ace.20132>.
- Smythe, S., 2018. Adult learning in the control society: digital era governance, literacies of control, and the work of adult educators. *Adult Educ. Q.* 68 (3), 197–214. <https://doi.org/10.1177/0741713618766645>.
- Statistics Canada, 2017. Immigration and Ethnocultural Diversity: Key Results From the 2016 Census. <https://www150.statcan.gc.ca/n1/daily-quotidien/171025/dq171025b-eng.htm?ind=14428-1&indgeo=0>. (Accessed 17 March 2020).
- UIL (UNESCO Institute for Lifelong Learning), 2016. Recommendation on Adult Learning and Education. UIL, Hamburg. Available at: <http://unesdoc.unesco.org/images/0024/002451/245179e.pdf>.
- UN (United Nations), 2015. *The Millennium Development Goals Report 2015*. UN, New York. http://www.undp.org/content/dam/undp/library/MDG/english/UNDP_MDG_Report_2015.pdf. (Accessed 8 November 2018).
- UN (United Nations), 2016. Sustainable Development Goal 4: Targets and Indicators [Dedicated Webpage]. UN, New York. <https://sustainabledevelopment.un.org/sdg4#targets>. (Accessed 8 November 2018).
- UNESCO (United Nations Educational, Scientific and Cultural Organization), 2012. UNESCO Guidelines for the Recognition, Validation and Accreditation of the Outcomes of Non-Formal and Informal Learning. UNESCO and UIL, Paris.
- UNESCO (United Nations Educational, Scientific and Cultural Organization), 2015. Recommendation on Adult Learning and Education (2015). UNESCO/UIL, Paris/Hamburg. <http://unesdoc.unesco.org/images/0024/002451/245179e.pdf>. (Accessed 25 November 2017).
- UNESCO (United Nations Educational, Scientific and Cultural Organization), 2016. Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4. Towards Inclusive and Equitable Quality Education and Lifelong Learning Opportunities for All. Education 2030. UNESCO, Paris. <https://unesdoc.unesco.org/ark:/48223/pf000-02456-56>. (Accessed 25 February 2019).
- Walia, H., 2010. Transient servitude: migrant labour in Canada and the apartheid of citizenship. *Race Class* 52 (1), 71–84. <https://doi.org/10.1177/0306396810371766>.
- Webb, S., 2015. The feminisation of migration and the migrants VET policy neglects, the case of skilled secondary women migrants in Australia. Special Issue – Gender in Vocational Education and Training: Pathways, Process and Practice *J. Vocat. Educ. Train.* 67 (1), 26–46. <https://doi.org/10.1080/13636820.2014.922117>.
- Webb, S., Faine, M., Pardy, J., Roy, R., 2017. The role of VET in the (dis)placing of migrants' skills in Australia. *J. Vocat. Educ. Train.* 69 (3), 351–370. <https://doi.org/10.1080/13636820.2016.127839>.

Global Education Monitoring Report

Robyn Read^a and Aaron Benavot^b, ^a Independent Researcher, Toronto, ON, Canada; and ^b School of Education, University at Albany-SUNY, Albany, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	553
The historical roots of global reports monitoring education	553
Enter the Education for All Global Monitoring Report	555
Shifting policy regimes: the emergence of Global Education Monitoring Report	557
Critiques and evaluations of the GEM Report series	561
Looking ahead at monitoring challenges and pandemic recovery	562
References	563

Introduction

Historically, international organizations have sought to influence global policy through thematic reports, many organized in a series.¹ These signature publications help justify shifts in international policy priorities, serve as mechanisms to monitor progress toward specific development goals, and seek to shape the direction of national education policies through an evidence-based analysis of priority areas. Many of these annual “flagship” publications include a focused literature review as well as an in-depth look at a special theme—all supported by empirical evidence displayed through extensive tables and figures both in the text and the appendices (Cussó and D’Amico, 2005; Spaulding and Chaudrhuri, 1999).

Global actors have increasingly adopted this model and staked their claim to global thought leadership through the launching of annual or periodic reports. To be sure, influence over policy or practice depends in no small measure on the report’s accessibility—long, highly technical, and research-focused publications curtail readership (Doemeland and Trevino, 2014). And yet, even technical reports can drive political discourse—consider recent controversies involving figures in the World Bank’s *Doing Business Report* (Nicola, 2021).

This chapter focuses on the *Global Education Monitoring Report* (or GEM Report)—formerly the *Education for All Global Monitoring Report* (GMR)—a special report series created and supported by the international education community as a monitoring and evaluation tool to track global progress toward the six Education For All (EFA) goals following the World Education Forum in Dakar in 2000. With time, the GEM Report² has become one of the most influential evidence-based publications in the field of international education development. While certainly not without criticism, the GEM Report series has been referred to as “required reading for anyone interested in education for development, as well as for educational policy-makers in developing countries and for donor agencies” (Brown, 2006, p. 480). Starting in 2002, this annual publication has informed national, regional and global policy makers on a diverse array of education issues and influenced international approaches to educational aid. Drawing on extensive quantitative and qualitative evidence, the Reports have tracked country progress (or lack thereof), highlighted critical challenges and hurdles, and become part of the international education architecture.

This chapter explores the development of the GEM Report series and its use as a reference for policymaking and monitoring progress toward international goals among aid agencies, non-governmental organizations, and developing countries (Daniel, 2010; Packer, 2008). While a key condition for the Report series’ success is its editorial independence from UNESCO, its host institution, the emergence and evolution of these reports are inextricably linked to UNESCO’s role and impact as a knowledge actor in the field of international education development. Therefore, first, we situate the GEM Report series within the historical evolution of UNESCO and the EFA movement. Next, we provide insight into the nature, structure, financing, and production of the report series. We then discuss critiques and evaluations of the reports. Lastly, given the rapidly changing contexts of global education due to the impact of the global COVID-19 pandemic, we consider the future role of the Report series in an increasingly uncertain and disruptive policy landscape.

The historical roots of global reports monitoring education

After two brutal World Wars, the United Nations Educational, Scientific and Cultural Organization (UNESCO) was formed to unite national governments and other international actors under a mandate of peacebuilding through international cooperation including, “full and equal opportunities for education for all” (UNESCO, 1946). Ratified in 1946, UNESCO directed the world’s

¹For example, the Food and Agricultural Organization’s, *The State of Food and Agriculture*, published on semi-regular basis since 1947; the World Bank’s *World Development Report* (1978–present); UNICEF’s *State of the World’s Children Report* (1980–present); the United Nation’s (UN) Development Program’s *Human Development Report* (1990–present); or the UN’s Educational, Scientific and Cultural Organization’s (UNESCO) *World Survey of Education* (1955–1971), *International Yearbook of Education* (1946–1969 ... 1979–1994), and the *World Education Report* (1991, 1993, 1995, and 1998).

²The term “GEM Report” refers to both the earlier series of EFA Global Monitoring Reports as well as the rebranded GEM Reports.

commitment to a global goal of “education for all.” Through its policies and international activities, UNESCO launched an “attack upon ignorance by helping all member states who desire such help to establish a minimum Fundamental Education³ for all their citizens” (Memorandum on the education program of UNESCO, Paper No. 1 prepared by the Education Staff of the Preparatory Commission, 1946 as quoted in Boel, 2016, p. 154). Since its founding, UNESCO has framed educational development as a technical and policy matter, centering scientific research and technical advice as the crucial ingredient to improve schooling around the world (Huxley, 1946). Given this underlying theory of change, UNESCO positioned itself as an innovator, standard-setter, and capacity builder in the education sector, leading the field in both the production and dissemination of global education knowledge (Menashy and Manion, 2016).

UNESCO’s early outreach programs focused on providing technical assistance in education. For example, it sent field officers to newly independent countries to help establish and expand education systems and literacy programs (Duedahl, 2016). Between 1946 and 1958 UNESCO was involved in activities related to fundamental education in at least 62 countries (Boel, 2016, p. 168). These included extensive fieldwork in the Democratic Republic of Congo in support of its emerging education system (Dikay, 2016), and a mass literacy campaign in Indonesia helping more than 10 million people gain literacy skills over the 12-year period (Boel, 2016). In the 1950s, UNESCO mobilized efforts to make good on country commitments to universal education by convening regional conferences on compulsory school legislation (e.g., in Bombay 1952, Cairo 1955, Lima 1956), and universal access to education (e.g., in Karachi 1960, Addis Ababa 1961, Santiago 1962, Tripoli 1966). These early efforts contributed to countries adopting compulsory attendance laws throughout the world. Today nearly all countries have established such laws and more than 75% of all countries require 9 years or more of compulsory education (UNESCO, 2020).

In addition to technical assistance, legislation and policies to help universalize access to basic education, UNESCO focused on the production and dissemination of educational statistics and research (Huxley, 1946). While UNESCO had emerged as a global leader in educational development by 1960, its position as a thought leader was contested. Some UNESCO-supported education initiatives were met with resistance (Boel, 2016). The use of research, data, and evidence to inform educational policymaking, far from politically neutral acts, raised concerns. Documented tensions between UNESCO’s political and technical roles in the international arena impacted both UNESCO’s funding and its ability to see its mission through to fruition (Burnett, 2011; Jones, 1988; Menashy and Manion, 2016; Mundy, 1999, 2006; Mundy and Manion, 2015). By the 1980s, many in the international community began to question UNESCO’s ability to provide rigorous, reliable technical advice; concurrently, several multilateral organizations, including the World Bank, UNICEF, and the OECD, began developing their own pool of education experts to inform education policy and programming (Burnett, 2010; Cussó and D’Amico, 2005).

And yet, despite an expanding technical evidence base, international donors in education did not consolidate around a common policy agenda. During the latter quarter of the 20th century, the assessment of education priorities changed rapidly and diverged widely, resulting in a general lack of consensus as to “what works” in development education (Mundy, 2006). In addition, the divisive politics of the cold war limited opportunities to align bilateral support to newly established education systems. Donor governments often used education as a tool to support their geopolitical agendas and encouraged neoliberal education policies in the Global South (Dale, 2005; Magrath, 2014; Robertson and Dale, 2006). As a result, instead of a globally united movement toward EFA, educational development projects were often short-term, small to moderate sized, bilateral transactions, many of which worked at cross-purposes (Mundy, 2006).

Fragmented efforts in support of universal education shifted in the 1980s as a global consensus around education began to emerge. Under the leadership of James Grant (then UNICEF Executive Director), the World Bank, UNDP, UNESCO and UNICEF formally established and funded an Inter-Agency Commission (IAC) on EFA, which together with an Executive Committee and a small Secretariat housed at UNICEF headquarters in New York, was charged with coordinating a global meeting of international actors working toward EFA (Little and Miller, 2000). The resulting World Conference on Education for All, held in Jomtien, Thailand in May 1990, marked a major breakthrough for the EFA movement—in tone, organization and intended impact.

The Jomtien meeting brought together representatives from 155 UN member states, delegates from 33 multilateral institutions, and representatives of 125 non-governmental organizations (NGOs), who contributed to a framework for action around the overarching goal of EFA (Packer, 2008). Multilateral and bilateral agencies connected with networks of NGOs to revitalize efforts to build global consensus around educational priorities. Following the meeting, the IAC Secretariat was moved to a new home at UNESCO headquarters with funding from UNESCO, UNICEF, UNDP, and the World Bank. Instead of focusing on action, the IAC was asked to play an advisory role to the international community—provide general guidance on successful approaches and promising new strategies, report on actual progress and discuss issues of widespread concern (Little and Miller, 2000).

³The notion of fundamental education, which was understood to be “at the heart of the work of UNESCO” in the early 1950s, was broad in meaning and scope and transformative in intent. The first of 12 *Monographs on Fundamental Education*, published in 1949, stated that “the aim of all education is to help men and women to live fuller and happier lives.” The idea of fundamental education was to enable communities to change living conditions through education in a wide range of areas, such as health education, domestic and vocational skills, knowledge and understanding of the human environment—including economic and social organization, law and government—and what was called “the development of qualities to fit men to live in the modern world, such as personal judgment and initiative, freedom from fear and superstition, sympathy and understanding for different points of view.” <https://en.unesco.org/news/fundamental-education-pioneer-concept-jens-boel-explains-why>.

The placement of the IAC within UNESCO could be interpreted as a sign of the global community's faith in UNESCO as a major producer and disseminator of educational data and statistics. But in fact, UNESCO came under considerable criticism in the 1990s by UNICEF and the World Bank regarding the quality, reliability, and relevance of its educational statistics (Cussó and D'Amico, 2005). In the wake of budget cuts, restructuring and belt tightening, the Division of Statistics at UNESCO headquarters was severely weakened. An external panel of consultants recommended that the Division needed greater autonomy and funding, and in 1999 UNESCO decided to relocate the division to Montreal and rename it the UNESCO Institute for Statistics to address these concerns (Cusso, 2006). Actual progress toward EFA during this decade was limited and uneven. Structural adjustment policies by the IMF undercut national allocations to education and needed funding from the donor community did not materialize (Delamónica et al., 2001). Moreover, collaborations among multilateral agencies (i.e., UNESCO, UNICEF, UNDP and the World Bank), national governments and NGOs were the exception, not the rule (Little and Miller, 2000). Thus, while Jomtien may have united the global community around a common EFA Framework, in practice international aid and assistance to national education sectors in the Global South remained splintered. Implementation of national EFA plans was uneven.

The establishment of a global EFA discourse and an emerging architecture for international aid to education did have some salutary benefits. More governmental and non-governmental actors embraced international development targets and EFA strategies to achieve them (Mundy, 2006). Multilateral organizations (e.g., the World Bank, UNDP, UNICEF, and the OECD) individually and collectively continued to raise the overall profile of EFA and keep it on the global development agenda (Edwards et al., 2018). Toward the end of the 1990s the IAC brought together more than 180 countries for a series of regional and sub-regional meetings to assess global progress toward EFA. The results of this assessment were reported in both a global synthesis document (Skillbeck, 2000), as well as a statistical report (UIS, 2000), and served to re-emphasize the significance of, and broad support for, the EFA agenda.

In 2000, more than 1100 members of the global education community (including representatives from donor agencies, multilateral organizations, NGOs, civil society groups, and 164 national governments) once again came together at the World Education Forum (WEF) in Dakar, Senegal to reconfirm their mutual commitments to EFA (Benavot et al., 2016). This time the international education community was determined to harness existing momentum and ensure that the WEF resulted in actionable targets and clear agendas to guide efforts going forward (Edwards et al., 2018). A key outcome of the Dakar meeting was the co-creation of clearly formulated, time-bound and measurable global education goals. This became the basis of *The Dakar Framework for Action*, officially adopted at the end of the Forum. The six Dakar goals consisted of: (1) Expanding early childhood care and education; (2) Providing free and compulsory good quality primary education for all; (3) Promoting the acquisition of life-skills by adolescents and youth; (4) Expanding adult literacy; (5) Eliminating gender disparities in education; and (6) Enhancing educational quality. In addition to goal setting, the Dakar Framework established mechanisms to move the EFA agenda forward and hold global stakeholders (i.e., international agencies, governments, donor partners, civil society, and the private sector) accountable for meeting the six EFA goals.

While many assumed that the World Bank would be charged with responsibility over the new EFA agenda, the participants of the Dakar Forum selected UNESCO for the key coordination role, including the creation of a new EFA High-Level Group (HLG) composed of Ministers, UN and bilateral agencies, and civil society representatives to be housed at UNESCO (Edwards et al., 2018). The Dakar Framework mandated that the HLG was to be "... informed by a monitoring report from the UNESCO International Institute for Educational Planning (IIEP), the UNESCO International Bureau of Education (IBE), the UNESCO Institute for Education (UIE) and, in particular, the UNESCO Institute for Statistics (UIS), and inputs from Regional and Sub-regional EFA Forums" (UNESCO, 2000, p. 10).

Enter the Education for All Global Monitoring Report

Monitoring EFA progress and holding countries and donors accountable to their EFA commitments were central concerns during the first annual meeting of the HLG, held in Paris in October of 2001. The Dakar Follow-up Unit at UNESCO was asked to prepare a short report to facilitate discussion among HLG members. This initial EFA monitoring report, entitled *Monitoring Report on Education for All: 2001*, provided: (1) an overview of the fundamental operating assumptions underlying the strategies outlined in the Dakar Framework and the structures that were agreed upon in order to implement the Dakar Framework; (2) an overview of global progress toward the EFA goals supported by data from UNESCO; (3) the current status of planning for EFA at both the national and international level, including an overview of international cooperation and resource mobilization toward EFA and the role of non-state actors; and (4) suggestions for moving the EFA agenda forward.

The initial report was not well received by the HLG (Edwards et al., 2018). Many characterized it as lacking adequate depth and analysis. Many HLG members viewed this EFA report—in addition to an earlier UNESCO monitoring report submitted during the Mid-Decade Meeting on EFA in Amman Jordan (UNESCO, 1996)—as falling short of the kind of authoritative monitoring report that honestly assesses EFA progress and provides direction to countries and the international community of needed actions to meet their EFA commitments (Universalia, 2006). The HLG decided to mandate a new report to be submitted at future meetings: "An authoritative, analytical, annual EFA Monitoring Report should be produced drawing upon national data—quantitative and qualitative—and assessing the extent to which both countries and the international community are meeting their Dakar commitments" (UNESCO, 2001, p. 29). To improve the report's quality and relevance, the HLG stipulated that UNESCO should urgently convene

key partners to discuss how the report should best be prepared, managed, and resourced, and that UIS should be given an important role in the development of the report (UNESCO, 2001).

Three months later, UNESCO convened a planning meeting of international stakeholders and education experts. Participants agreed that the new EFA Global Monitoring Report must be owned by, and report to, the EFA community (Universalia, 2006). Members of the planning meeting decided that while UNESCO would host the Report series, by providing office space and publishing services, the new Report team would be editorially and financially independent from UNESCO, guided by an international Advisory Editorial Board (Packer, 2008; Universalia, 2006).

In cooperation with the head of the EFA Observatory at the UNESCO Institute of Statistics, the Dakar Follow-Up Unit quickly pulled together a small team of 11 education experts to work under the guidance of a high-profile director, Professor Christopher Colclough, who had demonstrated experience bridging research and policy in the field of development education (Edwards et al., 2018). There was no formal governance structure for these reports. The Report Director was given the ultimate responsibility of conducting and coordinating research for each report in addition to drafting its content.

The Director worked with its newly established Editorial Board to formulate a vision statement for the Report series:

“An authoritative annual report, which holds the global community to account for commitments made at the World Education Forum. It will chart progress against the six Dakar goals and targets, including the Millennium Development Goals (MDGs) for education, highlight effective policies and strategies, alert the global community to emerging challenges and promote international action and co-operation. An indispensable advocacy and technical tool for everyone involved in promoting Education for All.”—Internationally Agreed Vision Statement, Global Monitoring Report on EFA, January 2002.

Universalia (2006, p. 5).

Members of the Editorial Board also underscored the critical importance of the Report’s editorial independence. A “key outcome” of the Board’s second meeting (July 2002) made this point crystal clear:

Independence: the independence of the Report is crucial, and it will therefore not be a report “from UNESCO,” but rather “to UNESCO.”

Under severe time constraints, the first edition of the Education for All Global Monitoring Report (GMR) was published in late 2002, defining indicators and measures, and analyzing global progress toward each of 6 EFA goals.

Whatever vision the HLG may have had when it established the Report series, the Reports took on a dynamic life of their own. Between 2002 and 2015 twelve Reports were published (see Table 1). Noted for their crisp and accessible writing style, careful

Table 1 The Education for All Global Monitoring Reports (2002–2015) and Global Education Monitoring Reports (2016–present)^a.

Report year	Report theme	Report director	Pages
Education for All Global Monitoring Report			
2002	Education for all: is the world on track?	Christopher Colclough	310
2003–2004	Gender and education for all: the leap to equality	Christopher Colclough	372
2005	Education for all: the quality imperative	Christopher Colclough	428
2006	Education for all: literacy for life	Nicholas Burnett	448
2007	Strong foundations: early childhood care and education	Nicholas Burnett	405
2008	Education for all by 2015: will we make it?	Nicholas Burnett	448
2009	Overcoming inequality: why governance matters	Kevin Watkins	477
2010	Reaching the marginalized	Kevin Watkins	525
2011	The hidden crisis: armed conflict and education	Kevin Watkins	431
2012	Youth and skills: putting education to work	Pauline Rose	480
2013–2014	Teaching and learning: achieving equality for all	Pauline Rose	496
2015	Education for all 2000–2015: achievements and challenges	Aaron Benavot	517
Global Education Monitoring Report			
2016	Education for people and planet: creating sustainable futures for all	Aaron Benavot	620
2017–2018	Accountability in education: meeting our commitments	Aaron Benavot/Manos Antoninis	508
2019	Migration, displacement and education: building bridges, not walls	Manos Antoninis	438
2020	Inclusion and education: all means all	Manos Antoninis	424
2021–2022	Non-state actors in education: who chooses? Who loses?	Manos Antoninis	474

^aIn 2001, in the wake of the Dakar EFA conference UNESCO asked an internal team led by Abhimanyu Singh to prepare a preliminary monitoring report (called the Monitoring Report on Education for All). This 51-page report can be considered as a forerunner of the EFA Global Monitoring Report series, even though it was not editorially independent of UNESCO.

argumentation, peer-reviewed analysis and extensive presentation of diverse evidence, early editions set a high standard for an informative and analytical global monitoring report.

Altogether, between 2002 and 2022 seventeen GEM Reports were produced. The reports drew on current statistics and other data sources to monitor progress in relation to international targets. They highlighted international trends in education, identified effective policies and reforms, assessed political commitments, discussed challenges and constraints, and promoted specific international strategies and cooperation. Fourteen of the 17 reports explored a crucial theme, understood to foster—or hinder—global educational progress (Table 1). These themes included gender quality, teaching and learning, adult literacy, inequality and marginalization, armed conflict, accountability, migration and displacement and non-state actors.

It was the GEM Report Advisory Board (initially called the Editorial Board), and not the Director, that determined the theme of each Report. The Advisory Board, which increased in size over the years, included funders of the report as well as representatives of multilateral and bilateral agencies, NGOs, civil society organizations, academic experts, and directors of UNESCO institutes (UNESCO, n.d.). In 2022 the Advisory Board included three dozen members and a dozen observers. While each Advisory Board provided advice on different report-related matters, its main statutory responsibility is singular: to select the theme of a future report. Once the theme is confirmed, the Report team immediately begins working on a comprehensive concept note and a new report cycle is initiated.

On average, it takes 18 months to produce a report—from concept note to finished publication (Universalia, 2010). Since 2002 the size of the team involved in the development, production, and dissemination of the Report has doubled—from about a dozen members to twice that number today. This increase mainly reflects the elaboration and diversification of website content and the preparation of several specialized publications—namely, a youth report, a gender report, focused language summaries, and, recently, a regional report. The main report relies on dozens of external experts who are commissioned to prepare thematic or monitoring related background papers, selections of which are integrated into the draft text. While the team follows UNESCO's administrative procedures for budgeting, contracts processing, and staffing (Edwards et al., 2018), it maintains its editorial autonomy and follows a flexible managerial and operational style to meet shifting and multiple production schedules (Carpenter et al., 2014).

It could be argued that many UNESCO colleagues were initially ambivalent toward the relatively well-resourced and editorially independent Report series, but this has given way to a more positive attitude toward the content and messages of each Report. Most UNESCO field and regional offices organize annual launch events, to which key government officials and other education stakeholders are invited. These events, at which key Report findings and policy messages are presented and discussed, help strengthen ties between UNESCO officials and various education stakeholders in the country or region.

In terms of UNESCO's formal role toward the GEM Report, little has changed. UNESCO provides a small annual financial contribution to the report (about \$50K USD). A more significant contribution is in-kind: administrative support at headquarters; data and technical support from UIS; some communication support involving press releases and media tracking; report distribution to UNESCO offices; and on-going engagement with UNESCO regional, cluster and national offices to conduct events and distribute reports. National Commissions to UNESCO are involved in the dissemination of the report in their respective countries and communicate key messages to national government officials and other stakeholders (Universalia, 2010).

Since 2002 bilateral agencies and international organizations have collectively provided funds for the research, drafting, production and distribution activities undertaken by the Report team. While this pooled funding model has helped ensure the editorial independence of the Report, it does involve a measure of unpredictability, undermining attempts at long-term planning. For example, report funding varied from \$5,360,000 in 2006, \$7,166,000 in 2011, and \$3,395,000 in 2013 (values adjusted to 2013 US dollars) (Carpenter et al., 2014; Universalia, 2010, p. 36). Typically, two to four major donors cover half of the total budget; the remainder is funded by a shifting collection of smaller contributors (Ipsos MORI, 2018).

Over time, the number of funders of the Report series has increased about threefold: from an average of 5–6 funders in the initial years to 16–19 funders in recent years (See Table 2). Six bilateral agencies have been regular contributors to the GEM Report over the entire 20-year period: France, Germany, Ireland, Norway, Sweden and Switzerland. The Netherlands and the UK have also been significant, albeit inconsistent, donors. Many other agencies and foundations have been irregular or one-time contributors. Agreements between a donor and the report team typically cover only 1 or 2 years, which means that the team must constantly engage in fundraising efforts, which ultimately detract from long-term planning. Despite the ebb and flow of funding over the years, the GEM Report team has managed to consistently produce high-quality, insightful reports, which are read or reviewed by diverse audiences: ministers and ministerial officials, international aid agencies, non-governmental organizations, as well as researchers and students (Packer, 2008; Daniel, 2010; Ipsos MORI, 2018).

Shifting policy regimes: the emergence of Global Education Monitoring Report

The evidence suggests that the Report series has evolved from a straightforward EFA monitoring tool to an integral pillar in the international education architecture (Burnett, 2019). The Report positioned itself as part of a broad epistemic community of actors involved in international development and education. In doing so, it sought to reflect the views and perspectives of a wide-ranging and oftentimes contentious community of stakeholders.

Table 2 Global Education Monitoring Report funders 2002–2020.

Funders	Education for All Global Monitoring Report (GMR)												Global Education Monitoring Report (GEM Report)				
	2002	2003/04	2005	2006	2007	2008	2009	2010	2011	2012	2013/14	2015	2016	2017/18	2019	2020	2021/22
Australia									✓	✓			✓	✓	✓	✓	✓
Bill and Melinda Gates Foundation														✓			✓
Canada					✓								✓	✓	✓	✓	✓
Denmark				✓		✓	✓	✓	✓	✓							
Education Above All																✓	✓
Education International											✓						
European Union																	✓
Finland										✓	✓			✓	✓		
France		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Germany		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hewlett Foundation													✓	✓	✓	✓	✓
Ireland		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Israel				✓													
Italy																✓	✓
Malala Fund														✓		✓	✓
MasterCard Foundation									✓			✓		✓	✓	✓	✓
Open Society Foundations											✓		✓	✓	✓	✓	✓
Netherlands	✓	✓		✓		✓		✓	✓	✓	✓	✓		✓	✓	✓	✓
Norway		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Results for Development Institute													✓				
Sweden	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Switzerland				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
United Kingdom	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
GCE-US												✓					
UNGEI														✓	✓	✓	✓
UNESCO	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
World Food Program																	✓
Total number of funders	4	7	5	11	9	10	9	9	11	11	12	13	12	16	15	18	17

Sources: 2016 Report to funders, GEM Report Advisory Board Meeting; and then GEM Reports 2017–2021/22.

In practice this meant commissioning background papers from experts and organizations from around the world. These externally commissioned papers, together with in-house literature reviews, tapped into a vast pool of existing academic and gray literature that examines national and regional progress, especially in particular areas like gender, quality, inequality, learning, governance and sustainability and informs the global education agenda. The team's engagement with experts and organizations created a knowledge network comprised of diverse global education stakeholders, which can be examined through the application of bibliometric analysis to the references included in the published Reports.

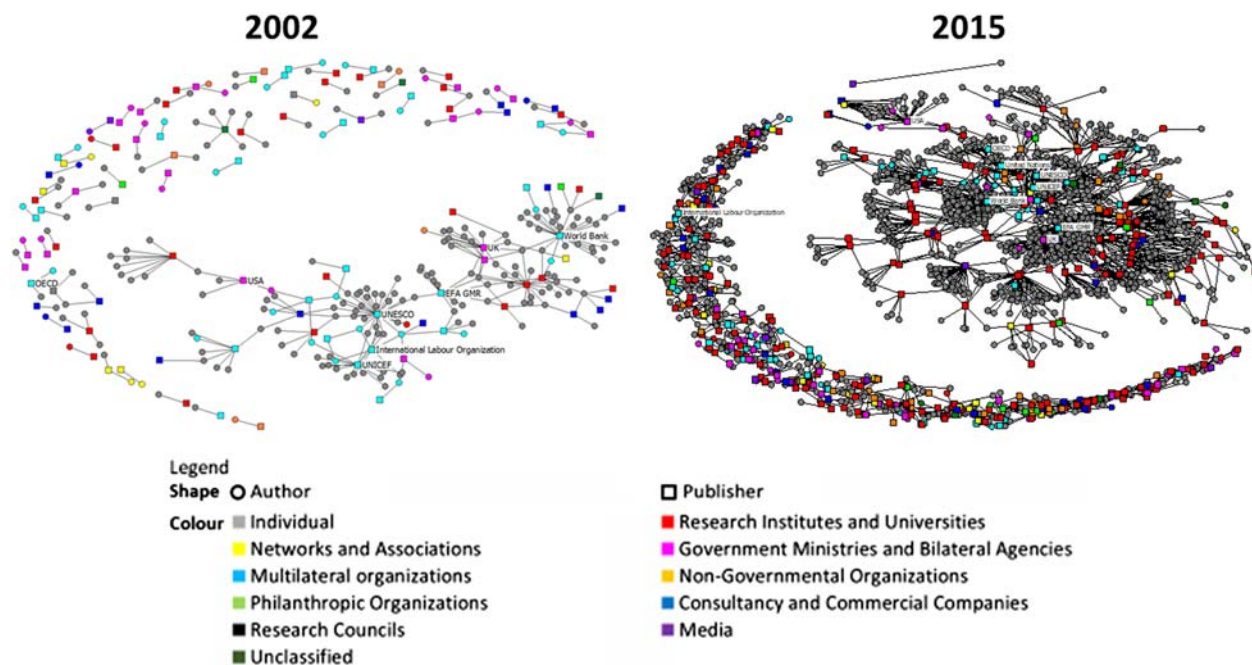
References in the first 12 Reports illustrate the sheer amount of knowledge produced by key actors in the international education policy field. Analysis of the 12 reports published between 2002 and 2015 found the Report team had cited a combined total of 8271 references, incorporating the perspectives of 6833 authors, including 6239 (91%) individuals and 595 (9%) organizations (Read, 2017). Across the combined bibliographies of the 12 Reports, eight publishers were referenced 100 or more times—accounting for 39% of all references during that period (Table 3). By comparison, 248 publishers (15.1%) were referenced only once. But even among this elite group of eight publishers, three key organizations were disproportionately cited, including the World Bank (listed as publisher for 825 references), UNESCO (709 references), and the GEM Report itself (listed as publisher of 653 references, mostly background papers commissioned for each Report), with the OECD coming in as a distant fourth (298 references) (Read, 2017).

However, while this analysis shows that the Reports have been influenced by a handful of key thought leaders in the field of global education, it also shows that the Reports have, in some sense, lifted all boats in the field. A comparison between the citation networks of the 2002 and 2015 reports (Fig. 1) clearly reveals the growth of the report's international knowledge network, with more actors and more types of actors consistently being cited in the reports during this period. The actors included multilateral and bilateral organizations, leading NGOs, think-tanks, philanthropic foundations, private companies, research teams and individual experts. By referencing wide-ranging research studies, data sources and voices, the GEM Reports helped increase the visibility and legitimacy of thousands of individual experts and smaller civil society and non-governmental

Table 3 Top publishers referenced 100 or more times in EFA Global Monitoring Reports published between 2002 and 2015.

<i>Publishing organization</i>	<i># Of times referenced</i>
World Bank	825
UNESCO	709
EFA GMR	653
OECD	298
UNICEF	208
United Nations	183
UK	171
USA	139

Source: Read (2017).

**Fig. 1** Education for all Global Monitoring Report citation networks 2002 and 2015.

organizations in the field. To the extent that the Reports created spaces for smaller actors to engage with thought leaders in open, independent, and informed debates, this helped solidify the Reports' status as an essential source on the state of global education (Edwards et al., 2018).

By the time the international education community came together at the World Education Forum in Incheon, South Korea to adopt a post-2015 education agenda, the Report series had become a common reference point for multiple stakeholders. The Report team had earned a reputation for being an independent, reliable, and authoritative source for accurate data and relevant evidence on EFA progress. At the Incheon meeting, it was proposed that the EFA Global Monitoring Report be rebranded as the Global Education Monitoring Report and given a formal mandate to monitor the emergent global goal on education (SDG 4) and its 10 targets. Participating delegates—namely, education ministers, heads and members of delegations, heads of agencies, officials of multilateral and bilateral organizations and representatives of civil society, the teaching profession, youth, and the private sector—had few caveats in adopting paragraph 18 of the Incheon Declaration, which states:

the Education for All Global Monitoring Report be continued as an independent Global Education Monitoring Report (GEM Report), hosted and published by UNESCO, as the mechanism for monitoring and reporting on the proposed SDG 4 and on education in the other proposed SDGs, within the mechanism to be established to monitor and review the implementation of the proposed SDGs.

UNESCO et al. (2015, p. vi).

UNESCO, in consultation with the Report Director, supported the rebranding of the Global Education Monitoring Report and embraced its new focus on monitoring all education-focused SDG targets, which were officially adopted by the UN General Assembly in September 2015. The institutional standing of the GEM Report was further acknowledged in the *Education 2030 Framework for Action* endorsed by ministries of education at UNESCO's General Conference in November 2015 (Jamil, 2016). The Education 2030 Framework document stated:

The EFA Global Monitoring Report will be continued in the form of the Global Education Monitoring (GEM) Report. It will be prepared by an independent team and hosted and published by UNESCO. The Director of the team is appointed by the Director-General of UNESCO. Attention will be paid to geographical balance in its Advisory Board. The GEM Report will be the mechanism for monitoring and reporting on SDG 4 and on education in the other SDGs, with due regard to the global mechanism to be established to monitor and review the implementation of the 2030 Agenda for Sustainable Development. It will also report on the implementation of national and Education 2030 Framework for Action international strategies to help hold all relevant partners to account for their commitments as part of the overall SDG follow-up and review.

UNESCO et al. (2015, pp. 28–29).⁴

The position of the GEM Report series in the post-2015 global education landscape was now secure. National policymakers, donors, international agencies and NGOs had voted their preference for an independent monitoring and accountability mechanism, endorsing the Report's model of financial and editorial independence. Under its new SDG mandate, the monitoring agenda of the GEM Report significantly expanded, with a more extensive and complex set of education targets and indicators to review. And yet, despite its expanded SDG mandate, the GEM Report faced an uncertain future, since many donors were either reluctant or slow in committing funds to the GEM Report series. As a result, the team was forced to allocate considerable time and effort to secure additional donors and increased funding from existing donors, even to maintain previous funding levels (Ipsos MORI, 2018). Between 2015 and 2017 the GEM Report received annual funding contributions from donors that ranged from \$3,000,000 to \$6,000,000 US dollars (Ipsos MORI, 2018, p. 10).

In these years the GEM Report adapted its monitoring toolkit to the broader and more ambitious SDG framework. It laid the groundwork of an expanded monitoring strategy in the pivotal 2016 GEM Report, carefully reviewing proposals for global indicators of SDG4 targets, as well as the validity, reliability and relevance of existing and new data sources.

The output of the GEM Report also evolved. In addition to the main report, which had been regularly published in almost all UN languages, the Executive Summary was translated each year into 10–15 additional language versions. Drawing on recommendations of the external evaluations, the GEM Report decided to diversify its knowledge products—for example, gender reports, youth reports, policy papers, infographics and, since 2019, thematic regional reports—to enhance the dissemination and influence of each Report's conclusions and policy recommendations. The team also expanded its online monitoring resources: PEER (Profiles Enhancing Education Reviews) compiles information on education laws and policies to help countries compare and learn from others; WIDE (The World Inequality Database on Education) visualizes data showing inequalities by groups, within and between countries; SCOPE (Scoping Progress in Education) provides visualizations explaining the data complexities behind SDG 4; and VIEW (Visualizing Indicators of Education for the World) compiles multiple data sources to look at completion rates over time by country and region.

These publications and resources, available on the Report website, were highlighted in more than 100 launch events held each year in locations around the world. It is estimated that 2 million copies of the report series have been downloaded from the website since the team began tracking website traffic. Launch events attracted relatively large audiences and typically included senior government officials, ministers of education, researchers, teachers, youth and representatives from civil society and the private sector (GEM Report, 2016). The report communications team collaborated strategically with many partners to extend messaging—for example, UNESCO field offices, other UN agencies (UNICEF, UNHCR, ILO), the Global Partnership for Education, the Global Campaign for Education, Education International, Malala Fund, Pratham and Education Cannot Wait (Ipsos MORI, 2018).

The team's communication strategy evolved to reach broader audiences. It included a newsletter, social media campaigns, photo competitions, and interviews and collaborative pieces with news media. The team (with some support from UNESCO) tracked media references of the Report—for example, in the Guardian, Al Jazeera, The Times Higher Education Supplement, BBC, The Dawn, Le Monde, RFI, and The Times of India among others. In 2013 report media coverage reached over 31 million people in 100 different countries (Redman, 2020).

In March 2010 the GEM report established a successful presence in the education blogosphere—the World Education Blog (<https://world-education-blog.org/>). Over the next 11 years (until the summer of 2021 when it moved to a new location), this blog-site published 1060 blogs that attracted 2.5 million views and 1.5 million visitors. The editorial independence of the World Education Blog allowed it to publish diverse individual and organizational viewpoints on global education issues, as well as pieces that amplified findings, interpretations and policy recommendations found in the main GEM Report or in related publications (e.g., policy papers, Gender Review, Youth Report).

⁴It is important to note that role of UNESCO's Institute for Statistics in preparing internationally comparable data for multiple end users, including the GEM Report, was also strengthened at the two 2015 meetings. The autonomy and evolution of UIS clearly deserve closer scrutiny following the adoption of the Education 2030 Framework for Action.

These communication and dissemination efforts raised awareness of the status of EFA or SDG progress as well as the nature and scale of the challenges facing national, regional and international actors seeking to advance global education agendas.

Critiques and evaluations of the GEM Report series

Prior to publication each GEM Report is critically reviewed by a group of academics and subject experts. After publication many GEM Reports have been reviewed and critiqued in select academic journals. Most scholarly critiques revolve around specific indicators or arguments advanced in a particular report. For example, [Alexander \(2015\)](#) notes that the EFA indicators measuring quality education cover a narrow range of variables, thus raising questions about the fit of the EFA monitoring process to encourage countries to transform approaches to teaching and learning. Alexander criticizes the top-down application of research for education monitoring and argues instead for a more radical, creative and inclusive approach that defines teaching and learning indicators from the bottom up. [Lewin \(2011\)](#) also takes issue with several metrics found in the Reports. Specifically, he challenges the Education for All Development Index (EDI) used in the earlier Report series, noting that aggregated composite indexes like the EDI are difficult to interpret and of limited use in measuring educational progress.

[Gordon \(2010\)](#) notes that while some aspects of the 2005 Report's *Framework for Understanding Education Quality* (see [UNESCO, 2005](#), p. 36) are helpful, they fail to problematize the broader historical and socio-economic contexts in which educational processes are situated. Gordon argues that the lack of an explicit value basis in the Report Framework limits the articulation of deeper critiques that go to the heart of the provision of good quality education, particularly in fragile and conflicted affected countries. In a more recent critique, [Magee \(2020\)](#) argues that the 2019 GEM Report on *Migration, Displacement and Education* does not go far enough in holding the global community to account for their commitments to refugees and the displaced. "If publications like the GEM report are going to play a successful role in accountability, much greater emphasis needs to be placed on monitoring international commitments, working toward a more just distribution of global wealth and education's role in achieving it" ([Magee, 2020](#)).

In addition to these scholarly critiques, there have been four external evaluations of the Reports—in 2006, 2010, 2014, and 2018—conducted by expert teams through an open tender process. These evaluations assessed the overall quality of the reports; the efficiency of the production and distribution processes; the effectiveness of the team's communication, media, and other outreach work; and, notably, the impact of the reports in terms of global reach as well as uptake in policy and practice. Overall, the evaluations found that the reports conform to their core mandate and are based on well-grounded research and analysis presented in engaging ways.

A recurrent critical theme in several evaluations is that the Report's global nature limits its policy relevance to specific national or local contexts. This lack of granularity tends to undermine the Report's influence and impact. While Directors have no intention to revise the Report's core *global* mandate, they have taken steps to address this concern—for example, assigning members of the drafting team to follow policy developments in specific regions; preparing succinct regional overviews of EFA progress; commissioning background papers that focus on a particular case or region; and using boxes to illustrate a broader point by focusing on a particular case. The decision to develop a full-blown regional report is the most recent indication of a shift in this direction.

All evaluations make clear that the editorial independence of the Report series has enhanced its perceived credibility as well as its rigor, insight and quality. As one Advisory Board Member noted: "The one quality that GEM Report has is that it is seen as a credible and trusted voice, and it can only be seen as trusted if it is seen as being independent of influence, and that must include UN influence" ([Ipsos MORI, 2018](#), p. 19). Although UNESCO hosts and publishes the Report as its "flagship" publication in education, UNESCO leadership has no say over the Report's content, which is determined by the Director. The 2018 evaluation supported this point: in a survey of stakeholders, more than 80% of respondents from UNESCO and agencies dependent on UNESCO agreed that the Report is an independent source of information ([Ipsos MORI, 2018](#), p. 19). And yet, this independence from UNESCO comes at a cost, since the Report's continued viability as a global public good depends on the generosity of external donors.

Between 2015 and 2020 the GEM Report received annual funding contributions from donors that ranged from just under \$4 million to almost 6.5 million US dollars for each year of the report, with an average of \$6.1 million per year between 2017 and 2020.⁵ The increased financial stability of the Report series in recent years, while important, involves considerable transaction costs. The time and effort needed for fundraising is significant ([Ipsos MORI, 2018](#)). If the funding levels of donors diminish, for whatever reason, the resulting financial instability could compromise the quality of the Report, and possibly jeopardize its editorial independence. New donors may come with strings attached—seeking to further their own agenda by demanding greater influence over the Report's content ([Ipsos MORI, 2018](#)).

In terms of distribution, the evaluations indicate that the Reports could benefit from a more strategic collaboration with UNESCO, particularly its field offices, to better distribute the Report and disseminate its messages and recommendations to policy makers on the ground. In-country investigations revealed that UNESCO field offices rarely develop strategies to distribute reports, track its targeted recipients, or develop feedback mechanisms to share stakeholder views with the Report team ([Carpenter et al., 2014](#)). Indeed, relying on field offices results in significant geographical gaps in distribution, particularly in Eastern Europe and

⁵A table listing the annual contributions of each funder was provided to the authors by the GEM Report Director.

Central Asia. Most requests for hard copies are concentrated in Europe and North America, whereas digital copies receive more downloads in lower income countries (Carpenter et al., 2014).

With regards to impact, the external evaluations used different methods to gauge the Report's influence. They measured the number of times that a GEM Report is mentioned in the media. They analyzed report launch strategies and distribution lists. They interviewed select international, regional and national stakeholders, who are involved in international development and education. For example, a survey of 1228 global education stakeholders conducted for the 2018 evaluation found different audience groups use the GEM Report for different purposes. The survey found most respondents used the GEM Report as a reference to support their existing work or study (76%), reference to identify good practice for policy (59%), or as a source of personal development and training (54%). Fewer respondents reported using the GEM Report as a source of evidence to influence strategy, programming, or policy in their organization (48%), or as an advocacy tool outside of their organization (32%) (Ipsos MORI, 2018, p. 46).

Overall, stakeholder-respondents find the reports to be relevant to education priorities in the international context, as well as relevant to their work on the ground. That said, many lack awareness of the critical role of UIS in sourcing and validating internationally comparable data. This means that some respondents interpret data shortcomings to be fault of the Report itself (Carpenter et al., 2014). For example, some critiqued the GEM Report for reporting statistical data that were not up-to-date—a needed condition for effective decision making, particularly in the Global South (Carpenter et al., 2014; Ipsos MORI, 2018; Universalia, 2006). The 2014 evaluation quoted an Indian official who indicated that they rarely use data from the Report because it is out-of-date; they prefer annual education data generated by more reliable government sources (Carpenter et al., 2014, p. 18). Another issue was the use of differing methodologies—national vs. international—to estimate specific data points. For example, Viet Nam disputed school drop-out figures in the 2008 Report since their Ministry used a different methodology from UIS (Carpenter et al., 2014).

The evaluations also use bibliometric methods to assess the citation impact of the reports. Such methods provide evidence of a substantial growth in citations of the reports in scholarly and academic publications. Each of the Reports published between 2002 and 2010 garnered more than 100 citations in publications included in the SCOPUS database, with a substantial increase in the number of times the reports are cited after 2009 (Carpenter et al., 2014). Most of these citations came from the authors affiliated with US or UK based research institutions, but also authors based in India, South Africa, Ghana, and other East African countries. By contrast, few authors from Southeast Asia, Central Asia, or Latin America cite the report in publications archived by SCOPUS (Carpenter et al., 2014, p. 32). 257 academic publications cited the 2015–2017/18 reports, including scholarly articles, books and book chapters, proceedings papers, editorial materials, letters and reviews (Ipsos MORI, 2018, p. 15). And the GEM Report remained the most widely cited global monitoring report on education (Ipsos MORI, 2018, p. 39).

With regards to the report evidentiary base, the 2014 evaluation noted that these "... sources are dominated by English language publications and research organizations in the global North" (Carpenter et al., 2014, p. 13). Specifically, 84% of all background reports commissioned between 2011 and 2014 were produced by authors with organizational affiliations in Europe and North America. By contrast, only 6% were prepared by authors with affiliations with East Asia, 4% with South and West Asia, 3% with Sub-Saharan Africa, 2% with Latin America, and only 1% with an organizational affiliation with the Arab States (Carpenter et al., 2014). This pattern reinforces the uneven geographical distribution of expertise mobilized for and reflected in the reports. Despite efforts to address this issue—through in-house literature reviews and by establishing links with Global South research institutions as potential contributors and research partners—the reports have clearly struggled to integrate Global South and Indigenous voices into the reports. Thus, while representatives from NGOs, multilaterals, and research institutions and the evaluation teams felt that the commissioning of background reports for the GEM Report is invaluable to the education for development field, "... as there are few other sources of funding to stimulate such focused research in the education sector [and the] commissioning work of the Report has potential to build research capacities in all regions of the world" (Carpenter et al., 2014). And yet there is a clear need to make sure that Global South authors have sufficient opportunity to participate in GEM Report think-pieces, consultations, and background papers to improve geographic representation of the Report, enhance ownership, and localize findings (Ipsos MORI, 2018, p. 71).

Finally, the evaluations revealed tensions between the Report's role as a technical document providing analytical data to inform decision making, and its role as an advocacy document alerting the global education community to areas of concern and advancing specific policy recommendations (Universalia, 2006, 2010). Some stakeholders believe the Report should be a prescriptive document, spotlighting key advocacy messages to achieve real change and commitment from international stakeholders. Others insist the Report must be an "objective" document that maintains neutrality in its effort to monitor the situation in some sub-sector of education, arguing that a prominent advocacy role could compromise the objectivity and credibility of the reports.

Looking ahead at monitoring challenges and pandemic recovery

The post-2015 global monitoring landscape is complex and challenging. UN Members States agreed on the global indicators for each SDG targets in 2017 after 2 years of negotiation. And "the global indicator framework ... is a voluntary and country-led instrument International organizations shall consult ... countries to produce and validate ... estimates before publication ..." (UN, 2017). While countries were encouraged to submit Voluntary National Reviews, these reviews varied in scope and detail, and rarely provided data on all education targets.

SDG monitoring data were not readily available in most national education statistical agencies. The new 11 global indicators referred to undeveloped measurements, or datasets with partial global coverage. Indeed, missing data and proxy measures have become a global monitoring norm in the post-2015 period.

Another important shift involved calls to use “nationally determined benchmarks” to track country progress of SDG4 targets. The Education 2030 Framework for Action (adopted in Incheon) called on countries to establish “appropriate intermediate benchmarks (e.g., for 2020 and 2025)” for SDG indicators, seeing them as “indispensable for addressing the accountability deficit associated with longer-term targets” (UNESCO, 2015, para 28). The 2030 Framework implicitly recognized the fact that, although countries committed themselves to the same education targets and global indicators, they wanted their progress to be assessed based on an existing starting point and according to known (historical) rates of progress. In short, rather than using a common endpoint within a set time frame to assess the distance a country needs to move to achieve a specific goal (as was the case under EFA), national benchmarks would capture the contribution each country would be prepared to make to the global education agenda. Getting countries to engage in the SDG4 benchmark-setting process has not been easy. Apart from EU and CARICOM members, who had already committed themselves to clear targets, only half of all remaining countries either submitted national SDG4 benchmarks or committed to do so in the coming year (UNESCO, 2021).

The unprecedented ambition of the SDG4 agenda has certainly experienced a significant deflation with the onset of the Covid-19 pandemic. Countries have closed schools, disrupting teacher-student interactions and the regularity of school-based routines. Pre-existing global inequalities have widened and entirely new gaps have been created. Learning loss, lower retention rates, and a significant increase in out-of-school children and adolescents have all become tangible outcomes of the pandemic, even if it subsides. Policy responses to COVID-19 have demonstrated just how much global leaders rely on tables, figures, and current statistics for guidance in making critical and time-sensitive decisions. A lack of reliable, real-time data has been a key barrier to understanding the full educational impact of COVID-19. Some estimates suggest that the disruptions caused by COVID-19 may have set the world back many years in terms of educational gains. Getting education systems back on track will be difficult (UN, 2021).

One could argue that it is precisely during unprecedented and disruptive times that trusted, independent sources of data, evidence and analysis are needed to ensure smart and effective decisions. The GEM Report would appear to be well-positioned to adjust its evolving monitoring agenda and address unforeseen challenges such as those currently facing national authorities, donors, NGOs and other education advocates due to the COVID pandemic. Pressures on national budgets, partly due to increased health outlays, have reduced funds for international development assistance. By 2020 it was clear that squeezed budgets in OECD countries had translated into reduced financing for global education (UNESCO, 2020a). In response, the GEM Report platform prepared a series of policy papers providing revised funding estimates required to meet the SDG4 targets by 2030 (UNESCO, 2020b), as well as guidance to inform decision makers seeking financing solutions (UNESCO, 2021). These publications, along with previously published GEM Reports, helped inform the direction of the Global Education Meeting held in Paris France (UNESCO, 2020b), ultimately resulting in the 2020 Paris Declaration—a global call to invest in the futures of education.

By continuing to be responsive to the global education community as a whole, rather than the interests of individual agendas, the GEM Report can strengthen its role as an integral pillar of the global education architecture. Over the past 20 years the Report series has leveraged its independence to bring diverse ideas and perspectives to global education actors and policy makers. In the coming years, the headwinds facing the implementation of the 2030 education agenda are acute. For many this is a time to pause, take stock, and re-prioritize. A global education monitoring report series of high quality—one that is nimble, but steady; one that both informs and provokes—as new strategies are considered is a crucial asset. Whether this global good will be resilient or squandered remains to be seen.

References

- Alexander, R., 2015. Teaching and learning: the quality imperative revisited. *Int. J. Educ. Dev.* 40, 250–258.
- Benavot, A., Antoninis, M., Bella, N., Delprato, M., Härmä, J., Jere, C., et al., 2016. Education for all 2000–2015: the influence of global interventions and aid on EFA achievements. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley Blackwell, West Sussex, UK, pp. 241–258.
- Boel, J., 2016. UNESCO's Fundamental Education Program, 1946–1958: vision, actions and impact. In: Duedahl, P. (Ed.), *A History of UNESCO: Global Actions and Impacts*. Palgrave Macmillan, New York, NY, pp. 153–167.
- Brown, L., 2006. EFA global monitoring report: literacy for life. *Int. J. Educ. Dev.* 27, 473–482.
- Burnett, N., 2010. How to develop the UNESCO the world needs: the challenges of reform. *J. Int. Coop. Educ.* 13 (2), 89–99.
- Burnett, N., 2011. UNESCO education: political or technical? Reflections on recent personal experience. *Int. J. Educ. Dev.* 31 (3), 315–318.
- Burnett, N., 2019. Invited Essay: it's past time to fix the broken international architecture for education. *Int. J. Educ. Dev.* 68, 15–19.
- Carpenter, J., Daniel, P., Grout-Smith, J., Ojwang, A., Tanner, S., Wood, J., 2014. External Evaluation of the Education for All Global Monitoring Report: Final Report. London.
- Cussó, R., D'Amico, S., 2005. From development comparatism to globalization comparativism: towards more normative international education statistics. *Comp. Educ.* 41 (2), 199–216. <https://doi.org/10.1080/03050060500037012>.
- Cussó, R., 2006. Restructuring UNESCO's statistical services—the “sad story” of UNESCO education statistics: 4 years later. *Int. J. Educ. Dev.* 26, 532–544.
- Dale, R., 2005. Globalisation, knowledge economy and comparative education. *Comp. Educ.* 41 (2), 117–149.
- Daniel, J., 2010. Reflections on UNESCO's World Education Reports and the Global Monitoring Reports on Education for All. *NORRAG News*. Retrieved from: <http://www.norrag.org/en/publications/norrag-news/online-version/a-world-of-reports-a-critical-review-of-global-development-reports-with-an-angle-on-education-and-training/detail/reflections-on-unesco-world-education-reports-and-the-global-monitoring-rep>.
- Delamonica, E., Mehrotra, S., Vandemoortele, J., 2001. Education for All is Affordable: A Minimum Global Cost Estimate (UNICEF Staff Working Paper, Evaluation, Policy and Planning Series No. EPP-01-001). Florence, Italy.

- Doemeland, D., Trevino, J., 2014. Which World Bank Reports are Widely Read? (English). Policy Research working paper; no. WPS 6851. World Bank Group, Washington, D.C. Retrieved from: <http://documents.worldbank.org/curated/en/387501468322733597/Which-World-Bank-reports-are-widely-read>.
- Duedahl, P., 2016. Part III: experts on the ground. In: Duedahl, P. (Ed.), *A History of UNESCO: Global Actions and Impacts*. Palgrave Macmillan, New York, NY, p. 154.
- Edwards, D.B., Okitsu, T., Da Costa, R., Kitamura, Y., 2018. Organizational legitimacy in the global education policy field: learning from UNESCO and the Global Monitoring Report. *Comp. Educ. Rev.* 62 (1), 31–63. <https://doi.org/10.1086/695440>.
- Global Education Monitoring Report, 2016. *Past Achievements, Future Visions: 2016–2018: A Strategy Document*. UNESCO, Paris.
- Global Education Monitoring Report, 2021. *How Committed? Unlocking Financing for Equity in Education (Policy Brief 44)*. UNESCO, Paris. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000375326>.
- Global Education Monitoring Report, 2022. *The Global Education Monitoring Report. In Brief*. Retrieved from: <https://en.unesco.org/gem-report/about>.
- Gordon, C., 2010. Reflecting on the EFA Global Monitoring Report's framework for understanding quality education: a teacher's perspective in Eritrea. *Int. J. Educ. Dev.* 30 (4), 388–395. <https://doi.org/10.1016/j.ijedudev.2009.12.007>.
- Huxley, J., 1946. *UNESCO: Its Purpose and Philosophy*. Paris.
- Ipsos MORI, 2018. *Evaluation of the Global Education Monitoring Report*.
- Jamil, R., 2016. *Reflections: The First Global Education Monitoring Report's Advisory Board Meeting*. Retrieved from: <https://en.unesco.org/gem-report/reflections-first-global-education-monitoring-reports-advisory-board-meeting>.
- Jones, P.W., 1988. *International Policies for Third World Education: UNESCO, Literacy and Development*. Routledge, London and New York.
- Lewin, K.M., 2011. *Taking Targets to Task Revisited: How Indicators of Progress on Access to Education Can Mislead*. Falmer, UK.
- Little, A., Miller, E., 2000. *The International Consultative Forum of Education for All 1990–2000: An Evaluation*.
- Magee, A., 2020. *Commentary on the 2019 Global Education Monitoring Report on Migration*. Retrieved from: <https://www.norrag.org/commentary-on-the-2019-global-education-monitoring-report-on-migration-by-arran-magee/>.
- Magrath, B., 2014. *Advocacy as Political Strategy: The Emergence of an "Education for All" Campaign at ActionAid International and the Asia South Pacific Association for Basic and Adult Education*. University of Toronto.
- Menashy, F., Manion, C., 2016. The historical evolution and current challenges of the United Nations and global education policymaking. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley Blackwell, Sussex, UK, pp. 319–334.
- Mundy, K., Manion, C., 2015. Education for All initiative: history and prospects post-2015. In: McCowan, T., Unterhalter, E. (Eds.), *Education and International Development: An Introduction*. Bloomsbury, London, UK, pp. 49–68.
- Mundy, K., 1999. Educational multilateralism in a changing world order: UNESCO and the limits of the possible. *Int. J. Educ. Dev.* 19 (1), 27–52.
- Mundy, K., 2006. Education for all and the new development compact. *Int. Rev. Educ.* 52 (1), 23–48.
- Nicola, F.G., 2021. Scandal involving World Bank's "Doing Business" index exposes problems in using sports-like rankings to guide development goals. *Conversation*. Retrieved from: <https://theconversation.com/scandal-involving-world-banks-doing-business-index-exposes-problems-in-using-sportlike-rankings-to-guide-development-goals-169691>.
- Packer, S., 2008. The education for all global monitoring report: a mid-term assessment. *Prospects* 38, 287–293.
- Read, R., 2017. *Knowledge Counts: A Bibliometric Analysis of the Education for All Global Monitoring Reports*. University of Toronto.
- Redman, K., 2020. *The GMR—and How It Has Contributed to EFA*. Retrieved from: <https://en.unesco.org/gem-report/gmr—and-how-it-has-contributed-efa>.
- Robertson, S.L., Dale, R., 2006. New geographies of power in education: the politics of rescaling and its contradictions. In: Kassem, D., Mufti, E., Robinson, J. (Eds.), *Education Studies: Issues and Critical Perspectives*. Open University Press, Berkshire, UK, pp. 221–232.
- Skillbeck, M., 2000. *Education for All: Global Synthesis*. Paris. Retrieved from: <http://unesdoc.unesco.org/images/0012/001200/120058e.pdf>.
- Spaulding, S., Chaudhuri, R., 1999. UNESCO's world education report: its evolution, strengths and possible futures. *Int. J. Educ. Dev.* 19, 53–63.
- UIS, 2000. *Education for All 2000 assessment: Statistical Document*. Paris. Retrieved from: <http://unesdoc.unesco.org/images/0012/001221/122124e.pdf>.
- UNESCO, UNICEF, World Bank, UNFPA, UNDP, UN Women, UNHCR, 2015. *Education 2030, Incheon Declaration and Framework for Action*. World Education Forum, pp. 1–107, 2015.
- UNESCO, 1946. *Constitution*. UNESCO, Paris. Retrieved from: http://portal.unesco.org/en/ev.php-URL_ID=15244&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- UNESCO, 1996. *The Mid-Decade Review of the International Consultative Forum on EFA*. Amman.
- UNESCO, 2000. *The Dakar Framework for Action*. UNESCO, Paris. Retrieved from: http://www.unesco.at/bildung/basisdokumente/dakar_aktionsplan.pdf.
- UNESCO, 2001. *High-Level Group on Education for All. First Meeting*, Paris.
- UNESCO, 2005. *Education for All Global Monitoring Report: The Quality Imperative*. UNESCO, Paris. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000137333>.
- UNESCO, 2015. *Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4 Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning Opportunities for All*. Retrieved from: http://uis.unesco.org/sites/default/files/documents/education-2030-incheon-framework-for-action-implementation-of-sdg4-2016-en_2.pdf.
- UNESCO, 2020a. *What You Need to Know About the Right to Education*. Retrieved from: <https://en.unesco.org/news/what-you-need-know-about-right-education>.
- UNESCO, 2020b. *Background Document: 2021 Global Education Meeting—High-Level Segment Invest in Education—a Global Mobilization for COVID-19 Recovery*.
- UNESCO, 2021. *Global Education Monitoring Report 2021/2: Non-state Actors in Education: Who Chooses? Who Loses?* UNESCO, Paris. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000379875>.
- United Nations, 2017. *Resolution Adopted by the General Assembly on 6 July 2017: 71/313. Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable Development*. United Nations, New York. Retrieved from: <https://undocs.org/A/RES/71/313>.
- United Nations, 2021. *The Sustainable Development Goals Report 2021*. United Nations, New York. Retrieved from: <https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021.pdf>.
- Universalia, 2006. *Formative Review of the Education for All Global Monitoring Report Volume I—Summary Report*.
- Universalia, 2010. *Final report. Evaluation of the Global Monitoring Report, vol. I*. Universalia Management Group, Montreal.

NGOs & international development: ongoing trends and new architectures

Sahara Pradhan, Verity Norman-Tichawangana, and Sangeeta Kamat, International Education Program, College of Education, University of Massachusetts at Amherst, Amherst, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	565
A Gramscian conceptual framework	566
Critical scholarship on NGOs	566
NGOization	567
The problematic promise of PPPs	568
NGOs and the new aid architecture	568
Implications for the NGO sector	569
A time of reckoning for the NGO sector	570
References	571

Introduction

Following early scholarship on NGOs in the 1990s, the NGO sector has undergone dramatic changes that call for fresh approaches to research and theorization. Robert Arnove's edited volume *Philanthropy and Cultural Imperialism: the Foundations at Home and Abroad* (1982) marks the first significant study of the non-profit sector in the field of international education. Since then, scholarship on the topic has expanded greatly with exciting new research from the field and from disciplines of anthropology, sociology and political science. One of the challenges of theorizing the NGO sector is its highly pluralistic character that encompasses multiple organizational forms with diverse missions and histories. The diversity of the NGO sector has generated several acronyms to distinguish between different types of NGOs such as CBOs (community-based NGOs), GROs (grassroots organizations), INGOs (international NGOs) and even GONGOs (Government NGOs).

The heterogeneity of the NGO sector notwithstanding, a persistent and common feature is the power inequity between Northern and Southern NGOs. At a basic level, the hierarchies are constituted by donor or aid INGOs based in the Global North and recipient or beneficiary NGOs located in the Global South. The former determines the flow of funding to the latter, creating a dependence on external aid and donors for Southern NGOs to staff and sustain their work. A critical body of scholarship has examined these power differentials and argued that NGOs are part of the problem rather than the solution to global poverty and inequality in the ways that they reproduce relations of patronage, neocolonialism and imperialism. These critiques are also reflected in discussions that have taken place within INGOs of ways to restructure and reform the NGO sector to be more equitable and reverse power imbalances between Northern and Southern NGOs. OXFAM and Action Aid for example, are two prominent INGOs that more than a decade ago made operational changes to allow for greater autonomy and leadership to their regional offices in the South (Scott-Villiers, 2002; Eade, 2004).

More recently, in a context where the Black Lives Matter movement has brought into focus racial capitalism on a global scale, calls to "decolonize" the NGO sector are on the rise by activists and NGO workers, including by international NGO staff based in the North (Busby, 2021; DeLuca and Williams, 2020). These recent developments require a careful accounting of NGOs to ask what "decolonization" would mean for different kinds of NGOs. The essay maps important trends in the NGO sector and the undeniable influence of neoliberal capitalism on the internal workings of the sector in the North and the South. We discuss key themes in the critical scholarship on NGOs that inform theory, policy and practice in the sector. We highlight significant challenges posed by the corporatization of aid and philanthropy that have transformed the development aid architecture with implications for the NGO sector worldwide. The following section outlines the conceptual approach that informs our analysis of the NGO sector. We explain how Gramsci's theorization of state and civil society offers a more appropriate framework to theorize the NGO sector in contrast to liberal political theory that is unable to adequately explain the evolution and nature of the sector. The third section elaborates upon critical scholarship on the NGO sector starting in the nineties and highlights the important contributions of the multi-disciplinary and geographically diverse critiques of the sector. The fourth section is on the rise of venture philanthropy as a recent phenomenon that has radically disruptive consequences for the NGO sector and its functioning. Drawing on existing literature and our own research, we discuss the neoliberalization of the NGO sector in greater detail as well as its latest iteration as venture philanthropy has led to a new architecture of development and restructured the NGO sector. For one, we argue that this mode of philanthropy has enabled powerful corporations to become agenda setting actors, subordinating NGOs to implementation agencies with limited leverage in the global arena that nonetheless play a complex role in generating "consent" and legitimacy for a corporatized social order. In conclusion, we critically examine calls to decolonize the NGO and aid sector to understand its limits and potential for more egalitarian and just practices, structures and relations. What does a demand to "decolonize" imply for INGOs such as USAID, DfID and IRC that are closely tied to imperial states and their interests? What would a genuine decolonization entail for the NGO sector more broadly? These are questions we engage to conclude the essay.

A Gramscian conceptual framework

The global reach of the NGO sector, governing almost every aspect of development, has drawn the attention of scholars and researchers for more grounded and rigorous theorizing that is both contextually specific while also offering general insights about the phenomenon. Liberal political theory that analyzes state and civil society as separate and distinct spheres tends to locate NGOs as part of civil society, separate and autonomous from state and market (the latter also conceptualized as part of civil society). Research studies within this framework study NGO relations and processes as voluntaristic and self-driven, emphasizing individual organizations and leaders as the primary actors that shape the field, at times in fraught relations but in a general equilibrium with state and market forces. An example of such work can be found in the writings describing BRAC and Mohammed Yunus that promote micro-credit NGOs as a universal and normative model to address global poverty (Yunus, 2007). A second conceptual approach to the study of NGOs is neo-institutionalism that applies the principles of organizational analysis to examine the rules, norms and institutional dynamics that shape the work of NGOs and the actions and choices of its actors (Edwards and Fowler, 2002; Lewis, 2007; Schoenefeld, 2021).

Critical scholars of NGO studies (including authors of this essay) posit that these theoretical frameworks are limited in their capacity to account for the growth and influence of the NGO sector or explain the dynamics by which the sector has emerged as a powerful actor that shapes global policy on health, education, and the economy. We propose a dialectical approach that considers NGOs as one albeit important node in the “structuring structures” of states, finance capital, multilateral institutions of the World Bank and the IMF and a diverse range of social movements. This departs from the orthodoxy that locates NGOs exclusively as an actor in civil society understood as separate from and often antagonistic with the state. Instead, we suggest a Gramscian approach in which the sector is theorized as part of the broader political economy of global capitalism where state and civil society are in a dialectical relation, with the former representing relations of coercion and the latter the arm of consent (Gramsci, 1971). As organizations of civil society, NGOs function more often than not as this arm of consent and perform the important function of legitimating the dominant political and economic order. Moreover, the nature of the state and the overall policy and social movement context has a determinate influence on the nature, size and efficacy of the NGO sector that makes regional specificity a consequential factor in one’s analysis. As such, the role of NGOs must be theorized not in static and ahistorical terms but as contextual and in a dynamic relationship with other aspects of the economy and society (Kamat, 2013). As we illustrate in this essay, changes in national contexts and the global political economy have consequences for the “relative autonomy” (to borrow Bourdieu’s formulation) of NGOs that range from limiting aid and funding to NGOs to transforming the nature and scope of their work. Following Gramsci’s thesis that hegemony is never a complete exercise, we also highlight the mobilization of oppositional discourses that seek to counter and resist the widespread neoliberalization and corporatization of NGOs.

Critical scholarship on NGOs

The international political order of the 21st century is dominated by the “neoliberal turn” (Harvey, 2006) that propagates the doctrine of market principles in all sectors of society. Neoliberalism as a global policy regime is traced to the Thatcherism and Reaganism of the 1980s, and the growth of the NGO sector is understood in large part as an outcome of the austerity politics that was inaugurated in this period. As policies of deregulation and privatization were imposed on the global South by World Bank/IMF structural adjustment programs (SAPs), the state relinquished its responsibilities and obligations in providing healthcare, education, and other social services. As privatization and deregulation became the norm, impoverishment and prolonged suffering increased for so many living in the so-called developing world, thereby creating conditions ripe for expanded NGO intervention and “development” (Hancock, 1989; Kajimbwa, 2006). NGOs quickly filled this space of providing public services that previously were the responsibility of the state.

Since then, a rich body of scholarship has evolved that theorizes how NGOs are a key feature of neoliberal globalization disputing the characterization of NGOs as simply expressions of democracy (Bernal and Grewal, 2014; Elyachar, 2005; Ferguson, 2006; Fisher, 1997; Kamat, 2002, 2004; Karim, 2011; Mercer, 2002; Schuller, 2013; Sharma, 2008; Silova and Steiner-Khamisi, 2008). Several scholars have documented how NGOs have proliferated under neoliberalism while intensifying competition for funding and international aid. NGO-led social and educational programs increasingly adhere to market principles and are tied to corporate actors (Bhanji, 2008, 2012; Srivastava, 2010; Moeller, 2020; Tarlau and Moeller, 2019). With funding from northern multinational organizations such as the World Bank (WB), the International Monetary Fund (IMF), the United Nations (UN), and other international donor agencies flowing to the global South, neocolonial dependency is maintained under the guise of democratization. Scholars have compared the work of the NGO sector to that of missionaries and colonial voluntary organizations and have argued that NGOs function as a tool for geopolitical and economic control (Manji and O’Coill, 2002; Petras, 1999). Accompanying this disparity of economic control and who determines the funding priorities have been calls for further investigation into where that funding goes, and who it benefits (Barghouthi, 1993; Maran 2006). In this way, NGOs represent an important vector in legitimating and consolidating neoliberalism and neocolonialism in structural and ideological ways.

In much of the literature of the 1990s, NGOs were increasingly positioned as watchdogs of the state and as part of an “emerging civil society,” as the official agenda of development aid and development discourse shifted toward promoting “democracy” and “good governance” (Mercer, 2002; Choudry and Kapoor, 2013). This period marked an intensification of neoliberalism, and development was seen as emerging from the triad of a balanced relationship between “government, market and the third sector,”

ushering in the empowerment of NGOs as powerful agenda-setting actors on the global stage (Lewis and Kanji, 2009). In an environment shaped by global trade, neoliberal policy regimes and national policy frameworks, NGOs became effective mediators of the “new policy agenda,” pursuing goals that coalesced with global capitalist interests (Fisher, 1997; Kamat, 2004). Backed by substantial foreign aid, NGOs became seen as “agents of democracy” that would usher in Western-type liberal capitalist regimes to so-called “transitional” countries. Based on their in-depth study of NGOs in post-Soviet contexts, Silova and Steiner-Khamisi argued that in these contexts, NGOs played “complementary, cooperative, or surrogate” roles (Silova, 2008, p. 27) in relation to the state, and that this role was one of navigating a messy middle ground—tasked by international donors to promote democracy and raise marginalized voices, while also contending with the on-the-ground realities of working with authoritarian states (Silova, 2008).

In the aftermath of the US invasion of Afghanistan, the role of the NGO sector has evolved further to that of “force multiplier,” in the words of Colin Powell, playing a critical role as part of the US government’s arsenal of war (Greenburg, 2017). In the so-called “War on Terror” the line between military, development contractors and NGOs has become increasingly fuzzy. Literally embedded in military Green Zones, the work of NGOs in Afghanistan, even accessing the sites they’re working with, is dependent on and inextricably intertwined with the military-industrial complex in which aid programs function as part of the security apparatus (Duffield, 2012). As development became increasingly subordinated to the US government’s defense and diplomacy agenda, there has also been a marked expansion of multinational companies and private contractors like Chemonics receiving a major share of development aid compared to the NGO sector (Bilmes, 2021; Nagaraj, 2015; Rohde, 2012).

NGOization

As NGOs have become an integral factor in hegemonizing global capitalism, the neoliberalization of the sector has restricted their capacity to engage in activism and advocacy. The term “NGOization” refers to the professionalization and depoliticization of social change programs and efforts that is increasingly evident in different parts of the world (Bernal and Grewal, 2013; Choudry and Kapoor, 2013; Ismail and Kamat, 2018; Kamat, 2003; Schuller, 2013, 2017; Jaoul, 2018). While NGOs have supported people’s movements and social justice struggles, the depoliticization of the sector brings NGOs into lockstep with neoliberal globalization, bending to market principles of efficiency, productivity and measurable targets. Scholars have argued that the professionalization and depoliticization of NGOs undermines autonomous people’s movements and shrinks the space for activism and movement organizations in the global south (Benin, 2014; Choudry and Kapoor, 2013; Ismail and Kamat, 2018; Jalali, 2013). Albeit, in contexts of illiberal states, NGOs may also be a source of support for people and offer a space for building democracy reflecting the dynamic and dialectical formation of the sector (Alvarez, 2009; Arafa, 2017; Silova and Steiner-Khamisi, 2008).

One of the key manifestations of the neoliberal mandate of efficiency are the short-term funding cycles that prioritize development goals directed by the global North, creating short-sighted priorities that measure targets and outcomes, rather than long-term development goals (Zencirci and Herrold, 2022). While NGOs are idealized as apolitical entities that “do good” without the constraints of government bureaucracy or the greed of the market, the requirements of short-term projects and funding cycles and narrow targets fosters an “audit culture” that undermines development as a long-term, integrated and intersectional process (Choudry and Kapoor, 2013; Vannier, 2010). A host of “buzzwords” and terminology such as “participation,” “empowerment,” and “local community” accompany donor-led goals, bringing into view the paradoxes of NGO work that continue to impose top-down solutions while co-opting spaces of contestation and resistance (Bernal and Grewal, 2014; Cooke and Kothari, 2001; Cornwall, 2007; Jaoul, 2018; Kamat, 2004; Leve and Karim, 2001; Reimann, 2006; White, 1996).

The NGOization of development has exacerbated social and economic inequalities at multiple scales. For one, it has produced a concentration of relatively well-paid professionals, often from the global North, in metropolitan regions of the global South. While Northern expats are highly paid and develop top-down programmatic responses far from the sites of implementation, staff from the South are often devalued in both their compensation and experience relative to their Northern counterparts. This notwithstanding, the NGO sector comprises a significant segment of the professional middle class in poor economically precarious countries. As a Zimbabwean NGO colleague of one of the authors’ noted: “Well, our work may not lead to alleviating poverty, but at the least it has changed the lives of the NGO workers and we now constitute the ‘middle class’ in Harare!” As local professionals opt for better-paid, externally funded, albeit unstable jobs in the NGO sector, NGOs also drain the public sector of teachers, doctors, nurses, social workers and other professionals. Thus, rather than contributing to the strengthening of public services and institutions, NGOs too often contribute to its opposite, expanding the market for private actors and subcontracting that undermines the goal of sustainable development as such (Ferguson, 2006; Kamat, 2004; Mercer, 2002; Schuller, 2017; Simmons, 1998).

“NGOization” has also significantly restructured how civil society interacts with the state and the market, whether in the case of guaranteeing the right to education, addressing climate crisis or gender justice movements. For example, a number of scholars have shown that while women’s NGOs gave visibility and strength to feminist struggles, as neoliberalism created more hardship for poor women, many of these same feminist NGOs became more engaged in survival struggles, rather than emancipatory ones (Alvarez, 1999, 2009). Over time, women’s struggles transformed into efforts to simply integrate poor women into the market primarily through micro-finance and debt entrapment (Elyachar, 2005; Karim, 2011; Roy, 2010). While in some cases NGOs may be allied to or emerge from social movements (Waghmore, 2012), one of the foremost consequences of NGOization has been the displacement of political and social movements from the public sphere as professional well-funded NGOs take their place (Choudry and Kapoor, 2013; Bernal and Grewal, 2014; Dauvergne and LeBaron, 2014).

The problematic promise of PPPs

While NGOs have historically partnered with government entities or global donor organizations, in the new millennium, public-private partnerships (PPPs) have emerged that signal a closer alliance between the private sector and NGOs. Arguably, the rationale for PPPs is the private sector has greater capacity and resources that should be leveraged for social good. The underlying assumption is that the private sector, NGOs and the state have a common interest in public welfare and development (Kamat, 2014). In her book "Shock Doctrine," Naomi Klein (2007) foregrounds the problem of the state outsourcing public services to the private sector—there are few or no consequences for private sector withdrawal, particularly during a time of disaster or crisis. This ability to withdraw from the partnership with few or no consequences also highlights the imbalance in the relationship or partnership between the private and public sectors. If the public is the "customer" so to speak, how do they negotiate with the private sector and stand up for what is rightfully theirs when these services disappear or are not of an adequate quality? Further, who is benefiting from these partnerships? In her study of PPPs in South Africa Mirafteb (2004) documents how PPPs function as privatization by stealth and aptly terms it as the "trojan horse of development" (see also, Robertson et al., 2012; Srivastava and Oh, 2012; Srivastava, 2010). In the authors' own experience as the executive director of a Kenyan NGO (Norman-Tichawangana), it was transparently evident that the PPP project disproportionately benefitted the private company rather than the poor women identified as the ostensible beneficiaries of the project. It was manifestly evident as well that with the help of the NGO, the company would be able to tap into an extensive rural poor market in the future. The following section elaborates upon the evolution of the PPP model in which finance and technology companies supersede NGOs in the design and delivery of education products and services in ways that portend a radical restructuring of education and development on a global scale.

NGOs and the new aid architecture

Intensifying the neoliberal reforms of the prior decades, the architecture of NGO funding began to further shift in the early 1990s. Three decades later, it is increasingly dominated by a mode of philanthropy known as philanthrocapitalism or venture philanthropy, which characterizes large swaths of the non-profit sector. Venture philanthropy signals a dramatic shift in NGOs, not only because of the sheer growth in scale and volume of philanthropic spending (McGoey, 2012; Olmedo, 2019) but also because it is thoroughly guided by logics of governance that simultaneously corporatize the economic, political and cultural spheres and concentrate enormous power in the hands of corporate actors to set the global agenda. It also introduces a new "common sense" that coalesces the idea that corporate interests are seamlessly aligned with, rather than in contentious and contradictory relation to the public good, and that corporate actors and the market are increasingly better positioned than states to effect social change, leveraging NGOs to do so. These logics thus have far reaching implications that have further restructured the NGO sector and escalated the neoliberal, corporatist ideologies and pressures that govern their functioning. This section provides a brief overview of Venture Philanthropy in the education sector and examines the implications it has precipitated for NGOs writ large.

Venture Philanthropy is known by many names but was first popularized as *Philanthrocapitalism* by Bishop and Green in 2008 in their book subtitled *How the Rich Can Save the World*. They argued for a "truly fair social contract" which "might actually allow the rich who give generously more grace than their peers to indulge themselves free from criticism in their yachts, space flights, and Rod Stewart-serenaded birthday parties" (p. 484). At its heart, venture philanthropy espouses the idea that ultra wealthy actors can do "well" financially by doing "good" (Bishop and Green, 2008). In doing so, "philanthrocapitalism" seeks to obscure the relationship between poverty and wealth and eschews their inextricable connection within an unequal global capitalist system wherein the wealthy "have captured so much of the benefit of the growth of the global economy" (p. 542).

Venture philanthropy is distinctive in part because it is not only funded by billionaire venture capitalists but because it employs the principles and practices of private investing, fundamentally transforming traditionally non-profit sectors like education (Olmedo, 2019; Saltman, 2010; Tarlau and Moeller, 2019). While scholars have shown how this mode of philanthropy represents a continuation of strategic corporate giving in ways that benefit their businesses (Tarlau and Moeller, 2019) as well as prior models of Corporate Social Responsibility that provide brand and PR benefits (Bhanji, 2008), VPs stand apart because they model their interventions on venture capital, market competition and business principles that are driven by economic incentive, and treat funding as financial "investments," seeking multiple kinds social and financial "returns" (Baltodano, 2017; Hites, 2019; Kumashiro, 2012; Saltman, 2010; Scott, 2009; Tarlau and Moeller, 2019). It is these profit motivations and mechanisms of venture philanthropy as well as the constitutive aim to make philanthropy a "lucrative industry in itself" (McGoey, 2014, p. 111) that distinguish it "in shape and essence" (p. 128) from earlier modes of philanthropy and charity (Olmedo, 2019). Scholars have highlighted in particular that while traditionally, the "legitimacy of philanthropy stemmed from the strategic ignorance of both donors and recipients of the opportunities for economic advancement" (p. 194) under this "new" mode of philanthropy, self interest is no longer disguised but is rather valorized and championed as not only justifiable but as a "prerequisite," that will ostensibly serve the wider public, indeed the world (McGoey, 2012).

In many ways, venture philanthropy is fundamentally contradictory because it is defined by what Lipman (2015) describes as "dialectic of disinvestment and wealth concentration" (p. 242). Scholars have brought attention to the extractive, colonial and imperial modes of dispossession and wealth accumulation that undergird "venture philanthropy." Embedded in venture philanthropy's very existence are systemic and structural inequities (e.g. mining for oil and minerals, mass deforestation, violation of human rights, environmental degradation and destruction, extraction of data/surveillance, exploitation of labor and digital labor

on platform economies, etc.) that maintain grotesque stratifications between the “givers” and the “given to” and primarily disenfranchise working class and communities of color both in the Global North and South. For example, the Gates Foundations’ revenue sources include endowments invested in companies like Dutch Shell and Exxon Mobil Corp “which are responsible for polluting the region [Nigeria] beyond levels permitted in Europe or North America” (McGoey p. 190) or corporations that have fueled health epidemics (such as McDonalds and Coca-Cola). Gates has also made his own personal fortune in part, amassed through exploitative labor practices. As disparities between the rich and poor have grown obscenely wider, corporations who fund social interventions have also deliberately resisted taxation responsibilities, which creates further social exclusions and most impacts those that corporate philanthropies claim to “give back” to (Baltodano, 2017; Barkan, 2011; Gurn, 2016; Lipman, 2015; Scott, 2009). In many ways, venture philanthropy disguises profit making investments as philanthropy (Saltman, 2018) and are directly implicated in creating and maintaining the inequities they claim to alleviate.

Implications for the NGO sector

In changing the locus of funding within the NGO sector, venture philanthropy is enacting what Lipman has called the “deeply colonial project” of shifting power and decision making further away from communities, giving corporate actors the enormous latitude to remake public systems in ways that exclude the most disenfranchised communities (Lipman, 2015). In doing so, they are also transforming the architecture and governing logics of the NGO sector.

Among the foremost challenges posed by this mode of “giving” is that it further deepens the anti-democratic tendencies of NGOs that maintain an unequal international global political economy. Venture philanthropy enables the billionaire class to disproportionately exert influence on local and national legislation and on public policy, with little to no public oversight or accountability (Baltodano, 2017; Saltman, 2010) at the same time as the claims of effectiveness and efficiency of foundation interventions have been contested (Srivastava and Oh, 2012). Such enormous power in the hands of venture philanthropists signals a further neoliberal structuring of society that erodes public accountability and profoundly risks any possibility of an independent social and advocacy mission for NGOs. One such example is the World Health Organization’s malaria program that is so heavily dominated by funding from the Gates Foundation that it risks compromising the WHO’s independence and as such, public health as a whole (McGoey, 2012). This was also demonstrated in the context of the ongoing Covid-19 pandemic, during which Bill Gates, who came out in favor of vaccine patents, “mobilized his immense personal wealth and power to ensure that the interests of for-profit drug companies prevail over global public health” (Savage, 2021).

Venture philanthropy also represents a more seamless suturing together of corporate interest and NGO sector as corporate sector actors themselves enter into NGOs and the space they occupy and as NGOs function more and more in corporate-like ways. Unlike earlier, traditional “hands off” forms of philanthropy, venture philanthropists play a more “hands on” role in getting directly involved in their investments—in goal setting, decision making, evaluation, etc., giving them greater control over the reform process (Kumashiro, 2012; Saltman, 2010) and in precipitating large scale transformations. This in turn has placed disproportionate emphasis on measurable “data” and quantification within NGOs, a device that dominates the corporate sphere. Consequently, sectors like education are read in data-driven, narrow, decontextualized ways and draw on technocratic rationales to define “success” in terms of “return for investment” or “value,” amplifying tendencies that accompanied the NGOization of the sector. As such, NGOs are often answerable to metrics and performance indicators not only to donor NGOs but to corporate actors that seek tangible, quantifiable improvements geared toward marketization, entrepreneurship and commodification. As corporatization in the NGO sector becomes more hegemonic, a qualitative shift has taken place as NGOs become thoroughly guided by the same logics that drive venture capital and the boundaries between business and philanthropy become more and more indistinguishable. The role of the state in the corporatization of the NGO sector also cannot be understated, as governments have facilitated or directly funded corporations, forming a powerful nexus of “state-market” hybrids with which NGOs interact in complex ways (McGoey, 2014).

Many VP-backed initiatives have also been able to garner grassroots “consent” and “technical consensus” from social and institutional actors by channeling the enormous financial resources, media power and networks at their disposal (Tarlau and Moeller, 2019). Examples, among numerous others, include the Gates Foundation’s single-handed role in “bankrolling” and creating wide political support for enacting the Common Core standards in the US, which demonstrated VPs extraordinary influence in shaping public policy, swiftly after which “Microsoft announced that it was joining Pearson, the world’s largest educational publisher, to load Pearson’s Common Core classroom materials on Microsoft’s tablet, the Surface,” allowing Microsoft to compete with Apple for school district monies (Layton, 2014). However, Tarlau and Moeller (2019) argue, writing in the context of curricular reform in Brazil, that their philanthropic influence is not simply geared toward profit maximizing but is rather, “an attempt by foundation and corporate leaders to garner power and influence on different scales, and re-make public education in their own image” (Tarlau and Moeller, 2019, p. 3). This disproportionate and hegemonic influence held by venture philanthropists over local, national and global agendas as such requires critical reflection on who and what constitutes the “grassroots.” As Mitchell and Lizotte (2014) have shown in the case of Seattle, corporate actors partake in a form neoliberal governance that relies on the moral authority of the philanthropic “gift” which creates a sense of ideological obligation toward the “givers.” What appears to be grassroots, participatory democracy upon closer observation can be traced to strategically funded institutional networks cultivated by philanthropic funding. Leveraging these networks to generate “consent,” venture philanthropy quickly brings even genuine grassroots demands that emerge from the systematic defunding of public systems and services into the narrow purview of market based solutions, furthering the consequences of “NGOization.”

Today, many corporate entities themselves are simultaneously helping aid and philanthropy initiatives directly while also funding a network of grassroots NGOs and conducting political lobbying to shape public policies that are conducive to their agenda. This both limits NGOs to act in the global arena and constrains their latitude in advocacy, reducing them to implementing organizations or subcontracting agencies that represent corporate and corporate foundation interests. Venture philanthropy seeks to solidify a “common sense” that corporate actors are better suited than states and NGOs to effect change, and so leverage them to stabilize their positions and legitimate global capitalism. NGOs play an important role in generating consent and consensus for a reconfiguration of relationships, institutions and subjectivities in order to facilitate market incentives to drive social change in service of a “socio/neoliberal” order (Olmedo, 2019).

A time of reckoning for the NGO sector

In closing, we wish to draw attention to two contradictory trends whose effects on the NGO sector needs further research and analysis. One, the rise of authoritarian rule in erstwhile democracies such as India, Brazil and the Philippines that have imposed punitive laws which radically delimit the scope and functioning of domestic and international NGOs. Parallel to this is the rise of militant right-wing and protectionist politics in donor nations such as France, UK and the US that impacts the extent and scope of aid to NGOs in the South (Davidson and Saull, 2017). A second contrasting trend is the centering of race and settler colonialism as a global issue, sparked by the long summer of Black Lives Matter protests in the US that resonated around the world. The NGO sector is responding to this moment, as evidenced by calls to review race relations within the sector and appeals to “decolonize” the NGO sector. Following a dialectical approach, it is important to reckon with the ways in which NGOs are likely to be impacted by these conflicting global tendencies. Equally important is to examine the ways in which NGOs address these emergent tendencies and how the sector can be supported as well as held accountable in this critical conjuncture.

The NGO sector appears to be one of the first casualties of democratically elected authoritarian governments such as India, Brazil and the Philippines, that represents a crisis of democracy with global repercussions. Governments in these countries have targeted reputed international NGOs such as Amnesty International, Human Rights Watch, Ford Foundation and Greenpeace and blocked their funding, harassed staff and closed down operations in the country (Amnesty International, 2019; Sherwood, 2015). The right wing governments have unleashed vicious campaigns of disinformation attacking NGOs as the source of political conflicts in the country. For example, in response to the widespread condemnation of his destructive policies in the Amazon, Bolsonaro has led media campaigns to attack NGOs as the source of political conflict (Watts, 2019). Similarly, Duterte alleged that human rights NGOs that raised the alarm on extra-judicial killings by his government, are linked to drug lords in the country (Human Rights Watch, 2018). Under Modi’s rule, in the last ten years, over 20,000 NGOs have had their licenses revoked prohibiting them from receiving foreign aid (Swart, 2020). Advocacy NGOs such as Amnesty International, Ford foundation, Greenpeace have been forced to close their offices and others to drastically reduce their in-country operations. In several countries, NGOs that depend on foreign aid to provide services such as education and health care to poor and minority communities have been surveilled, harassed and shut down (Arafa, 2017; Hecht, 2018; Mellen and Lynch, 2017). On the other hand, NGOs that are aligned with the states’ religious supremacist agenda in India and Pakistan have thrived in this period. In other words, the space for NGOs that challenge the human rights abuses of authoritarian governments has drastically shrunk in this decade. This is a moment of reckoning for the NGO sector that is contemporaneous with the crisis of democracy and raises questions about whether NGOs can fulfill their role as frontline defenders in the face of unprecedented repression from authoritarian states.

The second issue we seek to highlight is concerned with struggles against racism and settler colonialism that have precipitated a “reckoning” in the NGO sector and its related academic disciplines. Led by the Standing Rock Sioux tribe and Native Nations along with non-Native allies, grassroots protests that began in 2016 against the Dakota Access Pipeline gained extraordinary momentum and drew attention to the pervasive and powerful nexus of settler colonialism, white supremacy and racial capitalism that continues to create disproportionate harm and injustice for indigenous peoples. In 2020, the Black Lives Matter movement further exploded the myth of a post-racial US and has forced a racial reckoning within the NGO sector, including among established donor NGOs in the North. Following the protests instigated by the police murder of George Floyd, there were a flurry of statements that condemned racism and white supremacy within the NGO sector, including public criticisms of internal racism at prominent INGOs and bilateral aid organizations such as USAID, DFID, IRC and Doctors Without Borders (Busby, 2021; De Luce and Williams, 2020; McVeigh, 2020; Syal and Waterson, 2020), often coming from within the internal ranks of their organizations. Peace Direct’s (2021) recent report *Time to Decolonize Aid* also discusses the continued colonial systems, dynamics and epistemologies that continue to structure NGOs and INGOs, both between donors concentrated in the North and Southern NGOs as well as within the working cultures of these organizations under a guise of “neutrality.” As calls to action reverberated throughout the NGO sector and the field of development, important conversations about race, capital and power have resurfaced as foundations and grant-makers have been challenged by movements to interrogate both internal institutional dynamics and how they relate to recipients of aid in the South.

While the recent call for “decolonizing aid” seems to signal a more urgent appeal for progressive reforms within the NGO sector, it is also the evolution of an ongoing conversation around racial and geopolitical inequalities within the sector. In 1985, the charitable Live Aid concerts headed by Bob Geldof and other celebrities to raise funds for those affected by the famine in Ethiopia, set the stage for a culture of charitable giving that implored those in the global North to feel pity for starving children in “Africa” while failing to engage with the structural inequalities that created these conditions (Müller, 2013). In the context of the current moment, several INGO and NGO leaders are seeking to name and acknowledge systemic racism and engage structural and geopolitical

inequalities that shift power and go beyond earlier tokenistic and technical “fixes” that have been prevalent in the sector. The UK INGO Oxfam has sought to reframe the debate on aid in more fundamental ways by shifting the lens from poverty to wealth inequality and racial capitalism, recognizing that poverty cannot be addressed without also attending to how wealth is generated. Citing statistics such as the “world’s 2153 billionaires have more wealth than the 4.6 billion people who make up 60% of the planet’s population” they have begun calling for “deliberate inequality-busting policies” (Oxfam, 2020). Taking the lead from Vil-lanueva’s (2021) book *Decolonizing Wealth*, some foundations have begun to shift their practices to provide operational and unrestricted, multi-year and long-term funds to NGOs and sought to redistribute power and decision making to BIPOC led organizations and communities (DevEx, 2020). Other foundations and grantmakers have also sought to be movement aligned and movement accountable, and BIPOC have called for white leaders to be co-conspirators and shift the white “gaze,” and reconceptualize aid as reparations in the context of historical and ongoing dispossessions, challenging the notion of “helping” that is so constitutive of the philanthropic industry (Decolonizing Wealth Project, 2021).

Laudable as these efforts are, the speed and fluency with which “decolonization” has quickly been taken up by the NGO sector ought to give us pause. As Tuck and Yang (2012) caution, the “easy absorption, adoption, and transposing of decolonization is yet another form of settler appropriation” (p. 3). For the INGOs and aid organizations that represent imperial interests such as USAID and DFID, a genuine “decolonizing” would require thoroughly breaking from the fundamentals of the aid and development industry, including from the military-aid-industrial complex, from neoliberalism and financialized capitalism and from the role NGOs play in enforcing and sustaining deeply rooted structures of domination. Particularly for donor organizations and institutions located in the North, decolonization would require recognizing the connections between settler colonialism “here” and imperialism “elsewhere,” and as such, the *actual* “repatriation of indigenous land and life” (Tuck and Yang, 2012, p. 21). For academic fields and institutions implicated in colonial histories and benefitting from settler colonialism, it entails a “systemic overhaul” and destabilization of existing academic systems as well as reparations, epistemological & ontological disruptions and interrogations of the policies and practices of Northern states. As international development and education scholars have pointed out, fields like international education “are falling short of any genuine decolonization efforts” and questioned whether “decolonization” is “within our grasp” at this juncture (Ghaffar-Kucher, 2019, p. 2). For recipient NGOs in the global south, re-evaluating partnerships and forging ever important alliances with social movements, unions, and activist organizing to simultaneously challenge hegemony of Northern NGOs as well as authoritarian regimes and foster alternative institutional and political possibilities will remain an ongoing task. While continuing the indispensable work of forging solidarities and alliances, it will also be incumbent on those in the sector not to homogenize the specificity of movements and histories, lineages and demands, including demands to recognize the “incommensurability” of decolonization with other movements of social justice (Tuck and Yang, 2012). Bearing in mind, as Ruth Gilmore highlights that the majority of philanthropy stems from “repositories of twice-stolen wealth—(a) profit sheltered from (b) taxes” and in the case of the US, “thrice-stolen, considering the foundation of the American state on the theft of Indigenous lands” (Kanj, 2021), in this moment of reckoning, how would understandings of justice and NGO strategies shift if NGOs were accountable to this reality? Calls to “decolonize” NGOs and aid warrant asking questions with no easy resolutions and contending with the epistemological and material transformations that consequences of decolonization entail. With a deeper structural critique in mind, we ask, what is the NGO sector ready to relinquish in order to break from reproducing the settler colonial project and imperial modes of dispossession and injustice to stand with those whose interests they seek to represent?

We offer these reflections and their implications for NGOs that merit further discussion among scholars and practitioners. As various crises of neoliberalism coalesce and precipitate social and political upheaval, paving the way for authoritarian regimes, NGOs play complex and contradictory roles where they can be further incorporated in service of global capital and the dispossession and expropriation that are constitutive of it. Whether progressive NGOs will be able to maintain “relative autonomy” (Bourdieu, 1983) and be allied with critical struggles and demands for racial and economic justice will depend on whether the sector is able to navigate an increasingly treacherous and challenging environment to enable a transformative politics.

References

- Alvarez, S., 1999. Advocating feminism: the Latin American feminist NGO boom. *Int. Fem. J. Polit.* 1 (2), 181–209.
- Alvarez, S., 2009. Beyond NGO-ization? Reflections from Latin America. *Development* 52 (2), 175–184.
- Amnesty International, 2019. Global Assault on NGOs Reaches Crisis Point. <https://www.amnesty.org/en/latest/news/2019/02/global-assault-on-ngos-reaches-crisis-point/>.
- Arafa, M., 2017. The tale of post-Arab Spring in Egypt: the struggle of civil society against a Janus-faced state. *Indiana Int. Comp. Law Rev.* 27 (1), 43–78.
- Arnove, R. (Ed.), 1982. *Philanthropy and Cultural Imperialism: The Foundations at Home and Abroad*. Indiana University Press.
- Baltodano, M.P., 2017. The power brokers of neoliberalism: philanthrocapitalists and public education. *Pol. Futures Educ. Internet* 15 (2), 141–156.
- Barghouthi, M., 1993. North-South relations and the question of aid. *Dev. Pract.* 3 (3), 204–208.
- Barkan, J., 2011. Got dough? How billionaires rule our schools. In: *Dissent Magazine*.
- Benin, J., 2014. Civil society, NGOs and Egypt’s 2011 popular uprising. *S. Atl. Q.* 113 (2), 396–406.
- Bernal, V., Grewal, I. (Eds.), 2014. *Theorizing NGOs: States, Feminisms and Neoliberalism*. Duke University Press.
- Bhanji, Z., 2008. Transnational corporations in education: filling the governance gap through new social norms and market multilateralism? *Glob. Soc. Educ.* 6 (1), 55–73. <https://doi.org/10.1080/14767720701855618>.
- Bhanji, Z., 2012. Microsoft corporation: a case study of corporate-led PPPs in education. In: *Public Private Partnerships in Education*. Edward Elgar Publishing, Cheltenham, UK.
- Bilmes, J., 2021. Where did the \$5tn spent on Afghanistan and Iraq go? Here’s where. *Guardian*. <https://www.theguardian.com/commentisfree/2021/sep/11/us-afghanistan-iraq-defense-spending>.
- Bishop, M., Green, M., 2008. *Philanthro-capitalism: How the Rich Can Save the World*, first US ed. Bloomsbury Press.

- Bourdieu, P., 1983. The field of cultural production, or: the economic world reversed. *Poetics* 12, 311–356.
- Busby, M., 2021. David Miliband charity pushes “white supremacy culture,” workers allege. *Guardian*. <https://www.theguardian.com/global-development/2021/jun/19/david-miliband-charity-international-rescue-committee-white-supremacy-allegations>.
- Choudry, A., Kapoor, D. (Eds.), 2013. *NGOization: Complicity, Contradictions and Prospects*. Zed Books.
- Cooke, B., Kothari, U. (Eds.), 2001. *Participation: The New Tyranny?*. Zed Books.
- Cornwall, A., 2007. Buzzwords and fuzzwords: deconstructing development discourse. *Development in Practice*, 17:4-5, 471–484.
- Dauvergne, P., LeBaron, G., 2014. *Protest Inc.: The Corporatization of Activism*. Polity Press, Cambridge.
- Davidson, N., Saul, R., 2017. Neoliberalism and the far-right: a contradictory embrace. *Crit. Sociol.* 43 (4–5), 707–724.
- De Luce, D., Williams, A., 2020. USAID Staff Demand Action From Agency Leaders Over “Systemic Racism”. *NBC News*. <https://www.nbcnews.com/politics/politics-news/usa-id-staff-demand-action-agency-leaders-over-systemic-racism-n1230316>.
- Decolonizing Wealth Project, 2021. #PhilanthropySoWhite 2021. <https://www.youtube.com/watch?v=29YBL-6udc0>.
- DevEx, 2020. The Future of Humanitarian Action: Decolonizing Humanitarian Aid. September 14. <https://www.devex.com/news/sponsored/the-future-of-humanitarian-action-decolonizing-humanitarian-aid-98077>.
- Duffield, M., 2012. Challenging environments: danger, resilience and the aid industry. *Secur. Dialogue* 43 (5), 475–492.
- Eade, D., 2004. International NGOs and unions in the South: worlds apart or allies in the struggle? *Dev. Pract.* 14 (1–2), 71–84.
- Edwards, D., Fowler, A., 2002. *The Earthscan Reader on NGO Management*. Routledge.
- Elyachar, J., 2005. *Markets of Dispossession: NGOs, Economic Development, and the State in Cairo*. Duke University Press, Durham, NC.
- Ferguson, J., 2006. *Global Shadows: Africa in the Neoliberal World Order*. Duke University Press.
- Fisher, W.F., 1997. Doing good? The politics and antipolitics of NGO practices. *Annu. Rev. Anthropol.* 26 (1), 439–464.
- Ghaffar-Kucher, A., 2019. CIE methodology and possibilities of other futures. In: *Interrogating and Innovating Comparative and International Education Research*. Brill, pp. 195–205.
- Gramsci, A., 1971. In: Hoare, Q., Nowell Smith, G. (Eds.), *Selections From the Prison Notebooks of Antonio Gramsci*. International Publishers, New York.
- Greenburg, J., 2017. Selling stabilization: anxious practices of militarized development contracting. *Dev. Change* 48 (6), 1262–1286.
- Gurn, A.M., 2016. Courting corporate philanthropy in public education: multi-disciplinary literature review of Public–Private Partnerships (PPPs) in urban public schooling. *Sage Open* 6 (2).
- Hancock, G., 1989. *Lords of Poverty: The Freewheeling Lifestyles, Power, Prestige and Corruption of the Multibillion Dollar Aid Business*. Atlantic Monthly Press.
- Harvey, D., 2006. Neo-liberalism as creative destruction. *Geogr. Ann.* 88 (2), 145–158.
- Hecht, M., 2018. Pakistan expels NGOs in crackdown on dissent. *Nonprofit Q.* <https://nonprofitquarterly.org/pakistan-expels-ngos-in-crackdown-on-dissent/>.
- Hites, S., 2019. Between social duty and the greed of giving: on philanthrocapitalism and gift-patriotism. *Comp. Lit. Stud.* 56 (3), 469–486.
- Human Rights Watch, 2018. Philippines Tars Rights Groups With “Drug Lords” Smear. March 26. <https://www.hrw.org/news/2018/03/26/philippines-tars-rights-groups-drug-lords-smear>.
- Ismail, F., Kamat, S., 2018. The future of NGOs: reinvention, incorporation, critique? *Crit. Sociol.* 44 (4–5), 569–577.
- Jalali, R., 2013. Financing empowerment? How foreign aid to Southern NGOs and social movements undermines grassroots mobilization. *Sociol. Compass* 7 (1), 55–73.
- Jaoul, N., 2018. Politics against the grain: the dalit movement of Uttar Pradesh in the throes of NGOization. *Crit. Sociol.* 44 (4–5), 611–627.
- Kajimbwa, M., 2006. NGOs and their role in the global South. *Int. J. Not Profit Law* 9 (1).
- Kamat, S., 2002. *Development Hegemony: NGOs and the State in India*. Oxford University Press, New Delhi.
- Kamat, S., 2003. The NGO phenomenon and political culture in the third world. *Development* 46 (1), 88–93.
- Kamat, S., 2004. The privatization of public interest: theorizing NGO discourse in a neoliberal era. *Rev. Int. Polit. Econ.* 11 (1), 155–176.
- Kamat, S., 2013. Preface. In: Choudry, A., Kapoor, D. (Eds.), *NGOization Complicity, Contradictions and Prospects*. Zed Books, pp. viii–xii.
- Kamat, S., 2014. New development architecture and postpolitics in the global South. In: Wilson, J., Swynedouw, E. (Eds.), *Spectres of the Political: The Post-political and Its Discontents*. University of Edinburgh Press, Edinburgh, Scotland.
- Kanji, A., 2021. For “Love”: Charity-Washing Colonialism, Fascism and Genocide. *Al Jazeera*. <https://www.aljazeera.com/opinions/2021/3/12/for-love-charity-washing-colonialism-fascism-and-genocide>.
- Karim, L., 2011. *Microfinance and Its Discontents: Women in Debt in Bangladesh*. University of Minnesota Press, Minneapolis, MN.
- Klein, N., 2007. *The Shock Doctrine: The Rise of Disaster Capitalism*. Knopf Canada.
- Kumashiro, K., 2012. When Billionaires Become Educational Experts. *AAUP*. <https://www.aaup.org/article/when-billionaires-become-educational-experts#.YvBpZqZPD4>.
- Layton, L., 2014. How Bill Gates pulled off the swift common core revolution. *Wash. Post*. Retrieved from https://www.washingtonpost.com/politics/how-bill-gates-pulled-off-the-swift-common-core-revolution/2014/06/07/a830e32e-ec34-11e3-9f5c-9075d5508f0a_story.html.
- Leve, L., Karim, L., 2001. Politics of the poor? Grassroots political mobilization and NGOs in Bangladesh. *Polit. Leg. Anthropol. Rev.* 24 (1), 97–107.
- Lewis, D., Kanji, N., 2009. *Non-governmental Organizations and Development*, first ed. Routledge.
- Lewis, D., 2007. *The Management of Non-governmental Development Organizations*, second ed. Routledge.
- Lipman, P., 2015. Capitalizing on crisis: venture philanthropy’s colonial project to remake urban education. *Crit. Stud. Educ.* 56 (2), 241–258.
- Manji, F., O’Coill, C., 2002. The missionary position: NGOs and development in Africa. *Int. Aff.* 78 (3), 567–583.
- Maran, M., 2006. *The Road to Hell: The Ravaging Effects of Foreign Aid and International Charity*. Free Press.
- McGoey, L., 2012. Philanthrocapitalism and its critics. *Poetics* 40 (2), 185–199.
- McGoey, L., 2014. The philanthropic state: market–state hybrids in the philanthrocapitalist turn. *Third World Q.* 35 (1), 109–125.
- McVeigh, K., 2020. Médecins Sans Frontières is “institutionally racist,” say 1,000 insiders. *Guardian*. <https://www.theguardian.com/global-development/2020/jul/10/medecins-sans-frontieres-institutionally-racist-medical-charity-colonialism-white-supremacy-msf>.
- Mellen, R., Lynch, C., 2017. Inside Turkey’s NGO purge. *Foreign Pol.* <https://foreignpolicy.com/2017/08/03/inside-turkeys-ngo-purge/>.
- Mercer, C., 2002. NGOs, civil society and democratization: a critical review of the literature. *Prog. Dev. Stud.* 2 (1), 5–22.
- Mirafteb, F., 2004. Public-private partnerships: the trojan horse of neoliberal development? *J. Plann. Educ. Res.* 24, 89–109.
- Mitchell, K., Lizotte, C., 2014. The grassroots and the gift: moral authority, American philanthropy and activism in education. *Foucault Stud.* 18, 66–89.
- Moeller, K., 2020. Accounting for the corporate: an analytic framework for understanding corporations in education. *Educ. Res.* 49 (4), 232–240.
- Müller, T.R., 2013. The long shadow of band aid humanitarianism: revisiting the dynamics between famine and celebrity. *Third World Q.* 34 (3), 470–484.
- Nagaraj, V.K., 2015. “Beltway bandits” and “poverty barons”: for-profit international development contracting and the military-development assemblage. *Dev. Change* 46 (4), 585–617.
- Olmedo, A., 2019. *Venture Philanthropy and Education Policy-Making: Charity, Profit, and the So-Called “Democratic State”*. The Wiley Handbook of Global Educational Reform, pp. 47–69.
- Oxfam, 2020. World’s billionaires have more wealth than 4.6 billion people. January 20. <https://www.oxfam.org/en/press-releases/worlds-billionaires-have-more-wealth-46-billion-people>.
- Peace Direct, 2021. *Time to Decolonise Aid: Insights and Lessons From a Global Consultation*. Peace Direct, London.
- Petras, J., 1999. NGOs in the service of imperialism. *J. Contemp. Asia* 29 (4), 429–440.
- Reimann, K., 2006. A view from the top: international politics, norms and the worldwide growth of NGOs. *Int. Stud. Q.* 50, 45–67.

- Robertson, S., Mundy, K., Verger, A., Menashy, F. (Eds.), 2012. *Public Private Partnerships in Education: New Actors and Modes of Governance in a Globalizing World*. Edward Elgar Publishing.
- Rohdes, D., 2012. Visit Afghanistan's "little America," and see the folly of for-profit war. *Atlantic*. <https://www.theatlantic.com/international/archive/2012/06/visit-afghanistans-little-america-and-see-the-folly-of-for-profit-war/257962/>.
- Roy, A., 2010. *Poverty Capital: Microfinance and the Making of Development*. Routledge, London.
- Saltman, K.J., 2010. *The Gift of Education: Public Education and Venture Philanthropy* (Education, Politics and Public Life). Palgrave Macmillan, New York, NY.
- Saltman, K.J., 2018. *The Swindle of Innovative Educational Finance*. University of Minnesota Press, Minneapolis.
- Savage, L., 2021. Bill Gates Chooses Corporate Patent Rights Over Human Lives. Retrieved from. <https://jacobinmag.com/2021/04/bill-gates-vaccines-intellectual-property-covid-patents>.
- Schoenefeld, J., 2021. Interest groups, NGOs or civil society organisations? The framing of non-state actors in the EU. *Voluntas* 32, 585–596.
- Schuller, M., 2013. *Killing With Kindness: Haiti, International Aid and NGOs*. Rutgers University Press, New Brunswick, NJ.
- Schuller, M., 2017. Haiti's "Republic of NGOs". *Curr. Hist.* 116 (787), 68–73.
- Scott, J., 2009. The politics of venture philanthropy in charter school policy and advocacy. *Educ. Pol.* 23 (1), 106–136.
- Scott-Villiers, P., 2002. The struggle for organisational change: how the ActionAid accountability, learning and planning system emerged. *Dev. Pract.* 12 (3–4), 424–435.
- Sharma, A., 2008. *Logics of Empowerment: Development, Gender, and Governance in Neoliberal India*. University of Minnesota Press, Minneapolis, MN.
- Sherwood, H., 2015. Human rights groups face global crackdown "not seen in a generation". *Guardian*. <http://www.theguardian.com/law/2015/aug/26/ngos-face-restrictions-laws-human-rights-generation>.
- Silova, I., Steiner-Khamisi, G. (Eds.), 2008. *How NGOs React: Globalization and Education Reform in the Caucasus, Central Asia and Mongolia*. Kumarian Press.
- Silova, I., 2008. Contested alliances: international NGOs and authoritarian governments in the era of globalization. *Curr. Issues Comp. Educ.* 10 (½), 26–31.
- Simmons, P.J., 1998. Learning to live with NGOs. *Foreign Pol.* 112, 82–96.
- Srivastava, P., Oh, S.A., 2012. Private foundations, philanthropy and partnership. In: *Education and Development: Mapping the Terrain*. Public Private Partnerships in Education.
- Srivastava, P., 2010. Public-private partnerships or privatisation? Questioning the state's role in education in India. *Dev. Pract.* 20 (4–5), 540–553.
- Swart, M., 2020. Indian Law on Foreign Funding a "Tool to Silence" Civil Society. *Al Jazeera*. <https://www.aljazeera.com/news/2020/11/13/restrictions-on-funding-of-indian-civil-society>.
- Syal, R., Waterson, J., 2020. Racism endemic at DfID, staff claim. *Guardian*. <https://www.theguardian.com/world/2020/aug/26/racism-endemic-at-dfid-staff-claim>.
- Tarlau, R., Moeller, K., 2019. "Philanthropizing" consent: how a private foundation pushed through national learning standards in Brazil. *J. Educ. Pol.* 35 (3), 337–366.
- Tuck, E., Yang, W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Vannier, C.N., 2010. Audit culture and grassroots participation in rural Haitian development. *PolAR* 33, 282–305.
- Villanueva, E., 2021. *Decolonizing Wealth: Indigenous Wisdom to Heal Divides and Restore Balance*. Berrett-Koehler Publishers, Chicago.
- Waghmore, S., 2012. Beyond depoliticization? Caste, NGOs and Dalit land rights in Maharashtra, India. *Dev. Change* 43 (6), 1313–1336.
- Watts, J., 2019. Jair Bolsonaro Claims NGOs Behind Amazon Forest Fire Surge—But Provides No Evidence. *Guardian*. August 21. <http://www.theguardian.com/world/2019/aug/21/jair-bolsonaro-accuses-ngos-setting-fire-amazon-rainforest>.
- White, S., 1996. Depoliticising development: the uses and abuses of participation. *Dev. Pract.* 6, 6–15.
- Yunus, M., 2007. In: Jolis, A. (Ed.), *Banker to the Poor: Micro-lending and the Battle Against World Poverty*, first ed. 1999. Public Affairs.
- Zencirci, G., Herrold, C., 2022. Project-Think and the fragmentation and defragmentation of civil society in Egypt, Palestine and Turkey. *Non-profit and Voluntary Sector Quarterly*, 51 (3), 545–565.

Education International and its global contribution to educational debates

Jim Baker, Education International, Brussels, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	574
What is Education International?	575
History of EI	575
Education International and multi-lateral institutions	576
Trade unions	577
Trade unions and the teaching profession	577
Global framework of professional teaching standards	579
Issues of educational policy and practice	579
Uses of technology	579
Educational governance	580
Shortage of teachers	580
Teacher well being	580
The future of the teaching profession	581
Inclusion, equity, human rights, and democracy	581
Equality and education	581
Gender	581
Indigenous peoples	582
LGTQ+	582
Violence, harassment, and safe schools	582
Child labor	582
Academic freedom	583
Trade union rights and democracy	583
Education, trade unions, and democracy	583
The pandemic, education, and the social contract	584
References	585

Introduction

The COVID-19 pandemic has not changed the fundamental challenges to education. However, like in other areas, it has aggravated long-standing problems and placed many issues, especially inequality, in relief. The pandemic has deepened existing inequalities, and, in some cases, it has reversed progress. It has, for example, jeopardized important, if inadequate advances in areas like the education of girls and the elimination of child labor. The pandemic and associated crises have also served as an excuse for many governments to push the achievement of the Sustainable Development Goals (SDG) including goal 4 on education, further down the road. Achieving the goal by 2030 would represent a major shift in investment toward the public good and would enable education to come closer to playing its irreplaceable role in equipping young people to cope and thrive in a rapidly changing world.

One of the crises exacerbated by the pandemic is exclusion, whether it is persons with disabilities, refugees and migrants, indigenous peoples, LGBTI persons, religious minorities, or many others. It has boosted an already unacceptable level of violence against women, intensified discrimination, and dampened opportunities. Given these challenges, education needs to address and remedy inequalities in education, but also maximize its equality impact on the larger community. However, the impact of the pandemic goes far beyond equality. The democracy crisis has gotten worse as well. In 2020, Freedom House published a report ([Repucci and Slipowitz, 2020](#)) that showed that in 80% of countries, the situation of human rights and democracy had weakened. That was particularly for freedom of expression and freedom of association. Trade unions, including education unions, have been engaged in the fight for democracy in repressive countries and others where democracy is weak. Many activists have been imprisoned. Some have been killed.

Although this chapter will not dwell in any detail on the pandemic, it is relevant and will remain relevant to ongoing concerns in education long after the virus has subsided. Those lasting impacts will be reflected in the state of the education system. It will include the challenges already mentioned, but also the question of how to deal with the quantum leap in the use of education technology caused by the rapid adoption of distance learning, most often without consultation with teachers and their organizations. One of the challenges will be how to take advantage of technology to reduce workloads and streamline administration rather than using it in pedagogy in ways that would damage the quality of teaching and compromise the autonomy and role of professional teachers.

The lasting impact of the pandemic will depend on whether the many lessons from it have been learned and can improve education systems. That means, among other things, lifting the status of teachers and raising the priority of financing education which has

become urgent after so many years of neglect. This paper provides an account of the ways in which Educational International interprets the deep challenges facing education around the world, seeking to put the voices of the teachers at the heart of the global conversation and advocacy for educational reform. As a non-government organization (NGO), representing the profession, it contests the understanding of education represented by some leading intergovernmental and corporate organizations.

What is Education International?

Education International (EI) is an international professional trade union organization. It is a global union federation (GUF) with 383 member organizations in 178 countries and territories. They, in turn, represent more than 32 million teachers and other education staff. EI is one of the largest of the eight Global Union Federations, which represent workers by sector and occupation. In addition to GUFs, there are two international trade union bodies linked with GUFs that represent multi-sector national centers. They are the International Trade Union Confederation (ITUC) and the Trade Union Advisory Committee to the OECD (TUAC). Their role and membership are different from GUFs. The ITUC and TUAC share common values with the GUFs, but all are autonomous from the others. They cooperate, including in the Council of Global Unions (CGU), but are governed by their member organizations according to their respective Constitutions.

The highest authority in EI is the Congress, which meets every four years. All member organizations are invited to attend. The Congress elects the president, vice-presidents (one from each region), the general secretary and the other members of the Executive Board. The Executive Board governs EI between Congresses. Regional structures and other bodies also contribute to its governance (2021). The general secretary is the only member of the Executive Board who is a full-time employee of the organization. He/she is assisted by secretariat staff at the Brussels headquarters and at regional offices.

To respond to specific needs of education sectors, meetings are held, research performed, and information is circulated, and the organizations' views are advocated in the international community. The bulk of members are in primary and secondary education. They are the best organized and have the highest employment. Other sectors include early childhood education, higher education and research, technical and vocational education and training, and Educational Support Personnel. Democratic procedures are laid out in the Constitution and By-Laws, which are adopted and amended by Congress. There are also By-laws for the regional organizations.

Education International is financed by fees paid by member organizations based on the number of members. Contributions per member are not the same for all organizations, however. There is a scale based on GDP so that fees correspond to ability to pay. Membership fees finance the secretariat and much of EI's work. However, some member organizations make additional contributions for solidarity and capacity building work in to support education trade unions in low-income countries. That is an important part of the work of Education International. This support includes trade union education and other capacity building programmes or solidarity assistance when member unions confront emergency situations. Most of the resources required for capacity building are provided by national education unions in Europe, North America, Australia, Japan, South Africa, and others that maintain their own international solidarity funds or have access to public development resources.

Support is partly financed by EI's solidarity fund to which member organizations contribute on a voluntary basis. This fund also receives a percentage of EI's annual income. There are appeals to member organizations for emergency relief. The EI solidarity fund also provides assistance when member unions and their members are struck by disasters such as armed conflicts, terrorist attacks or natural catastrophes such as earthquakes and floods. One major example is from 2004 when Indonesia, Sri Lanka and Thailand were struck by a tsunami taking the lives of some 2500 educators and destroying many school buildings, Education International provided financial support to thousands of individual teachers and financed the re-construction of schools in Atjeh and Sri Lanka. This effort was also supported financially by Oxfam Netherlands.

History of EI

Although teacher organizations were already created in a few countries in the 19th century, international teacher organizations go back to the first decades of the 20th century (Windmuller et al., 2019). At that time, most of their members were professional associations, not trade unions. However, there were global trade union organizations in other sectors as well as global bodies representing national centers. Their evolving structures and histories are relevant to the development of global trade unionism in education. Of the modern Global Union Federations, the oldest is the International Transport Workers' Federation (ITF), which was founded in 1896. In transportation, global connections of workers are related to the fact that most of its sectors cross borders. The maritime industry was always global. Railways, civil aviation, and road transport are also international.

For teachers and others working in education, there were not those clear global links in the early years. In addition, for many, including governments, trade unions were seen as being for manual workers and not professionals. At that time, some teachers did not consider that it was appropriate for them to be in trade unions. The founding organizations of EI were the International Federation of Free Teacher Unions (IFFTU) and the World Confederation of Organizations of the Teaching Profession (WCOTP), both with roots going as far back as before WWI.

In 1945, the World Federation of Trade Unions (WFTU) was formed by national centers (Windmuller et al., 2019: 72). It created sectoral structures called Trade Union Internationals (TUI), which were subservient to it. When it became clear that the WFTU was, in fact, controlled by the Communist Party of the Soviet Union, the strongest International Trade Secretariat (the name for GUFs

until 2002), the ITF, decided to remain independent and keep control of the organization in the hands of their affiliated unions. Independent and democratic trade union national centers left the WFTU and founded the International Confederation of Free Trade Unions (ICFTU) in 1949.

Prior to the formation of EI in Stockholm in January of 1993, there were four global trade union organizations for educators, the social democratic IFFTU, the Christian World Confederation of Teachers (WCT), the Communist World Federation of Teachers' Unions (FISE—French initials) and the non-aligned WCOTP (Windmuller et al., 2019: 75). IFFTU and the WCOTP agreed that the new international would be composed of free and independent trade unions and associations. All existing member organizations would become members of the new organization and were expected to subscribe to the international's principles and goals. However, EI created an independent Committee of Experts that could examine member organizations to ensure that they met the criteria for membership as well as, where needed, examining trade unions seeking to join the organization. That Committee is unique among the global trade union organizations. Based on the work of that Committee, some of the original EI members were expelled.

With the collapse of the Soviet bloc in 1989–1990, many teacher organizations in the former communist countries became democratic and independent from government and left the FISE. The European members of the WCT joined EI in 2004 followed by the rest of the organization's members (Windmuller et al., 2019: 88). The WCOTP long saw itself as a professional organization and IFFTU was a trade union organization from its origin. However, already some years before the formation of EI, many of the professional organizations at the national level also saw themselves as trade unions. Many trade unions also focused on professional issues and on quality education and equitable school systems. Some of the European teacher organizations became members of both the IFFTU and the WCOTP prior to the merger. This dual function, to defend the wages, conditions, and rights of workers in education as well as the status of the profession and the promotion of quality education, is one of the strengths of EI and something that makes it different from most other GUFs.

Education International and multi-lateral institutions

An important job of Education International is speaking for its member organizations and the people that they represent. It is their global voice. Although member organizations are often involved in international work, it is organized and coordinated by EI. In the United Nations system, EI intervenes with several parts of the UN in addition to many of its specialized agencies. That includes the UN General Assembly and the UN Special Envoy for Global Education. A major focus of that work has been the adoption of the Sustainable Development Goals for 2030, especially SDG4 on education.

Among the UN specialized agencies, there is extensive cooperation and involvement with UNESCO on the full range of education issues for which they are responsible. EI participates in many consultations and Conferences. EI and its predecessor organizations have also worked with the International Bureau of Education, a center of excellence in curriculum and related matters founded in 1925, which is now part of UNESCO (IBE-UNESCO).

EI works closely with the International Labor Organization (ILO) on education sector issues, but also on more general trade union concerns. It uses its complaint mechanisms when there are government attacks on freedom of association and collective bargaining. Another important agency is the Office of the High Commissioner on Human Rights (OHCHR) in its different fora and activities, including its extensive system of special rapporteurs. Among them are the Special Rapporteurs on the Right to Education. Other agencies where there is some contact and cooperation include UNICEF, UNGEI, WHO, UNDP, UNFPA, UNHCR, and UN Women.

There are several non-UN international organizations that are important to the work of Education international too. They include the Organization of Economic Development and Cooperation (OECD). The OECD conducts several extensive surveys and performs research in many areas. They also work closely with education ministries. EI works with the OECD on their full range of education activities. EI chairs the committee that is responsible for education in the Trade Union Advisory Committee of the OECD as well as maintaining its own committee and other direct contacts and cooperation. The OECD and EI co-organize the annual International Summits on the Teaching Profession. In cooperation with ITUC and TUAC EI also contributes to the preparation of the G7 and the G20. EI consults with the Education Ministers of the G20. Another body that is important to education policy and funding is the Global Partnership for education (GPE). GPE works with 76 lower-income countries to help them build stronger, more effective education systems. It is a multi-stakeholder initiative largely funded by governments. EI serves on the Board of Directors and works closely with the secretariat and other partner organizations.

The World Bank is the largest donor to education. Among the intergovernmental bodies. However, it has a long record of supporting privatization of education and of encouraging the involvement of nonprofessional teachers, including volunteers and "informal" teachers. They, often in agreement with the IMF, have not supported adequate funding for public education and have tried to limit salaries for teachers. Even as EI's relationship with World Bank is contested one it meets the Ban regularly, sometimes with the assistance of the ITUC/Global Unions Washington, DC office. It also works with its member organizations on World Bank issues. They, in turn talk with their governments. As an example, the US member organizations and the government with the assistance of Congress were able to get the World Bank to suspend direct or indirect investments in private, for-profit pre-primary, primary, and secondary schools (Edwards, 2020a,b).

As trade agreements have expanded to services, EI and Public Services International (PSI) have worked to ensure that agreements do not open public services to private investment and control or insulate privatized services from becoming publicly controlled or limit the right to regulate of governments. A major trade agreement on services was negotiated before the World Trade Organization

(WTO) was established in 1995 under the General Agreement on Tariffs and Trade (GATT). The General Agreement on Trade in Services (GATS) is far-reaching. While there continue to be concerns about the impact of GATS on education and health care and other issues, however some of the most dangerous provisions that would have undermined public services and government sovereignty were not included. There have been no service agreements negotiated in the WTO, however, there is a “plurilateral” negotiation underway since 2013, the Trade in Services Agreement (TISA). The agreement has been criticized for being negotiated in secret. Opposition has grown and the process may not yield results.

Although the WTO does not set standards on intellectual property, it is responsible for the Trade-Related Aspects of Intellectual Property Rights (TRIPS), which came into effect in 1995. Intellectual property protections became an issue related to the production of antiviral medications to be used against HIV/AIDS. The controversy led to the adoption of the Doha Declaration in 2001 that provides for flexibility in public health crises. The crisis in availability of COVID-19 vaccines led over 100 nations to call for a temporary waiver of intellectual property protections at the WTO, which would allow for low-cost production in developing countries. EI and other Global Unions have been actively supporting the waiver. It is an issue that is likely to continue to arise with other vaccines and life-saving medications. Although the WTO acts to protect intellectual property rights, it does not adopt the standards. They are made by the World Intellectual Property Organization (WIPO). EI has supported education and research exceptions for reasonable use for copyrighted materials in education. That is already provided in national law in many countries, but in others, teachers risk running afoul of the law if they use copyrighted materials in the classroom. With distance learning, copyright restrictions often become even more complicated in education.

EI and its regional organizations also work with regional inter-governmental bodies on education and other issues. Institutional structures are further developed in Europe than in other regions. Relevant bodies include the European Commission, the Council of Europe (which reaches beyond the European Union), the Bologna process in the European Higher Education Area and the Erasmus program, also in higher education.

Trade unions

Like all free and independent trade unions, education unions are an instrument for workers to defend their interests as they define them. They are distinguished from other groups, corporations, or individuals that express views or advocate policy because they are representative and have a legitimacy that only a democratic mandate can provide. Leaders and members of governing structures are elected based on one member, one vote. This gives workers a collective power that they would not have as isolated individuals. That enables them to gain respect and protect their human dignity.

Trade unions also share a central concern for human rights. Although they support the full range of human rights, they focus on “enabling rights”. Such rights include freedom of association, the right to organize, and the right to collective bargaining. They are enabling because they make it possible for workers to obtain other human rights, including health and safety protections, child-labor and forced labor free workplaces, and to confront all kinds of discrimination. Freedom of expression is related to the exercise of those rights and is also an enabling right. The human right to education is an important and relevant enabling right. It opens doors for students and enables them to effectively exercise other rights as well as having better access to opportunities and more fully participate in their communities.

The mandate of Global Union Federations, including EI, comes from their member organizations at national level. That means that the priorities of EI are based on national preoccupations. Education is controlled by authorities at national level (in some case, by regional or local bodies); however, global integration and rapid communication has meant that developments in one country affect others to a larger extent than in the past. Good and bad ideas are more contagious than was once the case. The job of EI is not the same as that of national unions, but EI is an extension of them and carries their priorities into the international arena. A good flow of information and cooperation are important to understand the impact of other global forces on national systems of education.

Long before the establishment of teacher unions, the trade union movement supported free, universal, public education. It was seen as an important means to increase social justice and to reduce the penalties that children may suffer because their parents are poor, are not educated, or do not speak the national language or do not speak it very well. In countries that have a strong tradition of immigration, education has been a key factor in the integration of new populations. In the United States, for example, the first inter-sectoral trade union body was created in 1827 in Philadelphia. One of their first policy positions was on free public education. They declared that they sought, “a system that will fit the children of the poor as well as the rich to become our future legislators, a system that will bring the children of the poor and the rich to mix together as a band of Republican brothers”.

Trade unions and the teaching profession

There are many issues that define the “status” of the teaching profession. They include a common understanding of the purpose of education, professional autonomy, participation of teachers and their organizations in systems of education, and good-quality training in preparation for and during teaching careers. They also encompass fair treatment and salaries and conditions that are equivalent to similarly qualified professionals.

Schleicher (2018: 1), Director for the Directorate for Education and Skills of the OECD, outlined some key elements of being a professional in his foreword to the report for 2018. He wrote:

A 'professional' is someone who is able to use their expertise to make decisions about their work, who has the autonomy to act on those decisions, and who contributes to the body of professional knowledge. Such decisions and actions are grounded in a specialised set of knowledge and skills, stemming from both quality training and from the constant collaboration with peers and other stakeholders. For teachers and school leaders, professionalism means not only conducting their work in an effective manner, but working to improve their skills, collaborating with colleagues and parents, and thinking creatively about the challenges they face.

In 1966, the international Labor Organization (ILO) and UNESCO (1966) adopted the *Recommendation concerning the Status of Teachers* that covers primary and secondary school teachers. Some three decades later, UNESCO (1997) adopted the *Recommendation concerning the Status of Higher-Education Teaching Personnel*. The Recommendations are similar, and their application by national governments is overseen by a group of experts, the Joint ILO/UNESCO Committee of Experts on the Application of the Recommendations concerning Teaching Personnel (CEART). A publication covering both Recommendations is available with a Users' Guide (ILO, 2008).

An important contribution to the drafting of the 1966 Recommendations was made by William George Carr, the Executive Director of the US National Education Association (NEA), which was in the WCOTP at the time and is now an EI member organization. For the 1997 higher education recommendation, the lead was taken by Donald Savage, who was Executive Director of the Canadian Association of University Teachers (CAUT) working with other university associations and unions who all became members of EI.

The development of these two fundamental recommendations also demonstrates that EI is not just a relatively small secretariat under a General Secretary. It is also the leadership, talent and commitment of member organizations that make things happen at the global level. The two Recommendations do not cover early childhood education; however, their approach and contents influenced the development in 2013 of ILO Policy Guidelines on the promotion of decent work for early childhood education personnel. EI made an important contribution to the Guidelines and served on the group of experts that developed them.

In far too many countries, early childhood education (ECE) is considered to be childcare and educators are not able to acquire the necessary qualifications to enable them to fully contribute to the educational development of children prior to their entry into kindergarten and primary school. ECE teachers are disproportionately women and often have low wages and poor working conditions. EI and its member organizations have been using the ILO Guidelines to make the case to elevate the status of ECE educators. The Guidelines also help to argue that there should be public provision and regulation of ECE.

In the 1966 Recommendation (page 5), the rights and responsibilities of teachers are placed in the context of the mission and purpose of education. The first paragraph of the "Guiding Principles" in the Recommendations, based on the UN Covenant on Economic, Social, and Cultural Rights, reads.

Education from the earliest school years should be directed to the all-round development of the human personality and to the spiritual, moral, social, cultural and economic progress of the community, as well as to the inculcation of deep respect for human rights and fundamental freedoms; within the framework of these values the utmost importance should be attached to the contribution to be made by education to peace and to understanding, tolerance, and friendship among all nations and among racial or religious groups.

This broad definition of education, based on development of the purpose and meaning of education is the basis of the remaining provisions of the Recommendation that seek to ensure that teachers have the qualifications and latitude to help their students develop, not just to find their places in the job market, but also to live rewarding, rich lives, contribute to their communities, and be informed, responsible, and active citizens.

Education International, as part of its work since its inception, seeks education that delivers for "for the all-round development of the human personality" and society, and has resisted narrowing the scope of education to service to the economy. Focusing on narrow criteria for student success, excessive reliance on standardized tests, and constraining the role of teachers in responding to the needs of students and their communities, where that is the case, has diminished the role of education in building healthy and happy students and coherent and tolerant societies that accept diversity and build social peace.

EI promotes the two Recommendations with member organizations so that they can address professional issues with their governments based on global consensus recommendations. EI also does a regular survey of its member organizations on status issues which is submitted, with a separate set of views and recommendations to CEART at its meetings every three years. Examples of the survey are available from 2015 (EI, 2015), and also from 2018 and 2021. The most comprehensive EI submission to CEART, in addition to the surveys, was in 2015. It places modern issues in the context of the principles of the recommendations.

An example of something that emerged long after 1966 was the use of ICT and artificial intelligence in the classroom. However, the Recommendation states:

Since teachers are particularly qualified to judge the teaching aids and methods most suitable for their pupils, they should be given the essential role in the choice and adaptation of teaching material, the selection of textbooks, and the application of teaching methods, within the framework of approved programmes, and with the assistance of the educational authorities." It goes on to say, "Teachers and their organizations should participate in the development of new courses, textbooks, and teaching aid.

Global framework of professional teaching standards

In many countries, professional teaching standards have been adopted. They are often, but not always, developed with teacher organizations. In recent decades, there has been a tendency to judge performance of students and teachers in a way that limits rather than expands the impact of the teaching profession. Many of these “reforms” have been picked up at the international level. However, they are systems of quantification and standardization rather than professional standards.

Real and workable professional teaching standards need to reflect the experience and needs of the profession. The “standards light” approach has too often left teachers out of the picture. The teaching profession should play a leading role in the design and implementation of the policies and practices that will produce a school environment that is conducive to quality teaching and learning. EI took the initiative to create a global framework for such standards.

They were developed through a process of gathering information on national standards and practices and extensive consultations with member organizations. As part of that effort, a Unite for Quality Education and Leadership Conference was held in Rotterdam at the beginning of May of 2017 with representatives of member organizations and experts. This process led to drafting of standards with UNESCO. These standards, the Global Framework of Professional Teaching Standards (UNESCO, 2017) were agreed at the EI Eighth Congress in Bangkok in 2019. They have become the basis for discussion of national standards in many countries.

Given the growing internationalization of the discussions on the teaching profession, EI has emphasized many professional issues at the global level. An important focus on the profession has been annual International Summit on the Teaching Profession International Summit on the Teaching Profession (ISTP) held with the OECD since 2010 (Schliecher and Edwards, 2020). The International Summits, co-hosted by the OECD and EI, bring together Education Ministers with EI member organizations. ISTP meetings have changed the international dialog, have affected the work of the OECD, and have highlighted issues too long ignored, including the well-being of teachers and students.

In addition to generating an international discussion on the teaching profession, the process often enhances national social dialog. There are national discussions on issues considered by the summit and an attempt is made to reach a global consensus on priority issues. In some cases, this consultation process has helped to restore dialog between teacher unions and governments. A range of issues that have been explored at ISTP. Described below is the approach taken by EI, on various long-standing as well as some emerging issues.

Issues of educational policy and practice

Uses of technology

The question of the introduction of new technology and its positive and negative effects on the teaching profession and the status of teachers was already a concern before the pandemic, however, the widespread use of distance learning, especially in developed countries, accelerated due to the need to continue education to the greatest degree possible during lockdowns. However, consultation with teachers and their organizations on this question was exceptional.

The introduction and use of education technology or “edtech” have raised several issues relevant to teachers and their organizations. Among them are:

- The widespread exclusion of teachers and their organizations from discussions and decisions related to the use of learning technology and their effects on education, on students, and on the profession of teaching.
- The involvement, on a massive scale, of private corporations in the development and use of teaching materials and methods.
- The weakening of the responsibility of the State in the design and delivery of education services and, by extension, of the community and public governance.

The key to the introduction and use of new teaching technologies is participation of teachers through their organizations. For example, teachers are often concerned about excessive workload. Much of that burden is not directly related to teaching but is on administrative tasks. Technology might make it possible to reduce that burden. By shifting some administrative tasks to technology. Social dialog can help to determine where the use of new technologies can enhance the profession and where it diminishes it and the status of teachers and, therefore, undermines the quality of education.

There has already been bad experience with the development and use of IT-dependent teaching or measurement methods like standardized tests, most often developed by large corporations with little or no involvement of teacher organizations. That has also been true of assessment tools and other methods used nationally or internationally. These processes have contributed to de-professionalization. There is also a danger that new technology might be introduced as a way of reducing teaching and support staff even if it damages the quality of education. Again, the pandemic underlined the importance of social relationships in effective and good quality education.

The expanded participation of private firms, which may become a permanent feature of education, brings the danger that the vision of education as a market opportunity that has characterized school privatizations will extend to public schools, which will, in effect, become semi-privatized even if they remain, technically, public. As major source of profits of many of the high-tech companies come from the collection use and sale of data rather than from education services provided, there are important privacy issues as well. Research performed for EI has looked at edtech in primary and secondary education (Williamson and Hogan, 2020) as well as in higher education (Williamson and Hogan, 2021). That research was difficult not only because of the complexity

of the issues, but also because of lack of transparency, particularly on the private side. Understanding processes requires access to information, much of which takes the form of protected intellectual property.

Educational governance

Governance issues are relevant to the joint public-private role in education policy and delivery that comes with edtech (Edwards, 2021). One of the advantages of public schools is that transparent governance provides the possibility of public control of education and helps to ensure that it is rooted in the community. In democracies, making education a private as well as a public good, may put education beyond the reach of citizens.

Classic privatization remains an issue and there are still global companies trying to tap into public resources to make money by taking over the operation of schools. However, the edtech danger shows that there are also hidden forms of privatization or semi-privatization. Even if schools remain fully public, they may create competing schools with charter schemes tend to increase segregation. New Public Management, without necessarily any private investment, corrupts public education with market thinking and places distance between the practice of education and its mission. It may lead to segregation that public schools were intended to combat (Edwards, 2019a,b).

Shortage of teachers

The status of teachers is also a key factor in the shortage of teachers. Talented individuals may avoid careers in teaching because they see limited possibilities to develop as professionals and can be appreciated and better compensated in other lines of work. Far too many teachers are leaving the profession early, frustrated, stressed, and, in some cases “burned out”. There is also dissatisfaction due to inadequate training and preparation. A key factor in developing creative approaches to teaching and exploring pedagogy, but also substance is through initial teacher training, continuous, and in-service training. Improving teaching is also helped by good collaboration with other teachers. All these factors affect the attractiveness of the profession.

The EI 2018 survey report highlights the gross inadequacy of teacher training. EI member organizations reported that only 30% of respondents indicate that there is access to continuous professional development. They also reported that training was of poor quality and that teachers often had to pay for it and, in many cases, pursue training in their own time. This is an additional burden especially for teachers who are often underpaid and already have a heavy workload. The 2021 survey report confirms long-term concerns about continuing professional training in terms of quality, availability, and cost. The report states, “In too many cases, CPD was perceived to be of poor quality, not directly relevant to the issues that teachers were facing and came at personal financial cost without clear career benefits.”

In 2016, not long after the adoption of the sustainable development goals, UNESCO estimated that the world would need 69 million new teachers if it was to meet the SDG4 on education by 2030. The Organization of Economic Development and Cooperation (OECD) speaks of a “deep teacher-recruitment crisis”. A 2020 issue of OECD Observer says:

Education is the cornerstone of the knowledge-based society. But will quality teachers be available to provide it, or is the profession cracking under the strain of low salaries, an ageing workforce and demand for ever more complex teaching abilities?

If the teaching profession is to become attractive enough that young people want to join it and stay and so that experienced people do not leave early, major shifts in funding and support for teachers and their profession are required.

Teacher well being

In some countries, especially those employing “reforms” that include narrowing of subjects and promote evaluation based on simple factors like test scores have increased stress for teachers. However, they have also created a school environment that is less friendly to learning. Although most teachers are still satisfied with their profession and do not report debilitating stress to the OECD TALIS Survey (2018), they report that “acute stress at work is also strongly associated with teachers’ job satisfaction and their intention to continue teaching: 18% of teachers report experiencing a lot of stress in their work, and 49% report that having too much administrative work is one of the main sources of stress.”

Among EI member organizations, stress is reported especially from teachers in places that have pushed furthest with “reforms” including England, the US and Japan. The school is a community. It is a world of relationships among teachers, other school staff, administrators, students, and parents. Through those relationships, stress is easily transmitted. So is comfort and confidence. A school community functions best if it is stable healthy, and safe. School and education are part of life not preparation for life. The more that education is channeled into narrow paths that may be intended to serve the economy, but do not serve other human needs, the lower the quality will be and the less it will inspire young people and will, among other things, do a disservice to industry and the economy.

The future of the teaching profession

There has been a global discussion about the future of work in general. Much of it has been around the impact of technology on employment. It is a subject of interest for teachers and others in education because of the impact on teaching, but also because education will be a critical factor in adapting to changes in other sectors. EI has been working on the general challenge as well as on education specific issues. A resolution was adopted by the EI 8th World Congress in Bangkok, 2019 on the future of the Teaching profession (EI, 2019d). Based on that resolution, a Future of Work Reference Group was established composed of representatives of member organizations and experts.

The future, including the status of the teaching profession will depend on the direction that the evolution of education takes. It is a discussion leading to policy that is too important for teachers and their organizations to be excluded. They should play a role in shaping the choices related to the use of new technologies. However, the survey conducted for the reference group (Coldclough, 2020) showed a striking lack of consultation by education authorities. 45% of respondents said that they were not consulted at all and 29% said that they were only consulted “on a few aspects”.

The International Labor Organization established a Commission on the Future of Work. Its report called for a fundamentally different approach to technological and other changes. Rather than scrambling to adapt to whatever technological changes arrive, they called for a human-centered and long-term approach that takes advantage of a period of transition to try to design into the change process inclusion and social justice. EI welcomed the work of the Commission and its far-reaching recommendations (Edwards, 2019a,b). Given the speed of changes in the world of work, including in education, and the global nature of many of these changes, the efforts of EI and other Global Union Federations are the beginning of a much longer evolution. They will likely become a permanent, global trade union focus.

Inclusion, equity, human rights, and democracy

Although the situation varies by country, many have experienced a concentration of wealth and income. Even some countries that have made great progress against poverty are seeing unfair distribution. Some countries have growing numbers of “working poor”, often those in precarious and insecure work.

In its 2020 World Social Report published by UN Department of Economic and Social Affairs (DESA, 2020), it is reported that inequality is growing for over 70% of the world’s population. This is true in developed as well as developing countries. Inequality destabilizes societies, leading to despair and conflict, creating a climate of polarization and radicalization, and undermining democracy.

Equality and education

Education is part of the larger community and is affected by inequalities in society. Part of its mission is to help those who are disadvantaged in different ways. To the extent possible, inequities should be eliminated in systems of education. Economic inequality, by itself, may cause social, family, and individual problems, however, it is often compounded by other factors such as gender, ethnicity (especially for indigenous peoples), national origin, caste, religion, migrant, or refugee status (Edwards, 2020a,b), disability, or sexual orientation. In some cases, discrimination becomes more severe when there is a combination of characteristics (intersectionality). To help to address equity issues in education in a systematic and organized way, EI is working with member organizations on developing equity audits (EI, 2020) as tools to identify and remedy injustices.

Discrimination and exclusion are not new. However, they have been aggravated by politicians and extremist groups that generate hatred and disinformation. There are many examples from the past, some of which have led to violence and even genocide. These tendencies are often found in schools. At the same time, schools are among the rare places in the community where teaching materials and methods can reduce tensions and build understanding.

Gender

There has been work on how girls and women are viewed in curriculum. It is not just how they are treated in society, but how negative images may be perpetuated. In addition to dealing with equity issues in the context of education, EI addresses inclusion issues in systems of education. Often, women are heavily concentrated in early childhood education and primary schools and least present in higher education, research and TVET. EI has also promoted equity and inclusion in education trade unions. It has promoted greater women’s participation in trade union leadership. That includes through work in the World Women’s Conference and women’s networks.

Another priority for EI is the education of girls. In many countries, there has been a pattern of boys going to school and girls staying home to do chores or be child laborers. Or girls might be allowed to attend primary, but not secondary schools. Although there has been some progress in recent decades, every time there is an economic crisis, for example, with the impact of the pandemic on the economy, girls’ education suffers. SDG 4.5 commits to eliminating gender disparities in education by 2030. That is likely to be an unmet goal. Long and short term, changing the situation will require serious investment as well as further changes in attitudes.

Indigenous peoples

There are education issues like access and quality, but including indigenous history and culture in education as well as protection of the right of indigenous communities (and others) to have education in their own languages. Indigenous peoples not only report discrimination, but harassment and, in some cases, violence from surrounding communities and governments. There is an indigenous people's caucus at every Congress (EI, 2019a), which examines issues and makes recommendations.

The attitude toward indigenous peoples also reflects history that is often written about them, but without them. Such attitudes toward indigenous peoples or others are found in curriculum, teaching materials and approaches to issues. They perpetuate versions of history written by the most powerful. Corrective efforts are addressed in what is called, "decolonizing education" (EI, 2019b).

LGBTQ+

Another area where EI has played a pioneering role in the global trade union movement is on justice for LGBTQ+ people. It is more than just fighting discrimination, which in some countries is extreme and includes legal and police harassment and violence against LGBTQ+ people, especially teachers. EI and its regional organizations are working with member organizations to try to ensure that LGBTQ+ people feel welcome in their trade unions and in leadership and do not suffer from any forms of discrimination. There is an LGBTQ+ caucus (EI, 2019c) meeting held prior to each EI Congress.

Violence, harassment, and safe schools

EI has been preoccupied with violence and harassment in schools. Much of the focus has been on women and girls, which has been an especially serious problem. EI has issued a resolution on eliminating all forms of sexual harassment and Sexual Violence in Education Unions. However, over the years, the situation seems to have deteriorated. This reflects how common it has become, even though it is still under-reported. Efforts by EI have included an extensive program of training on School Based Gender Related Violence in cooperation with the United Nations Girls' Education Initiative (UNGEI).

However, it is also a problem for adults, including teachers and other education staff. That is one reason that EI, in cooperation with other Global Unions, campaigned for ILO Convention 190, which was adopted in 2019 and entered into force in 2021. C190, which applies to women and men, is a far-reaching Convention that has a reach beyond most ILO Conventions. It applies not only to employers, but to sub-contractors or others performing services as well as third parties, which might include, for example, customers in private companies or parents or other third parties in education.

The concept of "safe schools" is broad. It includes occupation of attacks on schools by military forces of governments or other armed forces. EI has been supporting government commitments to the Safe Schools Declaration (GCPEA, 2021) an intergovernmental political commitment to protect students, teachers, schools, and universities from the worst effects of armed conflict. Terrorists have attacked schools. Some have targeted teachers, especially women teachers. In some countries, especially the United States, schools have been attacked by armed individuals or groups, often using automatic weapons.

Quality education cannot take place in situations where students are afraid of being in schools or going to and from them. This applies to the danger of violence, but also harassment, including harassment or attacks on them because of what they are (e.g., LGBTQ+) or what they say. The school community should be a safe place to learn, discuss, and disagree without danger, fear, or stress. In some cases, it can serve as a safe place for students who are subject to dangers at home or in their neighborhoods.

The human rights of children are elaborated in the UN Convention on the Rights of the Child adopted in 1989 (UN, 1989) and entering into force in 1990. It mentions education, but also a wide range of other rights that are relevant to educators and education systems. Unfortunately, the Convention is often violated as children are too often subject to abuse of their fundamental rights. EI has explored ways of preventing such violation.

Child labor

Child labor was the scourge of the industrial revolution in the 19th century. One of the earliest targets of social legislation and of the early trade unions, was the banning of child labor. The ILO adopted Convention 138—Minimum Age Convention, 1973 (No. 138) as well as C182—Worst Forms of Child Labor Convention, 1999 (No. 182). However, child labor remains a serious problem. The ILO estimates that 160 million children were victims of child labor in 2020. EI has been involved in the global fight against child labor including, for example, through joining with Nobel Peace Prize winner Kailash Satyarthi in his efforts to call attention to child labor and eliminate it as well as cooperating with the ILO. EI has concentrated on the impact of child labor on education and has worked with its member organizations to get children out of the workplace and in schools. Unfortunately, progress quickly erodes in times of economic crisis as parents take their children out of school so that they can help support the family. Such a setback was experienced during the COVID-19 pandemic.

EI operates child labor programs with member organizations in Africa, Asia, Europe, and Latin America. The emphasis has been on getting children into school, creating child labor free zones, and implicating teachers, parents, and communities in the effort to ensure that children have a future through education rather than being exploited as workers and denied the possibility to develop. The success of the efforts has depended on trade unions working with and training teachers to persuade parents to put their children into school as well as working with other groups in communities.

Academic freedom

Academic freedom is, at the same time, a professional, a human right, and a democracy issue. In higher education, the professional right is, in effect, weakened by the growth in precarious work for teachers and researchers and by a weakening of systems of governance involving teaching personnel, which, in some cases, may also be related to hesitation to express one's views. Being assertive on academic freedom issues may mean that contracts are not renewed. In other words, employment security with decent conditions, collegiality, good governance, and academic freedom are all linked.

In many countries, there has been an alarming reduction in public funding for independent research. That can make universities dependent on private funds where private parties will determine what is researched. That can mean that universities and research institutions are not able to provide sufficient, independent research to help inform political discussion and decisions on public policy.

A combination of these factors may help to explain why the Academic Freedom Index (EI, 2021b) shows significant slippage in the respect for academic freedom in recent years. Examples of problems with academic freedom on all Continents were presented in a virtual conference in 2021. Academic freedom is a concern in lower grades, in secondary and primary schools. As teachers are examining issues that are relevant to public debate, many governments prefer indoctrination to education. However, the danger is not limited to governments. It can also come from active groups in the community, often extreme, that seek to control the content of teaching right down to the words used by teachers in the classroom.

Political issues may even include the communication of scientific facts. The teaching of evolution is still banned in some countries. Similarly, climate change deniers, sometimes with the backing of fossil fuel companies, may put pressure on schools and on politicians responsible for education to alter the facts or to present a "balanced" picture where both facts and alternative facts are taught. EI has long been worked on climate change issues. It is engaged in a major effort to ensure that teaching climate change (EI, 2021a) is part of school curriculum and that teachers are free and trained to teach the issues. However, it is more than building understanding of the issues. It is also encouraging young people to become active citizens, to organize and participate in the debate so as to influence government decisions.

Trade union rights and democracy

Teachers are often the victims of violation of their trade union rights. This may take the form of assassinations, imprisonment, torture, or they may be denied their rights in situations where fear is systematically organized. In many countries, they do not have the right to form independent trade unions, often due to larger bans on organizing by public employees. Even where they have the right to organize, they may be denied the right to collective bargaining or may be subjected to limits on the subjects that can be negotiated, referred to as the "scope of bargaining".

Teachers are targets for attacks. Just as with journalists, it is not only related to their trade union activity, but also the sensitivity of their profession for governments. Both journalists and teachers have a public role. Many governments would prefer to have teachers function as agents for their political agendas rather than as educators. Another reason may be that teacher unions, where they are allowed to exist, are often among the largest unions in a country and have members in both urban and rural areas. In many countries, they are also respected members of the community and considered to be leaders. Defending the rights of teachers and other education staff to bargain is a central priority of EI. That defense is not only in the form of support in the country, but also mobilizing support among member organizations, including in the form of urgent action appeals and solidarity appeals on the internet, and using international complaint procedures.

The most effective international organization for the defense of trade union rights is the International Labor Organization (ILO). It has adopted the crucial conventions on trade union rights, including C87 Freedom of Association and Protection, 1948, C98 Right to Organize and Collective Bargaining Convention 1940 and C151 Labor Relations (Public Service) Convention 1978. If the rights under these conventions are not protected by governments, regardless of whether they have been ratified, complaints can be made by EI to the Committee of Freedom of Association (International Labor Organization, n.d.). In addition, complaints on ratified conventions can be made to the on the Committee of Experts on the application of Application of Conventions and Recommendations. A selection is made of decisions by the Experts to be reviewed by the Committee on the Application of Standards (CAS) of the International Labor Conference. At the meeting of that committee, governments cited by the Experts are summoned to appear and questioned on their conduct. It is the only place in the world where governments are called upon to publicly respond to allegations of violation of those fundamental human rights.

Education, trade unions, and democracy

Human rights, including trade union rights, cannot be effectively protected in the absence of democracy. The struggle between totalitarianism and authoritarianism on the one hand and democracy on the other defined much of the history of the 20th century. The fall of fascism and of Communism in the Soviet bloc generated a wave of democracy on all continents. Trade unions and education are among the bulwarks of democracy. In the first decades of the 21st century, disturbing signs emerged of the weakening of democracy even in some long-standing, stable democracies. As trade unions, and collective bargaining weakened in most countries, democracy suffered.

In addition to being democratic forces, trade unions are democracies inside democracies that help deepen it beyond free elections. As, the focus of education in many countries, turned to test scores and other measurable achievements, there was less attention devoted to building the competencies of democracy of young people. Principal among those competencies is critical thinking, the ability to discuss and argue, and active citizenship.

These problems were aggravated by the development of social networks combined with the weakening of the media, especially ethical media. These networks enabled disinformation as well as hate and bigotry to be instantly transmitted around the world. They fueled polarization and radicalization where opponents became enemies. This massive disinformation, in some cases, led to violence. That added another necessary skill. People needed to be equipped to navigate the tsunami of information, develop the ability to be skeptical and to filter out, to the extent possible, lies and distortions.

The concern about the erosion of democracy among education trade unions led to a major discussion of democracy at the 8th World Congress of EI, held in Bangkok in July of 2019. EI published a book, "Education and Democracy: 25 Lessons from the Teaching Profession" (Hopgood and van Leeuwen, 2019) that was launched at the Congress and introduced by EI General Secretary Emeritus Fred van Leeuwen (2019) who stressed in his remarks that education systems have an important responsibility to protect and promote democracy. The book was available in English, French, and Spanish and has, subsequently been translated into Polish, German, Czech, Russian, and Dutch. It is without copyright and can be used, in part or in its entirety, in classrooms and in internal and public discussions.

In addition to trying to help stabilize democracies where they have been historically strong, trade unions are often in the front lines of fights where dictatorships have emerged or consolidated and space for democratic organization may be limited. The fight for democracy, for education and trade unionism is one fight. As expressed by David Edwards on the International Day for Democracy, "the fight for democracy is a war of ideas and we, as educators and trade unionists are on the front lines of that war. The time for complacency, if there ever was one, is over".

The pandemic, education, and the social contract

COVID-19 has killed people, devastated families, and created panic and desperation. It has also produced or deepened other crises. Responses have gone from generosity and assertion of community and caring on the one hand to meanness, bigotry, and violence on the other. The pandemic has also shown that global problems need global solutions but, at the same time, it has been used to promote nationalist populism. The conflict of these conflicting visions is global. Part of the job of EI is to ensure that issues and challenges that are global receive global as well as national responses.

Given the contrasting lessons of the pandemic, there is a stark choice to make to shape the direction of the future. Will it be a future of isolated screen-bound individuals, or will it be social beings embracing community, cooperation, and collective action? Will people be trapped in nationalism and fear of the unknown or will they reach out to, engage with, and learn from others? Will it be solidarity or selfishness?

Trade unions are collective. In fact, all democracies, by their very nature, are collective. It is in that secure environment that individual rights and diversity can flourish. Atomization can never effect change. For society to progress it must, first, exist. Schools are a microcosm of the community. Some of the lessons of shutdowns and remote learning may apply to larger society. Children, away from school, were lonely. Mental illness and suicides went up. Violence in homes flared, and the development of healthy personalities was retarded.

When children returned to school, in most cases, laughter came back as did the joy of learning. People have re-discovered that social relationships facilitate learning and development. The social contract is the context for the opportunity to change the way we live and make societies and schools more human centered. Children can be helped to become active citizens rather than just passive consumers. A social contract is about human beings not things. Education is such a contract which engages students, teachers, parents, and the community. It is built around common values. That is the point of reference from which "reforms" and changes in social relationships and teaching methods should be judged.

UNESCO, in its 2021 Futures of Education report, explains, "A social contract is 'an implicit agreement among members of a society to cooperate for shared benefit. A social contract is more than a transaction as it reflects norms, commitments and principles that are formally legislated as well as culturally embedded. The starting point is a shared vision of the public purposes of education'" (UNESCO, 2021: 2).

Education International is working to fundamentally change the philosophical environment of education so that it responds better to the aspirations of young people, the needs of society, and the common good. That means not only re-examining the learning environment that is created, but also the contribution made by education to social justice, integration, and tolerance. It may contribute to the economy, but it should also make it easier for people to live happy, full lives and to build solid, resilient, and sustainable democracies with social cohesion.

It also means rejecting so-called reforms of recent decades that have damaged the quality of education and the lives of students and teachers and restoring the joy of learning. The social body of schools has rejected the forced transplant of market dogma into school management and the competition and extreme individualism that comes with it. American writer Henry David Thoreau wrote, "the mass of men lead lives of quiet desperation". That does not have to be the case. Children should not leave school, as too many do, with feelings of fear and panic, beaten down when they should be spreading their wings. It should be recognized that school is part of life and not a boot camp to prepare for life. An education social contract can help young people lead whole, satisfying lives of joyful aspirations where hope vanquishes fear and where success is not judged by the vastness of one's investment portfolio, but by the reach of one's heart.

References

- Colclough, 2020. Teaching With Tech: The Role of Education Unions in Shaping the Future. https://eiwebsite.blob.core.windows.net/uploads/20201119_204902_2020_EI_Research_TeachingWithTech_ENG_final.pdf?sv=2019-10-10&ss=b&srt=o&sp=rdx&se=2030-05-26T22:00:00Z&st=2020-05-26T14:11:47Z&spr=https,http&sig=fqlBEld9c06/PzqL90FD54UfvT33KDBvH/hM9wsvLA%3D.
- DESA (Department of Economic and Social Affairs), 2020. Inequality in a Rapidly Changing World. United Nations. <https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/02/World-Social-Report2020-FullReport.pdf>.
- Education International, 2015. Education Unions Voices on the Status of Teachers and the Teaching-Profession. <https://www.ei-ie.org/en/item/21308:education-unions-voices-on-the-status-of-teachers-and-the-teaching-profession>.
- Education International, 2019a. Indigenous Peoples' Caucus: "Our Languages Haven't Died, They Are Asleep and We Need to Reawaken Them." <https://www.ei-ie.org/en/item/22925:indigenous-peoples-caucus-our-languages-havent-died-they-are-asleep-and-we-need-to-reawaken-them>.
- Education International, 2019b. LGBTI Caucus Examines Victories & Setbacks, Mobilisation Strategies to Advance Human Rights. <https://www.ei-ie.org/en/item/22919:lgbti-caucus-examines-victories-setbacks-mobilisation-strategies-to-advance-human-rights>.
- Education International, 2019c. Resolution On: Decolonising Education. <https://www.ei-ie.org/en/item/22951:resolution-on-decolonising-education>.
- Education International, 2019d. Resolution On: The Future of the Teaching Profession. <https://www.ei-ie.org/en/item/23043:resolution-on-the-future-of-the-teaching-profession>.
- Education International, 2020. Auditing Equity in Light of the COVID-19 Pandemic. https://eiwebsite.blob.core.windows.net/uploads/20201217_101520_EquityAuditWebinar_Report_EN.pdf?sv=2019-10-10&ss=b&srt=o&sp=rdx&se=2030-05-26T22:00:00Z&st=2020-05-26T14:11:47Z&spr=https,http&sig=fqlBEld9c06/PzqL90FD54UfvT33KDBvH/hM9wsvLA%3D.
- Education International, 2021a. Education International Manifesto on Quality Climate Change Education for All. <https://www.ei-ie.org/en/item/24244:education-international-manifesto-on-quality-climate-change-education-for-all>.
- Education International, 2021b. Global Index Finds Most Countries Do Not Respect Academic Freedom and Shows Signs of Decline. <https://www.ei-ie.org/en/item/24856:global-index-finds-most-countries-do-not-respect-academic-freedom-and-shows-signs-of-decline>.
- Edwards, D., 2019a. On Life and the Market. <https://www.ei-ie.org/en/item/23180:on-life-and-the-market-by-david-edwards>.
- Edwards, D., 2019b. Work With a Human Face. <https://www.ei-ie.org/en/item/22722:work-with-a-human-face-by-david-edwards>.
- Edwards, D., 2020a. Global Refugee Forum: Time for Immediate and Sustainable Action. <https://www.ei-ie.org/en/item/23177:global-refugee-forum-time-for-immediate-and-sustainable-action-by-david-edwards>.
- Edwards, D., 2020b. World Bank Shifts Position on Funding for Most Private-for-Profit Schools. <https://www.ei-ie.org/en/item/23306:world-bank-shifts-position-on-funding-for-most-private-for-profit-schools-by-david-edwards>.
- Edwards, D., 2021. On Ed-Tech, the Public Good and Democracy. <https://www.ei-ie.org/en/item/24809:on-edtech-the-public-good-and-democracy>.
- GCPEA, 2021. The Safe Schools Declaration: An Inter-governmental Political Commitment to Protect Students, Teachers, Schools, and Universities From the Worst Effects of Armed Conflict. <https://ssd.protectingeducation.org/>.
- Hopgood, S., van Leeuwen, F., 2019. On Education & Democracy—25 Lessons From the Teaching Profession, Written With Jim Baker, Felisa Tibbitts and Jelmer Evers. Creative Commons. https://issuu.com/educationinternational/docs/eiwc8_oneducationanddemocracy.
- International Labour Organization, 2008. The ILO/UNESCO Recommendation Concerning the Status of Teachers (1966) and the UNESCO Recommendation Concerning the Status of Higher-Education Teaching Personnel (1997) With a User's Guide.
- International Labour Organization, (n.d.). The Committee on Freedom of Association. <https://www.ilo.org/global/standards/applying-and-promoting-international-labour-standards/committee-on-freedom-of-association/lang-en/index.htm>.
- International Labour Organization, 2013. ILO Policy Guidelines on the Promotion of Decent Work for Early Childhood Education Personnel.
- OECD, 2020. OCED Observer. https://www.oecd-ilibrary.org/economics/oecd-observer_15615529/volume?vol=2020.
- Repucci, S., Slipowitz, A., 2020. Democracy Under Lockdown: the Impact of COVID-19 on the Global Struggle for Freedom. Freedom House.
- Schleicher, A., 2018. Annual Report of the Directorate for Education and Skills. OECD.
- Schliecher, A., Edwards, D., 2020. Ten Years of ISTP: What Governments and Unions and Can Achieve Together. <https://www.ei-ie.org/en/item/23380:10-years-of-istp-what-governments-and-unions-can-achieve-together-by-andreas-schleicher-and-david-edwards>.
- UNESCO, 1966. Recommendation Concerning the Status of Teachers. UNESCO.
- UNESCO, 1997. Recommendation Concerning the Status of Higher-Education Teaching Personnel. UNESCO.
- UNESCO, 2017. Global Framework of Professional Teaching Standards. UNESCO.
- UNESCO, 2021. Reimagining Our Futures Together: A New Social Contract for Education. <https://unesdoc.unesco.org/ark:/48223/pf0000379707>.
- UNESCO Institute of Statistics, 2016. The World Needs Almost 69 Million New Teachers to Reach the 2030 Education Goals. <https://unesdoc.unesco.org/ark:/48223/pf0000246124>.
- United Nations, 1989. Convention on the Rights of the Child. <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>.
- van Leeuwen, F., 2019. Remarks on Education and Democracy. YouTube.
- Williamson, B., Hogan, A., 2020. Commercialization and Privatization in/of in the Context of Covid-19. <https://www.ei-ie.org/en/item/23425:edtech-pandemic-shock-new-ei-research-launched-on-covid-19-education-commercialisation>.
- Williamson, B., Hogan, A., 2021. Post-pandemic Reform of Education: Market-First or Purpose-First Digital Transformation. <https://www.ei-ie.org/en/item/23685:post-pandemic-reform-of-higher-education-market-first-or-purpose-first-digital-transformation-by-ben-williamson-and-anna-hogan>.
- Windmuller, J., Pursey, S., Baker, J., 2019. The International Trade Union movement. In: Blanpain, R. (Ed.), Comparative Labour Law and Industrial Relations in Industrialized Market Economies. Kluwer Law International.

Global education philanthropy

Kenneth J. Saltman, University of Illinois at Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Venture philanthropy	587
Philanthrocapitalism	588
Digital directions of global education philanthropy	588
The Lego Foundation	589
Conclusion: need to shift philanthropy into the domain of the public	590
References	590

Prevalent liberal and conservative perspectives on global philanthropy in education presume that philanthropy operates in a largely charitable fashion that aims to expand educational access, inclusion, and opportunities for education to be an engine of economic mobility and national economic expansion. In these views, corporate and non-profit foundations, non-governmental organizations, state-sponsored organizations, and supra-national organizations may differ in their approaches and strategies but share a largely beneficent intent framed as universally beneficial to recipients. A less widely held critical perspective on global philanthropy focuses on how philanthropy, despite often being well-intentioned, tends to represent and promote the interests of ruling classes and dominant cultural groups that are in a position to fund, establish, and set the agendas of philanthropic foundations. In the critical perspective philanthropic projects and trends need to be understood as adjuncts to and instruments of global capital accumulation strategies and cultural ideological efforts that seeks to produce knowledge, ideology, organizational forms, and common sense in ways that naturalize and celebrate the social dominance of a small number of people over the mass of humanity. The dominant views presume an ameliorative role of philanthropy that mitigates the exceptional instances of individuals and groups who fail to succeed in an otherwise functional and potentially fully inclusive and beneficial global order. This global order is presumed by philanthropists to be ideally composed of capitalist economy paired with liberal democratic political arrangement. Philanthropic efforts often position education as the means for inclusion into such an arrangement despite the fact that such an arrangement appears incapable of keeping half the world in poverty and the other half sickened by maladies of excessive consumption.

Globally, education is positioned as a solution to non-educational public problems such as class inequality. Global education philanthropy appears largely incapable of providing the means to avert the catastrophic specters of dire public problems: global warming, the increasing risk of nuclear catastrophe, record levels of economic inequality and wealth concentration, rising authoritarianism and anti-democratic culture. Indeed, some of the largest educational philanthropy trends of the past two decades have aggressively promoted educational reforms that undermine the capacity of education to foster democracy, equality, civic life, collective agency, and public good. Global philanthropists in education work to concentrate governance and control over schools, knowledge, teaching, and curriculum in the hands of elites while undermining knowledge and learning as a means for people to comprehend and intervene against these enormous public problems. Instead global philanthropies have largely furthered the four decade neoliberal trend of treating education an investment opportunity through various forms of privatization and commercialism while fostering an ideological project that emphasizes standardization and homogenization of knowledge, depoliticization of education, the framing of knowledge as a deliverable commodity, and the reduction of schooling to its market uses (Saltman, 2010). Both the profit drive and ideological projects of global philanthropy have increasingly blurred the lines between non-profit giving and for profit business positioning business itself as philanthropy. Most recently philanthropy is playing an outsized role in education fostering digital capitalism, data extraction and commerce, and the intertwined project of “impact investing” and “innovative finance.”

In what follows here I provide an overview of how philanthropy in education has shifted over time to further the interests of a transnational capitalist class whose interests have been globalized. According to recent studies of global philanthropy the bulk of current global philanthropy has developed in the past 25 years, is focused on education, is financially dominated by US foundations, and these foundations are predominantly grant-giving as opposed to running in house programs (Fischer, 2018). According to UBS's Global Philanthropy Report done by Harvard researchers the key factors driving the expansion of philanthropy are: “global economic growth and the enormous increase in private wealth accumulation; persistent economic and social inequalities; and governmental and private efforts to encourage and support philanthropic institutions and giving” (Fischer, 2018). Academic, finance sector, and popular press accounts of global philanthropic growth tend not to recognize that there is a causality between worsening economic inequality and the upward concentration of wealth and the capacity for philanthropy. In other words, in order for philanthropists to give they must first take. The public, especially in the US, forgoes public money in taxes to financially support private foundations. Moreover, what capitalists who become philanthropists take is the uncompensated labor of workers as represented in the stagnation of real wages. Such a dynamic was evident during the heyday of the industrial economy as the captains of industry violently broke strikes to keep workers from organizing for higher pay, benefits, and shorter hours. These same bosses such as Andrew Carnegie led the creation of what is known as “scientific philanthropy.” According to Carnegie in *The Gospel of Wealth* the

crucial aims of such early philanthropy were to fund and support public sector knowledge making institutions such as libraries and universities in order to bolster meritocratic ideologies and also undermine support for radical movements such as socialism that would have redistributed control over capital from owners to workers. Despite these conservatizing tendencies of traditional philanthropy, they nonetheless represented an investment in and commitment to public knowledge-making institutions and were guided by a commitment to a collective sense of the public. They also largely provided funding to local institutions while allowing those involved in the public institutions to govern and administer the use of the funds. Public knowledge-making institutions such as schools, libraries, and museums are subject to contestation and struggle by classes and cultural groups vying for social hegemony. These aspects differed from the next major phase of philanthropy.

Venture philanthropy

In the early 1990s a new form of philanthropy called Venture Philanthropy began in Silicon Valley and was initially involved in democratic party politics. Modeled on venture capital and the finance sector, venture philanthropy radically recast the model of philanthropy and its purposes. Venture philanthropy posed philanthropy as if it is a business investment and sought to maximize “return on investment.” Though venture philanthropies are non-profit organizations and mostly seek to influence public institutions such as public schools they embrace a private sector framing and set of assumptions. Grant recipients ought to be forced to compete for scarce funding and should make system reforms along the lines favored by venture philanthropists. These reforms were decidedly neoliberal privatization- and deregulation- oriented reforms. For example, The big three venture philanthropies Gates Foundation, Walton Foundation, and Broad Foundation each sought to promote educational privatization with different emphases. Gates aggressively pushed charter school expansion of publicly funded yet privately managed schools and also led the standardization and homogenization of curriculum through the Common Core State Standards. Walton played an outsized role in promoting school vouchers for parents to opt out of public schooling and use public money to consume private schools. Broad promoted and developed the privatization and corporatization of educational leadership among other projects. These and other venture philanthropists “leveraged” their foundation money to influence states, districts, and schools to compete against each other for money by embracing market-based reforms. On the model of investment capital the “successful” school models would be “scaled up” while public schools would be allegedly forced to compete against privatized schools. Even though the total amount of money going from foundations to schools was small relative to total spending on public education, the influence was enormous as cash strapped schools and districts competed to out-comply with the venture philanthropists’ mandates such as lifting the caps on allowable charters in a district. The venture philanthropy model became integrated into the US Department of Education under the Obama administration with figures from the Gates Foundation setting up competitions to embrace privatization under the promise of getting scarce federal money.

Venture philanthropy was significant for shifting governance over public education away from the public and into the hands of the billionaire donors who established the foundations. It was also significant for influencing the public to accept a reframing of public education as being less about collective responsibility for all students and families and more a private consumable commodity for individual families to game in order to compete for shrinking access to jobs and future economic opportunity. While venture philanthropists spent billions of dollars on these projects, in traditional measures of test-based performance, vouchers performed poorly and charters did no better and sometimes worse than traditional public schools. Nonetheless, these reforms continue despite a lack of such empirical evidence for their success. Market ideology has been the driving force for these reforms. Many of these reforms have done measurable damage to public education. For example, in the US charter school expansion resulted in exacerbated racial segregation, push outs and exclusion of students in need of special education and language services. Gates and Zuckerberg have funded Bridge International Academies, a for profit educational management organization in India, Kenya, Uganda, Liberia, and Nigeria that has been criticized for human rights abuses.

... the United Nations, the African Commission on Human and Peoples’ Rights and the UK parliament’s International Development Committee have raised concerns about the quality of BIA’s educational programme, its relationship with governments, its compliance with government regulations, and relatively high cost of fees.

Doward (2020).

As well, the neoliberal education restructuring promoted by venture philanthropists bolstered excessive testing and anti-intellectual and anti-critical forms of teaching and approaches to curriculum, and expanded repression especially targeting working class and poor students and non-White students. Knowledge and learning within the logic of these reforms was framed not as a means of understanding and collectively governing the world but rather as a means of accumulating decontextualized knowledge that could be exchanged for education promotion to be then exchanged for economic opportunity. Though the standardization, homogenization, anti-intellectual approaches to teaching and learning were positioned by venture philanthropists as apolitical, neutral, and universally-beneficial they represented the particular interests and perspectives beneficial to donors. An important consequence of the privatizations promoted by venture philanthropy has been that they close down public space as contested terrain. As public schools become privately managed and owned public struggles over governance, curriculum, and pedagogy become foreclosed.

Philanthrocapitalism

As the influence and impact of venture philanthropy peaked around 2010 contemporary educational philanthropy took a slightly different direction. The Chan Zuckerberg Initiative, Omidyar Network, and Emerson Collective established Limited Liability Companies comprised of both for profit businesses and non-profit businesses and called them educational philanthropies. Philanthrocapitalist LLCs such as CZI and Emerson invest in for profit education companies particularly in the technology sector. Both have invested in adaptive learning technologies. The legal structure of an LLC allows the company to evade scrutiny the way non-profits cannot. The movement of money and financial data can be kept secret by these organizations thereby keeping the public from knowing what these businesses are doing and how activities deemed “philanthropic” are actually profit seeking and at odds with any stated mission. If venture philanthropists hijacked public governance by allowing super-rich individuals and corporations to take governance control over educational priorities and public resources to foster a radical neoliberal agenda, philanthrocapitalism takes this a step further. Philanthrocapitalists also hijack public governance but do so while engaging in direct school commercialism and privatization activities. They own and direct education businesses that seek to contract with public schools, districts, and governments and thereby capture public dollars and obscure how much money gets directed from educational resources and processes and skimmed out as business and investor profit under the guise of charity.

Global philanthropy in education needs to be understood as part of what has been termed the Global Education Reform Movement. This movement comprised of supra national organizations, corporate foundations, philanthropies, consultancies, think tanks, and for-profit businesses promotes the market-based restructuring of education around the world. The aim is to transform public education as a public good into a private consumable service. The agenda has also centrally been about promoting a vision for public education dominated by corporate culture and business uses for education. In rich countries the privatization agenda has for decades focused on promoting vouchers, charters, “neo-voucher” scholarship tax credits, and the standards and accountability movement that revives industrial efficiency models of schooling. In poor countries the agenda has been to avert the development of public schooling and instead support privatized directions for education typified in the push by the World Bank for pay for fee education rather than free universal public education.

In the past decade two global trends have been integral to the directions of global educational philanthropy: impact investing and other so-called innovative finance schemes and the vast expansion of digital educational products and services. One of the most significant impact investing schemes is known as Social Impact Bonds or Pay for Performance. The Rockefeller Foundation is among the largest global players in Social Impact Bonds. These arrangements, facilitated by philanthropies, bring together public sector actors such as states or cities with large investment banks. They finance educational and other social service projects vastly inflating the costs of the service for the public entity on the premise that if metrics show ineffective performance the public entity will not have to pay. These arrangements are framed by proponents as promoting accountability and cost savings. However, evidence suggests that the programs selected and involvement of vested interests assure that programs are nearly always paid out to investors (Saltman, 2018). Sometimes investment banks like Goldman Sachs can double their money at public expense as they did with an early childhood program in Chicago (Saltman, 2018).

Digital directions of global education philanthropy

Impact investing in which for profit investing is muddled with philanthropy has become increasingly interwoven with the expansion of digital forms of educational commercialism. The expansion of the Social and Emotional Learning trend in education is a case in point. Social Emotional Learning has expanded ostensibly to counter excessive testing and the dehumanizing tendencies of the standards and accountability movement that was promoted in large part by venture philanthropist Bill and Melinda Gates Foundation.

Globally, educational privatization is expanding in new directions characterized by the trade in data and digital technologies, the making of students and teachers into data engines, and the use of data and digital technologies to expand impact investment schemes (Williamson and Piattoeva, 2018; Saltman and Means, 2019). The quantification of so-called academic achievement in the form of standardized testing has long been big business benefitting particularly test and textbook companies (Strauss, 2015). Education and technology sectors have been converging as test and textbook publishing education companies like Houghton-Mifflin, ETS, Pearson, and Kaplan have increasingly become technology companies, and technology companies have increasingly become education companies (Saltman, 2016). However, the quantification and datafication of affect, student behavior, learned dispositions particularly for self-control, and so-called “soft-skills” under the rubric of Social Emotional Learning (SEL) are at the center of the new forms of technology-oriented privatization on a global scale (World Economic Forum, 2020; Williamson and Piattoeva, 2018). Such social quantification projects in SEL employ artificial intelligence education technologies such as adaptive learning technologies, biometric pedagogy devices like real-time webcam systems, video games, and avatars. The increasing turn to harnessing the daily activities of students, teachers, and administrators as data manufacturers via technology and the financial securitization of this data builds on longstanding dominant trends of profit seeking in education. That is, such technologies displace the labor costs of teachers and allow corporations to capture public education tax dollars by replacing the work and meaning making activities of teachers with corporate technology in the form of contracting (Barshay, 2019, p. 6). The trade in and devising of impact investment financial instruments based in student data-making represents a new major source of school profit seeking (Barshay, 2019). Many venture philanthropy players like the New Schools Venture Fund have shifted into the convergence of digital commerce and resilience discourse. For example, New Schools Venture Fund which was a promoter

of charter schools is now financing tech start ups that sell social emotional learning programs like the for profit companies Class Dojo, Centervention, and Amira.

This steady expansion of digital profiteering in public schooling needs to be understood not merely as a consequence of new technologies but rather as part of a broader class-based global project for hegemonic control by the transnational capitalist class (Robinson, 2018, pp. 11-29). As William I. Robinson argues, crises of capital overaccumulation and state legitimacy have resulted in strategies for “militarized accumulation” to expand markets into hitherto uncolonized places including the lifeworld and subjectivity (Robinson, 2018, pp. 11-12). So called innovative finance schemes have long been a means to justify this agenda by shifting the discussion of public goods and services into the language and logic of business and markets. Within this logic public schools have been declared failures that need to “go out of business” while schemes for competition, choice, and turnarounds have been pushed by philanthropies as the solution.

The recent rapid growth of two major new forms of educational privatization are in fact converging: (1) impact investing schemes that promote privatization of education such as venture philanthropy and social impact bonds, and (2) data producing educational technologies. The specific form of digital capitalist pillage of the lifeworld takes shape as what Couldry and Meijas refer to as the industries of social quantification (Couldry and Meijas, 2019). These two forms of privatization are coming together in part through the quantification of affect and behavior particularly through the discourses of resilience, Social Emotional Learning, and play based learning that are being promoted by supra-national organizations, corporations, investors, and philanthropic foundations. Resilience is a class based ideology that suggests that poverty and inequality cause stress and trauma in poor students. The solution in the discourse of resilience is not to end poverty and inequality by for example redistributing control and ownership of capital from the 0.1% to everyone. Rather the solution is to give individual coping strategies for students to develop character traits such as persistence or grit and tools for emotional self-management. There is a big business in resilience and impact investment “philanthropies” such as New Schools Venture Fund form a support structure enabling commercialism, resource-skimming, and the roll out of the resilience ideology. The activities of one such philanthropic foundation, The Lego Foundation, typifies the turn to quantifying student behavior, affect, and educational activity in order to create commercial opportunities for corporations under the aegis of care and charity.

The Lego Foundation

The Lego Foundation and the Real Play Coalition, of which it is a part, have been working with the World Economic Forum (WEF) and the Organization of Economic Cooperation and Development (OECD) to support the expansion of the measurement, quantification, and assessment of play. The Lego Group, headquartered in Denmark, is the largest toy company in the world earning about \$2 billion a year through sales of its plastic interlocking blocks but also its Legoland theme parks, series of highly lucrative Hollywood movies, and other increasingly digital merchandise. The Lego Foundation, a not for profit philanthropic organization, funded by the company, plays a prominent role in promoting play based learning internationally, promoting the idea of an international educational skills crisis that ought to be met with play based learning, and encouraging the measurement and assessment of play and the inclusion of such assessment in the OECD Pisa international educational comparison tests and the linkage of quantified play to the Millennium Development Goals promoted by the World Bank (Parker et al.). Lego Foundation produces a body of literature framing play based learning as fostering a specific brand of creativity that is defined by school to work skills aligned with the broader instrumental discourses of Career and Technical Education and STEM. Though it employs a number of seemingly progressive pedagogical terms the pedagogy delinks student learning and experience from the broader social structures, forces, and antagonisms that make learning situations meaningful and it defines agency through immediate problem solving skills and against self and social interpretation that could form the basis for social intervention. The approach is similar to that of other resilience pedagogies ranging from Grit to meditation programs that aim to foster in students a focus on subjective experience delinked from the objective conditions that produce subject positions and the culture within which these pedagogies become meaningful.

The quantification of play for the Lego group at the very least then has two important dimensions. One the one hand by inserting play into the global accountability and standards movement spearheaded by supra-national organizations OECD and WEF that represent multi-national capital, Lego aims to position their business and their products as essential to educational provision if not renewal, human development, economic growth through human capital, and individual skills that contribute to that economic growth.¹ By promoting play as quantifiable, measurable, and integral to global educational comparisons, Lego stands to legitimate itself and its profit seeking activities in these regards while increasing the likelihood of greater revenue. If quantified play can be successfully integrated into global comparative educational accountability standards, then play will be measured along with other academic achievement on national tests that form the basis for international comparisons. There are rewards and punishments for nation-states for complying with or failing to comply with OECD/WB mandates for education. Nationally, efforts will be made to increase test scores by focusing on the skills and dispositions that will be measured. Having successfully lobbied for the incorporation of play based learning as a measurable skill, Lego will be ready to provide the curriculum materials in the form of Lego products. The Lego Foundation is spending a lot of money and working with a lot of other organizations to accomplish this legitimization project: Project Zero at Harvard University, collaborations with OECD, WEF, and the Real Play Coalition to name a few of the most

¹The Lego Foundation's reports and advocacy materials make these aims explicit and the OECD's reports include the play based learning agenda as part of its neoliberal human capital development perspective.

significant. Secondly, the agenda for the quantification of play creates the conditions for the increasing convergence of education and media entertainment corporations around digital learning products. This convergence of education and digital media entertainment naturalizes as necessary for learning and human development gamified forms of pedagogy typified in digital Social Emotional Learning apps, biometric measurement systems, personalized learning products, and the interface of physical toys like Legos and drones with digital apps. As well, this convergence creates the conditions for the ever-greater digital measurement and surveillance of children's activity and human activity, the normalization and "making innocent" of total surveillance and commercialization of that surveillance through the discourse of childhood innocence.² Such digital surveillance and commercialism includes children's use of screens on various devices, webcams but also the use of biometric tracking devices that are now being put on children in some schools throughout the day. The social control and commercial extraction/digital representation possibilities of datafication are part of a broader trend toward the "Internet of Things" IOT aiming to integrate the physical world with digital industry (Alexander, 2018; Dyer-Witthford et al., 2019). The Covid 19 pandemic both accelerated the trends toward digital educational commerce and brought to the fore contests over the longstanding agendas of some philanthropies.

Conclusion: need to shift philanthropy into the domain of the public

In nations that give tax subsidies to philanthropy what ought to be recognized is that rather than generosity and charity these subsidies to the rich need to be seen as give-aways of public governance and public resources that could allow for greater citizen control over education and allow for money that is needed to support historically underfunded schools. As well, the ideological tendencies of most philanthropy need to be scrutinized for the key questions: Whose purposes for education are being promoted? Whose knowledge and curriculum is being promoted through these endeavors? The most charitable act philanthropists could do would be to work to change the systems that amass wealth and exacerbate inequality, and spread inequality and immiseration—that is, the very mechanism that allow philanthropists to be philanthropists in the first place.

References

- Alexander, J., 2018. *Means Learning to Save the Future*. Routledge, New York.
- Barshay, J., October 21, 2019. Impact funds pour money into ed tech businesses. *The Hechinger Report*.
- Couldry, N., Meijas, U.A., 2019. *The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press, Stanford.
- Doward, J., April 11, 2020. Stop spending millions on for-profit schools abroad, campaigners tell UK. *The Guardian*.
- Dyer-Witthford, N., Kjoson, A.M., Steinhoff, J., 2019. *Inhuman Power: Artificial Intelligence and the Future of Capitalism*. Pluto Press, London.
- Fischer, M.S., April 27, 2018. Global philanthropy sector is growing fast: Harvard Study. *Think Advisor* 2.
- Parker, et al., pp. 1–9.
- Robinson, W.I., 2018. *Into the Tempest*. Haymarket, Chicago.
- Saltman, K.J., 2010. *The Gift of Education: Public Education and Venture Philanthropy*. Palgrave, New York.
- Saltman, K.J., 2016. *Scripted Bodies: Corporate Power, Smart Technologies, and the Undoing of Public Education*. Routledge, New York.
- Saltman, K.J., 2018. *The Swindle of Innovative Educational Finance*. University of Minnesota Press, Minneapolis.
- Saltman, K.J., Means, A.J., 2019. *The Wiley Handbook of Global Educational Reform*. Medford, MA: Wiley Blackwell.
- Strauss, V., March 30, 2015. Big education firms spend millions lobbying for pro-testing policies. *Washington Post* (Valerie Strauss detailed the millions of dollars educational publishers invested into lobbying for testing and the billions they reaped).
- Williamson, B., Piattoeva, N., 2018. Objectivity as standardization in data-scientific education policy, technology, and governance. *Learn. Media Technol.* 64–76.
- World Economic Forum, 2020. *Schools of the Future: Defining New Models of Education for the Fourth Industrial Revolution*, pp. 1–33. www.weforum.org.

²On the discourse of childhood innocence and its political and pedagogical dimensions. See Henry Giroux *Mickey Mouse Monopoly*.

For profit corporations and education reform: global dimensions and impacts

Patricia Burch^a and S. Srinivasa Rao^b, ^aCenter on Education, Policy, Equity and Governance, Rossier School of Education, University of Southern California (USC), Los Angeles, CA, United States; and ^bZakir Husain Center for Educational Studies (ZHCES), School of Social Sciences (SSS), Jawaharlal Nehru University (JNU), Delhi, India

© 2023 Elsevier Ltd. All rights reserved.

When, how and why did for-profit corporations get involved in the field of education?	591
How is the role of FPCs in education being interpreted in education research?	592
What are the effects of the involvement of FPCs in education in terms of geopolitics and equity?	592
Conclusion and future research	594
References	594

When, how and why did for-profit corporations get involved in the field of education?

Since the 19th century, corporate actors have been involved in the “reforms” of primary and secondary education worldwide (Arno, 1980; Carnoy, 1975; Cuban, 2004; Hall, 2008; Lagemann, 1992; Molnar, 1996; Shipps, 2006; Watkins, 2001). The role and reach of corporate actors in education has evolved dynamically with social and historical (geopolitical context) (Au and Ferrare, 2015; Fabricant and Fine, 2015; Lipman, 2011).

By for-profit education corporation (fpec), we mean, following Moeller (2020), “actors, processes, networks, styles, and power relations related to businesses or the for-profit sector” as well as rulemaking and legislation related to the for-profit sector. Moeller (2020) includes in this definition the following types of organizational actors: corporations (Burch and Good, 2014; Koyama, 2010; Robertson et al., 2002); corporate foundations (Moeller, 2018; Schwittay, 2006); private foundations (Reckhow and Synder, 2014; Scott, 2009; Tarlau and Moeller, 2020; Tompkins-Stange, 2016); nongovernmental organizations (NGOs) (Kamat, 2004); religious organizations, think tanks, and lobbying groups (Anderson and Donchik, 2016); parent associations (Posey-Maddox, 2016); and global and national forums (Ball, 2012). We extend this definition to include state/government actors, at the national, regional, subregional and local level whose activities incentivize (via rewards and sanctions) corporate activity.

Globally, the types of reforms in which corporate actors have participated fall into the following main categories: charter schools, school vouchers, educational tax credits, philanthropy, for-profit schools, public-private partnerships, and education reform. These and other reform activities can be further differentiated as market or quasi-market reforms (Whitty, 1997). In quasi market education reforms, education might be publicly financed, with the government or state contracting with private agencies for service delivery as is the case of charter schools. Alternately, we also see for-profit corporate activity in reforms which shifts funding and control of schools to for-profit corporations. For-profit corporate actors can incentivize families to exit the public system through home schooling and/or the establishment and creation of new types of schools, through activities such as school vouchers and charter schools (Gunter et al., 2017; Levin, forthcoming). Reforms that involve corporate actors sit at the nexus of a fundamental tension in education: the public and private purposes of education (Levin, 1987). The public purpose is to prepare students to participate in democratic practices that protect freedoms. The private purpose concerns the rights of families to decide how their children are educated and the ideas to which they are exposed. Political and social crises (such as the pandemic) create opportunities for policy entrepreneurs to capitalize on the tensions between public and private goals of education for commercial or political gain, examples of which are discussed below.

It is important to note that the political systems and policies in place at the nation-state level, can influence corporate reform activities and expansion, but do not determine the presence or absence of activity even by law. One contemporary example of the former concerns recent actions by Chinese government to regulate the creation and operation of on-line after school tutoring companies under the publicly stated principle of reducing educational inequalities. The companies, with revenues in billions of dollars, were criticized by the Chinese government as widening income inequalities between well to do Chinese families with the income to purchase after school tutoring (and expected educational advantages) for their children and lower income families who lacked the means to do so. Government policies at least for the short term have generated substantive changes in the composition of the tutoring marketplace (in China and globally).

The involvement of corporate actors in education reform tends to result from one of the following privatizing trends a shift in the activities from public to private, including deregulation and globalization, and/or a shift in the production of goods and services from the public to private sector (Starr, 1988). These shifts may be the result of intentional government action, changes in the demands/purchasing behavior of individuals and organizations that may or may not be linked to government action. In either case, the shifts are always interdependent with broader trends in the institutional environment of education reform which can be understood as linked to broader ideological and geopolitical debates and enduring conflicts over the purposes of education. Locating the drivers of corporate activity in education requires moving beyond the basic binary of government and market and traditional conceptions of societal functions. As noted above, government actors are corporate actors with the investment capital to create entirely new markets in education from which both financial and political benefits flow back to governments. For profit

corporate entities that identify legally as corporations are political and government actors whose rule-making (particularly in the case of multi-national companies or companies that dominate regional/national markets) set policy for public access to services.

How is the role of FPCs in education being interpreted in education research?

There has been much good work synthesizing the nature and types of impacts of for-profit corporate activity based on the literature (c.f. Moeller, 2020; Verger, 2020). One research stream has focused on corporate actors as extensions of broader market logics in education and the effects of these market logics on educational policy and practice. This research has concluded that corporations act as carriers of corporate style reform which are modeled on the logics of efficiency and accountability (c.f. Burch, 2009; Molnar, 1996; Shipps, 2006). Through their influence, corporate actors shape policy and practice, conceptions of teachers' work, what counts as reform among other issues (Ball and Youdell, 2007; Ball, 2012; Lipman, 2004). From this work, we learn that while corporate activities move "globally," they manifest locally and are historically situated (Burch, 2021; Burch and Miglani, 2017; Miglani and Burch, 2021; Engel, 2008; Engel and Frizzell, 2015; Maxwell et al., 2020). And in this regard, the influence of corporate education actors is contingent on local action. Local actors—individual and organizational—rarely adopt reforms whole cloth. The impacts of for-profit corporate entities are mediated locally, as reforms are partially adopted and transformed, ignored or in other ways and at other times resisted and/or rejected (Burch and Good, 2014; Lipman, 2004; Münch, 2020).

A second stream of research has focused on whether and how the influence of these corporations has been "better" more effective/efficient in terms of student outcomes relative to public (non-corporate provision of services). This research has concluded that public models of service delivery as compared to private models tend to be different in terms of policy goals but both tend to be informed by costs. For-profit organizations seeking to make a profit may try to operate schools or make and administer tests inexpensively so they can make a profit. Public agencies may not be driven by profit but they can be driven by costs and their decisions to outsource to corporations may be influenced by the decision to reduce their own accountability for outcomes (c.f. Adamson et al., 2016; Carnoy, 2000; Chubb and Moe, 1990; Cuban, 2004; Levin, 2001; Lubienski, 2005; Scott, 2005).

A third (more recent) line of research has focused on the role of corporate actors within broader policy networks and streams and the role of power as deployed by corporate actors within these networks (c.f. Ball, 2012; Hogan, 2015; Verger et al., 2016; Gulson et al., 2017). As part of a broader body of work (in policy sociology) and on policy networks, this research examines how corporate actors within policy networks participate in the production of education policy (c.f. Ball, 2012; Ball et al., 2017; Münch, 2020). This work has emphasized the interdependence and intersectionality between corporate actors and state actors including international organizations, states and schools (c.f. Au and Ferrare, 2015; Lingard, 1993). Corporations as part of broader education reform networks operate locally and globally and across different parts of the world (c.f. Bartlett and Vavrus, 2015).

Policy entrepreneurs are integral to the activities of for-profit education corporations. For example, Nambissan and Ball (2010) highlight the role of "policy entrepreneurs" such as James Tooley, a British academic, who acted as a key nodal intellectual in choice advocacy as well as the for-profit education for the education of the poor in India. As a former consultant of the International Finance Corporation, Tooley directed the global study of investment opportunities for private education in developing countries for the IFC—the private finance arm of The World Bank." Research for advocacy by Tooley and others was undertaken in poor settlements of Hyderabad in the newly formed southern state of Telangana in the early 2000s (Tooley, 2009; Smith et al., 2005). As the state-run schools were not performing, claims were made that the poor willingly chose and paid for private education. The findings of the study were used by profit think tanks, media, research cohorts in academic institutions and in academic journals, and the multinational organizations, with policy entrepreneurs acting as intermediaries.

Thus, rather than separate, the research enterprise is a part of for-profit corporate activity in education. To offer another example, the corporate driven international school quality surveys such as PISA and TIMS have tended to depict government schools specifically in the global south as inefficient and of low quality and in doing so, to make a serious case for corporate involvement. In India, the role of Annual Survey of Education Reports (ASER) is a case in point. This corporate driven large-scale survey of learning achievements in its initial years was used to survey government run schools and concluded that the state schools were poorly performing and inefficient. More recently, in response to criticism (Vellanki, 2015), the ASERs have included private schools. While the ASER (2015) suggests that there have been substantial differences in learning achievements between the government and the private schools, the data presents a much more mixed account. The ASER report (2018) used point in time data to make the case that private schools were better than public schools, and the report's authors interpreted the data to support the expansion of low-cost schools.

What are the effects of the involvement of FPCs in education in terms of geopolitics and equity?

The intent and effects of involvement of for-profit corporations has been different in diverse national contexts. The differences exist in terms of not just the spread of educational services provided and produced, but also in terms of the role they play in education policy making or in transferring some of the policy ideas from one national context to another. Nambissan and Ball (2010) argue that the biggest contribution of for-profit corporations in education across the world is that they disseminate education policies through policy transfers and policy entrepreneurship and in the end result in what they call "policy convergence."

If we consider privatization of education as a critical response to the so-called failure of public education systems, then the starting point of privatization of educational provisioning has occurred in different periods of time in different national contexts. For example, while the contemporary privatization agenda became significant at the end of the 1970s and the beginning of the 1980s in the United Kingdom and the United States (Verger, 2020), their actual arrival in the global “South” has happened after almost two decades, mostly in the 1990s or in the early years of the new millennium, by which time the experiences of those initial experiments in the global “North” have yielded not much success. The UK and the US are “reform laboratories from which experiments are exported around the world” (Nambissan and Ball, 2010: p. 325), through the Transnational Advocacy Networks (TANs), which “change national government behavior through the exchange of norms, ideas and discourses” (Nambissan and Ball, 2010: p. 326).

What is significant to note is that the language of the FPCs in the emerging neoliberal contexts of the global “South” and the global “North” is qualitatively distinct. At times, the failed policies of the North made their entry into the South and survived due to the highly unequal hierarchical social as well as educational order in those societies. A pro-profit idea like the voucher system is one such example which has not worked in one context (the global “North”), but is applied in another context (global “South”) with the promise of mitigating social and educational inequities and the provision of good quality education. In India, the advocates of for-profit education (Das, 2020; Kathpalia, 2016; Kingdon, 2015) call for dropping the constitutional ideal of socialism as it disallows a private schools from making profit and thereby what they claim will harm the interests and aspirations of poor parents and children. Advocates claim that the single change in designation from “non-profit” to “profit” sector could create new market mechanism. Investments would flow rapidly into education, improving choice and quality (Das, 2020). This kind of advocacy, drawn from the corporate world and the FPCs, seek to advance for profit interests in the name of reforms, without considering concerns of socio-cultural and political economic contexts of societies in the “South.”

To be specific, the most significant feature of advocacy for reforms in the global “South,” particularly in India and Sub-Saharan Africa is that there is a clear absence of policy framework (even if it exists, it is fraught with a total lack of regulation), which enable proliferation of schools outside the public sector, which then results in mushrooming of unorganized, unregulated and unrecognized educational market for the poor in the name of low-cost private schools or budget schools. Phillipson (2008) foresees the emergence of a private education “that charges a fee that is low enough to be affordable by the poor families” (p. 1). He argues that this low-cost private education is seen as a “popular alternative to the public education system” (2008: p. 1) in countries like India, Nigeria and Uganda.

While it is difficult to precisely define the private-ness of the low-cost private school, Phillipson does point to what it is not, which help us understand what it actually is:

“It is not a school run by a religious organization for charitable religious organization for the furtherance of a particular set of moral values or beliefs. It is not a school offering an educational advantage to its pupils and charging a high price for a privilege of gaining access to it ... it is not a school set up by the local community until the government agrees to take over ownership. In contrast to these distinctions, the low-cost private school is a school that has been set up and is owned by an individual or individuals for the purpose of making profit” (2008: p. 1).

In India, the low-cost schools are generally not recognized and therefore are not listed in the annual census of educational institutions by large scale surveys like DISE and ASER. Moreover, policy makers and political class are complicit in facilitating the growth of for-profit, low-cost schools (Nambissan, 2016). These are usually run at what Phillipson (2008) calls “subsistence level” education and “a pre-occupation with economic survival permeates all activities in such schools” (p. 2). The growth of this sector seems to be the cause and consequence of public school closures post 2014 policy of rationalization of teacher deployment from schools with low enrollments to those with high enrollments, which led to decline in the government schools (Rao, 2017; Phillipson, 2008). In other words, the declining enrollments in government schools is a cause and consequence of not just the governmental policies, but it also indicates a preference for the low-cost schools. In Nigeria and Uganda, in fact, there has been a governmental support to these schools through a provision of matching grants and funds and of instructional materials (Phillipson, 2008).

For parents, “choosing a school is not always a one-off event” (Hill et al., 2011: p. 103). It involves a process of “symbolic expression of identity” (Gurney, 2017: p. 19). They often contemplate, revisit and review their original schooling choices if they perceive that the benefits are no different from their previous choice. Several studies, using Hirschman’s (1970) ideas of exit and voice, explain this process in different contexts. This, some researchers believe, has produced a differentiated school market (Hill et al., 2011; Srivastava, 2013), which in turn produces a stratified education system which internally differentiates poor in terms of their accessing low-cost private schools and those accessing a functional government school in the neighborhood. Thus, increased investment and expanded educational opportunities for the poor through low-cost private schools across the global “South” is not necessarily a panacea for social inequality, disadvantage and poverty (Hill et al., 2011; Nambissan, 2016).

Further, as mentioned earlier, though quality of education is said to be the reason for parental choice for low-cost private schools, the education provided in these schools is found to be of low quality (Kar and Sinha, 2021; Nambissan, 2012; Goyal and Pandey, 2012; Karopady, 2014). Yet, FPCs’ push for this sector of schooling in the Global South can be viewed as a play to secure entry into a potential market that survives on extracting from the poor households for whom accessing education is a necessary condition for escaping miseries of poverty and social disadvantage and to realize aspirations for social mobility. The poor parents see these low-cost private schools as easy and cheap options that can improve schooling as well as living standards for their children (Kumar, 2010).

Thus, the participation of corporate actors and the market-based approaches in both developed as well as developing countries have not made a huge difference to the prevailing education systems that are already unequal (Lahoti and Mukhopadhyay, 2019; Nambissan, 2016). At best, they created differences between sub-regional systems within the same education system and widened

inequities more prominently. Thus, an important consequence of this discourse is the emergence of inequities in terms of national systems and the regional and sub-regional systems and among the social group differences. Differential levels of privatization and corporatization and the widening of inequalities within the regional and sub-regional systems within the same national contexts are well documented (ASER, 2015, 2018). This is made even more complex by ideological orientations of the federal or Union and state/provincial governments, due to which the policy frameworks vary largely.

The FPC's intent in creating more choices for quality and for efficient delivery of provision of schooling has just not to be read as widening of choices, but must be seen as widening of inequities among pockets of affluence and pockets of deprivation; rural and urban spatial contexts; the social, racial, ethnic, gender, caste, tribal and religious minorities. As Kamat et al. (2016) argue, "corporations are primarily accountable to their stockholders and work first and foremost to create returns for investors, which are not necessarily aligned with those of the customers, i.e., the students, their families or their communities" (p. 100). They also endorse the view that "any system of education that increasingly relies on private provision to fulfill the right to education, creates a cycle wherein the poorest and most marginalized have diminishing access to quality education" (Kamat et al., 2016: p. 101). Evidence points to an increased socio-economic segregation as a result of school choice for poor and the disadvantaged (Kar and Sinha, 2021; Kumar, 2010). In such contexts, does diversity get augmented, as Steiner-Khamsi (2018) asks, amid what she paraphrases as a situation wherein "businesses seeing like a state" and "governments calculating like businesses" (p. 382).

The most important consequence of FPC involvement in education is surely not one of bridging educational inequalities in access to quality education, rather to make profits at the cost of the poorest of the poor, which then further widens inequalities as FPCs tend to exclude those who cannot afford and cannot succeed in terms of what they set as standards. State-run education shrinks and gets restricted to the social groups that are most disadvantaged socially economically and are vulnerable and under-represented. As FPC's push low-cost schools in the global south and use research as a political tool to legitimate this expansion, concerns mount about the continued role of the state in the provision of basic education.

Conclusion and future research

Empirical research on the role of for-profit corporations in education reform converges around the following points. First, the phenomenon is best understood as a broad social field that involves not only private for-profit actors but also government and state actors. Government and state actors incentivize much of the work, as opposed to serving a primary regulative function. Further, the distinction between for profit and not for profit actors means little in a context in which large not for profit actors operate as not for profits in name only. These entities operate independently as part of broader policy networks and in this work the role of government actors, not for profit actors and for-profit corporations can be indistinguishable with governments creating new markets and for-profit corporations setting policy. The drivers of this activity are global and as such connected to the rise of global trends such as neoliberalism. However, the manifestations of these trends are often "local"—the meaning and significance of "global" reform activities such as charter schools are locally situated and culturally determined. In all instances, the conflicts around this activity reflect a broader set of tensions around the public democratic purposes of public education and the private interests and liberties of families (and by extension for profit corporations selling choices to these families. The positive impacts of for-profit corporate activity around the claimed goals of increased efficiencies and increased equity are minimal, though still debated given that research on these activities itself is connected to rather than independent from the corporate enterprise. There are significant commercial and political interests at stake in the expansion of for-profit education corporations, particularly at a moment in which the democratic purposes of schools have come under attack and where multi-national corporations, for example in the online-tutoring industry, have felt the sting of government policy.

More serious research and policy attention is needed on the impacts of unregulated school sector that seeks out the poor. It also is important to further explore how the differentiated educational environments that this sector has created produce social hierarchies and internal differentiation in otherwise homogeneous neighborhoods. Research also needs to be directed toward the classroom processes within these schools to examine if these schools indeed are different from the state schools, which has been the sole selling point for the FPCs for entering into the low-cost schooling market. This research must attend not just to impacts but also to what Verger et al. (2017) call, "market making" and their differential social impacts on nations that are yet to universalize provision of their elementary education.

References

- Adamson, F., Astrand, B., Darling-Hammond, L. (Eds.), 2016. *Global Education Reform: How Privatization and Public Investment Influence Education Outcomes*. Routledge.
- Anderson, G.L., Donchik, L.M., 2016. Privatizing schooling and policy making: the American Legislative Exchange Council and new political and discursive strategies of education governance. *Educ. Pol.* 30 (2), 322–364.
- Arnove, R. (Ed.), 1980. *Philanthropy and Cultural Imperialism: The Foundations at Home and Abroad*. G. K. Hall.
- ASER, 2015. *Trends Over Time 2006–2014*. ASER Centre.
- ASER, 2018. *Annual Status of Education Report (Rural)*. ASER Centre.
- Au, W., Ferrare, J.J. (Eds.), 2015. *Mapping Corporate Education Reform: Power and Policy Networks in the Neoliberal State*. Routledge.
- Ball, S.J., Youdell, D., 2007. *Hidden Privatisation in Public Education*. Education International, Brussels.
- Ball, S.J., Junneman, C., Santori, D., 2017. *Edu.net: Globalisation and Education Policy Mobility*. Routledge, New York.

- Ball, S.J., 2012. *Global Education, Inc.: New Policy Networks and the Neo-liberal Imaginary*. Routledge.
- Bartlett, L., Vavrus, F., 2015. Transversing the vertical case study: a methodological approach to studies of educational policy as practice. *Anthropol. Educ. Q.* 45 (2), 131–147.
- Burch, P., Good, A.G., 2014. *Equal Scrutiny: Privatization and Accountability in Digital Education*. Harvard Education Press.
- Burch, P., Miglani, N., 2017. Technocentrism and social fields in the Indian EdTech movement: formation, reproduction and resistance. *J. Educ. Pol.* 33 (5), 590–616.
- Burch, P., 2009. *Hidden markets: The new education privatization*. The Critical Social Thought Series. Routledge, New York.
- Burch, P., 2021. *Hidden Markets: The New Education Privatization (Critical Social Thought Series)*, second ed. Routledge.
- Carnoy, M., 1975. *Schooling in a Corporate Society: The Political Economy of Education in America*. McKay.
- Carnoy, M., 2000. School choice? Or is it privatization? *Educ. Res.* 29 (7), 15–20.
- Chubb, J.E., Moe, T.M., 1990. *Politics, Markets, and America's Schools*. Brookings Institution Press.
- Cuban, L., 2004. *The Blackboard and the Bottom Line*. Harvard University Press.
- Das, G., 2020. India is Free, Its Schools are Not. *The Times of India*. February 6, Thursday.
- Engel, L.C., Frizzell, M.O., 2015. Competitive comparison and PISA bragging rights: sub-national uses of the OECD's PISA in Canada and the USA. *Discourse* 36 (5), 665–682.
- Engel, L.C., 2008. Breaking the centre-periphery mould: exploring globalisation and educational governance in Catalonia. *Glob. Soc. Educ.* 6 (3), 265–279.
- Fabricant, M., Fine, M., 2015. *Changing Politics of Education: Privatization and the Dispossessed Lives Left behind*. Routledge.
- Goyal, S., Pandey, P., 2012. How do government and private schools differ? *Econ. Polit. Wkly.* 47 (22), 67–76.
- Gulson, K.N., Lewis, S., Lingard, B., Lubienksi, C., Takayama, K., Webb, P.T., 2017. Policy mobilities and methodology: a proposition for inventive methods in education policy studies. *Crit. Stud. Educ.* 58 (2), 224–241.
- Gunter, H.M., Hall, D., Apple, M.W. (Eds.), 2017. *Corporate Elites and the Reform of Public Education*. Routledge.
- Gurney, E., 2017. Choosing schools, choosing selves: exploring the influence of parental identity and biography on the school choice process in Delhi, India. *Int. Stud. Sociol. Educ.* 26 (1), 19–35.
- Hall, C., 2008. Making colonial subjects: education in the age of empire. *Hist. Educ.* 37 (6), 773–787.
- Hill, E., Samson, M., Dasgupta, S., 2011. Expanding the school market in India: parental choice and the reproduction of social inequality. *Econ. Polit. Wkly.* 46 (35), 98–105.
- Hirschman, A.O., 1970. Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States. Harvard University Press.
- Hogan, A., 2015. Network ethnography and the cyberflaneur: evolving policy sociology in education. *Int. J. Qual. Stud. Educ.* 29 (3), 381–398.
- Kamat, S., Anne Spreen, C., Jonnalagadda, I., 2016. Profiting From the Poor: The Emergence of Multinational Edu-Businesses in Hyderabad, India. *Education International*.
- Kamat, S., 2004. The privatization of public interest: theorizing NGO discourse in a neoliberal era. *Rev. Int. Polit. Econ.* 11 (1), 155–176.
- Kar, B., Sinha, S., 2021. Disadvantages and segregation through school choice: elementary education in India. *Econ. Polit. Wkly.* 56 (12), 33–39.
- Karopady, D.D., 2014. Does school choice help rural children from disadvantaged sections: evidence from longitudinal research in Andhra Pradesh. *Econ. Polit. Wkly.* 49 (51), 46–53.
- Kathpalia, V., 2016. Schooled in Socialism. *The Times of India*. February 23, Tuesday.
- Kingdon, G., 2015. Schooling Without Learning. *The Times of India*. August 26, Wednesday.
- Koyama, J., 2010. *Making Failure Pay: For-Profit Tutoring, High-Stakes Testing, and Public Schools*. University of Chicago Press.
- Kumar, S.M., 2010. Is there a case for school vouchers? *Econ. Polit. Wkly.* 45 (7), 41–45.
- Lagemann, E.C., 1992. *The Politics of Knowledge: The Carnegie Corporation, Philanthropy, and Public Policy*. University of Chicago Press.
- Lahoti, R., Mukhopadhyay, R., 2019. School choice in rural India: perceptions and realities in four states. *Econ. Polit. Wkly.* 54 (49), 51–57.
- Levin, H.M., 1987. Education as a public and private good. *J. Public Manag. Anal.* 6 (4), 628–641.
- Levin, H.M., 2001. Studying privatization in education. In: Levin, H.M. (Ed.), *Privatizing Education: Can the School Marketplace Deliver Choice, Efficiency, Equity, and Social Cohesion?*.
- Lingard, B., 1993. The changing state of policy production in education: some Australian reflections on the state of policy sociology. *Int. Stud. Sociol. Educ.* 3 (1), 25–47.
- Lipman, P., 2004. *High Stakes Education: Inequality, Globalization, and Urban School Reform*. Psychology Press.
- Lipman, P., 2011. *The New Political Economy of Urban Education: Neoliberalism, Race, and the Right to the City*. Taylor & Francis.
- Lubienksi, C., 2005. Public schools in marketized environments: shifting incentives and unintended consequences of competition-based educational reforms. *Am. J. Educ.* 111 (4), 464–486. <https://doi.org/10.1086/431180>.
- Maxwell, C., Yemini, M., Engel, L.C., Lee, M., 2020. Cosmopolitan nationalism in the cases of South Korea, Israel and the US. *Br. J. Sociol. Educ.* 41 (2), 845–858.
- Miglani, N., Burch, P., 2021. Education Reform Imaginaries: Mapping -scapes of Philanthropic Influence. *Discourse: Studies in the Cultural Politics of Education* 42 (5), 682–698. <https://doi.org/10.1080/01596306.2020.1836747>.
- Moeller, K., 2018. *The Gender Effect: Capitalism, Feminism, and the Corporate Politics of Development*. University of California Press.
- Moeller, K., 2020. Accounting for the Corporate: An Analytic Framework for Understanding Corporations in Education. <https://doi.org/10.3102/0013189X20909831>. (Accessed 27 April 2020).
- Molnar, A., 1996. *Giving Kids the Business: The Commercialization of America's Schools*. Westview Press.
- Münch, R., 2020. *Governing the School Under Three Decades of Neoliberal Reform: From Educocracy to the Education-Industrial Complex*. Routledge.
- Nambissan, G.B., Ball, S.J., 2010. Advocacy networks, choice and private schooling of the poor in India. *Glob. Netw.* 10 (3), 324–343.
- Nambissan, G.B., 2012. Private schools for the poor: business as usual? *Econ. Polit. Wkly.* 47 (41), 51–58.
- Nambissan, G.B., 2016. *Poverty, Markets and Elementary Education in India, TRG Poverty and Education Working Paper Series*, Max Weber Stiftung.
- Phillipson, B. (Ed.), 2008. *Low-Cost Private Education: Impacts on Achieving Universal Primary Education*. Commonwealth Secretariat, London.
- Posey-Maddox, L., 2016. Beyond the consumer: parents, privatization, and fundraising in US urban public schooling. *J. Educ. Pol.* 31 (2), 178–197.
- Rao, S.S., 2017. Editorial. *Contemporary Education Dialogue* 14, 1–4.
- Reckhow, S., Snyder, J.W., 2014. The expanding role of philanthropy in education politics. *Educ. Res.* 43 (4), 186–195.
- Robertson, S.L., Bonal, X., Dale, R., 2002. GATS and the education service industry: the politics of scale and global reterritorialization. *Comp. Educ. Rev.* 46 (4), 472–496.
- Robertson, S.L., Mundy, K., Verger, A. (Eds.), 2012. *Public Private Partnerships in Education: New Actors and Modes of Governance in a Globalizing World*. Edward Elgar Publishing.
- Schwittay, A., 2006. *Digital Citizens, Inc.: Producing Corporate Ethics, Flexible Networks and Mobile Entrepreneurs in the Global Marketplace* [Unpublished doctoral dissertation]. University of California, Berkeley.
- Scott, J. (Ed.), 2005. *School Choice and Diversity: What the Evidence Says*. Teachers College Press.
- Scott, J., 2009. The politics of venture philanthropy in charter school policy and advocacy. *Educ. Pol.* 23 (1), 106–136.
- Shippis, D., 2006. *School Reform, Corporate Style: Chicago 1880–2000*. University Press of Kansas.
- Smith, F., Hardiman, F., Tooley, J., 2005. Classroom interaction in private schools serving low-income families in Hyderabad, India. *Int. Educ. J.* 6 (5), 607–618.
- Srivastava, P., 2013. Low-fee private schooling: issues and evidence. In: Srivastava, P. (Ed.), *Low-Fee Private Schooling: Aggravating Equity or Mitigating Disadvantage*, pp. 7–35.
- Star, P., 1988. The meaning of privatization. *Yale Law Pol. Rev.* 6 (1988), 6–41.
- Steiner-Khamisi, G., 2018. Businesses seeing like a state, governments calculating like a business. *Int. J. Qual. Stud. Educ.* 31 (5), 382–392.
- Tarlau, R., Moeller, K., 2020. “Philanthropizing” consent: how a private foundation pushed through national learning standards in Brazil. *J. Educ. Pol.* 35 (3), 337–366.
- Tompkins-Stange, M.E., 2016. *Policy Patrons*. Harvard Education Press.
- Tooley, J., 2009. *The Beautiful Tree: A Personal Journey Into How the World's Poorest People are Educating Themselves*. Penguin Viking.
- Verger, A., Fontdevila, C., Zancajo, A., 2016. *The Privatization of Education: A Political Economy of Global Education Reform*. Teachers College Press.

- Verger, A., Steiner-Khamisi, G., Lubienski, C., 2017. The emerging education industry: analysing market making in education through market sociology. *Glob. Soc. Educ.* 15 (3), 325–340.
- Verger, A., 2020. Privatisation. In: Ritzer, G., Rojek, C. (Eds.), *The Blackwell Encyclopaedia of Sociology*. John Wiley & Sons Ltd., pp. 1–3
- Watkins, W.H., 2001. *The White Architects of Black Education: Ideology and Power in America, 1865–1954*. Teachers College Press.
- Whitty, G., 1997. Creating quasi-markets: a review of recent research on parental choice and school choice in three countries. *Rev. Res. Educ.* 22, 3–47.

The geopolitics of climate change and education policy

Marcia McKenzie, Melbourne Graduate School of Education, University of Melbourne, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	597
Micro, meso, and macro sociopolitics of climate change inaction	597
Where and how climate change can be included in education policy	598
Conclusion	599
Acknowledgment	600
References	600

Introduction

Definitions of “geopolitics” center the study of geographical or spatial factors in relation to power and politics, noting the varied ways the term has been used across disciplines and historically. Despite the term apparently being coined in aid of imperialist and racist expansion, it has been at least partially recuperated inside and outside of geography through more critical geopolitical approaches, including in relation to accelerated globalization in the late 20th century (Dodds, 2007). Human-induced global climate change has resulted from the mobilities and associated carbon emissions of that globalization – including the spread of market-based cultural systems, and increased mobilities enabled by fossil fueled technological and travel infrastructures (Urry, 2007). “Uneven development” notwithstanding (Smith, 1984), global fossil fuel driven economic and technological systems have ushered in a new geological epoch. Humans are now changing the climate of the planet, and along with it, the planet’s varied geographies and their habitability for humans and other species.

The latest Intergovernmental Panel on Climate Change report (IPCC, 2022) offers a grim summary of the geopolitical impacts we are already locked into over the next decade – millions more living in extreme poverty, lack of food and water security, sea level rises, increased mortality due to extreme weather events, biodiversity loss, and more. It also outlines the prospects of not taking significant future action to reduce and adapt to climate change, including the loss of habitability of the planet for increasing numbers of species, including humans. Despite this, many communities and countries remain stuck in various forms of climate denial and inaction, supported by policy and political systems that are failing to address climate change within the timeframes needed (Ford and Norgaard, 2020).

Education and communication are articulated in the original United Nations Framework Convention on Climate Change (UNFCCC, 1992) and again in the Paris Agreement (UNFCCC, 2015) as key to developing the social and political will for increased action. Despite this, many national education remain overfocused on student academic achievement and workforce readiness (Gorur, 2017), prioritizing individual and national economic gain over planetary health. Like rearranging deck chairs on the Titanic, our collective inability to process and respond more adequately to the global climate crisis including in education is a global issue in urgent need of political and policy action. This paper overviews some of the sociopolitical factors at the heart of this inadequate response to the geopolitical shifts of climate change, and implications for education policy.

Micro, meso, and macro sociopolitics of climate change inaction

In Rich’s (2013) novel *Odds Against Tomorrow*, two main characters have inverse reactions to the physical and social realities of the world post dramatic climate change. Mitchell Zukor addresses his “real, hot, viral” fear (p. 25) by researching and predicting possible climate change disaster scenarios. In contrast, the ill Elsa Bruner builds an idyllic eco-retreat, living in a “permanent condition of hopeful, childlike, brainless bliss” (p. 62). Through their written correspondence, Rich’s characters weigh the relative merits of cultivating rational fear versus fearless curiosity in the face of death. In contrast to these opposing reactions to climate catastrophe, in the novel most New Yorkers “carried on with their usual activities, pantomiming quotidian normalcy, as if nothing were wrong” (p. 102).

In her research on living in denial of climate change, Norgaard (2011; with Ford and Norgaard, 2020) investigates psychological and sociological responses to climate change. She adopts the term “double reality” to describe the disjuncture between the collectively constructed sense of normal everyday life, and the troubling knowledge of the causes and impacts of climate change. Like Nathaniel Rich’s New Yorkers, the participants in Norgaard’s (2011) study manifested signs of such a double reality, as they recognized the impacts of climate change, but nonetheless remained in a state of inaction and denial. A number of other studies have found that increased levels of knowledge about climate change have a negative effect on concern and sense of responsibility (Kellstedt et al., 2008; Krosnick et al., 2006, cited in Norgaard, 2011, p. 2). Some have linked this with the inability to perceive issues over which one has little control, as matters of concern. For example, Lifton (1982) researched “psychic numbing” as a response to the possibility of nuclear catastrophe – with numbing coming not from a particular traumatic event, but from a crisis of meaning: “We know that our lives can end at any moment, yet we live as though we do not know this.” (Norgaard, 2011, p. 4). As a result we

perform “double lives,” or what Norgaard calls a “double reality,” in which everyday normal life coexists with the troubling knowledge of climate change and its impacts on human and other life forms.

Norgaard draws on the sociology of cognition to explore how we are socialized to pay attention to different information on climate change through shaped systems of perception and attention: whether climate change “is considered morally offensive or not is a function of whether it is inside or outside socially defined limits of concern” (Norgaard, 2011, pp. 5–6). This social shaping of our understandings and thought patterns has been termed the “social organization of denial,” in other words, “we learn what to ignore, and only then does its irrelevance strike us as natural or “logical” (Zerubavel, 1997, p. 47). Culture shapes such social organization, not by providing guiding values, but through assemblages of “symbols, stories, rituals and world-views,” including via emotion and affect (Norgaard, 2011). In the case of climate change, individuals collectively distance themselves from information because of cultural norms of emotion, conversation, and attention, and through using an existing cultural repertoire of strategies in the process, such as the perpetuation of national or cultural narratives which are incongruent with action on climate change.

Power and politics drive these habits of attention and action that constitute our cultural norms of social organization. As Norgaard (2011), suggests, “Cognition, awareness, and denial emerge from a process that spans micro-, meso-, and macrolevels of social structure. At the microlevel, we see the role of individual emotions and interactions. At the mesolevel of culture, norms of behavior from conversation, feelings, and attention shape and reinscribe what is considered “normal” to think about, talk about, and feel. And these local cultural norms are in turn connected to larger political economic relations.” In her book on “neoliberal exceptionalism and the culture of uncare,” Weintrobe (2021) suggests that the roots of the climate crisis lie in part in a psychology of exceptionalism that, she argues, became more pervasive due to the spread of neoliberal politics globally over the past half century. This includes characteristics such as seeing oneself or one’s group as ideal, as entitled to the power and possessions one craves, and not feeling morally troubled about this (p. 19). She describes this as enabling increasing numbers of people to live in a “bubble-like psychic retreat from reality” (p. 23). This “group denial” can be attributed with being responsible for the slow and inadequate response to climate change (Weintrobe, 2021).

Other studies have identified that this group denial is often along political party lines – with political affiliation and ideology being the most significant predictors of climate change belief and action (e.g., Kahen, 2012). Still only 60% of the United States (US) population believes climate change is mostly caused by human activities (Sinatra and Hofer, 2021), with 2021 Gallup Poll findings showing a widening gap between Democrats and Republicans: 67% of Democrats “now consider global warming a serious threat to themselves,” versus 11% of Republicans (a 56 point difference, up from a 38 point difference in 2016, and 16 points in 2001) (McKenzie, 2022). While many studies on the increasing polarization in views on climate change have focused on the US, other countries also have growing segments of climate contrarianism and other forms of science denial, including along political party lines (Urman, 2020). For example, in Australia, Liberal and National party supporters are less likely to accept the scientific consensus on anthropogenic climate change than those who identify with the Green and Labor Parties (Tanter, 2021). The growing polarization, including on climate change, has been linked to a range of factors, including increased globalization and economic disparity, a growing financial elite and associated think tanks, and increased misinformation shared through social media platforms and communities (Moffit, 2016; Zembylas, 2021). For example, contrarian blogs and social media sites have been found to be linked to the funding flows and messaging of prominent climate denialist think tanks and private donors (Coan et al., 2021), with public perception on climate change in part shaped through these digital spaces (McKenzie, 2022; Papacharissi, 2015).

All of this also applies to students and the communities at school and out-of-school that they belong to; and likewise the micro, meso, and macro forms of social organization, or political affiliations and “denial groups” that families, teachers, and education policy makers are embedded within and how these shape their responses to climate change. Pilot research suggests that the personal sentiments and beliefs of policy makers may be particularly key to whether climate change is included in education policy (McKenzie, 2017; Pitton and McKenzie, 2020). More broadly, there has been minimal research to date on how policy makers’ social and affective orientations to issues may influence their policy making. Cooper (2009) also pointed to this gap of attention in policy studies, suggesting a need for more research on the “messy” emotional influences on policy making. He suggested this could include a focus on the ways that broader societal structures of feeling shape the orientations of policy makers to the issues on which they are decision making (Williams, 1961; see also Armstrong, 2004). Some studies have explored whether there is shared “habitus” or more common dispositions among policy makers. For example, Lingard et al. (2015) identified that common qualities in a sample of education policy makers were being middle class, high modernist, cosmopolitan, and scientific in orientation. A few other studies have examined the broader sociopolitical factors that may influence policy making, such as fear for the future or anxiousness about economic precarity (Borén et al., 2021), as well as the use of emotion in policy texts (Jupp et al., 2016). However, the education policy literature has yet to fully examine the sociopolitical lives of policy makers for how these may influence their policy making, including in relation to climate change.

Where and how climate change can be included in education policy

Focusing on forms of climate change denial and inaction more broadly, Candis Callison (2014, 2020) discusses how, why, and when climate change *does* come to matter and what that looks like across a range of populations and geographies. In climate change ethnographies of northern Inuit communities, a business sector, a Christian association, and other forms of social organization, she found that when “climate change” is engaged, it occurs through the frames and language of that demographic. Thus, rather than treating climate change as a matter of fact, or as merely coalesced “scientific consensus,” Callison inquires into how various groups

translate the issue of climate change into their own vernacular, defining climate change within and for their own group. This involves, according to Callison, “an assemblage of institutions, material training, disciplining mechanisms, and modes of speech” (p. 15). She writes, “The fact of climate change does not advocate, on its own, for taking action now or in the future – it is the presentation of the facts, their socialization and communality, and inherent moral and ethical dictates that determine the need for immediate action” (p. 15). Through these means, Callison suggests climate change can shift from being a “matter of fact” to being a “matter of concern” (Latour, 2004). In other words, climate change is more likely to matter to individuals, including policy makers, when it is framed in the values, narratives, and priorities of the affinity groups to which they belong. Likewise, Ford and Norgaard (2020) point to how dialog and other cultural tools can decrease the dissonance experienced in encountering climate change, and thus open up greater possibility for social engagement.

While still an emerging field, research suggests that climate change education can engage these considerations through a more holistic approach. This entails including a focus on knowledge, or the cognitive (e.g., climate science literacy), but also on socio-political considerations (e.g., group denial, climate grief) and an action-orientation. Understanding the action that can be taken, and becoming involved in that action, can help counter emotions of denial, fear, or loss. A range of interdisciplinary research also suggests that framing climate change education messages in ways that are locally relevant and meaningful can help in overcoming climate change denial or inaction (Hargis and McKenzie, 2021). Such framings are important given that social group and political affiliation are the largest predictors of climate denial and inaction (Kahan et al., 2012). A review of the literature of climate change education also found that deliberative discussions, addressing misconceptions, and implementing action projects are promising strategies for climate change education (Monroe et al., 2019). A whole institution approach to climate change education involves incorporating climate change education within overall institutional governance (e.g., strategic plans), teaching and learning (e.g., curricula across all subjects), facilities and operations (e.g., reduce building and data energy emissions), community partnerships (e.g., collaborations with municipalities or Indigenous communities), and research (at the higher education level) (Hargis and McKenzie, 2021).

Addressing colonialism and climate justice are also central to climate change education. This includes given the land-based impacts of extractivism, and that those most impacted by climate change globally are often those who have contributed the least. Elsewhere, myself and Cree colleague Alex Wilson have outlined the importance of land-informed policy in education when engaging with climate and sustainability issues (McKenzie and Wilson, 2022). Rather than simply turning to globally mobile approaches, such as the UN Sustainable Development Goals (SDGs) which in some cases can prevent or undermine progress, local approaches can be land-informed, including in relation to Indigenous knowledge and priorities. Versus 17 SDGs, from which an institution can pick and choose which Goals to prioritize and report on for ranking systems, locally developed Indigenous knowledge principles can be more holistic and interwoven. This can offer a policy path that is described by Indigenous policy makers as both decolonizing and furthering of environmental sustainability, as well as advancing Indigenous cultural and language survivance and resurgence (McKenzie and Wilson, p. 13; see also McKenzie et al., 2021).

While education is increasingly recognized as key to achieving the necessary climate action globally within the necessary timeframes (IPCC, 2022; UNFCCC, 1992, 2015), still relatively few countries currently have a focus on climate change in their education systems, or education in their climate change approaches, indicating a critical gap in education and environment policy making (McKenzie, 2021; UNESCO, 2021a). For example, national curriculum frameworks from 100 countries (52% of UNESCO member states) were analyzed for inclusion of climate change content, with almost half (47%) of the countries making no reference to climate change (UNESCO, 2021a). When climate change content was mentioned, the inclusion was often cursory. A systematic review of existing studies also suggests that climate change education is not often included in education policy and curricula (UNESCO, 2021b). When addressed, it is often through soft governance mechanisms, such as curricular guides, versus policy or curriculum (UNESCO, 2021b). The type of inclusion tends to focus on learning “about” climate change (within the subjects of science and social studies) or technological solutions, rather than developing sociopolitical and action capacities (UNESCO, 2021b). This study also suggests that “justice” and “Indigenous knowledge” were infrequently engaged in relation to climate change in the curriculum and policy documents reviewed (UNESCO, 2021b). Likewise, a recent review of education policy and education standards across all 50 US states and Washington DC found that only four states (Indiana, Minnesota, North Carolina, Oregon) linked Indigenous knowledge with climate change education content (MECCE Project and NAAEE, 2022).

Conclusion

While countries are making progress toward including climate change in education policy, much more work remains to be done to further the quantity and quality of climate change education globally. To date, there has been a lack of research to understand the reasons for these gaps in education policy making on climate change (Aikens et al., 2016). However, given the broader literature on the sociopolitical factors in climate change belief in action, it seems engaging further with the micro, meso, and macro understandings and contexts of policy makers seems key to increasing the extent and quality of focus on climate change in education policy.

While global policy initiatives are championing an increased focus on education policy making on climate change (UNESCO and UNFCCC, 2016), the national and subgroup politics of policy makers and schools are also important. Broader trends in political polarization, social media influences, and science denial are barriers to policy makers, communities, and families acting to include climate change in education policy and practice. Education, in turn, is essential for helping build the social and political will in the broader public to address this urgent issue of the 21st century.

Acknowledgment

Support from the Social Sciences and Humanities Research Council of Canada is acknowledged, as well as the contributions of Ph.D. Research Assistant Kristen Hargis, Project Manager Nicola Chopin, and other members of the MECCE Project research team (www.mecce.ca).

References

- Aikens, K., McKenzie, M., Vaughter, P., 2016. Environmental and sustainability education policy research: a systematic review of methodological and thematic trends. *Environ. Educ. Res.* 22 (3), 333–359. <https://doi.org/10.1080/13504622.2015.1135418>.
- Armstrong, D., 2004. Emotions in organisations: disturbance or intelligence? In: Huffington, C., Halton, W., Armstrong, D., Pooley, J. (Eds.), *Working below the Surface: The Emotional Life of Contemporary Organisations*. Karnac Books, pp. 11–28.
- Borén, T., Grzyś, P., Young, C., 2021. Policy-making as an emotionally-charged arena: the emotional geographies of urban cultural policy-making. *Int. J. Cult. Pol.* 27 (4), 449–462. <https://doi.org/10.1080/10286632.2020.1792891>.
- Callison, C., 2014. *How Climate Change Comes to Matter: The Communal Facts of Life*. Duke University Press.
- Callison, C., 2020. The twelve-year warning. *ISIS. J. Hist. Sci. Soc.* 111 (1), 129–137. <https://doi.org/10.1086/707823>.
- Coan, T., Boussalis, C., Cook, J., Nanko, M.O., 2021. Computer-assisted detection and classification of misinformation about climate change. *SocArXiv*. <https://doi.org/10.31235/osf.io/crxfm>.
- Cooper, A., 2009. “Be quiet and listen”: emotion, public policy and social totality. In: Sclater, S.D., Jones, D.W., Price, H., Yates, C. (Eds.), *Emotion: New Psychosocial Perspectives*. Palgrave Macmillan UK, pp. 169–182.
- Dodds, K., 2007. *Geopolitics: A Very Short Introduction*. Oxford University Press.
- Ford, A., Norgaard, K.M., 2020. Whose everyday climate cultures?: environmental subjectivities and invisibility in climate change discourse. *Clim. Change* 163, 43–62.
- Gorur, R., 2017. Statistics and statecraft: exploring the potentials, politics and practices of international educational assessment. *Crit. Stud. Educ.* 58 (30), 261–265.
- Hargis, K., McKenzie, M., 2021. Responding to climate change education: a primer for K–12 education. Sustainability and Education Policy Network. University of Saskatchewan.
- IPCC, 2022. IPCC, 2022: climate change 2022: impacts, adaptation, and vulnerability. In: Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegria, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), *Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press (in press). SPM.D.5.3.
- Jupp, E., Pykett, J., Smith, F.M. (Eds.), 2016. *Emotional States: Sites and Spaces of Affective Governance*. Routledge.
- Kahan, D.M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L.L., Braman, D., Mandel, G., 2012. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nat. Clim. Change* 2, 732–735.
- Latour, B., 2004. Why has critique run out of steam? From matters of fact to matters of concern. *Crit. Inq.* 40, 225–248.
- Lifton, R.J., 1982. *Indefensible Weapons: The Political and Psychological Case Against Nuclearism*. Basic books.
- Lingard, B., Sellar, S., Baroutsis, A., 2015. Researching the habitus of global policy actors in education. *Camb. J. Educ.* 45 (1), 25–42. <https://doi.org/10.1080/0305764X.2014.988686>.
- McKenzie, M., 2017. Affect theory and policy mobility: challenges and possibilities for critical policy research. *Crit. Stud. Educ.* 58 (2), 187–204. <https://doi.org/10.1080/17508487.2017.1308875>.
- McKenzie, M., 2021. Climate change education and communication in global review: tracking progress through national submissions to the UNFCCC Secretariat. *Environ. Educ. Res.* <https://doi.org/10.1080/13504622.2021.1903838>.
- McKenzie, M., 2022. Global biopolitics of climate change: affect, digital governance, and education. In: Rizvi, F., Lingard, B., Rinne, R. (Eds.), *Reimagining Globalization and Education*. Routledge.
- McKenzie, M., Lewis, S., Gulson, K., 2021. Matters of (im)mobility: beyond fast conceptual and methodological readings in policy sociology. *Crit. Stud. Educ.* 62 (3), 394–410.
- McKenzie, M., Wilson, A., 2022. Sustainability as wild policy: mobile SDG interventions and land-informed policy in education. *Discourse: Stud. Cult. Politic. Educ.* <https://doi.org/10.1080/01596306.2022.2091519>.
- MECCE Project & NAAEE, May 2022. Mapping the landscape of K-12 climate change education policy in the United States Monitoring and Evaluating Climate Communication and Education Project and North American Association of Environmental Education. Retrieved from: <https://mecce.ca/publications/mapping-k12-climate-change-education-policy-in-the-us/>.
- Monroe, M.C., Plate, R.R., Oxarart, A., Bowers, A., Chaves, W., 2019. Identifying effective climate change education strategies: a systematic review of the research. *Environ. Educ. Res.* 25 (6), 791–812.
- Norgaard, K.M., 2011. *Living in Denial: Climate Change, Emotions, and Everyday Life*. MIT Press.
- Papacharissi, Z., 2015. *Affective Publics: Sentiment, Technology, and Politics*. Oxford University Press.
- Pitton, V.O., McKenzie, M., 2020. What moves us also moves policy: the role of affect in mobilizing education policy on sustainability. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2020.1852605>.
- Rich, N., 2013. *Odds against Tomorrow*. Macmillan.
- Smith, N., 1984. *Uneven Development: Nature, Capital, and the Production of Space*. Verso.
- Tanter, B., 2021. Climate Change Knowledge and Political Identity in Australia. *Sage Open*, pp. 1–11. <https://doi.org/10.1177/21582440211032673>.
- UNESCO, 2021a. *Learn for Our Planet: A Global Review of How Environmental Issues are Integrated in Education*. United Nations Educational, Scientific, and Cultural Organisation, Paris, France.
- UNESCO, 2021b. *Getting Every School Climate-Ready: How Countries are Integrating Climate Change Issues in Education*. United Nations Educational, Scientific, and Cultural Organisation, Paris, France.
- United Nations Education, Scientific and Cultural Organization, United Nations Framework Convention on Climate Change, 2016. *Action for Climate Empowerment: Guidelines for Accelerating Solutions through Education, Training and Public Awareness*. Author. <https://unesdoc.unesco.org/ark:/48223/pf0000246435.locale=en>.
- Urman, A., 2020. Context matters: political polarization on Twitter from a comparative perspective. *Media Cult. Soc.* 42 (6), 857–879. <https://doi.org/10.1177/0163443719876541>.
- Urry, J., 2007. *Climate change & society*, Polity.
- Weintrobe, S., 2021. *Psychological Roots of the Climate Crisis: Neoliberal Exceptionalism and the Culture of Uncare*. Bloomsbury.
- Williams, R., 1961. *The Long Revolution*. Penguin.
- Zembylas, M., 2021. *Affect and the Rise of Right-Wing Populism*. Cambridge University Press.
- Zerubavel, E., 1997. *Social Mindscales: An Invitation to Cognitive Sociology*. Harvard University Press.

Ecopedagogy and ecopedagogical literacy

Greg William Misiaszek^{a,b}, ^a Beijing Normal University, Faculty of Education, Institute of Educational Theories, China; and ^b Paulo Freire Institute, UCLA, United States

© 2023 Elsevier Ltd. All rights reserved.

Essence of ecopedagogy	602
World-Earth de-distancing	602
Freirean teaching for environmental praxis	602
Generative themes and dialogue	603
Utopias and hope	603
Reinventions	604
Ecopedagogy's uniqueness as an environmental pedagogy	604
Ecopedagogical literacy	604
Ecopedagogically reading <i>d/Development</i> and sustainability	605
Ecopedagogically reading <i>g</i> lobalizations	607
Ecopedagogically reading citizenships (local-to-planetary)	607
Ecopedagogically reading epistemologies and ecolinguistics	608
Ecopedagogically reading postdigitalism and post-truthism	609
Conclusion: critical comparing environmental pedagogies	610
References	611

Ecopedagogy is a collection of environmental pedagogies that emerged from, and is grounded in, the work of Paulo Freire (Kahn, 2010; Misiaszek, *in press*; Misiaszek, 2012; Gadotti, 2008c; Gutiérrez and Prado, 1989). Ecopedagogical teaching is for students to better understand the inseparability of environmental and social violence (socio-environmental connections); the politics of formal, nonformal, and informal education that falsely separates the two; and that “sustainability” must be planetary in scope (i.e., sustainability that includes all of Earth) (Gadotti, 2008a, 2008b; Misiaszek, 2020b).¹ As Freirean pedagogues teach through dialectically problematizing why oppressions take place (Gadotti, 1996; Morrow and Torres, 2019), ecopedagogues center problematizing on why unsustainable environmental violence takes place. The following two questions are key to continuously pose and re-pose contextually in ecopedagogical teaching. What are the local-to-global politics of why unsustainable environmental violence occurs? Who benefits and who suffers from false socio-environmental splitting? Inquiries on identifying negative consequences must encompass non-anthropocentric analysis with all of Nature, including when the term “who” is used. Coinciding with Freire (2000) problematizing the politics of education that justify oppressive acts, especially banking education models, ecopedagogues deconstruct how politics of (environmental) education justify acts of unsustainable environmental violence.

This chapter has three objectives. First, to define what is ecopedagogy and ecopedagogical literacy as an overall pedagogical aim of its teaching. The former will be done by explaining ecopedagogical teaching to critically unpack why environmentally violent acts occur, its planetary inclusiveness, and its Freirean essence. How ecopedagogy differs from other environmental pedagogies such as environmental education (EE) and education for sustainable development (ESD) will also be discussed. The importance of recognizing that these and other environmental pedagogies have diverse pedagogical framings and objectives that coincide and conflict with ecopedagogical ones will be explained. The latter aspect of defining ecopedagogical literacy coincides with this chapter’s second objective of delving into some key themes essential for teaching students to ecopedagogically read. These themes, as follows, were selected to coincide with this encyclopedia section’s topic “globalization and shifting geopolitics of education”: reading “development” and sustainability, reading processes of globalization, reading citizenship (local-to-planetary citizenship), reading through diverse ways of knowing and the politics of global epistemological dominance, and reading within our increasingly digitalized world. The third objective is arguing the need for ecopedagogies for transformative praxis from education for crucial paradigm shifts leading toward globally all-inclusive environmental justice *with* planetary sustainability.

¹The article “the” will not be used with “Earth” to not linguistically objectify Earth and will be upper-case. Nature will be upper-cased also with the same reasoning.

Essence of ecopedagogy

Ecopedagogy specifically focuses upon humans' interactions with the rest of Nature that leads to injustices and unsustainability. Sustainability is planetary in scope, in that balance and wellbeing is for all humans and all that exists outside the anthroposphere (i.e., all of Nature). There are endless complications in achieving planetary balance and is likely better conceptualized as a utopian goal, akin to a horizon or the North Star (see [Teodoro and Torres, 2007](#)), to give direction but highly unlikely to be fully achieved (utopia and education will be discussed more later).

Environmental issues are centered but are not isolated in ecopedagogy from all other happenings in our world. In the same way that teaching through critical race theories focus on racial issues specifically but do not ignore the intersectionalities with other forms of oppressions ([Ladson-Billings, 2009](#)), ecopedagogical work is largely done at these intersections. Falsely justifying environmental violence is largely due to systematically taught socio-historical prejudices, such as racism, and often leads to other forms of prejudices. In other words, learnt othering deprioritizes environmental oppressions of the Other, and objectifying the rest of Nature instills its existence as solely for humans' utilization and commodification. The ecopedagogical aspect of focusing on "globally all-inclusive" justice means that neither socially constructed differences (e.g., culturally, tradition-wise, linguistically, epistemologically) nor physical distances can justify unsustainable environmental oppressions onto others. This is especially important for socio-historically othered populations (e.g., Indigenous, global South, non-whites; these populations will be termed as upper-cased Other(s) in this chapter).

Planetary sustainability *with* goals of global socio-environmental justice is limitlessly complicated and often these two objectives can conflict with one another. For example, unjust food issues can sometimes be alleviated in the short-term by unsustainable farming practices (see [Epstein-HaLevi et al., 2018](#)) or environmental regulations can result in temporary labor injustices. Although these complex socio-environmental conundrums have no simple solutions, democratic, critical dialogue in ecopedagogical spaces on the often-hidden consequences and politics of environmental violence unveils the importance, crucial targeting, and challenges (i.e., "limit situations" in Freirean terminology) to construct possible solutions for needed environmental transformation. This includes interrogating previously taught ideologies that *distance* environmental harms and, in turn, deprioritize negative affects upon the othered "them" and/or objectified nature,² as will be discussed next.

World-Earth de-distancing

Teaching to disrupt education that *distances* us from one another and from the rest of Nature (world-Earth distancing) is essential for any environmental pedagogy to be successful ([Misiaszek, 2018](#)). Distancing is done in various aspects such as geographically, time-wise, and othering. Examples of world-Earth distancing, respectively, are the following: exportation of environmentally harmful production and recycling to the global South; ignoring global warming because effects will occur in the "distant" future; and dehumanizing othered populations, such as through Orientalism ([Said, 1979](#)), that justifies socio-environmental oppressions placed upon those seen as not fully human – i.e., the Other.

Teaching for World-Earth distancing falsely detaches us (i.e., humans) from the rest of Nature and, in turn, devalues and objectifies all which is not human. Neoliberalism, arguably the most prominent tool for world-Earth distancing, commodifies all of Nature as the possibilities for, as Postma argued (2006), profiteering without concern for injustices or unsustainability felt beyond one's own private sphere. Within public spheres, neoliberalism's innate connections to sustaining hegemony above all else, especially global hegemony ([Chomsky, 1999](#); [Harvey, 2005](#)), distances environmental ills placed upon others and the rest of Nature. Within ecopedagogy, how neoliberalism devalues injustices and unsustainability from environmental violence and the politics of neoliberal ideological indoctrination in education is occurring, are central concerns.

Freirean teaching for environmental praxis

Freire wrote on ecopedagogy in his later works (see [Freire, 1997, 1998, 2004](#)) and several Freireans have constructed ecopedagogies by reinventing his work. Reinvention was a key tenet of his work and will be discussed more later in this chapter. His unfinished writings before his untimely death was to focus on ecopedagogy and several passages from this work were posthumously published in his book *Pedagogy of Indignation* ([Freire, 2004](#)) ([Misiaszek and Torres, 2019](#)). Fully elaborating on Freire's essence within ecopedagogy is impossible within any publication length; this chapter will focus on emergent ecopedagogical praxis from teaching through democratic problem-posing dialogue, utopic hope, and contextual reinventions.

Goals of students' actions emergent from teaching vary between pedagogies. Freirean teaching is action-based pedagogies with the goal of ending oppressions ([Gadotti, 1996](#)); ecopedagogues teach not just for understanding environmental ills but for critical reflexivity that lead students to determine necessary actions to disrupt the ills ([Gadotti, 2008b](#)). Teaching for environmental wellbeing is the focus of all environmental pedagogies, although pedagogical goals for praxis is often superficial or absent. Many environmental pedagogies are absent of this described, or similar, goals of ecopedagogical praxis because critical reflexivity through dialectics, diverse epistemologies, and theorizing helps to actively counter hegemony that requires unsustainable environmental

²The lowercasing of nature here indicates the objectification of Nature.

violence to exist (Misiaszek, 2020b). Non-critical, non-dialectic *banking education* models that Freire (1998, 2000) vehemently opposed would be such pedagogies.

Ecopedagogues teach students how they can delve into the political aspects of environmentally violent acts for transformative praxis that takes aim at the deeper roots of these actions. This includes unpacking the politics of (anti-)environmental teaching that influences students' interpretations of environmental violence.

Coinciding with Freirean pedagogy (Freire, 1992), ecopedagogues teach through reflexivity that all aspects of societies are transformable (e.g., economic, political, social). For example, countering commonsense that development "naturally" adheres to neoliberalism needs critical problematizing in education (Torres, 2011), including ecopedagogical problematizing. All schooling is for individual and societal development, but how does dominant ideologies of "development" often justify environmental violence? Closely linked to this problem-posing of development is education for citizenship. How does citizenship lead to solidarity but also deems "non-citizens" as Others (e.g., (neo)colonized, Indigenous persons) and/or geographic solidarities (e.g., city-state, nation-state) leads to largely ignoring distanced ill effects from unsustainable environmental violence (Misiaszek, 2015, 2018). Teaching for both development and citizenship will be further problematized throughout this chapter. Many shallow environmental pedagogies help to sustain/intensify environmental violence by diverting attention away from its root causes and negative effects, as well as instilling that radical transformations are unthinkable (Misiaszek, 2020d). *Pedagogies on environmentalism* may do this purposely or not.

Pedagogies on environmentalism are different from environmental pedagogies, as they might have anti-environmental goals. For example, a transnational fossil fuel corporation giving underlying messages that global warming is still debatable or giving funding for selective teaching resources and educational research that delegitimize environmental concerns worldwide. Or a course on sustainability that deprioritizes environmental wellbeing in the Global South by prioritizing the global North's *Development* (upper-casing of "D" will be discussed later in this chapter).

Generative themes and dialogue

Freire (2000) argued that praxis emerges from critically determining what is currently happening, what "should" be happening, and what are the obstacles/barriers of not achieving the latter, in which he named *limit situations*. This is accomplished through problem-posing education in which both teachers and students democratically determine *generative themes* to teach and learn through, such as pollution which may be affecting those directly in a particular ecopedagogical space. The epistemologies, meanings, and experiences of everyone in the space is brought into the teaching and learning while problem-posing the themes. In addition, bringing in others' diverse knowledges and understandings from outside the learning space is also essential, especially from those who are most negatively affected and whose voices are too-often not heard. For example, bringing in the voices of Indigenous Amazon populations when discussing the local-to-planetary effects of deforestation that, in part, increases global warming.

Ecopedagogical spaces of democratic dialogue is opposite to banking spaces where students are taught *at* because only the teacher has legitimate, useful knowledge. Ecopedagogues create safe spaces so students can have what Friere (1985a) termed as *authentic dialogue*, rather than being meekly malleable by only repeating the teachers' ideologies, knowledges, and opinions. Opposite to systematically structured passive, falsely neutral classrooms that are too prevalent in schooling (Apple, 2004), ecopedagogical spaces, coinciding with critical/Freirean spaces, are full of conflict but are structured to remove any and all types of violence. The diversity of understandings, epistemologies, and experiences deepens and widens everyone's learning within the spaces. Conflict stems from the humanness of our individualized personhood in authentically problematizing the themes together, but also emergent commonalities help to understand others' perspectives. As will be discussed later toward the end of this chapter, intensifying post-truthism threatens such rich conflictual dialogue.

Utopias and hope

Teaching to imagine futures of environmental justice and sustainability is essential for praxis because it allows imaginaries of "what 'should' be happening," as discussed previously.³ Freire (1992, 2000) emphasized the need to teach through reflexivity that our world is *unfinished*, as opposed to banking educators teaching that the world is *finished* with current oppressions as "natural" and non-transformable. A single unsustainable future of socio-environmental injustices and unsustainability instilled within banking (environmental) teaching ceases all possibilities of truly transformative praxis. For example, neoliberal-based teaching normalizes the prioritization of profiteering within one's own private sphere above all else, without environmental concerns or cares (Postma, 2006). Ecopedagogical praxis cannot occur because thinking and imagining outside of neoliberal ideologies is viewed as abnormal and foolish.

Hope and utopia in (environmental) teaching is not "fanciful" dreaming but rather imagined possibilities (or *imaginaries*) grounded in critically understanding the current realities of oppressions (Romão, 2007; Freire, 1992), as well as planetary unsustainability. Reflexivity on the difficulties and hardships that are involved in overcoming limit situations of environmental ills is essential within (ecopedagogical) praxis. The use of diverse critical theories provides lenses to identify and better understand the links between environmental violence and socio-historical oppressions and the difficulties in eliminating limit situations while constructing utopian imaginaries. Freire (2000) expressed that our histories construct our futures; however, he is emphasizing that reflexivity must be through socio-historical lenses rather than any notion of fatalism.

³Futures is pluralized to emphasize the utopian notion that there is more than a single possible future. See <https://en.unesco.org/futuresofeducation/>.

How can students begin to understand unequal environmental ills in the global South without teaching histories of colonialization that have and continue to exploit natural resources and export environmental ills for sustaining Northern profiteering and hegemony? Simply, we cannot. These oppressive processes continue today by globalizations that can be rightfully termed as neocolonialism. Teaching to systematically use critical theories deepens and widens students' understandings of environmental violence through diverse positionalities, histories, and contexts, as well as understanding Nature holistically through non-anthropocentric, planetarism (e.g., Red Pedagogy, see [Grande, 2009](#); ecofeminism, see [Warren, 2000](#)).

Reinventions

Reinvention is necessary so that the "old is capable of remaining new when it remains faithful through time to the experience of original and founding intuitions and inspirations" ([Freire, 1998](#)). Freire argued that reinvention leads to transformative reflexivity and resulting praxis "through the construction of a new nexus between core generative concepts, the diagnosis of emergent historical realities, and testing and refinement of such reconfigurations through new experimental practices" ([Morrow, 2019](#)). These aspects are essential for reinventing Freire in (re)constructing (eco)pedagogies from his work and constructing praxis from teaching. An example of the latter is [Freire \(1985a\)](#) response of countering neoliberal globalization is not "taking power" but 'reinventing power', in which [Morrow \(2019\)](#) views challenges as "twofold: constructing new kinds of more humanizing social relations, which in turn presupposes a better analytical understanding [of] domination and its relation to different forms of power." These aspects of reinventing power can be linked to Southern Theories work on power dynamics of epistemologies of the South/North which will be discussed later in this chapter.

[Gadotti \(2008a\)](#) wrote that the implementation of the UN *Decade of Education for Sustainable Development* (DESD, 2005–2014) must have ecopedagogical foundations, but that contextual reinvention is paramount. He opposed a rigid structure or, worse, a universal step-by-step teaching approach for "successful" implementation which is almost never successful. There is no easy method to Freirean (eco)pedagogy and Freire himself stopped using the term conscientization (*conscientização* in Portuguese) when he began witnessing people using as a non-contextual, over-simplified "method" ([Schugurensky, 2011](#)). This argument for needing contextuality holds true for success in achieving the current UN *Sustainable Development Goals* (SDGs) ([Klees, 2017](#); [Misiaszek, 2021](#)). Contextual reinventions are particularly important for othered populations in which, for example, neocoloniality through globalizations has systematically delegitimize their pedagogies and epistemologies, and de-prioritize "their" injustices and local environmental unsustainability in favor of hegemonic Development.

Ecopedagogy's uniqueness as an environmental pedagogy

Every pedagogy has specific foundations and ecopedagogy is no different. Although EE has specific foundations, it is frequently incorrectly used as interchangeable with the umbrella term of "environmental pedagogies." EE and ESD are the most common environmental pedagogies which have diverse, often-contradictory pedagogical practices, objectives, and goals between and within them. For example, [Sauvé \(2005\)](#) maps out 15 currents of EE with at least one-third coinciding with ecopedagogy. [Jin-kling and Wals \(2008\)](#) map out a contested terrain of ESD – from teaching through critical, democratic dialogue aligning closely with ecopedagogical foundations to teaching for enforcing sustainability that resembles, they argue, "Big Brother" in Orwell's 1984 novel. There are numerous other environmental pedagogies worldwide, including ecopedagogy, that both work within EE and ESD, and are stand-alone pedagogies. It is important to stress that EE and ESD, especially models grounded within Northern epistemologies, are not the only environmental pedagogies and legitimizing them globally as the only ways for environmental teaching is problematic. This includes Western-centric analysis that ignores histories of Southern and Indigenous environmental pedagogies by insinuating that Rachel Carson's book *Silent Spring* (1962) was the main catalyst that initiated environmental teaching everywhere.

ESD largely emerged from various criticisms of EE including the following: not "practical" enough, too conservation-based which ignored social issues, ignored environmental violence connections to social injustices, ignored urban environmental concerns, and disregarded unequal environmental impacts on marginalized populations that, in turn, sustains socio-historical oppressions (e.g., racism, neocoloniality, patriarchy, heteronormativity). It depends on what current(s) are being analyzed when discussing if, and how, these aspects are absent in EE. ESD specifically links human actions to sustainability, including environmentalism (i.e., environmental sustainability) as one of its various aspects categorizing and measuring (un)sustainability overall. This was, in part, a pedagogical move toward more concretely linking environmental violence with social violence.

Ecopedagogical literacy

As [Freire \(2000\)](#) focused on literacy teaching for reading the word to read the world, ecopedagogical literacy is widening this to reading the world as *part* of Earth.

The act of reading cannot be explained as merely reading words since every act of reading words implies a previous reading of the world and a subsequent rereading of the world. There is a permanent movement back and forth between “reading” reality and reading words ... reading the word is not only preceded by reading the world, but also by a certain form of writing it or rewriting it ... of transforming it by means of conscious practical action. For me, this dynamic movement is central to literacy.

(Freire, 1985b).

Ecopedagogical literacy is the ability to critically deconstruct the often-hidden politics of environmental violence that continues injustices and unsustainability (Misiaszek, 2018). It is reading environmental violence through local-to-global-to-planetary perspectives for deepened localized comprehension of environmental issues *with* widened understandings beyond the Local (e.g., national, global, and planetary perspectives) (Misiaszek, 2018, 2022). Ecopedagogical reading aligns with Freirean bottom-up approaches in which praxis to end unsustainable environmental violence must emerge *from* and be guided *by* those who are oppressed by the violence in question (see Gadotti, 1996; Freire, 2000).

There are three important aspects of ecopedagogical literacy needing to be pointed out. First, Freirean bottom-up approaches do not promote the absence of outside/widened understandings and perspectives beyond the Local, but rather opposes “top-down” approaches that ignore and silence the Local. In short, ecopedagogical reading of environmental violence is through local *and* widened understandings and epistemologies. For example, reading the polluting of a river is understanding the environmental impacts from local populations down-river from the pollution source together with experts in fields such as environmental sciences and policies, biology, and chemistry. Better understanding the effects from the pollution and reasons why it happened, as well as possible solutions to stop it, emerges from diverse voices and the taught ability to read the authenticity and politics of the voices to listen to or not.

Second, there are not binarily-determined oppressed:oppressor classifications for specific individuals or populations, but rather it is relational with everyone being both oppressed *and* an oppressor (Torres, 2015). Reading to determine the differing degrees of environmental oppressions due to local-to-global politics is crucial, as well as recognizing that ecopedagogical reading is without endpoints. Environmental ills will ultimately affect everyone negatively, but Beck’s (1992) statement that “hunger is hierarchical, smog is democratic” helps to initially unpack the complexities of differing positionalities, hegemonies, and environmental violence when determining degrees of oppression. However, Beck’s statement needs further unpacking. How does air pollution affect a professor like myself who can work from home and afford air filtration systems differ from migrant workers who work outside 10+ hours per day? What are the socio-historical privileges that I have being a white, heterosexual, male from the global North (e.g., from histories of coloniality, Northern epistemological dominance) that lessens the negative aspects of the same environmental event upon myself as compared to others? Reversely, how am I being oppressed by, for example, global neoliberal systems. Critically reading contextual oppressed: oppressor relativity is essential when characterizing a population or process, including how related terms are discussed or written. Although I do not emphasize this repeatedly in this chapter or any of my work, this aspect is entrenched whenever writing “the oppressed(ors),” oppressive, empowerment, or similar terminologies.

A common question is if the “bottom” is powerless how can they enact change through liberatory praxis from education? Freire (2000) answered this dilemma by utilizing, in part, Hegel’s Master-Slave dialect (Gadotti, 1996), arguing that “only as they discover themselves to be ‘hosts’ of the oppressor can they contribute to the midwifery of their liberating pedagogy.” Anti-environmental pedagogies succeed because their teaching justifies environmental injustices and unsustainability, and systematically veils oppressors as, corresponding to Freire’s quote, parasites. Thus, becoming a “pedagogy of the (environmentally) oppressed” – corresponding to Freire’s (2000) most famous book title and its overall theme. Resistance is lessened by teaching that oppressions are “natural” and “deserving,” which can only be countered in teaching for students’ conscientization of this perverted educational goal (Apple, 2004; Giroux, 2001).

Third, reading the limitations of ecopedagogical reading is crucial. Coinciding with feminist standpoint theories (Harding, 2006), ecopedagogical reading requires reflexivity that an absolute understanding of oppressions outside of one’s own experiences and positionalities is impossible. A US professor spending every waking moment reading, visiting, and conducting research on environmental violence affecting children in South Africa will not have the exact same understandings as a child living in Johannesburg. In addition, epistemological reflexivity is essential in ecopedagogical reading that coincides with Southern Theorists’ (Santos, 2018; Assié-Lumumba, 2017; Takayama et al., 2017) work, which will be discussed more later.

Ecopedagogically reading *d/D*evelopment and sustainability

Critically reading the “D” is paramount in determining what are the pedagogical groundings and goals of specific ESD models. Problematizing what is meant by “development” is central in ecopedagogy. Coinciding with Hardings’ (2006) usage of uppercased Sciences to denote globally dominant, top-down sciences as opposed to lowercased sciences that are bottom-up (e.g., Indigenous sciences) opposite to these characteristics, differentiating lower/uppercasing of *d/D*evelopment can be utilized in similar ways. Below are two passages that begin to describe *d/D*evelopment differences which will be unpacked throughout the rest of this chapter.]

[*d*/*Development* and ecopedagogy:] four defining factors of *esD* [and *Development* (capital and italicized '*D*')]: (1) neoliberal economics as the sole factor of development analysis; (2) deprioritizing economic justice concern by ignoring how development processes sustain/increase hegemony; (3) deprioritizing planetary sustainability for Earth's balance; and (4) local framings of development are disregarded for globally constructed ones 'from above' (e.g. Western *Development* models), denoted by the lower-case, underlined, and italicized '*d*'. [*"D"* or "*d*" not underlined and italicized does not specify the type of development.]

(Misiąszek, 2020c).

I utilize the lower-cased *development* and upper-cased *Development* to indicate, respectively, empowering versus oppressive, holistic versus hegemonic, just versus unjust, sustainable versus unsustainable, and many other opposing framings of who is included within "development" and framings of *d*/*Development* goals. There are no absolute origins or framings differentiating between *d*/*Development*, but rather the essence and outcomes of their framings ... Constructs of *development* that counters growth towards and emergent from *Development*.

(Misiąszek, 2020b).

Some key problem-posing questions are the following: whose development is focused upon and not; is the rest of Nature outside of humans considered in development and, if so, to what extent; who defines development that is globally characterized as "legitimate," and what aspects of development are prioritized or not (Misiąszek, 2020c, 2022)? Such questioning is increasingly important and complicated within our globalizing world as "local happenings are shaped by events occurring many miles away and vice versa" (Giddens, 1990).

It is important to note that oppression/empowerment is contextually framed and relative, aligning with the previous argument that oppressed/oppressor categorization is not binary. However, ecopedagogical reading is grounded on the goal for globally all-inclusive justice and planetary sustainability, including the reading of development. Although relative, ecopedagogically reading *development* is guided by this utopic goal in which critical reading others' (de-)development and (un)sustainability for all of Nature is crucial for praxis? This is essential in ecopedagogically reading development, as well as all terms linked to (anti-)environmentalism.

Ecopedagogical teaching of sustainability actively disrupts "commonsense" that justifies actions result in negatively affecting anyone worldwide and devasts Nature, and the accompanying fatalism encouraging the notion that "this is just how the world is." Ecopedagogues teach to critically understand sustainability as plural in goals, scope, and measurement; the politics of and between these three; and the connections with development. Planetary sustainability is the ultimate goal of environmentalism and its teaching, rather than ending all environmental violence. Although this may sound initially counterintuitive, its logic is from the ecopedagogical assumption that all human acts of environmental violence occur to benefit some population(s) (Misiąszek, 2018). For example, deep-ocean drilling for a thick, viscous substance would be a fool's task unless that substance provided energy when refined. Ecopedagogical reading includes determining if acts of violence are sustainable or not, and, inseparable to this, who suffers from them. There are continuums of environmentally violent actions, such as from deep-sea oil drilling given in the previous example, to turning on a computer powered by fossil fuels when plugged in or by a battery composed of mined natural resources.⁴ The reason for writing the phrase "unsustainable environmental violence" throughout this chapter is due to this logic.

Terms of sustainability framings vary, including the following: *environmental sustainability* which environmental pedagogies focus upon; *planetary sustainability* that grounds the importance of sustainability for all of Earth including and beyond anthropocentric concerns (framing most utilized within ecopedagogy); and *sustainable development* to balance acts of "progress" within the anthroposphere (specific goals of ESD, ESD is environmental if also prioritizing one or both of the other sustainability framing(s)). These and similar sustainability terminologies are often incorrectly used interchangeably; ecopedagogues teach to read what specifically is (de)prioritized and (de)legitimized when sustainability is uttered, and the underlying politics. Stating "sustainability" or even a specific sustainability framing is insufficient in ecopedagogical work, as sustainability also forms contested terrains within and between them. What is the scope (local-to-planetary) of sustainability goals? Who is included and who is not when measuring sustainability? Is social justice *with* peace included in being "sustainable?" How does economic and labor justice factor in on sustainability, or not, as Apple has emphasized that we should be reflexive of these and other justice issues connected with energy sources every time we turn on a light switch.

Planetary sustainability is balancing our subjectively constructed world (i.e., anthroposphere) with the objective, static laws of Nature guiding the rest of Earth.⁵ Freire (2000) argued that only humans can cognitively develop and transform the world with the ability for socio-historical reflexivity and forming praxis from dreaming, while the rest of Nature reacts, adapts, and evolves. The rest of Nature returns to equilibrium while humans, often to our own detriment, can counteract Nature's balance in actions determined by our reflexivity. For example, hurricanes and tornadoes result from Nature returning to equilibrium from unbalanced atmospheric pressure systems. This is not considered an act of environmental violence because there is not cognitive reflexivity that cause the high-speed winds. In the same vein, a non-human animal attacking a hiker is done instinctually from possibly being hungry or fears of being threatened, but without the reflexive abilities to understand how the hiker will suffer. This scenario is not one of injustice

⁴Michael Apple discussed these continuums of environmental violence in lecture series at Beijing Normal University (BNU), China in September 2013 and November 2019.

⁵Objectivity refers to the laws themselves and not humans' interpretations of them, which can be subjective.

upon the hiker because all non-human beings lack the reflexive ability of being (un)just (see Warren, 2000 on ecofeminism). These two examples become acts of environmental violence when our actions increase the frequency and/or intensity of these happenings, such as global warming increasing the number and strength of storms or altering animals' environments that cause them to attack humans more often. Ecopedagogues teach to unpack the connections of our actions with these and similar occurrences are too-often dismissed as "just a part of Nature."

Lastly, reading to determine the local-to-global constructed *baselines of sustainability* is crucial, as well as critically comparing the commonalities and differences between baselines' goals and politics. If today's socio-environmental conditions form the sustainability baseline – no further unsustainability will happen beyond what is happening today – then current oppressions from environmental violence will only continue. This is far from leading to all-inclusive environmental justice and planetary sustainability.

Ecopedagogically reading globalizations

The politics of global governance is difficult to unravel and hegemony is often sustained by globalizing agents veiling their influences,⁶ especially in education (Dale, 2007). Teaching to better understand how globalization influences environmental (in)justice and (un)sustainability is extremely complicated because the root sources, processes, and politics of environmental violence are purposely hidden. For example, transnational corporations such as ExxonMobil or Nestlé will not publicly state their unsustainable environmental actions, but they will tout their superficial environmental projects and funding as a smoke-and-mirrors technique. Today's intensifying globalization forms a contested terrain of processes *from above* and *from below* (i.e., plural globalizations as termed by Torres (2009)) which need students to be taught how to critically unpack who actually benefits and suffers from globalizations, as well as all the associated politics (Kellner, 2002; Torres, 2009). For example, reading through theoretical lenses of globalizations, ecofeminism, and ecoracism lenses helps to better understand how marginalized populations will be overwhelmingly affected more by global warming with an estimated 25 million to one billion environmental migrants by 2050 (International Organization for Migration (IOM), 2014).⁷ According to the IOM (2014) the actual number within this astonishing wide-range will depend on "economic growth and development, population growth, and governance" – all requiring ecopedagogical reading on globalizations' effects on these three.

Although globalization from above is diverse, there are certain common goals of the processes overall (Torres, 2009; Stromquist, 2002), including politics influencing environmental teaching (Klees, 2017; Misiaszek, 2022). Neoliberal globalization, which is the most common process from above, has individual profiteering, sustaining current hegemony, and monetarily measuring and ranking all aspects of society, including education, as its goals. Neoliberalism is complex, multifaceted, and has local-to-global-reaching tentacles (see Harvey, 2005; Klees, 2008). Reading to deconstruct globalizations that intensify anti-environmentalism is essential for praxis that disrupts them, as well as determining possibilities of reconstructing the processes to become globalizations "from below" for environmentalism (Misiaszek, 2012; Kahn, 2010). It must be emphasized that there is no simple approach to ecopedagogical teaching and reading, including false notions that everything from the Local or Indigenous is environmental and, opposingly, that all which comes from international agents of globalization is simply anti-environmental. Rigorous, critical reading and re-reading is needed.

Ecopedagogically reading citizenships (local-to-planetary)

Similar to globalizations, local-to-national-to-global-to-planetary citizenship spheres (i.e., citizenships) form contested terrains, especially the coinciding and conflictive intersectionalities between citizenships (Torres, 2017b; Misiaszek, 2015, 2018). Framing citizenships is well beyond geographic location, especially in education when discussing refugees (McNamara, 2007), Indigenous populations and coloniality (Grande, 2009; Abdi et al., 2015; Memmi, 1967; Said, 1979), and Nature beyond the anthroposphere (Gadotti, 2008b; Gutiérrez and Prado, 1989; Misiaszek, 2018). Problematizing how students are taught to determine their "fellow citizens" and how to treat everyone else (i.e., non-citizens), including the rest of Nature, begins to unpack the roles of citizenship education on (anti-)environmental actions. Most global citizenship education (GCE) models have everyone worldwide as fellow global citizens; however, if justice and sustainability are centralized pedagogical goals, or even goals at all, is dependent on the model in question (Misiaszek, 2018; Abdi et al., 2015). GC/E has numerous categorizations, including Lynette Shultz (2007) categories of teaching for neoliberalist, radical, and transformationalist GC. The two latter models can be viewed, and will be termed in this chapter, as *critical GC/E* models. Planetary citizenship education models inherently have planetary all-inclusiveness and these goals as foundational.

There is an innate conundrum in using citizenship terminology with being "global" or "planetary." Lee et al. (2004) argues that historically citizenship education was for populations to be more homogeneous (e.g., same ethical, morality, value groundings) as opposed to critical GCE which is centered on valuing cultural heterogeneity. In addition to this, planetary citizenship includes the quite radical incorporation of non-humans and Earth holistically. Torres (2017b) partially addresses this conundrum by arguing that there are shared *global commons* valued by everyone worldwide and provides the following three:

⁶Scholars differ on who are the agents of globalization. Stromquist (2002) names the following five: central and weak states, transnational corporations, mass media, global financial institutions, and non-governmental organizations (NGOs).

⁷*Environmental refugees* is arguably a better term despite the political pressures for the United Nations not to officially use this term which calls for ecolinguistic analysis in ecopedagogical spaces.

[1] ... our planet is our only home, and we have to protect it through a global citizenship sustainable development education, moving from diagnosis and denunciation into action and policy implementation ... [2] global peace is an intangible cultural good of humanity with immaterial value ... [3] seeking to fulfill their individual and cultural interests and achieving their inalienable rights to life, liberty and the pursuit of happiness ... The great question about peace is how to cultivate the spirit of solidarity across the lines of difference.

The first global common is directly linked to ecopedagogy, but all three are inseparable from each other and have endless socio-environmental aspects throughout them. [Torres \(2017b\)](#) argues that GC/E will ultimately fail if we do not have globally common goals rooted in justice, peace, and sustainability for our world. In addition, [Torres \(2017b\)](#) views that GCE cannot be successful if it entirely counters traditional citizenships, but, rather it must add value to State citizenships (e.g., national citizenship).

Teaching critical GCE *with* ecopedagogy helps students to unlearn othering that is inherent in most sub-global citizenships' ideologies and connects the rest of Earth. Ecopedagogy is essential within GCE, in part, because it provides pedagogical tools for deepening and widening students' reflexivity for better environmental praxis to become active critical global citizens ([Misiaszek, 2018, 2015](#)). The reasons for needing critical GCE within ecopedagogy stems from the necessity of reading conflictive citizenships' framings that influence (anti-)environmental acts as discussed throughout this chapter ([Misiaszek, 2018, 2015](#)).

Ecopedagogical teaching for world-Earth de-distancing is inseparable with critically teaching citizenships. Teaching to become a planetary citizen is most directly connected to such de-distancing, as can be read by [Gadotti's \(2011\)](#) quote below.

[Planetary citizenship is] based on a unifying vision of a planet and a worldwide society. It manifests itself in different expressions: 'our common humanity', 'unity in diversity', 'our common future' and 'our common homeland'. Planetary citizenship expresses a set of principles, values, attitudes and behaviours that brings a new perception of the Earth as a single community. Frequently associated with the concept of 'sustainable development', it is much broader than simply a relation with the economy.

([Gadotti, 2011](#)).

Gadotti has stated that if Earth holistically is a citizen as most planetary models indicate, then Earth is the most oppressed and voiceless citizen. Considering Earth is a citizen, as well as other conceptual aspects of planetary citizenship, is at odds with dominant Northern epistemologies that objectifies and commodifies the rest of Nature, a topic expanded upon in the next section.

[Gutiérrez and Prado \(1989\)](#), who wrote the first book on ecopedagogy, argue below that forming new ways of building widened collectivist relationships in planetary citizenship is both needed for actively striving for utopian goals and also elicits utopianism.

The meaning itself, the sense that personal or collectively we can guide our path, it is the cornerstone on which significance must rest of planetary citizenship. This rationale it opens wide the door of possible and hopeful relationships. Our existence should no longer be an a priori imperative of obligations to fulfill, but a being open to the new reality, full of suggestions and possibilities. This dimension of openness and flexibility puts us in front of an education conceived as the creation of new and possible relationships.

Ecopedagogues facilitating imaginaries of, and emergent from, planetary citizenship constructs help to reinvent relationships with one another and the rest of Earth that is essential for praxis toward a just, peaceful, and sustainable planet ([Misiaszek and González, forthcoming, 2022](#)).

Ecopedagogically reading epistemologies and ecolinguistics

Ecopedagogical reading includes determining what epistemologies are used for teaching the environmental aspect at hand, what epistemologies are absent, what are the resulting (anti-)environmental understandings and actions from these epistemologies, and the local-to-global politics surrounding all this ([Misiaszek, 2020b](#))? Problematizing dominant epistemologies of the North is particularly ecopedagogically important because, as briefly stated previously, their grounding of socio-historical oppressions are obstacles to liberation and planetary sustainability ([Misiaszek, 2019](#) citing [Santos, 2018](#)).

Epistemologies of the South exist to counter Northern (epistemological) dominance, rather than becoming globally dominant themselves, for epistemological contextual relevance and diversity, as [Santos \(2018\)](#) terms as *ecologies of knowledges* which include "knowledges born in struggles." This coincides with [Friere's \(2000\)](#) arguments discussed previously that (pedagogical) liberation must be taught through the perspectives of those who struggle ([Gadotti, 1996](#)). [Tierney \(2018\)](#) argues that teaching and research should have goals of "global meaning making" which, in part, is "not displacing one epistemology with the other but a matter of positioning Indigenous knowledge as significant or primary, with the possibility of it being separate, fused, or integrated with Western ways of knowing." Teaching through epistemological modeling as expressed here is crucial for all environmental pedagogies. For example, ecopedagogically reading Northern groundings of disciplines is accomplished by unpacking their epistemological absences ([Santos](#) named these as *sociology of absences* in his own field) for determining necessary epistemological *emergences* of disciplines (*sociology of emergence*) necessary for liberatory praxis ([Santos, 2018](#)).

Epistemologies that fertilize "commonsense" which justify unsustainable environmental violence and instills hubris that our wants and desires can be fulfilled without considering the rest of Earth, is particularly important to be ecopedagogically read.

This latter aspect of hubris was discussed in Illich's last chapter of his book *Deschooling Society* (1983), which helps ground much of ecopedagogical scholarship. He wrote that *classic* [wo]man[/they] would challenge Nature but would know there would be consequences compared to *contemporary* [wo]man[/they] who falsely believe their actions will be absent of consequences. Ecopedagogical work aligns with Santos (2018) problematizing the processes of (critical) reflexivity when grounded by epistemologies that justify (environmental) struggles of the Other (Misiaszek, 2020b). Juxtaposing Said's (1979) work on Orientalism as justifying coloniality, Northern epistemologies in our globalizing world are, in part, justifying globalizations akin to neocoloniality through environmental violence. This includes languages and linguistics.

Use of language and forms of linguistics have direct and indirect political ideological underpinnings, including how environmentalism, Nature beyond humans, and world-Earth (de)distancing is (mis)communicated in education. Ecolinguistics has many framings; the following two quotes describe some of its overlapping tenets.

[a normative ecolinguistics framework] considers relationships of humans not just with other humans but also with the larger ecological systems that all life depends on.

(Stibbe, 2014).

[ecolinguists view] language is not an autonomous faculty, but it is rather related to other human functions that interconnect with many domains of human life like economy and politics and with other disciplines such as psychology and biology. ... relatively captivated by the idea of inter-relating a linguistic structure that is linguistic units and inter-relations of these units with environment, and here the extra-relations of language are conceived in a dynamic and a global way.

(Demi, 2008).

Ecolinguistics is a crucial tool in ecopedagogical reading to deconstruct communication used to further anthropocentrism, condone unsustainable environmental violence upon others, and objectify and commodify Nature, and vice-versa with ecopedagogical work aiding in deconstructing the politics of language that condones environmental ills (Misiaszek, 2021). For example, reading the language and use of language surrounding differing GCE models leading to (anti-)environmentalism. Dewi (2018) exemplifies this below by calling for GCE intersectionalities within EFL teaching.

EFL [(English as a Foreign Language)] learners by way of metacognitive classroom implementation should aim at respecting the Earth (thus Life itself) in all its diversity as our common home in this interdependent world. It is important to first discuss some theoretical concepts of global citizenship in view of Ecohumanism.

This is one of numerous examples of teaching language and linguistics through GCE models that align with the essence of criticality, ecopedagogy, and planetary citizenship.

Ecopedagogically reading postdigitalism and post-truthism

We now live in a postdigital age, where human destinies cannot be thought of without technologies. In a way, this was always the case – homo sapiens has survived because of technologies such as fire, clothing, and agriculture. But in our world, technologies permeate all aspects of our being. "The postdigital is hard to define; messy; unpredictable; digital and analog; technological and non-technological; biological and informational. The post-digital is both a rupture in our existing theories and their continuation" (Jandrić et al., 2018).

(Jandrić and Hayes, 2022).

The quote above focuses on postdigitalism, but it is also crucial that the larger aspects of media including and beyond new technologies are ecopedagogically read as aligning with and/or countering environmentalism. Teaching through lenses of media culture theories "deepens the potential of literacy education in critically analyzing relationships between media and audiences, information, and power," as well as constructing "alternative media production" to counter oppressions and domination (Kellner and Share, 2007). This unpacking and reinventing of media is essential in ecopedagogy (Kahn, 2010; Misiaszek, 2012). Due to this chapter's limited space, I will narrow discussions of this topic to the Internet and specifically how it has been the pedagogical tool of intensifying post-truthism today, as well as essential in disrupting false truths.

Utilizing false truths to publicly justify socio-environmental oppressions is not new; however, how the Internet to spreads false/truthful information is unique. For example, the single goal of social media to get users to keep scrolling down, no matter posts' truthfulness or falseness, encourages systematic targeting the specific audience to promote false "alternative facts" to for maximum political gains. Targeting the audience is made exponentially more precise with social media's greatest commodified asset – user data. This can be witnessed by Trump's successful US election in 2016 and the spread of false election fraud against his 2020 campaign loss, global warming denial, and Covid-19 misinformation that has led to deadly consequences. Ecopedagogical reading to identify and disrupt post-truthism that justifies Development, if successful, results in *conscientization* that de-development actually

results for the targeted populations which is often opposite to what is being presented to them online (Misiaszek, 2020a). Trump's (environmental) policies will most hurt his hardline supporters who do not have six more zeros on their personal bank statements. This aligns with Marx's argument on how best to dominate the subaltern is to systematically have them believe they are benefiting from happenings in which they really suffer from most.

Two important aspects should be emphasized here. First is that "disrupting" includes radically reinventing technologies to actively counter anti-environmentalism, including technologies initially created for doing the opposite (Kahn, 2010; Jandrić and Ford, 2022). Second is recognizing the differences between ecopedagogical work that problematizes and deconstructs dominant Truths, as opposed to post-truthism as not "being truth-seeking but rather to fit ideological ends, as well as influencing socio-environmental governance" (Misiaszek, 2020a). Ecopedagogues teach through the aspect that much of our human world as subjectively constructed, especially oppressions which are not fatalistically fixed but transformable. Ecopedagogical overall goals is for worldwide justice and planetary sustainability; however, false truths cannot be utilized to achieve this goal. For example, knowingly exaggerating the effects from an anti-environmental act is not ecopedagogical work even if it might result in students' increased sense of environmentalism's importance. The next and last aspect on this topic will be how post-truthism disrupts critical dialogue and, in turn, critical/ecopedagogical teaching.

Constructive dialectical conflict, which is essential to understand (in)justices through diverse perspectives, epistemologies, experiences, etc. (Gadotti, 1996; Apple, 2004), is impossible without agreed upon truth-seeking foundations. Post-truthism diminishes agreed-upon foundations necessary for critical dialogue to occur (à la Habermas, Foucault, etc.). Dialogue with foundations of false truths is not critical and will only lead false understandings and misguided actions. An environmental example would be trying to discuss global warming without agreeing that rising temperatures are occurring or even that Earth is spherical in shape (i.e., "flat-Earth" conspiracies), making the more "common" argument of being caused by human actions or not impossible to critically discuss.

Conclusion: critical comparing environmental pedagogies

There are two crucial aspects to reemphasize before concluding this chapter on the importance of comparative and international education (CIE) in using and implementing ecopedagogical teaching and research globally. First is that ecopedagogy is both a separate pedagogy and can be groundings/currents of other environmental pedagogies. Second, ecopedagogies are complex and require contextual reinventing, as required for all (environmental) pedagogies if explicitly stated or not (Gadotti, 1996, 2008a). This chapter provides the key groundings in which reinventing the plurality of "being ecopedagogical" emerges from; however, this chapter is far from being all-inclusive. Critical comparative methodologies are necessary to construct:

...effective, action-based environmental pedagogies by deconstructing anti-environmental ones, including those falsely cloaked as being "environmental", and the politics of their existence ... this deconstructing is CIE's most essential roles for identifying and disrupting taught fatalistic ideologies that normalize socio-environmental injustices, planetary unsustainability, and the devaluing of the rest of Earth.

(Misiaszek, 2022).

The adjective *critical* to describe the comparisons being carried out is essential because CIE has a contested terrain in which most comparisons were, and unfortunately too-frequently continues to be, benchmarked by Northern pedagogical practices grounded upon Northern epistemologies (Assié-Lumumba, 2017; Takayama et al., 2017).

Global peace, socio-environmental justice, and planetary sustainability are inherent critical CIE goals (Torres, 2017a). In the larger CIE field, what approaches, models, and epistemological groundings are being utilized in comparing methodologies is essential to deconstruct, as well as the envisioned endpoints (Assié-Lumumba, 2017; Takayama et al., 2017; Shultz, 2019). Critical CIE and ecopedagogies are interdependent as argued below.

(CIE) is essential for constructing effective, action-based environmental pedagogies by deconstructing anti-environmental ones, including those falsely cloaked as being "environmental", and the politics of their existence.

(Misiaszek, 2022).

...[ecopedagogy and] ecopedagogical reading within CIE is essential both within more methodological work by CIE scholars and all learning spaces, coinciding with socio-environmental inseparability ... [r]eading to critically compare what is taught on environmentalism, how it is taught, who decides what is being taught, and the politics ...

(Misiaszek, 2022).

All the critical aspects unpacked and problematized throughout this chapter on ecopedagogy and ecopedagogical literacy are parallel concerns for critical CIE (and *should be* for all of CIE models throughout).

References

- Abdi, A.A., Shultz, L., Pillay, T. (Eds.), 2015. *Decolonizing Global Citizenship Education*. Sense Publishers, Rotterdam.
- Apple, M.W., 2004. *Ideology and Curriculum*. Routledge, New York, NY.
- Assié-Lumumba, N.D.T., 2017. The Ubuntu paradigm and comparative and international education: epistemological challenges and opportunities in our field. *Comp. Educ. Rev.* 61, 1–21.
- Beck, U., 1992. *Risk Society: Towards a New Modernity*. Sage Publications, London.
- Carson, R., 1962. *Silent Spring*. Houghton Mifflin, Boston, MA.
- Chomsky, N., 1999. *Profit Over People: Neoliberalism and Global Order*. Seven Stories Press, New York, NY.
- Dale, R., 2007. Globalization and rescaling of educational governance: a case of sociological ectopia. In: Torres, C.A., Teodoro, A. (Eds.), *Critique and Utopia: New Developments in the Sociology of Education in the Twenty-First Century*. Rowman & Littlefield Publications, Lanham, MD.
- Demi, A., 2008. The ecolinguistic paradigm: an integrationist trend in language study. *Int. J. Soc. Cult. Lang.* 21–30.
- Dewi, N., 2018. Ecohumanism in teaching poetry for EFL students in Indonesia. *GEMA Online J. Lang. Stud.* 18.
- Epstein-Halevi, D.Y., Misiaszek, G.W., Kelly, H., Shah, S., Mugarura, C., Walsh, L., 2018. Building eco-social resilience in the face of climate change and drought: how permaculture pedagogy and praxis can benefit rural communities and their environment. In: Leal Filho, W. (Ed.), *Climate Change Management*. Springer, London.
- Freire, P., 1985a. *The Politics of Education: Culture, Power, and Liberation*. Bergin & Garvey, South Hadley, MA.
- Freire, P., 1985b. Reading the world and reading the word: an interview with Paulo Freire. *Lang. Arts* 62, 15–21.
- Freire, P., 1992. *Pedagogy of Hope*. Continuum, New York, NY.
- Freire, P., 1997. *Pedagogy of the Heart*. Continuum, New York, NY.
- Freire, P., 1998. *Pedagogy of Freedom: Ethics, Democracy, and Civic Courage*. Rowman & Littlefield, Lanham, MD.
- Freire, P., 2000. *Pedagogy of the Oppressed*. Continuum, New York, NY.
- Freire, P., 2004. *Pedagogy of Indignation*. Paradigm Publishers, Boulder, CO.
- Gadotti, M., 1996. *Pedagogy of Praxis: A Dialectical Philosophy of Education*. SUNY Press, Albany.
- Gadotti, M., 2008a. *Education for Sustainability: A Critical Contribution to the Decade of Education for Sustainable Development*. Paulo Freire Institute, São Paulo.
- Gadotti, M., 2008b. *Education for Sustainable Development: What We Need to Learn to Save the Planet*. Instituto Paulo Freire, São Paulo.
- Gadotti, M., 2008c. What we need to learn to save the planet. *J. Educ. Res. Sustain. Dev.* 2, 21–30.
- Gadotti, M., 2011. Adult education as a human right: the Latin American context and the ecopedagogic perspective. *Int. Rev. Educ.* 57, 9–25.
- Giddens, A., 1990. *The Consequences of Modernity*. Stanford University Press, Palo Alto, CA.
- Giroux, H.A., 2001. *Theory and Resistance in Education: Towards a Pedagogy for the Opposition*. Bergin & Garvey, Westport, CT.
- Grande, S., 2009. Red pedagogy: Indigenous theories of redistribution (a.k.a sovereignty). In: Apple, M.W., AU, W., Gandin, L.S.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, NY.
- Gutiérrez, F., Prado, C., 1989. *Ecopedagogia e cidadania planetária*. (Ecopedagogy and planetarian citizenship). Cortez, São Paulo.
- Harding, S.G., 2006. *Science and Social Inequality: Feminist and Postcolonial Issues*. University of Illinois Press, Urbana, IL.
- Harvey, D., 2005. *A Brief History of Neoliberalism*. Oxford University Press, Oxford.
- Illich, I., 1983. *Deschooling Society*. Harper Colophon, New York, NY.
- International Organization for Migration (IOM), 2014. *IOM Outlook on Migration, Environment and Climate Change*. International Organization for Migration, Geneva.
- Jandrić, P., Ford, D.R., 2022. *Postdigital Ecopedagogies: Genealogies, Contradictions, and Possible Futures*. Springer, London.
- Jandrić, P., Hayes, S., 2022. Postdigital critical pedagogy. In: Abdi, A., Misiaszek, G.W. (Eds.), *Palgrave Handbook on Critical Theories of Education*. Springer, Cham.
- Jandrić, P., Knox, J., Besley, T., Ryberg, T., Suoranta, J., Hayes, S., 2018. Postdigital science and education. *Educ. Philos. Theor.* 50, 893–899.
- Jickling, B., Wals, A.E.J., 2008. Globalization and environmental education: looking beyond sustainable development. *J. Curric. Stud.* 40, 1–21.
- Kahn, R., 2010. *Critical Pedagogy, Ecoliteracy, and Planetary Crisis: The Ecopedagogy Movement*. Peter Lang, New York, NY.
- Kellner, D., 2002. Theorizing globalization. *Socio. Theor.* 20, 285–305.
- Kellner, D., Share, J., 2007. Critical media literacy is not an option. *Learning Inquiry* 1, 59–69.
- Klees, S.J., 2008. A quarter century of neoliberal thinking in education: misleading analyses and failed policies. *Glob. Soc. Educ.* 6, 311–348.
- Klees, S.J., 2017. Will we achieve education for all and the education sustainable development goal? *Comp. Educ. Rev.* 61, 425–440.
- Ladson-Billings, G., 2009. Race still matters: critical race theory in education. In: Apple, M.W., AU, W., Gandin, L.S.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, NY.
- Lee, W.O., Grossman, D.L., Kennedy, K.J., Fairbrother, G.P. (Eds.), 2004. *Citizenship Education in Asia and the Pacific: Concepts and Issues*. Kluwer Academic, Hong Kong.
- Mcnamara, K.E., 2007. Conceptualizing discourses on environmental refugees at the United Nations. *Popul. Environ.* 29, 12–24.
- Memmi, A., 1967. *The Colonizer and the Colonized*. Beacon Press, Boston, MA.
- Misiaszek, G.W., 2012. Transformative environmental education within social justice models: lessons from comparing adult ecopedagogy within North and South America. In: Aspin, D.N., Chapman, J., Evans, K., Bagnall, R. (Eds.), *Second International Handbook of Lifelong Learning*. Springer, London.
- Misiaszek, G.W., 2015. Ecopedagogy and citizenship in the age of globalisation: connections between environmental and global citizenship education to save the planet. *Eur. J. Educ.* 50, 280–292.
- Misiaszek, G.W., 2018. *Educating the Global Environmental Citizen: Understanding Ecopedagogy in Local and Global Contexts*. Routledge, New York, NY.
- Misiaszek, G.W., 2019. The end of the cognitive empire: the coming of age of epistemologies of the South by Boaventura de Sousa Santos. *Comp. Educ. Rev.* 63, 452–454.
- Misiaszek, G.W., 2020a. Countering post-truths through ecopedagogical literacies: teaching to critically read “development” and “sustainable development”. *Educ. Philos. Theor.* 52, 747–758.
- Misiaszek, G.W., 2020b. *Ecopedagogy: Critical Environmental Teaching for Planetary Justice and Global Sustainable Development*. Bloomsbury, London.
- Misiaszek, G.W., 2020c. Ecopedagogy: teaching critical literacies of ‘development’, ‘sustainability’, and ‘sustainable development’. *Teach. High. Educ.* 25, 615–632.
- Misiaszek, G.W., 2020d. Reconstructing shallow environmental pedagogies into critical, transformative environmental learning spaces for praxis. In: Abdi, A. (Ed.), *Critical Theorizations of Education*. Brill, The Netherlands.
- Misiaszek, G.W., 2021. An ecopedagogical, ecolinguistic reading of the Sustainable Development Goals (SDGs): what we have learned from Paulo Freire. *Educ. Philos. Theor.* 1–15.
- Misiaszek, G.W., 2022a. Contested terrains of environmental pedagogies: comparing ecopedagogy, education for sustainable development (ESD) and environmental education. In: Comparative and International Education Society (CIES) Annual Meeting (Minneapolis, Minnesota, USA).
- Misiaszek, G. W. in press. *Freire and Environmentalism: Ecopedagogy*. London, Bloomsbury.
- in press Misiaszek, G.W., 2022. Contested terrains of environmental pedagogies: comparing ecopedagogy, education for sustainable development (ESD) and environmental education. In: Misiaszek, L., Arnone, R.F., Torres, C.A. (Eds.), *Comparative Education: The Dialectic of the Global and the Local (Volume 2)*, fifth ed. Rowman & Littlefield Publications, Lanham, MD.
- forthcoming Misiaszek, G.W., González, C.O., 2022. Utopia, ecopedagogy, and citizenships: teaching for socio-environmental justice, development, and planetary sustainability. In: Tarozzi, M., Bourn, D. (Eds.), *Pedagogy of Hope for Global Social Justice*. Bloomsbury, London.

- Misiaszek, G.W., Torres, C.A., 2019. Ecopedagogy: the missing chapter of pedagogy of the oppressed. In: Torres, C.A. (Ed.), *Wiley Handbook of Paulo Freire*. Wiley-Blackwell, New Jersey.
- Morrow, R.A., 2019. Paulo Freire and the "logic of reinvention": power, the state, and education in the global age. In: Torres, C.A. (Ed.), *Wiley Handbook of Paulo Freire*. Wiley-Blackwell, New Jersey.
- Morrow, R.A., Torres, C.A., 2019. Rereading Freire and Habermas: philosophical anthropology and reframing critical pedagogy and educational research in the neoliberal anthropocene. In: Torres, C.A. (Ed.), *Wiley Handbook of Paulo Freire*. Wiley-Blackwell, New Jersey.
- Postma, D.W., 2006. Why care for nature?. In: *Search of an Ethical Framework for Environmental Responsibility and Education*. Springer, New York, NY.
- Romão, J.E., 2007. Sociology of education or the education of sociology? Paulo Freire and the sociology of education. In: Torres, C.A., Teodoro, A. (Eds.), *Critique and Utopia: New Developments in the Sociology of Education in the Twenty-First Century*. Rowman & Littlefield, Lanham, MD.
- Said, E.W., 1979. *Orientalism*. Vintage Books, New York, NY.
- Santos, B.D.S., 2018. The End of the Cognitive Empire: The Coming of Age of Epistemologies of the South. Duke University Press, Durham, NC.
- Sauvé, L., 2005. Currents in environmental education: mapping a complex and evolving. *Can. J. Environ. Educ.* 10, 11–37.
- Schugurensky, D., 2011. *Paulo Freire*. Bloomsbury, London.
- Shultz, L., 2007. Educating for global citizenship: conflicting agendas and understandings. *Alberta J. Educ. Res.* 53, 248–258.
- Shultz, L., 2019. Addressing rising nationalism and populism through global citizenship education. In: Peters, M.A. (Ed.), *Encyclopedia of Teacher Education*. Springer Singapore, Singapore.
- Stibbe, A., 2014. An ecolinguistic approach to critical discourse studies. *Crit. Discourse Stud.* 11, 117–128.
- Stromquist, N.P., 2002. *Education in a Globalized World: The Connectivity of Economic Power, Technology, and Knowledge*. Rowman & Littlefield Publishers, Lanham, MD.
- Takayama, K., Sriprakash, A., Connell, R., 2017. Toward a postcolonial comparative and international education. *Comp. Educ. Rev.* 61, S1–S24.
- Teodoro, A., Torres, C.A., 2007. Introduction: critique and utopia in the sociology of education. In: Torres, C.A., Teodoro, A. (Eds.), *Critique and Utopia: New Developments in the Sociology of Education in the Twenty-First Century*. Rowman & Littlefield Publications, Lanham, MD.
- Tierney, R.J., 2018. Toward a model of global meaning making. *J. Literacy Res.* 50, 397–422.
- Torres, C.A., 2009. *Globalizations and Education: Collected Essays on Class, Race, Gender, and the State*. Teachers College Press, New York, NY.
- Torres, C.A., 2011. Public universities and the neoliberal common sense: seven iconoclastic theses. *Int. Stud. Sociol. Educ.* 21, 177–197.
- Torres, C.A., 2015. Oppressor and oppressed: logical dialectical categories? Tribute to Paulo Freire. *Sinética* 45. Available: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2015000200008&nrm=iso. (Accessed 31 January 2018). Accessed.
- Torres, C.A., 2017a. The state of the art in comparative education and WCCES at a crossroads in the 21st Century. *Global Comparative Education* 1, 20–32.
- Torres, C.A., 2017b. *Theoretical and Empirical Foundations of Critical Global Citizenship Education*. Routledge, New York, NY.
- Warren, K.J., 2000. *Ecofeminist Philosophy: A Western Perspective on What It Is and Why It Matters*. Rowman & Littlefield Publications, Lanham, MD.

What works or what matters? Contested ideas of girls' education, gender equity and global connection

Elaine Unterhalter, Centre for Education and International Development (CEID), Institute of Education, University College London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Conclusion
References

619
619

In 2021, as leaders of the G7 club of the richest and most powerful countries in the world traveled to the UK for a summit meeting, the UK Prime Minister made a speech about girls' education in which he termed interventions to promote girls' schooling:

the silver bullet, this is the magic potion, this is the panacea. This is the universal cure, this is the Swiss Army knife, complete with allen key and screwdriver and everything else that can solve virtually every problem that afflicts humanity.

Prime Minister's Office (2021).

While Boris Johnson's rhetoric is noted for bombastic flourishes, this perception that educating girls will solve all global problems of the world, and, by implication, that becoming educated is girls' special responsibility, has been a prominent theme in pronouncements on girls' education since the 1960s. What is notable about this approach to girls' education is the ways in which this aspiration has been instrumentalized to drive forward particular conceptions of development, and associated ideas about a global order.

This notion that girls' education is "what works" is articulated in many scholarly works and programmes of development agencies, as I outline below. In this article I counterpose this approach to girls' education with a different formulation, linked with feminism, rights, gender equality, intersectionality, critical decolonial perspectives and a view of the global as a terrain formed by historical injustices that must be addressed. I term this contrasting perspective "what matters." I first present the two very different formulations of how girls' education has been understood highlighting some of the different conceptions of gender, education and the global in play. I then go on to develop a third position, which I term, "what connects", and attempt to theorize this looking at associated ideas about gender, institutions and the form of global relationships. The argument I have made (Unterhalter, 2009, 2021) is that the sharp divergence between "what works" and "what matters" is itself problematic. If something matters, like feminism or intersectionality, we need to make it work to transform injustices. If something works, we need to be clear about why, normatively, this approach to working matters. How something works, and whether this is good or just is as important as the efficiency or effectiveness of the ways of working. In developing the position associated with "what connects" I suggest a normatively inflected view of girls' education, intersectionality, and the global as a sphere of shared obligation. In so doing, I highlight some of the analytical, policy and practice issues, that come to the fore.

The mainstream approach to girls' education, articulated in the G7 Declaration on Girls Education of 2021, exemplifies "what works." The Declaration notes:

12 years of safe and quality education for all children, and specifically girls, is one of the most cost-effective and impactful social and economic investments governments and donors can make.

G7 (2021).

The Declaration see girls' education as a particular kind of development intervention and investment, that is "cost effective and impactful" linking the process of education and differently situated people who learn, play, teach, administer, plan, research, or dream with narrowly bounded, precisely formulated inputs. The "what works" approach is associated with particular kinds of programmes which focus on a single intervention, and a narrow menu of research methods, most notably randomized control trials, to identify and evaluate interventions (Duflo et al., 2007; Bannerjee and Duflo, 2019). A particular feature of development assistance associated with "what works" links it to advancing global agendas that do not undermine existing geo political and economic arrangements (Fejorskov, 2022). The methods and form of analysis have been criticized for "piecemeal analysis, ad hoc resort to theory, indifference to history and context, and methodological fundamentalism" in relation to the ways they have tried to understand women's empowerment (Kabeer, 2020). Further critiques by the World Bank economist, Martin Ravillon (2020), note the limitations of the approach to inform policy, given that the preference for RCTs as the single effective method is unclear on a priori grounds, raises ethical objections with regard to who does and does not receive the intervention, and risks distorting the evidence-base informing policymaking, given that an insistence on RCTs generates selection bias in what gets evaluated. With regard to girls'

education, many gaps in evidence to inform policy are noted (Psaki et al., 2021). Many of these do not easily lend themselves to investigation through RCTs.

Despite these kinds of critiques from many directions, the stress on girls' education as a development intervention that "works" abound (Sperling and Winthrop, 2015; Agapitova and Moreno, 2017). In statements, such as the G7 Declaration, what is not noted is that girls are human beings, not development interventions. They may or may not engage in education for a wide range of reasons associated with individual, social, institutional, cultural, health and geo political reasons, and access to school education and enhanced learning outcomes are only one stand in the lives they lead. The dynamics shaping these individual and social processes may align with the social and economic investments of governments and donors, but they may not. For the "what works" perspective the rights, needs, capabilities and demands of girls are less significant than how their education works to serve other goals. Although voices of particular girls who chime well with this agenda, may well be amplified on certain occasions, processes of participation and critical reflection on development initiatives are generally not part of the "what works" menu (Jones et al., 2020; De Haan et al., 2020).

Although the prominence given to "what works" agendas, started to emerge from around 2000, the singling out of girls' education, as a strategy that "works" goes back to the 1960s. Initially claims about the significance of girls' education were linked to population reduction (Hartmann, 1987; Sheikh and Loney, 2018). Considerable research from the 1960s and 1970s, drawing on demographic studies showed that in societies where girls remained in school until 18, delaying their first pregnancy, there was a demographic transition to smaller families, associated with enhanced participation by women in the work force (Caldwell, 1980; Schulz, 2002), although a number of commentators raised questions regarding which way the lines of causality ran, and whether, what was significant about the relationship with education was what was learned, or merely entailed attending a school, or living in the vicinity of a school (Axinn and Barber, 2001; Cruz and Ahmed, 2018). This theme that the education of girls is important, as a form of contraception, has persisted and emerged as fears mounted about the dire outcomes of the climate crisis. In a number of accounts of how to respond to the climate girls' education is portrayed as one of the most effective responses governments can take, because educated women and girls are portrayed as having reduced fertility rates, slowing population growth, limiting green house gas emissions and competition for increasingly scarce resources (Drawdown, 2020; Volsett et al., 2020). Estimates have been made that providing for universal education, and bringing all girls into school, will have the result that the population in 2050 could be smaller by 1.5 billion people than if girls continued to have limited access to education. One estimate is that this "fix" would mean that by 2100 the global population could comprise 5.7 billion fewer people (Lutz et al., 2017). A further study identified a strong positive association between the average number of years' schooling a girl receives in her country and the country's ND-GAIN index, a measure of a country's resilience to climate disasters (Kwaak and Braga, 2017). This study found that for every additional year of schooling for a girl, her country's resilience to climate disasters could be expected to improve by 3.2 points (p. 19).

The focus on girls' early marriage and fertility is one stream of the "what works" framing of the significance of girls' education, but another is to see educated girls as an important source of labor in particular forms of manufacturing and service. In the 1990s, girls' education was identified as the most cost effective and efficient intervention that could be taken to reduce poverty. A classic World Bank publication (King and Hill, 1993), with a Foreword by Lawrence Summers, then Chief Economist at the World Bank, promoted girls' education as *the* best investment to address poverty. Summers had developed his analysis in a World Bank paper, where he noted

underinvestment in girls is not an ineluctable consequence of poverty, nor is it made necessary by any religious or cultural tradition. It is an economic problem that results from a vicious cycle caused by distorted incentives. The expectation that girls will grow to do little other than serve their husbands reduces parent's incentive to invest in their daughter's human capital. Uneducated women then have few alternatives and so the expectation becomes self fulfilling, trapping women in a continuous circle of neglect increasing educational opportunities for girls offers the best prospect for cutting into this vicious cycle. As an economic investment, increased outlays directed at educating girls may well yield the the highest return of all investments available in developing countries considering both private benefits and returns to other family members....Experience suggests that female education programs are relatively inexpensive compared to other development investments and could quickly increase female enrollment ratesMajor initiatives to increase female education have the potential to transform society over time. If a larger fraction of girls had gone to school a generation ago millions of infant deaths each year could have been averted and tens of millions of families could have been healthier and happier.

Summers (1992).

This early formulation of "what works" links girls' education with enhancing human capital, yielding high returns for inexpensive outlays, preventing infant mortality, and enhancing the health of populations. Nowhere does his summary mention how educating girls responds to rights, needs or aspirations of the girls' themselves. They only appear in shadowy form with regard to some generalized form of private benefits.

The focus in the seminal World Bank study (King and Hill, 1993) was on addressing barriers to girls' access to schooling, not on looking at experiences, content, pedagogies, hierarchies, exclusions, or school related gender based violence. In later versions of this concern, the focus would come to be on learning outcomes for girls, and failures with ensuring this some of the reasons for education not delivering the predicted outcomes with regard to ending poverty and promoting economic growth. One of the clearest expressions of this view are in the working papers associated with the RISE (Research on Improving Systems of Education) program. A key distillation of this view in a paper on women's literacy (Kaffenberger and Pritchett, 2021) draws on cross national comparable data sets with information on schooling, assessments of literacy, and life outcomes for 50 countries. The analysis shows women's

education, comprising 6 years of schooling and the acquisition of literacy, is positively associated with reduced fertility and child mortality, increased empowerment (assessed through responses to questions on a DHS survey about family decision making and perceptions of intimate partner assault and violence) and financial practices with regard to women's independent use of money. The analysis established that positive life outcomes, associated with literacy and schooling, was larger than the standard estimates of the gains from schooling alone. The analysis here is more nuanced than that promoted by Summers (1992), because positive outcomes are noted in relation to women themselves, but the implications of these, suggested by Kaffenberger and Pritchett (2021), are for decisions to be taken about education systems and investments in learning outcomes, rather than addressing gender inequalities or building an infrastructure around women's rights.

This idea that girls' education was a kind of injection that could be given against poverty, high infant mortality or failures in economic growth has been prevalent in policy documents of the World Bank for thirty years, as noted by a number of critical commentators (Unterhalter, 1996; Peppin Vaughan, 2010; Stromquist, 2012; Monkman, 2021). This has shaped the programmes of many large donor agencies, associated with governments and the private sector (Unterhalter, 2017; Moeller, 2018; Dejaeghere, 2017). The RISE program (eg. Kaffenberger and Pritchett, 2021) has argued to a shift in focus on increasing girls' learning attainment, and focused outcomes on features of women's lives, but has not gone further to suggest more widely framed visions of development or rights.

This focus on the way in which girls' education was a "tunkey" policy for development came to frame the ways in which development was monitored, as the number of girls' enrolled in schools was seen to signal the extent to which other development aspirations could or could not be realized. In influential frameworks for monitoring and evaluating global development, most notably the MDGs (Millennium Development Goals) adopted by nearly all the government of the world in 2000, indicators for what was termed the global goal of "gender equality and women's empowerment" centered exclusively on gender parity in education, that is girls' enrollment in school as a proportion of boys. (Black and White, 2004; Unterhalter, 2008a; Sen and Mukherjee, 2014).

As global policy attention widened from the rather narrow perspective of the MDGs to a broader framing associated with the SDGs (Sustainable Development Goals), the "what works" framing of girls' education came to be associated with particular large programmes, often involving the private sector in PPPs (public private partnerships), which made claims about the ways in which girls' education worked and supported empowerment (Fennell, 2012; Unterhalter, 2017; Evans and Yuan, 2022). These often rested on very limited evidence, or on studies that showed, that how girls education worked was often highly contingent on the relationships in particular contexts, the policy and practice approaches of governments and how these were enacted, and the complexities of power in households, communities, institutions, and political economies, stretching from the global to the local (Unterhalter et al., 2014; Parkes et al., 2016; Unterhalter and Robinson, 2020; Pankhurst, 2022). Despite this expansive critical literature, and even a clear delineation from within a form of "what works" analysis of areas where there was limited evidence (Psaki et al., 2021, pp. 65–67, 74–75), confident tropes associated with "what works" around girls' education continued to be expressed, and were notable during the COVID 19 pandemic. For example, a 2021 Report of the Global Education Evidence Advisory Panel identifying "Great buys" and "good buys", that is intervention programmes that a team of experts viewed as highly cost-effective, with a strong evidence base to guide initiatives to build back better, drew confident conclusions with regard to girls' education and enhanced learning outcomes, even though, the team acknowledged the evidence base was extremely limited, and in the case of the data on Afghanistan, was about to be overturned by political events:

While some effective access programs target only girls, most of the interventions referenced in this note have impacts for both girls and boys, with some general access interventions proving particularly beneficial for girls (for example, in Afghanistan). Most evaluations of programs that did not specifically target girls have found bigger impacts on girls' access than on boys' (where girls were the disadvantaged gender), and general interventions deliver gains for girls in access and learning that are comparable to gains from girl-targeted interventions. But given that girls' participation rates remain lower than those of boys in 53 developing countries, with particularly large disparities in West Asia and Sub-Saharan Africa, more research is needed on the effectiveness of programs that specifically target girls, especially in areas where girls are far behind boys.

GEEAP (2020).

The "what works" position in girls' education can be understood as primarily a techno-rationalist formulation of scaling up from the effectiveness of the intervention proved in one site, and therefore understood as replicable globally, despite many studies that call this into question either because of differences in context (Gibbs et al., 2021), limited research (Psaki et al., 2021) or, as learned in other sectors, the need to document and track the forms of intervention over decades before conclusions can be drawn (Barrientos, 2018).

Comparison in this formulation of "what works" is a form of governance by numbers (Grek, 2020). Numbers of girls enrolled in school or attaining specific learning outcomes are required to validate other development initiatives, as exemplified in the "Good buys" and "Great buys" briefing paper quoted above. These monitoring frameworks place technical experts, who design the RCTs, or the surveys to collect information on enrollment or attainment, at the center of evaluation. For these technical experts gender is understood unproblematically as a form of social relationship linked with biology. The many personal and political debates concerning gender and forms of social justice are not relevant.

In the literature on "what works" the learning crisis is a key organizing idea. The depiction of "a learning crisis" through which children are enrolled in school, but not attaining set levels, emerged at the same time as the "what works" framework (Pritchett, 2019; Languille, 2014; Unterhalter, 2019a; Sriprakash et al., 2020). Indeed, one framing supports the other as the Kaffenberger

and Pritchett (2021) paper illuminates. Thus, the argument runs, the learning crisis, variously associated with poor teaching, and lack of government accountability to deliver education, is preventing girls from learning at school, and limiting the effectiveness of the “magic potion” associated with girls’ education. It was partly the search for a “what works silver bullet” to make the “silver bullet” of girls’ education effective, in the face of the learning crisis, that drove the thinking behind the “what works” approach to education systems (Psaki et al., 2021).

An issue for the “what works” approach has remained how to make sense of the complexities of context, which have often made it so difficult to replicate an intervention in a different locale to that it was trialled in. While some writers associated with this position acknowledge this concern (Psaki et al., 2021), others metonymically read off particular markers associated with socio-economic status, rural or urban location, or the success or failure of education systems to expand learning outcomes to signal aspects of trying to fit in context (GEEAP, 2020; Kaffenberger and Pritchett, 2021).

This narrow conception of girls’ education has been challenged from many directions. While “what works” rests on linking girls’ education to the way it could achieve objectives beyond education, the contrasting position “what matters” rests on articulating a number of different kinds of visions, and showing how education of women and girls was a central component of the vision and how it would be achieved.

From the 1960s women’s rights activism in the global north and south raised issues of reproductive rights and the ways in which poverty and patriarchy impacted on women’s lives, despite education provision. These themes gained prominence as global women’s networks were built linked to UN sponsored conferences (Antrobus, 2008; Zwingel and Zwingel, 2016). A sharp critique developed of the kind of project Summers was promoting, which Devaki Jain, a feminist activist and analyst, called “the ‘catching-up-with-men’ approach”, noting these “pushed poor women into deeper corners” (Antrobus, 2015; Jain, 2012). In 1983 Jain and colleagues embarked on a collective writing and networking initiative, which generated the DAWN (Development Alternatives of Women for a New Era) manifesto in 1985 (Sen and Grown, 1987). This was as central to the “what matters” paradigm as Summers’ paper was to “what works” and was equally influential, although on a different constituency. For DAWN it was essential that struggles and contributions of poor women in different regions of the world were taken into account in any development policy or practice. Their experiences, insights and struggles to overturn intersecting inequalities were what mattered. The DAWN manifesto noted how what were defined as “women’s issues” by mainstream development policy and practice, linked with macro-economic structures and global crises. The analysis looked at the relationship of gender and class and critiqued existing economic and population control programmes, and approaches, such as the basic needs approach for failing to document or engage with the concerns of poor women, shaped by colonial heritages and resource inequalities. The argument in the DAWN Manifesto noted how a range of systemic crises and reproductive failures shaped the experiences of poor women who confronted continuing struggles to access food, fuel and water, the consequences of national debt and the lack of public service. Many experienced the effects of war, violence and a surge in repressive patriarchal cultures. To challenge and overturn poverty and gender subordination the DAWN manifesto mapped out alternative visions, strategies and methods. The vision they outlined was:

We want a world where the inequality based on class, race and gender is absent from every country and from the relationships between countries.
Sen and Grown (1987, p. 80).

Strategies to undertake this were to be formulated in discussion with women’s movements, and long term strategies to dismantle forms of inequality were to be undertaken together with short term actions to counter contemporary crises. Education appears as a key means to understand and document the forms and effects of inequalities and to mobilize for change. The limitations imposed by illiteracy are noted and the needs for mass education programmes are outlined linked with Freirean consciousness raising (Sen and Grown, 1987, p. 88). Thus what matters is centring the perspectives and experiences of poor women and their concern to transform intersecting inequalities. Education projects are part of this process, and need to be closely connected with this vision.

Other critical positions with regard to “what works” were being formulated by scholars writing about feminisms in different regions, and taking context as key shaper of perspective and analysis, not a background idea. For example, in Africa, the work of a group of writers have come to be grouped together linked by concerns with African feminisms which express ideas about what matters centered on a range of different social and political positions linked with power, powerlessness, and processes of change (Mama, 2011). Nkealah (2016, p. 62) distils African feminist concerns with what matters linked with “indigenous feminist models’ considering an African cultural perspective, an African geo-political location, and an African ideological viewpoint.” Kwachou (2023, forthcoming) notes many different strands in African feminist work on education highlighting language, values, negotiations with men, and a different vision of equity. Here education is a site to realize African feminist visions and practices.

A third perspective on “what matters” starts off not from a particular perspective, as affirmed by DAWN and proponents of African feminisms, but from the assertion of particular normative positions associated with feminism, rights or capabilities. While many of these did not have a specific focus on girls’ education, implications could be distilled from the analysis and linked with a critical perspective on “what works”, generating ideas about what matters. Much of my work on this theme has taken this path. Drawing on Molyneux’s (1984) distinction between practical gender needs and strategic gender interests, I wrote a number of critiques of the World Bank perspectives on girls’ education as an inoculation against poverty, highlighting how concerns with gender equality and a transformation of unjust institutions was occluded (Unterhalter, 1996, 2005a, 2008a; Unterhalter, 2000). In other analyses I noted the significance of women’s rights, including the right to education, and considered how analyses

of women's needs, rights and capabilities could be distinguished and mapped onto different approaches and policies (Unterhalter, 2003a, 2005b, 2005c, 2012). I considered normative work on advancing women's empowerment, and the implications of this for thinking about girls' education (Unterhalter, 2008b, 2013; Unterhalter and North, 2018; Unterhalter, 2019b). In this writing what matters is expressed normatively, and girls' education is one facet of the realization of a particular normative position.

A critical perspective on this version of what matters, notes that feminism is not a single narrative or shaped by a single stream of normative ideas. From this viewpoint, these normative visions are often co-opted by the leadership of powerful states, and become a rallying cry for versions of white saviourism (Zakaria, 2021). Shenila Khoja-Moolji shows how "what works" assumptions and versions of ideas about rights were formulated both by the political leadership of Muslim states and by aid organisations, drawing on a homogenized grouping of Muslim women and girls, oppressed everywhere by religion, tradition, and patriarchy, only able to fulfill their potential if "saved" by modern education. Her book challenges these positions centring women's historical narratives which show them to be "fully human and political subjects" (Khoja-Moolji, 2018, p. 22). Here what matters is not the abstractions of analysis but the lived experiences and actions against injustice, that need to be listened to by researchers or academics with humility, paying attention, not imposing frameworks. (Kadiwal, 2021).

Although there are considerable differences between the articulations of "what matters" I have outlined, there are some common features. In all there is a thinking about education systems linked with transformation of injustice, and affirmation of rights, and connection with the demands of feminist, anti-racist, and climate justice activists. In this analysis girls' education matters because of the way it is woven in *both* with the injustices of racism, colonialism, patriarchy, neoliberalism and climate catastrophe, *and* with strategies to bring about change. Girls education is not a "turnkey" intervention but part of both the problem of many faceted injustices, and a small element of many complex processes to address these. Unlike "what works" which has come to center on one crisis – the learning crisis – the "what matters" paradigm delineates multiple crises – environmental, economic, expanding inequalities, violence, colonial relationships from the past and present. There is no signature research method associated with this work, but many studies of girls' education draw on history, literature, and qualitative sociological and anthropological research. Women's studies has been a key generator of theories and epistemological frames.

In "what works" the notion of the global is remarkably frictionless. Interventions can be scaled up or exported around the world. By contrast from the "what matters" perspective, the global is marked by uneven and unjust power between and within countries and regions, between groups of people, and between historical processes. For all the differences in articulation of "what matters", there is a coalescence of views about the need to change these structures, institutions and relationships which form inequality. However, there is considerable difference with regard to how to do this. In the DAWN perspective, it is the link with feminist women's movements representing poor women that is a central to the vision, the strategy and the method. For writers engaging with African feminisms actions are guided by a historically and culturally located place. When the starting place is a normative frame, actions steer toward its guiding light, while for writers articulating a decolonial feminist position it is critical engagement with positionality, epistemic violence, exclusions and distortion that shapes action.

In what works comparison between countries in implementing girls' education is noted with regard to policies, enrollment or attainment. By contrast in "what matters" comparison appears as an exercise either in ideal theory, drawing out the normative idea – gender equality, women's rights, intersectionality – or as paying attention to the lived reality associated with poor women, place, position or representation.

Much of my writing on girls' education has been from within the "what matters" paradigm, and I remain steeped in its politics and analysis. But, taking a somewhat critical perspective on some of my own work, I have for some years been concerned to work to move the locus of engagement of analysts concerned with "what matters" from a margin, to some of the realms of policy making associated with power. I recognize this is a difficult task, and that acting as an ally requires humility, listening, critical reflection, which are not easy to practise sometimes under pressures to deliver to high standards of academic rigor or framed through a particular genre of policy discourse. There is also an ever present problem of co-option, which has dogged work on women's rights and empowerment (Unterhalter, 2019a,b). In developing the position I have termed "what connects" I have tried to develop a normatively inflected view of girls' education, intersectionality, and the global as a sphere of shared obligation identifying certain key sites of engagement. Here I set out some of the analytic spine of the position I am trying to develop around "what connects" in girls' education and provide one example of trying to put this process of acting as an ally into practice.

The idea of connecting ideas and actions through praxis is mapped explicitly in the DAWN manifesto (Sen and Grown, 1987), and appears in some versions of African feminism, particularly nego-feminism as articulated by Nnaemeka (2004) and in some of the work on feminism and decoloniality (eg. Harcourt, 2016; Okech, 2020). This form of connection is different from the mainstream version of "what works" because it is not the outcome that is important, regardless of the process, but how the outcome realizes the ideal. For example, enrolling all girls in school under a repressive, authoritarian government that maintains high levels of poverty and insecurity, suppresses women's rights to own property, move about freely, or challenge violence, might work to increase learning outcomes, reduce population, and engage women in work, giving them access to financial products, but it would not undo unjust structures or institutions of gender inequality. By contrast, leading a mass literacy campaign amongst girls with material developed by a local feminist movement, raising consciousness about why and how gender injustices were kept in place, and what examples of successful struggles for change in provision of say water, food, and fuel had been, and how these could be sustained, would be an instance of connecting ideas and actions for change.

The position I am terming "what connects" draws on a number of formulations of connection. These include ideas of praxis which I have expanded to draw in reflections on working in comparative education and some theorisations of intersectionality. In recent work I have drawn on analyses formulated by Jagdish Gundera (2000, 2014) to expand an approach I have termed

reflexive comparative education through which I attempt to connect concerns with the local, the national, the global and the trans-local. Gundara, I have written, outlines methods to combine some of the insights from comparative education and education and international development, - methods that are "heterodox and multivalent, translocational, and critical" (Unterhalter, 2020). Gundara's themes and work, I consider express some elements of connection that are highly salient to thinking about girls' education. Gundara draws on history, both his own autobiography, and the history of the layers of colonial education in Kenya where he grew up. He thus highlights a strand that is key in what connects, to understand the implications of past actions, for example how girls' education policies were implemented, and work to change them, if they have been associated with war, violence, and a maintenance of inequalities, drawing on a hope that this can happen. Gundara, I have written (Unterhalter, 2020, p. 18), delineates a form of listening to others' perspectives, asserting and examining universal values, charting a different course to the destructive, often hate-filled paths on offer, challenging the maintenance of boundaries, and drawing on some of the insights of diaspora and exile. Gundara's work expresses "concerns with lived experience, trying to think through and enact the intercultural" alert to visual arts, literature and oral narratives, reading, the global, the national and the local, their interconnections, injustices, and disjunctures (Unterhalter, 2020, p. 18). Gundara's work thus expresses many forms of connection, but in his professional engagement he worked to put these ideas into practice, joining committees, shaping policy. Thus one stream animating the position of "what connects" is connecting to history, to place, to people taking account of all the complexity and drawing on the insights this offers to guide action in research, policy and practice.

A second seam of thinking in "what connects" draws on my engagement with the ideas around intersectionality, itself an expression of connection. Intersectionality is widely used as an analytic tool for institutions and organisations to address inclusion, fairness, and think about equity, and has been used in a number of studies of gender and of education (Collins and Bilge, 2020, pp. 2–3; Anthias, 2013; Harris and Leonardo, 2018; Crimmins, 2019). The term intersectionality is generally associated with the work of Kimberlé Crenshaw (2017) who argued, in the context of the USA, that African American women needed to take account politically and analytically of the ways in which they were subordinated as women, as black women, and as black women workers. The idea of intersectionality has widespread currency in some parts of the world amongst activists, academics and policy communities, although the idea of intersectionality is somewhat fluid, with a number of different interpretations (Collins and Bilge, 2020). Drawing on McCall's (2005) discussion of intersectionality which distinguishes between intra-categorical, inter-categorical, and anti-categorical formulations, a recent paper I co-wrote, commissioned for UNESCO, identifies three forms of intersectionality present in discussions of girls' education (Unterhalter et al., 2020). These entail, firstly, a policy approach that highlights descriptions of overlapping differences within groups defined by gender, which we term descriptive intersectionality, and which is sometimes adopted in the what works perspective. We also identify a second form of intersectionality, which we term institutional and normative intersectionality. This entails undertaking an analysis of the ways that formations of power and powerlessness operate in and through education. A third form of intersectionality, discursive intersectionality we characterize as engaging with discursive critique of the ways ideas of gender and inclusion are formulated in education policy and practice. Institutional and discursive intersectionality are both features of "what matters." The argument we made in the UNESCO background paper was the need to bring together the three forms of intersectionality in analysis. This process I think could be taken forward as a form of practice, linked to ideas about understanding histories, facilitating participation, listening, detailed analysis, critique, attending to experience and voice and thinking about the capability for connection.

The AGEE (Accountability for Gender Equality in Education) project, which I lead, could be seen as an instance of "what connects." AGEE is an innovative project, funded primarily by the UK ESRC (Economic and Social Research Council) and FCDO to develop an indicator framework for gender equality in education through critical participatory discussion at local,¹ national, and international levels. AGEE began as a partnership between researchers at universities in Malawi, South Africa and the UK, working with selected UN agencies, bilateral donors and a range of government and civil society organisations, and is currently expanding the range of projects and partnerships. It is an attempt to link together groups with and without power around a project concerned with gender, rights and capabilities (Unterhalter et al., 2022). In some of the work for the project we have tried to engage with, many of the critical comments about governance by indicators (Grek, 2020). The analysis we have made within the project is that from the perspective of "what connects" and the ideas about articulation (Unterhalter, 2019b; Unterhalter et al., 2022) that it draws on, indicators are important to ensure robust and reliable ways of identifying and evaluating inequalities and their underlying causes in order to build more just and resilient societies. How we do this, however, is not straightforward and presents many conceptual and practical challenges around understanding, accessing and utilizing information. The AGEE Framework for evaluating gender equality in education was launched in 2021 and is designed for use at national and international levels to support diagnostic and evaluative work regarding local and global commitments to equality, social justice and sustainable development. It is intended for use in the following settings:

1. International cross country comparison of gender equality in education using a set number of common indicators across countries; this will underpin steps to develop a composite indicator
2. National dashboards to be used supporting work on gender equality linked to Education Sector Plans or other national accountability processes; indicators for domains in the AGEE Framework will be determined through national consultations and reflections

¹Project reference: ES/P005675/1. AGEE has also received support from UCL Covid Support Scheme, the Department of Education Practice and Society (EPS), UCL, and worked in partnership with UNESCO, UNGEI, ActionAid and GPE on specific projects.

3. Project dashboards to be used drawing on the AGEE Framework for diagnostic, monitoring or evaluation work on projects concerned with girls' education and gender equality in education
4. Dashboards for use with mobile or displaced populations, supporting work with a focus on emergencies, conflict and peacebuilding
5. Dashboards for diagnostic, monitoring or evaluation work in neighbourhoods looking at addressing gender and intersecting inequalities in and through schools, looking at schools and their communities.
6. Dashboards for diagnostic, monitoring or evaluation work on gender in higher education institutions or networks of higher education institutions

Current phases of work entail a participatory process for developing criteria for the cross national comparative indicators, the generation of a shortlist of indicators to use on the cross national dashboard, undertaking participatory processes for developing national dashboards, and work in single institutions on generating customized dashboards. It can be seen that in the practice of the project there is an attempt to move between sites that are global, national and local, connecting insights, and forms of practice.

Conclusion

This article has outlined two major positions on girls' education, development and the global order. The "what works" position which stresses a techno-rationalist view of managing education, a vision of education and development that stresses efficiency rather than any larger concern with addressing poverty, inequality and injustice. This approach does not explicitly seek to change the major injustices and disjunctures in the global order. The questions this brings to the fore are how can girls' education as a global policy idea and practice be steered more toward concern with women's rights, gender equality, and redressing global injustices.

I contrast this "what works" position on girls' education with a position I term "what matters." This position links normative visions for supporting rights, capabilities and attending to the demands of poor women, with girls' education, a participatory and transformational approach to education and development, and an intention to engage with the transformation of global injustices. The questions this raises are whether the policies and visions advocated, link concretely with actions for change.

Building from critiques of "what works" and ideas informed by "what matters" I formulate a third position on girls' education and global policy in terms of "what connects." This position drawing on theories of praxis, the overlap in interpretations of intersectionality and some theorisations of articulation, stresses the importance of normative ideas put into practice and the global as a site of shared obligation to address poverty and intersecting inequalities. This surfaces ideas about who formulates policy and guides practice, and how normative ideas are assessed in very specific local contexts. Girls' education is not a "universal cure", it is an area where different pathways are struggled over and evaluated, and where active processes of engagement are needed to listen to and work with the demands of the poorest and most discriminated against. Supporting rights needs, capabilities, equality and justice all matter, and multiple connections are required to make the politics, policy and practice associated with this work.

References

- Agapitova, N., Navarrete Moreno, C., 2017. Educate Girls. World Bank, Washington DC.
- Anthias, F., 2013. Intersectional what? Social divisions, intersectionality and levels of analysis. *Ethnicities* 13 (1), 3–19.
- Antrobus, P., 2008. *The Global Women's Movement: Origins, Issues and Strategies*. Bloomsbury Publishing.
- Antrobus, P., 2015. DAWN, the third world feminist network: upturning hierarchies. In: Bakshi, R., Harcourt, W. (Eds.), *The Oxford Handbook of Transnational Feminist Movements*. University Press, Oxford/New York, pp. 163–165.
- Axinn, W.G., Barber, J.S., 2001. Mass education and fertility transition. *Am. Socio. Rev.* 481–505.
- Banerjee, A., Duflo, E., 2019. *Poor Economics*. Penguin Books, London.
- Barrientos, A., 2018. The Informational Basis of Emerging Social Assistance in Low and Middle-Income Countries. GDI Working Paper, p. 23.
- Black, R., White, H., 2004. Targeting development. In: *Critical Perspectives on the Millennium Development Goals*. Taylor and Francis, Abingdon.
- Caldwell, J.C., 1980. Mass education as a determinant of the timing of fertility decline. *Popul. Dev. Rev.* 225–255.
- Collins, P.H., Bilge, S., 2020. *Intersectionality*. John Wiley & Sons.
- Crenshaw, K.W., 2017. *On Intersectionality: Essential Writings*. The New Press, New York.
- Crimmins, G. (Ed.), 2019. *Strategies for Resisting Sexism in the Academy: Higher Education, Gender and Intersectionality*. Springer.
- Cruz, M., Ahmed, S.A., 2018. On the impact of demographic change on economic growth and poverty. *World Dev.* 105, 95–106.
- de Haan, A., Dowie, G., Mariara, J., 2020. To RCT or not, is not the question: methods for policy-relevant research on gender equality. *World Dev.* 127, 104794.
- DeJaeghere, J., 2017. *Educating Entrepreneurial Citizens: Neoliberalism and Youth Livelihoods in Tanzania*. Taylor & Francis, London.
- Drawdown, 2020. Available at: <https://drawdown.org/solutions/family-planning-and-education>.
- Duflo, E., Glennerster, R., Kremer, M., 2007. Using randomization in development economics research: a toolkit. In: *Handbook of Development Economics*, vol. 4, pp. 3895–3962.
- Evans, D.K., Yuan, F., 2022. What we learn about girls' education from interventions that do not focus on girls. *World Bank Econ. Rev.* 36 (1), 244–267.
- Fejerskov, A., 2022. *The Global Lab: Inequality, Technology, and the Experimental Movement*. Oxford University Press.
- Fennell, S., 2012. Why girls' education rather than gender equality? The strange political economy of PPPs in Pakistan. In: Robertson, S., Mundy, K., Verger, A., Menashy, F. (Eds.), *Public Private Partnerships in Education*. Edward Elgar Publishing, Cheltenham, pp. 259–276.
- G7, 2021. declaration on girls' education in communique from G7 foreign and development ministers meeting. London. Available at: <https://www.gov.uk/government/publications/g7-foreign-and-development-ministers-meeting-may-2021-communique/declaration-on-girls-education>.
- Geep, 2020. Cost Effective Approaches to Improve Global Learning. Recommendations of the Global Evidence Advisory Panel. World Bank/DFID, Washington/London. Available at: <https://documents1.worldbank.org/curated/en/719211603835247448/pdf/Cost-Effective-Approaches-to-Improve-Global-Learning-What-Does-Recent-Evidence-Tell-Us-Are-Smart-Buys-for-Improving-Learning-in-Low-and-Middle-Income-Countries.pdf>.

- Gibbs, E., Jones, C., Atkinson, J., Attfield, I., Bronwin, R., Hinton, R., Savage, L., 2021. Scaling and 'systems thinking' in education: reflections from UK aid professionals. *Compare* 51 (1), 137–156.
- Grek, S., 2020. Prophets, saviours and saints: symbolic governance and the rise of a transnational metrological field. *Int. Rev. Educ.* 66 (2), 139–166.
- Gundara, J.S., 2000. *Interculturalism, Education and Inclusion*. Sage.
- Gundara, J.S., 2014. Global and civilisational knowledge: eurocentrism, intercultural education and civic engagements. *Intercult. Educ.* 25 (2), 114–127.
- Harcourt, W. (Ed.), 2016. *The Palgrave Handbook of Gender and Development: Critical Engagements in Feminist Theory and Practice*. Springer.
- Harris, A., Leonardo, Z., 2018. Intersectionality, race-gender subordination, and education. *Rev. Res. Educ.* 42 (1), 1–27.
- Hartmann, B., 1987. *Reproductive Rights and Wrongs: The Global Politics of Population Control*. South End Press, Harper and Row, NY.
- Jain, D., 2012. Passion: driving the feminist movement forward. *J. Wom. Hist.* 24 (4), 201–207.
- Jones, N., Presler-Marshall, E., Kassahun, G., Kebede Hateu, M., 2020. Constrained choices: exploring the complexities of adolescent girls' voice and agency in child marriage decisions in Ethiopia. *Prog. Dev. Stud.* 20 (4), 296–311.
- Kabeer, N., 2020. Women's empowerment and economic development: a feminist critique of storytelling practices in "randomista" economics. *Fem. Econ.* 26 (2), 1–26.
- Kadiwal, L., 2021. Feminists against fascism: the Indian female Muslim protest in India. *Educ. Sci.* 11 (12), 793.
- Kaffenberger, M., Pritchett, L., 2021. Effective investment in women's futures: schooling with learning. *Int. J. Educ. Dev.* 86, 102464. <https://doi.org/10.1016/j.ijedu-dev.2021.102464>. ISSN 0738-0593. Available at:
- Khoja-Moolji, S., 2018. *Forging the Ideal Educated Girl: The Production of Desirable Subjects in Muslim South Asia*. University of California Press, Oakland.
- King, E.M., Hill, M.A. (Eds.), 1993. *Women's Education in Developing Countries: Barriers, Benefits, and Policies*. World Bank Publications.
- Kwaak, C., Braga, A., 2017. *Three Platforms for Girls' Education in Climate Strategies*. Brookings, Washington, DC.
- Kwachou, M., 2023. 'In response to Acker: black and African feminist theory on gender and education'. *Comp. Educ.* 59 (1).
- Langille, S., 2014. Quality education through performativity: 'Learning crisis' and technology of quantification in Tanzania. *Int. J. Educ. Dev.* 39, 49–58.
- Lutz, W., Butz, W.P., Samir, K.C., 2017. In: *World Population & Human Capital in the Twenty-First Century: An Overview*. Oxford University Press, Oxford.
- Mama, A., 2011. What does it mean to do feminist research in African contexts? *Fem. Rev.* 98 (1_Suppl. 1), e4–e20.
- McCall, L., 2005. The complexity of intersectionality. *Signs* 30 (3), 1771–1800.
- Moeller, K., 2018. *The Gender Effect: Capitalism, Feminism, and the Corporate Politics of Development*. Univ of California Press.
- Molyneux, M., 1984. Mobilisation without emancipation? Women's interests, state and revolution in Nicaragua. *Crit. Soc. Pol.* 4 (10), 59–71.
- Monkman, K., 2021. Gender equity in global education policy. In: *Oxford Research Encyclopedia of Education*. Oxford University Press, Oxford. <https://doi.org/10.1093/acrefore/9780190264093.013.1334>. Available at:
- Nkealah, N., 2016. (West) African feminisms and their challenges. *J. Lit. Stud.* 32 (2), 61–74.
- Nnaemeka, O., 2004. Nego-feminism: theorizing, practicing, and pruning Africa's way. *Signs* 29 (2), 357–385.
- Okech, A., 2020. African feminist epistemic communities and decoloniality. *Crit. Afr. Stud.* 12 (3), 313–329.
- Pankhurst, C., 2022. *Girls' Education and Climate Change: A Critical Review of the Literature*. AGEE working paper, AGEE Project. University College London.
- Parkes, J., Heslop, J., Ross, F.J., Westerveld, R., Unterhalter, E., 2016. *A Rigorous Review of Global Research Evidence on Policy and Practice on School-Related Gender-Based Violence*. UNICEF, New York.
- Peppin Vaughan, R., 2010. Girls' and women's education within unesco and the World Bank, 1945–2000. *Compare* 40 (4), 405–423.
- Prime Minister's Office, 10 Downing Street, 2021. PM speech at the global education summit: 29 July 2021. Available at: <https://www.gov.uk/government/speeches/pm-speech-at-the-global-education-summit-29-july-2021>. (Accessed 28 June 2022). Accessed:
- Pritchett, L., 2019. Understanding the politics of the learning crisis: steps ahead on a long road. In: *The Politics of Education in Developing Countries*. Oxford University Press, pp. 197–209.
- Psaki, S., Haberland, N., Kozak, M.L., Woyczynski, L., 2021. *Girls' Education Roadmap: 2021 Report*.
- Ravallion, M., 2020. Should the Randomistas (Continue to) Rule? (No. w27554) National Bureau of Economic Research.
- Schultz, T.P., 2002. Why governments should invest more to educate girls. *World Dev.* 30 (2), 207–225.
- Sen, G., Grown, C., 1987. *Development Crises and Alternative Visions: Third World Women's Perspectives*. Monthly Review Press, New York.
- Sen, G., Mukherjee, A., 2014. No empowerment without rights, no rights without politics: gender-equality, MDGs and the post-2015 development agenda. *J. Hum. Dev. Capab.* 15 (2–3), 188–202.
- Sheikh, S.M., Loney, T., 2018. Is educating girls the best investment for South Asia? Association between female education and fertility choices in South Asia: a systematic review of the literature. *Front. Public Health* 6, 172.
- Sperling, G.B., Winthrop, R., 2015. *What Works in Girls' Education: Evidence for the World's Best Investment*. Brookings Institution Press.
- Sripakash, A., Tikly, L., Walker, S., 2020. The erasures of racism in education and international development: Re-reading the 'global learning crisis'. *Compare* 50 (5), 676–692.
- Stromquist, N.P., 2012. The gender dimension in the World Bank's education strategy: assertions in need of a theory. In: *The World Bank and Education*. Brill, pp. 159–172.
- Summers, L., 1992. 'Investing in All the People. Educating Women in Developing Countries' EDI Seminar Paper Presented at Development Economics Seminar. World Bank.
- Unterhalter, E., 1996. States, households and the market in World Bank discourses, 1985–1995: a feminist critique. *Discourse: Studies in the Cultural Politics of Education* 17 (3), 389–401.
- Unterhalter, E., 2000. Transnational visions of the 1990s. Contrasting views of women, education and citizenship. In: Arnot, M., Dillabough, J.A. (Eds.), *Challenging Democracy: Feminist Perspectives on the Education of Citizens*. Routledge, London, pp. 87–102. Arnot, M. and Dillabough, J.A.
- Unterhalter, E., 2003. The capabilities approach and gendered education. An examination of South African complexities. *Theor. Res. Educ.* 1 (1), 7–22.
- Unterhalter, E., 2005c. Fragmented frameworks: researching women, gender, education and development. In: Aikman, S., Unterhalter, E., 'Introduction', Aikman, S., Unterhalter, E. (Eds.), *Beyond Access: Developing Gender Equality in Education*. Oxfam Publishing, Oxford, pp. 15–35.
- Unterhalter, E., 2005a. Global inequality, capabilities, social justice and the Millennium Development Goal for gender equality in education. *Int. J. Educ. Dev.* 25, 111–122.
- Unterhalter, E., 2005b. Global values and gender equality in education: needs, rights and capabilities. In: Fennell, S., Arnot, M. (Eds.), *Gender Education and Equality in a Global Context*. Routledge, London, pp. pp31–46.
- Unterhalter, E., 2008a. *Gender, Schooling and Global Social Justice*. Routledge, Abingdon.
- Unterhalter, E., 2008b. The capability approach and gendered education: some issues of operationalisation in the context of the HIV/Aids epidemic in South Africa. In: Alkire, S., Comin, F., Quisilbash, M. (Eds.), *The Capability Approach: Concepts, Measures and Applications*. Cambridge University Press, Cambridge.
- Unterhalter, E., 2009. Translations and transversal dialogues: an examination of mobilities associated with gender, education and global poverty reduction. *Comp. Educ.* 45 (3), 329–345.
- Unterhalter, E., 2012. *Mutable Meanings: Gender Equality in Education and International Rights Frameworks*.
- Unterhalter, E., 2013. Connecting the private and the public: pregnancy, exclusion and the expansion of schooling in Africa. *Gend. Educ.* 25 (1), 75–90.
- Unterhalter, E., 2017. A Review of public private partnerships around girls' education in developing countries: flicking gender equality on and off. *J. Int. Comp. Soc. Policy* 33 (2), 181–199.
- Unterhalter, E., 2019a. The many meanings of quality education: politics of targets and indicators in SDG 4. *Global Policy* 10, 39–51.
- Unterhalter, E., 2019b. Balancing pessimism of the intellect and optimism of the will: some reflections on the capability approach, gender, empowerment, and education. In: Clark, D., Frediani, A., Biggeri, M. (Eds.), *The Capability Approach, Empowerment and Participation: Concepts, Methods and Applications*. Palgrave, London, pp. 75–100.
- Unterhalter, E., 2020. An other self? Education, foreignness, reflexive comparison and capability as connection. *Comp. Educ.* 56 (1), 3–19.

- Unterhalter, E., 2021. Addressing intersecting inequalities in education. In: McCowan, T., Unterhalter, E. (Eds.), *Education and International Development: An Introduction*, second ed. Bloomsbury, London, pp. 149–168.
- Unterhalter, E., North, A., 2018. *Education, Poverty and Global Goals for Gender Equality: How People Make Policy Happen*. Routledge, Abingdon.
- Unterhalter, E., Robinson, L., 2020. Precise evasions. PPPs, aid and education. In: Gideon, J., Unterhalter, E. (Eds.), *Critical Reflections on Public Private Partnerships*. Routledge, Abingdon, pp. 98–114.
- Unterhalter, E., North, A., Arnot, M., Lloyd, C., Moletsane, L., Murphy-Graham, E., Saito, M., 2014. *Education Rigorous Literature Review: Interventions to Enhance Girls' Education and Gender Equality*. DFID, London.
- Unterhalter, E., Robinson, L., Balsera, M., 2020. *The Politics, Policies and Practices of Intersectionality: Making Gender Equality Inclusive and Equitable in and through Education*. Background paper prepared for the Global Education Monitoring Report Gender Report. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000374495>.
- Unterhalter, E., Longlands, H., Peppin Vaughan, R., 2022. Gender and intersecting inequalities in education: reflections on a framework for measurement. *J. Hum. Dev. Capab.* <https://doi.org/10.1080/19452829.2022.2090523>.
- Vollset, S.E., Goren, E., Yuan, C.W., Cao, J., Smith, A.E., Hsiao, T., Murray, C.J., 2020. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. *Lancet* 396 (10258), 1285–1306.
- Zakaria, R., 2021. *Against White Feminism: Notes on Disruption*. WW Norton & Company.
- Zwingel, S., Zwingel, 2016. *Translating International Women's Rights*, vol. 300. Palgrave Macmillan UK, London.

Global connective media: YouTube as an educational infrastructure

Inés Dussel^a and Patricia Ferrante^b, ^a DIE/Cinvestav, CDMX, Mexico; and ^b FLACSO Argentina, Ciudad Autónoma de Buenos Aires, Argentina

© 2023 Elsevier Ltd. All rights reserved.

Introduction	622
YouTube as a global connective platform	623
YouTube as a knowledge platform: popularity algorithms and the logics of searching	624
Using YouTube videos for school study: alignments and recalibrations	625
Students' uses of YouTube: the quest for "fixating knowledge in one's mind" and "getting it right"	625
Teachers' uses of YouTube: the uphill battle for attention and knowledge	626
Concluding remarks	627
Acknowledgment	628
References	628

Introduction

The expansion of digital media, including platforms, social media, and smartphones, brought a new wave of globalization in the realm of culture and technologies (Jin, 2020). Production, consumption, and distribution of cultural goods have mostly shifted toward the digital domain, with platforms as key players. Education and everyday school practices have not been spared by this trend, both because of an increasing platformization of teaching, learning and managerial practices (van Dijck et al., 2018) and because search engines, apps and platforms are now a significant part of the lives of students and teachers and have a role in defining what counts as knowledge and learning (Livingstone and Sefton-Green, 2016; Alirezabeigi, 2021; Perrotta et al., 2021, among many others).

At the same time, it is important to locate schools within media settings that predate digitalization. In the *longue durée*, schooling practices have always been entangled with media practices that involve print, sound, and other technologies that have shaped the ways in which knowledge is produced and circulates across bodies and spaces (Caruso, 2019; Day Good, 2020). These media were also global: atlases, journals, textbooks, globes, blackboards, desks, radio, and film, were often internationally produced and sold, but they were also picked up by local business and adapted to vernacular practices. Thus, the opposition of the global and the local should be unpacked and problematized. Even if events only happen at the material-local scale, they are always constituted by global flows, be it of seeds, waters, animals or human beings, knowledges, or technologies (Galison, 2016), and what is considered global can also be analyzed as a particular site with longer networks and more powerful mediators (Strathern, 1996).

Digital media, then, are not the first nor the last global media, and the extent to which they are "global" merits further analysis (Burrell, 2012; Srinivasan, 2017). Yet they have introduced some breakthroughs in the production and circulation of knowledge that are having ample effects on educational practices. One main breakthrough was to configure new economies of time, quicker, shorter, more intense, and with a capacity to calibrate or coordinate larger areas of human and non-human activity (Agar, 2006; Crary, 2013). Another significant change was the growing datafication of these activities, which made human sociality increasingly technologically mediated. Datafication allows what van Dijck (2013) has termed as connectivity (the use of data made by tech companies to make recommendations and feed our digital practices), different from connectedness (what users make public). Connective digital media are today deeply enmeshed in our social fabric, and corporate forces are dominant in shaping the meanings not only of what constitutes public issues but also of intimate practices such as friendship and sexuality (van Dijck, 2013; Zuboff, 2019).

Connective, that is, commercialized global digital media also impact on how schools are shaped and situated in other geopolitical spaces, bringing into the classroom global narratives that flow through platforms that set different rules and arenas for participation and for knowledge practices (Williamson and Hogan, 2020; Dussel and Fuentes Cardona, 2021). Nowadays, education and everyday classroom activities are entangled in a corporate media and technology realm dominated by a handful of powerful companies, mostly based in the United States. It is the world of GAFAM, or more recently AMAMA (as Google and Facebook have changed to Alphabet and Meta), in which any exchange throughout digital media is at some point related to Alphabet, Apple, Amazon, Meta, or Microsoft (van Dijck et al., 2018). Education and everyday school practices are part of what has been called platform capitalism (Smieck, 2018) or surveillance capitalism (Zuboff, 2019), either by institutional decisions in the buying or using of digital devices and platforms, or through users' daily digital practices (Selwyn et al., 2022).

In the following pages, we discuss the spread of digital connective media in schools through an analysis of the uses of YouTube for school study in a secondary school in Argentina, which is part of a broader research project described below. While in previous studies we approached digital media mostly as a means for cultural production and participation and were concerned with identity, voice, and community (Dussel and Benasayag, 2019; Winocur and Dussel, 2021), we would like to look at global connective media as part of the shift to consider digital infrastructures and materialities (Plantin et al., 2018; Perrotta et al., 2021). We are interested in

understanding how digital platforms “leave their marks on the knowledge they help us produce” (Gitelman, 2014, p. 74); this is visible in the uses of YouTube for school study, in which algorithmic logics and school pedagogies might offer competitive orderings. We argue that in these practices of study with global connective media students unfold a set of practices that, mostly inadvertently, tends to be regulated by platforms’ protocols and users’ practices rather than by school curriculum or pedagogies.

Throughout the chapter, we aim to identify and describe changing conceptions of school knowledge in a media landscape “swamped with information, much of which we are in no position to assess” (Appadurai, 2020). Knowledge practices seem to be shifting toward data management and connectivity; knowledge is measured no longer with criteria related to their precision and truthfulness toward facts (Poovey, 2017), but according to affective qualities such as smartness, sentience, and efficiency for becoming sticky and viral (Halpern, 2014; Jenkins et al., 2013). We are interested in discussing the kinds of knowledge that are mobilized in students’ and teachers’ engagements with the video platform, and the autonomy they can exercise in relation to algorithmic platforms.

To approach these issues, we start with an analysis of YouTube as a digital connective platform, which is followed by a discussion of its educational uses, as they appear in the literature and in documents issued by Alphabet, its owner. We then present some findings from our research project with secondary school teachers and students in the city of Buenos Aires. Finally, we offer some concluding remarks on what these findings might mean in terms of the knowledge flows and practices configured by global connective media in their intersections with schools.

YouTube as a global connective platform

YouTube is part of a larger ecology of digital connective media but has some specificities that are worth of further analysis. The launch and extremely fast expansion of YouTube changed drastically the access to audiovisual content and distribution. YouTube presented itself originally as a social network for sharing amateur video and its first slogan was “*Your digital video repository*,” later replaced by a more accurate one in relation to the practices taking place in YouTube: “*Broadcast yourself*.”

The platform was only 18 months old when a short amateur video showed YouTube founders, Jawed Karim, Steve Chen, and Chad Hurley, announcing that Google was buying the platform for u\$1.65 billion dollars, turning it into one of the fastest success cases in Internet history and signaling the shift toward video internet, a kind of consumption that would become dominant. By 2020, YouTube have more than 2 billion users, one-third of the total of connected people in the world, and 1 billion hours are seen per day; it should be remembered that this does not include China, due to the resistance of the Chinese government to allow Google to operate in its territory.

Both in terms of quantity and quality, YouTube radically changed the ways of accessing, producing, and sharing video content, not only for amateur videos, but for the whole entertainment and film industries that found in the platform a powerful channel to distribute and control their content (Strangelove, 2010). Without a calendar, schedule, or limits to the times a video can be seen, it marked the end of the dominance of broadcasting and started what philosopher Bernard Stiegler called the carnival of the new screen, in which users have cardinal access to discrete (that is, severed from the flow of broadcasting) content units at their own whim (Stiegler, 2009), and that advanced some sort of business model that led to streaming services. Also, considering that users can decide what and when to access content, viewers’ attention becomes the center of a battle to capture a “time of available brains,” that is, a disposition to pay attention and produce more traceable interactions with the platforms (Stiegler, 2009, p. 44). Consistent with forms of digital citizenship that were turning carnivalesque (Papacharissi, 2010), YouTube promotes new formats and genres such as YouTubers and other influencers that rely on their popularity on the platform, based on the amount and duration of viewing and on social media reactions that are datafied and monetized.

As Snickars and Vonderau (2009) point out, the peculiarity of YouTube is the way in which the platform managed to navigate between community and commerce. If something defines the platform’s identity is that it is part of the industry and part of many communities, which still use it as a venue to upload content and visibilize themselves. The relationship between YouTube and audiovisual industries was initially tense, as corporations saw the platform as an infringement of property rights (van Dijk, 2013), but due to its success YouTube became an ally in promotional campaigns and as another screen in which artists and products wanted to be seen. Thus, even if YouTube was born as a user generated content (UGC) platform, today it is part of the conglomerate of global entertainment platforms in which the distinction between amateur and professional content can be considered irrelevant (Scolari and Fraticelli, 2017).

As a symptom of its success, watching YouTube is now synonymous with watching TV for a large part of the population and the platform is always one of the three most visited websites wherever it operates (Cunningham and Craig, 2020). In that respect, YouTube is one of the winners of the “battle for attention” denounced by Stiegler (2009). This is not by chance; along its history, YouTube adjusted its algorithm to be more effective. The breakthrough came in 2016, when the company set in motion what they called Deep Neural Networks, which made use of data from previous behavior and tailored its recommendations with greater precision (Covington et al., 2016). As it will be shown below, the platform’s recommendations feed is one that students value and use frequently in their school assignments.

In relation to its global reach, Cunningham and Craig (2019) state that YouTube exhibits much greater content, creator, linguistic, and cultural diversity than any other global media titan, maybe because its requirements for uploading content are lower than in platforms where writing is still key and because it has become a privileged venue for media activism (Meikle, 2020). Even when YouTube headquarters are US based, 80% of its traffic comes from other countries (Cunningham and Craig, 2019). Since

2013, YouTube offers the possibility of auto-captioning and translating contents into 14 languages, at least until December 2021. Yet this diversity should not be overstated: the availability of global content is filtered by algorithms that establish priorities and hierarchies, in which commercial interests play a significant role, and ultimately this circulation is determined by data profiles and data traces monopolized by the company, in a process that can be considered as a new colonialism, a data colonialism (Couldry and Mejías, 2019).

Are nation-states still relevant in the corporate media landscape that YouTube has helped to shape? National frontiers and local communities appear less important as youth identities are forged in consumption styles and media fandom across the globe in seemingly more appealing ways than older political or legal constituencies (Rizvi, 2012). Following Castells (2013) and Jin (2020), global media is not freed from the national state, but defined by the global distribution enabled by the existence of digital networks and by the directions of flow of people, culture, and capital. And while the nation state and national interests are still significant actors in some spaces (as the example of China or Iran show), their possibilities of setting guidelines and directions are limited “in a de-territorialized world marked by conflict over internet cables, advanced technologies, and market access” (Moi-sio, 2018, p. 5). Yet perhaps “de-territorial” is not the best way to describe this world, as the practices that define it are concrete and material, such as mining for minerals critical to producing devices, setting offshore factories with cheap labor to assemble them, or producing or consuming content. This hyper-connected world is shaped by asymmetries related to how these global flows are made local and recirculated, and the unequal and uneven power relations of key agents. Also, it is important to note that in youth’s engagements within global media, nations continue to be significant levels “of place and affiliation” (McLeod, 2009, p. 284), and schools may play a role in expanding or shrinking the width and density of knowledge flows, as will be seen below.

YouTube as a knowledge platform: popularity algorithms and the logics of searching

In this section we would like to analyze YouTube as part of broader knowledge infrastructures, both in terms of educational content and of how knowledge is defined and circulates. As said before, YouTube is not the first nor the last educational media. Yet, even if instructional videos were contemporary to the birth of cinema and expanded with domestic video devices (Orgeron et al., 2012), there is no doubt that their availability and consumption grew enormously with the expansion of YouTube. The most popular formats are graphic animations with a voice-over that presents curricular content, like the instructional videos produced and posted by Khan Academy, an online learning provider developed by an NGO yet integrated into a larger commercial distribution circuit in which Google-YouTube are key players (Williamson and Hogan, 2020).

Another popular format are videos uploaded by what in Spanish are called *edutubers*, that is YouTube influencers who teach school content mostly through short videos and who are subjected to the same dynamics of monetization and popularization than other producers (Scolari and Fraticelli, 2017; López Aguilar et al., 2020). Most of them are male, confirming that digital platforms are far from an equal playing field (Noble, 2018; Eubanks, 2018; O’Neil, 2016; Regueira et al., 2020). It is also very likely that the prevalence of male YouTubers might be related to the amount of hate speech female YouTubers receive in comparison to males (Döring and Mohseni, 2020). In Argentina, the most popular YouTube educational videos are in Spanish, but not always locally produced. *EduTubers* from Colombia, Mexico, and Spain share the top positions with content produced by local teachers or TV channels that massively distribute their contents throughout the platform (López Aguilar et al., 2020).

Considered as a knowledge infrastructure, YouTube is not only a venue for displaying educational content created by users but is also becoming a key player in the definition of what counts as education and how knowledge is organized and hierarchized. “Learning” is listed within YouTube’s exploring categories, competing with live videos, videogames, movies, and sports. Under this tab, users can find “how-to” videos and school content videos that include cultural content, marketing courses, or tutorials for the most diverse range of needs. Also, during the COVID-19 pandemic YouTube created a hub for educational content, Learn@home, consisting of recommendation lists of videos with school contents. In these lists, state educational agencies and traditional TV channels were lined up with Khan Academy, TEDx, and amateur creators successfully turned into *edutubers*. Interestingly, research conducted before the pandemic showed that in scientific videos, institutional channels, even run by prestigious research such as the Max Planck Institute in Germany and state research institutes in France, fared worse than YouTube influencers in views, subscribers, and engagement (Donhauser and Beck, 2021; Debove et al., 2021); the mix-up of amateurs and institutions seems to be a good move to capture viewers’ attention, but shifts power and authority to new knowledge figures who are validated by their popularity.

However, what is perhaps YouTube’s most decisive intervention on the infrastructures of knowledge is its algorithm for searching. As part of Alphabet, YouTube shares with Google a searching platform that establishes patterns out of words and produces a list of results hierarchized by its popularity (most visited sites) among other criteria that are opaque due to the proprietary quality of the platform’s algorithm. Word-searching becomes the key to producing knowledge, which appears as morsels of information that can be tagged and indexed. As Appadurai and Morozov state, “searchability has the indirect effect of eliminating the need for reading, cross-referencing, recall, retrieval and memory,” and endows “the digital search with oracle qualities” (Appadurai and Morozov, 2021, p. 3). It seems that humans can access or acquire knowledge without having to appeal to any other intellectual and ethical considerations. This reduction of knowledge to data retrieval and management seems to carry several risks, among them a delegation of knowledge practices to the platforms and the marginalization of alternative epistemologies and ways of knowledge that might not appear in the top 10 results but that could enrich our lives and make our cultures more plural and democratic.

In the last years, researchers have started to analyze how YouTube videos are used in schools. [Moghavvemi et al. \(2018\)](#) and [Zhou et al. \(2020\)](#) state that students assign a positive value to studying with YouTube videos, because they feel they enhance their learning experience with visual aids or because they help them consolidate or sediment their learning. [Gil-Quintana et al. \(2020\)](#), in a quantitative and qualitative study with Italian teenagers which use YouTube videos to study mathematics, found that while YouTubers are more valued than teachers as communicators, most students still prefer to interact with teachers and peers in their learning processes; in their study, students with lower academic grades use YouTube videos more frequently, and have more followers in social media platforms, showing a certain disparity between school capital and social media capital. In Argentina, in 2018 one of us conducted qualitative research on uses of digital devices by university students from public and private institutions in the city of Buenos Aires ([Ferrante and Igarzábal, 2018](#)), and found a frequent use of YouTube for studying, mostly during exams' periods. Students privileged TED talks or short animated videos over conferences or keynotes for understanding social or biological processes but also complex theoretical concepts. They searched with words related to the content they were studying and adding "explanation" as a keyword; the videos were selected with criteria such as popularity and recommendations by the platform. As it will be shown, these steps are practically the same used by secondary school students in their navigation of the platform.

Using YouTube videos for school study: alignments and recalibrations

We would like to briefly present some of the reflections about uses of YouTube in schools that are emerging in an ongoing research project on digital inequalities in schools (RED, see acknowledgments). We conducted our research in a private secondary school located in the City of Buenos Aires that receives middle and high-income students; we were planning to conduct research in low-income schools, but it was interrupted due to the pandemic and will be resumed in 2022. Secondary schooling lasts 5 or 6 years in Argentina; in the city of Buenos Aires, the duration is 5 years and we focused on the 3rd (middle) and 5th (last) year of schooling, which generally include students of 15–18 years of age.

Through online surveys and on-site interviews, we looked at how teachers (3) and students (69) use YouTube for school purposes and inquired about their search criteria, video valuation, and other critical aspects in this use. The teachers taught Introduction to social sciences (3rd year) and Social research projects (5th year), and frequently used digital media in their teaching. The students were 15–18 years old, and 34 were female, 31 male, and 4 non-binary or preferred not to tell; 63 students owned a smartphone and 61 had a computer or laptop for their exclusive use.

The school was founded in 1988 by a group of progressive educators and is focused on arts and communication. Its pedagogies lean toward social constructivism and critical thinking, and the curriculum is flexible and open to responding to students' interests and needs. All teachers are university graduates, with degrees in science, humanities, or social sciences, and some of them are pursuing or have finished graduate studies. All classrooms have Internet connectivity and are equipped with screens and projectors; students bring their cellphones and most have laptops at home. As per reported in the interviews, most teachers use audiovisual materials in their presentations, including films, and ask students to make digital productions as school assignments. Considering its material and human resources, this school seems to be in a good position to exercise some autonomy in relation to what is promoted by algorithmic platforms. However, as it will be argued, this has proven to be more difficult than expected.

Students' uses of YouTube: the quest for "fixating knowledge in one's mind" and "getting it right"

Unsurprisingly, students reported a massive use of YouTube: only one in 69 responses said they (gender unknown) doesn't use YouTube at all, and most use it more than once a day in their mobile phones. They tend to watch YouTubers and music videos; less than half reported watching funny videos, videogames or video gamers, tutorials and sports or series recaps. Other frequent but less popular choices are Japanese anime, films, conferences, documentaries, and TV shows.

When they watch for leisure, students said that they do not perform any search but rely on what the platform offers in their personalized profiles; they also watch videos that are shared by friends or that just "pop" into their mobiles, analogous with what [Boczkowski et al. \(2018\)](#) describe as incidental news consumption. This describes how the logic of algorithm operates: searching is not necessary because the platform "knows" what one likes to see, value and share. Students said that they like the recommendations made by the algorithm as they save them time; they considered it good that the platform anticipates their interests and see no harm in it.

However, when using YouTube for studying, students report that they do search, using the words that are related to the content they need to study plus the word "explanation:" this word seems to be good to find educational videos, a thread that will be followed later. Among the results, students reported to choose videos according to the "usefulness" of its content for acquiring knowledge or improving test scores, but they also stick to what appears in the first five options or look for popular videos with many likes, which give primacy to the hierarchies set by the knowledge infrastructure of the platform. Only 9 students said they consider the source or the author of the video as a criterion to select it.

Duration of the videos is another key issue for selecting what to see for studying. In the survey, most students said that they watch videos that last between 5 and 10 min; only 7 students watch videos of 30 min or more. Most students reported that they value videos whose narrative is precise and gets "straight to the point." In the interviews, students related the short duration of videos

with their ability to summarize content, and said that if faced with longer videos, they use the forwarding key to find the content they are looking for. Few students mentioned rewinding and watching videos again for further or better understanding; as Appadurai (2020) remarks, users do not value cross-referencing or further readings.

In relation to their genre, the most popular videos for studying are *edutubers*, animations, presentations, and to a lesser extent video lessons and documentaries. While students mentioned *edutubers*, they were not able to refer to their names, in a similar line than what Gil-Quintana et al. (2020) found in Italian teenagers. The only exception was an *edutuber* called #JULIOPROFE, a Colombian teacher turned entrepreneur with more than 4.72 million followers who publishes short and unsophisticated videos, usually lasting no more than 5 min, that include curricular content for mathematics and physics. In the interviews, students referred to “YouTube videos,” irrespective of who produced them; the platform becomes a kind of authorial voice that blurs their qualities and specific traits (Doueiri, 2010). Asked again about naming specific YouTubers or sources, students mentioned English and Spanish channels and a Mexican math *edutuber*; students could not identify any Argentinean channel or *edutuber* as useful for studying, although their teachers frequently share videos from local sources. Their references seem to come from the platform itself and from the global flow of materials organized by its algorithms; national boundaries do not seem to organize their virtual libraries.

Once they select videos, how do students judge their value for school study? 61 students considered that a good video is the one that “explains well,” again referring to explanation as a pedagogical ideal; they also valued that the video is directly related to what teachers said in class and privileged its ability to summarize content. The videos that they find the most useful are the ones that help them “fixate” concepts, ideas, or processes in their minds and eventually help them get better test scores.

In these students’ knowledge practices, YouTube educational videos do not seem to bring in constructive pedagogies or arouse curiosity, even though the school’s ethos and curriculum promotes these ideals. On the contrary: these videos are brought for specific uses that are related to improving their academic performance and are mostly used when studying for tests. The relationship to knowledge that these assumptions carry is one in which knowledge is a pre-defined entity, a finished and stable product that is out there and that is imprinted on the mind. Once it is imprinted, it stays there, thus the metaphor of fixating (Charlot, 2007). In their video preferences, students seem to reproduce what Jacques Rancière calls as “explicative order,” in which learners, deemed unable to find it for themselves, are given an explanation “for comprehension to take place.” In this relationship, “the words of the master must shatter the silence of the taught material” (1991, p. 4), closing any chance that students raise questions and compare the answers for themselves. For Rancière, the explicative order always reinstates the distance between teacher and student and is an order of *enforced stultification* (p. 5).

What emerges in the survey and interviews are practices that are aligned with the platform’s algorithms: preferences for influencers, short duration, popularity as judged by viewers and likes, and transnational media content. It is noteworthy that these knowledge practices are not the ones that are prevalent in this school, at least as seen and heard in interviews and classes; students seem to evoke and reenact prior study practices that consider knowledge as a finished, highly stable product that must be imprinted or inserted in their minds -the “explicative order” denounced by Rancière (1991). In that respect, YouTube videos contribute to reinforce this conception of knowledge and provide a shortcut that make their study practices more time-efficient (“get to the point”), but not necessarily richer or denser.

Teachers’ uses of YouTube: the uphill battle for attention and knowledge

In this school, teachers frequently use videos in their classes to illustrate content and provide other sources for curriculum work. Among the genres, they bring movies (complete or fragments), documentary style TV shows (mostly produced by the Argentinean public cultural channel *Encuentro*), TED talks and, to a lesser extent, animated short videos. The use of videos grew exponentially during the pandemic, as teachers had to teach online and in hybrid classrooms; they used the school’s learning platform to distribute content but also shared links to YouTube when available. This selection of videos was cherished by students, who said that these are good choices because “teachers know best,” “they search better” and “are more trustworthy.” Students also valued that these videos were directly connected to the lesson plan and the required readings; they saw them as consistent with the curriculum. However, they recognized that the videos selected by their teachers never pop up or appear in their own searches, as they are not popular, and these videos are not the ones they use for revising or “fixating” knowledge.

Teachers said that they are aware that students use YouTube videos for studying and that duration and popularity are the main criteria at play in their selection, whose shortcomings they recognize. In relation to duration, teachers shared with students the unwritten imperative that a video must be short to be good and capture students’ attention. Videos are, for them, conveyors of information; they don’t discuss their style or genres or promote visual literacy. Teachers reported that they feel the pressure to produce or share only short videos, because they wanted to promote students’ engagement and “avoid boredom.”

In this sense, teachers are also engulfed by an economy of attention that is characterized by fragmentation, and in which speed is much more valued than depth (Crary, 2013). These short videos were used for starting a lesson or triggering a discussion, but teachers recognized that subjects cannot be taught in a few minutes and that the videos need to be integrated into larger sequences. Moreover, they considered that their short duration can play against a better understanding of complex phenomena; as one teacher said: “you cannot explain the Spanish Civil War in 10 min, you just can’t.”

Even more than duration, teachers are mostly concerned about media content that might be inappropriate or ideologically questionable, such as racism or sexism. Aware of these problems, the school promotes projects to foster critical digital literacies. However, one teacher narrated an activity that she had designed to improve critical research skills, which was turned upside down by her students:

Gender is a key issue for students, so I asked them to bring to the classroom sexist memes that they found in social media, expecting them to do some research in their own profiles and decide which images were offensive. Instead, they went to Google and typed the words “meme” and “machismo,” and brought the first images provided by the search engine. They did not think why those images were sexist, they just googled and copied the results (Female teacher, Social Research projects).

These practices show the extent to which students delegate decisions to algorithms: the critical gaze is defined by the very platform the activity was supposed to challenge. Certainly, students might have good reasons not to share with their teacher what goes on in their social media profiles, but teachers find it difficult to interrupt this delegation of knowledge and affirm projects in which students are more reflective of their media uses.

Reenacting a position that was already prevalent in critical media literacy, teachers assume a critical ideological stance toward digital media, but it is only recently that they have started to promote further reflections on the architecture of knowledge and the effects of popularity algorithms, and particularly on what Google and YouTube offers them based on their previous searches. Issues around datafication are starting to figure prominently in their lessons. One teacher highlighted that he tries to teach students how to work with sources, design research, and go beyond the first results they find in Google; however, he was not optimistic: *“I always have the feeling that I can’t struggle against this, it’s just too big.”* The image of the uphill battle emerged consistently in the interviews with teachers, and it is probably an accurate metaphor for the difficulties individual teachers are facing when confronted to the knowledge practices that connective global media are promoting.

Concluding remarks

There is no doubt that global connective media are becoming relevant actors in education, providing not just a virtual library in which to find educational content or apps to provide tailor-made support, but also datafying practices and shaping what counts as knowledge in schools and elsewhere. In our research on uses of YouTube for school study, it was clear that the algorithms for searching contribute to morselizing knowledge into parcels that must be “found,” conceived as complete and discrete chunks that provide finished and stable knowledge. Videos are valued because they are short and get to the point, and because they were liked and viewed by other users. Through its algorithm, YouTube privileges affective engagement and quick interactions; it confirms the shift from valuing critical distance, reflection, and argumentation as knowledge practices in favor of immediate emotional adhesion and automated responses to manage content (van Dijck, 2013; Poovey, 2017; Appadurai and Morozov, 2021).

The school in which we did our research tries to mobilize progressive pedagogies to overcome the traditional “explicative order” and counterbalance digital platforms, but teachers feel like Sisyphus, carrying a stone deemed to fall on themselves. Students search for short videos that provide explanations; while they value what their teachers offer, they use other criteria in their autonomous searches, following the first videos on the list or the comments made by other viewers. Videos are not criticized or studied as texts; both for teachers and students, they appear as conveyors of information. Interestingly, students prefer to watch *edutubers*, whose videos are not highly produced but that “get to the point” that the curriculum demands. Their choices are more aligned with the “explicative order” of traditional schooling, now amplified by global connective media, than with their own progressive school, which values inquiry, research, and creation.

More research should be done on whether different teaching strategies would lead to other results. It is not surprising that, even within progressive pedagogies, students remain concerned about tests, which continue to be a pivotal center of institutionalized schooling. But it is also important to acknowledge that there is always a tension in emancipatory pedagogies that try to move away from the “explicative order” denounced by Rancière (1991) and still seek to provide some ordering or cultural hierarchy that risks reinstating an explicative order (Bingham and Biesta, 2010). These debates and quandaries should be brought to the fore in critical digital pedagogies to open up more questions and research projects on what these pedagogies are achieving.

Also, there are methodological limitations in our own approach. It is possible that students learn more than they claim from their teachers, and that they are able to use critical skills and strategies learned in school in other contexts; algorithm learning might not be as powerful as it seems if a different angle is taken. However, the difficulties voiced by these teachers should be taken seriously. As Lisa Gitelman says, “search engines and platforms [are] helping to change how we know and what we know” (2014, p. 80), and while datafication is celebrated by some because it makes life easier, its epistemic and political costs remain opaque and are not present on the public agenda.

The final point we would like to make is related to this opaqueness. Students and teachers did not refer to public policies or to public discussions; these issues seem to be confined to teachers’ individual orientations and resources, and to students’ personal tastes and convenience. Users’ practices are following the protocols of global connective media, and school pedagogies, even when they are willing to, are not powerful enough to recalibrate them toward other goals and practices that might uncouple these

practices from the platforms' business models and reorient them toward the common good (Appadurai and Morozov, 2021). Teachers and students alone cannot solve this; unless these trends and challenges are turned into public issues and are taken by larger constituencies such as universities and curriculum makers, this will continue to be an uphill battle.

Acknowledgment

This research was possible thanks to the generous support of Rijksbanken Jubileum Fond, which funds the project RED (Reconfigurations of Educational (In)Equalities in a Digital World) (GI19-1500), on which we are inscribed. We are grateful to our colleagues for the continuous conversations and discussions on digital media, schools, and inequalities. We also want to thank Benito Castro Pérez, who is conducting a similar study in Mexican schools and helped us refine our arguments.

References

- Agar, J., 2006. What difference did computers make? *Soc. Stud. Sci.* 36 (6), 869–907.
- Alirezabeigi, S., 2021. In Between Continuity and Rupture: On Taskified Classroom Practices. Ph.D. Dissertation. Faculty of Psychology and Educational Sciences, KU-Leuven.
- Appadurai, A., Morozov, E., 2021. Digital challenges. In: Document Produced for the Futures of Education Initiative (Mimeo). March 2021.
- Appadurai, A., 2020. Gradual Learning, the Urgency of Knowledge and the Connectivity of Humanity. UNESCO Futures of Education Ideas LAB. Available at: <https://en.unesco.org/futuresofeducation/appadurai-gradual-learning-urgency-knowledge>.
- Bingham, C., Biesta, G., 2010. Jacques Rancière: Education, Truth, Emancipation. Bloomsbury Publishing, London.
- Boczkowski, P.J., Mitchelstein, E., Matassi, M., 2018. "News comes across when I'm in a moment of leisure": understanding the practices of incidental news consumption on social media. *New Media Soc.* 20 (10), 3523–3539. <https://doi.org/10.1177/1461444817750396>.
- Burrell, J., 2012. Invisible Users. Youth in the Internet Cafés of Urban Ghana. The MIT Press, Cambridge, MA.
- Caruso, M., 2019. Geschichte der Bildung und Erziehung. Medienentwicklung und Medienwandel. Verlag Brill-Schöningh, Paderborn.
- Castells, M., 2013. Communication Power. Oxford University Press, Oxford, UK.
- Charlot, B., 2007. La relación con el saber. Libros del Zorzal, Buenos Aires.
- Couldry, N., Mejias, U., 2019. The Costs of Connection. How Data is Colonizing Human Life and Appropriating it for Capitalism. Stanford University Press, Palo Alto, CA.
- Covington, P., Adams, J., Sargin, E., 2016. Deep Neural Networks for YouTube Recommendations. Google. Mountain View, California. <https://doi.org/10.1145/2959100.2959190>.
- Crary, J., 2013. 24/7: Late Capitalism and the Ends of Sleep. Verso, London.
- Cunningham, S., Craig, D., 2019. Social Media Entertainment. The New Intersection of Hollywood and Silicon Valley. New York University Press, New York.
- Cunningham, S., Craig, D., 2020. Global social media entertainment. In: Shimpach, S. (Ed.), *The Routledge Companion to Global Television*. Routledge, New York, pp. 49–59.
- Day Good, K., 2020. Bring the World to the Child. Technologies of Global Citizenship in American Education. The MIT Press, Cambridge, MA.
- Debove, S., Füchslin, T., Louis, T., Masselot, P., 2021. French science communication on YouTube: a survey of individual and institutional communicators and their channel characteristics. *Front. Commun.* 6, 612667. <https://doi.org/10.3389/fcomm.2021.612667>.
- Donhauser, D., Beck, C., 2021. Pushing the Max Planck YouTube channel with the help of influencers. *Front. Commun.* 5, 601168. <https://doi.org/10.3389/fcomm.2020.601168>.
- Döring, N., Mohseni, R., 2020. Gendered hate speech in YouTube and YouTube comments: results of two content analyses. *Stud. Commun. Media* 9 (1), 62–88.
- Doueih, M., 2010. La gran conversión digital. Fondo de Cultura Económica, Buenos Aires.
- Dussel, I., Benasayag, A., 2019. Digital media production outside the school. Youth knowledge and cultural participation in Argentina and Mexico. In: Sefton-Green, J., Erstad, O. (Eds.), *Learning Beyond the School: International Perspectives on the Schooled Society*. Routledge, London, pp. 133–150.
- Dussel, I., Fuentes Cardona, G., 2021. Los grupos de WhatsApp y la construcción de nuevas ciudadanías en las escuelas: Un estudio de caso con docentes en Ciudad de México, vol. 42. *Educación e Sociedade*, CEDES-UNICAMP, Brasil, pp. 1–18. <https://doi.org/10.1590/es.251642>.
- Eubanks, V., 2018. Automating Inequality. How High-Tech Tools Profile, Police and Punish the Poor. St. Martin's Press, London.
- Ferrante, P., Igarzábal, B., 2018. Experiencias culturales de la juventud en un mundo conectado. Paper presented to Congreso del Centro Estudios sobre Medios y Sociedad en Argentina (MESO) 2018. Universidad de San Andrés, Buenos Aires, Argentina.
- Galison, P., 2016. Limits of localism: the scale of sight. In: Doniger, W., Galison, P., Neiman, S. (Eds.), *What Reason Promises. Essays on Reason, Nature, and History*. De Gruyter, Berlin/Boston, pp. 155–170.
- Gil-Quintana, J., Malvasi, V., Castillo-Abdul, B., Romero-Rodríguez, L.M., 2020. Learning leaders: teachers or YouTubers? Participatory culture and STEM competencies in Italian secondary school students. *Sustainability* 12, 7466. <https://doi.org/10.3390/su12187466>.
- Gitelman, L., 2014. Searching and thinking about searching JSTOR. *Representations* 127 (1), 73–82.
- Halpern, O., 2014. Beautiful Data. A History of Vision and Reason Since 1945. Duke University Press, Durham, NC & London.
- Jenkins, H., Ford, S., Green, S., 2013. Spreadable Media. Creating Value and Meaning in a Networked Culture. New York University Press, New York.
- Jin, D.Y., 2020. Globalization and Media in the Digital Platform Age. Routledge-Taylor & Francis, New York and London.
- Livingstone, S., Sefton-Green, J., 2016. The Class. Living and Learning in the Digital Age. New York University Press, New York.
- López Aguilar, J.L., Maza-Córdova, J., Tusa, F., 2020. Educar en el contexto digital: el reto de ser edutuber. *Revista Ibérica de Sistemas e Tecnologías de Informação* E25, 188–200.
- McLeod, J., 2009. Youth studies, comparative inquiry, and the local/global problematic. *Rev. Educ. Pedagog. Cult. Stud.* 31 (4), 270–292.
- Meikle, G. (Ed.), 2020. *The Routledge Companion to Media and Activism*. Routledge, New York and London.
- Moghavvemi, S., Sulaiman, A., Jaafar, N.I., Kasem, N., 2018. Social media as a complementary learning tool for teaching and learning: the case of YouTube. *Int. J. Manag. Educ.* 16 (1), 37–42. <https://doi.org/10.1016/j.ijme.2017.12.001>.
- Moisio, S., 2018. *Geopolitics of the Knowledge-Based Economy*. Routledge, London & New York.
- Noble, S., 2018. Algorithms of Oppression. How Search Engines Reinforce Racism. New York University Press, New York.
- O'Neil, C., 2016. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishers, New York.
- Orgeron, D., Orgeron, M., Streible, D. (Eds.), 2012. *Learning With the Lights Off: Educational Film in the United States*. Oxford University Press, New York.
- Papacharissi, Z., 2010. A Private Sphere. Democracy in a Digital Age. Polity Press, Cambridge, UK.
- Perrotta, C., Gulson, K.N., Williamson, B., Wittenberger, K., 2021. Automation, APIs and the distributed labour of platform pedagogies in Google Classroom. *Crit. Stud. Educ.* 62 (1), 97–113. <https://doi.org/10.1080/17508487.2020.1855597>.
- Plantin, J.C., Lagoze, C., Edwards, P.N., Sandvig, C., 2018. Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media Soc.* 20 (1), 293–310.
- Poovey, M., 2017. Why post-factuality is so difficult to fight. *Bildungsgeschichte* 7 (2), 220–223.
- Rancière, J., 1991. The ignorant schoolmaster. In: Ross, K. (Ed.), *Five Lessons in Intellectual Emancipation*. Stanford University Press, Stanford, CA.

- Regueira, U., Alonso-Ferreiro, A., Da-Vila, S., 2020. Women on YouTube: representation and participation through the Web scraping technique. *Comunicar* 28 (63). <https://doi.org/10.3916/C63-2020-03>.
- Rizvi, F., 2012. Mobilities and the transnationalization of youth cultures. In: Lesko, N., Talburt, S. (Eds.), *Keywords in Youth Studies. Tracing Affects, Movements, Knowledges*. Routledge, London, pp. 191–213.
- Scolari, C., Fraticelli, D., 2017. The case of the top Spanish YouTubers: emerging media subjects and discourse practices in the new media ecology. *Convergence* 25 (3), 496–515.
- Selwyn, N., Cumbo, B., Pangrazio, L., 2022. Data classes: an investigation of the people that “do data” in schools. In: Pangrazio, L., Sefton-Green, J. (Eds.), *Learning to Live With Datafication: Educational Case Studies and Initiatives From Around the World*. Routledge, London and New York. <https://doi.org/10.4324/9781003136842-4>.
- Snickars, P., Vonderau, P. (Eds.), 2009. *The YouTube Reader*. National Library of Sweden, Stockholm.
- Srinivasan, R., 2017. *Whose Global Village? Rethinking How Technology Shapes Our World*. New York University Press, New York.
- Smieck, N., 2018. *Capitalismo de plataformas*. Caja Negra Editora, Ciudad Autónoma de Buenos Aires.
- Stiegler, B., 2009. The Carnival of the new screen: from hegemony to isonomy. In: Snickars, P., Vonderau, P. (Eds.), *The YouTube Reader*. National Library of Sweden, Stockholm, pp. 40–59.
- Strangelove, M., 2010. *Watching YouTube: Extraordinary Videos by Ordinary People*. University of Toronto Press, Toronto.
- Strathern, M., 1996. Cutting the network. *J. Roy. Anthropol. Inst.* 2 (2), 517–535.
- van Dijck, J., Poell, T., De Waal, M., 2018. *The Platform Society. Public Values in a Connective World*. Oxford University Press, Oxford, UK.
- van Dijck, J., 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, Oxford, UK.
- Williamson, B., Hogan, A., 2020. Commercialisation and Privatization in/of Education in the Context of Covid-19. *Education International Research*. Available at: https://issuu.com/educationinternational/docs/2020_eiresearch_gr_commercialisation_privatisation?fr=sZDJkYjE1ODA2MTQ.
- Winocur, R., Dussel, I., 2021. Memes production as parodic activism: inclusion and exclusion in young people's digital participation in Latin America. In: Frau, D., Kotilainen, S., Pathak-Shelat, M., Hoechsmann, M., Poyntz, S. (Eds.), *Handbook of Media Education Research*. WileyBlackwell Publishers, London, pp. 33–45.
- Zhou, Q., Lee, C.S., Sin, S.-C.J., Lin, S., Hu, H., Bin Ismail, F.F.M., 2020. Understanding the use of YouTube as a learning resource: a social cognitive perspective. *Aslib J. Inf. Manag.* 72 (3), 339–359. <https://doi.org/10.1108/ajim-10-2019-0290>.
- Zuboff, S., 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs, New York.

Global competition over skills

Sin Yee Koh, Institute of Asian Studies, Universiti Brunei Darussalam, Gadong, Bandar Seri Begawan, Brunei Darussalam

© 2023 Sin Yee Koh. Published by Elsevier Ltd. All rights reserved.

Introduction	631
Key global developments and debates	631
The attraction and retention of skills through selective migration policies	632
Skilled immigration policies	633
Diaspora strategies and return migration policies	634
The production of skills through education and training	634
TVET	634
Student migrants as future skilled migrants	635
Emergent research agendas	635
Whither “skills” and “talent”	635
The migration industry	636
The COVID-19 pandemic	636
Conclusion	637
References	637
Further reading	639
Relevant website	639

Glossary

- Brain circulation** The propagation of skills and knowledge through the migration of skilled individuals
- Brain drain** The loss of human capital through emigration
- Brain gain** The gain in human capital through immigration
- Diaspora strategies** Policies and programs adopted by the country of origin to engage with and mobilize exterritorial citizens and emigrants
- Human capital** The skills and resources that an individual or the collective labor force possesses, that has contributive economic value
- Knowledge economy** An economic mode of production centered on information, skills, and know-how as opposed to capital or labor. It is fueled by innovation, research, and technological advancements
- Labor market test** A government requirement where the employer must show that no resident labor is available for the position, in order to justify the need to recruit skilled immigrant labor
- Migration industry** An assemblage of middlemen actors that are driven by their commercial interests in the facilitation and management of migration
- Skilled migrants** Individual migrants who possess the educational and training qualifications, work experiences, and other conditions of immigration set by the host country
- Skills shortage list or skilled occupation list** A list of skills or skilled occupations (that have been identified by the host country as essential, critical, or desirable), used for the selection of skilled immigrants
- Talents** High performing individuals who possess the skills and abilities that are desired by the employing institution, sector, and/or country

Nomenclature

- ADB** Asian Development Bank
- GTCI** Global Talent Competitive Index
- ILO** International Labor Organization
- IOM** International Organization for Migration
- MNC** Multinational Corporation
- OECD** Organization for Economic Co-operation and Development
- PBS** Points-Based System
- STEM** Science, Technology, Engineering, and Mathematics

TNC Transnational Corporation
 TVET Technical and Vocational Education Training
 TYTS Thousand Youth Talents Scheme

Introduction

Countries around the world are increasingly competing for skilled human capital. This global competition over skills or “global race for talent” (Shachar, 2006; Shachar and Hirschl, 2013) have been pursued through policies and programs in the areas of migration as well as education and training. On the one hand, migration policies and programs target skilled immigrants, skilled diasporas, and returnees. On the other hand, education and training policies and programs target nationals and international students turned skilled migrants. Regardless of the specific target groups, the fundamental aims of such policies and programs are consistently similar: to attract, retain and produce skilled human capital that are necessary to sustain the country’s economic growth and global economic competitiveness.

This article provides a critical review of the rationale, strategies and consequences of the global competition over skills. First, it outlines the key global developments and debates that contextualize countries’ competition for skilled human capital. Second, it reviews policies and programs that focus on the attraction and retention of skilled human capital. This includes skilled migration policies, as well as diaspora strategies and return migration policies. Third, it reviews policies and programs that focus on the production of skilled human capital. This includes technical and vocational education training (TVET) programs, and higher education student immigration policies (including post-study work visas). Fourth, it discusses emergent research agendas, including debates around how “skills” and “talent” are defined, the emergence of the migration industry that facilitates and shapes skills mobility, and the impact of the COVID-19 pandemic on the global competition for skills.

In this article, the term “migration” is used in reference to all types of migration flows, including immigration and emigration. The term “immigration” is used in specific reference to inward migration to host countries, while “emigration” is used in specific reference to outward migration from sending countries. The term “skilled migration policies” encompasses policies that attract and encourage immigration, circular migration, and return migration of skilled individuals. The term “skilled immigration policies” refer specifically to policies concerning the inward mobility of skilled individuals to the host country.

Key global developments and debates

The debates on the migration of skilled human capital started emerging in the mid-1960s following the large scale emigration of British scientists and engineers to Canada and the USA in the 1950s–60s (Jöns and Cranston, 2020). Then, the focus was on brain drain, where the en masse emigration of skilled individuals was seen as the irreversible loss of human capital from the origin to the destination country. During the 1970s–90s, academic debates focused on brain drain from developing countries to developed countries, especially in certain skill intensive professional sectors such as the medicine, nursing, science, and engineering (e.g., Portes, 1976; Joyce and Hunt, 1982). The term brain drain thus became associated with the permanent loss of scientific human capital, with negative economic development effects for the origin countries that are primarily developing countries in the Global South.

It was not until the 1990s that the issue of skilled migration started to be framed as the *global* competition over skills—that is, it is no longer seen as a sole problem of one-directional brain drain from developing countries. This development has been largely a response to broader trends including economic globalization, the development of the knowledge economy, and the rise and expansion of multidirectional international migration flows that are no longer confined to South–North flows. Under the broader context of economic globalization, there has been increased integration of national economies into the global economy. Alongside this, the emergence and expansion of the knowledge economy saw the turn toward intangible skills, knowledge, know-how, and technological innovations as means of production rather than generic (unskilled or less-skilled) labor force and tangible raw materials. The knowledge economy germinated in developed countries, but has rapidly expanded to middle-income developing countries that are seeking to gain competitive footholds in the global economy. This development has, to some extent, encouraged reverse brain drain flows to such developing countries, thereby constituting a loss for the developed countries that were hosting those skilled individuals (Ong et al., 1992).

As part of economic globalization and the rise of the knowledge economy, multinational corporations (MNCs) and transnational corporations (TNCs) are locating their headquarters and other functional offices in cities and countries where they have access to the necessary skilled human resources for their operations. These include locations in the Global South where wages and operating costs are comparatively lower. The geographical spread in the sites of employment that are linked to the global knowledge economy further stimulated multidirectional international migration flows. Such migration flows could be produced and managed internally within the corporation (e.g., through intra-company transfers), or stimulated through external hiring. For example, Koser and Salt (1997: 286) noted that TNCs, through their internal and external labor markets, employ their necessary skilled manpower

and deploy them to any of their sites worldwide. Moreover, skilled manpower that have originally migrated internally within a corporation could become subsequently available to other employers locally and globally (Findlay et al., 1996). Hence, the skilled labor demands of MNCs and TNCs in turn stimulate multidirectional international migration flows of skilled manpower.

The push toward the development of skills for countries' economic competitiveness has also been actively promoted by international and regional organizations such as the Asian Development Bank (ADB), the International Labor Organization (ILO), the Organization for Economic Co-operation and Development (OECD), and the World Bank (e.g., ADB, 2007; ADBI et al., 2014; Dundar et al., 2014; Lall, 1999). In a working paper for the OECD, Salt (1997) argued that a highly selective skilled migration regime that effectively manages inflows and outflows of skilled migrants is essential for a country's economic development. This emphasis on the link between the mobility of skilled migrants and global competitiveness is further promoted in an OECD report published a decade later (OECD, 2008). Nevertheless, a key developmentalist stance underwire much of international and regional organizations' advice regarding the importance of skills for countries' economic development. For example, a common trope is that "[t]he immigration policies of developed countries should facilitate movement; yet, they should incorporate mechanisms that encourage developing country economic growth" (Lowell and Findlay, 2002: 2) such as encouraging return migration through the temporary skilled worker visas to prevent brain drain from the origin developing countries.

In contrast to the debates on the detrimental impacts of brain drain to developing countries prior to the 1990s, after the turn of the century the migration of skilled individuals was no longer seen as a one-off occurrence that entailed an absolute gain or loss on the part of the receiving or sending country. Docquier (2014), for example, argued that some level of skilled emigration can generate positive feedback for sending countries (e.g., through remittances, circular migration, the presence of diasporas in destination countries)—although the extent to which this is possible is largely country specific. Indeed, with the emergence of concepts such as brain circulation and brain exchange, scholars, policymakers and international organizations have argued that skilled individuals can contribute to the development of their origin countries remotely, or through occasional return visits (Agunias and Newland, 2012; Saxenian, 2005). Such a perspective informed the development of diaspora strategies or diaspora engagement policies, where the governments of origin countries introduce policies to actively engage with their emigrant skilled populations for skills development and skills transfers (de Haas, 2006; Kuznetsov, 2006a; Wescott and Brinkerhoff, 2006). While return migration (also conceptualized as reverse brain drain) constitutes one of such diaspora strategies, the fundamental aim is to encourage the circulation of skills, knowledge exchange, and technological transfers rather than corporeal mobility per se.

The turn of the century has also witnessed the development of annual indices and rankings that chart the global economic competitiveness of countries and cities in talent attraction. For example, the Global Talent Competitive Index (GTCI), produced by INSEAD (a renowned global graduate-only business school) since 2013, is an annual benchmarking index of 125 countries (INSEAD, 2020). The GTCI is produced based on an analysis of input factors (talent enablers, attraction, growth, retention) and output factors (vocational and global knowledge skills) in each indexed country. Another annual index is the IMD World Talent Ranking, produced by the Institute for Management Development (IMD) World Competitiveness Center (a research center under a global business school based in Switzerland and Singapore) since 2014 (IMD, 2020). This index measures three main factors: (1) the investment in the development of resident talent; (2) the extent to which a country appeals to skilled migrants; and (3) the availability and readiness of skills in the country's talent pool. The existence of these indices and rankings further reinforce and legitimize the global competition over skills. Although developed countries in the Global North still dominate the top positions in such lists, (sector-specific) talent hubs in developing and periphery countries in Asia and the Middle East (Ewers, 2017; Ewers et al., 2021; Ortega et al., 2020) have started to emerge as "rising stars in the global race for skill" (Cerna and Czaika, 2021).

On the one hand, the global competition for skills has been produced in response to national, regional, and global demands for skilled human capital which may not be readily available locally. On the other hand, the increase in the number and proportion of educated populations has also meant that more individuals are aspiring for work and life opportunities that may not be available in their countries of origin. As a result of both these demand and supply factors, a "global skills market" has emerged (Czaika, 2018: 3–7).

The attraction and retention of skills through selective migration policies

As a result of the broader trends and developments outlined earlier, countries are increasingly competing with each other in attracting and retaining their skilled workforce. The global competition for skills is aptly captured in Kapur and McHale's (2005) book title: "Give us your best and brightest". In the book, the authors suggested that countries enter into this global competition by increasing the focus on skills in their immigration policies. And indeed, since the 2000s increasing numbers of developed and developing countries have designed and implemented selective skilled immigration policies (Czaika, 2018: 1). To be sure, *attracting and retaining* skilled human capital in the country's labor force through immigration is often a more economically efficient manner to boost productivity and generate growth. This is in relative comparison to strategies aimed at *producing* skilled human capital through education and training, which often require investments over a longer timeframe before the outputs can be reaped. Moreover, the investments into education and training, undertaken by the origin country, may eventually lead to the loss of the trained human capital through emigration to the destination country (i.e., brain drain).

The migration based strategies undertaken by countries to attract and retain skilled human capital can be broadly categorized into three areas of focus, each targeting a specific group of skilled migrants. First, with regards to readily available skills, countries typically utilize skilled migration policies targeted at importing sector-specific skilled migrant labor. Second, with regards to the

transfer of knowledge, skills and networks, sending countries could utilize diaspora engagement policies targeted at their emigrant and diaspora populations. Finally, with regards to skilled human capital in-the-making, countries could turn to student immigration policies (including post-study work visas) that target student turned skilled migrants. The last group will be discussed in the next section, on the production of skills through education and training.

Skilled immigration policies

It has been pointed out that there is often ambiguity between “skilled” and “highly-skilled” migration in academic and policy discussions (Parsons et al., 2020; Weinar and Klekowski von Koppenfels, 2020). For example, in a working paper discussing the conceptualization of highly-skilled migration, Parsons et al. (2014) highlighted that there are three discordances that complicate any neat and objective definition of this type of migration. In particular, they highlighted that there is firstly a “definitional discordance” where an individual’s status as a highly-skilled migrant is dependent on the destination country’s definition; secondly an “occupational discordance” as countries may have competing occupational categorizations that may not be directly comparable, thus affecting the individual’s classification as highly-skilled or otherwise; and thirdly a “policy discordance” due to the intrinsic contradictions in the country’s treatment of a specific group across different sets of policies (Parsons et al., 2014: 4–5). In sum, how skilled and highly-skilled migrants are being defined through policy terms are highly dependent on changing national priorities as well as global politico-economic dynamics.

As “skill” may be difficult to define and objectively measured, countries often use indicators as proxies of selection in their skilled immigration policies. Commonly used proxies include an individual’s highest level of educational or technical qualification, occupational classification, or minimum salary (Czaika, 2018; Parsons et al., 2014). With regards to assessing potential migrants’ qualifications and experiences, host countries may look to internationally recognized formal education and training. With regards to occupational classification, host countries may establish and maintain their own national classification systems, such as skills shortage lists, skilled occupation lists, and priority employment sectors. Additionally, some destination countries may require labor market tests to ensure that employment of skilled migrants do not reduce the employability of local residents. The key characteristic of such selection instruments is that they can be tweaked and modified to suit the destination country’s changing labor market needs.

Skilled and highly-skilled immigration policies can be generally categorized into three broad policy categories: (1) skill-selective admission policies; (2) post-entry policy instruments; and (3) bilateral labor agreements (Czaika and Parsons, 2017: 610; Table 1). Each policy category has its own function and rationale. Briefly, the first category filters and selects for skilled human capital for immigration entry, while the second and third categories contribute toward the attractiveness of the host country for skilled migrants in the short, medium and long term. Given the increasing competition for skills, it is perhaps the second and third categories that constitute the determinative factors to the attraction and retention of skilled human capital. Indeed, Papademetriou et al. (2008) have suggested that skilled and highly-skilled migrants consider a country’s “total immigration package” (i.e., the sum total of the country’s immigration regulations and conditions) when making migration decisions. For example, potential skilled migrants

Table 1 Categories of skilled and highly-skilled immigration policies.

<i>Skill-selective admission policies</i>	<i>Post-entry policy instruments</i>	<i>Bilateral labor agreements</i>
Points-based system (PBS): selects potential skilled migrants for entry using a scoring system where points are allocated to a set of factors (e.g., education level, language fluency, existing job offer, connection with the host country, age, occupational sector, etc.). Used in Australia, Canada, the UK, and the USA. Skills shortage list: potential skilled migrants who are qualified to work or train in an occupation in the list are eligible for entry through this immigration route. The host country may maintain different lists (e.g., temporary/long-term list, regional list, employer nominated list). Used in Australia and New Zealand. Labor market test: mechanism that ensures that skilled migrants are only admitted after employers have demonstrated their unsuccessful recruitment of national workers. Used in most of the European Union member states.	Access to permanency rights: possibilities for skilled migrants (and their family members or dependents) to access rights to permanent settlement in the host country. This includes permanent residence, citizenship, spousal work rights, children’s education, family reunification, etc. Financial incentives: special financial arrangements for skilled migrants, such as tax exemptions, tax breaks, preferential tax rates, allowances, etc.	Mutual recognition of qualifications: regulation of the reciprocal recognition of diplomas and credentials by sending and host countries. Portability of social benefits: regulation of the equal or indiscriminated ability for skilled migrants to retain access rights to social benefits such as pension, healthcare, and other accrued benefits when they move across countries. Double taxation: agreements that prevent the double taxation of income, wealth, property, and inheritances for skilled migrants who may hold multiple residences.

may take into consideration the host country's other immigration and settlement possibilities that implicate they and their family members' future livelihoods and mobilities. As [Bailey and Mulder \(2017\)](#) have pointed out, skilled and highly-skilled migrants' migration decisions must also be contextualized to their lifecourse trajectories, in addition to their evaluations of policies and institutions that support their migration plans and professional development.

Diaspora strategies and return migration policies

While skilled and highly-skilled immigration policies are used to target immigrants, diaspora strategies and return migration policies are used to target skilled emigrants and members of the diaspora. This set of policies and programs are typically used by sending countries with significant numbers of skilled emigrants and diasporas, and that may not be as competitive as other host countries in attracting skilled immigrants. This is especially the case for developing, middle-income and upper middle-income countries with skilled emigrants and diasporas in the STEM (Science, Technology, Engineering, and Mathematics) fields. As mentioned earlier, diaspora strategies include those that encourage permanent, temporary or circular returns, as well as those that facilitate offshore knowledge exchange and other contributions.

Like skilled immigration policies, there is "no 'one size fits all' policy [through] which countries attract talent" ([Cerna and Czaika, 2021: 18](#)). It therefore follows that diaspora strategies must be designed to suit to each specific context. [Kuznetsov \(2006b: 233–236\)](#) has suggested that the degree to which sending countries are able to engage with and mobilize their diasporas in the global competition for skills is dependent on the interactions between two factors: (1) country conditions; and (2) the size and sophistication of the skilled diaspora. For example, diaspora skills and knowledge transfers may be hampered by the lack of a robust investment and innovation ecosystem in a sending country, even with the presence of a relatively large, mature, and well organized diaspora with the desire to assist in the country's skills development. Hence, the type of diaspora strategies and return migration policies must be appropriately designed and implemented to suit each country's peculiarities.

East Asian and Middle Eastern industrializing countries such as China, South Korea, and Israel have been most prominent in their use of diaspora strategies and return migration policies targeting skilled emigrant and diaspora populations ([OECD, 2008: 48–49](#); [Welch and Hao, 2016](#); [Zweig, 2006](#)). Under President Park Chung-Hee's leadership, South Korea implemented systemic reverse brain drain policies during the 1960s–90s, which led to significant repatriation of Korean scientists and engineers from their host countries ([Yoon, 1992](#)). Crucially, South Korea's success in this regard was the result of an organized state-led effort that provided a conducive domestic environment (e.g., research and development institutional development, legal and administrative reform) with conditions that empowered the returnee skilled human capital to pursue their research and innovation (e.g., compensation and benefits, guarantees of research autonomy) ([Yoon, 1992](#)). Similarly, China's return migration policies, such as the Thousand Youth Talents Scheme (TYTS) launched in 2011, have been relatively successful as the schemes have been accompanied by supportive domestic environments for professional development, especially since the turn the 21st century ([Li et al., 2018](#); but see [Zweig and Wang, 2013](#)). Another key factor that contributed to the relative success of China's skilled return migration schemes is the existence of coordination between central and local governments in the implementation of the schemes, from the recruitment of talents to the support provided to them after their return to China ([Li et al., 2019](#)).

In contrast, Israel has adopted an approach that primarily utilized ethnonationalism as a discursive strategy to entice the return of its highly skilled diasporas. As [Cohen \(2009\)](#) has argued, this discursive approach effectively excludes members of the diaspora who do not fit into the state's target population (see also [Koh, 2015](#) on the Malaysian case). Moreover, such an approach may overly focus on invoking ethnonationalism, without due consideration of individual skilled diasporas' personal/professional circumstances and other non-ethnonationalistic decision making factors ([Cohen and Kranz, 2015](#)). Most importantly, unlike the case of South Korea and China explained above, this approach has focused almost exclusively on the act of return while neglecting what comes after the skilled diasporas' return.

The production of skills through education and training

The previous section has discussed the strategies used by countries to attract and retain "ready made" (i.e., already trained) skilled human capital. This section shifts the focus to the strategies used by countries to produce skilled human capital through education and training programs and policies. These include TVET programs aimed at developing domestic human capital, and higher education student migration policies (including post-study work visas) aimed at attracting student migrants as future skilled migrants.

TVET

In a briefing paper for the ADB, [Jagannathan and Geronimo \(2013\)](#) highlighted that TVET programs should be integrated into the skills development strategy for developing countries in the Asia-Pacific. As they argued, the coordinated development of technical and transferable skills that serve interrelated industries can increase the economic efficiency and competitiveness of the industry cluster as a whole. Crucially, skills that are needed by these developing economies include advanced skills that are required for the knowledge economy, as well as medium skills that are required for labor-intensive sectors.

Indeed, there are differences in the level of maturity and sophistication in each country's knowledge economy. Some countries may be more advanced in this area, thereby requiring skilled human capital that possess more advanced and specialist skills and

knowledge. Some countries may be less advanced in its knowledge economy, with some components of the national economy being reliant on labor-intensive production. This means that such countries would need to focus on the development of medium level technical and vocational skills, rather than that of advanced knowledge skills. Regardless of the exact composition of the country's economy and industry sectors, what is important is that the global, regional and national markets are rapidly changing. By extension, labor market needs are also constantly changing. To respond to the changing skills demand in the labor markets, countries would need a robust national education system that include public and private education and training provisions, as well as better integration between skills production and skills utilization (see [Froy et al., 2012](#) for an analysis of the experiences of Italy and the UK; and [Yonezawa and Kosugi, 2006](#) for an analysis of Japan's experience).

Student migrants as future skilled migrants

There are close links between higher education and the production of skilled human capital and their post-qualification labor mobilities. A report on labor migration, skills and student mobility in Asia noted that the facilitation of student mobility is often seen as a key element in the broader promotion of skills development and mobility ([ADB et al., 2014: 15](#)). Indeed, many host countries see student migrants as future skilled migrants ([Ziguras and Law, 2006](#)). This assumption of a more or less linear and predictable transition from higher education to skilled labor has informed the design of post-study skills retention policy instruments such as job search visas (also known as post-study work visas), waiver of labor market tests, bonus points for switching into other visa categories, and facilitated access to permanent residency and citizenship (see [ADB et al., 2014: 23–26](#)).

As [Mosneaga \(2014\)](#) has rightly pointed out, students sit at the intersection between skills production and skills attraction. Classified as “students”, they are seen as in the process of gaining skills; and it is only upon the completion of their education or training that they can attain the status of “skilled migrants”. This clear categorical separation, however, obscures the reality that “students are not only students, and skilled workers are not only workers: the roles change during the time they spend in the host country” ([Eskelä, 2013: 150](#)). Moreover, some studies have found that international student migration becomes a channel for entry into temporary and precarious low-wage or low-skilled jobs in the host country (e.g., [Liu-Farrer, 2009](#); [Robertson, 2016](#)), rather than into the skilled labor market.

Despite the contradictory outcome where international students do not turn into the skilled human capital that the host countries expect them to be, many host countries have designed and implemented post-study skills retention policy instruments, especially during the 2000s to early-2010s. And indeed, the existence of such policy instruments may well be the key attractor for potential international student migrants who are planning ahead for their post-study careers and migration plans. [Baas \(2010\)](#), for example, found that Indian student migrants have chosen Australia as their study destination because of the clear-cut rules for permanent residency application after graduation. This has developed to the extent that the India-Australia student migration route became a “PR factory” (i.e., a factory line production of permanent residents) ([Baas, 2010](#)). Here, again, is an example of the contradictory outcome of a post-study skills retention policy that have been appropriated by different actors for their respective purposes.

This mismatch between expectation and outcome of post-study skills retention policies, as well as shifting government stances on immigration, have led to some host countries reforming their post-study work visas such that these no longer provide an automatic pathway to migration and settlement. In 2013, the Australian government introduced the revised temporary graduate visas (subclass 485) which aimed at decoupling international student immigration from skilled immigration ([Spinks, 2016](#); see also [Tran et al., 2020](#)). Similarly, the UK's post-study work visa was abolished in April 2012 amidst a hostile immigration environment (see [Moskal, 2016: 128–129](#)). In October 2019, the British government announced the revised post-study work visa (i.e., the Graduate route visa), primarily aimed at increasing the number of international students in the UK ([UK Home Office, 2021](#)) rather than retaining graduates in the skilled workforce. These examples suggest that host countries such as Australia and the UK are shifting away from post-study skills retention policies.

Emergent research agendas

Whither “skills” and “talent”

Although there has been a proliferation of policies and programs in the global competition for skills over the past decades, there is nevertheless no universal agreement on the definitions of “skills” and “talent”. Many scholars have pointed out that “skills” and “talent” are illusive concepts that are socially constructed, power laden, and context specific ([Boucher, 2020](#); [Skeldon, 2018](#); e.g., [Williams and Baláz, 2008: 26](#)). As a result, attempts to delineate common categories and typologies of “skills” or “talent” are never sufficiently comprehensive nor easily comparable across different country contexts. For example, [Solimano's \(2008\)](#) six-type talents categories (technical talents; scientists, academics and international students; professionals in the health sector; entrepreneurs and managers; professionals in international organizations; and cultural talents) missed other types of talents (e.g., sports talents) that may be valued by a particular host country (see [Shachar and Hirschl, 2013](#)). More recently, [Cerna and Chou \(2019: 821\)](#) noted that “definitions of talent differ by discipline, organization, policy sector, country, as well as over time”. Afterall, “skills” and “talent” can be turned into rhetorical devices and policy instruments by countries to select and reject individuals on the basis of their changing labor and other needs.

Parsons et al. (2020) have pointed out that “skill” is implicit in skilled immigration policies but not explicitly delineated. On the one hand, the ambiguities surrounding definitions of “skills” and “talent” can translate into downstream implications for policy design, implementation, and evaluation (Cerna and Chou, 2019). On the other hand, a lack of transparency and consistency in the definitions of “skills” and “talent” raises important ethical and political questions for the global competition for skills. These questions include: “Who are the arbitrators of skill? What constitutes skill? And how is skill constructed in the migration process and in turn, how does skill affect the mobility?” (Liu-Farrer et al., 2020: 4, original emphasis). Indeed, skilled migration has never been equally fluid and frictionless for every individual. Recent contributions have highlighted that the gender aspects of skilled migration, including biases in skills definition and recognition, remain under-researched (Boucher, 2021; Dodson, 2021). Gender may also interact with other axes of intersectionalities (e.g., age, race/ethnicity, religion, sexuality) in producing diverse and stratified skilled migration outcomes and experiences (Boucher, 2020; Kofman, 2014).

To this end, Raghuram (2021) has called for attention to the spatio-temporalities of skills, that is, the time and place (including sites and networks) where skills are produced and reproduced. As she persuasively argued, attention to the “politics and ethics of skilling” can shed light on the varied mechanisms and outcomes of the spatialities of skills which “prescribe, produce, have prevailed and precluded some people from skilled mobility” (Raghuram, 2021: 5, my emphasis). According to her, such an approach offers three productive possibilities. Firstly, it enables the analysis to get past the limits of methodological nationalism that have informed the debates on brain drain versus brain gain. Secondly, it enables a shift beyond the national frame, by taking into consideration the structural conditions that have shaped the historical developments of skilled migration and policies in and across specific sites. Finally, it allows for an interrogation of collective responsibility to the inequalities that have been produced and intertwined into skilled migration, beyond attributing those inequalities to individuals’ personal attributes and characteristics. In sum, this approach could offer productive insights to the question of distributive justice underlying the global competition for skills.

The migration industry

The intensified global competition for skills has led to the development of a highly diversified and transnational migration industry that caters to countries’ need to attract, retain, and produce skilled human capital. The migration industry can be understood as “the ensemble of entrepreneurs who, motivated by the pursuit of financial gain, provide a variety of services facilitating human mobility across international borders” (Hernández-León, 2008: 154). In human resource management, companies may outsource their talent search to third party companies with the relevant networks to assist the companies in identifying and recruiting the “right” candidate for the position. Similarly, countries may also outsource their talent search to actors in the migration industry.

There may also be cases where the migration industry has developed as a response to market demands and operates independently without any direct collaboration with the host countries. For example, Baas (2019) has noted the emergence of an education-migration industry as a response to the increasing alignment between policies aimed at attracting international student migrants and policies aimed at attracting skilled migrants. As he explained (Baas, 2019: 224):

... the industry ... can be observed to carefully navigate and negotiate a fine line between supply and demand. ... the industry often appears to operate in an in-between space where on the one hand they can be observed to cater for international students’ demands and expectations of onward migration or even permanent residency, while on the other hand paying careful attention to and making a business out of the nation’s rules and regulations as well as policy adjustments and interventions.

At the same time, Cohen (2021) has also observed the development of a return migration industry as a result of Israel’s focus on enticing the return of its skilled diaspora. Specifically, he observed the proliferation of non-state actors that have emerged to capitalize on delivering a suite of services that are needed by the returning skilled migrants and the Israel state. These include “[t]alent acquisition specialists, [international] tax advisors, real estate brokers” (Cohen, 2021: 4101). Crucially, the return migration industry is involved in two stages of the return skilled migration process: firstly, the recruitment or mobilization phase, where the skilled individual is being enticed to return; and secondly, the reintegration phase, where the skilled individual and the accompanying family members are being provided with assistance in their relocation and resettlement into Israel (e.g., children’s education, housing, tax planning).

The COVID-19 pandemic

Since early 2020, the COVID-19 pandemic has generated worldwide disruptions and changes to the normalcy of human socio-economic life. In order to contain the spread of the virus, countries have imposed border lockdowns and travel restrictions, as well as introduced additional health and quarantine requirements for travelers and migrants. These changes have created a new global mobility environment with significant impacts on the ease and possibilities of human movement and skills mobility. The pandemic has also impacted upon labor market demand and supply, especially in terms of skills shortages in certain sectors and professions. Overall, this global health crisis has led to debates about what are “essential jobs”, who are “essential workers”, and what are “essential mobilities” (see Salazar, 2021). These pandemic generated debates would necessarily shape the positions taken by each sending and destination country in the new landscape of the global competition for skills.

With regards to impacts of the COVID-19 pandemic on labor markets, Gamlen (2020) has noted rising unemployment and looming economic depression in the USA and the European Union. As he suggests, these macroeconomic conditions would lead to a reversal of the decades-long trend of migrant labor dependency in key migrant destination countries. The time lag in labor demand and supply may lead to a situation where “migrants cluster in essential jobs at the high- and low-skill ends of the skill spectrum” (Gamlen, 2020: 3), as well as under employment and skills mismatch of both resident and migrant labor. As anti-immigrant sentiments grow, governments worldwide may turn to more restrictive immigration policies.

The restrictions on migration and human movement during the pandemic has meant that skills mobility and knowledge exchange through previously existing routes have been suspended or disrupted. Importantly, while these restrictions may be temporary or short-term in terms of their duration of implementation, it is likely that they will lead to significant and longer-term implications for the production and circulation of skilled human capital. For example, travel restrictions could stifle scientific careers and limit the growth of global scientific networks (Woolston, 2021). Border closures and fully online teaching in higher education institutions may compel potential international student migrants to reconsider their migration and post-study career plans. Indeed, the increasing availability and acceptance of online and distance learning options are expected to shape international student migration flows in the post-COVID world (Mittelmeier et al., 2021). As international students could be future skilled migrants, changes in international student mobility trends will significantly shape the global skills landscape in the mid- and long-term.

Conclusion

Debates surrounding the global competition for skills has evolved from a focus on brain drain versus brain gain, to brain circulation and exchange, and most recently to the race for talents. On the one hand, countries compete with each other to attract and retain what appears to be scarce human capital resource that is in limited supply. This strategy seeks to address skilled labor demand in the short term, especially in specific sectors. On the other hand, countries implement education and training programs to produce skilled labor supply. This strategy seeks to address skilled labor demand in the medium to long term.

While the global competition for skills has been in existence since the mid-twentieth century, there is thus far no universal and objective agreement on the what constitutes “skills” and who is a “talent”. “Skills” and “talent” are socially constructed notions that are context specific, and that are embedded within the spatio-temporalities of their creation and existence. To balance the desires to attract the “best and brightest” (Kapur and McHale, 2005) with any apprehensions toward the influx of foreigners and immigrants, countries implement selective skilled immigration policies that generally attract for entry yet prohibit permanent settlement. Recent trends show that this restrictive immigration stance have also been extended to international student migrants, who were previously viewed as future skilled human capital by their host countries.

Recent research and debates have raised questions about the ethics and politics of “skills” and “talent” in skilled migration policies (e.g., Liu-Farrer et al., 2020). Some scholars have highlighted the need to critically consider how axes of intersectionalities shape different skilled migration outcomes and experiences, as well as how inequalities and inequities have been inscribed into skilled migration policies (e.g., Boucher, 2021; Dodson, 2021; Raghuram, 2021). Another emergent research strand is the emergence of the migration industry that oils the global competition for skills (e.g., Baas, 2019; Cohen, 2021). As the war for talent intensifies, there is a need to understand and evaluate the roles that such a migration industry might play in facilitating as well as shaping skilled migration flows. Finally, the COVID-19 era has ushered in new trends concerning education, migration, and work. We need more systematic and comparative research in order to fully understand the impact of the pandemic for the global competition for skills.

References

- Agunias, D.R., Newland, K., 2012. *Developing a Road Map for Engaging Diasporas in Development: A Handbook for Policymakers and Practitioners in Home and Host Countries*. Migration Policy Institute; International Organization for Migration, Geneva; Washington, D. C.
- Asian Development Bank (ADB), 2007. *Moving Toward Knowledge-Based Economies: Asian Experiences*. Technical Notes No. 091107. Asian Development Bank, Manila.
- Asian Development Bank Institute (ADBI), Organization for Economic Co-operation and Development (OECD), International Labour Organization (ILO), 2014. *Labor Migration, Skills & Student Mobility in Asia*. Asian Development Bank Institute, Tokyo.
- Baas, M., 2010. *Imagined Mobility: Migration and Transnationalism among Indian Students in Australia*. Anthem Press, London; New York.
- Baas, M., 2019. The education-migration industry: international students, migration policy and the question of skills. *Int. Migrat.* 57 (3), 222–234. <https://doi.org/10.1111/imig.12540>.
- Bailey, A., Mulder, C.H., 2017. Highly skilled migration between the Global North and South: gender, life courses and institutions. *J. Ethnic Migrat. Stud.* 43 (16), 2689–2703. <https://doi.org/10.1080/1369183X.2017.1314594>.
- Boeri, T., Becker, S.O., 2012. *Brain Drain and Brain Gain: The Global Competition to Attract High-Skilled Migrants*. Oxford University Press, Oxford.
- Boucher, A.K., 2020. How “skill” definition affects the diversity of skilled immigration policies. *J. Ethnic Migrat. Stud.* 46 (12), 2533–2550. <https://doi.org/10.1080/1369183X.2018.1561063>.
- Boucher, A.K., 2021. Gender bias in skills definition, labor market dynamics and skills recognition. In: Mora, C., Piper, N. (Eds.), *The Palgrave Handbook of Gender and Migration*. Springer International Publishing, Cham, pp. 187–201.
- Cerna, L., Chou, M., 2019. Defining “talent”: insights from management and migration literatures for policy design. *Pol. Stud. J.* 47 (3), 819–848. <https://doi.org/10.1111/psj.12294>.
- Cerna, L., Czaika, M., 2021. Rising stars in the global race for skill? A comparative analysis of Brazil, India, and Malaysia. *Migrat. Stud.* 9 (1), 26–41. <https://doi.org/10.1093/migration/mnaa009>.

- Cohen, N., Kranz, D., 2015. State-assisted highly skilled return programs, national identity and the risk(s) of homecoming: Israel and Germany compared. *J. Ethnic Migrat. Stud.* 41 (5), 795–812. <https://doi.org/10.1080/1369183X.2014.948392>.
- Cohen, N., 2009. Come home, be professional: ethno-nationalism and economic rationalism in Israel's return migration strategy. *Immigr. Minorit.* 27 (1), 1–28.
- Cohen, N., 2021. Israel's return migration industry. *J. Ethnic Migrat. Stud.* 47 (17), 4100–4117. <https://doi.org/10.1080/1369183X.2020.1752638>.
- Czaika, M., Parsons, C.R., 2017. The gravity of high-skilled migration policies. *Demography* 54 (2), 603–630. <https://doi.org/10.1007/s13524-017-0559-1>.
- Czaika, M., 2018. Highly-skilled migration: introduction and synopsis. In: Czaika, M. (Ed.), *High-Skilled Migration: Drivers and Policies*, first ed. Oxford University Press, Oxford, pp. 1–19.
- de Haas, H., 2006. *Engaging Diasporas: How Governments and Development Agencies Can Support Diaspora Involvement in the Development of Origin Countries*. International Migration Institute, University of Oxford, Oxford.
- Docquier, F., 2014. The brain drain from developing countries. *IZA World Labor* 31, 1–10. <https://doi.org/10.15185/izawol.31>.
- Dodson, B., 2021. Gender and gender relations in skilled migration: more than a matter of brains. In: Mora, C., Piper, N. (Eds.), *The Palgrave Handbook of Gender and Migration*. Springer International Publishing, Cham, pp. 203–220.
- Dundar, H., Millot, B., Savchenko, Y., et al., 2014. Building the Skills for Economic Growth and Competitiveness in Sri Lanka. The World Bank, Washington, D.C.
- Eskelä, E., 2013. Migration decisions of skilled migrants: international degree students in an offbeat destination. *Nord. J. Migrat. Res.* 3 (3), 145–152.
- Ewers, M.C., Khattab, N., Babar, Z., et al., 2021. Skilled migration to emerging economies: the global competition for talent beyond the West. *Globalizations* 1–17. <https://doi.org/10.1080/14747731.2021.1882816>.
- Ewers, M.C., 2017. International knowledge mobility and urban development in rapidly globalizing areas: building global hubs for talent in Dubai and Abu Dhabi. *Urban Geogr.* 38 (2), 291–314. <https://doi.org/10.1080/02723638.2016.1139977>.
- Findlay, A.M., Li, F.L.N., Jowett, A.J., et al., 1996. Skilled international migration and the global city: a study of expatriates in Hong Kong. *Trans. Inst. Br. Geogr.* 21 (1), 49–61.
- Froy, F., Giguère, S., Meghnagi, M., 2012. *Skills for Competitiveness: A Synthesis Report*. OECD Local Economic and Employment Development (LEED) Working Papers No. 2012/09. OECD, Paris.
- Gamlen, A., 2020. Migration and Mobility After the 2020 Pandemic: The End of an Age? International Organization for Migration (IOM), Geneva.
- Hernández-León, R., 2008. *Metropolitan Migrants: The Migration of Urban Mexicans to the United States*. University of California Press, Berkeley, London.
- INSEAD, 2020. *The Global Talent Competitiveness Index 2020: Global Talent in the Age of Artificial Intelligence*. INSEAD, Fontainebleau Cedex, France.
- Institute for Management Development (IMD), 2020. *IMD World Talent Ranking 2020*. Institute for Management Development, Lausanne.
- Jagannathan, S., Geronimo, D., 2013. Skills for Competitiveness, Jobs, and Employability in Developing Asia-Pacific. ADB Briefs No. 18. Asian Development Bank, Manila.
- Jöns, H., Cranston, S., 2020. Brain drain. In: Kobayashi, A. (Ed.), *International Encyclopedia of Human Geography*, second ed. Elsevier, pp. 385–389.
- Joyce, R.E., Hunt, C.L., 1982. Philippine nurses and the brain drain. *Soc. Sci. Med.* 16 (12), 1223–1233. [https://doi.org/10.1016/0277-9536\(82\)90147-2](https://doi.org/10.1016/0277-9536(82)90147-2).
- Kapur, D., McHale, J., 2005. *Give Us Your Best and Brightest: The Global Hunt for Talent and Its Impact on the Developing World*. Center for Global Development, Washington, D. C.
- Kofman, E., 2014. Towards a gendered evaluation of (highly) skilled immigration policies in Europe. *Int. Migrat.* 52 (3), 116–128. <https://doi.org/10.1111/imig.12121>.
- Koh, S.Y., 2015. State-led talent return migration program and the doubly neglected “Malaysian diaspora”: whose diaspora, what citizenship, whose development? The doubly neglected “Malaysian diaspora”. *Singapore J. Trop. Geogr.* 36 (2), 183–200. <https://doi.org/10.1111/sjtg.12107>.
- Koser, K., Salt, J., 1997. The geography of highly skilled international migration. *Int. J. Popul. Geogr.* 3 (4), 285–303. [https://doi.org/10.1002/\(sici\)1099-1220\(199712\)3:4<285::aid-ijpg72>3.0.co;2-w](https://doi.org/10.1002/(sici)1099-1220(199712)3:4<285::aid-ijpg72>3.0.co;2-w).
- Kuznetsov, Y. (Ed.), 2006a. *Diaspora Networks and the International Migration of Skills: How Countries Can Draw on Their Talent Abroad*. World Bank Institute, Washington, D. C.
- Kuznetsov, Y., 2006b. Leveraging diasporas of talent: toward a new policy agenda. In: Kuznetsov, Y. (Ed.), *Diaspora Networks and the International Migration of Skills: How Countries Can Draw on Their Talent Abroad*. World Bank Institute, Washington, D. C., pp. 221–237.
- Lall, S., 1999. *Competing With Labor: Skills and Competitiveness in Developing Countries*. Issues in Development Discussion Paper No. 31. International Labour Organization, Geneva.
- Li, M., Yang, R., Wu, J., 2018. Translating transnational capital into professional development: a study of China's Thousand Youth Talents Scheme scholars. *Asia Pac. Educ. Rev.* 19 (2), 229–239. <https://doi.org/10.1007/s12564-018-9533-x>.
- Li, W., Bakshi, K., Tan, Y., et al., 2019. Policies for recruiting talented professionals from the diaspora: India and China compared. *Int. Migrat.* 57 (3), 373–391. <https://doi.org/10.1111/imig.12456>.
- Liu-Farrer, G., Yeoh, B.S.A., Baas, M., 2020. Social construction of skill: an analytical approach toward the question of skill in cross-border labor mobilities. *J. Ethnic Migrat. Stud.* 1–15. <https://doi.org/10.1080/1369183X.2020.1731983>.
- Liu-Farrer, G., 2009. Educationally channeled international labor mobility: contemporary student migration from China to Japan. *Int. Migrat. Rev.* 43 (1), 178–204. <https://doi.org/10.1111/j.1747-7379.2008.01152.x>.
- Lowell, B.L., Findlay, A., 2002. *Migration of Highly Skilled Persons From Developing Countries: Impact and Policy—Synthesis Report*. International Migration Papers No. 44. International Labour Office, Geneva.
- Mittelmeier, J., Gunter, A., Raghuram, P., Rienties, B., 2021. Migration intentions of international distance education students studying from a South African institution: unpacking potential brain drain. *Glob. Soc. Educ.* <https://doi.org/10.1080/14767724.2021.1947202>.
- Moskal, M., 2016. International students pathways between open and closed borders: towards a multi-scalar approach to educational mobility and labor market outcomes. *Int. Migrat.* 55 (2), 126–138. <https://doi.org/10.1111/imig.12301>.
- Mosneaga, A., 2014. Student migration at the global trijunction of higher education, competition for talent and migration management. In: Tejada, G., Bhattacharya, U., Khadria, B., et al. (Eds.), *Indian Skilled Migration and Development: To Europe and Back*. Springer India, New Delhi, pp. 87–111.
- OECD, 2008. *The Global Competition for Talent: Mobility of the Highly Skilled*. OECD Publishing, Paris.
- Ong, P.M., Lucie, C., Evans, L., 1992. Migration of highly educated Asians and global dynamics. *Asian Pac. Migrat. J.* 1 (3–4), 543–567. <https://doi.org/10.1177/011719689200100307>.
- Ortiga, Y.Y., Chou, M.-H., Wang, J., 2020. Competing for academic labor: research and recruitment outside the academic center. *Minerva* 58 (4), 607–624. <https://doi.org/10.1007/s11024-020-09412-7>.
- Papademetriou, D.G., Somerville, W., Tanaka, H., 2008. *Talent in the 21st Century Economy*. Migration Policy Institute, Washington, D. C.
- Parsons, C.R., Rojon, S., Samanani, F., et al., 2014. *Conceptualising International High-Skilled Migration*. IMI Working Papers Series No. 104. International Migration Institute (IMI), Oxford.
- Parsons, C.R., Rojon, S., Rose, L., et al., 2020. High skilled migration through the lens of policy. *Migrat. Stud.* 8 (3), 279–306. <https://doi.org/10.1093/migration/mny037>.
- Portes, A., 1976. Determinants of the brain drain. *Int. Migrat. Rev.* 10 (4), 489. <https://doi.org/10.2307/2545081>.
- Raghuram, P., 2021. Interjecting the geographies of skills into international skilled migration research: political economy and ethics for a renewed research agenda. *Popul. Space Place* 27 (5), e2463. <https://doi.org/10.1002/psp.2463>.
- Robertson, S., 2016. Intertwined mobilities of education, tourism and labor: the consequences of 417 and 485 visas in Australia. In: Dickie, M., Gozdecka, D., Reich, S. (Eds.), *Unintended Consequences: The Impact of Migration Law and Policy*. ANU Press, Canberra, pp. 53–80.
- Salazar, N.B., 2021. Existential vs. essential mobilities: insights from before, during and after a crisis. *Mobilities* 16 (1), 20–34. <https://doi.org/10.1080/17450101.2020.1866320>.
- Salt, J., 1997. *International Movements of the Highly Skilled*. OECD Social, Employment and Migration Working Papers No. 3. OECD, Paris.
- Saxenian, A., 2005. From brain drain to brain circulation: transnational communities and regional upgrading in India and China. *Stud. Comp. Int. Dev.* 40 (2), 35–61. <https://doi.org/10.1007/bf02686293>.

- Shachar, A., Hirschl, R., 2013. Recruiting super talent: the new world of selective migration regimes. *Indiana J. Global Leg. Stud.* 20 (1), 71–107.
- Shachar, A., 2006. The race for talent: highly skilled migrants and competitive immigration strategies. *N. Y. Univ. Law Rev.* 81, 148–206.
- Skeldon, R., 2018. High-skilled migration and the limits of migration policies. In: Czaika, M. (Ed.), *High-Skilled Migration: Drivers and Policies*, first ed. Oxford University Press, Oxford, pp. 48–64.
- Solimano, A., 2008. The international mobility of talent and economic development: an overview of selected issue. In: Solimano, A. (Ed.), *The International Mobility of Talent: Types, Causes, and Development Impact*. Oxford University Press, Oxford, pp. 21–43.
- Spinks, H., 2016. Overseas Students: Immigration Policy Changes 1997–2015. Research Paper Series 2015–16. Parliamentary Library, Canberra.
- Tran, L.T., Rahimi, M., Tan, G., Dang, X.T., Le, N., 2020. Post-study work for international graduates in Australia: opportunity to enhance employability, get a return on investment or secure migration? *Glob. Soc. Educ.* 18 (5), 495–510. <https://doi.org/10.1080/14767724.2020.1789449>.
- UK Home Office, 2021. Graduate Route to Open to International Students on 1 July 2021, 4 March. <https://www.gov.uk/government/news/graduate-route-to-open-to-international-students-on-1-july-2021>.
- Weinar, A., Klekowski von Koppenfels, A., 2020. *Highly-Skilled Migration: Between Settlement and Mobility*. Springer International Publishing.
- Welch, A., Hao, J., 2016. Global argonauts: returnees and diaspora as sources of innovation in China and Israel. *Glob. Soc. Educ.* 14 (2), 272–297. <https://doi.org/10.1080/14767724.2015.1026249>.
- Wescott, C.G., Brinkerhoff, J.M. (Eds.), 2006. *Converting Migration Drains Into Gains: Harnessing the Resources of Overseas Professionals*. Asian Development Bank, Manila.
- Williams, A.M., Baláz, V., 2008. *International Migration and Knowledge*. Routledge, Abingdon.
- Woolston, C., 2021. How COVID-19 has stomped on scientists' travel plans. *Nature* 593, 613–615. <https://doi.org/10.1038/d41586-021-01389-x>.
- Yonezawa, A., Kosugi, R., 2006. Education, training, and human resources: meeting skill requirements. In: Shibata, T. (Ed.), *Japan, Moving Toward a More Advanced Knowledge Economy: Volume I: Assessment and Lessons*. The World Bank, Washington, D. C., pp. 105–126.
- Yoon, B.-S.L., 1992. Reverse brain drain in South Korea: state-led model. *Stud. Comp. Int. Dev.* 27 (1), 4–26. <https://doi.org/10.1007/BF02687102>.
- Ziguras, C., Law, S.-F., 2006. Recruiting international students as skilled migrants: the global “skills race” as viewed from Australia and Malaysia. *Glob. Soc. Educ.* 4 (1), 59–76.
- Zweig, D., Wang, H., 2013. Can China bring back the best? The Communist Party organizes China's search for talent. *China Q.* 215, 590–615. <https://doi.org/10.1017/S0305741013000751>.
- Zweig, D., 2006. Competing for talent: China's strategies to reverse the brain drain. *Int. Lab. Rev.* 145, 65–90. <https://doi.org/10.1111/j.1564-913X.2006.tb00010.x>.

Further reading

- Cohen, N., 2017. Diaspora strategies: actors, members, and spaces. *Geogr. Compass* 11 (e12308), 1–13. <https://doi.org/10.1111/gec3.12308>.
- Gureyev, V.N., et al., 2020. Review and analysis of publications on scientific mobility: assessment of influence, motivation, and trends. *Scientometrics* 124 (2), 1599–1630. <https://doi.org/10.1007/s11192-020-03515-4>.

Relevant website

2020 Global Talent Competitive Index, <https://gtcistudy.com/>.

Public diplomacy and international higher education¹

Eva Hartmann, Faculty of Education, University of Cambridge, Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Educating the elites	640
Diplomacy and the liberal world order	641
The Cold War and public diplomacy	641
The changing higher education landscape	642
The power of educational diplomacy	643
A sign of power or vulnerability?	644
Soft power as a category of practice	644
Toward critical geopolitics	644
References	645

Glossary

International higher education, geopolitics, cross-border mobility of students, liberal world order, China as an emerging power, decoloniality, soft power, International Relations.

Governments increasingly see international students as “a major contributor to (...) soft power” (House of Lords, 2014: 98), a key feature of the country’s “rayonnement international” (Perraton, 2014), as their “lifelong friends (...) in both spirit and deed” (DAAD, 2018: 3). In line with this political interest, scholars of international higher education (IHE) have started to pay more attention to the link between IHE and foreign policies. This contribution provides an overview of this emerging field of study and explores the different angles scholars take to examine the diplomatic dimension of IHE. I am particularly interested in how scholars account for power relations, and will show how the field would benefit from taking better account of the nature of global politics and the world order that has emerged since the Second World War. A more nuanced notion of power is pivotal for coming to grips with recent transformations of the global order. These changes put into question some of the conclusions drawn by scholars of educational diplomacy. Against the backdrop of my critique, I will outline an avenue the field could take with a view to developing a critical perspective that includes insights provided by decolonial studies.

Educating the elites

The emerging literature on educational diplomacy is characterized by diversity, with scholars exploring from different angles the link between international higher education and international relations, understood in a broad sense (Knight, 2019). Some studies relate educational diplomacy to transnational learning and highlight the importance of intercultural competences (Bayat, 2014). They focus on dialogue and the capacity to engage diverse participants in a social process (e.g., Özel, 2021; Bizzell et al., 2018). Other scholars foreground power in their analysis, some of them drawing on International Relations to highlight the link between global power and the role of education in socializing elites in countries of the South (Ikenberry and Kupan, 1990). At the center of these studies is the impressive number of heads of state who were educated in countries that are most influential in global politics. In 2018, 58 out of the 352 serving heads of state or government had been educated in the United States, 57 in the United Kingdom and 40 in France (Hillman, 2018).

The socialization of elites in other countries has a long tradition. It goes back to the British and French Empire that allowed the elites in their colonies to enroll in their universities (Pervillé, 1984). Interestingly enough this educational strategy was not applied to all colonies, as John Ikenberry and Charles Kupan show in their studies of the British Empire (Ikenberry and Kupan, 1990). They point out how the strategy was predominantly used in India and informed by an assimilation strategy. In a famous quotation Thomas Babington Macaulay, a British historian and Whig politician of the 19th century, underlined the importance of creating “a class of persons Indian in blood and color, but English in tastes, in opinions, in morals and in intellect. To that class we may leave it to refine the vernacular dialects of the country, to enrich those dialects with terms of science borrowed from the Western nomenclature, and to render them by degrees fit vehicles for conveying knowledge to the great mass of the population” (Macaulay, 1835 (1957)).

¹I would like to thank Fazal Rizvi, Stefano Guzzini and an anonymous reviewer for their comments on a previous draft. The intensification of our work has put peer review under enormous pressure. I am therefore especially grateful to them for taking the time to provide very substantial feedback that helped me to further develop my thoughts. The usual disclaimer applies.

This quotation also highlights the key role of the “colonial subjects” (Bhabha, 2004) in translating Western knowledge and its nomenclature into the different languages of the colonized. The one-directionality of the translation mirrors the power that helped establish English as a lingua franca (Ives, 2009). However, the act of translation also gives the translator some leeway. It requires important decisions since many terms are untranslatable, as Barbara Cassin’s monumental *Dictionary of Untranslatables* well documents (Cassin et al., 2014, see, also Syrotinski, 2019). Translators needed to be creative to overcome the incommensurability and to ensure that the translation made sense “in the vernacular dialects of the country”. This brings the agency of the colonial subjects to the fore and highlights the colonisers’ dependency on them to ensure their claim to power. The freedom of interpretation has increased with the independence of the former colonies. At the same time the education of the elites in the South by the dominant countries has gained further momentum, as the numbers already mentioned indicate. Scholars who reduce educational diplomacy to cross-border mobility of students are not able to account for this complex interaction between power and freedom that is at the center of the post-war global order, as I will show in the next section.

Diplomacy and the liberal world order

The term diplomacy, as we understand it today, is closely linked to an international liberal order that emerged in the 18th century in the aftermath of the Peace of Westphalia ending the Thirty Years’ War in 1648 (Milton, 2020). The end of the Cold War turned the liberal order, which used to be only hegemonic in the West, into a global one (Ikenberry, 2012). Taking a critical stance, Inderjeet Parmar suggests understanding this order as “imperialism by another name” (Parmar, 2018). However, putting the liberal order on an equal footing with imperialism without further qualification risks underestimating the complex relationship between freedom and dominance that is at the heart of the liberal world order.

International law and the rules of Westphalian system require all members of the society of states to renounce some of their liberty in return for a global order that gives them security and peace. Examples would be the prohibition of some types of weapons, or global trade regulations established by the World Trade Organization or regional institutions, both of which have major implications for the domestic level. At the same time the society of states acknowledges the exclusive sovereignty of individual states over their territory. This principle of sovereignty is still at the heart of today’s post-war order and enshrined in the United Nations Charter, which states that “nothing ... shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state” (UN, 1945: 2(7)).

With increasing global dependence, the definition of what should be considered within a domestic jurisdiction has become increasingly difficult to determine. Diplomacy, which previously meant the prince’s authorization to travel, has become a key means to mediate the tension between global law and order and the sovereignty of the individual states. It is an “established method of influencing the decisions and behavior of foreign governments and peoples through dialogue, negotiation, and other measures short of war or violence” (Marks and Freeman, 1998).

In a liberal vein, some studies of educational diplomacy underline the importance of IHE in enhancing dialogue and negotiations with a view to increasing compliance with international norms and standards to enhance world peace and prosperity. They express the hope that university-to-student dialogue will help build trust, enhance knowledge transfer and reach development goals (Hone, 2014). This dialogue shall create “lifelong friends (...) in both spirit and deed” (DAAD, 2018: 3), as the German Academic Exchange Service put it in one of its documents. Conversely, a realist perspective foregrounds the great inequality between the different members of the society of states, with the United States claiming leadership since the rise of the Pax Americana after World War II.² However, the particularity of the liberal order, with sovereignty as one of its key pillars, makes it impossible for the leading country to just impose its will on other countries, notwithstanding its significant military and economic power. Ikenberry identifies key conditions that the United States need to fulfill to keep the global leadership they gained after World War II. They need to provide – or at least to be seen as providing – public goods in the area of security, and if possible in other areas (Ikenberry, 2012: 281). The second variable is “the ability of the United States to actually agree to reductions in its rights and privileges, sharing rule and authority” (282). This self-restraint on the part of the most powerful, or at least the perception of it, has become a vital incentive for other governments to accept the leadership of the liberal Leviathan, the US. The third variable that Ikenberry identifies is the willingness of other major states, including rising powers such as China and India, to operate within the liberal international order. They may seek to increase their stakes. As long as they accept the principles and values of the order, the global leadership of the US will continue. It is this third dimension of power that is at the center of educational diplomacy studies.

The Cold War and public diplomacy

The role of international higher education gained momentum with the broadening of diplomatic relations in the 1960s. Edmund Gullion, a US career diplomat, popularized the term public diplomacy for this transformation in 1965 (see Cull, 2020). This type of diplomacy no longer builds on exclusivity and secrecy but rather seeks to influence public opinion in other countries. Besides

²For a good overview on different notions of power that the different schools of International Relations have developed, see Guzzini, 2021. Power in World Politics, DIIS WORKING PAPER 2021: 17. Copenhagen: Danish Institute for International Studies..

education and cultural exchange, international broadcasting, and most recently social media have also become key channels in this context (Kukkola, 2018).

The rise of public diplomacy was closely related to the Cold War. Both East and West established international student and academic exchange and scholarship programmes to attract international students (see, e.g. Tsvetkova, 2008). In the 1970s, cultural exchange and cross-border mobility of academics and students came to be seen as a vital instrument to ease tension (*détente*) (Kremenyuk, 1995). The Helsinki Process played a vital role in this context, led by the Conference for Security and Cooperation in Europe (CSCE), with a view to facilitating the dialogue between East and West. The process culminated in 1975 in the Helsinki Final Act, signed by the Soviet Union, all European States and the United States. The act comprised a call for “a more exact assessment of the problems of comparison and equivalence of academic degrees and diplomas” (OSCE, 1975: 53). It was the United Nations Educational, Scientific and Cultural Organization (UNESCO) that was assigned the task of providing the legal instrument for this recognition regime geared to facilitating the exchange of students, teachers and scholars of all participating countries (Hartmann, 2015). The hope was to turn all of them into quasi-diplomats helping to mitigate the tension between East and West. The involvement of the US in this process was the main reason for including this country in the UNESCO definition of the European Region in 1976 (Daniel, 2003). Only a few years later, in 1978, the UNESCO Convention on the Recognition of Studies, Diplomas and Degrees in Higher Education was adopted (UNESCO, 1979). This historical legacy of the conventions is the reason for the geopolitical scope of the revised UNESCO convention, the Convention on the Recognition of Qualifications concerning Higher Education in the European Region, which was adopted in 1997 and became the main legal framework for the Bologna Process aimed at establishing a European Higher Education Area (Hartmann, 2010). The United States has not ratified the convention so far but is one of its signatory countries as they were also heavily involved in the drafting of the revised convention in the 1990s. The legacy of the Cold War on today’s geopolitical arrangements of cross-border mobility arrangements has received surprisingly little attention so far in IHE studies and is certainly worth further exploration.³ This study would also need to pay attention to scholarship programmes that have been in place for a long time, like the ones of the Ford Foundation for instance. It would be interesting to scrutinize whether the way how these programmes were informed by the geopolitical order of the Cold War still influences their set-up today.

The changing higher education landscape

The public diplomacy of IHE gained further momentum with a major increase in international students. Their number has risen from one million in the mid-1970s to over five million today (OECD, 2020). The liberal Leviathan, the US, has become the main destination in the post-war order and attracted 1,075,496 international students in the pre-pandemic year 2019/20 (Redden, 2020, for a historical account, see Bevis and Lucas, 2007). However, the number of new international students enrolling at US colleges and universities has been in decline in recent years.⁴ Australia has clearly benefited from this decline and managed to steadily increase its intake of international students in the pre-pandemic period, reaching 956,773 in 2019/20, with an average annual enrolments growth rate of 10.3% per year over the preceding five years (Australian Government, 2020). The rise of Australia as a major destination country in turn pushed the UK down to third place, with 538,615 international students in 2019/20 of which 73% were from non-EU countries.⁵ China has almost caught up with the UK in recent years, attracting 492,185 international students (see, eg. Marginson, 2018; Altbach, 2012; Hong, 2014; Wu, 2019; Rana, 2016). The rise of China indicates an important change in the geography of attraction. It mainly recruits students from Asia (60%) and much less from Africa (17%) and Europe (15%). France is close on China’s heels with 358,000 international students. Germany has become another key destination with 319,000 in 2019/20, a 6% rise in comparison to the previous year (DAAD, 2021: 60). But smaller emerging economies have also entered the global market. Malaysia, for instance, has undertaken major efforts to attract international students, in particular from Muslim-majority countries like the Gulf Region, in direct competition with Iran, a country that also seeks to recruit students from Malaysia (Banikamal and Ra’ees, 2018). We can thus identify some important geopolitical transformations in terms of attractiveness and an intensified competition between the countries (McGill Peterson, 2014). Prominent sending countries like China have also become important receiving countries (eg. Lo and Pan, 2021; Ham and Tolentino, 2018; Jain, 2020).

The diversification of attractiveness coincides with a broader geopolitical transformation of the global order with the rise of China as an economic and political power, followed by other emerging economies claiming more global weight, and the United Kingdom reimagining its global role after BREXIT (De Ville and Siles-Brügge, 2019; Kettle, 2020; Yu, 2019). But how exactly are global politics and international higher education interrelated? Is the increased competition between student destinations a reflection of broader political changes or does it trigger the transformation of the geopolitical order itself? The majority of contributions

³I am aware that geopolitics is a contested term, coined by Rudolf Kjellén and laying the foundations for the German *Geopolitik* (see, DODDS, K., KUUS, M. & SHARP, J. (eds.) 2013. *Introduction: Geopolitics and its Critics from: The Ashgate Research Companion to Critical Geopolitics*, London: Routledge.). It had been avoided therefore by scholars of International Studies for good reasons, to have a comeback only recently. My understanding of geopolitics draws on the perspective developed by Richard Ashley and his critical social theory of international politics. This perspective seeks to further develop a realist study of global power to overcome its narrow notion of power and to examine the creation of geographical and social space as well as borders (see, ASHLEY, R. 1987. *The Geopolitics of Geopolitical Space: Toward a Critical Social Theory of International Politics*. *Alternatives*, XII, 403–434).

⁴See also <https://www.iie.org/Research-and-Insights/Open-Doors/Fall-International-Enrollments-Snapshot-Reports> [last accessed 1/11/2021].

⁵<https://www.universitiesuk.ac.uk/universities-uk-international/explore-uuki/international-student-recruitment/international-student-recruitment-data> [last accessed 1/11/2021].

to education diplomacy studies fail to address this pivotal question. They tend to assume that international students enhance the global power of the host country, rather than empirically examine the link. In the next section I will turn to contributions that scrutinize this link empirically and will discuss the insights they provide, as well as some of the shortcomings of this research.

The power of educational diplomacy

Educational diplomacy studies draw heavily on Joseph Nye's notion of soft power. Nye popularized the term to explore US influence in a post-Cold War world order characterized by a shift toward public diplomacy. Privileging a behavioral notion of power, he defines soft power as "the ability to affect others to get the outcomes one wants" (Nye, 2021: 197). At the center is a power that Max Weber describes as authority, or "the probability that a command with a given specific content will be obeyed by a given group of persons" (Weber, 1978: 52). This power contrasts with power as coercion, or "the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance" (Weber, 1978: 53). There might be good reasons why people obey a rule. James Scott distinguishes between a thick and a thin notion of acquiescence (Scott, 1990: 72). In the first case people obey because they actively believe in the values underpinning the command and hence their subordination. They may do it because they have internalized the norms, or think they are in their own best interest. The thin notion of acquiescence comes into play where people comply out of resignation.

Transposing these ideas of power to the international level raises important questions given the nature of the international order and the leeway national states have, backed by the principle of sovereignty. Why should nation states deem norms and values established by the liberal Leviathan, the USA, and increasingly the EU, as so authoritative that they adhere to them even if they may be not in their interest? Why should they, as Amelia Hadfield puts it in her comparison of the soft power of the EU and the Commonwealth, "desire similar goals and outcomes to the initiating actor" (Hadfield, 2015:3)? Drawing on Robert Dahl's concept of power, Nye links soft power to the charm and charisma of A, which can convince B to do what A wishes (Dahl, 1957: 206). Accordingly, soft power is "the ability to get what you want through attraction rather than coercion or payment. It arises from the attractiveness of a country's culture, political ideals and policies" (Nye, 2004: 204). Global leadership thus depends not only on military and economic capacity but also on being seen as a model for the rest of the world. Nye refers to the combination of the two types of power as smart power. To convince other societies of the attractiveness of values and norms requires more than propaganda or public relations, he argues. "Simple propaganda often lacks credibility and thus is counterproductive as public diplomacy. (...) Conveying information and selling a positive image is part of it, but public diplomacy also involves building long-term relationships that create an enabling environment for government policies." (Nye, 2008: 101) This is a very important point, as it rejects a narrow notion of power that pays little attention to its relational nature. Studies that link the branding activities of a countries to educational diplomacy are at particular risk of conflating intention and effect (e.g. Raev and Minkman, 2020). There is no guarantee that the branding of a region or a higher education landscape will have the desired effects. It depends on how it is perceived at the other end.

More promising approaches explore whether study abroad increases students' appreciation of the host country's values and its claim to be a model for the world. Surveys and qualitative interviews have been widely used to determine such changes in students' minds (El-Nawawy, 2006). Many studies merely focus on the change in students' attitudes. However, a positive appraisal of the host country expressed by students on their return may simply reflect the great time the students had abroad, as Ane Bislev rightly points out (Bislev, 2017). It does not necessarily lead to more fundamental changes in terms of values. Other surveys seek to identify more fundamental value changes. Jian Guan and his research team examined changes in value between Chinese students who stayed at home and compared them to those who went abroad by using the Chinese Value Survey (Guan and Dodder, 2001). Their study shows rather mixed results. Some returning students had actually become more committed to their own culture instead of embracing US values and norms. Carol Atkinson identifies three factors that need to be in place in order to ensure that students embrace the host country's values and institutions. Firstly, they need to interact extensively with the host society. Secondly the international and domestic students and institutions need to develop a sense of community. Little interaction or negative experiences like racism can cause negative feelings toward the former host country. These two dimensions points out how the government's foreign policy depends on support from within its society to ensure the effectiveness of its strategy. This perspective brings the heterogeneity of actors involved to the fore. However, it continues underestimating the diversity of interests and issues at stake. Most studies assume that the host countries have one coherent IHE strategy in place. They fail to explore the extent to which ministries may differ in terms of the type of students they like to attract and, more generally, how closely they like to tie IHE to foreign policy goals. There might also be important differences between the strategy the government seeks to pursue in comparison to the one of the university associations. The third factor that determines the effectiveness of norm implementation according to Atkinson is the likelihood that the students obtain an influential position upon return. In other terms, the employability of international graduates back home influences the extent to which the values and norms of the former host country gain soft power (Atkinson, 2010: 2). This third factors cast lights on another kind of dependence of the host country; the economic development of the sending country. The employability of international graduates has indeed become a major concern in particular in the UK in recent years (e.g., HEPI, 2021; UUKI, 2020). It seems that the British government no longer trust that its culture, political ideals and policies are enough to further increase the number of international students in line with its 2019 International Education Strategy (HM Government 2019). But such increase does not come without risks.

A sign of power or vulnerability?

In 2019, a report published by an Australian think tank warned of the financial risks that an over-reliance on Chinese students entailed, only a few months before the Covid-19 pandemic forced the government to close the border (Babones, 2019). The worst-case scenario had become reality, triggering a major contraction of Australian higher education (Terry, 2021). Similar concerns about overreliance have been raised in other important destination countries since then. Many undergraduate courses in the UK are no longer financially viable without the revenue generated by international students, as the provost of University College London, Michael Spence, warned in summer 2021 (see Adams, 2021). In the United States, doctoral studies in particular have become dependent on the revenues generated by international students (Cantwell, 2015). In other words, a high number of international students may not necessarily be an expression of the country's global power (Rizvi, 2019). It may rather be an indicator of an important shift in economic power with major consequences for political leadership claims. The major branding activities may actually be a sign of these countries' awareness of their fragility rather than an expression of their power, in contrast to what some scholars have suggested (Lomer et al., 2018). The reason why educational diplomacy studies tend to ignore such signs of vulnerability is closely related to Nye's notion of soft power, as I hope to show in the next section.

Soft power as a category of practice

Todd Hall makes a strong case for soft power being more a category of practice than an analytical category (Hall, 2010). It has become a term that helps politicians to frame their international goals. The British House of Lords, for instance, set up a Select Committee on Soft Power and the UK's Influence in 2013 to discuss how the government could increase the country's soft power (House of Lords, 2014). In a recent reflection on the term, Nye expressed pride in the fact that his notion of smart power had been used by Hillary Clinton while she was Secretary of State (Nye, 2021: 204). A political consultancy and public relations agency, Portland, set up in 2001 by a former adviser to the former British Prime Minister, Tony Blair, has further popularized the notion by publishing regular rankings of countries' soft power.⁶ A number of other rankings of soft power have emerged since then (e.g., Brand Finance). In most of these rankings the Anglo-American countries score highly. Educational diplomacy studies that use soft power in an uncritical way risk becoming part of this propaganda strategy. At best they provide a justification for increasing the intake of international students, but at the price of losing their analytical power that is able to do justice to the complexity of educational diplomacy, the current transformation and the fact that it is always a two-way process. Only a few scholars have paid attention to how overseas students could also increase the soft power of the sending countries (see, e.g., Bislev, 2017: 88).

Toward critical geopolitics

I would like to conclude by identifying some avenues for future educational diplomacy studies that reduce the risk of reifying existing power relations. A first step would be to go beyond a state-centric perspective and to develop a more sophisticated understanding of power that not only enlarges the range of actors it pays attention to (e.g., Hartmann, 2021). It also seeks to highlight the complex relationship between intentions and effects of actions. Following Steven Lukes, we could understand power as a capacity: "(O)bserving the exercise of power can give evidence of its possession and counting power resources can be a clue to its distribution, but power is a capacity, and neither the exercise nor the vehicle of that capacity." (Lukes, 2005: 479) Some capacities may ensure that activities have the intended effect. However, actions can also trigger chains of unintended consequences. Not accounting for unintended consequences risks overestimating the capacity of a powerful actor to control a complex world. A case in point would be the fact that overseas students may also increase the soft power of their own countries.

We can further widen our understanding of the capacity by drawing on Susan Strange's notion of structural power that seeks to complement a behavioral notion of power (Strange, 1988: 29). Strange reminds us not just to explore who has the power, but also to examine its enabling conditions. Her notion of structural power highlights four closely interrelated dimensions, all of which have a close link to higher education and science in a broad sense. The first dimension is the ability to protect others from threats, or rather the perception of this ability. Ikenberry refers to this ability as public goods that the Leviathan must provide to gain leadership in a liberal world order (Ikenberry, 2012: 281). Strange's second structural power is the ability to produce. This dimension addresses issues of competition and the means of production and their ownership. A more critical perspective, drawing on Marx, would foreground the ability to make profits and exploit unpaid labour and nature more generally (Harvey, 2004). The third dimension that Strange identifies is linked to finance and, more precisely, the control of investments and credit. The fourth structural power is the capacity to develop, acquire and deny the access of others "to a kind of knowledge respected and sought by others; and whoever can control the channels by which it is communicated to those given access to it, will exercise a very special kind of structural power" (Strange, 1988: 31).

⁶<https://softpower30.com> [last accessed 1/11/2021].

Higher education and research play a vital role in this context. They create and disseminate knowledge that is relevant for all four dimensions. Drawing on Antonio Gramsci's theory of hegemony and Nicos Poulantzas' theory of the capitalist state we get, in addition, a sense of their mediation role (Hartmann, 2015, Gramsci, 1999 (1971): 217, Thomas, 2007). This perspective understands intellectuals as part of a hierarchical system that represents and mediates between different interests that are loosely associated with different social classes. The mediation by way of this system of intellectuals ensures that some of the allies' interests are included in the economic and political strategies of the ruling power, which in turn provides the strategies with the legitimation and soft power they need to be effective. Transposing this idea to the international level raises interesting questions as to how a global system of intellectuals might look like, the type of different interests they represent, the hierarchy between them, and the mediation they engage with. Drawing on critical geography we can understand the system of intellectuals as a specific social space (Ashley, 1987). It is this mediation role that I suggest putting at the center of educational diplomacy. The role of science and research and not only education suggest linking educational diplomacy to scientific diplomacy, a term that has gained in popularity in recent years (Turekian, 2018). Mediation takes place in all four dimensions that Strange identifies. Academics play a vital role in mediating knowledge that in turn supports different geopolitical claims of military and technological superiority. Although some of these mediations, notably where they involved scientific collaboration between Western and Chinese universities, have come in for criticism. A recent example is the US-Chinese scientific collaboration that caused a public outcry as a link to Chinese military manufacturers was suspected (Tylecote and Clark, 2021). But intellectuals also mediate significantly between theoretical concepts informing the organisation of the economy, investment, and credits, hence Strange's second and third dimensions, with important consequences for the world economy and its value chains. Most evident is the mediating role of universities and academics for Strange's fourth dimension of power; i.e. the capacity to develop, acquire and deny the access of others to knowledge. The interesting question is, of course, who indirectly benefits from the mediation, given the hierarchy built into the system of intellectuals. What are the consequences of the most recent effort of the European Union to strengthen academic freedom by making its respect a precondition for its funding of international scientific collaboration (European Parliament, 2018)? How do decolonial studies transform the hierarchy built into the mediation? These studies have gained momentum and attention in recent years. They highlight the inequalities built into academic knowledge creation, as well as the access to it and empower new types of knowledge, often framed as local or indigenous knowledge (see, e.g., Santos, 2014; Connell, 2007; Mignolo, 2021).

Examining the role of higher education and academic knowledge for international relations is thus a very topical endeavor. However, it requires a broader analytical perspective than most educational diplomacy studies offer to gain a better understanding of the mediation between different ways of knowledge and reality definition. Without strengthening its analytical power, the field of research risks becoming part of the host countries' propaganda, as I have argued in this contribution.

References

- Adams, R., 2021. Tuition fees from UK students fail to cover costs of undergraduate courses at UCL University College London and other institutions must rely on fees from international students, provost says. *The Guardian*, 11 September 2021. <https://www.theguardian.com/education/2021/sep/11/tuition-fees-uk-students-fail-to-cover-costs-undergraduate-courses-ucl>.
- Altbach, P.G., 2012. China Has Right Approach to Foreign Collaboration. *Inside Higher Ed: The World View*. February 19, 2012. <https://www.insidehighered.com/blogs/world-view/philip-altbach-china-has-right-approach-foreign-collaboration>.
- Ashley, R., 1987. The geopolitics of geopolitical space: toward a critical social theory of international politics. *Alternatives* XII, 403–434.
- Atkinson, C., 2010. Does soft power matter? A comparative analysis of student exchange programs 1980–2006. *Foreign Pol. Anal.* 6, 1–22.
- Australian Government, 2020. 2019 International Student Data.
- Babones, S., 2019. The China Student Boom and the Risks it Poses to Australian Universities. *Analysis Paper 5*. Centre for Independent Studies, Sydney.
- Banikamal, A., Ra'ees, W., 2018. Iran's educational diplomacy in the Muslim world: activities of Al-Mustafa International University in Malaysia and Afghanistan. *Insight Turkey* 20, 165–188.
- Bayat, M., 2014. Global education diplomacy and inclusion in developing countries: lessons from West Africa. *Child Educ.* 90, 272–280.
- Bevis, T.B., Lucas, C.J., 2007. *International Students in American Colleges and Universities: A History*. Palgrave Macmillan, New York, NY.
- Bhabha, H., 2004. *The Location of Culture*. Routledge, London.
- Bislev, A., 2017. Student-to-student diplomacy: Chinese international students as a soft-power tool. *J. Curr. Chines Aff.* 46, 81–109.
- Bizzell, B.E., Counts Kahila, R., Talbot, P.A.T. (Eds.), 2018. *Cases on Global Competencies for Educational Diplomacy in International Settings*. IGI Global, Pennsylvania.
- Cantwell, B., 2015. Are international students cash cows? examining the relationship between new international undergraduate enrollments and institutional revenue at public colleges and universities in the US. *J. Int. Stud.* 512–525.
- Cassin, B., Pter, E., Lezra, J., Wood, M. (Eds.), 2014. *Dictionary of Untranslatables: A Philosophical Lexicon*. Princeton University Press, Princeton.
- Connell, R., 2007. *Southern Theory: The Global Dynamics of Knowledge in Social Science*. Polity Press, Cambridge.
- Cull, N.J., 2020. Public diplomacy before Gullion: the evolution of a phrase. In: Snow, N., Cull, N.J. (Eds.), *Routledge Handbook of Public Diplomacy*. London Routledge, New York.
- DAAD, 2018. *The DAAD Alumni Concept – Goals and Focuses*.
- DAAD, 2021. *Wissenschaft weltweit Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland und weltweit*. DAAD, Bonn.
- Dahl, R.A., 1957. The concept of power. *Syst. Res. Behav. Sci.* 2, 201–215.
- Daniel, J., 2003. Higher education: past, present, and future - a view from UNESCO. *High Educ. Eur.* 28, 21–26.
- de Ville, F., Siles-Brügge, G., 2019. The impact of Brexit on EU trade policy. *Polit. Govern.* 7, 7–18.
- Dodds, K., Kuus, M., Sharp, J. (Eds.), 2013. *Introduction: Geopolitics and its Critics from: The Ashgate Research Companion to Critical Geopolitics*. Routledge, London.
- El-Nawawy, M., 2006. US public diplomacy in the Arab world the news credibility of Radio Sawa and Television Alhurra in five countries. *Glob. Media Commun. Assoc. Inform. Syst.* 2, 183–203.

- European Parliament, 2018. European Parliament Recommendation of 29 November 2018 to the Council, the Commission and the Vice-President of the Commission/High Representative of the Union for Foreign Affairs and Security Policy on Defence of Academic Freedom in the EU's External Action (2018/2117(INI)).
- Gramsci, A., 1999 [1971]. *Selections from the Prison Notebooks of Antonio Gramsci*. ElecBook, London.
- Guan, J., Dodder, R.A., 2001. The impact of cross-cultural contact on value and identity: a comparative study of Chinese students in China and in the U.S.A. *Mank. Q.* 41, 271–288.
- Guzzini, S., 2021. Power in World Politics, DIIS WORKING PAPER 2021: 17. Danish Institute for International Studies, Copenhagen.
- Hadfield, A., 2015. Commune consensus: a soft power comparison of the commonwealth and the European Union. *Round Table: Commonwealth J. Int. Aff.* 104, 397–412.
- Hall, T., 2010. An unclear attraction: a critical examination of soft power as an analytical category. *Chin. J. Int. Polit.* 3, 189–211.
- Ham, M., Tolentino, E., 2018. Socialisation of China's soft power: building friendship through potential leaders. *China Int. J.* 16, 45–68.
- Hartmann, E., 2010. The United Nations Educational, Scientific and Cultural Organisation: pawn or global player? *Glob. Soc. Educ.* 8, 307–318.
- Hartmann, E., 2015. The educational dimension of global hegemony. *Millennium* 44, 89–108.
- Hartmann, E., 2021. The internationalisation of further education: between geoeconomics and geopolitics. In: Parreira Do Amaral, M., Thompson, C. (Eds.), *Geopolitical Transformations in Higher Education, Imagining Fabrication and Contesting Innovation*. Springer, Cham.
- Harvey, D., 2004. The "new" imperialism: accumulation by dispossession. In: Panitch, L., Leys, C. (Eds.), *Socialist Register* (London).
- HEPI, 2021. Paying More for Less? Careers and Employability Support for International Students at UK Universities. Higher Education Policy Institute, Oxford.
- Hillman, N., 2018. UK Slips behind the US, Which Takes the Number One Slot, for Educating the World's Leaders. Higher Education Policy Institute, Oxford.
- HM Government, 2019. International Education Strategy: global potential, global growth, March. Department for Education, Department for International Trade, London.
- Hone, K., 2014. Education Diplomacy – towards a Common Understanding. Network for international policies and cooperation in education and training).
- Hong, N.Y., 2014. EU-China education diplomacy: An effective soft power strategy? *Eur. Foreign Aff. Rev.* 19, 155–172.
- House of Lords, 2014. House of Lords Soft Power and the UK's Influence Committee, Persuasion and Power in the Modern World, 28 March 2014, HL Paper 150 of session 2013–14.
- Ikenberry, J.G., 2012. *Liberal Leviathan: The Origins, Crisis, and Transformation of the American World Order*. Princeton University Press, Princeton, NJ.
- Ikenberry, J.G., Kupan, C.A., 1990. Socialization and hegemonic power. *Int. Organ.* 44, 283–315.
- Ives, P., 2009. Global english, hegemony and education: lessons from gramsci. *Educ. Philos. Theor.* 41, 661–683.
- Jain, R., 2020. China's strategic foray into higher education: goals and motivations vis-à-vis Nepal. *Dipl. Statecr.* 31, 534–556.
- Kettle, M., 2020. Dawn of Asian Century Puts Pressure on EU to Choose Sides, Says Top Diplomat. *The Guardian*, 25 May.
- Knight, J., 2019. *Knowledge Diplomacy in Action*. British Council, London.
- Kremenyuk, V., 1995. The Soviet Union and Europe in historical perspective. In: Gottstein, K. (Ed.), *Tomorrow's Europe. The Views of Those Concerned*. Campus Verlag, Frankfurt/M.
- Kukkola, J., 2018. Russian Cyber Power and Structural Asymmetry, pp. 362–368.
- Lo, T.Y.J., Pan, S., 2021. The internationalisation of China's higher education: soft power with 'Chinese characteristics'. *Comp. Educ.* 57, 227–246.
- Lomer, S., Papatsiba, V., Naidoo, R., 2018. Constructing a national higher education brand for the UK: positional competition and promised capitals. *Stud. High Educ.* 43, 134–153.
- Lukes, S., 2005. Power and the battle for hearts and minds. *Millennium* 33, 477–493.
- Macaulay, T.B., 1835 [1957]. *Minute of 2 February 1835 on Indian Education*. Harvard University Press, Cambridge.
- Marginson, S., 2018. The UK in the global student market: second place for how much longer?. In: *Centre for Global Higher Education Research Findings*. Centre for Global Higher Education.
- Marks, S., Freeman, C.W., 1998. *Diplomacy*. Encyclopaedia Britannica.
- McGill Peterson, P., 2014. Diplomacy and education: a changing global landscape. *Int. Higher Educ.* 75, 2–3.
- Mignolo, W.D., 2021. *The Politics of Decolonial Investigations*. Duke University Press, Durham.
- Milton, P., 2020. The Mutual guarantee of the peace of Westphalia in the law of nations and its impact on European diplomacy. *J. Hist. Int. Law* 22, 101–125.
- Nye, 2004. *Soft power: the means to success in world politics*. Public Affairs, New York.
- Nye, J.S., 2008. Public diplomacy and soft power. *Ann. Am. Acad. Polit. Soc. Sci.* 616, 94–109.
- Nye, J.S., 2021. Soft power: the evolution of a concept. *J. Polit. Power* 14, 196–208.
- OECD, 2020. *Education at a Glance 2020: OECD Indicators*. OECD, Paris.
- OSCE, 1975. *Conference on Security and Co-operation in Europe. Final Act*.
- Özel, M., 2021. International exchange programmes in the context of educational diplomacy: the example of the Erasmus+ program. *Turkish online J. Des. Art commun.* 11, 943–966.
- Parmar, I., 2018. The US-led liberal order: imperialism by another name? *Int. Aff.* 94, 151–172.
- Perraton, H.D., 2014. *A History of Foreign Students in Britain*. Palgrave Macmillan, Basingstoke.
- Pervillé, G., 1984. *Les étudiants algériens de l'université française, 1889-1962: populisme et nationalisme chez les étudiants et intellectuels musulmans algériens de formation française*. Éditions du CNRS, Paris.
- Raev, A., Minkman, E., 2020. Emotional Policies: Introducing Branding as a Tool for Science Diplomacy. *Humanities and Social Sciences Communications*, p. 7.
- Rana, K.S., 2016. India and China: soft power in an Asian context. In: Chitty, N., Ji, L., Rawnsley, G.D., Hayden, C. (Eds.), *The Routledge Handbook of Soft Power*. Routledge, London.
- Redden, E., 2020. For five consecutive years, the number of new international students Inside Higher Ed. November 16, 2020. <https://www.insidehighered.com/news/2020/11/16/survey-new-international-enrollments-drop-43-percent-fall>.
- Rizvi, F., 2019. Marketisation Weakens the Public Diplomacy Role of HE. *University World News*, 12 October.
- Santos, B.D.S., 2014. *Epistemologies of the South: Justice against Epistemicide*. Routledge, London.
- Scott, J.C., 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. Yale University Press, New Haven CT.
- Strange, S., 1988. *States and Markets. An Introduction to International Political Economy*. Blackwell Pub, London.
- Syrotinski, M., 2019. Postcolonial untranslatability: reading Achille Mbembe with Barbara Cassin. *J. Postcolonial Writ.* 55, 850–862.
- Terry, D., 2021. Australia needs a new model of international education for the post-Covid era. *Times Higher Education*. August 26, 2021. <https://www.timeshighereducation.com/opinion/australia-needs-new-model-international-education-post-covid-era>.
- Thomas, P.D., 2007. Gramsci and the intellectuals: modern prince versus passive revolution. In: Bates, D. (Ed.), *Marxism, Intellectuals and Politics*. Palgrave Macmillan, Basingstoke.
- Tsvetkova, N., 2008. International education during the cold war: soviet social transformation and American social reproduction. *Comp. Educ. Rev.* 52, 199–217.
- Turekian, V., 2018. The evolution of science diplomacy. *Glob. Policy* 9, 5–7.
- Tylecote, R., Clark, R., 2021. Inadvertently Arming China? the Chinese Military Complex and its Potential Exploitation of Scientific Research at UK Universities (Civitas).
- UN, 1945. *United Nations Charter*. United Nations.
- UNESCO, 1979. *Convention on the Recognition of Studies, Diplomas and Degrees in Higher Education in the States Belonging to the Europe Region*, Adopted at Paris, 21 December 1979, UN Treaty Series NO. 20966.

- UUKI, 2020. Supporting International Graduate Employability: Making Good on the Promise. Univesities UK International. January.
- Weber, M., 1978. *Economy and Society*. University of California Press, Berkeley.
- Wu, H., 2019. Three dimensions of China's "outward-oriented" higher education internationalization. *High Educ.* 77, 81–96.
- Yu, F.-L.T., 2019. Neo-mercantilist policy and China's rise as a global power. In: Yu, F.-L.T., Kwan, D.S. (Eds.), *Contemporary Issues in International Political Economy*. Palgrave Macmillan, Singapore.

Populism and education

Edda Sant, Manchester Metropolitan University, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	648
What is populism?	648
Who are “the populists”?	649
Is populism a democratic threat?	650
Is populism a consequence of educational failure?	651
Should education struggle against populism?	652
Education as a struggle against anti-rationalism	652
Education as a struggle against economic deprivation	653
Should education learn from populism?	653
Learnings from populism for educational theory	653
Learnings from populism for educational practice	654
Learnings from populism for educational research	655
Conclusions	655
References	655

Introduction

Populism—a “vague” term referring to political practices that polarize society into two opposite groups, the elite and the people—very often functions as a “dirty” word to disqualify others. Many politicians including Donald Trump (former United States President), Boris Johnson (United Kingdom First Minister), Hugo Chavez (former Venezuela President) or Narendra Modi (Indian Prime Minister) have been described as populists often in an attempt to compromise their public image. Populists are told to be insincere and manipulative, and their supporters are portrayed as ignorant or low educated. Thus, the question of populism often has triggered educational debates. Education (or its lack) has been seen as both a cause and a solution to populism.

However, political and educational theory and research provide a far more nuanced image of populism. The question of populism is indeed very contested. There is no clear answer to what populism is, how populism relates to education and whether education should somehow react to the rise of populist leaders. This article will interrogate different controversies associated with populism and its relation to education: (1) what is populism?; (2) who are the “populists”?; (3) is populism a democratic threat?; (4) is populism a consequence of an educational failure?; (5) should education struggle against populism? and (6) should education learn from populism? This paper aims to shed light on populism and its implications for education.

What is populism?

There is no agreement on what populism is. In the introduction of this article, I defined populism as a vague term referring to political practices that polarize society into two opposite groups, the elite and the people, where the people are considered the source of democratic legitimacy. In populism, the will of the people is deemed to be preferable and non-negotiable. I suspect this definition could be shared among most of those studying populism. However, outside this generic definition, political scientists disagree with practically everything else.

Some academics argue that “the people” of populism is always a homogeneous and anti-pluralist group. According to Müller (2017), the key populist feature is their intention to destroy their enemies so only “the people have access to power. Others, in contrast, explain that “the people” can be heterogeneous providing they are united in a cause. Analyzing the case of SYRIZA in Greece, Stavrakakis and Katsambekis (2014) argue that, for populism to exist, the people only need to be united in their political struggle against the elite.

Some authors argue that populist movements always display authoritarian and charismatic leadership, and the leader presents themselves as the one who unites the people and represents its true nature (Weyland, 2017). This perspective is, nevertheless, challenged by many (see. e.g., Moffitt and Tormey, 2014). Most political scientists appear to agree that there are too many examples of populist movements (e.g., Tea Party in the United States; BREXIT in the United Kingdom) without a clear leader.

A key issue is whether populism is always right-wing orientated. Some argue that populism tends to be associated with extreme-right ideologies that show no consideration for those who do not belong to the majoritarian ethnic group (Müller, 2017). Such is the case, for instance, of Marine Le Pen in France or Pauline Hanson in Australia. This explanation also triggers associations between populism, fascism and Nazism and favors the pejorative use of the term populism. However, this perspective is highly contested; there is a wide range of left-wing populists who openly challenge xenophobia and racism (e.g., Pablo Iglesias; Evo Morales; Bernie

Sanders). Many argue that populism is very often combined with other ideologies. This is why we can have right-wing and left-wing populists. In some cases, populism is combined with nativism and/or racist attitudes resulting in right-wing or far-right-wing populism. In other instances, populism is combined with social-democratic values resulting in left-wing or far left-wing populism (Mudde and Rovira Kaltwasser, 2017).

The final disagreement relates to the performative style portrayed by populists. According to Ostiguy (2017), populists always present themselves as outsiders, challenge political and social correctness, and favor ordinary individuals' lived experiences over technical and academic expertise. However, while this perspective clearly defines some populist leaders such as Thaksin Shinawatra in Thailand or Abdoulaye Wade in Senegal, it does not define them all. Alexis Tsipras (SYRIZA leader), Preston Manning in Canada or Alberto Fujimori in Peru responded to more conventional political behaviors, and Pablo Iglesias (former Spanish Vice-President) is an academic himself.

Who are “the populists”?

It is complicated to find out political leaders or movements who are unanimously defined as populists. As the term populism is often used pejoratively, those defined as populist tend to deny such accusations or pass the label of populist to someone else. Michael Gove, former secretary of education in the United Kingdom very often described by the press as populist, once congratulated his country for getting rid of populist parties.

The word “populist” does not have the same connotations across time and space. Populism derives from the Latin *populi* (people), term that encompassed three different meanings: the people holding political power, the people as the nation, and the “common people” (in opposition to the ruling elite) (Canovan, 2004). Besides, in languages such as English, “the people” also can be used as synonymous of humanity. The way we understand populism is highly conditioned by our linguistic, cultural, and historical context.

Political scientists, nevertheless, tend to agree that the first manifestations of “modern” populism can be tracked to 19th century Russia. The *narodniki* was a group of young urban students who romanticized the countryside and tried to mobilize its peasants to overthrow the Tsar regime (Rovira Kaltwasser et al., 2017). Just a few years later, in 1890s United States, the People's Party became relatively successful in the rural Southern and Western states with a discourse against the corruption of Washington politics, the manipulation of the press, and the economic inequality between “paupers and millionaires” (cited in Laclau, 2007, p. 201).

Not without controversy, empirical research suggests that there have been manifestations of populism across the different continents. In Asia, Prime Minister of Thailand Thaksin Shinawatra (2001–2006) gained support from rural poor, styling himself as a man “of the people”, wearing informal clothes, and talking about his sex life in local slang (Moffit and Tormey, 2014). Indian Prime Minister Narendra Modi (2014–) is often defined as the Asian populist *par excellence*. Modi extensively uses social media and presents Hindu people in cultural opposition to the English-speaking, Westernized and secular elites (Rao, 2018).

In Australia and New Zealand, two politicians are more often described as populists. Pauline Hanson, the extreme right Australian One Nation Party, repeatedly appeared in reality shows presenting herself as a “fish and chip shop lady” and complaining against Asian immigration, multiculturalism and aborigines' rights (Moffit and Tormey, 2014). In New Zealand, Winston Peters secured political power as Deputy Prime Minister and Treasurer (1996–1998) and Minister of Foreign Affairs (2005–2008) (Moffit, 2017) with a discourse against intellectual elites and immigrants.

Africa has seen different populist waves. According to Resnik (2017), African populism was initially connected to the *coups* in the 1980s where different military figures such as Captain Thomas Sankara in Burkina Faso (1983–1987) and Flight Lieutenant Jerry Rawlings (1981–2001) in Ghana orchestrated “popular” revolutions against civilian governments defined as corrupted elites. More recently, others have capitalized the continent's economic and demographic changes, particularly the increasing urbanisations and inequality experienced by African youth. For instance, in Senegal, Abdoulaye Wade obtained the Presidency (2000–2012) after having been proclaimed *Président de la rue* (“President of the street”)—label he gained through campaign rallies in which he aligned himself with the poor in opposition to the political elite.

Populism has been more widely studied in the Americas and Europe. North America has seen numerous cases of unquestionable populism. For instance, in the United States, a range of political leaders including Ronald Reagan (1911–2004) and Sarah Palin (1964–) have been defined as populists for their appeal to the “common man” to struggle against the “corrupt elites” of Washington (Lowndes, 2018). In the 2016 Presidential elections, two populist figures, one in the right (Donald Trump) and one in the left (Bernie Sanders), (more or less) successfully challenged non-populist candidates. They mobilized a huge majority of American voters (Hawkins et al., 2017). Trump successfully appealed to the “common” Americans, against Muslims, women, Latin American immigrants and his political opponents overall (Inglehart and Norris, 2016). Bernie Sanders, in contrast, was closed to be nominated Democratic candidate, with a discourse attacking “big corporations, big donors, and big banks” (Inglehart and Norris, 2016, p. 5). While many argue that Trump is undoubtedly populist, there is academic and public debate about whether Sanders is or is not (Sant, 2021).

A relatively high number of Latin American political movements, parties and leaders have been qualified as populists. During the 1930s and the 1940s, leaders such as Juan Perón (Argentina) and Getúlio Vargas (Brazil) successfully obtained power after performing various collective actions, including mass rallies. Whether later politicians in early 1990s (e.g., Carlos Menem in

Argentina and Alberto Fujimori in Peru) were populists is still controversial. According to [De La Torre \(2017\)](#), these political leaders were “neoliberal populists”. They blamed traditional politicians for the economic situation of their countries and, once in power, imposed austerity and privatizations reforms based on the recommendations of the International Monetary Fund (IMF). At the beginning of the 21st century, Hugo Chávez, Evo Morales, and Rafael Correa became the images to illustrate left populism: they “defended the people” against neoliberal practices and the “intrusion” of IMF, the World Bank, and the US government in Latin American politics ([Panizza, 2008](#)). Some years later, Jair Bolsonaro capitalized on political uncertainty in Brazil, and with a discourse framed by racism, homophobia, sexism and military authoritarianism won the elections in October 2018 ([Gledhill, 2019](#)).

In contemporary Europe, who is and who is not populist has been a long controversy. Were the 1920s and 1930s totalitarian movements (i.e., Nazism and fascisms) populists ([Müller, 2017](#))? While many academics appear to disregard such association, the potential links between fascism, racism and populism have nurtured the European demonization of populists ([Mudde and Rovira Kaltwasser, 2012](#)). Indeed, many political scientists argue that, beyond the *narodniki*, modern populism is a relatively new phenomenon in Europe. In the 1990s, populist parties emerged in Central and Eastern Europe as a reaction of the “losers” of the liberal politics introduced during post-communist transition against the elites who implemented these reforms ([Stanley, 2017](#)). Later parties have further used the uncertainties created by rapid changes to gain the electorate. For instance, Law and Justice (*Prawo i Sprawiedliwość*, PiS) drawn upon the frustration of those experiencing the grievances of economic crisis, liberal policies, political corruption, and Russian intrusion ([Sztompka, 2016](#)). South and Western Europe have also experienced a rise of populist movements. For instance, in Greece and Spain, left-wing movements have secured some political power. The Greek coalition SYRIZA gained power in 2015 with a discourse against the corrupted elites and more, particularly against the politics of austerity fostered by the *troika* (European Commission, European Central Bank and IMF) ([Stavrakakis and Katsambekis, 2014](#)). In Spain, *Podemos* (We can) entered the coalition in 2019 with a discourse against the political, economic and judicial elites accused of corruption and keeping Franco’s dictatorship alive ([Franzé, 2017](#)).

Meanwhile, right-wing populist parties have also expanded their number of electors. In The Netherlands and France, Geert Wilders and Marine Le Pen’s National Rally (former *Front National*) foster a populist discourse against Muslim migration. The Brexit Referendum (the 2016 United Kingdom European Union membership referendum) is often considered a populist movement. Simultaneously, so-called populist infiltrated mainstream parties. Boris Johnson (UK First Minister) and Jeremy Corbyn (former Labor party leader) were the two leading candidates in the 2019 United Kingdom general election. While many saw in Johnson the clear image of populism, many others questioned that Corbyn was a populist himself. And yet, the opening lines of Corbyn’s first speech of the 2017 campaign were,

“The dividing lines in this election could not be clearer from the outset ... It is the establishment versus the people and it is our historic duty to make sure that the people prevail ... We don’t fit in their cosy club. We’re not obsessed with the tittle-tattle of Westminster or Brussels. We don’t accept that it is natural for Britain to be governed by a ruling elite, the City and the tax-dodgers, and we don’t accept that the British people just have to take what they’re given, that they don’t deserve better.”

Quoted by [Flinders \(2018, p. 233\)](#).

Who is and who is not a populist is not an easy question to answer.

Is populism a democratic threat?

One of the main reasons populism matters for education is that populism is often conceived as a democratic threat. Democracy is an ambivalent term with many meanings, but authors usually define democracy in terms of equality and popular sovereignty ([Sant, 2019a, 2021](#)). Many academics in education believe that populism challenges equality; they argue that “the people” is simultaneously constructed against the elite and the immigrant Other who is racialized and considered an enemy that needs to be expelled (e.g., [Osler and Starkey, 2018](#)). If that is the case, populism uncontestedly damages the democratic principles of equality. However, the way this perspective conceptualizes populism is challengeable. According to [Stavrakakis and Katsambekis \(2014\)](#), this perspective fails to understand that populism is not necessarily associated with nativist or far-right-wing attitudes. Many populists struggle against racism and inequality. Extreme right politics are a democratic threat—regardless of whether these politics are populists or not, the relationship between populism and democracy is more complicated.

In many countries, democracy is often understood as synonymous of liberal or representative democracy. Liberal democracy favors both equality and popular sovereignty (democracy) and individual liberty, separation of powers and the rule of law (liberalism). Populism emphasizes popular sovereignty (the will of the people) but not through the liberal principles of separation of powers and the rule of law. For instance, in the United Kingdom, after the results of the Brexit Referendum, many of those supporting to leave the European Union argued that the people had decided and the parliament should stop discussing and “just” deliver Brexit. [Rovira Kaltwasser \(2012\)](#) explains, “[p]opulism assumes that once ‘the people’ have spoken, nothing should constrain the implementation of its will” (p. 195). Most academics appear to agree that populism privileges the rule of the people and that this can represent a threat to the existing authority. However, this does not make populism necessarily

a democratic threat but rather a threat to liberal democracy. [Martinelli \(2016\)](#) argues, “[p]opulism is not anti-democratic (...), but it implies an illiberal version of democracy” (p. 13).

More widely, [Canovan \(1999\)](#) and [Arditi \(2004\)](#) define populism as a shadow of democracy. In their perspective, any democratic system requires a balance between ideal and pragmatic politics. Ideal democracy responds to the rule of the people—the “utopian” situation in which there is no gap between public and power, where the people, as a political community, have all control to self-govern, and where self-government totally satisfies all members of this community. Pragmatic politics, in contrast, focuses on solving everyday disputes through institutions—the rule of law—that give answers to such disputes. Populism, as mentioned, considers the rule of the people is non-negotiable and, in doing so, utilizes the internal mismatch between ideal and pragmatic democracy. Populists draw upon the fact that pragmatic politics are seen as elitist and irresponsible because they cannot deliver the will of the people in a fast and satisfactory manner. Populism appeals to the ideal face of democracy and questions the irresponsible elitism of its pragmatic mechanisms. From this perspective, populism is not a democratic threat, but it evidences how the rule of the people and the rule of law can challenge each other.

Is populism a consequence of educational failure?

Social and political scientists have emphasized the importance of understanding the role education plays in populism. In 2016, [Runciman \(2016\)](#), a well-known professor of politics, wrote a newspaper article entitled “[h]ow the education gap is tearing politics apart”. Indeed, empirical research in sociology and political science has suggested that those graduated from colleges and universities rarely vote populist leaders. Those with lower or no educational qualifications are more likely to support populist alternatives ([Goodwin and Heath, 2016](#); [Inglehart and Norris, 2016](#)). This relationship has been proved in different contexts, including Belgium, The Netherlands, the United Kingdom and the United States. For instance, in the Brexit referendum, while over two-thirds of those without qualifications voted to leave, only a quarter of those with postgraduate qualifications did so ([Hobolt, 2016](#)). The twenty “most educated” areas voted to remain, whereas fifteen of the twenty “least educated” areas voted to leave ([Goodwin and Heath, 2016](#)). In the 2016 US Presidential elections, Trump enjoyed a narrow margin of four points over Hillary Clinton among white college-educated, but a vast advantage of 40 points among the non-college-educated ([Schaffner et al., 2018](#)). It has been argued that even considering other factors (e.g., age, gender, region, ethnicity, social class), education remains one of the most (if not the most) important cleavages ([Goodwin and Heath, 2016](#); [Hobolt, 2016](#); [Runciman, 2018](#); [Schaffner et al., 2018](#)).

However, this relation appears to be more complex. In the United Kingdom, [Dorling et al. \(2016\)](#) have suggested that, as for the Brexit vote, it is very difficult to separate educational attainment from age. A master’s degree in 2016 could be equivalent to a good post-16/pre-university qualification on 1946. The Brexit vote cannot fully be examined without considering historical dynamics and the hidden experiences of those who did not vote. Several left-wing populist European parties have the support of a highly educated population ([Ramiro, 2016](#)). For instance, [Ramiro and Gomez \(2017\)](#) have shown how a sector of *Podemos* electorate in Spain comprises university-educated young professionals in situations of economic insecurity. In 2017 United Kingdom general elections, the Labor party experienced a surge of young university-educated voters who felt appealed by the populist nature of Corbyn’s discourse ([Sloam and Henn, 2019](#)). In the United States, there are not apparent educational differences between those who supported Bernie Sanders or those who supported other (non-populist) Democratic candidates ([Dyck et al., 2018](#)). Similar demographic patterns have been found in numerous Latin American countries ([Remmer, 2012](#)).

The relation between populism and education appears to be context-dependent, but this has not stopped academics to explore why this relationship exists. There are four different hypotheses. First, there is a cognitive hypothesis. [Inglehart and Norris \(2016\)](#) explicitly refer to knowledge and cognitive skills to justify this relation. Similarly, [Spruyt et al. \(2016\)](#) explain, “[t]his relationship is to a certain extent mediated by the indicators that refer to the more evolved cognitive aspects people acquire through education” (p. 8). There is a second status hypothesis. Some argue that this relation is a consequence of the status effect of education; those who have a higher educational qualification are more likely to have better career opportunities, higher salaries and overall to gain higher socio-economic status ([Inglehart and Norris, 2016](#); [Spruyt, 2014](#)). Achieving a higher degree makes one less likely to be/become left behind by neoliberal globalizations ([Teney et al., 2014](#)). Meanwhile, non-college educated (often male) feel resentment against higher educated elites for their lost of status ([Cohen, 2019](#)). The cognitive and status hypothesis implicitly see populism as a result of educational failure: populists have not gained an education and therefore have not acquired knowledge and skills and the jobs and income expected to accompany those learnings.

The third hypothesis is the socializing hypothesis. Some have explained this potential division in terms of the socializing effects of schooling practices ([Inglehart and Norris, 2016](#); [Spruyt et al., 2016](#); [Spruyt, 2014](#)). Those who spend longer in education are more likely to be socialized into normative frameworks that might challenge populist discourses. There is also the possibility (i.e., fourth hypothesis) that individuals might have educational identities and that these identities might fuel/prevent alignment with populist discourses ([Runciman, 2018](#); [Spruyt, 2014](#)). [Bovens and Wille \(2017\)](#) and [Spruyt \(2014\)](#) defend the conflict hypothesis: the existence of a perceived educational conflict between higher and lower educated. They explain the rise of populism as a revolt of the lower educated that might be seen as the emergence of “the people” as a collective actor ([Laclau, 2007](#)). In contrast with the cognitive and the status hypothesis, the socializing and the conflict hypothesis offer a much more nuanced account signaling that populist

and no-populists might have different life (including educational) experiences that might favor them to see the world in different ways. [Runciman \(2018\)](#) explains,

"One of the largest divides in a twenty-first-century Western democracy is produced by education. Whether or not someone went to college is a more significant determinant of how they are likely to vote than age, class or gender. (...) The educated are a tribe. They stick together. They may tell themselves that they do this because they have a better understanding of how the world works. But that's what makes them so alienating to the other side: they appear to mistake their tribalism for superior wisdom" (pp. 163–164).

These two hypotheses also help to explain why the potential relationship between populism and education outlined above is quite "clear" in some cases, and it is not that clear in others. Here, the sum of different life experiences (with education being an essential part of them) creates our "tribes", and we vote the same that members of our "tribes" regardless of whether we vote populist or not populist alternatives.

Should education struggle against populism?

Many believe that education should struggle against populism. Often driven by concerns about the health of liberal democratic systems (see above), different policymakers and academics have emphasized the role of education in counter-acting populism (see [Sant and Brown, 2020](#)). Two main alternatives have been put forward: a struggle against anti-rationalism and a battle against economic deprivation.

Education as a struggle against anti-rationalism

Many educational policymakers, academics and practitioners have seen education as a potential cure for populism ([Sant and Brown, 2020](#)). For instance, the former Secretary of Education of Spain, Iñigo Méndez de Vigo argued that Europe faced three challenges: terrorism, populism and nationalism. He further explained, "[t]hese could fracture Europe if we fail to invest in education and the fundamental values of social harmony and democracy" ([La Moncloa, 2017](#)). Tom Boland, former chief of the Higher Education Authority (HEA) in Ireland, also wrote in 2016,

"We need the capacity for critical thought and analysis, and we need academics committed to questioning and testing received wisdom, putting forward new ideas and stating controversial and unpopular opinions, and we need government and the institutions to work constructively, mutually to support our democratic society. Populism contains a special threat to all of that. But 'isms' that would destroy the values, rights and freedoms of Western democracy have been defeated before. Higher education must be at the forefront of the struggle."

[Boland \(2016\)](#).

Underpinning these understandings lie different political and educational assumptions: populists are always anti-pluralist, authoritarian, racist, and they represent a democratic threat. The alignment with populist discourses is irrational (cognitive hypothesis above). Populism is here seen as an individual deficit of those who, lacking critical thinking and knowledge, get obsessed with conspiracies, manipulated by unsincere leaders or let their emotions prevail over their reason ([Mudde, 2010](#)). From this perspective, education is essential. Students should learn critical thinking and critical literacy to counter the anti-rationalist (populist) threat. Literacy and critical literacy skills are often identified as the key to uncover populist rhetoric and communication, so students do not feel attracted by populist discourses (see, e.g., [Heggart and Flowers, 2019](#)).

While this perspective is implicit in many educational accounts of populism, it is not undisputable ([Sant and Brown, 2020](#)). There is first an issue with the cognitive assumption underpinning this perspective. As we have seen above, very highly educated people support populist discourses. Knowledge and literacy skills might explain the alignment with populism to some extent, and in some cases, but it is not a philosopher's stone. Why people vote or vote not for populist alternatives is likely to be a more complex issue.

Further, the binary divide between cognition and emotion that this perspective assumes, with populists leaving their emotions to prevail and non-populists using knowledge and reason, is problematic in many ways. Psychologists [Pons et al. \(2010\)](#) have explained that we do not yet know where cognition begins and emotions finishes, and how they relate to each other. Post-structuralists, post-colonialist and feminists have much evidenced how cognitive hierarchies based on reason/emotion divides historically served to reproduce power relations ([Hoppers, 2001](#)). Mentions to irrationality have often been used to lesser others, being those often women, working-class or colored or colonized peoples. Thus, some might argue that we should be suspicious against any form of education that aims to educate the "irrational".

This perspective also assumes that all populist movements are anti-pluralist, authoritarian and racist. As we have seen above, this is also a matter of debate. Even if we just consider far-right-wing forms of populism, whether or not authoritarian and racist attitudes would be changed through an education focused on knowledge, critical thinking and literacy is questionable. [Waisbord \(2018\)](#) explains,

“Facts never change the unfalsifiable premise of populism—the eternal division of ‘pure people’ and ‘evil elites’. This conception of politics turns into a political fantasy that cannot ever be proven wrong (Müller, 2017). Populism dismisses facts that challenge overriding narratives. No matter what happens, populism obstinately clings to the notion that elites are always in power and continue to distort the truth through their institutions. Populism can never be corrected by its critics. Critics can never offer facts that question, challenge, or complement populist assertions.” (p. 10).

Similar arguments have been made by those conducting empirical research. According to Miller-Idriss and Pilkington (2017), educational interventions that aim to reach far right populations can only succeed by acknowledging the affective dimension. Thus, academic proposals that separate cognition from emotion and only focus on facts and knowledge might not succeed in their purpose.

Finally, it can also be argued that this perspective self-defeats itself in presenting populism as a democratic threat that needs to be counter-acted in the educational arena. As we have explored above, whether or not populism is a democratic threat depends on how we understand democracy. This perspective, is suspected, assumes that populism is a threat to democracy because, in conceptualizing democracy as a liberal democracy, it privileges the rule of law over the rule of the people. The problem here is that this perspective appears to protect the rule of law against the rule of the people itself. We might or might not like who has been elected, but the power of this person/party arises from the people’s will. Some might wonder, is it democratic to attempt to educate people to vote anti-populist options, so previous institutions are secured?

Education as a struggle against economic deprivation

Other policymakers, academics and practitioners understand populism is a consequence of economic deprivation and economic deprivation is a consequence of a lack of education. Here, a different kind of educational failure causes populism. Populists have not been able to keep up to date on the skills and knowledge modern society require. For instance, in an interview, Tony Blair, former UK Prime Minister explicitly discussed the rise of populism and argued the answer was not to “blame migrants” but to “get the education and skills” that could lead to better jobs and opportunities (Coughlan, 2016).

This perspective relies on the status hypothesis outlined above. According to Kriesi (2014), global dynamics of economic and cultural globalizations have fostered the existence of a bank of “losers of globalization” who are particularly appealed by populist discourses. Since the end of the second World War, western societies have transitioned to a post-industrial economy where blue-collar male-dominated jobs have decreased and lost economic value (Cohen, 2019). The 2008 financial crisis accentuated economic inequalities in Western societies, increasing the number of “losers” or workers who, after decades of delocalizations, feel that their knowledge and skills are no longer economically valuable. Sztompka (2016) explains how these “losers” are often “less educated, manual workers but also several branches of more skilled workers whose training and skills have turned obsolete” (2016, p. 82). This has led to emotions of resentment, economic insecurity and politico-socio-economic status anxieties that have been exploited by some populist leaders (Cohen, 2019).

While there is much value on the losers of globalization analysis, this analysis is not exempt from critique (see Sant and Brown, 2020). Empirical research often suggests that support for populism is conditioned by education, age, gender and ethnicity but whether or not economic deprivation is correlated to populism is highly questionable (Inglehart and Norris, 2016). Power relations are multi-faceted; some might not have high educational qualifications and yet have good incomes. Many others in a situation of economic deprivation are migrants themselves who would not “blame migrants”. Further, not all education and skills lead to better jobs and opportunities as shown, for instance, by the highly qualified—but poorly paid -electorate of Spanish *Podemos*.

Further, it can be argued that this perspective also ignores that, in capitalist societies, educational qualifications are positional, not absolute. What matters to get a better job is whether you are better qualified than others, not your qualification overall. Thus, while education is often presented as a consumption good that generate returns for progression in the social ladder, the correlation between education and economic status are often challenged by sociological analysis (e.g., Brown et al., 2010). It does not matter if people are better educated; there will always be those who are “lower” educated than others. We can further educate new generations, but within capitalism, there will always be people who will not gain higher-skilled jobs because they will be others more qualified. If we take this critique into account, providing more education to all might not necessarily address the economic inequalities that trigger some forms of populism.

Should education learn from populism?

Some academics in education have questioned whether populism could be meaningful for educational practice and research. Philosophers and political scientists including Ernesto Laclau, Chantal Mouffe, Yannis Stavrakakis, Aletta Norval, Jason Glynos and David R. Howarth have considered that the study of populism provides insight knowledge on how politics function. Following their example, some educational academics have questioned whether the way populism operates can tell us something about education itself.

Learnings from populism for educational theory

Four different main concepts have been borrowed from political theory examining populism: demands, chains of equivalences, antagonism and empty signifiers.

Demands are requests that people pose to those on power. We always have requests over issues—these demands are not necessarily conscious, but we feel we want or need something, and we ask for it. In many occasions, those on power satisfy these demands. But in many other instances, these demands are not met, which triggers the beginning of political movements. Populism is a consequence of many unsatisfied demands with people feeling that those on power are irresponsible.

Chains of equivalences (also equivalential chain, discursive chain). People have different demands: what I want is not necessarily the same as my neighbor wants. We are different (the logic of difference). However, on some occasions, different demands can articulate each other as they all request something from those on power. My neighbor and I might want different things from our local authority, but we both are unhappy with them (the logic of equivalence), and thus, we create an alliance. Chains of equivalences are alliances across people holding different demands when people or groups provisionally put aside their differences in support the achievement of the broader aim. For instance, Bernie Sanders mobilized several heterogeneous groups, including Occupy Wall Street and Black Lives Matter. While each group had different particular demands, they allied—they create a chain of equivalences—in support of Sanders.

Antagonism. Demands are always posed to someone else, and when this someone else is irresponsible to our demands, we might see them as antagonists. In the case of populism, the antagonist is the elite: ‘the people’ are allied against the elite. My neighbor and I might have very different demands, but we are united against our antagonist local authority. This alliance—against the elite—is strong enough for us to mobilize together. While Occupy Wall Street and Black Lives Matters are very different movements with different demands, they both oppose existing inequalities in American society and those who perpetuated such disparities. Donald Trump, himself owner of a big corporation who holds support from white supremacist movements, represents the antagonist to be defeated.

Empty signifiers are signifiers (words, images, signs, etc.) that can take different meanings. “The people”, of populism, is a clear empty signifier: it means different things to different people. For instance, while Donald Trump often defines the American people as the common American who wants little interference from Washington, Bernie Sanders often describes them as oppressed by economic powers such as those of Wall Street. When different groups compete to give meaning to this signifier, the empty signifier becomes a *floating signifier*. In political discourses (populist or not), “the people” is often a floating signifier.

Learnings from populism for educational practice

Education in the “Mode of Populism” (Mårðh and Tryggvason, 2017) aims to explore how these different conceptual tools can be helpful for pedagogical practice. Education in the “Mode of Populism” does not seek to educate young people to be populist or anti-populists but rather consider how, what we know from populism, might facilitate more democratic educational practices.

In this perspective, Szkudlarek (2011) argues that critical curricula content such as “democracy”, “culture”, “nation”, “science” or “the people” are empty signifiers very often in dispute. Thus, educational practices intentionally or not might be participating in the struggle to define what each of these signifiers mean. Educators could facilitate that young people themselves participate in determining the meaning of these contested ideas (see, e.g., Sant and Hanley, 2018). For instance, teachers could ask, who is “the people”?

It has also been argued that a different approach to political demands can be helpful to mobilize young people. My own work (Sant, 2019b, 2021; Sant et al., 2021) following that of others (Mårðh and Tryggvason, 2017; Ruitenberg, 2009; Zembylas, 2020) has questioned whether educators should facilitate that students can publicly demonstrate their demands and affective investments. Exploring how young people feel about different contradictory issues could foster educational and political engagement and allow students to experience how those perceived as “antagonists” in some political struggles can be allies in others (Sant et al., 2021).

Mårðh and Tryggvason (2017), Snir (2017) and Sant (2021) have also argued that learning from populism might offer opportunities to generate collective identities. After exploring students’ demands, teachers could facilitate that students articulate among themselves and others in equivalence (Snir, 2017). The role of educators here would be to create spaces of discussion where students (and teachers) can raise their demands, solidarize with others, and build new political projects.

Finally, it has also been questioned whether educators could utilize the tools above to interrupt challenging populist discourses. In line with Michael Apple’s (2015) “politics of interruption” (174), educators could attempt to challenge existing political alliances and create new ones so students position themselves against oppressive regimes. Thiel (2019) has argued that higher educators could mobilize a discourse of “the people” presenting themselves in alliance with university students against existing oppressive structures of governance. In my work, I have questioned whether educators could appeal to some egalitarian demands within populist movements (e.g., political and economic equality) to stop young people feeling attracted to extreme right political options (Sant, 2019b).

These proposals, nevertheless, also hold significant challenges. In one extreme, there is the risk of educators perpetuating existing anti-democratic threats by making room for some challenging discourses. What would happen, for instance, if students would antagonistically present themselves against each other or against those who are not in a position of defending themselves? Indeed, as Mårðh and Tryggvason (2017) argue, educational practices need to be assessed in relation to their intentions and their consequences. If in an attempt to make pedagogical practices more democratic, educators do indeed facilitate the spread of extreme-right discourses, this would certainly be counter-productive.

In the other extreme, there is the risk of educators attempting to socialize students on the political options they find preferable. If educators do indeed interrupt existing political alliances and create new partnerships with their students, this might be ethically and legally problematic. Many countries have legislation that stop teachers and other educators from “indoctrinating” children and

young people in a partisan matter. Further, in the same way as above, we could wonder, is it ethically appropriate for a teacher to attempt to educate young people so they support their own political perspective?

Learnings from populism for educational research

Educational researchers have also considered how the tools derived from the study of populism can be used for discourse analysis. The terms demands, chains of equivalence, empty and floating signifiers have been used to explore the complex way in which people understand ideas such as “the nation” (Sant and Hanley, 2018), “the people” (Sant, 2019b), “democratic education” (Sant, 2019a), education for global citizenship (Mannion et al., 2011; Pais and Costa, 2020), “educational assessment” (Thiel, 2019) or “good teaching” (Trent, 2011). When explored as empty signifiers, these concepts can be examined in all their complexity, considering the demands that aim to give meaning to the signifier and how these demands relate to each other in chains of equivalences. The use of empty signifiers can facilitate uncovering how ideas overlap, evolve, contradict and articulate to each other. The drawback of this tool for educational research, is its highly situated nature. In the same way it is challenging to generalize anything related to populism. Any findings derived from educational research on empty signifiers are very likely to be context-dependent and ungeneralizable to other times and places.

The conceptual tool of “antagonism” has also been used for educational research. For instance, in my work, I have explored whether university students conceptualize “global citizenship” in all-encompassing or on antagonistic terms (Sant, 2018). Others have also used the tool of antagonism to examine how religious education is understood in teacher education practices (Carlsson, 2020) or how UK university students and lectures position themselves in relation to others (Lapping, 2005; Thiel, 2019). This conceptual tool also has its drawbacks. Similarly as above, the use of antagonism as a tool for discourse analysis is context-dependent, and researchers cannot assume that antagonistic relations found exist elsewhere or will always exist. Also, it can be argued that the use of antagonism as a conceptual tool presupposes a binary understanding of societies that, while implicit in Western cultures, does not exist elsewhere (see, e.g., Sant, 2018, 2021). Thus, educational researchers might need to be careful to not imposing a “binary tool” that might tell more about their assumptions than about the data they are indeed examining.

Conclusions

Populism is not an easy concept. Academic theory and research on populism offer a far more nuanced image of populism than the one we sometimes encounter in the press or everyday speech. What populism is and whether or not populism is bad for democracy is a matter of debate. The same can be said for education. Populism cannot be easily used to justify why educational interventions are needed. Whether or not populism is caused or can be corrected via education is questionable. Further, whether it should indeed be “corrected”—via education or any other social intervention—is a question that demands ethical consideration. Simultaneously, the study of populism has offered educators some possibilities to consider for educational theory, practice and research. However, as populism itself, these possibilities are not straightforward: they demand contextualization and careful ethical debate.

References

- Apple, M.W., 2015. Understanding and interrupting hegemonic projects in education: learning from Stuart Hall. *Discourse* 36 (2), 171–184.
- Arditi, B., 2004. Populism as a spectre of democracy: a response to Canovan. *Polit. Stud.* 52 (1), 135–143.
- Boland, T., 2016. The Threat of Populism to Higher Education. *University Times*. Accessed at: <http://www.universitytimes.ie/2016/12/the-threat-of-populism-to-higher-education/>. (Accessed 6 August 2020).
- Bovens, M., Wille, A., 2017. *Diploma Democracy: The Rise of Political Meritocracy*. Oxford University Press.
- Brown, P., Lauder, H., Ashton, D., 2010. *The Global Auction: The Broken Promises of Education, Jobs, and Incomes*. Oxford University Press, Oxford.
- Canovan, M., 1999. Trust the people! Populism and the two faces of democracy. *Political Stud* 47 (1), 2–16.
- Canovan, M., 2004. Populism for political theorists? *J. Polit. Ideol.* 9 (3), 241–252.
- Carlsson, D., 2020. Representation and safe space: conflicting discourses in RE teacher education supervision. *Br. J. Relig. Educ.* 42 (1), 36–44.
- Cohen, J.L., 2019. Populism and the politics of resentment. *Jus Cogens* 1 (1), 5–39.
- Coughlan, S., 2016. Education System “Toughest to Reform”—Tony Blair. *BBC News*. Accessed at: <https://www.bbc.co.uk/news/education-35779235>. (Accessed 6 August 2020).
- De la Torre, C., 2017. Los populismos refundadores: promesas democratizadoras, prácticas autoritarias. *Nueva Sociedad* 267, 129.
- Dorling, D., Stuart, B., Stubbs, J., 2016. Don't mention this around the Christmas table: Brexit, inequality and the demographic divide. *LSE Eur. Polit. Policy Blog*. Retrieved from: <https://blogs.lse.ac.uk/eurompblog/2016/12/21/christmas-table-brexit-inequality-demographic-divide/>.
- Dyck, J.J., Pearson-Merkowitz, S., Coates, M., 2018. Primary distrust: political distrust and support for the insurgent candidacies of Donald Trump and Bernie Sanders in the 2016 primary. *Polit. Sci. Polit.* 51 (2), 351–357.
- Flinders, M., 2018. The (anti-) politics of the general election: funneling frustration in a divided democracy. *Parliam. Aff.* 71 (Suppl. 1), 222–236.
- Franzé, J., 2017. La trayectoria del discurso de Podemos: del antagonismo al agonismo. *Spanish J. Polit. Sci.* 44, 219–246.
- Gledhill, J., 2019. Capitalism, the politics of hate, and everyday cosmopolitanisms. *Critiq. Anthropol.* 39 (3), 310–328.
- Goodwin, M.J., Heath, O., 2016. The 2016 referendum, Brexit and the left behind: an aggregate-level analysis of the result. *Polit. Q.* 87 (3), 323–332.
- Hawkins, K.A., Dudley, R., Tan, W.J., 2017. Made in USA: populism beyond Europe. In: Martinelli, A. (Ed.), *Beyond Trump: Populism on the Rise*. Italian Institute for International Political Studies, Milan, Italy, pp. 93–110.

- Heggart, K.R., Flowers, R., 2019. Justice citizens, active citizenship, and critical pedagogy: reinvigorating citizenship education. *Democr. Educ.* 27 (1), 2.
- Hobolt, S.B., 2016. The Brexit vote: a divided nation, a divided continent. *J. Eur. Publ. Pol.* 23 (9), 1259–1277.
- Hoppers, C.A., 2001. Indigenous knowledge systems and academic institutions in South Africa. *Perspect. Educ.* 19 (1), 73–86.
- Inglehart, R., Norris, P., 2016. Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash. HKS Working Paper No. RWP16–026.
- Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P., 2017. Populism: an overview of the concept and the state of the art. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from: <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhdb/9780198803560.001.0001/oxfordhdb-9780198803560-e-34>.
- Kriesi, H., 2014. The populist challenge. *West Eur. Polit.* 37 (2), 361–378.
- Laclau, E., 2007. *On Populist Reason*. Verso, London.
- Lapping, C., 2005. Antagonism and overdetermination: the production of student positions in contrasting undergraduate disciplines and institutions in the United Kingdom. *Br. J. Sociol. Educ.* 26 (5), 657–671.
- Lowndes, D.T.J., 2018. Populism and race in the United States from George Wallace to Donald Trump. In: De La Torre, C. (Ed.), *Routledge Handbook of Global Populism*, pp. 204–214.
- Mannion, G., Biesta, G., Priestley, M., Ross, H., 2011. The global dimension in education and education for global citizenship: genealogy and critique. *Glob. Soc. Educ.* 9 (3–4), 443–456.
- Märth, A., Tryggvason, Á., 2017. Democratic education in the mode of populism. *Stud. Philos. Educ.* 36 (6), 601–613.
- Martinelli, A., 2016. Populism and the crisis of representative democracy. In: Martinelli, A. (Ed.), *Beyond Trump: Populism on the Rise*. Italian Institute for International Political Studies, Milan, Italy, pp. 13–32.
- Miller-Idriss, C., Pilkington, H., 2017. In search of the missing link: gender, education and the radical right. *Gend. Educ.* 29 (2), 123–146.
- Moffitt, B., 2017. Populism in Australia and New Zealand. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from: <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhdb/9780198803560.001.0001/oxfordhdb-9780198803560-e-5>.
- Moffitt, B., Tormey, S., 2014. Rethinking populism: politics, mediatization and political style. *Polit. Stud.* 62 (2), 381–397.
- Moncloa, L., 2017. Iñigo Méndez de Vigo Commits to Including Fundamental and Democratic EU Values in Education. Accessed at: https://www.lamoncloa.gob.es/lang/en/gobierno/news/Paginas/2017/20171121_educationvalues.aspx. (Accessed 6 August 2020).
- Mudde, C., Rovira Kaltwasser, C. (Eds.), 2012. *Populism in Europe and the Americas: Threat or Corrective for Democracy?*. Cambridge University Press.
- Mudde, C., Rovira Kaltwasser, C., 2017. *Populism: A Very Short Introduction*. Oxford University Press.
- Mudde, C., 2010. The populist radical right: a pathological normalcy. *West Eur. Polit.* 33 (6), 1167–1186.
- Müller, J.W., 2017. *What is Populism?* Penguin UK.
- Osler, A., Starkey, H., 2018. Extending the theory and practice of education for cosmopolitan citizenship. *Educ. Rev.* 70 (1), 31–40.
- Ostiguy, P., 2017. Populism: a socio-cultural approach. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from: <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhdb/9780198803560.001.0001/oxfordhdb-9780198803560-e-3>.
- Pais, A., Costa, M., 2020. An ideology critique of global citizenship education. *Crit. Stud. Educ.* 61 (1), 1–16.
- Panizza, F., 2008. Fisuras entre populismo y democracia en América Latina. *Stockholm Rev. Latin Am. Stud.* 3, 81–93.
- Pons, F., de Rosnay, M., Cuisinier, F., 2010. Cognition and Emotion. *Encyclopedia of Education*.
- Ramiro, L., Gomez, R., 2017. Radical-left populism during the great recession: podemos and its competition with the established radical left. *Polit. Stud.* 65 (1_Suppl. I), 108–126.
- Ramiro, L., 2016. Support for radical left parties in Western Europe: social background, ideology and political orientations. *Eur. Polit. Sci. Rev.* 8 (1), 1–23.
- Rao, S., 2018. Making of selfie nationalism: Narendra Modi, the paradigm shift to social media governance, and crisis of democracy. *J. Commun. Inq.* 42 (2), 166–183.
- Remmer, K.L., 2012. The rise of Leftist–Populist Governance in Latin America: the roots of electoral change. *Comp. Polit. Stud.* 45 (8), 947–972.
- Resnik, D., 2017. Populism in Africa. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from: <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhdb/9780198803560.001.0001/oxfordhdb-9780198803560-e-4>.
- Rovira Kaltwasser, C., 2012. The ambivalence of populism: threat and corrective for democracy. *Democratization* 19 (2), 184–208.
- Rutenberg, C.W., 2009. Educating political adversaries: Chantal Mouffe and radical democratic citizenship education. *Stud. Philos. Educ.* 28 (3), 269–281.
- Runciman, D., 2016. How the Education Gap is Tearing Politics Apart. *The Guardian*. Retrieved 22 May 2020 from: <https://www.theguardian.com/politics/2016/oct/05/trump-brexit-education-gap-tearing-politics-apart>.
- Runciman, D., 2018. *How Democracy Ends*. Profile Books.
- Sant, E., Brown, T., 2020. The fantasy of the populist disease and the educational cure. *Br. Educ. Res. J.* <https://bera-journals.onlinelibrary.wiley.com/doi/pdf/10.1002/berj.3666>.
- Sant, E., Hanley, C., 2018. Political assumptions underlying pedagogies of national education: the case of student teachers teaching “British values” in England. *Br. Educ. Res. J.* 44 (2), 319–337.
- Sant, E., McDonnell, J., Pashby, K., Menendez Alvarez-Hevia, D., 2021. Pedagogies of agonistic democracy and citizenship education. *Educ. Citizsh. Soc. Justice* 16 (3), 227–244.
- Sant, E., 2018. Global citizens, “not Mancunians or Athenians”. *Citizen. Teach. Learn.* 13 (3), 273–292. https://doi.org/10.1386/ctl.13.3.273_1.
- Sant, E., 2019a. Democratic education: a theoretical review (2006–2017). *Rev. Educ. Res.* 89 (5), 655–696.
- Sant, E., 2019b. History, myths and democratic education: secondary students’ discursive construction of Catalan nationhood. *Pedagog. Cult. Soc.* <https://www.tandfonline.com/doi/pdf/10.1080/14681366.2019.1703793>.
- Sant, E., 2021. *Political Education in Times of Populism. Towards a Radical Democratic Education*. Palgrave MacMillan/Springer Nature, Cham, Switzerland. Retrieved 08 August 2021. <https://link.springer.com/book/10.1007/978-3-030-76299-5>.
- Schaffner, B.F., MacWilliams, M., Nteta, T., 2018. Understanding white polarization in the 2016 vote for president: the sobering role of racism and sexism. *Polit. Sci. Q.* 133 (1), 9–34.
- Sloam, J., Henn, M., 2019. *Youthquake 2017: The Rise of Young Cosmopolitans in Britain*. Springer Nature, p. 129.
- Snir, I., 2017. Education and articulation: Laclau and Mouffe’s radical democracy in school. *Ethics Educ.* 12 (3), 351–363.
- Spruyt, B., Keppens, G., Van Droogenbroeck, F., 2016. Who supports populism and what attracts people to it? *Polit. Res. Q.* 69 (2), 335–346.
- Spruyt, B., 2014. An asymmetric group relation? An investigation into public perceptions of education-based groups and the support for populism. *Acta Polit.* 49 (2), 123–143.
- Stanley, B., 2017. Populism in central and Eastern Europe. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from: <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhdb/9780198803560.001.0001/oxfordhdb-9780198803560-e-6>.
- Stavrakakis, Y., Katsambekis, G., 2014. Left-wing populism in the European periphery: the case of SYRIZA. *J. Polit. Ideol.* 19 (2), 119–142.
- Szkudlarek, T., 2011. Semiotics of identity: politics and education. *Stud. Philos. Educ.* 30 (2), 113–125.
- Sztopka, P., 2016. The “demand side” of populism: the case of Poland. In: Martinelli, A. (Ed.), *Beyond Trump: Populism on the Rise*. Italian Institute for International Political Studies, Milan, Italy, pp. 73–92.
- Teney, C., Laceywell, O.P., De Wilde, P., 2014. Winners and losers of globalization in Europe: attitudes and ideologies. *Eur. Polit. Sci. Rev.* 6 (4), 575–595.

- Thiel, J.J., 2019. The Production of "Student Voice" and Its Effects on Academia. Doctoral dissertation, Manchester Metropolitan University.
- Trent, J., 2011. "Four years on, I'm ready to teach": teacher education and the construction of teacher identities. *Teach. Teach.* 17 (5), 529–543.
- Waisbord, S., 2018. The elective affinity between post-truth communication and populist politics. *Commun. Res. Pract.* 4 (1), 17–34.
- Weyland, K., 2017. Populism: a political-strategic approach. In: Kaltwasser, C., Taggart, P., Espejo, P., Ostiguy, P. (Eds.), *The Oxford Handbook of Populism*. Oxford University Press. Retrieved 22 May. 2020, from <https://www-oxfordhandbooks-com.manchester.idm.oclc.org/view/10.1093/oxfordhb/9780198803560.001.0001/oxfordhb-9780198803560-e-2>.
- Zembylas, M., 2020. The affective modes of right-wing populism: Trump pedagogy and lessons for democratic education. *Stud. Philos. Educ.* 39 (2), 151–166.

The China challenge; China's challenges

Anthony Welch, School of Education and Social Work, University of Sydney, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

The China challenge	658
Genesis and development	658
Science and innovation	659
Western incursions	659
An enduring legacy: from Republic to People's Republic	660
China's challenges	661
Reform and opening era	661
Rural-urban differences	661
Private education	662
Migrant children	662
Guanxi	662
Corruption	662
US-China tensions	663
Philosophical-practical contradictions	663
References	664

The China challenge

As with several other Asian systems, China's education system has ancient roots and modern branches. Its rich legacy of past traditions haunts its present, as it strives to strengthen education, reduce inequalities, develop world class universities, and challenge the best of the West. At the same time, while spectacular progress has been made in both quantitative and qualitative terms, it confronts several key challenges that arguably inhibit its further development.

Genesis and development

While the emphasis in what follows is on the modern system of education, it is not possible to understand key contemporary developments without seeing these in a longer historical perspective. China's rich history of ideas, embracing intersecting frameworks such as Buddhism, Daoism, and Confucianism, and major figures such as Lao Tze (老子) [571-501? BCE], Confucius (孔夫子) [551-479 BCE], and Mencius [372-289 BCE] (孟子), helped forge what can be called the Chinese model of education. First developed several centuries BCE, during the Zhou dynasty, it was finally formalized during the Tang dynasty (618–907). Centered on a Confucian worldview, it stressed self-cultivation at the individual level, and active ethical engagement and governance at the societal level, encapsulated in the idiom “sageliness within and kingliness without” (*nei sheng wai wang*, 内圣外王) (Zha, 2022; see also Yang, 2020). The model, based on Confucian ethics and a curriculum based on the *Four Books* and *Five Classics*, was widely influential in East Asia, including Japan (briefly), Korea, as well as Southeast Asia's Singapore, and Việt Nam where Confucius was celebrated as the teacher of a thousand sovereigns. Persisting in the region until the early 20th century, elements survived the *anti-Confucius* campaign of the 1970s, while, more recently, elements of Confucian thought have again been pressed into service in education. The Confucian tradition in education continues to exert a powerful influence on the contemporary world of Chinese education governance (Yang, 2020).

China's institutional legacy is just as strong, beginning with the first Asian institution of higher learning, the *Jixia Academy*, established in 360 BCE, during the Warring States period of what is now China. Its teachers included Mencius (孟子), Hsun Tzu (孫子) [c. 310-220 BCE], Zou Yan (鄒衍) [305-240 BCE] and Shen Dao (慎到) [350-275 BCE]. The *keju*, or imperial examination system was another important innovation, that opened the door to official positions, to students who survived the rigorous, formal set of written examinations. The examinations were conducted at three levels—district, province and national—and quite a number fell at the first hurdle. The highest degree, known as *Jinshi* (進士) was required for the highest office, whereas the basic degree of the *Shengyuan* (生員) was insufficient to secure a post, because it was so oversupplied. The Confucian institutional model, including the core principle regarding the value of knowledge and self-cultivation, became an important base of regional higher education (Hayhoe, 2013). Versions of the system spread to parts of both East and Southeast Asia (Welch, 2019). Residues remain in Singapore (Tan, 2012). The *Taixue* (太学) sometimes called the Imperial Academy, was founded in the Han dynasty (206 BCE-220 AD), when it had as many as 30,000 students and teachers. Renamed the *Guozijian* (國子) during the Sui dynasty, it was considered the pinnacle of educational establishments. Based on a curriculum of Confucianism and Chinese literature, it supplied well educated, senior administrators for the imperial government, and endured until 1905.

Two further problems weakened the system. Wealthy merchant families could educate their sons, thereby gaining preferment at the expense of poorer families. Secondly, corruption ensured that in some cases it was also possible to purchase degrees. Becoming even more formal and increasingly orthodox during the Ming dynasty (1368–1644), the model survived until the end of the sclerotic Qing dynasty at the beginning of the 20th century. By then, however, the system was being openly criticized. On the one hand, the formal knowledge taught was narrow and exclusive: its traditional curriculum ossified, and eschewed the study of modern subjects such as science, technology and mathematics that in the West were seen as a key pillar of modernity. In China, the civil service examination system that persisted until 1905 was increasingly seen as stifling national scientific and technological advancement, and at least in part responsible for China's humiliating failure to resist Western imperialist incursions and plunder in the 19th century (Ironically, however, the modern Chinese schooling system of the time, established before the May Fourth movement, was in large part based on emulating the foreign systems, notably German and Japanese). On the other hand, it was also inefficient: some 2 or 3 million applicants sat the exams each year, with a failure rate of around 99%.

A second, less well-known institutional innovation was seen by some as “the highest institution of learning” (Ding, 1996, 231). It had greater, if fragile, autonomy, than the *Keju*, according to Hayhoe. Originating during the Tang dynasty, the *Shuyuan*, (书院) or academy, was a non-state institution of learning, comprising a community of scholars and teachers focused on self-study and research. *Shuyuan* were often sited in quiet rural settings, to avoid worldly distractions. Famous examples included the Bailudong (White Deer Grotto), Yuelu, and Taishi Academies, versions of which spread to Korea where they were known as *Seowon*. They endured until the dying days of the Qing dynasty, in 1898, and influenced 20th century reformers such as Cai, Yuanpei and Hu Shi, as well as Mao Ze Dong (Ding, 1996).

For the masses, village schools of varying quality were not uncommon. Poorer communities struggled to sustain such institutions. Poorer families were also less likely to educate their daughters than sons, based on the Chinese convention which, upon marriage, transferred a daughter to the control of the husband's family. Hence, to educate a daughter was tantamount to “weeding the field of another man.” Rote learning was ubiquitous, based on texts such as the *Sanji Jing* (Three Character Classic), followed by the *Qianzi Wen* (Thousand Character Classic).

Science and innovation

Chinese science and discovery also flourished, including the four great inventions—papermaking, printing, gunpowder, and the compass—that represent major contributions to world civilization. These inventions were only a part of China's rich array of scientific and technological inventions, however (USC, n.d.). It is no surprise that *Zhongguo* (the Middle Kingdom) saw itself as the center of the world, with, at various times, little interest in the world outside (Kwong, 2015). Ancient maps tended to portray China at the center, with other categories radiating outwards toward the outermost “barbarians.” (Jacques, 2012)

But by the onset of what is broadly known in the West as the Scientific Revolution, from around the 16th century, the flowering of Chinese inventiveness had largely waned. Its education system, too, which had been formalized in the Tang dynasty, had begun to ossify, at much the same time as China began to need to negotiate the crossroads of modernity (Kwong, 2015).

Western incursions

The final decades of the Qing dynasty (1642–1911) included a longstanding debate about the value of Western knowledge, that, while evident as early as the long stay of foreigners such as Matteo Ricci (1552–1610), continues in some form to this day. To some extent, it hinged around different stances regarding the Chinese aphorism *Zhōngxué Wéi Tǐ, Xīxué Wéi Yòng* (中学为 体, 西学为用) that can be broadly translated as Chinese learning for fundamental principles; Western learning for practical application.

The forced opening (and plundering) of China by Western imperial powers from the time of the Opium Wars of the 1840s, not only convinced many in China that the traditional form of education, among other elements of Chinese society, was no longer adequate. It also led to new institutional forms, since part of the nearly 200 unequal treaties imposed on China at the time included freedom for Western missions. Prior to the Opium Wars, foreigners were banned from missionary activities. The American magazine *Puck* satirized the missionary zeal in a pointed cartoon of 1895. Showing the interchange of missionaries and Chinese authorities as occurring under the looming cannons of Western military forces, its caption read “According to the ideas of our missionary maniacs, the Chinaman must be converted, even if it takes the whole military and naval forces of the two greatest nations of the world to do it” (cited Welch, 2014b).

In blissful ignorance of centuries of Catholic activity in China, Protestant missionary schools were established, with the enthusiastic support of networks in the US, who supplied staff, money, and Chinese-language publications (sometimes in local dialects, rather than imperial Mandarin). Beyond their particular missionary remit, they had wider educational aims: to translate foreign texts, help introduce new ideas in mathematics, science, and engineering, and develop institutional innovations such as schools for the blind and industrial schools. Just thirteen American colleges were established, (although their influence was broader than their mere number), while an 1877 report listed 30 boys' and 38 girls' boarding schools, enrolling around 600 boys, and approximately 800 girls (totaling 1400 pupils in a population of 400 million). In addition, there were 82 girls' day schools and 177 boys' day-schools (Welch, 2014b: 11, citing Du Bose, 1879). The British Church Missionary Society in Fujian Province founded over a hundred village day schools (Welch, 2014b: 22) While for missionaries, the prime purpose was evangelism, local pupils often saw it differently. Those who learned English could often make money by interpreting, or engaging in trade with foreigners

(sometimes pilloried as 鬼佬 *Gweilao*, or foreign devils). Chinese families, too, saw such schooling as a means of personal and familial advancement. The missionaries' evangelical aspiration remained largely unfulfilled.

The end of the 19th century also saw the establishment of the country's first Western style universities. Tianjin University (天津大学) (formerly Peiyang University) (1895) and Peking University (北京大学) (1898) were prominent early examples (Hayhoe, 1996). Each has already celebrated their 120th anniversary. Each also revealed foreign influences in their early development. At Peking University, Cai Yuanpei's experience of both the German and French models of higher education led him to introduce forms of faculty governance, democratic management, and principles of freedom of thought. Tianjin university was more influenced by US higher education, via the work of the educator and diplomat Charles D. Tenney, Peiyang's first President. Its early graduates could enroll directly into either Harvard or Yale.

An enduring legacy: from Republic to People's Republic

While the rigid formal mode of education that was associated with the old imperial system fell into abeyance, it nonetheless left an enduring legacy of respect for learning and scholars, as even Western educators still acknowledged in the early 20th century: "Study in China is considered a vocation, the maxim being common among the people that 'Study is the highest pursuit a man can follow'" (cited Welch, 2014a: 25).

Inspired by democratic principles, Republican China (1912–1949) saw some significant initiatives introduced. This included reforms to schooling, the practice of which was seen as productive of a modernizing democratic polity. Not all reforms, however, came to fruition. The ambitious School Reform Decree of 1922, for example, came up against a society that was "intensely undemocratic to begin with" (Keenan, 1974: 225).

Western educational influences during the Nationalist government were repudiated after the victory of Communist forces, and proclamation of the Peoples Republic of China (PRC) in 1949. Gu usefully divides education in the PRC into three stages, with distinct emphases: Learning from the USSR (1949–1958); Leftist education (1958–1978); and Reform and Opening (1978–) The second stage, he argues, should be further sub-divided into three periods: 1958–1966, comprising the education revolution, 3-year adjustment and class struggle; 1966–1976, resulting in the educational ravages of the Cultural Revolution; and 1976 to 1978, called "bringing order out of chaos" (Gu, 2015: 195; Welch, 2019).

The early years of the People's Republic, from 1949, marked major changes in education, not least through a stress on public education. Many private schools and universities were simply closed or transformed into public institutions, a major shift given that private higher education institutions (HEIs), for example, had comprised 39% of overall enrollments. Nankai and Xiamen universities were prominent examples of transformation into public HEIs, a process which had been completed by 1956. The scale of the national challenge was illustrated by a 1951 conference devoted to Engineering education, that identified the following major issues:

- imbalanced geographical distribution;
- dispersion of teachers and facilities leading to uneconomic practices;
- impractical specialties that failed to train specialized personnel;
- and a shortage of students (Yang, 2000: 321).

Given the huge scale of needed reforms, and the background of widespread illiteracy, the Cold War era, ravages inflicted by the Nationalist-Communist Civil War, and China's lack of much international support, it was no surprise that the first few years (from 1949) were characterized by the dependence on institutional models, translated textbooks, forms of pedagogy, and some teachers and "advisors," from the much-more-established Union of Soviet Socialist Republics (USSR). In China this era was sometimes characterized as "Learn from Big Brother." Russian replaced English as the major foreign language, "spare-time" schools and other special schools were instituted for peasants and workers, including their families, and higher education was reorganized on Soviet lines. International scholarly networks, too, were bounded by the USSR and its eastern-European satellites:

study abroad—unless undertaken in the then Soviet Union or Eastern Europe—was generally regarded as politically and culturally suspicious. Those who had studied in the West ... were denounced as "running dogs" of imperialism and traitors.

Rhoads (2011: 6).

The major ideological break with the USSR from the mid-late 1950s presaged major changes, however, including in education. This included a much more discriminating attitude toward Soviet educational theory and practice, adopting only what was found to be useful in the Chinese context. Open criticisms of Soviet errors by some Chinese students in the then Union of Soviet Socialist Republics (USSR) led to some being expelled in 1963 for spreading ideologically incorrect views (You, n.d.) More broadly, an emphasis on ideology grew in education. Workers and peasants were admitted to spare time schools based on their political orthodoxy, rather than formal qualifications.

The emphasis on ideology only intensified during the Cultural Revolution (1966–76). Education reform was seen as a catalyst for cultural revolution and class struggle (Pepper, 1996: 352 et seq.). Educators and intellectuals were sent down to the countryside,

to learn from peasants. Many endured years of painful privation; some died or committed suicide (Ai, 2021; Welch, 2019). Universities were closed, and a generation lost the opportunity for further study. Not until 1976 did universities open again, and by then it was too late for many. Not for nothing did Gu characterize the immediate aftermath of the Cultural Revolution as “bringing order out of chaos” (Gu, 2015: 195).

China's challenges

China's rich history of respect for learning, together with influential pedagogical and institutional models created a sound educational basis, enduring for millennia and emulated regionally. Even after institutions such as the *Shuyuan* and the imperial examination system were abandoned at the end of the Qing dynasty, a legacy of respect for learning and education persists broadly, including among the worldwide Chinese diaspora, where students of Chinese background often outperform locals (Mullis et al., 2008; Zhao and Singh, 2011). International achievement tests, such as the OECD's Program of International Student Achievement (PISA) further confirm high levels of achievement among the Chinese student cohort (OECD, 2018). Chinese teachers of mathematics are now sought internationally, and leading Chinese universities now take their place among the world's best (Yuan and Huang, 2019; ARWU, 2021). China is now the world's largest source of research publications, its leading scientists now routinely figure among the world's highly cited scholars, and it has now become a major knowledge partner of a range of countries, both within its region and more broadly (SupChina, 2022; Cheng, 2021; Welch, 2014c, 2021b). It was also (pre-COVID) one of the world's leading destinations for international students, with 500,000 enrollments, and its large and diverse worldwide knowledge diaspora is making substantial contributions to research, both in China and many other parts of the world (Yang and Welch, 2010; Welch and Hao, 2016). China's universities, particularly those in its border regions, have forged important relations with ASEAN neighbors, and China has been generous with scholarships both ASEAN scholars, and to the developing world, to study at its universities (Wen and Hu, 2016; Welch, 2014c, 2018).

Together these achievements make for a remarkably impressive achievement. And China's restless ambition seeks to achieve more, including goals to move to an innovation-based knowledge society, with world-leading universities and research. At the same time, however, the country confronts a range of challenges that impinge on these lofty ambitions.

Reform and opening era

The decade long Cultural Revolution had laid waste to China's education system. The Third Plenary Session of the Eleventh Central Committee of the CPC (1978) faced the urgent challenge of reconstructing and improving education. The Cultural revolution's legacy was catastrophic: a shortage of teachers, as well as the destruction of institutions and scholarly resources of all kinds. Regulations governing educational institutions were revised and re-established, and order restored. Restoring the national tertiary education entrance exam (*Gaokao*) symbolized restoration and reconstruction. Educational funding was set to rise, as were education costs borne by families—the latter gradually. Some decentralization was introduced, and the doctrine of principal accountability (*Xiaozhang Fuzezhi* 校长负责制) introduced. More emphasis was given to market principles from the 1990s, with the aim of developing a socialist market economy with Chinese characteristics.

Paralleling China's monumental economic achievements, that saw its economy grow by an average of 10% in the decades following 1978, educational achievements were also impressive. From the 1980s, adult literacy rates rose from 65% to the current 96%. Over the same period, the university participation rate by high school graduates rose from 20% to closer to 60%.

But such success led to other difficulties. China now produces more than 11 million university graduates annually, who struggle for advantage within China's highly competitive job market. Graduate unemployment and underemployment has become a significant issue, especially for those from second tier, and regional universities. Such students fare less well, and Lian Si's so-called “Ant Tribe” (*Zhuyin Fuhao*, or 蚁族) phenomenon entered popular discourse from 2009. It described a situation where, facing unemployment or low wage and part-time employment, impoverished and either unemployed or underemployed students, often from rural or second-tier universities, crowded together in meager living spaces in China's major cities. “They share every similarity with ants. They live in colonies in cramped areas. They're intelligent and hardworking, yet anonymous and underpaid” (China Daily, 2010).

Rural-urban differences

Forms of inequality in education begin much earlier, with rural children often taught by less qualified teachers, and facing less well-equipped schools. Their chances of gaining access to one of China's forty or so elite 985—now known as Double First Class or *shuangyiliu* (双一流)—universities (mostly located in China's major cities) are much lower than children who have grown up in such cities and attended a well-equipped and well-staffed school. Rural-urban inequalities have tended to widen over recent decades, despite efforts to reduce the difference. According to one survey, relatively high dropout rates from primary and Junior High Schools meant that a mere 6% of rural children entered Senior High Schools, only half of whom graduated. Two percent obtained access to tertiary education. For urban students, the respective rates were 63%, 63% and 54% (Zhang et al., 2015: 200).

Some forms of decentralization were introduced into education, but this too, had some problematic outcomes. In practice, decentralized operation and management (*Fenji Banxue Fenji Guanli* 分级办学 分级管理) meant that funding compulsory schooling

became the responsibility of township-level governments, thus widening the gap between richer and poorer communities, and rural and urban settings.

Private education

After pilot programs in 1993, some private (*minban*) schools were introduced in 1994, although in accordance with socialist ideology the term “civilian run and public subsidized (*minban gongzhu* 民办公助)” was preferred to “private.” There are in addition hundreds of *duli xueyan* (privately run, “second-tier” colleges, only able to offer degrees via their major “mother” public university). In practice, these commonly comprise a partnership between a major public HEI and a private entrepreneur, for whom developing and selling housing around the HEI was likely to be much more lucrative than income from student fees. Tensions developed between public HEI partners more interested in quality, and private entrepreneurs more interested in profits. Examples of corruption touched recruitment, fees and financial administration (UWN, 2014; Welch, 2020). Private enrollment at the tertiary level now accounts for around 13.5% of the total (5,871,139 of a total enrollment of 43,367,394 (PROPHE, 2020: 12; Welch, 2021a), but concerns regarding the quality of *minban* reflect widespread community understanding that this is second-choice higher education. Although fees can be quite high, no such institution comes close to achieving the status or achievements of China’s major public universities, and concerns have been voiced about the impact of privatization: “privatization favors only the elite and wealthy families who are privileged to start with because of their social status and financial capability” (Zha, 2006: 63).

Migrant children

A further form of differentiation is seen in the schooling afforded to migrant children. Wholesale migration of migrant laborers from rural areas to China’s major cities in Eastern provinces (164 million rural workers migrated to China’s cities between 1990 and 2010), led to a growing pool of “left behind” children (*liúshǒu’értóng* or 留守儿童), numbering around 69 million (mainly situated in China’s South and Central regions). This is despite major nationwide poverty reduction measures. Gradual amelioration of some of the more restrictive elements of the household registration (*hukou* 户口) system, particularly after the 1996 *Measures for the Education of School-age Children and Adolescents among the Urban Migrant Population (Trial)* allowed more migrant children with rural *hukou*, who had accompanied their migrant laborer parents into China’s big cities, to gain access to schooling. Different cities adopted varying strategies, some relying more on public schools, while others relied more on the private sector. While 9 years of schooling is free and mandatory, some research shows that 36 million migrant children still do not have access to public schools, while the quality of informal schools that sprang up to cater to migrant children in China’s major cities is variable, and not always well regulated (Liu, 2019).

Guanxi

Guanxi (关系) describes the network of social relations that largely govern Chinese lives. Not entirely dissimilar to the Western concept of social capital, it encompasses the mutually beneficial system of trusted relationships that are depended on by individuals to live by and advance one’s career or interests. Effectively, personal bonds must first be established before an effective reciprocal relationship can be inaugurated. To some extent, such relations are grounded in Confucian assumptions about familial and hierarchical social relations, reciprocity and trust and the concept of face (*面子*, *miànzi*). While *guanxi* was set originally within a set of ethical principles, its form is at times abused to operate in unethical ways. It can be used to extract social or economic advantage in social, political or business environments, such that interpersonal obligations take precedence over civic duties. In education, it has been associated with non-merit-based awards of research grants, and promotions, for example (Welch, 2020). In education, for example, “... if someone has a good *guanxi* with the official in charge, the official could interpret the rules in the way that is favorable for that person” (Guan, 2010: 2). This can lead to exploitation of younger staff, often female. And in relation to research, it “encourages researchers not to devote their time on research but on schmoozing with those in power. In some cases, bribery and corruption also take place” (Qiu, 2014: 162).

Corruption

Corruption exists in most societies, to greater or lesser degree. But, like numerous other Asian nations, China does not score well on measures of corruption, defined by Transparency International (TI), as the abuse of entrusted power for private gain, either material or personal. In education, this embraces such practices as appointments or promotions gained through nepotism or cronyism, paying a bribe to secure a place in a private university, or an individual pocketing funds to which they were not entitled. In higher education, for example, it includes the issuing of fraudulent certificates (for payment); inflated or illegitimate procurement contracts, (for new university infrastructure, or building maintenance); accepting or paying bribes or exploiting students for sex in return for entry into university, passing a course, or granting a degree; offering or accepting payments for higher grades; levying private payments for additional tuition, without which the student is unable to pass the exam; claiming fake qualifications; engaging in cronyism/nepotism; pursuing plagiarism; falsifying research results; passing off student research as that of the professor, misusing research funds; or paying bribes to secure academic posts, get promotion, or gain program or institutional accreditation

(Welch, 2020). Such forms of academic corruption all occur in China; they weaken higher education and erode trust, both in the institutions in which they occur, and more broadly in society.

US-China tensions

Increasingly rivalrous and rancorous US-China relations in recent years, arose from the so-called Trade War, that soon came to be increasingly acknowledged as a Technology War, Culture War, and even a Cold War. The impact on education was seen starkly in moves toward de-coupling of international research networks, and restrictions on international staff and student mobility. Collectively, these curtailments were said to erode national and international innovation. Motivated by techno-nationalism, and a fear that China might out-innovate the US, America moved to deny visas to 3000–4000 Chinese researchers and post-graduates in key high-technology areas (robotics, new materials, artificial intelligence, quantum information, neuroscience), revoked the visas of a further 1000 researchers and students, and prosecuted ten Chinese-American scientists (Zhu, 2021).

Pressure by the US on its allies to align themselves led to increasing pressure on bi-lateral research collaboration with China, the securitization of related policy arenas, rising *anti-Chinese* sentiments and attacks, and heightened nervousness and concern among thousands of Chinese researchers in a range of countries, who feared they were being singled out (Raby, 2020; Brophy, 2021; Welch, 2021b). Japan, Germany, Canada and Australia each introduced a range of measures to audit and curtail China research relations. The passage of the bi-partisan *US Competition and Innovation Act*, by the US Senate in mid 2021 directly aimed to counter China's technological rise. This was compounded by a White House decision to increase the number of Chinese firms subject to a black ban on US investment, that in turn, followed earlier actions to ban or restrict Chinese investment in US technology firms and encourage US firms to move production of key tech components out of China (Evans, 2020; Franck, 2021a,b). Staff and student mobility were further reduced by a range of travel restrictions introduced by China and many other countries from 2020 in response to the COVID pandemic. In a number of Chinese universities, international student enrollments fell to a small fraction of pre-COVID numbers.

Philosophical-practical contradictions

To some extent, contemporary contradictions in Chinese education can be said to stem from tensions between two co-existing traditions, each of which is deeply rooted in Chinese history: "the Confucian core notion of eternal self-perfection and spiritual freedom, and a politicized and instrumental Chinese higher education in reality" (Zha, 2022, see also Yang, 2020). These tensions were sharpened after 2012, through the increasing re-assertion of the Communist Party of China's (CPC's) leading role, including in many facets of education. In a series of guidelines, such as "Opinions concerning Further Strengthening and Improving Propaganda and Ideology Work in Higher Education Under New Circumstances" (2015), educational institutions were instructed to hew more to the party line and eschew Western ideas and textbooks. Additional measures were successively instituted to enhance ideological commitment (such as regular meetings on ideology, especially "Xi Jinping thought," for which teachers and academics needed to submit regular "reports," and enhanced tests of political orthodoxy introduced into appointments and promotions procedures). While the technological and natural sciences were less affected, the impact on the social sciences and humanities was greater. According to one analyst, the challenge is to "to uproot Chinese academia from a culturally nepotistic and administrative society, where academic ethics and quality can yield to non-academic factors and the power of 'circles'" (Xu, 2021; 487).

But the moves to re-assert Party leadership were somewhat at odds with performance pressures to publish more internationally, especially in leading English language journals. Such demands were at their most intense at leading Chinese universities where young staff, often returnees with their Ph.D. from a leading western university, were expected to produce a minimum number of such articles, in order to maintain their position and promotion momentum. Older staff, usually male, who had worked all their lives in the Chinese system and had limited English language facility, were exempt, (but would likely be replaced by younger, bilingual staff upon retirement (Welch, 2021c).

In a sense, such pressures represented a new form of the age-old dilemma *Zhōngxué Wéi Tǐ, Xīxué Wéi Yòng* (中学为体, 西学为用). As the Beirut-based sociologist Sari Hanafi explains, the old maxim of Publish or Perish now exists in a new form, especially for academics working in non-English systems: Publish Globally and Perish Locally; or Publish Locally and Perish Globally. In practice, younger Chinese scholars are under pressure to do both. On the one hand international publications are needed, both to gain promotion internally as well as to establish their international profile. But as one young female at a double first class (*shangliyi* 双一流) university pointed out, "If I only publish in English, no-one in China would know who I am" (Welch, 2021c: 378). The intellectual leaders in her field—largely older male academics—mostly published only in Chinese and would not know of her international publications. In effect, she needed to respond to the dual pressures of producing publications in leading international journals, as well as major domestic journals in Chinese. Some Chinese scholars have expressed concerns over the potential for "self-colonization" (Dang, 2005: 68, cited Xu et al., 2021)—due to the over-emphasis on international publications, Chinese HSS (Humanities and Social Science) academics may "accept Western concepts and theoretical frameworks without critical scrutiny and creative thinking" (Deng, 2010: 182), and focus more on Western issues or selectively on Chinese issues of interest to Western countries (Zhu, 2009, cited Xu et al., 2021: 471). In a response to such bias toward Western publications, three major universities, Renmin, Nanjing and Lanzhou withdrew participation in international rankings in 2022, after a speech by President Xi at Renmin, that argued that China must "blaze a new path to developing world-class universities based on Chinese conditions and with Chinese characteristics" (UWN, 2022).

China's current dramatic rise in education and research merits the term renaissance, given the importance of its original model of education. Its quantitative and qualitative achievements are nothing short of extraordinary. It remains to be seen whether further progress ensues at the same pace, or whether ideology limits such development, as it has done on occasion before. A synthesis of Eastern and Western ideas and institutions, incorporating the best of the Chinese tradition, holds great promise for the world, however elusive thus far.

References

- Academic Ranking of World Universities [ARWU], 2021. 2021 Academic Ranking of World Universities. <https://www.shanghairanking.com/rankings/arwu/2021>.
- Ai, W., 2021. 1000 Years of Joys and Sorrows. Bodley Head, London.
- Brophy, D., 2021. China Panic. Australia's Alternative to Paranoia and Pandering. LaTrobe University Press, Melbourne.
- Cheng, Y., 2021. The US is Building Walls Around Science, and We're All Poorer for It. <https://www.vice.com/en/article/xgzbqg/us-china-relations-science-rivalry-gang-chen-mit>.
- China Daily, 2010. China's "Ant Tribe" Still Struggling. December 14.
- Dang, S., 2005. Can American standards set the highest evaluation benchmark for Chinese Social Sciences?—take SSCI as an example [Meiguo biao zhun neng chengwei Zhongguo Renwenshe chengguo de zuigao biao zhun ma?—Yi SSCI weil]. Soc. Sci. Forum 4, 62–72.
- Deng, Z., 2010. Westernization of the Chinese Social Sciences: the case of legal science (1978–2008). In: UNESCO, International Social Science Council (Eds.), World Social Science Report 2010: Knowledge Divides. UNESCO & International Social Science Council, Paris, pp. 182–183.
- Ding, G., 1996. The Shuyuan and the development of Chinese universities in the early twentieth century. In: Hayhoe, R., Pan, J. (Eds.), East West Dialogue in Knowledge and Higher Education. Routledge, London, pp. 218–244.
- Evans, P., 2020. Techno-nationalism in China-US relations: implications for universities. East Asian Pol. 12 (2), 80–92.
- Franck, T., 2021a. Democrats, GOP Team Up on Bill Targeting China as U.S. Suffers Microchip Shortage. CNBC. <https://www.cnbc.com/2021/05/25/chip-shortage-democrats-gop-team-up-to-target-china.html>.
- Franck, T., 2021b. Biden Prohibits U.S. Investment in 59 Chinese Companies Allegedly Tied to Military, Surveillance. <https://www.cnbc.com/2021/06/03/biden-prohibits-us-investment-in-59-chinese-companies.html>.
- Gu, M., 2015. Influences on Chinese education exerted by Soviet education. Cultural Foundations of Chinese Education. Brill, Leiden, pp. 195–216.
- Guan, J., 2010. Guanxi: the key to achieving success in China. Sino-Platonic Papers 217, 1–11.
- Hayhoe, R., 1996. China's Universities 1895–1995 A Century of Cultural Conflict. Garland Publishers, New York/London.
- Hayhoe, R., 2013. Hopes for Confucian pedagogy in China? J. Curric. Stud. 46 (4), 313–319.
- Jacques, M., 2012. When China Rules the World: The Rise of the Middle Kingdom and the End of the Western World. Penguin.
- Keenan, B., 1974. Educational reform and politics in early Republican China. J. Asian Stud. 33 (2), 225–237.
- Kwong, L., 2015. What's in a name: Zhongguo (or "middle kingdom") reconsidered. Hist. J. 58 (3), 781–804.
- Liu, M., 2019. Chinese migrant children: do they have the right to education? In: Rehman, J., Shadid, A., Foster, S. (Eds.), The Asian Yearbook of Human Rights and Humanitarian Law, vol. 3. Brill, Leiden, pp. 303–330. Law, Gender and Sexuality.
- Mullis, I., Martin, M., Foy, P., 2008. TIMSS 2007 International Mathematics Report: Findings From IEA's International Mathematics and Science Study at the Fourth and Eighth Grades. Centre for the Study of Testing, Evaluation and Educational Policy, Boston College, Massachusetts.
- Organisation for Economic Cooperation and Development [OECD], 2018. PISA Results From 2018 Country Note. Beijing, Shanghai, Jiangsu and Zhejiang. https://www.oecd.org/pisa/publications/PISA2018_CN_QCI.pdf.
- Pepper, S., 1996. Radicalism and Education Reform in Twentieth Century China. Cambridge University Press, Cambridge.
- Program on Private Higher Education [PROPHE], 2020. How Covid-19 Puts Private Higher Education at Especially High Risk—and Not: Early Observations Plus Propositions for Ongoing Global Exploration. PROPHE Working Paper 22. https://prophe.org/cache/2254169_WP22-PROPHE-COVID.pdf.
- Qiu, J., 2014. China's funding system and research innovation. Natl. Sci. Rev. 1 (1), 161–163.
- Raby, G., 2020. China's Grand Strategy and Australia's Future in the New Global Order. Melbourne University Press, Melbourne.
- Rhoads, E., 2011. Stepping Forth Into the World. The Chinese Educational Mission to the United States 1872–1881. Hong Kong University Press, Hong Kong.
- SupChina, 2022. The Sad State of US-China Scientific Collaboration. <https://supchina.com/2022/05/12/the-sad-state-of-u-s-china-scientific-collaboration/>.
- Tan, C., 2012. "Our shared values" in Singapore: a Confucian perspective. Educ. Theor. 62 (4), 449–463.
- University of Southern California [USC], (n.d.). Top 20 Ancient Chinese Inventions. <https://china.usc.edu/sites/default/files/forums/Chinese%20Inventions.pdf>.
- University World News, 2014. Independent Colleges—a Hybrid Response to Massification. <https://www.universityworldnews.com/post.php?story=20141106074227990>.
- University World News, 2022. Three Major Universities Quit International Rankings. <https://www.universityworldnews.com/post.php?story=20220511170923665#:~:text=The%20three%20prestigious%20universities%2C%20Renmin,autonomy%E2%80%9D%20and%20%E2%80%9Ceducation%20with%20Chinese>.
- Welch, A., 2014a. China-ASEAN relations in higher education: an analytical framework. In: Yonezawa, Y.K., Meerman, A., Kuroda, K. (Eds.), Emerging International Dimensions in East Asian Higher Education. Springer, Dordrecht, pp. 103–120.
- Welch, I., 2014b. Culture and Protestant Mission Schools in Nineteenth Century China. Working Paper, ANU. <https://openresearch-repository.anu.edu.au/bitstream/1885/11462/1/Welch%20Culture%20and%20mission%20schools%202014.pdf>.
- Welch, A., 2014c. Richer relations? Four decades of ASEAN-Australia relations in higher education. In: Wood, S., He, B. (Eds.), The Australia-ASEAN Dialogue: Tracing 40 Years of Partnership. Palgrave Macmillan, New York, pp. 145–166.
- Welch, A., 2018. China's Southern Borderlands and ASEAN higher education; a cartography of connectivity. In: Meusburger, P., Heffernan, M., Suarsana, L. (Eds.), Geographies of the University. Springer, Dordrecht, pp. 567–602.
- Welch, A., 2019. Higher education in Asia. In: Rury, J., Tamara, E. (Eds.), The Oxford Handbook of the History of Education. Oxford University Press, Oxford.
- Welch, A., 2020. Governance and corruption in East and South-East Asian higher education. In: Capano, G., Jarvis, D. (Eds.), Convergence and Diversity in the Governance of Higher Education. Cambridge University Press, Cambridge, pp. 355–395.
- Welch, A., 2021a. Private higher education in East and South-East Asia: growth, challenges, implications, UNESCO global monitoring of education report 2021. In: Non-state Actors in Education. Who Chooses, Who Loses? UNESCO, Paris.
- Welch, A., 2021b. A plague on higher education: Camus, COVID and Culture Wars in Australian. High. Educ. Q. <https://doi.org/10.1111/hequ.12377>.
- Welch, A., 2021c. In league? The brave new world of higher education. In: Welch, A., Li, J. (Eds.), Measuring Up. How University Rankings and League Tables are Re-shaping Knowledge Production in the Global Era. Palgrave Macmillan, London, pp. 375–386.
- Welch, A., Hao, J., 2016. Global argonauts: returnees and diaspora as sources of innovation in China and Israel. Glob. Soc. Educ. 14 (2), 272–297.
- Wen, W., Hu, D., 2016. China's approach to HE regional cooperation with ASEAN. In: Hawkins, J., Neubauer, D. (Eds.), Palgrave Handbook of Asia Pacific Higher Education. Palgrave Macmillan, London, pp. 173–182.
- Xu, X., Oancea, A., Rose, H., 2021. The impacts of incentives for international publications on research cultures in Chinese humanities and social sciences. Minerva 59, 469–492.

- Yang, R., 2000. Tensions between the global and the local: a comparative illustration of the reorganisation of China's higher education in the 1950s and 1990s. *High. Educ.* 39, 319–337.
- Yang, R., 2020. Political culture and higher education governance in Chinese societies: some reflections. *Front. Educ. China* 15 (2), 187–221.
- Yang, A., Welch, A., 2010. Globalisation, transnational academic mobility and the Chinese knowledge diaspora: an Australian case study. *Discourse* 31 (5), 593–607.
- You, L., (n.d.). Caught in the Split: Chinese Students in the Soviet Union. pp. 1960–1965. <https://www.wilsoncenter.org/publication/caught-the-split-chinese-students-the-soviet-union-1960-1965>.
- Yuan, H., Huang, X., 2019. China-England mathematics teacher exchange and its impact. *Front. Educ. China* 14 (3), 480–508.
- Zha, Q., 2006. The resurgence and growth of private higher education in China. *High. Educ. Perspect.* 2 (1), 54–68.
- Zha, Q., 2022. Revisiting the discourse of a Chinese model of the university: a Confucian-Legalist legacy impact perspective. In: Côté, J., Pickard, S. (Eds.), *Routledge Handbook of the Sociology of Higher Education*, second ed. Routledge, London and New York.
- Zhang, D., Li, X., Xue, J., 2015. Education inequality between rural and urban areas of the People's Republic of China, migrants' children education, and some implications. *Asian Dev. Rev.* 32 (1), 196–224.
- Zhao, D., Singh, M., 2011. Why do Chinese-Australian students outperform their Australian peers in mathematics? *Int. J. Sci. Math. Educ.* 9, 69–87.
- Zhu, J., 2009. Academic evaluation, academic journals and academic internationalisation—re-thinking the fervour of humanities and social sciences internationalisation [Xuepingjia, xueshuqikan yu xueshu guojihua: Dui renwenshehuikexue guojihuarechao de lengsikao]. *J. Tsinghua Univ.* 5, 126–137.
- Zhu, L., 2021. Chinese-American Academics on Edge. <https://global.chinadaily.com.cn/a/202106/09/WS60bffdcb31024ad0bac4723.html>.

China's engagement with Southeast Asia in higher education

Kejin Zhu^a and Rui Yang^b, ^a Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong; and ^b Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong

© 2023 Elsevier Ltd. All rights reserved.

Introduction	666
China as a host to Southeast Asian students	666
Confucius Institutes in Southeast Asia	668
China's higher education aid to Southeast Asia	669
Sino-ASEAN higher education partnership	670
Concluding remarks	671
References	671

Introduction

Southeast Asia has long been one of China's priorities for regional cooperation. Many analysts have pointed out that the strengthened economic ties contribute most to China's recent growing influence in Southeast Asia. China has continued to be ASEAN's largest trading partner for 12 years in a row. In turn, the Association of Southeast Asian Nations (ASEAN) countries became China's biggest trading partner for the first time in 2020 (China SCIO, 2021). Since the introduction of the Belt and Road Initiative (BRI), China's presence in Southeast Asia has been increasingly featured by a growing focus on further and strategic promotion of its economic, diplomatic, social and cultural connections with countries along the Maritime Silk Road in this region.

Compared with the heated discussions on China's recent economic power enhancement, China's diplomatic efforts in Southeast Asia in and through higher education remain very little documented. Among different levels of education, higher education is most influential in terms of its potential role as a carrier of cultural and political values. China's higher education diplomacy is thus an integrated and significant part of its soft power projection. When used tactfully, approaches to higher education engagement can effectively shape positive perceptions among international audiences, facilitate public diplomacy through transmission of the cultures and values of the host country, and exert influence on future elites of other nations (Yang, 2007).

China engages comprehensively with Southeast Asia in various dimensions of social and economic development. Despite that higher education is an integral part of the broader engagement, the literature on the relations between China and Southeast Asia in higher education is still much lacking. To address the gap, this chapter focuses on some major aspects of the comprehensive and complex relations including international students, language training, development aid, and partnerships in institutions and/or programs.

China as a host to Southeast Asian students

The number of ASEAN students enrolled in Chinese universities has skyrocketed during the past twenty years (See Fig. 1). In 2000, the number of the students from ASEAN countries totaled 5391, constituting 10.34% of the total international enrollments.¹ Since the BRI was launch in 2015, the enrollment of ASEAN students in Chinese universities has been undergoing a huge surge and reached 98,823 in 2018, accounting for 20.08% of the total foreign students.²

Apart from the consecutive growth, the group of Southeast Asian students also exhibits other characteristics. Firstly, a substantial number of Southeast Asian students tend to choose vocational tertiary education. In 2013, a total of 1337 ASEAN students studied in vocational tertiary institutions in China, accounting for 55.7% of all international enrollments in China's vocational tertiary institutions (Wen, 2016). Furthermore, until 2021 five annual China-ASEAN Vocational Education Exhibitions and Forums (MOE, 2021) have been co-organized by the Guangxi provincial government and MOE, and over 350 higher institutions participated in the discussions (MOE, 2018).

Secondly, non-degree and self-funded program enrollments have constituted a considerable proportion of the whole ASEAN student group. A substantial portion of Southeast Asian students have been enrolled in short-term exchange programs in Chinese universities. Since China has been forging tight economic ties with Southeast Asia, more self-funded international students consider a degree earned in Chinese universities or non-degree study experience in China as a stepping-stone to better job prospects at their

¹Authors' own calculation based on the data retrieved from Department of International Cooperation and Exchanges of Ministry of Education of China (2000).

²Authors' own calculation based on the data retrieved from Department of International Cooperation and Exchanges of Ministry of Education of China (2018).

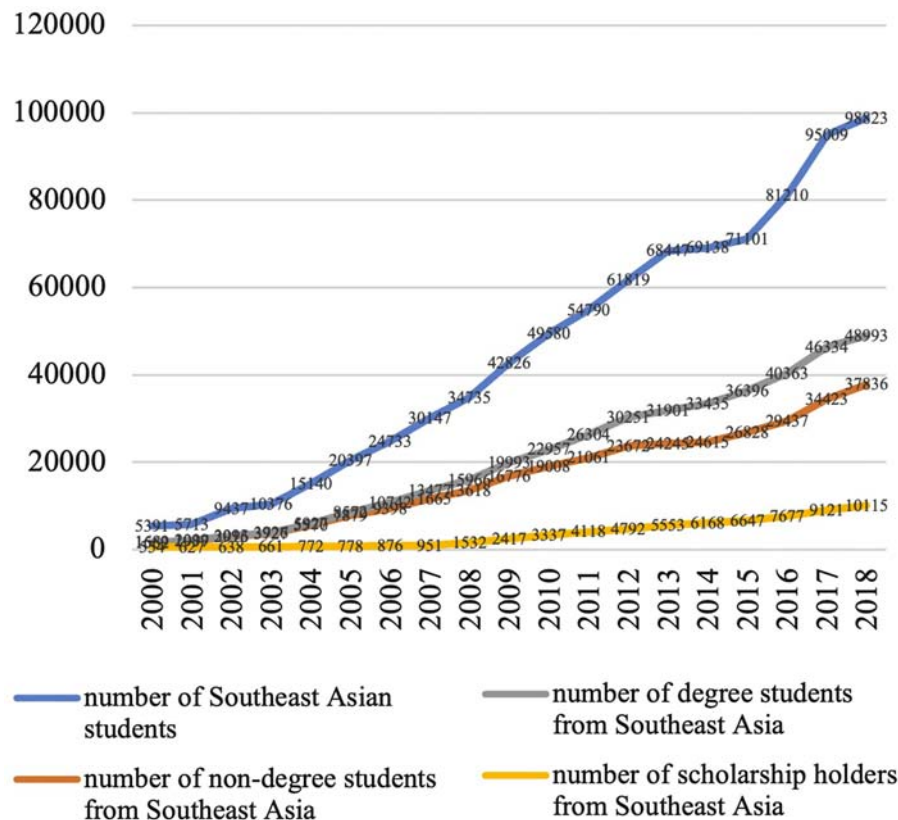


Fig. 1 The number of Southeast Asian students in Chinese universities (2000–2018). Authors' own design based on the data retrieved from *Statistics Book of International Students Studying in China* Department of International Cooperation and Exchanges of Ministry of Education of China (2000–2018).

home countries (Tan, 2018). As Fig. 1 shows, the pace of growth of total Southeast Asian students has always been higher than the increase of scholarship holders from ASEAN countries.

Meanwhile, it is noteworthy that an increasing number of Southeast Asian degree students tend to prefer STEM higher education programs because of the latest outstanding performances of Chinese universities in global rankings of academic subjects. Times Higher Education World University Rankings 2021 by Subject put Peking University in 9th in engineering, surpassing the National University of Singapore, the Imperial College London, and Yale University. Tsinghua University ranked 10th in QS World University Rankings by Subject 2017–2019: Engineering and Technology. It moved up to 9th in 2020, beyond Harvard University, California Institute of Technology, and the University of Tokyo.

In general, the proportion of international students in China in science and engineering programs increased from 8.55% in 2008 to 16.69% in 2018 while the proportion of international students in Chinese language programs decreased from 55.74% to 37.68%.³ This is a new development from the previous situation that international students in China concentrated extremely on Chinese language programs (Kuroda, 2014). The results of the latest survey by Obed et al. (2020), based on 1450 current international enrollments, also corroborate such an argument. They report that African and Asian students in China have patronized STEM majors, with Engineering topping the list, whereas more American and European students in China have flowed into social science disciplines.

Due to geographical proximity, southwestern provinces in China have been accredited as the “bridgehead” in China's foreign policy toward Southeast Asia, contributing to a sustaining increase of ASEAN students in the region. In 2017, 14,500 international students had been studying in Guangxi. Among them no less than 65% were from ASEAN countries. The size was tripled within the past decade (MOE, 2018). Lower living expenses are another influencing factor that Guangxi and Yunnan, as less-developed parts of China, have attracted a large number of Southeast Asian students (Yang, 2012).

It seems that diplomatic events have been closely associated with China's initiatives for improving student mobility with ASEAN countries. In 2004, China and ASEAN announced their first *Plan of Action to Implement the Joint Declaration on ASEAN-China Strategic Partnership for Peace and Prosperity* for the period 2005–2010. This plan firstly contained guidance of educational exchange and cooperation. Meanwhile, China promised to provide additional long-term and short-term scholarships to ASEAN countries, especially to

³Authors' own calculation based on the data retrieved from Department of International Cooperation and Exchanges of Ministry of Education of China (2008, 2018).

those that were underdeveloped (ASEAN-China, 2004). From then on, the number of international students from Southeast Asian countries began to demonstrate a modest growth.

On the 15th Anniversary of the China-ASEAN dialog relations in 2006, the *Joint Statement of China-ASEAN Commemorative Summit* was declared. In terms of socio-cultural cooperation, China launched a prestigious China-ASEAN Scholarship (ASEAN-China, 2006). Guiyang Declaration was announced at the First China-ASEAN Education Ministers Roundtable in 2010, agreeing to faithfully implement the “Double 100,000 Plan” to bring the number of Chinese and ASEAN students in each other’s universities to 100,000 respectively by 2020. China would offer 10,000 China-ASEAN Scholarships to Southeast Asian countries per year (Xinhua News, 2011). Next year, Guangxi allocated 15 million RMB as annual funding for Guangxi Government Scholarship for ASEAN Students (China News, 2013). Since then, the number of ASEAN students studying in China has experienced an accelerated increase.

China’s scholarships for Southeast Asian students were also frequently packaged in the speeches delivered by Chinese leaders at key diplomatic occasions. Premier of the State Council Li Keqiang’s address on the 16th China-ASEAN (10 + 1) Leaders’ Meeting promised that China would offer 15,000 Chinese Government Scholarships to ASEAN countries during the next three to five years. Later in 2016, his address at the first Lancang-Mekong Cooperation⁴ (LMC) Leaders’ Meeting proposed to offer 18,000 person-year scholarships and 5000 training opportunities to Mekong countries in the next three years (Li, 2016).

Particularly, with the unveiling of the BRI, new scholarships at national, provincial, city, university, enterprise levels have been set up to attract students from countries along the Belt and the Road. The representative national one is the Chinese Government Scholarship-Silk Road Program, which offers 10,000 annual scholarships to international students from countries along the Belt and the Road, including all ASEAN countries. It turns out that the number of scholarship holders from Southeast Asia rose to 10,115 in 2018, tripled the number in 2010 before the commencement of “Double 100,000 Plan” (See Fig. 1).

Confucius Institutes in Southeast Asia

As China is forging tight economic, cultural and political connections with Southeast Asia, there is an increasing need for people with sufficient Chinese language proficiency and knowledge about Chinese society. Confucius Institutes (CI) and Confucius Classrooms have thus become effective platforms to cultivate “zhihua youhua” (with knowledge about China and a positive impression of China) graduates who are also motivated by better job prospects. Since the first CI established in 2004, 548 CIs and 1193 Confucius Classrooms have been fast established in 146 countries and regions by 2018 (Chen, 2020). By 2020, 39 CIs and 34 Confucius Classrooms scatter in ten Southeast Asian countries as indicated in Fig. 2 below.

The development of CIs in Southeast Asia demonstrates similar features with the CIs in other regions but also follows their own trajectory. Different from other language training institutes such as the British Council, CIs operate in a unique way by pairing leading universities in the host countries with the Chinese domestic ones. Moreover, Hanban, a government unit mainly managed by MOE, was responsible for offering startup funding, allocating of teaching staffs and materials, and coordinating the two sides (Hartig, 2012).

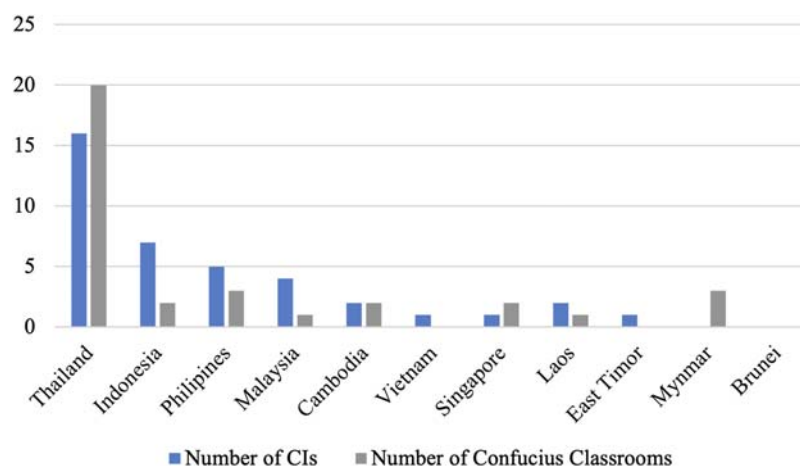


Fig. 2 Number of CIs and confucius classroom in Southeast Asia. Authors’ own design based on data retrieved from Chen (2020).

⁴Lancang-Mekong Cooperation refers to the new sub-regional cooperation mechanism initiated by China, Myanmar, Laos, Thailand, Cambodia, and Vietnam to focus on political and security issues, economic and sustainable development, and social, cultural and people-to-people exchanges of Lancang-Mekong countries via concrete and pragmatic projects.

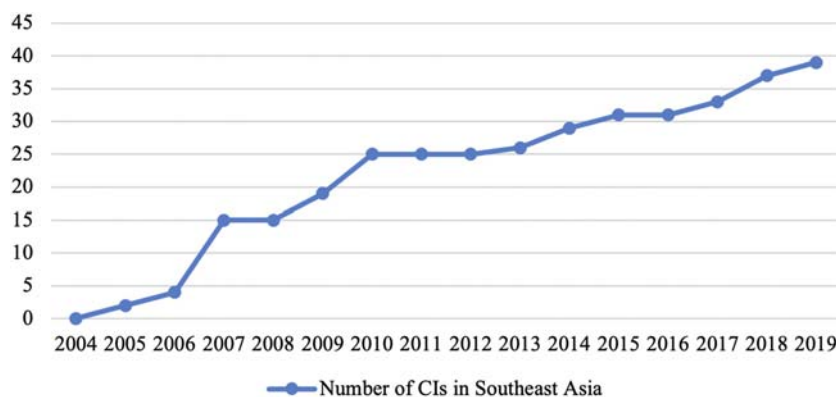


Fig. 3 Growth of CI numbers in Southeast Asia from 2006 to 2019. Authors' own design based on data retrieved from [Chen \(2020\)](#).

It needs to note that because of intensified criticism on CI's possible political and ideological intentions from mainly Western politicians, Hanban is no longer the host of CIs since June 2020. Instead, the Chinese International Education Foundation (CIEF), a non-profit charitable social organization founded by 27 Chinese universities, enterprises, and NGOs, is now in charge of the management of CIs (CIEF, 2021).

The overwhelming majority of partner universities of CIs in ASEAN countries is local higher education institutions in Guangxi province. Eight Guangxi borderland universities have participated in the set-up of CIs in four Southeast Asian countries, reflecting the influence of China's "bridgehead" diplomacy. Regarding the development of CIs in Southeast Asia, the periods of 2008–2010 and of 2015–2019 witnessed two fast growth points (as Fig. 3 shows), corresponding respectively to the announcement of "Double 100,000 Plan" in 2010 and the launch of BRI in 2015 as aforementioned.

The CI at the Royal Academy of Cambodia serve as an illustrative example. Starting from its inauguration, it has been endowed with the mission of symbolizing the good relationship between China and Cambodia. In December 2009, then Vice President Xi Jinping kicked off his official visit to four Asian countries, including Japan, South Korea, Myanmar, and Cambodia. Xi attended the opening ceremony of the first CI in Cambodia and delivered a public speech. He recalled the long history of friendly communication between China and Cambodia and defined the CI as a milestone in Sino-Cambodian people-to-people and cross-cultural exchanges to strengthen the current cooperative political and economic partnership between the two countries (Xinhua News, 2009).

To some extent, the CI at the Royal Academy of Cambodia has been acting like a local diplomat through offering language teaching services to the Cambodian government officials. There were 23 Chinese teaching bases and two CI classrooms supported by the CI at Royal Academy of Cambodia until December 22, 2019. Altogether 12 Chinese centers were set up in high-level government and military offices (Confucius Institute at Royal Academy of Cambodia, 2019). For classes taught at the Prime Minister's Office and the Senate as well as the Ministry of National Defense, the CI at the Royal Academy of Cambodia would arrange the state-sponsored teachers with rich language teaching experience.

Additionally, the CI at the Royal Academy of Cambodia has coordinated with China's Embassy in the Kingdom of Cambodia to organize social and academic activities. A typical example is a project that translated and published *Xi Jinping: The Governance of China I (Khmer version)* by the CI at the Royal Academy of Cambodia. The book launch was held at the Prime Minister's Office on April 11, 2017, witnessed by the Cambodian Prime Minister Hun Sen and the Director of the State Council Information Office of China. Other activities include seminars on Sino-Cambodian relations and BRI, "Chinese Bridge" Match (Chinese Language Proficiency Competition for Foreign College Students) in Cambodia with themes related to "Building a Community with a Shared Future for Humanity," Chinese Essay Competition about "China in My Heart," theme music composition of documentary film named *The Story of Time: Sino-Cambodian Friendship*.

China's higher education aid to Southeast Asia

With its burgeoning volume of development aid to developing countries during the past two decades, China's distinct donor logic has triggered intense debate among scholars, containing both criticism and defense. Critics share the viewpoint that despite the generosity and non-conditionality of most Chinese aid projects, China's true motivation still focuses on gaining access to raw materials (Brautigam, 2009; Tan-Mullins et al., 2010). In contrast, defenders usually refer to China's "win-win" donor logic that emphasizes on equal partnerships and mutual benefits (Yuan, 2011). They point out the advantages of adding an alternative besides Western donors especially in consideration of the China's development model as a useful reference point for developing countries (Wang, 2014).

Higher education aid is a key ingredient of China's development assistance to consolidate South-South cooperation, which is traditionally unfolded in the large number of scholarships for foreign students from less developed countries, institution or

program partnerships, and degree training or short-term human capacity building programs for both students and university professionals in Chinese universities and post-secondary vocational colleges. The scholarly debates, however, predominantly surround China's higher education engagement with Africa in forms of long- and short-term training/scholarship programs, CIs, stand-alone projects such as school construction, and 20 + 20 university partnership scheme.

In comparison, however, China's higher education development aid to the developing countries of ASEAN has been much less discussed. Given that China has employed not only development assistance but also non-development aid, low-interest loans, and investment agreements, its foreign aid has already reached a gigantic size and exerted a tangible influence in many Southeast Asian countries including Myanmar, Cambodia, Laos, Vietnam, Thailand, Indonesia, and Philippines (Lum et al., 2008). China's aid to higher education development in Southeast Asia consists of scholarships, training seminars for professionals, and school building projects.

For instance, the Ministry of Commerce of China (MOFCOM) Scholarship is a China Aid project to build intellectual capacity through master's and doctoral programs. It tends to elaborately design the curriculum with lessons from China's development model and foreign policy. In 2020, there were 32 universities designated by MOFCOM to recruit students. Seven programs in public administration and economics explicitly outlined their objectives as cultivating future leaders of developing countries via learning from China's development practices. The compulsory, elective courses and field trips contained in their syllabus were closely related to China's perspectives on public policy and development strategy.

The Academy for International Business Officials (AIBO, also known as Training Center of the Ministry of Commerce) is the only educational and training institution to administrate and coordinate all China-Aid training programs. It has been committed to telling China's stories to seminar trainees. The revised version of the *Work Manuals for Implementation and Management of MOFCOM's China-Aid Training Programs* released on 2010 listed the specific training requirements about China's national conditions. It requires that the training program's providers ought to choose site visits and field trips that promote both international communication and cooperation and trainees' recognition and understanding of China's national conditions (AIBO, 2012).

The Institute of South-South Cooperation and Development (ISSCAD) was another flagship China Aid project with funding from the Department of Foreign Assistance of MOFCOM. Based at Peking University, ISSCAD was initiated by President Xi Jinping on the High-Level Roundtable on South-South Cooperation at the 70th Anniversary Celebration of the United Nations in 2015. Aiming to share the development experience, it facilitates capacity building with mid-level to senior officials, managers and researchers from developing countries, including less-developed ASEAN countries. Its curriculum consists of China Immersion Program (China's development theory and practice) as the core course. It covers a variety of themes on economic system reform and development planning, China's opening-up and globalization and BRI in its one-semester-long China Immersion Program (ISSCAD, 2018).

Starting from 2015, the LMC was China's another sub-regional diplomacy in Southeast Asia with flagship higher education aid projects contained. The *Five-Year Action Plan of Lancang-Mekong Cooperation (2018–2022)* further supported China to establish vocational education partnership with Mekong countries. Premier Li Keqiang's address at the second LMC Leader's Meeting suggested that China would extend agricultural cooperation with countries along the Mekong River (Li, 2018). Within such a policy context, most recent higher education infrastructure and short-term training aid projects in Mekong countries related to agriculture and vocational training. The China-funded University of Kratie was the first public comprehensive university focusing strongly on agronomy in the Northeastern Cambodia. With 10,000,000 US dollars from China to construct its teaching building, laboratory, lecture hall, library, and cafeteria, its operation was officially started on April 25, 2018.

Sino-ASEAN higher education partnership

Along with China's successful developmental trajectory of moving from periphery to center, many Chinese universities have also climbed up in global ranking systems. Their performances have attracted an increasing number of Southeast Asian institutions to form partnerships or joint programs with them. Two higher education institutions located in Chongqing and Jiangsu respectively are cooperated jointly by Chinese and ASEAN counterparts. A total of 36 joint programs between China and ASEAN countries in the form of "3 + 1" "2 + 2" or "2 + 1" are running in Chinese universities. It merits attention that 25 cooperative programs are established with Singapore, ten with Malaysia, and one with Thailand (Wen, 2016).

Besides the joint programs operating within China, Chinese universities have recently begun to follow "go global" strategies to establish overseas campuses and joint programs in ASEAN countries. According to the data released by *Guangming Daily* (2016), Chinese tertiary institutions had already set up 98 overseas higher education programs, among which three MOE-approved international branch campuses had been established in Southeast Asian countries, including Soochow University (Laos), Xiamen University (Malaysia), Bangkok Business School of Yunnan University of Finance & Economics (Thailand). He and Wilkins (2018) have discussed the impacts of China's education-related soft power aspects on international branch campuses in Southeast Asia. Their interview data identify the necessity of the Chinese language, and cultural history and heritage, as well as China's exponential economic growth as driving forces for local students to choose Chinese universities' overseas campuses.

Moreover, borderland Chinese universities have initiated transnational joint programs in less developed ASEAN countries. For example, the "2 + 2" program launched by Honghe University (Yunnan) to operate in Vietnam was endorsed by MOE and the Ministry of Education and Training of Vietnam in 2011 (Honghe University Website, 2019). Additionally, the higher education forums and associations between China and ASEAN countries are also active in seeking more collaboration opportunities, including

annual China-ASEAN University President Forum, the ASEAN-China Academic Cooperation and Exchange Program, the Association of Pacific Rim Universities, and Universitas 21.

Especially for latecomer countries without well-developed higher education resources, higher education partnership with China usually serves as an icebreaker to strengthen diplomatic relations. For instance, the RUPP (Royal University of Phenom Penh)-DHY Chinese Language Center was co-established by Dali University, Honghe University, Yuxi Normal University, and RUPP in 2007. The incentive of the three Yunnan local universities was the guiding foreign policy of "bridgehead" strategy. Along with the announcement of China and ASEAN Free Trade Area, Yunnan stood out as the bridgehead facing Southeast Asia. Yunnan Provincial Party Committee and People's Government released the *Advice on Accelerating the Implementation of "Going Out" Strategy by Universities to Improve Higher Education Internationalization Level* in 2006 (MOE, 2011). Through promoting higher education internationalization, it fulfills "good neighbor" and "peaceful development" foreign policies. Related initiatives consist of the establishing ten CIs and Chinese language centers in Southeast Asia and South Asia (Yunnan Provincial Party Committee and People's Government, 2006).

During his investigation in Yunnan in 2009, China's then President Hu Jintao officially stated a plan to set up Yunnan as the bridgehead. Later, the State Council released the *Advice on Supporting Yunnan Province to Accelerate the Construction of Being the Key Bridgehead Facing the Southeast*. It clarified that bridgehead construction would fully carry out the guiding principle of China's diplomacy of "building friendship and partnership with neighboring countries" and the foreign policy of "an amicable, secure and prosperous neighborliness." One of the detailed projects was to support the education exchange and cooperation between higher education institutions in Yunnan and Southeast Asia (State Council of China, 2011).

Besides being the icebreaker for "bridgehead" strategy at the beginning, higher education partnership has continued to reinforce diplomatic relations between China and Southeast Asia. Taking RUPP-DHY Chinese Language Center again as an example, the Yunnan delegation led by Governor Li Jiheng and Vice Governor Gu Zhaoxi of the Yunnan province visited the Greater Mekong Sub-region from March 25 to April 6 in 2012. Before paying a visit to Cambodia, they were informed that there had been a successful higher education partnership program at RUPP. Then, the Education Department of Yunnan assisted Dali University to apply for the donation of Chinese language laboratory.

On 6 April, the governor and the vice governor attended the China (Yunnan Province) and Cambodia Economic and Trade Cooperation Conference. The governor stressed the significance of enhancing people-to-people cultural communication between Cambodia and Yunnan through educational collaboration in his keynote speech. After conference, the Vice Governor visited RUPP and unveiled the nameplate of Language Laboratory sponsored by Yunnan Provincial Government with a donation of devices of more than 200,000 RMB value (Yunnan Daily, 2012). While the inauguration of language laboratory has laid a solid foundation for more far-reaching higher education partnership between both sides, it is a clear diplomatic symbol of Yunnan to deepen its amicable neighborhood and economic and trade cooperation with Cambodia.

Concluding remarks

The BRI serves as a powerful catalyst to deepen the influence of both economic and soft power of China in Southeast Asia. Thereinto, higher education plays an active role, echoing China's peripheral diplomacy to further diffuse its cultures and values. Remarkably, the ascending performance of Chinese universities in STEM has become an emerging attraction to Southeast Asian neighbors. Meanwhile, China's achievements in scientific research and development break the fault line of its economic development model. It is not unexpected to find that the export of China's development model has recently become an indivisible part of its higher education aid projects. Although many approaches including student mobility, language training, educational development assistance, and university partnership have all tried to incorporate China's foreign policies across time, a systematic and coherent mechanism of China's higher education engagement with Southeast Asia has not been formulated yet.

References

- AIBO, 2012. Work Manuals for Implementation and Management of MOFCOM's China-Aid Training Programs. Retrieved from. <https://www.china-aibo.cn/article/content/view?id=318>.
- ASEAN-China, 2004. Plan of Action to Implement the Joint Declaration on ASEAN-China Strategic Partnership for Peace and Prosperity (2005–2010). Retrieved from. https://asean.org/?static_post=plan-of-action-to-implement-the-joint-declaration-on-asean-china-strategic-partnership-for-peace-and-prosperity.
- ASEAN-China, 2006. Joint Statement of China-ASEAN Commemorative Summit. Retrieved from. <http://jm.chineseembassy.org/eng/xw/t280186.htm>.
- Brautigam, D., 2009. *The Dragon's Gift: The Real Story of China in Africa*. Oxford University Press.
- Chen, L., 2020. The influencing factor of the distribution of CIs in Southeast Asia (in Chinese). *Lit. Educ.* 6, 44–49.
- China News, 2013. Guangxi Province Set Up Scholarship for ASEAN International Students. Retrieved from. <https://www.chinanews.com/xxsh/2013/07-02/4995018.shtml>.
- China SCIO (The State Council Information Office of the People's Republic of China), 2021. China-ASEAN Trade Skyrockets by 85 Times in Three Decades. Retrieved from. http://english.scio.gov.cn/pressroom/2021-07/30/content_77662122.htm.
- CIEF, 2021. About CIEF. Beijing. Retrieved from. <https://www.cief.org.cn/jj>.
- Confucius Institute at Royal Academy of Cambodia, 2019. The 10th Anniversary of Confucius Institute at Royal Academy of Cambodia. Retrieved from. <http://cirac.kh.chinesecio.com/zh-hans/node/3081>.
- Department of International Cooperation and Exchanges of Ministry of Education of China, 2000–2018. *Laihua liuxuesheng jianming tongji* [Statistics Book of International Students Studying in China]. Department of International Cooperation and Exchanges of Ministry of Education of China.

- Guangming Daily, 2016. Overseas Education Programs and Institutions: Going Out and Going Stably. Retrieved from. <http://edu.people.com.cn/n1/2016/1129/c1006-28904841.html>.
- Hartig, F., 2012. Confucius institutes and the rise of China. *J. Chin. Polit. Sci.* 17, 53–76.
- He, L., Wilkins, S., 2018. The return of China's soft power in South East Asia: an analysis of the International Branch Campuses established by three Chinese Universities. *High Educ. Pol.* 1–17.
- Honghe University Website, 2019. International Project. Retrieved from. <http://en.uoh.edu.cn/info/1078/1055.htm>.
- ISSCAD, 2018. Master and Doctoral Programs in National Development. Retrieved from. <http://img.bimba.pku.edu.cn/resources/file/9/10/2018/04/19/2018041917253318.pdf>.
- Kuroda, C., 2014. The new sphere of international student education in Chinese higher education: a focus on English-medium degree programs. *J. Stud. Int. Educ.* 18, 445–462. <https://doi.org/10.1177/1028315313519824>.
- Li, K., 2016. Li Keqiang's Address at the First Lancang-Mekong Cooperation Leaders' Meeting. Retrieved from. http://www.lmcchina.org/eng/2016-03/25/content_41449867.html.
- Li, K., 2018. Li Keqiang's Address at the Second Lancang-Mekong Cooperation Leaders' Meeting. Retrieved from. http://www.lmcchina.org/2018-01/11/content_41447216.htm.
- Lum, T., Morrison, W., Vaughn, B., 2008. China's "Soft Power" in Southeast Asia. CRS Report for Congress RL34310. Congressional Research Service, Washington, DC, 4 January.
- Ministry of Education (MOE), 2011. HE of Yunnan Province Develops Fast, Values Quality, Demonstrates Distinctiveness and Strong Activities. Retrieved from. http://www.moe.gov.cn/s78/A08/moe_745/201103/t20110314_115858.html.
- Ministry of Education (MOE), 2018. Guangxizhuangzuzhiqu fahui quweiyoushi zhuoli dazao mianxiang dongmeng jiaoyu hezuojiaoliu xingao [Guangxi Province Uses Its Geographical Advantages to Strengthen Cooperation With ASEAN]. Retrieved from. http://www.moe.gov.cn/jyb_sjzl/s3165/201812/t20181227_365124.html.
- Ministry of Education (MOE), 2021. Guangxi tuidong zhijiao neihan fazhan fuwu yidaiyili [Guangxi Province Promotes Vocational Education Development to Serve the BRI]. Retrieved from. http://www.moe.gov.cn/jyb_xwfb/s5147/202108/t20210830_555861.html.
- Obed, L., Banda, L., Liu, J., Zhou, W., 2020. Trending Majors Among International Students in China: A Case of African Students. <https://doi.org/10.6084/m9.figshare.13269533>.
- State Council of China, 2011. Released Advice on Supporting Yunnan Province to Accelerate the Construction of Being the Key Bridgehead Facing the Southeast. Retrieved from. http://www.gov.cn/zwqk/2011-11/03/content_1985444.htm.
- Tan, V., 2018. Why More Southeast Asian Students Are Choosing China for Higher Education. Channel News Asia. Retrieved from. <https://www.channelnewsasia.com/news/asia/why-more-southeast-asian-students-are-choosing-china-for-higher-10042118>.
- Tan-Mullins, M., Mohan, G., Power, M., 2010. Redefining "aid" in the China–Africa context. *Dev. Change* 41 (5), 857–881.
- Wang, L., 2014. Internationalization with Chinese characteristics: the changing discourse of internationalization in China. *Chin. Educ. Soc.* 47 (1), 7–26.
- Wen, W., 2016. China's approach towards HE regional cooperation with ASEAN. In: Collins, C.S., Lee, M.N.N., Hawkins, J.N., Neubauer, D.E. (Eds.), *The Palgrave Handbook of Asia Pacific Higher Education*. Palgrave Macmillan, New York, pp. 173–182.
- Xinhua News, 2009. Vice President Xi Jinping Attended the Inauguration of the First Confucius Institute in Cambodia. Retrieved from. http://www.gov.cn/dhd/2009-12/22/content_1494281.htm.
- Xinhua News, 2011. China-ASEAN Cooperation: 1991–2011. Retrieved from. http://www.asean-china-center.org/english/2011-11/16/c_131249656_8.htm.
- Yang, R., 2007. China's soft power projection in higher education. *Int. High. Educ.* 46, 24–25.
- Yang, R., 2012. Internationalization, regionalization and soft power: China's relations with ASEAN member countries in higher education. *Front. Educ. China* 7 (4), 486–507. <https://doi.org/10.3868/s110-001-012-0025-3>.
- Yuan, T., 2011. Chinese Educational "Aid" to Africa: A Different "Donor Logic"? (Ph.D. thesis), e-thesis. Available online at. <https://research-information.bris.ac.uk/ws/portals/portal/34506403/540880.pdf>.
- Yunnan Daily, 2012. China (Yunnan Province) and Cambodia Economic and Trade Cooperation Conference. Retrieved from. http://news.yntv.cn/content/15/15_528166.shtml.
- Yunnan Provincial Party Committee, People's Government, 2006. Advice on Accelerating the Implementation of "Going Out" Strategy by HE Institutions to Improve HE Internationalization Level. Retrieved from. https://code.fabao365.com/law_296421.html.

Datafication

Matthew Thorpe^a and Sam Sellar^b, ^a Manchester Metropolitan University, Manchester, United Kingdom; and ^b University of South Australia, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	673
A brief overview of datafication in education	674
Datafication in higher education	675
Emerging challenges	677
Deprofessionalization	677
Inequalities	678
Privatization	678
Regulation	679
Conclusion	679
References	679

Introduction

Datafication is a new word for a process with a long history: the translation of things into quantitative information. The term came into popular usage around 2013 with the publication of Mayer-Schonberger and Cukier's (2013) book, *Big data: A revolution that will transform how we live, work, and think*. The term is closely linked to the emergence of big data analytics and thus the production, analysis and use of *digital* data. This article will focus specifically on developments in the datafication of education over the past two decades. In particular, we consider how datafication has contributed to change in higher education over this period and identify three emergent challenges and one effect that it produces for educators and policymakers.

In order to discuss datafication, it is necessary to begin by considering data as the root word and the product of the process. Data now generally refers to digital data, but the term has a longer and broader usage extending back to the early eighteenth century. Data has referred to different things over this period, but it has always retained the sense of "that which is given prior to argument" (Rosenberg, 2013, p. 36). Data are distinct from facts, evidence and truth. In the form of principles and premises, or quantitative information produced through observation, experiment or digital capture, data have always provided a basis for analysis and argument from which facts and evidence may emerge. As Rosenberg (2013) writes, "[e]lectronic data, like the data of the early modern period, is given. It may be that the data we collect and transmit has no relation to truth or reality whatsoever beyond the reality that data helps us to construct" (p. 37). Datafication is a process that generates ever-increasing volumes of material for interpreting, representing and remaking our worlds.

Mayer-Schonberger and Cukier (2013) define datafication as "taking information about all things under the sun ... and transforming it into a data format to make it quantified" (p. 15). The emergence of datafication as a distinct phenomenon is closely linked to digitalization and significant increases in computational capacities that enable greater volumes and varieties of data, indeed all data in some cases, to be captured and analyzed increasingly quickly (Kitchin, 2014). The constructive nature of datafication is exemplified in the use of large digital data sets to perform predictive analytics, which have driven recent excitement about the promise of big data. Datafication enables us to make predications and shape the future, beyond simply generating information about the past or present. However, this world making capacity of big data brings new risks. Mayer-Schonberger and Cukier (2013) argue that "big-data predictions render the future less open and untouched ... our future seems already sketched in faint traces that are discernible to those with the technology to make them apparent" (p. 194). Datafication is thus not only a relatively new digital technology, but also a thoroughly political and economic one that grants those who are able to capture and analyze data the ability to turn potential futures into a set of probable and actual ones.

As with practically every aspect of contemporary life, datafication has had a significant impact in education (e.g. Bradbury and Roberts-Holmes, 2017; Hartong and Förschler, 2019; Sellar, 2015; Williamson, 2017; Wyatt-Smith et al., 2021). For the purposes of this article, we will limit our discussion to formal education, and we will use the specific example of higher education to trace the development of datafication over the past decade and into the near future. Williamson (2017, p. 5) defines datafication in education as "the transformation of different aspects of education (such as test scores, school inspection reports, or clickstream data from an online course) into digital data". Recent studies of this phenomenon build on a longer tradition of research into the role that quantification has played in transforming policy and governance (e.g. Gorur, 2014; Grek, 2009; Lawn, 2011; Lingard, 2011; Ozga, 2009). Over the past two decades, many government departments and educational institutions have moved beyond routinely collecting data to monitor the performance of students, educators and institutions, and are seeking to embed digital datafication as a foundational element of organizational and educational strategy. Private actors, particularly the education technology sector, now also accrue vast quantities of digital data through the use of their products and services. While test scores and attendance records have been used to quantify educational engagement and performance since the 19th century (Stephens, 1998), digital data are

increasingly shaping educational experience in fundamental ways through, for example, the predictions that guide policymakers and the algorithms that underpin personalized learning.

Datafication is both a product of, and a resource for, imagined futures. Drawing on the work of Sheila Jasanoff, Williamson (2017) argues that it is in sociotechnical imaginaries – “the visions and values that catalyze the design of technological projects” (p. 16) – that “transformative scientific ideas, technological objects and social norms become fused in practice and help to sustain social arrangements or create new rearrangements in cultures, institutions and routines” (p. 16). Datafication is driven by desire for data and by technologies that enact this desire, which together contribute to arranging education in new ways (Thompson and Sellar, 2018).

The COVID-19 pandemic is accelerating the use of digital platforms in education, and thus also the datafication of education, because the use of these platforms generates large volumes of data. Many people quickly came to rely upon new technologies to teach, learn and meet after the pandemic caused the closure of schools and universities, and the EdTech industry saw this as an opportunity to disrupt education systems and embed datafication through the use of their platforms (Williamson and Hogan, 2020). Moreover, investment in the EdTech sector doubled in 2020 and the profitability of this sector is predicated on the collection, aggregation and monetization of data through the extraction of economic rents (Komljenovic, 2021). Datafication will drive the next generation of commercialization in education, and perhaps even the disruption of public education systems. In this context, addressing the emergent challenges that datafication creates for education has become even more urgent.

The datafication of education has occurred in symbiotic relationship with the rise of “evidence-based policy” approaches, and debates about its risks and benefits often trace well-worn lines of contestation between qualitative and quantitative research paradigms. These debates pose datafication as an epistemological problem, raising questions about the most reliable ways to generate evidence and facts from data. Datafication is also commonly posed as an ontological problem when it is critiqued as a technology that converts human activity and experience into digital data that encourage instrumental-rationalist framings of education and its subjects (i.e. human capital). While both of these problems are important and demand attention, our primary aim in this article is to provide an organized overview of new forms of digital datafication, focusing on the practices and practical effects of data, such as deprofessionalization, privatization and the creation of new inequalities.

This article is divided into three main sections. We begin by providing a brief overview of the evolution of datafication across different education sectors and we consider some of the impacts it has had on policy and pedagogy. This discussion is necessarily limited, but provides a general perspective on datafication in education and its antecedents. We then turn to the example of higher education and discuss how technological advances, such as the Internet of Things (IoT) and the use of Virtual Learning Environments (VLEs), have provided the conditions for higher education institutions (HEIs) to pursue new data-driven practices. These include using data to track and support students with their studies and to pursue interventions to improve learning outcomes and wellbeing. The use of data to guide interventions is linked to the capacity to predict student behaviors and outcomes, and we consider the predictive use of data and its use to train algorithms that automate aspects of education, for example students engaging with chatbot services. The final section of the article explores three emergent challenges and one effect that need to be considered as the datafication of education intensifies: (1) the deprofessionalization of educator’s work; (2) the creation of new inequalities; (3) the privatization of education; and, as a result, (4) the need for new modes of regulation.

A brief overview of datafication in education

Schools and universities have a long history of translating things and activities into quantitative formats. Spelling tests were developed during the late 19th century, closely followed by IQ tests in early 20th century (Symonds, 1940). The first international large-scale assessment (ILSA) was developed in 1950s and national assessment programmes emerged during the latter part of the 20th century, often linked to new accountability agendas. From the 1990s, the scale and scope of ILSAs has grown and large-scale assessments have evolved to quantify a more diverse set of student attributes (Sellar and Lingard, 2014). National and international assessments of students have contributed to growing volumes of information that are used to shape policy and practice. These assessment data have been combined with other data relating to student behavior (including attendance) and demographic data to monitor and manage education systems. However, the emergence of new modes of datafication involves larger volumes of digital data being generated more quickly, stored in databases that are increasingly joined up, and used for a range of new purposes.

The datafication of education extends from early years settings to post-secondary education. Children are now exposed to significant processes of datafication and teachers in early years and primary settings are now under significant pressure to record data regarding children’s development (Bradbury and Roberts-Holmes, 2017). The “data double” of the child becomes a yardstick for assessing how they are progressing, often substituting for an assessment of the individual themselves (Bradbury, 2019). This attraction to detailed recording of data on student progress has influenced the pedagogical approaches that are adopted in the classroom, because those that are easily measured lend themselves to datafication and educational imaginaries in which student and teacher performance is governed through data. This can lead to teachers feeling conflicted when their educational values and professional knowledge are at odds with the practices that datafication encourages or requires (Roberts-Holmes, 2015). A culture of performativity has also emerged as the measurement of children and young people in schools has intensified, resulting in teachers’ work being evaluated and rewarded based on data (Ball, 2003). Cultures of performativity have intensified in schooling with publication of league tables linked to parental school choice and school funding. New modes of accountability linked to performance data have had a significant impact on how the purposes of education are imagined and how judgments are made about educational practices.

As young people move into secondary schools and through to post-secondary education, processes of datafication can become increasingly normalized and opaque. Datafication increases as students generate more digital trace data as they interact with a variety of education institutions, platforms and applications. Students also engage with a growing number of formative and summative assessments as they progress through the stages of education, creating possibilities for using data from past assessments to predict future performance and guide “early interventions”. In this example, we can see how datafication and prediction shape evaluations of students’ potential and pedagogical decision-making. Indeed, datafication increasingly informs pedagogy, curriculum and assessment, and in the process can augment, discredit or replace teacher professional judgment, which provides a basis for teaching as a profession.

Over the past decade, we have also observed the promotion within school settings of personalized learning that is tailored to the needs of the individual student (Williamson, 2016). This approach is premised on the view that real-time data generated through the learning process can enable curriculum and pedagogy to be adjusted to meet individual abilities and needs, enabling learners to maximize their potential to achieve successful outcomes. Parents can also more readily access data on their child’s progress and students are increasingly expected to reflect upon their own digital footprint and set educational targets on this basis, providing an example of how the quantified self is enjoined to continually monitor and modify their behavior based on measurements and targets (Lupton, 2016).

In higher education, processes of datafication have been lauded as a means for students to gain greater insight into their learning and wellbeing. Institutions also see other benefits from datafication, including, for example, intervening with students to address engagement in more targeted ways that may translate into improved retention rates. Data are also valued as market intelligence and as a means to improve recruitment and admissions processes (Jisc, 2021). The increased scale of university student bodies and the research community within the higher education sector has allowed it to pioneer new approaches to datafication in education. These developments are often described using the label of learning analytics, and many HEI’s now have staff and student facing applications that quantify learning and progress. HEI’s also support research communities that explore what can be datafied and how such “intelligence” can be actioned. At the same time, there is a significant number of scholars offering a critical commentary on practices of datafication across all education sectors, and working to develop critical data literacy for both staff and students. Datafication is often controversial and increasingly subject to competing educational imaginaries, representing a tension between the perceived usefulness of data-driven “business intelligence” and critical perspectives that highlight risks of various kinds.

Datafication in higher education

Our discussion of datafication in higher education will focus on developments over the past two decades linked to the growth of online learning and digitalization. Distance and e-learning experienced significant growth in the 2000’s (Radford, 2011). These developments represent a significant break from face-to-face teaching and are a key milestone in the growth of capacities for datafication across the sector. E-learning led to the introduction of a variety of digital applications and platforms, including the widespread introduction of VLEs in colleges and universities. While the introduction of e-learning was not inspired initially by imaginaries of datafication, it produced a shift that enabled educational processes and events to be more readily captured as digital data.

Distance and e-learning approaches thus created fertile conditions for the growth of datafication and new possibilities for data use. The digital footprint created by students leaving traces when they interact with virtual learning environments and other applications could be analyzed for insights into student learning. For example, clickstream data captures time-stamped records of click events as students interact with online resources and activities. Such data are valued for their potential to provide different insights than those made possible by intermittent assessments and teacher observation. However the usefulness of such data has also been described as “partial and noisy” (Baker et al., 2020). Indeed, the tension between the promises of datafication and its technical limitations has often resulted in sociotechnical imaginaries outrunning actually existing capabilities.

Increased awareness of the potential usefulness of digital data can be traced back to the rise of two closely related areas: educational data mining (EDM) and learning analytics. EDM and learning analytics share common assumptions about the value of data-driven approaches to education. The two fields emerged at a similar point in time, with the first EDM international conference being held in 2008 and the first international learning analytics conference taking place in 2011. Both fields have much in common, but there are important differences: learning analytics aims to use data to aid human judgment while EDM focuses on automated data analytics and adaption of learning processes without the need for human intervention (Siemens and Baker, 2012). Both approaches have accelerated efforts to capture and use a greater volume and variety of data in higher education.

In the last decade, we have observed growth in datafication beyond digital traces left by students as they interact with educational applications on individual computers. This expansion is linked to the development of the Internet of Things (IoT): the ever-expanding network of objects that can communicate and exchange data over the internet through sensors and software (Rouse, 2019). The network effects created by the move to online learning and the automatic data capture that occurs through the use of digital devices marks an important point in the rise of big data and the datafication of human activity, because the IoT enables data to be generated and recorded without the need for human data entry (Ashton, 2009). The IoT also provides new possibilities for data to be aggregated in novel and abstract ways without human involvement. Kitchen (2014) identifies “relationality” as one of seven characteristics of big data and this describes how data sets are joined up across a wide variety of interoperable devices, platforms and applications. For example, the growth in use of handheld devices, wearables and sensing technologies makes it possible

to capture, combine and analyze huge volumes of data relating to a wide range of student activities. In addition to recording activity on personal computers, the use of wearable cameras, biosensors (recording heartbeat, body temperature, skin conductivity etc.), location tracking and eye tracking is being pursued as a means to provide better educational experiences for students (Blikstein, 2013). New possibilities and challenges are created by the increasing ubiquity of often invisible technologies capable of recording and sharing data.

The emergence of the IoT has also changed the nature of physical spaces in higher education. Easterling (2012, p. 1–2) describes the IoT as “... a world embedded with so many digital devices that the space between them consists not of dark circuitry but rather the space of the city itself”. The HE sector has embraced this development in its pursuit of smart campuses, which can be seen as a subset of the Smart City (Thrift, 2014). The concept of the Smart City describes the use of data collected via various methods to manage assets, resources and services efficiently and to improve the experience of the population living in urban areas (Lai et al., 2021). Examples of specific uses of such data streams can range from managing traffic flows to detecting residents smoking in unauthorized zones. Like datafication more broadly, discussion of Smart Cities has tended to be either “idealistic and/or technical, or critical but lacking in nuance and/or empirical evidence” (Kitchin, 2015, p. 135).

There are many parallels between the use of data in Smart Cities and the rise of the smart campus. For example, in the UK the Joint Information Systems Committee, now known as Jisc, support HE institutions and the Further Education (FE) sector with network and IT services. Jisc have recently argued for the need to evolve the existing infrastructure to support smart campuses. Jisc’s senior co-design manager, James Clay, described the shift as follows:

The smart campus is already here; the technology, sensors and data analysis capability are all available, but it isn’t all joined up and so has limited scope in terms of what we can learn and how we can use the knowledge. What we need now is the intelligent campus, where data from the physical, digital and online environments can be combined and analyzed, opening up vast possibilities for more effective use of learning and non-learning spaces. (Owen, 2018, online)

The intelligent campus relies on more joined-up data to provide the informational material that can shape decision-making processes. Examples include using occupancy data from classrooms and common areas to make effective use of campus space, which could enable staff and students to identify the busiest and quietest times to access the library (Sutjarittham et al., 2019). In conjunction with capturing data on the physical aspects of the campus, providing more relevant messaging through digital signage has also been explored to enable more targeted and real-time communications designed to improve student experience (Owen, 2018). Jisc have also emphasized the need to join up data sets that use RFID tracking of physical spaces correlated with attendance data, VLE data, student records data and data from other online services.

As datafication becomes more ubiquitous, many in HE hope that greater insights can be provided. In addition to measuring behavioral data relating to student learning, attendance or presence in physical spaces, data are now being used to consider other aspects of students’ lives. EDM techniques are being used to conduct sentiment analyses based on data from student assignments, contributions to online discussion and social media footprints (Altrabsheh et al., 2013). These analyses can be used to make judgments about how students are feeling rather than what they are doing, focusing on which aspects of the curriculum students enjoy or whether they feel stressed about their studies. The aim of these approaches is to help staff understand their students better and thereby improve their experience. The learning analytics field has argued that areas such as sentiment analysis, alongside practices such as discourse analysis, Social Network Analysis (SNA) and emotion analysis, provide data that can be used to create a more intimate picture of a student’s affective states during their studies (Sclater et al., 2016).

One rationale for using data to capture more than just digital traces relating to academic performance is the potential to support students’ mental health and wellbeing. It is believed that data can be used to provide enhanced insight in this area, allowing educators, counselors and support services to intervene with students who may be struggling (Rogers, 2019). There have been tragic occurrences in schools, colleges and universities where students have experienced mental health issues, and these cases are not always easy to detect. This has led to campaigning for further access to and utilization of data so that early interventions can take place (Hall, 2018). Although there are grounds for further extraction and analysis of data in these areas, there are also concerns about data accuracy and false positives when targeting students who may be identified for intervention. There are also questions about how such data can be effectively used, and consideration must be given to how tutoring, support and counseling services should react to such insights (Ahern, 2018; Hughes et al., 2018).

The focus on pre-emptive intervention reflects the predictive capabilities that big data can provide (Davis, 2018; Raley, 2013). In recent years, there has been a trend toward future-oriented governance that seeks to anticipate and pre-empt problems (Anderson, 2010; Massumi, 2007). It is no longer sufficient to know what happened or what is happening; rather, policymakers and organizational leaders want to know what will happen next. Within education, predictive analytics and modeling have become more widely used to identify “at risk” students by correlating demographic data with learning data (Kuzilek et al., 2015). These approaches aim to enable educators and support staff to intervene in a more targeted and appropriate manner, leading to increased student engagement and outcomes. Scholars have critiqued the use of datafication practices for the purposes of prediction on the grounds of data accuracy and bias, however even where data may be considered accurate, predictive analytics runs the risk of establishing a form of determinism (Kitchen, 2014; Raley, 2013). For example, people who are predicted to achieve low outcomes as a learner may be prescribed less challenging curriculum and enact the role they have been assigned.

The next step beyond prediction is prescription and automation (Sclater et al., 2016; Sellar and Gulson, 2021). This is exemplified by educational recommender systems that provide the “best” content for students based upon algorithmic calculations. Work in this area is still emergent, but it forms the basis for the pervasive discourse that education technology can personalize learning. This

rationale is inspired by the function of digital platforms such as Netflix, Facebook or Amazon, which serve content for us to consume, purchase or watch based upon datafication of our previous activity. The pursuit of more customized approaches to pedagogy, tailored to the individual learner, have received significant interest. Others in the educational community have raised concerns about the narrow view of learning embedded in the notion of personalization, arguing that this lessens opportunities for collaborative social learning (Harsim, 2017).

The growth of datafication has also created the conditions for the application of artificial intelligence (AI) in education (Williamson and Eynon, 2020). AI, and specifically machine learning, enables the automation of educational processes using algorithms trained on large data sets. A common example is chatbot services that enable students to ask questions and receive answers generated by algorithms. Such services are now being used in a variety of sectors, and this includes the use of intelligent personal tutoring systems (Luckin et al., 2016). Similar AI techniques are being used to facilitate and prompt student discussion during collaborative exercises. AI is also being trialled as a way of offering real time feedback to students on assignment work (Luckin et al., 2016).

Views on the benefits and risks of embedding AI into the fabric of education are mixed. Advocates claim AI will revolutionize our educational institutions, with automation leading to increased efficiency by helping teachers to “... maximize student success and prepare them for the future” (Ayoub, 2020, online). From this perspective, AI can save teacher time by taking over tasks such as providing written feedback on assessments and answering student queries. This will ostensibly enable teachers to spend more time with students engaging in aspects of teaching and learning that cannot be automated. However, others argue that increased automation could easily go beyond complementing educator’s work and lead to the deprofessionalization of teaching and a lack of autonomy for teachers when making decisions about what is best for learners (Schiff, 2021). The question of whether it is possible or desirable for the role of the teacher to be replaced by artificial intelligence has also been raised (Selwyn, 2019). Concerns about technological unemployment created by automation have already been realized in some areas, and while teaching is predicted to be less susceptible to automation (Frey and Osborne, 2017), there is a growing concern that the cost efficiencies gained with the use of AI will be used to reduce the need for human teachers. Some scholars predict that it is a case of when, and not if, academic work, including teaching, will become automated (Ivanov, 2016).

Datafication is now thoroughly intertwined with the development of education technology and provides a basis for the next generation of digital platforms in which AI will be embedded. There is a growing body of literature investigating the use of education technology and raising critical questions; for example, whether it is possible to capture sufficient data and analyze it appropriately to ensure that automated pedagogy is properly contextualized and accessible for all (Bradbury, 2019; Selwyn, 2019; Williamson, 2016). In addition to these technical concerns about the increasing use of education technology and AI, we must ask fundamental questions about the purposes of education and the forms that we want our education systems to take. Do we want data analytics and AI to be deployed in education as a means of automation, or should it be utilized to augment human intelligence (Harasim, 2017). The latter perspective encourages exploration of how we can use AI appropriately, ethically and cautiously in a manner which aligns with societal and community views about the purposes of education.

Emerging challenges

As we have shown, the emergence of datafication in education serves as a critical link between longstanding data-driven accountability agendas in schooling (Lingard et al., 2017; Ozga, 2009), new approaches to data analytics and the increasing application of AI in education. The challenges and effects raised by datafication are diverse and complex. Here we focus on three challenges and one resulting effect:

1. The risk of data-driven automation *deprofessionalizing* teachers;
2. The creation of new *inequalities* for students;
3. The making of new data-driven markets, products and services that contribute to the *privatization* of education; and, as a result,
4. The need for new *regulatory* approaches in response to these challenges.

In our discussion of the three challenges, we broaden our perspective on datafication to include education technologies that generate or require large volumes of data. Datafication goes beyond explicit uses of data by policymakers and educators; it is now embedded in the development and use of digital platforms that hold the potential to disrupt education in significant ways.

Deprofessionalization

Datafication has long contributed to the deprofessionalization of teachers’ work by contributing to neoliberal modes of accountability that have replaced societal trust in teachers as professionals (Ranson, 2003). However, the role of datafication in the creation of new education technologies, including AI, further increases the risk of deprofessionalization. Educators are often disconnected from the design and function of digital platforms and applications that shape pedagogy, curriculum and assessment. These platforms can enter learning spaces as “black boxes” that influence, or even substitute for, aspects of professional judgment and decision-making.

The professional autonomy of teachers to shape the learning experience of students is a longstanding principle enshrined in the 1966 UNESCO and International Labor Organization (ILO) *Recommendation Concerning the Status of Teachers*, which states that teachers:

should be given the essential role in the choice and the adaptation of teaching material, the selection of textbooks and the application of teaching methods ... [and] be free to make use of such evaluation techniques as they may deem useful for the appraisal of pupils' progress. (UNESCO/ILO, 1966, p. 8)

The growth of education technologies developed by agencies, and in contexts, disconnected from teachers' professional judgment undermines this principle. Digital platforms employed in education are not inherently problematic developments for education systems, the teaching profession or students; indeed, they may provide important benefits and opportunities. However, their growth and spread does undermine the profession if teacher judgment does not inform the design and use of new technologies, and when the function of these technologies limits teachers' ability to choose and adapt teaching methods, materials and assessments. Moreover, the sociotechnical imaginaries driving datafication frequently represent the teaching profession as anachronistic and in need of new technologies developed by commercial actors (Williamson and Hogan, 2020).

Data-driven developments in education technology thus change the conditions and capacities for teacher professional autonomy. The data generated by these technologies may be used to evaluate teacher performance, further contributing to the deprofessionalizing effects of accountability agendas. Finally, while automation is considered to present a lower risk of technological unemployment in education, it will change the nature of teachers' work and will have some impact on the number and nature of jobs in the sector. To address this challenge, the teaching profession and education researchers must actively promote and explore complementary uses of data-driven technologies that enhance the work of human educators while challenging the view that it is desirable to replace narrow elements of this work that can be automated. At the same time, educators and researchers should identify how professional judgment can be re-connected to the development and use of new data-driven technologies.

Inequalities

Datafication has the potential to exacerbate existing inequalities that shape educational opportunities for learners and as well as to create new inequalities linked to access and resource allocation. Selwyn et al. (2020) argue that future developments in education technology have significant potential to create new forms of digital inclusion and exclusion. As with other aspects of education, "[t]hose individuals who are well-resourced and have strong educational backgrounds are likely to benefit the most from digital education" (p. 2). Lack of access to expensive devices required to use the digital platforms and applications driven by the datafication of education may create new forms of exclusion. Data analytics enabled by these platforms and applications may also influence the allocation of resources to students and the pedagogies and curricula to which they are exposed.

Datafication may also exacerbate longstanding biases in education and introduce new ones when algorithms are trained on biased data sets and make decisions about students' current and future performance. Slade and Prinsloo (2013) warn that institutions must be mindful of stereotypes and prejudices that algorithms may perpetuate, suggesting that students must be allowed the freedom to act outside of their data profiles to avoid determinism and self-fulfilling prophecies. Similarly, The Open University (no date, online) states in its *Policy on Ethical use of Student Data for Learning Analytics* that "[s]tudents should not be wholly defined by their visible data or our interpretation of that data". This principle reflects concerns about transparency and "black boxing" (Knox, 2017) of algorithmic prediction and decision-making. Students and teachers have the right to understand how data has been aggregated and analyzed to ensure that decisions can be understood contextually and challenged where necessary (Slade and Prinsloo, 2013; Sclater, 2016).

Current datafication practices offer a partial representation of students learning, progress and well-being. Recognizing this partiality and potential bias is a critical element of staff and student data literacy, which is important to ensure that datafication does not exacerbate educational inequalities. Education researchers must hold new technologies up to critical scrutiny to identify risks created in their development and use.

Privatization

The privatization of education can occur endogenously and exogenously (Ball and Youdell, 2008). First, private sector approaches can be imported into the public sector, in order to make education more like a business (privatization of education). Second, the public sector can be opened up to private actors, involving them in the design, management and delivery of different aspects of public education (privatization in education). Often these two approaches go hand in hand. The EdTech industry is increasingly contributing to privatization in education (Hogan and Thompson, 2020; Williamson and Hogan, 2020). This occurs when schools, universities and governments purchase their products or contract out services to EdTech companies through business-to-business transactions.

However, a new type of privatization of education is also emerging, and we argue that it is sufficiently different from warrant its own description: the Uberization of education. Uberization describes the disruption of an existing industry and its business models through the introduction of digital platforms that enable peer-to-peer transactions. Unlike other forms of privatization, Uberization does not seek to alter the model of public education; rather, private actors aim to break this model by introducing new ways to sell education directly to customers (business to customer). Uberization is not new, of course, but the disruption that Uber caused to the taxi industry has not yet occurred in education. With the unbundling of higher education (McCowan, 2017) and the growth of business to customer markets and business models premised on monetizing data and extracting economic rents from users, this form of privatization is likely to increase, particularly in higher education (Komljenovic, 2021). Datafication is central to this development.

The private sector can, and does, make valuable contributions to public education. These contributions can be beneficial if they are transparent and support equitable approaches to sustaining public education for all students. Importantly, and linked to the challenge of deprofessionalization, these contributions should be shaped by teacher professional judgment and held to the profession's standards of practice. However, this is not the case with many new digital platforms, particularly when the education provided by these platforms is developed and delivered by teachers who are not qualified, or these platforms are designed to function in ways that are inscrutable to human users.

Regulation

Broad societal concerns about the rise of datafication, predictive analytics and AI have led to debates about data regulation and ethical use of data. New regulatory frameworks, such as the General Data Protection Regulation (GDPR) in Europe, have been introduced to ensure that the capture, processing, and use of personal data protect rights to privacy. New codes of conduct and ethical frameworks have also been published in the higher education sector to support ethical data use (Tsai and Gasevic, 2017). The recommendations in these documents generally focus on transparency, consent, privacy and access (Slade and Prinsloo, 2013; Sclater, 2016; Open University, no date). A key principle evident in these policies is ensuring that students know how they are being datafied, for what purposes and how they can “opt-out” if they wish to do so.

Regulatory change is also being pursued in relation to the use of AI. For example, the AI Now Institute's regular reports have made recommendations designed to prevent AI from contributing to widening or creating inequalities and exacerbating discriminatory practices in society (Crawford et al., 2019). Bias in facial recognition technology and its inaccuracy in identifying people of color represents one high profile example of such discrimination (Lohr, 2018). The AI Now Institute's reports offer not only critique, but a vision for AI and a set of ethical standards that should guide its use.

Data ownership and the distribution of value generated from data is another area that requires further regulation (Savona, 2019). Savona (2019, p. 18) argues that “data use by non-owners might legitimately require their owner to be compensated, and the non-owner to abide to specific duties”. Currently, only a small number of actors have the ability to extract financial value and benefit from data, often through the creation of assets that are protected by intellectual property rights, and this power imbalance must be readdressed as the economic value of data increases. While attention has focused on questions of data privacy, discussion about the fair distribution of data's value represents a relatively new area of research and policy concern and more work is needed in this area.

Conclusion

In this article, we have provided an overview of datafication in education using higher education as our primary example. While acknowledging the important antecedents to datafication in the 19th and 20th centuries, we have surveyed developments in digital education over the past two decades during which digital datafication has accelerated rapidly. Data-driven education policy and digital datafication have had pernicious and perverse effects in education. Our aim in this article has been to provide an organized overview of new forms of digital datafication, focusing on the challenges it creates, such as deprofessionalization, privatization and the creation of new inequalities, and the need for new kinds of regulation. In particular, we have considered how data-driven changes in education arise not simply from new technological capacities, but from the desires and imaginaries that drive the development of these capacities.

The datafication of education is happening simultaneously with datafication in other sectors and within society at large. This is not a new or isolated trend and it is likely to be irrevocable. The three key challenges that we have discussed—deprofessionalization, inequalities, and privatization—do not cover the full range of issues raised by datafication, but we argue that they are primary areas of concern today. Approaching the problem of datafication from epistemological perspectives on the evidence generated from data or the risks and benefits of technology for humanity remains important, however we believe that the problem also requires constructive critique and experiments with critical approaches to datafication that aim to strengthen public education and counter the harmful effects of surveillance capitalism (Zuboff, 2019). For both pragmatic and strategic reasons, we argue that opposing datafication and the growth of data-driven technologies is likely to be unsuccessful; more creative and engaged approaches are needed to push these developments in directions that can benefit educators, students, communities and societies.

References

- Ahern, S.J., 2018. The potential and pitfalls of learning analytics as a tool for supporting student wellbeing. *J. Learn. Teach. High. Educ.* 1 (2), 165–172.
- Altrabsheh, N., Gaber, M.M., Cosea, M., 2013. SA-E: sentiment analysis for education. *International Conference on Intelligent Decision Technologies* 255, 353–362.
- Anderson, B., 2010. Preemption, precaution, preparedness: anticipatory action and future geographies. *Prog. Hum. Geogr.* 34 (6), 777–798.
- Ashton, K., 2009. That ‘internet of things’ thing. *RFID* 22 (7), 97–114.
- Ayoub, D., 2020. Unleashing the Power of AI for Education. MIT Technology Review. Retrieved on 13 July 2021 from: <https://www.technologyreview.com/2020/03/04/905535/unleashing-the-power-of-ai-for-education/>.
- Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, C., Rodriguez, F., Warschauer, M., Smyth, P., 2020. The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: opening the black box of learning processes. *Int. J. Educ. Technol. High. Educ.* 17 (1), 1–24.
- Ball, S.J., Youdell, D., 2008. *Hidden Privatisation in Education*. Education International, Brussels.

- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Blikstein, P., 2013. Multimodal learning analytics. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pp. 102–106.
- Bradbury, A., Roberts-Holmes, G., 2017. The Datafication of Primary and Early Years Education: Playing with Numbers. Routledge, London and New York.
- Bradbury, A., 2019. Datafied at four: the role of data in the 'schoolification' of early childhood education in England. *Learn. Media Technol.* 44 (1), 7–21.
- Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kazianus, E., Kak, A., Mathur, V., McElroy, E., Nill Sánchez, A., Raji, D., Lisi Rankin, J., Richardson, R., Schultz, J., Myers West, S., Whittaker, M., 2019. AI Now 2019 Report. AI Now Institute, New York. Retrieved on 13 July 2021 from: https://ainowinstitute.org/AI_Now_2019_Report.html.
- Davis, M., 2018. Top 10 Moments from Gartner's Supply Chain Executive Conference. Retrieved on 6th January 2020. <https://blogs.gartner.com/matthew-davis/top-10-moments-from-gartners-supply-chain-executive-conference/>.
- Easterling, K., 2012. An internet of things. *E-flux J.* 31, 1–8.
- Frey, C.B., Osborne, M.A., 2017. The future of employment: how susceptible are jobs to computerisation? *Technol. Forecast. Soc. Change* 114, 254–280.
- Gorur, R., 2014. Towards a sociology of measurement in education policy. *Eur. Educ. Res. J.* 13 (1), 58–72.
- Grek, S., 2009. Governing by numbers: the PISA 'effect' in Europe. *J. Educ. Pol.* 24 (1), 23–37.
- Hall, M., 2018. Jisc and the mental health crisis in our universities. Retrieved on 21st January 2020. <https://www.hepi.ac.uk/2018/07/17/jisc-mental-health-crisis-universities/>.
- Harasim, L., 2017. *Learning Theory and Online Technologies*. Routledge.
- Hartong, S., Förtschler, A., 2019. Opening the black box of data-based school monitoring: data infrastructures, flows and practices in state education agencies. *Big Data Soc.* 6 (1), 1–12.
- Hogan, A., Thompson, G. (Eds.), 2020. *Privatisation and Commercialisation in Public Schooling: How the Public Nature of Schooling is Changing*. Routledge, London and New York.
- Hughes, G., Panjawni, M., Tulcidas, P., Byrom, N., 2018. Student mental health: the role and experiences of academics.
- Ivanov, S.H., 2016. Will robots substitute teachers?. In: *Proceedings of 12th International Conference of Modern Science, Business and Education*. Varna University of Management, Varna, pp. 42–47.
- Jisc, 2021. The Future of Student Recruitment. Retrieved on 17 March 2021. <https://www.jisc.ac.uk/reports/the-future-of-student-recruitment>.
- Kitchin, R., 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE, Los Angeles.
- Kitchin, R., 2015. Making sense of smart cities: addressing present shortcomings. *Camb. J. Reg. Econ. Soc.* 8 (1), 131–136.
- Knox, J., 2017. Data power in education: exploring critical awareness with the "Learning Analytics Report Card". *Telev. New Media* 18 (8), 734–752.
- Komljenovic, J., 2021. The Rise of Education Rentiers: Digital Platforms, Digital Data and Rents. *Learning, Media and Technology*, pp. 1–13. <https://doi.org/10.1080/17439884.2021.1891422>.
- Kuzilek, J., Hlosta, M., Herrmannova, D., Zdrahal, Z., Wolff, A., 2015. OUAlyse: analysing at-risk students at The Open University. *Learning Analytics Review LAK15-1*, 1–16.
- Lai, C.S., Lai, L.L., Lai, Q.H., 2021. *Smart Grids and Big Data Analytics for Smart Cities*. Springer International Publishing.
- Lawn, M., 2011. Governing through data in English education. *Educ. Inq.* 2 (2), 277–288.
- Lingard, B., Sellar, S., Lewis, S., 2017. Accountabilities in schools and school systems. In: Noblitt, G. (Ed.), *Oxford Research Encyclopedia of Education*. Oxford University Press, New York, NY, pp. 1–28.
- Lingard, B., 2011. Policy as numbers: Ac/counting for educational research. *Aust. Educ. Res.* 38 (4), 355–382.
- Lohr, S., 2018. Facial recognition is accurate, if you're a white guy. *N. Y. Times* 9 (8), 283.
- Luckin, R., Holmes, W., Griffiths, M., Forcier, L.B., 2016. *Intelligence unleashed: an argument for AI in education*.
- Lupton, D., 2016. *The Quantified Self*. Polity Press, London.
- Massumi, B., 2007. Potential politics and the primacy of preemption. *Theory & Event* 10 (2). <https://doi.org/10.1353/tae.2007.0066>.
- Mayer-Schonberger, V., Cukier, K., 2013. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt, Boston & New York.
- McCowan, T., 2017. Higher education, unbundling, and the end of the university as we know it. *Oxf. Rev. Educ.* 43 (6), 733–748.
- Open University, n.d.. Policy on Ethical Use of Student Data for Learning Analytics. Retrieved on 13 July 2021 from <https://help.open.ac.uk/documents/policies/ethical-use-of-student-data/files/22/ethical-use-of-student-data-policy.pdf>.
- Owen, J., 2018. What Makes an Intelligent Campus? Education Technology. Retrieved on 13 July 2021 from: <https://edtechnology.co.uk/comments/what-makes-an-intelligent-campus/>.
- Ozga, J., 2009. Governing education through data in England: from regulation to self-evaluation. *J. Educ. Pol.* 24 (2), 149–162.
- Radford, A.W., 2011. Learning at a Distance: Undergraduate Enrollment in Distance Education Courses and Degree Programs. *Stats in Brief*. NCES 2012-154. National Center for Education Statistics.
- Raley, R., 2013. *Dataveillance and countervailance*.
- Ranson, S., 2003. Public accountability in the age of neo-liberal governance. *J. Educ. Pol.* 18 (5), 459–480.
- Roberts-Holmes, G., 2015. The 'datafication' of early years pedagogy: 'If the teaching is good, the data should be good and if there's bad teaching, there is bad data. *J. Educ. Pol.* 30 (3), 302–315.
- Rogers, C., 2019. Why data is an essential part of student wellbeing. Retrieved on 17th January 2020. <https://edtechnology.co.uk/comments/why-data-is-an-essential-part-of-student-wellbeing/>.
- Rosenberg, D., 2013. Data before the fact. In: Gitelman, L. (Ed.), *"Raw Data" is an Oxymoron*. The MIT Press, Cambridge, MA & London, pp. 15–40.
- Rouse, M., 2019. Internet of things" IoT Agenda. Retrieved on 5th October 2019. <https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>.
- Savona, M., 2019. *The Value of Data: Towards a Framework to Redistribute it*.
- Schiff, D., 2021. Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI Soc.* 36 (1), 331–348.
- Sclater, N., Peasgood, A., Mullan, J., 2016. *Learning Analytics in Higher Education*. Jisc, London.
- Sclater, N., 2016. Developing a code of practice for learning analytics. *J. Learn. Anal.* 3 (1), 16–42.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. *Br. Educ. Res. J.* 40 (6), 917–936.
- Sellar, S., 2015. Data infrastructure: a review of expanding accountability systems and large-scale assessments in education. *Discourse* 36 (5), 765–777.
- Sellar, S., Gulson, K.N., 2021. Becoming information centric: the emergence of new cognitive infrastructures in education policy. *J. Educ. Policy* 36 (3), 309–326.
- Selwyn, N., Hillman, T., Eynon, R., Ferreira, G., Knox, J., Macgilchrist, F., Sancho-Gil, J.M., 2020. What's next for Ed-Tech? Critical hopes and concerns for the 2020s. *Learn. Media Technol.* 45 (1), 1–6.
- Selwyn, N., 2019. *Should Robots Replace Teachers?* Polity Press, Cambridge, UK and Medford, MA.
- Siemens, G., Baker, R.S.D., 2012. Learning analytics and educational data mining: towards communication and collaboration. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 252–254.
- Slade, S., Prinsloo, P., 2013. Learning analytics: ethical issues and dilemmas. *Am. Behav. Sci.* 57 (10), 1510–1529.
- Stephens, W.B., 1998. School attendance and literacy: 1750 to the later nineteenth century. In: *Education in Britain, 1750–1914*. Palgrave, London, pp. 21–39.
- Sutjarittham, T., Gharakheili, H.H., Kanhere, S.S., Sivaraman, V., 2019. Experiences with IoT and AI in a smart campus for optimizing classroom usage. *IEEE Int. Things J.* 6 (5), 7595–7607.
- Symonds, P.M., 1940. Trends in educational research. *J. Educ. Res.* 34 (4), 300–302.
- Thompson, G., Sellar, S., 2018. Datafication, testing events and the outside of thought. *Learn. Media Technol.* 43 (2), 139–151.
- Thrift, N., 2014. The 'sentient' city and what it may portend. *Big Data Soc.* 1 (1), 2053951714532241.

- Tsai, Y.S., Gasevic, D., 2017. Learning analytics in higher education—challenges and policies: a review of eight learning analytics policies. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference. Association for Computing Machinery, New York, pp. 233–242.
- UNESCO and International Labour Organisation, 1966. Recommendation Concerning the Status of Teachers. UNESCO, Paris.
- Williamson, B., Eynon, R., 2020. Historical threads, missing links, and future directions in AI in education. *Learn. Media Technol.* 45 (3), 223–235.
- Williamson, B., Hogan, A., 2020. Commercialisation and Privatisation In/of Education in the Context of Covid-19. Education International, Brussels.
- Williamson, B., 2016. Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments. *J. Educ. Pol.* 31 (2), 123–141.
- Williamson, B., 2017. *Big Data in Education: The Digital Future of Learning, Policy and Practice*. Sage, London.
- Wyatt-Smith, C., Lingard, B., Heck, E. (Eds.), 2021. *Digital Disruption in Teaching and Testing: Assessments, Big Data, and the Transformation of Schooling*. Routledge, New York and London.
- Zuboff, S., 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books, London.

Issues: neocolonialism and global languages

Suzanne Majhanovich, Faculty of Education, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Background and context	682
Colonialism's contributions to economic hegemony	682
Neocolonialism: the goal of economic hegemony continues	683
Education as part of the colonialist imperialist project	685
Neo-colonialism and education	686
How global languages and European education systems support colonialist hegemony	686
The power of language in the neocolonial project	687
Global English supports and promotes neoliberalism	688
Final thoughts and summary	689
References	689

Did Colonialism ever end? In our complicated globalized world operating on the basis of international treaties and agreements, remnants of the past colonial system are evident everywhere. In this article, I review the development of colonialism with its economic ramifications. As the great European powers granted independence to their former colonies, a new world order arose, one that has been labeled Neocolonialism with an attendant imbalance of power and stark inequalities between the developed and developing world. The wealthy powerhouses of the world continue to retain some form of control over the developing world usually in the form of economic influence. The connection of globalization and neo-liberalism to the new imperialist project is discussed. The place of education in the development of neocolonialism is also considered along with the role of powerful European languages, primarily English, as a means to making neocolonialism possible and sustaining it.

Background and context

From time immemorial, powerful states have exerted imperialistic control over weaker ones, expanding their territories, creating colonies, using their superior military force to subjugate the people for use as slave labor or being pressed into service for the armed forces. The examples of the Chinese, Roman and Mongol Empires come to mind when one thinks of widespread early colonialization.

The economic drivers of colonialism cannot be underestimated. European countries recognized that territories in Asia and Africa possessed wealthy commodities and resources that they coveted. The European nations began setting up trading posts as early as the 16th Century, as for example the British East India Company established in 1600 or the Dutch East India Company, set up in 1602 in Indonesia. Other European trading nations like France, Belgium, Portugal, Spain and Denmark among others began their inroads into Asia, South America and Africa through the establishment of such trading companies. Soon the process of colonialization had taken root. Colonialism basically involves conquest of a region by another to allow for exploitation of the conquered territory's resources to the profit of the conqueror. The rules and policies of the colonizer are imposed on the conquered. As [Horvath \(1972\)](#) explains, colonialization is a "form of domination—the control by individuals or groups over the territory and/or behavior of other individuals or groups" (p. 46). Some of the best-known sociologists and historians of colonialism such as [Achebe \(1958, 2000\)](#), [Fanon \(1967\)](#), [Nyerere \(1968\)](#), [Guha \(1998\)](#), and [Mann \(2006\)](#) have highlighted the pervasive nature of colonialism, its psychological and cultural effects that have had a strong impact on education, politics and the economy of the colonized territories. This has been carried over into the times of neocolonialism as will be discussed below. At this point, I will focus on economic outcomes of colonialism.

Colonialism's contributions to economic hegemony

Nowadays we speak of globalization, neoliberalism and imperialism driven by the force of the world markets. And yet the global reality today had its roots in the great trading companies of the 16th, 17th and 18th centuries that colonized the world outside of Europe. Much colonial enterprise was controlled not directly by powerful European states but rather by the trading companies they had set up and given free rein to infiltrate the economies all over the world with the goal of enriching the European nations.

Mercantilism was the operating economic principle whereby a colonized nation's economy was regulated and controlled by the mother nation in order to increase state power of the colonizer at the expense of rival powers. As defined by the Encyclopedia Britannica:

Mercantilism contained many interlocking principles. Precious metals, such as gold and silver, were deemed indispensable to a nation's wealth. If a nation did not possess mines or have access to them, precious metals should be obtained by trade. It was believed that trade balances must be "favourable," meaning an excess of exports over imports. Colonial possession should serve as markets for exports and as suppliers of raw materials to the mother country. Manufacturing was forbidden in colonies, and all commerce between colony and mother country was held to be a monopoly of the mother country (McKenna, Encyclopaedia Britannica on-line).

Spain profited handsomely from its colonized areas in the New World. Britain likewise expanded its empire worldwide. Indeed, at one time it could be said that the sun never set on the British Empire. The favored commodities such as gems and precious metals came from Africa and Asia. The sugar plantations in the West Indies also provided a valued commodity such that at one point, Britain considered trading its colony Canada to the fledgling United States for an island in the West Indies with sugar plantations!

By the 18th Century much of the world had been colonized by Britain and other European powers. However, the situation could not last and after the American Revolution, 1775–83, Britain lost its 13 American colonies but retained Canada and several islands in the Caribbean in the so-called new world. Spain and Portugal lost their South American colonies early in the 19th Century. Still, colonialism was alive and well in Africa and Asia. One has only to consider the Berlin Conference of 1884 where Otto von Bismarck convened representatives of the European nations to decide upon the division of African territories to fall under the power of the European powers. The African people had no say in the way their homelands were partitioned (Heath, 2020). The Europeans continued to exploit the people and riches of Africa and Asia while claiming to have accepted the challenge of civilizing the so-called Dark Continent to bring it up to the enlightened state the Europeans enjoyed. Following the exhortation of Kipling's famous poem, "The White Man's Burden" (Kipling, 1899), the imperialists tried to justify their incursions into lives of the colonized as a moral duty of the naturally superior white men to educate the child-like ignorant natives inculcating European values and culture. The colonizers either were unaware of or simply dismissed the indigenous ways of educating the youth in the colonies and considered the local cultures inferior.

Neocolonialism: the goal of economic hegemony continues

After WWII, the European colonizing nations began to grant independence to their former colonies. New fledgling nations gained political independence but were bound by economic ties still controlled by the colonizing powers. In this way, Neocolonialism took the place of the former colonialism. In Africa, at an All African People's Conference (AAPC), the term "Neocolonialism" was first defined in the "1961 Resolution on Neocolonialism". As Afisi (2010) explains, "[Neocolonialism] was described as the deliberate and continued survival of the colonial system in independent African states, by turning these states into victims of political, mental, economic, social, military and technical forms of domination carried out through indirect and subtle means that did not include direct violence" (Afisi, <https://iep.utm.edu/neocolon/>) The same could be said for the former colonies in Asia.

It was Kwame Nkrumah who used the term in his 1963 preamble of the Organization of African States (OAU) Charter and later, as the title of his 1965 book, *Neo-Colonialism, the Last Stage of Imperialism*. (cited in Afisi, 2010).

In his 1965 book, Nkrumah states:

The neo-colonialism of today represents imperialism in its final and perhaps its most dangerous stage

The essence of neo-colonialism is that the State which is subject to it is, in theory, independent and has all the outward trappings of international sovereignty. In reality, its economic system and thus its political policy is directed from outside.

(Nkrumah, 1965).

Others have used the term neo-colonialism to decry the continued imperialism of European powers even after independence to former colonies had been granted. Jean-Paul Sartre, for example in *Colonialism and Neocolonialism* (1964), criticized France for its continued influence in the governments and policies of its ex-colonies, particularly Algeria and called for their emancipation from French interference.

It was largely through economic pressure that former colonial powers retained control over former colonies. In many ways nothing had changed and the practice of mercantilism continued.

To illustrate the similarity of neo-colonial practices to the former mercantile ones, Nkrumah explains:

.... Neo-colonialist control is exercised through economic or monetary means. The neo-colonial State may be obliged to take the manufactured products of the imperialist power to the exclusion of competing products from elsewhere. Control over government policy in the neo-colonial State may be secured by payments towards the cost of running the State, by the provision of civil servants in positions where they can dictate policy, and by monetary control over foreign exchange through the imposition of a banking system controlled by the imperial power.

(Nkrumah, 1965, p. 3).

He also points out that the control may actually be exercised not by the former colonizing states but rather by consortia of financial interests. As will be discussed below, the influence of such organizations as the World Bank and International Monetary Fund with their headquarters in Washington, DC and the World Trade Organization, head office in Geneva, Switzerland are the powers today that drive the neo-colonial project. In addition, growing world powers like China are also exerting their neocolonial influence in previously colonized nations in Africa. Hence, under neo-colonialism, rather than having colonized states subject to their European mother states, the former colonies are now controlled by powerful economic interests whether they be the large global financial powers like the USA and China, or the financial organizations like the World Bank, International Monetary Fund, World Trade Organization or Southeast Asian Bank, or transnational corporations. These powerful states and organizations continue their imperialist agenda through development aid provided to ex-colonies with conditionalities that undermine the political and cultural organisms of the developing countries (See [Serequeberhan, 1998](#)). In recent years many have questioned the actual benefit of aid to African countries and find it often to have been a pernicious element that has destroyed the indigenous education, local cultures and economies, and plunged the nations into extreme poverty ([Easterly, 2006](#); [Easterly and Williamson, 2011](#); [Geo-Jaja and Majhanovich, 2017](#)).

In the 1980s when much of the world was in economic crisis, many former colony states, particularly in Africa were in grave need of financial support. The World Bank and International Monetary Fund stepped up to provide financial backing but saw the opportunity to reform the economies of the developing countries to bring them into line with neo-liberal economic policies. The Structural Adjustment Programs (SAPs) they developed required that in order to secure a loan, the debtor countries would have to accept certain economic reforms to bring their countries into line with free-market systems. The conditions of the loan included some combination of the following features:

- Devaluation of the local currency to reduce balance of payment deficits;
- Reduction of public sector employment, and other spending to reduce budget deficits;
- Privatization of state-owned enterprises and deregulation of state-controlled industries;
- Deregulation in order to attract foreign investment;
- Closing tax loopholes; improve tax collection. ([Halton, 2020](#))

Although neoliberal adherents of such policies saw such adjustments as a means to make the developing countries self-sufficient, there were severe criticisms of SAPs. African critics have seen these economic adjustments as nothing less than a tool of neocolonialism whereby the poor countries are left at the mercy of exploitative investment by multinational corporations. The impoverished developing nations experience increased dependency on the richer nations and international monetary organizations while falling into ever deepening poverty (See [Global Issues, 1998](#) www.globalissues.org/article/3/structural-adjustments-a-major-cause-of-poverty). Although the WB and IMF did recognize that SAPs were not working and were in fact deleterious to the nations they were supposed to be helping, their replacement program called The Poverty Reduction Strategy Papers (PRSP) did not solve the problems and, indeed, exacerbated them through the continued imposition of conditionalities:

The PRSP process is simply delivering repackaged structural adjustment (SAPs). It is not delivering poverty-focused development plans and it has failed to involve civil society and parliamentarians in economic policy discussions. (*PRSPs just PR say civil society groups, Bretton Woods Project Update # 23, June/July 2001*, cited in [Global Issues](#) article on Structural Adjustment, 1998).

Countries that have experienced the deleterious effects of SAPs include Algeria, Benin, Bolivia, Ecuador, India, Niger, Nigeria, Pakistan, Russia, Sudan, Uganda and Venezuela among others ([Global Issues, 1998](#)). Apparently, India leads the world with the number of SAPs arranged. We note that the majority of the countries negatively affected by SAPs are former colonies. The fact that there are also non-former colonies confirms that neocolonialism now extends beyond the formerly colonized world.

The problem of policy rhetoric versus hard reality regarding the effect of SAPs has been well documented, particularly for Africa (See [Geo-Jaja and Magnum, 2001, 2003](#); [Geo-Jaja, 2004](#)). [Abdi \(2013\)](#) has written at length on the effect of SAPs on Africa:

One can look at what SAPs have achieved for Africa in the past 30 or so years, and from any pragmatic perspective, the picture is anything but encouraging. So much so that one might not help but assume the intentions of the IFI's (International Financial Institutions) were never formulated to fulfil their rhetorical representations. Indeed, the very nature of SAPs i.e., what they are at their core, is not compatible with the basic structures and

practices of African life. At their core, SAPs are one important part of globalization, but at a sub-level, they are to fulfill the requirements of the dominant neo-liberal political and economic agenda where the practices of supply side economics in which in the context of non-interventionalist and unregulated market driven relationships, supply creates its own demand (p.354).

Modern globalization has facilitated the growth of neo-colonialism. Indeed, some equate globalization with the new form of neocolonialism (Mazrui, 2002). As Michael Maduagwu has asserted “globalization is only the latest stage of European economic and cultural domination of the rest of the world which started with colonialism, went through imperialism and has now arrived at the globalization stage” (Maduagwu, 1999, p. 65). As Grainger (2017) comments, intuitively one might think positively about modern globalization in the way links have been developed across the planet, and travel, trade and information exchange have all been facilitated. However, easier access has also facilitated corporate power grabs by transnational companies resulting in an extreme rise in wealth inequality. Modern globalization favors economic liberalism but not for all. Bourdieu (2001), cited in Phillipson, 2006 warned of the hegemonic effects of globalization when he stated:

“Globalization” serves as a password, a watchword while in effect, it is the legitimacy mask of a policy aiming to universalize particular interests, and the particular tradition of the economically and politically dominant powers, above all the United States, and to extend to the entire world the economic and cultural model that favours these powers most, while simultaneously presenting it as a norm, a requirement (p.84/Phillipson, p. 352).

A particularly pernicious element of modern globalization is that the colonized countries under neo-colonialism do accept the imposed inequality as a norm. Perhaps they have no choice but to accept what the dominant powers are proposing. Thus, they have willingly entered into the programs of the developed North that have undermined their ability for development, have left them heavily indebted and impoverished with their indigenous culture and languages threatened. The operations and heritage of European colonialism that we experience today has been termed by Quijano as the “Coloniality of Power”. He writes, “Europe’s hegemony over the model of global power concentrated all forms of the control of subjectivity, culture, and especially knowledge and the production of knowledge under its hegemony” (Quijano, 2000). As such the racial, political, economic, cultural, and linguistic orders of the colonized are considered inferior, and certainly of lower value to that of the colonizing powers, European and American (See Quijano, 2007).

Education as part of the colonialist imperialist project

In colonial times, schools in the image of their European counterparts were set up. The issue of education in colonial India provides a salient example. Thomas B. Macaulay’s “Minute Upon Indian Education” of February 2, 1835 argued for English as the medium of instruction in all schooling taught by trained Indian English-speaking teachers. The goal was “to form a class who may be interpreters between us and the millions whom we govern,—a class of persons Indian in blood and color, but English in tastes, in opinions, in morals and in intellect” (Macaulay, Minute, Section 34, 1835). Despite some opposition, Macaulay’s arguments prevailed with the result that traditional Indian education, vocational and science education were eliminated. To this day the medium of instruction in Indian schools is English. Macaulay reflected the racist beliefs of the imperialists of his times declaring for the supremacy of English (or other European languages) over the vastly inferior languages of the colonized “barbarians”, and the intrinsic superiority of European literature and culture. As smugly stated in the “Minute”, “a single shelf of a good European library [is] worth the whole native literature of India and Arabia” (Section 10, Minute, 1835).

It is clear that the goal was not really to educate and enlighten the colonized population but rather to prepare intermediaries among them who would prepare the locals for the menial jobs needed to sustain the economic endeavors of the ruling class.

In addition, colonial education often had the effect of at best undermining, at worst destroying local cultures and languages. One of the most egregious examples comes from the residential school system set up in Canada in 1883 to “solve the Indian problem” through the education of indigenous children. In reality, policy makers sought to assimilate indigenous people through the eradication of cultural (and by extension linguistic), spiritual and economic foundations of First Nation societies (Hill, 2004). There is incontrovertible evidence of the devastation visited upon indigenous peoples in North America including “loss of language, grade retardation, high dropout rates, rampant physical, sexual and emotional abuse, alienation and intergenerational communication breakdown” reported in extensive research (Binda and Lall, 2013, p. 16; see also Aboriginal Healing Foundation, 2009, Canadian Council on Learning, 2007; Royal Commission on Aboriginal Peoples, 1996; Assembly of First Nations, 1994).

Generations of native children were forcibly removed from their parents and transported to far off residential schools where they were educated in the ways of the dominant culture and were severely punished if they dared to speak their native language or practice cultural rituals. For much of the time when residential schools existed, it was various religious orders that ran the institutions, inflicting the most severe disciplinary practices against the helpless children when they strayed from the rigid code of behavior. The children were often malnourished, physically abused, received minimal medical care, and many died. Eventually the government took control of the schools from the orders after becoming aware of the atrocities committed. However, this

system continued operating for over a century still dedicated to assimilation of the native children, still forbidding the use of the indigenous languages and cultural customs. The last residential school in Canada only closed in 1996. The abuse and indignity that the children suffered in these schools has left an indelible mark on them and their descendants. Only recently has the Canadian government begun to address the wrongs perpetrated on the native people through the residential school system. The recent discovery of hundreds of unmarked graves of indigenous children on the sites of former residential school provides a powerful and sobering reminder of the violence perpetrated under the colonial imperialist project to assimilate/eliminate the indigenous populace (See [Aljazeera, 2021](#)). The Canadian Truth and Reconciliation Commission, modeled on the commission set up in South Africa to address the outcomes of the Apartheid system, was set up in 2007 and over the following 6 years traveled all over Canada and took depositions from 6500 witnesses. In 2015 the final report was tabled, with 94 “calls to action” which focused on language and cultural issues ([Government of Canada, 2019](#)). It is noteworthy that the Apartheid system apparently was based on the residential school system in Canada ([Engler, 2013](#)). In any case, these systems reflect colonialist goals of suppressing the colonized and assimilating them to European customs, and values to the detriment of local cultural, spiritual and economic foundations.

Neo-colonialism and education

It is generally accepted that education provides the mechanism whereby people can reach their full potential to take a functional place in society. However, when the formal schooling such as that imposed on formerly colonized nations and their disadvantaged populace serves to inculcate the values and culture of dominant powers to the exclusion, suppression or degradation of local ways of knowing and learning, then the recipients of this form of schooling are badly served. Several problems arise when those organizing the educational systems choose to ignore or disparage indigenous educational practices.

Indeed, many have argued that systems of modern western education with its origins in 16th Century Europe, instead of providing a liberating force of enlightenment have actually served to secure an imperial or colonial world system whereby the values of the European world are privileged over all others and the people of the subject states are inculcated into these understandings of the world, eschewing their native ways of knowing.

As [Baker \(2009\)](#) has argued,

Contrary to the deeply held nationalist assumptions and benevolent liberal intentions among contemporary educators, administrators and policy makers, modern western education systems are part of an imperial/colonial world historical system. State schooling reproduces the dominant social imaginary reflected in the geoculture of the modern/colonial world system. Schooling practices are largely constituted by western knowledges and Eurocentric assumptive worldviews that systemically maintain hierarchical, unequal and violent (symbolic and material) global power/knowledge relations. Through the incorporation and stratification of individuals within particular ethno-cultural intersubjective structures of knowledge and understanding characterized broadly as Eurocentrism, schooling reproduces the modern/colonial world system. State educational institutions are part of a systemic civilizational apparatus in the ongoing redemptive mission of western expansion and domination (pp. 52–53).

The imposition of a European schooling system upon first the colonies, and later former colonial states and developing nations has not had a positive result. To begin with, much of the subject matter imported directly from the European contexts is not at all relevant to the learners. Readers containing stories about European Christian holidays and other activities and pastimes, typical for Europeans or Americans, for example, mean little to children who do not belong to that religious faith nor are aware of the climate and culture and values of the developed North. Moreover, often the curriculum offered has been designed to train only a small percentage of the locals as future elites who would serve as intermediaries between the colonial/global masters and the locals who would only be prepared as unskilled laborers. Since the schooling provided was inevitably in a colonial, European language with little chance of learning in the only language most children understood, namely, their local language, it was ensured that most would never be able to complete the program offered and dropouts would be rampant. It is after all difficult if not impossible to learn concepts when being taught in a language one does not understand (See [Babaci-Wilhite et al., 2012](#); [Majhanovich, 2013](#)).

How global languages and European education systems support colonialist hegemony

There have been years of research that demonstrate that success in schooling is linked to being taught in one’s home language. Some who have focused on the issue in the African context include [Babaci-Wilhite \(2013\)](#); [Benson \(2008\)](#); [Bjork Holmarsdottir \(2009\)](#); [Brock-Utne \(2005, 2007a, 2007b\)](#); [Kimizi \(2009\)](#); [Mbekwa and Nomlomo \(2013\)](#); [Prah and Brock-Utne \(2009\)](#). Similarly, [Kosonen \(2005, 2013\)](#) working in the South East Asian context has argued for the importance of providing particularly primary education in the native language. The issue is not the total removal of English and replacement with local languages; rather it is recommended that early education be taught through the medium of the local languages and English be introduced later, taught as a subject, so that the children can use it in communication and possibly when entering the working world where the global language English is the main medium of communication. However, the prevailing trend continues to privilege European languages

for schooling, particularly English, knowledge of which is viewed as a highly valued commodity. As Pennycook has articulated so well, English being inextricably bound to social, political, economic and cultural contexts worldwide, is part of a power dynamics that reflects the neocolonial world of today. As a result of this, the colonized actually have no choice but to learn English in order to participate in the workings of the modern world (Pennycook, 1998, 2017).

The argument is made that being schooled in English will result in competence in the language and will provide a ticket to a highly paid job in the global market. Despite evidence to the contrary, local parents believe that education in the European languages is preferable to learning in a language their children understand. European-style education, they believe, will help to pull them out of poverty and advance in the globalized world. That this is not the case has been widely demonstrated (see Geo-Jaja, 2013; Mignolo, 2007; Mignolo and Tlostanova, 2007 among many others). Those from the WB who promote European-style education for their client states in the developing world apparently believe that European style education for developing countries is the way out of poverty as well. Accepting education packages published in the developed North were often requirements in the structural adjustment packages negotiated as conditions of much needed financial loans to developing countries. A cynical interpretation might see these educational provisions as a basis for ensuring the continuation of the modern neocolonial condition and perpetuating the inequalities that inevitably result when local cultures, languages and ways of knowing are rejected and in some cases destroyed.

One is struck by the arrogant assumptions of those who impose Western style education on developing areas and are not in the least interested in the opinions and indeed experience of the receivers of this “largesse”. The assumptions about education as it is understood in the developed North and West as the only possibility, all others being inferior or non-existent appear in reports from the OECD and do not apply only to formerly colonized regions. Silova, Millei and Piattoeva, in a recent article entitled “Interrupting the Coloniality of Knowledge Production in Comparative Education: Post-Socialist and Post-Colonial Dialogs after the Cold War” (2021) detail how they as experienced and qualified educational researchers working in post-Soviet countries on curriculum and educational reform routinely had their well-founded, knowledge-based findings dismissed by Western “experts”. Silova et al. quote from OECD reports showing general lack of confidence in the ability of local educators to produce meaningful reform in education claiming that locals were unable to produce meaningful educational research that could lead to educational transformation (p. 9). They as western-trained scholars and educators with first-hand experience in the post-Soviet countries were deemed incompetent because they had different ideas about how to reform and improve education. Silova et al. (2021) comment that the coloniality of the OECD reports is striking. Neocolonialism is alive and thriving all around the world.

The power of language in the neocolonial project

Language is the carrier of culture and values. When educational systems all over the world are offered in one of the prominent European languages, primarily in English but also in French or Spanish, it is inevitable that the values, ways of knowing and perceiving the world as well as the ideologies of the countries of these powerful languages are transmitted to the students (Pennycook, 2017). English holds a particularly privileged position in this regard. The Medium of instruction for elementary and secondary education in especially the formerly British colonized world is commonly English. As Brock-Utne (2013) comments, “no secondary school in Africa has an African language as the language of instruction” (p. 77). Majhanovich (2014) has outlined the proliferation of English education in countries in Africa and the Asia-Pacific region observing that both former colonies of Britain as well as countries that were never British colonies have embraced English education especially at the post-secondary level. She points out the importance accorded to English education particularly in the areas of Business, Science and Technology, and Medicine. Many universities in English speaking nations, especially the US, Australia and the UK, but also to a lesser extent Canada have set up off-shore campuses in Asia and elsewhere to offer these programs with implicit neo-liberal and market-oriented content. However, a large number of the universities in Japan, South Korea, Cambodia, Taiwan and China offer parts of their program in English. In China, for example, the value of knowing English is accepted. English is part of China’s “Four Modernizations” program which has focused on modernizing agriculture, industry, national defense and science and technology along with the promotion of English teaching throughout the country. Instruction in English begins in the third grade; students in secondary school have to work through nine proficiency levels of English finally passing a College English Test to secure entrance to university. Even in University, students may take up to 10% of their credits in English (Feng, 2009). In former colonies such as Hong Kong, Malaysia or Singapore, there is even more evidence of English throughout the educational levels (See Majhanovich, 2014).

Of the powerful European languages, English has emerged as the favored tongue and is often referred to as the *lingua franca* of the world as it is the language most commonly used in cross-cultural or inter-community communication (See for example Hülbauer et al., 2008 or Phillipson, 1992 on Linguistic Imperialism). English is spoken as a first or learned language by more people around the world than any other language. We should not overlook in this regard, the notion of world Englishes as proposed by Kachru (1985). His three concentric circles model depicts first the inner circle for countries such as the UK, USA, Australia, New Zealand and Canada where English is the primary language; the outer/middle circle which includes countries which are former British colonies and where English can be considered as a second language, and the expanding circle for countries like China and Russia with no colonial connection to the UK but where English as a foreign language is spoken in the world business context. Kachru’s model aptly depicts the hegemonic power of English and its connection to the modern neocolonial reality. It goes without saying that those countries in the inner circle enjoy a certain advantage when advancing their political and economic agenda.

The power of the English language is demonstrated not only through its connections to economic and political transactions, it also by necessity is embedded in educational systems globally. It is the language of the academe: as mentioned above, many tertiary level education programs, especially in Business, Science and Technology, and Medicine are offered in English in the top universities globally. The top tiered journals, ranked on international indexes publish primarily in English. Even the UNESCO based Journal, *The International Journal of Education* which publishes abstracts in English, French, German, Spanish and Russian publishes these days almost exclusively in English. At International Conferences, the language of communication is almost always English perhaps with handout summaries provided in the language of the country where the conference is being held. Having to write and present academic research in English presents unfair challenges to those from Kachru's expanding circle (Kachru, 1985), non-English academics and researchers who must learn to write according to norms of Standard English, norms which are often alien to the researchers' culture and methodological practices. It is not surprising that many have objected to the inequitable situation that non-English academics must face. Flowerdew (2007) calls it plainly inequality; Parijs (2007) clearly resents the "free ride" that English speaking researchers are given when it comes to publishing their findings. Yukio Tsuda (1997) laments the many negative consequences for academics as a result of English dominance, namely:

- (1) Linguistic inequality to a great disadvantage of the speakers of languages other than English;
- (2) Discrimination against the non-English speaking people and those who are not proficient in English; and
- (3) Colonization of the consciousness of and identification with the English, its culture and people (Tsuda, 1997, p. 22).

Tsuda has recognized the key role of the English language in the continued colonization (or neo-colonization) of the globe. It is necessary to learn and be proficient in English in order to succeed in academia today, but in so doing one comes to accept the values conveyed through the language and begins to identify with them, thus, becoming colonized by and subject to all that the English-speaking world represents.

Global English supports and promotes neoliberalism

The mission statements of universities, but also policies related to elementary and secondary education are imbued with neo-liberal slogans. The English language has become commodified according to linguists like Heller (2010). She has examined the discourses used in educational and fiscal policies and finds evidence of neoliberalism threaded through them. Whereas traditionally education was engaged in to achieve enlightenment and intellectual growth, now it is a commodity to be bought, its outcomes to prepare one as a member of the global market. Abelman et al. (2009), cited in Pillar and Cho, 2013, p. 30 comment on student identity today in a neoliberal world, noting, "the new model student is an autonomous student-consumer who is responsible for managing his or her own lifelong creative capital development" (p. 232). As early as secondary school, students are encouraged to prepare a "business" plan where they list the courses they will take to prepare themselves to participate in the global market. Economic new-speak has colonized policy making across governments. Holborow (2006) cites Hasan (2003) where she comments on how the English language has been affected under the neoliberal agenda:

She talks of "glibspeak" [which] consists of turning the semantics of ordinary English upside down and globalizing new concepts which are friendly to the ideology of capitalism ... she observes that political words such as *equality*, *freedom*, *liberalization*, and *non-discrimination* are redolent with ideological shifts. She also charts the process of "re-semantization" by drawing attention to the ideological meanings which have attached to *globalization* only recently—like "lower costs of production", "international expansion of companies" and appropriate take-overs

(Hasan, 2003, p. 437 cited in Holborow, 2006 p. 90).

In this way we can see how the English language is a co-opted partner in the neoliberal/neocolonial globalization project.

English dominance and hegemony are reflected everywhere in the globalized world. Indeed, Majhanovich (2013) has argued that the English language as the *de facto lingua franca* of the world contributes vastly to sustaining the neoliberal/neocolonial agenda. English is the working language of a majority of global organizations. It is interesting that the European Union, although recognizing 24 official languages typically works in English, French and German with two-thirds of the policy drafts in English. Its motto "One Europe" is in English only (Phillipson, 2003, 2006). Now that the UK has withdrawn from the European Union, there has been talk of changing the main working language to something else, perhaps Latin, but most wish to continue using English as the most convenient language understood by the majority. The working language of such organizations as the World Bank, World Trade Organization, and the Association of South East Asian Nations is English. These organizations promote privatization, deregulation, decentralization, and *laissez-faire* economic policies, all facets of the neo-liberal agenda that sees market-oriented policies as the operating principles of the globalized world. In order to succeed and compete in the global market, countries have to embrace the tenets of neo-liberalism as expressed by the powerful English-speaking nations of the world. In this way then, the powerful European languages, and especially English, contribute to the continuation of neocolonialism around the globe.

Final thoughts and summary

In this paper the rise of European imperialism and colonialization was reviewed, showing their connection to economic hegemony. Neocolonialism as defined by Nkrumah in 1965 continued the goals of colonialism, especially those related to economic control by world powers and transnational institutions of former colonies and developing nations. It is a violent history where those in power have not hesitated to sacrifice the colonized, their cultures and languages in the goal of imposing European standards and practices to ensure profit and economic dominance.

Education for the colonies as funded by the former imperialist powers played a large role in the colonialist project and this has continued in the neocolonial age. One cannot underestimate the role of global languages, particularly English, in supporting and promoting the particular neoliberal *Weltanschauung* prevalent today. Indeed, policy documents reflect the way that the English language is used to promote the neoliberal agenda, forcing all areas of society to conform to its goals.

What response can be mounted to counteract neocolonialism that is so pervasive and has resulted in so much inequality and disparity of wealth worldwide? Clearly, every attempt at decolonization must be undertaken to work on tearing down the structures of colonization that continue to prevail. Alternative collective approaches similar to those used by indigenous societies could be entertained. Efforts to honor and resuscitate indigenous languages and cultures should be undertaken. Policies that work to sustain the neoliberal market agenda should be carefully examined and revised to eliminate the inherent injustices they promote. Finally, we need to be sensitive to the way the English language is evolving to serve as the handmaiden of the globalizing world powers.

References

- Abdi, A., 2013. Intensive globalizations of African education: re-interrogating the relevance of Structural Adjustment Programs (SAPs). In: Hébert, Y., Abdi, A. (Eds.), *Critical Perspectives on International Education*. Sense Publishers, Rotterdam, pp. 349–362.
- Abelmann, N., Park, S.-J., Kim, H., 2009. College rank and neo-liberal subjectivity in South Korea: the burden of self-development. *Inter Asia Cult. Stud.* 10 (2), 229–247.
- Aboriginal Healing Foundation, 2009. *Response, Responsibility and Renewal. Canada's Truth and Reconciliation Journey*. Author, Ottawa.
- Achebe, C., 1958. *Things Fall Apart*. Heinemann, London.
- Achebe, C., 2000. *Home and Exile*. Oxford University Press, New York.
- Afisi, O.T., 2010. Neocolonialism. *Internet Encyclopaedia of Philosophy*. Retrieved from: <https://iep.utm.edu/neocolon/>.
- Aljazeera, 2021. More Unmarked Residential School Graves Discovered in Canada. <https://aljazeera.com/news/2021/7/13/more-unmarked-residential-school-graves-discovered-in-canada#~:text=>
- Assembly of First Nations, 1994. *Breaking the Silence: An Interpretive Study of Residential School Impact and Healing*. Author, Ottawa.
- Babaci-Wilite, Z., Geo-JaJa, M.A., Lou, S., 2012. Education and Language. A human right for sustainable development in Africa. *Int. Rev. Educ.* 5, 619–647.
- Babaci-Wilite, Z., 2013. A study of escalating debates on the use of a global or local language in education: Tanzania and Malaysia. In: Napier, D., Majhanovich, S. (Eds.), *Education, Dominance and Identity*. Sense Publishers, Rotterdam, pp. 121–132.
- Baker, M., 2009. Situating Modern Western Education Within the Modern/Colonial World System. Retrieved from: <https://www.academia.edu/1518842/Situating-Modern-Western-Education-within-the-Modern-colonial-World-System>.
- Benson, C., 2008. *Language "Choice" in Education*. University of Cape Town, Cape Town. PRAESA Occasional Papers No. 30.
- Binda, K.P., Lall, M., 2013. Decolonizing indigenous education in Canada. In: Napier, D., Majhanovich, S. (Eds.), *Education, Dominance and Identity*. Sense Publishers, Rotterdam, pp. 11–27.
- Bjork Holmardsottir, H., 2009. Cultural identity: implementing a new language in education policy in South African schools. In: Zajda, J., Daun, H., Saha, L. (Eds.), *Nation Building, Identity and Citizenship Education: Cross Cultural Perspectives*. Springer, Dordrecht, pp. 155–170.
- Bourdieu, M., 2001. *Contre-feux 2: Pour Un Mouvement Social Européen (Counter-fire 2: For a Social European Movement)*. Raisons d'agir, Paris, p. 84.
- Brock-Utne, B., 2005. Language-in-education policies and practices in Africa with a special focus on Tanzania and South Africa—insights from research in progress. In: Lin, A.M.Y., Martin, P.W. (Eds.), *Decolonization, Globalisation, Language-in-Education Policy and Practice*, pp. 173–194 (Clevedon: Multilingual Matters).
- Brock-Utne, B., 2007a. Learning through a familiar language versus learning through a foreign language—a look into some secondary school classrooms in Tanzania. *Int. J. Educ. Dev.* 27, 487–498.
- Brock-Utne, B., 2007b. Language of instruction and student performance: new insights from research in Tanzania and South Africa. *Int. Rev. Educ.* 53, 509–530.
- Brock-Utne, B., 2013. Language and liberation. Language of instruction for mathematics and sciences. In: Benson, C., Kosonen, K. (Eds.), *Language Issues in Comparative Education. Inclusive Teaching and Learning in Non-dominant Languages and Cultures*. Sense Publishers, Rotterdam, pp. 77–93.
- Canadian Council on Learning, 2007. *Redefining how Success is Measured in First Nations Inuit and Métis Learning*. Canadian Council on Learning, Ottawa. <https://www.afn.ca>.
- Easterly, W., Williamson, C., 2011. Rhetoric versus reality: the best and worst of aid agency practices. *World Dev.* 39 (11). <https://doi.org/10.1016/j.worlddev.2011.07.027>.
- Easterly, W., 2006. *The White Man's Burden: How the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. The Penguin Press, New York.
- Engler, Y., 2013. Our shame: Canada supported apartheid South Africa. In: *Canadian Dimension*. <https://canadiandimension.com/articles/view/our-shame-canada-supported-apartheid-south-africa>.
- Fanon, F., 1967. *Black Skin/White Masks*. Grove Press, New York.
- Feng, A., 2009. English in China: convergence and divergence in policy and practice. *AILA Rev.* 20, 85–102.
- Flowerdew, J., 2007. The non-anglophone scholar on the periphery of scholarly publication. *AILA Rev.* 20, 14–27.
- Geo-JaJa, M.A., Majhanovich, S., 2017. The Promise and pitfalls of development aid: the elusive goal of aid that helps people help themselves. *World Stud. Educ.* 18 (1), 37–59.
- Geo-JaJa, M.A., Mangum, G., 2001. Structural adjustment as an inadvertent enemy of human development in Africa. *J. Black Stud.* 32 (1), 30–50.
- Geo-JaJa, M.A., Mangum, G., 2003. Economic adjustment, education and human resource development in Africa: the case of Nigeria. *Int. Rev. Educ.* 49 (3–4), 309–326.
- Geo-JaJa, M.A., 2004. Decentralization and privatization of education in Africa. Which option for Nigeria? Special Issue of *Int. Rev. Educ.* 50 (3–4), 309–326.
- Geo-JaJa, M.A., 2013. Education localization for optimizing globalization's opportunities and challenges in Africa. In: Majhanovich, S., Geo-JaJa, M.A. (Eds.), *Economics, Aid and Education. Implications for Development*. Sense Publishers, Rotterdam, pp. 159–181.
- Global Issues, 1998/2001. <http://www.globalissues.org/article/3/structural-adjustment-a-major-cause-of-poverty>.
- Government of Canada, 2019. *Truth and Reconciliation Commission of Canada*. Queen's Printer, Ottawa.
- Grainger, A., 2017. *Neocolonialism: The Dark Side of Globalisation*. Uncommon Ground Media. <https://uncomongroundmedia.com>.
- Guha, R., 1998. *Dominance Without Hegemony: History and Power in Colonial India*. Harvard University Press, Cambridge, MA.

- Halton, C., 2020. Structural adjustment. In: Investopedia. <https://investopedia.com>.
- Hasan, R., 2003. Globalization, literacy and ideology. *World Englishes* 22 (4), 433–448.
- Heath, E., 2020. Berlin Conference of 1884–1885. Oxford African American Studies Center. <https://doi.org/10.1093/acref/978019530173191349265>. Extract.
- Heller, M., 2010. The commodification of language. *Annu. Rev. Anthropol.* 39, 101–114.
- Hill, K., 2004. First Nations Languages and Education in Ontario. <http://www.edu.gov.on.ca/eng/document/curricul.native/index.html>.
- Holborow, M., 2006. Ideology and language: interconnections between neo-liberalism and English. In: Edge, J. (Ed.), *(Re)locating TESOL in an Age of Empire*. Palgrave Macmillan, Houndmills, Basingstoke, Hampshire, pp. 84–103.
- Horvath, R.J., 1972. A definition of colonialism. *Curr. Anthropol.* 13 (No 1), 45–57.
- Hülbauer, C., Böhringer, H., Seidlhofer, B., 2008. Introducing English as a lingua franca (ELF): precursor and partner in intercultural communication. *Syner. Eur.* 3, 25–36. No.
- Kachru, B., 1985. Standards, codification, and sociolinguistic realism: the English language in the outer circle. In: Quirk, R., Widdowson, H. (Eds.), *English in the World: Teaching and Learning of Language and Literature*. Cambridge University Press, Cambridge, pp. 11–30.
- Kimizi, M., 2009. From a Eurocentric to an afrocentric perspective on language of instruction: a view from within. In: *Multilingualism: A Paradigm Shift in African Language of Instruction Policies*. Centre for Advanced Studies of African Society (CASAS), Cape Town, pp. 195–219.
- Kipling, R., 1899. The White Man's Burden: The United States and the Philippine Islands" Poem by Rudyard Kipling to Celebrate the Diamond Jubilee of Queen Victoria originally published in *McClure's Magazine* February 1899.
- Kosonen, K., 2005. Education in local languages: policy and practice in South-east Asia. In: UNESCO. *First Language First: CommunityBased Literacy Programmes for Minority Language Context in Asia*. UNESCO, Bangkok, pp. 239–258.
- Kosonen, K., 2013. The use of non-dominant languages in education in Cambodia, Thailand and Vietnam: two steps forward one step back. In: Benson, C., Kosonen, K. (Eds.), *Language Issues in Comparative Education: Inclusive Teaching and Learning in Non-dominant Languages and Cultures*. Sense Publishers, Rotterdam, pp. 39–58.
- Macaulay, T.B., 1835. Minute Upon Indian Education. <http://www.mssu.edu/projectsouthasia/history/primarydocs/education/Macaulay001.htm>.
- Maduagwu, M., 1999. Globalization and its challenges to national cultures and values: a Perspective from Sub-Saharan Africa. In: Paper Presented at the International Roundtable on the Challenges of Globalization (University of Munich, 18-19, March, 1999. <https://i-p-o.org/Maduagwu>.
- Majhanovich, S., 2013. How the English language contributes to sustaining the neoliberal agenda: another take on the strange non-demise of neoliberalism. In: Majhanovich, S., Geo-JaJa, M.A. (Eds.), *Economics, Aid and Education. Implications for Development Implications for Development*. Sense Publishers, Rotterdam, pp. 79–96.
- Majhanovich, S., 2014. Neo-liberalism, globalization, language policy and practice issues in the Asia-Pacific region. *Asia Pac. J. Educ.* 34 (2), 168–183.
- Mann, C., 2006. 1492: New Revelations of the Americas Before Columbus. Vintage Press, New York.
- Mazrui, A.M., 2002. The English language in African education: dependency and decolonization. In: Tollefson, J.W. (Ed.), *In Language Policies in Education: Critical Issues*. Lawrence Erlbaum Associates, Inc, New Jersey, pp. 267–282.
- Mbekwa, M., Nomlomo, V., 2013. Voices from the classroom: teacher and learner perceptions on the use of the learners' home language in the teaching and learning of school mathematics and science. In: Napier, D., Majhanovich, S. (Eds.), *Education, Dominance and Identity*. Sense Publishers, Rotterdam, pp. 133–152.
- Mignolo, W.E., Tlostanova, M., 2007. The logic of coloniality and the limits of postcoloniality. In: Krishnaswamy, R., Hawley, J.C. (Eds.), *The Postcolonial and the Global: Connections, Conflicts, Complicities*. University of Minnesota Press, Minneapolis, pp. 109–123.
- Mignolo, W.E., 2007. Delinking: the rhetoric of modernity, the logic of coloniality and the grammar of decoloniality. *Cult. Stud.* 21 (2–3), 449–514.
- Nkrumah, K., 1965. *Neo-Colonialism: The Highest Stage of Imperialism*. Heinemann, London.
- Nyerere, J., 1968. *Freedom and Socialism: Writing and Speeches, 1965–1967*. Oxford University Press, London.
- Parijs, P. van, 2007. Tackling Anglophones' free ride: fair linguistic cooperation with a global lingua franca. *AILA Rev.* 20, 72–86.
- Pennycook, A., 1998. *English and the Discourses of Colonialism*. Routledge, London.
- Pennycook, A., 2017. *The Cultural Politics of English as an International Language*. Routledge Linguistics Classics, NY. NY.
- Phillipson, R., 1992. *Linguistic Imperialism*. Oxford University Press, Oxford.
- Phillipson, R., 2003. English for the globe or only for globe-trotters? The world of the EU. In: Mair, C. (Ed.), *The Politics of English as a World Language*. Editions Rodopi B.V, Amsterdam-New York, pp. 19–30.
- Phillipson, R., 2006. Language policy and linguistic imperialism. In: Ricento, T. (Ed.), *An Introduction to Language Policy. Theory and Method*. Blackwell Publishing, Oxford, UK, pp. 342–361.
- Piller, I., Cho, J., 2013. Neo-liberalism as language policy. *Lang. Soc.* 42, 23–44.
- Prah, K.K., Brock-Utne, B., 2009. Multilingualism—An African Advantage. A Paradigm Shift in African Languages of Instruction Policies. Centre for Advanced Studies of African Society (CASAS), Cape Town.
- Quijano, A., 2000. Coloniality of power and eurocentrism in Latin America. *Int. Sociol.* 15 (2), 215–232. <https://doi.org/10.1177/0268580900015002005>.
- Quijano, A., 2007. Coloniality and modernity/rationality. *Cult. Stud.* 21 (2–3), 168–178. <https://doi.org/10.1080/09502380601164353>.
- Royal Commission on Aboriginal Peoples, 1996. Report of the Royal Commission on Aboriginal Peoples. Queen's Printer, Ottawa.
- Rouse, M. & Florian, L., 2012. Inclusive practice project: Final report. Aberdeen: University of Aberdeen. https://www.academia.edu/4432038/The_Inclusive_Practice_Project_Final_Report
- Serequeberhan, T., 1998. Philosophy and post-colonial Africa. In: Chukwudi Eze, E. (Ed.), *African Philosophy: An Anthology*. Blackwell Publishers Ltd, Massachusetts, pp. 9–22.
- Silova, I., Millei, Z., Piattoeva, N., 2021. Interrupting the coloniality of knowledge in comparative education: post-socialist and post-colonial dialogues after the cold war. In: *Contesting Coloniality Special Issue*. <http://www.journals.uchicago.edu/doi/abs/10.1086/690458> 26.
- Tsuda, Y., 1997. The hegemony of English and strategies for linguistic pluralism: proposing the ecology of language paradigm. In: Smith, L., Formen, M. (Eds.), *World Englishes 2000. Selected Essays*. College of Languages, Linguistics and Literature: University of Hawaii and East-West Center, Honolulu, HI, pp. 21–31.

The mangle of contemporary geopolitics

Radhika Gorur^a and Joyeeta Dey^b, ^a Deakin University, Australia; and ^b University of Melbourne, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	691
Posthumanism—a Copernican revolution	692
Education, the new geopolitics and posthumanism	693
Posthumanism and STS in contemporary education	694
Reassembling “context”	694
Datafication	695
Education in the Anthropocene	695
Looking ahead	695
References	696
Further reading	696

Introduction

The recent ransomware attack on Colonial Pipeline, the largest petroleum pipeline in the US, by cyber criminals acting on their own or as agents of a hostile government, illustrates the complexity of contemporary geopolitics. Wars can now almost be waged by a handful of actors attacking private companies, acting alone or on behalf of governments, by crippling critical infrastructures. Cambridge Analytica’s interferences in the 2016 elections in the US had already demonstrated how easily national and international politics can be interfered with, in a significant way and with enormous consequences, by distant private actors whose actions may go unnoticed for a considerable time. Amassing wealth and a nuclear arsenal does not provide any protection against such actions. The display of tanks and missiles and other weapons in parades to commemorate key national celebrations make little sense in these times. Power can no longer be assumed to be held by any nation in any stable and sustained way—power can now be assembled in small scatterings almost anywhere through new forms of action by unknown and invisible actors. A Cambridge Analytica or a Facebook can flex its muscles to create ripples anonymously across the world without visible bloodshed.

In another bizarre turn, as the world struggled with COVID, vaccine shortages and severe economic hardships for many, a couple of billionaires breached another boundary—the earth’s atmosphere—to have a joyride in space. Another arena that was once exclusively the purview of governments and a matter of national pride is now a competition between private individuals. The concentration of wealth in the hands of a few players has had an incredible influence on geopolitics. Commentators have wryly noted how easy it is for Bill Gates to commandeer a meeting with any head of state and to divert massive funds toward his pet projects at the expense of the priorities of international aid organizations. In the process, national priorities and agendas end up being reformulated. Private actors also became involved in the evacuation of Afghanis and foreign nationals from Taliban-held Afghanistan to fill the deficit in the efforts of the most powerful nation on earth. The lines between state and non-state actors are increasingly blurred, and it is becoming harder to identify those capable of mobilizing harm, challenging national and international regulations and rendering even the most powerful nations vulnerable in new ways.

In a completely different way, the COVID-19 pandemic has also tested the extent and limits of sovereignty. While the virus has utterly disregarded national boundaries and traveled freely and promiscuously, it is shutting borders to human travelers. The global shortage of the vaccine has once again disadvantaged the poorest nations, and yet it is clear that no country can be safe until most of the world is vaccinated. Even mighty economies have not been immune to the destructive power of this invisible menace.

The pandemic is linked to another significant contemporary issue—climate change—an imminent existential threat which is already literally redefining geographies. To begin with, the climate crisis is planetwide—and it requires cooperation among nations, rather than competition and warmongering. Economic might, so central to global geopolitics, now has to be reconsidered as a goal, since flooding the planet with more goods that it requires comes with a climate price tag, including actual flooding and other disasters (and the danger of becoming a global pariah). Climate change has summoned nations around the world to a series of summits and induced various agreements and commitments. The crisis has reorganized the list of “good” and “bad” nations not only in terms of their human rights records etc., but also in terms of their contribution to global warming. Most recent reports signal we are past the point of no return and that the destruction and havoc we have wreaked is now impossible to reverse. If humans exercised their agency wantonly in destroying the planet, cutting down forests and filling the atmosphere with harmful emissions, nature is reasserting its agency.

The rise of big data and algorithmic decision-making has further complexified an already very complex scenario, where structures of power and the ways in which agency is allocated or appropriated are no longer clear or predictable. Structural and infrastructural tropes—organization “trees” to denote the structures of power and responsibility, for example—are giving way to more complex

maps of widespread, shifting, networks. The distributed nature of action challenges ideas of hegemonic power amassed at traditional centers of power (Gorur, 2019).

At the same time, technological tools—particularly social media, big data tools and AI-enabled technologies—have magnified the reach of otherwise unremarkable actors—as a result, not only have actors proliferated, they are also very diverse—ranging from individuals to distributed and “headless” groups, to intelligence communities of nations acting behind the scenes to machines learning as they go. The (apparently) abundant and free knowledge available to us is no antidote since our access to information is, increasingly, being mediated by algorithms designed by faceless individuals for commercial or political gain. A few clicks, “likes” and downloads configure the contours of the world we come to know. Again, nowhere is this more poignantly evident than with Cambridge Analytica’s manipulation of information streams via Facebook, or the hold of conspiracy theories that is drawing in not just the fringe elements, but, increasingly, mainstream society. At the same time, the technologies powering an uberized economy are extractive and exploitative—capitalizing on the unsuspecting labor of consumers who give up their data with little knowledge or persuasion, as well as an army of highly regulated individuals employed under intolerable conditions (Crawford, 2019).

These rearrangements are upending the established geopolitical order, heightening a sense of unpredictability, uncertainty and vulnerability and creating intolerable inequities and anxieties. The comfortable projections of econometricians, pollsters and intelligence officials are now less reliable with national economies and politics impacted by distant and unexpected actors and phenomena. Intergovernmental organizations such as the WHO and the UN have been rendered less effective in maintaining moral and practical sway over nations, thanks largely to Trump’s white anting of these agencies, and institutions that held for over 200 years, providing a sense of stability, have demonstrated that they, too, are vulnerable.

Contemporary geopolitics makes compellingly obvious the key roles played by nature and technology in the ordering of societies. Any attempt to examine and theorize geopolitics today must therefore take seriously the agency of these non-human actants. But the acknowledgment of the agency and power of non-human elements such as technology and nature poses a huge challenge for social theory, which, since enlightenment, has been predominantly human-centric. Privilege, interest and intention no longer suffice as explanations. Nor do sweeping theories such as neoliberalism. It is not that old and recognizable structures of hierarchy and power have disappeared. Rather, simultaneously, we now have, alongside the established structures, novel, agile, light and unpredictable ways in which power is mobilized. To properly account for the increasingly critical role of non-human entities in contemporary geopolitics, we require post-humanist theories.

The next section provides a brief overview of posthumanism and following that, we explore how contemporary geopolitics are materializing in education contexts, and how posthumanist theories, particularly STS, might be usefully deployed in studying contemporary education phenomena.

Posthumanism—a Copernican revolution

Decentering the human in contemporary social theory is no less a revolution than Copernicus had provoked six centuries ago, decentering the earth with his heliocentric model of the universe. “Posthumanist” theories, broadly speaking, seek to accommodate the undeniable participation of non-human actors such as technology and nature, and the new modes of exercising power in more-than-human societies. Rosi Braidotti, one of the leading scholars of posthumanism, argues that we are now witnessing a convergence of two major movements—posthumanism, which rejects “the humanist ideal of ‘Man’ as the allegedly universal measure of all things,” and post-anthropocentrism, “which rejects species hierarchy and human exceptionalism” (Braidotti, 2019, p. 32). The project of posthumanism is a political one. It is committed to empirical descriptions that recognize the polycentric and immanent assemblages that produce social phenomena. In this way, Braidotti argues, critical posthumanism provides critical as well as productive ways to study the “posthuman era” which she describes as a combination of the Anthropocene, technological advances and the social and economic inequities that characterize the current moment.

Old geopolitical orders function on an explicitly humanist politics premised on human exceptionalism that highlight the moral supremacy and agency of humans. They restored control over human life to rational human thought and action, in a reaction to the religious authoritarianism and the superstition that marked Medieval Europe. Humans were differentiated from other animals and organisms based on their consciousness, their capacity for benevolence, their language, and so on. Humans were seen as a rational, agentic individuals (Nayar, 2013).

However, this modernistic notion of “pure” and unambiguous human individuals has always been a fiction, as Latour (1993) has argued. Nayar (2013) provides a series of examples from texts by authors such as Ursula Le Guin, Octavia Butler, Kazuo Ishiguro and Marge Piercy, as well as movies such as the Terminator films and Gattaca, where the boundaries of “So-called ‘normal’, self-contained and sovereign humans” that have “specific biological features and abilities, sex, form and functions” is severely troubled. At the same time, he notes that animal-human assemblages abound in fables—from animals that exhibit human capacities to humans that turn into werewolves and so on. These presentations force a reconsideration of the fiction that has come to be defined and essentialized as a “normal” human—and to question the “particular physiology, anatomy, intellectual ability and consciousness” that has come to mark “normalcy” (Nayar, 2013, p. 11). Moreover, stem-cell engineering, cloning, artificial intelligence (AI) and other technologies have translated these fictional hybridities into real life, although, even without these interventions, various forms of microbes colonize our bodies, and the human body is thus already a hybrid assemblage.

The construction of what constitutes a “normal” human comes with repercussions for those cast outside that norm. So, troubling the boundaries of “normal humans” is a political project. Women have historically been left out in considerations of the human in

many societies, seen as lacking in intellect, agency, and so on, as well as “prisoners, comatose patients and asylum inmates, along with ethnic/linguistic/religious minorities, differently abled and developmentally challenged people” and inter- and trans-gender individuals. Particular definitions of “normal humans” provides license to treat those outside these norms as “less than human” (Nayar, 2013).

Posthumanism decenters the placement of humans above other life forms, and at the same time rejects the view of humans as autonomous and fully defined individuals. Instead, it treats “the human itself as an assemblage, co-evolving with other forms of life, enmeshed with the environment and technology” (Nayar, 2013, p. 13). This creates fundamental shifts in how society—and with it, contemporary geopolitics—is understood.

Critical posthumanism may be regarded as emerging from a range of what Braidotti (2019) refers to as interdisciplinary “studies” which are often marginalized in contemporary cognitive capitalism—gender studies, queer studies, postcolonial studies and science and technology studies (STS). With the explicit politics of troubling binaries such as nature/culture and human/non-human, post-humanist approaches are well suited to the study of new geopolitics, highlighting issues related to the inequities that arise from humanistic normativities, and taking into account the work and agency of non-humans such as nature and technology.

STS is one of many posthumanistic approaches. Originating in the 1970s, and inspired by Kuhn, STS scholars began to study science as a social practice, as well as the impact of science and technology on society. Studies of laboratory science (e.g., Latour, 1987; Latour and Woolgar, 1979) and of epistemic practices (Knorr-Cetina, 1999), for example, focused on how objectivity and rationality are fabricated through a series of material-semiotic practices.

The second “S” in STS is sometimes used to denote “society,” and STS has also engaged in the studies of how social orders are produced, and in the working of politics and governance. The rationality, objectivity, individualism, universalism, accountability and so on associated with liberalism are also challenged by STS scholars. Although STS has made a relatively recent entry into education, there is a growing body of scholarship in the field (Gorur et al., 2019). In the next section, we provide brief understandings of how STS contributes to education policy research within contemporary geopolitics.

Education, the new geopolitics and posthumanism

As a key technology for maintaining social order, education has been a site of interest for the global community, particularly since the post-war period of reconstruction and following the independence of many colonized nations in the 1960s. Ensuring that all children have access to education has been considered a moral imperative. Since education is key to securing jobs and gaining access to a better quality of life, not having access to such education results in poverty and leads to inequities and discontent in society—potentially a threat to global security. Intergovernmental organizations such as UNESCO, the World Bank and UNICEF have invested in mapping and measuring literacy and numeracy across the globe for several decades. Major global reform plans such as Education for All, the Millennium Development Goals and the Sustainable Development Goals (SDGs) have brought different nations into relations of obligation and accountability to the global community. Low- and middle-income nations have been sites of intervention for the global community for decades. International funding has often come with strings attached in the form of new accountability measures to be adopted. In recent decades, there have been attempts to recast the hierarchical relations between the donors and the recipient nations, with a stress on the harmonization of goals and a participatory partnership in which nations are not merely “done to” but actively support international agencies in developing an agreed agenda and ensuring they develop the capacity to achieve the agreed goals (Gorur and Addey, 2021). These changing relations are one aspect of contemporary geopolitics in education.

In education, too, there has been a huge proliferation in the number and range of actors—particularly non-state actors. NGOs and INGOs abound in low-income nations, intervening in a number of ways—directly supporting teaching in classrooms, developing curricular resources, providing equipment such as sports materials or computers, supporting teacher education, providing scholarships and so on. There are also a range of organizations training parents to engage in school management committees, advocating for particular changes (such as education in the mother tongue), encouraging parents to send girls to school, setting up programs to prevent students dropping out and so on. These organizations also hold events to raise funds and awareness for their programs, often targeting the same group of citizens. In countries like Cambodia which have a strong presence of aid agencies and over 2000 NGOs as well as national donors and INGOs, development partners aligned their programs and efforts through monthly Education Sector Working Groups led by UNESCO and UNICEF. Boundaries between national and international, and state and non-state actors and activities begin to become fuzzy in these efforts to harmonize and coordinate efforts.

The phenomenon of global actors interacting and influencing national policies is not limited to low- and middle-income nations. The SDGs are a commitment made by almost all the nations of the world, and there is a voluntary submission to UNESCO’s requirements of transparency and accountability by all the signatories. Moreover, the OECD has been influencing high-income nations, in particular, with its own metrics, particularly the Program for International Student Assessment (PISA) as well as various forms of analysis and attendant policy advice.

A range of private, for-profit actors are also very active in the education space, acting in variety of arrangements, including as outsourced providers or in partnership with the public sector. Beyond the very substantial presence of private schools and colleges in the education systems of most nations, there are charter schools, voucher systems and other alliances through which the boundaries between “public” and “private” start to blur.

A key feature of education of recent decades has been the massive increase in datafication, as nations have sought to govern education through mechanisms of legibility, transparency, and accountability, at ever-expanding scales and in ever-increasing detail. New and extended infrastructural networks are enabling “intimate accounting” (Asdal, 2011; Gorur, 2016, 2018), whereby the center is no longer a distant entity seeing and steering from a distance, but a much more immediate presence, penetrating institutions such as schools and forcing them to expose their most intimate details on public sites in the name of transparency, as is the case with Australia’s MySchool (Gorur, 2013). The gaze is at once both intimate and disinterested, creating a range of new anxieties that leave institutions constantly scrambling to appear pleasing, compliant and attractive in the face of extreme scrutiny. What is interesting about this intimate accounting is the production of “non-authority” whereby bodies of authority seemingly give up their powers to inspect and regulate, instead relying on institutions to set their own targets and to regulate themselves through public disclosure of their data (Asdal, 2011). This kind of regulation is also evident at a global scale in wake of the declaration of the SDGs in which nations are subjecting themselves to intimate scrutiny through projects of global metrics.

More recently, there has been a steep rise in use of digital technologies in the education space. Following COVID, the use of online teaching has of course become extremely widespread (but by no means universal—the availability of these technologies cannot be taken for granted for a sizable proportion of the population)—but this is only a very small part of the datafication evident in education. Platforms and apps are prolific in every conceivable activity—as we elaborate in the next section.

Finally, nature and climate change are increasingly becoming actors in education. Students have been striking around the world demanding that the adults in charge do a better job of protecting the planet. Meanwhile, questions are being raised not only about whether the curriculum taught currently in schools does enough to create the necessary understandings about sustainability, but whether systems premised on market-based theories of competition are suitable at a time when collectivism rather than individual gain is the need of the hour. Education thus reflects many of the same ripples that are causing turmoil in global geopolitics.

Posthumanism and STS in contemporary education

Some of the key contributions of STS scholars to the study of emerging geopolitics in education are theoretical and empirical accounts of how power is being mobilized and circulated. We provide a flavor of these studies in this section.

Reassembling “context”

A range of studies in education have dealt with the contemporary phenomena of neoliberalism and globalization. But, as Latour (2005) and Callon et al. (2011) have emphasized, for STS scholars, these phenomena are not explanations of other social phenomena, but are themselves what need to be explained. In other words, they are not transcendental phenomena driving all other events, but are assembled and maintained through a range of practices on multiple sites (see also Ong, 2007). By making these phenomena objects of research, STS approaches render them less monolithic, less black boxed, and more amenable to challenge (Law, 2008).

In contemporary education, “context” comes up in two different ways. On the one hand, global reform agendas and efforts to create global comparability demand a consideration of the diversity of contexts—even if only to neutralize it through various technologies of standardization. Large scale standardized assessments, for example, do not entail the same survey being administered all over the world. Rather, the surveys are tailored to local “contexts” so that the variations help create instruments that work in a similar way across varied contexts. At least, this is the effort. “Context” also became a concern for global development organizations from the late 1900s, when they found that many of their high-investment projects, particularly in Africa, simply collapsed once the international team had left the scene. The realization that projects conceptualized elsewhere and plunked into local “contexts” that operate very differently are prone to failure has prompted a great deal of attention to “capacity building” activities that attempt to reshape the context to support such projects, even as projects are seen as adapting to the local contexts (Addey and Gorur, 2020). Indeed, today funding is dependent on recipient nations agreeing to participate in “capacity building,” and the World Bank sets aside a third of its development budget for capacity building activities. What “capacity building” does is essentially engineer the context so that the reform can take root and flourish—it creates the assemblage that holds the reform efforts together. This effort is a kind of “politics by stealth” that appears to recolonize even as it proclaims decolonization (Gorur and Addey, 2021).

Once context is understood not merely as a backdrop within which processes like datafication and standardization are to be “read” (i.e., context as a “container”), and is instead conceptualized as dynamic and co-produced with the phenomenon being studied, the construction of “context” is rendered into something that can be empirically examined in useful ways. This approach also un-black boxes such phenomena as globalization and which are often taken for granted as monolithic categories, and makes it possible to follow the quotidian and mundane practices (Woolgar and Neyland, 2013) through an attention to the empirical that is not preceded by theoretical orthodoxies. While critics of STS contend that this approach diminishes the significance of pre-existing power structures, STS scholars argue that their insistence on immanence and emergence represents a more hopeful and productive politics which moves existing power structures from the realm of the un-displaceable to “matters of concern”—matters that become available for intervention. As Gorur et al. argue:

now, perhaps more than ever before, STS is needed to make sense of, and to respond to, the unique pressures facing the field of education. As we grapple with increasing inequities, growing standardisation and internationalisation, the deprofessionalisation of teachers, the entrenchment of detrimental neoliberal practices of governance and accountability, the erosion of institutions and the increasingly unquestioning trust in numbers (Porter, 1995), STS provides the means for engaging with the ontological politics of these phenomena.

Gorur et al. (2019).

The politics performed by STS is thus different from the conventional “critical sociologies”—it is “politics by other means” (Latour, 1988).

Datafication

Data, and increasingly, big data and algorithmic governance are shaping education policies at the national and global levels. The global south is seeking to embrace AI as a means to “leapfrog” over infrastructural deficits. In India, AI is being deployed for personalized learning, automated career advice, intelligent tutoring, and automated decision making for such consequential decisions as allocating education loans. These developments raise a range of concerns, particularly those related to loss of privacy, fairness and the perpetuation of bias (Chamuah and Ghildiyal, 2020). Moreover, such moves could shift attention from the paucity of the more basic physical and digital infrastructure which continue to vex those at the ground level. Ambitious digital datafication projects in India, for example, are disproportionately burdening teachers and principals who must find workarounds for the lack of electricity or slow internet (Gorur and Dey, 2020).

Digitization and datafication are also linked to privatization. Among the many issues raised by privatization are how it might clash with the notions of equity, the right to education and education as a public good, especially in the global south; and the colonization of the education space by private actors, including global multinational companies, through such technologies as the Google Classroom. In higher education, platformization is combining with more traditional forces of neoliberalism to reshape universities—perhaps in ways—and with effects—that were never properly anticipated or desired, and with downstream consequences that have not yet been fully imagined or realized (Komljenovic, 2020). Williamson (2020), Selwyn (2019) and others have raised serious issues relating not only to such issues as bias and the reproduction and amplification of structural inequities using AI in education, but also the entry of particular psychological, neurological and emotional logics that are built into AI-mediated platforms and apps that are now ever more widely used in schools and universities.

Education in the Anthropocene

Finally, STS scholars in education are now also engaging with the role of education in the Anthropocene (e.g., Silova, 2020). Whereas earlier efforts focused on including environmental education in schools encouraging schools and universities to practice waste reduction etc., some scholars are taking a broader view and demonstrating how the market-orientation of education in neoliberal times—perpetuated through rankings and competition—fosters an individualistic and greedy, consumptive society. They argue that teaching environmental consciousness within a system that encourages consumption is farcical, and seek other ways—such as systems premised on the collective good—as means to reimagine education in the Anthropocene.

Looking ahead

Although STS-informed posthumanist studies are becoming more accepted in education, and the body of work in this area grows everyday, the core commitments of these approaches remain controversial and still encounter some resistance. The analytical symmetry toward human and non-human agency remains a sticking point, and is often taken as a naïve and anthropomorphic approach to the non-human. At the same time, posthumanistic understandings of “humans” (as human-non-human assemblages) are also hotly contested by critics. At the heart of this is the issue of intention. Critical sociologists, in particular, hold a view of power which seeks to hold wrongdoers responsible (Gorur, 2021). They fear that treating humans and non-humans symmetrically will shift intentionality from humans to non-humans, thus letting off the human perpetrators too lightly.

For STS scholars who treat “individuals” as assemblages of human and non-human actors, the concern is reversed. They fear that focusing on human actors alone would elide the very practical ways in which human intentions are carried through instruments and technologies that present themselves as “neutral” because they are non-human. Unless the ways in which power is mobilized and circulated through assemblages of both human and non-human actors is fully described, the practical means of countering such power will be hidden from view. Even if a particular minister is removed from office, the innocuous instruments that carry their intentions into practices would continue to mobilize that minister’s influence. For the STS scholar with an interest in studying Trumpism, the algorithms of Cambridge Analytica, the complex systems by which Facebook decides when it should intervene, and the symbolic and material affordances of the MAGA slogans and hats are all as integral to Trumpism as Trump himself. A Trump devoid of all these accoutrements would not be the same force, irrespective of his intentions and his growing ambitions. In fact, the Twitter-less Trump in Mar-a-Lago cannot mobilize action in the same way as the Trump who could sign executive orders in the Oval Office or speak from the bully pulpit in the White House Rose Garden. Material entities

vastly multiply human reach and agency, and a lack of attention to these actants not only tells an incomplete story, but it also reduces the scope for intervention.

Increasingly, the obvious ways in which such “actors” as climate change and AI are reconfiguring our lives is making the claim of posthumanist theorists seem less bizarre and more acceptable to more scholars. Although AI today increasingly act invisibly and undetectably, Siri and Alexa talk to us, obey our commands, and run errands for us—and the idea that they are agentic or can “act” and do things is less difficult to swallow, even for scholars from humanistic traditions. This is ironic since it is the anthropomorphism of non-humans that is making their agency more plausible to skeptics. However, while the more obvious “human-like” nature of AI and the increasingly pressing insistence with which climate change is demanding recognition as an agent of change may have made posthumanistic approaches more palatable, it is not clear to what extent the ontological politics of STS and other post-humanistic approaches will come to be accepted—not only by scholars, but also by practitioners and policy makers in education. Currently, the post-humanist diaspora itself appears fractured—seeking unity rather than accepting diversity, seeking to centralize epistemic value to its own particular strand rather than accepting the distributed and polycentric nature of power that is one of the core commitments of these theories.

References

- Addey, C., Gorur, R., 2020. Translating PISA, translating the world. *Comp. Educ.* 56 (4), 547–564.
- Asdal, K., 2011. The office: the weakness of numbers and the production of non-authority. *Account. Org. Soc.* 36 (1), 1–9.
- Braidotti, R., 2019. A theoretical framework for the critical posthumanities. *Theor. Cult. Soc.* 36 (6), 31–61.
- Callon, M., Lascoumes, P., Barthe, Y., 2011. Acting in an Uncertain World: An Essay on Technical Democracy. Inside Technology.
- Chamuah, A., Ghildiyal, H., 2020. AI and Education in India, vol. 10. Friedrich Ebert Stiftung & Tandem Research.
- Crawford, K., 2019. Halt the use of facial-recognition technology until it is regulated. *Nature* 572 (7771), 565–566.
- Gorur, R., Hamilton, M., Lundahl, C., Sundstrom Sjodin, E., 2019. Politics by other means? STS and research in education. *Discourse Stud. Cult. Pol. Educ.* 40 (1), 1–15.
- Gorur, R., Addey, C., 2021. Capacity building as the “third translation”: the story of PISA-D in Cambodia. In: Grek, S., Maroy, C., Verger, A. (Eds.), *The World Yearbook of Education 2021*. Routledge, London & New York, pp. 72–87.
- Gorur, R., Dey, J., 2020. Making the user friendly: the ontological politics of digital data platforms. *Crit. Stud. Educ.* 62 (1), 67–81.
- Gorur, R., 2013. My school, my market. *Discourse Stud. Cult. Pol. Educ.* 34 (2), 214–230.
- Gorur, R., 2016. The performative politics of NAPLAN and my school. In: Thompson, G., Sellar, S., Lingard, R. (Eds.), *National Testing and Its Effects: Evidence From Australia*. Routledge, London, pp. 30–43.
- Gorur, R., 2018. Escaping numbers? Intimate accounting, informed publics and productions of authority and non-authority. *Sci. Technol. Stud.* 31 (4), 89–108.
- Gorur, R., 2019. Old power, new power and ontological flattening. In: Mølsted, C., Pettersson, D. (Eds.), *New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research*. Routledge, London and New York, pp. 66–82.
- Gorur, R., 2021. Opening the black box of peer review. In: Addey, C., Piattova, N. (Eds.), *Intimate Accounts of Education Policy Research*. Routledge, London & New York, pp. 62–76.
- Knorr-Cetina, K., 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press, Cambridge, Mass.
- Komljenovic, J., 2020. The future of value in digitalized higher education: why data privacy should not be our biggest concern. *High Educ.* <https://doi.org/10.1007/s10734-020-00639-7>.
- Latour, B., Woolgar, S., 1979. *Laboratory Life: The Social Construction of Scientific Facts*. Sage Publications, Beverly Hills.
- Latour, B., 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press, Cambridge, Mass.
- Latour, B., 1988. *The Pasteurization of France*. Harvard University Press, Cambridge, MA.
- Latour, B., 1993. *We Have Never Been Modern*. Harvard University Press.
- Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.
- Law, J., 2008. On sociology and STS. *Sociol. Rev.* 56 (4), 623–649.
- Nayar, P.K., 2013. *Posthumanism*. Polity Press, London.
- Ong, A., 2007. Neoliberalism as a mobile technology. *Trans. Inst. Br. Geogr.* 32 (1), 3–8.
- Porter, T.M., 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press, Princeton.
- Selwyn, N., 2019. *Should Robots Replace Teachers? AI and the Future of Education*. John Wiley & Sons.
- Silova, I., 2020. Anticipating other worlds, animating ourselves: an invitation to comparative education. *ECNU Rev. Educ.* 3 (1), 138–159.
- Williamson, B., 2020. New digital laboratories of experimental knowledge production: artificial intelligence and education research. *Lond. Rev. Educ.* 18 (2), 209–220.
- Woolgar, S., Neyland, D., 2013. *Mundane Governance: Ontology and Accountability*. Oxford University Press, Oxford.

Further reading

- Critical Posthumanism, Braidotti, R., 2018. A theoretical framework for the critical posthumanities. *Theor. Cult. Soc.* 36 (6), 31–61.
- Special Issue of *Discourse: Cultural Politics of Education* 40 (1)—Politics by Other Means? STS and Research in Education.

FOURTH EDITION

International Encyclopedia of Education

Diversity, Democracy, and Social
Justice in Education

VOLUME
2

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Roland Sintos Coloma
Jeong-eun Rhee



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 2

**Diversity, Democracy, and
Social Justice in Education**

VOLUME 2 EDITORS

Roland Sintos Coloma and Jeong-eun Rhee



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

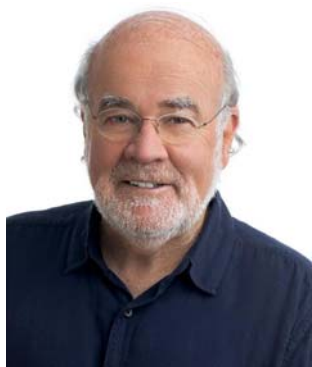
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



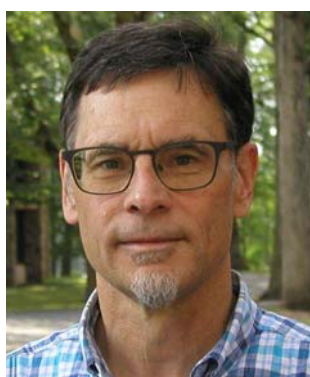
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 2: Diversity, Democracy, and Social Justice in Education

Coloma Roland Sintos and Jeong-eun Rhee

© 2023 Elsevier Ltd. All rights reserved.

Introduction: Struggling Against Virus of Hate

We continue to live, work, teach, learn, and advocate in challenging times. We write our Introduction to this encyclopedia volume on the third year of the COVID-19 global pandemic that has devastated many families, communities, and countries. The pandemic's latest official toll count includes over 565 million confirmed cases and over 6.3 million deaths around the world while true deaths are estimated around 15 million ([World Health Organization, 2022](#)). The pandemic has exacerbated and put into sharp relief existing social inequalities between and within countries, particularly along singular and intersecting lines of socioeconomic class, gender, and race ([Miguel and Mobarak, 2021](#)). The most adverse effects of the pandemic have negatively impacted lower-middle-income and low-income countries located in so-called developing and least developed nations in the Global South. In so-called developed countries in the Global North, like the United States, low-income and poor communities—especially those inhabited predominantly by Indigenous and racialized minority populations—have disproportionately higher rates of COVID-19 morbidity and mortality. During the pandemic, a larger share of women became unemployed or underemployed, thereby disrupting their professional and economic advancement and further distressing female-headed households ([OECD, 2020](#)). When the lockdowns took effect, individuals with limited or no available child or elderly care, personal transportation, Internet, and/or computers struggled to maintain or gain employment and to access services and resources, including medical and mental health support as well as food and household essentials. Hence, the global pandemic revealed in glaring ways the ongoing and deepening fissures of social injustice along national and class lines that are also highly gendered and racialized ([World Health Organization, 2022](#)).

During public health crises that adversely impact the lives, employment, education, and well-being of many individuals, families, and communities, a different form of virus also has been surging: Racism. In the United States where we live and work as Education university professors of Asian descent, anti-Asian racism and violence have become more rampant and visible ([Coloma et al., 2021](#)). Nearly 11,500 incidents of assaults, harassment, and civil rights violations against Asian Americans have been reported from March 2020 to March 2022 ([Stop AAPI Hate, 2022](#)). Asian American women, elderlies, and/or lesbians, gays, bisexuals, transgenders, and queers (LGBTQs) have been targeted, and many other incidents are not reported. Amid the global COVID-19 pandemic, Asian American and Asian international students, faculty, and staff experience racist discrimination and microaggression in US colleges and universities. As plans for returning to campus unveil for the Fall 2022 term, many continue to be gravely concerned about our safety and well-being in light of bigoted rhetoric and actions.

Since US higher education is our professional context, we begin this Introduction with the ongoing COVID-19 pandemic, its global ravages and consequences, and the increasing and more visible anti-Asian racism and violence in the United States. Our editorial position for this volume is that an approach to issues of diversity, democracy, and social justice in education needs to be contextualized and historicized. In this sense, it is important to note that the entire encyclopedia—and, for our discussion, this volume—have been further developed, written, and published during a worldwide public health, economic, and social crisis that we have not experienced during our lifetime. Unfortunately, it may not be surprising that the already vulnerable and marginalized are the ones disproportionately burdened by the pandemic. At the same time, Asians—especially those in the diaspora—have become targets of pandemic-fueled hate and harm. As Asian Americans with filial and transnational ties to the Philippines and South Korea, we struggled and continue to struggle with public and educational spaces where our physical and emotional well-being could be quickly compromised. Therefore, we discuss how our positionalities are intricately intertwined with ways we have worked on this volume.

Asians and Asian Americans are largely invisible in US higher education ([Teranishi, 2010](#)). (Please note that we as editors will use the pronoun “we” in addressing issues relevant to Asians and Asian Americans in the United States as part of a collective embodied experience and racialized solidarity). We appear in statistical reports and publicity photos on institutional diversity and demographics, yet receive little, if any, culturally relevant curriculum, programming, and services. We bolster student enrollment, selectivity, and graduation data and generate revenue through tuition dollars, especially from Asian international students who are aggressively recruited to come to the United States. When the Trump administration restricted visas and enrollment of international students in July 2020, more than 200 institutions of higher education as well as 70 higher education associations wrote and testified to challenge this policy. Yet many colleges and universities do not provide us with culturally and linguistically affirming resources to enhance our learning and campus life. Asian and Asian American faculty substantially contribute to institutional missions of scholarship and teaching, but are scarce in leadership positions in higher education. In fact, the lack of meaningful response from many institutions on the rising anti-Asian violence and racism raises serious questions if they truly care for their campus community members of Asian heritage.

Stereotyped as model minority, Asians and Asian Americans are often perceived as academically successful and self-reliant. Hence, the misconception is that we do not need policies, curriculum, and sociocultural, linguistic, and mental health support in US higher education. We are more likely than white and other minority students to have considered or attempted suicide, but receive mental health services at a lower rate ([Chen et al., 2019](#)). Where Asian Americans are visible is in a spin-off narrative of the model minority stereotype that pits us against other racialized minority groups. The ongoing debates and lawsuits on

affirmative action and race-conscious admissions point to the number of Asian American students in highly selective universities and position us against other minorities, particularly African Americans, even though a majority of Asian Americans support race-conscious equity policies (Lee et al., 2021). Such ill-informed zero-sum game and wedge politics are used to perpetuate anti-Black racism and to erase the white supremacist strategy of divide-and-conquer. They also fail to account for the long history of Asian-Black coalitions for racial justice in the United States and globally (Boggs, 2016; Zia, 2000).

To this day, Asian Americans are perceived as perpetual foreigners who do not belong in the United States. Regardless if we are recent immigrants or citizens whose families have lived in the United States for generations, our status as legitimate Americans remains conditional, suspect, or invalidated. As perpetual foreigners and outsiders, Asian Americans become expedient scapegoats for US domestic and international woes. Trump's failure to contain the COVID-19 pandemic resorted to a tactic of deflecting attention and describing the virus as "Chinese virus" and "kung flu." Many agree that Trump's repeated use of racist monikers has contributed to increased anti-Asian bigotry and attacks. Words matter. When conveyed by those in control and authority, racist discourses wield a powerful influence over individual beliefs, perceptions, and actions. Through racist eyes, all Asians and Asian Americans are the same: Yellow Perils (Tchen and Yeats, 2014).

Yet Asian Americans are extremely diverse. Estimated at 22.86 million in 2019, we constitute 6.8% of the US population and are the fastest growing racialized group (Budiman and Ruiz, 2021). With about 14% of US residents born outside of the country, Asia is the second largest source of immigrants, following Latin America. The top 10 Asian ethnic groups by order of their current population size in the United States are: Chinese, Asian Indian, Filipino, Vietnamese, Korean, Japanese, Pakistani, Thai, Hmong, and Cambodian, who are also implicated in complex inter/intra-ethnic geopolitical histories. Differences within and between these groups range from generations, socioeconomic statuses, and educational levels to languages, religions, and political perspectives. In fact, Asian Americans have the largest socioeconomic divide among all racial groups (Hassan and Carlsen, 2018). Yet in spite of our heterogeneity, Asian Americans are largely considered a monolithic minority group.

Stereotyping Asian Americans as model minorities who could allegedly accrue academic and economic capital unlike other racialized minorities does not mean that Asian Americans—and those who are perceived as Asians—are less subject to racist suspicions, microaggressions, and outright violence. This pandemic has repeated histories of the ways anti-Asian racism disenfranchises our status and belonging as Americans in the United States. United Nations Secretary-General António Guterres (2020) indicated that the "pandemic continues to unleash a tsunami of hate and xenophobia, scapegoating and scare-mongering," and consequently called on political leaders, educational institutions, media companies, and civic groups to "act now to strengthen the immunity of our societies against the virus of hate."

Our involvement with local and other initiatives to address anti-Asian racism as well as racial discrimination against other minoritized groups established part of our intellectual, political, and educational frames in conceptualizing this encyclopedia volume. A larger part also drew from our long-standing personal, professional, and academic interests in diversity, democracy, and social justice matters from transnational and postcolonial perspectives (e.g., Coloma, 2017; Rhee, 2009, 2013). As researchers and educators who work to challenge Euro-American intellectual, political, socioeconomic, and educational hegemony, we tasked ourselves to formulate an encyclopedia volume that could represent, as much as possible, a wide range of critical voices from different fields of education and beyond and from various institutions and countries, especially from the Global South and areas where English is not a dominant language. Toward these ends, we were guided by feminist works on recognition, redistribution, and refusal, which coincidentally mapped on to our conceptualizations of diversity, democracy, and social justice in education (Fraser, 2003; Simpson, 2017).

Nancy Fraser (2003) examines historical and contemporary struggles for social justice from the vantage points of recognition and redistribution. On the one hand, the goal of the politics of recognition is to promote "a difference-friendly world, where assimilation to majority or dominant cultural norms is no longer the price of equal respect" (p. 7). On the other hand, the goal of the politics of redistribution is to advance "a more just distribution of resources and wealth" (p. 7). Conventionally understood as a marked distinction between identity politics (based on race, gender, sexuality, etc.) and class politics, Fraser does not view the politics of recognition and redistribution as separate and incommensurable. Rather, she considers them as interconnected and complementary. Following Fraser, we consider the ways diversity, democracy, and social justice in education can be understood through recognition and redistribution.

In the politics of recognition, the identification of identities or subject positions functions as important markers when addressing not only visible and not-so-visible differences but also varying power relations embedded in dynamics of oppression, privilege, and complicity at the individual, interpersonal, and institutional levels (National Equity Project, n.d.). Such identification is tied to who is represented or who is at the table, if you will, when attending to certain social, institutional, and educational issues and concerns. In such inquiries, educators, researchers, and advocates generally take up the perspectives of race and ethnicity, Indigeneity, gender, sexuality, socioeconomic class, dis/ability, language, religion and spirituality, nationality and citizenship, migration and diaspora, and geography (urban, suburban, and rural as well as Global North and Global South)—a list which informed our initial conceptualization of this volume's first two sections on Key Concepts and on Identity and Difference (more on these sections below). Such perspectives have been examined, expounded, and extended by theoretical frameworks, such as critical race theory, ethnic studies, multiculturalism, feminist theory, queer theory, trans theory, gender and sexuality studies, Marxism, disability studies, crip theory, diaspora studies, transnationalism, postcolonial theory, decolonization, and settler colonialism, among many others. While we recognize the salience of single issue/lens analysis, we are also committed to intersectional analysis inspired by feminist of color and queer of color theorizing that accounts for the ways identities, experiences, and social and institutional conditions are intertwined. In our case, we are not only Asian American, queer man or heterosexual woman, able-bodied, raised Christians, married,

urban-dwellers, or US citizens, as if such identities can be separate. We are a combination of these identities and so much more. Some of these identities can be detected in visible ways, while others cannot be. Yet these subject positions shape how we see the world and how others view and treat us. Hence, in the politics of recognition, it is important to have various identities represented to lift up diverse voices and perspectives, especially when some have been historically and continuously marginalized, dismissed, or erased.

In the politics of redistribution, the availability, allocation, and dissemination of material and symbolic opportunities and resources for educational and community needs is the central priority for educators, researchers, and advocates and is largely considered through equal or equitable mechanisms (National Equity Project, n.d.). On the one hand, equality means that individuals and groups are provided with the same opportunities and resources regardless of historical and current needs, conditions, and preferences. On the other hand, equity considers the historical and ongoing circumstances of individuals and groups, so that they are provided with opportunities and resources needed to reach their desired goals and highest potential. The persistent debates between equality and equity informed our preliminary considerations of this volume's next two sections on Primary and Secondary Education as well as Adult and Higher Education as we tackled formal institutional spaces of education (more on these sections below). Moreover, these debates manifested in governmental responses to the COVID-19 global pandemic, especially in accessing and distributing vaccines. Wealthy countries in North America, Europe, and the Gulf have been ordering, bulk-buying, and hoarding vaccines, prioritizing the needs of their own citizens, a political-economic and public health move dubbed as "vaccine nationalism" (Bollyky and Bown, 2020). These countries have struck deals and contracts with pharmaceutical companies manufacturing COVID-19 vaccines with orders that often surpass the dosage needs for vaccine and booster shots of their populations. Since the start of the pandemic, wealthy countries, such as the United States, have disproportionately received far more vaccines, compared to low-income and middle-income countries. Instead of equal or equitable access and distribution of life-saving resources, they display and exercise outright greed and selfishness in the name of national safety, well-being, and self-protection, while disregarding the lives and conditions of the world's majority, as if all our lives, well-being, and quotidian realities are not globally interconnected. They leverage their national wealth and global status to acquire resources that enable their people to live while letting others die.

When we bring together the politics of recognition and redistribution to analyze the global supply and distribution of life-saving vaccines against the COVID-19 pandemic, a long-standing pattern emerges, once again: benefitting are wealthy countries, mostly in the Global North, with predominantly white populations; and continuing to suffer are low-income and middle-income countries, mostly in the Global South, with predominantly Black and Brown populations. The consequences of the global pandemic have also disproportionately burdened women, people with disabilities, as well as communities and countries with limited or no access to computers and broadband technologies. Ultimately, the ongoing global pandemic has made visible and exacerbated existing inequities based on individual and intersecting fault-lines along geography, socioeconomic status, race, gender, dis/ability, and technology. However, must we accept such realities or inevitabilities? Must we consent to how the world continues to operate, especially who benefits and who suffers in these conditions? How might we refuse such seemingly universalizing norms?

Moving beyond the politics of recognition and redistribution, Audra Simpson offers refusal as a necessary critical intellectual, methodological, and even pedagogical intervention to matters related to diversity, democracy, and social justice in education. For Simpson (2017), "refusal" rather than recognition is an option for producing and maintaining alternative structures of thought, politics, and traditions away from and in critical relationship to states" (p. 19). Grounded in historical and ongoing struggles for Indigenous sovereignty, onto-epistemology, and revitalization, Simpson's work has inspired us to further promote Indigenous and decolonizing thought and praxis in local and global movements and to raise questions to what has been normalized as universal and status quo. Such questioning has led us as editors to consider informal and nonformal education spaces as sites for teaching and learning, which manifested in this volume's last two sections on Informal Education and on Activism and Social Movement (more on these sections below). Simpson's concept of refusal has also led us to inquire into the ways the very concepts of the encyclopedia volume may be already embedded within hegemonic Euro-American-centric use and understanding. For example, Indigenous philosophy and politics call attention to the conventional use and understanding of citizen and citizenship as defined by settler colonial nation-states responsible for Indigenous genocide, land theft, and broken treaties. They challenge who gets to be defined as proper citizens with rights and under whose terms, as seen in settler colonial nation-states such as the United States, Canada, and Australia. What are the terms and definitions? And under what and whose terms? These questions further complicate even critical terms, such as diversity, democracy, and social justice, when they are primarily conceptualized from Euro-American and Global North perspectives and do not fully account for the ways Indigenous and Global South frameworks make sense of and utilize them. As an international encyclopedia, this is certainly an important area that we hope could be taken up further in future volumes. As editors, we acknowledge our own conceptual limitations and blind spots and mark particular topics that we yearned to include in this volume but hopefully could be addressed in future iterations.

The encyclopedia volume on Diversity, Democracy, and Social Justice in Education contains 46 chapters, not including this Introduction. The chapters have been organized into six thematic sections on Key Concepts, Identity and Difference, Primary and Secondary Education, Adult and Higher Education, Informal Education, and Activism and Social Movement. We offer a brief overview of each section below.

The first section on Key Concepts serves as the overarching guide for the entire volume by identifying and defining important terms and providing insightful examples and case studies. Beginning with concepts underpinning this volume—diversity, democracy, and social justice—this section takes up additional concepts, such as culture, decolonial, citizenship, migration, human rights,

international development, ethnonationalism, and globalism, and elaborates on their definitions, uses, and applications through examples from Global North and Global South contexts. Future volumes could add specific chapters on other concepts, such as anti-oppression, inclusion, ontology, epistemology, methodology, pedagogy, and even praxis as part of this section. The volume's second section on Identity and Difference builds upon the first section by explicating how the key concepts are connected to particular markers of identity and difference. Hence, this section begins with an explanation on identity and its connections to diversity and education. It is followed by chapters on intersectionality, race and gender, feminism (especially on feminisms of color), girls with disabilities in the Global South, Indigeneity, and political orientation. Largely due to the proliferating mobilization of intersectional analysis at local, regional, national, and global levels, it has become a significant approach utilized by researchers, educators, and advocates alike. For this section, future volumes could add specific chapters on race and ethnicity (especially as understood in various geographical contexts), socioeconomic class, gender and gender identity, sexuality and sexual orientation, dis/ability, linguistic minority, religion and spirituality, geographical location (urban, suburban, rural), and geopolitics (North–South and South–South).

The volume's next sections emphasize formal institutions and processes of education. The third section on Primary and Secondary Education attends to cultures, policies and politics, curriculum, teaching and learning, students and educators, and broader social contexts of schools. Starting with curriculum and policies on global and international education, the section elaborates on school curricula, such as Afrocentric education, comprehensive sexuality education, and multilingual education; school disciplinary areas, such as literacy, ecology, and environmental education; school experiences, such as those of LGBTQ students and educators; and school contexts, such as elite schools and participatory democracy. Future volumes would do well adding chapters on multicultural education, Indigenous education, ethnic studies in schools, special education and disability studies, religious schooling, and career/technical education. The fourth section on Adult and Higher Education underscores the education of adults in colleges, universities, and workplaces as well as the wider global contexts of such education. With teaching responsibilities in the field of teacher education, we also included the preparation of preservice and in-service teachers in this section. Hence, the chapters in this section tackle work and learning, minorities and the knowledge-economy, global university rankings, the interconnectedness of global and Indigenous education, and critical topics for teacher education, such as Truth and Reconciliation research, racial literacy, and diversity discourses in non-Euro-American contexts.

The volume's last two sections draw inspiration from informal and nonformal spaces and actors of education, thereby disrupting mainstream understanding of schools, colleges, and universities as the only legitimate institutions and mechanisms of teaching and learning. The fifth section on Informal Education offers robust chapters on community-based education, media education, arts-based education, and museum education. It accounts for women and youth as informal educators and attends to virtual spaces, environmental citizenship, and shadow (tutoring) education where informal teaching and learning take place. Future volumes could elaborate on additional topics, such as mothers/parents as informal educators, popular culture, health, and wellness, especially informal and nonformal education during global pandemics. The sixth and final section highlights Activism and Social Movement as important educational sites and mechanisms for diversity, democracy, and social justice. Similar to Informal Education, Activism and Social Movements often fall outside of formal institutionally sanctioned realms of curriculum, teaching, and learning. Although many schools, colleges, and universities may include mission statements on diversity, equity, and inclusion as well as pedagogical priorities to critical thinking and community engagement, they generally do not condone activism and social movements deemed radical or disruptive of the status quo. Hence, when students and educators participate in advocacy and organizing, especially to address inequities and injustices in schools, higher education, and society at large, they inevitably encounter conflicts with those in positions of power and authority. As the chapters in this section reveal, some activists may emphasize local issues, but their impact and consequences resonate more expansively at the state, national, and even global levels. Other activists may take up an educational topic as their major concern, yet it is intrinsically connected to broader social dynamics and vice versa. Authors in this section address a remarkable range of activism and social movements, from Black Lives Matter and educational access for undocumented migrants, to young women, college students, and teachers fighting for rights, resources, and legitimacy. They offer concrete and informative examples of activist strategies, personal and collective narratives, successful campaigns and difficult lessons learned, internal and external pressures and conflicts, and the joys and challenges of advocating for genuine change and transformation. Future volumes could provide examples of other social movements, especially from Indigenous and Global South contexts, so that we could gain lessons of education and solidarity from activists and educators "from below" (Della Porta et al., 2006).

As this volume's editors, we are deeply grateful to the encyclopedia's editors-in-chief Fazal Rizvi, Rob Tierney, and Kadriye Ercikan for inviting us to take the lead on this important collection on Diversity, Democracy, and Social Justice in Education, which will be the first of its kind in the encyclopedia's history and various editions. We appreciate their inspiring and compelling vision for the entire international encyclopedia, and their ongoing guidance and support as we undertook this daunting initiative during a very challenging period. Paula Davies and the technical and production team from Elsevier are the consummate professionals who, with meticulous care and good humored patience, have truly been instrumental in the beautiful and timely completion of this volume (and this entire encyclopedia). Lastly, we could not have undertaken this volume without the brilliance and commitment of the 80 authors—scholars, educators, students, and activists from around the world—who willingly shared their wisdoms, experiences, and perspectives with us as editors and now with you as readers. When we initially conceptualized this volume, we charted a rather ambitious agenda with a robust list of chapter topics and potential authors, reflecting educational questions, sites, and voices from various countries across continents. The mighty 80 who generously accepted our invitations submitted and revised their

chapters and worked through feedback and follow-up, while grappling with the quotidian uncertainties and dire repercussions of the COVID-19 global pandemic on their personal and professional lives. What the authors gave us, ultimately, are precious gifts of their knowledge and time. And we are forever thankful and blessed for their gifts. It is our sincere hope that what we are able to offer here individually and collectively can invite, incite, and inspire ongoing praxis of diversity, democracy, and social justice in education and beyond.

References

- Boggs, G.L., 2016. *Living for Change: An Autobiography*. University of Minnesota Press.
- Bollyky, T.J., Bown, C.P., 2020. The tragedy of vaccine nationalism: only cooperation can end the pandemic. *Foreign Affairs* 99 (5), 96–108.
- Budiman, A., Ruiz, N.G., 2021. Asian Americans are the fastest-growing racial or ethnic group in the US. Pew Res. Center. Retrieved on July 22, 2022 from. <https://www.pewresearch.org/fact-tank/2021/04/09/asian-americans-are-the-fastest-growing-racial-or-ethnic-group-in-the-u-s/>.
- Chen, J.A., Stevens, C., Wong, S.H.M., Liu, C.H., 2019. Psychiatric symptoms and diagnoses among U.S. college students: a comparison by race and ethnicity. *Psychiatr. Serv.* 70 (6), 442–449.
- Coloma, R.S., 2017. “We are here because you were there”: on curriculum, empire, and globalization. *Curric. Inq.* 47 (1), 92–102.
- Coloma, R.S., et al., 2021. Reckoning with anti-Asian violence: racial grief, visionary organizing, and educational responsibility. *Educ. Stud.* 57 (4), 378–394.
- della Porta, D., Andretta, M., Mosca, L., Reiter, H., 2006. *Globalization From Below: Transnational Activists and Protest Networks*. University of Minnesota Press.
- Fraser, N., 2003. Social justice in the age of identity politics: redistribution, recognition, and participation. In: Fraser, N., Honneth, A. (Eds.), *Redistribution or Recognition?: A Political-Philosophical Exchange*. Verso, pp. 7–108.
- Guterres, A., 2020, May 20. Tweet: #COVID19 Does Not Care Who We Are, Where We Live, or What We Believe. Retrieved on July 22, 2022 from. <https://twitter.com/antonioguterres/status/1258613180030431233?s=20>.
- Hassan, A., Carlsen, A., 2018. How “crazy rich” Asians have led to the largest income gap in the US. *New York Times*. Retrieved on July 22, 2022 from. <https://www.nytimes.com/interactive/2018/08/17/us/asian-income-inequality.html>.
- Lee, J., Wong, J., Ramakrishnan, K., 2021. Asian Americans Support for Affirmative Action Increased Since 2016. Retrieved on July 22, 2022 from. <http://aapidata.com/blog/affirmative-action-increase/>.
- Miguel, E., Mobarak, A.M., 2021. The economics of the COVID-19 pandemic in poor countries. *Annu. Rev. Econ.* 14. <https://doi.org/10.1146/annurev-economics-051520-025412>.
- National Equity Project. (n.d.). Homepage. Retrieved on July 22, 2022 from <https://www.nationalequityproject.org/>.
- OECD, 2020. Women at the core of the fight against COVID-19 crisis. In: *OECD Policy Responses to Coronavirus (COVID-19)*. <https://doi.org/10.1787/553a8269-en>. Retrieved on July 22, 2022 from.
- Rhee, J.E., 2009. International education, the new imperialism, and technologies of the self: branding the globally educated elite. *Multicult. Educ. Rev.* 1 (1), 55–82.
- Rhee, J.E., 2013. The neoliberal racial project: the tiger mother and governmentality. *Educ. Theory* 63 (6), 561–580.
- Simpson, A., 2017. The ruse of consent and the anatomy of “refusal”: cases from indigenous North America and Australia. *Postcolonial Stud.* 20 (1), 18–33.
- Stop AAPI Hate, 2022. Two Years and Thousands of Voices: What Community-Generated Data Tells Us About Anti-AAPI Hate. Retrieved on July 22, 2022 from. <https://stopaapihate.org/year-2-report/>.
- Tchen, J.K.W., Yeats, D. (Eds.), 2014. *Yellow Peril!: An Archive of Anti-Asian Fear*. Verso.
- Teranishi, R.T., 2010. *Asians in the Ivory Tower: Dilemmas of Racial Inequality in American Higher Education*. Teachers College Press.
- World Health Organization, 2022. WHO Coronavirus (COVID-19) Dashboard. Retrieved on July 22, 2022 from. <https://covid19.who.int/>.
- Zia, H., 2000. *Asian American Dreams: The Emergence of an American People*. Farrar, Straus and Giroux.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 2

Cash Ahenakew

Department of Educational Studies, University of British Columbia, Ponderosa Commons North (Oak House), Vancouver, BC Canada, Musqueam Territory

Vanessa Andreotti

Department of Educational Studies, University of British Columbia, Ponderosa Commons North (Oak House), Vancouver, BC Canada, Musqueam Territory

Awo Okaikor Aryee-Price

Education for Liberation Network, United States

Sanzhar Baizhanov

Michigan State University, East Lansing, MI, United States

Abigail B Bakan

Department of Social Justice Education, Ontario Institute for Studies in Education (OISE), University of Toronto, Toronto, ON, Canada

Bianca J Baldrige

Harvard University, Cambridge, MA, United States

Jeff Bale

University of Toronto, Toronto, ON, Canada

Jason Beech

School of Education, Culture and Society, Faculty of Education, Monash University, Clayton, VIC, Australia

Deevia Bhana

Gender and Childhood Sexuality, University of KwaZulu-Natal, School of Education, Durban, South Africa

Megan Boler

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Mark Bray

Centre for International Research in Supplementary Tutoring (CIRIST), Faculty of Education, East China Normal University, Shanghai, China; and The University of Hong Kong, Pokfulam, China

Nazir Carrim

Wits School of Education, University of the Witwatersrand, Johannesburg, South Africa

Jasone Cenoz

University of the Basque Country, UPV/EHU, Donostia-San Sebastian, Spain

Yoonjung Choi

Ewha Womans University, Seoul, South Korea

Elaine Coburn

International Studies, Centre for Feminist Research, York University, Toronto, ON, Canada

Barbara Comber

University of South Australia, Education Futures, Adelaide, SA, Australia

Jerusha O Conner

Villanova University, Villanova, PA, United States

Jonathan W Crocker

College of Education, Texas Christian University, Fort Worth, TX, United States

Cindy Cruz

University of Arizona, Tucson, Arizona, United States

George JS Dei

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Brandon Edwards-Schuth

Washington State University, Pullman, WA, United States

Joy Esboldt

Graduate School of Education, University of California, Berkeley, CA, United States

Hoda Gharib

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Ratna Ghosh

McGill University, Faculty of Education, Montreal, QC, Canada

Durk Gorter

University of the Basque Country, UPV/EHU-Ikerbasque, Basque Foundation for Science, Donostia-San Sebastian, Spain

Alison Happel-Parkins

University of Memphis, Memphis, TN, United States

Tang T Heng

Nanyang Technological University, National Institute of Education, Singapore

Annette Henry

Language and Literacy Education, Faculty of Education, Institute of Gender, Race, Sexuality and Social Justice, University of British Columbia, Vancouver, BC, Canada

Li-Ching Ho

University of Wisconsin-Madison, Madison, WI, United States

Adam Howard

Education Department and Interdisciplinary Studies, Waterville, ME, United States

Lisa Howell

University of Ottawa, Ottawa, ON, Canada

M Francyne Huckaby

College of Education, Texas Christian University, Fort Worth, TX, United States

Salma Ismail

Faculty of Humanities, School of Education, University of Cape Town, Cape Town, South Africa

Susan Jagger

Toronto Metropolitan University, Toronto, ON, Canada

Carl E James

York University, Toronto, ON, Canada

Adeerya Johnson

Institute for Gender, Race, Sexuality and Social Justice, Faculty of Arts, University of British Columbia, Vancouver, BC, Canada

Denisha Jones

Sarah Lawrence College, Bronxville, NY, United States

Arlo Kempf

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Jane Edith Kenway

Monash University, Melbourne University, VIC, Australia

Rita Kohli

University of California, Riverside, CA, United States

Marissa Largo

Visual Art and Art History, York University, Toronto, ON, Canada

Andrew Leland

Kean University, Union, NJ, United States

Zeus Leonardo

Graduate School of Education, University of California, Berkeley, CA, United States

Leonel Lim

Nanyang Technological University, National Institute of Education, Singapore

Helen Longlands

IOE, UCL's Faculty of Education and Society, Department of Education, Practice and Society, University College London, London, United Kingdom

John Lupinacci

Washington State University, Pullman, WA, United States

Luis Fernando Macías

California State University, Fresno, CA, United States

Alma Maldonado-Maldonado

Departamento de Investigaciones Educativas, CINVESTAV, Mexico City, Mexico

Melinda M Mangin

Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

Aparna Mishra Tarc

York University, Toronto, Canada

Erica Neeganagwedgin

Western University, London, ON, Canada

Nicholas Ng-A-Fook

University of Ottawa, Ottawa, ON, Canada

Xuan Thuy Nguyen

Institute of Interdisciplinary Studies, Carleton University, Ottawa, ON, Canada

Amy North

IOE, UCL's Faculty of Education and Society, Department of Education, Practice and Society, University College London, London, United Kingdom

Bathseba Opini

Educational Studies, Faculty of Education, University of British Columbia, Vancouver, BC, Canada

Corinna D Ott

University of California, Riverside, CA, United States

Wanda S Pillow
*Gender Studies, School for Cultural & Social
 Transformation, University of Utah, Salt Lake City, UT,
 United States*

Raechele Pope
*University at Buffalo, Amherst, NY,
 United States*

Murray Print
University of Sydney, Sydney, NSW, Australia

Hua Que
*Faculty of Education, Memorial University of
 Newfoundland, St. John's, NL, Canada*

Stewart Riddle
*School of Education, University of Southern Queensland,
 Springfield Central, QLD, Australia*

Whitney Neumeyer Roach
*College of Education, Texas Christian University, Fort
 Worth, TX, United States*

Chazz Robinson
University at Buffalo, Amherst, NY, United States

Tricia Seow
*National Institute of Education, Nanyang Technological
 University, Jurong West, Singapore*

Riyad A Shahjahan
*Michigan State University, East Lansing, MI, United
 States*

Hongxia Shan
*The Faculty of Education, University of British
 Columbia, Vancouver, BC., Canada*

Sandra L Sirota
*Human Rights Institute, University of Connecticut,
 Storrs, CT, USA*

Nerida Spina
*Queensland University of Technology,
 Faculty of Education, Kelvin Grove, QLD,
 Australia*

Sharon Stein
*Department of Educational Studies, University of
 British Columbia, Ponderosa Commons North (Oak
 House), Vancouver, BC Canada, Musqueam Territory*

Melissa Marini Švigelj
*University of California at Santa Cruz, Santa Cruz,
 California, United States*

Paul Tarc
Western University, London, ON, Canada

Felisa L Tibbitts
*Netherlands Institute of Human Rights (SIM),
 Department of Law, Economics and Governance,
 Utrecht University, Utrecht, The Netherlands*

Sandra Annette Toro
*United States Department of Education, Office of
 Elementary and Secondary Education (OESE),
 Washington, DC, United States*

Amanda Trigiani
*Ontario Institute for Studies in Education, University of
 Toronto, Toronto, ON, Canada*

Rita Turner
Independent Scholar, Baltimore, MD, United States

Daniella Santos Vieira
*Bowling Green State University, Bowling Green, OH,
 United States*

Lisa D Weems
*Miami University of Ohio, Department of Educational
 Leadership, Oxford, OH, United States*

Handel Kashope Wright
*University of British Columbia, The Center for Culture,
 Identity and Education, Faculty of Education,
 Vancouver, BC, Canada*

Yao Xiao
*University of British Columbia, The Center for Culture,
 Identity and Education, Faculty of Education,
 Vancouver, BC, Canada; and The Faculty of Education
 and The Social Justice Institute, University of British
 Columbia, Vancouver, BC., Canada*

This page intentionally left blank

CONTENTS OF VOLUME 2

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 2: Diversity, Democracy, and Social Justice in Education</i>	<i>xix</i>
<i>Contributors to Volume 2</i>	<i>xxv</i>

KEY CONCEPTS

Diversity <i>Nazir Carrim</i>	1
Democracy and education <i>Stewart Riddle</i>	8
Social justice and education: addressing inequities in schooling <i>Elaine Coburn and Carl E Jame</i>	13
Culture <i>Yao Xiao and Handel Kashope Wright</i>	21
Decolonial and education <i>Aparna Mishra Tarc</i>	31
Citizenship <i>Abigail B Bakan</i>	38
Curriculum, citizenship, migration <i>Yoonjung Choi</i>	46
Human rights education <i>Felisa L Tibbitts and Sandra L Sirota</i>	53
Education, equality and international development <i>Helen Longlands and Amy North</i>	65
Knowledge/power and education: ethnonationalism, globalism, and social justice <i>Alma Maldonado-Maldonado</i>	74

IDENTITY AND DIFFERENCE

Education, identity and diversity <i>Ratna Ghosh</i>	83
White woman: or, the abused abuser's role in U.S. educational stratification <i>Zeus Leonardo and Joy Esboldt</i>	93
Feminist genealogy & policy futures <i>Wanda S Pillow</i>	104
Indigeneity <i>Erica Neeganagwedgin</i>	109
Intersectionality in the Caribbean and Africa <i>Annette Henry, Bathseba Opini, and Adeerya Johnson</i>	116
Decolonizing research with girls with disabilities in the Global South <i>Xuan Thuy Nguyen</i>	124
Political orientation <i>Murray Print</i>	134

PRIMARY AND SECONDARY EDUCATION

K-12 global/international education: dancing with "diversity, democracy, and social justice" <i>Paul Tarc</i>	139
Policy transfer in education <i>Jason Beech</i>	151
Afrocentric education: traditions, politics and possibilities <i>George JS Dei and Arlo Kempf</i>	161
Comprehensive sexuality education: young masculinity, cultural relevance and context matters <i>Deevia Bhana</i>	171
Queer in school: the experiences of youth and educators in the United States <i>Andrew Leland and Melinda M Mangin</i>	179
Multilingualism at school and multilingual education <i>Jasone Cenoz and Durk Gorter</i>	188
Primary and secondary education: literacy and social justice <i>Nerida Spina and Barbara Comber</i>	195
Ecocritical pedagogies and curriculum <i>John Lupinacci, Brandon Edwards-Schuth, Alison Happel-Parkins, and Rita Turner</i>	202
Exploring currents and confluences in environmental education <i>Susan Jagger</i>	210
Reading studies of elite schools through an intersectional lens <i>Jane Edith Kenway and Adam Howard</i>	220
Education and participatory democracy <i>Whitney Neumeyer Roach, Jonathan W Crocker, and M Francyne Huckaby</i>	227

ADULT AND HIGHER EDUCATION

Work and learning: power, politics and approaches <i>Hongxia Shan, Hua Que, and Yao Xiao</i>	236
Minoritized individuals and knowledge-economy <i>Chazz Robinson and Raechele Pope</i>	244
Power/knowledge: global and Indigenous education <i>Sharon Stein, Cash Ahenakew, and Vanessa Andreotti</i>	251
Global university rankings and geopolitics of knowledge <i>Riyad A Shahjahan and Sanzhar Baizhanov</i>	261
Truth and then reconciliation research: an emerging field of educational studies <i>Lisa Howell and Nicholas Ng-A-Fook</i>	272
Teacher education for social change: shifting from race-evasiveness to racial literacy <i>Corinna D Ott and Rita Kohli</i>	283
Diversity beyond its discourse: examining the discourse of diversity in non-Anglo-Saxon contexts <i>Tang T Heng and Leonel Lim</i>	289

INFORMAL EDUCATION

Community-based education <i>Bianca J Baldrige</i>	297
Media education: history, frameworks, debates and challenges <i>Megan Boler, Amanda Trigiani, and Hoda Gharib</i>	301
Arts-Based Educational Research <i>Marissa Largo</i>	313
Museum education: building the capacity of society and its members <i>Sandra Annette Toro</i>	320
Women's informal learning in poor communities, social movements and virtual spaces <i>Salma Ismail</i>	323
Informal environmental citizenship education <i>Li-Ching Ho and Tricia Seow</i>	330
Youth as educators <i>Melissa Marini Švigelj and Cindy Cruz</i>	337
Shadow education <i>Mark Bray</i>	346

ACTIVISM AND SOCIAL MOVEMENTS

From #BlackLivesMatter to Black Lives Matter at school: how the fight for Black Liberation entered the schoolhouse <i>Denisha Jones and Awo Okaikor Aryee-Price</i>	352
Young women and activism and dissident global citizenship <i>Lisa D Weems</i>	363

Undocumented activism: a bate-papo~testimonio <i>Luis Fernando Macías and Daniella Santos Vieira</i>	374
Language, race, and legitimacy: the Toronto Teachers' Federation's battle against heritage-language programs in Toronto schools, 1982–1986 <i>Jeff Bale</i>	385
Social movements and college student activism <i>Jerusha O Conner</i>	398

Diversity

Nazir Carrim, Wits School of Education, University of the Witwatersrand, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Diversity and race relations	1
Diversity and inclusivity	3
Difference	4
Difference ontologically	4
Difference epistemologically	5
Toward a more socially just world	6
References	6
Relevant websites	7

Introduction

This chapter looks at the development of diversity from its race relations original emphasis and the ways in which it has developed. Diversity in the 21st century is more than a race relations issue. Diversity is now viewed not only in terms of race, but includes ethnicity, class, gender, region, sexual orientation, ability, and religion, inter alia. As will become clearer in this chapter, diversity now is more about recognizing difference, and working constructively with difference in order to move toward a more social just and inclusive world. As will also be clearer later in this chapter, it is, thus, important to also recognize that diversity needs to be viewed as being linked to difference and inclusivity.

This chapter is written from a South African perspective. The South African experience provides an important perspective of the dangers in using diversity and how it can construct the basis for segregation, as was the case under apartheid in South Africa. While this is the perspective and lens from which this chapter is written, it engages with the way diversity has also been used in other country contexts.

The first part of this chapter deals with diversity and race relations. The second part deals with diversity and inclusivity, and the third part deals with diversity and difference. The conclusion of this chapter argues that it will be more productive and constructive to deal with difference, since the continued use of diversity will perpetuate the historical baggage of meaning that has come to inform diversity.

Diversity and race relations

Diversity in South Africa has a specific meaning due to the history of apartheid. The racist segregation of South Africans was justified by apartheid ideologues and in policies and legislation by reference to cultural and ethnic diversity (Rose and Tunmer, 1975). Apartheid claimed that four “race” groups existed: “whites”, “coloreds”, “Indians” or “Asians”, and “blacks”. The differences among these “race” groups were argued to have been not due to “race” but due to cultural differences among them. Apartheid justified itself on the basis of cultural diversity (Carrim, 1995, 2018; Cross, 2004; Verwoerd, 1953).

Two important things emerge from the above. First, diversity was linked to culture and “race”. Second, are the problems related to defining people in racial terms.

Apartheid South Africa used diversity as linked to culture and “race” as a means to deal with race relations. South Africa was not alone in doing this. Apartheid was officialized in 1948 when the white National Party government came into power. Soon after World War II, the United Kingdom, for example started dealing with “race” relations issues to deal with what it then called the colored immigration problem (Rex and Mason, 1986). Approaches to dealing with “race” relations issues included “assimilationism”, “integration”, “multiculturalism”, and “anti-racism” (Carrim, 1995). Dealing with the issues related to “race” relations, thus, became linked to culture and how to deal with the diversity of cultures became the focus.

Apartheid South Africa’s answer to this issue was to organize South Africa systemically and systematically on the basis of racist segregation. The Population Registration Act ensured that people were classified as white, colored, Indian/Asian, or black. The Group Areas Act ensured that people lived in areas that were designated for their racial groups. The different education Acts ensured that schools catered for differently classified racial groups in segregated spaces. The entire edifice of Apartheid was thus constructed to ensure segregationism and racism and it justified this on the basis of recognizing differences in cultures among South Africa’s people. The then Prime Minister of South Africa H.F Verwoerd’s speech in the “white” Senate of parliament said the following to justify Bantu Education (an education for “blacks”).

It is the policy of my department that education should have its roots entirely in the Native areas and in the Native environment and Native community. There Bantu education must be able to give itself complete expression and there it will have to perform its real service. The Bantu must be guided to serve his own community in all respects. There is no place for him in the European community above the level of certain forms of labour. Within his own community, however, all doors are open.

For that reason, it is of no avail for him to receive a training which has as its aim absorption in the European community while he cannot and will not be absorbed there. Up till now he has been subjected to a school system which drew him away from his own community and practically misled him by showing him the green pastures of the European but still did not allow him to graze there. This attitude is not only uneconomic because money is spent on education which has no specific aim, but it is even dishonest to continue with it. The effect on the Bantu community we find in the much-discussed frustration of educated Natives who can find no employment which is acceptable to them. It is abundantly clear that unplanned education creates many problems, disrupts the communal life of the Bantu, and endangers the communal life of the European.

For that reason, it must be replaced by planned Bantu Education. In the Native territories where the services of educated Bantu are much needed, Bantu education can complete its full circle, by which the child is taken out of the community by the school, developed to his fullest extent in accordance with aptitude and ability and thereafter returned to the community to serve and to enrich it.

Verwoerd (1953, Speech as Minister of Native Affairs, June 7, 1954).

The link between racist segregationism and cultural diversity was made explicit by Apartheid. Cross (2004) described apartheid as an extreme form of multiculturalism. Its extremeness was due to the blatant inequality that Apartheid constructed between “whites”, on the one hand, and “coloreds”, Indian/Asians, and blacks, on the other hand; and the extreme segregationism it imposed. For Apartheid, it was separate and unequal.

The second issue raised above was the problem of defining people in racial terms. Apartheid defined people as “white”, “colored” Indian/Asian, and “black”. Colored was meant to refer to people of mixed descent. Yet, although “white”, “colored” and “black” referred to skin pigmentation, “Indian/Asian” are geographical descriptions and not a description of color or skin pigmentation. In the Senate speech of Verwoerd referred to earlier one will also note that “whites” are referred to as “Europeans”.

Despite the inconsistencies of these racial/racist descriptions of South Africans, when it comes to differences in cultures, two things seem to happen. The first is that “whites” are projected as if they have a homogenous culture. Differences between the French, Germans, Dutch, etc. are ignored. Second, it is only when it comes to people who are not “white” that cultural differences are seen to be at issue. This leads to the important argument made by anti-racists and anti-apartheid (and anti-colonial) activists in their responses to diversity and viewing diversity as linked to cultures.

Anti-apartheid activists, and anti-racists, in other parts of the world noted that dealing with “race” relations as if they are about differences in cultures ignores the reality of racism. Looking at diversity of cultures does not address the power relations in society which privilege “white” people. It ignores the systemic ways in which racism is institutionalized and the ways in which a sense of “white” superiority and entitlement are constructed. Diversity of cultures which focus on differences in language, beliefs, customs, traditions, tastes, food, and dress do not address the racism of societies. Instead, diversity of cultures is an ideological mechanism which allows racism to continue (Bhaham et al., 1992; Gillborn, 1990). This was abundantly clear under Apartheid which constructed a racist society in the interests of “white” people could ideologically justify this by suggesting it was a way to recognize people’s cultural diversity. Diversity as a reference to cultures as a way to deal with “race” relations is a way to obscure the systemic and institutionalized construction of racism in people’s lives.

The “new” South Africa has on its coat of arms the phrase “United in our Diversity” (www.sahistory.org.za/chapter/coat-of-arms). It is important to elaborate on why this is the case. Although Apartheid perverted and subverted cultural diversity to suit its racist plans, it would be incorrect to suggest that there is no cultural diversity among people who have been defined racially. The “United in our Diversity” is meant to recognize that people have a right to their own cultures, languages, cultures, styles, dress, and food, etc. and these forms of diversity are a source of strength and resource for the country. It also recognizes that all people are entitled to such cultural rights on the basis of equality.

The problems with viewing diversity as a matter of “race” relations and consequently in terms of differences in cultures is that it tends to essentialize people in these terms. It suggests that people are only what cultures they belong to, and the experience of Apartheid shows that this can be done in a way that is profoundly unjust, unequal, and racist. It also shows that people are more than what they are defined as being racially or culturally. As Vertovec (2015) puts it:

Why is there so much interest in ‘diversity’ now? There are many reasons offered. Perhaps the foremost reason is the success, increasingly and commutively since the 1960s, of the key struggles of the identity-based public/political campaigns among women, African Americans, lesbian and gay, age- and disability-based movements.

Vertovec (2015: p. 1).

In this regard, people are more than just their “race” and culture. They are also classed, gendered, have different sexual orientations and with different abilities and of different ages. It is when people are viewed beyond “race” and cultural terms, that these other forms of identities that are also part of who people are become more visible and important.

Diversity and inclusivity

The “new” democratic South Africa sees itself as a non-racial, non-sexist, and democratic South Africa, and in its Constitution, it recognizes the following identities as part of the Bill of Rights.

The state may not unfairly discriminate directly or indirectly against anyone on one or more grounds, including race, gender, sex, pregnancy, marital status, ethnic or social origin, colour, sexual orientation, age, disability, religion, conscience, belief, culture, language, and birth

South African Constitution (1996).

This clause in the South African Constitution’s Bill of Rights recognizes more identities than the [Universal Declaration of Human Rights \(1948\)](#). What is significant about this is that it is another view of diversity that goes beyond simply recognizing people racially or culturally, while still recognizing these. “Race” and culture are a part of people’s identities, but people are also gendered, classed, speak different languages, have different abilities, religions, political and sexual orientations, age and in different regions, following on the provisions of the South African Bill of Rights. When diversity is viewed in these ways, diversity is more about inclusivity.

The idea of inclusivity has been prominent in the debates and developments of inclusive education. In inclusive education, the struggles of disabled people have been highlighted to indicate the need to mainstream provision for disabled learners in general schooling, and not to keep them segregated from so-called normal schools ([Slee, 2011](#)). Although debates continue between the special education and social models approaches to inclusive education ([Oliver and Barnes, 2010](#)), the Salamanca Statement of 1994 states:

every child has a fundamental right to education, and must be given the opportunity to achieve and maintain an acceptable level of learning, ● every child has unique characteristics, interests, abilities and learning needs, ● education systems should be designed and educational programmes implemented to take into account the wide diversity of these characteristics and needs, ● those with special educational needs must have access to regular schools which should accommodate them within a child-centred pedagogy capable of meeting these needs, ● regular schools with this inclusive orientation are the most effective means of combating discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all; moreover, they provide an effective education to the majority of children and improve the efficiency and ultimately the cost-effectiveness of the entire education system.

UNESCO (1994, Salamanca Statement, p. 1).

In this Statement, although its focus is on people with disabilities and special needs, it points importantly to diversity and “uniqueness” as part of people and the recognition thereof is necessary for “building an inclusive society”. This viewing of people’s diversity as necessary for inclusivity forces seeing diversity and inclusivity as more than referring to “race” and culture; and, in the case of inclusive education as more than just dis/ability.

Within inclusive education [Oliver and Barnes \(2010\)](#) note that the influence of critical theory, and its links with emphasis on transformation and emancipation were crucial in expanding the understanding of diversity and inclusivity. In inclusive education, the influence of critical theory has been significant in arguing for non-reductionist forms of analysis and to recognize the multiplicity of ways in which people are excluded, either directly or indirectly, from accessing social goods they are entitled to, in this case from education. Inclusive education, then began to develop a social model approach to inclusive education which included people with disabilities but also included people who are raced, gendered, classed, have different sexual orientations and ages, and so on. The social model of inclusive education argued that all people, no matter who they are, should have equitable and just access to education to allow for their own development. This is also captured in the South Africa White Paper 6 ([DBE, 2001](#)) which is the policy for inclusive education in South Africa. In White Paper 6, the notion of barriers to learning is put forward, and it indicates that it is not only people with disabilities who experience barriers to learning, but other people too. These other people include people who experience barriers to learning due to “race”, class, gender, region, age, etc., while still acknowledging the barriers to learning that people with disabilities experience.

There are two important issues that emerge from this. One is the recognition that people have multiple identities which need to be recognized and which are in fact contained in the Bill of Rights and/or the Universal Declaration of Human Rights. Secondly, it also indicates that for inclusive education, and thereby also inclusivity, to be realized it is crucial to recognize all such identities that constitute people. One cannot only recognize gender, for example, and not also look at class, region, language, religion, etc. This is important because the developments in inclusive education, and consequently the development of inclusivity, is that it is counterfactual to suggest that people are only one identity, and that they can be essentialized in terms of that identity. Instead, what this development points to is the importance of recognizing the multiplicity and complexity of people’s identities, and the crucial need

to move away from reductionist tendencies which promote the idea that people can be defined in terms of one single thing, and their senses of self and being can be reduced and fixed in these ways.

Diversity, then, as [Vertovec \(2015\)](#) indicated and quoted above, is now more than just a matter of “race” relations. It is about the multiplicity and multi-layered identities that people possess, and that this is crucial for a development of inclusivity, in inclusive education or for the development of an inclusive society and an inclusive world. The importance of this recognition is now also acknowledged by the Sustainable Development Goals—SDGs—of the United Nations ([UN, 2015](#)) and recently by the World Economic Forum—WEF—(2021), where it is noted that for a sustainable future a move toward a more inclusive world is critical, and this inclusive world should include all people in terms of the multiple and complex identities they possess individually and collectively.

Goal 8 of the SDGs states “Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all”. Goal 9 states “Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation”, and Goal 11 states, “Make cities and human settlements inclusive, safe, resilient and sustainable” ([UN, 2015](#)). The [WEF September \(2021\)](#) Briefing Paper uses the word “inclusive” in its vision of a future sustainable and equitable and better future for all.

Diversity then has morphed into inclusivity. However, the basic argument about inclusivity is the recognition of difference. This means that inclusivity acknowledges that people are different within themselves and in relation to others. It is important to elaborate on how difference is crucial for understanding inclusivity.

Difference

Difference is not a simple matter. It is not simply saying that all people are different. Difference is an ontological and epistemological claim, and it is important to understand difference in these terms ([Carrim, 2018](#)).

Difference ontologically

Ontologically, it is important to understand that difference is brought into play the instance when an identity emerges. A thing becomes what it is the instance it separates itself from all other things to become an individual entity. Nothing can be said to exist if it is not distinguishable from other things. If it is not distinguishable from other things, it becomes the same as other things since there is no distinction between it and other things. A tree is not a river, for example, because the tree separates itself from a river in order to be a tree. As another example, a child separates itself from its mother at birth to become an individual in its own right. The ontological basis for all things to come into existence as an entity of its own, thus, depends on separating itself from all other things, distinguishing itself from others and being different from all else.

Identities, then, come into existence because they differentiate themselves from all other things to become what it is. Identities, by definition are rooted ontologically in difference.

It is also important to note that once an identity comes into existence it does not remain the same. It changes, it grows, it develops. In this regard, identities are fluid and change all the time. Identities, then, come into being and they change once in existence. Being then is also simultaneously about becoming ([Nietzsche, 1968](#)).

The implications of recognizing difference ontologically are significant. Three such implications will be dealt with here: (1) social categories; (2) dominance of particular forms of identity; and (3) importance of multiplicity within and across people.

Recognizing difference within social categories is significant because it throws light on the unhelpful and counter productive categorization of people in terms of social categories. For example, to simply view women as women fails to consider that there are different ways in which women experience and define themselves as women. Women who are “black” do not experience being women as a “white” woman does, neither does a working-class woman experience being a woman as a middle-class woman does, nor does a rural women experience being a woman as an urban woman does. To homogenize and essentialize all women in terms of the social category of women is, thus, problematic if not short-sighted and counterfactual. Women may be distinguishable from men, for example, but it is erroneous to suggest that all women are, therefore, the same.

Critical theory and postmodernist forms of thinking have highlighted the importance of recognizing difference within social categories, and have, thus, argued for intersectionality ([Hall et al., 1992](#)). Intersectionality connects social categories to demonstrate the multiplicity and complexity of the forces that impact people. A woman, with intersectionality, is a woman who is also classed, raced, with different sexual and political orientations, lives in different regions and is of different ages and ability. The same holds true for other social categories too.

One could ask why, then, do particular forms of identity or social category become dominant. There are two ways in which one could respond to this. The first is to suggest that particular forms of identity or social category become dominant because in specific country contexts particular configurations of power emphasize one form of identity or social category over others because it is the basis for ideological construction and public imagination of those who are in political and economic power. In Laclau’s and Mouffe’s (see [Snir, 2017](#)) terms it is a dominant articulating principle. [Mamdani \(2018\)](#) indicates this as the specific configurations of power in particular country contexts.

The second response to why particular forms of identity or social category become dominant is because of ideological constructions which create imaginary senses of belonging uphold and perpetuate hegemonic orders. [Hall et al. \(1992\)](#) put it in the following way:

Within us are contradictory identities, pulling in different directions, so that our identifications are being continuously shifted about. If we feel we have a unified identity from birth to death, it is only because we construct a comforting story or 'narrative of the self' about ourselves. The fully unified, completed, secure and coherent identity is a fantasy.

[Hall et al. \(1992: 277\)](#).

Following on [Hall et al. \(1992\)](#), it can be said that dominant articulations of particular forms of identity or social categories are a result of political and ideological machinations which construct narratives that chimerically project some identities and social categories as more important than others and do so by silencing and invisibilizing other forms of identities and social categories which although deep in people's experiences get ignored. Such constructions are fantasies and construct comforting narratives of self (where self can be individual or collective).

Difference epistemologically

If one accepts that people are different and they are constantly in a state of becoming it follows that people will have different views and understanding of themselves, others and the world, and these views and understandings will change too. Knowledge/s that people have are, however, not what characterizes dominant world views. Instead, coming from colonized people and women and black people, as well as people from different religions and different forms of sexual orientations and dis/abilities are marginalized or ignored. Whose views dominate and how knowledge is viewed, then, become an important question.

Writing from the context of the United States of America [Apple \(2018\)](#) elaborated extensively on how dominant ideologies inform knowledge systems in the world, especially in school curricula. Apple noted that "white", European, and Western, heterosexist, male and Christian views dominate knowledge systems. The views of "the other" are either illegitimized, marginalized or simply ignored. This is also promoted widely by media, and do not only happen in schools. The impact of this is that it suggests that Western, European, "white", Christian, heterosexist and male views are correct and better than "other" worldviews. [Visvanathan \(1997\)](#), writing from an Indian perspective and on the basis of experiences of colonialism, describes this as "cognitive injustice", [Bourdieu \(1976\)](#) in looking at experiences of the working class in Algeria, described it as "symbolic violence" and [Spivak \(1988\)](#), also writing from an Indian perspective and from experiences of colonialism described as "epistemic violence". The point that is made by all of them is that when dominant knowledge systems legitimize themselves by illegitimizing "other" worldviews and knowledge systems it creates violence on the "other", and erroneously (perhaps even deliberately) projects its own knowledge system as "the truth".

Although there is no space in this chapter to engage with the many and complex arguments about truth claims which have preoccupied philosophers over the ages, it is important to clarify that an epistemology that recognizes difference is not necessarily equal to saying anything goes, and all views are valid. All views no matter what they are or where they come from need to make sense, that is they need to be logical and the argument they put forward needs to be coherent. The importance of articulating views in logical and coherent ways is argued by [Nietzsche \(1968\)](#) and by critical realists ([Sayer, 2000](#); [Potter and Lopez, 2001](#)). [Nietzsche \(1968\)](#) pointed to the importance of logic and argument, [Sayer \(2000\)](#) and [Potter and Lopez \(2001\)](#) argued that although epistemological relativism is important to consider there is still a way in which judgmental rationality can be exercised which can interrogate the basis on which claims are made, whether they make logical sense, are coherent in their argument, and if evidence is provided in sound ways to support claims made. In this light, then, it is important to realize that an epistemology of difference does not mean that any view, no matter how incoherent and illogical it may be, should be accepted.

[Spivak \(1988\)](#) points to another important dimension to releasing what she calls subjugated knowledges. Subjugated knowledges are those knowledge forms that have been illegitimized in the construction of dominant knowledge systems. Spivak in looking at the practice of "sati/satee" in India shows that when subjugated knowledges are released, they may not lead to more forms of inclusivity or justice. They may in fact reinforce dominant knowledge systems. Sati which is a ritual in India where a widowed woman will place herself on the pyre on which her recently deceased husband is being cremated, allows herself to burn herself alive because it is believed that she is of no worth if without a husband. Spivak shows that this ritual is deeply embedded in traditional, indigenous knowledge systems and argues that if such views are rearticulated as a way of resisting dominant knowledge systems, it will have drastic and damaging consequences for the struggles of women to emancipate themselves from gender and sexual oppression. The reason why this account of sati by Spivak is important is that it points crucially to the importance and need of interrogating all knowledge systems and to evaluate the extent to which they do or do not assist in recognizing difference among people, and how they assist in recognizing the equality of all human beings, rather than putting into place re/configurations of unequal and unjust systems of dominance and subordination.

Inclusivity, then, entails difference, and it is important to view difference in ontological and epistemological terms which considers the multiplicity and plurality of being human as well as the varied and varying ways in which human beings make sense of themselves, others, and the world. Such knowledge system upon which human beings understand themselves and their world may be relative, but they need to meet the conditions of logic, coherence of argument, and supporting claims made, for them to enter meaningfully into dialogs with other knowledge systems in order to arrive at a better understanding of phenomena.

Toward a more socially just world

This chapter has focused on the diversity and linked it to inclusivity and difference. Drawing on the South African experience of apartheid, it has shown that diversity could have profoundly negative consequences. What diversity is understood to mean, how it is framed, and the implications it has need to be understood and not be taken for granted.

It has been argued that diversity has historically been linked to questions of race relations. In so doing it has meant and been framed as referring to recognition of different “races” and ethnic or cultural groups. The apartheid experience in South Africa showed that this could be used in ways that segregate people in racist systems which is based on inequality. Although, the post-apartheid South Africa recognizes diversity in positive terms and in its Constitution as does the United Nations Universal Declaration of Human Rights, which recognize diversity on the basis of human rights and equality, the dominant view of diversity remains couched in racial, ethnic, and cultural terms.

It has also been argued the diversity has developed as linked to inclusivity. With the influences of critical theory, postmodernism and critical realism, the importance of moving away from reductionist forms of thinking that tend to homogenize and essentialize people’s senses of themselves and which lapse into mono-causal explanations. It has been pointed out that multiplicity, plurality, and difference are crucial to consider. In this regard, it has been demonstrated that viewing people as having multiple and multi-layered aspects to their own selves and in relation to others, and the ways in which different aspects of themselves intersect with each other is crucial. Inclusivity was shown to recognize that diversity is not only about race, ethnicity, or culture, but also about gender, class, region, sexual and political orientation, etc. Such complexity and multiplicity of human beings and their contexts broaden the meaning and frame of diversity.

Inclusivity was also shown to imply difference. In this regard, it has been argued that difference, if viewed ontologically and epistemologically, provide a useful way to take forward diversity and inclusivity so that they can speak to the complexities and multiplicities that have come to the fore in postmodernist conditions and in the 21st century.

The suggestion made in this chapter is that difference, more than diversity, may provide a more useful way in which to recognize the plurality of dimensions that characterizes human beings’ own senses and experience of themselves, in their relations with others and the world. For a more sustainable future that is based on inclusivity, it has been suggested that working with and through difference may be more useful than holding onto diversity, which due to its historical connotations and meaning, have tended to function ideologically to not bring into focus the systemic, institutional, and systematic ways in which systems of domination and subordination are reproduced. Moving toward more socially just and sustainable futures require justice for all on the basis of equality, not just in terms of race, ethnicity, and culture, as diversity implies. Justice for all in a socially just order of the future will need to engage with the complexities, multiplicities and difference that exist within people, between people and the world. Recognizing and working with difference in constructive and productive ways, it has been argued, is needed to move away from reducing and fixing people into one or another social category or narrative which strip people of the complex beings that they are and the equally complex world that they inhabit.

References

- Apple, M., 2018. *Ideology and the Curriculum*, fourth ed. Taylor and Francis, New York.
- Braham, P., Rattansi, A., Skellington, R., 1992. *Racism and Antiracism: Inequalities, Opportunities and Policies*. Sage, London.
- Bourdieu, P., 1976. The school as a conservative force. In: Dale, R., Esland, G., Macdonald, M. (Eds.), *Capitalism and Schooling*. Open University with Routledge and Kegan Paul, London.
- Carrim, N., 1995. From “race” to ethnicity: shifts in the educational discourses in South Africa and Britain in the 1990s. *Compare* 25 (1), 17–33.
- Carrim, N., 2018. Difference in current post-apartheid education. In: Osman, R., Walton, E. (Eds.), *Teacher Education for Diversity: Conversations From the Global South*. Routledge, Oxon, pp. 153–167.
- Cross, M., 2004. Institutionalising campus diversity in South African Higher Education: a Review of diversity scholarship and diversity education. *High Educ.* 47, 387–410.
- Department of Basic Education, 2001. *Education White Paper Six: Special Needs Education: Building an Inclusive Education and Training System*. Government Press, Pretoria.
- Gillborn, D., 1990. *Race, Ethnicity and Education: Teaching and Learning in Multi-ethnic Schools*. Routledge, London.
- Hall, S., Held, D., McGrew, T., 1992. *Modernity and Its Futures*. Polity Press, Cambridge.
- Mamdani, M., 2018. *Citizen and Subject*. Princeton University Press, Maryland.
- Nietzsche, F., 1968. *Will to Power*. Translated by Walter Kaufmann. Vintage, New York.
- Oliver, M., Barnes, C., 2010. Disability studies, disabled people, and the struggle for inclusion. *Br. J. Sociol. Educ.* 31 (5), 547–560.
- Potter, G., Lopez, J., 2001. *After Postmodernism: The New Millennium*. Athlone, London.
- Rex, J., Mason, D. (Eds.), 1986. *Theories of Race and Ethnic Relations*. Cambridge University Press, Cambridge.
- Rose, B., Tunmer, R., 1975. *Documents in South African Education*. Donker, Johannesburg.
- Sayer, A., 2000. *Realism and Social Science*. Sage, London.
- Slee, R., 2011. *The Irregular School*. Routledge, London.

- Snir, I., 2017. Education and articulation: Laclau's and Mouffe's radical democracy in schools. *Ethics Educ.* 12 (3), 351–363.
- Spivak, G.C., 1988. Can the sub-altern speak? In: Nelson, C., Grossberg, L. (Eds.), *Marxism and the Interpretation of Culture*. Macmillan, London, pp. 24–68.
- South African Constitution, 1996. Bill of Rights. Government Press, Pretoria (Chapter 2) Clause 9, Section 3.
- UNESCO, 1994. Salamanca Statement and Framework for Action on Special Needs Education. UNESCO, Spain.
- United Nations Universal Declaration of Human Rights, 1948.
- United Nations, 2015. Sustainable Development Goals.
- Vertovec, S., 2015. Introduction to diversity studies. In: Vertovec, S. (Ed.), *Routledge International Handbook of Diversity Studies*. Routledge, London, pp. 1–20.
- Verwoerd, H.F., 1953. Senate Speech. www.sahistory.org.za/archive.
- Visvanathan, S., 1997. *A Carnival for Science: Essays on Science, Technology, and Development*. Oxford University Press, London.
- World Economic Forum, 2021. Briefing Paper: Shaping an Equitable, Inclusive, and Sustainable Recovery: Acting Now for a Better Future. WEF, Davos. September.

Relevant websites

- www.sahistory.org.za/chapter/coat-of-arms (Accessed September 2016).
- <https://www.sahistory.org.za/archive/hendrik-verwoerd-10-quotes-hendrik-verwoerd-politics-web-20-september-2016> (Accessed September 2021).
- <http://www.waynevisser.com/report/sdgs-finalised-text> (Accessed October 2021).

Democracy and education

Stewart Riddle, School of Education, University of Southern Queensland, Springfield Central, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	8
Contemporary challenges for democracy and education	9
Democracy and education as neoliberal commodities	10
Toward education for democracy	11
References	12

Introduction

Democracy is simultaneously an ancient idea and an unfinished experiment in social and political governance. Although the Athenians are generally credited with conceiving of the notion of δημοκρατία (democracy)—by which the people (δῆμος; demos) ruled (κρατία; cracy) the polis—there have been different forms of democracy that have arisen ever since people first formed together in communities. Despite its ubiquity as a political form and an ideal of collective commitment to civic engagement, democracy has long been criticized for its shortcomings. For example, Plato (2007) famously railed against the failings of Athenian democracy in *The Republic*, claiming that it enabled tyranny by the mob through which base impulses drove decisions, rather than a commitment to the common good of the polis. However, Churchill (2008) countered that “no one pretends that democracy is perfect or all-wise. Indeed, it has been said that democracy is the worst form of government except for all those other forms that have been tried from time to time” (p. 574).

Despite its long history as a mode of social and political organization, democracy remains a contested concept, which “is at the center of struggles over what the goals of education should be” (Apple, 2011, p. 23). Since American pragmatist philosopher, John Dewey wrote about democracy and education more than a century ago, the two have been intimately entwined in debates about the purpose of schooling and other sites of education in preparing young people for the world. Dewey (1916) argued that “democracy is more than a form of government; it is primarily a mode of associated living, of conjoint communicated experience” (p. 101). Through democracy, people are able to realize the full potential of their collective interest and break down social barriers such as class, race, and gender. However, “democracy is not something out there waiting to be reached. Rather it is in the work itself as we create ways to promote human dignity, equity, justice, and critical action” (Beane and Apple, 2007, p. 25).

In his writings on education and democracy, Dewey (1899) argued that schools should provide the opportunity for all young people to be educated into the democratic practices and social belonging of the community. He claimed that “what the best and wisest parent wants for his own child, that must the community want for all of its children” (Dewey, 1899, p. 19), through a balance between individualism and socialism. In the contemporary policy landscape of neoliberalism, in which the individual reigns supreme over any concern for the collective good, such a sentiment stands in stark contrast to the educational realities of many young people who are disenfranchised and marginalized due to their skin color, sexual orientation, religious or spiritual beliefs, cultural histories or socioeconomic status.

Mouffe (1993) argued that “far from being the necessary result of a moral evolution of humankind, democracy is something uncertain and improbable and must never be taken for granted. It is an always fragile conquest that needs to be defended as well as deepened” (p. 6). This chapter seeks to provide an understanding of the relationship between education and democracy, given the complex set of contemporary challenges that face young learners in the 21st century. In doing so, this chapter argues that we—educators, researchers, policymakers, parents, learners, and community members—should work toward an education that is for democracy.

Democratic education is not simply an educational ideal, it is also a deeply political commitment to a collective ethics of responsibility to each other (Gutmann, 1999). In the social order, there is little more important than the nurture and education of the young. As Dewey (1916) claimed, “each generation is inclined to educate its young so as to get along in the present world instead of with a view to the proper end of education: the promotion of the best possible realization of humanity” (p. 111). Democratic education is not interested in the simple maintenance of the status quo, but rather in the active reconfiguration of society to be more inclusive and embracing of all, especially those who are marginalized and disenfranchised.

However, as Carr (2011) maintained, “the problem is not that democracy is not a worthy concept but, rather, that the type of democracy that has received normative, relativistic salience is, in many regards, anti-democratic” (p. 4). Contemporary liberal democracies act as little more than fronts for corporate oligarchies and massified individualistic consumerism (Rancière, 2006), in which the *thick* version of democracy gives way to a commodified *thin* version of democracy writ small as *choice* (Apple, 2018), whether via an electoral experience that provides little, if any, differentiation between political candidates, and market-driven product consumption and individual competition, which leaves little time for the consideration of public life and its concerns for the welfare of others.

There is an argument that neoliberalism has fostered a thin version of minimalist democracy, in which exclusion is the prime function of contemporary social, economic, and political discourses, which “depends not only on the predominance of the economic in the political arena but also on an elitist conception of citizen participation (electoral, sporadic and easily manipulated by spectacle and demagoguery)” (Szolucha, 2016, p. 20). However, it does not need to be so. Indeed, as Biesta (2014) claimed, democratic possibility enables the rethinking of democratic ideas, subjectivities, and configurations in different ways, and that the treatment of democracy “as an ongoing and never-ending experiment is a way to capture this idea” (p. 6).

Contemporary challenges for democracy and education

The first two decades of the 21st century have ushered in an era of intense crisis for democracy and education. There are complex and interconnected challenges that have emerged, which threaten the foundations of liberal democratic societies. Despite the argument in the late-twentieth century that globalized liberal capitalist democracy had triumphed over all other political and economic forms, culminating in the “end of history” (Fukuyama, 1989), there has since been a rapid fraying of democracy around the world. This can be observed in the rise of neofascist populism in many countries, including the US, Brazil, and Turkey. As Moffitt (2016) contended, “populism is back—and it is back with a vengeance. What was once seen as a fringe phenomenon relegated to another era or only certain parts of the world is now a mainstay of contemporary politics across the globe” (p. 2).

The rise of alternative facts, fake news and rightwing extremism has posed a significant challenge to the democratic practices and public institutions of many countries. This can perhaps best be observed in the phenomenon of Trumpism, a four-year experiment in demagoguery, which culminated in claims fueled on social media and by Trump himself that the 2020 US Presidential election was stolen by Biden and the Democrats. Gardels and Berggruen (2019) warned of the increasing deleterious effects on civic discourse from growing spread of misinformation, arguing that “despite expectations that the Internet Age would create an informed public more capable of self-government than ever before in history, fake news, hate speech, and ‘alternative facts’ have seriously degraded the civic discourse” (pp. 2–3).

At the same time that democracy as a form of political order is rapidly unraveling in liberal capitalist democratic societies, growing social and economic inequality pose additional risks to social cohesion. There are shocking wealth disparities between the elite and the general population in countries that for much of the twentieth century prided themselves on egalitarianism and equal opportunity. The myth of meritocracy has been laid bare through the rapid stripping out of the middle-class and the accelerated growth of the underemployed and unemployed as a disposable surplus population (Srnicek and Williams, 2016). Billionaires fire themselves into space, while millions live with the daily threat of malnutrition and starvation. Dorling (2019) argued that the top 1% “contribute to rising inequality, not just by taking more and more, but by suggesting that such greed is justifiable and using their enormous wealth to promote that concept” (p. 22).

Runaway climate change continues apace, while governments and other political, social and economic institutions seem unable or unwilling to address the urgency of the need to rapidly decarbonize and decouple from energy reliance on fossil fuels. The powerful oil lobby, buoyed by a dependent globalized economy and compliant governments, continues to wreak havoc on the natural systems of this finite planet. It is a poorly kept secret that powerful corporate and political “elites see equality as a danger” (Schostak, 2019, p. 1103), and as such, consider democracy itself to be a threat to their continued privilege and power.

The first couple of years of this decade have seen political, social and economic systems entirely overwhelmed by the challenges posed by the COVID-19 global pandemic. Entire economies ground to a halt, widespread restrictions on freedom of movement were imposed, relaxed, and imposed again. Those who had the material and cultural resources to do so, bunkered down safe in the knowledge of having access to supplies and high-speed internet connections to ensure a smooth transition to home-based working and learning. At the same time, those already in precarious situations found that the ground had given way, without adequate access to shelter, food and healthcare. Even a working internet connection became a necessary condition of daily engagement with the worlds of work and education. For many millions, such luxuries were simply not available to them. It is no coincidence that those who were hardest hit by the pandemic were the working poor, those living with disabilities, and other vulnerable and marginalized groups in society. The pandemic has increased the pace of wealth distribution toward the few wealthy elites, while many continue to suffer the consequences of predatory capitalism, globalization and increased precarity, to say nothing of the enormous toll of the global COVID-19 public health crisis.

At every stage, with every challenge, democratically elected governments have failed to adequately respond in meaningful ways. Trust in governments and public institutions continues to wane, as people live with the daily realities of rising inequality, social instability and political corruption (Dorling, 2019). The rise of populism is one effect of a situation that seems inescapable; people cry out for the strongman to do something, anything, to fix the problem. However, history has shown multiple times over that populism, particularly the militarized, rightwing version that appears to be taking hold in many supposedly liberal democracies, never ends well for the people who support its rise.

Rapid changes in technology also pose a significant challenge to democracy in the 21st century. Bloom and Sancino (2019) argued that smart technologies (e.g., robotics, artificial intelligence, social media, automation, and the internet of things) are disrupting politics and economics around the world, with the effects on democracy and education yet to be realized. They warn of the potential for technological advances to intensify inequality and the concentrated power and wealth of a globalized corporate elite, while also suggesting that there is potential for the radical transformation of “democracies and civic societies, creating ones that are more responsive, egalitarian and accountable” (Bloom and Sancino, 2019, p. 5).

Democracy and education as neoliberal commodities

Much of the contemporary policy landscape is underscored by the mantra of neoliberalism, which turns everything—including education and democracy—into markets, then seeks to make those markets wider (Connell, 2013). The defining characteristics of neoliberalism are competition, individual choice, and free markets. Citizens are recast as consumers, who make choices within markets whether that be for education, work or purchasing goods and services. The rapid globalization of capital—human and economic—since the 1980s has driven neoliberal policy thinking into most corners of the world. Some of the largest supranational institutions have their main goal to support and perpetuate neoliberal ideals, including the World Trade Organization, the World Bank, and the International Monetary Fund.

Rancière (2006) argued that in contemporary capitalist societies that purport to liberal democratic modes of governance, democracy has been co-opted within the formation of modern society as the “apolitical life of the indifferent consumer of commodities” (p. 29). Further, Rancière (2006) claimed that mass individualism and oligarchic interests drive the forms of subjectivation within contemporary democratic societies, in which democratic passions are redirected toward private pleasure and people are encouraged to disengage from public life. Democracy is reduced to the possessive individualism of the citizen as consumer, which has implications for de-racing, de-classing and de-gendering issues of injustice and inequality, which are reframed as individual problems rather than social justice issues (Apple, 2006).

McLaren and Farahmandpur (2005) argued that neoliberal reforms over the past several decades have weakened social support structures such as public health and education, undermining social justice and the collective and individual rights of people to engage in meaningful work and education. Instead, the power of corporate oligarchs has eclipsed the democratic sphere, where international corporations have more economic and political power than democratically elected governments. Perhaps the most perverse outcome is the corruption of democratic processes to the market, in which the election cycle becomes one more carefully controlled public relations exercise by corporate powers and the representatives of their commercial interests.

Giroux (2016) argued that “under the auspices of a diverse group of extremists, including political, religious, and market fundamentalists, political and educational culture is being transformed into the discourse of privatization, consumerism, market-based choice, the spectacle of celebrity, and the revived ethics of social Darwinism” (p. xxvii). Neoliberalism recasts education as “human capital formation”, through which students learn the skills and attitudes to be productive members of a globalized workforce, which produces increasing profits for the market economy (Connell, 2013). Democracy and education is reduced to the simple notion of choice in the education marketplace, in which parents and their children make decisions about where they will get the best return on their investment. It is particularly visible in the influence of corporate interests, such as those which “provide for-profit education in the form of ventures such as online charter schools, marketed to parents as a more democratic, personalized and accessible educational opportunity for their children” (Riddle and Heffernan, 2018, p. 321).

Further, neoliberalism works to ensure a thin, marketized and commodified version of democracy, which promotes possessive individualism and competition, rather than a fuller democracy, which enables collective participation and critical engagement with society’s institutions and public decision-making apparatuses (Apple, 2013, 2018). While liberty and equality have always been understood as necessary elements of democracy, contemporary neoliberal formations have skewed democracy “toward a particular understanding of individual liberty at the expense of social equality” (Szolucha, 2016, p. 23).

Even liberty and individual freedom become reduced to a superficial form of choice in a market in neoliberalism, stripped of its emancipatory politics and ethics. Monbiot (2016) claimed that “freedom of the kind championed by neoliberals means freedom from competing interests. It means freedom from the demands of social justice, from environmental constraints, from collective bargaining and from the taxation that funds public services. It means, in sum, freedom from democracy” (p. 12).

Biesta (2014) warned that “by replacing democracy with choice, by letting the logic of the market into the public domain, and by giving up on the idea that democracy is ultimately about transformation, the possibilities for the enactment of democratic citizenship begin to disappear” (p. 9). The weakening of democratic practices and opportunities for people to meaningfully engage with each other in public spaces and working toward the common good enables what Schostak and Goodson (2020) described as the *propaganda society*, which relies on a compliant and acquiescent public that is educated according to the desired outcomes of the elites, rather than a democratic mode of education that produces critical and creative citizens who can work together in multiple ways to ensure a rich, sustainable society.

Apple (2006) argued that public institutions should sit at the heart of caring, inclusive democratic societies. As such, capitalist market relations should support and extend democratic institutions to ensure “a fuller and thicker participatory democratic polity and daily life” (Apple, 2006, p. 103). However, a functioning democracy requires an educated citizenry (Carr, 2020), whose capacities for critical engagement and participation are supported by their education, which enables the sharing of diverse perspectives within an inclusive society.

The question is, “can normative, hegemonic democracy effectively grapple with social inequalities? Neoliberalism, tightly controlled elections offering few options, and an educational system that is structured more for the job market than social justice appear to be obstacles” (Carr and Thésée, 2019, p. 245). What are viable alternatives to normative, hegemonic democracy as it is currently shaped by neoliberal policies within globalized capitalism? Is it possible to ensure a more transformative education that supports meaningful democracy, especially for peoples who are presently marginalized and excluded from substantial engagement and participation in society? One approach to conceptualize potential solutions to the challenges facing education and democracy in the 21st century is to consider the multiple ways in which education can be *for* democracy.

Toward education *for* democracy

There is little doubt that education is dangerous (Connell, 2012), not only to totalitarian and authoritarian regimes, but even to the powerful elites within supposedly democratic states in the contemporary globalized and interconnected world. Similarly, democracy is viewed not only as a dangerous concept, but even more so, a dangerous practice (Riddle and Apple, 2019). As Rancière (2006) argued, those in power have a vested interest to hold democracy at bay, to restrain the impulse of the masses to collectivize and work toward a common good. How quickly it would become evident that capitalism's production of a minority wealthy elite and a majority of people living at or just above subsistence levels is an unsustainable proposition for a healthy, functioning society. While billionaires jaunt into sub-orbital space, millions of people live in conditions unworthy of the term "living". Corporate greed and a compliant political class ensure that the destruction of natural ecosystems continues unabated, despite the clear and present threat of climate change. Unsurprisingly, the working poor, unemployed, disadvantaged, marginalized, and those cast aside by society are the ones most at risk of displacement and dispossession in the coming years as rising sea levels, increased extreme weather events and reduced arable land and drinkable water begin to take their toll.

Given the challenges facing societies in the 21st century, what could and should be the role of education in preparing young people for a changing world with complex problems that require new ways of thinking, working, learning, and living together? An education that is for democracy can provide young people with the critical and creative capacities to engage in collaborative and collective responses to these complex challenges. Education must sit at the center "of our efforts to reconfigure the social apparatus in more equitable and socially just ways, for both the sustainability of our communities and the planet" (Riddle, 2019, p. 3).

Giroux (2010) argued that education is never neutral because it attempts to teach young people to "inhabit a particular mode of agency, enable them to understand the larger world and one's role in it in a specific way, define their relationship, if not responsibility, to diverse others, and experience in the classroom some sort of understanding of a more just, imaginative, and democratic life" (p. 718). An education that is for democracy recognizes that young people's agency in the world depends on their interrelations within societies, but also importantly within ecosystems and the finite planet upon which we all rely for our continued existence.

The search for a more democratic society through education "is tantamount to seeking the truth. Never comfortable, nor easily achieved, such a proposition requires a multitude of measures as well as the belief that people can, ultimately, function together without self-destruction" (Carr, 2011, p. 262). Additionally, radically inclusive forms of education, which bring people together in communities of learning, enable a rich diversity of perspectives, voices, and understandings to be included in the formation of multiple democratic spheres of being and belonging (Schostak, 2019).

Neither democracy nor education are ideal in their conception and practical realization, but that is no reason to decide that they cannot work more fully toward the realization of an inclusive and caring society (Gutmann, 1999). Apple (2011) claimed that democratic education is not simply about education, but at its core regards the kind of society we want and the politics that might help us achieve those ends. Similarly, Carr (2020) argued that our present context of multiple political, economic, and public health crises requires us "to re-imagine a more meaningful, critically engaged democracy" (p. 4). Education provides an important site of preparation for engagement in such meaningful democratic practice.

While schools and other sites of education can be co-opted into the existing power relations of contemporary neoliberalism—with public schools producing compliant workers for a globalized marketplace and elite private schools producing future political and corporate leaders—clearly, it does not have to be this way. There are many examples of educators, students, their families, and communities engaging in collective acts of resistance and reimagining education in local and global contexts (e.g., Apple, 2018; Riddle and Apple, 2019; Riddle et al., 2021). These multiple and varied enactments of democratic modes of living together offer all of us hope for a future that is more sustainable, inclusive and caring, especially for those who are least advantaged by the present system.

There are multiple examples of youth-led movements through which young people have agentic expression through political resistance, civil disobedience and the reclaiming of schools and other sites as democratic spaces as part of a public pedagogy (Brady, 2006; Giroux, 2003, 2004). The aim of such grassroots democratic mobilizations is not to cause anarchy or destabilization, but to reimagine public institutions, power and politics in more democratic ways (Mouffe, 2019). The Occupy movement provides a key example of radical democratization through acts of sustained public resistance to neoliberalism and corporate control of public institutions (Amsler, 2015; Szolucha, 2016). Similarly, the School Strike for Climate movement demonstrates the potential power of democracy by young people, who are usually excluded from the spheres of political and economic influence (Mattheis, 2020). There are many other examples of both localized and globally active youth-led movements, which have galvanized around issues of gender and sexuality, diversity and inclusion, racism, gun control and climate activism. Each democratic movement and act of public pedagogy represents an important commitment to reimagining democracy in ways that are more inclusive, sustainable and representative of diverse communities.

However, these acts of engaged reimagining of education and democracy need to be underwritten by much wider societal and institutional reform at the systemic level. Otherwise, the complex array of challenges posed by rising social and economic inequality, increasing social unrest and destabilization, political extremism and the influence of misinformation, fear, distrust, and hate, alongside the calamitous effects of catastrophic climate change, corporate environmental vandalism, and ecosystem collapse, will make such efforts pale against the enormity of the task. An education that is truly for democracy in the 21st century requires a radical reimagining of the processes, power relations and institutional structures, right down to the foundations of liberalism, capitalism, and democracy itself. The stakes are simply too high to do otherwise.

References

- Amsler, S.S., 2015. *The Education of Radical Democracy*. Routledge.
- Apple, M.W., 2006. Educating the "Right" Way: Markets, Standards, God and Inequality, second ed. Routledge.
- Apple, M.W., 2011. Democratic education in neoliberal and neoconservative times. *Int. Stud. Sociol. Educ.* 21 (1), 21–31. <https://doi.org/10.1080/09620214.2011.543850>.
- Apple, M.W., 2013. *Knowledge, Power and Education*. Routledge.
- Apple, M.W., 2018. *The Struggle for Democracy in Education: Lessons From Social Realities*. Routledge.
- Beane, J.A., Apple, M.W., 2007. The case for democratic schools. In: Apple, M.W., Beane, J.A. (Eds.), *Democratic Schools: Lessons in Powerful Education*, second ed. Heinemann, pp. 1–29.
- Biesta, G., 2014. Learning in public places: civic learning for the twenty-first century. In: Biesta, G., De Bie, M., Wildemeersch, D. (Eds.), *Civic Learning, Democratic Citizenship and the Public Sphere*. Springer, pp. 1–11.
- Bloom, P., Sancino, A., 2019. Disruptive Democracy: The Clash Between Techno-populism and Techo-democracy. Sage.
- Brady, J.F., 2006. Public pedagogy and educational leadership: politically engaged scholarly communities and possibilities for critical engagement. *J. Curric. Pedagogy* 3 (1), 57–60. <https://doi.org/10.1080/15505170.2006.10411575>.
- Carr, P.R., Thésée, G., 2019. It's Not Education that Scares Me, It's the Educators: Is There Still Hope for Democracy in Education, and Education for Democracy? Myers Education Press.
- Carr, P.R., 2011. *Does Your Vote Count? Critical Pedagogy and Democracy*. Peter Lang.
- Carr, P.R., 2020. Shooting yourself first in the foot, then in the head: normative democracy is suffocating, and then the coronavirus came to light. *Postdigit. Sci. Educ.* <https://doi.org/10.1007/s42438-020-00142-3>.
- Churchill, W., 2008. *Churchill by Himself: The Definitive Collection of Quotations*. Public Affairs Books.
- Connell, R.W., 2012. Just education. *J. Educ. Pol.* 27 (5), 681–683. <https://doi.org/10.1080/02680939.2012.710022>.
- Connell, R.W., 2013. The neoliberal cascade and education: an essay on the market agenda and its consequences. *Crit. Stud. Educ.* 54 (2), 99–112. <https://doi.org/10.1080/17508487.2013.776990>.
- Dewey, J., 1899. *The School and Society: Being Three Lectures*. University of Chicago Press.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. The Macmillan Company.
- Dorling, D., 2019. *Inequality and the 1%*. Verso.
- Fukuyama, F., 1989. The end of history? *Natl. Interest* 16, 3–18.
- Gardels, N., Berggruen, N., 2019. *Renovating Democracy: Governing in the Age of Globalization and Digital Capitalism*. University of California Press.
- Giroux, H.A., 2003. Public pedagogy and the politics of resistance: notes on a critical theory of educational struggle. *Educ. Philos. Theor.* 35 (1), 5–16. <https://doi.org/10.1111/1469-5812.00002>.
- Giroux, H.A., 2004. Public pedagogy and the politics of neoliberalism: making the political more pedagogical. *Pol. Futures Educ. Internet* 2 (3–4), 494–503. <https://doi.org/10.2304/pfie.2004.2.3.5>.
- Giroux, H.A., 2010. Rethinking education as the practice of freedom: Paulo Freire and the promise of critical pedagogy. *Pol. Futures Educ. Internet* 8 (6), 715–721.
- Giroux, H.A., 2016. *Schooling and the Struggle for Public Life: Democracy's Promise and Education's Challenge*, second ed. Routledge.
- Gutmann, A., 1999. *Democratic Education*. Princeton University Press.
- Mattheis, N., 2020. Unruly kids? Conceptualizing and defending youth disobedience. *Eur. J. Polit. Theor.* <https://doi.org/10.1177/1474885120918371>.
- McLaren, P., Farahmandpur, R., 2005. *Teaching Against Global Capitalism and the New Imperialism: A Critical Pedagogy*. Rowman & Littlefield Publishers.
- Moffitt, B., 2016. *The Global Rise of Populism: Performance, Political Style and Representation*. Stanford University Press.
- Monbiot, G., 2016. *How Did We Get Into This Mess? Politics, Equality, Nature*. Verso.
- Mouffe, C., 1993. *The Return of the Political*. Verso.
- Mouffe, C., 2019. *For a Left Populism*. Verso.
- Plato, 2007. *The Republic*. Penguin Classics.
- Rancière, J., 2006. *Hatred of Democracy*. Verso.
- Riddle, S., Apple, M.W. (Eds.), 2019. *Re-imagining Education for Democracy*. Routledge.
- Riddle, S., Heffernan, A., 2018. Education and democracy for complex contemporary childhoods. *Glob. Stud. Childhood* 8 (4), 319–324. <https://doi.org/10.1177/2043610618817370>.
- Riddle, S., Heffernan, A., Bright, D. (Eds.), 2021. *New Perspectives on Education for Democracy: Creative Responses to Local and Global Challenges*. Routledge.
- Riddle, S., 2019. Democracy and education in local–global contexts. *Int. Educ. J.* 18 (1), 1–6.
- Schostak, J., Goodson, I., 2020. *Democracy, Education and Research: The Struggle for Public Life*. Routledge.
- Schostak, J., 2019. "Towards a society of equals": democracy, education, cooperation and the practice of radical inclusion. *Int. J. Incl. Educ.* 23 (11), 1103–1115. <https://doi.org/10.1080/13603116.2019.1629161>.
- Smicek, N., Williams, A., 2016. *Inventing the Future: Postcapitalism and a World Without Work*. Verso.
- Szolucha, A., 2016. *Real Democracy in the Occupy Movement: No Stable Ground*. Routledge.

Social justice and education: addressing inequities in schooling

Elaine Coburn^a and Carl E. James^b, ^a International Studies, Centre for Feminist Research, York University, Toronto, ON, Canada; and ^b York University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Theorizing social justice and education from above	13
Theorizing socially just education from below	15
Critiques of social justice and education	16
Socially just education in our times	17
References	18

Contemporary debates pertaining to social justice and education remain a major concern among educators advocating for equitable education. We maintain that the moral and political urgency of arguments related to social justice and education arises, ironically, in the absence of social justice in and through formal education systems. Scholars and policy makers, educators and community activists are motivated to theorize about social justice in order to understand and then challenge injustices. Debates about social justice and education are thus rooted in a normative desire to challenge unjust relationships (or inequities) while creating right relationships in and through formal education systems.

Unlike legal justice focused on the individual rights-bearing persons and associated juridical remedies, social justice is concerned with morally and politically desirable social relationships writ large. The satisfaction of individual rights to education through the juridical system may thus be part of but is not commensurate with social justice in education, which depends on just relationships more than the realization of individual rights. Socially just education thus counters unintentional and structural reproduction of inequities (Young, 1988: 272–265). Yet theories of social justice in education are sufficiently diverse and contested that even this capacious conceptualization of social justice may be challenged, notably by those who anchor their understandings of a socially just education system in the ethical stance “that individuals are the ‘primary objects of moral concern’” (Brighouse and Swift quoted in Walker and Unterhalter, 2007: 2). A central aim of this contribution is therefore to explore debates about what social justice is and how these inform varied understandings of socially just education systems.

To those ends, we first conceptualize social justice theorizing from above and consider how three influential Western philosophical arguments about social justice have informed understandings of what socially just education does (or should do). Next, we turn to intellectual traditions from below that emphasize the necessity of theorizing historically specific injustices in education, especially class inequities, racisms and colonialism, rooted in insights gained from the lived experiences of the most marginalized and racialized individuals. Then, we briefly consider some arguments against socially just education from the left and from the right, before concluding with a brief survey of contemporary concerns about social justice and education in our times.

Theorizing social justice and education from above

Twenty years ago, Gewirtz (1998) expressed concern that social justice is under-conceptualized in education policy discussions, even among those who emphasize their commitments to social justice in education. Although there continue to be expressions of concern that social justice is a “*feel good*” concept, where the affective appeal disguises a lack of theoretical precision (Brennan and Naidoo, 2008: 287, italics in original), today, Gewirtz’ observation is much less true. Instead, as Wilson-Strydom (2015: 144) argues, there is recognition that “it is necessary to interrogate the notion of social justice, what it means, and why it is helpful in understanding inequalities in a[n] ... education context.”

Thus, theorizing the relationship between social justice and education (or in education) is often explicitly situated within broader Western philosophical debates about justice and what social justice entails and demands. In particular, theories of justice respectively elaborated by Rawls (1972) concerning questions of distribution, by Young (1990) and Fraser (1995), concerning questions of difference, and by Sen (1993) and Nussbaum (2011), concerning capabilities, have been influential in theorizing social justice in education (for instance, Bell, 2016; Eisenberg, 2006; Nussbaum, 2006; Polat, 2011; Vincent, 2003; Walker and Unterhalter, 2007; Wilson-Strydom, 2015; Zajda et al., 2006).

Contemporary Western philosophical theories of social justice have traditionally been concerned with questions of the fair distribution of goods, where liberal philosopher Rawls’ has offered the most influential account of distributive justice (1972).¹ Following Rawls’ emphasis on “conditions of fair equality of opportunity” (Rawls cited in Taylor, 2004: 334) and granting that education is itself a good, the question for educators concerned with social justice as distribution is then “who gets its benefits?”

¹Explorations less frequently turn to classical as well as contemporary philosophers. Zajda et al. (2006: 9) briefly invoke Plato, Aristotle and Aquinas, for instance, in thinking through diverse, Western philosophical conceptualizations of justice and their implications for socially just education.

(Connell, 1993: 12) or alternatively stated, who has the right to education? (Connell, 1992: 133–5; Gewirtz, 1998: 470–471). Benefits are understood as immediate (since education is a good in itself) and as leading to other benefits (like better-paying jobs for the educationally credentialed). Consequently, studies concerned with social justice as a question of distribution often focus on the availability of education, as well as who has access to an education and educational certification and who doesn't (Oduaran, 2006). Ultimately, the broad or universal availability of education and access to education are taken to be a measure of a just society, or, as one United Nations' publication emphasizes, "the availability of educational opportunities and ensuring the ability to take advantage of them is a fundamental component of any equitable community" (Matsuura, 2006: v).

The distributive justice approach to education informs normative arguments in favor of universal access to education, as well as empirical studies assessing access to education, like graduation rates at high school or university by gender, race and/or class. Nationally-focused studies of distributive justice in education, for instance, assess Black students' "access and ... fair chance of success" in higher education in post-apartheid South Africa (Sehoole and Adeyemo, 2016), inquire into access to post-secondary education in the province of Ontario, Canada by historically underrepresented Black, Latino and South Asian youth (James and Parekh, 2021), and consider the availability of early childhood education to disadvantaged (low-income) families in Chile (Barco and Carrasco, 2020). A characteristic question that scholars of education concerned with distributive justice seek to answer is thus: "Who goes to university [or elementary or secondary school], and do all have an equal chance to attend?" (Weingarten, 2021: 30–43). Often, there is a particular concern for the most disadvantaged, drawing on John Rawls' argument that measures of distributive justice should support the "least favored" (Connell, 1992: 139–40; Giroux, 2003: 10; Taylor, 2004: 334).

Other scholars concerned with social justice in education challenge an exclusive focus on the distributive dimension of education, understood as a good. In so doing, they follow theorizations of justice as demanding recognition, put forward by philosophers Iris Marion Young (1990) and Fraser (1995), among others (Huttunen and Heikkinen, 2004). As Eisenberg (2006: 9) observes, this approach is motivated by "the concern ... that, as important as universal accessibility to education (or any other public good) is, it is not, by itself, enough to ensure that all individuals are treated justly by institutions or, in the case of education, that the system is doing its utmost to secure social justice." Instead, justice demands a politics of distribution *and* recognition, notably including the positive recognition of cultural differences in the school or university classroom (Vincent, 2003: 3–4). Arguments here emphasize that educational content must be "inclusive", "diverse" and "multicultural", responding to the principle that socially just education recognizes cultural differences as valuable (Connell, 1992: 139). The question is not only access to education but to an education that allows for the recognition of a range of cultural meanings, against what Young refers to as "cultural imperialism" (Young, 1990: 36) or the imposition of the standards of a particular culture as if these were universally desirable. This dual concern with distribution and difference is captured in conceptualizations of socially just education as challenging "inequities and marginalization" (Theoharis, 2010: 333); in other words, focused on redistributing resources, especially to the most disadvantaged, while also positively recognizing hitherto absented knowledges and cultures in the classroom.

As Eisenberg (2006: 8) explains, "a politics that welcomes difference" may be understood as one critical element in a broader politics that seeks to challenge oppression, along but not limited to dimensions of both (re)distribution and recognition. Drawing on Young's (1988) well-known analytical framework delineating "the five faces of oppression", Gewirtz (1998) then offers an expansive list of questions that *anti*-oppressive social justice educators must ask:

"How, to what extent and why do education policies support, interrupt or subvert:

1. Exploitative relationships (capitalist, patriarchal, racist, heterosexist, disablist, etc.) within and beyond educational institutions?
2. Processes of marginalization and inclusion within and beyond the education system?
3. The promotion of relationships based on recognition, respect, care and mutuality or produce powerlessness (for education workers and students)?
4. Practices of cultural imperialism? And which cultural differences should be affirmed, which should be universalized and which rejected?
5. Violent practices within and beyond the education system?" (Gewirtz, 1998: 482).

In this view, a socially just education counters exploitation, marginalization, powerlessness, and (as noted earlier) practices of cultural imperialism and violence. In so doing, socially just (here *anti*-oppressive) education seeks to transcend a range of historically and locally situated inequities, related to class, gender, race, sexuality, disability and more, as experienced by actors in formal education systems, including both educators and students.

Other scholars turn to philosopher Sen's (1993) "capabilities approach", as well as Nussbaum's (2011) further development of this perspective (Wilson Strydom, 2015) to establish right relationships in education systems. For Sen, justice demands the enhancement of individual capabilities, so that each person may become and do in ways that are meaningful in their own critical self-evaluation (Sen, 1979). The question is thus less concerned with distribution and more focused on functional abilities, a matter of "asking the question 'what can people do?' (rather than how much do they have)" (Terzi, 2005: 387). From this perspective, socially just education supports children, youth and adults in developing the capacities they need to freely "choose and ... achieve" the ends each individual values (Walker and Unterhalter, 2007: 5). The capabilities approach to social education is less focused on redistribution and recognition, given inequitable economic and imperialistic cultural contexts, and more focused on individual freedoms to act.

The capabilities approach has been influential, especially in studies of education and development. At the turn of the millennium, the United Nations' Human Development Report (2000: 34) emphasized that, "The rights to food, health, education and privacy—as rights to capability building—were fundamental building blocks of the Universal Declaration of Human Rights."

Building capabilities through education is seen as critical to the exercise of human freedoms, and as Nussbaum (2006) argues, for healthy democratic decision-making, as well. Thus, a sound, socially just education supports students in developing the “habits of mind” that enable critical reflection (Nussbaum, 2006: 387), in a context of respect for all peoples and persons, cultivated through deliberative reasoning and an imaginative ability to empathize with others and express oneself. Socially just education informs democratic participation by supporting the free agency of individuals and respect for the free agency of others.

These three distinctive approaches to theorizing social justice in education variously emphasize the fair distribution of resources, the need for recognition of difference and to counter oppression, and the importance of supporting the development of individual capabilities. There are tensions across these approaches. Notably, the capabilities approach focusses on individual agency and freedom while distributive justice and recognition theories place a greater emphasis on relationships and challenging broad structures of inequity and oppression (Dejaeghere, 2020). Yet all three are rooted in Western philosophical accounts of justice, where these paradigms are diverse, sometimes conflicting but—it is increasingly recognized—far from universal. Next, we explore paradigms of social justice and education where the intellectual genealogies draw their strength and inspiration, in an immediate way, from critical engagement with a range of marginalized actors in Western nations and social movements in the global South. Such approaches may invoke general concepts, like oppression, but they insist on the historical and local specificities, that is, “the concrete situation” (Freire, 2000: 50) of socially just education.

Theorizing socially just education from below

Published in English in 1970, nearly two decades before Iris Marion Young’s (1988) theorization of the “five faces of oppression”, Paulo Freire’s *Pedagogy of the Oppressed* (2000) is one of the most-read texts in the field of education. Innovating from a Marxist intellectual heritage and developed when he was a teacher in Brazil and then in exile, Freire emphasizes that his approach is a response to the lived experiences of peasants and working-class people: “Thought and study alone did not produce *Pedagogy of the Oppressed*; it is rooted in concrete situations and describes the reactions of laborers (peasant or urban) and of middle-class persons whom I have observed directly or indirectly during the course of my educative work” (Freire, 2000: 37). If philosophical theorizing about social justice in education is often, even usually, inspired by injustice and social struggles, Freire makes this explicit as the motive informing a pedagogical project that seeks to be politically and existentially emancipatory for those who are oppressed under current economic and political relationships. In Brazil, Freire sought to counter oppression emerging from stark wealth inequalities, especially unequal access to land, and widespread political oppression and human rights abuses under military dictatorship.

In Freire’s (2000: 52) conception, the pedagogy of the oppressed is a “critical intervention in reality” through which the oppressed cease to be objects and instead become subjects who act in the world, freeing themselves from unjust relationships of domination and fatalistic resignation to the world-as-it-is to become fully human actors creating the world-as-it-may-be. In contrast to what he describes as the “banking” approach to education, in which knowing teachers deposit lessons in the heads of ignorant students, Freire (2000: 90) maintains that socially just pedagogy demands a shared commitment by both teachers and students “to the common task of learning and acting” rooted in genuine dialog carried out with humility and love. Dialogical, critical pedagogical practices encourage students and teachers to develop the freedoms necessary to create new worlds that transcend existing dehumanizing relationships of domination. As Freire explains, socially just education thus depends upon “a serious and profound effort at *conscientização*—by means of which the people, through a true praxis, leave behind the status of *objects* to assume the status of historical *Subjects*” (Freire, 2000: 160, italics in original).

If Freire responded to the concrete historical circumstances of peasants and workers in Brazil, in the 1960s, in Australia, Canada, New Zealand, Sweden and other “first-world” nations, Indigenous peoples were increasingly assertive about the need to challenge education systems participating in the genocide of Indigenous peoples. Since, there has been a long reckoning with the central role of education systems in these nations’ oppression of Indigenous peoples (Andersen et al., 2008; Battiste, 2013; Kuokkanen, 2005; Smith, 1993, 2000). Today, in Canada, for instance, the broader public has had to grapple with the aftermath of the Truth and Reconciliation Commission (2015), which investigated the residential schools that operated from 1867 to 1996. When civil servant Duncan Campbell Scott made attendance at residential school mandatory for Status Indians in 1920, he explained frankly that his objective was “to continue until there is not a single Indian in Canada that has not been absorbed into the body politic and there is no Indian question” (MacDonald, 2019: 71). Indigenous children recognized as “status Indians” under the 1876 *Indian Act* were separated from their families and forbidden from speaking their languages and carrying out traditional practices; many were subject to physical and sexual abuse, in chronically underfunded schools that provided poor housing and, often, inadequate food (MacDonald, 2019; Starblanket, 2018). Indigenous communities have long maintained that the residential schools are sites of mass graves, a fact now more widely accepted as radar technologies have recently confirmed the existence of the graves of hundreds of children, including 250 graves at the Kamloops Indian Residential School in the province of British Columbia and 751 graves at Cowessess First Nation in the province of Saskatchewan (Somos, 2021). In this and other national contexts of contemporary colonialism, Indigenous communities have argued that socially just education must respond to the specific, genocidal harms of settler colonial education systems.

For Indigenous actors, this demands both *decolonization*, challenging colonial education that values Eurocentric knowledges while eliminating Indigenous knowledges, and *indigenization*, the recentering of Indigenous ways of knowing, including Indigenous languages, in formal education from primary school through to university (Battiste, 2013; Gaudry and Lorenz, 2018; Munroe et al.,

2013; Tuck and Yang, 2019). Rooted in commitments to “land, sovereignty and Indigenous perspectives” (Tuck and Yang, 2019: xi), indigenization is about creating “educational spaces for Indigenous peoples, by Indigenous peoples” (Tuck and Yang, 2019: xi). This includes a range of initiatives, among them, language and cultural immersion classes for adult Mohawk learners in Canada (Iehnhotonkwash/Bonnie Jane Maracle, 2002), Sámi cultural learning through extracurricular arts classes in Sweden (Kallio and Lämsman, 2018), land-based pedagogies across diverse Indigenous polities in Canada and the United States (Ballantyne, 2014; Simpson, 2014; Tuck et al., 2014; Wildcat et al., 2014), and decolonizing scholarly inquiry, for instance, through Kaupapa Maori research methods in Aotearoa/New Zealand (Smith, 2008). Institutionally, in universities, indigenization has been pursued by developing Indigenous Studies programs and scholarship, including dedicated associations, like the North American Indigenous Studies Association (NAISA) and peer-reviewed journals, like the *International Journal of Critical Indigenous Studies*, *He Pukenga Korero: A Journal of Maori Studies* and *Wicazo Sa Review* (Andersen, 2020), which seek to critically engage with diverse Indigenous knowledges, on their own terms. For many Indigenous actors, restoring socially just education thus demands the transformation of educational institutions, simultaneously challenging deep-rooted genocidal and assimilationist aims in colonial education systems and creating new spaces within schools and universities for Indigenous people-specific and place-specific ways of being, knowing and doing.

Like Indigenous scholars, *anti-racism* educational experts have challenged existing formal education systems as foundationally unjust, hence requiring strategic interim measures and broader transformation. In particular, they have drawn attention to the ways that supposedly value-neutral pedagogy and scholarship reproduce Eurocentric, patriarchal understandings of the world and rewards middle-class White Anglo-saxon representations of appropriate behaviors in the classroom, from primary school through to graduate education (Dei, 1993; Dua and Lawrence, 2000; Henry et al., 2017; James and Howard, 2021; Troyna, 1987). In Canada, critical race scholars point, for instance, to the underrepresentation of racialized faculty in universities, especially Indigenous and Black professors in the senior ranks and senior administration (Henry et al., 2017; Mohamed and Beagan, 2019), while documenting the ways that many racialized children, especially Indigenous and Black children, are “pushed out” of the education system (Abdi and Ibrahim, 2016; Codjoe, 2001; Henry, 1993). Although Canada is often portrayed as a tolerant, multicultural society in comparison with the United States, *anti-racism* scholars concerned with the reproduction of racisms, observe that inclusion of racialized others is often tokenistic (James, 2001; James and Howard, 2021). Within this context, “visible minorities” and Indigenous peoples are socially constructed as multicultural others, while the educational cultures of the White majority are deemed neutral or universal, even when they narrowly valorize knowledges from those of White European and North American heritage (Rahim, 1990; Jay, 2003; Writer, 2008).

Consequently, *anti-racism* scholars have argued that socially just education requires schooling and universities that specifically redresses race inequities. Afrocentric education, emphasizing the diversity of Black and African knowledges (Asante, 1991; Dei, 1994; Kempf and Sefa Dei, 2020; Hampton, 2010; Thompson and Thompson, 2008), specifically centers “African people’s contributions to Canada, North America and the world” (James, 2010: 205). At the same time, unlike contemporary education which presumes race equity and tolerance, sometimes under the guise of “multiculturalism”, Africentric education in Canada explicitly recognizes historical and ongoing *anti-Black* racisms, so aiming to “challenge and problematize the stereotypes and myths about African peoples held by others” (James, 2010: 205). If such an education supports Black and African migrant students against *anti-Black* racism, in the Canadian context, critical race scholars and *anti-racism* activists emphasize that all youth benefit from schooling that explicitly grapples with race inequities, including as they are reproduced in educational content, pedagogical and learning styles. To those ends, *anti-racism* scholars emphasize that socially just education requires well-financed, institutional strategies to specifically support racialized students, notably including Black youth disproportionately pushed out of education system in nations like Canada (Howard and James, 2019).

Traditions of socially just education system, from *anti-oppressive*, decolonizing and *anti-racism* perspectives, draw their strength directly from broader social struggles. Freire drew on peasant and worker movements in elaborating a pedagogy of the oppressed that he hoped would be existentially liberating for those living in oppressive economic and political circumstances. Indigenous resistance and resurgence movements against colonialism inform challenges to a genocidal, assimilationist education systems, while imagining new/old ways of learning through diverse Indigenous worldviews. *Anti-racism* struggles against normalized and naturalized White supremacy motivate meaningful transformation, over tokenistic inclusion, in education, so as to recognize the diversity of human knowledges and to challenge historical and ongoing racisms. *Anti-oppressive*, decolonizing and *anti-racism* perspective have been developed paying attention to intersectionality in terms of critical disability models of education (Brown et al., 2021; Goodley, 2007; Liasidou, 2014; Parekh and Brown, 2020), LGBTQ+ and queer challenges to inequities in the classroom (Brockenbrough, 2015; MacGillivray, 2000), as well as a range of global South pedagogies—for instance, African educational systems rooted in the concept of *ubuntu* (Ngubane and Makua, 2021; Waghid et al., 2018)—which seek to re-imagine socially just education against the concrete, specific inequities of any given historical moment and place, by working in relationships with and for those from marginalized communities.

Critiques of social justice and education

If there are differences about what socially just education means, in theory and practice it might seem difficult to imagine critiques. After all, who could be against socially just education? Yet critiques of social justice in education have been made from both the right and from the left, or, alternatively stated, from both conservative and reform thinkers.

Inspired by positivist knowledge paradigms, conservative critics emphasize that education is concerned with sharing truth. Truth is established through rigorous argument, and, in the sciences, formally conceptualized in hypotheses that are then tested against empirical evidence. From this perspective, conjugating social justice and education represents a lamentable politicization of universal, value-neutral, rigorous, scientific knowing (Coburn et al., 2013). If conservative critics acknowledge that teachers and students alike may hold political and moral commitments to social justice, they maintain that, ideally, these are bracketed from free, open classroom discussions, in which the unifying principle is neither shared politics nor shared morality but a common commitment to rigorously specified argument. From this perspective, the idea of “social justice teaching” (Sleeter, 2015) threatens to replace meaningful pedagogy with the brainwashing of vulnerable students. Similarly, since rigorous learning is an universal endeavor, Indigenous and African-centered education are vigorously opposed as separatist initiatives that dangerously undermine the common educational basis for social cohesion and the democratic *polis*.

In contrast, reform critics raise quite different concerns about the idea of socially just (formal) education. Among these, the most thorough-going is arguably Ivan Illich’s (1973) critique (1973), since he argues that education systems should be abolished, in the name of social justice. In Illich’s view, education has become a service industry that disempowers the vast majority, especially the poor, who turn to professional institutions to become educated, so coming to distrust their own critical capacities for learning. In this context, in Illich’s estimation, “neither individual learning nor social equality can be enhanced by the ritual of schooling” (1973: 18). More educational institutions, even if they are reformed in important ways, will not result in more skills acquisition and knowledge creation. Instead, what is required is the de-schooling of society and the rejuvenation of pedagogy through informal mutual education among lay people. In this vision, “education for all means education by all” (Illich, 1973: 11), as ordinary people come together to casually develop and share skills and knowledges. For Illich, socially just education demands an end to formal schooling, and associated professionalization and credentialisms, from kindergarten to university. This opens up possibilities for a society that will enable learning as part of a new lifestyle in which we are “spontaneous, independent, yet related to each other” (Illich, 1973: 24), as learning is decoupled from formal, bureaucratic and top-down formal education systems.

Those wishing to defend the importance of socially just education must answer both conservative and reform critiques. Against conservatives, those concerned with socially just education question the underlying epistemological claims that education is about learning universal truths. As Kincheloe (2004: 53) observes, to be realistic, knowledge claims must instead “tak[e] into account the situatedness of the researcher and the researched [and the teacher and the taught]—where they are standing or are placed in the social cultural, historical, philosophical, economic, political, and psychological web of reality.” This parochializes all claims to knowledge, while insisting on a realistic appraisal of how all human beings seek out truth within given historical moments and unequal social relationships. There is no “god’s trick” (Haraway, 1988) that allows human beings access to truth outside of the social and material world in which we all live. Drawing on this insight, defenders of Indigenous and African-centered education observe that supposedly value-neutral classrooms are, in fact, White and often patriarchal, both with respect to curricular content and regarding embodied authority. Rather than replacing universal knowledge with particular knowledges, as conservatives charge, these programs instead counter the inequitable marginalization of Indigenous and African epistemologies and methodologies. In short, for defenders of social justice education, conservatives misconstrue social justice education as an attack on value-neutral education, but this assessment is rooted in the epistemologically erroneous assumption that knowledge can be separated out from social values and political context.

Against Illich, it might be argued that mutual learning in spontaneous groups may lead to strong hierarchies under charismatic leaders, whereas the formality of educational institutions may help curb the personal sway of individual teachers in ways that protect the young and vulnerable. Further, learning through spontaneous affinity groups may encourage people to associate in homophilic networks rather than encountering people from diverse social locations and with a range of political views. This may lead to narrow forms of learning, reproducing in-group thinking rather than producing challenging, critical knowledges among heterogeneous, often unequal publics. In this view, formal, institutionalized and, at its best, universally accessible education may be protective against the dangers of charismatic leaders and the reproduction of narrow, segmented, informal publics rooted in (real or perceived) shared interests. Ultimately, this defense of socially just formal education systems depends upon broad access to high-quality education.

Socially just education may seem difficult to critique, since social justice may appear to be a self-evident good. Nonetheless, there are persistent debates of this kind, especially concerning what conservatives see as the dangerous politicization of education and, less usually, critiques that argue for the abolition of formal education as a means to achieve truly socially just learning.

Socially just education in our times

We have so far presented different visions of socially just education, observing how these are informed by broader philosophical and social concerns with justice, especially the desire to end injustices. We have raised common criticisms against socially just education, from both conservative and reform minded scholars and activists, as well as responses to those criticisms. We conclude here by considering the contemporary moment’s broad range of struggles for social justice in education.

Today, scholarly concern for social justice in education is wide ranging, including, for instance:

- An analysis of barriers to full participation in education by those with non-normative bodies and psyches, in a social justice approach to disability that rejects identification of disability with “deficit, limitation or inability” (Evans et al., 2016: xiii):

- An exploration of queer pedagogy disrupting heteronormativity in education systems by advocating for equity across “all categories of (a)gender performances” (Miller quoted in Pennell, 2020: 2299);
- A call to invest in care for others as a purposeful subversion of neoliberal education, which commodifies teaching and learning as market transactions that are “fixed, priced and traded”, leaving little time and resources for the attentiveness to others that social justice demands (Hölscher, 2018: 44);
- Invitations to the Black Lives Matter movement to remember and participate in “righteous fights for racial justice” (Dixon, 2018: 240) and go beyond demands for inclusion in White educational institutions, instead contributing to radical “imaginations of an emancipated Black future” (Dancy et al., 2018: 190);
- An emphasis on the necessity of re-centering environmental and climate justice, as a social justice issue, in social work education (Beltrán et al., 2016).

As this cursory summary suggests, those concerned with social justice in education recognize that these questions are deeply intertwined with broader question of inequities across society as a whole. These include inequities related to Indigeneity, disability, lesbian, gay and queer emancipation, the neoliberal expansion of marketized relationships, *anti-Black* racism, and demands for ecological justice including that which is related to the existential crisis of climate change.

To those following Rawls, Fraser and Sen, achieving social justice in education in the context of these inequities demands attentiveness to the ways they are structured in and through the maldistribution of resources, failures to recognize the insights and knowledges of minoritized peoples, and inadequate attention to building human capacities. Yet as *anti-oppressive*, *anti-colonial* and *anti-racism* struggles, among others, have insisted, meaningful political action further requires grappling with the historically specific ways these relationships of oppression are maintained through class inequities, political dictatorships, historical and ongoing colonialisms, and White supremacy, as well as ableism, heterosexism, and climate injustice. In short, if education systems are neither separate from broader social injustices nor from the ways these inequities manifest in particular historical moments, then social justice in education is bound up, in profound ways, with the questions: what does the just society looks like and how can we act together, politically, to achieve it?

References

- Abdi, A.A., Ibrahim, A. (Eds.), 2016. *Education of African Canadian Children: Critical Perspectives*. McGill-Queen's Press-MQUP.
- Andersen, C., Bunda, T., Walter, M., 2008. Indigenous higher education: the role of universities in releasing the potential. *Aust. J. Indig. Educ.* 37 (1), 1–8.
- Andersen, C., 2020. The institutional and intellectual trajectories of indigenous studies in North America: harnessing the “NAISA effect”. In: *Routledge Handbook of Critical Indigenous Studies*. Routledge, pp. 9–22.
- Asante, M., 1991. The Afrocentric idea in education. *J. Negro Educ.* 60, 170–180.
- Ballantyne, E.F., 2014. Dechinta Bush University: mobilizing a knowledge economy of reciprocity, resurgence and decolonization. *Decolonization* 3 (3), 67–85.
- Barco, B., Carrasco, A., 2020. Access and equity policies in early childhood education: the case of Chile. *Early Years* 40 (4–5), 399–414.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. UBC press.
- Bell, L.A., 2016. Theoretical foundations for social justice education, 25–54. In: Adams, M., Bell, L.A., Goodman, D.J., Joshi, K.Y. (Eds.), *Teaching for Diversity and Social Justice*, third ed. Routledge, London.
- Beltrán, R., Hacker, A., Begun, S., 2016. Environmental justice is a social justice issue: incorporating environmental justice into social work practice curricula. *J. Soc. Work. Educ.* 52 (4), 493–502.
- Brennan, J., Naidoo, R., 2008. Higher education and the achievement (and/or prevention) of equity and social justice. *High. Educ.* 56, 287–3.
- Brockenbrough, E., 2015. Queer of color agency in educational contexts: analytic frameworks from a queer of color critique. *Educ. Stud.* 51 (1), 28–44.
- Brown, R.S., Parekh, G., Abdulkarim, F., 2021. *Special Education in the TDSB: A Re-examination of System Trends*. Technical Report. York University. Prepared for the Toronto District School Board, Toronto: Ontario.
- Coburn, E., Moreton-Robinson, A., Sefa Dei, G., Stewart-Harawira, M., 2013. Unspeakable things: indigenous research and social science. *Socio. La nouvelle revue des sciences sociales* 2, 331–334.
- Codjoe, H.M., 2001. Fighting a “public enemy” of Black academic achievement—the persistence of racism and the schooling experiences of Black students in Canada. *Race Ethn. Educ.* 4 (4), 343–375.
- Connell, R.W., 1992. Citizenship, social justice and curriculum. *Int. Stud. Sociol. Educ.* 2 (2), 133–146. <https://doi.org/10.1080/0962021920020202>.
- Connell, R.W., 1993. *Schools and Social Justice*. Temple University Press, Philadelphia.
- Dancy, T.E., Edwards, K.T., Davis, J.E., 2018. Historically white universities and plantation politics: anti-Blackness and higher education in the Black Lives Matter era. *Urban Educ.* 53 (2), 176–195.
- Dei, G.J.S., 1993. The challenges of anti-racist education in Canada. *Can. Ethn. Stud.* 25 (2), 36.
- Dei, G.J.S., 1994. Afrocentricity: a cornerstone of pedagogy. *Anthropol. Educ. Q.* 25 (1), 3–28.
- Dejaeghere, J.G., 2020. Reconceptualizing educational capabilities: a relational capability theory for redressing inequalities. *J. Hum. Dev. Capab.* 21 (1), 17–35.
- Dixon, A.D., 2018. “What’s going on?”: a critical race theory perspective on Black Lives Matter and activism in education. *Urban Educ.* 53 (2), 231–247.
- Dua, E., Lawrence, B., 2000. Challenging white hegemony in university classrooms: whose Canada is it? *Atlantis* 24 (2), 105–122.
- Eisenberg, A., 2006. Education and the politics of differences: Iris Marion Young and the politics of education. *Educ. Philos. Theor.* 38 (1), 7–23.
- Evans, N.J., Broido, E.M., Brown, K.R., Wilke, A.K., 2016. *Disability in Higher Education: A Social Justice Approach*. John Wiley & Sons.
- Fraser, N., 1995. From redistribution to recognition? Dilemmas of justice in a “post-socialist” age. *N. Left Rev.* 212, 68–93.
- Freire, P., 2000. *Pedagogy of the oppressed*. Continuum International Publishing Group, New York.
- Gaudry, A., Lorenz, D., 2018. Indigenization as inclusion, reconciliation, and decolonization: navigating the different visions for indigenizing the Canadian Academy. *Alternative* 14 (3), 218–227.
- Gewirtz, S., 1998. Conceptualizing social justice in education: mapping the territory. *J. Educ. Pol.* 13 (4), 469–484. <https://doi.org/10.1080/0268093980130402>.
- Giroux, H.A., 2003. Public pedagogy and the politics of resistance: notes on a critical theory of educational struggle. *Educ. Philos. Theor.* 35 (1), 5–16.
- Goodley, D., 2007. Towards socially just pedagogies: deleuzoguattarian critical disability studies. *Int. J. Incl. Educ.* 11 (3), 317–334.
- Hampton, R., 2010. Black learners in Canada. *Race Class* 52 (1), 103–110.

- Haraway, D., 1988. Situated knowledges: the science question in feminism and the privilege of partial perspective. *Fem. Stud.* 14 (3), 575–599.
- Henry, F., Dua, E., James, C.E., Kobayashi, A., Ali, P., Ramos, H., Smith, M.S., 2017. *The Equity Myth: Racialization and Indigeneity at Canadian Universities*. University of British Columbia Press, Vancouver.
- Henry, A., 1993. Missing: Black self-representations in Canadian educational research. *Can. J. Educ.* 206–222.
- Hölscher, D., 2018. Caring for justice in a neoliberal university. *S. Afr. J. High. Educ.* 32 (6), 31–48.
- Howard, P.S.S., James, C.E., 2019. When dreams take flight: how teachers imagine and implement an environment that nurtures Blackness at an Africentric school in Toronto, Ontario. *Curric. Inq.* 49 (3), 313–337.
- Huttunen, R., Heikkinen, H.L.T., 2004. Teaching and the dialectic of recognition. *Pedagog. Cult. Soc.* 12 (2), 163–174.
- Iehnotonkwas/Bonnie Jane Maracle, 2002. Adult Mohawk language immersion programming. *McGill J. Educ.* 37 (.3), 387.
- Illich, I., 1973. *Deschooling Society*. Harrow Books.
- James, C.E., Howard, P.S.S., 2021. The education system of Canada: multiculturalism, multicultural education, and racialized students in Canada. In: Jorntz, S., Parreira do Amaral, M. (Eds.), *The Education Systems of the Americas*. Global Education Systems. Springer, Cham. https://doi.org/10.1007/978-3-030-41651-5_44.
- James, C.E., Parekh, G., 2021. Fixed trajectories: race, schooling and graduation from a Southern Ontario university. *Can. J. High. Educ.* 51 (4), 67–84.
- James, C.E., 2001. Multiculturalism, diversity, and education in the Canadian context: the search for an inclusive pedagogy. In: Grant, C.A., Lei, J. (Eds.), *Global Constructions of Multicultural Education: Theories and Realities*, pp. 175–204.
- James, C.E., 2010. Multicultural education in a color-blind society, pp. 205–224. In: Grant, C.A., Portera, A. (Eds.), *Intercultural and Multicultural Education: Enhancing Global Interconnectedness*. Routledge, London.
- Jay, M., 2003. Critical race theory, multicultural education, and the hidden curriculum of hegemony. *Multicult. Perspect.* 5 (4), 3–9.
- Kallio, A.A., Lämsman, H., 2018. Sami re-imaginings of equality in/through extracurricular arts education in Finland. *Int. J. Educ. Arts* 19, 7.
- Kempf, A., Sefa Dei, G.J., 2020. Afrocentric education in North America: an introduction. In: Abidogun, J., Falola, T. (Eds.), *The Palgrave Handbook of African Education and Indigenous Knowledge*. Palgrave Macmillan.
- Kincheloe, J., 2004. The knowledges of teacher education: developing a critical complex epistemology. *Teach. Educ. Q.* 49–66.
- Kuokkanen, R., 2005. Láhi and attáldat: the philosophy of the gift and Sami education. *Aust. J. Indig. Educ.* 34, 20–32.
- Liasidou, A., 2014. Critical disability studies and socially just change in higher education. *Br. J. Spec. Educ.* 41 (2), 120–135.
- MacDonald, D.B., 2019. *The Sleeping Giant Awakens: Genocide, Indian Residential Schools, and the Challenge of Conciliation*. University of Toronto Press.
- MacGillivray, I.K., 2000. Educational equity for gay, lesbian, bisexual, transgendered, and queer/questioning students: the demands of democracy and social justice for America's schools. *Educ. Urban Soc.* 32 (3), 303–323.
- Matsuura, K., 2006. Foreword. pp. v–vi. In: Oduaran, A., Bhola's, H.S. (Eds.), *Widening Access to Education as Social Justice : Essays in Honor of Michael Omolewa*. Springer, The Netherlands.
- Mohamed, T., Beagan, B.L., 2019. "Strange faces" in the academy: experiences of racialized and Indigenous faculty in Canadian universities. *Race Ethn. Educ.* 22 (3), 338–354.
- Munroe, E., Borden, L., Murray Orr, A., Toney, D., Meader, J., 2013. Decolonizing aboriginal education in the 21st century. *McGill J. Educ.* 48 (2), 317–337. <https://doi.org/10.7202/1020974ar>.
- Ngubane, N.I., Makua, M.J., 2021. Intersection of Ubuntu pedagogy and social justice: transforming South African higher education. *Transform. High. Educ.* 6, 1–8.
- Nussbaum, M.C., 2006. Education and democratic citizenship: capabilities and quality education. *J. Hum. Dev.* 7 (3), 385–395.
- Nussbaum, M., 2011. *Creating Capabilities*. Harvard University Press, Cambridge, MA.
- Oduaran, A., 2006. Conceptualizing the widening of access to education as social justice, pp. 69–82. In: Oduaran, A., Bhola's, H.S. (Eds.), *Widening Access to Education as Social Justice : Essays in Honor of Michael Omolewa*. Springer, The Netherlands.
- Parekh, G., Brown, R.S., 2020. Naming and claiming: the tension between institutional and self-identification of disability. *Can. J. Disabil. Stud.* 9 (5), 347–379. <https://doi.org/10.15353/cjds.v9i5.701>.
- Pennell, S.M., 2020. Queer theory/pedagogy and social justice education. In: *Handbook on Promoting Social Justice in Education*, pp. 2291–2311.
- Polat, F., 2011. Inclusion in education: a step towards social justice. *Int. J. Educ. Dev.* 31 (1), 50–58.
- Rahim, A., 1990. Multiculturalism or ethnic hegemony: a critique of multicultural education in Toronto. *J. Ethn. Stud.* 18 (3), 29.
- Rawls, J., 1972. *A Theory of Justice*. Clarendon Press, Oxford.
- Sehoole, C., Adeyemo, K.S., 2016. Access to, and success in, higher education in post-apartheid South Africa: social justice analysis. *J. High. Educ. Afr.* 14 (1), 1–18.
- Sen, A., 1979. Equality of what?, pp. 197–220. In: McMurrin, S. (Ed.), *Tanner Lectures on Human Values*. Cambridge University Press, Cambridge.
- Sen, A., 1993. Capability and well-being. In: Nussbaum, M., Sen, A. (Eds.), *The Quality of Life*. Oxford University Press.
- Simpson, L.B., 2014. Land as pedagogy: nishnaabeg intelligence and rebellious transformation. *Decolonization* 3 (3), 1–25.
- Sleeter, C.E., 2015. Deepening social justice teaching. *J. Lang. Lit. Educ.* 42 (6), 512–535.
- Smith, L.T., 1993. Getting out from down under: Maori women, education and the struggles for Mana Wahine. In: Arnot, M., Weil, K. (Eds.), *Feminism and Social Justice in Education: International Perspectives*. The Falmer Press, London, pp. 58–78.
- Smith, G.H., 2000. Maori education: revolution and transformative action. *Can. J. Native Educ.* 24 (1), 57–72.
- Smith, L.T., 2008. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books Ltd.
- Somos, C., 2021. Residential School Graves "Totally Pushed in the Wayside" in Shadow of Election Indigenous Leaders Say. <https://www.ctvnews.ca/politics/federal-election-2021/residential-school-graves-totally-pushed-to-the-wayside-in-shadow-of-election-indigenous-leaders-say-1.5550433>. (Accessed 17 August 2021).
- Starblanket, T., 2018. *Suffer the Little Children: Genocide, Indigenous Nations and the Canadian State*. Clarity Press.
- Taylor, R., 2004. Self-realization and the priority of fair equality of opportunity. *J. Moral Philos.* 1 (3), 333–347.
- Terzi, L., 2005. A capability perspective on impairment, disability and special needs: towards social justice in education. *Theor. Res. Educ.* 3 (2), 197–223.
- Theoharis, G., 2010. Disrupting injustice: Principals narrate the strategies they use to improve their schools and advance social justice. *Teach. Coll. Rec.* 112 (1), 331–373.
- Thompson, D., Thompson, J., 2008. Afrocentric schools and the politics of invisibility. *Arden* 1 (1), 45–49.
- Troyna, B., 1987. Beyond multiculturalism: towards the enactment of anti-racist education in policy, provision and pedagogy. *Oxf. Rev. Educ.* 13 (3), 307–320.
- Truth and Reconciliation Commission, 2015. *Honouring the Truth, Reconciling the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. McGill-Queen's University Press, Montreal and Kingston.
- Tuck, E., Yang, K.W., 2019. Series editors' introduction. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge.
- Tuck, E., McKenzie, M., McCoy, K., 2014. Land education: indigenous, post-colonial, and decolonizing perspectives on place and environmental education research. *Environ. Educ. Res.* 20. Number 1, 2 January 2014, pp. 1–23(23).
- Vincent, C., 2003. *Social Justice, Education and Identity*. Routledge.
- Waghid, Y., Faiq, W., Waghid, Z., 2018. *Rupturing African Philosophy on Teaching and Learning: Ubuntu Justice and Education*. Springer.
- Walker, M., Unterhalter, E., 2007. The capability approach: its potential for work in education, pp. 1–18. In: Walker, M., Unterhalter, E. (Eds.), *Amartya Sen's Capability Approach and Social Justice in Education*. Palgrave Macmillan, New York.
- Weingarten, H.P., 2021. *Nothing Less Than Great: Reforming Canada's Universities*. University of Toronto Press, Toronto.
- Wildcat, M., McDonald, M., Irlbacher-Fox, S., Glen, C., et al., 2014. Learning from the land: Indigenous land based pedagogy and decolonization. *Decolonization* 3 (3).
- Wilson-Strydom, M., 2015. University access and theories of social justice: contributions of the capabilities approach. *High. Educ.* 69 (1), 143–155.

- Writer, J.H., 2008. Unmasking, exposing, and confronting: critical race theory, tribal critical race theory and multicultural education. *Int. J. Multicult. Educ.* 10, 2.
- Young, I.M., 1988. Five faces of oppression. *Phil. Forum* 19 (4), 270–290.
- Young, I.M., 1990. *Justice and the Politics of Difference*. Princeton University Press. Princeton.
- Zapata-Díaz, J., 2008. Social justice: education and social justice. *Int. Rev. Educ.* 52 (1/2), 9–22.

This page intentionally left blank

Culture

Yao Xiao and Handel Kashope Wright, University of British Columbia, The Center for Culture, Identity and Education, Faculty of Education, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	21
Culture in sociology and anthropology of education	21
Multicultural education	23
Cultural studies and education	25
Culture and indigeneity: culturally rooted education, land education, and an ethics of incommensurability	26
Conclusion	28
References	28

Introduction

Culture, according to British cultural studies founding father, Raymond Williams (1983), traverses linguistic, social, and intellectual movements in English and European systems of thoughts—signifying multiple, complicated, intertwining, and sometimes problematic meanings and practices of “cultivation”, “civilization”, “art and learning”, “material production”, a whole way of life, dynamics of domination/subordination, and various conceptual developments in academic disciplines such as anthropology (pp. 87–93). In *An aesthetic education in the era of globalization*, South Asian feminist postcolonialist theorist, Gayatri Chakravorty Spivak (2012) refers to culture as “a package of largely unacknowledged assumptions, loosely held by a loosely outlined group of people, mapping negotiations between the sacred and the profane, and the relationship between the sexes ... But these presuppositions also give us the wherewithal to change our world, to innovate and create” (p. 120). Taking culture more specifically as a site for critical, transgressive pedagogical process, bell Hooks (1994) explores African American representations and speaks affirmatively to “the heartbeat of cultural revolution” where cultural criticism can be an educational project for critical consciousness in liberatory ways, “only if we start with a mind-set and a progressive politics that is fundamentally anticolonialist, that negates cultural imperialism in all its manifestations” (p. 7). Smith et al. (2019) (who are from New Zealand/Aotearoa, Canada and the United States/Turtle Island) collectively assert that “Water is life; Land is our first teacher” (p. 1) as a foundational cultural statement agreed upon by far flung Indigenous peoples on the inherently culture-based relationship between Indigenous people and the Land, and as part of the bedrock of the work in Indigenous and decolonizing education. These examples illustrate that culture is both related to pedagogy and itself pedagogical, foundational to making meaning of the everyday and the sacred across time and space.

Culture is not only profound and deeply consequential: Williams (1958/1989) has also declared, rather cryptically, that “culture is ordinary, in every society and every mind” (p. 4). This is a declaration that is deceptively simple. What Williams does through this statement is to challenge the idea of culture as the purview of the elite (e.g., to be cultured is to be middle class or upper class, educated and cosmopolitan; cultural forms are elite forms such as ballet, opera and high cuisine, much of which is expressed in foreign languages). Instead of accepting the taken for granted conception of “culture” as special, rare, and extraordinary, Williams argues that working class culture and popular culture are also forms of culture and should be valued and even studied as such. Culture, he argues, is in fact possessed by all in society, and in every society; it is expressed not only in foreign languages deemed elite but in every local language and accent. Given Williams’ reconception, to know factory floor culture is as valid as knowing boardroom culture; breakdancing, hip hop and spoken word are as worthy of study as ballet and opera and religious incantation; peasant food is as interesting as gourmet cuisine; Kikuyu is as valid a medium for culture in Kenya as English. In other words, culture is everywhere, possessed and expressed by everyone: in short, culture is ordinary. To be educated in the sense of this redefinition of culture is to acknowledge both elite and popular culture, and at times to obfuscate the boundaries between them (think of Pavarotti’s collaborations with Sting for example or the rock group Queen’s “Bohemian Rhapsody”).

As these varied, nuanced conceptualizations illustrate, culture is a complex concept, a veritable floating signifier, definitions of which shift with geographical and especially disciplinary and discursive context. This essay explores the intricate relationship between culture and the field of education, or more specifically the place of the multiple, complex concept of culture and the cultural in the making of forms of critical education. The essay emphasizes the socio-political dynamics of that relationship, in four forms of critical, equity and social justice oriented education, namely critical sociology and anthropology of education, critical multicultural education, cultural studies of education, and Indigenous and decolonizing education.

Culture in sociology and anthropology of education

A substantial part of sociological studies in education addresses questions about and is related to the politics of culture concerning access, reproduction, domination, subordination, and agency (e.g., Apple et al., 2010; Bloomfield and Fisher, 2016; Koh et al., 2020). Cultures, especially school cultures, mediate the process of students’ identity-making, boundary-making and social-

material networks (e.g., McDougall et al., 2004; Carter, 2006; Perrotta et al., 2018), while being part and parcel of local educational discourses, policies, programs, and practices (e.g., Smith, 2012a; Phipps and Young, 2015; Toh, 2020). The concept “cultures of schooling”, for example, examines the ways in which aspects of sociocultural identity such as race, social class, gender, sexuality, ability, migration history, citizenship, etc. contribute to social stratification and operate in constituting school choice, family-school partnership, admission policy, the assemblage of digital technology in learning, as well as students’ experiences of alienation or inclusion in school. For Koh et al. (2020), “critiquing dominant cultures of schooling” is a sustained critical enterprise in sociology of education, especially in the current context of the COVID-19 pandemic and persistent racial injustice affecting societies worldwide (p. 199).

One of the most influential theories on how culture mediates social structure and agency in education is “cultural capital” (Bourdieu and Passeron, 1977). With its micro manifestation as “habitus”, cultural capital is a theory proposed by French sociologist Pierre Bourdieu to understand and critique the process of social domination and reproduction. Most manifest in education (especially in the formal educational system), cultural capital is a form of capital that reflects social class interests and oftentimes the tacit processes through which a hierarchy of cultures (e.g., bodies of knowledge, skillsets, abilities, proficiencies and ways of learning) is systemically reproduced and where people possess, appropriate and utilize these aspects of culture differentially and strategically. Culture—capitalized for material and symbolic privileges and used to mark inter-group distinctions—is thus inextricably linked to social status, power, and domination. It has long been argued that schooling and students’ success are linked to the societal norms that define and evaluate cultural capital—some cultural forms or indeed “cultures” have higher exchange value convertible to a multitude of economic goods and social network resources, while other cultural forms and “cultures” are relegated to lower exchange value or devalued through the educational system (e.g., DiMaggio, 1982). Cultural capital theory has also been applied in the socio-geographical studies of education, beyond the conventional notion of “society” within national borders, especially in dealing with international (middle class and elite) migration, transnational family strategies, international education, and global class mobility (e.g., Waters, 2006; Igarashi and Saito, 2014; Zhang and Xu, 2020).

In extensions of Bourdieu’s primarily social class based conception of cultural capital, some sociological studies in education have revealed that cultural capital relates not only to social class, but also to gender (Dumais, 2002; Sheng, 2012) and race (Lareau and Horvat, 1999; Wallace, 2017; Richards, 2020). Going even further and utilizing a Critical Race Theory (CRT—which originated in critical legal studies in the United States, e.g., Crenshaw et al., 1995) lens to understand inextricable layers of racialized subordination based on intersections of race with gender, class, immigration status, surname, phenotype, accent and sexuality, Yosso (2005) challenged traditional interpretations of Bourdieusian cultural capital theory. In doing so, he introduced an alternative concept “community cultural wealth,” which decenters White middle class cultures and has the potential to transform the schooling process, especially for communities of color. Yosso (2005) presented at least six forms of capital, most of which go unacknowledged or unrecognized, that comprise community cultural wealth, namely aspirational capital, linguistic capital, familial capital, social capital, navigational capital, and resistant capital.

This Critical Race Theory (CRT) informed concept of community cultural wealth not only signifies a departure from the conventional conception of “cultural capital”, but also marks an important development in anthropology of education, especially in re-assessing “culture”—which has long been approached as a whole way of life in anthropology, has problematic colonial entanglements that serve the overall colonizing agenda (Mudimbe, 1988) and the interest of the “colonial administrator” (Hall, 2016, p. 15), and remains a highly contested concept among anthropologists of education (Niesz, 2001; Levinson and Pollock, 2011). To challenge and dismantle the colonial gaze of “othering”, some have pointed out that anthropology of education should contribute (and indeed is contributing) to CRT (Villenas, 2019), and some have envisioned a future of research and praxis guided by CRT and critical language studies (Deyhle and McCarty, 2007). For example, the life work of Lakota scholar Beatrice Medicine has challenged racism in formal and informal systems of learning, and considered Indigenous community-based education, especially the school-based promotion of Indigenous languages, as essential for Indigenous self-determination and cultural survival (Medicine and Jacobs, 2001). As Deyhle and McCarty (2007) argue, learning from the legacy of Beatrice Medicine enables anthropologists of education to view culture in the triangular power dynamics of race, language, and community, and requires a commitment to the understanding that the relationship between culture and education cannot be adequately understood without seriously confronting racism, linguisticism, and the effects of colonization.

With this critical stance, three interrelated approaches in anthropology of education have marked developments in re-assessing “culture”, especially in ways that represent a more asset, resource, and community based approach to culture. One example which has emerged in teacher education in the United States and was proposed by Moll et al. (1990) emphasizes “Funds of Knowledge (FoK)”. Moll et al. (1990) documented how mainstream teachers in the United States failed to acknowledge and build upon the wealth and strength of cultures in working class Latinx (then predominantly named “Hispanic”) communities, in particular their families and household activities as important social and cognitive resources for education. Moll (2005) and Hogg (2011) exhort teachers to seriously engage aspects of students’ FoK across various ethnoracial communities, including family culture, peer culture, community culture, popular culture, place-based knowledge, etc.

Another approach is what has been identified as “Community-Based Educational Spaces (CBES)”, exemplified by Lee et al. (2020)’s ethnographic project with a grassroots, community-based, nonprofit organization in building cross-racial solidarity between low-income Southeast Asian American youth and Black youth. Lee et al. (2020) argue that CBES, which have long been utilized among immigrant communities beyond traditional schools, could play a unique role in teaching youth to engage in anti-racist work by providing opportunities to have conversations on and make sense of the similarities and differences between anti-Blackness and anti-Asian-ness, and to build Black and Asian solidarity with a focus on addressing white supremacy.

A third approach focuses on the concept and practices of “culturally sustaining pedagogy (CSP)”, which was developed in the context of changing demography toward a majority multilingual society of color in the United States and involved advocating for more complex pedagogical approaches to the nuanced, fluid and transnational dynamics between race, ethnicity, language, and cultural ways of being (Paris, 2012; Paris and Alim, 2014). CSP has been utilized in ethnographic studies of racialized immigrant youth in schools (Jaffe-Walter and Lee, 2018) and racialized immigrant workers in adult education programs (Emerick et al., 2020), and has also been adapted as “culturally sustaining/revitalizing pedagogy (CSR/P)” —in the context of ongoing colonization and colonial erasure—employed to understand the linguistic, cultural, and educational realities of Native American communities and Native Hawaiian communities (McCarty and Lee, 2014; Kaomea et al., 2019). For Paris and Alim (2014), CSP is distinct from the past formulation of resource pedagogies in three aspects. First, they emphasize that culture is living and evolving—thus to be culturally sustaining is to support not only heritage but also the dynamic, fluid cultural practices of young people and their communities. Second, CSP recognizes that linguistic and cultural flexibility has an instrumental purpose for both students of color and White students, as multicultural, multilingual skills are increasingly linked to access and power in US and global contexts. Third, CSP promotes a critical analysis of oppressive practices (e.g., misogyny, heteronormativity) within the very cultural formations it seeks to sustain, as well as within the dominant culture.

As these new approaches emphasize a reflexive empowerment of immigrant, racialized and marginalized communities, and as culturally sustaining pedagogy is considered “a powerful next-generation articulation” (Lee, 2017, p. 261) of culturally relevant pedagogy (Ladson-Billings, 1995, 2014) aligned with multiculturalism and multilingualism in educational perspectives and practices (Paris and Alim, 2014), it makes sense to turn to the field of multicultural education, which also involves sociocultural identity and diversity but operates not along disciplinary lines but rather as a philosophy and praxis that addresses representation across the board in educational settings.

Multicultural education

As a philosophy and policy framework, multiculturalism is a popular approach to addressing sociocultural diversity in society and in education, especially in certain Western countries such as Canada (with interculturalism in Quebec), the United States, and Australia. While some philosophers (e.g., Kymlicka, 2012) trace multiculturalism’s origin back to post-World War II, West-led definition of human rights discourse, the movements of multiculturalism in the West were full of national variations, regional metamorphosis, ruptures, and contradictions. In Canada, multiculturalism originated in attempts to bridge the political (and cultural) divide between Anglophone and Francophone Canada, within a bilingual framework, and operated “as a successor regime to race relations which in turn was an improvement on assimilation and explicitly exclusivist and racist immigration policies that reflected Eurocentric conceptions of the nation” (Wright, 2012, p. 104; Taylor, 2012). In Australia, multiculturalism was first presented as the basis for migrant settlement and social-cultural policy in the 1970s, and conceived as a shift from the White Australia policy to a more integration-oriented discourse in recognition of diverse immigrants’ cultural identities (Parliament of Australia, 2010). Multicultural education in this context represents politics of differences that challenge and unsettle the Anglo-Celtic, Europe-centric reproduction of Australian-ness and Australian nation-building (Hodge and O’Carroll, 2006; Ang, 2009). “The multicultural question” in the postcolonial, diasporic formation of British-ness (Hall, 2000) with debates on “multicultural citizenship” related to ethnoracial and religious differences (Modood, 2013) in UK, and the declared “failure” of multiculturalism in Europe resulting in a turn to interculturalism (Chin, 2017), are but some other examples.

Less known is the fact that multiculturalism and multicultural education are not the exclusive purview of the West. Singapore for example, made official and legal commitments to multiculturalism (Keong, 2013), with policies in language and education, and with a particular governance framework of postcolonial “multiracialism” currently structured around the constitutive ethnoracial groups: Chinese-Malay-Indian-Others (CMIO) (Huat, 2003; Goh, 2017). In South Africa, the post-apartheid discourse of “Rainbow Nation” is in operation—from the politics of truth and reconciliation (Aiken, 2016) to the architecture of places and space (Raedt, 2012) to the realities of continued racial inequalities in schools (Teeger, 2015) and to the “sandwiched” position of some minority communities such as the Indian South Africans (Vahed and Desai, 2017).

The formation of multicultural education in the United States exemplifies the complexity, capacity and changes of multicultural education, reflected by the work of many critical educators such as James Banks and Christine Sleeter—two of the most prominent proponents in articulating American multicultural education. Banks (2013) traces the construction and historical development of multicultural education with its embryo emerged from the Civil Rights Movement where debates on culture, schooling, and education were inseparable from African American struggles around racial segregation and citizenship rights. The subsequent development of ethnic studies in the 1970s involved political analysis of multiple racialized and ethnic groups, which called for and became the bedrock of multiethnic education. As multiethnic education through the 1980s expanded to include the intersected analysis of social class, gender, religion, sexuality, exceptionality, etc., broader conceptions of diversity education started to converge into new movements of multicultural education since the late 1980s. A central goal of multicultural education, then, was to transform the school and other educational institutions so that students from diverse groups will have an equal chance to experience educational equality, success, and mobility (Banks, 1992). Other prominent scholars in multicultural education (e.g., Sleeter, 1996, 2009, 2013; May and Sleeter, 2010) have argued that critical multicultural education involves social activism, requires teaching for social justice in classrooms across disciplines, and includes developing teachers’ epistemological sophistication in teacher education. Culture, in this vein, is neither essentialized nor depoliticized. Instead, cultures and identities are framed in

the context of unequal power relations, “as multilayered, fluid, complex, and encompassing multiple social categories, and at the same time as being continually reconstructed through participation in social situations” (May and Sleeter, 2010, p. 10).

One movement in multicultural education is propelled by educational projects based on the cultural difference paradigm, in contrast to cultural deprivation paradigm and rejecting the idea that students of color have cultural deficits. Cultural conflicts and how the educational system responds to such conflicts—rather than intrinsic cultural deficits and individual efforts—explain students’ “(under)achievement”. In this paradigm, educators insist that schools must change in ways that respect, reflect, and responded to the rich cultural strengths of students from diverse ethnoracial groups and use teaching strategies that are consistent with their cultural characteristics, resources, and meanings (Banks, 2013). These changes and strategies—emerged primarily from the contexts of teaching/learning with racialized and marginalized students—are known variously as “equity pedagogy” (Banks and Banks, 1995; Banks, 2013), “culturally relevant pedagogy” (Ladson-Billings, 1995, 2014), and “culturally responsive pedagogy” (Gay, 2000, 2010). For example, culturally relevant pedagogy—emerging from research and interviews with teachers known for their pedagogical excellence in a predominantly African American, low-income elementary school district (Ladson-Billings, 1995)—becomes a model that simultaneously supports students’ academic achievement, affirms students’ cultural identity, competence and integrity, and enables students to critically analyze social inequalities that are relevant to their communities and belongings.

Another related movement in multicultural education is around re-evaluating cultural assumptions, frames of reference, and the value-laden construction of knowledge. Specifically, curricula are reconceptualized in ways that help students understand how knowledge is created, by critically evaluating different kinds of knowledge. These kinds of knowledge—contextualized and scaffolded for students’ learning—include personal/cultural knowledge (e.g., experience from home, families, and community cultures), popular knowledge (e.g., media and pop culture), mainstream academic knowledge (traditional Western-centric in history and the behavioral and social science), transformative academic knowledge (e.g., challenging/substantially revising established canons), and school knowledge (e.g., representations in textbook and lectures by teachers). At a more general level, multicultural education is a process that involves personal, institutional, and systemic changes, through constant inter-subjective construction of knowledge, explication of positionalities and cultural assumptions, creating institutional cultures that value equity and diversity, and critical inter-cultural understanding against discrimination and oppression. Banks (1994), for example, elaborated five dimensions in multicultural education, including content integration, the knowledge construction process, prejudice reduction, an equity pedagogy, and an empowering school culture and social structure. One key goal of multicultural education, according to Banks (1994, p. 5), is to help students “develop the knowledge, skills, and attitudes needed to participate in reflective civic action” in a pluralistic democratic society. In this vein, different approaches to citizenship education, analysis of global and regional migration, as well as international comparative studies in diversity education, have also become the concerns of multicultural education (e.g., Banks, 2009).

This international comparative approach was utilized by Cha et al. (2018) in their exploration of the uneven landscapes of multicultural education in contexts of what they called the “Asia Pacific”, with contributions from scholars doing research projects on diversity and education in Australia, Cambodia, Hong Kong, Indonesia, Japan, Laos, Mainland China, Singapore, South Korea, and Thailand. While the selected societies differ significantly in sizes, political structures, interpretations and issues of diversity, migration patterns, and the (existence/development/lack of) official multiculturalism policies, Cha et al. (2018)’s collection of essays does not prescribe a unified approach to multicultural education but rather signify the contextual geographical sensitivity to diversity and how multicultural education (entangled with different states) operates (or remains absent/unarticulated) in realms such as “conceptualization and perspectives”, “teaching, learning, and curriculum”, “leadership and policy”, and “equity and social justice in the context of diversity” (p. 6).

Examining the status quo and possible future of multicultural education internationally and in conjuncture with policies/practices of multiculturalism, Wright et al. (2012), scholars from Canada, Australia and the UK respectively, discuss how multiculturalism is infused with pedagogical dimensions such as youth identities, global intellectual equality, media pedagogy, multicultural literacy, the critical reading of intersected class elitist, white supremacist, and patriarchal power blocs, as well as potential alternatives to multicultural education. Although caught between “historical and contemporary dissent from the political left and right; continuing post 9/11, post 7/7 bombings backlash on the one hand and the related rise of alternatives like cosmopolitanism and interculturalism on the other” (Wright et al., 2012, p. 11), multicultural education remains an ongoing political project that enables us to learn about student culture, identity, power differentials, diverse epistemologies, and the (globalized but geopolitically divided?) world.

Given the lay of land internationally as articulated by Wright et al. (2012), what kinds of politics on cultural diversity are viable and inspirational at the present global historical moment (characterized by continued neoliberalism and neo-colonialism and a global pandemic)? While multiculturalism and multicultural education in their various forms and along the political spectrum (liberal, celebratory multicultural education remains dominant but there are also conservative, critical and revolutionary forms) might have some answers, we could also turn to the more consistently critical and educationally open discourse of cultural studies and cultural studies of education for more pointed responses to the status quo. As long ago as the 1990s and in the US context, for example, bell Hooks (1994) connected critical educational projects to cultural studies for its transdisciplinary stance and political sensitivity to the everyday extramural sites of learning. In doing so she is influenced by prominent British cultural studies figure, Stuart Hall who declared that “the work that cultural studies has to do is to mobilize everything it can find in terms of intellectual resources in order to understand what keeps making the lives we live, and the societies we live in, profoundly and deeply anti-human” (Hall, 1989, as quoted in bell Hooks, 1994, p. 4). Taking this connection seriously, we examine the relationship between cultural studies and education in the next section.

Cultural studies and education

Considering “cultures” as whole ways of life, as Raymond Williams insisted, and whole ways of communicating, Stuart Hall (1992) locates cultural studies at “the dirty crossroads where popular culture intersects with the high arts; that place where power cuts across knowledge or where cultural processes anticipate social change” (p. 10). In *The Long Revolution*, Raymond Williams (1961/2011) examines cultural changes in the interrelations between democratic movements, industrial developments, and extended communications, including an analysis of the development of “major cultural institutions, from education to the press, ... and the relation between certain forms in art and the general development of the society” (p. 14). These and other questions on an expanded sense of “culture”—(to include popular culture), social class (with emphasis on the working class), knowledge, power, representation (especially of “the other”), communication (as complex process), community (making of rather than given), the postcolonial, the diasporic, race (especially racialized groups and addressing racism), ethnicity, gender (especially women and feminism), hegemony (and counter hegemony), and social change—constituted British Cultural Studies and were manifest in much of the work by the Center for Contemporary Cultural Studies (CCCS). While cultural studies as an intellectual movement has been historicized and contextualized with multiple origins and focuses beyond the British context (e.g., Wright, 1998), its entanglement with, origin in, and relevance to education (including extramural adult education) requires elaboration. As prominent American cultural studies figure, Lawrence Grossberg (1997) once pointed out, education is “one of the most pressing, promising, and paradoxical sites of cultural studies” (p. 374).

Wright and Maton (2004), Canadian and British scholars respectively, come together to offer insightful ways to recognize that history and reconstitute a future of cultural studies in/and education globally and locally (their collection included essays from all over the world from the UK to Brazil, Belgium to South Africa). Education is one of the foundations of cultural studies – for example, all the recognized founding figures of British cultural studies (including Richard Hoggart, Raymond Williams, E. P. Thompson, and Stuart Hall) started their careers and intellectual projects, in extramural adult education with working class learners. From the 1970s onwards, education was an integral aspect of the work of the Center for Contemporary Cultural Studies at the University of Birmingham (which is widely recognized as the original institutional establishment of cultural studies). From the 1980s to the turn of the century in Canada and the United States, there were movements of commingling between cultural studies and education, characterized by the cultural studies’ institutionalization in postsecondary educational programs, and its intellectual relationship with “critical pedagogy”—movements of critical literacy and anti-oppressive education perhaps most ardently articulated by Paulo Freire in the Brazilian context since the 1960s, with impacts on and in dialogue with critical educators and activists around the world (e.g., Horton and Freire, 1990; Hooks, 1994; Slater et al., 2002). Following the development of cultural studies and education, Wright and Maton (2004) point out the relevance of cultural studies to educational policy, institutionalization in higher education, and new directions in critical education: through “the institutionalization of cultural studies in the national academy; the institutionalization of cultural studies in education; and literacy, reading and cultural pedagogy” (p. 82). In another special issue, Hong Kong scholars Chan and Hui (2011) also demonstrated and aimed to further develop the productive interface between cultural studies and education in various contemporary social contexts (including essays from India, Hong Kong, Taiwan, Britain, and Mainland China). With three themes identified as “education as cultural process”, “rethinking cultural practices for education”, and “the work of educators as cultural mediators” (p. 1), Chan and Hui (2011) highlighted how the interdisciplinarity of cultural studies and localized cultural research have influenced the process of policy planning, educational reforms, and pedagogical practices, as cultural studies itself worked as a critical project mediating the relations “between school education, modernization and the emerging hegemony of the modern nation in the global capitalist context” (p. 5).

While it could be argued that the relationship between cultural studies and education in the British context started as far back as the founding fathers of cultural studies work in extramural education and the education origins of institutional cultural studies, in the North American context, that relationship was articulated somewhat later by critical education theorists such as Henry Giroux’s intervention in the field of cultural studies. Giroux (1992) called for cultural workers (e.g., teachers, intellectuals, cultural critics) to cross borders between critical pedagogy (a critical approach to education) and cultural studies—for example through a hybrid discourse of a “pedagogy of representation”. Giroux’s call to imbricate the worlds of critical pedagogy and cultural studies was well received in some quarters in both fields but also resisted in other quarters in the two fields. For example, some in critical education associated cultural studies with “elitism” and a “claim to the moral high ground of a politics of inclusivity” (Steinberg and Kincheloe, 1997, p. 6). While the institutionalization of cultural studies in higher education has happened, forms of curriculum, pedagogy and evaluation associated with (at least) the rhetoric of cultural studies systematically favor students from the new middle class, and students tend to view cultural studies as an entryway to careers into what they consider the more prestigious world of the cultural industries rather than a means of undertaking serious engagement with critical pedagogy and social justice education in the supposedly more mundane world of education. Understandably, questions have been raised about the revolutionary potential of cultural studies—“if radical it must remain marginal, if effective it must become institutionalized” (Maton and Wright, 2002, p. 385). If we are to consider the presence of cultural studies in the academy, the picture is decidedly mixed. In terms of the disciplinarity, cultural studies has influenced several academic disciplines and achieved legitimacy and even considerably high status itself in higher education, having evolved into something of a (post)discipline itself, even as it has retained its characteristic of being interdisciplinary. On the other hand, in its epistemological and political articulations as a global discourse, cultural studies has not achieved sustained transnationalism and inclusion and adequate representation of marginal schools (Tomaselli and Wright, 2008; Wright, 2016; Wright and Xiao, 2020).

Finally, taking cultural studies seriously in education involves (or should involve) a critical departure from European/Euro-American cultural superiority and Western-centric cultural imperialism, and making space for the emergence and prevalence of alternative polycentric ways of knowing and being in the world. For Kenyan writer and cultural critic [Ngugi \(1986\)](#), language carries the culture, and the “mental universe of the colonized” was dominated in part by the use of colonizers’ language and literature which “resulted in the alienation of a child from the child’s history, geography, music and other aspects of culture” (pp. 36–37). Ngugi holds, therefore, that to write in African languages in the African context is to engage in an anti-colonial struggle. In the context of African epistemologies and Blackness, [Wright and Xiao \(2021\)](#) call for “a parallel process of localization, Africanization, Third Worldization of knowledge” in order to decolonize and replace both the Eurocentric body of knowledge and, more importantly, Eurocentric epistemology (pp. 42–43).

Examining modernity and coloniality in the context of Latin America, Walter [Mignolo \(2007, 2009, 2011\)](#) argues that it is important to delink from colonialist hegemonic conceptions of epistemological order, world order, “modernity”, “underdevelopment”, and “otherness”, which collectively position the so-called “underdeveloped world” as having produced only “culture” but not science nor philosophy ([Mignolo, 2007](#), p. 273). He also advocates that the Third World eschew the Cartesian “worldview” for a more holistic “world sensing” not just through the eyes and the mind but by utilizing all the senses and embodied knowledge. For Taiwanese cultural studies scholar, Kuan-Hsing [Chen \(2010\)](#), examining the changing representations of “Asia” in its own contingent historical, cultural, and geopolitical meanings enables deeper understandings of “Asia as method”, a method that invites the intellectual undoing of Cold War logics, continuation of decolonization, and a move toward global deimperialization. While these ideas have been taken up in various educational contexts (e.g., [Carter-Ényì and Carter-Ényì, 2019](#); [De Lissovoy, 2019](#); [Lin, 2012](#); [Lee, 2019](#); [Martín-Díaz, 2017](#)), decolonization is an incomplete process and demands serious understanding of Indigenous approaches to culture. The next section will then focus on the relationship between culture and Indigeneity, involving culturally rooted education, land education, and an ethics of incommensurability.

Culture and indigeneity: culturally rooted education, land education, and an ethics of incommensurability

In *Decolonizing methodologies*, Maori scholar Linda Tuhiwai [Smith \(2012b\)](#) emphasizes the political nature of knowledge production, and the importance for Indigenous communities—while undoing the penetration and surveillance of Western imperialist gaze—to reclaim an ethical theorizing of origins as Indigenous people with various aspects of linguistic and material culture. Culture, in this vein, is necessarily part of the Indigenizing and decolonizing movements in “struggling to make sense of our own [Indigenous] world while also attempting to transform what counts as important in the world of the powerful” (p. 88).

One movement relates to projects about “culturally rooted education” ([Alfred, 2011](#); [Hare et al., 2011](#)) for Indigenous cultural pride, heritage, and revitalization. In the face of ongoing colonialism that erases Indigenous ontological and epistemological presences, Taiaiake [Alfred \(2011\)](#) from Kahnawá:ke in the Mohawk Nation insists that “being Indigenous and being a warrior are one and the same” (p. 7), and culturally rooted education—as a central strategy of anticolonial struggles for Indigenous futures—means to teach Indigenous youth their histories and languages, to see, feel, and think with Indigenous eyes, hearts, and minds. Through Indigenous-led scholarship, pedagogy, and community initiatives, culturally rooted education is manifested in projects such as visual pedagogy for urban Indigenous youth to expose the legacy of the Indian Residential School System, wholistic educational practices provided by non-profit Indigenous youth organization, culturally relevant assessment model for Indigenous students in college, and Indigenous young women circles of leadership for cultural and linguistic revitalization ([Hare et al., 2011](#)). Further in the direction of reclaiming Indigenous cultural and epistemological power in/for education, [Archibald Q’um Q’um Xieem \(2008, 2019\)](#) proposes and elaborates on a foundation of “Indigenous Storywork”, that Indigenous story making, storytelling, and the cultural meanings of stories are valuable knowledge and serious intellectual work, and that decolonizing education involves ethical, pedagogical, and methodological principles working with Indigenous peoples, knowledges, and stories. Originally based on the multi-year research process with Coast Salish/Stó:lō Elders, cultural knowledge holders, and educators from Indigenous communities in British Columbia, Canada, Indigenous Storywork has been taken up by educators across Canada, Australia, and Aotearoa New Zealand, and has been utilized in diverse disciplines (e.g., mathematics education, literary studies, films aesthetics, geography, gender studies, legal practices, etc.).

Land education emerges as a concurrent and sometimes interrelated movement. The interrelated notions that “Water is life; Land is our first teacher” ([Smith et al., 2019](#), p. 1) invite people to think about the deeply cultured relationships between Indigenous people and the Land, as well as the non-anthropocentric, land-based context of learning. Land education not only values ecological interdependence and land as everyday pedagogy, but also recognizes land as alive and relating to Indigenous stewardships, sovereignties and futurities. It is a critical intervention in and departure from conventional environmental education and place-based education that do not seriously address colonialism and in particular settler colonialism and tend to “skip ahead to some neutralized ahistorical, guilt-free, pain-free, romanticized version of environmental education” ([Tuck et al., 2014](#), p. 14). As an Indigenous philosophical construct, Land (capital “L”) is more than physical geographic space or the diaphanousness of inhabited memories—instead, “Land is spiritual, emotional, and relational; Land is experiential, (re)membered, and storied; Land is consciousness—Land is sentient. Land refers to the ways we honor and respect her as a sentient and conscious being” ([Styres, 2019](#), p. 27). With this understanding, land education opens up educational spaces that respect, recognize, and center Indigenous culture (languages, histories, identities, knowledges, cosmologies, and relationships to the places) in ways that affirm Indigenous ontological and epistemological accounts of the land ([Smith et al., 2019](#)). Noting how Indigenous peoples and cultures are constantly

struggling against colonial violence, erasure and reduction, Alex Wilson from the Opaskwayak Cree Nation uses the “cultural iceberg” analogy to call for a deeper understanding of Indigenous cultures in connection to the land: while many people’s knowledge of Indigenous cultures is restricted to visible and material aspects (i.e., the tip of the iceberg being severed from its bottom), decolonization requires restoring the relationships between visible and material culture and “the deep knowledge, value systems, philosophies, cosmologies and other invisible aspects” of Indigenous cultures—the immense bulk of iceberg which cannot be seen, restored, and recognized without “land-based knowledge” (Wilson and Laing, 2019, p. 136).

“Literacies of Land” (Styres, 2019) is an inspiring idea of how Indigenous cultures and land are intimately interrelated, through landscapes, soundscapes, languages, histories, fauna, flora, spirituality, cosmology, wellness, aesthetics, activities, and stories. Connecting language and land-based representation, Rasmussen and Akulukjuk (2009) insist that when discussing environmental education, the crucial question is, “what language does the environment speak?”: in Nunavut, “the land speaks Inuktitut” (p. 285), as language is not simply developed in isolation in human brains, but “in relationship to land and water” (as quoted in Tuck et al., 2014, p. 19). Within diverse Indigenous communities, language is also deeply and foundationally related to specific places, territories, embodiment, and senses of belonging. This explains, for example, why Swampy Cree member and language speaker Kevin Burton felt dislocated and disembodied in making a Vancouver-based documentary (where he translated the place names in Coast Salish territories of the Musqueam/Tsleil-wau-tuth/Squamish to his Cree language), and how he felt more culturally congruent and integral in making a second documentary based in his Cree home territories, using his Cree language. Beyond language, literacies of land enables a holistic, Indigenous-centered reading of and engagement with the land: for example, the non-academic and informal education common among Diné people and their horses (John, 2019), Torres Strait Islanders and their knowledge connected to coral reefs and totemic sea animals (Whitehouse et al., 2014), and Māori people’s nurturing of the harakeke plantation that links together daily life, medicine, wellness, family relationships, plant conservation, and cosmology (Pihama et al., 2019). In Styres (2019)’s words, “Land as a decolonizing praxis informs pedagogy through storied relationships. These stories are etched into the essence of every rock, tree, animal, pathway, and waterway (whether in urban or rural/natural or built environments) in relation to the Indigenous people who have existed on the land since time immemorial” (p. 29).

In Indigenous and decolonizing education, praxes of “culture” involve both epistemological/pedagogical inclusion and political/pedagogical incommensurability. There are notable inclusive, cross-cultural approaches, such as concept of “Cultural Interface” in the Australian curriculum framework (Nakata, 2011), the idea of “Two-Eyed Seeing” shared by Mi’kmaw Elder Albert Marshall in Canada and utilized in science education, environmental education and medical education (e.g., Bartlett et al., 2012), the “Meeting of Knowledges” projects in Brazil and Columbia (de Carvalho et al., 2016), and the different institutional models of “Intercultural University” in Mexico and Ecuador (Cortés, 2017; Martín-Díaz, 2017). In this direction, one of the forefronts is science education at the intersection of Indigenous cultures, community knowledge, and holistic transdisciplinary learning (e.g., Newberry and Trujillo, 2019). While genuine recognition of Indigenous cultures means a rejection of Western-only or Western-centric sciences, it is not a rejection of sciences. To speak with Cajete (1994, 2018), to understand the foundations of Indigenous science, one must participate with the natural world and “must become open to the roles of sensation, perception, imagination, emotion, symbols and spirit, as well as that of concept, logic and rational empiricism” (Cajete, 2018, p. 16).

In a radical vein, “decolonization is not a metaphor” (Tuck and Yang, 2012) is a statement not so much on inclusion/integration, but really on why decolonization must involve sustained praxes (instead of easy adoption of the decolonizing discourse by educational projects) that will lead to the repatriation of Indigenous land and life, as well as center the cultural and political futurities of Indigenous peoples (instead of settlers). Taking decolonization seriously, Tuck and Yang (2012) forward “an ethic of incommensurability” that recognizes what is distinct and what is sovereign for projects of decolonization, which could be incommensurable with some objectives in social justice, critical methodologies, or approaches that decenter settler perspectives. This has important implication for the meanings and methods of education for decolonization. For example, reflecting on critical theory, critical pedagogy, and Indigenous worldviews/pedagogies, Grande (2010) advocates for “Red Pedagogy” that does not retain core Western assumptions, does not center Marxist theory, and does not retain measures of anthropocentrism. Although anti-capitalism and anti-colonialism are not mutually exclusive, Red Pedagogy centers struggles around the broader construct of colonization, in a way that “shifts the pedagogical goal from the critique and transformation of capitalist social relations of production (i.e., democratization), toward the critique and transformation of colonialist relations of exploitation (i.e., sovereignty)” (Grande, 2010, p. 203). Reviewing discourses of decoloniality and in particular the praxes of decolonization in higher education, Wright and Xiao (2021, p. 31) distinguish between “decolonization light” and “true decolonization”: while the former is a softer, reformist approach that includes Indigenous perspectives and cultures in the environmental studies curriculum and assumes epistemological commensurability, the latter is a more radical approach that centers Indigenous epistemological and ontological accounts of land, and assumes ontological incommensurability between settlers and Indigenous peoples in an era of colonialism, especially when it comes to cosmology and the relationship to the land.

Outlining 25 Indigenous projects involving Indigenous and non-Indigenous peoples around the world, Smith (2012a) shows how decolonizing knowledge, research and education might look like for Indigenous communities, such as storytelling to sustain oral cultures and as a culturally appropriate way of representation; repatriating artifacts, remains and other cultural materials stolen or removed and taken overseas; organizing survivance (in the form of story, popular music, cultural events, etc.) that accentuates the degree to which Indigenous peoples and communities have retained cultural and spiritual values and authenticity in resisting colonialism; revitalizing Indigenous languages, arts and cultural practices through school-based, media-based, and community-based education; changing the practices of mainstream science teaching in schools to counter the history of hostile attitudes toward Indigenous cultures and the ways Indigenous students learn; understanding the beauty of Indigenous knowledge; and sharing knowledge

with appropriate cultural protocols/expectations. With these and various other projects, Smith (2012b) cogently reminds us that the imperatives of Indigenous research, learning and knowledge creation should be clear: “the survival of peoples, cultures and languages; the struggle to become self-determining, the need to take back control of our [Indigenous] destinies” (p. 238).

Conclusion

This essay has pointed to the complex relationship between culture and education, especially in the articulation of critical, social justice oriented forms of and approaches to education. In terms of the disciplines, it has focused on how culture intersects with and operates within critical streams in the sociology and anthropology of education. While sociology and anthropology are the disciplines in which that interrelation is most dominant, the relationship between culture, the cultural and education exists in other disciplinary sub-fields of education as well. For example, psychology of education includes concepts such as “cultural development”, “cultural security”, and “cultural barriers” (e.g., Hviid and Mårtens, 2019); in history of education the view of culture as “a society’s way of life and way of thinking” (Giorgetti et al., 2017, p. 2) is important for some progressive historians, and in geography of education there are new perspectives on the cultures of workplace, technology, training, landscape, and the meanings of ‘progress’ (e.g., Mills and Kraftl, 2016). Beyond individual disciplines, culture plays a strong role in the articulation of interdisciplinary and transdisciplinary work in education and the essay has addressed the relationship between culture and education and the formation of cultural studies of education as a primary example. Much of the work addressed is about how education addresses identity and sociocultural diversity in ways that embrace representation and social justice. The essay went beyond the disciplinarity therefore to address culture and the cultural in the form of multiculturalism and multicultural education. Another example taken up was the ways in which Indigeneity specifically is being asserted in education in the forms of Indigenous and decolonial education (including land education and the ethics and politics of incommensurability).

The essay opened not only with the complexity of conceptualizing culture but also with the deceptively basic idea that culture involves a whole way of life and a holistic approach to communicating and learning and this reconception of culture carries through much of the essay in the ways in which the relationship between culture and education is articulated. Paradoxically complex and basic, emphasis has been, throughout, on culture as critical and on critical education, on approaches to both that assume that neither culture nor education are neutral. In this sense the essay has emphasized that culture and education and the relationship between them are not free from but are in fact sites of engagement of power, politics, ethics, meanings, worldviews. Engaging the relationship between culture and education at the current historical juncture involves a background of or better yet active work on and hope for representation of and justice for the historically colonized, oppressed, marginalized, trivialized and exploited. It is in short, education for equity and social justice-at the local, national and global levels.

References

- Aiken, N.T., 2016. The distributive dimension in transitional justice: reassessing the South African Truth and Reconciliation Commission's ability to advance interracial reconciliation in South Africa. *J. Contemp. Afr. Stud.* 34 (2), 190–202.
- Alfred, T., 2011. Guest editorial. *Can. J. Native Educ.* 34 (1), 7–8.
- Ang, I., 2009. Beyond multiculturalism: a journey to nowhere? *Humanit. Res.* 15 (2), 17–21.
- Apple, M.W., Ball, S.J., Gandin, L.A. (Eds.), 2010. *The Routledge International Handbook of the Sociology of Education*. Routledge, New York.
- Archibald, Q'um Q'um Xiem, J., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. UBC Press, Vancouver.
- Archibald, Q'um Q'um Xiem, J., Lee-Morgan, J.B.J., De Santolo, J., 2019. *Decolonizing Research: Indigenous Storywork as Methodology*. Zed Books, London.
- Banks, J.A., 1992. Multicultural education: approaches, developments, and dimensions. In: Lynch, J., Modgil, C., Modgil, S. (Eds.), *Cultural Diversity and the Schools, Education for Cultural Diversity: Convergence and Divergence*, vol. 1. The Falmer Press, London, pp. 83–94.
- Banks, J.A., 1994. The canon debate, knowledge construction, and multicultural education. *Educ. Res.* 22 (5), 4–14.
- Banks, J.A. (Ed.), 2009. *The Routledge International Companion to Multicultural Education*. Routledge, New York.
- Banks, J.A., 2013. The construction and historical development of multicultural education, 1962–2012. *Theor. Pract.* 52 (Suppl. 1), 73–82.
- Banks, C.A.M., Banks, J.A., 1995. Equity pedagogy: an essential component of multicultural education. *Theor. Pract.* 34, 152–158.
- Bartlett, C., Marshall, M., Marshall, A., 2012. Two eyed seeing and other lessons learned within a co-learning journey of bringing together Indigenous and mainstream knowledges and ways of knowing. *J. Environ. Stud. Sci.* 2 (4), 331–340.
- Bloomfield, V.E., Fisher, M.E. (Eds.), 2016. *LGBTQ Voices in Education: Changing the Culture of Schooling*. Routledge, New York.
- Bourdieu, P., Passeron, J., 1977. *Reproduction in Education, Society and Culture*. Sage, London.
- Cajete, G., 1994. Look to the Mountain: An Ecology of Indigenous Education. Kivaki Press, Skyland.
- Cajete, G., 2018. Native science and sustaining Indigenous communities. In: Nelson, M.K., Shilling, D. (Eds.), *Traditional Ecological Knowledge Learning from Indigenous Practices for Environmental Sustainability*. Cambridge University Press, New York, pp. 15–26.
- Carter, P.L., 2006. Straddling boundaries: identity, culture, and school. *Sociol. Educ.* 79 (October), 304–328.
- Carter-Ényi, Q., Carter-Ényi, A., 2019. Decolonizing the mind through song. *Perform. Res.* 24 (1), 58–65.
- Cha, Y.K., Ham, S., Lee, M. (Eds.), 2018. *Routledge International Handbook of Multicultural Education Research in Asia Pacific*. Routledge, London.
- Chan, S., Hui, P., 2011. Introduction: education and cultural studies. *Cult. Stud.* 25 (1), 1–8.
- Chen, K.H., 2010. *Asia as Method: Toward Deimperialization*. Duke University Press, Durham.
- Chin, R., 2017. *The Crisis of Multiculturalism in Europe: A History*. Princeton University Press, Princeton and Oxford.
- Cortés, L.S.M., 2017. Indigenous youth graduating from intercultural universities: capability building through intercultural higher education in Veracruz, Mexico. *J. Intercult. Stud.* 38 (2), 155–169.
- Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), 1995. *Critical Race Theory: The Key Writings that Formed the Movement*. The New Press, New York.
- de Carvalho, et al., 2016. The Meeting of Knowledges as a contribution to ethnomusicology and music education. *World Music* 5 (1), 111–133.

- De Lissovoy, N., 2019. Decoloniality as inversion: decentring the west in emancipatory theory and pedagogy. *Glob. Soc. Educ.* 17 (4), 419–431.
- Dayhle, D., McCarty, T.L., 2007. Beatrice medicine and the anthropology of education: legacy and vision for critical race/critical language research and praxis. *Anthropol. Educ. Q.* 38 (3), 209–220.
- DiMaggio, P., 1982. Cultural capital and school success: the impact of status culture participation on the grades of U.S. high school students. *Am. Socio. Rev.* 47 (2), 189–201.
- Dumais, S., 2002. Cultural capital, gender, and school success: the role of habitus. *Sociol. Educ.* 75 (1), 44–68.
- Emerick, M.R., Hoffman, B.Y., Kanno, Y., 2020. Teaching Hispanic restaurant workers: translanguaging as culturally sustaining pedagogy. *Anthropol. Educ. Q.* 51 (3), 304–332.
- Gay, G., 2000. *Culturally Responsive Teaching: Theory, Research, and Practice*. Teachers College Press, New York.
- Gay, G., 2010. *Culturally Responsive Teaching: Theory, Research, and Practice*, 2nd ed. Teachers College, New York.
- Giorgetti, F.M., Campbell, C., Arslan, A., 2017. Culture and education: looking back to culture through education. *Paedagog. Hist.* 53 (1–2), 1–6.
- Giroux, H., 1992. *Border Crossings: Cultural Workers and the Politics of Education*. Routledge, New York.
- Goh, D., 2017. Diversity and Nation-Building in Singapore. Global Centre for Pluralism. Retrieved from: https://www.pluralism.ca/wp-content/uploads/2017/10/Singapore_Case_Note_ENG.pdf.
- Grande, S., 2010. Red pedagogy. *Counterpoints* 356, 199–207.
- Grossberg, L., 1997. *Bringing it all back home: essays on cultural studies*. Duke University Press, Durham.
- Hall, S., 1989. The origins of cultural studies (Accessed 18 March 2022).
- Hall, S., 1992. Cultural Studies and Its Theoretical Legacies. In: Grossberg, L., Nelson, C., Treichler, P.A. (Eds.), *Cultural Studies*. Routledge, London, pp. 277–286.
- Hall, S., 2000. Conclusion: the multicultural question. In: Hesse, B. (Ed.), *Un/Settled Multiculturalisms: Diasporas, Entanglements, Transruptions*. Zed Books, London, pp. 209–241.
- Hall, S., 2016. *Cultural Studies 1983: A Theoretical History*. Duke University Press, Durham.
- Hare, J., Archibald, J., Fellner, K., Christian, D., 2011. Editorial: Indigenous youth as the new warriors. *Can. J. Native Educ.* 34 (1), 1–6.
- Horton, M., Freire, P., 1990. *We Make the Road by Walking: Conversations on Education and Social Change*. Templeton University Press, Philadelphia.
- Hviid, P., Mårtens, M., 2019. *Culture in Education and Education in Culture Tensioned Dialogues and Creative Constructions*. Springer, Cham.
- Hodge, B., O'Carroll, J., 2006. *Borderwork in Multicultural Australia*. Allen & Unwin, Sydney.
- Hogg, L., 2011. Funds of knowledge: an investigation of coherence within the literature. *Teach. Educ.* 27 (3), 666–677.
- Hooks, B., 1994a. *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, New York.
- Hooks, B., 1994b. Introduction: the heartbeat of cultural revolution. In: *Outlaw Culture: Resisting Representations*. Routledge, New York, pp. 1–8.
- Huat, C.B., 2003. Multiculturalism in Singapore: an instrument of social control. *Race Class* 44 (3), 58–77.
- Igarashi, H., Saito, H., 2014. Cosmopolitanism as cultural capital: exploring the intersection of globalization, education and stratification. *Cult. Sociol.* 8 (3), 222–239.
- Jaffe-Walter, R., Lee, S.J., 2018. Engaging the transnational lives of immigrant youth in public schooling: toward a culturally sustaining pedagogy for newcomer immigrant youth. *Am. J. Educ.* 124, 257–283.
- John, K.D., 2019. Rez ponies and confronting sacred junctures in decolonizing and Indigenous education. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, London, pp. 50–61.
- Kaomea, J., et al., 2019. Reclaiming, sustaining and revitalizing Hawaiian education through video-cued Makawalu ethnography. *Anthropol. Educ. Q.* 50 (3), 270–290.
- Keong, C.S., 2013. Multiculturalism in Singapore. *Singapore Acad. Law J.* 25 (1), 84–109.
- Koh, A., Maxwell, C., Engel, L., Yemini, M., 2020. Critiquing dominant cultures of schooling as a sustained critical enterprise in sociology of education. *Int. Stud. Sociol. Educ.* 29 (3), 199–203.
- Kymlicka, W., 2012. *Multiculturalism: Success, Failure, and the Future*. Transatlantic Council on Migration, Migration Policy Institute, Washington, DC.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lareau, A., Horvat, E.M., 1999. Moments of social inclusion and exclusion: race, class, and cultural capital in family-school relationships. *Sociol. Educ.* 72 (1), 37–53.
- Lee, C.D., 2017. An ecological framework for enacting culturally sustaining pedagogy. In: Paris, D., Alim, H.S. (Eds.), *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York, pp. 261–273.
- Lee, Y., 2019. A critical dialogue with 'Asia as method': a response from Korean education. *Educ. Philos. Theor.* 51 (9), 958–969.
- Lee, S.J., et al., 2020. Asians for Black lives, not Asians for Asians': building Southeast Asian American and Black solidarity. *Anthropol. Educ. Q.* 51 (4), 405–421.
- Levinson, B., Pollock, M., 2011. A companion to the anthropology of education. Wiley-Blackwell, Oxford.
- Lin, A., 2012. Towards transformation of knowledge and subjectivity in curriculum inquiry: insights from Chen Kuan-Hsing's "Asia as Method". *Curric. Inq.* 42 (1), 153–178.
- Maton, K., Wright, H.K., 2002. Returning cultural studies to education. *Int. J. Cult. Stud.* 5 (4), 379–392.
- May, S., Sleeter, C. (Eds.), 2010. *Critical Multiculturalism: Theory and Praxis*. Routledge, London.
- Martin-Diaz, E., 2017. Are universities ready for interculturality? The case of the intercultural University 'Amawtay Wasi' (Ecuador). *J. Lat. Am. Cult. Stud.* 26 (1), 73–90.
- McCarty, T., Lee, T., 2014. Critical culturally sustaining/revitalizing pedagogy and Indigenous education sovereignty. *Harv. Educ. Rev.* 84 (1), 101–124.
- McDougall, J., et al., 2004. High school-aged youths' attitudes toward their peers with disabilities: the role of school and student interpersonal factors. *Int. J. Disabil. Dev. Educ.* 51 (3), 287–313.
- Medicine, B., Jacobs, S. (Eds.), 2001. *Learning to Be an Anthropologist and Remaining "Native": Selected Writings*. University of Illinois Press, Urbana.
- Mignolo, W.D., 2007. Introduction: coloniality of power and decolonial thinking. *Cult. Stud.* 21 (2–3), 155–167.
- Mignolo, W.D., 2009. Epistemic disobedience, independent thought and decolonial freedom. *Theory Cult. Soc.* 26 (7–8), 159–181.
- Mignolo, W.D., 2011. Geopolitics of sensing and knowing: on (de)coloniality, border thinking and epistemic disobedience. *Postcolonial Stud.* 14 (3), 273–283.
- Mills, S., Kraft, P., 2016. Cultural geographies of education. *Cult. Geogr.* 23 (1), 19–27.
- Modood, T., 2013. *Multiculturalism: A Civic Idea*, second ed. Polity, Cambridge.
- Moll, L.C., 2005. Reflection and possibilities. In: Gonzalez, N., Moll, L.C., Amanti, C. (Eds.), *Funds of Knowledge: Theorizing Practices in Households, Communities and Classrooms*. Lawrence Erlbaum, New Jersey, pp. 275–287.
- Moll, L.C., Velez-Ibanez, C., Greenberg, J., Rivera, C., Andrade, R., Dworin, J., et al., 1990. *Community Knowledge and Classroom Practice: Combining Resources for Literacy Instruction: Technical Report (OBEMLA Contract No. 300-87-0131)*. University of Arizona College of Education and Bureau of Applied Research in Anthropology, Tucson, AZ.
- Mudimbe, V.Y., 1988. *The Invention of Africa: Gnosis, Philosophy and the Order of Knowledge*. Indiana University Press, Bloomington.
- Nakata, M., 2011. Pathways for Indigenous education in the Australian curriculum framework. *Indigenous Educ.* 40, 1–8.
- Newberry, T., Trujillo, O.V., 2019. Decolonizing education through transdisciplinary approach to climate change education. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, London, pp. 131–145.
- Ngugi, W.T., 1986. *Decolonising the Mind: The Politics of Language in African Literature*. James Currey, London.
- Niesz, T., 2001. Using "culture" in anthropology and education. *Anthropol. News* 46–47.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Paris, D., Alim, H.S., 2014. What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harv. Educ. Rev.* 84 (1), 85–100.
- Parliament of Australia, 2010. *Multiculturalism: A Review of Australian Policy Statements and Recent Debates in Australia and Overseas*. https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1011/rp101111p06.
- Perrotta, C., Bailey, C., Garside, C., 2018. Culture, technology and local networks: towards a sociology of 'making' in education. *Camb. J. Educ.* 48 (5), 553–569.

- Phipps, A., Young, I., 2015. Neoliberalisation and 'lad cultures' in higher education. *Sociology* 49 (2), 305–322.
- Pihama, L., Campbell, D., Greensill, H., 2019. In: Archibald, J., Lee-Morgan, J.B.J., De Santolo, J. (Eds.), *Decolonizing Research: Indigenous Storywork as Methodology*. Zed Books, London, pp. 137–150.
- Raedt, K., 2012. Building the Rainbow Nation: a critical analysis of the role of architecture in materializing a post-apartheid South African identity. *Afr. Focus* 25 (1), 7–27.
- Rasmussen, D., Akulukjuk, T., 2009. My father was told to talk to the environment first before anything else": arctic environmental education in the language of the land. In: McKenzie, M., Hart, P., Bai, H., Jickling, B. (Eds.), *Fields of Green: Restorying Culture, Environment, and Education*. Hampton Press, Cresskill, pp. 285–298.
- Richards, B.N., 2020. When class is colorblind: a race-conscious model for cultural capital research in education. *Sociol. Compass* 14 (7), e12789.
- Sheng, X., 2012. Cultural capital and gender differences in parental involvement in children's schooling and higher education choice in China. *Gen. Educ.* 24 (2), 131–146.
- Slater, J.J., Fain, S.M., Rossatto, C.A., 2002. *The Freirean Legacy: Educating for Social Justice*. Peter Lang, New York.
- Sleeter, C., 1996. *Multicultural Education as Social Activism*. SUNY Press, Albany.
- Sleeter, C., 2009. Developing teacher epistemological sophistication about multicultural curriculum: a case study. *Action Teach. Educ.* 31 (1), 3–13.
- Sleeter, C., 2013. Teaching for social justice in multicultural classrooms. *Multicult. Educ. Rev.* 5 (2), 1–19.
- Smith, L.T., Smith, S.H., 2012a. *Decolonizing Methodologies: Research and Indigenous Peoples*, second ed. Zed Books, London and New York.
- Smith, L.T., Smith, S.H., 2012b. Reading sociology into scholarship on school-based sex education: structure and culture. *Sociol. Compass* 6 (8), 614–623.
- Smith, L.T., Tuck, E., Yang, K.W., 2019. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, London.
- Spivak, G.C., 2012. *An Aesthetic Education in the Era of Globalization*. Harvard University Press, Cambridge, MA.
- Steinberg, S., Kincheloe, J. (Eds.), 1997. *KinderCulture: The Corporate Construction of Childhood*. Westview Press, Boulder.
- Styres, S., 2019. Literacies of land: decolonizing narratives, storying, and literature. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, London, pp. 24–37.
- Taylor, C., 2012. Interculturalism or multiculturalism? *Philos. Soc. Critic.* 38 (4–5), 413–423.
- Teeger, C., 2015. Ruptures in the Rainbow Nation: how desegregated South African schools deal with interpersonal and structural racism. *Sociol. Educ.* 88 (3), 226–243.
- Toh, G., 2020. Schooling (for) Japanese children in cosmopolitan Singapore: building bridges and erecting barriers. *Int. Stud. Sociol. Educ.* 29 (3), 274–292.
- Tomaselli, K., Wright, H.K., 2008. Editorial statement: African cultural studies. *Cult. Stud.* 22 (2), 173–186.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Tuck, E., McKenzie, M., McCoy, K., 2014. Land education: Indigenous, post-colonial, and decolonizing perspectives on place and environmental education research. *Environ. Educ. Res.* 20 (1), 1–23.
- Vahed, G., Desai, A., 2017. Stuck in the middle? Indians in South Africa's fading rainbow. *South Asian Diaspora* 9 (2), 147–162.
- Villenas, S.A., 2019. The anthropology of education and contributions to critical race studies. *Equity Excell. Educ.* 52 (1), 68–74.
- Wallace, D., 2017. Reading 'race' in Bourdieu? Examining Black cultural capital among Black Caribbean youth in south London. *Sociology* 51 (5), 907–923.
- Waters, J.L., 2006. Geographies of cultural capital: education, international migration and family strategies between Hong Kong and Canada. *Trans. Inst. Br. Geogr.* 31 (2), 179–192.
- Whitehouse, H., et al., 2014. Sea Country: navigating Indigenous and colonial ontologies in Australian environmental education. *Environ. Educ. Res.* 20 (1), 56–69.
- Williams, R., 1958/1989. Culture is ordinary. In: *Resources of Hope: Culture, Democracy, Socialism*. Verso, London, pp. 3–14.
- Williams, R., 1961/2011. *The Long Revolution*. Parthian, Cardigan.
- Williams, R., 1983. *Keywords: A Vocabulary of Culture and Society*. Oxford University Press, New York.
- Wilson, A., Laing, M., 2019. Queering indigenous education. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, London, pp. 131–145.
- Wright, H.K., 1998. Dare we de-centre Birmingham? troubling the 'origin' and trajectories of cultural studies. *Eur. J. Cult. Stud.* 1 (1), 33–56.
- Wright, H.K., 2012. Between global demise and national complacent hegemony: multiculturalism and multicultural education in a moment of danger. In: Wright, H.K., Singh, M., Race, R. (Eds.), *Precarious International Multicultural Education: Hegemony, Dissent and Rising Alternatives*. Sense Publishers, Rotterdam/Boston/Taipei, pp. 103–113.
- Wright, H.K., 2016. What has African cultural studies done for you lately? autobiographical and global considerations of a floating signifier. *Crit. Arts* 30 (4), 478–494.
- Wright, H.K., Maton, K., 2004. Cultural studies and education: from Birmingham origin to glocal presence. *Rev. Educ. Pedagog. Cult. Stud.* 26 (2–3), 73–89.
- Wright, H.K., Xiao, Y., 2020. African cultural studies: an overview. *Crit. Arts* 34 (4), 1–31.
- Wright, H.K., Xiao, Y., 2021. Decolonisation and higher education: theory, politics and global praxis. *Postcolonial Direct. Educ.* 10 (1), 23–50.
- Wright, H.K., Singh, M., Race, R. (Eds.), 2012. *Precarious International Multicultural Education: Hegemony, Dissent and Rising Alternatives*. Sense Publishers, Rotterdam/Boston/Taipei.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Zhang, S., Xu, C.L., 2020. The making of transnational distinction: an embodied cultural capital perspective on Chinese women students' mobility. *Br. J. Sociol. Educ.* 41 (8), 1251–1267.

Decolonial and education

Aparna Mishra Tarc, York University, Toronto, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	31
Decolonial and education	32
Toward decoloniality	34
Mobilizing the decolonial	34
Conclusion	36
References	36
Further reading	37
Relevant websites	37

Glossary

Anti-colonial is the precursor to all decolonizing movements. Anti-colonial evokes political social movements resisting imperialism. Anticolonial movements in former colonized nations worked to overthrow the direct rule of European powers over much of Africa, Asia, and South America.

Colonialism is an (en)forced practice of domination, which involves the subjugation of one people to another.

Colonialization is a state-initiated process by which a political power from one territory exerts control over a different territory. It involves unequal power relations and includes policies and/or practices of acquiring full or partial political control over other people or territory, occupying the territory with settlers, and exploiting it economically and socio-culturally.

Colonial Education is a key mechanism of colonialism that inculcates the children of colonizers in the epistemological institutions housing and circulating systems of power and knowledge of the colonizer.

Colonial violence is the forceful and systemic mental, psychological and physical means of subjugation of one people over another.

Decolonial is a movement to dismantle structures and institutions of oppression garnering the ongoing subjugation colonized peoples.

Decoloniality refers to the departure from the logic, metaphysics, ontology and matrix of power created by the massive processes and aftermath of colonization and settler-colonialism.

Decolonization is long-term process involving the bureaucratic, cultural, linguistic and psychological divesting of colonial power. The term is currently used in the context of Indigenous struggles of self-determination in settler colonial nations.

Epistemicide refers to the systemic silencing, elimination and destruction of ontological, religious, cultural and land-based knowledges of colonized peoples.

Emancipation is the act of freeing oneself or one's people of political, social, legal oppression (i.e., slavery).

Liberation is the act of freeing oneself or one's people of colonial oppression.

Settler colonialism is distinct from the global colonial projects of Empire in that it refers to the specific take-over of Indigenous lands for the purpose of replacing these populations with colonial subjects.

Speculative Future(s) speculate on silenced, erased and subjugated histories to imagine another future world of consentience and cohabitation.

Traumatic colonial history refers to the present-day transgenerational effects of those witnessing the violence, mass degradation, killing, and destruction wrought by previous colonial powers on colonized people.

Introduction

The history of anti-colonial, decolonial, and post-colonial movements arise with liberatory movements in the colonies during and after the fall of European colonialism. Anti-colonial struggles in the former colonized nations reached their height after the Second World War. Colonial occupation and conquest were enacted across continents in the global South and in the Americas. In other words, across the world, wherever European powers invade, occupy, and settle lands that were not originally their own, calls for purging the power and effects of colonization persist. Anticolonialism, like other movements, refers simultaneously to the historical political movement, critical apparatus and schools of thought and action generated from the collective revolt, struggle, and

reclamation of land and way of life of colonized people. In the contemporary moment, struggles continue to eradicate colonial rule and influence in Global South and settler colonial nations.

Anti-colonialism forms the basis for decolonial and postcolonial intellectual movements. Anticolonial critique predates contemporary theories of globalization that present it as a new development of history. Despite its manifestation in local contexts, colonialism has always been a form of globalization requiring the flows of ideas, people and capital to operate. With contemporary globalization, local calls and struggles to decolonize human systems, cultures and relations consolidate into a global movement, particular in a time of pandemic, war and ongoing environmental catastrophe.

As a political project, anti-colonial movements were unevenly successful in ejecting the presence of colonial rulers from their territory. As with decolonial processes, anti-colonial movements are particular to geographic region. These nations continue to mitigate the after-effects of colonial modernization institutionally, legally, educationally and socio-culturally. However, these mitigations are tempered by the necessity or compulsion to compete in the global economy (Soudien and Cornelise, 2000). Moreover, often colluding with colonial power, national elites of the “de-colonized” states continue to uphold colonial institutions of education. Although anti-colonial struggles continue in some parts of the world, for the most part, many scholars and activists has turned to decolonial strategies.

Education is both a mechanism of colonization and liberation for colonized peoples. On the one hand, policies such as residential schooling, assimilation and eradication of native epistemologies continue to hold and have ongoing and deep consequences for selves and societies subject to the educational dehumanization of colonial systems. The supremacy of the white and European cultures continues to be circulated and reinforced through English as the lingua franc and language of capital, ideas, culture, and communication. On the other hand, colonized peoples have recognized the importance of enacting an education that dismantles and puts forth Indigenous and other visions by which to learn, live and be. The recent reclamation of language and educational practices (i.e., powwow) for Indigenous peoples is one such educational horizon that might bring about new forms of humanizing education.

This entry will discuss some of the strategies that proponents of decolonial frameworks advance to combat the violent and powerful cognitive, linguistic and cultural effects of colonial education European colonizers, and particularly, the British Empire. The global “Commonwealth” policy was particularly successful in using education as instrument of knowledge/power to secure the assimilation of first colonial and now world subjects. Decolonial approaches to global British/European metaphysical Empire grounding modern education proves to be the primary site of inquiry, retrieval, speculation, and emancipation for those calling for alternative futures through education in formerly colonized nations (Asher, 2009a,b).

Decolonial and education

The worldwide movements to decolonize nations colonized by European powers were initiated by formerly colonized peoples in Asian and Africa post World War II. These movements were liberatory, seeking to reclaim nation states by, at times, violently taking back political power from the colonial elite and proxies. These liberation movements failed as their leaders did not consider the long-term and lasting effects of soft state apparatus: law, language, education, and culture invading the populations at large. They somewhat naively believed, as Scott (2004) remarks, that a change in power would automatically lead to a change in society. However, a century of colonial rule instituting religious and modern epistemologies in the hearts and minds of colonized subjects would require more than a replacement of heads of states of the modern nation. Or, as Nugugi (2018) insists: “History moves on, theories of liberation march alongside it but without our languages, we will remain trapped in, what literary critic Adam Beach calls the English metaphysical empire” (np).

Education, through language and literacy, is a key mechanism of colonization, particularly for the British Empire. Educational colonization enacts an insidious and onto-epistemological form of rule over the mind and bodies of its subjects. The effects of colonial education are long lasting as they garner assimilation of native populations into the ways of life of the colonizer such that the colonial subject no longer recognizes what Franz Fanon (1961/2004) conceives as “his true nature” (p. 198).

As John Willinsky (1999) documents in *Education After Empire*, winning over the hearts and minds of colonial subjects through cognitive, affective, linguistic, religious and cultural force was a formative pedagogical instrument of the British Empire. Education formed a significant aspect of colonial policy in India, as documented in Maccauley’s minutes of 1816 where the policymaker espoused the “intrinsic superiority of Western literature” over that of ancient Sanskrit and sacred texts foundational to a centuries long civilization in the South Asian subcontinent.

Maccauley’s (1835) imposition of the “English language and literature” into the education of Indians was designed intentionally to assimilate the minds and habits of Indians. “We must,” he writes,

at present do our best to form a class who may be interpreters between us and the millions whom we govern, –a class of persons Indian in blood and colour, but English in tastes, in opinions, in morals and in intellect. To that class we may leave it to refine the vernacular dialects of the country, to enrich those dialects with terms of science borrowed from the Western nomenclature, and to render them by degrees fit vehicles for conveying knowledge to the great mass of the population.

A version of this class of interpreters, what are now teachers and educators, converting native populations into “Englishmen” was replicated with peoples in every Commonwealth outpost in the world, including Indigenous peoples of the Settler Americas. Supplementing the conversion of colonized peoples, was the enslavement and indenturement of the poor and dispossessed for capital gain. With the rise of the American Empire and its media, cultural and commercial technologies of education, English and racist capitalism was exported across the world such that Anglo/American culture has instituted itself as superior to all others as McCauley terribly designed. The effect of this formation of an elite class of English subjects who gain privilege from an English education and its privileging of white cultures (Adichie, 2009) set the course for the triumph of the superiority of the West, whiteness, and English-Medium systems of education across the world.

At its core, decolonial education addresses the dismantling of colonial metaphysical and cultural systems imposed on and internalized by a peoples through instruments of power. It does this in two ways: first, decolonial education seeks to counter the dehumanizing logics of subjectification used to dismiss and silence the knowledge and way of life of colonized people. Second, decolonial education seeks to recover these subjugated and erased knowledges and ways of life as educative in of itself. Decolonial education is an intentional project that aims to emancipate colonial subjects by liberating schooling institutions used as apparatuses of colonial power. In every formerly colonized nation, there are ongoing efforts to decolonize schools, universities and other public educational sites. Decolonial interventions cross a range of institutional, curricular and socio-emotional registers of modern public and higher education.

Decolonial struggles manifest differently in localized nation-state contexts subject to myriad and shifting forms of imperialist, settler and empire colonialism. For countries colonized by the British Empire, decolonizing is especially a double-bind: the linguistic and cultural tools and English education affords its subjects are now highly coveted capacities and knowledges across the world and, yet, one’s assent to an English education seals one’s implicatedness in ongoing structures of oppression. Influential children, who often have no say in the matter of their schooling, are implicated in colonial, and now capitalist and commodifying systems through compulsory public English medium schooling (Coetzee, 2021).

Decolonial education is put forth by many scholars as a means of mitigating the colonial-capitalist foundations underpinning modern education. Approaches by educational scholars and researchers are largely curricular, aimed at countering canonical and representational knowledge as a means to decolonize education. For example, in *Decolonizing Education*, Marie Battiste (2019) calls for the revitalization of Indigenous principles and worldviews of knowledge in educational classrooms.

In her book *Red Pedagogy*, Sandy Grande (2015) looks to Indigenous and decolonizing methods to counter and purge colonial imposition on Indigenous knowledge and way of life. In the historical context of Anti-Black racism, George Dei (2006) has called for a retheorization of “blackness” through indigeneity to interrupt categorical understandings of curriculum grounded in racialization. Nina Asher (2009a,b) looks to rethinking one’s selfhood and homeland as a means to delink self from one’s Western education. And Bettina Love (2019) calls for abolition as an epistemological framework that can bring about educational freedom for Black people and their descendants in the Americas colonized through enslavement.

Recently, the term decolonial is reclaimed by Indigenous scholars as politics rather than as epistemological intervention. Seeing that movements in mainstream educational institutions have failed, Tuck and Yang (2012) sought to reinvigorate the term by calling for what anti-colonial scholars did a century before: demanding the repatriation of land to Indigenous peoples. Tuck and Yang argue that “decolonization is not a metaphor” claiming: “When metaphor invades decolonization, it kills the very possibility of decolonization; it recenters whiteness, it resettles theory, it extends innocence to the settler, it entertains a settler future” (p. 3). Tuck and Yang rally those who use the term loosely to reserve and reclaim it for Indigenous struggles to reclaim land appropriated by settlers.

In “Slavery is a Metaphor,” Garba and Sorentino (2020) argue that Tuck and Yang’s (2012) bid to reclaim and resituate decolonizing in the native-settler colonial nexus occludes the existence of the historical, material and political ongoing realities of the descendants of enslavement and racialized people subject to anti-Black racism. Garba and Sorentino (2020) position highly politicized identity-based arguments as dangerously a-historical and ironically a-political.

By centralizing political ontology precisely as the problem of the relationship between history and contingency, materiality and the symbolic, we attempt to account for the genesis and structure of modernity, and propose slavery as the code that functions as “the a priori or ground of the history to which it gives rise.

(p. 11).

Contra to Tuck and Wang, Garba and Sorentino suggest the uncertainty of metaphor carried by decolonial movements opens to a way forward for an education that liberates. They lean on Mignolo’s (2007) argument of delinking colonial epistemology from epistemology proper as educational in of itself. Eschewing decolonizing as an enactment of the reclamation of land and undoing of settler innocence, Garba and Sorentino call for engaging the failures of decolonizing movements to expunge from epistemologies the logics of Anti-Black and other racisms. They suggest a rethinking of the metaphoric and material conditions of enslavement in relation to the ongoing effects of colonialism founded in the violent capture of racialized peoples elided with colonial notions of property as land.

Toward decoloniality

As Garba and Sorentino (2020) predict, decolonializing has become hyper-metaphoric in educational and mainstream North American culture. Without dismantling or abolishing the metaphysical and Anti-Black moorings of neoliberal capitalisms, decolonializing movements like the anticolonial ones before them, are stuck in a war over positions and representation (Gilroy, 2014; Mishra Tarc, 2015). Decolonizing movements continue to fail to address the ongoing racist and material effects of colonization and harness the imaginative power of rethinking an education that frees rather than subjugates. Rather than advancing a radical education, the term decolonizing is disturbingly taken up in many neoliberal and managerial campaigns as a rhetorical substitute for social change and justice. Equity, Diversity, Inclusion and Decolonizing (EDID) mandates, warriors, and slogans abound in educational and corporate institutions. Ironically making decolonizing a literal politics of land reclamation that is cut off from education, has made what was once an epistemological and educational emancipatory project, easily accommodated and mobilized by capitalist politicians, institutional leaders and neoliberal PR firms.

As calls for EDID as decolonization mount in pluralistic nations, the histories of struggle underpinning specific decolonial movements are emptied of meaning. To recover the decolonial is to return to an understanding of the purpose of education itself which is onto-epistemologically colonial in Western contexts. The educational end to make children particularly human has a long history of signification that requires rethinking education in terms of what Franz Fanon positioned as rereading the human toward the pedagogical praxes of generating new socialities. A hermeneutics of the human requires a reading and rereading of the self under formation as caught in colonial and racist histories. Pedagogy refers to a sociogenic analysis and intervention into colonial subjectification as an ongoing praxis of undoing colonial structures. For Fanon, to decolonize was to begin the process of creating a new species of men (sic) and forms of life in the quest for being human in the ways each are all uniquely meant to be. This struggle for humanizing and the humanity of us is political, onto-epistemological, social and pedagogical, in other words, educational.

Following Fanon, Achille Mbembe (2020) suggests we might view decolonization not as politics or project, but “as an act of refusal turned into an act of self-assertation, an act of rebellion turned into an act of refoundation,” one “imagined as a kind of relation to the future” (p. 11). The future here is not teleological or projected on to unborn children as per our modern understandings. Rather this horizon of the possible aligns to the category of the future as advanced by Indigenous and Black scholars. The future as human potentialities are “another name for the force of self-creation and invention” (p. 17) in search of new modes of beyond human existence. As Mishra Tarc (2015) argues in *Literacy of the Other*, this relational pedagogical force of self and societal determination, in the first place, grounds the impulse for resistance and revolt of the colonizing impositions of the other. It is this specter of another future that sustains movements to decolonize the mind and societies from the violences of Eurocentric modernity, racism and other colonial systems.

The decolonial in education has yet to find full form in the way first envisioned by Fanon (1961/2004): a collective fight as a species to reclaim and remake our humanity from the forces of domination and oppression that continue to subjugate and confine one’s mind and our human potential to the violence of, supremacist and racist metaphysics of being. In many ways, Fanon’s (1961/2004) plea to colonized subjects to understand “The European game has finally ended” and to “find something different” toward the intentional and profoundly pedagogical struggle to “set a new history of the human” (p. 311) has yet to be initiated. Scholars of education and the academy prefer to wage territorial and turf knowledge wars rather than turn their critical view to the world wracked with disasters produced by the delusional supremacist logics of Empire itself (this “preference” of course is itself structured by the pedigree and class positioning of university scholars). In the moment of the pandemic, world wars and climate catastrophe perhaps we are more directly witnessing the undeniable consequences of maintaining Eurocentric modernity to the point that our humanity may not recover and find its way back to what Fanon (1961/2004) calls its Originary native being and citizenship.

Mobilizing the decolonial

Acknowledging historical failures, wrong turns, and that “decolonization is a messy, uneven process”, Andreotti et al. (2015) look to decoloniality and its tools as pedagogical interventions to the established “vocabularies” and vernaculars of modernity (p. 22). Following Black and Indigenous critical educational scholars engaged in the futures of the decolonial, Andreotti et al. call for speculative approaches grounded in specified histories of colonial oppressions. They find the constructing of social cartographies that can present numerous ways of attending to “the violences of modernity” that might, in turn, lead to a “mapping of different visions of decolonizing in higher education” (p. 22). Andreotti et al. emphasize that the pedagogical provides the possibilities of non-normative or interruptive knowledges, ideas, encounters and experiences unlike projects of education that follow the colonial and instrumental impulse to assimilate, habituate or form subjects into particular ways of thinking and being.

Ngugi’s (1981) call to “decolonize the mind” is a precursor to the decoloniality movements advanced by Walter Dignolo and others in the South. Decoloniality is primarily an educational means by which to undo colonial structures deeply embedded in the social-political, institutional, and cultural organization of formerly colonial people. According to Dignolo (2017):

Decoloniality means first to delink (to detach) from that overall structure of knowledge in order to engage in an epistemic reconstitution. Reconstitution of what? Of ways of thinking, languages, ways of life and being in the world that the rhetoric of modernity disavowed and the logic of coloniality implement (np).

For Mignolo, decoloniality does not reject modern knowledge as these disciplines provide critical means by which examine colonialization, settler colonialism and its intersections with white cultural superiority and racial capitalism. Mignolo (2011) instead calls for sustained and critical engagement of the procedures, discourses, culture and other colonial mechanism continuing to steer hierarchies of economic being across the globe today. As he indicates:

although knowledge-making is a common human endeavour ... the racialization of places and people in the formation and transformation of the colonial matrix of power not only established hierarchical ranking between languages and categories of thought, but also built economic and political structures of domination and oppression based on the geopolitical and hierarchical organization of knowledge (p. 141).

To truly embark on decolonizing colonial structures from modern institutions of knowledge and education, Mignolo recommends that we de-link from universalizing Eurocentric epistemologies, to engage other knowledge, perspectives and cosmologies as a “pluriverse” (see also Grosfoguel, 2007).

Decolonial scholars and researchers examine the modes of epistemicide colonial knowledges enact on their subjects through the erasure of linguistic and cultural knowledge systems to lift up those systems as key to the revitalization of a de-linked knowledge (Mignolo, 2007; Grosfoguel, 2007). Or as Mignolo (2011) insists: “The decolonial option does not simply protest the contents of imperial coloniality; it demands a delinking of oneself from the knowledge systems we take for granted (and can profit from) and practicing epistemic disobedience” (p. 122). Along with encouraging disobedience, myriad approaches to the decolonial reject monolingualism and culturalisms and the epistemological and national methodologism(s) tied to territorial modernity and or specified supremacist logics of English and/or European metaphysics. Rather, the decolonial disrupts and destabilizes that link to insert colonized and erased ways of life for modern potential and possibilities of reviving old-new worlds eradicated by Empire.

As with Mignolo, Sylvia Wynter (in Snaza and Mishra Tarc, 2019) turns to education and study to call for decoloniality as the key mechanism by which the colonial structures of knowledge/power can be exposed, unraveled and weakened. She argues that racialization founding European colonization remains its most pernicious effect, one that requires a systemic dismantling from the concepts, institutions, and practices upholding mythologies and ideas of the white man as the human par excellence. Wynter insists that categorical norms and names for the human along registers of phenotype, sex, and ability as contributing dehumanize those who identify and are visibly identified by hierarchical markers of social difference. Wynter (2003) calls for a direct and studious confrontation with the histories of colonialism and their vicious after-effects through sustained critique of the ever morphing racist-capital systems of knowledge-power keeping white, male supremacy intact.

Challenges abound for Indigenous and Black scholars of education engaging a reclamation of eradicated knowledge through decolonial frameworks. Unlike formerly colonized subjects residing in nations with long oral, symbolic, and documented traditions of languages and educational systems, these systems were near-terminated and finitely eradicated for Indigenous and Black peoples in Settler colonies. These peoples are put upon to excavate reclaim, revive language and cultures expunged from their cultural memories through cultural and linguistic epistemicide (Santos, 2014).

Another challenge that continues to negatively impact Indigenous and other formerly colonized communities, Dwayne Donald (2021) argues, the legitimizing of objective, documentary, and scientific methods to determine the worth of knowledge. For Indigenous communities where words do not identify and mark but rather embody philosophical ideas and knowledge, Western knowledge continues to institute diminishment and erasure of relational thought-being systems Indigenous people rely on to understand and live in the world. This thought violence to entire communities of Indigenous people goes unrecognized in modern systems privileging instrumentalization and valuation of knowledge. The eradication of the onto-epistemological or relational knowledge has brought about an ethnic cleansing of Indigenous and other worldviews and knowledge systems (Santos, 2014), and also the very means by which Indigenous encounter others, beyond human others, and the world.

An ongoing challenge faced by Indigenous and Black schools is an over-dependence on the archives, methods, and products of Western modernity to excavate and reconstruct their eradicated experiences and knowledges. In his work on Black speculation and futurity, cultural critic Mark Fisher (2014) opts for epistemologies not tied to identities and institutions, or nations and histories, and turns to the ephemera of speculation and hauntology as frameworks for critical and scholarly enquiry. For Fisher the only way to decolonize the future is to look for its appearance in histories of the past. Embracing a similar politics of the future, Ferreira Da Silva (2014) turns to art and aesthetics to speculate the possibility of conceiving knowledge as “poethics”, a mode of epistemology that resists and exposes the Western epistemological default of positioning the racialized subject as other rather than tethered to the colonial one.

Alongside these “future-looking” and “disobedient” gestures to decolonize are student grassroots movements. In March 2015 students at University of Cape Town sparked the #RhodesMustFall movement after they organized to take down a statue of apartheid proponent, white supremacist and imperialist Cecil Rhodes. Soon after in June 2017 educational authorities agreed to a student request to take another statue of Rhodes down at Orion Campus at Oxford University in the UK (Lewis, 2019). On May 2011, students at Ryerson University in Toronto penned a letter to university officials stating they would refer to the university as X after refusing to belong to an institution named after the colonialist proponent of Indigenous residential schooling (Toronto Star, 2021). Through these revolts against the symbols of a colonialism not yet past, decolonializing becomes part of a larger futurist movement of public pedagogy, a living collective form of education through grassroots calls for redress.

A myriad of scholars and students engage new frameworks of the future to chart decolonial pathways that do not replicate Western systems and traditions of knowledge that they argue continue to do harm to the earth and its inhabitants. In a collaborative piece Andreotti et al. (2016) call for a new charting of cartography through sociality rather than territoriality as way forward. Stein et al. (2020) similarly embark on a study of decolonial futures through a myriad of speculative and post-human frameworks, as means of generating thinking that exceeds the man-centric systems of knowledge grounding Western colonial knowledge systems. The numerous un-named student groups calling to decolonize the universities are also leading the way in scholarly communities. These new intentional and ground-up collectivities call for a future rid of colonial logics past to advance the decolonializing impulse as the key movement of thought and of thinking the possibilities of “another world” for a scorched and reeling earth and all its inhabitants (Roy, 2003).

Conclusion

The decolonial in education is at a crossroads between reckoning with the effects of the traumatic colonial past and charting new forms of “beyond human” educational modes, forms, forums, and imaginaries. As the political, social and existential effects of colonial powers come to roost in this volatile and precarious historical moment, we are seeing increased need and calls for collective thinking and action between educational scholars, students and activists. These collectives coalesce around the need for an education to dismantle the habits of mind, the violent logics of property and the commodity and racial identification conceived and enacted by European empire.

Clearly, in education there is a need to move the decolonial beyond localized politics and projects to a larger, relational world-view such as those advanced by Fanon, Mignolo and Sylvia Wynters. Many educational scholars such as Snaza and Mishra Tarc (2019), Love (2019), and Andreotti et al. (2016) recognize the need to ground the decolonial in its historical antecedents and deploy it through speculative approaches oriented by new onto-epistemologies that can withstand the violence of the past fueling and coupled with uncertainties of the present day. Forging new educational institutions, arrangements and collectives the decolonial can guide our thinking, knowing and learning between the past and dead futures to engage a metaphysics of being untethered to race, patriarchal heteronormativity, capital and territory.

References

- Adichie, C.N., 2009. The Danger of a Single Story. Ted Talk.
- Andreotti, V., Stein, S., Akenakew, C., Hunt, D., 2015. Mapping interpretations of decolonization in the context of higher education. *Decolonization* 4 (1), 21–40.
- Andreotti, V., Stein, S., Pashby, K., Nicolson, M., 2016. Social cartographies as performative devices in research on higher education. *High Educ. Res. Dev.* 35 (1), 84–99.
- Asher, N., 2009a. Decolonization and education: locating pedagogy and self at the interstices in global times. *Counterpoints* 369, 67–77.
- Asher, N., 2009b. Writing home/decolonizing text (s). *Discourse* 30 (1), 1–13.
- Battiste, M., 2019. *Decolonizing Education: Nourishing the Learning Spirit*. UBC Press.
- Coetzee, J.M., 2021. Growing up with the children's Encyclopedia. *Engl. Afr.* 48 (3), 7–28.
- Da Silva, D.F., 2014. Toward a Black feminist poethics: the quest(ion) of blackness toward the end of the world. *Black Scholar* 44 (2), 81–97.
- Donald, D., 2021. We need a new story: walking and the wāhkōhtowin imagination. *J. Can. Assoc. Curric. Stud.* 18 (2), 53–63.
- Dei, G.J.S., 2006. Mapping the terrain—towards a new politics of resistance. In: *Anti-colonialism and Education*. Brill, New York, pp. 1–23.
- Fisher, M., 2014. *Ghosts of My Life: Writings on Depression, Hauntology and Lost Futures*. John Hunt Publishing, New York.
- Fanon, F., 2004. *The Wretched of the Earth*. 1961. Trans. Richard Philcox. Grove Press, New York, p. 6.
- Garba, T., Sorentino, S.M., 2020. Slavery is a metaphor: a critical commentary on eve tuck and K. Wayne Yang's “decolonization is not a metaphor. *Antipode* 52 (3), 764–782.
- Grande, S., 2015. *Red Pedagogy: Native American Social and Political Thought*. Rowman & Littlefield.
- Grosfoguel, R., 2007. The epistemic cultural turn. *Cult. Stud.* 21 (2–3), 211–223.
- Lewis, G., 2019. On Black Pain/Black Liberation and Rise of Fallism. Blog of the APA. Access at: <https://blog.apaonline.org/2019/03/19/on-black-pain-black-liberation-and-the-rise-of-fallism/>.
- Love, B.L., 2019. *We Want to Do More than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon Press.
- Maccauley, T.B., 1835. Minute by the Honorable T.B. Maccauley. Access at: http://www.columbia.edu/itc/mealac/pritchett/00generallinks/maccauley/txt_minute_educati_on_1835.html.
- Mignolo, W., 2007. Delinking. *Cult. Stud.* 21 (2), 449–514.
- Mignolo, W., 2011. *The Darker Side of Western Modernity: Global Futures, Decolonial Options*. Duke University Press, Durham.
- Mignolo, W., 2017. Interview- Walter Mignolo/Part 2: Key Concepts. Access at: <https://www.e-ir.info/2017/01/21/interview-walter-mignolopart-2-key-concepts/>.
- Mmembe, A., 2020. Out of the Dark Night. *Essays of Decolonization*. Duke University Press, Durham, NC.
- Mishra Tarc, A., 2015. *Literacy of the Other. Renarrating Humanity*. Suny Press, New York.
- Ngugi, T., 1981. *Decolonizing the Mind. The Politics of Language in African Literature*. Heineman, London.
- Ngugi, M., 2018. What decolonizing the mind mean today?. In: *The Work of Linguistic Decolonization Can't Be Done by Writers Alone*. Lit Hub. Access at: <https://lithub.com/mukoma-wa-ngugi-what-decolonizing-the-mind-means-today/>.
- Roy, A., 2003. Another World is Not Only Possible, She is on Her Way. Access here. <https://www.workersliberty.org/story/2017-07-26/world-social-forum-arundhati-roy>.
- Santos, B., 2014. *Epistemologies of the South. Justice Against Epistemicide*. Routledge, New York.
- Scott, D., 2004. *Conscripts of Modernity. The Tragedy of Colonial Enlightenment*. Duke University Press, Durham, NC.
- Snaza, N., Mishra Tarc, A., 2019. To wake up our minds”: the re-enchantment of praxis in Sylvia Wynter. *Curric. Inq.* 49 (1), 1–6.
- Soudien, C., Cornille, C., 2000. South African higher education in transition: global discourses and. In: Stromquist, N., Monkman, K. (Eds.), *Globalization and Education: Integration and Contestation across Cultures*. R & L Education, New York.
- Stein, S., Andreotti, V., Suša, R., Amsler, S., Hunt, D., Ahenakew, C., et al., 2020. Gesturing towards decolonial futures: reflections on our learnings thus far. *Nordic J. Compar. Int. Educ.* 4 (1), 43–65.

Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.

Wynter, S., 2003. Unsettling the coloniality of being/power/truth/freedom: towards the human, after man, its overrepresentation—an argument. *CRN. Centen. Rev.* 3 (3), 257–337.

Further reading

Andreotti, V., Stein, S., Sutherland, A., Pashby, K.L., Susa, R., Amsler, S., 2018. Mobilising different conversations about global justice in education: toward alternative futures in uncertain times. *Policy Pract.* 26, 9–41.

Barker, J., Byrd, J.A., Goldstein, A., Grande, S., Bernal, J., DeVore, R., et al., 2021. Catastrophe, care, and all that remains. *Soc. Text* 39 (4), 27–53.

Da Silva, D.F., 2016. On Difference Without Separability. Catalogue of the 32a São Paulo Art Biennial, “Incerteza Viva” (Living Uncertainty).

Desai, K., Sanya, B.N., 2016. Towards decolonial praxis: reconfiguring the human and the curriculum. *Gend. Educ.* 28 (6), 710–724.

Mignolo, W.D., 2021. *The Politics of Decolonial Investigations*. Duke University Press.

McKittrick, K. (Ed.), 2015. *Sylvia Wynter: On Being Human as Praxis*. Duke University Press, Durham.

Snaza, N., Mishra Tarc, A., 2019. “To wake up our minds”: the re-enchantment of praxis in Sylvia Wynter. *Curric. Inq.* 49 (1), 1–6.

Smiet, K., 2022. Rethinking or delinking? Said and Mignolo on humanism and the question of the human. *Postcolonial Stud.* 25 (1), 73–88.

Relevant websites

Common Worlds Collective. <http://commonworlds.net/>.

Gesturing Towards Decolonial Futures Project. [http://blogs.ubc.ca/towardsdecolonialfutures/Decolonizing Teacher Education](http://blogs.ubc.ca/towardsdecolonialfutures/Decolonizing%20Teacher%20Education). <https://indigenizinglearning.educ.ubc.ca/>.

Citizenship

Abigail B. Bakan, Department of Social Justice Education, Ontario Institute for Studies in Education (OISE), University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	38
The Canadian State and multiculturalism	39
Rhetoric versus reality: racism and dispossession	39
From immigration to security to pandemic politics	41
Conclusion: implications for education for citizenship	43
References	44

Introduction¹

Citizenship is commonly understood as a legal category. For example, the name of the state that defines one's passport is often seen as an important signifier of identity.

Legal citizenship status is certainly significant, as this is often pivotal in determining certain rights and protections in relation to states. Legal citizenship status can be one signifier for national identity, including feelings of patriotism that have proven to be central to many political acts such as wars, occupations, settler colonialism and border controls (Anderson, 2006; Wolfe, 2016). However, a focus on national identity regarding citizenship risks blurring the complexity of differences and the intersections within, across and between identities (Crenshaw, 1991; Carastathis, 2016).

Legal status is actually only part of how we think about the many facets of citizenship, including borders, belonging, relationships, and interactions with other identities or peoples who are non-citizens. Citizenship, and its relationship to identity and difference, suggests a series of processes more complex and varied than a passport, or other symbols like flags or currencies. This complexity is indicated when we consider certain questions, such as why everyone, or every nationality, does not have a recognized passport, and many lack legal status in a given country. Other related questions include: who is deemed a citizen of a state and who is not? How are citizens of the same state treated differently, and why? Why are some citizenship statuses considered important, while others are treated as diminished or even threatening? And how and why are some populations denied rights of citizenship through violence and dispossession?

Importantly, citizenship for some always occurs against a background of non-citizenship for others. Citizenship elevates borders and related power relations when the line between citizenship and non-citizenship is defined (Stasiulis and Bakan, 2005). Citizenship, therefore, is actually better understood as a spectrum of negotiated and contested rights and obligations among people and states. This spectrum includes features of both state repression and protections, which are changing and unstable. The spectrum of citizenship and non-citizenship, varies greatly over time, place and populations. This reveals tensions between the universalistic discourses and discriminatory impact of policies and practices of modern democratic states (Stasiulis and Bakan, 2005). Legal citizenship in states in the global north – such as Canada, the United States, or the states of Western Europe – for example, are considered to represent political memberships that are elevated in a global hierarchy, whereas citizenship in countries of the global south are different, often altered by patterns of colonialism and imperialism. Moreover, within states of the global north, large sections of the populations – including for example Black, Indigenous and racialized populations, sexual and gender minorities, and those living with disabilities – also have differential access to rights of citizenship.

The spectrum of citizenship and non-citizenship generates various types of differences, and multiple identities. In what follows, drawing on the example of Canada as a case study, the discussion addresses contradictions and tensions regarding the citizenship/non-citizenship spectrum from the perspective of differences and identities. Canada is a notable example because it is widely seen as a state modeled on successful, inclusive citizenship. In fact, such a rhetorical promise of universality stands in contradiction to reality, and the gap between the promise and experience of multiple exclusions reveals the ongoing character of Canada as a colonial settler state.

While Canada, like all states, indicates unique features and a specific history of economic, political and social experiences, as a liberal democratic state which belies the promise of inclusive citizenship it can serve as an important entry point to wider comparisons. The focus of this chapter is specifically on the exclusionary impact of racism in Canada, grounded in a political economy of exploitation and colonialism (Bakan and Dua, 2014; Dei et al., 2000; Dei and Kempf, 2013; Galabuzi, 2006; Bannerji, 2014; Thobani, 2007; Razack et al., 2010). The Indigenous and Black populations of Canada are particularly impacted (Monture-Angus, 1995;

¹Some parts of this chapter are based in Abigail B. Bakan (2016) "Multiculturalism and its Contradictions: Education for Citizenship and Social Justice in Canada", in *The Palgrave International Handbook of Education for Citizenship and Social Justice*, eds. Andrew Peterson, Robert Hattam, Michalinos Zembylas and James Arthur (UK: Palgrave), 347–368.

Canada RCAP, 1996; TRC, 2015; McKittrick, 2006; McKittrick and Woods, 2007; Walker, 2010), as well as, most notably in the years since September 11, 2001 (9/11), those ascribed to be associated with the Muslim faith or of Arab origin (Abu-Laban and Dha-moon, 2009).

This discussion proceeds by first considering the Canadian state as an example of “multiculturalism”. The rhetoric of multiculturalism will then be considered against the realities of difference and oppression, with a focus on racism and dispossession. Then, turning to the contemporary period, with a focus on immigration and the period from 9/11 through the pandemic of the 21st century, the ways in which certain crises reveal and exacerbate these contradictions will be addressed. The chapter will conclude with a consideration of the implications for education for citizenship.

The Canadian State and multiculturalism

The context in which contested notions of citizenship, identity and power emerge is significant. Canada is a federated state, where the federal, central government shares power according to specific terms with elected provincial parliaments or assemblies. The central parliamentary government in the nation’s capital city, Ottawa, resides in the province of Ontario, and manages the overarching affairs of citizenship according to designated powers. Subject to this central authority, other governing powers specific to certain regions rest with Canada’s provinces and territories (Franks, 1987; Tully, 1995). Provinces have specific regulatory authorities regarding education, employment, health care, language and residence.

Multiculturalism was adopted as the official policy of the Canadian federal government in 1971. This was preceded by the report of the Royal Commission on Bilingualism and Biculturalism from the 1960s (Haque, 2012); the policy adoption was followed by the passage into law of the Canadian Multiculturalism Act in 1988. The Multiculturalism Act also followed the repatriation of the Canadian Constitution in 1982, which included a Charter of Rights and Freedoms (Canada, 1982). Significantly, the Multiculturalism Act notes in the preamble that the Canadian state recognizes diversity among its citizens – specifically regarding race, ethnic or national origin, religion and color – and indicates a commitment to preserving and enhancing this multicultural heritage. The Multiculturalism Act also commits the state to working to achieve “equality of all Canadians”, in every aspect of citizenship, including the economic, social, political and cultural dimensions (Canada, 1988, preamble, 2).

The language presented by the Canadian state, both domestically and on the world stage, is bold in its claimed assurance of inclusiveness and equity, noting that multiculturalism is “fundamental” and “encourages racial and ethnic harmony and cross-cultural understanding, and discourages ghettoization, hatred, discrimination and violence” (Canada, 2022). The Act has also been the subject of intense critique, noting the gap between the bold rhetoric and the experiences of ongoing and intense discrimination within Canada.

A number of examples indicate the gap between promises and reality. Millions of Canada’s citizens have left their homes of origin and all that is familiar to them to seek permanent residence in Canada. Canada has one of the highest rates of immigration in the world, measured in terms of both overall and net migration (IOM, 2022). Moreover, Canada also has one of the highest rates of naturalization of any country, with a large majority of those who immigrate successfully obtaining legal status of Canadian citizenship (Picot and Hou, 2011a, p. 171). In terms of formal educational success, second-generation immigrants to Canada measure highly compared with international indicators (Picot and Hou, 2011b).

These numbers, however, belie experiences of systemic, structural discrimination evidenced in the formal education system. This contradictory pattern, where the promise of inclusion for diverse communities is halted by powerful barriers to equity in education, is characteristic from primary through to university levels, with lifelong implications (Cunningham, 2007, p. 153; Razack, 1998; Maynard, 2017). The official context of formal inclusion associated with multiculturalism serves to mask continuing and deepening exclusionary barriers to equal citizenship for marginalized sectors of Canadian society. Historic contextual factors include formal colonialism, racial slavery and the genocidal residential schooling of Indigenous children. The contemporary context continues to enforce patterns of racism and colonialism, even in an age of redress and a political climate of recognition (Coulthard, 2014). The patterns of discrimination in Canadian citizenship practices remain deeply entrenched, based on class, ability, gender, language and ethnicity, among other factors (Razack, 1998, 2015; Stasiulis and Bakan, 2005; Titchkosky, 2003; Little, 1998; Malhotra and Rowe, 2014).

The contemporary period can be traced to Canada’s policies during World War II, through to the period of the “war on terror” following 9/11 and related policies, and the contemporary period of COVID-19 pandemic.

Rhetoric versus reality: racism and dispossession

The specific historical context of Canada’s claim to multiculturalism is significant. Canada was established as a white settler state, officially founded in 1867, where British subjects were privileged to be eligible for the legal status of citizens (Abele and Stasiulis, 1989; Stasiulis and Jhappan, 1995; Thobani, 2007).

Imperialism is a system of state power based on empire, where the rulers of one state own and control the land, resources, rights to population and products of labor of another state, whether through formal policy and politics or military force, or both. As Edward Said notes, “[m]odern imperialism was so global and all-encompassing that virtually nothing escaped it ... [and] the nineteenth century context over empire is still continuing today” (Said, 1993, p. 68). Colonialism, the policy of imperialism based on dispossession and colonization, takes various forms. It continues today even where formal governmental autonomy has been achieved in former colonies (Coulthard, 2014).

Multicultural policy followed from debates about the nation's linkages to European empire and colonialism. The Royal Commission on Bilingualism and Biculturalism was commissioned in 1963, under the Liberal government of Prime Minister Lester B. Pearson, in response to a rise of French-Canadian nationalist sentiment in the province of Quebec (Gagnon, 2004). Among the Commission's key recommendations was that Canada adopt both French and English as official languages; this recommendation was accepted and made into law under the administration of federal Liberal Prime Minister Pierre Elliot Trudeau (father of Canadian Liberal Prime Minister Justin Trudeau) in 1971. Second in size only to Ontario, and constituting about one-third of the Canadian population, Quebec is the only one of ten provinces where French is the first language for 80% of the population. While recognizing the significance of Canada's historically excluded and oppressed Francophone population, the Commission's report came under heavy criticism for its Eurocentric focus on those of English and French descent as foundational to the Canadian nation. Multiculturalism emerged as a counter-point to "biculturalism", largely because "non-British, non-French, and non-Aboriginal Canadians – especially those of Ukrainian origin – challenged the symbolism of a bicultural and bilingual Canada" (Abu-Laban, 2007, p. 140). As multiculturalism advanced as a dominant governmental initiative, and as a public narrative of Canadian nation-building, its limitations – absencing Indigenous rights and languages, absencing Canada's Black population and privileging European colonial settlement – have become the subject of substantive critique, in scholarship, policy and social movement activism (Stasiulis, 1988; Tully, 1995; Day, 2000; Mackey, 2002; Galabuzi, 2006; Abu-Laban, 2007; Thobani, 2007; Kernerman, 2005; Dhamoon, 2009; Haque, 2012; Coulthard, 2014).

The Canadian state's close association with British imperial policy is well-documented and widely recognized, originating from a federation of colonial territories in North America. Indeed, Canada is, to this day, a constitutional monarchy. The connections to the British Crown are symbolically central to national cultural signifiers including anthems, currency, cultural practices and holidays, citizenship law and the structure of parliamentary governance. This fact of Canada's foundational history is not subject to contention. However, the details of the colonial experience and the implications of this association and its continued impact on contemporary issues of social justice and citizenship are key issues of ongoing debates, influencing social, economic, cultural and political spheres.

The process of settlement included centuries of conflict among European empires. This included competition over North America with France, whose prisoners were among those exiled to colonial North America where the Catholic Church was a primary actor (see Gagnon, 2004). The Indigenous peoples of the land of Turtle Island (North America, before contact) encountered these European strangers variously over the period from first encounters, with some sections of French settlers interacting with Indigenous peoples, emerging into the population of Canada's Métis (Lawrence, 2004; Lischke and McNab, 2007). However, current historical studies demonstrate that violent European colonial dispossession, theft and occupation of Indigenous land, suppression of language and culture, violence toward children, and the denial of rights were foundational to the Canadian state project (Tully, 1995; Green, 2001; Coulthard, 2014; Simpson, 2014; Maynard, 2017; TRC, 2015).

The context of Canada's history regarding European colonialism, white colonial settlement and severe oppression of the Indigenous population is shared with its neighbor to the south, the United States (USA). The 13 colonies that rebelled against the Kingdom of Britain in the American Revolutionary War (1775–83) were not joined by the colonies that were to form Canada, despite rebel invasions in the regions of Quebec and Nova Scotia (Ryerson, 1975, 1983). The federal Dominion of Canada was formed through the confederation of several North American British colonies in 1867, just two years after the US Civil War. Three British colonies (the United Province of Canada, New Brunswick and Nova Scotia) became the first four provinces of Canada (Ontario and Quebec, New Brunswick and Nova Scotia), later to be joined by others to constitute the federated state of Canada today.

A common national narrative in Canada is that only in the USA was slavery instituted, while Canada was a place of freedom. This is a myth, as slavery was widespread in Canada. While Canada organized slavery differently from the US in terms of political economy, in both states institutionalized slavery was central to the settler state accumulation projects (Maynard, 2017). Particularly relevant to the historical context regarding social justice in Canada, the United States incorporated and recognized plantation slavery in its core political economy until the end of the American Civil War (1861–65). But the construction of national historical memory is a central feature of modern nation-making (Hobsbawm, 1992). The mythologized notion that Canada has a history of freedom from slavery is consistent with this type of constructed historical memory, one that fits well with the advancement of contemporary multiculturalism.

The historical record, however, runs contrary to the dominant narrative. The colonial territory that is now Canada was, in fact, formally committed to slavery until 1834, when abolition in the British colonial empire was decreed, and enforcement of slave freedom was limited (Winks, 1997; Walker, 2010; Clarke, 2006; Cooper, 2005). Indeed, Canada's first Prime Minister, Sir John A. MacDonald, was a firm supporter of the US South, which was committed to plantation slavery, in the Civil War. A lawyer by profession, MacDonald was the hired advocate for the Copperhead conspirators – an organization which advocated for the victory of the US South by committing vigilante acts of arson and raids on northern cities (Ryerson, 1983, pp. 334–335; Bakan, 2008, pp. 18–19).

Canada's racist practices toward Black residents – many of whom emigrated originally to Canada from the United States as slaves of British loyalists – as well as toward more recent immigrants from the Caribbean, Africa and other regions, is well-documented (Tennyson, 1990; Winks, 1997; Henry and Tator, 2002; Stasiulis and Bakan, 2005; Cooper, 2005; Galabuzi, 2006; McKittrick, 2006; Walker, 2010; Maynard, 2017). Anti-black racism was, and is, consistent with colonial patterns of dispossession and occupation, and the historic commitment to "civilization" as an explicitly European, white settler project.

Canada's original Indigenous population, was, and continues to be, victim to extreme racism and violence. The system of forced residential schooling for Indigenous children has, after considerable denial, finally been acknowledged, most notably in an extensive report by the Truth and Reconciliation Commission of Canada (TRC) (TRC, 2015). Following decades of active organizing among survivors of the residential schools, in 2008 then Prime Minister of Canada Stephen Harper formally apologized to the survivors of this abusive system. The Truth and Reconciliation Commission report concluded, however, that this "promise of reconciliation" had "faded" through inaction (TRC, 2015, p. 8).

The main finding of the TRC was that the residential school system was consistent with the central goals of Canada's policy toward Indigenous, or Aboriginal, peoples, and this was a policy best described as "cultural genocide" (TRC, 2015, p. 1). The schools, as the report noted, were not about education as we commonly understand its meaning. Instead, they were "created for the purpose of separating Aboriginal children from their families" in order to weaken their relationships, and "to indoctrinate children into a new culture – the culture of the legally dominant Euro-Christian Canadian society" (TRC, 2015, p. v). The schools, more like prisons for children, were in place for over a century. They were already established by churches in Canada prior to the country's origin in 1867, and were supported through the 1870s by the Roman Catholic and Protestant missionaries as the settlement project moved across the west and north of the continent. In 1883, the policy of the residential schools was formally endorsed and advanced with federal government support (TRC, 2015, p. 3).

Canada's inclusive multiculturalism stands in an uncomfortable, and contradictory, relationship with a long-denied history and continued daily violence and discrimination against the original inhabitants of Turtle Island (North America) (TRC, 2015, p. 3). Many issues regarding Indigenous peoples in Canada, not only historic but ongoing in present times, have remained unaddressed and unrecognized for decades. For example, a campaign to seek justice for the victimization of Indigenous women, who have gone missing or been found murdered in their hundreds, was met with minimal state and media attention (Razack et al., 2010; Amnesty International Canada, 2022). Only in 2015, after a mass public campaign and the election of Liberal Prime Minister Justin Trudeau, did the federal government announce a commitment to advance a national public inquiry. The findings were published following a contentious process fraught with considerable controversy (Canada, 2019). Ultimately, there were extensive consultations through the National Inquiry into Missing and Murdered Indigenous Women and Girls (MMIWG), including 2386 participants. Among these were 1484 family members and survivors of violence who participated in 15 community hearings and provided testimony; 819 individuals sharing artistic expressions; and the contributions of 83 elders, knowledge keepers and officials (Canada, 2019). Between 1980 and 2012, Indigenous women were estimated to comprise 16% of all female homicides, despite representing only 4% of the female population. This is a number reported by the Royal Canadian Mounted Police (RCMP) and considered to be an underestimation (Canada, 2019, p. 55). Indeed, the Final Report of the National Inquiry concluded that the Canadian state's acts against Indigenous people, "including violence against Indigenous women, girls and 2SLGBTQIA people", were consistent with the definition of "genocide" (Canada 2019, p. 54).

For too many, however, the consultations were far too late, as many lives were lost in the years of advocacy when the government refused to even acknowledge the reality of the violence (Palmater, 2021). This is illustrative of a pattern of reluctant, partial and limited accommodation to long-standing demands for Indigenous rights, forwarded in the language of recognition. The pattern marks a form of colonial politics that continues, and has become the subject of considerable analysis and challenge (Coulthard, 2014; Simpson, 2014). There are numerous other examples to confirm this pattern of inclusive rhetoric belied by experiences of overt racism and dispossession, only a few of which can be addressed in this brief discussion. This brings us to consider the wider systemic features associated with citizenship, identity and difference.

From immigration to security to pandemic politics

Canada's original infrastructure is seeped in colonialism and racism. The extensive railroad system, built by the Canadian Pacific Railway through the early 1880s, depended on tens of thousands of laborers from China who worked in dangerous conditions without legal rights. Fearing that these workers, almost all male, would immigrate to Canada permanently, a prohibitive head tax was levied by the federal government, through the Chinese Immigration Act of 1885. Decades later, the Act was superseded by the Chinese Immigration Act of 1923, also known as the Chinese Exclusion Act (Fernando, 2011). This act was repealed in 1948. Only in 2006, however, after decades of organizing and demanding redress, was a formal apology and some compensation extended by Prime Minister Stephen Harper (Prime Minister's Office's, 2006).

While the case of the Chinese head tax is one example, the practice of formal state restrictions of unwanted immigrants is widespread. It has affected numerous communities in Canada and internationally. Taking another example, in May of 1914, a ship arrived on Canada's western shore, in the city of Vancouver, British Columbia, with 376 Sikh, Hindu and Muslim passengers from the Punjab, India. The Komagatu Maru was not allowed to disembark, despite a concerted legal challenge raised by Vancouver's South Asian community. After being docked with its passengers in the harbor for two months, the ship was forced to return to India. Immigration from India was virtually frozen for the next 25 years (Miki, 2004, p. 22; Dhamoon et al., 2019; Coloma and Pon, 2017). Anti-Asian racism has continued in Canada, increasing in times of crisis such as the COVID-19 global pandemic (Claxton et al., 2021).

Immigration controls have often tightened in times of crisis, on claimed grounds of security, or protection of limited employment opportunities. But what has also increased is discrimination against those considered to "look like" immigrants, regardless of legal citizenship status. During World War II, when Canadian soldiers participated with the Allied powers, Japanese Canadians were treated as enemies at home. Between 1941 and 1943, a series of federal orders-in-council denied equal rights to Japanese Canadians.

These rulings required the registration, fingerprinting, forced movement to designated “protected” areas, confiscation of property and the deportation of Japanese Canadians. These actions were ostensibly on grounds of security, on the unfounded assumption that sympathy with Japan and the Axis powers could be assumed by racial profiling. Significantly, no charges were laid, but trials or hearings were held ([Canadian Japanese Internment Camps, 2011](#); [Johnson and Enomoto, 2007](#)). Neither was this experience of racial profiling, however, invented when security threats were asserted. There is a long earlier history of provincial and federal legal restrictions of Asian Canadians, including the 1897 British Columbia Elections Act which denied the franchise to citizens of “Asiatic” origin ([Fernando, 2011](#)).

In 1988, after decades of organizing to demand an apology and compensation, the National Association of Japanese Canadians saw the federal government, under the administration of Prime Minister Brian Mulroney, sign the Japanese Canadian Redress Agreement. This agreement was signed in the aftermath of the 1982 Constitution Act, which included the Canadian Charter of Rights and Freedoms. The Charter was explicitly noted in the redress agreement, which was understood to afford minority communities the right to protection against the recurrence of discriminatory treatment ([Miki, 2004](#), pp. 317–318; [CRRF, 2022](#)).

The wartime history of Canada holds other examples of such contradictions in terms of social justice and citizenship. The public claim to be committed to protecting democracy from Hitler’s Nazis during World War II was not matched by compassion for victims of anti-Jewish racism, or anti-Semitism, at home. In fact, the Canadian government’s policy toward the plights of Jewish refugees fleeing the Nazi Holocaust was summarized in the statement of Frederick Charles Blair, the director of the Immigration Branch of the wartime government of Liberal Party Prime Minister William Lyon Mackenzie King, as “None is too many” ([Abella and Troper, 1983](#)).

Over the course of the post-war years, Canada’s immigration policy was amended to meet the demands of a growing labor market. But still, this was not universal, as some sectors, such as foreign live-in domestic care workers, saw working conditions in Canada decline as other migration options for women workers from countries in the global South tightened ([Bakan and Stasiulis, 1997](#); [Stasiulis and Bakan, 2005](#)).

The Canadian economy as a whole has greatly benefited from the labor of millions of new immigrants, while changes in Canada’s political economy have affected the contemporary context ([Kellogg, 2015](#)). The 1960s saw a period of reform in immigration policy, with high demand for foreign labor, particularly skilled labor. Considered on a world scale, Canada is a nation where there is a relatively high level of acceptance of new immigrants ([Kymlicka, 2010](#), p. 7). Yet, these indicators belie other trends in immigration, including restrictive policies in other states globally, and unequal conditions among immigrants to Canada based on class, gender and race ([Galabuzi, 2006](#); [Stasiulis and Bakan, 2005](#)). Though certain categories of immigration, particularly those geared to family unification, opened opportunities to communities from countries of the global South that had historically been unwelcome to reside permanently in the country, this was neither universal nor consistent over time.

Canada’s immigration and citizenship policies have generally tightened in the 21st century. Issues regarding immigration and citizenship have been influenced by public debates on security. Racial profiling of Black and Indigenous Canadians, regardless of citizenship status, continues despite documentation of clear violation of principles of human rights ([TRC, 2015](#); [Maynard, 2017](#)). Neither is there a linear progression of increasing rights over time. The Canadian government readily allied itself with the United States and other Western powers following the 9/11 attack on the Pentagon and the World Trade Center. However, policy amendments relied on racial and cultural stereotypes, rather than protection of Canadian civil and human rights ([Thobani, 2002](#)). A significant example is the case of Maher Arar, a dual Canadian and Syrian citizen, who, while traveling in 2002, was detained by US authorities on suspicion of being a member of Al Qa’ida, the terrorist organization associated with the 9/11 attacks. The details of Arar’s experience are significant. His story struck a chord with many immigrant and Canadian Muslim families in Canada, who expected to be welcomed and recognized as part of the multicultural state ([Dhamoon and Abu-Laban, 2009](#)). Arar had immigrated to Canada from his native Syria at the age of 17, where he received a post-secondary education and, by the age of 32, was established as a highly skilled telecommunications engineer. He was married and the father of two children. On September 26, 2002, Arar was traveling on a Canadian passport, returning to his home in Canada by way of New York, following a family holiday in Tunisia. He was detained in the Kennedy Airport in New York by US officials and held in solitary confinement until October 2, 2002. Despite the interventions of his family and a US-based lawyer, US Immigration and Naturalization Services officials advanced accusations that Arar was a member of Al Qa’ida. On the basis of these unfounded accusations, he was deported and remained in prison and tortured in Syria until October 5, 2003 ([Abu-Laban and Nath, 2007](#)). Only following a concerted campaign led by his wife, Monia Mazigh, and other supporters, was Arar released and returned to his home in Ottawa, Canada. Public support to investigate the case and clear Arar’s name continued, resulting in several formal investigations ([Abu-Laban and Nath, 2007](#)). As a result of the findings, Arar was eventually cleared of all wrongdoing. In 2007, then Canadian Prime Minister Stephen Harper issued an official apology and financial compensation for this act of extraordinary rendition to another state, in recognition of the damage caused to Arar and his family.

While security claims have continued to be adopted to justify discrimination against racialized groups rendering more difficult the transition for immigrants to full legal citizenship status, all forms of inequality have been revealed and intensified in the conditions of the pandemic virus COVID-19. The global pandemic, declared by the World Health Organization in March 2020 ([WHO, 2020](#)), has revealed this pattern internationally, and Canada has been no exception. While the virus threatens lives universally, those in precarious work and living situations, with limited access to healthcare, and suffering from poor health associated with discrimination and poverty have face heightened risks. Internationally, there has been a marked rise in anti-Asian racism, including racial profiling, hate crimes and social media targeting ([He et al., 2021](#); [Gover et al., 2020](#)). Moreover, persistent social and economic inequalities have proven life-threatening in pandemic conditions, indicating ethnic minorities, the poor and the elderly are

particularly vulnerable (Ali et al., 2020). Black, Indigenous and people of color populations have disproportionately suffered from COVID-19 in Canada as well as internationally (Dei and Lewis, 2020). Black families face significantly higher rates of poverty in Canada, with approximately 50% of the Black Canadian population “categorized as low income, with that number jumping to 65% for new Black immigrants” (UN General Assembly Human Rights Council 2017, p. 12).

Conclusion: implications for education for citizenship

These historical and contemporary contexts regarding identities, difference and citizenship suggest important implications for education.

Canada’s national project has been strongly aligned with the achievements of British colonialism. At the same time, the country’s growth has been centrally dependent on immigration, largely from nations in the global South, to fuel its expanding labor force and population. An historic reliance on an assimilationist conception of citizenship was expressed largely through the education system (Heath, 2002). Loyalist notions of Christianity, patriarchy and state authority were carried by educational theorists such as Egerton Ryerson (1803–82), after whom a major Canadian university was named (see Pearson, 1988; Dellaire and Ellis, 2021). Indeed, Ryerson’s commitment to public education for children of Canadian settlers was coincident with his advocacy for residential schools for Indigenous children, understood as two sides of the colonial project (Dellaire and Ellis, 2021). Schools have therefore been sites of deep contention, with teachers, students, parents and scholars leading movements for increased access and equality for marginalized communities, and for redress and measures to mitigate acknowledged patterns of discrimination (Banks, 2004). Following consistent pressure raised within and beyond Ryerson University for over a decade, a university task force to investigate the legacy and implications recommended changing the name of the institution, to no longer recognize a person so deeply implicated in the system of residential schools and related discriminatory and harmful acts. In 2022, the university adopted a new name, as Toronto Metropolitan University. This is a limited reform, but it is reflective of the pattern of transformation following concerted efforts and resistance.

Formal citizenship education has expanded in Canada. While ideals of social justice, diversity and inclusion are more common in educational curricula, these goals are compromised by an emphasis on a singular, uniform Canadian national identity (Pashby, 2014), one which continues to perpetuate systemic discrimination (James, 2007; Mujawamariya, 2007; Dei et al., 1997). Scholars of social justice have forwarded innovative theorizations to address these contradictory trends, positing the concept, for example, of negotiated citizenship (Stasiulis and Bakan, 2005). Citizenship, as argued here, needs to be understood as more than legal status alone, existing “on a spectrum, involving a pool of rights that are variously offered, denied, or challenged, as well as a set of obligations that are unequally demanded” (Stasiulis and Bakan, 2005, p. 2). Viewed in this way, citizens and non-citizens can be seen to continually challenge various restrictions and barriers, while also collectively negotiating to advance rights and freedoms.

Within education broadly defined to embrace experiential learning beyond the school classroom, there are also contested ideas. The adoption of formal multiculturalism policy in the 1970s and the expansion of constitutional rights in the 1980s are notable markers. The contradictions between policy and reality however, have generated significant resistance, and this is a story that is still in the telling. Canada is the chosen home of millions of people from around the world, while it also remains the place of residence to Indigenous peoples from time immemorial. For those born in other nations, obtaining legal citizenship in Canada requires meeting a series of complex requirements, ultimately including an affirmation of an oath of loyalty. This oath calls for a promise to “bear true allegiance” to Her Majesty Queen Elizabeth II, “Queen of Canada”, as well as to her heirs and successors, and to observe the laws of the land. The oath, along with various other rights and responsibilities of citizenship, is described in Discover Canada, the study guide for those seeking citizenship published by Citizenship and Immigration Canada (Canada, 2022). While the content of this text emphasizes multiculturalism, gender equity – including same-sex marriage and women’s rights – and diversity, the commitment to the Queen as the sovereign personified reveals the contradictory character of Canadian citizenship and the continued colonialism in theory and practice. This complex spectrum demands continued study and research, and a central commitment to the theory and practice of social justice education.

While the country case study of Canada is specific, the experiences of colonialism, dispossession, racism, exploitation, sexism and gender discrimination, are painfully common among nations and populations across the globe. Indeed, the case of Canada is notable, arguably, as a “best case” scenario of modern neoliberal capitalism in terms of its promise of inclusive democracy and equality. Canada is widely taken as a model state, raised as an ideal type to be emulated and with exemplary lessons that can be shared internationally (Stasiulis and Bakan, 2005). This is demonstrated, for example, in the elevation of a school of Canadian political theory, drawn from examples of multiculturalism in Canada and generalized to apply to liberal democratic states as a model of governance (see for example Kymlicka, 1995, 2001, 2010; Taylor, 2004, 2009). Canada is framed as a place of refuge, and certainly it is the home to many thousands who have fled onerous and often life-threatening conditions. It is precisely this elevated status, however, that compels a closer look at the realities of oppressive and colonial practices in Canada, not least to expose the realities and intransigence of systemic discrimination. As the discussion presented above has demonstrated, the rhetoric outstrips, and covers for, a very different reality, and merits continued action and support for profound social transformation. Surely, a better world is possible.

References

- Abella, I., Troper, H., 1983. *None Is Too Many: Canada and the Jews of Europe 1933–1948*. Lester and Orpen Dennys, Toronto.
- Abele, F., Stasiulis, D., 1989. Canada as a 'White Settler Colony': what about natives and immigrants? In: Clement, W., Williams, G. (Eds.), *The New Canadian Political Economy*. McGill-Queen's University Press, Montreal.
- Abu-Laban, Y., 2007. Political science, race and public policy. In: Orsini, M., Smith, M. (Eds.), *Critical Policy Studies*. UBC Press, Vancouver, pp. 137–157.
- Abu-Laban, Y., Nath, N., 2007. From deportation to apology: the case of Maher Arar and the Canadian State. *Can. Ethnic Stud.* 39 (3), 71–97.
- Abu-Laban, Y., Dhamoon, R., 2009. Dangerous (internal) Foreigners and nation-building: the case of Canada. *Int. Polit. Sci. Rev.* 30 (2), 163–183.
- Ali, S., Asaria, M., Stranges, S., 2020. COVID-19 and inequality: are we all in this together? *Can. J. Public Health* 111, 415–416.
- Amnesty International Canada, 2022. No More Stolen Sisters. <http://www.amnesty.ca/our-work/campaigns/no-more-stolen-sisters>. (Accessed 5 August 2022).
- Anderson, B., 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. revised edition. Verso, London and New York.
- Bakan, A.B., 2008. Reconsidering the underground railroad: slavery and racialization in the making of the Canadian state. *Social. Stud.* 4 (1), 3–29.
- Bakan, A.B., Dua, E. (Eds.), 2014. *Theorizing Anti-racism: Linkages in Marxism and Critical Race Theories*. University of Toronto Press, Toronto.
- Bakan, A.B., Stasiulis, D. (Eds.), 1997. *Not One of the Family: Foreign Domestic Workers in Canada*. University of Toronto Press, Toronto.
- Banks, J.A., 2004. Teaching for social justice, diversity, and citizenship in a global world. *Educ. Forum* 68 (4), 296–305.
- Banerji, H., 2014. Marxism and anti-racism in theory and practice: reflections and interpretations. In: Bakan, A.B., Dua, E. (Eds.), *Theorizing Anti-racism: Linkages in Marxism and Critical Race Theories*. University of Toronto Press, Toronto, pp. 127–141.
- Canada, Government of 1982, "Constitution Act." Available at: <https://caid.ca/ConstAct010208.pdf>. (Accessed 26 August 2022).
- Canada, Government of 1988. "Canadian Multiculturalism Act." Available at: <https://laws-lois.justice.gc.ca/eng/acts/c-18.7/page-1.html>. (Accessed 26 August 2022).
- Canada, Government of June 3, 2019. *Reclaiming Power and Place: The Final Report of the National Inquiry into Missing and Murdered Indigenous Women and Girls*, vol. 1 and 2. Available at: <https://www.mniwg-fada.ca/>. (Accessed 4 July 2022).
- Canada, Government of 2022. Immigration and Citizenship. Available at: <https://www.canada.ca/en/services/immigration-citizenship.html>. (Accessed 26 August 2022).
- Canada, Royal Commission on Aboriginal Peoples (RCAP), 1996. *The Report of the Royal Commission on Aboriginal Peoples*, 5. Commission, Ottawa. Available at: <https://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/royal-commission-aboriginal-peoples/Pages/final-report.aspx>. (Accessed 1 February 2022).
- Canadian Japanese Internment Camps, 2011. *Facts and Figures*. Available at: <https://canadianjapaneseinternmentcamps.wordpress.com/2011/05/11/facts-figures/>.
- Canadian Race Relations Foundation (CRRF), 2022. "About Us". <https://www.crrf-rcr.ca/en/about/crrf-about> (Accessed 5 August 2022).
- Carastathis, A., 2016. *Intersectionality: Origins, Contestations, Horizons*. University of Nebraska Press, Lincoln and London.
- Clarke, G.E., 2006. Foreword. In: Cooper, A. (Ed.), *The Hanging of Angélique: The Untold Story of Canadian Slavery and the Burning of Old Montreal*. Harper Collins, Toronto, pp. xi–xviii.
- Claxton, N., Fong, D., Morrison, F., O'Bonsawin, C., Omatsu, M., Price, J., Kaur Sandhra, S., 2021. Challenging Racist "British Columbia": 150 Years and Counting. University of Victoria Research Project and Canadian Centre for Policy Alternatives (BC Office), British Columbia.
- Coloma, R.S., Pon, G., 2017. *Asian Canadian Studies Reader*. University of Toronto Press, Toronto.
- Coulthard, G.S., 2014. *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition*. University of Minnesota Press, Minneapolis.
- Crenshaw, K., 1991. Mapping the margins: identity politics, intersectionality, and violence against women. *Stanford Law Rev.* 43 (6), 1241–1299.
- Cooper, A., 2005. *The Hanging of Angélique: The Untold Story of Canadian Slavery and the Burning of Old Montreal*. Harper Collins, Toronto.
- Cunningham, F., 2007. The University and social justice. *J. Acad. Ethics* 5, 153–162.
- Day, R., 2000. *Multiculturalism and the History of Canadian Diversity*. University of Toronto Press, Toronto.
- Dhamoon, R., Abu-Laban, Y., 2009. *International Political Science Review*, 30 (2), 163–183.
- Dhamoon, R., Bhandar, D., Mawani, R., Kaur Bains, S., 2019. *Unmooring the Komagata Maru: Charting Colonial Trajectories*. UBC Press, Vancouver.
- Dei, G.J.S., Kempf, A., 2013. *New Perspectives on Africa-Centred Education in Canada*. Canadian Scholars' Press, Toronto.
- Dei, G.J.S., Mazza, J., McIsaac, E., 1997. *Reconstructing 'Drop-Out': A Critical Ethnography of the Dynamics of Black Students' Disengagement from School*. University of Toronto Press, Toronto.
- Dei, G.J.S., James, I.M., Karamanchery, L.L., Wilson, S.J., Zine, J., 2000. *Removing the Margins: The Challenges and Possibilities of Inclusive Schooling*. Canadian Scholars' Press, Toronto.
- Dei, G.J.S., Lewis, K., 2020. COVID-19 and the Black Experience: Keeping the Ember Burning. *Antipode: A Radical Journal of Geography*. Symposium on Black Humanity: Bearing Witness to COVID-19 (December 4). <https://antipodeonline.org/2020/12/04/black-humanity-bearing-witness-to-covid-19/>.
- Dellaire, J., Ellis, C., 2021. *Standing Strong Task Force Report and Recommendations*. Ryerson University, Toronto.
- Dhamoon, R., 2009. *Identity/Difference Politics: How Difference Is Produced, and Why it Matters*. UBC Press, Vancouver.
- Fernando, S., 2011. *Race and the City: Chinese Canadian and Chinese American Political Mobilization*. UBC Press, Vancouver.
- Franks, C.E.S., 1987. *The Parliament of Canada*. University of Toronto Press, Toronto.
- Gagnon, A., 2004. *Québec, State and Society*. Broadview Press, Peterborough.
- Galabuzi, G.-E., 2006. *Canada's Economic Apartheid: The Social Exclusion of Racialized Groups in the New Century*. Canadian Scholars' Press, Toronto.
- Gover, A.R., Harper, S.B., Lynn, L., 2020. Anti-Asian hate crime during the COVID-19 pandemic: exploring the reproduction of inequality. *Am. J. Crim. Justice* 45, 647–667.
- Green, J., 2001. Canaries in the mines of citizenship: Indian women in Canada. *Can. J. Polit. Sci.* 34 (4), 715–738.
- Haque, E., 2012. *Multiculturalism within a Bilingual Framework: Language, Race and Belonging in Canada*. University of Toronto Press, Toronto.
- He, B., Ziemis, C., Soni, S., Ramakrishnan, N., Yang, D., Kumar, S., 2021. Racism Is a Virus: Anti-Asian Hate and Counterspeech in Social Media during the COVID-19 Crisis" arXiv (Cornell University). <https://arxiv.org/abs/2005.12423>.
- Heath, J., 2002. Citizenship education and diversity. *Educ. Can.* 42 (3), 1–7.
- Henry, F., Tator, C., 2002. *Discourses of Domination: Racial Bias in the Canadian English-Language Press*. University of Toronto Press, Toronto.
- Hobsbawm, E., 1992. *The Invention of Tradition*. Cambridge University Press, Cambridge.
- International Organization for Migration (IOM), 2022. *World Migration Report 2022*. IOM, Geneva, Switzerland. Available at: <https://publications.iom.int/books/world-migration-report-2022>. (Accessed 2 February 2022).
- James, C., 2007. Negotiating school: marginalized students' participation in their education process. In: Johnson, G.F., Enomoto, R. (Eds.), *Race, Racialization and Antiracism in Canada and Beyond*. University of Toronto Press, Toronto, pp. 17–36.
- Johnson, G.F., Enomoto, R. (Eds.), 2007. *Race, Racialization and Antiracism in Canada and Beyond*. University of Toronto Press, Toronto.
- Kellogg, P., 2015. *Escape from the Staple Trap: Canadian Political Economy after Left Nationalism*. University of Toronto Press, Toronto.
- Kernerman, G., 2005. *Multicultural Nationalism: Civilizing Difference, Constituting Community*. UBC Press, Vancouver.
- Kymlicka, W., 1995. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Clarendon Press.
- Kymlicka, W., 2001. *Politics in the Vernacular: Nationalism, Multiculturalism, and Citizenship*. Oxford University Press, Oxford UK.
- Kymlicka, W., 2010. *The Current State of Multiculturalism in Canada and Research Themes on Multiculturalism 2008–2010*. Minister of Public Works and Government Services of Canada, Ottawa.
- Lawrence, B., 2004. "Real" Indians and Others: Mixed-Blood Urban Native Peoples and Indigenous Nationhood. UBC Press, Vancouver.
- Lischke, U., McNab, D.T. (Eds.), 2007. *The Long Journey of a Forgotten People: Studies in Métis Identities and Family Histories*. Wilfrid Laurier University Press, Waterloo.

- Little, M.J.H., 1998. *No Car, No Radio, No Liquor Permit: The Moral Regulation of Single Mothers in Ontario, 1920–1997*. Oxford University Press, Don Mills.
- Mackey, E., 2002. *The House of Difference: Cultural Politics and National Identity in Canada*. University of Toronto Press, Toronto.
- Malhotra, R., Rowe, M., 2014. *Exploring Disability, Identity and Disability Rights through Narratives: Finding a Voice of Their Own*. Routledge, Oxfordshire/New York.
- Maynard, R., 2017. *Policing Black Lives: State Violence in Canada from Slavery to the Present*. Fernwood, Halifax.
- McKittrick, K., 2006. *Demonic Grounds: Black Women and the Cartographies of Struggle*. University of Minnesota Press, Minneapolis.
- McKittrick, K., Woods, C. (Eds.), 2007. *Black Geographies and the Politics of Place. Between the Lines Press, Boston*.
- Miki, R., 2004. *Redress: Inside the Japanese Canadian Call for Justice Vancouver: Raincoast Books*.
- Monture-Angus, P., 1995. *Thunder in My Soul: A Mohawk Woman Speaks*. Fernwood Publishers, Halifax.
- Mujawamariya, D., 2007. Multicultural education: teacher candidates speak out. In: Johnson, G.F., Enomoto, R. (Eds.), *Race, Racialization and Antiracism in Canada and Beyond*. University of Toronto Press, Toronto, pp. 37–50.
- Palmer, P., 2021. National Action Plan and Federal Pathway Will Not End Genocide of Indigenous Women and Girls. PamPalmer.com/blog (June 3). <https://pampalmer.com/2021/06/national-action-plan-and-federal-pathway-will-not-end-genocide-of-indigenous-women-and-girls/>. (Accessed 4 July 2022).
- Pashby, K., 2014. Discovering, recovering, and covering-up Canada: tracing historical citizenship discourses in K-12 and adult immigration citizenship education. *Can. J. Educ.* 37 (2), 1–26.
- Pearson, C.D., 1988. Egerton Ryerson's Canadian liberalism. *Can. J. Polit. Sci.* 21 (4), 771–793.
- Picot, G., Hou, F., 2011a. Naturalisation: A Passport for Better Integration of Immigrants? *Citizenship Acquisition in Canada and the United States: Determinants and Economic Benefit*. OECD Publishing, Paris.
- Picot, G., Hou, F., 2011b. *Preparing for Success in Canada and the United States: The Determinants of Educational Attainment Among the Children of Immigrants*. Statistics Canada, Social Analysis Division, Ottawa.
- Prime Minister's Office, Canada, 2006. Press Release: Prime Minister Harper Offers Full Apology for the Chinese Head Tax. Available at: <https://www.canada.ca/en/news/archive/2006/06/prime-minister-harper-offers-full-apology-chinese-head-tax.html>. (Accessed 2 February 2022).
- Razack, S., 1998. *Looking White People in the Eye: Gender, Race and Culture in Courtrooms and Classrooms*. University of Toronto Press, Toronto.
- Razack, S., 2015. *Dying from Improvement: Inquests and Inquiries into Indigenous Deaths in Custody*. University of Toronto Press, Toronto.
- Razack, S., Smith, M., Thobani, S., 2010. *States of Race: Critical Race Feminism for the 21st Century*. Between the Lines, Toronto.
- Ryerson, S.B., 1975. *The Founding of Canada: Beginnings to 1815*. Progress Publishers, Toronto.
- Ryerson, S.B., 1983. *Unequal Union: Confederation and the Roots of Conflict in Canada*. Progress Publishers, Toronto.
- Said, E.W., 1993. *Culture and Imperialism*. Vintage, Random House, New York.
- Simpson, A., 2014. *Mohawk Interruptus: Political Life across the Borders of Settler States*. Duke University Press, Durham.
- Stasiulis, D.K., 1988. The symbolic mosaic reaffirmed: multiculturalism policy. In: Graham, K.A. (Ed.), *How Ottawa Spends: 1988–89*. Carleton University Press, Ottawa, pp. 81–112.
- Stasiulis, D.K., Bakan, A.B., 2005. *Negotiating Citizenship: Migrant Women in Canada and the Global System*. University of Toronto Press, Toronto.
- Stasiulis, D., Jhappan, R., 1995. The fractious politics of a settler society: Canada. In: Stasiulis, D., Yival-Davis, N. (Eds.), *Unsettling Settler Societies: Articulations of Gender, Race, Ethnicity and Class*. Sage, London, pp. 95–131.
- Taylor, C., 2004. *Modern Social Imaginaries*. Duke University Press, Durham and London.
- Taylor, C., 2009. *A Secular Age*. Harvard University Press, Cambridge Mass, USA.
- Tennyson, B.D., 1990. Canada and the Commonwealth Caribbean. In: Tennyson, B.D. (Ed.), *Canadian-Caribbean Relations: Aspects of a Relationship*. Centre for International Studies, Cape Breton, pp. 1–57.
- Thobani, S., 2002. War Frenzy. *Atlantis A Women's Stud. J.* 27 (1), 5–11.
- Thobani, S., 2007. *Exalted Subjects: Studies in the Making of Race and Nation in Canada*. University of Toronto Press, Toronto.
- Titchkosky, T., 2003. Governing embodiment: technologies of constituting citizens with disabilities. *Can. J. Sociol.* 28 (4), 517–542.
- Tully, J., 1995. *Strange Multiplicity: Constitutionalism in an Age of Diversity (The Seeley Lectures)*. Cambridge University Press, Cambridge.
- Truth and Reconciliation Commission of Canada (TRC), 2015. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. Available at: <https://www.rcaanc-cimac.gc.ca/eng/1450124405592/1529106060525#chp2>. (Accessed 2 February 2022).
- United Nations (UN) General Assembly Human Rights Council, September 11–29, 2017. *Report of the Working Group of Experts on People of African Descent on its Mission to Canada*.
- Walker, B., 2010. *Race on Trial: Black Defendants in Ontario's Criminal Courts*. University of Toronto Press and Osgoode Society for Canadian Legal History, Toronto.
- Winks, R., 1997. *A History of Blacks in Canada*. McGill-Queen's University Press, Montreal.
- World Health Organization (WHO), March 11, 2020. WHO Director-General' Opening Remarks at the Media Briefing on COVID-19. Available at: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>. (Accessed 2 February 2022).
- Wolfe, P., 2016. *Traces of History: Elementary Structures of Race*. Verso, London and New York.

Curriculum, citizenship, migration

Yoonjung Choi, Ewha Womans University, Seoul, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Introduction	46
Theoretical discourses on global citizenship	47
Curricular practices of global citizenship education	48
Moving forward during and after COVID-19	50
Conclusion	50
References	51

Introduction

The world has observed the expanding scale and pace of international migration for the past decades, although this is not phenomenon unique to this time. As of 2019, the total number of international migrants is 271.6 million, which accounts for 3.5% of the world's population ([International Organization for Migration, 2019](#)). The number of international migrants has increased approximately by 50 million every ten years in the 21st century—173.6 million in 2000, 220.8 million in 2010, and 271.6 million in 2020 ([United Nations, Department of Economic and Social Affairs, Population Division, 2019](#)).

Migration has become a major indicator of the ever-increasing interdependencies and complexities of global society. The expansion of global trade, integration of labor markets, and proliferation of supranational governing organizations collectively signaled the acceleration of large-scale migration along with the diminishing power of individual nation-states and increasingly permeable borders. Given that migrants have often been subjected to the unwarranted accusation that they are a source of social problems and a threat to national security, their historical and contemporary contributions to countries of origin as well as countries of destination have long been overlooked. Migration has been proven to play a significant role in economic growth, technological innovation, social dynamism, and cultural diversity across the world ([International Organization for Migration, 2019](#)).

In light of the increasing migration flow worldwide, a sense of urgency about reconceptualizing what it means to be a citizen, and how to implement citizenship education has grown. Ideas regarding citizenship and citizenship education, “once solely and firmly the domain of national citizens,” have come to transcend national boundaries, being “reconstituted around a constellation of multiple affiliations including race·ethnicity, gender, place, and class to name a few” ([Gaudelli, 2009](#), p. 69). Assimilationist practices rooted in the dominant form of citizenship education in the 20th century requesting migrants to abandon their ethnic, linguistic, and cultural characteristics for the sake of unity and social cohesion of the host society, and resulting in the marginalization and discrimination of migrant learners in school turned out to be dysfunctional ([Banks, 2017](#); [Coloma, 2017](#); [Suárez-Orozco, 2019](#)). Nation-centric approach to citizenship education has been warned to be harmful to domestic-born mainstream learners as well who might lose significant opportunities to explore some of the most pressing issues at home and abroad and thus fail to develop into informed, competent, and caring global citizens ([Patterson and Choi, 2018](#); [Subedi, 2013](#)).

Global citizenship education has evolved as a discursive space, curricular approach, and practical suggestion for responding to the renewed educational agenda in the era of globalization ([Gaudelli, 2016](#)). From scholars and educators to governments and corporations, many diverse educational stakeholders have discussed the responsibilities of schools and the roles of school curriculum in preparing the next generation to survive and thrive as global citizens. Mindful that school curriculum is a manifestation of the expectations for its citizens and a carrier of official knowledge in a society, (re)orienting what it should look like and how it should work in the critical time of globalization has become a significant task in many countries ([Apple, 2014](#); [Ross, 2014](#)). Despite decades-long scholarly discussions, multi-level policies, and onsite practices, the ultimate question of global citizenship education “whether or not it should be taught and if so, then how” has remained unanswered ([Goren and Yemini, 2017](#), p. 170). The term “global citizenship education” has been utilized with different political, economic, sociocultural aims of multiple stakeholders, which in turn allowed confusion, skepticism, tokenism, and/or disorientation of global citizenship education in practice ([Patterson and Choi, 2018](#)).

The COVID-19 outbreak which plunged the world into fear and confusion added another layer of uncertainty and complexity to this issue. Unlike prior predictions of ever accelerating globalization and increased migration flow, the unexpected worldwide pandemic resulted in border closures, local lockdowns, travel restrictions, and quarantine requirements across the globe. A historical low for international migration was observed in 2020, while the expansion of human mobility has been put on hold ([OECD, 2020](#)). Discourses on migration, citizenship, and school curriculum, collectively as well as respectively, presuppose that human beings live as members of society through relationships with others. At a time when schooling has been disrupted and distancing has become the new norm, educators’ role in rethinking longstanding approaches to citizenship education and constructing more resilient, sustainable communities from the local to the global level have become more critical than ever before.

This chapter explores the intersectionalities of migration, citizenship, and curriculum, placing global citizenship education at the center of its discussion. Through a review of literature on global citizenship education published in the past decade, this article examines the following inquiries: From a theoretical standpoint, what scholarly approaches have emerged in the discourse of global citizenship education? From an empirical view, have official curriculum in different parts of the world portrayed global citizenship, and if so how? Given the COVID-19 pandemic causing a dramatic change and challenge in the realities of migration and globalization in the recent years, uncertainties, complexities, and controversies related to discourses on citizenship education and curriculum are discussed as thoughts for future research.

Theoretical discourses on global citizenship

The growing scale and pace of global migration is a key indicator demonstrating the world's heightening interdependencies, complexities, and transformations. The diversified facets of globalization including globally integrated labor markets, innovative technologies in media and communication, and convenient and affordable cross-border transportation have widely led the escalation of migration across the globe (Suárez-Orozco and Suárez-Orozco, 2010). People living outside their country of birth reached approximately 272 million as of 2019, a number close to the national population of Indonesia (International Organization for Migration, 2019). 38 million of these are international migrants below the age of 20 (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Further attention must be given to the fact that the number of refugees and asylum seekers has continuously risen (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Unresolved issues from local to global including, but not limited to, political insecurity, socioeconomic inequalities, human rights violations, and environmental degradation, have repeatedly caused the forcible displacements of people.

Scholarly discourses on citizenship education have considerably been diverse, convoluted, and contentious. Given the complex, fluid dimensions of globalization and migration, there seems to be a general consensus on the need to expand traditional notions of citizenship beyond national boundaries and to prepare the next generation to live as a global citizenry (Banks et al., 2016; Hahn, 2020). Global citizenship education emerged as a noble idea, worthy goal, and agile response to the urgent agenda of reformulating citizenship education, drawing a wide range of interest and popularity among the general public, scholars, and policy makers in places around the world (Cho and Mosselson, 2017; Yemini et al., 2018). Yet, what an ideal global citizen should look like, how global citizenship can be defined, and what how to whom schools should teach for it have been envisioned differently by a range of educational stakeholders with diverging ontological, epistemological, and ideological stances (Oxley and Morris, 2013). A nationalist hegemony in company with free-market values emphasizes nurturing "productive, self-employable, and competent beings" (Choi and Kim, 2020, p. 2) ready for competition in the globally integrated market as well as framing global citizenship as an extension of national citizenship. On the other hand, critiques call for a systematic reexamination of global hegemony and its associated inequity, oppression, and dehumanization (Banks et al., 2016; Dill, 2013; Evans, 2015; Giroux, 2013). A seemingly apolitical, controversy-free conceptualization of global citizenship education by UNESCO, "building the knowledge, skills, values and attitudes that learners need to be able to contribute to a more inclusive, just and peaceful world" (UNESCO, 2015, p. 15), has been widely cited in scholarly publications, policy documents, and curriculum standards. Yet the term "global" itself has long been a target of criticism by conservatives as a radical, value-laden concept that uproots national sovereignty and state authority (Lourenço, 2018). Scholars and educators have eagerly sought to unpack the multiple meanings, purposes, conceptions, and approaches of and surrounding global citizenship education, drawing upon alternative as well as complementary lenses in efforts to propose more nuanced theoretical frameworks and contextualized curricular models.

This article outlines some of the most influential, widely-used typologies of global citizenship education. In what follows, Andreotti's (2006) soft vs. critical framework, Gaudelli's (2009) heuristics of global citizenship, and Oxley and Morris's (2013) typology identifying eight conceptions of global citizenship education are introduced in sequential order. These broadly recognized, highly-cited works offer a useful overview of diverse theories and curricular approaches shaping the contemporary discussions and practices of citizenship education in the age of globalization and migration.

Andreotti (2006) identifies a twofold framework of soft vs. critical global citizenship education, contrasting opposing views, ideals, and approaches in the field. Soft global citizenship education, currently the dominant form in practice, seeks to raise awareness of the world and cultural diversity at large. It suggests economic progressivism as a solution to the problems in question, while referring to the individual level of moral responsibility and humanity as a ground for change (pp. 46–47). Critical global citizenship education advocates empowering learners to reflectively analyze the Western, colonial legacies of global power relations and intellectual production, and to participate in informed, responsible, and ethical actions for problem solving (pp. 46–47).

Gaudelli (2009) provides a heuristics of global citizenship, identifying five distinctive approaches that shape contemporary scholarly discussions and curricular implementations. These five approaches center around national, neoliberal, Marxist, world justice and governance, and cosmopolitan conceptions. According to this model, the national (building on inviolable sovereignty of national governments and promoting a sense of shared national belonging) and neoliberal (rooted in nationalism yet framing people as transformative agents who can make the most of the global economy, thus emphasizing utility-based knowledge) approaches have been hegemonic forms of citizenship discourses (pp. 69–72). Meanwhile, the Marxist (advocating empowerment of the oppressed as a resolution to the exploitative nature of globalization and featuring critical pedagogy as its method), world justice and governance (prioritizing universal humanity and peace by means of codified human rights and institutionalized civil

actions), and cosmopolitan (seeking further imaginary, dialogic, and esthetic modes of discussion of civil identities, cultural pluralism, and moral values) views, though being marginal in practice, have emerged as meaningful counter-narratives to the former ones (pp. 73–77).

Oxley and Morris (2013) propose a comprehensive typology of eight conceptions of global citizenship in an effort to map out disparate theoretical underpinnings in the field. This typology features two broad types of citizenship ideals, each with four distinctive conceptions. The cosmopolitan type, often interpreted as a neo-imperial form of global citizenship, includes the following: political (envisioning a well-ordered world through the strengthening and democratizing of supranational institutions), moral (promoting global ethic on the basis of universal humanitarianism), economic (focusing on international development in favor of capitalism, neoliberalism, and/or corporate philanthropic activities), and cultural (celebrating multiculturalism and pursuing the globalization of arts, media, languages, sciences, and technologies) conceptions (pp. 305–311). The advocacy type refers to a more holistic, anti-individualistic views, which incorporates social (being attentive to the intersectionalities of multicultural identities and marginalized voices), critical (calling for action to challenge prevailing inequalities, discriminations, and oppressions around the world), environmental (supporting sustainable development based on human responsibility toward nature), and spiritual (finding meanings in the non-scientific, immeasurable aspects of human relations; namely, love, care, and emotions) conceptions (pp. 311–315).

Such scholarly endeavors contributed to enrich our understanding of the global citizenship education theories that have emerged in the context of increased migration and globalization. These frameworks, synthesizing and visualizing multifariously intricate arguments and imaginaries in the discursive space of global citizenship education, have been applied extensively as useful analytic lenses into educational policies and curricular practices. They have also served as guiding principles for many subsequent models in the field (Goren and Yemini, 2017). Complications within and across these multiple models are evident, yet interpretable as encouragement, instead of trouble, aiming toward more open debates and inclusive conversations in the scholarship.

Although drawing upon diverse methodological approaches and utilizing different methods of categorization, these aforementioned typologies in common feature the limits and dangers of neoliberal rhetoric and nationalist visions of citizenship. Scholars have warned that such “subtractive” (Valenzuela and Brewer, 2010) notions not only fail to reflect the realities of an integrated world and people, but also limit the opportunities of all learners to thrive as informed, competent, and caring participants of the globe (Banks, 2016). A self-reflective criticism of the scholarship itself has also been identified; namely, that it is geared toward a Western, Northern, liberal democratic subject as the desirable norm of global citizenship given the overrepresentation of existing literature from English-speaking, Western countries (Pashby et al., 2020). Subsequent theoretical research has tended to put emphasis on the need for (post-)critical approaches and post-colonial imaginaries toward global citizenship education in practice (Andreotti, 2014; Stein, 2015).

Curricular practices of global citizenship education

In conjunction with the expansion of scholarly discussions, the ways in which global citizenship education is articulated through school curriculum has received a growing empirical interest. A large volume of prior research on such inquiry has dominantly been conducted in Western societies that have been traditionally desirable destinations for global migrants; namely, Europe and North America. In recent decades, global migration patterns have become more complex with a broader regional scope, while the division between sending and hosting countries has collapsed (International Organization for Migration, 2019). As the citizenship status and self-defined identities of migrants have become more diverse and fluid, growing attentions to school curriculum for global citizenship have been observable in a wider range of regions around the world (Alviar-Martin and Baidon, 2016; Goren and Yemini, 2016; Koh, 2017; Pan, 2011; Patterson and Choi, 2018; Suárez, 2008; Quaynor, 2015). The recent scholarship has largely advocated localized, contextualized global citizenship curriculum design and implementation taking into account diverse historical backgrounds, political systems, religious beliefs, economic interests, geopolitical locations, public sentiments each society has been situated within. Meanwhile, concerted international efforts toward curriculum design to ensure shared goals of peace, equality, cultural diversity, and sustainable development of the globe has been emphasized by supranational institutions (UNESCO, 2020). Topics addressed in research have become more diverse as well. Whether and to what extent global citizenship related knowledge, disposition, skills, values and attitudes are portrayed in formal school curricula has been a significant line of inquiry (Choi and Kim, 2020; Goren and Yemini, 2017; Ideland and Malmberg, 2014). In what ways migration is discoursed in the curriculum and such curriculum (dis)serves the needs and interests of migrants has been an issue of particular interest in the context of proliferating human migration (Quaynor and Murillo, 2018). A burgeoning number of empirical studies focusing on the enacted dimension of curricula, have investigated the ways in which global citizenship curriculum is experienced in the everyday realities of school (Goren and Yemini, 2016; Kim, 2019; Hahn, 2020; Rappoport, 2010; Rios-Rojas, 2018; Verma, 2017).

This section presents a sketch of literature dealing with the place of global citizenship education in the formal school curriculum around the world. Despite the significance of understanding the enacted aspects of curriculum implementation, this article sheds light on the official curriculum (e.g., policy documents, local-regional-national curriculum standards, and textbooks) to make better sense of how global citizenship education is intended and narrated in the prescriptive curriculum that mirrors the official knowledge and desired civic virtue of a given society (Apple, 2014). Drawing upon some of the most recent empirical research on the formal curriculum practices of global citizenship education around the world, this section discusses the current research findings in the field.

The ways in which the formal curriculum (dis)employs and (dis)articulates global citizenship have been actively investigated in research across the world. A significant number of studies has been published in the contexts of and by scholars from the so called 'first world', Western, English-speaking, democratic societies (i.e., the US and Canada), which have typically been the most popular destinations for global migrants. These studies in large have explored the portrayal of global citizenship and the treatment of non-Western regions and people in school curricula (Rapoport, 2009; Swanson and Pashby, 2016; Subedi, 2013). In the recent decade, various curricular approaches to global citizenship education in Western societies with diverse linguistic, political, geographical, and/or academic backgrounds for example, Australia and New Zealand in Oceania (Peterson et al., 2018) and European countries including Finland (Andreotti et al., 2015), Sweden (Ideland and Malmberg, 2014), the Netherlands (Weiner, 2018), Poland (Leek, 2016), Austria, Bulgaria, the Czech Republic, Ireland, Italy, Latvia, Spain, and Portugal (Tarozzi and Inguaggiato, 2018) have been reported. Diverse cases from and between other regions including South Korea (Choi and Kim, 2020), Singapore, and Hong Kong in Asia (Alviar-Martin and Baidon, 2016); Ghana in Africa (Arnot et al., 2018); and Argentina, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, Nicaragua, Peru, and Venezuela in Latin America (Poorter and Aguilar-Forero, 2020; Sant and González Valencia, 2018) have also been examined. Collectively, these examinations have enriched our understanding of how nations around the world shape global citizenship discourses via the school curriculum. Such line of inquiry has been of particular interest in countries where government-led curricular initiatives with high levels of prescription strongly impact everyday teaching and learning practices in schools.

A growing body of empirical evidence from the previous literature has argued that the position of global citizenship in formal school curricula around the world has been ambiguous, fragile, and troubling. Although there have been widely-shared commitments to global citizenship education, official school curricula in many developed nations have reportedly been riddled with simplistic, superficial portrayals of the world; apolitical approaches to highly controversial global issues; and a subtractive, instead of inclusive, ethos of citizenship with strong nationalist and neoliberal agendas (Goren and Yemini, 2017; Peterson et al., 2018). National exceptionalism along with dichotomous distinctions between tradition vs. civilization, dirtiness vs. purity, chaos vs. order, ignorance vs. morality, and helped vs. helping seems still apparent, while stories of contemporary migration as part of global and national citizenship have been largely excluded (Ideland and Malmberg, 2014; Quaynor and Murillo, 2018; Weiner, 2018). When these narratives are included, discussions of migration tend to be couched in historical and abstract terms, celebrating contributions of migrants who have seamlessly assimilated themselves into host societies (Patterson and Choi, 2018, p. 488). Concerned scholars warn that the current form of curricula is likely to inhibit opportunities for critical thinking and reflective actions, "dissuading students' analysis of their own complicity in global injustices and the moral dimensions of living in an interrelated world" (Patterson and Choi, 2016, p. 170).

The distinctive issues and concerns of the abovementioned lesser-studied regions deserve further attention. For instance, studies on South Korea, Singapore, and Hong Kong have identified a shared concern regarding the limitations of the state-approved citizenship education curriculum, which employs a narrow sense of nationalism and relies on an unquestioned utilitarian agenda (Alviar-Martin and Baidon, 2016; Choi and Kim, 2020). For these societies that share a common experience of colonial rule, that have adopted a Western model of development, and that have grown into economically affluent cosmopolitan cities, focusing their curricula on building a unitary national identity and nurturing citizens with marketable knowledge and skills in order to compete in the global age could be part of a strategic survival method, contextualized decision-making, and/or a consequence of the colonial past (Patterson and Choi, 2018).

The conflicting ethos between Western notions of the competitive, individualized, active citizen and community-oriented civic identities and virtues based on indigenous cultures appear at the heart of the postcolonial dilemmas in Ghana and other places in Africa (Arnot et al., 2018). "The mixing of national unity and patriotism with the principles of children's rights and equity and with neo-liberal agendas around individualization, enterprise and competing environmental and development goals" (p. 123) has been an uneasy task that the Ghanaian school curriculum has juggled with.

In Latin America, curricular approaches to global citizenship education have been ideologically and geopolitically divided. The concept of global citizenship is only explicitly discussed in the educational policies of some pro-globalization and market-oriented countries including Chile, Colombia, and Peru. Other countries such as Cuba and Brazil with their proclaimed notions of anti-imperialism have challenged the dominance of Western civilization discourses and the prominent understanding of global citizenship conceptualized by supranational authorities founded on imperial legacies. Instead, these countries have advocated the critical examination of the global power relations, reflecting regional voices and demands as alternatives (Sant and González Valencia, 2018).

The influence and enforcement of national policies, textbooks, and curriculum standards may vary from country to country. Yet, the limitations, dilemmas, and challenges evident in the official curriculum for global citizenship are most likely to impact how the curriculum is enacted in the realities of school. This is particularly true given the current emphasis on the importance of the accountability system and high-stakes testing (Evans, 2015). Reflecting on how the wars, imperialism, and colonial history of the 20th century affect global citizenship education and its associated curricular practices in the 21st century and searching for the commonalities and differences within and between regions having diverse historical, political, economic, and sociocultural contexts through a comparative lens can offer meaningful insights and lessons for educators envisioning how the curriculum should move forward.

Moving forward during and after COVID-19

Citizenship, curriculum, and migration, together and apart, are context-sensitive subjects that reflect both the specificities of locals and complexities of global circumstances. In anticipation of deepening globalization and expanding cross-border migration, continuous engagements in research on the fluid concept of global citizenship education along with praxis-oriented efforts to implement contextually meaningful, nuanced curricula for diverse learners have strongly been called for. The COVID-19 pandemic brought unexpected challenges, further uncertainties, and complex questions to the already confusing, divergent field of citizenship education. Throughout and since the COVID-19 outbreak, the world has experienced the most severe restrictions on and the most rapid decline in global human mobility in modern times (Chamie, 2020). As of 2020, 120 countries have reportedly ordered border shut-downs with travel restrictions on foreign arrivals (Gamlen, 2020). Anti-migrant, xenophobic sentiments appeared to intensify while migrants were accused of accelerating the spread of the infectious disease and further burdening the shrinking economy (OECD, 2020). If and how the pandemic signifies “a temporary trough in global population movements, or the end of the most recent age of migration” (Gamlen, 2020, p. 2) is a pressing question that has been receiving close attention from academia as well as the general public.

World-wide school closures as a consequence of the pandemic have reportedly yielded inequities in public education and constricted the meaning of learners as social beings and learning as a social interaction. As of 2020, over 190 countries closed all of their primary and secondary schools, though in varying periods of time, affecting 1.5 billion students and 63 million teachers (Bryant et al., 2020, April 21; Waters, 2020). Physical distancing within and outside the classroom was encouraged on days when students attended school in person. It has not been unusual that students wearing masks sit at desks with partitions as far away as possible from other classmates, and/or engaging in individual study without social contact. The pandemic has caused much more than simply the discontinuity of public schooling and the expansion of online classes. Self-directed learning emerged as trendy jargon given the expansion of remote learning, yet too often resulted in the learning loss, achievement gaps, and social isolation of already underserved students in vulnerable situations including, students of color, migrant youth, linguistic minorities, and children in lower-income families (Dorn et al., 2020 December 8; Sonnemann and Goss, 2020).

Questions, concerns, and possibilities regarding the direction of citizenship education and the roles of curriculum have risen amid the greater uncertainties and weaker sense of belonging produced by COVID-19. At a time when the fragility of linear progress and modernity has been greatly tested, will global citizenship education within and across nations remain consistent with strong positivistic, colonial assumptions? Or will it shift toward a more critical examination of human development? While borders are closed and migrants' movement is restricted under the authority of individual states, how should the ideas of citizen, citizenship, and citizenship education be revisited? Are school curricula in many countries around the globe going to fortify their existing nationalistic, patriotic focus, or will they become more inclusive and globally-oriented, emphasizing cosmopolitanism to overcome this global crisis together? In what ways will COVID-19 be taught as a global issue in post-pandemic era classrooms? How much will the treatment of and response to this issue vary in school curricula from region to region? If and how will the portrayal of migration and migrants, which have largely been marginalized or instrumentally used if included in the curricula, be transformed? Will COVID-19 fuel the proliferation of online-based, AI-assisted remote learning? Or will it restore the longstanding values of community building, social interaction, and collaborative learning in citizenship education classrooms? In the post pandemic era, will global citizenship education become an exclusive domain of affluent learners with access to advanced communication technologies and able to travel internationally? Or will it be more accessible to individuals from underprivileged regions, providing a wider range of information and greater exposure to the world?

The roles and responsibilities of scholars, policy makers, and practitioners to continuously reflect on the past, question the present, and imagine the future of global citizenship education appear more urgent than ever before in order to responsibly cope with this unforeseen global crisis and to carefully prepare for the post-pandemic era.

Conclusion

This chapter concludes with a call for educators' responsible participation in the discursive spaces of global citizenship education and its curricular implementations. In the body of existing literature on the nexus between migration, citizenship education, and curriculum, there seems to be a disparity between a critical, post-colonial focus in its theoretical arguments and, on the other hand, the nationalistic, neoliberal rhetoric dominant in the realities of curricular practice (Goren and Yemini, 2017). Citizenship education and school curricula displaying their commitment to global citizenship in many parts of the world have reportedly exhibited an overriding apolitical approach to the most controversial global issues, negligence in terms of the global citizenry's civic obligations by emphasizing national demands, and a subtractive lens with either-or distinctions limiting the voices of the marginalized. Scholars have criticized the failure of governments, policies, and official curricula to provide more democratic, justice-oriented visions of citizenship education for schools, while also searching for context-specific explanations and alternatives. This article invites committed educators at all levels to persistently question and revisit global citizenship education to ensure more humane, democratic, and ethical implementations for all learners. Moving forward, diversified curricular imaginations and pedagogical experimentations of what is possible and how it can be done from various ontological, epistemological, theoretical, and methodological lenses will be as significant as investigating the failures and complexity of curricula.

Globalizing scholarship on global citizenship education itself is also strongly encouraged. Mindful of the continuity of modern, colonial assumptions and strong influences of the “first world” in the scholarship which can be “inherently violent and unsustainable” (Pashby et al., 2020, p. 160), voices and stories from diverse regions and people that have traditionally been lesser-known in particular, are welcomed to be highlighted. It is imperative for scholarship on global citizenship education to offer more open, inclusive, and cosmopolitan discursive spaces that can enable in-depth explorations of how we teach and learn beyond Western-centric assumptions and colonial legacies. Comparative analyses on the intersections between these explorations can provide valuable insights regarding the concerted efforts that can be made by responsible educators around the world to foreground justice, equity, and, diversity in global citizenship education. Such endeavors will be even more meaningful during and after the era of deepening global instabilities and uncertainties caused by COVID-19, when the importance of international cooperation and solidarity, the protection of universal human rights and safety, and the construction of a healthy and sustainable global society are more critical than ever before.

References

- Aviari-Martin, T., Baidon, M., 2016. Context and curriculum in two global cities: a study of discourses of citizenship in Hong Kong and Singapore. *Educ. Pol. Anal. Arch.* 24 (58), 1–27.
- Andreotti, V., Biesta, V., Ahenakew, C., 2015. Between the nation and the globe: education for global mindedness in Finland. *Glob. Soc. Educ.* 13 (2), 246–259.
- Andreotti, V., 2006. Soft versus critical global citizenship education. *Policy Pract.* 3 (Autumn), 40–51.
- Andreotti, V., 2014. Critical and transnational literacies in international development and global. *Sisyphus* 2 (3), 32–50.
- Apple, M.W., 2014. *Official Knowledge: Democratic Knowledge in a Conservative Age*, third ed. Routledge, New York.
- Arnot, M., Casely-Hayford, L., Yeboah, T., 2018. Post-colonial dilemmas in the construction of Ghanaian citizenship education: national unity, human rights and social inequalities. *Int. J. Educ. Dev.* 61 (C), 117–126.
- Banks, J.A., Suárez-Orozco, M.M., Ben-Perez, M. (Eds.), 2016. *Global Migration, Diversity, and Civic Education: Improving Policy and Practice*. Teachers College Press, New York.
- Banks, J.A., 2016. Civic education in the age of global migration. In: Banks, J.A., Suarez-Orozco, M.M., Ben-Peretz, M. (Eds.), *Global Migration, Diversity, and Civic Education: Improving Policy and Practice*. Teachers College Press, New York, pp. 29–52.
- Banks, J.A. (Ed.), 2017. *Citizenship Education and Global Migration: Implications for Theory, Research, and Teaching*. The American Educational Research Association, Washington, DC.
- Bryant, J., Chen, L., Dorn, E., Hall, S., 2020. School-System Priorities in the Age of Coronavirus. Retrieved from: <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/school-system-priorities-in-the-age-of-coronavirus>.
- Chamie, J., 2020. International migration amid a world in crisis. *J. Migration Human Secur.* 8 (3), 230–245.
- Cho, H.S., Mosselson, J., 2017. Neoliberal practices amidst social justice orientations: global citizenship education in South Korea. *Compare* 48 (6), 1–18.
- Choi, Y., Kim, Y., 2020. Deconstructing neoliberalism in global citizenship discourses: an analysis of Korean social studies textbooks. *Crit. Stud. Educ.* 61 (4), 464–479.
- Coloma, R.S., 2017. “We are here because you were there”: on curriculum, empire, and global migration. *Curric. Inq.* 47 (1), 92–102.
- Dill, J.S., 2013. *The Longings and Limits of Global Citizenship Education: The Moral Pedagogy of Schooling in a Cosmopolitan Age*. Routledge, New York.
- Dorn, E., Hancock, B., Sarakatsannis, J., Viruleg, E., 2020. COVID-19 and Learning Loss—Disparities Grow and Students Need Help. Retrieved from: <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help>.
- Evans, R.W., 2015. *Schooling Corporate Citizens: How Accountability Reform Has Damaged Civic Education and Undermined Democracy*. Routledge, New York.
- Gamlen, A., 2020. Migration and Mobility After the 2020 Pandemic: The End of an Age? IOM Publications. Retrieved from: <https://publications.iom.int/system/files/pdf/migration-and-mobility.pdf>.
- Gaudelli, W., 2009. Heuristics of global citizenship discourses towards curriculum enhancement. *J. Curriculum Theor.* 25 (1), 68–85.
- Gaudelli, W., 2016. *Global Citizenship Education: Everyday Transcendence*. Routledge, New York.
- Giroux, H.A., 2013. Can democratic education survive in a neoliberal society? In: Reitz, C. (Ed.), *Crisis and Commonwealth*. Lexington Book, Plymouth, pp. 137–152.
- Goren, H., Yemini, M., 2016. Global citizenship education in context: teacher perceptions at an international school and a local Israeli school. *Compare* 46 (5), 832–853.
- Goren, H., Yemini, M., 2017. Global citizenship education redefined—a systematic review of empirical studies on global citizenship education. *Int. J. Educ. Res.* 82, 170–183.
- Hahn, C.L., 2020. Educating citizens in an age of globalization, migration, and transnationalism: a study in four European democracies. *Theor. Res. Soc. Educ.* 48 (2), 244–284.
- Ideland, M., Malmberg, C., 2014. “Our common world” belongs to “us”: constructions of otherness in education for sustainable development. *Crit. Stud. Educ.* 55 (3), 369–386.
- International Organization for Migration, 2019. *World Migration Report 2020*. International Organization for Migration, Geneva.
- Kim, Y., 2019. Global citizenship education in South Korea: ideologies, inequalities, and teacher voices. *Glob. Soc. Educ.* 17 (2), 177–193.
- Koh, S.Y., 2017. *Race, Education and Citizenship: Mobile Malaysians, British Colonial Legacies, and a Culture of Migration*. Palgrave Macmillan, New York.
- Leek, J., 2016. Global citizenship education in school curricula: a Polish perspective. *J. Soc. Stud. Educ. Res.* 7 (2), 51–74.
- Lourenço, M., 2018. Global, international and intercultural education: three contemporary approaches to teaching and learning. *Horizon* 26 (2), 61–71.
- OECD, 2020. What is the Impact of the COVID-19 Pandemic on Immigrants and Their Children? OECD Policy Responses to Coronavirus (COVID-19). Retrieved from: https://read.oecd-ilibrary.org/view/?ref=137_137245-8saheqv0k3&title=What-is-the-impact-of-the-COVID-19-pandemic-on-immigrants-and-their-children%3F.
- Oxley, L., Morris, P., 2013. Global citizenship: a typology for distinguishing its multiple conceptions. *Br. J. Educ. Stud.* 61 (3), 301–325.
- Pan, S., 2011. Multileveled citizenship and citizenship education: experiences of students in China’s Beijing. *Citizen. Stud.* 15 (2), 283–306.
- Pashby, K., Costa, M., Stein, S., Andreotti, V., 2020. A meta-review of typologies of global citizenship education. *Comp. Educ.* 56 (2), 144–164.
- Patterson, T., Choi, Y., 2016. Global citizenship education: new pathways in practice. *J. Res. Curriculum Instr.* 20 (3), 169–172.
- Patterson, T., Choi, Y., 2018. Global citizenship, migration and national curriculum: a tale of two nations. *Br. J. Educ. Stud.* 66 (4), 477–496.
- Peterson, A., Milligan, A., Wood, B.D., 2018. Global citizenship education in Australasia. In: Davies, I., Ho, L., Kiwan, D., Peck, C.L., Peterson, A., Sant, E., Waghid, Y. (Eds.), *The Palgrave Handbook of Global Citizenship and Education*. Palgrave Macmillan, London, pp. 3–20.
- Poorter, J.D., Aguilar-Forero, N., 2020. The emergence of global citizenship education in Colombia: lessons learned from existing education policy. *Compare* 50 (6), 865–883.
- Quaynor, L., Murillo, A., 2018. Migration and implications for global citizenship education: tensions and perspectives. In: Davies, I., Ho, L., Kiwan, D., Peck, C.L., Peterson, A., Sant, E., Waghid, Y. (Eds.), *The Palgrave Handbook of Global Citizenship and Education*. Palgrave Macmillan, London, pp. 425–438.
- Quaynor, L., 2015. “I do not have the means to speak”: educating youth for citizenship in post-conflict Liberia. *J. Peace Educ.* 12 (1), 15–36.
- Rapoport, A., 2009. A forgotten concept: global citizenship education and state social studies standards. *J. Soc. Stud. Res.* 33 (1), 91.
- Rapoport, A., 2010. “We cannot teach what we don’t know”: Indiana teachers talk about global citizenship education. *Educ. Citizen Soc. Justice* 5 (3), 179–190.
- Rios-Rojas, A., 2018. Conditional citizens, suspect subjects: producing “illegality” and policing citizens in a citizenship education classroom in Spain. *Curric. Inq.* 48 (1), 70–94.
- Ross, E.W. (Ed.), 2014. *The Social Studies Curriculum: Purposes, Problems, and Possibilities*, fourth ed. SUNY Press, New York.

- Sant, E., González Valencia, G., 2018. Global citizenship education in Latin America. In: Davies, I., Ho, L., Kiwan, D., Peck, C.L., Peterson, A., Sant, E., Waghid, Y. (Eds.), *The Palgrave Handbook of Global Citizenship and Education*. Palgrave Macmillan, London, pp. 67–82.
- Sonnemann, J., Goss, P., 2020. COVID Catch-Up: Helping Disadvantaged Students Close the Equity Gap. Grattan Institute, Victoria.
- Stein, S., 2015. Mapping global citizenship. *J. Coll. Character* 16 (4), 242–252.
- Suárez, D.F., 2008. Rewriting citizenship? Civic education in Costa Rica and Argentina. *Comp. Educ.* 44, 485–503.
- Suárez-Orozco, M.M., Suárez-Orozco, C., 2010. Children of migrant populations. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, 4 vols. Elsevier, Oxford, pp. 629–635.
- Suárez-Orozco, M. (Ed.), 2019. *Humanitarianism and Mass Migration: Confronting the World Crisis*. University of California Press, Oakland.
- Subedi, B., 2013. Decolonizing the curriculum for global perspectives. *Educ. Theor.* 63 (6), 621–638.
- Swanson, D.M., Pashby, K., 2016. Towards a critical global citizenship? A comparative analysis of GC education discourses in Scotland and Alberta. *J. Res. Curriculum Instr.* 20 (3), 184–195.
- Tarozzi, M., Inguaggiato, C., 2018. Implementing global citizenship education in EU primary schools: the role of government ministries. *Int. J. Dev. Educ. Global Learn.* 10 (1), 21–38.
- UNESCO, 2015. *Global Citizenship Education: Topics and Learning Objectives*. UNESCO, Paris.
- UNESCO, 2020. *Understanding GCED in Asia-Pacific Pacific: A How-To Guide for "Taking It Local"*. UNESCO, Thailand.
- United Nations, Department of Economic and Social Affairs, Population Division, 2019. *International Migrant Stock 2019* (United Nations Database, POP/DB/MIG/Stock/Rev.2019). Retrieved from: <https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp>.
- Valenzuela, A., Brewer, C., 2010. Submerged in a neoliberal utopia: disruption, community dislocation and subtractive citizenship. In: VanHeertum, R., Torres, C., Olmos, L. (Eds.), *Educating the Global Citizen: Globalization, Educational Reform, and the Politics of Equity and Inclusion*. Bentham Science Publishers, Oak Park, pp. 28–43.
- Verma, R., 2017. *Critical Peace Education and Global Citizenship: Narratives From the Unofficial Curriculum*. Routledge, New York.
- Waters, A., 2020. Inequities in access to education: lessons from the COVID-19 pandemic. *Brown Univ. Child Adolesc. Behav. Lett.* 36 (8), 8.
- Weiner, M.F., 2018. Curricular alienation: multiculturalism, tolerance, and immigrants in Dutch primary school history textbooks. *Humanity Soc.* 42 (2), 147–170.
- Yemini, M., Goren, H., Maxwell, C., 2018. Global citizenship education in the era of mobility conflict and globalisation. *Br. J. Educ. Stud.* 66 (4), 1–10.

Human rights education

Felisa L. Tibbitts^a and Sandra L. Sirota^b, ^a Netherlands Institute of Human Rights (SIM), Department of Law, Economics and Governance, Utrecht University, Utrecht, The Netherlands; and ^b Human Rights Institute, University of Connecticut, Storrs, CT, USA

© 2023 Elsevier Ltd. All rights reserved.

Introduction	53
Goals and concepts of human rights education	53
Purpose and structure of the chapter	54
Background on HRE	54
Definition of HRE	54
History of HRE	54
Legal and normative dimensions of human rights	55
Spread of HRE	55
Impacts of HRE	56
Human rights education as an evolving practice	56
Evolving theories of human rights	56
Transformative and emancipatory HRE methodologies	56
HRE in the era of Black Lives Matter and the Covid-19 pandemic	57
Critiques of human rights and HRE	57
Universalism	57
Hegemonic, western-centered and “top down”	58
Possible solutions for human rights critiques	58
HRE in action	59
HRE in National Plans of Action (Mexico)	59
HRE in the K-12 schooling sector (South Africa, USA)	59
HRE in higher education	59
HRE in the non-formal education sector	61
Amnesty International (global)	61
Women for Women’s Human Rights (Turkey)	62
HRE in community development	62
Tostan (Senegal)	62
Human Rights Cities (global)	62
Concluding thoughts	63
References	63

Introduction

Before, I was not aware of my rights. Now, I am aware of all of them. I have the courage to stand up for myself and to ask for my rights when I am denied them.

Premalatha Tamilselvan (HREA et al., 2012)

Premalatha Tamilselvan is a young woman from India who participated in a human rights class in her school as a child. Organized by the non-profit organization, People’s Watch, with support from UN Human Rights, this program trained teachers to bring human rights education to their classrooms. For Premalatha, who is a member of the Dalit community, once known as “Untouchables” and the lowest caste in India, the human rights class was life changing. Now, at the age of 21, she plans to become a lawyer in order to advocate for the human rights of others (Office of the United Nations High Commissioner for Human Rights, 2020a).

Goals and concepts of human rights education

Human rights education (HRE) is a newly established field of educational theory and practice with increasing attention and significance across the globe. This effort, which has gained momentum since the early 1990s, has generated a growing body of educational theory, practice and research that often intersects with activities in other fields of educational study in schools, such as citizenship education, peace education, anti-racism education, Holocaust/genocide education, education for sustainable development and education for intercultural understanding. However, HRE is not only aimed at the formal education sector but has deep

roots in non-formal education, such as non-profit organizations and after school clubs, and also takes place in the training of professionals, such as social workers, teachers and law enforcement officials.

Inter-governmental organizations, including the United Nations, and regional human rights entities, such as the Council of Europe, have promoted HRE in policies and practices. In (UNGA, 2005), the United Nations General Assembly created the permanent and ongoing World Programme for Human Rights Education and, in 2011, passed the Declaration on Human Rights Education and Training. From the UN point of view, knowledge about human rights by all actors in society is fundamental to the virtuous cycle of people knowing and claiming their rights, and governments being held accountable for their human rights promises.

Purpose and structure of the chapter

This chapter shares a brief history and current overview of HRE today, with snapshots of HRE in practice in various contexts across the globe. First explaining HRE through definitions and models of teaching and learning, it then examines HRE with a critical eye. Following this, in connecting theory to practice, the chapter offers examples of HRE in formal, non-formal, and professional settings, as a means of instilling values, sharing knowledge, and training in skills to advocate for human rights.

Background on HRE

Definition of HRE

The most widely accepted definition of HRE is that offered by the United Nations, whose General Assembly passed in December 2011 a Declaration on Human Rights Education and Training with the following language (Article 2):

1. Human rights education and training comprises all educational, training, information, awareness-raising and learning activities aimed at promoting universal respect for and observance of all human rights and fundamental freedoms and thus contributing, inter alia, to the prevention of human rights violations and abuses by providing persons with knowledge, skills and understanding and developing their attitudes and behaviors, to empower them to contribute to the building and promotion of a universal culture of human rights.

This paragraph reaffirms the UN's long-standing definition that HRE has a place in all forms of education and training, including the formal, nonformal and informal sectors. As with other educational processes, HRE has components of knowledge, skills and attitudes, which should be consistent with recognized human rights principles and empower individuals and groups to address oppression and injustice.

The [Office of the United Nations High Commissioner for Human Rights \(2011\)](#) offers further explanation of what HRE encompasses:

- (a) Education *about* human rights, which includes providing knowledge and understanding of human rights norms and principles, the values that underpin them and the mechanisms for their protection;
- (b) Education *through* human rights, which includes learning and teaching in a way that respects the rights of both educators and learners; and
- (c) Education *for* human rights, which includes empowering persons to enjoy and exercise their rights and to respect and uphold the rights of others (Article 2, para 2) (Office of the United Nations High Commissioner for Human Rights, 2011).

This paragraph reflects the evolution of HRE practice, as education *about*, *through* and *for* human rights affirms the full spectrum of learner goals in accordance with knowledge/understanding, values, capacities and actions, with a framework of human dignity and personal empowerment. This new, extended definition also draws attention to teaching and learning processes and reaffirms the outcomes of HRE as being oriented towards taking action “for” human rights.

History of HRE

With the adoption by the United Nations of the Universal Declaration of Human Rights (UDHR) in 1948, HRE was considered crucial to the realization of human rights for all people. In the preamble to the UDHR, it states,

Whereas a common understanding of these rights and freedoms is of the greatest importance for the full realization of this pledge,

Now, Therefore THE GENERAL ASSEMBLY proclaims THIS UNIVERSAL DECLARATION OF HUMAN RIGHTS as a common standard of achievement for all peoples and all nations, to the end that every individual and every organ of society, keeping this Declaration constantly in mind, shall strive by *teaching and education* to promote respect for these rights and freedoms ... (emphasis added). (United Nations, 1948)

Soon after the adoption of the UDHR, the United Nations launched educational campaigns to promote rights. In 1951, UNESCO launched an educational campaign to dispel false and discriminatory information regarding race. UNESCO's (1974) Recommendation concerning Education for International Understanding, Cooperation and Peace, and Education Relating to Human Rights and Fundamental Freedoms, officially called for education to advance human rights. Four years later, the first official definition of HRE was created by the International Federation of Teachers in 1978, which stated:

Human rights education and teaching must aim at: (i) Fostering the attitudes of tolerance, respect and solidarity inherent in human rights; (ii) Providing knowledge about human rights, in both their national and international dimensions, and the institutions established for their implementation; (iii) Developing the individual's awareness of the ways and means by which human rights can be translated into social and political reality and at both the national and the international levels.

The definition has evolved since 1978, as noted above. Yet, both definitions share the common emphasis on fostering respect, sharing knowledge about human rights, and raising awareness about strategies to promote human rights from both an interpersonal and legal perspective.

Legal and normative dimensions of human rights

The goals of HRE are oriented around the elimination of human rights violations. Consistent with this higher aim, HRE can be oriented towards changes in the public domain (the behavior of governments) but also changes in the private domain (the behavior of individuals).

In achieving these goals, HRE has both normative and legal dimensions. The legal dimension incorporates sharing content about international human rights standards as embodied in the UDHR and other human rights treaties and covenants to which countries subscribe. These standards encompass civil and political rights, as well as social, economic, and cultural. In recent years, environmental and collective rights have been added to this evolving framework. This law-oriented approach recognizes the importance of monitoring and accountability in ensuring that governments uphold the letter and spirit of human rights obligations. Through the lens of the legal standards, it is governments (signatories to treaties) that are ultimately responsible for preventing such abuses, both through their own behavior but also through their ability to influence the actions of citizens whose conduct may be infringing on the rights of others.

At the same time, HRE is a normative and cultural enterprise. Through the lens of social change, the goals and process of HRE can also be oriented toward the hearts, heads and hands of ordinary people. The process of HRE is intended to be one that provides skills ("hands"), knowledge ("head"), and motivation ("heart") to individuals to transform their own lives and realities so that they are more consistent with human rights norms and values.

Human rights education in school settings is adapted to the age of learners and the conditions of national/local educational policies and schools. Developmental and conceptual frameworks for HRE have been developed by the United Nations, human rights and education scholars, and non-governmental organizations (NGOs). These frameworks assist in setting goals for HRE, illustrating both what it shares and what it adds to other educational approaches that address values such as social justice.

Human rights themes and content in school curricula can take the form of cross-cultural topics mandated by educational policy or it can be integrated within existing subjects, such as history, civics/citizenship education, social studies and humanities. HRE can also be found in arts programs and non-formal clubs and special events occurring in school settings.

In addition to primary and secondary schools, HRE is organized in settings of higher education; in training programs for professionals such as the police, prison officials, the military, and social workers; for potentially vulnerable and minoritized populations; as part of community development programs; and in public awareness campaigns.

Spread of HRE

As part of the expanded human rights movement, HRE has proliferated and differentiated since the 1990s. In fact, the expansion of HRE has benefited from several worldwide developments. These include not only the worldwide human rights movement but also the expansion of mass education and globalization (Ramirez et al., 2006, as quoted in Russell and Suarez, 2017).

What evidence do we have of the proliferation of HRE? Civil society networks serve an important role in disseminating ideas and norms about human rights (Keck and Sikkink, 1998). The only study on this subject indicated that the number of organizations dedicated to HRE quadrupled between 1980 and 1995 (Ramirez et al., 2006). These developments mirrored similar increases in the number of international NGOs during that time.

Human rights discourse has entered into the realms of the nation-state. The establishment of national human rights institutions is one sign of state commitment to human rights. An analysis of national constitutions from 189 countries through 2005, found that 60% mention the term "human rights" at least once (Beck et al., 2012), although we cannot say, of course, if and how the presence of national human rights institutions or human rights in national constitutions reflects actual state behavior or the emphasis on HRE. For some countries, the link is clearly stated. For example, the South African Constitution mandated the creation of the South African Human Rights Commission, which includes education as one of its core responsibilities.

The institutionalization of HRE is also reflected in the proliferation of HRE academic publications and curricula. A 2017 study carried out by Russell & Suarez found “human rights education” in the titles of 474 publications in 2013, as compared with 51 in the year 2000. In the area of curricula, more than 83 countries across different regions of the world have adopted HRE in legislation, policy documents and curricula since the 1990s (Moon, 2009; as quoted in Russell and Suarez, 2017, p. 31). Studies of textbooks have also shown a dramatic increase in the number of times that human rights is mentioned, with increases most pronounced in Africa, Asia and the West, and least pronounced – though still improved – in Eastern Europe and the Middle East (Russell and Suarez, 2017, p.33).

Hundreds of human rights-related teaching materials have been developed worldwide for use in classrooms and schools, and many of these are widely available, free of charge, on the Internet. The On-Line Resource Center of Human Rights Education Associates (HREA), HRE USA, and Novamerica, serve as good examples. As the COVID-19 pandemic of 2020 forced many education courses online, access to HRE in the virtual space, through webinars, workshops, and academic and community courses, increased remarkably that year. Moreover, bibliographies and descriptive databases of HRE materials are available through key human rights organizations as well as United Nations-related agencies (Office of the United Nations High Commissioner for Human Rights 2005; Amnesty International, 2020b; HREA, n.d.; Council of Europe, n.d.).

Impacts of HRE

The literature on HRE has gravitated towards empowerment as a key feature of successful programming. At the same time, empowerment is a multifaceted and nuanced concept that is difficult to define in concrete and observable terms. Empowerment methodologies can be easily linked with skill development. Having the opportunity to develop concrete skills, such as organizational, advocacy or leadership skills, can be considered a form of “instrumental empowerment” (Ross et al., 2011).

Knowledge itself can be a form of empowerment, for example, learning about the law and how to use it to protect one’s rights. Reflecting and recognizing that one’s personal values are consistent with those contained in international human rights standards or that one’s personal experiences of discrimination are shared by others, can also be empowering.

At the University of San Francisco, graduate students, who were also educators, enrolled in an HRE course in which they contributed to research on HRE based on their own teaching experiences. One teacher described how her students responded with feelings of empowerment when she taught HRE in her class. In particular, students whose histories were not reflected in the history textbooks found the means to challenge oppression through HRE (Adamian, 2015). Another teacher shared that her students found hope in human rights. Some of them have gone on to study human rights law and international relations, and be activists for human rights in college (Arduini, 2015). A third teacher found that HRE fosters critical thinking and action. One of her students shared that HRE allowed them to identify rights violations, recognize privilege, and understand their own ability to influence society to promote human rights (Padilla, 2015).

In India, an organization, People’s Watch, has trained teachers in HRE in nearly 4000 schools. Students shared that they were personally transformed by HRE. They used their knowledge to raise awareness about human rights, report abuses, and even to stop human rights abuses themselves (Bajaj, 2011a,b).

These studies provide just a few examples of the impact HRE programs may have on people’s lives. While additional research is needed, it is clear that HRE has the potential to empower people to take action to protect human rights.

Human rights education as an evolving practice

Evolving theories of human rights

Scholars and practitioners have identified principles of good practice for HRE, but there are no ready-made formulas. For each context, for each set of learners, HRE is customized. A term that has come to be used in the field of HRE is vernacularization or localization. It is this adjustment to content and methodologies in programming that allows the message of human rights to be brought closer to people in their daily lives. Bajaj (2011a,b) distinguishes different versions of HRE by ideology, context, constituency and locale. Tibbitts developed the HRE models of Values & Awareness - Socialization, Accountability-Professional Development and Activism-Transformation according to learner groups and strategies for promoting social change (Tibbitts, 2002, 2017).

This flexibility makes HRE a fascinating area of study. It can also create challenges in relation to developing a core literature that applies across contexts. In other words, HRE that is carefully and strategically designed to account for context and environmental specificities means that the results of studies on how learners experience HRE may provide insights locally but often have limited explanatory power elsewhere. For example, in an analysis of HRE content in South African textbooks post-apartheid, it was found that much of the content and exercises emphasized personal and communal values (Russell et al., 2019). The prominence of communal values is not a common theme across all HRE content, although it is becoming more prevalent.

Transformative and emancipatory HRE methodologies

For this reason, interactive, learner-centered methods are widely promoted. The following pedagogical methods are representative of those promoted by HRE advocates. These methods are applicable to all types of HRE but are most comprehensively implemented in popular education learning models.

Justice-oriented: centering marginalized voices, social and economic rights (in addition to civil and political rights), and the analysis of systemic injustice through lenses of critical, antiracist, multicultural, and/or social justice education ([Gibson and Grant, 2017](#));

Decolonizing: centering the contributions of formerly colonized peoples, acknowledging and rectifying the wrongs committed against them, and rejecting the dominance of Eurocentric ways of teaching and learning ([Zembylas, 2018](#))

Experiential and activity-centered: involving the solicitation of learners' prior knowledge and offering activities that draw out learners' experiences and knowledge;

Problem-posing: challenging the learners' prior knowledge;

Participative: encouraging collective efforts in clarifying concepts, analyzing themes and doing the activities;

Dialectical: requiring learners to compare their knowledge with those from other sources;

Analytical: asking learners to think about why some conditions are unjust and how they came to be so;

Healing: promoting human rights in intra-personal and inter-personal relations;

Strategic thinking-oriented: directing learners to set their own goals and to think of strategic ways of achieving them; and

Goal and action-oriented: allowing learners to plan and organize actions in relation to their goals. ([ARRC, 2003](#))

Transformative methodologies encompass and extend methodologies of instrumental empowerment. Both sets of methodologies are intended to cultivate agency in the learner. However transformative methodologies are different in two respects. The first is that the agency of the learner is cultivated with the explicit aim of social transformation through human rights activism. HRE that prepares learners to organize human rights awareness-raising or campaigning can be associated with transformative methodologies, though this can still be considered a form of instrumental empowerment.

The second way in which transformative methodologies are different than empowerment methodologies is that they can also explicitly foster personal transformation, aligned with the concept of "intrinsic empowerment" ([Ross et al., 2011](#)). Transformative and emancipatory learning approaches, drawing from critical pedagogy, invite a critical reflection on power and oppression in one's local environment, usually as part of a close community of learners. Any subsequent reshaping of one's understanding of the world can result in taking actions to combat one's own oppression in one's family and immediate environment, consistent with wider processes of (privately experienced) social change in a society. When organized on a widespread basis with and for persons belonging to oppressed groups, such personal transformations are the basis of human rights activism.

The specific methodologies of transformative and emancipatory learning are associated with critical pedagogy and [Paulo Freire \(1968\)](#). The HRE literature is strongly rooted in critical pedagogy, which encourages learners to think critically on their situation, recognize connections between their individual problems and the social contexts in which they live and to take action against oppression. Critical pedagogy was and continues to be associated with the HRE Transformation Model ([Bajaj, 2011a,b](#); [Tibbitts, 2002, 2017](#)), as this model is explicitly oriented toward a form of empowerment related to overcoming internalized oppression.

Critical pedagogies offer a means to address human rights violations due to systemic discrimination, that is, discrimination that is enmeshed in the systems and values of society. Anti-racist education is an example of a critical and justice-oriented pedagogy that can support learners and educators in disrupting racism at the individual, societal, and global levels. It provides them with the tools to examine racism on multiple levels, including their own position in perpetuating or dismantling it ([Gibson and Grant, 2017](#); [Love, 2019](#)).

HRE in the era of Black Lives Matter and the Covid-19 pandemic

Current events have human rights dimensions and the application of the human rights lens allows for conversations about structural violence and injustices, especially those affecting vulnerable groups. The Black Lives Matter Movement is a human rights movement that calls for racial justice. The movement is not only calling for the end to discrimination against individuals, but a dismantling of the structures in society that perpetuate discrimination. To address injustices fully, HRE must teach about both forms of discrimination as well, and strategies to advocate for change. Teaching about current events such as Black Lives Matter can help learners to understand the need for this dual approach to human rights advocacy.

The global health crisis of Covid-19 has revealed underlying inequalities – both within and across countries – that are affecting people's access to health care including the Covid-19 vaccine, quality education and even food and water. At the time this chapter was being written, new HRE curriculum was under development to support teachers in introducing human rights themes related to the pandemic, foster the sharing of student experiences during the pandemic, and promote taking action to address injustices revealed in the local community and on the global level ([Generation Human Rights and HREA, 2020](#)).

Critiques of human rights and HRE

Universalism

Human rights education (HRE) is one approach in secular education that has had to grapple with the concept of universal values. A traditional HRE approach promotes the universal values of the United Nations and the position that human rights are "universal, interrelated, indivisible and interdependent" ([UN, 1993](#); [OHCHR, 2022](#)). Numerous HRE resources are explicitly oriented toward the transmission of human rights values and standards. The intention is that the human rights framework will be used as a lens for

analyzing human and government behavior and as a basis for improving the realization of human rights in the everyday lives of people.

Hegemonic, western-centered and “top down”

Human rights education needs to be taught in a way that allows for questioning and dialogue. Some HRE scholars have critiqued a “declarationist” approach to HRE as promoting the values embodied in international human rights standards as absolutist, negating the possibility of genuine “dialogue” with learners in regard to their existing value systems (Keet, 2012). At best, such an HRE approach fails to fully foster the critical capacities of learners; at worst, it promotes a hegemonic, Western-centered values system. For HRE to truly be empowering and transformational, it should not be taught in a way that replicates the very rights abuses it aims to dismantle. For that reason alone, HRE necessitates the inclusion of dialogue, deliberation, and debate.

HRE lessons and curriculum aligned with the traditional, or declarationist, approach draw on some of the techniques associated with the psychological theory of intuitive ethics as well as the values clarification approach, a pedagogy that provides learners with opportunities to explore and develop their own value system (Lee, 2001, p. 33). At their core, such values clarification activities presume that the learners will already possess values that are consistent with human rights standards and respect for human dignity. From this perspective, we can see that HRE indirectly endorses the proposal that there are “common values” that operate cross-culturally. In practice, these kinds of HRE activities are also able to bring out some complexity, such as the potential tension between multiple human rights (for instance, freedom of expression in a situation where hate speech is being used).

Some scholars have identified HRE teaching and learning processes that do not allow for a genuine critique of the human rights system as a whole and have therefore promoted political correctness (Baxi, 2009; Keet, 2014). One may use even values-clarification techniques instrumentally for validating the preordained universal values of human rights. Such concerns build on those presented by relativists that the human rights framework reflects an overly dominant Western influence in relation to the ideas of the Enlightenment and the historical context of the immediate post-WWII period. Critics of a traditional “transmission” approach to HRE argue that it does not allow learners to consider these concerns and is therefore contributing to a non-reflective socialization process. These concerns are also amplified in environments where there is sensitivity to the term “human rights” for reasons related to local political dynamics.

Possible solutions for human rights critiques

Educators, scholars, and practitioners have proposed solutions to these critiques that are primarily pedagogical. Some now promote the critical pedagogy approach – established by Paulo Freire and foundational to the popular education version of HRE – for analyzing the human rights framework as a whole, not only the structures in which learners live and work. This approach can explicitly reveal and examine the Western origins and potential hegemonic influences of the human rights system, while in a learning context dedicated to applying teaching and learning systems to promote and protect human dignity. The focus then becomes not whether the human rights framework is the preferable value system to apply, but “what is to be done” to bring about change – using the value system and language indigenous to and embraced by the learner (and presumably consistent with humanistic values).

The literature calls this approach “critical HRE” (CHRE) in order to emphasize that the critique extends internally to the human rights system itself, as well as to society (Keet, 2015; Bajaj 2011a,b). A more recent development is the use of the term “transformative HRE” (THRED) (Bajaj et al., 2016), emphasizing the role of pedagogy in promoting social change and consistent with the transformative model of HRE (Tibbitts, 2002, 2017). Some scholar-practitioners have moved away from use of the term “HRE” altogether and instead promote the concept of “humanizing pedagogy” or the “pedagogy of mutual vulnerabilities” (Salazar, 2013; Keet et al., 2009). Some view such approaches as helping to ensure the emancipatory and liberating potential of HRE.

Another pedagogical response actively seeks to bridge local and transnational vernacular rights cultures. Critical interpretive pedagogy proposes more specifically that human rights must be interpreted within a plurality of cultural, religious and philosophical doctrines (Adami, 2014; Al-Duraweesh and Snauwaert, 2013). Rather than simply applying a universal set of abstract principles, rights are actively interpreted, shaped and transformed. They are thus “sutured with the multiple histories, cultures and contexts in which they are formed, and reflect a history of struggles that is both local and transnational” (Dunford and Madhok, 2015). This quasi-universalistic approach is a practical accommodation to the criticisms of universal values.

In complement with these pedagogical methods, the decolonization of HRE may ensure that HRE does not perpetuate the very oppression it seeks to undo. Decolonizing HRE means to not favor Eurocentric epistemologies to the detriment of others, particularly those from the Global South. Decolonization could include the acknowledgment and attribution of contributions to the human rights framework by non-Western countries and individuals and to share content and knowledge from non-Western scholars, practitioners, and activists. It goes beyond the classroom content, though, to include who is teaching, how they teach, and the environment in which they teach.

The HRE models and pedagogies inform teaching and learning at the practical level. HRE is implemented in a wide range of settings and contexts with diverse groups of learners. To be effective, educators must adapt their pedagogical approaches and educational content based on these variations. In the next sections, we share examples of this variety of approaches and content of HRE in action.

HRE in action

HRE is offered in formal settings such as primary, secondary, and higher education and as part of professional development. It is offered non-formally by NGOs and after-school clubs. It is even present in informal settings such as in community conversations, as public art, and around the dinner table. The online space, in part spurred on by the COVID-19 pandemic, is now flooded with HRE courses, workshops, lessons, and other resources.

HRE in National Plans of Action (Mexico)

In national curricula, HRE is included at different levels. At least 20 countries have created National Plans of Action for Human Rights Education ([Office of the United Nations High Commissioner for Human Rights, 2020a,b](#)). Another 36 countries address HRE in their National Plans of Actions for Human Rights. The inclusion of HRE in a national level plan is one sign of a country's commitment to the right to HRE. Mexico offers one example of this. The first country to include economic and social rights in its constitution in 1917 ([Sikkink, 2019](#)), today in Mexico, HRE is offered through NGOs as well as local and national commissions for human rights in order to work towards ending current and preventing future human rights violations ([Martinez Sanz, 2018](#)). Between 2013 and 2015, the National Human Rights Commission reported that public organizations had carried out at least 20,000 HRE related activities in the country. This number does not include non-profit organizations and other private sector activities ([Martinez Sanz, 2018](#)).

Even countries that do not have plans at the national level may still include HRE in their formal and non-formal educational institutions. For example, in the United States, where education is decentralized, at least 42 of the 50 states have required or recommended HRE standards in their social studies curricula ([Sirota, 2017](#)). Through these diverse and varied offerings, HRE reaches individuals and communities around the world, as described in the following examples.

HRE in the K-12 schooling sector (South Africa, USA)

In South Africa, HRE can be found in the secondary school textbooks for the subjects of social studies, history, and life orientation. In social studies and history, the HRE content is integrated into other material, such as World War II. In life orientation, human rights is offered as its own topic, though requiring only 7 h of instruction in 10th and 11th grades, and 4 h in 12th grade. In the history and social studies textbooks, HRE focuses on knowledge and values. In contrast, the life orientation textbooks include lessons to ensure learners not only know their rights and critically consider their values, but also how to advocate for human rights ([Russell et al., 2019](#)). At present, a plan is underway to include HRE in a new national teacher training program, Teaching for All, in South Africa. ([British Council, 2020](#)).

In the United States, as education is decentralized, each state sets its own curriculum standards. While HRE is in social studies curriculum standards in at least 42 states, only two states include HRE at every grade level – New Jersey and Mississippi ([The Advocates for Human Rights, 2016](#)). Where national and state mandates are lacking, local school districts may choose to include HRE more robustly than required. In Manchester, Connecticut, for example, HRE is now a required high school course in order for students to graduate ([Faipier, 2018](#)). Some schools that do not offer standalone human rights classes invest in human rights by supporting teachers' professional development through HRE organizations such as RFK Human Rights. Through the organization's program, Speak Truth to Power, RFK Human Rights has trained over 9000 educators around the world ([RFK Human Rights, 2020](#)).

HRE in higher education

Human rights centers (global, USA)

At the higher education level, as in many other areas of education, HRE has a growing presence. Human rights and HRE degrees are offered at the bachelors, masters, and doctoral levels. Universities and colleges also offer human rights as an undergraduate double major or a minor. They offer human rights concentrations or a certificate. At even more universities and colleges, at a minimum, students can enroll in individual human rights courses. In addition to degrees and courses, in recent years two special issues on HRE were published in the journals *Prospects* and the *Journal of Human Rights*. Two new journals dedicated to HRE were created – the *International Journal of Human Rights Education*, out of the University of San Francisco and the *Human Rights Education Review*, out of the University of South-Eastern Norway.

College and university human rights centers are often integral to the promotion of HRE on campus, and in the surrounding communities. In addition to overseeing and coordinating majors, minors, concentrations, and certificates, they provide offerings to the broader community within and around the institution through public events, workshops and classes. In recent years, human rights educators and center directors have established networks to support each other's work and further develop the field of HRE. In the United States, the University and College Consortium for Human Rights Education (UCCHRE) was founded in 2016. The mission of the consortium is "to further human rights learning, research, policy, and practice within and across university and college communities." UCCHRE's effort is another example of the growing cooperation in and professionalization of the field.

Practical human rights courses in higher education offer the opportunity for students to contribute to advocating for human rights by working on specific issues and learning from community members. These opportunities may be offered by professors in partnership with civil society organizations. Professor Bill Simmons teaches a course at the University of Arizona known as Problem-based Learning without Borders (PBL). He emphasizes the course's versatility: it can be taught at the undergraduate or graduate level, online and offline, with any group of students and in large and small classes. Course meetings may even be held at

a community-based setting. Students research local or global issues, working with and learning from the community-based group. Past courses have focused on human rights topics such as mass incarceration in Arizona, femicide in Mexico, and women's rights in South Sudan. The students conduct research, offer strategies to achieve the community's goals, and work with community partners to strengthen and implement those strategies. Professor Simmons' students have directly contributed to efforts to secure human rights, such as when their work was used in a legal case on femicide in Mexico which was brought in front of the Inter-American Court of Human Rights (Simmons, 2020).

Scholars at risk Student Advocacy Seminars (global, USA)

Opportunities for gaining practical skills and experiences, and contributing to the field of human rights may be initiated by NGOs in partnership with institutions of higher education. The organization Scholars at Risk (SAR) offers an example of this. SAR is an NGO that since 2000 has worked with colleges and universities around the world to host scholars whose academic freedom is in jeopardy in their home countries. In 2014, SAR launched Student Advocacy Seminars. Through these seminars which are run around the world, students spend a semester conducting research and advocating to secure the freedom of a scholar who has been wrongfully imprisoned, in violation of their human rights, often due to advocating for democracy. Students gain knowledge and skills in promoting human rights as they simultaneously apply theory to practice in advocating for the freedom of an imprisoned scholar (Scholars at Risk Network, 2020).

Universities and colleges can be resources for the general public to access HRE. The University of Arizona, University of Connecticut, Columbia University, the University of Graz (Austria), Université catholique de Louvain (Belgium) and Leiden University (The Netherlands), among others, offer human rights courses and workshops that are open to the public on current issues such as systemic racism, migrant and refugee rights and the COVID-19 pandemic, and on skills-building such as conducting research or teaching human rights.

HRE in teacher training (USA)

As demonstrated, HRE is included in regional, national, and local curriculum in primary, secondary, and higher education. Teachers receive training in HRE through multiple avenues. Teacher education in human rights may take place in universities and colleges, through non-profit organizations, at the level of the national government, or in professional development programs in K-12 schools.

Teacher education in human rights is a fairly small field at present, though in a number of places, it is readily available and robust. In the northeastern US, the University of Connecticut offers HRE to pre-service teachers and current high school teachers. For example, teachers may become certified as Early College Experience, or ECE teachers, thereby being able to teach human rights to students in their high school, who receive college credit for these classes. After their initial coursework, in order to maintain ECE certification, teachers must participate in HRE workshops at least once every two years. The number of teachers who have become ECE certified is growing. As of 2020, 25 teachers across the state were certified, with six more teachers in the midst of the certification process (Glenn Mitoma, personal communication, August 7, 2020).

HRE in the training of social workers and psychologists (Netherlands, USA, Europe)

At the heart of the UDHR is the "equal unalienable worth of a human being". The fields of both social work and psychology are essentially about the problems of human beings and human relationships. In the field of psychology, the person, couple or family unit is the unit of change whereas in social work the community and society (in addition to the person and family) can be the focus area.

In recent years, practitioners and trainers in these fields have begun to integrate an explicit human rights dimension into their work. Human rights is linked with the core dignity of each client and potentially offers a universal framework of ethics that is rooted in international law. The human rights value system is in line with the "capabilities approach" argued by Sen (1999), who sees capabilities as essential for freedom to being human, to develop and for the realization of human rights. Moreover, the human rights approach for social work explicitly recognizes community organizing as a form of community social work (Staples, 2012) that "connects people on the basis of a sense of common experiences, such as exclusion, poverty or other forms of social injustices" (Rayneart et al., 2010, p. 7).

In the social work sector, the International Federation of Social Workers has endorsed human rights as a core competency. The working group of social work trainers in the Netherlands and Flanders developed a manifesto in 2015 calling for the inclusion of HRE within the training of social workers (Raynaert et al., 2010). As of 2015, the Council on Social Work Education has as its Competency 3: Advance Human Rights and Social, Economic and Environmental Justice as a required goal for schools of social work in order to be accredited (Council on Social Work Education, 2015, p.7). This competency creates an incentive for more schools of social work to integrate these principles within their education programs. The National Association of Social Workers (2006) in the US also has a Policy on Human Rights.

Scholarship related to human rights and social work is growing, including the launch of the *Journal of Human Rights and Social Work*. This scholarship includes studies related to the application of human rights to social work policies and practices.

The Institute for Social Work at the Utrecht University of Applied Sciences (Netherlands) provides an example of the infusion of HRE. The first year of social work education curriculum now integrates human rights themes repeatedly within the program, Principles of Social Work. This approach cements several theme-centered programs where HRE is delivered in a more reflective and experiential way. The recently developed second-year bachelor pursues the same didactic structure of first year Social Work. The second

year comprises four theme-centered programs (1) social security and safety, (2) social (in)justice (community level), (3) participation, and (4) empowerment (individual level) (Reynaert et al., 2010, p. 3).

Over the years, both international and national psychology associations have endorsed the application of ethics and, specifically, human rights norms in the professional practices of researchers, practitioners and educators within psychology. This includes the following: The International Association of Applied Psychology (IAAP), the International Association for Cross-cultural Psychology (IACCP), and the International Union of Psychological Science (IUPsyS) developed and endorsed The Universal Declaration of Ethical Principles for Psychologists (UDEPP) (2008). The European Federation of Psychologists' Associations (EFPA) established a Board of Human Rights and Psychology in 2013. In addition, there are numerous national associations, like the American Psychological Association (APA), which have issued statements and principles related to human rights, psychologists and codes of conduct.

Thematic areas where HRE might be integrated within psychology education include, but are not limited to:

- International and regional human rights standards related to psychology (e.g., the UDHR and the rights of vulnerable groups)
- History of psychology/psychologists and human rights (e.g., psychological approaches to intergroup contacts, social inclusion, conflict and discrimination)
- Human rights as part of ethics/codes of conduct training
- Advocacy and root cause analysis within practicums (e.g., addressing political, social and economic conditions violating human rights and contributing to human suffering)
- Foundation of psychology – Feminist, cross-cultural and decolonial perspectives (Tibbitts and Hagenaars, 2020, pp. 254–259).

The Global Network of Psychologists for Human Rights (GNPHR) (<http://humanrightspscychology.org/>) was launched in 2020, and the website is hosted by the International Council of Psychologists (ICP, Inc). The website contains an evolving resource section to support the integration of human rights within psychology education and training. In this fast-developing area of HRE work, two publications came out in 2020 (P. Hagenaars et al., 2020; Rubin and Flores, 2020) and more are anticipated (Global Network of Psychologists for Human Rights, 2020).

HRE in the non-formal education sector

Amnesty International (global)

Amnesty International is the largest membership-based human rights organization in the world with over 2 million members and 5 million supporters (Amnesty International website). Amnesty's primary focus is research and campaigning to influence governments to abide by human rights obligations. Founded in 1961 in the United Kingdom initially to campaign for the release of political prisoners, the organization began to integrate HRE within its activities in the mid-1990s (Amnesty International, 2020a).

One key program is the "Write for Rights" (W4R) campaign, which takes place annually around Human Rights Day on December 10th. Human Rights Day commemorates the day when the Universal Declaration of Human Rights was adopted in 1948, shortly after the founding of the United Nations. Amnesty encourages people to show their solidarity with the international human rights movement by engaging in the W4R campaign.

A major part of the W4R campaign consists of a letter-writing marathon which involves millions of people around the globe. In 2017, more than 6 million actions were taken in support of the individuals and groups featuring in the campaign (Amnesty International, 2018, p. 3). Every year, supporters write two letters. "One is to the person in authority – it could be a king/queen, president or head of police – who can help make change happen. The other letter is to the person or group of people we are fighting for (and/or their family) to express support so they know we will never forget them" (Amnesty International, 2018, p. 5).

For example, one of the human rights defenders that Amnesty included in their 2018 campaign was Atena Daemi from Iran, who was serving a seven-year prison sentence in Tehran's Evin prison. She had suffered inhumane conditions and been denied medical treatment as retaliation for defending human rights, in particular for speaking out against the death penalty (Amnesty International, 2018, p. 13).

Another key HRE program carried out by Amnesty International is "Human Rights Friendly Schools", which began in 2009 and involves 16 countries. Amnesty staff and volunteers work with whole schools that "places human rights at the heart of the learning experience and makes human rights an integral part of everyday school life" (Amnesty International, n.d.). Beginning with a problem-seeking analysis of school conditions using a human rights lens, projects are undertaken to improve the realization of the principles of equality, dignity, respect, non-discrimination and participation. All aspects of school life are eligible for improvement: decision making in the school, discipline policies, relationships, curriculum, extra-curricular activities and the school's physical environment. Amnesty's Human Rights Friendly Schools program has a toolkit and builds on the literature of the whole school approach and whole school development.

A recently developed HRE area for Amnesty is its *Human Rights Academy*, an online platform for human rights learning. The initiative was spearheaded by the Amnesty team in the Middle East and North Africa in the wake of the Arab Spring, in order to support human rights defenders in the region and elsewhere. The multi-lingual learning platform includes self-paced online learning (for example, on different human rights areas, such as the rights of refugees), advanced capacity building on documenting human rights violations as well as online tools to facilitate youth activism (Amnesty International, 2020b).

Women for Women's Human Rights (Turkey)

The majority of HRE efforts carried out in the non-formal education sector are through local or national organizations. One organization with a long record of carrying out HRE is Women for Women's Human Rights (WWHR)- New Ways, which operates in Turkey. WWHR – New Ways is a non-governmental organization (NGO) that organized a legal literacy program for adult women through a 14-year partnership with the Turkish General Directorate of Social Services (GDSS). WWHR prepared GDSS staff to deliver a 16-week Human Rights Education Program (HREP) for women who come to local community centers to receive services. Based on both a legal rights and gender perspective, the 16-week program seeks to empower women to know and claim their rights both in the private sphere (including familial relations, sexual and reproductive rights, gender sensitive parenting) and the public sphere (including economic rights, political rights, organizing, and access to justice).

According to WWHR literature, the specific objectives of HREP are:

- Enable women to gain awareness of their rights as equal citizens;
- Raise women's consciousness of how customary practices, that is, 'unwritten' laws, often limit or violate their rights;
- Serve as a catalyst for women's organizing efforts on the local and national levels and support grassroots organizing initiatives;
- Support women in developing strategies for the implementation of their legal rights in daily life (Tibbitts, 2016, pp. 2–3).

As of fall 2013, nearly 9300 women had participated in the program. The results of an impact assessment study in 2012 demonstrated impacts both on the learners and facilitators. Areas of impact included their identity; knowledge, skills, and attitudes; and behaviors in their family and in the workplace. These results are demonstrated not only in the private domain, but in the women's professional roles (Tibbitts, 2016, pp. 16–17).

HRE in community development

Tostan (Senegal)

HRE programs can be embedded within community development programs and linked with skills related to self-reliance. A community-based approach to HRE can involve processes such as mapping power and leadership in a community, holding discussions about the needs of members, and deciding collaboratively which ones to prioritize. In the absence of effective state services, community members address their own problems, which may or may not involve resorting to changing formal or customary laws but will almost certainly involve human rights values. In the non-formal HRE program in Senegal, the NGO Tostan integrates HRE within their Community Empowerment Program (Cislighi et al., 2016).

Tostan is a three-year, education program that puts rural communities in charge of their own futures. The Community Empowerment Program (CEP) involves human rights-based education classes and the establishment of a Community Management Committee (CMC).

Each village runs two classes of 25–30 participants – one for adults and one for adolescents – that meet three times per week over the course of the program. These are facilitated by someone from the same ethnic group but a different village, thus someone fluent in the local language and familiar with local customs. The classes draw on “modern nonformal education techniques as well as traditional African oral traditions like theater, storytelling, dance, artwork, song, and debate” (Tostan, 2020).

In addition to the CEP classes, communities establish a CMC that is responsible for implementing development projects designed by the community. Trained by Tostan, these are democratically-selected 17-member committees, of which nine members are women (Tostan, 2020).

Human Rights Cities (global)

The global Human Rights Cities movement similarly focuses on change at the city, town and village levels. A Human Rights City may be defined as a municipality that has adopted human rights principles and laws as guiding norms of governance. In practice, it provides for an “inclusive civic engagement with city management, where people have the power to decide on policy proposals and politicians assume the role of policy implementers, hence fostering participatory democracy with solidarity and social justice. Such a commitment translates locally into an improvement in the quality of life of all residents” (Human Rights Cities, 2020).

There are human rights cities in every major global region, which are also networked. Examples include Korogocho (Kenya), Nuremberg (Germany), Porto-Alegre (Brazil), Rosario (Argentina), and Walewale (Ghana). In the cases of these cities, each process has been tailor-made, is context specific and based on individual commitments. Different degrees of commitment and responsibility exist among human rights cities, which may vary from the declaration of intention to the effective implementation of a process where residents' civil, political, economic, social and cultural rights are at the center of the city's governance and public policies (Human Rights Cities, 2020).

Human rights cities initiatives involve some degree of public awareness raising for HRE. A human rights city initiative that is formalized can involve a commitment to the Universal Declaration of Human Rights as a norm of governance, combined with other human rights conventions such as the Convention on the Rights of the Child (1989) or the Convention on the Elimination of All Forms of Racial Discrimination (1969). Key is “the establishment of a socio-political process where the community and the municipality (usually represented by a governing body) cooperate to integrate the fundamental values as a priority in the planning and governance of the city” (Human Rights Cities, 2020).

Within the community development framework, HRE fosters the changed behavior of both state- and non-state actors, that is, everyday people in their daily lives. The norms for such changed behavior can be fed through the cross-cutting (human rights)

values of non-discrimination, equality, inclusion and participation, as well as the norms associated with the human rights of specific marginalized groups.

Concluding thoughts

The Universal Declaration of Human Rights was created in 1948. Despite the fact that it has been translated into more languages than any other document in the world, over 70 years later, people around the globe still lack access to human rights education. Yet, as demonstrated, promising HRE initiatives can be found flourishing in schools, communities, and professional spaces, empowering both learners and educators. In order for it to continue to flourish, just as the UDHR has been reinterpreted over time to now include such rights as the internet and same sex marriage, human rights educators must continue to reflect on and critique HRE to ensure it adapts with the times.

References

- Adami, R., 2014. Human rights for more than one voice: rethinking political space beyond the global/local divide. *Ethics Global Polit* 7 (4), 163–180.
- Adamian, A.S., 2015. Bringing to life human rights education in the science classroom. In: Katz, S., McEvoy Spero, A. (Eds.), *Bringing Human Rights Education to US Classrooms*. Palgrave Macmillan US, New York, NY, pp. 65–86.
- Al-Duraweesh, F., Snauwaert, D.T., 2013. Towards a hermeneutical theory of international human rights education. *Educ. Theor.* 63 (4), 389–412.
- Amnesty International (nd) Human Rights Cities and insert before (Accessed 14 April 2022). <https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.amnesty.org%2Fen%2Fhuman-rights-education%2Fhuman-rights-friendly-schools&data=05%7C01%7CMRW-EDU4%40elsevier.com%7C8981dc92e7b948566b2908da787fd0e1%7C9274ee3f94254109a27f9fb15c10675d%7C0%7C0%7C637954787708896625%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAilCJQjoiV2luMzIlCjBtIlk1haWwILCjXVCl6Mn0%3D%7C3000%7C%7C%7C&sdata=m2vK7uiYSrDmEDfw%2Bs2g099ZAwizGxTnLz0Ne0fKu6g%3D&reserved=0> <https://www.amnesty.org/en/human-rights-education/human-rights-friendly-schools>
- Amnesty International, 2018. *Write for Rights 2018 – A Human Rights Education Toolkit for Educators*. Amnesty International, London.
- Amnesty International, 2020a. Human Rights Education. www.amnesty.org/en/human-rights-education/human-rights-friendly-schools/#:~:text=Amnesty%20International%20supports%20schools%20and%20their%20wider%20communities,and%20principles%20into%20all%20areas%20of%20school%20life. (Accessed 7 July 2020).
- Amnesty International, 2020b. Human Rights Academy – MENA. <https://www.amnesty.org/en/latest/education/2019/07/human-rights-academy-mena/>. (Accessed 7 July 2020).
- Arduini, B.J., 2015. Teaching the crucible through human rights. In: Katz, S., McEvoy Spero, A. (Eds.), *Bringing Human Rights Education to US Classrooms*. Palgrave Macmillan US, New York, NY, pp. 149–167.
- ARRC [Asia-Pacific Regional Resource Center for Human Rights Education], 2003. What is human rights education?. In: *Human Rights Education Pack*. ARRC, Bangkok, Thailand.
- Bajaj, M., 2011a. Human rights education: ideology, location, and approaches. *Hum. Right Q.* 33, 481–508.
- Bajaj, M., 2011b. *Schooling for Change: The Rise and Impact of Human Rights Education in India*. Bloomsbury, New York.
- Bajaj, M., Cislighi, B., Mackie, G., 2016. *Advancing Transformative Human Rights Education: Appendix D to the Report of the Global Citizenship Commission*. Open Book Publishers, Cambridge, UK. <https://doi.org/10.11647/obp.0091.13>.
- Baxi, U., 2009. *Human Rights in a Posthuman World*. Oxford University Press, New Delhi.
- Beck, C.J., Drori, G.S., Meyer, J.W., 2012. World influence on human rights language in constitutions. *Int. Sociol.* 27 (4).
- British Council, 2020. Teaching for All. Retrieved January 31, 2021 from. <https://www.britishcouncil.org/programmes/education/teaching-all>.
- Cislighi, B., Gillespie, D., Mackie, G., 2016. *Values Deliberation and Collection Action: Community Empowerment in Rural Senegal*. Palgrave Macmillan, New York.
- Council of Europe (n.d). Education for Democratic Citizenship and Human Rights Education - Resources. <https://www.coe.int/en/web/edc/resources>. (Accessed 8 June 2022).
- Accessed.
- Council on Social Work Education, 2015. 2015 Educational Policy and Accreditation Standards for Baccalaureate and Master's Social Work Programs. Council on Social Work Education, Alexandria, VA. In: www.csw.org/getattachment/Accreditation/Accreditation-Process/2015-EPAS/2015EPAS_Web_FINAL.pdf.aspx.
- Dunford, R.F., Madhok, S., 2015. Vernacular rights cultures and the 'right to have rights'. *Citizen. Stud.* 19 (6–7), 605–619.
- Faipier, D., December 9, 2018. Advancing Human Rights Education in Connecticut 70 Years After UDHR. *UConn Today*. Retrieved from. <https://today.uconn.edu/2018/12/advancing-human-rights-education-connecticut-70-years-udhr/>.
- Freire, P., 1968. *Pedagogy of the Oppressed*. Seabury Press, New York.
- Generation HR and HREA, 2020. UNmasking- the Pandemic Curriculum. <https://www.generationhumanrights.org/>. (Accessed 10 September 2020).
- Gibson, M.L., Grant, C.A., 2017. Historicizing critical educational praxis: a human rights framework for justice-oriented teaching. In: *Human Rights Education: Theory, Research, Praxis*, p. 225.
- Global Network of Psychologists for Human Rights (GNPHR) (2020). <http://humanrightspychology.org/>. (Accessed 10 September 2020).
- Hagenaars, P., Plavšić, M., Sveaas, N., Wagner, U., Wainwright, T., 2020. *Human Rights and Human Rights Education for Psychologists*. Routledge, New York.
- HREA, Soka Gakkai International, OHCHR, 2012. *A Path to Dignity: The Power of Human Rights Education*. <http://pathtodignity.org/>. (Accessed 10 September 2020).
- Human Rights Education Associates (HREA)(n.d.) Online Resource Center. (Accessed 16 April 2022).
- Human Rights Cities, 2020. What is a Human Rights City? <https://humanrightscities.net/what-we-do/>. (Accessed 3 August 2020).
- Keck, M.E., Sikkink, K., 1998. *Activists Beyond Borders*. Cornell University Press, Ithaca.
- Keet, A., 2012. Discourse, betrayal, critique: the renewal of human rights education. In: Roux, C. (Ed.), *Safe Spaces: Human Rights Education in Diverse Contexts*. Sense Publishers, Rotterdam, pp. 7–27.
- Keet, A., 2014. Plasticity, critical hope and the regeneration of human rights education. In: Bozalek, V., Leibowitz, B., Carolissen, R., Boler, M. (Eds.), *Discerning Critical Hope in Educational Practices* (69–81). Routledge, London and New York.
- Keet, A., 2015. It is time: critical human rights education in an age of counter-hegemonic distrust. *Educ. Change.* 19 (3), 46–64.
- Keet, A., Zinn, D., Porteus, K., 2009. Mutual vulnerability: a key principle in a humanizing pedagogy in post-conflict societies. *Perspect. Educ.* 27 (2), 109–119.
- Lee, W.O., 2001. Moral perspectives on values, culture and education. In: Cairns, J., Lawton, D., Gardner, R. (Eds.), *World Yearbook of Education 2001: Values, Culture and Education*. Kogan Page, London, pp. 27–45.
- Love, B.L., 2019. *We Want to Do More than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon Press.
- Martinez-Sanz, G., 2018. Human rights education and training programs in Mexico: a cross-case analysis of practitioners' professional knowledge and practices. *Mex. Law Rev.* XI (1), 103–130.
- Moon, R., 2009. *Teaching World Citizenship: The Cross-National Adoption of Human Rights Education in Formal Schooling*. Doctoral dissertation. Stanford University, Stanford, CA.

- Office of the United Nations High Commissioner for Human Rights, 2005. World Programme for Human Rights Education. United Nations, Geneva, Switzerland.
- Office of the United Nations High Commissioner for Human Rights, 2011. United Nations Declaration on Human Rights Education and Training. GA 66/127. United Nations, Geneva.
- Office of the United Nations High Commissioner for Human Rights, 2020a. From Student to Human Rights Defender: The Power of Human Rights Education. Retrieved from www.ohchr.org/EN/NewsEvents/Pfrom:fages/Powerofhumanrightseducation.aspx. October 30, 2020.
- Office of the United Nations High Commissioner for Human Rights, 2020b. National Plans of Action. <https://www.ohchr.org/EN/Issues/Education/Training/Pages/NationalActionPlansHumanRightsEducation.aspx>. (Accessed 1 September 2020).
- OHCHR, 2022. What are Human Rights? <https://www.ohchr.org/en/what-are-human-rights>. (Accessed 8 June 2022). Accessed.
- Padilla, L., 2015. Reframing a community college social problems course through a human rights perspective. In: Katz, S., Spero, A. (Eds.), *Bringing Human Rights Education to US Classrooms*. Palgrave Macmillan US, New York, NY, pp. 169–185.
- Ramirez, F.O., Suarez, D., Meyer, J.W., 2006. The worldwide rise of human rights education. In: Benavot, A., Braslavsky, C. (Eds.), *School Knowledge in Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education*. Comparative Education Research Centre (CERC) and Springer, Hong Kong, Hong Kong.
- Rayneart, D., Van de Veld, S., Bie, M.B., 2010. Children's rights education and social work: contrasting models and understanding. *Int. Soc. Work* 53 (4), 443–456.
- RFK Human Rights, 2020. Speak Truth to Power. Retrieved from. <https://rfkhumanrights.org/work/teaching-human-rights>.
- Ross, H.A., Shah, P.P., Wang, L., 2011. Situating empowerment for millennial schoolgirls in Gujarat, India and Shaanxi, China. *Fem. Form.* 23 (3), 23–47.
- Russell, S.G., Suarez, D., 2017. Symbol and substance: human rights education as an Emergent Global Institution. In: Bajaj, M. (Ed.), *Human Rights Education: Theory, Research, Praxis*. University of Pennsylvania Press, Philadelphia.
- Rubin, N.S., Flores, R. (Eds.), 2020. *Cambridge Handbook on Psychology and Human Rights*. Cambridge University Press, Cambridge, UK.
- Russell, S.G., Sirota, S.L., Ahmed, A.K., 2019. Human rights education in South Africa: ideological shifts and curricular reforms. *Comp. Educ. Rev.* 63 (1), 1–27.
- Salazar, M., 2013. A humanizing pedagogy: reinventing the principles and practices of education as a journey towards liberation. *Rev. Res. Educ.* 37 (1), 121–148.
- Scholars at Risk Network, 2020. Student Advocacy Seminars. <https://www.scholarsatrisk.org/actions/student-advocacy-seminars/>. (Accessed 10 September 2020).
- Sen, A., 1999. *Development as Freedom*. Allen Lane, London.
- Sikkink, K., 2019. *Evidence for Hope: Making Human Rights Work in the 21st Century*, vol. 28. Princeton University Press.
- Simmons, W., 2020. Problem Based Learning Without Borders. Retrieved from. <https://www.williampaulsimmons.com/problem-based-learning-without-borders>. (Accessed 1 November 2020).
- Sirota, S., 2017. The inconsistent past and uncertain future of human rights education in the United States. *Prospects* 47 (1–2), 101–117.
- Staples, L., 2012. Community organizing for social justice: grassroots groups for power. *Soc. Work Groups* 35, 287–296.
- The Advocates for Human Rights, 2016. Human Rights in State Social Studies Standards: An Analysis. Report prepared for Human Rights Educators USA.
- The National Association of Social Workers, 2006. International Policy on Human Rights. *Social Justice*. socialworkers.org.
- Tibbitts, F., 2002. Understanding what we do: emerging models for human rights education. *Int. Rev. Educ.* 48 (3–4), 531–549.
- Tibbitts, F., 2016. Women's human rights education trainers in Turkey: situated empowerment for social change. *J. Peace Educ.* 13 (1), 41–59.
- Tibbitts, F., 2017. Evolution of human rights education models. In: Bajaj, M. (Ed.), *Human Rights Education: Theory, Research, Praxis*. University of Pennsylvania Press, Philadelphia, pp. 69–95.
- Tibbitts, F., Hagenaars, P., 2020. Chapter 18. Planning human rights education for psychologists. In: Hagenaars, P., Plavšić, M., Sveaas, N., Wagner, U., Wainwright, T. (Eds.), *Human Rights and Human Rights Education for Psychologists*. New York: Routledge.
- Tostan, 2020. Community Empowerment Program. <https://www.tostan.org/programs/community-empowerment-program/program-structure/>. (Accessed 13 August 2020).
- UNESCO, 1974. Recommendation Concerning Education for International Understanding, Cooperation and Peace, and Education Relating to Human Rights and Fundamental Freedoms. Retrieved from. http://portal.unesco.org/en/ev.php-URL_ID=13088&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- United Nations, 1948. Universal Declaration of Human Rights. GA 217 A. United Nations, Geneva.
- UN, 1993. Vienna Declaration and Programme of Action. United Nations, Geneva.
- Zembylas, M., 2018. Reinventing critical pedagogy as decolonizing pedagogy: the education of empathy. *Rev. Educ. Pedagog. Cult. Stud.* 40 (5), 404–421.

Education, equality and international development

Helen Longlands and Amy North, IOE, UCL's Faculty of Education and Society, Department of Education, Practice and Society, University College London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	65
Framing education and international development	65
Quality, inequality and equality in education	68
Education, gender equality and measurement	69
Complex crises, emerging issues and challenges	71
Conclusion	72
References	72
Books	72
Articles in edited books	72
Articles in journals	72

Introduction

In this chapter we explore the relationship between education and development and how this has been considered in research and policy. We discuss key issues regarding how this relationship has been understood, conceptualized and monitored. In doing so, we pay particular attention to debates regarding gender equality and girls' education, and how these have been positioned in relation to a wider development agenda within global policy frameworks (such as the United Nation's (UN) Millennium Development Goals and subsequent Sustainable Development Goals), and in the work of international development organizations.

The chapter is organized as follows: first, we provide a brief overview of some of the key debates regarding the relationship between education and international development, and their position within global policy. We then turn our attention to key issues in relation to concerns with equality, inequalities and girls schooling, and look at approaches to measuring and monitoring these. Finally, we consider the implications of the current global context, and multiple, intersecting forms of crisis, for efforts to promote equitable education for all, in the context of wider global aspirations for sustainable development.

Framing education and international development

The 2030 Agenda for Sustainable for Development, which was adopted by all UN member states at the UN Summit for Sustainable Development in New York in September 2015, sets out a global vision for achieving "sustainable development in its three dimensions—economic, social and environmental" (UN, 2016), and identifies 17 Sustainable Development Goals (SDGs), which are seen as key to realizing this. Education holds a central place within the 2030 Agenda. It appears both as a goal in its own right—Goal 4 sets out 7 targets to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (see Fig. 1)—and as a key component of the targets of many of the other goals. Education is thus positioned both as an end and a means to development, and viewed as critical for achieving a wider vision of "ending poverty and hunger", "protecting the planet from degradation", ensuring "that all human beings can enjoy prosperous and fulfilling lives" and fostering "peaceful, just and inclusive societies" (UN, 2016).

The 2030 Agenda does not represent the first time that education has been positioned in this way within global policy on development. In the year 2000, when leaders from around the world came together in New York under the umbrella of the UN to agree the Millennium Development Goals (MDGs), which, as the predecessors to the SDGs, set out a series of goals and targets to be reached by the year 2015, education was similarly central, appearing both as an individual goal—Goal 2, to achieve universal primary education—and as a key target for Goal 3 on gender equality. Moreover, as noted by Unterhalter and North (2017, p. 2), "schooling and strategies to engage with poverty had featured prominently in post-colonial policy discussions around the world since the end of World War 2". However, while policy frameworks such as the SDGs and MDGs appear to reflect a growing consensus regarding the centrality of education within a wider global development agenda, how both education and development have been defined, and how the relationship between them is understood, is often highly debated and subject to multiple contestations and interpretations.

Development, as a term, is particularly contested. McCowan and Unterhalter (2022, p. 5) note that the emergence of the term "development" is sometimes associated with "the process of change of the less wealthy and often newly independent countries of Africa, Asia and Latin America, in conjunction with an increasingly elaborate architecture of development agencies", which arose in the period following World War 2. However, while conceptualizations of development, linked to ideas about "progress" and processes of industrialization and economic growth in lower-income countries in the global South, remain dominant in mainstream development literature, these have been challenged by a number of different groups and perspectives. Post-development,

The Sustainable Development Goals
Goal 1: End poverty in all its forms everywhere
Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
Goal 3: Ensure healthy lives and promote well-being for all at all ages
Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
Goal 5: Achieve gender equality and empower all women and girls
Goal 6: Ensure availability and sustainable management of water and sanitation for all
Goal 7: Ensure access to affordable, reliable, sustainable and modern energy for all
Goal 8: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
Goal 10: Reduce inequality within and among countries
Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable
Goal 12: Ensure sustainable consumption and production patterns
Goal 13: Take urgent action to combat climate change and its impacts
Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development
Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
Goal 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development

Fig. 1 The 17 Sustainable Development Goals (SDGs).

post-colonial, and decolonial scholars (e.g., Escobar, 2011; Sachs, 1997; Kapoor, 2008), for example, have argued for the need to understand development as a “eurocentric and hierarchic construct defining non-Western, non-modern, non-industrialized ways of life as inferior” (Ziai, 2017, p. 2547), and emphasized its role in the maintenance of colonial hierarchies and relationships (McCowan and Unterhalter, 2022). Meanwhile, although ideas regarding rights, equity and capabilities, as advanced in the work of scholars such as Amartya Sen (1999), have, since the 1990s, increasingly formed part of official development discourse, such normative concerns often sit uneasily alongside models and measures of development based on GDP and economic growth within the work of international agencies and other development actors (Chase and North, 2022). Education, too, can be interpreted in a range of different ways. While the term is often associated with the formal schooling of children (reflected in the focus on universal primary education in the MDG framework), a more expansive understanding of education also includes attention to learning that takes place in a wide range of formal, non-formal and informal settings (reflected in the emphasis in SDG 4 on *lifelong learning opportunities for all*), and which encompasses a range of different forms of knowledge, including non-western and indigenous knowledges (Chase and North, 2022).

These very different conceptualizations of what development and education are and mean have significant implications for understanding the relationship between education and development and how this has been conceptualized. Unterhalter (2022, p. 20), in outlining the emergence and evolution of the field of education and international development, identifies two contrasting approaches, which reflect “clear disputes with regard to how development is understood and what approaches to education policy, practice and research are entailed”. The first “mainstream” account of the education and international development field is linked to the emergence of development economics as a distinctive disciplinary area from the 1950s. Unterhalter argues that, in this account, development is understood primarily in terms of economic development, and education—mostly conceived of in terms of formal schooling—“is viewed as an aid to economic growth, which itself is initially linked with enhancing industrialization” (Unterhalter, 2022, p. 20). In this view, education is linked to supporting the “needs of the economy through building human capital and enhancing market relations” (Unterhalter, 2022, p. 22) and, more recently, “with enhancing relations of good governance, building institutions, and sometimes with addressing poverty” (Unterhalter, 2022, p. 22). Unterhalter describes such a concern with “making education effective”, which is often associated with the work of international organizations, including those associated with the UN as well as the World Bank and bi-lateral donor agencies, as the “what works” agenda, due to its core focus on “what works” for effective delivery of education and associated development outcomes.

A second, more critical account “draws out how ideas about education and development had been a strand in the expansion of colonial or imperial rule for many centuries” and sees education as a “key location for political, cultural and social contestation” (Unterhalter, 2022, p. 21). Education and international development as a field is thus linked with “many areas of subaltern and dominant knowledge, contested claims to authority, to land, to struggles over rights” (Unterhalter, 2022, p. 22). Here, education

is not simply understood in terms of the effective or efficient delivery of schooling but is a “contested terrain in which different values ... [and] stories of pain, dispossession, inequality and the effects of violence are articulated” (Unterhalter, 2022, p. 22). In this more critical view, the relationship between education and development is therefore a complex one, with attention paid to how different educational spaces and systems may play a role in reproducing or contributing to the transformation of different forms of inequalities and exclusion. Unterhalter refers to this understanding of education and international development as being associated with a “what matters” agenda, linked to concerns with supporting human rights, decolonial struggles, sustainability, and the promotion of gender and other forms of equality.

When, as 2015 approached, the global community reviewed progress in relation to the MDG goals and targets, it became apparent that, despite important gains in terms of expanding access to primary schooling around the world, and improvements in terms of the numbers of boys and girls enrolled in schooling, wider inequalities within and beyond education persisted. It was clear that the strategy of simply getting children into schools on its own was not enough to transform the wider dynamics of inequality and poverty, and the relationships that produce them, in order to deliver “development”. Rather, there was, it appeared, a need to move beyond a focus on “quick wins” (and “what works”) in terms of getting children into school, and to instead pay closer attention to work across sectors, to understanding context and the social and political relationships and dynamics of power in which educational systems and actors are located, and to engaging with the content of teaching and learning and the reproduction of inequalities within the education system itself. To some extent, this is reflected in the more expansive goals and targets associated with the 2030 Agenda, which pay attention to different spaces and levels of education, and place particular emphasis on issues of quality. SDG 4 (see Fig. 2) speaks to the need to consider different forms of inequality, and for educational spaces to be safe and inclusive. It also highlights issues of content, indicating that all learners should “acquire the knowledge and skills needed to promote sustainable development”. However, what this means in practice and how the aspirations set out in the

SDG 4: Ensure inclusive and quality education for all and promote lifelong learning:

Targets:

1. By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes
2. By 2030, ensure that all girls and boys have access to quality early childhood development, care and preprimary education so that they are ready for primary education
3. By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university
4. By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship
5. By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations
6. By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy
7. By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development

Means of implementation:

- a. Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all
- b. By 2020, substantially expand globally the number of scholarships available to developing countries
- c. By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States

See: <https://sdgs.un.org/goals/goal4>

Fig. 2 SDG 4 targets and means of implementation.

SDG targets can and should be implemented and monitored, particularly in complex and resource-poor contexts, remains an area of much debate.

In the next sections we turn our attention to some of these debates about quality, equality, and measurement, and consider, in particular, the focus on girls' education and gender equality, and how these have been positioned in relation to wider discussions regarding education in and for international development.

Quality, inequality and equality in education

Issues of power and powerlessness are embedded within the complexity of the field of education and international development outlined above. And concomitant concerns around "quality", "equality" and "inequality" in relation to education arise in varying, often contrasting and contested, ways from these longstanding and emerging debates about the meaning and nature of "development" and the associated positioning and role of education. Within the "mainstream" account of education and development, quality education has predominantly been conceptualized in instrumental terms, more closely linked with notions of efficiency and effectiveness and a focus on outputs and outcomes within formal schooling (Smith and Benavot, 2022; Wulff, 2020), underpinned by the "what works" agenda. From this perspective, equality in education has largely focused on equality in numbers, particularly in terms of equal access to education and the development (or not) of adequate skills for labor market participation through education (Sayed and Moriarty, 2020).

By contrast, within the second more complex and critical account of development, ideas around what quality education entails are generally more holistic and less linear, incorporating the intrinsic value of education and a focus on inputs, activities and experiences of education. Quality education in this account also connects with a broader understanding of education beyond formal schooling, reflecting the "what matters" agenda. It is thus a more contested notion because different things often matter differently to different people within and across different contexts. Likewise, equality in this account is also a more complex and contested notion, linked with ideas around rights, social justice, inclusivity, identity, gender equality and women's empowerment, and decolonization (UNESCO, 2005; Wulff, 2020; Unterhalter, 2022). Of course, because the distinctions within and between these different accounts of education and international development are not necessarily clear cut or disconnected from one another, understandings of and concerns about quality and equality can also incorporate ideas linked both with "what works" and "what matters" (Unterhalter, 2022, p. 27).

In whichever way education and development is conceptualized, recognizing the multiple and multi-dimensional nature of inequalities in education is important because the different ways in which these inequalities operate have different implications for policy and practice (Stewart, 2002, 2011; Picketty, 2019; Unterhalter and North, 2017). Thus, we can think about inequalities as being horizontal (relating to groups, cultures and associated forms of inclusion and exclusion in education), vertical (relating to distributions of resources, such as how much schooling people receive) or process (relating to learning and teaching interactions or an individual's ability to exercise agency) (Unterhalter et al., 2022a). We can also think about inequalities as being intersectional, whereby inequalities linked with different forms of social structure, division and discrimination, such as class, caste, gender, poverty, ethnicity or disability, can combine or intersect and amplify each other (Crenshaw, 1991). And such insight can be crucial for determining those who are most disadvantaged or marginalized (Kabeer and Santos, 2017). Crossing through these different groupings, we can identify inequalities of opportunities in terms of access to and participation in the various spaces of formal, non-formal and informal education; inequalities in experiences within education, linked with teaching and learning practices, relationships and features of quality; and inequalities of outcomes, whereby, for example, some people attain less well or better than others, or experience discrimination in the labor market. Inequalities can also operate at different but often interconnected levels – at the individual, institutional, community, regional, national or international level, depending, for example, on the efficiencies, capabilities and financing of schools and education systems (Unterhalter, 2022). And while some inequalities are relatively easy to identify (e.g., inequalities in access to formal schooling) and/or may be temporary, other inequalities (e.g., inequalities in people's experiences in education, including their experiences of violence, or their ability to exercise agency) are less straightforward (Parkes et al., 2020b) and/or may be deeply embedded and persistent.

Across all countries, gender inequalities in education are among some of the most persistent as well as the most complex and challenging forms of inequality to identify and address. Not only has this been recognized as a major obstacle for sustainable development and peaceful societies (see SDG 2030 Agenda; UNDP, 2019, p. 147), but it is also a grave matter of social injustice. Gender equality has thus been a driving principle in international development policy and practice for many decades, with girls' education, in particular, often hailed a silver bullet for development, linked to the belief that if you "educate a girl, you educate a nation" (see, for example, the 2021 G7 Declaration on girls' education (UK Government, 2021; UNGEI, 2022).

Within global policy, the Universal Declaration of Human Rights first articulated the right to education, irrespective of sex, in 1948. This was reaffirmed by the UN General Assembly in the 1989 Convention on the Rights of the Child and by the international education community in both the 1990 World Declaration on Education for All and the 2015 Incheon Declaration, which supports the 2030 SDG Agenda. Key global declarations in support of gender equality in development more broadly include the 1979 Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) and the 1995 Beijing Declaration and Platform for Action, agreed in Beijing at the Fourth World Conference on Women, which is still widely considered one of the world's most comprehensive and progressive proposals for advancing women's rights and gender equality (UN Women, 2014).

Nevertheless, despite these significant policy agreements, the achievement of substantive gender equality in and through education remains stubbornly elusive.

The many debates about definitions of gender and intersecting inequalities make the challenges for achieving transformational change in both the education field and the wider international development field even more acute. During the MDG era, the provision of free primary education to support children to attend and stay in school was a major strategy for addressing inequalities in education, including gender inequalities. Gender equality in education, meanwhile, was correspondingly seen as an issue of getting equal numbers of girls as boys in school, with an assumption that a number of beneficial social, health and economic outcomes associated with girls' education would follow (Unterhalter et al., 2022a)—an assumption supported by various quantitative studies which have outlined the beneficial effects of expanding girls' enrollment on a range of development outcomes including economic growth, fertility rates and infant mortality (e.g., Klasen, 2002; Psacharopoulos and Patrinos, 2018). Popular policies aiming to support girls' education have included conditional and unconditional cash transfers, scholarships and stipends, and free school meals (see Kremer et al., 2009; Baird et al., 2011; Baulch, 2011; Unterhalter et al., 2014; Sperling and Winthrop, 2016). Gender equality, in this sense, is narrowed down to the target of gender parity, which measures the numbers of girls/women as a proportion of the number of boys/men in different aspects of education, such as enrollment, attendance, attainment or teacher training, and provides some insight into the allocation and use of educational resources (North and Longlands, 2017). But, as we discuss below, while important, gender parity alone cannot measure the complexity of gender inequalities in and through education and is thus not sufficient to adequately inform strategies for the achievement of more substantive gender equality (Bardhan and Klasen, 1999; Unterhalter, 2015; UNESCO, 2016; Unterhalter et al., 2022a).

By 2015, although 66% of countries had achieved gender parity in primary education, 5 million more primary aged girls than boys were still out of school (UNESCO, 2016). Moreover, the focus on universal primary education was not accompanied by a similar push to provide quality education or quality educational experiences, or to ensure inequalities at other levels of education were addressed. Improvements in adult literacy rates between 2000 and 2015 (the end of the MDGs), for example, were negligible, despite being a key goal agreed by the international community—in 2014, it was estimated that 758 million adults still lacked basic literacy skills, 63% of whom were women (UNESCO, 2016; North and Longlands, 2017). By 2015, it was thus clear that significant barriers and crosscutting issues internal and external to education had not been addressed and were having the greatest impact on the most marginalized children and adults, particularly women and girls (North and Longlands, 2017).

In developing insight into the failures of the MDG period, a number of recent comprehensive reviews on interventions for gender equality and girls' education (e.g., Unterhalter et al., 2014; Sperling and Winthrop, 2015; Parkes et al., 2016; Perezniето et al., 2017) highlight the significance of the disconnect between the gender and education policy community and scholars and activists working around gender equality and women's rights, which resulted in a squandered opportunity to develop more nuanced and sophisticated understanding of inequalities in and through education and effective ways to address them (North and Longlands, 2017; Unterhalter et al., 2022a). As we explore in the next section, this has significant implications for informing development planning for policy and practice, and for identifying the kinds of information and data that are needed to monitor and evaluate progress toward sustainable and equitable development.

Education, gender equality and measurement

Detailed information drawing on a range of data is central for assessing whether progress toward achieving gender equality and sustainable development within and beyond education has or has not happened and for holding those in power accountable (Longlands et al., 2021). Yet data and measurement frameworks (including the SDGs) that have been used to monitor change in these processes have remained disappointingly inadequate due to lack of resources, definitional challenges, and debates over both the problems and the solutions (Unterhalter, 2017; Wulff, 2020). What gender equality in education means, how it is understood and articulated through policy, and how it can be measured is far from straightforward. As a result, gender parity has remained the most widely used measure to describe features of gender and education (Unterhalter et al., 2022a).

Gender parity is conceptually straightforward, simple to understand and communicate, and largely relies on administrative data, such as enrollment or completion rates, that are relatively easily and routinely collected by schools or education authorities. As such, it has been an important metric for governments and practitioners for monitoring and communicating improvements in education participation, progression and academic achievement: aspects linked with the "what works" development agenda in particular. National education sector plans (ESPs), international assessment frameworks,¹ and the SDG indicators for the targets associated with SDG 4 all use gender parity as a key indicator, despite SDG 4 presenting a broader vision of gender equality in education beyond parity. But as has been widely documented, gender parity is limited in what it can tell us about the many diverse forms and sources of gender discrimination and injustice in education systems linked with, for example, policies, pedagogy, curricula and violence, or the ways in which gender intersects with disability, poverty, ethnicity or other dimensions of disadvantage – issues that impact on people's, particularly women's and girls', day to day lives and their experiences and opportunities in and through education (Parkes, 2015; UNESCO, 2020; Unterhalter et al., 2022a).

¹These include PISA (Programme for International Student Assessment) and PIRLS (Progress in International Reading and Literacy Study).

Adequately identifying and measuring forms of gender (in)equality in education beyond gender parity is thus vital for gaining more comprehensive insight into gender inequalities in education in order to meaningfully inform effective development strategies for quality education linked both with “what works” and “what matters” agendas. Yet, as Unterhalter et al. (2022a) outline, designing a measurement framework that can do this is challenging and requires engaging with significant conceptual, political and practical difficulties. Concepts are required that can review a wide range of interconnecting relationships between individuals, institutions and processes, and consider both education systems as well as the ways in which education connects with other areas of human development (Unterhalter et al., 2022a). Moreover, the multiplicity of sites of gender inequality, complex connections between forms of inequality and the need to build alliances to support change at multiple levels, requires a measurement framework that can be multifaceted and responsive to different individuals, institutions and education systems within and across diverse contexts. And in order to develop such a framework, a wide range of debate and discussion is important for understanding these inequalities locally and nationally, for gaining insight into how gender intersects with other forms of inequality in these settings, and for assessing relevant data for identifying appropriate indicators that can measure change.

The Accountability for Gender Equality in Education (AGEE)² project, coordinated by a team of researchers at University College London, has been engaging with these complexities to develop an innovative and comprehensive indicator framework for measuring gender inequalities and progress toward gender equality in education and sustainable development. The work of the AGEE project draws on the concepts and principles of the capability approach, which support alternative understandings and measures for gender inequalities in education beyond gender parity. The capability approach is a way of evaluating equality and social justice developed by Amartya Sen (e.g., 1999, 2011), Martha Nussbaum (e.g., 2011) and other human development scholars, such as Ingrid Robeyns (e.g., 2017). While this approach has informed additional key measurement frameworks within the field of international development, such as the Human Development Index, that attempt to capture a more holistic insight into intersectional inequalities,³ the AGEE project presents the first time such a framework has been developed specifically in the field of education.

The capability approach recognizes the effects of social, environmental, institutional, and cultural structures on individual people’s agency, choices and wellbeing and is concerned with evaluating what people are able to do and the lives they are able to lead. In terms of gender equality and education, this approach entails paying attention to the complexities of the physical, political and social environment and resource distribution, and the differences in how these can be converted into individual freedoms and opportunities within and through education (Unterhalter et al., 2022a). In order to understand these complexities, a key principle of the capability approach is that collective deliberation and democratic participation must be essential components of any development process.

Thus, the AGEE project team, in consultation with key stakeholders from national and local governments, bilateral and multilateral organizations, civil society organizations, researchers, practitioners and activists, has developed the AGEE Framework. This measurement framework is designed to facilitate, at local, national and international levels, multilayered analysis of gender inequalities in and through education, and the conditions that underpin them. This represents significant progress from the basic and often distorting analyses of gender inequalities in education that have relied on gender parity measures.

The AGEE Framework⁴ identifies six key interconnected domains, in which indicators are selected that take account of conditions both inside and outside schools linked with opportunities, freedoms and agency, as well as resources and measures of subjective wellbeing. The AGEE domains are: Resources (goods and services required for a gender equitable education system); Values (the context of how policy is understood, operationalized, or contested); Opportunities (factors that can constrain or enable gender equalities in education and the ways in which these are differently converted into actual opportunities by individuals and/or institutions); Participation in Education (gender differences in the capability to participate and progress in all levels of education); Knowledge, Understanding and Skills (learning areas and levels for girls and boys in relation to formal instruction and informal formation of attitudes), and; Outcomes (gender equality in, and support for, areas of human development, including girls’ and women’s health, employment, access to support to protect from violence, political participation, ownership of assets, and legal standing). Combined, the six AGEE domains allow for assessment of horizontal, vertical, and process inequalities, as well as how successful strategies and interventions for addressing problems have been (Unterhalter et al., 2022a).

While work on the AGEE Framework and associated indicator selection is ongoing, gathering information across each of the six domains can help inform more thoughtful and locally appropriate development strategies that can capture issues distinct to education systems as well as issues that arise through the ways in which education connects with other areas of human development (Unterhalter et al., 2022a). Such engagement with the complexities of identifying, measuring and addressing the many forms of gender inequality and associated forms of injustice within and beyond education is a significant leap forward for the international development community. It is a vital component for ensuring the many international and national agendas and promises for equitable education, lifelong learning, gender equality and sustainable development are fulfilled.

²See: gendereddata.org.

³These frameworks include: the Human Development Index (HDI) (<https://hdr.undp.org/en/content/human-development-index-hdi>), the Gender Inequality Index (GII) (hdr.undp.org/en/content/gender-inequality-index-gii), the Multidimensional Poverty Index (MPI) (<https://ophi.org.uk/multidimensional-poverty-index/>), the OECD’s Social Institutions and Gender Index (SIGI) (genderindex.org) and the Women’s Empowerment in Agriculture Index (WEAI) (<https://www.ifpri.org/project/weai>).

⁴See <https://www.gendereddata.org/framework/>.

As we explore in the next section, the various environmental, political, economic and health related challenges the world currently faces makes the urgency and significance of this work particularly acute.

Complex crises, emerging issues and challenges

The increasingly complex current global context, and the different forms of intersecting crisis, associated with, for example, COVID-19, the impact of climate change, and crises linked to conflict, displacement, or poverty in particular settings, have significant implications for the delivery of education and for monitoring and evaluating progress, especially in resource-poor contexts, which are often the targets of development interventions. The existing literature points to a number of ways in which a range of economic, political, climate and health crises have contributed to exacerbating gender inequalities in and beyond education. Studies in refugee and conflict-affected contexts, for example, have highlighted high risks of sexual violence and school drop-out among girls displaced by conflict (Falb et al., 2016; Schulte and Rivzi, 2012); while research following the 2004 Indian Ocean Tsunami highlighted increased barriers to girls' participation in schooling in affected areas (Zulfika et al., 2009; Ariyabandu, 2006).

More recently, an emerging body of research has explored the interactions between the gendered impacts of climate change and girls' schooling. Several studies have pointed to the significant implications of the changing climate for girls' and women's participation in education (see, for example, Rao et al., 2019a,b). This is highlighted in a recent review of the literature on girls' education and climate change (Sims, 2021), which suggests that climate change "threatens to reverse developmental gains made in access to education over the past couple of decades" (p. 3). The report highlights the multiple ways in which climate change, and associated extreme weather events, may disrupt learning, particularly in low- and middle-income contexts, and notes the "disproportionate impact on girls as barriers to their education are multiplied by the effects of climate change and environmental degradation" (p. 3). Such gendered barriers are linked to a range of factors including "an increase in household chores and responsibilities, which are typically allocated to girls in contexts where there are gendered norms; decisions to withdraw girls from education, prioritizing their male siblings' education; and, in some contexts, difficulties in managing menstrual hygiene" (Sims, 2021). Such impacts are felt most acutely by poor girls; the report (citing Atkinson and Bruce, 2015) notes that it is estimated that at least 200 million adolescent girls living in the poorest communities face heightened risk from the effects of climate change.

The global health crisis associated with the COVID-19 has also had significant—and gendered—implications for education in low- and middle-income country contexts. While research is still emerging about the impacts of COVID-19 on education around the world, research concerned with experiences of previous health crises, including those linked to Ebola and HIV and Aids, have highlighted the gendered nature of their impacts on aspects of education and schooling for girls. For example, the HIV health crisis (c. 1990–2010) brought attention to School Related Gender Based Violence (SRGBV), highlighting the potential for schools to intensify gendered forms of inequality as well as build equalities (Aikman et al., 2008; Morrell et al., 2009; Parkes, 2015). Research on the Ebola crisis, meanwhile, has highlighted increases in schoolgirl pregnancy, and incidences of violence against women and girls (Unterhalter et al., 2022b).

A systematic review of the evidence on the impact of school closures and children being out of school for extended periods due to environmental disasters, pandemics and violent conflict "points to these processes deepening gendered and intersecting exclusions and vulnerabilities" with disruptions in learning having particularly adverse effects on children from the most vulnerable socio-economic groups (Unterhalter et al., 2022b, p. 9), and initial research on the educational impacts of the Covid-19 pandemic, and the impact of school closures in particular, suggests similarly gendered patterns. Research commissioned by UNESCO (2021) on the impact of school closures due to Covid-19 in Bangladesh, Kenya, Mali and Pakistan, for example, found that school closures have gendered effects in relation to four key areas. Firstly, participation in remote learning was affected by gendered household demands, with girls experiencing particular increases in domestic responsibilities; secondly, girls' learning was constrained by a gendered digital divide, with girls less likely than boys to have access to the Internet or a mobile phone; thirdly, available data suggests gender disparities in return to school, with girls in Kenya less likely to return than their male counterparts; finally, the report highlights the pandemic's impacts on wellbeing beyond school, noting higher levels of stress, anxiety and depression reported among girls. A World Bank study (Dang et al., 2021), which analyses household level data from six countries in sub-Saharan Africa, similarly highlights how the crisis has deepened educational inequalities, as the poorest households were least able to provide learning activities for their children during school closures, although it does not consider the gendered dynamics of this.

It is clear that the complex and intersecting impacts of the current global context present significant challenges for the equitable delivery of education in resource poor settings, many of which were already grappling with existing crises. However, there has been limited work on joined-up actions, monitoring or research, which attend to the interactions between different forms of crisis and their wider implications for education and development processes. Paying closer attention to how we understand and monitor the impacts of these intersections is clearly critical in order to "build back better" and realize the potential of education to contribute to transforming gendered and other inequalities, as envisaged in the 2030 Agenda. This is essential if the global community is to improve on its track record with the MDGs and meet the more expansive aspirations for sustainable development and global flourishing set out in the SDGs.

Conclusion

This chapter has provided some insight into the ways in which, over the past few decades, the relationship between education and development has been conceptualized and approached through policy, practice and research. In doing so, it has outlined some of the longstanding and emerging challenges linked with identifying, understanding and addressing inequalities associated with education, particularly gender inequalities—challenges that are not easily overcome. As the contemporary period of complex crisis makes strikingly clear, however, the significance of education for reframing how we think about and value each other and the world we live in, and for addressing the many short- and long-term impacts of inequalities for current and future generations, must not be underestimated. The challenge going forward is to develop coordinated, collaborative and multi-layered approaches that build on examples of good practice and take account of what research and experience tells us to develop strategies that can truly deliver substantive and sustainable change and turn global and national commitments on equality in and through education into meaningful action and practice on the ground. This requires engaging with different viewpoints, building alliances, and channeling sufficient resources into local, national and international efforts, so that rather than continuing to simply focus on changing what is easy, we can together understand and change what really matters.

References

Books

- Alkman, S., Unterhalter, E., Boler, T. (Eds.), 2008. *Gender Equality, HIV, and AIDS: A Challenge for the Education Sector*. Oxfam, Oxford.
- Escobar, A., 2011. *Encountering Development*. Princeton University Press, Princeton.
- Kapoor, I., 2008. *The Postcolonial Politics of Development*. Routledge, London.
- Morrell, R., Epstein, D., Unterhalter, E., Bhana, D., Moletsane, R., 2009. *Towards Gender Equality: South African Schools During the HIV and AIDS Epidemic*. University of KwaZulu-Natal Press, Scottsville, South Africa.
- Parkes, J. (Ed.), 2015. *Gender Violence in Poverty Contexts: The Educational Challenge*. Routledge, Abingdon.
- Picketty, T., 2019. *Capital and Ideology*. Harvard University Press, Cambridge, MA.
- Robeyns, I., 2017. *Wellbeing, Freedom and Justice: The Capability Approach Re-examined*. Open Book Publishers, Cambridge.
- Sachs, W. (Ed.), 1997. *The Development Dictionary: A Guide to Knowledge as Power*. Zed Books, London.
- Sen, A., 1999. *Development as Freedom*. Alfred Knopf, New York.
- Sen, A., 2011. *Peace and Democratic Society*. Open Book Publishers, Cambridge.
- Schulte, J., Rivzi, Z., 2012. *Search of Safety and Solutions: Somali Refugee Adolescent Girls at Sheder and Aw Barre Camps, Ethiopia*. Women's Refugee Commission, New York.
- Unterhalter, E., North, A., 2017. *Education, Poverty and Global Goals for Gender Equality: How People Make Policy Happen*. Routledge, London and New York.
- Wulff, A. (Ed.), 2020. *Grading Goal Four: Tensions, Threats and Opportunities in the Sustainable Development Goal on Quality Education*. Brill Publishers, Leiden, Netherlands.

Articles in edited books

- Chase, E., North, A., 2022. Introduction. In: North, A., Chase, E. (Eds.), *Education, Migration and Development: Critical Perspectives in a Moving World*. Bloomsbury Publishing, London.
- McCowan, T., Unterhalter, E., 2022. Introduction. In: McCowan, T., Unterhalter, E. (Eds.), *Education and International Development: An Introduction*. Bloomsbury Publishing, London.
- North, A., Longlands, H., 2017. Gender, poverty and educational equality. In: Scheulka, M.J., Johnstone, C.J., Thomas, G., Artilles, A.J. (Eds.), *The Sage Handbook of Inclusion and Diversity in Education*. Sage, London.
- Sayed, Y., Moriarty, K., 2020. SDG4 and the "education quality turn": prospects, possibilities and problems. In: Wulff, A. (Ed.), *Grading Goal Four: Tensions, Threats and Opportunities in the Sustainable Development Goal on Quality Education*. Brill Publishing, Leiden, Netherlands.
- Smith, W.C., Benavot, A., 2022. Quality education and global learning metrics. In: McCowan, T., Unterhalter, E. (Eds.), *Education and International Development: An Introduction*, second ed. Bloomsbury, London.
- Unterhalter, E., 2022. Histories of the field of education and international development. In: McCowan, T., Unterhalter, E. (Eds.), *Education and International Development: An Introduction*. Bloomsbury Publishing, London.

Articles in journals

- Ariyabandu, M.M., 2006. Gender issues in recovery from the December 2004 Indian Ocean tsunami: the case of Sri Lanka. *Earthq. Spectra* 22 (3_Suppl. 1), 759–775.
- Atkinson, H.G., Bruce, J., 2015. Adolescent girls, human rights and the expanding climate emergency. *Ann. Glob. Health* 81 (3).
- Baird, S., McIntosh, C., Ozler, B., 2011. Cash or condition? Evidence from a cash transfer experiment. *Q. J. Econ.* 126 (4), 1709–1753.
- Bardhan, K., Klasen, S., 1999. UNDP's gender-related indices: a critical review. *World Dev.* 27 (6), 985–1010.
- Baulch, B., 2011. The medium term effect of the primary education stipend in rural Bangladesh. *J. Dev. Effect.* 3 (2), 243–262.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299.
- Dang, H.-A., Oseni, G., Zezza, A., Abanokova, K., 2021. Impact of COVID-19 on Learning: Evidence from Six Sub-Saharan African Countries. The World Bank, Washington. Available online: World Bank Document. (Accessed 1 February 2022).
- Falb, K.L., Tanner, S., Ward, L., Erksine, D., Noble, E., Assaznew, A., et al., 2016. Creating opportunities through mentorship, parental involvement, and safe spaces (COMPASS) program: multicountry study protocol to protect girls from violence in humanitarian settings. *BMC Publ. Health* 16 (1), 231.
- Kabeer, N., Santos, R., 2017. *Intersecting Inequalities and the Sustainable Development Goals: Insights from Brazil*. International Inequalities Institute, LSE, Working Paper 14. The London School of Economics and Political Science, London.
- Klasen, S., 2002. Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *World Bank Econ. Rev.* 16 (3), 345–373.
- Kremer, M., Miguel, E., Thornton, R., 2009. Incentives to learn. *Rev. Econ. Stat.* 91 (3), 437–456.
- Longlands, H., Unterhalter, E., Peppin Vaughan, R., 2021. COP26 and gender equality in education: mobilising data for sustainable change. In: CEID Blog Series. Centre for Education and International Development, University College London. Online at: <https://blogs.ucl.ac.uk/ceid/2021/11/09/longlands-et-al/>.
- Parkes, J., Datzberger, S., Howell, C., Knight, L., Kasidi, J., Kiwanuka, T., Nagawa, R., Naker, D., Devries, K., 2020b. *Young People, Inequality and Violence During the COVID-19 Lockdown in Uganda*. UCL, London.

- Parkes, J., Heslop, J., Westerveld, R., Unterhalter, E., 2016. A Rigorous Review of Global Research Evidence on Policy and Practice on School-Related Gender-Based Violence. UNICEF, New York.
- Parkes, J., Ross, F.J., Heslop, J., 2020b. The ebbs and flows of policy enactments on school-related gender-based violence: insights from Ethiopia, Zambia, Cote d'Ivoire and Togo. *Int. J. Educ. Dev.* 72, 102–133.
- Pereznielo, P., Magee, A. and Fyles, N., 2017. Evidence Review: Mitigating Threats to Girls' Education in Conflict-Affected Contexts: Current Practice. New York: UNGEI and ODI.
- Psacharopoulos, G., Patrinos, H.A., 2018. Returns to Investment in Education: A Decennial Review of the Global Literature. Policy Research Working Paper; No. 8402. World Bank, Washington, DC.
- Rao, N., Lawson, E.T., Raditloaneng, W.N., Solomon, D., Angula, M.N., 2019a. Gendered vulnerabilities to climate change: insights from the semi-arid regions of Africa and Asia. *Clim. Dev.* 11 (1), 14–26.
- Rao, N., Mishra, A., Prakash, A., Singh, C., Qaisrani, A., Poonacha, P., Bedelian, C., 2019b. A qualitative comparative analysis of women's agency and adaptive capacity in climate change hotspots in Asia and Africa. *Nat. Clim. Change* 9 (12), 964–971.
- Sims, K., 2021. Education, Girls' Education and Climate Change. K4D Emerging Issues Report 29. Institute of Development Studies.
- Sperling, G. and Winthrop, R., 2015. What Works in Girls' Education: Evidence for the World's Best Investment. Washington, DC: Brookings Institution Press.
- Sperling, G.B., Winthrop, R., 2016. What Works in Girls' Education: Evidence for the World's Best Investment. Brookings Institution Press, Washington, DC.
- Stewart, F., 2002. Horizontal Inequalities: A Neglected Dimension of Development. Working Paper No. 1. Centre for Research on Inequality, Human Security and Ethnicity (CRISE), Oxford.
- Stewart, F., 2011. Horizontal Inequalities as a Cause of Conflict: A Review of Crise Findings. World Bank, Washington, DC.
- UK Government, 2021. Prime Minister Appoints New Envoy on Girls' Education: 16 January 2021. Press Release. Prime Minister's Office, 10 Downing Street, online at: <https://www.gov.uk/government/news/prime-minister-appoints-new-envoy-on-girls-education-16-january-2020>.
- UN, 2016. Transforming Our World: The 2030 Agenda for Sustainable Development. Available online: <https://sdgs.un.org/2030agenda>.
- UN Women, 2014. Introduction. Beijing Declaration and Platform for Action; Beijing+5 Political Declaration and Outcome. Online at: <https://www.unwomen.org/en/digital-library/publications/2015/01/beijing-declaration>.
- UNDP, 2019. Human Development Report 2019: Beyond Income, Beyond Averages, Beyond Today: Inequalities in Human Development in the 21st Century. United Nations Development Programme, New York.
- UNESCO, 2005. EFA Global Monitoring Report 2004/2005: Education for All—the Quality Imperative. UNESCO, Paris.
- UNESCO, 2016. Global Education Monitoring Report, Gender Review: Creating Sustainable Futures for All. UNESCO, Paris.
- UNESCO, 2020. 'A New Generation: 25 Years of Efforts for Gender Equality in Education', Global Education Monitoring Report Gender Review. Paris: UNESCO.
- UNESCO, 2021. When Schools Shut: Gendered Impacts of Covid-19 School Closures. UNESCO, Paris. Available online: UNESCO 379270eng.pdf. [unesci.org](https://www.unesco.org). (Accessed 31 January 2022).
- UNGEI, 2022. Unlocking the Gender Transformative Power of Education: Why Do We Need Gender Equality in Education? Online at: <https://www.ungei.org>.
- Unterhalter, E., 2015. Measuring Gender Inequality and Equality in Education, Concept Note for UNGEI Workshop. UN Girls' Education Initiative, London and New York.
- Unterhalter, E., 2017. Negative capability? Measuring the unmeasurable in education. *Comp. Educ.* 53 (1), 1–16.
- Unterhalter, E., Longlands, H., Peppin Vaughan, R., 2022a. Gender and intersecting inequalities in education: reflections on a framework for measurement. *J. Hum. Dev. Capab.* Online at: <https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F19452829.2022.2090523&area=0000000000000001>.
- Unterhalter, E., Howell, C., Mutanga, O., Robinson, L., North, A., 2022b. A Systematic Review of the Evidence on the Gendered Impacts of School Closures. UNESCO, Paris.
- Unterhalter, E., North, A., Arnot, M., Lloyd, C., Moletsane, R., Murphy-Graham, E., Parkes, J., Saito, M., 2014. Girls' Education and Gender Equality: Education Rigorous Literature Review. DfID, London.
- Ziai, A., 2017. Post-development 25 years after the development dictionary. *Third World Q.* 38 (12), 2547–2558.
- Zulfika, S.F., Soliman, I., Tamatea, L., Ware, H., 2009. Barriers to Girls' Accessing Primary Schooling in Post-Tsunami Sri Lanka: A Critical Ethnography of the Kalmunai Education Zone.

Knowledge/power and education: ethnonationalism, globalism, and social justice

Alma Maldonado-Maldonado, Departamento de Investigaciones Educativas, CINVESTAV, Mexico City, Mexico

© 2023 Elsevier Ltd. All rights reserved.

Power	74
Knowledge	75
Ethnonationalism	77
Globalism (or globalization)	78
Social justice	79
Case studies	79
Andres Manuel López Obrador proposals	80
Gabriel Boric's proposals	80
References	81

This entry problematizes the relationships between education, knowledge and power but locates this debate in current relevant social processes, such as globalism, nationalism (ethnonationalism) and social justice. The present discussion provides concrete ways to analyze power in education. Power is assumed to be everywhere, but a guide on how to understand the operationalization of power is a necessary contribution, particularly in the field of education. The inclusion of this entry in the section on “diversity, democracy and social justice in education” provides different angles on the realities of territories where diversity is not always accepted, where democracy is not always an ideal and where social justice seems as though a utopic fantasy.

Ideologies and politics need to be taken seriously when educational change is at stake. The entry first presents the conceptual puzzle on education, knowledge, power, globalism, (ethno)nationalism and social justice. Then, it suggests ways to analyze their relationships and hierarchies. Finally, two case studies are presented and contrasted to show the applications and implications of this entry's content.

Power

Every social interaction is affected by power, and education is not an exception. *Discipline and Punish: The Birth of the Prison* by Michel Foucault was not the first book to discuss the issues of power and education, but it was a crucial text that broadened this debate. Other authors with critical views, the large majority of them from a post-Marxist approach, continued the conversation. These authors include Louis P. Althusser, Samuel Bowles, Herbert Gintis, Michael Apple, Thomas S. Popkewitz and Stephen J. Ball. What has been made clear is that a theory on power in education is missing, and what has been written needs to be updated and presented in a more organized manner.

Power can be identified in different ways as force, potency, act, possibility, faculty, capacity, action, authority, reactance, punishment and coercion. Also, power is related in one way or another to legitimacy, conflict, resistance and violence.

According to Giddens (1984), “Power’ is very often defined in terms of intent or the will, as the capacity to achieve desired and intended outcomes. Other writers by contrast, including both Parsons and Foucault, see power as above all a property of society or the social community” (p. 15). This assumes power in both senses, and it emphasizes the exercise of power as a human quality and a force present in all human activities. More importantly, power could help knowledge and education become spaces of social transformation.

In fact, few things are more powerful than education and knowledge. The act of deciding what should be taught to children and youngsters makes education one of the most important social activities. It is hard to find another more important activity in society. The intention of curriculum content is to establish what knowledge is valid, legitimate and useful for youth. Current debates have to do with what topics of sexuality, religion, race or history must be taught at schools, at what levels and in what types of institutions. Each country is waging its own internal battles over the degree to which its education is technological, general, Western oriented or liberal (Carnochan, 1993). Ideology and politics play major roles in such curriculum decisions. Additionally, these discussions have to do with levels (primary, secondary, high school, higher education) and age, up to the point of what professions should be taught, what disciplines must be included, or in a broader sense, what type of training citizens require.

This entry suggests a way to organize the debates on power and education. First, this work proposes four scales to understand power in education: agents, symbols, structure and processes (see Fig. 1). *Agents* include teachers, students, superintendents, deans, academics, unions, leaders, collectivities, professional associations, accreditors, academic associations, institutions, governments, countries, politicians, guilds, states, nation states, decision makers, bureaucracies, elites, civil society and stakeholders. *Symbols* include ideas, theories, policies, knowledge, ideologies, creativity, narratives, ideals, dreams, traditions, emotions, values, morals,

culture and prestige. *Structure* can include colonialism, language, rules, normativity, law, context, legacy, legitimacy, networks, economy, political systems, technology, gender, curriculum, degrees, credentials, money, infrastructure and disciplines. Finally, *processes* include learning, teaching, researching, evaluating, shaping behaviors and disciplining.

The inclusion of an element on one scale does not mean that this element cannot be part of another scale. In this scheme, agents and structure are obvious elements, since they are part of a classic sociological debate, but the inclusion of processes and symbols is an addition. The idea behind including processes is that they are exercised by agents in particular structures, but a process is difficult to capture as a single moment or circumstance. The same situation occurs with symbols. An element identified as a structure can have a symbolic heavy load, but it is important to try to differentiate between the different spaces occupied. In the case of symbols, this scale of power is less analyzed in education but one of the most powerful. This is problematic since symbols cannot be objective or measured; most symbols do not contain single meanings. Symbols are essentially embedded in the cultural background.

Second, there are some questions that will help reflect on power in education. [Zelditch and Walker \(2000\)](#) pointed out that the “adequate description of power relation will typically refer to: (1) the bases of power; (2) the means of power; (3) the strength of power; (4) the costs of power; (5) the amount of power; (6) the scope of power; (7) the domain of power” (p. 1457) ([Fig. 2](#)). This applies to education at its different levels (primary, secondary, tertiary, etc.), types (formal, informal and non-formal) and modalities (traditional face-to-face learning, distance learning, blended learning, home schooling, etc.) ([Fig. 2](#)). In the case of education, each of these aspects has different possibilities and could respond to a different question. The bases of power in education may respond to the following questions: What role does education play in society? What are the sources of legitimacy, and what is the origin of the power that every actor has in educational processes? The means of power in education are related to the following questions: What does power in education symbolize? What are the different ways power is held (prestige, authority, discipline or recognition)? Does the strength of power have to do with the different intensities of the forms of power in education? The cost of power is related to the following question: What does it take to hold power in the educational context? What is the price that different agents have to pay to hold such power? The amount of power in the context of education may respond to the following question: How much power can someone exercise? Is it possible to establish certain scales to determine the quantity for power? The issue of soft power may be incorporated here. The scope of power in education should answer the following: What is the range of power that an agent, a structure, a symbol or a process can maintain? Finally, the domain of power has to do with the spaces where power can be exercised. Can it be an institution, a country, an educational level, an educational system or a curriculum?

Power is everywhere but naming where and in what ways it is present is required. In this case, once there is a way to organize the debate on power and education, the second aspect is to consider the purpose of using power. Ideally, power must be used to challenge the present reality of education and to improve people's conditions.

Knowledge

The relationship between power and knowledge has been widely discussed in the literature. [Foucault \(1992\)](#) was one of the first to elaborate on the complexity of this relationship (as cited in [Eribon, 1992](#)). Knowledge is one of the most important human

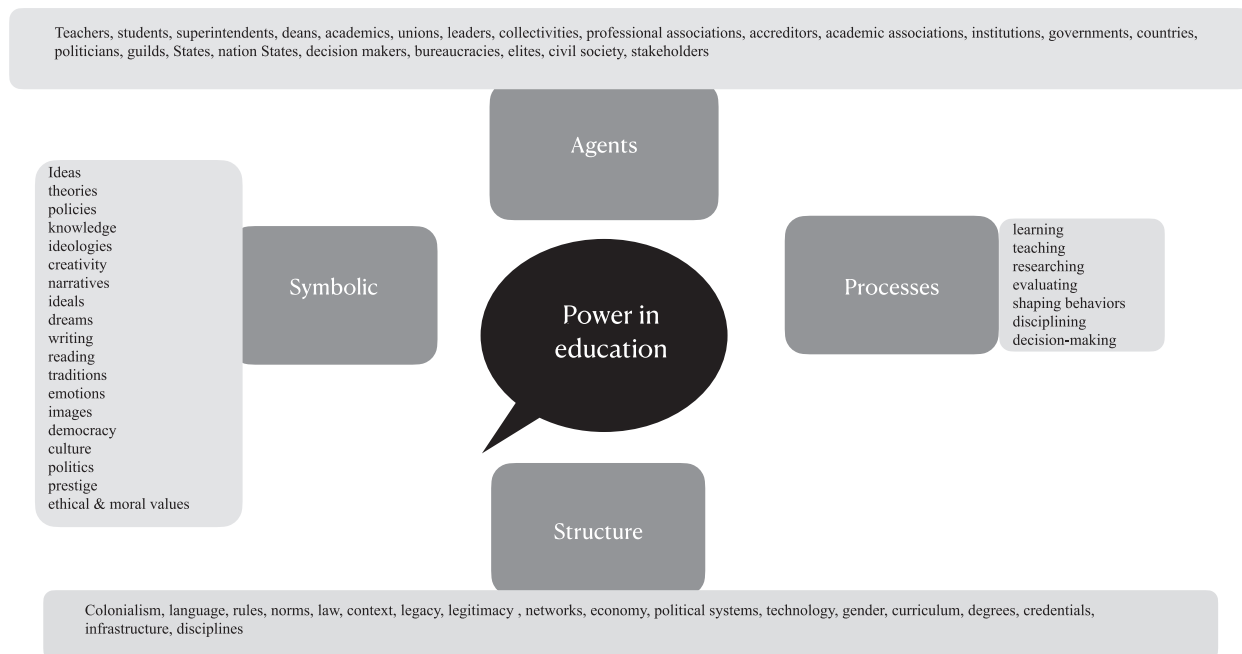


Fig. 1 Scales of power in education.

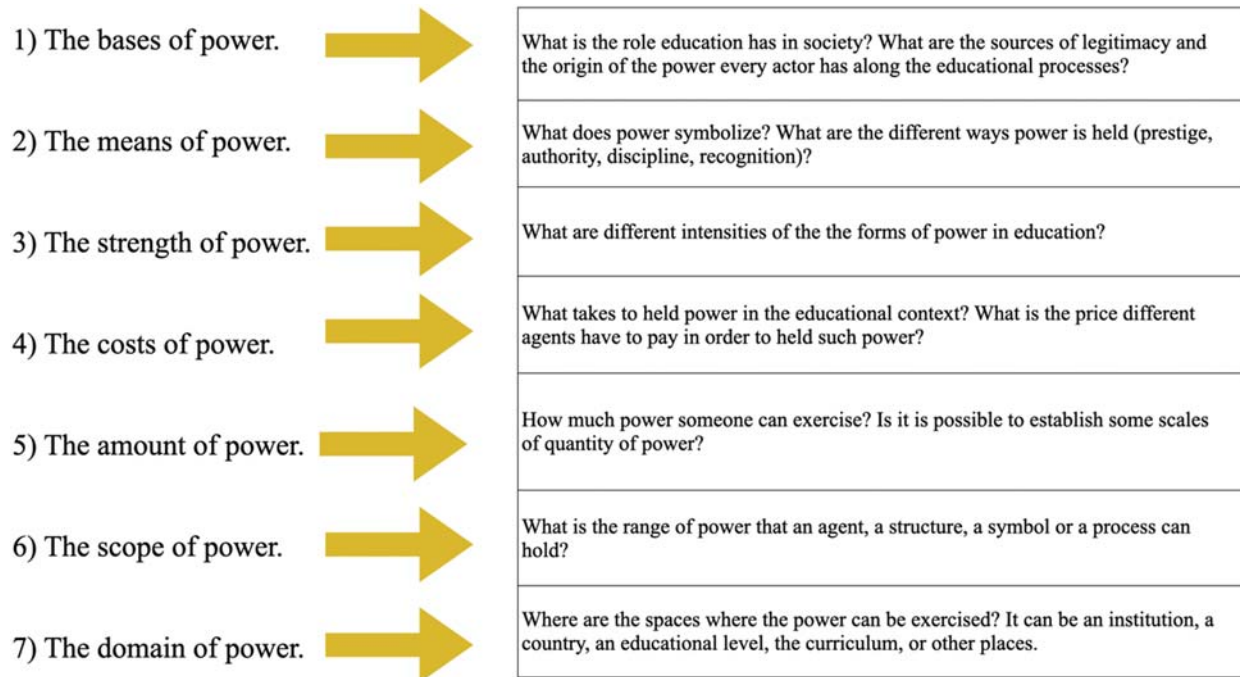


Fig. 2 Ways to analyze power in education. Based on Zelditch and Walker (2000).

characteristics, as it is attached to power in most forms. There are different types of knowledge, and it has a lot to do with legitimacy. Currently, possessing knowledge means acquiring power on different scales.

According to Plato, knowledge does not only function as reason or the state of pure intelligence; true knowledge also means the love to know and to acquire wisdom (Xirau, 1990). Knowledge is considered a commodity and a public good. This has also occurred with education, where depending on the level, education can be considered a public good, a commodity, a public service or a right. One difference between knowledge and education is accessibility. Knowledge has the quality to be a public good because its use does not prevent other people from acquiring it. Stiglitz (1999) pointed out that “a public good has two critical properties, non-rivalrous consumption—the consumption of one individual does not detract from that of another—and non-excludability—it is difficult if not impossible to exclude an individual from enjoying the good” (p. 1).

Knowledge can be considered a public good and subject to commodification as an educational product and as a product of academic research (Kauppinen, 2014). Indeed, the idea of knowledge as a commodity has become more common. The increased prevalence of this idea is related to the increase in market-oriented public policies. There is knowledge that represents more value than others. In particular, applied knowledge allows innovation and helps to develop industries. As an example of this value, the World Trade Organization (WTO) estimates that the revenue produced by the commercial transactions of intellectual property produced about \$429 billion (Maldonado-Maldonado and Cortes-Velasco, 2021). Countries tend to calculate intellectual property payments and receipts, which give an excellent idea of their relevance as not only knowledge consumers but also producers. However, this results in a paradox when considering the fact that “Knowledge knows no borders. It is freely circulated internationally, even when governments attempt to restrict its flow. The development of an international network of knowledge distribution is by no mean a new phenomenon” (Altbach, 1987, p. 65). On the one hand, the revenue produced by selling and buying knowledge is an economic measure and part of an international competition, especially among the richest and most powerful countries in the world, but on the other hand, knowledge has been part of international exchange and an element of international cooperation, and some of the most important scientific discoveries in recent times were made possible by historically established international collaborations.

In fact, the Organization for Economic Co-operation and Development (OECD) has coined the term “knowledge-based economy” to describe current economies’ dependence on knowledge production; the knowledge-based economy refers to a certain type of knowledge, particularly applied knowledge related to innovation and technology. According to the OECD (1996), there have been “trends in advanced economies toward greater dependence on knowledge, information and high skill levels, and the increasing need for ready access to all of these by the business and public sectors”.

Other international organizations have preferred to use the concept of “knowledge society”, which are

[...] Capabilities to identify, produce, process, transform, disseminate and use information to build and apply knowledge for human development. They require an empowering social vision that encompasses plurality, inclusion, solidarity and participation.

UNESCO (2005).

Knowledge is a social construction, and “knowledge is all around us” (Sloman and Fernbach, 2017, p. 13). These authors mentioned, “We have access to more knowledge today than ever before We live in a community of knowledge” (Sloman and Fernbach, 2017, p. 13). In many ways, the power of knowledge resides in such a collective characteristic. This is the same reason knowledge is used to label societies or economies that depend on its production to subsist, develop and make broader the differences among countries.

On the role of basic scientific knowledge and what it takes to produce it, Merton (1996) pointed out the following:

[...] Science is a deceptively inclusive word which refers to a variety of distinct though interrelated items. It is commonly used to denote (1) a set of characteristic methods by means of which knowledge is certified; (2) a stock of accumulated knowledge stemming from the application of these methods; (3) a set of cultural values and mores governing the activities termed scientific; or (4) any combination of the foregoing (p. 267).

Nowadays, knowledge is no longer produced exclusively at higher education institutions; there are new actors that include industries in the private sector, private laboratories or other private enterprises. However, basic science is still mostly produced at public and private higher education institutions where autonomy, academic freedom, graduate students, postdocs, financing, resources, infrastructure, the academic community and other elements facilitate the development of academic research and science. Furthermore, the knowledge production ecosystem goes beyond higher education institutions, research laboratories or centers; “It broadly includes government science ministries and research councils, certain enterprises and other private bodies, and supporting infrastructure” (OECD, 1996, p. 21).

According to Clark (1983), higher education is constituted by groups of people who carry knowledge. The more the knowledge is multiplied, refined and segmented, the more it becomes something peculiarly slippery and invisible (Clark, 1983). In the same sense, Altbach (1987) pointed out that “the educational system is one of the most important consumers and creators of knowledge” (p. xi). In particular, higher education institutions are part of the chain of knowledge distribution and some of the main publishers of scientific papers. Furthermore, “Universities are publishers, and such are direct contributors to the knowledge dissemination system. The most direct and obvious publishing activity of a university is that of the university press” (Altbach, 1987, p. 69).

Aside from scientific knowledge, there are other types of knowledge, including popular, Indigenous, intuitive and theological. Knowledge can be partial, limited, unlimited, unconditional or conditional. The OECD (1996, p. 12) highlighted four ways to refer to knowledge: know-what (facts), know-why (principles and laws of nature), know-how (skills or the capability to do something) and know-who (about the subject who knows how to do that).

There are important debates against the discrimination of certain types of knowledge, particularly Indigenous knowledge in colonized countries. Decolonial positions (Bidaseca, 2016; Dussel, 1988; De Sousa Santos, 2021) suggest that maintaining a single idea about knowledge excludes other cultural alternatives or possibilities in terms of its definition, validity and accuracy. The debate in this case has to do with whom decides on what type of knowledge is legitimate. According to Young (1995), “colonial discourse analysis was initiated as an academic sub-discipline within literary and cultural theory by Edward Said’s *Orientalism* in 1978” (p. 159). It seems to prevail a powerful Eurocentric vision about how to define knowledge, modernity and peripheries (Dussel, 1998, pp. 62–63).

One criticism decolonial authors have made has to do with the kind of voices that are represented in the debates and who has assumed the representation of “others”. As Young (1995) pointed out, “In recent years there has also been a growing unease about the tendency of anti-Eurocentric writing to homogenize not just the “Third World” but also the category of “the West” as such—by writers who are of course, very often themselves the product of that same ‘West’” (p. 165).

The dominant role of scientific knowledge over other types of knowledge cannot justify, by any means, the attack science has faced lately from religious fundamentalists, populist governments, dogmatic ideologies, authoritarian regimes, ethnonationalists and extremists, among others (Deaton, p. 326).

Ethnonationalism

The concept of ethnonationalism links the idea of the nation to a particular ethnic identity. This implies the possibility that nations share a common heritage. However, this assumption is problematic in different ways, particularly when it is used to justify wars, invasions or interventions in public policies to stop immigrants with different ethnic backgrounds from getting into a particular country because the dominant powers are afraid of what “the other” represents.

Ethnonationalism might imply the use of the same language, historical background, religion or territorial limits. It is even more problematic if it implies what can be considered a common genetic ancestry. In the preservation of these elements, the idea of ethnonationalism tends to exclude or reject the idea of multi-ethnic societies. Defining a single ethnic identity challenges the reality of geographical mobility and its impact on migration.

Nationalism is a movement defined by the construction of historical and cultural identities. Adding the ethnic aspect constrains this identity on racial or genetic characteristics that are biologically unsustainable. To provide a clear example, Nazism and other similar extreme movements have been nourished by ethnonationalist ideologies. Ethnonationalism can inspire xenophobic

responses. In less extreme cases, ethnonationalistic ideologies can provoke a hostile external environment, such as in a country like Korea, where they are surrounded by important geopolitical powers (Lee, 2009).

Young (1995) summarized how problematic the ideology of race is, as well as its implications and origins:

The ideology of race, a semiotic system in the guise of ethnology, “the science of races”, from the 1840s onwards necessarily worked according to a doubled logic, according to which it both enforced and policed the differences between the whites and the non-whites, but at the same time focused fetishistically upon the product of the contacts between them (p. 181).

Young (1995) also indicated that the idea of racism operates by “the determination of degrees of deviance in relation to the White-Man face” (p. 180), which means that the parameter of comparison, ethnically speaking, is a Eurocentric, malecentric and Whitecentric figure (always in prejudice to those who are different, to otherness). A particular ethnic group is more likely to be racist if this group feels a threat to its dominance with the incursion of other ethnic groups (Connor, 1984).

Similar to nationalism, ethnonationalism is closely related to the concept of the “self-determination” of nations (Connor, 1973). Self-determination was a way to affirm the existence of new countries. Similarly to defining a nation based on cultural heritage, defining it based on ethnic heritage is also problematic. As pointed out by Harari (2013), one key question is what legacy the nation decides to take, unless people are able to accept that it is normal to tend to choose the side of the victors or losers.

Migration, which is one key element of globalism, represents a major threat to ethnonationalism since migrants not only bring different cultures, but they look different in most cases. Migrants from developing countries are ethnically diverse compared to the traditional populations in developed countries, which are the places to which they migrate.

Globalism (or globalization)

Globalism or globalization contradicts the idea of ethnonationalism, and in general, it challenges the idea of nationalism. Globalization has to do with the diminishing of national borders, the regional integration of countries, the increase of commercial, educational, economic and political exchanges and individuals’ realisation of these processes and actions in consequence (Waters, 1995). Globalization has erased the borders between them and us, the near and the distant (Lipovetsky and Juvin, 2010). Globalism imposes the vision that the planet is one. There are global problems, global solutions, global organizations and global non-governmental organisations (NGOs) among other actors. Integration between regions and nations has increased in recent decades. Mass media, technology, transport and pop culture have accelerated the idea of worldwide interconnection (Murray and Overton, 2006).

Globalism does not mean the end of the state or the end of the idea of the nation, but as Castells (2000) argued, it does imply its transformation. This is where the element of sovereignty becomes relevant. What are the changes that national states are able to implement to become part of global processes? What are the types of concessions governments are able to make, and how do they exercise their power?

Globalism has winners and losers, and for a long period of time, many problems in the world have been attributed to this axis. This is explained by the fact that economic globalization has caused other forms of global processes, such as cultural or political globalization, to collapse (Noya and Rodríguez, 2010). The failures of the global market economy are considered a consequence of globalization. “Markets are supposed to be stable, but the global financial crisis showed that they could be very unstable, with devastating consequences” (p. xi), as Stiglitz (2012) pointed out. Castells (2000) clarified that the “global economy was not created by markets, but by the interaction between markets and governments and international financial institutions acting on behalf of markets—or of their notion of what markets ought to be” (p. 135). In any case, Deaton (2013) is very conclusive in this respect: “Globalization has made the world more unequal, that while the rich have been presented with new opportunities for getting richer, the poor of the world have gained little” (p. 257).

One of the main challenges of globalization did not come from left-wing movements but right-wing leaders, particularly populist leaders like Donald Trump (USA), Boris Johnson (UK), Jair Bolsonaro (Brazil), Narendra Modi (India) and Andrés López Obrador (Mexico). These leaders are nurtured by nationalism. Their position against globalization is to blame the global integration of the planet as the main reason for the national crisis in their countries and as a barrier against their progress.

Stiglitz (2002) argued that globalism is neither good nor bad. There are some positive aspects that have resulted from this phenomenon. For example, a global civil society has been formed to defend a global agenda, such as the fight against climate change. He sees the role of knowledge as another component that enables nations to get some advantages from globalism: “Globalization has reduced the sense of isolation felt in much of the developing world and has given many people in the developing countries access to knowledge well beyond the reach of even the wealthiest in any country a century ago” (Stiglitz, 2002, p. 4). In the same sense, Castells (2000) understands that “global technological development needs the connection between science, technology and the business sector, as well as with national and international policies” (p. 126).

Nevertheless, globalism is responsible, in part, for inequalities around the world. This is where the last concept in this entry enriches the debate.

Social justice

The scarcity of all types of goods and resources underlines the problem of what the best way to decide the fair distribution of them is. Depending on the goods and resources, scarcity can be strong or weak, natural, almost natural or artificial (Elster, 1992). The pursuit of justice is derived from the idea of fairness. There is a matter of deciding who should aspire to receive what benefits, what the role of social institutions is in making sure there is a fair distribution and what injustice a society is capable of admitting.

When the process of globalism started, it was believed that a new moral planetary consciousness was born. The possibility of creating global justice seemed realistic (Noya and Rodríguez, 2010); however, this illusion soon vanished. The question asked by Amartya Sen (2009) is very pertinent: “What international reforms do we need to make the world a bit less unjust?” (p. 25). In looking to answer this question, it can be noted that egalitarian values clash with libertarian values when interventions about how to mitigate inequalities have to be defined, including who needs to intervene to reduce gaps in equality in society (normally would be the state). Atkinson (2015) suggested that “the achievement of an equitable society depends to a considerable degree on the effectiveness of the public administration and the quality of its dealings with citizens” (p. 122). In the same sense, Stiglitz (2012) stated that “we have to rein in the excesses at the top, strengthen the middle, and help those at the bottom. Each goal requires a program of its own. But to construct such programs, we have to have a better understanding of what has given rise to each facet of this unusual inequality” (p. 29).

The ideal of social justice represents the possibility of change. Collier (2008) called attention to the bottom billion of the poorest people of the world who are trapped by socioeconomic and political conditions. The debate has focused more on inequality as the main problem to solve than on the face of social justice.

“Citizenship premium” is how Milanovic (2017) refers to the advantage someone has when the person is born in a rich country. The issue is how to avoid providing advantages to particular people just because they are luckier to be citizens of a country than others from a different nation or because they are born into a wealthier family than someone else. An accident, or fate, cannot be the only reason why some have better possibilities for development and better opportunities than others.

According to Milanovic (2017), malign and benign forces reduce inequality. Among the malign forces are wars or civil conflicts, and the benign forces include social pressure, increased education, aging and technological change—in other words, issues related to knowledge production. Piketty (2014) goes in the same direction:

Historical experience suggests that the principal mechanism for convergence at the international as well as the domestic level is the diffusion of knowledge. In other words, the poor catch up with the rich to the extent that they achieve the same level of technological know-how, skill, and education, not by becoming the property of the wealthy. The diffusion of knowledge is not like manna from heaven: it is often hastened by international openness and trade (autarky does not encourage technological transfer). Above all, knowledge diffusion depends on a country's ability to mobilize financing as well as institutions that encourage large-scale investment in education and training of the population while guaranteeing a stable legal framework that various economic actors can reliably count on. It is therefore closely associated with the achievement of legitimate and efficient government. (p. 71)

Most authors who discuss inequality agree with the importance of education, technology and knowledge to narrow inequity gaps. However, few have suggested specific ways in which this can be made possible. Even less of them have debated on the structural elements that preside over people's access to education, knowledge or technology, or recent topics such as the role of meritocracy and the real chances of having equal opportunities or an even starting point.

Atkinson (2015) wondered how much inequality is acceptable, and this is where the idea of social justice acquires relevance. Additionally, this is where power, knowledge and education became transcendent in looking for a fairer world. Sen (1999) would refer not only to looking for a better distribution of primary goods but to looking to provide capabilities to all humans, bolstering their chances to develop their potential.

Case studies

Chile and Mexico are two Latin American middle-income countries. Both are members of the OECD, and both have democratic regimes. However, there are many differences between the two. First, Chile has a population of about 19.1 million, and Mexico has a population of about 128.9 million. There are considerable geographical, economic and historical differences, but they also share the Latin American cultural identity, language, problems and challenges, and both nations were colonized by the Spaniards. In recent times, both countries have faced major political shifts with the election of left-wing presidents, but with different characteristics. Gabriel Boric, the president of Chile, was born in 1986. Andrés Manuel López Obrador was born in 1953. Boric was elected president when he was 36 years old. By the time López Obrador was 33 years old, Boric had just been born. Boric had a trajectory into politics as a student activist at the University of Chile, and López Obrador was a social activist part of the official political party that had been in power for 71 years (Partido Revolucionario Institucional, or PRI).

Both presidents have a strong agenda regarding social justice issues. This partially explains their electoral victories. Inequality is a major characteristic of most Latin American countries. There are very important segments of the population living in poverty. When Boric, López Obrador and other candidates in countries such as Colombia, Brazil, Argentina or Peru speak, they are talking

to the most vulnerable populations who have witnessed injustices, lived under the oppression of the elites, suffered scarcity and seen how they have been left behind.

Generally, political campaigners make promises, this is not new. In the case of the promises these candidates and later presidents have made, only their higher education agendas are considered here with an important clarification: by the time this entry was being written, López Obrador was in the middle of his term (third year out of six), and Gabriel Boric was just at the beginning (first year out of four), so it is hard to compare their results; instead, it is fairer to compare their proposals.

Andres Manuel López Obrador proposals

López Obrador made four key higher education-related proposals as a presidential candidate. The first was free education at all levels, including higher education and graduate education. Second, entrance examinations to higher education institutions should be eliminated. Third, which was the only one included in the official document “Nation Project 2018–2024”, autonomy should be provided to private higher education institutions. The fourth was the opening of 100 new universities (Universidades para el Bienestar Benito Juárez García). These universities were planned to be built in the poorest areas of the country (Rodríguez and Maldonado-Maldonado, 2019).

Gabriel Boric’s proposals

Boric’s presidential campaign identified six important issues. The first had to do with forgiving student debt and establishing a credit system in the short term to transition to universal free higher education. The second included rebuilding and expanding public free and quality higher education, including increasing public funds. Third, the goal was to improve free education and financing policies. Fourth was a new policy to promote access and permanence, which included the establishment of affirmative action policies. Fifth was to conceptualise and improve the comprehensive quality of higher education and refine gender policies. Sixth was transforming technical and professional education, modernising these fields and eliminating the for-profit technical and professional education (Boric, 2022).

In accordance with the four scales of power in education—agents, structure, symbolic and processes—Chile and Mexico represent two cases where the new leaderships are aiming to transform the social reality and looking to create a context where social justice should prevail. The four scales help clarify where power and education struggles take place. The agents are relevant and as follows: two leftist presidents, opposition leaders, the senate and legislatures, sectors traditionally excluded from higher education, elites, higher education institutions, academics, accreditation bodies or professional associations, among others. The processes refer to the different moments when these change projects are developing with respect to decision-making processes. The symbolic has to do with the ideologies and new narratives that have been built, such as changing the higher education systems in both countries and challenging the traditional practices of higher education institutions that have created major inequalities in access. The structures seem the most difficult change at a significant scale because they refer to the laws, norms and political and economic systems in these countries. Dissecting the different elements helps to understand the dimensions of the problem when it is a matter of creating change.

The scales, together with the questions proposed by Zelditch and Walker (2000), are as follows: What role does education play in Chile and Mexico? What are the sources of legitimacy for the two relatively new governments led by López Obrador and Boric? How do prestige, authority, discipline and recognition take place in the educational context of each society? How much power can both leaders exercise to push for changes in education? Among other questions, they can serve as a good guide for exploring the situation in each nation.

On the one hand, López Obrador decided to stay at the scale of the symbolic—the narratives and the ideals—but did not look to change structures. The policies that have been implemented are hardly going to change the stratification of the system or reduce the inequalities in access to the tertiary education sector in Mexico. For example, the so-called 100 new universities cannot be considered a success, as they do not have enough resources, the program offered are very limited, they are not located in the poorest regions of the country and they do not have enough faculty, students and infrastructure to operate, but they have been promoted as an important governmental achievement. There are a lot of concerns about the outputs of the policies implemented by the López Obrador government, particularly because the promise of free higher education seems unreachable in a country that has not looked for fiscal reform (despite the fact that his government does have the majority to pass it) and whose economic policies are continuing to reduce public spending in education and science.

On the other hand, Gabriel Boric is looking to change the taxation system and the country using a previously decided process to reform its constitution in a process that had been initiated before the presidential election. The current constitution is a relic of Augusto Pinochet, the dictator who ruled Chile for 17 years. After several political movements (some led by student activists such as Boric), most Chileans decided that on September 4th 2022, there would be a plebiscite to vote for the new constitution. The project includes 388 articles, which, if approved, would be one of the largest constitutions in the world (Propuesta: Constitución política de la República de Chile, 2022). The Constitution proposal establishes that the State will ensure access to higher education for all persons who fulfill the requirements. Access will be governed by equity and inclusion principles, focusing on historically

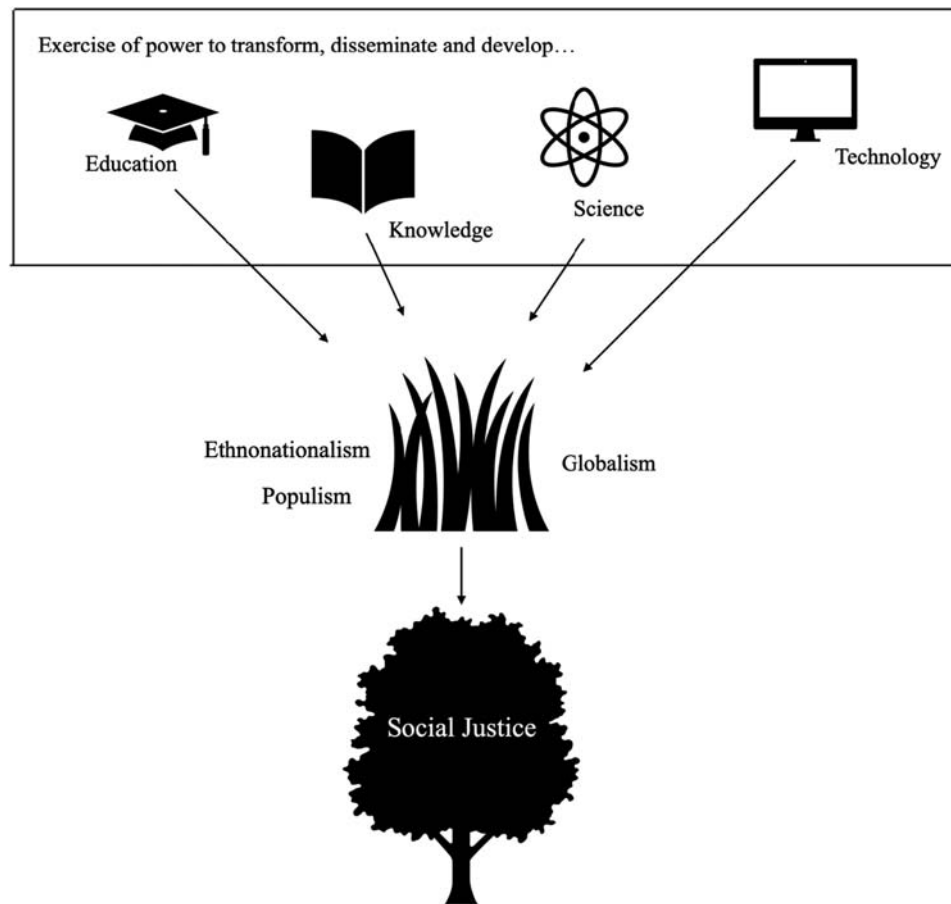


Fig. 3 Knowledge/power and education: ethnonationalism, globalism and social justice.

excluded groups. The fiscal reform, which looks to tax the rich to finance social rights, does not have an easy road to approval, since Boric does not have the majority in Congress. His government is in the transition phase from symbolic to structural changes.

The four scales proposed in this entry do not suggest that there is a route to follow to instigate an educational revolution, but they serve as a map to see how an educational system (in this case) operates and how to understand the roles of agents or the timing of the processes.

Populist narratives are very attractive. Promising to modify higher education entrance requirements to end the exclusion that many segments of the population have historically experienced is also appealing. Some of these narratives oppose globalism, considering that it has done more harm than good, and use nationalist and ethnonationalist ideologies to support their causes. Nevertheless, if legal, economic or political structures are not changed (because governments do not want or because there are impossibilities to do it), the symbolic scale helps a candidate get elected or get public support, but it does not help to reach social justice. A review of the other scales would help us comprehend the possibilities of change in education and how the exercise of power would make an impact.

In sum, the opportunities to exercise power to transform education and make it something more than a reproductive instrument, the dissemination of knowledge as a public good and the development of scientific research, innovation and technologies could be a way to face ethnonationalist, populist and dogmatic regimes, to find better ways to take advantage of globalism, to reduce inequalities and to search for a world where social justice is the common characteristic and not a mere accident that happens here or there occasionally. The idea behind integrating all these elements in this entry, including two concrete cases of its usefulness is to propose an approach to continue with this discussion. It is relevant to reflect on each element separately and then suggest a way to put together all the elements to move forward (see Fig. 3). The ultimate purpose is to create a more democratic and a fairer world.

References

- Altbach, P.G., 1987. *The Knowledge Context: Comparative Perspectives on the Distribution of knowledge*. SUNY.
- Atkinson, A.B., 2015. *Inequality: What Can Be Done?* Harvard University Press.
- Bidasaca, K., 2016. *Genealogías críticas de la colonialidad en América Latina, África, Oriente*. Universidad Nacional de San Martín, IDEAS, Clacso.

- Boric, G., 2022. Programa de gobierno. Apruebo dignidad. Boric presidente. [https://s3.amazonaws.com/cdn.boricpresidente.cl/web/programa/Plan+de+gobierno+AD+2022-2026+\(2\).pdf](https://s3.amazonaws.com/cdn.boricpresidente.cl/web/programa/Plan+de+gobierno+AD+2022-2026+(2).pdf).
- Carnochan, W.B., 1993. *The Battleground of the Curriculum. Liberal Education and American Experience*. Stanford University Press.
- Castells, M., 2000. *The Rise of the Network Society*. Blackwell Publishing.
- Clark, B.R., 1983. El sistema de educación superior. Una vision comparative de la organización académica. Nueva Imagen-Universidad Futura-UAM.
- Collier, P., 2008. *The Bottom Billion. Why the Poorest Countries Are Failing and What Can Be Done About it*. Oxford University Press.
- Connor, W., 1973. The politics of ethnonationalism. *J. Int. Aff.* 27 (1), 1–21.
- Connor, W., 1984. Eco-or ethno-nationalism? *Ethn. Racial Stud.* 7 (3), 342–359. <https://doi.org/10.1080/01419870.1984.9993449>.
- De Sousa Santos, B., 2021. Decolonizar a la Universidad: El desafío de la justicia cognitive global. Clacso.
- Deaton, A., 2013. *The Great Scape*. Princeton University Press.
- Dussel, E., 1998. Ética de la liberación en la edad de la globalización y de la exclusion. Trotta.
- Elster, J., 1992. Justicia local. De qué modo las instituciones distribuyen bienes escasos y cargas necesarias. Gedisa.
- Eribon, D., 1992. Michel Foucault. Anagrama.
- Foucault, M., 1992. *Microfísica del poder*. Las Ediciones de la Piqueta.
- Giddens, A., 1984. *The Constitution of Society. Outline of the Theory of structuration*. Polity Press.
- Harari, Y.N., 2013. *From Animals Into Gods. A Brief History of Humankind*. Harper.
- Kauppinen, I., 2014. Different meanings of 'knowledge as commodity' in the context of higher education. *Crit. Sociol.* 40 (3), 393–409. <https://doi.org/10.1177/0896920512471218>.
- Lee, Y.-K., 2009. Migration, migrants and contested ethnonationalism in Korea. *Crit. Asian Stud.* 41 (3), 363–380. <https://doi.org/10.1080/14672710903119727>.
- Lipovetsky, G., Juvén, H., 2010. El Occidente globalizado. Un debate sobre la cultura planetaria. Anagrama.
- Maldonado-Maldonado, A., Cortes-Velasco, C., 2021. Why researcher matters: Latin America facing world-class universities and rankings. In: Hazelkorn, E., Mihut, G. (Eds.), *Research Handbook on University Rankings, Theory, Methodology, Influence and Impact*. Edward Elgar Publishing.
- Merton, R.K., 1996. *On Social Structure and Science*. The University Chicago Press.
- Milanovic, B., 2017. *Desigualdad mundial. Un nuevo enfoque para la era de la globalización*. México: Fondo de Cultura Económica.
- Murray, W.E., Overton, J., 2006. *Geographies of Globalization*. Routledge.
- Noya, J., Rodríguez, B., 2010. *Teorías sociológicas de la globalización*. Tecnos.
- OECD, 1996. *The Knowledge-Based Economy (OECD/GD(96)102)*. OECD.
- Piketty, T., 2014. *Capital in the Twenty-First century*. The Belknap Press of Harvard University Press.
- Propuesta. Constitución Política de la República de Chile, 2022. <https://www.chileconvencion.cl/wp-content/uploads/2022/07/Texto-CPR-2022.pdf>.
- Rodríguez, R., Maldonado-Maldonado, A., 2019. Mexico: Higher Education Under Populism. *Inside Higher Education*. <https://www.insidehighered.com/blogs/alma-maldonado-maldonado-and-roberto-rodriguez-g%C3%B3mez>.
- Sen, A., 1999. *Development as Freedom*. Oxford University Press.
- Sen, A., 2009. *The idea of Justice*. Allen Lane and Penguin Group.
- Sloman, S., Fernbach, P., 2017. *The Knowledge Illusion. Why We Never Think Alone*. Riverhead Books.
- Stiglitz, J.E., 1999. Knowledge as a global public good. In: Kaul, I., Grunberg, I., Stern, M. (Eds.), *Global Public Goods*. Oxford University Press.
- Stiglitz, J.E., 2002. *Globalization and Its Discontents*. W.W. Norton & Company.
- Stiglitz, J.E., 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W.W. Norton & Company.
- UNESCO, 2005. *Towards Knowledge Societies (UNESCO World Report)*. UNESCO.
- Waters, M., 1995. *Globalization*. Routledge.
- Xirau, R., 1990. *Introducción a la historia de la filosofía*. UNAM, México.
- Young, R.J.C., 1995. *Colonial Desire: Hybridity in Theory, Culture and Race*. Routledge.
- Zelditch, M., Walker, H.A., 2000. The normative regulation of power. *Advances in Group Processes (Advances in Group Processes, Vol. 17)*. Emerald Group Publishing Limited, Bingley, pp. 155–178. [https://doi.org/10.1016/S0882-6145\(00\)17007-9](https://doi.org/10.1016/S0882-6145(00)17007-9).

Education, identity and diversity

Ratna Ghosh, McGill University, Faculty of Education, Montreal, QC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	83
Difference and diversity	84
Difference	84
Diversity	85
Identity	85
Identity crisis	86
Education's role in the construction of identity in a diverse world	88
Teachers' impact on identity development	89
Implications for teacher education	90
Conclusion	90
References	91

Introduction

The pandemic has been a watershed in our history that will change how we live, work, and relate to each other as human beings as well as how we think of ourselves after months of confinement and continuing uncertainty. The pandemic has revealed the effects of inequality that exist among nations and within societies, and exposed major cracks in societal institutions all over the world. The lingering effects of imperialism, slavery, settler colonialism, and genocide continue through the uneven distribution of wealth, the inequality of living conditions among people of the world, climate change, and devastation (Douglass and Cooper, 2020) as well as an increase in militarization, corruption, violence, war, and human migration (Amin, 2020). In 2018, 70.8 million people were forcibly displaced around the world (United Nations High Commissioner for Refugees [UNHCR], 2019). In 2019, there were 272 million international migrants—3.5% of the world's population—living in a country where they were not born (International Organization for Migration [IOM], 2019). While the overwhelming number of people leave their home for work, family, or study, many are forced to leave to escape violence, conflict, and climate change.

Movement of people has been greatly disrupted due to responses of societies to the pandemic and the unequal ability of countries to control the COVID-19 virus and its variants. While this is likely to be temporary, migration patterns have become more complex and are likely to continue to increase in volume, scope, and diversity of populations. This is largely due to globalization and the technological revolution which have dramatically increased worldwide interconnectedness. Technological advances have facilitated travel as well as communication and education. The international movement of people has increased diversity and its complexity in societies around the world, and consequently in school populations. This super-diversity (Vertovec, 2007) involves various forms of affiliation and belonging, and different experiences due to multiple identities and their intersections. Super-diversity cannot be understood as static, because individuals' identities are changing and fluid.

Diversity implies difference, and difference is defined through identity. The interrelationship of difference and identity has played a significant role in the development of Western philosophy, and since the 1960s the topic has been prominent in cultural studies. Traditional views of identity consider the concept as fixed and unitary. Poststructuralist and postmodern perspectives view identity as being constructed relationally, and therefore, the concept has meaning only in relation to difference (Ghosh and Galczynski, 2014). The meanings given to differences based on race, ethnicity, culture, religion, gender, sexual orientation, dis/ability and other such constructs are influenced by history and structures of power. Given the changing nature of structures of power, who and what is different is always shifting. For example, one group that is seen less favorably today may be supplanted by another group which then becomes marginalized. Therefore, difference and identity are constructions that are always in a state of flux and shifting and "must be continually reinvented within new and changing circumstances" (Dolby, 2000, p. 901).

In the contemporary world, *globalization* implies global flows of people, technology, images, ideas, and money which can cause tensions between cultural homogenization and heterogenization (Appadurai, 1996). Perhaps the greatest danger of globalization is loss of identity. On the one hand, globalization results in homogenization by spreading a consumer culture that is ubiquitous and erodes local cultures and identities through mass media by "saturating the cultural space of a nation" (Schiller, 1985, p. 18). On the other hand, this cultural assault causes anxieties in individuals and groups who want to be recognized and assert their distinctive identity or culture. So, there is a tension between homogenization or sameness, and heterogenization or openness to difference due to diversity. This tension is present in all societies that have been touched by globalization and has given rise to identity politics resulting in a wide range of political activities based on different groups' experiences of injustice. With the spread of democracy, education, and awareness of human rights, the demand for recognition of different cultural identities by all groups is rising.

This chapter looks at ways educators can rethink the concepts of difference, diversity, and identity given this background of tensions caused by the global movement of people. Even in countries where there have been minimal demographic changes,

the push for inclusion has forced people to rethink existing differences among people in their societies, especially people with physical and mental challenges. The United Nation's Sustainable Development Goals (2015b) have emphasized inclusive education which focuses on those who have hitherto been excluded from societal institutions, including educational institutions, because of their differences. The chapter begins with a discussion of difference and diversity, followed by an exploration of the concept of identity. It shows the role of education in the development of identity and some of the problems that may arise when identity development is neglected or obstructed. It briefly touches on educational programs that could support positive identity development and ends with the implications of identity in teacher education.

Difference and diversity

Difference

There is an astonishing variety of difference among human beings in every aspect whether it be individual abilities and physical characteristics or social and geographical positioning. Difference is a comparative term. The Oxford Dictionary defines it as "that which distinguishes one thing from another." Several questions arise from this definition: Different from whom? Different in what way? In the context of individual and cultural identity, difference is measured against a norm and the norm is defined by the dominant group which has power and whose characteristics become the standard. Those who do not fit the model are different and become *the other* (Ghosh and Abdi, 2013).

Who is seen as different? It is *the other* who is different: the woman is different from the man; gays and lesbians from heterosexuals (Pandey, 2010). In the United States, Canada, Australian, and European contexts the black or brown person is different from the white person. In the context of education, the black student is different and the white student is the norm. Difference here is an expression of dominance and subordination, hierarchy in social, political, and economic power. In South Asian countries, the lower castes are different. Thus, the meanings given to differences is a way of legitimating, normalizing, and reinforcing hierarchical power relations (Ghosh, 2012).

But, as Hall (1991) has pointed out, difference is not a static, fixed reality; difference is always constructed. Markers of difference have mostly been race/ethnicity, gender, culture, religion, language, socio-economic class, caste in India/South Asia, sexual orientation/identification, and disabilities. Difference is based on socially created categories built on physical (race/ethnicity), biological (gender based on sex), and socio-economic (class) characteristics.

Evolutionary theory points out that group living is an adaptive strategy for the human species and critical to reproduction and human survival against hostile environments (Van Vugt and Schaller, 2008). Van Vugt & Schaller point out that groups evoke feelings of loyalty and solidarity leading to tribalism and ethnocentrism in which people are aware of within group physical differences but attach minimal importance to them. However, these feelings of togetherness also lead to rivalry and exclusion of other groups resulting in racism, hate, and exclusion of the other.

Racism, which is extremely damaging to one's identity, is generally considered "an assertion of Western cultural superiority" (D'Souza, 1995, p. 537) that is a direct result of colonialism. The classification of people into hierarchies of "races" by Europeans during colonization was based on select physical characteristics and social constructs. Similar to racism rooted in colonialism, the caste system in India is rooted in Hindu texts and is a hierarchal form of social stratification. It is hereditary, and the markers of difference are derived from religion and cultural notions of purity and pollution rather than physical differentiation. Although it originated in the Hindu scriptures, it has penetrated all religious groups across South Asia. Its impact is enormously damaging to the identity and all aspects of the lives of those who are subject to *casteism*, or discrimination based on lower caste status. The Indian Constitution prohibits discrimination based on caste but legislation has not removed caste discrimination in society, which is often violent and prevalent in many parts of the country. Similarly, the concept of race is very much alive in countries of the Global North, as highlighted by the Black Lives Matter (BLM) movement. Racism exists even when it is illegal, and despite lack of scientific basis for it. "Caste and race are interchangeable in terms of their potential for human wrongdoing or evil" (Thirumal, 2021). It should be noted that there are ostracized groups in other parts of the world such as the Burakumin in Japan and Korea, Osu in Nigeria, and similar communities in other parts of Africa (EPW Engage, 2021).

Biologically, "humans share 99.9% of the same DNA. The physical differences we observe (e.g., skin color, eye color, hair texture, bone structure) are known as clines, which do not qualify humans to be distinguished into different biological races of people" as stated in National Human Genome Research Institute (2018) and Sussman (2014) (Svetaz et al., 2020, p. 313). Nobel laureate Toni Morrison explained, "There is no such thing as race. None. There is just a human race—scientifically, anthropologically. Racism is ... a social construct ... it has a social function, racism" (2014). So, while race has no scientific basis, it is a powerful social and political force, and racism is prevalent in many areas of the world, greatly affecting the identity and every aspect of the lives of those who experience it.

Concepts of race, ethnicity, gender, and class are social constructions. They are inherently political because they result from the power to exclude and serve to form boundaries by separating groups of people whose identities are defined by the groups in power to safeguard their position of privilege. As sociological ideas, these concepts have changed over time and are not fixed. Although racism now is seen as discrimination directed mainly at non-white people, this was not always the case. For example, Eastern Europeans were subjected to racism in Canada at the turn of the 20th century because they were seen as being of a different race and ethnicity by Anglo-Saxon settlers (Ghosh and Galcynski, 2014). Viewing whites as one racial group is a recent phenomenon, and Banks (1995) points out that in the 19th and early 20th centuries in North America some white races were considered superior

to others. The classifications of race, ethnicity, gender, and class have no relevance either as a basis for assignment of rights or as grounds for prejudice and discrimination. They indicate changing power and social relations (for example, women did not always have the right to vote). They are inherently political, because they veil domination and exclusion, these classifications are symbolic of some form of underlying power struggle (Goldberg, 1992). They are also ideological because they provide the basis for action (Buchignani, 1980). Most importantly, these classifications impact identity development and the life chances of individuals and groups. As globalization and other factors impact diversity across the globe, hierarchies based on perceived differences of race, ethnicity, gender, and class continue to develop and fluctuate.

Diversity

Statistics indicate that there is tremendous mixing of the world's population. The plurality of human identities makes human beings "diversely different" (Sen, 2006, p. xiv). Globalization, wars, and natural disasters have caused large scale demographic shifts, thus increasing diversity in societies. Data on international migration shows increase in the last five decades: there were 281 million international migrants in the world in 2020 with Europe and Asia hosting 87 and 86 million respectively and North America 59 million (IOM, 2022). Of these, North America and Europe added 16% and 12% international migrants to their total populations but Asia and Europe took in the largest numbers (IOM, 2022). United States has been the main country of destination for international migrants since 1970 while India is now the largest country of origin. More than 40% of international migrants worldwide were from six countries in Asia in 2020.

Societies are now recognized to be diverse in multiple ways. Diversity implies variety but it means different things to different people, such as ideas, characteristics, values, or other factors. Demographic diversity usually suggests differences in ethnicity, culture, socio-economic status, gender, religion, language, and sexual identity and preference. Demographic diversity mainly results from the global movement of people across national borders. Immigrants have tended to concentrate in a shrinking pool of prime destination countries. Czaika and de Haas offer several reasons for this despite the globalization of migration; certain countries, for example, favor skilled or wealthy groups reflecting the "asymmetric nature of globalization processes in general" (2015, p. 319). With aging populations and rapidly declining birthrates in countries of the Global North, the need for skilled and non-skilled labor is met through immigration and refugee resettlements. Many of these countries have changed their criteria for selection of immigrants which were based on race/ethnicity and country of origin to a focus on skills and education so that immigrants are arriving from non-Western cultures and belong to religions other than Christianity (Ghosh and Abdi, 2013). In addition to immigrants, 26 million refugees worldwide, all from countries of the Global South, are fleeing war, violence, persecution, food shortages and climate change. Although 85% refugees are hosted in countries of the Global South, including Turkey, Jordan, Lebanon, Sudan, Pakistan, and others, many are also going to countries of the Global North. Germany is host to the largest number of refugees among countries of the Global North and countries assisting in the resettlement of refugees include the USA, Canada, the UK, France, and Norway, among others (Moutafis, n.d.).

As such, the global movement of people in homogenous societies have made them increasingly complex. With immigration from other European, later Asian, African and Latin American nations, countries in the Global North have become more and more diverse in ethnicity and culture, religion, language and socio-economic status in addition to contemporary recognition of gender (cisgender and LGBTQ+) identities. So, *super-diversity* (Vertovec, 2007) implies more than a focus simply on ethnicity and culture not only due to the variety of categories but also because of the complex intersection of differences in social class, gender, religion, language, etc. which result in very different experiences for any one individual or group. Differences in people's localities (historical, global) and positionalities (race, gender, language, religion, class, sexual orientation) influence diversity because where and how a person is situated in a society leads him or her to live through a particular set of experiences and to encounter distinctive power relations (Ghosh, 2019).

Identity

Human development is influenced by the context and the socio-cultural experiences of people for the growth of moral, social, emotional, and cognitive capacities. Although there are numerous studies of child and adolescent development in many non-western countries, they are usually based on western theories (Bansal, 2022). Indigenous models of development such as the Brahmanical-Hindu world view in India, Bansal points out, takes "a largely upper caste ... male centric view" (p. 42) and does not reflect the experiences of most people such as women and those who belong to other castes or those outside the caste system.

A fundamental characteristic that makes us human is our continuous search to understand who we are as individuals, what we want, what we may become. Who am I? Identity is a constructed sense of self. It is a subjective sense of oneself influenced by personal characteristics which are significant for an individual. It is who I am, my individuality. That is individual identity. Collective identity is how individuals define themselves as belonging to a group based on markers such as race, ethnicity, culture, gender, socio-economic class, religion, dis/ability etc. From the psychological perspective, individual identity refers to a person's self-concept: how I see myself in terms of the markers of difference, my attributes (self-image), how much I value myself and how I think others see me (self-esteem). Among many social identity theories, Tajfel and Turner (2001) see self-concept made up of two essential parts: personal identity (what makes one unique) and social identity (belonging to groups).

The drive for identity is a fundamental human need. Identity is extremely complex. Although we tend to identify ourselves mostly in terms of our physical characteristics (e.g., race/ethnicity), biological (sex), social (e.g., socio economic class/caste), and cultural relations (influenced by our history and social/economic conditions), we have multiple affiliations and identities. Moreover, our identities are not fixed but are influenced by our social realities which are changing. How others identify us is also complex because those shape how we identify ourselves. Self-concept and identity are formed through social interactions and social distance because of how others perceive us and how we see ourselves in relation to our history, culture, class, gender, etc.

In his influential essay, *The Politics of Recognition* (1994), political philosopher Charles Taylor points out: “Identity is partly shaped by recognition or its absence, often by the misrecognition of others ... Non-recognition or misrecognition can inflict harm, can be a form of oppression, imprisoning someone in a false, distorted, and reduced mode of being” (p. 75). As a social-psychological construct, identity is shaped by those around us, with media and technology, through government policy, our histories, and cultural expression.

In his seminal work on identity, Hall (1991) points out that difference and identity are not given; they are produced in relationship with one another. One’s identity is constructed only in relationship with another. Difference is also produced in relationship with another. Since one’s identity is defined through difference, the question then is how difference is linked to power and how it is produced and reproduced.

Post-colonial discourse disputes that there is any culture that is pure or essential. The word ‘hybridity’ borrowed from biology, where it indicates the merger of two genetic streams, had a negative connotation in colonial times. In post-colonial discourse, however, it connotes the bridging of two or more cultures. In a globalized world “all forms of culture are continually in a process of hybridity” (Rutherford, 1990, p. 211). The cosmopolitan hybrid identity is positioned not in any one culture but in a third space to include the other in a transcultural, cosmopolitan situation. Attributed to Bhabha (1994), this postcolonial theory of identity suggests that the hybrid third space is an ambivalent site where cultural meaning and representation are not fixed. By being positioned in the juxtaposition of cultures the hybrid identity can mediate between many cultures to translate and negotiate difference.

Identity is the basis of discrimination because racism, sexism, and homophobia for example, are built on differences from the norm, which is based on the characteristics of the group in power, and how people are identified in terms of race, gender, or sexual identity in comparison to that norm. Everyone has multiple identities, and these identities intersect in complex ways. Although Kimberlé Crenshaw (1989) drew attention to how social identities are configured by intersecting issues of power and oppression in the context of legal studies, it was not till a decade later in 2000 that intersectionality became an analytical lens to investigate discrimination in the social sciences. The dynamics of intersecting layers of multiple identities captures the complexity of identities by acknowledging the reality of people’s experiences, especially those who are subjected to multiple forms of oppression (McCall, 2005).

Identity crisis

Perhaps the most profound impact on individual and collective identities has been caused by colonialism. Briefly, colonization by European countries in Asia, Latin America and Africa had a very negative impact on the identity of people of the colonized countries by instilling a sense of shame and inferiority in them vis-à-vis the white colonizer. Fanon (1952) describes how the inferiority complex is internalized by the colonized through the creation of a false identity in the hope that imitating the white colonizer would give them equality, while their true identities remain hidden behind the white masks. Like Fanon, Rabindranath Tagore, in India, was more concerned about the deformation of the mind and Indian identity due to colonial education than he was about colonial rule (Ghosh et al., 2010).

Similarly, in settler colonial countries the Indigenous populations faced the same denigration of their cultures. In some cases, as in Canada and the US, the attempt was to wipe out the Indigenous identities of children through residential schools where students were subjected to extreme sexual and physical violence. But while the colonial project led to mimicry (Bhabha, 1994), or imitation of the colonizer and created a false identity in the colonized, settler colonialism forced assimilation on Indigenous people in an attempt to kill their cultural identity.

In post-colonial societies, the problem of coloniality (the identities associated with European domination) persists despite the end of colonial rule. However, in the settler colonial countries, the challenge is also the settler identity. For example, in Canada, the English population has become very diluted due to immigration. There are complexities in the “Canadian” identity which is further complicated by the French province of Quebec which has asserted its own identity through the assertion of the French language. While ideologically, Canada has created a social and national identity through its Multicultural Policy (1971) to unite its population, tensions remain with Quebec and Indigenous groups.

In postcolonial countries like India, as in other South, Southeast and East Asian countries which are collectivist societies, identity development takes place amid contradictions of tradition and modernity, traditional values and modern values, parochial culture and cosmopolitan culture, family and peer group that create confusion and inner conflict about individual’s identities (Bansal, 2022). Urban youth have access to technologies and a range of social media, consumer goods and “guiltless indulgence” and are identified by brand names (Bansal, 2022, p. 49). They are exposed to a number of alternatives and what Bansal calls “enterprise culture” (p. 50) characterized by confidence, initiative, and risk-taking. Rural youth (the majority) and low caste groups who are poor are aware how their poverty is excluding them from the opportunities offered by globalization. They experience a diminished sense of self-worth, and develop a negative identity (lazy, angry, violent delinquent) to resolve a crisis an identity (Bansal, 2022).

Children who experience caste discrimination grow up with a “stigmatized identity” (Thorat, 1979, p. 74), and yearn for equal treatment.

Identity crises are also caused by globalization, technological advances, and natural disasters when individuals and groups attempt to make sense of rapid changes and try to figure out who they are. The latest crisis is the COVID-19 pandemic and its variants which have brought a feeling of uncertainty and helplessness causing many people to question their identity and re-evaluate their life's path. This is an identity crisis.

With the announcement of a public health emergency by the World Health Organization, countries around the globe hurriedly took aggressive measures to curb the spread of the virus: lockdowns, social distancing and isolation, refusing visitation to loved ones who were dying in large numbers in hospitals, canceled events, closed institutions including schools, colleges, universities, offices and businesses, all have challenged our identities. But these measures had very different impact on the populations of different countries. The virus has infected people indiscriminately and different countries dealt with their situations very differently. In India, for example, 1.38 billion people were locked down within 4 h “The lockdown to enforce physical distancing had resulted in the opposite—physical compression on an unthinkable scale ... the poor are sealed into cramped quarters in slums and shanties” (Roy, 2020, para. 29). In contrast, “... the wealthy and middle classes enclosed themselves in gated colonies” (Roy, 2020, para. 22). The marginalized were more affected due to lack of resources and inability to safely distance themselves even in the most affluent countries. In the US, for example, while overwhelmed hospitals and overworked medical staff dealt with lack of equipment and chaos the majority of deaths due to COVID-19 were people living in poorer and more urban counties (Holmes Finch and Hernández Finch, 2020). In individualistic cultures, such as the US, media reports indicated that a sizable number of people refused to get vaccinated whereas in many poor countries in Africa vaccines were in short supply.

Collective identity—social relations with others—is challenged when isolation and social distancing cause depression and other mental health issues. The experiences are quite different for people working in urban areas in secondary (e.g., manufacturing), tertiary (e.g., service industries) and quaternary (e.g., technology and scientific fields) sectors from those in the primary sector (e.g., agriculture). Individual identity is challenged when people spend all their time at home. When individuals are constantly close to family and working from home, they start to re-evaluate their lives. In countries like Canada and US, with the disruption in normal routines obliging people to spend time alone and bringing them face to face with mortality through media reports of alarming number of deaths, loss of friends and relatives has made people ask existential questions about what is important in their lives, their goals, and priorities. As the economy and educational institutions reopen, there is a shortage of people in several service area jobs because some people are not returning to their jobs according to media reports. In countries of the South, the majority of the population is in the primary sector, and they have fallen into greater debt and poverty. For example, the havoc caused by the pandemic is expected to double India's poverty in which the rural areas are most affected (Ram and Yadav, 2021).

The global increase in rates of domestic violence, and the urgent need to focus on mental health and well-being in adults and children, both result from identity crises. Although our identities change through our daily interactions throughout our lives, the pandemic has forced us to break with the past because it has touched every area of our lives, all our identities. The vaccine has brought some normalcy, but the new normal will be very different. We will need to rebuild our identities. It will mean that the vaccinated will have the privilege to live safely by following strict regulations but know that responsibilities are part of the deal. Those who challenge the regulations, for example, by refusing to wear a mask, have difficulty coping with the challenge to their identities, unable to acknowledge that the world has changed, that their lives will no longer be the same.

COVID-19 has had a tremendous effect on education, both good and bad. Everywhere, educational institutions had to be closed suddenly and the business of educating students led educators all over the globe to resort to online learning. The delivery of education has changed dramatically. Instant communication and social media affect the way we communicate and impact the way we develop our identities. The various online teaching platforms enabled instruction when it was essential to stay at home. Yet, losing the opportunity to be close to others in a regular classroom affects the relational aspect of identity development and this is worse when some people do not show their faces when they are online. Privacy issues prevent asking students to show their faces in an online class, and by neglecting to identify their presence much is lost in student-teacher relations. While online delivery was the only option for at least a year, with educational institutions gradually opening to face-to-face instruction the new normal will have blended and hybrid modes of delivery. In this too, the marginalized suffer due to lack of computers and bandwidth, not only in economically poor countries where even electricity is sometimes lacking, but also in rural and poor neighborhoods in affluent countries where internet reception may be problematic, and bandwidth is not enough for several activities such as video conferencing.

Technology has had unprecedented significance during the pandemic, particularly in education. Technology affects the way people communicate at every level and therefore impacts the way people form their identities, both individual and collective. The pandemic has underlined the significance of social cohesion and collaboration in an interdependent world and drawn attention in education to the need for student and teacher well-being and care through a focus on individual identity and compassion for the other (Ghosh, 2022). This focus on student and teacher well-being and mental health, and to importance of social and emotional learning (SEL) is a positive outcome of the pandemic. There was a remarkable increase in collaborative webinars during the pandemic. The increase in educational reach to students who would otherwise not be able to participate, (especially in countries of the Global South and in rural and remote areas in the Global North) is a very good outcome. There was also an immense increase in online programs, platforms, and teaching methods to improve online teaching.

Education's role in the construction of identity in a diverse world

The socialization of children takes place through formal as well as informal means. Many countries, both in the geographical North and the South, have universal education through the efforts of the United Nations and its member countries that follow Education for All—EFA, Millennium Development Goals (MDG) and Sustainable Development Goals (SDG) (UNESCO, 1990; United Nations, 2015a,b). This implies that in most countries young people spend at least 10 years of their lives in schools. During the developmental age children seek to uncover their own identity, form a world view, bolster their self-confidence and search for meaning in their lives. So, the school (formal education) is an important place where identity is formed.

Education is a complex concept and is a double-edged sword. It can be used to promote positive individual identity development and thus positive self-esteem, but it can also have negative effects on identity development of children and youth who then develop low self-esteem. Education's role can be significant in the learning process because the beliefs, values, and feelings transmitted through the curricula, school environment, and classroom interactions shape the construction of each person's identity and self-concept. These values involve the whole community, teachers, administrators, parents, and others. Ideally, the community strives to provide a safe space for students and a sense of empowerment, as well as to develop critical understanding of how social relations are maintained through interaction with others (Ghosh, 2021).

As countries become more diverse, the diversity in society is reflected in schools resulting in discrimination, hate, and violence based on how students are identified. Education plays an important role in the development of racial, cultural, ethnic and gender identities (DeCuir-Gunby, 2009). Discrimination based on race, for example as in the US, exerts a profound effect on youth and children's identity development, their health, and lived experiences because, "Racial-ethnic identity is a multidimensional construct" (Svetaz et al., 2020, p. 310). While identities are constructed from the earliest stages of a child's life, Svetaz and colleagues point out that identity development is a key component of adolescence when a sense of self is crystalized (2020, p. 310). Even when racism and other forms of discrimination are not explicit, "implicit bias is a critical factor contributing to disparities" (Svetaz et al., 2020, p. 315) such as trauma. Moreover, positive well-being is an important predictor of school success (Dyrbye et al., 2006) and poor mental health can cause negative feelings about one's identity such as exclusion (Knauss et al., 2015) or the experience of a threat (Nelson, 2009) resulting in poor performance in school.

Most Western countries (such as US, Canada, Australia) are increasingly diverse in terms of culture, religion, and language in addition to existing socio-economic and gender divisions. Education can be an equalizing force, but it can also perpetuate inequities for those whose identities are marginalized due to discrimination and non-recognition so that they become disengaged (Ghosh et al., 2016). In modern democratic societies, the challenge to schools and teachers is to provide fair and equitable learning opportunities and experiences to diverse student populations that enable them to participate fully in society. This is especially difficult in cultures which have had assimilationist policies and now need inclusive, critical multicultural policies which are radically different. So, the need to deal equitably with students who are different in ethnicity, culture, language, religion becomes imperative. The shift from assimilation into the host society to multicultural and inclusive education involving space for other cultural identities is a paradigm shift. In European countries, the US, Canada, Australia, and New Zealand, the inclusion of non-Western identities—in terms of race/ethnicity, culture, religion, language—can cause nationalistic sentiments giving rise to hate and violence against groups seen as the "other".

In many countries of the Global South, like those in South Asia, different challenges such as unprecedented demand for education, intense competition, high attrition, and generally low-quality education cause tensions in educational opportunities and impact the identities of those who are underemployed and unemployed among large sections of both educated and uneducated people. In the school system, socio-economic class, gender, caste, religion, and physical/mental challenges are markers of difference which exclude children. These inequities exist both in the Global North and South but are larger in scale in the South and therefore, more critical.

Students who don't feel their identities are validated by the curriculum and school environment can become disengaged and are often marginalized. They may withdraw from participating actively in school or even get drawn to anti-social activities like joining gangs, getting involved in violence, and on occasion engaging in extremism to find validation when morally disengaged. Mental health problems not only make school experiences difficult but can lead to self-injury and even suicide. Some root causes that can trigger an individual to disengage or turn to violence are alienation, marginalization, insult or misrecognition of one's individual or collective identity, a search for validation, and the failure to understand the historical development of discrimination and other global inequalities.

What can education do? The concept of education has changed from a process of transferring knowledge from the teacher to the student (teacher-centered), to learning as the construction of knowledge by the student (student-centered). Education that involves the construction and dissemination of knowledge and values through dialogue is likely to encourage students to develop positive identities because it allows for the freedom to question, negotiate and relate one's personal experiences to the learned concepts. On the other hand, education in which knowledge is transmitted to learners through practices such as rote learning implies lack of opportunity to question or challenge ideas and develop individual identities.

Educational content and process should focus on the validation of the cultural, social, and gender differences of students and the development of their individual identities. Education can attempt to combat extremist behavior through developing values and understanding, but educational strategies are long term and aim to accomplish long-lasting objectives. It is not only the content that is taught which should reflect the diversity of the students, but the attitude of the classroom teacher as well as that of the school

administration toward students that are not from the dominant group whether it is in terms of race/ethnicity, culture, gender, socio-economic class, religion or ability.

Teachers' impact on identity development

Verhoeven et al. (2019) conducted a literature review to examine the role of schools in adolescents' identity development. Although this review of 111 studies looked only at studies published in journals in the Global North, it still gives insight into school's impact on identity development. Their findings indicated that while a stable identity is important for being resilient, reflective, and autonomous in making important life decisions with a sense of competence, this was becoming more challenging due to individualization, emancipation, and migration. By introducing new ideas, activities, and possibilities in school, teachers can help adolescents to explore the identity implications. The results of Verhoeven and colleagues literature review show that:

- (1) teachers may unintentionally affect students' identity development through the hidden curriculum. This happens through (a) selection processes for various school activities and differentiation such as grouping students. For example, the identity exploration of high-status groups are positive about their future, whereas those in lower groups are fearful of who they might become; (b) teacher strategies unintentionally inform students' identity development; (c) teacher expectations are very significant, and teachers tend to have expectations based on ethnicity and gender. So, certain groups of students may experience different opportunities which have a positive or negative impact on their identity development; (d) peer norms also affect identity development especially when some identity positions are stigmatized.
- (2) Teachers can foster identity development organizing Explorative Learning Experiences in which intentional and planned learning experiences are offered. (a) In-Breadth Exploration or exposure to on-site, hands-on unfamiliar learning experiences and content allow students to try out new identity positions and new interests as well as identify new talents; (b) In-Depth Exploration enables students to delve deep into their subjects of interest and on-site, hands-on experiences such as meeting with role models encourage them to develop a keen interest in making identity commitments. This is particularly important for STEM subjects (c) Reflective Exploration through reading and discussing topics of interest to students whereby dialogue could lead them to search for a balance in their dreams and societal expectations.

What programs exist to negotiate difference? Marsden (2017) identifies three academic approaches: the moral, social justice, and rights-based approaches. There has been rapid development in programs and courses to deal with the challenge of diversity in schools. Multicultural/intercultural/diversity and inclusive education, values education, human rights concepts, and anti-racist education increasingly permeate the curriculum. Human rights and citizenship education, social justice-based education, and values education can provide a foundation for recognizing the dignity and rights of human beings. However, if they are focused on abstract ideas of rights and values without relating student experiences to the complex social, economic, and political issues, if they are not embedded in the curriculum through an ethics of care (Noddings, 2013), students might develop "tolerance" for the other from a superior and superficial point of view rather than a caring way of relating to others. Nel Noddings' (2013) theory of an ethics of care focuses on an education based on compassion and care in relationships. Caring relationships focus on developing people as caring and relational, which plays a crucial role in developing a person's identity. What is essential is the development of relational and critical thinking. Focus on the relational aspect of human identity is important in schools since identity is formed through relationships. Jackson's (1997, 2004) interpretive approach incorporates dialogical methods, critical pedagogy, and cooperative learning with a focus on the importance of students' experiences, the context, and the curriculum. This approach helps students situate their identities through a comprehensive understanding of the complex issues related to their lives (Ghosh et al., 2016). Multiple perspectives are crucial because the failure to explore various interpretations could prevent students from engaging critically with absolute claims to truth and what is good. The interpretive approach has significant implications for teacher education programs.

In traditional educational contexts, teaching methods such as lecturing and rote learning that discourage student participation and critical thinking work against developing engaged citizens. That is why using ineffective pedagogical methods to teach social justice issues does not have the desired effect of promoting social inclusion and building on students' identities. Education programs can inadvertently close minds if they are uncritical in nature and feed into existing biases and prejudices.

With the shift in the conception of Multicultural Education from a liberal perspective to Critical Multicultural Pedagogy, the focus has moved from folklore, cultural preservation, and prejudice reduction strategies to inclusion of the identities (representing the diversity of society), anti-discrimination (such as racism) and such democratic values as equality, fairness, and inclusion. The structural barriers in society and school stifle students' identities due to the social constructions of race and gender, religion, and physical/intellectual challenges. These inequalities are intergenerational and not only affect their identities but also their identity aspirations and life chances.

Critical Multicultural pedagogy aims to develop an understanding of oppressive structures through critical examination of colonialism, slavery, imperialism, and globalization which have created the current social and economic inequalities globally (Ghosh, 2022). It also looks at contributions of various cultures (including indigenous cultures) and civilizations to the development of science and technology. The focus is on the development of critical consciousness of oneself and of others. It incorporates social and emotional aspects of learning with cognitive aspects of learning and being. The idea is to help develop students' individual identities with critical thinking for social consciousness of rights and responsibilities in society and the world we live in. Most

importantly, the negative effects of unequal treatment and the benefits of social justice and diversity affect all, not only the other (Ghosh, 2022). Critical Multicultural Education (CME) is a worldview that should permeate the curriculum in all subjects.

Dominated by a neo-liberal ideology and the commodification of education as a result of globalization, the instrumentalist form of education has obscured identification with humanistic values of compassion, an ethics of care, collective rights, human dignity, and worth. The focus of this form of education is on development of work skills and training, rather than critical thinking and global citizenship values such as human rights, social justice, interdependence, intercultural communication and social cohesion. Economic and technological skills are essential for living, but cognitive skills and social and emotional skills can also be developed alongside those in schools. What makes education different from training are the long lasting human values it instills. These skills are important for good citizenship if we seek a future which is less unequal.

The pandemic has shown that inequality in the world is not only a moral issue; it is an economic issue, an environmental issue and most importantly, an educational issue. Online learning is not accessible for those who suffer from poverty and homelessness, for those who live in many countries of the South and in remote areas of countries of the North, and for those who are physically challenged and marginalized.

Implications for teacher education

Recognition and inclusion are fundamental to human dignity. To have a truly inclusive classroom in which each student is validated, recognized, and feels cared for as well as challenged means understanding difference through a critical multicultural, anti-racist, and social justice perspective. The change from an assimilationist system of teaching and learning to an inclusive and critical multicultural perspective implies a paradigm shift in teacher education programs, a total reversal of one's worldview.

The pandemic can be a turning point for educational institutions to reinvent their pedagogies and structures. We have the opportunity in this defining moment to reconceptualize the content and methods of teaching but most of all, we need a dramatic change in worldview. The need is a philosophical rethinking and reconstruction, a new, inclusive world view that permeates all courses and activities in teacher education programs. To validate identities and provide students the opportunity to develop strong identities teachers must themselves develop a critical consciousness, take multiple perspectives, provoke students to challenge the dominant norm, and disrupt the way knowledge is processed. It is of course imperative for teachers to be introspective and examine their own ways of thinking, their biases and prejudices, and reflect on their own worldviews before they facilitate students in examining their beliefs. This is no easy task; nor is it packaged in a course. It is a way of thinking, a philosophy of life and of teaching. Pre-service and in-service teachers need to develop critical thinking and a critical awareness of how they teach. Teachers need to see the importance of global interdependence, the complexity as well as the potential of diversity, and the necessity of encouraging students to express their identities and experiences to make lessons more relevant to students' lives in a student-centered philosophy of teaching (Ghosh, 2019).

Inclusive classrooms imply making all students feel that they are part of the classroom, the school, and the society to avoid the varied effects of alienation, marginalization, and isolation (Ghosh, 2019). Teachers who deal with children from different ethnic/race, class, gender, religious, language, sexual orientations and physical or mental challenges must aspire to provide equal educational opportunities to all children irrespective of their social category. They cannot be fair to all students by ignoring their difference (e.g., being color-blind) because society is not color-blind; it privileges certain ethnic groups. It is fairness, not blindness to differences or color blindness, that is a fundamental principle of justice (Appiah, 1992). Fairness is not equal treatment but equity or equality of opportunity: it is to recognize difference without allowing it to categorize people (Ghosh and Galczynski, 2014).

Conclusion

One of the purposes of education is to enable citizens to develop strong identities to prevent and counter any action that stifles human dignity. Schools play a significant role in the construction of children's identity. Self-concept and identity are formed through social interactions and social distance as a result of how others perceive them and how they see themselves in relation to their history, culture, class, gender, etc. Identity construction has tremendous implications for educational success because it is related to self-esteem, empowerment, school achievement, and ultimately to one's life chances. Multiple diversities make identity development complex. Davies (2009) suggests the encouragement of hybrid identities so that children can see themselves as unique and distinctive with a variety of identities.

In these deeply controversial times, amid contradictory discourses, the pandemic has forced us to collectively question our identities. Face to face with an invisible enemy which is attacking people indiscriminately and devastating the most vulnerable, we are together questioning all aspects of what used to be. One of the most far-reaching global catastrophes of our time has provided an opportunity to recreate our collective identity as one human race and re-envision a world that is less divisive and more humane with less inequality, poverty, discrimination, and violence for a sustainable and peaceful planet.

References

- Amin, T., 2020. How Climate Change and Colonialism Are Spurring Mass Migration: The Violent Roots of Today's Unprecedented Displacement. Minneapolis Institute of Art. <https://new.artsmia.org/stories/how-climate-change-and-colonialism-are-spurring-mass-migration-the-violent-roots-of-todays-unprecedented-displacement/>.
- Appadurai, A., 1996. *Modernity at Large: Cultural Dimensions of Globalization*. University of Minnesota Press.
- Appiah, K.A., 1992. In My Father's House: Africa in the Philosophy of Culture. Oxford University Press.
- Banks, J.A., 1995. Multicultural education and curriculum transformation. *J. Negro Educ.* 64 (4), 390–400. <https://doi.org/10.2307/2967262>.
- Bansal, P., 2022. Development in adolescence: an introduction. In: Parta, S. (Ed.), *Adolescence in India: Issues, Challenges and Possibilities*. Springer, pp. 37–68.
- Bhabha, H., 1994. *The Location of Culture*. Routledge.
- Buchignani, N., 1980. The social and self-identities of Fijian Indians in Vancouver. *Urban Anthropol.* 9 (1), 75–97. <http://www.jstor.org/stable/40552897>.
- Crenshaw, K., 1989. Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist policies. *Univ. Chicago Leg Forum* 1, 139–167. <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>.
- Czaika, M., de Haas, H., 2015. The globalization of migration: has the world become more migratory? *Int. Migrat. Rev.* 48 (2), 283–323. <https://doi.org/10.1111/imre.12095>.
- Davies, L., 2009. Educating against extremism: towards a critical politicization of young people. *Int. Rev. Educ.* 55, 183–203. <http://www.jstor.org/stable/40270074>.
- DeCuir-Gunby, J.T., 2009. A review of the racial identity development of African American adolescents: the role of education. *Rev. Educ. Res.* 79 (1), 103–124. <https://doi.org/10.3102/0034654308325897>.
- Dolby, N., 2000. Changing selves: multicultural education and the challenge of new identities. *Teach. Coll. Rec.* 102 (5), 898–912. <https://doi.org/10.1111/0161-4681.00083>.
- Douglass, K., Cooper, J., 2020. Archaeology, environmental justice, and climate change on islands of the Caribbean and southwestern Indian Ocean. *Proc. Natl. Acad. Sci. U. S. A.* 117 (15), 8254–8262. <https://doi.org/10.1073/pnas.1914211117>.
- D'Souza, D., 1995. Is racism a western idea? *Am. Scholar* 64 (4), 517–539. <https://www.jstor.org/stable/41212409>.
- Dyrbye, L.N., Thomas, M.R., Shanafelt, T.D., 2006. Systematic review of depression, anxiety, and other indicators of psychological distress among US and Canadian medical students. *Acad. Med.* 81 (4), 354–373. <https://doi.org/10.1097/00001888-200604000-00009>.
- EPW Engage, August 20, 2021. Caste and race: discrimination based on descent. *Econ. Polit. Wkly.* <https://www.epw.in/engage/article/caste-and-race-discrimination-based-descent>.
- Fanon, F., 1952/2008. *Peau noire, masques blancs*. Seuil. Translated as *Black Skin, White Masks* (R. Philcox, Trans.). Grove Books.
- Ghosh, R., 2012. Diversity and excellence in higher education. *Comp. Educ. Rev.* 56 (3), 349–365. <https://doi.org/10.1086/666545>.
- Ghosh, R., 2019. Teacher education in a globalized world. In: Mueller, J., Nickel, J. (Eds.), *Globalization and Diversity in Education: What Does It Mean for Canadian Teacher Education?* Canadian Association for Teacher Education, pp. 416–443.
- Ghosh, R., 2021. Can education contribute to social cohesion? In: Lindsay, B. (Ed.), *Reflecting on Comparative and International Education: Leading Perspectives from the Field*. Palgrave, pp. 87–104.
- Ghosh, R., Naseem, A., Vijh, A., 2010. Tagore and education: gazing beyond the colonial cage. In: Abdi, A. (Ed.), *Decolonizing Philosophies of Education*. Sense Publishers, Rotterdam, pp. 59–71.
- Ghosh, R., 2022. Critical multicultural education as a platform for social justice education in Canada. In: Abdi, A. (Ed.), *Social Justice Education in Canada*. Canadian Scholars Press.
- Ghosh, R., Abdi, A., 2013. *Education and the Politics of Difference: Canadian Perspectives*, second ed. Canadian Scholars' Press.
- Ghosh, R., Galczynski, M., 2014. *Redefining Multicultural Education: Inclusion and the Right to Be Different*, third ed. Canadian Scholars' Press.
- Ghosh, R., Manuel, A., Chan, A., Dillimulati, M., February 2016. *Education and Security: A Global Literature Report on Countering Violent Extremism*. Tony Blair Faith Foundation (now Tony Blair Institute for Global Change), United Kingdom.
- Goldberg, L.R., 1992. The development of markers for the Big-Five factor structure. *Psychol. Assess.* 4 (1), 26–42. <https://doi.org/10.1037/1040-3590.4.1.26>.
- Hall, S., 1991. The local and the global: globalization and ethnicity. In: King, A.D. (Ed.), *Culture, Globalization and the World-System*. Department of Art and Art History, State University of New York at Binghamton, pp. 18–39.
- Holmes Finch, W., Hernández Finch, M.E., 2020. Poverty and Covid-19: rates of incidence and deaths in the United States during the first 10 weeks of the pandemic. *Front. Sociol.* 5 (47). <https://doi.org/10.3389/fsoc.2020.00047>.
- International Organization for Migration, 2019. *World Migration Report 2020*. IOM, Geneva, Switzerland. <https://publications.iom.int>.
- International Organization for Migration, 2022. *World Migration Report 2022*. <https://worldmigrationreport.iom.int/wmr-2022-interactive/>.
- Jackson, R., 1997. *Religious Education: An Interpretive Approach*. Hodder and Stoughton.
- Jackson, R., 2004. *Rethinking Religious Education and Plurality: Issues in Diversity and Pedagogy*. Routledge Falmer.
- Knauss, M.A., Günther, K., Belardi, S., Morley, P., von Lersner, U., 2015. The impact of perceived ethnic discrimination on mental health depends on transcultural identity: evidence for a moderator effect. *BMC Psychol.* 3 (30). <https://doi.org/10.1186/s40359-015-0088-x>.
- Marsden, S.V., 2017. Negotiating difference in education. In: Panjwani, F., Revell, L., Gholami, R., Diboll, M. (Eds.), *Education and Extremisms: Rethinking Liberal Pedagogies in the Contemporary World*. Routledge, pp. 177–190.
- McCall, L., 2005. The complexity of intersectionality. *Signs* 30 (3), 1771–1800. <https://doi.org/10.1086/426800>.
- Morrison, T., 2014. Interview With Stephen Colbert. *The Colbert Report*. The Daily Show, Comedy Central, New York. <https://www.dailymotion.com>.
- Moutafis, G., (n.d.). *The World's Refugees in Numbers*. Amnesty International. Retrieved [insert date] from <https://www.amnesty.org>.
- Nelson, T.D., 2009. *Handbook of Prejudice, Stereotyping, and Discrimination*. Taylor & Francis.
- National Human Genome Research Institute, September 7, 2018. *Genetics vs. Genomics Fact Sheet*. <https://www.genome.gov/about-genomics/fact-sheets/Genetics-vs-Genomics>.
- Noddings, N., 2013. *Caring: A Relational Approach to Ethics and Moral Education*, second ed. University of California Press.
- Pandey, G., 2010. Politics of difference: reflections on Dalit and African American struggles. *Econ. Polit. Wkly.* 45 (19), 62–69. <https://www.jstor.org/stable/27807002>.
- Ram, K., Yadav, S., 2021. The impact of COVID-19 on poverty estimates in India: a study across caste, class and religion. *Contemp. Voice Dalit* 1–15.
- Roy, A., April 3, 2020. The pandemic is a portal. *Financial Times*. <https://www.ft.com/content/10d8f5e8-74eb-11ea-95fe-fcd274e920ca>.
- Rutherford, J., 1990. *Identity, Community, Culture, Difference*. Lawrence and Wishart.
- Schiller, H.I., 1985. Electronic information flows: new basis for global domination? In: Drummond, P.E., Paterson, R. (Eds.), *Television in Transition*. BFI Books, pp. 11–20.
- Sen, A., 2006. *Identity and Violence: The Illusion of Destiny*. Norton and Company.
- Sussman, R.W., 2014. *The Myth of Race: The Troubling Persistence of an Unscientific Idea*. Harvard University Press.
- Svetaz, V.M., Coyne-Beasley, T., Trent, M., Wade Jr., R., Ryan, M.H., Kelley, M., Chulani, V., 2020. The traumatic impact of racism and discrimination on young people and how to talk about it. In: Ginsburg, K.R., McClean, Z.B.R. (Eds.), *Reaching Teens: Strength-Based, Trauma-Sensitive, Resilience-Building Communication Strategies Rooted in Positive Youth Development*, second ed. American Academy of Pediatrics [https://www.adolescenthealth.org/SAHM_Main/media/Anti-Racism-Toolkit/Traumatic-Impact-of-Racism-on-Young-People-\(1\).pdf](https://www.adolescenthealth.org/SAHM_Main/media/Anti-Racism-Toolkit/Traumatic-Impact-of-Racism-on-Young-People-(1).pdf).
- Tajfel, H., Turner, J., 2001. An integrative theory of intergroup conflict. In: Hogg, M.A., Abrams, D. (Eds.), *Intergroup Relations: Essential Readings*. Psychology Press, pp. 94–109.
- Taylor, C., 1994. Politics of recognition. In: Goldberg, D. (Ed.), *Multiculturalism: A Critical Reader*. Blackwell Publishers, pp. 75–106.
- Thorat, S.K., 1979. Passage to adulthood: perceptions from below. In: *Identity and Adulthood*, pp. 65–81.
- Thirumal, P., April 2021. Dalit Perspective on (Isabel Wilkinson's) Caste and the Lies that Divide Us. *Academia Letters*. <https://doi.org/10.20935/AL891>. Article 891.
- UNESCO, 1990. *World Declaration for Education for All*. https://www.campaignforeducation.org/wp/wp-content/uploads/2018/04/JOMTIE_E.pdf.

- United Nations High Commissioner for Refugees, 2019. Global Trends. Forced Displacement in 2018. UNHCR, Geneva, Switzerland. <http://www.unhcr.org/statistics>.
- United Nations, 2015a. The Millennium Development Goals Report. <http://www.un.org>.
- United Nations, 2015b. Transforming Our World: The 2030 Agenda for Sustainable Development. <https://sustainabledevelopment.un.org>.
- Van Vugt, M.V., Schaller, M., 2008. Evolutionary approaches to group dynamics: an introduction. *Group Dynam.* 12 (1), 1–6. <https://doi.org/10.1037/1089-2699.12.1.1>.
- Verhoeven, M., Poorthuis, A.M.G., Volman, M., 2019. The role of school in adolescents' identity development: a literature review. *Educ. Psychol. Rev.* 31, 35–63. <https://doi.org/10.1007/s10648-018-9457-3>.
- Vertovec, S., 2007. Super-diversity and its implications. *Ethn. Racial Stud.* 30 (6), 1024–1054. <https://doi.org/10.1080/01419870701599465>.

White woman: or, the abused abuser's role in U.S. educational stratification

Zeus Leonardo and Joy Esboldt, Graduate School of Education, University of California, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

"White woman" as an ideal type: the conceptual framework of racial abuse as a gendered project	94
Historical construction of white woman as ideal type	96
Historical archetype meets present day	98
References	101

The US and other parts of the world are experiencing a racial upheaval. During former US president Trump's run up to the election and the following 4 years as president, the US nation felt the racial divide perhaps only comparable in recent memory to the Nixon era. Before that, we may have to recall the Civil War. For many people, Trump's rise conjured up an old America that he and his supporters wanted to recapture, emblazoned on red MAGA hats to "Make America Great Again." His performance of toxic masculinity, curtsying to white nationalism, and exhibition of personal narcissism were difficult to miss, even to the casual observer who may or may not have voted for him.

However, a key ingredient to the formation of the New or Alt-Right is the participation of white women, without whom it would be facile to pin the upheaval on Trump, despite that he may wish for that. Pointing out a couple instances should suffice for now. White women's vote for Trump lagged behind their vote for Hillary Clinton but not as significantly as would have been predicted for the first woman presidential candidate. The appointment of several white women to important posts in Trump's administration enabled Trump to construct a Rightist (some call neo-fascist) regime that reached across the gender aisle. This included Secretary of Education, Betsy DeVos of the famous Michigan billionaire family at the head of Amway company. Last, on January 6, 2021, just two weeks before the handover of power to President-elect Joe Biden and Vice President-elect Kamala Harris, the first Black and South Asian woman to hold that title, the Capitol insurrectionists stormed congress to try to stop the process of officially counting the electoral votes that would eventually seal Trump's fate as a one-term president. At that siege, there were many white women, including those who breached the Capitol building and Ashli Babbitt, the young veteran woman who was fatally shot in a skirmish with a Capitol police officer (Barry et al., 2021).

In more "mundane" instances, the white woman phenomenon, dubbed "Karen," makes the headlines when white women summon the authorities to criminalize otherwise harmless behavior from Blacks. "Karen" is a stereotyped white woman with passive-aggressive tendencies. For example, holding an outdoor barbeque, like an incident at Oakland, CA's Lake Merritt, prompts a white woman to file a complaint with the police department (Wood, 2019). In another incident, Amy Cooper reported a "threatening" Black man, Christian Cooper, who requested that she leash her dog in Central Park, NY. Amy Cooper has since publicly apologized, faced legal charges for falsely accusing Christian Cooper, fired from her job at an insurance company, and a mental illness defense for her racist rant was been launched (Nir, 2020). The Cooper episode happened in the middle of the Black Lives Matter (BLM) Movement and the public recognition that Black men stand the chance of being killed by police officers for countless non-violent offenses, as with the case of George Floyd who wrote a bad check and was dead by the end of the day. For that matter, Black people are accused of non-offenses, as with a white woman who accused a Black teenager of stealing her cell phone in a hotel where the young man and his father were staying. Naively, which should not be confused with innocently, hailing the police is a deadly decision, weaponized by Amy Cooper and other white women who seek the protection of the racialized nation state and its racial apparatuses, including its repressive forms (see Leonardo, 2005, 2015).

It should not be lost on the reader that Amy Cooper's "insanity defense" fits two profiles. One, it reinforces an older idea of racism as evidence of an irrational mind, rather than part of whites' rational and racial interests. Two, it deploys a patriarchal image of women as hysterical or emotional to the point of madness. Of course, we may simply excuse the entire defense as self-serving and convenient, which may be the case. But consistent with our tack in this chapter, we emphasize Amy Cooper and her precursors' racialized and gendered history, like Carolyn Bryant who was directly responsible for Emmett Till's lynching. In a cruel twist of fate, Bryant has since confessed that certain parts of her accusations of Till were false (Tyson, 2017). A little too late.

As it concerns Cooper's mental incapacity defense, it may work to soften the vitriol she received since the Central Park episode, but she is diminished by her association with being insane. A white man using a similar escape hatch is more difficult to imagine, perhaps only as a last-ditch effort, because a racial patriarch is assumed to be rational even when he is racist. Although admittedly from a different era, Carolyn Bryant's husband, Roy Bryant, Emmett Till's murderer, would not have admitted to the madness of his behavior. It was considered a rational response to a perceived racial threat. On the other hand, Amy Cooper may be dismissed by the casual observer as an irrational, hysterical woman. This political trade may preserve her racial advantage but subverts her gender interests as a member of the "second sex" (de Beauvoir, 2009/1949). This brief portrait of white women's participation in race

projects happens in the backdrop of a patriarchal system that asks for their loyalty to whiteness. It serves as the poignant example of the compromises an injured group makes with whiteness in exchange for a modicum of protection from the injuries they face as a targeted group within other social relations, like gender, or class, as with the white working class. As this chapter points out, this imperfect union does not ameliorate, but rather deepens, white women's real and perceived injuries.

Education is acknowledged as a racial project on one hand, or a gender project on the other. Both lines of research have a long-standing presence in educational research and the disparities we find in each relation is remedied by suggestions around anti-racism and anti-sexism in schools, respectively. However, for the most part their engagement has remained separate from each other. Powerful explanations exist (see Crenshaw, 1991; Hooks, 1984; Collins, 2009; Pateman and Mills, 2007; Villenas, 2010; Coloma, 2012) but they have not been able to shift the general rule that race and gender analyses are developed independently of each other within education research. As a result, educational scholars lack a robust framework for understanding racialized sexism or sexualized racism as points of intersection in order to arrive at a more complete picture of subordination as it concerns students of color of any gender. For example, Critical Race Theory (henceforth CRT) in education has advanced an elegant empirical and theoretical study that favors race and subsumes a study of gender within the general predicament of racialization. Said another way, CRT in education is first and foremost a race intervention. On the other hand, feminist work in education has long been weak at addressing racial subordination in exchange for a singular attention to gender with the occasional synthesis with either race or class. With gender as primordial, race-evasive analysis constructs race as somebody else's problem, usually people of color.

Recent research in education has begun to focus on the fact that the vast majority of teachers are white, most of them being white women. This relatively well-known data point has not spurred on a concerted effort to understand the particularities of what white women bring to the raced and gendered dimensions of social life, considered as entangled within the educational enterprise. However, this bundled analysis is becoming more difficult to neglect, especially as educators predict that students of color soon will become the majority population in public schools. Said another way, more students of color will be taught by a majority white woman teaching population. With white women at the helm of most classrooms, racism does not appear general or ungendered, but arguably in a feminine form (see Leonardo and Boas, 2013, 2022; Leonardo and Gamez-Djokic, 2019; Miller and Lensmire, 2021). Although they do not call the shots in a patriarchal system, white women sometimes pull the racial trigger.

We submit the thesis that a feminized or euphemized form of racism exists in schools. White women teachers' interactions with students of color necessitate a deeper look at, or a critical theory of, the ideal type in classrooms: "white woman." Admittedly, there are different ways to be a white woman, but our task is to understand what has continued and continues to structure this educational archetype, including the role that it has played historically in how students of color experience education. As we show, white women occupy a contradictory position as both oppressor and oppressed, abuser and abused, injurer and injured. In other words, white women are targets of sexism and patriarchy even as they embody white privilege and enact racism in education. This double role is not a controversial claim but points out that an appropriate understanding of white women's place in education needs to account for both dynamics of their subordination by gender and superordination by race. We weave several lines of thought to support this analysis. In particular, we draw on historical scholarship on whiteness to overview the construction of the white woman teacher ideal type as well as critical race and intersectional scholarship, which offer insights into the contemporary structures of teaching linked to this ideal type. Additionally, and foremost, we leverage Thandeka's (2000) book, *Learning to be White*, for its generative framework of "white racial abuse," which we explain in the next section.

"White woman" as an ideal type: the conceptual framework of racial abuse as a gendered project

Few publications in the last several years have been more provocative than Reverend Thandeka's someday to be classic text, *Learning to be White*. Although it is more common and commonsensical to regard people of color as the actual victims of racial abuse, Thandeka inverts the analysis to suggest that the process of becoming white takes the form of abuse, in both the clinical and symbolic sense. At first glance, this innovation is provocative yet counter-intuitive, for it suggests a first or primary racial injury, which occurs in the psycho-social development of a child who transforms from a human to white child through cultural, psychological violence. Bearing the marks of a universal human, not unlike other children, the white child begins the narrowing process that negates their universality in exchange for the mythical status of becoming white. This is different from other forms of specificity that all humans experience as they enter an ethnic, a cultural, or national milieu, a process that integrates them into those formations, admittedly having their own set of problems. However, for the child who becomes white, they must go through *dis-integration*. Their core sense of self undergoes a fragmentation, and they sacrifice self-love for group loyalty to whiteness, treason to which either incurs more abuse or excommunication from the only community they have known thus far. Throughout the many stages of development from childhood to adulthood, being white is akin to picking up the shards of identity leftover from the broken mirror of whiteness. In modern terms, whites assume the tragic figure that would dwarf the fallen heroes in Homeric poems.

As part of partaking in whiteness, whites sacrifice human essentials, such as unconditional love from caretakers, freedom from emotional abuse, and curiosity about the other. Overall, becoming white is a process of negation and separation, sedimented in the human subject in what Thandeka calls white racial shame, or *shame*, plainly speaking. Opposite the assumption that whiteness cultivates a narcissistic self-love at the cost of loathing nonwhite others, or *blancophilogenesis* (see Leonardo, 2015), Thandeka shifts our

understanding of whiteness as *prima facie* evidence of a fundamental self-hate. In fact, in one of Thandeka's case studies of Podhoretz, a Jewish man struggling with his second-class status as a "facsimile white," Thandeka writes, "In this skewed scenario, the Negro anti-Semite becomes a mask: the blackface of Podhoretz's own white shame. His rage against the Negro masks his own white shame (or his own Jewish anti-Semitism) by making it unrecognizable" (p. 33). The convenient reasoning is that as a non-WASP, Podhoretz shields anti-semitism with feelings of anti-blackness, from which Thandeka departs by suggesting that it is instead a lifelong *intra*racial struggle within whiteness projected onto Black people. Rather than a victorious moment, becoming white is a collection of moments of defeat experienced by the child, usually perpetrated by their caretakers who should also be understood as victims of having become white. It is triumphalistic at best. This arrested development continues well into adulthood, that is, unless whites experience another upheaval, this time anti-racist.

Our main purpose in this section is, first, to introduce and interpret Thandeka's abuse framework, and second, to illuminate what appears to be an underdeveloped gender dimension to her framework. White racial abuse is a psycho-social explanation of white development at the hands of caretakers. Although it is reasonable to assume that Thandeka uses caretakers in a broad sense that includes family, teachers, the social welfare system, even one's government, the primary unit of care remains the family. Here, Thandeka does not offer care (or its cognate, love) as only a positive practice, desire, or principle. By care, she means a relation that one enters whereby a caretaker, such as parents or guardians, provides for or protects the cared-for, in this case children (concerning whites of any age, we may say they are children of the white nation). This caring act is not an unconditional good and when race enters the picture, may qualify as its very opposite: violence (cf. [Leonardo and Porter, 2010](#)). Contrary to traditional race understanding, Thandeka reinterprets the newly inducted white person as racial violence's primary target, its first victim.

There are other moments in history when claims to caring for others resulted in violence. Colonialism is one such case, or when European powers claimed to save natives from themselves by caring for their soul in the form of religious conversion (see [Coloma, 2012](#)). Caring for native souls arguably represents the highest form of concern for others, if you follow the colonizer's rationale. The logical outcome of colonial care, at least to colonizers, is gaining salvation in God's kingdom. Religious conversion and other colonial events in history happen through masculine coercion but they just as often transact through feminine persuasion, imperial western women being the favored agent of colonialism's kinder and softer form. Education was often a central, feminine mechanism of this process ([Jacobs, 2009](#)), which we discuss more fully in the following section. In short, caring becomes the primary cultural role of women, mothers, nuns, and in the case of contemporary, public education, white women teachers. Again, this does not mean that white women somehow "control" the process, because they are under the surveilling eyes of the patriarchal gaze. Thus, a caring framework is more convincing and complete when a gender component is integrated into it, something to which education is not a stranger (see [Noddings, 1984](#); [Beck, 1992](#); [Valenzuela, 1999](#)). Appropriating Thandeka's elegant framework for education necessitates integrating a gender analysis to the assertion of white racial abuse and the shame that accompanies it for the specificities of studying what [Althusser \(1971\)](#) once called the educational ideological state apparatus (ISA), reframed as a racialized ideological state apparatus (RISA) (see [Leonardo, 2005](#)) or the educational racial contract (ERC) (see [Leonardo, 2013; 2015](#)).

But there is more to it. If taken literally, then caring's ironic, violent consequences are somehow associated with white women, the primary caretakers of white children, both in the family and schools. As victims of structural gender abuse, women are targets of patriarchal violence and therefore enter the process of racial abuse as double subjects: the *abused abuser*. Any accounting for their role in caring for or raising white children includes accounting for their subjective reduction as the "caring gender" within structural, patriarchal relations. This move does not deny their part and participation in racism. On the contrary, white women, in this case teachers, must reflect on their racial compromise with whiteness, the privileges of which compensate for their gender injury by perpetrating a different kind of violence, this time directed at people of color of any gender. To Thandeka, the original abuse and the ensuing sense of being under perpetual siege or attack experienced by white humans, explains their tragic inability to understand their racial predicament. To this, we add that, in schools, understanding white women teachers' racial development is simultaneously an accounting for their gendered participation in the racial army.

To be clear, caring here should be understood as an empty signifier with a neutral status, that is, until it is filled with a political project. This is not unlike the dialectics of violence one finds in Fanon's *The Wretched of the Earth*, where violence possesses a cleansing effect, partly in its ability to inject a sense of self-respect in Black natives who become *actional* humans through the catharsis of anti-colonial violence (see [Fanon, 2005](#); for an explanation on the dialectics of violence, see [Leonardo and Porter, 2010](#)). Likewise, with respect to race history, caring is not the fuzzy, warm feeling that humans normally understand by or associate with that term. In other words, as an empty category, caring may be paired with racism, as was the case during colonization when nuns and priests "cared" for the natives' soul. Likewise, when filled with a white supremacist project, caring within the white family, schools, neighborhoods facilitates the subjective transformation of human beings into white people, which is an assault on their core sense of self. Racialized caring relation is the beginning of a lifelong process of feminized racial violence that most whites are ill-equipped to understand when they repress what it means to have undergone said transformation. White children's first and last teacher is the mother figure by virtue of a sexual contract in place ([Pateman, 1988](#)), which relegates women to the role of mothering, even if not all women are in fact mothers.

Like classical abuse, white racial injury is assimilated into the white child's psyche as a confusing set of orientations to self and others. And like such abuse, the abused often becomes an abuser of others as the cycle repeats itself if more humans are to become white, generation after generation. Focusing on childhood memories has been key to even a popular cultural understanding of psychology and psycho-therapy, so formative childhood development is for the adult that humans eventually become, racial or

otherwise. For white adults, who comprised Thandeka's participants, childhood memories of becoming white were ghostly and haunting. This racial remembering is responsible for the deep shame that follows whites well into adulthood. But their white memories were evident and Thandeka needed only to scratch the racial surface in order to elicit stories and emotions, ranging from white lessons over dinner table conversations, bringing home a Black friend, running into Black people on the street, a Black family moving into the neighborhood, or the arguably worst of all taboos, interracial dating. This means that the "white pride" calling that far Right groups use to recruit members is not that ironic, if we consider that white humiliation is part of becoming white.

Of all these emotions, Thandeka elevates the feeling of *shame* as one of the most debilitating feelings that whites experience. It is a stain that whites carry with them basically all of their life, a burdened or heavy feeling that something about the self they have become is not quite right. Its social version is felt as embarrassment because whereas it is quite natural to share childhood memories with friends or strangers, white racial memories, as recollections of abuse, come with (1) a sense of sinful confession (2) distress from potential judgment and (3) evidence of their unlovability and (4) fear of exile from one's community (on white exile, see also Leonardo, 2020). Wanting integration, they fear further disintegration; desiring connection, they are paralyzed by separation anxiety but fulfill it as the only way they have known in order to survive; craving authenticity, whites give up the reality principle for a mystical understanding of the racial world. Or as Thandeka distinguishes, feelings of guilt are evidence that one has done something wrong, whereas shame is the acknowledgment that one is fundamentally wrong.

With all of these emotions in play, it is not surprising that many of Thandeka's white respondents end up in tears. However, she reminds us that they weep not primarily because they feel shame over their ostensible racism toward people of color but over whites' own fragmentation. Theirs are not mere crocodile tears, but represent the fundamental loss associated with becoming white. Of course, white loss is a conflictual phrase since to people of color, the loss is more accurately called a *trade*, and a trade up at that. In other words, whites traded their ethnicity and language for social promotion to whiteness; for many people of color, including immigrants, such a loss simply amounts to a demotion since they received no white dividends at the end of the process. The former is a sign of shame, the latter a sham. This fact notwithstanding, white trade comes with structural value, so it comes as no surprise that it is coveted by previously non-white subjects even as it exacts its cost and demands its pound of white flesh. This is where Thandeka departs from Du Bois' (1935/1998) phrase and later Roediger's (1991) appropriation of "wages of whiteness," or the public and psychological advantages of being or perceived as white. Consistent with Thandeka's framework that stresses the cost of becoming white, Thandeka pivots from a different preposition in what she calls "wages *for* whiteness" (p. 79; italics in original). This deceptively simple prepositional replacement is replete with meaning, such that it emphasizes the cost of becoming white for the person who is abused into whiteness. Whereas "wages" is more often interpreted as the financial dividends like salaries, it may also take on a second meaning as it evokes the internal war that a white person wages by becoming white, a storm that rages in their fractured identity (see Levine-Rasky, 2016).

Teachers are often referred to as *in loco parentis* (see Dreeben, 1968) insofar as they take on the role of surrogate parents for the youth. More to our point, because most teachers are white women, they more accurately assume the specific role of *in loco mater* in public schools. This means that their *idealized* reproductive role in the nuclear family as mothers extends from the dinner chairs to the classroom seats. Their racialized emotional labor, or care, makes them central figures in the lives of both white and nonwhite children. As it concerns the former, the racial abuse that white children experience goes well beyond the private and includes the public sphere. As an extension of the matriarchal figure in white children's intimate lives, white women teachers continue the shame of whiteness that children carry, the disappointing of which would lead to a thousand small deaths in millions of small humans. In order to understand more fully the toll that racial abuse takes on white women and the children they teach, it is necessary to excavate the history that produced them as subjects, which is our task in the following section.

Historical construction of white woman as ideal type

In this section, we examine the racialized and gendered ideologies of the white woman archetype and their articulation with structures of teaching in order to examine the contemporary education landscape. Like other work on the historical construction of whiteness, Thandeka explores the classed compromises and costs involved in "the right to be white" (p. 87) and tangentially notes the subjugated position of white women. Although Thandeka focuses on the system of white racial abuse used to exploit the white working class, we illuminate the insights her framework might offer for a gendered analysis of whiteness.

Thandeka draws from legal race scholars to map the historical construction of and the socialization into whiteness. Keeping her focus ultimately on the psycho-social rather than the material, Thandeka argues that law constructs whiteness, and race more broadly, by concretizing abstract ideas and beliefs through a system of discriminatory laws, whereby material differences across racial identities are created and then maintained. As these laws create walls between racial groups, they also abuse lower status whites into a psychological allegiance to the elite, offering not protection but rather a restraint on the amount abuse one faces. While Thandeka develops a historical overview on the racial compromise of working-class whites (read: men), with caution not to flatten dynamics of class and gender into replica relations of power, here we develop we want to further develop this historical analysis in relation to gender and the racial compromises white women enter.

White women's legal identity was always constructed through their relationship with men and their rights and privileges through their gendered participation in the white racial army. Understanding the gendered and sexual politics of race laws created and used by white elite men, illuminates how white women have been produced as disintegrated subjects and socialized into the archetype of the abused abuser, at times through their explicit inclusion and at times through their absence. For example, in the fifty-two

prerequisite cases on racial differentiation leading up to the US Supreme Court cases of *Ozawa* and *Thind*, the judiciary logic never engages or names any women (Lopez, 2006). In *White by Law*, Lopez (2006) acknowledges this absence of women as evidence of the context that doubly restricted white women "as individuals and less than individuals (that is, as wives)" (p. 34). In other words, white women's whiteness became legally constructed not through their independent whiteness, but through their dependent relationship with white men (see also Mohanram, 2007). Yet, white womanhood and ideologies of femininity have been explicitly intertwined with the legal construction of race since the early US colonial laws. Thandeka notes how America's naturalization laws, which until 1931 automatically stripped white women of her US citizenship when marrying a "nonwhite" noncitizen, offer an example of how "a white is created by law" (p. 85) and how whiteness is a legal object (Harris, 1993) that may be gained and lost. In the 1600s, Maryland and Virginia made it illegal for whites and Black people to marry. Bans against interracial partnership along with the matrilineal inherited status of children functioned to restrict and regulate white women's intimate relationships, police and punish Black men, and legitimate and legalize the rape of enslaved Black women. With westward expansion, the US government broadened marriage restrictions to other non-white groups beyond Blacks, like Filipinos.

Entangled with ideas of the innocence and purity of whiteness, these laws codified gendered ideologies of race and racialized ideologies of gender. In the process, white women were defined as the symbolic opposite of Black women (Collins, 2009; Lugones, 1994). White innocence (Dyer, 1997; Painter, 2010), simultaneously gendered, cast white women as pure, at risk of corruption, or fallen victims, directly relying on the foil of Black women as strong, queer, and/or jezebels. In addition, Black men were constructed as criminal or contagion, such as the myth of the Black male rapist (Davis, 1990; Wells-Barnett, 1900). Gendered racism afforded white men a more self-righteous morality and constructed its complementarity in the form of a white "damsel in distress," needing a white male savior for protection and justifying violence against Black men. Repression and violence against indigenous people were carried out under a similar justification: to protect white women against threats of "Indian depredations" (Jacobs, 2009, p. 22).

Although governed by laws they did not themselves create, white women throughout history have worked on both sides of racism: participating in and being against it. In other words, the racial compromise white women strike with whiteness today is not new. Recent historical scholarship refutes conceptions that white women have been passive bystanders to racial violence and reveals the ways they worked for and found power within whiteness, not *despite* their gendered position in whiteness but precisely *because* of it. For example, Jones-Rogers (2019) documents various legal and extra-legal ways white women successfully wielded power to profit from slavery in the American South, despite, and at times because of, their lower gender status position, what Stepan (1986) calls the "lower race" of gender" (p. 264). In addition to full participation in the brutalities of the slave market, white women capitalized on specific feminine forms of participation in slavery, including making and regulating the market of wet nursing and the motherly duty of socializing white children into future positions of racial power. But again, whiteness exacts its moral price and white women give up profound parts of becoming human in exchange for becoming white humans. As noted, whereas Du Bois and Roediger prefer to describe this process as the accumulation of privilege, captured above by the "wages of whiteness," Thandeka (2000) turns the proposition on a different preposition with "wages for whiteness." Whites gain material wealth through whiteness, but they lose something not wholly insignificant (or more significant, depending on one's perspective). They kill their spirit, in the process. More than a pound of flesh, losing a kilo of soul turns whites into incomplete and sometimes empty receptacles. Although they may think such material wealth justifies the means, whites are the poorer for it.

This feminized participation in race projects has frequently been intertwined with schooling, both in the socialization or care-taking of children into whiteness and in the violence directed at nonwhite others. Starting in the late 1880s, Jacobs observes that "progressive" white women leveraged their gendered-racialized subjectivities as "better mothers" and morally superior caretakers at a political and interpersonal level to advocate for indigenous child removal, Indian Boarding school enrollment, and ultimately land dispossession across the United States and Australia. Twentieth century imperial feminism in the Philippines similarly forged openings for gendered authority in schooling as a mechanism for "civilization" through mixed claims of motherly benevolence and need to discipline people of color and/or indigenous peoples (Coloma, 2012). Similarly, the expansion of Common Schools in the States structured racialized-gendered ideologies and invested them with the white woman archetype of the "good teacher" (see Broderick and Leonardo on the problem of "goodness," 2016). As Common Schools expanded across the US during the 90th century in part as a way to "control," "save," and "civilize" poor not-yet-white, immigrant youth and freed Black children, white women quickly became the ideal teacher as they supplied a cheaper source of labor than their male counterparts, assumed inherently to work better with small children, and facilitated men's move into more "respectable" positions of administration and leadership (Apple, 1988; Kaestle, 1983; Tyack, 1974).

Yet, this expansion of white women's reproductive work to include the school, did not alleviate their gender injury. Despite these desirable traits, white women teachers still held a troubled position. Stereotypes that reinforced ideas of white women as caring, nurturing mothers also positioned them as weaker, non-rational actors (Ehrenreich, 2002). White women were regarded as limited in their capacity, never in full control, and needing a more rational male principal or superintendent overseer, particularly in the upper grades. Women politically active in teacher organizing were seen as a threat to the social order, departing from their core characteristics of submissive, self-sacrificing, and pious, therefore dangerous to the school system and unable to care for children (D'Amico, 2017). This gendered-racial occupational history solidified into ideas of professionalism in teaching, which mirror forms of managerial control more than the professionalism of autonomy, expertise, and status frequent in white-male-dominated industries, undergirding teacher-blaming discourses in contemporary reforms (D'Amico Pawlewicz, 2020). In short, as teaching became known as a white woman's profession, notions of professionalism were in turn restricted and always partial due to the gendered composition of teachers.

This historical white feminization of teaching structured racialized-gendered ideas of “good teaching” ultimately produced the white woman as the ideal teacher archetype, well positioned to carry out racialized care’s violent work. Understandings of “good” teaching and “good” teachers (see [Ladson-Billings, 2009](#)), as well as the structures that accompany these ideas, both produce white woman as key actors within the racialized ideological state apparatus of schooling and shapes the exclusion and displacement of Black teachers. Against the white woman norm, substantiating the work of Black teachers was deemed disruptive. For example, following the desegregation of schools post-*Brown v. Board of Education*, the authority and positioning of Black educators was challenged when white parents deemed them unfit to teach their children. Strategic federal, state, and local policies across the US in the years following contributed to Black teachers’ massive, forced pushout from the profession even when they were better educated than white counterparts ([Fultz, 2004](#); [Tillman, 2004](#); [Walker, 2001](#)). Many of those who stayed in teaching entered hostile environments that did not recognize or validate their expertise. For many Black children this meant an interruption in an epistemology of teaching whereby good teaching was previously rooted in deep pedagogical and political commitment to the community.

On the other hand, considering the histories of white women in the Civil Rights Movement illuminates limitations and contradictions of white women’s anti-racist efforts given the compromises they enter and make with whiteness. Most white women and their social organizations involved with the civil rights movement were described as far from radical ([Evans, 1979](#); [Laville, 2017](#); [Little, 2009](#); [Newman, 1999](#); [Olson, 2001](#); [Schultz, 2001](#)). They were often recruited through religious or campus organizations understood as an extension of moral or religious gendered care work, and not necessarily political. Frequently with more status, wealth, and relationship privileges to lose than their Black counterparts, most white women’s participation was what Irons calls “low-risk institutional activism” (1998, p. 698). Although Irons speaks specifically about women in Mississippi, we can draw from this to attend to the contradictions in anti-racism from doubly positioned white women.

This limited or low-risk participation allowed white women a sense of contribution to the movement while not risking their economic, social, or political positions. As [Little \(2009\)](#) and [Evans \(1979\)](#) note, moderate or even conservative institutions became instruments against white supremacy and a space for developing radical consciousness and practice against racism and sexism for some white women. Furthermore, despite frequently functioning at the peripheries of Black organizing, white women’s greatest tool, or their influence within their white familial, social, and religious spheres, was noteworthy in the larger movement. However, while white women’s low-risk anti-racism may have strained or even challenged some relationships, it frequently sought to maintain the modicum of protection afforded to them by whiteness, often through white men’s patronage (thus, the phenomenon of “Karen” has a longer history than we acknowledge). As Irons documents, when their husbands faced backlash or social pressures at work due to white women’s involvement, many women stepped down from highly scrutinized efforts and adopted less public roles in the work, sensing their “internal nonwhite zone... under siege” ([Thandeka, 2000](#), p. 23). In other words, white women’s loyalty to whiteness comes with the cost of a double alienation from an authentic relationship with one’s truer self and with non-white others, “where impulses to community with persons beyond the pale are slaughtered” ([Thandeka, 2000](#), p. 24). While relationships with Black women were tentative at best and troubled at worst, many white women and their social organizations still expressed beliefs of racialized authority and/or struggled with their organizational histories of segregation even if they expressed desires to overcome them.

Rejecting missionary activism rooted in ideas of Lady Bountiful or fair play segregationist stances, a small number of white women embraced a more radical mission and stance. Rooted in deep interracial friendships and accountability, activist training, and intraracial networks of support, their work was frequently in the background, such as raising money, creating platforms, and recruiting within white communities to support Black-led organizations ([Olson, 2001](#); [Schultz, 2001](#)). The involvement of these more radical women came with labels of social deviance as “race traitors” and “communists,” with real personal, physical, and financial risk. These experiences may even invoke feelings of shame as white women who develop race consciousness must confront their own complicity. [Thandeka \(2000\)](#) explains, “The experience of shame is thus a negative self-exposure, a revelation of forbidden desires. The self exposed is incongruous with itself. It is seen as who it is not supposed to be. It feels what it is not supposed to feel. It is aware of what it is not supposed to know ... Shame acts as a kind of Check-point Charlie” (pp. 13–27). White shame is a constitutive part of becoming and being white, most intensely felt when white women, in this case, become at odds with their whiteness as they establish solidarity with people of color. Although scholars have argued that white women’s radical race consciousness created a unique threat for whiteness because of the pretenses of protecting white womanhood to justify racial violence (see [Tyson, 2000](#)), at the heart of these white women’s effectiveness was a deep participation in the work, networks, and leadership of Black folks and a revocation of group loyalty to whiteness. In other words, this work upends individualistic notions of care and instead roots care in a political stance and radical accountability.

Historical archetype meets present day

In the preceding sections, we provided a conceptual understanding of white womanhood through Thandeka’s framework of racial abuse, specifically white women as *abused abusers*. Next, we traced the history of policies, practices, and ideologies that incorporated white women in the patriarchal-white supremacist project and the unique role they played in its structure. In this final section, we examine the white woman as an archetypal positioning within contemporary education, blurring the

line, or perhaps tracing the lineage from how things were and how things are now. We draw on critical race and intersectional scholarship to evidence of the evolution and persistence of racialized-gendered ideologies and structures of an abused-abuser ideal type in contemporary patterns and accounts of teaching. While the gendered nature of teaching as “women’s work” still shapes the structures of deprofessionalization in teaching, we can also connect lines of how public-school teaching continues to keep racialized “others” out and care for white women, instructionally and affectively, even within career conditions of constraint.

First, while our focus in this section centers around contemporary structures and ideals that protect and position white women’s race work, a double analysis also acknowledges the contemporary structural mechanisms of sexism and patriarchy within the gendered profession of education, and the compromised position of teachers today. Recent education reforms have ushered a focus on standardization, regulation, and marketization that further disempower teachers. Across competing reforms, there has been an anti-union trend and a loss of union-bargained agreements within a larger context of disinvestment and austerity (Horsford et al., 2019). As standardized testing takes on more focus and higher stakes, teachers’ work is more controlled, narrowly quantified, and deskilled. The most recent and profound inequalities that have been exacerbated through the Coronavirus Pandemic, among other dynamics, serve as a powerful reminder that women are still more easily pushed out of the workforce, burdened with expectations of and exploitation through their reproductive labor (Federici, 2014).

Within race and gender-neutral policies and structures, strong evidence of the white woman archetype endures. There are several ways this archetype endures instructionally, or how white women receive disproportionate care as it relates to their instructional recruitment, retainment evaluation, and assignment. Following the pattern of Black teacher displacement post *Brown v. Board*, as discussed above, recent years reflect a similar dismissal of the expertise of teachers of color, albeit shrouded through a race-evasive logic in a neoliberal era (Scott, 2011). The massive firing and loss of bargaining power of veteran Black teachers in New Orleans post Hurricane Katrina, and the subsequent replacement by a majority-white, novice teaching force through state partnerships with organizations like Teach for America (TFA), provide perhaps the most obvious example (Buras, 2016). In New Orleans Parish Public prior to the storm, Black teachers made up 73% of teachers in 2004–2005, nearly mirroring student demographics. Post-Katrina, that number decreased by 71% the first full year following, and by 2014, Black teachers made up less than half the teaching body (Cook and Dixon, 2013). Rather than a mournful loss of professional and community expertise, for many policy leaders and reform advocates, the devastation from Katrina was viewed as an unprecedented opportunity for education redesign through characterization (Anderson and Dixon, 2016). A similar combined prioritization of education market logics and low cultural valuation of Black teachers has been traced to the declining Black teacher percentages in other cities such as Chicago, D.C., and Philadelphia (Delpit, 2012; White, 2016).

These patterns suggest the culture of community, place-based expertise, and racial justice activism frequently descriptive of veteran Black teacher, evade dominant understandings of what characteristics make a good teacher across the United States. In pursuit of market-based educational reform and increased performance on standardized tests, they are disregarded. The pervasive narrative found within the recruitment logic of programs like TFA, is that good teachers should be selected and recruited from the “best and the brightest” (Milner and Howard, 2013, p. 537) at predominantly elite, white and middle/upper-class institutions, even if they are not going to remain in the teaching profession (Trujillo et al., 2017). In addition to recruiting within predominantly white spaces, these narratives and the practices connected with them also narrow the scope of “good teaching” to prioritize particular forms of knowledge, such as curriculum or subject-based knowledge, and enable a more short-term or contract-based workforce. While these recruiting approaches may not have a particular gendered focus, their efforts align with the larger deprofessionalizing reforms shaped by the gendered roots of the profession.

Likewise, although widespread contemporary discourse, including from TFA, has acknowledged the need for more teachers of color, these calls do not necessitate a challenge to the positioning of the white woman archetype. While there is a robust and important body of work about the effectiveness and importance of teachers of color, calling for more teachers of color is not always connected to broader efforts toward racial justice (Philip and Brown, 2020). Isolated or simplified calls for more teachers of color can act as a “symbolic buffer” (White, 2016, p. 5) or discursive alibi for institutions and/or leaders, limiting redress of racial violence or transformative change. While TFA boasts a more racially diverse corps membership since shifting recruitment approaches in recent years, there is less evidence that such changes will address the relationship of their diversifying corps demographic to the realities that teachers of color face either more broadly or specific to the region of their placement. Across these accounts, shared lip service to and even action driven agreement around the need for more teachers of color has not required challenging racialized-gendered structures and ideologies of “good teaching” that constructed and undergird the white woman ideal type.

Even in traditional certification routes, teachers of color face persistent structural barriers across the stages of a teaching career. Read another way, white women are still disproportionately cared for within these structures that are still designed for them in their double role. Within a marginalized profession, white women do not face the same discrimination and barriers toward certification and once in the profession encountered by teachers (candidates) of color (Petchauer et al., 2018; Rogers-Ard et al., 2012; Sleeter, 2017). Like other academic arenas governed by standardized tests and “neutral” evaluation protocols, certification exams are plagued by exclusionary white, Eurocentric norms and standards, and reviews of portfolios are subject to reviewer racial bias.

In addition to shaping working environment and school culture, the white woman ideal type endures in teaching practices, structures, and evaluation. Substantial research has documented the cultural bias in teacher performance assessment. Presented as race-evasive and the universal standard, ideas of “good teaching” and “best practices” continue to be grounded on unnamed white-feminized scripts and norms (Acosta, 2019; Endo, 2019). This does not mean individual white women teachers are in charge of the process. But over time as this arrangement plays out, a gendered white systematization of “good” teaching functions as a diminished form of privilege across multiple systems for those that conform and are read as such. A brief digression into empirical literature offers evidence of the institutionalization of white feminine scripts and norms through locally determined systems of teacher evaluation (Campbell, 2020; Drake et al., 2019; Grissom and Loeb, 2017), national criteria for good teaching such as National Board for Professional Teaching Standards (Irvine and Fraser, 1998), federal indicators of “highly qualified teachers” (Acosta, 2019), in standardized performance assessments (Petchauer et al., 2018) and in hiring (D’amico et al., 2017). Like much policy constructed and implemented in this era, rarely is this overt. Rather, as norms and systems interact with one another they become solidified in particular privileges.

We offer two brief examples. First, recent evidence on the race gap in teacher performance finds that teachers’ observations scores fail to account for school and classroom context-specific factors resulting in disproportionate lower rankings of Black teachers (Steinberg and Sartain, 2020). Second, beginning to counter the paucity of research on education policy’s impact attentive to teachers’ multiple race-gender identities, Campbell (2020) presents compelling evidence that Black women teachers receive lower observation ratings across and within schools than their white women counterparts, even when accounting for effectiveness, and were twice as likely to be placed on punitive development plans. While this literature is relatively new, as analysis of race and gender still largely remain separate, this work underscores how ideas of good teaching and the accompanying structures (including certification, assignment, and evaluation) are anchored in a sociohistorical construction of teaching as white women’s work and grounded in the historical construction of white womanhood.

Furthermore, the white woman archetype affectively and organizationally shapes teaching assignments and training. As teachers enter the classroom for the first time, white women matching the archetype in their gendered and racial presentation are read as simply as a “new teacher,” their gendered-racial identity subsumed into the norm. What type of teacher they will become remains unknown, to be shaped by the individual, placement, and experiences. On the other hand, teachers of color frequently encounter preconceived, essentialist teaching roles, work that they are expected to conduct because of the color of their skin and their presumed predisposition to perform particular tasks. In these “pedagogical kind[s],” (Brown, 2012, p. 297) teachers’ of color subjectivities, pedagogies, and expectations have been constructed by white colleagues and systems prior to entering the classroom. For example, Black men are normalized as disciplinarians, their value dependent on their ability to govern Black boys and Latino men are asked to embody and signal a “corrective representation” of masculinity in their positions as representational role models (Brockenbrough, 2012; Singh, 2018).

Building off of Patricia Hill Collins’ work on the controlling images of Black women as Sapphire or Matriarch, Acosta (2019) describes how Black women teachers must fight widespread caricatures of superhero or bodyguard. These caricatures objectify Black women as the other and can lead to disproportionate work overload and the undermining of their effectiveness (see also Mensah, 2019). While much of this scholarship explores the unofficial ways in which these roles reproduce, scholarship on teacher sorting reveals that teachers of color are more likely to be assigned to lower-track classrooms than their white colleagues (Kalogrides et al., 2013). These pre-set roles undermine teachers of color’s professional expertise and frequently positions them to save or discipline students of color rather than lead or co-create rigorous learning environments. Scholarship on these pre-set roles also attend to how teachers of color are also called on to save their white women colleagues from students of color. While white women may be positioned as the “better mother,” colleagues of color are asked to step in and rescue white women from students that they deem hard to teach.

This affective rescuing of white women gets invoked not only in their instructional teaching roles from their students of color, but affectively in their professional development and learning spaces as well from their colleagues of color. Even as contemporary initiatives may strive to engage deliberately with race, across teacher education and teacher professional development, the protection of whiteness, and in turn white women, remains centered. Most multicultural teacher education courses “reflect White sensibilities” (Sleeter, 2017, p. 158) in their curricular content and pedagogical approach. Racial analyses in these teacher learning settings are commonly power-evasive and rarely lead to profound shifts in teacher beliefs or practice. Much like the institutional or discursive alibi of calling for more teachers of color sans a structural change, such power-evasive trainings can blunt broader racial justice efforts and even reify institutional norms. Candidates and teachers may learn about race or racism, but white standards, identities, and communities remain intact.

Ultimately, structures of whiteness and race-gendered privileges are maintained. Even teacher trainings that take a structural approach to racial justice struggle to shift beyond prioritizing white women’s emotional needs, sensitivities, and consciousness (Kohli, 2016; Matias, 2016). Within these spaces, a plethora of accounts tell the story of a white woman feeling uncomfortable with the discussion of racial injustice, sometimes breaking down in tears and leaving the conversation, whether physically exiting or by shifting the focus and content of the conversation to center their emotions. In this well-rehearsed scene, white women’s tears function as an affective technology of to reinforce the rationality and positioning of white men to come in and save them whiteness (Hook, 2005; Leonardo and Zembylas, 2013). The parallels to the earlier examples of Amy Cooper are not lost here, where racism takes on a feminine form of helplessness, this time not enacted against students but colleagues of color.

In this chapter, we argued for greater appreciation of the role and history of white women in education. As the majority of the US teaching force, white women represent the primary subjects in the enactment of whiteness with children of color. If it were as simple as that, we could stop short of a deeper analysis. As targets of patriarchal power, white women occupy a structural position with a profound sense of ambivalence. In short, they are abused abusers whose measure of racial power and privilege is compromised by their position in gender hierarchy. However, this dual position holds out the possibility that white women teachers may yet recognize their own injury and use it as a basis to fight against racial injustices by becoming “bad” white subjects, that is, by resisting the injuring of others. That this has not happened, by and large, speaks to the difficulties of transferring insights about one form of abuse to understanding another set of relations, the complications among diverse and sometimes divergent political interests, and the challenge of a more complete social analysis. In our own attempts, we tried to unveil the contradictions found in the archetype of the “white woman,” whose whiteness is both specific to the institution of schooling as it is currently structured and the place of white women within the evolution of whiteness.

References

- Acosta, M.M., 2019. The paradox of pedagogical excellence among exemplary Black women educators. *J. Teach. Educ.* 70 (1), 26–38. <https://doi.org/10.1177/0022487118808512>.
- Althusser, L., 1971. *Lenin and Philosophy*. B. Brewster (Trans.). Monthly Review Press.
- Anderson, C.R., Dixon, A.D., 2016. Down by the riverside: a CRT perspective on education reform in two river cities. *Urban Educ.* 51 (4), 363–389. <https://doi.org/10.1177/0042085916638749>.
- Apple, M., 1988. *Teachers and Texts: A Political Economy of Class and Gender Relations in Education*. Routledge.
- Barry, E., Bogel-Burroughs, N., Philipps, D., 2021. Woman Killed in Capitol Embraced Trump and QAnon. *The New York Times*. <https://www.nytimes.com/2021/01/08/us/who-was-ashli-babbitt.html>.
- Beck, L., 1992. Meeting the challenge of the future: the place of a caring ethic in educational administration. *Am. J. Educ.* 100, 454–495. <https://doi.org/10.1086/444025>.
- Brockenbrough, E., 2012. Agency and abjection in the closet: the voices (and silences) of Black queer male teachers. *Int. J. Qual. Stud. Educ.* 25 (6), 741–765. <https://doi.org/10.1080/09518398.2011.590157>.
- Broderick, A., Leonardo, Z., 2016. What a good boy: the deployment and distribution of “goodness” as ideological property in schools. In: Connor, D., Ferri, B., Annamma, S. (Eds.), *DisCrit: Critical Conversations across Race, Class, & Dis/Ability*. Teachers College Press, pp. 55–67.
- Brown, A.L., 2012. On human kinds and role models: a critical discussion about the African American male teacher. *Educ. Stud.* 48 (3), 296–315. <https://doi.org/10.1080/00131946.2012.660666>.
- Buras, K.L., 2016. The mass termination of Black veteran teachers in New Orleans: cultural politics, the education market, and its consequences. *Educ. Forum* 80 (2), 154–170. <https://doi.org/10.1080/00131725.2016.1135379>.
- Campbell, S.L., 2020. Ratings in black and white: a quantcrit examination of race and gender in teacher evaluation reform. *Race Ethn. Educ.* 1–19. <https://doi.org/10.1080/13613324.2020.1842345>.
- Collins, P.H., 2009. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, third ed. Routledge. First published in 1991.
- Coloma, R.S., 2012. White gazes, brown breasts: imperial feminism and disciplining desires and bodies in colonial encounters. *Paedagog. Hist.* 48 (2), 243–261. <https://doi.org/10.1080/00309230.2010.547511>.
- Cook, D.A., Dixon, A.D., 2013. Writing critical race theory and method: a composite counterstory on the experiences of black teachers in New Orleans post-Katrina. *Int. J. Qual. Stud. Educ.* 26 (10), 1238–1258. <https://doi.org/10.1080/09518398.2012.731531>.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and the violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299. <https://doi.org/10.2307/1229039>.
- Davis, A., 1990. *Women, Race, and Class*. Random House.
- de Beauvoir, S., 2009. *The Second Sex*. C. Borde and S. Malovany-Chevallier (Trans.). Random House. First published in 1949.
- Delpit, L., 2012. “Multiplication Is for White People”: Raising Expectations for Other People’s Children. *The New Press*.
- Drake, S., Auletto, A., Cowen, J.M., 2019. Grading teachers: race and gender differences in low evaluation ratings and teacher employment outcomes. *Am. Educ. Res. J.* 56 (5), 1800–1833. <https://doi.org/10.3102/0002831219835776>.
- Dreeben, R., 1968. *On What is Learned in School*. Addison-Wesley.
- Du Bois, W.E.B., 1998. *Black Reconstruction in America, 1860–1880*. The Free Press. First published in 1935.
- Dyer, R., 1997. *White*. Routledge.
- D’Amico, D., 2017. An uneasy union: women teachers, organized labor, and the contested ideology of profession during the Progressive Era. *Labor* 14 (3), 35–54. <https://doi.org/10.1215/15476715-3921321>.
- D’Amico, D., Pawlewicz, R.J., Earley, P.M., McGeehan, A.P., 2017. Where are all the Black teachers? Discrimination in the teacher labor market. *Harv. Educ. Rev.* 87 (1), 26–49. <https://doi.org/10.17763/1943-5045-87.1.26>.
- D’Amico Pawlewicz, D., 2020. *Blaming Teachers: Professionalization Policies and the Failure of Reform in American History*. Rutgers University Press.
- Ehrenreich, N., 2002. Subordination and symbiosis: mechanisms of mutual support between subordinating systems. *UMKC Law Rev.* 71 (2), 251–324.
- Endo, R., 2019. Male of color refugee teachers on being un/desirable bodies of difference in education. *Equity Excell. Educ.* 52 (4), 448–464. <https://doi.org/10.1080/10665684.2019.1684220>.
- Evans, S., 1979. *Personal Politics: The Roots of Women’s Liberation in the Civil Rights Movement and the New Left*. Vintage Books.
- Fanon, F., 2005. *The Wretched of the Earth*. R. Philcox (Trans.). Grove Press. Originally published in 1961.
- Federici, S., 2014. *Caliban and the Witch*, second ed. Autonomedia.
- Fultz, M., 2004. The displacement of Black educators Post- *Brown*: an overview and analysis. *Hist. Educ. Q.* 44 (1), 11–45. <https://doi.org/10.1111/j.1748-5959.2004.tb00144.x>.
- Grisson, J.A., Loeb, S., 2017. Assessing principals’ assessments: subjective evaluations of teacher effectiveness in low- and high-stakes environments. *Educ. Finance Policy* 12 (3), 369–395. https://doi.org/10.1162/EDFP_a_00210.
- Harris, C., 1993. Whiteness as property. *Harv. Law Rev.* 106 (8), 1710–1791. <https://doi.org/10.2307/1341787>.
- Hook, D., 2005. Affecting whiteness: racism as technology of affect. *Int. J. Critical Psychol.* 16, 74–99. <https://doi.org/10.1177/0959354311433142>.
- Hooks, B., 1984. *Feminist Theory: From Margin to Center*. South End Press.
- Horsford, S.D., Scott, J., Anderson, G.L., 2019. *The Politics of Education Policy in an Era of Inequality*. Routledge.

- Irons, J., 1998. The shaping of activist recruitment and participation: a study of women in the Mississippi Civil Rights Movement. *Gend. Soc.* 12 (6), 692–709. <https://doi.org/10.1177/089124398012006006>.
- Irvine, J.J., Fraser, J., 1998. "Warm demanders": do national certification standards leave room for the culturally responsive pedagogy of African American teachers? *Educ. Week* 17 (35), 56.
- Jacobs, M., 2009. *White Mother to a Dark Race: Settler Colonialism, Maternalism, and the Removal of Indigenous Children in the American West and Australia, 1880–1940*. University of Nebraska Press.
- Jones-Rogers, S., 2019. *They Were Her Property: White Women as Slave Owners in the American South*. Yale University Press.
- Kaestle, C., 1983. *Pillars of the Republic*. Hill and Wang.
- Kalogridis, D., Loeb, S., Bételle, T., 2013. Systematic sorting: teacher characteristics and class assignments. *Sociol. Educ.* 86 (2), 103–123. <https://doi.org/10.1177/0038040712456555>.
- Kohli, R., 2016. Behind school doors: the impact of hostile racial climates on urban teachers of color. *Urban Educ.* 53 (3), 307–333. <https://doi.org/10.1177/0042085916636653>.
- Ladson-Billings, G., 2009. "Who you callin' nappy-headed? A critical race theory look at the construction of Black women. *Race Ethn. Educ.* 12 (1), 87–99. <https://doi.org/10.1080/13613320802651012>.
- Laville, H., 2017. *Organized White Women and the Challenge of Racial Integration, 1945–1965*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-49694-8>.
- Leonardo, Z., 2005. Through the multicultural glass: althusser, ideology, and race relations in post-civil rights America. *Pol. Futures Educ. Internet* 3 (4), 400–412. <https://doi.org/10.2304/pfie.2005.3.4.400>.
- Leonardo, Z., 2013. The story of schooling: critical race theory and the educational racial contract. *Discourse* 16 (4), 470–488. <https://doi.org/10.1080/01596306.2013.822624>.
- Leonardo, Z., 2015. Contracting race: writing, racism, and education. *Crit. Stud. Educ.* 56 (1), 86–98. <https://doi.org/10.1080/17508487.2015.981197>.
- Leonardo, Z., 2020. *Eduard Said and Education*. Routledge.
- Leonardo, Z., Boas, E., 2013. Other kids' teachers: what children of color learn from white women and what this says about race, whiteness, and gender. In: Lynn, M., Dixon, A. (Eds.), *Handbook of Critical Race Theory and Education*, first ed. Routledge, pp. 313–324.
- Leonardo, Z., Boas, E., 2022. Other kids' teachers: what children of color learn from white women and what this says about race, whiteness, and gender. In: Lynn, M., Dixon, A. (Eds.), *Handbook of Critical Race Theory and Education*, second ed. Routledge, pp. 153–165.
- Leonardo, Z., Gamez-Djokic, 2019. Sometimes leaving means staying: race and white teachers' emotional investments. *Teachers College Yearbook* 121 (13), 1–22. <https://www.tcrecord.org/PrintContent.asp?ContentID=22986>.
- Leonardo, Z., Porter, R.K., 2010. Pedagogy of fear: toward a Fanonian theory of "safety" in race dialogue. *Race Ethn. Educ.* 13 (2), 139–157. <https://doi.org/10.1080/13613324.2010.482898>.
- Leonardo, Z., Zembylas, M., 2013. Whiteness as technology of affect: implications for educational praxis. *Equity Excell. Educ.* 46 (1), 150–165. <https://doi.org/10.1080/13613324.2010.482898>.
- Levine-Rasky, 2016. *Whiteness Fractured*. Routledge.
- Little, K., 2009. *You Must Be From the North: Southern White Women in the Memphis Civil Rights Movement*. University of Mississippi Press.
- Lopez, I.H., 2006. *White by Law*. New York University Press.
- Lugones, M., 1994. Purity, impurity, and separation. *Signs* 19 (2), 458–479.
- Matias, C.E., 2016. *Feeling White*. SensePublishers.
- Mensah, F.M., 2019. Finding voice and passion: critical race theory methodology in science teacher education. *Am. Educ. Res. J.* 56 (4), 1412–1456. <https://doi.org/10.3102/0002831218818093>.
- Miller, E., Lensmire, T., 2021. Mammies, brute Negroes, and white femininity in teacher education. *Curric. Inq.* <https://doi.org/10.1080/03626784.2020.1860643>.
- Milner, H.R., Howard, T.C., 2013. Counter-narrative as method: race, policy and research for teacher education. *Race Ethn. Educ.* 16 (4), 536–561. <https://doi.org/10.1080/13613324.2013.817772>.
- Mohanram, R., 2007. *Imperial White: Race, Diaspora, and the British Empire*. University of Minnesota Press.
- Newman, L.M., 1999. *White Women's Rights: The Racial Origins of Feminism in the United States*. Oxford University Press.
- Nir, S.M., 2020. How 2 Lives Collided in Central Park, Rattling the Nation. *The New York Times*. <https://www.nytimes.com/2020/06/14/nyregion/central-park-amy-cooper-christian-racism.html>.
- Noddings, N., 1984. *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press.
- Olson, L., 2001. *Freedom's Daughters: The Unsung Heroines of the Civil Rights Movement from 1830 to 1970*. Simon & Shuster.
- Painter, N.I., 2010. *The History of White People*. W. W. Norton & Company.
- Pateman, C., 1988. *The Sexual Contract*. Stanford University Press.
- Pateman, C., Mills, C., 2007. *Contract and Domination*. Polity Press.
- Petchauer, E., Bowe, A.G., Wilson, J., 2018. Winter is coming: forecasting the impact of edTPA on Black teachers and teachers of color. *Urban Rev.* 50 (2), 323–343. <https://doi.org/10.1007/s11256-018-0453-1>.
- Philip, T.M., Brown, A.L., 2020. We All Want More Teachers of Color, Right?: Concerns About the Emergent Consensus. National Education Policy Center.
- Roediger, D., 1991. *The Wages of Whiteness*. Verso.
- Rogers-Ard, R., Knaus, C.B., Epstein, K.K., Mayfield, K., 2012. Racial diversity sounds nice; systems transformation? Not so much: developing urban teachers of color. *Urban Educ.* 48 (3), 451–479. <https://doi.org/10.1177/0042085912454441>.
- Schultz, D.L., 2001. *Going South: Jewish Women in the Civil Rights Movement*. New York University Press.
- Scott, J.T., 2011. Market-driven education reform and the racial politics of advocacy. *Peabody J. Educ.* 86 (5), 580–599. <https://doi.org/10.1080/0161956X.2011.616445>.
- Singh, M.V., 2018. Role models without guarantees: corrective representations and the cultural politics of a Latino male teacher in the borderlands. *Race Ethn. Educ.* 21 (3), 288–305. <https://doi.org/10.1080/13613324.2017.1395330>.
- Sleeter, C.E., 2017. Critical race theory and the whiteness of teacher education. *Urban Educ.* 52 (2), 155–169. <https://doi.org/10.1177/0042085916668957>.
- Steinberg, M.P., Sartain, L., 2020. What explains the race gap in teacher performance ratings? Evidence from Chicago public schools. *Educ. Eval. Pol. Anal.* <https://doi.org/10.3102/0162373720970204>, 016237372097020.
- Stepan, N., 1986. Race and gender: the role of analogy in science. *Hist. Sci. Soc.* 77 (2), 261–277.
- Thandeka, 2000. *Learning to Be White*. Bloomsbury Academic.
- Tillman, L.C., 2004. (Un)intended consequences?: the impact of the Brown v. Board of education decision on the employment status of Black educators. *Educ. Urban Soc.* 36 (3), 280–303. <https://doi.org/10.1177/0013124504264360>.
- Trujillo, T., Scott, J., Rivera, M., 2017. Follow the yellow brick road: teach for America and the making of educational leaders. *Am. J. Educ.* 123 (May), 353–391. <https://doi.org/10.1086/691232>.
- Tyack, D., 1974. *The One Best System: A History of American Urban Education*. Harvard University Press.
- Tyson, T., 2000. Dynamite and "the silent South": a story from the second reconstruction in South Carolina. In: Dailey, J., Gilmore, G.E. (Eds.), *Jumpin' Jim Crow: Southern Politics From Civil War to Civil Rights*. Princeton University Press, pp. 275–297.
- Tyson, T., 2017. *The Blood of Emmett Till*. Simon & Shuster.
- Valenzuela, A., 1999. *Subtractive Schooling: US-Mexican Youth and the Politics of Caring*. State University of New York Press.

- Villenas, S., 2010. Thinking Latina(o) education with and from Chicana/Latina feminist cultural studies: emerging pathways—decolonial possibilities. In: Leonardo, Z. (Ed.), *Handbook of Cultural Politics and Education*. SensePublishers, pp. 451–476.
- Walker, V.S., 2001. African American teaching in the South: 1940–1960. *Am. Educ. Res. J.* 38 (4), 751–779. <https://doi.org/10.3102/00028312038004751>.
- Wells-Barnett, I.B., 1900. *On Lynchings*. Dover Publications Inc.
- White, T.C., 2016. Teach for America's paradoxical diversity initiative: race, policy, and Black teacher displacement in urban schools. *Educ. Pol. Anal. Arch.* 24, 16. <https://doi.org/10.14507/epaa.24.2100>.
- Wood, C., 2019. Oakland Comes Out to Celebrate Bbq'n While Black. *Berkeleyside*. <https://www.berkeleyside.org/2019/07/22/oakland-comes-out-to-celebrate-bbqn-while-black>.

Feminist genealogy & policy futures

Wanda S. Pillow, Gender Studies, School for Cultural & Social Transformation, University of Utah, Salt Lake City, UT, United States

© 2023 Elsevier Ltd. All rights reserved.

A reminder of what policy studies is & can be	104
Feminist genealogy as policy studies approach	105
Policy studies present futures	106
References	107

The coronavirus (COVID-19) pandemic has exposed the fragility of the US economy and the immense power disparities and systemic disadvantages built into our social contract

(The Roosevelt Institute Covid-19 Response, 2020).

These are important first steps to colonizing space and other worlds. Humankind will explore and colonize our solar system and then reach for the stars
(The Gateway Foundation).

Writing and thinking about policy studies during a global pandemic, multiple occurrences of windstorms, earthquakes, fires, flooding, and extreme heat or cold disasters, amid racist attacks and continued murder of Black persons by police, social justice movements countered by extremist tactics, months of attempts by an acting US President to overthrow democratic elections and disrupt a peaceful transfer of power resulting in the January 6, 2021 insurrection at the US Capitol—all occurring in a chorus of loud actions and messages of hate, violence, and conspiracy theories—has been, to put it mildly, overwhelming. Any of these contexts would cause reflection and analysis. When these contexts are occurring simultaneously and intersectionally, ruptures in the infrastructure of U.S. democracy, economy, and access to basic necessities of everyday life (such as clean water, food, lodging) are exposed. These exposures and dichotomies have drawn direct attention to the “systemic disadvantages built into our social contract,” a social contract based on settler colonial logics, logics not limited to Earth.

These conditions have opened up conversations across progressive and mainstream institutions and puts questions about “Who are we? What do we expect from society, ourselves, and our institutions?” to the forefront of policy studies. That is to the forefront of the social values of society, to the roles of research, the policy process, policy actors, and discourse (Ball, 1993, 2015; Rizvi and Lingard, 2010). Initially, this chapter was titled “doing policy, differently.” However, after a pandemic + year of thinking about institutional to departmental level policy needs and possibilities, I suggest it is not always about doing things differently, racing toward the next in theory, ideas, or analytics, but often it is about doing our work deeply—attending to context, process, and using strong theory and analysis. Part of doing work deeply is reminding ourselves what policy studies is, as well as what are policy possibilities, and policy studies responsibilities/response-abilities. I present feminist genealogy as an informative approach to policy studies and spend the body of the paper discussing what feminist genealogy is and can do. This is policy studies written toward remembering and a commitment to decolonial reparations for present futures (Pillow, 2017a; Webb and Gulson, 2015).

A reminder of what policy studies is & can be

Policy studies encompasses analysis of the process of policymaking, that is, the policy process itself, as well as analysis of the contents and outcomes of policy. Social and education policy reflect social values, and as a field policy studies is influenced by a range of paradigms and methods. Across any approach, policy studies has theoretical, methodological, and practice responsibilities. From a critical, feminist, poststructurally influenced lens (Pillow, 2015, 2017b, 2019), some key points of responsibility include:

- policy is about social regulation; thus, we can examine who is being regulated, how, and why, and to what effects/affects
- policy can be a source of social change and we can examine change for the values of equity, equality, social justice, reparative transformation
- policy analysis can address social injustice and policy debts (Ladson-Billings, 2006; Pillow, 2017b)
- policy studies can be dedicated to a “widened range of policy possibilities” (Roosevelt Institute Covid-19 The Roosevelt Institute Covid-19 Response, 2020).

This is a way of thinking about policy studies wholistically—considering policy philosophically as well as applied and asking how definitions of policy problems and choice of methodologies impact the shape of policy discourse and enactments (Ball, 1994; Gale, 2001, 2003; Pillow, 2004). As a field, policy studies has a responsibility to know its own history. Asking for example: How have policies supported division and invoked harm and how has a policy studies theoretical and methodological arena reproduced inequities and limited policy possibilities. I argue that with the depth of excellent policy studies work (e.g., Ball, 1993; Gillborn and

Youde, 2000; Hawkesworth, 1994; Myers, 2002; Myers and Ha, 2019; Rizvi and Lingard, 2010; Stone, 1998), policy scholars and policy actors have a “response-ability.” That is we have the resources to be able to do policy differently. Thinking about policy studies as a field with theoretical and ethical responsibilities as well as a field with the ability to apply knowledge activates policy studies.

Toward this goal, the London School of Economics and Political Science (LSE) description of social policy is very useful (Platt, 2021):

Social policy is concerned with the ways societies across the world meet human needs for security, education, work, health and wellbeing. Social policy addresses how states and societies respond to global challenges of social, demographic and economic change, and of poverty, migration and globalisation. Social policy analyses the different roles of: national governments, the family, civil society, the market, and international organisations in providing services and support across the life course from childhood to old age. These services and support include child and family support, schooling and education, housing and neighbourhood renewal, income maintenance and poverty reduction, unemployment support and training, pensions, health and social care. Social policy aims to identify and find ways of reducing inequalities in access to services and support between social groups defined by socio-economic status, race, ethnicity, migration status, gender, sexual orientation, disability and age, and between countries. At LSE, social policy is explicitly *international, interdisciplinary and applied*.

Thinking with and enacting the LSE definition requires response-ability. The past three decades of policy studies scholarship have demonstrated that policy studies is compelling and provocative. Close investigations of epistemologies, ontologies, methodologies have provoked examination of how we ask questions, what approaches are engaged, and how we might differently engage knowing/seeing/witnessing (Pillow, 2019).

The above LSE definition summarizes where policy studies is. Policy studies is theorized as temporal—always addressing past, present, futures (Pillow, 2015). Policy studies is on the ground, embodied, lived experiences, and materialities. Policy is action, text, discourse; interactive and intra-active. Policy is simultaneously theoretical and applied. Policy studies methodology of the policy process itself and/or a focus on policy outcomes offers built in spaces for critical reflection and reflexivity. Given policy is a “reflection of social values” (Rizvi and Lingard, 2010), understanding and defining policy studies in these ways is crucial to developing policy studies as a field that demonstrates a response-ability to invoke social values.

Feminist genealogy as policy studies approach

Feminist genealogy arose from my thinking with poststructural feminisms, Women of Color (WOC) feminisms, and Michel Foucault. This approach was developed as a necessity. In the 1990s, education policy studies was closely indebted to quantitative supremacy and a “rational decision making” model of reasoning (Stone, 1998). Challenges from critical, feminist, poststructural thinking were just beginning to be discussed at conferences and published in essays and volumes (Ball, 1994; Marshall, 1997; St. Pierre and Pillow, 2000; Troyna, 1994). It was a decade-plus of heated debated, often referred to paradigm wars (Denzin and Lincoln, 2008). However scholars often presented in siloed spaces presenting to and talking with those mostly in paradigmatic agreement. Certainly spaces to think with like-minded critically engaged colleagues are necessary to produce excellent scholarship (and these spaces are crucial to BIPOC, queer, under-represented persons and scholarship), but at conferences I found myself running between postie sessions, policy sessions, and sessions deeply engaged with race or intersectional WOC theory. The postie sessions mostly ignored race and systemic colonialism and treated questions of either as simply identity critiques (Pillow, 2019); mainstream policy sessions avoided feminist, race, and poststructural thinking; and the race/queer/woc spaces provided provoking analysis and discussion that fed me personally and professionally.

Academia’s disciplinary structure—a structure that desires purity logics (Lugones, 1994)—discounts interdisciplinary scholar’s splintered experiences in courses, conferences, and publishing. Certainly this is changing with programs like LSE and the growth of interdisciplinary policy studies journals and publication outlets. Yet, policy studies as a field, and policy in development and implementation, is immersed in colonial logics leading to inherent systemic biases, including a centering of assumptive heterosexuality, normative gender binary, and whiteness. Given this basis, there is a responsibility to acknowledge that policy studies needs the theoretical insights and shifts, the analytic skills of WOC, race, and queer theory in order to produce policy studies that is not mired in its own white-out ontology (Pillow, 2019). Feminist genealogy is one way to approach this responsibility.

My first moves toward feminist genealogy were deeply influenced by feminist theories of the body and embodiment (Pillow, 1997). How does policy account for, think about the bodies it seeks to define, regulate, change, control, or access? How do these bodies experience policy? This two-pronged interest drove my curiosity about both the “inscriptive” and “lived body” (Grosz, 1995) and what would it require for policy studies to consider how it thinks about inscribed bodies as well as what it would look like to analyze policy from/with corporeality of lived bodies. My key interest was: “How does paying attention to bodies change what we look at, how we look, what we ask, and what we choose to represent?” (Pillow, 1997, p. 349).

This question and feminist genealogy’s theorizations and methodological explorations are indebted to the young mother’s I spent time with and learned from. Focusing on pregnant/mothering students access to K-12 education as an education and social/legal policy issue presented many challenges and ironies including assumptions amid absences and sexualizations amid

discomfort and silences, as well as deeply racialized good girl/bad girl discursive constructions reinforced in research, policy, practice, and social attitude. Turning to the body, I argued

... can work to make explicit what is both overexposed and obscured elsewhere. Both of the exposures presented point to the inability of our theories, practices, and programs to deal with certain bodies, in this case, the pregnant teen body.

Pillow (1997, p. 360).

My close observation of pregnant/mothering bodies in schools, demonstrated how bodies are defined, assumed, and treated and these lived experiences necessitated further understanding of how and why certain bodies are treated differently. It became clear that to understand what was occurring to deflect and limit pregnant/mothering students access to public education in the 1990s–2000s, required an understanding and analysis of socio-historical racialized constructions of “unwed mothers” and how these constructions prevail in present and future ideologies of who the pregnant/mother student is and thus, how she should be treated and what she deserves (or does not deserve). That is, how certain bodies are thought is linked to past constructions, which impact present and future possibilities.

Increasingly, I saw history as key to any policy study of the present, and by 2003, I introduced “feminist genealogy” as a policy studies methodology. Feminist genealogy incorporates qualitative and cultural history methodologies with close attention to bodies and at the time what I termed “race feminisms.”

Feminist genealogy builds from the work of Foucault read through and with race-feminisms and focuses specific attention upon the discursively structured raced, gendered and sexed body. Feminist genealogy embodies policy analysis, forefronting and refiguring conceptualizations of bodies, power and knowledge in policy studies.

Pillow (2003, p. 146).

While I introduced feminist genealogy as methodology, it was and is also a theoretical approach. Feminist genealogy offers ways of “doing” policy research and analysis—close meticulous observation of how power enacts on bodies combined with socio-historical analysis of how and why regulatory, discursive power defines bodies in certain ways—and simultaneously feminist genealogy is a way to interrogate policy studies itself; it offers a way to observe and critique for example assumptions involved in determining what is a policy problem or how a policy issue is defined or how to interpret policy data.

In these ways, feminist genealogy is a constant check and challenge on the policy studies scholar, requiring us to continually question our own assumptions. Feminist genealogy thus “reinterpret(s) power, truth, ontology and subjectivity”; offers an “historical, discursive critique” of “what appears natural and necessary”; while locating “the work of genealogy as critique within the body, a methodology that embodies policy analysis” (Pillow, 2003, p. 147).

Feminist genealogy acknowledges the embedded relationality between theory, research, and practice and is a way to undo policy studies while doing policy studies. As noted above, my turn to and development of feminist genealogy was due to lacks and gaps in policy studies; policy studies theories and approaches were actively working against what I was witnessing in public schools. While I continued to work on policy and legal change for pregnant/mothering students to access education in the US, feminist genealogy also allowed me to shift my lens of analysis from young mothers as the problem, to the arena of policy studies itself. Policy studies had a problem—theoretically and methodologically—and alongside other critical, race-feminist, poststructural work, I sought to actively shift how we think policy studies:

An embodiment exposed through feminist genealogy alters what is meant by policy studies. It opens spaces for questioning the construction of the policy process, now policy studies becomes in essence the subject of analysis and challenges analytic methodological practices. Embodied analysis provides a way to speak up and speak back to the paradoxes in existing policies and expose the ineffectiveness of reliance upon assumptive, seemingly natural.

Pillow (2003, p. 156).

Policy studies present futures

As an arena of study, policy studies has a responsibility to investigate itself as a field of knowledge, research, and practice and to expand how, who, and what is theorized and how and what research is authorized. I suggest that feminist genealogy is a useful approach toward these needs. Partly what feminist genealogy can do is identify how policy studies determines and enacts what matters—what bodies matter, what issues matter, what knowledge matters, what research matters—across decades.

As Trouillot (1995) asks: “If some events cannot be accepted even as they occur, how can they be assessed later? Can historical narratives convey plots that are unthinkable in the world within these narratives take place? How does one write a history of the impossible?” (p. 73). In this I hear a challenge for feminist genealogy, a challenge to both identify and see what is hegemonically unthinkable and to identify the processes, the plots, that reproduce the unthinkable. This requires as Freeman (2010) argues an affective temporal drag of data; “a way of forcing the present to touch its own disavowed past or seemingly outlandish possible

future" (p. 78). Data as "temporal drag" offsets past-present-future absences, while also creating a drag effect and affect on how data is thought, read, and discovered in observation, in data sets, or in the archive. I raise these points to highlight that feminist genealogy, like any approach, will reproduce what and how it is thought. Feminist genealogy is only as "good," is only as useful at analyses of power, as the theorizations and provocations it is thought with.

In early work, alongside feminist poststructural scholars, I often called for thinking otherwise. However, looking back I see how these "otherwise" calls could and often did simply reproduce existing standards and structures of power. Thinking otherwise assumed a "homogenous empty time" (Ferguson, 2007), assumed that there was not knowledge and theory available to undo the identified power arrangements (the colonial logics) when in fact there was a wealth of theorizing available to "think otherwise"—it was just that this work was not part of fem-postie theorizing and not authorized in those spaces. This is a reflective looking back, which is part of our work as scholars, and in this reflexivity I am not prescribing thinking with only certain critical theories. I am identifying a gap—an ontological white-out—in that heady insightful work produced in policy studies in 1990–2010, a gap that continues but does not need to be repeated.

As theorizing continues to be on the move, policy studies would benefit from increased interdisciplinary thinking—whether thinking with critical theories in humanities, physical sciences, earth sciences, geography, etc. In order for policy studies to theorize and analyze what decolonial otherwise spaces might be, the field will continuously need to address its ontological white-outs and for that policy scholars need the depth of critique and knowledge available in Women of Color, decolonial indigenous feminism, interdisciplinary Race theory, and Queer of Color critique (Magubane, 2016).

A policy studies committed to thinking otherwise requires specific deep engagement with postcolonial studies, decolonial feminisms, and indigenous critical theory and any and all theory that helps see ontological white-outs. The intention is to have "thinking otherwise" provoked into thinking approaches that have theoretical impetus while directly addressing materialities of social injustice. Here for example, thinking otherwise becomes reparational transformation, decolonial praxis, decoloniality (Byrd, 2016; Dharmoon, 2015; Figueroa, 2015). Policy studies—across theory, research, and in practice—has a primary role in addressing societal needs, including what decoloniality does, and what can decoloniality look like in policy studies? What are elements of decolonial policy in theory, research, and practice? What are processes of decolonial policy? Feminist genealogy with its close analysis of the present as indicative of the past while articulating potential futures is a useful way for policy studies to engage with decoloniality and decolonial futures. This work is not only a global issue but a cosmological, galactic issue as governments, entrepreneurial groups, and technologies readily deploy ontologies of "space colonization" relying on nationalistic colonial discourses and practices to dominate space including unknown worlds (<https://orbitalassembly.com/>; https://www.nasa.gov/centers/hq/library/find/bibliographies/space_colonization).

Whether and how feminist genealogy is thought and utilized is up to present and future scholars, but analyses of power is at its core and feminist genealogy is not about stagnant replications of critique. In addition, the impetus of feminist genealogy to necessarily think with feminist, race, queer+ studies because these lenses are needed in order to see and interrogate colonial logics is also at the core of feminist genealogy and I suggest necessary to policy studies. The ongoing challenges of recognizing and re-theorizing colonial logics and colonial white-outs is the responsibility and the response-ability of policy studies; a response-ability to past, present, and futures.

References

- Ball, S.J., 1993. What is policy? Texts, trajectories and toolboxes. *Discourse* 13 (2), 10–17.
- Ball, S.J., 1994. *Education Reform: A Critical and Post-structural Approach*. Open University Press, Buckingham.
- Ball, S.J., 2015. What is policy? 21 years later: reflections on the possibilities of policy research. *Discourse* 36 (3), 306–313.
- Byrd, J.A., 2016. Still waiting for the "post" to arrive: Elizabeth Cook-Lynn and the imponderables of American Indian Postcoloniality. *Wicazo Sa Rev.* 31 (1), 75–89.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2008. *Collecting and Interpreting Qualitative Materials*. SAGE, Thousand Oaks, CA.
- Dharmoon, R., 2015. A feminist approach to decolonizing anti-racism: rethinking transnationalism, intersectionality, and settler colonialism. *Feral Femin.* 4, 20–37.
- Ferguson, R.A., 2007. The relevance of race for the study of sexuality. In: Haggerty, G.E., McGarry, M. (Eds.), *A Companion to Lesbian, Gay, Bisexual, Transgender, and Queer Studies*. University of California-Irvine: Blackwell Press, pp. 109–123.
- Figueroa, Y.C., 2015. Reparation as transformation: radical literary (re)imaginings of futurities through decolonial love. *Decolonization* 4 (1), 41–58.
- Freeman, E., 2010. *Time Binds: Queer Temporalities, Queer Histories*. Duke University Press, Durham.
- Gale, T., 2001. Critical policy sociology: historiography, archaeology and genealogy as methods of policy analysis. *J. Educ. Pol.* 16 (5), 379–393.
- Gale, T., 2003. Realising policy: the who and how of policy production. *Discourse* 24 (1), 51–65.
- Gillborn, D., Youdell, D., 2000. *Rationing Education: Policy, Practice, Reform and Equity*. Open University Press, London.
- Grosz, E., 1995. *Space, Time, and Perversion*. Routledge, London.
- Hawkesworth, M., 1994. Policy studies within a feminist frame. *Pol. Sci.* 27, 97–118.
- Ladson-Billings, G., 2006. From the achievement gap to the education gap: understanding achievement in U.S. schools. *Educ. Res.* 35 (7), 3–12.
- Lugones, M., 1994. Purity, impurity, and separation. *Signs* 19 (2), 458–479.
- Magubane, Z., 2016. America's sociology racial ontology: remembering slavery, deconstructing modernity, and charting the future of global historical sociology. *Cult. Sociol.* 10 (3), 369–384.
- Marshall, C. (Ed.), 1997. *Feminist Critical Policy Analysis I: A Primary and Secondary Schooling Perspective*. Falmer Press, London.
- Myers, S.L., Ha, I., 2019. *Race Neutrality: Rationalizing Remedies to Racial Inequality*. Lexington Books.
- Myers, S., 2002. Analysis of race as policy analysis. *J. Pol. Anal. Manag.* 21 (2), 169–190.
- Pillow, W.S., 1997. Exposed methodology: the body as a site of deconstruction. *Int. J. Qual. Stud. Educ.* 10 (3), 349–363.
- Pillow, W.S., 2003. "Bodies are dangerous": using feminist genealogy as policy studies methodology. *J. Educ. Pol.* 18 (2), 145–160.
- Pillow, W.S., 2004. *Unfit Subjects: Educational Policy and the Teen Mother*. Routledge Press, New York.

- Pillow, W.S., 2015. Policy temporality & marked bodies: feminist praxis amongst the ruins. *Crit. Stud. Educ.* 56 (1), 55–70.
- Pillow, W.S., 2017a. Imagining (data) differently. In: Gale, T., Gulson, K. (Eds.), *Education Policy & Social Inequality*. Springer Press, London, pp. 133–151.
- Pillow, W.S., 2017b. Policy studies debt: a feminist call to expand policy studies theory. In: Diem, S., Young, M. (Eds.), *Critical Approaches to Education Policy Analysis*. Springer Press, New York, pp. 261–274.
- Pillow, W.S., 2019. Epistemic witnessing: theoretical responsibilities, decolonial attitude and lenticular futures. *Int. J. Qual. Stud. Educ.* 32 (2), 118–135.
- Platt, L., 2021. What is Social Policy? International, Interdisciplinary and Applied. <https://www.lse.ac.uk/social-policy/about-us/What-is-social-policy>.
- Rizvi, F., Lingard, B., 2010. *Globalising Education Policy*. Routledge, New York.
- St Pierre, E., Pillow, W.S. (Eds.), 2000. *Working the Ruins: Feminist Poststructural Theory and Methods in Education*. Routledge, New York.
- Stone, D.A., 1998. *Policy Paradox: Political Reason*. Scott, Foresman & Co, Boston, MA.
- The Gateway Foundation, <https://gateway.spaceport.com/>.
- The Roosevelt Institute Covid-19 Response, April 15, 2020, <https://rooseveltinstitute.org/roosevelt-institute-covid-19-response/>.
- Trouillot, M.F., 1995. *Silencing the Past: Power and the Production of History*. Beacon Press, Boston.
- Troya, B., 1994. Critical social research and educational policy. *Br. J. Educ. Stud.* 42 (2), 70–84.
- Webb, P.T., Gulson, K.N., 2015. Policy scientificity 3.0: theory and policy analysis in-and-for this world and other-worlds. *Crit. Stud. Educ.* 56 (1), 161–174.

Indigeneity

Erica Neeganagwedgin, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	109
Positionality on indigeneity	109
Indigenous land knowledge and connections	110
Indigenous education systems	112
Indigenous philosophies of reciprocity	112
Conclusion	113
References	114

Introduction

Indigeneity, according to [Steeves \(2018\)](#), is not as simple as an opposition to identity erasure or resistance against colonization. Rather, Indigeneity is composed of many histories and experiences. This chapter does not attempt to define Indigeneity, but to discuss and explain Indigeneity in the contexts of relationships. As [Baird \(2016, p. 501\)](#) explains, Indigeneity does not mean the same to everyone. Rather, the particular social, cultural, political, and historical contexts of different places influence the ways in which “indigeneity is being conceptualized and practiced.” For example, [Grydehøj et al. \(2018, p. 17\)](#) speak to the notion of “Indigenous” and the challenges in avoiding labels and categories imposed by colonization. Therefore, the term Indigenous would not be used by everyone. Similarly, the idea of Indigeneity is going to be understood in different ways and contexts by some people.

The United Nation states that Indigenous peoples are inheritors and practitioners of unique cultures and ways of relating to people and to their environments. Despite colonial devastations, Indigenous people continue to retain social, cultural, economic, and political characteristics that are distinct ([United Nations, 2009](#)). Indigenous Peoples, while unique, share collective ancestral ties to their Lands and natural resources from which they often are, or have been, displaced. Their lands and resources are linked to their identities, cultures and livelihoods, as well as their physical and spiritual well-being ([World Bank, 2021](#)). Indigenous Peoples are currently 6.2% of the world’s population, and they come from all regions of the globe ([United Nations Development Programme, 2021](#)). [Steeves \(2018, para. 1\)](#), points out, “diverse Indigenous communities weave Indigeneity through a multifaceted array of space and time to revive identities and cultural practices and to regain or retain land, human rights, heritage, and political standing.” The Declaration on the Rights of Indigenous Peoples (2007) affirms, and reaffirms, the unique status, distinctiveness, rights, and protection of Indigenous peoples globally. Regardless of who Indigenous People are and where Indigenous People live, there are common bonds and understanding in the Indigenous worlds that unite them.

[Radcliffe \(2017, p. 221\)](#), describes Indigeneity as a cross-disciplinary concept. The author adds that “indigeneity refers at its broadest to the quality of being Indigenous. Indigeneity can be defined as the socio-spatial processes and practices whereby Indigenous People and places are determined as distinct (ontologically, epistemologically, culturally, in sovereignty, etc.) to dominant universals.” [Matute \(2021, p. 257\)](#) examines what it means to be Indigenous in a globalized world and looks at the power relations. Furthermore, the author contends that Indigenous peoples “display... a multiplicity of indigeneities, considering different interpretations and experiences. ...” [Jaffee and John \(2018, p. 1412\)](#) remind us that the term Indigenous was used by Indigenous peoples in their quest to be recognized within the United Nation. They assert that, although there are numerous Indigenous People who have different languages, cultures and come from many different places, Indigenous People “stress a connection to ancestral Lands and a non-Western ontology which frames their kinship languages” and the ways in which their societies are organized and structured. [Dei and Jaimungal \(2018, as cited in Dei, 2016, p. 61\)](#) reports there is general agreement that there are “contestations and contentions of Indigenous and Indigeneity”, and that “there is neither a singular category which encompasses Indigenous, nor Indigeneity.” Indigeneity is therefore “constantly being redefined and negotiated,” based on the power relations at work in particular moments and time ([Postero, 2013, p. 112](#)). This chapter draws on the work of Indigenous scholars and discusses Indigeneity in the context of relationships to Land, Connection to Indigenous ways of knowing through Education, and the Indigenous world-views of reciprocity.

Positionality on indigeneity

As a rural Island person from the Caribbean, and someone who identifies as Taino, I have always understood my Indigeneity even if it had been unspoken. Macintosh suggests that the enduring existence of Indigenous people in the Caribbean into the new millennium continues to challenge popular writing and standard scholarship ([McIntosh, 1999, para 3](#)). For example, generations of

people for the past five hundred or more years grew up hearing and believing that Taino people ceased to exist after first coming in contact with Europeans. This became normalized in school texts for hundreds of years, and Taino people remain invisible today. Yet many Indigenous people in the Caribbean continue to express their Indigeneity in their own ways. Grydehøj et al. (2018, p. 14) state, Island peoples “are re-engaging and reworking their self-determination, economies, identities and relations to their territories and to mainland and global society” There is a resurgence and a reassertion of Indigeneity in the Caribbean. This is just one example of that resurgence and what I describe as “singing it out loud.” Yet, as noted, there is no one single universal form of Indigeneity. Based on my lived experiences and knowledges in a Caribbean, and specifically Yamaye, context, I understand what it means to be Indigenous.

Raddcliffe (2018, p. 443) points to the impact of colonialism, and how power dynamics are encountered in the everyday lives of Indigenous peoples, and the ways in which Indigeneity is expressed socially and intersectionality ... My lived experiences inform my Indigeneity which was given to me by my ancestors. I think about the plants that provide nourishment, including Rosemary, leaf of life, countless tropical fruits, and my relationship to my ancestral lands. The land was constantly giving to us, to our family, and we in turn love the land. Growing up in a rural part of the Caribbean Island, Land was of utmost significance, and it remains that way today. The many plants, shrubs, and food vegetation that we get from the Land is at the foundation of Indigeneity, as is the red dirt, the soil that is part of the Land where I come from. Kimmerer (2018, p. 28) explains that “plants are also recognized as our oldest teachers” and adds that plants have been around for a long time. “They know how to make food and medicine out of light and air and then they give it away. They unite the earth and the sky and exemplify the virtues of generosity, providing us with all that we need to live.”

Bishop (2001, para 1) shares the following story which also grounds my understanding of an Indigenous world. She writes about listening to the story of an Elder from the Kalahari Desert in Botswana. The Elder shared stories “about the waving savannah grasses and the endless sky of his homeland and told me that he could never leave because the land knew him”. This resonates with me because not only does the land knows me, but the land remembers me.

The critical importance of Land is intrinsically understood in my family. The Land feeds us spiritually—so much so that even when away from the Land selling it is not an option. Rather, it is imperative to have someone care for and tend to it. I have lived the experience of learning from the Land and having that connection and relationship with it. For example, corn, casava, sweet potato and beans are only a few of the food systems in the context of subsistence living that reinforce my connection to my ancestral Lands. Our Taino ancestors had food sovereignty which they derived from their Lands. We continue to eat the food that our ancestors ate. Most of these foods, given to us by the Land, remain central sources of nourishment within the contemporary context despite the more than 500 years of colonial and imperial domination of our Lands and People. In the Islands, the food that comes from the earth sources provides for the community, which is a form of recognition that the Land gives to the people. Assing (2016, para 3) explains Caribbean peoples’ own Indigenous pre-Columbian ecology, including the forests and rivers, were teeming with more life than we know to exist today. Even the sea that separates these islands was full of food. Further, Bishop (2001, Para 2) notes how the Kalahari Elder’s statement about his ancestral Lands is, political, emotional, and philosophical. It is at the heart of Indigenous Peoples’ understanding of their world. Yet, Bishop (2001, Para 2), writes “Unfortunately, the depth of this belief for traditional Indigenous people has rarely been fully appreciated”.

To examine “what is Indigeneity?” Matute (2021, p. 261) writes that it is necessary to listen to the narratives, histories, perspectives, and horizons from those who live it. The next section of this chapter looks more indepth at Land and Indigenous principle of relationships to get a meaningful understanding of the significance of Land and its connection to Indigeneity.

Indigenous land knowledge and connections

As previously emphasized, Indigenous people throughout the world have their own worldviews and cultural practices. Nevertheless, many Indigenous people have been impacted by Land loss, dispossession, theft, and displacement. However, through Indigenous peoples’ interconnection and interrelatedness with their Lands, they have found ways to challenge Western thoughts and practices (Jaffee and John, 2018, p. 1410).

Raddcliffe (2017) observes that Land is also reflected in the understanding of Indigeneity and Indigenous Peoples’ spiritual and intimate relationship with their Lands. The United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP, 2007) Article 25 affirms Indigenous peoples’ right to

maintain and strengthen their distinctive spiritual relationship with their traditionally owned or otherwise occupied and used lands, territories, waters and coastal seas and other resources and to uphold their responsibilities to future generations in this regard.

The United Nations Department of Economic and Social Affairs (2009, p. 52) report on statement of the world’s Indigenous people reinforces the commonality in worldviews of Land to Indigenous peoples globally. It stated that *Land is the foundation of the lives and cultures of Indigenous peoples all over the world.*

Another example that demonstrates the wholistic relationship of Indigenous people with the Land is the work of [Restoule et al. \(2013\)](#) which discusses a research project that looks at “honoring Mushkegowuk Cree view of Land, Environment and life in Fort Albany First Nation in Canada (p. 68).” They note that, “the term Paquataskamik is about relationship to the Lands which comes from Indigenous laws and governance, and is grounded in Indigenous Land Knowledges that have always existed (pp. 78–79).”

Significantly, [Dei and Jaimungal \(2018, p. 61\)](#) speaks about Indigeneity,

As a process of identity formation, Indigeneity is also about politics and resistance. Indigeneity is developing an Indigenousness or an Indigenous consciousness of who one is and the importance of using such Indigenous identity to resist colonial, and on-going, colonial impositions.

Dei adds, “Indigeneity has a close relationship to Land. It is about the politics of the Land that gives a people their identities.” Further showing that Land is the center for the lives, cultures, and identities of Indigenous Peoples. The United Nations Department of Economic and Social Affairs (2009) asserts that Indigenous peoples globally express and enact their inherent rights over their Lands, their territories, and natural resources to governments. Due to imperialism and colonialism, both historically and today, “Indigenous peoples have been dispossessed of their Lands or faced the threat of dispossession and forced removal, leading to increased poverty, erosion of cultures and even outright extinction or complete assimilation” (p. 54). For many Caribbean Indigenous peoples, the impact of colonialism since 1492 has been devastating. However, Taino People, and Indigenous People globally, continue to rebuild their governance systems. We are reminded that Land is where the ancestors are buried and where sacred places are visited and revered ([United Nation, 2009, p. 53](#)).

[Dei and Jaimungal \(2018, p. 18\)](#) speak to the ways in which Land is lived and felt and of its connection to Indigeneity.

The Land is also intricately tied to Indigeneity and the saliency of Indigeneity and its connection to the Land is crucial to highlight. Land is taken both concretely and metaphorically, allowing bodies to implicate space and learning. Land evokes more than a physical presence. It is a spiritual place and a spiritually centred understanding of social space.

This draws me back to reflect more on the story shared earlier by the Kalahari Elder and the intimate relationship with, and emotional attachment to, Land. Additional examples, according to [Koot et al. \(2019, p. 342\)](#), show that belonging, as it relates to Land, is of key importance to “Indigenous’ people of southern Africa, who articulate their Indigeneity in many ways” They note that “perhaps none more potently than through their struggle to demonstrate their unique connections to and interdependence with the land within these neoliberal contexts.”

Indigenous Peoples are close to the land, and many still depend on it for their livelihood ([Cajete, 2020, p. 3](#)). Throughout the world, Indigenous people continue to exercise their rights to their Lands. For example, [Koot et al. \(2019, p. 342\)](#) point to the Hai//om San who, they explain, were evicted from Etosha National Park. [Jaffee and John \(2018, p. 1407\)](#) discuss how “Land appropriation, resource extraction, linguistic genocide, forced removal, erasure, and devastation by settlers invariably wreaks havoc on the land, spirit, livestock, and bodies of Indigenous peoples.”

Yet, another example of the ways in which Indigenous people are violently displaced occurred with the removal of Batwa families from the Kahuzi-Biega Forest in the Congo. There were promises of compensation for the families, but those promises were not honored. The Batwa peoples were traumatized and were not allowed to even hunt in the park. The literature shows that they now lack food, medicine plants, “and the forest is no longer their place of worship”. As the [Report of the African Commission \(2005, p. 22\)](#) explain, “The Batwa have been culturally and psychologically shattered by the loss of their forests” ([Report of the African Commission, 2005, p. 22](#)). The Land and its forest, the ancestral home of the Batwa is critically linked to their identity and Indigeneity as Indigenous people. These examples demonstrate the dangerous and capitalistic characteristics of colonialism and the capitalistic obsession for Indigenous lands. It works similarly globally. [Barker and Battell Lowman \(2015, para 3\)](#) explain, “Settler colonial societies around the globe tend to rely on remarkably similar spatial constructs, power structures, and social narratives.” They state, “Beginning with terra nullius—the perception that Lands in long-term use by Indigenous peoples are empty or unused—settler colonization proceeds to carve up Indigenous-held lands into discrete packets of private property.”

Meanwhile a [Final Report \(2021\)](#) for the Athabasca Chipewyan First Nation provides detailed accounts of the history of the Wood Buffalo National Park in Alberta, Canada, and the relations with the Denésuliné. The report details the systemic removal of the Denésuliné from their ancestral lands. According to the report, once evicted, many people’s homes were burned down; they lost cabins and belongings” (p. 74). Like Indigenous People elsewhere, the Denésuliné’s relationship with their lands is foundational to intergenerational transmission of their knowledge systems and their lives (p. 25). Even when Indigenous People have been displaced from their Lands or have been away from it, many still feel the connection in their heart. As an example, [Warren \(2016, p. 697\)](#) speaks to the cultural and spiritual significance of the Land to the Mapuche people. The author asserts, although they will move to cities, Mapuche youth “see their relationship with nature in the countryside as what differentiates them from non Mapuche Chileans”. It is through the Land that teaching, and learning take place in many Indigenous worlds. Land and learning epitomize and personify Indigenous Education systems and their Indigeneity ([Neeaganagwedgin, 2021](#)).

Indigenous education systems

Indigenous Lands and education systems are interrelated and attached. Land is embedded in relationships with the people, plants, and other living beings. Indigenous knowledge and Education come from the Lands. According to Bishop (2020, p. 6), knowledge exchange has been happening for generations and in doing so it has ensured a continuous sustenance and abundance for people. Bishop (2020, p. 8) writes “There is no universal version of Indigenous education sovereignty, nor should there be.” Bishop states that this is because, Indigenous Peoples have many forms of education which are connected to their Indigeneity. Most importantly, for Cajete (2005, p. 71), the role and importance of the extended family, the clan and the community provide context and serve as the source of teaching. Bishop (2020, p. 7) asserts that Indigenous Education involves “intergenerational relationality; knowledge sharing through Indigenous processes.” Meanwhile, Alfred and Cornstassel (2005, p. 598) discuss how settlers continue to follow the colonial legacy of their ancestors to erase Indigenous Peoples, histories, and geographies, which is at the center of Indigenous identities. For example, Canada’s residential school system lasted for hundreds of years. Its intention was to erase Indigenous peoples from their identity and sever their ties to their ancestral lands and their Indigeneity. However, Indigenous Peoples have their own education systems which they have had since time immemorial, and certainly prior to the European sense of entitlement to Indigenous Lands through the idea of *Terra nullius*, or “the empty lands”. The settlers’ desire for Indigenous Lands historically was relentless and remain that way today. It was revealed through the European education imposed on Indigenous children and families. According to the previously mentioned Final Report (2021), Indigenous children often faced extreme hunger, disease, and physical, mental, emotional, and sexual abuse (2021, p. 39). This systemic abuse, and erasure is part of the structure of Canada’s residential schools designed to inflict genocide on Indigenous People. Sium et al. (2012, pp. I–XIII) discuss how colonialism and white supremacy have viewed Indigeneity as “different and threatening” and as something to marginalize and erase out of existence.

Moquino and Kitchen (2020) state that Canada’s residential schooling served to violently assimilate Indigenous people. They make the connection between Land theft and the assault on Indigenous sovereignty. The Canadian system continue to erase Indigenous youth through its contemporary education system which remains structured on Eurocentric models and approaches.

The education system today is still centered on the historical continuation of settler-colonial logic of elimination (Wolfe, 2006). This means that the school and education system is immersed in a system that expects everyone else to fit into it as it incessantly reproduces itself with the intent to erase Indigenous youth. Moquino and Kitchen (2020, para 20) state that the colonial schooling has functioned to maintain settler colonialism through epistemicide—the attempted killing of knowledge.

It is important to understand the purpose of education, who education is for, and which forms of education counts. Is education for assimilationist purposes? This is an important question that must be critically considered when thinking about both education and schooling (Report of the African Commission, 2005, p. 55). In 2020, The World Health Organization declared a pandemic. Bishop (2020) suggests that the coronavirus pandemic has changed how education is carried out globally. As Bishop states, it does not mean that education has not remained Eurocentric. Rather, the author points out that the response to the pandemic has shown that systemic transformation can happen, and it can be done quickly (p. 6).

Indigenous People’s right to their education are affirmed in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP, 2007), Article 14, Number 2, states the following unequivocally. Indigenous individuals, particularly children, have the right to all levels and forms of education of the State without discrimination (p. 13). Indigenous people’s rights to their education are on an individual and collective level. Many Indigenous nations continue to implement their own education systems, and Cajete calls for the creation of community-based curricula (Cajete, 2020, p. 7), and specifically curricula that reflect Indigenous peoples’ multiple worldviews. The author provides an example of the Medicine Wheel, Corn, or Tree of Life symbolism which have deep metaphorical meanings that frame essential goals and visions of Indigenous education and sustainability.

These wholistic models of education are designed according to Indigenous ethics and models. Cajete (2020, p. 7) exclaims, Indigenous educators work with Indigenous communities and students around place Land and project-based curriculum models. Indigenous education involves developing relationships between teacher and learner and ongoing commitment and that relationality is reciprocal (Bishop, 2020, p. 10). Harris and Wasilewski (2004, p. 494) say of this reciprocity that “overlapping relationships will always take care of us. This was why there were no orphans in Comanche society”. Indigenous Education means the family is at the center of learning, teaching and relationships. Harris and Wasilewski write children were the responsibility, not just of the mother, but of the mother’s entire family and, ultimately, of the community. This is another example of responsibility emerging from relationships. This is what Cajete refers to as a “dynamic flow of interrelationship” (2015, p. 365). Indigenous Lands and education systems are interconnected and entrenched in the understanding of Land as a wholistic Knowledge system that embodies the concept of reciprocity.

Indigenous philosophies of reciprocity

Similarly, Reciprocity, like Land and education, involves exchanges and is part of Indigenous worldviews, governing systems, and Indigeneity. Reciprocity is a cyclical obligation. It emphasizes the fact that, in nature or natural law, there is a circular model approach. For instance, the seasons, life, as well as the dynamics between any two entities in relationship with each other. Once we have encountered another, we are in a relationship with them write (Harris and Wasilewski, 2004, p. 493).

According to [Cajete \(2020, p. 9\)](#), Indigenous worldview is relational. Moreover, building upon and extending reciprocity is an essential process for developing relationships. Reciprocity is then interconnected with relationships. [Alfred and Cornthassel \(2005, p. 609\)](#) assert, Indigenous identity, and interconnection, which is historical, ceremonial, and involves languages and Lands, is also about relationships and kinship networks. Reciprocity is structured within kinship which is connected to Indigenous Land and education systems. Moreover, it is through these relationships that [Bishop \(2020, p. 5\)](#) writes about Indigenous sovereignty being enacted in everyday lives through these interconnecting relationships which are immersed in giving and taking and the principle of respect. In Canada, despite the history of, and contemporary reality of, settler colonialism, as well as the history of the Canadian state's genocidal residential school systems, Indigenous people continue to transmit their knowledges ancestrally and generationally. This includes models of reciprocity of how to live the "good life" for which Indigenous peoples globally have their own terms in their own languages. This exemplifies the values of relationships and the significance of community to Indigenous peoples. What's more, intergenerational reciprocity is a reminder of Indigenous people's obligations to one another—not simply to give back, but to listen, to share, and to give forward ([Bishop, 2020, p. 11](#)).

[Postero \(2013, p. 442\)](#), writes that to understand Indigeneity as embodied in a reciprocal relationship with other-than-human beings raises questions and challenges of what constitutes knowledges and forms of knowledges. Additionally, reciprocity is about sharing, and Indigenous lived experiences are connected to Indigenous education. For instance, Indigenous lived experiences focus on the ways in which Indigenous bodies—in all their diversity—are treated in the colonial present, and how bodily experiences come to express the boundary between indigeneity and non-indigeneity ([Postero, 2013, p. 437](#)).

As Alfred and Cornthassel assert (2005, p. 609):

it is the need to maintain respectful relationships that guides all interactions and experiences with community, clans, families, individuals, homelands, plants, animals etc. in the Indigenous cultural ideal. If any one of these elements of identity, such as sacred history, is in danger of being lost, unified action can be taken to revitalize and restore that part of the community by utilizing relationships, which are the spiritual and cultural foundations of Indigenous peoples.

Reciprocity is learned through Land ethics and is lived, observed, and taught by family and the land itself symbolizes Indigenous pedagogy. Indigenous youth today practice ceremonies and Indigenous models of education regardless of where in the world they live, as they did prior to contact with European colonizers and settlers. Youth learn about the Land and plants through their everyday lived experiences. For that reason, reciprocity among people, as an Indigenous foundation, is intimately connected, interactive and interdependent on one another. It is this deep respect for relationships that embeds family relations, and community ([Cajete, 2015, p. 365](#)). These relationships inform Indigenous knowledge transfer systems generationally, but also as a site of memory within the body. [Dei and Jaimungal, 2018, p. 63](#)) emphasizes Indigenous critical perspectives, practices, epistemologies, roles, and spaces for sharing and producing local cultural knowledges and their importance to community and relationships.

For [Cajete \(2015, p. 367\)](#) the life of Indigenous community is mutually reciprocal and interdependent with the living communities in the surrounding natural environment. Although Indigenous peoples come from many different languages and cultures, these similarities in their focus on interrelationships are embodied within. Indigenous societies and knowledge systems have developed to sustain reciprocal relationships between culture and nature, and therefore they use scientific approaches that are rigorous in their own methods and rely on long-term observations ([Johnson et al., 2015, p. 8](#)). This intricate and dynamic relationship demonstrates Indigenous worldviews alive and in action. Reciprocity includes additional core Indigenous principles and philosophies. For example, [Harris and Wasilewski \(2004, p. 492\)](#) refer to these core values as the Four "R" s which include the relationship, responsibility, reciprocity, and redistribution. These Four "R" s personify Indigenous ethics that are central to Indigenous ways of life. Moreover, the authors speak about relationships as the kinship obligations to the multiple forms of life that exist and interrelate with one another.

[Cajete \(2005, p. 70\)](#) discusses the processes and principles of reciprocity between humans and all living beings and reinforces the idea that reciprocity is also grounded in relationality. Indigenous peoples, through their lived experiences and worldviews, understand how relationships shape and influences responsibilities. The relationships are grounded in Indigenous ethics of relationships and are connected to the concept of Indigeneity ([Harris and Wasilewski, 2004, p. 493](#)). Given the history and impact of the devastation of colonialism and the ways in which its destructive history and current actions have created serious implications for Indigenous peoples Throughout the world, there has been a reassertion and regeneration of Indigenous lifeways. [Kimmerer \(2018, p. 41\)](#) speaks of restoring reciprocal relationships and reengaging people with the Lands that continue to provide sustenance. This understanding is related to Indigenous identity and is rooted in the essence of the Indigeneity of Indigenous peoples globally.

Conclusion

This chapter has provided a discussion of the concept of Indigeneity, and the importance of relationships, Lands, Indigenous Education, and reciprocity to Indigeneity. Although Land is of critical significance to Indigenous Peoples regardless of where they live, many Indigenous people have been systemically removed and displaced from their Lands. Indigenous People have been impacted in different ways and at different times by this violent and systematic separation and displacement from their Lands.

As Radcliffe, 2017 observed, settler colonialism is ongoing, and it is not just a historical or single event. Consequently, Indigenous people continue to be impacted in multiple devastating ways which go beyond geography. Reciprocity with Indigenous educational approaches and the Land have also been impacted because settlers, historically and in current times, continue to have a deep obsession with, and desire for, Indigenous lands. Yet, Indigenous people are finding their own ways to regenerate and rebuild their Lands and pedagogical systems which are grounded within the principle of reciprocity.

Indigenous Peoples come from multiple places and understandings of connectedness and relationality. Therefore, it is recognized that to speak of Indigeneity is to also acknowledge that Indigenous people around the world are not uniform or static (Romm, 2015, p. 420). Indigenous People come from dynamic and fluid cultures. According to Alfred and Cornstassel (2005, p. 597), "the world's indigenous peoples in their cultures, political-economic situations, and in their relationships with colonizing settler societies are varied." However, their determination, "to survive as distinct peoples on foundations constituted in their unique heritages, attachments to their homelands, and natural ways of life is what is shared by all Indigenous peoples."

No matter where Indigenous People are in the world, it is essential and necessary to listen to them. Listening is an Indigenous educational model and approach which binds Land and reciprocity. To develop reciprocal relationships requires deep listening. To understand the Land means to be able to listen to its stories. Indigeneity begins with listening to, and honoring, Indigenous voices. Indigenous people centering Indigenous voices are creating and constituting, reciprocity, wherever in the world they might be.

References

- Alfred, T., Cornstassel, J., 2005. Being indigenous: resurgences against contemporary colonialism. *Gov. Oppos.* 40 (4), 597–614. <https://doi.org/10.1111/j.1477-7053.2005.00166.x>.
- Assing, T., 2016. Dining like the ancestors. *Caribbean Beat Mag.* <https://www.caribbean-beat.com/issue-141/dining-like-ancestors#ixzz72uPXOWfl>.
- Baird, I.G., 2016. Indigeneity in Asia: an emerging but contested concept. *Asian Ethn.* 17 (4), 501–505. <https://doi.org/10.1080/14631369.2016.1193804>.
- Barker, A., Battell Lowman, E., 2015. Settler Colonialism, in *Global Social Theory*. Available at: <https://globalsocialtheory.org/concepts/settler-colonialism/>.
- Bishop, K., 2001. This land knows me: indigenous land rights. *Cultural Survival Q. Mag.* CSQ Issue 25-1 Available at <https://www.culturalsurvival.org/publications/cultural-survival-quarterly/land-knows-me-indigenous-land-rights>.
- Bishop, M., 2020. Indigenous education sovereignty: another way of "doing" education. *Crit. Stud. Educ.* 1–16. <https://doi.org/10.1080/17508487.2020.1848895>.
- Cajete, G., 2005. American Indian epistemologies. *New Directions for Student Services* 109, 69–78. <https://doi.org/10.1002/ss.155>.
- Cajete, G.A., 2015. Indigenous education and the development of indigenous community leaders. *Leadership* 12 (3), 364–376. <https://doi.org/10.1177/17427150150160412>.
- Cajete, G.A., 2020. Indigenous science, climate change, and indigenous community building: a framework of foundational perspectives for indigenous community resilience and revitalization. *Sustainability* 12 (22), 9569. <https://doi.org/10.1080/17508487.2020.1848895>.
- Dei, G.J., 2016. Conceptualizing indigeneity and the implications for indigenous research and African development. *Confluence* 2. Available at: <https://scholarworks.iu.edu/iupjournals/index.php/confluence/article/view/532>.
- Dei, G., Jaimungal, C.S., 2018. Indigeneity and decolonial resistance: an introduction. In: Dei, G.J.S., Jaimungal, C. (Eds.), *Indigeneity and Decolonial Resistance: Alternatives to Colonial Thinking and Practice*. Myers Education Press, Maine, pp. 11–35.
- Final Report, 2021. A History of Wood Buffalo National Park's Relations With the Denésuliné. Willow Springs Strategic Solution; Athabasca Chipewyan First Nation (ACFN). Available at: <https://www.theglobeandmail.com/files/editorial/politics/WBPNP.pdf>.
- Grydehøj, A., Nadarajah, Y., Markussen, U., 2018. Islands of Indigeneity: Cultural Distinction, Indigenous Territory and Island Spatiality. <https://doi.org/10.1111/area.12520>.
- Harris, L.D., Wasilewski, J., 2004. Indigeneity, an alternative worldview: four R's (relationship, responsibility, reciprocity, redistribution) vs. two P's (power and profit). *Sharing the journey towards conscious evolution. Syst. Res. Behav. Sci.* 21 (5), 489–503. <https://doi.org/10.1002/sres.631>.
- Jaffee, L., John, K., 2018. Disabling bodies of/and land: reframing disability justice in conversation with indigenous theory and activism. *Disabil. Global South* 5 (2), 1407–1429.
- Johnson, J.T., Howitt, R., Cajete, G., Berkes, F., Louis, R.P., Kliskey, A., 2015. Weaving indigenous and sustainability sciences to diversify our methods. *Sustain. Sci.* 11 (1), 1–11.
- Kimmerer, R., 2018. Mishkos Kenomagwen, the lessons of grass: restoring reciprocity with the good green earth. In: Nelson, M.K., Shilling, D. (Eds.), *Traditional Ecological Knowledge Learning From Indigenous Practice for Environmental Sustainability*. Cambridge University Press, pp. 27–57.
- Koot, S., Hitchcock, R., Gressier, C., 2019. Belonging, indigeneity, land and nature in southern Africa under neoliberal capitalism: an overview. *J. South Afr. Stud.* 45 (2), 341–355. <https://doi.org/10.1080/03057070.2019.1610243>.
- Matute, I.D., 2021. Indigeneity as a transnational battlefield: disputes over meanings, spaces and peoples. *Globalizations* 1–17. <https://doi.org/10.1080/14747731.2020.1771962>.
- McIntosh, I.S., 1999. The indigenous people of the Caribbean. *Cultural Survival Q. Mag.* CSQ Issue 23-4. Available at <https://www.culturalsurvival.org/publications/cultural-survival-quarterly/indigenous-people-caribbean>.
- Moquino, T., Kitchens, K., 2020. Solidarity Necessitates Reciprocity: Confronting Settler Colonialism in Antiracist Education. *Embracing Equity.Org.* Available at: <https://embracingequity.org/blog/2020/9/22/solidarity-necessitates-reciprocity>.
- United Nation, 2009. State of the World's Indigenous People. Department of Economic and Social Affairs. https://www.un.org/esa/socdev/unpfii/documents/SOWIP/en/SOWIP_web.pdf.
- Neeganagwedgin, E., 2021. Reflections on Land Back and Education, vol. 1. Rooted: An Indigenous Law Publication, pp. 43–47. Issue 2.
- Postero, N., 2013. Introduction: negotiating indigeneity. *Lat. Am. Caribb. Ethnic Stud.* 8 (2), 107–121. <https://doi.org/10.1080/17442222.2013.810013>.
- Radcliffe, S.A., 2017. Geography and indigeneity I. *Prog. Hum. Geogr.* 41 (2), 220–229. <https://doi.org/10.1177/0309132515612952>. <https://doi.org/10.1177/0309132515612952>.
- Radcliffe, S.A., 2018. Geography and indigeneity II: critical geographies of indigenous bodily politics. *Prog. Hum. Geogr.* 42 (3), 436–445. <https://doi.org/10.1177/0309132517691631>.
- Report of the African Commission's Working Group of Experts on Indigenous Populations/Communities, 2005. Available at https://www.igwia.org/images/publications/African_Commission_book.pdf.
- Restoule, J.-P., Gruner, S., Metatawabin, E., 2013. Learning from place: a return to traditional Mushkegowuk ways of knowing. *Canadian J. Educ.* 36 (2), 68–86. Available at: <https://journals.sfu.ca/cje/index.php/cje-rce/article/view/1570>.
- Romm, N.R.A., 2015. Reviewing the transformative paradigm: a critical systemic and relational (indigenous) lens. *Syst. Pract. Action Res.* 28 (5), 411–427. <https://doi.org/10.1007/s11213-015-9344-5>.
- Sium, A., Desai, C., Ritskes, E., 2012. Towards the "Tangible Unknown": decolonization and the indigenous future. *Decolonization* 1 (1), I–XIII. Retrieved from: <http://transnationaldecolonialinstitute.wordpress.com/decolonial-aesthetics/>.

- Steeves, P.F., 2018. Indigeneity. *Anthropology*. Oxford Bibliographies. <https://doi.org/10.1093/OBO/9780199766567-0199>.
- United Nations, 2007. United Nation Declaration on the Rights of Indigenous Peoples (UNDRIP). UN. https://www.un.org/esa/socdev/unpfii/documents/DRIPS_en.pdf.
- United Nations Development Programme, 2021. 10 Things to Know About Indigenous Peoples by United Nations Development Programme on Exposure. UNDP. <https://stories.undp.org/10-things-we-all-should-know-about-indigenous-people>.
- Warren, S.D., 2016. Indigenous in the city: the politics of Urban Mapuche identity in Chile. *Ethn. Racial Stud.* 40 (4), 694–712.
- Wolfe, P., 2006. Settler colonialism and the elimination of the native. *J. Genocide Res.* 8, 387–409. <https://doi.org/10.1080/14623520601056240>.
- World Bank, 2021. Indigenous Peoples. World Bank. Available at. <https://www.worldbank.org/en/topic/indigenouspeoples>.

Intersectionality in the Caribbean and Africa

Annette Henry^a, Bathseba Opini^b, and Adeerya Johnson^c, ^a Language and Literacy Education, Faculty of Education, Institute of Gender, Race, Sexuality and Social Justice, University of British Columbia, Vancouver, BC, Canada; ^b Educational Studies, Faculty of Education, University of British Columbia, Vancouver, BC, Canada; and ^c Institute for Gender, Race, Sexuality and Social Justice, Faculty of Arts, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Intersectionality: a black feminist concept	116
A brief overview	116
Intersectionality across disciplines	117
Intersectionality in education	117
The Caribbean	118
Africa	120
Conclusion	121
References	121

There is no such thing as a single-issue struggle Lorde

Almost all social science handbooks and encyclopediae contain a chapter on intersectionality (Dill and Kohlman, 2012; Grant and Zwier, 2017; Sibbett, 2020). A number of systematic reviews also exist on the subject (Harris and Leonardo, 2018; Nichols and Stahl, 2019; Robert and Yu, 2018). In December 2013, the United Nations General Assembly adopted Resolution 68/237 which proclaimed 2015 to 2024 to be the International Decade for People of African Descent, with the theme “People of African Descent: Recognition, Justice and Development.” The theme emphasizes promoting greater knowledge, recognition of, and respect for the culture, history and heritage of people of African descent, including through research and education (UN, 2013). Considering this declaration and the global Black Lives Matter movement, we sought to focus on intersectionality research in education by people of African descent in Africa and the Caribbean whose scholarship addresses matters of Black life and well-being from an education standpoint. Furthermore, given the international charge of this volume, this chapter is devoted to an exploration of whether and how intersectionality is used in the educational research literature of these two regions as a way of inclusively and comparatively internationalizing the discussion of the concept and in this way probing some of the gaps, limitations and possibilities of intersectionality.

In our initial meetings regarding this project, we posed the question of whether intersectionality is an ‘Americocentric’ term used in social justice studies in education in the Northern hemisphere, and, perhaps even an imperial imposition of ways of knowing and being onto Global South regions. We posed the question, “What might be the correspondences in predominantly Black countries? In other words, how might African continental and African Caribbean researchers understand and theorize the complex dimensions and processes of their lives, in their own contexts? What might be the (equivalent or unique) terminologies specific to those settings? What kinds of intersectionalities can we talk about in these regions?

Intersectionality: a black feminist concept

A brief overview

In this section we review the concept of intersectionality, developed in the United States by critical feminist legal scholar Kimberl Crenshaw. The concept was first advanced in two law review articles (Crenshaw, 1989, 1991) whereby Crenshaw highlighted the ways in which African American women were misrepresented or erased in the legal system. This theorization critiqued linear single-issue agendas and additive approaches to Black women’s lives (Bilge and Denis, 2010; Bowleg, 2008). The concept stressed the multidimensionality of African American women’s experiences that include race, gender, and class and other dimensions and how they intersect in distinct ways in societal of marginalization and subordination. The dynamism and simultaneity of these social identities were not being taken into account in legal cases, causing further systemic inequalities for Black women. Intersectionality, then, was advanced as a way of analyzing these interconnections between inequalities in a complex system of power. Intersectional frameworks do not envision oppressions as competing, but rather as working mutually in and through all social relations and systems of power (Collins, 2017). African American scholar Patricia Hill Collins stresses the importance of a historical understanding of the theory and practice of intersectionality (e.g., Collins, 2000, 2015; Collins and Bilge, 2016). That is, although the term is attributed to Crenshaw who provided an analysis in her legal writing, intersectionality has been a fundamental part of Black women’s epistemology, activism, and intellectual tradition. In the United States, the “Ain’t I a woman” speech of abolitionist and former slave Sojourner Truth (1797–1883), at a women’s rights convention in Ohio in 1851 and the writings of Maria Stewart (1803–79) are two early examples of what we now understand to be an intersectional analysis. Likewise, the *Combahee River Collective Black Feminist Statement* (1977) and other Black feminist scholars wrote about the intersecting multiple oppressions and

recognized the ways in which multidimensionality and simultaneity of their lived realities in a complex system of power relations before Crenshaw's analysis, often using descriptions such as "interlocking" or "intersecting" (Collins, 1987; Davis, 1981; Zinn and Dill, 2016). Indeed, Black women throughout the African diaspora are able to describe their lived realities of power and subordination, regardless of social location. This concept has provided a tool for Black feminist researchers to broaden as well as sharpen their analyses of Black women's lives and realities and analyze the ways that their realities may vary across identities (such as class, sexuality, disability) and race and contexts (such as urban, rural, poor). Theories know no borders and thus the concept has been embraced across the English-speaking African diaspora as well as across international research communities. Time and space do not allow for a discussion of the ways that intersectionality has been "stretched" and reshaped across geographic contexts, nor its critiques (see Evans, 2016; for a discussion in the UK, especially among young people.).

Intersectionality across disciplines

Since the 1980s, intersectionality has increasingly been applied to academia, applied professions, and justice-oriented praxes to critically analyze and change systems and processes of inequity and marginality. Intersectionality has been subject to much critique, such as lacking theoretical clarity (McCall, 2005) becoming diluted and depoliticized (Bilge and Denis, 2010), being coopted by White women, thereby erasing Black women's theoretical and political traditions (Mugabo, 2022; Nash, 2017). Canadian sociologist Sirma Bilge (2020) writes that intersectionality has become, "an empty shell onto which scholars of all stripes can project their own concerns and feel legitimate to do so" (p.1). At the 2014 American Studies Conference, on a panel called "Kill this keyword," intersectionality was declared "dead" as one of the keywords to be eliminated. This act is reminiscent of other "deaths," such as the death of the subject, in which Nancy Hartsock suggested, "Why is it, exactly at the moment when so many of us who have been silenced begin to demand the right to name ourselves, to act as subjects rather than as objects of history, that just then the concept of subjecthood becomes 'problematic'?" (Hartsock, 1983, p. 196). Indeed, Nash (2019) suggested that in essence the declaration to kill intersectionality was a declaration to kill Black women. Canadian Black feminist Délice Mugabo (2022) sums it up: "In other words, Black women's intellectual labor must be erased for intersectionality to reach its latent potential for white feminism."

Intersectionality in education

In education, intersectionality as a critical analytical tool is applied to unearth and challenge inequities for feminist researchers (Henry, 2015; Nyangweso, 2017). For example, Henry (2015) contends despite her leadership position as a department head, unlike her white male predecessors, her race, gender and perceived "outsider" status positioned her as inferior and incompetent according to her white colleagues in a white supremacist institution where systemic racism pervaded daily routines and practices. An intersectional analysis allowed her to prise apart the complex system of power relations in which she was embroiled.

Intersectionality is often employed as part of a larger Critical Race Theory (CRT) framework. CRT originated in the 1970s. Activist lawyer Derrick Bell and a group of mostly Black legal scholars at the time were interested in the ways that race and power were structured and routinized into every facet of life in the United States. Since that time, it has evolved and been adapted to many marginalized groups. Offshoots such as Critical Latinx Theory or LatCrit (Martinez, 2019), Critical Asian Theory or AsianCrit (Iftikar and Museus, 2018), Critical Disability Theory or DisCrit (Annamma et al., 2013) have emerged as specific applications. Likewise, critical scholars who study race outside of the United States have adapted and reshaped Critical Race Theory for their contexts. The tenets of CRT are easily adaptable across white supremacist capitalist societies all of which are characterized by structural and institutional racism.

Education researchers use intersectional frameworks and CRT (which includes intersectionality) to examine inequalities across multiple social contexts and identities and how these influence educational opportunity and disparities in distinctive ways (Núñez, 2014; Harris and Leonardo, 2018). For example, James's (2021) analysis of Black male student success in Toronto, Canada takes into consideration constructs such as class, nationality, ethnicity, school environment, community influences, parental education and whether the youth are first or second generation. Critically probing the simultaneity of all these dimensions and how they played out in the boys' lives affords a more complex understanding of why one boy became a successful lawyer and the other, a bus driver. As Patricia Hill Collins and Valerie Chepp (2013, p.58) write: "[T]hese constructs signal an intersecting constellation of power relationships that produce unequal material realities and distinctive social experiences for individuals and groups positioned within them." Similarly, in their systematic review of intersectionality in higher education, Nichols and Stahl (2019) concur: "when considering the workings of multiple systems of (dis)advantage across gender, race and class, it is not possible to simply focus on academic participation and achievement in isolation from the social and personal aspects" (p.7).

Intersectional educational research is found, for the most part, in qualitative studies (Covarrubias and Vélez, 2013). Quantitative studies may present certain challenges regarding intersectionality—perhaps the reason for terms such as "intersectionality-informed" quantitative approaches (Rouhani, 2014). Bowleg (2008) argues that statistical methods, even those that test interactions, were not designed with the study of intersectionality in mind (p. 320). It is important to acknowledge that one's choice of methods is determined by the kinds of research questions being asked, the underlying epistemological underpinnings, as well as one's view of factors such as the nature of knowledge, and the purposes and meanings of research. Nonetheless, educational researchers have been stretching the boundaries of intersectionality using Critical Race Theory studies in education to examine inequalities using quantitative methods, hence the range of terms QuantCrit (Gillborn et al., 2018) QuanCrit (López et al., 2018), and CritQuant (Sullivan et al., 2010). In her 2018 article in the *American Educational Research Journal* (AERJ), Sung Tae Jang (2018) advances a quantitative

intersectionality analysis of South East Asian females' educational outcomes in the United States. She argues that the use of numbers and statistical inferences reveal structural inequalities deeply associated with power relations between different social categorizations and that "In this regard, critical quantitative researchers have an essential and powerful role in creating conditions of equity that attend to intersectionality: They can inform and alarm policymakers and school leaders about structural inequalities as well as realize educational equity and social justice" (p.1299).

Cho et al. (2013) laud "the remarkable degree of theoretical and methodological engagement that the concept of intersectionality has invited among feminist and antiracist scholars around the globe" (p. 787). Tormos (2017) writes that "intersectionality has gone global" (p.1). How global is the concept of intersectionality? It was indeed our quest to examine whether intersectionality was even a concept in Global South education research as we were not familiar with its usage in these regions in the educational literature. We set out to explore research articles between 1990 and 2020, given the date of Crenshaw's seminal articles in 1989 and 1990. More challenging were decisions about criteria for inclusion. For example, do we include diasporic authors who may have left their birth country and who write about education in their homelands? Epistemic questions were always at the fore: the term comes out of African-American reality. Can we apply African American terminology to the African diaspora with their varied histories of capitalism, slavery and colonialism? We wondered whether we would find sufficient studies, at first, and whether we were imposing American conceptual patterns of thinking on global contexts. We also wondered, whether it was imperative for authors to use the word "intersectionality"? (See Roth (2013) who contends that the concept is deeply Euro- and US.-centric can be seen as an instance of epistemic Occidentalism and thus must be rethought for use in transnational contexts and for other locations of knowledge production).

We discuss our methods in the respective sections of the Caribbean and Africa as we were examining different parts of the globe, different, journals, and regions with different racializing histories.

The Caribbean

It is important to acknowledge the diversity and hybridity of the Caribbean. This majority Black region consists of 15 nation states: It has endured various histories of European enslavement and colonialism. It is comprised of various ethnicities, races, cultures, European languages, Creoles and identities. It is, thus, a multiracial, multilingual, multi-faith region that has undergone inter-racial mixing through waves of migrants. It is comprised of many groups such as Blacks, South Asian, Chinese, Jews, Arabs, Europeans to name a few. Similar to many nation states with a history of colonialism, its Indigenous populations are insufficiently acknowledged.

The inclusion criteria for the Caribbean articles were that they had to be written by African Caribbean people who lived and worked in the English-speaking Caribbean and who wrote about African Caribbean educational issues using intersectionality. That is, papers by African Caribbean people living outside of the Caribbean would be excluded in understanding current trends in Caribbean countries. Admittedly, this criterion was problematic by definition, a limitation of the research that risks essentialism. Firstly, as noted above, due to the multiraciality of the Caribbean, the notion of "Black" is quite variegated, and often politicized. Thus, how can we always know how authors identify themselves in this polyracial region? Identities are intertwined with other social categories, so that, for example, race, skin color, religion and gender are intimately articulated within structured hierarchies in Caribbean countries (Law and Tate, 2015; Thomas, 2011). Secondly, due to scholarly collaborations and global mobility, scholars who reside in the Caribbean collaborate globally and these international collaborations produce exciting new understandings and thus, as mentioned above borrow and re-shape theories from North American and British contexts. In fact, in conducting research for this essay, we were struck by the amount of Caribbean/diasporic research collaborations between Caribbean-based researchers at the University of West Indies and their Caribbean colleagues in other countries, especially in the United States. All this to say, it is important to acknowledge up front, the impossibility of total accuracy and to focus on the larger goal of exploring if and how intersectionality was a tool for Caribbean educational research.

The initial searches in databases such as Summon, EBSCO Academic Search Premier, ERIC and Google Scholar revealed that few studies published in the Caribbean were accessible through the databases. For example, an EBSCO search for articles using the descriptors "Intersectionality," Caribbean," and "education" written after 1990 that explicitly acknowledged intersectionality in the title or abstract or in the main argument yielded 20 articles (This was one of many descriptor searches.) Only one study was conducted in the Caribbean that is, Ellis's insightful (2018) intersectional study of boys' achievement in a Jamaican classroom in *Gender and Education*. Another study, Khabir's autoethnography of her four months in Brazil was published in the *Journal of Education and Development (JEDI)* a Caribbean journal. Our aim was to feel the pulse of Caribbean-based scholarship Neither study met the criteria, as the former was a doctoral student in Canada and thus influenced by the Canadian scholarly conversations. The second was a narrative of an African American woman. Based upon the available search findings, a decision was made to focus on the educational studies in *The Caribbean Journal of Education (CJE)* and *The Journal of Education and Development in the Caribbean (JEDIC)*. We had access to the abstracts through the University of West Indies sites, <https://www.mona.uwi.edu/soe/publications/cje>, and <https://www.mona.uwi.edu/soe/publications/jedic>. Unfortunately, the searchable software program only allowed access to abstracts rather than access to articles. These journals were chosen for accessibility and because they represent two long-standing, well-regarded peer-reviewed Caribbean educational journals in which Caribbean-based and international authors publish their empirical research. This choice created an immediate concern of not representing the range of methods and studies produced in the Caribbean, as authors may choose to publish in other Caribbean-based journals outside of education or in international education journals. Intersectionality as a search term was absent in these journals. Thus, we employed other key terms such as "gender"

and “ethnicity” singularly and in combination: “gender + ethnic(ity)”, “religion” and “boys and achievement, socioeconomic and Religion”. Indeed, after finding authors’ work with these descriptors, we cross-checked with Google Scholar and library databases for 1) the article in question, or a variation thereof in another journal, b) other works by these authors that could ascertain their conceptual and analytic frameworks; and c) authors’ ResearchGate and [Academia.edu](https://www.academia.edu) accounts for versions of select papers. These steps were added to find full-length articles, and not to rely solely on the abstracts. Furthermore, along with cross-checks from co-authors, they acted as a trustworthiness check, due to the inability to access more than the abstracts in CJE and JEDIC outside of the Caribbean.

The above challenges were two of many systemic challenges in the attempt to gather Caribbean research studies in education. Databases in five library systems in the USA and Canada were searched. Libraries lack subscriptions to journals in the Global South. This has far-reaching consequences for scholars world-wide. Researchers in developing and Global South countries often rely on databases, journals and research from North American and European countries, yet researchers in North American and European countries are less interested in their Global South colleagues’ research. The institutional Whiteness manifest in library holdings dampens intellectual dialogue between Global North and South regions and represents a powerful hierarchy that determines whose knowledge is worthwhile, what methods are ‘rigorous’ ‘scientific’ and ‘valid’ and whose work gets reviewed and published ([Galvan, 2015](#); [Olson, 2004](#)). Therefore White norms remains comfortably intact ([Yancy, 2012](#)). Certainly, libraries have budgets and cannot possibly subscribe to every possible journal. However, Black knowledges are low on the hierarchy of what is deemed most valuable knowledge.

These systemic challenges represent the ever-present legacies of White supremacy and colonialism in Education. In sum, these institutional challenges exemplify intersectionality in everyday life—How social inequalities are reproduced through power relations that is, the ways that capitalism, racism and Whiteness and the knowledge economy produce political and intellectual domination in people’s lives and have larger consequences of people in the Caribbean.

All abstracts from CJE (1997–2021) and JEDIC (1974–2021) were searched for “intersectionality” as they were readily available online. These two majority-Black peer-reviewed journals have steadfastly published Caribbean-based empirical studies using a range of methodologies - quantitative, qualitative and mixed methods research and conceptual frameworks. With the exception of Khabir’s intersectionality article mentioned earlier which does not meet inclusion criteria, the preponderance of articles in the journals employed classical empirical research designs and did not use intersectionality as a conceptual or methodological tool in the ways that it has been formulated by Crenshaw or CRT educational research. Indeed, studies examined multiple factors that influence student achievement. For the most part, these multifactorial studies were quantitative (descriptive and correlational studies examining factors influencing phenomena), such as the relationship between parenting style and adolescent communication ([Anderson and Cole, 2017](#)), associations between variables such as gender, ethnicity, socioeconomic status, and disability in the examination of literacy in underperforming youth ([Smith et al., 2020](#)). Explicit analyses of power relations were not part of these theoretical frameworks. Only one qualitative article “A Critical Multicultural Analysis” of Jamaican Children’s Literature explicitly acknowledged the oppressive power structures of race and class and the legacies of slavery and colonialism as a central organizing principle ([Richards, 2013](#)).

By comparison, *The Caribbean Review of Gender Studies (CRGS)*, published at the University of West Indies, at the St. Augustine Campus in Trinidad and Tobago, was more likely to include interdisciplinary scholars as well as education scholars who utilize intersectional frameworks and who live and work in the Caribbean (e.g., [Stewart, 2019](#)) or the Caribbean diaspora (e.g., [Douglas, 2019](#)). This journal was not included in our study as it is not an educational research journal. However it points to the reality that in general, Caribbean authors publish in journals with aims and scopes that appeal to particular audiences, many of which as noted above, are international. Our narrowly defined research domain is not representative of the range of frameworks in which Caribbean education researchers are carrying out their research, nor does it examine the scope of intersectionality research employed by scholars in the Anglophone Black Caribbean diaspora. Instead, it exemplified the studies published in its two major education journals as an exploration of intersectionality in this region and gave an overview of the types of educational studies conducted.

In sum, based on a review of CJE and JEDIC abstracts, at the time of writing, intersectionality as a method of critical inquiry had not been readily employed in Caribbean educational research journals in the ways that it has been defined earlier. However, many Caribbean colleagues are engaged in transnational research connections, it will mostly likely be adopted in educational research since it is currently being used in other Caribbean scholarship (e.g., [Bailey and Ngwenyama, 2012](#); [Stewart, 2019](#)).

Despite the fact that the studies examined cannot be considered “intersectional” by Crenshaw’s definition, nonetheless, Caribbean scholars (and most people living in the Caribbean) are well aware of the legacies of European colonialism in contemporary life: poverty, class, race, colorism, gender are implicated in violence, inequality, oppression and denial of equal opportunity. These very inequities underpin much of the current educational research agendas in the Caribbean using traditional research designs and methods. Significantly, the Caribbean has had a long history of anticolonial struggles and a long tradition of Black radical activist-intellectuals who have critiqued these issues who have made significant contributions in the Caribbean and its diaspora such as Marcus [Garvey \(1983\)](#), Walter [Rodney \(1973\)](#), Claudia Jones ([Boyce Davies, 2007](#)) and Andaiye ([Trotz, 2020](#)), to name a few.

Tormos’s claim that “Intersectionality has gone global,” might very well point to the Eurocentricity of her statement. One might find it encouraging that all that is done in the North is not blindly adopted in South. That said, it is promising to re-imagine studies of social and educational stratification in Caribbean-specific ways to examine how power relations multiply marginalize children’s achievement, for example, and how power operates through structural (e.g., the school) and cultural (churches, families) institutions as well as identities such as sex, gender, class, disability, language, ethnicity, race, and other dimensions. Critically analyzing

the constructs as interdependent and mutually constitutive in context-specific analyses would involve in a reworking of intersectionality from Caribbean-informed perspectives.

Africa

Africa is a continent consisting of 54 countries with over 3000 ethnic communities who speak close to 2000 languages. We recognize the diversity in pre-European contact and colonial histories, education systems and structures as well as the lived experiences in the different countries that make up the African continent. We draw on examples of works of Black African scholars from English-speaking African countries including Ghana, Kenya, Nigeria, South Africa, Uganda, and Zimbabwe to inform our analysis. This implies that the analysis offers perspectives from contexts where English is the “official” lingua franca while scholarship in Indigenous African languages is not represented and is systematically ignored in scholarly research and publications. Nonetheless it would be important to capture the literature from these contexts as their perspectives are seldom heard. Moreover, we did not review articles published in French, Portuguese or Spanish. It is important to recognize that intersectional thinking is embedded in the worldviews of Black African peoples even though it is not specifically named as intersectionality. Black African scholars have built rich scholarship that points to the connections between different identity markers including gender, social class, race, ethnicity, religion, sexuality, as well as location – rural-urban differences – and how these shape people’s lived experiences in the continent. This literature from 1990 to 2010, in the context of our analysis, does not employ intersectionality theory but uses terms such as connects, interconnections, links pointing to intersections, multiple disadvantages, relationships between, nexus, and linkages in their discussions and analysis. (See for example, [Amadiume, 1987](#); [Assie-Lumumba, 1997](#); [Beoku-Betts, 1976/2005](#); [Chilisa and Ntseane, 2010](#); [Ogundipe-Leslie, 1994](#); [Okech, 2020](#); [Oyewumi, 1997](#); and [Tamale, 2000](#)).

Other existing works, largely by feminist scholars, have debated about whether western-originating theories and educational curricula can address the complex realities facing the people of Africa, and especially Black African women. (See [Atanga, 2013](#); [Gaidzanwa, 2013](#); [Mikell, 1997](#); [Mama, 2011](#); [Sudarkas, 1986](#)). While these works reveal their analytic approach of intersectionality, [Nchanji et al. \(2018\)](#) in their commentary observed that they see potential in research naming an intersectionality approach in Africa. Gender theorists, [Nchanji et al. \(2018\)](#) noted that the sociopolitical, economic and cultural environment and the diversity of identities and social categories in Africa make intersectionality research analysis relevant, while bearing in mind that Africa displays commonalities and uniqueness across the 54 countries ([Tshabangu, 2018](#)). Black African knowledges should not become subsumed by western theories. The application of intersectionality in Africa should include the specificities of contexts to minimize the risk of epistemological colonization.

In the last 10 years, Black African scholars in Africa and in the Global North started to apply and name intersectionality theory in studies on food, water and energy ([Ahmed and Mzimela, 2018](#)), land issues ([Akanle, 2017](#)); domestic work ([Maqubela, 2016](#)); crime and incarceration ([Agboola and Rabe, 2018](#)); healthcare ([Marindo, 2017](#)); race and racism in education ([Moagi, 2016](#)); and refugee experiences ([Yacob-Haliso, 2016](#)). In 2017, the journal *Agenda* dedicated Volume 31, Issue 1, to the theme of “Considering Intersectionality in Africa.” This is significant because it was the first time that there was a call for papers specifically employing and naming intersectionality in the African context. Following the call, twelve of the fifteen articles in the issue were by authors based in South Africa. Five of the articles were sole-authored or co-authored by Black Africans. While this outcome was probable, given that the editors of the issue, [Meer and Muller \(2017\)](#), are based in South Africa, it was also a lost opportunity to capture intersectionality’s theoretical nuances, and research reports from other African countries, and particularly those by Black African scholars.

Our analysis focused on applications of intersectionality theory in education in African contexts. We conducted systematic searches of the EBSCO, ERIC Education, and Google Scholar databases through the University of British Columbia library using the terms “education,” “Africa,” and “intersectionality.” We searched these descriptors in the abstracts, titles of published articles and keywords. The initial search yielded 548 articles. We then cross-checked the results for peer reviewed articles, those published between 1990 and 2020 and those written in the English language. We excluded books, book chapters, book reviews, dissertations and reports. We narrowed the search further by geographical location (sub-Saharan Africa), articles authored by Black African scholars located in Africa, and a focus on education, which yielded 33 articles. A close reading of the 33 abstracts revealed that only 13 articles used “intersectionality” as an analytical framework. Next, we read the full text version of the 13 articles paying attention to the inclusion and exclusion criteria. We found that only two of the articles focused on education and were written by Black African scholars who were based in Africa ([Moagi, 2016](#); and [Msibi, 2019](#)). The other 11 articles, although written by Black African scholars, were in the field of Gender Studies, Health Sciences, and Political Science. We also conducted a similar search using the African Journals Online database, which is an open access platform of African-published scholarly journals, and found the same two articles. The two empirical articles showed that intersectionality is used as a framework to complicate and emphasize equity and social justice research and practice.

[Moagi \(2016\)](#) studied how gender and racial relations are played out by high school girls in one South African high school. Moagi employed intersectionality to understand links between class, ethnicity and gender in a school setting. The framework helped to create awareness among the high school girls about the structures that influence racial, ethnic and gender identity, and the ways these structures interact with learners from different races and ethnicities in the school setting. [Msibi \(2019\)](#) researched different ways that Black South African male teachers who engage in same-sex relations negotiate and manage their identities in school contexts characterized by heteronormativity (p.389). The use of intersectionality allowed for unpacking the ways in which multiple forces impact the negotiation and management of identity by the male teachers interviewed [Msibi \(2019\)](#) drew on [Connell’s \(2014\)](#)

works on decolonization of thought, research and application to argue for a localized approach that would allow for understanding of ways in which different identity markers shape the teachers' experiences, and not assume homogeneity across contexts.

It is evident that intersectionality theory is applied to educational research in Africa. However, an accurate scope of its use in education remains to be determined precisely due to our inability to access journal articles from Africa through library sources, as referred to earlier. Many journals were based in South Africa, and much of the scholarship was by White South Africans. Other studies and literature were authored or co-authored with non-Black scholars from the Global North; so they did not meet our criteria for analysis of scholarship by Black Africans based on the continent of Africa.

In sum, the challenges facing the education sector in Africa are best understood and addressed in that context, although there are similar issues in other parts of the world. Applying intersectionality as an analytical framework necessitates that attention be paid to the historical, social, political and economic contexts of the different countries within the continent. It also requires a critical examination of the western-informed methodological approaches privileged in research.

Conclusion

To conclude, we have attempted to give an overview of the historical roots and uses of intersectionality in education with a focus on the Caribbean and Africa as much as possible, given the limitations of textual accessibility mentioned above. While some racialized scholars argue that intersectionality has strayed far from its original intent (Mugabo, 2022) others argue that it has become a hegemonic concept (Galindo, 2011) especially in gender studies. Nonetheless educational researchers, mostly in the Global North, are conceptualizing the innovative possibilities that the theory affords. Intersectional analyses can allow researchers to examine and understand the complex ways people experience systemic marginalization and oppression in various cultures and deepen analyses of educational processes.

Theories are culturally based, and shaped in particular contexts. Wiredu (1996) argues for methodological decolonization. Most research methods are steeped in ways of thinking and knowing that were not developed by people of African descent and therefore do not fully speak to their lived experiences. Although a Black feminist concept, any educational applications in Caribbean and African countries, which are majority Black regions, would have to take into consideration the understandings of how African Indigeneity, race, racialization, imperialism and colonialism are at work in these contexts and their interrelatedness with, for example, class, color, and ethnicity as well as other social identities. We leave this task to our colleagues in these regions, lest intersectionality become yet another colonizing discourse.

References

- Agboola, C., Rabe, M., 2018. Intersectionality and crime: reflections from female ex-inmates in South Africa. *Acta Criminol.: South. Afric. J. Criminol.* 31 (1), 1–18.
- Ahmed, F., Mzimela, J., 2018. An intersectional lens on the food-energy-water Nexus: on rural women's livelihood vulnerability reflection from Ndwedwe-Cibane, KwaZulu-Natal. *Afr. Insight* 48 (3), 73–87.
- Akanle, O., 2017. The intersectionalities of land grabs in Nigeria: engaging the new scramble for African lands. *Afr. Insight* 46 (4), 41–53.
- Amadiume, I., 1987. *Male Daughters, Female Husbands: Gender and Sex in an African Society*. Zed Books.
- Anderson, S., Cole, S., 2017. The relationship between adolescent aggression, parenting styles and family structure in urban inner-city Jamaica. *J. Educ. Devel. Caribbean* 16 (2), 1–19.
- Annamma, S.A., Connor, D., Ferri, B., 2013. Dis/ability critical race studies (DisCrit): theorizing at the intersections of race and dis/ability. *Race Ethn. Educ.* 16 (1), 1–31.
- Assié-Lumumba, N.T., 1997. Educating Africa's girls and women: a conceptual and historical analysis of gender inequality. In: Imam, A.M., Mama, A., Sow, F. (Eds.), *Engendering African Social Sciences*. CODESRIA, pp. 297–316.
- Atanga, L.L., 2013. African feminism? In: Atanga, L.L., Ellece, S.E., Lithosseliti, L., Sunderland, J. (Eds.), *Gender and Language in Sub-saharan Africa: Tradition, Struggle and Change*. John Benjamins, pp. 301–314.
- Bailey, A., Ngwenyama, O., 2012. Integration and Communication in the Diaspora: Intersectionality and Community Facilitation through ICTs. *AMCIS 2012 Proceedings*. <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1526&context=amcis2012>.
- Beoku-Betts, J., 1976/2005. Western perceptions of African women in the 19th and early 20th century. In: Cromwell, A. (Ed.), *Readings in Gender in Africa*. Indiana University Press, pp. 20–31.
- Bilge, S., Denis, A., 2010. Introduction: women, intersectionality and diasporas. *J. Intercult. Stud.* 31 (1), 1–8.
- Bowleg, L., 2008. When Black + lesbian + woman =/ Black lesbian woman: the methodological challenges of qualitative and quantitative intersectionality research. *Sex. Roles* 59, 312–325.
- Chilisa, B., Ntseane, G., 2010. Resisting dominant discourses: implications of Indigenous, African feminist theory and methods for gender and education research. *Gend. Educ.* 22 (6), 617–632.
- Cho, S., Crenshaw, K.W., McCall, L., 2013. Toward a field of intersectionality studies: theory, applications, and praxis. *Signs: J. Women Cult. Soc.* 38 (4), 787–810.
- Collins, P.H., 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge.
- Collins, P.H., 2015. Intersectionality's definitional dilemmas. *Annu. Rev. Sociol.* 41, 1–20. <https://doi.org/10.1146/annurev-soc-073014-112142>.
- Collins, P.H., 2017. On violence, intersectionality and transversal politics. *Ethn. Racial Stud.* 40 (9), 1460–1473.
- Collins, P.H., Bilge, S., 2016. *Intersectionality*. John Wiley.
- Collins, P.H., Chepp, V., 2013. Intersectionality. In: Waylen, G., Celis, K., Kantola, J., Weldon, S.L. (Eds.), *The Oxford Handbook of Gender and Politics*. Oxford University Press.
- Combahee River Collective, 1977/1982. The Combahee River collective. "A Black feminist statement.". In: Hull, G., Scott, P., Smith, B. (Eds.), *All the Women Are White, All the Men Are Black but Some of Us Are Brave*. Feminist Press.
- Connell, R., 2014. Using southern theory: decolonizing social thought in theory, research and application. *Plan. Theory* 13 (2), 210–223.
- Covarrubias, A., Velez, V., 2013. Critical race quantitative intersectionality: an anti-racist research paradigm that refuses to "let the numbers speak for themselves.". In: Lynn, M., Dixon, A.D. (Eds.), *Handbook of Critical Race Theory in Education*. Routledge, pp. 270–285.

- Grenshaw, K., 1989. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *Univ. Chic. Leg. Forum.* 139–167.
- Grenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 1241–1299. <https://doi.org/10.2307/1229039>.
- Davies, C.B., 2007. *Left of Karl Marx: The Political Life of Black Communist Claudia Jones*. Duke University Press.
- Davis, A., 1981. *Women, Race and Class*. Random House.
- Dill, B.T., Kohlman, M.H., 2012. Intersectionality: a transformative paradigm in feminist theory and social justice. In: Hesse-Biber, S.N. (Ed.), *Handbook of Feminist Research: Theory and Praxis*. Sage Publications, pp. 154–174. <https://doi.org/10.4135/9781483384740>.
- Douglas, J., 2019. Black women's activism and organisation in public health-struggles and strategies for better health and wellbeing. *Caribb. Rev. Gen. Stud.* 13, 51–68.
- Ellis, E.G., 2018. Gender complexities in Jamaican secondary education: class culture and the 'elephant' in the room. *Gen. Educ.* 30 (8), 1019–1031.
- Evans, E., 2016. Intersectionality as feminist praxis in the UK. In: *Women's Studies International Forum*, 59. Pergamon, pp. 67–75.
- Gaidzanwa, R.B., 2013. African Feminism, pp. 1–10. <http://www.zmdt.org.za/wp-content/uploads/2018/10/Africa-Feminism-Rudo-Gaidzanwa.pdf>.
- Galindo, M., 2011. Lecture "El paradigma de la interseccionalidad en América Latina" at the Lateinamerika-Institut of the Freie Universität Berlin. November 17, 2011, cited in.
- Galvan, A., 2015. Soliciting Performance, Hiding Bias: Whiteness and Librarianship. <http://www.inthelibrarywiththeleadpipe.org/2015/soliciting-performance-hiding-bias-whiteness-and-librarianship/>.
- Garvey, M., 1983. The Marcus Garvey and Universal Negro Improvement Association Papers, Volume XI: The Caribbean Diaspora, 1910–1920, vol 11. Duke University Press.
- Gillborn, D., Rollock, N., Warrington, P., Demack, S., 2016. Race, Racism and Education: Inequality, Resilience and Reform in Policy and Practice. http://soc-for-ed-studies.org.uk/documents/GillbornD-et-al_Race-Racism-and-Education.pdf.
- Gillborn, D., Warrington, P., Demack, S., 2018. QuantCrit: education, policy, "Big Data" and principles for a critical race theory of statistics. *Race Ethn. Educ.* 21 (2), 158–179.
- Grant, C.A., Zwi, E., 2017. *Intersectionality and Education*. Oxford University Press.
- Harris, A., Leonardo, Z., 2018. Intersectionality, race-gender subordination, and education. *Rev. Res. Educ.* 42 (1), 1–27.
- Hartsock, N., 1983. *Money, Sex and Power: Toward a Feminist Historical Materialism*. Longman.
- Henry, A., 2015. 'We especially welcome applications from members of visible minority groups': reflections on race, gender and life at three universities. *Race Ethn. Educ.* 18 (5), 589–610.
- Iftikar, J.S., Museus, S.D., 2018. On the utility of Asian critical (AsianCrit) theory in the field of education. *Int. J. Qual. Stud. Educ.* 31 (10), 935–949.
- James, C., 2021. More than brains, education and hard work. The aspirations and career trajectories of two African Canadian male youth. *Colour matters: Essays on the experiences, education, and pursuits of Black youth*. University of Toronto Press, pp. 150–168.
- Jang, S., 2018. The implications of intersectionality on Southeast Asian female students' educational outcomes in the United States: a critical quantitative intersectionality analysis. *Am. Educ. Res. J.* 55 (6), 1268–1306. <https://doi.org/10.3102/0002831218777225>.
- Law, I., Tate, S.A., 2015. *Caribbean Racisms: Connections and Complexities in the Racialization of the Caribbean Region*. University of Leeds.
- López, N., Erwin, C., Binder, M., Chavez, M.J., 2018. Making the invisible visible: advancing quantitative methods in higher education using critical race theory and intersectionality. *Race Ethn. Educ.* 21 (2), 180–207.
- Mama, A., 2011. What does it mean to do feminist research in African contexts? *Fem. Rev.* e4–e20. <https://nigs.paginas.ufsc.br/files/2017/07/fr201122a-AMINA-MAMA-Feminist-Research-in-Africa.pdf>.
- Maqubela, L.N., 2016. Changing motherhood and the shifting social networks-of-care within Black families in the post-Apartheid South Africa. *Gender Behav.* 14 (2), 7225–7234.
- Marindo, R., 2017. Gendered epidemics and systems of power in Africa: a feminist perspective on public health governance. *Afr. Dev.* 42 (1), 199–219.
- Martinez, M., 2019. Latina educational trajectories to the academy. *J. Transform. Lead. Policy Stud.* 8 (1), 30–58.
- McCall, L., 2005. The complexity of intersectionality. *Signs. Journal of women in culture and society* 30 (3), 1771–1800.
- Meer, T., Müller, A., 2017. Considering intersectionality in Africa. *Agenda* 31 (1), 3–4.
- Mikell, G. (Ed.), 1997. *African Feminism: The Politics of Survival in Sub-saharan Africa*. University of Pennsylvania Press.
- Moagi, L.A., 2016. Confronting the intersectionality of gender and race: a case study of Black girls in two multiracial high schools in post-apartheid South Africa. *Gender Behav.* 14 (2), 7321–7328.
- Msibi, T., 2019. Passing through professionalism: South African Black male teachers and same-sex desire. *Sex. Educ.* 19 (4), 389–405.
- Mugabo, D., 2022. Intersectionality in Blackface: when post-racial nationalism meets Black feminism. In: Ibrahim, A., Kitossa, T., Smith, M., Wright, H. (Eds.), *The Nuances of Blackness*. University of Toronto Press.
- Nash, J.C., 2017. Intersectionality and its discontents. *Am. Q.* 69 (1), 117–129.
- Nchanji, E., Nkengla, L., Ajambo, S., 2018. Complexity of Intersectionality: Relevance to African Research. <https://dev.ckm.ili.org/cgiar-gender/grit-intersectionality-african-research/>.
- Nichols, S., Stahl, G., 2019. Intersectionality in higher education research: a systematic literature review. *High. Educ. Res. Dev.* 38 (6), 1255–1268.
- Núñez, A.M., 2014. Employing multilevel intersectionality in educational research: Latino identities, contexts, and college access. *Educ. Res.* 43 (2), 85–92.
- Nyangweso, M.A., 2017. Intersectionality, education, and advocacy against sexual violence. *J. Fem. Stud. Relig.* 33 (1), 180–182. <http://www.jstor.org/stable/10.2979/jfemistudrel.33.1.24>.
- Ogundipe-Leslie, M., 1994. *Re-creating Ourselves: African Women and Critical Transformations*. African World Press.
- Okech, A., 2020. African feminist epistemic communities and decoloniality. *Crit. African Stud.* 12 (3), 313–329.
- Olson, H.A., 2004. The ubiquitous hierarchy: an army to overcome the threat of a mob. *Libr. Trends* 52 (3), 604–616.
- Oyewumi, O., 1997. *The Invention of Women: Making an African Sense of Western Gender Discourses*. University of Minnesota Press.
- Richards, M., 2013. A critical multicultural analysis of children's literature. *J. Devel. Educ. Caribbean* (1/2), 54–83.
- Robert, S.A., Yu, M., 2018. Intersectionality in transnational education policy research. *Rev. Res. Educ.* 42 (1), 93–121.
- Rodney, W., 1973/1981. *How Europe Underdeveloped Africa*. Howard University Press.
- Roth, J., 2013. Entangled Inequalities as Intersectionalities: Towards an Epistemic Sensibilization. *Desigualdades.net. Research Network on Interdependent Inequalities in Latin America*.
- Rouhani, S., 2014. Intersectionality-informed quantitative research: a primer. *Am. J. Publ. Health* 103 (6), 1082–1089.
- Sibbett, L., October 27, 2020. Intersectionality in U.S. Educational Research. *Oxford Research Encyclopedia of Education*. <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-403>.
- Smith, P., Warican, Kumi-Yeboah, A., Cheema, J., Allyene, M., 2020. Disrupting (Mis)representation in the literacy achievement of underperforming youth. *J. Educ. Devel. Caribbean* 18 (2), 111–165.
- Stewart, S., 2019. Navigating the academy in the post-diaspora: Afro-Caribbean feminism and the intellectual and emotional labour needed to transgress. *Caribbean Rev. Gender Stud.* 13, 147–172.
- Sudarkasa, N., 1986. The status of women in Indigenous African societies. *Fem. Stud.* 12 (1), 91–103.
- Sullivan, E., Larke, P.J., Webb-Hasan, G., 2010. Using critical policy and critical race theory to examine Texas' school disciplinary policies. *Race, Gender & Class* 72–87.
- Tamale, S., 2000. *When Hens Begin to Crow: Gender and Parliamentary Politics in Uganda*. Fountain Books.
- Thomas, D., 2011. *Excetional Violence: Embodied Citizenship in Transnational Jamaica*. Duke University Press.
- Tierney, R.J., 2018. Toward a model of global meaning making. *J. Literacy Res.* 50 (4), 397–422.

- Tormos, F., 2017. Intersectional solidarity. *Politics, Groups, and Identities* 5 (4), 707–720.
- Trotz, D.A. (Ed.), 2020. *The Point Is to Change the World: Selected Writings by Andaiye*. Pluto Books.
- Tshabangu, I., 2018. The intersectionality of educational inequalities and child poverty in Africa: a deconstruction. *Educ. Res. Pol. Pract.* 17 (1), 69–82.
- United Nations General Assembly, 2013. Resolution 68/237 Proclamation of the International Decade for People of African Descent. <http://www.ny.un.org/doc/UNDOC/GEN/N13/453/67/PDF/N1345367.pdf?OpenElement>.
- Wiredu, K., 1996. *Cultural Universals and Particulars: An African Perspective*. Indiana University Press.
- Yacob-Haliso, O., 2016. Intersectionality and durable solutions for refugee women in Africa. *J. Peacebuilding Dev.* 11 (3), 53–67.
- Yancy, G., 2012. *Look, a White!: Philosophical Essays on Whiteness*. Temple University Press.
- Zinn, M.B., Dill, B.T., 2016. Theorizing difference from multiracial feminism. In: *Race, Gender and Class*. Routledge, pp. 76–82.

Decolonizing research with girls with disabilities in the Global South

Xuan Thuy Nguyen, Institute of Interdisciplinary Studies, Carleton University, Ottawa, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction: contested discourses over “girls with disabilities”	124
Contested accounts in disability studies	125
Research as a decolonial praxis: experiences from the TDKRA project	126
Findings	127
Understanding complex experiences of girls with disabilities from the Global South	127
Strengths and capacities	128
Discussion	130
Conclusion	131
Acknowledgment	132
References	132

Introduction: contested discourses over “girls with disabilities”

Globally, girls with disabilities have taken up a significant part of the disability populations. However, they have been invisible within education. While empirical evidence regarding girls with disabilities remain to be scarce and inconsistent, data from the World Report on Disability shows that female with disabilities reported 41.7% primary school completion, compared to 52.9% of female without disabilities (World Health Organization and World Bank, 2011). Their invisible girlhoods have been intensified within the context of the Global South. Stienstra (2015) observes:

We know little about the intersections of being a child, a girl, and disabled, let alone what it is like in the global South, being Indigenous, or coming from a racialized minority (p. 54).

International interventions on the rights of women and girls with disabilities in recent years have raised some awareness on their barriers to education (United Nations Girls' Education Initiative and Leonard Cheshire Disability, 2017), lack of access to employment, health care, and sexual and reproductive rights (Office of the United Nations High Commissioner for Human Rights, 2017; Human Rights Watch, 2020), and discrimination against their disability and gender (Shah and Bradbury-Jones, 2018). However, these global approaches tend to sideline epistemological and political issues regarding who these girls are, how we come to know them, and what it means to include them in humanitarian and human rights discourses.

Questions regarding how we come to know of “girls with disabilities” are epistemological. Despite the inclusion of “girls with disabilities” within the United Nation's rights-based agenda (United Nations, 2006), these girls are usually silenced within human rights and humanitarian campaigns. There has been little question regarding how to engage these girls in expressing their views, raising their consciousness, and fostering their political agency. Disabled girls are usually depicted as both hyper-visible and invisible within the transnational spaces such as social media and humanitarian discourses (Erevelles and Nguyen, 2016). Specifically, while some disabled girls have been “included” into the public spaces, thus become visible, and sometimes hyper-visible within these spaces, they have been invisible from political activism and other participatory youth-lead agendas (Nguyen et al., 2021b).

The epistemological and political endeavor to disrupt normative assumptions of what it means to be “a girl” has been critical for girlhood studies. Scholars of girlhood studies are “interested in fragmenting the idea of ‘the girl’ ... and draws attention to differences among them” (Gonick and Cannon, 2014, p. 1). They highlight the heterogeneity of girlhoods, including how these girls' identifications and social locations vary across a vast array of race, gender, sexuality, ethnicity, age, and nationality. However, girlhood scholars tend to exclude experiences of disability; and where included, “disability” is primarily discussed as a bio-medical category without interrogating its socio-political implications. Thus, even within this critical scholarship, disability is still seen as a peculiar *thing* attached to their otherwise “normal” able-bodied bodies.

In a report on behalf of the Women Refugee's Commission for the journal of Girlhood Studies, Pearce et al. (2016) posit that humanitarian programs should see these girls as “girl first” (p. 129) rather than their disabilities. They explain that these girls wanted to be identified as daughters, sisters, and friends—that is, as individuals with performative gendered roles and activities. While it is critical to understand how these girls want to be represented, their lived experiences and embodied social relations relative to the contexts where they came from are generally unknown. Here, disabled girls stand in a sharp contrast to the Western framings of girlhood as agentic, independent, and capable of creating changes (Erevelles and Mutua, 2005). This narrative of disabled girlhood justifies ableism - that is, by valuing able-bodiedness and able-mindedness as the descriptive standard of girlhood, this approach to girlhood has justified discrimination against disabled bodies and minds by failing to recognize diverse girls' experiences while

reinforcing ableist assumption about disability. Interestingly, while disabled girls—as a social group—are culturally and politically diverse, this global discourse of “girls with disabilities” tends to frame these girls as a homogenized social grouping in ways that fail to recognize their complex social situations. Indeed, audience in the Global North comes to know about disabled girls in the Global South “from the center” (Connell, 2007, p. 45) where their individualized narratives are usually served as a rhetorical tool for objectifying their vulnerabilities. By erasing the cultural, historical, and epistemological differences which constitute their lived experiences, such discourses have reproduced the “coloniality of being” (Maldonado-Torres, 2007), where disabled girls’ differences are classified, hierarchized, and objectified from the Western perspective.

The “coloniality of being” refers to the consequences of colonization on the lived experience of colonized people and their language in ways that unveil the dark side of modernity (Maldonado-Torres, 2007). Coloniality is different from colonialism in that coloniality refers to long patterns of power that emerged as a consequence of colonialism, but extends well beyond the political, economic, and administrative structures of colonialism. Coloniality defines culture, labor, intersubjective relations, as well as knowledge production through the constructions of categories such as race and gender (Maldonado-Torres, 2007; see also Mignolo, 2018). Thus, research about disabled girls in the Global South can become a form of “coloniality of power” (Quijano, 2007, p. 171) when it fails to elucidate how Black, racialized people, women, girls, and disabled children have been historically constituted as colonial difference, and how actions taken by these groups can serve as a radical point of departure from the epistemic and political violence which marks them invisible.

This approach reinforces epistemic injustice against disabled bodies from the global South, where knowledge about them was “collected, classified, and then represented back to the West” (Smith, 1999, p. 1) through the research practices which reinforce the “coloniality of power.” The intersection between colonial and ableist ideologies creates categories of normative girlhood as able-bodied while erasing difference from the South.

Contested accounts in disability studies

In contrast to humanitarian and academic discourses on girlhoods and rights, political interventions of disability studies manifestos are critical for disrupting able-bodied and able-minded ideologies that constitute disability and difference (Titchkosky, 2011; Garland-Thomson, 2018). These accounts challenge modernist constructions of disability as an individualized problem constituted by the able-bodied norms (Davis, 1995) and create alternative ways of re-imagining disability in cultural, political, and institutional discourses and practices. However, as Jaffee and John (2018) observe, the genealogies and narratives of the field have continued to center canonical texts. As such, disability studies has preserved the very structure it aims to challenge. There is little interest in these bodies of scholarship to learn from or engage with forms of knowledge from the Global South, especially within accounts not traditionally seen as “disability studies.” In fact, the uncritical transfer of Western disability studies into the South has been problematic because it fails to engage with distinctive historical and conceptual frameworks in relation to theories and praxis from the South (Anand, 2015; Mehrotra, 2020; Staples, 2020). By assuming that North American and Eurocentric approaches of disability studies are universal and generally applicable to Southern spaces through Global North interventions (Swartz, 2018), Western disability studies reinforces unequal power relationships between the Global North and South.

Within these contested power relations, disability studies has become unsettling. As a result, questions regarding the scholarship, more specifically, who produces it and from which particular spaces and social location(s), where it is read, and by whom are inherently political (Nguyen, 2021). Researchers from the Global North tend to interpret disability experiences in the South from their locations. This is problematic because it fails to engage with such social experiences from perspectives of disabled people from the Global South. This Eurocentric approach in disability studies reinforces epistemic violence on Southern epistemologies (Grech, 2015). As feminist and post-colonial disability studies scholar Kim (2017) argues, the intellectual’s social location and the politics of enunciation is inherently political. At issue, she asks: “Who speaks of these differences to whom, for what purposes, and in what location?” (p. 230).

Coming from distinctive spaces in the Global South, decolonial disability studies (decolonial DS) is a collective body of knowledge that emerges from the South as a way of reclaiming its existence beyond Western theories and epistemologies (Nguyen, forthcoming). It seeks to critically engage with Western disability studies while at the same time resisting its hegemonic structures that constitutes its difference. Furthermore, decolonial DS offers epistemic and material resources for interrogating individualized concepts of ability and disability, as not natural endowment of human essence, but rather, as consequences of the colonial appropriation of extraction of the Other’s productivity and the disablement of the majority world (Dirth and Adams, 2019). By embodying intellectual projects from the Global South, decolonial DS enables disability scholars to re-visit the limits of Western disability studies from the spaces of bodies and minds historically colonized by Western knowledge and power.

Decolonial DS is critical for research with girls with disabilities in the Global South as it provides an alternative approach to knowledge production in ways that reclaim forms of knowledge that have been absent from Western ideologies. This approach not only rejects the ableist methods to studying “girls with disabilities” as framed in Western social sciences research, but also, its colonial impositions in transnational contexts. Thus, a decolonial DS approach tackles the gaps in research on girls with disabilities by first, addressing the absence of knowledge on girls with disabilities in the Global South and highlighting the epistemologies of disabled girls in transnational spaces; secondly, it challenges the ableist and colonial approaches in research on girls with disabilities by engaging with concrete experiences of these girls in Southern contexts; and finally, it creates decolonial spaces for disabled girls in the Global South to build their leadership and identities in ways that reclaim their existence. In so doing, decolonial DS is a form of enunciation that opens up an opportunity for transforming the Western knowledge system from the standpoints of disabled girls in the Global South.

This chapter aims to reshape the ways in which research *on* girls with disabilities has been conducted in the Global South and proposes a decolonial approach to working *with* these diverse groups. I use the term "girls with disabilities" and "disabled girls" interchangeably to reflect diverse social locations of girls with impairments in the Global South and to respect how these girls prefer to be named from their locations. Drawing on a 4-year research project, Transforming Disability Knowledge Research and Activism (TDKRA), funded by the Social Sciences and Humanities Research Council of Canada (2016–2020) in Vietnam, I will examine how a decolonial approach to research with girls with disabilities can be useful in empowering these girls to re-imagine their spaces, social locations, and express their perspectives on disability through multiple forms of representation. I ask a number of questions: What does it mean to engage girls with disabilities in the Global South? What approaches and theories can be used that consider the lived experiences of these girls? How can we meaningfully draw communities in the Global South in this conversation while resisting the coloniality of power that has driven our research agenda?

Girls with disabilities in Vietnam are situated within a specific context of rapid social change. However, the deep-seated roots of exclusion of these girls could be traced back to the French constructions of Indigenous knowledge systems, where disabled girls were seen as in need of being civilized and normalized (Nguyen, 2015). While the colonial and imperialist regimes were overturned in Vietnam's transition into socialism, colonial practices are still inherent within the nation's constructions of "the race" and its "ideal" populations, where disabled girls are seen as a contrast to the body politics of the nation-state (Nguyen and Stienstra, 2021). At the same time, their situations have been materially shaped by historical war with the United States, which exposed many families to Agent Orange to high level of traumas and suffering because of the intergenerational impacts of the dioxin (Ngo et al., 2013). The socio-economic reforms seen within Vietnam's Doi Moi policy, decentralization and privatization of services, and ongoing practices of state control and population surveillance further reinforced their exclusion. Despite rights-based discourse fostered by the United Nations and non-governmental organizations (NGOs), there is ongoing exclusion and marginalization of women and girls with disabilities within the neo-liberal and neo-colonial context of development.

In what follows, I will use my research experience with girls with disabilities within the TDKRA to demonstrate how a decolonial approach to research with girls with disabilities can be conducted within contested power relations in the Global South. It offers an alternative approach to engaging with girls with disabilities in socio-political spaces that open up new possibilities for re-imagining difference from disabled girls' standpoints. I argue that a decolonial approach to research with girls with disabilities is critical for deepening our transnational conversations in ways that create more inclusive spaces for these girls to tell their stories and to resist exclusion. I posit that this conversation is critical for education research because it creates a meaningful approach for decolonizing hegemonic forms of Western knowledge and invites new ways of framing social justice in the Global South through decolonial DS.

Research as a decolonial praxis: experiences from the TDKRA project

Transforming Disability Knowledge, Research and Activism is a collaborative research project that aims to engage with disabled women and girls in Vietnam. The research is rooted in its purpose of unsettling research and activism by using research as a form of activism that challenges the gap in knowledge with and about girls with disabilities in the Global South (Nguyen et al., 2019). To engage these girls, we used a participatory arts-based approach along with traditional methods such as in-depth interviews and focus groups. While historically visual approaches such as photography were used as a colonial tool to objectify colonized subjects (Rose, 2001), participatory visual methods (PVM) offer a way of shifting traditional research relations through the creation of visual artifacts that are produced by historically colonized and oppressed groups in order to raise their voices about issues that concern them, and in so doing, counter dominant narratives embedded within male, heterosexual, able-bodied, and cis-gendered perspectives (Mitchell, 2011). In the TDKRA project, we worked with girls and women with disabilities in Vietnam using different PVM methods such as drawing, cellphilm (a short video made from cellphone), and photography to engage them in co-producing knowledge about their lived experiences and setting a stage for their activism.

This project was implemented in three communities, including Bac Tu Liem (Hanoi—North of Vietnam), A Luoi (Thua Thien Hue—Center of Vietnam), and Binh Thuy and Ninh Kieu (Can Tho — South of Vietnam). These communities are distinct in terms of their socio-economic, ethnic, and political locations. We engaged 85 girls and women with disabilities in three disadvantaged communities in Vietnam. The girls (10–21 years old) played the role of knowledge producers while the women (20–53 years old) participated in the research training as co-facilitators and mentors to the girls. The girls were also key informants in the project. They worked with the women to share stories about their lives, construct consciousness about their experiences, and develop strategies for connecting and mobilizing their knowledge as a form of collective activism. The majority of our participants came from very disadvantaged situations.

The decolonial aspects of TDKRA were manifested in different stages of the project planning, implementation, writing, and engaging the local communities. Specifically, we sought to decolonize the hegemonic structures of Western research by involving and partnering with Disabled People's Organizations (DPOs) and their communities in all stages of the research. Girls and women with disabilities were recruited through the local DPOs to ensure that DPOs directly involved in the research process as a part of the decolonizing methodologies (Smith, 1999). We collaborated with DPOs through the development of the grant proposal and collaborative agreements such as funding and fieldwork planning, ensuring that the local partners can meaningfully participate and have a voice in the planning and implementation process. While the collaboration was at times challenging, our engagement with the girls and women with disabilities and their organizations was central to relationship building and empowerment of the local communities.

Over the course of 4 years, we collected 55 interviews, 115 drawings, 82 photo images, 111 poster-narratives, 293 audio recordings of the workshop sessions; and 17 cellphilms produced by the girls the workshop sessions. We also co-produced a documentary film with the participation of the girls and women with disabilities across all three communities. We used critical engagement, deep listening, trustworthiness, transparency, and reflexivity as the main criteria for our qualitative analysis (Denzin, 2009). From a critical approach to qualitative inquiry, data is not neutral; rather, it is constructed through “complex interpretive processes [that] shape how evidence is turned into data, and how data, in turn, are coded, categorized, labeled, and assembled into data banks” (Chamaz, as cited in Denzin, 2009, p. 146). This reflects the need to see data as socially constructed and that the context of research is critical for interpreting its outcomes.

Data was transcribed and translated into English from Vietnamese and local dialects (Pa Co and Ta Oi). It was coded and categorized using Nvivo software. We tested out a set of codes which had been used in an earlier project on the educational rights of girls with disabilities in Vietnam (Nguyen et al., 2015) by piloting a number of interviews and visual data. We confirmed and collapsed the codes and sub-codes by considering the significance of empirical evidence. We then identified emerging themes through the data. Additionally, the interviews, captions and visual narratives (drawings, photos, cellphilms) were further analyzed using textual and discourse analysis to make sense of their meaning and implications. The context in which the participants produced knowledge in relation to their class, gender, ethnicities, disabilities, and impairments is critical for the politics of interpretation. We then triangulated the data with our fieldnotes to evaluate the trustworthiness of different types of data.

The next section draws on the results from the aforementioned interviews and visuals (i.e., drawings and cellphilms) to showcase the ways in which participatory visual research, informed by decolonizing methods, can speak to and engage with the diverse perspectives of girls with disabilities in the Global South.

Findings

Understanding complex experiences of girls with disabilities from the Global South

The findings of TDKRA reflect the global dynamics of impairment that have been embodied within the participants’ personal experiences. For example, some disabled girls and young women in the project experienced intergenerational traumas due to being exposed to Agent Orange—one of the most toxic herbicides sprayed by American forces during the US- Vietnam War. They perceived disability as an illness that could be passed on their younger generations. As GWD 11 shared her story: “My brother is disabled because of inheritance of poison from father to child. My father too, inheritance from mother to child” (Fieldwork in A Luoi, 2018).

As a consequence of societal prejudice against their disabilities, some participants seemed to have internalized ableism by assuming that their disability is an illness that could pass on to their future generations. A young woman viewed disability as “a disease,” as manifested in the intergenerational impacts of Agent Orange: “[disability is] ... a disease, just like a normal person who has a catastrophe. There are very few inborn people with disabilities. The parents who suffered from Agent Orange may give birth to the children being paralyzed or being disabled in the leg and arm” (Fieldwork in A Luoi, 2018). At the same time, they shared experiences being discriminated in schools and public spaces. An ethnic girl with disabilities said that her friends used “unacceptable words” such as “crazy” when teasing her. Another ethnic girl shared her story in her minority dialect:

Facilitator: How do Kinh people treat you? Have they ever teased or discriminated against you?

Min: Yes.

Facilitator: What did they say about you?

(Interpreter): They said I was crazy.

Ethnic minority girls with disabilities experienced multiple oppressions due to institutional racism, ableism, and classism. A girl revealed that others treated her as “ugly”: “They spoke badly about me Like I was disabled, I was ugly” (GWD 10). Other girls expressed vulnerabilities when being looked down by their peers: “I feel hurt ... I feel ashamed” (GWD 11). A girl with epilepsy, That Tha, shared the same feeling: “I felt timid, sad. Sometimes I cried.” The intersections of multiple disadvantages constituted their experiences with vulnerabilities.

At the same time, the participants expressed their capacity to counter such oppressions. For example, they resisted ableism by disrupting what their peers said about them:

Facilitator: Is there anyone who speaks badly of you, tease you?

Moon: There are some older students who often said my eyes were like a devil.

Facilitator: When they said that, how did you feel?

Moon: I said: "Although my eyes are not beautiful, my soul is more beautiful than others."

When Moon first joined the project, she expressed feelings of being different and inferior to her able-bodied friends. Here it seems that Moon wanted to resist the ableist attitude by saying that "my soul is more beautiful than others." Her experience is multi-layered, and sometimes conflictual, suggesting that we must interpret the experiences of girls with disabilities through complex social context and relationships with their facilitators. In this case, Moon's resistance against ableism needs to be contextualized within her relationships with her facilitator, who is also a woman with disabilities. Their connections and shared experiences with ableism in school enabled her to frame her discourse while resisting ableist ideology on what it means to be "beautiful." The participants' collective experiences with exclusion and powerlessness were connected with one another in the space where they felt they are not the only one who experienced discrimination. They found the same stories told by their peers. As such, our workshop became a space for them to share their stories collectively, which stands in contrast to the individualized perspective of disability. They also identified positive experiences. For example, a girl who chose her nickname as Nho Mit Uot shared that "Because everyone finds my family poor, they treat me well." Be Cao, a visually impaired girl, talked about her relationship with her mother and how this transformed her sense of the self:

Be Cao: I used to think that I couldn't do anything because I am disabled. But I can earn for my daily living and support my mother.

Facilitator: Nothing we cannot do, remember that. Be strong.

Be Cao: Yes.

Facilitator: How did you feel when you were helped?

Be Cao: I felt so happy because she understood me.

The wording "we" and self-affirmation by the facilitator who is also a woman with disabilities is important and speaks to the contextual and social relations aspect of decolonial DS. It suggests that disabled girls and women can transform their situation through their local disability networks. The participants also expressed a strong interest in participating in social activities as fully capable members of society. Their desire to be seen and heard and to foster social changes within their communities can be further exemplified through their visual productions, which suggest their ways of seeing by capitalizing their strengths and capacities.

Strengths and capacities

A decolonial disability studies approach requires that we resist the tendency of Western research to homogenize these girls' experiences because this would perpetuate their representations as suffering bodies while abstracting their experiences from specific socio-political contexts. In other words, we must critically engage with the local knowledge by creating decolonial spaces for participants to shape their knowledge. The use of creative methods to amplify the voices and capacities of the participants is critical for

generating alternative forms of knowledge co-production in ways that disrupt colonialist, ableist, and male patriarchal perspectives on their dis/abilities.

In a cellphilm entitled “Don’t look down on children with disabilities!”, produced by her group, Be Cao chose to participate and re-imagine her experience from the powerlessness she witnessed when being mocked and made fun of by her peers. This two-minute cellphilm represents the experiences of blind students in school. After their brainstorming, the group assigned the roles of team members. Be Cao played the main character and Moon was the film director. The first scenario shows two disabled girls holding their hands to go to school. Be Cao was playing her everyday experience in navigating her way to school. Another girl with visual impairment, Bup Be, was shown to be walking in the front so that Be Cao could hold on to her shoulder to navigate the direction. The girls did not hold any walking canes in their road. In the next scenario, the film showed some sighted girls walking by, pointing at the disabled girls, and making fun of them. When Bup Be told Be Cao that someone was teasing them, she raised her voice clearly and loudly: “Why are you laughing at me? You shouldn’t act like that!” This “speaking back” moment was quite powerful as it disrupts ableist attitude on their difference and create a space for them to articulate their voice and exercise activism. In their message about the film: “Don’t look down upon children with disabilities! Please care for us!”, it seems that the group wanted to use their cellphilm as a symbolic representation of their collective activism (See also [Nguyen et al., 2021b](#)). That is, disabled students can react determinately and powerfully when they are given a tool to do so.

Talking about her gift to the community, Be Cao said that she wanted to use her voice to share her opinion with everyone. Instead of speaking directly, song lyrics are one way of delivering her messages. Her song, “Live like flowers,” provides some beautiful lyrics on how girls with disabilities wanted to transform their lives by desiring disability differently. Other girls used their drawings to express their desire to participate in the public spaces, design clothes, care for their families and friends, enjoy sports, and develop their leadership skills (see also [Nguyen et al., 2021a](#)) ([Fig. 1](#)).



Fig. 1 Drawing titled “Park,” created by That Tha during a workshop in Hanoi in 2018.

In her drawing, That Tha draws a picture titled “Park” to represent the visibility of a girl with disabilities in the public space. The picture shows a disabled girl wearing a red shirt seating in a wheel-chair in the park. She is surrounded by her friend, land, and trees. There is a football net behind her. Through this picture, it seems that That Tha wants to express her desire to be visible in the public space where she can participate in everyday activities. Another deaf girl, Little Cat, draws her school with friends. Her use of bright colors in the drawing suggests her optimism for participating in an inclusive school ([Fig. 2](#)).

Little Cat draws a two-storey building surrounded by green trees and blue sky. There are many students in the schoolyard and in the classrooms. At the center of the drawing are the images of two students—one of whom is her self-representation marked by her own signature. It seems that she viewed her relationship with friends and schooling as something vital to the schooling. As she talks about her picture using her sign language: “[this drawing shows that] I hang around and play with new friends. Going to school in Can Tho is fun.” Here, she re-imagines schools as a space for connections and building relationships. This reflects the centrality of social relationship in schooling in their lives of Deaf girls in the community. At the same time, this space is politically constructed because it has been re-imagined from the perspective of a Deaf girl. In so doing, her Deafness is made visible to the school community through her drawing. Thus, the use of art for oppressed groups such as Deaf girls to imagine themselves and their sense of belonging helps to unsettle the regulatory practices of schooling by allowing these girls to re-imagine schools from their own standpoints.

In another Community-Asset Mapping workshop organized by our team members, participants were invited to locate their local spaces, including their homes, markets, schools, communities, and their friends’ homes. Using enlarged printed maps of their own communities so that they can identify their location, the participants used Lego figures in order to build their community such as homes, bridges, football field. Being able to involve in these playful activities, participants were invited to re-image their spaces within their communities and create connections with their locations ([Fig. 3](#)).



Fig. 2 Drawing “Students in my school” by Little Cat during a workshop in Can Tho in 2018.



Fig. 3 A Community-Asset Mapping workshop in Can Tho in 2018.

The playfulness of this activity offers an important opportunity for the girls to re-imagine their local space in relationship to others. The participants were excited to locate their space where they feel belonging. The community then became the most important social network for participants to re-locate themselves. They were then invited to identify their “gifts” as a part of their strengths for creating a more inclusive community. The girls chose to depict their strengths as being able to take care of their family chores, take care for their family members, being able to participate in social activities, having artistic skills such as drawing, or supporting their peers. In so doing, the participants were able to re-evaluate their own values and resist ableist values about disability which many of them may have internalized.

Discussion

The findings from TDKRA show that girls with disabilities are diversely represented through their class, gender, ethnicity, age, and disability. Their experiences are complex and multi-faceted. This was reflected in their understanding of disability, which is not monolithic but instead influenced by a variety of factors including how they have been treated by others. For example, some girls internalized ableism because of attitudinal and institutional oppressions which position as inferior to others. In the context of Agent Orange, some disabled girls saw their disability as a disease that needs to be cured, while at the same time highlighting the lack of

access to health care and social services. This reflects the challenges of disability communities in the Global South to resist ableist ideologies and practices that have been deeply entrenched and naturalized within the context of the Global South due to colonial and imperialist impacts.

In fact, while discussions on disability in communities in the South have emerged over the last three decades, they continue to be driven by the charity and medical approaches. Little has been discussed in relation to socio-political forces, which created impairment on bodies (Connell, 2011). This was displayed through the ongoing and intergenerational impacts of Agent Orange on disabled women and girls, and future generations which solidified medical perceptions that disability is a problem with the body (Nguyen and Stienstra, 2021). This context illuminates the ways in which imperialism caused impairments to Indigenous bodies through forms of governance which are epistemically and politically governed by the colonial structures and ideologies which create categories of difference (See also, Meekosha and Soldatic, 2011). This governing structure resulted in their constructions of class, race, gender, ability and disability (read as illness, disease, or abnormal bodies)—marked as deviant from the normative category of human beings as defined from a Western gaze and its Eurocentric ideologies (Smith, 1999).

At the same time, participants also demonstrated their strengths and capacities to contribute to their families and communities through the structure of their social lives. Some girls felt included in schools because others understood their struggles and supported them. In their visual productions, they drew their images in relation to friends, teachers, parents, schools, the classrooms, the playgrounds, along with cultural surroundings such as home, land, and the environment as spaces in which they feel connected. They also recognized the importance of engaging and creating collective support networks such as organizations of people with disabilities.

At the center of this epistemology is the concept of relationality. According to Indigenous epistemologies, relationality means “all things exist in relatedness” (Tynan, 2021, p. 601). Relationality speaks to how we relate to other human and non-human entities through our interdependence and mutual obligations. It does not only shape our ways of knowing and being but also shapes our ethical practice. Relationality rejects the idea of Otherness inherent in the hierarchical practice of the coloniality of being which positions the Other outside of the matrix of relatedness. Instead, relationality is conceptualized as “... that which brings us into existence through practice and process” and which such existence “is premised on the practice of our relational reality” (p. 601).

The concept of relationality could be used to understand non-Western epistemologies when conducting research in relation to disability in the global South. For example, Nguyen and Stienstra (2021) used this concept to demonstrate how non-Western epistemologies could generate a different system of thoughts which respect and value difference. Drawing on Ubuntu epistemology, they argue that we are always in a relational cycle with everything on the Earth (Mucina, 2013). We are connected to one another despite our differences. In fact, we depend on our differences in order to survive and sustain ourselves. Thus, while the disabled girls experienced intersectional forms of oppression, they also expressed a desire to provide reciprocal care for their family members, friends, and communities. For example, a disabled girl said that she wanted to find a job to earn money so that she could buy a house for her father before he dies. This reflects the ways in which social-relational care and cultural reciprocity are seen as central to their worldviews.

This socio-relational approach illuminates the ways in which educational research could be reframed from non-Western epistemologies and from the standpoint of girls with disabilities (Nguyen, 2016). A decolonial approach in social and educational research with girls with disabilities in the Global South requires that we critically engage with complex, and sometimes conflictual experiences of girls with disabilities in specific communities in the Global South as a form of countering the colonial imagination of difference and of reclaiming re-existence. It challenges epistemic injustice that has been framed as universal by the Western empire. This art of research requires us to begin the work of undoing and unthinking the coloniality of power which shapes how we come to know and learn about the Other (Ndlovu-Gatsheni, 2020).

Within this decolonial lens then, a participatory arts-based approach that engages with girls with disabilities is important because it reframes the colonial structure of research—that is, the ways in which research tends to be done *about* and not *with* or *from* the Global South. With basic filming skills and techniques, the girls were able to illuminate different aspects of unequal power relations which they endure in their everyday life. Their cellfilms suggest that the girls were well aware of the structures of inequalities in relation to disability, gender, and ethnic discrimination. In so doing, this approach worked to empower these girls to frame their own stories and reclaim their existence.

Conclusion

Writing from their locations as disability scholars situated in the Global North and South, Goodley and Swartz (2016) state that it is difficult for theorists from the Global South to engage in these conversations given the power of academics from the Global North. They argue that it is important for those generally excluded from these discussions to actively participate. However, it is not easy for the girls and young women with disabilities in the Global South to be involved in these conversations. In order to do that, we must interrogate epistemological privileges of Western and Eurocentric ideologies in knowledge production and redefine traditional approaches to qualitative research methods such as interviews or focus groups, which sustain power dynamics between Western researchers and participants from the Global South. Such approaches tend to marginalize and exclude disabled activists and scholars from the South from these conversations.

TDKRA offers an example for how research with girls with disabilities can offer an alternative approach for engaging disability and girlhood in the Global South through creative and participatory arts-based methods. The use of decolonial DS and participatory visual research empowered participants to re-imagine their spaces, relationships, and social identities in ways that challenge Northern narratives of disabled girlhoods' sufferings. In so doing, this chapter demonstrates how research *with* communities and participants in the Global South can be conducted as a form of "epistemological decolonization" (Quijano, 2007, p. 177), despite being imbued with multiple forms of struggles. Rather than offering a *universal* truth about the lives of disabled girls, decolonial DS is *an option* (Mignolo, 2018, p. 115) for educational research to resist coloniality; in so doing, it enables us to reclaim the voices, power, and identities of disabled and marginalized communities in the Global South.

Acknowledgment

I would like to thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for its financial support. I thank Meysoon Amin for her thoughtful comments and assistance towards the completion of this chapter.

References

- Anand, S., 2015. Corporeality and culture: theorizing difference in the South Asian context. In: Rao, S., Kalyanpur, M. (Eds.), *South Asia and Disability Studies: Redefining Boundaries & Extending Horizons*. Peter Lang Publishing, New York, pp. 154–170.
- Connell, R., 2007. *Southern Theory: Social Science and the Global Dynamics of Knowledge*. Polity Press, Oxford, England.
- Connell, R., 2011. Southern bodies and disability: re-thinking concepts. *Third World Q.* 32, 1369–1381.
- Davis, L.J., 1995. *Enforcing Normalcy: Disability, Deafness and the Body*. Verso Books, London.
- Denzin, N.K., 2009. The elephant in the living room: or extending the conversation about the politics of evidence. *Qual. Res.* 9, 139–160.
- Dirth, T.P., Adams, G.A., 2019. Decolonial theory and disability studies: on the modernity/coloniality of ability. *J. Soc. Polit. Psychol.* 7, 260–289.
- Erevelles, N., Mutua, K., 2005. "I am a woman now!" Rewriting cartographies of girlhood from the critical standpoint of disability. In: Bettis, P., Adams, N.G. (Eds.), *Geographies of Girlhood: Identities In-Between*. Routledge, New York, pp. 127–134.
- Erevelles, N., Nguyen, X.T., 2016. Disability, girlhood, and vulnerability in transnational contexts. *Girlhood Stud.* 9, 3–20.
- Garland-Thomson, R., 2018. Critical disability studies: a knowledge manifesto. In: Kent, M., Ellis, K., Robertson, R., Thomson, R.G. (Eds.), *Manifestos for the Future of Critical Disability Studies*. Routledge, Oxon, pp. 11–19.
- Gonick, M., Gannon, S., 2014. *Becoming Girl: Collective Biography and the Production of Girlhood*. Women's Press, Toronto.
- Goodley, D., Swartz, L., 2016. The place of disability. In: Grech, S., Soldatic, K. (Eds.), *Disability in the Global South. International Perspectives on Social Policy, Administration, and Practice*. Springer, Cham, pp. 69–83.
- Grech, S., 2015. Decolonising Eurocentric disability studies: why colonialism matters in the disability and global South debate. *Soc. Ident.* 21, 6–21.
- Human Rights Watch, 2020. Disability is Not Weakness: Discrimination and Barriers Facing Women and Girls With Disabilities in Afghanistan. Available at: <https://www.hrw.org/report/2020/04/28/disability-not-weakness/discrimination-and-barriers-facing-women-and-girls>. (Accessed 17 September 2020).
- Jaffee, L., John, K., 2018. Disabling bodies of/and land: reframing disability justice in conversation with indigenous theory and activism. *Disability and the Global South* 5 (2), 1407–1429. Retrieved from https://disabilityglobalsouth.files.wordpress.com/2018/08/dgs_05_02_04.pdf.
- Kim, E., 2017. *Curative Violence: Rehabilitating Disability, Gender, and Sexuality in Modern Korea*. Duke University Press, Durham.
- Maldonado-Torres, N., 2007. On the coloniality of being: contributions to the development of a concept. *Cult. Stud.* 21, 240–270.
- Meekosha, H., Soldatic, K., 2011. Human rights and the Global South: the case of disability. *Third World Q.* 32, 1383–1397.
- Mehrotra, N., 2020. *Disability Studies in India: Interdisciplinary Perspectives*. Springer, Singapore.
- Mignolo, W., 2018. The decolonial option. In: Mignolo, W., Walsh, E. (Eds.), *On Decoloniality: Concepts, Analytics, Praxis*. Duke University Press, Durham, pp. 105–244.
- Mitchell, C., 2011. *Doing Visual Research*. Sage Publications, Los Angeles.
- Mucina, D., 2013. Ubuntu orality as a living philosophy. *J. Pan Afr. Stud.* 6, 18–36.
- Ndlovu-Gatsheni, S.J., 2020. The cognitive empire, politics of knowledge and African intellectual productions: reflections on struggles for epistemic freedom and resurgence of decolonisation in the twenty-first century. *Third World Q.* 42, 882–901.
- Ngo, A.D., Brolan, C.E., Fitzgerald, L., Pham, V., Phan, H., 2013. Voices from Vietnam: experiences of children and youth with disabilities, and their families, from an Agent Orange affected rural region. *Disabil. Soc.* 28, 955–969.
- Nguyen, X.T., Stienstra, D., 2021. Engaging girls and women with disabilities in the global South: beyond cultural and geopolitical generalizations. *Disabil. Glob. South* 8, 2035–2052.
- Nguyen, X.T., Mitchell, C., De Lange, N., Fritsch, K., 2015. Engaging girls with disabilities in Vietnam: making their voices count. *Disabil. Soc.* 30, 773–787.
- Nguyen, X.T., Stienstra, D., Gonick, D., Do, M., Huynh, N., 2019. Unsettling research versus activism: how might critical disability studies disrupt traditional research boundaries? *Disabil. Soc.* 34, 1042–1061.
- Nguyen, X.T., Dang, T.L., Mitchell, C., 2021a. How can girls with disabilities become activists in their own lives? Creating opportunities for policy dialogue through "knowledge mobilisation spaces". *Agenda* 35, 67–79.
- Nguyen, X.T., Gonick, M., Bui, T., 2021b. Engaging girls with disabilities through cellphilm: reflections on participatory visual research as a means of countering violence in the Global South. *Childhood* 28, 380–394.
- Nguyen, X.T., 2015. *The Journey to Inclusion*. Sense Publishers, Rotterdam.
- Nguyen, X.T., 2016. Girls with disabilities in the Global South: rethinking the politics of engagement. *Girlhood Stud.* 9, 53–71.
- Nguyen, X.T., 2021. Critical disability studies at the edge of global development: why do we need to engage with Southern theory? In: DRIEGER, D. (Ed.), *Still Living the Edges: A Disabled Women's Reader*. Inanna, Toronto, pp. 349–371.
- Office of the United Nations High Commissioner for Human Rights, 2017. Report on Sexual and Reproductive Health and Rights of Girls and Young Women With Disabilities. Available at: <https://www.ohchr.org/EN/Issues/Disability/SRDIsabilities/Pages/ReproductiveHealthRights.aspx>. (Accessed 18 November 2021).
- Pearce, E., Paik, K., Robles, O.J., 2016. Adolescent girls with disabilities in humanitarian settings: "I am not 'worthless'—I am a girl with a lot to share and offer". *Girlhood Stud.* 9, 118–136.
- Quijano, A., 2007. Coloniality and modernity/rationality. *Cult. Stud.* 21, 168–178.
- Rose, G., 2001. *Visual Methodologies: An Introduction to the Interpretation of Visual Materials*. Sage, Thousand Oaks, CA.
- Shah, S., Bradbury-Jones, C., 2018. *Disability, Gender and Violence Over the Life Course: Global Perspectives and Human Rights Approaches*. Routledge, England.

- Smith, L.T., 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books, England.
- Staples, J., 2020. Decolonising disability studies? Developing South Asia-Specific approaches to understanding disability. In: Mehrotra, N. (Ed.), *Disability Studies in India*. Springer, Singapore, pp. 25–41.
- Stienstra, D., 2015. Trumping all? Disability and girlhood studies. *Girlhood Stud.* 8, 54–70.
- Swartz, L., 2018. Representing disability and development in the global south. *Med. Humanit.* 44, 281–284.
- Titchkosky, T., 2011. *The Question of Access: Disability, Space, Meaning*. University of Toronto Press, Canada.
- Tynan, L., 2021. What is relationality? Indigenous knowledges, practices and responsibilities with kin. *Cult. Geogr.* 28, 597–610.
- United Nations Girls' Education Initiative, Leonard Cheshire Disability, 2017. *Still Left Behind: Pathways to Inclusive Education for Girls With Disabilities*. Available at. <https://www.ungei.org/publication/still-left-behind>. (Accessed 19 November 2020).
- United Nations, 2006. *Convention on the Rights of Persons With Disabilities*. Available at. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>. (Accessed 21 November 2021).
- World Health Organization and World Bank, 2011. *World Report on Disability*. Available at. http://whqlibdoc.who.int/publications/2011/9789240685215_eng.pdf. (Accessed 5 February 2012).

Political orientation

Murray Print, University of Sydney, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Youth electoral study	134
Political orientation influences	135
Informal curriculum: student governance	136
Political party orientation	137
Conclusion	138
References	138

Under the broad theme of diversity, democracy and social justice in education, the theme for this volume, it is possible to explore a vast array of sub-themes. This chapter examines identities and differences: political orientation through the vehicle of the school, the school curriculum and a national assessment of student knowledge and understanding of civic and political knowledge, skills and values. In the context of Australia, which can serve as an example to others, this analysis examines the role of the Australian Curriculum (AC) in influencing student political orientation and some measures of that orientation through the National Assessment Program Civics and Citizenship (NAPCC).

What does political orientation mean? By policy orientation we mean an orientation that characterizes the political thinking of a group or nation. Simplistically this can be left–right analysis, but more realistically there are four recognized political orientations - liberalism, conservatism, libertarianism and populism. In addition, political orientation can refer to the willingness to participate in politically, that is, do people exhibit a willingness to participate, or not, in what they consider to be political activities such as enrolling to vote and the act of voting itself (especially where voting is voluntary).

With youth the situation is different, evolving and more dynamic than with older groups as young people experience their transitional search for identity. Transitioning from child to teenager to adult in a relatively short time is, in itself, a potentially traumatic, incomplete experience. Overlaying an emerging political identity and orientation on this physical, emotional and psychological transition can complicate the process profoundly. Consequently many to most young people as they emerge into adulthood at the nominal age of eighteen years lack a political orientation or are trying to identify themselves beyond the influences of family and friends.

This chapter examines youth political orientation through the lenses of the school and what evidence we have of political orientations and where and how they are formed. In the context of Australian schools there exists a widespread understanding that teachers and schools will not deliberately promote any particular political party orientation/ideology. However, through the formal and informal curricula in schools, as well as the hidden curriculum (Print, 1993), there are multifarious opportunities for students to acquire, modify or consolidate a political orientation. And with senior school students especially (as most Australian senior students turn eighteen in their final school year), there are opportunities for young people to express those orientations in real forms such as voting in elections, joining political parties/groups and participating in political activities as an adult.

Age has not prevented thousands of Australian school students from demonstrating their political orientation on major international issues. In particular Australian students have joined international movements to protest about climate change and the role that the Australian governments have played. This form of political orientation is essentially non-partisan and in this case reflects widespread concern about the future of the earth. In late October 2021, for example, thousands of Australian school students took to the streets and the internet to demand action on climate change. The demonstration has been one of several over recent years on essentially the same theme that reflects, as some see, a lack of appropriate curriculum experience on climate change (Gobby and Varylan, 2021; also see <https://www.schoolstrike4climate.com>). Essentially a national student demonstration demanding action on climate change has occurred annually since 2018 and continued during the years of the pandemic.

Youth electoral study

Most information we have about student political orientation and identity comes from time-constrained surveys of young people or media, especially newspaper, coverage. Typically they are conducted at a politically-charged time and so enhanced expectations, actions and emotions ensue. There is a research study that was conducted in conjunction with the Australian Electoral Commission that sought to understand, among other issues, student political orientation.

The Youth Electoral Study (YES) was conducted nationally in Australia from 2004 to 2010 and included participants from all six Australian states and the two territories. A sample of 4758 secondary school students participated in the national survey were in the fifteen to seventeen age range while hundreds more were included in focus groups conducted in schools around the country. The approximately equal female and male students studied at government, independent and Catholic schools in metropolitan and rural locations around Australia with a slight emphasis in non-government school participation.

A key source of qualitative data came from the 16 electoral divisions (from 150 nationally) selected as research sites. These covered the main categories of electoral divisions – inner city, mid city-suburban, outer suburban, rural city, rural town and remote. Qualitative data were collected over a five year period with youth aged 17–25 in school and non-school sites, the latter including universities, TAFE, unemployed and employed groups. Reports of the project may be found on the Australian Electoral Commission's website under the heading "research." Other publications from the project are cited below and the reference list.

An initial key finding of the study was that all students knew a minimum of three issues related to themselves, their political orientation and voting. First was that voting in Australian elections, at least at national and state levels, was compulsory. Second that the voting age was eighteen years old. And third students knew that there were penalties involved in not voting when it was compulsory. They demonstrated little understanding of what the penalties were (mostly an inexpensive fine) or that they were entitled to register to vote when they turned seventeen years of age (AEC, 2009).

Most students had not made substantial political orientations in terms of political ideologies or political parties (Print, 2007; Saha and Print, 2010). Those who had were significantly influenced by their family, especially their parents, and were likely to be left of center in political orientation. This group tended to be the same very small percentage of students who could be considered politically active and hence displayed a political orientation.

The general lack of political orientation can be identified in the data on voting from the Australian Electoral Commission. Their data, for example, revealed that less than 10% of 17 year olds were registered to vote with the AEC and when they could vote at 18 years, nearly 60% actually voted. These percentages must be remembered in the context of compulsory voting from 18 years and with sanctions, and that all student knew the voting age was 18 and that voting in Australia, especially at federal and state elections, was compulsory (AEC, 2009).

Most students recalled that they found the study of government rather boring, if they could remember at all! (AEC, 2009; Saha et al., 2007). The last time most students studied about the government and the political system was in Year Nine or Ten when they were typically fifteen years old. Less than a quarter of students found the study of government interesting (AEC, 2009), reflecting a limited perspective on political orientation. However, "... students who could remember studying about the government, and also could remember the name of the subject and the year in which they took it, are slightly more likely to say that they will vote at 18, even if they did not have to." (AEC, 2009, p. 12). And the more interested students are in the political system and government, the more committed they are to voting.

A general feature of schools and the school curriculum in Australia was that students found extremely limited opportunities to learn about politics through the formal curriculum or school subjects (Print, 1993; Saha and Print, 2010). Typically after Year 10 students had extremely limited opportunities to study any aspect of political systems in the final two years of school (Saha et al., 2007).

The international research reveals mixed results for student political orientation and participation in school contexts and many studies are conducted through forms of survey research. For example the large (28 countries) IEA Civic Education Study (CivEd) investigated high school student knowledge, attitudes and behavior in school contexts especially where students participated in formal and informal civic education activities. The international study showed that formal civic learning, as measured by students' civic knowledge, was positively related to the likelihood that students intended to vote in the future as a form of political engagement (Torney-Purta et al., 2001). The later and larger (38 countries) International Civic and Citizenship Education Study (ICCS), however, found that while civic knowledge was positively associated with future electoral participation and legal protest, it correlated negatively with expected active political participation and illegal protest (Schulz et al., 2010).

An example of many studies at a national level is the review of American youth political research by William Galston where he relied on survey research to identify young Americans' civic/political knowledge and engagement. He noted, as have others such as the famous study by Putnam (2000), the decline in political engagement among American youth has been in decline, especially as measured by voting in elections. More concerning perhaps was that trust in government and most major social institutions was low and had remained so for some time. Further, despite large increases in formal education in the US, measures of political knowledge have remained flat for decades. Studies and commentaries by Campbell (2010), Harris et al. (2010), Geboers et al. (2015), and Barrett (2017), across different countries, reveal a similar results though in more recent years an increasing willingness of young people to demonstrate their political orientation through direct action.

Political orientation influences

The Youth Electoral Study found that parents were the most significant source for influencing student views on politics. They were more influential than siblings, peers, teachers, schools, community and religion (AEC, 2009). The study revealed that the characteristics of the parents (education levels, discussions on politics) were more important in determining whether the student intended to vote and their political orientation, rather than the other variables that could apply. "In the family, the extent to which the student learns about politics from their parents, and the extent to which the student claims to be influenced about voting by parents, both appear very important in affecting the student's commitment to voting" (AEC, 2009, p. 36).

Teachers were also influential though in a non-partisan way. Teachers in Australian schools are expected to refrain from espousing their personal political orientations and so to not influence student political orientations. Students stated that their teachers generally did not impart their personal political views though some were supportive of broader political agendas such as anti-Iraq war involvement.

Interestingly school type made a difference to student political knowledge and commitment to voting with students attending independent schools higher in both than catholic and government school students. However, when both attendance at independent school and family characteristics were considered, the influence of private schools on voting commitment disappeared, whereas with respect to political knowledge, it diminished, but remained important (AEC, 2009).

Students agreed by way of explanation, at the moment, they had little in the way of responsibilities to society and their family other than succeeding at school. Politics in general, especially if requiring a political orientation, was not high on their agendas. Of those who did express political orientations, most would classify themselves as supporting democracy, liberalism and were "left-wing" though only a handful had joined a political party (AEC, 2009).

However, there are issues that students feel very strongly about and are prepared to take direct action about, with resulting major demonstrations held in school hours. These are instances of political orientation in practice though they are rarely aligned with any political party or political ideology. For example, over the past two decades there have been major student demonstrations against Australia's participation in the Iraq war, about support for climate change action, about government responses to the global financial crisis, and recently demanding action against domestic violence, child abuse and the need for consent (<https://www.schoolstrike4climate.com>).

What inspires such actions and movements is a widespread concern among many young people about their future and the future of the world. They have, though various sources including mainstream media, social media and school experiences, come to understand the seriousness of these issues and the associated need to speak up and democratically protest about perceived inaction. Interestingly those students who protested in recent years, have now come of voting age and in the most recent Australian general elections the conservative government was soundly defeated, producing its worst election vote in decades. A key issue in that election was how the government appeared to be not listening carefully to concerns about issues such as climate change.

And what role have schools played in these actions? In the student strike against climate inaction by governments most secondary schools in Australia were remarkably supportive. Many schools "allowed" students to attend demonstrations, some even supported by teachers, while others ignored student non-attendance on a school day. The existing Australian Curriculum, employed across all Australian 9450 schools (though with some modifications in a few states) does address key issues that concern students, their future, the future of Australia and the world. The informal curriculum in schools provides a further, perhaps more tangible, opportunity for enhancing student political understanding and engagement.

Informal curriculum: student governance

In the literature on the informal curriculum discussion has ranged about the effectiveness and impact of activities identified as within the informal curriculum such as student governance, especially student councils (Print, 2009; Print et al., 2002; Saha and Print, 2010). YES identified a relationship between participation in student governance and political engagement which, in turn, requires a political orientation. The study data made it clear that participating in school elections had a beneficial positive effect on both electoral enrollment behavior and the voting intentions of secondary school students. "For students who are 17 years old, the relationship between voting and standing for office are much more strongly related to intention to vote than for actually having enrolled..." (AEC, 2009, p. 32).

The Australian study found that students who participated in school elections were more knowledgeable and prone to engage in the political realm (Saha and Print, 2010). Consequently these students developed a political orientation either overtly and directly or indirectly and unknowingly. YES found that voting in school elections was positively related to feeling prepared to vote as an adult, to being committed to vote when 18, to political knowledge, and to engagement in forms of peaceful activism. The more active role of running for student government office was further related to political knowledge and participation in peaceful activism. These results reinforced the findings of adult retrospective studies, showing that participation in school elections served as a beneficial experience in the preparation of students for life as an active adult citizen with a political orientation.

There is a body of literature that argues that by experiencing various political activities, and acquiring political knowledge, a person acquires a self-identity of which political orientation is a necessary part (Barrett, 2017; Geboers et al., 2015; Harris et al., 2010). It is, however, a complex and messy process, but one where school experience can contribute. For example, by participating in a school election, a student would self-identify as a "voter," which makes it easier to vote again later in life as a political orientation has been established. Furthermore, students who vote, or run for office, in school elections begin to think that they can make decisions in school settings, and that they possess power to determine, at least in part, their own political environment in the future, thereby consolidating a sense of political orientation. At this point nascent party political orientations may also become apparent.

Similarly, in a significant study in the United States David Campbell (2010) identified that a sense of civic duty could spring from one's early influences such as the school. The degree to which people feel a sense of civic obligation, including a political orientation, Campbell claimed, stems, in part, from their adolescent experience. The socialization process in the context of a community with strong civic norms would lead people to be civically engaged in adulthood and hence manifesting a form of political orientation. Campbell (2010) demonstrated how the civic norms manifest within the high school would impact on individuals' civic involvement and political orientation even a decade or after those individuals had completed secondary education.

In conclusion the Youth Electoral Study provided evidence to support the notion that school elections may be the "cradle" of democracy, since it is here that students experience a fundamental democratic practice and early signs of political orientation (Saha

and Print, 2010). YES demonstrated empirically that students who do vote or run for office, are also more prone to feel prepared to vote as adults, to actually intend to vote, to know more about politics, and to have already experienced some form of political activism, such as attending rallies or letter writing (AEC, 2009; Saha and Print, 2010). These behaviors represent forms of political orientation both in a general sense of orientation and in the sense of political party orientation.

Political party orientation

In western democracies political parties appear to be a fixed feature of the political landscape. This reflects the democratic world where governments are based on political parties who control a majority in parliaments. Political parties play a role in identifying political philosophies and policies that affect the political orientation of young people. Examining the political orientation of young people toward political parties is important if one wants to study young people and voting. In Australia we are interested to identify when young people vote for the first time that they know enough about political parties, and about their own personal philosophy, to make a reasoned choice in their voting. Since parents are the most important early political socialisers of their children, we need to know how much of young persons' political party orientation is related to the leanings of their parents. Further we want to know whether the political orientation makes a difference in young people's voting intentions.

Using the data from the YES research of young Australians it is possible to identify their nascent political orientation. In very small numbers the political orientation of young people is clear and unequivocal. In Australia voting is compulsory for all citizens aged eighteen years and older. So if young people are required to vote it is reasonable to expect that they will begin to differentiate between the parties as they approach voting age. Yet slightly more than half of the young people in the study did not feel confident in their knowledge of political parties to vote in a national election (AEC, 2009).

Youth political orientation is acquired through multifarious sources over time so that by the time students are in the final years of school in Australia, and are able to vote, they invariably exhibit some form of political party orientation however minor. In the YES survey nearly half the students identified a political party orientation while the remainder were unsure which political party met their political needs (Saha et al., 2007). And those who indicated a political party orientation were more likely to vote in the future (Saha et al., 2007). In terms of political parties and party differences, it is clear that the Liberal/National parties enjoyed the highest level of voting commitment by the youth who identify with them, followed by the Labor Party and the Green Party. What is more significant in the Australian context, is how voting commitment declined for young people who have either rejected a party identity, or do not, for whatever reason, identify with any party. Only one-third of these students would vote in the hypothetical situation if they did not have to (Saha et al., 2007).

Why do young Australians demonstrate a political party orientation? In the YES research many students named the same political party orientation as they perceived their parents favored. This outcome was facilitated through family discussions and comments about political parties, figures and policies in daily interactions. A sense of family "inheritance" of party identity can therefore be identified and can be explained in many ways, for example by means of political party cues that young people pick up from their parents, or through the perceived link between the social status of the party and that of the family and even the school.

The formation of party identity among young people, that is their political party orientation, will continue as long as the Australian political system is based on party politics, a Westminster parliamentary government, and voting focused more on political parties than on individuals. The YES research has demonstrated a high degree of party inheritance attributed to the family, though other factors at play. They study identified that more than half of the young people did not name a party or demonstrate clear political party orientation. Further, it appears that some young people deliberately aligned themselves with a party other than that of their parents (Saha et al., 2007).

In their national study of older youth participatory practices and political knowledge Reichert and Print (2018, 2019) used items drawn from the National Assessment Program Civics and Citizenship (NAPCC). The study utilized appropriate items on political knowledge and attitudes and found that young people's commitments to traditional forms of political participation were declining, which challenges arguments supporting political orientation, particularly political party orientation. By contrast, political orientation of young people was demonstrated by engagement in socio-political forms of participation around key and pressing phenomena such as climate change which were not identified as political party oriented (Gobby and Varylan, 2021). In recent times these issues have led to widespread student action such as school student strikes in favor of greater direct political action to combat climate change (<https://www.schoolstrike4climate.com>).

Significantly their study found participation in party-political activities appears to require high levels of efficacy and motivation (Saha et al., 2007). Socio-political activities might be more intrinsically rewarding than party-political activities that often require stronger personal investments for young people. This study suggests that higher levels of knowledge are associated with more frequent exposure to political news on television only among highly educated respondents who report average or below average interest in politics. This association might reflect knowledge that is absorbed while watching entertaining news as this association only appears among highly educated young adults (Saha et al., 2007).

The NAPCC itself, a national sample assessment of school students in Years six and ten, identifies some items in civics and citizenship that require students to reveal their general political orientation and to a minor extent their political party orientation. A commonly agreed theme of NAPCC has been that civics and citizenship education aims to enable students to become informed and active citizens. From its inception, NAPCC has consequently collected data on students' knowledge and understanding of civics and citizenship content as well as the attitudes, values and behaviors that relate to participatory citizenship. In turn these data provide an

insight into student political orientation. For example, a key goal of the NAPCC is to ensure students develop knowledge, understanding and appreciation of the values, principles, institutions and practices of Australia's system of democratic government and law, and the role of the citizen in Australian government and society. This goal facilitates an understanding of student political orientation.

In terms of a general political orientation and the knowledge required to support it the report found that at the national level in 2016, only 38% of Year 10 students achieved at or above the proficient standard. This standard is what would be expected of a student to achieve and the 2016 result for Year 10 students is significantly lower than the percentage achieved in each of the two previous cycles (44% in 2013 and 49% in 2010) (ACARA, 2017, p. 41). Surprisingly the Year 10 NAPCC data show at the national level that political knowledge, which is considered a necessary precursor to political orientation, more than 60% of students did not achieve at the proficient standard. This is a sad indictment of preparation for becoming an informed adult citizen in Australian democracy.

Students in 2016 also tended to have higher levels of interest in civic issues, greater belief in the value of civic action, and stronger expectations to engage in civic activities in the future, particularly compared to 2010 (ACARA, 2017, p. 117). Interestingly 63% of Year 10 students identified supporting a political party as an important citizenship behavior, while only 51% supported discussing politics. However, addressing issues stated above, such as protecting the environment, promoting human rights and personal efforts to protect natural resources received around 80% support (ACARA, 2017, p. 142) and in the process demonstrated evidence of a general political orientation.

Conclusion

The concept of political orientation among young people is complex and dominated by two perspectives. First a general understanding of political orientation that reflects the general level of interest and participation in politics. A second perspective considers young people's political party orientation and the degree to which, if any, they are involved in political party activities. Generally Australian school students have demonstrated substantial general political orientation but relatively low political party orientation. On specific issues, however, as the NAPCC data reveal, young Australians demonstrate clear evidence of political orientation.

Between the two key perspectives there is increasing evidence of student political orientation toward specific phenomena or key issues about which young people have strong views. The strength of these views may involve participating in general political behavior such as demonstrations and the use of social media to influence governments and the community in general. For example, many young people have evidenced very strong views and behavior on perceived federal government inaction on climate change in Australia. Mostly this political orientation is general in nature though targeted to a specific issue rather than follow political party policy and ideology. However, there is evidence that a small number of young people in schools do exhibit political party orientation though this is not promoted or endorsed by the school.

References

- AEC, 2009. Youth Schools and Learning about Politics. *Youth Electoral Study*. AEC, Canberra.
- Australian Curriculum and Assessment Reporting Authority, 2017. NAP Sample Assessment: Civics and Citizenship Report, 2016. ACARA, Sydney.
- Barrett, M., 2017. Young people's civic and political engagement and global citizenship. *UN Chronicle* LIV, 44–47.
- Campbell, D., 2010. *Why We Vote: How Schools and Communities Shape Our Civic Life*. Princeton University Press, Princeton.
- Geboers, E., Geijssels, F., Admiraal, W., ten Dam, G., 2015. Citizenship orientations and knowledge in primary and secondary education. *Soc. Psychol. Educ.* 18 (4), 749–767.
- Gobby, B., Varylan, G., 2021. Retrieved from. <https://theconversation.com/curriculum-is-a-climate-change-battleground-and-states-must-step-in-to-prepare-students-172392>. (Accessed 27 February 2021).
- Harris, A., Wyn, J., Younes, S., 2010. Beyond apathetic or activist youth. 'Ordinary' young people and contemporary forms of participation. *Young* 18 (1), 9–32.
- Print, M., 1993. *Curriculum Development and Design*. Allen & Unwin, Sydney.
- Print, M., 2007. Learning political engagement in schools. In: Print, M., Saha, L., Edwards, K. (Eds.), *Youth Participation in Democracy*. Sense, Amsterdam, pp. 95–112.
- Print, M., 2009. Connecting youth political participation and civic education in school. In: Print, M., Milner, H. (Eds.), *Civic Education and Youth Political Participation*. Sense Publishers, Rotterdam.
- Print, M., Ørnstrøm, S., Skovgaard Nielsen, H., 2002. Education for democratic processes in schools and classrooms. *Eur. J. Educ.* 37 (2), 193–210.
- Putnam, R.D., 2000. *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, New York.
- Reichert, F., Print, M., 2018. Civic participation of high school students: the effect of civic learning in school. *Educ. Rev.* 70, 318–341.
- Reichert, F., Print, M., 2019. Participatory practices and political knowledge: how motivational inequality moderates the effects of formal participation on knowledge. *Soc. Psychol. Educ.* 22, 1085–1108. <https://doi.org/10.1007/s11218-019-09514-5>.
- Saha, L., Print, M., Edwards, K. (Eds.), 2007. *Youth Participation in Democracy*. Sense Publishers, Amsterdam.
- Saha, L., Print, M., 2010. Schools, student elections and political engagement: the cradle of democracy? *Int. J. Educ. Res.* 49, 22–32.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D., Losito, B., 2010. ICCS 2009 International Report: Civic Knowledge, Attitudes, and Engagement Among Lower-Secondary School Students in 38 Countries. IEA, Amsterdam.
- Torney-Purta, J., Lehmann, R., Oswald, H., Schulz, W. (Eds.), 2001. *Citizenship and Education in Twenty Eight Countries: Civic Knowledge and Engagement at Age Fourteen*. IEA, Amsterdam. Retrieved from. <https://www.schoolstrike4climate.com>.

K-12 global/international education: dancing with “diversity, democracy, and social justice”

Paul Tarc, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	139
Terminology and trajectories: distinctions, overlaps and percolating flashpoints	140
International education	140
Global education	141
Variants of GE	144
Critical faultlines and tensions of GE	145
The tension between strategic and dedicated cosmopolitan sovereignties and subjectivities	145
The demands on education to make a difference and considering the educational of education	146
Conclusion: prospects for diversity, democracy and social justice	146
References	147

Introduction

This entry provides a synthetic overview of (historically-inflected) visions and practices of global/international education, oriented to the domain of k-12 schooling in the 21st century. Reflecting the English-medium literature/scholarship examined (and my social/epistemic location as a White Anglo-Canadian researcher of, and practitioner in, global/international education), the perspective offered is Anglo-Western. It is important to note, however, that non-Western and non-Anglo regions have shaped/co-constituted “Western” conceptions of global/international education; moreover, across most parts of the world, Western modes of schooling (and being) have been dominant referents for educational initiatives (i.e. “education for development,” “study abroad” ...), under the broader sweep of 20th century Euro-American hegemony. Therefore, this Anglo-Western centered conception of global/international education is informed by, and will likely have significant reverberations with, global/international education in multiple societal and educational jurisdiction contexts. Finally, in terms of the literature cited in this entry, I did not attempt a neutral search strategy to filter the available literature. Rather, toward offering a coherent and compelling (if partial) accounting of global/international education, I purposefully leaned into my own reading bibliography and past writing, as well as into my experience as a researcher and educator in the area of global/international education.

Given the thematic of this volume, the entry attends to the idealist visions of global/international education, such as educating for “international understanding” and “planetary consciousness,” and how these visions (are to) find expression in k-12 school practices. As will be explicated in the next section, *global* education and *international* education (among other terms) can be differentiated. These differences are important and need to be considered in contextually-sensitive ways. *Nevertheless*, under interpenetrating colonial histories ascribing countries as “developed” or “under-developed” and more recent processes of (neoliberal) globalization and the associated global educational policy convergence (Rizvi and Lingard, 2010), which includes the rise of global/international education as an “expedient” (Tarc, 2013a), there are *converging* pressures on, and thereby overlapping uses and tensions of, global/international education in these early decades of the 21st century.

Further, the various terms/signifiers of global and international educational discourse often work *performatively* as partly “floating signifiers” (Manion et al., 2011; Tarc, 2019), referring to an overlapping set of features ascribed to global and international education. In very abbreviated terms, these features are founded on the historical/deeper aims and functions of (1) a border-crossing/inter-national and/or holistic/planetary orientation and (2) an idealist pedagogical vision/belief that bringing in or having this supra-national(ist) orientation can support education’s normative goal of making a “better” (less violent, more egalitarian, ...) world (Hanvey, 1976; Heater, 1980; Tarc, 2009). In terms of the *education* in global/international education, liberal-humanist and child-centred pedagogies, reflecting the progressive internationalist viewpoints of advocates of global/international education, have been the dominant forms (Hughes, 2012; Pike and Selby, 1988; Tarc, 2009); there have also been calls for more *critical* orientations of global/international education such as the Marxist-inspired “peace education” of the 1970s (Heater, 1980) and more recent, critically-informed articulations of “global citizenship education” (Abdi et al., 2015; Andreotti, 2014; Andreotti and de Souza, 2012; Lapayese, 2003; Marshall, 2011; Pashby and Sund, 2020; Stein, 2021; Tarc, 2012, 2015). Rather than prioritizing a semantic analyses on the different terms used or on the boundary making of academic discourse communities, the purpose of this article is to provide an historical overview and schema to interpret the multiple manifestations of global/international education more conceptually (see also Tarc, 2019). Accordingly, this entry digs into the deeper aims and evolving tensions as (partly) constitutive features of K-12 global/international education converging in the early decades of the 21st century.

Given the *shared* core impulses and features, as well as the slippages of the terms used in educational discourses, I take “global/international” as an umbrella term, inclusive of a set of other terms, such as “global citizenship education,” and

education for “intercultural competence.” Further, for readability, I shorten “global/international education” to “global education” and use the acronym GE when referring to global education as the global/international education *umbrella* term. Following the next section discussing terminology, where I differentiate some of the distinctive inflections and uses of the variant terms, in this entry GE broadly signifies a whole set of older and newer terms related to global and international education. Although GE has a longer history (Good, 2020; Hill, 2002; Méras, 1932), this entry considers the global education trajectory over the last fifty or so years. Additionally, the focus on the k-12 level of schooling also has guided the selection of literature informing this entry.

Finally, the discussion moves from a synthetic and historically-informed overview to a postulation of the converging core flashpoints and tensions of global/international education as they have evolved into the contemporary moment. These flashpoints/tensions include the following: the tensions between strategic and dedicated cosmopolitan sovereignties and subjectivities; *how* GE can make a difference, *educationally*; and how this *difference made* relates to the project of making a “better world.”

Terminology and trajectories: distinctions, overlaps and percolating flashpoints

Reflecting heightened transnational flows and interdependencies, the adjectives “global,” “international,” “intercultural” and “cosmopolitan” have become ubiquitous across many educational discourses and levels of schooling. The purpose of this section is to first summarize key semantic distinctions primarily between “global” and “international” education. The second objective is to illustrate a kind of convergence around their performative meanings and uses. This convergence allows for a “frame” for illuminating, researching and potentially steering variants of GE across multiple contexts (Tarc, 2019).

International education

In Tarc’s (2007a) conceptual genealogy of the “international” in the International Baccalaureate, he finds that the *international* in international education carries both a literal and idealist meaning (see also entry *International Baccalaureate: Meanings and uses in a globalizing world* in Elsevier’s companion volume “Globalization and the Geopolitics of Education,” Tarc, 2022). On the one hand, inter-national education, *literally* signifies educational activities that cross or connect up across national borders. Here we have the example of students who “study abroad.” Yet the actual curricula they study may be a wholly national or state curriculum. On the other hand, international education, at least for the last century, has also signified an outward-looking, progressive and child-centered education for “international understanding” that seeks to move beyond a nationalistic frame, toward the making of a “better world” in a context of heightened interdependencies (Tarc, 2009). Here the “international” in international education brings both an international-humanist orientation to education as well as a pedagogical methodology that is progressive and child centered. This conception of the core visions of international education aligns well with Stier’s (2004) depiction of three foundational ideologies of the internationalization in *higher education* (IHE) as “instrumentalism,” “idealism” and “educationalism,” and with Hill’s (2002) differentiation of the “ideological,” “utilitarian” and “pedagogical” rationales for international education in the *field of international schools*. In real world manifestations of international education, the literal and idealist/aspirational agendas become entangled (Tarc, 2009). For example, the context of crossing geographic borders and encountering the unfamiliar can provoke the kind of international understanding to which the idealist agendas are oriented. So to, those individuals inclined to international and intercultural understanding may be more open to, and involved in, cross-border activities (Tarc, 2013a).

To differentiate international education from global education, it is important to emphasize how, semantically, international education represents an intervention to, or departure from, what can be called national education, or education within a nation-state-centric frame. And, again, the intervention can be more of an *idealist* challenge to a nationalist/parochial education or it can be more about mobilities—teachers, students and educational programs that “travel” or connect *across* political borders. In the first sense, international education is to open up learners’ awareness of, and responsibility toward, more distant others in an interconnected, interdependent and fraught world. Developing a “world history” syllabus, re-vizing ethnocentric textbooks and “bringing the world” into the classroom through visual aids are all 20th century examples (Good, 2020; Tarc, 2009).

In the second sense, international education can refer to any kind of education that takes place outside of one’s country or that comes into one’s country from the outside. In the “developing country” context, international education often refers to a “Western education,” that can advance social mobilities under the conditions of global capitalism and the hegemonic status of colonial languages (especially English). And so coming to “study abroad” in a university in the West, irrespective of the curriculum or learning, represents “international education.” And providing (qualificatory and credentializing) systems that supports the mobilities of students, such as the Erasmus program for higher education in Europe, or the International Baccalaureate Diploma Program as a secondary school leaving diploma, also represent international education. And again, when these mobilities happen in educational contexts, the more idealist notions tend to surface. For example in the context of pragmatic IHE (i.e. for revenue generation), as Western university instructors respond to the backgrounds and needs of incoming international students by internationalizing curricula or employing culturally-relevant pedagogy or responding to students’ experiences of discrimination, the *idealist* goals of supporting intercultural dialogue, respect and understanding (or in today’s vocabulary—“diversity, democracy and social justice”) reveal themselves.

Global education

The modifier “global” in global education, at least in theory, is less tied to the category of the nation and to *political* borders, and signifies more “systems” or holistic approaches privileging “planetary” perspective, local-global relations and transdisciplinarity (see below). Global education often emphasizes a transdisciplinary or thematic approach to curriculum, appropriate to engaging inter-connected issues and challenges; accordingly it also tends to favor child-centred, “constructivist” pedagogical approaches. The “global” thus more loosely refers to learning about the “wider world” as opposed to the “international” which takes the level of the nation-state (within an international order) as its point of departure. Some educationalists place global education more centrally in the domain of k-12 schooling and international education in the higher education domain (Pike, 2015); this differentiation has likely been precipitated by the internationalization of higher education movement, largely driven by the recruitment of high-tuition paying international students and the branding of universities as “world class,” which has accelerated from the 1990s into the 21st century under neoliberalisation processes (de Wit and Altbach, 2021; Knight, 2003; Tarc, 2013a).

A number of authoritative typologies and reviews of global education are instructive for understanding contemporary meanings and uses of GE. These late-20th and early-21st century conceptions of global education illustrate distinctive emphases from international education, but, at the same time, there is much overlap. Although I do not by any means provide a comprehensive listing, the following small set of influential accounts well illustrates the main contours of *global education* as represented (primarily) in the Anglo-West. Robert Hanvey’s (1976) *An Attainable Global Perspective* continues to be an influential framing of global education by global education scholars and practitioners (Hicks, 2003). Hanvey (1976) identifies five dimensions of awareness and knowledge that are “requirements” for a global perspective as follows:

- *Perspective Consciousness*, underscores the need to recognize the concept that everyone’s perspective is shaped by subtle influences and that others may have different perspectives.
- *State of the Planet Awareness*, examines the problems and solutions for increasing the ability of individuals to intelligently interpret information about world conditions.
- *Cross Cultural Awareness*, describes the different degrees of cross-cultural awareness and the necessity to reach a stage beyond empathy where one has the capacity to imagine oneself in a role within the context of a foreign culture.
- *Knowledge of Global Dynamics*, analyzes the world as an interdependent system where the issue of growth may be the predominant contemporary problem.
- *Awareness of Human Choices*, emphasizes that increased global perspective will require difficult value decisions about the solutions to our world problems. (my emphases, p. i)

Derek Heater’s (1980) *World studies: Education for international understanding in Britain*, provides a comprehensive accounting of the post-WWII “World Studies” trajectory in Britain. Heater’s definition of “world studies,” as his overarching term, is stated as follows:

Studies that draw upon a number of disciplines, mainly in the field of the social studies, whose subject matter is not limited by any geographical restrictions, taking the whole world as the source of material for investigation. Programmes of study maybe interdisciplinary and holistic in organisation or composed of separate subjects or fields of study whose syllabuses are constructed on the basis of global scope. The selection of material from the great potential range may be determined by any or all of the following organizing principles:

1. Societies and cultures other than those of the student.
2. Issues and problems of current significance whose impact extends beyond a single continent.
3. The interrelatedness of life, events and institutions beyond the confines of any nation. The objectives of these studies include an appreciation of the need to view many issues in a global context, the enhancement of tolerance between peoples and a desire for a more just world society. (p. 5)

Heater wades through the “jungle of nomenclature” (p. 27) to identify and unpack six major strands of or variations to his generic term “world studies,” as follows: (1) the “traditional” education for international understanding [often termed “international education”]; (2) global education (favored in the US), (3) peace education, (4) development studies, (5) environmental education and (6) multicultural education. Heater provides a comprehensive UK-centred accounting of these various strands and unpacks their educational implications and manifestations. His book offers a weighty accounting of global/international education in the *pre-neoliberal* globalization era.

David Hicks (2003) provides a useful and more contemporary review of the UK global education movement across “thirty years.” He claims that the UK “variant of global education” (p. 266) began with the “world studies project” under Robin Richardson’s lead. From Richardson’s work, the “World Studies 8–13 project” became a “highly successful national curriculum project” that developed five themes for classroom inquiry, namely: “(i) ourselves and others; (ii) rich and poor; (iii) peace and conflict; (iv) our environment; (v) the world tomorrow” (Hicks, 2003, p. 267). Also influenced by Richardson, Graham Pike and David Selby “further developed the conceptual map of the field” (p. 267). In their *Global teacher, Global learner*, Pike and Selby (1988) first mark out “four dimensions of globality” as the *spatial*, the *temporal*, the *issues* and the *human potential* dimensions (cited in Hicks, 2003, p. 267). Additionally Pike and Selby (1988) adapt Hanvey’s model in their articulation of five “aims for global education” as follows: “Systems consciousness,” “Perspective consciousness,” “Health of planet awareness,” “Involvement consciousness and preparedness” and “Process mindedness” (cited in Hicks, 2003, p. 268).

Ultimately, Hicks states his belief that “it is possible to identify the core elements that are required for any endeavor to be labeled as global education” (p. 270). Hick’s model echoes the dimensional model of Pike and Selby and is condensed into the following “core elements:”

1. *Issues dimension*—This embraces five major *problem* areas (and *solutions* to them): inequality/equality; injustice/justice; conflict/peace; environmental damage/care; alienation/participation
2. *Spatial dimension*—This emphasizes exploration of the *local–global connections* that exist in relation to these issues, including the nature of both interdependency and dependency
3. *Temporal dimension*—This emphasizes exploration of the interconnections that exist between *past, present and future* in relation to such issues and in particular scenarios of preferred futures
4. *Process dimension*—This emphasizes a participatory and experiential pedagogy which explores differing *value perspectives* and leads to politically aware local–global *citizenship*. (p. 271)

Global educators will recognize the concepts, dimensions and issues represented in Hicks model. It is fairly open and able to accommodate specificities across time and space, such as new forms or “inequality/equality” or “alienation/participation.” The “process dimension” continues to be particularly salient for GE as does the desire that global education pedagogy “leads to politically aware local–global citizenship.” What is meant by “politically aware local–global citizenship,” in different spatial contexts and from different theoretical perspectives, is a key flash-point emerging in recent writing in GE, and one to which I will return.

Before continuing this skeletal trajectory of *ideas of global education* it is worth noting that such conceptions are formed *out of* the conditions of policy-making, and *from* the on-the-ground experiences and practices of educators. In the first sense, the larger political and cultural debates (such as the *A Nation at Risk* trope emerging in the early 1980s in the United States) enable and constrain attention to, and resourcing of, GE (Tarc, 2009). So global education *policy making* has been formed by these larger politics (Heater, 1980; Tye, 2014). Thus, one can see the rise and fall of GE policy initiatives across time in parallel to the ebb and flow of countries internationalism (Tarc, 2009). GE practices and research studies, in turn, are also enabled and constrained by the larger socio-political contexts and accompanying policy reforms.

This entry does not detail the growing (libraries of) resources and prescriptions for k-12 global education (Tye, 2014) or map how global education is taught in k-12. But, it is important to understand the *experiential* and *practice-based* foundations of GE. Enduring commitments to GE and practices of individual “champion” educators in k-12 and higher education settings have directly informed, and continue to inform, the models and heuristics of GE developed. In terms of lived experience, many of the influential 20th century global educator scholars, as well as current theorists of global citizenship education, have themselves articulated how their experiences—of cross-cultural immersion, encounters of war and injustice, study and teaching abroad, Peace Corp, North-South (colonially-inflected) relations—have informed their commitments to and scholarship in GE (for example, Becker in Kirkwood-Tucker, 2018; Jefferess, 2008; Merryfield, 2018; Selby, 2008; Tarc, 2013b). Tang and colleagues’ (forthcoming) analysis of the “making of distinguished global scholars” in the domain of US.-based global education is particularly instructive.

Kenneth Tye’s (2014) *Global education: A worldwide movement. An update*, brings us closer to the contemporary moment and attempts to illuminate the global education movement on a wider scale. Tye provides an update on his previous research study surveying how global education is understood and taken up more globally. In the original study (Tye, 1999), he encounters a number of limitations in gathering the data, but manages to collect at least some data from 24 countries. Beyond the non-comprehensiveness of the data collected, Tye (2014) remarks on the conception that “global education” is tied to Western coloniality and that some countries in the South might have opted out of participating in the original study for this reason. On his original study he reports (Tye, 2003):

The major finding from my study was that throughout the world, schooling is still seen as a major force in the building of national loyalties. This is true despite the inexorable movement toward regional and international cooperation and the growing interconnectedness of the global systems ... (p. 165)

In his update (2014) to “assess the [current] state of the field,” Tye adapts his methodological approach. He explains as follows:

Rather than repeat the cumbersome methodology of the previous study, and because there now is a substantial international literature about global education, it was thought that a meta-analysis of the field presented in the literature might better suffice to describe what is currently happening in the world with regard to global education. (p. 855)

In his synthesis of global education “as a world wide movement,” Tye presents a synthesized definition of global education and identifies three main “barriers” to “the advancement of global education.” On the former, he writes:

To begin with, and while there are some disagreements, there is a definition that is felt to be suitable by a large number of global educators, as follows:

- Global education involves learning about those problems and issues that cut across national boundaries, and about the interconnectedness of systems – ecological, cultural, economic, political, and technological.
- Global education involves perspective taking – seeing things through the eyes and minds of others.
- Global education involves taking individual and collective action for social justice and the creation of a better world.

In addition, it is generally agreed that issues to be considered are ecology and environmental studies and education for sustainability; human rights education; education for peace and conflict resolution; and multicultural/intercultural education. (p. 867).

Tye’s fairly comprehensive, and still relevant, definition/conception of global education will be familiar across multiple contexts. It includes (1) knowledge and awareness of international issues and of the interconnectedness of systems, (2) perspective taking and (3) investments and action toward making a “better world.”

From this short and partial review of 20th to 21st century global education discourse, one can see, again, how each author cited adjusts their definition in relation to previous definitions. One can see how key constructs, such as Hanvey’s (1976) “perspective consciousness,” become slightly altered or becomes subsumed as another category or meta-term. Thus, it is evident that these definitions overlap considerably (and with those of international education), despite changing vocabulary and evolving iterations of high-level frameworks. Conceptions of global education find altered emphases over time, reflective of the changing understandings of the world, of the nation-state and of education. One such turn-of-the-century change has been a more *overt* emphasis on social justice and demands that education more directly “make a difference” (Tarc, 2012). Reflecting on his edited collection, *The Bloomsbury handbook of global education and learning*, Douglas Bourn (2020) remarks:

GE ... can mean a number of different things but what ... [brings] them together here has been a common thread of *seeking global social justice*, of challenging dominant new-liberal orthodoxies and offering different solutions and ways of working. (my emphasis, p. 453)

While this reflection is based upon the individual foci of the handbook’s contributors, it is also the case that in GE discourses, there has been an intensification of rhetorical commitments to social justice, equity and “acting to make a difference”; this emphasis on equity reflects the larger mainstreaming of progressive discourse under liberalism and some versions of neoliberalism (Fraser, 2019; Tarc, forthcoming, 2015, 2012; Tye, 2014). Despite this more recent *explicit* emphasis on (global) social justice, one can also identify in 20th century global education discourse, an *enduring* set of topics, themes, analytic registers and orientations for education to be more outward looking, relevant and less parochial; to be grounded in understandings of evolving worldly conditions and human development; and, to be in pursuit, if circuitously, of making a “better world.”

In Heater’s (1980) and Hanvey’s (1976) high-level models, there is no explicit mention of pedagogy or the *process* by which learners are to acquire a “global perspective” or *study* “world studies.” One would have to look further in these texts to encounter their conceptions of pedagogical processes and tensions. Heater (1980) discusses, for example, how GE advocates had to navigate the effects of the larger culture wars periodically criticizing GE as indoctrination, unrigorous and/or unpatriotic. These criticisms and debates endure into the present day with evolving dynamics (Bunnell, 2012; Chalfant and Samuels, 2020; Veugelers, 2007). On the other hand, Selby and Pike (2000) and Hicks (2003) include the “process dimension”—“participatory and experiential pedagogy” (Hicks, 2003, p. 171)—as a key tenet of global education. This process dimension continues to be strongly accentuated by most contemporary advocates of global education as Oxfam, the International Baccalaureate Organization, Unesco, etc. The question of process or pedagogy is important in the sense that models, typologies and prescriptions for global education have to be translated and re-contextualized in school sites. The more ambitious goals of progressive or critical GE require much more than consolidating the content to be digested or following a new curricular recipe. In this sense GE frameworks are more entry points and guiding principles than “teacher proof” lessons. Alongside the “main barriers” to global education that I turn to next, an additional barrier is one of school pedagogy, how the progressive and critical idealist goals of GE are to manifest in classroom teaching and learning. I consider this additional tension or flashpoint in the discussion section.

The “main barriers” to global education of “nationalism, neoliberalism and confusion of meaning” that Tye (2014, p. 867)—among multiple others (Goren and Yemini, pp. 180–181; Jones, 2000; O’Sullivan, 1999; Pike, 2000, 2015; Tarc, 2013a, 2009)—identifies, are also instructive to understanding the current state of GE. “Confusion of meaning” is a concern; it can impede both inter-cultural/societal dialogue, inter-disciplinary collaboration as well as global education policy and curriculum development; however, having a fixed definition seems neither tenable nor desirable. Instead, one can identify and work with multiple, overlapping conceptions of global education *analytically* as a starting point for discussion and curricular development. In this entry, and elsewhere (Tarc, 2019, 2022) I have offered one type of approach.

In the end, new models, typologies or terms such as “glocal” can sometimes work as productive shortcuts to (in this case) signify the inter-penetrating relationality of the global and the local and how the national level mediates (or not); however, given the ongoing necessity of translations across linguistic, cultural and geographic contexts and ever-changing global dynamics we end up returning to more core, but also not fixed, concepts as the “global” and “national” always in need of further scrutiny (see also Sassen, 2006, 2009, on dis-assembling meta-categories). For example, both the Co-vid 19 pandemic and the emerging invasion of Ukraine are altering thinking on the *qualities* of the porousness of political borders and of the “post” of post-Westphalian

geopolitical order. In my perspective, the other two barriers Tye identifies are perhaps more salient and particular to GE; they are discussed below.

Variants of GE

Before concluding this section on terminology, it is worth mentioning a few other streams or “vocabularies” of GE. First, coming out of more psychological-based “adaptation” studies is the stream of “intercultural competence” that has generated considerable literature (Bennett, 2015; Savicki, 2008), including critique (Dervin et al., 2012; Marginson and Sawir, 2011). This stream has found traction in contexts of postmodern performativity (Lyotard, 1984) privileging, in this case, the *measuring* of “intercultural competence” (Hammer et al., 2003). Educating for “intercultural competence” can be seen as an antecedent to educating for “global competence,” which is the term of choice in the OECD’s recent, and influential intervention (Engel et al., 2019), *Preparing our Youth for an Inclusive and Sustainable World: The OECD PISA Global Competence Framework* (OECD, 2018). By including “global competence” within its Program for International Student Assessment (PISA) regime, the OECD brings a “harder” modality to the “soft” transnational governance of GE (Tarc, 2021). Another influential transnational actor in global education, the IBO, which has expanded well beyond its genesis in the international school sector (Tarc, 2009; Resnik, 2012), centers its pedagogical aspirations on the development of “international mindedness” (Hacking et al., 2016; IBO, 2017; Singh and Qi, 2013).

Coming out of the resurgence of cosmopolitanism in academic discourses in the 1990s (Cheah and Robbins, 1998) “cosmopolitan” has also become a modifier—of education, of the teacher and of learning—under heightened transnationalism (Hansen, 2011; Luke, 2004; Rizvi, 2009; Tarc, 2013b). Although a contested concept, “rooted” or “vernacular” cosmopolitanism can be understood as upholding a common and relationally-constituted world, without neutralizing difference or cultural rootedness (Appiah, 2006; Hall, 2002) nor over-privileging the nation-state level in determining human relations and subjectivities (Rizvi, 2009). In this sense, a cosmopolitan lens is potentially more nuanced for studying educational processes and effects in a more globalizing and globalized world. Rizvi’s (2009) discussion of the “epistemic virtues” of historicity, relationality, criticality and reflexivity as pedagogic modes of “cosmopolitan learning,” (CL) foregrounds the learner’s developing awareness of their “situatedness” in a hyper-connected and uneven world; CL represents a promising version of a (phenomenal and ethical) pedagogy for GE (Tarc and Budrow, 2022).

Additionally, versions of multicultural education and citizenship education are being re-articulated to respond to heightened global migrations and cultural flows and connections (Abu El-Haj, 2007; Banks, 2017; Davies et al., 2005); contributions in James Banks’s (2017) edited collection, *Citizenship education and global migration: Implications for theory, research, and teaching*, are salient examples. There are also important cautions on conceptions of scaling up citizenship and “responsibility to others” to the global scale, as well as critiques of conflating multicultural education and GE (Davies et al., 2005; Idrissi et al., 2020; Rorty, 1998).

Most broadly, there is a coalescing overlap between more justice-oriented global education (as stated previously) and less nation-state-centric conceptions of teaching for diversity, equity, anti-racism and social justice (see for example, Schecter and James, 2021; Shaklee and Baily, 2012). This overlap represents one significant *subset* of the larger “global turn” in education (Manion et al., 2011), which also includes more instrumental and neoliberal-inflected versions of GE (Andreotti et al., 2010; Andreotti et al., 2016); these neoliberal versions are particularly pronounced in policy statements calling for “global competencies” and “21st century” learning (i.e. OECD, 2018; OME, 2015). In this sense, GE has significantly expanded and multiple academic, NGO and practice-based discursive communities are engaged in defining and re-defining its contours and boundaries.

Arguably, the most notable *contemporary* variant of GE is global citizenship education (GCE), which has emerged as a new field of research (see Andreotti, 2014; Andreotti and de Souza, 2012; Goren and Yemini, 2017; Oxley and Morris, 2013; Pashby et al., 2020; Shultz, 2007). In relation to other terms, such as educating for “global competence,” GCE tends toward a more justice-oriented approach (Oxfam, 2006; Tarc, 2012; UNESCO, 2013). As with older versions of GE, GCE can be framed from different ideologies or perspectives, such as neoliberal, liberal, critical/social justice and post-critical (Andreotti et al., 2016; Tarc, 2012). Still, the use of the adjective “critical” to modify GCE, emphasizes that there are less desirable versions of GCE, such as those advancing neoliberal (education as “human capital development” and “as capital accumulation”) paradigms (Ledger et al., 2019; OECD, 2018; Ong, 1999; Tarc, 2021). GCE has been advanced in key policy statements as the UN Sustainability goals (UNESCO, 2016), as well as within state ministry policies (Franch, 2020).

Goren and Yemini (2017) highlight a significant gap between critical theorizations of GCE and how GCE is taken up in schools. More bluntly, research that qualifies what teachers are already doing and thinking about in the name of GE (as an evolving set of schooling practices in a dynamic world) is scarce (Merryfield, 1998; Tarc, 2011; Pashby et al., 2021). While there are important efforts to understand how schools are taking up GE, “educational content” or documented *intentions* at policy and curricular levels tend to be the “proxies” analyzed over classroom *practices*; a good example is the UNESCO (2019) commissioned study, *Educational Content up Close: Examining the Learning Dimensions of Education for Sustainable Development and Global Citizenship Education*. Here, the objects analyzed are limited to “national laws, policies and the intended curriculum” (p. 1).

The more specific gap between *critical* GCE theorizing and school/classroom practices highlighted by Goren and Yemini (2017) is particularly timely as calls intensify for education to more substantively respond to a world threatened by multiple crises; see UNESCO’s (2021) recent report *Reimagining our futures together: A new social contract for education* as a significant example. Multiple educational researchers are discussing this gap; they highlight, for example, how demands for critical (Marshall, 2011; Tarc, 2015), “heroic” (Lyons and Tarc, forthcoming), and “redemptive” (Estellés and Fischman, 2021) versions of GCE/GE fail to “fully tak[e] into account the [political, social and] economic contexts, and pedagogic realities and traditions within schools” (Marshall,

2011, p. 422). Consequently, what emerges is a pattern of *not finding* in schools the critical idealizations desired, and instead encountering the more neoliberal-individualist forms (as raising funds for charity), and/or encountering liberal/moral (vs. political) forms of GCE (Evans et al., 2009; Franch, 2020; Mundy and Manion, 2008; Tarc, 2015; Veugelers, 2010; Westheimer and Kahne, 2004).

This enduring gap is further precipitated by instrumental notions of teacher practice (as implementation) and the intended and unintended scholasticism of academic research. To address this gap, academic theories on the limits and (im)possibilities of critical GE ought to be grounded in empirical research that substantively attends to the “lifeworld”/pedagogical conditions of schools, classrooms and teacher praxes (Marshall, 2011; Tarc, 2011). Although the pedagogical conditions of schools and classrooms are structured by larger forces that need to be robustly *theorized*, schools are the sites where GE pedagogies happen and matter. Critical theories that ignore or “black-box” these sites risk producing idealizations of little relevance. *Grounded* research addressing the gap is very limited but emerging (for example, Pashby et al., 2021).

Unesco (2021) in the document cited above that calls for a critical “social contract for education,” are clearly acknowledging this gap and the need for research to be grounded in the conditions of school teaching and practitioner knowledge. They write:

Research and knowledge about the futures of education begins with the work teachers perform, and indeed, many of the elements of a new social contract for education may already exist in the transformative pedagogy many teachers are practicing. Teachers’ work as knowledge producers and pedagogical pioneers must be recognized and supported, assisting them to document, share, and discuss relevant research and experience with their fellow educators and schools in formal and informal ways. *Universities and higher education can imagine new institutional configurations that enable sustained research and professional relationships with teachers in support of their profession-wide knowledge production.* (my emphasis, pp. 150–151).

Indeed we must find multiple routes for understanding and steering the educational processes and practices that (might) be effectual in such calls for educational reform toward sustainable and just human and planetary futures. The next section identifies a few specific lacunae for further inquiry.

Critical faultlines and tensions of GE

Above I have presented an outline of older and newer conceptions/articulations of GE found in the English-medium literature as relevant to k-12 schooling. As the world (d)evolves and the literature grows it is difficult to develop generative syntheses of such an expansive area without giving up nuance and particularity. The above more general mapping situates and foreshadows this second, more particular contribution of my entry. In this final discussion, I revisit and briefly discuss a few *fault lines* and *tensions* of k-12 GE that stand out for me as a GE researcher and educator. GE is dynamic and embedded in larger social, political and economic conditions. It is an intervention (of one sort or another) to state or “traditional” education, carrying both progressive and critical visions that are entangled with normalizing (nationalist and statist) discourses and practical contingencies. Thus, GE has been and continues to be “dancing with diversity, democracy and social justice.” The faultlines and tensions presented in this section link up with Tye’s “barriers” of nationalism and neoliberalism mentioned above. They also arise from a consideration of the qualities and (aporetic) inter-relations of education, politics and the (motivations for the) making of a better world.

The tension between strategic and dedicated cosmopolitan sovereignties and subjectivities

Retrospectively reflecting on the highest aims of international education, founding Director General of IB, Alec Peterson (1987), emphasizes the altruistic over pragmatic “intention” of international education. He states:

[International education’s] intention is not simply to help the next generation to know better their enemies or their rivals, but to understand and collaborate better with their fellow human beings across frontiers. (p. 195)

His statement reflects the longstanding hopes of GE advocates; it also highlights an enduring and indeed intensifying tension on how IE is actually manifest in the world. GE’s altruistic intention has always been somewhat tempered by pragmatic considerations, particularly from states’ nationalist agendas (for example, “soft power”). However, in recent decades, this tension between the idealist and instrumental uses of international education has become especially heightened. With GE as expedient under neoliberalism, we have economic instrumentalities dominating and thereby constraining the idealist vision at both the levels of state sovereignty and citizenship. As Ong (2006) illustrates, both sovereignties and citizenships are stretching across territories (see also Sassen, 2006), as both citizens and states *flexibly* navigate their participation in local to global economies (Ong, 1999). States understand that citizens and diasporic groups can contribute to the “nation” regardless of where they reside (Rizvi, 2005); and, middle and upper class families engage in social class making on a global scale (Ball, 2010; Ong, 2006; Rizvi, 2007; Tarc et al., 2019; Wu and Tarc, 2019).

Although GE is fundamentally an intervention against nationalism’s “disfiguring” of education (Berle in Mayer, 1968, p. vi), it nonetheless is enacted in conditions structured by state nationalism. It is not surprising, thus, to find nationalist agendas at work in GE enactments (Leach, 1969; Sidhu, 2006; Tye, 2014). For example studies find GE in US social studies curricula imbued with “American exceptionalism” (Gaudelli, 2003) and IHE in Japan and Singapore as “being driven by nationally-centred rationales” (Sanders, 2019, p. 413). Thus governments forge their internationalist projects with national interests in mind. So again, while university presidents talk about developing “global citizens” who can have a positive impact in the world, state policies frame international students as revenue generation and potential human capital.

In terms of GE and its mainstreaming under the ascendancy of neoliberalism, many states have increasingly focused on governing subjects for economic competitiveness over “flag waving” patriots (Tarc, 2009; Green, 1997). In this sense, neoliberal ideology has increasingly taken over, at least in liberal/pluralistic contexts as the core “barrier” to GE and, in turn, has been an ongoing target of critique (Pike, 2015; Richardson, 2008; Tarc, 2012). The most basic problem of the neoliberal imaginary is much deeper than the GE’s dependency on pragmatic or financial logistics, but that the very meanings of education as well as the cosmopolitan goals of GE become transformed by economic/exchange logics (Gorur, 2016; Rizvi, 2007; Borg and Mayo, 2006; Tarc, 2012). More recently still, with the rise of populist nationalisms (Geiselberger, 2017) and the attendant challenge to progressive neoliberalism (Fraser, 2019), the longer-standing barrier of an “inward” nationalism has regained ground.

Individual users of GE are also implicated in the tension between “strategic” and “dedicated” cosmopolitanism (Mitchell, 2003). On the one hand, GE is to ideally foster open-mindedness, intercultural understanding, critical thinking and global responsibility in learners; on the other hand, GE is used as a form of cosmopolitan cultural capital that individuals can mobilize within their schooling-career trajectories for social advantage in competitive and hierarchically-formed societies. So this tension of the altruistic versus instrumental agenda of GE plays out on the ground at the most micro levels. Yet, at this ground level, these agendas might not *always* conflict where educators and students engage in GE with “eyes wide open” to its faultlines and tensions.

The demands on education to make a difference and considering the educational of education

Another main fault line emerging in GE (among other progressive and critical models of education) is the increasing demands placed on education to “make a difference.” Tarc (2015) examines how the “active” in global citizenship education is discursively situated in the conjunctures and disjunctures of critical/social justice desires, liberal/experiential pedagogies and neoliberal processes. From a (progressive) neoliberal frame he finds a seamless *conjuncture* with social justice rhetoric, active learning pedagogies and individualist/market forms of acting (see also Tarc, forthcoming). The *disjuncture* is strongest from a critical perspective, where education is expected to “make a difference” at the level of “root causes” through student acting as pre-political subjects (Tarc, 2015). With (Arendt, 2006/1954) and Topolski (2008) he questions whether one can *practice* active citizenship in the absence of citizenship rights, which thereby suggests schooling’s aim might be tempered by a focus on *learning* over “action” and that, therefore, “experiential education” in political citizenship represents a kind of aporia. Students are in and of the world and becoming subjects inside and outside of schools; from a post-humanist perspective, they also have political *effects* as “actants.” However, if they are to act as political *subjects*, being seen and heard in public (Arendt, 2006/1954), their citizenship education might be thus limited as a preparatory/future-oriented education.

Lyons and Tarc (forthcoming) also take up the question of how GE might be “make a better world” in the context of an IB education in elite international schools. They challenge both neoliberal conceptions as well as more critical but “heroic” conceptions. On the latter they suggest that despite the increasing demands for education to make a difference under ongoing and looming global crises, the “educational of education” remains a “weak” and existentially “risky” endeavor (Biesta, 2014; Di Paolantino, 2018; Rogoff, 2010; Tarc, 2007b). DiPaolantino (2018) warns against what a “heroic” and “strong language” of education can produce, as follows:

Believing that education miraculously will correct or offer the answer to problems which society itself has failed to properly remedy (such as ensuring economic prosperity for all, fixing unemployment, creating social justice), we quickly come to feel betrayed and turn away in disgust from education. Our “strong language” of education actually sets us up to judge education according to a language and set of criteria foreign to education, consequently impoverishing our thinking about the educational in education. (p. 55)

By “asking too much of education” we risk losing focus on what education might *actually* do and *how* (through what modalities or mechanisms). The response to endemic depoliticization and global crises is not to politicize education, but to study and enhance enactments of, and experiments in, education from critical/social justice orientations.

Conclusion: prospects for diversity, democracy and social justice

It is important to understand that the (more desirable) idealist/aspirational visions of international education are not expressed in social/political/cultural vacuums. Thus, GE has always been inflected by the larger nationalist and colonial imaginaries activated through state policy and normalizing discursive regimes (Good, 2020; Kapoor, 2008; Sidhu, 2006). What this means is that even more well-intended pedagogical acts of encountering, understanding, caring for and/or helping Others in inter-societal

contexts can be (and are) strained or undermined by colonizing/civilizing geopolitics and mentalities of supremacy, domination and control (Jefferess, 2016; Tarc, 2013c). For this reason, critical theorizations often name and critique these potentially colonizing manifestations of GE.

Indeed, beyond the neoliberal forms, making GE more critical/justice-oriented is a major thrust in academic and educational domains as I have emphasized. Further, critical theorizations are beginning to be translated into pedagogical resources. Andreotti's (2016) heuristic for GCE teachers, for example, orients teachers and students to be attentive to uncritical/neocolonial forms of engagement, othering and self-rationalization as signaled by the acronym HEADS UP: Hegemony, Ethnocentrism, Ahistoricism, Depoliticization, Self-congratulatory tendencies, Uncomplicated solutions and Paternalism (p. 108). For another example, Tarc (2013c) employs the parable of *The Elephant and the Blind Men* to outline a set of inherent difficulties for intercultural understanding. With HEADS UP, Andreotti is further reflexive in discussing the potential limits of these terms such as, how challenging “hegemony” ought not be aimed at “creating new hegemonies” (p. 109). Moreover, as Andreotti acknowledges, even as a critically-framed heuristic, GE reflective practitioners need to have autonomy in adapting and using the heuristic and need to reflect on its enactment in the scene of classroom pedagogy. She states: “the lessons that emerge from the failure and the limits of these devices are [also] essential for deepening our understanding of educational contexts and of the process of pedagogical articulation itself” (p. 107). HEADS UP, as Andreotti shares, comes out of a particular location addressing a set of specific colonially-inherited North-South relationalities. Therefore, the heuristic represents a generative curricular object for engagement by teachers and students rather than a set of new truths to digest. Notably, researchers have begun to examine the use of HEADS UP in school classrooms (Pashby et al., 2021).

Of course forward-looking, critical and justice-oriented pedagogies emerge from *within* “national” trajectories such as citizenship education and multicultural education. In many ways, GE overlaps with these trajectories, as stated above. In our current context, despite the many challenges inherent to enacting GE discussed thus far, what GE offers these similar educational trajectories/discourses is its worldly, intercultural orientation. The radical “deparochializing” impulse of GE can support researchers’ educators’ and learners’ potential attunement to wider and deeper relationalities; a multiplicity of criss-crossing but distinctive histories, Southern theories and alternative modernities and cosmologies (Grosfoguel, 2020; Mignolo, 2007).

Thus, GE can potentially attend to diverse diversities/racializations/racisms at a more inter-societal level and challenge methodological nationalism. Admittedly the possibilities for diversity, democracy and social justice are set against the weight of history, the dramatic inequalities between groups, and the instrumental and competitive/sorting uses of education. GE ultimately compels learners to engage and learn from multiple stories/perspectives (Adichie, 2009), as well as be reflexive of the social and epistemic locations from which we have and take a perspective (de Souza, 2012) and their connectedness to “elsewheres.” Finally, it compels us to consider how the inherited categories we employ might be globalized to challenge nation-centric methods, which fail to satisfactorily attend to the complex relationalities of the world, education and the human.

References

- Abdi, A.A., Shultz, L., Pillay, T. (Eds.), 2015. *Decolonizing Global Citizenship Education*. Brill, Leiden; Netherlands.
- Abu El-Haj, T.R., 2007. “I was born here, but my home, it’s not here.” Educating for democratic citizenship in an era of transnational migration and global conflict. *Harv. Educ. Rev.* 77 (3), 285–316.
- Adichie, C.N., 2009. The Danger of a Single Story. TED Conferences. Available at: https://www.ted.com/talks/chimamanda_ngozi_adichie_the_danger_of_a_single_story#t-15570.
- Andreotti, V., 2014. Soft versus critical global citizenship education. In: *Development Education in Policy and Practice*. Palgrave Macmillan, London, pp. 21–31.
- Andreotti, V., De Souza, L.M.T.M. (Eds.), 2012. *Postcolonial Perspectives on Global Citizenship Education*. Routledge, New York.
- Andreotti, V., Jefferess, D., Pashby, K., Rowe, C., Tarc, P., Taylor, L., 2010. Difference and conflict in global citizenship in Higher Education in Canada. *Int. J. Dev. Educ. Global Learn.* 2 (3), 5–24.
- Andreotti, V., Stein, S., Pashby, K., Nicolson, M., 2016. Social cartographies as performative devices in research on higher education. *High Educ. Res. Dev.* 35 (1), 84–99.
- Appiah, K., 2006. *Cosmopolitanism: Ethics in a World of Strangers*. W.W. Norton, New York.
- Arendt, H., 2006/1954. The crisis in education. Chap. 5. In: *Between Past and Future*. Penguin Books, New York.
- Andreotti, V., 2016. The educational challenges of imagining the world differently. *Can. J. Dev. Stud./Rev. Can. Etudes. Dev.* 37 (1), 101–112.
- Ball, S., 2010. Is there a global middle class? The beginnings of a cosmopolitan sociology of education: a review. *J. Comp. Educ.* 69 (1), 137–161.
- Banks, J.A. (Ed.), 2017. *Citizenship Education and Global Migration: Implications for Theory, Research, and Teaching*. American Educational Research Association, Washington, DC.
- Bennett, J.M. (Ed.), 2015. *The SAGE Encyclopedia of Intercultural Competence*. Sage Publications.
- Biesta, G., 2014. *The Beautiful Risk of Education*. Paradigm, St Paul, MN.
- Borg, C., Mayo, P., 2006. Learning and Social Difference: Challenges for Public Education and Critical Pedagogy. Paradigm Publishers, Boulder.
- Bourn, D. (Ed.), 2020. *The Bloomsbury Handbook of Global Education and Learning*. Bloomsbury, London.
- Bunnell, T., 2012. *Global Education under Attack: International Baccalaureate in America*. Peter Lang, Frankfurt.
- Chalfant, M., Samuels, B., 2020. Trump Announces Commission to Further ‘patriotic Education’. *The Hill*. Available at: <https://thehill.com/homenews/administration/516941-trump-announces-commission-to-further-patriotic-education>.
- Cheah, P., Robbins, B. (Eds.), 1998. *Cosmopolitics: Thinking and Feeling beyond the Nation*. University of Minnesota, Minneapolis, pp. 91–114.
- Davies, I., Evans, M., Reid, A., 2005. Globalizing citizenship education? A critique of ‘global education’ and ‘citizenship education’. *Br. J. Educ. Stud.* 53 (1), 66–89. Available at: <https://www.jstor.org/stable/1556020>.
- de Wit, H., Altbach, P.G., 2021. Internationalization in higher education: global trends and recommendations for its future. *Policy Rev. High. Educa.* 5 (1), 28–46.
- de Souza, L.M.T., 2012. Engaging the global by resituating the local: (Dis)locating the literate global subject and his view from nowhere. In: *Postcolonial Perspectives on Global Citizenship Education*. Routledge, New York, pp. 80–96.
- Dervin, F., Paatela-Nieminen, M., Kuoppala, K., Riitaoja, A.L., 2012. Multicultural education in Finland: renewed intercultural competences to the rescue? *International J. Multicult. Educ.* 14 (3).

- Di Paolantonio, M., 2018. ¿Puede el Propio Pensamiento de Educación Romper Esquemas? Análisis de 'Un Manifiesto para la Educación' de Biesta y Säfström./Can the Very Thought of Education Break Bricks? A Commentary on Biesta and Säfström's A Manifesto for Education. *Praxis Educativa* 22 (2), 47–57. <https://doi.org/10.19137/praxiseduca-va-2018-220206>.
- Engel, L.C., Rutkowski, D., Thompson, G., 2019. Toward an international measure of global competence? A critical look at the PISA 2018 framework. *Glob. Soc. Educ.* 17 (2), 117–131.
- Evans, M., Ingram, L.A., MacDonald, A., Weber, N., 2009. Mapping the “global dimension” of citizenship education in Canada: the complex interplay of theory, practice and context. *Citizen. Teach. Learn.* 5 (2), 17–34.
- Estellés, M., Fischman, G.E., 2021. Who needs global citizenship education? A review of the literature on teacher education. *J. Teach. Educ.* 72 (2), 223–236.
- Franch, S., 2020. Global citizenship education: a new ‘moral pedagogy’ for the 21st century? *Eur. Educ. Res. J.* 19 (6), 506–524.
- Fraser, N., 2019. *The Old Is Dying and the New Cannot Be Born: From Progressive Neoliberalism to Trump and beyond*. Verso Books, New York.
- Gaudelli, W., 2003. *World Class: Teaching and Learning in Global Times*. L. Erlbaum Associates, Mahwah, N.J.
- Geiselberger, H. (Ed.), 2017. *The Great Regression*. Polity Press, Cambridge.
- Good, K.D., 2020. *Bring the World to the Child: Technologies of Global Citizenship in American Education*. MIT Press, Cambridge, MA.
- Goren, H., Yemini, M., 2017. Global citizenship education redefined – a systematic review of empirical studies on global citizenship education. *Int. J. Educ. Res.* 82, 170–183.
- Gorur, R., 2016. Seeing like PISA: a cautionary tale about the performativity of international assessments. *Eur. Educ. Res. J.* 15 (5), 598–616.
- Green, A., 1997. *Education, Globalization, and the Nation State*. St. Martin's Press, New York.
- Grosfoguel, R., 2020. Decolonizing post-colonial studies and paradigms of political-economy: transmodernity, decolonial thinking, and global coloniality. *Transmodernity J. Peripher. Cult. Prod. Luso-Hisp. World* 1 (1), 2011. Available at: <https://escholarship.org/uc/item/21k6t3fq>. (Accessed 29 December 2020).
- Hacking, E.B., Blackmore, C., Bullock, K., Bunnell, T., Donnelly, M., Martin, S., 2016. *The International Mindedness Journey: School Practices for Developing and Assessing International Mindedness across the IB Continuum*. Department of Education, University of Bath, Bath, UK.
- Hall, S., 2002. Political belonging in a world of multiple identities. In: Vertovec, S., Cohen, R. (Eds.), *Conceiving Cosmopolitanism: Theory, Context, and Practice*. Oxford University Press, Oxford, pp. 25–31.
- Hammer, M.R., Bennett, M.J., Wiseman, R., 2003. Measuring intercultural sensitivity: the intercultural development inventory. *Int. J. Intercult. Relat.* 27 (4), 421–443.
- Hansen, D.T., 2011. *The Teacher and the World: A Study of Cosmopolitanism as Education*. Routledge, New York.
- Hanvey, R.G., 1976. *An Attainable Global Perspective*. Centre for War/Peace Studies, New York.
- Heater, D.B., 1980. *World Studies: Education for International Understanding in Britain*. Harrap, London.
- Hicks, D., 2003. Thirty years of global education: a reminder of key principles and precedents. *Educ. Rev.* 55 (3), 265–275.
- Hill, I., 2002. The history of international education: an International Baccalaureate perspective. In: Hayden, M., Thompson, M., Walker, G. (Eds.), *International Education in Practice: Dimensions for National & International Schools*. Kogan Page, London, pp. 18–35.
- Hughes, C., 2012. Child-centred pedagogy, internationalism and bilingualism at the international school of Geneva. *Int. Sch. J.* 32 (1), 71–79.
- IBO, 2017. *What is an IB education?* Cardiff, Wales: international Baccalaureate Organization. Available at: <https://ibo.org/globalassets/what-is-an-ib-education-2017-en.pdf>. (Accessed 3 September 2020).
- Idrissi, H., Engel, L., Pashby, K., 2020. The diversity conflation and action ruse: a critical discourse analysis of the OECD's framework for global competence. *Comp. Int. Educ.* 49 (1), 1–18.
- Jefferess, D., 2016. Cosmopolitan appropriation or learning? Relation and action in global citizenship education. In: *Globalization and Global Citizenship: Interdisciplinary Approaches*. Routledge, pp. 103–114.
- Jefferess, D., 2008. Global citizenship and the cultural politics of benevolence. *Critical Literacy: Theor. Pract.* 2 (1), 27–36.
- Jones, P.W., 2000. Globalization and internationalism: democratic prospects for world education. In: Stromquist, N.P., Monkman, K. (Eds.), *Globalization and Education: Integration and Contestation across Cultures*. Rowman & Littlefield, Lanham, MD, pp. 27–42.
- Kapoor, I., 2008. *The Postcolonial Politics of Development*. Routledge, New York.
- Kirkwood-Tucker, T.F., 2018. Civic education for a global nation: James M. Becker. In: Kirkwood-Tucker, T.F. (Ed.), *The Global Education Movement: Narratives of Distinguished Global Scholars*. IAP, pp. 1–36.
- Knight, J., 2003. Updated internationalization definition. *Int. High. Educ.* 33 (3), 2–3.
- Lapayese, Y.V., 2003. Toward a critical global citizenship education. *Comp. Educ. Rev.* 47 (4), 493–501.
- Leach, R.J., 1969. *International Schools and Their Role in the Field of International Education*. Pergamon Press, Oxford.
- Ledger, S., Thier, M., Bailey, L., Pitts, C., 2019. OECD's approach to measuring global competency: powerful voices shaping education. *Teach. Coll. Rec.* 121 (8), n8.
- Luke, A., 2004. Teaching after the market: from commodity to cosmopolitanism. *Teach. Coll. Rec.* 106 (7), 1422–1443.
- Lyons, J., Tarc, P. (forthcoming). How Might IB Classroom Pedagogy ‘Make a Better World?’ (Toward) Illuminating a Promising IBDP Teacher Praxis.
- Lytard, J.F., 1984. *The Postmodern Condition: A Report on Knowledge*. University of Minnesota Press, Minneapolis.
- Manion, G., Biesta, G., Priestley, M., Ross, H., 2011. The global dimension in education and education for global citizenship: genealogy and critique. *Glob. Soc. Educ.* 9 (3–4), 443–456.
- Marginson, S., Sawir, E., 2011. *Ideas for Intercultural Education*. Palgrave MacMillan, New York.
- Marshall, H., 2011. Instrumentalism, ideals and imaginaries: theorising the contested space of global citizenship education in schools. *Glob. Soc. Educ.* 9 (3–4), 411–426.
- Mayer, M., 1968. *Diploma: International Schools and University Entrance*. Twentieth Century Fund, New York.
- Méras, E.A., 1932. World-mindedness. *J. High Educ.* 3 (5), 246–252.
- Merryfield, M., 2018. Cross-cultural experiences in the making of a global educator. In: Kirkwood-Tucker, T.F. (Ed.), *The Global Education Movement: Narratives of Distinguished Global Scholars*. IAP, pp. 121–142.
- Merryfield, M.M., 1998. Pedagogy for global perspectives in education: studies of teachers' thinking and practice. *Theor. Res. Soc. Educ.* 26 (3), 342–379.
- Mignolo, W.D., 2007. Delinking. *Cult. Stud.* 21 (2), 449–514.
- Mitchell, K., 2003. Educating the national citizen in neoliberal times: from the multicultural self to the strategic cosmopolitan. *Trans. Inst. Br. Geogr.* 28 (4), 387–403.
- Mundy, K., Manion, C., 2008. Global education in Canadian elementary schools: an exploratory study. *Can. J. Educ.* 31 (4), 941–974.
- OECD, 2018. *Preparing Our Youth for an Inclusive and Sustainable World: The OECD PISA Global Competence Framework*. OECD, France. Available at: <https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf>.
- OME (Ontario Ministry of Education), 2015. *The Ontario Strategy for K-12 International Education*. Available at: http://www.edu.gov.on.ca/eng/parents/international_education.html.
- O'Sullivan, B., 1999. Global change and educational reform in Ontario and Canada. *Can. J. Educ.* 24 (3), 311–325.
- Ong, A., 2006. *Neoliberalism as Exception: Mutations in Citizenship and Sovereignty*. Duke University Press, New York.
- Ong, A., 1999. *Flexible Citizenship: The Cultural Logics of Transnationality*. Duke University Press, Durham, NC.
- Oxley, L., Morris, P., 2013. Global citizenship: a typology for distinguishing its multiple conceptions. *Br. J. Educ. Stud.* 61 (3), 301–325.
- Oxfam, 2006. *Education for Global Citizenship: A Guide for Schools*. Oxfam, London. Available at: <http://www.oxfam.org.uk/education/gc/>. (Accessed 15 August 2010).
- Pashby, K., da Costa, M., Stein, S., Andreotti, V., 2020. A meta-review of typologies of global citizenship education. *Comp. Educ.* 56 (2), 144–164.
- Pashby, K., Da Costa, M., Sund, L., Corcoran, S.L., 2021. Resourcing an ethical global issues pedagogy with secondary teachers in Northern Europe. In: *Teaching and Learning Practices that Promote Sustainable Development and Active Citizenship*. IGI Global, pp. 47–66.

- Pashby, K., Sund, L., 2020. Decolonial options and foreclosures for global citizenship education and education for sustainable development. *Nordic Journal of Comparative and International Education (NJCIE)* 4 (1), 66–83.
- Peterson, A.D.C., 1987. *Schools across Frontiers: The Story of the International Baccalaureate and the United World Colleges*. Open Court, La Salle, IL.
- Pike, G., 2015. Re-Imagining global education in the neoliberal age: challenges and opportunities. In: Reynolds, R., Bradbery, D., Brown, J., Carroll, K., Donnelly, D., Ferguson-Patrick, K., MacQueen, S. (Eds.), *Contesting and Constructing International Perspectives in Global Education*. SensePublishers, Rotterdam, pp. 11–25.
- Pike, G., 2000. Global education and national identity: in pursuit of meaning. *Theor. Pract.* 39 (2), 64–73.
- Pike, G., Selby, D., 1988. *Global Teacher, Global Learner*. Hodder & Stoughton, Sevenoaks, CA.
- Resnik, J., 2012. The denationalization of education and the expansion of the International Baccalaureate. *Comp. Educ. Rev.* 56 (2), 248–269.
- Richardson, G., 2008. Caught between Imaginaries: global citizenship education and the persistence of the nation. In: O'Sullivan, M., Pashby, K. (Eds.), *Citizenship Education in an Era of Globalization: Canadian Perspectives*. Sense Publishers, Rotterdam, pp. 55–64.
- Rizvi, F., 2009. Towards cosmopolitan learning. *Discourse: Studies in the Cultural Politics of Education* 30 (3), 253–268. <https://doi.org/10.1080/01596300903036863>. Available at:
- Rizvi, F., 2007. Internationalization of curriculum: a critical perspective. In: Hayden, M., Levy, J., Thompson, J.J. (Eds.), *The Sage Handbook of Research in International Education*. Sage, London, pp. 390–403.
- Rizvi, F., 2005. Rethinking brain drain in the era of globalization. *Asia Pac. J. Educ.* 25 (2), 175–193.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, New York.
- Rogoff, I., 2010. Turning. In: O'Neill, P., Wilson, M. (Eds.), *Curating and the Educational Turn*. de Appel/Open Editions, Amsterdam, pp. 32–46.
- Rorty, R., 1998. Justice as a larger loyalty. In: Cheah, P., Robbins, B. (Eds.), *Cosmopolitics: Thinking and Feeling beyond the Nation*. University of Minnesota, Minneapolis, pp. 45–58.
- Sanders, J.S., 2019. National internationalisation of higher education policy in Singapore and Japan: context and competition. *Compare* 49 (3), 413–429.
- Sassen, S., 2009. Saskia Sassen: digging in the shadows. In: Kenway, J., Fahey, J. (Eds.), *Globalizing the Research Imagination*. Routledge, New York, pp. 115–133.
- Sassen, S., 2006. *Territory, Authority, Rights: From Medieval to Global Assemblages*. Princeton University Press, Princeton, NJ.
- Savicki, V. (Ed.), 2008. *Developing Intercultural Competence and Transformation*. Sterling Stylus, VA.
- Schechter, S.R., James, C.E. (Eds.), 2021. *Critical Approaches toward a Cosmopolitan Education*. Routledge, New York.
- Selby, D., 2008. Designing transformation in international education. In: Savicki, V. (Ed.), *Developing Intercultural Competence and Transformation*. Stylus, Sterling, VA, pp. 1–10.
- Selby, D., Pike, G., 2000. Global education: relevant learning for the twenty-first century. *Convergence* 33 (1), 138.
- Shaklee, B.D., Bailly, S. (Eds.), 2012. *Internationalizing Teacher Education in the United States*. Rowman & Littlefield, Lanham, MD.
- Shultz, L., 2007. Educating for global citizenship: conflicting agendas and understandings. *Alberta J. Educ. Res.* 53 (3).
- Sidhu, R.K., 2006. *Universities and Globalization: To Market, to Market*. L. Erlbaum Associates, Mahwah, NJ.
- Singh, M., Qi, J., 2013. *21st Century International Mindedness: An Exploratory Study of its Conceptualization and Assessment*. Centre for Educational Research School of Education. University of Western Sydney, Australia.
- Stein, S., 2021. Reimagining global citizenship education for a volatile, uncertain, complex, and ambiguous (VUCA) world. *Glob. Soc. Educ.* 1–14.
- Stier, J., 2004. Taking a critical stance toward internationalization ideologies in higher education: idealism, instrumentalism and educationalism. *Glob. Soc. Educ.* 2 (1), 1–28. <https://doi.org/10.1080/1476772042000177069>. Available at:
- Tang, H., Tarc, P. (forthcoming). What Makes a 'Distinguished Global Scholar' in Global Education? Transformative Experiences Toward Global Mindedness.
- Tarc, P., Budrow, J., 2022. Seeking the cosmopolitan teacher: Internationalising curricula in a Canadian preservice teacher education program. *Teachers and Teaching: Theory and Practice*.
- Tarc, P. (forthcoming). The Capture of International Education by Progressive Neoliberalism: Illuminations, Qualifications and Educating beyond. In: Sharma, A., Wurzburg, E., Schmeichel, M., New York: Routledge.
- Tarc, P., 2022. Encyclopedia entry - international Baccalaureate: meanings and uses in a globalizing world (Volume: globalization and the geopolitics of education). In: Tierney, R., Rizvi, F., Ercikan, K., Smith, G. (Eds.), *International Encyclopaedia of Education, fourth ed.* Elsevier, Amsterdam.
- Tarc, P., 2021. Transnational governing for the pedagogical ideals of K-12 international education: contrasting PISA and IB. *Educ. Rev.* 74 (1), 93–113. <https://doi.org/10.1080/00131911.2021.1965095>.
- Tarc, P., 2019. Internationalization of education as an emerging field of study? A conceptualization of international education for cross-domain analyses. *Pol. Futures Educ. Internet* 17 (6), 732–744.
- Tarc, P., 2015. What is the 'active' in 21st century calls to develop 'active global citizens'? Justice-oriented desires, active learning, neoliberal times. In: Harshman, J., Augustine, T., Merryfield, M. (Eds.), *Research in Global Citizenship Education*. Information Age Publishing, Charlotte, NC, pp. 35–58.
- Tarc, P., 2013a. The rise of international education: expanded opportunities, new complications. In: Tarc, P. (Ed.), *International Education in Global Times: Engaging the Pedagogic*. Peter Lang, New York, pp. 1–18.
- Tarc, P., 2013b. Fostering cosmopolitan literacies: toward educating international education. In: Tarc, P. (Ed.), *International Education in Global Times: Engaging the Pedagogic*. Peter Lang, New York, pp. 97–120.
- Tarc, P., 2013c. The challenge of learning across difference: employing 'the elephant and the Blind men'. In: Tarc, P. (Ed.), *International Education in Global Times: Engaging the Pedagogic*. Peter Lang, New York, pp. 19–38.
- Tarc, P., 2012. How does 'global citizenship education' construct its present? The crisis of international education. In: Andreotti, V., Menezes de Souza, L.M. (Eds.), *Postcolonial Perspectives on Global Citizenship Education*. Routledge, London, pp. 105–123.
- Tarc, P., 2011. Bridging (representations of) the 'global present' with the life-world of the classroom: researching toward 'learning to teach' global citizenship education. In: Shultz, L., Abdi, A., Richardson, G. (Eds.), *Global Citizenship Education in Post-secondary Institutions: Theories, Practices, Policies*. Peter Lang, New York, pp. 64–75.
- Tarc, P., 2009. *Global Dreams, Enduring Tensions: International Baccalaureate in a Changing World*. Peter Lang, New York.
- Tarc, P., 2007a. What is the "international" of the International Baccalaureate? Towards a Periodization of IB in the World. ProQuest.
- Tarc, P., 2007b. Teaching for dispositions? Old debates, new orthodoxies: hanging onto a 'knowledge' approach. *Journal of Educational Controversy* 2 (2).
- Tarc, P., Mishra Tarc, A., Wu, X., 2019. Anglo-Western international school teachers as global middle class: portraits of three families. *Discourse: Studies in the Cultural Politics of Education* 40 (5), 666–681. <https://doi.org/10.1080/01596306.2019.1570635>.
- Topolski, A., 2008. Creating citizens in the classroom: Hannah Arendt's political critique of education. *Ethical Perspect.: Journal of the European Ethics Network* 15 (2), 259–282.
- Tye, K.A., 1999. *Global Education: A Worldwide Movement*. Interdependence Press, Orange, CA.
- Tye, K.A., 2003. Global education as a worldwide movement. *Phi Delta Kappan* 85 (2), 165–168.
- Tye, K.A., 2014. Global education: a worldwide movement. An update. *Pol. Futures Educ. Internet* 12 (7), 855–871.
- UNESCO, 2021. *Reimagining Our Futures Together: A New Social Contract for Education*. UNESCO, Paris.
- UNESCO, 2019. *Educational Content up Close: Examining the Learning Dimensions of Education for Sustainable Development and Global Citizenship Education*. UNESCO, Paris.
- UNESCO, 2016. *Education for People and Planet: Creating Sustainable Futures for All. Global Education Monitoring Report 2016*. UNESCO, Paris. Available at: <https://gem-report-2016.unesco.org/en/chapter/target-4-7-sustainable-developmentand-global-citizenship/>. (Accessed 10 May 2018).

- UNESCO, 2013. Global Citizenship Education: Preparing Learners for the Challenge of the 21st Century.
- Veugelers, W., 2007. Creating critical democratic citizenship education: empowering humanity and democracy in Dutch education. *Compare* 37 (1), 105–109.
- Westheimer, J., Kahne, J., 2004. What kind of citizen? The politics of educating for democracy. *American Educational Research Journal* 41(2), 237–269.
- Veugelers, W., 2010. Moral values in teacher education. In: Veugelers, W. (Ed.), *Glob. Soc. Educ.* vol. 9 (3–4), 473–485 (2011).
- Westheimer, J., Kahne, J., 2004. What kind of citizen? The politics of educating for democracy. *Am. Educ. Res. J.* 41 (2), 237–269.
- Wu, X., Tarc, P., 2019. Chinese international students in a Canadian private secondary school: becoming flexible citizens? *Compare*. <https://doi.org/10.1080/03057925.2019.1684242>.

Policy transfer in education

Jason Beech, School of Education, Culture and Society, Faculty of Education, Monash University, Clayton, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Travellers and reformers	151
Post-World War II: international organizations and human development	153
Global policy spaces: international large-scale assessments and the global educational market	155
Conclusions	157
References	158

Policy transfer has been a central theme in the systematic study of education since the first efforts at creating mass education systems, and continues to be a key issue in the fields of education policy and comparative and international education. On the one hand, policy makers and those in the business of advising them have always looked at ideas, policies and practices across their borders with a focus of finding practical solutions to the universal challenge of education. On the other hand, from a more analytic perspective, researchers have been trying to understand processes through which ideas about education (including policies, institutional forms, pedagogic theories and practices) move from one context to another and are transformed as they move. Policy transfer has also been associated with geopolitics and power in education, contributing to colonial and post-colonial epistemic domination through the construction and diffusion of hegemonic—mostly Western—views of knowledge and education.

In this chapter I will provide a historical overview of the theme of policy transfer from the 19th to the 21st Century with a focus on both practical and analytical orientations. I will start by examining the period of the travellers that in the 19th Century wandered around the world looking for examples and models that could assist them in the mission of founding mass education systems in their home nations. This style of doing comparative education continued well into the 20th Century and is still even widespread, for example in contemporary study trips to Finland and other PISA winners. However, a different form of doing policy transfer emerged in the second half of the 20th Century with the consolidation of international organizations. In the second section I will analyze how international organizations contributed to the creation of a global policy space in which ideas about what it means to be an educated person, and how education should be organized, financed and evaluated are being produced and reproduced, influencing local educational practices. In the third section I identify another discontinuity in modes of policy transfer with the creation of the Program for International Student Assessment and networked policy spaces. New actors, such as for-profit corporations, and some established ones, influence education policies and practices in different ways.

The division of the chapter into these three sections is purely analytical and a way of organizing the narrative into three periods that represent ways in which ideas about what it means to be an educated person and how education should be organized circulate and are transformed. In no way, the distinction between each section is clear cut and definitive in empirical terms. For each period I will mainly focus on three themes: who are the main actors that participate in processes of policy transfer, understanding the processes through which ideas about education move and are recontextualized, and the kind of theoretical approaches that were/are used in the literature to interpret policy transfer in education.

It is also important to note that this chapter is written from a particular perspective; one of a South American academic that has studied and now works in English speaking contexts. As this text engages with the global diffusion of discourses and power related to education, it is critical to consider that there is no such thing as universal common world (Latour, 2004). How we see and interpret the world and the idea of “the global” will be affected by perspective, historical experience and material conditions. For example, experiences of colonialism, post-colonialism and the diffusion of Western modern institutions have been very different in the Americas, when compared to Asia and Africa (Moraña et al., 2008). Thus, the picture that is represented in this text, would be very different if written by scholars coming from those historical experiences. Language of course is another lens, and my linguistic limitations force me to access and use only references in Spanish and English.

Travellers and reformers

Voyages have been a fundamental part of the process of creation of mass educational systems around the world. During the 19th Century many administrators and intellectuals traveled around the world (or some parts of the world) looking for models that could be used in their own nations to develop an education system. The world expansion of the Lancasterian school method is one of the first examples of international education policy transfer since the creation of nation states (Caruso and Roldán Vera, 2005). The method started with Andrew Bell in Madras in 1790 and Joseph Lancaster in London in 1798. It expanded rapidly to schools in Britain, Africa, Australia and the United States, and after the fall of Napoleon it was adopted in almost all Western European countries. The diffusion of the Lancasterian model can also be seen as an example of how processes of policy transfer relate to geopolitics and power relations in education. A model that was created in colonial India by a British educator was transferred to Europe, to post-colonial independent states in the Americas and to other colonial sites such as Africa. Yet, the pace of the

world-wide diffusion and adoption of the system was as impressive as was the speed at which it declined (Caruso and Roldán Vera, 2005).

It was also through travel and international exchanges that what we know as modern schooling and its particular grammar were created. These were times in which mass education systems were seen as key elements in the promotion of nationalism in Europe and in the newly independent states in the Americas. The rupture between the colonial system and independence in the Americas was quite weak in terms of the distribution of power among people with different epistemic traditions. Creole oligarchies replaced the Spanish and the Portuguese authorities, but the overall colonial social structure was continued, keeping indigenous peoples and epistemologies in subaltern and marginal positions (Dussel, 2008; Moraña et al., 2008). This was the case in the development of educational systems that tended to follow European models and portrayed European peoples and epistemes as superior in both sides of the Atlantic. In this way, the politics of colonialism influenced modern systems of education, as well as the relationship between them in most parts of the world.

Cousin in France, Mann in Massachusetts, Shuttleworth in England, Tolstoy in Russia and Sarmiento in Argentina (among many others) were pioneers of comparative education in the XIX Century. They were in most cases hired by their governments to travel around the world to find examples that could help in the creation of mass education systems in their home country (Beech, 2006). Based on linear notions of progress, these travellers and reformers believed in the evolution of educational systems. Thus, they thought that by borrowing policies and institutions from other more advanced systems they could avoid their mistakes and take advantage of their achievements in their own search for an ideal education system. For example, after his tour through some selected European countries, Horace Mann, in his *Report to the Board of Education of the State of Massachusetts*, noted:

if we are wise enough to learn from the experience of others, rather than await the infliction consequent upon our own errors, we may yet escape ... those calamities under which some other communities are now suffering. On the other hand, ... there are many things abroad which we, at home, should do well to imitate; things, some of which are here, as yet, mere matters of speculation and theory, but which, there have long been in operation, and are now producing.

Cited in Noah and Eckstein (1969, pp. 17–18).

Much of the work of these public intellectuals was used in their own countries as the basis for educational reform. The Guizot Law of 1833 that established the French system of primary education was based on Cousin's *Report on the State of Public Instruction in Prussia* (Noah and Eckstein, 1969). Similarly, Mann's Report was very influential on education policy in Massachusetts (Holmes, 1981). However, these comparativists were not always in favor of transferring solutions from other countries. Tolstoy, for example, was very critical of education in Western Europe and asserted that any imitation of European institutions and legislation related to compulsory education would be a step backwards and not forwards for Russia (Hans, 1963). Many of these educators and reformers also debated the need to adapt education institutions to their new contexts. Mann, for example, noted that "if Prussia can pervert the benign influences of education to the support of arbitrary power, we surely can employ them for the support and perpetuation of republican institutions" (Noah and Eckstein, 1969, p. 23). In this way, institutions and pedagogic practices were seen as potentially neutral technologies that could be used in different contexts with very different objectives and philosophies (Noah and Eckstein, 1969).

The assumption that there are practical lessons to be learned from the study of foreign education system continued to guide research and reform up to the present days, as exemplified in the case of the US by books such as Rickover's *Swiss Schools and Ours: why theirs are better* (1962), the United States Department of Education's *Report on Japanese Education Today* (1987), Stevenson and Stigler's *The Learning Gap: why our schools are failing and what we can learn from Japanese and Chinese education* (1993), and Chubb & Moe's *A Lesson in School Reform from Great Britain* (1992) that suggested that the US should follow the market oriented logic that Thatcher had introduced in the UK. The interest in the education system of Finland after its success in PISA is a more recent example of how research in education still follows an overall approach that believes in the feasibility of education borrowing as a strategy for reform and improvement.

However, a different perspective on policy transfer emerged in the field of comparative education and consolidated with the work of authors like Ushinsky in Russia and Sadler in England in the early 1900s. Sadler's (1979) famous lecture at the Guildford Educational Conference started with a question in the title: "How Far Can We Learn Anything of Practical Value From the Study of Foreign Systems of Education?" Using the metaphor of a garden he suggested that it was not possible to cut parts of different foreign plants and hope that when placed together in a home garden a new kind of plant would be born. What is especially relevant about Sadler is that he defined a more analytical approach to comparative education and policy transfer, moving away from purely practice-oriented aims. He stressed the importance of considering education practices and institutions as embedded in a given political, economic and cultural context and emphasized the relevance of historical forces in shaping these institutions. He argued that studying foreign systems of education was useful and important. Yet the benefit of this outward look was not practical but analytical. By studying other systems of education, it was possible to better understand one's own system.

In this way, the interest in processes of policy transfer in education was divided into two main strains. A more practical oriented approach aimed at finding solutions to educational challenges by looking at examples abroad, and a more analytic strain that aimed at understanding processes of policy transfer, how ideas about education moved and the effects they generated as they were transformed in different contexts. Overall, processes of policy transfer were interpreted as responding to the following pattern: (1) a local problem was identified; (2) solutions were sought in foreign educational systems; and (3) a "tested" institution or educational

practice (that had worked or was believed to have worked) was adapted to the new context and then implemented. (4) These processes occurred in a chronological order as described above (Beech, 2011).

These processes of policy transfer were embedded within the spatial geopolitics of colonialism, including the transfer of educational models aimed at those populations that in colonial classifications were considered to be inferior. Steiner-Khamisi and Quist (2000) examine how the Hampton-Tuskegee model of industrial education that was devised in the US as a mode of education for African Americans that lived in the segregated South was imported to Africa where the British relabeled the model “adapted education,” to focus it on agricultural and manual training of Africans.

The most salient debates in the literature, regarding the interpretation of processes of policy transfer have been based on the extent to which they contribute to a growing isomorphism of education systems across the world. From a mostly functionalist perspective, what is known as neo-institutionalist approaches argue that the creation of mass elementary schools (and further educational developments) does not respond to specific political processes in different countries, but rather to the emergence of a world culture. Meyer and Ramirez (2000) describe this process as the “world institutionalization of education”:

The logic of rationalized modern mass and elite education has always meant that a high degree of international homogenization was involved. Within the framework of a world society the process of becoming a nation-state and of competing with other nation-states led to the adoption of remarkably similar technologies, such as education. This process has produced pressures towards institutional isomorphism throughout the nineteenth and twentieth centuries (p. 130).

Based on this approach, quantitative studies analyzing data from many countries have found for example that primary school curricula are very similar in most parts of the world (Benavot et al., 1991).

From a different perspective, scholars based on culturalist approaches have criticized the interpretations of neo institutionalists, mainly arguing that by looking at data at a very high level of generality, they fail to capture processes of indigenization or recontextualization through which education policies, ideas and institutions are transformed as they interact with local cultures. Using qualitative methodologies, authors such as Schriewer (2016), Steiner-Khamisi (2014) and Anderson-Levitt (2003) acknowledge the existence of global models that influence educational policies and practices, but they argue that these models are re-signified, indigenized and transformed at national and local levels, and consequently they question the idea that education policies and practices around the world tend to isomorphism.

The debate on isomorphism versus recontextualization would gain much force in the second half of the 20th century as international organizations were created and assumed the role of promoting educational policies and ideas globally.

Post-World War II: international organizations and human development

During the postwar period processes of policy transfer in education were significantly affected by the creation and consolidation of international organizations such as UNESCO, the World Bank and the OECD, as they became key actors in promoting ideas about education. The notion of development, post-colonial processes in Africa and Asia and the Cold War influenced the geopolitics of policy transfer in this period, while research in policy borrowing in education focused on the work of these organizations, on new postcolonial theoretical approaches and, later, on understanding processes of globalization.

UNESCO was created after the War with the motto “since wars begin in the minds of men, it is in the minds of men that the defenses of peace must be constructed.” As an organization for culture and education, its involvement with education policies and the transfer of educational knowledge has always been at the center of UNESCO’s agenda (Mayor and Tanguiane, 1997). The original aim of the World Bank, founded in 1944, was to provide loans to countries that had been stricken by the War, mostly for infrastructure. As this mission became redundant, it shifted its aim toward development and poverty alleviation, focusing on poor economies. Since its first loan for education in 1963, the World Bank became in the 1990s the largest single source of external financing for education in “developing countries” (World Bank, 1995, p. 14). However, the Bank’s role is not only financial. It also defines itself as a “Knowledge Bank” that creates, shares and applies knowledge about education (and other social policies). Furthermore, the Bank also promotes a certain vision of education when it associates its lending programs with certain conditions. Thus, when “client countries” receive a loan for educational purposes from the World Bank, this act is not only a transfer of funds, but it is also a policy transfer. The OECD was also created to administer American and Canadian aid under the Marshall Plan for reconstruction of Europe after World War, and later shifted its focus. One of the most important objectives of OECD is to search for certain universal values, rules and policies and promote them among member and non-member countries. However, the OECD has no prescriptive mandate over its member countries (Papadopoulos, 1994). As part of this role, the OECD positions itself as an organization that can understand emerging challenges and identify and promote the kind of policies that should be used to tackle them. Thus, the “transfer” of “cutting edge” educational knowledge has been for many decades one of the main self-proclaimed roles of UNESCO, the World Bank and OECD.

The creation and expansion of these organizations was strongly linked to the notion of development applied to social issues that became fundamental to understanding the geopolitics of education and the expansion of global policy spaces in education in the Post War period. In 1949 President Truman presented a vision for international development that would become a landmark for the work of international organizations.

... we must embark on a bold new program for making the benefits of our scientific advances and industrial progress available for the improvement and growth of underdeveloped areas (...) I believe that we should make available to peace-loving peoples the benefits of our store of technical knowledge in order to help them realize their aspirations for a better life.

Truman (1949).

The notion of development became a potent discursive construction in the geopolitics of policy borrowing in education and provided a form of continuation to the logic of colonization (Escobar, 1995). The binary developed/underdeveloped became a key feature in the circulation of knowledge and power in education, as countries were classified as belonging to one of the two poles, which implied being on the donor or the recipient side in the flows of education policy transfers. At the same time, post-colonial processes in Africa and Asia meant that many newly created countries were faced with the economic and technical challenges of creating their state bureaucracies, including their education systems. It was international organizations (and some national development agencies, such as USAID) that led the efforts to promote mostly Western technical knowledge about how to administer education to developing countries. In this way, Western education systems became a norm and a model for the newly independent nations and contributed to the perpetuation of colonial domination through other forms.

The agenda focused on promoting macro level organizational techniques, such as educational planning. The recipients were mostly high-level officials from the Ministries of Education that were offered technical reports and participated in training and regional or international meetings. By promoting developmental and technocratic views, international agencies were advocating a number of abstract universal social technologies that—in the logic of these agencies—could be used to improve education in most contexts (Beech, 2011).

Another relevant geopolitical force during the period was the Cold War. The two superpowers competed for hearts and minds around the world in a cultural war that was fought through propaganda, cultural influence and through education. The Soviet Union, for example, influenced the education systems of China and Cuba, while the US had a strong influence in the development of education in Brazil that, for example, adopted the model of the Comprehensive High School (Cowen and Figueiredo-Cowen, 1989). While the superpowers were fighting to control the world militarily and politically, education also became one of the arenas of the dispute. For example, the launch of the Sputnik by the USSR was conceptualized as an educational crisis in the US. The education system of the Soviet Union attracted the attention of US researchers, as illustrated by Trace's (1961) "What Ivan knows that Johnny doesn't." However, the "solution" to the lack of creativity promoted in schools was found in the child-centered English Infant Schools that became a source of inspiration for school reforms in many states in the US (Ravitch, 1983).

With the end of the Cold War, the notion of globalization gained the spotlight and had a significant influence in further legitimizing the work of international organizations in education. The late 1980s and early 1990s were read as a moment of enormous social change that required a substantial reengineering of education systems. Technological changes such as the ubiquitous diffusion of the personal computer and the incipient Internet, post Fordism and knowledge economies, postmodernity and respect for diversity were some of the drivers of changes in education, alongside the fall of the Soviet Union. The idea that education systems had been created for another era and should be completely reinvented was widespread. In this context, international organizations positioned themselves discursively as the experts that could predict future changes and offer a recipe for education systems to adapt. On the other hand, the notion of globalization implied that all educational systems were facing similar challenges and that solutions should also be global in nature, reinforcing the discursive power of international organization.

Giddens's (2000) lectures that were published under the title "Runaway World: How globalization is shaping our lives" captured the spirit of the times. In response to a rapidly changing world that was seen as out of human control, international organizations promoted a "Universal Model of Education for the Information Age" (Beech, 2011). The logic was that since global changes could not be controlled, the best that educators could do was to be flexible and adaptable to the new social conditions, educating young people to also be flexible and adaptable to change. Lifelong learning was seen as a fundamental technology for people to be able to learn and relearn as would be necessary to respond to economic changes, themselves driven by technological advancements. Overall, the model was based on promoting decentralization and school autonomy, lifelong learning, a curriculum based on the development of competencies, child centered pedagogies that redefined teachers as facilitators of learning, and standardized evaluation systems based on performance indicators that connect the policy producing center with schools and teachers that should be granted more autonomy to attain pre-defined goals.

Although this model was presented as being universal, it was devised from a Western epistemic and political perspective with little attention to non-Western forms of knowledge and ways of life. This model, which with some variations was promoted by UNESCO, the World Bank and the OECD, became a norm that could be used to identify the problems with education systems all over the world and, at the same time, it offered the solution to those problems (Beech, 2011). Another significant shift for understanding processes of policy transfer is that the model did not only provide a template for changing educational governance at a macro level, but also focused on prescribing quite detailed curricular changes and pedagogical approaches, a shift in focus that would be intensified in the future.

The emergence of international organizations as key actors in promoting policy transfer in education had a strong impact in the ways in which policy ideas travel and are recontextualized. It also generated different kinds of interpretations in the literature.

Noah and Eckstein (1969), for example, defined the initial efforts of international organizations as an advancement toward what they envisaged as a scientifically legitimated comparative education: “The work of these organizations is in the hands of specialists. Thus, what began as philanthropy has ended with professionalism” (Noah and Eckstein, 1969, p. 82). Others were much more critical. Based on *dependencia* theories stemming from South America, Carnoy’s (1974) *Education as Cultural Imperialism* saw international organizations as mechanisms of imperial domination that promoted capitalist hierarchies. A similar logic was followed by Altbach and Kelly (1978) in *Education and Colonialism*. These texts, among others, introduced the problem of agency into discussions about educational transfer. They implicitly distinguished between voluntary and non-voluntary policy transfer, depending on who was the prime agent of the process and brought into the fore the ethical issues of colonization and post-colonial dominance in the circulation of knowledge about education around the world.

The work of international organizations in education became the focus of much research since the 1990s, as they gained power as central actors in the globalization of education. Overall, studies were focused on understanding how these organizations contributed to reshaping educational policy spaces and the flow of power in education (Ball, 1998; McNeely, 1995; Shahjahan, 2012), on analyzing the work of one of these agencies or the interactions between them (Jones, 2006; Sellar and Lingard, 2013; Beech, 2009b; Bonal, 2002), and on examining the processes through which the recommendations of these agencies were recontextualized in different settings (Waldow et al., 2014; Ydesen, 2021). Many of these approaches had a very critical view of the ways in which these organizations imposed educational principles and perspectives in non-Western contexts (Shahjahan, 2016; Sguissardi and Amaral, 2000).

What is clear is that the involvement of international organizations in promoting education policies implied a significant change in how ideas about education traveled. The first, and most noticeable, shift away from trans-national patterns of policy transfer is that policies do not move from one educational system to another, but from international agencies to specific educational systems. International organizations are abstracted from actual educational policies and practices. They aim at identifying a series of universal educational principles that could be applied in most education systems to improve education. Thus, what is transferred is not a tested education policy, but rather a set of abstract universal social technologies (Beech, 2011).

The second shift in patterns of policy transfer is linked to the sequencing of the process. In trans-national patterns of educational transfer, following a chronological order, a problem was identified, then a solution was sought in a foreign system, and finally a “tested” social technology was transferred. When international organizations promote certain educational principles, they simultaneously define the problems of an educational context and offer the solutions to these problems. So, for example, what I have labeled the Universal Model of Education for the Information Age does not only include the solutions to most educational problems, but it also defines the problems in the education systems to which it travels. This abstract universal model is offered as a norm against which the adequacy of existing educational policies in a given context can be measured. Once the “problems” have been identified, there are only a limited number of themes that can be debated, and a limited number of policy options that can address them (Beech, 2011).

Another important shift to consider, is that since the 1990s international organizations included within its focus much more detailed prescriptions related to curricular content and pedagogic strategies. This trend toward a focus on the classroom and the work of teachers would be further intensified in the 2000s.

Global policy spaces: international large-scale assessments and the global educational market

The next discontinuity in processes of policy transfer in education can be associated with the creation of International Large-Scale Assessments (ILSA), the consolidation of global policy spaces in which new kind of actors participate in promoting the global circulation of ideas, and, among these new actors, the involvement of for-profit corporations that compete for financial benefits in what they see as a global educational market.

The creation of ILSAs contributed to a systematic reference to compare education systems based on student learning outcomes. Although the first ILSAs started in the 20th Century it was the launching of the OECD’s Program for International Student Assessment (PISA) that had the biggest impact on changing the landscape of global policy spaces in education. The first ILSAs assessed students in terms of their local curriculum, and consequently the comparison between countries did not necessarily show that students in one country had learned more or better than the other. PISA assesses students in terms of a universal definition of what a student should have learned at age 15 to be prepared to fully participate in the knowledge economy. In this way, PISA became a universal measure of education quality that was accepted by educators, political leaders and researchers as an objective indicator of education quality, validating international comparisons and global rankings.

The position of an education system in the PISA ranking became a fundamental indicator of the quality of that education system and in this way, winners and losers were identified (Novoa and Yariv-Mashal, 2014; Kim, 2017). This had several effects in terms of processes of policy transfer in education. The notion of “PISA shock” was coined to describe the situation generated by lower than expected results in places such as Germany (Waldow, 2009) and Japan (Takayama, 2008, 2012) among many others (Pons, 2012). The reaction, in many cases, was to conceptualize PISA winners, most notably Finland, as the solution to the educational problems that PISA had made visible. The other side of the crises was the portraying of the winners, with Finland in particular as the poster boy that attracted massive attention from educational researchers and policy makers (Baroutsis and Lingard, 2017). Educational Tourism to “go and see” Finnish education to distil and copy their “best practices” was reminiscent of the times of the travellers and reformers of the 19th Century.

Through its global rankings, PISA also contributed to the further consolidation of the performative logic (Lyotard, 1984) in global education. As the desired output of education was standardized around the world and defined in terms of a numerical measurement, education policy was recast as a series of interventions to improve these outcomes. Although it is clear that even with all its technological sophistication PISA cannot measure all aspects that would constitute a good education, the improvement of PISA scores is increasingly being defined as the aim of education. In this sense, as PISA and its rankings gained legitimacy it became a driver of educational reforms aimed at aligning curricular definitions and assessment practices with its standardized tests. Efficiency became an end in itself. The curriculum focuses on the subjects that are assessed in PISA, and pedagogic practices are aimed at training for the test.

Through PISA, the OECD has become the main reference in global educational debates, promoting a market-oriented logic that places individuals as the focus of educational strategies and narratives about human progress and global understanding. Furthermore, springing from the success of PISA, the OECD has created new projects to extend its influence to other countries through the creation of PISA for Development (Addey, 2017; Auld et al., 2019), and to other areas of education with an assessment for teachers (TALIS), one for five year-old children to assess the quality of early childhood education (Urban, 2017), and the Program for the International Assessment of Adult Competencies (PIAAC) (Schleicher, 2008). Yet, in parallel to the growing influence of the OECD, new types of actors started to get involved in the business of promoting ideas about education globally.

Global policy spaces in education have become a complex array of competition and collaboration between different institutions, such as international organizations, private for-profit corporations, global philanthropies, NGOs, consultants, regional blocks, universities that participate in different networks with the aim of influencing education policies in new ways (Beech, 2009a). Policy consultancy, the provision of educational services, evaluation, learning technologies and pedagogical consultancy have all become commoditized. Knowledge on how to improve education is produced in order to be sold and is bought in order to augment performance in standardized tests. Alliances between international organizations, for profit corporations and philanthropies further contribute to the marketization of education, and to blurring distinctions between public/private, for profit/non-for profit and global/local.

The case of Pearson can illustrate some of the complexities of global policy spaces. Pearson is the biggest global corporation in education (Santori et al., 2016). It defines itself as the world's leading learning company. It has more than 20,000 employees and provides its services in more than 20 countries. One of the many services it offers is education policy consultancy to governments and pedagogic materials and training for schools. Pearson hired Michael Barber as its lead education adviser. Barber had previously worked for McKinsey (another for profit consultancy) and was responsible for reports (Barber and Mourshed, 2007; Mourshed et al., 2010) that became very influential in education policy in many parts of the world. Prior to working for McKinsey, Barber had been "Prime Minister's Chief Adviser on Delivery" during the Blair government in the UK (Coffield, 2012). In 2014, Pearson partnered with the OECD to develop the technical aspects of PISA. This partnership improves the market value of its products. If Pearson is developing the PISA tests and governments measure their educational success in terms of their performance in these tests, policy solutions sold by Pearson become more attractive as they are the ones who "know how to do it."

Low fee private schools constitute another business opportunity into which Pearson has invested. It created the Pearson Affordable Learning Fund (PALF) with the vision "to help millions of children in the world access a quality education in a cost effective, profitable and scalable manner" (Pearson, 2016). PALF is a venture fund through which Pearson invests in local edupreneurs in countries such as the Philippines, Nigeria, Ghana and South Africa, with a focus on high quality, for-profit education solutions for the low-income segment in the developing world (Santori et al., 2016). Chains of low fee private schools for the poorest populations in the world have become a growing feature of the globalization of education and the transfer of education models across space. The ethical implications of these initiatives have been heatedly debated in the literature (Srivastava, 2016; Nambissan and Ball, 2011; Edwards et al., 2017; Tooley, 2013), and in global civil society (OXFAM, 2019). Some of the ethical contradictions of these moves can be seen in the voice of Pearson itself, as it notes that the creation of the PALF is based on a "moral imperative" to improve the quality of learning in the poorest regions of the world and contributing to the achievement of the UN Sustainable Development Goals. This moral imperative was addressed through the creation of a for profit venture fund in response to what the company sees as a "vital market opportunity" (Santori et al., 2016). The discursive rationality behind this approach is clear in the words of Pearson's CEO:

We believe that a free-enterprise model of low-cost schools—schools affordable for many of the poorest—may be the best chance to provide both benefits.

Cited in Santori et al. (2016).

The logic is that for a social problem to be solved, such as access of the poorest populations of the world to education, it has to be commodified to attract investments searching for profit. Yet, low fee private schools for the poor are not only promoted by Pearson. They are part of a complex network in which the boundaries between international organizations, philanthropies, for profit corporations and entrepreneurs and, sometimes, state institutions are difficult to delimit (Ball and Junemann, 2012). These global policy spaces are dominated by market ideologies that permeate the definition of education itself. In this way, the market-oriented, individualized logic of an education for economic competition dominates education policy debates, while it also perpetuates and extend geopolitical differences and post-colonial practices in education.

Another relevant change in processes of policy transfer in education and in the influence of global actors on education policies is the shift toward a focus on the classroom and teaching practices. McKinsey and the OECD have been very influential in promoting the idea that teachers are the most important factor in improving the quality of education (Barber and Mourshed, 2007; Mourshed et al., 2010). This view had a strong impact on education policies all over the world, but also at the level of global actors themselves. The OECD is producing lesson plans for teachers (Vincent-Lancrin et al., 2019) and the World Bank (2019) offers a classroom observation tool to assess how teachers use time and the kind of classroom culture and socio-emotional skills they promote. What is visible is a further move toward a focus on pedagogic practices in global actors that aim at influencing schools and teachers directly, even sidelining state institutions. The creation of PISA for Schools, a version of the conspicuous OECD assessment that can be contracted by individual schools, is another example of this kind of move (Lewis, 2017).

These major developments in global policy spaces have attracted the interest of researchers and have resulted in rich empirical and theoretical contributions to the understanding of the complexity of global influences in education.

One of the most prolific areas of research that benefited from the availability of PISA data that can be used to compare the quality of education outcomes across space and time is the Economics of Education. Furthermore, PISA data in conjunction with other socio-economic indicators has been used in efforts both to explain why certain education systems perform better than others (Hopfenbeck et al., 2018), and to analyze how the quality of education is related to other outcomes, such as economic growth (Hanushek and Woessmann, 2007). These works have been quite influential in legitimating educational reforms and much of the proposals of the World Bank and the OECD. However, both the search for factors that explain success and the causal inference that connects education attainment with economic performance have been object of criticisms for making lineal unfalsifiable claims (Auld and Morris, 2016), and for having statistical problems (Komatsu and Rappleye, 2017).

PISA has also become a central object in the fields of comparative education and education policy. It is not easy to map all of the different approaches to PISA in the literature, but overall, some authors have aimed to understand how PISA and other ILSAs impacted in modes of governance in education (Sellar and Lingard, 2013; Gorur, 2014), while others have criticized PISA mostly in relation to its rankings and the ways in which the OECD and specific governments have used politically the results of the exams (Takayama, 2008; Gorur and Wu, 2015; Waldow, 2009).

Discussions of the new spatialities of education and the search for new theoretical approaches to understanding the circulation of knowledge and power about education globally have also flourished. The use of relational notions of space that see space not as a container of social processes, but rather as being constructed and reconstructed by social interactions has been an important development, complemented by approaches based on the mobility turn (Urry, 2012) and cosmopolitan sociology (Beck, 2006). Overall, the authors following these perspectives have made an effort to open up the notion of globalization and understand the complexity of power relations in the educational field. One of the most important moves in this respect has been the awareness of the need to rethink borders and binary definitions such as global/local, macro/micro, public/private, among others.

In rethinking education space, the notion of networks has been at the center of the literature, both in terms of theoretical and methodological approaches. So, for example, the idea of networked governance has been widely used to explain the ways in which state actors link with non-state actors to govern education (Ball, 2009; Lingard and Sellar, 2013). These actors connect through networks that have been described as heterarchical—horizontal and unstable networks that are contrasted with more stable and hierarchical relationships in bureaucratic logics of government. Methodologies such as social network analysis and network ethnography have been used to map and interpret flows and movement in global education policies (Avelar and Ball, 2019). Actor-Network Theories and other approaches related to Science and Technology Studies have also been used in very different ways to explore the global governance of education and its interaction with local education policies and practices, adding an emphasis on materiality and the heuristic device of assemblages to understand power and change (Gorur, 2011; Beech and Artopoulos, 2021; Piattoeva, 2018). A final line of prolific and very productive line of work has been based on a critical political economy approach to the globalization of education (Novelli et al., 2014; Robertson and Dale, 2015).

Conclusions

In this chapter I have shown how educators around the world have collaborated through processes of policy transfer in sharing ideas on how to better educate the younger generations. Processes of knowledge sharing have taken different forms throughout history, but overall have contributed to the construction of certain global common ways of approaching education.

From travellers making long journeys to find the holy grail of reason and science used for building mass education systems as a mode of government, to Pearson and its associates addressing the moral imperative of educating the poorest population of the world while making monetary profit. What is clear is that the movement of ideas about education has followed the fault lines of geopolitical power: From religious institutions in pre-modern periods (not covered in this chapter), to colonialism, the modern Nation-State and to global for profit corporations in the 21st Century.

As much as policy transfer can be seen as a process of collaboration, it is also a mode of competition and colonization. Global policy actors sometimes collaborate in a complex array of fluid and unstable networks, but they also compete to influence education policies in different places. Furthermore, through mechanisms such as PISA, some of these global actors promote competition among education systems that see their position in the global ranking of education quality as a measure of their failures and success. For example, in a quite remarkable statement, the 2013 Australian Education Act defined the objective to be in "the top 5 highest

performing countries based on the performance of school students in reading, mathematics and science” by 2025 (cited in Gorur and Wu, 2015).

The rationale for promoting competition within and among education systems is well known in the market oriented logic that was introduced in the governance of education in the dictatorship of Pinochet in Chile and the governments of Thatcher and Reagan in the UK and US. This logic that has now become dominant in global policy spaces assumes that competition results in overall improvement.

However, at times in which humanity is facing huge challenges, such as the climate crisis, growing inequalities, the weakening of democracy, growing military conflicts and the covid pandemic, it seems that a push to be more effective and efficient at delivering the same kind of education could be at least problematic. For example, countries with higher scores in PISA are also the ones with the most carbon emissions (Komatsu and Rapple, 2018). Correlation does not imply causality, but in this case it might be an indication that delivering faster and more efficiently the current education paradigm, itself associated with a non-sustainable mode of living, might not be the best way to address these challenges.

On the contrary, to successfully address the challenges we are facing we will need to collaborate both within our communities and at a global scale. That is why we need to promote a kind of education in which children and young people learn to live and collaborate with others that are different, rather than promote competition and individualization as the reigning values in education. In doing so, it might be important to pay attention and draw upon non-Western approaches to knowledge and education.

With this in mind, UNESCO created two years ago the Futures of Education initiative that has recently launched a major report with the title *“Reimagining our futures together: a new social contract for education?”*. The report seeks to promote a global debate on how knowledge and learning can shape the future of humanity and the planet. It is based on the conviction that education does more than respond to a changing world. It has the potential to transform the world. The report is both an attempt to regain UNESCO’s leading position in global educational debates and to reaffirm its own historical more humanistic perspective to balance the dominating market driven logic. The report defines education as a public endeavor and a common good while it appeals to an education based on the principles of cooperation, collaboration and global solidarity instead of competition and individuality. Furthermore, the Report’s critique of standardized tests and rankings is probably its most firm counterpoint with the logic promoted by PISA.

The debates about these two very different ways of approaching the relation between education and the future of humanity will probably take the center stage in global policy spaces in the next few years. These debates highlight how education is more than just teaching and learning. Education has the ability to shape the future and our approaches to challenges such as climate change, growing inequalities and the pandemic. Whether young people face such challenges collectively or individually; in competition or through collaboration, will surely be influenced by the kind of education principles promoted in global policy spaces.

References

- Addy, C., 2017. Golden relics & historical standards: how the OECD is expanding global education governance through PISA for development. *Crit. Stud. Educ.* 58, 311–325.
- Altbach, P., Kelly, G., 1978. *Education and Colonialism*. Longman, New York and London.
- Anderson-Levitt, K., 2003. *Local Meanings, Global Schooling: Anthropology and World Culture Theory*. Palgrave Macmillan, New York.
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education “best practice”. *Comp. Educ.* 52, 202–229.
- Auld, E., Rapple, J., Morris, P., 2019. PISA for development: how the OECD and World Bank shaped education governance post-2015. *Comp. Educ.* 55, 197–219.
- Avelar, M., Ball, S.J., 2019. Mapping new philanthropy and the heterarchical state: the mobilization for the National Learning Standards in Brazil. *Int. J. Educ. Dev.* 64, 65–73.
- Ball, S.J., 1998. Big Policies/Small World: an introduction to international perspectives in education policy. *Comp. Educ.* 34, 119–130.
- Ball, S.J., 2009. Privatising education, privatising education policy, privatising educational research: network governance and the “competition state”. *J. Educ. Pol.* 24, 83–99.
- Ball, S.J., Junemann, C., 2012. *Networks, New Governance and Education*. The Policy Press, Bristol.
- Barber, M., Mourshed, M., 2007. *How the World’s Best-Performing School Systems Come Out on Top*. McKinsey & Company.
- Baroutsis, A., Lingard, B., 2017. Counting and comparing school performance: an analysis of media coverage of PISA in Australia, 2000–2014. *J. Educ. Pol.* 32, 432–449.
- Beck, U., 2006. *Cosmopolitan Vision*. Polity.
- Beech, J., 2006. The theme of educational transfer in comparative education: a view over time. *Res. Comp. Int. Educ.* 1, 2–13.
- Beech, J., 2009a. Policy spaces, mobile discourses, and the definition of educated identities. *Comp. Educ.* 45, 347–364.
- Beech, J., 2009b. Who is strolling through the global garden? International agencies and educational transfer. In: Cowen, R., Kazamias, A.M. (Eds.), *International Handbook of Comparative Education*. Springer Netherlands, Dordrecht.
- Beech, J., 2011. *Global Panaceas, Local Realities*. Peter Lang Verlag, Berlin, Germany.
- Beech, J., Artopoulos, A., 2021. Actor-network-theory and comparative and international education. In: *The Bloomsbury Handbook of Theory in Comparative and International Education*, p. 429.
- Benavot, A., Cha, Y.-K., Kamens, D., Meyer, J.W., Wong, S.-Y., 1991. Knowledge for the masses: world models and national curricula, 1920–1986. *Am. Socio. Rev.* 56, 85–100.
- Bonal, X., 2002. Plus ça change ... the world bank global education policy and the post-Washington consensus. *Int. Stud. Sociol. Educ.* 12, 3–22.
- Carnoy, M., 1974. *Education as Cultural Imperialism*. Longman, New York and London.
- Caruso, M., Roldán Vera, E., 2005. Pluralizing meanings: the monitorial system of education in Latin America in the early nineteenth century. *Paedagog. Hist.* 41, 645–654.
- Chubb, J., Moe, T., 1992. *A Lesson in School Reform From Great Britain*. The Brookings Institution, Washington DC.
- Coffield, F., 2012. Why the McKinsey reports will not improve school systems. *J. Educ. Pol.* 27, 131–149.
- Cowen, R., Figueiredo-Cowen, M., 1989. Educational excellence. The case of Brazil. *High. Educ. Pol.* 2, 14–17.
- Dussel, E., 2008. Philosophy of liberation, the postmodern debate, and Latin American studies. In: Morafía, M., Dussel, E., Jáuregui, C.A. (Eds.), *Coloniality at Large: Latin America and the Postcolonial Debate*. Duke University Press, Durham NC.
- Edwards, D.B., Klees, S.J., Wildish, J.L., 2017. Dynamics of low-fee private schools in Kenya: governmental legitimation, school-community dependence, and resource uncertainty. *Teach. Coll. Rec.* 119, 1–42.
- Escobar, A., 1995. *Encountering Development: The Making and Unmaking of the Third World*. Princeton University Press, New Jersey.
- Giddens, A., 2000. *Runaway World: How Globalization is Reshaping Our Lives*. Routledge, London.
- Gorur, R., 2011. ANT on the PISA Trail: following the statistical pursuit of certainty. *Educ. Philos. Theor.* 43, 76–93.

- Gorur, R., 2014. Towards a sociology of measurement in education policy. *Eur. Educ. Res. J.* 13, 58–72.
- Gorur, R., Wu, M., 2015. Leaning too far? PISA, policy and Australia's "top five" ambitions. *Discourse* 36, 647–664.
- Hans, N., 1963. *The Russian Tradition in Education*. Routledge & Paul Kogan, London.
- Hanushek, E.A., Woessmann, L., 2007. *The Role of Education Quality for Economic Growth*. World Bank, Washington DC.
- Holmes, B., 1981. *Comparative Education: Some Consideration of Method*. Allen & Unwin, London.
- Hopfenbeck, T.N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., Baird, J.-A., 2018. Lessons learned from PISA: a systematic review of peer-reviewed articles on the programme for international student assessment. *Scand. J. Educ. Res.* 62, 333–353.
- Jones, P., 2006. *World Bank Financing of Education: Lending, Learning and Development*. Routledge, London.
- Kim, J.-H., 2017. The traveling of PISA: fabricating the Korean global citizen and the reason of reforms. In: Popkewitz, T.A., Diaz, J., Kirchgaser, C. (Eds.), *A Political Sociology of Educational Knowledge: Studies of Exclusions and Difference*. Taylor & Francis Group, London, UK.
- Komatsu, H., Rappleye, J., 2017. A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth. *Comp. Educ.* 53, 166–191.
- Komatsu, H., Rappleye, J., 2018. Will SDG4 Achieve Environmental Sustainability. Center for Advance Studies in Global Education, Arizona State University (ASU).
- Latour, B., 2004. Wose cosmos, which cosmopolitics? Comments on the peace terms of Ulrich Beck. *Common. Knowl.* 10, 450–462.
- Lewis, S., 2017. Governing schooling through "what works": the OECD's PISA for schools. *J. Educ. Pol.* 32, 281–302.
- Lingard, B., Sellar, S., 2013. Globalization, edu-business and network governance: the policy sociology of Stephen J. Ball and rethinking education policy analysis. *Lond. Rev. Educ.* 11, 265–280.
- Lyotard, J.F., 1984. *The Postmodern Condition: A Report on Knowledge*. Manchester University Press, Manchester.
- Mayor, F., Tanguiane, S., 1997. *UNESCO, an Ideal in Action: The Continuing Relevance of a Visionary Text*. UNESCO Pub, Paris.
- Mcneely, C.L., 1995. Prescribing national education policies: the role of international organizations. *Comp. Educ. Rev.* 39, 483–507.
- Meyer, J., Ramirez, F., 2000. The world institutionalization of education. In: Schrewer, J. (Ed.), *Discourse Formation in Comparative Education*. Peter Lang, Frankfurt am Main.
- Moraña, M., Dussel, E., Jauregui, C.A., 2008. Colonialism and its replicants. In: Moraña, M., Dussel, E., Jauregui, C.A. (Eds.), *Coloniality at Large: Latin America and the Postcolonial Debate*. Duke University Press, Durham NC.
- Mourshed, M., Chijioke, C., Barber, M., 2010. *How the World's Most Improved School Systems Keep Getting Better*. McKinsey & Company.
- Nambissan, G.B., Ball, S.J., 2011. Low cost private schools as part of the solution for education for all. In: Lall, M., Nambissan, G.B. (Eds.), *Education and Social Justice in the Era of Globalisation: Perspectives From India and the UK*. Routledge, India.
- Noah, H., Eckstein, M., 1969. *Towards a Science of Comparative Education*. Macmillan, London.
- Novelli, M., Higgins, S., Ugur, M., Valiente, O., 2014. The Political Economy of Education Systems in Conflict-Affected Contexts: A Rigorous Literature Review.
- Novoa, A., Yaviv-Mashal, F., 2011. Comparative research in education: a model of governance or a historical journey. In: Fenwick, T., Mangez, E., Ozga, J. (Eds.), *World Yearbook of Education 2014: Governing Knowledge Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education*. Routledge, London.
- Oxfam, 2019. *False Promises How Delivering Education Through Public-Private Partnerships Risks Fueling Inequality Instead of Achieving Quality Education for All* OXFAM Briefing Paper. OXFAM, Oxford.
- Papadopoulos, G., 1994. *Education 1960–1990: The OECD Perspective*. OECD, Paris.
- Pearson, 2016. *The Pearson Affordable Learning Fund: Myths + Facts* [Online]. Available: https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/get-the-facts/PALF_Myths_Facts_March_2016.pdf.
- Plattoeva, N., 2018. How can transnational connection hold? An actor network theory approach to the materiality of transnational education governance. In: *Education Governance and Social Theory: Interdisciplinary Approaches to Research*. Bloomsbury, London, pp. 103–122.
- Pons, X., 2012. Going beyond the "PISA shock" discourse: an analysis of the cognitive reception of PISA in six European countries, 2001–2008. *Eur. Educ. Res. J.* 11, 206–226.
- Ravitch, D., 1983. *The Troubled Crusade: American Education*. Basic Books, New York, pp. 1945–1980.
- Rickover, H., 1962. *Swiss Schools and Ours: Why Theirs are Better*. Brown, New York.
- Robertson, S.L., Dale, R., 2015. Towards a "critical cultural political economy" account of the globalising of education. *Glob. Soc. Educ.* 13, 149–170.
- Sadler, M., 1979. How far can we learn anything of practical value from the study of foreign systems of education? In: Higginson, J. (Ed.), *Selections From Michael Sadler: Studies in World Citizenship*. DeJall & Meyor, Liverpool.
- Santori, D., Ball, S.J., Junemann, C., 2016. Financial markets and investment in education. In: Verger, A., Lubienski, C., Steiner-Khamsi, G. (Eds.), *World Yearbook of Education 2016: The Global Education Industry*. Routledge, London.
- Schleicher, A., 2008. Piac: a new strategy for assessing adult competencies. *Int. Rev. Educ.* 54, 627–650.
- Schriewer, J., 2016. *World Culture Re-contextualised: Meaning, Constellations and Path Dependencies in Comparative and International Education Research*. Routledge, London and New York.
- Sellar, S., Lingard, B., 2013. The OECD and global governance in education. *J. Educ. Pol.* 28, 710–725.
- Sguissardi, V., Amaral, N.C., 2000. O Banco Mundial revisa posições: quem há de pagar a conta?.
- Shahjahan, R.A., 2012. The roles of international organizations (IOs) in globalizing higher education policy. In: Smart, J.C., Paulsen, M.B. (Eds.), *Higher Education: Handbook of Theory and Research*, vol. 27. Springer Netherlands, Dordrecht.
- Shahjahan, R.A., 2016. International organizations (IOs), epistemic tools of influence, and the colonial geopolitics of knowledge production in higher education policy. *J. Educ. Pol.* 31, 694–710.
- Srivastava, P., 2016. *Questioning the Global Scaling Up of Low-Fee Private Schooling: The Nexus Between Business, Philanthropy and PPPs*. Education Publications.
- Steiner-Khamsi, G., 2014. Cross-national policy borrowing: understanding reception and translation. *Asia Pac. J. Educ.* 34, 153–167.
- Steiner-Khamsi, G., Quist, H.O., 2000. The politics of educational borrowing: reopening the case of Achimota in British Ghana. *Comp. Educ. Rev.* 44, 272–299.
- Stevenson, H., Stigler, J., 1993. *The Learning Gap: Why Our Schools are Failing and What Can We Learn From Japanese and Chinese Education*. Summit Books, New York and London.
- Takayama, K., 2008. The politics of international league tables: PISA in Japan's achievement crisis debate. *Comp. Educ.* 44, 387–407.
- Takayama, K., 2012. Bringing a political "bite" to educational transfer studies: cultural politics of PISA and the OECD in Japanese education reform. In: *World Yearbook of Education 2012*. Routledge.
- Tooley, J., 2013. Challenging educational injustice: "Grassroots" privatisation in South Asia and sub-Saharan Africa. *Oxf. Rev. Educ.* 39, 446–463.
- Trace, A., 1961. *What Ivan Knows that Johnny Doesn't: A Comparison of Soviet and American School Programs*. Harper, New York.
- Truman, H., 1949. Inaugural Address of President Harry S. Truman [Online]. Available: https://avalon.law.yale.edu/20th_century/truman.asp. (Accessed March 2022).
- United States. Office of Educational Research and Improvement, 1987. In: Team, J.S. (Ed.), *Japanese Education Today*. United States Department of Education, Washington DC.
- Urban, M., 2017. We need meaningful, systematic evaluation, not a preschool PISA. *Glob. Educ. Rev.* 4, 18–24.
- Urry, J., 2012. *Sociology Beyond Societies: Mobilities for the Twenty-First Century*. Routledge.
- Vincent-Lancrin, S., González-Sancho, C., Bouckaert, M., Luca, F.D., Fernández-Barrerra, M., Jacotin, G., Urgel, J., Vidal, Q., 2019. *Fostering Students' Creativity and Critical Thinking: What It Means in School*. OECD Publishing, Paris.
- Waldow, F., 2009. What PISA did and did not do: Germany after the "PISA-shock". *Eur. Educ. Res. J.* 8, 476–483.
- Waldow, F., Takayama, K., Sung, Y.-K., 2014. Rethinking the pattern of external policy referencing: media discourses over the "Asian Tigers" PISA success in Australia, Germany and South Korea. *Comp. Educ.* 50, 302–321.
- World Bank, 1995. *Priorities and Strategies for Education: A World Bank Sector Review*. World Bank, Washington DC.

- World Bank, 2019. Teach: Helping Countries Track and Improve Teaching Quality. World Bank. Available. <https://www.worldbank.org/en/topic/education/brief/teach-helping-countries-track-and-improve-teaching-quality>.
- Ydesen, C., 2021. Globalization and localization in the shaping of the Danish public education system: recontextualization processes in four historical educational reforms. In: Zhao, W., Tröhler, D. (Eds.), *Euro-Asian Encounters on 21st-Century Competency-Based Curriculum Reforms: Cultural Views on Globalization and Localization*. Springer Singapore, Singapore.

Afrocentric education: traditions, politics and possibilities

George J.S. Dei and Arlo Kempf, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction and aims	161
Authors' self locations	162
Origins and historical overview of Afrocentric education	162
Struggling for Afrocentric schooling	164
Major tenets & philosophies of Afrocentric education	165
Contemporary implications and offerings of Afrocentrism for educational change	167
Conclusion	169
References	169

Introduction and aims

This article offers an overview and introduction to Afrocentric education, describing the central tenets of Afrocentric education as well as its political and philosophical origins. We also investigate the early purposes and contemporary implications of Afrocentric education (a term which, for the purposes of this chapter, we define to include Africentric, African-centered, and African-focused education). At its core, Afrocentric education offers educational possibilities which lie beyond, and which belie mainstream multicultural approaches to teaching and learning. Recognizing that schooling in colonial, settler colonial, and so-called post-colonial contexts frequently constitutes and is constituted by White supremacist and colonial power relations, Afrocentric education offers a critical rejection of colonial knowledge, ways of knowing, pedagogies, and educational systems. As an alternative which supports the spiritual, psychological, and academic well-being and learning of African and African Diasporic learners, Afrocentric education centers African and African Diasporic knowledges, cultures, histories, epistemologies, pedagogies, and politics.

Although Afrocentrism and Afrocentric education may be sometimes grouped within broader categories such as inclusive education, multicultural education, critical education for social justice, etc, we do well to recognize that Afrocentric education holds at its core a fundamental rejection of assimilationist ideals in education. While these broader categories are no doubt areas of study and practice which aim to better serve all students from pre-K to university contexts, at the core of most mainstream moves to equity in education lies a hope that a core curriculum which serves and includes all students is possible—or at the very least desirable. Afrocentric education focuses instead on the contemporary education and well-being of African and African-descended peoples specifically, in recognition that African and African Diasporic knowledges, pedagogies, and traditions can powerfully support the learning and well-being of African and African-descended children and youth in formal education settings, and that mainstream education (including decades of equity and inclusion initiatives) has persistently failed African-descended students.

As a starting point, Asante (1991) argues that Afrocentricity questions “the imposition of the White supremacist view as universal and/or classical (Asante, 1990); demonstrates the indefensibility of racist theories that assault multiculturalism and pluralism; and projects a humanistic and pluralistic viewpoint by articulating Afrocentricity as a valid, nonhegemonic perspective” (p. 173). With roots dating back well over a century, Afrocentricity, broadly understood, has a history which spans continents and generations. In elementary and secondary schooling, Afrocentric education is taken up in a variety of ways in a variety of places. Most notably, schools in the US saw a rise in independent Afrocentric schools beginning in the 1970s. Additionally, a number of schools in the US and Canada offer Afrocentric programming, and often associate culturally affirming school practices with Afrocentric schooling. At the upper secondary and post-secondary levels, Afrocentricity is an important area of study, and is often braided with other critical analytical frameworks.

This chapter focuses generally on Afrocentric education as a philosophical approach in education. After offering remarks below on our subject location as authors, the first section of this chapter provides an overview of the historical origins of Afrocentric education, noting a varied a complex history with roots in a number of traditions. The next section takes up the key purposes, tenets, and philosophies of Afrocentric education, with a look at key contributions from leading scholars. We offer the basic tenets of Afrocentric schooling with hopes that these are instructional for moving the debate along in terms of concrete ways for counter visioning primary and secondary education. Finally, we discuss the contemporary implications and offerings of Afrocentrism for educational change in the early-mid 21st century. We take up what it means for Afrocentric schooling and education to advocate for social change, and how we can collectively ensure our young learners practicalize and actualize what they learn in school in terms of Afrocentric teachings. We also argue for the potential of pedagogies employed in Afrocentric schools to be applied in mainstream education contexts to serve all students. The contributions of Afrocentric education are indeed an important focal and entry point for critical educators in an age of widespread and highly public struggles for racial justice.

Authors' self locations

We start this paper with our respective voices before we join in a conversation.

George Dei: I acknowledge that I am writing from the Treaty Lands and Territory of the Mississaugas of the New Credit First Nation and the Dish with One Spoon Wampum, what is now known as Canada. The Land has always been a source of Earthly teachings and blessings. I also relate to this Land from my African Indigeneity as a member of the traditional Asakyiri clan of the Ghana, my birthplace. The Asakyiri clan is one of the eleven major clans in the Asante kingdom of Ghana, and our spirit animal is the vulture. Asakyiri mythology states my ancestors chose the vulture as our spirit animal because of our collective belief in making use of that which has been discarded. This, along with the vulture's teachings of being cognizant of the downtrodden and forgotten and how not to be, has advanced my appreciation of life and the privileges I have been granted. I see the vulture as an important source of learning about the sanctity of life and activity.

Teaching and learning demands the sanctity of activity as our coming to know only makes sense with connections of the social, physical, and metaphysical realms of human existence. We are in relation with the entire universe as "fundamentally spirit and spiritual with material manifestations" (Carroll, 2014, p. 259), stressing the importance of community cultural values and prioritizing group relations. This is a philosophical worldview that crystalizes in the nexus of body, mind, soul, and spirit in the coming to know, understand, and act within worlds. Education is more than an activity of the body and mind. To engage in holistic learning, we must engage the soul and spirit of learners. My scholarly work through the African-centered lens has led me to appreciate social justice and equity work as flowing from spirituality. Thus, I locate my classroom teachings and scholarship within the context of an African identity constructed in history, culture, and politics laced within spiritual ontologies. I validate Afrocentrism as an appreciation of the teachings of African culture, history and Indigeneity as important starting points. This validation has served as foundational knowledge with which to interrogate schooling and education, and to think through new educational futures. I see the Afrocentric paradigm pioneered by global influential scholar Molefi Asante and others as offering a significant counter reading of what we understand to be schooling for particular [but not exclusively] youth of African descent. Of course, no knowledge is immune to interrogation and criticisms. I also believe that we cannot dismiss a counter-reading of our world and the possibilities of what Afrocentric education can offer in the education of young learners of today. In my journey in Afrocentric scholarship, I have worked with Arlo Kempf in mutually reinforcing and beneficial ways to contribute to the literature on Afrocentrism from the Canadian context.

Arlo Kempf: Acknowledging that where we come from informs where we are, George and I are writing from different places. I am a White settler Canadian. On my father's side, my ancestors came to North America from what is now Germany. On my mother's side, our people come from Oklahoma—after searching through a series of adoptions, stories, lies, and aspirations, as best I can tell my ancestors were likely Scottish and Cherokee—I'll never be sure, and as I have never been claimed by any group, I make no claims to any group. I was born on Kanien'kehá:ka land in Tiohtià:ke (commonly known as Montreal). My parents were immigrants from the United States. We later moved to Treaty 13 territory, and I grew up there, on the land of the Mississaugas of the New Credit in what is commonly known as Toronto, where I work and live today. Settler Colonial momentum is a river into which Canadian Settlers are born, or into which they migrate. As I have written before (Purton et al., 2020), stepping out of the water requires first noticing its ceaseless movement—the way it flows around bends and breaks, over rocks and through low hanging brush. If one is lucky enough to notice, it then takes some doing to get out of its rushing waters. I'm not there yet, but I work to find, at the very least, an eddy on the margins of the river from which to catch my breath and step out of the flow. I write from this context of complication, privilege, and humility—a positionality that informs my work as a scholar and my actions as a citizen, family member, descendant, and sapiens. I have been speaking and writing about these issues for about a decade; privileged to do so primarily with my co-author who is, in his own right, a founding scholar of Afrocentric thinking in education. I thus come to discussions of Afrocentricity and Afrocentric education typically when invited—guided by an intellectual humility, informed by my positionality, and my tremendous respect for the African and African-descended scholars and activists who have guided and continue to advance this important work.

Origins and historical overview of Afrocentric education

We locate Afrocentric education outside of any particular national, state, provincial, or district context, as a transhistorical framework that resists and forgoes Eurocentric teaching and learning in favor of education that focuses on the implications and applications of African scientific and cultural knowledges, epistemologies, and axiologies in order to support the academic, spiritual, and psychological success of African-descended students. The vast majority of contemporary academic work on Afrocentricity focuses on charter schools in the United States. Although we offer brief commentary on US Afrocentric schools below, US Afrocentric charter schooling should not be taken as an exemplar or definitive example in determining our understandings of Afrocentric education. Afrocentric education is an applied terrain of Afrocentric theory and activist struggle. We thus locate the historical origins of Afrocentric education more broadly, within critical African-centered knowledge, theoretical approaches, and activism for racial justice.

In the US context, this work goes back well over a century to works of African-American scholar, poet, and activist W.E.B. Du Bois (1868–1963). He was among the earliest African-American anti-racist scholars and public intellectuals, and was the first African-American to earn and receive a doctorate from Harvard University. Du Bois' early theorizations of the race problem in the US remain

a guide for contemporary anti-racist scholarship (see Du Bois, 1996), and his contributions remain a central inspiration for contemporary Afrocentric theory. As co-founder of the National Association for the Advancement of Colored People (NAACP), Du Bois was also the editor-in-chief of the NAACP's revolutionary journal, *The Crisis: A Record of the Darker Races*, founded in 1910 and still in circulation as of 2020.

Du Bois theorized the need for African peoples to reconceptualize race at the structural, institutional, and individual levels in order to mitigate the impacts of quotidian racism and to challenge systemic and structural racial oppression. Du Bois was also among the first scholars to consider and theorize the epistemics of race and positionality with regard to the lived reality of race in the US generally and the experiences of African-Americans specifically. Coining the term "double consciousness," Du Bois articulated the processes by and through which African-Americans developed a sense of self through both self-perception and the process of being perceived and thereby understood by White Americans. Writing in 1903, he argued:

It is a peculiar sensation, this double-consciousness, this sense of always looking at one's self through the eyes of others, of measuring one's soul by the tape of a world that looks on in amused contempt and pity. One ever feels his two-ness—an American, a Negro; two souls, two thoughts, two unreconciled strivings; two warring ideals in one dark body, whose dogged strength alone keeps it from being torn asunder.

Du Bois (1996, p. xxxiv).

These early articulations of race and racialized consciousness and experience have provided a foundation for subsequent academic and activist anti-racist work, and have underpinned US civil rights movements for the past 75 years. This includes later formulations of Afrocentricity which draw heavily on notions of raced consciousness and raced experience.

Writing around the same time, Liberian-American Edward Blyden rejected Eurocentric readings of African culture as lacking or incomplete. He argued that African societies were epistemically and culturally complete and whole, and suggested this understanding was absent in mainstream teachings about Africa to Africans in the Diaspora (Blyden, 1994). Coalescing in New York City (as early progenitors of the Harlem Renaissance), African-American thinkers and leaders like Arthur Schomburg, John Edward Bruce, and others promoted the inherent value of African knowledge and culture. In 1911, Bruce and Schomburg founded the Negro Society for Historical Research in Yonkers, New York. Four years later in Chicago, Illinois, Carter G. Woodson, William B. Hartgrove, George Cleveland Hall, Alexander L. Jackson, and James E. Stamps established the Association for the Study of Negro Life and History on October 2, 1915.

Following these early works, in 1933, Carter G. Woodson offered a foundational contribution to the education question with his coining of the term "miseducation" to identify the process of methodically denying African Americans access to education in which they might find themselves well-represented in their full humanity (Woodson, 1933). Woodson located this miseducation at the heart of the cultural, economic, and social oppression facing African-Americans, and noted the disproportionate value placed on Eurocentric teachings in mainstream schooling as harmful for Black children.

At roughly the same time, the work and influence of Marcus Mosiah Garvey (1887–1940) was on the rise. Garvey was a Diasporic anti-racist organizer, leader, and theorist whose work was key for later conceptions of Afrocentricity. With powerful critiques of the structural, institutional, and individual racism facing Africans in Africa and the Diaspora, Garvey offered a revolutionary articulation of the realities of racism with an emphasis on the agentive capacities of Africans on the continent and throughout the Diaspora. Garvey was the founder Universal Negro Improvement Association and African Communities League (UNIL-ACL) which originated the popular Back-to-Africa movement. By 1918, Garvey had founded the ground-breaking newspaper *Negro World* published by and for African-American and African peoples. The Back-to-Africa movement promoted the notion that Africans in the Diaspora should return to the continent to avoid and reject the racial prejudice, discrimination, and exclusion, which were rampant in Diasporic contexts.

Membership of the Universal Negro Improvement Association numbered in the millions by 1920 as Hill notes (1983, p. 36) with precise numbers unknown; while the circulation of *Negro World*, grew to between 50,000 and 200,000 (Hill, 1983), also with precise numbers unknown. Although the newspaper did not constitute a formal curriculum in a schooling sense, this widely read periodical offered a powerful rejection of mainstream cultural and historical analysis. At the time, as is the case in some contexts as of this writing, mainstream curricula (both formal and informal) available to African-Americans was frequently White supremacist, colonial, and exclusive of the experiences of African and African-descended peoples. *Negro World*, like the NAACP's *The Crisis* offered an African-centered reading of the world.

This intellectual and activist trajectory continued into the latter half of the 20th century. The US Black Power and Black Nationalist movements of the 1960s and 1970s can be understood as a continuation of the work of early 20th century thinkers and activists, all of whom drew on Indigenous African epistemologies, axiologies, and knowledges. In some cases, the earlier work of scholars such as Du Bois and Garvey ended up influencing African intellectuals, including Chinua Achebe, Cheikh Anta Diop, and others, who would in turn inspire the critical Afrocentric work in the US, including that of intellectuals like Molefi Kete Asante, Maulana Karenga, Kofi Lomotey, Mwalimu J. Shujaa, and others who would establish the contemporary foundations of Afrocentricity discussed in the next section.

As we note elsewhere (Dei and Kempf, 2013), Diasporic Africans have continuously resisted oppression in both action and theory when faced with racial educational injustice. Indeed, the contemporary discussion and practice of Afrocentricity and Afrocentric education stand on the foundations erected by those who have come before. African-centered education builds on the work and struggles of radicals, activists, poets, scholars, and regular people in Africa and the Diaspora who have fought against

oppression in their own ways. These figures include Frederick Douglass, Harriet Tubman, Booker T. Washington, Langston Hughes, Frantz Fanon, Aimé Césaire, Amílcar Cabral, Rosa Parks, Viola Desmond, Martin Luther King Jr., James Baldwin, Malcolm X, Nelson Mandela, and countless others.

Struggling for Afrocentric schooling

Scholarship on Independent Black Institutions, Charter Schools, and Black Supplementary schools (as they are called in the United Kingdom) offers important analysis and documentation of the early educational activism of Africans in the Diaspora. The works of [Anderson \(1988\)](#), [Butchart \(1980\)](#), [Ratteray \(1992\)](#), and [Watkins \(2001\)](#) detail resistance by African-descended peoples in and to colonial schooling contexts, while [Ratteray and Shujaa \(1987\)](#) trace this tradition of resistance to the 1790s and the global system of trade in enslaved people. Contemporary activism for racial justice in education, and some conceptions of Afrofuturism ([Ladson-Billings, 2018](#)), can be understood in this continuum of resistance, as well.

In discussing the US context, [Lee \(1992\)](#) and [Reay and Mirza \(1997\)](#), [Mirza and Reay \(2000\)](#) investigate African-American activists' efforts for self-education during the period of slavery (1619–roughly 1883) as well as during the post-Reconstruction and post-slavery period (1865–1877). [Mirza and Reay \(2000\)](#) chart the establishment and presence of Black supplementary schools in the US, the UK, and elsewhere through the 20th century. [Lawson Bush \(2004\)](#) examines Independent Black Institutions (IBIs) in the US as far back as the end of slavery in 1863. These writings paint a powerful portrait of African agency and self-determination in education, despite facing significant racial injustice at the structural, systemic, institutional, and individual levels. Referencing [Anderson \(1988\)](#), Bush writes:

White northern missionaries were ... surprised by the will of the formerly enslaved Africans to educate themselves. Books or fragments of books were seen in the hands of African American men, women and children everywhere they travelled in the South.

[Bush \(2004, p. 387\)](#).

As we have noted ([Dei and Kempf, 2013](#)), Bush's work locates historical discourses, such as the Back-to-African movement, Pan-African movement, and Black Consciousness movements (and their leaders), as having tremendously impacted the establishment of many early African-centered institutions. In addition, [Rashid and Muhammad \(1992\)](#), in tracing the history of the Sister Clara Muhammad school (among the first Islamic schools emphasizing African histories and cultures), discusses the impact of Marcus Garvey's Universal Negro Improvement Association. Its motto of Black self-determination and racial pride became the foundational principle upon which the institution's pedagogy of liberation emerged as encapsulated in the message of knowing the self, loving the self, and doing for self ([Rashid and Muhammad, 1992](#)).

The development of schooling for racial justice generally and for Afrocentric education in particular has been an uphill battle against settler colonial and tacit White supremacist state bureaucracy. Due to changes to funding and political will, African-centered programs and institutions appear to have increased and decreased in number periodically, including a notable reduction in the US from 1890 to 1935, before rising again beginning in the 1980s. As mentioned, the US currently has more Afrocentric schools and programs than any other diasporic jurisdiction. Although exact numbers are unavailable, the number of schools seems to have been at their highest in 1999 at about 400 nationally ([Cohen, 2016](#); [Murrell, 1999](#)). By contrast, Canada has only a handful of schools and programs.

In addition to a lack of state and national data on Afrocentric schooling, numbers are also difficult to establish because not all schools that offer Afrocentric programming identify as Afrocentric schools. Not all Afrocentric schools offer the same degree of Afrocentric content: with some offering immersive approaches, and some offering culture affirming approaches, and most with local definitions and criteria for operationalizing such terms. There are different types of US Afrocentric schools, including private, public, charter, magnet, and other designations from pre-Kindergarten to upper secondary. Afrocentric schools frequently face a myriad of political issues, including chronic underfunding, under-resourcing, as well as racist public [mis]perception ([Dei and Kempf, 2013](#)). Given this variety, we cannot paint all Afrocentric schools with the same brush. However, we can see them all as manifestations of a larger Black struggle for Afrocentric agency in education and as part of the Black struggle for the spiritual, psychological, and academic well-being of Black children.

As we argue elsewhere ([Dei and Kempf, 2013](#)), this human thirst for learning explains the epistemological focus of the colonial project from literacy laws in the US, to forced segregation in schooling in Canada, to the forcible exclusion of African peoples from formal education in the Caribbean during the early European colonial era. In this sense, Afrocentric education as an idea, should not be understood outside of the context of this long struggle. To wit, we must avoid understandings of Afrocentricity and Afrocentric education that take an evental approach (looking only at individual moments, contexts, etc.), and instead invoke a longer view which recognizes the historical trajectory of scholarly activism and activist scholarship, by individuals, communities, and institutions fighting for racial justice centered on Afrocentric teachings and learnings.

Although Afrocentric education is reliant on particular philosophical tenets, it has grown out of the soil of resistance by African peoples to colonial and White supremacist education. So, in addition to the public and well-documented work of Du Bois, Garvey, and others we mention above, Afrocentric education also can be linked to tireless efforts of enslaved Africans who risked personal harm to themselves and family in order to hide books, who struggled to read hidden in the dark, and who made countless sacrifices

in order to achieve even basic formal education for themselves and their children. We now move to a discussion of the major tenets and philosophies of Afrocentric education.

Major tenets & philosophies of Afrocentric education

Although there is no singular philosophical proposition which defines Afrocentric education, or discrete list of beliefs offering a snapshot of Afrocentricity in its totality, there are central figures and themes without which an understanding of Afrocentric education would be impossible. Broadly, Afrocentric education can be understood as providing:

an avenue for ensuring the holistic education of African children in a way that nurtures their successful social, spiritual, and academic development [by aiming to] unite academic excellence and positive character, in order to foster an environment where interdependence, community and reciprocity are part and parcel the learning process.

Kempf (2015, p. 176).

Building on generations of activist and scholarly work dating back to the late 19th century, contemporary Afrocentric education scholarship in the US began to flourish alongside the establishment and development of Afrocentric programs, centers, and schools in the 1970s and 1980s.

The major tenets and philosophies of Afrocentric education emerged from the foundational works of Molefi Kete Asante, Maulana Karenga, Kofi Lomotey, Akwasi Akoto, Marianna Ani, Carol D. Lee, Peter C. Murrell, and Mwalimu J. Shujaa, to name a few. These authors work in the tradition of critical resistance to and rejection of mainstream Eurocentric schooling and its harmful impacts on Black children in terms of spiritual, academic, and psychological health. Their works are the building blocks of a counter vision of education, which centers, celebrates, and interrogates continental African as well as Diasporic African culture, epistemology, ways of being, and experiences.

The works of Molefi Asante, Kofi Lomotey, Maulana Karenga and Mwalimu Shujaa offer the central theorizations of contemporary conceptions of Afrocentric education. Asante, in particular, provides the core formulation of contemporary Afrocentric education in three seminal pieces: *Afrocentricity: The Theory of Social Change* (1980); *The Afrocentric Idea* (1987); and with specific significance for education, *The Afrocentric Idea in Education* (1991). The push for Afrocentric schooling has always been an act of resistance, and Asante's scholarship was no exception. Kendi (2016) locates the rise of Afrocentric theory in the 1980s alongside and in relation to Black studies programs, hip-hop pedagogy/education programs, and other African-centered academic offerings in US universities, in the context of a rising conservative movement focused in part on the promotion of tropes of Black criminality (p. 442).

Building upon the work of Woodson (1933), in *The Afrocentric Idea in Education* (1991), Asante identifies "two propositions that make up the background of the theoretical and philosophical issues" for which he is arguing (p. 170). First, he argues, "education is fundamentally a social phenomenon whose ultimate purpose is to socialize the learner; to send a child to school is to prepare that child to become part of a social group" (Asante, 1991). Second, he states, "schools are reflective of the societies that develop them (i.e., a White supremacist-dominated society will develop a White supremacist educational system)" (Asante, 1991). Recognizing the inherent White supremacy and Eurocentricity characterizing much mainstream teaching and learning, Afrocentric educational approaches "provide students the opportunity to study the world and its people, concepts, and history from an African world view" (Asante, 1991, 171). In the Afrocentric classroom, the Afrocentric teacher will not:

marginalize African American children by causing them to question their own self-worth because their people's story is seldom told. By seeing themselves as the subjects rather than the objects of education-be the discipline biology, medicine, literature, or social studies-African American students come to see themselves not merely as seekers of knowledge but as integral participants in it (Asante, 1991, 171).

Lee's (1992) writing on Afrocentric schools in the US offers an exceptionally clear and comprehensive summary of the aims of Afrocentric education. She offers an eight-point summary of these aims as follows:

1. Legitimize African stores of knowledge
2. Positively exploit and scaffold productive community and cultural practices
3. Extend and build upon indigenous African languages
4. Reinforce community ties and idealize [the concept of] service to one's family, community, nation, race and world
5. Promote positive social relationships
6. Impart a worldview that idealizes a positive, self-sufficient future for one's people without denying the self-worth and right to self-determination of others
7. Support cultural continuity while promoting critical consciousness
8. Promote the vision of individuals and communities as producers rather than as simply consumers (Lee, 1992, pp. 165–166).

Key here is the braiding of academic learning, morality, and ethics as equally significant pedagogical foundations within the Afrocentric approach. This is guided by the Kemetic proposition that moral social practice is necessary for the development of humanity (For more on Kemetic philosophy, Asante's *The Egyptian Philosophers: Ancient African Voices from Imhotep to Akhenaten* [2000], and Maulana Karenga's *Maat: The Moral Ideal in Ancient Egypt—A Study in Classical African Ethics* [2004] are essential and highly informative.)

Afrocentric education “aims to unite academic excellence and positive character” and to foster an environment where interdependence, community and reciprocity are part and parcel the learning process (Lee, 1992, citing Karenga, 1991). Afrocentric education is thus informed by Karenga's (1991) notions of Kwanzaa, which include faith, unity, purpose, creativity, cooperative economics, self-determination, and collective work and responsibility. The operationalization of these key principles is required for the successful realization of an Afrocentric educational program. It follows that the instructor or teacher is as important as the curriculum in the context of Afrocentric education. As Shujaa (1994) states:

When discussing African-centered education ... more emphasis should be placed on pedagogy than on curriculum. My reasoning for this is that pedagogy conveys the importance of the teacher to the education process while curriculum is too often reduced to documentation It is the African centeredness of the teacher's thinking that determines the African-centeredness of the teaching, (p. 265).

Thus, for Afrocentric education to succeed, as we have argued elsewhere (Dei and Kempf, 2013), teachers and educators must live, breathe, and be constantly engaged in the social philosophy and praxis of Maat (Lee, citing Karenga, 1991). Thus, there is a moral expectation of teachers. Reflecting on Karenga, Kempf (2011) provides the following overview of Maatian ethics:

Etymologically Maat is traced to notions of straightness and evenness. Drawing from a substantial literature on Maat, Karenga reveals its meaning as one signifying a guiding force, a cosmic order, truth, ideal wisdom, a metaphysical ideal and an epistemological ideal. As a wide-ranging concept, it expresses itself in four domains: (1) the universal—the totality of ordered existence; (2) the political domain—regulating justice and injustice; (3) the social domain—relationships and duty in the context of community; and, (4) the personal domain in which—following the rules and principles of Maat is to realize concretely the universal order in oneself (p. 102).

Maatian ethics are then a shared responsibility of the Afrocentric educator and the student studying within an Afrocentric program. This shared ethical stance is thus a required mutual starting point for Afrocentric teaching and learning, which recognizes the social playing field is not level and schooling is no exception; that learners' schooling identities are shaped by experiences, history, cultures, and socialization processes; that there is a link between these identities—race, class, gender, sexuality, disability, and other factors and how as learners and educators we come to know and engage schooling; and that every student's identity is relevant, in a productive sense, to his, her, or their learning.

From this philosophical launching pad, we can dive into the core approaches of Afrocentric education. Building on Dei et al.'s earlier (2000) writing on Afrocentric education, we sketch nine spheres of implementation for Afrocentric education: representation, language, family/community and school partnerships with a focus on shared responsibility, co-operative education, equity and values education, agency for transformation, Indigenous/community knowledges, spirituality, and learning for climate justice.

- (1) *Representation* refers to the visual culture/landscape of the school, knowledge representation of different bodies, and physical representation of diverse Black and other racialized educators and administrative staff who are well-grounded in their communities. This includes digital media and online learning activities and platforms, as well as race-conscious approaches to media literacy.
- (2) *Language* refers to the teachings of African and other Indigenous languages apart from the dominant local, regional, or national language. This domain engages questions of linguistic and the central importance of sustaining language, as well as an anti-colonial reading of the politics of language use.
- (3) *Family/Community and school partnerships with a focus on shared responsibility* refers to meaningful engagement with local communities as significant and valuable partners in the delivery of education to youth. This element recognizes the diversity in and between communities, and includes an explicit commitment to engaging the politics of a given school community context.
- (4) *Co-operative education* refers to education that stresses co-operative learning, multiple forms of student excellence, and doing away with hierarchical structures of schooling. This includes attention to the politics of cooperative vs. competitive approaches.
- (5) *Equity and values education* refer to the linkage of issues of power and social difference with schooling processes and knowledge production, and develops the learner's understanding of local agency.
- (6) *Agency for Transformation* refers to the development of students' ability to engage their communities and their broader world, as empowered and critically consciousness transformative agents. Building on spheres three and five, students are emboldened and empowered to understand and undertake responsibilities to his/her/their communities in pursuit of transformation for social justice.
- (7) *African Indigenous/community knowledges* refers to cultivating multi-centric knowledge systems in schools by affirming off-school and local cultural resource knowledge bases, as well as lifelong learning. This includes a recognition of the parallels and

distinctions of European colonial and settler colonial race paradigms which illustrates opportunities for solidarity and collaboration between African Indigenous peoples and other Indigenous groups.

- (8) *Spirituality* refers to evoking learners' understandings of the self, inner strength/environments and connections to the group and outer environments, as well as stressing the body-mind-soul interface, and the society-nature-culture nexus. This may have specific and/or general links to extant religious beliefs or practices of students and school families, but is not bound to or guided by any formal religious approach or tenet. This domain recognizes and nourishes the interlocking relationships which braid between academic, psychological, spiritual, physical, and cultural health.
- (9) *Learning for climate justice* refers to developing learners' critical understanding of global climate crises and the politics of climate crisis aggravation and mitigation. With clear emphases on African and African Diasporic perspectives, solutions, and struggles in global perspective, as well as attention to the relationship between colonialism and environmental destruction, this sphere promotes Indigenous African and other Indigenous and anti-colonial approaches to sustainability and climate justice.

Adapted from Dei et al. (2000).

Successful Afrocentric education across these nine spheres relies on the learner and the teacher being affirmed in all of their identities in order for holistic education to proceed. Among the most important tasks of Afrocentric education is to expose learners to the realities of society in a supportive environment that empowers today's youth to deal with societal pressures (racism, sexism, employment, family support, housing, etc.). This means teaching about race and racism, gender, disability, age, sexuality, language, religion, class, and other markers of social identity. The African continent itself cannot be understood through cultural reductionist paradigms. Similarly, the African Diaspora refers to a plurality of people, cultures, knowledges, histories, relations, and experiences. With this in mind, African-centered learning draws from a dynamic and fluid understanding of *Africa* and *African*, an understanding which evolves from the past, present and future, and which is constituted by African peoples of today as well as those who have come before.

There are both challenges and opportunities inherent to educational work in these domains. Challenges arise around a host of issues, including a scarcity of time and resources for teachers, questions of representation, debates on perspective and on what constitutes authentic African and/or Afrocentric knowledge, classroom community building, differences of opinion and experience, and the tensions that naturally arise in participatory education when working to honor a variety of voices and experiences which may include string rejections of formal education and the teacher. These challenges are not unique to Afrocentric education and, with the exception of time and resources, can be addressed over time through pedagogical practice.

We have attempted to offer a general overview of key philosophical tenets of Afrocentric education. Just as we caution against understanding Afrocentric education based on a particular school or formation, we suggest a look at its broad historical trajectory and situation. We also emphasize that, even guided by the principles above, Afrocentric education cannot be developed or imposed absent a robust engagement with the local context in which it is taking place. Afrocentric education begins with a recognition of contextual specificity—of the political, social, historical, geographical, and cultural contexts of a given educational setting. While Afrocentric education is (a priori) a site of struggle for racial educational justice, the specifics of different sites and the local contours of the struggle will necessarily vary. It follows that so too must each approach to Afrocentric teaching and learning. The tenets and philosophies outlined here are a starting point for engagement and application with specific local contexts. We turn now to a discussion of the contemporary implications and offerings of Afrocentrism for educational change.

Contemporary implications and offerings of Afrocentrism for educational change

Although we support the call for Afrocentric education, we do not propose to simply sweep away the current school system. We critique mainstream educational systems as a way to think through new and radical possibilities. We argue that the current school systems for most Diasporic Africans has been built on unstable foundations and could easily crumble unless we respond to its challenges. Clearly, there are redemptive qualities of conventional schooling and education, and we are not willing to give up hope and possibility of transformation. But we all (whether learners, educators, administrators, policy makers, etc.) must be prepared to listen to the uncomfortable voices asking for change because the school system is not working for all. The approach to new educational futures may well lie outside the current school system in terms of creating new forms of schooling and education (e.g., Afrocentric schools). However, it is an idea that we encourage rather than dismiss. We support the establishment and maintenance of new and existing Afrocentric schools within local vicinities and outside mainstream schools, beyond just additional curricular programming. These schools should be publicly funded because they serve the common good of the community. They should be neither implemented nor approached as separate, independent private schools. In a climate where a number of our youth do not find themselves in mainstream schools, we have a societal responsibility to put into action counter measures to address schooling and educational problems.

We also want to work with an "and/with" approach that looks at possibilities of educational transformation from within. This is not without its challenges and risks of disappointment, as this has unfortunately been the story of Black, Indigenous, and racialized communities. Afrocentric teachings are not just for Black and African youth. These teachings are useful for all students attending mainstream schools, as well. We want to turn the frustration of Black and African parents and caregivers and the long history of inaction on the concerns of Black education from an "either/or" situation to an "and/with" solution. On the one hand, we advocate for and support the creation of Afrocentric stand-alone schools in our communities as a possible way out of the educational "crisis"

for our youth while utilizing the lessons to improve mainstream schooling. On the other hand, we support a parallel educational approach. Absent of a complete revolution that tears down the current school system, we advance an argument for integrating Afrocentric teachings into mainstream schools to serve the needs of all students.

Historically, legitimate concerns have been expressed about the ability of current schooling to educate all learners for social success. We suggest Afrocentric education should seek and remain open to, new educational futures and approaches that can engage Afrocentric teachings, with particular implications for young learners of today. We acknowledge some attempts being made for our educational system to respond to calls for social justice, race equity, inclusion, and diversity. These attempts have manifested as modifications in body and knowledge representation, and in some situations an “overhaul” of school climate, environment, culture, and the socio-organizational life of schools. We are aware of the push for educational expertise that extends beyond critical Black and African Studies, to the paradigm of critical analyses of anti-Black racism and colonial violence against Black, Indigenous, and racialized learners (see [Ladson-Billings, 2018](#); [Roane, 2017](#), [Kendi, 2016](#), and others). We cannot discount the possibility of ideas to bring about educational change. And, we see power in the ideas promoted by Afrocentric education. These are the profound ideas of the Land and Earthly teachings of relationality, reciprocity, community building, mutual interdependence, social responsibility, and the pursuit of radical inclusion through a reclamation of local Indigeneities. The reclamation of cultural knowledges is the bedrock for the pursuit of decolonial and anti-colonial prisms of teaching, learning, and administration of education. Afrocentric education calls for more than responding to diversity, social justice, and equity in schooling. It is about re-envisioning education with a mindset and political practice to ensure that young learners are able to meet the changing needs of a globalized world.

The question of what concrete steps need to be taken to respond to primary and secondary education through Afrocentric schooling is a fundamental one. Schools must center local cultural knowledges in classroom teachings. History must be taught as a totality of all learners’ lived experiences. Education must be about Land and Earthly teachings, bodies, spirits, and minds in interactions. Educators and learners must be positioned to develop a more insightful, thought-provoking, in-depth, diverse thinking of schooling and education that have local, national, and international appeal and to promote global conversations about educational theory, practice, and policy changes.

Afrocentric education calls for asking “outside the box questions” to inform a theorization of the meaning and purpose of schooling. Learners must be equipped to ask questions that not only speak to their intellectual agency to design their own futures (e.g., public advocacy questions), but also implicate themselves in what is required to make educational change possible. Learners and educators must be engaged to develop lessons of our shared and contingent histories, experiences, and identities (which are not necessarily singular) to create a different path of educational futurity for all. Schooling must be pursued as an approach to education about building relationality, sharing, and connections to break the cycle of hierarchies and merit badges that reward individual competition over group efforts. The connections between schools and local communities must be strengthened in ways that bring parents, Elders, and local community workers into the schooling sphere to enhance youth learning. Elders can teach through stories, cultural songs, and mythologies. Education can also honor the contributions of our diverse communities in terms of global histories and world civilizations in ways that give power and agency to the learner to pursue individual goals and aspirations through collective and shared dreams. We can no longer continue to marginalize Afrocentric learning in the global politics of knowing, learning, teaching, and governing. We must rethink and restructure education to engage multiple knowledge systems.

Afrocentric education helps learners reclaim their humanity and anti-colonial consciousness. This means teaching about the colonial hierarchies of schooling, including the myriad oppressions, such as anti-Black racism rooted in an ideology of racial hierarchies and a denial and devaluation of African and Black life. We see White supremacy, anti-Blackness, and anti-Black racism as intertwining forces that have succeeded in denying Black humanity and presenting Black/African peoples as a “problem people/population” (see also [Hayes, 2003](#), p. 85). The rallying cry of “Black Lives Matter” is not new. It has been picked up by many, however, given growing frustration and anger with the ways in which Black bodies are easily dispensable and made disposable. In the Black Lives Matter social movement, questions of racial justice have intersected with other forms of social oppression, including disparities and discrimination in health, employment, housing, education, as well as land and environmental causes. Particularly for Black and African bodies, the collective consciousness around these issues must be strengthened through an Afrocentric prism.

We do not present Afrocentrism as simply an “academic theory.” It must be a weapon of educational change, recognizing young learners as the future with a collective capacity of the self-generated knowledge necessary for decolonization. The teachings of Afrocentric schooling spelled out in this chapter suggest making African knowledges about community, relationality, reciprocity, culture, spirituality, history, and Indigenous sciences core ideas in the pre-K-12 curriculum and beyond. This entails a multi-pronged approach to integrating and infusing Afrocentric ideas in all subject fields (i.e., arts, humanities, social studies, and physical sciences) and teaching subjects holistically (including, for example, analyses of power and the relevance of engaging first-hand accounts and stories of actual lived experiences of peoples). There are omissions and hidden histories that speak about social injustices, violence, and local resistances. These can be uncovered when we teach about colonization and enslavement critically to reveal White/colonizer privilege and violence as well as African and Indigenous resistance ([Dei, 2021](#)). This will require an Afrocentric schooling focus on precolonial Africa and cartographies of Indigeneities, African knowledge and philosophies.

In pursuing Afrocentric knowledge in schooling, educators must develop explicit lesson plans and curriculum units using diverse knowledge sources, local fables, stories, proverbs, and tales, as well as literature units on Black authors and engage the work of artists, educators, politicians, and community activists—young and old ([Dei and McDermott, 2019](#)). We need Afrocentric narration of histories from anti-colonial prisms and to center Black and African male, female, cisgender, queer, etc. identities and voices in telling their own stories ([Dei, 2021](#)). Education can de-root learners from culture, history, and politics. To challenge the

depersonalization of conventional schooling, students should be able to center themselves in their own learning. Afrocentric schooling has the capacity to help African learners reinvent their Africanness in a global Diasporic context. Furthermore, we need education to embolden learners to acknowledge and validate their respective histories, cultures, and identities as intertwined and interconnected. It is not enough to “learn to be Black or African.” It is more important for the learner to think through and to act on one’s Blackness, Africanness, and other forms of identity for collective social change.

Learners must be equipped to challenge White Eurocentric cultural norms, ideals, and social values in defining the global humanity and, in general, what it means to be “human” in a way that subverts the ideological tenets of Western individualism, freedom, liberty and choice. A neo-liberal corporate agenda has fostered a climate of competitiveness and a culture of hierarchy that works against the view of “schooling as community” and cultivating and a “community of learners.” The political economy of education must move beyond material, physical, and economic determinations to consider matters of the heart, spirit, and mind. This holds the key to educational and social transformation.

Conclusion

The future of Afrocentricity lies in educators and students realizing the need for a cultural and spiritual regeneration—a sort of ancestral rebirth and a way to reconnect consciously with a cultural heritage. Western education is at times overly skeptical about spirituality, and thereby risks interpreting spiritual ontologies as mythologies and non-scientific knowledge. We do so at our collective risk. We must promote multiple knowledge forms if we are to develop counter measures of schooling and education to address the current challenges faced by our school system. For many learners, spiritual and emotional knowings are relevant knowledge for their self and collective identifications. African and Indigenous learning has long been anchored in an ancestral rebirth of local communities across multiple geo-spaces on the continent and in global diaspora embracing histories, cultures, politics and identities. The ability to sustain Afrocentric education as part of the counter-visioning of schooling will depend on the commitment to a philosophy that centers the African humanity and its dreams and aspirations.

References

- Anderson, J., 1988. *The Education of Blacks in the South, 1860–1935*. University of North Carolina Press, Chapel Hill.
- Asante, M.K., 1980. *Afrocentricity: The Theory of Social Change*. Amulefi Publishing Company, Buffalo, NY.
- Asante, M.K., 1987. *The Afrocentric Ideal*. Temple University Press, Philadelphia.
- Asante, M.K., 1990. *Kemet, Afrocentricity and Knowledge*. Africa World Press, Trenton, NJ.
- Asante, M.K., 1991. The Afrocentric idea in education. *J. Negro Educ.* 60 (2), 170–180.
- Asante, M.K., 2000. *The Egyptian Philosophers: Ancient African Voices From Imhotep to Akhenaten*. African American Images, Chicago.
- Blyden, E.W., 1994. *African Life and Customs*. Black Classic Press.
- Bush, L., 2004. Access, school choice and independent black institutions: a historical perspective. *J. Black Stud.* 34 (3), 386–401.
- Butchart, R.E., 1980. *Northern Schools, Southern Blacks, and Reconstruction: Freedmen’s Education, 1862–1875*. Greenwood, Westport, CT.
- Carroll, K.K., 2014. An introduction to African-centered sociology: worldview, epistemology, and social theory. *Crit. Sociol.* 40 (3), 257–270.
- Cohen, R., 2016. The Afrocentric education crisis. *Am. Prospect*. Retrieved from: <https://prospect.org/education/afrocentric-education-crisis/>.
- Dei, G.J.S., Kempf, A., 2013. *New Perspectives on African-Centered Education in Canada*. Canadian Scholars Press, Toronto.
- Dei, G.J.S., McDermott, M., 2019. *Centering African Proverbs, Indigenous Folktales and Cultural Stories in Curriculum: Units and Lesson Plans for Inclusive Education*. Canadian Scholars Press, Toronto.
- Dei, G.J.S., James, C., James-Wilson, S., Karumanchery, L., Zine, 2000. *Removing the Margins: The Challenges and Possibilities of Inclusive Schooling*. Canadian Scholars Press, Toronto.
- Dei, G.J.S., 2021. *The Black Scholar Travelogue in Academia*. Forthcoming.
- Du Bois, W.E.B., 1996. *Souls of Black Folk*. Penguin Classics.
- Hayes, F., 2003. Review of: Clarence Shole Johnson, 2003 “Cornel west and philosophy: the quest for social justice”. In: McClendon, J., Yancy, G. (Eds.), *Newsletter on Philosophy and the Black Experience*, vol. 2. APA Newsletters, New York: Routledge, 2.
- Hill, R.A., 1983. *The Marcus Garvey and Universal Negro Improvement Association Papers*. University of California Press, Berkeley.
- Karenga, M., 1991. *African Culture and the Ongoing Quest for Excellence: Dialogue, Principles, Practice*. The Black Collegian Archives. Retrieved February 29, 2012 from: <http://www.black-collegian.com/african/karenga.shtml>.
- Karenga, M., 2004. *Maat: The Moral Ideal in Ancient Egypt: A Study in Classical African Ethics*. Routledge, New York.
- Kempf, A., 2011. North African knowledges and the western classroom: situating selected literature. In: Wane, N., Kempf, A., Simmons, M. (Eds.), *The Politics of Cultural Knowledge*. Sense Publishers, New York, pp. 93–110.
- Kempf, A., 2015. African centered education, Canada. In: Shujaa, M. (Ed.), *Encyclopedia of African Cultural Heritage in North America*. Sage Publications, p. 176.
- Kendi, I.X., 2016. *Stamped From the Beginning: The Definitive History of Racist Ideas in America*. Bold Type Books, New York.
- Ladson-Billings, G., 2018. *Wakanda, Zumunda, and Afromantic: critical race theory and Utopian blackness*. In: Keynote, Center for Integrative Antiracism Decolonizing Conference, Toronto.
- Lee, C.D., 1992. Profile of an independent Black institution: African-centered education at work. *J. Negro Educ.* 61, 161–177.
- Mirza, H.S., Reay, D., 2000. Spaces and places of black educational desire: rethinking black supplementary schools as a new social movement. *Sociology* 34 (3), 521–544.
- Murrell, P.C., 1999. Chartering the village: the making of an African-centered chartered school. *Urban Educ.* 33, 565–583.
- Purton, F., Styres, S., Kempf, A., 2020. Speaking back to the institution: teacher education programs as sites of possibility. In: Cote-Meek, S., Moeke-Pickering, T. (Eds.), *Decolonizing the Academy*. Canadian Scholars Press, Toronto.
- Rashid, H., Muhammad, Z., 1992. The sister Clara Muhammad schools: pioneers in the development of Islamic education in America. *J. Negro Educ.* 61 (4), 178–185.
- Ratteray, J., Shujaa, M.J., 1987. *Dare to Choose: Parental Choice at Independent Neighborhood School*. Institute for Independent Education, Washington DC.
- Ratteray, J., 1992. Independent neighborhood schools: a framework for the education of African Americans. *J. Negro Educ.* 61 (4), 139–147.

Reay, D., Mirza, H.S., 1997. Uncovering genealogies of the margins: black supplementary schooling. *Br. J. Sociol. Educ.* 18 (4), 477–499.

Roane, J.T., 2017. Pedagogy for the World: Black Studies in the Classroom and Beyond. *Black Perspectives*. January 8. Retrieved from. <https://www.aaihs.org/pedagogy-for-the-world-black-studies-in-the-classroom-and-beyond/>.

Shujaa, M.J., 1994. Afrocentric transformation and parental choice in African–American independent schools. In: Shujaa, M.J. (Ed.), *Too Much Schooling, Too Little Education: A Paradox of Black Life in White Societies*. Africa World Press, Trenton, NJ, pp. 361–376.

Watkins, W.H., 2001. *The White Architects of Black Education: Ideology and Power in America, 1865–1954*. Teachers College Press, New York.

Woodson, C.G., 1933. *The Mis-education of the Negro*. Booksellers & Associates.

Comprehensive sexuality education: young masculinity, cultural relevance and context matters

Deevia Bhana, Gender and Childhood Sexuality, University of KwaZulu-Natal, School of Education, Durban, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Masculinity, culture and context	172
Theorizing masculinity beyond hegemony: thinking through assemblages	173
Community, culture, ceremonies: masculinity and heterosexuality	173
Conclusion	176
References	177

Since the 2018 United Nations Educational, Scientific and Cultural Organization (UNESCO) technical report on comprehensive sexuality education (CSE), there has been a global reckoning with issues around gender and sexuality (De Leon and Allegrante, 2020; Goldfarb and Lieberman, 2021; Krebbekx, 2019; Miedema et al., 2020). CSE it is argued remains an important aspect in a multi-levelled approach required to address gender relations of power, sexuality, sexual health and wellbeing (UNESCO, 2018). Schools have been identified as significant sites for the delivery of CSE across the young life course from early childhood through to young adulthood (Keogh et al., 2021; Le Mat et al., 2019). In the context of young people's continued vulnerability to early childbearing, HIV, inconsistent contraceptive use, gender and sexual violence, CSE remains an important component in school-based interventions to address young people's sexual risk (Goldfarb and Lieberman, 2021). The case for CSE remains compelling especially in relation to enduring gendered patterns where hegemonic notions of masculinity place limits on what is possible to become (Rollston et al., 2020). Beyond a risk reduction focus, CSE is expected to provide young people with knowledge, value and skills to understand and address relationship dynamics, sexual consent, enjoyment and pleasure while affirming young people's sexuality and desires (Kantor and Lindberg, 2020; Ketting et al., 2021). Advocates of CSE suggest that expansive topics in CSE including sexual debut, sexual pleasure and enjoyment are affirming of young people's sexuality (Herat et al., 2018; Kirby, 2011; Vanwesenbeeck, 2020). Notably, this expanding focus on gender and sexuality diversity also elaborates CSE's core content that addresses inclusion, respect and sexual justice (Herat et al., 2018; Miedema et al., 2020; Vanwesenbeeck, 2020). CSE is expected to promote positive ideas toward gender and sexuality diversity and an approach that recognizes the contextual basis and cultural relevance in the teaching and learning of new values and skills that support rights-based discourses. As UNESCO (2018, p. 16) notes CSE involves the development of values that will capacitate young people to "realize their health, well-being and dignity; develop respectful social and sexual relationships; consider how their choices affect their own well-being and that of others; and, understand and ensure the protection of their rights throughout their lives." As Kirby (2011, p. 12) states:

As the world is confronted with the largest generation ever of adolescents and young people, there is no time to lose in investing in their future by providing them with the information and skills needed to ensure their sexual and reproductive health over a lifetime.

While CSE's global commitment toward gender justice, rights and inclusivity are widely supported (Santelli et al., 2018), there are many challenges. Across the globe debates about how best to intervene in school-based CSE and protect young people's sexual health and well-being continues (De Leon and Allegrante, 2020; Iyer and Aggleton, 2015).

At the heart of CSE is sex and putting sex and sexuality together with children and young people remains problematic, despite evidence of the merit of CSE (Herat et al., 2018; Ramaswamy, 2021; Vanwesenbeeck et al., 2019). Deep-seated anxiety and discomfort prevail about young people's association with sexuality as taboo and culturally shameful. Conservative forces, religious norms, abstinence-only education and the idea of the corrupting force of sexual pleasure and enjoyment are antithetical to CSE. In this crucible, an idealized notion of childhood sexual innocence is pre-eminent shaping and reinforcing conservatism and resistance to CSE (Schalet, 2011). Blistering debates ensued over CSE in light of UNESCO's (2018) core content that seeks to enhance young people's knowledge of, sexual diversity, sexual pleasure and masturbation beyond the risk-focused sexuality education. Strong opposition from well-funded organizations such as *Focus on the Family* in the US (Moore et al., 2021), assembled together to attack CSE. This attack was premised upon the presumption of childhood sexual innocence, the core content of CSE was deemed to promote promiscuity, tainting healthy sexual development through its focus on the erotic aspects of sexuality.

In South Africa, the setting for this chapter, *#LeaveOurKidsAlone* comprising teachers and parents aligned with dominant narratives opposed to CSE. Chaskalson et al. (2019) point to the vitriolic ranting, the fears and panic produced through this hashtag, in relation to young people's exposure to "satanically" motivated education. CSE is not easy to implement. Facilitating CSE is difficult

as it necessarily engages with difficult conversations about issues such as gender inequalities, violence and male culpability in it, homophobia, sexual diversity and pleasure. The affective force of these issues is heightened in relation to young age and young people precisely because sex, sexuality and sexual knowledge are highly controversial issues (Egan, 2013). CSE is “work never done, unfinished business” requiring attention to the “cultural, and political” and the social and ideological contexts through which it is situated (Vanwesenbeeck, 2020, p. 1).

Despite the widespread recognition of the need for CSE to be culturally relevant and contextually specific, the interaction between young boys and masculinity appears to be overlooked in debates and interventions around CSE. More specifically, the ways in which boys in rural primary school contexts make sense of masculinity remains largely absent in thinking about the contextual significance of these meanings to the development of expanded understandings of gender and sexuality diversity. As UNESCO (2019) notes, CSE should be scientifically accurate, age and developmentally incremental so that knowledge may be progressively built upon to create learned behavior and attitudes, culturally relevant, human rights based, transformative, contextually, and culturally appropriate. Notwithstanding these imperatives, a major gap in understanding the situatedness of gender and sexuality remain especially among younger boys in primary school settings. This chapter makes a case for addressing how young boys produce masculinity as a culturally exalted expression of manhood while limiting the expressions of gender and sexuality diversity. “CSE has to be culturally relevant and inclusive is increasingly recognized as a fundamental aspect of supporting young people to live healthy sexual lives,” as noted by Coultas et al. (2020, p. 1).

By drawing from new feminist materialism (Fox and Alldred, 2013; Fox and Bale, 2018), the chapter shows how different expressions, forms and content assemble together to both constrain and enable masculinity. Before engaging with the masculinity assemblage, the chapter considers the context within which masculinity is produced.

Masculinity, culture and context

This chapter brings together boys’ account of masculinity within a distinctively rural context where compliance with cultural norms is more evident than in urban settings (Hunter, 2010) and through which a culturally hegemonic form of masculinity is exalted (Connell, 1995). Customary practices have a strong foothold in the construction of masculinities in South Africa. Attention has been given to the ways in which masculinity, heterosexual domination, misogyny and violence cohere around hegemonic masculinity (Morrell et al., 2012). An understanding of hegemonic masculinity draws mainly from Connell (1995, p. 77) and defined as “the configuration of gender practice which embodies the currently accepted answer to the problem of legitimacy of patriarchy”. Hegemonic masculinity, following Connell (1995) can be viewed as a dominant version masculinity within a socio-cultural setting where male power is exerted over women and non-gender-conforming men. While there has been critique against this understanding of power based on domination and subordination and the typologies created between hegemonic masculinity and other forms of masculinity (Connell and Messerschmidt, 2005), hegemonic masculinity has shaped much of South African research on men and masculinities (Jewkes and Morrell, 2012). In this research careful consideration is given to the dynamic nature of masculinity and the value of understanding its cultural contextualization (Mfecane, 2018). Drawing on the concept of masculinity, it is possible to understand male power as an investment in customary practices and patriarchal norms which has produced inequitable gender practices and disparagement of gender non-conforming. In this regard and beyond the static account of domination and subordination, attention to the contextualization of race, class, sexuality, gender and cultural norms in the production of masculinity remain important work in delineating the specific experiences of masculinity (Ratele, 2016).

Within the KwaZulu-Natal region of South Africa, customary practices are actively negotiated in relation to gender and generational hierarchies. Historically power was vested in a system of chieftaincy where age and gender coalesced to produce reverence for older men who held power. In this system younger men, women and children were deemed to have less control in relation to gender and generational hierarchies (Hunter, 2010). These cultural practices are not static but in rural contexts they remain more widespread than in urban settings where fragile family structures has loosened the grip on these norms (Hunter, 2010). In many rural contexts, customary practices are underlined by male power, heterosexual domination and gender inequalities. *Hlonipha*, for instance is based on the concept of respect and deferential disposition toward the male head of the household. Masculine power derives from the system of chieftaincy, the historical emphasis placed on male economic power which was leveraged through cattle, the creation of a household, having many children and polygamy (Moore, 2015). Marriage was negotiated through the payment of *ilobolo* (bridewealth) where 11 cows were usually paid by the groom to the virgin bride’s family (Wickstrom, 2010). Virginity was thus prized. Hegemonic versions of masculinity have historical and material roots and takes shape at the intersection of heterosexuality, gender and cultural norms.

In describing masculinity in relation to these intersecting forces Mfecane (2018) suggests that the ways in which masculinity is produced in local settings is based upon customary practices. Indeed, cultural expectations around *hlonipha* (respect), *ilobolo* (bridewealth) and male role and expectations within rural settings provide a rich context of the ways in which boys negotiate masculinities. Within these broader social and cultural settings, patriarchal norms regulate sexuality in ways that sustain the notion of homosexuality as “un-African” (Epprecht, 2018; Msibi, 2018). While this essentialist understanding of sexuality has been critiqued, what is necessary here is that these binary positions between heterosexuality and homosexuality serve to uphold hegemonic masculinity. As will be shown in this chapter, an investment in culture and cultural practices is a signifier of masculinity and the presumption of heterosexuality and through which non-gender conforming others are subjugated.

It is in this context, that CSE operates. If CSE is to be driven by local realities then attention to the ways in which masculinity and sexuality are negotiated and grappled with remain important work. Significantly, the cultural setting and context is also one in which traditional customary practices sit alongside opportunities for the expression of rights. Cultural norms are changing and changeable and the transformative potential in thinking about masculinities and culture beyond the heterosexual imperative remain possible.

Theorizing masculinity beyond hegemony: thinking through assemblages

In this chapter I seek to expand how we currently theorize and understand the operation of young masculinity by highlighting data from focus groups with primary school black boys in a rural school in KwaZulu-Natal province of South Africa. The boys are aged between 11 and 13 years of age and emerge from rural settings which are historically produced in relation to race and class. While [Connell \(1995\)](#) has provided important understandings of masculinity within frameworks of social constructivist thinking, such structures in the absence of materiality seek to restrain the production of knowledge when the social and material world are separated. Drawing from new feminist materialism, this chapter illuminates the connection between the human and meta-physical world through the operation of assemblage theorizing ([Van der Tuin, 2011](#)) in order to account for the socio-material-cultural connections that are made in the production of masculinity. Cultural norms such as *inhlonipha*, practices, sexuality, rituals, *ilobolo*, masculinity, rurality, ideas, and thought are all in a continuous entanglement with the human body. Beyond the binary established between hegemony and subordination, a materialist perspective suggests that all elements, expressions, forms and context intra-act in the production of sexuality and masculinity. Material agency is not specific to boys' agency but operates within an assemblage as entangled and interconnected actors that support meaning about masculinity. Within new feminist materialism, an engagement of matter through social actions and events constitutes the assemblage that consequently produces capacities to affect or be affected by other forms and elements ([Deleuze and Guattari, 1987](#)). The term assemblage as coined by Deleuze and Guattari illustrates the rhizomic network of social life which is made up of "semiotic flows, material flows, and social flows simultaneously" ([Deleuze and Guattari, 1980/1987](#): 22–23). The assemblage is dynamic and is constituted in relation to moment to moment formations of elements and bodies which create flexibility and movement. When boys connect ideas, culture and masculinity with the body, they do so in ways that draw from the assemblage that is contextually situated. Therefore, local practices around men and masculinity remains powerful in affecting how power is validated. However, since the assemblage is unstable, it can also open up possibilities to change and become something other. We can thus understand how boys make sense of masculinity as a process of becoming going through material, social and semiotic flows of movement. Matter and meaning are thus entangled with materiality and the body. The continuous state of flux within an assemblage is therefore illustrative that aspects of masculinity cannot be reduced to the body but is produced relationally as the assemblage navigates to encounter other matter within its context. All matter within the assemblage has agential capacity to affect ([Barad, 2003](#): 181). Affects link matter to other matter. While the assemblage is dynamic and changing, the network based on its particular form and elements can also reinforce connections. In this chapter, the reinforcement of hegemonic masculinity based on cultural norms and practices are part of the habitual connections being made to constrain the assemblage and restrict masculinity to male power, hegemony and heterosexuality.

[Hickey-Moody \(2019\)](#) understands masculinity as the operation of affects, capacities and assemblages within [Butler's \(1988, 1993\)](#) notion of performativity. In other words, masculinity can be seen as a performance—a repeated stylization of the self- or as noted above as habitual. This does not mean that masculinity is fixed. Rather it is more-than-human and exists in relation to different sets of ideas, thoughts, bodies, culture, and practices. Affects challenge the habitual and repetitive stylization of identity and thus through affect assemblages are changeable within a particular context. In this regard [Hickey-Moody \(2019: 50\)](#) suggests that masculinities "are a creative, non-human force with no allegiance to the male body other than its capacity to affect it or be affected by it". The boys in this study are always becoming in/through the assemblage and are able to affect rather than limited by the embodiment of masculinity. By focusing on what boys can do and how they are constrained, the chapter entwined within a new materialist lens opens up possibilities for thinking about ways to reconfigure masculinity and sexuality in ways that support gender equitable relations and change.

In situating CSE within this complex making of masculinity and sexuality, the chapter develops an understanding of the entanglement of CSE with boys' actual lives and through which we can begin to understand how important it is to work through the constant process of becoming, the constraints and habituation in meaning and also the possibilities for rupturing the assemblage. This theorizing will offer new complexity in thinking about heterosexuality, CSE and gendered binaries within localized settings. Simply put, CSE changes from one setting to another and in each new assemblage different things manifest to shape sexual futures. In doing so, the chapter argues for cultural and contextual relevance in understanding CSE and for imperative of different strategies and support to advance the goals of CSE.

Community, culture, ceremonies: masculinity and heterosexuality

In this section, following a new feminist materialist insight, I seek to show how the assemblage operates in the production of heterosexual masculinity constraining the spaces for the expression of gender and sexuality diversity. Boys' ideas, thoughts and bodies are caught up within the cultural assemblage—which affect and is affected by a range of other elements within the assemblage. In

describing a “real” man in their rural community, the boys brought together ideas around gender discourses and hegemonic masculinity, associated with being dignified and manly. Specifically, as I will show the boys in this context, drew attention to cattle and livestock, herding, cultural dances, invoking the spirits of ancestors and slaughtering goats as culturally normative of masculinity and heterosexuality.

Following Hickey-Moody, masculinity is more-than-human. Building on new materialist (van der Tuin, 2011) ideas concerned with how matter interacts with thoughts, ideas and how it interacts, affects and is affected by other elements, masculinity is seen as a becoming rather than a fixed construct or restricted to its embodiment. To shed light on the interactive elements within the assemblage, the boys draw attention to particular capacities that engender particular ways of being based on hegemonic notions of masculinity:

Andile: He has to follow tradition and practice our culture ... and also be able to take care of his family ... and be able to do traditional rituals.

Bonginkosi: Men need to do these things because it was done this way by our forefathers ... and how men show that they are men.

Buyisiwe: Yes, I also agree that it is a law that is placed by the chief and the people who lived in this community before we were born.

Within the cultural assemblage, tradition, cultural practices, reverence for ancestors and the power of the chief are brought together to reinforce a version of masculinity based on hegemony. As Hickey-Moody (2019, p. 1) indicates masculinity is “taught and learned through embodied and symbolic sets of practices that take place in a range of places and are distributed across often quite complex networks”. In South Africa, Mfecane (2018) delineates these complex networks that include connections with ancestors suggesting the significance of spiritual identities within a broader set of gendered networks. This gendered network operates alongside kinship relations, traditional rituals, respect for the chief of the area and relations with the ancestors. The assemblage thus ignites a complex network through which hegemonic notions of masculinity are legitimated while excluding the possibility of becoming other. As Connell (1995) says, this idea of a culturally exalted form of masculinity generates pressure and anxiety about failing to achieve this idealized version of masculinity:

Dingani: For me I would say it means being a good example within your community and be a good example at home ... to keep your dignity as a man when you are amongst other men ... it means behaving like a man ... and also not being a *stabane*[gay] ... you must not behave like a *stabane* because behaving like that will not preserve our culture ... do the things that are expected to be done by men in the community.

Failing to be a “real” man has consequences and is stigmatized. In this assemblage, Dingani, brings together community prescriptions around gendered conduct with male dignity and culture while setting real men with dignity in opposition to “*stabane*”. In other words, the affectivity of these relations produces particular capacities that engender Zulu masculinity while disparaging “*stabane*”. Butler (1993) states that there is nothing real about masculinity, neither is masculinity based on individual agency. What is evident through the boys’ responses is that masculinity an “an embodied practice of citation” (Butler, 1988, p. 13). The cultural citations suggest the deep connection with context and the ways in which masculinity is produced in relation to community and cultural norms. The assemblage thus operates as a complex set of elements that come together, collectively to perform a version of masculinity that is reproduced in relation to cultural ideals. As Hickey-Moody (2019) indicates, masculinity can be viewed as an assemblage connecting people, matter, ideas, context and acts that are seen to be masculine. These materialities are relational and thus masculinity is contingent on matter, bodies, things and ideas in the assemblage where *stabane* is viewed as weak and othered:

Mandlenkosi: They [community members] don’t see him as a complete man ... they are suspicious of him ... they don’t regard him as a dignified man ... they see him as a weak man ... the people from that community don’t trust him ... they don’t trust that he is a real man ... he is a weak man like a *stabane* ... I think they don’t invite him during important cultural ceremonies in the community because he does not do what other men do ... He ends up not knowing his culture because ... because he doesn’t partake in cultural ceremonies with other men and also ... and also if he has children his children grow up not knowing their culture.

Within the masculinity-assemblage, community, culture, ceremonial practices, social standards around gender, ideas about weakness incapacitate particular identities such as “*stabane*”. In other words, becoming a man is part of the assemblage in which the capacity to be is intensified through community solidarity, cultural affiliation with ceremonial practices and strength. The assemblage produces a narrow conception of masculinity while aggregating thoughts, ideas and bodies within normative account of manhood. In doing so, the assemblage reinforces male power, disparages alternative forms of sexuality, and constraints possibilities to become other when heterosexual masculinity is upheld.

Throughout the study, the boys referred to the persistent power of Zulu cultural norms in the community for explaining the legitimacy of heterosexual masculinity. Culture, community and gender norms coalesce to produce “authentic” Zulu masculinity

Bhekani: I would say being a man means having a lot of cattle ... being able to herd ... and him being able to take care of this livestock ... I think he must know how to do our cultural dances like *indlamu* so that people can see that there is a man in that family ...

Cattle, care, masculinity, cultural dances and heterosexual family unit interacts with each other, affects and is affected by other materialities to produce manhood and real Zulu masculinity. In other words, Bhekani puts to work the gender binary within the assemblage in order to manifest real heterosexual masculinity. In making these associations between human and more-than-human elements, social-material-cultural relationalities are produced reinforcing Zulu masculinity in opposition to *stabane*.

While there is nothing essentialist about Zulu communities, the assemblage is produced in relation to a strong sense of community and interdependence which subjects masculinity to a particular kind of gendered patterning. In many rural contexts in South Africa, interdependency ties communities together underpinned by the African philosophy of *ubuntu*—a humanity toward other people in the community (Piper, 2016). The operation of culture, gender and community entanglement within the assemblage suggests how masculinity is made to be intelligible through the cultural norms which binds people together in a common sense of humanity. The assemblage is thus an intricate complex network of relations and affects and is affected by other elements within it. Being a real man is intertwined and interdependency operates within a web of relations that reflect the broader community. A person’s humanity rather than existing in individualist ways, show strong collectivist characteristics (Mkhize, 2006). As Rudwick and Posel (2014, p. 123) suggest, “the proverb *umuntu ngumuntu ngabantu* (a person is a person because ... other people) captures the importance of relationships with family and friends in Zulu society ... the frequent saying *bazothini abantu* (what will the people say) describes a deep concern with other peoples’ thoughts and opinions.” Underlined by an interweaving connection between family, community, culture and gender, the boys explicitly link men and masculinity with ascribed gender roles and with cultural norms sanctioned and regulated by broader community. As a theoretical tool new feminist materialism accounts for the varied and intricate ways through which human and more-than-human forms come together to produce meaning that is tied to the embodiment of masculinity alongside larger cultural contexts.

A unifying feature of much of this, is the power men wield within gender relations and in the context of Umgababa it is clear the power they wield in community dynamics. As Hearn and Morrell (2012) note, gender operates within social groups. In the context of Umgababa, community values, as the boys elucidate prescribe particular patterns of behavior which include participating in cultural ceremonies and rituals informed by male power. The boys use cultural norms as a mechanism to reinforce heterosexual masculinity, and defended through rituals and practices evident in the rural community. Furthermore, these norms are also magnified by Dingani through sentiments such as “you mustn’t do anything that is different” not doing anything that is different within this closely-knit rural community, maintains communal customs and renders a male individual “a dignified man”.

Bongani in the transcript below provides further clarity in this regard:

Bongani: I would say he has to know how to speak to the ancestors ... he must also know how to slaughter a goat and also attend cultural ceremonies ... because I saw it being done that way here in the community and also being done by the men in my family as well ... and also because things need to be done that way in order to preserve our culture.

The preservation of culture, the deep co-existence and collective relations between family and community are described by Bongani informed by and premised upon heterosexual masculinity. The cultural ritual in speaking to ancestors is associated with several aspects of daily life including birth, death, illness in seeking protection, guidance and validation from ancestors for community health and wellbeing (Mfecane, 2018). According to Mfecane (2018) patriarchal ideals assign older males within families the responsibility for “speaking” to the ancestors. Messages from the ancestors to the community are relayed through dance, songs and invoking the spirits of the ancestors by uniting the concrete real world with the spiritual world based on bringing wealth, health and well-being to the community (Mfecane, 2018). The male role in renewing community well-being is interdependent upon heterosexuality, male power and reproduction of cultural norms.

Another participant, Mandla referred to a man’s role during funeral proceedings to highlight male responsibility in facilitating numerous cultural aspects that are involved with conducting a funeral:

Mandla: For me he needs to know how to dig a grave when there is a funeral and also know our cultural funeral proceedings.

Despite women performing a pivotal role during the funeral proceedings, the cultural assemblage places men at the forefront to lead funeral planning and processions. The powerful role of men during funeral proceedings is assigned importance due to the physical exertion involved in digging the grave, if the deceased is to be buried in a rural area (Ngubane, 2004). Furthermore, men

are also assigned the leadership role during funeral proceedings because they are responsible for slaughtering animals (goat or ox/cow) before the night vigil, on the day of the funeral and a few days after the funeral for the cleansing ceremony (Ngubane, 2004).

The assemblage connects boys, matter, ideas, gendered discourses, communities, contexts, ceremonies, culture and cultural practices within an interlocking grid through which Zulu masculinity is produced. As Hickey-Moody (2019) suggests masculinity, matter and thought are not separate from each other.

Ayanda: ... he has to have a wife ... and also be able to have children ... he must have money for *ilobolo* [bridewealth] ... a man must follow the rules that were made by our elders ...

Ayanda brings attention to the heterosexual compulsion, marriage and cultural norms involving *ilobolo*. *Ilobolo*, as a customary practice informs marriages and is deeply rooted in the performance of heterosexuality where male economic power and provider masculinity is validated. Customary marriage practices while changing revolves around negotiation of *ilobolo* where the groom is expected to negotiate a payment (either in cash or cattle) with the family of the bride. Thus, the cultural assemblage invokes provider masculinity, kinship relations, and respect for elders ensuring the reproduction of cultural norms while limiting the capacity to become other.

While the assemblage is constraining, it is also in flux and changing. As noted by Deleuze and Guattari (1987), affects produced in the assemblage can “affect and be affected” producing shifts and changes. In this regard what can boys do in challenging the cultural assemblage through which idealized versions of masculinity is produced. Within new materialism, affect is a powerful concept which suggests that bodies, ideas, thoughts are changeable and emerge within social-material forces.

Bandile: ... there are many people that live a different life than the way others live ... and some people don't do cultural or traditional rituals because they don't believe in them ... There is nothing wrong if they don't do it because everyone must do what they want to do ... they must not be forced to do things that they don't want to do.

Bongani: It's their [gay people's] right to do things that they like or want to do ... they shouldn't be forced to do things that they don't want to do because they also have rights ... so the other men that are different should also be allowed to live their lives even though they are seen as different.

Following Deleuze and Guattari (1987), Hickey-Moody (2019) suggests that masculinities can be seen as creative with a capacity to affect and be affected. What is evident in the above is the flux and change. Within an assemblage that is constraining on capacities to become other, restricting the expression of gender and sexuality diversity, rights and instability of culture ignite new ways of thinking as this thinking aggregates matter with other matter in the assemblage (Deleuze and Guattari, 1987). Cultural norms aggregates bodies within a defined understanding of masculinity and therefore constraints. However, this is not fixed. Ideas around rights, the dynamic nature of culture point to the flexibility in thinking about masculinity and the heterosexual imperative as unstable. Masculinity is fluid and relational and becoming other interacts with rights discourses to expand sexuality in ways that go beyond heteronormativity. In other words, bodies are always becoming in/around assemblages (Renold and Ringrose, 2011) which create capacities to affect and be affected.

Conclusion

By addressing, young boys in the primary school as they make meaning of masculinities, the chapter has highlighted the cultural relevance and contextual specificity in addressing gender and sexual norms in CSE. The analyses has shown how habitual and repetitive patterns within the assemblage exalt a culturally specific form of masculinity which draws from a range of ideas, thoughts, gendered discourses, bodies, cultural practices, ceremonies, rituals and includes the broader social and cultural structures through which “real” men are idealized. By drawing on theoretical tools from new feminist materialism, the chapter has shown how ideas, thoughts, rituals, bodies, ceremonies, cultural practices, chieftains, ancestors, cattle produce specific capacities that make specific identities possible while limiting others. Being a real man thus emerges from the complex production of affective relations within the assemblage (Fox and Bale, 2018). Specifically, the evaluation of the assemblage shows stereotypical views portrayed around men and masculinity because definitions hinge upon male strength, power, male as the head of the household while associating cultural rituals, practices, ceremonies to masculinity thereby subordinating both femininity and expressions that go beyond heteronormativity. Stabane, was deployed as a stigmatizing category and remained othered within a cultural collectivity based on the heterosexual imperative. Culture was used as strategy within the assemblage to reinforce claims to an “authentic” Zulu masculinity by aggregating bodies, rituals and cultural practices into prescriptive understandings of gender and sexuality. Thus, the potential to become other is limited. However, the assemblage is always in flux and different elements, forms of expression and content has the potential to move bodies in different ways. In some instances, the assemblage was ruptured by rights discourses and ideas around

difference and the flexibility around culture suggesting small openings toward new ways of thinking about men, masculinity, heterosexuality and culture.

For CSE, this study illustrates the importance of understanding the contextual specificity and cultural relevance in primary schoolboys' sense making of men, masculinity and sexuality. I argued that local concerns, sensitivities and affects around culture and customary practices shapes how CSE is viewed and is an "absolute necessity for multicomponent approaches as part of CSE implementation" (Vanwesenbeeck et al. (2019, p. 297). Boys' negotiations of gender and masculinity need to be centered in understanding the cultural relevance of CSE. The approach taken in this study, drawing from new feminist materialism, is that masculinity is the capacity to affect and be affected. CSE is produced within an assemblage where masculinity is not located in genital capacities but in relation to ideas, gender discourses, chiefs, heterosexual family, ancestors, cultural practices, ceremonies, cattle, and these different forms and elements come together with their own capacities to affect one another. For example, a man is expected to marry, have children, perform cultural practices and participate in ceremonies in ways that delineate specific capacities associated with being a man. An approach to CSE that recognizes these multifarious elements (human and more-than-human) is necessary to illustrate the complex range of things and ideas that come together to give meaning to masculinity. In this way, culturally relevant interventions will recognize how particular assemblages are produced under specific contexts and circumstances. This will require questioning of and rethinking the social-material-cultural nexus through which the assemblage is formed. Questions about particular cultural norms which are both problematic for women and expressions of gender and sexuality diversity will be required. While CSE cannot tackle culture on a large scale, more can be done to recognize the ways in which homophobia, gender inequalities and male domination are reproduced as young boys attach culture to masculinity which offer little scope to expand potentials for becoming other. CSE must be understood in local settings by creating the space for boys (and girls) to develop their own narrative, ideas and themes that they consider to be important.

References

- Barad, K., 2003. Posthumanist performativity: toward an understanding of how matter comes to matter. *Signs* 28, 801–831.
- Butler, J., 1988. Performative acts and gender constitution: an essay in phenomenology and feminist theory. *Theat. J.* 40, 519–531.
- Butler, J., 1993. *Bodies that Matter: On the Discursive Limits of "sex"*. Routledge, London.
- Chaskalson, J., Bhana, D., Brouard, P., Hodes, P., Ngabaza, S., Silbert, P., Zungu, N., 2019. Knowledge is power: the case for comprehensive sexualities education in South Africa. In: *Daily Maverick: Maverick Citizen*. <https://www.dailymaverick.co.za/article/2019-11-26-knowledge-is-power-the-case-for-comprehensive-sexualities-education-in-south-africa/>.
- Connell, R.W., Messerschmidt, J.W., 2005. Hegemonic masculinity: rethinking the concept. *Gend. Soc.* 19, 829–859.
- Connell, R.W., 1995. *Masculinities*. Polity Press, Cambridge.
- Coultras, C., Campbell, C., Mohamedi, R., Sanga, U., 2020. Comprehensive sexuality education is "not for us": rethinking "cultural relevance" through young Tanzanians' identifications with/against intervention knowledge. *Soc. Sci. Med.* 265, 113239.
- De Leon Jr., R., Allegrante, J.P., 2020. Strengthening comprehensive sexuality education Anew: a commentary. *J. Sch. Health* 90, 825–829.
- Deleuze, G., Guattari, F., 1980, 1987. In: Massumi, B. (Ed.), *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press, Minneapolis.
- Deleuze, G., Guattari, F., 1987. *A Thousand Plateaus*. A&C Black, London.
- Egan, R.D., 2013. *Becoming Sexual: A Critical Appraisal of the Sexualization of Girls*. Polity press, Cambridge.
- Epprecht, M., 2018. Two decades of sexuality research in Africa south of the Sahara. *Sexualities* 21, 1276–1281.
- Fox, N.J., Alldred, P., 2013. The sexuality-assemblage: desire, affect, anti-humanism. *Socio. Rev.* 61, 769–789.
- Fox, N.J., Bale, C., 2018. Bodies, pornography and the circumscription of sexuality: a new materialist study of young people's sexual practices. *Sexualities* 21, 393–409.
- Goldfarb, E.S., Lieberman, L.D., 2021. Three decades of research: the case for comprehensive sex education. *J. Adolesc. Health* 68, 13–27.
- Hearn, J., Morrell, R., 2012. Reviewing hegemonic masculinities and men in Sweden and South Africa. *Men Masculinities* 15, 3–10.
- Herat, J., Plesons, M., Castle, C., Babb, J., Chandra-Mouli, V., 2018. The revised international technical guidance on sexuality education- a powerful tool at an important crossroads for sexuality education. *Reprod. Health* 15, 1–4.
- Hickey-Moody, A., 2019. *Deleuze and Masculinity*, vol. 210. Palgrave Macmillan, Cham, Switzerland.
- Hunter, M., 2010. *Love in the Time of AIDS: Inequality, Gender, and Rights in South Africa*. University of KwaZulu-Natal Press, Pietermaritzburg.
- Iyer, P., Aggleton, P., 2015. Seventy years of sex education in Health Education Journal: a critical review. *Health Educ. J.* 74, 3–15.
- Jewkes, R., Morrell, R., 2012. Sexuality and the limits of agency among South African teenage women: theorising femininities and their connections to HIV risk practices. *Soc. Sci. Med.* 74, 1729–1737.
- Kantor, L.M., Lindberg, L., 2020. Pleasure and sex education: the need for broadening both content and measurement. *Am. J. Publ. Health* 110, 145–148.
- Keogh, S.C., Leong, E., Motta, A., Sidze, E., Monzón, A.S., Amo-Adjiei, J., 2021. Classroom implementation of national sexuality education curricula in four low-and middle-income countries. *Sex. Educ.* 21, 432–449.
- Ketting, E., Brockschmidt, L., Ivanova, O., 2021. Investigating the "C" in CSE: implementation and effectiveness of comprehensive sexuality education in the WHO European region. *Sex. Educ.* 21, 133–147.
- Kirby, D., 2011. *The Impact of Sex Education on the Sexual Behaviour of Young People*. Population Division Expert Paper No. 2011/12. United Nations Department of Economic and Social Affairs, New York, NY.
- Krebbek, W., 2019. What else can sex education do? Logics and effects in classroom practices. *Sexualities* 22, 1325–1341.
- Le Mat, M.L., Kosar-Altiynelken, H., Bos, H.M., Volman, M.L., 2019. Discussing culture and gender-based violence in comprehensive sexuality education in Ethiopia. *Int. J. Educ. Dev.* 65, 207–215.
- Mfecane, S., 2018. (Un) Knowing Men: Africanising Gender Justice Programmes for Men in South Africa. University of Pretoria: CSA&G Press.
- Miedema, E., Le Mat, M.L., Hague, F., 2020. But is it comprehensive? Unpacking the "comprehensive" in comprehensive sexuality education. *Health Educ. J.* 79, 747–762.
- Mkhize, N., 2006. African traditions and the social, economic and moral dimensions of fatherhood. In: Richter, L., Morrell, R. (Eds.), *Baba: Men and Fatherhood in South Africa*. HSRC Press, Cape Town, pp. 183–198.
- Moore, E.V., Hirsch, J.S., Spindler, E., Nalugoda, F., Santelli, J.S., 2021. Debating sex and sovereignty: Uganda's new national sexuality education policy. *Sex. Res. Soc. Pol.* 1–11.
- Moore, E., 2015. Forms of femininity at the end of a customary marriage. *Gend. Soc.* 29, 817–840.
- Morrell, R., Jewkes, R., Lindegger, G., 2012. Hegemonic masculinity/masculinities in South Africa: culture, power, and gender politics. *Men Masculinities* 15, 11–30.

- Msibi, T., 2018. *Hidden Sexualities of South African Teachers: Black Male Educators and Same-Sex Desire*. Routledge, New York.
- Ngubane, S., 2004. Traditional practices on burial systems with special reference to the Zulu people of South Africa. *Indilinga* 3, 171–177.
- Piper, B., 2016. International education is a broken field: can ubuntu education bring solutions? *Int. Rev. Educ.* 62, 101–111.
- Ramaswamy, H.H.S., 2021. Will comprehensive sexuality education (CSE) help in youth development? *Health Educ.* 121, 140–149.
- Ratele, K., 2016. *Liberating Masculinities*. HSRC Press, Cape Town.
- Renold, E., Ringrose, J., 2011. Schizoid subjectivities? Re-theorizing teen girls' sexual cultures in an era of "sexualization". *J. Sociol.* 47, 389–409.
- Rollston, R., Wilkinson, E., Abouelazm, R., Mladenov, P., Horanieh, N., Jabbarpour, Y., 2020. Comprehensive sexuality education to address gender-based violence. *Lancet* 396, 148–150.
- Rudwick, S., Posel, D., 2014. Contemporary functions of ilobolo (bridewealth) in urban South African Zulu society. *J. Contemp. Afr. Stud.* 32, 118–136.
- Santelli, J.S., Grilo, S.A., Choo, T.H., Diaz, G., Walsh, K., Wall, M., Hirsch, J.S., Wilson, P.A., Gilbert, L., Khan, S., Mellins, C.A., 2018. Does sex education before college protect students from sexual assault in college? *PLoS One* 13, e0205951.
- Schalet, A.T., 2011. Beyond abstinence and risk: a new paradigm for adolescent sexual health. *Women Health Issues* 21, S5–S7.
- United Nations Educational, Scientific and Cultural Organization (UNESCO), 2018. *International Technical Guidance on Sexuality Education—An Evidence-Informed Approach*. UNESCO, Paris.
- United Nations Educational, Scientific and Cultural Organization (UNESCO), 2019. *Policy Paper 39: Facing the Facts: The Case for Comprehensive Sexuality Education*. UNESCO, Paris.
- van der Tuin, I., 2011. New feminist materialisms. *Women Stud. Int. Forum* 34, 271–277.
- Vanwesenbeeck, I., Flink, I., van Reeuwijk, M., Westeneng, J., 2019. Not by CSE alone... furthering reflections on international cooperation in sex education. *Sex. Educ.* 19, 297–312.
- Vanwesenbeeck, I., 2020. Comprehensive sexuality education. In: *Oxford Research Encyclopedia of Global Public Health*. Oxford University Press, Oxford, UK.
- Wickström, A., 2010. Virginity testing as a local public health initiative: a "preventive ritual" more than a "diagnostic measure". *J. Roy. Anthropol. Inst.* 16, 532–550.

Queer in school: the experiences of youth and educators in the United States

Andrew Leland^a and Melinda M. Mangin^b, ^a Kean University, Union, NJ, United States; and ^b Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Terms and concepts	179
LGBTQIA people as deviant and diseased	180
Legal criminalization	181
Medical pathologization	181
Implications for schools	181
Queer educators	182
School curricula	182
Conclusion	182
Increased visibility and protections for LGBTQIA people	182
Positive, public visibility	183
Expanded legal protections	183
Implications for schools	184
Queer educators	184
Queer students	184
Conclusion	185
Looking forward: toward a queer affirming future	185
References	186

The acronym LGBTQIA (lesbian, gay, bisexual, transgender, queer, intersex, agender/asexual) has become increasingly commonplace in the context of schools, if not fully embraced as a legitimate focus of educational research and attention. A comprehensive discussion of LGBTQIA school-related topics could fill an entire set of encyclopedias. Therefore, to adequately address the topic and avoid overgeneralizing, we focus on contemporary public schooling in the United States. This chapter responds to the questions: How have LGBTQIA students and educators experienced schooling in the United States? And, what are the outcomes of those school experiences for LGBTQIA individuals and communities?

In contrast to some other nations where education is under federal control, the United States' education system is controlled by states and localities. In the United States, the federal context provides some broad mandates, such as prohibiting discrimination on the basis of sex under Title IX of the Education Amendments of 1972, while still allowing for state discretion over education decision making. This state-based system results in variations across states; however, there are also observable nation-wide trends. We identify and describe two broad societal trends that have influenced the treatment and experiences of LGBTQIA students and educators in the United States over approximately one hundred years of contemporary schooling: (1) the treatment of LGBTQIA people as deviant and diseased and (2) increased visibility and protections for LGBTQIA people.

This chapter is limited in two regards. First, the need for brevity constrains the chapter's depth and breadth. Second, the extant scholarship on LGBTQIA school experiences has primarily focused on cisgender white men, eclipsing the experiences and histories of Transgender, Black, Indigenous, and People of Color (BIPOC). Despite these limitations, this encyclopedia chapter can serve as an introductory resource that invites further examination of how LGBTQIA identities intersect with other marginalized identities within the categories of ability, citizenship, class, ethnicity, race, and religion.

We begin by defining key terms and concepts as a way to reduce misconceptions about LGBTQIA people. Next, we describe each of the two societal trends that have influenced LGBTQIA students and educators. The first and most enduring trend reflects longstanding efforts to portray and treat LGBTQIA individuals as deviant and diseased. The second trend is characterized by increased visibility and protections for LGBTQIA people. We conclude with a discussion of what it might look like if schools were places of equity and liberation for LGBTQIA students and educators.

Terms and concepts

Misconceptions about LGBTQIA people can lead to mistreatment and discrimination. To reduce misinformation, we start by explaining the difference between sex, gender, and sexual orientation, three concepts that are commonly conflated or confused.

The predominant view of biological sex in the United States emphasizes two binary categories, male and female, which are commonly conveyed as mutually exclusive and easily identified based on external genitalia. From a scientific standpoint, however, sex is actually a complex combination of hormones, chromosomes, and anatomy. People with physiological characteristics that do not align neatly with the biology ascribed to males and females are known as intersex. *Intersex* (I) is a broad category that includes

people with such physical, hormonal, and/or chromosomal characteristics, such as hyperandrogenism or XXY chromosomes. Intersex people account for approximately 1.7–2% of the population (Blackless et al., 2000), similar to the percentage of redheads. Despite the variations that occur across humans, legal documents in the United States almost exclusively require either F (female) or M (male) sex designations. As a result, sex takes on both medical and legal aspects that are regulated by federal and state-level government entities. As a medico-legal category, the criteria used to define and determine sex can differ based on legal context and not solely on the basis of biology or medical expertise.

Gender is an internal sense of identity and, like sex, it is commonly understood as a binary—girl/woman or boy/man. Expectations about how boys and girls should express their gender are highly dependent on societal norms and can change across culture and time. When an infant's sex is assigned at birth, gender norms are also presumed. For example, it is generally expected that babies assigned male at birth will develop gender identities as boys and men. People who experience congruence between their sex assigned at birth and their gender are known as *cisgender*. *Cisnormativity* is the belief that cisgender identities are preferable over other gender identities. People who experience incongruence between their sex and gender are broadly understood as *transgender* (T). Transgender, sometimes written as *trans*, is an umbrella term that can include a wide variety of gender identities and expressions. Some trans people conform to binary gender norms albeit for a different sex from the one they were assigned at birth. Some transgender individuals challenge gender norms, expressing themselves in ways that defy stereotypes for boys/men and girls/women. Recent research suggests that 0.6% of people 13 and up and 1.4% of people 13–17 identify as transgender (Herman et al., 2022). People for whom gender is not a salient identity may identify as *agender* (A).

Transgender and gender-expansive people refer to themselves using a wide range of terms, including but not limited to *AFAB* (assigned female at birth), *AMAB* (assigned male at birth), *enby*, *nonbinary*, *gender fluid*, *genderqueer*, *agender*, *transmasculine*, *transfeminine*, *gender expansive*, or simply boy, man, girl, or woman. The language people use to represent their gender identity is highly personal and it is important not to make assumptions about someone's gender identity based on their gender expression—outwardly visible behaviors and appearance such as clothing, hairstyle, or mannerisms. For more information on gender see works by Lee Airtion (2018), *Gender: Your Guide: A Gender-Friendly Primer on What to Know, What to Say, and What to Do in the New Gender Culture*; Melinda Mangin (2020) *Transgender Students in Elementary School: Creating Affirming and Inclusive School Cultures*.

Sexual orientation refers to patterns of romantic, emotional, and physical attraction that contribute to sexual identity. Attraction often depends on how we “read” another person or interpret their gender expression, not necessarily based on their biological sex. In the United States, some common terms used to describe sexual orientation include *gay* (G), *lesbian* (L), *bisexual* (B), *asexual* (A), and *straight*. The term *gay* describes someone who is attracted to people of the same sex/gender. *Gay* can be used for both men and women; however, sometimes women prefer the term *lesbian*. The use of *homosexual* for gay people should be avoided because of its linkage to homophobic ideologies and stereotypical notions such as sexual promiscuity among gay men. *Bisexual* refers to individuals who are attracted to both males/men and females/women. *Asexual* describes a person who is not sexually attracted to other people but may still have a sexual libido. *Straight* refers to someone who is attracted to people whose sex/gender differs from their own. This is formally known as heterosexuality. *Heteronormativity* is the belief that heterosexuality is preferable over other sexual orientations. Another increasingly common term is *queer* (Q). Originally, queer was a synonym for “peculiar” and was used as a slur for gay and gender-expansive people throughout the twentieth century. In the 1980s LGB activists deliberately began to use queer to describe themselves. Reclaiming words that were once pejorative is a way for marginalized groups to assert socio-political agency. Today, LGBTQIA people increasingly identify as queer and in this chapter we use queer interchangeably with LGBTQIA.

In addition to LGB, other more nuanced terms are increasingly being used to describe sexual orientation. The growth in language related to sexual identities reflects new understandings that binary representations of sex and gender fail to convey the richness of human experience. For example, *pansexual* describes someone who is attracted to the entire array of people regardless of their sex or gender identity. *Polysexual* people are attracted to multiple genders but not all genders. *Demisexual* refers to someone who experiences sexual attraction only to people with whom they have an emotional connection.

The acronym LGBTQIA is an alliance of marginalized identities from across each of these three concepts: sexual orientation, gender, and sex. Alliances can help oppressed groups of people build social, political, and legal influence. In cisnormative and heteronormative contexts, where cisgender and heterosexual identities are consistently privileged over LGBTQIA identities, such alliances convey a shared history of oppression and can help protect against hostility and discrimination. For people with multiple marginalized identities from across categories like ability, citizenship, class, ethnicity, race, and religion, the adverse effects of cisnormative and heteronormative contexts are magnified. To be queer in school means facing the same cisnormative and heteronormative beliefs that characterize our larger society.

LGBTQIA people as deviant and diseased

In the United States in the 19th and 20th centuries, the legal and medical fields played a key role in reinforcing cisnormativity and heteronormativity. In this section, we describe how the legal criminalization and medical pathologization of LGBTQIA people

positioned them as deviant and diseased and rendered queer people invisible, restricting their access and agency in society. In this section, we also show that positioning queer identities as deviant and diseased ultimately influenced efforts to erase queer people from public schools throughout the twentieth century.

Legal criminalization

Identities and expressions other than cisgender and heterosexual have been criminalized for most of the United States' history. Legal scholar William N. Eskridge describes this history in *Dishonorable Passions: Sodomy Laws in America, 1861–2003* (2008). He explains that the legal system of the American colonies, and later the United States, included laws that attempted to regulate sexual intercourse as an act of procreation between married heterosexual couples. For example, anti-sodomy statutes prohibited sexual intercourse between consenting adult men and vagrancy laws prohibited engaging in or soliciting sex with others in public. These laws were increasingly enforced beginning in the late 1800s when purity movements, or efforts to control sexual behavior, gained traction. Eskridge explains that, by the 1920s, arrests increased tenfold in many areas, as did the widespread policing of sexual behavior particularly around sexual interactions between men.

By the late 1940s and 1950s, panic around sexual acts between two men became even more heightened with what has been referred to as the Lavender Scare. Public discourse and dominant narratives presented men who had sex with men as immoral, indecent and as a threat to national security. McCarthyist style policing ensued. Politicians, journalists, and citizens endeavored to find and remove these men from the public, particularly in the workplace. One such purge included President Dwight D. Eisenhower's signing of Executive Order 10450, which allowed the federal government to deny employment to lesbian and gay people (Johnson, 2009).

The 1960s and 1970s ushered in the gay rights movement and shaped anti-sodomy laws across the country. Beginning in Illinois in 1961, a number of states passed legislation to remove penalties associated with consensual sodomy. Nineteen additional states followed suit through the 1970s. At the same time, anti-sodomy laws remained a criminal offense in other states, and some even revised definitions so the law only applied to same-sex sexual acts. These state-based differences continued until 2003, when the Supreme Court of the United States ruled anti-sodomy laws unconstitutional in the landmark case, *Lawrence v. Texas* (2003). Despite this ruling, numerous states have not repealed their anti-sodomy laws, which reflects and contributes to the ongoing discourse of deviance associated with same-sex, sexual acts.

Medical pathologization

Throughout the twentieth century, the criminalization of queer people was bolstered by the medical field. Medical organizations played a significant role in crafting deviant classifications for queer people and served as mechanisms of surveillance and punishment within and beyond medical institutions. Prominent organizations such as the American Psychological Association (APA) and the World Health Organization (WHO) classified LGBTQIA people as mentally ill and diseased. This pathologization resulted from measuring LGBTQIA identities against a set of norms that were built upon narrow understandings of the human body, sexual orientation, and gender identity.

For example, same-sex and same-gender attraction, have undergone several iterations of being classified as a mental disorder by the Diagnostic and Statistical Manual of Mental Disorders (DSMs). The DSM classified homosexuality as a sociopathic personality disturbance in 1952, a deviation from normal sexual behavior in 1968, and an ego-dystonic disorder in 1973—a classification that was not completely removed until 1987. Similarly, the ninth revision of the International Classification of Diseases and Related Health Problems (ICD-9), developed by the World Health Organization (WHO), listed homosexuality as a mental disorder in 1977, and this designation was not removed until the ICD-10 in 1990.

The DSM and ICD have also contributed to the pathologization of gender identities other than cisgender and of sex traits that fall outside the narrow, subjective criteria for male and female. The DSM currently lists "gender identity disorder" with diagnostic criteria that includes "a strong and persistent cross-gender identification" and "persistent discomfort with his or her sex or sense of inappropriateness in the gender role of that sex." Similarly, the latest revision of the ICD provides "gender incongruence" as a diagnostic category, which is defined as "marked and persistent incongruence between an individual's experienced gender and the assigned sex." With regard to sex traits, the latest versions of the ICD and DSM continue to pathologize intersex individuals as having disorders of sex development (DSD). Such a diagnosis commonly results in high-risk, genitalia-altering surgeries on non-consenting infants and children that can lead to long-term physical and psychological damage (Jordan-Young, 2010).

The unjust classification of LGBTQIA people as diseased has resulted in unnecessary and harmful medical intervention that has perpetuated the stigma of queer people, contributing to what Subini Annamma (2017) has described as a pedagogy of pathologization or the social construction of criminal identities based on deviant labels.

Implications for schools

The legal and medical positioning of queer people as deviant and diseased significantly influenced the treatment of queer people in schools. As a result, several mechanisms were developed to keep queer people and their identities out of education. These included efforts to remove and bar queer educators, as well as school curricula that perpetuated deficit perspectives or ignored queer people entirely.

Queer educators

Throughout the latter half of the 20th century, dominant discourses falsely positioned homosexuality as an ideology that could be indoctrinated, placing vulnerable student populations at risk. As a result, queer educators were perceived as a threat and education policies explicitly denied rights to queer educators or offered little to no employment protections (Blount, 2006; Lugg, 2006). Schools aggressively policed sexual orientation by identifying, removing, and barring queer educators from practice. Although queer educators were the target of such policing, all educators needed to prove their sexual orientation by maintaining social and cultural norms that corresponded with their gender identities (Kissen, 1996; Woods and Harbeck, 1992). For example, heterosexual marriage requirements and evidence of an athletic background served as “tests” intended to root out gay men from administrative roles. Interestingly, lesbians were rarely subjected to the same requirements, as they were not as prominent in popular thought and because women were paid less than men in the same profession (Blount, 2006).

Failure to meet heteronormative standards resulted in the loss of employment, public shame and ridicule, and in some cases criminal prosecution. In many regions across the country, queer educators were fired and sometimes had their professional credentials revoked. When dismissals were challenged at district, local, or state levels, courts often decided against queer educators on the grounds that educators were expected to serve as moral exemplars for students, and being queer violated notions of morality (Graves, 2009). State legislation also supported legal forms of discrimination. Some states attempted to pass laws that forbade the employment of lesbian and gay educators, while nearly half of US state statutes legally allowed employment discrimination on the basis of sexual orientation and gender identity (Graves, 2009).

School curricula

Negative beliefs about queer people as deviant and diseased also lead to their exclusion from school curricula. Throughout the latter half of the 20th century, materials with references to queer topics, figures, accomplishments and historical events rarely occurred (Lipkin, 1999). The few states and school districts that pursued LGBTQIA-related materials were often met with resistance from school boards, community members, policymakers, and educators (Biegel, 2010; Lipkin, 1999). One exception was the Massachusetts’ Safe Schools Program established in 1993 in the wake of increased suicides and other risk factors for queer youth (Lipkin, 1999).

Efforts to erase and stigmatize queer experiences and identities from school curricula were especially fervent in relation to sex education. During the early 1900’s, concerns about adolescent sexuality led to two divergent views, one that supported sex education as a way to bring attention to a so-called sexual crisis, and a second that feared sex education would contribute to immoral activity (Moran, 2009). Both viewpoints equated sexuality with heterosexuality. Exceptions existed but in ways that warned adolescents, particularly boys, about gay men being sexual predators. Efforts to include positive depictions of gay and lesbian people in sex education increased somewhat in the 1970s, coinciding with the gay rights movement and the APA’s removal of homosexuality as a mental illness from the DSM; however, these reforms gained little traction. In the 1980s queer inclusion in sex education became synonymous with HIV/AIDS education (Moran, 2009), further reinforcing negative beliefs about queer people as diseased and deviant.

By the end of the twentieth century, most sex education programs continued to portray heterosexuality and cisgender bodies as normative and advocated for abstinence only until marriage. Even schools with comprehensive sex education programs, which included information on contraception, continued to discourage sexual activity and stigmatized queer people by linking them to HIV/AIDS instruction (Kendall, 2012).

Conclusion

The legal criminalization and medical pathologization of LGBTQIA people reflected and reproduced heteronormative and cisnormative societal beliefs. The effects on LGBTQIA people have been dire, legitimizing a discourse of deviance and disease that pervades our organizations and institutions. As a result, policies and practices systemically work to perpetuate discrimination against LGBTQIA people, denying them the same agency and access to the same benefits as heterosexual and cisgender people. As a microcosm of culture and society, our educational systems have similarly contributed to this discrimination, resulting in the near absolute erasure of LGBTQIA individuals from our schools for most of US history. LGBTQIA educators and students alike have been forced into the “closet,” forced to deny their identities, experiences, and families or risk certain harassment, unemployment, and discrimination. This struggle for visibility is one that continues in the 21st century.

Increased visibility and protections for LGBTQIA people

The turn of the 21st century marked a time of increased negotiations between society at large and LGBTQIA activists who increasingly demanded equal rights in all aspects of public life. Following the civil rights protests of the 1960s and the AIDS crisis of the 1980s, America seemed to develop greater empathy for LGBTQIA individuals. Beginning in the 1990s, positive public visibility of LGBTQIA people increased and, in the 2000s, numerous rulings provided vital legal protections. At the same time, this steady advancement has been tempered by backlash and persistent bias. In this section, we discuss these influential societal shifts and the impact they have had on queer people in schools.

Positive, public visibility

The 1990s marked the beginning of a gradual shift in the national dialog about sexual orientation in the United States. Long-standing narratives that marked homosexuality as immoral and deviant were countered by increased positive public representations of gay and lesbian people. Attention to gender and positive representations of transgender people followed.

Mainstream media played a key role in shifting public beliefs about gay and lesbian people. Many of these efforts focused on normalization. Popular television series such as *Roseanne* and *Friends* broadcast episodes with positive portrayals of gay characters and topics (see: *Roseanne*, “Don’t Ask, Don’t Tell,” 1994 and *Friends*, “The One With the Lesbian Wedding,” 1995). In 1997, lead character Ellen DeGeneres of *Ellen* came out to an audience of 42 million viewers, a performance that earned her a Peabody award (see “The Puppy Episode”). Gradually, television programs featured gay lead characters in shows like *Will and Grace* (1998), *Queer as Folk* (2000), and *The L Word* (2000). In 2009, *Modern Family* featured two dads raising their daughter, a move that helped to normalize same-sex families as being like any other family.

Gradually, terms like gay, lesbian, and queer became part of the mainstream lexicon and opinion polls indicated increased acceptance of gays and lesbians. Despite positive advancements, public backlash was also on display. Homophobic hate crimes continued and anti-gay groups worked to intimidate and undermine LGBTQIA advocacy efforts. One example of anti-gay organizing is the Westboro Baptist Church (WBC), which is categorized as a hate group by the Southern Poverty Law Center (SPLC). In 1991, the Westboro Baptist Church (WBC) organized its first picketing ministry—protests designed to raise awareness of the church’s beliefs. The group gained notoriety in 1998 when members protested the funeral of Matthew Shepard, a gay college student brutally murdered in an anti-gay hate crime. WBC protesters target businesses and individuals that are perceived as supporting LGBTQIA rights and they commonly carry signs proclaiming, “God Hates Fags” and “Thank God for AIDS.” The WBC claims to have held 40,000 protests (SPLC, n.d.).

Although increased public visibility of LGBTQIA people has been viewed as a positive advancement, some queer activists have raised concerns that queer people’s acceptance is often conditional, granted only when queer people conform to dominant ideologies around gender, sex, or sexual orientation. The idea that marginalized peoples’ acceptance is contingent upon their compliance with mainstream society’s expectations is referred to as *respectability politics* (Higginbotham, 1993). *Respectable queerness* follows the same logic in which acquiring a respectable social identity means performing in ways that replicate heteronormativity and cisnormativity, such as assimilating to heterosexual constructions of domesticity (e.g., marriage, monogamy, procreation) or adhering to binary gender norms (Joshi, 2012). This creates a hierarchical system of privilege that illustrates respectable people as all-American, White, privately affectionate, and family oriented (Ho, 2019). Those who fail to meet respectability criteria are rendered voiceless, powerless, and invisible (Cohen, 1999; Ho, 2019; Story, 2016). Thus, advancements in LGBTQIA visibility have not extended equally to all queer people.

Expanded legal protections

The 2000s also marked a time of expanded legal protections for LGBTQIA people. Legal decisions focused primarily on whether to include LGBTQIA people in laws and policies that already protected cisgender and heterosexual people. One such decision was *Lawrence v. Texas* (2003), which eliminated sodomy laws and extended constitutional protections for privacy and consensual intimate contact to individuals regardless of sex, gender, and sexual orientation. Around the same time, some states expanded statutes to protect individuals based on gender identity and sexual orientation in policies that include housing, employment, and biased-motivated crimes. Laws around marriage and domestic partnership also changed.

Many of the new protections resulted from federal legal rulings. In *United States v. Windsor* (2013) the Supreme Court struck down the Defense of Marriage Act’s definition of marriage as being between a man and a woman. In *Obergefell v. Hodges* (2015), the court ruled that all states must recognize marriages between same-sex couples. It is important to note that the decision in *Obergefell* was largely justified by an emphasis on the importance of children being raised in family units with two parents, and thus, the *Obergefell* case hinged on respectability politics. In contrast, *Bostock v. Clayton County* (2020) helped to broaden legal protections beyond cisnormative and heteronormative beliefs. The *Bostock* ruling granted employment protections to LGBTQIA people under Title VII of the Civil Rights Act of 1964, which prohibits discrimination based on sex. Previously, Title VII had been used to advance equal opportunities for women, conflating “woman,” a gender category, with “sex.” Successful application of Title VII in the *Bostock* ruling expanded sex-based discrimination to include discrimination against LGBTQIA people whose identities are socially constructed in relationship to medico-legal definitions of sex.

Efforts to expand legal protections throughout the 2000s have not always been successful and some protections have been undermined. For example, almost half of all US states include broad statutory religious exemption laws, making religious organizations exempt from laws that conflict with their doctrines or beliefs. Some states specifically allow organizations with religious affiliations to deny services such as healthcare or adoption to LGBTQIA people. Supreme Court rulings have similarly threatened the rights of LGBTQIA people to access the same services as heterosexual and cisgender people. In *Masterpiece Cakeshop Ltd. v. Colorado Civil Rights Commission* (2018), in a 7-2 decision, the Supreme Court ruled in favor of a business owner who refused to create a cake for a same-sex wedding celebration. The 2022 Supreme Court decision to overturn *Roe v. Wade*, which protected the constitutional right to an abortion, was accompanied by calls for the court to reexamine other decisions on topics not explicitly addressed in the constitution. This includes the 2015 *Obergefell v. Hodges* decision, which argued for marriage as a fundamental right and liberty that extends to same-sex marriage. These legal machinations demonstrate the precarity of LGBTQIA rights.

More generally, it is important to note that advancements in visibility and legal protections have not been experienced evenly across all LGBTQIA people. Activism aimed at normalizing gay people often occurred at the expense of gender expansive and transgender people. Intersex people remain largely invisible and nonbinary people continue to fight for sex/gender markers that adequately convey their identities and experiences.

Implications for schools

The experiences of LGBTQIA people in schools at the turn of the 21st century parallel many of the developments in larger society. As public visibility and legal rights increased, so did the visibility of LGBTQIA students and educators and the calls for protection and inclusion. These efforts have been integral in developing safe and representative spaces for queer people. In many cases, they have also been met with pushback, revealing gaps in policy and practice that fail to include and affirm queer people in PreK-12 education.

Queer educators

Many scholars have documented the benefits of having openly queer educators in schools. Research from GLSEN (formerly the Gay, Lesbian and Straight Education Network) substantiates multiple positive benefits for students and overall climate (Kosciw et al., 2018). Out and visible queer educators help affirm the sexual orientation and/or gender identity of LGBTQIA students. Queer students tend to feel safer when an out adult works in their school and advocates for equity and social justice. Out teachers lead to decreased instances of student harassment and violence based on sexual orientation and gender identity. Despite these benefits, lack of adequate supports can inhibit educators from being out and visible.

By the 21st century, protections for queer educators varied considerably. A 2015 report documented inconsistencies across state policies, namely, that nearly half of US states did not include sexual orientation or gender identity in employment non-discrimination statutes (Wright and Smith, 2015). The report also used survey data from LGBT educators to document differential experiences across contexts and found that the extent to which schools protected, affirmed, and included queer identities created significant differences in school climate. In schools where policies existed for reporting harassment, professional development addressed queer identities, and curricula and resources were queer-inclusive, LGBT educators were less likely to fear and experience backlash about their sexual and/or gender identities. Differences in protections across the United States persisted until 2020 when the Supreme Court ruling in *Bostock v. Clayton County* (2020) granted employment protections to LGBTQIA people.

Although increased protections exist, coming out as a queer educator presents several challenges. Histories of and experiences with discrimination and oppression may instill feelings of fear and rejection. Many queer educators continue to worry about employment retaliation—either as it relates to job security or confronting harassment and instances of violence from colleagues, students, families, and community members. These risks are further amplified for queer educators who belong to and/or work in regional and cultural communities in which greater shame and stigma are associated with being queer (Brockenbrough, 2018; Cooper, 2016; Tooms, 2007). As a result, many queer educators remain closeted to some extent—either by not disclosing or making careful decisions about what to disclose to whom. Educators who remain closeted often engage in nondisclosure strategies, such as avoiding conversations about relationships or performing their gender in ways to appear non-queer. These strategies require having “dual identities,” where they are out in their private life but closeted as professionals (Cooper, 2016). This approach can jeopardize educators’ sense of authenticity, particularly when teaching and serving as role models for youth.

Queer students

The end of the twentieth century also saw growing efforts to create safe spaces in schools for gay and lesbian students (Griffin and Ouellett, 2003). The impetus came, in part, from a burgeoning body of research aimed at documenting the overall wellbeing of queer youth. In 1989, the U.S. Department of Health and Human Services released a report showing that gay and lesbian youth were three times more likely to die by suicide than heterosexual youth (Remafedi, 1994). Additional studies indicated that gay and lesbian youth experienced higher rates of victimization from peers in school settings (D’Augelli and Dark, 1994). In 1999, GLSEN released their first school climate survey, which exposed the depth of bullying directed at queer students, which included verbal harassment, physical assault, and sexual assault with limited intervention from faculty or staff (Kosciw and Cullen, 2002).

In response to these data, 21st century queer rights activists and their allies have fought for national attention, inciting calls to action and demanding safer school environments (Martin and Strom, 2019). One approach has been to lobby for state- and district-level anti-bullying policies. In 2011, 46 states had anti-bullying statutes but only a quarter of states had statutes with protections for sexual orientation and gender identity or expression (Stuart-Cassel et al., 2011). By 2019 all US states had adopted anti-bullying laws (U.S. Department of Health and Human Services, 2018); however, policies alone may be insufficient (Lugg and Murphy, 2017). For example, an examination of district-level data shows that only 1 in 10 school districts had anti-bullying policies that explicitly protected LGBT students (Kull et al., 2015).

Another approach has focused on addressing the systemic marginalization of LGBTQIA youth in schools through school-sponsored organizations (Mayo, 2004; Payne and Smith, 2013; Sadowski, 2016). Some schools developed student-led Gay-Straight Alliances or Gender and Sexuality Alliances (GSAs) as a way to create safe spaces for queer youth and their allies to meet, discuss interests, and receive support (Macgillivray, 2007). Research shows that students in schools with GSAs encounter fewer instances of bullying, harassment, and violence than students in schools without GSAs (Kosciw et al., 2018; Toomey et al., 2011). At the same time, GSAs have been critiqued as being composed predominantly of White queer youth or White allies (McCreedy, 2010; Pascoe, 2007). The presence of GSAs in schools has also been the source of considerable pushback and even legal

action. Detractors have argued, for example, that GSAs make school personnel feel uncomfortable (*Vanguard High School Gay-Straight Alliance v. Yancey et al.*) and, in the case of *Gonzalez v. School Board of Okeechobee County*, that a GSA would be incompatible with abstinence-only policies (American Civil Liberties Union, 2015). Ultimately, the courts have consistently ruled in favor of GSAs, holding that the 1984 Federal Equal Access Act guarantees students in secondary public schools the right to form GSAs.

In recent decades, advocates have also called for increased LGBTQIA visibility in school curricula. LGBTQIA curricular inclusion is linked to perceptions of and experiences with school safety (O'Shaughnessy et al., 2004; Russell et al., 2006). Queer students who learn in environments with less inclusive curricula are more likely to be victimized and harassed based on their sexual orientation and gender identity (Kosciw et al., 2018). They are also more likely to consider not graduating from high school (Palmer et al., 2016). California became the first state to mandate a more LGBTQIA-inclusive curriculum through the 2011 FAIR Education Act; however, by 2020, only three other states (New Jersey, Illinois, and Colorado) had followed suit and six states continued to deny more inclusive curricula through "No Promo Homo" laws that ban teaching about "homosexuality" as an "acceptable life-style" (Murphy and Lugg, 2016). Thus, few PreK-12 schools in the United States include LGBTQIA topics in curriculum and those instances are generally confined to sexual education classes, and to a lesser extent in language arts and history (Kosciw et al., 2018). Unfortunately, much of the curriculum and related texts have perpetuated deficit perspectives, such as stigmas around HIV/AIDS for gay men (Fields, 2008; Kendall, 2012; Snapp et al., 2015). Furthermore, inclusion rarely moves beyond treating sex, gender, and sexual orientation as rigid binaries and typologies (Kendall, 2012), which can have detrimental academic, social, and emotional effects on queer students (Kosciw et al., 2018).

Despite clear, consistent, and widespread research that justifies LGBTQIA inclusion in schools, queer visibility remains elusive. Education scholar Elizabeth J. Meyer (2015) suggests three reasons why schools have been slow to implement LGBTQIA policy and practice: (1) resistance stemming from the religious right; (2) fears that inclusion would lead to discussions of sexual behavior; and, (3) a lack of courage among practitioners, politicians, and professional organizations. Some scholars have also suggested that implementation of LGBTQIA policy and practice in PreK-12 schools may stem from the absence of LGBTQIA topics in training programs for teachers and school leaders (Jennings, 2012; Leland, 2020; Murray, 2014; Payne and Smith, 2018). As such, real improvements in school safety for LGBTQIA students have not kept pace with the need or with growing societal acceptance.

Conclusion

Some progress has occurred throughout the 20th and 21st centuries in terms of inclusion and visibility, more often, in broader social and political contexts than in schooling and education. Public visibility through media representations helped to shift mainstream beliefs about queer people. Major Supreme Court rulings found sodomy laws unconstitutional, granted equal marriage rights regardless of sex, and increased employment protections on the basis of sexual orientation and gender identity. At the same time, heightened backlash, including from the conservative US Supreme Court, appears imminent.

In our schools and educational settings, inclusion and protection have been slow to take shape. Queer educators continue to fear backlash. Policies and practices vary widely in their ability to protect queer students. School curricula continues to erase LGBTQIA people and related topics.

Looking forward: toward a queer affirming future

Existing scholarship on queer students and educators in the United States shows how constructions of sex, gender, and sexual orientation have normalized heterosexuality and cisgender bodies and how educational institutions reinforce these norms to the detriment of LGBTQIA people. Certainly, in the last twenty years there has been an increase in protections and visibility, but significant harmful limitations remain. Improvements to policy and practice have not occurred uniformly across the United States, with rural and suburban areas demonstrating less progress. Changes to policy and practice are often symbolic representations that lack the resources and accountability mechanisms needed to make meaningful improvements. The activism of queer people of color and their contributions to dismantling forms of oppression have been overlooked and underrepresented in accounts of LGBTQIA experiences.

Moving forward, we argue for educational policies and practices that affirm queer people, their identities, and their rights. This work extends from prior efforts to make queer people more visible and school spaces more inclusive. Visibility and inclusion should not be symbolic, or token gestures, nor should they depend on respectability politics. Rather, the advancement of LGBTQIA rights should reflect genuine, critical engagement that is sustained and embedded throughout educational institutions and the policies that drive them. Queer-affirming schools should acknowledge the full spectrum of queer subjectivities and recognize systems of privilege around sex, gender, and sexual orientation. They should challenge norms that privilege heterosexuality and cisgender identities, and disrupt the regulatory policing of sexual orientation, gender expression, and gender identity. Schools should also adopt approaches to inclusion that attend to the range of social and cultural issues that intersect with queer identities. Affirming queer people requires simultaneous efforts to interrogate forms of oppression and marginalization around race, ethnicity, nationality socioeconomic status, ability, and religious beliefs that are likely to exist in school policies and practices.

Moving forward, a queer-affirming future can only be accomplished through the involvement and collaboration of multiple education stakeholders. This includes researchers whose scholarship could expand upon the state of queer visibility and inclusion

in schools; policymakers whose efforts could challenge and rewrite policies that drive educational systems and structures; and practitioners whose instruction and leadership could create more inclusive school climates for queer students and educators. Further, the work of researchers, policymakers, and practitioners should mutually inform each other to help make sense of existing knowledge and practice, but also to pursue underexplored areas and find potential solutions to problems. By engaging in these critical conversations, we will gain a richer, deeper, and more accurate picture of the landscape of schooling and education for queer students and educators.

References

- Airton, L., 2018. *Gender: Your Guide: A Gender-Friendly Primer on What to Know, What to Say, and What to Do in the New Gender Culture*. Simon and Schuster, Avon, MA.
- American Civil Liberties Union, 2015. GSA court victories: a guide for LGBT high school students. <https://www.aclu.org/other/gsa-court-victories-guide-lgbt-high-school-students>.
- Annamma, S.A., 2017. *The Pedagogy of Pathologization: Dis/abled Girls of Color in the School-Prison Nexus*. Routledge, New York, NY.
- Biegel, S., 2010. *The Right to Be Out: Sexual Orientation and Gender Identity in America's Public Schools*. University of Minnesota Press, Minneapolis, MN.
- Blackless, M., Charuvastra, A., Derryc, A., Fausto-Sterling, A., Lauzanne, K., Lee, E., 2000. How sexually dimorphic are we? Review and synthesis. *Am. J. Hum. Biol. Offic. J. Hum. Biol. Assoc.* 12 (2), 151–166.
- Blount, J.M., 2006. *Fit to Teach: Same-Sex Desire, Gender, and School Work in the Twentieth Century*. SUNY Press.
- Bostock v. Clayton County, 590 U.S. (2020).
- Brockenbrough, E., 2018. *Black Men Teaching in Urban Schools: Reassessing Black Masculinity*. Routledge, New York, NY.
- Cohen, C.J., 1999. *The Boundaries of Blackness: AIDS and the Breakdown of Black Politics*. University of Chicago Press, Chicago, IL.
- Cooper, H.A., 2016. *The Teacher's Closet: Lesbian and Gay Educators in Georgia's Public Middle Schools*. Peter Lang, New York, NY.
- D'Augelli, A.R., Grossman, D., Starks, E., 1994. Lesbian, gay, and bisexual youths. In: Eron, L.D., Gentry, J.H., Schlegel, P. (Eds.), *Reason to Hope: A Psychosocial Perspective on Violence and Youth*. American Psychological Association, Washington, DC, pp. 177–196. <https://doi.org/10.1037/10164-008>.
- Eskridge, W.N., 2008. *Dishonorable Passions: Sodomy Laws in America, 1861-2003*. Penguin Random House, New York, NY.
- Fields, J., 2008. *Risky Lessons: Sex Education and Social Inequality*. Rutgers University Press, New Brunswick, NJ.
- Graves, K., 2009. *And They Were Wonderful Teachers: Florida's Purge of Gay and Lesbian Teachers*. University of Illinois Press, Chicago, IL.
- Griffin, P., Ouellett, M., 2003. From silence to safety and beyond: historical trends in addressing lesbian, gay, bisexual, transgender issues in K-12 schools. *Equity Excell. Educ.* 36 (2), 106–114. <https://doi.org/10.1080/106656803030508>.
- Herman, J.L., Flores, A.R., O'Neil, K.K., 2022. How many adults and youth identify as transgender in the United States? The Williams Institute, Los Angeles, CA.
- Higginbotham, E.B., 1993. *Righteous Discontent: The Women's Movement in the Black Baptist Church, 1880–1920*. Harvard University Press, Cambridge, MA.
- Ho, J.A., 2019. Queer sacrifice in masterpiece Cakeshop. *Yale J. Law Fem.* 31, 249–335.
- Jennings, T., 2012. Sexual orientation topics in educational leadership programmes across the USA. *Int. J. Incl. Educ.* 16 (1), 1–23.
- Johnson, D.K., 2009. *The Lavender Scare: The Cold War Persecution of Gays and Lesbians in the Federal Government*. University of Chicago Press, Chicago, IL.
- Jordan-Young, R.M., 2010. *Brain Storm: The Flaws in the Science of Sex Differences*. Harvard University Press, Boston, MA.
- Joshi, Y., 2012. Respectable queerness. *Columbia Hum. Right Law Rev.* 43 (2), 415–467.
- Kendall, N., 2012. *The Sex Education Debates*. University of Chicago Press, Chicago, IL.
- Kissen, R.M., 1996. *The Last Closet: The Real Lives of Lesbian and Gay Teachers*. Heinemann Educational Books, Portsmouth, NH.
- Kosciw, J.G., Cullen, M.K., 2002. *The GLSEN 2001 National School Climate Survey: The School-Related Experiences of Our Nation's Lesbian, Gay, Bisexual and Transgender Youth*. GLSEN, New York, NY.
- Kosciw, J.G., Greytak, E.A., Zongrone, A.D., Clark, C.M., Truong, N.L., 2018. *The 2017 National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual, Transgender, and Queer Youth in Our Nation's Schools*. GLSEN, New York, NY.
- Kull, R.M., Kosciw, J.G., Greytak, E.A., 2015. From Statehouse to Schoolhouse: Anti-bullying Policy Efforts in U.S. States and School Districts. GLSEN, New York, NY.
- Lawrence v. Texas, 539 U.S. 558 (2003).
- Leland, A.S., 2020. Doing family differently in school-to-home contexts: gay fathers at the schoolhouse door. *Educ. Adm. Q.* <https://doi.org/10.1177/0013161X20922582>.
- Lipkin, A., 1999. *Understanding Homosexuality, Changing Schools*. Westview Press, Boulder, CO.
- Lugg, C.A., 2006. Thinking about sodomy: public schools, legal panopticons, and queers. *Educ. Pol.* 20 (1), 35–58. <https://doi.org/10.1177/0895904805285374>.
- Lugg, C.A., Murphy, J.P., 2017. The shifting political winds: LGBTQ students, educational policy and politics, and the dilemmas confronting street level bureaucrats. In: Russell, S.T., Horn, S. (Eds.), *Sexual Orientation, Gender Identity, and Schooling: The Nexus of Research, Practice, and Policy*. Oxford University Press, New York, NY, pp. 238–254.
- Macgillivray, I.K., 2007. *Gay-straight Alliances: A Handbook for Students, Educators, and Parents*. Routledge, New York, NY.
- Mangin, M.M., 2020. *Transgender Students in Elementary School: Creating an Affirming and Inclusive School Culture*. Harvard Education Press, Cambridge, MA.
- Martin, A.D., Strom, K.J. (Eds.), 2019. *Exploring Gender and LGBTQ Issues in K-12 and Teacher Education: A Rainbow Assemblage*. Information Age Publishing, Charlotte, NC.
- Masterpiece Cakeshop, Ltd. v. Colorado Civil Rights Commission, 584 U.S. (2018).
- Mayo, C., 2004. Queering school communities. *J. Gay Lesb. Issues Educ.* 1 (3), 23–36. https://doi.org/10.1300/J367v01n03_04.
- McCready, L.T., 2010. *Making Space for Diverse Masculinities: Difference, Intersectionality, and Engagement in an Urban High School*. Peter Lang, New York, NY.
- Meyer, E.J., 2015. The personal is political: LGBTQ education research and policy since 1993. *Educ. Forum* 79 (4), 347–352. <https://doi.org/10.1080/00131725.2015.1069514>.
- Moran, J.P., 2009. *Teaching Sex: The Shaping of Adolescence in the 20th Century*. Harvard University Press, Cambridge, MA.
- Murphy, J.P., Lugg, C.A., 2016. No Promo homo Policies. *The SAGE Encyclopedia of LGBTQ Studies*. SAGE, Thousand Oaks, CA, p. 816.
- Murray, O.J., 2014. *Queer Inclusion in Teacher Education: Bridging Theory, Research, and Practice*. Routledge, New York, NY.
- Obergefell v. Hodges, 576 U.S. 644 (2015).
- O'Shaughnessy, M., Russell, S., Heck, K.C.C., Laub, C., 2004. *Safe Place to Learn: Consequences of Harassment Based on Actual or Perceived Sexual Orientation and Gender Non-conformity and Steps for Making Schools Safer*. California Safe Schools Coalition 4-H Center for Youth Development. Retrieved from. <http://www.casafeschools.org/>.
- Palmer, N.A., Greytak, E.A., Kosciw, J.G., 2016. *Educational Exclusion: Drop Out, Push Out, and the School-To-Prison Pipeline Among LGBTQ Youth*. GLSEN, New York, NY.
- Pascoe, C.J., 2007. *Dude, You're a Fag: Masculinity and Sexuality in High School*. University of California Press, Berkeley, CA.
- Payne, E., Smith, M., 2013. LGBTQ kids, school safety, and missing the big picture: how the dominant bullying discourse prevents school professionals from thinking about systemic marginalization or...why we need to rethink LGBTQ bullying. *QED: A Journal in GLBTQ Worldmaking* 1–36.
- Payne, E., Smith, M., 2018. Refusing relevance: school administrator resistance to offering professional development addressing LGBTQ issues in schools. *Educ. Adm. Q.* 54 (2), 183–215.
- Remafedi, G., 1994. *Death by Denial: Studies of Suicide in Gay and Lesbian Teenagers*. Alyson Publications, New York, NY.
- Russell, S.T., Kostroski, O., McGuire, J.K., Laub, C., Manke, E., 2006. *LGBT issues in the curriculum promotes school safety*. (California safe schools Coalition research Brief #4). California safe schools Coalition. Retrieved from. <http://www.casafeschools.org/>.

- Sadowski, M., 2016. *Safe Is Not Enough: Better Schools for LGBTQ Students*. Harvard Education Press, Cambridge, MA.
- Snapp, S.D., Burdge, H., Licona, A.C., Moody, R.L., Russell, S.T., 2015. Students' perspectives on LGBTQ-inclusive curriculum. *Equity Excell. Educ.* 48 (2), 249–265.
- Southern Poverty Law Center (n.d.). Westboro Baptist Church. Retrieved from. <https://www.splcenter.org/fighting-hate/extremist-files/group/westboro-baptist-church>.
- Story, K.A., 2016. On the cusp of deviance: respectability politics and the cultural marketplace of sameness. In: Johnson, E.P. (Ed.), *No Tea, No Shade: New Writings in Black Queer Studies*. Duke University Press, Durham, NC, pp. 362–379.
- Stuart-Cassel, V., Bell, A., Springer, J.F., 2011. Analysis of State Bullying Laws and Policies. Office of Planning, Evaluation and Policy Development, U.S. Department of Education.
- Toomey, R.B., Ryan, C., Diaz, R.M., Russell, S.T., 2011. High school gay–straight alliances (GSAs) and young adult well-being: an examination of GSA presence, participation, and perceived effectiveness. *Appl. Dev. Sci.* 15 (4), 175–185. <https://doi.org/10.1080/10888691.2011.607378>.
- Tooms, A., 2007. The right kind of queer: fit and the politics of school leadership. *J. Sch. Leader.* 17 (5), 601–630. <https://doi.org/10.1177/105268460701700503>.
- United States v. Windsor, 570 U.S. 744 (2013).
- U.S. Department of Health and Human Services, 2018. Laws, policies and regulations. Retrieved from. <https://www.stopbullying.gov/laws/index.html>.
- Woods, S.E., Harbeck, K.M., 1992. Living in two worlds: the identity management strategies used by lesbian physical educators. In: Harbeck, K.M. (Ed.), *Coming Out of the Classroom Closet: Gay and Lesbian Students, Teachers and Curricula*. Harrington Press, Nyack, NY, pp. 141–166.
- Wright, T.E., Smith, N.J., 2015. A safer place? LGBT educators, school climate, and implications for administrators. *Educ. Forum* 79 (4), 394–407. <https://doi.org/10.1080/00131725.2015.1068901>.

Multilingualism at school and multilingual education

Jasone Cenoz^a and Durk Gorter^b, ^a University of the Basque Country, UPV/EHU, Donostia-San Sebastian, Spain; and ^b University of the Basque Country, UPV/EHU-Ikerbasque, Basque Foundation for Science, Donostia-San Sebastian, Spain

© 2023 Elsevier Ltd. All rights reserved.

Multilingualism at school	188
Types of bilingual and multilingual education	189
Traditional trends in multilingual education	190
New trends in multilingual education	190
Translanguaging	191
Pedagogical translanguaging	192
Conclusions	192
Acknowledgments	193
References	193
Relevant website	194

Multilingualism at school

Multilingualism can be defined as “the ability of societies, institutions, groups and individuals to engage, on a regular basis, with more than one language in their day-to-day lives” (European Commission, 2007: 6). Nowadays multilingualism has spread in the educational context because of historical, social, political and economic reasons.

Multilingualism in education takes many forms depending on factors such as linguistic diversity, the aims of the educational program or the characteristics of the languages involved. In many contexts, a great number of schoolchildren have different home languages from the ones used at school. Today, school classes are more linguistically diverse than in the past, particularly in the case of big cities in Europe and North America, due to the mobility of the population. In a globalized world, even young learners come to the classroom with rich and dynamic trajectories that are influencing their language learning process. These students have to learn the language(s) of the host country and other languages in the curriculum, which in many cases are different from the ones spoken at home.

Multilingualism in education is often related to the mobility of the population but it can also be related to the use of two or more languages in the same region. In these contexts, the situation of the languages is often asymmetrical and one language, the majority language, is dominant while the other(s) have a minority status. Students who speak the minority language at home are also in most cases multilingual because the majority language is widely used and it is necessary to learn it.

The position of minority languages in education can range from a strong position of having the minority language as the main language of instruction to a very weak position of making no provision for the minority language in the curriculum. Educational policies on the role of regional minority languages in the curriculum can change over time. For example, the status of minority languages in Spain changed in the last decades of the 20th century. After many years of exclusion, Basque, Catalan and Galician became co-official languages in their regions in the late 1970s and early 1980s and were strongly reinforced because of their use in education. Minority languages are defined by their status as compared to the majority language. A minority language such as Catalan with approximately 10 million speakers has a minority status as compared to Spanish in Spain but has more speakers than other European state languages such as Danish, Greek or Swedish. As Gorter (2001) explains, the weak status of a minority language and its use at school may be associated with the limited number of speakers (for example, Navajo in the US or Frisian in the Netherlands), or their limited functions in society even when the demography is strong (for example, Quechua in Peru, Bolivia and Ecuador). The status of the minority language can have an impact on its use in education but other factors such as the linguistic distance between the minority and the majority language can also be important. When the majority and minority languages are typologically related, such as Catalan and Spanish or Frisian and Dutch, it is easier to use the minority language in education because there are fewer comprehension problems. When there is typological distance between the majority and the minority language, such as Māori and English or Basque and Spanish, more effort is needed to use the minority language in education.

Multilingualism in education does not necessary imply multilingual education. Due to the existence of regional minority languages or the mobility of the population, students may have different home languages but this situation does not mean that multilingual competence is necessarily considered a goal in education. Multilingual education can be defined as “the use of two or more languages in education provided that schools aim at multilingualism and multiliteracy” (Cenoz and Gorter, 2015: 2). Therefore, we refer to multilingual education when multilingualism is one of the educational goals in a school, region or state. In this case the educational system makes provisions for teaching languages either as subjects or integrating language and content. This does not necessary mean that the students’ first language(s) (L1) are among the languages to be learned.

Multilingual education can be aimed at speakers of majority or minority languages. Students with the majority language as a first language are sometimes enrolled in educational programs with a minority language as the language of instruction. This is common in some regions such as Wales or the Basque Country where the minority language has become the main language of instruction. These programs are at the same time enrichment programs for students with the minority language as a home language and immersion programs for speakers of the majority language as home language.

Apart from minority and majority languages, nowadays multilingual education in different parts of the world is characterized by the presence of English in the curriculum. This is the case not only when English is the majority language in countries such as the UK, the US, Canada, Australia or New Zealand but also in countries where English is a second or foreign language. English is taught as a subject in schools in most countries in the world and it is increasingly used as a language of instruction at all levels of education. English is more important in the curriculum than any other foreign language in the past. English is introduced earlier and it is often used as an additional language of instruction.

Sometimes the term “bilingual education” is used when there are two languages in the curriculum and multilingual education when there are more than two languages. Bilingual programs involving minority languages in non-English speaking countries have moved from bilingual to multilingual education because English has been added to the curriculum from an early age (Gorter et al., 2014a,b). Nowadays, educational policies including minority languages often include three languages which are English, the majority language and the minority language, except in cases where English is the majority language as in Welsh or Māori schools. Some examples of this situation include: Frisian, Dutch and English in Friesland, the Netherlands (Gorter, 2005); Basque, Spanish and English in the Basque Country, Spain (Cenoz, 2009) or Mongolian, Chinese and English in the Inner Mongolian Autonomous Region of China (Yi and Adamson, 2019).

Multilingual education aims at multilingualism for students who speak different home languages and learn additional languages at school. Globalization, the mobility of the population or the protection of minority languages are important factors for the development of multilingual education. Multilingual education is inclusive education that acknowledges and promotes linguistic diversity.

Schools that aim at multilingualism can be found in different parts of the world and involve different types of languages, different pedagogies and sociolinguistic contexts. This situation creates the need for new approaches that can deal with the diversity of students and the increasing role of languages in the school curriculum.

Types of bilingual and multilingual education

As it has been said, the distinction between bilingual and multilingual education usually refers to the number of languages in the curriculum but the term “multilingual education” is sometimes used as a generic term.

There are ways in which the different types of bilingual and multilingual education can be analyzed. Baker and Wright (2017) refer to bilingual education and distinguish strong and weak forms of bilingual education according to the language background of the child, the language of the classroom, and the linguistic, societal and educational aims. Weak forms of bilingual education do not have bilingualism as their aim and are submersion, segregationist, transitional, mainstream and separatist. Strong forms of bilingual education foster bilingualism and are immersion, maintenance and heritage language bilingual education, dual language bilingual education and bilingual education in majority languages. The typology is very comprehensive and most educational systems could be included in one or more of these categories but as Baker and Wright (2017: 193) acknowledge “there are multitudinous sub-varieties” and “not-all real-life examples will fit easily into de classification”.

The number of types and sub-varieties can entail a substantial increase when three or more languages are considered. Cenoz (2009) proposed a different perspective to describe, compare and evaluate different types of multilingual education in the *Continua of Multilingual Education*. This model represents the features of multilingual education as continua rather than in closed categories and looks at linguistic, sociolinguistic and educational factors. The idea is that when two or more multilingual programs are compared along these three dimensions they are represented in different positions along continua.

When analyzing the characteristics of multilingual programs using the *Continua*, the linguistic factor is language distance and the ends of this continuum are “more distant versus less distant”. Linguistic distance is relative and the home languages and the languages in the curriculum can be related to each other to different degrees not only because of origin but also because of contact. Linguistic distance can influence the design of the different programs and the need to devote more or fewer hours to the different languages either as subjects or as languages of instruction. It can also have some influence on the degree of multilingualism of the teachers. For example, French is closer to English than to Japanese but there are also words that French and Japanese or English and Japanese share. Linguistic distance is important when examining multilingual education because languages that are closer share more aspects of their syntactic structure, discourse, phonology or lexis and may be easier to learn.

At the sociolinguistic level, the variables to consider are the use of the languages by students at home and in their social networks, the number of speakers of the different languages in the community, the status of the language, their use in the media or in the linguistic landscape outside school. Different programs can be represented as “more or less multilingual” along the continua for these dimensions. Multilingual programs are more multilingual when students use more languages and when the schools are located in a context that is more linguistically diverse.

The educational variables that can indicate if a program is more or less multilingual are: subject, language of instruction, teacher and school context. The multilingual programs being compared can be represented as “*more or less multilingual*” along the continua for these dimensions. ‘Subject’ refers to the use of different languages as school subjects and the intensity and grades in which the languages are taught. ‘Language of instruction’ considers the teaching of content subjects through different languages. ‘Teacher’ looks at the multilingual profile of the teachers and their specific training in multilingual education. ‘School context’ refers to the diversity of languages used among teachers, with supporting staff, with parents or in the linguistic landscape of the school (see Cenoz and Gorter, 2019). The continua of multilingual education model highlights the interaction of linguistic, sociolinguistic and educational variables.

The *Continua of Multilingual Education* is a tool to compare different multilingual education programs but could also be used for comparing the degree of multilingualism in the linguistic, sociolinguistic and educational dimensions at two points in time so as to see progress in multilingual education. The use of continua to compare different multilingual programs is a dynamic approach without fixed categories and uses open categories allowing a wider diversity of situations, particularly when three or more languages are involved.

Traditional trends in multilingual education

Even if multilingual education aims at developing competences in several languages the traditional approach has been to focus on one language at a time and to separate languages at school. Some teachers who are multilingual are told or decide themselves to pretend they are monolingual so as to use only the target language. Some schools have tried to have a designated classroom to teach only one language and have posters and charts on the walls only in one language. In some English classes students even use English names instead of their real names. Pedagogical practices have been based on isolating languages by using only the target language and avoiding translation or any use of the first language(s). These traditional policies are associated with monolingual ideologies and the idea that the best way to develop multilingualism is through monolingualism. The monolingual native speaker is seen as the ideal goal to be achieved and learners are seen as inferior to the native speaker because even if they make progress in most cases they will still be different from native speakers of the target language.

Traditional monolingual views are widely spread in multilingual education. Using only the target language in class has been justified as a way to avoid interference from other languages and also because this policy of separation can provide more exposure to the target language and opportunities to use it. There is also the idea that both teachers and students have to work harder if only the target language is used in the classroom. However, monolingual ideologies and policies of language separation have also been criticized (Cummins, 2017). One criticism is the idea is that multilingual speakers are different from monolinguals and should not be considered deficient when compared to monolingual speakers because their language competence should not be evaluated using monolingual standards (Cook, 2013). Moreover, multilingual speakers have more resources at their disposal and they can use elements only from one language or from two or more languages depending on the communicative context and the linguistic skills of their interlocutors.

New trends in multilingual education

In the last years there have been theoretical and pedagogical proposals that go against the strict separation of languages in multilingual education and the reference to the ideal monolingual speaker (Cummins, 2017; Cenoz and Gorter, 2020). These ideas can be summarized with reference to the model “Focus on Multilingualism” which considers three dimensions: the multilingual speaker, the multilingual repertoire and the social context (Cenoz and Gorter, 2014).

The multilingual speaker. As we have seen, multilingual speakers are different from monolingual speakers. The idea is that multilingual speakers cannot be considered “*failed native speakers*” because they often are “*highly skilled communicators who make use of their multilingual resources*” (Jenkins et al., 2011: 284). In fact, multilingual speakers can communicate more effectively in more situations and with a larger number of interlocutors. An important characteristic of multilingual speakers is that their linguistic trajectories are richer and more dynamic than those of monolingual speakers. Quite often as the result of mobility or being speakers of minority languages, multilingual speakers have experienced many different linguistic situations that have shaped the way they communicate (see also Canagarajah, 2018).

Moreover, multilingual speakers navigate between languages and their discourse practices are different from those of monolinguals because of the resources they can use (Mauranen, 2018). The type of competence a multilingual speaker has in one language is difficult to isolate from the other languages because all the languages are intertwined. Multilingual speakers’ competence is fluid because there are no hard boundaries between languages. Multilingual competence is also dynamic because it can change over time according to the multilingual speaker’s experience with the different languages. An important point related to the characteristics of multilingual speakers is that their competence is different and the monolingual perspective of achieving perfect knowledge in each of the languages is not only unreachable but unfair (Grosjean, 2010). Multilingual ideologies consider that multilingual speakers are not the sum of several monolingual speakers because they have a different type of competence that cannot be compared to that of monolinguals (Cenoz and Gorter, 2020).

The whole multilingual repertoire. The whole multilingual repertoire refers to the multilingual speaker's linguistic knowledge related to all the languages s/he has experience with (see also [Hall, 2019](#)). According to [Cenoz and Gorter \(2014, 2015\)](#), multilingual speakers can be more effective learners and users of a target language if they are allowed to use resources from their whole linguistic repertoire. Multilingual speakers have a rich repertoire that includes not only linguistic elements such as vocabulary or syntactic rules but also their whole trajectories as language learners and language users.

The whole multilingual repertoire can be regarded as prior knowledge based on previous experience and the use of this knowledge is crucial for learning ([Bransford et al., 2000](#)). In the case of multilingualism, languages support one another at the cognitive level. Research on third language acquisition shows that bilinguals have advantages when learning additional languages because of their enhanced metalinguistic awareness and their expanded linguistic repertoire ([Cenoz, 2013](#); [Jessner and Cenoz, 2019](#)). [Cummins and Persad \(2014\)](#) highlight the importance of integrating new input into existing cognitive structures in the case of students with diverse linguistic backgrounds because previous knowledge is the foundation for interpreting new information. Monolingual ideologies and practices prevent students from using the resources they already built over the years as a result of their multilingualism. New trends in multilingual education such as Focus on Multilingualism propose the activation of the whole linguistic repertoire because it can be a rich resource for learning. In fact multilingual learners can develop metalinguistic awareness by comparing elements of their different languages they know at different levels (phonetic, lexical, morphosyntactic, pragmatic, discursive) and use their resources cross-linguistically.

The social context. Multilingual speakers use the resources in their multilingual repertoire in different ways depending on the social context. The teaching and learning of languages needs to take into account the characteristics of the social context and the communicative needs students have. It is not the same to learn a language that is needed for many purposes in everyday life or a language that is only needed to read occasionally some reports. At the same time, it is important to consider that multilingual speakers have their own communities of practice and that students can use elements from different languages in a natural communicative way, provided that these elements are understood by their interlocutors. Multilingual pedagogies imply that the languages in the students' multilingual repertoire and other languages in society are taken into account without establishing hard boundaries between languages. The social context also implies that students develop their language awareness about the values and roles of different languages in their community and in other contexts.

Focus on Multilingualism "proposes soft boundaries between languages and relates the way multilingual students use their communicative resources in spontaneous conversation to the way languages are learned and taught at school" ([Cenoz and Gorter, 2015](#): 8). The idea is to consider the learner as a multilingual speaker who is allowed to use resources in the multilingual repertoire in different ways depending on the characteristics of the social context. The idea is to activate students' multilingual resources so that they can benefit more from multilingualism.

The Council of Europe proposed the Common European Framework of Reference (CEFR) in 2001 and already then considered that the aim is not to take the languages in isolation and to have the ideal native speaker as a model. A recent development that softens boundaries between languages can be found in the Common European Framework of Reference Companion (2018). This volume includes descriptors of plurilingual and pluricultural competence. These descriptors may not be completely ready to be used in different contexts yet but imply an important step toward considering the whole multilingual repertoire. An example of a B2 level descriptor for "Building on plurilingual repertoire" is the following ([Council of Europe, 2018](#): 162):

Can make use of different languages in his/her plurilingual repertoire during collaborative interaction, in order to clarify the nature of a task, the main steps, the decisions to be taken, the outcomes expected.

As it can be seen using more than one language in the same interaction is accepted in this descriptor.

A multilingual perspective in education goes against long-lasting traditions of language separation and monolingual ideologies and practices. As [Kramsch \(2012\)](#) says they can be considered revolutionary because they challenge the foundation of theoretical and applied linguistics and language teaching.

Translanguaging

The concept most widely used when describing new trends that reflect the shift from monolingual to multilingual ideologies in education is translanguaging. The origin of translanguaging can be found in bilingual education in Wales where Welsh and English are the languages of instruction. It refers to the pedagogical practice of alternating the languages used for input and output in a systematic way ([Lewis et al., 2012](#)). For example, students read a text in English and write a summary or discuss the text orally in Welsh. The idea is that translanguaging requires a deeper understanding because the languages reinforce each other when both are used.

The concept of translanguaging has extended beyond the specific pedagogical practices of alternating the languages used for input and output. [García \(2009: 45\)](#) defines translanguaging as "multiple discursive practices in which bilinguals engage in order to make sense of their bilingual worlds". According to this definition translanguaging refers to spontaneous practices in communication

that have not been designed as a teaching strategy and that could take place inside or outside the classroom. Cenoz and Gorter (2021) explain that translanguaging is an umbrella term that embraces different theoretical and pedagogical approaches. They are different perspectives but they all go against monolingual ideologies of language separation (Canagarajah, 2011; García and Li, 2014).

In some educational contexts translanguaging can be beneficial for speakers of minority languages because it can empower students if they are allowed to use the discursive practices they use for communication outside the school (Otheguy et al., 2015). However, in some contexts translanguaging can be felt as a threat for the use of minority languages. Jones and Lewis (2014) explain that in the Welsh context, translanguaging can also pose a risk in the case of minority languages because it may trigger the use of the majority language. In a study on teachers' beliefs and attitudes, Arocena, Cenoz and Gorter (2015) confirmed that some teachers in the Basque Country thought that translanguaging could weaken the position of Basque, the minority language. Cenoz and Gorter (2017) explain that in some contexts translanguaging could potentially empower majority language students and that is necessary to allocate specific spaces for using the minority language.

Translanguaging in this context, where everybody is proficient in Spanish and only one-third of the population is proficient in Basque, can be seen as a threat for Basque. It is not a threat for Spanish because it is the powerful language and the influence goes from the majority to the minority language. In fact, in this asymmetrical relation, students can use Spanish with very limited traces of Basque but because of the strength of Spanish in the social context, when they speak Basque they tend to use many elements from Spanish. When this is the case translanguaging does not empower speakers of the minority language because their language could disappear in the near future. Translanguaging practices can also be positive in the context of minority languages but it is necessary to develop strategies to maintain and develop minority languages (Cenoz and Gorter, 2017).

Pedagogical translanguaging

In the previous section we have seen that there are different perspectives when discussing the concept of translanguaging because of the complex and diverse contexts of multilingual education. However, in all cases translanguaging implies that the students' linguistic repertoire is an asset and that languages should not be taught as separate entities.

Pedagogical translanguaging implies softening the hard boundaries that have been traditionally built between languages in school contexts. Pedagogical translanguaging is planned by the teacher and can refer to the alternation of language in the input and the output but also includes other planned strategies based on the use of the students' multilingual repertoire as a resource (Cenoz and Gorter, 2021). Pedagogical translanguaging is different from spontaneous translanguaging which is not planned but occurs naturally among multilinguals in context where boundaries between languages are fluid. Multilingual speakers translanguage spontaneously in some contexts but they are not always fully aware of the resources they have as multilinguals and how they can use them to learn languages and academic content. Pedagogical translanguaging aims at developing metalinguistic awareness so that students become aware of their own linguistic resources and use languages more efficiently. Pedagogical translanguaging can be used to develop vocabulary, grammar, pragmatics, pronunciation or discourse.

In the case of vocabulary, by working across languages students can get awareness of their prior knowledge so as to compare cognates, derivatives or compounds in different languages. For example, in a project conducted in the Basque Country, primary school students took pictures of shops in their town and were asked to compare the elements in words like of "bookstore" in three languages (Basque, Spanish and English) so as to see how in some cases languages that are not typologically related such as Basque and English can share the same structure: liburu + denda in Basque, book + store in English. The results of this study indicate that pedagogical translanguaging can have a positive influence on vocabulary development (Leonet et al., 2020; see also Lyster et al., 2013). Pedagogical translanguaging practices to develop vocabulary are not limited to the languages of the curriculum and other home languages can also be brought to the classroom. Some studies have shown how the activation of cognates, that is words that are close in meaning, spelling and pronunciation, can also be very positive (Arteagoitia and Howard, 2015; Simpson Baird et al., 2016; Cenoz et al., 2021).

It may be more challenging to use resources from the multilingual repertoire when languages are linguistically distant but pedagogical translanguaging is not limited to vocabulary development and can be applied to pragmatic and discourse competence. For example, the structure of a letter or an argumentative text can be compared across languages so that students develop metalinguistic awareness and use their multilingual resources to the full extent.

Pedagogical translanguaging implies different teaching strategies and practices and the use of different languages in the same class. It goes against the strong tradition of language separation. It considers students as multilingual speakers who can use their resources to understand more deeply how languages are used and to make links between prior knowledge and new knowledge.

Conclusions

School populations tend to have greater linguistic diversity now than in the past years and at the same time in some contexts one of the aims of education is the development of multilingualism. These trends are related to the mobility of the population, the recognition of some minority languages and the spread of English as the main language of international communication. However there are still many challenges that multilingualism in school contexts faces. One is the different status of the languages. Multilingualism

in international languages is highly prestigious but multilingual schoolchildren who speak other languages are seen as a problem. In some contexts the idea is to forget about those other languages and to adopt a monolingual perspective that does not recognize schoolchildren as multilingual. The recognition of linguistic diversity and multilingual students' linguistic trajectories through multilingual education is a priority in education. It is also a priority to promote multilingual education that incorporates minority languages. Multilingual education goes beyond language classes because it encourages education that contribute to social justice.

In this article we have seen that it is possible to adopt a "Focus on multilingualism" that recognizes students as multilingual who have a multilingual repertoire that can be used in class. In this way multilingual speakers can have more resources at their disposal because they link their prior knowledge even in languages that are not in the curriculum to new languages and academic content.

Traditional approaches are being replaced by multilingual practices such as pedagogical translanguaging that can help learners to activate their own resources by softening language boundaries. The use of linguistic resources from different languages can develop metalinguistic awareness and a deeper understanding of the way languages work. Pedagogical translanguaging has a great potential not only to enhance the acquisition of different languages but also to support learning content.

Acknowledgments

This work was supported by the Agencia Estatal de Investigación, Spain [Grant Number FEDERPID2019-105818GB-I00] and the Basque Government [Grant Number DREAM IT-1225-19].

References

- Arteagoitia, I., Howard, L., 2015. The role of the native language in the literacy development of Latino students in the US. In: Cenoz, J., Gorter, D. (Eds.), *Multilingual Education: Between Language Learning and Translanguaging*. Cambridge University Press, Cambridge, pp. 61–83.
- Baker, C., Wright, W.E., 2017. *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters, Bristol.
- Bransford, J.D., Brown, A.L., Cocking, R.R. (Eds.), 2000. *How People Learn: Brain, Mind Experience and School*. National Academies Press, Washington, DC.
- Canagarajah, S., 2011. Codemeshing in academic writing: identifying teachable strategies of translanguaging. *Mod. Lang. J.* 95, 401–417.
- Canagarajah, S., 2018. Materializing "competence": perspectives from international STEM scholars. *Mod. Lang. J.* 102, 1–24.
- Cenoz, J., 2009. *Towards Multilingual Education*. Multilingual Matters, Bristol.
- Cenoz, J., 2013. The influence of bilingualism on third language acquisition: focus on multilingualism. *Lang. Teach.* 46, 71–86.
- Cenoz, J., Gorter, D., 2014. Focus on multilingualism as an approach in educational contexts. In: Blackledge, A., Creese, A. (Eds.), *Heteroglossia as Practice and Pedagogy*. Springer, Dordrecht, Netherlands, pp. 239–254.
- Cenoz, J., Gorter, D., 2015. Towards a holistic approach in the study of multilingual education. In: Cenoz, J., Gorter, D. (Eds.), *Multilingual Education: Navigating Between Language Learning and Translanguaging*. Cambridge University Press, Cambridge, pp. 1–15.
- Cenoz, J., Gorter, D., 2017. Minority languages and sustainable translanguaging: threat or opportunity? *J. Multiling. Multicult. Dev.* 38, 901–912.
- Cenoz, J., Gorter, D., 2019. Educational policy and multilingualism. In: Singleton, D., Aronin, L. (Eds.), *Twelve Lectures on Multilingualism*. Multilingual Matters, Bristol, pp. 101–132.
- Cenoz, J., Gorter, D., 2020. Teaching English through pedagogical translanguaging: breaking away from monolingual ideologies. *World Englishes* 39, 300–311.
- Cenoz, J., Gorter, D., 2021. *Pedagogical Translanguaging*. Cambridge University Press, Cambridge, UK.
- Cenoz, J., Leonet, O., Gorter, D., 2021. Developing cognate awareness through pedagogical translanguaging. *Int. J. Biling. Educ. BiLing.* <https://doi.org/10.1080/13670050.2021.1961675>.
- Cook, V., 2013. Multi-competence. In: Robinson, P. (Ed.), *The Routledge Encyclopedia of Second Language Acquisition*. Routledge, New York, pp. 447–451.
- Council of Europe, 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press, Cambridge.
- Council of Europe, 2018. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume With New Descriptors*. Council of Europe, Strasbourg.
- Cummins, J., 2017. Teaching for transfer in multilingual school contexts. In: García, O., Lin, A., May, S. (Eds.), *Bilingual Education: Encyclopedia of Language and Education*. Springer, Berlin, pp. 103–115.
- Cummins, J., Persad, R., 2014. Teaching through a multilingual lens: the evolution of EAL policy and practice in Canada. *Educ. Matter.* 2, 3–40.
- European Commission, 2007. *Final Report: High Level Group on Multilingualism*. European Communities, Luxembourg. Retrieved from: http://ec.europa.eu/education/policies/lang/doc/multireport_en.pdf.
- García, O., 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell, Chichester, UK.
- García, O., Li, W., 2014. *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan, New York, NY.
- Gorter, D., 2001. A Frisian update of reversing language shift. In: Fishman, J.A. (Ed.), *Can Threatened Languages Be Saved? Reversing Language Shift: A 21st Century Perspective*. Multilingual Matters, Clevedon, pp. 215–233.
- Gorter, D., 2005. Three languages of instruction in Fryslân. *Int. J. Sociol. Lang.* 171, 57–73.
- Gorter, D., Zenotz, V., Cenoz, J. (Eds.), 2014a. *Minority Languages and Multilingual Education*. Springer, Berlin.
- Gorter, D., Zenotz, V., Etxague, X., Cenoz, J., 2014b. Multilingualism and European minority languages: the case of Basque. In: Gorter, D., Zenotz, V., Cenoz, J. (Eds.), *Minority Languages and Multilingual Education: Bridging the Local and the Global*. Springer, Berlin, pp. 278–301.
- Grosjean, F., 2010. *Bilingual: Life and Reality*. Harvard University Press, Cambridge, Mass.
- Hall, J.K., 2019. The contributions of conversation analysis and interactional linguistics to a usage-based understanding of language: expanding the transdisciplinary framework. *Mod. Lang. J.* 103, 80–94.
- Jenkins, J., Cogo, A., Dewey, M., 2011. Review of developments in research into English as a lingua franca. *Lang. Teach.* 44, 281–315.
- Jessner, U., Cenoz, J., 2019. Teaching English as a third language. In: Gao, X. (Ed.), *Second Handbook of English Language Teaching*. Springer, Berlin, pp. 155–172.
- Jones, B., Lewis, G., 2014. Language arrangements within bilingual Education. In: Thomas, E.M., Mennen, I. (Eds.), *Advances in the Study of Bilingualism*. Multilingual Matters, Bristol, pp. 141–170.
- Kramsch, C., 2012. Authenticity and legitimacy in multilingual SLA. *Crit. Multiling. Stud.* 1, 107–128.
- Leonet, O., Cenoz, J., Gorter, D., 2020. Developing morphological awareness across languages: translanguaging pedagogies in third language acquisition. *Lang. Aware.* 29, 41–59.
- Lewis, G., Jones, B., Baker, C., 2012. Translanguaging: origins and development from school to street and beyond. *Educ. Res. Eval.* 18, 641–654.

- Lyster, R., Quiroga, J., Ballinger, S., 2013. The effects of biliteracy instruction on morphological awareness. *J. Immers. Content-Based Lang. Educ.* 1, 169–197.
- Mauranen, A.K., 2018. Second language acquisition, world Englishes, and English as a lingua franca (ELF). *World Englishes* 37, 106–119.
- Otheguy, R., García, O., Reid, W., 2015. Clarifying translanguaging and deconstructing named languages: a perspective from linguistics. *Appl. Ling. Rev.* 6, 281–307.
- Simpson Baird, A., Palacios, N., Kibler, A., 2016. The cognate and false cognate knowledge of emergent bi-literate Latino preschoolers. *Lang. Learn.* 66, 448–470.
- Yi, Y., Adamson, B., 2019. English in a Mongolian ethnic minority primary school. In: Liyanage, I., Walker, T. (Eds.), *Multilingual Education Yearbook 2019*. Springer, Cham, Switzerland.

Relevant website

Multilingual Education: Let's Make the Most of Multilingualism, <https://www.youtube.com/watch?v=vg8i2UHAf8s>.

Primary and secondary education: literacy and social justice

Nerida Spina^a and Barbara Comber^b, ^a Queensland University of Technology, Faculty of Education, Kelvin Grove, QLD, Australia; and ^b University of South Australia, Education Futures, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	195
Social justice in schools: teachers as researchers	195
Teacher-research: exploring social justice through critical literacy	196
Vivian Vasquez	196
Marianna Souto-Manning	196
Rebecca Rogers	197
Gerald Campano	198
Rob Simon	199
Amparo Clavijo	199
Jennifer Alford	199
Concluding remarks	200
References	200
Further reading	201
Relevant websites	201

Introduction

Literacy and language have long been implicated in oppression, privilege and the functioning of democracy. Take, for example, the following line from an English teachers' handbook produced in Victorian England, which suggested that the study of English would "promote sympathy and fellow feeling among all classes" (cited in [Eagleton, 1983](#), p. 20). Eagleton suggests that these efforts were ideological, in that they sought to impart a respect for middle-class values and achievements, to avert political action by working-classes, by building solidarity with middle-classes. The proscribed reading of texts which promote the interests of ruling classes has been problematized by those who instead to view education as a site for social justice. Educators continue to question taken-for-granted rhetoric and assumptions around literacy and schooling and, at the classroom level, many teachers focus on issues of justice and fairness through pedagogies which incorporate critical literacies – studies of textual practices which question the relations of language and power.

Social justice in schools: teachers as researchers

In this entry, we examine a range of practices that educators have designed and enacted as they work toward everyday justice goals in schools. We recognize that some of the most important work undertaken remains invisible to those beyond the classroom or local community, because it is not necessarily the subject of research or publication. We explore a selection of international studies where teacher-researchers have published, or worked alongside an academic collaborator to publish accounts of such practice. We highlight work from across different stages of schooling and from different places to examine what was attempted and why, and what was accomplished.

In mapping the terrain, we begin with an introduction to teacher-research, most notably that of Cochran-Smith and Lytle. Drawing on the work of Antonio Gramsci's writing in Italy leading up to World War I, [Cochran-Smith \(1991\)](#) argued some decades ago that action "is everyone's responsibility, and that each individual, no matter how apparently powerless [is] accountable for the role [they] play or fail to play in the larger political struggle" (p. 279). Using Gramsci's words as a call to address issues of social justice, Cochran-Smith inspired a renewed focus in which teachers are called to inquire into their own practice with justice and fairness in mind. This work also draws on the influences of Kurt Lewin's action research, in the 1940s, which encouraged ordinary people to inquire into their everyday lives.

Australian researcher Garth [Boomer \(1985\)](#) contrasted what he called "big R Research" - which he saw as a "postgraduate luxury," with "small-r research" that teachers could undertake in their own classrooms and schools. Rather than seeing issues of justice as something "out there," teacher inquiry entails teachers adopting an inquiry stance ([Cochran-Smith and Lytle, 2001](#)). Cochran-Smith and Lytle argue that teachers should not only see their work as related to high quality teaching and learning, but also understand that teaching must be part of broader movements for social change and justice. Adopting such a stance in the teaching of literacy requires teachers to be curious, to find moments of dissonance, and to explore both big-P and little-p politics with students as a means of addressing social injustice.

We (the authors) have acted as critical friends and co-researchers with teacher-researchers as they grapple with issues such as racial and gender based oppression, and the effects of poverty. Drawing from critical, Frierian, feminist and poststructuralist theoretical perspectives, Australian teacher-researchers in the eighties were among the first to investigate ways of working with theory in classroom curriculum and pedagogy (Comber, 2018; Luke, 2018). Since that time, many activist teacher/researchers and academics have continued to work together using critical literacy approaches to talk back to existing power structures, and to empower disenfranchized communities and students. Here, we note that as authors, we have both lived and worked in Australia, which is predominately an anglophone country, despite the diverse languages spoken by Indigenous peoples prior to and since colonization. Many of the scholars whose works are explored in this entry work in colonized countries, with historical hegemony and dynamics in relation to English language usage. We hope, however, that the tools and discussion in this chapter provide a framework for further problematizing literacy, including the hegemonic use of English in many educational settings. The section that follows provides some examples of writing that has emanated from internationally located research that considers the literate experiences of diverse and often disenfranchized students, and, is active in engaging teachers into the process of inquiry.

Teacher-research: exploring social justice through critical literacy

Vivian Vasquez

A pioneer of critical literacy approaches in early childhood settings in Canada, Vivian Vasquez documented her work with young children about their inquiries into everyday schooling practices which they saw as unfair. Initially influenced by the critical literacy approach of Australian teacher-researcher, Jennifer O'Brien, in fostering children's questions about the gendered nature of textual representations in books designed for young children, Vasquez (2007) developed her own distinctive pedagogical repertoire, which extended beyond texts to wider cultural and institutional practices.

Vasquez defines critical literacy as a curriculum that "arises from the social and political conditions that unfold in communities in which we live" (2007). Challenging notions that young children are not interested or capable of thinking about diversity, justice and power, her work as an activist scholar has shown how critical literacy can be lived in classrooms, so as to embed critical perspectives into conversations with young children as they learn to think about the world around them. In other words, rather than retaining a focus on critical text analysis, she has encouraged educators to engage young children in substantive discussions that allow them to see how they are positioned in everyday practices, and how both language and imagery positions readers in very particular ways. Drawing on the work of Comber et al. (2007), she has promoted critical literacy practices that extend conceptualizations of literate practices, for example by podcasting as a means of encouraging students to conceptualize texts more broadly than books (e.g., Vasquez, 2010).

Vasquez has used this approach of engaging children in substantive conversation about issues of race, class and gender to make power-relations and socio-political systems visible. As she writes, "Students who engage in critical literacy from a young age are likely going to be better able to contribute to a more equitably and socially just world by being better able to make informed decisions regarding such issues as power and control" (2016, p. 4).

Drawing attention to critical literacies as a means of revealing the exercise of power, Vasquez is also known for her work in popular culture and texts. In thinking about how teachers of young children might make use of popular texts in critical literacy lessons and conversations, Vasquez (2016) points out the importance of considering "how [we can] strike a balance between creating spaces or opportunities for childhood popular culture discourses and everyday texts in the classroom and engage children in working critically with these texts while at the same time making sure not to co-opt their interests. (p. 116). By "co-opting," Vasquez is drawing attention to how critical literacy conversations can draw on children's passions, and, create spaces for substantive conversations that have relevance to their lives, without making these texts "any less desirable" by turning them into schoolwork.

In her seminal paper "What Pokemon Can Teach Us about Learning and Literacy" (Vasquez and Smith, 2003), she considers how young children might be engaged in critical literacies that are both powerful and pleasurable. To analyze the usefulness of popular culture texts in early years education, she explains how her young nephew has engaged in the discourses and complex literate practices required to engage with the Pokemon series of texts. In describing these practices and literacies, the paper highlights the investment of time and energy required by teachers to break the code of a popular culture text, so as to be in a position to talk about, and, engage with students' interests.

Marianna Souto-Manning

In part, the nature of acquiring literacy skills has meant that scholarship in the early years of schooling has focused on decoding and coding print media through teaching spelling conventions, sight words, phonics and so on. Building on the work of Freire, Souto-Manning's scholarship offers practical examples of how literacy can be positioned as a tool for redressing injustice, by acknowledging the cultural and social realities of children and their families. Souto-Manning, an American scholar, has based much of her work documents the inquiries of teachers (including her own inquiries), which often reveal the disjunctures between the

home-cultures of students and socially-dominant cultures of school. As she points out in her work, these early schooling experiences are frequently the first time that children are exposed to dominant cultural discourses and practices, although they are typically expected to embody these values and ways of being.

Souto-Manning has worked with teacher-researchers to adopt a “humble stance” that “[blurs] the boundaries between teacher and learner” (Souto-Manning and Mitchell, 2010, p. 272). For instance, one teacher inquiry (Souto-Manning and Mitchell, 2010) documented the reflections of a teacher-researcher who had sought to include the diverse children in her class by talking about the countries Santa might visit at Christmas as a means of introducing discussion of an array of holidays such as Kwanzaa, Hanukkah, and Chinese New Year. However, the teacher’s reflections showed that children’s perspectives were still largely absent, and that by inviting a parent with insider-knowledge to participate in classroom activities, children were able to learn about each other, and the diversity of traditions, stories and practices that exist beyond the dominant social culture in which the school was embedded. Pursuing genuine participation, and building agency among students, families and communities through literacy is a way of working toward both recognitive and representative justice (Fraser, 2008).

Souto-Manning and Yoon (2018)’s work is an example of how justice has been taken up through literacy inquiry, in the early years of schooling. Drawing on Freire’s emancipatory philosophies, the authors produced *Rethinking Early Literacies: Reading and Rewriting Worlds*, in which they argue that early childhood settings are often where minorities are disenfranchised through practices that fail to attend to power relations, with minority families being blamed for failing to attend to school readiness, or to ensure children coming to school having acquired basic literacy skills. Instead, they argue, reading (and literacy) can be particularly alienating in early years settings, when the dominant focus of literacy instruction is on letter and sound recognition. Disrupting typical education discourses that use standardized assessment in a dominant language (typically English), this approach works toward justice by refusing to accept multilingual students as deficit, because of their proficiency in the dominant language. Souto-Manning’s use of Freire as a decolonizing approach illustrates how even very young children can engage in literacy in ways that value their own languages and circumstances.

This approach of inquiring into issues of justice and diversity has been explored in other ways by Souto-Manning, for example in how LGBTQ+ issues are concealed by dominant literacy practices in schools. Her work with elementary teacher Jill Hermann-Wilmarth (2008) contested a video that suggested diverse sexualities were not relevant in elementary schooling. Reflecting on her experiences as a pregnant, lesbian elementary school teacher, Hermann-Wilmarth drew attention to the ways in which gay and lesbian issues can be silenced through dominant heteronormative discourses and practices in schools. Through critical inquiry using a queer theoretical lens, they proposed practical ways that teachers might become more inclusive of all students in their literacy programs. Writing of her early experiences teaching with a child with lesbian parents, Souto-Manning wrote in her journal that, “I don’t really know what to do It is not fair for me to remain silent. I have to speak up, but how?” (p. 269). She ultimately talked with the child’s parents, reflected on possible approaches (such as reading books to her class that had diverse representations of families) as part of her inquiry into her own practice. Seeing herself as a deliberate questioner, she investigated how her silence had been a form of complicity, and used the opportunity to learn from her students and their families. Sharing stories about her own gay uncles, she expanded the selection of texts available in her classroom, and used books that included diverse family representations as conversation starters.

Souto-Manning’s approach of “kidwatching” (2009) encapsulates how teachers can become curious in learning more about their students and communities, with justice in mind. She gives an example of her work with a student, George, who came to her elementary class having been described a “behavior problem” and who was notorious among first-grade teachers. Rather than accepting this view of George, Souto-Manning inquired into her own practice, troubling the notion of African-American Vernacular English as problematic form of communication in the classroom, and instead seeing how teachers’ insistence on Standard American English created a situation in which teachers and peers chastised and corrected George each time he spoke, leading to frustration, and ultimately, his exclusion from class, and from learning. Through inquiring into her practice (and that of her colleagues), Souto-Manning came to realize that linguistic patterns in Standard American English such as asking rhetorical questions such as, “would you like to sit down?” were problematic. George understood these kinds of rhetorical questions as questions, rather than commands or directions, leading many teachers to think of him as disobedient. Listening *to* George, rather than listening to others talk *about* George, Souto-Manning noticed his eagerness to interact, inquire and participate in class discussions. Other work by Souto-Manning and colleagues has focused on engaging young children in discussions about inequities in the nature of employment and work, race, diverse sexualities, as well as examining the intersectionalities between many of these issues.

Rebecca Rogers

In her presidential address to the Literacy Research Association, American scholar Rebecca Rogers (2018) noted that “literacy research has an entangled history with racial justice” (p. 24) because it has been part of systems that have subjugated and marginalized peoples, restricting access to education, health care, housing and other privileges. For this reason, she argued that literacy research must continue to examine how literacy shapes the flow of knowledge and power; and take care to avoid further reproducing unequal and unjust power relations. Rogers’ used critical race theory to compare her own family history with that of one of her research participants, a Black woman, June and her family: she found that while her family benefitted from access to

education and property, June's family were excluded from education by literacy assessments that "became a proxy for discrimination" after Jim Crow laws were abolished (p. 29). Rogers' work (e.g., 2011) with June and her family provides moving accounts of the subtle practices in schooling (such as being tracked into special education programs) can operate to perpetuate racial injustice.

Rogers has published extensively, with much of her work using critical discourse analysis, to provide practical examples of how teacher-researchers might better understand diversity and use critical literacies to further social justice agendas. Acknowledging the micro-politics and action that form everyday life is important because it acknowledges the subtle ways that power can be exerted and challenged in classrooms, playgrounds and beyond. In her work with Labadie and Wetzel (Labadie et al., 2012), she shows how a teachers introduction to a book can become a time where readers are encouraged to disrupt common-sense assumptions embedded in texts; to consider diverse perspectives; and to tackle issues of power and justice. While the texts that teachers select in their classrooms are important, this work highlights the importance of how texts are taken up in classrooms, for example by allowing time for students to ponder unequal relations and silences in texts. Examining the micro-practices of teaching reveals how moments such as lesson introductions that are often overlooked, or seen as a moment for a teacher to impart their own lesson structure, can become a time for sparking critical thinking about issues of justice and diversity. This work shows how activist teacher-researchers can use small moments in lessons to create "an ongoing discussion and reflection about a book instead of only an opening" (p. 127) in ways that recognize diversity and explore locally relevant questions. This attention to everyday doings in classrooms, including how texts are taken up in lessons, requires a stance in which there is a consistent exploration of issues of privilege, diversity, power and justice.

Gerald Campano

Gerald Campano's work in critical literacy centers around the way immigration can affect the educational experience of students. In the field of critical literacy, Campano's work has shown the value incorporating the experiences of immigrant students into curriculum choices. Campano, who teaches in the United States, describes himself as a proud third generation Pinoy with ancestral roots in the Mindanao region of the Philippines. His work shows the value of teachers who are intentional in encouraging migrant and refugee students to share their narratives, and to feel that their personal cultural identities are represented and valued both in school and in the wider community.

Campano has drawn heavily on the scholarship of Marilyn Cochran-Smith and Susan Lytle, describing Lytle as his mentor and the center of his work (Campano, 2019). Their inquiry-based approach has therefore been important in Campano's approach toward agentic, participatory inquiry, as he works alongside students and community members. Campano therefore writes from his own experience, as well as the lived experiences of the immigrant students he interviews. By making visible the embodied knowledges and experiences of immigrant communities, his work in critical literacy is based on the view that it is imperative that teachers and communities build a shared vision of justice and rights.

His work also encourages students to engage in critical inquiry themselves, questioning ideals they had previously taken for granted, particularly relating to the issue of race, class and immigration (Beach et al., 2015, p. 142). 'Immigrant Stories and Literacy', for instance, is a compilation of critical stories of the classroom life of immigrant students from Campano's fifth-grade class. In it, he reflects on his early question: "What would happen if I invited children from immigrant and migrant backgrounds to read, write, and speak from their own experiences and the realities of their lives?" (p. 31). A key argument in Campano's work is that it is possible for teachers to make space for children to draw on their cultural traditions, and to resist the dominant culture of standardization and quantification that has characterized much of literacy instruction. This approach recognizes the important contributions of students who might normally be marginalized by literacy instruction, and instead, to offer space to for students to generated sophisticated insights about issues of justice that are based on their own life experience.

Campano's work draws on foundational critical literacy theorization, such as that of New London Group. For example, in his paper with Low, "The image becomes the weapon" (Low and Campano, 2013) examines the privileging of canonical texts, before looking at how this is taken up in the classroom with students:

what we have learned... is the need to view all students as cosmopolitan intellectuals... who have the capacities to think transculturally and across genres, making any texts they may encounter potential grist for their creative mills. The canon is not solely a mechanism to be reproduced; it is also critically reinvented by our students (p. 31).

In his analysis of the kinds of texts that US schools use to explore the experiences of immigrants, he shows how the dominant views around the importance of an "English only" approach silences the experiences and views of multilingual students; and their hybrid language practices and experiences. His analysis (Ghiso and Campano, 2013) of the book *One Green Apple* in which a young South Asian girl talks about going to school "where I will learn to speak English," shows the conflation between schooling and learning the English language. Instead, he argues that teachers adopting critical literacy approaches must learn to listen to community members and students, which requires substantial time and "intellectual and empathetic labor" (p. 55). This important work shows how justice requires a collaborative approach between teachers, students and communities, in which teachers take account of local actualities, such as community experiences of multilingualism and multiculturalism.

Rob Simon

Rob Simon at the University of Toronto's Ontario Institute for Studies in Education has co-authored with Campano to examine critical literacy among adolescents. Partnering with teacher researchers, the pair published *Activist literacies: Teacher research as resistance to the 'normal curve'* (2013). This work examines the institutional constructions of "gifted and talented" and "special needs" in two schools from vastly different socio-economic and cultural regions in the US. Their use of critical disability scholarship frames their understandings of how activist literacy educators can (and must) contest dominant discourses, for example in understanding "normal," bell-curved constructions of student achievement. Conceptualizing disability as a social construction rather than a medical, individually embodied pathology provides a lens for teacher-researchers to resist assumptions about students, and to "take social location—their own and their students'—seriously" (p. 25). This view challenges dominant views of literacy that rely on deterministic understandings of young people as literate beings based on narrowly constructed literacy assessments. Instead, Simon and Campano's work demonstrates how teacher/researchers can assume that there is much more to learn about students, and work together to develop rich understandings of students' literate practices and needs. Drawing on the work of Cochran-Smith and Lytle (2015), they explain how this theorization "from the thick of things" (p. 22) helps teachers to understand students' everyday lives.

Adopting the view that teaching literacy should not be seen as an intervention performed by teachers, means looking for moments of dissonance that exist in everyday life to inquire into rich lives of diverse groups of students, to question existing power structures, and to reveal and contest the inequities that affect students and their communities. Simon (2012) has documented how pre-service teachers can engage in the practice of learning from and alongside students, with justice in mind. His account (Simon, 2012) of a pre-service teacher inquiring into the challenges faced by an adolescent student, Yusef, shows how teachers can become "connoisseurs" of students' complex identities and practices. In so doing, he reveals how adopting an inquiry stance can change how teachers see students, often moving from deficit views toward more complete understandings of the out-of-school, multiliterate practices of young people.

Amparo Clavijo

Amparo Clavijo is another scholar who has drawn on Comber's work in critical literacy in her explorations of community-based pedagogies, particularly in Colombian public schools. This body of work examines issues of power through having students consider local issues that affect their lives and neighborhoods.

Her most recent research focuses on Community Based Pedagogies and Literacies in Language Teacher Education. She has published articles and book chapters about Colombian public school teachers using community pedagogies with students to inquire about social and cultural issues that affect their neighborhoods. Her work draws teachers into critical literacy, through encouraging them to consider their own personal and professional journeys. For instance, in a chapter she co-authored with Ramírez Galindo (2018), the authors used self-study to consider how community pedagogies are enacted in socioeconomically and culturally diverse schools in Colombia. In particular, they considered how national language policies and practices in Colombia that draw on foreign curricula, textbooks and standards can impose social and cultural norms that do not include local knowledges and practices.

Through her journey from the oldest daughter growing up in a patriarchal culture, to her experiences as a Fulbright scholar and Colombian language teacher-educator, Clavijo reflects on how she brought together the theoretical traditions she has used in her approach to critical literacy. Her experiences are reflective of a Western-centric focus that has dominated critical literacy research and practice. For instance, she explains (2018, p. 43) that she was not introduced to Paulo Freire's work that promoted emancipatory pedagogies in rural Brazil until she attended a doctoral program in the United States. This reality—in which she suggests local scholarship in teaching and teacher-education must be strengthened—is part of Clavijo's critical literacy agenda. Her promotion of Freirian pedagogies has since been used to respect the languages and cultures of Indigenous children and communities. This is not to say that Clavijo's work is not underpinned by broader scholarship on critical literacy that meets the needs of disadvantaged and marginalized communities and students. For instance, she has drawn on Comber and Kamler's (2004) notion of "turnaround pedagogies" with teachers and pre-service teachers, to encourage an approach to literacy that is intentional in refusing to begin from a view of children and families as being in deficit.

Clavijo and colleagues developed their approaches using the theoretical work of a range of critical theorists from Butler to Gramsci, Foucault and the New London Group. Her attention to reflexivity, self-consciousness, and "reading the community critically" (Medina et al., 2015) demonstrates how such scholars are able to work toward justice in diverse settings, by using a diverse range of theoretical approaches to make sense of their worlds and that of their students, so that critical literacy is used to forward liberatory agendas.

Jennifer Alford

Jen Alford is an Australian teacher and scholar whose work in critical literacy focuses on the classroom practices of teachers as they embed critical literacy into their work with adolescent English language learners. With similarities to other critical literacy advocates such as Vasquez and Campano, Alford's work shows that English language learners in Anglophone countries and schools remain capable of engaging with texts, and adopting critical readings. Alford's work (e.g., 2022) is unique in that

it provides a detailed look at *secondary* school aged students who are learning an additional language. As she describes, the literature on these adolescent learners is sparse, despite the ‘assessment stakes [being] higher in pre-tertiary [years of schooling]’ (p. 17). For example, in the state of Queensland where her work is primarily conducted, externally administered and graded English examinations are worth 25% of every students’ grades in their final year of schooling. As Alford explains, high stakes assessments such as this have the potential to impact on teachers’ willingness to adopt critical literacy approaches. Her work draws attention to the hegemonic nature of standard English in schooling, particularly for adolescent learners. As she points out, it is rare for Australian teachers to draw on learners’ range of linguistic resources, instead insisting on English as the dominant mode of communication. Her work shows how critical literacy can be used to disrupt this practice by validating all languages in the classroom (Alford, 2019). She shows how the explicit teaching of language structures as part of additional language instruction provides an opportunity to problematize and contest linguistic choices. She also argues for the importance of pedagogical practices such as the use of translanguaging, which integrates learners’ knowledge of multiple languages, allowing them to move between languages to maximize communication: ‘It is also a pedagogical approach teachers can adopt that makes full use of learners’ language repertoires. Learners can research topics in their own languages, discuss texts and issues in their own languages in classrooms, and use their own languages in the production of texts’ (p. 180). Her work demonstrates that critical literacy in classes with adolescent learners is especially significant, given that they are often at the forefront of political change. The past decade has been characterized by numerous movements such as those seeking reform on climate change, domestic violence, gun control and policing reform have included activism by young people. As Alford (2022) writes, ‘these events are timely reminders of the keen interest adolescents have in ideological matters of concern to their own futures. Critical literacy provides a way for them to engage with these matters and find much-needed solutions. In a post-truth world where much uncertainty prevails, critical literacy is a basic for all young people learning to be literate’ (182).

Concluding remarks

The current period is marked by political upheaval in countries around the globe, a proliferation of mis/disinformation promulgated on social media platforms, and an increasing rejection of rising inequality and injustice. As individual countries and societies deal with their own inequality and political upheavals, we cannot minimize global inequality and inequity between countries and regions. We note that we have situated this discussion from our own context of Australia, an Anglophone country with a westernized mass education system, and experiencing similar political tussles to other western countries. While our conversations are grounded in this context, we hope to have provided a sense of what is possible when teachers, students and families leverage critical literacy to disrupt dominant power structures. We argue that critical literacy has been a vital tool in equipping young people to question the taken-for-granted truths that circulate and maintain oppressive power structures. Amidst a global pandemic and an increasing frustration over continued injustices perpetrated on Black and LGBTIQ+ lives, critical literacy is continuing to play a part in how educators help students to engage with the world around them. The examples of scholarship provided in this entry are by no means definitive, but rather have been used to demonstrate the depth of work being carried out by teachers and researchers with the view of adopting a stance in their approach toward literacy that is intentional in tackling issues of justice. While new public management practices have entered literacy education with the aim of bolstering and competition using literacy data collected through high-stakes standardized tests, the teachers and researchers who engage in critical inquiry have shown how these goals can be displaced by the use of range of data to inquire into issues of power, oppression and justice; and to signal a focus on democratic traditions and goals.

References

- Alford, J., 2022. Critical Literacy With Adolescent English Language Learners: Exploring Policy and Practice in Global Contexts. Taylor & Francis Group, New York.
- Alford, J., 2019. Teaching critical literacy with EAL/D learners: making the most of diverse languages. In: Henderson, R. (Ed.), *Teaching Literacies—Pedagogies and Diversity*, second ed. Oxford University Press, Oxford, pp. 315–338.
- Beach, R., Campano, G., Borgmann, M., Edmiston, B., 2015. *Literacy Tools in the Classroom: Teaching Through Critical Inquiry, Grades 5-12*. Teachers College Press, New York.
- Boomer, G., 1985. *Fair Dinkum Teaching and Learning: Reflections on Literacy and Power*. Boynton/Cook Publishers Inc, Upper Montclair, NJ.
- Campano, G., 2019. *Immigrant Students and Literacy: Reading, Writing, and Remembering*. Teachers College Press, New York.
- Clavijo, A., Ramírez Galindo, M., 2018. Using self-study to examine our research and teaching practices as EFL teacher educators in Colombia. In: Sharkey, J., Peercy, M.M. (Eds.), *Self-Study of Language and Literacy Teacher Education Practices: Culturally and Linguistically Diverse Contexts*. Emerald Group Publishing, Bingley, pp. 37–53.
- Cochran-Smith, M., 1991. Learning to teach against the grain. *Harv. Educ. Rev.* 61, 279–311.
- Cochran-Smith, M., Lytle, S.L., 2001. Beyond certainty: taking an inquiry stance. In: Lieberman, A., Miller, L. (Eds.), *Teachers Caught in the Action: Professional Development that Matters*. Teachers College Press, New York, pp. 45–58.
- Cochran-Smith, M., Lytle, S.L., 2015. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York.
- Comber, B., Kamler, B., 2004. Getting out of deficit: pedagogies of reconnection. *Teach. Educ.* 15, 293–310.
- Comber, B., Nixon, H., Reid, J. (Eds.), 2007. *Literacies in Place: Teaching Environmental Communications*. Primary English Teachers Association, Newtown, NSW.
- Eagleton, T., 1983. *Literary Theory: An Introduction*. John Wiley & Sons, Hoboken.
- Fraser, N., 2008. *Scales of Justice: Reimagining Political Space in a Globalizing World*. Columbia University Press, Boston.
- Ghiso, M.P., Campano, G., 2013. Ideologies of language and identity in U.S. children’s literature. *Bookbird* 51, 47–55.
- Labadie, M., Wetzel, M.M., Rogers, R., 2012. Opening spaces for critical literacy: introducing books to young readers. *Read. Teach.* 66, 117–127.
- Low, D.E., Campano, G., 2013. The image becomes the weapon: new literacies and canonical legacies. *Voices Middle* 21, 26–31.

- Luke, A., 2018. *Critical Literacy, Schooling, and Social Justice*. Routledge.
- Medina, R.A., Ramirez, L.M., Clavijo, A., 2015. Reading the community critically in the digital age: a multiliteracies approach. *ISLS Read. Lang. Stud.* 5, 45–66.
- Rogers, R., 2011. The sounds of silence in educational tracking: a longitudinal, ethnographic case study. *Crit. Discourse Stud.* 8, 239–252.
- Rogers, R., 2018. Literacy research, racial consciousness, and equitable flows of knowledge. *Literacy Res.* 67, 24–43.
- Simon, R., 2012. "Without comic books, there would be no me": teachers as connoisseurs of adolescents' literate lives. *J. Adolesc. Adult Literacy* 55, 516–526.
- Simon, R., Campano, G., 2013. Activist literacies: teacher research as resistance to the "normal curve." *J. Lang. Literacy Educ.* 9, 21–39.
- Souto-Manning, M., Mitchell, C.H., 2010. The role of action research in fostering culturally-responsive practices in a preschool classroom. *Early Child. Educ. J.* 37, 269–277.
- Souto-Manning, M., Yoon, H.S., 2018. *Rethinking Early Literacies: Reading and Rewriting Worlds*. Routledge, New York.
- Vasquez, V., 2007. Using the everyday to engage in critical literacy with young children. *N. Engl. Read. Assoc. J.* 43, 6–11.
- Vasquez, V., 2010. *Getting Beyond "I Like the Book": Creating Space for Critical Literacy in K-6 Classrooms*, second ed. International Reading Association, Washington, DC.
- Vasquez, V., Smith, K., 2003. What Pokemon can teach us about learning and literacy. *Lang. Arts* 81, 118–125.
- Vasquez, V.M., 2016. *Critical Literacy Across the K-6 Curriculum*. Taylor & Francis, New York.

Further reading

- Bentley, D., Souto-Manning, M., 2016. Toward inclusive understandings of marriage in an early childhood classroom: negotiating (un)readiness, community, and vulnerability through a critical reading of King and King. *Early Years* 36, 195–206.
- Blackburn, M., 2010. *Acting Out: Combating Homophobia through Teacher Activism*. Teachers College Press, New York.
- (Illustrator) Bunting, E., Lewin, T., 2006. *One Green Apple*. Houghton Mifflin, New York.
- Comber, B., 2013. Schools as meeting places: critical and inclusive literacies in changing local environments. *Lang. Arts* 90, 361–371.
- Fitzgerald, F.S., 1925. *The Great Gatsby*. Charles Scribner's Sons, New York.
- Gee, J.P., 2003. *What Video Games Have to Teach Us About Learning and Literacy*. Macmillan, New York.
- Hayik, R., 2015. My critical literacy journey in a Middle Eastern EFL classroom: insights and challenges. In: Yoon, B., Sharif, R. (Eds.), *Critical Literacy Practice*. Springer, Singapore, pp. 95–109.
- Janks, H., 2005. Deconstruction and reconstruction: diversity as a productive resource. *Discourse* 26, 31–43. <https://doi.org/10.1080/01596300500040078>.
- Janks, H., 2013. Critical literacy in teaching and research. *Educ. Inquiry* 4, 225–242. <https://doi.org/10.3402/edui.v4i2.22071>.
- Kirkland, D., 2013. *A Search Past Silence: The Literacy of Young Black Men*. Teachers College Press, New York.
- Rogers, R., Wetzel, M.M., 2013. *Designing Critical Literacy Education Through Critical Discourse Analysis: Pedagogical and Research Tools for Teacher-Researchers*. Routledge, New York.
- Rogers, R., Mosley Wetzel, M., O'Daniels, K., 2016. Learning to teach, learning to act: becoming a critical literacy teacher. *Pedagogies* 11, 292–310.
- Sharkey, J., Clavijo Olarte, A., Ramirez, L.M., 2016. Developing a deeper understanding of community-based pedagogies with teachers: learning with and from teachers in Colombia. *J. Teach. Educ.* 67, 306–319.
- Souto-Manning, M., Cheruvu, R., 2016. Challenging and appropriating discourses of power: listening to and learning from early career early childhood teachers of color. *Equity Excell. Educ.* 49, 9–26.
- Souto-Manning, M., Hermann-Wilmarth, J., 2008. Teacher inquiries into gay and lesbian families in early childhood classrooms. *J. Early Child. Res.* 6, 263–280. <https://doi.org/10.1177/1476718X08094450>.
- Souto-Manning, M., Martell, J., 2017. Committing to culturally relevant literacy teaching as an everyday practice: it's critical! *Lang. Arts* 94, 252–256.
- Souto-Manning, M., Rabadi-Raol, A., 2018. (Re)centering quality in early childhood education: toward intersectional justice for minoritized children. *Rev. Res. Educ.* 42, 203–225.
- Souto-Manning, M., 2009. Acting out and talking back: negotiating discourses in American early educational settings, 179, 1083–1094.
- Simon, R., Hicks, B.L., Walkland, T., Gallagher, B., Evis, S., Baer, P., 2019. Breaking gender expectations: adolescents' critical rewriting of a trans young-adult novel. In: Guzzetti, B., Dunkerly-Bean, J., Bean, T. (Eds.), *Literacies, Sexualities, and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, pp. 141–152.
- The New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66, 60–93.
- Vasquez, V.M., Janks, H., Comber, B., 2019. Critical literacy as a way of being and doing. *Lang. Arts* 96, 300–311.

Relevant websites

- <https://podcasts.apple.com/us/podcast/critical-literacy-in-practice-clip-podcast/id1059193459>, Podcast of Dr Vivien Vasquez.
- <https://ncte.org/blog/2019/07/critical-literacy/>, National Council of Teachers of English: Key Aspects of Critical Literacy.
- <http://addressinginjustices.com/gender-is-like-an-ocean>, Project that Documents the Addressing Injustices Project Collaboration Between Youth, Educators and Academics (Including Rob Simon).
- <https://thelearningexchange.ca/videos/jerome-harste-and-vivian-vasquez-critical-literacy/>, Interview With Jerome Harste and Vivian Vasquez: Critical Literacy.

Ecocritical pedagogies and curriculum

John Lupinacci^a, Brandon Edwards-Schuth^a, Alison Happel-Parkins^b, and Rita Turner^c, ^a Washington State University, Pullman, WA, United States; ^b University of Memphis, Memphis, TN, United States; and ^c Independent Scholar, Baltimore, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	202
Introducing an ecocritical framework	203
Anthropocentrism as it contributes to human-supremacy: teaching in support of an Earth Democracy	206
Ecocritical curriculum and pedagogies: teaching for a just and sustainable future	207
References	209
Further reading	209

Glossary

Power/Knowledge Knowledge is not neutral nor universal, it is instead understood as historically specific and always implicated in questions of social power—Power and knowledge are mutually constitutive (Barker and Jane, 2016; Foucault, 1980)

Ecocritical Drawing on a similar definition as ecocriticism—which is the study of literature and the environment from an interdisciplinary point of view; ecocritical refers to interdisciplinary work that: is critical of Western industrial culture (specifically Modernity) and its impacts on social and environmental systems; examines the relationships between language, culture, and value-hierarchies prevalent in Western industrial culture; and working actively to identify how to teach or share skills and habits of mind that support social justice and sustainability

Ecojustice Education Educational efforts that recognize that both ecological degradation and social justice issues have deep cultural roots in Western industrial colonial cultures, and that educators can—and ought to—center the health of local and global ecosystems as essential to all life, while recognizing the cultural assumptions that undergird socially constructed value-hierarchies that reproduce racism, sexism, classism, ableism, speciesism, etc. (Martusewicz et al., 2021)

Equity Referring to fairness and justice and important to differentiate from equality. Equity frameworks acknowledge diversity of needs

Individualism The deep belief that individual human beings are independent autonomous beings and that individuals in pursuit of their own self-interests leads to a greater society. This deep cultural belief falsely positions competition as human nature. It is important to differentiate individualism from the important concept of individuality which is the recognition that everyone has unique differences in how they identify and contribute in society

Anthropocentrism A logic structure of thinking that situates, or positions, human-beings as the center or top of a socially constructed value-hierarchy of all things. This deep belief becomes the basis for human-supremacy and is inextricable from the logics of racism, sexism, classism, ableism, etc.

Interdependence The idea that all beings are related and have mutual responsibilities and contributions to one another. In other words, we are all dependent on one another and the diversity of life that contributes to a balance in the ecosystems within which we all live. This concept can be understood through the Buddhist concept of dependent

Earth Democracy A decision-making system that empowers local communities to value and acknowledge diversity together with an intrinsic worth of all species in all decisions. All decisions are made by and with consideration for all that is impacted by the decision being made and in protection of the basic foundations of life for the present as well as for future generations (Shiva, 2015)

Bioregion A locale defined by characteristics of an ecosystem rather than by man-made divisions or state borders

Introduction

Considering the difficult conditions for diversity and life on the planet due to climate change (IPCC, 2022)—and the connected consequences including, but not limited to: poverty, famine, pandemics (Shah, 2020) and increased violent conflict as a result of devastating inequity—educators as leaders are presented with the challenge of rethinking education and doing so with close attention to what can be done differently in curriculum and pedagogy. In this era, one in which global climate change is threatening the very existence of human and more-than-human communities, as ecocritical teacher educators and curriculum scholar-activists we turn our attention toward the possibilities of radically reconsidering the purpose, and the associated procedures and practices, of not only teacher education but also of PreK-12 schooling.

Critical scholarship in education over the past few decades has a strong tradition of calling into question the relationship between knowledge, power, and curriculum. Building on critiques of how neoliberal policies support and promote particular kinds of “knowledge” (rooted in capitalism, patriarchy, and white cis hetero-human supremacy) and regimes of truth as a core foundation of curriculum content and pedagogy, we are interested specifically in how these regimes of truth are learned, and hopefully unlearned, via disrupting dominant curriculum through the employment of an ecocritical approach in education. While there are a handful of important radical frameworks in critical education that contribute to, and fall under, what we define in this encyclopedia entry as *ecocritical*, we want to be clear that the ecocritical framework we share here—and the associated pedagogical practices and recommendations—draws on a variety of radical traditions in educational scholarship that offer up possibilities of critically addressing the cultural roots of social justice and sustainability in classrooms, schools, neighborhoods, and communities.

Responding to the systemic violence and exploitation perpetuated by the current dominant social, economic, and environmental contexts in North America (similar to other nation states embedded in Western industrial culture), ecocritical educators examine and address how it is that schools in Western industrial societies create, support, and sustain the habits of mind that rationalize, justify, and (re)produce unjust social suffering and devastating degrees of environmental degradation inseparable from a devaluing and loss of human rights. Ecocritical work, like many other critical frameworks, considers how power and knowledge (power/knowledge) are intimately and inextricably connected (Foucault, 1980), emerging from specific power relations to form one of several truths. Given this, Truth (emphasis of capital T) does not exist, but instead truth is a cultural “good” describing the world with practical consequences (Barker and Jane, 2016; Rorty, 2009)—an epistemological product of specific group(s) interests always immersed within (and influencing) the social context. Adding to the critical tradition to examine, and trouble the Eurocentric notion of one Truth, in favor of a diversity of cultural truths that contributes to a dynamic knowledge system emerging from democratic and scientific consensus; an ecocritical framework places specific attention on the problematic notion that humans exist as separate from and superior to not only one another but also from the ecosystems the planet’s capacity to support human life as part of diversity of species.

When faced with such a challenge, an ecocritical framework can be understood as a line of inquiry through which scholar-activist educators ask the questions: *How is it that exploitation is rationalized, justified, and/or (re)produced by teachers, the curriculum, and schools in general?* Furthermore, scholar-activist educators working through such inquiries are also committed to turning the critical lens inward and asking: *What can be done in order to teach in support of alternatives to Western industrial culture?* In an attempt to address these questions, we introduce what we are referring to as an ecocritical framework for curriculum and propose engaging in pedagogical practices as a way to explore the possibilities and potentialities for diverse critical ecological perspectives in education.

An ecocritical framework addresses how education is influenced by systems of exploitation and violence, systems which rely on a refusal to acknowledge and embrace mutuality and interdependence (Lupinacci and Happel-Parkins, 2015, 2016). Ecocritical scholars in education use diverse critical lenses for addressing and rethinking dominant cultural frameworks, but certain principles remain at the center of the work. Specifically rooted in and simultaneously pushing the boundaries of the critical tradition, teacher educators positioned within the ecocritical movement recognize that social and environmental justice are inseparable and inextricably linked, and that these injustices rely on the perpetuation of cultural habits—like domination, individualism, and consumerism. They also acknowledge that such habits are deeply rooted in value-hierarchized, often dualistic, social thought—dualisms like culture/nature, male/female, mind/body, and reason/emotion—that influence collective attitudes and unexamined behaviors. To dismantle such injustices, then, these scholar-educators analyze the culturally-constituted value hierarchies our society is reproducing. This approach also includes exploring alternative knowledges and ways of recognizing and understanding differences that move beyond the limitations of Eurocentric (or Western industrial) thought and the complex tensions, double-binds, and even contradictions that exist within our modern cultures and so many of the short-term solutions. For example, we often find ourselves consuming goods in an effort to improve our lives and even support “green” practices, but our relationship to these goods, to the people who create them, and to the beings who supply the materials, remains invisible and unquestioned. Our methods of consumption and participation in society are deeply influenced by our culture’s hierarchies of value. As such, scholar-activist educators ask: *How are our social practices and relationships dependent on privilege and exploitation, and in what ways might these practices and relationships be restructured to be more inclusive without exploitation and exclusion?*

Introducing an ecocritical framework

An ecocritical framework strives to organize scholar-activist educators to learn from, connect with, and engage in identifying and critically examining the role that education both plays, and ought to play, in transitioning toward supporting diverse, socially just, and sustainable communities. Drawing from EcoJustice Education (Martusewicz et al., 2021), and stemming from the growing field of ecocritical work in social and cultural foundations of education, we summarize the movement’s aims in three aspects. Ecocritical perspectives help educators to critically and ethically (1) examine Western industrial culture, in relationship to the global spread of hyper-consumerism and its impacts on social and environmental systems; (2) examine value-hierarchized dualisms, and the associated behaviors and mechanisms that contribute to inequities such as racism, classism, sexism, disableism, anthropocentrism, etc. in connection with degradation of the environment; (3) examine and identify how to teach or share skills, and habits of mind, that support socially just and environmentally sustainable communities (Lupinacci and Happel-Parkins, 2016). Furthermore, the ecocritical framework we share here is informed not only by an EcoJustice Education perspective but also from EcoFeminism (Plumwood, 1993, 2002), Ecopedagogy (Kahn, 2010), Ecocritical Pedagogies (Lupinacci et al., 2019), Critical Place-based Ed

(Gruenewald, 2003; Gruenewald and Smith, 2014), and Indigenous and Land Pedagogies (Basso, 1996; McCoy et al., 2017; Tuck et al., 2014).

Simply put, through an ecocritical framework, educators work to support learners in recognizing two conflicting and foundationally different worldviews—on the one hand, ecological worldviews of diversity, interdependence and interspecies equity (see Fig. 1); and on the other, a human-centered worldview informed by human-supremacist hierarchal capitalist, racist, and patriarchal relationships (see Fig. 2).

It is important for educators to simultaneously learn to recognize how diverse worldviews rooted in interdependence and multispecies exist and have much to offer while also learning to recognize how the power/knowledge relationship works in a capitalist, patriarchal, white cis-hetero human supremacist cultural worldview, and how this particular worldview shapes knowledge and epistemological understandings of life and relationships. We assert that it is important within ecocritical efforts to engage in pedagogical work to disrupt the latter framework through examining how such a worldview (as illustrated in Fig. 2 below) might be reconstituted—via education or what we refer to later in this entry as *ecocritical pedagogies*—in ways that are local and in support of living systems.

Working through and within an ecocritical framework often include students and teachers examining how knowledge systems and the associated behaviors and mechanisms of human domination—in relationship to language, culture, and power—are culturally mediated, and how PreK-12 educators can play a role in reconstituting those understandings, behaviors, and mechanisms. Furthermore, with(in) such a framework, it is important to recognize that the role of teacher can be taken on by more-than-human members of any learning community. By highlighting how many dominant cultural belief systems, root assumptions, and narratives currently destroying the planet and prospects of peace are constructed and not simply “natural,” educators can help students develop critical perspectives on these root beliefs. This undertaking also allows for the exploration of alternative belief systems and metaphors that facilitate different kinds of relationships with other people, other beings, and the land.

An ecocritical framework also illuminates the systematic economic and political restructuring of lives that, while it might address immediate needs of human or environmental systems, ultimately often perpetuates unjust social suffering and extreme environmental degradation. Rethinking education through an ecocritical framework require moving beyond a simple critique of technological development that portrays technology, or humans, as the problem or the solution. For example, on one hand being critical of the dominant systematic economic and political restructuring of lives in current hyper-consumerist Eurocentric cultures includes also recognizing how technologies like pace-makers and vaccines have been, and probably ought to be, valued contributions to addressing human suffering. However, ecocritical work requires that teachers and students also ask how these valued technologies may or may not be sustainable and equitable via renewable resources and fair labor. Ultimately, addressing unjust suffering in society is important, but it must also include an awareness and analysis of the impact of any proposed solution on several generations of both human and more-than-human communities. What has become commonplace over the past century,



Fig. 1 Multispecies equity worldview: diversity, interdependence, & equity.



Fig. 2 Anthropocentric cultural worldview: individualism & human-supremacy.

and extending into the current, is the intentional restructuring of relationships to control and commodify lives in order to maintain and manufacture markets. For example, food and water are life-sustaining elements necessary for supporting healthy communities. However, the relationships to these “resources” have been enclosed—that is, they have been monetized or understood as commodities to be earned and purchased. This iteration of capitalism as a “supply and demand” economic system predicated on exploitation works to enclose living systems and can be understood as the globalizing force to commodify and privatize that which is common and public. Critiquing this kind of enclosure means reimagining the sorts of relationships humans can have with the more-than-human world and with each other, rather than simply trying to make human behavior “less bad” by slightly reigning in current destructive practices.

In short, an ecocritical framework calls for pedagogical efforts that center student learning on recognizing the importance of examining intellectual, environmental, and cultural practices and traditions in regard to how they either support or undermine living systems together with whatever content is being taught. Whether examining discursive practices or economic structures while learning mathematics, language arts, science, or social studies, a key feature of ecocritical pedagogies is the recognition that human knowledge systems are culturally constructed, have consequences for all living beings, and can be re-imagined in transformative ways (Turner, 2015). Another distinguishing aspect of ecocritical pedagogies is that whatever the lesson or activity, students and teachers together address the powerful role that their culture plays in the development of themselves, their values, and their diverse relationships. Such a framework examines, explores, and proposes diverse and collaborative pedagogical projects that respond to current dominant belief systems, and works to ensure that any responses are necessarily collaborative with diverse cultures in ways that are local, situational, and in support of decentralized living systems. In the classroom, this might include examining clips from news media to determine what root metaphors are at work in the way our culture communicates about the more-than-human-world. It might mean exploring personal relationships with nonhuman beings through art or creative writing. It might mean comparing historical writings about race, gender, and species. Or it might mean using interdisciplinary knowledge to investigate a local issue of environmental justice.

A primary premise in ecocritical work in education that differentiates this approach from most other critical frameworks is the explicit recognition of the entanglement of human supremacy and Eurocentric culture—that is, the embedded worldviews and belief systems originating in Western traditions of thought, which have colonized much of the world. Ecocritical educators assert that situated at the root of social and ecological injustice is a fundamental—and problematic—assumption that humans, as a species, are understood (or self-identify) as superior to and somehow separate from all other living beings and non-living things. Thus, guiding an ecocritical framework is the understanding that the manifestation of a human-supremacist worldview is culturally constructed and inextricable from current Eurocentric industrial dominant cultural assumptions about race, class, gender, ability, age, and so forth. A foundational tenet in ecocritical work in education is that cultural habits of mind in dominant Eurocentric industrial culture are based on a system of human-supremacy—stemming from anthropocentrism—and that such a perspective is ubiquitous throughout dominant colonial culture and informs how we as humans in such a culture learn to interpret and assign value to differences.

Anthropocentrism as it contributes to human-supremacy: teaching in support of an Earth Democracy

Anthropocentrism, or human-centeredness, is a belief system that privileges humans over other species and functions to maintain the superiority of human existence by marginalizing and subjugating anything nonhuman (Plumwood, 1993, 2002). Radical ecocritical scholars and educators take the position that anthropocentrism is so pervasive in the metaphors by which we teach and learn in Western industrial schooling that, in order to interrupt the dominant belief system, educators must first learn to name and recognize anthropocentrism as problematic in connection with what can be done to rethink the vital relationships upon which we all depend. Given how dominant and foundational anthropocentrism is in schools and society, we suggest that teachers and teacher educators can—and do—play a key role in addressing and interrupting anthropocentrism, and that they must furthermore consider what a non-anthropocentric education might look like.

Ecocritical work continues and pushes the boundaries of the work of social justice scholar-activists within critical education spaces by making clear the connections between human-supremacy, patriarchy, racism, and other forms of domination like ableism, ageism, and classism. Ecocritical scholars understand that each of these value-hierarchized structures of domination, based in binary hierarchical thinking, mutually reinforce one another, while they obscure the fundamental interdependence and interrelationship of all beings. These hierarchies inform dominant cultural assumptions in Western industrial culture, and they are all based on normalized logics of domination, which means that they are inseparable and intersect in complex and contextualized ways. For example, the water crisis in Flint, Michigan or Indigenous Water Protectors in the global north (LaDuke, 2020) protesting the oil pipeline illustrate the complex intersections of anthropocentrism, racism, capitalism, etc.

Another key aspect of ecocritical work is to radically reconsider decision-making practices commonly referred to as democracy. Democracy is too often experienced as a decision-making process in which elected representatives make centralized decisions with an emphasis on voting which hinges on a majority rules mentality. This kind of decision-making (majority rules), while argued as efficient, is not at all democratic and in fact is highly exclusionary based on race, class, gender, ability, and species. Directly stated, democratic decision-making ought to be understood and practiced as: decisions being made by all those directly impacted by the decisions.

Most examples of democracy that are discussed are highly anthropocentric, and so as ecocritical educators it is important to interrupt/disrupt the power/knowledge idea that only humans (and too often only certain groups of human beings) matter in the decisions being made. We build into an ecocritical pedagogy learning about decision-making rooted in consent, consensus, and with respect for a diversity of cultures and species. An ecocritical approach to such a shift focusses on local decision-making that is responsive to and responsible for the sustained health of cultural diversity and biodiversity in the present and for several future generations. Working to interrupt and expose the diverse ways that anthropocentrism works to undermine our capacities and facilities for recognizing and valuing diversity in our decision-making can begin with relocating decision-making practices from distant centralized capitals often associated with colonial constructions of the nation-state and organizing more bioregionally with attention to diverse cultures and species locally.

With so many democracies around the planet eroded by corporations and private interests rooted in neoliberal capitalism, we have found Shiva's (2015) conceptualization of Earth Democracy to be of most use in rethinking the role of education. Specifically, Shiva (2015) helps us to think about how we can create educative spaces that support the development of a global citizenry that is not a monoculturalization of Eurocentric ideals but rather contributes to a future inclusive of a planet of diverse democracies not reliant on nation-states or centralized control. Shiva (2015) defines Earth Democracy as "consisting of the local, decentralized communities which include (and protect) all diversity in nature and culture" (pp. 8–9). Through the practice of a living democracy, decisions are made not only by those making the decisions at the most local levels, but in defense for more-than-human-species. This decentralized approach, described as localization by Shiva (2015) is:

Based on the interdependence between nature and culture, humans and other species, local and global, micro and macro, thus entailing an obligation of responsibility and eco-citizenship. Localization treats every place as the center of the world, placing every person, every being as the center of ever widening circles of compassion and care (p. 73).

Thus, Earth Democracy prescribes "self-rule and self-governance" to the communities and ecosystems which they are a part of (2015, p. 9).

As history continues to repeat itself in vicious cycles of tyrannical governments usurping power from local democracies—often in the proclamation of democratization—an example of an ecocritical position includes asking the question: *Are there any potential solutions to environmental and social injustices among nation-states that have been more than symbolic? And, Are there possibilities for organizing decision-making with emphasis and priority on diversity, social justice, and healthy ecosystems?* Contrary to models of decision-making rooted in an Anthropocentric Cultural Worldview (see Fig. 2), central to an ecocritical framework is turning local educational efforts and attention toward asking: *How might we embrace and value diversity in our social decision-making structures? And, What might it be like to learn from and with local systems of shared governance which nurtures ecojustice, political autonomy, and accountability through bottom-up democracies?*

Shiva's (2015) concept of Earth Democracy challenges the dominant form of social organizing globally, the nation-state model, with globally individuated populations governed largely by a political minority (be it a few representatives or even a single figure).

In an Earth Democracy model, which aligns with a Multispecies Equity Worldview perspective (see Fig. 1) put into practice, political power begins and emanates from the bottom up. This emerges where, “Authority is delegated to more distant levels of government on the principle of subsidiarity: things are most effectively done at the level closest to where the impact is felt” (2015, p. 57). This change from globalized market economies (based on greed and violence) to a “living economy” (based on compassion) cultivates better capacities of humans and the nurturing of living democracies for diverse species (2015, p. 75). Given such attention on addressing how a human supremacist Eurocentric worldview undermines diversity and democracy we turn to pedagogical suggestions that understood through an ecocritical framework contribute to what we refer to as ecocritical pedagogies—or efforts in education that directly emphasize learning to value the interconnectedness of all species and the interrelated ways that biological and culturally diversity are primary to how organize and learn to make decisions together as human cultures together with our local ecosystems.

Ecocritical curriculum and pedagogies: teaching for a just and sustainable future

Now more than ever, educators need to engage in pedagogies which encourage recognition of (and resistance to) all forms of domination. Doing so furthers the potential of fostering spaces of learning together to recognize harmful assumptions and actions which undergird social and ecological injustice. We admire, value, and are firm supporters of a shared commitment to respond to the undeniable atrocities that we—as humans—enact on one another. But we also believe that these atrocities are inextricably connected to the cruelties we perpetuate against non-human animals and the environment. None of these atrocities occur in isolation; they are all interconnected and mutually reinforcing. Hence, the importance of focusing on the social and ecological: diverse bioregions. To confront human supremacy together with other forms of supremacy in education, educators must work as committed allies to those suffering while challenging and confronting the systemic roots of oppression on our respective fronts. An *ecotistical* approach focusing on the pedagogies of diverse bioregions can help highlight the interconnectedness of the social and environmental.

Drawing on the ecocritical framework introduced in this entry, we propose that an ecocritical pedagogy might be best understood as pedagogies of diverse bioregions. Such a relocation of learning from the abstract idea of ecocritical to the concrete concept of a bioregion draws on these aforementioned ecocritical traditions and invites educators to address social and environmental injustices occurring at the local and global through a multispecies equity worldview perspective (See Fig. 2). Doing so always requires identifying social and material relations and challenging oppressive cultural frameworks which impact *all* species and places. Oppressive cultural frameworks like those emerging from pervasive and extreme anthropocentrism, must always be considered as connected to or closely related to upholding racism, cis normative heteronormativity, sexism, classism, disableism, and so on. The models for thinking through these oppressive mindsets in educational contexts are useful when we are teaching with folx within cultures that predominantly are based in what we presented as an Anthropocentric Cultural Worldview (See Fig. 2). We want to be clear that what we present for teaching purposes as a multispecies equity worldview perspective represents what many diverse Indigenous cultures have as a fundamental understanding of self in relationship to community. We aim our work at exposing how in particular how Western industrial cultures deeply entrenched in patriarchal, white, and human supremacism combines with hypercapitalism and ableist heteronormative individualism to constitute deeply problematic relationships among humans and also among humans and the more-than-human world. In other words, we want to be clear that we acknowledge that diverse Indigenous cultures do not need this process and that what we share is not in any way new. Rather, our hope is that the kind of work we are sharing and proposing aides in developing the awareness, knowledge, and abilities for educators in Western industrial culture to practice recognizing diverse knowledge systems and intentionally working to value different variations of a multispecies equity worldview perspective that for centuries have sustained human communities living in accordance with the diverse species, geographies, and climates in their bioregions. We hope that this kind of work better prepares white colonial patriarchal educators, or educators steeped in Western industrial cultures, to better partner with diverse educators and worldviews that share a commonality in that they do not place humans as singular beings who assume or strive to maintain supremacy.

Ecocritical pedagogies of diverse bioregions also require a fluidity of learning and doing in both the material and the cultural. Recognizing and protecting local watersheds from profit margins such as Elwha Dam Indigenous protests (Norman, 2018) and Indigenous resistance in the Brazilian Amazon (Langlois, 2022); engaging in guerilla gardening (Reynolds, 2014) to circumvent private property and food deserts—these are just some examples of engaging in pedagogies of diverse bioregions. For educators and teacher educators, we can begin to challenge oppressive cultural discourses in our classrooms. Below are some preliminary suggestions that are emerging, and always situationally evolving, from ongoing work with diverse scholar-activist educators (Edwards-Schuth and Lupinacci, 2021; Lupinacci et al., 2018, 2019, 2023) that help us to continue to move toward unlearning assumptions and practices in education rooted in an Anthropocentric Cultural Worldview in favor of recognizing and valuing how we might begin to utilize an ecocritical framework toward diverse pedagogies of bioregions in our classrooms, relocating decision-making and power to be local and in support of living systems.

- Engage in teaching and learning that explores diverse projects to rethink the dominant assumptions influencing how we, as humans, construct meaning and thus how we might learn to challenge the power/knowledge patterns in our communities through rethinking how we relate to each other and the more-than-human world. Further, make the commitment to critically and ethically examine how, as teachers, we individually and collectively understand educating, organizing, and taking action

toward supporting healthy communities that include all beings and the intrinsic value of recognizing, respecting, and representing the right of all beings to belong to and live peacefully within an ecological system. For example, critically engage in questioning how we language our world. *What does it mean to refer to stacks of trees or the waterways on the planet as “natural resources”? What does it look like to advocate for the diverse species of our bioregions? What language further perpetuates social inequalities and undermines human rights in our communities?* Furthermore, as educators, think about how we might frame lessons that include ecocritical questions that disrupt dominant worldview assumptions. Asking: *What does it mean to be human beings in our diverse communities of life? Who/what benefits and who/what suffers? How are learning relationships in our classrooms influenced by value-hierarchies and cultural assumptions? What does it mean to teach toward the abolition of Superior/inferior (Either/or) thinking and decision making and in support of challenging human supremacy in favor of anti-racist and feminist teachings as part of any relationship to places or transition to pedagogies of bioregions?*

- **Engage in critical and ethical examinations of how dominant power/knowledge schemas shape our understandings of community and how such work might offer alternative to how we come to learn about one another, make decisions, and live together with diverse species in our bioregions.** As notions of community are all too often defined in terms of white supremacy, patriarchy, and human-centered exclusion, it is important to work to reconsider community in terms of who and what is included in our definitions of this construct and how those definitions contribute to either supporting or undermining the right of all beings to coexist in peace. *Who and what might we be ignoring when we think about who is considered in decisions in our neighborhood community? What animals and plants live and make homes in our community, and how are we interrelated to them? What are we doing to practice reciprocity with our diverse human and more-than-human neighbors?* As educators we can work with our students to learn deeply about the local and to imagine how decisions might be made that consider more broadly how important it is to center diverse voices, ideas, abilities, and experiences given a historical dominance of white male colonial voices. Consider engaging students in learning to listen and be responsive to diverse language systems like breathing, smelling, listening to diverse sounds, gestures, and recognizing the diverse weather, climate, water, soils, birds, insects, fungi, forests, and other mammals and animals in our bioregions. Specifically, work to identify—or seek out—a more-than-human teacher in a local bioregion (something/one you learn from and intentionally engage in a learning relationship). At first, this is just about making a commitment to learning from this different kind of teacher–student relationship in a way that interrupts habits and assumptions of anthropocentrism and human-supremacy. Then, journal over the course of the semester and share your experiences with how you learned from, and learned to listen to, your ecological surroundings. Commit to learning about the ways in which oppressed communities have suffered and survived the extreme violence of white-male heteronormativity and to not reproducing or existing in complicity with these atrocities.
- **Engage in examining community in terms of inclusion and the diverse ways in which our living relationships can be recognized, respected, and represented through teaching and learning that interrupts and exposes dominant power/knowledge assumptions and practices in favor of learning to value and implement changes that welcome a sense of belonging, justice, and democracy for all members rooted in supporting systems of bioregions working cooperatively and with solidarity across bioregional differences.** Specifically, engage in recognizing the role activist networks play in alleviating and eliminating unjust suffering in our communities. Build networks of solidarity with these organizations. *How can single-issue or locally focused social justice groups make alliances with other social and/or environmental justice groups?* For example, in our community, we ask *what commonalities and bridges might there be between Water protectors, ANTIFA, Planned Parenthood, the Black Lives Matter movement, and the Sierra Club?* Internationally, we ask *what might efforts to address global climate change have in common solidarities with human rights issues especially with rises in the forced displacement of women and children as refugees from global conflict, famine, and fascism.* And, furthermore: *How can we as educators connect with and start conversations between these kinds of organizations, our classrooms, and communities?* This starts with learning about who is already attempting these conversations and how to support their attempts. Additionally, think about shifting classroom communities in order to encourage collaboration rather than competition and focus lessons on fostering and developing skills of community collaboration rooted in mutual aid and interdependence. School and community gardens can be a site of such teaching and learning. *How might we learn to teach from the Black Lives Matter at Schools resources? Or, what might it be like to teach or be a student in Rojava and what might we learn about other movements in different places that help us locally to resist the antidemocratic violence often coupled with nation-state identities?*
- **Engage in supporting the diverse approaches to taking up resistance and healing from the power/knowledge curriculum of an Anthropocentric Worldview rooted in Individualism & Human Supremacy and, in solidarity, show respect for epistemologies that differ from the current dominant discourses of Western industrial culture.** Support the ways in which diverse forms of resistance work to challenge value-hierarchized dualisms that perpetuate value-hierarchized thinking. For example, explore the ways in which local groups in your community are fighting against past and present acts of colonization, both in the US and internationally. Imagine how educators might teach lessons that challenge and question progress rooted in hyper-consumerism and global market ideology. For example, have students explore barter and trade in efforts to make explicit that economic systems do not need to be rooted in capitalism. Have students identify non-monetized activities that exist in the community. Additionally, have students explore the speciesist, racial, and gendered dimensions of non-monetized caring work and how that work together with commitments to antiracism and feminism upholds and strengthens communities and democratic possibilities bioregionally.

In conclusion, above all, and in addition to the attempts to disrupt institutionalized Western industrial culture that is perpetuated through and within educational spaces—we urge that we all do what we can as teachers, teacher educators, education researchers,

curriculum and pedagogy scholars, or simply just concerned citizens addressing what can be done given a global crisis. Following what we have shared as possible starting points for what we call ecocritical pedagogies we summarize all such efforts into a necessary commitment to the daily effort of turning our attention to: learning from the local and building critical friendships with other humans and non-human animals, engaging with adversaries, and sharing stories of hope and resilience while critiquing local and international relations of domination and hierarchies in favor of bioregional Earth Democracies. We believe that it is through diverse, international coalitions in solidarity of this work that we will be able to meaningfully change. In such volatile, authoritarian times with the dangers and implications of climate change imminent, it is important that critical educators challenge dominant perceptions of what currently constitutes schooling, education, and knowledge to collectively imagine with open hearts and minds what is possible when we commit to learning to be members of bioregional Earth Democracies that value diversity and the interdependencies of our living worlds and work together to shift what we learn and practice in schools toward a practice of rethinking, reconnecting, and revitalizing democratically, locally, and in support of living systems.

References

- Barker, C., Jane, E.A., 2016. *Cultural Studies: Theory and Practice*, 5th. SAGE.
- Basso, K.H., 1996. *Wisdom Sits in Places: Landscape and Language Among the Western Apache*. UNM Press.
- Edwards-Schuth, B., Lupinacci, J., 2021. Pedagogies of diverse bioregions: an ecotistical move from ego to eco. *EuropeNow* 45.
- Foucault, M., 1980. In: Gordon, C. (Ed.), *Power/Knowledge: Selected Interviews and Other Writings 1972–1977*. Pantheon Books.
- Gruenewald, D.A., 2003. The best of both worlds: a critical pedagogy of place. *Educ. Res.* 32 (4), 3–12.
- Gruenewald, D.A., Smith, G.A., 2014. Place-Based Education in the Global Age: Local Diversity. In: Routledge.
- IPCC, 2022. IPCC Sixth Assessment Report: Climate Change 2022. Retrieved June 28, 2022, from Intergovernmental Panel on Climate. https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_FinalDraft_FullReport.pdf.
- Kahn, R., 2010. *Critical Pedagogy, Ecoliteracy, & Planetary Crisis: The Ecopedagogy Movement*. Peter Lang.
- LaDuke, W., 2020. *To Be a Water Protector: The Rise of the Windigo Slayes*. Fernwood Publishing.
- Langlois, J., 2022. An Amazon Defender Stands Up for Her Land and Her People. *Yale Environment* 360. <https://e360.yale.edu/features/an-amazon-defender-stands-up-for-her-land-and-her-people>.
- Lupinacci, J., Happel-Parkins, A., 2015. Recognize, resist, and reconstitute: an ecocritical framework in teacher education. *SoJo J.* 1 (1), 45–61.
- Lupinacci, J., Happel-Parkins, A., 2016. (Un)learning anthropocentrism: an ecojustice education framework for teaching to resist human-supremacy in schools. In: Rice, S., Rud, A. (Eds.), *The Educational Significance of Human and Non-human Animal Interactions: Blurring the Species Line*. Palgrave, pp. 13–30.
- Lupinacci, J., Happel-Parkins, A., Turner, R., 2018. Ecocritical scholarship toward social justice and sustainability in teacher education. *Issues Teach Educ.* 27 (2).
- Lupinacci, J., Happel-Parkins, A., Turner, R., 2019. Ecocritical pedagogies for teacher education. In: Peters, M. (Ed.), *Encyclopedia of Teacher Education*. Springer.
- Lupinacci, J., Happel-Parkins, A., Turner, R. (Eds.), 2023. *Ecocritical Perspectives in Teacher Education*. Brill | Sense.
- Martusewicz, R., Edmundson, J., Lupinacci, J., 2021. *EcoJustice Education: Toward Diverse, Democratic, and Sustainable Communities*, 3rd. Routledge.
- McCoy, K., Tuck, E., McKenzie, M., 2017. *Land education: Rethinking pedagogies of place from Indigenous, postcolonial, and decolonizing perspectives*. Routledge.
- Norman, E., 2018. Toward a global water ethic: learning from Indigenous communities. *Ethics Int. Aff.* 32 (2), 237–247. <https://doi.org/10.1017/S0892679418000333>.
- Plumwood, V., 1993. *Feminism and the Mastery of Nature*. Routledge.
- Plumwood, V., 2002. *Environmental Culture: The Ecological Crisis of Reason*. Routledge.
- Reynolds, R., 2014. *On Guerrilla Gardening: A Handbook for Gardening Without Boundaries*. Bloomsbury Publishing.
- Rorty, R., 2009. *Philosophy and the Mirror of Nature*, 81. Princeton University Press.
- Shah, S., April 03, 2020. Think Exotic Animals are to Blame for the Coronavirus? Think Again. *The Nation*. Retrieved November 24, 2020, from. <https://www.thenation.com/article/environment/coronavirus-habitat-loss/>.
- Shiva, V., 2015. *Earth Democracy: Justice, Sustainability, and Peace*. North Atlantic Books.
- Tuck, E., McKenzie, M., McCoy, K., 2014. Land education: indigenous, post-colonial, and decolonizing perspectives on place and environmental education research. *Environ. Educ. Res.* 20 (1), 1–23.
- Turner, R., 2015. *Teaching for Ecojustice: Curriculum and Lessons for Secondary and College Classrooms*. Routledge.

Further reading

- Bowers, C.A., 2013. *In the Grip of the Past*. Eco-Justice Press LLC.
- Conklin, C., Psaros, M., 2021. *The Atlas of Disappearing Places: Our Coasts and Oceans in the Climate Crisis*. The New Press.
- Shiva, V., 2020. *Reclaiming the commons: biodiversity, indigenous knowledge, and the rights of mother earth*. Synergetic press.

Exploring currents and confluences in environmental education

Susan Jagger, Toronto Metropolitan University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Conceptions and currents of environmental education	210
The naturalist, value-centered, and holistic currents	211
Nature preschools and forest kindergartens	212
The conservationist/resourcist, problem-solving, scientific, and sustainability currents	212
EcoSchools programs	213
The systemic, humanist/mesological, bioregional, ethnographic, and eco-education currents	214
Community mapping	215
The praxic, socially critical, and feminist currents	215
School garden and nature-based participatory research with children	216
Moving forward with and in the currents	217
References	218

The field of environmental education is complex and complicated. It is inherently cross-disciplinary, drawing from across the sciences, social sciences, and humanities as it invites us to explore, meditate on, and critique our knowing and being in the world. Environmental education is reflective of who we are, where we are in space and time, what our relationships are (and are not) with the worlds that we live within, and what we hope for in the future. The inescapable complexity and ever-changing nature of environmental education cannot be framed in a snapshot but instead necessitates our consideration of the histories, philosophies, beliefs, and understandings of the land that we live with and in. In doing so, environmental education invites us to engage in learning and acting that recognizes and celebrates diversity across human and more-than-human worlds and is motivated by social and ecological justice.

This chapter¹ explores the dynamic cartography of environmental education, both historically and contemporarily, and brings together its varied landscapes of curriculum, pedagogy, and discourse. Further, it illustrates these intersecting and interacting approaches in practice and invites us to immerse ourselves in both the calm and turbulent waters of environmental education.

Conceptions and currents of environmental education

*environ*²—to circle, encompass, from French *en-* in + *viron* circle

educate—bring up, from Latin *educere* bring out, lead forth

In order to guide our navigation of environmental education's ever-changing terrain, it can be helpful to use Sauvé's (2005) currents of environmental education.³ Sauvé identifies 15 categories that capture the theoretical and practical commonalities of the diversity of environmental education programs and approaches. She distinguishes between those currents with longer histories in teaching and learning and those more contemporary currents that we see guiding our environmental curriculum and instruction (see Table 1). While distinctions are made, it is important to highlight that we can identify threads of both older and newer currents in present environmental education (Sauvé, 2005).

Sauvé's discussion of the currents of environmental education, and their respective pedagogies and practices, are inherently shaped by our conceptual foundations of environment—how do we recognize environment? She outlines six conceptions of environment that underpin the curriculum and pedagogy of environmental educators (Sauvé, 1996). *Environment as nature* valorizes a pure natural space that is separate from humans. It is through immersion within these spaces that we are able to learn about and appreciate how nature works and how we might live in respectful relation to it. In contrast, *environment as resource* views the natural world as providing us with the raw materials to support our desired quality of life. Recognizing that some resources are in limited supply and being depleted, we are encouraged to be mindful of and manage our usage of natural resources that the environment presents us with. Next, and related, the *environment as a problem* focuses on the biophysical environment and threats to its health (e.g., pollution, deforestation, etc.). Here, we are tasked with devising strategies to solve these problems and mitigate their negative impact on us. *Environment as a place to live* situates us within the local environment inclusive of biophysical, socio-cultural, historical, political, and technological features. We recognize it as *our* environment; it is a space that we dwell within and are

¹This chapter was written while the author was living and learning on and with the traditional lands of many nations including the Mississaugas of the Credit, the Anishnabeg, the Chippewa, the Haudensaunee, and the Wendat peoples. Toronto is now home to many First Nations, Inuit, and Métis peoples.

²Etymological roots of key terms were located using the *Oxford Dictionary of English Etymology* and the *Online Etymology Dictionary*.

³Please see Sauvé's (2005) article, "Currents in Environmental Education: Mapping a Complex and Evolving Pedagogical Field," for her in depth tracing of environmental education's envisioning and practices.

Table 1 Sauvé's 15 currents of environmental education.

<i>Currents with longer histories</i>	<i>Currents that are more contemporary</i>
Naturalist current	Holistic current
Conservationist/resourceist current	Bioregionalist current
Problem solving current	Praxic current
Systemic current	Socially critical current
Scientific current	Feminist current
Humanist/mesological current	Ethnographic current
Value-centered current	Eco-education current
	Sustainable development/sustainability current

Adapted from Sauvé (2005).

members of. *Environment as biosphere* embraces the environment, and the earth, as an interdependent and conscious system. Extending from theories including James Lovelock's Gaia Theory, acknowledgment of the environment as such a system necessitates living and being together in ways that respect and care for all things. Finally, *environment as a community project* invites us to get involved and participate in our communities and environment. This conception brings together socially critical activism and political action as it calls for our informed citizenry (Sauvé, 1996). Taken together, these six conceptions of environment loosely capture the significant epistemological identities that ground environmental education curriculum and pedagogy and act as the upwellings at the source of Sauvé's currents of environmental education.

Below, and with these conceptions in mind, I have explored the mixing of traditional and more contemporary currents of environmental education to highlight the connections and interactions that present themselves in their pooling and the related motivations, curricula, and pedagogies that flow out of these mixings. It is important to note that while I have brought together the following currents, these are certainly not the only groupings possible, indeed we might explore poolings of many different currents and their resulting ebbs and flows.

The naturalist, value-centered, and holistic currents

nature—essential qualities or innate character, from Latin *nasci* to be born

value—worth, price, from Latin *valere* be strong, be well

holism—unification of separate parts, from Greek *holos* whole

The *naturalist current* is, for most people, what comes to mind first when they think of environmental education. It is motivated by our connections with the natural world and includes cognitive, experiential, affective, spiritual, and artistic understandings and connections to nature (Sauvé, 2005). Historically, we can locate the headwaters of this current in the Romantic movement and in the theorizing of Jean-Jacques Rousseau. In *Émile*, Rousseau asserted that the child must learn from experiencing and exploring their environment. These statements were later supported by Louis Agassiz, a Swiss-American biologist and educator who contended that students should learn directly from nature. Specific to the early years, Friedrich Fröbel, the creator of the first kindergarten, and Maria Montessori similarly emphasized the importance of nature being integrated into curriculum and instruction to support children's cognitive and natural development. We might even locate the early springs of the naturalist current in Ancient Greece, when scholars and learners attended Plato's Academy and gathered in discussion and debate in its gardens and grounds.

In the 19th century, what we now identify as *environmental education* was included within *nature study*. Jackman's (1891) *Nature Study for Common Schools* encouraged taking school children outside to learn across the disciplines through first-hand explorations of the natural world. Key within nature study was bringing the learner into intimate and emotional contact with nature; it was an integrated and place-based educational movement. Nature study was prevalent in early childhood education well into the 1920s until the strain of the Great Depression in the 1930s and its eventual and complete overshadowing by the rise and rigor of science education in the 1940s and 1950s and nature study's lack of congruence with the increasing rigidity of school structure and organization (Disinger, 1983/2001).

Naturalist currents resurfaced again as nature and earth-focused approaches to environmental education were taken up in the 1970s (see, for example, van Matre, 1990). Here, and along with time spent in the natural world, we see the early flows of value-centered and holistic currents of environmental education. *Value-centered currents* carry with them environmental values, morals, or ethics (see, for example, Brennan and Lo, 2010) that are recognized to underpin positive environmental actions and relationships with and in the natural world. One philosophy relevant, albeit controversially, to environmental education is deep ecology. Founded by the late Arne Naess, a Norwegian philosopher, and inspired by the writings of environmental activists of the 1960s, deep ecology asserts the inherent value of all living things on earth regardless of their direct worth to humans (Naess, 1973). It highlights the complex interconnectedness of all members of the earth's ecosystems and maintains that these relations necessitate our respect and care for all of our human and more-than-human neighbors. As such, deep ecology takes up a holistic view and understands that the health and wellbeing of individuals in ecosystems depends on the health and wellbeing of ALL

members of that system (Naess, 1973). It follows that the *holistic current* extends and applies this understanding of connectedness as it embraces the inherent socio-environmental relationships with those experienced realities of the learner, including the sensory, cognitive, affective, and esthetic experiences of the environment. Knowing and being is embodied and guided by subject-object relations: physical experience with and in environment is central (Sauvé, 2005).

Nature preschools and forest kindergartens

We can see the naturalist, value-centered, and holistic currents flowing through and converging in many environmental education approaches, programs, and practices. Significant to these currents are the contributions of Richard Louv and his tireless advocacy on the importance of nature in children's lives (Louv, 2005). Drawing on trends in childhood health and development (i.e., increasing obesity, attention disorders, and mental health concerns), Louv (2005) maintains that children are suffering the effects of *nature-deficit disorder* and their disconnect from the natural world. His remedy is for children to return to nature, to freely and immersively explore and relocate for themselves the sense of awe and wonder in the natural world, its inhabitants, and its phenomena. These acts can not only improve and support the health and wellbeing of children but also of adults accompanying those explorations and of environments through children's growing understandings and connections to natural places and spaces and resultant positive environmental attitudes and actions. Nature preschools and forest kindergartens take up Louv's prescription as they afford young children the opportunity to be regularly immersed in the natural world's curriculum. Extending from the early kindergartens and learning environments of Fröebel and Montessori and inspired by Steiner's Waldorf kindergartens and the Reggio Emilia approach, nature preschools and forest kindergartens are common through Europe and are increasingly found in Asia and North America (see, for example, Sobel, 2016, 2020; Elliot, 2019; Elliot and Krusekopf, 2017). While context may vary, what remains consistent across these learning environments is their active situation of curriculum and pedagogy in children's observations, actions, emotions, questions, and conversations as they explore the environment around them. The natural world provides the context for learning and also a second teacher. Nature preschools and forest kindergartens, while presenting learners and educators with risky play situations, can in fact foster safety awareness and skills development while building a community of learners. The teachable moments and sensory pauses of awe and wonder abound in the dynamic outdoor curricular spaces, situating knowledge and nurturing connections to place and histories (Elliot, 2019).

The conservationist/resourcist, problem-solving, scientific, and sustainability currents

conserve—preserve safely, from Latin *conservare*, from *com* with, together + *servare* keep watch, maintain

resource—means of supplying a want or deficiency, from Latin *resurgere* rise again

problem—difficult question, from Greek *problema*, from *pro-* forward + *ballein* to throw

solve—loosen, unbind, from Latin *solvere* unfasten, free, pay, from *se-* + *leu* to loosen, divide, cut apart

science—knowledge, especially of a technical kind, from Latin *scire* to know, *scindere* to cut, divide, from Greek *skhizein* to split

sustain—hold the course of, keep in being, from Latin *sustinere*, *sub-* up from below + *tenere* to hold

develop—unfold, lay open, from Old French *desveloper*, from *des-* undo + *-voloper* wrap up

We can locate the headwaters of the *conservationist/resourcist current*, as well as the problem solving, scientific, and sustainability currents of environmental education, in the beginning of the 20th century when much of the western world was embroiled in war or suffering the economic impacts of those years of conflict as well as extended periods of drought. The *conservation education* movement of the early 1900s focused on the preservation of both quantity and quality of natural resources for human use; these included forests, soils, water, animals, minerals, oils, and plants, all of which provided directly or indirectly products to meet human needs and wants. From a position of conservation, natural resources and the natural world were to be managed and used wisely so that they could be preserved for use by future generations. The conservation movement originated largely from government interests and resulted in legislation, enforcement, and finally science-focused education. Related, and later in the 20th century, *resource-use education*, the social studies partner of conservation education, centered on geography and economics where conservation education focused on natural history and resource management education presented a professional relationship between humans and the environment through teaching of topics including soil conservation, water management, urban and regional planning, and landscape design. Overall, conservation education focused on resources rather than education more broadly (Disinger, 1983/2001).

Beginning in the 1950s and gaining momentum in the mid to late 1960s, with shifting social and political climates and an increasing awareness of environmental issues (e.g., pesticide use, water and air pollution) motivated in part by the contributions of environmental writers including Rachel Carson and Aldo Leopold, there was a significant renewal of interest in bringing environmental well-being and justice to the forefront of discourses. This interest catalyzed conversations on the place of environmental teaching and learning and the situation of environmental education as a separate discipline (Disinger, 1983/2001). This also brought conservation education into the *problem solving current* in environmental education. Stapp's 1969 definition of environmental education attempted to extend previous thinking about environmental education into future directions that addressed and acted upon environmental concerns and issues. He identified environmental education as being "aimed at producing a citizenry that is knowledgeable concerning the biophysical environment and its associated problems, aware of how to help solve these problems, and motivated to work toward their solution" (Stapp, 1969/2001, p. 34). Objectives within

environmental education included an “understanding that man (*sic*) is an inseparable part of a system, consisting of man, culture, and the biophysical environment, and that man has the ability to alter [those] relationships” (Stapp, 1969/2001, p. 34), an appreciation of the natural and man-made biophysical environment and its place in society, an understanding of biophysical environmental problems “confronting man” and how these problems can be solved by responsible citizens and governments, and a concern for environmental quality that motivates citizens to participate in environmental problem solving (Stapp, 1969/2001). Stapp’s early definition of environmental education recognizes that humans are part of the environment but in its wording and priorities, clearly views environmental education, and the environment itself, through an anthropocentric lens mirroring the sentiments shared by Aristotle over 1000 years earlier, “nature has made all things specifically for the sake of man.” This lens magnifies the importance of “man” while failing to capture in its field of view the very relationships that humans have with the environment that cause the problems themselves.

Concurrent, and in complement, to the problem-solving currents of environmental education flow the waters of the *scientific current*. In K-12 education, environmental education has typically been siled within science education where it can form the hook with which to pique the interest of learners in thinking about and doing science (see, for example, Gough, 2002; Sauvé, 2005). Science education in the 20th century has, for the most part, been motivated by the need for countries to become and remain competitive through advancements in innovation (Wals et al., 2014). This is clear in the acceleration of science and technology, and in turn privileging of mathematics and science in schooling, following World War II and into the 1960s as the United States battled with the former USSR for victory in the space race. As such, science education’s goals focused on developing scientific content knowledge and science process skills in learners. In environmental education and in order to solve environmental issues and problems, science was understood to be the means and the method of finding solutions and within K-12 education, we can see the flow of the scientific current through environmental education curriculum and instruction. For example, while Ontario secondary schools used to offer elective grade 10 and 12 Environmental Science courses in the 1980s and early 1990s, environmental education is now in theory taken up across all subject areas (Ontario Ministry of Education, 2017) but in practice this tends to be heavily weighted in science. We similarly see this persistence of the scientific current in both focus and method of studies on K-12 environmental education. Ardoin et al. (2018) review and analysis of research on environmental education and K-12 student outcomes highlights this, indicating that 68% of studies measured knowledge outcomes and 78% of studies followed a quasi-experimental design using pre-post measures or multiple groups (Ardoin et al., 2018). Development of scientific knowledge and the use of the scientific method strongly feature in environmental education curriculum and research.

What is fundamentally problematic in the relationship of science to environmental education is the disconnect between foundational worldviews. Whereas science is underpinned by a strong empiricist and positivist tradition through which knowledge is obtained and truth realized through an objective and rational scientific method and that dominates formal education, environmental education instead presents multiple and alternative frameworks for knowing and being. These worldviews necessarily consider social, ethical, ecological, epistemological, and ontological questions and issues. Sustainability education, or education for sustainable development, has been proposed as a bridge crossing both science and environmental education waters (Holdsworth et al., 2012).

The *sustainable development/sustainability current* emerged in the 1980s and gained strength through the 1990s and into the 2000s. This current’s flow can be traced back to the meetings and resultant reports of the United Nations and UNESCO in the 1970s, 80s, and 90s, namely the Belgrade Charter (UNESCO/UNEP, 1975), the Tbilisi Report (UNESCO/UNEP, 1977), the Brundtland Report (United Nations, 1987), and Agenda 21 (United Nations, 1992). Agenda 21 was shared following the 1992 United Nations Conference on Environment and Development (UNCED) and its 40 chapters addressed the pressing environmental issues (e.g., loss of biological diversity; population growth, poverty, and inequality; climate change) and the critical importance of looking ahead and preparing for future environmental challenges. Chapter 36 of Agenda 21—Education, Training, and Public Awareness—carried forward Brundtland’s earlier assertion that education can enable students with the environmental and ethical knowledge, skills, values, attitudes, and behaviors to support sustainable development (Palmer, 1998). The declaration of the United Nations Decade of Education for Sustainable Development (2005–2014) further honed the global focus in environmental education on sustainability.

EcoSchools programs

One group of educational programs that the conservationist/resourcist, problem-solving, scientific, and sustainable development currents flow through are the Eco-Schools programs. In response to the United Nations’ calls to action outlined above, Eco-Schools began in 1994 with a few participating schools in Europe and has since expanded to 59,000 schools across 68 countries (Eco-Schools, n.d.). Many projects take up the 3R’s of reduce, reuse, and recycle in their focus including waste reduction, energy and water conservation, and school ground greening as schools aim for bronze, silver, gold, or platinum levels of certification (Eco-Schools Canada, 2019). EcoSchools programs and projects can provide learners with unique opportunities to learn about the environment and environmental issues and take positive actions to improve their environment. It must be noted though that many of these projects and school-based initiatives (e.g., lights out in classrooms, using both sides of paper) may in fact be motivated by economic benefit. Also local and place-specific knowledge and relationships may be lost in a prescriptive and hierarchical uptake of sustainability and sustainable development (see, for example, Jickling and Wals, 2008).

The systemic, humanist/mesological, bioregional, ethnographic, and eco-education currents

system—organized or connected group of objects, from Greek *systema* organized whole, *syn-* together + *-sta-* to stand, make or be firm

human—of or belonging to man, from Latin *humanus* of man, also philanthropic, kind, gentle, polite

meso—middle, from Greek *mesos* middle, in the middle

-logy—a speaking, discourse, from Greek *logos* word, speech, statement

bio—from Greek *bios* life

region—kingdom, realm, land, province, from Latin *regere* to direct, guide, rule

ethnography—description and classification of the races of mankind, from Greek *ethnos* nations + *-graphia* study

ecology—the study of relations of living things to their environments, from Greek *oikos* house, dwelling place + *-logia* study of

In contrast to the often divisive and disconnected nature and effects of the conservationist/resourceist, problem solving, scientific, and sustainability currents, the collective of the systemic, humanist/mesological, bioregional, ethnographic, and eco-education currents tend to embrace a more connected and reciprocal approach that extends from earlier naturalist, value-centered, and holistic currents. The waters of the *systemic current* present a “big picture” of environment and environmental education that is inclusive of biophysical and social elements and their interrelations. It draws on human ecology, it invites enlightened decision-making, and its locus of attention is the ecosystem (Sauvé, 2005). Here we can see a bringing together of the holistic and problem-solving currents: understandings of natural systems and situations deeply guide decision making.

Following the systemic current’s connectedness, the *humanist/mesological current* flows through the intersection and exchange of nature and culture: it circulates at the middle of biophysical, historical, cultural, political, economic, and emotional aspects of the environment (Sauvé, 2005). Rather than separate and untouched, this current recognizes the impressions of human movement left in nature and identifies environments to be landscapes shaped by these relations. This includes city landscapes and garden landscapes to name but a few. Like earlier currents, the humanist/mesological embraces a cognitive approach but also sensory, affective, and creative engagements with and in environments and understandings gained through these experiences allow for our enriched environmental relations and interventions (Sauvé, 2005).

The *bioregional current* flows alongside the humanist/mesological current but runs deeper, circulating and taking up the natural and cultural elements of place. In contrast to sustainable development, which has grown out of international conferences and considerations, the bioregionalist current’s early waters have been at the grassroots level and encouraged by the ideas and actions of both philosophers and practitioners who emphasized the need to dwell thoughtfully and responsibly in our life place for the benefit of all systems of life (MacGregor, 2005). It carries with it the geography and the consciousness of a place and in its flow, often across political boundaries, shares the wisdom of how to live with and in that place. Bioregions typically feature a common identity in human membership and residents often live in ways that embrace and celebrate the natural communities and lands of the region. Within the bioregional current, place is defined by its natural and social systems and is grounded in its natural and cultural histories. Members of bioregions *dwell* in those places: they hold a sense of belonging and participate in reciprocal and caring relationships with its ecological and social environments (Sauvé, 2005). An example is shared by MacGregor (2005) in their discussion of the bioregional curriculum created with Cascadia, a unique bioregion in the Pacific Northwest of North America encompassing British Columbia, Washington, Oregon, and Idaho, as the focus of learning.

The *ethnographic current* focuses the opening of environmental knowledge and relationships to ways of knowing and being beyond the Western scientific worldview. This emphasis on the cultural dimension of environmental kinship recognizes and respects the wisdom of places and the narratives, traditions, and practices that teach those understandings. Looking to Indigenous worldviews, we see the shared centrality of both content and process in knowledge and the inextricable connectedness of all beings with each other and with the land. This is, for example, taken up in the theory of Tsawalk, the Nuuchah-nulth perspective that everything is one, and comes from their creation stories and is realized in their traditional knowledge and wisdom (Atleo, 2004). Recognizing that everything is connected necessitates our caring for others and the land as we care for ourselves. Given the holistic, situated, and deep knowledge and process in indigenous ways of knowing, as noted by Cajete and Pueblo (2010), traditional indigenous education “is an environmental education process that can have a profound meaning for the kind of modern education requirement required to face the challenges of living in the world of the 21st century” (p. 1128). When taken up in environmental education, we see the active embodiment of holistic indigenous philosophies as learners engage in ceremony, visioning, storytelling, observing, experiencing, reflecting, and creating (Simpson, 2002). Elders are acknowledged as keepers and teachers of tradition and culture and are consulted as experts in all stages of programming and indigenous languages allow for deeper connections to wisdom. Land is central to traditional knowledge and wisdom and opportunities for learners to connect to the land emotionally, spiritually, physically, and intellectually are inherent in teaching and learning. Finally, space for resistance and ongoing decolonization are essential for learners to critically examine dominant worldviews and reclaim and renew cultural knowledge and traditions (Simpson, 2002).

While we recognize different cultural ways of knowing and being, this does not necessitate rejecting other knowledge and wisdom. Instead, we might take up Two-Eyed Seeing, “learning to see from one eye with the *strengths* of Indigenous knowledges and ways of knowing, and from the other eye with the *strengths* of Western knowledges and ways of knowing, and using both these eyes together, for the benefit of all” (Bartlett et al., 2012, p. 335). This embrace of multiple perspectives at once validates non-Western knowledge and being while also allowing for the complement of different perspectives and interdependence of all beings to deepen and enhance our learning and living. An exceptional text that brings together Indigenous wisdom with western science in

its reminder to us and celebration of the gift of our relationship with the living world is Robin Wall Kimmerer's *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge, and the Teachings of Plants* (2013).

The *eco-education current* focuses on the development of the self within the context of the environment. It realizes the etymological genealogy of education itself as a bringing forth and leading out of the self through and with our relationship with the environment. This interaction is foundational to our eco-formation and our eco-ontogenesis. Eco-formation is the process by which the individual develops as they are reciprocally shaped and transformed by the environments that they relate with; it is our being-in-the-world (Cottureau, 1999 as cited in Sauvé, 2005). Similarly, eco-ontogenesis centralizes the place of environmental relations with the growth of the individual and identifies the unique periods of development of this relationship through childhood, youth, and adulthood (Berryman, 2003 as cited in Sauvé, 2005). If we recognize one of the goals of education itself to be active and informed membership in communities, societies, and environments, then it follows that environmental education and the nurturing of a respectful and meaningful relationship with the more-than-human world is an important part of all learners' curriculum.

Community mapping

Community mapping is a pedagogical tool that allows learners and educators to engage in environmental learning and teaching through their active and reflective explorations of place. Broadly, mapwork is means of making sense and meaning of the world and maps created are themselves representations of our knowledge, understandings, experiences, and histories of different spaces. Community maps can take many forms and the form that the map takes depends on the mapmakers. I collaborated with a British Columbian grade four teacher and her class on a community mapping project of their local provincial park, Sandy Beach.⁴ Our project was guided by place-based education, a situated and experiential approach to teaching and learning that evolves from and examines the particularities of a place, celebrates the diversity of places, informs how we live and interact within the world on many levels—socially, culturally, politically, ecologically—and highlights the interconnectedness and interplay of these systems in its pedagogy (Smith, 2007, 2012). Through its curriculum and instruction, it brings together place with the self and the community (Woodhouse and Knapp, 2000). Our three-month project focused on the local history of the park, its natural history and ecosystems, Indigenous peoples' relationships with the land, and the children's own personal connections to and stories of place. Over the duration of the project, we visited the park on three field trips, enjoyed visits to the school from the local museum director and the school's First Nations liaison teacher, and engaged in various mapping activities once or twice per week through the collaboration. As we focused our gaze on Sandy Beach, we explored and reflected on the park through our active mapwork of photography, storytelling, illustrations, poetry, and sculpture.

Through their mapwork, the students' understandings of and connections to the park and to the local environment grew and deepened (Jagger, 2013, 2016). They recognized the park's biodiversity, the adaptive nature and relationships of its plants and animals, and the rhythms of its ecosystems. The students' deep knowing of nature was supported by community mapping and nurtured through their observing, situated knowing, identifying, restoring, and transforming (Jagger, 2013). The students' attitudes and feelings about the park and the land similarly developed; they expressed care for humans and the more-than-human world, noting that the park was a safe and caring space for everyone. They also recognized the relationships between humans and the environments that we live with and in, how our actions can negatively impact other members of our environmental community, and how they might support the environment's health. Personal histories and narratives also featured significantly in the students' mapwork and in their discussions; they felt a connection to the park and recognized that they were a part of the park ecosystem (Jagger, 2016).

The praxic, socially critical, and feminist currents

praxis—practice, exercise, from Greek *praxis* doing, action

social—living with others, from Latin *socius* companion, ally, follower

critical—censorious, inclined to find fault, from Greek *kritikos* able to make judgments

feminism—qualities of females, from Latin *femina* woman, female, she who suckles

The *praxic current* in environmental education is one of action and reflection. Unlike those currents that call first for the development of knowledge and skills ahead of action, the praxic current recursively engages the learner in reflection on action and in action guided by reflection (Sauvé, 2005). This dialogic process mirrors that of participatory action research, the purpose of which is to change or improve the quality of the lives and experiences of people, communities, and environments through its practical and collaborative research process. In the praxic current's flow through environmental education, we see learners and educators affecting change through their acting and learning through, by, and for the environment. As it upsets the linearity of knowledge and skills preceding action, the praxic current and its shaping by action research also shifts the dominant relationship of educator as expert and student as novice in teaching and learning to one of equality as both educator and student are recognized as co-learners bringing unique perspectives, ideas, and questions to their explorations.

⁴All identifying names of people and places have been anonymized to protect the identity of study participants.

The *social critical current* is fed by the headwaters of critical theory, a philosophical approach that assesses and critiques society and culture to reveal power structures in place that oppress and upset those structures to emancipate individuals, communities, societies, and cultures (Sauvé, 2005). Education is reflective of those forces in effect in society that actively empower and disempower. Critical pedagogy supports learners in their reading of the world, their recognition of the social, political, and economic contradictions that create oppression and perpetuate the status quo, and their transformation and liberation of the oppressed in actions against dominant realities (Freire, 1970/2000). In environmental education, the social critical current considers and analyzes the social structures, roles, and practices that create and perpetuate environmental situations and crises. These include the explicit and implicit intentions, positions, values, and actions of those responsible for decision making. It also identifies who holds power, and who does not, and how those imbalances impact our interactions and influences on the environment. For environmental educators, the social critical current invites their active and ongoing reflection on their beliefs and practices and holds them accountable to address power imbalances to ensure that their planned curriculum is deeply and authentically realized (Sauvé, 2005).

Gruenewald (2003) proposes a critical pedagogy of place as he brings together “the best of both worlds” in his analysis and synthesis of critical pedagogy and place-based education. The resulting proposition of a critical pedagogy of place asserts the goals of decolonization and reinhabitation, aims that at once retain the emphases of critical pedagogy’s challenge of the taken for granted and oppressive assumptions and actions in education and culture and place-based education’s rich embeddedness of active citizens in their local biophysical, social, historical, economic, and political environments. In a critical pedagogy of place, decolonization and reinhabitation are reciprocal and simultaneous moves. Decolonization recognizes and upsets the persistent patterns of dominant and oppressive ideas and ideologies in schooling and culture through an unlearning of that curriculum and a re-educating that embraces socially and ecologically just and transformative ways of knowing and being. Noted by Bowers (2001), our active decolonization must not only reject dominant oppressions and reimagine ways of knowing and being, but it must also include a recovery and renewal of traditional and harmonious cultural ways of living and being. Here is where we see the reciprocal and supporting action of reinhabitation. Within this, we are challenged to “identify, recover, and create material spaces and places that teach us how to live well in our total environment” (p. 9)—to reinhabit the land. This involves living well in place, *dwelling*, knowing our place, and maintaining a reciprocal relationship with that place. Together, the complimentary and inextricably linked acts of decolonization and reinhabitation can allow for the critical exploration of the shared space of education, culture, and environment (Gruenewald, 2003). In practice, we can see this in a critical environmental education that recognizes the environment as socially situated, expands conceptions and perceptions of environment, identifies positionalities, is guided by learners’ lived experiences, critically reviews curriculum, diversifies the environmental canon, embraces people of color as environmental role models, and appreciates shared goals (Stapleton, 2020).

The *feminist current*, like the social critical current, brings with it change in viewpoints and values. It focuses on those so often unspoken power relationships that privilege one member over all others. Historically, feminism has critiqued and deconstructed the binary of man/woman but this has been extended to look more broadly at power relations, uncovering and actively upsetting oppressions based on gender, race, class, ability, and orientation (Barrett, 2005). The dominance of women is mirrored in the dominance of the environment, of Mother Earth. As such, the feminist current, and its distributaries of ecofeminism and new material feminism to name but a few, further considers the concepts of nature and environment, opening them as well to review of relations of domination and control and in their deconstruction, blur the perceived boundaries of human/nature and create “a single mutually interdependent interactive, mutually influential, and mutually interdependent post-human nature” (Merchant, 2016, p. 161, as cited in Gough and Whitehouse, 2018, p. 343). This is particularly relevant when we consider connecting children with nature in a wild and untouched state, nature outside of the city and away from human influence; this is an impossible task as human relations are now inherent in nature’s being, in postnature and post-human-nature (Gough and Whitehouse, 2018). Like some earlier and persistent currents and in opposition to rational approaches to environmental concerns, the feminist current seeks to nurture and care for others—both human and more-than-human beings and to (re)establish respectful and reciprocal relationships that are enhanced by the realization of the individual’s strengths and talents (Sauvé, 2005).

School garden and nature-based participatory research with children

The confluence of praxis, social critical, and feminist currents is clear in participatory research with children and youth in environmental education. The United Nations Convention of the Rights of the Child (UNCRC) declared that “in all actions concerning children ... the best interests of the child shall be a primary consideration” (United Nations, 1989, Article 3) and that children should be allowed to express their own views and be heard in discussions that affect them (United Nations, 1989, Article 12). This has direct implications for education and in turn for curriculum, pedagogy, and research in environmental education. Children are active, creative, and social change agents who are experts into their own lives; they offer to unique perspectives and ideas that cannot be realized through their typical role of *researched* rather than *researcher* (Barratt Hacking et al., 2012; Mayall, 2000, 2002).

Over the past 10 years, I have collaborated with elementary school students in participatory environmental education research: I partnered with a group of grade six students to explore their school garden experience and understandings and am currently working a longitudinal study of ecological citizenship and children’s participation in school garden and nearby nature-based curriculum and research. In these studies, students participated in all stages of research from brainstorming research questions to devising data collection and analysis methods to sharing findings with peers and their community. The students’ research questions and topics ranged from grade 1 explorations of plants, worms, and what people like, eat, and want to grow in the school garden to deeper

studies of birds, frogs, adaptations of plants and animals, community composition, and differences in neighbourhoods' environments and green spaces in grade 2 to increasingly complex research by grade 3 students on caring for their school garden, helpful and harmful relationships with nature, how living and nonliving things work together, waste and keeping our world healthy, local and natural history of the community, plant and animal growth, and activism and action to help the school garden and environment. The students created tools and conducted data collection including surveys, interviews, observations, and experiments and analyzed their findings by looking for commonalities and patterns, analogous to thematic coding (see, for example, [Jagger and Hayward, 2019](#)). At the end of their studies, students shared their findings with the peers and the community through posters, plays, comic books, and presentations, to name but a few.

What has been clear throughout these research engagements with children in environmental education is their growing understanding of and connection to the environment and the more-than-human world. Students' direct and active engagement with the environment in their school garden and nearby nature shaped their thinking about the land and its inherent relations ([Jagger and Hayward, in press](#)). Along with their developing knowledge about the plants, animals, and ecosystems, students increasingly recognized the blurred intersections of nature and culture, social injustices and their realization in environments, and the strength of their voices and actions to make positive changes in their communities ([Hayward and Jagger, 2019](#)). The complement of the students' active participation in environmental curriculum and research has supported their developing and enacted ecological citizenship.

Moving forward with and in the currents

Along with many other things, both uplifting and inspiring and negative and discouraging, what the very recent past has shown us is that our thinking and in turn our actions have deep and extensive impacts on those in our immediate communities and in our broader systems. We have seen the rapid and devastating spread of a deadly virus around the world. We have witnessed brutal violence on plain citizens at the hands of authority. We have come together in support of each other and have diverged in seeking out our individual interests at the expense of others. All of these actions, whether looking out for self-interest or caring for the community, have an effect on others now and in the future. And this extends to the more than human world. In restricting our movements and staying home to reduce the spread of COVID-19, we have seen changes in environments. With the lockdown of cities and towns in Spring 2020, air quality improved in many areas as emissions reduced and concentration of air pollutants decreased ([Kumari and Toshniwal, 2020](#)). Wildlife seems to be returning, or becoming more prevalent, in urban areas (see, for example, [Lowrie, 2020](#)). And, while the lagoons and canals of Venice are not actually hosting dolphins, reduced boat traffic has cleared its waters ([Daly, 2020](#)). What is clear is our inherent and inextricable connection to and influence on those who we share environments with.

Environmental education and its currents are similarly related. Each current flows into the others, creating deep and rich curricular confluences. The currents shape each other and respond to the movements of seemingly parallel or perpendicular streams and to obstructions in their paths. [Sauvé's \(2005\)](#) currents of environmental education are listed linearly and organized as those with longer traditions and those more recently emerged. However, as she notes, all of the currents coexist and present themselves to us as inspirations for engagement in environmental curriculum and pedagogy ([Sauvé, 2005](#)). The shared histories, philosophies, and movements of the currents detailed in this chapter and the confluences that I have explored highlight their relatedness. As we continue to be inspired and guided by the currents in our work as educators and as citizens, it is important to observe in our thinking and acting their presence, as roaring torrents or as trickling brooks, and reflect these relations in our planned and lived curriculum and pedagogy. Nature kindergartens, EcoSchools projects, community mapping, and participatory research on school gardens and nearby nature are inherently connected: their motivations, histories, understandings, and outcomes overlap, build, and grow from each other.

While we meditate on the movement of the currents, it is also critical, now more than ever, that we respect and allow for the pooling and circulating of environmental education's waters on and with the land. As educators and as learners, we must open up environmental education to read, reflect, and respond to the stories and relationships which it inherently is shaped by and shapes. Here, our curriculum and pedagogy must actively and consistently critique dominant worldviews and problematize and break down those power inequities that privilege based on social position, identity, and contexts. A good place to start is with the foundations of our environmental education practice and our conceptions of environment. While [Sauvé's \(1996\)](#) six conceptions of environment—environment as nature, resource, problem, place to live, biosphere, and community project—provide us with a launching point into the currents of environmental education, we must also consider the environment as a space of interacting and conflicting personal, community, and cultural histories and relations. For example, many environmental justice activists see the environment as not a space to be protected but rather as something to be protected from ([Allen et al., 2007](#)). Environmental racism's (not always) quiet persistence is clear here where marginalized communities historically and continue to be disproportionately impacted by environmental hazards including toxic waste, unsafe drinking water, polluted air, and pesticide use (see, for example, [Bullard, 1990](#); [Mohai et al., 2009](#)). All too often environmental education takes up a romantic view of children learning and being in nature, following approaches that embrace the values and ideals of western and white middle class while overlooking or ignoring the perspectives of people of color and the tensions and injustices that marginalized individuals, communities, societies, and cultures lived, and continue to live, with ([Nxumalo and Ross, 2019](#); [Stapleton, 2020](#)). Noted by [Gough \(2013\)](#) in her reasoning for developing a feminist perspective in environmental education, it is essential "to recognize the complexity of human roles and

relationships with respect to environments and that there are multiple subjectivities and multiple ways of knowing and interacting with environments" (p. 376).

Orr (1992) poignantly penned that "all education is environmental education. By what is included or excluded, emphasized or ignored, students learn that they are a part of or apart from the natural world" (p. 90). The environment is at once nature, a resource, a problem to be solved, a place to live, the biosphere, a community project, and a historical, cultural, and political space and environmental education should recognize these conceptions, their inherent connections, and their shaping of us, our communities, and the land that we live with and in. Perhaps we ought to dig down into the roots of ecology, *oikos*, and focus on our shared environment or home and caring for all of our relations of home. Through caring for and empowering all, we can work toward a world in which environmental health and well-being and environmental justice are symbiotic and concomitant.

References

- Allen, K., Daro, V., Holland, D., 2007. Becoming an environmental justice activist. In: Sandler, R., Pezzullo, P. (Eds.), *Environmental Justice and Environmentalism: The Social Justice Challenge to the Environmental Movement*. MIT Press, Cambridge, MA, pp. 105–134.
- Ardoin, N., Bowers, A., Roth, N.W., Holthuis, N., 2018. Environmental education and K-12 student outcomes: a review and analysis of research. *J. Environ. Educ.* 49 (1), 1–17.
- Atleo, E.R., 2004. *Tsawalk: A Nuu-Chah-Nulth Worldview*. UBC Press, Vancouver, BC.
- Barratt Hacking, E., Cutter-Mackenzie, A., Barratt, R., 2012. Children as researchers. In: Stevenson, R.B., Brody, M., Dillon, J., Wals, A.E.J. (Eds.), *International Handbook of Research in Environmental Education*. Routledge, New York, pp. 438–458.
- Barrett, M.J., 2005. Making (some) sense of feminist poststructuralism in environmental education research and practice. *Can. J. Environ. Educ.* 10, 79–93.
- Bartlett, C., Marshall, M., Marshall, A., 2012. Two-eyed seeing and other lessons learned within a co-learning journey of bringing together indigenous and mainstream knowledges and ways of knowing. *J. Environ. Stud. Sci.* 2 (4), 331–340.
- Bowers, C.A., 2001. *Educating for Eco-Justice and Community*. The University of Georgia Press, Athens, GA.
- Brennan, A., Lo, Y., 2010. *Understanding Environmental Philosophy*. Routledge, New York.
- Bullard, R., 1990. *Dumping in Dixie: Race, Class, and Environmental Quality*. Westview Press, Boulder, CO.
- Cajete, G.A., Pueblo, S.C., 2010. Contemporary indigenous education: a nature-centered American Indian philosophy for a 21st century world. *Futures* 42, 1126–1132.
- Daly, N., 2020. Fake Animal News Abounds on Social Media as Coronavirus Upends Life. *National Geographic*. <https://www.nationalgeographic.com/animals/article/coronavirus-pandemic-fake-animal-viral-social-media-posts>.
- Disinger, J., 1983/2001. Environmental education's definitional problem. In: Hungerford, H.R., Bluhm, W.J., Volk, T.L., Ramsey, J.M. (Eds.), *Essential Readings in Environmental Education*, second ed. Stipes Publishing, Champaign, IL, pp. 17–31.
- EcoSchools Canada, 2019. 2019–2020 Certification Guide for Bronze, Silver, Gold, and Platinum Schools. Retrieved from. <https://ecoschools.ca/wp-content/uploads/2019/09/Standard-and-Platinum-Certification-Guide-2019-2020.pdf>.
- Eco-Schools, (n.d.). Eco-Schools. Retrieved from <https://www.ecoschools.global/>.
- Elliot, E., Krusekopf, F., 2017. Thinking outside the four walls of the classroom: a Canadian nature kindergarten. *Int. J. Early Child.* 49 (5), 1–15.
- Elliot, E., 2019. There are relationships beyond the classroom: a nature kindergarten. In: Jagger, S. (Ed.), *Early Years Education and Care in Canada: A Historical and Philosophical Overview*. Canadian Scholars Press, Toronto, ON.
- Freire, P., 1970/2000. *Pedagogy of the Oppressed*. Continuum, New York.
- Gough, A., Whitehouse, H., 2018. New vintages and new bottles: the "Nature" of environmental education from new material feminist and ecofeminist viewpoints. *J. Environ. Educ.* 49 (4), 336–349.
- Gough, A., 2002. Mutualism: a different agenda for environmental and science education. *Int. J. Sci. Educ.* 24 (11), 1201–1215.
- Gough, A., 2013. Researching differently: generating a gender agenda for research in environmental education. In: Stevenson, R., Brody, M., Dillon, J., Wals, A. (Eds.), *International Handbook of Research on Environmental Education*. Routledge, New York, pp. 375–383.
- Gruenewald, D.A., 2003. The best of both worlds: a critical pedagogy of place. *Educ. Res.* 32 (4), 3–12.
- Hayward, M., Jagger, S., 2019. Enacting Ecological Citizenship: Children's Engagement With and in School Garden-Based Curriculum and Research. Paper Presented at the Canadian Society for the Study of Education (CSSE) Annual Conference, Vancouver, BC.
- Holdsworth, S., Thomas, I., Hegarty, K., 2012. Sustainability education. In: Stevenson, R.B., Brody, M., Dillon, J., Wals, A.E.J. (Eds.), *International Handbook of Research in Environmental Education*. Routledge, New York, pp. 349–358.
- Jackman, W., 1891. *Nature Study for Common Schools*. Henry Holt and Company, New York.
- Jagger, S., Hayward, M., 2019. Ecological Citizenship in the Early Years: Nurturing Children's Growing Participation in Environmental Curriculum and Research. Paper Presented at American Educational Research Association (AERA) Annual Conference, Toronto, ON.
- Jagger, S., Hayward, M. (in press). Children's embodiment of a land ethic. In: Trifonas, P.P., Jagger, S. (Eds.), *International Handbook of Curriculum Theory, Research, and Practice*. Dordrecht, ND: Springer.
- Jagger, S., 2013. "This is more like home": knowing nature through community mapping. *Can. J. Environ. Educ.* 18, 173–189.
- Jagger, S., 2016. "It's more like what you think of land": bringing together community and education through mapping. *LEARNing Landscapes* 10 (1), 105–124.
- Jickling, B., Wals, A.E.J., 2008. Globalization and environmental education: looking beyond sustainable development. *J. Curric. Stud.* 40 (1), 1–21.
- Kimmerer, R.W., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge, and the Teachings of Plants*. Milkweed Editions, Minneapolis, MN.
- Kumari, P., Toshniwal, D., 2020. Impact of lockdown on major cities across the globe during COVID-19 pandemic. *Urban Clim.* 34, 100719. Retrieved from. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7562773/>.
- Louv, R., 2005. *Last Child in the Woods: Saving Our Children From Nature-Deficit Disorder*. Algonquin Books, Chapel Hill, NC.
- Lowrie, M., 2020. Canadians Reporting More Wildlife Sightings During COVID-19 Isolation. CTV News. <https://www.ctvnews.ca/health/coronavirus/canadians-reporting-more-wildlife-sightings-during-covid-19-isolation-1.4902658>.
- MacGregor, J., 2005. Curriculum for the bioregion: learning to live sustainably in our "life places". *Appl. Environ. Educ. Commun. Int. J.* 4, 239–243.
- Mayall, B., 2000. The sociology of childhood in relation to children's rights. *Int. J. Child Right.* 8 (3), 243–259.
- Mayall, B., 2002. *Toward a Sociology for Childhood: Thinking From Children's Lives*. Open University Press, Maidenhead, UK.
- Mohai, P., Pellow, D., Roberts, J.T., 2009. Environmental justice. *Annu. Rev. Environ. Resour.* 34, 405–430.
- Naess, A., 1973. The shallow and the deep, long-range ecology movement. A summary. *Inquiry* 16, 95–100.
- Nxumalo, F., Ross, K.M., 2019. Envisioning Black space in environmental education for young children. *Race Ethn. Educ.* 22 (4), 502–524.
- Ontario Ministry of Education, 2017. *The Ontario Curriculum Grades 9-12: Environmental Education Scope and Sequence of Expectations*. Queen's Printer, Toronto, ON. Retrieved from. http://www.edu.gov.on.ca/eng/curriculum/secondary/environmental_ed_9to12_eng.pdf.
- Orr, D.W., 1992. *Ecological Literacy: Education and the Transition to a Postmodern World*. State University of New York Press, Albany, NY.

- Palmer, J.A., 1998. *Environmental Education in the 21st Century: Theory, Practice, Progress, and Promise*. Routledge, London.
- Sauvé, L., 1996. Environmental education and sustainable development: a further appraisal. *Can. J. Environ. Educ.* 1, 7–31.
- Sauvé, L., 2005. Currents in environmental education: mapping a complex and evolving pedagogical field. *Can. J. Environ. Educ.* 10, 11–37.
- Simpson, L., 2002. Indigenous environmental education for cultural survival. *Can. J. Environ. Educ.* 7 (1), 13–25.
- Smith, G.A., 2007. Place-based education: breaking through the constraining regularities of public school. *Environ. Educ. Res.* 13 (2), 189–207.
- Smith, G.A., 2012. Place-based education: practice and impacts. In: Stevenson, R.B., Brody, M., Dillon, J., Wals, A.E.J. (Eds.), *International Handbook of Research in Environmental Education*. Routledge, New York, pp. 213–220.
- Sobel, D., 2016. *Nature Preschools and Forest Kindergartens: The Handbook for Outdoor Learning*. Redleaf Press, St. Paul, MN.
- Sobel, D., 2020. *The Sky Above and the Mud below: Lessons From Nature Preschools and Forest Kindergartens*. Redleaf Press, St. Paul, MN.
- Stapleton, S.R., 2020. Toward critical environmental education: a standpoint analysis of race in the American environmental context. *Environ. Educ. Res.* 26 (2), 155–170.
- Stapp, W.B., 1969/2001. The concept of environmental education. In: Hungerford, H.R., Bluhm, W.J., Volk, T.L., Ramsey, J.M. (Eds.), *Essential Readings in Environmental Education*, second ed. Stipes Publishing, Champaign, IL, pp. 33–36.
- UNESCO/UNEP, 1975. The Belgrade Charter: A Framework for Environmental Education. <https://unesdoc.unesco.org/ark:/48223/pf0000017772>.
- UNESCO/UNEP, 1977. Intergovernmental Conference on Environmental Education—Final Report. https://www.gdrc.org/uem/ee/EE-Tbilisi_1977.pdf.
- United Nations Sustainable Development, 1992. Agenda 21. <https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf>.
- United Nations, 1987. Report of the World Commission of Environment and Development: Our Common Future. <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>.
- United Nations, 1989. Convention of the Rights of the Child. <https://www.ohchr.org/en/professionalinterest/pages/crc.aspx>.
- Van Matre, S., 1990. *Earth Education: A New Beginning*. The Institute for Earth Education, Warrenville, IL.
- Wals, A.E.J., Brody, M., Dillon, J., Stevenson, R.B., 2014. Convergence between science and environmental education. *Science* 344, 583–584.
- Woodhouse, J.L., Knapp, C.E., 2000. Place-Based Curriculum and Instruction: Outdoor and Environmental Education Approaches. Retrieved From ERIC Database (ED448012).

Reading studies of elite schools though an intersectional lens

Jane Edith Kenway^a and Adam Howard^b, ^a Monash University, Melbourne University, VIC, Australia; and ^b Education Department and Interdisciplinary Studies, Waterville, ME, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	220
Intersectional analysis: conceptual clarifications	220
Critical studies of elite schools	222
Approaches with a restricted purchase	222
Diminishment	222
Formulaic	222
Micro-analytic	223
Approaches with greatest purchase	223
Praxis	223
Relationality	224
Social contexts	224
Complexity	224
Concluding thoughts	225
References	225

Introduction

Intersectional analysis is increasingly deployed as a way of examining the complex unjust configurations of the social. Critical studies of elite schools have similar concerns but have, largely, not consciously deployed intersectional analysis. On first glance, this seems strange. But perhaps it is not as strange as it first appears. Intersectional analysis focusses on variations of subjugation. Critical analysis of elite schools focusses on variations of domination—on the dominant not the dominated. Another reason that intersectional analysis has been overtly absent from studies of elite schools is that these two fields of inquiry have very different roots. The former has its roots in Black feminism. Studying those who subjugate has not been part of its remit. The latter's roots are largely in variations of class analysis. While, feminist analysis has a long-standing presence in studies of elite schools, Black feminism has very little presence (see however, [Horvat and Antonio, 1999](#)). Nonetheless, even despite this apparent lack of connection, we suggest that many recent studies of elite schools have a broad intersectional sensibility—even though it is unacknowledged.

Our overall purpose here is dredge this sensibility to the surface by reading studies of elite schools through an evolving intersectional lens. We begin with a short history of intersectionality, identify the central debates about it, clarify some of its different expressions and note the current conundrums that it faces. We then focus on those studies of elite schools which can be read as adopting weaker forms of intersectional analysis, followed by a focus on those that can be read as adopting stronger versions. We conclude by acknowledging some of the dangers of what we have done.

Intersectional analysis: conceptual clarifications

What is intersectionality? Is it a “a theory, metaphor, paradigm,” and what are “its precise methodological attributes” ([Davis, 2008](#))? Given its contested history, rapid and wide-ranging spread and its “ambiguity and open-endedness” ([Davis, 2008](#), p. 67), it is difficult to pin down. Nonetheless, for our purposes, some initial statements are necessary. Our focus is on intersectional analysis not intersectionality more broadly.

Broadly, intersectional analysis arises from certain feminist dissatisfaction with the ways in which women's inequality and oppression are understood in everyday society—particularly by white liberal feminists. It seeks to move beyond any notion that society consists of distinct silos of power (e.g., of race, gender, class, sexuality, ability, nationality, religion and age)—only one of which is primary. In other words, it challenges approaches to analysis that are one-dimensional—those that focus on a single axis of power. These are regarded as essentialist, reductionist and universalizing.

Again, very broadly, intersectional analysis tries to address such limitations by stressing the manner in which multiple inequalities and oppressions occur together—but not just in an additive way. Hence it mobilizes such language as intersecting, overlapping, criss-crossing and cross-roads. Its metaphors include “interlocking oppressions,” “mutually-reinforcing oppressions” and the “matrix of domination.” As a slogan it is a short hand acknowledgment that women differ and that subordinated groups of women reject the suppression of their uniqueness, concerns and preferences. In its strongest iteration it looks at, “The joint operations of major systems of oppression that form a complex social structure of inequality” ([Collins and Bilge, 2016](#), p. 70).

That said, there are many narratives and debates about what intersectional analysis is, and stands for, who and where it comes from, how it evolved and spread and where it might go next. As most commentators demonstrate, there is nothing singular or linear about any such matters. Intersectional analysis is complex, heterogeneous and dynamic (Jibrin and Salem, 2015; Tomlinson, 2013). It defies closure.

Further, intersectionality is nomadic. Its roots, primarily in USA black feminism in 1970s and 80s, continue to be recognized as foundational (particularly via the work of the Combahee River Collective (1977), Collins (1989) and (Crenshaw (1989, 2004)). Crenshaw actually coined the term and initially focused on the links between race and gender. Some who preceded her were concerned about “triple oppression”—classism, sexism and racism (Gordon, 2016). Intersectionality spread rapidly in the 2000s and came to include a wider range of identities, issues and causes. Further it has taken many routes across institutions, sectors and nations and into a wide range of disciplines with diverse theoretical and methodological paradigms. Consequently, it mutated—as do most traveling theories (Davis, 2008).

Cho et al. (2013) describe the intersectionality literature arising from these routes as “centrifugal.” As such they alert us to the manner in which it is both altered by the protocols of its new locations and, in turn, alters them. It informs, and is informed by, other fields and locations and thus enriches scholarship, overall, as it travels.

But Cho, Crenshaw and McCall also point out that much, maybe most, intersectional scholarship is “centripetal.” It brings together *inter- and trans-disciplinary* scholarship (women’s/gender/feminist studies and women-of-color feminism, as well as queer, critical race and post-colonial studies). This scholarship refuses to be beholden to, and disciplined by, the disciplines or by the codes of a specific field. In drawing into its ambit, a variety of theories and methods, and in bringing together unorthodox lines of inquiry, intersectionality is unconventional, insurgent and disorderly. It insists on being, and remaining, a space of disruption and transformation.

This insistence can be juxtaposed against attempts to unify and mandate the terrain of intersectionality scholarship—particularly with regard to who are its legitimate practitioners and subjects. For example, American Black women and Black feminism have been central to the development of intersectional analysis. Thus, questions arise about whether Black women should be its only practitioners and whether it should, or can, be deployed to address expressions of social injustice other than race and gender (Drayton, 2014). The “whitening of intersectionality” (Bilge, 2013) is a concern here. But so too is its Westernism given that women from the Global South are often excluded from USA’s Black Feminism. To address this issue, Grabe and Else-Quest (2012) argue for “Trans-national intersectionality [which] places importance on the intersections among gender, ethnicity, sexuality, economic exploitation, and other social hierarchies in the context of empire building or imperialist policies characterized by historical and emergent global capitalism” (p. 159).

Intersectional analysis has attracted many debates and criticisms. But given that the concept has been deployed in the various ways, noted above, critiques of it may well be focusing on certain iterations and not others. Further, criticisms will depend on where the critic is coming from, politically and geographically, and when they were written. We comment on some such criticism here in order to make clear both the variations of intersectionality that we distance ourselves from and those we identify with. In the process we note some of the dilemmas faced by those who seek to deploy intersectional analysis.

We distance ourselves from three widely adopted, but less widely criticized, approaches. The first involves *diminishment*. These approaches are evident, for example, in various policy and management documents and popular texts. Therein intersectionality has become a slogan or buzzword—a catch-all, catch-phrase. These texts co-opt and domesticate the concept thereby trivializing and depoliticizing it—treating it as akin to domesticated notions of difference, diversity and inclusivity, for example. As Gordon (2016) observes “intersectionality can signal a kind of pluralism, in which identities are represented as equivalents on a field of competing interest groups” (p. 347). Such incorporation also raises praxis questions. For example, how readily can intersectionality be deployed in political practice? And how is solidarity to be expressed when multiple axes of oppression and injustice are invoked?

The *formulaic* is the second problematic approach. Here the main effort goes into multi-dimensional mapping, plotting or graphing. As the number of axes expands the utility of the metaphors diminishes. Indeed, the individual can become the ultimate subject if all possible axes are deployed. Further, such exercises tend to offer portrayals that lack dynamism and thus a sense of history and context. Without history and context, it becomes difficult to decide on what, and how many, axes to cover and in what combinations. What is it necessary to focus on and on what grounds? What can, legitimately, be implied, or left out, and why?

Micro-analytic approaches also require critical attention. They usually make identities, or identity categories, a central starting point and adopt an additive approach. Ironically, even despite the emphasis on mutual constitution these identities can appear to remain primary. This is why intersectionality has been criticized as a refined version of identity politics. A further concern, we have, is that the primary focus is on injurious experiences. This is a problem if such experiences are divorced from such broader structural processes as capitalism, nationalism, colonialism and imperialism and their associated regimes of power.

Some of intersectionality’s key proponents have argued, for quite a while, that it has reached a crucial turning point. For example, Cho et al. (2013) call for a “broadened field of vision” that encompasses both the centrifugal and the centripetal. Importantly, they argue,

If intersectionality is an analytic disposition, a way of thinking about and conducting analyses, then what makes an analysis intersectional is not its use of the term “intersectionality,” nor its being situated in a familiar genealogy, nor its drawing on lists of standard citations. Rather, what makes an analysis intersectional—whatever terms it deploys, whatever its iteration, whatever its field or discipline—is its adoption of an intersectional way of thinking about the problem of sameness and difference and its relation to power. This framing—conceiving of categories not as distinct but as always

permeated by other categories, fluid and changing, always in the process of creating and being created by dynamics of power—emphasizes what intersectionality does rather than what intersectionality is.

Cho et al. (2013, p. 795).

Intersectionality, they say, must expand “to include the integration of projects that bring crucial theoretical, methodological, and substantive resources to studies of intersectionality” (2013, p. 795). Collins and Bilge (2016) make similar calls. They observe that intersectionality is “at a crossroads,” call for a more “expansive view” and invite more innovative approaches. Collins and Bilge insist that intersectional analysis must address “social inequality, power, relationality, social context and complexity” (pp. 25–30). And they also insist that intersectionality must be about “praxis.” They offer compelling empirical examples of such analysis.

Critical studies of elite schools

After a period of relative lack of attention paid to elites and their schooling, the study of elite schools has gained new momentum over the past 15 years (cf Kenway and Koh, 2015). For example, the growing body of work in North America (e.g., Khan, 2011; Weis et al., 2014) illustrates the various ways in which elite schools undertake contemporary class-making. Similarly, in Europe, there has been an increased focus on students at elite schools (e.g., Draelants and Darcy-Koechlin, 2011) and elite schools (e.g., Brooks and Waters, 2015; Courtois, 2018) to understand more fully their various strategies for reproducing power and privilege. And, as various overviews (e.g., Barrera et al., 2021), and edited collections (e.g., van Zanten et al., 2015; Maxwell and Aggleton, 2017; Koh and Kenway, 2016), show English language studies of elite schools now include many studies of elite schools in the global South.

This body of work has addressed a wide range of topics, including the historical rise of elite schools across the world, the diverse experiences of students, and the various outcomes of elite schooling. While class remains the dominant conceptual focus in tackling these topics, researchers have given increased attention to intersections. Much more than their predecessors, these (mostly ethnographic) more recent studies have explored the intersections of class primarily with femininities and masculinities (e.g., Charles, 2014; Chase, 2008; Proctor, 2011) and, to a lesser extent, race and ethnicity (e.g., Ayling, 2019; Gaztambide-Fernández, 2009; Kuriloff and Reichert, 2003; Warikoo, 2016). While such researchers don’t tend to name what they do as intersectional analysis they nonetheless conduct a variation of it. And, their variations can be judged as having varying degrees of analytical purchase according to the criteria we noted above.

Approaches with a restricted purchase

Diminishment

Scholars’ increased calls for the need to explore the complicated intersections of different categories of identities in the study of elite education often fall into the trap of trying to address everything without addressing anything very deeply. One of the dangers of such efforts is to curtail a theorization of that which is being investigated. For example, Noblit (2007), in his call for emphasizing intersectionality in rethinking class urges us to take the “weight off” class in considering the various way in which class structures people’s everyday experiences. He argues that class must be “re-accomplished through other forms of difference” (p. 343). The question then becomes: How far can class be reconstituted through other forms of difference before our analysis and discussion are not really about social class anymore? Noblit’s rightly identifies the need to explore the complicated intersections of different categories of identity; however, his analysis falls short in acknowledging fully the distinct nature of class. In trying to cover everything, Noblit ends up covering nothing in depth. Class, in particular, gets ultimately lost in the kind of intersectional analysis of class that he proposes.

Another (related) danger of attempting to address everything is that it inhibits theorization of that which is being studied. For example, “privilege” has become a dominant concept in the wave of recent research (Kenway and Koh, 2015). Emerging from anti-racist feminist thought (McIntosh, 1988) the concept of privilege recognizes the various forms of advantage and the status and the power that are attached to them. Although originally developed to encourage intersectional analysis, the dominant ways of thinking about privilege are often limited to identifying advantages and status. In the study of elite education, this concept, more often than not, becomes another way of focusing on class instead of facilitating intersectional analysis. For the most part, privilege has become a limiting and vague concept in this body of work. Despite efforts to reconceptualize privilege to make it more theoretically useful for the study of elites and elite education (e.g., Howard, 2008), the prominent uses of this concept has led more to description than theorization.

Formulaic

Too often, intersectionality is addressed in additive or formulaic ways that make it empty of meaning. One such way has been to explore the intersectionality of social and cultural identifications without drawing on the necessary conceptual resources to do such analysis at more than a surface level. These explorations remain incomplete and oftentimes spurious in their failure to fully attend to the complexities involved in these intersections.

In their critique of Jack's (2019) research on the experiences of low-income, students of color at an elite university, Gaztambide-Fernández and Angod (2019) draw attention to how Jack, similar to other researchers who study elite education, bypasses the significance of race and, in so doing, ends up primarily focusing on class. They argue, "While he does 'mention' race ... Jack avoids any analysis of whiteness or how white supremacy shapes whether, how, and which students get to develop a 'sense of belonging' within an elite academic space" (p. 724). Although we contend that Jack does more than simply mention race in his work, Gaztambide-Fernández and Angod make a critical point about the necessity of using conceptual resources, such as whiteness, white supremacy, and colonialism, to analyze the intersections between, in this case, eliteness, class, and race. Consequently, they conclude, Jack fails to offer a thorough theorization of these intersections. However, Gaztambide-Fernández and Angod's overall argument that "eliteness approximates whiteness" limits the very theorization of intersections that they claim is needed in the works of Jack and other scholars studying elite education. The concept of whiteness has more limited purchase when engaging in intersectional analysis of elite schooling in contexts where whiteness is not as useful or relevant as a conceptual resource as it is in the global North.

Micro-analytic

In the surge of studies on elite schooling, researchers tend to utilize conventional ethnographic approaches. These ethnographic studies have yielded profound insights about the educational structures, practices, experiences, meanings and processes of elite schooling. However, researchers tend to stretch the limits of ethnography beyond its capacity by studying sites rather than larger flows and fields, bounded groups rather than boundaries and practices rather than processes. Studies of elite schools usually focus on the institutions themselves—their curricula and practice, the entangled linkages between secondary and tertiary provision, and how students experience such elite spaces. Although this body of work has offered valuable insights into the everyday experiences of those within these elite institutions, there is limited purchase in exploring how these experiences reflect something larger than what is happening within these institutions being studied.

Further, many recent studies in elite schools are nation bound and consequently subscribe to "methodological nationalism"—the misguided assumption that emphasizes the cultural and political form of the nation state in analysis of the modern world (Kenway and Koh, 2015). Because of the various forces of globalization, as Koh and Kenway (2012) remind us, "it must be acknowledged that a proper understanding of elite schools cannot be gleaned from within the site [or indeed the state] alone" (p. 337). However, most researchers avoid connecting what is happening within the particular site under study to the changing nation state within which their site is located and larger global shifts.

A further issue arises with regard to micro-analytic studies conducted in the global South. This relates to the intersection between theory and location. Many such studies draw on theoretical resources developed in the global North. This raises questions about methodological Westernism and traveling theory—about whether Southern theory (Connell, 2007) better assists us to understand elite schooling in the global South. A micro-analytical focus seems to preclude such questions.

Approaches with greatest purchase

Clearly, then, some versions of intersectional analysis have had only limited purchase for studies of elite schools. This is particularly so when they focus excessively on identity and/or the micro at the expense of the systemic and structural, and/or when they subscribe to methodological nationalism and methodological Westernism.

However, and in contrast, Collins and Bilge's core concerns, identified above, do speak to some existing streams of research on elite schools. Certainly, many studies of elite schools address social inequality and power. While "relationality, social context and complexity" and praxis are less well addressed they are not absent.

Praxis

Growing strand of research on elite schools is concerned about *praxis*—it is action-oriented. In recent collection, for example, Swallow and Spikes (2021) bring together a diverse group of practitioners and scholars researching, theorizing, and enacting social justice education in elite schools. Drawing on multiple perspectives and experiences, researchers and practitioners explore various ways that privileged students can be educated to encourage them to take active steps toward reducing and disrupting oppression. But some caveats are in order. If such action-oriented approaches seek to put into practice the liberal-progressive mantra that elite schools preach, rather than challenging their "disavowals" (Kenway and Lazarus, 2017), they will be weak. More broadly, to the extent that such research, either explicitly or implicitly, subscribes to weaker versions of intersectionality its *praxis* will also be weak. Stronger versions of intersectionality have more to offer but these are under-developed.

Most examples of stronger versions focus on action-oriented work conducted outside the context of elite schools. But they have important implications for such schools. Stoudt et al. (2012), for example, use critical participatory action research (CPAR) to integrate the "enactments of privilege" in youth experiences of educational, criminal, and health (in)justice. This participatory method enabled the research team and those within the community to challenge injustices, connect empirically and politically the privileged and the disadvantaged, and contest, reframe, and transform privilege. More than just a method to generate data, they used CPAR to engage in an intersectional analysis for studying, problematizing, questioning, analyzing, and challenging power dynamics. Benjamin et al. (2021) also use CPAR and intersectional feminist approaches in an effort to facilitate a process that opens

possibilities for Jewish teen girls with racial and class privilege involved in a community-based organization to develop an emerging critical consciousness.

Relationality

We have argued elsewhere (Howard and Kenway, 2015), that many studies of elite schools offer a rather restricted reading of *relationality*. Relationships within classes (class, race, gender) often take precedence over relationships between classes. Such intra-class analysis involves a form of analytical enclosure.

In contrast, a strong intersectional sensibility requires, and also complicates, inter-class analysis. It leads, for instance, to the sort of study conducted by Madrid in Brazil wherein the relationships between ruling class men and the various women in their lives are, he found, filtered through class and gender. Madrid (2013) observes,

ruling class men related differently to ruling and working class women. In the domestic sphere, the relationship between men and mothers and wives is different from the one they have with *nanas*. This dichotomy is extended to sexuality, where there are different practices depending on whether they are relating to a similar or a different woman. However, the relation with ruling class women also varies. For instance ruling class mothers have a different role in their sons' masculinity construction depending on which figures of motherhood they are associated with (pp. 222–223).

A relational focus also draws attention to the extractive economy of elite schools in national education systems—to the manner in which they wrest financial and reputational resources away from those systems and schools that serve the children of the “lower” classes. For example, Boden et al. (2020) illustrate how taxation and charity law and the British state's political apparatuses have, over extended time, led to major tax breaks for elite and other private schools and to the depletion of financial resources for government schools that serve the bulk of the population.

Social contexts

Obviously, attention to *social contexts* takes studies beyond the micro analytic approaches noted above. It invites considerations of what contexts matter, how and why. In our view, to study elite schools *in situ* is to explore the intersections between history and geography, temporality and locality, and to identify the ways in which these play out in elite schools and education systems. Dunwill's (2021) study illustrates this point. Her methodology involved four intersecting “angles of scrutiny”—“globalization, European integration, socialist heritage and tradition.” Together, these help her to explain “the multidimensional and fundamental changes that have occurred during the last 30 years of post-communist transformation” (Dunwill, 2021, p. 42) and their implications for new elite formation and new elite schools in post-communist Poland. In another study, Howard and Maxwell (2021) identify the push and pull of the transnational and national constantly at work in cosmopolitan practices that are a central part of the class-making project of elite schools. Through such practices, elite schools facilitate the process of conferring cosmopolitan status, which their students can recognize, name and draw on as needed, as they work to reproduce and extend their elite positions.

Complexity

An insistence on *complexity* involves addressing the structural and systemic and their everyday expressions. Either can be a starting point but one without the other is insufficient. An intersectional approach at the systemic level requires looking at and beyond the state, at and beyond the West, at and beyond the present. While it points to various entanglements it also disentangles.

We adopt this approach in *Class Choreographies* as we unravel the links between elite schools located in seven former British colonies, the former British empire, and globalization. As we say,

Colonialism, capitalism and Christianity are the mighty Cs of empire. They are entwined entities sharing an uneasy coexistence ... Tightly braided together like strands of the ropes used on early merchant ships, the expansion of the British Empire cannot be disentangled from forces of capitalism. Capitalist expansion is threaded with labour (in the form of slaves, convicts and indentured labourers) and materials traded as commodities, exported from and imported to the colonies; while colonial missions of conquest cannot be unravelled from colonial “civilizing” (through Christianity) missions.

Kenway et al. (2018, p. 49).

We show how this history is currently expressed in the routine life of these schools as they attempt to connect to contemporary asymmetrical globalization while also responding to the requirements of the neo/post-colonial states in which they operate.

Complexity requires a methodology that can give it due attention. One such methodology is employed by Prosser (2014) who develops a “constellation” approach by “arranging ideas and themes in a non-deterministic fashion.” In his study of an elite school in Argentina he looks at the “local and global determinants” of the school's status including the manner in which “the history of global capitalism coincides with Argentine nationalism, British ‘informal’ imperialism and Presbyterian educationalism” (Prosser, 2014). Another methodology that addresses complexity is “multi sited global ethnography” as developed by Kenway (2015),

Kenway et al. (2018). Howard (2008) has also employed this approach in a study of the ways in which students, attending elite schools, located in six countries across the world, were cultivated into particular subjects through practices and meanings attached to their global citizenship education programs. Howard identifies how these programs are adopted to align the schools in global terms as part of their larger strategy of making and remaking elites. Although elite schools position themselves as increasingly global, Howard finds that national interests remain vital in their class-making strategies (e.g., Howard et al., 2020).

Concluding thoughts

There are dangers to what we have done here. We have shown that studies of elite schools have been deploying versions of intersectional analysis without necessarily naming it as such. We suggested that the reasons include their different roots and starting points, even though, as indicated, strong intersectional studies of elite schools have certainly included post-colonial thinking. Nonetheless certain absences in studies of elite schools remain—black feminism particularly. In this respect, to claim that studies of elite schools have an implicit intersectional sensibility runs the risk of implicitly “whitening” the concept. In contrast, consciously naming it might foreground Black feminism. This could lead to important new lines of inquiry about elite schools around the globe particularly if colonialism and capitalism are analyzed from a Black feminist perspective and if the concerns of girls and women from the global South are addressed. Epstein’s (2014) *Race-ing class ladies: lineages of privilege in an elite South African school* exemplifies the possibilities here. It shows how the “legacy of colonizer and colonized,” and the racialized politics of South Africa’s past and present, intersect with notions of femininity to “whiten” black girls at an elite girls’ school (p. 244).

That said, a benefit of claiming intersectionality, in the manner we have, is that weaker and stronger versions have become evident. Thus, those who undertake future intersectional analysis of such schools can get a sense of preferable starting points—specifically according to the criteria laid down by Collins and Bilge (2016).

On the other hand, if intersectional analysis has reached a turning point, if it is to broaden its field of vision, if it is to integrate fresh foci and resources, then it might include, within its remit, studies of elite educational institutions and elites. Undoubtedly such foci would be new as would some of the analytical directions and resources we have identified.

References

- Ayling, P., 2019. *Distinction, Exclusivity and Whiteness: Elite Nigerian Parents and the International Education Market*. Springer, Singapore.
- Barrera, J., Falabella, A., Ilabaca, T., 2021. “The untouchables”: elite schools, privileges, and new scenarios. *Pensamiento Educativo* 58 (1), 1–16.
- Benjamin, B.C., Proweller, A., Catlett, B., Jacobs, A., Crabtree-Nelson, S., 2021. Intersectional feminist and political education with privileged girls. In: Swalwell, K., Spikes, D. (Eds.), *Anti-oppressive Education in “Elite” Schools: Promising Practices and Cautionary Tales From the Field*. Teachers College Press, New York, pp. 169–180.
- Bilge, S., 2013. Intersectionality undone: saving intersectionality from feminism intersectionality studies. *Du Bois Rev.* 10, 405–424.
- Boden, R., Kenway, J., James, M., 2020. Private schools and tax advantage in England. *Crit. Stud. Educ.* <https://doi.org/10.1080/17508487.2020.1751226>.
- Brooks, R., Waters, J., 2015. The hidden internationalism of elite English schools. *Sociology* 49 (2), 212–228.
- Charles, C., 2014. *Elite Girls’ Schooling: Social Class and Sexualised Popular Culture*. Routledge, London.
- Chase, S., 2008. *Perfectly prep: gender extremes at a New England prep school*. Oxford University Press.
- Cho, S., Crenshaw, K., McCall, L., 2013. Toward a field of intersectionality studies: theory, applications, and praxis. *Signs* 38 (4), 785–810.
- Collins, P., 1989. The social construction of black feminist thought. *Signs* 14 (4), 745–773.
- Collins, P., Bilge, S., 2016. *Intersectionality*. Polity, Cambridge.
- Combahee River Collective, 1977. *The Combahee River Collective Statement*. Retrieved from: <http://circuitous.org/scraps/combahee.html>.
- Connell, R., 2007. *Southern Theory: The Global Dynamics of Knowledge in Social Science*. Polity Press, Cambridge.
- Courtois, A., 2018. *Elite Schooling and Social Inequality: Privilege and Power in Ireland’s Top Private Schools*. Palgrave MacMillan, London.
- Crenshaw, K., 1989. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *Univ. Chicago Leg. Forum* 140, 139–167.
- Crenshaw, K., 2004. Intersectionality: The Double Bind of Race and Gender, Interview With Kimberlé Crenshaw. American Bar Association. Retrieved from: www.americanbar.org.
- Davis, K., 2008. Intersectionality as buzzword: a sociology of science perspective on what makes a feminist theory successful. *Fem. Theor.* 9 (1), 67–85.
- Draelants, H., Darchy-Koechlin, B., 2011. Flaunting one’s academic pedigree? Self-presentation of students from elite French schools. *Br. J. Sociol. Educ.* 32 (1), 17–34.
- Drayton, T., 2014. 10 Things White Feminists Should Know to Better Understand Intersectionality. *The Frisky*. Retrieved from: <https://thefrisky.com>.
- Dunwill, A.M., 2021. *Elite Schools in Post-communist Poland: Four Angles of Scrutiny*. Ph.D. Dissertation. Monash University, Melbourne, Australia.
- Epstein, D., 2014. Race-ing class ladies: lineages of privilege in an elite South African school. *Glob. Soc. Educ.* 12 (2), 244–261.
- Gaztambide-Fernández, R., 2009. *The Best of the Best: Becoming Elite at an American Boarding School*. Harvard University Press, Cambridge, MA.
- Gaztambide-Fernández, R., Angod, L., 2019. Approximating whiteness: race, class, and empire in the making of modern elite/white subjects. *Educ. Theor.* 69 (6), 719–743.
- Gordon, L., 2016. “Intersectionality,” socialist feminism and contemporary activism: musings by a second-wave socialist feminist. *Gend. Hist.* 28 (2), 340–357.
- Grabe, S., Else-Quest, N., 2012. The role of transnational feminism in psychology: complementary vision. *Psychol. Women Q.* 36 (2), 158–161.
- Horvat, E.M., Antonio, A.L., 1999. “Hey, those shoes are out of uniform”: African American girls in an elite high school and the importance of habitus. *Anthropol. Educ. Q.* 30 (3), 317–342.
- Howard, A., 2008. *Learning Privilege: Lessons of Power and Identity in Affluent Schooling*. Routledge, New York.
- Howard, A., 2020. Globally elite: four domains of becoming globally-oriented within elite schools. *Educ. Rev.* <https://doi.org/10.1080/00131911.2020.1805412>.
- Howard, A., Kenway, J., 2015. Canvassing conversations: obstinate issues in studies of elites and elite education. *Int. J. Qual. Stud. Educ.* 28 (9), 1005–1032.
- Howard, A., Maxwell, C., 2021. Conferred cosmopolitanism: class-making strategies of elite schools across the world. *Br. J. Sociol. Educ.* 42 (2), 164–178.
- Howard, A., Pine, J., Muench, W., Peck, S., 2020. Securing eliteness: class strategies of an elite school in Chile. *Glob. Soc. Educ.* 18 (5), 564–577.
- Jack, A., 2019. *The Privileged Poor: How Elite Colleges are Failing Disadvantaged Students*. Harvard University Press, Cambridge, MA.
- Jibrin, R., Salem, S., 2015. Revisiting intersectionality: reflections on theory and praxis. *Trans-Scripts* 5, 7–24.
- Kenway, J., Koh, A., 2015. Sociological silhouettes of elite schooling. *Br. J. Sociol. Educ.* 36 (1), 1–10.

- Kenway, J., Lazarus, M., 2017. Elite schools, class disavowal and the mystification of virtues. *Soc. Semiotic*. 27 (3), 265–275.
- Kenway, J., 2015. Ethnography “is not what it used to be”: rethinking space, time, mobility and multiplicity. In: Honig, M. (Ed.), *MultiPluriTrans: Emerging Fields in Educational Ethnography*. Transcript Publishers in collaboration with Columbia University Press, US, Bielefeld, Germany.
- Kenway, J., Fahey, J., Epstein, D., Koh, A., McCarthy, C., Rizvi, F., 2018. *Class Choreographies: Elite Schools and Globalisation*. Palgrave MacMillan, London.
- Khan, S., 2011. *Privilege: The Making of an Adolescent Elite at St. Paul's School*. Princeton University Press, Princeton, NJ.
- Koh, A., Kenway, J., 2012. Cultivating national leaders in an elite school: deploying the transnational in the national interest. *Int. Stud. Sociol. Educ.* 22 (4), 333–351.
- Koh, A., Kenway, J., 2016. Introduction: reading the dynamics of educational privilege through a spatial lens. In: Koh, A., Kenway, J. (Eds.), *Elite Schools: Multiple Geographies of Privilege*. Routledge, London, pp. 1–17.
- Kuriloff, P., Reichert, M., 2003. Boys of class, boys of color: negotiating the academic and social geography of an elite independent school. *J. Soc. Issues* 59 (4), 751–769.
- Madrid, S., 2013. *The Formation of Ruling Class Men: Private Schooling, Class and Gender Relations in Contemporary Chile*. PhD Dissertation. University of Sydney, Australia.
- Maxwell, C., Aggleton, P. (Eds.), 2017. *Elite Education: International Perspectives*. Routledge, London.
- McIntosh, P., 1988. *White Privilege and Male Privilege: A Personal Account of Coming to See Correspondences Through Work in Women's Studies*. Working Paper 189. Wellesley College Center for Research on Women, Wellesley, MA.
- Noblit, G., 2007. *Class-déclassé*. In: Van Galen, J., Noblit, G. (Eds.), *Late to Class: Social Class and Schooling in the New Economy*. State University of New York Press, Albany, NY, pp. 313–346.
- Proctor, H., 2011. Masculinity and social class, tradition and change: the production of “young Christian gentlemen” at an elite Australian boys' school. *Gend. Educ.* 23 (7), 843–856.
- Prosser, H., 2014. Reach for the stars: a constellational approach to ethnographies of elite schools. *Glob. Soc. Educ.* 12 (2), 275–289.
- Stoudt, B.G., Fox, M., Fine, M., 2012. Contesting privilege with critical participatory action research. *J. Soc. Issues* 68 (1), 178–193.
- Swalwell, K., Spikes, D. (Eds.), 2021. *Anti-oppressive Education in “Elite” Schools: Promising Practices and Cautionary Tales From the Field*. Teachers College Press, New York.
- Tomlinson, B., 2013. To tell the truth and not get trapped: desire, distance, and intersectionality at the scene of argument. *Signs* 38 (4), 993–1017.
- Warikoo, N., 2016. *The Diversity Bargain: And Other Dilemmas of Race, Admissions, and Meritocracy at Elite Universities*. University of Chicago Press, Chicago.
- Weis, L., Cipollone, K., Jenkins, H., 2014. *Class Warfare: Class, Race, and College Admissions in Top-Tier Secondary Schools*. University of Chicago Press, Chicago.
- van Zanten, A., Ball, S., Darchy-Koechlin, B. (Eds.), 2015. *Elites, Privilege and Excellence: The National and Global Redefinition of Educational Advantage*. Routledge, London.

Education and participatory democracy

Whitney Neumeyer Roach, Jonathan W. Crocker, and M. Francyne Huckaby, College of Education, Texas Christian University, Fort Worth, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	227
Interrupting the democratic ideal	228
Schooling and democratic citizenship	230
Whiteness against participatory democracy	231
Schooling as affective violence	232
Conclusion: the body as a site of resistance	233
References	234

Introduction

Systems of education are a predicate of U.S. democracy and essential to the development of informed and thoughtfully engaged civic and political citizens (Kridel, 2010). In terms of participatory democracy, where democratic actors are positioned to take a strong role in shaping the political future, education plays a prefigurative role where students can begin to engage in democratic processes although they are unable to vote. Westheimer and Kahne (1998, p. 11, 13) note that “a democracy cannot function meaningfully without informed and critical analysis of issues and social problems”. This requires “that our youngest citizens regain a sense of civic commitment and the belief that they have the capacity to work with others to improve society”.

We take issue with the notion that students need to regain a sense of civic commitment as if prior generations had this felt obligation and present generations do not. We also take issue with the confining definitions of citizenship which denies political agency according to a hierarchical, nationalistic order. Instead, we view a radical, embodied form of participatory democracy as a pathway for political engagement and increased awareness of social justice issues as always already present in the lives of minoritized youth and undocumented democratic participants. In short, by actively participating in democratic mechanisms of change, students can begin to see themselves for what they already are: powerful voices and embodied actors in a social web of political possibility.

Historically, the role of education within the US has been to prepare students to engage in democratic processes in the future once they reach voting age. (Biesta, 2016, pp. 123–124) notes that education along these lines functions as “education for democracy,” which resists the active promotion of democracy in schools and instead focuses on the knowledge, skills, and disposition to engage in democracy in the future. Often individualistic in its approach, education for democracy actively works against the collectivism and activism needed for political change in a democracy. An alternative to an individualist, preparatory paradigm in democratic education is “education through democracy” (Biesta, 2016, pp. 124–125) which emphasizes Hannah Arendt’s (2013) political conception of democratic subjectivity—*vita activa*—to “bring something new into the world”. In this case, students are not asked to wait to deliberate or participate in the political process and teachers are not asked to simply prepare students for future democratic engagement. Students are asked to take up the difficult task of actively participating in “democratic life itself” and teachers are asked to reflect on the “democratic quality” of the learning environment to support students’ direct experience with democratic processes. We wish to highlight the various obstacles to student participation in the political process that multiple models of democracy maintain when radical participation is ignored through sociolegal, external, and internal exclusion.

First, the competitive representative model of democracy is exclusionary for students who are not yet of voting age. This model relies on the judgment of professional politicians who are not obligated to consider K12 students as constituents. As a result, before these students reach voting age, their criticisms, opinions, and suggestions for policy-making and political action can be easily disregarded based on ageist notions of political legitimacy. This model effectively removes students as authorized bearers of political responsibility as the primary mechanism for political accountability—the vote—is not within their reach. Second, the competitive representation model links socioeconomic advantage (e.g., the ability to donate to political organizations, exclusionary tactics based on race, class, gender, sexuality, and ability) with political advantage (i.e., privileged access to democratic mechanisms) and thereby reproduces political inequality (Landemore, 2017). For this reason, as youth are pushed to the margins due to their class-, race-, sex-, gender-, and/or ability-position they take on a minoritized status in terms of economic power and social power which directly impedes their political power. It is easy to imagine how these sociolegal obstacles will remain in place until (1) students’ socioeconomic status is fundamentally altered through the development of anti-classist, anti-racist, anti-sexist, pro-queer, and anti-ableist political practices and (2) politically-engaged youth are recognized as legitimate political actors both socially and legally. Here, we reach a political bind for minoritized youth’s political power: they are unable to equally participate in the competitive representative model until elected representatives alter the broader sociopolitical and socioeconomic reality to allow for their engagement. Due to the structure of the US Constitution, we will indeed be left with some degree of competitive representation as a necessary mechanism of our national democratic machinery. Indeed, alternative models of democracy are not in competition with one another. Any advocacy on our part for participatory forms of democratic engagement should not be read as a desire

to supplant other forms that democratic action might take. Here, we wish to emphasize that voting and formal political representation—although one necessary and principal branch of our political efforts—is not the only means through which voters and non-voters, and students, in particular, can have a voice.

Unlike the competitive representative model that asks voters to merely elect representatives and move on with their lives, we recognize that ongoing deliberation and participation is a necessary catalyst for social justice. At this point, we arrive at two main, somewhat overlapping alternatives: deliberative democracy and participatory democracy. The deliberative model (Cohen, 1989; Fishkin, 1991; Gutmann and Thompson, 2009; Habermas, 2015) focuses on reasonableness, consensus building, and public discourse as the primary means through which publics are involved in the democratic process. One of the main criticisms of the deliberative model is what Biesta (2009, p. 105) calls “entry conditions for participation in deliberation”. He notes that it “is not so much a question about who should be included,” rather, “it is first and foremost a question about who is able to participate effectively in deliberation”. Iris Marion Young (2002, p. 53, 55) points to two forms of exclusion: external and internal. External forms of exclusion are those that “keep some individuals or groups out of the fora of debate or processes of decision-making, or which allow some individuals or groups dominative control over what happens in them”. In schools, external forms of exclusion are outlined in behavior and conduct policies which administrators and teachers are granted the authority to enforce. In fact, the hierarchical nature of the educational structure grants less authority to keep some individuals and groups out of the decision-making process as workers are placed further down the organizational structure. Students are generally placed at the bottom of this hierarchy with limited democratic control over what happens in their learning environment. Students’ physical bodies are highly regulated by authority figures under constant threat of disciplinary consequences. Internal forms of exclusion operate when “the terms of discourse make assumptions some do not share, the interaction privileges specific styles of expression, [or] the participation of some people is dismissed as out of order”.

As we highlight in subsequent sections, internal forms of exclusion in schools align with the demand to adhere to performative Whiteness, which is evaluated under Western, Eurocentric rubrics that (re)order the student body under threat of academic and social failure. These demands to perform Whiteness across curricular content and pedagogical strategies are modeled and evaluated by authorities as a means to exclude students who are affectively misaligned with the performative rubrics of Western idealism. While these forms of external and internal exclusion are bound to an exclusionary, Otherizing past and present, it is important to note that youth participation in democratic deliberation is externally and internally denied through ageist notions of youths’ lack of knowledge, reasonableness, rationality, or a legitimate legal claim to politics in general. Although other countries across the globe are experiencing anti-democratic, authoritarian shifts (Morgan, 2021) we focus here on the United States and its ongoing anti-democratic sentiment. As we will cover in detail in subsequent sections, these forms of ageist exclusion function as a more widely accepted stand-in for racist, sexist, anti-queer, and ableist exclusionary tactics by dominant sociopolitical groups and educational authorities.

Interrupting the democratic ideal

In the quest to (re)produce ideal democratic citizens, US education systems enact immense violence against those deemed non-ideal and therefore out of order. As these systems trudge forward in stagnant practices, the diversity in student body makeups, and quite literally that of individual student bodies, exist in direct opposition to schooling’s routine ordering of the body along the lines of race, gender, sexuality, ability, and class. These myriad students challenge and obstruct violent practices through embodiments, identities, and ways of being that exemplify resistance, challenging cultural flattening and erasure. Emerging alongside the civil rights movement of the 1960s, participatory democracy is an amalgamation of complicated ideas imbued by a radical reformulation of traditional democratic approaches (Bobo et al., 2004). With its three primary themes: (1) grassroots organizing and people’s involvement in decision making, (2) minimizing hierarchies and professionalization throughout social change organizations, and (3) direct action to dismantle structures of injustice, participatory democracy provides a robust framework for understanding, legitimizing, and arousing public voices. Emphasizing collective public participation, participatory democracy frames calls of involvement as an obligation of joint citizenship (Bobo et al., 2004). Although conflicting histories credit the origins of participatory democracy to either Ella Baker or Tom Hayden, Baker, by way of personal embodiment, highlights transformative democratic possibilities, value in lived knowledge, and a countering of Whiteness’ normative assumptions. The activism of the Student Nonviolent Coordinating Committee (SNCC) (Ransby, 2003), a student-led independent organization of the 1960s attributed to Ella Baker and her philosophy of social collectives, offers a prime example of participatory democracy in action. Additionally, Baker’s role in forming and implementing participatory democracy contextually anchors embodiment itself as a liberatory practice alongside the ongoing struggles against heteropatriarchy, racism, sexism, anti-queerness, and Whiteness (Bobo et al., 2004).

By asking people to make decisions about issues that impact their lives, participatory democracy rejects self-governance traditions and replaces obstacles of socially embedded hegemonies (Kridel, 2010). A participatory democracy model—constructed to counter indifference, fear, and tolerance of autonomy while shaping collective goals—requires (re)imagined, inclusive foundations restructured through unconventional explorations and actions (Bobo et al., 2004). Because participatory democracy is but a small part of an overall strategy of empowerment, it requires full engagement with dialog wherein individuals listen attentively with a preparedness to experience change and speak with hopes of being heard (Kridel, 2010). This preparedness is not only an ability to experience personal change in the future, to be heard at some point, or to practice one’s distant engagement with the social or political environment but an affective orientation toward present work on one’s self, one’s expression, and one’s environment.

Various interventions across discourses and within US education systems require re-signifying to symbolize attempts by those who have been historically and regularly suppressed by unyielding discriminatory binaries and associated expectations meant to reify ongoing subjugation. To this point, the goal of participatory democracy is not to idealize the possibility of a “fully functioning person [who] arrives at participatory democracy,” but rather to recognize a journey made possible through curriculum (Kridel, 2010), “where one practices a pedagogy of the question[and] one teaches and one learns [with] one another” (Freire, 1993, pp. 18–20). Showcasing these intellectual pursuits, however, are not merely for that of an omnipotent Truth, but rather for a “status of truth and the economic and political role it plays” in determining one’s place in society (Foucault and Faubion, 2000, p. 132).

A shift toward the democratization of schooling is made possible by way of democratic discussions and actions anchored in queries of academic promises and canons of life—what it means to have and live a good life and how to do so, how the pursuit of a fulfilling life is never-ending, how this pursuit might contribute to equity, happiness, and love, and what ongoing solidarity with others, both locally and globally makes possible (Kridel, 2010). It is not the emancipation of “truth from every system of power” this endeavor seeks, but the detachment of “power for truth from the forms of hegemony, social, economic, and cultural, within which it operates at the present time” (Foucault and Faubion, 2000, p. 133). To serve the public good through participatory democracy, systems of education must intervene in hegemonic issues of heteropatriarchy, the prominence of oppression originating in theocratic orientations, and the prevalence of “middle-class values” embedded throughout (null) curricular assumptions of race, gender, and sexuality (Hobbel and Chapman, 2009).

Interventions against such stagnant ideologies require a re-envisioning of participation, possibility, and democracy that signify an excess to its signifier and an ongoing politics to subvert and obstruct hegemonic discourses of an ideal and constructs of normalcy (Kridel, 2010). To “challenge the normalization of the White worldview of knowledge while affirming the personal and family histories of dark people,” a re-envisioning via collective participation by those so often excluded from discourses of democracy—social outliers whose very existence challenges Truth narratives of embodied worth and presumed normalcy is needed (Love, 2019, p. 136). Individuals who sit in opposition to hegemonic heterosexuality, gender conformity, Whiteness, and class norms resist through their very survival as outliers. These subjects embody resistance to traditional historical domination and, as this work proposes, are categorically queer;

not because of their sexual practice, identity, or performance but because they, as well as other young and poor folks of color, operate in the world as queer subjects: the targets of racial normalizing projects intent on pathologizing them across the dimensions of race, class, gender, and sexuality, simultaneously making them into deviants while normalizing their degradation and marginalization until it becomes.

(Cohen, 1997)

Reconstructing notions of queerness as a radical collective of social outliers proposes a politics wherein one’s comrades are determined by their relation to power rather than by their embodiment, identity, or privilege (Cohen, 1997). The radical potential of queer politics as transformative recognizes knowledge as embodied (Love, 2019) and residing with those who, because of their marginalized social positions, manifest spaces in direct opposition to dominant hegemonic exclusionary norms (Cohen, 1997).

By way of a queer participatory democracy to challenge socially constructed assumptions of normalcy and liberation, dreaming of impossible freedoms is made possible. This orientation toward freedom, known as “Freedom Dreams,” are, as Love (2019, p. 101) explains, “critical and imaginative dreams of a collective resistance”. Imagining a collective countering to White hegemony, as radical and impossible as it might seem, is the first step in direct action against oppressions. Baker’s practice of participatory democracy models freedom dreaming in her quest for collective action, offering a liberatory and inclusive democracy made up of social outliers (Ransby, 2003, p. 310). Valuing critical, cultural, embodied ways of knowing as knowledge, which has long been excluded from traditional systems of US education, asks educators to (re)evaluate whose knowledge is of most worth, whose truths are told, and what places exist for authentic learning and examination of one’s own experiences?

One such reprieve, *homeplace*, coined and defined by bell hooks (1989, p. 384), is

the construction of a safe place where black people could affirm one another and by so doing heal many of the wounds inflicted by racist domination. We could not learn to love or respect ourselves in the culture of white supremacy, on the outside; it was there on the inside, in that ‘homeplace,’ most often created and kept by black women, that we had the opportunity to grow and develop, to nurture our spirits.

By way of a queer participatory democracy and freedom dreaming, a collective might create space, critique, tear down, and “rebuild a school system that loves all children and sees school as children’s homeplaces” (Love, 2019, p. 102).

Amending hooks’ description of homeplace, Johnson (2016, p. 46) offers a new idea of *homespace*, wherein “the spaces we create are not necessarily sited in a particular “place” but rather made or assembled through our interactions within it”. Applying home-space and freedom dreaming to existing exclusionary entities, such as schools, offers an opportunity to democratically challenge educational Whiteness through engagement with radically queer participation, wherein the journey of education becomes a liberatory conduit.

Schooling and democratic citizenship

Distinct from education itself, schooling is imbued with heteropatriarchal, classist, ableist, white-supremacist democratic foundations, wherein those categorized as non-ideal by Eurocentric rubrics affectively experience schooling as a site of immense violence (Love, 2019; Vaught, 2014; Wozolek, 2020). Contemporary understandings of K-12 education cite inclusion, individuality, and opportunity, while infrastructure, curricula, and educators reinforce discriminatory abuses. Hence, schooling negates obligations to acknowledge, adjust, and atone for historical inequities enacted upon oppressed populations. Instead, marginalized students are hyper-visible, over-policed, devalued, and gaslighted when attempting to negotiate the ever-present violence of educational endeavors—experiences inscribed by heteropatriarchal, ableist, and classist histories of exclusion and abuse. Along these lines, schooling is closely connected to the ways in which students learn to become engaged in democracy. An important portion of non-ideal categorization of democratic participants is not only how students are taught to participate democratically but also who has worth in the democratic process.

Perhaps the most powerful area of critique for educational researchers is citizenship education. When educators take up questions about what it means to be a citizen, how one becomes a citizen, or what constitutes citizenship more generally, they may look to legalist frameworks which are the most common orientation in state social studies standards (Stuteville and Johnson, 2016). In other words, who is or is not a citizen depends on their legal status that confirms a claim of participatory power through the state. Instead, the term citizen might be up for critical redefinition outside the confines of legal definition in the classroom. Markowitz (2015) argues that the current citizenship scheme at the federal level has profound negative consequences for immigrants, the communities they live in, and a democratic society more broadly. For example, absent a state-affirmed citizenship status, the government has no obligation to ensure their “fair and humane treatment” (p. 911). In addition, undocumented people are emotionally, socially, economically, and politically deprived as a result of their legal status. In order to combat the ongoing marginalization of undocumented people in the US, he claims there is a justified legal argument to be made that individual states can confer the rights of states citizenship without the support of the federal government. This would accomplish a movement of these people from a “semipermanent underclass” (p. 915) to the democratic participants they already are whether this status is conferred legally or not.

The role that education can play in this critique of citizenship is to deemphasize the legalistic definition of the term and, instead, foreground the knowledge, skills, values, and dispositions (Johnson and Morris, 2010) which constitute a citizen extralegally. Accordingly, this necessitates a critique of the participatory citizen in a participatory democracy. Westheimer and Kahne (2004) outline three categories of what good citizens are and what they do. At one end of the spectrum, there is the personally responsible citizen who “acts responsibly in [their] community” by doing things like obeying laws and volunteering their time (p. 241). These people understand their role as one defined by their compassionate, honest, and responsible character. At the other end of the spectrum is the justice-oriented citizen which is often “least commonly pursued” (p. 242). These people work to address systemic injustices in their community and emphasize the root causes of the social justice problems (pp. 242–243). Between these two orientations is the participatory citizen—those who are active participants in the existing structures of democracy. By participating in “collective, community-based efforts” (p. 242) to make their communities better, participatory citizens need knowledge of how communities are organized and governmental agencies function (p. 240).

For more than a century, schooling has negotiated the task of bringing together groups of students from unequal social classes and introducing them to these democratic processes and conceptions of citizenship. Through its original purpose, the assimilation of Others into “the superior culture—White, Anglo-Saxon American” (Ladson-Billings, 2018, p. 6), schooling actively eradicates “the socially inferior ... [by preserving] ‘old stock’ American values” (Winfield, 2007, p. 100). Vaught (2014, p. 153) defines educational institutions as state apparatuses that maintain White liberal initiatives, perpetuating “systemic racialized, gendered, sexualized and aged power relations” while (re)producing ideological models of democracy. One such ideological model is related to character education and an emphasis on good citizenship and personal responsibility predicated on obedience to law and order. Along these lines, those who are considered socially inferior in terms of character are those who undermine these “old stock” American values when they refuse to acquiesce to the so-called “superior culture”. With its complex exclusionary histories, schooling does not support the very existence of students who do not meet this democratic ideal couched in a White, able-bodied, heteronormative model, and instead, schools remain stagnant in discussions and actions regarding the presence of non-ideal democratic participants (Love, 2019). What may be considered non-ideal in the citizenship framework is any participatory action that undermines the common way of engaging in the democratic process or, even worse, one that attempts to get to the root of social justice issues like income inequality, racism, sexism, homophobia, and so on.

As institutions of education “legitimize a hierarchical social order based on privileged and marginalized identities” (Rawlings, 2019, p. 701), they encompass various far-reaching democratic controls to reify and position individuals and identity groups based on differences. One mechanism of this control disproportionately criminalizes Black students in particular (Walton, 2018) through white-supremacist policies deemed necessary for maintaining safety within the institution (Rawlings, 2019). Black youth, when identified as undisciplined, are the perfect candidates to become schooling’s scapegoat and “source of societal ill” (Vaught, 2014, p. 157). Inequities, both implicit and explicit, ensure a system that “creates marginalized student and teacher populations through everyday teaching and learning” (Wozolek, 2015, p. 7). Identifying and Othering of marginalized students (Apple, 2006; Pinar, 1998) is further reified within systems of education wherein “historical and political structures of Anglo, heteronormative power are continually maintained” (Wozolek, 2015, p. 8). Mandated curricula normalize what Wozolek (2015, p. 14) refers to as “contemporary iterations of educational lynching,” thereby reinforcing students’ feelings of inadequacy and inferiority. As

such, educational institutions delegitimize undesirable democratic participants by denoting success and achievement explicitly within the context of White heteronormativity (Grady et al., 2012). The citizenship framework is often considered progressive in the sense that students move from personal responsibility to participatory to social-justice orientations (Blevins et al., 2020) which points to the fact that the personal responsibility orientation is the default position. It follows that one of the most powerful ways students' perceptions of their current or future political agency is undermined is to attack their character and some innate ability—couched in race, gender, sexuality, and/or ability—to be responsible citizens.

Today, we confront the historical obfuscation of intellectual, social-justice orientations as a result of Othering through both subtle and unconcealed mechanisms of silencing and eradication. Dismantling such embedded structures is possible by way of radical democratic participation and advocacy for social-justice orientations in the classroom. Excluding political undesirables' intellectual traditions amounts to de facto exclusion from fundamental democratic processes which rely upon deliberation and participation in public fora. For example, the privileging of "white scholarly discourse" as a means for "examining Black experiences" in academia (Grant et al., 2015, p. xvii), the democratic control of educational systems by White architects to ensure a "social order" based on racial supremacy (Watkins, 2001, p. 181), and the parallel movement of race-based worth, eugenics, and education (Winfield, 2007) point to an ever-present structure of schooling that measures, sorts, rewards, and punishes folk based on a rubric of democratic citizenship. These notions of whose scholarly work, whose leadership and guidance, and which people are more orderly in their race as well as their "class, habit, and behavior" (Winfield, 2007, p. 159) depend upon a White, heteropatriarchal, and anti-democratic network of knowledge and practice to ensure that the Other is always perceived as deficient under the criteria of democratic citizenship. In the following section, we consider Whiteness as the fundamental ordering mechanism that ranks and sorts human beings in preparation to be marginalized by anti-democratic structures in and out of schools.

Whiteness against participatory democracy

As a constructed yet meaningful social category, Whiteness is held together through its malleability. By maintaining an internal heterogeneity, hegemonic Whiteness endures alongside other possibilities of subordinate and marginalized racialized Whiteness (Lewis, 2004). Agential and pliant, Whiteness ensures its high social station "vis-à-vis interracial boundaries" (Hughey, 2010, p. 1292) while simultaneously exploiting internal variances of social class, gender identity, age, and sexuality (Omi and Winant, 2014). To maintain unequal systems of political power within schooling, practices and policies connect understandings of the ideal to notions of Whiteness and reify marginalization and racialization (Carter Andrews et al., 2019). Whiteness, defined by Sarah Ahmed (2007, p. 15), is "an ongoing and unfinished history which orientates bodies in specific directions".

Reaching far beyond limited racial identity constructs, abstractions of Whiteness attribute varied appearances, behaviors, and performances with a collective social privilege to validate selected actions and bodies (Ahmed, 2007). However, this collective privilege, which presents as natural formations of power, is not universal and can be countered through a collective recognition of and resistance against hegemonic ordering. When embodying Whiteness, one carries with them "the privilege of being perceived as innocent despite overwhelming evidence that signals otherwise" (Carter Andrews et al., 2019, p. 2538). In spaces awash in Whiteness, such as educational institutions, White bodies find comfort in this extension of their shape (Ahmed, 2007). Moreover, attempts at reconciling Whiteness in education against the illusions of social justice result in an educational dissonance for unwelcomed students who must navigate the precariousness of their identity against the ordering mechanism of Whiteness. However, for academic social justice to prove successful, an understanding of how affective Whiteness violently structures Western thoughts (Derida, 1976) and an understanding of what is and is not ideal under a anti-democratic rubric is paramount.

Whiteness's subtlety justifies its social control and distance (Lewis, 2004), while a guise of neutrality allows it to operate as a measure of humanity (Brodkin, 1998) against those who deviate from anticipated democratic norms—ordering them as Other. Affect also orders and positions, not by way of cognitive understanding, but rather through embodied experiences in relation and response to a (dis)orientation and (mis)alignment with surroundings (Ahmed, 2006). Socially and physically orienting inhabitants, spaces take on Whiteness' qualities, welcoming those who align while manifesting hostility for Others. As structures and sites of power "acquire the 'skin' of the bodies that inhabit them," locations of authority maintain exclusion—making certain they are (un)inhabitable for some" (Ahmed, 2007, p. 157). Within these democratic institutions of Whiteness, it is embodied Whiteness that goes unnoticed; Whiteness is the unremarkable presence upon which all others appear hyper-visible, deviant, or as those who have, in fact, deviated from socially negotiated practices (Ahmed, 2007). In addition to embodied racialization, Whiteness's abstraction serves to recognize a compulsion of order as one which demands the (de)valuing of bodies. When wielded as an ordering device, "Whiteness modifies and regulates bodies, assuring they 'line up' and fade into a model 'sea of Whiteness'" (Ahmed, 2007, p. 159). Those who fall into order, performing and modeling Whiteness, journey uninterrupted, while those who cannot—or will not—model Whiteness are met with suspicion, slowed, or stopped. The refusal to model Whiteness results in forms of internal exclusion from democratic debate (Young, 2000). Those who do not succeed in maintaining a performative alignment to Whiteness are systematically ignored, vilified, or expelled from public discussion.

Another crucial element in this matrix of power is that Whiteness serves as a foundation for normative citizenship that violently deploys "the logics of comparative anti-intersectionality" (Brandzel, 2016, p. 34). These logics assert themselves in "state-led practices of anti-intersectionality" (p. 23), which are streamlined through institutions of formal education via hidden and overt forms of curriculum. Along these lines, the school is but one site where the "state reproduces hegemonic epistemologies of identities that are normative, single-axis, and comparatively valued against other categories of identity" (p. 23). Whiteness is but one root of this

comparative epistemic project alongside others, as we have mentioned, that intersect with gender, sexuality, ability, and class hierarchies. The construction of the ideal participatory citizen through education centers around these kinds of anti-intersectional assimilation (Abowitz and Harnish, 2006). Broader notations of assimilationist citizenship coupled with White epistemologies conveyed through state education are keys to accessing the participatory model of democracy—to become an idealized participant in an idealized democratic system. Therefore, citizenship in the US conceptually narrows who participates and how they engage the model of democracy the state has created and maintains. Schools are a functional arm of this mechanism that inculcates children into its system beginning at the age of three or four years old and continues through the first two decades of a person's life.

As White orientations situate academic spaces, it ensures only those who acquiesce to Whiteness belong. Thereby granting access to individuals insofar as they cooperate through performative habituations to Whiteness—an act executed by those with varying racial identities. Indistinguishably tethered to white-supremacy, Whiteness assures Whites' racial position atop the social hierarchy via absolute contamination (Tsing, 2015) of (inter)actions (Bonilla-Silva, 2012). Whiteness is, as Pfeil (2007) suggests, the privileged mode of participating within systems of White heteropatriarchy and is conferred by those who experience the standpoint of external racialized Whiteness. Internalized Whiteness is enacted unconsciously through habitual actions serving to reify white-supremacist ideals to exclude all other ways of knowing and being (Grimes, 2017). While White racialized bodies present visible Whiteness, ascribing to an external reproduction of Whiteness with ease, they too must attend to the task of White performance and ordering. The added burden of "acting White" for those marginalized by racism, however, has profoundly violent psychological consequences, including that of cultural dissonance (Ogbu, 2004). This violence epistemologically silences social-justice counter-narratives by isolating their experiences from any connections to racism, discrimination, or dispossession (Carter Andrews et al., 2019).

With its regular application and practice, Whiteness inscribes actions to define appropriate embodiment through spaces already awash in White ideals. Individuals ensure a "correct and competent" Whiteness via performances that align with oppressive symbolic practices (Hughey and Byrd, 2013). These "socially negotiated practices," (Lewis, 2004, p. 634) such as "talking White" (Boulton, 2016, p. 5) or not adhering to time constraints "carry a variety of material rewards and penalties" (Lewis, 2004, p. 634) including access to or exclusion from political engagement. Additionally, policies demanding students request permission to attend to basic bodily needs like eating, drinking, or using the restroom strip them of their agency and dignity. In its elusiveness, Whiteness is not recognized by those who inhabit its manifestation. Instead, Whiteness "trails behind" them, orienting the body within White spaces, shaping "what bodies can do" (Ahmed, 2007, p. 157). Thereby ensuring those with White racialized bodies can "[hijack] the institutional structure of society for private benefit" (Jaycox, 2017, n.p). Compulsive ambient practices of ideal Whiteness permeate historical logic by dictating who can take up space and what they can do (Ahmed, 2007). Functioning as a *trompe l'oeil*, counter-Whiteness's epistemological silencing works to mask discrimination (Carter Andrews et al., 2019) while constructions of Whiteness materialize as obstacles for democratic participation for those who do not enact it (Wozolek, 2015/2020). These perceptions of political deficiency are not merely reproduced in abstract, bureaucratic workings of educational institutions. These violences have real, material impacts on students' bodies as they navigate Whiteness-as-order in schools, are routinely disciplined for transgressing the boundaries Whiteness establishes, and come to develop an affective resistance to behavior within the democratic order.

Schooling as affective violence

Amid the deluge of idyllic clichés about education—that it can save us, that it can promote socio-economic equity, that it can promote democratic activism—there are more insidious consequences for students who experience schooling as a violent encounter with powerful institutions and authorities. These consequences may be difficult to come to terms with, especially for those of us who are embedded in these institutions who would like to think of the teaching profession as an exemplar of care for all children, their academic achievement, and their social and economic advancement. However, attractive as these clichés are as they work to inspire teachers and administrators to keep fighting for social justice, learning to engage in radically democratic teaching practices, and implementing liberatory pedagogies, these goals remain unachievable if educators continue to ignore the harmful effects of schooling and refuse to undermine them. To acknowledge these platitudes as such would mean that the increasingly diverse but still predominantly White, middle-class teaching force and their administrative counterparts (Ingersoll et al., 2018, pp. 14–16) would be required to confront the color-evasive (Annamma et al., 2017) approach to their practice instead of "[ensuring] the maintenance of the status quo, to ensure that power, the culture of power, remains in the hands of those who already have it" (Delpit, 1988, p. 285). Harber (2004, p. 40) notes that this refusal to see the violence of schooling is the result of a "tendency for *ought* to come before *is* and for it to have an idealized, taken-for-granted image as inherently good and beneficial". This *ought* is an *ought* about character which sorts for race, ability, class, sexuality, and gender; ignoring the complex relationality of individual and group identity and the violence that is socially reproduced through schooling. Instead, it is crucial that educators become willing to interrogate these violences and work to uproot them on their own campuses and in their own classroom practices.

Drawing on Salmi's (2000) work on violent educational formations, Harber (2004, pp. 44–45) describes *repressive* violence as "the deprivation of fundamental human rights" and alienating violence as "the deprivation of higher rights" that exists in schools. These less visible forms of violence lend themselves to indirect violence as "a violation of the right to survival" as well as actual, physical, direct violence in schools and society as "deliberate injury to the integrity of human life". In other words, it is clear that the repression of students' freedom of thought, speech, and equality originate within an alienating social environment

saturated with classism, transphobia, racism, sexism, Islamophobia, and anti-queerness in schools. These exclusionary boundaries are a catalyst for other purportedly more serious forms of violence like physical and psychological bullying, hate crimes like assault and murder, and other forms of mistreatment.

Violent schooling cannot be solved by addressing these violences individually but by actively confronting the authoritarian structure of schools—where children have little to no power over what they learn or how they learn it and where disciplinary mechanisms demand thoughtless obedience to authority and rigid compliance with policies. Along these lines, the material spaces of schools must be interrogated in parallel with the material bodies within these spaces and, in particular, their ability to engage in democratic processes. When students resist democratic disciplinarity and foreclosure by making their bodies, knowledges, and ways of being visible and vocal, they risk an encounter with the institutional violence of schooling and exclusion from decisions that influence their lives. In short, to be present as they are, embodied as they are actively confronts the anti-democratic ought of schooling. Instead, students should be seen as the legitimate democratic actors that they are rather than as schools and educational authorities dictate they ought to be.

Understood as an ever-evolving entity, assemblages of violence include and are informed by historical and current context “constructed through violent inter-actions that touches one’s body and/or is taken against one’s ways of knowing and being” (Wozolek, 2018, p. 15). Through coded language, access, assumptions, and White ideals, assemblages of aggressions manifest in violence toward unwelcomed bodies. For those who do not embody racialized Whiteness, regardless of their White performativity, its mandated reproduction violently displaces identity causing psychological, cultural dissonance (Ogbu, 2004). Importantly, these violent assemblages are not merely coincidental with a historical legacy of political dispossession through violence that continues despite efforts to resist it. Instead, these violences are ongoing material violations that operate on a “logic of injury in addition to a process of subordination” (De Lissovoy, 2015, pp. 63–64). This means that the demands to reproduce Whiteness and the racism that must accompany this project are both “active and willful—not simply required of Whites but even enjoyed by them”. as they are granted access to forms of democratic participation and deny participatory access to Others. Schools, which function as spaces for students, in varying degrees, to comply with or resist a forced assent to Whiteness, are also spaces where “categories of deviance” are reproduced and “tools for injury” are deployed against politically minoritized students.

The very embodiment of those deemed Other in these limiting spaces affectively codes how marginalized participants negotiate lineages of historically frozen yet ever-changing abuses. As agential (Rosiek, 2017), discrimination remains embedded historically while malleable to present-day practices, adapting to current racist, sexist, classist, ableist, and anti-queer hegemony. Affect works alongside this transformative and evasive Whiteness as a spillway for the past to be of current significance. Situated within an ongoing past, historical rememberings are persistently woven throughout present-day encounters and influence future endeavors (Haugeland, 2013). Connecting to affectivity recognizes contextual nuance beyond that which is physical to situate the self in such a way as to acknowledge the ever-present “always more” alongside a lack of understanding, the “always less” of intelligibility (Slaby, 2017). That is to say, students, educators, administrators, and community members must labor, albeit in ways not always perceived, by continuously tuning into their surroundings. While one absorbs, adapts, and processes the world around them, they remain in or out of tune, embroiled in multifaceted and untold ways (Slaby, 2017).

Seen and understood through the white-supremacist gaze of education, White students are viewed as the “unequivocal embodiment of innocence” while Black students are understood as “dangerous, promiscuous, sassy, threatening, overly independent, and needing less nurture, protection, support, and comfort” (Carter Andrews et al., 2019, p. 2539). The correlation of Whiteness to “normal” underpins notions of Whiteness to safety (Blount, 2000, p. 92). Conversely, as Blackness is deemed dangerous, a vision of “black innocence is a structural impossibility” (Wun, 2016, p. 744). Reinforced by policy, Black children who “exist as a trope in liberal state schools” are already recognized as discipline subjects to alleviate the school of responsibility by prioritizing individualization (Vaught, 2014, p. 156). In doing so, Black students face negative experiences linked to behaviors coded as individual inadequacies. Thereby declaring that these “unfit subject[s] are not conceived of as unfit through the systems of power and domination that produce them, such as misogyny and homophobia,” but rather by an innate deficit (Vaught, 2014, p. 156).

Cultural reproduction within educational institutions (Bourdieu, 1971; Nash, 1990) establish and reinforce “White middle class cultural norms” as a “foundation for a reproduction of privileged cultures in the larger education system” (Jennings & Lynn, 2005, p. 21). Schooling mandates the adoption of “stereotypical, demeaning, and pejorative representations of ways of being” wherein notions of intelligence as White property dehumanize and typecast racially marginalized students who are perceived as unintelligent and disorderly (Carter Andrews et al., 2019). Notions of beauty, intelligence, and worth constructed within Western heteropatriarchal parameters ignore Blackness (Collins, 2004) and presume their danger by exclaiming they are not well suited or welcome within the field of education. These notions are not only expected to be learned and recited but also practiced and embodied to the extent that marginalized individuals repeatedly assent to Whiteness, as they are implored to leave their non-normativity behind but remain perceived as Other despite ongoing violent adaptation to anti-democratic ways of knowing and being.

Conclusion: the body as a site of resistance

While structural change is paramount, it is unlikely to occur in any significant form in the immediate future unless teachers begin to work to offer students opportunities to engage in democratic processes. To undermine Whiteness in education and resist

exclusionary, anti-democratic ordering mechanisms, responsibility must fall to those with authority—not to students who, as this work has demonstrated, embody the very notions of resistance through their educational and political survival. Privileged educators as co-conspirators (Love, 2019), such as those who are heterotypical or embody Whiteness in light-skinned-European bodies, must labor alongside so-called non-ideal students and colleagues. They should sacrifice their privilege to enact antiracist, pro-queer, fully inclusive democratic practices. One small step educators may be able to take is to take up a social-justice orientation toward citizenship and democratic participation. Individual educators' recognition of Whiteness as an ordering mechanism, denoting who is granted political access and participation, must be deconstructed to recognize Eurocentric, heteropatriarchal, ableist, classist foundations (Derrida, 1976). An acknowledgment of such, however, demands a participatory collective to resist disproportionate, privileged access to democratic mechanisms imbued through Whiteness.

To be hospitable toward those who are perceived as non-ideal, educators are invited to imagine schooling as a radically inclusive entity, one which authentically welcomes all visitors and encourages democratic participation. Welcoming those coded as non-ideal challenges administrative and social policies dictating how students will experience schooling and see themselves as political subjects. Unquestionable educational hospitality, while an imperfect, contradictory, and idealistic goal, "emerges from the conflict between what we imagine and what we can do; our commitment to justice and human rights does not, and indeed cannot lie flush with social practices" (Gilbert, 2014, p. 93). This dissonance between social hegemony and administrative policies forces educators to evaluate their roles as citizens within a larger society, as actors of democracy, and as culpable for educational violences. If we are to be authentically welcoming of all, a radically queer acceptance "demands that we accept what is not yet intelligible" (Gilbert, 2014, p. 84) before inquiring if or how they fit into White democratic ideals.

This hospitality, this welcoming, is not only a matter of warmly appreciating, respecting, or accepting all but, as Bettina Love (2019, p. 9, 54) reminds us, that "to want freedom is to welcome struggle". As co-conspirators, we welcome the necessary struggle involved in moving beyond merely surviving to joyful thriving in and out of educational institutions and especially in public political fora. We want to reiterate that anti-racist teaching "is not just about acknowledging that racism exists but about consciously committing to the struggle of fighting for racial justice". If youth are not granted access to the machinery of democracy, then this college struggle for racial justice will be restricted due to their lack of involvement in ongoing activist projects. We want to reemphasize that the historical and long-lasting impacts of democratically supported Whiteness have not only sculpted education in its image but continue to violently shape the material bodies and minds of students, teachers, administrators, and communities through political exclusion. Any attempt to resist these institutionalized racisms, classisms, sexism, ableisms, and heterosexisms must be prior to words, policies, and actions—it should originate in ourselves and become manifest in our material bodies.

At its core, anti-racist and abolitionist education is about "the freedom to create your reality, where uplifting humanity is at the center of all decisions" as much as it is about "tearing down old structures and ways of thinking" (Love, 2019, pp. 88–89). These notions of critical deconstruction, freedom to (re)create, and ability to engage politically are embodied possibilities for transformation of ourselves, others, and society. Transformation at any level is often a slow, arduous, and iterative process in liberatory projects. We must be critical and hospitable as we work to fundamentally reposition our educational institutions as sites of resistance, homeplaces, and homespaces counter to Whiteness, anti-democratic frameworks, and affective violence.

References

- Abowitz, K.K., Harnish, J., 2006. Contemporary discourses of citizenship. *Rev. Educ. Res.* 76 (4), 653–690.
- Ahmed, S., 2006. Orientations: toward a queer phenomenology. *GLQ A J. Lesbian Gay Stud.* 12 (4), 543–574. <https://doi.org/10.1215/10642684-2006-002>.
- Ahmed, S., 2007. A phenomenology of whiteness. *Fem. Theor.* 8 (2), 149–168. <https://doi.org/10.1177/1464700107078139>.
- Annamma, S.A., Jackson, D.D., Morrison, D., 2017. Conceptualizing color-evasiveness: using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethn. Educ.* 20 (2), 147–162. <https://doi.org/10.1080/13613324.2016.1248837>.
- Apple, M.W., 2006. *Educating the 'Right' Way: Markets, Standards, God, and Inequality*. Routledge, New York, NY. <https://doi.org/10.4324/9781315022949>.
- Arendt, H., 2013. *The Human Condition*. University of Chicago Press, Chicago, IL.
- Biesta, G., 2009. Sporadic democracy: education, democracy, and the question of inclusion. In: Katz, M.S., Verducci, S., Biesta, G., Education (Eds.), *Democracy, and the Moral Life*. Springer, New York, NY, pp. 101–112. https://doi.org/10.1007/978-1-4020-8626-7_8.
- Biesta, G., 2016. *Beyond Learning: Democratic Education for a Human Future*. Routledge, New York.
- Blevins, B., Bauml, M., Scholten, N., Smith, V.D., LeCompte, K.N., Magill, K.R., 2020. Using inquiry to promote democratic citizenship among young adolescents during summer civics camps. *Citizen. Teach. Learn.* 15 (3), 271–295.
- Blount, J.M., 2000. 'Spinsters, bachelors, and other gender transgressors in school employment', 1850–1990. *Rev. Educ. Res.* 70 (1), 83–101. <https://doi.org/10.3102/00346543070001083>.
- Bobo, J., Hudley, C., Michel, C., 2004. *The Black Studies Reader*. Psychology Press. <https://cpb-us-e1.wpmucdn.com/blogs.uoregon.edu/dist/8/11384/files/2020/01/Mueller-Ella-Baker-and-origins-of-participatory-democracy.pdf>.
- Bonilla-Silva, E., 2012. The invisible weight of whiteness: the racial grammar of everyday life in contemporary America. *Ethn. Racial Stud.* 35 (2), 173–194. <https://doi.org/10.1080/01419870.2011.613997>.
- Boulton, C., 2016. Black identities inside advertising: race inequality, code switching, and stereotype threat. *Howard J. Commun.* 27 (2), 130–144. <https://doi.org/10.1080/10646175.2016.1148646>.
- Bourdieu, P., 1971. Systems of education and systems of thought. In: Young, M.F.D. (Ed.), *Knowledge and Control: New Directions for the Sociology of Education*. Methuen, London, pp. 189–207.
- Brandzel, A.L., 2016. *Against Citizenship: The Violence of the Normative*. University of Illinois Press.
- Brodin, K., 1998. *How Jews Became White Folks and what that Says About Race in America*. Rutgers University Press, New Brunswick, NJ.
- Carter Andrews, D.J., Brown, T., Castro, E., Id-Deen, E., 2019. The impossibility of being 'perfect and White': black girls' racialized and gendered schooling experience. *Am. Educ. Res. J.* 56 (6), 2531–2572. <https://doi.org/10.3102/0002831219849392>.

- Cohen, C.J., 1997. Punks, bulldaggers, and welfare queens: the radical potential of queer politics? *GLQ: A Journal of Lesbian Gay Studies*. 3 (4), 437–465.
- Cohen, J., 1989. Deliberation and democratic legitimacy. In: Hamlin, A., Pettit, P. (Eds.), *The Good Polity: Normative Analysis of the State*. Basil Blackwell, Oxford.
- Collins, P.H., 2004. *Black Sexual Politics: African Americans, Gender, and the New Racism*. Routledge, New York. <https://doi.org/10.4324/978020309506>.
- De Lissovoy, N., 2015. *Education and Emancipation in the Neoliberal Era: Being, Teaching, and Power*. Springer, New York. <https://doi.org/10.1057/9781137375315>.
- Delpit, L., 1988. The silenced dialogue: power and pedagogy in educating other people's children. *Harv. Educ. Rev.* 58 (3), 280–299. <https://doi.org/10.17763/haer.58.3.c43481778r528gw4>.
- Derrida, J., 1976. The supplement of copula: philosophy before linguistics. *Ga. Rev.* 30 (3), 527–564.
- Fishkin, J.S., 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. Yale University Press, New Haven, CT.
- Foucault, M., & Faubion, J. D. 2000. *Power: The Essential Works of Michel Foucault 1954–1984*. Vol. 3. Three, ed. J. Faubion, trans. R. Hurley et al. London: Allen Lane.
- Freire, P., 1993. *Pedagogy of the City*. Burns & Oates, London.
- Gilbert, J., 2014. *Sexuality in school: The limits of education*. U of Minnesota Press.
- Grady, J., Marquez, R., McLaren, P., 2012. A critique of neoliberalism with fierceness: queer youth of color creating dialogues of resistance. *J. Homosex.* 59 (7), 982–1004. <https://doi.org/10.1080/00918369.2012.699839>.
- Grant, C.A., Brown, K.D., Brown, A.L., 2015. *Black Intellectual Thought in Education: The Missing Traditions of Anna Julia Cooper, Carter G. Woodson, and Alain Leroy Locke*. Routledge, New York. <https://doi.org/10.4324/9780203081303>.
- Grimes, K. M., 2017. Breaking the body of Christ: the sacraments of initiation in a habitat of white supremacy. *political theology*, 18(1), 22–43.
- Gutmann, A., & Thompson, D., 2004. *Why deliberative democracy?*. Princeton University Press.
- Habermas, J., 2015. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. John Wiley & Sons, Hoboken, NJ.
- Harber, C., 2004. *Schooling as Violence: How Schools Harm Pupils and Societies*. Routledge, New York. <https://doi.org/10.4324/9780203488423>.
- Haugeland, J., 2013. *Dasein Disclosed*. Harvard University Press, Cambridge. <https://doi.org/10.4159/harvard.9780674074590>.
- Hobbel, N., Chapman, T.K., 2009. Beyond the sole category of race: using a CRT intersectional framework to map identity projects. *J. Curriculum Theor. (JCT)* 25 (2), 76–89.
- hooks, b., 1990. *Yearning: Race, Gender and Cultural Politics*. South End Press, Boston, MA.
- Hughey, M.W., 2010. The (dis)similarities of white racial identities: the conceptual framework of 'hegemonic whiteness'. *Ethn. Racial Stud.* 33 (8), 1289–1309. <https://doi.org/10.1080/01419870903125069>.
- Hughey, M.W., Byrd, W.C., 2013. The souls of white folk beyond formation and structure: bound to identity. *Ethn. Racial Stud.* 36 (6), 974–981. <https://doi.org/10.1080/01419870.2013.753153>.
- Ingersoll, R.M., Merrill, E., Stuckey, D., Collins, G., 2018. *Seven Trends: The Transformation of the Teaching Force—Updated October 2018* (CPRE Research Reports).
- Jaycox, M.P., 2017. Black Lives Matter and Catholic whiteness: a tale of two performances. *Horizons* 44 (2), 306–341. <https://doi.org/10.1017/hor.2017.121>.
- Jennings, M.E., Lynn, M., 2005. The house that race built: critical pedagogy, African-American education, and the re-conceptualization of a critical race pedagogy. *Educ. Found* 19, 15–32.
- Johnson, E.P. (Ed.), 2016. *No Tea, No Shade: New Writings in Black Queer Studies*. Duke University Press, Durham, NC. <https://doi.org/10.1515/9780822373711>.
- Johnson, L., Morris, P., 2010. Towards a framework for critical citizenship education. *Curric. J.* 21 (1), 77–96.
- Encyclopedia of Curriculum Studies. In: Kridel, C. (Ed.), 2010, vol. 1. Sage Publications, Thousand Oaks, CA. <https://doi.org/10.4135/9781412958806>.
- Ladson-Billings, G., 2018. The social funding of race: the role of schooling. *Peabody J. Educ.* 93 (1), 90–105. <https://doi.org/10.1080/0161956x.2017.1403182>.
- Landmore, H., 2017. Deliberative democracy as open, not (just) representative democracy. *Daedalus* 146 (3), 51–63. https://doi.org/10.1162/daed_a_00446.
- Lewis, A.E., 2004. What group? Studying whites and whiteness in the era of 'color-blindness'. *Socio. Theor.* 22 (4), 623–646. <https://doi.org/10.1111/j.0735-2751.2004.00237.x>.
- Love, B.L., 2019. *We Want to Do More Than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon Press, Boston, MA.
- Markowitz, P.L., 2015. Undocumented no more: the power of state citizenship. *Stanford Law Rev.* 67, 869–915.
- Morgan, M., 2021. *Understanding the Global Rise of Authoritarianism*. Freeman Spogli Institute for International Studies, Stanford University. <https://fsi.stanford.edu/news/understanding-global-rise-authoritarianism>.
- Nash, R., 1990. Bourdieu on education and social and cultural reproduction. *Br. J. Sociol. Educ.* 11 (4), 431–447. <https://doi.org/10.1080/0142569900110405>.
- Ogbu, J.U., 2004. Collective identity and the burden of "acting White" in Black history, community, and education. *Urban Rev.* 36 (1), 1–35. <https://doi.org/10.1023/b:urre.0000042734.83194.f6>.
- Omi, M., Winant, H., 2014. *Racial Formation in the United States: From the 1960s to the 1990s*, third ed. Routledge, New York. <https://doi.org/10.4324/9780203076804>.
- Pfeil, M.R., 2007. The transformative power of the periphery: can a white US Catholic opt for the poor? In: Cassidy, L.M., Mikulich, A. (Eds.), *Interrupting White Privilege: Catholic Theologians Break the Silence*. Orbis Books, Maryknoll, NY, pp. 127–146.
- Pinar, W.F., 1998. *Queer Theory in Education*. Taylor & Francis, Nahwah, NJ.
- Ransby, B., 2003. *Ella Baker and the Black Freedom Movement: A Radical Democratic Vision*. The University of North Carolina Press, Chapel Hill, NC.
- Rawlings, V., 2019. 'It's not bullying', 'It's just a joke': teacher and student discursive manoeuvres around gendered violence. *Br. Educ. Res. J.* 45 (4), 698–716. <https://doi.org/10.1002/berj.3521>.
- Rosiek, J., 2017. Agential realism and educational ethnography: guidance for application from Karen Barad's new materialism and Charles Sanders Peirce's material semiotics. In: Beach, D., Marques da Silva, D.S., Bagley, C. (Eds.), *The Wiley Handbook of Ethnography and Education*. Wiley, London, p. 403. <https://doi.org/10.1002/9781118933732.ch17>.
- Salmi, J., 2000. *Violence, Democracy and Education: An Analytical Framework* (LCSHD Paper Series).
- Slaby, J., 2017. More than a feeling: affect as radical situatedness. *Midwest Stud. Philos.* 41, 7–26. <https://doi.org/10.1111/misp.12076>.
- Stuteville, R., Johnson, H.J., 2016. Citizenship education in the United States: perspective reflected in state education standards. *Adm. Issue. J.* 6 (1), 99–117.
- Tsing, A.L., 2015. *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton University Press, Boston, MA. <https://doi.org/10.1515/9781400873548>.
- Vaught, S.E., 2014. Structural liberalism and anti-bullying legislation. *Equity Excell. Educ.* 47 (2), 152–166. <https://doi.org/10.1080/10665684.2014.900393>.
- Walton, E., 2018. Habits of whiteness: how racial domination persists in multiethnic neighborhoods. *Sociol. Race Ethnic.* 7 (1), 1–15. <https://doi.org/10.1177/2332649218815238>.
- Watkins, W.H., 2001. *The White Architects of Black Education: Ideology and Power in America, 1865–1954*. Teachers College Press, New York.
- Westheimer, J., Kahne, J., 2002. Education for action: preparing youth for participatory democracy. In: Hayduk, R., Mattson, K. (Eds.), *Democracy's Moment: Reforming the American Political System for the 21st Century*. Rowman & Littlefield, pp. 91–107.
- Westheimer, J., Kahne, J., 2004. What kind of citizen? The politics of educating for democracy. *Am. Educ. Res. J.* 41 (2), 237–269.
- Winfield, A.G., 2007. *Eugenics and Education in America: Institutionalized Racism and the Implications of History, Ideology, and Memory*. Peter Lang, New York.
- Wozolek, B., 2015. Schooling racialized bodies: curriculum at the intersection of visibility and absence. *Int. J. Curric. Soc. Just.* 1 (1), 7–17.
- Wozolek, B., 2018. Gaslighting queerness: schooling as a place of violent assemblages. *J. LGBT Youth* 15 (4), 319–338. <https://doi.org/10.1080/19361653.2018.1484839>.
- Wozolek, B., 2020. Hidden curriculum of violence: affect, power, and policing the body. *Educ. Stud.* 56 (3), 269–285. <https://doi.org/10.1080/00131946.2020.1745808>.
- Wun, C., 2016. Unaccounted foundations: black girls, anti-black racism, and punishment in schools. *Crit. Sociol.* 42, 737–750. <https://doi.org/10.1177/0896920514560444>.
- Young, I.M., 2000. *Inclusion and Democracy*. Oxford University Press. <https://doi.org/10.1093/cdj/37.1.113>.

Work and learning: power, politics and approaches

Hongxia Shan^a, Hua Que^b, and Yao Xiao^c, ^aThe Faculty of Education, University of British Columbia, Vancouver, BC., Canada; ^bFaculty of Education, Memorial University of Newfoundland, St. John's, NL, Canada; and ^cThe Faculty of Education and The Social Justice Institute, University of British Columbia, Vancouver, BC., Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	236
The changing world of work	236
Studies of work and learning: three main traditions	237
The conventional approaches to work and learning	237
Critical approaches to work and learning	238
The practice-based approaches to work and learning	240
Discussion and conclusion	240
References	241

Introduction

Work and learning lie at the center of fields, such as workplace learning, vocational education, human resources management, organizational learning, industrial relations, and labor education. While work is commonly associated with waged employment, in this chapter, we address work broadly as activities that are socially organized to achieve goals that are work related. The definition of learning varies widely across the literature. When considering the unit of analysis for learning, i.e. the analytical focus of researchers, often the projected sites of intervention in studies of work and learning, three major approaches can be identified. The first is the conventional approach that focuses on the enhancement of individual skills and competency. The second is the critical approaches that challenge social hierarchies and structural injustice through the promotion of personal empowerment, conscientization, and solidarity. The third is the practice-based approaches that seek to understand learning as the effects of sociocultural and sociomaterial practices. This chapter provides an overview of the major approaches to work and learning, giving special attention to the trends of development in each approach. It starts with an overview of the changing context of work. It then introduces these approaches, with attention to their theoretical premises, units of analysis as well as the goals they ascribe to education and learning. The paper ends with a recap of this chapter and a discussion of the development of work and learning research.

The changing world of work

[Modernity] is a paradoxical unity, a unity of disunity: it pours us all into a maelstrom of perpetual disintegration and renewal, of struggle and contradiction, of ambiguity and anguish. To be modern is to be part of a universe in which, as Marx said, "All that is solid melts into air."

Berman (1982, p. 15).

The history of modern societies is largely a history of contradiction driven by capitalist expansion, and it has always involved changes to traditions and traditional forms of culture, economic structure, and social relations. The scale and speed of changes the world has experienced since the 1960s and 1970s, however, are phenomenal. Globally, European occupation of former colonies came to an end as a result of the independence movements in Asia and Africa. Meanwhile, the established capitalist west was struck by a massive economic recession coupled with social unrest. To salvage itself from the social and economic crisis, the western world chose to embark on economic globalization, which marked the commencement of an era of flexible capitalism (Sennett, 1998) or flexible accumulation (Harvey, 1990).

Riding on swift technological development, and facilitated by the ideological spread of neoliberalism, economic globalization is a multifaceted and developing story. It is known for instigating a race to the bottom, with manufacturing and sometimes service jobs outsourced to countries and areas that "need" to be developed and modernized, and that offer lower labor costs and higher tax incentives for foreign investment. It is also associated with the accelerated mobility of people on a global scale, as countries ratchet up efforts to attract skilled workers, which leads to brain drain from developing countries to the developed west. Today, we are experiencing not only the mobility of goods, services, and capital, but also that of people, culture, ideologies, technologies, and other non-human things/beings. Hannam et al. (2006) observed that these multiple and intersecting mobilities "seem to produce a more 'networked' patterning of economic and social life, even for those who have not moved" (p. 2). This is not to say that the world is flat

now. It is rather to suggest that the geopolitical configuration of globalization has been continuously (re)configured. Indeed, the western centric order of globalization has been challenged. One example in this regard is the Indigenous resistance against (neo)colonialism, displacement and deprivation of land (Tuck and Yang, 2012; Smith et al., 2019). It has also been threatened by the rise of emerging economic powers, such as Brazil, Russia, India, and China. The global pandemic of COVID-19 has also underscored the vulnerability of a world on the move. Clearly, while western powers still dominate, borders of all kinds have become permeable, and porous, as much as they are entrenched.

In the midst of all the global changes, the world of work continues to shift. In the economic west, with the decline of the manufacturing sector, countries have seen an expansion of the services sector. In developing countries, the quality of employment remains poor, and vulnerable groups such as women and people with less education are less likely to have access to quality jobs (United Nations, 2020). This structural change in the west has given rise to the illusion that the postindustrial countries have entered an era of knowledge economy. Much earlier, in his forecast for a postindustrial society, Bell (1976) suggested that sciences and technology would facilitate the expansion of the services sector and would lead to an occupational shift toward professional and technical class. Despite this prediction and all the celebration of technological advancement, studies show that services professions are segmented, if not polarized. While occupations premised on scientific and professional knowledge are growing in importance, they remain smaller in the aggregate than other parts of the economy (Brint, 2001; Brown and Lauder, 2006; Popiel, 2017). In other words, if the knowledge economy is a utopia, not everyone has entered it. Also of note, while the value of knowledge economy is rarely debated, the current COVID-19 pandemic, which has significantly slowed down our economic lives, has provided us with a reality check. Among other things, it has certainly taught us to re-evaluate the value of essential work, including the work of cleaners, servers, retail workers, transit workers, health aides, and other workers who keep our lives going and yet whose work has rarely been appreciated.

Given the structural changes, the organization of employment patterns has shifted as well. Postindustrial countries have moved from a system of standardized lifelong employment to a system of boundary-less and protean careers. In place of standard, stable, and lifelong employment, we have seen the normalization of “non-traditional”, “non-standard”, “project-based” and “contract-based” employments. This has to do with the perceived needs for corporations to become more responsive to the demands of the market, which may easily fluctuate. It also has to do with the retreat of the welfare state, which means less social protection for workers. These non-standard employments, particularly those at the bottom of the economy, typically come with “limited social benefits, and statutory entitlements, job insecurity, low wages and high risks of ill-health” (Vosko, 2006, p. 3). Racialized workers, migrants, and women are much more likely to take up positions that are feminized, insecure, low-paying, and precarious. In the middle of all the changes, unions across OECD countries since the 1980s have shrunk in both membership and influence (Wallerstein and Western, 2000; Vachon et al., 2016).

It is impossible not to address the impacts of technological changes when discussing the changing world of work. When taking on the standpoint of capital and the management, it is easy to be drawn to the productive benefits of new technologies. Technological advancements, particularly what Schwab (2016) called the fourth industrial revolution, a digital revolution marked by nanotechnology, quantum computing, and new processes, seems to promise more meaningful work, and perhaps a new and flattened work order (Gee et al., 2019) where workers or knowledge workers may exercise more discretion and control at work. When taking the perspective of labor or Marxist labor process theory, however, questions have long been raised about whether technological change has led to workers’ autonomy at work (e.g., Braverman, 1974; Rifkin, 1995). Instead of flattening workplace hierarchy, much research has shown that technological change does not necessarily bring with it workplace democracy (Beirne and Ramsay, 2019). Instead, it is implicated in the tightening of managerial manipulation despite the use of rather democratic language. Moreover, it also diverts our attention from the kind of deprivation and systematic exploitation and injustice that people in the developing world experience (Gee et al., 2018).

Studies of work and learning: three main traditions

Against the backdrop laid out above, work and learning have attracted increasing attention across fields. Taken together, related research typically focuses on how individuals, groups, and organizations learn, how knowledge is “transferred” and managed, and how educators, organizations, and institutions might better organize, educate, and support learning across settings. Notably, researchers of work and learning are rarely impartial observers. Rather, they conduct research with diverse interests: some to improve pedagogies of teaching; some to enhance productivity; some to contribute to the well-being of individual workers; and others to challenge the capitalist system of exploitation. It follows that the unit of analysis that researchers focus on varies. Based on the unit of analysis or the projected site of intervention, the literature evinces three traditions: the conventional approaches that focus on individual acquisition of skills and competencies; the critical approaches that believe in the need to challenge existing power hierarchies and structural injustice through collective empowerment and consciousness raising; and the practice approaches that see learning as a derivative effect of social practices.

The conventional approaches to work and learning

The conventional approaches are variously informed by behaviorism, cognitivism, and constructivism, and they focus on how individuals may learn to enhance their skills, competencies, employability, and human capital. Behaviorism assumes that learning is

about effecting observable changes in human behaviors. However, observable behaviors alone are not reliable in predicting job performance. As such, attention has been directed to unobservable cognitive activities, such as thinking, reflecting, knowing, and understanding (Belanger, 2011). For example, drawing on organizational psychology and management theory that focuses on organizational effectiveness and individual well-being, Chris Argyris and Donald Schön (1974, 1978) emphasized that human reasoning underlies individual behaviors within organizations. To give an example, when people join an organization with a defensive mode of reasoning, they tend to get into routines that preempt embarrassment or threats, may negatively affect on organizational performance and effectiveness. What is needed is for management to critically examine staff's defensive reasoning and help them take on productive reasoning (Argyris, 2006). Schön (1992) further highlighted the importance of reflection in how professionals solve workplace challenges. He opposed the dominant technical rationality or "the application of scientific theory and technique to the instrumental problems of practice" (p. 30). Instead, he stressed that contexts of application are often characterized by "uncertainty, uniqueness, instability, and value conflict" (p. 42). To deal with uncertainty, reflection-in-action is the key, which allows individuals to think about and criticize their prior understanding of the phenomenon and, perhaps to generate a new understanding of it as well (Schön, 1992).

The theory of constructivism is "an approach to learning that holds that people actively construct or make their own knowledge and that reality is determined by the experiences of the learner" (Elliott et al., 2000, p. 256). A multitude of adult learning theories, such as experiential learning and transformative learning, have been influenced by constructivism. To give an example, Marsick and Watkins (2018), who have contributed to the conceptualization of informal and incidental learning, "focused on how individuals learn by engaging an ongoing, not necessarily sequential, interactive cycle of learning from experience that started with an unfamiliar challenge or problem at work that triggered informal or incidental learning" (pp. 10–11). Their theory then evolved to emphasize not only individual agency in meaning construction, but also the power of social interaction in shaping learning.

In all cases, conventional approaches to work related learning and training focus on individuals' acquisition of skills and competencies so that they fit into the normative expectations at work. In the last few decades, a variety of occupational matrices of skill and competency have been created as the cornerstone of instructional designs, prior learning assessment, foreign qualification recognition, professional education, and occupational certification. A common trend with these matrices is that they have covered not only hard and technical skills and professional specific competency, but also soft skills, and individual dispositions. While this expanded discourse of skill seems to be self-undermining (Payne, 2000), it correlates with, and indeed supports, the shift of industrialized nations toward a service-based economy, which capitalizes on the commodification of personal emotions, attitudes, and ways of presentation.

Notably, these matrices of skills and competencies have reshaped the social and power relations among learners/workers, employers, educators, and administrators (Jackson, 1993). To start with, they seem to be providing employers, the managerial staff, and administrators with "objective" reference points, and quality assurance vis-a-vis a diverse workforce. This dovetails well with the movement toward new managerialism that stresses observable benchmarks for performance, efficiency and effectiveness. Over time, employers and industries have developed more control over what counts as valuable skills and competencies. In contrast, educators' power and discretion in teaching have been circumscribed to the extent that they are expected to deliver prescribed competencies (Hodge, 2016). This type of control has been referred to as controlled vocationalism (Bates et al., 1998, p. 115). Not only does it shift the control of the content away from the educators, but it ties the goal of training and education to the instrumentalist needs of the economy. As a result, educators might be subjected to constant surveillance. They may be pressured to deliver programs in accordance with competency standards. In neoliberal economies, where instructors are faced with increased job insecurity, controlled vocationalism may also mean that "[i]nstructors ... fearfully develop a self-control or self-censorship that steers away from teaching and learning of social justice" (Abendroth, 2014, p. 20).

With all the emphasis on individual skills and competencies, there is also an expectation for individual workers to mold themselves into an entrepreneurial self, i.e., a subject who takes responsibility for becoming successful at work and in life. This ideology has underpinned many adult education or training programs. For instance, in Maitra's study (2018) of career training programs offered by immigrant services organizations in Canada, she found that the programs tend to emphasize the importance of self-sufficiency and self-endeavor for immigrants to become a competitive and productive worker in the labor market. What often escapes our attention is that the discourse of entrepreneurialism is based on "subtle domination and discrimination inducing individuals into internalizing certain discourses of success and economic productivity irrespective of the actual structures of domination" (Maitra, 2018, p. 209). Such programs are not positioned to address systematic social barriers that the learners might be faced with (e.g., Warriner, 2016). In a similar study in the Canadian context, Shan (2015) also observed that when re-training and re-education overlook the social and ideological processes that shape immigrant women's training and career "choices", they would act "as a 'sorting' process" that continuously enters immigrant women into a labor market "stratified by gender, race, ability and other axes of differentiation" (p. 187).

Critical approaches to work and learning

The development of critical perspectives on work and learning, especially since the 1980s, could be understood as a systematic exploration of the relationships between labor, learning, and the knowledge economy in the context of global capitalism (Bratton et al., 2004). Critical researchers believe that maldistribution of economic goods and unequal distribution of power along intersectional lines of social differences, such as gender, race, ability, and sexuality, are the main source of social injustice, conflicts, and violence. Equalizing the distribution of economic goods and recognizing cultural differences are proposed as the major remedies

to advance social justice (Fraser, 2004). Critical studies of work and learning often center analysis of conflicts, power, inequality, agency, and ethics (Malloch et al., 2011). Related research is variously informed by historical and dialectical Marxism (e.g., McLaren and Jandrić, 2018), Freire's pedagogy of the oppressed (1970), Habermas' idea of communicative action (e.g., Welton, 1991, 2014), Gramsci's notions of "hegemony" and "common sense" (Carpenter et al., 2013), critical race theory (e.g., Crenshaw, 1991, 2011; Closson, 2010), and much more. While it is impossible to provide an exhaustive survey of critical approaches, we focus on three new trends of development in the critical tradition.

The first trend could be characterized as "intersectionality at work" – also the title of Tariq and Syed's work (Tariq and Syed, 2017) on South Asian Muslim women's experiences of employment and leadership in the UK, which pointed to a range of challenges (and sometimes opportunities) that the women faced at the intersection of gender, religion, race, and ethnicity in workplaces. Intersectionality thinking has been applied in the study of work and learning across geographical sites, identity spectrums, and workplace pathways/hierarchies – from Black women's entrepreneurial experiences in Canada (Knight, 2016), to disabled women's workplace well-being in the US (Brown and Moloney, 2019); from ethnic minority women's organizational leadership in South Africa (Carrim and Nkomo, 2016), to both men and women employees' work in a prestigious accounting firm in Mexico (Ruiz-Castro and Holvino, 2016). In these works, intersectionality helps researchers grapple with various dimensions of workplace learning, including but not limited to identity work, discrimination, career impediment, and critique of global capitalism. Despite the wide range of application of intersectional thinking, there is no consensus on what constitutes intersectionality. In Patricia Collins' (2015) language, intersectionality work can be considered as a "knowledge project" (p. 1) suited to the analysis of interlocked social inequalities, power differences, and structural loopholes.

The second trend is the critical engagement with transformative learning in the study of work and learning. In general, this engagement concerns the transformative potentials of workers' identification in relation to the multifaceted formation of learning spaces. Along with van Dellen and Cohen-Scali (2015), we argue for a more holistic conception of "workplaces as learning spaces", and promote learning spaces that "matter in an affective, cognitive and social manner" (p. 733) so as to reduce alienation that workers may experience. In this regard, we are pleased to note that spirituality has become a focal point of consideration in workplace learning practices. For example, Groen and Kawalilak (2016) used autobiographical reflections to affirm a deeper, more spiritual and holistic practice of transformative learning beyond the analytic, rational transformation of individual learners/workers – a practice that moved people toward the creation of more relational and caring workplaces in higher education. In another example, Tolliver's study (Tolliver, 2016) of African epistemology and diasporic activism demonstrated how transformative learning infused with workplace spirituality could "provide powerful incentives for employees not only to 'see' social justice issues but also to commit to addressing instances of social injustice within the workplace as well as in the world" (p. 81). Knowing the necessity to challenge structural racism and white supremacy in both formal and informal learning (Brookfield, 2019), we ask how this "spirituality" turn could become explicit antiracism strategies.

The third trend in the development of the critical approaches is for researchers and scholars to confront white supremacy through decolonization. This trend has gathered momentum in particular in the studies of immigrants' work and learning experiences. For example, primarily through an examination of the employment experiences of transnational migrants, Maitra and Guo (2019) defined projects of decolonization in lifelong learning in four dimensions: recognizing the colonial power differentials embedded in knowledge-making; challenging the hegemony of Western knowledge, education, and credential recognition; planning and designing learning curricula based on non-Western knowledge system; decolonizing our minds as lifelong learners, practitioners and policy makers. In another example, Morrice's (2019) drew on Santos' (2016) notion of abyssal thinking and exposed how a West-centric, colonial, restrictive migration-education nexus maintains traditional colonial power divisions, exclusion, and control. She instead called for a global, structural reimagining of educational spaces that recognize "the particularity of subjectivities and struggles of migrants, and others in the global South", and advocated that such struggles should be accompanied by the production of "new forms of emancipatory knowledge" (p. 30).

Of note, the move toward decolonization has recently joined hands with the social and intellectual movement of indigenization. A body of research has emerged that centers Indigenous worldviews, recognizing Indigenous authorities on land stewardship, knowledge systems, and promoting cultural practices beyond the settlers' purview and beyond 'acceptance' or liberal inclusion. For example, affirming Maori authority in Aotearoa New Zealand workplaces, Huygens (2006) insisted that decolonization requires a move from organizational accountability (practices within a particular institution) to constitutional accountability (practices connected to the larger society). Working as Indigenous persons (Plains Cree, Saulteaux, and Mi'kmaq) in the Canadian context of distance adult education, Kovach and Montgomery (2010) pointed out the need for instructors and staff within the organizations to "design decolonizing curricula that engages Indigenous knowledges and upholds oral culture" (p. 27). While there has always been a crucial connection between labor education, union activism, and labor movements (Spencer, 2002), Ng and Huang (2012) worked with labor educators and activists to explore how a critical, decolonial approach to relationship building – through a "pedagogy of solidarity" and the work of "solidarity circle" (pp. 533–535) – could connect labor educators, activists, and movement organizers from Indigenous and racialized communities.

We emphasize that not all narratives *about* or *including* Indigenous people are critical. One vital question remains: how do we ensure the efforts and processes to facilitate Indigenous participation, knowledge, and learning in workplace do not repeat a paradigm where Indigenous learners and voices are included, and yet the larger oppressive structures are left unchallenged? Furthermore, as critical researchers work with Indigenous worldviews, they may necessarily might experience cognitive dissonance. For instance, as Grande (2008) pointed out, "[w]hile it is evident that revolutionary critical theory holds great promise, because it also retains core Western assumptions, it also stands in tension with those central to indigenous pedagogies. Specifically, the radical notion of

'democratization' does not theorize the difference of indigenous sovereignty ... and insofar as the discourse of revolutionary critical pedagogy is informed by Marxist theory, it retains a measure of anthropocentrism that belies indigenous views of land and 'nature' (pp. 234–235). In this context, Grande (2008) asked critical researchers to work with she called "red pedagogy", which is to create a space of reflexive engagement between critical theory and Indigenous worldviews.

The practice-based approaches to work and learning

If the conventional approaches are interested in fostering learning so that individuals keep up with the changing needs of the working world, and the critical approaches are invested in challenging the oppressive social structures, the practice-based approaches focus on mapping social practices that give rise to learning as a derivative effect (Fenwick and Nerland, 2014). The practice-based approaches to learning are informed by what Schatzki (2001) called "the practice turn" and the practice-based learning theories fall largely into two categories: the sociocultural and the sociomaterial. The former include situated learning, community of practices (Lave and Wenger, 1991; Wenger 1998) and cultural historical activity theories (Engeström, 2001). Informed by social cognitivism and social constructivism, situated learning examines learning as a process of social participation in the world. In Lave and Wenger's study of apprenticeship of trades people (1991), learning is understood to be a process of legitimate peripheral participation in a community of practices (CoPs), i.e., a group of people engaged in a common domain of practice, sharing ideas, and looking for solutions to problems (Wenger et al., 2002). The notion of CoP has been widely taken up in managerial and organizational studies that aim to refine practices, facilitate knowledge transfer and promote practice innovation (e.g., Bolisani and Scarso, 2014; Pyrko et al., 2016).

While CoP addresses people's learning experiences in relation to their membership within CoP, it does not attend to the power differences that may impact our access to CoP (Hodkinson and Hodkinson, 2004). Further, with its sole emphasis on the cultural conditioning of learning, it may have occluded the political and economic relations that shape CoP. In these respects, Engeström's (2001) CHAT usefully extends the illustrative power of CoP. CHAT is influenced by Vygotsky's cultural historical studies of psychology, which is a "psychologically relevant application of dialectical and historical materialism" (Vygotsky, 1978, p. 6). It is attentive to work practices as a cultural and historical system pivoted on an object, i.e., a social concern that focuses our attention, engages our efforts, and generates individual and collective actions and interactions. It also directs attention to the interrelations among individual subjects, objects and mediators of learning, in relation to the rules, community, and division of labor (Fenwick et al., 2011). While addressing multiple voices and interests, CHAT is also apt at identifying contradictions in practice as driving forces of change. Engeström et al. (2014), for instance, elaborated on how transformative agency and expansive learning evolve as individuals work through contradiction, raise questions, experiment with solutions, and expand objects and practices.

While sociocultural theories approach practices as activities that are purposeful, ordered, and regular, sociomaterial perspectives, exemplified by complexity theory, actor network theories, and assemblage thinking, help us zoom into practice as it is produced from moment to moment (Fenwick and Nerland, 2014). Developed initially in sociology of sciences and technology, sociomaterial perspectives have been used to study professional learning and professional work that are highly populated with texts, tools, and digital technologies (Fenwick et al., 2012; Reich and Hager, 2014). Typically, sociomaterial research stays away from reifying or naturalizing established sociological explanatory categories or schemes. It is rather interested in following the actors empirically, and hence is sensitive to the emergent sociomaterial organization of work and life, which may defy predictability and stretch human imaginations at all times (Fenwick et al., 2011). From the sociomaterial perspectives, non-human things are no longer treated as the contexts for learning and work or mediators of human action but as agentic actors. Star's notion of boundary object is illustrative in this regard. According to Star (2010), a boundary object is "something people (or, in computer science, other objects and programs) act toward and with. Its materiality derives from action, not from a sense of prefabricated stuff or 'thing'-ness" (p. 603). Boundary objects are known for their interpretive flexibility which arises organically in response to the requirement for information and work (Fenwick et al., 2011). Looking to the formation of sociomaterial relations focuses researchers' attention on the exterior relationality among heterogeneous entities and on the provisional sociomaterial ordering that may give rise to opportunistic, innovative, and creative forms of learning and engagement (Johnsson and Boud, 2010).

Of note, issues of social justice and power differences, especially those associated with gender, race, class, and other relations of differences, are not typically the focus of practice-based studies of work and learning. What is more, in most of the sociomaterial studies, power is not considered the cause of action, but rather the consequence of it (Latour, 1986). Latour's translation model of power, for instance, shows that power is not possessed by particular groups. Nor is it what one does to another. It is rather something that is worked on or up within a chain of actions where each actor shapes it as it circulates. In this view, power is produced "by enrolling actors in a given political and social scheme" (Latour, 1986, p. 264). In other words, power is modified as humans and non-human things hang together and perform themselves into being. While not all practice-based research has a social agenda, we believe that there is potential to mobilize this performative view of power for a more critical and nuanced understanding of work and learning.

Discussion and conclusion

Learning has accrued more attention today given the changing context of work since the 1970s. In the era of neoliberalism, learning seems to have become a lifeline for people to keep up with all the changes in the working world. Correspondingly, there has been an

increasing number of studies that focus on the learning needs and experiences of different groups, often marginalized groups in relation to their work experiences. This chapter reviews the scholarship of work and learning across fields based on the unit of analysis, or the analytical focus of researchers. It points to three traditions of scholarship. The first, also the dominant one, is the conventional approach that focus on individuals' development of skills, competencies, and dispositions. In this tradition, individuals are treated as generic learners and workers who could be better positioned in the labor market through continuous investment in learning. While there is no doubt that learning is instrumental to improving individuals' employability in the labor market, a few questions need to be asked. The first is whether all individuals are uninhibited entrepreneurial learners and workers, regardless of backgrounds. The second is whose interests and perspectives are reflected in the kind of skills and competency that are valued within vocations. The third is how controlled vocationalism (Bates et al., 1998) may limit the work of educators and, associated with this, how we might be able to expand the goal of education beyond instrumentalism.

The second tradition is the critical perspectives that direct attention to how the realities of work and learning are inseparable from unequal power relations and systems of domination. Related research suggests that struggles for good or meaningful workplace learning are indeed struggles for liberation. Such an understanding is based on various approaches: systematic exploration of learners/workers' uneven positionalities linked to their intersecting subjectivities and social differences; reflections on how holistic personal and relational transformative learning could lead to more equitable, caring workplaces; and practices and visions of decolonization that not only recognize the diverse knowledge systems of indigenous people, but also require further critical studies of work, learning, and justice from non-Eurocentric, non-anthropocentric perspectives.

The third is the practice-based approaches to learning, which is a relatively new trend that have gathered momentum in the last twenty years. Research in the former two traditions suggest that learning, training, and education are often positioned to fulfill two objectives: one is to prepare men and women for their work roles, and the other is to promote independent and critical thinking of individuals as social beings (Spencer and Kelly, 2013; Morrice et al., 2018). The practice-based approaches do not necessarily commit to either objective. Theoretically, these approaches are good at tracing the sociocultural and sociomaterial relations that give rise to human and organizational learning. In a world that is marked with multiple mobilities, practice-based approaches provide a multitude of heuristics that can help trace, pragmatically, how the working world is continuously configured and reconfigured.

By no means are the research traditions presented in this chapter exhaustive of all approaches taken to the study of work and learning. What this entry aims to accomplish is to lay out, in broad strokes, the main contours of research concerning work and learning, giving special attention to the unit of analysis or the imagined site of intervention, and the intended goals of education and learning. Of note, the three traditions of work introduced are not necessarily distinct from one another. For instance, there is no prevention for conventional approaches can be used to help workers inquire into the nature of their working conditions, especially the exploitative social relations that workers need to navigate. Meanwhile, critical researchers should not be surprised if workers do not consider transformative learning or a shift of consciousness good learning at all times. For some people, acquisition of particular instrumental skills might be of more value (at particular times) (Newman, 2012). As well, while social justice and equity are not necessarily the central concern of practice-based research, there is potential for researchers to mobilize practice-based thinking to address the changing forms, means, flows and politics of learning in a world where power works in complex ways.

It needs to be noted that this entry draws on research and literature based mostly on the western context. It is assumed that the issues we related about work and learning in this chapter might also be present in other parts of the world. We assume that this chapter might provide a heuristic for researchers to examine work related issues in other contexts. At the same time, we acknowledge that there is also space for researchers in the west to learn from other places. For alternative work to emerge beyond the existing epistemological horizons, researchers may consider comparing not only issues of work and learning, but also methodologies, and theoretical approaches mobilized in related research between the postindustrial west and the rest of the world.

References

- Abendroth, M., 2014. Adult Education at Risk: Fronts of Resistance to Neoliberalism [online]. Available at: <https://newprairiepress.org/aerc/2014/papers/1/>. (Accessed 15 June 2020).
- Argyris, C., 2006. *Reasons and Rationalizations the Limits to Organizational Knowledge*. Oxford University Press, Oxford.
- Argyris, C., Schön, D., 1974. *Theory in Practice: Increasing Professional Effectiveness*. Jossey-Bass, San Francisco.
- Argyris, C., Schön, D., 1978. *Organizational Learning: A Theory of Action Perspective*. Addison-Wesley, Reading.
- Bates, I., Bloomer, M., Hodkinson, P., Yeomans, D., 1998. Progressivism and the GNVQ: context, ideology and practice. *J. Educ. Work* 11 (2), 109–125.
- Beirne, M., Ramsay, H., 2019. *Information Technology and Workplace Democracy*, second ed. Routledge, London.
- Belanger, P., 2011. *Theories in Adult Learning and Education*. Barbara Budrich Publishers, Opladen and Farmington Hills.
- Bell, D., 1976. *The Coming of the Post-industrial Society: A Venture in Social Forecasting*. Basic Books, New York.
- Berman, M., 1982. *All that Is Solid Melts into Air: The Experience of Modernity*. Penguin Books, New York.
- Bolisani, E., Scarso, E., 2014. The place of communities of practice in knowledge management studies: a critical review. *J. Knowl. Manag.* 18 (2), 366–381.
- Bratton, J., Mills, J., Sawchuk, P., 2004. *Workplace Learning: A Critical Introduction*. Garamond Press, Auroa.
- Braverman, H., 1974. *Labour and Monopoly Capitalism: The Degradation of Work in the 20th Century*. Monthly Review Press, New York.
- Brint, S., 2001. Professions and the "knowledge economy": rethinking the theory of post-industrial Society. *Curr. Sociol.* 49 (4), 101–132.
- Brookfield, S.D., 2019. *Teaching Race: How to Help Students Unmask and Challenge Racism*. Jossey-Bass, San Francisco.
- Brown, P., Lauder, H., 2006. Globalization, knowledge and the myth of the magnet economy. *Glob. Soc. Educ.* 4 (1), 25–57.

- Brown, R.L., Moloney, M.E., 2019. Intersectionality, work, and well-being: the effects of gender and disability. *Gen. Soc.* 33 (1), 94–122.
- Carpenter, S., Ritchie, G., Mojab, S., 2013. The dialectics of praxis. *Social. Stud.* 9 (1), 1–17.
- Carrim, N.M.H., Nkomo, S.M., 2016. Wedding intersectionality theory and identity work in organizations: South African Indian women negotiating managerial identity. *Gen. Work. Organ.* 23 (3), 261–277.
- Closson, R.B., 2010. Critical race theory and adult education. *Adult Educ. Q.* 60 (3), 261–283.
- Collins, P.H., 2015. Intersectionality's definitional dilemmas. *Annu. Rev. Sociol.* 41 (1), 1–20.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299.
- Crenshaw, K., 2011. Twenty years of critical race theory: looking back to move forward. *Conn. Law Rev.* 43 (5), 1253–1354.
- Elliott, S.N., Kratochwill, T.R., Cook, J.L., 2000. *Educational Psychology: Effective Teaching, Effective Learning*. McGraw-Hill College, Boston.
- Engeström, Y., 2001. Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14 (1), 133–156.
- Engeström, Y., Sannino, A., Virkkunen, J., 2014. On the methodological demands of formative interventions. *Mind Cult. Activ.* 21 (2), 118–128.
- Fenwick, T., Edwards, R., Sawchuk, P., 2011. *Emerging Approaches to Educational Research: Tracing the Socio-Material*. Routledge, London.
- Fenwick, T., Nerland, M., 2014. *Reconceptualising Professional Learning: Sociomaterial Knowledges, Practices and Responsibilities*. Routledge, London.
- Fenwick, T., Nerland, M., Jensen, K., 2012. Sociomaterial approaches to conceptualising professional learning and practice. *J. Educ. Work* 25 (1), 1–13.
- Fraser, N., 2004. To interpret the world and to change it: an interview with Nancy Fraser. *Signs* 29 (4), 1105–1124.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Herder and Herder, New York.
- Gee, J.P., Hull, G., Iankasheer, C., 2018. *The New Work Order: Behind the Language of the New Capitalism*, second ed. Routledge, New York.
- Grande, S., 2008. Red pedagogy: the un-methodology. In: Denzin, N.K., Lincoln, Y.S., Smith, L.T. (Eds.), *Handbook of Critical and Indigenous Methodologies*. Sage, Thousand Oaks, pp. 233–254.
- Groen, J., Kawallak, C., 2016. Creating spaces for transformative learning in the workplace. *N. Dir. Adult Cont. Educ.* 2016 (152), 61–71.
- Hannam, K., Sheller, M., Urry, J., 2006. Editorial: mobilities, immobilities and moorings. *Mobilities* 1 (1), 1–22.
- Harvey, D., 1990. *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*. Blackwell Publishers Inc, Cambridge.
- Hodge, S., 2016. Alienating curriculum work in Australian vocational education and training. *Crit. Stud. Educ.* 57 (2), 143–159.
- Hodkinson, H., Hodkinson, P., 2004. Rethinking the concept of community of practice in relation to schoolteachers' workplace learning. *Int. J. Train. Dev.* 8 (1), 21–31.
- Huygens, I., 2006. Discourses for decolonization: affirming Maori authority in New Zealand workplaces. *J. Community Appl. Soc. Psychol.* 16 (5), 363–378.
- Jackson, N., 1993. If competence is the answer, what is the question? *Aust. N. Z. J. Vocat. Educ. Res.* 1 (1), 46–66.
- Johnsson, M., Boud, D., 2010. Toward an emergent view of learning work. *Int. J. Lifelong Learn.* 29 (3), 355–368.
- Knight, M., 2016. Race-ing, classing and gendering racialized women's participation in entrepreneurship. *Gen. Work. Organ.* 23 (3), 310–327.
- Kovach, M., Montgomery, H., 2010. What kind of learning? For what purpose? Reflections on a critical adult education approach to online Social Work and Education courses serving Indigenous distance learners. *Crit. Soc. Work* 11 (1), 27–41.
- Latour, B., 1986. The powers of association. In: Law, J. (Ed.), *Power, Action, Belief: A New Sociology of Knowledge?* Routledge & Kegan Paul, London, pp. 264–280.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Maitra, S., 2018. Crafting 'enterprising' workers through career training programs among Canada's South Asian professional immigrants. *Stud. Educ. Adults* 49 (2), 196–213.
- Maitra, S., Guo, S., 2019. Theorising decolonisation in the context of lifelong learning and transnational migration: anti-colonial and anti-racist perspectives. *Int. J. Lifelong Educ.* 38 (1), 5–19.
- Malloch, M., Cairns, L., Evans, K., O'Connor, B. (Eds.), 2011. *The SAGE Handbook of Workplace Learning*. Sage, London.
- Marsick, V.J., Watkins, K.E., 2018. Introduction to the special issue: an update on informal and incidental learning theory. *N. Dir. Adult Cont. Educ.* 2018 (159), 9–19.
- McLaren, P., Jandrić, P., 2018. Paulo Freire and liberation theology: the Christian consciousness of critical pedagogy. *Vierteljahrsschr. Wiss. Pädagogik* 94 (2), 246–264.
- Morrice, L., 2019. Abyssal lines and cartographies of exclusion in migration and education: towards a reimagining. *Int. J. Lifelong Educ.* 38 (1), 20–33.
- Morrice, L., Shan, H., Sprung, A., 2018. Migration, adult education and learning. *Stud. Educ. Adults* 49 (2), 129–135.
- Ng, W., Huang, J., 2012. Pedagogy of solidarity: educating for an interracial working class movement. *J. Workplace Learn.* 24 (7/8), 528–537.
- Newman, M., 2012. Calling transformative learning into question: some mutinous thoughts. *Adult Educ. Q.* 62 (1), 36–55.
- Payne, J., 2000. The unbearable lightness of skill: the changing meaning of skill in UK policy discourses and some implications for education and training. *J. Educ. Pol.* 15 (3), 353–369.
- Popiel, P., 2017. "Boundaryless" in the creative economy: assessing freelancing on Upwork. *Crit. Stud. Mass Commun.* 34 (3), 220–233.
- Pyrko, I., Dorfler, V., Eden, C., 2016. Thinking together: what makes communities of practice work? *Hum. Relat.* 70 (4), 389–409.
- Reich, A., Hager, A., 2014. Problematising practice, learning and change: practice-theory perspectives on professional learning. *J. Workplace Learn.* 26 (6/7), 418–431.
- Rifkin, J., 1995. *The End of Work: The Decline of the Global Labor Force and the Dawn of the Post-market Era*. G. P. Putnam's Sons, New York.
- Ruiz-Castro, M., Holvino, E., 2016. Applying intersectionality in organizations: inequality markers, cultural scripts and advancement practices in a professional service firm. *Gen. Work. Organ.* 23 (3), 328–347.
- Santos, B.D.S., 2016. *Epistemologies of the South: Justice against Epistemicide*. Routledge, London.
- Schatzki, T., 2001. Introduction: practice theory. In: Cetina, K.K., Schatzki, T.R., Savigny, E. (Eds.), *The Practice Turn in Contemporary Theory*. Routledge, London, pp. 1–14.
- Schön, D., 1992. *The Reflective Practitioner*. Routledge, London.
- Schwab, K., 2016. *The Fourth Industrial Revolution*. World Economic Forum, Geneva.
- Sennett, R., 1998. *The Corrosion of Character: The Personal Consequences of Work in the New Capitalism*. Norton, New York.
- Shan, H., 2015. Complicating the entrepreneurial self: professional Chinese immigrant women negotiating occupations in Canada. *Glob. Soc. Educ.* 13 (2), 177–193.
- Spencer, B. (Ed.), 2002. *Unions and Learning in a Global Economy: International and Comparative Perspectives*. Thompson Educational Publishing, Toronto.
- Spencer, B., Kelly, B., 2013. *Work and Learning: An Introduction*. Thompson Educational Publishing, Toronto.
- Smith, L.T., Tuck, E., Yang, W., 2019. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, New York and London.
- Star, S.L., 2010. This is not a boundary object: reflection on the origin of a concept. *Sci. Technol. Hum. Val.* 35 (5), 601–617.
- Tariq, M., Syed, J., 2017. Intersectionality at work: South Asian Muslim women's experiences of employment and leadership in the United Kingdom. *Sex Roles* 77 (7–8), 510–522.
- Tolliver, D., 2016. Spirit-ness at work: connections between workplace spirituality, transformative learning, and social justice. *N. Dir. Adult Cont. Educ.* 2016 (152), 73–84.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- United Nations, 2020. *World Economic Situation and Prospects* [online]. Available at: https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2020_Full-Report.pdf. (Accessed 25 April 2021).
- van Dellen, T., Cohen-Scali, V., 2015. The transformative potential of workplace learning: construction of identity in learning spaces. *Int. Rev. Educ.* 61 (6), 725–734.
- Vachon, T.E., Wallace, M., Hyde, A., 2016. Union decline in a neoliberal age: globalization, financialization, European Integration, and union density in 18 affluent democracies. *Socius* 2, 1–22.
- Vosko, L.F. (Ed.), 2006. *Precarious Employment: Understanding Labour Market Insecurity in Canada*. McGill-Queen's University Press, Montreal and Kingston.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge.
- Wallerstein, M., Western, B., 2000. Unions in decline? What has changed and why. *Annu. Rev. Polit. Sci.* 3 (1), 355–377.

- Warriner, D.S., 2016. 'Here, without English, you are dead': ideologies of language and discourses of neoliberalism in adult English language learning. *J. Multiling. Multicult. Dev.* 37 (5), 495–508.
- Welton, M., 1991. Shaking the foundations: the critical turn in adult education theory. *Can. J. Stud. Adult Educ.* 5, 21–42.
- Welton, M., 2014. The educator needs to be educated: reflections on the political pedagogy of Marx, Lenin and Habermas. *Int. J. Lifelong Educ.* 33 (5), 641–656.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, Cambridge.
- Wenger, E., McDermott, R., Snyder, W.M., 2002. *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Harvard Business School Press, Boston.

Minoritized individuals and knowledge-economy

Chazz Robinson and Raechele Pope, University at Buffalo, Amherst, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	244
Knowledge economy through a racial capital lens	245
Human capital in the knowledge economy	245
The educated and minoritized communities	246
Low income/working class minoritized individuals and knowledge economy	246
Racially minoritized individuals and the knowledge economy	247
Transforming the knowledge economy	248
Theories for change	248
Conclusion	249
References	249

Introduction

College attendance and graduation have many benefits and are linked to higher income, more civic engagement, and enhanced health (Terenzini et al., 2001). The assumption is that the knowledge attained through formal education will lead to more opportunities and a better life. This has not always been the case, in fact before the 1970s, the Fordism movement and traditional manufacturing was the basis for many to achieve success and happiness in the United States and Europe economies, while educational attainment was not viewed as important as it is today (Boczy et al., 2020). The idea of educational attainment was prompted by the transition from an industrial society, centered in the Fordism movement, to a more technologically driven and international economy (Esping-Andersen, 1999). This transitional period moved the economy away from traditional industry to more information and communications technology (Hope and Martelli, 2019). Additionally, Hope and Martelli (2019) explain,

the information and communication technology revolution and globalization allowed top managers, CEOs, and entrepreneurs to apply their talents to a much wider pool of resources to reach a substantially larger audience than had been possible in previous generations (p. 237).

This technological revolution created a global expansion where knowledge creation and obtainment became the central source for economic prosperity.

Knowledge being seen as an economic tool has been explored through the knowledge economy framework. The term “knowledge economy” was coined by Peter Drucker, an Austrian writer on management, economics, and organizational theory, in his 1969 book, *The Age of Discontinuity*. Knowledge economy can be defined as “an economic system in which the production of goods and services is based principally on knowledge-intensive activities that contribute to a rapid pace of advancement in technical and scientific innovation as well as accelerated obsolescence” (Powell and Snellman, 2004, p. 199). This positions knowledge as a resource and as a production tool used to boost the economy; therefore, the more knowledge obtained, the more likely one may be of use within a knowledge-based society. Drucker (1992) elaborated on the knowledge economy theory in an essay where he connected economic productivity directly to knowledge, stating there was a significant shift underway to what he called knowledge society. Drucker explained that in this new society, knowledge was a more impactful resource than land, labor, or financial income and envisioned a move from a capitalistic society to a post-capitalist society and then to a knowledge society. Through a Drucker lens, the advancement of knowledge is the primary driver of economic success.

This framing of the knowledge economy is a globalized concept that has spread internationally. Advanced democracies in regions such as Western Europe and Asia-Pacific support the belief that the knowledge economy has created a demand for college educated individuals with additional skillsets (Hope and Martelli, 2019). The global demand for college educated individuals further underscores how educational attainment is central to the knowledge economy and how the knowledge economy has become a global framework that shapes policies. Naidoo (2003) explains how national governments are reforming higher education to adjust for postindustrial societies. The preparation for this new knowledge-based society has led to national government policy and research influencers, such as the World Bank and the Organization for Economic Co-operation and Development (OECD), supporting the knowledge economy framework.

The World Bank, home to 189-member countries, relies on the knowledge economy framework and has used it to shape its initiatives (Weber, 2011). In these initiatives, the World Bank (2010) emphasizes four pillars of the knowledge economy: economic and institutional regime, education and skills, information and communication infrastructure, and innovation systems. Weber (2011) explains, “Education and skills in the World Bank schema ... plays a critical role in innovation system since educational institutions and technology incubators connected to universities are often the sources of new technologies, patents, and novel ideas” (p. 2590). Martens (2007) adds the OECD has, “become one of the most significant players in matters of education through soft

tools" (p. 44). It is important to highlight that both the World Bank and OECD are central to the promotion of the knowledge economy globally and those organizations have been incremental in creating dialogue, research and policies that influence how societies are shaped to view the purpose of education. Positioning knowledge as a tool that can be leveraged by individuals for more economic mobility seems promising, but some critical scholars highlight that race must be central to the conceptualizing of any institutions within the United States context.

Knowledge economy through a racial capital lens

Drucker's assumption that formal education in the US would become the key to navigating in a knowledge economy and the pursuit of education would increase economic mobility was correct overall, however scholars highlight that embedded in the ethos of schools and every other facet of the American system is race (Robinson, 1983). To conceptualize the connections between race and systems within America Cedric Robinson, author of *Black Capitalism*, argued that capitalism is racial in its origin which created the social structures in which Americans operate in today. Benson and Dumas (2021) add, "Under the conditions of racial capitalism, capitalist domination is interrelated and mutually constitutive with White supremacist structures and discourses around the differential valuing—and therefore treatment—of human life" (p. 77). Robinson's account of racial capitalism works well in framing the US educational system that was historically built for those who were white and wealthy (Thelin, 2011). Under this condition education within the knowledge economy must be analyzed through the intersections of race and class or risk leaving minority students out of the educational conversation. Robinson (1983) purported that, "the development, organization, and expansion of capitalist society pursued essentially racial directions, so too did social ideology. As a material force, then, it could be expected that racialism would inevitably permeate the social structures emergent from capitalism" (p. 2). Focusing Robinson's argument on education, means that we must acknowledge educational systems are racial and capitalistic structures, which may challenge the linear view of the educational system that was proposed by Drucker.

Using race as an analytic, prominent scholar Jean Anyon (2014) contended education is not linear at all and educational systems reproduce inequalities for low-income minorities. This differs quite differently from Drucker's view in the sense that education does not equal even opportunity for everyone nor does every person experience education in the same way, and that education is not creating a more equal society. While the knowledge economy can be discussed on a global level due to its wide usage, this article will examine the concept of knowledge economy through the lens of the United States, as a case study, while making careful consideration of the global perspective. Furthermore, this chapter argues that the intersections of race and class highlight how US education within the knowledge economy will not always be a linear process for the marginalized and an equity lens is needed when conceptualizing and utilizing the knowledge economy framework locally and globally.

Human capital in the knowledge economy

To fully analyze the knowledge economy a careful look at the theoretical underpinnings is necessary. The theoretical underpinnings of the knowledge economy highlight several theories; however, human capital theory appears to be driving the main discourse on the knowledge economy within the US and globally, with top organizations like the World Bank and the OECD using it in tandem with the knowledge economy to shape practical uses within society (Takayama, 2013; Weber, 2011). Harvard economist Claudia Goldin (2016) explains that the human capital definition centers on the idea that investments in people, such as education, increase workers' productivity and skillsets. This theory relates to the knowledge economy because it highlights the common assertion that the more educated a person becomes the more, they will be desired for their specific knowledge and increase their human capital. Korres (2008) argued that, "a lack of investment in human capital ... prevents poor countries from catching up with rich ones. Educational attainment and public spending on education are correlated positively to economic growth" (p. 714). This connection directly places economic value on knowledge obtained. The connection between human capital theory and the knowledge economy becomes more evident when exploring the assumptions within the theory of human capital.

Human capital theory assumes that investment in education is necessary to acquire skills and training which, in turn, will increase individual capital (Blundell et al., 1999). Tan (2014) adds,

These knowledges and skills will increase his or her productivity in the workplace. This increased productivity will bring a higher salary to the individual since the wage of a person, in the ideal labor market, is determined by the person's productivity. Therefore, people would invest in education up to the point where the private benefits from education are equal to the private cost (p. 413).

This conceptualization of the knowledge economy is being globalized, with organizations indicating that human capital is accumulated by investing in education so that individuals can compete in the new society and poverty can be ended (World Bank, 2018). While the World Bank has aligned their organizational goals around the knowledge economy and has developed a metric system to examine how countries are meeting their practical goals in the knowledge economy, they acknowledge that before and even more so during COVID season, low-income societies, and gender minoritized people are struggling within the knowledge economy (World Bank, 2018). While the World Bank made references to gender minoritized individuals such as women, missing

were also the experiences of sexual minoritized people. Similar, there have been reports that students in the United States who identify as low income, international, and/or racially minoritized have been most impacted by COVID restrictions (Means and Neisler, 2021). These initial findings could explain why knowledge economy and its theoretical underpinnings to human capital theory both have been criticized for lacking a cultural perspective on how people navigate the knowledge economy (Weber, 2011). Connecting this critique to today's time shows the inconsistencies in the knowledge economy regarding minority groups.

The educated and minoritized communities

Drucker's early assertions that knowledge would be valued by society has found substantial support; however, research has shown that when focusing in on minoritized students many are having drastically different outcomes. Research has shown how educational systems still amplify barriers for some low income racially minoritized students in the US. (Ardoin and Martinez, 2019; Park, 2018). Despite this, the Pew Research Center (2018) report that the Millennial and Gen Z generations with the US are the two most educated generations ever, with Gen Z projected to unseat Millennials as the biggest pursuers of college education at fifty nine percent. Additionally, Gen Z is the most racially diverse of any generation with almost half of them identifying as people of color. Globally, the pursuit of education is on the rise as well, with newer generations seeking education at higher levels than previous generations and overall, the world is more college educated than ever before (Roser and Ortiz-Ospina, 2016). In addition, Gen Z is projected as being more likely to pursue education than the labor force within the US (Pew Research Center, 2018). Evident is that, within the US, these past two generations are much more racially diverse. That means that educational institutions, in the US, are projected to see a higher level of minoritized populations. Educational research demonstrates that minoritized students, such as international, low income, students of color, and gender and sexual minoritized individuals face great obstacles in the educational systems (Ardoin and Martinez, 2019; Collins, 1990; Duran, 2018; Park, 2018; Stewart, 2017). This further emphasizes that minorities are struggling in the knowledge economy.

On the surface, the assumptions of the knowledge economy seem to be working accordingly, with knowledge attainment as higher than it has ever been, yet a more critical look paints a picture of inequity. As previously discussed, the knowledge economy is predicated on knowledge attainment, which means educational institutions are paramount to creating a more educated public. Yet, during the process of schooling, in places like the United States, educational opportunity has been marred with inequities revolving around minoritized populations. Scholars have long researched the topic of educational inequity, for example, Patricia Hill Collins (1990), in her book, *Black Feminist Thought*, discussed how American educational systems are configured on racism, genderism, and classist practices and policies. Additionally, she acknowledged that every educational context in the United States is often situated within a macro-level Eurocentric past that has oppressed people of color, women, and poor individuals. Collins's assertions agree with Robinson's (1983) argument that race has permeated all social structures in the US. Given that the knowledge economy is predicated upon educational institutions those who are marginalized are likely to face barriers navigating the knowledge economy. Identities within the US such as race, class, and gender, may not have the same meaning or impact in other countries, however it should be noted that international students are attending US colleges/universities at large rates, and many international institutions are working diligently to shape their institutions and educational policies to be competitive in national rankings (Carey, 2018). This further emphasizes why developing a more critical perspective on the knowledge economy and understanding minoritized individuals within the knowledge economy is essential on a local and global level.

Understanding minoritized individuals and the knowledge economy requires examining the historical legacy of schooling in the United States. Thelin (2011) explained how the American classical college was originally based on the European model, and the purpose of higher education institutions was to educate the elite and religious leaders, which consisted of white male elites. This is also on par with Robinson's racial capitalism in which he argued that before the creation of a capitalistic structure people were being racialized in Europe which fed into today's social structure. If students were not upper-class white men, educational attainment was limited (Thelin, 2011). Central to the knowledge economy framework is that educational attainment is necessary for economic success and prosperity. Yet, from a historical standpoint race and class were two factors embedded in the building of the US educational system which could impact minority student educational experiences. Torraco (2018) highlights that educational inequity is exclusively tied to economic inequity, and for students to obtain an education, they must deal with rising costs which exclusively impedes low-income families. This further emphasizes the intersection between race and class in shaping educational experiences. In the next section race and class will be discussed separately for deep analytic purposes, however the intersection of both identities can be explored to examine theory.

Low income/working class minoritized individuals and knowledge economy

Class structures have long existed in many countries, with much research showing that social class has always impacted educational systems (Ardoin and Martinez, 2019). As previously mentioned, the knowledge economy is predicated on increasing levels of education, and the theory most used in tandem when discussing the knowledge economy, human capital theory, attests that as education expands for individuals so does their access to opportunities (Becker, 1964; Carnevale et al., 2013; Torraco, 2018). However, many findings indicate that low income and wealthy families differ on educational attainment due to the resources available to wealthy students throughout their lives, such as traveling, summer enrichment camps, and special lessons (Dickert-Conlin,

2007; Torraco, 2018). While this highlights some of the personal struggles for low-income individuals, it is necessary to analyze educational systems to better understand the class structures within institutions. While there are promising low-income students who succeed in education, data shows they are less likely to attend college and less likely to succeed toward degree completion (Tsoi-A and Bryant, 2015). Despite those barriers for many low-income students, the Pew Research Center (2015) reported that the enrollment of low-income students jumped from 12% to 20% between 1996 and 2016. These findings indicate that overall, there is a rise in the enrollment of low-income students in colleges/universities. However, upon arrival, many low-income students struggle (Barnwell, 2017), which is why it is worth analyzing.

For example, the myth of meritocracy, that working hard and achieving a higher education always leads to continuing improvement and prosperity is limited for those from lower social class backgrounds who do not have access to the same cultural capital as those from higher social class backgrounds (Demetry et al., 2015; Liu, 2011). In recent decades, these disparities have become more pronounced. This is a critical consideration for individuals from lower SES backgrounds because it exposes how even when trying to achieve and succeed in the knowledge economy, low-income students may struggle within the system itself. Yet despite the disparities within higher education, attending college is still promoted as the pathway to economic prosperity. As students strive through the knowledge economy, they have vastly different experiences based on their identities, such as class and race (Bottiani et al., 2016; Wilkins, 2014). In part this may be due to how many educational systems are more geared toward middle to upper-class students which reinforces class inequities for low-income students (Martin, 2012). This educational system being geared toward those from middle to upper class create challenges for low-income students who often have vastly different experiences due to their social and cultural backgrounds.

Some classic scholars argued that education was far from a consensual space and embedded in the hidden curriculum was classism that shape how students were taught and categorized (Berstein, 1971; Bowles and Gintis, 1976; Bourdieu and Passeron, 1990). Specifically, they explain that working class students were groomed to enter the labor market based on how well they behaved and assimilated to these hidden expectations. Just as many students strive to navigate the educational curriculum, there is also a process by which students learn hidden norms and values of their discipline (Alsubaie, 2015; Margolis, 2001; Karnieli-Miller et al., 2011). Educational institutions not only teach the primary curriculum, but they also inherit the teaching of the hidden social and cultural expectations and norms prevalent in educational systems (Apple, 2000; Margolis, 2001). These expectations, if followed, usually raise the cultural capital of the students, however, there can often be a disconnect between the social and cultural expectations of the educational system and what minoritized students have inherited as their own social and cultural aspects.

Yosso (2005) explains that cultural capital can be passed down based on family ties or school, however, Yosso further explains that cultural capital that marginalized students bring into the educational system may not be valued equally. For example, Anthony Jack (2019) in his book, *The Privileged Poor*, compared the stories of three groups of students from the same elite school to examine the role that cultural capital plays in the ability of students to navigate an elite university. One of Jack's most notable findings was that students of the same low-income background differed completely in their ability to navigate based on the cultural capital they gained from attending a college prep school versus a public school. The students who attended a prep school were able to adapt to the cultural and social expectations easier than the students who didn't attend a college prep school. The students who did not attend a prep school had less cultural capital in elite universities because they had not learned the dominant cultures social and cultural expectation to the same level as the prep school students. That's one example of many that highlights how educational systems can create systems of inequality within the knowledge economy and foster barriers for marginalized students. Additionally, the recent study conducted by Jack (2019) shows the issue of educational inequality has persisted throughout the years and continues to raise concerns. While income is a major issue in education, race is also a major factor that has also been an issue of concern as students navigate the knowledge economy.

Racially minoritized individuals and the knowledge economy

Centering on class is one way to examine the inequities present within the knowledge economy, but researchers have explored how race is also central to understanding the knowledge economy (Park, 2018). Given the rise today in support for the Black Lives Matter and Stop Asian Hate movements, studying how racism impacts students is very important. Park (2018) reminds us that American colleges were once made up of only White men but growing numbers of students of color have drastically changed campuses. Additionally, Park contends "the inequality that stratifies students" pathways to college does not end once classes start" (p. 3). This analysis positions race as a central factor in obtaining the education and skills necessary to compete in the knowledge-based society. Concepts such as social class may be more identifiable globally; however, race is central to certain areas such as the United States and with rising international enrollment it is crucial to understand how race impacts the knowledge economy.

While examining social class may be more common globally, the concept of race is increasingly central in U. S. culture. In fact, the human capital theory used to operationalize the knowledge economy was originally questioned for its racial undertones. Theodore Schultz, a prominent economist at the time human capital theory was being created, explained during a presidential address at the American Economic Association, that people should not be seen as "capital" or property and assets due to the connection that has to slavery (Schultz, 1961). In Becker's book, *Human Capital: Theoretical and Empirical Analysis, with Special Reference to Education*, Becker later admitted that he was also worried about using the phrase human capital due to the predictable critiques it would receive (Becker, 1964). So even with race being central to U. S. society, the theory that many countries abroad have embedded within their policies to promote the knowledge economy has been questioned for its racial and culture implications.

For example, the United Kingdom is facing issues within their educational institutions as they work to analyze how race is impacting their systems of education (Arday and Mirza, 2018). Therefore, the concept of race can become a globalized term that has the power to influence any country, and how it impacts the knowledge economy should be deeply explored.

While progressing through the knowledge economy and obtaining education and skills seems simple, race has impacted how students access education and what they experience when they are admitted to educational institutions (Harper et al., 2011; Kniffin, 2007; Park, 2018). Early research has shown that racially minoritized students tend to attend college at a lower rate than their peers even though their overall enrollment has increased. Additionally, while racially minoritized Black students may be attending college more frequently at predominantly white institutions, they are still very unrepresented in certain majors such as the STEM field (Hurtado et al., 2010). While some have identified a multitude of reasons why STEM fields are known for the underrepresentation of Black students, such as precollege preparation, some researchers have found that racial discrimination in the workplace and toxic racial climates in college to also be a major factor (Lee et al., 2020; Pew Research Center, 2018). Present in these findings are pressing issues that the knowledge economy and the theory of human capital can't account for, such as racial trauma and lack of opportunity given to students based on race. Additionally, even when racial minoritized individuals are given access to new work and economic capital after their obtaining of new education and skills, like human capital theory assumes, they still face problematic work environments. Overall, human capital theory which is the basis for the knowledge economy seems to lack in taking account for racialized spaces and experiences.

Transforming the knowledge economy

While this chapter focused primarily on race and class for the purpose of this chapter it is crucial to reiterate there so many identities that students can carry, and they are all connected to a students intersectional experience. Limpangog (2016) reminds us that years after the impactful work of scholars like Patricia Hill Collins, scholars have expanded the marginalized communities that should be studied, such as those with disabilities, transgender and gender nonbinary people, and trafficked people. For example, if many scholars who study gender remind us of the educational spaces also have a deep problematic history in relation to gender and sexuality (Collins, 1998; Duran, 2018; Johnson and Quayle, 2017; Strayhorn and Mullins, 2012). According to the Association of Governing Boards of Colleges and Universities (Trammel, 2014), there are many LGBTQ students in colleges and universities, yet research on those students is lacking (Duran, 2018). This would mean that there are groups of students navigating the knowledge economy that are often forgotten and their stories are often untold. Lange et al. (2019) explain that many LGBTQ students deal with oppressive collegial environments, yet they are often undiscussed and not included in these spaces. Additionally, Lange et al. highlight how students of color who also identify as LGBTQ may be subjugated to additional oppressive systems, such as racism, genderism, and sexism.

To understand these groups new theories, research, and perspectives must be used to critique and analyze the knowledge economy based on diverse minoritized communities. The knowledge economy and the theory (human capital theory) being used to operationalize its assumptions around the world has been critiqued for its prescriptive and "one size fits all" approach (Weber, 2011). Additionally, it has been criticized for seeing human beings as economic tools and dismissing what it means to truly be a human being (Tan, 2014). Despite these critiques, the knowledge economy and human capital theory have become the dominant discourse around the world and have shaped perceptions and practices regarding education (Tan, 2014). However, when we look at minoritized individuals within the knowledge economy, there are clear spaces where many are marginalized within the society and experiencing additional barriers. Tan (2014) states that majority of the critiques on the theory supporting knowledge economy fail because,

Majority of criticisms directed at HCT are rather tantalizing. These critiques give the impression that they have a better alternative model for education policies. But sometimes it is quite noticeable that these criticisms are mostly driven by an ideological zeal with full of resentment just to attack the dominant school of thought with no present alternative at hand (p. 437).

While this assertion is fair and highlights how human capital theory will likely remain the dominant discourse, it still lacks in understanding minoritized experiences. Society is becoming more diverse than ever so if human capital theory remains the dominant discourse, it is equally important to have additional theories that highlight a demographic of individuals who are coming into colleges at an increasing rate. New theories are necessary to build new perspectives and policy changes for education.

Theories for change

Inequality within education has been a topic that has been long discussed within research. Building on Marx's theories, Pierre Bourdieu once argued that other forms of capital existed outside of just economic. Arguably one of his most used theoretical concepts is his examination of what he called "cultural capital." Cultural capital, as defined by Bourdieu, is a person's knowledge and intellectual skills that provide them with the privilege of achieving higher social status (Bourdieu and Passeron, 1977). In addition,

Bourdieu and Passeron argued that if one was born into a family of higher class they would have access to middle to upper-class knowledge and be able to utilize it to increase their social mobility through formal education. This seminal theory has been used for years to explain why certain groups advance in society and while others fall behind. However, Bourdieu in his best efforts failed to explain the racial and cultural factors associated within this process.

Yosso (2005) extended on Bourdieu's work on cultural capital, however she explained the issue in Bourdieu's explanation is that it looked at marginalized populations from a deficit viewpoint and assumed people from marginalized communities needed to "catch up" to those of middle to upper class. However, Yosso contends that people of color have different cultural tools that should be valued by institutions and not create the perception that people of color have less knowledge. In essence, Yosso flipped the argument toward the inequality and inequity present within the systems of schooling and added race as central to educational systems process of valuing and devaluing certain knowledges (Delgado Bernal and Villalpando, 2002; Ladson-Billings, 2000). Yosso formulated what she called a, "community cultural wealth model" that provides a more asset-based view of what students of color bring to their college campuses and how their own cultural aspects are strengths not deficits. Her work is considered a seminal piece in advancing cultural capital and will be used in this session to highlight how this same framework can be applied to advancement in student affairs for Black professionals and students. Yosso's work is one example how we must advance theories that have shaped the perceptions and views toward educational systems.

A theory that has become increasingly popular is the theory of intersectionality coined by Dr. Kimberlee Crenshaw. Intersectionality is focused on identifying, discussing and addressing the ways that systems of inequity, including sexism, racism and class bias, intersect to produce complex relations of power and (dis)advantage (Cho et al., 2013; Crenshaw, 1991). Additionally, this theory encourages us to consider the context individuals are placed in and the power dynamics present within a given system, such as a school. While intersectionality has been used by some scholars as sole theoretical framework (Davis, 2008), other scholars in fields like public health have used intersectionality as a more analytical framework or paradigm or in tandem with another theory (Bowleg, 2012). This highlights how a theory such as human capital theory can be looked at from an intersectional perspective. While these are only a few examples, it shows how main theories used to shape educational systems and how people view the purpose of education need to adapt as society has changed substantially.

Conclusion

Education was once viewed as the perfect means to solving inequities within society. Horace Mann famously said, "Education, then, and beyond all other divides of human origin, is a great equalizer of conditions of men ... the balance wheel of the social machinery". The idea of a knowledge economy concept was present when the purpose of education was first being conceptualized. However, it is very clear that education is not the great equalizer and that minority students are still experiencing great obstacles as they navigate the knowledge economy. While this chapter could not cover all identities race and class alone show the potential problem with viewing education as linear. This Americanized way of thinking has also shaped education globally and this could create a misunderstanding of how education is viewed worldwide. It is paramount that we produce and use theories that also highlight minoritized individuals and communities so we can better understand their experiences and create safer environments. Additionally, more disaggregated quantitative data is necessary to highlight the many experiences minority students are having. Also, qualitatively there should be a continuous highlighting of minority student experiences. Society and educational institutions are becoming increasingly diverse, and systems of thoughts must shift to meet the needs of the new student demographics. Understanding the future leaders of tomorrow should be on the agenda of every educational system.

References

- Alsubaie, M.A., 2015. Hidden curriculum as one of current issue of curriculum. *J. Educ. Pract.* 6, 125–128.
- Anyon, J., 2014. *Radical Possibilities: Public Policy, Urban Education, and a New Social Movement*, second ed. Routledge, New York.
- Apple, M.W., 2000. *Official Knowledge*. Routledge, New York, NY.
- Arday, J., Mirza, H.S., 2018. *Dismantling Race in Higher Education; Racism Whiteness and Decolonising the Academy*. Palgrave Macmillan, London, UK.
- Ardoin, S., Martinez, B., 2019. *Straddling Class in the Academy: 26 Stories of Students, Administrators, and Faculty from Poor and Working-Class Backgrounds and Their Compelling Lessons for Higher Education Policy and Practice*. Stylus Publishing.
- Barnwell, P., 2017. Expanding Access to the Nation's Top Colleges: across the country, selective colleges struggle to recruit and enroll talented, low-income students. But an emerging set of strategies is creating new pathways for admission. *Educ. Update* 59 (6), 2–6.
- Becker, G., 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Harvard University Press, Cambridge, MA.
- Benson, J., Dumas, M.J., 2021. Building out the edges: reading racial capitalism into Jean Anyon's political Economy of urban education. *Teach. Coll. Rec.* 123 (14), 72–94.
- Bernstein, B., 1971. *Class, Codes and Control*, vol. 1. Routledge, New York, NY.
- Blundell, R., Dearden, L., Meghir, C., Sianesi, B., 1999. Human capital investment: the returns from education and training to the individual, the firm and the economy. *Fisc. Stud.* 20 (1), 1–23.
- Boczy, T., Cefalo, R., Parma, A., Skovgaard Nielsen, R., 2020. Positioning the urban in the global knowledge economy: increasing competitiveness or inequality. *Soc. Incl.* 8, 194–207.
- Bottiani, J.H., Bradshaw, C.P., Mendelson, T., 2016. Inequality in black and white high school students' perceptions of school support: an examination of race in context. *J. Youth Adolesc.* 45 (6), 1176–1191.
- Bourdieu, P., Passeron, J.-C., 1977. *Reproduction in Education, Society and Culture*. SAGE Publication, London.
- Bourdieu, P., Passeron, J., 1990. *Reproduction in Education, Society, and Culture*. R. Nice, Trans., second ed. Sage Publication, London.

- Bowleg, L., 2012. The problem with the phrase women and minorities: intersectionality-an important theoretical framework for public health. *Am. J. Publ. Health* 102 (7), 1267–1273.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Basic Books, New York, NY.
- Carey, K., 2018. Introduction: a different kind of college ranking. *Wash. Mon.* 50 (9/10), 15–18.
- Carnevale, A., Smith, N., Strohl, J., 2013. *Recovery: Projections of Jobs and Education Requirements through 2020*. Center on Education and the Workforce, Georgetown University, Washington.
- Cho, S., Crenshaw, K.W., McCall, L., 2013. Toward a field of intersectionality studies: theory, applications, and praxis. *Signs* 38 (4), 785–810.
- Collins, P., 1990. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Unwin Hyman, Boston.
- Collins, P.H., 1998. It's all in the family: intersections of gender, race, and nation. *Hypatia* 13 (3), 62–82.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299.
- Davis, K., 2008. Intersectionality as buzzword: a sociology of science perspective on what makes a feminist theory successful. *Fem. Theor.* 9 (1), 67–85.
- Delgado Bernal, D., Villalpando, O., 2002. An apartheid of knowledge in academia: the Struggle Over the 'legitimate' Knowledge of Faculty of Color, Equity and Excellence in Education 35 (2), 169–180.
- Demetry, D., Thirk, J., Fine, G.A., 2015. Strategic poverty: how social and cultural capital shapes low-income life. *J. Consum. Cult.* 15 (1), 86–109.
- Dickert-Conlin, S., 2007. *Economic Inequality and Higher Education: Access, Persistence, and Success*. Russell Sage Foundation, New York, NY.
- Drucker, P.F., 1992. *Post-Capitalist Society*. Harper Business, New York, NY.
- Duran, A., 2018. Queer and of color: a systematic literature review on queer students of color in higher education scholarship. *J. Divers. Higher Educ.*
- Esping-Andersen, G., 1999. *Social Foundations of Postindustrial Economies*. Oxford University Press.
- Goldin, C., 2016. *Human Capital, Handbook of Cliometrics*. Springer Verlag, Heidelberg, Germany.
- Harper, S.R., Davis, R.J., Jones, D.E., McGowan, B.L., Ingram, T.N., Platt, C.S., 2011. Race and racism in the experiences of black male resident assistants at predominantly white universities. *J. Coll. Student Dev.* 52 (2), 180–200.
- Hope, D., Martelli, A., 2019. The transition to the knowledge economy, labor market institutions, and income inequality in advanced democracies. *World Polit.* 71 (2), 236–288.
- Hurtado, S., Newman, C.B., Tran, M.C., Chang, M.J., 2010. Improving the rate of success for underrepresented racial minorities in STEM fields: insights from a national project. *N. Dir. Inst. Res.* 148, 5–15.
- Jack, A.A., 2019. *Privileged Poor: How Elite Colleges Are Failing Disadvantaged Students*. Harvard University Press, Cambridge, MA.
- Johnson, A., Quayle, S., 2017. Queering Black Racial Identity Development. *J. College Stud. Dev.* 58, 1135–1148.
- Karnieli-Miller, O., Vu, R.T., Holtman, M.C., Clyman, S.G., Inui, T., 2011. Medical students' professionalism narratives: a window on the informal and hidden curriculum. *Correct. Acad. Med.* 86 (1), 29.
- Kniffin, K., 2007. Accessibility to the PhD and Professoriate for First Generation College Graduates: Review and Implications for Students, Faculty, and Campus Policies. *American Academic*.
- Korres, G.M., 2008. *Technical Change and Economic Growth: Inside the Knowledge-Based Economy*, second ed. Routledge, London, UK.
- Ladson-Billings, G., 2000. Racialized discourses and ethnic epistemologies. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 257–277.
- Lange, A.C., Duran, A., Jackson, R., 2019. The state of LGBT and queer research in higher education revisited: current academic houses and future possibilities. *J. Coll. Student Dev.* 60 (5), 511–526.
- Lee, M.J., Collins, J., Harwood, S., Mendenhall, R., Hunt, M.B., 2020. "If you aren't White, Asian or Indian, you aren't an engineer": racial microaggressions in STEM education. *Int. J. STEM Educ.* 7, 1–16.
- Limpangog, C.P., 2016. Matrix of domination. In: Naples, N. (Ed.), *Wiley-Blackwell Encyclopedia of Gender and Sexuality Studies*, pp. 1–3.
- Liu, W.M., 2011. *Social Class and Classism in the Helping Professions: Research, Theory, and Practice*. Sage Press, Thousand Oaks.
- Margolis, E. (Ed.), 2001. *The Hidden Curriculum in Higher Education*. Routledge, New York.
- Martens, K., 2007. How to become an influential actor: the 'comparative turn' in OECD education policy. In: Martens, K., Rusconi, A., Leuze, K. (Eds.), *New Arenas of Education Governance*. Palgrave, New York, NY, pp. 40–56.
- Martin, N., 2012. The privilege of ease: social class and campus life at highly selective, private universities. *Res. High. Educ.* 53 (4), 426–452.
- Means, B., Neisler, J., 2021. Teaching and learning in the time of COVID: the student perspective. *Online Learn.* 25 (1), 8–27.
- Naidoo, R., 2003. Repositioning higher education as a global commodity: opportunities and challenges for future sociology of education work. *Br. J. Sociol. Educ.* 24 (2), 249–259.
- Park, J.J., 2018. *Race on Campus: Debunking Myths with Data*. Harvard Education Press, Cambridge.
- Pew Research Center, 2015. *Most Say Government Policies since the Recession Have Done Little to Help the Middle Class, Poor*. Pew Research Center, Washington, DC.
- Pew Research Center, 2018a. *Early Benchmarks Show 'Post-Millennials' on Track to Be Most Diverse, Best-Educated Generation yet*. Pew Research Center, Washington, DC.
- Pew Research Center, 2018b. *Women and Men in STEM Often at Odds over Workplace Equity*. Pew Research Center, Washington, DC.
- Powell, W., Snellman, K., 2004. The knowledge economy. *Annu. Rev. Sociol.* 30, 199–220.
- Robinson, C.J., 1983. *Black Marxism: The Making of the Black Radical Tradition*. University of North Carolina Press.
- Roser, M., Ortiz-Ospina, E., 2016. *Global Rise of Education*. Retrieved from: <https://ourworldindata.org/global-rise-of-education>.
- Schultz, T.W., 1961. Investment in human capital. *Am. Econ. Rev.* 51, 1–17.
- Stewart, D.-L., 2017. Trans'versing the DMZ: a non-binary autoethnographic exploration of gender and masculinity. Special issue: What's transgressive about trans* studies in education now? *Int. J. Qual. Stud. Educ.* 30 (3), 284–304.
- Strayhorn, T., Mullins, T.G., 2012. Investigating Black Gay Male Undergraduates' Experiences in Campus Residence Halls, 39. *The Journal of College and University Student Housing*, pp. 140–161.
- Takayama, K., 2013. OECD, 'Key competencies and the new challenges of educational inequality. *J. Curric. Stud.* 45 (1), 67–80.
- Tan, E., 2014. Human capital theory: a holistic criticism. *Rev. Educ. Res.* 84 (3), 411–445.
- Terenzini, P., Cabrera, A.F., Bernal, E., 2001. *Swimming against the Tide: The Poor in American Higher Education*. The College Board, New York, NY.
- Thelin, J., 2011. *A History of American Higher Education*. John Hopkins University Press, Baltimore, MD.
- Torraco, R., 2018. Economic inequality, educational inequity, and reduced career opportunity: a self-perpetuating cycle?. In: *New Horizons in Adult Education and Human Resource Development*, 30, pp. 19–29.
- Trammel, J.B., 2014. *LGBT Challenges in Higher Education Today: 5 Core Principles for Success*. Association of Governing Boards of Colleges and Universities. Retrieved from: <https://www.agb.org/trusteeship/2014/5/lgbt-challenges-higher-education-today-5-core-principles-success>.
- Tsoi-A-Fatt Bryant, R., 2015. *College Preparation for African-American Students: Gaps in the High School Educational Experience*. CLASP, Washington, D.C.
- Weber, A., 2011. The role of education in knowledge economies in developing countries. *Proc. Social Behav. Sci.* 15, 2589–2594.
- Wilkins, A.C., 2014. Race, age, and identity transformations in the transition from high school to college for black and first-generation white men. *Sociol. Educ.* 87 (3), 171–187.
- World Bank Institute, 2010. *Measuring Knowledge in the World's Economies: Knowledge Assessment Methodology and Knowledge Economy Index*. Knowledge for Development (K4D) Program, Washington.
- World Bank, 2018. *About the Human Capital Project*. Retrieved from: <https://www.worldbank.org/en/publication/human-capital/brief/about-hcp>.
- Yosso, T., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.

Power/knowledge: global and Indigenous education

Sharon Stein, Cash Ahenakew, and Vanessa Andreotti, Department of Educational Studies, University of British Columbia, Ponderosa Commons North (Oak House), Vancouver, BC Canada, Musqueam Territory

© 2023 Elsevier Ltd. All rights reserved.

Body text	251
Colonial divisions	252
Resisting the colonial divide, reclaiming non-western education	253
Different approaches to global and Indigenous education	255
Global and Indigenous education in the context of emerging challenges	257
Ecological well-being	258
Cognitive well-being	258
Affective well-being	258
Relational well-being	258
Economic well-being	259
Acknowledgment	259
References	259
Relevant websites	260

Body text

Because education is widely understood to be a central site for the reproduction of society, questions about the content, format, and purpose of education are highly contested. This includes questions such as: What/whose knowledge has value, what is education for, and who should decide?

While the ethical and political nature of these questions (and their answers) are rarely acknowledged in mainstream conversations, Indigenous, decolonial, postcolonial, and critical race scholars have all drawn attention to the ways that education has been and continues to be mobilized to entrench colonial hierarchies premised on white, male, western supremacy and the denigration of Indigenous and other non-western ways of knowing, being, and relating. Following [Sandy Grande \(2004\)](#), mainstream education in the west can more precisely be called “whitestream” education, in that is not only often dominated by white people, “but also principally structured on the basis of white, middle-class experience, serving their ethnopower interests and capital investments” (p. 125).

These scholars also emphasize that while whitestream education (especially western schooling) has often served to reproduce colonial divides, other approaches to education can challenge those hierarchies, mobilize global and social change, and nurture non-western knowledge systems, thereby nurturing other possibilities for existence. Here, the distinction offered by many of these scholars between education and schooling becomes useful, wherein “schooling” is used to describe the specific institutionalized form of education that became dominant first in the industrializing west, and then throughout the world largely through colonial processes and globalizing forces ([Ball, 2012](#); [Rizvi and Lingard, 2013](#); [Spring 2015](#)). Meanwhile, “education” can be understood much more broadly as guided learning toward particular directions, which varies widely depending on the context ([Andreotti and Ahenakew, 2013](#)).

The struggle over the direction and purpose of education is heightened in current debates about the role of education in preparing young people to face today’s many overlapping social and ecological challenges. In particular, there are significant questions about whether education should be a site for social reproduction or social transformation; and if the latter, what this transformation should be and how it should happen. In this text, we contextualize the systemic and historical grounds of these discussions in relation to global and Indigenous education specifically. We draw particular attention to the ways that western schools and universities have entrenched universalizing and Eurocentric knowledge systems that invisibilize or invalidate other knowledge systems, and the people who produce, hold, and embody those knowledges ([Sousa Santos, 2007](#); [Willinsky, 1998](#)). In doing so, we also review the landscape of contemporary efforts to create more space for global and Indigenous education, both in the context of western educational institutions and beyond. As part of this review, we offer a social cartography that seeks to support more generative engagements between different approaches to global and Indigenous education. Finally, we conclude the text by offering a second cartography that emerged from our work in the areas of global and Indigenous education, and that seeks to prepare learners to face the complexity, uncertainty, and plurality of today’s world with maturity, humility, discernment, and accountability ([Stein, 2021](#)).

Before we proceed, we note that there is a paradoxical dimension to any effort to challenge the universalism of western knowledge in educational contexts. On the one hand, it is necessary to at least temporarily center that knowledge in order to critically trace its colonial origins and implications; if we do not do this work of deconstructing Eurocentric frames, then they are likely to be naturalized and reproduced. On the other hand, critiquing western universalism takes space away from engaging other ways of knowing

and being, and their associated forms of education. We have sought to balance both of these imperatives while also addressing head on the complexities that emerge at the interface of different knowledge systems with unequal systemic power, and the common challenges and paradoxes that characterize efforts to include or integrate other ways of knowing into western-centric educational contexts. This text is informed by our different positionalities as a white settler (Sharon), an Indigenous person (Cash), and a racialized immigrant who is an Indigenous descendant (Vanessa), as well as our current geographic location in what is currently known as Canada. Finally, this text is inspired by our ongoing collaborations with each other and many others, especially our work as part of the Gesturing Towards Decolonial Futures Collective. The collective is made up of researchers, artists, educators, students, activists, and Indigenous knowledge keepers who are interested in how we might imagine “education otherwise” in relation to both global and Indigenous education. Rather than try to imagine alternative forms of education from where we currently stand, the collective seeks to develop “alternative approaches to alternatives” (Sousa Santos, 2007) through pedagogical experiments that gesture toward possibilities that are viable, but unimaginable within mainstream institutions. This approach presumes that if alternatives are not to reproduce colonial patterns by assuming a “single path forward” and making harmful projections about the future, then they cannot be predetermined in advance. Rather, they must be collaboratively woven as we move together in the context of relationships premised on trust, respect, reciprocity, accountability, and consent (Whyte, 2020).

Colonial divisions

Knowledge, power, and education have become deeply entangled through the historical and ongoing structures of western colonization. As John Willinsky (1998) writes,

Much of the knowledge achieved through conquest and colonization was understood to legitimate the political and cultural domination of imperialism. The resulting perspective on the world formed an educational legacy that we have now to reconsider. (p. 3)

In fact, Ramon Grosfoguel (2013) argues that western knowledge and thus western education was only able to reach its current position of global hegemony because of the violence that was enacted through European colonialism and slavery, starting in the 15th century. In other words, the *epistemicides* of non-western knowledges were only possible because of the *genocides* of non-western peoples. At the same time, these genocides and epistemicides were accompanied by *ecocides* that were rationalized through western knowledge that framed the earth and “nature” (other-than-human beings) as resources to be extracted for human use, rather than as living entities with whom humans are related and to whom they have reciprocal obligations (Davis and Todd, 2017).

One of colonialism’s most significant educational legacies is the ongoing hierarchical division of the world into two sets of people: those who are deemed the “leaders” of humanity, and those deemed the “followers.” Those designated as “leaders” of humanity are presumed to have the ability and the authority to determine: a universal direction for human progress and change; the legitimacy of knowledge (what counts as legitimate, good, important, normal); the allocation of resources; and the rules of global engagement, among many other things. Generally, this designation is reserved for those who are located in the Global North and/or those who are deemed to be white. Meanwhile, those designated as the “followers” of humanity are presumed to lack the relevant knowledge and capacity for political leadership, economic autonomy, and epistemic and moral authority, and are thus structurally relegated to positions of servitude and subordination. Generally, this designation is imposed on those located in the Global South and/or those who are deemed to be non-white.

This racial and colonial division of humanity has been contested from its very beginnings by those whom it assigns to positions of subjugation. Often in response to this resistance, the exact form of this division has shapeshifted over time and place. Nonetheless, these basic divides continue to naturalize the hegemony of the systems and subjectivities of the supposed leaders of humanity. This includes an economic system rooted in global capitalism, a political system rooted in nation-states, a knowledge system rooted in universal reason, and subjectivity premised on unrestricted autonomy and individualism. Generally speaking, the education system that accompanies these systems supports their reproduction. Collectively, these systems can be said to constitute a “capitalist/patriarchal western-centric/Christian-centric modern/colonial world system” (Grosfoguel, 2013, p. 82), which we refer to throughout this text as the “modern/colonial system.”

The relationship between power and knowledge that characterizes this modern/colonial system is usefully illustrated by Boaventura de Sousa Santos’s (2007) notion of abyssal thinking. Sousa Santos argues that modern western thinking is characterized by an “abyssal line” that divides reality into two distinct parts. On “this side” of the colonial divide is what is real, imaginable, and desirable, while what lies on “the other side” is deemed unintelligible and insignificant, framed as the product of superstition and a lack of universal reason. While this outright dismissal is harmful, even when “this side” of the line acknowledges and tries to access “the other side”, it is often mis-interpreted in ways that are also harmful. For example, western frames of knowledge generally cannot comprehend that other-than-human beings are understood within many Indigenous knowledge systems as teachers (Marker, 2004; Simpson, 2014). This is illustrated in a school strategy document for Indigenous engagement in Canada, which featured a quote from an Indigenous elder telling a story about being chased by a bear up a tree, affirming that this experience made it clear to him how the bear is also a teacher. However, in the document that reported this story, the elder’s account is interpreted as an affirmation of the value of environmental education for Indigenous children and of the need for knowledge about dangerous

animals that can pose a threat to one's safety in natural spaces. In this interpretation, the voice of the elder was grafted onto an anthropocentric (human-centered) ontology in which knowledge is produced and shared solely by humans, nature and culture are separate spheres, and relationships between humans and animals are mediated by human knowledge that is produced for human safety.

Reality on "this [modern western] side" of the abyssal line does not simply ignore, belittle, or mis-read reality on "the other side"; it is structured by the colonial divide itself. "This side" is presumed to hold totalizing, universal knowledge and truth, which is only legible in contrast to its presumed opposite, that is, partial or particular knowledge and truth on the "other side" (of non-western knowledge). As [Ahenakew \(2016\)](#) argues, "The denial and denigration of non-western ways of knowing has been part and parcel of European colonialism and a primary means by which the universality of Western knowledge was asserted" (p. 327). Furthermore, this abyssal line has not only epistemological but also material implications, as the "denial and denigration" of non-western knowledge has been "used as a justification for the dispossession, destitution, and genocide of populations who were perceived to be lacking knowledge of universal worth (but who occupied lands of strategic importance)" (p. 327). Thus, the abyssal line that constitutes western knowledge refuses the equality of other knowledges, and denies the full humanity of the people who hold those knowledges, in a process that [Sousa Santos \(2007\)](#) calls "epistemicide."

The effect of the abyssal line on education, particularly in western schooling contexts, is significant. Not only did imperial worldviews inform western schooling from the outset, but western powers have also imposed western knowledges and educational systems onto non-western peoples in both direct and indirect ways. In what is currently known as Canada, over 150,000 Indigenous children were separated from their families and forced attend to residential schools from the 1880s until as recently as 1996 ([Truth and Reconciliation Commission of Canada, 2015](#)). Residential schools were framed as part of a benevolent "civilizing mission" that would "save" Indigenous peoples from ignorance and religious heresy by assimilating them into white middle-class Christian society; the schools are now recognized as a form of cultural genocide. The first Prime Minister of Canada, John A. Macdonald, told the Canadian House of Commons in 1883, "When the school is on the reserve the child lives with its parents, who are savages; he is surrounded by savages, and though he may learn to read and write his habits, and training and mode of thought are Indian. He is simply a savage who can read and write ... Indian children should be withdrawn as much as possible from the parental influence, and the only way to do that would be to put them in central training industrial schools where they will acquire the habits and modes of thought of white men." Apart from the deep paternalism and western supremacy that oriented residential schools, they can also be understood as part of a wider colonial effort to alienate Indigenous peoples from their ancestral lands, communities, and cultures, and thus, secure the future of settler colonial society on those lands.

Similar schools were created in what are currently known as the United States and Australia. In fact, Canadian residential schools were likely influenced by US boarding schools, first developed by former US Army general Richard Pratt, who famously stated that the intention of the schools was to "Kill the Indian, save the man." This sentiment was echoed in other educational projects throughout the US and its territories, as well as in other colonial contexts throughout the world. For instance, in British colonized India, [Thomas Macaulay \(1835\)](#) wrote, "We must at present do our best to form a class who may be interpreters between us and the millions whom we govern; a class of persons, Indian in blood and color, but English in taste, in options, in morals, and in intellect" (p. 729).

Although these more explicit strategies of colonization through education are increasingly rare, the colonial division of humanity and subtler forms of assimilation continue to structure education throughout much of the world ([Ahenakew, 2016](#)). For instance, many national and global educational reform efforts portray Indigenous and racialized peoples through pathologized and deficit-based framings ([Ahenakew, 2017](#); [Ahenakew and Naepi, 2015](#); [Ladson-Billings, 2006](#); [Tuck, 2009](#)). While often framed as benevolent interventions, these reforms tend to measure Indigenous and racialized peoples against white, Eurocentric norms, thereby re-universalizing western ways of knowing and being, and denigrating other possibilities for education, and existence. This approach is guided by the same colonial divisions that were instituted centuries ago, in which non-white people are presumed to be less developed along a universal, Eurocentric path of human progress led by white people (the "leaders of humanity"). In addition to reproducing racist hierarchies, this framing deflects attention away from the fact that the educational experiences of both dominant and marginalized communities are deeply shaped by systemic inequalities that are rooted in the compounded social, political, and economic impacts of centuries of European colonization and slavery.

Resisting the colonial divide, reclaiming non-western education

As Konai [Thaman \(2008\)](#) writes, speaking of educational reforms that are framed as "transferring" purportedly universal forms of knowledge and education from the Global North to the Global South,

People who participate in these interventions rarely ask: How do people in this community, this place, conceptualize wisdom, learning and knowledge? Nor do they wonder if the values inherent in and propagated by their agendas are shared by the majority of people whose lives are meant to be improved as a result of their intervention. Few even realize the ideological and philosophical conflicts associated with different perceptions of championed ideas, leaving many communities confused and, in some cases, angry. (p. 3)

For as long as western powers have sought to impose and universalize their worldview, many of the communities on the receiving end of this imposition have resisted it (Morreira et al., 2020). In doing so, they have insisted on the need to contest western epistemic universalism and the hegemony of western schooling, and support and sustain educational efforts that center non-western ways of knowing and being (without then assuming that those ways of knowing and being are static or unchanging). Resistance, revitalization, and resurgence efforts in education have been theorized and documented by educational scholars writing from the peripheries of the white, western hegemonic center, including: Black feminist perspectives, such as those of Gloria Ladson-Billings (1998, 2006), bell hooks (1994), Jessica Harris and Lori Patton Davis (2019); Indigenous education scholars, including Eve Tuck (2009), Linda Tuhiwai Smith (2012), and Leilani Sabzalian (2019); those who have critiqued the colonial dimensions of international and global education, like Leon Tikly (2004), Keita Takayama (2018), and Vanessa Andreotti (2011); and various educational scholars of post-colonialism, such as Ashis Nandy (2000), Roland Sintos Coloma (2009), and Riyad Shahjahan (2014), and decolonization, such as La Paperson (2017), George Sefa Dei (2008), and Rubén Gaztambide-Fernández (2012). Thus, while below we review three specific examples of efforts to resist the coloniality of western education and support other educational possibilities, we emphasize that there are countless other examples that we could draw on. These three examples were chosen in order to illustrate a range of possible strategies for resisting and reimagining education, including how these possibilities are shaped by socially and historically situated contexts and analyses.

Our first example is *Indian Control over Indian Education*, a position paper written in 1972 by the National Indian Brotherhood (now known as the Assembly of First Nations) in Canada. As Michelle Pidgeon (2016) notes, the document sought to assert Indigenous peoples' right to determine their own education "across the lifespan (e.g., from early childhood to post-graduate education)" (p. 78). The paper also articulated a specific orienting purpose for Indigenous peoples' education, noting "We believe in education: as a preparation for total living, as a means of free choice of where to live and work, as a means of enabling us to participate fully in our own social, economic, political, and educational advancement" (National Indian Brotherhood, 1972, p. 3). 50 years after *Indian Control Over Indian Education* was published, Indigenous peoples in Canada and elsewhere are still pursuing diverse strategies to secure access to forms of education that "center Indigenous knowledge as the source for content, curricula, and pedagogy" (Pidgeon et al., 2013, p. 28).

Indigenous communities have for centuries sustained Indigenous educational practices in the face of colonial efforts to eradicate them. As Leanne Simpson (2014) argues, it is not necessary to entirely forgo western schooling in order to emphasize the importance of Indigenous forms of education. Yet, she also asks, "Don't we deserve learning spaces where we don't have to address state learning objectives, curriculum, credentialism and careerism, where our only concern for recognition comes from within?" (p. 23). Indeed, Indigenous peoples and communities often relate to education in layered ways that are always shaped by the larger colonial context. In some cases, they focus on interrupting the coloniality of whitestream schools and carving out Indigenous spaces in those institutions. In other cases, they focus on reinvigorating land-based Indigenous educational practices that exceed what is possible within western schooling. While these strategies might appear on the surface to be contradictory, Indigenous peoples' approaches to education are shaped by a paradoxical set of imperatives: the practical imperative to survive within western societies that are based on an unsustainable single story of human progress achieved through social-economic accumulation that treats the earth as a resource to be exploited; and the existential imperative to keep alive alternative possibilities for noncapitalist existence based on reinvigorated ancestral ways of knowing and being that recognize the earth as a living entity and relation (Ahenakew, 2016). This paradox is further heightened in the context of contemporary social and ecological crises, including climate change and biodiversity loss, political instability, and economic austerity and precarity. In this context, the crises and potential collapse of our existing system only heightens the importance of sustaining alternative possibilities. At the same time, those possibilities have become sites of extraction and appropriation for non-Indigenous peoples and corporations seeking to reinvigorate that system using the commodified knowledges and practices of Indigenous peoples, for instance in relation to sustainable planning practices and plant medicines (Whyte et al., 2018).

Our second example of resistance to whitestream educational norms comes from Manish Jain's (2013) writings about global education. Jain critiques the UNESCO "Education for All" campaign, suggesting that it is more accurately described as "McEducation for All" as it imposes a western, capitalist approach to education onto the non-western world as if it were the only viable option. He argues that Education for All "tells us that we must all walk on a single, universal, linear, standardized path of education and development, which is dictated by the logic of the industrial-military system" (p. 85). He suggests that while education is often framed as the solution to many contemporary problems, dominant forms of education are in fact the root of many of these problems. Jain advocates for an approach to global education that supports efforts to "reclaim and revitalize what is commonly called 'informal education', basically all of the natural learning that happens in everyday life that does not fit neatly into institutionalized categories of formal and non-formal education" (p. 84). In his own practice, Jain seeks to reinvigorate informal education in India through his involvement with Swaraj University, which he identifies as "India's first gift culture university dedicated to regenerating the local economy, local culture and local ecology" (pp. 89–90).

Swaraj can be understood as part of the larger global deschooling movement, which seeks to reclaim education from its institutionalized, and especially westernized, forms. This includes the Ecosities Alliance, of which Swaraj is a part, which seeks to reimagine higher education by asking "What might the university look like if it were at the service of our diverse ecologies, cultures, economies, spiritualities, and Life within our planetary home?." Jain (2016) has articulated a "Declaration of Decolonizing Education" to help orient this work, which includes the following statements:

I can no longer accept a narrative of education, which teaches me that my village grandmother was illiterate, primitive, backward, stupid, uneducated, underdeveloped, uncivilized and not capable of managing their own affairs.

I can no longer accept a narrative of education that sees my links to my land, to my local languages, to my seeds, to my rivers, to my trees, to my histories and herstories, to my body, to my inner voice, to the spirit world, to my community all as a barrier to modernization and development which must at best be destroyed if we are to progress, and at worst be condemned to a multicultural day festival in school.

I can no longer accept a narrative of education that teaches me that physical work in the fields, in my home and in my community is drudgery and my children's definition of "happiness" lies in drinking Coca Cola, eating at McDonalds, using Fair and Lovely face whitening creams and chatting on Facebook.

I can no longer accept a narrative of education that teaches me that I have to compete against others in my community and against peoples from other countries to survive.

I can no longer accept a narrative which teaches me that learning is a commodity (along with the air, water, land, food) and that knowledge is the property of individuals through copyrights and patents.

In our third and final example, [Django Paris and H. Samy Alim \(2014\)](#) write about the importance of "culturally sustaining pedagogies" (CSP) that decenter white middle class norms of educational achievement, and "foreground the heritage practices of community of color", while also "taking into account contemporary/evolving community practices" (p. 85). CSP is rooted in what Paris and Alim call "asset pedagogies" that have sought to position "the linguistic, literate, and cultural practices of working-class communities – specifically poor communities of color – as resources and assets to honor, explore, and extend" (p. 87). This approach seeks to both refuse the pathologization of non-white communities, and the denigration of their knowledges, but also to specifically challenge the notion that equity in education is equivalent to getting "working class students of color to speak and write more like middle-class white ones" (p. 87). [Ahenakew and his colleagues \(2014\)](#) raise similar concerns. While affirming the contingent importance of inclusion-based efforts to address colonization in whitestream institutions, they warn that "if we try to provincialize Western thought within the institutions (e.g. nation-states, universities, schooling, etc.) that were created to naturalize it, we will need to remain within its language, epistemology and ontology, even when we claim to be doing the opposite" (p. 220). This points to the complexities and tensions arise within and between different approaches to engaging global and Indigenous education in the context of whitestream institutions.

Different approaches to global and Indigenous education

In many whitestream educational institutions, there is growing interest in incorporating global and/or Indigenous perspectives into curriculum, pedagogy, and institutional culture. Often, these institutions are seeking to restore or maintain their perceived benevolence and relevance in response to public critiques about their whiteness or Eurocentrism. While this has presented important opportunities to denaturalize and decenter western knowledge, it has also raised considerable complexities and tensions in practice. In particular, Indigenous and other non-western knowledges are often engaged in selective, tokenistic, and decontextualized ways that do not disrupt business as usual; meanwhile, critiques that challenge the presumed benevolence, universality, and continuity of the modern-colonial system are generally excluded or minimized.

One way of considering how educational institutions incorporate global and Indigenous perspectives is to ask how they approach changes to the modern/colonial system that, up until now, has largely excluded and invalidated these perspectives within formal education. Approaches to change can range from thinking that the system is working optimally as is, to thinking that the system has a good structure but requires some degree of revision or "reform," to thinking that the system is so fundamentally flawed that it cannot possibly be reformed and ultimately we will need a different system, or set of systems, in its place. In this section, we review four different approaches to reform.

In many whitestream schools and universities, there is still no perceived imperative to rethink educational practice or reform the modern/colonial system (i.e. *no reforms*). Increasingly common is an approach that presumes that this system requires *minor reforms*, such as small policy shifts and curriculum changes. Others suggest that *major reforms* are needed in order to center non-western perspectives and redistribute resources to marginalized students, researchers, and communities. Still others argue that these institutions might ultimately be unable to transform their colonial foundations, and therefore are *beyond reform* (see [Andreotti et al., 2015](#)). In order to consider the possibilities and limitations of each of these approaches, in the remainder of this section we present a social cartography of four different approaches to education, emphasizing how each approach engages the modern/colonial

Table 1 Different approaches to education in relation to the modern/colonial system.

	<i>Education for system maintenance (no reform)</i>	<i>Education for system expansion (minor reform)</i>	<i>Education for system transformation (major reform)</i>	<i>Education for system hospicing (beyond reform)</i>
Purpose of education	Transmit knowledge and values to ensure the reproduction of the modern/colonial system	Expand access to the modern/colonial system, and ensure wider participation in its improvement	Empower learners to organize in pursuit of greater social, political, and economic justice	Prepare learners with the capacity to face complexity and uncertainty with maturity, humility, discernment, and accountability
Approach to knowledge	Western knowledge is universally valuable; other knowledges are insignificant	Western knowledge is universal, but should be supplemented by local knowledges	Western knowledge should be decentered, non-western knowledges should be centered	All knowledges are partial, situated, and have contextual (rather than universal) relevance
Approach to global education	Developing cultural competencies and global awareness in order to foster good diplomatic and business relations, and competitive advantages	Facilitating inclusion in and informed engagement with institutions to promote global peace, equality, empathy, mutual understanding	Fostering advocacy for and political solidarity with historically and systemically marginalized populations against systemic global injustice	Engaging with the complexities, paradoxes, uncertainties, difficulties of social and ecological challenges without seeking simple solutions, and without allowing complicity in systemic harm to disappear from view
Approach to Indigenous education	Indigenous knowledge has no proper place in the context of western schooling; Indigenous peoples should adapt to existing institutions	Select elements of Indigenous culture can be used to support Indigenous success and foster non-Indigenous empathy	Indigenous education should be implemented in ways that center and prioritize Indigenous self-determination and sovereignty	Indigenous peoples have a right to determine their own education; Indigenous education has much to teach non-Indigenous people
Imaginary of hope and justice	Through continuous innovation we will increase efficiency, effectiveness and competitiveness of existing educational institutions	Through small, consensus-based changes to policy and practice, we will ensure continuous improvement to existing forms of education	Through collective action, we will radically transform the existing educational institutions to work for more people, and enact redress for harm	We need to hold space for “education otherwise,” that is, education outside of what is currently imaginable within the modern/colonial system
Theory of change	Human rationality and ingenuity will enable us to engineer system continuity even in the face of change	Prepare more informed and engaged citizens and workers to adapt in face of disruption	Exposure to critiques of inequity will lead people to transform their values & actions	The limits of the current system will compel us to consider other possibilities for knowing and being
Example	Common Core State Standards Initiative	UNESCO Futures of Education Initiative	<i>Indian Control over Indian Education</i>	The University of the Forest

system (i.e. no, minor, major, or beyond reform), how each approaches global and Indigenous education in particular, and examples of each (summarized in [Table 1](#)). Readers might also wish to revisit the three examples of resistance to whitestream education reviewed in the previous section, and consider where they would place each example in relation to this cartography, noting that some might fall at the interface of different approaches.

Social cartographies are not meant to offer a comprehensive representation of reality, but rather to highlight tensions and contradictions and to make visible perspectives that are being erased. Further, positions in a cartography are not meant to indicate static stances or identities, as one's approach might shift from one context to another depending on what is possible and legible. As thought experiments, cartographies invite readers to think *with* rather than *about* them, and prompt the possibility of dynamic movement without directing people toward a particular end. For instance, in relation to this cartography of education, we invite readers to consider: What are the assumptions behind each position? What does each position enable and preclude, and what does each position invisibilize in order to maintain its own coherence? What other approaches might be missing? How can we engage in generative conversations across these different perspectives?

An approach to education that seeks *system maintenance* emphasizes the need to ensure the continuity of “business as usual,” including ongoing innovation to engineer increased efficiency and effectiveness. Ultimately, education is understood as a means to reproduce the modern/colonial system as is. In this approach, western knowledge is understood to be universally valuable and relevant, regardless of the context. The approach to global education from this position emphasizes the development of cultural competencies that can enable good geopolitical diplomacy, secure strategically important international business relationships, and maximize individual and national economic competitiveness. The approach to Indigenous education from this position sees little value in Indigenous ways of knowing and being, and no place for them within whitestream educational institutions; it is presumed that Indigenous people should integrate (assimilate) into dominant society, apart from perhaps preserving a few cultural traditions in their own spaces. Alongside system expansion, system maintenance is probably the most widely accepted approach to education in the context of whitestream institutions.

Education based on *system expansion* generally assumes that minor reforms to the current educational system are all that is required, with the goal of spreading the perceived gifts of western knowledge and education as widely and democratically as possible. Thus, this approach seeks tweaks to policy and practice that expand access to this knowledge and education. Global education from this position is focused on fostering peace, equality, and mutual understanding, while Indigenous education is approached as a means to supplement western knowledge with elements of Indigenous culture in order to both support Indigenous communities to succeed in whiteman institutions, and support western students to develop empathy for Indigenous communities. The result is that Indigenous knowledges are selectively engaged and approached through the frames of western knowledge and white middle-class sensibilities, rather than on their own epistemological and ontological terms (Ahenakew, 2016). Ultimately, this approach to education does not decenter western knowledge, challenge western paternalism, nor question the desirability of the continuity of the modern/colonial system.

A *system transformation* approach to education actively seeks to decenter western knowledge, and often seeks to center non-western knowledges in its place. There is a commitment to developing forms of education that can support learners to transform the world toward greater justice, especially in ways that center the needs, perspectives, and knowledges of marginalized peoples. Global education is framed as an opportunity to foster solidarity between communities across differences and in pursuit of systemic changes, shifting power relations, and redress for historical and ongoing injustice. Indigenous education in this approach should center the needs and priorities of Indigenous communities; whether this happens in formal institutions or alternative educational contexts is secondary to the question of whether or not it supports Indigenous peoples' political and intellectual sovereignty and self-determination. This approach to education is generally dismissed within whiteman discussions and institutions, although it might be granted space at the margins.

Finally, a *system hospicing* approach to education suggests that while reforms to existing institutions might be important in the short-term, ultimately, these institutions are rooted within a modern/colonial system that is inherently violent and unsustainable. From this analysis comes the understanding that we need a form of education that can prepare us to face the possible end of the existing system – that is, to “hospice” it – and create space for something else to emerge in its place. The metaphor of hospicing derives from a sense that our current modern/colonial system is dying, and those of us who are part of this system have a responsibility to support it to “die well”. This support entails learning from the system's gifts as well as its many mistakes, and cleaning up its messes (including working to redress the harm it has caused). Hospicing is different from efforts to place this system on life support through various reforms, or efforts to force its end prematurely. The work of system hospicing is also accompanied by the work of “midwifing”, or supporting the birth of something different. Education from this approach emphasizes the indispensability and insufficiency of all knowledges, while noting that Indigenous and other non-western communities have much to teach western communities that exceeds what is currently taught in western education. However, it cautions that these engagements will reinscribe the abyssal line if they do not interrupt extractive, appropriative, and tokenistic modes of relationship, and instead operate from consent, reciprocity, and respect (Whyte, 2020). Thus, in relation to both global and Indigenous education, this approach to education emphasizes the generative possibilities that can emerge at the interface between different knowledge systems, as well as the potential to reproduce colonial patterns at this interface. This approach to education is not only marginalized within whiteman institutions, but also challenges existing epistemic frames to the extent that it is largely unintelligible to many within them, which is why we have offered an extended engagement with it here. Given that both hospicing and midwifing are complex, difficult, and at times painful processes, this approach to education is generally of most interest to those who have reached the limits of various efforts to reform the modern/colonial system, and are thus open to learning (and unlearning) that asks questions with no clear answers and invites deep inquiry with no clear solutions or end point. Rather than transmitting specific content or prescribing a pathway to a predetermined end, an education rooted in system hospicing supports learners to develop the maturity, humility, discernment, and accountability that they will need in order to sustain hospicing and midwifing work in the long-term.

Global and Indigenous education in the context of emerging challenges

The current landscape is characterized by numerous overlapping social and ecological challenges, including climate crises and biodiversity loss, economic austerity, precarity and instability, mental health crises, global health pandemics, social fragmentation, political polarization, and more. We conclude this text by asking what kind of education could prepare learners to face these challenges in generative ways. In this, we are inspired by Zygmunt Bauman (2000), who over 20 years ago described education as “preparation for life.” Without prescribing a single approach to this preparation, we propose some provisional possibilities based on our own research and practice.

Education, in its different modalities (formal schooling, higher education, informal education, alternative education, etc.), has historically been tasked with steering learning toward ensuring continued human survival as well as the reproduction of cultural norms and ideals. However, this double mandate becomes paradoxical when the reproduction of dominant cultural ideals poses a threat to human survival. This paradox is illustrated by Luis Prádanos (2020), who asked, in a recent piece about the future of education: “[I]s it really smart to educate people to technologically and theoretically refine a system that operates by undermining the conditions of possibility for our biophysical survival?” Prádanos argues that it is unwise to approach education in a way that presumes the continuity of our existing system, because the continuation of that system will ultimately cause us to exceed the limits of the planet. In addition to questions of ecological unsustainability, we face enduring challenges that are rooted in the historical and ongoing maintenance of a modern/colonial system characterized by a Eurocentric and anthropocentric hierarchy that

legitimizes violence against both humans and other-than-human beings. Thus, we add to Prádanos's inquiry another question: Do we really want to educate people to technologically and theoretically refine a system that is sustained through colonial violence?

As the limitations of our existing system and its educational institutions become increasingly evident, more people are seeking out educational alternatives. Some of those experimenting with these alternatives are oriented by a system transformation approach to education, while others are oriented more by a system hospicing approach, or an intention to skip the hospicing altogether and jump right into a predetermined replacement system. Some of these efforts can be considered alternative approaches to education, while a few might be "alternative approaches to alternatives." In other words, these educational alternatives are oriented by different desires, investments, and horizons of hope, and thus, it is unlikely that we will arrive at a coherent alternative educational vision. Indeed, we suggest that a single vision is undesirable, as it would lead to implementing a new totalizing universal in place of the old, and suppressing dissenting voices in the process. However, we close this text by inviting readers to sit with the following question: What approaches to global and Indigenous education could enable us to stay grounded and accountable in the face of unprecedented and unpredictable change, and respond to whatever arises without becoming overwhelmed, or immobilized? We offer one partial, provisional answer to this question through another social cartography that outlines a set of basic dispositions that might orient us away from forms of education that are premised on the reproduction of the enduring colonial divide that characterizes western-centric schooling, and toward forms of education that foster a deepened sense of responsibility for shared existence on a finite planet, across species and across generations.

The "In Earth's CARE" framework emerged through our transnational collaborations with Indigenous communities in Canada, Peru, Mexico, and Brazil (Stein et al., 2020). In this way, it can be understood as relevant to both global and Indigenous education. This framework is informed by alternative economies and ecologies, Southern epistemologies, and initiatives that highlight the importance of teachings from grassroots resistance and land-centered movements, including Black, Indigenous, landless, peasant, and Quilombola struggles. This pedagogical framework was produced with an intention to identify the limits of the modern/colonial system without overdetermining what an alternative system, and attendant forms of education, might look like. It proposes a set of dispositions for a pedagogy of everyday life organized around five dimensions of well-being that are represented in the acronym of the framework: E – Ecological; C – Cognitive; A – Affective; R – Relational; and E – Economic. We review these below:

Ecological well-being

- Enabling education from a space where humans are not separated from the land/planet;
- Grappling with the complexities of addressing differential complicities in ecological harm;
- Reflecting on the challenges of shared existence on a finite planet from different perspectives, including those of other-than-human beings; and,
- Developing resiliency for the slow and challenging work that needs to be done in order to address the accumulated effects of centuries of ecological destruction.

Cognitive well-being

- Thinking in multiple layers to face and embrace complexities, tensions, and paradoxes at the interface between different knowledge systems;
- Connecting dots in the bigger picture to understand how different struggles are connected;
- Examining the different layers of our systemic complicity in harm in deeper ways; and,
- Making space for and relating differently to the unknown and the unknowable.

Affective well-being

- Learning to hold space for internal complexities and contradictions;
- Developing the capacity to be uncomfortable and to accept uncertainty and internal contradictions without becoming irritated, overwhelmed, or immobilized;
- Identifying and starting to unravel projections, fears, (in)securities, anxieties, and perceived entitlements at individual and collective levels; and,
- Nurturing a sense of responsibility that is not dependent on choice, convictions, or convenience.

Relational well-being

- Developing stamina to engage with difficult issues and conversations without relationships falling apart and without investments in harmful continuities;
- Exploring possibilities for relating not grounded on shared meaning, identity or conviction;
- Experiencing the difficulties and complexities of ethical engagements and solidarity from a space of accountability; and
- Learning collectively through difficult events with humility, generosity and patience.

Economic well-being

- Identifying and interrupting patterns of consumption (of knowledges, experiences and relationships) as a mode of relating to the world, in order to give and receive differently;
- Decentering one's desires and centering collective needs by committing to do what is needed rather than what one wants to do;
- Challenging patterns of entitlement coming from social, economic and/or racial hierarchies; and
- Learning to practice economies based on abundance, reciprocity, and redistribution, rather than self-interest or utility maximization.

In this text we have sought to identify the historical context and ongoing conditions that shape both global and Indigenous education. In addition to identifying a modern/colonial system that shapes dominant forms of education, especially in the context of western schooling, we have examined different efforts to critique and reform these forms of education. We suggest that in light of today's many emerging, overlapping social and ecological challenges, there is a need to consider the gifts and limitations of each of these and other approaches to education. We have particularly emphasized forms of education that are rooted in and/or inspired by Indigenous and other non-western knowledge systems, and that gesture beyond investments in the continuity of the existing modern/colonial forms of education. These approaches are often marginalized in whitestream schools and universities, and indeed are even dismissed as unintelligible due to the abyssal line that invisibilizes anything that is not articulated within and affirming of western values and frames. However, as the continuity of the modern/colonial system becomes increasingly uncertain, the imperative to create space for other educational possibilities becomes increasingly evident.

Acknowledgment

We would like to acknowledge that this text was completed on the traditional, ancestral, and unceded territory of the Musqueam people. We support the struggle of the Musqueam people, and all Indigenous peoples, to have their lands returned to them. This text was made possible through the funding of the UBC Hampton Fund Research Grant, "Challenges and Complexities of Decolonizing Higher Education: A Global View", the SSHRC Insight Grants, "Re-imagining Aboriginal Education for Shared Futures" and "Ethical Social Innovation for Decolonial Global Futures", and the SSHRC Insight Development Grant "Towards the Ethical Integration of Different Knowledge Systems: Lessons from STEM and Health Fields."

References

- Ahenakew, C., 2016. Grafting Indigenous ways of knowing onto non-Indigenous ways of being: the (underestimated) challenges of a decolonial imagination. *Int. Rev. Qual. Res.* 9 (3), 323–340.
- Ahenakew, C.R., 2017. Mapping and complicating conversations about Indigenous education. *Diaspora Indig. Minor. Educ.* 11 (2), 80–91.
- Ahenakew, C., Andreotti, V.D.O., Cooper, G., Hireme, H., 2014. Beyond epistemic provincialism: de-provincializing Indigenous resistance. *AlterNative* 10 (3), 216–231.
- Ahenakew, C., Naepi, S., 2015. The difficult task of turning walls into tables. In: *Sociocultural Realities: Exploring New Horizons*. Canterbury Press, pp. 181–194.
- Andreotti, V., 2011. *Actionable Postcolonial Theory in Education*. Springer.
- Andreotti, V., Ahenakew, C., 2013. Educating. In: Matthewman, S., West-Newman, C., Curtis, B. (Eds.), *Being Sociological*. Palgrave MacMillan, New York, NY, pp. 233–250.
- Andreotti, V., Stein, S., Ahenakew, C., Hunt, D., 2015. Mapping interpretations of decolonization in the context of higher education. *Decolonization* 4 (1), 21–40.
- Ball, S.J., 2012. *Global Education Inc: New Policy Networks and the Neo-Liberal Imaginary*. Routledge, New York, NY.
- Bauman, Z., 2000. Education under, for, and despite postmodernity. In: Brown, A., Schemmann, M. (Eds.), *Language, Mobility, Identity: Contemporary Issues for Adult Education in Europe*. Lit Verlag, Munster, pp. 27–43.
- Coloma, R.S., 2009. *Postcolonial Challenges in Education*. Peter Lang.
- Davis, H., Todd, Z., 2017. On the importance of a date, or, decolonizing the Anthropocene. *ACME* 16 (4), 761–780.
- Gaztambide-Fernández, R.A., 2012. Decolonization and the pedagogy of solidarity. *Decolonization* 1 (1), 41–67.
- Grande, S., 2004. *Red Pedagogy: Native American Social and Political Thought*. Rowman & Littlefield.
- Grosfoguel, R., 2013. The structure of knowledge in westernised universities: epistemic racism/sexism and the four genocides/epistemicides. *Hum. Architect. J. Sociol. Self Knowl.* 1, 73–90.
- Harris, J.C., Patton, L.D., 2019. Un/doing intersectionality through higher education research. *J. High Educ.* 90 (3), 347–372.
- Hooks, b., 1994. *Teaching to Transgress: Education as the Practice of Freedom*.
- Jain, M., 2013. McEducation for All: whose agenda does global education really serve? *Crit. Lit. Theor. Pract.* 7 (1), 84–90.
- Jain, M., 2016. Declaration of Decolonizing Education. Retrieved from: <https://medium.com/the-emperor-has-no-clothes/declaration-of-decolonizing-education-37ddc32a4cde>.
- la paperson, 2017. *A Third University Is Possible*. University of Minnesota Press, Minneapolis, MN.
- Ladson-Billings, G., 1998. Just what is critical race theory and what's it doing in a nice field like education? *Int. J. Qual. Stud. Educ.* 11 (1), 7–24.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in US schools. *Educ. Res.* 35 (7), 3–12.
- Macaulay, T.B., 1835. On Indian education. Minute of the 2nd of February, 1835. In: Young, G.M. (Ed.), *Macaulay: Prose and Poetry*. Harvard University Press, Cambridge, MA, pp. 721–729.
- Marker, M., 2004. Theories and disciplines as sites of struggle: the reproduction of colonial dominance through the controlling of knowledge in the academy. *Can. J. Native Stud.* 28 (1/2).
- Morreira, S., Luckett, K., Kumalo, S.H., Ramgotra, M., 2020. Confronting the complexities of decolonising curricula and pedagogy in higher education. *Third World Thematics: A TWQ Journal* 1–18.
- Nandy, A., 2000. Recovery of indigenous knowledge and dissenting futures of the university. In: *The University in Transformation: Global perspectives and the future of the university*, pp. 115–123.
- National Indian Brotherhood, 1972. *Indian Control over Indian Education*. Available at: <https://blogs.ubc.ca/indigenouseducation/2013/01/14/indian-control-of-indian-education/>.
- Paris, D., Alim, H.S., 2014. What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harv. Educ. Rev.* 84 (1), 85–100.
- Pidgeon, M., 2016. More than a checklist: meaningful Indigenous inclusion in higher education. *Soc. Incl.* 4 (1), 77–91.

- Pidgeon, M., Muñoz, M., Kirkness, V.J., Archibald, J.A., 2013. Indian control of Indian education: reflections and envisioning the next 40 years. *Can. J. Native Educ.* 36 (1), 5–35.
- Prádanos, L., 2020. How Did This Class Prepare You for Extinction? Resilience. Retrieved from. <https://www.resilience.org/stories/2020-04-21/how-did-this-class-prepare-you-for-extinction/>.
- Rizvi, F., Lingard, B., 2013. *Globalizing Education Policy*. Routledge, New York, NY.
- Sabzalian, L., 2019. *Indigenous Children's Survivance in Public Schools*. Routledge.
- Sousa Santos, B.D., 2007. Beyond abyssal thinking: from global lines to ecologies of knowledges. *Binghamton University Review* 30 (1), 45–89.
- Sefa Dei, G.J., 2008. Indigenous knowledge studies and the next generation: pedagogical possibilities for anti-colonial education. *Aust. J. Indig. Educ.* 37 (1), 5–13.
- Shahjahan, R.A., 2014. From 'no' to 'yes': postcolonial perspectives on resistance to neoliberal higher education. *Discourse* 35 (2), 219–232.
- Simpson, L.B., 2014. Land as pedagogy: Nishnaabeg intelligence and rebellious transformation. *Decolonization: Indigeneity, Education & Society* 3 (3), 1–25.
- Smith, L.T., 2012. *Decolonizing Methodologies*. Zed Books, London and New York, NY.
- Spring, J., 2015. *Globalization of Education: An Introduction*, second ed. Routledge, New York, NY.
- Stein, S., 2021. Reimagining global citizenship education for volatile, uncertain, complex and ambiguous (VUCA) times. *Global. Soc. Educ.* 19 (4), 482–495.
- Stein, S., Andreotti, S., Hunt, D., Suša, R., Amsler, S., Ahenakew, C., Jimmy, E., Cardoso, C., Siwek, D., Cajkova, T., Pitaguary, B., Pataxo, U., Valley, W., D'Emilia, D., Okano, H., 2020. Gesturing towards decolonial futures: reflections on our learnings thus far. *Nord. J. Comparat. Int. Educ.* 4 (1), 43–65.
- Takayama, K., 2018. Beyond comforting histories: the colonial/imperial entanglements of the international institute, Paul Monroe, and Isaac L. Kandel at Teachers College, Columbia University. *Comp. Educ. Rev.* 62 (4), 459–481.
- Thaman, K., 2008. Preface. In: Andreotti, V., Mario, L., de Souza, T. (Eds.), *Learning to Read the World through Other Eyes*. Global Education Derby, Derby, p. 6.
- Tikly, L., 2004. Education and the new imperialism. *Comp. Educ.* 40 (2), 173–198.
- Truth, and Reconciliation Commission of Canada, 2015. *Canada's Residential Schools: The Final Report of the Truth and Reconciliation Commission of Canada*. Available at. <https://publications.gc.ca/site/eng/9.807830/publication.html>.
- Tuck, E., 2009. Suspending damage: a letter to communities. *Harv. Educ. Rev.* 79 (3), 409–428.
- Whyte, K., 2020. Too late for indigenous climate justice: ecological and relational tipping points. *Wiley Interdisciplinary Reviews: Clim. Change* 11 (1), 1–7.
- Whyte, K.P., Caldwell, C., Schaefer, M., 2018. Indigenous lessons about sustainability are not just for “all humanity”. In: *Sustainability: Approaches to Environmental Justice and Social Power*, pp. 149–179.
- Willinsky, J., 1998. *Learning to Divide the World: Education at Empire's End*. University of Minnesota Press, Minneapolis, MN.

Relevant websites

Critical Internationalization Studies Network. <http://criticalinternationalization.net>
Gesturing Towards Decolonial Futures. <http://decolonialfutures.net>
In Earth's Care Global Justice Framework. <https://blogs.ubc.ca/earthcare/>

Global university rankings and geopolitics of knowledge

Riyad A. Shahjahan and Sanzhar Baizhanov, Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

What is geopolitics of knowledge (GoK)?	261
Brief overview of the big three GURs	262
ARWU	262
THE	263
QS	263
GURs and universality of quality and excellence	264
GURs and colonial knowledge/power relations	265
Concluding remarks	268
Acknowledgment	268
References	269

When we recently did a quick google search on “world university rankings,” over 250 million sites emerged. On top of this list, we found global university rankers’ websites, such as *Times Higher Education’s World University Rankings (THE)*, *QS’ World University Rankings*, *US News World Report’s Best Global Universities Report*, *Shanghai Ranking’s World University Rankings portal*, and so on. Other sites that emerged were university websites highlighting their rankings performances, such as, “Texas AandM Rises In World University Rankings, Now In Top 4%”, or “QS ranks MIT the world’s No. 1 university for 2020–21”, and “U of T ranked among world’s top 25 universities in latest QS World University Rankings.” We also found links to various news media outlets, like the British Broadcasting Corporation (BBC) reporting that “Oxford top of global university rankings,” or the Malaysian New Straits Times, noting that “Malaysian research universities in top 200 of QS World University Rankings.” Given the huge coverage of university rankings across mass media globally, including institutional websites, news media, and social media (Stack, 2016), it’s not surprising that the phrase “world university rankings” has produced over a quarter billion hits in our google search. But, why do university rankings get so much notice? What informs our curiosity and dependence on them?

Over the past two decades, global university rankings (GURs) have received unprecedented attention from so many higher education stakeholders, ranging from policy makers and scholars to parents, students, and the general public. In our own personal interactions with higher education (HE) stakeholders across the globe, including North America, East Asia, South Asia, Central Asia, Latin America, and/or Africa, the topic of global university rankings always surfaces. It seems like a privilege for those who can ignore it (Sauto-Otero and Enders, 2017). GURs introduce new complex dynamics to higher education policy and governance, and the effects of these processes are various and context-dependent.

Since the emergence of the Academic Ranking of World Universities (ARWU), the first GUR in 2003, there has been a growing robust literature on global rankings in the past decade. Other GURs include the Times Higher Education World University Rankings (THE), the QS World University Rankings (QS), and U-Multirank among others. This literature includes discussions of GURs’ methodological limitations, GURs’ impact on faculty, institutions, policy makers, and studies on how GURs influence institutional decision making (Erkilä, 2014; Hazelkorn, 2018; Lo, 2011; Stack, 2016; Yudkevich et al., 2015). Taken together, this work has provided a rich analytical and empirical description of the rise and impact of GURs on higher education institutions (HEIs) in the context of globalization.

Considering the above GUR trends and scholarly debates as the wider context, this article provides a critical overview of GURs, using a geopolitics of knowledge (GoK) lens (Mignolo, 2011). More specifically, we will first illuminate what we mean by GoK and how it applies to GURs. We next highlight and introduce three major university rankings that have had a significant global impact, such as ARWU, THE and QS, also considered the Big Three in global higher education (Stack, 2020). Using the existing literature, we then highlight how these three rankings perpetuate GoK. We illuminate two ways in which GoK manifests through GURs, such as (a) projecting universality of quality and excellence, and (b) reproducing colonial knowledge/power relations. We argue that critically examining the GoK underlying GURs is integral to articulating and challenging the global politics of knowing and being in contemporary HE. We suggest that to achieve diversity and global justice in education, we need to challenge and address the global politics of knowing and being by a) tackling the colonial knowledge/power relations and b) denaturalizing the terms of existence framed by modern/colonial ways of being informing the principle of rankings.

What is geopolitics of knowledge (GoK)?

We draw foremost upon Walter Mignolo’s (2003a, 2011) GoK to frame our analysis. GoK signifies how all knowledge systems originate in geographic and social contexts and are situated within historically and transnationally constituted power relations. Mignolo noted that “local histories are everywhere but ... only some local histories are in a position of imagining and implementing global designs” (Delgado et al., 2000, p. 8). Geopolitics of knowledge is derivative of ongoing coloniality—an ongoing logic of

domination underlying imperial powers (whether Spanish, Portuguese, British, and, later on, American) and a Eurocentric process of expansion of knowing, being, and representation (Mignolo, 2011). Furthermore, coloniality and modernity are constitutive of each other: there is no modernity without coloniality. We use GoK in the HE context in two ways.

First, by GoK, we refer to a set of knowledge/power relations privileging a certain gaze or representation of the world deemed universal, delocalized, and applied unquestioningly. Here, we will demonstrate that GURs reflect and perpetuate an epistemic gaze (i.e., criteria of rankings, quality higher education, etc.) that originates in geographic and social contexts. Yet, the positionality of the epistemic gaze is simultaneously rendered invisible by its worldwide normalization (Mignolo, 2003a). Modern epistemology has thus perpetuated coloniality by delocalizing and detaching its concepts (e.g., rationality, logic, and epistemic virtues) from local histories and presenting them as applicable to everyone (Mignolo, 2003b; Smith, 1999). In terms of our analysis, this means that only certain local knowledge systems, always already derivatives of particular historical-material conditions, have the social privilege to shape global thinking. For instance, the GoK continues in contemporary higher education, as the “Westernized” university and liberal arts models have taken root across the globe (Godwin, 2015; Grosfoguel, 2013; Mignolo, 2003a), and Americanized global templates enfolding concepts of autonomy, competition, and markets are transmitted and assimilated across the globe (King, 2009), even if the ways they are taken up are influenced by context. Similarly, transnational actors, like W.B. (World Bank) or OECD (Organization for Economic Cooperation and Development), perpetuate GoK through their epistemic activities as they perpetuate a “policy by numbers” culture—representations of policy realities through performance indicators (Rizvi and Lingard, 2010)—as they introduce their policy ideas and numbers to policymakers (Shahjahan, 2016).

As we will demonstrate, GURs embody a quality assurance culture derived from Anglo-European knowledge systems, which becomes a global template in the GoK (Lewis, 2009; Shahjahan, 2013). We demonstrate how GURs privilege and valorize templates (i.e., globally competitive HE institutional models, policy by numbers, and discourse of quality) that derive from historically epistemically privileged positions, thereby reproducing GoK. The installment of ranking criteria and language (albeit modified) continue to perpetuate GoK as they are viewed as universal or objective recipes.

Second, we draw on a GoK lens to unearth and signify the “geopolitical configuration of the world” (Mignolo, 2011, p. 118) informing existing hierarchical knowledge systems (e.g., institutional types, disciplines, and languages), and the psycho-social dynamics that feed the desires for emulating metropolitan centers of knowledge (e.g., subjectivity). According to Frantz Fanon (1968), those have been historically allocated in the zone of non-being (i.e., marginalized or considered sub-human in the Global North or South) are forced to emulate those in the zone of being, just to be or belong, which, in turn, legitimates the being-ness of those in the zone of being. Similarly, as Ashis Nandy (1983) reminds us, the coloniality of knowledge “perpetuates itself by inducing the colonized, through socio-economic and psychological rewards and punishments, to accept new social norms and cognitive categories” (p. 3). Within the GoK, those who occupy zones of being have the epistemic privilege to articulate global designs (Shahjahan and Morgan, 2016). Yet, privileged epistemic centers of knowledge also comprise privileged institutions within peripheral¹ zones, such as the University of Cape Town in South Africa or Nazarbayev University in Kazakhstan.

GoK facilitates the production of ranking subjects (i.e., institutions, policy makers, academics and students) who internalize, assume, and take for granted coloniality of knowledge (i.e., dominant knowledge systems) as neutral, superior, and good for everyone, as it colonizes their ways of knowing and being. GoK thus plays out as peripheral nations, HEIs, faculty, and students (who previously occupied zones of non-being), mimic (albeit differently) the characteristics inscribed within the GURs criteria, to enter the competition even if many have no chance of winning. GURs and the templates that go with it thus gains legitimacy and reach via peripheral players’ desire for belonging in an international community and desire to gain epistemic privilege (Shahjahan and Morgan, 2016). In this regard, peripheral players reinforce the universality of “university rankings” as a means to assert their competitiveness and strive for the opportunity to attain global stature (Ntshoe and Letseka, 2010). GURs offer subjects an orientation, or a path, that is “followed by being created and are created by being followed. The more people follow the path, the clearer it becomes, the easier it is to follow”, thus narrowing the “set of possibilities” (Schmitz, 2014, p. 103). GoK helps us go beyond colonial epistemes, to articulate and challenge why HEIs, nation-states, faculty, and students, particularly from zones that are marginalized, nevertheless willingly join this “rankings” academic olympiad. Overall, centering the GoK lens shifts attention from knowledge itself to, who, when, why, and where knowledge is produced (Mignolo, 2011), which is why we next situate the Big Three GURs and answer these W questions.

Brief overview of the big three GURs

ARWU

ARWU, also known as the Shanghai Ranking, was created by Shanghai Jiao Tong University (SJTU) in 2003 as part of the government’s “985” Project to encourage Chinese universities to pursue world-class status. ARWU commenced by Professor Nian Cail Liu, a chemistry professor at SJTU, to achieve an aim “to supply a benchmark to measure China’s progress in the university arena” (Stack, 2016, p. 34). To compare Chinese institutions to the “world-class universities” in the US and worldwide, ARWU developers used Euro-American-centric data and categories in which US research universities stand out as proxies of excellence (Ordorika and

¹By peripheral zones or players, we are referring to those geographic regions and actors, usually former colonies or colonized subjects marked by underdevelopment or cultural differences (i.e. “low-middle income country” or “developing economy”), who continue to experience the brunt of coloniality, leading to large inequities and are allocated at the margins of knowledge production due to colonial epistemes, leading to epistemic dependence (Mignolo, 2002; Grosfoguel, 2002).

Table 1 Global university ranking criteria.

ARWU ²	THE (World university rankings) ³	QS (World university rankings) ⁴
(1) Quality of faculty 40% The staff of an institution winning Nobel prizes and fields medals 20% highly cited researchers 20% (2) Research output 40% Papers published in nature and science 20% Papers indexed in science citation index-expanded and social science citation index 20% (3) Quality of education 10% Alumni of an institution winning Nobel prizes and fields medals 10% (4) Per capita academic performance 10% The weighted scores of the above five indicators divided by the number of full-time equivalent academic staff 10%	(1) Research 30% Reputation survey 18% Research income 6% Research productivity 6% (2) Teaching 30% Reputation survey 15% Staff-to-student ratio 4.5% Doctorate-to-bachelor's ratio 2.25% Doctorates-awarded-to-academic staff ratio 6% Institutional income 2.25% (3) Citations 30% (4) International outlook 7.5% Proportion of international students 2.5% Proportional international staff 2.5% International collaboration 2.5% (5) Industry income 2.5%	(1) Academic reputation 40% Teaching quality Research quality (2) Faculty/student ratio 20% (3) Citations per faculty 20% (4) International student ratio 5% (5) International faculty ratio 5% (6) Employer reputation 10%

Lloyd, 2015), such as publications in English, number of Fields Medals and Nobel Prize winners, and publications in Science and Nature (see Table 1). ARWU currently publishes four types of rankings to meet diversified needs for the global comparison (see Table 2). The administrative operations today are provided by Shanghai Ranking Consultancy that defines itself as a fully independent organization focused on higher education information with authority to publish and copyright. Initially developed for the Chinese context to build world-class universities, ARWU lacked a specific audience (Lim and Oeberg, 2017). ARWU's primary focus was on universities and government policymakers (Stack, 2016). The Center for World-Class Universities, housing ARWU, provides consultancy services for governments and universities (Downing and Ganotice, 2016). ARWU has organized information sharing and network facilitating services through the Annual International Conference on World-class Universities since 2005.

THE

Since the company's beginning in the UK in 2004, THE was a subsidiary of TES Global, a UK digital education company focusing on student placement services and recruitment with a focus on MBA programs (Stack, 2013). Most recently, TES Global's owner, TPG Capital, a US private equity firm, sold THE to Inflexion, a UK private equity firm (THE Reporters, 2019). Despite the change in corporate structure and ownership, THE's mission remains the same: to "help the world's universities achieve excellence" (THE, 2019). To do this, a primary operation is ranking universities around the world. THE's first ranking, the QS-THES ranking, was a collaboration between QS and THE. After disagreements on methodology, when QS methodology favored the sciences, and their reputation survey was highly problematic, the two rankers split in 2009, and THE began producing its own World University Ranking (Stack, 2016). Now, THE offers world, regional, subject, teaching, and impact rankings (THE, 2019). In 2014, THE ended its partnership with data-supplier Thomson Reuters and has since used Elsevier as their supplier of ranking data (Hanafi and Boucherie, 2018; THE, 2014). THE provides 12 types of rankings to meet diverse needs for global comparison (see Table 2). THE also offers other HE products and services oriented toward senior academics and sector "thought leaders," including news media, hosting conferences, branding, and data analytics (Lim and Oeberg, 2017; Shahjahan et al., 2020, 2022). THE consultancy provides strategic, data-driven guidance to universities and governments globally.

QS

QS is a global HE ranking edu-business based in the UK. In 1990, founder and current CEO, Quacquarelli Symonds, created QS as a startup business during his MBA from the Wharton School of the University of Pennsylvania (QS, 2019a). The company offered guides for international students interested in elite US universities. By 1994, QS was a UK corporation connecting university

²ARWU. (2019). *Methodology* [Online]. Academic Rankings of World Universities. Available at: <http://www.shanghairanking.com/ARWU-Methodology-2019.html> (Accessed 25 August 2020).

³THE (2020a). *THE World University Rankings 2020: methodology* [Online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2020-methodology> (Accessed 25 August 2020).

⁴QS. (2020b). *QS World University Rankings methodology* [Online]. QS Topuniversities. Available at: <https://www.topuniversities.com/qs-world-university-rankings/methodology> (Accessed 19 August 2020).

⁵ARWU. (2020). *Rankings* [Online]. Academic Rankings of World Universities. Available at: <http://www.shanghairanking.com/ARWU2020.html#> (Accessed 25 August 2020).

⁶THE (2020b). *World University Rankings* [Online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings> (Accessed 25 August 2020).

⁷QS. (2020c) *University Rankings* [Online]. QS Topuniversities. Available at: <https://www.topuniversities.com/university-rankings> (Accessed 25 August 2020).

Table 2 Types of rankings by GURs.

ARWU ⁵	THE ⁶	QS ⁷
Global rankings		
Academic ranking of world universities	World university rankings	QS world university rankings
Rankings by subject/field		
• Academic ranking of world universities by subject (54 subjects) • Academic ranking of world universities by field (5 fields) • Shanghai ranking's global ranking of sport science schools and departments	• World university rankings by subject (11 subjects)	• QS world university rankings by subject (48 subjects) • QS business masters rankings • QS global MBA rankings
Rankings by region/location		
	• US college rankings • Japan university rankings • Europe teaching rankings • Latin America university rankings • Asia university rankings • Emerging economies university rankings	• QS US university rankings • QS university rankings Asia • QS university rankings Arab region • QS university rankings Latin America • QS university rankings emerging europe and central Asia • QS university rankings BRICS • QS university rankings by location • QS best student cities
Other categories of rankings		
	• THE young university rankings • World reputation rankings • Wall street journal/times higher education college rankings • impact rankings	• QS graduate employability rankings • QS top 50 under 50 • QS system strength rankings • QS star rating system • QS IGAUGE rating system

admissions offices and prospective international students around the world. Now, QS self-describes as a leading higher education marketing company (QS, 2019b). QS lists many influential media partners, including the Guardians, IELTS, British Council, Times of India, and the New York Times. QS's mission is "to enable motivated people anywhere in the world to fulfill their potential through educational achievement, international mobility, and career development" (QS, 2019a). After ending its ranking collaboration with THE in 2009, QS developed a multitude of rankings, including its World University Ranking and rankings by subject, graduate employability, age of university, cities, and region (Elsevier, 2015). The current set of rankings includes 17 types, comparing various aspects of higher education institutions and systems (see Table 2). Ranking data primarily comes from reputation surveys and Elsevier's Scopus database (Elsevier, 2015; QS, 2019a). QS also offers institutional products and services, including consulting, marketing, and data analytics. And, like THE, QS took several decades to achieve this operational portfolio. QS oversees international conferences for higher education leaders, including Reimagine Education, EduData Summit, QS Apple, and a range of subject-focused summits held with partner universities.

As our Table 2 highlights, the variety and types of rankings has mushroomed among the Big Three GURs (particularly QS and THE), ranging from rankings that are global to subject-field oriented, to geographic-specific (i.e., regional or national), toward more technical oriented rankings (e.g., impact, teaching, or graduate employability), or specific institutional rankings (e.g., Young university). Such a growing variety of rankings reflects rankers "seeking ways to become attractive" to a variety of local markets, and have accommodated the latter "by creating new rankings that focus only on a smaller subset of the global rankings" (Lim and Oeberg, 2017, p. 103). Now that we have situated the Big Three GURs, in the following sections, we demonstrate how these GURs perpetuate GoK in two ways: (a) universalizing particular languages, and (b) reproducing colonial knowledge/power relations.

GURs and universality of quality and excellence

GURs perpetuate the geopolitics of knowledge through a universality project. This happens by universalizing particular languages of quality and excellence. By language, we're not simply referring to a communication medium, but also an onto-epistemic grammar that undergirds what is desirable, knowable, and valid as ways of knowing. We view language here as a system of representation, or conceptual map, that helps constitute the meaning of the world (Hall, 1997). Here we demonstrate how universality of quality and excellence is reproduced by furthering totalizing the English language, a "policy by numbers" heuristic, rankings language, and proxies of quality.

GURs feed the GoK by further universalizing the English language as the bastion of academic language. Many note that HEIs derived from English-speaking countries and/or former English colonies perform the best in GURs. For instances, Kaba (2012) noted that "six out of every ten universities on the top 200 list [in THE rankings] were located in countries that were at one time

partly or fully colonized by the United Kingdom" (p. 30), and Shin and Harman (2009) found research universities concentrated in English-speaking and more economically developed countries. GURs reinforces the universality of English in several ways. While GURs derive much of their legitimacy from claims of objectivity, GURs rely heavily on internationally recognized measures of research production, such as the number of scholarly articles included in the Thomson Reuters' Web of Science or Elsevier's Scopus databases, which are heavily biased toward English-language publications (Ordorika and Lloyd, 2015) (see Table 1). A dominant GURs culture results in devaluing research in the national language, as citation indexes referenced in GURs, are based predominantly on publications in English Language and does not acknowledge vernacular languages, especially in social sciences and humanities (Lo and Liu, 2021; Ishikawa, 2009). This has led to certain aspiring globally competitive nations and their HEIs to engage in English language initiatives, like English skills development in academic writing among faculty (e.g., Russia), or having more programs that utilize English, and/or wide-scale implementation of English in their institutions (e.g., Malaysia, China, and/or Japan) (Azman and Kutty, 2016; Dunrong, 2017; Pavlyutkin and Yudkevich, 2016).

GURs also foster GoK by universalizing a "policy by numbers" heuristic. By perpetuating the enumerative logic of counting and classification (Shahjahan, 2016), GURs feed colonial epistemological governance (Appadurai, 1996). The appeal for GURs finds fertile ground in the context of a transnational drive for evidence-based decision-making (Erkkila, 2014). Amidst this policy scape, GURs establish and assume a universal language of excellence/quality facilitated by a "practice of quantification" that is "applicable to all universities worldwide irrespective of location and sociocultural context" (Brankovic et al., 2018, p. 277). Rankings also serve as calculative practices making characteristics, actions, and positions of individuals and institutions legible, thereby actively extending a numerical imagination of the world (Collins and Park, 2016). GURs serve policy-making by numbers through framing decision-making and resource allocation (Hazelkorn, 2018). As such, GURs have increasingly shifted from being a statistical-technical phenomenon, influencing institutional decision-making, toward a policy instrument driving competition at institutional, national, and global levels (Hazelkorn, 2018). Such statistics and book-keeping can create an "iron cage" that leaves no other option but to submit to their calculative logic—limiting the realm of politics and ethics (Erkkila, 2014).

Tied to epistemological governance, GURs perpetuate the GoK by universalizing the ranking language as proxies of quality. Such language is underpinned by a zero-sum game of winners and losers with no total gain or loss, in which resources gained by one party (i.e., ranking position) are matched by corresponding losses to another party (Lee et al., 2020a,b). GURs thus presume a nonexistent zero-sum game by arbitrarily creating limited spots for top ranks (Lee et al., 2020a,b). Using tables, GURs also allocate universities hierarchically according to "how much excellence" they have, and in so doing, GURs suggest "that the vertical slots allocated merely reflect "the objective reality" when it comes to performance: (a) not all universities can be equally good and (b) the better a university, the higher its slot" (Brankovic et al., 2018, p. 280). GURs thus indexes global higher education in a language of single logic and reductionism. By signifying qualities as numbers, rankers validate their zero-sum statement in unambiguous terms and provide a means of public comparison that is audience-friendly (Brankovic et al., 2018, p. 274). As we will demonstrate, such ranking language is furthered in global HE through the mediatization of GURs among the mass media and institutional branding efforts.

GURs also universalize by fixing the meaning of quality, thus perpetuating the GoK. The locus of articulation of quality is slowly shifting from local HE to the global (that derive from particular local histories). As we demonstrated earlier, each GUR has its own history and particular vision of university excellence, advancing that notion, and their methodological construction favors certain types and nationalities of institutions (Lim and Oerberg, 2017, p. 100). First, GURs presume that the highest standard of quality of performance is something that can be universally defined, objectively captured and ascribed to HEIs of all stripes, and further measured, compared and communicated to an audience (Brankovic et al., 2018, p. 277). Second, GURs' notions of quality are diffused because of the geopolitics of HE, whereby GURs play an important role in an academic olympiad. National pride and reputation also include the quality of human capital formation, which are predetermined by the leading positions of the leading national universities in GURs (Avrilev and Efimova, 2013). Thus, the state's position in the world largely depends on the positions held by its leading universities in the world. As such, the notion of quality higher education is being transformed from something being institutionally led to being driven and regulated by the state, and now a critical part of international HE, and rankings become convenient report cards (Hazelkorn, 2018). GURs are perceived as an international measure of quality, especially in developing countries, particularly given some nation states don't have HE quality assurances in place, hence GURs serve as those tools (Hazelkorn, 2018). Third, GURs count what is measured rather than measure what counts. In other words, as mentioned earlier, the availability of data becomes the means of defining quality, such as only those indicators which are available and can be measured and counted are used (e.g., research, reputation, learning environments, etc.) (Altbach and Hazelkorn, 2018; Kehm, 2014). In particular, GURs put emphasis on research performance and then make statements about the quality and performance of the institution as a whole (Kehm, 2014).

GURs and colonial knowledge/power relations

Here we demonstrate how GURs reinforce existing hierarchical knowledge systems and produce subjectivity intertwined with colonial/knowledge power relationships. We show how GURs orient HE stakeholders (e.g., government, institutions, faculty, leadership, and students) toward certain parameters of success.

GURs perpetuate the GoK utilizing the stakeholders' aspiration for Anglo-American types of institutions. The Big Three have demonstrated consistently similar top ten institutions in each of their annual rankings. The two of the 10 universities came from Southern England (Oxford and Cambridge), and the rest came from the USA (i.e., Caltech, Stanford, Berkeley, Columbia, Yale, Princeton, MIT, Harvard, and University of Chicago). These wealthy institutions have endowments and access to major media centers (Stack, 2020). While some universities are public, all are research-intensive, with diversified funding, entrepreneurship culture, and wide international collaboration. They perpetuate an Anglo-American university model, as an increasing number of countries attempt to imitate the HE system of the US and UK (Kaba, 2012). For instance, China planned to have 20 top universities that match MIT; Nigeria hoped for 20 world-class universities (WCUs); Sri Lanka wants at least one WCU; Japan's Ministry of education has targeted 30 universities becoming world-class institutions (Douglass, 2014). Beyond exemplifying Anglo-American institutional types, GURs privilege metropolitan centers of knowledge within nation-states. For instance, QS Best Student Cities rankings (see Table 2) distinguishes large urban, world-class cities (e.g., Tokyo, Paris, Berlin, Beijing, and Seoul) with highly ranked institutions at the expense of less developed cities (with exceptions in the US) (Shattock, 2017).

Governments play a major role in reproducing the GoK. Governments strive to belong to the "international" community, assert their "humanity", or demonstrate "quality" through WCU initiatives. Interestingly, the concept and diffusion of WCU was an idea popularized at the global scale by a World Bank staff member Jamil Salmi, and its Tertiary Education Program (Salmi, 2009). Salmi's (2009) book, *The Challenge of Establishing World Class Universities*, published by the World Bank, shaped national understandings of what "top" universities ought to do and their special role for economic development (Allen, 2019; Lim and Oerberg, 2017). Such a WCU concept both relied on and shaped the criteria of different rankers, thus reproducing GoK. As mentioned above, many governments and institutions worldwide heavily invest in WCU initiatives within their respective higher education sectors to "catch up" with Anglo-American institutional models (Dembereldorj, 2018; Collins and Park, 2016; Kim et al., 2018; King, 2009; Lee and Naidoo, 2020; Lo and Liu, 2021; Makinen, 2021). Government-driven WCU initiatives range from mergers of smaller institutions to larger ones (e.g., France and Russia), internationalization efforts (e.g., Japan and South Korea), funding initiatives focusing on an exclusive number of national universities to attract academic stars and increase their publication outputs in indexed journals (e.g., China and Korea), build up research infrastructure (through research centers) (e.g., Germany and Japan), and/or through competitive student admissions, infrastructure construction and seeking strategic partnerships with highly ranked institutions (e.g. Malaysia or Kazakhstan). Furthermore, other countries have attracted WCUs from elsewhere to create "hubs" to set up operations locally and establish joint ventures with their domestic counterparts (e.g., Singapore) (King, 2009). In developing countries and middle-income countries, the priority has been to build research-intensive universities (Dembereldorj, 2018; Kiraka et al., 2020; Lee and Naidoo, 2020). For instance, Nazarbayev University, in Kazakhstan, was created by the decree of the president N. Nazarbayev (Ruby et al., 2017; Sabzalieva, 2017), and similarly the Malaysian government invested in the Universiti Sains Malaysia to attain international stature (Lee and Naidoo, 2020). Based on the examples, national policymakers are active participants in accepting and reinforcing the Anglo-Saxon model of higher education. While national actors, such as Russia, India, Brazil, and China, express concerns over the hegemonic influence of GURs, their HE sectors nonetheless reflect the influence of ranking's paradigm (Lloyd and Ordorika, 2021; Makinen, 2021; Shahjahan et al., 2021). In short, governments representing former imperial powers, colonies, and/or those desiring to take on the new imperial role (Rhee, 2009), seek to improve their global status through WCU initiatives tied to GURs.

Coupled with the national priorities to establish WCUs, ranking agencies advise governments globally co-partnering in HE policy interventions amplifying the GoK. *THE*, for instance, provides advising to governments on the development of institutional programme activity supporting the United Nations Sustainable Development Goals and World Bank Tertiary Education program (THE, 2019). UNESCO, together with the Organization for Economic Co-operation and Development (OECD) and the World Bank, organized the Global Forum on Rankings and Accountability in Higher Education: Uses and Misuses in 2011. The organizers pursued the goal to build capacity for responsible development, dissemination, and informed use of rankings (Marope et al., 2013). Therefore, the governments' cooperation with ranking agencies to reform higher education systems represent higher education policy interventions promoting dominant policy scripts, best practices (Kehm, 2014) and crypto-currency of "world-classness" (Hazelkorn, 2018, p. 19). The cases of Denmark and India demonstrated examples when ranking agencies become "closely involved into national policy discourses, and moved beyond a technical practice measuring performance to become a technology of government with only limited control" (Lim and Oerberg, 2017, p. 104). Similarly, in the case of Kazakhstan, ranking logic was introduced and supported by the QS agency who was actively involved in the promotion of their ranking system among national universities (Anafinova, 2020). Finally, GURs through affective infrastructure, including HE gatherings, virtual interactions, fact-finding visits, and fairs, shape the trends mobilizing *emoscape* binding national HE orientations for best practices (Shahjahan et al., 2020, 2022). As the affect circulates through government affective infrastructures, rankings accumulate value, furthering its power to shape policy. For example, QS rankings have eleven high-level international events scheduled in 2020–21, including QS Maple conference, Edu-Data Summit, and others that will take place online due to the COVID-19 virus (QS, 2020a).

GURs foster monocultures of institutional missions, causing deterioration of public good values and institutional autonomies and reproduce GoK. GURs establish a universal framework of comparison through rankings, such as their global rankings, but also through regional rankings (e.g., QS University Rankings by region, or THE Asia University rankings) (see Table 2). As a result, increasingly elite and selective universities need to perceive themselves as global actors, rather than just national institutions connected to local traditions and histories (Brankovic et al., 2018). Homogenizing institutional goals on the global scale at the expense of local agendas and missions, GURs inadvertently reproduce "cultural imperialism" depriving individual institutions and higher

education systems of autonomy and legitimacy in the eyes of the public stakeholders (Lloyd and Ordorika, 2021; Ishikawa, 2021). For instance, the Pontifical Catholic University in Chile increasing quest for global ranking positioning has led to increasing internationalization efforts, which clash with other university missions such as national and local community building (Veliz and Marshall, 2021). Furthermore, in a recent study of strategic plans among 78 Institutions from 33 countries, typifying high ranked, mid/low ranked, and unranked institutions, Lee et al. (2020a,b) found that top ranked HEIs are less explicit about their locally focused third mission compared to mid/low and unranked institutions. In short, GURs tend to reconfigure the mission from a national institution serving the public toward a corporate entity guided principally by global performance (Collins and Park, 2016, p. 126), that in turn help reproduce a global elite class.

GURs reorient periphery faculty success strategies toward metropolitan centers of knowledge. As Table 1 highlights, GURs conduct an institutional evaluation based on cumulative faculty research productivity and citation indices (see also Allen, 2019; Stratilatis, 2014). More specifically, the weight of indices measuring faculty research productivity ranges between 20% and 40% among the Big Three (see Table 1). Further, institutions recognize faculty with a higher number of research publications and faculty winning Nobel Prizes and Fields Medals, rather than their colleagues primarily focusing on teaching (see Table 1). The institutional evaluation of GURs results in faculty reputation hierarchization. The institutions use impact factor as a criterion in faculty hiring, promotion, distribution of research funding, and administration of institutional evaluations (Stergiou and Lessenich, 2014). As a result of reputational hierarchization and structural changes, faculty demonstrate increasing rank-seeking behavior to become more visible and consolidate individual benefits and social capital (Hazelkorn, 2018). Thus, GURs reconfigure faculty subjectivity through modification of work orientations and create a situation of competition between faculty within an institution and between faculty representing different institutions globally.

GURs undervalue research contribution of faculty publishing in non-English language regions. This academic misrepresentation constitutes the second tier of faculty hierarchization based on language differences between the center and periphery reproducing colonial/knowledge power relationships (Ordorika and Lloyd, 2015). Faculty hierarchization becomes further exacerbated through disciplinarian relatedness of faculty to “internationally competitive natural sciences” and diminishing “parochial social sciences” (Ishikawa, 2009, p. 170). In particular, unlike their colleagues in STEM, faculty in the social sciences and humanities are stigmatized as less technologically oriented, scientifically inefficient, stifling international reputation and return on investment (Lloyd and Ordorika, 2021; Ishikawa, 2021; Lo and Liu, 2021). The hierarchization of academic disciplines shifts the balance of power and distribution of funding between natural sciences/engineering and humanities/social sciences faculty. Taken together, GURs’ approach disadvantages faculty publishing in the national languages and in social sciences and humanities. These phenomena have been defined as subaltern positions (Grosfoguel, 2002) or subaltern sides of colonial difference (Delgado et al., 2000). Faculty in the subaltern positions experience unconscious self-colonization, submit to the dominant paradigm, and unquestioningly adapt knowledge to local narrow utilitarian needs (Mignolo and Tlostanova, 2012).

Institutional leaders confirm they take rankings seriously and incorporate them into strategic planning processes, institutionalizing principles of GoK. University leadership are rarely just passive recipients of rankings, but they monitor and leverage them to their strategic advantage, especially when the ranking benefits their institutional visibility (Veliz and Marshall, 2021). For instance, 84% of surveyed higher education institutions from 41 different countries have established a formal internal mechanism to monitor institutional rank on a regular basis (Hazelkorn, 2008). Furthermore, in 40% of these cases, the Vice Chancellor, President, or Rector coordinated the tasks (Hazelkorn, 2008). In addition to the input metrics, administrators create narratives legitimizing and manipulating the rankings to better position institutions in the global lists (Heffernan and Heffernan, 2018). In some cases, the institutional leaders had to pay attention to rankings since the institutional performance in rankings influenced their salary, career growth, or justified claims on resources (Rauhvargers, 2013). While rankings have no norm-driving authority over university officials, they increasingly frame institutional policy making (Erkilla, 2014). Thus, university officials further reinstall policy scripts of GOK through marketing, institutional policies, and information sharing using GURs.

Media and mediatization principles operating in the context of GURs represent an instrument significantly contributing to the reproduction of GoK. Mass media heavily advertise GURs globally (Stergiou and Lessenich, 2014). Furthermore, in an era of the digital revolution, mass media provide instant planetary communication of ranking information among major stakeholders (Shahjahan et al., 2020, 2022). The global impact of GURs is based upon the use of several mediatization principles such as: (a) temporal aspect of repeated publication, (b) simplicity of visual order and quantified comparisons of league tables, and (c) regular publications in different social fields on a continuous basis (Brankovic et al., 2018; Stergiou and Lessenich, 2014; Yudkevich et al., 2015). Furthermore, national media play an important role in localizing GURs, allowing GURs to persevere despite their questionable legitimacy in various national contexts, thus globalizing higher education policy (Barats, 2020; David, 2016; Shahjahan et al., 2021). Through the mediatization principles, GURs artificially construct a social situation of competition between universities for the status of global excellence, which is visible, easily comparable, and continuously dynamic.

Finally, HEIs contribute to media circulation of rankings through their branding efforts (e.g., university websites, social media, and brochures) that reproduce specific policy scripts (Stack, 2016). HEIs are not only objects of ranking evaluation, but they find opportunities to regain some power by creating discourses to represent themselves in a positive light (Heffernan and Heffernan, 2018; Lo and Liu, 2021). HEIs are increasingly contracting commercial rankers such as THE and QS for services to improve their rankings, such as: advertising, access to additional ranking data and analytics, consulting (typically for branding, quality-improvement, and reporting strategies), and access to events and workshops (Chirikov, 2021; Shahjahan et al., 2020, 2022). To this end, for instance, Russian universities spent \$3,871,378 on QS services during the period between 2013 and 2020 (Chirikov, 2021, p. 5). Universities also engage in language games to counterbalance the scientific data of rankings through legitimizing and

manipulating their ranking (Brankovic et al., 2018; Heffernan and Heffernan, 2018). For instance, the National University of Singapore ranked 24th in the *THE* world rankings represented itself as “Top in Asia,” and repositioned institutional ranking as being number one at the regional level (National University of Singapore, 2016). Unfortunately, in some cases, HEIs seek to “game the system” and engage in unethical practices to improve their public image (Ordorika and Lloyd, 2015). By adopting ranking language games, HEIs nurture the second cycle of ranking mediatization by further intensifying global competition not only among already ranked institutions but also among all non-participating institutions in rankings.

Concluding remarks

As we demonstrated, GURs have introduced new complex dynamics into global higher education that stifle attaining diversity and global justice in education. As the HE sector continues to be globalized (e.g., increased connections, exchanges, and movements, albeit unevenly distributed), and the sector becomes more complex and difficult to tame and name, GURs become convenient tools to inform decision making across stakeholders and scales. Many “crave” the simplicity and certainty such rankings provide. Yet such “cravings” are not simply natural or inevitable. Such desires for rankings are historically contingent, political, and are produced by various actors, including rankers themselves. Given its unprecedented attention, we demonstrated how GURs reproduce GoK in two ways. First, they perpetuate GoK through universalizing particular languages of excellence and quality, ranging from the English language in scholarship and teaching and learning to policy by numbers, and/or proxies of quality. Second, GURs reinforce existing hierarchical knowledge systems and produce subjectivity intertwined with colonial/knowledge power relationships. We highlighted how HE stakeholders (e.g., government, institutions, faculty, leadership, and students) acquiesce to such GURs, as GURs provide an orientation to be followed and perceived as successful. In short, GURs perpetuate the GoK by normalizing a principle of calculability (Welsh, 2019), reconfiguring social relations and knowledge practices, and shaping subjectivity using the enticing soft-power of modernity/coloniality (Lo, 2011; Shahjahan et al., 2017).

The GoK frame allows us to foreground the colonial logics underpinning the *principle* of rankings (Shahjahan et al., 2017; Welsh, 2019) that has grave consequences for attaining global justice in education. As we demonstrated, coloniality/modernity assumptions such as universality, causality, rationality, and Enlightenment notions of progress, underlie the practice of GURs. A GoK frame gestures us toward raising questions about ontological securities, status, and privileges afforded to us via HEIs and their rankings. It also cautions against universalist tendencies of such solutions that trap us back in modernity logics (Shahjahan et al., 2017), that have proved historically to have grave consequences for the colonized Other. GURs not only impact the structural changes in global HE, but they also have important consequences for ontological governance, in constituting and reproducing particular ways of knowing and being, i.e., a calculating way of being (Shahjahan, 2020). For instance, the use of GURs raises the question: why do we increasingly crave and depend on selective “data” to orient our realities and decision making? As such, GURs simply and complexly reproduce a way of knowing and being that views quantifiable “data” as a means to secure worthiness and assurance in the quest to control one’s futurity, and is embedded in affective and short-term temporal investments.

Finally, in contesting the global politics of knowing and being in contemporary global HE, we cannot ignore GURs, but we cannot bend to its will either. Our analysis of GURs from the GOK standpoint highlights the importance of the colonial knowledge/power relations constituting global politics of knowing and being. By “attending to what is habitual and routine in ‘the what’ of the world” (i.e. rankings, or notions of excellence or quality, or “world class”), our analysis highlights how we continue to be “stuck” in colonial/modernity paradigms, and to decolonize education, requires moving beyond “wishing for new tricks” (Ahmed, 2007, p. 165). Amid globalization of HE tied to asymmetrical movements, networks, and untethered economies furthering colonial power relations and cultural imperialism (Sidhu, 2006; Shahjahan, 2019), the global politics of knowing and being cannot be separated from the global divisions of labor, racial/ethnic hierarchies, and hegemonic modern epistemologies. GURs hinder global justice in education through establishing and maintaining reputation hierarchies, linguistic and social-ideological biases, and concentrations of financial resources. The standard of humanity has shifted from a colonial era’s theological model (Does one have a soul?) to the secular scientific model (Can one think?) to the developmental model (Is one civilized?) to today’s global model (are we world class?) perpetuating colonial difference (Shahjahan, 2016). Often conflated with diversity, democracy, and social justice projects in education, our analysis suggests moving beyond simple or drastic interruptions of the business as usual, and/or institutional/social reform focused on redistribution, recognition, or representation (Andreotti et al., 2018). Instead, to achieve “diversity” and “justice” in education, we need to challenge and address the global politics of knowing and being by denaturalizing “forms of existence framed” by modernity/colonial ways of being (Andreotti et al., 2018). In short, we suggest considering alternative ways of knowing and being, that are rendered invisible through modernity/coloniality paradigms such as the principle of rankings.

Acknowledgment

We would like to thank Ryan Allen, Hoda Abdelrahman, Kairat Kurakbayev, Naseeb Bhargal, and editors of this volume, for their comments on earlier drafts of this manuscript.

References

- Ahmed, S., 2007. A phenomenology of whiteness. *Fem. Theor.* 8 (2), 149–168. <https://doi.org/10.1177/1464700107078139>.
- Allen, R.M., 2019. What do students know about university rankings? Testing familiarity and knowledge of global and domestic university league tables in China. *Front. Educ. China* 14 (1), 59–89.
- Altbach, P.G., Hazelkorn, E., 2018. Measuring education quality in global rankings: what's the likelihood? *Int. High. Educ.* 95, 12–14.
- Anafinova, S., 2020. The role of rankings in higher education policy: coercive and normative isomorphism in Kazakhstani higher education. *Int. J. Educ. Dev.* 78, 102246.
- Andreotti, V., Stein, S., Sutherland, A., Pashby, K., Susa, R., Amsler, S., 2018. Mobilising different conversations about global justice in education: toward alternative futures in uncertain times. *Pol. Pract.* 26, 9–41.
- Appadurai, A., 1996. *Modernity At Large: Cultural Dimensions of Globalization*, vol. 1. University of Minnesota Press, Minneapolis.
- ARWU, 2019. Methodology [Online] Academic Rankings of World Universities. Available at: <http://www.shanghairanking.com/ARWU-Methodology-2019.html>. (Accessed 19 August 2020).
- ARWU, 2020. Rankings [Online] Academic Rankings of World Universities. Available at: <http://www.shanghairanking.com/ARWU2020.html#>. (Accessed 25 August 2020).
- Avralev, N.V., Efimova, I.N., 2013. University rankings as a tool to enhance competitiveness, clustering and transnational governance of higher education in the context of globalization. *Middle East J. Sci. Res.* 16 (3), 357–361.
- Azman, N., Kutty, F.M., 2016. Imposing global university rankings on local academic culture: insights from the National University of Malaysia. In: Yudkevich, M., Altbach, P.G., Rumbley, E.R. (Eds.), *The Global Academic Rankings Game*. Routledge, Abingdon Oxfordshire, pp. 97–123.
- Barats, C., 2020. Dissemination of international rankings: characteristics of the media coverage of the Shanghai Ranking in the French press. *Palgrave Comm.* 6 (1), 1–11. <https://doi.org/10.1057/s41599-020-0440-5>.
- Brankovic, J., Ringel, T., Werron, T., 2018. How rankings produce competition: the case of global university rankings. *Z. Soziol.* 47 (4), 270–288.
- Chirikov, I., 2021. Does Conflict of Interest Distort Global University Rankings? (No. qt8hk672nh) Center for Studies in Higher Education, UC Berkeley. <https://escholarship.org/uc/item/8hk672nh>.
- Collins, F.L., Park, G.S., 2016. Ranking and the multiplication of reputation: reflections from the frontier of globalizing higher education. *High Educ.* 72 (1), 115–129.
- David, M., 2016. Fabricated world class: global university league tables, status differentiation and myths of global competition. *Br. J. Sociol. Educ.* 37 (1), 169–189. <https://doi.org/10.1080/01425692.2015.1096190>.
- Delgado, L.E., Romero, R.J., Mignolo, W., 2000. Local histories and global designs: an interview with Walter Mignolo. *Discourse* 22 (3), 7–33.
- Dembereldorj, Z., 2018. Review on the impact of world higher education rankings: institutional competitive competence and institutional competence. *Int. J. High. Educ.* 7 (3), 25–35.
- Douglass, J.A., 2014. Profiling the Flagship University Model: An Exploratory Proposal for Changing the Paradigm From Ranking to Relevancy, CSHE. University Berkeley [online]. Available at: <https://escholarship.org/uc/item/8kn1m9dz>. (Accessed 24 August 2020).
- Downing, K., Ganotice Jr., F.A. (Eds.), 2016. *World University Rankings and the Future of Higher Education*. IGI Global, Hershey.
- Dunrong, B., 2017. Global rankings and world-class university aspirations in China. In: Yudkevich, M., Altbach, P.G., Rumbley, L.E. (Eds.), *The Global Academic Rankings Game: Changing Institutional Policy, Practice, and Academic Life*. Routledge, Abingdon Oxfordshire, pp. 57–78.
- Elsevier, 2015. QS Extends Its Collaboration With Elsevier for Its World University Ranking [Online]. Available at: <https://www.elsevier.com/about/press-releases/science-and-technology/qe-extends-its-collaboration-with-elsevier-for-its-world-university-rankings>. (Accessed 19 August 2020).
- Erkkilä, T., 2014. Global university rankings, transnational policy discourse and higher education in Europe. *Eur. J. Educ.* 49 (1), 91–101.
- Fanon, F., 1968. *Black Skin, White Masks*. Grove Press, New York.
- Godwin, K.A., 2015. The counter narrative: critical analysis of liberal education in global context. *New Global Stud.* 9 (3), 223–243.
- Grosfoguel, R., 2002. Colonial difference, geopolitics of knowledge, and global coloniality in the modern/colonial capitalist world-system. *Review* 203–224.
- Grosfoguel, R., 2013. The structure of knowledge in westernised universities: epistemic racism/sexism and the four genocides/epistemicides. *Hum. Architect. J. Sociol. Self Knowl.* 1, 73–90.
- Hall, S. (Ed.), 1997. *Culture, Media and Identities. Representation: Cultural Representation and Signifying Practices*. Sage Publications.
- Hanafi, S., Boucherie, S., 2018. Discover the Data Behind the Times Higher Education World University Rankings [Online]. Available at: <https://www.elsevier.com/connect/discover-the-data-behind-the-times-higher-education-world-university-rankings>. (Accessed 19 August 2020).
- Hazelkorn, E., 2008. Learning to live with league tables and ranking: the experience of institutional leaders. *High Educ. Pol.* 21 (2), 193–215.
- Hazelkorn, E., 2018. Reshaping the world order of higher education: the role and impact of rankings on national and global systems. *Pol. Rev. High. Educ.* 2 (1), 4–31.
- Heffernan, T.A., Heffernan, A., 2018. Language games: university responses to ranking metrics. *High Educ. Q.* 72 (1), 29–39.
- Ishikawa, M., 2009. University rankings, global models, and emerging hegemony: critical analysis from Japan. *J. Stud. Int. Educ.* 13 (2), 159–173.
- Ishikawa, M., 2021. Between local distinction and global reputation: university rankings and changing employment in Japan. In: Stack, M. (Ed.), *Global University Rankings and the Politics of Knowledge*. University of Toronto Press, pp. 153–171.
- Kaba, A.J., 2012. Analyzing the Anglo-American hegemony in the "times higher education" rankings. *Educ. Pol. Anal. Arch.* 20 (21), 1–53.
- Kehm, B.M., 2014. Global university rankings—impacts and unintended side effects. *Eur. J. Educ.* 49 (1), 102–112.
- Kim, D., Song, Q., Liu, J., Liu, Q., Grimm, A., 2018. Building world class universities in China: exploring faculty's perceptions, interpretations of and struggles with global forces in higher education. *Compare* 48 (1), 92–109.
- King, R., 2009. *Governing Universities Globally: Organizations, Regulation and Rankings*. Edward Elgar Publishing.
- Kiraka, R., Maringe, F., Kanyutu, W., Mogaji, E., 2020. University league tables and ranking systems in Africa: emerging prospects, challenges and opportunities. In: Mogaji, E., Maringe, F., Hinson, R. (Eds.), *Understanding the Higher Education Market in Africa*. Routledge Studies in Marketing. Routledge, Abingdon Oxfordshire.
- Lee, J.T., Naidoo, R., 2020. Complicit reproductions in the global south: courting world class universities and global rankings. In: Rider, S., Peters, M.A., Hyvönen, M., Besley, T. (Eds.), *World Class Universities*. Springer, Singapore, pp. 77–91.
- Lee, J., Liu, K., Wu, Y., 2020a. Does the Asian catch-up model of world-class universities work? Revisiting the zero-sum game of global university rankings and government policies. *Educ. Res. Pol. Pract.* 1–25.
- Lee, J.J., Vance, H., Stensaker, B., Ghosh, S., 2020b. Global rankings at a local cost? The strategic pursuit of status and the third mission. *Comp. Educ.* 56 (2), 236–256. <https://doi.org/10.1080/03050068.2020.1741195>.
- Lewis, R., 2009. Quality assurance in higher education – its global future. In: OECD (Ed.), *Higher Education to 2030, Globalisation*, vol. 2. OECD Publishing, Paris, pp. 323–355. <https://doi.org/10.1787/9789264075375-13-en>.
- Lim, M.A., Williams Ørberg, J., 2017. Active instruments: on the use of university rankings in developing national systems of higher education. *Pol. Rev. High. Educ.* 1 (1), 91–108.
- Lloyd, M., Ordorika, I., 2021. International university ranking as cultural imperialism: implications for the global south. In: Stack, M. (Ed.), *Global University Rankings and the Politics of Knowledge*. University of Toronto Press, pp. 25–49.
- Lo, W.Y.W., Liu, S., 2021. Are global university rankings still important? Perspectives from greater China. In: Hazelkorn, E. (Ed.), *Research Handbook on University Rankings: History, Methodology, Influence and Impact*. Edward Elgar, Cheltenham.

- Lo, W.Y.W., 2011. Soft power, university rankings and knowledge production: distinctions between hegemony and self-determination in higher education. *Comp. Educ.* 47 (2), 209–222.
- Mäkinen, S., 2021. Global university rankings and Russia's quest for national sovereignty. *Comp. Educ.* 1–18.
- Marope, P.T.M., Wells, P.J., Hazelkorn, E., 2013. Rankings and Accountability in Higher Education: Uses and Misuses. *E-book library* [online]. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000220789>. (Accessed 25 August 2020).
- Mignolo, W., Tlostanova, M., 2012. Knowledge production systems. In: *The Encyclopedia of Global Studies*. Sage Publications, Thousand Oaks, pp. 1005–1010.
- Mignolo, W., 2002. The geopolitics of knowledge and the colonial difference. *S. Atl. Q.* 101 (1), 57–96.
- Mignolo, W., 2003a. Globalization and the geopolitics of knowledge: the role of the humanities in the corporate university. *Nepantla Views South* 4 (1), 97–119.
- Mignolo, W., 2003b. *The Darker Side of the Renaissance: Literacy, Territoriality, and Colonization*. University of Michigan Press, Ann Arbor.
- Mignolo, W.D., 2011. *The Darker Side of Western Modernity*. Duke University Press Books, Durham, NC.
- Nandy, A., 1983. *The Intimate Enemy: Loss and Recovery of Self Under Colonialism*. Oxford UP, Delhi.
- National University of Singapore, 2016. Nus is Top Asian University and 24th in the World in Latest Times Higher Education World University Rankings [Online]. Available at: <https://news.nus.edu.sg/press-releases/nus-top-asian-university-and-24th-world-latest-times-higher-education-world>. (Accessed 19 August 2020).
- Ntshoe, L., Letseka, M., 2010. Quality assurance and global competitiveness in higher education. In: Portnoi, L.M., Rust, V.D., Bagley, S.S. (Eds.), *Higher Education, Policy, and the Global Competition Phenomenon*. Palgrave Macmillan, New York, pp. 59–71.
- Ordorika, I., Lloyd, M., 2015. International rankings and the contest for university hegemony. *J. Educ. Pol.* 30 (3), 385–405.
- Pavlyutkin, I., Yudkevich, M., 2016. The ranking game on the Russian battlefield: the case of the higher school of economics. In: Yudkevich, M., Altbach, P.G., Rumbley, L.E. (Eds.), *The Global Academic Rankings Game*, first ed. Routledge, Abingdon Oxfordshire, pp. 171–194.
- Q.S., 2020a. Higher Education Conferences [Online]. QS Topuniversities. Available at: <https://www.qs.com/solutions/events/higher-education-conferences/>. (Accessed 24 August 2020).
- QS, 2019a. About US [Online]. QS. Available at: <https://www.qs.com/about-us/>. (Accessed 19 August 2020).
- QS, 2019b. Quacquarelli Symonds, Ltd [Online]. QS. Available at: <https://www.qs.com>. (Accessed 19 August 2020).
- QS, 2020b. Qs World University Rankings Methodology [Online]. QS Topuniversities. Available at: <https://www.topuniversities.com/qs-world-university-rankings/methodology>. (Accessed 19 August 2020).
- QS, 2020c. University Rankings [Online]. QS Topuniversities. Available at: <https://www.topuniversities.com/university-rankings>. (Accessed 25 August 2020).
- Rauhvargers, A., 2013. *Global University Rankings and Their Impact: Report II*. European University Association, Brussels.
- Rhee, J.E., 2009. International education, the new imperialism, and technologies of the self: branding the globally educated elite. *Multicul. Educ. Rev.* 1 (1), 55–82. <https://doi.org/10.1080/23770031.2009.11102863>.
- Rizvi, F., Lingard, B., 2010. *Conceptions of Education Policy. Globalizing Education Policy*. Routledge, Oxon.
- Ruby, A., Kuzhabekova, A., Lee, J., 2017. Recruitment of international faculty: the case of Nazarbayev university. In: Yudkevich, M., Altbach, P.G., Rumbley, L.E. (Eds.), *International Faculty in Higher Education Comparative Perspectives on Recruitment, Integration, and Impact*. Routledge, New York, pp. 150–173.
- Sabzalieva, E., 2017. The policy challenges of creating a world-class university outside the global 'core'. *Eur. J. High Educ.* 7 (4), 424–439.
- Salmi, J., 2009. *The Challenge of Establishing World-Class Universities*. E-book library [online]. Available at: <https://openknowledge.worldbank.org/bitstream/handle/10986/2600/476100PUB0FREN10101OfficialUseOnly1.pdf?sequence=1>. (Accessed 25 August 2020).
- Schmitz, S., 2014. Affect/emotion: orientation matters. A conversation between Sigrid Schmitz and Sara Ahmed. *FZG–Freiburger Zeitschrift für Geschlechter Studien* 20 (2), 97–108. <https://doi.org/10.3224/fzg.v20i2.17137>.
- Shahjahan, R.A., Morgan, C., 2016. Global competition, coloniality, and the geopolitics of knowledge in higher education. *Br. J. Sociol. Educ.* 37 (1), 92–109.
- Shahjahan, R.A., Blanco-Ramirez, G., Andreotti, V., 2017. Attempting to imagine the unimaginable: a decolonial reading of global university rankings (GURs). *Comp. Educ. Rev.* 61 (1), 51–73.
- Shahjahan, R.A., Sonneveldt, E.L., Estera, A.L., Bae, S., 2020. Emoscapes and commercial university rankers: the role of affect in global higher education policy. *Crit. Stud. Educ.* 1–16.
- Shahjahan, R.A., Estera, A.L., Bae, S., Sonneveldt, E.L., 2022. Imagining 'Asian' higher education: visual campus gaze and global university rankings (GURs) websites. *Compare: A Journal of Comparative and International Education*, 52 (1), 129–146.
- Shahjahan, R.A., Bylsma, P.E., Singai, C., 2021. Global university rankings as 'sticky' objects and 'refrains': affect and mediatization in India. *Comp. Educ.* 1–18. <https://doi.org/10.1080/03050068.2021.1935880>.
- Shahjahan, R.A., 2013. Coloniality and a global testing regime in higher education: unpacking the OECD's AHELO initiative. *J. Educ. Pol.* 28 (5), 676–694.
- Shahjahan, R.A., 2016. International organizations (IOs), epistemic tools of influence, and the colonial geopolitics of knowledge production in higher education policy. *J. Educ. Pol.* 31 (6), 694–710.
- Shahjahan, R.A., 2019. From 'geopolitics of being' towards inter-being: envisioning the 'in/visibles' in the globalization of higher education. *Youth Globaliz.* 1 (2), 282–306. <https://doi.org/10.1163/25895745-00102005>.
- Shahjahan, R.A., 2020. On a 'calculating way of being': the perils of datafication on re/connecting community. In: *Keynote Presented at the Annual Meeting of European Educational Research Association. EERA, Glasgow, UK*.
- Shattock, M., 2017. The 'world class' university and international ranking systems: what are the policy implications for governments and institutions? *Pol. Rev. High. Educ.* 1 (1), 4–21.
- Shin, J.C., Harman, G., 2009. New challenges for higher education: global and Asia-Pacific perspectives. *Asia Pac. Educ. Rev.* 10 (1), 1–13.
- Sidhu, R.K., 2006. *Universities and Globalization: To Market, to Market*. Routledge.
- Smith, L.T., 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books Ltd.
- Souto-Otero, M., Enders, J., 2017. International students' and employers' use of rankings: a cross-national analysis. *Stud. High Educ.* 42 (4), 783–810.
- Stack, M.L., 2013. The Times Higher Education ranking product: visualising excellence through media. *Glob. Soc. Educ.* 11 (4), 560–582.
- Stack, M., 2016. Who is watching the watchdogs? The business of rankings. In: Stack, M. (Ed.), *Global University Rankings and the Mediatization of Higher Education*. Palgrave Macmillan, London, pp. 31–50.
- Stack, M., 2020. Academic stars and university rankings in higher education: impacts on policy and practice. *Pol. Rev. High. Educ.* 4 (1), 4–24.
- Stergiou, K.I., Lessenich, S., 2014. On impact factors and university rankings: from birth to boycott. *Ethics Sci. Environ. Polit.* 13 (2), 101–111.
- Stratilatis, C., 2014. University rankings and the scientification of social sciences and humanities. *Ethics Sci. Environ. Polit.* 13 (2), 177–192.
- THE Reporter, 2019. New Owner for Times Higher Education [online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/news/new-owner-times-higher-education>. (Accessed 19 August 2020).
- THE, 2014. Times Higher Education Announces Reforms to Its World University Rankings [online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings/news/times-higher-education-announces-reforms-to-world-university-rankings>.
- THE, 2019. About Us [online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/about-us>.
- THE, 2020a. The World University Rankings 2020: Methodology [online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2020-methodology>. (Accessed 25 August 2020).

- THE, 2020b. World University Rankings [online]. The World University Rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings>. (Accessed 25 August 2020).
- Véliz, D., Marshall, P., 2021. Influence of global rankings on strategic planning from the perspective of decision-makers: A case study of a Chilean research university. *High Educ. Q.* <https://doi.org/10.1111/hequ.12333>.
- Welsh, J., 2019. Ranking academics: toward a critical politics of academic rankings. *Crit. Pol. Stud.* 13 (2), 153–173.
- Yudkevich, M., Altbach, P.G., Rumbley, L.E., 2015. Global university rankings: the “Olympic Games” of higher education? *Prospects* 45 (4), 411–419.

Truth and then reconciliation research: an emerging field of educational studies

Lisa Howell and Nicholas Ng-A-Fook, University of Ottawa, Ottawa, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

An age of apology: sorry as a neoliberal paradigm	273
Neoliberalism and redress: historicizing past injustice	273
Truth and then reconciliation at the limits of settler colonialism	274
Beyond apologies: truth before reconciliation within higher education	275
#Reconciliation: not a lesson plan	275
What we have learned: the perfect stranger and the contradictions of Canadianness	276
Toward understanding colonialism as culture, unlearning and ethical relationality	277
Professional unlearning opportunities: truth and then reconciliation with teachers	278
Restoring truths and then reconcilia(c)tion as future educational research	279
References	280

Discussions within the field of recognition and reconciliation politics, however, rarely treat emotions like anger and resentment even-handedly.

Coulthard (2014, p. 111)

To do this, we need a new social contract for education that can repair injustices while transforming the future.

UNESCO (2021)

What is truth? And, who gets to story historical truths? Or, even contemporary truths for that matter? Several nation-states have created a discursive regime supported by their educational systems and respective curriculum to teach us how to forget historical truths. The media and politicians spin truths to serve their versions of the public good. And yet, several Indigenous Peoples across the world continue to experience the intergenerational historical and contemporary harms of European settler colonial violence in Australia, Brazil, Canada, New Zealand, Australia, South Africa, and United States. That is a truth. In turn, Indigenous Peoples¹ have been strategically calling for settler colonial governments and their citizens to recognize and acknowledge the ongoing cultural, ecological, economic, emotional, material, and psychological legacies of such historical and contemporary harms. And yet, settler nation states have been slow to recognize, respond, or react in terms of their different relational commitments and/or promises to different Indigenous Peoples. And when “they” do, it often begins with official state sponsored apologies followed by commitments to policy enactments of truth, and then reconciliation education.

Part of such “a contemporary international trend,” is as Czyzewski (2011) writes, “the notion that a more inclusive citizenship, as well as reconciliation, can be achieved through truth-telling and subsequent public education about ‘past’ injustices” (p. 1). Indeed, truth-telling, redressing history harms, and public education have become overarching objectives of various international Indigenous-settler reconciliation initiatives, including those of the TRC in Canada (2015), the TRC in South Africa (Regan, 2010), the Waitangi Tribunal in Aotearoa, New Zealand (Joseph, 2008), and the Human Rights and Equal Opportunity Commission in Australia (Simon, 2013). In Canada, First Nations, Inuit, and Métis Survivors of colonial injustices have long-called for the truths that they have courageously shared to be taught in schools across the country. In response, there is a growing reconceptualization of how historical truths “ought” to be taught in response to such historical harms (Regan and Craft, 2020; Madden, 2019; Wallace-Casey, 2022; Llewellyn and Ng-A-Fook, 2020). And yet, how might such truth-telling be re-storied as part of a new social contract for education?

Today, several government institutions and citizens in different parts of the world continue to deny or distance themselves from the violent history and legacy of settler colonial harms. Such active distancing from historical “truths” affords the intergenerational beneficiaries of settler colonialism, including some educators, to superficially recognize the lived experiences of Indigenous Peoples (Coulthard, 2014; Ng-A-Fook, in-press). “The contemporary context of reconciliation is also an opportunity,” as Stein reminds us, “to reconsider what the Canadian university should or could be, including in ways that go beyond defenses of the ‘public good’ of higher education if that ‘public’ is still presumed to be the citizenry of an otherwise unchanged colonial society” (p. 159). Consequently, there is much work to be done within and across different educational systems. In response, what we have witnessed since the turn of the 21st century, is an emerging field of educational studies that has and is focusing on truth and then reconciliation education. Several of these scholars and, their educational research, seek to create curricular and pedagogical opportunities for living a new social contract toward repairing historical and contemporary injustices while transforming and reimagining our future relations

¹We use the term Indigenous internationally, “First Nations, Inuit, and Métis” to describe people who are Indigenous to Canada, and specific names of Nations and communities when possible. The word “Indigenous,” though currently in popular usage in Canada, is latin-derived and not Indigenous itself (Peters and Mika, 2017). Furthermore, the term Indigenous tends to homogenize rather than recognize. We have kept the term Indigenous when it is used in a citation or reference.

beyond settler colonial moves to innocence. In this chapter then, we offer a partial and situated overview of the possibilities and limitations of such scholarship in response to what others have called a neoliberal *Age of Apology* here in Canada (Henderson and Wakeham, 2013). We conclude with a brief discussion about different research taking place among A Canadian Curriculum Theory Project research team in response to the emerging field of truth, and then reconciliation education research.

An age of apology: sorry as a neoliberal paradigm

Responding to the growing demand from communities for redress, states have increasingly relied on transitional justice mechanisms—such as truth commissions—to engage in action of performative morality as they seek to contend with their own histories of violence.

George (2020, p. 87)

Since the end of the second world war, political apologies have become increasingly common (Organick, 2020), with the *Institute for the Study of Human Rights* (2022) recording 815 political apologies during this time period. The fact that 79% of these apologies have been issued since 1996 has prompted many to refer to the last 25 years as “the age of apology” (Sanz, 2012, p. 3). Political apologies are referred to as state apologies, national apologies, historical apologies and a myriad of other terms. They are often made by politicians on “behalf,” as Sanz (2012) suggests, “of and directed to identifiable communities for past wrongdoings by the modern-state” (p. 10). State sanctioned apologies seek to address reparations of past harms (Edwards, 2010), re-righting past beliefs and their corresponding harmful actions (Nobles, 2008), and to promote peace, mutual confidence, and stabilization (Sanz, 2012). While many governments offer state sanctioned apologies, others in turn also establish commissions, which have focused on restorative justice, accountability, human rights, and compensation (Ibhawoh, 2019). Indeed, over the past three decades, more than forty countries have established truth (and reconciliation) commissions such as, but not limited to, South Africa, Australia, Sierra Leone, Chile, the Philippines, South Korea, Canada, and Rwanda (Ibhawoh, 2019). In the aftermath of the 1994 Rwandan genocide for example, in which an estimated 800,000 members of the minority Tutsi population were killed by Hutu extremists, the government attempted to promote a policy of national unity and reconciliation (Jessie, 2012). However, as Jessie (2012) cautions, many survivors feel that:

... reconciliation with their neighbours, many of whom were allegedly complicit, if not directly responsible, for the harms visited upon them during the genocide, is impossible. In particular, survivors argue that there can be no reconciliation until those responsible for murdering their missing loved ones reveal the locations where their bodies have been deposited so they can be identified and reburied with respect according to Rwandan tradition. (p. 7)

The ongoing reconciliatory challenges in Rwanda remind us that lived experiences of the intergenerational survivors of genocide must be central to truth, and then ensuing reconciliation efforts. However, in many ways, this current age of apology and the very “ideas” of truth and then reconciliation have become a social paradigm that often fails to directly address the harms and concerns of survivors (Henderson and Wakeham, 2013). In turn, settler movements for redress have become entrenched in 21st century frameworks of neoliberal settler colonial democracies.

Neoliberalism and redress: historicizing past injustice

Liberal reconciliation frames are designed to settle the settler within, without, in fact, requiring the settler to acknowledge themselves as a settler or to reflect on a relational and inherited coloniality with Indigenous peoples.

MacDonald (2020, pp. 7–8)

Truth and reconciliation commissions have sought to address genocide and crimes against humanity by settler-nation states². Thus, state apologies are often officiated and announced by governments in relation to their past actions by, and are therefore orchestrated and informed by intergenerational neoliberal settler colonial worldviews. Such movements, as McCready (2009) writes, are a “cultural force that perpetuates individualism, cynicism and the frantic guarding of social privilege” (p. 163). Moreover, he reminds us that in Canada specifically, neoliberalism has manifested in multicultural policies that deflect the demands of marginalized groups, emerging as a “highly effective governmental technology for managing difference” (p. 166). In relation to this nation-state context, discourses of “racelessness” perpetuate what Goldberg (2007) defines as the “attempt to go beyond-without (fully) coming to terms with-racial histories and their accompanying racist inequities and inequities” (p. 217). Thus,

²In this article, the terms “settler” and “settler Canadian” are used to describe peoples who came to Canada with agency and choice, to benefit economically from the land through different past and present colonial policies. Therefore, Black Canadians, as the descendants of enslaved peoples, as well as immigrants and refugees fleeing war, discrimination, injustices, and oppressions are not constituted as settler colonial Canadians. For more discussion on this, see Thomas (2019).

neoliberalism defines and limits the ways that reconciliation, redress, and future relations between Indigenous peoples and Canadians are able to take (McCready, 2009). “Expressing sympathy for a historic wrong comfortably locates colonialism in the past rather than the present and is,” as McCready (2009) maintains, “an act that confirms a dearly held narrative of the evolution of settler Canadian attitudes from regrettable original violence, through past “misunderstandings” and unfortunate “mistakes” to present day post-racial equity” (p. 167). Canadians have continuously witnessed the historicizing of past wrongs by successive federal governments. Over the past 20 years alone, Canada has issued 39 national apologies for historical injustices, including for the internment of Japanese-Canadians during the Second World War, the head tax imposed on Chinese immigrants between 1885 and 1923, and for the 1914 Komagata Maru incident, in which several hundred Indian immigrants were turned away from Vancouver and returned to India, where some were killed, and others arrested (Institute for the Study of Human Rights, 2021).

More recently, the Canadian government’s apologies have been focused on reconciliation with Indigenous peoples, including official apologies for the Indian Residential Schooling system,³ the mistreatment of Inuit with tuberculosis in the mid-20th century, and for the forced relocations of hundreds of Inuit people during the 1950s. On June 25, 2021, Prime Minister Justin Trudeau once again apologized for government policies that forced First Nations, Inuit, and Métis children to attend the Residential Schooling system. This latest apology came after the confirmation of 751 unmarked graves at the site of the former Marieval Indian Residential School on Cowessess First Nation. This occurred less than two months after the remains of 215 children were identified and recovered from the site of the former Kamloops Residential School in Tk’emlups te Secwépemc First Nation (Leichnetz, 2021). In his ensuing public address, Prime Minister Trudeau apologized to members of the Cowessess community and Treaty Four communities, stating that:

We are sorry. It was something that we cannot undo in the past, but we can pledge ourselves every day to fix in the present and into the future. That means recognizing the harms, the impacts, the inter-generational trauma, the cycles of challenges that far too many Indigenous peoples face in this country because of actions that the federal government and other partners deliberately and willingly undertook.

Cpac (2021, 03:36)

His apology was immediately met with calls for accountability and action, including from Pam Palmater, a Mi’kmaq lawyer and professor, and author of *Warrior Life: Indigenous Resurgence and Resistance*. Palmater reminded Canadians, “Apologies are important for those who need healing or closure ... but it’s also about providing financial supports, human resources, equipment, and inter-jurisdictional cooperation that will be needed to identify, locate, assess and return these children home” (Gilmore, 2021, paras. 10–12). In other words, reconciliatory gestures such as apologies must include meaningful financial and political commitments toward future policies and actions that result in redress, equity and justice for those harmed.

Truth and then reconciliation at the limits of settler colonialism

The mobilization of truth and reconciliation by the settler colonial state forces us to question their capacity to creatively build and envision decolonial Indigenous futures.

George (2020, p. 108)

While many Canadians have largely accepted symbolic forms of redress, such as government apologies, many remain aloof—even hostile—to calls for contemporary justice and equity (McCready, 2009). For example, when former Conservative Senator Lynn Beyak stood up on behalf of settler colonialism to mount an abhorrent defense of the Residential Schooling system in the Canadian Senate, many Canadians wrote to her office in support of her racist beliefs (Howell and Ng-A-Fook, 2022). These letters, which were published on her Senate of Canada website expressed “harmful anti-Indigenous stereotypes and tropes of an imagined ‘lazy Indian who doesn’t pay taxes,’ the ‘corrupt chief who misspends federal dollars,’ and the ‘chronic whining of people who can’t get over the past’” (Howell and Ng-A-Fook, 2022, p. 11). It took 2 years for the Senate Ethics commissioner to order the letters removed, and a further 2 years for Beyak to resign (Tasker, 2021).

Residential School denialism and its denialists, such as but not limited to Senator Beyak, is not limited to politicians. “Residential school denialists,” write Heath Justice and Carleton (2021) “include missionary apologists, writers and academics, right-wing and anti-Indigenous editorialists and relatives of residential school staff who uncritically refer to personal memory and work to defend their family reputations” (para 7). Justice Murray Sinclair has said that residential school denialism in Canada is on the rise, cautioning that different political and/or educational stakeholders are “rushing reconciliation and leaving truth behind” (Heath Justice and Carleton, 2021, para. 9). Consequently, as Heath Justice and Carleton (2021) reminds us, truth must come before we can address reconcilia(c)tions together inside, or outside, the contexts of the public schooling curriculum.

³Please note that the Indian Residential Schooling System is capitalized and referred to as a schooling system throughout the article to place emphasis on the fact that it was a system, with systemic, intergenerational, and ongoing impacts on First Nations, Inuit, and Métis survivors, families, and communities.

Beyond apologies: truth before reconciliation within higher education

The context of reconciliation has created an ambivalent opportunity for addressing the enduring coloniality of Canadian higher education.

Stein (2019, p. 168)

We are now close to 7 years since the close of *Canada's Truth and Reconciliation Commission* (TRC), and currently, only 11 of the 94 Calls to Action have been completed (Jewell and Mosby, 2021). Moreover, despite the recent litany of apologies, public outpourings of grief, the half-mast flags, and the establishment of a National Day of Truth and Reconciliation in 2021, settler colonialism continues to pervade political, social, and legal aspects of Canadian public policy and parlance. In fact, our apologetic Prime Minister spent the first National Day of Truth and reconciliation on the beaches of Tofino, British Columbia, rather than with residential school survivors and communities. He subsequently apologized for this decision, explaining that he made a regrettable mistake (Tunney, 2021).

Systemic changes are needed to move beyond the age of apology (Blackstock, 2021; Joseph, 2018). One has to look no further than to the vast overrepresentation of First Nations, Inuit and Métis peoples who are incarcerated, impoverished, food insecure, un/underemployed, and affected by suicide, violence, the child welfare system, addictions, and intergenerational trauma caused by colonialism to understand this. Chief Commissioner Murray Sinclair has often called education the "key to truth and reconciliation" stating that education is both the cause for, and at the heart of, addressing "truth," and sustaining "reconciliation" across the school curriculum in Canada (TRC, 2015). Education has therefore once again become the site where transformative change might occur. As non-Indigenous teacher educators and researchers, how the TRC *Calls to Action* for curricular and pedagogical transformation are being addressed, or not, within Faculties of Education is a significant ongoing inquiry for us to question, deconstruct, and reconstruct together.

Since 2015, universities, school districts, schools, education workers, teachers, and administrators have mobilized to respond to the Calls put forth by the TRC, specifically calls 62 and 63, which directly call upon provincial and territorial ministries, faculties of education, and K-12 administrators and teachers to integrate Indigenous knowledge and pedagogies across the school curriculum and teacher education programs. Although many provinces and territories have implemented curriculum that responds to the TRC Calls to Action (Kairos, 2018), educators often report that they do not have the skills, knowledge, or confidence to teach about the history of the Indian Residential Schooling system, Indigenous knowledges, or reconciliation (Dion, 2014; Howell, 2017; Madden, 2019; Smith, 2017). Recognizing that teacher education is crucial in this process, faculties of education across the country began to make commitments toward addressing truth, and then reconciliation (Cote-Meek, 2018; Madden, 2019; Universities Canada, 2015). Many of these commitments have included mandatory teacher education courses in First Nations, Inuit, and Métis perspectives, as well as recruiting and supporting Indigenous faculty, increasing Indigenous knowledge in courses, and in addressing institutional racism and settler privilege (Association of Canadian Deans of Education, 2017). Subsequently, a new field of study and research in truth and reconciliation education has emerged, led by Indigenous and non-Indigenous scholars, researchers, and teacher educators in areas of curriculum, teacher education, historical consciousness, Indigenous education, and teacher professional learning. We will examine some of the research that has emerged from this field in recent years later in this paper.

Despite the ongoing work of educators and scholars, settler colonial conceptions of reconciliation remain contested by several Indigenous scholars, individuals, and communities because of the ongoing daily non-Indigenous institutional, societal, and individual acts of anti-Indigenous racism (Coulthard, 2014; Refugee research, 2017; Vowel, 2016; Wakeham, 2019). Teachers and leaders in educational institutions must therefore continue to question and address how teacher education programs continue to be largely framed by settler colonial worldviews, histories and perspectives (Howell et al., 2021). In this section we briefly describe some of these limitations.

#Reconciliation: not a lesson plan

Reconciliation is more than a single event. To be fully realized, everyone will need to engage in genuine forms of learning, relationship building, and advocacy.

MacDonald and Markides (2018, p. 230)

Marie Battiste a prominent Mi'kmaq scholar in Canada, insists that pervasive of Eurocentrism within the higher education and public schooling system facilitates the exclusion of Indigenous peoples. In *Decolonizing Education: Nourishing the Learning Spirit*, Battiste (2013) calls on Settler Canadians to:

... imagine the consequences of a powerful ideology that positions one group as superior and gives away First Nations peoples' lands and resources and invites churches and other administrative agents to inhabit their homeland, while negating their very existence and finally removing them from the Canadian landscape to land no one wants. (p. 23)

The epistemological and material legacies of a settler colonial ideological state apparatus continue to reproduce and sustain national myths here in Canada, such as but not limited to “multiculturalism, peace-keeping, socially-progressive politics, and hard-earned prosperity” (Davis et al., 2017, p. 44). Non-Indigenous Canadians citizens are able to put forth stories within the public school curriculum that continue legitimize the intergenerational privileges we have incurred as settler beneficiaries (Regan, 2010). In turn, settler colonial “benevolent” and “paternalistic” stories continue to persist in our history curriculum here in Canada. They persist in government policies and politics of recognition directed at First Nations, Métis, and Inuit communities (Coulthard, 2014). It persists for several Canadian citizens in our historical public settler consciousness (Carleton, 2021; Howell and Ng-A-Fook, 2022). How might we then, challenge such historical accounts in ways that seek to trouble the fetishization of such settler colonial “benevolence” in relation to how each of us has, and continues, to benefit from such broken promises? To do so, as Battiste (2013) maintains, requires meaningful systemic educational reform that confronts the “hidden standards of racism, colonialism, and cultural linguistic imperialism in the modern curriculum” (p. 29). Our commitments toward truth and reconciliation often neglect these necessary confrontations, focusing instead on abstract notions of ways to “decolonize” and “indigenize.”

In response, Tuck and Yang (2012) remind us how many of the current conceptions of decolonization continue to be put forth as a metaphor that enable settlers to “equivocate contradictory decolonial desires” that turns “decolonization into an empty signifier to be filled by any track toward liberation” (p. 7). Likewise, Hare (2020) stresses that reconciliation in teacher education must challenge conceptions of a professional settler consciousness and move teacher candidates toward action and accountability. Otherwise, as she cautions, reconciliation will remain a “trendy topic as it has been cast in social media and scholarly writing” (p. 23). Similarly, Madden (2019) draws our attention to the hashtag “#ReconciliACTION” (p. 285). She reminds us that the TRC *Calls to Action* cannot be reduced to pedagogical approaches or “best” teaching practices, but rather a multi-dimensional approach that critiques the “construction and enactment of reconciliation” (p. 286).

In “Teaching truth before reconciliation,” Jurgens (2020) stresses that “histographies of Indigenous Nations are severely impeded by the colonial project ... within schools, millennia of Indigenous histories remain forgotten, positioned as bygone relics, as prehistory, or as folktales deemed irrelevant” (p. 123). Jurgens reminds us that educational institutions themselves bear responsibility for (re)righting these “historical amnesias” to move toward a critical historical consciousness in schools that disrupts the status quo, and whereby educators discuss “settler colonialism, Eurocentrism, racism, gender violence, and white supremacy” in their classrooms (p. 129). How might educators unlearn and learn to identify and name ongoing colonialism in the justice system, healthcare system, child welfare system, education system, and connect these abuses with past institutionalized violence toward Indigenous peoples? Here Jurgens cautions,

When we do not educate to get to the root of the Canadian problem, we foreclose opportunities to come to a true understanding of our past ... Working toward reconciliation means the classroom discourse must center on unpacking these issues underpinning the continual reproduction of harms against Indigenous peoples. (p. 130)

The thinking that these scholars bring forth teach us that the work of disrupting, unlearning⁴ and learning settler colonial world-views as part of our teacher education curriculum is not simply about transforming teaching practices, embracing decolonization strategies, or committing to certain reconcilia(c)tions. In the next section, we discuss what we have learned from some of the research done within Faculties of Education across Canada.

What we have learned: the perfect stranger and the contradictions of Canadianness

Learning about settler colonialism as an ongoing structure and set of inequitable relations is akin to coming out of a fog, where settlers begin to realize that Indigenous people have been present all along, in our schools, neighbourhoods, and cities but we settlers could not (would not) see all that we did not know. It is hidden from us and we hide it from ourselves.

Rice et al. (2022, p. 23)

Long before the TRC closing ceremonies and the Calls to Action, there were teachers and teacher educators across Canada committed to disrupting the normative narratives of Canadian history in their classrooms. Much of this early work focused on understanding the obstacles that teachers faced when learning about Indigenous perspectives and history. From this early work, a growing body of research emerged on what is often termed “teacher resistance” in regards to teaching Indigenous content and perspective (Aitken and Radford, 2018; Dion, 2014, 2016; Rice et al., 2022; Tupper, 2011, 2012). Several studies identify barriers for teachers doing reconciliation work, including a lack of confidence, perceived absence of resources, minimal time, curriculum restraints, and fear of offending and fear of cultural appropriation (Dion, 2014; Howell, 2017; Smith, 2017).

⁴While several institutions across Canada are calling for the “decolonizing” and “indigenizing” of “the academy” (Gaudry and Lorenz, 2018), these terms are both ambiguous, intricate, and raise many questions. Is it possible for Canadians to “decolonize” while living on, and profiting from the lands many of our relatives recently colonized? Moreover, what does it mean to “Indigenize” and who is doing the indigenizing? In our view, the concept of “unlearning,” as a process toward truth, and then, reconciliation, is a more explicit and accurate description of the work ahead of us. It draws attention that the work we must do is about us.

Susan Dion is a Potawatomi Lenapé scholar who considers the resistances among settler educators to teaching Indigenous perspectives and how they are revealed. After extensive work with teacher candidates, [Dion \(2016\)](#) put forward what she calls, “the allure of the perfect stranger” to explain teacher resistance (p. 178). Dion posits that when faced with teaching Indigenous perspectives, teachers often insist:

I can’t teach this content. I know nothing about Indigenous people, I have no friends that are Indigenous, I didn’t grow up near a reserve, I didn’t learn anything when I was in school. I am a perfect stranger to Indigenous people. (p. 178)

The perfect stranger positionality permits, as [Dion \(2016\)](#) stresses, teacher candidates to excuse themselves of responsibility and prevents them from feelings of discomfort. Furthermore, staying within the space of the perfect stranger enables teacher candidates to remain respectful, hardworking, and benevolent Canadians. Indeed, narratives of “Canada the Good” provide the foundation for how Canadian “students read and racialize the story of the territory on which they live and their relationship with First Nations peoples” ([Tupper, 2011](#), p. 41). For many teacher candidates, these narratives have been upheld and validated by the curricular and epistemological exclusion, nullification, and dismissal of Indigenous peoples throughout their years in the Canadian schooling system. Thus, the identity of the “perfect stranger” is formed and informed by the narratives we learn, and don’t learn throughout our lives.

Consequently, teacher candidates feel entitled to absolve themselves of responsibility toward teaching Indigenous perspectives and remain isolated from Indigenous people. Moreover, by assuming the role of the perfect stranger they avoid facing the contradictions of their histories as Canadians. Donald (personal communication, February 16, 2019) explains:

If somebody has a Canadian identity, and that’s how they understand themselves, and they look at an Indigenous person, then part of the acknowledgment is that the Indigenous person is from here. They have deep roots in this place ... So with that admitted, then that means that “Canadian” must be foreign to this place. And that’s a very strange contradiction, because how can a Canadian be a foreigner in Canada? This results in a “psychosis-like state” where Indigenous people can’t possibly be seen as in relation to Canadians.

The contradictions within such intergenerationally inherited and sustained beliefs about what it means to be “Canadian” challenge our status quo of comfort and privilege. [Brant \(2017\)](#) identified several ideologic barriers among the teacher candidates she interviewed. She found “a pedagogy of Whiteness which manifests itself as a challenge to mandatory Indigenous education and an active distancing in attempts to absolve oneself of responsibility in relation to Indigenous peoples” (p. 41). In turn, as Brant reminds us, this not only caused serious impediments to reconciliatory efforts within teacher education, they also risked “impeding further reconciliation in all future educational spaces” (p. 41). In [Aitken and Radford’s \(2018\)](#) research with teacher candidates at the University of Ottawa and Bishops University, proclivities to protecting whiteness and defending Settler identity were apparent in several ways. When teaching about the Indian Residential schooling system, for example, teacher candidates focused on “respect and sensitivity toward their own students, rather than the victims of colonization” (p. 45). Furthermore, teacher candidates suggested that the best ways to go about teaching “truths” and “reconciliation” was through the non-confrontational transmission of information and facts, the use of neutral language, and by “making it lighter” (p. 45). Teacher candidates spoke passionately about the importance of “avoiding approaches that would elicit emotions, particularly shame and guilt” in their students (p. 45). Moreover, their focus on knowing what to do, rather than needing to learn, is an important finding.

For [Watson and Currie-Patterson’s \(2018\)](#), their teacher education students wanted to learn about Indigenous peoples rather than learning about their own colonialism. “Students appeared to be preoccupied,” they tell us, “with ‘the cultural other’ and how to ‘deal with’ the other in the classroom” (p. 366). Similarly, in [Howell’s \(2017\)](#) research, teachers expressed comparable resistances and rhetoric in statements such as “no one taught me how to teach this” (p. 89). And, “I just don’t have the resources to teach it” (p. 101). The perfect stranger was also invoked when teachers spoke of the “fear of offending ... as this history is not my own” (p. 102). Such truth and education research demonstrates that fear, discomfort, and colonial logics are substantial impediments to teacher candidates unlearning colonialism. This will affect how, and if, they take up the relationships between Indigenous peoples and Canadians in their own classrooms.

In the next section, we examine responses to the concerns that this research has revealed. How might teacher educators work to unsettle the colonial blockages that are occurring in teacher candidates? What are the pedagogical and curricular choices we can make that might challenge resistances and provoke curiosity, relationality, and openness to unlearning?

Toward understanding colonialism as culture, unlearning and ethical relationality

The journey towards reconciliation begins with listening—not in a reactive or shallow way that results in guilt or outrage but in a deeply reflective way—where the listener continually asks the question: what action can I take to reconcile?

Poitras Pratt and Danyluk (2019, p. 8)

We have suggested that settler colonialism and its psychosis of denial, as a cultural, pedagogical, curricular practice, pose a serious challenge to the existing historical consciousness and worldviews of teacher candidates. “If teacher education systems in Canada continue to rely on existing dominant knowledge systems and governing structures to respond to the Calls to Action,” state Daniels et al. (2018) “they will fail to address them in the spirit and intent with which they were issued” (p. 202). We suggest that the spirit and intentions of the TRC *Calls to Action* (2015) could be enacted within a teacher’s professional standards of practice and ethics of care when Faculties of Education begin to address ethical relationality as part of our professional and personal curricular and pedagogical inquiries. Donald (2009) describes ethical relationality as:

... an ethic of historical consciousness. The ethic holds that the past occurs simultaneously in the present and influences how we conceptualize the future. It requires that we see ourselves as related to, and implicated in, the lives of those that have gone before us and those yet to come. It is an ethical imperative to recognize the significance of the relationships we have with others, how our histories and experiences are layered and position us in relation to each other, and how our futures are similarly tied together. (p. 7)

There are many curricular choices that teacher educators can make when attempting to address “truth,” and then “reconciliation,” across their teacher education programs (Brant-Birioukov et al., 2020; Madden, 2019). As Mitchell and Tupper’s (2017) research illustrates, when teachers experience “beyond the classroom” relational opportunities for learning, they are encouraged to “contextualize colonial events and policies beyond residential school experiences as isolated events” (p. 11). Teacher Education and its respective curriculum inquiries grounded in ethical relationality appear to increase the affective engagements that enable teacher candidates to visualize themselves in relation to people and places, “moving past the assignment of guilt to a consideration of present-day responsibility” (Mitchell and Tupper, 2017, p. 11). In turn, Mitchell and Tupper describe a noticeable increase in meaningful learning when teachers have opportunities to visit former residential school sites, meet with Elders and survivors, hear stories of Indigenous resistance and resilience, and “overtly learn about whiteness and its restructuring forces” in order to make explicit what often remains invisible (p. 20). Mitchell and Tupper noted that “critical consciousness was fostered through their (the teachers’) engagement with colonialism ... in effective, meaningful ways as opposed to traditional practices that had left them believing they had ‘learned this already’” (p. 17).

Likewise, Smith and Ng-A-Fook (2017) discovered that when teacher candidates were able to speak and listen to Elders and Residential School system survivors, they became “historical subjects during the encounters with the past lives of others ... learning how to reread, rewrite, and redress a more nuanced storied account” (pp. 81–82). During this oral history project, teacher candidates were called upon by Elders’ restorying to unlearn the different curricular ways in which settler colonialism has framed our historical consciousness and in turn the stories we teach ourselves about being “good” Canadian citizens. Yet, emotional responses to the stories of survivors and Elders alone do not facilitate change in teacher candidates (or in any of us). To become agents of change, we need to not only *feel* the weight of historical and contemporary injustices, we must also recognize ourselves as *implicated* in the relationships and in terms of the professional and personal responsibilities that such renewed relationships involve (Brant-Birioukov et al., 2020).

Rice et al. (2022) found that when engaging Indigenous and non-Indigenous participants in collaborative story-telling, it was important to “scaffold spaces for listening differently ... by creating opportunities for settlers to listen and learn from Indigenous perspectives, they became better equipped to revise their relationalities with Indigenous peoples” (p. 21). Teacher education programs cannot only be concerned with teacher candidates learning about the truths that the TRC final report revealed, but also in terms of “how they might be prepared to take up the responsibility of reconciliation in their own teaching practices” (Aitken and Radford, 2018, p. 42). Aitken and Radford’s research suggests that the opportunities for responsibility and relationality are nurtured when teacher candidates are given

... extended periods of time in a course for bearing witness to others’ testimonies of loss, studying the historical context of residential school and cultural genocide, being called on to create learning situations for the greater public, as well as thinking about future students. (p. 47)

Encouraging teachers to engage in ongoing reflective writing, whether it is the use of extensive digital blogging (Aitken and Radford, 2018), letters to survivors after hearing testimonies (Howell, 2017; Smith, 2017) or settler life writing (Tupper, 2019), holds possibilities for reckoning, revisiting, unlearning, and learning anew. As scholars and teacher educators we recognize that such pedagogical and curricular unlearning represents privileged professional learning for us, or a potential transformation of our settler colonial historical consciousness, but does not necessarily address or change the actual ongoing lived experiences for First Nation, Inuit, and Métis communities. And yet, for several teacher candidates, such calls toward unlearning within teacher education programs, might be the first time they have been invited to question their complicity with colonial systems and its intergenerational violence. Unlearning occurs through confronting settler colonialism, deconstructing and reconstructing our collective historical consciousness, and positioning ourselves in relation the both.

Professional unlearning opportunities: truth and then reconciliation with teachers

... pedagogical approaches that aim to transform Indigenous-settler relations must engage with and unsettle white settlers’ emotional attachments to dominant narratives of Canada and to their own kinship stories of origins, and call into question their assumed right to material benefits accrued from colonization.

Rice et al. (2022, p. 16)

The TRC *Calls to Action* initiated the unprecedented development of teacher resources aimed at teaching toward reconciliation. However, there is an absence of professional learning opportunities for practicing teachers (Blackstock et al., 2018). Most existing professional learning opportunities are offered through teachers' unions and focus on resource sharing and workshops. For example, the Elementary Teacher's Federation of Ontario (ETFO) offers online education resources and workshops. Such resources support their members in "learning instructional practices to integrate Aboriginal content and perspectives in a positive approach and ensure that First Nations, Métis and Inuit people's issues and worldviews are authentically reflective and connected to the content at hand" (n.d.). Likewise, the Alberta Teachers Association (ATA), the BC Teacher's Federation (BCTF) and many others offer similar workshops and resources for teachers. Many other organizations across the country provide professional learning for teachers, such as the First Nations Child and Family Caring Society (Caring Society), The Royal Canadian Geographic Society, Legacy of Hope, and Project of Heart,⁵ to name a few (Howell, 2017). However, there is scant research on the impacts of these truth and then reconciliation professional learning opportunities, and many teachers continue to feel that they have "so many resources but not enough knowledge, time or understanding" (p. 132). Below, we describe research currently being conducted that seeks to reframe how teachers learn about and teach toward truth and reconciliation in their classrooms.

Since September 2018, we have been researchers on the Social Sciences and Humanities Research Council (SSHRC) funded project called, *Just Because we're small doesn't mean we can't stand tall: Reconciliation education in the elementary school classroom*. The research seeks to understand better the impacts of reconciliation education on students' and teachers' experiences, with a goal of improving teacher professional learning. The research team is a multidisciplinary partnership between the University of Ottawa's Faculty of Education and the Caring Society. Through our partnership, we have engaged teachers in numerous professional learning opportunities, such as resource development and facilitation, and the Spirit Bear Retreat for Teacher Professional Learning which was held in August of 2021. During the 2-day virtual retreat, participants learned from several First Nations, Inuit, and Métis scholars, students, and activists who focused on building and restor(y)ing Canadian-Indigenous relations. In a research report recently published (Howell and Brittain, 2022), feedback from educators who attended the retreat reveals the transformative nature of teachers experiences:

I just wanted to take a moment to tell you how deeply moved and impressed I was by what was shared at the retreat. I learned so much. It was an open-hearted, open-minded, movement to reflecting and respecting the wisdom of Indigenous Elders while acknowledging the limitations of being mired in the context of current colonized, westernized school structures. It renewed hope. It was nothing short of transformative.

Professional learning opportunities for teachers such as the type of complex engagements that Hare (2019) maintains have the potential to reduce the proximity of space in relationships between Indigenous peoples and Canadians. Moreover, Hare stresses that rather than create professional knowledge that perpetuates superficial understandings, teacher educators must invite them to engage in complex pedagogical spaces that nurture reflexivity. The work and research we are engaged in with the Caring Society nurtures complexity by welcoming teachers into a meaningful, sustained professional learning community called the Spirit Bear Virtual School for Teachers. The Spirit Bear school provides opportunities for teachers to learn from First Nations, Inuit, and Métis peoples, unlearn colonialism, engage in holistic pedagogies, and co-create learning resources based on their lived experiences with students in their classrooms.

Restorying truths and then reconcilia(c)tion as future educational research

At different universities across Canada, we are taking small steps towards bridging our relations across such distances. And one small step calls on us to re-story the narrative compositions of historical injustices within teacher education.

Brant-Birioukov et al. (2020, p. 127)

As academic institutions that seek and share knowledge and truth as their *raison d'être*, universities are inherently implicated in the reproduction, or the unsettling, of how we ethically situate ourselves. In an ideal world, an increased level of awareness would prompt the deliberate and strategic prioritization of Indigenous perspectives in post-secondary settings.

Poitras Pratt and Gladue (2022, p. 114)

When educators honor the historical and contemporary possibilities of truth and then reconciliation education, it provides a strategic curricular and pedagogical move toward re-storying our collective pasts in a way that both strengthens Indigenous communities and politicizes the nation-state's settler colonial creation stories as ones of forgetting as much as remembering selective historical truths (see Brant-Birioukov et al., 2020; Donald, 2019; Stein, 2019). Our research team at *A Canadian Curriculum Theory Project* is current collaborating on several different federally funded research projects that are committed to extending promising

⁵Project of Heart is an arts-based inquiry project that facilitates learning about residential schools through research, connecting with survivors, and social justice actions. Thousands of Canadians across the country have taken part since its creation in 2007 by Ottawa teacher Sylvia Smith and her students.

practices toward truth, and then reconciliation education research as an emerging field of study. Importantly, all projects seek to unlearn and learn with and from Indigenous Peoples.

For example, *Mobiliser les histoires autochtones pour le changement social* | Building Connections: Mobilizing Indigenous Histories for Social Change sought to learn more about the processes of revitalization from different international Indigenous scholars, community leaders, Knowledge Keepers, and Elders who in turn have developed their own history and educational memory projects. This recorded speakers series sought to mobilize the knowledges and experiences of Indigenous communities and settler institutions to identify promising practices for the development of transformational approaches to re-storying Indigenous histories in relation to history education both inside and outside of a settler colonial schooling system.

Another project, *Co-Curricular-Making: Honoring Indigenous Connections to Land, Culture, and the Relational Self* is a national partnership grant between First Nations communities, Indigenous and non-Indigenous educational researchers who are collaborating to articulate reciprocal curricular pathways for educators and their students; to enhance understandings of Indigenous cultures and supports related to these groups; and to mobilize local, place-based, land-based First Nations ways of knowing and being. Part of this partnership project aims to study the curricular and pedagogical conditions that both challenge and sustain our capacities to participate in reconciling conversations, grapple with personal narratives, and foreground truth, and then reconcilia(c)tions inside and outside of our different teacher education programs.

Finally, Just because we're small doesn't mean we can't stand tall: Reconciliation education in the elementary classroom seeks to understand how teachers in the Ottawa-Gatineau region of Ontario/Quebec draw on the First Nations Child and Family Caring Society's (Caring Society) reconciliation-based campaigns in their elementary classrooms. We interviewed educators involved in the campaigns to examine the impacts such truth, and then reconciliation education has on elementary students and teachers. In turn, we are collaborating with the Caring Society, teachers, Elders, and Knowledge Keepers to co-create, support, and sustain a professional learning network for truth and then reconciliation education.

Certainly, there continues to be challenges and limitations toward the ways in which truth and then reconcilia(c)tion education and its research is conceived and lived across the different settler colonial educational systems. In terms of our research within the contexts of higher education, and more specifically teacher education, we continue to be implicated in the reproduction of settler colonialism even while attempting to unsettle taken-for-granted historical and/or contemporary truths. And yet, our research team remains committed as educational researchers, citizens, and fellow human beings toward working with First Nations, Inuit, and Métis peoples. Perhaps then we might move past an *Age of Apology*, and toward living, restorying and reimagining narrative compositions of past, present, and future relations as Indigenous and non-Indigenous citizens living on these lands.

References

- Aitken, A., Radford, L., 2018. Learning to teach for reconciliation in Canada: vulnerability, resistance and stumbling forward. *Teach. Teach. Educ.* (75), 40–48.
- Association of Canadian Deans of Education, 2017. ACDE Responds to the TRC. <https://csse-scee.ca/acde/wp-content/uploads/sites/7/2020/10/Media-Statement-ACDE-Responds-to-the-TRC.pdf>.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. Purich, Saskatoon.
- Blackstock, C., Bennet, S., Ng-A-Fook, N., 2018. Just Because We are Small Doesn't Mean We Can't Stand Tall: Reconciliation Education in the Elementary Classroom. Unpublished SSHRC Proposal. McGill University, Montreal, QC.
- Blackstock, C., 2021. Screaming into silence. MacLean's. <https://www.macleans.ca/opinion/residential-schools-survivors-cindy-blackstock/>.
- Brant, K., 2017. "But How Does This Help Me?": (Re)thinking (Re)conciliation in Teacher Education [Unpublished Master's Thesis]. University of Ottawa.
- Brant-Birioukov, K., Ng-A-Fook, N., Llewellyn, K.R., 2020. Restorying settler teacher education: truth, reconciliation, and oral history. In: Llewellyn, K.R., Ng-A-Fook, N. (Eds.), *Oral History, Education, and Justice: Possibilities and Limitations for Redress and Reconciliation*. Routledge, pp. 108–131. <https://doi-org.proxy.bib.uottawa.ca/10.4324/9781315179278>.
- Carleton, S., 2021. I don't need any more education': Senator Lynn Beyak, residential school denialism, and attacks on truth and reconciliation in Canada. *Settl. Colon. Stud.* 1–21. <https://doi.org/10.1080/2201473X.2021.1935574>.
- Cote-Meek, S., 2018, Nov. 6. Making a long-term commitment to Indigenous education. University Affairs. <https://www.universityaffairs.ca/opinion/from-the-admin-chair/making-a-long-term-commitment-to-indigenous-education/>.
- Coulthard, G.S., 2014. *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition*. University of Minnesota Press.
- CPAC, June 25, 2021. PM Trudeau Addresses Unmarked Graves at Residential Schools. https://www.youtube.com/watch?v=-_AhHQk_zLk.
- Czyzewski, K., 2011. The truth and reconciliation commission of Canada: insights into the goal of transformative education. *Int. Indig. Policy J.* 2 (3), 1–12. Retrieved from. <http://ir.lib.uwo.ca/iipj/vol2/iss3/4>.
- Daniels, L., Deer, S., Donald, D., Low, B., Wiseman, D., 2018. Editorial: taking up the calls to action of the TRC in teacher education. *McGill J. Educ.* 53 (2), 202–212.
- Davis, L., Hiller, C., James, C., Lloyd, K., Nasca, T., Taylor, S., 2017. Complicated pathways: settler Canadians learning to re/frame themselves and their relationships with Indigenous peoples. *Settl. Colon. Stud.* 7 (4), 398–414.
- Dion, S., 2014. inVISIBILITY: Indigenous in the city—Indigenous artists, Indigenous youth and the project of survivance. *Decolonization* 3 (1), 159–188.
- Dion, S., 2016. Mediating the space in between: voices of Indigenous youth and voices of educators in the service of reconciliation. *Can. Sociol. Assoc.* 53 (4), 468–473.
- Donald, D., 2009. Forts, curriculum, and Indigenous Métissage: imagining decolonization of Aboriginal-Canadian relations in educational contexts. *First Nations Perspect.* 2 (1), 1–24.
- Donald, D., 2019. Homo economicus and forgetful curriculum. In: Tomlins-Jahnke, H., Styres, S., Lilley, S., Zinga, D. (Eds.), *Indigenous Education: New Directions in Theory and Practice*. University of Alberta Press, pp. 103–125.
- Edwards, J.A., 2010. Apologizing for the past for a better future: collective apologies in the United States, Australia, and Canada. *South. Commun. J.* 75 (1), 57–75.
- Gaudry, Lorenz, D., 2018. Indigenousization as inclusion, reconciliation, and decolonization: navigating the different visions for indigenizing the Canadian academy. *Alternative* 14 (3), 218–227. <https://doi.org/10.1177/1177180118785382>.
- George, R.Y., 2020. A move to distract: mobilizing truth and reconciliation in settler colonial states. In: Regan, P., Craft, A. (Eds.), *Pathways of Reconciliation: Indigenous and Settler Approaches to Implementing the TRC's Calls to Action*. University of Manitoba Press, pp. 87–116.
- Gilmore, R., June 25, 2021. Trudeau sorry for "incredibly harmful" residential schools as advocates call for accountability. *Global News*. <https://globalnews.ca/news/7980719/residential-schools-trudeau-apology-cowessess-751-unmarked-graves/>.

- Goldberg, D., 2007. Raceless states. In: Enomoto, R., Johnson, G.F. (Eds.), *Race, Racialization, and Antiracism in Canada and Beyond*. University of Toronto Press, pp. 206–232.
- Hare, J., June 2, 2019. Enhancing Indigenous Perspectives Through Collaborative Inquiry in Teacher Education Coursework. Keynote at Congress 2019. UBC.
- Hare, J., 2020. Reconciliation in teacher education: hope or hype? In: Phelan, A., Pinar, W., Ng-A-Fook, N., Kane, R. (Eds.), *Reconceptualizing Teacher Education: A Canadian Contribution to a Global Challenge*. University of Ottawa Press, pp. 19–38.
- Heath Justice, D., Carleton, S., August 5, 2021. Truth before reconciliation: 8 ways to identify and confront Residential School denialism. Conversation. <https://theconversation.com/truth-before-reconciliation-8-ways-to-identify-and-confront-residential-school-denialism-164692>.
- Henderson, J., Wakeham, P., 2013. *Reconciling Canada: Critical Perspectives on the Culture of Redress*. University of Toronto Press.
- Howell, L., 2017. *Reconciliation in Action and the Community Learning Centres of Quebec: The Experiences of Teachers and Coordinators Engaged in First Nations, Inuit and Métis Social Justice Projects* [Unpublished Master's Thesis]. University of Ottawa.
- Howell, L., Brittain, M., 2022. Research Report: Just Because We're Small Doesn't Mean We Can't Stand Tall: Education for Reconciliation in the Elementary Classroom. First Nations Child and Family Caring Society. <https://incaringssociety.com/publications/research-report-just-because-were-small-doesnt-mean-we-cant-stand-tall-reconciliation>.
- Howell, L., Ng-A-Fook, N., 2022. A case of Senator Lynn Beyak and anti-Indigenous systemic racism in Canada. *Can. J. Educ.* 45 (1), 1–45.
- Howell, L., Brant-Birioukov, K., Ng-A-Fook, N., 2021. National Day for Truth and Reconciliation: universities and schools must acknowledge how colonial education has reproduced anti-Indigenous racism. Conversation. <https://theconversation.com/national-day-for-truth-and-reconciliation-universities-and-schools-must-acknowledge-how-colonial-education-has-reproduced-anti-indigenous-racism-123315>.
- Ibhawh, B., January 23, 2019. Do truth and reconciliation commissions heal divided nations? Conversation. <https://theconversation.com/do-truth-and-reconciliation-commissions-heal-divided-nations-109925>.
- Institute for the study of human rights, 2022. Political Apologies. <http://www.humanrightscolumbia.org/ahda/political-apologies>.
- Jessie, E., 2012. Promoting Reconciliation Through Exhuming and Identifying Victims in the 1994 Rwandan Genocide. *Africa Initiative and The Centre for International Governance Innovation*. https://books.scholarsportal.info/uri/ebooks/ebooks0/gibson_cpcc/2012-11-09/1/10585931.
- Jewell, E., Mosby, I., 2021. Calls to Action Accountability: A 2021 Status Update on Reconciliation. Yellowhead Institute. <https://yellowheadinstitute.org/trc/>.
- Joseph, R.A., 2008. Jade door: reconciliatory justice as a way forward citing New Zealand experience. In: Brant-Castellano, M., Archibald, L., Degagné, M. (Eds.), *From Truth to Reconciliation: Transforming the Legacy of Residential Schools*. Aboriginal Healing Foundation Research Series, pp. 207–229.
- Joseph, B., 2018. 21 Things You May Not Know About the Indian Act. Page Two Books, Inc.
- Jurgens, E., 2020. Teaching truth before reconciliation. In: Regan, P., Craft, A. (Eds.), *Pathways of Reconciliation: Indigenous and Settler Approaches to Implementing the TRC's Calls to Action*. University of Manitoba Press, pp. 119–141.
- Kairos, 2018. Winds of Change: 2018 Report Card. <https://www.kairoscanada.org/what-we-do/indigenous-rights/windsofchange-report-cards>.
- Lechnitz, J., 2021. Canada's colonial reckoning. *Int. Polit. Soc.* <https://www.ips-journal.eu/topics/democracy-and-society/canadas-colonial-reckoning-5253/>.
- Llewellyn, K., Ng-A-Fook, N., 2020. Oral history and education: hopes for addressing redress and reconciliation. In: Llewellyn, K., Ng-A-Fook, N. (Eds.), *Oral History, Education, and Justice: Possibilities and Limitations for Redress and Reconciliation*. Routledge, pp. 1–10. <https://doi.org/10.4324/9781315179278>.
- MacDonald, J., Markides, J., 2018. Called to action: dialogue around praxis for reconciliation. *McGill J. Educ.* 53 (2), 213–232.
- MacDonald, D.B., 2020. Paved with comfortable intentions: moving beyond liberal multiculturalism and civil rights frames on the road to transformative reconciliation. In: Regan, P., Craft, A. (Eds.), *Pathways of Reconciliation: Indigenous and Settler Approaches to Implementing the TRC's Calls to Action*. University of Manitoba Press, pp. 3–34.
- Madden, B., 2019. A de/colonizing theory of truth and reconciliation education. *Curric. Inq.* 49 (3), 284–312.
- McCready, A.L., 2009. Redressing Redress: the Neoliberal Appropriation of Redress in the anti-Native Backlash at Caledonia. *Engl. Stud. Can.* 35 (1), 161–190. <https://doi.org/10.1353/esc.0.0170>.
- Mitchell, T., Tupper, J., 2017. *Highschool Teachers Working Toward Reconciliation: Examining the Teaching and Learning of Residential Schools*. McDowell Foundation. http://mcdowellfoundation.ca/isf/uploads/2016/07/Project_270_Working_Towards_Reconciliation.pdf.
- Ng-A-Fook, N., (in press). Reconstructing curriculum studies in Canada: life writing, settler colonialism, truth and then reconciliation. In: Pinar, W.F., Phelan, A. (Eds.), *Curriculum Studies in Canada: Present Preoccupations*. University of Toronto Press.
- Nobles, M., 2008. *The Politics of Official Apologies*. Cambridge University Press.
- Organick, A.G., 2020. Non-apology in the age of apology. *Denning Law J.* 31 (1), 149–164. <https://doi.org/10.5750/dlj.v31i1.1793>.
- Peters, M.A., Mika, C.T., 2017. Aborigine, Indian, indigenous or first nations? *Educ. Philos. Theor.* 49 (13), 1229–1234.
- Poitras Pratt, Y., Danyluk, P.J., 2019. Exploring reconciliatory pedagogy and its possibilities through educator-led praxis. *Can. J. Scholarsh. Teach. Learn.* 10 (3), 1–16. <https://doi.org/10.5206/cjsotl-rcacea.2019.3.9479>.
- Poitras Pratt, Y., Gladue, K., 2022. Re-defining academic integrity: embracing indigenous truths. In: Eaton, S.E., Hughes, J.C. (Eds.), *Academic Integrity in Canada: An Enduring and Essential Challenge*. Springer, pp. 103–123. https://doi.org/10.1007/978-3-030-83255-1_5.
- Refugeereseach, June 15, 2017. Part 2: Can Reconciliation Take (Up) Place? [Video]YouTube. <https://www.youtube.com/watch?v=Eu-opkwFw0s>.
- Regan, P., 2010. *Unsettling the Settler Within: Indian Residential Schools, Truth Telling, and Reconciliation in Canada*. UBC Press.
- Regan, P., Craft, A., 2020. *Pathways of reconciliation: Indigenous and settler approaches to implementing the TRC's calls to action*. University of Manitoba Press.
- Rice, C., Dion, S.D., Fowlie, H., Breen, A., 2022. Identifying and working through settler ignorance. *Crit. Stud. Educ.* 63 (1), 15–30. <https://doi.org/10.1080/17508487.2020.1830818>.
- Sanz, E., 2012. *National Apologies: Mapping the Complexity of Validity*. The Centre for Peace and Conflict Studies. <http://www.centrepeaceconflictstudies.org/publications/browse/national-apologies-mapping-the-complexities-of-validity>.
- Simon, R.I., 2013. Towards a hopeful practice of worrying: the problematics of listening and educative responsibilities of Canada's truth and reconciliation commission. In: Henderson, J., Wakeham, P. (Eds.), *Reconciling Canada: Critical Perspectives on the Culture of Redress*. University of Toronto Press, pp. 129–142.
- Smith, S., 2017. Teachers' Perceptions of Project of Heart, an Indian Residential School Education Project [Unpublished Master's Thesis]. University of Regina.
- Smith, B., Ng-A-Fook, N., 2017. Doing oral history education toward reconciliation. In: Llewellyn, K., Ng-A-Fook, N. (Eds.), *Oral History and Education*. Palgrave Macmillan, pp. 65–86. https://doi.org/10.1057/978-1-349-95019-5_4.
- Stein, S., 2019. "Truth before reconciliation": the difficulties of transforming higher education in settler colonial contexts. *High. Educ. Res. Dev.* 39 (1), 156–170. <https://doi.org/10.1080/07294360.2019.1666255>.
- Tasker, J.P., January 25, 2021. Lynn Beyak, the senator who defended residential schools, is resigning. CBC News. <https://www.cbc.ca/news/politics/beyak-retirement-1.5886435>.
- Thomas, A.R., February 15, 2019. Who is a Settler, According to Black and Indigenous Scholars? Vice. <https://www.vice.com/en/article/gyaj4/who-is-a-settler-according-to-indigenous-and-black-scholars>.
- Truth and Reconciliation Commission of Canada, 2015. *Final Report of the Truth and Reconciliation Commission of Canada. Volume 1: Summary. Honouring the Truth, Reconciling the Future*. James Lorimer & Co. Ltd.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Tunney, C., October 6, 2021. Trudeau says he regrets travelling to Tofino, B.C., on 1st national day for truth and reconciliation. CBC News. <https://www.cbc.ca/news/politics/trudeau-tofino-trip-truth-reconciliation-1.6201475>.
- Tupper, J., 2011. Disrupting ignorance and settler identities: the challenges of preparing beginning teachers for treaty education. *Education* 17 (3), 38–55.
- Tupper, J., 2012. Treaty education for ethically engaged citizenship: settler identities, historical consciousness and the need for reconciliation. *Citizen. Teach. Learn.* 7 (2), 143–156.

- Tupper, J., 2019. Cracks in the foundation: (re)storying settler colonialism. In: Llewellyn, K., Ng-A-Fook, N. (Eds.), *Oral History, Education, and Justice: Possibilities and Limitations for Redress and Reconciliation*. Routledge, pp. 88–104.
- UNESCO, 2021. *Reimagining Our Futures Together: A New Social Contract for Education*. United Nations Educational, Scientific and Cultural Organization.
- Universities Canada, June 29, 2015. *Universities Canada Principles on Indigenous Education*. News Brief. <https://www.univcan.ca/media-room/media-releases/universities-canada-principles-on-indigenous-education/>.
- Vowel, C., 2016. *Indigenous Writes : A Guide to First Nations, Métis, and Inuit Issues in Canada*. HighWater Press.
- Wakeham, P., 2019. Outsourcing reconciliation: the Government of Canada's #IndigenousReads campaign and the appropriation of Indigenous intellectual labor. *Stud. Am. Indian Lit.* 31 (1–2), 1–30.
- Wallace-Casey, C., 2022. Teaching and learning the legacy of residential schools for remembering and reconciliation in Canada. *Hist. Educ. Res. J.* 19 (1), 4. <https://doi.org/10.14324/HERJ.19.1.04>.
- Watson, K., Currie-Patterson, N., 2018. Responding to the calls to action: reflections on teaching mandatory indigenous education to teacher candidates in Ontario. *McGill J. Educ.* 53 (2), 362–371. <https://doi.org/10.7202/1058403ar>.

Teacher education for social change: shifting from race-evasiveness to racial literacy

Corinna D. Ott and Rita Kohli, University of California, Riverside, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Racism and race-evasiveness in teacher education	283
Race-evasive curriculum in teacher education	284
Race-evasive approaches to white student resistance in teacher education	284
Racial literacy in teacher education	285
Teacher education programs must actively support diversification	285
Teacher education programs must actively shift their ideological commitments	286
Discussion	286
References	287

Teacher education is a powerful socializing institution that shapes who K-12 educators are and how they approach teaching. In the United States, teacher education programs were established 200 years ago, in the 1820s, with the growth of normal schools. They were modeled after European teacher training institutions, guiding educators on how to engage students, utilize limited resources, and create a positive learning environment within the newly emerging systems of education (Ducharme and Ducharme, 2022). For many working class and students of Color, these education systems were designed as a form of control, socializing students to their place within the hierarchy of American society; and teachers were instrumental to this project (Gratz, 2009). In Nebraska, the Teachers' Committee was cited as stating,

How can we have a national spirit in a Commonwealth where there is an infusion of the language and blood of many nations, unless there is a very strong effort made to socialize the different elements and weld them into a unified whole? It therefore becomes evident how important it is that the teacher be an American in sympathy, ideals, training and loyalty.

Gratz (2009, p. 55).

They were calling for teachers who embodied nationalism and assimilationist goals. By the mid-20th century, normal school teacher training programs had expanded into 4 year colleges and universities across the nation (Ducharme and Ducharme, 2022).

In 2022, where over half of public school students in the nation are students of Color, teacher education programs carry many of the same principles and structures of their origins. Today, 70% of teacher candidates, 87% of adjunct instructors, and 91% of tenured/tenure track instructors are white (King and Hampel, 2018) and predominantly monolingual, and teacher education programs continue to be critiqued for maintaining and reproducing whiteness. Scholars such as Sleeter (2001, 2016), Cross (2005), and Souto-Manning (2019) have argued of an ever-present whiteness in teacher education, where programs are complicit in "White ways and systems of knowing which ... further White interests through the invisibility and/or normalizing of systemic racism" (p. 100). And while this manifests through student resistance or teacher educator race-evasiveness, it is important to understand the complexity of racism in teacher education as a structural issue (Kohli and Pizarro, 2022).

Kohli et al. (2021) argue that there must be a systematized effort to address racism and whiteness within the multiple facets of teacher education, from admissions processes and the recruitment of teacher candidates, to the curriculum and pedagogy of coursework, field placements, and the retention of teacher candidates of Color. Building from higher education research on racial climate, they offer a five dimensional model for a healthy racial climate in teacher education programs, that includes a shift from a race-evasive approach—where programs ignore and invisibilize the prevalence of racial inequity and racism—to one of racial literacy—where program leaders, teacher educators, and teacher candidates are able to effectively identify racism and are committed to disrupting it. In this entry, we build upon this model and posit that for teacher education programs to become pivotal spaces of change and transformation they must operate in ways that embody racial literacy.

Racism and race-evasiveness in teacher education

Teacher education has historically been complicit in upholding and maintaining forms of racism through the policies and practices it enacts. Today, rather than overt discrimination, teacher education engages in subtle and more sophisticated practices that are just as effective in maintaining the racial status quo of white superiority and dominance (Bonilla-Silva, 2018). One of the primary tools

to upholding the racial status quo is what Bonilla-Silva labeled color-blindness,¹ and what we refer to as race-evasiveness—claims of “not seeing” race as a central issue that directly impacts the lives of people of Color. Instead, this race-evasiveness shifts the blame onto historically marginalized people of Color for their lack of “success” or achievement, while rendering the oppressive role of systemic and institutional forces as invisible. In this section, we point to how racism and persistent whiteness is maintained in teacher education programs through race-evasive approaches to curriculum and white student resistance, resulting in the (re) production of a predominantly white teaching force and dominant ideologies of whiteness.

Race-evasive curriculum in teacher education

Although many teacher education programs espouse the value of developing teachers who are committed to enacting social change in their classrooms, there appears to be a clear disconnect between this rhetoric and the actual practices of these programs. The use of terms such as “diversity,” “multiculturalism,” and “inclusion” often prevail within teacher education, obscuring the pervasive forces of whiteness, racism, and power that shape policies and practices across the educational system (Cross, 2005). Thus, while teacher education programs may emphasize developing teachers who are critically attuned to issues of educational in/equity in their mission, many of these programs often privilege white, middle-class norms and values (Marom, 2019) while simultaneously erasing the histories, perspectives, and contributions of teacher candidates of Color (Brown, 2014; Gomez et al., 2008; Nguyen, 2008).

For example, most teacher education programs are required by credentialing entities to attend to racial, cultural, and/or linguistic diversity. However, they often present these topics in a shallow manner, frequently held in one “multiculturalism” course, and absent an analysis of power (Amos, 2016; Sleeter, 2001). Sleeter (2016) asserts that separating “diversity” into one course within a teacher education program “limits how teacher preparation programs are holistically designed to prepare teachers for the diverse students in schools.” (p. 5). Additionally, Gorski (2009) found that while these “diversity” courses were designed to prepare teacher candidates’ cultural sensitivity, tolerance, and multicultural competence, in an analysis of multicultural course syllabi, the majority of them “failed to frame multicultural education as a political movement concerned with social justice, as an approach for comprehensive reform, as a critical analysis of power and privilege, or as a process for eliminating educational inequities” (p. 316).

This race-evasive approach to the content of teacher education results in teacher candidates with little awareness surrounding issues of discrimination and racism, how these issues manifest and operate with schools and society at large, and how these forces are reproduced in day-to-day life (Sleeter, 2008). When unchallenged, beliefs in meritocracy and individualism also uphold a myth of equality that can result in deficit thinking of students of Color, blaming them for their “failures” in schools (Cherng and Halpin, 2016; Sleeter, 2008; Tenenbaum and Ruck, 2007; Valencia, 2010). Thus, as programs engage in race-evasive practices, clustering content about race, culture, and language into one course without limited structural analyses, they stunt teacher candidates’ needed capacity to read racism, understand its impact, and disrupt its harm (Sleeter, 2001; Picower, 2009).

Race-evasive approaches to white student resistance in teacher education

While teacher education programs have attempted to recruit more teachers of Color into their programs, they ultimately fail to adequately address and combat the pervasive culture of whiteness embedded within their programs, thus disproportionately centering the needs and comfort of white students (Brown, 2014). This results in teacher candidates of Color being rendered to the margins and frequently subjected to racism throughout their programs and student teaching placements from peers, instructors, and school placement staff (Kohli, 2021; Rodriguez-Mojica et al., 2020). For instance, the Latinx students in Irizarry’s (2011) study were silenced by culturally insensitive white professors in their teacher education courses, and expressed feeling marginalized, rather than valued, within the teacher education community. Teacher candidates of Color who do feel comfortable enough to share their opinions in their classes have felt that their white peers do not take them seriously, were rendered irrelevant (Amos, 2010), and were fearful of further ostracization from their white peers (Amos, 2010, 2016). Similarly, Chinese-American students in Hernández Sheets and Chew’s study (2002) resorted to silence in order to cope with the discomfort they felt as a result of the disparaging comments made by White students in their classes. This frequent silencing and fear of retaliation from White students results in teacher candidates of Color forcibly assimilating to the whiteness of their program in order to complete it (Amos, 2016). Other teacher candidates seriously consider leaving their programs, or are pushed out altogether (Kohli, 2018, 2021).

Within majority white teacher educators and administrative leadership, when programs try to present a more critical focus, it is typically teacher educators of Color hired to supplement the race-evasive curriculum (Kohli and Pizarro, 2022; Ladson-Billings, 2005). As teacher educators of Color teach about race and racism, white teacher candidates enact numerous strategies of resistance (Amos, 2016; Chang-Bacon, 2021). Research has shown that they discredit the impact racism has had on people of Color in the United States (Garrett and Segall, 2013), engage in emotional outbursts (Amos, 2010; DiAngelo, 2006; Hernández Sheets and Chew, 2002), and express anger in course evaluations (Truong et al., 2014; Graves, 2017; Han and Leonard, 2017). Scholars have argued that white teacher candidates use the anonymity of these evaluations to reaffirm their dominance and power in these programs (The Teacher of Color Collective and Souto-Manning, 2021) by questioning the competency, pedagogy, and emotional

¹We build from and honor the pivotal work of Bonilla-Silva (2018) to describe an avoidance discourse about race and racism as a mechanism that maintains racism. Critical disability scholars have argued that the use of the term “colorblind” is ableist, calling for more precision in our language (see Annamma et al., 2017). Aligned with their use of color-evasive, we use the term race-evasive (Harrison, 1998) as a more precise way to name this phenomenon.

stability of teacher educators of Color generally, and Black women teacher educators in particular (Evans-Winters and Twyman Hoff, 2011; Kohli and Pizarro, 2022; Sleeter, 2016).

While many white teacher educators and teacher candidates are enacting harm that impacts teacher educators of Color, teacher candidates of Color, and eventually K-12 students of Color, Kohli and Pizarro (2022) remind us that teacher education programs are making choices through their embedded policies and practices. They are actively admitting cohorts that lack diversity, have limited understandings of inequity, and often carry deficit frames of communities of Color. Thus, it is important to highlight how programs remain culpable in neglecting and perpetuating racial harm within their programs and in K-12 schools.

Racial literacy in teacher education

Given the clear prevalence of racism, whiteness, and the racial harm that is fostered through teacher education, Kohli et al. (2021) argue that to move toward a healthy racial climate in teacher education where the well-being of communities of Color is prioritized, programs must operate with racial literacy. Building from the work of Guinier (2004) and Twine (2004), we contend that racial literacy equips people with a set of discursive practices to name, critically analyze, and act upon the ways in which racism informs the structural, political, and economic circumstances of the US. Racial literacy involves the possession of a “racial grammar and vocabulary that facilitates a conversation about race, racism, and antiracism” (Twine and Steinbugler, 2006, p. 344) and is a practice that is continually taught and learned throughout the course of one’s life.

If teacher education programs want to produce a teaching force responsive to racial inequities, they must develop educators who are critically oriented, culturally sustaining (Paris and Alim, 2017), and are prepared to critique and take action against the racism within schools. The more experiences teachers have developing racial literacy, the more likely they are to create conditions in their classrooms and curriculum that are supportive of students of Color’s realities (Rogers and Mosley, 2008). However, for teacher education to be a lever of transformation and social change, Kohli et al. (2021) maintain that leaders must work toward not only developing the racial literacies of teacher candidates, but must also ensure that racial literacy is used to guide the various facets of the program. We argue that programs must actively support diversification and they must actively shift the ideologies taught and embodied in teacher education programs.

Teacher education programs must actively support diversification

US public schools were established and continue to exist as sites of racial inequity and racial harm, where students of Color have had their language, culture, and knowledge systems demeaned, ignored, or erased. Within this context, there have always existed teachers of Color who have resisted the assimilationist goals of schools and who have served as an important resource for students of Color and their families. It has been noted that teachers of Color are more likely to have high classroom expectations (Villegas and Irvine, 2010), create cultural bridges with families (Gomez et al., 2008), teach in culturally responsive and sustaining ways (Kohli, 2021), and recognize and resist racism (Villegas and Irvine, 2010)—all conditions that are supportive of the success and well-being of students of Color. For example, a Black student with at least one Black teacher is shown to have higher growth in their reading abilities and is less likely to be suspended or referred out of class (Lindsay and Hart, 2017).

Unfortunately, teachers of Color are underrepresented nationally, making up just over 20% of K-12 educators within public schools serving over 50% students of Color. Considering their significant role in the success of students of Color, yet their low representation in the field, the need to diversify the teaching force has come into focus for many teacher education programs. Undergraduate education programs should consider partnering with high schools and campus summer programs to create diverse pipelines; graduate programs should recruit from student organizations and majors that draw students of Color; and programs at both levels should provide resources that mitigate known barriers to their matriculations, such as supporting aspiring teachers through the prerequisites requirements and the high cost of tuition (Kohli, 2021).

Another proven method to building culturally responsive teachers into communities of Color is “grow your own” programs, which seek to build a diverse “homegrown” teaching force who are already committed to serving their communities (Haddix, 2017; Valenzuela, 2017). By providing a variety of financial, academic, and social supports to prospective teacher candidates who may otherwise be barred from joining the teaching profession, grow your own programs have been particularly successful at recruiting and retaining linguistically diverse, first-generation, historically marginalized teacher candidates of Color from local communities and schools (Gist et al., 2019). Grow your own programs emphasize equitable approaches and critical perspectives that combine the cultural wealth and knowledge of these prospective teachers, culturally responsive and sustaining practices, and social justice pedagogies (Valenzuela, 2017) to build the racial literacy of these teacher candidates. The combination of these practices can reduce high rates of teacher turnover in these areas, as developing teachers of Color are already familiar with the communities that they serve and can offer direct and tailored instruction to students that teachers outside the community may not be able to. As a result, students of Color in these schools have access to stable, experienced teachers who are committed to engaging them in sustaining ways, share similar backgrounds and experiences, and are deeply committed to their learning and well-being.

However, in a field where the majority of teacher educators and teacher candidates are white, and curriculum rarely interrogates oppression because it is structured to privilege white comfort, recruiting teacher candidates of Color into programs is not enough. As some programs have focused on diversifying their students, they have made very few other shifts, consequently enlisting already

minoritized educators into hostile climates not designed for their success. Thus, it is important for teacher education program leaders to establish and maintain learning environments that nurture the well-being of teacher candidates of Color.

Rather than viewing teacher candidates of Color as an additive change on top of existing structures that include predominantly white teacher candidates, programs need to re-think the overall composition of programs. The racial dispositions of peers have a great impact on the well-being of teachers of Color (and the well-being of K-12 students of Color); thus, to avoid harmful interactions, programs should consider “asset framing of communities of Color” a core competency for admissions. The structures of the program must also consider the epistemologies and ontologies of communities of Color, fostering a collaborative environment where teacher candidates are cultivated to collectively plan and see each other as resources in the profession.

To realize these goals, programs must ensure that teacher education faculty and teacher mentors are racially, culturally, and linguistically diverse; have advanced racial literacies; and are ready to actively resist harm. Programs must also be prepared to respond to racial harm as it predictably arises. In addition to having staff who competently teach about race and racism, programs must also have established policies, protocols, and practices to directly address racial injustice within courses, in school placements, and between students and/or staff.

Teacher education programs must actively shift their ideological commitments

Traditional forms of curriculum in teacher education programs often center technical forms of teaching rather than fostering environments for students to engage in critical conversations surrounding race, power, and other forms of injustice in schools. Yet, if teacher education programs are truly committed to serving as levers of transformation, in addition to actively supporting diversification, they must also shift their ideological commitments. This can occur through the implementation alternative forms of critical curriculum and pedagogy, offering students access to critical theory, readings from theorists who are Black, Indigenous, and people of Color, and experiences that explicitly strengthen their skills to navigate the racial climates of K-12 schools. Rather than reinforcing race-evasive perspectives rooted in whiteness, these critical pedagogical approaches are anchored in visions of justice, equity, and the belief that education is inherently political (Souto-Manning, 2010), and thus require that teacher education programs build upon the strengths, backgrounds, skills, and knowledge of communities of Color. Two such examples are culturally sustaining pedagogies and community-based teacher education.

Culturally sustaining pedagogies (CSP) seek to “preserve and maintain linguistic, literate, and cultural pluralism as part of the democratic project of schooling” (Paris, 2012, p. 95). In doing so, CSP has the power to transform the practice of teaching so that teachers entering classrooms treat their students of Color with dignity and actively work to sustain their strengths, humanities, and communities. Adopting CSP in teacher education requires programs to actively assess if their teacher candidates are teaching in ways that are culturally sustaining toward students and communities of Color, or if they are continuing to produce teacher candidates who view these communities from deficit perspectives (Ladson-Billings, 2017). The ultimate goal of implementing CSP in teacher education should be to develop teachers who are racially literate and who are actively empowered and well-equipped to cultivate racial literacy among their future students; this should result in moving their students to take action against the injustices present in their daily lives. Further, by centering elements of CSP within their programs, teacher education programs can also sustain the assets and humanities of teacher candidates of Color and cultivate their racial literacies, rather than continuing to render them to the margins of their programs.

An additional approach that has the potential to shift the ideological commitments of programs is community-based teacher education, which seeks to prepare future teachers to work in historically marginalized communities. Community-based teacher education builds partnerships between universities, communities, and schools to develop teachers who are equity-focused, critically conscious, and socially just (Zygmunt and Clark, 2021). Teacher candidates work in solidarity with members of the community to understand issues of inequity that impact the schooling contexts they are a part of. As they immerse themselves in the community, these prospective teachers learn to “understand systems of power and privilege and are compelled to work collaboratively toward a socially just society in order to rectify conditions that disadvantage populations of children” (Zygmunt and Clark, 2021, p. 5). With the proper support, teacher candidates and community members can work together to actively take action against the injustices and inequitable conditions that disadvantage students in schools.

Discussion

Teacher education programs continue to implicitly replicate whiteness since their inception more than 200 years ago, much to the detriment of communities of Color. While forms of racism are less explicit now than before, racism is still deeply embedded within the policies and practices of teacher education, ranging from admissions processes to the centering of whiteness in curriculum and subsequent replication of ideologies rooted in white supremacy. These practices that maintain racial inequities in both teacher education programs and K-12 schools are maintained through pervasive forms of race-evasiveness, where teacher educators and program leaders overlook their roles in the reproduction and development of predominantly white, monolingual teachers who lack necessary skills to interrogate systemic racism.

To maintain a healthy racial climate where teachers of Color, teacher educators of Color, and K-12 students of Color are not continuously harmed by race-evasive practices, teacher education programs must make significant structural changes that intentionally cultivate racial literacy among faculty, staff, and teacher candidates. These structural changes must ensure that all teacher

education faculty share the responsibility of reflecting on and challenging racism, and support students of Color as they navigate these programs. This begins by requiring faculty and program leadership to understand the historical and contemporary experiences and strengths of communities, teacher candidates and faculty of Color. Racial literacy should also be a core competency for new faculty hires, evaluated through rubrics that gauge understandings of structural racism and a commitment to challenging it. In addition, racial literacy growth should be part of the evaluative process for all staff, faculty, and administrators in teacher education (Kohli et al., 2021).

Teacher education programs that seek to prioritize the racial literacy development of their teacher candidates must critically assess how they prepare their candidates to teach in liberatory ways for racial and social justice, rather than continue to view communities of Color as deficient and in need of “fixing.” Cultivating racial literacy in teacher candidates requires them to be able to identify, respond to, and counter forms of racism that they might encounter in a variety of contexts (Price-Dennis and Sealey-Ruiz, 2021). Teacher education programs must provide candidates with the space and support to critically discuss and reflect on their experiences with race, racism, and power, and interrogate how these forces shape the schooling contexts they are a part of. As teacher education programs assist teacher candidates in the process of cultivating their racial literacy, they become equipped with the “skills to use the power they have and the agency they will need to examine and discuss race and racism, to adopt anti-racist stances ... and to disrupt racialized educational inequities” (Rolón-Dow et al., 2021, p. 666). Shifting away from race-evasiveness toward racial literacy in teacher education requires a serious commitment from the various program stakeholders to actively identify and dismantle the long-standing practices and policies that prioritize whiteness.

References

- Amos, Y.T., 2010. “They don’t want to get it!” Interaction between minority and white pre-service teachers in a multicultural education class. *Multicult. Educ.* 17 (4), 31–37.
- Amos, Y.T., 2016. Voices of teacher candidates of color on white race evasion: “I worried about my safety!”. *Int. J. Qual. Stud. Educ.* 29 (8), 1002–1015.
- Annamma, S., Jackson, D.D., Morrison, D., 2017. Conceptualizing color-evasiveness: using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethn. Educ.* 20 (2), 147–162.
- Bonilla-Silva, E., 2018. *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*, fifth ed. Rowman & Littlefield Publishers, Maryland.
- Brown, K.D., 2014. Teaching in color: a critical race theory in education analysis of the literature on preservice teachers of color and teacher education in the US. *Race Ethn. Educ.* 17 (3), 326–345.
- Chang-Bacon, C.K., 2021. “We sort of dance around the race thing”: race-evasiveness in teacher education. *J. Teach. Educ.* 1–15.
- Cherng, H.S., Halpin, P.F., 2016. The importance of minority teachers: student perceptions of minority versus white teachers. *Educ. Res.* 45 (7), 407–420.
- Cross, B.E., 2005. New racism, reformed teacher education, and the same ole’ oppression. *Educ. Stud.* 38 (3), 263–274.
- DiAngelo, R.J., 2006. White fragility: I’m leaving. In: Armstrong, D., McMahon, B. (Eds.), *Inclusion in Urban Educational Environments*. Information Age Publishing, pp. 213–240.
- Ducharme, E., Ducharme, M., Teacher Education Historical Overview. Accessed February 1, 2022, <https://education.stateuniversity.com/pages/2479/Teacher-Education.html#:~:text=The%20first%20formal%20teacher%20preparation,at%20least%20one%20of%20them>.
- Evans-Winters, V.E., Twyman Hoff, P., 2011. The aesthetics of white racism in pre service teacher education: a critical race theory perspective. *Race Ethn. Educ.* 14 (4), 461–479.
- Garrett, H.J., Segall, A., 2013. (Re)consideration of ignorance and resistance in teacher education. *J. Teach. Educ.* 64 (4), 294–304.
- Gist, C.D., Bianco, M., Lynn, M., 2019. Examining grow your own programs across the teacher development continuum: mining research on teachers of color and nontraditional educator pipelines. *J. Teach. Educ.* 70 (1), 13–25.
- Graves, D., 2017. “Written all over my face”: a Black man’s toll of teaching white students about racism. In: Picower, B., Kohli, R. (Eds.), *Confronting Racism in Teacher Education: Counternarratives of Critical Practice*. Routledge, New York, pp. 60–66.
- Gratz, D.B., 2009. The Peril and Promise of Performance Pay: Making Education Compensation Work. *R&L Education*.
- Gomez, M.L., Rodriguez, T.L., Agosto, V., 2008. Who are Latino prospective teachers and what do they bring to U.S. schools? *Race Ethn. Educ.* 11 (3), 267–283.
- Gorski, P., 2009. What we’re teaching teachers: an analysis of multicultural teacher education courses. *Teach. Teach. Educ.* 25, 309–318.
- Guinier, L., 2004. From racial liberalism to racial literacy: Brown v. Board of Education and the interest-divergence dilemma. *J. Am. Hist.* 91 (1), 92–118.
- Han, K.T., Leonard, J., 2017. Why diversity matters in rural America: faculty of color challenging whiteness. *Urban Rev.* 49, 112–149.
- Haddix, M.M., 2017. Diversifying teaching and teacher education: beyond rhetoric and toward real change. *J. Lit. Res.* 49 (1), 141–149.
- Harrison, F.V., 1998. Introduction: Expanding the discourse on “race”. *Am. Anthropol.* 100 (3), 609–631.
- Hernández Sheets, R., Chew, L., 2002. Absent from the research, present in our classrooms: preparing culturally responsive Chinese American teachers. *J. Teach. Educ.* 53 (2), 127–141.
- Irizary, J.G., 2011. En la lucha: the struggles and triumphs of Latino/a preservice teachers. *Teach. Coll. Rec.* 113 (12), 2804–2835.
- King, J.E., Hampel, R., 2018. *Colleges of Education: A National Portrait*. American Association of Colleges for Teacher Education, Washington, DC.
- Kohli, R., 2018. Behind school doors: the impact of hostile racial climates on urban teachers of color. *Urban Educ.* 53 (3), 307–333.
- Kohli, R., 2021. *Teachers of Color: Resisting Racism and Reclaiming Education*. Harvard Education Press, Massachusetts.
- Kohli, R., Pizarro, M., 2022. The layered toll of racism in teacher education on teacher educators of color, first ed. AERA Open 8, 1–12.
- Kohli, R., Dover, A.G., Jayakumar, U.M., Lee, D., Henning, N., Comeaux, E., Nevárez, A., Hipolito, E., Carreno Cortez, A., Vizcarra, M., 2021. Toward a healthy racial climate in teacher education: centering the well-being of teacher candidates of color. *J. Teach. Educ.* 73 (1), 52–65.
- Ladson-Billings, G., 2005. *Beyond the Big House: African American Educators on Teacher Education*. Teachers College Press, New York.
- Ladson-Billings, G., 2017. The (r)evolution will not be standardized: teacher education, hip hop pedagogy, and culturally relevant pedagogy 2.0. In: Paris, D., Alim, H.S. (Eds.), *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York, pp. 141–156.
- Lindsay, C.A., Hart, C.M.D., 2017. Exposure to same-race teachers and student disciplinary outcomes for Black students in North Carolina. *Educ. Eval. Pol. Anal.* 39 (3), 485–510.
- Marom, L., 2019. Whiteness and teacher education. In: Peters, M.A. (Ed.), *Encyclopedia of Teacher Education*. Springer Singapore, pp. 1–5.
- Nguyen, H.T., 2008. Conceptions of teaching by five Vietnamese American preservice teachers. *J. Lang. Ident. Educ.* 7 (2), 113–136.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Paris, D., Alim, H.S., 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York.
- Picower, B., 2009. The unexamined whiteness of teaching: how white teachers maintain and enact dominant racial ideologies. *Race Ethn. Educ.* 12 (2), 197–215.
- Price-Dennis, D., Sealey-Ruiz, Y., 2021. *Advancing Racial Literacies in Teacher Education: Activism for Equity in Digital Spaces*. Teachers College Press, New York.
- Rodriguez-Mojica, C., Rodela, K.C., Ott, C., 2020. “I didn’t wanna believe it was a race issue”: student teaching experiences of preservice teachers of color. *Urban Rev.* 52 (3), 435–457.

- Rogers, R., Mosley, M., 2008. A critical discourse analysis of racial literacy in teacher education. *Ling. Educ.* 19 (2), 107–131.
- Rolón-Dow, R., Flynn, J.E., Mead, H., 2021. Racial literacy theory into practice: teacher candidates' responses. *Int. J. Qual. Stud. Educ.* 34 (7), 663–679.
- Sleeter, C.E., 2001. Preparing teachers for culturally diverse schools: research and the overwhelming presence of whiteness. *J. Teach. Educ.* 52 (2), 94–106.
- Sleeter, C., 2008. Preparing White teachers for diverse students. In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D.J., Demers, K.E. (Eds.), *Handbook of Research on Teacher Education*. Routledge, New York, pp. 559–582.
- Sleeter, C.E., 2016. Critical race theory and the whiteness of teacher education. *Urban Educ.* 1–15.
- Souto-Manning, M., 2010. *Freire, Teaching, and Learning: Culture Circles Across Contexts*. Peter Lang Publishing Inc, New York.
- Souto-Manning, M., 2019. Toward praxically-just transformations: interrupting racism in teacher education. *J. Educ. Teach.* 45 (1), 97–113.
- Tenenbaum, H.R., Ruck, M.D., 2007. Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *J. Educ. Psychol.* 99 (2), 253–273.
- The Teacher of Color Collective, Souto-Manning, M., 2021. On the mis-education of teachers of color: a letter to teacher educators. *J. Teach. Educ.* 73 (1), 66–80.
- Truong, K.A., Graves, D., Keene, A.J., 2014. Faculty of color teaching critical race theory at a PWI: an autoethnography. *J. Crit. Thought Prax.* 3 (2).
- Twine, F.W., 2004. A white side of black Britain: the concept of racial literacy. *Ethn. Racial Stud.* 27 (6), 878–907.
- Twine, F.W., Steinbugler, A.C., 2006. The gap between whites and whiteness: interracial intimacy and racial literacy. *Du Bois Rev.* 3 (2), 341–363.
- Valencia, R.R., 2010. *Dismantling Contemporary Deficit Thinking: Educational Thought and Practice*. Routledge, New York.
- Valenzuela, A., 2017. Grow your own educator programs: a review of the literature with an emphasis on equity-based approach. IDRA EAC-South.
- Villegas, A.M., Irvine, J.J., 2010. Diversifying the teaching force: an examination of major arguments. *Urban Rev.* 42 (3), 175–192.
- Zygmunt, E., Clark, P., 2021. Educator preparation and the rhetoric of "with liberty and justice for all.". In: Clark, P., Zygmunt, E., Tancock, S., Cipollone, K. (Eds.), *The Power of Community Engaged Teacher Preparation: Voices and Visions of Hope and Healing*. Teachers College Press, New York, pp. 1–11.

Diversity beyond its discourse: examining the discourse of diversity in non-Anglo-Saxon contexts

Tang T. Heng and Leonel Lim, Nanyang Technological University, National Institute of Education, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	289
Diversity as a social construction	289
Diversity discourse in Anglo-Saxon contexts	290
Diversity discourse beyond Anglo-Saxon contexts	290
The Singapore context	291
Background of study	292
Insights from Singapore teachers	292
Diversity as academic readiness	292
Curricular and pedagogical modifications around academic differences	293
Revelations of dominant societal narratives and concerns	293
Diversity as a social construction	294
The need to transform discourse and curriculum around diversity	295
References	295

Introduction

The term “diversity” has gained credence not only in educational settings in North America but beyond, as international institutions, corporations, and organizations increasingly reference the concept in their ethos and policies (Beech and Lista, 2012). Discourse around diversity, as with any other discourse, is subjected to recontextualization and transformation when it moves across borders and becomes varyingly interpreted by local actors (Bernstein, 2000; Lim, 2016a,b). This chapter aims to expand discourse around diversity beyond what is often referenced in Anglo-Saxon contexts, particularly in the US. By drawing upon literature beyond Anglo-Saxon contexts and sharing highlights from the authors’ previous case study of Singapore teachers’ perceptions of and attention to diversity, this chapter argues for diversity discourse to extend beyond Anglo-Saxon perspectives to engender more inclusive conversations and to understand its transformative possibilities and limits on the curriculum. To this end, diversity needs to be discussed vis-à-vis local contexts, by foregrounding how social, political, cultural, and geo-historical factors shape local definitions of diversity.

Diversity as a social construction

Diversity departs from difference. Whereas difference addresses the quality of variance, diversity is socially constructed with meanings and valuation assigned to group differences (Ghosh, 2018; Swartz, 2009; Lawson et al., 2013). Ethnicity, for instance, is a way of categorizing a group of people who identify with each other based on shared attributes or ways of living, hence the existence of *different* ethnic groups. However, it takes on another meaning when, upon being cataloged, evaluated and hierarchized, discriminatory behaviors or prejudicial attitudes are attached to *diverse* groups. How differences are conceptualized and socially constructed thus serves to divide people (Ghosh, 2018). To illustrate further, when African Americans are referred to as *diverse*, it reflects views of what is considered normal in the US context, namely White, Anglo-Saxon perspectives. Simultaneously, it speaks to the history of the US (briefly condensed) as founded and governed by White colonialists and missionaries, with slavery legalized to support the colonial elite’s interests, and the consequent marginalization of descendants of Africans who were forced to migrate.

Contemplating diversity as a social construction thus necessitates a scrutiny of the context within which diversity is invoked; history, politics, economics, culture, society and education shape and are shaped by the construction of social meanings, giving unique interpretations of diversity in various contexts.

(Heng and Lim, 2021, p. 2).

Diversity, therefore, is part of a “hegemonic discourse” (Swartz, 2009, p. 1045). It is not possible to speak of diversity alone. Diversity draws attention to what is different from the norm, through a deficit lens, with the norm yielding dominance over what is diverse (Lawson et al., 2013). Given how diversity discourse acquires unique meanings across different contexts, the next section sketches its development within Anglo-Saxon contexts. Particular attention is paid to the US as its demography, history,

politics, as well as English language, intersect to dominate the movement of discourse around diversity (Bank, 2012; Beech and Lista, 2012).

Diversity discourse in Anglo-Saxon contexts

School desegregation efforts, as with the US Civil Rights movement in the 1960s, drew attention to diversity within educational settings (Banks and Banks, 2019; Gollnick and Chinn, 2017). Consequently, educational initiatives—like multicultural education, ethnic studies, school desegregation research—flourished in attempts to include marginalized or absent groups of students into schools or the curriculum (Banks and Banks, 2010; Gollnick and Chinn, 2013). Driven by a vision of society that celebrates human rights, social justice, and freedom for all ethnic groups, these initiatives privileged cultural pluralism (Kallen, 1924).

Contemporary discourse on diversity in the US has since evolved from a narrow focus on ethnicity or culture, to encompass a wider variety of identity markers (Nieto, 2009). Through the years, as the scope of multicultural education broadened beyond ethnic studies and detecting textbook bias into other themes, like prejudice reduction, intercultural competence, and equity pedagogy (Bennett, 2001), so has the range of identity markers that have been associated with diversity. The widening range of identity markers reflects developments in the US society in face of the complex interplay of demographic, immigration, civil rights, political, and educational issues as well as the radical dissent and contestations characteristic of US history (Nieto, 2009). A cursory scan of textbooks in the US related to diversity sees diversity as frequently associated with race, culture, ethnicity, gender, class, language, sexual orientation, nationality, religion, physical and mental ability, exceptionality, immigration status (Vavrus, 2014; Banks and Banks, 2019; Gollnick and Chinn, 2017; Nieto and Bode, 2018; Sleeter and Grant, 2009).

Likewise, diversity discourse in Anglo-Saxon contexts such as Australia (Inglis, 2009; Ghosh, 2018), Canada (Joshee, 2009; Ghosh, 2018), and the UK (Tomlinson, 2009), make similar references to overlapping markers, like ethnicity, culture, and language. Gay's (2010) analysis of US textbooks used in teacher education courses on diversity found that textbooks tend to reference diversity predominantly around ethnicity. Similarly, studies on teachers' struggles with diversity in Australia (Hollinsworth, 2016; Mills, 2008), Canada (Parhar and Sensoy, 2011; Jacquet, 2008), the UK (Dunne and Gazeley, 2008), and the US (Castro, 2010; Ladson-Billings, 2006; Gay, 2010), appear to cluster around issues associated with White middle-income teachers' tensions in understanding their students' cultural, linguistic, and class challenges.

The predominant focus on issues of culture, language and class in research set within Anglo-Saxon contexts may seem natural to some. Given that diversity is a social construction that is typically construed in opposition to the dominant group, such a predominance echoes Ladson-Billings' (1999) observation that diversity "is that 'thing' that is other than White and middle class" (p. 219). Departure from Anglo-Saxon norms is seen "as deficit, as divisive, and as illegitimate" (Swartz, 2009, p. 1045), reflecting what is valued by the dominant group in society. Diversity, thus, needs to be "managed and domesticated" (Beech and Lista, 2012, p. 376); yet, as is evident in the subsequent section, diversity discourse beyond Anglo-Saxon contexts reflects different inflections departing from the discourse dominant in Anglo-Saxon contexts.

Diversity discourse beyond Anglo-Saxon contexts

In recent years, a growing body of work documenting diversity has emerged beyond Anglo-Saxon contexts (Beech and Lista, 2012). Hastened by educational policy borrowing and promoted by multi-lateral agencies and international education organizations, such as United Nations Educational, Scientific and Cultural Organization (UNESCO) and World Bank, diversity discourses have gained attention worldwide. While diversity discourses beyond Anglo-Saxon contexts similarly revolved—to some extent—around issues of ethnicity (Tsuneyoshi, 2018; Chang and McLaren, 2018; Jackson, 2018; Bhowmik et al., 2018), gender (Chang and McLaren, 2018), and class (Jackson, 2018; Chang and McLaren, 2018), nuances are evident given recontextualization through the local peculiarities of different countries.

In Japan, for instance, school children returning from abroad, are seen as diverse even though they share the same racial phenotype. These returnees prompted measures to "save" (Tsuneyoshi, 2018) them as they are seen as lacking language proficiency and the norms of acceptable behavior in both schools and the Japanese society. Likewise, mainland Chinese students who commute into Hong Kong for school—or, otherwise known as cross-border students—and new arrivals from mainland China are deemed diverse as linguistic minorities even though they are considered Han Chinese (Chang and McLaren, 2018). Yet, compared to cross-border students and families, ethnic minorities—like South Asians who, through voluntary and involuntary migration associated with British colonialism, have grown up in Hong Kong—are seen as even more diverse, lacking in Han Chinese culture and language (Bhowmik et al., 2018; Jackson, 2018).

In South and South-East Asia, diversity takes on other unique connotations. In India, diversity discourse tends to cluster around discussions of caste and religion, reflecting the dominance of the Brahmins and orthodox Hindus, as well as language and class, reflecting the British colonial legacy's valuation of the English language and class system (Ghosh, 2018). In Malaysia, religion and ethnicity are primarily used to identify who is diverse (Stephenson and Rajendram, 2019; Giacchino-Baker, 2000). Non-Muslims are deemed as "second-class citizens" (Giacchino-Baker, 2000) who are excluded from the economic privileges of being a *Bumiputra*, or son of the soil, as with non-Malay ethnic groups. This mechanism arose from a desire to respect pre-colonial indigenous populations and Malay immigrants from Indonesia and to recalibrate the economic imbalance, across the Malays, Chinese, and other ethnic groups, that the British had engineered (Stephenson and Rajendram, 2019).

Beyond Asia, Beech and Lista's (2012) work reveal how multi-lateral agencies' (such as UNESCO and World Bank) assumption of diversity, breaks down across contexts as it becomes conceptually ambiguous when positioned in different "cultural, ideological, socio-economic and political contexts" (p. 374). Citing Bolivia as an example, Beech and Lista showed how diversity departs from the European experience where immigrants of varying ethnicities are deemed diverse. Instead, indigenous communities marginalized through colonization are deemed "diverse" by Creole elites (descendants of European colonizers) despite the former's presence predating the latter's. UNESCO's "respect for diversity" was recontextualized as reclaiming the educational rights of Bolivia's indigenous population, even though they comprise more than 60% of the population. The notion of diversity, in the Bolivian example, reflects a complicated history of European colonization and dominance over indigenous communities, with dominance accorded to the groups wherein power resides, not necessarily the groups where numbers overrule.

These examples illustrate how "diversity acquires unique meanings in different contexts" (Heng and Lim, 2021) and, inadvertently, speak to the composition and concerns of dominant groups in these societies. The nuances in how diversity is interpreted departs from Anglo-Saxon perspectives and remind us of the need to foreground local peculiarities. Yet, conversations around and use of the term diversity, such as definitions and systems of classifications, have subliminally continued to center around Anglo-Saxon contexts (Beech and Lista, 2012; Yiu, 2019). Alternative ways of seeing diversity are marginalized (Yiu, 2019) or misrepresented (Yuan et al., 2020). Multi-lateral agencies and international education organizations' promotion of respect for diversity, for instance, reflects assumptions of Anglo-Saxon countries around ethnicity and culture and creates ambiguity and unintended consequences in destination countries (Beech and Lista, 2012).

Ironically, the diversity discourse set in Anglo-Saxon contexts risks becoming the dominant voice, with discourse departing from Anglo-Saxon contexts classified as alternative or diverse perspectives. A discourse that initially emerged out of a desire to include those that are marginalized and absent, with the potential to transform students' lives when enacted in the curriculum, now increasingly becomes mainstreamed with the power to silence others. While this discourse is beneficial in problematizing what is different and valued, in addition to setting examples for how dialogue and actions can ensue, it becomes limiting—when it is implicitly assumed to resonate in other contexts—and omits alternative voices. Given these partialities of the field, this chapter aims to expand and decenter the conversation around diversity beyond Anglo-Saxon contexts and to add to the nascent but growing body of knowledge on diversity worldwide. To this end, the chapter leverages a case study that examines how teachers in Singapore understood and addressed diversity in their classroom.

The Singapore context

Founded by the British in 1819, modern Singapore is home to immigrants from neighboring regions and afar. Singapore has historically been multicultural, attracting immigrants from China, South Asia, the Malay Peninsula, Riau Archipelago, and Europe (Lian, 2016). Contemporary Singapore has a population of 5.6 million with its resident population classified into four groups: Chinese (74%); Malay (13%); Indian (9%); and other ethnic groups (3%) (Department of Statistics, 2020). Multiculturalism, henceforth, is enshrined in its national policies, narrative, and ideology post-independence in 1965. Singapore's national pledge, for instance, declares that "regardless of race, language or religion," all citizens are united to "build a democratic society, based on justice and equality". Faced with a young nation comprising immigrants from different regions, the ruling party adopted multiculturalism in a bid to rapidly establish the country's national identity (Lian, 2016).

Over the years, the concept of multiculturalism has worked its way into governance policies, and official as well as private discourses (Lian, 2016; Ho, 2009). One example of multiculturalism's influence on policies includes the bilingual education policy where students learn their mother tongue during second language lessons and all lessons are taught in English to foster communication across ethnic groups. Another is the National Education program, a citizenship curriculum aimed at cultivating national identity and cohesion through Singapore's core values of multiracialism, meritocracy and multireligionism (Loh, 2014). These core values are promoted through the social studies curriculum as well as events associated with the National Education program, like Racial Harmony Day (where students learn about different cultures and dress up in ethnic costumes).

Despite the pervasiveness and emphasis on multiculturalism, scholars have critiqued the limitations of state-sanctioned policies and the concept of multiculturalism (Loh, 2022). For instance, Singaporeans and teachers are reluctant to join in conversations around topics like race and religion as they deemed the topics excessively sensitive (Alviar-Martin and Ho, 2011), and state-sanctioned diversity markers have seemingly erased differences that fall outside of those markers (Abu Bakar et al., 2021; Ho and Alviar-Martin, 2010; Lim and Tan, 2018). Teachers, as with students, tended to interpret diversity along state-sanctioned markers such as language, religion, and race, and are unaware of (or, hesitant to articulate) other conceptions of diversity.

Likewise, the concept of meritocracy has been subjected to criticism. Meritocracy works in tandem with multiculturalism: in a meritocratic system, all are equal to compete since a multicultural society promises a level playing field where no one culture is discriminated against (Koh, 2014; Lim and Apple, 2015; Lim and Tan, 2018). In the case of Singapore, merit has been associated with academic merit as its pioneering leaders argued that, being resource poor and small, investing in human capital was its most viable route to survival. Thus, education served the multi-pronged purposes of enhancing human capital, developing Singapore's economy, fostering national unity, leveling the society and promoting national identity (Gopinathan and Mardiana, 2013). Academic merit, arguably, seen as both objective and non-discriminatory, has consequently been woven into national and educational policies around tracking, promotion, access to scholarships, and prestigious civil service appointments, among others (Gopinathan and Mardiana, 2013; Lim, 2013). In enhancing social mobility through differential resource access, academic merit is thus

increasingly seen as a “positional good” (Kwek et al., 2019, p. 90). Yet, meritocracy appears to have created some unintentional consequences. Scholars pointed out that national standardized examinations have become so high stakes that they constrain teachers’ willingness to embrace more learner-centered pedagogy and assessment approaches (Heng et al., 2021; Heng and Song, 2020, 2021). Further, and especially pertinent to the present chapter, to the extent that the state’s discourse of meritocracy—particularly in its being putatively blind to individual’s social locations (class, ethnicity, gender, language, religion, etc.)—has evacuated a concerted focus on a politics of difference: teachers struggle with recognizing other kinds of diversity in addition to adopting more innovative and subtle forms of multicultural education (Alviar-Martin and Ho, 2011; Lim and Tan, 2018).

These details about Singapore set the stage for illuminating how diversity discourse outside the Anglo-Saxon context can differ markedly. The next sections sketch a study of ten Singaporean teachers’ perceptions of diversity and the curricular and pedagogical practices they undertake to address diversity. The research utilized data from a subset of a larger study of the same teachers’ perceptions and implementation of differentiated instruction (DI). As respect for student diversity is a core philosophy underpinning differentiated instruction, an educational approach championing instruction that is responsive to student needs (Tomlinson, 2014), a section of the study’s data—which was previously published (Heng and Lim, 2021) and is summarily treated here—lent itself readily to addressing the goals of this chapter.

Background of study

Ten participants from ten public schools were recruited via criterion sampling. Participants taught the nationally-examinable subjects of English, Mathematics, and Science, half taught in primary and half secondary schools, and all had teaching experience over two years. As part of the qualitative research data collection process, participants each completed a background questionnaire, three interviews (one each at the start and end of 2018) and four observation cycles.

The first interview elicited participants’ perspectives of student differences as well as their teaching and learning philosophies, the second, their insights on patterns observed over 39 lesson observation cycles. Each observation cycle comprises a lesson observation together with pre- and post-lesson observation interviews. During each cycle, two researchers kept qualitative notes, completed observation protocols separately, and met to derive consensus observation protocols based on their notes. Consensus observation protocols were subsequently analyzed, with the ratings and patterns reported to participants, prior to the second interview, as a form of member check and to collect data of their interpretations. Using Plano Clark and Ivankova’s (2015) method of “embedding” (p. 140)—i.e. including a secondary method of quantitative scales through the observation protocol in an otherwise qualitative approach—trustworthiness of the research was strengthened.

Inductive data analysis was performed, allowing for codes to emerge, as opposed to using pre-figured codes (Crabtree and Miller, 1999) that could carry inflections of Anglo-centric identity markers. For instance, when analyzing participants’ perspectives of student differences, recurring codes such as “mindset,” “motivation,” “disposition,” were grouped under “attitude” while “family background” or “single family” were sorted “family.” To further center participants’ voices and enhance trustworthiness, data were also compared across interviews, lesson observations, and analytic memos (Shenton, 2004). (For more details of the study’s research design and method, please refer to Heng and Lim, 2021).

Insights from Singapore teachers

Participants primarily saw student differences in terms of their academic strengths. This was evident in interviews—where they comprehended student differences predominantly as “ability,” “family,” and “disposition”—and in classroom observations, where curricular and pedagogical decisions revolved around students’ academic readiness. (Pseudonyms have been used below to protect participants’ identities.)

Diversity as academic readiness

In the first interview, participants were asked to state three important differences they saw across their students. The question deliberately used the word “difference,” as opposed to diversity, in case participants repeated oft-read Anglo-Saxon definitions of diversity. Participants’ top three responses banded around “ability,” “attitude,” and “family.” Participants defined ability as students’ “aptitude,” whether they “grasp things quickly” (Jon) and if they have “the ability in the mind to get or not get it” (Stella). Attitude was construed as students’ innate psychological strength or predisposition and discussed vis-à-vis educational outcomes. Participants talked about students differing in terms of their “motivation ... the attitude toward learning ... [with some] who are just very motivated to do something ... and [others] don’t budge” (Ping). They clarified that academic ability, while important, needs to be substantiated by “effort,” with effort articulated as “the will” versus a “mindset of giving up” and “determination ... [or] a drive that you want certain things” (Sharon). Relatedly, participants brought up family in terms of its structure—“single” versus “solid nuclear families”—and the “type of support parents give and values they instil” (Alfan). “Some kids come in advantaged ... [because] they have been propped up by mother or father or tuition,” (Matilda), others are “latchkey” as their “parents are busy working” (Zia). Participants primarily associated family with the level of academic support given to their children, explaining that family wealth, education or structure does not matter as much as whether “they [parents] will nag their kid” (Sharon) or support their children to study.

Despite the fact that the interview question was intentionally left open as “differences” participants saw, the congregation of participants’ responses around academic outcomes or studying was revealing. Ability was seen as “academic” ability (or how fast children

learn); attitude as the willingness to put in effort in learning; family as the level of support parents gave their child in studying. The gravitation toward differences around academic readiness and attribution of readiness to the individual or family—not wider institutions, cultures, or structures—is telling about the kinds of narratives and concerns dominating the larger society (and which will be discussed further below).

Curricular and pedagogical modifications around academic differences

Participants' modification of curriculum and pedagogy similarly reflected an emphasis on students' academic differences. Participants most commonly collected information on student readiness, and was scored 1.8 out of 4 (0 indicating no evidence observed in class and 4 indicating strong evidence). Readiness, according to Tomlinson (2014), is students' point of entry in relation to the knowledge or skill they are supposed to learn. A more ready student is seen as having mastered content or skill, while a less ready student needs more support and opportunities to master the learning objectives. While participants showed the strongest evidence collecting information on student readiness, they showed weaker evidence collecting information on student interest (0.6) and learning profile (0.8). Student interest is defined as what students exhibit affinity to or are curious about, while learning profile involves the ways in which students learn (Tomlinson, 2014).

Similarly, participants adjusted curriculum and pedagogy based on student readiness. When adaptive teaching was used, participants showed the strongest evidence in modifying instruction based on student readiness (0.8), followed by interest (0.6), and learning profile (0.4). Participants conducted assessment to determine academic readiness (e.g., administered Kahoot quiz or diagnostic worksheets to determine students' learning gaps), gave tiered assignments (i.e. assigned more challenging or extension tasks to academically-stronger students and more scaffolded tasks to weaker students), and focused more on academically-weaker students ("I help my rainbow group more because they're weaker" Zia).

Even then, adaptive teaching is infrequent as shown in the less than 1 scores across readiness, interest, and learning profile, pointing to a low appreciation of each child as an individual or recognition of different students' needs and how their needs could be addressed through the curriculum and instruction. Conversations in class tended to revolve around exam concerns, like "it's [the question] the easiest to score ... it's three marks ... a high value question" (Sharon) or teaching students procedural shortcuts that can help them score during exams. Some participants explained that choice was not offered to students as for "students who come from households that are not very structured ... it would be chaos ... it is very difficult for me to trust them," reflecting deficit perspectives around their students' academic capacities and family support for learning.

Revelations of dominant societal narratives and concerns

When participants were asked why readiness (or "academic ability") was privileged in how participants understood and responded to differences in class, they alluded to the national narrative of meritocracy, multiculturalism, equality, on top of a strong assessment and accountability culture.

Participants felt that highlighting and attending to student differences is contrary to the ethos of meritocracy where everyone begins from an equal footing and the most (academically) talented is sieved out. Participants pointed out that students and parents are also accustomed to the "fair as equal" mindset, thus prompting them to veer toward ensuring that students are "given the same things." Fairness occurs when "every child is given the same, regardless of their race or nationality" (Adel) so that meritocracy can perform its function.

The idea of meritocracy means that everyone has to be *given the same things* [emphasis added] whether or not they are able to perform. It's really up to them, it's their ability... I mean you and I are given the same things, so if you cannot perform then maybe it's because you're not working hard enough, or you're just lazy.

Alfan

To participants, examinations provide students an "equal playing field" as the outcomes (in this case, academic merit) can be "quantified" and evaluated against an "objective baseline." They acknowledged that standardized assessments are necessary to uphold a meritocratic and hierarchical society, where the "society is made up of many different levels Grading is an important thing to be able to sift out the different individuals to fit the different roles for society" (Adel). While some acknowledged that the playing field may not be as fair as it seemed, they rationalized that academic merit determined by standardized examinations was the best option until fairer alternatives present themselves.

Participants also exhibited discomfort in acknowledging and addressing differences, particularly along racial or cultural differences, revealing the effect of coupling meritocracy with multiculturalism. They worried about being seen as discriminatory and alluded to the sensitivities of discussing certain topics in Singapore. For instance, when asked why differences are rarely discussed in the classroom, Matilda explained that it is "very sensitive" to talk about differences "because we're very conscious of not offending and really don't know how other people will interpret what we say. In particular, she divulged that addressing cultural differences may result in her students thinking she is "racist." Participants' discomfort in discussing differences as it could be seen as "racist" and their perceptions of standardized exams as the most "equal playing field" reveal the constraining effect of meritocracy around discourses of differences, echoing other studies within the Singapore context (Abu Bakar et al., 2021; Ho and Alviar-Martin, 2010; Lim and Tan, 2018). Such assumptions and concerns obscure further investigations on the role of structural, cultural, and

institutional factors in students' lives and learning. This brings to question the plight of students with less economic, cultural, or linguistic capital, or whose ethnicity, nationality, or other identities may predispose them to less than optimal learning conditions.

Additionally, participants alluded to the tight coupling of the success-as-academic-merit mindset with Singapore's history, economic development, education approach, as well as assessment and accountability culture. They argued that since exams "sift[ed] out the different individuals to fit the different roles in society" (Adel) and "get [students] to a place that they want" (Sharon), there is an "intense pressure for academic excellence" (Stella), with teachers seeing their responsibility as helping students achieve the best results. This, they felt, was a legacy of Singapore's policies post-independence, as the national narrative revolved around the development of human resources as the only recourse for Singapore's development.

Our Singapore society has been defined and constructed [this way] ever since our independence ... we have only our manpower to rely on, because of limited resources, we need to achieve maximum efficiency for our students But eventually when the economy progressed ... the focus has changed ... but ... society hasn't quite moved on with the times, because we are still obsessed with the paper chase. Even the teachers also felt that, a measure of their own competency ... is defined as a success of students.

Alfan

Accountability, participants felt, is tied to the "performance of the class", and "results still matter, because everybody has their own professional reputation to uphold, nobody wants to be known as the teacher who cannot procure results" (Jon). While participants admitted that they were not averse to attending to students' needs beyond their academic readiness, they acknowledged that this has to be done post-exams, where "kids can be more creative" in the "freer, more relaxed environment ... where they don't have exams at the back of the mind" (Zia). They argued that student-centered approaches like differentiated instruction were "time consuming" (Adel/Jon/Matilda/Sharon/Stella) and impractical in light of the heavy syllabus. Meng Yong went as far as to admit that since success was defined through academic merit, "wider character development ... it's been sacrificed" in pursuit of academic excellence, while Sharon rued that the Ministry of Education's new policies for students to "find joy in learning rather than to just study for grades" is unrealistic as the "harsh reality" is that "post-secondary education really rely on the numbers [results]." According to participants, the tight coupling of the success-as-academic-merit mindset with Singapore's history, education, assessment and accountability culture constrain teachers' willingness to engage in educational approaches that value students' beyond their academic talents or that pays attention to other aspects of students' differences. Consequently, other aspects of students, for instance, their race, class, gender, or nationality, are bracketed away in the unparalleled focus on academic readiness (or merit).

Diversity as a social construction

This study of how teachers in Singapore perceive and respond to differences in the classroom illuminates the nuances around how diversity is seen in different sociocultural contexts. Teachers saw differences among students mainly through their academic readiness and related aspects, like student motivation and parental support. This predisposition toward academic readiness implicitly reveals the assumptions that dominate in society: that academic merit is to be expected and valued, and a lack thereof suggests variance from the norm. Diversity in this case, did not revolve around markers such as culture, language, and class—markers typically referenced in Anglo-Saxon scholarship. Diversity, in this case, also departed from those previously raised in Singapore scholarship, likely because this study involved teachers teaching a range of academic subjects, as opposed to only social studies teachers (Alviar-Martin and Ho, 2011), or teachers who were selected based on their understanding of culturally relevant pedagogy (Lim and Tan, 2020). Unlike previous studies where Singapore teachers expressed cognizance around state-sanctioned markers like race, language, and religion, teachers in this study were predominantly focused on seeing difference as academic readiness, to the extent of ignoring arguably less contentious aspects such as student interest.

Students, and parents, are seen unfavorably when they do not value or invest effort in academics, reflecting a deficit perspective and the wider hegemonic discourse (Lawson et al., 2013; Swartz, 2009). Participants' perspective of academic readiness as a combination of "ability," "attitude," and "family" allude to an unconscious predisposition to attribute differences to the individual or family, rather than the wider societal, institutional, or cultural influences. Alfan's observation that the meritocratic ideology places full responsibilities on the individual illuminates how academic weakness becomes privatized. The individualization and psychologization of differences in students' academic readiness and absence in interrogation of structural impacts on students' learning corroborate findings in extant literature around the privatization of academic underperformance (Jiang, 2016; Lim and Tan, 2018). The national narrative of meritocracy, in conjunction with non-discriminatory policies, appears to result in the perception of and attendance to student differences through varying levels of academic merit, echoing scholars' observations of the unintended consequences and constraints of meritocratic principles and policies (Alviar-Martin and Ho, 2011; Koh, 2014; Lim and Tan, 2018). This is further compounded by the strong assessment and accountability system, where high-stakes standardized exams determine students' futures and, consequently, evaluations on teachers' performances. To the extent that the above conditions perhaps suggest a weaker set of transformative possibilities for teaching practices that respond to diversity, it is important to point out that in less than liberal "strong" states that have historically assumed a larger presence in charting their society's moral and social orders, curriculums and their transformative potential have always been closely regulated by the state (Lim and Apple, 2016). Indeed, it may be that under conditions of meritocracy that constrain the elaboration of a discourse of social differences, the transformative potential of curriculums and pedagogies that seek to respond to diversity will remain rooted in dominant understandings of merit and ability.

Seen this way, in understanding diversity primarily in terms of academic readiness and subsequently differentiating their instruction to better cater to these differences so that students may be better prepared for the high-stakes examinations, the Singapore teachers presented were both working with and against meritocracy and its contradictory ideals around non-discrimination/egalitarianism and academic elitism (Lim, 2013).

The need to transform discourse and curriculum around diversity

This study illuminates how diversity is a social construction that is established within the wider and/or dominant context. Diversity, thus, needs to be understood as a relational concept and used with greater lexical precision. In using the phrase “diversity”, it cannot be assumed that readers or the audience recognize what it stands for as meanings of diversity are decoded and recoded across space and time. By continuing to talk about diversity in a taken-for-granted manner, diversity discourse ends up being normed against Anglo-Saxon discourse, even though markers typically associated with diversity—such as culture, class, language—are unique to Anglo-Saxon contexts. As the case of Singapore unveiled, presented teachers’ conceptions of diversity departed from culture, class, and language, clustering, instead, around the presence or absence of academic merit and its related effort. In sum, we argue that diversity needs to be articulated vis-à-vis its context. That is, when using the term “diversity,” it is essential to locate it within the discourse and practices in its context and foreground the history, culture, and modes of social organization to develop more accurate and encompassing understanding. For instance, in discussing diversity, it would be helpful to approach it as “diversity in Singapore/the US/context” or to include brief contextual insights, as opposed to merely referencing “diversity.” Analogically, these contextual references are akin to positionality statements in qualitative research, revealing authors’ awareness of the parameters of their lenses and their research’s transferability. The omission of contextual references and assumption that readers understand the term “diversity” per authors’ intent can potentially estrange and exclude readers from other contexts; it presumes that the authors’ context is the global context. Thus, the foregrounding of context makes diversity discourse more inclusive.

Consequently, diversity discourse, as well as curriculums around diversity, need to be deparochialized or provincialized to reflect its relational nature (Lingard, 2006; Takayama, 2016). Diversity-related curriculum, textbooks, and courses, for instance, would benefit from foregrounding how the concept is seen not just in the context or country of use, but beyond. Incorporating examples of how notions of diversity vary around the world, engenders greater sensitivity to different contexts and issues at play, reminding readers of ways of being and thinking they are unaccustomed to. The progression in genres of research related to multicultural education (Bennett, 2001) and incorporation of diversity-related courses in teacher education programs point to increasing recognition of the need to address diversity-related issues. Given the importance of these issues, more needs to be done to decenter the discourse so as to achieve epistemological inclusivity and greater inter-cultural understanding beyond Anglo-Saxon confines.

References

- Abu Bakar, M., Teng, S.S., Layne, H., Kaurani, S.N., 2021. Navigating diversities: experiences of youths in one Singapore school. *J. Multicult. Discourses* 1–20.
- Alviar-Martin, T., Ho, L.-C., 2011. “So, where do they fit in?” Teachers’ perspectives of multi-cultural education and diversity in Singapore. *Teach. Teach. Educ.* 27, 127–135.
- Banks, J.A., Banks, C.A.M., 2010. *Multicultural Education: Issues and perspectives*. John Wiley & Sons, Hoboken, NJ.
- Banks, J.A., Banks, C.A.M., 2019. *Multicultural Education: Issues and Perspectives*. John Wiley & Sons.
- Beech, J., Lista, E., 2012. Flowing discourses and border crossing: the slogan of ‘respect for diversity’ in Latin America. In: *World Yearbook of Education 2012*. Routledge.
- Bennett, C., 2001. Genres of research in multicultural education. *Rev. Educ. Res.* 71, 171–217.
- Bernstein, B., 2000. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Rowman & Littlefield, Lanham, MA.
- Bhowmik, M.K., Kennedy, K.J., Hue, M.-T., 2018. Education for all—but not Hong Kong’s ethnic minority students. *Race Ethn. Educ.* 21, 661–679.
- Castro, A.J., 2010. Themes in the research on preservice teachers’ views of cultural diversity: implications for researching millennial preservice teachers. *Educ. Res.* 39, 198–210.
- Chang, B.B., McLaren, P., 2018. Emerging issues of teaching and social justice in Greater China: neoliberalism and critical pedagogy in Hong Kong. *Pol. Futures Educ. Internet* 16, 781–803.
- Crabtree, B.F., Miller, W.L., 1999. *Doing Qualitative Research*. Sage Publications.
- Dunne, M., Gazeley, L., 2008. Teachers, social class and underachievement. *Br. J. Sociol. Educ.* 29, 451–463.
- Gay, G., 2010. *Culturally Responsive Teaching: Theory, Research, and Practice*. Teachers College Press, New York.
- Ghosh, R., 2018. Multiculturalism in a comparative perspective: Australia, Canada and India. *Can. Ethnic Stud.* 50, 15–36.
- Giacchino-Baker, R., 2000. New perspectives on diversity: multicultural metaphors for Malaysia. *Multicult. Perspect.* 2, 8–13.
- Gollnick, D.M., Chinn, P.C., 2013. *Multicultural Education in a Pluralistic Society*. Pearson Higher, Upper Saddle River, NJ.
- Gollnick, D.M., Chinn, P.C., 2017. *Multicultural Education in a Pluralistic Society*, tenth ed. Pearson Higher Ed.
- Gopinathan, S., Mardiana, A.B., 2013. Globalization, the state and curriculum reform. In: Deng, Z., Gopinathan, S., Lee, C.E. (Eds.), *Globalization and the Singapore Curriculum*. Education Innovation Series. Springer, Singapore. https://doi.org/10.1007/978-981-4451-57-4_2.
- Heng, T.T., Lim, L., 2021. Diversity, difference, equity: how student differences are socially constructed in Singapore. *Camb. J. Educ.* 1–20.
- Heng, T.T., Song, L., 2020. A proposed framework for understanding educational change and transfer: insights from Singapore teachers’ perceptions of differentiated instruction. *J. Educ. Change* 21, 595–622.
- Heng, T.T., Song, L., 2021. At the intersection of educational change and borrowing: teachers implementing learner-centred education in Singapore. *Compare. A Journal of Comparative and International Education*. <https://doi.org/10.1080/03057925.2021.1910014>.
- Heng, T.T., Song, L., Tan, K., 2021. Understanding the interaction of assessment, learning and context: insights from Singapore. *Educ. Res.* 63, 65–79.
- Ho, L.-C., 2009. Global multicultural citizenship education: a Singapore experience. *Soc. Stud.* 100, 285–293.
- Ho, L.-C., Alviar-Martin, T., 2010. Between self and state: Singapore teachers’ perspectives of diversity. *J. Int. Soc. Stud.* 1, 20.
- Hollinsworth, D., 2016. Unsettling Australian settler supremacy: combating resistance in university Aboriginal studies. *Race Ethn. Educ.* 19, 412–432.

- Inglis, C., 2009. Multicultural education in Australia: two generations of evolution. In: Bank, J. (Ed.), *The Routledge International Companion to Multicultural Education*. Routledge, New York.
- Jackson, L., 2018. Who is Hong Kong? Diversity in postcolonial Hong Kong curriculum. In: *Routledge International Handbook of Multicultural Education Research in Asia Pacific*. Routledge.
- Jacquet, M., 2008. The discourse on diversity in British Columbia public schools: from difference to in/difference. In: Gérin-Lajoie, D. (Ed.), *Educators' Discourses on Student Diversity in Canada: Context, Policy, and Practices*. Canadian Scholars, Toronto.
- Jiang, H., 2016. Revisiting individualism and collectivism: a multinational examination of pre-service teachers' perceptions on student academic performances. *Intercult. Educ.* 27 (1), 101–110.
- Joshee, R., 2009. Multicultural education policy in Canada: competing ideologies, interconnected discourses. In: Bank, J. (Ed.), *The Routledge International Companion to Multicultural Education*. Routledge, New York.
- Kallen, H.M., 1924. *Culture and Democracy in the United States*. Boni & Liveright, New York, NY.
- Koh, A., 2014. Doing class analysis in Singapore's elite education: unravelling the smokescreen of 'meritocratic talk'. *Glob. Soc. Educ.* 12, 196–210.
- Kwek, D., Miller, R., Manzon, M., 2019. 'Bridges and ladders': the paradox of equity in excellence in Singapore schools. *Equity Excell.* 87–108. Springer, Singapore.
- Ladson-Billings, G., 2006. It's not the culture of poverty, it's the poverty of culture: the problem with teacher education. *Anthropol. Educ. Q.* 37, 104–109.
- Ladson-Billings, G.J., 1999. Chapter 7: preparing teachers for diverse student populations: a critical race theory perspective. *Rev. Res. Educ.* 24, 211–247.
- Lawson, H., Boyask, R., Waite, S., 2013. Construction of difference and diversity within policy and practice in England. *Camb. J. Educ.* 43, 107–122.
- Lian, K.F., 2016. Multiculturalism in Singapore: concept and practice. In: *Multiculturalism, Migration, and the Politics of Identity in Singapore*. Springer.
- Lim, L., 2013. Meritocracy, elitism, and egalitarianism: a preliminary and provisional assessment of Singapore's primary education review. *Asia Pac. J. Educ.* 33, 1–14.
- Lim, L., 2016a. Knowledge, Control and Critical Thinking in Singapore: State Ideology and the Politics of Pedagogic Recontextualization. Routledge, New York.
- Lim, L., 2016b. Globalization, the strong state and education policy: the politics of policy in Asia. *J. Educ. Policy* 31 (6), 711–726. <https://doi.org/10.1080/02680939.2016.1181790>.
- Lim, L., Apple, M.W., 2015. Elite rationalities and curricular form: "Meritorious" class reproduction in the elite thinking curriculum in Singapore. *Curric. Inq.* 45 (5), 472–490. <https://doi.org/10.1080/03626784.2015.1095622>.
- Lim, L., Apple, M.W., 2016. *The Strong State and Curriculum Change: Assessing the Possibilities and Politics of Education Change in Asia*. Routledge, New York.
- Lim, L., Tan, M., 2018. Meritocracy, policy and pedagogy: culture and the politics of recognition and redistribution in Singapore. *Crit. Stud. Educ.* 1–17.
- Lim, L., Tan, M., 2020. Meritocracy, policy and pedagogy: culture and the politics of recognition and redistribution in Singapore. *Crit. Stud. Educ.* 61 (3), 279–295.
- Lingard, B., 2006. Globalisation, the research imagination and deparochialising the study of education. *Glob. Soc. Educ.* 4, 287–302.
- Loh, P.Y., 2014. Racial Harmony Day [Online]. Available: https://eresources.nlb.gov.sg/infopedia/articles/SIP_965_2004-12-31.html.
- Loh, S.H., 2022. The continued relevance of multiculturalism: dissecting interculturalism and transculturalism. *Ethn. Racial Stud.* 45, 385–406.
- Mills, C., 2008. Making a difference: moving beyond the superficial treatment of diversity. *Asia Pac. J. Teach. Educ.* 36, 261–275.
- Nieto, S., 2009. Multicultural education in the United States: historical realities, ongoing challenges, and transformative possibilities. In: *The Routledge International Companion to Multicultural Education*. Routledge.
- Nieto, S., Bode, P., 2018. *Affirming Diversity: The Sociopolitical Context of Education*. Pearson, USA.
- Parhar, N., Sensay, O., 2011. Culturally relevant pedagogy redux: Canadian teachers' conceptions of their work and its challenges. *Can. J. Educ.* 34, 189–218.
- Plano Clark, V.L., Ivankova, N.V., 2015. *Mixed Methods Research: A Guide to the Field*. Sage publications.
- Shenton, A.K., 2004. Strategies for ensuring trustworthiness in qualitative research projects. *Educ. Inf.* 22, 63–75.
- Sleeter, C.E., Grant, C.A., 2009. *Making Choices for Multicultural Education: Five Approaches to Race, Class, and Gender*. Wiley, New York.
- Statistics, D.O., 2020. Singapore Population and Population Structure Statistics. Available: <https://www.singstat.gov.sg/>.
- Stephenson, G.K., Rajendram, S., 2019. "This barrier between:" the ethnic divisions of higher education in Malaysia and the United Arab Emirates. *High Educ.* 77, 889–903.
- Swartz, E., 2009. Diversity: gatekeeping knowledge and maintaining inequalities. *Rev. Educ. Res.* 79, 1044–1083.
- Takayama, K., 2016. Provincializing and globalizing critical studies of school knowledge: insights from the Japanese history textbook controversy over "comfort women". In: Lim, L., Apple, M.W. (Eds.), *The Strong State and Curriculum Change: Assessing the Possibilities and Politics of Education Change in Asia*. Routledge, New York, pp. 161–179.
- Tomlinson, C.A., 2014. *The Differentiated Classroom: Responding to the Needs of All Learners*. ASCD, Virginia.
- Tomlinson, S., 2009. Multicultural education in the United Kingdom. In: Bank, J. (Ed.), *The Routledge International Companion to Multicultural Education*. Routledge, New York.
- Tsuneyoshi, R., 2018. Understanding multicultural Japan: four frameworks. In: *Routledge International Handbook of Multicultural Education Research in Asia Pacific*. Routledge.
- Vavrus, M., 2014. *Diversity and Education: A Critical Multicultural Approach*. Teachers College Press.
- Yiu, L., 2019. Developing a harmony-based approach to student diversity: multiculturalism in Taiwan's schools. American Educational Research Association, Toronto, Canada.
- Yuan, M., Sude, Wang, T., Zhang, W., Chen, N., Simpson, A., Dervin, F., 2020. Chinese minzu education in higher education: an inspiration for 'western' diversity education? *Br. J. Educ. Stud.* 68, 461–486.

Community-based education

Bianca J. Baldridge, Harvard University, Cambridge, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Historical context of community-based education	297
Features of community-based education: key findings	298
Pedagogical and philosophical diversity	298
Relationships, social capital, and advocacy	298
Activism and organizing	298
Considering sociopolitical context	299
Challenges in the field	299
Possibility and future(s): research, policy, and practice	299
References	300

Glossary

Community-based Educational Spaces (CBES) Community-based educational spaces include after-school programs, community-based youth organizations, youth centers, etc. The term “education” is used to decenter “schools” and center the learning and development within communities or informal learning contexts

Out-of-School Time (OST) Out-of-school time refers to activities children and youth participate in during non-school hours. OST includes programming that takes place before, during, or after school and on weekends, and during the summer

Youth Work Youth work is widely recognized as a set of coordinated activities engaging youth in many formal and informal contexts supporting young people’s academic and social development

Youth Worker Youth workers are adults who guide, educate, mentor, and nurture youth through various stages of development. Youth workers may also be called mentors, youth advocates, adult allies, community-based educators, youth specialists, etc.

In the public imagination, education is often synonymous with formal schooling. Yet children and youth have always learned within informal settings like community-based educational spaces (CBES). Community-based education occurs in many contexts, including families, faith-based spaces, non-profit organizations, museums, libraries, detention centers, parks, and neighborhood centers (Fusco, 2012; McLaughlin, 2000; Rogoff et al., 2016; Heath and McLaughlin, 1994). Community-based educational spaces (Baldridge, 2018) consist of afterschool and out-of-school programs and a wide range of nationally-based and grassroots organizations that guide students through academic, social, and emotional development. Distinguishing community-based education from school-based education is vital to center the agency and knowledge of communities in young people’s education.

The presence of community-based education helps to dislodge the myth that learning and education only occur in traditional school contexts. “Youth work”—a set of activities to support the care, education, and development of youth typically in non-school settings, is central to community-based education. Although youth work happens in various settings, community-based educational spaces are common sites for youth work.

Historical context of community-based education

The decline of child labor and the rise of compulsory schooling in the United States during the 19th Century created the need for afterschool child care. Newly arrived immigrants and children living in poverty were the primary participants of afterschool care in settlement housing, faith-based institutions, and early Boys Clubs (Halpern, 2002). Programming included English language development, play, citizenship education, and workforce skill development (Halpern, 2002). Non-European youth were primarily excluded from these early afterschool spaces.

During Black Migration, Black families settled in northern cities, placing them in contact with European immigrant communities. However, Black youth were excluded from afterschool clubs where European immigrant youth were present. Black communities were forced into specific neighborhoods in cities and were excluded from established clubs. Youth work took place within churches and other community programs established by the Black community (Baldridge, 2019; McKenzie, 2004). Other racially minoritized and marginalized communities facing oppression established spaces of affirmation and refuge for young people in their communities. Ultimately, what is recognized today as community-based education or youth work has always occurred within racially minoritized communities. The labels and terms we use today to describe education in communities outside of traditional schools may be new, but the work itself has always been embedded within historically marginalized communities.

Economic disinvestment in urban communities and lack of social safety nets characterized the late 1970s and 1980s. During this time, raced and classed-based policies criminalized Black and Latinx youth as problems to be controlled and contained (Ginwright, 2010). As such, community-based programs for youth emerged significantly during this time and engaged Black and Latinx youth in cities (Baldridge, 2019; Ginwright and Cammarota, 2002)—and in some cities, Southeast Asian and Pacific Islander youth (Kwon, 2013). The framing of community-based programs targeting these youth populations was often rooted in racist paternalism. Due to economic disinvestment and inadequate schooling, these programs were vital to youths' academic development, social support, safety, and well-being (Baldridge, 2019; Ginwright, 2010).

In 1994, The Clinton Administration signed into legislation the 21st Century Community Learning Centers (21CCLC) grants. This legislation provides opportunities for federal funding for afterschool programs to support college and career readiness. Schools and organizations that meet the criteria can apply for 21CCLC funding. Organizations apply for these funds while also relying on donors, local city support, and philanthropic foundations.

Features of community-based education: key findings

Research in the United States and the United Kingdom demonstrate that youth work and education in community-based spaces support youths' academic growth, facilitates greater attachment to school, builds youths' self-efficacy, and raises their critical consciousness, among other things (de St Croix, 2018; Ginwright and Cammarota, 2002; Hirsch et al., 2011; Vadeboncoeur, 2006). Besides supporting young people's academic identities, community-based education connects youth with caring adults to foster social, political, and cultural development. To be sure, the comprehensive approach to youth development within community-based educational spaces can be distinct from the formal and often punitive nature of traditional school spaces. Research studies with youth participants in community-based educational spaces reveal that many prefer attending after-school programs at community centers over schools because they feel treated with dignity and respect by adult youth workers (Baldridge et al., 2017; Douglas and Peck, 2013). Cultivating relationships between youth and adults grounded in respect and dignity for young people's humanity is an essential feature of community-based education.

Pedagogical and philosophical diversity

Community-based educational spaces are philosophically and pedagogically diverse (Baldridge, 2018). The structure of community-based education depends on its leadership, staff/youth workers, funding status and sources. Even though there are distinct differences in how community-based educational spaces carry out their work with youth, what is expected is the centering of youth and comprehensive youth development approaches (Brion-Meisels et al., 2020). It is also important to understand that leaders and youth workers in these settings approach their work seriously and are grounded in pedagogical practices that are rigorous and systematic.

Relationships, social capital, and advocacy

A key feature of youth work occurring within community-based education is the relationship building between youth and adult youth workers. The relationships that form in these spaces are often rooted in mutual respect and radical care and the humanity of young people and their multiple needs (Brion-Meisels et al., 2020; Ginwright, 2010; Paris and Winn, 2013). Research studies show that community-based education fosters social capital building in many forms, including institutional advocacy (Stanton-Salazar, 2011). Further, social capital building in community-based education affirms cultural wealth (Yosso, 2005) and supports cultural capital acquisition. For example, in programs that support youth through college and career access, youth workers broker connections with industries youth seek to be a part of; and provide material resources to support college and career readiness and lead youth to others who can support their transitions (Nelson, 2016). Some scholars have also extended how social capital manifests in urban community-based educational spaces. In a study with Black high school-aged youth in Oakland, California, Ginwright (2007) found that youth cultivated what he terms *critical social capital*—an intergenerational relationship between youth and adults in which youth identify and organize against sociopolitical problems in their schools and communities.

Activism and organizing

The acquisition of critical consciousness and sociopolitical development are precursors to engaging in activism and organizing (Kirshner and Ginwright, 2012; Watts and Hipólito-Delgado, 2015) and are vital features of some community-based educational spaces especially engaging Black, Indigenous, Latinx, and Asian and Pacific Islander youth. Community-based programs focused on racial identity development, gender justice, queer justice, racial and ethnic coalition building have been exemplars of engaging youth in organizing and activism (Kirshner, 2015). Research demonstrates that schools rarely engage youth in critical conversations about the world; whereas, community-based educational spaces are celebrated as flexible and unencumbered by the constraints of schools, to engage youth in "reading the world" (Freire, 1970/2018), to process and identify sociopolitical problems, and to act upon them (Kwon, 2013; Nygreen et al., 2006).

Considering sociopolitical context

Community-based educational spaces are celebrated for being settings where youth can thrive. They are often positioned in stark contrast to school settings where racialized and gender and sexually expansive youth experience harm. Yet, community-based educational spaces are not exempt from the same structural forces that shape the context of schooling (Baldridge, 2020a). The sociopolitical context of community-based education includes macro-level structural forces like racism, anti-blackness, capitalism, settler colonialism, heteronormativity, to name a few. These forces inform the construction of these spaces, philosophical approaches to youth development, funding patterns, pedagogical practices, as well as leaders, youth workers, and youth participants' experiences (Baldridge, 2020a; Fusco et al., 2013; Kwon, 2013; Lardier et al., 2018).

This sociopolitical landscape also includes how discourse and trends in educational policy shape how community-based education is delivered. For example, market-based approaches to an educational policy like No Child Left Behind (2002) and its emphasis on standardization and accountability—or Race to the Top (2009)—centered on competition—has led to success being measured by test scores, grades, and an organization's ability to “reduce achievement gaps” (Baldridge, 2014; Nygreen, 2017). More recent research has revealed how educational policy and educational philanthropy have shifted how community-based educational spaces are funded (Baldridge, 2019). For instance, some have argued that with the onset of privately-run charter management organizations in an era of neoliberal education reform, funders who have traditionally funded youth development organizations are now widely donating to charter schools instead of community-based educational spaces (Baldridge, 2014).

The reality of racism, white supremacy, heteronormativity, and capitalism, among other forces of domination, informs these spaces' legibility for racially minoritized youth. In other words, these forces can shape how community-based education is situated within the landscape of education. A reality of this context's impact is that community-based educational spaces can be positioned to “fix” or “save” youth from their circumstances. This framing is dangerous as it obscures the role of structural inequality and oppression that shapes young people's lives—especially young people experiencing multiple layers of racial, gender, class, and sexual oppression.

Challenges in the field

A strength of community-based education is its diversity; however, its heterogeneity also makes the field complex and challenging. The landscape of community-based education and youth work is vast, making it challenging to define. The sector is not one thing but a collection of many approaches, settings, and institutional arrangements that make up the field (Baldridge et al. 2022). The sector includes programming in city government agencies, cultural institutions, grassroots organizations, nation-wide programs, colleges and universities, and many other spaces that engage youth. Although these spaces serve youth, their ties, funding sources, and approach to youth development and education vary.

Because learning and education are analogous to schooling in the public imagination, community-based education and youth work are often overshadowed by learning in traditional school settings (Baldridge, 2020a). As such, community-based programs are undervalued as learning sites, and the educators within them are often delegitimized as educators (Vasudevan, 2019). The lack of a solid infrastructure does not exist; and, youth work does not typically require credentials and advanced schooling to enter the profession. Further, youth work does not fit neatly into one discipline and lacks a professional association as other disciplines have (Fusco, 2012; Yohalem, 2003). Although this may be a challenge for the field, some scholars recognize that requiring specific credentials may cause stratification and inequality. Many youth workers do not hold advanced degrees or even college diplomas but are highly competent and dedicated community-based educators (Baldridge, 2020b; Heathfield and Fusco, 2016; Watson, 2012). At the same time, research has shown that youth workers desire more support, professional learning communities, and opportunities to improve their practice (Yohalem and Pittman, 2006; Vasudevan, 2019).

Community-based education is marked by high turnover, and it is well-known that youth workers receive low wages and work long hours. In some studies, community-based youth workers from racially minoritized backgrounds report having few opportunities to be promoted to leadership positions (Baldridge, 2020b; Fusco, 2012; Pittman et al., 2004; Vasudevan, 2019). Generally, the field is characterized by instability and lacks strong infrastructure. In many cases, the flexibility offers greater choice and breadth for youth. However, the flexibility and instability have real material consequences for leaders and youth workers in community-based educational spaces.

Possibility and future(s): research, policy, and practice

Despite the challenges facing community-based education, there are immense possibilities within the field across research, policy, and practice. Decades of research reveal that young people—especially the most marginalized—find refuge, healing, radical care, and sometimes, tangible material benefits from their participation. Research continues to demonstrate the extraordinary impact of community-based education on young people's academic, social, political, and activist identities. Although the field's heterogeneity creates conundrums and tension for practitioners and researchers, it also provides ample opportunity to meet youths' specific needs.

Given the unstable infrastructure of community-based education, programs are often in financial jeopardy as funding patterns are uneven. Future policy directions can be instrumental in legislating more funding for community-based educational spaces or simply avoiding financial cuts to 21st Century Community Learning Grants and supplementary education at the federal level. Research continues to push the field in important directions concerning understanding youth workers and youth participants' experiences. More recent research is advancing the field in critical ways that raise questions about the broader sociopolitical landscape of the field—including forces of domination, educational policy, and philanthropy, and how it shapes the construction of community-based educational programs, youth and youth worker experiences. Research in this domain is necessary to uncover and disrupt processes that perpetuate harm to ensure that community-based education supports comprehensive, humanizing, and dignity-based programs for youth.

Community-based education is constantly in flux, but its future is bright. With additional research, political and public buy-in, supportive policy initiatives, and a solid infrastructure that values and protects youth workers and community-based leaders, community-based education can be valued as a legitimate space of learning and development for youth.

References

- Baldrige, B.J., Beck, N., Medina, J., Reeves, M., 2017. Toward a new understanding of community-based education: the role of community-based educational spaces in disrupting inequality for minoritized youth. *Rev. Res. Educ.* 41 (1), 381–402.
- Baldrige, B.J., DiGiacomo, D., Kirshner, B., Mejias, S., Vasudevan, D., 2022. New Directions for Learning and Development Outside of School: A Critical Bifocal Approach to Integrating Sociopolitical Context With Innovative Program Designs. Wallace Foundation Report.
- Baldrige, B.J., 2014. Relocating the deficit: reimagining Black youth in neoliberal times. *Am. Educ. Res. J.* 51 (3), 440–472.
- Baldrige, B.J., 2018. On educational advocacy and cultural work: situating community-based youth work[ers] in broader educational discourse. *Teach. Coll. Rec.* 120 (2), 1–28.
- Baldrige, B.J., 2019. Reclaiming Community: Race and the Uncertain Future of Youth Work. Stanford University Press.
- Baldrige, B.J., 2020a. The Youthwork Paradox: a case for studying the complexity of community-based youth work in education research. *Educ. Res.* 49 (8), 618–625.
- Baldrige, B.J., 2020b. Negotiating anti-black racism in liberal contexts: the experiences of black youth workers in community-based educational spaces. *Race Ethn. Educ.* 23 (6), 747–766.
- Brion-Meisels, G., Jessica, F., Vasudevan, D., 2020. At our best: Building youth-adult partnerships in out-of-school time settings. Information Age Publishing.
- de St Croix, T., 2018. Youth work, performativity, and the new youth impact agenda: getting paid for numbers? *J. Educ. Pol.* 33 (3), 414–438.
- Douglas, T., Peck, C., 2013. Education by any means necessary: peoples of African descent and community-based pedagogical spaces. *Educ. Stud.* 49, 67–91.
- Fusco, D., Lawrence, A., Matloff-Nieves, S., Ramos, E., 2013. The accordion effect: is quality in afterschool getting the squeeze? *J. Youth Dev.* 8 (2), 4–14.
- Freire, P., 1970/2018. *Pedagogy of the oppressed*. Bloomsbury, USA.
- Fusco, D., 2012. On becoming an academic profession. In: Fusco, D. (Ed.), *Advancing Youth Work: Current Trends Critical Questions*. Routledge, New York, NY, pp. 111–126.
- Ginwright, S., 2010. *Black youth rising: activism and radical healing in urban America*. Teachers College Press, New York.
- Ginwright, S., Cammarota, J., 2002. New terrain in youth development: the promise of a social justice approach. *Soc. Justice* 29 (4 (90)), 82–95.
- Ginwright, S.A., 2007. Black youth activism and the role of critical social capital in Black community organizations. *Am. Behav. Sci.* 51 (3), 403–418.
- Halpern, R., 2002. A different kind of child development institution: the history of after-school programs for low-income children. *Teach. Coll. Rec.* 104, 178–211.
- Heath, S., McLaughlin, M., 1994. The best of both worlds connecting schools and community youth organizations for all-day, all-year learning. *Educ. Adm. Q.* 30, 278–300.
- Heathfield, M., Fusco, D., 2016. Honoring and supporting youth work intellectuals. In: Pozzoboni, K., Kirshner, B. (Eds.), *The Changing Landscape of Youth Work: Theory and Practice for an Evolving*. Information Age Publishing, pp. 127–146.
- Hirsch, B.J., Deutsch, N.L., DuBois, D.L., 2011. *After-School Centers and Youth Development: Case Studies of Success and Failure*. Cambridge University Press.
- Kirshner, B., Ginwright, S., 2012. Youth organizing as a developmental context for African American and Latino adolescents. *Child Dev. Perspect.* 6 (3), 288–294.
- Kirshner, B., 2015. *Youth Activism in an Era of Educational Inequality*. New York University Press, New York, NY.
- Kwon, S.A., 2013. *Uncivil Youth: Race, Activism, and Affirmative Governmentality*. Duke University Press.
- Lardier, D.T., Herr, K.G., Garcia-Reid, P., Reid, R.J., 2018. Adult youth workers' conceptions of their work in an under-resourced community in the United States. *J. Youth Stud.* 21 (8), 1029–1044.
- McKenzie, B., 2004. Religious social networks, indirect mobilization, and African-American political participation. *Political Res. Q.* 57 (4), 621–632.
- McLaughlin, M.W., 2000. *Community Counts: How Youth Organizations Matter for Youth Development*.
- Nelson, I.A., 2016. *Why Afterschool Matters*. Rutgers University Press.
- Nygreen, K., Kwon, S., Sanchez, P., 2006. Urban youth building community: social change and participatory research in schools, homes, and community-based organizations. *J. Community Pract.* 14 (1–2), 107–123.
- Nygreen, K., 2017. Troubling the discourse of both/and: technologies of neoliberal governance in community-based educational spaces. *Pol. Futures Educ. Internet* 15 (2), 202–220.
- Paris, D., Winn, M.T., 2013. *Humanizing Research: Decolonizing Qualitative Inquiry With Youth and Communities*. Sage, Thousand Oaks, CA.
- Pittman, K.J., Irby, M., Yohalem, N., Wilson-Ahlstrom, A., 2004. Blurring the lines for learning: the role of out-of-school programs as complements to formal learning. *N. Dir. Youth Dev.* 2004 (101), 19–41.
- Rogoff, B., Callanan, M., Gutierrez, K.D., Erickson, F., 2016. The organization of informal learning. *Rev. Res. Educ.* 40 (1), 356–401.
- Stanton-Salazar, R.D., 2011. A social capital framework for the study of institutional agents and their role in the empowerment of low-status students and youth. *Youth Soc.* 43 (3), 1066–1109.
- Vadeboncoeur, J.A., 2006. Chapter 7: engaging young people: learning in informal contexts. *Rev. Res. Educ.* 30 (1), 239–278.
- Vasudevan, D., 2019. "Because We Care": Youth Worker Identity and Persistence in Precarious Work. Unpublished Dissertation. Harvard Graduate School of Education.
- Watson, V., 2012. *Learning to Liberate: Community-Based Solutions to the Crisis in Urban Education*. Routledge.
- Watts, R.J., Hipolito-Delgado, C.P., 2015. Thinking ourselves to liberation? Advancing sociopolitical action in critical consciousness. *Urban Rev.* 47 (5), 847–867.
- Yohalem, N., 2003. Adults who make a difference. In: *Community Youth Development: Programs, Policies, and Practices*, pp. 358–372.
- Yohalem, N., Pittman, K., 2006. Putting youth work on the map. *Forum for Youth Investment*, Washington DC.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.

Media education: history, frameworks, debates and challenges

Megan Boler, Amanda Trigiani, and Hoda Gharib, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Context and significance	301
The digital media landscape and its challenges for media education	301
History of media education	303
Media literacy education frameworks	304
Critical media literacies	304
Advertising literacy	305
Information literacy	305
News literacy	306
Digital literacies	306
Civic media literacies	307
Gaps & future directions	307
Centering affective dimensions	307
Assessing the needs of marginalized and immigrant communities	307
Adults as target population for disinformation media literacy	308
Checklists, fact-checking & information evaluation frameworks	308
Teaching lateral reading and process thinking	308
Engaging mindfulness techniques	308
Conclusion	309
References	309

While media education has a long history and takes shape under many names, schools of thoughts, and approaches, the fundamentals remain constant, aptly captured by the basic definition of media literacy which originated in the early 1990s, now updated, by the US based National Association for Media Literacy Education (NAMLE): “a framework to access, analyze, evaluate, create and participate with messages in a variety of forms—from print to video to the Internet. Media literacy builds an understanding of the role of media in society as well as essential skills of inquiry and self-expression necessary for citizens of a democracy” (medialit.org, accessed March 1, 2022). This entry is organized into Context and Significance, History of Media Education, Media Literacy Frameworks, and Gaps and Future Directions.

Context and significance

Each year, the media landscape poses daunting new challenges to 21st century education. In our era of ubiquitous media, people around the world spend more time online every year; almost 60% of the world’s population are now using the internet ([Johnson, 2021](#)). As of January 2020, 49% of the world’s population is using social media platforms with the average usage time rising to almost two and a half hours per day, which is expected to increase ([Statista, 2022](#)). Social media users are growing at a faster rate than pre-pandemic with the global total increasing at a rate of 13 new users every second ([DataReportal, 2022](#)).

Before the advent of the internet, exposure to media messages was limited to certain times and places. News was restricted to the daily newspaper or the evening television program, and to films available only in movie theaters. Media was unidirectional, with messages transmitted solely from producers to audiences who consumed content, but had little opportunity to respond or produce themselves.

The advancement of information and communication technologies makes it easier not only to access but to produce and circulate various forms of media to users around the world, radically changing not only the entire media ecosystem, but society as a whole ([Sagan and Leighton, 2010](#)). The early stages known as “Web 1.0” allowed internet users to create their own webpages and post their own content online, but these were usually static rather than interactive ([O’Reilly, 2007](#)). While often noted as an “industry-driven phenomenon,” the interactive and user-centered nature of Web 2.0 and emergence of social networking platforms now drive the participatory web ([Ellison and Boyd, 2013](#)). The digital revolution has altered the very nature of communication with users being one click away from connecting with anyone around the world.

The digital media landscape and its challenges for media education

Some describe the contemporary landscape as an “infodemic”—an overabundance of information that compromises logical thinking ([Durodolu et al., 2021](#); [World Health Organization, 2020](#)). Public opinion is increasingly influenced more by “emotions

and personal belief” than “objective facts,” resulting in what some call a “**post-truth**” era (Oxford Dictionaries, n.d.), a term named word of the year in 2017, its usage having risen 3000% in a single year (Boler and Davis, 2021; McIntyre, 2018). Media educators, librarians, and information professionals are left asking: how can we discern a shared reality amidst not only a deluge of “accurate” information sources but in this time of heightened deception by “pervasive and deliberate spread of misinformation geared towards deceitfulness”? (Durodolu et al., 2021, p. 175).

The proliferation of mis- and dis-information has become a key focus of educators and librarians. **Misinformation** refers to information that is unintentionally inaccurate or misleading, while **disinformation** is a type of intentionally deceptive information (Hwang et al., 2021) usually designed to achieve specific political ends (Benkler et al., 2018). Disinformation is often discussed as “**fake news**”, a buzz word that many scholars now deem as “vague and value-laden” (Marwick, 2018, p. 476), “used differently by different discourse communities” (Comber et al., 2018, p. 97). Defined as “false, often sensational, information disseminated under the guise of news reporting,” in 2017 the term “fake news” was named “Word of the Year” by Collins Dictionary—a year in which its usage rose by 365% (Flood, 2017). While false and sensationalized news are not modern phenomena by any means, as histories of news media and propaganda evidence, they are produced and circulated with the touch of a button (Boler and Davis, 2021; Osborne, 2018; Schudson, 1995).

The confusion surrounding “fake news” has led some to advocate the term “**problematic information**” to describe the range of false and misleading information that exists on a continuum with mainstream partisan media (Marwick and Lewis, 2017; Jack, 2017). While some forms of falsified news stories are particularly dangerous, all types play a key role in information warfare. A 2018 review of 34 articles suggested a typology of **six types of “fake news”**: news satire, news parody, fabrication, manipulation, advertising, and propaganda (Tandoc et al., 2018). Propaganda increasingly takes “new digital forms” that are pervasive and accelerate distribution, and continue to “blur the lines between entertainment, information, and persuasion” (Hobbs, 2017, p. 28).

Importantly, these new forms of propaganda are primarily utilized by the **far right**. Far-right extremist groups have developed “techniques of ‘attention hacking’ to increase the visibility of their ideas” (Marwick and Lewis, 2017, p. 2), creating conspiracy, community and spectacles using techniques that range from exploiting outrage to using memes, bots, social media and algorithmic recommendation systems to increase the visibility of their ideas (Farkas et al., 2018; Conway et al., 2019).

The evolving digital media landscape is shaped by numerous elements that have redefined our relationship to information and media.

Bots are automated software agents, programmed to collect data and conduct searches and speed up processes of information collection from the internet. Bots can radically influence public opinion, when used “to inflate the numbers of followers a politician has; to spread propaganda; to subtly influence political discourse; and to aggregate and broadcast content” (Marwick and Lewis, 2017, p. 38; Geiger, 2018).

Algorithms refer to the code that instructs the computer to perform specific tasks, playing a key role in determining what information is considered most relevant to us (McKelvey, 2014; see also Gillespie and Seaver, 2016; Brogan, 2016). Algorithms are increasingly recognized as embedded with “the agenda, biases, and vulnerabilities of the programmer” (Cohen, 2018, p. 145) raising concerns about **algorithmic governance** (Ziewitz, 2016). Algorithmic bias results, for example, in reinforcing stereotypes and social biases of race, gender, sexuality and ethnicity, privileging whiteness and emphasizing racist and gender stereotypes in searches (Noble, 2018).

Clickbait refers to the deliberate use of misleading headlines and thumbnails of content on the Web, whose main purpose is to attract attention and encourage clicks using sensationalist and emotional approaches to presenting information, thus easily spreading conspiracy theories and disinformation (Zannettou et al., 2019).

Echo chambers, or **filter bubbles**, describe the phenomenon of being enclosed within online communities of like-minded individuals, and thus not encountering alternative views and sources; as well, they reflect the degree to which algorithms shape and narrow what we see and who we interact with only (Rushkoff, 2010). Debates continue regarding the degree to which echo chambers may increase social and political polarization and extremism (Pariser, 2011; Flaxman et al., 2016).

While **propaganda** and news media have always sought to hook our emotions as a means of grabbing attention, the targeting of emotion and affect is prevalent within the ongoing media battles to shape, influence, and control public and political opinions. Viral media relies on the targeting of strong emotions such as rage, hate, and excitement (Crockett, 2017). **Polarization** refers in part to the gap in ideological perspective and values held by opposing parties (PEW Research, 2021), and stark disagreements regarding issues (like the Covid-19 pandemic) and differing levels of trust in media (Dimock and Wike, 2020). Polarized politics are increasingly linked to social media expressions of emotion and identity, especially surrounding race, immigration, and gender (Benjamin, 2019).

Confirmation bias refers to the tendency to search for and interpret information in a way that supports our prior beliefs, expectations, or hypotheses (Halpern, 2013). Confirmation bias is driven by **motivated reasoning** in which emotional biases lead to making decisions based on desirability rather than accuracy (Westen et al., 2006).

Distrust of the foundational institutions of democracy—government, media, and experts—by citizens from both of the political aisle has led to divisive online discourses and hyper-partisanship (Rainie et al., 2019). Public confidence in scientists and medical professionals continues to decline sharply: “confidence in medical scientists to act in the best interests of the public” dropped from 40% of American adults in 2020 to 29% in 2021 (Kennedy et al., 2022; Oreskes, 2021). Public confidence in the military, K-12 school principals, and police officers to act in the public’s best interest has also declined (Kennedy et al., 2022).

Skepticism directed at scientific authority is apparent in mis- and disinformation campaigns in the context of climate change and the COVID-19 pandemic. For example, climate change misinformation campaigns falsely frame the issue as a debate by

suggesting there is still doubt about the scientific findings (Damico et al., 2018); “elite-driven counter-movement” linked to conservative think tanks cultivates public skepticism regarding environmental protection (Jacques et al., 2008; Oreskes, 2021).

In sum, the digital media revolution has proven to be a double-edged sword: it gives everyone a voice, makes “everyone a journalist,” and provides means to catalyze protest movements like Black Lives Matter and #metoo (Boler, 2008; Leong et al., 2019; Ratto and Boler, 2014). But it also enables increased police surveillance of activists and BIPOC (Walsh and O’Connor, 2019; Benjamin, 2019), false information published with the touch of a finger, and the rise of surveillance capitalism, a “new logic of accumulation ... [that] aims to predict and modify human behavior as a means to produce revenue and market control” (Zuboff, 2015, p. 75).

For librarians and educators tasked with providing guidance in evaluating authority and credibility, this increasing distrust poses significant challenges (Nichols, 2017; Saunders and Budd, 2020). What counts as authoritative and credible information if there is no agreement about expertise? How can media education find the right balance given the need to critically assess the authority of information, while also maintaining a respect for expertise? How much responsibility lies with educators, versus with governments, social media platforms, advertisers?

History of media education

The rise of mass media in the 1920s represents a key moment for the roots of media literacy as radio, film and journalism became increasingly accessible. Across myriad countries, film-related education and outreach began through formal and informal educational contexts: from the Cinema Club movement in Paris (1922), to the British Film Institute (1933) which provided extensive educational outreach, and in Russia as well, though Stalinism quickly put this on hold.

Worries about propaganda emerged in the build up to the Second World War alongside the rise of journalism, radio, and broadcast television news. Edward Bernays, known as the “father of public relations,” contributed what essentially became the handbook for principles of propaganda (Bernays, 1928).

Educational efforts to counter propaganda took root in the 1930s, with the 1937 establishment of the Institute for Propaganda Analysis (IPA), an organization of journalists and educators which can be seen as an early arm of media literacy. Influenced by the study of Nazi propaganda, educators tended to see audience members as passive victims of mass media. The overlap of contemporary approaches to media education and the “ABC’s of Propaganda Analysis” developed in 1939 and the 2007 Core Principles of Media Literacy Education developed by the “National Association of Media Education and Literacy” is noted by Hobbs and McGee (2014).

A significant strand of what would come to be called the “protectionist” and “inoculation” theory of media education surfaced early on, with the 1938 UK Spens Report depicting media and popular culture as a contagious threat to civilized values. The inoculation approach was also reflected in what is termed the Leavisite tradition of Britain, which would be radically challenged by British cultural studies in the 1980s. F.R. Leavis’s 1942 curriculum sought to “inoculate” students from the cheap commercialized content of “lowbrow” popular culture that was “infecting” Britain’s “high cultural” values of civilized literature (Masterman, 1980).

The debates and discourses of “protectionism” and “inoculation” of youth against media’s ill effects overlaps with other significant media studies debates, such as those surrounding technological determinism and the “media effects” model. In Canada, Marshall McLuhan’s radical theories (e.g., McLuhan, 1964) developed the thesis of the media effects model. It is not just the content of media, he argued, but their form itself that communicates the message: the medium is the message. In part thanks to McLuhan’s influence, Canada was an early leader in media education.

The uses and gratifications approaches to understanding human relationship to media that took hold in the 1960s and ’70s in America contrasted with models of passive viewership, introducing a functionalist view and “cognitive revolution” in media theory. While the media effects model privileged the perspective of the communicator and medium, uses and gratifications focused on audience perspective, and the “social and psychological needs” “satisfied by exposure to mass media” (Ruggiero, 2000, p. 6).

The 1970s and ’80s saw the influence of Paolo Freire’s critical pedagogy which provided the seeds for critical media literacy (CML), and his philosophy and literacy praxis which involved critical consciousness and learners “reading the world” (not just the word). In 1980 Len Masterman’s highly influential text *Teaching the Media* was published, a new touchstone for media education around the world.

In 1982, UNESCO hosted a symposium with educators, communicators and researchers from nineteen countries, who came together for a five-day international symposium, the result of which was the unanimously issued Grunwald Declaration on Media Education, which urged educators, parents, communicators to accept the central and pervasive role media had come to play in our lives. Lamenting the absence of formal and informal efforts in media education, the Report advised: “Political and educational systems need to recognize their obligations to promote in their citizens a critical understanding of the phenomena of communication Responsible educators will not ignore these developments, but will work alongside their students in understanding them and making sense of such consequences” (Grunwald Declaration, 1983).

Debates that emerged in the 1980s continue to this day: Can media education protect students from the pernicious effects of media depictions of violence, pornography, body image, and harmful stereotypes around race, class, gender, sexual orientation, etc.? Do violent video games promote and cause violent behaviors? Is Instagram causing tweens and teens to strive toward impossible body images and social status, promoting eating disorders and self-harming behaviors?

Other debates emerged in the 1990s, for example in the context of the first media education conference in the United States in Rowe (1995). Conference events and sessions revealed a significant divide in educators and scholars' perceptions regarding the aims of media literacy, coalescing around such questions as: How "evil" is mass media and popular culture, and should educational efforts accept financial support from media industries (see Huesmann et al., 1983)? Should media literacy primarily be the province of schools (Aufderheide, 1993; Alvermann and Hagood, 2000; Hobbs, 2008)? Should media production be an essential feature of media literacy education (Lewis and Jhalley, 1998; Hobbs et al., 2013)? To what extent should media literacy have an explicit political and ideological agenda? (Hobbs, 1999).

Debates at the 1995 conference regarding the extent to which media literacy should have explicitly political and ideological agendas resulted in the conference splitting into two distinct organizations—one with a more activist and political view, the Action Coalition for Media Education (ACME) and the other, now known as the National Association for Media Literacy Education, NAMLE which saw media education as more politically neutral.

Sparked by the Universal Declaration of Human Rights, in 2011, the European Commission called for "strategic media education in compulsory schooling" (McDougall et al., 2015, p. 5), seeking to put into practice the human right "to seek, receive and impart information and ideas through any media and regardless of frontiers" in order to "empower citizenries all around the world to have the full benefits of this fundamental right" (UNESCO, 2011, p. 4). The UK has consistently ensured a focus on media literacy; for example, in the 1950s establishing two renowned sources of media education early on British Society for Education in Film and Television and commencing the publication of *Screen*, academic journal of film and television studies (see <https://www.frankwbaker.com/mlc/media-literacy-history/>). Hungary, Slovakia and Czech Republic were early leaders integrating media literacy into secondary school curricula (Federov, 2014). Australia has long advanced media literacy curricula, dating back to the 1940s (Corser et al., 2021). In Canada, the Association of Media Literacy was founded in 1978 by Barry Duncan; the province of Ontario was the first province to mandate high school media literacy curricula in 1987, and it added this for K-8 in 2006 (Jolls and Wilson, 2014; 2021). Research found that the United States is the "leading exporter of media" yet having the fewest media education programs within schools (Kubey, 1998). The 2020 U S Media Literacy Policy Report reveals that in the US, "only 14 states have taken substantial legislative action for media literacy education" (Hobbs and Jensen, 2009; Media Literacy Now, 2021).

Media literacy education frameworks

Conceptualizations of literacy have undergone several transformations in the second half of the 20th century and into the 21st century that have largely been mimicked by media literacy and its sub-types (Ashley, 2019; Cappello, 2017; Pangrazio, 2016; Zarouali et al., 2019). Literacy was traditionally understood as being the ability to read, write, and understand text or arithmetic (Cappello, 2017; Malmelin, 2010), an "umbrella concept that comprises different forms of reading and interpretation" and involves an "individual's personal ability to understand different kinds of signs and symbolic systems, and ... produce different kinds of messages by using these symbolic systems" (Malmelin, 2010: 132). In the late 1970s, scholars from New Literacy Studies argued that literacy is not a neutral practice that is linked only to psychological abilities, but rather a situated practice that is closely tied to political, institutional, and cultural relationships and practices (Cappello, 2017). Later, technological advancements raised questions about whether it was enough for literacy to be concerned with written text or if it should also include the "reading" and production of images, videos, music, and other forms of communication (Hobbs, 1997; Alvermann and Hagood, 2000; Gee, 2003). In 1996, The New London Group advocated the "multiliteracies" approach to account for the way learners navigate multiple literacy methods including visual, spatial, and multimodal (New London Group, 1996; Luke, 2005). At this same time, the growing media literacy movement defined literacy as "the ability to access, analyze, evaluate, and communicate messages in many forms," which also reflects recognition of multiliteracies (Hobbs, 1997, p. 169). It is from this definition that media literacy became defined as the ability to access, evaluate, analyze, and produce media of a diverse kind (Aufderheide, 1992).

The following descriptions of variations of media literacy each emphasize different aspects of media or different approaches, philosophies, or pedagogical aims.

Critical media literacies

CML, also known as critical media education, works from the basic definition of media literacy: the ability to access, analyze, evaluate, and communicate information (NAMLE, 2007). However, CML moves beyond the arguably more "instrumental" forms of media literacy to incorporate analysis of the political, economic, cultural, and historical contexts of media messages: media messages are neither neutral nor do they exist in a vacuum. All media serve as agents of socialization that promote values, cultures, and interpretations reflecting the interests of their producers and funders (Domine, 2011; Ashley, 2019). The question of who controls the media, who benefits from the meanings and values embedded in representations, are driving questions of critical media education.

CML is tasked with providing skills to deconstruct the symbols, signs, codes and conventions that produce meanings that become "naturalized" as taken-for-granted dominant cultural "truths" reinforced by mass media and stereotyped representations. CML education involves "cultivating skills in analyzing media codes and conventions, abilities to criticize stereotypes, dominant values, and ideologies, and competencies to interpret the multiple meanings and messages generated by media texts" (Kellner and Share, 2005, p. 372). Teaching students to be critical of media representations and discourses that (re)produce oppressive

norms prepares learners to navigate media while understanding broader contextual issues of power, ideology, and intent (Funk et al., 2016; see also Luke, 1997; Buckingham, 1996).

CML has roots in numerous traditions: the Frankfurt School (Adorno and Horkheimer, 1944); critical literacy and pedagogy traditions (Freire and Macedo, 1987; Luke, 1997); and the Birmingham School of Cultural Studies and the work of Stuart Hall (1980). Critical and radical pedagogy are grounded in Gramsci's conception of "hegemony," which describes how the dominant class maintains control by "naturalizing" ideologies. Freire's emphasis on critical consciousness or "Conscientization" understands literacy as far more than a set of skills but rather a way of "reading the world," of understanding how hegemony works—how socio-economic and political powers construct our very perception of the world (Garcia et al., 2013; Freire and Macedo, 1987; Freire, 1970). Neo-Marxist scholars (such as Henry Giroux, Peter McLaren, Stanley Aronowitz, Colin Lankshear) theorized the field of critical pedagogy. The work of Stuart Hall and the Birmingham School of Cultural Studies significantly shapes CML, understanding of meaning making "as a cultural struggle between dominant readings, oppositional readings, or negotiated readings" (Kellner and Share, 2007b, p. 8; see also Hall). CML also adopted the cultural studies understanding of viewers not simply as passive spectators or "victims" of media but as "active constructors of reality" (Kellner and Share, 2007a, p. 63).

Significantly, CML counters the longstanding "protectionist" anxieties of media studies, instead embracing both the pleasures of and the need to deconstruct popular cultural texts. Although popular culture and mass media may contribute to social problems, such media need to be more fully contextualized (Kellner and Share, 2005, p. 60; see also Buckingham, 2002; Jenkins, 2009). Early on, CML embraced media arts production as a means for young people and marginalized communities to create their own representations, and engage media in struggles against oppressions (Alvermann and Hagood, 2000; Morrell et al., 2015). By 2015, scholars and educators recognized young people's engagement as producers, "spending the majority of their time as consumers and producers of media," which of course has both "positive and potentially harmful effects" (Morrell et al., 2015).

While critical media education focuses centrally on deconstructing stereotypes, not surprisingly scholars and teachers point to the need for strategies specifically tuned to the needs ranging from a focus on equity, anti-colonialism, critical race, and native and indigenous media literacy. A 2020 issue of the *International Journal of Multicultural Education* contributes 11 articles providing diverse approaches to "critical race media literacy," which constitutes "a frame through which teachers and students can confront racist, colonial, classist, ableist, and heteronormative representations in media; interrogate the messages represented in problematic media messages; and navigate the complexities of representation in the classroom and society" (Hawkman and Shear, 2020, p. 2; see also Yosso, 2002; Scharrer and Ramasubramanian, 2015). Work on media education and decolonization is more recent (Romero Walker, 2021). A 2020 special issue of *The International Journal of Critical Media Literacy* provides cutting-edge work on CML and hip hop education (Graves et al., 2020).

Advertising literacy

Advertising first came under intense social scrutiny and concern in the late 1950s when market researcher James Vicary fraudulently claimed to have successfully increased the sales of products like Coca-cola and popcorn through the use of subliminal cues (Mal-melin, 2010). Contemporary advertising literacy (AL) focuses centrally on consumer's ability to recognize and critically evaluate advertisements (Rozendaal et al., 2011; Hudders et al., 2017) and its goal is to empower individuals with the knowledge, skills, and attitudes to protect themselves against harmful effects of advertising and be critical consumers (Jeong et al., 2018). The core AL skills include understanding persuasive intent and tactics, recognizing advertisements when they are present, and activating the right knowledge and attitude structures to critically evaluate advertising (Rozendaal et al., 2011).

The limitations associated with peoples' ability to recognize, and subsequently resist, advertising is a growing concern in the 21st century. The shift from print and TV advertisements to those embedded into other forms of media have made recognizing the presence of advertising an even greater challenge than before (Hudders et al., 2017; Shade and Chan, 2020). Newer forms of marketing (e.g., stealth marketing) that are subtle, emotionally engaging, and embedded into other forms of media, have become ubiquitous yet paradoxically more difficult to spot (Calvert, 2008; Lampert et al., 2021; Sweeney et al., 2021). It is on these grounds that advertising forms such as native advertising and influencer marketing have been criticized as potentially unethical, particularly for children (Lampert et al., 2021; Sweeney et al., 2021).

Information literacy

Developed and engaged predominantly within the area of library and information science, information literacy (IL) has a long and venerable history. IL is the ability of an individual to "(1) recognize their information needs; (2) locate and evaluate the quality of information; (3) store and retrieve information; (4) make effective and ethical use of information; (5) apply information to create and communicate knowledge" (Catts and Lau, 2008).

Unlike other forms of media education, IL has had advocacy and visibility at the national and governmental level for decades (Sample, 2020). As early as the 1970s, librarians began emphasizing its importance, particularly as a civic skill within democratic societies. In 1989, the American Library association's Final Report outlined IL as central to the education of citizens in a well-functioning democracy: "All men are created equal but voters with information resources are in a position to make more intelligent decisions than citizens who are information illiterates. The application of information resources to the process of decision-making to fulfill civic responsibilities is a vital necessity" (O'Connor, 2009).

The 1989 Final Report of the American Library Association Presidential Committee on Information Literacy further highlighted the need for students to “play an active role in knowing, identifying, finding, evaluating, organizing and using information” (Mackey and Jacobsen, 2011). In 1991, the Second White House Conference on Libraries and Information Services (WHCLIS) recommended that the US government establish a National Coalition for Information Literacy that would involve not only schools, libraries, and parents but labor, industry, government, and the public to develop “a strategic plan for the general development of skills required for information literacy” (Behrens, 1994, p. 319). The American Association of School Librarians (AASL) in 1994 adopted official national standards, subsequently adopted by the National Forum on Information Literacy: (1) defining the need for information; (2) initiating the search strategy; (3) locating the resources; (4) assessing and comprehending the information; (5) interpreting the information; (6) communicating the information; and (7) evaluating the product and the process (Doyle, 1994). Across the Atlantic, the SCONUL Seven Pillars Model for Information Literacy was developed in 1999 by the Society of College, National and University Libraries in the United Kingdom (Mackey and Jacobson, 2011). In addition to the US and UK, the United Kingdom, Australia, and Mexico adopted national IL standards. Tribal college and university (TCU) libraries Diné College in Tsailé, Arizona have developed an approach to IL that integrates non-traditional Indigenous knowledges and sources, specifically the four educational principles constituting the Diné traditional living system, *Sq’ah Naagháí Bik’eh Hózhóón* (Sorrell, 2019).

News literacy

It wasn’t until the early 21st century that news literacy education became recognized as a sub-type of media literacy (Ashley, 2020). After 2016, with the rising anxieties regarding “post-truth” media and “fake news,” global interest in news literacy has increased dramatically (Kajimoto and Fleming, 2019). While definitions of news literacy and media literacy overlap significantly (e.g., Fleming, 2014; Sivek, 2018; Swart, 2021), news literacy is centrally concerned with “actionable information” (Klurfeld and Schneider, 2014) and focuses on the “veracity and reliability of information by scrutinizing the origin of particular information” while media literacy tends to emphasize deconstructing media messages and meanings “created with that information” (Kajimoto and Fleming, 2019, p. 7; see also Maksl et al., 2015; Tully et al., 2021; Vraga et al., 2021). News literacy curricula are inspired by three educational foci: distinguishing between fact and fiction, understanding the democratic mission and freedom of the press, and more recently, using the “markers of high-quality journalism” to help students understand the “principles of solid, fact-based news storytelling compared with the increasing popular and politically polarized cable news sphere” (Kajimoto and Fleming, 2019, p. 4).

“Critical news literacy” is a nascent area, likely to explode given the increasingly urgent need. Arguably, news literacy has thus far taken a rather narrow approach, focusing on spotting hoaxes, checking facts and sources, and verifying claims. As numerous scholars argue, it’s necessary to think like sociologists (rather than journalists as per tradition of news literacy): in addition to critically analyzing individual news messages and articles, we must also understand market forces, editorial processes, and journalistic routines and practices that shape news reporting and their effect on people and societies (Ashley, 2020; Vraga et al., 2021; Kajimoto and Fleming, 2019). The five “C’s” framework of news literacy—context, creation, circulation, consumption, and content—additionally emphasizes “news literacy behaviors,” such as fact-checking or verifying claims, seeking news and political content from a variety of sources, and conscious processing of news messages (Tully et al., 2021).

Digital literacies

Originally used in the 1990s to refer to the ability to read and understand hypertext, the term “digital literacy” has come to mean much more in present times. In 1997, Glister defined digital literacy as “the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers” (Reddy et al., 2020, p. 8). Although distinguishing it from media literacy has often proven complicated (Pangrazio, 2016), digital literacy pays particular attention to personal and technological competencies for living in a digital society, while media literacy focuses on critical engagement with mass media messages. Contemporary digital literacy includes themes of online safety, access and navigation skills, online relationships, and digital rights and responsibilities. From its conception, digital literacy scholars advocate for seeing technology as more than simply machines used for information seeking, but as “cultural forms” that provide “new ways of mediating and representing the world and of communicating” (Buckingham, 2006, p. 74).

Digital literacy advocates that users be actively engaged while contributing to and creating online spaces (Knobel and Lankshear, 2008). While earlier conceptions of media literacy focus on the importance of individuals as critical consumers of media, digital literacy emphasizes consumers as participants as well. Participating in meaning making requires that users develop technical skills and knowledge of evolving digital tools to become active participants in and producers of the digital world (Erstad, 2008). Digital literacy, therefore, privileges making rather than just consuming, interactive construction rather than deconstruction, and pushes beyond critique to creation and transformation (Martin and Grudziecki, 2006).

A criticism of digital literacy is that it neglects questions of power and ideology—issues central to CML—with its focus on digital tools. Critical digital literacies essentially combines CML and DL, expanding digital literacy to incorporate critical thinking, anti-oppression, and multimodal formats of expression and representation. It can be defined as the “ability to read and create digital texts in a reflective way to identify and ameliorate the power, inequality and injustice in human relationships” (Hughes and Morrison, 2014, p. 1).

Civic media literacies

One of the newest forms of media literacy is civic media literacy (Kahne et al., 2012; Mihailidis, 2019). As its name suggests, it merges traditional media literacy education with civic literacy education, focusing on the provision of skills and knowledge necessary for students to be informed and engaged citizens (e.g., voting, being informed, knowing how to exercise rights at all levels of civic life from local to global). It seeks to move beyond critique to emphasize active civic participation.

Civic media literacy emerged from critiques that media literacy education does not focus enough on developing citizens' civic skills. Mihailidis (2019) argues that media literacy has prioritized the teaching of content-based skills such as deconstructing messages, critical thinking, and critical media production as well as knowledge of media systems and political and historical contexts but has left a gap "between skill attainment and civic outcomes [that] is apparent in the lack of action-taking or civic participation embedded in media literacy practices" (Mihailidis, 2019, p. 3).

As civic responsibilities increasingly move beyond in-person or offline forms of engagement, proponents of civic media literacy urge media education to catch up (Hodgin and Kahne, 2018). Media and digital literacy tend to neglect the importance of online circulation as a form of civic engagement (Middaugh, 2019), neglecting "discussion about how acts of circulation harm or help the goals of representation and productive public discourse" (p. 26). Audiences retain responsibility as "gatekeepers of information" (Roschke, 2018) who filter, verify, and spread information. Media literacy education is called to bring to the forefront the everyday engagement that is often local, nonlinear, and intentionally or unintentionally civic (Middaugh, 2018). Additionally, civic media literacies question how and why we see and share certain content over others, and how we come to know the difference between reliable, truthful, and/or problematic information (Damico and Panos, 2018; Hodgin and Kahne, 2018).

Gaps & future directions

Centering affective dimensions

Media literacy has historically focused exclusively on cognition, ignoring the affective dimensions of production and consumption. A review of contemporary scholarship on media education reveals an absence of sustained attention to the role of emotion and affect within media. However, scholarship in media education increasingly recognizes the urgent need of attending to emotion and affect, and challenging the rationalist traditions that have shaped media literacy approaches and theories (Sivek, 2018).

A review of nine handbooks on media education and "fake news" published between 2017 and 2018 reveals a glaring absence of attention to the central role of emotion within contemporary media, most neglecting any mention of emotion or affect and instead mentioning "cognitive bias" (Boler, 2019). Media education is thus challenged to consider in greater depth how social media and problematic information hook our emotions, and how contemporary forms of propaganda and marketing target and weaponize our emotions through the potent marriage of big data and cognitive psychology (Boler and Davis, 2021).

One way for media education to address the ways in which our emotions and affects are increasingly targeted may be engaging a pedagogy of discomfort to help us recognize when and how we are "hooked" or "triggered" by media, and to disrupt emotional habits and attachments (see Boler, 1999). Adapted to media education, a pedagogy of discomfort invites students to apply critical inquiry and skepticism to self-reflexive practices: What emotional investments do we have in particular cultural narratives and myths? What is our motivation for circulating the news we share? When and how are we influenced by emotional peer-pressures to adopt particular beliefs and identities? A pedagogy of discomfort provides a starting place to invite both students and educators to re-evaluate the unconscious emotional habits, patterns, and assumptions that shape our complex, affective relationships perhaps to media ranging from news media (Sivek, 2018) to representations of race, gender, and sexuality (Nadan and Stark, 2017; Gachago et al., 2013; Diallo and Friberg, 2021; MacEntee and Madrona, 2015).

Assessing the needs of marginalized and immigrant communities

An underdeveloped area of media education scholarship is one that considers the specific needs of BIPOC and new immigrants when dealing with affective, social, and structural dimensions of problematic information. As the need for media literacy and digital skills becomes more prevalent in everyday working life, especially as a result of the COVID-19 pandemic, the gap between global online and offline populations keeps widening. UNESCO's Broadband Commission for Sustainable Development (2020) estimates that nearly 3.6 billion people remain totally unconnected. Even countries with a high overall connectivity rate, gaps in availability, affordability, and digital skills exist in remote and marginalized communities, which often coincides with other disparities such as income, gender and race-based inequality.

Although it is not new for media educators to consider the targeting of racial and social minorities and examine stereotypes presented in the media, the current context poses particular challenges of navigating a high volume of often conspicuous problematic information, and its impact on specific communities cannot go unnoticed (see especially the section on "Global Youth Cultures" in Frau-Meigs et al., 2020). Marginalized communities are disproportionately impacted by disinformation surrounding divisive issues like immigration. Far right communities are using media in new ways to circulate hate speech through manipulated images, memes, images targeting immigrants and marginalized groups. The impact of Islamophobic disinformation dates back to the aftermath of 9/11 (Zine, 2004), however, sites like YouTube have made the circulation of Islamophobia disinformation online, radicalization, and hate speech more accessible (Lewandowsky and Yesilada, 2021). Media and information literacy should address social

exclusion of marginalized groups such as *trans* folks and refugees, which is perpetuated through sensationalist, emotion-targeting disinformation (Yosso, 2002; Krutkowski et al., 2020; Leurs et al., 2018).

Media literacy programming should consider the unique needs of immigrants who experience language barriers to traditional interventions, have a multilingual media diet, and may have cultural biases to specific information (Samtani, 2019; Katz, 2010). Educators encounter the challenge of accessing local resources that are not Western-centric to teach about problematic information, such as a program in Singapore (Dahri and Richard, 2018).

Adults as target population for disinformation media literacy

Media literacy education has largely been directed toward a youth population within primarily middle and high school settings. There is a critical need for adult-directed media education, especially outside of formal education settings (Roschke, 2018). Adults face unique barriers when adapting to evolving technologies and often require assistance setting up and using their new devices (Anderson and Perrin, 2017). Adults are described as a population that is particularly susceptible to disinformation and scamming due to an overall lack of digital literacy skills (Lee, 2018; Brashier and Schacter, 2020) and are therefore also more likely to share unverified information (Khan and Idris, 2019). Despite the “gray divide” that describes older adults as less involved and skilled with digital media, Quan-Haase et al. (2018) reiterate that they are not disconnected, rather they are users with unique needs to consider.

Checklists, fact-checking & information evaluation frameworks

A widely used information evaluation framework is the CRAAP test (Currency, Relevance, Authority, Accuracy, Purpose), created by the Meriam Library, California State University & Chico (2010). The CRAAP framework is often suggested as a crucial component in media education by librarians and educators around the world. In South Africa, 44 out of 48 of the librarian guides addressing disinformation contain links to evaluation tools such as the CRAAP framework and “How to Spot Fake News” infographic, created by the International Federation of Library Associations (2020) (Bangani, 2021). The same trend exists in the United States of America where all 21 librarian guides evaluated suggest checklist approaches to detecting fake news, mirroring suggestions provided by other American scholars, librarians and educators (Batchelor, 2017; Lim, 2020; Stein-Smith, 2017). The CRAAP checklist is also utilized and promoted in Canada (Hanz and Kingsland, 2020), Pakistan (Naeem and Bhatti, 2020), the UK (Clough and Closier, 2018), and Ireland (De Paor and Heravi, 2020).

There are other suggested adaptations to this approach in response to specific challenges of problematic information. One suggestion is the “CRAAP-based” RADAR framework, which focuses on Rationale, Authority, Date, Accuracy, and Relevance (Neely-Sardon and Tignor, 2018). Another proposed checklist is the IMVA/IN framework (Independent, Multiple, Verify, Authoritative/Informed, Named) created by the News Literacy Project (Dahri and Richard, 2018). After describing the CRAAP test as inefficient, Caulfield suggests the use of SIFT (Stop, Investigate the source, Find trusted coverage, and Trace claims) (Ojala, 2019). Lastly, the CARS checklist is also proposed as a framework to scan for Credibility, Accuracy, Reasonableness, and Support (Jacobson, 2017; Gardner, 2017).

A third group of scholars suggests that these types of information evaluation frameworks may be ineffective and unrealistic in the context of the fast-paced infodemic (McGrew et al., 2017). While scrolling, users often spend little time reading the details of the texts and are unlikely to assess what they are reading against any checklist. Others express concern that the checklist approach “bypasses the need for critical thinking” (Saunders and Budd, 2020) and limits opportunities for open-ended, inductive, and exploratory approaches to understanding a media text (Shenton, 2021).

Teaching lateral reading and process thinking

In contrast to the “checklist” solution to problematic information, many educators and librarians are advocating for the use of “lateral” reading strategies (Caulfield, 2017; McGrew et al., 2017). Lateral reading is the process of searching for information to contextualize or validate a source, “digging deep and wide” (Miller, 2018). As opposed to checklists that tend to focus on the singular text in isolation, lateral reading encourages leaping off the landing page to open new tabs instead of digging into the appearance and content features of a website.

While evaluating information sources, professional fact checkers themselves will almost immediately open up a series of new tabs along the browser’s horizontal axis, drawing on the resources of the Internet to learn more about a site and its claims, according to Wineburg and McGrew (2019). The web of relational information on the internet is used intentionally to guide the assessment of credibility as well as evaluate a websites’ reputation to decide the trustworthiness of the source (Faix, 2018). In addition to lateral reading, educators and librarians suggest that media education teaches research skills; students should engage with the research process while also learning how experts engage in research (see *News literacy*).

Engaging mindfulness techniques

In today’s media environment where entertainment, advertising, news, and memes are all integrated into one platform and endless scrolling and engagement is heavily facilitated by these same platforms, finding ways to encourage people to “stop and think” is paramount to the long-term success of media literacy efforts (Rozendaal et al., 2011)—hence, the need for mindfulness.

Although mindfulness has been referenced by media scholars in the past (Serafin, 2007), it has surfaced as a necessary quality to have within our fast-paced culture and in response to our overflow of emotionally-heightened information. Drawing on Buddhist teachings, some scholars suggest that a first step to healthy news consumption is developing awareness of when, where, and how we are exposed to news as well as attending to how we read and process news and information (Sivek, 2018). The Thai Civic Agency and Thai educators have incorporated principles of mindfulness into their curriculum as well as instruction on how media can influence our emotional states (Kleebugung, 2017).

Users are invited to read with mindfulness so as to not make snap judgments when reading about complex matters while engaged in news media (Berkman, 2021), integrating small pauses, breaths and moments of reflection and mindfulness of emotions, feelings, and thoughts. "Click restraint" is also suggested for social media users to slow down and carefully review search results before clicking (Wineburg and McGrew, 2019). Media literacy programs can encourage slow reading to regulate information flow and help maintain information equilibrium (Tsvetkova, 2017).

Conclusion

Despite media literacy's decades-long history, the evolving nature of the media landscape challenges media education to develop ever more savvy and creative approaches. Media literacy education is a dynamic theoretical field with myriad of practices, and will continue to adapt to the evolving technological and media environments, to develop citizens who are empowered, confident, and skilled in media critique and production that is relevant and applicable to its time.

Questions continue to arise: who is responsible for creating a healthy and productive media ecosystem? Should governments regulate social media platforms or does responsibility lie with the platforms to develop safeguards against misinformation and commercial exploitation? Is it up to the social media user to fact-check before consuming information, and/or should journalists develop fact-checking systems? How much responsibility lies with educators and parents?

Quite often, hopes are pinned on media education as the solution to the infodemic. However, as detailed above, it is not only citizens that affect the media ecosystem; rather this system is best understood through a socioecological lens. Forces at all levels of society—from the individual level to social and technological environment up to the policy, cultural, and ideological levels—play a part in how the media ecosystem is built and operates and how it affects and is affected by citizens. Certainly educators have an important role to play in empowering citizens and providing them with the knowledge and skills necessary to engage critically and effectively with media. But as well, responsibility for the spread and dissemination of disinformation and emotional exploitation lies as well with advertisers, social media platforms, governments and political figures. Appropriate steps are needed at all levels of society to build a healthy and sustainable media environment.

References

- Adorno, T.W., Horkheimer, M., 1944. *Dialectic of Enlightenment*, vol. 15. Verso.
- Alvermann, D.E., Hagood, M.C., 2000. Critical media literacy: research, theory, and practice in "New Times". *J. Educ. Res.* 93 (3), 193–205.
- Anderson, M., Perrin, A., 2017. Technology Use Among Older Adults. PEW Research Center. Available from: <https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/>.
- Ashley, S., 2019. News Literacies. *The International Encyclopedia of Media Literacy*, pp. 1–11.
- Ashley, S., 2020. *News Literacy and Democracy*. Routledge.
- Aufderheide, P., 1992. Cable television and the public interest. *J. Commun.* 42 (1), 52–65.
- Aufderheide, P., 1993. *Media Literacy: A Report of the National Leadership Conference on Media Literacy*. Aspen Institute, Communications and Society Program, Washington, DC.
- Bangani, S., 2021. The fake news wave: academic libraries' battle against misinformation during COVID-19. *J. Acad. Librarian* 47 (5), 102390.
- Batchelor, O., 2017. Getting out the truth: the role of libraries in the fight against fake news. *Ref. Serv. Rev.* 45 (2), 143–148.
- Behrens, S.J., 1994. *A Conceptual Analysis and Historical Overview of Information Literacy*.
- Benjamin, R., 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Social Forces.
- Benkler, Y., Faris, R., Roberts, H., 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.
- Berkman, R., 2021. Mindfulness, nudges, and empathy: new media literacies. In: *Online Searcher*, vol. 45, p. 3.
- Bernays, E.L., 1928/2005. *Propaganda*. IG Publishing.
- Boler, M., 1999. *Feeling power: Emotions and education*. Routledge.
- Boler, M. (Ed.), 2008. *Digital Media and Democracy: Tactics in Hard Times*. MIT Press.
- Boler, M., Davis, E., 2018. The affective politics of the "post-truth" era: feeling rules and networked subjectivity. *Emot. Space Soc.* 27, 75–85.
- Boler, Megan. "Digital Disinformation and the Targeting of Affect: New Frontiers for Critical Media Education." *Research in the Teaching of English* 54.2 (2019): 187-91.
- Boler, M., Davis, E., 2021. *Affective Politics of Digital Media: Propaganda by Other Means*. Routledge.
- Brashier, N.M., Schacter, D.L., 2020. Aging in an era of fake news. *Curr. Dir. Psychol. Sci.* 29 (3), 316–323.
- Brogan, J., February 2016. Your Algorithms Cheat Sheet: Who Are the Leading Thinkers in Algorithms? What Are the Big Ethical Debates? And What's Procedural Generation? Slate. Available from: <https://slate.com/technology/2016/02/algorithms-101-a-cheat-sheet-to-the-terminology-the-ethical-debates-and-more.html>.
- Buckingham, D., 1996. Critical pedagogy and media education: a theory in search of a practice. *J. Curric. Stud.* 28 (6), 627–650.
- Buckingham, D., 2002. *Teaching Popular Culture: Beyond Radical Pedagogy*. Routledge.
- Buckingham, D., 2006. Defining digital literacy—what do young people need to know about digital media? *Nordic J. Dig. Lit.* 1 (4), 263–277.
- Calvert, S.L., 2008. Children as consumers: advertising and marketing. *Future Child.* 18 (1), 205–234.
- Cappello, G., 2017. Literacy, media literacy and social change. Where do we go from now? *Ital. J. Sociol. Educ.* 9 (1), 31–44.
- Catts, R., Lau, J., 2008. *Towards Information Literacy Indicators*. UNESCO, Paris. Available from: <https://www.storre.stir.ac.uk/bitstream/1893/2119/1/cattsandlau.pdf>.
- Caulfield, M., 2017. *Web Literacy for Student Fact-Checkers*. Washington State University.

- Clough, H., Closier, A., 2018. Walking the talk: using digital media to develop distance learners' digital citizenship at the Open University (UK). *Ref. Libr.* 59 (3), 129–133.
- Cohen, J., 2018. Exploring echo-systems: how algorithms shape immersive media environments. *J. Media Lit. Educ.* 10 (2), 139–151.
- Comber, B., Janks, H., Hruby, G.G., 2018. Texts, identities, and ethics: critical literacy in a post-truth world. *J. Adolesc. Adult Lit.* 62 (1), 95–99.
- Conway, M., Scrivens, R., Macnair, L., 2019. Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends. International Centre for Counter-Terrorism.
- Corser, K., Dezuanni, M., Notley, T., 2021. How news media literacy is taught in Australian classrooms. *Aust. Educ. Res.* 1–17.
- Crockett, M.J., 2017. Moral outrage in the digital age. *Nat. Human Behav.* 1 (11), 769–771.
- Dahri, R.B.M., Richard, H.C.Y., 2018. Librarians joining the fight against fake news: a NUS case study. *Singapore J. Library Inf. Manag.* 47, 15–24.
- Damico, J.S., Baildon, M., Panos, A., 2018. Media literacy and climate change in a post-truth society. *J. Media Lit. Educ.* 10 (2), 11–32.
- Damico, J.S., Panos, A., 2018. Civic media literacy as 21st century source work: future social studies teachers examine web sources about climate change. *J. Soc. Stud. Res.* 42 (4), 345–359.
- DataReportal, 2022. Digital 2022 Global Digital Overview. Available from: <https://datareportal.com/reports/digital-2022-global-digital-overview>.
- De Paor, S., Heravi, B., 2020. Information literacy and fake news: how the field of librarianship can help combat the epidemic of fake news. *J. Acad. Librarian* 46 (5), 102218.
- Diallo, O.K., Friberg, N.M., 2021. Subverting the white cis gaze: towards a pedagogy of discomfort, accountability and care in the anthropology classroom. *Teach. Anthropol.* 10 (4), 17–35.
- Dimock, M., Wike, R., 2020. America is Exceptional in the Nature of Its Political Divide. PEW Research Center. Available from: <https://www.pewresearch.org/fact-tank/2020/11/13/america-is-exceptional-in-the-nature-of-its-political-divide/>.
- Domine, V., 2011. The coming of age of media literacy. *J. Media Lit. Educ.* 3.
- Doyle, C.S., 1994. Information Literacy in an Information Society: A Concept for the Information Age. Diane Publishing.
- Durodolu, O.O., Ibenne, S.K., Dube, T.V., 2021. New mental costumes: leveraging ACRL's value of information frame in dealing with infodemics in a post-truth era. *Coll. Res. Libr. News* 82 (4), 175–176.
- Ellison, N.B., Boyd, D., 2013. Sociality through social network sites. In: *The Oxford Handbook of Internet Studies*, pp. 151–172.
- Erstad, O., 2008. Trajectories of remixing: digital literacies, media production, and schooling. In: Lankshear, C., Knobel, M. (Eds.), *Digital Literacies: Concepts, Policies and Practices*.
- Faix, A., 2018. Teaching Online Research in the "Fake News" Era. Association Supporting Computer Users in Education.
- Farkas, J., Schou, J., Neumayer, C., 2018. Cloaked Facebook pages: exploring fake Islamist propaganda in social media. *New Media Soc.* 20 (5), 1850–1867.
- Fedorov, A., 2014. Media education literacy in the world. *Eur. Res.* 67 (1–2), 176–187.
- Flaxman, S., Goel, S., Rao, J.M., 2016. Filter bubbles, echo chambers, and online news consumption. *Publ. Opin. Q.* 80 (S1), 298–320.
- Fleming, J., 2014. Media literacy, news literacy, or news appreciation? A case study of the news literacy program at Stony Brook University. *J. Mass Commun. Educ.* 69 (2), 146–165.
- Flood, A., November 2, 2017. Fake news is "Very Real" world of the year for 2017. *The Guardian*. Available from: <https://www.theguardian.com/books/2017/nov/02/fake-news-is-very-real-word-of-the-year-for-2017>.
- Frau-Meigs, D., Kotilainen, S., Pathak-Shelat, M., Hoechsmann, M., Poyntz, S.R. (Eds.), 2020. *The Handbook of Media Education Research*. John Wiley & Sons.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Routledge, pp. 131–139.
- Freire, P., Macedo, D., 1987. *Literacy: Reading the Word and the World*. Routledge.
- Funk, S., Kellner, D., Share, J., 2016. Critical media literacy as transformative pedagogy. In: Yildiz, M.N., Keengwe, J. (Eds.), *Handbook of Research on Media Literacy in the Digital Age*. IGI Global, Hershey, PA, pp. 1–30.
- Gachago, D., Ivala, E., Condy, J., Chigona, A., 2013. Journeys across difference: pre-service teacher education students' perceptions of a pedagogy of discomfort in a digital storytelling project in South Africa. *Crit. Stud. Teach. Learn.* 1 (1), 22–52.
- Garcia, A., Seglem, R., Share, J., 2013. Transforming teaching and learning through critical media literacy pedagogy. *Learn. Landsc.* 6 (2), 109–124.
- Gardner, L., 2017. Information literacy, a revised plan. *Sch. Libr. J.* 63 (1), 20.
- Gee, J.P., 2003. What video games have to teach us about learning and literacy. *Comput. Entertainment* 1 (1), 20.
- Geiger, R.S., 2018. The lives of bots. In: Lovink, G., Tkacz, N. (Eds.), *Critical Point of View: A Wikipedia Reader*. Institute of Network Cultures, AMS, the Netherlands.
- Gillespie, T., Seaver, N., December 15, 2016. Critical Algorithm Studies: A Reading List. Social Media Collective. Retrieved March 5, 2022, from: <https://socialmediacollective.org/reading-lists/critical-algorithm-studies/>.
- Graves, D.A., Leigh Kelly, L., McArthur, S.A., 2020. Introduction: teaching for critical consciousness at the intersection of critical media literacy and hip hop education. *Int. J. Crit. Media Lit.* 2 (1), 1–8.
- Grünwald Declaration on Media Education, 1983. Challenge of Media Education. Center for Media Literacy. Available from: <https://www.medialit.org/reading-room/challenge-media-education-grunwald-document>.
- Hall, S., 1980. Encoding/decoding. In: Hall, S., Hobson, D., Lowe, A., Willis, P. (Eds.), *Culture, Media, Language*. Hutchinson, London, pp. 128–138.
- Halpern, D.F., 2013. *Thought and Knowledge: An Introduction to Critical Thinking*. Psychology Press.
- Hanz, K., Kingsland, E.S., 2020. Fake or for real? a fake news workshop. *Ref. Serv. Rev.* 48 (1), 91–112.
- Hawkman, A.M., Shear, S.B., 2020. "Who made these rules? We're so confused": an introduction to the special issue on critical race media literacy. *Int. J. Multicult. Educ.* 22 (2), 1–4.
- Hobbs, R., 1997. Expanding the concept of literacy. In: *Media literacy in the information age*, vol. 6, pp. 163–183.
- Hobbs, R., 1999. The seven great debates in the media literacy movement. *J. Commun.* 48 (1), 16–32.
- Hobbs, R., 2008. Debates and challenges facing new literacies in the 21st century. In: Drotner, K., Livingstone, S. (Eds.), *International Handbook of Children, Media and Culture*. SAGE, London, pp. 431–447.
- Hobbs, R., 2017. Teaching and learning in a post-truth world. *Educ. Leader* 75 (3), 26–31.
- Hobbs, R., Donnelly, K., Friesem, J., Moen, M., 2013. Learning to engage: how positive attitudes about the news, media literacy, and video production contribute to adolescent civic engagement. *Educ. Media Int.* 50 (4), 231–246.
- Hobbs, R., Jensen, A., 2009. The past, present, and future of media literacy education. *J. Media Lit. Educ.* 1 (1), 1.
- Hobbs, R., McGee, S., 2014. Teaching about propaganda: an examination of the historical roots of media literacy. *J. Media Lit. Educ.* 6 (2), 56–66.
- Hodgin, E., Kahne, J., 2018. Misinformation in the information age: what teachers can do to support students. *Soc. Educ.* 82 (4), 208–212.
- Hudders, L., De Pauw, P., Cauberghe, V., Panic, K., Zarouali, B., Rozendaal, E., 2017. Shedding new light on how advertising literacy can affect children's processing of embedded advertising formats: a future research agenda. *J. Advert.* 46 (2), 333–349.
- Huesmann, L.R., Eron, L.D., Klein, R., Brice, P., Fischer, P., 1983. Mitigating the imitation of aggressive behaviors by changing children's attitudes about media violence. *J. Pers. Soc. Psychol.* 44 (5), 899.
- Hughes, J.M., Morrison, L., 2014. The impact of social networking and a multiliteracies pedagogy on English language learners' writer identities. *Writ. Pedagog.* 6 (3), 607–631.
- Hwang, Y., Ryu, J.Y., Jeong, S.-H., 2021. Effects of disinformation using deepfake: the protective effect of media literacy education. *Cyberpsychol. Behav. Soc. Netw.* 24 (3), 188–193.
- Jack, C., 2017. Lexicon of lies: terms for problematic information. *Data Soc.* 3 (22), 1094–1096.
- Jacques, P.J., Dunlap, R.E., Freeman, M., 2008. The organisation of denial: conservative think tanks and environmental scepticism. *Environ. Polit.* 17 (3), 349–385.
- Jacobson, L., 2017. The smell test (Cover story). *Sch. Libr. J.* 63 (1), 24–28 (Library, Information Science & Technology Abstracts).

- Jenkins, H., 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. The MIT Press, p. 145.
- Jeong, S.H., Yum, J.Y., Hwang, Y., 2018. Advertising literacy education?: effects of component types and the moderating role of age. *J. Advert.* 47 (4), 347–361.
- Johnson, J., September 2021. Worldwide Digital Population as of January 2021. Statista. Available from: <https://www.statista.com/statistics/617136/digital-population-worldwide/>.
- Jolls, T., Wilson, C., 2014. The core concepts: fundamental to media literacy yesterday, today and tomorrow. *J. Media Lit. Educ.* 6 (2), 68–78.
- Kahne, J., Lee, N.J., Feezell, J.T., 2012. Digital media literacy education and online civic and political participation. *Int. J. Commun.* 6, 24.
- Kajimoto, M., Fleming, J., 2019. News literacy. In: *Oxford Research Encyclopedia of Communication*.
- Katz, V.S., 2010. How children of immigrants use media to connect their families to the Community: the case of Latinos in South Los Angeles. *J. Child. Media* 4 (3), 298–315.
- Kellner, D., Share, J., 2005. Toward critical media literacy: core concepts, debates, organizations, and policy. *Discourse* 26 (3), 369–386.
- Kellner, D., Share, J., 2007a. Critical media literacy: crucial policy choices for a twenty-first century democracy. *Pol. Futures Educ. Internet* 5 (1), 59–69.
- Kellner, D., Share, J., 2007b. Critical media literacy, democracy, and the reconstruction of education. *Media Lit.* 3–23.
- Kennedy, B., Tyson, A., Funk, C., 2022. Americans' Trust in Scientists, Other Groups Declines. PEW Research Center. Available from: <https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/>.
- Khan, M.L., Idris, I.K., 2019. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behav. Inf. Technol.* 38 (12), 1194–1212.
- Kleebung, N., 2017. The Challenge of Media Literacy Education in Thailand. In: *International Handbook of Media Literacy Education*, pp. 43–53.
- Klurfeld, J., Schneider, H., 2014. News Literacy: Teaching the Internet Generation to Make Reliable Information Choices. Brookings Institution Research Paper.
- Krutkowski, S., Taylor-Harman, S., Gupta, K., 2020. De-biasing on university campuses in the age of misinformation. *Ref. Serv. Rev.* 48 (1), 113–138.
- Knobel, M., Lankshear, C., 2008. Digital literacy and participation in online social networking spaces. *Digital Lit.* 11, 249–278.
- Kubey, R., 1998. Obstacles to the development of media education in the United States. *J. Commun.* 48 (1), 58–69.
- Lampert, C., Schulze, A., Dreyer, S., 2021. "This might be advertising": perception, understanding and handling of online advertising by children. *Medienpädagogik* 43, 1–18.
- Lee, N.M., 2018. Fake news, phishing, and fraud: a call for research on digital media literacy education beyond the classroom. *Commun. Educ.* 67 (4), 460–466.
- Leong, C., Pan, S.L., Bahri, S., Fauzi, A., 2019. Social media empowerment in social movements: power activation and power accrual in digital activism. *Eur. J. Inf. Syst.* 28 (2), 173–204.
- Leurs, K., Omerović, E., Bruinenberg, H., Sprenger, S., 2018. Critical media literacy through making media: a key to participation for young migrants? *Communications* 43 (3), 427–450.
- Lewandowsky, S., Yesilada, M., 2021. Inoculating against the spread of islamophobic and radical-islamist disinformation. *Cogn. Res.* 6.
- Lewis, J., Jhally, S., 1998. The struggle over media literacy. *J. Commun.* 48 (1), 109–120.
- Lim, S., 2020. Academic library guides for tackling fake news: a content analysis. *J. Acad. Librariansh.* 46 (5), 1–11.
- Luke, C., 1997. Media literacy and cultural studies. In: *Constructing Critical Literacies: Teaching and Learning Textual Practice*, pp. 19–49.
- Luke, C., 2005. Cyber-schooling and technological change: multiliteracies for new times. In: *Multiliteracies: Literacy and Learning*. Routledge, pp. 77–98.
- MacEntee, K., Mandrona, A., 2015. From discomfort to collaboration: teachers screening cellphims in a rural South African school. *Perspect. Educ.* 33 (4), 42–56.
- Mackey, T.P., Jacobson, T.E., 2011. Reframing information literacy as a metaliteracy. *Coll. Res. Libr.* 72 (1), 62–78.
- Maksl, A., Ashley, S., Craft, S., 2015. Measuring news media literacy. *J. Media Lit. Educ.* 6 (3), 29–45.
- Malmelin, N., 2010. What is advertising literacy? Exploring the dimensions of advertising literacy. *J. Vis. Lit.* 29 (2), 129–142.
- Martin, A., Grudziecki, J., 2006. DigEuLit: concepts and tools for digital literacy development. *Innov. Teach. Learn. Inf. Comput. Sci.* 5 (4), 249–267.
- Marwick, A.E., 2018. Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technol. Rev.* 2 (2), 474–512.
- Marwick, A., Lewis, R., 2017. *Media Manipulation and Disinformation Online*. Data & Society Research Institute, New York, pp. 7–19.
- Masterman, L., 1980. *Teaching About Television*. Macmillan International Higher Education.
- McDougall, J., Berger, R., Fraser, P., Zezulkova, M., 2015. Media literacy, education & (civic) capability: a transferable methodology. *J. Media Lit. Educ.* 7 (1), 4–17.
- McGrew, S., Ortega, T., Breakstone, J., Wineburg, S., 2017. The challenge that's bigger than fake news: civic reasoning in a social media environment. *Am. Educat.* 41 (3), 4–9.
- McKelvey, F., 2014. Algorithmic media need democratic methods: why publics matter. *Can. J. Commun.* 39 (4), 597–613.
- McIntyre, L., 2018. *Post-truth*. MIT Press, London, EN.
- McLuhan, M., 1964. *Understanding Media. A Signet Book*.
- Media Literacy Now, 2021. *U.S. Media Literacy Policy Report 2021*. Media Literacy Now.
- Media Smarts, 2021. *Media Education in Ontario*. Available at: <https://mediasmarts.ca/teacher-resources/digital-media-literacy-outcomes-province-territory/media-education-ontario>.
- Middaugh, E., 2018. Civic media literacy in a transmedia world: balancing personal experience, factual accuracy and emotional appeal as media consumers and circulators. *J. Media Lit. Educ.* 10 (2), 33–52.
- Middaugh, E., 2019. More than just facts: promoting civic media literacy in the era of outrage. *Peabody J. Educ.* 1–15.
- Mihalidis, P., 2019. Civic media literacies. In: *The International Encyclopedia of Media Literacy*, pp. 1–10.
- Miller, A., 2018. THERE'S SO much there!: helping kids conquer the internet & save democracy. *Knowl. Quest.* 47 (1), 24–30. Library & Information Science Abstracts (LISA).
- Morrell, E., Duenas, R., Garcia, V., Lopez, J., 2015. *Critical Media Pedagogy: Teaching for Achievement in City Schools*. Teachers College Press.
- Nadan, Y., Stark, M., 2017. The pedagogy of discomfort: enhancing reflectivity on stereotypes and bias. *Br. J. Soc. Work* 47 (3), 683–700.
- Naeem, S.B., Bhatti, R., 2020. The Covid-19 'infodemic': a new front for information professionals. *Health Info. Libr. J.* 37 (3), 233–239.
- National Association for Media Literacy Education (NAMLE), November 2007. *The Core Principles of Media Literacy Education*. Available from: <https://namle.net/resources/core-principles/>.
- Neely-Sardon, A., Tignor, M., 2018. Focus on the facts: a news and information literacy instructional program. *Ref. Libr.* 59 (3), 108–121.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Nichols, T., 2017. *The Death of Expertise: The Campaign Against Established Knowledge and Why It Matters*. Oxford University Press.
- Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2 (2), 175–220.
- Noble, S.U., 2018. *Algorithms of Oppression*. New York University Press.
- O'Connor, L., 2009. Information literacy as professional legitimization: the quest for professional jurisdiction. *Library Rev.* 58 (4), 272–289.
- Ojala, M., 2019. Mindfulness, media, and misinformation in the digital era. *Inf. Today* 36 (9), 8–9. Library & Information Science Abstracts (LISA).
- O'Reilly, T., 2007. *What is Web 2.0: design patterns and business models for the next generation of software*. Commun. Strat. (1), 17.
- Oreskes, N., 2021. *Why Trust Science?* Princeton University Press.
- Osborne, C.L., 2018. Programming to promote information literacy in the era of fake news. *Int. J. Leg. Inf.* 46 (2), 101–109. Library & Information Science Abstracts (LISA).
- Oxford Dictionaries, (n.d.). *Post-Truth*. Available from: <https://www.oxfordlearnersdictionaries.com/us/definition/english/post-truth>.
- Pangrazio, L., 2016. Reconceptualising critical digital literacy. *Discourse* 37 (2), 163–174.
- Pariser, E., 2011. *The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think*. Penguin.
- PEW Research Center, 2021. *Political Polarization*. Available from: <https://www.pewresearch.org/topic/politics-policy/political-parties-polarization/political-polarization/>.
- Quan-Haase, A., Williams, C., Kicevski, M., Elueze, I., Wellman, B., 2018. Dividing the grey divide: deconstructing myths about older adults' online activities, skills, and attitudes. *Am. Behav. Sci.* 62 (9), 1207–1228.
- Rainie, L., Keeter, S., Perrin, A., 2019. *Trust and Distrust in America*. PEW Research Center. Available from: <https://www.pewresearch.org/politics/2019/07/22/trust-and-distrust-in-america/>.

- Ratto, M., Boler, M. (Eds.), 2014. *DIY Citizenship: Critical Making and Social Media*. MIT Press.
- Reddy, P., Sharma, B., Chaudhary, K., 2020. Digital literacy: a review of literature. *Int. J. Technoethics* 11 (2), 65–94.
- Romero Walker, A., 2021. Using critical media literacy to create a decolonial, anti-racist teaching philosophy. *J. Media Lit. Educ.* 13 (2), 86–93.
- Roschke, K., 2018. Gatekeeping Practices of Participants in a Digital Media Literacy Massive Open Online Course (MOOC). Arizona State University.
- Rowe, M., 1995. First National Media Literacy Conference – Boone, NC – 1995. Center for Media Literacy. Available from: <https://www.medialit.org/reading-room/first-national-media-literacy-conference-boone-nc-1995>.
- Rozendaal, E., Lapierre, M.A., Van Reijmersdal, E.A., Buijzen, M., 2011. Reconsidering advertising literacy as a defense against advertising effects. *Media Psychol.* 14 (4), 333–354.
- Ruggiero, T.E., 2000. Uses and gratifications theory in the 21st century. *Mass Commun. Soc.* 3 (1), 3–37.
- Rushkoff, D., 2010. *Program or Be Programmed: Ten Commands for a Digital Age*. OR Books.
- Sagan, P., Leighton, T., 2010. The Internet & the future of news. *Daedalus* 139 (2), 119–125.
- Sample, A., 2020. Historical development of definitions of information literacy: a literature review of selected resources. *J. Acad. Librarian* 46 (2), 102116.
- Samtani, H., 2019. Lost in misinformation. *Sch. Libr. J.* 65 (10), 32–35 (Library, Information Science & Technology Abstracts).
- Saunders, L., Budd, J., 2020. Examining authority and reclaiming expertise. *J. Acad. Librarian* 46 (1), 1.
- Scharrer, E., Ramasubramanian, S., 2015. Intervening in the Media's Influence on Stereotypes of Race and Ethnicity: The Role of Media Literacy Education.
- Schudson, M., 1995. *The Power of News*. Harvard University Press.
- Serafin, G.M., 2007. Media mindfulness. *Media Liter.* 178–186.
- Shade, L.R., Chan, S., 2020. Digital privacy policy literacy: a framework for Canadian youth. In: Frau-Meigs, K., Pathak-Shelat, M., Hoechsmann, M., Poyntz, S.R. (Eds.), *The Handbook of Media Education Research*, pp. 327–338.
- Shenton, A.K., 2021. Developing inductively-inspired frameworks for evaluating information—Part One. *Sch. Libr.* 69 (1), 8–10.
- Sivek, S.C., 2018. Both facts and feelings: emotion and news literacy. *J. Media Lit. Educ.* 10 (2), 123–138.
- Sorrell, R., 2019. From JSTOR to Jini: incorporating traditional knowledge in teaching information literacy at tribal colleges. *Tribal Coll. J. Am. Indian Higher Educ.* 31 (1).
- Statista Research Department, March 2022. Global Social Networks Ranked by Number of Users 2022. Available from: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
- Stein-Smith, K., 2017. Librarians, information literacy, and fake news. *Strategic Libr.* 37, 1–4.
- Swart, J., 2021. Tactics of News Literacy: How Young People Access, Evaluate, and Engage With News on Social Media. *New Media & Society*.
- Sweeney, E., Lawlor, M.A., Brady, M., 2021. Teenagers' moral advertising literacy in an influencer marketing context. *Int. J. Advert.* 41 (1), 1–24.
- Tandoc Jr., E.C., Lim, Z.W., Ling, R., 2018. Defining “fake news” a typology of scholarly definitions. *Digital Journalism* 6 (2), 137–153.
- Tsvetkova, M., 2017. The speed reading is in disrepute: advantages of slow reading for the information equilibrium. *Eur. J. Contemp. Educ.* 6 (3).
- Tully, M., Maksl, A., Ashley, S., Vraga, E.K., Craft, S., 2021. Defining and Conceptualizing News Literacy. *Journalism*.
- UNESCO, 2011. Media and Information Literacy Curriculum for Teachers. United Nations Educational, Scientific and Cultural Organization. Available from: <http://www.unesco.org>.
- Vraga, E.K., Tully, M., Maksl, A., Craft, S., Ashley, S., 2021. Theorizing news literacy behaviors. *Commun. Theor.* 31 (1), 1–21.
- Walsh, J.P., O'Connor, C., 2019. Social media and policing: a review of recent research. *Sociol. Compass* 13 (1), e12648.
- Westen, D., Blagov, P.S., Harenski, K., Kilts, C., Hamann, S., 2006. Neural bases of motivated reasoning: an fMRI study of emotional constraints on partisan political judgment in the 2004 US presidential election. *J. Cognit. Neurosci.* 18 (11), 1947–1958.
- Wineburg, S., McGrew, S., 2019. Lateral reading and the nature of expertise: reading less and learning more when evaluating digital information. *Teach. Coll. Rec.* 121 (11), 40.
- World Health Organization, 2020. Infodemic management: a key component of the COVID-19 global response. *Wkly. Epidemiol. Rec.* 95 (16), 145–160.
- Yosso, T.J., 2002. Critical race media literacy: challenging deficit discourse about Chicanas/os. *J. Popular Film Televis.* 30 (1), 52–62.
- Zannettou, S., Sirivianos, M., Blackburn, J., Kourtellis, N., 2019. The web of false information: rumors, fake news, hoaxes, clickbait, and various other shenanigans. *J. Data Inf. Qual.* 11 (3), 1–37.
- Zarouali, B., De Pauw, P., Ponnet, K., Walrave, M., Poels, K., Cauberghe, V., Hudders, L., 2019. Considering children's advertising literacy from a methodological point of view: past practices and future recommendations. *J. Curr. Issues Res. Advert.* 40 (2), 196–213.
- Ziewitz, M., 2016. Governing algorithms: myth, mess, and methods. *Sci. Technol. Hum. Val.* 41 (1), 3–16.
- Zine, J., 2004. Anti-islamophobia education as transformative pedagogy: reflections from the educational front lines. *Am. J. Islam Soc.* 21 (3), 110–119.
- Zuboff, S., 2015. Big other: surveillance capitalism and the prospects of an information civilization. *J. Inf. Technol.* 30 (1), 75–89.

Arts-Based Educational Research

Marissa Largo, Visual Art and Art History, York University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	313
What is arts-based educational research?	313
Historical and theoretical foundations and characteristics of ABER	314
Qualities and rigors of ABER	315
Paradigm shifts in ABER	316
ABER and critical race theory	317
ABER and New Materialism	317
Implications of ABER for diversity, democracy, and social justice	318
References	318

Introduction

Within the field of qualitative research, Arts-based educational research (ABER) methodologies provide artists, scientists, and practitioners with interpretive modes for conducting interdisciplinary and artful research. Arts-based educational research adopts esthetic modalities in order to identify inquiry, artfully gather or create data, and formulate theories of pedagogical phenomena in order to inform praxis, curriculum, and/or educational policies. This chapter will begin with a definition of the methodological field, a delineation of the historical and theoretical foundations of ABER as well as its dominant characteristics, followed by a discussion of its qualities and rigors. Next is a paradigm analysis and the shifts that are currently taking place in the field, such as critical methodological interventions in Critical Race Theory and New Materialism. Finally, the significance of ABER for social justice education will be articulated. Existing at the intersections of art, research, and education, ABER adopts esthetic modes of knowledge creation to generate understanding about human experience.

What is arts-based educational research?

Before articulating its imbrications with educational research, an understanding of arts-based research must first be established. *Arts-based research* (ABR) is emergent critical theorizing led by the researcher's artistic practice. As a practice-based form of inquiry, it is carried out through creative materials and techniques that give rise to data of individual or social significance. According to art education scholar James Haywood Rolling, ABR is a "flexible architecture of practice-based theory building methodologies" (1990). ABR has been taken up by a wide variety of disciplines such as social work, healthcare, and more prolifically, public pedagogy and education (Wang et al., 2017). For Rolling, arts-based research becomes arts-based *educational* research when it intersects with concerns around curriculum and educational praxis (2010, p. 104).

For the past four decades in the North American context (dominantly the US and Canada), qualitative researchers in the social sciences have recognized the limits of existing methodologies in capturing all the complexities, affects, and nuances of lived experience. Educational research, which is traditionally embedded within the social sciences, began to embrace artistic approaches to inquiry that blurred the disciplinary boundaries with the arts and humanities. ABER methodologies are simultaneously "practice, process, and product" (Sinner et al., 2006, p. 1225) and is an umbrella term that includes a host of artistic research practices. Some popular and current methods in ABER include, but are not limited to, the literary arts (Barone, 2000), the whole spectrum of visual arts, music, performance (Garoian and Gaudelius, 2004), walking methods (Springgay and Truman, 2018), photographic methods such as Photovoice (Wang and Buriss, 1997), audio arts such as podcasting, and comics. Along with these artistic practices, ABER encompasses a range of methods from qualitative research that may be mobilized at any point throughout the process from the data collection, analysis, and the representation of results (Leavy, 2009). The desired outcomes for ABER are to positively impact educational policy and praxis while enhancing perspectives through esthetic inquiry and challenging the accepted norms of social science inquiry. This is achieved by employing embodied, sensorial, and esthetic ways of knowing and making throughout the entire research process.

Educational methodologist Lather (2006) maps out 4 major epistemological paradigms in qualitative pedagogical research: *positivist*, *interpretivist*, *critical theory*, and *deconstructivist*. Using Lather's delineations, arts-based educational research paradigms may be considered. Firstly, qualitative research within the positivist paradigm presupposes that there is an objective external reality that is knowable. Alternatively, ABER supports multiple perspectives and interpretations or "polyphonous" outcomes that disrupt singular grand narratives (Rolling, 2013) and exceed "positivism's explanatory power" (Lather, 1990). ABER methodologies posit that knowledge and power are relational and are experienced differently based on one's subjectivity. Secondly, ABER is a form of

interpretive research that embraces polyvocality and multiple truths that come to be understood through dialogic discourse. Thirdly, ABER in more recent iterations, adopts critical theoretical architectures, which are intent on the goal of emancipation.

The social goals of critical theory mobilized through ABER methods is to spur imaginative solutions for institutional transformation and positive social change. Finally, *deconstructivist* paradigms in ABER often adopt poststructural theoretical discourses, such as posthumanism, that critique modernism's Enlightenment logics.

Historical and theoretical foundations and characteristics of ABER

Many of the foundational underpinnings of ABER may be traced back to the theories of American educational philosopher John Dewey (1859–1952). Dewey is credited with characterizing learning as an experience. He theorized that humans construct meaning based on an array of sensory experiences, which inevitably informs learning, famously stating that “education is life itself.” Further, Dewey advocated for student-centered approaches in education that connected curriculum with the lived experiences of learners. Dewey's theorizing has had a profound impact on pedagogical theory that challenged dominant banking approaches to education at that time.

Beginning in the 1970s, educational researchers began adopting esthetic forms in their research (Eisner, 1976; Greene, 1975). Largely influenced by Dewey's work on art and pedagogy, Stanford art education scholar Elliot Eisner furthered theory-building and methodological innovations in early arts-based inquiry and education. Eisner (1991) developed *educational criticism*; a methodology informed by art criticism and grounded in the arts and humanities. Applied to social science inquiry, the researcher, or “educational critic” analyzes educational experiences in a similar fashion as the criticism of fine art.

By the 1990s, arts-based forms of educational research were founded. Some of these practices included various forms of writing such as narrative inquiry, autobiography, poetry, fiction, and creative non-fiction. Fine art practices such as dance, creative movement, theater, new media, many of the visual arts disciplines, and music were also fused with longstanding social science research methods. According to Sinner et al. (2006), arts-based research in education involved “processes, forms (or structures), and approaches of creative practices in academic scholarship” (p. 1223). This interdisciplinary approach mobilizes a vast range of artistic methods with the goals of social science research to develop emergent methodologies in qualitative research in education.

Early artistic research largely involved the literary arts. Barone and Eisner (1997) first proposed “arts-based educational research” to describe creative imbrications of the literary arts (short stories, poetry, and theater) and educational research. They further theorized that ABER is characterized by the presence of ambiguity, accessible academic writing, and expressive work that demonstrated a high level of empathy for the participants of the study. In the context of fictional narratives in artistic research, art education scholar Leavy (2016) posits that there are three main evocations: *understanding*, *imagination*, and *empathy*. These evocations can be seen in various other ABER methods, including performance-based approaches. While other forms of educational research may seek certainty and definitive conclusions, ABER instead provides the “enhancements of perspectives” through esthetic form (Barone and Eisner, 2006, p. 96). These perspectives are often polyvocal and support a vast range of interpretations.

Barone and Eisner (2006) propose two qualifying characteristics: first, ABER revolves around “artistic activity” the objective of which is intended to expand understanding of human experience particularly educational phenomena. Secondly, ABER evidences “esthetic qualities or *design elements*” that underlie the research process and its adopted form. These design elements are not limited to the literary arts, but also extend to non-textual artistic forms, such as music, dance, and the visual and performing arts. The scholars claim the prominence of the design elements in the research determines the extent to which the research may qualify as arts-based (p. 95). The design elements are specific to the artistic medium and they speak of an attentiveness to *format*, *language*, *empathetic understanding*, and *virtual realities* (pp. 96–98). Further, they identify three genres of ABER: *narrative construction and storytelling*, *educational connoisseurship and criticism*, and *non-literary forms* (p. 98). The first genre explores the ways in which fiction and non-fiction writing can be a form of educational research. Informed by earlier theorizing mentioned above, educational connoisseurship is honed when a researcher has “refined his or her understanding and perception of a domain that the meanings the individual is able to secure are both complex and subtle” (p. 100). Cahnmann-Taylor and Siegesmund (2008) declare that visuals are not just for representing research findings, but have the potential to be “generative,” or in other words, visuals have become a mode for *creating data* (*italics added*) (p. 99). Visual ABER methods generate data, produce meaningful conclusions, and can represent them in video, performance, dance, film, music, collage, photography, and other forms without privileging written or spoken language. These methods account for the cognitive plurality that exists among educational researchers and audiences.

Eisner (2008 in Cahnmann-Taylor) outlines tensions that emerge from arts-based research in education. ABER yields a multiplicity of interpretations. For researchers who seek to build common conclusions, this can be seen as a limitation as much as a possibility. In the specificity of many ABER case studies, research reproducibility and generalizability may not be achievable. Eisner also identifies arts-based researchers' aspiration for esthetic form may override epistemological concerns, or in other words, materiality and form may take primacy over knowledge creation. Further, artistic inquiry, which can be abstract and esoteric at times, if exists within the social sciences, must be grounded in a legible discourse. Lastly, artistic innovation may be at odds with the imperatives of research to make an impact on society (p. 25). These tensions in the early articulations of ABER point to the emerging field's need to define the parameters of arts-based educational research and its position in the social sciences and its relationship to artistic practice.

In attempts to define the field, scholars have identified and named commonalities and themes among ABER projects. Cahnmann-Taylor (2008) asserts that ABER offers “a healthy synthesis of approaches to collect, analyze, and represent data in ways that paint a full picture of a heterogeneous movement to improve education” (p. 4). She names two predominant strands

in arts-based educational research methodologies, namely (1) *hybrid forms* of artistic and scientific research and (2) *art of scholarship's sake* (p. 8). The former refers to the application of artistic methods to respond to educational inquiry and meld artistic and social sciences tools and approaches. Hybrid forms affirm artistic inquiry as a valid form of research and imbue educational scholarship with expression and artistry. The latter comprises of research activity that resides more within the realm of the fine arts. Practitioners of this theme often have specialized education in a creative discipline along with training in social science. As art for scholarship's sake is preoccupied with the forward-facing aspects of the project, artist researchers apply esthetic forms to the mobilization of the study's findings to affect and connect with publics (2008, p. 10).

Emerging from a group of researchers led by Canadian art education scholar Rita Irwin at the University of British Columbia is another hybrid arts-based approach to educational inquiry called *a/r/tography* (de Cosson and Irwin, 2004). This approach acknowledges that the *a/r/tographer's* simultaneous roles as artist and researcher co-constitute their subjectivity and in turn, inform their analysis and interpretations. *A/r/tographers* engage in artmaking and writing in a way in which the writing is not illustrative of the art, but instead, extends its meanings. While *a/r/tographers* are known to employ standard forms of data collections, such as interviews and participant observation, the production of meaning is not always based on "data." Rather, meaning is also produced in conjunctions with artistic creation and interpretation (Le Blanc and Irwin, 2019). Findings from the inquiry process take on literary, visual, or performative formats and exceed exclusive text-based representation.

Given that education acknowledges multimodal ways of learning, it is fitting that educational research should also embrace capacious methods that have capacity to capture the affective, experiential, and esthetic dimensions of educational phenomena.

Qualities and rigors of ABER

Debates around artistic merit in determining research impact often arise in the field. Cahnmann-Taylor (2008) calls for a "tough critical community" of arts-based educational researchers. She claims that specific training and a culture of critique is required to measure and maintain quality in ABER. In rubrics of quality, she distinguishes the "amateur" from "excellence" (p. 11) and proposes setting up criteria to determine the characteristics of outstanding examples of artistic research in education (p. 12). Similarly, Piirto (2002) calls for extensive preparation of arts-based researchers and recommends at least undergraduate study in ABER and successful peer-review of published artistic scholarship are required for practice.

On the opposite side of the debate are scholars who agree that one does not need to be a professionally trained artist to engage in ABER (McDermott, 2010; Mulvihill and Swaminathan, 2020). Mulvihill and Swaminathan (2020) emphasize that mastery of an artistic discipline is not necessary to practice ABER, but rather, the prerequisites are, "to be interested, bold, and adventurous while being prepared for continuous learning and challenging oneself" (p. 2). It is not unheard of for social scientists to seek out the support of arts-based research specialists to assist them in pushing the esthetic boundaries of their methods.

Rather than a preoccupation with expertise and esthetic quality, art educator and researcher Morna McDermott proposes three questions to assess the success of ABER projects: *What purpose does it serve? Whom does it serve? Who does it serve? Does it affect change?* (2010, p. 7). By focusing on the mobilization of the artistic research and its ability to inspire action and change, researchers may develop an analytic for evaluation.

There is no doubt in the importance of arts-based educational researchers to engage in ongoing reflexive practice and to critically assess the rigor of their methods (Bresler et al., 2010). While some believe that rigor lies in the evaluation of the art produced or in the experience and training of the artist-researchers, other assert that space needs to be made for non-experts, such as youth, who are often the participants of such studies. Smithbell (2010) states that questions of rigor and expertise remain elusive: "there are no definitive answers to vital questions such as these. So, as new knowledge is generated, new and more complex questions develop. It is just this reflexivity, triggered by these new questions, that supports the rigor of arts-based research" (p. 1600).

Critiques of ABER include questions of validity and how it may be complicated by creativity's insistency on novelty, which arises occasionally, at the expense of utility. As artistic research prides itself on innovation, its newness must not eclipse its application in enhancing educational experience. As Eisner has said, if "arts-based research culminates in little more than a delightful poetic passage or a vivid narrative that does little educational work it is not serving its function" (2008, p. 23). The main goal of arts-based research in education is to esthetically illuminate polyvocal interpretations and to understand multifaceted lived experiences to affect diverse audiences and ultimately, instigate positive change in educational policy and practice.

Academic and artistic elitism continue pervade the field. In speaking back to exclusionary opinions, McDermott colorfully explains her anti-elitist belief that researchers and participants do not require specialized art training to engage in ABER:

[E]ven though I am tone deaf doesn't mean I don't sing in the shower. So, in doing arts-based educational research, I have to consider that while dancing in my living room lacks a quality of skill and rigor to constitute "good" arts-based educational research, I equally contend that in order to successfully work within this field I need not be Madonna either (2010, p. 6).

While amusing, this comment points to the democratic potential of ABER to invite researchers and participants alike to tap into their creativity to produce knowledge and culture as the impulse to express, to make beautiful, and to make meaning has long been a cornerstone of human experience—regardless of skill level. Given that art and the academy have long histories of exclusionary

practices and structures, such as constructing master narratives that exalt the dominant culture and writing out marginalized communities from official archives, it is vital for ABER to continually examine its gatekeeping practices.

ABER discourse will often lay claim to its “outsider” status; that due to its embrace of researcher’s subjectivity and its focus on artistic methods and affect, it challenges traditional research’s reproduction of hierarchical social relations. However, ABER’s self-proclamation of academic marginality does not move itself to innocence as it still holds potential for reproducing oppressive patterns. As white supremacy is endemic in our research institutions, so too is it evident in much of the ABER research. Researchers’ privilege has been chronically undertheorized, particularly around race, and it is only recently that whiteness is becoming a central analytic in art education scholarship (Denmead, 2021; McGarry, 2021). For this reason, art education and research has been critiqued for normalizing white privilege (Desai, 2010; Acuff, 2019), for the exaltation of Western art forms (Acuff, 2014, 2019), and its avoidance of difficult questions around sexuality, gender, race, and class (Desai, 2003).

Paradigm shifts in ABER

By conducting a paradigm analysis of arts-based inquiry and its pedagogical implications, American art education scholar Rolling (2010) determines that there are 6 arts-based theoretical models: *poststructural*, *prestructural*, *performative*, *pluralistic*, *proliferative*, and *postparadigmatic* (2010, p. 106). *Poststructuralism* acknowledges that our ideas are situated within discursive fields that require constant reinterpretation of the research subject and the researchers’ subjectivity. *Prestructural* practice refers to lived experience as research that supports phenomenological understandings. Often improvisatory, *performative* practice enacts new knowledge to an audience and mirrors the ways in which education is unscripted and idiosyncratic. *Pluralistic* practice describes the wide array of perspectives and methods that are brought together under the umbrella of arts-based inquiry. *Proliferative practice* describes an expanding praxis-led inquiry that arises from trial-and-error and is a “de/re/construction of research methodology” (p. 108). Finally, *postparadigmatic* theoretical model of arts-based inquiry characterizes a field that is rife with co-existing contrasting paradigms (p. 109) and is a growing field characterized by what methodologist Lather (2006) describes as paradigmatic “profusion” and “epistemological diversity” (p. 36). Rolling’s paradigm analysis provides a useful framework for understanding the theoretical trends of the already-existing and emergent forms of arts-based creation of knowledge.

Since its inception, ABER has sought to transform knowledge and culture, and in recent studies has been explicitly mobilized for social justice outcomes (Keifer-Boyd, 2011). In the last two decades, the field has seen an increased desire to engage with critical theory. This has led to intersections between arts-based research and critical pedagogy among scholars who work for social and political action. Gaining traction during heightened surveillance in post-9/11 society, critical ABER seeks to dismantle oppressive structures through “reflexive critique in the practice of critical research” (Finley, 2017, p. 562). Drawing from critical pedagogues such as Paolo Freire and Henry Giroux, critical arts researchers such as Garoian and Gaudelius (2004) theorize that art production, through the deconstruction of visual codes and messaging, can serve as a form of critique of day-to-day realities and a democratic form of resistance. To further distinguish research aspirations for social transformation, *critical arts-based education research* was coined. According to Finley (2017), critical arts-based educational research, “makes intentional use of imagination. It is a performative research methodology that is structured on the notion of possibility, the *what might be*, of a research tradition that is postcolonial, pluralistic, ethical, and transformative in positive ways” (p. 561). Proponents see their interventions as challenging colonial ways of thinking and being and suggest non-dualist relationalities between “researcher/researched, subjectivity/objectivity, critical and emotional thought to elaborate and create theoretical understandings” (Finley, 2017, p. 562). Armed with the ethical and political commitment to destabilize traditional power dynamics embedded in research and with its motivations to spur social transformation, critical ABER is well suited for engagement with marginalized students, audiences, and communities.

Mulvihill and Swaminathan (2020) suggest that ABER has the potential to “decenter power hierarchies while broadening our ability to accentuate a wider spectrum of human experiences” (p. 2). The need for non-hierarchical power relations is especially necessary when conducting research with marginalized communities. Taking on the poststructural paradigm, educational research with and through the arts has been known to adopt participatory art making methods to consolidate individual and collective goals for empowerment (Wang and Burris, 1997; Largo, 2012). Photovoice, for example, adopts principles and strategies of participatory action research. It blurs the hierarchies between the researcher and the subject by having the community members define the research questions and through photography, they form visual responses to their inquiries. Participants take the lead in articulating possible solutions through the accessible visual method of digital photography. Not only are the participants activated, but so too is the audience. By sharing the results of the photovoice with the public and policymakers, the ultimate goals of community empowerment & social change can be achieved (Wang and Burris, 1997).

Also embracing participatory action research, community-based art education scholar Largo (2012) contends that the researcher, also acting as the artist/educator, must be attentive to the politics of representation when working with marginalized communities. In her study with Filipino youth activists, she determined that there are two ways in which the art educator may engage with a community group: (1) as a facilitator of the collaborative creative process, and/or (2) as a “gatherer” who collects material produced by the community members and consolidates them. Ongoing reflection and feedback from the community members is essential in ensuring that representation occurs on their terms.

ABER and critical race theory

In an increasing degree, the *critical* in critical ABER gestures to radical pedagogical research that seeks to generate liberatory knowledge and culture. In more recent interventions, critical ABER has been infused with the emancipatory ambitions of critical race theory (CRT). According to Ladson-Billings (1998), critical race theory is a theoretical framework that seeks to understand systemic oppressions in society that privilege some at the expense of others. Considering growing political conservatism in the US that seeks to eliminate the study of race and society within education, the fusion of ABER and CRT allows for responsive, justice-driven, and multidisciplinary inquiry that centers resistance politics through the deployment of counternarratives and the critical imagination.

In the wake of violence connected to anti-Black racism in the US and anti-Asian racism exacerbated by the COVID-19 pandemic, approaches to arts-based educational research have merged with CRT and discourses in anti-oppression. US and Canadian education have long been critiqued as overwhelmingly Eurocentric and colonial, despite the growing diversity of these countries and in turn, schools (Acuff, 2019). In public schools and higher education, educators are confronting systemic racism and the hidden curriculum that perpetuates white supremacy. US-based art education scholar Desai (2009) asserts that educators may mobilize the arts in institutional change-making as they “develop a politics of the imagination in schools and universities that strategically uses the power of the image to ‘unframe’ the serious issues that we face today by asking critical questions that envision alternative just futures” (p. 25).

Elucidating pluralistic and performative paradigms, Bagley and Castro-Salazar (2012) fuse ABER with Critical Race Theory (CRT) in their theater-based research with Americans of Mexican descent. They apply performance methods to center the life experiences of those marginalized in formal schooling and in society at large. Their critical approach to ABER has the clear political intent of highlighting the experiences of those who have been disavowed by dominant society because of their race and citizenship. Most importantly, they engage in a purposeful discussion around the educational privileges that are afforded to white society, citing the lower attainment of college education among students of Mexican descent compared to non-racialized Americans. Operating from the CRT framework, Bagley and Castro-Salazar collaborate with practicing artists, musicians, poets, dancers, and photographers who enliven the collected life history data through their respective art forms (p. 240).

Informed by CRT and critical whiteness studies, art educator and teacher educator McGarry (2021) acknowledges her privilege as a white cisgender woman engaged in narrative inquiry around the impacts of whiteness on teaching and learning. Through the creation and performance of a fictionalized script, McGarry problematizes her discomfort around race and racism and critiques her privilege as a white teacher. She challenges the persistent notion of whiteness as neutral in studio practice (as on a mixing palette) and recognizes that the classroom is not a politically neutral site. Through the construction of counternarratives, McGarry rehearses listening to her students of color and working for their liberation.

ABER methods have also been used in teacher education programs for learning about histories of racism in education. Adopting the performative theoretical model, Fierros (2009) integrated performance ethnography as an arts-based instructional, research, and assessment method for preservice teachers in their study of the impact of the 1954 *Brown v. Topeka Board of Education*—a landmark case which declared racial segregation in public schools unconstitutional—on members of their university community. As an experiential approach to engaging with subject matter, student teachers’ theatrical performances were informed by ethnographic data gathered through first-hand interviews with community members. The ensuing performances grappled with matters of privilege, race, and institutional racism. The process was as important as the product; it served as an alternative form of assessment that permitted students to “move beyond the familiar and into a challenging set of authentic experiences that result in a deeper understanding and a greater acceptance of the unfamiliar” (Fierros, 2009). As performance ethnography involves collaboration, interviewing, and the recitation of a script based on lived experience, student teachers achieve an affective, embodied, and multisensory education in equity that allowed them to encounter and empathize with the “unfamiliar.”

This politics of the imagination fuels critical ABER’s interventions in racial justice in education by engaging with the complex intersectional lived experiences of those who have been written out of dominant narratives, by disrupting cultural norms and ideologies, and by communicating findings to publics with accessibility and affective allure. At an increasing rate, methodological innovations in critical ABER are turning to anti-oppressive and decolonial theories and practices to generate new forms of public discourse.

ABER and New Materialism

Proliferative and deconstructivist genres of critical ABER are arising under the framework of “New Materialism.” Hood and Kraehe (2017) argue that Western thinking’s insistency on human-centered research forecloses possibilities in art-based educational research. Firstly, there is a dualism between human life and the non-human, in which non-human things are seen as inactive and humans have agency over them. Secondly, there is an absence of theorizing on the contributions that materials and objects have in the creation of art. New materialist scholars propose that critical ABER account for our “interconnectedness with things—lingering within wonderment, struck by the power of the things that pull us out of our routines” (p. 33). New materialist thought presupposes that all matter has agency and makes space for emergent methods in art education practice as research.

Many artists have long been known to hold a special consideration for the essence of material. New materialist approaches to research are attentive to this ethical impulse and thus, create speculative inquiries based on these considerations. The emergent nature of this kind of research often generates new questions and directions throughout the process and often resists prescriptive

and positivist processes. In considering the more-than-human point-of-view in critical arts-based educational research, new possibilities for inquiry and praxis surface.

Paradigmatic profusion in ABER continues to evolve. This methodological diversity includes emergent forms of research-creation in education such as the integration of participatory esthetics and socially engaged art (LeBlanc, 2019), sustainability art and land-based pedagogies (Hauk and Kippen, 2017), and walking methods (Springgay and Truman, 2018; Patterson, 2021). Demonstrating a postparadigmatic era in which various theoretical models coexist, these new directions are attentive to esthetic provocations generated by space, ethical engagement with ecologies, and non-humancentric perspectives that consider the agency of land, and the politics of movement and embodiment.

Implications of ABER for diversity, democracy, and social justice

In social justice education, new knowledge that emerges from marginalized positionalities decenters dominant perspectives. Within the margins reside transformative potential to reconfigure what we know and how we know it. Arts-based methodologies, as they impact educational inquiry and curriculum, unsettle former ways of conducting research and knowledge production, and thus, provide new frameworks for theory-building and praxis. They capture points of view that may otherwise go unnoticed by more conventional research methods.

When applied to social justice education, arts-based inquiry allows for subjects to engage in affective, experiential, and embodied ways of knowing and making, often revealing truths that cannot be captured through positivist research approaches. As practices of representation, education and educational research have the capacity to replicate existing oppressive power relations or contest them. By circumventing the use of specialized language and modes of data gathering that alienate non-academic audiences, art-based educational research attempts to upend longstanding power imbalances between researcher and subject and between teacher and student. Further, ABER often operates within participatory and collaborative modes, allowing participants to directly engage in defining the research problem, data gathering, and the mobilization of findings to an audience. This democratic approach to research can be considered empowering, especially to historically marginalized communities who are granted a greater amount of agency at every stage of the research process. By fusing scholarly investigations with artistic expression, participants connect personal experiences with broader social concerns, such as inequity and oppression. For these reasons, critical researchers find great value in artful inquiry for the opportunity to bring in those who have been excluded into meaningful educational reform and social transformation.

An undeniable benefit of ABER is its potential for wider knowledge mobilization beyond the academy. It provides an accessible and engaging means of presenting findings to the general public, thus transcending educational and linguistic obstacles that characterize much of mainstream research (Wang and Burris, 1997). For example, in communicating research findings to classroom practitioners and students, esthetic modalities are more likely to garner successful reception than an academic article laden with specialized language. The esthetic quality of the research process and outcomes draws on affect and “awaken” audiences to the unnoticed (Barone, 2002). This in turn, begets empathetic responses that have the potential to inform action and positive social change. The accessibility and allure of many ABER projects make these methodologies enticing for researchers who seek to make meaningful impacts on the communities to whom they are accountable.

References

- Acuff, J.B., 2014. (Re)constructing cultural conceptions and practices in art education: an action research study. *Vis. Arts Res.* 40 (2), 67–78.
- Acuff, J.B., 2019. Whiteness and art education. *J. Cult. Res. Art Educ.* 36 (1 & 2), 8–12.
- Bagley, C., Castro-Salazar, R., 2012. Critical arts-based research in education: performing undocumented histories. *Br. Educ. Res. J.* 38 (2), 239–260. <https://doi.org/10.1080/01411926.2010.538667>.
- Barone, T., 2000. Chapter 8: using the narrative text as an occasion for conspiracy (1990). *Counterpoints* 117, 137–160.
- Barone, T., 2002. Who cares? A play about passion in teaching and in the researching of teaching. In: Mirochnik, E., Sherman, D.Y. (Eds.), *Passion and Pedagogy: Relation, Creation, and Transformation in Teaching*. Peter Lang, New York, pp. 75–94.
- Barone, T., Eisner, E.W., 1997. Art-based educational research. In: Jaeger, R.M. (Ed.), *Complementary methods for research in education*. American Educational Research Association Press, Washington, DC., pp. 73–116. Submitted for publication.
- Barone, T., Eisner, E., 2006. Arts-based educational research. In: Green, J., Camilli, G., Elmore, P. (Eds.), *Complementary methods in research in education*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 95–109.
- Bresler, L., Cahnmann-Taylor, M., Erickson, F., Garoian, C., Ho Sik Ying, P., Siegesmund, R., 2010. On Rigor in Arts-Based Research. Presented at the Congress of Qualitative Inquiry. University of Urbana-Champaign, Illinois.
- Cahnmann-Taylor, M., 2008. Arts-based research: histories and new directions. In: Cahnmann-Taylor, M., Siegesmund, R. (Eds.), *Arts-Based Research in Education: Foundations for Practice*. Routledge, New York, NY, pp. 3–15.
- Cahnmann-Taylor, M., Siegesmund, R., 2008a. Arts-Based Research in Education: Foundations for Practice. Routledge, New York, NY.
- Cahnmann-Taylor, M., Siegesmund, R., 2008b. The tensions of arts-based research in education reconsidered: the promise for practice. In: Cahnmann-Taylor, M., Siegesmund, R. (Eds.), *Arts-Based Research in Education: Foundations for Practice*. Routledge, New York, NY, pp. 231–246.
- De Cosson, A., Irwin, R.L., 2004. *A/r/tography: Rendering Self Through Arts-Based Living Inquiry*. Pacific Educational Press, Vancouver.
- Denmead, T., 2021. Time after whiteness: performative pedagogy and temporal subjectivities in art education. *Stud. Art Educ.* 62, 130–141. <https://doi.org/10.1080/00393541.2021.1896252>.
- Desai, D., 2003. Multicultural art education and the heterosexual imagination: a question of culture. *Stud. Art Educ.* 44, 147–161. <https://doi.org/10.1080/00393541.2003.11651735>.

- Desai, D., 2009. Imagining justice in times of perpetual war: notes for the classroom. *J. Curric. Pedagogy* 6, 6–26. <https://doi.org/10.1080/15505170.2009.10411734>.
- Desai, D., 2010. The challenge of new colorblind racism in art education. *Art Educ.* 63, 22–28.
- Eisner, E. (Ed.), 1976. *The Arts, Human Development, and Education*. McCutchan, Berkeley, CA.
- Eisner, E.W., 1991. *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*. Macmillan, New York, NY.
- Eisner, E., 2008. Persistent tensions in arts-based research. In: Cahnmann-Taylor, M., Siegesmund, R. (Eds.), *Arts-Based Research in Education: Foundations for Practice*. Routledge, New York, NY, pp. 16–27.
- Fierros, E.G., 2009. Using performance ethnography to confront issues of privilege, race, and institutional racism: an account of an arts-based teacher education project. *Multicult. Perspect.* 11 (1), 3–11. <https://doi.org/10.1080/15210960902743896>.
- Finley, S., 2017. Critical arts-based inquiry: performances of resistance politics. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*. SAGE Publications.
- Garoian, C.R., Gaudelius, Y.M., 2004. Performing resistance. *Stud. Art Educ.* 46, 48–60. <https://doi.org/10.1080/00393541.2004.11650068>.
- Greene, M., 1975. Curriculum and consciousness. In: Pinar, W. (Ed.), *Curriculum Theorizing: The Reconceptualists*. McCutchan, Berkeley, CA, pp. 295–298.
- Hauk, M., Kippen, A.R., 2017. Just sustainability arts: a vibrant convergence. *Interdiscipl. Environ. Rev.* 18, 3–4.
- Hood, E.J., Kraehe, A.M., 2017. Creative matter: new materialism in art education research. *Teach. Learn. Art Educ.* 70, 32–38. <https://doi.org/10.1080/00043125.2017.1274196>.
- Keifer-Boyd, K., 2011. Arts-based research as social justice activism: insight, inquiry, imagination, embodiment, relationality. *Int. Rev. Qual. Res.* 5 (1), 3–19.
- Ladson-Billings, G., 1998. Just what is critical race theory and what's it doing in a nice field like education? *Int. J. Qual. Stud. Educ.* 11, 7–24. <https://doi.org/10.1080/095183998236863>.
- Largo, M., 2012. From the pearl of the orient to uptown: a collaborative arts-based inquiry with Filipino youth activists in Montreal. In: *Filipinos in Canada*. University of Toronto Press, Toronto, pp. 243–264. <https://doi.org/10.3138/9781442662728-017>.
- Lather, P., 1990. Reinscribing otherwise: the play of values in the practices of the human sciences. In: Guba, E. (Ed.), *The Paradigm, Dialog*. Sage Publications, Inc., Newbury Park, CA, pp. 315–332.
- Lather, P., 2006. Paradigm proliferation as a good thing to think with: teaching research in education as a wild profusion. *Int. J. Qual. Stud. Educ.* 19 (1), 35–57.
- Leavy, P., 2009. *Method Meets Art: Arts-Based Research Practice*. Guilford Press, New York, NY.
- Leavy, P., 2016. *Fiction as Research Practice: Short Stories, Novellas, and Novels*. Routledge, New York. <https://doi.org/10.4324/9781315428499>.
- LeBlanc, N., 2019. The “dream scene”: exploring participatory socially engaged art practices through arts-based educational research. *Int. J. Art Des. Educ.* 38 (3), 562–571. <https://doi.org/10.1111/jade.12244>.
- LeBlanc, N., Irwin, R., 2019. A/r/tography. *Oxford Research Encyclopedia of Education*. <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-393>. (Accessed 17 November 2021).
- McDermott, M., 2010. Outlaw arts-based educational research. *J. Curric. Pedagogy* 7 (1), 6–14. Available at. http://resolver.scholarsportal.info/resolve/15505170/v07i0001/6_oaer.xml.
- McGarry, K., 2021. A white palette gathering: a literature synthesis in one act. *Art Res. Int.* 6, 2021. <https://doi.org/10.18432/ari29555>.
- Mulvihill, T.M., Swaminathan, R., 2020. *Arts-Based Educational Research and Qualitative Inquiry: Walking the Path*, first ed. Routledge, Milton. <https://doi.org/10.4324/9781315143361>.
- Patterson, J.A., 2021. Walking ABER: of pedagogy, practice, and potentiality. *J. Curric. Pedagogy* 18, 83–95. <https://doi.org/10.1080/15505170.2020.1832001>.
- Piirto, J., 2002. The question of quality and qualifications: writing inferior poems as qualitative research. *Int. J. Qual. Stud. Educ.* 15, 431–445. <https://doi.org/10.1080/09518390210145507>.
- Rolling, J., 2010. A paradigm analysis of arts-based research and implications for education. *Stud. Art Educ.* 51 (2), 102–114.
- Rolling, J.H., 2013. *Arts-Based Research Primer*. Peter Lang Publishing, New York.
- Sinner, A., Leggo, C., Irwin, R.L., Gouzouasis, P., Grauer, K., 2006. Arts-based educational research dissertations: reviewing the practices of new scholars. *Can. J. Educ.* 29, 1223–1234. <https://doi.org/10.2307/20054216>.
- Smithbell, P., 2010. Arts-based research in education: a review. *Qual. Rep.* 15 (6), 1597+. Available at. https://link.gale.com/apps/doc/A242016921/AONE?u=yorku_main&sid=bookmark-AONE&xid=6ba9ea60. (Accessed 2 November 2021).
- Springgay, S., Truman, S.E., 2018. *Walking Methodologies in a More-Than-Human World*. Routledge.
- Wang, C., Burris, M.A., 1997. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ. Behav.* 24, 369–387. <https://doi.org/10.1177/109019819702400309>.
- Wang, Q., Coemans, S., Siegesmund, R., Hannes, K., 2017. Arts-based methods in socially engaged research practice: a classification framework. *Art Res. Int.* 2 (2), 5–39.

Museum education: building the capacity of society and its members

Sandra Annette Toro, United States Department of Education, Office of Elementary and Secondary Education (OESE), Washington, DC, United States

© 2023 Elsevier Ltd. All rights reserved.

Museum education and the idea of building capacity	320
Before the pandemic: museums as learning institutions helping to build capacity	320
During the pandemic: challenges and responses	321
Moving forward: convergence	321
References	322
Relevant websites	322

Museum education and the idea of building capacity

Museums are the learning institutions of democratic societies, fostering the development and growth of people of all ages and introducing learners to the scientific, historical, art, living, and other collections and resources of society. Museums, including art and science museums, historic houses, aquariums, zoos, and botanical gardens, provide physical and virtual spaces for people to think and act individually and collectively. It is in museums where families, peer groups and friends, school groups, and individuals can choose what, when, and how they learn as well as gather for leisure-time activities and events (Falk and Dierking, 2012). Pre-COVID-19, in the United States alone, 900,000,000 people visited museums annually while museums invested \$2 billion on education activities each year (AAM, 2021).

This essay examines how museums have and continue to “build capacity” among both museum visitors and staff before, during, and after the COVID-19 pandemic, building upon John Dewey’s view of capacity as “an ability, a power; and by potentiality potency, force,” (Dewey, 1916, p. 41) and “a force positively present—the ability to develop” (p. 42). For Dewey, the older and more experienced and knowledgeable members of a society determine how younger or more immature citizens should develop and grow in a positive sense, so that those more inexperienced members can eventually and collectively help society, itself develop further.

Before the pandemic: museums as learning institutions helping to build capacity

The emergence of museums in society is linked to goals of education and access (Protas, 2019), whether we understand museum education as beginning with the 1781 opening of the Belvedere Museum in Vienna or the educational work and outreach of John Cotton Dana and Louise Connelly in Newark, NJ, in the early 1900s (Mulryan, 2001). Museums play a key role in helping a society figure out how to facilitate learning by creating spaces and programs that fit the needs of people in surrounding communities and building museum visitors’ capacities for thinking, reasoning, doing or making, and being overall productive members of society through exhibitions, educational programs, and experiences designed to facilitate and foster the acquisition of a range of knowledge and skills.

Prior to March 2020 and the spread of the infectious disease COVID-19, visitors had come to understand that museums could be used to support their own curiosity and interest; support the learning needs of others (like children); build personal or professional identities; refresh and rejuvenate themselves; and explore and do novel things (Falk and Dierking, 2019). This understanding on the part of visitors reflects that educators in museums had, in many cases, moved away from a transmission model of education to one that incorporated theory and research from different fields of study (Zimmerman and Martell, 2008).

Scholars like Weil (1995) and Martin and Toon (2010) had called for museum staff to view visitors not as empty vessels to be filled with knowledge but as equal contributors to meaning making in museum spaces, particularly through the use of narrative and storytelling. In fact, museum staff came to use a range of theoretical and conceptual approaches to teach visitors, including social constructivism, identity formation, interest, social learning theory, and experiential learning, among others (Hohenstein and Mousouri, 2017).

Much research in science museums examining varied approaches to teaching found evidence supporting, among other aspects of science learning: participating in scientific activities and learning practices with others and using scientific language and tools; and thinking about oneself as a science learner and developing an identity as someone who knows about, uses, and sometimes contributes to science (NRC, 2009, p. 4). Meanwhile, though some educators and other museum professionals saw themselves as neutral guardians of cultural heritage (Audunson et al., 2019), other museum staff hoped to push past “a preoccupation with education, entertainment and consumption,” to bring about important changes in society (Janes and Sandell, 2019, p. 1).

As a profession, people who worked and taught in museums prior to the pandemic were encouraged to face their unconscious biases; not hinder progress by debating definitions; understand inclusion as central to the effectiveness and sustainability of

museums; see systemic change as vital to long-term, genuine progress; and view empowered, inclusive leadership as essential at all levels of museum operations (AAM, 2018). In short, both museum visitors and workers in museums built capacity, including intellectual, social, and for the general benefit and well-being of society.

In 2019, the United States Federal agency, the Institute of Museum and Library Services (IMLS), and the non-profit organization, Partners for Public Good (PPG), conducted a mixed-methods research study focused on this idea of building capacity on the part of the museum as an institution. IMLS and PPG defined museum capacity as “the generation of resources or support intended to help an institution enhance its ability to fulfill its mission or purpose (i.e., any activity or support that is focused on the health and sustainability of the museum rather than on specific exhibitions or programs,” (Hausmann and Hackney, 2021, p. 7). And, though museum missions are as diverse as their collections, study results included that, for all museums, capacity building must link directly to audience responsiveness; is driven by those with power such as funders and museum board members; and is complex and requires a holistic change management approach and commitment from leadership and staff. Researchers also found that museums experience different barriers to participating in capacity building due to such factors as funding and staff time; a perception of museum uniqueness may be a barrier to the application of capacity building best practices from other sectors; and infrastructure, including availability of resources, within the museum sector is insufficient to build museums’ adaptive and relational capacities and address diversity, equity, accessibility, and inclusion (DEAI) issues (Hausmann and Hackney, 2021).

When the COVID-19 situation occurred, however, research and capacity building efforts were stalled or halted as museums faced challenges, including how to remain operational, keep staff employed, and decide whether to suspend all educational programming or make it remote (Greenberg, 2021). Museum professionals had to ask themselves, “What does it mean to be a museum or a library when no one is coming through the doors?” (IMLS, 2021, p. 6).

During the pandemic: challenges and responses

During the pandemic, challenges to museum education in the United States included state governments forbidding in-person field trips for school groups (Young, 2021) and mass layoffs of museum educators (Fonda, 2021.) Museum educators who remained employed were forced to contend with the “digital divide” and reexamine how existing technology platforms and software could help all members of a community; how preexisting museum education program models could be applied to digital settings; to consider forming digital access partnerships with other organizations to support museum content delivery; and to develop educational programs based on the specific needs of community groups and families, particularly those with limited access to technology (Bailey and Sautner, 2021). Some museums repurposed and adapted distance learning workshops and other materials developed prior to the pandemic (Eyl and Stirn, 2021) while other museums helped educate the public about pandemics and the history of vaccine distributions (Øverland, 2021) or used arts and culture-based approaches to increase COVID-19 vaccine confidence (University of Florida, 2021). Innovative programming paradigms designed to address the new, pandemic-based capacity building needs of visitors emerged alongside new museum workplace practices that adhered to local, state, and federal mask and quarantine mandates and related limitations such as timed ticketing, reduced hours, fewer staff, and the use of hoteling as opposed to assigned workspaces for museum professionals (Merritt, 2021).

One key area of exploration for museum workers, especially in the United States, catalyzed in response to tragic events, including the murder of George Floyd—an unarmed Black man killed by a Minnesota police officer during an arrest—and coincided with the emergence of the pandemic, shining a light on issues of race and racial dynamics. Scholars called for the use of theory, including critical race theory, to inform understanding of museums’ rootedness in colonialism and imperialism (Moore, 2021). Theory could help shape daily practices in museums, including “creating spaces, opportunities, advisory capacities, and more to elevate the presence, power, and voices of historically marginalized communities” to share authority and ways of knowing (Moore, 2021). Museum professionals revisited past efforts addressing structural racism in museums and called for museums to listen to the experiences, data, and insights of people of color; and create a new paradigm addressing inclusivity of staffing, collections, and interpretation as well as equality of access to resources. A goal for the professional would be to develop a philosophically-based action plan to be morally and culturally responsive while empathic, respectful, and inclusive so as to transform institutional policies in positive ways (Museums & Race: Transformation and Justice, 2021). Civic-minded museum staff could therefore push themselves and visitors to make society better, shunning pervasive elitism, and including all voices.

Moving forward: convergence

As each country works to get out from under the pandemic situation and away from the impacts and effects of COVID-19, dialogs around equity, inclusivity, accountability, allyship, and the well-being of all of society’s members (Falk, 2021) are driving new approaches to capacity building in museums (See list of Relevant Websites for examples). These discussions align with Dewey’s thoughts about capacity development as strength-rather than deficit-based, because movement is being made to recognize and embrace all people as equal contributors, as both meaning makers and actors with the power to grow individually and as a collective whole. Museums provide safe if not always neutral spaces where people can seek answers to questions not only about specific disciplines like science and history but engage in conceptual work about what it means to be a productive, contributing member of society. The museum field is advancing toward a new understanding that agency on the part of individuals and institutions alike

transcends disciplinary boundaries as well as the limitations of physical spaces. Finding solutions, for example, to the challenges of climate change, racism, and inequality and oppression will happen through a convergence of museums, museum visitors, and museum professionals that is epistemological and ontological.

References

- American Alliance of Museums, 2018. Facing Change: Insights From the American Alliance of Museums' Diversity, Equity, Accessibility, and Inclusion Working Group. AAM, Washington.
- American Alliance of Museums, 2021. Museum Facts & Data. Accessed online at. <https://www.aam-us.org/programs/about-museums/museum-facts-data/>.
- Audunson, R.A., Hobohm, H., Tóth, M., 2019. ALM in the public sphere: how do archivists, librarians and museum professionals conceive the respective roles of their institutions in the public sphere? *Inf. Res.* 24 (4).
- Bailey, E., Sautner, K., 2021. How Can Education Programs Help Bridge the Digital Divide? Accessed online at. <https://www.aam-us.org/2021/11/19/how-can-education-programs-help-bridge-the-digital-divide/>.
- Dewey, J., 1916. *Democracy and Education*. The Free Press, New York.
- Eyl, J., Stirn, L., 2021. Teaching beyond the apocalypse: taking virtual learning from pivot to priority. In: Young, T. (Ed.), *Creating Meaningful Museum Experiences for K-12 Audiences*. Rowman & Littlefield, Lanham, pp. 201–208.
- Falk, J., 2021. *The Value of Museums: Enhancing Societal Well-Being*. Rowman & Littlefield Publishers, Lanham.
- Falk, J.H., Dierking, L.D., 2012. *The Museum Experience Revisited*. Routledge, Walnut Creek.
- Falk, J.H., Dierking, L.D., 2019. Reimagining public science education: the role of lifelong free-choice learning. *Discip. Interdiscip. Sci. Educ. Res.* 1, 10. <https://doi.org/10.1186/s43031-019-0013-x>.
- Fonda, S., 2021. Building the museum educators of the future: a new model for staffing school programs. In: Young, T. (Ed.), *Creating Meaningful Museum Experiences for K-12 Audiences*. Rowman & Littlefield, Lanham, pp. 91–100.
- Greenberg, E., 2021. A challenging proposition: the financial realities of museum education. In: Young, T. (Ed.), *Creating Meaningful Museum Experiences for K-12 Audiences*. Rowman & Littlefield, Lanham, pp. 29–38.
- Hausmann, T., Hackney, S., 2021. Market Analysis and Opportunity Assessment of Museum Capacity Building Programs.
- Hohenstein, J., Moussouri, T., 2017. *Museum Learning: Theory and Research as Tools for Enhancing Practice*. Routledge.
- Institute of Museum and Library Services, 2019. *Understanding the Social Wellbeing Impacts of the Nation's Libraries and Museums: 2021 Report*. IMLS, Washington.
- Janes, R.R., Sandell, R., 2019. *Museum Activism*. Routledge, New York.
- Martin, L., Toon, R., 2010. Narratives in a science center: interpretation and identity. *Curator* 48 (4), 407–425.
- Merritt, E., 2021. Pandemic Disruption as an Opportunity for Change: Intuit. Accessed online at. <https://www.aam-us.org/2021/08/17/pandemic-disruption-as-an-opportunity-for-change-intuit/>.
- Moore, P., 2021. A Liberatory Framework: Critical Race Theory Can Help Museums Commit to Anti-racism and Combat Anti-blackness. Accessed online at. https://www.aam-us.org/2021/01/01/a-liberatory-framework-critical-race-theory-can-help-museums-commit-to-anti-racism-and-combat-anti-blackness/?gclid=EAlaIqobChMI36OyhuC19AIVUNyGCh1IVgYUEAAYASAAEgJpl_D_BwE.
- Mulryan, M.J., 2001. John Cotton Dana, John Dewey, and the Creators of the Newark Museum: A Collaborative Success in the Art of Progressive, Visual Instruction. Seton Hall University Dissertations and Theses (ETDs). <https://scholarship.shu.edu/dissertations/2414>.
- Museums & Race: Transformation and Justice, 2021. Accessed online at. <https://museumsandrace.org/>.
- National Research Council, 2009. *Learning Science in Informal Environments*. The National Academies Press, Washington.
- Øverland, C., 2021. All Hands on Deck: How Museum Staff Boosted Wisconsin's Pandemic Response. Accessed online at. <https://www.aam-us.org/2021/09/08/all-hands-on-deck-how-museum-staff-boosted-wisconsins-pandemic-response/>.
- Prottas, N., 2019. Where does the history of museum education begin? *J. Mus. Educ.* 44 (4), 337–341. <https://doi.org/10.1080/10598650.2019.1677020>.
- University of Florida, 2021. Vaccine Confidence Arts Response Repository. Accessed online at. <https://www.vaccinate.arts.ufl.edu/vaccine-confidence-repository>.
- Weil, S., 1995. *A Cabinet of Curiosity: Inquiries Into Museums and Their Prospects*. Smithsonian Institution Press, Washington.
- Young, T., 2021. Unpacking the field trip. In: Young, T. (Ed.), *Creating Meaningful Museum Experiences for K-12 Audiences*. Rowman & Littlefield, Lanham, pp. 7–19.
- Zimmerman, H., Martell, S.T., 2008. From the guest editors. *J. Mus. Educ.* 33 (2), 101–106.

Relevant websites

- Conversations With Crosby: Smithsonian Secretary Lonnie Bunch on American Heritage, Culture, and the Humanities: MLS Director Crosby Kemper and Lonnie G. Bunch III, the 14th Secretary of the Smithsonian Institution, Discuss the Depth and Diversity of American Stories, What Being a Custodian of Cultural Heritage Means Today, and How Museums and Libraries Can Help Strengthen the Fabric of Our Communities at a Critical Point in the Nation's History. <https://www.imls.gov/events/conversations-crosby-smithsonian-secretary-lonnie-bunch-american-heritage-culture>.
- Museum Expert is a Reimagining of How Museums and Experts Can Work Together in New Ways. <https://www.museumexpert.org/>.
- Museum Hue is the Leading Organization Dedicated to Advancing Black, Indigenous, and Other People of Color in the Cultural Field With Over 400 Institutional Members, Representing Cultural and Academic Institutions Across the United States. <https://www.museumhue.com/>.

Women's informal learning in poor communities, social movements and virtual spaces

Salma Ismail, Faculty of Humanities, School of Education, University of Cape Town, Cape Town, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	323
Conceptual framework	324
Power in the learning process	324
Learning in community development	325
Learning in social movements	326
Learning in virtual communities	327
Conclusion	328
References	328
Relevant websites	329

Introduction

The concern and context of this chapter is women's learning in poor communities in an informal context and poses questions for social transformation. The purpose of this chapter is to explore the interconnectedness between learning and social change in different sites of informal learning and reviews women's learning in community development, social movements and virtual spaces.

The conceptual and analytical framework of this discussion is located within a theoretical paradigm of learning which is multi-disciplinary and includes adult education, development studies, feminist and popular pedagogy and draws on sociology and psychology. In this conceptual framework, women are not characterized as a universal group, they have competing ideologies which are shaped by socio-economic forces, history, context, and agency which influence their learning and life chances which are the opportunities to increase one's position in the social class structure.

Formal education is defined as long-term, planned, certified, takes place in a formal institution, is hierarchically structured and involves the transmission of knowledge which is chronologically graded. It includes education from primary schools through to university and includes professional training programs. On the other hand, informal learning is usually unplanned, short-term and is not certified.

The context of this discussion is situated within poor communities. I draw on various international examples to illustrate (1) women's learning in communities, (2) in social movements and (3) virtually. However, poor women often experience oppression resulting from patriarchy which may affect their ability to access formal learning. Patriarchy (social relationships in which men hold the power) takes different forms but usually includes gender-based violence (GBV), such as sexual, physical or emotional abuse. This was illustrated by the formation of the #MeToo movement when women across race, gender, class, and ethnic divides stood up, made themselves visible and added their voices against sexual exploitation.

Rather than education, which is associated with formal institutions, women in poor communities receive informal learning which is defined as learning in the everyday, in the family and while engaging in activities such as community development and social movements. Informal learning is not accredited and institutionalized therefore it is difficult to measure as there are no set tangible outcomes. Women receive tacit knowledge through socialization and life experiences as well as from others in communities of learning (Foley, 1999; Lave and Wenger, 1991). Brookfield (2001: 149), drawing on Malcolm Knowles (1975), maintains that informal learning is self-directed learning that often takes place through groups of adults united by a common purpose who exchange information, ideas, skills, and knowledge with other members in the group. This knowledge sharing often includes problem solving and the creation of new modes of practice or new forms of knowledge. Walters (1998: 436) defines informal learning as "often unplanned, incidental learning which occurs while organizing community soup kitchens through savings clubs and religious organizations".

In many countries, poor women have been excluded from formal education. The reasons for their exclusion include apartheid legislation, as was the case in South Africa; a lack of money; difficulties accessing educational facilities; and social/cultural or religious practices and prejudices which dictate that women must stay at home to care for the family (Ismail, 2006:12). For example, in 2021, after the United States of America's (USA) withdrawal from Afghanistan, the Taliban closed many schools for girls (<https://www.usip.org/publications/2022/04/talibans-ban-girls-education-Afghanistan>).

When formal education is lacking, women learn in informal ways that are geared to their everyday realities and support them in their daily struggles. In many countries, informal learning has developed so that excluded groups have opportunities to learn. Non-formal programs funded by the World Bank or UNESCO provide a systematic transmission of knowledge and skills which add to women's informal learning (<https://www.worldbank.org/en/topic/gender/overview#2>; <https://en.unesco.org/themes/education-and-gender-equality>).

In some cases, the adult learning programs use popular and feminist pedagogies which have led to social activism and political education (Walters, 1989).

Many women have proven that they can learn and perform equally in all major scientific disciplines. However, the patriarchal belief in women's cognitive inferiority persists as well as the prejudice that women are unable to rise above the practical and emotional preoccupations of everyday life and are characterized as irrational, passionate and sensuous (Pateman, 1990). On the contrary, radical popular education and feminist pedagogy have had an important impact on these prejudices by examining the political identities of women, focusing on questions of transformation, the role of the collective in mobilizing for action and the link between learning and changes in political consciousness. In the next sections I discuss the conceptual framework and women's learning in informal sites.

Conceptual framework

Popular education and feminist pedagogy draw on Paulo Freire's (1983) theory of challenging dominant ideologies of learning. Underlying both theories are certain assumptions about experiences as the basis of knowledge, oppression, consciousness raising, and how education can lead to women's empowerment.

Popular education is a term used in informal contexts to define educational practices which aim to challenge injustices and oppression. A common understanding of popular education is that its main objective is to enable participants to understand the contexts in which they live and act (Walters and Manicom, 1996; Weiler, 1991). The curriculum is drawn from the concerns and interests of the participants, and the critical moment is when participants can name the issues and analyze the forces within themselves and other dominant forces that prohibit progress (Arnold et al., 1991:25–29). Then a plan of action is developed based on a social analysis of the context and political forces (Walters, 1998). Feminist pedagogy challenges the male bias in popular education with a focus on challenging women's subordinate position (Walters, 1998). It includes issues of identity and culture and has theorized the different meanings and forms of power.

These pedagogies focus on the experiences of the learners in all considerations of teaching and learning. Learners analyze their experiences by reflecting, evaluating, and reconstructing them in order to create knowledge and make meaning. There is a strong organic connection between education (gaining new knowledge) and personal experience.

There is a dialectic between action and reflection. In the process of action and reflection, people's attitudes and values change, and their experiences are used to raise awareness and lead to a change in consciousness and the material position of women (Jarvis, 1987; Weiler, 1991). The methodologies used in popular and feminist pedagogy include creating a democratic space to offset power between the educator and participants and participation through group discussions and workshops where the content considers the needs and interests of the participants (Ismail, 2006:29).

Popular and feminist education is not neutral but political therefore there are debates about women's empowerment and conscientization which lead women to examine power relationships, to take charge and show agency in dismantling oppressive structures. However, sometimes women may choose to act strategically and continue with the norm. The feminist contribution is to make the personal political and to challenge patriarchal social relations (Walters and Manicom, 1996). The next section discusses how issues of power are addressed in the learning process.

Power in the learning process

The way in which popular and feminist educators have managed to address power relationships between themselves and learners is by sharing their knowledge with learners who apply it to their situation in a process which does not deny inequality of knowledge but which is based on co-operative and democratic principles of power. This they argue (Rowlands, 1998; Rogers, 1992; Kaplan, 1997) can be achieved if the attitudes of the educator include complete respect for each individual and for the group, humility, a willingness and eagerness for learning to be mutual and commitment to the empowerment process. Jo Rowlands (1998) argues that the dominant understanding of power within the social sciences has been that of "power of" or "power over", whereas a feminist understanding should be a dynamic one, which conceptualizes power as a process rather than a particular set of results. She conceptualizes empowerment as a process that cannot be done to women but must emerge from them. In Rowlands' (1998) study of how to assess the impact of empowerment among women in an educational program in Honduras, she concluded that empowerment took different forms in different spaces of women's lives. She argues that the process was personal and unique even though one woman may go through the same experiences as the collective. In turn, the personal empowerment of women with their husbands and other family relationships can be different from the personal and collective experiences (Ismail, 2006).

In a recent feminist book on this issue edited by McGee and Pettit (2020), power is conceptualized as "iterative, intersectional, and multidimensional, departing from the dualist 'agency vs structure' lenses that tend to prevail in theoretical debates" (McGee and Pettit, 2020: 4). These writers see power as visible (open commands and laws), hidden or shadow power (the setting of agendas of what can be said and done and what cannot be said or done) and, above all, invisible power. Invisible power includes "the socialized norms, beliefs and behavior shaping the boundaries of what is considered politically possible" (McGee and Pettit, 2020 cited in Rogers, 2020: 170). The "Me Too Movement" highlighted the invisible power that allowed men to continue to sexually abuse women even when these offenses were disclosed and found to be true. Spivak (1994) posed the question whether researchers can speak for poor women without violating their reality. Her article, "Can the subaltern speak?" (1994) has provoked a new consciousness among researchers and development practitioners to examine their assumptions, to remain vigilant about the

purposes of research and how women are represented in publications. In all the different learning situations discussed below, and those not highlighted in this chapter, the issue of power is ever present.

Learning in community development

Rogers (1992: 20–21) explains that learning is the process of making changes to knowledge, skills, attitudes, value systems and behavior and that learning includes goals, purposes, intentions, choices, and decision making. According to Anderson (1983), community is a sociological term that is not tangible and may also be defined as imaginary. It is also a term that often invokes sameness, a communal spirit, a geographical space, a sense of belonging, a language or a culture that unites people. It is a contested concept and often hides different forms of oppressions. Similarly, development is a term that has lost its promise of poor countries catching up with the rich (Kothari et al., 2019). Kilgore (1999) argues that theories of learning in community can “provide a framework in which we examine how people construct shared visions of social justice and learn and act together to promote these shared visions” (Kilgore, 1999: 201).

Women are usually the key actors in community issues as their lives are affected by the everyday struggles of survival. However, within communities, there are power struggles and hierarchies, such as those in patriarchal communities, in which women suffer various forms of gender oppression. These include structures such as chiefs in rural areas and local government authorities, landlords or gangsters who exert power and control over sections of a community in an urban environment. Even within communities, there are struggles for social change and for the same resources. There are also economic divisions, unequal gender dynamics, power struggles over leadership and corruption, or mismanagement of funds which impact the solidarity of the collective.

The book, *Playing with Fire* (Sangtin writers and Nagar, 2006), tells the stories of seven women village level activists from different religions and castes in rural India, who were employed by the Nari Sumata Yojana—a Non-Governmental Organization (NGO) which seeks to empower rural women of the Dalit castes (lowest castes). In this book, these women analyze how the different layers of the social units oppress women, firstly at the unit of family life, secondly at the community level and then thirdly at the wider social and global levels. Their reflections illustrate that the dominant ideologies that oppress women are interconnected and therefore it is difficult to tackle one oppression at a time. The women write that each social unit is filled with “poisons of casteism, classism, communism, and sexism” and that “these blend into one another [...] [so] that their bitterness remains intact on our tongues” (Sangtin writers and Nagar, 2006: 16). Their stories illustrate that communities may also be a prison for those who go against community norms.

The feminists, Weiler (1991) and Thompson (2000), argue that women in a single community do not form a homogenous category with a single ambition and consciousness but may be divided among themselves. Third World postcolonial feminists, Alexander and Mohanty (1997), argue that geography, genealogy, and colonial histories are important factors that impact on agency and interpersonal relationships while the postmodern feminist, Mouffe (1992), believes that there is not one absolute discourse or truth to development and equality in society, nor one linear path to it.

Feminists, Lee and Weeks (1991), maintain that the structure of women's lives militates against community building as women may be married and obey the demands of their husbands, they may work in a paid labor force or they may be preoccupied with domestic labor and family commitments. These arguments are valid but the history of women's movements and liberation struggles indicates a rich culture of women participating in community and mass democratic struggles (Foley, 1999; Ismail, 2015).

Poor communities have a history of working with “outsiders” that include state community workers, social activists, development workers or global organizations, such as UNESCO, that come from a range of political persuasions. The outside agencies' relationships and interactions with communities have implications for how learning is supported and how society is changed as the role of outsiders is to construct new ways in which the community sees itself. In this process, the members of the community learn that they need to become active citizens for the realization of their democratic rights (Shaw and Martin, 2000: 409). This is both a political and an intellectual challenge and allows for “those who have experienced social exclusion to have a voice” (Fiona Williams cited in Shaw and Martin, 2000: 410).

Many global institutions working in development contexts, such as UNESCO and UNICEF, often showcase uncritical and romantic stories of women in development on their websites which further entrenches the view that communities are homogenous. However, there are many examples worldwide to illustrate that women do work together across many differences as shown below.

In 2020, during the pandemic brought on by the COVID-19 virus, many women played critical roles in their communities by providing food, water, care, and information on health practices to vulnerable persons. Although women are on the frontlines in their homes, communities, and health care facilities, they are often excluded from the community, governance structures and national decision-making processes that determine their response (Fuhrman and Rhodes, n.d.).

The South African Homeless Peoples Federation is an umbrella body for the collection of savings of groups of poor African women living in informal settlements. These small daily savings, together with a state subsidy, have enabled many women to build their own homes. The women say that the savings form the cement that binds them together in solidarity. The women, firstly, learned to save and then went on to learn other skills, such as brick-laying, designing, and building houses. In the process, they also gained political skills to negotiate with government for land and housing subsidies. These learning outcomes took them beyond their own communities as they formed alliances across the globe and exchanged knowledge and skills with other women as far afield as India, Brazil, and the Philippines. They connected intimately with the Slum Dwellers in India (SDI) and others in the Asian Coalition for Housing Rights (Ismail, 2015). These exchanges of knowledge, skills and political strategies can be compared to

Alexander and Mohanty's (1997: xix) "transnational border crossings" in which women from similar backgrounds and colonial histories in different geographical spaces engage in feminist praxis. In these cross-cultural exchanges, local knowledge becomes international knowledge that counters the limitations of local knowledge and opens up spaces for poor people to become experts for each other, locally and internationally (Mayo, 2005: 113).

The examples above illustrate Lave and Wenger's (1991) notion of learning in "communities of practice" as they foreground learning rather than teaching and argue that learning is best viewed as a social practice. A "community of practice" is formed by a community's pursuit of engagement with the rest of the world. If the experience is meaningful, Lave and Wenger (1991) conclude that people are learning. For the community, learning is the vehicle for its development and transformation. In this process, people act as a resource, exchange information, help to make sense of situations, share new ideas, keep each other company and add life to each other's working days. Critical to this theory is identity formation as Lave and Wenger (1991) show that learners are engaged in the co-construction of knowledge through participating in a continuing dialogue between themselves, the other people they are interacting with and the systems, technologies, and the wider culture of that setting. There is an experience of meaning because, for them, learning is a social phenomenon and reflects the social nature of human beings.

The various responses to the pandemic in which women took the lead, cited in this chapter, are illustrative of Lave and Wenger's (1991) theory as, in many communities, women responded to the crisis to ensure that their communities were safe, had basic needs and were organized. For example, in Kerala in India, a state with a high literacy rate, empowered women's self-help groups, such as Kudumbashree, formed thousands of WhatsApp groups connecting poor neighborhoods. These WhatsApp groups were used to educate on key safety measures as advocated by the government during lockdown. They also initiated community kitchens, with the support of local community organizations and NGOs, to provide free meals to many vulnerable people (World Health Organization [WHO]; JASS, 2021).

Some feminist literature discusses women's learning from a psychological point of view. This literature suggests that there is a reciprocal relationship between individuals and their communities (Kilgore, 1999). Benky et al. (1986) describe how women, who come from disadvantaged backgrounds and seek medical advice in a health clinic, learn more by listening and receiving knowledge from other poor women who devote themselves to the care and empowerment of others by listening and understanding or by teaching others what they know, and that this experience gives them more confidence. Benky et al. (1986: 228) model shows that learning helps women "toward community, power, and integrity [and that] such an education facilitated the development of women's minds and spirits".

In a different study of working-class women and their relationships to knowledge, Wendy Luttrell (1988) argues that knowledge, which develops in the collective and from practice, is often more reliable than expert knowledge. Luttrell (1988) believes that common sense knowledge is experienced differently by white and black women. Patricia Hill Collins (1995) expresses a similar view and says the way in which black women claim their knowledge as "mother wit" is perceived, in part, as their ability to work hard and to obtain material things without a man's support. Often women refer to their common-sense knowledge as intuitive and stemming from feelings. Black women also regard their ability to deal with racism as another form of real intelligence, which they share with black men against the ignorance of whites. This relates to Luttrell's (1988) distinction between knowledge and wisdom. She says that black women are reminded of their collective identity as oppressed women daily and this mitigates the reminder of their individual identities as women (Barr, 1999: 108–111). These solidarities were expressed in the "Black Lives Matter" social movement in America when black women and men marched against police brutality and the indiscriminate murders of black people by the police.

Learning in social movements

Women's movements, like most social movements, are issue based and are associated with a broad range of struggles that include national liberation, human rights, the democratization of authoritarian regimes, working class struggles, gender specific grievances and concerns over basic needs. In these social movements, women are organized around a range of issues that include: political and legal rights; violence against women; reproductive choices; abortion; sexual freedom; employment opportunities and discrimination; dowry deaths; access to land, houses and finance; sustainable development; and the environment. An example is the story told in the film, *Women Hold Up the Sky* produced in 2019 by the WoMin Alliance, of how women activists, affected by mining and other forms of large-scale extractives in South Africa, Uganda, and the Democratic Republic of Congo (DRC), are deeply engaged in resistance and an active struggle to take back control of their land, their rights, their bodies, their lives, their livelihoods and their environments (WoMin, 2019).

Social movements generally lose momentum and may dissolve once the issues under protest are resolved (Tarrow, 1994). These movements may not be concerned with the capture of state power and revolutions. However, women may be consciously or unconsciously building a countervailing power to dominant state power.

Eyerman and Jamieson (1991) argue that social movement activists learn by doing—people learn by moving from discovery to articulation to application and to becoming part of an organization. The knowledge that is created is therefore a collective process. The movement creates a space for new kinds of knowledge and ideas, and new relationships to develop. The creation of knowledge as well as identity formation is central to the movement. Knowledge is created in a series of meetings, between meetings, and, most importantly, between movements and their opponents. The new ideas produced on specific issues and topics are shared by the participants. The focus is on creating a movement identity, which is a cognitive praxis. The actors taking part in the process of

articulating a movement identity are called “movement intellectuals” (Eyerman and Jamieson, 1991). Illustrative of these theories are the new environmental and climate change social movements which are youth led, for example, by Greta Thunberg who started skipping school in 2018 to strike for climate action outside the Swedish parliament in Stockholm. The movement quickly went global, and she became the symbol for youth activism against a greedy corporate world (Marris, 2019).

The “Me too Movement” in which women shared their experiences of sexual abuse and harassment in the workplace is another example of a movement in which women from all spheres of life and across classes, races, genders and ages found a common cause, broke their silence, and collectively supported one another to envision a society in which patriarchy is dismantled.

Newman (2005) outlines how to use information and knowledge as resources for critical skills such as challenging power relationships. The Treatment Action Campaign (TAC) in South Africa exemplifies this characterization. The TAC built a movement which focused on removing the negative stigma attached to HIV/AIDs sufferers by brandishing the words “HIV Positive” on T-shirts which provoked questions, raised awareness, empowered the wearers and created solidarity between those infected and affected by HIV (Ismail, 2017: 161). The TAC took the government and huge pharmaceutical companies to court to provide affordable medicines and, in this process, embarked on an education campaign to counter the messages from government that HIV/AIDS is primarily a disease of poverty which can be cured by eating garlic and potatoes.

Foley (1999) and Wildermeersch and Jansen (1997), through case study research, conclude that learning cannot be determined by measuring cognitive skills as learning is not only about acquiring skills and knowledge but is also about change in attitude and behavior and does not involve a timetable or organization of time. Learning is national and international as social activism moves knowledge beyond its national borders but they caution that not all learning is progressive or linear.

An example of transnational learning or learning across borders from others was present in the Dakota pipeline protests in which Native-American women were at the frontline to protect the rivers against the oil corporates and gained global support against oil corporations and the USA government. In Mexico, Karen Thunder learned that pipelines were coming down from Canada and then through the United States to Mexico. As she learned about pipeline struggles across the borders, she realized that this was a global issue. In her travels after Standing Rock, Thunder listened to stories about rivers and lakes that need to be defended from oil/gas and mining extraction, known as extractivism (Barnett, 2019).

Learning in virtual communities

Virtual communities began with the concept of networking which was used mainly to co-ordinate a more efficient use of resources but is now increasingly used to exchange information, as an advocacy tool and for learning about or from different social and economic struggles.

Acharya and Robinson-Pant (2017), who research women's literacy and health programs in Nepal, emphasize the need to look beyond planned development programs as people's lives and literacy learning opportunities are rapidly changing due to new technologies, which have increased mobility and globalized markets. They argue that the boundary of literacy learning has expanded with the development of technologies and can no longer be confined to formal or non-formal education programs. Access to mobile phones and computers has expanded people's choices and use of communication media, even in remote areas. Communication choices and cultures have become more personalized and covert, offering new spaces for women and men to develop relationships and transforming indigenous social structures. Acharya and Robinson-Pant (2017) cite the following examples from their research with the Gurung community in Nepal:

The Gurung community has a specific cultural practice known as “Rodhighar” particularly organized to facilitate socialization among young men and women. Today, access to mobile phones offers young people the opportunity for virtual socialization, replacing indigenous social practices like the “Rodhighar” but there is some evidence that mobile phones are having a negative impact on gender equality and health enhancement. Recent research on the misuse of mobile phones by young boys and adult men has shown serious implications for women's mental and physical health.

Another illustration from Tanzania in Africa was the development of a mobile application (app) for informal women traders to market their agricultural products. This innovation means that the women do not have to travel to informal markets or go door to door to sell their produce. However, the new technology does require a substantial investment in a smart phone, batteries and internet connectivity, all of which are difficult for some women to purchase (Kaoinga et al., 2016).

In 2020, the world was caught in multiple crises that included the COVID-19 pandemic, economic recession, climate change, environmental degradation, and racism. During this time, learning about the virus took place mainly informally and virtually as formal education institutions were closed, and people were not in physical contact. People learned informally about the origins of COVID-19, how it spread, its impact, the possibility of a vaccine, the treatment of patients and how countries reacted to contain the spread of the virus. There was a mass of information, some of it based on science (mainly from the World Health Organization), some on experiences of living with the virus, in addition to misinformation and alternative views that the pandemic was a conspiracy. There was also widespread use of community radio and television in countries in which internet connectivity was poor. Posters and pamphlets were put up in public places to alert people to wear masks, sanitize hands and to practice social (physical) distancing.

Social media to disseminate information used included Facebook, Twitter, WhatsApp, Instagram, Webinars and Watch parties. Statistics of the spread of the pandemic globally were provided graphically in dynamic ways. However, these forms of online learning highlighted the digital divide and therefore calls were made for Digital Social Justice, especially in rural areas.

Women-led trade unions and NGOs ensured that workers had personal protective equipment (PPE) and many civic organizations urged the state to supply PPEs at schools and public facilities when they reopened. Women in Informal Employment: Globalizing and Organizing (WIEGO, 2022) argues that many informal laborers are “essential workers”, including waste-pickers and street vendors selling affordable foods and should be included in the free distribution of PPEs. In Mexico, a feminist trading platform supported the exchange of goods and services for vulnerable workers, with a deeper objective of fostering women’s solidarity (International Institute for Environment and Development [IIED], 2020). In Sudan, MANSAM, a coalition of women’s organizations and civil society groups, took on the pandemic by making masks, donating needed supplies, distributing posters with essential information and raising awareness about the virus via social media (Diop, 2020).

There is evidence that the COVID-19 pandemic disproportionately affected women and girls therefore it was important that COVID-19 responses were gendered and that women and girls were able to participate in making the decisions that affected them (JASS, 2021).

Conclusion

The theoretical and conceptual framework used in this chapter drew largely on theorists from feminist, popular education and critical perspectives to discuss women’s informal learning in contexts such as communities, social movements and virtual spaces. The analytical concepts identified were experiential learning, reflection, action, participatory democracy, conscientization and empowerment. It emphasized that women and communities are not homogenous categories and that learning needs to be analyzed on the micro-level in terms of wider socio-economic and political forces. It characterized informal learning as occurring on many levels and that it was tacit, incidental, active, planned, interpretive, instrumental, critical, intellectual, emotional, and physical. Learning in the various spaces outlined is a collective experience that involves an activity toward a common purpose, creates new identities and shifts social learning to its transformative power, and points to possibilities to intervene and make things better (Ismail, 2006:54). In the informal learning context women learn by reflecting on lived experience and, in these ways, generate new knowledge and understanding that may challenge oppressive relations. The literature suggests that learning, empowerment, and consciousness raising are not linear processes but are dependent on their contexts that are contested concepts. Learning, in this paradigm, is a central element in development and social change.

The complex picture of social, political, economic, and environmental change that emerges from this chapter suggests that social transformation—particularly in gender relations and roles—comes about when people become “co-creators” of knowledge, learning informally with and from their peers, and choosing when and how to engage with new knowledge and skills (Acharya and Robinson-Pant, 2017: 15).

The importance of this review for educators who teach in diverse contexts characterized by differences in race, gender, ethnicity, nationalities, religions, and world views toward democracy and social justice is to be conscious of the histories of their learners, the power dynamic between educator and learner, to work from the personal experience of the participants, allow them to co-design the curriculum and see them as co-producers of knowledge. Educators can take heed from the late bell hooks’ (1994) notion of an engaged pedagogy, based on Freire’s liberatory pedagogy. In summary she urges educators who work in diverse learning spaces to take risks, to share their personal experiences with learners, to help students think critically, to transgress dominant ideologies, to care about each learner, to educate with their hearts and minds and to act against injustices for freedom.

References

- Acharya, S., Robinson-Pant, A., 2017. Women, literacy, and health: comparing health and education sectoral approaches in Nepal. *Compare* 49 (2), 211–229.
- Alexander, M.J., Mohanty, C. (Eds.), 1997. *Feminist Genealogies, Colonial Legacies, Democratic Futures*. Routledge, New York.
- Anderson, B., 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Versa, New York.
- Arnold, R., Burke, B., James, C., Martin, D., Thomas, B., 1991. *Educating for Change*. Between the Lines and the Doris Marshall Institute, Toronto.
- Barnett, T., 2019. *Standing Rock: Three Years and Still Fighting*. <https://www.resilience.org/>.
- Barr, J., 1999. *Liberating Knowledge: Research, Feminism and Adult Education*. NIACE, London.
- Belenky, M.F., Clinchy, B.M., Goldberger, N.R., Tarule, J.M., 1986. *Women’s Ways of Knowing, the Development of Self, Voice and Mind*. Basic Books, New York.
- Brookfield, S., 2001. *Understanding and Facilitating Adult Learning: A Comprehensive Analysis of Principles and Effective Practices*. Open University Press, Milton Keynes.
- Diop, B., June 17, 2020. The Impact of COVID-19 Pandemic on the Women, Peace and Security Agenda in Africa. <https://www.accord.org.za/>.
- Eyerman, R., Jamieson, A., 1991. *Social Movements: A Cognitive Approach*. Polity Press, Cambridge.
- Foley, G. (Ed.), 1999. *Learning in Social Action: A Contribution to Understanding Informal Education*. Zed Books, London.
- Freire, P., 1983. *Pedagogy of the Oppressed*. Continuum, New York.
- Fuhrman, S., Rhodes, F. (n.d.). The Absence of Women in COVID-19 Response. <https://insights.careinternational.org.uk/>.
- Hill Collins, P., 1995. The social construction of black feminist thought. In: Guy-Sheftall, B. (Ed.), *Words of Fire: An Anthology of African American Feminist Thought*. The New Press, New York, pp. 337–357.
- hooks, b., 1994. *Teaching to Transgress-Education as the Practice of Freedom*. Routledge, New York.
- Ismail, S., 2006. “A Poor Women’s Pedagogy” - An Exploration of Learning in a Housing Social Movement. PhD Thesis. University of Cape Town, Cape Town.
- Ismail, S., 2015. *The Victoria Mxenge Housing Project: Women Building Communities Through Social Action and Informal Learning*. UCT Press/Juta, Cape Town.
- Ismail, S., 2017. Contours of radical pedagogy: saving lives and building solidarity. In: Von Kotze, A., Walters, S. (Eds.), *Forging Solidarity: Popular Education at Work*. Sense Publishers, Amsterdam, pp. 159–170.
- International Institute for Environment and Development (IIED), 2020. *Emerging Lessons From Community-Led COVID-19 Responses in Urban Areas*. <https://www.iied.org/>.

- Jarvis, P., 1987. *Adult Learning in Social Context*. Groom Helm, London.
- JASS, 2021. *Dialogue 2: Learning From the Crisis*. <https://justassociates.org/>.
- Kaingo, A.F., Montero, C.S., Mbini, E.M., 2016. Investigating Women Entrepreneurs' Usage of Mobile Technology for Accessing Markets in Iringa, Tanzania. Paper presented at Conference Interactive Mobile Common, Technology and Learning in San Diego, USA.
- Kaplan, T., 1997. *Crazy for Democracy: Women in Grassroots Movements*. Routledge, London.
- Kothari, A., Salleh, A., Escobar, A., Demaria, F., Acosta, A., 2019. *Pluriverse: A Post-development Dictionary*. Tulika Books, Chennai, India.
- Kilgore, D., 1999. Understanding learning in social movements: a theory of collective learning. *Int. J. Lifelong Educ.* 18 (3), 191–202.
- Knowles, M.S., 1975. *Self-Directed Learning: A Guide for Learners and Teachers*. Prentice Hall, Englewood Cliffs, NJ.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Lee, B., Weeks, W., 1991. Social action theory and the women's movement: an analysis of assumptions. *Community Dev. J.* 26 (3), 1–3.
- Luttrell, W., 1988. Different women's ways of knowing: a feminist perspective on the intersection of gender, race and class in the social construction of knowledge. In: Zukas, M. (Ed.), *Proceedings From the Transatlantic Dialogue*. School of Continuing Education, University of Leeds, Leeds, pp. 249–254.
- Marris, E., 2019. Why Young Climate Activists Have Captured the World's Attention. <https://www.nature.com/>.
- Mayo, M., 2005. *Global Citizens: Social Movements and the Challenge of Globalization*. Zed Books, London.
- McGee, R., Pettit, J. (Eds.), 2020. *Power, Empowerment, and Social Change*. Routledge, London.
- Mouffe, C. (Ed.), 1992. *Dimensions of Radical Democracy: Pluralism, Citizenship and Community*. Verso, London.
- Newman, M., 2005. Popular teaching, popular learning and popular action. In: Crowther, J., Galloway, V., Martin, I. (Eds.), *Popular Education: Engaging the Academy: International Perspectives*. NIACE, Leicester, pp. 22–31.
- Pateman, C., 1990. *The Disorder of Women*. Stanford University Press, California.
- Rogers, A., 1992. *Adults Learning for Development*. Cassel Education Limited and Education for Development, London.
- Rogers, A., 2020. Two books about power: a comparative review. *Int. Rev. Educ.* 66 (2–3). <https://doi.org/10.1007/s11159-020-09838-x>.
- Rowlands, J., 1998. A word of the times, but what does it mean? Empowerment in the discourse and practice of development. In: Afshar, H. (Ed.), *Women and Empowerment: Illustrations From the Third World*. Macmillan, London, pp. 11–34.
- Sangtin writers, Nagar, R., 2006. *Playing With Fire: Feminist Thought and Activism Through Seven Lives in India*. University of Minnesota Press, Minneapolis.
- Shaw, M., Martin, I., 2000. Community work, citizenship, and democracy: re-making the connections. *Community Dev. J.* 35 (4), 401–413.
- Spivak, G., 1994. Can the subaltern speak? In: Williams, P., Chrisman, L. (Eds.), *Colonial Discourse/Post-colonial Theory: A Reader*. Columbia University Press, New York.
- Tarrow, S., 1994. *Power in Movement: Social Movements, Collective Action and Politics*. Cambridge University Press, Cambridge.
- Thompson, J., 2000. *Women, Class, and Education*. Routledge, London.
- Walters, S., Manicom, L. (Eds.), 1996. *Gender in Popular Education. Methods for Empowerment*. Zed Books, London.
- Walters, S., 1989. *Education for Democratic Participation: An Analysis of Self-Education Strategies Within Certain Community Organisations in Cape Town in the 1980s*. University of the Western Cape, Centre for Adult and Continuing Education, Bellville.
- Walters, S., 1998. Informal and non-formal education. In: Stromquist, N.P., Monkman, K. (Eds.), *Women in the Third World: An Encyclopaedia of Contemporary Issues*. Garland Publishing Inc, New York, pp. 436–443.
- Weiler, K., 1991. Freire and a feminist pedagogy of difference. *Howard Educ. Rev.* 61 (4).
- Widermeersch, D., Jansen, T., 1997. Strengths and limitations of social learning as a key concept for adult and continuing education in reflexive modernity. In: *Proceedings of the 27th Annual Scrutor Conference*. Birbeck College, London.
- Women in Informal Employment: Globalizing and Organizing (WIEGO), 2022. <https://www.wiego.org/>.
- Womin, 2019. *Women Hold Up the Sky* [Film]. <https://womin.africa/multimedia/>.
- World Health Organization (WHO), 2020. Responding to COVID-19 - Learnings From Kerala. <https://www.who.int/india/news/>.

Relevant websites

- Dakota Standing Rock Pipeline Protests: <https://www.resilience.org/>.
- Gendered Response to COVID: <https://justassociates.org/>.
- Jass Associates: <https://www.justassociates.org/>.
- Taliban and Girls Education: <https://www.usip.org/publications/2022/04/talibans-ban-girls-education-afghanistan>.
- The World: <https://theworld.org/stories/2016-10-31/>.
- UNESCO: <https://en.unesco.org/>.
- UNESCO: <https://en.unesco.org/themes/education-and-gender-equality>.
- UNICEF: <https://www.unicef.org/>.
- Women Hold Up the Sky: <https://womin.org.za/>.
- Women Led Responses to COVID: <https://www.iiid.org/>.
- World Bank: <https://www.worldbank.org/en/home>.
- World Bank: <https://www.worldbank.org/en/topic/gender/overview#2>.
- World Health Organization: <https://www.who.int/>.

Informal environmental citizenship education

Li-Ching Ho^a and Tricia Seow^b, ^a University of Wisconsin-Madison, Madison, WI, United States; and ^b National Institute of Education, Nanyang Technological University, Jurong West, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Environmental citizenship education	330
The rights and responsibilities of environmental citizenship	330
Public and private realms of environmental citizenship	331
Spatial and temporal dimensions of environmental citizenship	331
Informal and non-formal environmental education	332
Informal learning about the environment	332
Evaluating the efficacy of informal environmental education	333
Case studies of informal environmental citizenship education	333
Case study 1—Women in Agroecology Leadership for Conservation, Guatemala	333
Case study 2: Roots and Shoots, China	334
Case study 3: the Wabanaki Youth Science (WaYS) program, United States	334
Conclusion	335
References	335

Environmental learning can occur in a myriad of ways and settings. Yet, relatively few scholars and educators have focused on how learning *about* and *for* the environment can be facilitated in informal learning contexts. The notion of environmental citizenship education, similarly, is one that has remained in the margins of the field, in part because of its amorphous, ambiguous, and interdisciplinary nature, and its overlap with other concepts such as ecological citizenship and sustainability citizenship. In this chapter, we start by providing an overview of the concepts of environmental citizenship education and informal education. We then introduce three case studies of informal environmental education in Guatemala, China, and the United States.

Environmental citizenship education

The idea of environmental citizenship is highly contested in part because the very concept of *citizenship* is controversial. As Dobson and Bell (2006, p. 4) have pointed out, there is “no determinate thing called ‘environmental citizenship.’” Even though there has been increasing attention paid to environmental citizenship within the fields of environmental science, social studies, and civic education, there is still significant uncertainty and debate about its core elements. Unsurprisingly, therefore, scholars have adopted different definitions of environmental citizenship. For example, it has been conceptualized as a matter of justice, a commitment to the common good, and a radical challenge to social injustices (Barry, 2006; Dobson, 2007). Others, however, have characterized environment citizenship in terms of environmental rights, relationships, and obligations (Bell, 2004; Warwick, 2012).

Apart from the different scholarly conceptions of environmental citizenship highlighted above, various national and international government and non-governmental organizations have also developed their own definitions. For example, the European Network for Environmental Citizenship (2018: para. 1) characterized education for environmental citizenship as cultivating “a coherent and adequate body of **knowledge** as well as the necessary **skills, values, attitudes and competences** that an environmental citizen should be equipped with in order to be able to act and participate in society as an **agent of change** in the **private and public sphere**, on a **local, national and global scale**, through **individual and collective actions**.” Notably, the organization has also emphasized the importance of considering democratic action, structural causes of environmental issues, as well as inter- and intra-generational justice.

In this section, we focus on three key elements highlighted in the preceding paragraphs: (a) rights and responsibilities of environmental citizenship; (b) public and private environmental citizenship; (c) spatial and temporal dimensions of environmental citizenship.

The rights and responsibilities of environmental citizenship

A liberal understanding of environmental citizenship emphasizes how citizens have substantive and procedural environmental rights such as a right to environmental goods including water and air that meets particular health and legal standards, and a right to participate in developing policies and regulations that establish those standards (Bell, 2005; Schild, 2016). The Aarhus Convention, which entered into force in 2001, is a prominent example of a legally binding international

agreement that has established environmental rights for individuals and groups. The three rights that the Convention provides for include: (a) the right of access to environmental information held by public authorities; (b) the right of the public to participate in making decisions with regard to the environment; and (c) the right to challenge or contest decisions made by public authorities (United Nations Economic Commission for Europe, n.d.). Significantly, this international agreement imposes obligations on governments to be accountable, transparent, and responsive toward the public with regard to environmental issues.

The civic republican conception of environmental citizenship, on the other hand, argues that citizenship is not just about rights and entitlements but also requires a reciprocal commitment to fulfilling obligations and duties (Smith and Pangsapa, 2008). From this perspective, environmental citizenship is thus primarily concerned about “doing, acting, and performing duties and public acts ... for the common/public good” (Barry, 2006: p. 27). This definition offered by Barry highlights several key elements including civic virtues, civic responsibilities and service, and a commitment to the common good. These virtue- and duty-based understandings of environmental citizenship can frequently be found in school curricula and in public statements offered by governments and international organizations. For example, the Government of Canada’s (2012: Citizenship Responsibilities, para. 6) overview of Canadian citizens’ responsibilities states: “Every citizen has a role to play in avoiding waste and pollution while protecting Canada’s natural, cultural and architectural heritage for future generations.” In a similar vein, the European Network for Environmental Citizenship (2018: para. 1) defines environmental citizenship as “the exercise of environmental rights and duties” and these duties include being “agents of change” as well as taking action to solve and prevent environmental problems.

Civic republican and liberal traditions of citizenship have greatly influenced environmental education despite the attempts of some scholars to push back against largely Eurocentric assumptions of such traditions. Scholars in Latin America, for instance, have argued that it is important to go beyond the traditional Western focus on rights and responsibilities to focus on a conception of “citizenship” that is “more integrally ecological” (Latta and Wittman, 2010: p. 113). This alternative conception needs to acknowledge indigenous worldviews, consider different ways of knowing, and recognize the “materiality and subjectivity of nature” (p. 112).

Public and private realms of environmental citizenship

Second, environmental citizenship encompasses *both* the public and private realms. The fluid and multidimensional nature of environmental issues means that the distinction between public citizen and private citizen is blurred. As Dobson and Bell (2006: p. 7) have pointed out, “‘private’ actions can have important public consequences.” For example, an ostensibly private consumer decision like purchasing a car or driving to work (rather than taking public transport) can have public environmental and social consequences such as increased air and noise pollution, and greater fossil fuel consumption. Similarly, other private actions such as deciding whether to install solar panels or home insulation may have significant public consequences in terms of reducing energy use and wastage.

Some scholars have suggested, therefore, that environmental citizenship should address both public and private sphere environmentalism (Stern, 2000). *Public* environmental citizenship activities can incorporate different kinds of environmental activism ranging from low-commitment activities such as lobbying, petitioning, and donating money, to actively participating in environmental organizations and movements. There are also non-activist types of public environmental behaviors and these include articulating support for environmental policies or willingness to pay higher taxes. *Private* sphere environmentalism, on the other hand, is limited to individual behaviors in the private sphere such as recycling and purchasing energy efficient products. The environmental consequence of each of these behaviors is small but when aggregated, these can potentially have a significant impact (Stern et al., 1999).

Most environmental education programs have tended to focus on the private sphere and behavioral modification of individuals (Berkowitz et al., 2005). However, as Chawla and Cushing (2007) and others have argued, the effectiveness of private sphere actions is limited unless it includes collective organizing for changes in the public sphere. This is because collective political engagement can move business and government to remove barriers to action in private life. For instance, consumers cannot opt for public transport if the infrastructure is not made available by businesses and the state. Consequently, the goal for environmental education should not simply be environmental action but should instead be the “most strategic actions” in the public and private spheres (Chawla and Cushing, 2007: p. 438). Therefore, drawing on the work of Bandura (1997), Chawla and Cushing (2007: p. 446) also suggest that education for the environment should not just focus on developing individual competence, but also *collective* competence, defined as “the belief held by members of a group that they can coordinate their actions effectively and accomplish shared goals through unified efforts.” This can include having opportunities to make meaningful and successful contributions to community projects as well as learning how to plan, coordinate, and collaborate with diverse others.

Spatial and temporal dimensions of environmental citizenship

Citizenship is, as Shin and Bednarz (2019: p. 1) remind us, “inherently spatial” and this has been further exacerbated by the increasing ubiquity of geospatial technologies that have changed the nature of our relationships with and understandings of the physical world. In addition, given the increasingly complex and multidimensional nature of environmental issues, it is important

for students to be conscious of the *spatial* dimensions of citizenship (e.g., Cogan and Kubow, 1997; Shin and Bednarz, 2019). While some environmental citizenship education programs explicitly or implicitly assume what Lorimer (2010: p. 313) characterizes as “a commitment to a global spatial imaginary” such as open borders, global networks, and mobility of people and ideas across international borders, other programs are significantly more local and limited in scope.

It is important to consider the spatial dimensions of environmental citizenship programs because research has shown critical spatial thinking and awareness of socio-spatial processes can help students acquire deeper understandings of the impact of historical and present-day boundaries, spatial exclusion, and spatial inequalities (Gordon et al., 2016). In addition, students’ willingness to engage in broader and more global forms of civic action can be shaped by the kinds of knowledge, experiences, and interactions that they have had at different levels. Some studies, for instance, have shown that students who were exposed to a curriculum that focused on personally relevant local problems and issues and how the local government addressed these problems learned more about how social change occurs (Kahne et al., 2006).

Concurrently, students also need to be able to understand the *temporal* dimensions of environmental citizenship, namely how past actions can affect current experiences, and how present-day decisions can have an impact on future generations. Scholars such as Anand and Sen (2000) have, for example, written about the relationship between environmental sustainability and intragenerational/intergenerational justice and equity. They argue:

We cannot abuse and plunder our common stock of natural assets and resources leaving the future generations unable to enjoy the opportunities we take for granted today. We cannot use up, or contaminate, our environment as we wish, violating the rights and the interests of the future generations. That universalism also requires that in our anxiety to protect the future generations, we must not overlook the pressing claims of the less privileged today. A universalist approach cannot ignore the deprived people today in trying to prevent deprivation in the future (p. 2030).

In other words, it is vital for environmental citizenship education to help students understand the necessity of fair and equitable allocation of environmental resources between the present and future generations.

The spatial and temporal dimensions of environmental citizenship thus suggest that it is essential for students to recognize the *interrelated* dimensions of environmental concerns and problems. Collective action problems, for instance, are a useful way of illustrating both the spatial and temporal dimensions of environmental issues and how these dimensions demonstrate the interrelated and interdependent nature of such environmental concerns. According to Ostrom (2010: p. 155), collective action problems are “settings where decisions about costly actions are made independently but outcomes jointly affect everyone involved.” Environmental issues such as the management of fishing resources and extraction of water resources for farming or industrial use are collective action problems that can affect not only the local people living in the vicinity but can also have negative effects on people physically located far away. These collective action problems can also simultaneously affect both this generation and future generations. For example, even though an individual may personally benefit if he/she chose to exceed recommended catch limits for fish or use more water from a river for irrigating his/her crops, if everyone chose to maximize their short-term benefits by taking such actions, the existing and future entire natural world (not just humans) will be worse off.

Informal and non-formal environmental education

While scholars and educators have consistently acknowledged the importance of informal and non-formal learning, much of the attention within education research has been focused on formal learning environments. Scholars such as Scribner and Cole (1973), however, have argued for an increased emphasis on studying learning outside the formal education system because schooling in the industrialized world is mainly focused on didactic and rational approaches to learning whereas the bulk of learning that people go through takes place in everyday situations. Following Scribner and Cole, other scholars have also attempted to document the effectiveness of learning in societies which do not have formalized education structures (e.g., Greenfield and Childs, 1991; Saxe, 1988).

Learning is, as Falk and Dierking (2000: p. 136) point out, a context-based process that requires making explicit connections across time and understanding the intricate relationships among individuals, communities, and the physical environment. They write: “(Learning is) relating past experiences to the present, connecting what is happening in the present to what has happened in the past ... Learning is a dialog between the individual and his or her social/cultural and physical environment.” When applied to environmental education, this implies that both formal and informal kinds of learning are important for helping young people acquire environmental knowledge and providing them with an intellectual foundation for thinking about how humans interact with the environment. In order to support effective environmental education, therefore, the educational infrastructure should include both formal schooling as well as other social and cultural resources that support informal environmental experiences (Falk, 2005; St. John and Perry, 1993).

Informal learning about the environment

While environmental educators, researchers and policy makers have historically focused their attention on formal schooling, there has been a growing appreciation for the role of other non-school based resources for learning (Falk, 2005). These include

physical settings such as museums, nature parks, zoos and aquariums, as well engaging with media like nature documentaries on television, news about the environment in the newspapers, radio, or on the Internet. Additionally, learning may also occur through participation in environmental organizations such as the World Wildlife Fund, nature societies, or outdoor adventure clubs.

The definition of informal learning varies within the literature but in general, researchers have sought to define learning based on the inputs and outcomes of learning, as well as the degree of control the learner has over his/her own learning. Heimlich (1993), drawing on the Lifelong Learning Model developed by Mocker and Spear (1982), has suggested that learning settings can be differentiated based on who controls the objectives and means of learning. *Formal learning* occurs when institutions control both the objectives and means of learning, while *informal learning* involves individual control of the objectives but not necessarily the means of learning. For instance, Falk (2005: p. 272) has argued that many informal learning opportunities involve “carefully crafted pedagogy or andragogy,” and therefore are not completely without institutional facilitated “instruction.”

Heimlich (1993) has also proposed the category of *self-directed learning*, where individuals control both the objectives and means of learning. Similarly, Falk and Dierking (2000) have defined such individual control over learning as *free-choice learning*. Free-choice learning, they argue, emphasizes the underlying motivation and interest of the learner (rather than the objectives of the institution in which the informal learning takes place), and can occur in any sequence, at any time, and with complete learner volition. Therefore free-choice learning in environmental education leverages on the learner’s interest in the topic or setting, and allows the learner to select from a range of learning options. For example, Riemer et al. (2014) have proposed a model that allows participants the means to define the context of their participation and enact meaningful change in their own lives and/or the lives of their community.

Evaluating the efficacy of informal environmental education

Scholars have found that encountering nature in an informal setting often leads to greater environmental understandings and an increase in the adoption of environmentally sustainable practices. Numerous studies, for example, have noted the link between outdoor experiences and improvements in environmental knowledge (Hart, 1979; Lisowski and Disinger, 1991; Moore, 1986; Sobel, 1993). Other scholars such as Ballantyne and Packer (2005) have also observed how informal learning can impact the emotional engagement of learners. Unlike formal schooling, informal learning involves participants in meaning-making which occurs in context (Nunes, 2010) and thus learning about the environment will be imbued with “situation-specific meaning” (Falk and Dierking, 2000: p. 461). In addition, informal learning experiences can also contribute to children’s feelings of connection with nature, develop their knowledge and values related to the protection of nature (Palmberg and Kuru, 2000), and provide learners with positive experiences of how to apply the knowledge they learn to their own lives (Alsop and Watts, 1997; Ballantyne et al., 2001).

An additional advantage of informal learning is that it extends a society’s learning infrastructure for environmental education. For instance, informal learning can potentially help mitigate the limitations of the physical classroom (Schwarz and Stolow, 2006) or shortages in budgets for the provision of learning resources in schools. The kinds of learning experiences offered in informal settings are also potentially more diverse and provide spaces in which “the very different and competing discourses of the school system and the everyday world are reconciled” (Stockmayer et al., 2010: p. 26). Arguably, therefore, informal learning in environmental education has the potential to allow learners who have poor access to formal education resources to develop the knowledge, skills, and dispositions for transformational environmental action.

Case studies of informal environmental citizenship education

In this section, we highlight and analyze three diverse case studies of informal environmental citizenship education located in different national and cultural contexts—Guatemala, China, and the United States.

Case study 1—Women in Agroecology Leadership for Conservation, Guatemala

The first case study highlights an example of *formal learning* in an informal environmental citizenship education program. This case illustrates how the institution—the US-based Community Cloud Forest Conservation organization—determines both the objectives and means of learning for the participants. The Women in Agroecology Leadership for Conservation (WALC) program is a multi-dimensional 25-day environmental education and leadership program for young unmarried indigenous Q’eqchi’ Maya women aged 12–28 years living in the central highlands of Guatemala. The program has three requirements for the participants: (a) Complete the 25-day program; (b) Continue their formal education in schools; (c) Establish a small agroforest parcel of 16 trees as well as heirloom and traditional crops (Community Cloud Forest Conservation, n.d.). Upon completion of the program, the participants are awarded scholarships of approximately US\$135 for use toward school fees, books, and other expenses (Bowen and Miller, 2018). The curriculum, lessons and small group activities, conducted in the local language, are mostly led by peer leaders or staff members who were previously participants of the program.

The program appears to prioritize private sphere environmentalism by focusing on developing the participants’ individual capacities and understandings, although there is some attention paid to how these participants can become community leaders.

The program is principally centered on women's rights and empowerment, environmental conservation, as well as economic security. For example, the participants study a range of interrelated topics such as conservation, agroecology, sustainability, nutrition, and health, and they are taught about traditional Q'eqchi' crops, traditional medicinal plants, and reforestation strategies. In order to enhance their economic contributions to their families, students are taught practical skills such as candle-making and the production of dried fruit and jam for sale or for future home consumption (Bowen and Miller, 2018).

This program also focuses on the responsibilities of the environmental citizen, in particular, how the participants can contribute to ensuring the health and well-being of the cloud forest ecosystem. The interactive lessons held at the NGO's agroecology center include sustainable farming strategies such as the use and production of organic fertilizer, planting and grafting fruit trees, and managing vegetable gardens and tree nurseries (Bowen and Miller, 2018). Notably, in order to address the problem of deforestation, the program also emphasizes environmental stewardship and conservation biology classes that teach the participants about protecting local ecosystems and respecting wildlife (Briggs et al., 2019). These new skills and understandings are collectively designed to help the participants confront some of the socio-economic and ecological challenges that they face in their daily lives including gender prejudice and loss of biodiversity due to deforestation.

Case study 2: Roots and Shoots, China

Roots & Shoots (R&S) was established in Tanzania by Jane Goodall, the United Nations Messenger for Peace and renowned chimpanzee researcher, as an environmental and humanitarian program that empowers youth to participate in sustainable development. It operates in almost 100 countries, championing the development of ethical and sustainable environmental values and behaviors through community service programs. R&S is a good example of *informal learning* as it allows youths to be engaged in the community and develop projects that are relevant to their communities. While projects are identified and developed by youths, they nevertheless operate within an institutionally structured 4-step framework: (a) Understand one's own strengths and weaknesses; (b) Observe the community to identify areas of need; (c) Plan a service-learning campaign to address the need; and (d) Reflect on the impact of the campaign (Roots and Shoots, 2018). At the same time, program officers provide input on the feasibility of the proposals and guidance during the execution of the projects. This recalls Falk's (2005) assertion that informal learning activities may provide individuals with freedom insofar as objectives of learning are concerned, while crafting the means of learning.

Johnson et al. (2007) studied R&S clubs operating in schools in Shanghai, Beijing, and Wenzhou, China, and noted that youths had opportunities to participate in learning opportunities beyond the formal school curriculum. For instance, they went on field-trips to natural environments and attended art exhibitions which sought to raise environmental awareness. These students conducted research in collaboration with local communities to raise awareness of, and/or provide solutions to environmental issues. Similar to other R&S programs, these R&S clubs in China are premised on inculcating civic responsibility and working toward the common good. In fact, the authors found in their study that the impact was greatest in the domains of civic and social responsibility, as self-reported by youths and observed by their program coordinators. R&S also provided the social infrastructure to extend participants' personal environmental interests into the public sphere. For instance, respondents were highly motivated by the opportunity to establish and maintain contact with like-minded peers, and valued the connections they made with the local community. Moreover, while most of their actions were focused locally, R&S provided links to partner clubs in other parts of China, as well as the world.

Interestingly, the weakest benefits from the program was in the area of knowledge gains, particularly global environmental knowledge. This is unsurprising given the locally oriented nature of the R&S projects. Building on Schwarz and Stollow's (2006) assertion that informal learning can overcome the shortcomings of formal education, we argue that formal schooling can also buttress environmental citizenship education by helping students connect their experiential learning in the here and now to larger environmental issues at the international scale, and to understand the human-environmental interactions across time.

Case study 3: the Wabanaki Youth Science (WaYS) program, United States

The Wabanaki Youth Science (WaYS) program is another example of *formal learning* in an informal environmental citizenship education program. A collaboration between the University of Maine and the Native American Maine Wabanaki community, the program seeks to help young Wabanaki students (grades 6–12) to become future environmental leaders for their communities. The WaYS model is unique in its explicit attempt to fuse Western science knowledge with Traditional Ecological Knowledge (TEK) in order to meet the cultural needs of the students and their communities (Carr and Ranco, 2017).

The WaYS program includes three components: (a) Week-long summer earth camps and weekend mini camps; (b) TEK programs at teen centers; and (c) short- and long-term student internships (Carr et al., 2017). These three components, taught by university faculty, tribal cultural knowledge keepers, and other environmental leaders, are designed to provide the youth with opportunities to engage with the program over several years, and possibly even at the university level. Even though the WaYS program is fairly structured, the Native youth also have an opportunity to make decisions about the kinds of environmental issues that they are interested in and also identify additional programming that meets their needs. For example, the mini camp idea was initiated by the youth who wanted to be engaged year-round in activities (e.g., foraging, fire-building, and shelter-building), and not just the summer months (Carr et al., 2017).

The WaYS program incorporates both the private and public elements of environmental citizenship because it does not just focus on equipping individual students with environmental knowledge and skills, but also explicitly addresses the broader environmental

needs of the tribal communities. During the summer earth camp, for example, students are taught topics such as basket weaving, medicinal plants, tidal ecology, and the impact of climate change on fish habitats, while also working with tribal elders and natural resource experts on environmental issues that are particularly relevant to the tribal community (Carr et al., 2017; University of Maine, 2015). Importantly, the spatial and temporal dimensions of citizenship are also particularly prominent in this program with its emphasis on the traditional cultural heritage of the students, the teaching of cultural science, and the place-based nature of the learning.

Collectively, these three cases illustrate the different dimensions of environmental citizenship education and informal education—the rights and responsibilities of environmental citizenship, the public and private dimensions of environmental citizenship, and the spatial and temporal dimensions of environmental citizenship. Finally, these three case studies of environmental learning featuring indigenous youth and participants in non-Western settings also highlight the increasing need for scholars and educators to look beyond current dominant conceptions of environmental citizenship to incorporate unique and traditional ways of knowing and doing that are more culturally relevant.

Conclusion

The knowledge, competences, and attitudes necessary for individuals and communities to advocate for the environment at various scales is key to the development of just and equitable societies. While the focus of much educative effort remains on schools and formal education, we have highlighted how informal environmental education can extend the educational resources of a community beyond the formal school curriculum in transformative ways. Moreover, such community-based environmental education encourages learners to identify and act on issues that are culturally and contextually meaningful to their communities. The case studies highlighted in this chapter suggest how informal environmental education can strengthen learners' understandings of individuals' rights and responsibilities toward society, foster procedural knowledge of how to organize and work collectively toward a common goal, and/or connect local perspectives to broader environmental issues across space and time.

References

- Alsop, S., Watts, M., 1997. Sources from a Somerset village: a model for informal learning about radiation and radioactivity. *Sci. Educ.* 84 (5), 658–679.
- Anand, S., Sen, A., 2000. Human development and economic sustainability. *World Dev.* 28 (12), 2029–2049. [https://doi.org/10.1016/S0305-750X\(00\)00071-1](https://doi.org/10.1016/S0305-750X(00)00071-1).
- Ballantyne, R., Packer, J., 2005. Promoting environmentally sustainable attitudes and behavior through free-choice learning experiences: what is the state of the game? *Environ. Educ. Res.* 11 (3), 281–295.
- Ballantyne, R., Fien, J., Packer, J., 2001. Program effectiveness in facilitating intergenerational influence in environmental education: lessons from the field. *J. Environ. Educ.* 32 (4), 8–15.
- Bandura, A., 1997. *Self Efficacy: The Exercise of Control*. W. H. Freeman and Company, New York.
- Barry, J., 2006. Resistance is fertile: from environmental to sustainability citizenship. In: Bell, D., Dobson, A. (Eds.), *Environmental Citizenship: Getting From Here to There?* MIT Press, Cambridge, MA, pp. 21–48.
- Bell, D.R., 2004. Creating green citizens? Political liberalism and environmental education. *J. Philos. Educ.* 38 (1), 37–53.
- Bell, D.R., 2005. Liberal environmental citizenship. *Environ. Polit.* 14 (2), 179–194. <https://doi.org/10.1080/09644010500054863>.
- Berkowitz, A.R., Ford, M.E., Brewer, C.A., 2005. A framework for integrating ecological literacy, civics literacy, and environmental citizenship in environmental education. In: Johnson, E., Mappin, M. (Eds.), *Environmental Education and Advocacy: Changing Perspectives of Ecology and Education*. Cambridge University Press, Cambridge, pp. 227–266.
- Bowen, D.S., Miller, A.L., 2018. Education, leadership, and conservation: empowering young Q'eqchi' women in Guatemala. *Int. J. Educ. Dev.* 59, 28–34. <https://doi.org/10.1016/j.jiedudev.2017.10.012>.
- Briggs, L., Stedman, R., Krasny, M., 2019. Place attachment and social-ecological system sustainability examined through the voices of indigenous Guatemalan women. *Sustain. Sci.* 14, 655–667. <https://doi.org/10.1007/s11625-018-0634-6>.
- Carr, T., Rancho, D.J., 2017. Citizen science and traditional ecological knowledge—values of inclusion in the Wabanaki Youth Science Program. *Maine Policy Rev.* 26 (2), 86–88.
- Carr, T., Kenefic, L.S., Rancho, D.J., 2017. Wabanaki Youth in Science (WaYS): a tribal mentoring and educational program integrating traditional ecological knowledge and Western science. *J. For.* 115 (5), 480–483.
- Chawla, L., Cushing, D.F., 2007. Education for strategic environmental behavior. *Environ. Educ. Res.* 13 (4), 437–452.
- Cogan, J.J., Kubow, P.K., 1997. Multidimensional citizenship: educational policy for the 21st century. In: Final Report of the Citizenship Education Policy Study Project. Japan. Community Cloud Forest Conservation, (n.d.). Women in Agroecology Leadership for Conservation (WALC). <http://cloudforestconservation.org/our-work/walc/>.
- Dobson, A., Bell, D., 2006. Introduction. In: Dobson, A., Bell, D. (Eds.), *Environmental Citizenship*. MIT Press, Cambridge, MA, pp. 1–18.
- Dobson, A., 2007. Environmental citizenship: towards sustainable development. *Sustain. Dev.* 15, 276–285.
- European Network for Environmental Citizenship, 2018. Defining “Education for Environmental Citizenship”. <https://enec-cost.eu/our-approach/education-for-environmental-citizenship/>.
- Falk, J.H., Dierking, L.D., 2000. *Learning From Museums: Visitors Experiences and the Making of Meaning*. AltaMira Press, Lanham, MD.
- Falk, J.H., 2005. Free-choice environmental learning: framing the discussion. *Environ. Educ. Res.* 11 (3), 265–280.
- Gordon, E., Elwood, S., Mitchell, K., 2016. Critical spatial learning: participatory mapping, spatial histories, and youth civic engagement. *Child Geogr.* 14 (5), 558–572. <https://doi.org/10.1080/14733285.2015.1136736>.
- Government of Canada, 2012. Discover Canada—Rights and Responsibilities of Citizenship. <https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/discover-canada/read-online/rights-responsibilities-citizenship.html>.
- Greenfield, P.M., Childs, C.P., 1991. Developmental continuity in biocultural context. In: Choen, R., Siegel, A.W. (Eds.), *Context and Development*. Erlbaum, New Jersey, pp. 135–159.
- Hart, R., 1979. *Children's Experience of Place*. Knopf, New York.
- Heimlich, J.E., 1993. Nonformal Environmental Education: Toward a Working Definition. The Environmental Outlook. ERIC/CSMEE Informational Bulletin, Columbus, OH.

- John, S., Perry, D., 1993. A framework for evaluation and research: science, infrastructure and relationships. In: Bicknell, S., Farmelo, G. (Eds.), *Museum Visitor Studies in the 90s*. Science Museum, London, pp. 59–66.
- Johnson, L.R., Johnson-Pynn, J.S., Pynn, T.M., 2007. Youth civic engagement in China: results from a program promoting environmental activism. *J. Adolesc. Res.* 22 (4), 355–386.
- Kahne, J., Chi, B., Middaugh, E., 2006. Building social capital for civic and political engagement: the potential of high-school civics courses. *Can. J. Educ.* 29 (2), 387–409.
- Latta, A., Wittman, H., 2010. Environment and citizenship in Latin America: a new paradigm for theory and practice. *Eur. Rev. Lat. Am. Caribb. Stud.* (89), 107–116.
- Lisowski, M., Disinger, J.F., 1991. The effect of field-based instruction on student understandings of ecological concepts. *J. Environ. Educ.* 23 (1), 19–23.
- Lorimer, J., 2010. International conservation "volunteering" and the geographies of global environmental citizenship. *Polit. Geogr.* 29 (6), 311–322. <https://doi.org/10.1016/j.polgeo.2010.06.004>.
- Mocker, D.W., Spear, G.E., 1982. *Lifelong Learning: Formal, Nonformal, Informal, and Self-Directed*. Information Series No. 241. ERIC Clearinghouse on Adult, Career, and Vocational Education, Columbus, OH.
- Moore, R.C., 1986. *Childhood's Domain: Play and Space in Child Development*. Croom Helm, London.
- Nunes, T., 2010. Learning outside of school. In: Peterson, P., Baker, E., McGrath, B. (Eds.), *International Encyclopedia of Education*, fifth ed. Academic Press, Oxford, pp. 457–463.
- Ostrom, E., 2010. Polycentric systems for coping with collective action and global environmental change. *Global Environ. Change* 20 (4), 550–557. <https://doi.org/10.1016/j.gloenvcha.2010.07.004>.
- Palmberg, I.E., Kuru, J., 2000. Outdoor activities as a basis for environmental responsibility. *J. Environ. Educ.* 31 (4), 32–36.
- Riemer, M., Lynes, J., Hickman, G., 2014. A model for developing and assessing youth-based environmental engagement programs. *Environ. Educ. Res.* 20 (4), 552–574.
- Roots, Shoots, 2018. Formula Toolkit. https://news.janegoodall.org/wp-content/uploads/2018/05/Toolkit_.pdf.
- Saxe, G.B., 1988. Candy selling and math learning. *Educ. Res.* 17 (6), 14–21.
- Schild, R., 2016. Environmental citizenship: what can political theory contribute to environmental education practice? *J. Environ. Educ.* 47 (1), 19–34.
- Schwarz, E., Stolor, D., 2006. Twenty-first century learning in afterschool. *N. Dir. Youth Dev.* 110, 81–99.
- Scribner, S., Cole, M., 1973. Cognitive consequences of formal and informal education. *Science* 182 (4112), 553–559.
- Shin, E.E., Bednarz, S.W., 2019. Conceptualizing spatial citizenship: citizenship through geography. In: Shin, E.E., Bednarz, S.W. (Eds.), *Spatial Citizenship Education*. Routledge, Abingdon, pp. 1–9.
- Smith, M.J., Pangsap, P., 2008. *Environment and Citizenship: Integrating Justice, Responsibility and Civic Engagement*. Zed Books, New York.
- Sobel, D., 1993. *Children's Special Places: Exploring the Role of Forts, Dens and Bush Houses in Middle Childhood*. Zephyr Press, Brookline, MA.
- Stern, P., Dietz, T., Abel, T., Guagnano, G., Kalof, L., 1999. A value-belief-norm theory of support for social movements: the case of environmentalism. *Hum. Ecol. Rev.* 6 (2), 81–97.
- Stern, P.C., 2000. Toward a coherent theory of environmentally significant behavior. *J. Soc. Issues* 56 (3), 407–424.
- Stockmayer, S.M., Rennie, L.J., Gilbert, J.K., 2010. The roles of the formal and informal sectors in the provision of effective science education. *Stud. Sci. Educ.* 46 (1), 1–44.
- United Nations Economic Commission for Europe, (n.d.). Introduction: UNECE Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters. <https://unece.org/environment-policy/public-participation/aarhus-convention/introduction>.
- University of Maine, 2015. Camp Merges Environmental Science, Traditional Native Culture. <https://umaine.edu/news/blog/2015/07/14/finding-ways/>.
- Warwick, P., 2012. Climate change and sustainable citizenship education. In: Arthur, J., Cremin, H. (Eds.), *Debates in Citizenship Education*. Routledge, New York, pp. 132–145.

Youth as educators

Melissa Marini Švigelj^a and Cindy Cruz^b, ^a University of California at Santa Cruz, Santa Cruz, California, United States; and
^b University of Arizona, Tucson, Arizona, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	337
What is informal learning?	338
A brief genealogy of informal education concerns and technology	339
Youth as perpetual educators	340
Youth, media, technology, and social movements	341
#Hashtag	342
Conclusion	343
References	344
Further reading	345

Glossary

Youth Individuals who are approximately age twenty-five or younger

Technology Tools or devices used to communicate and create that require knowledge and skill for application, yet are always tied to social relations and institutions

Media Any tool or medium used for broad communication purposes

Social Movement Actions taken for justice to form a collective and coalitions for social, political, or economic change through strategizing and direct actions

Justice Equitable access to opportunities and resources for living a healthy, joyful, and peaceful life for all species

Educators Those of all species who humbly share their knowledge

Learning Receiving and accumulating knowledge that is shared or learned through experience

Introduction

As early as 1932, Vygotsky “identified the immense value of everyday, informal learning for young children” (Hedges, 2021, 1). Yet as Cohen noted in 2007, “The educational activities and settings defined as ‘informal’ or ‘non-formal’ are extremely diverse ... [and] The seemingly impossibility of isolating traits clearly distinguishing informal from formal education have fueled an ongoing debate” (p. 1). According to Rogoff et al. (2016), it is unfortunate that most of the attempts at defining or describing “informal learning” actually involve identifying how it is not “formal learning.” There are also descriptions and definitions of “informal,” “non-formal,” and “formal” learning within the thousands of available articles and books focusing on these topics that might be considered similar or contradictory at times (Cohen, 2007).

In 1982, La Belle called for an approach that examines formal and informal learning through a less distinct lens, arguing “that educational resources must be viewed as interacting modes of emphasis rather than as discrete entities” (p. 159). Additionally, La Belle asserts that various forms of “education may exist simultaneously, sometimes in concert with one another and sometimes in conflict” (1982, 162). Indeed, a 2012 article by Cartwright is titled “Informal education in compulsory schooling in the UK,” offering an example of exactly what La Belle asserted and exemplifying how “the two are constantly intermingled” and “how they are constantly interacting in the classroom” (Scribner and Cole, 1973, 553; 558). Furthermore, Callanan et al. note that a formal-informal dichotomy is too simplistic, and the authors highlight other theoretical and empirical works that make similar claims (2011). Thus, informal learning presents itself as an elusive, relational, contextual, and malleable concept.

In this entry, we use explorations of scholarly literature and online archives to situate youth alongside technologies and communication mediums as innovators and educators for social and political change. We begin by offering suggested criteria from the literature across disciplines and fields for distinguishing between what is labeled as *informal* versus *formal* learning or education and with this initial acknowledgment that the two concepts continuously commingle (Scribner and Cole, 1973). Next, we offer a brief genealogy of the fears often associated with new technologies and their potential to provide what adults often categorize as undesirable informal learning opportunities for children and youth. Then, we provide examples of young people as educators strategically utilizing available technologies to educate others, motivate direct actions, and build memberships and coalitions for justice. Finally, we suggest the need for expanded discussions concerning the purpose of learning or education in society and other notions for contemplation.

Throughout the remainder of this entry, we choose to engage with the words “informal learning” rather than “informal education” because the former indicates an action or a doing rather than the representation of an institution or system which might be more commonly associated with the noun “education.” Additionally, when referring to “youth” we consider recent research which has found that adult and teen brains work differently. Thus, “youth” or “young people” in this paper is referring to anyone who is approximately 25 years-old or younger in alignment with studies from neuroscientists who have found that “The rational part of a teen’s brain isn’t fully developed and won’t be until age 25 or so” (Campellone and Turley, n.d.). This word-choice determination was made alongside the understanding that “Childhood and youth are complex terms: socially constructed, historically contingent and variously located (Holloway and Valentine, 2000)” (Mills and Kraftl, 2014, 2). We are conceiving of the word “technology” as consistent with Raymond Williams’s view:

A technology ... is, first, the body of knowledge appropriate to the development of such skills and applications and, second, a body of knowledge and conditions for the practical use and application of a range of devices ... What matters in each stage, is that a technology is always, in a full sense, social. It is necessarily in complex and variable connection with other social relations and institutions ... (1981/1989, 173).

What is informal learning?

The difficulty of distinguishing between various types of learning is partially because “learning has been conceptualized quite differently by behaviorist, cognitive, developmental, and educational theorists” (Callanan et al., 2011, 646). However, scholars have long recognized that learning is embedded in social and cultural contexts and that people spend most of their time learning outside of school environments. In 1934, Thrasher wrote that most of what young people learn is not “acquired through the formal processes of education as exemplified in schools and other educational agencies” (p. 471). Rather, Thrasher contends that the majority of educational processes youth encounter “are the result of the universal process of social interaction” (1934, 471). This section investigates and organizes some of the themes and criteria present when scholars are defining and describing informal learning.

In 1950, Parker stated that “one’s informal education commences practically at birth ... with such rapidity [of development] that by the time [they] first reach what we think of as school-age, [they] already, thru wholly informal education, are a relatively accomplished young citizen” (p. 90). As an experienced educator in the Laboratory Schools at the University of Chicago, Parker would likely have observed many of the influences that accompany students learning yet exist outside of schools. Thus, researchers and educators have long recognized what was quantified at the National Science Foundation’s Center for Learning in Informal and Formal Environments in 2005: 81% of learning opportunities for children are provided outside of school settings (Rogoff et al., 2016).

Although there is not a standard definition of informal learning accepted and utilized across all areas of scholarship and approaches, most definitions offer similar characteristics. Often, informal learning is described as part of everyday activities in which people engage (Geerds et al., 2015). Coombs and Ahmed offer the following definition of informal education: “informal education is ‘the lifelong process by which every person acquires and accumulates knowledge, skills, attitudes and insights from daily experiences and exposure to the environment’” (Coombs and Ahmed, 1974, 161). In a review of the book, *Using informal education: An alternative to casework, teaching and control* (1990), Rogers presents some key characteristics of informal learning which are presented in the book he is reviewing. These include “the everyday life experience of the participants, not a set and codified curriculum” (Rogers, 1991, 105). Livingston also emphasizes the lack of mandated curriculum as a characteristic of informal learning stating, “Informal learning is any activity involving the pursuit of understanding, knowledge, or skill that occurs without the presence of externally imposed curricular criteria” (Livingstone, 2006, 206).

In addition to informal learning being flexible and having the process and content controlled by the learner, some argue that it occurs in an informal setting such as a space not specifically designed for schooling (Rogers, 1991). Informal learning practices occur “in the everyday spaces of human activity ... which lie outside the institution of schools and other formal learning sites” (Kauko and Wilkinson, 2018, 1176). Erstad recommends research approaches to learning that examine the inter-relationship between home, school, and community because “it offers a way of bridging the binary opposition between formal and informal learning” (2012, 30). However, Rogoff et al. contend that they believe “how learning is organized and supported is more important than where learning occurs” (2016, 357–358). Furthermore, Livingston argues that informal learning can be done anytime, anywhere, and with anyone (2006, 217). Although defining informal learning appears to be challenging and at times perhaps contentious, it seems “clear that informal learning occurs in many varieties across multiple communities of practice” (Callanan et al., 2011 646).

Drawing from a review of literature on informal learning, Callanan et al. present five dimensions that they associate with informal learning. These dimensions of informal learning include that it is non-didactic, highly socially collaborative, embedded in meaningful activity, initiated by learner’s interest or choice, and removed from external assessment (Callanan et al., 2011). Foley also emphasizes the social aspect of informal learning, describing it as people teaching and learning from each other “naturally and socially in workplaces, families, community organizations, and social action” (1999, 7).

According to Vizcaino-Verdu et al. cyberspace, digital capabilities, and virtual opportunities have created additional social spaces for informal learning and collective pedagogies with “a countless number of communities based on unlimited types of interests in which infinite forms of social relations develop” (Bautista-Sancho, 2012; 124)” (Vizcaino-Verdu et al., 2019, 94). Exploring queer identity through informal teaching and learning in online LGBTQ communities, Fox and Ralston assert that there is much informal

learning that occurs as users “seek out and establish connections with others via common nodes or common interests” (2016, 636). Thus, many of the informal learning dimensions identified by Callanan et al. appear across the literature with a particular emphasis on social interactions even as ways to develop communities and social interactions evolve over time alongside forms of communication and technologies.

A brief genealogy of informal education concerns and technology

As many scholars have noted, “For millennia, children have learned the skills, life ways, and philosophies of their communities informally, through their engagement in everyday family and community settings” (Rogoff et al., 2016, 362). However, studies of informal education in academia in the United States are a more recent phenomenon. Although the study of informal learning began during the same period as the origin of the American Educational Research Association, it was not until the 1920s that descriptions of how and why children were informally learning began to appear in anthropological publications. Shortly after in the 1930s, connections between anthropology and psychiatry developed (Rogoff et al., 2016) which seemingly increased interest in informal learning across disciplines. The multiple areas of study represented in the citations of this writing illustrate the trans-disciplinary and expansive interest in informal learning.

Many recent popular debates across disciplines center on young people’s online activities and focus on possibilities of harmful informal learning or negative activities. Alternatively, alarms about the harmful potentials within informal learning “have been criticized by some as being overly deterministic, as positioning young people as passive and ignoring the complexity of children’s and young people’s consumption of media” (Willet, 2008, 49). History can also demonstrate that student engagement with technologies and communication and student-led media can provide a platform for youth to educate others, “express their opinions, control their messages, and facilitate political participation” (Salamon, 2018, 4). This section reviews some of the technological advances that have ignited interest and concern among adults because of the technological capacities they create for informal learning opportunities among youth.

There is an interesting chronology of articles studying technological advances and mass media and their potential influence on the informal education of young people. Particularly, concerns continue to be raised that young people do “not yet possess fully the capacity for ‘adult discount,’ [and] do not yet have the background by which to discredit sufficiently” (Cressey, 1934, 506) those representations of life that technologies and mass media produce. For generations adults “have worried about whether exposure to ‘new’ forms of media, including the radio (Heisler, 1948; Longstaff, 1936), comic books (Thrasher, 1934), television (Maccoby, 1951), video games (Egli and Meyers, 1984), and violent media (Anderson et al., 2010; Ferguson and Kilburn, 2010), is ‘harming our children’” (George and Odgers, 2015, 832).

Funk (1998) argues that with the proliferation of radios in people’s homes in the 1920s and 1930s, “mass communication became mass education” (p. 43). The author also notes that “the history of the rapid expansion of the radio industry bears an uncanny resemblance to the all-consuming expansion of computing [and] telecommunications” (Funk, 1998, 31) in the late 20th and early 21st centuries. Along with adjusting to new technologies and forms of communication, the informal educational opportunities that might be available to young people through new technologies are often a source of discomfort for adults. Funk claims “In families where public information was usually screened by the parents and others in charge of moral conduct, many feared losses of control over the information their families heard from such an easily accessible source—the radio” (1998, 42). The author also describes a reciprocal relationship between what was learned in formal and informal education spaces and claims that speculating about the effects of one is not possible without including the other (Funk, 1998, 2016). Of course, as recording technologies improved and radios became more standard in automobiles, mostly white adults raised racialized concerns about the informal learning effects of exposing white youth to musical styles previously unknown to them (Clayton, 2016).

In 1934, Cressey expressed his concern that “The great frequency of cinema attendance by children and young people is not without its important educational effect ... [because] the motion-picture industry dispenses a great deal of informal education” (p. 505). Reflective of more contemporary criteria for identifying informal learning, Cressey notes that “In contrast to the traditional school, where motivation in learning arises extraneously, primarily through the teacher’s special efforts and skill, the cinema provides for many children a means, vicariously at least, by which learning may really be a natural result of interest and activity” (1934, 509). Preceding now commonly recognized socio-cultural traditions of study that emphasize context, Cressey avoids generalizing or promoting panic with regard to motion pictures explaining that “The full significance of the cinema, however, cannot be seen except by reference to the specific social backgrounds of each individual ... in terms of his own cultural heritages, his own dominant interests and values, and his own *axeological world* can we begin to see adequately what any photoplay may mean for a spectator” (1934, 11).

With the invasion of televisions into homes after World War II, concerns again arose but this time mostly regarding the effects that viewing violent television shows might have on children. Again, researchers attempted to assuage concerned adults explaining that “opinions as to the consequences [of televised portrayals of violence] are mostly based on unsound assumptions and glaring gaps in the variables considered. Perhaps the most serious of these gaps is the omission of the wider social context in which viewing takes place” (Edgar and Edgar, 1972, 602). Although many studies can show an influence of television on children, the authors argue that “not enough is known about the mechanisms involved” (Edgar and Edgar, 1972, 610) and they urge researchers to maintain an approach that incorporates “the total socialization context” (Edgar and Edgar, 1972). Alarm-raising about the effects of children and adolescents viewing violent television was only amplified when video-game playing became prominent among youth. Yet,

"Despite decades of research on the effects of violent video games on behavior the implications for the 'real-world' effects of violent video games remain highly contested" (Coyne and Stockdale, 2021, 11).

Unsurprisingly, the accessibility of information through the internet and the extensive use by youth of portable devices connected to the internet (George and Odgers, 2015) offers a seemingly infinite number of informal learning and teaching opportunities that have also given rise to concerns about young people's development. After interviewing 141 parents of young people, George and Odgers found that parents expressed certain common concerns regarding the influence of mobile technologies: "adolescents' safety (cyberbullying and online solicitation), social development (peer relationships, parent-child relationships, and identity development), cognitive performance, and sleep" (2015, 831). Even among the uninterrupted fearfulness about the influence of the latest media and technologies upon youth that has spanned generations, "few uniformly negative effects of new technologies on adolescent development have been documented. With some notable exceptions, many of the effects of new technologies on adolescents' lives have been, or are expected to be, positive" (George and Odgers, 2015, 848). Cruz's (2015) work with homeless LGBTQ youth and technology use found that SmartPhones facilitate social networks and the connections to street economies of exchange and bartering. Informal learning about social service networks, surveillance by police, and other necessities like food, shelter, and safer space are shared among communities of street youth through apps and other digital platforms. Generally, those studying the effects of the latest media and technologies emphasize that there are a multitude of factors outside of the technology that need to be considered before concluding whether young people are susceptible to harm from the informal learning and teaching opportunities the new technologies or media present (Boyd, 2014; Calvert et al., 2002; George and Odgers, 2015).

Youth as perpetual educators

Merely asserting that youth or children are naturally informal educators is perhaps stating the obvious, and there are blurred boundaries between formal and informal teaching as well as learning. For example, older siblings may engage in educational care-work as part of their "sibcare" obligations which likely places demands on older siblings to engage with informal and formal (school-related) teaching practices. This is particularly true if schools close because of a pandemic and subsequently implement synchronous distance learning during the school day while some parents are away from homes because they cannot work remotely. "Sibcare" is defined as "activities ranging from complete and independent full-time care of a child by an older child ... as well as simply 'keeping an eye out' for siblings" (Weisner and Gallimore, 1977, 169).

Often, educational caretaking (as a part of sibcare duties) is a transformative act that supports the family's future mobility (Lambek, 2011, 3). The pandemic that caused widespread school closures in 2020 and caused most schools to resort to forms of virtual remote learning likely increased the demand for siblings to do the educational caretaking in some families, especially in those that are under-resourced and/or in families where parents or guardians could not work remotely during school hours (Kirsch, 2020). Thus, this contextualization of sibcare illustrates both the commingling of informal and formal, portrays youth as educators, and includes a connection to the use of technology. This section continues with moments gathered from literature and online archives that illustrate how and when technologies and advances in communication have placed youth in positions as educators.

In this section, we are going to explore the use of technology by young people as a means or mode for informal teaching and learning with a particular focus on how technology has been used as a tool by youth to educate others during movements for social and political change. According to Buckingham, there is a long tradition in Cultural Studies of "empirical research on youth culture, dating back to the early 1970s, and it is mostly concerned with the ways in which young people appropriate cultural commodities and use them for their own devices" (2008, 4). Looking at young people is particularly useful because

communication between peers naturally increases in terms of both frequency and intensity during adolescence (Larson and Richards, 1989; Raffaelli and Duckett, 1989). Friendships and communication with peers are viewed as critical venues for the development of life-long social and relationship skills (for more discussion see for example Hartup, 1996; Hartup and Stevens, 1997; or Newcomb and Bagwell, 1995) and mobile technologies facilitate constant connectivity with peers and provide new tools for communication and informal learning.

George and Odgers (2015, 834).

Salamon notes that "Although students used media for activism in the 1960s, students now have more tools to quickly spread their messages widely and, in doing so, shape national conversations" (2018, 5). Additionally, Buckingham discusses how technology may be utilized as

... a force of liberation for young people—a means for them to reach past the constraining influence of their elders, and to create new, autonomous forms of communication and community. Far from corrupting the young, technology is seen to be creating a generation that is more open, more democratic, more creative, and more innovative than their parents' generation (2008, 13).

According to Mitra (2000, 2005; Mitra and Rana, 2001), minimally invasive education projects demonstrate that when youth are curious or motivated by something then informal learning will occur through self-instruction and peer-shared knowledge. Thus, amidst all of the adults' fears that accompany not knowing what new technologies and media might bring are also examples of

young people teaching and learning in innovative ways and examples of youth using technology to communicate on behalf of social and political change.

Youth, media, technology, and social movements

The Children's Crusade in May of 1963 in Birmingham, Alabama is (hopefully) a well-known example of how a local radio station was used for desegregation efforts when "black teenagers tuned into their favorite deejays [like 'Tall Paul' White] to hear coded messages telling them where and when to gather" (Boyle, 2005, 251) as they organized to protest Birmingham's violent discrimination and segregation. According to the Civil Rights Movement Archive, "A passion for freedom sweeps through Parker High [School] and the other Black schools of Birmingham and Bessemer, an emotional firestorm ignited by SCLC's young field workers, led by class presidents and prom queens, cheerleaders, and football players" (The Civil Rights Movement Archive, n.d.). Massive numbers of children (those under age eighteen and as young as seven) participated in acts of protest that made significant contributions to the eventual passage of the 1964 Civil Rights Act and radio was a significant instrument during their organizing.

A genealogy of struggles for equitable access to various forms of media or technology for formal and informal learning is also present among young people's social movements for justice. Buckingham cautions that "we should also not forget that access to these media—and the ways in which they are used—is partly dependent upon differences to do with factors such as social class, gender, and [race]" (2008, 5). The courage of the children in Birmingham during the civil rights movement educated and inspired direct action and protests in communities throughout the United States. In the couple of months following the children's actions in Birmingham, statisticians counted 758 protests in 186 cities (Birmingham The Children's Crusade, 1963). A perhaps less well-known story of resistance inspired by the Birmingham children is the bravery of the "Leesburg Stockade Girls." In Americus, Georgia, in July of 1963, fifteen girls ages 12–15 were viciously attacked by police and jailed for challenging segregation laws when they attempted to purchase movie tickets at the front entrance rather than in a back alley of the Martin Theater on Forsyth Street (Salahu-Din, 2019).

For forty-five days the girls were held twenty miles away from their families in a remote area of the woods in a Civil War era structure: the Leesburg Stockade. The authorities refused to disclose the girls' location to their families but a 21-year-old photographer for the Student Non-violent Coordinating Committee (SNCC), Danny Lyon, secretly found the girls, photographed them, and had the photos published in SNCC's newspaper, *The Student Voice*. The story along with the pictures the young photographer took spread, educated, and garnered national attention. The girls were finally released the same week as the bombing of the Sixteenth Street Baptist Church in Birmingham that killed four little girls who were in the basement of the church during Sunday morning services (Salahu-Din, 2019).

Although major media outlets had initially ignored what was happening in Birmingham until the children started their crusade in May of 1963, their courage led to almost 200 reporters covering events in the city. Images from Birmingham were broadcast nationally and globally via newspapers and televisions and inspired young people like the "Leesburg Stockade Girls." The horrific scenes of children in the supposed "land of the free" being attacked by police and police dogs as well as fire hoses also provoked disgust and condemnation of the United States from around the world. According to the online Civil Rights Movement Archives, "the disciplined courage of the protesters was itself a key component of the Movement's message" (n.d., n.p). Along a spectrum of informal and formal learning, through the media, and with available technologies and media communication to share their message, the youth created and spurred a movement that would educate others globally and compel United States officials to eventually respond to injustices (Civil Rights Movement Archives, Birmingham: The Children's Crusade, 1963).

Student journalists like those creating SNCC's newspaper, *The Student Voice*, use media as a key tool for informal teaching and learning, especially when mainstream media ignores students' actions and demands (Armstrong, 2018). According to Armstrong, student newspapers on college campuses "provide information about the world; fill the needs of the students; represent the college or university to the rest of the country; and act as an agent for change" (2018, 2). When students at Howard University in Washington, D.C. in 1968 demanded that the administration make the curriculum more relevant and asked for the inclusion of more student input concerning campus activities, it was the student newspaper, *The Hilltop*, that provided reporting on the protests that occurred until the administrative officials at Howard met some of the students' demands. One of the lead organizers, Adrienne Manns, was also the editor-in-chief of *The Hilltop*. Salamon (2018) explains that "Manns demonstrated that student journalists could draw on their experiences as activists, using media to tell alternative narratives, build public support and create change" (p. 3).

In more recent decades, mobile devices, computers, and the Internet have decreased barriers to communicating and organizing that young people may have encountered prior to the existence of this technology. According to Theocharis et al. "The internet transformed political action and long-established mechanisms of social movement organization, communication, and mobilization in various ways" (2015, 204). The latest technologies and communication forms "connect [young] people according to their interests and enable them to engage in all kinds of collective action, including learning. Schools no longer have the sole monopoly on learning most things" (Rhinegold, 2020, xv). These new modes and means for informal learning and rapid communication transform potential political and civic engagement for youth by providing convenient educational and mobilization tools. For example, research has shown that content circulation via Twitter related to Occupy Wall Street in 2011 was a highly significant informal learning medium for information distribution that enabled the widespread creation and mobilization of counterpublics (Penney and Dadas, 2014; Birmingham The Children's Crusade, 1963).

Since the late 1990s, the Internet has been considered as having the potential for “reinvigorat[ing] youth participation by creating new opportunities for engagement and allowing established forms of offline participation to be carried out with substantially lower costs (Bimber et al., 2012; Earl and Kimport, 2011)” (Theocharis and Quintelier, 2016, 820). In recent decades, social media has frequently been utilized for collective action. There is a significant amount of research exploring how social media use can mobilize activism although it is not always focused on youth as organizers and educators (Wilkins et al., 2019). Movements such as #BlackLivesMatter, the DREAMer movement, Occupy Wall Street, The People’s Climate March, March for Our Lives, the Arab Spring, and others illustrate how informal learning and teaching using digital tools, and internet-worked engagement and civic participation are contributing to the ways young people organize for social change in the 21st century.

For example, Beltrán (2021) investigates “how young immigration reform activists, ‘DREAMERS,’ have brought new, social-media-powered strategies to the fore and, in important ways, have displaced older models of political participation” (p. 50). Another example involves digital video production through a social justice framework. Researchers working with youth who were tasked with creating the videos describe this approach as “intentional, informal learning” (Goldman et al., 2008, 201). According to the authors of this video production study, the youth used technologies to communicate their social message, and “The collaborative video production was a medium through which young people developed ‘critical literacy’ and recognized their potentials for leadership, for peer-to peer-mentorship, and for social change” (Goldman et al., 2008, 202).

Spanning two continents, Tom examines “how youth from specific communities access knowledge which is denied to them in schools, give revolutionary voice to their realities, and broadcast perspectives on race, place and belonging” through the production of hip-hop which “becomes a medium to create alternative educational projects” (2016, 711). Again, the medium is the message and “Hip-hop occupies a space where educators/producers of rap and their students [consumers] can reclaim themselves and their communities from colonial meta-narratives of knowledge” (Tom, 2016, 730). The author of this study explains that the motivation to study youth as informal educators was partially inspired by Michael Apple’s observations that “schools preserve inequalities by affirming the cultural legitimacy of a particular knowledge (i.e., colonial knowledge) as legitimate at the level of curriculum, pedagogy and forms of classroom evaluation which may disqualify or marginalize other communities’ knowledges, histories, languages and cultures” (Tom, 2016, 712). Thus, hip-hop productions, performances, and the digital dissemination of the music by the youth Tom studied is an informal and alternative educational project with young people serving as educators while creating and leading.

#Hashtag

In more distant decades of the past, young people like those currently utilizing or directing the discourse on social media to educate others “might not have had the opportunity to participate in the public sphere” (Carney, 2016, 195–196) so widely. One strategy youth employ to garner attention and a following on social media is the use of hashtags (a word or phrase preceded by the pound # symbol). A hashtag on social media also “serves as an indication (for users and algorithms) that a piece of content relates to a specific topic or belongs to a category” (Olafson, 2021, n.p.), and attaching hashtags to social media posts indicates that a user is openly engaging in conversations associated with them. Hashtags are an informal method of trying to grow conversation and education around a topic and some young people deliberately use hashtags to promote information with social justice aims, to spread their messages, and to solicit new followers on social media (Anderson, 2016).

In response to the acquittal of Trayvon Martin’s murderer in 2013, and to uplift the voices of those often suppressed or missing in social movements (including black women, LGBTQ+, transgender, disabled, and undocumented Black peoples), the Black Lives Matter (BLM) movement was developed (Cox, 2017) and became widely recognized because of the hashtag “#BlackLivesMatter” on social media (Mundt et al., 2018). Between its initial appearance in July of 2013 and through the month of March in 2016, “the hashtag #BlackLivesMatter appeared on Twitter almost 11.8 million times” (Anderson, 2016, n.p.). On Twitter’s 10-year anniversary, the site published a list of the most used hashtags related to social causes with #Ferguson being the most used social-issue hashtag and #BlackLivesMatter being third (Anderson, 2016). Through an analysis of tweets in the early stages of the BLM movement, Wilkins et al. found that “social media can fulfill rhetorical functions for social movements in terms of dening the scope and agenda of the movement; for example, the issues with which the movement is concerned, who is or is not part of the movement, and the specific outgroups whose behavior the movement seeks to change” (Wilkins et al., 2019, 787).

BLM addresses many social justice concerns including women having control over their bodies, fighting against the Dakota Access Pipeline, for people who are undocumented and transgender equality, and stopping violence against Black, Indigenous, and additional People of Color (BIPOC) (Cox, 2017; Carney, 2016). On its website, Black Lives Matter describes itself as a movement to fight for freedom, liberation, justice, and the eradication of white supremacy. Scholarly analyses of BLM’s strategies and mission illustrate how social media as a platform provides “for continued involvement and reflection around issues related to race and policing and highlight[s] its significance for developing common understandings of ideology and a shared sense of movement identity” (Mundt et al., 2018, 2).

Social media allows users to actively engage and shape discussions, which also offers youth an opportunity to contest dominant ideologies and educate. Yet, as others have recognized, often students motivated to educate others on behalf of social justice concerns are “no more passionate than the young people of, say, the Black Lives Matter movement, [yet] it [becomes] clear that the political establishment [is] going to receive them a different way” (Witt, 2018b, n.p.). While social media sites might “serve as platforms for building collective identity and creating a cohesive movement ideology that might mobilize new participants or facilitate coalition-building” (Mundt et al., 2018, 3), scholars also note that there are possible limitations resulting from platform constraints or broader social, economic, and cultural factors (Freelon et al., 2016; Hintz, 2016; Poell, 2014; Tufekci, 2018; Youmans and York, 2012).

Another movement that developed and expanded using a hashtag involves a group of students who survived a 2018 mass shooting at Marjorie Stoneman Douglas High School (MSDHS) in Parkland, Florida. The students organized through their trauma and grief via the internet with the hashtags #MarchforourLives and #NeverAgain. Bent argues that “From their intentional disruption of thoughts and prayers rhetoric to their organized political action, March for Our Lives (MFOL) student-activists cut through business-as-usual politics” (2019, 56). Shortly after the shootings and through a digital and social media campaign, the activist students organized national school walkouts and a main rally in Washington, D.C. with approximately 800 rallies around the country occurring at the same time (Zoller and Casteel, 2021).

At the rallies and through various media outlets including social media, the surviving young people from MSDHS educated and shamed the nation. Alfonso Calderon, a co-founder of Never Again, explains how he is certain fellow Never Again founder, Emma Gonzalez, is accurately educating the world with her widely shared “B.S.” speech that was passionately and tearfully given at the D.C. rally, as well as through social media and other speaking engagements. Calderon states “Even before this happened, we already knew all the facts. We already knew everything ... if she were to say something that was non-factual, you know she would be highly scrutinized by literally everybody, including the President [Trump]” (Witt, 2018, n.p.). Additionally, perhaps benefiting from experience as a reporter for the high school newspaper, Never Again co-founder, David Hogg, appeared confidently asserting facts on national news broadcasts and demanding action from elected officials.

The internet and digital media communication platforms as well as other forms of media create new opportunities for youths to act as educators, to develop collective movements’ identities and to mobilize resources “for strengthening movements and creating coalitions (Earl and Kimport, 2014)” (Mundt et al., 2018, 12). This entry attempts to highlight some of the ways youth have utilized technologies and communications for informal teaching and learning during social justice movements. However, by illuminating youth as the educators, leaders, and main participants during various social movements, our intention is not to shift the focus away from the education and messages that the movements provide (Applegarth, 2020), nor are we gesturing toward the usefulness of comparing movements that emerge in different spatial, temporal, political, and cultural contexts. Central to the informal teaching and learning highlighted in this entry are significant social justice issues with embodied consequences such as gun violence, racism, coloniality, ableism, sexism, homophobia, and more. Thus, although the focus of this writing is on how youth engage with technologies and communication mediums for social movement making and informal teaching and learning, the information and demands from the movements should remain the priority. Furthermore, the internet “is a rapidly developing arena for political action, and there is a great deal of work still to be done to understand the relationship between [youths’] online and offline engagement” (Keating and Melis, 2017, 891).

Conclusion

Youth are educators from the moment they enter this world, even if at times they are only teaching themselves. Thus, to narrow the scope of this entry, we focused on how young people use technologies and forms of media communication to educate others about a need for social and/or political change. We began by highlighting scholarly discussions related to informal and formal learning. Then, we traced some of the responses of fearful adults regarding young people’s informal learning when “new” technologies and communication mediums emerge (i.e., the radio, movies, television, video games, the internet). Next, we provided examples of ways young people have utilized technologies and communication mediums available to them for informal learning while acting as youth educators and activists laboring on behalf of social justice causes. Through this tracing, we hope to elucidate how the apprehension or fears about youth engaging in deleterious informal learning via new technologies might be unwarranted or unnecessarily elevated at times. Furthermore, we agree with scholars such as Buckingham who assert that rather than corrupting young people, we can witness youth utilizing technology and media for democratic participation and social justice aims.

We end by suggesting that what we want informal or formal learning or education to do or be is the same as asking what we want society to do or be. We imagine a society that honors knowledges and ecologies across species and elements—a society that conceives of knowledge as non-hierarchical, multiple, malleable, and contingent. We also wonder if the time spent by adults who were panicking about new technologies might have been better used for collaborating with youth (like those who engage in participatory action research) to figure out how to make our world more just—or to create an entirely different and just world. Miles Horton, an educator who uses a form popular education with young people who want to educate for social change, argues in his autobiography, *A Long Haul*, that

If we are to have a democratic society, people must find or invent new channels through which decisions can be made ... the problem is not that people will make irresponsible or wrong decisions. It is, rather, to convince people who have been ignored or excluded in the past that their involvement will have meaning and that their ideas will be respected. The danger is not too much, but too little participation (1990, 134).

Exclusion and ignorance are not the foundations of liberation or a just society. Young people like those referred to in this entry are teachers and learners and seek to have meaningful engagements. It might be better for all of us if we find ways to include, encourage, and support youth working on behalf of justice.

Moving forward our questions are not related to whether informal and formal learning are separate, nor are our inquiries the result of an accumulation of speculations about the potential harms of new technologies—there is too much actual suffering happening in the immediate moment that requires our attention. Instead, we ask how we might use technologies and media to collectively collaborate for the creation of informal and formal learning and education that center on limiting fear and amplifying love and liberation? With youth as a resource, how, when, and where can we act transformatively with abolitionist intentions for freedom and justice?

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Anderson, M., 2016, Aug. 15. The Hashtag #BlackLivesMatter Emerges: Social Activism on Twitter. Social Media Conversations About Race, PEW Research Center. Retrieved from <https://www.pewresearch.org/internet/2016/08/15/the-hashtag-blacklivesmatter-emerges-social-activism-on-twitter/>.
- Anderson, C.A., Shibuya, A., Ihori, N., Swing, E.L., Bushman, B.J., Sakamoto, A., Rothstein, H.R., Saleem, M., 2010. Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin* 136 (2), 151–173. <https://doi.org/10.1037/a0018251>.
- Appelgarth, R., 2020. News that isn't new: March for our lives and media mobilization of historical precedent. *Rhetor. Rev.* 39 (2), 159–173. <https://doi.org/10.1080/07350198.2020.1727102>.
- Armstrong, K.D., 2018. *How Student Journalists Report Campus Unrest*. Lexington Books, Lanham, MD.
- Bautista-Sancho, 2012. Los cambios en la web 2.0: una nueva sociabilidad. *Estudios Sobre El Mensaje Periodístico*, 18. https://doi.org/10.5209/rev_ESMP.2012.v18.40917
- Birmingham—The Children's Crusade (April–May), 1963. The Civil Rights Movement Archive. <https://www.crmvet.org/tim/timhis63.htm#1963bhcc>.
- Beltran, C., 2021. "Undocumented, unafraid, and unapologetic": DREAM activists, immigrant politics, and the queering of democracy. In: Allen, D., Light, J. (Eds.), *From Voice to Influence*. University of Chicago Press, Chicago. <https://doi.org/10.7208/9780226262437-toc>, 180–104.
- Bent, E., 2019. Unfiltered and unapologetic: March for our lives and the political boundaries of age. *Jeunesse* 11 (2), 55–73. <https://doi.org/10.1353/jeu.2019.0017>.
- Boyd, D., 2014. *It's Complicated: The Social Lives of Networked Teens*. Yale University Press, New Haven, CT.
- Boyle, K., 2005. Radio, race, and the re-writing of civil rights. *Rev. Am. Hist.* 33 (2), 249–253. <http://www.jstor.org/stable/30031469>. (Accessed 15 June 2021).
- Buckingham, D., 2008. *Youth, Identity, and Digital Media*. MIT Press, Cambridge, MA.
- Campellone, J., Turley, R.K., (n.d.). Understanding the Teen Brain. Health Encyclopedia. University of Rochester Medical Center. Retrieved June 13, 2021 from <https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=1&ContentID=3051>.
- Callanan, M., Cervantes, C., Loomis, M., 2011. Informal learning. *WIREs Cognit. Sci.* 2, 646–655. <https://doi.org/10.1002/wcs.143>.
- Calvert, S.L., Jordan, A.B., Cocking, R.R. (Eds.), 2002. *Children in the Digital Age: Influences of Electronic Media on Development*. Praeger Publishers/Greenwood Publishing Group.
- Carney, N., 2016. All lives matter, but so does race: black lives matter and the evolving role of social media. *Humanity Soc.* 40 (2), 180–199. <https://doi.org/10.1177/0160597616643868>.
- Cartwright, I., 2012. CH 9: informal education in compulsory schooling in the UK: humanising moments, utopian spaces? In: Kraftl, P., Horton, J., Tucker, F.J. (Eds.), *Critical Geographies of Childhood and Youth: Contemporary Policy and Practice*. Policy Press.
- Clayton, J., 2016. *Uproot: Travels in 21st-Century Music and Digital Culture*. FSG Originals.
- Cohen, E.H., 2007. Researching informal education: a preliminary mapping. *Bull. Sociol. Methodol.* 93 (1), 70–88. <https://doi.org/10.1177/075910630709300106>.
- Coombs, P.H., Ahmed, M., 1974. *Attacking Rural Poverty: How Nonformal Education Can Help*. Johns Hopkins University Press.
- Cox, J.M., 2017. The source of a movement: making the case for social media as an informational source using Black Lives Matter. *Ethn. Racial Stud.* 40 (11), 1847–1854. <https://doi.org/10.1080/01419870.2017.1334935>.
- Coyne, S.M., Stockdale, L., 2021. Growing up with grand theft auto: a 10-year study of longitudinal growth of violent video game play in adolescents. *Cyberpsychol. Behav. Soc. Netw.* 24 (1), 11–16. <https://doi.org/10.1089/cyber.2020.0049>.
- Cressey, P.G., 1934. The motion picture as informal education. *J. Educ. Sociol.* 7 (8), 504–515. <https://www.jstor.org/stable/2961536>.
- Cruz, C., 2015. When does resistance begin. In: Conchas, G. (Ed.), *Cracks in the Schoolyard—Confronting Latino Educational Inequality*. Teacher's College Press, pp. 289–320.
- Edgar, P.M., Edgar, D.E., 1972. Television violence and socialization theory. *Publ. Opin. Q.* 35 (4), 608–612.
- Erstad, O., 2012. The learning lives of digital youth—beyond the formal and informal. *Oxf. Rev. Educ.* 38 (1), 25–43.
- Ferguson, Kilburn, J., 2010. Much Ado About Nothing: The Misestimation and Overinterpretation of Violent Video Game Effects in Eastern and Western Nations: Comment on Anderson et al. (2010). *Psychological Bulletin* 136 (2), 174–178. <https://doi.org/10.1037/a0018566>.
- Foley, G., 1999. *Learning in Social Action: A Contribution to Understanding Informal Education*. Global Perspectives on Adult Education and Training. St. Martin's Press Inc.
- Fox, J., Ralston, R., 2016. Queer identity online: informal learning and teaching experiences of LGBTQ individuals on social media. *Comput. Hum. Behav.* 65, 635–642. <https://doi.org/10.1016/j.chb.2016.06.009>.
- Freelon, D., McIlwain, C., Clark, M., 2016. Beyond the Hashtags: #Ferguson, #Blacklivesmatter, and the Online Struggle for Offline Justice. Retrieved from <https://ssrn.com/abstract=2747066>.
- Funk, C., 1998. The art in America radio programs, 1934–1935. *Stud. Art Educ.* 40 (1), 31–45. <https://doi.org/10.1080/00393541.1998.11650045>.
- Funk, C., 2016. *A Quick and Dirty Guide to Art, Music, and Culture*. The Ohio State University. <https://ohiostate.pressbooks.pub/artandmusicbiographies/>.
- Geerdts, M.S., Van de Walle, G.A., LoBue, V., 2015. Parent-child conversations about animals in informal learning environments. *Visit. Stud.* 18 (1), 39–63.
- George, M.J., Odgers, C.L., 2015. Seven fears and the science of how mobile technologies may be influencing adolescents in the digital age. *Perspect. Psychol. Sci.* 10 (6), 832–851.
- Goldman, S., Booker, A., McDermott, M., 2008. Mixing the digital, social, and cultural: learning, identity, and agency in youth participation. Youth, identity, and digital media. In: Buckingham, D., The John, D., Catherine, T. (Eds.), *MacArthur Foundation Series on Digital Media and Learning*. The MIT Press, Cambridge, MA, pp. 185–206. <https://doi.org/10.1162/dmal.9780262524834.185>.
- Hedges, H., 2021. Contemporary principles to lead understandings of children's learning: synthesizing Vygotsky, Rogoff, Wells and Lindfors. *Early Child Dev. Care.* 191 (7–8), 1056–1065.
- Hartup, 1996. The Company They Keep: Friendships and Their Developmental Significance. *Child Development* 67 (1), 1–13. <https://doi.org/10.2307/1>.
- Hintz, A., 2016. Restricting digital sites of dissent: commercial social media and free expression. *Crit. Discourse Stud.* 13, 325–340.
- Holloway, S.L., Valentine, G., 2000. Spatiality and the new social studies of childhood. *Sociology* 34 (4), 763–783.
- Horton, M., 1990. *The Long Haul: An Autobiography*. Anchor Books, Doubleday, New York.
- Kauko, M., Wilkinson, J., 2018. Praxis and language: teaching newly arrived migrant children to "live well in a world worth living in". *TESOL in Context* 27 (2).
- Keating, A., Melis, G., 2017. Social media and youth political engagement: preaching to the converted or providing a new voice for youth? *Br. J. Politics Int. Relat.* 19 (4), 877–894.
- Kirsch, Z., April 10, 2020. When Siblings Become Teachers: It's Not Just Parents Who Find Themselves Thrust Into the Demanding Role of at-Home Educators. 174. Retrieved online from <https://www.the74million.org/article/when-siblings-become-teachers-its-not-just-parents-who-find-themselves-thrust-into-the-demanding-role-of-at-home-educators/>.

- La Belle, T.J., 1982. Formal, nonformal, and informal education: a holistic perspective on lifelong learning. *Int. Rev. Educ.* XXVII, 159–175.
- Lambek, M., 2011. Kinship as gift and theft: acts of succession in Mayotte and ancient Israel. *Am. Ethnol.* 38 (1), 2–16.
- Larson, Richards, M., 1989. Introduction: the changing life space of early adolescence. *Journal of Youth and Adolescence* 18 (6), 501–509. <https://doi.org/10.1007/BF02139070>.
- Livingstone, D., 2006. Informal learning: conceptual distinctions and preliminary findings. *Counterpoints* 249, 203–227. <http://www.jstor.org/stable/42979596>. (Accessed 14 June 2021).
- Longstaff, 1936. Preliminary results of a study of mothers' opinions of children's radio programs. *Journal of Applied Psychology* 20 (3), 416–419. <https://doi.org/10.1037/h0061756>.
- Mills, S., Kraftl (Eds.), 2014. *Informal Education, Childhood and Youth*. Palgrave Macmillan.
- Mitra, S., Rana, V., 2001. Children and the Internet: experiments with minimally invasive education in India. *Br. J. Educ. Technol.* 32 (2), 221–232.
- Mitra, S., 2000. Minimally Invasive Education for Mass Computer Literacy. Presented at the CRIDALA 2000 Conference in Hong Kong, June 21–25.
- Mitra, S., 2005. Self organising systems for mass computer literacy: findings from the “hole in the wall” experiments. *Int. J. Dev. Issues* 4 (1), 71–81. <https://doi.org/10.1108/eb045849>.
- Mundt, M., Ross, K., Burnett, C.M., 2018. Scaling Social Movements Through Social Media: The Case of Black Lives Matter. *Social Media + Society*. <https://doi.org/10.1177/2056305118807911>.
- Olafsun, K., June 16, 2021. How to Use Hashtags in 2021: A Quick and Simple Guide for Every Network. Retrieved from. <https://blog.hootsuite.com/how-to-use-hashtags/>.
- Penney, J., Dadas, C., 2014. (Re)tweeting in the service of protest: digital composition and circulation in the occupy wall street movement. *New Media Soc.* 16, 74–90. <https://doi.org/10.1177/1461444813479593>.
- Poell, T., 2014. Social media activism and state censorship. In: Trotter, D., Fuchs, C. (Eds.), *Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in the Age of Facebook, Twitter and YouTube*. Routledge, London, England, pp. 189–206.
- Putnam Parker, E., 1950. Overcoming undesirable influences of informal education. *J. Geogr.* 49 (3), 89–99. <https://doi.org/10.1080/00221345008982449>.
- Raffaelli, Duckett, E., 1989. We were just talking - conversations in early adolescence. *Journal of Youth and Adolescence* 18 (6), 567–582. <https://doi.org/10.1007/BF02139074>.
- Rogers, A., 1991. *British Journal of Educational Studies*, vol. 39, pp. 105–107. <https://doi.org/10.2307/3120884>.
- Rogoff, B., Callanan, M., Gutierrez, K.D., Erickson, F., 2016. The organization of informal learning. *Rev. Res. Educ.* 40, 356–401.
- Salahu-Din, T., July 26, 2019. Hidden Herstory: The Leesburg Stockade Girls. National Museum of African American History and Culture. <https://nmaahc.si.edu/blog-post/hidden-herstory-leesburg-stockade-girls>.
- Salamon, E., March 2018. March for Our Lives Awakens the Spirit of Student and Media Activism of the 1960s. *The Conversation*. Retrieved from. <https://theconversation.com/march-for-our-lives-awakens-the-spirit-of-student-and-media-activism-of-the-1960s-93713>.
- Scribner, S., Cole, M., 1973. Cognitive consequences of formal and informal education. *Science* 182 (4112), 553–559. Stable URL. <http://www.jstor.org/stable/1737765>.
- Thrasher, F., 1934. Social backgrounds and informal education. *J. Educ. Sociol.* 7 (8), 470–484. <https://doi.org/10.2307/2961533>.
- Theocharis, Y., Quintelier, E., 2016. Stimulating citizenship or expanding entertainment? The effect of Facebook on adolescent participation. *New Media Soc.* 18 (5), 817–836. <https://doi.org/10.1177/1461444814549006>.
- Theocharis, Y., Lowe, W., van Deth, J.W., García-Albacete, G., 2015. Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Inf. Commun. Soc.* 18 (2), 202–220. <https://doi.org/10.1080/1369118X.2014.948035>.
- Tom, M.N., 2016. Contesting history and pursuing “other” knowledge: a study of hip-hop and non-formal education among Native American youth in San Francisco and black Portuguese youth in Lisbon. *Int. Rev. Educ.* 62, 711–731. <https://doi.org/10.1007/s11159-016-9598-x>.
- Tufekci, Z., May 19, 2018. Facebook's surveillance machine. *The New York Times*.
- Vizcaino-Verdu, A., Contreras-Pulido, P., Guzman-Franco, M.-D., 2019. Reading and informal learning trends on YouTube: the booktuber. *Media Educ. Res. J.* 59 (XXVII), 93–101.
- Weisner, T.S., Gallimore, R., 1977. My brother's keeper: child and sibling caretaking. *Curr. Anthropol.* 18 (2), 169–190.
- Wilkins, D.J., Livingstone, A.G., Levine, M., 2019. Whose tweets? The rhetorical functions of social media use in developing the Black Lives Matter movement. *Br. J. Soc. Psychol.* 58, 786–805. <https://doi.org/10.1111/bjso.12318>.
- Willett, R., 2008. Consumer citizens online: structure, agency, and gender in online participation. Youth, identity, and digital media. In: Buckingham, D., The John, D., Catherine, T. (Eds.), *MacArthur Foundation Series on Digital Media and Learning*. The MIT Press, Cambridge, MA, pp. 49–70. <https://doi.org/10.1162/dmal.9780262524834.049>.
- Williams, R., 1981/1989. *What I Came to Say*. Hutchinson Radius, London.
- Youmans, W.L., York, J.C., 2012. Social media and the activist toolkit: user agreements, corporate interests, and the information infrastructure of modern social movements. *J. Commun.* 62, 315–329.
- Zoller, H.M., Casteel, D., 2021. #March for our lives: health activism, diagnostic framing, gun control, and the gun industry. *Health Commun.* <https://doi.org/10.1080/10410236.2020.1871167>.

Further reading

- Cammarota, J., Fine, M. (Eds.), 2016. *Revolutionizing Education: Youth Participatory Action Research in Motion*. Routledge, CRC Press.
- Noguera, P., Cammarota, J., Ginwright, S., 2006. *Beyond Resistance! Youth Activism and Community Change: New Democratic Possibilities for Practice and Policy for America's Youth*. Routledge, CRC Press.
- YPAR Hub, 2015. YPAR Hub Resources May Be Reprinted or Adapted for Nonprofit Purposes by K-12 Schools, Colleges, or Universities, provided the Source is Accurately Quoted and Duly Credited. <http://yparhub.berkeley.edu/>.

Shadow education

Mark Bray^{a,b}, ^a Centre for International Research in Supplementary Tutoring (CIRIST), Faculty of Education, East China Normal University, Shanghai, China; and ^b The University of Hong Kong, Pokfulam, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	346
Definitions and scope	346
Scale and nature of shadow education	346
Drivers of demand	347
Diversity in supply	348
Social and political implications	349
References	349

Introduction

Private supplementary tutoring, operating in parallel to regular schooling, has a history as long as that of schooling itself. However, only in recent decades has it become a major phenomenon and attracted significant attention.

Publications in English and in some other languages commonly use the term ‘shadow education’ on the grounds that the curriculum of tutoring changes as the curriculum of regular schooling changes, and that as the school system grows so does its shadow. The metaphor was first used in Malaysia, Singapore and Japan in the early 1990s, and then the first global comparative study using the metaphor was published in 1999 by UNESCO’s International Institute for Educational Planning (Bray, 1999). A sequel to this book was published a decade later (Bray, 2009), and has been followed by various regional studies (e.g. Bray, 2021a,b; Bray and Hajar, 2023; Bray and Lykins, 2012). Other writings with a global focus include Aurini et al. (2013), Park et al. (2016), and Enrich (2021).

This entry to the encyclopedia commences with definitions and scope, and then moves to the scale and nature of shadow education. The following pair of sections focuses on drivers of demand and of supply, and the entry concludes with remarks about social and political implications.

Definitions and scope

The definition of shadow education presented by Bray (1999: 20) has been widely followed in the literature (e.g. Aslam and Atherton, 2014; Tan, 2009; Zhang and Yamato, 2018) and is employed in this paper. It has three components, namely:

- *Supplementation*. The focus is on subjects that are already covered in school, and does not include for example language classes for minority children whose families are anxious that new generations retain competence in languages not taught in their schools.
- *Privateness*. The focus is on tutoring provided in exchange for a fee. It does not include free tutoring by relatives or by teachers as part of their school duties.
- *Academic*. The focus is on subjects taught in schools, particularly languages, mathematics and other examinable subjects. It does not include musical, artistic or sporting skills that are learned primarily for pleasure and/or more rounded personal development.

This entry is concerned with the subjects learned in primary and secondary schooling. Shadow education does also exist at pre-primary and post-secondary levels. The phenomenon at these levels also deserves attention but it is excluded here in order to permit greater depth.

The forms of shadow education may be varied. Some tutoring is provided one-to-one in the homes of either the tutors or the clients. Other tutoring is in small groups, in large classes, or even huge lecture theaters with video screens to cater for overflows. Tutoring over the internet has become increasingly common, both on a one-to-one and one-to-group basis. Internet tutoring may include the dual-tutor model in which the main tutor is at a distance while a teaching assistant works with students in their groups, to make sure that they have understood and to follow up on activities.

Scale and nature of shadow education

Historically, shadow education has been most prominent in East Asia. South Korea has long been famous for its tutorial institutions called *hagwons* (Kim, 2016), and in 2019 according to official statistics 83.5% of elementary-school, 71.4% of middle-school, and

67.9% of general-high-school students were receiving private supplementary support (KOSIS, 2020). Neighboring Japan is equally famous for its *juku* (Enrich, 2018; Sato, 2012), and in recent decades shadow education has been very prominent in Hong Kong, Singapore and Taiwan (Bray and Lykins, 2012). Mainland China had a slower start, but until the promulgation of fierce regulations to dampen the industry (China, 2021, 2022) had caught up with and even overtaken other parts of the region (Feng, 2021; Zhang and Bray, 2021).

Other world regions also have countries in which the phenomenon is particularly visible. In Europe, Greece has long been the most prominent with its tutorial institutions known as *frontistiria* (Tsiloglu, 2005). In North Africa the counterpart country is Egypt (Hartmann, 2013); and elsewhere in Africa, Mauritius stands out for its long history of the phenomenon (Foondun, 1992). In general, however, shadow education across the world has not been of major significance until recent years. This is among the reasons for paucity of statistical information, especially data collected on a common yardstick across national boundaries.

One of the few cross-national statistical surveys of shadow education was undertaken by the Southern and Eastern Africa Consortium for Measuring Educational Quality (SACMEQ). Researchers in 15 education systems surveyed Grade 6 students during 2007 (SACMEQ, 2010). They found that shadow education was most prominent in Mauritius, with 74.6% of students receiving it, followed by Kenya (46.3%) and Uganda (25.1%). At the other end of the scale were Eswatini (1.1%), Lesotho (2.5%), Namibia (2.9%) and South Africa (4.0%). However, a repeat survey in 2013 showed significant increase. In Mauritius the proportion had risen to 81.4% (Dwarkan, 2017, p. 37), in Namibia it had risen to 5.8% (Shigwedha et al., 2015, pp. 35–36), and in South Africa it had risen to 29.1% (Chetty et al., 2017, p. 16).

Other indicative data can be derived from the Programme for International Student Assessment (PISA) operated under the auspices of the Organisation for Economic Co-operation and Development (OECD) (see e.g. Byun et al., 2018). These surveys have been conducted every three years since 2000. However, from the perspective of shadow education as defined here the surveys have limited usefulness first because of ambiguities in the questions, second because of problems with translations of the questions, and most of all because they have not clearly asked whether the tutoring was fee-free or fee-charging (Bray et al., 2020).

Even when they are available, national statistics on enrollment rates require much elaboration, in particular to include information on the types of shadow education, on the modes of delivery, and on the identities of the tutors and tutees. For example with reference to Egypt, Sieverding et al. (2019, pp. 567–568) distinguished between private lessons (*durus khosouseya*), on a one-to-one or small-group basis outside schools, and “fortification” lessons offered in schools but after hours for a fee. Data cited by Sieverding et al. indicated that 53% of primary, lower secondary and general upper secondary students received private lessons, most of which were in the students’ or teachers’ homes but some of which were in tutorial centers operated by companies. The in-school fortification lessons, received by about 10% of students across levels, were operated in larger classes and were commonly perceived to be inferior to private lessons.

Fig. 1 elaborates on possible modes of shadow education. It shows not only group sizes but also the media through which tutoring may be provided, several categories of providers, curricula, purposes, and levels of education.

Further variations are evident in the seasons and durations of tutoring. In Angola, within a sample of Grade 11 and 12 students surveyed by Chionga (2018), 28.3% received tutoring all the year round, 23.3% during parts of the school year, and 48.3% during the examination season. In Kenya, private tutoring has commonly been delivered not only after the school day and at weekends but also during school holidays (Mogaka et al., 2016).

Drivers of demand

Throughout the world, the greatest driver of demand for shadow education is social competition. This is especially evident at the season of watershed examinations that determine opportunities for further study and/or employment. Families increasingly consider schooling to be inadequate by itself to meet their aspirations, and view shadow education as a supplementary component to secure advantage. Then, as more and more families secure the supplementation, others have to join in or face being left behind.

To some extent, this pattern reflects global expansion of schooling in line with advocacy by UNESCO and others in the Education for All (EFA) movement (see e.g. UNESCO, 2000). During earlier periods of history, average levels of schooling were modest and in some societies even completion of lower secondary or even primary education distinguished the holders from the crowd. But as first primary and then lower secondary education became universal, upper secondary and then university education were needed to secure distinction from the crowd. Indeed some societies have reached the point at which a university degree is not enough and postgraduate qualifications are necessary.

These patterns have two implications. First, families that in previous eras would have left school, not aspiring to remain in the system because few possibilities were on the horizon, now find that they can compete with others and seek shadow education to assist in that competition. Second, because unit costs are greater at higher levels of education, government budgets are stretched. This may impact on the quality of education, causing frustrated families to seek private tutoring to bridge gaps (Bray, 2017).

Insofar as education is a positional good (Hollis, 1982), a major aim among parents is to secure better grades than other parents and then access to more prestigious schools and universities in order to enhance subsequent careers and life chances. Yet once one group of parents does this, others feel obliged to do the same until patterns build up to what Lampl (2017, p.2)

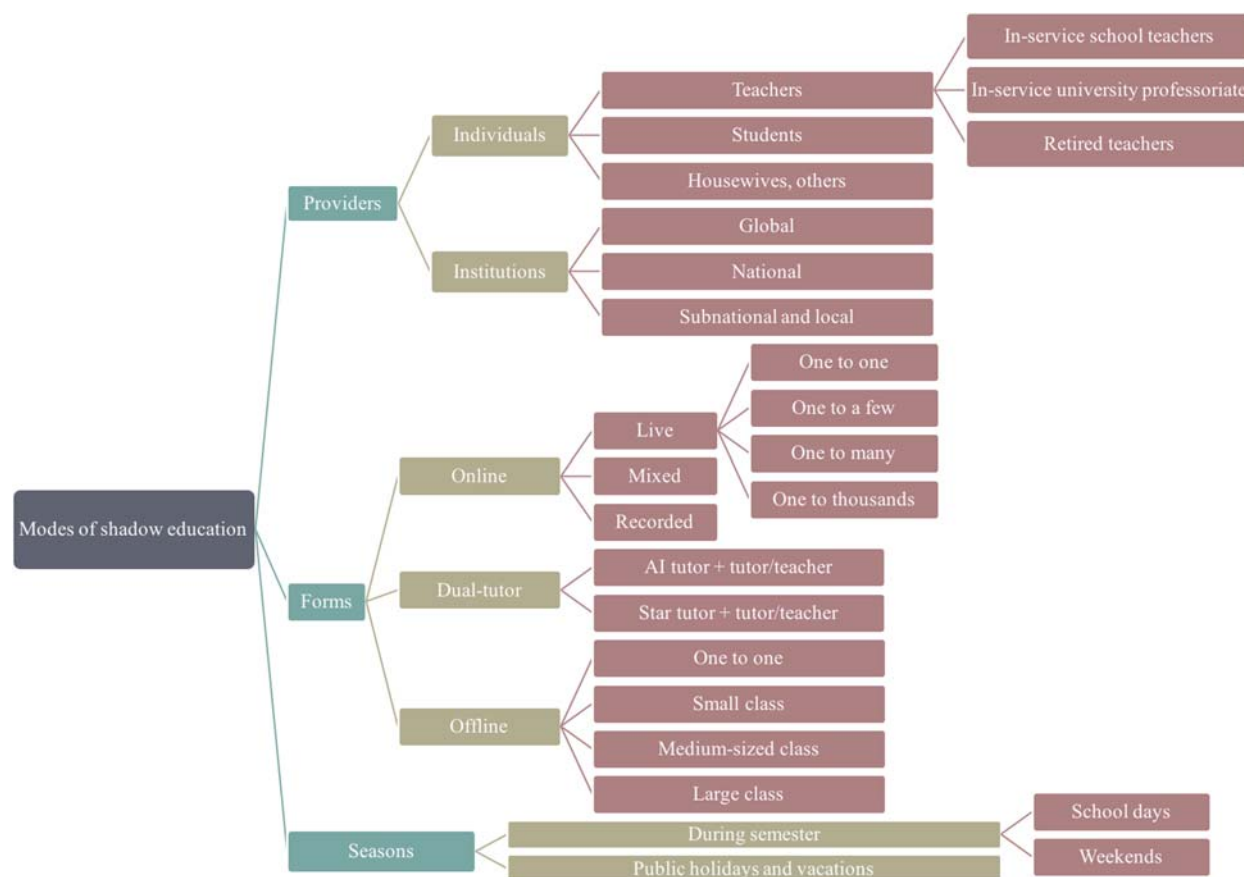


Fig. 1 Diversity of modes of shadow education delivery. Reproduced with permission from Zhang and Bray (2020, p. 328).

has called an educational ‘arms race’. This pattern to a significant extent underpins the expansion of private tutoring in England and Wales. Holloway and Kirby (2020, p. 171), presented data from Sutton Trust surveys of students aged 11–16. When asked in 2005 whether they had ever received paid private tutoring, 18.3% replied affirmatively. Eleven years later, a repeat survey found that 26.2% replied affirmatively, and that proportions were 43.8% in London.

The EFA movement has also contributed to the growth of shadow education insofar as the expansion of schooling was only achieved at the expense of qualitative decline. Many families perceive schooling to have inadequate quality, particularly when classes are large and teachers are not motivated.

Diversity in supply

In the commercial sector, much diversity is evident in the sizes, types and market reach of companies. Around the world, the majority of companies are small and serve their immediate localities. In China, for example, one report (MobData Research Institute cited by Zhang, 2019: 27) grouped industry operators into three categories: (i) nationwide companies operating with branches in first- and second-tier cities on the way to reaching third-tier cities, (ii) region-wide companies operating across one province or municipality, with branches mainly in first-, second- and third-tier cities, and (iii) smaller enterprises which ranged from one-person workshops to medium-sized companies. The first category at that time accounted for less than 5% of the market share, while the third category accounted for the majority.

In contrast to these local and national companies are multinationals such as Kumon which is headquartered in Japan. Kumon operates on a franchise model and has 4.3 million subject enrolments (i.e. if one person enrolled in three subjects, that would be three subject enrolments) in 50 countries (www.kumon.org). Kip McGrath is headquartered in Australia and operates through franchises in four continents. Alongside them are companies such as Sylvan and Oxford Learning.

The above remarks about the size of enterprises also have implications for urban/rural distribution of services. Cities usually have more higher-income families able to afford private tutoring, and also have sufficient density of population for companies to attract viable numbers of clients. Universities, which commonly provide many students seeking pocket money

from informal supply of tutoring, are also more likely to be located in cities. In this respect, shadow education tends to exacerbate urban/rural disparities in access to and use of educational provision. The extent of such disparities has been shown in diverse contexts. In Czech Republic, for example, 47.5% of the students in Prague sampled by Št'astný (2016: 20) were receiving tutoring, compared with 35.9% in the more rural Moravian-Silesian region. Similarly in Bangladesh, comparison of secondary students receiving private tutoring in English showed 84.7% doing so in Dhaka but only 60.7% in a rural area (Mahmud and Bray, 2017: 303).

Another major source of tutoring supply, more balanced across urban and rural areas, is the mainstream teaching profession. In many countries, regular teachers offer private tutoring as a way to supplement their incomes (Bray, 2021c). In some settings this is officially sanctioned, albeit with Ministry of Education regulations on which categories of teachers are permitted to tutor which categories of students, when and where (see e.g. Malaysia, 2006; Tunisia, 2015; United Arab Emirates, 2019). In more settings, however, such private tutoring by regular teachers is prohibited (see e.g. Egypt, 1998; Pakistan, 2019). The main social benefit of private tutoring by regular teachers is that it makes a resource available to students in need, and thus enhances their learning and human capital development. By supplementing the teachers' incomes, it also helps to retain teachers in the profession. However, such tutoring also raises the danger that teachers neglect their regular duties, for which their incomes are more or less guaranteed, in order to devote more time and effort to their private tutoring. Further, such tutoring can raise major issues of conflict of interest when teachers tutor students for whom they are already responsible during regular lessons. Writing about Cambodia, Dawson (2009) referred to 'the tricks of the teacher' which included deliberately withholding essential content during regular lessons in order to ensure that students enrolled in supplementary classes.

For these and other reasons, governments are increasingly paying attention to regulation of the tutoring sector (Bray and Kwo, 2014; Choi and Cho, 2016; Zhang, 2019). The school sector has long been regulated, albeit with gaps around the world in the enactment of these regulations, and now the tutorial sector is catching up. However, regulation for the sector is more challenging than for schooling because of its fluid nature. Much of the supply, especially by informal suppliers, is *ad hoc* and does not adhere to fixed timetables and locations. Especially challenging for regulation is provision via the internet, which may even be across national boundaries. For such reasons, some governments place emphasis on educating and empowering the consumers so that they can themselves demand appropriate behavior from tutoring providers (Bray and Kwo, 2014). Such bottom-up measures to promote accountability may be more effective than top-down regulations.

Social and political implications

The great global expansion of shadow education has brought increasing recognition of social and political implications alongside educational ones. These implications especially concern social inequalities, since it is obvious that more prosperous families can afford greater quantities and better qualities of private tutoring than can lower income families. Indeed in many contexts families increasingly feel that "schooling is not enough", and turn to the shadow sector to complement as well as supplement. Then the children whose families cannot afford tutoring, or cannot access it e.g. because they live in rural and remote locations, are left behind at the starting line.

These matters raise questions not only about regulation of the shadow education sector but also about possible enlargement of access. Some commentators (e.g. Lampl, 2017) have recommended government provision of vouchers to enable low-income families to invest in tutoring. However, such action would amount to official recognition that indeed schooling is not enough, and would be a further move toward privatization and marketization of the education sector in line with controversial ideologies of neoliberalism. Shadow education in some respects seems to challenge the institutions of schooling, in many settings employing untrained tutors yet displaying flexibility and tailored support in ways that families find attractive.

An alternative approach is suppression of the sector, exemplified by the Chinese government policy launched in 2021 (China, 2021, 2022). As noted by various commentators, however, this policy is tackling the symptoms rather than the underlying causes (Zhang, 2021). Comparable attempts in South Korea launched in 1980 had to be ameliorated first *de facto* and then officially in 2000 (Bray, 2009, pp. 47–54). Patterns elsewhere seem similarly to be more accommodating of shadow education, albeit with stronger regulatory approaches.

To address such matters effectively, a starting point is to take the topic more firmly 'out of the shadows'. Governments, academics and others can help by securing more data on its scale, nature and impact. And policy makers, administrators, schools and community bodies can usefully dialogue on what sorts of shadow education are likely to be of most benefit to whom and, by corollary, what sorts are likely to threaten wider goals of social and educational development.

References

- Aslam, M., Atherton, P., 2014. The shadow education sector in India and Pakistan: opening Pandora's box. In: Macpherson, I., Robertson, S., Walford, G. (Eds.), *Education, Privatisation and Social Justice: Case Studies From Africa, South Asia and South East Asia*. Symposium, Oxford, pp. 137–158.
- Aurini, J., Davies, S., Dierkes, J. (Eds.), 2013. *Out of the Shadows: The Global Intensification of Supplementary Education*. Emerald, Bingley.
- Bray, M., 1999. *The Shadow Education System: Private Tutoring and its Implications for Planners*. UNESCO International Institute for Educational Planning (IIEP), Paris.

- Bray, M., 2009. Confronting the Shadow Education System: What Government Policies for What Private Tutoring? UNESCO International Institute for Educational Planning (IIEP), Paris.
- Bray, M., 2017. Schooling and its supplements: changing global patterns and implications for comparative education. *Comp. Educ. Rev.* 61 (3), 469–491.
- Bray, M., 2021a. Shadow education in Europe: growing prevalence, underlying forces, and policy implications. *ECNU Rev. Educ.* 4 (3), 442–475.
- Bray, M., 2021b. Shadow Education in Africa: Private Supplementary Tutoring and Its Policy Implications. Comparative Education Research Centre, The University of Hong Kong, Hong Kong.
- Bray, M., 2021c. Swimming against the tide: comparative lessons from government efforts to prohibit private supplementary tutoring delivered by regular teachers. *Hung. Educ. Res. J.* 11 (2), 168–188.
- Bray, M., Hajar, A., 2023. Shadow Education in the Middle East: Private Supplementary Tutoring and Its Policy Implications. Routledge, London.
- Bray, M., Kobakhidze, M.N., Suter, L.E., 2020. The challenges of measuring outside-school-time educational activities: experiences and lessons from the Programme for International Student Assessment (PISA). *Comp. Educ. Rev.* 64 (1), 87–106.
- Bray, M., Kwo, O., 2014. Regulating Private Tutoring for Public Good: Policy Options for Supplementary Education in Asia. Comparative Education Research Centre, The University of Hong Kong, and UNESCO, Hong Kong and Bangkok.
- Bray, M., Lykins, C., 2012. Shadow Education: Private Tutoring and Its Implications for Policy Makers in Asia. Asian Development Bank and Comparative Education Research Centre, The University of Hong Kong, Mandaluyong City and Hong Kong.
- Byun, S.-Y., Chung, H.J., Baker, D.P., 2018. Global patterns of the use of shadow education: Student, family, and national influences. *Res. Sociol. Educ.* 20, 71–105.
- Chetty, M., Moloi, M.Q., Poliah, R.R., Tshikoro, J., 2017. The SACMEQ IV Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education. Department of Basic Education, Pretoria.
- China, CPC [Communist Party of China] Central Committee and the General Office of the State Council, 2021. Policy on Further Reducing the Workload of Students and the Burden of Out-of-School Training in the Compulsory Education Stage. Ministry of Education, Beijing [in Chinese]. https://mp.weixin.qq.com/s/-Qd8YcNaEx_3Y2qkTe3W0A.
- China, Ministry of Education, 2022. How to Do “Double Reduction” in 2022? Department of Supervision of Extracurricular Education and Training of Ministry of Education Issues Clear Card [in Chinese]. <https://mp.weixin.qq.com/s/PSM1yO4BzCvONMp9ztvZw>.
- Chionga, N.C.B., 2018. As Explicações: Seu Impacto Sobre o (In) Sucesso Escolar e o Alargamento da Classe Média [em Angola]. Masters dissertation. University of Minho.
- Choi, J., Cho, R.M., 2016. Evaluating the effects of governmental regulations on South Korean private cram schools. *Asia Pac. J. Educ.* 36 (4), 599–621.
- Dawson, W., 2009. The tricks of the teacher: shadow education and corruption in Cambodia. In: Heyneman, S.P. (Ed.), *Buying Your Way Into Heaven: Education and Corruption in International Perspective*. Sense, Rotterdam, pp. 247–264.
- Dwarkan, L., 2017. SACMEQ IV Study Mauritius: A Study of the Conditions of Schooling and the Quality of Education. Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), Port Louis.
- Egypt, Ministry of Education, 1998. Banning private tutoring. In: Ministerial Decree 592/1998. Ministry of Education, Cairo [in Arabic]. <https://docs.google.com/viewer?u=v&pid=sites&srcid=ZGVmYXVsdGRvbWVpbnxuYXNyZWxudWJhdGVhY2hlcnNuYW1lc3xneDozNTYzMWM4OTlwMGFkYTE0>.
- Enrich, S.R., 2018. Shadow Education and Social Inequalities in Japan: Evolving Patterns and Conceptual Implications. Springer, Cham.
- Enrich, S.R., 2021. Worldwide shadow education and social inequality: explaining differences in the socioeconomic gap in access to shadow education across 63 societies. *Int. J. Comp. Sociol.* 61 (6), 441–475.
- Feng, S., 2021. The evolution of shadow education in China: from emergence to capitalisation. *Hung. Educ. Res. J.* 11 (2), 89–100.
- Foordun, R., 1992. Private Tuition in Mauritius: The Mad Race for a Place in a “Five-Star” Secondary School. UNESCO International Institute for Educational Planning (IIEP), Paris.
- Hartmann, S., 2013. Education “home delivery” in Egypt: private tutoring and social stratification. In: Bray, M., Mazawi, A.E., Sultana, R.G. (Eds.), *Private Tutoring Across the Mediterranean: Power Dynamics, and Implications for Learning and Equity*. Sense, Rotterdam, pp. 57–75.
- Hollis, M., 1982. Education as a positional good. *J. Philos. Educ.* 16 (2), 235–244.
- Holloway, S.J., Kirby, P., 2020. Neoliberalising education: new geographies of private tuition, class privilege, and minority ethnic advancement. *Antipode: Rad. J. Geog.* 52 (1), 164–184.
- Kim, Y.C., 2016. Shadow Education and the Curriculum and Culture of Schooling in South Korea. Palgrave Macmillan, New York.
- KOSIS [Korean Statistical Information Service], 2020. Participation Rate on Private Education by School Level and Characteristics. http://kosis.kr/eng/statisticsList/statisticsListIndex.do?menuId=M_01_01&wcc=MT_ETITLE&parmTabId=M_01_01&statId=1963003&themald=#SelectStatsBoxDiv. (Accessed 23 August 2020).
- Lamp, S.P., 2017. Foreword. In: Jerrim, J. (Ed.), *Extra Time: Private Tuition and Out-of-School Study*, New International Evidence. Sutton Trust, London, p. 2.
- Mahmud, R., Bray, M., 2017. School factors underlying demand for private supplementary tutoring in English: urban and rural variations in Bangladesh. *Asia Pac. J. Educ.* 37 (3), 229–309.
- Malaysia, Government of, 2006. “Guidelines for Approval to Do Jobs Outside as Tutors or Part-Time Trainers”. Service Circular No. 1 of 2006. Ministry of Education, Putrajaya [in Malay].
- Mogaka, A.G., Gunga, S.O., Juma, S.N., Monanda, S.M., Owino, K.O., 2016. Factors that influence the need for private supplementary tuition in secondary schools: a case study of selected schools in Borabu District of Nyamira County, Kenya. *Int. J. Adv. Manag. Econ.* 4 (3), 9–20.
- Pakistan, 2019. Govt bars public school teachers from private tutoring. *Pakistan Today* 20. <https://www.pakistantoday.com.pk/2019/02/20/govt-bars-public-school-teachers-from-private-tutoring/>.
- Park, H., Buchmann, C., Choi, J., Merry, J.J., 2016. Learning beyond the school walls: trends and implications. *Annu. Rev. Sociol.* 42, 231–252.
- SACMEQ [Southern and Eastern Africa Consortium for Monitoring Educational Quality], 2010. How Widespread is the Provision of Paid Tuition in School Subjects?. SACMEQ Policy Issues 7. <http://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iii/policy-issue-series/007-sacmeqpolicyissuesseries-extralessons.pdf>.
- Sato, Y. (Ed.), 2012. One Hundred Years of Juku and Fifty Years of Juku Associations. Federation of Private Tutoring Associations, Tokyo [in Japanese].
- Shigwedha, A.N., Nakashole, L., Auala, H., Amakutuwa, H., Ailonga, I., 2015. The SACMEQ IV Project in Namibia: A Study of the Conditions of Schooling and the Quality of Education in Namibia. Ministry of Education, Arts and Culture, Windhoek.
- Sieverding, M., Krafft, C., Elbadawy, A., 2019. An exploration of the drivers of private tutoring in Egypt. *Comp. Educ. Rev.* 63 (4), 562–590.
- Štastrný, V., 2016. Private supplementary tutoring in the Czech Republic. *Eur. Educ.* 48 (1), 1–22.
- Tan, J., 2009. Private tutoring in Singapore: bursting out of the shadows. *J. Youth Stud.* 12 (1), 93–103.
- Tunisia, Ministry of Education, 2015. Décret gouvernemental n° 2015-1619 du 30 octobre 2015, fixant les conditions d’organisation des leçons de soutien et des cours particuliers au sein des établissements éducatifs publics. *Journal Officiel de la République Tunisienne* 88, 2641–2643, 3 November.
- Tsiloglou, L., 2005. Frontistaria in Greece: Their History and People. Kedros, Athens [in Greek].
- UNESCO, 2000. The Dakar Framework for Action – Education for All: Meeting Our Collective Commitments. UNESCO, Paris.
- United Arab Emirates [UAE], 2019. UAE to Legalise Private Tutoring in Order to Improve Standards in Schools. *The National*, 26 April. <https://www.thenational.ae/uae/education/uae-to-legalise-private-tutoring-to-improve-standards-in-schools-1.853516>.
- Zhang, W., 2019. Regulating private tutoring in China: uniform policies, diverse responses. *ECNU Rev. Educ.* 2 (1), 25–43.
- Zhang, W., 2021. Non-State Actors In Education: The Nature, Dynamics and Regulatory Implications of Private Supplementary Tutoring. Background Paper for Global Education Monitoring Report. UNESCO, Paris. <https://gem-report-2021.unesco.org/wp-content/uploads/2021/12/03-Wei.pdf>.

- Zhang, W., Bray, M., 2021. A changing environment of urban education: historical and spatial analysis of private supplementary tutoring in China. *Environ. Urbanization* 33 (1), 43–62.
- Zhang, W., Bray, M., 2020. Comparative Research on shadow education: achievements, challenges, and the agenda ahead. *Eur. J. Educ.* 55 (3), 322–341.
- Zhang, W., Yamato, Y., 2018. Shadow education in East Asia: entrenched but evolving private supplementary tutoring. In: Kennedy, K., Lee, J.C.K. (Eds.), *Routledge International Handbook on Schools and Schooling in Asia*. Routledge, London, pp. 323–332.

From #BlackLivesMatter to Black Lives Matter at school: how the fight for Black Liberation entered the schoolhouse

Denisha Jones^a and Awo Okaikor Aryee-Price^b, ^a Sarah Lawrence College, Bronxville, NY, United States; and ^b Education for Liberation Network, United States

© 2023 Elsevier Ltd. All rights reserved.

From the streets to the school house	353
Grassroots organizing for racial justice in schools	356
A curriculum for Black Lives	356
Struggling for union support and BLMAS in higher education	358
From a week of action to a year of purpose and a lifetime of practice	359
Year of Purpose self reflection questions	359
Black Lives Matter at school national days of action	360
First day: Black to School (varies depending on each district's designated first day of school)	360
October 14th: Justice for George Day	360
Principle: Restorative Justice	360
November 20th: Transgender Day of Remembrance	360
Principle: Trans Affirming	360
December 3rd: International Day of Persons with Disabilities	360
Principles: Globalism and Collective Value	360
January Queer Organizing Behind the Scenes	360
Principle: Queer Affirming	360
February 18th: Unapologetically Black Day	360
Principle: Unapologetically Black	360
March 6th: Student Activist Day	360
Principles: Loving Engagement and Empathy	360
April Revolutionary Black Arts	361
Principle: Intergenerational	361
May 3rd: Black Radical Educator Day	361
Principle: Black Villages	361
June 5th: #SayHerName Day	361
Principle: Black Women	361
June 19th: Education for Liberation Day	361
Principles: Black Families and Diversity	361
Last Day of School: A Day for Self-Reflection	361
Review all 13 principles	361
Conclusion	361
References	361

On February 26, 2012, Trayvon Martin, a 17-year-old Black boy, was walking home from a convenience store in Sanford Florida, when a violent vigilante racially profiled, stalked, confronted, and killed him. Nearly two months later, after numerous rallies demanding justice, Trayvon's killer was arrested and charged with second-degree murder. After a not guilty verdict was returned on July 13, 2013, two Black queer women, Alicia Garza and Patrice Khan-Cullors, began a series of social media posts that led to the creation of a modern-day movement for Black liberation. In response to a love letter to Black people posted by Garza, Khan-Cullors typed three words that would become an international rallying cry for racial justice: #BlackLivesMatter (Khan-Cullors and Bandle, 2017). As the killings of unarmed Black people continued to galvanize the US and the world, #BlackLivesMatter grew into an international grassroots movement fighting for Black liberation. From protests to teach-ins, the explosion of #BlackLivesMatter struck a chord in many Black people leading to the organizing work from the larger coalitional work of the Movement 4 Black Lives (M4BL).

In 2016 as the assault on Black life continued with the killing of Philando Castille and Alton Sterling, teachers in Seattle Washington organized the first #BlackLivesMatter day of solidarity in a public elementary school. Despite the attacks that ensued including a bomb threat to the school, the teachers at John Muir Elementary wore t-shirts, collaborated with Black-led organizations, and created a curriculum to support students' understanding of the #BlackLivesMatter movement. A month later, educators across Seattle spearheaded a city-wide Black Lives Matter at School day with over 3000 teachers participating in the second #BlackLivesMatter solidarity event in schools (Au and Hagopian, 2018). As word of these and other events spread, "... the 'Black Lives Matter' cry for freedom hopped on the yellow bus, walked through the school-house door, occupied the gymnasium, rallied in the auditorium, ripped up the textbooks, and took over the daily lesson plans" (Hagopian, 2020, p. 1).



(Au and Hagopian, 2018)

From the streets to the school house

The ushering of Black Lives Matter at School, as a national movement, came from the principled, strategic, and organizing efforts of educators from Philadelphia, Pennsylvania who witnessed the Day of Action that took place in Seattle and expanded it to a Week of Action. Members of the Caucus of Working Educators (WE) Racial Justice Organizing Committee (RJOC), inspired by the events in Seattle, came together to plan a week of action for students in Philadelphia. Tamara Anderson notes how their work was not prompted by a specific local event but rather a response to ongoing issues that plagued the Black and Latinx populations in their district (Anderson, 2020). They intentionally set out to make their event inclusive of issues beyond education such as housing, poverty, and fair wages. Thus they invited parents, college faculty, and local organizers as partners to help build what became the Black Lives Matter Week of Action (Anderson, 2020). As word of the week of action in Philadelphia spread they immediately received pushback. Those opposed to racial justice labeled #BlackLivesMatter as anti-police and a terrorist organization. As it became clear that some members of the caucus were confused about the purpose and not willing to embrace the week of action, the Philly organizers made a strategic decision to counter the attacks.

We responded to this pushback by re-centering the narrative on the truth of the Black Lives Matter movement, a grassroots movement for "freedom, liberation, and justice." We also faced fear with truth by using the blueprint of Black Lives Matter's thirteen guiding principles, which emphasize inclusive values, including restorative justice, globalism, and trans affirming, and, the power that resides in being unapologetically Black and no longer silent or invisible. The thirteen principles laid the groundwork for our planning and curriculum, and eventually, for the week going national.

Anderson (2020, p. 51).

Denisha Jones
January 21, 2017 · 🌐

Shout to all the Philly teachers making this happen. This is so badass!!! Thank you!!

**Black Lives Matter Week of Action
Calendar of Events***

*Check the full calendar for descriptions and more events at <http://tinyurl.com/BLMPhEd>
For the full list of Curriculum Resources, go to <http://tinyurl.com/BLMPhEdCurriculum>

Monday, 1/23	Tuesday, 1/24	Wednesday, 1/25	Thursday, 1/26	Friday, 1/27	Saturday, 1/28
Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment	Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment	Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment	Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment	Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment	Guiding Principle: Resistance, Justice, Economic, Learning, Empowerment
Curriculum Links: Curriculum Middle & High	Curriculum Links: Curriculum Middle & High	Curriculum Links: Curriculum Middle & High	Curriculum Links: Curriculum Middle & High	Curriculum Links: Curriculum Middle & High	Curriculum Links: Curriculum Middle & High
Meet your BLM 7 (Share and Retweet) Take a solidarity photo with everyone in their year!	#EducationPXL Kick-Off Meeting 5:00-6:00pm at D. King's Center (17th Street Store)	"How do we bridge the gap between parents/families and schools?" Power and Community Meeting 5:00pm at City Hall (10th and Market)	"The effects of mass incarceration on Black and Brown communities" Screening of the Film 13th and Talkback Discussion 5:00-6:00pm at Drexel Elementary (19th Street)	"Demosty Black Women and Black Girls: Monopoly, Slavery, and Power" Panel Discussion 5:00-6:00pm at UPenn Graduate School of Education Room 202 (37th Street)	LGBO Youth Conversation about "Patriot" and "Moonlight" 6:00-8:00pm at Temple University, Parkland Learning Center (20th Street)

Caucus of Working Educators
January 20, 2017 · 🌐

Your guide to #BLMPhEd Week, starting on Mon 1/23! There are courageous events, guiding principles, and activity ideas for every day. Download the PDF version ...
See More

(Jones and Hagopian, 2020, p. 201).

On January 21, 2017, the Caucus of Working Educators posted a link to their Black Lives Matter Week of Action Calendar of Events on Facebook. The Google doc compiled links to K-12 curriculum resources to teach the 13 Guiding Principles (13GPs) on different days of the week, and information on activities and events planned throughout the city. In addition to the Week of Action calendar, the Philly organizers prepared a FAQ guide with talking points for educators, families, and parents as leverage for teachers to receive support from school administrators (Anderson, 2020).



Photo taken at Free Minds Free People 2017

As word of the first Black Lives Matter Week of Action spread, the Philly organizers were invited to present their work at the Free Minds Free People conference in Baltimore in July of 2017. The Philly organizers would use the FMFP conference space to recruit more education organizers and build a base to move what they started with the Week of Action in their local community and move it onto a National stage. Attendees learned about their week of action, the 13 GPS, and the initial success and challenges they faced. A signup sheet for those interested in organizing a similar Week of Actions was sent around and the call to action made its way to other events such as the Charlottesville Teach-in where organizers co-planned and developed the relationships that would lead to the National BLM at School movement teach-in Charlottesville, Virginia. In October of 2017, the first national planning meeting met over zoom with representatives from New York, New Jersey, and Washington, resulting in the National Black Lives Matter at School (BLMAS) movement. The National Week of Action was officially named the Black Lives Matter at School Week of Action and would be held the first week of February across the country.

An all-Black national steering committee was formed to ensure the work of BLMAS would remain in the hands of Black educators, activists, and organizers. Though not officially affiliated with #BlackLivesMatter (BLM) chapters, BLMAS embraced the grassroots, Black feminist organizing styles of the larger movement. During the fall of 2017, the BLMAS steering committee organized virtual meetings to support the organizers across the country in developing their own Week of Actions (WOA) for February 2018.

Through the national steering committee, a national curriculum committee was created to assist teachers in developing lesson plans and other materials. Building off of curriculum resources created by the Seattle educators and Philly organizers, the national curriculum committee created accessible Google folders that housed a vast amount of K-12 and higher education lesson plans and resources. Local WOA organizers were invited to create and share their calendar events while the national steering committee organized outreach webinars, endorsements, and a website to share resources and promote events. On Monday, February 5, 2018, the first BLMAS WOA began in 20–30 cities across the US.

D.C. AREA EVENTS



WED JAN 24 6-8PM	#DCschoolsBLM Curriculum Share Inspired Teaching Demonstration School (200 Douglas St NE) Get lesson plans, share lesson ideas, address questions and concerns
SAT JAN 27 10-12PM	#DCschoolsBLM Curriculum Share Payden Center, Trinity Washington University (125 Michigan Ave NE) Get lesson plans, share lesson ideas, address questions and concerns
MON JAN 5- SAT JAN 10	Teaching for Change Workshops Classroom visits available on Black Muslims in the U.S., the Palenque maroon community, a go-go musician, and a SNCC veteran
MON FEB 5 6:30-8PM	Poor People's Campaign Mass Meeting All Souls Church, Unitarian (1500 Harvard St NW) Led by Poor People's Campaign: A National Call for Moral Revival
TUE FEB 6 6:30-8:30PM	Paul Ortiz Book Talk Busboys & Poets Takoma (235 Carroll St NW) Discuss <i>An African American and Latinx History of the United States</i>
WED FEB 7 6:30-8:30PM	#MoonlightInSchools Twitter Chat Join author @HankinsLatoya to watch "Moonlight" and discuss how to elevate queer and trans students of color in DC #MoonlightInSchools
THU FEB 8 5:30-7:30	Examining Racial Justice Interscholar Seminar Edmund Burke School (4101 Connecticut Ave NW) Intergenerational dialogue about topics related to Black Lives Matter
THU FEB 8 6-7PM	Book Panel Discussion with Teen Authors Mt.Pleasant Library (3160 16th St NW) Discussion of the Shoutmouse Press book, <i>The Day Tajon Got Shot</i>
FRI FEB 9 5:30-8PM	Love Poems to Black Women: An Open Mic Night Woodrow Wilson High School (3950 Chesapeake St NW) Join young artists to share poetry about the themes from the week

DC WOC Calendar 2018

BLACK LIVES MATTER

PRINCIPLES

<https://blacklivesmatter.com/about/what-we-believe/>

PRINCIPLE	DESCRIPTION
RESTORATIVE JUSTICE	We intentionally build and nurture a beloved community that is bonded together through a beautiful struggle that is restorative, not depleting.
DIVERSITY	We acknowledge, respect, and celebrate differences and commonalities.
EMPATHY	We practice empathy. We engage comrades with the intent to learn about and connect with their contexts.
LOVING ENGAGEMENT	We embody and practice justice, liberation, and peace in our engagements with one another
QUEER AFFIRMING	We foster a queer-affirming network. When we gather, we do so with the intention of freeing ourselves from the tight grip of heteronormative thinking, or rather, the belief that all in the world are heterosexual (unless s/he or they disclose otherwise).
TRANSGENDER AFFIRMING	We are self-reflexive and do the work required to dismantle cisgender privilege and uplift Black trans folk, especially Black trans women who continue to be disproportionately impacted by trans-antagonistic violence. We make space for transgender brothers and sisters to participate and lead.
UNAPOLOGETICALLY BLACK	We are unapologetically Black in our positioning. In affirming that Black Lives Matter, we need not qualify our position. To love and desire freedom and justice for ourselves is a prerequisite for wanting the same for others.
BLACK WOMEN	We build a space that affirms Black women and is free from sexism, misogyny, and environments in which men are centered.
BLACK FAMILIES	We make our spaces family-friendly and enable parents to fully participate with their children. We dismantle the patriarchal practice that requires mothers to work "double shifts" so that they can mother in private even as they participate in public justice work.
BLACK VILLAGES	We disrupt the Western-prescribed nuclear family structure requirement by supporting each other as extended families and "villages" that collectively care for one another, especially our children, to the degree that mothers, parents, and children are comfortable.
GLOBALISM	We see ourselves as part of the global Black family, and we are aware of the different ways we are impacted or privileged as Black people who exist in different parts of the world.
INTERGENERATIONAL	We cultivate an intergenerational and communal network free from ageism. We believe that all people, regardless of age, show up with the capacity to lead and learn.
COLLECTIVE VALUE	We are guided by the fact that all Black lives matter, regardless of actual or perceived sexual identity, gender identity, gender expression, economic status, ability, disability, religious beliefs or disbeliefs, immigration status, or location.

BLM 13GPS Poster

Abbreviated Calendar of Events (see [Facebook.com/BLMEDU](https://www.facebook.com/BLMEDU) for details) *

Monday 2/5	Tuesday 2/6	Wednesday 2/7	Thursday 2/8	Friday 2/9
★ 5:30-8 pm: Black Lives Matter in Education C.A.S.A. Middle School, Bronx NY ★ 6 pm: Tubman Teaching Artist, Lacesha Berry, one-woman show. Free, RSVP Queens Central Library Jamaica, NY ★ 6 pm Dance w/ArilJoy, Social fitness class hosted by ArilJoy Fitness. Free ★ 5-7 pm: Teach-in at the Museum of the City of New York: <i>Activist New York</i> 5th Ave & 103rd St, NY	★ 6:30- 8 pm Perspectives: Dance & Social Justice Partnered with Nimbus Dance Works Manhattan Country School, NYC ★ 5:30-7:30 pm Urban Word First Draft Youth Open Mic - BLM Edition. Free, RSVP 5 West 19th Street, 3rd Floor, NY, NY 1001 ★ 6:00-9pm: Panel Conversation and performance by Batala at Caribbean Cultural Center African Diaspora Institute	★ 6:30 pm The Intersection of High-Stakes...Testing & Race in NYC Schools I.C.E. Middle School NYC ★ 6 pm to 9 pm "They...who?" World-Making w/ Gender Exploration **QTPOC safe space Audre Lorde Project 147 West 24th St, NY ★ 6-8 pm: Teaching the Black Panthers - Zinn Education Project Grace Church High School, 40 Cooper Union, NYC, Rm. 205	★ 5:30-8 pm What is a Community School? Educator Resources and Panel Discussion Brooklyn Museum Teaching Demo ★ 5:30-7:30 pm Talking about Race in the Classroom Join Borders Crossers for a free 2-hr workshop. RSVP 96 Morton St, 7th Floor New York, NY 10014 ★ 5:30-7:30 pm Teachable Moments: Black Lives Matter storytelling event Free, RSVP Harlem Nights- 261 7th Ave NY, NY 10030	★ 6-7:30 pm Stories for Black Lives with Peggy Pettitt 1573 Madison Avenue ★ 7-9 pm Teen Dance Party Schomburg Center ★ 7-10 pm Chill & Design: QTBCPOC Love Edition **QTPOC safe space. Make your own unlimited jewelry designs all night. Tickets: FB Invite New Woman Space, 188 Woodpoint Rd. Brooklyn, NY 11211
★ February 3rd, 5-10pm: BLM Week of Action Kick-off with Urban Word NYC at Brooklyn Museum ★ February 4th, 6:30-8pm Arts Showcase & Screening Lion Theatre 410 W 42nd St. NY ★ February 10th, 10am-9pm: BLM Teen Conference @ Schomburg Center ★ February 12th, 6pm: "Reclamation: A Celebration of Art & Action" Exhibition at the Wild Project				★ See more 2/3 and other special events on Facebook page!

NYC WOC Calendar 2018

Grassroots organizing for racial justice in schools

A curriculum for Black Lives

The WOA focuses on teaching the 13GPs and engaging students and the community around the four national demands (See [Table 1](#)). Chris Rogers who transitioned from contributing to the Philadelphia curriculum to chairing the national curriculum committee shared the process of curating educational resources.

The educational resources that I and others developed in 2017 were grounded in a belief that centering the words and deeds of Black freedom fighters in the classroom would have a liberatory potential for all involved. In that initial year, we were able to set up a framework that performed what Barbara Ransby calls "political quilting" – bridging movement work across time and place, and nurturing relationships and mass awareness. We learned that by recognizing the inherent intellectual practice in diverse organizing traditions, we were building a rich repository to reimagine the purpose of education towards a praxis of liberation.

Rogers (2020, pp. 58–59).

Rogers goes on to note how the work of collecting primary and secondary sources from Black-led justice organizations amounted to a "curricular strike" where the corporate neoliberal curriculum is set aside to make space for communities of color to center their cultural knowledge and freedom dreams (Rogers, 2020, p. 59).

In January during the MLK holiday weekend, the national curriculum committee released a revised set of curriculum resources to support the WOA. The national curriculum included book lists, lesson plans, specialized syllabi, posters, and other resources for early childhood, elementary, and high school students. To help prepare educators to effectively use the curriculum materials, local cities held curriculum fairs before the WOA began. Teachers new to the BLMAS movement could come together to learn about the 13GPs, hear from the experience of teachers who participated in the past or created lessons, and collect curriculum resources. The collective work of curricular organizing stems from a long tradition of Black educators creating curriculum materials to subvert the dehumanization and erasure of Black lives (King, 2014).

Early in the growth of the national movement, the BLMAS curriculum was made accessible to young learners. When Laleña Garcia, a kindergarten teacher in New York City, first learned about the WOA, she began by translating the 13GPS into child-friendly

Table 1 Thirteen guiding principles (13GPs) by day.

Monday	Loving engagement, restorative justice, empathy
Tuesday	Globalism and diversity
Wednesday	Trans affirming, queer affirming, and collective value
Thursday	Intergenerational, black villages, black families
Friday	Black women and unapologetically black

Table 2 BLMAS 13GPs with *child-friendly language* (in italics).

Restorative justice is the commitment to build a beloved and loving community that is sustainable and growing. “We know that if you knock down someone’s block building, you have to help them rebuild it, you can’t just say, “Sorry,” and walk away. We also know that it’s important for kids to be able to make a better choice another time, and it’s grownups’ job to help them make better choices and to give them chances to do that. Another way to say that is restorative justice.”
Empathy is one’s ability to connect with others by building relationships built on mutual trust and understanding. “It’s so important to think about how other people feel, because different people have different feelings. Sometimes it helps to think about how you would feel if the same thing that happened to your friend happened to you. Another way to say that is empathy.”
Loving engagement is the commitment to practice justice, liberation and peace. “It’s so important to make sure that we are always trying to be fair and peaceful, and to engage with other people (treat other people) with love. We have to keep practicing this so that we can get better and better at it. Another way to say that is loving engagement.”
Diversity is the celebration and acknowledgment of differences and commonalities across cultures. “Different people do different things and have different feelings. It’s so important that we have lots of different kinds of people in our community and that everyone feels safe. Another way to say that is diversity.”
Globalism is our ability to see how we are impacted or privileged within the black global family that exists across the world in different regions. “Globalism means that we are thinking about all the different people all over the world, and thinking about the ways to keep things fair everywhere.”
Transgender affirming is the commitment to continue to make space for our trans siblings by encouraging leadership and recognizing trans-antagonistic violence, while doing the work required to dismantle cisgender privilege and uplift black trans folk. “Everybody has the right to choose their own gender by listening to their own heart and mind. Everyone gets to choose if they are a girl or a boy or both or neither or something else, and no one else gets to choose for them.”
Queer affirming is working toward a queer-affirming network where heteronormative thinking no longer exists. “Everybody has the right to choose who they love and the kind of family they want by listening to their own heart and mind.”
Collective value means that all black lives, regardless of actual or perceived sexual identity, gender identity, gender expression, economic status, ability, disability, religious beliefs or disbeliefs, immigration status or location, matter. “Everybody is important, and has the right to be safe and happy. Another way to say that is collective value.”
Intergenerational is a space free from ageism where we can learn from each other. “It’s important that we have spaces where people of different ages can come together and learn from each other. Another way to say that is intergenerational.”
Black families creates a space that is family friendly and free from patriarchal practices. “There are lots of different kinds of families; what makes a family is that it’s people who take care of each other. It’s important to make sure that all families feel welcome.”
Black villages is the disruption of western nuclear family dynamics and a return to the “collective village” that takes care of each other. “There are lots of different kinds of families; what makes a family is that it’s people who take care of each other; those people might be related, or maybe they choose to be family together and to take care of each other. Sometimes, when it’s lots of families together, it can be called a village.”
Black women is the building of women-centered spaces free from sexism, misogyny, and male-centeredness. “There are some people who think that women are less important than men. We know that all people are important and have the right to be safe and talk about their own feelings.”
Unapologetically black is the affirmation that black lives matter and that our love, and desire for justice and freedom are prerequisites for wanting that for others. These principles are the blueprint for healing and do not include nor do they support ignoring or sanitizing the ugliness and discomfort that comes with dealing with race and anti-race issues. “There are lots of different kinds of people and one way that we’re different is the color of our skin. It’s important to make sure that all people are treated fairly, and that’s why we, and lots of other people all over the country and the world are part of the Black Lives Matter movement.”

Source: Garcia (2020, pp. 141–144).

language. She notes, “I hoped that the kid-friendly language would provide a jumping-off point for educators who wanted to bring the principles into their classrooms while remaining age-appropriate for their young students” (Garcia and Clark, 2020, p. 135, see Table 2). In addition to articulating the 13GPs into kid-friendly language, Garcia worked with NYC art teacher Caryn Davidson to create a coloring book to engage preschool-age children. Together they published two activity books, one for preschoolers and another for K-fifth grade students that examine each principle and provide reflective questions, letters to adults, and a teacher’s guide.

In addition to the 13GPs, the national BLMAS steering committee issued four demands for all school districts in the US to work toward educational justice for Black lives.

1. End zero tolerance discipline and mandate restorative justice.
2. Mandate Black history and Ethnic Studies.
3. Hire more Black teachers.
4. Fund counselors, not cops.

When the first national BLMAS WOA launched in 2018, there were three national demands, however, after the assault of a Black girl by a school resource officer in South Carolina, the steering committee agreed to add the fourth demand. Some local WOC organizers have adopted demands specific to their communities.

Local WOA organizers plan events including youth-led rallies that focus on the demands. The national steering committee collaborates with multiple racial justice organizations to garner support for the demands. Since BLMAS went national, Seattle public schools adopted a resolution in support of ethnic studies. After the murder of George Floyd, students in Minneapolis and Oakland successfully worked to defund school police. Thus, centering the demands is an important component of the WOA and the national movement for racial justice in schools.



BLMAS Demands Poster by Caryn Davidson

Struggling for union support and BLMAS in higher education

In the United States, there are two national teachers' unions: The American Federation of Teachers (AFT) and the National Education Association (NEA). Teachers become members of either union depending on where they teach, and the union is authorized to engage in collective bargaining on behalf of the teachers. Five states prohibit collective bargaining while others permit and some mandate bargaining. Teachers' salaries, benefits, and working conditions are often negotiated through collective bargaining. Unions can also provide support for initiatives by endorsing them and even providing funds at the local level. Thus, early on in the BLMAS national movement, many educators attempted to gain union support through endorsements at the local and national levels. NEA was the first national union to endorse the WOA and encourage teacher participation at the local level. Beginning in 2020 NEA Ed Justice, the digital center for racial justice resources in NEA, created a map for teachers to list their BLMAS WOA events and maintains a Black Lives Matter @ School web resource page for members. Given this level of support, teachers who are members of NEA were able to successfully pass resolutions endorsing the WOA. These endorsements provide protection to teachers from attacks by anti-racial justice groups.

The AFT finally endorsed the BLMAS WOA in 2021 after many local chapters struggled to gain union support. In NYC, the WOA organizers failed to secure an endorsement from their local United Federation of Teachers (UFT) until the national endorsement from AFT came in 2021. Without union support teachers in NYC were at risk of being reprimanded for participating in the WOA. Myrie, a NYC teacher, and chapter leader shared how securing union support for BLMAS was crucial to supporting the struggles of BIPOC (Black, Indigenous, and People of Color) students. He notes

You see the difference in LA and Chicago unions as leadership and the rank and file immerse social justice demands in their contracts. This is the way it is done. The leadership of the LA and Chicago unions is part of UCORE, and hence the urgency to engage its membership about the students they serve and meeting students' and their families' needs.

Jones and Hagopian (2020, p. 105).

After AFT endorsed the BLMAS WOA, the NYC organizers continued to advocate for full support. As the attacks on Critical Race Theory (CRT) and the banning of books and teaching about racism continue to grow, WOA organizers must have full protection from their unions.

Early in the development of the national BLMAS movement, many higher education professors and researchers endorsed the WOA. A higher education committee was formed in 2018 and many colleges of education introduced the WOA to their teacher education students. Higher education WOA organizers often host panel discussions, film screenings, and curriculum fairs. Some even take up the demands as they relate to their institution. Dana Morrison, a professor at a college in Pennsylvania, worked with the Philadelphia organizers on a research project which led to her involvement in BLMAS. During a 2018 planning meeting, she began discussions with other faculty and graduate students which led to a partnership of 30 people from five local colleges and

universities. She notes how this “initial work brought together educators of all levels around a common concern, racial justice” and “re-centered [her] commitments as a critical educator and built relationships with other teachers in ways [she] never thought possible” (Jones and Hagopian, 2020, p. 197).

Higher education involvement came in multiple ways, including a Black Lives Matter Early Childhood Symposium at Bank Street College in New York. Takiema Bunche Smith was the Director of the Center on Culture, Race, and Equity, when she organized the first Early Childhood Symposium as part of the 2019 BLMAS WOA. After learning about the WOA, Takiema immediately recognized the need to center the youngest Black children in our fight for racial and education justice. Her goal was to situate the historical fight for Black freedom with an emphasis on Black children and how the denial of Black humanity starts at birth. She states, “When we think of a three-year-old, we may picture a child playing, but for Black children, their bodies have been seen as commodities and they’ve been denied that opportunity. During enslavement, Black children as young as three were made to work in the fields” (Jones and Hagopian, 2020, p. 153). Noticing how young Black children were often not included in these historical discussions, Takiema approached her college leadership with the idea of an early childhood symposium as part of the BLMAS WOA. Each symposium highlighted a different aspect of anti-Black racism in the early years and provided a rare professional development experience that centers Black children.

From a week of action to a year of purpose and a lifetime of practice

The global pandemic of COVID-19 necessitated the shutdown of the world as well as taking notice of all the ways systems have impacted people’s lives, particularly the lives of Black, Indigenous, Latinx, and other People of Color. The Movement for Black Lives uprisings of 2020 that swept the globe after the state-sanctioned and extrajudicial murders of George Floyd, Breonna Taylor, Tony McDade, Ahmaud Arbery, and others named and unnamed, ignited new urgency and radical possibilities for advancing abolitionist practice and uprooting institutional racism.

The uprising for Black lives in 2020 prompted the National Black Lives Matter at School movement to expand its proposed activities from a “*Week of Action*” to a “*Year of Purpose*,” and Lifetime of Practice in addition to the annual Week of Action held during the first week of February. The centerpiece of the *Year of Purpose* asks educators to reflect on their own work in relation to antiracist pedagogical and abolitionist practices, persistently challenging themselves to center Black lives in their classrooms. Additionally, as part of the *Year of Purpose*, educators are asked to participate in intentional days of action throughout the school year uplifting different intersectional themes vital to making Black lives matter in schools, communities, and beyond.

The formal and informal learning environments we aspire to create reflect a deep understanding of the experiences of Black children, families, and communities, as well as our own ongoing work of critical self-reflection and personal transformation. The questions we are compelled to ask ourselves in perpetuity are: Are we creating humanizing communities that respond to the concerns of our students? Are we committed to leveling up our expectations for Black students? The *Year of Purpose* compels educators to turn inward in order to reach outward, linking their efforts to broad, integrated movements for social justice and change.

As our ancestor, the Black lesbian warrior poet Audre Lorde, stated, “There is no such thing as a single-issue struggle because we do not live single issue lives.” This means we must commit to living our principles every day, in and out of our classrooms, within our homes, and with our communities. It is a commitment to the village.

Integral to the *Year of Purpose*, the excerpted questions chosen are meant to support educators—and parents who are educating their children at home during the pandemic—throughout the year. The questions below, as well as pieces from the paragraphs above written by Awo Okaikor Aryee-Price, Maria Fernandez, and Christopher Rogers, first appeared in the Education for Liberation Network’s educator plan book *Planning to Change the World: A Plan Book for Social Justice Teachers* (2019–2020). We invite educators and educators-in-training to meditate on the questions that follow, and—given that no such list can be comprehensive—to pose questions of their own. Only through deliberate reflection can we realign our teaching practices to meet our current challenges and invent new practices where there are none. Additional information about the *Year of Purpose* and opportunities to participate are available at BlackLivesMatterAtSchool.com.

Year of Purpose self reflection questions

1. What is our school’s relationship to Black community organizing? Do we have relationships with local movement organizers? Do they see our school as a place that believes in their mission? Do they see our school as a place to connect with local families?
2. How are school-wide policies and practices—especially disciplinary practices—applied across categories of race? Do problematic patterns emerge when we look at how policies are applied to Black students and when we also consider the intersections of gender, sexual orientation, and (dis)ability with Blackness?
3. How are the voices, accomplishments, and successes of Black folx uplifted in my lessons, units, and curriculum? Rather than focus on singular events or individuals, does my approach highlight the everyday actions and community organizing that will lead to change?
4. In what ways do our practices erase the histories of our students and prevent them from bringing their whole selves into the learning environment?
5. How do I understand the role that local/state laws and policies have on the educational experiences of my students? What is my role in working to change policies, regulations, and practices that harm Black students and families?

Black Lives Matter at school national days of action

In addition to self-reflection questions, the Year of Purpose encourages educators, parents, organizers, and community members to participate in the following days of action throughout the year. Each action is grounded in the Movement for Black Lives Principles that we adopted as well:

First day: Black to School (varies depending on each district's designated first day of school)

- Wear the Black Lives Matter at School shirt
- Review the Black Lives Matter at School reflection questions and write up an anti-racist action plan for the year
- Graffiti Wall: "What are we going to do differently this year to further the movement for Black lives in our school."
- Post a video to social media
- Twitter chat

October 14th: Justice for George Day

Principle: Restorative Justice

October 14th is George Floyd's Birthday. Justice for George is a day to remember George Floyd and call for the defunding of the police and the redirecting of those funds toward social programs and education.

November 20th: Transgender Day of Remembrance

Principle: Trans Affirming

November 20th,
Transgender Day of Remembrance and William Dorsey Swann Day.

December 3rd: International Day of Persons with Disabilities

Principles: Globalism and Collective Value

December 3rd is International Day of Persons with Disabilities Day. Harriet Tubman, and Fannie Lou Hamer are two disabled freedom fighters we revere, even as the disabilities they carried with them into struggle are not consistently lifted up as assets in their fight. To fight against societal ableism, we must celebrate our differences and understand how the lessons from the Disability Justice movement and Black disabled organizers teach us how to build inclusive, accessible movements.

January Queer Organizing Behind the Scenes

Principle: Queer Affirming

January- During January, we find it critical to lift up Bayard Rustin, one of the principal organizers behind the March on Washington which is crowned as one of MLK's lasting achievements. To be queer-affirming means lifting up our queer ancestors who were at the foundation of our movements throughout time. This deepens the purpose of MLK day to understand that no one person makes a movement, highlighting how MLK's legacy encompasses the contributions of many.

February 18th: Unapologetically Black Day

Principle: Unapologetically Black

During this day, we celebrate the life's work and contributions that Audre Lorde and Toni Morrison have made to our collective movements. They are not only freedom fighters, but they share the same birthday February 18th, and so we lift them up and celebrate them.

March 6th: Student Activist Day

Principles: Loving Engagement and Empathy

March 6th: Barbara Johns Black Student Activist Day is a day to celebrate Black student activists and organizers.

- Philly artist (M. Brian Bowens) who created this amazing solidarity piece to show unity and activism through compassion and empathy. <https://bbsketch.com/store>
- Grades K-2 Activities/readings
 - <https://kids.britannica.com/kids/article/Barbara-Johns/631154>

- Grades 3–6 Activities/readings
 - <https://junior.scholastic.com/pages/content-hubs/real-teens-of-history/lesson-plan-desegregation-barbara-johns.html>
 - <https://www.rif.org/literacy-central/book/girl-tar-paper-school-barbara-rose-johns-and-advent-civil-rights-movement>
- Grades 7–12 Activities/readings
 - <https://snccdigital.org/events/barbara-johns-leads-prince-edward-county-student-walkout/>
 - <https://www.zinnedproject.org/news/tdih/barbara-johns-leads-student-protest/>
 - <https://kinginstitute.stanford.edu/additional-activities>

April Revolutionary Black Arts

Principle: Intergenerational

April- During National Library Week, we seek to center the classic contributions of Black Writers and artists across the generations: Zora Neale Huston, Faith Ringgold, Alma Thomas, Augusta Savage, Jasmine Mans. How are the themes and radical vision that they brought to their art reflected in your classrooms and communities? How can young people extend on these legacies?

May 3rd: Black Radical Educator Day

Principle: Black Villages

On May 3rd, Septima Clark's birthday, we celebrate Black Radical Educator Day.

June 5th: #SayHerName Day

Principle: Black Women

June 5th, Breonna Taylor's Birthday–Day to call for justice for Breonna and uplift the #SayHerName movement.

June 19th: Education for Liberation Day

Principles: Black Families and Diversity

Juneteenth: Education for Liberation Day is a day to celebrate the struggle that brought down slavery and reflects on what must be done to win Black liberation.

Last Day of School: A Day for Self-Reflection

Review all 13 principles

On the last day of school, we ask that educators, students, parents, community people, and organizers lift up Reflection Day; a day to reflect on your year of antiracist teaching and abolitionist practices. This can be an activity that people do individually or within a group.

Conclusion

For the past five years, the organizers from the National Black Lives Matter at School have been able to build upon the work that Black grassroots organizers, past and present, have laid the foundation for with no institutional financial support or backing. This not only speaks to the passion, but the persistent and visionary spirit of radical Black educators and organizers who refuse to be deterred by the structural limitations designed to keep us hopeless. The seeds that were planted by the educators, parents, and community members from the Black Lives Matter at School Seattle events that led to the Philly Week of Action, then the forming of the National Black Lives Matter at School grew from the work of previous generations and will continue on for future generations.

References

- Anderson, T., 2020. From Philly with love: Black lives matter at school goes national. In: Jones, D., Hagopian, J. (Eds.), *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books, pp. 49–56.
- Au, W., Hagopian, J., 2018. How one elementary school sparked a citywide movement to make Black students' lives matter. In: Watson, D., Hagopian, J., Au, W. (Eds.), *Teaching for Black Lives. Rethinking Schools*, pp. 22–31.
- Garcia, L., 2020. How to talk to young children about the Black lives matter guiding principles. In: Jones, D., Hagopian, J. (Eds.), *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books, pp. 140–144.
- Garcia, L., Clark, R., 2020. NYC Teachers bring the Black lives matter at school week of action to the early years. In: Jones, D., Hagopian, J. (Eds.), *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books, pp. 134–139.

- Hagopian, J., 2020. Making Black lives matter at school. In: Jones, D., Hagopian, J. (Eds.), *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books, pp. 1–24.
- Jones, D., Hagopian, J. (Eds.), 2020. *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books.
- Khan-Cullors, P., Bandele, A., 2017. *When They Call You a Terrorist: A Black Lives Matter Memoir*. St. Martin's Press.
- King, L.J., 2014. When lions write history: black history textbooks, African American educators & the alternative Black curriculum in social studies education 1890–1940. *Multicult. Educ.* 22 (1), 2–10.
- Rogers, C., 2020. Organizing the national curriculum. In: Jones, D., Hagopian, J. (Eds.), *Black Lives Matter at School: An Uprising for Educational Justice*. Haymarket Books, Chicago, IL, pp. 200–207.

Young women and activism and dissident global citizenship

Lisa D. Weems, Miami University of Ohio, Department of Educational Leadership, Oxford, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: (re)framing young women and activism	363
Notes on transnational feminist approach	364
A snapshot of contemporary young women and activism—issues	365
Issues—violence against young women and women	366
Issues—reproductive justice, SOGI, and HIV/AIDS	366
Issues—racial justice activism	367
Environmental issues—climate justice & land sovereignty	368
Modes of activism among young women	368
Direct protest/action	368
Digital media and CyberActivism	369
Community collaboration/art installations	369
Dissent, dissidence, and (global) citizenship from below	370
References	371
Relevant websites	373

Introduction: (re)framing young women and activism

Young women are often used rhetorically to justify campaigns for larger political initiatives designated as democratic or progressive projects. This rhetorical framing positions young women as innocent, pure, and in need of protection (Sensoy and Marshall, 2010). In other words, young women are discursively constructed as the recipients or objects of global citizenship practices rather than active agents or subject-makers in cultural and political citizenship. Furthermore, racist, and imperialist narratives bifurcate young women into categories of those who need to be saved and those who save (Khoji-Moolji, 2018). Postcolonial feminists articulate how Western feminism perpetuates the narrative of “white women saving brown women from brown men” (Abu-Lugood, 2002; Spivak, 1988; Subedi and Maleku, 2021). Drawing from these critiques, this chapter seeks to (re)frame the position of young women within global discourses on democracy, citizenship, and social justice. Specifically, my account aims to provide a critical aperture in such discourses by illuminating how young women, as knowledge producers and cultural makers, participate in a wide range of activist initiatives that may be considered dissent.

In order to explore young women’s participation in global discourses on democracy, citizenship, and social justice, I want to unpack the relations between resistance, activism, and dissent. I argue that each is important, yet distinct, and deserve separate attention. I use the term resistance to describe an act, thought, or feeling that opposes another act, thought, or feeling. Resistance involves embodied dynamics that are historical, biographical, and geopolitical (Anzaldúa, 1987; Weems, 2015; Rhee, 2021). I assume that power is an expression of human agency where young women navigate specific ways of knowing and being through particular subject positions available to them through global and local discursive practices, structures, and relations (Foucault, 1984, 1989, 1991; Ong, 1996; Sandoval, 2000). More specifically, I use the term “resistance” to refer to acts that challenge, or contend with, particular “regimes of truth” as conceptualized in within disciplinary formations of power/knowledge (Foucault, 1984). As Foucault notes, power/knowledge is a product of language, institutional structures, subjectivity, and relations of power. Foucault suggests that institutional structures (such as schools, religion, politics, etc.) involve technologies of control and regulation as authoritative formations of disciplinary power/knowledge but which also produces the potential for disruption through resistance. In this sense, Foucault connects discursive, symbolic, and material conditions in theorizing power/knowledge that circulates through and among heterogeneous bodies, identities, talk, text, and images. Yet, as a productive and distributive force, resistance is possible through embodied transactions that challenge disciplinary and localized forms of power/knowledge. Thus, I employ Foucault’s emphasis on resistance within micro and meso-level contexts and places. Sometimes resistance is intentional; sometimes it may be an unconscious and/or visceral reaction to disciplinary habits and dispositions. The point is that is that resistance is neither inherently political nor made visible for public consumption.

In contrast, I conceptualize activism as physical acts, talk, text, and images that are *intentionally constructed within/through collectivist and/or community-based forms of resistance* that may occur through street protests, demonstrations, cyberactivism, etc., Grounded in particular socio-cultural, geopolitical, and economic contexts, activism “makes sense” through regional definitions, rituals, and forms of collective resistance that may be implicitly or explicitly designed for movement-building. In this way, activism involves an attunement to macro social and political sensibilities and practices. Yet, activism, like resistance, *does not necessarily signal a progressive oriented ideologies or invested in diversity, equity, and social justice agendas, especially given the growing rise of right-wing global populisms*. However, given my investment in a feminist transnational praxis, my analysis foregrounds young women’s activism for social justice based in global discourses on radical democracy.

I use the term dissent to refer to a particular type of active resistance that is intentionally constructed against some ideology, social practice, structure, etc., Perhaps the most common example of dissent is the policy of the US Supreme to allow Justices to express disagreement to a majority outcome of a legal decision. Given this etymology, I define dissent as those statements, images, physical and/or symbolic acts that are conscious and intentional. Dissent operates within, and responds to, an openly politicized decree, doctrine, policy, law and/or initiative *precisely* to raise awareness among a larger audience. Thus, while dissent always already operates within historical and socio-cultural spaces, it is designed to intervene in public and/or popular narratives/discourses involving cultural politics, community norms, and narratives of citizenship. Dissent reflects contested dynamics that traverse onto-epistemological, socio-cultural, and geopolitical boundaries and is aimed at symbolic as well as material change.

Novelist and activist Arundhati Roy characterizes the kind of urgent (and vigilant) disposition of transnational feminist praxis this way,

What we need to search for and find, what we need to hone and perfect into a magnificent, shining thing, is a new kind of politics. Not the politics of governance, but the politics of resistance. The politics of opposition. The politics of forcing accountability. The politics of slowing things down. The politics of joining hands across the world and preventing certain destruction. In the present circumstances, I'd say that *the only thing worth globalizing is dissent* (2002, n.p.).

Taking a cue from Roy, I argue that *dissent* is a *productive* form of power/knowledge/resistance that is both healthy and necessary for democratic societies toward achieving a more planetary vision of social justice and critical global citizenship (Andreotti, 2014).¹ Thus, my analysis documents the localized (yet traveling) forms of young women's activism and dissent against the intersectional oppression of young women across borders of tribal, cultural, and national citizenship. Grounded in critical discourses of education for empowerment, I believe that activism serves as a site of critical knowledge production and global citizenship (Curley et al., 2018).

I conceptualize feminist activism as acts, events, relationships, creative and cultural productions that are *intended* to promote critical consciousness, social change, political organizing, and/or movement building involving issues related to gender. More specifically, my analysis charts contemporary activism that might be construed as part of "transnational feminist praxis" (Grewal and Kaplan, 1994,99 2000; Alexander and Mohanty, 2010; Baksh and Harcourt, 2015; Sharoni et al., 2015). As such, I present some of the empirical international scholarship on young women and activism in formal and non-formal settings. The empirical data demonstrates how young women (*cis*, queer, *trans*) participate in strategic initiatives against gender-based violence, toward racial, reproductive, and environmental justice. I've selected these four areas to illustrate how young women engage in copious forms of social justice activism in transnational context. My analysis then concludes with a discussion of young women's activism as part of a more expansive worldview of education, pedagogy, and activism as global citizenship (Curley et al., 2018).

Notes on transnational feminist approach

Throughout this essay, I primarily focus on young women rather than girls. This focus is a result of my own experiences working with young women from diverse backgrounds in the United States, as well as my expertise in transnational scholarship on young women. My formal training as a language arts high school teacher in the US Midwest and community educator (with Planned Parenthood and the Youth Action Agency) was formative in understanding "young women" as a heterogeneous category nuanced by inequitable structures of racialization; economic and citizenship status; linguistic, cultural, and religious traditions; and educational opportunities. These experiences, coupled with my epistemological grounding in woman of color feminisms, crystallized my long-term interest and commitment to working with minoritized young women. Through active marginalization, members of a particular group become "minoritized" as not (quite) citizens—or "Others."

Post-colonial feminist theorist Gayatri Spivak coined the concept of "Othering" (1985) to describe the processes and outcomes of the marginalization of myriad persons or populations in a particular society (or geopolitical location). Othering is a form of cultural imperialism that bifurcates people into binary categories such as native/foreigner, good/evil; primitive/civilized, etc. Imperialist/racist/heteronormative constructions of the Other are embedded into governmental, corporate, and non-profit policies and discursive practices of subjectivity and citizenship. These cultural binaries and hierarchies lead to the institutionalization of Othering and minoritization. In these terms, young women across the globe are minoritized around issues of gender identity (Raub et al., 2016) and sexual orientation (Suen and Chan, 2021), whether they are *cis*- or transgender, because of prevailing heteropatriarchal logics, affects and structures (Minh-ha, 1989; Mohanty, 2003; Alexander, 2005; Reddy et al., 2009; Pin, 2018).

Drawing from scholarship in feminist transnational studies, this chapter focuses on the embodied and creative forms of activism and dissent deployed by young women in (g)local movements against colonial, capitalist, white supremacist, and heteropatriarchal social arrangements. My use of the term "transnational" and "young women of color" are not simply found descriptions; rather, I deploy their use in ways to suggest that myriad complex forms of multiracial and international solidarity are possible and desirable

¹Written in 2002, Roy's statement on dissent takes for granted that the politics of resistance works against domination and oppression. However, given the increase of cultural and political conflict within and across national boundaries over the past 20 years, it is no longer safe to assume that even resistance is automatically tied to democratic or social justice-oriented models of activism and dissent.

in part through an active rejection of the trope of “global sisterhood” (Banda, 2015; Mohanty, 2013). According to Mohanty, the discourse of global sisterhood attempts to center feminist struggles in the Global North articulated through institutionalized forms of feminist projects (such as North American women’s studies literature, feminist NGOs, etc.). In contrast, transnational feminist praxis can be thought of as a loose series of localized responses to “scattered hegemonies” (Grewal and Caplan, 1994). In other words, scattered hegemonies of patriarchy mean that women experience gender-based oppression differently, and in turn, produces activism that is both situated yet connected to a larger agenda that links women’s rights’ with ‘universal human rights’ even while the categories of women and the universal remain “sites of incessant contest and resignification” (Butler, 2002).

Transnational feminist praxis explores how sex and gender-based discrimination and oppression are further compounded by other forms of minoritization or what Patricia Hill Collins (1990) refers to as interlocking systems of oppression. For example, young women of African descent women are minoritized within and across national borders due to the legacies of the transatlantic slave trade, white supremacy, and contemporary global anti-Black racism. Similarly, since 9/11, there is an uptick in violence against ethnic and religious minorities within and across nation-states fueled by populist nationalism and weaponization of Islamophobia, xenophobia, im/migration politic in the name of the “war on terror” (Jones et al., 2017). More recently, there is a rise in anti-Christian sentiment and persecution in countries such as India, China, Nigeria, and North Korea despite the global dominance of Christianity (Sherwood, 2021). Systemic policies, and practices fuel discrimination and violence driven by “ethno-racial” politics within, across state, national and continental borders (Pon et al., 2013). The particularities of these interlocking systems of oppression are named and experienced differently across geopolitical borders (see discussion below). Nonetheless, Crenshaw’s notion of “intersectionality” is a helpful way to characterize the exponential ways in which poverty, patriarchy, and cultural imperialism work to both disenfranchise and demonize those who are Othered (Crenshaw, 1991).

This weaponization of fear and hate against the very bodies, identities, and communities of young women that are actively minoritized leads to an existence of “living hells” driven by “carceral capitalism” (Wang, 2018). While young women may be “protected” by national and international edicts and proclamation of “rights,” their meaningful participation in the classroom, workplace, government, media, medical, scientific, and technological dimensions of subjectivity and citizenship are effectively diminished and/or nullified. It is not simply that minoritized young women are overlooked or underserved. Their real and imagined lifeworlds are constrained by the contemporary landscapes of global and racial capitalism, militarized necropolitics, racial and religious extremism, and intersectional violence.

Against this backdrop, young women (especially minoritized young women) actively resist or contest the material conditions and cultural politics of their local community in a myriad of forms (physical demonstrations, boycotts, blogging, digital culture jamming, etc.). Young women participate in formal and informal forms of associations and organizing to address gender-related issues at the local, regional, and global arena. For the remainder of this chapter, my analysis presents a curated archive of young women’s activism based on two criteria. First, I foreground young women’s activism that operates from an intersectional perspective that allows for the analysis of interlocking systems of oppression such as race, gender, social-economic status, sexuality, and youth. While these categories of identity and difference are historically produced social constructs, they produce material consequences that are embedded in political and economic structures, as well as institutionalized cultural ideologies and policies. Thus, my primary interest is young women’s activism that challenges a “politics of domination” (Hooks, 1986) across multiple sites of knowing and being. Secondly, I foreground forms of activism that directly center issues facing young women as *embodied subjects* (such as violence based on sexuality and gender identity (SOGI), as well as struggles for justice where young women are disproportionately affected (such as reproductive, and climate-based justice).

In the next section, I provide a skeleton of some issues and modes of young women’s activism. The scope of my account is limited to initiatives related to four major areas; (1) violence against women; (2) reproductive justice (including issues related to Sexuality and Gender Orientation as well as HIV/AIDs); (3) racial justice; and (4) activism for climate justice. These four areas provide concrete examples of intersectional and transnational contemporary activism among young women. I conclude the section on issues with a provide a brief discussion of the need to shift discourses on young women and empowerment from “access to equity in schools”—toward a vision of activism as dissident global citizenship.

A snapshot of contemporary young women and activism—issues

One way to track young women’s global activism is to look for international registries of political and cultural organizations dedicated to the young women’s rights, equity, and justice. According to a survey conducted by FRIDA (a research arm of the Association for Women’s Rights in Development (AWID, n.d.)), Young Feminist Organizations (YFO) are relatively new and diverse but integrated into larger coalitions for women’s rights. The FRIDA report identified the following areas of primary concern for YFOs: “gender-based violence; sexual reproductive health and rights; and political and economic empowerment for women.” Gender-based violence was a key driver for young women’s activism across geographic locations whether in the form of sexual harassment, rape, domestic abuse, hate crimes, murder, or femicide. However, additional areas of immediate concern were also associated specifically with geopolitical locations. For example, in Sub-Saharan Africa, economic empowerment and HIV/AIDS prevention are pressing issues, while LGBTQI rights were reported as the most pressing challenges facing YFOs in countries within the Americas as well as the Middle East and Northern Africa. Young women in YFOs across Asia (Central, South, Southeast, and East), reported feeling unsafe to participate in formal social movements due to surveillance and physical intimidation by governmental and religious extremism.

Issues—violence against young women and women

The World Health Organization (WHO) considers violence against women to be a violation of women's rights and thus, human rights more broadly. According to their 2021 executive report, 1 in 3 women in the world experience either intimate partner violence or non-partner sexual violence or both. Thus, one of the most predominant issues among young women activists is to end one or more forms of violence against young women and women. Activists characterize violence in broad but concrete terms including but not limited to physical and emotional coercion and abuse, sexual harassment, and rape. In addition, young women are particularly affected by femicide, forced sterilization, forced marriage, honor killing, and state-sanctioned gender-based violence (such as the crisis of Missing and Murdered Indigenous Women, Girls and Two-Spirit persons in North America and the governmental failure to produce justice for their families). Furthermore, as Truong notes, neoliberal globalization has enflamed the rise in sex trafficking and commercial surrogacy (2015).

Organizations such as the Rape, Abuse, and Incest National Network in the United States provide direct services such as the sexual abuse hotline as well as act as a clearinghouse of resources for survivors, researchers, and policymakers. Similarly, the African American Policy Forum gathers and distributes contemporary statistics of violence against Black, and Latinx young women in the United States. Individual young Black women, such as Nupol Kiazulu and Naomi Wadler, both from the United States, are part of a new cadre of young feminists committed to intersectional activism. For example, during her speech at the 2018 March for Our Lives, Wadler detailed the impact of gun violence on Black and Latinx youth. As Wadler noted, young women of color are acutely aware of their precarity due to their multiple layers of oppression and physical proximity to gang-related, anti-Black racist, and police-based violence. Consider the tragic case of Oluwatoyin "Toyin" Salau, a young Black woman who organized a protest at the Tallahassee (Florida) Police Department after Tony McDade, a transgender man, was fatally shot by the police. As a result of her visibility, however, Salau was subject to physical and cyber-bullying and ultimately killed in a "domestic dispute." Unfortunately, Salau's case is far from the exception; organizing against gender-based violence, especially as a young woman of color, or as a queer or *trans* youth of color, makes one a target for further attack. Thankfully, there are a growing number of community-based organizations whose approach to ending GBV is to focus on creating space for survivors to process trauma and heal alongside other survivors (ALWH). Moreover, organizations such as the Chicago-based Girl/Friends Institute cultivates "the empowerment of young people that amplifies the voices and creative visions of those most vulnerable to violence: girls and women of color" (ALWH). Graduates of the 5-week summer Girl/Friends Institute then take peer education and social justice campaigns to local schools and community events.

Issues—reproductive justice, SOGI, and HIV/AIDS

Young women across the globe are at the forefront of organizing for comprehensive reproductive justice. Reproductive justice involves access to safer sex education as well as access to birth control, abortion, and prenatal care. For some young feminist activists, reproductive justice also includes access to affordable childcare and programs that de-stigmatize sex work and LGBTQ2S identities. Reproductive justice activists advocate for women's bodily autonomy—defined as the ability for individuals to make decisions about their bodies without restriction, fear, or coercion (Rodríguez, 2021; UNFPA, 2021). According to the 2021 State of the World Population Report, approximately 55% of women globally (despite a nation's wealth or status) do not possess bodily autonomy due to interpersonal relations, community-based norms, and/or legal restrictions. In response, the Center for Women's Global Leadership launched an annual initiative to end gender-based violence in 1991 called 16 Days of Activism. Over the past 30 years, over 6000 organizations across 187 countries have participated in the initiative that links the International Day Against Violence Against Women (November 25) to the International Human Rights Day (December 10) and include World AIDS Day (December 1st). 16 Days of Activism has spread globally to include partnerships with IGOs, NGOs, and local communities through street activism, press campaigns, and digital messaging. More recently, corporations and other organizations within civil society have offered their support such as the 2021 theme "Orange the World" through orange-colored lights and logos on buildings and billboards.

In the FRIDA report, one of the primary areas of concern were issues regarding sexual orientation and/or gender identity, and specifically, those particularly affecting Lesbian, Gay, Bisexual, Transgender, Queer, or 2Spirit persons. However, issues of sexual orientation and gender identity (SOGI), like educational issues, vastly differ from country to country. For example, 69 countries currently criminalize homosexuality including some that are accompanied by the death penalty. Most of the 69 countries that legally criminalize are in Africa and other parts of the Global South.

Beyond the criminalization of homosexuality, non-LGBTQ inclusive policies are scattered (distributed) across the Global North and South. Lack of protection in social policies (such as employment, housing, education, etc.) leads to harassment, violence, and hate crimes directed toward queer, *trans*, and nonbinary persons (Suen and Chan, 2021; Mampane, 2020). Across the globe, Black *Trans* women are not only highly targeted for gender and race-based violence; crimes against them are often left uninvestigated or unprosecuted, compounding the trauma for the person, family, and larger LGBTQ community.

In terms of intersectional activism on SOGI, I commend organizations or networks that proffer "support not stigma" for folks who are multiply marginalized such as LGBTQ2S youth who are sex workers (Maggies of Toronto or AF3IRM) or undocumented (Undocuqueer) or homeless (Benn, 2017). It is no surprise that these examples of highly visible community-based supports for LGBTQ2S youth are based in the Global North and reflects how agency and activism are a function of the material conditions that make possible certain political claims and forms of cultural citizenship (Gopinath, 2005; Queisada et al., 2015; Bapat, 2021). For example, Gopinath argued about the "impossibility" of the lesbian subject in 2005 India given the rise of Hindu

nationalism. Just 15 years later, the members of Queer Crescent, an organization based in Los Angeles, California, created a menu of cultural and religious resources for LGBTQI Muslims across the Islamic world. In other words, geopolitical privilege allows young activists to build more visible and funded community-based networks to support SOGI issues or issues that affect LGBTQ2S youth. Such community-based networks are likely to exist across the globe, yet given the political and cultural climate, they may be underground.

In many parts of the world, young women are disproportionately affected by HIV/AIDS and other infectious diseases due to inadequate access to safer sex education, preventative health care, and resources to affirm and accommodate living with HIV/AIDS (The Global Fund, 2019). As of December 2021, 53% of the 37.7 million persons living with HIV are women (WHO, 2021). So, while the numbers of infected persons reflect a “gender balance,” it is important to note that the most vulnerable populations include sex workers, gay men, men who have sex with men, transgender women, and IV drug users. Put simply, transmission is correlated with poverty, sexual violence, and transphobia. Thus, HIV/AIDS activists are currently working to decrease discrimination and social stigma and improve political, social, and economic conditions.

One area that has been significantly impacted by HIV/AIDS is Sub-Saharan Africa, where it is estimated that every week, 7000 girls and young women between the ages of 15–24 become infected.² In turn, young women in Sub-Saharan Africa are at the forefront of HIV/AIDS advocacy and services for children, youth, and adults (Yola et al., 2019; Hollis, 2019). Besides organizing for direct access to medical services, young women such as Jane Emmanuel (Tanzania) and Gogontlejang Phaladi (Botswana) are working to secure more comprehensive health services and education including a newer trend to include the psycho-social impact of HIV/AIDS on individuals, families, and communities. Emmanuel serves as a peer educator and support group facilitator for HIV + teens (Ped-AIDS, 2020). Phaladi, at just 5 years old, created the Gogontlejang Phaladi Pillar of Hope Organization (GPPH) as a charitable organization to support youth who are infected and/or orphaned by HIV/AIDS. With the help of her parents, Phaladi grew GPPH into a nongovernmental organization that focuses on human rights and gender equality. Since then, Phaladi has taken up leadership positions (such as a global ambassador or regional director) to both offer her expertise as well as serve as a visible role model for young African women to become entrepreneurs, philanthropists, and professionals. This type of advocacy is often linked to IGOs, NGOs, and other non-profit organizations and rely on their financial sponsorship for maintaining everyday services. Thus, the visions, goals, and aims of such organizations are often framed through international reform initiatives such as the United Nations Sustainability Development Goals (SDS) or the Partnership for Maternal, Newborn & Child Health (PMNCH). In contrast to the NGO model, there are also grassroots organizations mobilized by and for local HIV + individuals and coalitional networks such as the Brazil RNP+ and the Durbar Collective of sex workers in Calcutta, India (Chose, 2012). Both organizations either meet in person or connect through closed groups within social media platforms such as WhatsApp™. Both aim to empower HIV + individuals yet provide a level of discretion and protection within a larger context of political and religious intolerance where stigma has been weaponized to create place-based hotbeds of transmission and violence.

Issues—racial justice activism

Perhaps the most visible public campaigns led by young women in the 21st-century center racial justice: including countering anti-Blackness, colorism, casteism, and the particular issues that affect women at the intersections of class, gender, and race. In North America, for example, young Black women have been at the forefront of sustaining and progressing the Black Lives Matter (BLM) movement through local direct actions, protests, social media campaigns, and classroom projects (Troutman and Jimenez, 2016). Institutes like the African American Policy Foundation, founded by Dr. Kimberle Crenshaw, have played a pivotal role in creating a sustained movement for research and advocacy regarding young Black women (especially with regards to violence, education, and the Prison Industrial Complex (PIC) in the United States.

Often young women activists espouse a platform that understands power, privilege, and oppression as intersectional, institutionalized, and relational. Yet there are complexities to these claims. First, we know that all knowledge is partial and situated and that persons who benefit from hegemonic relations often do not recognize their limited reflexivity. For example, US Black Feminist scholar Jennifer Nash cautions that Black women are disappeared even as “intersectionality” becomes institutionalized in organizations of/for women. For Nash, the ongoing structure of institutionalized racism often leaves Black women either outside or marginalized within discourse practices of “Black Feminism” in academic circles.

A second complexity of the espoused platform of intersectionality is the important distinctions between various geopolitical definitions of which social categories are of most importance in terms of analyzing systems of power, privilege, and oppression. For example, in many countries in the Global North (United States, Canada, Australia, and the United Kingdom) young feminist activism draws from the US Black feminists such as Kimberle Crenshaw, Patricia Hill Collins, and bell hooks, to think about how the “holy trinity of oppression: race, class, and gender” particularly affects young women of the African diaspora (Johnson, 2018). Whereas, in the Global South, it is not uncommon to see explicit campaigns that connect femicide to Western imperialism, governmental censorship, militarism, environmental degradation, and economic instability (Antrobus, 2015).

To be sure, young activists draw from local and historically situated frames of difference (such as ethnicity, religion, caste, citizenship, and/or language) to understand the symbolic and material practices of power, privilege, and oppression they experience. In turn, activists mobilize by and through the rhetoric and platforms of local leadership and movements for equity and justice. For

²<https://www.theglobalfight.org/women-girls-and-hiv/>.

example, young feminist organizing in Northern and Western Africa is especially attuned to differences in ethnicity, religious affiliation (Arab, Muslim, Christian, Indigenous), and geographic status (rural/urban). Whereas young feminist organizing in Sub-Saharan, or South Africa often includes campaigns against gender-based violence, homophobia, and HIV/AIDS stigmatization as central to the health and wealth of the nation and its young population (Zerai et al., 2017) And, as Sharoni et al. (2015) note, (US-based) feminism can learn from the Boycott, Divestment, and Sanctions (BDS) movement for justice in Palestine to imagine and create issue-based coalitions (2015). This situated understanding of intersectionality, and social justice echos the transnational solidarity work of the BDS anti-apartheid movement in South Africa during the 1980s and the current anti-occupation alliance between Black Lives Matter and Palestinian activists (Davis, 2016; Salam, 2021). As Armstrong notes, a critique of Western feminism along the axis of South-South commonality has at times allowed for women's movements to take the shape of a provisional anti-imperialist feminist solidarity toward a "leftist women's internationalism" aimed to mobilize rural and urban peasants (2016, p. 308).

Environmental issues—climate justice & land sovereignty

Greta Thunberg (Sweden) may be the world's most identifiable young activist regarding climate change and environmental justice. In 2018, Thunberg then 14 years old, catalyzed an international "School Climate Strike" (SCS) where young people across the globe protested the lack of attention and action by elected officials regarding the deleterious effects of climate change. Thunberg has leveraged media visibility to publicly confront leaders from European parliaments and the United Nations and for the most part, received positive reception. And it is important to recognize that millions of young activists across the globe are also deeply involved in informal and formal movement-building for climate and environmental justice (especially Indigenous youth across the Americas, Asia, and Africa), sometimes with criminal prosecution and penalty. One such prominent figure is Disha Annappa Ravi (India) who mobilized the Fridays for Future (FFF) network to support the 2020–2021 Indian Farmers protests. Ravi was arrested for sedition for her involvement including the use of social media to create awareness about the protests. Similarly, several of the young women who took part in the frontline of the anti-Dakota pipeline (#NODAPL) initiative at the Standing Rock reservation in North Dakota, USA, faced police brutality and criminal charges simply for serving as water protectors as they are known in their community. As Tokata Iron Eyes from Respect our Water notes, indigenous leaders and tribal citizens are the most equipped to promote and educate the kind of activism environmental justice necessary to stop planetary degradation and recalibrate the inter-relationships between humans, animals, earth, sky, and water.

Modes of activism among young women

Young women engage in myriad modes of activities in the name of social change, equity, justice, and movement-building ranging from informal consciousness raising, education and formal leadership development. Whereas, the FRIDA report captured many of the YFOs that focus on capacity building, many of the young women discussed in this chapter could be characterized as young women who organize direct service, actions, demonstrations, or protests. This distinction is not intended to pit leaders against each other; rather my hope is to draw attention to the depth and breadth of activism among young women. Young women who are linked to colleges, universities, or other institutions that foster scholarship may be able to engage in participatory research with other young women, and women. These forms of university-community-based youth-led organizations can be extremely useful for two reasons. First, these partnerships often put the direct needs of the young women at the center of the research project. Second, such partnerships create the opportunity for capacity building by equipping young feminist leaders with research skills and building social capital (Edwards-Jauch, 2012; Wang, 2018; Bailly et al., 2020).

Direct protest/action

Within and across national and international boundaries, young women activists mobilize protests to raise awareness about state-sanctioned racialized, gendered, and sexual violence against Indigenous, Black, and racialized young women (Danforth, 2011; Garita, 2015; Kurian et al., 2015; Pastor and Hernandez, 2021; Phillips, 2020; Truong, 2015). From China to Mexico to South Africa and India, young women form grassroots-based collaborations for direct action. The most common form of direct action is a street march or protest to draw attention to sexual or gender-based violence (Molestane, 2018). For example, Zee Thomas, 15, of Nashville, Tennessee, worked with five of her friends to lead a march to denounce the police killing of George Floyd (Bennett, 2020). What she thought would elicit a small gathering turned into a crowd of 10,000. For sure, the Black Lives Matter movement has sparked the fire for young Black women to take their rightful place as the next generation of leaders for social change.

Young feminist protests or direct actions can include creative demonstrations, teach-ins, and workshops to illuminate the embodied nature of heteropatriarchy and the intergenerational effects of femicide. For example, young feministas in Latin America have responded to a five-year surge in femicides in multiple ways. It is estimated that 20 million women in Mexico participated in a work strike during the month of March 2020. Young feministas also staged mass protests at highly visible locations such as the Attorney General's office and covered the National Palace with red paint. Young feminist organizations such as Circulo Violeta (Mexico) and Colectiva Feminista in Construcion (Puerto Rico) combined direct tactics of traditional political activism with cultural and creative forms of resistance in the form of ceremonial remembrance (Chavez, 2021). It is not uncommon for protests to include

singing, dancing, spoken word, mural painting, and graffiti, as well as other sacred forms of ceremony across the Americas Caribbean to oppose and survive centuries of colonial rule. First Nations, Aboriginal, Native, and Indigenous leaders understand the importance of combining culture and creativity with the ongoing fight for justice and sovereignty. In *The Winter We Danced*, Sylvia McAdam (Saysewahem) notes, “our history is still unfolding; it’s led by our songs and our drums” (2014, p. 67). Similarly, a new generation of “image warriors,” such as Erin Konsmo (a self-described Prairie Queer of the Métis Nation) combine Indigenous land-based materials with Western art practices and technologies to promote awareness and connection through the online circulation of what might be considered protest art.

Digital media and CyberActivism

Digital technologies afford opportunities for radically conceptualizing how we think about the self, community, public as mediated spaces ripe for participatory action, and radical disruption of hegemonic images and narratives (Ong, 2006; Boler and Nitsou, 2014; Weems, 2009, 2014, 2018; Al Zidjali, 2017). Online activism, culture jamming, and digital dissent have proliferated in the 21st century with the growth of global infrastructures in information and communication technologies, as well as access to personal computers, and mobile devices. Greater access to the internet has allowed young women to produce and share content related to intersectional violence and oppression as it occurs in local settings (Mai and Laine, 2017; Schuster, 2013). In turn, activists have utilized the internet to connect with others to create solidarity and circulate testimonios, counternarratives, and strategies for action (Kimball, 2019; Tanner et al., 2020; Castilho and Romancini, 2017). For example, Janna Jihad and Ahed Tamami, two young women from the West Bank, utilize(d) Instagram, Snapchat, Facebook, and Twitter to report ongoing airstrikes and confinement of Palestinian citizens. Like other young female Palestinian women, Jihad and Tamami demonstrate their commitment to independence, activism, and social justice by documenting and drawing global attention to Israeli occupation and how they thrive under conditions of perpetual war, apartheid, and settler colonialism.

Some of the most successful examples of contemporary social movements that grew from effective online mobilization are #Black Lives Matter, #NuEnos, #MeToo, and #IdleNoMore. Each of these initiatives began with a single Twitter post that grew by instant recirculation to become crowdsourcing. The phenomenon known as “hashtag activism” can be traced back to the Arab Spring uprisings in the 2010s when governments across the Middle East cracked down on censorship and media coverage (Moss, 2012).

As Jackson et al. (2020) note, hashtag activism allows for “networks of racial and gender justice” advocates to connect, share information, organize, and document actions. For many people around the world, hashtag activism allows for social movements to flourish despite the rising climate of political and religious (conservative) extremism. Yet, it is important to remember that social media can also function as surveillance or to transmit misinformation. As Jillian York (2021) notes, the move to online or digital technologies as a platform for movement-building means that corporate definitions of free speech and acceptable content now operate alongside governmental standards of censorship and technologies of surveillance to uphold them.

Community collaboration/art installations

Besides street protests and media campaigns, another form of activism involves community-based cultural events, exhibits, or collaborative art installations (Matherly, 2013; Desai and Abeita, 2017). One such collaboration is the group Walking with Our Sisters, founded in 2012 by an intergenerational group of First Nations and Métis women to raise awareness of the violence against Indigenous women and the alarming silence among politicians, policymakers, and the media. The art installation consists of over eighteen hundred vamps (the top part of a traditional moccasin) hand-beaded by more than thirteen hundred artists from across North America. Many of the vamps are beaded together in community in activities that integrate the learning of traditional arts, cultural forms of association such as intertribal and intergeneration dialogue, sacred reflection, and raising political awareness. The vamps are created in intergenerational circles of First Nations, Métis, and Indigenous women and men. Similarly, natives and non-natives work together to prepare, transport, and install the exhibit to ensure it retains a spiritual and culturally appropriate form of display. Since 2012, the Walking with Our Sisters exhibit has traveled across provincial and tribal borders due to the collective power of crowdsourced funding. As they write, “There is power in numbers, and there is power in art” (WWOS, 2014).

One of the most potent aspects of the WWOS project (and similarly the Red Dress Project) is that it links remembrance, local funds of knowledge, cultural practices with political aims that are community oriented. Like much of the participatory research with young Indigenous women, the aim is to celebrate and sustain the resiliency of the group as much as the individual (de Finney, 2017; Dhillion, 2015). Much of the highly visible feminist scholarship on young women and activism conflates cultural resistance with intentional opposition to formal political authorities (Fine, 2014; Taft, 2011). While this slippage can be problematic, it, nonetheless, makes possible a linkage between cultural resistance and political activism.

Indigenous scholarship on young women and activism makes explicit the connection between well-being, community organizing, and cultural revitalization as collective forms of citizenship toward sovereignty that is gendered, sexed, and tethered to body autonomy (Simpson, 2013; Dhillion, 2015; Lesperance, 2016; de Finney, 2017; Monchalin et al., 2016; Belcourt, 2020). Furthermore, this political work often involves intimate labor in that it focuses on collective and intergenerational notions of subjectivity and culture. Tribal elders and scholars of Native/First Nations studies recognize (and highly recommend) community-based practices of memory work, creative scholarship, and artistic production as legitimate forms of political participation and activism. Thus, *citizenship from below* emphasizes collective and intergenerational (but not necessarily biological) notions

of kin, culturally safe education, community-based cultural practices, and organized demonstrations against multiple forms of violence. I turn now to a more in-depth discussion of dissent, dissidence, and the potential for global citizenship from below.

Dissent, dissidence, and (global) citizenship from below

Despite their relative absence in official archives of political activism, young women take up, deploy, and subvert dominant discourses of subjectivity, culture, belonging, and citizenship. My interest is to investigate "*citizenship from below*" (such as grassroots or youth-led activism by foregrounding how young women of color play a key role in feminist, anti-racist, de/colonial, and transnational activism (Weems, 2014, 2018)). As I have shown, young women activists challenge contemporary efforts of neo-colonialism related to reproductive justice, land use, militarization, the corporatization of higher education, loss of Indigenous language and knowledge, and gendered violence (Walking With Our Sisters [WWOS], 2014). Through divergent strategies, these practices operate toward a vision of global citizenship through education, peace, and interconnectivity (Cristaldi and Pampolini, 2016).

In the beginning of this chapter, I argued that dissent embodies political critique on a larger scale, but which entails *particular* audiences being attuned to historically situated issues within modern contexts of mediated cultural productions (Maira, 2010; Mai and Laine, 2017; Weems, 2018). For example, Bhatia (2021) illustrated how a 101-day sit-in protest by a few Muslim women in a small locality in Delhi (Shaheen Bagh) catalyzed a national network of online dissent against the 2019 Citizenship Amendment Act (CAA). According to Bhatia, the ruling Bhartiya Janata Party Act effectively weaponized Hindutva ideology and demonized immigrant Muslims as potential "threats" to national security (1). Bhatia analyzed how participation in both offline and online protests and activism that position female Muslim body as a site of power/resistance and dissent. In this example, (mass) media served as a vehicle to construct certain bodies and communities as violent for purposes of regulating citizenship and controlling mobility. Yet in turn, digital media also served as a vehicle to produce bodies and communities of dissent through social media through communication about and documentation of the #womenofshaheenbagh and the handle @Shaheenbaghoff1. Thus, dissent is not simply about one exercising one's agency within a limited horizon of actions but is actively designed to make a statement and intervention to a larger, broader public regarding those limited conditions and their lingering effects. In this way, dissent is about protesting limited and oppressive conditions, exercising agency through action, as well as participating in an intersubjective struggle for "the non-coercive re-arrangement of desires" (Spivak, 2004, p. 526).

Sparks theorizes dissent as a radical relational practice that involves an ongoing struggle to "reconstitute the very boundaries of the political" (1997, p. 75). It is this ongoing position of dissent, particularly by members who are marginalized as unequal members of society, that constitutes a courageous form of dissident citizenship. Sparks writes:

I conceptualize dissident citizenship as the practices of marginalized citizens who publicly contest prevailing arrangements of power using oppositional democratic practices that augment or replace institutionalized channels of democratic opposition when those channels are inadequate or unavailable. Instead of voting, lobbying, or petitioning, dissident citizens constitute alternative public spaces through practices such as marches, protests, and picket lines; sit-ins, slow-downs, and cleanups; speeches, strikes, and street theater (1997, p. 75).

For Sparks, dissident citizenship involves creative oppositional practices by folks who are often excluded from formal institutional politics and employs courage to challenge both the content and form of citizenship. Thus, dissident citizenship illustrates the performative and discursive aspects of power and cultural politics (Darder, 2011; Butler, 2015). Dissident citizenship is performative in that bodies animate the affective dimension of belonging and exclusion and the disjuncture between these spaces. Dissident citizenship is also discursive in that actors are both deconstructing and re-inscribing multiple scripts laid out in political, juridical, and socio-cultural narratives crafted by and for historically situated agents (Sparks, 1997, p. 87).

The model of dissident citizenship, while useful in theorizing young women's activism and dissent, may also be limited in understanding citizenship as subject making (Ong, 1996) and/or the emphasis on sovereignty among indigenous, decolonial and transnational feminist activism (Simpson, 2013). First, I remain persuaded by Caribbean theorist M. Jacqui Alexander (1994) that "not just any(body) can be a citizen," given the onto-epistemological assumptions of Western philosophy. In other words, the very constructs of self/Other and citizenship are entrenched in colonial discourses that render some bodies human and others nonhuman, as well as classify and regulate ways of thinking and being into bordered categories of intelligibility and desirability (Wynter, 2003; Cankar and Maira, 2005; Somers, 2008). Furthermore, vying for entry into the political realm via official citizenship may not only be impossible for young women of color, but it may also be undesirable in favor of more sustainable and ethical forms of personhood, belonging, and governance.³

Thus, we might conceptualize young women's activism and dissent as a series of "scattered belongings" in response to "scattered hegemonies" of heteropatriarchy, global racism, and colonial capitalism. Understanding activism and dissent as scattered

³It may be helpful here to remember Mohanty's caution to challenge the "feminist individualist subject of much of liberal feminist theory" free from the legacies of imperialism (2003, p. 80). Thus, I am cognizant of how transnational feminist analysis (including my own) tends to re-center the Global North, the United States, and its nationalist mythologies and demands of assimilation, even as I attempt to interrupt this analytic (Swarr and Nagar, 2012; Wright, 2000; Sharoni et al., 2021).

belongings requires multi-layered and nuanced investigations of how young women interpret, navigate, and contest the “politics of domination” in their everyday lives. As such, they work *within and against* transnational feminist discourses of consciousness, praxis, and empowerment to reducing social inequality (Bhavnani, 2006; Africa-Young, 2021; Maira, 2017). Like their elders, young women activists practice dissent as an expression of dissident global citizenship on multiple levels of scale. Yet, it is important to name and claim young women’s global dissent as “small, small incidents” (Van der Tolen and Bol, 2011) of critical and inter-generational social capital that is built by “witnessing” and collaboration (Liebert and Thompson, 2015, p. 94). To return to Roy’s statement, I agree that we need to globalize dissent. What is necessary, then, is a flexible model of dissident global citizenship, one that operates to build and sustain larger networks (or archives) of transnational feminist praxis constituted by human, material, discursive, and digital resources ready for situated deployment against the politics of domination.

Dissident global citizenship is not a pre-determined set of ideological beliefs of strategies that are universally imagined or practiced. Rather, dissident global citizenship *requires framing and theorizing* young women as active participants in intersectional struggles against patriarchy, racism, colonialism, global capitalism, homophobia, and transphobia. Young feminist activists’ foreground embodied protest, digital dissent, and creative forms of socio-cultural and political citizenship as they are highly attuned to the global, affective, and (mediated) dimensions of cultural politics and social transformation. Digital media and other popular cultural productions afford the current generation of young women’s activism to connect and amplify creative dissent and dissident global citizenship through the production, circulation, and distribution of gifs, memes, blogs, videos, sampled beats, etc., that challenge the permeation and permanence of multiple layers of intersectional violence (Weems, 2015; Kumari, 2021). Our work as (elder) transnational feminists is to seek out and sustain opportunities for connection, collaboration, and solidarity across differences to transform gendered relations of power and oppression toward the eradication of violence and reproductive, racial, environmental, and anti-colonial justice.

This essay explores recent issues, frameworks, and social practices involving young women, and activism in global context. My analysis is not intended to be a comprehensive review of the existing scholarship about young women and activism on an international scale. Such an endeavor is an impossible task for both logistical and conceptual reasons. Using a feminist transnational approach, I have demonstrated how young women across the globe both utilize and challenge Western modes of activism. Moreover, I have illuminated how young women enact dissent against a “politics of domination” as a form of critical global citizenship that we might think of as dissident global citizenship.

References

- Abu-Lughod, L., 2002. “Do muslim women really need saving?” Anthropological reflections on cultural relativism and its others. *Am. Anthropol.* 104 (3), 783–790. <https://doi.org/10.1525/aa.2002.104.3.783>.
- Africa-Young, N., 2021. From the personal to the political: young, Filipino, and feminist. *Asian J. Wom. Stud.* 27 (1), 109–129.
- Al Zidjaji, N., 2017. Memes as reasonably hostile laments: a discourse analysis of political dissent in Oman. *Discourse Soc.* 28 (6), 573–594.
- Alexander, M.J., 1994. Not just (any) body can be a citizen: the politics of law, sexuality, and postcoloniality in Trinidad and Tobago and the Bahamas. *Fem. Rev.* 48 (Autumn), 5–23.
- Alexander, M.J., 2005. *Pedagogies of Crossing: Meditations on Feminism, Sexual Politics, Memory, and the Sacred*. Duke University Press, Durham, NC.
- Alexander, M.J., Mohanty, C.T., 2010. Cartographies of knowledge and power: transnational feminism as transnational praxis. In: Swarr, A.L., Nagar, R. (Eds.), *Critical Transnational Feminist Praxis*. SUNY, Albany, NY, pp. 23–45.
- Andreotti, V., 2014. Soft vs. critical global citizenship education. In: McCloskey, S. (Ed.), *Development Education in Policy and Practice*. Palgrave Macmillan, New York, pp. 21–31.
- Antrobus, P., 2015. Dawn - the third world feminist network: upturning hierarchies. In: Baksh, W., Harcourt, W. (Eds.), *Oxford Handbook of Transnational Feminist Movements*. Oxford University Press, Oxford, UK, pp. 159–187.
- Anzaldúa, G., 1987. *Borderlands/La Frontera: The New Mestiza*. Aunt Lute Books, San Francisco, CA.
- AWID (Association for Women’s Rights Development and Young Feminist Activism Program), (n/d). Brave, Creative, Resilient: The Global State of Young Feminist Organizing. Retrieved August 20, 2021, from https://youngfeministfund.org/publications-and-resources/?_sft_issues=activism.
- Baily, S., Wang, G., Scott-Lavino, E., 2020. The inheritance of Activism: does social capital shape women’s lives? *Girlhood Stud.* 13 (2), 86–102. <https://doi.org/10.3167/ghs.2020.130208>.
- Baksh, R., Harcourt, W. (Eds.), 2015. *The Oxford Handbook of Transnational Feminist Movements*. Oxford University, London, UK.
- Banda, R.C., 2015. Every woman has a story: narratives of Southern African women. In: U.S. Higher Education. Unpublished Doctoral Dissertation. Miami University, Oxford, OH. Retrieved from: http://rave.ohiolink.edu/etdc/view?acc_num=miami1429373672.
- Bapat, S., June 8, 2021. Uplifting Gender Justice Activism. *Queer Crescent*. Proteus Fund. Retrieved from: <http://proteusfund.org/uplifting-gender-justice-activism-Queer-Crescent/>.
- Belcourt, B., 2020. *A History of My Brief Body*.
- Benn, A., 2017. The black queer youth initiative: challenging racism, stigma & exclusion. In: Abramovich, A., Shelton, J. (Eds.), *Where Am I Going to Go? Intersectional Approaches to Ending LGBTQ2S Youth Homelessness in Canada & the US*. Canadian Observatory on Homelessness Press, Toronto.
- Bennett, J., July 3, 2020. These teen girls are fighting for a more just future. *New York Times*. Retrieved from: <https://www.nytimes.com/2020/0626/style/teen-girls-black-lives-matter-activism.html>.
- Bhatia, K.V., 2021. The revolution will wear burqas: feminist body politics and online activism in India. *Soc. Mov. Stud.* <https://doi.org/10.1080/14742837.2021.1944850>.
- Bhavnani, K., 2006. Interconnections and configurations: towards a global feminist ethnography. In: Hesse-Biber, S. (Ed.), *Handbook of Feminist Research: Theory and Practice*. Sage, London, UK, pp. 639–650.
- Boler, M., Nitsou, C., 2014. Social media and global protest: hybrid social movement and politics as the “redistribution of the sensible”. In: McCaughey, M. (Ed.), *Cyberactivism 2.0*. Routledge, New York, NY, p. x.
- Butler, J., 2015. *Notes Toward a Performative Theory of Assembly*. Harvard University Press, Cambridge, MA.
- Cainkar, L., Maira, S., 2005. Targeting Arab/Muslim/South Asian Americans: criminalization and cultural citizenship. *Amerasia* 31 (3), 1–27.
- Castilho, F., Romancini, R., 2017. “Fight like a girl”: virtual bedroom culture in public school occupations in Brazil. *Catalan J. Commun. Cult. Stud.* 9 (2), 303–320.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity, politics, and violence against women of color. *Stanford Law Rev.* 43, 1241–1299.
- Cristaldi, M., Pamponini, G., 2016. Research and activism about girls’ education for global democracy: the case of the campaign “Etna, Volcano of Peace”, Catania, Italy. *Pol. Futures Educ. Internet* 14 (5), 578–589.

- Curley, S., Rhee, J., Subedi, B., Subreenduth, S., 2018. Activism as/in for global citizenship: putting un-learning to work towards educating the future. In: Davies, I., et al. (Eds.), *The Palgrave Handbook of Global Citizenship and Education*, pp. 589–606. https://doi.org/10.1057/978-1-137-59733-5_37.
- Danforth, J. (Ed.), 2011. *Feminism for Real: Deconstructing the Academic Industrial Complex of Feminism*. The Canadian Centre for Policy Alternatives, Ottawa, ON.
- Darder, A., 2011. *A Dissident Voice: Essays on Culture, Pedagogy, and Power*. Peter Lang, New York, NY.
- Davis, A.Y., 2016. *Freedom is a Constant Struggle: Ferguson, Palestine, and the Foundations of a Movement*. Haymarket Books, Chicago, IL.
- de Finney, S., 2017. Indigenous girls' resilience in settler states: honouring body and land sovereignty. *Agenda* 31 (2), 10–21.
- Desai, S., Abeita, A., 2017. Institutional microaggressions at a Hispanic serving institution: a Diné woman utilizing tribal critical race theory through student activism. *Equity Excell. Educ.* 50 (3), 275–289.
- Dhillon, J., 2015. Indigenous girls and the violence of settler-colonial policing. *Decolonization* 4 (2), 1–31.
- Edwards-Jauch, L., 2012. Action research on gender-based violence at the University of Namibia: results and methodological reactions. *Fem. Africa* 17, 102–113.
- Fine, M., 2014. An intimate memoir of resistance theory. In: Tuck, E., Yang, K.W. (Eds.), *Youth Resistance Research and Theories of Change*, pp. 46–58. New York.
- Foucault, M., 1984. Truth and power. In: Rabinow, P. (Ed.), *The Foucault Reader*. Pantheon Books, New York, NY, pp. 51–75.
- Foucault, M., 1989. The subject and power. In: Dreyfus, H.L., Rainbow, P. (Eds.), *Michel Foucault: Beyond Structuralism and Hermeneutics*. University of Chicago Press, Chicago, IL, pp. 208–222.
- Foucault, M., 1991. On governmentality. In: Burchell, G., Gordon, C., Miller, P. (Eds.), *The Foucault Effect*. University of Chicago Press, Chicago, IL, pp. 87–104.
- Garita, A., 2015. Moving toward sexual and reproductive justice: a transnational and multigenerational remix. In: Baksh, R., Harcourt, W. (Eds.), *The Oxford Handbook of Transnational Feminist Movements*. Oxford University Press, pp. 271–294.
- Ghose, T., 2012. Politicizing political society: mobilizing sex workers in Sonagachi, Calcutta. In: Loomba, A., Lukose, R.A. (Eds.), *South Asian Feminisms*. Duke University Press.
- Gopinath, G., 2005. *Impossible Desires: Queer Diasporas and South Asian Public Cultures*. Duke University Press, Durham, NC.
- Grewal, Kaplan, C. (Eds.), 1994. *Scattered Hegemonies: Postmodernity and Transnational Feminist Practices*. University of Minnesota Press, Minneapolis, ISBN 978-0816621385.
- Grewal, I., Kaplan, C., 2000. Postcolonial studies and transnational feminist practices. In: Hall, S., et al. (Eds.), *Culture, Media, Language: Working Papers From Cultural Studies*, 1972–79. Routledge, London, pp. 2–35. Retrieved from: <http://english.chass.ncsu.edu/jouvert/v5i1/grewal.htm>.
- Hollis, S., 2019. Empowering Women and Girls in the Fight Against HIV. Friends of the Global Fight Against AIDS, Tuberculosis, and Malaria. Retrieved from: <https://www.theglobalfight.org/women-girls-and-hiv/>. (Accessed 5 September 2021).
- hooks, B., 1986. Sisterhood: political solidarity between women. *Fem. Rev.* 23, 125–138. <https://doi.org/10.2307/1394725>.
- Jackson, S.J., Bailey, M., Welles, B.F., 2020. *#Hashtag Activism: Networks of Gender and Racial Justice*. MIT Press, Cambridge, MA.
- Johnson, E.P., 2018. *Black, Queer, Southern, Women: An Oral History*. Oxford University Press.
- Jones, H., Gunaratnam, Y., Bhattacharya, G., Davies, W., Dhalwal, S., Forkert, K., Jackson, E., Saltus, R. (Eds.), 2017. *Go Home? The Politics of Immigration Controversies*. Manchester University Press, Manchester, UK.
- Khoji-Moolji, S., 2018. *Forging the Ideal Educated Girl: The Production of Desirable Subjects in Muslim South Asia*. University of California Los Angeles Press.
- Kimball, G., 2019. Media empowers brave girls to be global activists. *J. Int. Wom. Stud.* 20 (7), 35–56. <https://vc.bridgew.edu/jiws/vol20/iss7/3>.
- Kumari, S., 2021. Meme culture and social media as gendered spaces of dissent and dominance. *J. Vis. Literacy*. <https://doi.org/10.1080/1051144X.2021.1974775>.
- Kurian, P.A., Munishi, D., Mundkur, A., 2015. The dialectics of power and powerlessness in transnational feminist networks: online struggles around gender-based violence. In: Baksh, R., Harcourt, W. (Eds.), *The Oxford Handbook of Transnational Feminist Movements*. Oxford University, London, UK, pp. 871–894.
- Lesperance, A., 2016. Sexy health carnival: one small part of Indigenous herstory. *Girlhood Stud.* 9 (2), 88–89.
- Liebert, R.L., Thompson, L., 2015. Reflexivity & (r)evolution: editorial reflections on the special feature. *Fem. Psychol.* 25 (1), 95–100.
- Mai, Y., Laine, S., 2017. Blogging activism of young educated and global women in Tunisia and Vietnam: a two-case study. *Partecipazion e Conflitto* 9 (3), 893–917.
- Maira, S., 2010. Citizenship and dissent: South Asian Muslim youth in the US after 9/11. *S. Asian Popular Cult.* 8 (1), 31–45.
- Maira, S., 2017. *The 9/11 Generation: Youth, Rights, and Solidarity in the War on Terror*. New York University Press, New York.
- Mampane, J.N., 2020. Susceptible lives: gender-based violence, young lesbian women, and HIV risk in a rural community in South Africa. *J. Int. Wom. Stud.* 21 (6), 249–264.
- Matherly, J., 2013. How many voices is one? Three uses of collage as activism for young women. *J. Lang. Literacy Educ.* 9 (1), 214–219.
- McAdam, S., 2014. Armed with nothing more than a song and a drum. The Kino-nda-niimi Collective. In: *The Winter We Danced: Voice From the Past, the Future, and the Idle No More Movement*. ARP Books, Winnipeg, pp. 65–66.
- Minh-Ha, T.T., 1989. *Woman, Native, Other: Writing Postcoloniality and Feminism*. Indiana University Press, Bloomington, IN.
- Mohanty, C.T., 2003. *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity*. Duke University Press, Durham, NC.
- Mohanty, C.T., 2013. Transnational feminist crossings: on Neoliberalism and radical critique. *Signs* 38 (4), 967–991.
- Moletsane, R., 2018. "Stop the war on women's bodies": facilitating a girl-led march against sexual violence in a rural community in South Africa. *Stud. Soc. Justice* 12 (2), 235–250.
- Monchalin, R., Lesperance, A., Flicker, S., Logie, C., Native Youth Sexual Health Network, 2016. Sexy health carnival on the powwow trail: HIV prevention by and for Indigenous youth. *Int. J. Indig. Health* 11 (1), 159–176.
- Moss, E., July–August, 2012. Voices silenced, tweets heard. *Foreign Serv. J.* 11–12.
- Ong, A., 1996. Cultural citizenship as subject-making. *Curr. Anthropol.* 37 (5), 737–762.
- Ong, A., 2006. *Neoliberalism as Exception: Mutations in Citizenship and Sovereignty*. Duke University Press, Durham, NC.
- Pastor, D., Hernández, A., 2021. Young women and decolonization changing history through struggle and resistance. *Cult. Surviv. Q.* 45 (3), 6–9.
- PedAIDS, 2020. Youth Activists: "Do not forget that HIV is still here". <https://www.pedaids.org/2020/08/11/youth-activists-do-not-forget-that-hiv-is-still-here/>. (Accessed 15 June 2021).
- Phillips, T., March 7, 2020. "This is our feminist spring!": millions of Mexican women prepare to strike over femicides. *The Guardian*. <https://www.theguardian.com/world/2020/mar/07/mexico-femicides-protest-women-strike>.
- Pon, Coloma, Kwak, Huynh, 2013. Asian Canadian studies now: directions and challenges. In: Coloma, R.S., Pon, G. (Eds.), *Asian Canadian Studies Reader*. University of Toronto, Toronto, ON, pp. 3–29.
- Quiesada, U., Gomez, L., Vidal-Ortiz, S., 2015. *Queer Brown Voices: Personal Narratives of Latino/a LGBT Activism*. University of Texas Press, Austin, TX.
- Rhee, J., 2021. *Decolonial Feminist Research: Haunting, Rememory, and Mothers*. Routledge, New York: NY.
- Rodriguez, L., December 8, 2021. 5 Amazing Young people fighting for women's right to control their own bodies. *Global Citizen*. Retrieved from: <https://www.globalcitizen.org/en/content/activists-fighting-womens-right-to-control-body/>.
- Roy, A., February 18, 2002. Shall we leave it to the experts: Enron's power project in India demonstrates who benefits from globalization. *The Nation*. Retrieved at: <https://www.thenation.com/article/archive/shall-we-leave-it-experts/>. (Accessed 5 January 2014).
- Salam, E., June 17, 2021. Black Lives Matters protesters make Palestinian struggle their own. *The Guardian*. <https://www.theguardian.com/world/2021/jun/16/black-lives-matter-palestinian-struggle-us-left>.
- Sandoval, C., 2000. *Methodology of the Oppressed*. University of Minnesota Press, Minneapolis, MN.
- Schuster, J., 2013. Invisible feminists? Social media and young women's political participation. *Polit. Sci.* 65 (1), 8–24.
- Sensory, Ö., Marshall, E., 2010. Missionary girl power: saving the Third World one girl at a time. *Gend. Educ.* 22 (3), 1–17.
- Sharoni, S., Abdulhabadi, R., Al-Ali, N., Eaves, F., Lentini, R., Siddiqi, D., 2015. Transnational feminist solidarity in times of crisis. *The Boycott, Divestment, and Sanctions Movement (BDS) and justice for/in Palestine*. *Int. Fem. J. Politics* 17 (4), 654–670.

- Sherwood, R., 2021. Christian persecution rises as persons refused aid in COVID crisis. The Guardian. <https://www.theguardian.com/world/2021/jan/13/christian-persecution-rises-as-people-refused-aid-in-covid-crisis-report>.
- Simpson, L.B., 2013. *Islands of Decolonial Love*. Arp Books, Winnipeg, MB.
- Somers, M., 2008. *Genealogies of Citizenship: Markets, Statelessness, and the Right to Have Rights*. Cambridge University Press, Cambridge, UK.
- Sparks, H., 1997. Dissident citizenship: democratic theory, political courage, and activist women. *Hypatia* 12 (4), 74–110.
- Spivak, G.C., 1985. The Rani of Sirmur: an essay in reading the archives. *Hist. Theor.* 24 (3), 247–272.
- Spivak, G.C., 1988. Can the subaltern speak? In: Grossberg, L., Nelson, C. (Eds.), *Marxism and the Interpretation of Culture*. University of Illinois Press, Urbana, IL, pp. 271–311.
- Spivak, G.C., 2004. Righting wrongs. *S. Atl. Q.* 103, 523–581.
- Subedi, B., Maleku, A., 2021. Bhutanese–Nepali young women's experiences with racism inside and outside of schools. *Educ. Stud.* <https://doi.org/10.1080/00131946.2021.1878176>.
- Suen, Y.T., Chan, R.C.H., 2021. "Gay Brain Drain": Hong Kong Lesbian, Gay, and Bisexual people's consideration of emigration because of non-inclusive social policies. *Sex. Res. Soc. Pol.* 18, 739–752.
- Swarr, A.L., Nagar, R. (Eds.), 2012. *Critical Transnational Feminist Praxis*. SUNY Press, Albany, NY.
- Taft, J., 2011. *Rebel Girls: Youth Activism and Social Change Across the Americas*. New York University, New York.
- Tanner, S., Kalm, N., Cull, H., Matthewson, G., Aisenberg, A., 2020. Disruption and design: crowdmapping young women's experience in cities. *Inst. Dev. Stud. Bull.* 51 (2), 113–128.
- The Global Fund, 2019. Step Up the Fight: Ending HIV Among Adolescents Girls and Young Women. <https://healtheducationresources.unesco.org/library/documents/step-fight-ending-hiv-among-adolescents-girls-and-young-women>. (Accessed 23 February 2020).
- Troutman, S., Jimenez, I., 2016. Lessons in transgression: #BlackGirlsMatter and the feminist classroom. *Meridians* 15 (1), 7–39.
- Truong, T., 2015. Human trafficking, globalization, and transnational feminist responses. In: Backsh, R., Harcourt, W. (Eds.), *The Oxford Handbook of Transnational Feminist Movements*. Oxford University Press, Oxford, UK, pp. 295–320.
- UNFPA, 2021. My Body is My Own: Claiming the Right to Autonomy and Self-Determination. State of the World Population Report (Retrieved from UNFPA.org/SWOP).
- Van der Molen, T., Bal, E., 2011. Staging "small, small incidents": dissent, gender, and militarization among young people in Kashmir. *Focaal* 60, 93–107.
- Walking with Our Sisters, 2014. The Project: A Commemorative Art Installation for the Missing and Murdered Indigenous Women of Canada and the United States. Retrieved on Dec. 8, 2015, from <http://walkingwithoursisters.ca>.
- Wang, Q., 2018. Young feminist activists in present-day China: a new feminist generation? *China Perspect.* 3, 59–68. <http://journals.openedition.org/chinaperspectives/8165>.
- Weems, L., 2009. M.I.A. in the global youthscape: rethinking girls' resistance and agency in postcolonial contexts. *J. Girlhood Stud.* 2 (2), 55–75.
- Weems, L., 2014. Refuting "refugee chic": transnational girl(hood) and the guerilla pedagogy of M.I.A. *Fem. Form.* 26 (1), 115–142.
- Weems, L., 2015. Rethinking resistance and intimacy. *Jeun. Young People, Texts, Cult.* 7 (2), 134–147.
- Weems, L., 2018. *Staging Dissent: Young Women of Color and Transnational Activism*.
- World Health Organization, 2021. Violence Against Women Prevalence, Estimates, 2018 Executive Summary. Retrieved from <https://www.who.int/publications/i/item/9789240026681>.
- Wright, H.K., 2000. Why write back to the new missionaries?: addressing the exclusion of (Black) others from discourses of empowerment. In: Dei, G.S., et al. (Eds.), *Power, Knowledge and Anti-Racism in Education: A Critical Reader*. Fernwood.
- Wynter, S., 2003. Unsettling the coloniality of being/power/truth/freedom: towards the human, after man, its overrepresentation – an argument. *Clim. Res.* 3 (3), 257–337.
- Yola, N., Mpongo, C., Latief, I., Bekker, L., 2019. World AIDS day reminds us to focus on communities. *Indian J. Med. Res.* 150, 518–520.
- Zera, A., Perez, J., Wang, C., 2017. A proposal for expanding endarkened transnational feminist praxis: creating a database of women's scholarship and activism to promote health in Zimbabwe. *Qual. Inq.* 23 (2), 107.

Relevant websites

Organizations and grassroots social movement organizing

- African American Policy Foundation (AAPF) Founded in 1996, The African American Policy Forum (AAPF) is an innovative think tank that connects academics, activists, and policymakers to promote efforts to dismantle structural inequality and promote human rights through intersectionality.
- AF3irm – Transnational Feminist Organization with chapters in CA, WA, NY, HI, MN, and NV.
- A Long Walk Home – (United States) community-based organization focused on survivors of sex- and gender-based trauma.
- Association for Women's Rights in Development (AWID) is a global, feminist, membership, movement-support organization working to achieve gender justice and women's human rights worldwide.
- BlankNoise.org –(India) is a community of 'Action Sherooks, Theyroes and Heroes', individuals and citizens united to eradicate gender-based and sexual violence. Initiated in 2003, as a response to street harassment as a final graduation project at the Srishti Institute of Art Design and Technology, Blank Noise has grown into a multi-city/country collective led by its Action Sherooks, Theyroes, and Heroes.
- Feminist Five (China).
- Coalition to Stop Violence Against Native Women (United States).
- Young Feminist Fund – FRIDA (Flexibility Resources Inclusivity Diversity Action) - provides young feminist organizers with the resources they need to amplify their voices and bring attention to the social justice issues they care about.
- HEART Women and Girls – Organization to "ensure that Muslims have the resources, language, and choice to nurture sexual health and confront sexual violence."
- First Nations-led organization for Indigenous Sovereignty on Turtle Island (North America).
- Me too. A global organization that grew from the #MeToo activism founded in 2006 by Tamara Lomax in the United States.
- Missing and Murdered Indigenous Women and Girls.
- Native Youth Sexual Health Network (NYSHN) (Canada) community-based organization.
- National Indigenous Women's Resource Center – (United States) Native-led organization to provide culturally appropriate resources and information regarding MMIWG2S.
- #NiUnaMenos (Latin American Social Movement that formed in Argentina in response to gender-based violence).
- Sisters Rising (Victoria, CAN) is an Indigenous community-based project for Indigenous girls, the youth of all genders, and communities speaking back against gendered and sexualized violence.
- Sisters Uncut.
- Warrior Women Project.
- <https://pmnch.who.int/about-pmnch/constituencies/adolescents-and-youth>.
- <http://www.afghanwomensnetwork.org/>.

Undocumented activism: a bate-papo~testimonio

Luis Fernando Macías^a and Daniella Santos Vieira^b, ^a California State University, Fresno, CA, United States; and ^b Bowling Green State University, Bowling Green, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	374
Literature overview	375
Architecting the DREAM	375
Felons not families	375
Trump administration	376
Undocumented activism globally	376
Methodology	376
Bate-papo and pláticas	376
Testimonio	377
Positionality	377
Guiding frameworks	377
Activism spectrum	377
Critical Race Theory	378
Activism, advocacy, allyship, usurping	378
Daniella's testimonio	378
Do you think undocumented students belong in college?	378
Hope, fear, and action	379
I trust that you will not post that anywhere	380
One last push ... and shove	380
Beyond ESU	381
Analysis	381
Existence as resistance: origins of an activist	381
Sacrifice in frontline activism	381
Usurping the moment and the movement	382
Perceiving "wins"	382
Interpreting signs of the university's commitment to diversity	383
The legacy of progress	383
Conclusion	383
References	384

Introduction

Undocumented people in the United States live intersectional lives (Collins, 1998) and when they engage in activism, their activism acts as a rotating epicenter (Montoya and Valdes, 2008) that is done in coalition with other social movements. Undocumented activism has also been inextricable to the struggle for equal educational opportunities. In the 1970s for example, Mexican families in Texas risked deportation fighting for their children's right to an education. That fight would become the landmark Supreme Court decision of Plyler v Doe (1982) which extends a free public compulsory education to students, regardless of their immigration status (Olivas, 2005).

For undocumented students, two laws passed in 1996 severely curtailed their pursuits of a higher education by disqualifying them from qualifying for in-state tuition (P.L.104-208-104th Congress: IIRAIRA, 1996) and excluding them from federal financial aid (H.R. 3734-104th Congress, PRWORA, 1996). The impact of these laws is evident when considering that out of the approximately 95,000 undocumented youth who graduate from high school yearly, only between 5 and 10% enroll in college (Zong and Batalova, 2019) with Latinxs students making up nearly 48% of undocumented college students (Feldblum et al., 2021).

Undocumented youth's fight for higher education has produced some of the most transformative and revolutionary acts of activism. Our co-author's experiences as an undocumented student activist at a predominately white institution (PWI) serve as an important contribution to the existing research literature. Daniella's story of activism is shared by way of a bate-papo~testimonio which is rooted in decolonizing methodologies. Her testimonio provides critical nuance that can generate necessary rectifying measures. Our use of testimonio is informed by Solórzano and Bernal (2001) who maintain that readers of a testimonio are put in a position akin to jury members in a courtroom. After bearing witness, they are under obligation to respond. They can choose to act on that obligation or not, but cannot ignore it.

Literature overview

In 1994, California's Prop 187 proposed denying all state-funded benefits (including public education) to undocumented people (Olivas, 2005; Villazor, 2019). Community members, many of whom were undocumented, correspondingly organized mass demonstrations in opposition. Youth's involvement in activism became stark on November 02, 1994, when approximately 10,000 students in Los Angeles walked out of school in protest (Pyle and Romero, 1994). While access to K-12 education remained protected, in 1996, two separate laws severely curtailed undocumented students' pursuits of a higher education by disqualifying them from qualifying for in-state tuition (P.L.104-208-104th Congress: IIRAIRA, 1996) and excluding them from federal financial aid (H.R. 3734-104th Congress, PRWORA, 1996).

Architecting the DREAM

In the late 1990s, civil rights associations and some state legislators coordinated efforts to promote an immigration reform law. Their public opinion research indicated that the least stigmatized immigrant group were young students who were raised in the country, identified as American, but lacked immigration status (Nicholls, 2013; Perez, 2009). Their first attempt at introducing an immigration reform bill focused on this group that they came to call Dreamers; a moniker that purposefully alluded to them as an embodiment of the country's storied meritocratic values (i.e., the American Dream). In 2001, the first DREAM Act was introduced into Congress. Had it passed, it would have allowed those eligible the ability to adjust their immigration status and become US citizens (S. 1291-107th Congress, DREAM, 2001-2).

Subsequently, civil rights associations helped form a network of local and state student groups and advocacy organizations called the "United We Dream Coalition" (Nicholls, 2013). This coalition trained youth in organizing, campaigning, and mass media communication. Undocumented youth were consulted on how to frame their personal stories as ones of assimilation, exceptionalism, and innocence whenever speaking with politicians, the public, and the media (Costanza-Chock, 2014). The network grew rapidly, particularly across university campuses in areas of the US like California, Arizona, Texas, and Chicago (Dougherty et al., 2010). As the network expanded, internal disagreement began over strategy and direction. Many felt that the focus on lobbying politicians was unproductive and that the DREAMer narrative was exclusionary of many and condemning of their parents. They were also critical of the network's leadership which was mostly non-undocumented people.

During this time, the Sensenbrenner Bill (H.R. 4437) was introduced which proposed a series of draconian efforts which included constituting undocumented people's presence in the US a felony. In 2006, mass protests against the bill began in Chicago and continued nationwide for a period of eight weeks (Pallares and Flores-González, 2010). Undocumented youth were at the center of many of the organizing efforts against H.R. 4437, which is estimated to have brought in between 1.6 and 2.2 million protesters across the United States (Fox, 2006).

Felons not families

Immigration reform was central to Barack Obama's 2008 presidential campaign. However, once elected, his administration instead dramatically increased deportations. Undocumented youth responded in a variety of ways. Groups broke off from the larger DREAM networks and abandoned the established strategy. These new groups favored highly visible acts of civil disobedience with the purpose of drawing attention to the protection of all undocumented people (Truax, 2015). This faction of activists embraced an organizational structure that allowed for more horizontal coordination among groups, rather than a traditional hierarchical model (Nicholls, 2013). They utilized emerging digital media outlets as a form of ground organizing (Seif, 2011). The message they crafted and promoted was their own: undocumented, unapologetic, and unafraid (Zimmerman, 2011). To emphasize this message they staged numerous national, state, and local protests such as long-distance walks, sit-ins, hunger strikes, rallies, and mock graduation ceremonies where they publicly proclaimed their undocumented status (Gonzales, 2008; Truax, 2015; Abrego and Negrón-Gonzales, 2020).

After mounting pressure from activists, another failed DREAM Act in 2010, and facing re-election, Obama announced the Deferred Action for Childhood Arrivals (DACA) program in 2012. DACA is not an immigration reform law in the way that the DREAM Acts would have been. It is a temporary discretion that grants recipients a valid social security number and work authorization for 2 years if they meet several very specific key requirements. The program has facilitated employment and post-secondary education for many recipients (Castrellón et al., 2017; Nienhuser et al., 2020; Olivas, 2020; Macías, 2018), but it has done little to placate activists' demands. Groups like National Immigrant Youth Alliance (NIYA) focused their activist efforts on the immigrant detention industrial complex (Truax, 2015). In 2012, NIYA members Marco Saavedra and Viridiana Martinez surrendered themselves to immigration officials in order to infiltrate the Broward County detention center to expose human rights abuses and halt deportations (Ibarra and Rivera, 2019). Afterward, Marco became part of the DREAM 9 activists who self-deported to Mexico and later presented themselves at the US border seeking asylum. This incredibly risky and highly public act was done to protest the record number of deportations and challenge Obama's position that only "felons, not families" were being deported.

Trump administration

In the United States, undocumented activism under the Trump administration remained consistent albeit under new threats. Trump's transition team projected deportations under the new administration would increase by 75% without adding personnel (Bennet, 2016). This projection alluded to ending DACA then using existing data (i.e., DACA recipients' personal information) and possibly subpoenaing school records to increase deportations.

Trump's racist-nativist political rhetoric and subsequent policies impacted negatively impacted Immigrant and Students of Color, throughout the US (Pérez Huber and Muñoz, 2021). Public schools and other institutions responded in a variety of ways, one of which was declaring themselves sanctuaries from immigration enforcement (Collet and Macías, 2021). For schools and universities, undocumented activists advocated creating sanctuary campuses which would include refusing to cooperate with subpoenas and immigration officials (Muñoz and Vigil, 2018; Allard, 2018). Educators and students were on the frontlines activists organized in various ways such as creating safe spaces, rapid response teams and other support networks in response to deportations (Crawford and Dorner, 2019). In 2019, Trump lost his presidential re-election campaign. He lost critical electoral votes in states like Arizona, Georgia, and Michigan where organizers and activists have been credited for mobilizing voters.

Undocumented activism globally

While this work is situated in the US context, it is imperative to note that immigration is a global occurrence. Displacement, changes in labor markets, climate change, political unrest, as well as the search for safety and/or increased opportunities drive people to migrate (Connor, 2016). However, increased immigration, particularly by people from the Global South immigrating to Western nations¹ has been met by increasingly draconian laws that illegalize their presence then surveil, identify, detain, and deport them (Connor, 2016; Gonzales et al., 2019).

Correspondingly, undocumented immigrants across Europe commonly experience racial aggressions, social stigma, as well as financial limitations, and governmental exclusions (Gonzales et al., 2019; Alfred, 2020; Madia, 2021). Undocumented students, in particular, face similar challenges to those in the US. Compulsory education is available to them in many European countries, but those with post-secondary aspirations are excluded from government educational subsidies and largely unable to afford the high foreign student tuition fees associated with their immigration status (Covington, 2008).

Thus, immigrant activists have organized various campaigns and tactics to assert their rights and sense of belonging across European nations (Swerts and Nicholls, 2021). In Ireland for example, the "Young, Paperless and Powerful" (YPP) campaign supports artistic outlets for undocumented youth to share their experiences and advocates for immigration reform in Ireland (Alfred, 2020). In Greece, activists occupied/squatted in an abandoned orphanage in Thessaloniki, Greece. Organizers of "Orfanotrofo" housing squat in 2015 provided meals, basic needs, and other supplies to immigrants and activists. This action was done, in part, as a response to the Greek government's hardline response to the nearly one million people seeking refuge from conflicts in Syria, Iraq, and Afghanistan (Karaliotas and Kapsali, 2021). More recently, undocumented activists held a two-month-long hunger strike across Belgium in 2021. The approximately 500-person hunger strike sought to bring attention to exploitation and exclusions they face in Belgium (Madia, 2021). Undocumented youth also wanted to draw attention to the financial unfeasibility of attending university. While these campaigns have had varying levels of success, their strengths demonstrate the power of organizing and alliance-building and the limitations signal the underlying issue of co-optation by middle-class allies (Karaliotas and Kapsali, 2021).

Methodology

A methodology takes into consideration the epistemological framing of the researcher and their relationship to the research inquiry (Fierros and Delgado Bernal, 2016). Ladson-Billings (2003) recognizes that epistemology is a person's "system of knowing" the world which is shaped by the conditions under which they live. As the authors of this work, we recognize that our ways of knowing are intrinsic to our respective mother tongues, cultural upbringing, and also experiences with immigration. Our lingua franca is English but we were raised in Mexican (Luis Fernando) and Brazilian (Daniella) households speaking Spanish and Portuguese, respectively. We grew up in mixed immigration status families but recognize important distinctions with respect to our immigration statuses. Our methodological approach reflects these rich epistemologies, which we are calling a methodology of bate-papo~testimonio.

Bate-papo and pláticas

Bate-papo is a Brazilian expression for ongoing, informal conversations shared between friends and family. The word "papo" refers to a bird's crow; while the "bate" means to strike. When chickens make noise, their throat appears to strike the inside of their neck. "Bate-papo" is the imagery of the chicken's throat movement. The expression is akin to saying, "let's get our throats moving". A bate-papo may vary from playful, as the name suggests, to solemn. Regardless of the tone, people exchange ideas, pass along

¹many of which were former colonial powers.

information, and build community. In this way, bate-papo is conceptually linked to pláticas which are also part of rich oral traditions within Latinx/Latin American cultures. Pláticas are a collaborative process of “sharing stories, building community, and acknowledging multiple realities and vulnerabilities” (Burciaga and Tavares, 2006). They take part through the “syncopated sounds of carcajadas, gritos, and, murmurings” (Chabram-Dernersesian and de la Torre, 2008). Pláticas are a space where Latinas are rightfully situated as experts embodying knowledge and producing theory about their own lived experiences (Gonzalez, 1999). The knowledge produced offers valuable insights for academic scholarship (Delgado Bernal, 1998, 2002). Fierros and Delgado Bernal (2016) outlined five principles of plática methodology. The most salient in this work are the principles that (i) draw from Chicana/Latina feminist theory and recognize (ii) situate platicadoras as co-creators/co-constructors of knowledge, (iii) solidify the connections between lived experiences and the research inquiry and (iv) underline the need for reciprocity and trust.

Nosso bate-papo centers Daniella’s experiential knowledge with undocumented activism. Its candid nature was only possible because of a friendship and trust that began when we were part of an informal student group fighting for undocumented student rights. For this work specifically, batemos um papo over six months via text message exchanges, collaborative online document editing, phone, and video calls. The prompts that shaped our conversations were rooted in Luis Fernando’s interest in learning more about Daniella’s experiences and perspectives. Our work also introduces new conceptual understandings of activism which are the product of what Latina Feminist theory recognizes as Daniella’s own articulation of a “theory in the flesh” (Moraga and Anzaldúa, 2015).

Testimonio

Nosso bate-papo served as the vehicle for Daniella’s testimonio about her different experiences with activism which may be considered this work’s findings. We chose to assign pseudonyms to the people and organizations referenced in Daniella’s testimonio to maintain their anonymity. We’re particularly mindful of protecting other undocumented immigrants mentioned in Daniella’s testimonio from having their status disclosed and exposing them to backlash from the powerful institutions and individuals cited in this testimonio.

Testimonios were born out of Latin American human rights struggles and are recognized as a verbal journey that denounces injustices and documents the experiences and acts of survival of oppressed groups (Solórzano and Bernal, 2001). Testimonios are also critical reflections of a personal experience which serve to shed light on injustices and also provide possibilities for rectification (Delgado Bernal et al., 2012). When employed by Latinas, testimonios as a methodology empowers testimonialistas to honor their complexities and reclaim their own story. The ~ (tilde) between pláticas~testimonios signifies the overarching process of reflexión—thus, the bridge between the two methodologies (Flores Carmona et al., 2018). Similarly, we embrace the bridge in nosso bate-papo~testimonio and continue the tradition of decolonizing methodologies that “teach, learn and form alliances based on our personal stories that are very much communal” (Delgado Bernal, 2009).

Positionality

I, Luis Fernando, am a first-generation US citizen who grew up in a mixed immigration status home. My family immigrated from Mexico to Southwest Texas when I was a child. Given the mixed statuses in my family, immigration was a pressing issue in my home and community. However, I have never been undocumented and thus recognize that I am the beneficiary of multiple privileges allotted to a US citizen.

I, Daniella, am a DACA recipient born in Brazil and raised in Massachusetts. As an undocumented immigrant, my life and scholarship have been directly shaped by immigration policy. The deportation protection I have as a DACA recipient enables me to disclose my status with a lesser degree of fear than I felt prior to having DACA. Both of my parents are naturalized US citizens. As a result, I feel privileged to be able to discuss my undocumented identity without putting my family at risk.

Guiding frameworks

Our work draws on two complementing frameworks to drive the discussion on undocumented activism. We rely on a conceptual framework that characterizes Black and Brown students’ activism as a spectrum of civic engagement and Critical Race Theory’s (CRT) tenet of interest convergence. These frameworks also inform our own theoretical contributions of what it means to engage in activism, advocacy, allyship and to usurp.

Activism spectrum

Logan et al.’s (2017) study on Black and Brown millennial students’ activism suggests that their characterization of activism falls within a spectrum, rather than a singular definition. That spectrum ranges from “existence as resistance” on one end and “activism consumes life” on the other extreme. “Existence as resistance” is defined as the mere act of existing as a marginalized person within a hostile environment. From this perspective, a student of Color attending a PWI engages in activism simply by virtue of their presence. Strategy and intentionality are not required. Toward the opposite end of the spectrum is “frontline activism” which the authors describe as activism with more intentionality and greater personal sacrifice. Frontline activists are strategic about their actions and

take on leadership roles in the various efforts to fight injustices experienced by their community. At the very end of the spectrum is “activism consumes life” which is a full-time commitment to visibly organizing and advocating against injustices. The factor that sets this form of activism apart from “frontline activism” is the recognition that only a few people have the time and resources available to have activism be their primary role in life. Logan et al. (2017) then posit that different students identify singularly within one of those definitions in the spectrum.

Critical Race Theory

Critical Race Theory (CRT) is an intellectual movement that recognizes People of Color’s experiences with racism and other forms of subordination as units of analysis to understand inequalities in the United States (Solorzano, 1997). One of CRT’s main tenets is interest convergence which proposes that there are moments when the interest of the dominant group temporarily aligns with that of minority groups that then allows for some measure of social change. Interest Convergence outlines that the systematically privileged dominant group will support justice only to the extent that it does not result in the forfeiture of their own power (Bell, 1980). Thus, seemingly opposing interests converge but the specific interest of a dominant group is always furthered.

Activism, advocacy, allyship, usurping

Rooted in Daniella’s experience with activism, we theorize what it means to engage in activism, advocacy, allyship, and to usurp. Different from Logan et al. (2017) we do not conceptualize these as categories that are static, hierarchical, or on a spectrum. Rather, they are actions that people take in relation to their position within certain social movements.

We posit that activism is an effort to organize and/or increase consciousness by someone who is directly impacted by the injustice(s) being fought. These actions may include leading a roundtable discussion, participating in a public protest, or attending support groups to interact with fellow members of a marginalized group. Critical to activism is some measure of personal sacrifice. Sacrifice may be understood as sharing vulnerable aspects of one’s identity or dedicating time and resources. A person’s wellbeing, relationships, and/or professional opportunities may also be a form of sacrifice.

Advocacy differs from activism in that advocacy is to take actions in support of activism. People who enact advocacy are not directly impacted by the injustice(s) being fought. Thus, they may leverage their privilege(s) to advance activism efforts. Advocacy may include contributing personal, professional resources or intervening in moments or hostile environments that are not safe for those directly impacted.

Allyship is to demonstrate support toward those who are facing injustice(s) without leveraging any privileges or contributing resources. Demonstrations of support in allyship may include nominal public donations, participating in protests, signing petitions and/or asking elected officials to support certain legislation. Allyship differs slightly from virtue signaling. While both are largely symbolic, we do not propose that allyship is done in order to garner praise or acknowledge one’s righteousness. A person may not have the knowledge, availability, or interest to engage in advocacy and thus their actions are limited to allyship.

Conversely, we posit that to usurp is in fact to take actions that garner praise and advance personal and professional ambitions at the expense of those directly impacted by the injustice(s). Usurping is not sacrifice nor leveraging privilege. There is a performative element of public support and/or labor taken in organizing efforts (i.e., rallies, petitions) yet those actions are taken for self-benefit. Usurping may be done willfully or unknowingly. Nevertheless, this form of co-optation is harmful because it exploits marginalized people, erases their efforts, and undermines their agency and purpose for activism.

Daniella’s testimonio

When I, Daniella, was in my senior year of high school, I became a DACA recipient which made it easier for me to apply to multiple colleges around the country. I was accepted to several universities but ultimately enrolled at Elm State University (ESU) because it was the only economically feasible option. All other colleges that accepted me flagged my application as ineligible for any financial aid due to my undocumented status. ESU, however, awarded me with a merit-based scholarship that was renewable for four years. I later learned that this scholarship award was only possible due to an administrative misstep that allowed my undocumented status to go undetected. This oversight, while financially beneficial, incentivized me to hide my status. I navigated my time in college in constant fear that ESU would find out about my status and revoke my scholarships. Each of the following vignettes provides a snapshot of my experiences as a college student which will subsequently be discussed in conjunction with the frameworks to understand them as activism.

Do you think undocumented students belong in college?

I was a sophomore at ESU when I first volunteered to help at the university’s Latinx Success Conference, which is geared toward encouraging high school students to pursue college. My assigned role for the event was minimal: arrange chairs and hand out lunch bags. The keynote speaker was an academic, a Latina who focused her speech on the dismal statistics surrounding high school and college graduation rates for Latinx students. Her speech felt discouraging, the opposite of what the event was intended to do. After the keynote, a panel of Latinx ESU students took the stage to answer questions from the audience. Eduarda, a graduate student, and

staff member, moderated the event, handing the microphone to students in the audience asking questions. As the event came to an end, the final question came from a high school teacher who asked the panelists “Do you think undocumented students belong in college?” I gathered that he asked the question because there were undocumented high school students in the room, who probably did not feel comfortable asking that question themselves. I watched the panelists awkwardly answer that they thought anyone could belong in college but that they were not undocumented. Their answers were vague. They stammered and shifted in their seats, noticeably uncomfortable. I empathized with undocumented students in the audience who were listening to these unhelpful answers, especially following the keynote speaker’s grim portrayal of Latinx student success. If I was hearing this as a high school student I would have felt left out and discouraged. My heart started to race as I looked around in disbelief. I felt an overwhelming sense of duty at that moment, almost like a rush. I then went up to the stage and asked Eduarda if I could speak. For the first time in my life, I shared publicly that I was undocumented. I took the microphone and said emphatically that we not only belonged in college but that we could thrive in higher education. I told the room that I was living my dream college experience: I had close friends, a prestigious, well-paid internship, and good grades. If I could do it, I felt that other undocumented students should be encouraged to do the same. I left the stage feeling accomplished and uncertain. I was surprised by my own courage to share my status but unsure if my approach had been effective in countering all of the negativity and uncertainty that lead up to that question.

Hours after the event, Eduarda messaged me and told me that a high school student in that room who was a DACA recipient came up to her crying afterward because I made her believe that she could go to college. This felt incredibly empowering and rewarding. The pride and joy inside me was nearly overflowing, but I felt that I had to keep it inside because I was not ready for anyone else to know my status. After the conference, Eduarda and I developed a close friendship. She would later introduce me to a group of undocumented ESU students.

Hope, fear, and action

Soon after the conference, I began meeting regularly with the other undocumented students on campus to socialize and build community. We playfully called ourselves the Undocumented Student Alliance (USA). We became close friends and co-constructed a space where we were able to discuss our undocumented experiences candidly with each other. We built community, and at times organized, but we were always strategic and reflective of each other’s comfort levels. We grew closer at the same time that the Donald Trump campaign was garnering more attention. It was a topic we often discussed because of the many anti-immigrant promises he had made. At first his candidacy seemed impossible, but we could already feel the damage his rhetoric was causing. There were Trump signs posted on dorm windows all over ESU, and incidents of anti-immigrant street harassment reminded us that we were not safe on our own campus. At USA meetings, we brainstormed both how to prepare for either outcome of the election, but specifically what we would do if Trump was elected.

Two weeks before the November 2016 election, Astrid, a local public radio reporter, reached out to some of us at USA. She asked us if we were interested in being interviewed about the impact of the election on the lives of undocumented immigrants. I was hesitant to participate because of the wide audience this interview would reach. I would be outing my status to listeners who, if I lost my DACA protections, as Trump had promised, could identify me for deportation. Also, one of my classmates or professors could hear the story and confront me about it and jeopardize my funding at the university. Despite these risks, I decided to go forward with the interview because my USA friends trusted the reporter who seemed to genuinely believe this story might inform voters about the need for immigration reform. Since I cannot vote, I viewed the interview as my way of participating in the election, my civic duty. I wanted there to be a historical record of my involvement in that unprecedented election and so I shared my full name and university affiliation in the interview.

On election day, I felt hopeful and confident that voters would reject the Trump agenda and elect the first woman president of the United States. I spent the day in anticipation of a historic outcome. I watched the election results come in over dinner with a family I met while canvassing that weekend. Once it seemed inevitable that Trump would win, they drove me home and I held back tears until I got inside the safety of my bedroom. Minutes after the television networks called the election for Trump, I received a Facebook message from two self-proclaimed organizers who found my name through the radio interview: Michael and Andrew, both white, U.S.-born citizens. They urged me to meet that night to plan an anti-Trump demonstration the next day. Their message read “we need to have an emergency meeting with undocumented people”. I shared with them that I was exhausted and was going to be with my USA friends to process the election results but that we could connect the following day to organize. They replied “don’t panic! Where are you meeting right now? I have some people that would like to go”. They were forceful and persistent so we met at around 1am at a USA friend’s house.

I introduced them to the other members, primarily the other women. Michael and Andrew immediately began talking about their very specific vision for how the rally would unfold. We would share our immigrant stories with the crowd and the media. Specifically, they wanted us to highlight our economic contributions to the country and our fear that Trump would enact his anti-immigrant campaign promises. They leveraged their network and circulated the rally invitation to a wide audience and ensured that there were plenty of media outlets there. The next night, around 200 people gathered near city hall to attend the rally. Michael and Andrew led the event and dictated the lineup of speakers who shared their reactions to the election. I did not receive media training nor had I been coached on how to protect my identity when speaking to random reporters. Nevertheless, when Andrew handed me the megaphone, it felt liberating to shout my immigrant story to a massive crowd. After the rally, Michael and Andrew praised me for moving the crowd to tears and called the event a success.

Immediately after that rally, Michael and Andrew recruited my undocumented peers and I for other events and demonstrations. It gradually became apparent that in community planning meetings for those events, the organizers (a majority white crowd) outnumbered and spoke over those of us in attendance who were more directly impacted by Trump's proposals and rhetoric. The resulting events were not led by us. We, mostly undocumented women, were singularly expected to attend events and enact an emotional performance highlighting our fears and vulnerability in a way that would fire people up at the expense of our safety. At the last event, I attended with them, Michael and Andrew took me to a city council meeting. They wanted to petition the council to draft protections for the city's immigrants during the Trump administration. They had their talking points prepared and my input was not valued. I was simply tasked with eliciting sympathy from the council members.

I trust that you will not post that anywhere

Months into the Trump presidency, I was feeling hopeless and scared. I had not seen any progress from the public rallies and events that I was a part of with Michael and Andrew. Thus, I shifted focus. I felt that seeking some sort of commitment from ESU to protect undocumented students seemed more viable at the time rather than hoping for larger scale political change. The radio interview, exclusively featuring four undocumented students from ESU, increased our visibility on campus. My undocumented student peers and I had been meeting on a more regular basis and focused on developing a list of demands for ESU which included revamping the admissions process so that undocumented students were not considered international students and making all eligible scholarships publicly available to undocumented students. Dr. West, a senior university official, agreed to meet with us and discuss these and other demands. Ysabel, an administrator who was positioned by the university as the de-facto contact for undocumented students, facilitated the meeting. She coached us on using language she thought would be well received by Dr. West.

Dr. West was always very measured and soft spoken, at times almost robotic. During the meeting, we expressed to him our fears that DACA would be rescinded and asked for his commitment that ESU would still honor in-state tuition rates and refuse to share our information with immigration authorities if that were the case. He responded diplomatically but was noncommittal. He said that he was open to having some privately funded merit-based scholarships available to undocumented students. However, he made it clear that the university could not make this consideration public or state it on their website because the board of trustees would not be supportive of it. When we discussed our need for having a staff member dedicated exclusively to serve as the main resource on campus for undocumented students, he surprisingly agreed. Dr. West never said "no" to any of our demands. He seemed receptive to everything. Ysabel encouraged us to see his response as "a win", even though it was unclear which of our demands would be directly met. At the close of the last meeting, she suggested we take a picture with him. Dr. West smiled brightly for the photograph but immediately after it said "I trust that you will not post that anywhere, right?"

That semester, three out of the six USA students at that meeting graduated and, as the group got smaller, the pressure being applied on the administration subsided. Ultimately, our specific personnel recommendations were overlooked as was the proposed title for this position. Rather than what we had petitioned for, ESU added the title "DACA Liaison" to an unrelated position that was already in the final stages of being filled in the Office of Student Affairs. The person that was hired did not have any professional or scholarly expertise on undocumented student support. A year after their hire, the Director of the Multicultural Center, who also lacks any formal training on undocumented student support, relieved this person of their DACA liaison duties and subsumed the role himself. As for Dr. West, he has since left ESU for an even more prominent role at one of the largest public university systems in the country; ironically one that has publicly marketed itself as a champion of undocumented students.

One last push ... and shove

Entering my final semester at ESU, commencement was constantly on my mind. I kept picturing my family at the university's stadium celebrating my achievements. My vision of that moment was coupled with concern that some of my family members would not be able to fully participate in the commencement ceremony because English is not their dominant language. I also realized that for many students (both domestic and global), commencement would be their family's first time on campus. Given ESU's prestige and global recognition, I figured they would have the resources needed to provide multilingual services for their commencement.

Five months prior to commencement, I contacted the Office of Special Events requesting the university to provide multilingual translations at commencement, especially in the languages most prominent among the student body. I did not receive a reply, but I kept checking in periodically. Four months after my initial email, an administrator in the Office of Special Events responded that a multilingual service would not be viable. The first excuse she stated was that no other university had ever done this. The second excuse was that there would be 88 countries represented at commencement and thus they could not guarantee translation for all the languages represented. Further, the administrator argued that by refusing to provide translation the school was being "mindful of our diverse student body" and that providing access to some, but not all, would be a "disservice to many."

Feeling frustrated, I reached out to Astrid, the reporter, and asked for her help doing a story on the university's response to my request. She referred me to her colleague, who agreed to not only interview me for a story but also reach out to the university to ask about this. In the meantime, I reached out to other peers including Eduarda and Luis Fernando, both of whom had some experience with professional translation and also had family members who would benefit from having a multilingual ceremony. We worked together to draft email responses challenging the administrator's response. We identified which languages we had been asking for based on our understanding of the languages represented in the local community and within the student body. The Office of Special

events finally agreed to a meeting with those of us who sent the email. During that meeting, we responded to their assertion that such an undertaking was impossible by presenting the research we had done regarding the audio technology needed and all of the logistical considerations. It was astounding to me that us students could research and develop a detailed plan to have a multilingual commencement, while the university, with all its resources, maintained that it was not possible. This meeting took place only two weeks before commencement, but I was not going to let the university use time as an excuse because I had first engaged them on the topic months prior. We were confident and assertive while speaking with the administrator and I made sure to point out that there would be a news story about the university's response to our demands.

After much effort on our parts, five months after my original email, the university agreed to provide live translation over a phone hotline in Spanish, Mandarin, Portuguese and Somali. The news story that ran soon after was first published under the title "After A Push—And Shove—From Students, University Will Offer Commencement Translations." A couple of days after the publication, the headline was changed to remove the words "And Shove." Later, during an event at commencement week, the Dean of Students said to families in attendance that the university was offering live translation at commencement for the first time in its 147-year history. She signaled that this was "a sign of the university's commitment to diversity." I was furious. There was no mention of the fact that it was a student-led effort and that we lobbied the university for months in order to make it happen. On commencement day, multiple family members of mine used the translation service in Portuguese. For us, it was a huge success and I cannot picture my commencement ceremony without it. Despite their proclaimed commitment to diversity, ESU has not offered multilingual translation services at commencement ceremonies since then.

Beyond ESU

After graduation, I leveraged what I learned at ESU and continued engaging in activism at my place of employment, in my community, and also at the national level. While I had some success advocating for my then employer to support DACA-recipient employees, the lack of national legislative progress pushed me to a state of hopelessness. During this time, I suffered a severe depressive episode and received intensive medical treatment for three months. Consequently, I retreated from activism for about a year and half. Currently, I am attending graduate school where I am focusing my emerging scholarship on the educational journeys of my undocumented community.

Analysis

We situate Daniella's testimonio in conversation with [Logan et al.'s \(2017\)](#) spectrum of civic engagement to examine the different types of activism she took on as an undocumented student. Equally, we identify and discuss our conceptualizations of activism, advocacy, allyship, and usurping throughout Daniella's experiences. Incorporation of CRT's interest convergence is key in our analysis because it provides a historical understanding of how institutions respond when the status quo is challenged.

Existence as resistance: origins of an activist

[Logan et al. \(2017\)](#) state that student activism is often a response to hostility experienced throughout the campus climate. Hostility can take many forms, including erasure. The erasure of undocumented student experiences at the Latinx Success Conference was stark. Latinxs make up approximately 48% of undocumented college students in the US ([Feldblum et al., 2021](#)), yet there were no events directly related to this prominent student population at the conference. The only time it was mentioned was toward the end of the ESU student panel when a teacher in the audience asked about it, seemingly, as a proxy question for his students in attendance. Even then, the student panel failed to answer the question. Daniella responded to this erasure by openly sharing her accomplishments as an undocumented college student. By sharing her success, she demonstrated to the audience (undocumented students in particular) that it was possible to do well at a prominent institution. Her actions countered the erasure at the event, the pessimism of its keynote speaker, and the apathy expressed by the student panel. Daniella's actions may be understood as "existence as resistance" because it affirmed the ability of undocumented students to be on ESU's campus. According to [Logan et al. \(2017\)](#), the very presence of marginalized people within a hostile context can be conceptualized as activism. This form of activism does not require intentionality or strategy. Daniella did not attend the event with plans of sharing her status. She felt compelled to act and risked the security of keeping her status a secret. Daniella points to her actions at the conference as the birth of her activism.

The high school teacher who prompted the question also played an important role. His proxy question introduced the topic of undocumented students in an event that had erased their presence. We consider this teacher's actions as that of an advocate. He wielded his relative privilege in that situation to speak up in support of undocumented students.

Sacrifice in frontline activism

USA students publicly disclosing status was a monumental step and a critical form of activism. By participating in the radio interview, they all sacrificed their safety (i.e., anonymity) in an increasingly anti-immigrant climate. In Daniella's case, sacrificing her anonymity also meant risking that administrators or staff at ESU would find out about her status and possibly revoke her

scholarships. Her willingness to sacrifice her financial support is emblematic of our position that sacrifice is intrinsic to activism. While not every member of the USA group chose to participate in the interview or later meet with Dr. West, we still recognize them as engaging in the “existence as resistance” form of activism.

Daniella felt a greater sense of protection from deportation for her (via DACA) and her family (via their US citizenship) in comparison to some of her USA peers. This protection coupled with the belief that she had a responsibility to act compelled her to do the radio interview and thus engage in “frontline activism” which is higher visibility and involves a greater degree of sacrifice and discomfort.

Usurping the moment and the movement

The radio interview exposed Daniella and others at USA to another risk: that of co-optation. Michael and Andrew were not members of USA, nor had they worked previously with the group in any capacity. Rather, they sought out Daniella immediately after the election outcome, prompted an “emergency meeting”, and invited themselves to where USA was gathering to recruit them to take part in a public rally. We find their approach inconsiderate and inappropriate. More specific to our analysis, their actions are emblematically opportunistic and a usurpation. They contacted Daniella during an emotional time fraught with uncertainty and disregarded her response under the pretext of having the duty to act immediately. Similarly, they neglected the group’s need to process their emotions and capitalized on that vulnerability to recruit speakers for their upcoming event; they usurped the moment.

At all subsequent planning events, Michael and Andrew demonstrated their unwillingness to share their positions of power with the USA members. They dictated what aspects of the USA members’ stories would be shared and overlooked suggestions from the undocumented activists on when, how, and where to mobilize. They pushed Daniella and her undocumented peers to be in constantly visible and high commitment roles, with no regard to the disruption the election results had brought to their lives. At public events, Michael and Andrew treated Daniella’s personal story as a commodity to be shared with anyone who wanted something to rally behind in their anti-Trump demonstrations. The two self-proclaimed organizers, however, never discussed any security or support services in the event of threats, harassment, or detention. They conveniently recognized the human appeal in her story but not the need to protect her humanity.

We now recognize Michael and Andrew’s ability to dedicate themselves full time to organizing is not indicative of their investment in undocumented people’s rights. Rather, it demonstrates the multiple privileges that they channeled for self-benefit. These privileges become starker when juxtaposing their actions with Daniella’s involvement and the risks associated with it. Daniella perceived her involvement as a matter of survival while Michael at one point bolstered that he had different sources of funding to support his organizing efforts.

During this time, Michael and Andrew fielded numerous media interviews which helped them grow their network and influence as organizers. They have since appeared on global news outlets and are now public figures in academia and in the nonprofit sector, respectively. Paradoxically, Daniella could not so much as highlight her involvement in the rallies and community meetings on her résumé because doing so would signal to possible employers that she is undocumented. In this way, Michael and Andrew also usurped any professional possibilities that could come from the leadership experiences of organizing. This dynamic is not exclusive to undocumented activism in US. As [Karaliotas and Kapsali \(2021\)](#) point out uneven power relations within immigrant rights movements globally often result in privileged allies excluding and speaking on behalf of the marginalized.

Daniella’s willingness to openly advocate for herself and her community was essential to the success of the rallies and community meetings but was never credited. Michael and Andrew had the opportunity to provide support and mentorship to undocumented activists as well as the privilege and resources to become effective advocates, but their actions demonstrate an unwillingness to channel those privileges for the benefit of others. Instead, they benefited from the sacrifice of activists and usurped any and all prestige that came as a result of their shared organizing efforts.

Perceiving “wins”

The radio interview with USA students was pivotal; it publicly put the onus on ESU to recognize its undocumented students and their fears if Trump was elected. Students seized this opportunity and secured a meeting with Dr. West to discuss the university’s role in supporting them and other undocumented students. Ysabel was instrumental, but not indispensable, in creating that meeting. Once at the meeting, her actions shifted to passive. When Dr. West evaded commitments or shifted the conversation, she remained quiet. She did not interject, advocate for a point the group was making, or redirect Dr. West back to the discussion. The extent of her contributions at the meeting was the prompting for a group picture. In this way, Ysabel’s actions are allyship. She publicly supported USA’s cause, had a minor level of involvement, but remained quiet at pivotal moments. Even though she urged the group to consider the meeting as “a win”, she did not channel any of her privileges to advocate for change.

Our analysis also maintains that ESU’s proposed support of undocumented students is a clear example of interest convergence. The group championed public institutional support for all undocumented students. Dr. West instead ceded to discreetly promote information about existing scholarships eligible to DACA students and create a largely symbolic DACA liaison position. ESU’s concessions yielded a measure of change but they were not systemic and largely served the university and Dr. West’s public image. The critical shortcomings with the concessions are that they exclude undocumented students who are ineligible for DACA and/or

failed to commit support if DACA status was rescinded. By not making any changes public, only students who know Ysabel (or any of the other few knowledgeable staff) were privy to that information. In this way, they have a hidden prerequisite for eligibility.

The DACA liaison position is characteristic of interest convergence. The title itself situates DACA students outside of the university and thus the need for a liaison. The position is also largely symbolic as evidenced by the lack of rigor in properly staffing it and the limited responsibilities it carried. In 2017, when universities nationwide were demonstrating their support for undocumented students as a response to the Trump administration's termination of DACA, ESU published a vague statement referencing the position as evidence of the university's support for DACA students but did not include eligible scholarships or other concrete supports.

Interpreting signs of the university's commitment to diversity

ESU's response to a multilingual commencement is another telling example of interest convergence. ESU agreed to the translation services Daniella proposed only because it could benefit the university's public image of inclusion and diversity. Months after Daniella's inquiry was initially ignored, the administration replied to her request as a "disservice to many" and rejected it. It was not until the reporter doing the news story reached out to ESU administrators for comment that they were open to discussion. We gather that the news story created public pressure which then facilitated the meeting with students. In this regard, we consider Astrid's actions to be advocacy because she leveraged some of her professional resources to assist Daniella. By connecting her to the appropriate reporter to pitch the story of the multilingual commencement, Astrid intentionally channeled her power in support of activism.

By the time the meeting with students was agreed to, all of the labor had already been completed. Daniella and her peers presented the results of their research and strategic plan. ESU did not have to dedicate any resources. Administrators simply had to agree to implement the plan. ESU then swiftly claimed ownership of the idea. The Dean of Students' public statement that this was a "sign of the university's commitment to diversity" is evidence of how interests converged only to benefit the university.

The legacy of progress

Regardless of setbacks, instances of usurping, and interest convergence, Daniella and the USA group's efforts have led to significant measures of change. The USA's news story raised critical awareness about the issues undocumented students face. They then leveraged that awareness into a meeting with Dr. West which created the DACA liaison position. Albeit a mostly symbolic position, it may serve as a starting point for undocumented applicants at ESU to gather information about resources. With institutional support, the position itself can grow into a viable asset for undocumented students at ESU.

The pressure for institutional support continues. For example, while writing this article, we learned that a new undocumented student group has formed at ESU. The Coalition of Undocumented and Allied Students (CUAS) is an officially recognized student group at the university. This designation grants them access to institutional resources which they can then utilize to promote initiatives and other support that was unavailable to the USA group. The group's title articulates its commitment to undocumented students at large (not exclusively DACA) which is indicative of progress. Lastly, while ESU has failed to continue offering multilingual ceremonies, they cannot deny their ability to have them. Anyone who wishes to advocate for them in the future can readily find the history of it being done at ESU. The ability to reference a change in the status quo is important when encountering institutions who resist change.

Conclusion

We conclude by offering ways the reader can act after bearing witness to Daniella's testimonio. People who have the desire to be involved in social causes should be mindful of the stories and contributions of those directly impacted by the injustices being fought. Self-awareness and critical reflection are paramount in preventing the passion felt toward a cause from being distorted into thinking of one's self as a savior to others.

For marginalized students feeling uneasy about engaging in public activism, remember that your presence in exclusionary institutions is contributing to changing the status quo. We encourage undocumented youth, in particular, to recognize themselves as powerful, knowledgeable, and in control of how they choose to engage in activism. Also, be mindful of how common usurping occurs and protect yourself by standing firm in your boundaries and convictions.

We encourage educators especially to channel personal privileges and institutional power in support of undocumented students. Educators are uniquely situated to be powerful advocates that can use their access to administrators, faculty, and staff to relay concerns and/or advocate faithfully on students' behalf. They can also mentor students and help them cultivate skills that they can then expand on in their roles as activists. While the actions of allies can be beneficial to social justice movements, our analysis demonstrates that advocates are far more impactful. When deliberating one's role in fighting against injustices, perhaps the great Brazilian pedagogue, Paulo Freire, stated it best "Liberation is a praxis: the action and reflection of men and women upon their world in order to transform it."

References

- Abrego, L.J., Negrón-Gonzales, G., 2020. *We are Not Dreamers: Undocumented Scholars Theorize Undocumented Life in the United States*. Duke University Press.
- Alfred, C., 2020. Inspired by the US, Europe's "paperless" Children are Speaking Out. *The Guardian*. Available at: <https://www.theguardian.com/world/2020/aug/03/europe-paperless-young-people-speak-out-undocumented-dreamers>. (Accessed 14 March 2022).
- Allard, E., 2018. "A promise to support us:" undocumented experiences on a sanctuary campus. *Assoc. Mex. Am. Educ. J.* 12, 53.
- Bell, J.D.A., 1980. *Brown v. Board of education and the interest-convergence dilemma*. *Harv. Law Rev.* 93.
- Bennet, B., 2016. Yes, Trump Can Boost Deportations and Gut the Dreamer Program for Young Immigrants. *Los Angeles Times*.
- Burciaga, R., Tavares, A., 2006. Our pedagogy of sisterhood: a testimonio. In: *Chicana/Latina Education in Everyday Life: Feminista Perspectives on Pedagogy and Epistemology*, pp. 133–142.
- Castrellón, L.E., Reyna Rivarola, A.R., López, G.R., 2017. We are not alternative facts: feeling, existing, and resisting in the era of Trump. *Int. J. Qual. Stud. Educ.* 30, 936–945.
- Chabram-Dermersian, A., De La Torre, A., 2008. *Speaking from the Body: Latinas on Health and Culture*. University of Arizona Press.
- Collet, B.A., Macías, L.F., 2021. Sacred mobilisations: sanctuary churches and sanctuary schools. *Compare* 1–18.
- Collins, P.H., 1998. It's all in the family: intersections of gender, race, and nation. *Hypatia* 13, 62–82.
- Connor, P., 2016. *International Migration: Key Findings from the US, Europe and the World*, vol. 15. Pew Research Center.
- Costanza-Chock, S., 2014. Out of the Shadows, into the Streets!: Transmedia Organizing and the Immigrant Rights Movement.
- Covington, B., 2008. Is postsecondary access for undocumented immigrants an important right-how the United States and Europe differ. *Geo. Immigr. L.J.* 23, 403.
- Crawford, E.R., Dorner, L.M. (Eds.), 2019. *Educational Leadership of Immigrants: Case Studies in Times of Change*. Routledge.
- Delgado Bernal, D., 1998. Using a Chicana feminist epistemology in educational research. *Harv. Educ. Rev.* 68, 555–582.
- Delgado Bernal, D., 2002. Critical race theory, latino critical theory, and critical raced-gendered epistemologies: recognizing students of color as holders and creators of knowledge. *Qual. Inq.* 8, 105–126.
- Delgado Bernal, D., 2009. Introduction: our testimonios as methodology, pedagogy, and a messy work in progress. In: *Unidas we heal: Testimonios of Mind/Body/Soul. Latinas Telling Testimonios*, pp. 4–6.
- Delgado Bernal, D., Burciaga, R., Flores Carmona, J., 2012. Chicana/latina "testimonios": mapping the methodological, pedagogical, and political. *Equity & Excell. Educ.* 45, 363–372.
- Dougherty, K.J., Nienhusser, H.K., Vega, B.E., 2010. Undocumented immigrants and state higher education policy: the politics of in-state tuition eligibility in Texas and Arizona. *Rev. High. Educ.* 34, 123–173.
- Feldblum, M., Hubbard, S., Lim, A., Penichet-Paul, C., Siegel, H., 2021. Undocumented Students in Higher Education: How Many Students are in US Colleges and Universities, and Who are They? *New American Economy and Presidents' Alliance on Higher Education and Immigration*.
- Fierros, C.O., Delgado Bernal, D., 2016. Vamos a platicar: the contours of pláticas as Chicana/Latina feminist methodology. *Chicana/Lat. stud.* 15, 98–121.
- Flores Carmona, J., Hamzeh, M., Bejarano, C., Hernández Sánchez, M., EL Ashmawi, Y.P., 2018. Pláticas~testimonios: practicing methodological borderlands for solidarity and resilience in academia. *Chicana/Lat. Stud.* 18, 30–52.
- Fox, J.A., 2006. *Invisible No More: Mexican Migrant Civic Participation in the United States*.
- Gonzales, R.G., 2008. Left out but not shut down: political activism and the undocumented student movement. *Nw. J. L. Soc. Pol'y* 3, 219.
- Gonzales, R.G., Sigona, N., Franco, M.C., Papoutsis, A., 2019. *Undocumented Migration*. John Wiley & Sons.
- Gonzalez, F.E., 1999. The Formations of "Mexicaneness: Trenzas de Identidades Múltiples". *The Development of Womanhood among Young Mexicanas: Braids of Multiple Identities*. University of California, Davis.
- Ibarra, C., Rivera, A., 2019. *The Infiltrators*.
- Karaliotas, L., Kapsali, M., 2021. Equals in solidarity: orfanotrofio's housing squat as a site for political subjectification across differences amid the "Greek crisis". *Antipode* 53 (2), 399–421.
- Ladson-Billings, G., 2003. Racialized discourses and ethnic epistemologies. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Landscape of Qualitative Research: Theories and Issues*. SAGE Publications, Thousand Oaks, Calif.
- Logan, G., Lightfoot, B.A., Contreras, A., 2017. Black and Brown millennial activism on a PWI campus in the era of Trump. *J. Negro Educ.* 86, 252–268.
- Macías, L.F., 2018. The scheme game: how DACA recipients navigate barriers to higher education. *Educ. Stud.* 54, 609–628.
- Madia, S.M., 2021. Why Undocumented Migrants Went on Hunger Strike in Belgium. *Refugees | Al Jazeera*. Available at: <https://www.aljazeera.com/features/2021/9/19/why-undocumented-migrants-went-on-hunger-strike-in-belgium>. (Accessed 14 March 2022).
- Montoya, M.E., Valdes, F., 2008. Latinas/os and the politics of knowledge production: latcrit scholarship and academic activism as social justice action. *Ind. L. J.* 83, 1197.
- Moraga, C., Anzaldúa, G., 2015. *This Bridge Called My Back: Writings by Radical Women of Color*. Suny Press.
- Muñoz, S.M., Vigil, D., 2018. Interrogating racist nativist microaggressions and campus climate: how undocumented and DACA college students experience institutional legal violence in Colorado. *J. Divers. High. Educ.* 11, 451.
- Nicholls, W., 2013. *The DREAMers: How the Undocumented Youth Movement Transformed the Immigrant Rights Debate*. Stanford University Press.
- Nienhusser, H.K., Reyna Rivarola, A., Salazar, C., 2020. Association for the Study of Higher Education (ASHE) & ACPA: College Student Educators International Joint Response to the Supreme Court of the United States Decision on the Deferred Action for Childhood Arrivals (DACA) Program and Position Statement on Undocumented College Students. ASHE and ACPA. <https://www.ashe.ws/Files/Position%20Taking/ASHE%20>.
- Olivas, M.A., 2005. 2 The Story of Plyler v. Doe: The Education of Undocumented Children and the Polity. *No Undocumented Child Left Behind*. NYU Press.
- Olivas, M.A., 2020. *Perchance to DREAM: A Legal and Political History of the DREAM Act and DACA*. NYU Press.
- Pallares, A., Flores-González, N., 2010. ¡Marcha! : Latino Chicago and the Immigrant Rights Movement. University of Illinois Press, Urbana, Ill.
- Perez, W., 2009. *We are Americans: Undocumented Students Pursuing the American Dream*. Stylus Publishing, LLC.
- Pérez Huber, L., Muñoz, S., 2021. *Why They Hate Us: How Racist Rhetoric Impacts Education*. Teachers College Press.
- Pyle, A., Romero, S., 1994. CALIFORNIA ELECTIONS/PROPOSITION 187: Student Walkouts Transform a Political Issue into a Cause. *Los Angeles Times*.
- Seif, H., 2011. "Unapologetic and unafraid": immigrant youth come out from the shadows. *N. Dir. Child Adolesc. Dev.* 2011, 59–75.
- Solórzano, D.G., Bernal, D.D., 2001. Examining transformational resistance through a critical race and LatCrit theory framework: Chicana and chicano students in an Urban context. *Urban Educ.* 36, 308–342.
- Solorzano, D.G., 1997. Images and words that wound: critical race theory, racial stereotyping, and teacher education. *Teach. Educ. Q.* 5–19.
- Swerts, T., Nicholls, W., 2021. Undocumented immigrant activism and the political: disrupting the order or reproducing the status quo? *Antipode* 53 (2), 319–330.
- Truax, E., 2015. *DREAMers: An Immigrant Generation's Fight for Their American Dream*. Beacon Press.
- Villazor, R.C., 2019. Reflecting on California and Prop. 187: from the anti-immigrant state to the sanctuary state. *UC Davis Law Rev.* 53, 2015.
- Zimmerman, A.M., 2011. A dream detained: undocumented Latino youth and the DREAM movement. *NACLA Rep. Am.* 44, 14–17.
- Zong, J., Batalova, J., 2019. *How Many Unauthorized Immigrants Graduate from US High Schools Annually?*.

Language, race, and legitimacy: the Toronto Teachers' Federation's battle against heritage-language programs in Toronto schools, 1982–1986

Jeff Bale, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Theoretical and methodological framing	386
Historical and policy context	386
The emergence of heritage languages as a new social “problem”	386
Early experiments in Toronto and new policies in Ontario	387
Round 1: The TTF endorses bilingual education	388
Round 2: The TTF's fight to end daytime HL classes	389
The Work Group on Third Language Instruction and its final recommendations	389
TTF engagement with the work group	391
Backlash	393
Conclusion	395
References	396
Primary sources	396
Secondary sources	397

On Friday, May 11, 1984, some 1350 members of the Toronto Teachers' Federation gathered to decide who their next president would be. As incumbent Jack McFadyen and contender Clem Marshall gave their speeches, one issue separating the candidates took center stage: heritage-language classes in the Toronto Board of Education, in particular the Board's plans to expand the number of elementary schools that offered such classes during the school day. Heritage-language (HL) classes were designed to teach languages such as Cantonese, Greek, Gujarati, Italian, Portuguese, Spanish, and Ukrainian, languages all widely spoken in Toronto. The provincial policy authorizing these classes allowed for programs to be held after school, on weekends, or during the school day by extending the timetable by 30 min. It was this latter option that had become so controversial; indeed, at the time of this meeting, the union was already three months into a work-to-rule job action to protest daytime HL classes.

At the meeting, McFadyen promised to toughen the union's objection to daytime HL classes, to “bursts of cheers and applause” from his audience (Brown, 1984c, p. A1). Not only would this issue be the sole non-monetary question the union would raise in forthcoming contract negotiations, he noted. But also, McFadyen stressed that by November 1984, the union would be in a legal position to go on strike to force an end to daytime HL classes. Had it occurred, this strike would have been the union's first. “We're sick of teacher bashers who say teachers are racist,” McFadyen argued. “We're not fighting heritage languages—we're fighting the fact the board slapped this on us without negotiating” (p. A1).

For his part, Marshall opposed both the on-going job action and threats of a strike. Born in Guyana and now a teacher of French as a second language, Marshall rooted his argument in his own lived experiences:

As one who has been in the so-called [multicultural] melting pot, I can tell you that melting pot can get too hot sometimes and a few of us have got burned. We've got to consider the heritage program in light of the greater good it provides.

Brown (1984c, p. A7; inserted word in the original).

Marshall argued daytime HL classes are “worth fighting for—not against—because it helps immigrant children gain self-esteem” (p. A7). For him, the job action and strike were bound to lose, because “parents are no longer behind us. We have lost them as allies with our behaviour” (p. A7). His comments received “sparse” applause. Indeed, he lost the election that night with 147 votes against McFadyen's 1368 (p. A7).

How is it that a modest program to support heritage languages would become the central issue animating a union election, even potentially triggering the union's first ever work stoppage? To what extent were McFadyen's claims accurate? Was the TTF only challenging a specific model of HL classes, and not the very idea of HL education itself? Had the Toronto Board of Education really just “slapped” this model on them without negotiations? Perhaps most complicated of all, were their protests evidence of racism? At first glance, these questions may appear too focused on a distant union scuffle from almost 40 years ago. By sitting with these questions, however, we instead see how they reveal fundamental dynamics about race/racism, language, and legitimacy that continue to shape school life in Toronto. Which languages—better asked, *whose* languages—belong in school? In which spaces, at what times, and to what ends? And perhaps most important: *who gets to decide*?

In this article, I analyze two distinct moments in the TTF's interventions around these questions. The first of these precedes the heated meeting described above by a decade, as the Board had just begun to study and take action around the gap between the Anglo

dominance of school culture and the curriculum, and the reality of racial and linguistic diversity among its students. As we will see, at that time the TTF stood firmly on the side of bilingual education models that had emerged in North America. Over the next decade, the Board would continue studying this issue, and the province itself would intervene in 1977 with its *Heritage Languages Program* policy authorizing HL classes. However, once actual HL programs existed, the TTF took radically different actions—reflected in John McFadyen’s words above and the “bursts of cheers and applause” from his union siblings. This time, teachers closed ranks and challenged the legitimacy of “outsiders” in determining what happens in Toronto schools. As the opening description suggests, this put teachers—who in the early 1980s were almost exclusively white and Anglophone—squarely at odds with demands made by racialized and multilingual parents and other community members. I analyze these contradictions so as to identify lessons from this conflict that might inform current efforts of the TTF’s successor, the Elementary Teachers of Toronto, to combat systemic racism in Toronto schools.

Theoretical and methodological framing

This paper tells one story and aims to connect it to larger questions of teacher-union activism to challenge systemic racism, both within its ranks as well as in schools. The paper is part of a larger project, funded by the Social Sciences and Humanities Research Council of Canada, that examines heritage-language policies as sites for regulating linguistic and racial difference in Canadian schools. The study draws on materialist, anti-racist (Flores and Chaparro, 2018) and political-economic (Bale, 2015) approaches to theorizing language-education policies. The global economic crisis in 2008 forced applied linguistics as a field to explore notions of class in the study of language (e.g., Block, 2014; Heller and Duchêne, 2012; Holborow, 2015). This so-called material turn has rekindled debate about the relationship between class and race as analytical categories, and how language interacts with class and race in lived experience. Following Anyon (2005), Flores and Chapparo (2018) proposed a materialist, anti-racist approach that is both historical and intersectional. This approach starts with the historical conditions that produce a given policy as a putative solution to a perceived social problem. Furthermore, their approach assumes that language-education policies only have meaning in relation to other policies related to immigration, housing, health, employment, etc. The political-economic approach I have developed (Bale, 2015, 2016) understands language-education policies as more than just ideological reflections of an underlying social order; rather, they also serve as venues in which race and class co-construct that social order in the first place. This framework contrasts with the historical-structural approach (Tollefson, 1991) used in most historical inquiry into language policy, which typically frames language policies as top-down mechanisms (Shohamy, 2006) that elites use to control society. Instead, a political-economic approach is analytically more open: it reveals how language-education policies can function both as tools for imposing a linguistic or racial order, and also as vehicles for contesting that order (see Bale, 2016).

The larger project adopts raciolinguistic perspectives (Flores and Rosa, 2015; Rosa and Flores, 2017) to theorize the relationship between language and race/racism. Raciolinguistics frames language and race as ideological categories that interact to structure society and maintain social inequalities. This perspective clarifies how systemic racism and (settler-)colonialism are constituent elements in co-constructing global capitalism. This framework reveals how contemporary language ideologies and language practices intersect with racism and colonialism in ways that often seem “natural” to us. That is, language, racism, and colonialism collaborate to organize and justify the unequal distribution of resources in society (Flores and McCauliffe, 2020). Flores (2016, 2019) has applied raciolinguistics to explore the continuity of *anti-Latinx* racism within bilingual-education programs in the United States, despite the liberationist social movements of Puerto Rican and Chicana activists in the 1960s–1970s that led to bilingual-education rights in the first place. In the project informing this paper, I use raciolinguistic analysis to explore a similar contradiction in Canada: how do we account for the *subordination* of the very language-education policies that were ostensibly designed to *support* the learning of minoritized and racialized languages?

The study is designed as a policy genealogy (Haque, 2019; Gale, 2001). Unlike traditional approaches to history, which link a present-day phenomenon to its origins in the past or view the present as the result of continuous development, genealogies instead focus on discontinuity (Gale, 2001). As Garland (2014) has described, genealogical research proceeds by “trac[ing] the erratic and discontinuous process whereby the past became the present: an often aleatory path of descent and emergence that suggests the contingency of the present and the openness of the future” (p. 372). By focusing on discontinuity, rather than seeking a linear progression to history, genealogical analysis assumes the *openness of history*. No matter how enduring they may appear, the structures and ideologies we live with today are the result of past conflict (Tamboukou, 1999). In this way, genealogical study of the past is a kind of “effective history” in the Foucauldian sense in two ways: (1) it reveals the power relations upon which the present relies, and the contingent, historically situated processes that became our present; and (2) it challenges us to think and act differently about the present (Fendler, 2008)—in this case, about racial and linguistic difference in schools.

Historical and policy context

The emergence of heritage languages as a new social “problem”

The Canadian government introduced the Official Languages Act (OLA) in 1969 and the Multiculturalism Policy in 1971. As Haque (2012) argued in her ground-breaking study of the Royal Commission on Bilingualism and Biculturalism (1963–69) that led to these two policies, official bilingualism was intended to resolve a fundamental contradiction underlying the Canadian state, by

reorganizing its white-settler identity to include both Anglophone and Francophone settlers on more equal terms. At the same time, official bilingualism antagonized another central contradiction underlying the Canadian state: both through the workings of the Royal Commission and in its central policy outcomes, official status for First Nations, Métis, and Inuit peoples and for non-Anglo, non-French immigrant communities was denied. Pierre Trudeau attempted to resolve this contradiction by proposing *multiculturalism within a bilingual framework*, that is, by allowing Indigenous peoples and immigrants to sustain their *culture* while denying them formal *language* rights. Yet, by elevating only two white-settler languages, official bilingualism formalized the hierarchy of minoritized and racialized languages that had long existed in the Canadian state (Haque, 2012; Haque and Patrick, 2015).

Within this context, *heritage languages* emerged as a new social problem recognizable to Canadians. The federal government's policies on multiculturalism and official bilingualism had the contradictory effect of raising expectations among Indigenous and immigrant communities for linguistic and cultural rights of their own, while simultaneously conferring official status on only two white-settler languages. At the grassroots level, various Indigenous, language-minoritized, and racialized communities pressured boards and ministries to transform curricula to better fit the linguistic and cultural reality of Canadian schools (see Wong and Guo, 2015). As one indication of this activity, HL education policies began to appear across the Canadian state: in Alberta in 1971, Ontario and Quebec in 1977, Saskatchewan in 1978, and Manitoba in 1979 (Cummins, 1983; Cummins and Danesi, 1990). These policies represented the provinces' first efforts to reconcile the reality of multilingual schools with an officially bilingual Canada. That is, HL policies sat—and remain—squarely at the heart of the contradiction between official bilingualism as an *imagined ideal* for Canadian national identity, and the *lived realities* of multilingual, racialized students in Canadian schools.

Despite their obvious significance, these policies have largely been ignored by scholars. The few exceptions include Cummins (1983) and Cummins and Danesi (1990), which provided a national overview, while Tavares (2000) analyzed the shift in the ideological framing of HL programs in Western Canada in the 1990s. In Ontario, besides the other publications that have emerged from this project (Bale, 2019; Bale and Kawaguchi, 2020; Kim et al., 2020), there exist only one dissertation (Berryman, 1986), one personal reflection (Wright, 1997), and one narrative inquiry involving HL instructors (Feuerverger, 1997). Otherwise, Ontario's HL policy and the programs it supported, when mentioned at all, are minor characters in larger stories about migration and multiculturalism (e.g., Gidney, 1999; Masemann, 1984; McCaskell, 2005; Vipond, 2017).

Early experiments in Toronto and new policies in Ontario

As indicated in Table 1, there was a great deal of experimentation in the 1970s involving linguistically- and culturally-relevant programming in Ontario generally and in Toronto specifically (Jean-Pierre and Nunes, 2015). This included private heritage-language classes established by various immigrant communities (Cummins, 1983) and Black Cultural Heritage programs designed to counter the racism that Black youth experienced at school (Bramble, 2000). The Toronto Board of Education also experimented with bilingual-education programs, and from 1974 to 82 struck three separate task forces to study the relationship between language, culture, race, and education (about which, more below). Although each of these “work groups” framed its inquiry differently—first by focusing on multiculturalism, then on racism as *race relations*, and finally on language—they all understood that policy and curriculum change must start from understanding the interconnections among language, race and culture.

Within the midst of this experimentation and sustained inquiry at the Board, the Ontario government initiated its *Heritage Languages Program* (HLP) in 1977. The government surprised many by proposing the policy in March, ahead of provincial elections on June 9. It then rushed the policy into place with a memo on June 15 to heads of Ontario school boards announcing the new program. This slapdash approach to policy-making has been interpreted as an attempt to expand electoral support for the governing Progressive Conservatives by bringing (European) immigrant communities into the fold (see Gidney, 1999).

Table 1 Key context and policy activity regarding race, language, and belonging in Toronto schools, early 1970s to early 1980s.

At the federal level		At the Toronto board of education	
Multiculturalism within a bilingual framework era begins	Early 1970s in Toronto	- Privately-run heritage-language programs - First black culture school, organized by the black education project (January 1970) - First board-sponsored pilot programs in transitional bilingual education for Cantonese and Italian	
	- Official Languages Act (1969) - Multiculturalism policy (1971)	Work group on Multicultural programs (1974–76)	<i>Bias of Culture</i> issue paper (1974) - Draft report (1975) - Final report (1976)
At the provincial level	Sub-Committee on race relations (1978–79)		- Draft report (1978) - Final report (1979)
Heritage Languages program begins (June 15, 1977)	Work group on third Language instruction (1981–82)		<i>Challenge of Languages</i> issue paper (1981) - Draft report (1981) - Final report (1982)

Indeed, the HLP was a kind of political compromise. On one hand, the policy, which funds 2.5 h per week of study in a heritage language, was widely and quickly implemented. By the 1980 school year, it supported more than 78,000 students across the province in 3400 classes learning 45 different languages, with most enrollments (55%) in Toronto's public and Catholic school boards (Report of the O.P.S.T.F. Task Force, 1983). On the other hand, many language-minoritized and racialized communities viewed the policy's stipulations as divisive, for example with the Ministry's refusal to include Black Cultural Heritage programs as part of HL funding streams (see Bale and Kawaguchi, 2020). Many also viewed the new policy as further subordinating the status of HLs at school. HLs could not be used to teach curriculum content. Because they were not part of the formal elementary-school curriculum, HL classes had to occur outside the regular school day. Instructors came from the respective HL community, but were rarely certified teachers and thus not members of teachers' unions. Moreover, children of any background could enroll, rendering the policy's objectives unclear: was it intended to sustain ethnolinguistic identity for children raised with a heritage language, or to provide an enrichment activity for Anglophones (see Cummins and Danesi, 1990)? Could it do both at once?

In every way, then, this policy positioned HLs at the margins of school life, structurally separate from the Ontario curriculum and the teachers responsible for teaching it. Language-minoritized and racialized communities responded with sustained advocacy to broaden the scope of HL programs. At the same time, (mostly white) Ontarians filtered their racialized grievances against the policy through tropes of taxpayer concern over public spending, program feasibility, and the fragility of Canada's fledgling official bilingualism (see Bale, 2019). In this way, the HLP became the flashpoint of conflict over how schools should respond to growing linguistic and racialized difference in Ontario. These conflicts waxed and waned from 1977 to 1987, when a minority Liberal provincial government introduced reforms to the HLP that solidified the structure of the program that we live with to this day.

The remainder of this paper focuses on perhaps the most heated and explicitly racialized conflict over the policy and its implementation in the Toronto Board of Education. The analysis is divided into two parts. First, I begin by tracing teacher-union support for bilingual-education programs. This support was expressed most clearly in the Board's first effort to study the relationship between language, culture, race and schooling from 1974 to 76. Second, and by contrast, the outcomes of the Board's third and final work group on this issue in 1981–82 were met with direct and sustained union resistance, to such a level that the union and its members were denounced as racists. The paper concludes with lessons of this analysis for current efforts in the teachers' union to combat systemic racism, both within its ranks and in Toronto schools more broadly.

Round 1: The TTF endorses bilingual education

In June 1974, the Toronto Board of Education established the first of what would be three separate task forces to study the relationship between language, race, and culture in its schools (see Table 1). The Work Group on Multicultural Programs would set the pattern for this work. Each task force would begin by issuing a discussion paper, and then solicit written feedback and organize public hearings on the paper. To ensure that consultations were broad, copies of the paper would be sent to schools, specific Board-employee groups and/or their organizations, parent and community organizations, politicians (provincial and federal) with ridings in Toronto, the media, etc. The work groups would also consult scholarly and practitioner publications from across Canada, the US and Europe related to their inquiry. Based on these consultations and research, the work groups would write a draft report with a series of recommendations for consideration. The draft report would then undergo a similar, if less extensive, consultation process, leading to a final report with recommendations submitted to the Board for deliberation and potential approval.

As the Work Group on Multicultural Programs got underway, it extended an invitation to the Toronto Teachers' Federation (TTF) to participate in the group's deliberations over multicultural programs. The invitation stressed that the work group was "most anxious that the advice and counsel of teachers be clearly heard" (Kerr, 1974, p. 1). The president of the TTF responded with thanks for such a proactive invitation, noting that TTF Executive Board member Muriel Anderson would serve as the liaison between the union and the work group (Wooding, 1974).

In January 1975, Anderson wrote the brief that would become the TTF Executive's interim position with respect to the work group's inquiry. The brief spilled little ink in defending bilingual-education programs, because "informed individuals are presently in favour of such programs in principle" (Anderson, 1975, p. 2). Yet, the brief acknowledged that putting such principles into action is more complicated; as we will see shortly, Anderson's comment was remarkably prescient. The brief endorsed two models of bilingual education to be implemented depending on student needs or interests: *transitional bilingual education* (wherein the less dominant language is used to teach curriculum content in the early grades while students become proficient in the dominant language; thereafter, only the dominant language is used); and what would now be understood as a weak form of *maintenance bilingual education* (wherein a few school subjects, but not all, are taught in the less dominant language for a longer period of time so that the student maintains proficiency in the less dominant language as well).

The brief's support for language-minoritized and racialized students is couched in terms of the "problems" their linguistic, cultural and racial otherness presents at school. While the brief endorsed different kinds of bilingual programming, the rationale for such programs was to best integrate linguistically-minoritized and racialized youth into the Canadian mainstream. This is wholly consistent with the liberal multicultural tradition of bilingual education that Flores (2016) has identified in the United States, which "position[s] bilingual education for Latino students within a larger project of assimilation into [dominant] relations of power" (p. 14). Indeed, the very euphemism common at this time, and used in the brief, to refer to linguistically and racially diverse youth—*New Canadians*—suggests this assimilationist perspective.

Yet even working within a liberal-multicultural framework, the brief was razor sharp in its analysis of the conditions that constrain bilingual education. Anticipating the central issues that would lead to explosive battles over HL classes and the provincial policy that funded them, the brief noted that for bilingual and other linguistically- and culturally-responsive programs to succeed, they require (1) a high degree of community involvement in programs and planning; (2) decentralization and greater autonomy for individual schools; (3) meaningful funding; and (4) qualified staff. On the latter point, it was here that the TTF first articulated its expectation that instructors in bilingual and HL programs be fully certified teachers, a position they would later (and quietly) drop. To underscore their stance, the brief cited minutes of Board meetings to stress that “the Toronto Board had not given high priority to bilingual programs but instead has adopted a passive role” (Anderson, 1975, p. 4). The brief concluded with two short recommendations: that bilingual programs as described in the brief be made available on a pilot basis in select schools to all students who require the support or simply wish to participate; and that the Board make this a high priority by allocating the appropriate funds and qualified staff to make it happen.

The brief was quickly adopted by the TTF Executive and later sent to area representatives, who were then to organize local discussions among TTF members about its substance. It is not clear at this stage in my research whether these local discussions ever happened. What is clear is that the brief stalled: as the Board’s work group was preparing its final report in October 1975, it became known that the TTF brief had only been accepted by the union as an *interim* position. Neither the substance of the brief nor its two recommendations were ever formally adopted as TTF policy (Chown, 1975). I am not yet able to determine why the brief stalled: because of internal union opposition to its content, an oversight in an otherwise busy organization, or perhaps a combination of the two? Indeed, a new TTF president had been elected in the spring of 1975, so perhaps this reflected a new stance toward the proceedings of the work group.

What I can note is the gap between the work group’s *draft* report, which included several recommendations for bilingual programs similar to what the TTF had proposed, and the *final* report, in which those recommendations were dropped. In the end, the work group itself settled on the compromise of offering HL classes that taught the language as a subject only (i.e., not used to learn curriculum content). This is the same model that the province would endorse the following year in its Heritage Languages Program policy. In other words, if there had not been broad support within the union for full bilingual programs, it seems that support was also lacking among work group members themselves, such that bilingual-education programs were sidelined. Not until the third and final Board task force in 1981–82 would this question be revisited again. It was this final round of deliberations on the relationship between language, race and culture that led to explosive conflicts in Toronto.

Round 2: The TTF’s fight to end daytime HL classes

If the Board thought the question of heritage languages, or of the broader relationship between race, language and culture in its schools, was settled with the recommendations approved from the first work group, it would soon be proven very wrong. Between 1977, when the province’s Heritage Languages Program appeared, and the spring of 1982, when the Board began to implement recommendations from its third and final task force, the question of where—indeed *whether*—heritage languages (HL) belonged in publicly-funded schools was hotly contested. The province’s new policy was greeted by scores of angry letters written mostly by white Ontarians to the Ministry of Education, while right-wing tabloids such as the *Toronto Sun* regularly denounced the policy and Toronto’s HL classes as a threat to Canadian unity (see Bale, 2019). In spring 1982, as the Board held public meetings on implementing the recommendations of its final work group, tensions spilled over; Fig. 1 presents a picture of parents and other community members at one such meeting.

The media described these meetings as “explosive confrontations” over a “language battle ready to explode” (Spears, 1982, p. A6), with “voices raised in anger against foreign tongues” (Miller, 1982, p. G6). Harish Shah, a parent of three children in Toronto schools, described racist threats made by hecklers after he spoke at one of these meetings: “They told me next time they’ll bring white sheets” (Maychak, 1982, p. A6). A board trustee reported receiving phone calls at home threatening his safety (Matas, 1982) after another of these meetings.

It was in this heated, highly racialized context that the Toronto Teachers’ Federation threw all its weight behind ending daytime HL classes. In the second part of my analysis, I discuss: (1) the third and final work group that set the policy stage for this battle; (2) how the TTF had engaged with the work group; and (3) the terms on which the union waged its battle against daytime HL classes.

The Work Group on Third Language Instruction and its final recommendations

The Work Group on Third Language Instruction, the last of the Board’s three task forces, issued by far the most comprehensive and progressive recommendations as related to the relationship between language, race and culture in Toronto schools. For example, in its recommendations, the work group used the language of *bi-* and *multilingualism* to avoid any sense of prescribed order in how multiple languages are learned or used by Toronto youth. It is worth quoting from the final report at some length to give a sense of the work group’s thinking. Fig. 2 presents the work group’s first two recommendations in full.

Two aspects of this excerpt require comment. By “comprehensive,” the work group meant that the Board should adopt a *single* policy to encompass the manifold dimensions of multilingualism in Toronto: learning English as an additional language or standard Canadian English as an additional dialect; learning about historical and contemporary Black language practices as part of Black Cultural Heritage programs; learning Indigenous languages; learning heritage languages; learning French as either a first or

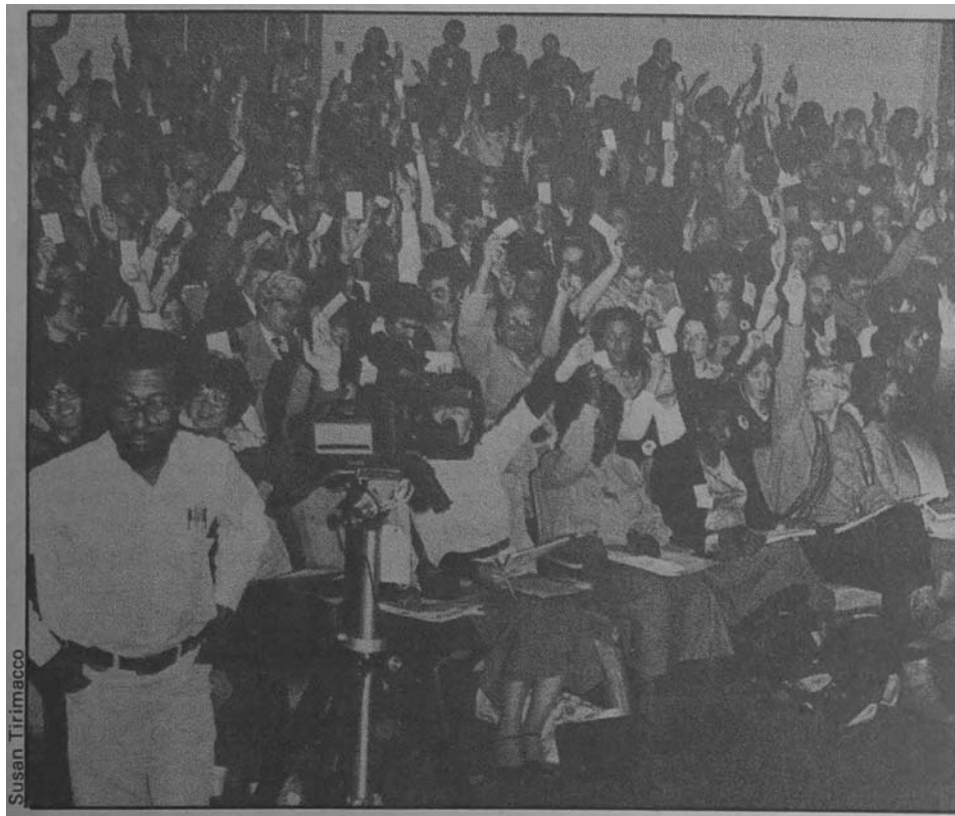


Fig. 1 Photo by Susan Tirimacco of a community consultation meeting over daytime HL classes, 1984; Davis (1984, p. 2).

additional language (including, as recommendation 6 would articulate, establishing Black Cultural Heritage programs in French); and, at the secondary level, learning modern and/or classical languages. Hitherto, the Board had treated each of these dimensions as discrete; indeed, while the Board's previous task forces addressed each program type in their reports, they addressed them as distinct issues affecting different student subgroups. By contrast, in this report these various dimensions of language education were construed as a single phenomenon of multilingualism that was relevant to all Toronto students. From the work group's perspective, Toronto schools therefore bore the responsibility to sustain these multilingual practices. Second, this excerpt indicates the work group's perspectives on *who* had the final say in determining what one's bi- or multilingualism should sound like. According to the logic of this final report, *bilingualism* would be a requirement for all students, in which case the Board assumed the authority to issue such a requirement. Whether students chose to "develop proficiency" in additional languages, (i.e., to expand their *multilingual* abilities) was ultimately up to them and/or their parents.

If this final work group adopted such inclusive and progressive positions with respect to multilingualism, it is in large part due to the activism of language-minoritized and racialized supporters of HL and Black Cultural Heritage programs. In the five years between the start of the province's HLP policy and this work group, the Board faced sustained pressure to allow HLs to be used as a medium of instruction in bilingual programming. For many advocates, if HLs were not used to teach curriculum, alongside English and French, they would always be relegated to the margins of school life, both symbolically and in practice. Bilingual HL programs however, would contravene Ontario's *Education Act*, which allows only English and French to be used to teach the curriculum. As such, there had been multiple efforts to amend the act, including 4 separate private members' bills in Ontario's legislature between 1978 and 1987. In each case, these efforts failed; to this day, the *Education Act* restricts the language of instruction to English and French only, with very few exceptions.

Second, the provincial HL policy had always allowed for daytime HL classes (by extending the school day by 30 min and slotting in HL instruction, usually around lunchtime). However, and perhaps by historical accident, most HL classes in the Toronto Board of Education were held after school or on weekends. Advocates pressured the Board to change this by expanding daytime HL programs. If HLs were not allowed for use in teaching curriculum content, then the next-best option for many advocates was to include them as a subject during the school day to bring HLs closer to the heart of school life. Finally, and by contrast, some advocates argued that mainstream schools would never be welcoming of linguistically- and culturally-responsive programs. Instead, they called for magnet schools, in which a single HL would be taught on a bilingual basis with English. The debates around this third demand were sharp and fall outside the scope of this paper. But they raised fundamental questions about whether inclusion and equity were indeed even possible in mainstream school environments. Taken together, this advocacy and the related debates directly

1. That the Toronto Board of Education develop and implement a comprehensive language policy.
2. That the Toronto Board of Education endorse and actively promote the improvement and co-ordination of its language programs and develop a language policy with the objective of providing programs for students at all levels of the Toronto school system to enable them to become multilingual, trilingual or bilingual adults for the direct purpose of:
 - (a) assisting all students to develop proficiency in English.
 - (b) assisting all students, including francophone and students of Native Ancestry, to hold safe from loss or deterioration, to develop and, in those instances which render it necessary, to relearn their original language or languages.
 - (c) assisting all students to attain a proficiency in at least one language other than their own in accordance with the definition of proficiency levels provided by the Ontario Ministry of Education.
 - (d) assisting all students to attain proficiency in one or more additional languages, in accordance with the definition of proficiency levels provided by the Ontario Ministry of Education, as they, or their parents or guardians if the students are minors, consider appropriate.

Fig. 2 First two recommendations from the Work Group on Third Language Instruction final report. Source: [Toronto Board of Education \(1982, p. 71\)](#).

shaped the work group's inquiry, as reflected in its final recommendations that were so comprehensive in nature and so focused on expanding daytime HL classes.

TTF engagement with the work group

The Work Group on Third Language Instruction was constituted somewhat differently from the Board's first task force. This time, TTF representative Imogene Walker was a member of the work group, not merely a liaison as Muriel Anderson had been with the first work group. Moreover, HL education had since become more institutionalized in the Board. The Board had established the Heritage Languages Advisory Committee, with representatives of various language-minoritized and racialized communities in Toronto, to serve as a liaison between these communities and the Board. Indeed, four of the 13 members of the work group were from this Committee. The work group also included Miriam DiGiuseppe, a widely respected member of the Italian-Canadian community who was the central Board administrator responsible for HL programs. Finally, the work group included other Board administrators, seconded staff, a representative of the principals' association, and Board trustees. In this way, the work group consisted of individuals representing the various communities and employee groups with a direct stake in the future of "third-language" instruction.

The TTF was directly engaged with the work group's inquiry from the get-go. As had become the pattern, the work group distributed a discussion paper, *The Challenge of Languages*, for which it sought written and oral feedback. In the TTF's response, the union continued to operate in a liberal-multicultural framework for linguistically-responsive programs, insofar as it understood maintenance of children's home languages primarily as a resource for acquiring English and integrating into the mainstream of Canadian life. However, the brief also called for an autonomous, community-driven approach to local school management and the need for schools and teachers to be responsive to community interests. Again, it is worth citing the brief's language at length to give a fuller sense of the union's stance. [Fig. 3](#) reproduces its opening paragraphs.

The balance of the brief makes good on the statement that "the issue is not whether English should be taught but how the mother tongue can be retained." The brief identified six profiles of linguistically- and culturally-diverse students in Toronto schools; it argued that while English-as-Second-Language programs might support the needs of *some* of these profile types, HL classes support

Toronto teachers' federation

222 Niagara Street

Toronto, Ontario

M6J 2L3

(416) 366-1681

TTF RESPONSE TO "THE CHALLENGE OF LANGUAGES"

The Toronto Teachers' Federation is committed to a decentralized, community-based process of education. The realities of our city, therefore, require that we as educators respect the concerns of the majority of the children in Toronto, whose first language is other than English.

Whereas it should remain our main objective to provide children with the self-esteem and basic skills to provide equality of educational opportunity, we must begin from the cultural and linguistic heritage of the children we teach. It is often argued that children who develop a greater understanding and pride in their ethnic background are better able to deal with the challenge of learning English.

It is for this reason that we can agree with the discussion paper premise that "the issue is not whether English should be taught but how the mother tongue can be retained."

Fig. 3 Opening paragraphs from the TTF brief to the Work Group on Third Language Instruction, 1982. [Toronto Teachers' Federation \(1981, p. 1\)](#).

all of them and thus should be expanded in Toronto schools. The brief also addressed questions of feasibility, given that so many languages were spoken across Toronto. It viewed HL classes offered after school or on Saturdays as the most practical option, even calling for the creation of itinerant HL positions to support HL classes in schools with smaller enrollments. However, the brief recognized that "many parents of children enrolled in after-school heritage programs are dissatisfied by the apparent lack of intensity" ([Toronto Teachers' Federation, 1981, p. 2](#)). As such, the brief concluded that the "real issue" (p. 2) facing the work group is how "to increase the amount of time children spend in third language instruction" (p. 3).

As its first recommendation, the brief argued that "where a significant majority of children come from one ethnic group," HL classes should be integrated into the school day. The brief understood that both school staff and the local community would have to concur on such a decision. The brief reiterated the union's call for HL instructors to be fully certified teachers "where possible" (p. 3) and paid a teacher's contractual salary. With respect to debates about setting up "heritage language schools" and other kinds of magnet programs, the brief sided squarely against such proposals.

Similar to the first work group, the final recommendations of the Work Group on Third Language Instruction were consistent with the spirit of the TTF brief. The work group rejected "heritage language schools" and other magnet programs. Instead, it focused primarily on how to increase the number of Toronto schools with daytime HL classes, and it laid out a framework for how schools could make a collective decision about whether to integrate HL classes into the school day. Furthermore, the final report outlined a plan to pilot bi- and trilingual programs, going so far as to call on the Director of Education to work with the Ministry of Education to amend the *Education Act* to allow for non-English and non-French languages to be used as medium of instruction. Finally, all 13 members of the work group signed the final report before it was distributed. Only one member, Antonio Vaz of the Heritage Languages Advisory Committee, signed "with reservations" ([Toronto Board of Education, 1982, p. v](#)). For her part, TTF representative Imogene Walker signed without indicating any reservations, suggesting that the union had accepted the work group's final recommendations.

Backlash

In the year between the work group's final report in March 1982 and the Board's implementation of its main recommendations in March 1983, the TTF's actions took a sharp turn, contradicting the very principles and recommendations it had just endorsed. The trigger for this turn was the Board's passage on March 10, 1983 of an "Operations Manual" to guide the process by which elementary schools could decide whether to initiate daytime HL classes. Two immediate issues came to the fore. First, instituting daytime HL classes would require extending the school day by 30 min. This meant teachers in schools with daytime HL classes would work 30 min longer with no additional compensation. Beyond the pay issue, it remained unclear which responsibilities, if any, teachers would have for those students *not* participating in HL classes during the extra 30 min. Second, this new procedure for deciding whether to initiate HL classes gave parents the ultimate say. If enough parents petitioned a school to consider daytime HL classes, the school would be required to hold a public meeting to discuss the program and how it works. A letter would then be sent home to all parents with similar information. Then, parents would vote; a simple majority in favor would trigger implementation of daytime HL classes. As Board employees, teachers were expected not to intervene, whether for or against, during this public consultation process. The Operations Manual thus construed this decision as one between the Board—represented by local school leadership—and parents in the respective community.

For the TTF, these were two steps too far; on April 11 it filed a grievance that would take three years to resolve. For the union it was unacceptable to have some of its members working longer hours than others without additional pay. But the larger issue related to the question of who gets to decide what happens in Toronto schools. The union argued that it had the right to negotiate the length of the school day and compensation with the Board directly; the Board took the position that it alone had the right, indeed the obligation, to implement Board-approved curriculum, such as daytime HL classes. In May, the grievance moved to mediation, and later to arbitration, not to be resolved for another three years. **Table 2** presents a timeline of the TTF's resistance to daytime HL classes from the filing of this grievance in April 1983 to the arbitrators' ruling in June 1986.

From the union's perspective, this was an urgent issue to resolve. As indicated in **Table 2**, by the end of the 1984–85 school year the potential existed for over one-half of Toronto's 110 elementary schools to be operating on an extended timetable with daytime HL classes. The union might have seen the writing on the wall and, for example, worked to ensure that all schools adopted daytime HL classes to remove the imbalance in work day and pay. Or, it might have supported efforts to amend the *Education Act* to allow HLs to be part of the formal elementary-school curriculum, thus removing the need to extend the school day by 30 min that created the disparity in the first place.

Instead—and under a new leadership with Jack McFadyen as president—the TTF threw all its weight into blocking daytime HL classes, contravening the very principles it had just articulated and agreed to through its participation in the Work Group on Third Language Instruction. As the vignette that opens this article suggests, McFadyen repeatedly claimed that the teachers' fight was against daytime HL classes only, not HL education overall. However, the union's words and deeds from 1983 to 86 indicate just the opposite. Closer examination of these claims reveals that the union was in fact challenging the legitimacy of HL education overall.

Central to the TTF's efforts to block daytime HL classes was questioning the legitimacy of "outsiders" to have a say in how local schools were run. Key examples of this stance appeared in a study conducted by the TTF's provincial "parent" union, the Ontario Public School Teachers' Federation (OPSTF), of the conflicts in Toronto over daytime HL classes. The OPSTF's final report from December 1983 contains multiple passages that suggested the TTF now rejected the "decentralized, community-based process of education" it had endorsed in March 1982. For example, the introduction to the report stated: "In the City of Toronto, the Board's School-Community Relations Department was becoming increasingly active in the ethnic communities. Increased interpretive

Table 2 Timeline of events in TTF's opposition to daytime Heritage Language classes.

June 1981–March 1982	• Work group on third language instruction • Among its many recommendations, it calls for expanded daytime HL classes • Lays out principles for how schools can decide on starting daytime HL classes
March–April 1983	• Board issues Operations manual to guide the decision-making process on whether a school should initiate daytime HL classes
Fall 1983	• TTF files a grievance against this procedural policy • Of Toronto's 110 elementary schools: • 12 already have daytime HL classes (in place since at least 1977) • 33 to use the new procedure to initiate the decision-making process • 13 announce they will begin that process in fall 1984 • OPSTF, the TTF's provincial parent union, conducts its own study of daytime HL classes
Jan.–Nov. 1984	• Final report issued in December • TTF begins its work-to-rule job action in January • In May, TTF votes to end the job action by September 1 • In November, TTF is in a legal position to strike and threatens to do so, but never did
June 1986	• TTF grievance from 1983 is settled by arbitration • Arbitrators side on every count with the board and against the TTF

services, multilingual communications, and the establishment of the powerful Heritage Languages Advisory Committee were significant developments" (Report of the O.P.S.T.F. Task Force, p. 2). Later in its overview of implementing daytime HL classes, the report noted, "In Toronto, at the prodding of the Heritage Languages Advisory Committee, the Director of Education 'designated' several additional schools" to implement the extended-day model (p. 11). Finally, in its analysis of three Toronto schools and how they deliberated over daytime HL classes, the report stated that voting over whether to implement daytime HL programs "has been influenced by lobbying" (p. 19). In this way, the OPSTF report suggested that the very structures the Board had set up to support linguistically- and culturally-inclusive programming had become too "powerful" and had overstepped their bounds by "prodding" and "lobbying" Board leadership and/or parents regarding daytime HL classes on local schools.

If the OPSTF report had been subtler, charges of outside meddling would become more direct as the TTF's work-to-rule job action began in January 1984. In reference to Alberto DiGiovanni, head of the Board's Heritage Language Advisory Committee, a *Toronto Sun* headline noted, "Director under fire as outsider" (McLeod, 1984, p. 3). Without attribution to any individual, the article stated "angry parents, trustees and teachers in the public school system want di Giovanni [sic] 'sent back to the [Catholic] school board where he belongs and can't create any more hassles for us'" (p. 3). The article noted that DiGiovanni lived in a suburb of Toronto, not in the city itself, while his wife served as a top administrator for Toronto's Catholic school system. Further, the article repeated a Board Trustee's claim that DiGiovanni might be reaping personal financial gain, insofar as Italian-language teaching resources he helped develop were being used in HL classes. This trifecta of insinuations—a foreigner both to Canada and to Toronto, Catholic, corrupt—reveal the anti-immigrant, indeed racialized, sentiment present in these claims that the TTF and others made about "outsiders."

The TTF also fought hard to limit the amount of control parents had over the decision-making process regarding daytime classes. In a letter of understanding between the TTF and Board issued shortly after the union's work-to-rule job action had begun, the two parties agreed to a number of compromises. One of these dealt with the voting process for deciding whether to initiate daytime HL classes. No longer was a simple majority enough to trigger these classes. Now, an "enriched majority" of parents must vote in favor of the program, and at least 50% of the school's student body must actually enroll in daytime HL classes for the program to go forward (Toronto Board of Education, 1984, p. 2). Exactly which percentage constituted an enriched majority was not stipulated. However, this interim agreement underscored the TTF's stance that parents cannot be trusted with decision-making authority over local schools. In the end, the concession failed on two counts. The TTF was still not satisfied with the decision-making process and continued with its work-to-rule job action. For their part, parents were so outraged by this enriched majority compromise that 16 members of the Heritage Languages Advisory Committee resigned the next day in protest (Brown, 1984b).

Furthermore, as the conflict over daytime HL classes progressed in 1982–84, one original TTF demand quietly fell away. No longer was the union insisting that HL instructors be fully certified teachers. Instead, the TTF demanded that its members have nothing to do with HL instruction in any way. The same letter of understanding referenced above reflected another area of interim agreement between the Board and the TTF, namely that classroom teachers "would not be required to have any involvement" with HL classes, would not be assigned any additional duties, and would only be required "to fulfill their professional commitment for remedial programming as may be appropriate" (Toronto Board of Education, 1984, p. 1). In this way, the letter of understanding between the Board and the union formalized the division between classroom teachers and HL instructors that continues to this day, positioning the latter as outsiders in the school community. And yet neither this formalized division nor the new definitions of "majority" from this letter of understanding were enough for the TTF to end its job action against daytime HL classes.

In a further move to challenge daytime HL classes, the union repeatedly argued that they posed a threat to other kinds of extra-curricular school programming, in particular after-school sports clubs and remedial help teachers offered to students. However, *how* the union made this argument suggested it was, in fact, questioning HL education overall. For example, in the OPSTF report introduced earlier, it raised the question of how schools support the "whole child." The report cited interviews with teachers who sponsored after-school sports and letters from provincial sports associations to stress that daytime HL classes crowded out space for these and other after-school programs. The report argued: "For Toronto schools which have prided themselves on addressing the needs of the whole child through many extra activities for students, as well as elaborate professional development activities for teachers, this extension [of the school day] erodes such beneficial activities" (p. 12). Later, it listed more specific implications: "This reduction of extended activities that address the need of the whole child may have serious effects on individual growth, team building, recreational opportunities, and school spirit" (p. 16). Such arguments were only possible by excluding HLs as another key part of supporting the "whole child," let alone building meaningful community and school spirit. In this way, the report subtly challenged to the very idea of HL education, not just the form it took in an extended school day.

Given these efforts—some directly aimed at ending daytime HL classes, others more indirectly challenging the legitimacy of HL education in Toronto schools—it is little wonder that the TTF and its members were accused of racism. Shortly after the TTF initiated its work-to-rule job action, parents denounced the union's actions. As one news report noted, "At a news conference yesterday, parents representing eight cultural communities—black [sic], Chinese, Greek, Italian, Portuguese, Asian [sic], Spanish, and French—called a two-week old boycott of after school programs 'racial discrimination' by teachers aimed at stopping heritage language programs" (Brown, 1984a, p. A6). At the news conference, Eduardo Sousa, the Portuguese representative on the Heritage Language Advisory Committee, said it plainly: "We believe the real reason behind the arguments used by the teachers is racism. They are discriminating against the teaching of any language and culture other than English" (p. A6). Sousa's comment provided further indication that many HL advocates interpreted the teachers' actions as rejecting the legitimacy of HL education overall, not merely the daytime model.

The union responded in two primary ways to deflect these charges of racism. First, it argued that the fact that teachers work with linguistically- and culturally-diverse children at all means they cannot be racist. Speaking to a reporter after a meeting of over 100 TTF representatives on February 17, 1984, TTF president McFadyen stated, “Teachers were outraged that anyone should accuse them of racism ... Teachers who work with children in immigrant communities are especially upset that their integrity and professional dedication have been challenged” (*Globe and Mail*, 1984, p. 5). Second, the TTF insisted that its concerns were contractual ones over working conditions, and not about the substance of daytime HL classes. The union used this “labour-dispute” logic as their primary strategy to deflect charges of racism. In its statement regarding the letter of understanding between the union and the Board in early mediation over the grievance it had filed, the TTF even went so far to demand that the Board “issue a joint public statement to the effect that, in the current dispute between the teachers and the Board, the Board is satisfied that the teachers’ actions have been motivated by concerns over working conditions and delivery of programs and not, as has been alleged, by racism or bigotry” (*Toronto Teachers’ Federation*, 1984, p. 2). What is revealing here is that the TTF looked to *the Board* to have its name cleared, and not to *parents* and other HL advocates. Indeed, a handwritten note on the Board’s copy of this TTF statement asked with incredulity, “TTF has issued writs? With respect to racism and bigotry?” (p. 2). In this way, the union insisted on separating notions of race and class and asserted their concerns were merely contractual grievances so as to deflect charges that their fight against daytime HL classes was racist.

At this stage of my research, I am not able to document what happened inside the TTF that led it to reverse the positions on bilingual education, HL classes, and decentralized schools rooted in community input that the union had developed as part of a decade of sustained public deliberation over the relationship between language, race and culture in Toronto schools. What is clear, however, is that once actual HL programs existed and once real mechanisms for greater parental input were in place, the union took a sharp turn to the right. In practice, the union questioned not only the legitimacy of linguistically- and culturally-inclusive programs such as HL classes, but also the very right of parents to have any meaningful say in what happens in their children’s schools. Moreover, at every turn, the union tried to mask the depth of its grievances by portraying them as limited to just a single model of HL education and focused on discrete issues of contract negotiations.

In the end, few were persuaded by these claims. Perhaps the sharpest assessment of the TTF’s backlash from 1983 to 1986 came from a progressive community newspaper, *Mudpie*, and one of its regular columnists, James McQueen. In April 1984, McQueen wrote:

And isn’t it an interesting list of new working class heroes—Bette Stephenson, Jack McFadyen, Nola Crewe, David Moll, and various elementary school principals. And isn’t the rhetoric equally inter-esting—“quality education”, “solidarity with teachers”, “the need for consultation”, “concern for the needs of the immigrant student”, the need for research. Gimme a break. McFadyen has accomplished exactly what he prom-ised to do when he ran for office. He’s wrestled control of the Toronto Teachers’ Federation from the socialist hoards. He’s driven a wedge between parents, the boards and teachers. He’s taken T.T.F. to bed with the right-wing educators in Toronto and he’s accomplished it all in the name of union solidarity and concern for the working conditions of teachers.

McQueen (1984, p. 22).

While McQueen hangs this right-wing backlash around McFadyen’s neck, it is important to understand the support he had among his union siblings. Recall that he was re-elected in spring 1984—that is, in the midst of this battle against daytime HL classes—by an overwhelming margin of 1368 votes to 147 for his rival. These votes in his favor accounted for 55% of the union’s entire membership of roughly 2500 Toronto elementary teachers.

Conclusion

On June 17, 1986, arbitration of the TTF’s grievance against how the Board implemented daytime HL classes was finally resolved. On every count, the arbitrators ruled against the TTF and for the Toronto Board of Education. By then, however, this outcome mattered little. Municipal elections in 1985 had ousted the “socialist hoards” McQueen referred to in his 1984 column—that is, Board Trustees aligned with the social-democratic New Democratic Party, were replaced with a decisively conservative school board. This new group set out to reverse many of the most progressive initiatives the Board had implemented after its first Work Group on Multicultural Programs in the mid 1970s (see *McCaskell*, 2005). Moreover, the sustained parent activism that had pushed the Board so far in its deliberations over linguistically- and culturally-inclusive programming was already beginning to wane (see *Dehli*, 1994). Finally, as mentioned early, a minority Liberal provincial government was elected in 1987; its reforms of the Heritage Languages Program, the policy behind the daytime conflict analyzed in this paper, solidified the structure of HL education in Ontario that we continue to live with to this day.

The decade of deliberation and conflict over HL programs and other kinds of linguistically- and culturally-inclusive programs represents one of the most important moments of *discontinuity* in the history of Ontario public education. As argued above, there had long existed a hierarchy of minoritized and racialized languages in Canada. During the 1960s, pressure from Quebec nationalism and the Red Power movement forced a break in the continuity of that hierarchy (see *Haque and Patrick*, 2015). Even if the federal government aimed to restore continuity by elevating English and French as official languages, the contradictions of multiculturalism within a bilingual framework in fact opened up significant political space to experiment with different kinds of

linguistically- and culturally-inclusive education. HL policies and the programs they funded were just one example of these options; but it is important to stress that they emerged in an era when so much seemed possible in terms of reinventing schools in ways that centered the linguistic and racial lives of the students who attended them. As this paper has argued, when this reinvention got too close to the Toronto Teachers' Federation, when this reinvention began to make demands on teachers themselves to change, the union reversed course on its own stances it had developed over 7 years, closed ranks, and threw all its weight behind ensuring schools changed as little as possible. It did so by counterposing the language of contract negotiations and working conditions to community-driven demands challenging racism and other kinds of discrimination in Toronto schools. This meant seeing parents as illegitimate outsiders, not as allies; indeed, as Clem Marshall argued in his campaign speech in 1984: "Parents are no longer behind us. We have lost them as allies with our behaviour" (Brown, 1984, p. A7). In the end, the union's backlash was an important part of bringing this moment of discontinuity to an end by the mid 1980s, leaving us with school structures that keep HLs at the margins of school life. These structures have changed little, if at all, since then.

This union scuffle may seem distant and focused on a school program that garners little attention in contemporary discourse or scholarship about education. Nevertheless, the lessons from this conflict are vital for us to study and reckon with. We are currently living through another moment of discontinuity as it relates to social and racial justice in Toronto schools. In July 2020, for example, the Ontario government proposed several initiatives to address systemic racism in the province's publicly funded schools (Office of the Premier, 2020). These initiatives, described by the government as intended to "help break down barriers for Black, Indigenous and racialized students" (p. 1), include banning suspensions in the early grades, ending the practice of streaming in Grade 9, and toughening sanctions against teachers who use racist or other discriminatory language. In announcing these measures, the education minister stressed, "To racialized students in Ontario: we see you" (p. 2). It would be hard for him not to: his announcement followed immediately after another explosion of activism against anti-Black racism in Toronto (Armstrong and Lorinc, 2020). As well, persistent activism by Black and Muslim parents in Peel District School Board, the province's second-largest board, has exposed the extent of racism within its schools, leading the minister to impose a change in board leadership last June (Rushowy, 2020).

Moreover, in June 2020, the Elementary Teachers of Toronto, the successor to the TTF, elected Jennifer Brown as president, the first Black woman ever to head the union. Her election is as important as the long-standing activism of rank-and-file ETT members to engage the union in social-justice, anti-racist and other kinds of equity work. Nevertheless, Brown has been a target of racist backlash for her activism and for speaking out against the discrimination that she and other Black teachers in Toronto schools continue to experience. As Brown said to the *Toronto Star*, "Some of the racism has come from my very own colleagues and fellow teachers. And I know that I'm not the only one facing racism, I know that many teachers and union members share the experience with bullying and discrimination because that's what I faced" (Gallardo, 2021, n.p.). The lessons from this conflict over HL education from some 40 years ago suggest that the fight against the racism that Brown faces both within the union and within Toronto schools is best waged by understanding how race and class collaborate in organizing school life, and thus must collaborate in informing social-justice teacher unionism. Moreover, this history makes clear that this fight against racism and other kinds of discrimination is not won through policy briefs and resolutions; those can be undone from one day to the next. Instead, this fight is based on sustained activism, both within the union, but—perhaps more important—alongside parents and students, not just as allies but rather as in solidarity based on the shared interests each group has in challenging racism and expanding social justice in Toronto schools.

References

Primary sources

- Anderson, M., January 9, 1975. Bilingual Education: An Interim Position Paper of the TTF Executive. Toronto District School Board Archive, Box 266/C5-1, Briefs (As of April 23, 1975)—Elementary—Secondary—Central.
- Brown, L., February 15, 1984a. Teachers' boycott based on racism, groups say. *Toronto Sun* A6.
- Brown, L., February 22, 1984b. Ethnic parents resign in heritage language row. *Toronto Star* A7.
- Brown, L., May 12, 1984c. Teachers back tough-talking leader. *Toronto Star* A1–A7.
- Chown, P.B., October 14, 1975. Letter to Dan Leckie, Trustee Ward 6. Toronto District School Board Archive, Box 270/C5-2, Briefs Received in Response to the Draft Report 1975, Letters to Be Answered.
- Davis, B., 1984. Toronto's heritage language dilemma. *Mudpie* 5 (7), 2.
- Globe & Mail, February 18, 1984. Teachers deny racism change: heritage language program. *The Globe and Mail* 5.
- Kerr, E., September 16, 1974. Letter to Chris Wooding, President of the Toronto Teachers' Federation. Toronto District School Board Archive, Box 273/C5-3, 2. Correspondence d. Work Group, 7. Correspondence — General, Others.
- McLeod, J., 1984. Director under fire as outsider. *Toronto Sun* Feb 13, 3.
- McQueen, J., April 1984. Toronto teachers led up the garden path? *Mudpie* 5 (4), 22.
- Miller, D., April 25, 1982. Voices raised in anger against foreign tongues. *Toronto Star* G6.
- Report of the O.P.S.T.F. Task Force on Heritage Language Programs, December 1983. Ontario Teachers' Federation Records, 8384.07.
- Spears, J., April 21, 1982. Language battle ready to explode, parents warned. *Toronto Star* A6.
- Toronto Board of Education, 1982. Towards a Comprehensive Language Policy: The Final Report of the Work Group on Third Language Instruction. Toronto Board of Education, Toronto.
- Toronto Board of Education, February 21, 1984. Letter of Understanding Committee on "Provision of Heritage Language Programs Within the Integrated/Extended Day". Archives of Ontario, B700347, RG 2-215 Ministry of Education Legislation and Legal Services Operational Files, Heritage Language Program 1978–1984 Ss9 (2).
- Toronto Teachers' Federation, 1981. TTF Response to the "Challenge of Languages". Toronto District School Board Archive. TDSB2003-0537.
- Toronto Teachers Federation, February 21, 1984. Letter of Understanding. Archives of Ontario, B700347, RG 2-215 Ministry of Education Legislation and Legal Services Operational Files, Heritage Language Program 1978–1984 Ss9 (2).
- Wooding, C., October 4, 1974. Letter to Ed Kerr of the Work Group on Multicultural Programs. Toronto District School Board Archive, Box 273/C5-3, 2. Correspondence d. Work Group, 7. Correspondence—General, Others.

Secondary sources

- Anyon, J., 2005. What "counts" as educational policy? Notes towards a new paradigm. *Harv. Educ. Rev.* 75, 65–88. <https://doi.org/10.17763/haer.75.1.g1q5k721220ku176>.
- Armstrong, L., Loric, J., June 19, 2020. Black Lives Matter protest issues 27 demands for reform in major anti-police protest. *Toronto Star*. Retrieved from. <https://www.thestar.com/news/gta/2020/06/19/we-demand-better-we-demand-our-lives-black-lives-matter-toronto-protesters-call-for-defunding-of-police.html>.
- Bale, J., 2015. Language policy and global political economy. In: Ricento, T. (Ed.), *Language Policy and Political Economy*. Oxford University Press, Oxford, pp. 72–96.
- Bale, J., 2016. In defense of language rights: re-thinking the rights-orientation from a political economy perspective. *Biling. Res. J.* 39, 231–237. <https://doi.org/10.1080/15235882.2016.1224208>.
- Bale, J., 2019. Racialized resistance to the heritage languages program in Ontario. In: Ricento, T. (Ed.), *Language Politics and Policies: Perspectives From the United States and Canada*. Cambridge University Press, Cambridge, pp. 215–233.
- Bale, J., Kawaguchi, M., 2020. Heritage-language education policies, anti-racist activism and discontinuity in 1970s and 1980s Toronto. *Crit. Inq. Lang. Stud.* 17, 5–25. <https://doi.org/10.1080/10100309230802486150>.
- Berryman, J., 1986. Implementation of Ontario's Heritage Languages Program as a Case Study of the Extended Day Model [Unpublished Dissertation]. University of Toronto, Toronto.
- Block, D., 2014. *Social Class in Applied Linguistics*. Routledge, New York.
- Bramble, M., 2000. Black education in Canada: past, present and future. In: Goldstein, T., Selby, D. (Eds.), *Weaving Connections: Education for Peace, Social and Environmental Justice*. Sumach Press, Toronto, pp. 99–119.
- Cummins, J., 1983. *Heritage Language Education: A Literature Review*. Ministry of Education, Ontario, Toronto.
- Cummins, J., Danesi, M., 1990. *Heritage Languages: The Development and Denial of Canada's Linguistic Resources*. Our Schools/Our Selves Education Foundation and Garamond Press, Toronto.
- Delhi, K., 1994. *Parent Activism and School Reform in Toronto: A Report*. Ontario Institute for Studies in Education, Toronto.
- Fendler, L., 2008. The upside of presentism. *Paedagog. Hist.* 44, 677–690. <https://doi.org/10.1080/00309230802486150>.
- Feuerverger, G., 1997. On the edges of the map: a study of heritage language teachers in Toronto. *Teach. Teach. Educ.* 13, 39–53.
- Flores, N., 2016. A tale of two visions: hegemonic Whiteness and bilingual education. *Educ. Pol.* 30, 13–38.
- Flores, N., 2019. Producing national and neoliberal subjects: bilingual education and governmentality in the United States. In: Rojo, L.M., del Percio, A. (Eds.), *Language and Neoliberal Governmentality*. Routledge, New York, pp. 49–68.
- Flores, N., Chaparro, S., 2018. What counts as language education policy? Developing a materialist anti-racist approach to language activism. *Lang. Pol.* 17, 365–384. <https://doi.org/10.1007/s10993-017-9433-7>.
- Flores, N., McCauliffe, L., 2020. "In other schools you can plan it that way": a raciolinguistic perspective on dual language education. *Int. J. Biling. Educ. Biling.* <https://doi.org/10.1080/13670050.2020.1760200>.
- Flores, N., Rosa, J., 2015. Undoing appropriateness: raciolinguistic ideologies and language diversity in education. *Harv. Educ. Rev.* 85, 149–171. <https://doi.org/10.17763/0017-8055.85.2.149>.
- Gale, T., 2001. Critical policy sociology: historiography, archaeology, and genealogy as methods of policy analysis. *J. Educ. Pol.* 16, 379–393. <https://doi.org/10.1080/02680930110071002>.
- Gallardo, C., May 25, 2021. This teachers' union president has endured years of anti-Black racism in TDSB; she has a plan for change. *Toronto Star*. Retrieved from. <https://www.thestar.com/news/gta/2021/05/25/this-teachers-union-president-has-endured-years-of-anti-black-racism-in-the-tdsb-she-has-a-plan-for-change.html>.
- Garland, D., 2014. What is a "history of the present"? On Foucault's genealogies and their critical preconditions. *Punishm. Soc.* 16, 365–384. <https://doi.org/10.1177/1462474514541711>.
- Gidney, R.D., 1999. *From Hope to Harris: The Reshaping of Ontario's Schools*. University of Toronto Press, Toronto.
- Haque, E., 2012. *Multiculturalism Within a Bilingual Framework: Language, Race, and Belonging in Canada*. University of Toronto Press, Toronto.
- Haque, E., 2019. A Foucauldian approach to language policy in Canada. In: Ricento, T. (Ed.), *Language Politics and Policies: Perspectives from the United States and Canada*. Cambridge University Press, Cambridge, pp. 232–243.
- Haque, E., Patrick, D., 2015. Indigenous languages and the racial hierarchisation of language policy in Canada. *J. Multiling. Multicult. Dev.* 36, 27–41. <https://doi.org/10.1080/01434632.2014.892499>.
- Heller, M., Duchêne, A., 2012. Pride and profit: changing discourses of language, capital and nation-state. In: Duchêne, A., Heller, M. (Eds.), *Language in Late Capitalism: Pride and Profit*. Routledge, New York, pp. 1–21.
- Holborow, M., 2015. *Language and Neoliberalism*. Routledge, London, UK.
- Jean-Pierre, J., Nunes, F., 2015. From integration to empowerment: multicultural education in the board of education in the city of Toronto, 1960–1975. In: Guo, S., Wong, L. (Eds.), *Revisiting Multiculturalism in Canada: Theories, Policies and Debates*. Sense Publishers, Rotterdam, pp. 271–288.
- Kim, H., Burton, J., Ahmed, T., Bale, J., 2020. Linguistic hierarchisation in education policy development: Ontario's Heritage Languages Program. *J. Multiling. Multicult. Dev.* 41, 320–332. <https://doi.org/10.1080/01434632.2019.1618318>.
- Masemann, V., 1984. Multicultural programs in Toronto schools. In: Mallea, J.R., Young, J.C. (Eds.), *Cultural diversity and Canadian Education: Issues and Innovations*. McGill–Queens University Press, Montreal and Kingston, pp. 349–369.
- Matas, R., April 23, 1982. Racist callers plague trustees over proposal. *The Globe and Mail* 4.
- Maychak, M., April 28, 1982. Heritage class foes accused of race bias. *Toronto Star* A6.
- McCaskell, T., 2005. *Race to Equity: Disrupting Educational Inequality*. Between the Lines Press, Toronto.
- Office of the Premier, July 9, 2020. Ontario Taking Bold Action to Address Racism and Inequality in Schools [Press Release]. Retrieved from. <https://news.ontario.ca/opo/en/2020/7/ontario-taking-bold-action-to-address-racism-and-inequity-in-schools-1.html>.
- Rosa, J., Flores, N., 2017. Unsettling race and language: toward a raciolinguistic perspective. *Lang. Soc.* 46, 621–647. <https://doi.org/10.1017/S0047404517000562>.
- Rushowy, K., June 23, 2020. Peel education director fired after criticism over board's handling of racism, dysfunction. *Toronto Star*. Retrieved from. <https://www.thestar.com/politics/provincial/2020/06/23/peel-education-director-fired-after-criticism-over-boards-handling-of-racism-dysfunction.html>.
- Shohamy, E., 2006. *Language Policy: Hidden Agendas and New Approaches*. Routledge, London.
- Tamboukou, M., 1999. Writing genealogies: an exploration of Foucault's strategies for doing research. *Discourse* 20, 201–217. <https://doi.org/10.1080/0159630990200202>.
- Tavares, A.J., 2000. From heritage to International languages: globalism and Western Canadian trends in Heritage language education. *Can. Ethnic Stud.* 32, 156–171.
- Tollefson, J., 1991. *Planning Language, Planning Inequalities*. Longman, New York.
- Vipond, R.C., 2017. *Making of a Global City: How One Toronto School Embraced Diversity*. University of Toronto Press, Toronto.
- Wong, L., Guo, S., 2015. Revisiting multiculturalism in Canada: an introduction. In: Guo, S., Wong, L. (Eds.), *Revisiting Multiculturalism in Canada: Theories, Policies and Debates*. Sense Publishers, Rotterdam, pp. 1–15.
- Wright, O.M., 1997. *International Languages in the Elementary School: The Heritage Languages Program—An Experiment in Cultural Accommodation*. Toronto Board of Education, Toronto.

Social movements and college student activism

Jerusha O. Conner, Villanova University, Villanova, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Distinguishing terms	398
An abbreviated historical overview of student activism	399
Fighting cuts to higher education in Europe	401
Divestment at University of Cape Town	401
#NotAgainSU	401
Tactics, issues, and campaigns of US college students	402
Effects on students	403
Effects on institutions	404
Conclusion	405
References	405

In popular imagination, the term “student activism” conjures up images of student protests and demonstrations, like the student-led protests in Tiananmen Square, China in 1989 (Brown, 2021), sit-ins like the iconic Greensboro, North Carolina lunch counter sit-in of 1960 in the US (Schmidt, 2018), or occupations of administrative buildings, as when students barricaded themselves in schools during the so-called “march of the Penguins” in Chile in 2006 (Cummings, 2015); however, direct action is but one form of student activism. Student activism can also involve closed-door meetings with decision makers and efforts to educate others or raise awareness through such activities as tabling, issuing reports, or creating online content. Some scholars have argued for a broad definition of student activism that encompasses service and volunteering work (Braxton, 2016) or unapologetically showing up in educational spaces designed and intended for others (Stewart and Williams, 2019); however, not all students who participate in service-learning or civic engagement activities, and not all students who hold minoritized identities in their schools or universities would consider themselves to be activists (Linder et al., 2019). Even students who attend a protest may not feel comfortable applying the label of activist to themselves. For these reasons, it is important when studying student activism to center the identity claims and understandings that the students themselves embrace. This entry starts, therefore, by considering how activists themselves define the terms frequently used to describe them and their work. It then offers an abbreviated overview of the history of college student activism, followed by snapshots of three recent campaigns. It then zeros in on the US context and offers a close look at the issues and tactics mobilizing contemporary college student activists in the US. The entry concludes with a brief review of the literature on the effects of college student activism, for both the students and their institutions, drawn primarily from US-based scholarship.

Distinguishing terms

According to a recent sample of self-identifying college student activists in the US, *activism* entails “taking action to effect change in an unjust status quo” (Conner, 2020). This definition operates as an umbrella that can cover a wide range of actions and ideologies about justice, including conservative ideologies, but it underscores activism as active engagement in change efforts.

Some self-identifying student activists prefer the term *organizer* to *activist*. These students distinguish *organizing* as a form of activism that is sustained rather than sporadic or spontaneous; rooted in relationships rather than in a particular issue; and determined to build the power of the people who are organizing together to address the issues that matter to them. On campus, organizers may work to build the power of a particular identity-based group, such as transgender students or students with disabilities (Warnock and Hurst, 2016), or they may organize a group with a shared position, such as graduate student Teaching Assistants or non-unionized custodial staff. Organizing campaigns may focus on a particular demand or set of demands, but they do not end after a campaign is won or lost, because the long-term goal is sustained power and a seat at the decision-making table.

In contrast to organizing groups in which the issues of concern bubble up organically from the community members, *social movement organizations* (SMOs) emerge in response to an issue and exist to do the work necessary to sustain the broader social movement focused on that issue. The Hong Kong pro-democracy student organization, Scholarism, offers a key example. Created in 2011 by students opposed to a proposed Moral and National Education curriculum, Scholarism went on to play a critical role in sparking the Umbrella Revolution movement of 2014, organizing a week-long boycott of classes by university and secondary school students in the lead up to sustained, large-scale protests for democratic reform in central Hong Kong (Macfarlane, 2017). While some student-led SMOs exist independently from higher education institutions, others have established a presence on college campuses. For example, the Sunrise Movement, a prominent climate justice SMO, has hubs on many US campuses. The National Rifle Association (NRA), similarly, has established NRA Collegiate Coalition chapters on campuses across the US to engage students in advocacy for gun rights.

Of course, students do not need SMOs to set up shop on their campuses in order to translate social movements into calls for institutional change. In some cases, existing clubs or organizations, like the Black Student Alliance or an LGBTQ+ club, may take up the issues resonating in broader society, connect them directly to student experiences of oppression or injustice on campus, and put forward demands for change. In other cases, students reacting to an incident on campus or a global current event may find one another serendipitously, begin organizing, and only then decide to incorporate as a formal club or give themselves a group name (Fisher, 2018). NotAgainSU, discussed later in this entry, offers a prime example of this phenomenon. It is also important to note that while many social movements emerge off campus and quickly make their way onto campus, some social movements are born on campus. One such example is the Fees Must Fall movement, which arose from the Rhodes Must Fall movement that started at the University of Capetown in South Africa in 2015 and spread to other campuses around the world (Luescher and Klemenčič, 2017). Finally, some student movements may occur entirely off campus and focus on issues that are irrelevant or tangential to education policy or practice. Broadly understood as a collective effort “by a large number of students to bring about or prevent change” (Gill and de Fronzo, 2009, p. 209), these movements may focus on an issue that concerns youth. They become student movements, rather than youth movements, when the participants strategically decide to mobilize their identity as students for political or moral leverage. For instance, during the Road Safety Movement in Bangladesh, students wore their school uniforms and backpacks to advance a nonpartisan identity as responsible, hard-working students who deserved to be protected from reckless drivers by greater safety measures (Mirza et al., 2022).

In short, student activism, student organizing, and student engagement in social movement work are not synonymous. Activism may entail organizing and/or participation in social movements, but it also encompasses a broader array of actions. According to self-identifying student activists, some of these actions may be “small, daily acts,” like refusing to consume animals’ bodies or being an “upstander,” who calls out microaggressions, and some may be unconventional forms of activism, like participating in student government, facilitating a dialog, or creating digital art to raise awareness (Conner, 2020). While some student activists understand activism as “consuming all daily life,” others recognize “existence as a form of resistance,” and still others see a wide spectrum in between these two poles (Logan et al., 2017).

A final point of distinction worth noting is between student activism and youth activism. As a subset of youth activism, student activism refers to activism in which students’ identities or concerns *as students* are salient or when students use their campuses as mobilizing bases or stages for their activism. Classmates who meet up at an off-campus protest, unrelated to their institution, who hold signs and repeat chants that do not invoke their identities as students, would be better described as youth activists than student activists. Youth activism is the broader category because it includes young people who are not enrolled in institutions of higher education as well as those who are. Of course, the moniker student may refer to those studying in primary and secondary schools in addition to those attending post-secondary institutions. While it is certainly true that secondary school students have played key roles in social movements throughout history [from the Child’s March during the US Civil Rights Movement (Hunter-Gault, 2013) to the worldwide school climate strikes of 2019], this entry focuses on the activism demonstrated by college and university students.

An abbreviated historical overview of student activism

The history of student resistance stretches back to the advent of the medieval university. Boren (2001) traces the roots of student activism to the early thirteenth century when European students began forming collectives. For example, in 1217 a group of university students in Bologna, Italy organized a strike to protest their teachers’ tendencies to arrive late to lectures, if they came at all. Most of the earliest examples of student agitation were directed against the towns in which the universities were situated, where locals tried to limit student activity; however, students have long engaged in power struggles with administrators and faculty over university policy and politics.

Throughout history, students have also played pivotal parts in revolutions, liberal democracy movements, and nationalist uprisings. The revolutions of 1848 in Europe saw students play key roles (Altbach, 1989). In Vienna, for example, the revolution was actually started by medical students, who petitioned the monarch for free speech, a free press, a constitution, and academic freedom. When their petition was not signed, they staged a demonstration that inspired the monarch to flee the country, and they then took over the management of the city for months (Robertson, 1969). Citing examples from India, Kenya, Burma, and Vietnam, Altbach (1989) argues that student activists were also influential and active in struggles for independence in Africa and Asia during the 20th century.

In 1918, the Córdoba Reform spread across Latin America after students at the National University of Córdoba in Argentina organized to demand curricular changes, free tuition, and student inclusion in governance (Walter, 1969). Using protests, strikes, and occupations, students prevailed in effecting changes to governance policies, and their success inspired similar student uprisings not only in other Latin American countries, but also in Asia and across Africa, resulting in students’ participation in decision making becoming institutionalized (Ntuli, 2019).

1919 saw several dramatic moments of student activism around the world. For example, in Germany, students’ councils organized a strike involving more than 85% of Berlin’s vocational school students. Not only did this strike result in agreements to end the use of corporal punishment by instructors and a change in the school schedule from evening to day classes, but it was also credited with altering the power relations between students and their teachers (Weipert, 2020). Meanwhile, on May 4th, 1919, students from Beijing University gathered to protest the Paris Peace Conference decision to grant Japan control over Chinese

provinces. After the government issued a violent crackdown on student demonstrators, the students responded by initiating a mass strike, which spread to universities and secondary schools throughout China, and inspired a parallel workers' strike, eventually leading the government to reject the Treaty of Versailles's pro-Japanese provisions (Boren, 2001).

The 1920s and 30s witnessed what some scholars have called "the dawn of widespread student activism" around the world (Rhoads, 1998, p. 33). In the US in 1932 alone, The National Student League staged several strikes, mass demonstrations, and creative protests directed both inward, toward campus matters, and outward, toward events happening in the world around them. They addressed campus specific issues, such as harsh student discipline, unfair censorship attempts, and unjust faculty dismissals, while also engaging in national politics, organizing students to stand in solidarity with striking Kentucky coal miners, for example (Cohen, 1993). In 1936, tens of thousands of students in India engaged in protests against the English language requirement for entrance to Indian universities; students also formed the All-India Students' Federation to advance the cause of Indian independence (Boren, 2001). In 1937, Brazilian students formed the União Nacional dos Estudantes (UNE) to advocate for the improvement of post-secondary education (Green and Skidmore, 2021).

War was also very much on the minds of student activists during this period. In the US in 1935, more than 150,000 students boycotted class in a strike for peace, and 1 year later, in 1936, upwards of 185,000 students took part in an anti-war demonstration (Boren, 2001). During World War II, Brazilian students in the UNE successfully pressured the country's president to join the allied forces (Green and Skidmore, 2021), and students at the University of Munich, in Germany, started the covert White Rose society, whose members risked arrest and execution to distribute leaflets exposing the horrors of the Nazi regime and encouraging resistance (Boren, 2001).

During the 1950s, student activists led decolonization struggles across the African continent (Ntuli, 2019), uprisings against authoritarian regimes in Asia, and helped to topple repressive governments throughout Latin America (Boren, 2001). In Europe, the Hungarian Revolution in 1956 was set into motion by students (Boren, 2001). In North America, however, student activism subsided for a time after the 1930s only to dramatically re-emerge during the 1960s. In fact, Williams (2013) has observed that it wasn't until the 1960s that the term "student activist" first appeared in print.

During the 1960s, student activists across the world not only demanded more relevant curriculum, a greater role in university governance, and stronger student rights, but also became involved in broader societal struggles for civil rights and liberation (Brooks, 2017). In the US, the Student Nonviolent Coordinating Committee, formed in April 1960, when Ella Baker convened students from lunch counter sit-ins and urged them to assume leadership in the struggle for racial equality and civil rights (Rhoads, 1998). The Free Speech Movement, born at UC Berkeley in 1964, mobilized students on many campuses to fight for greater freedom of expression and autonomy from restrictive and repressive university policies and "administrative intrusion" (Rhoads, 1998, p. 2). Student protests against the Vietnam War also surged during the latter half of the decade, including teach-ins on many campuses as well as large-scale demonstrations not only in the US, but in Canada, West Berlin, the UK, and Japan (Boren, 2001).

1968 was a particularly pivotal year, with student movements emerging in countries like Italy, Czechoslovakia, Argentina, and Japan. In Germany in 1968, the West German Student Movement mounted a powerful challenge to the political regime (Altbach, 1968), and in France in May 1968, student protestors set off a seven-week uprising against the De Gaulle government, American imperialism, and capitalism (Boren, 2001). Meanwhile, in Mexico City, in 1968, as the country prepared to host the Olympic Games, months of student protests against repression and violence culminated when troops opened fire on thousands of student protestors, killing what to this day remains an unknown number of youth, with estimates ranging from 4 to 3000 (Ordorika, 2021; Richman and Diaz-Cortez, 2008).

The 1970s witnessed the Soweto Uprising of 1976, in which school children in Soweto, South Africa went on strike to protest the Afrikaans Medium Decree mandating the use of Afrikaans for the teaching of mathematics and social studies. The strikes spread across schools, and as momentum built, the Soweto Students Representative Council's Action Committee organized a march and demonstration (Brown, 2016). On June 16th, approximately 20,000 students marched from their schools to the rally site. Police responding to one of the largest marches unleashed dogs and fired shots into the crowd, killing more than 100 students. This act of violence triggered international outrage, including condemnation by the UN, galvanized university students, and marked what some experts view as the beginning of the end of apartheid (Gallagher, 2002). Throughout the late 1970s and into the early 1990s, across the US, student activists pressured their universities to divest from investing in companies doing business in South Africa.

The second decade of the 21st century saw student activism once again surge around the world. 2011 saw the rise of movements in which students played prominent roles, including the Arab Spring (Altbach and Klemenčič, 2014), the international Occupy movement (Brooks, 2017), and the Chilean winter, in which students protested for free public education and an end to for-profit education (van de Velde, 2021). The Chilean student movement of 2011, which can be seen as a continuation of the rich legacy of Latin American student-led resistance against imperialism and colonialism, inspired similar student movements in Colombia, Brazil, and Mexico (Guzmán-Valenzuela, 2017). On the Asian continent, in what became known as the "sunflower movement," students in Taiwan occupied the legislature in Taipei for 24 days in March 2014 to protest a trade agreement with China that was being forced through parliament (Macfarlane, 2017). Encouraged in part by their peers in Taiwan, a few months later, students in Hong Kong boycotted classes to call for full democratic elections. Their action inspired hundreds of thousands of citizens to take the streets in what became known as the "Umbrella Revolution," because protestors used umbrellas to shield themselves from the tear gas and water cannons used by police to disperse them (Macfarlane, 2017; van de Velde, 2021). In the US, the killing of Michael Brown sparked mass unrest in Ferguson, Minnesota during the summer of 2015. The "Black Lives Matter"

chants that rung out during these protests accelerated a movement, born 2 years earlier after the killing of Trayvon Martin, that quickly gained ground on campuses across the US, as students returned in the fall of 2015.

Student activism garnered international attention again during the 2015–2016 academic year. Around the world, student activists staged dramatic acts of resistance to campus policies, protested guest speakers, and forced administrators to engage in long-overdue conversations about institutionalized racism, sexism, and xenophobia (Conner and Rosen, 2016). They demanded that school leaders issue statements, change policies, revise curricula, develop new diversity initiatives, and reconsider building names, campus iconography, terminology, and investment holdings. They called on governments to reverse fees and tuition increases and invest in public education as a public good (Luescher and Klemenčič, 2017). In many cases, they were successful in effecting change. In the years since, new cohorts of students have built on these efforts, continuing to push for deeper, more transformative institutional change as well as broader social change. Although the work continues to feel incremental and incomplete for students with bold post-neoliberal social imaginaries, their ambitions and radical hope are paving the way for new possibilities (Cole and Heineke, 2020).

Below, I offer examples of campaigns student activists have sustained or launched since 2015 on three different continents. These three campaigns speak to the range of issues that galvanize student activists, as well as the diversity of tactics they use, ranging from organizing protests and occupations to releasing reports and using social media and traditional media to drive attention to their causes. The cases show that student activism can target the state, the educational systems, or institutional specific policies, practices, and culture. Despite their points of distinction, the cases also collectively illustrate the tenacity and courage of student activists. Each campaign took careful coordination and ongoing commitment, and each was animated by a sense of moral urgency and collective responsibility. These three cases ultimately reveal that even though progress tends to be incremental and requires vigilance, student activism can lead to change.

Fighting cuts to higher education in Europe

As in the US and Korea (Shin et al., 2014), across Europe neoliberal reforms and austerity measures have led to substantial funding cuts in government spending on higher education, resulting in rising tuition costs. Students have responded by conducting research, organizing mass protests, and leveraging the media. For example, the European Student Union published a study based on findings from a survey of 38 national unions of students across Europe, which found that “underfunding of HEIs is leading to reductions in student services, growing tuition fees, ultimately damaging education quality” (O’Driscoll et al., 2015, p. 6). The report called for countries to “take on the responsibility to fund the reforms that they have (or should have) implemented” and to involve students “in all discussions and decision-making” (p. 7). In Denmark in October 2015, between 22,000 and 40,000 student protestors marched to the Copenhagen Parliament, chanting, “Stop the budget cuts in the billion class!” “Improve quality,” and “Invest in the future” (Myklebust, 2015). Thousands of students demonstrated again 2 years later in five Danish cities to protest the effects of continued cuts to education funding (Danish students to demonstrate, 2017). After continued outcry, a new Socialist Government in Denmark passed legislation in 2019 to end the requirement that universities cut their budgets by 2% each year (Matthews, 2019).

Divestment at University of Cape Town

In February 2017, amid mounting pressure from student activists, the Convocation, a statutory body representing alumni, vice chancellors, and academic staff, voted in favor of a resolution to divest the University of Cape Town from all fossil fuel holdings within 5 years (Giamporcaro and Dhlamini, 2017). This vote marked the culmination of a 4-year campaign by student activists to demand divestment. Also in 2017, student activists in Fossil Free UCT won the creation of the University Panel on Responsible Investment (UPRI), an advisory body that would make recommendations pertaining to the endowment. Despite these wins, no changes occurred. In fact, it took four more years and two more votes by the Convocation in favor of similar resolutions before the UPRI formally launched. Shortly after its launch, in May 2021, students from the Green Campus Initiative, Fossil Free South Africa, and the Climate Justice Charter Movement staged a protest renewing their calls for fossil fuel divestment. Part of their protest involved delivering a memo to the Chairperson of the UPRI, which included demands for “more transparent and accountable communication on where the institution’s fossil fuel investments lie” and a commitment “to complete decarbonization of [UCT’s] operations by 2040” (O’Regan, 2021). As they have sustained their campaign, the student activists have drawn inspiration from not only the Rhodes Must Fall and Fees Must Fall movements, but also the South African apartheid divestment efforts of US and UK students in the 1980s and 90s and the Black Lives Matter protests of the day (Giamporcaro and Dhlamini, 2017; Kusmer, 2021).

#NotAgainSU

In the fall of 2019, in the wake of documented racial incidents on campus that student activists felt were not appropriately addressed by the President and Chancellor, students at Syracuse University in New York launched a sit-in in the Health and Wellness complex, under the banner #NotAgainSU. The eight-day sit-in ended when 16 of their 19 demands were met by the administration, with the remaining three demands partially met. The University’s Office of Diversity and Inclusion posted each demand, the university’s response, and progress indicators online on a page devoted to concerns of the #NotAgainSU students. When new anti-Semitic

and racist images began circulating again on campus, however, the students understood this progress to be inadequate and launched a second occupation, this time of the main administrative building. This occupation lasted for 32 days and only ended when the Covid pandemic shut down the entire campus. During this occupation, the student activists revised 18 of their original demands, and added ten new demands. Administrators and protesters were in negotiations in early March, joined by professors, parents, and legal counsel, when the Covid pandemic forced the campus closure. On the 1-year anniversary of launch of sit in, Feb 17, 2021, the #NotAgainSU issued a letter, which centered an abolitionist framework and blasted surface-level reforms (“i.e., body cameras, diversity training, creating more committees, etc.”) as “ineffective because they will never solve any problems that deal with systemic issues because it does not begin to identify the root causes of racism and oppression.” The work of #NotAgainSU continues.

Tactics, issues, and campaigns of US college students

Contemporary college student activists in the US draw on a wide and ever-expanding repertoire of strategies and tactics, restlessly innovating in their efforts to effect change. In addition to “tried and true” practices, like protests, sit-ins, teach-ins, and strikes, students have deployed newer tactics, like “mic-checks” or echoed refrains during lectures by controversial speakers and die-ins. These tactics may be considered variations on a theme, in that they represent forms of protest, sit-ins, or even political theater. Conservative students have staged provocative events, such as affirmative action bake sales, in which they charge students of different races different prices and “Catch an Illegal Immigrant Day” (Cowen et al., 2019). The most notable tactical innovations, however, have come from students’ use of digital media, as students avail themselves of the tools of their day to make their voices heard. This is not to say that digital activism has taken the place of old-fashioned frontline activism. According to one study, protests and demonstrations remain the most common tactic used by student activists, reported by more than 80% of a sample of 237 student activists representing campuses in 47 US states (Conner, 2020); however, the novel use of digital media by Gen Z student activists continues to be noteworthy.

Data from recent studies of US student activists’ tactics highlights the emergence of digital activism in recent years. In a study of 79 4-year private US college campuses between 1989 and 2010, Barnhardt (2014) found that more than half the campuses reported experiencing petitions, rallies, and letter-writing campaigns, and a little more than a third reported experiencing student protests and demonstrations. The next most common tactics were leafletting and pamphleteering, experienced on about one-quarter of the campuses. Notably, Barnhardt’s list of tactics did not include online or social media activism. Five years later, in a 2015–2016 study of US student activists, Conner (2020) found that more than 60% of the sample reported engaging in online campaigns. And in 2020, during the height of the Coronavirus pandemic, a study of US climate justice and gun violence prevention student activists found that 81% had created online content for their movement, 86% had participated in an online event, such as a rally, town hall, or webinar, related to their movement, and 100% had signed an online petition for their movement within the last 6 months (Conner et al., 2021b). These studies illustrate the growing prominence of digital activism among US-based student activists since 2010.

Today’s student activists use a full suite of digital media tools, including videos, memes, and posts, to raise awareness, mobilize people and resources, apply institutional pressure, and increase accountability. In addition to using tools like Zoom, Groupme, and Slack to coordinate and organize, activists use TikTok, Instagram, and Twitter to record or express their activism (Davis, 2019). Some activists denounce what they call “performative activism,” which refers to superficial forms of activism that are demonstrated online or in meetings or classrooms in order to burnish one’s image, and which are not buttressed by offline investments of time, money, or labor for the causes the activist purports to care about. Despite these critiques, online activism has become influential and impactful.

For example, during the summer of 2020, as people around the world reacted to the murder of George Floyd and called for a racial justice reckoning, the Instagram Black@ stories of Black students’ experiences of racism and microaggressions on particular US campuses attracted widespread attention, forcing institutional responses. At the University of Southern California, for instance, the President initially responded to the hundreds of stories posted to the Black_at_usc Instagram page by calling for forums, surveys, and interviews. The students rebuffed this approach as woefully insufficient and insensitive (Navarro, 2020). Under mounting pressure, the President then wrote to the community announcing several more substantive changes, including the removal of a controversial name from a building, the recognition of Juneteenth as a holiday, the launch of a new office and means of reporting incidences of bias, the creation of a Community Advisory Board for the Department of Public Safety, and the establishment a Task Force on Racial Equity, Diversity, and Inclusion, charged with examining processes that perpetuate racism and inequity and issuing recommendations within the year. Although it comes with risks, such as cyberbullying and doxing, online activism can galvanize real world change. Students have wielded this tool strategically to force administrators’ hands, compelling them to act in order to protect the institution’s reputation, if not the students who are voicing their hurt.

Paralleling the rise of digital media has been the rise in coalitions among student activists. Students have formed coalitions both on a single campus (Mustaffa and Mayorga, 2019) and across campuses (Keeton et al., 2021). Some coalitions are focused on a singular issue, like ending sexual assault on campus (Brodsky, 2016) or fossil fuel divestment (Harper et al., 2022), while others aim to build solidarities across issues, reflecting students’ understanding of systems of oppression as interlocked and entangled. Coalitions have helped student activists to build their capacity as well as their power. At Princeton University, for example, in May 2021, a coalition of five student organizations (Divest Princeton, Natives at Princeton, the Princeton Indigenous Advocacy

Coalition, Students for Prison Education, Abolition, and Reform, and Change Princeton Now) released a joint demand letter, which they also displayed in an art installation on campus.

Although identity-based activism remains prominent on US college campuses (Linder et al., 2020), single-issue activism is rare among Generation Z student activists. Prefiguring this development, Lorde (1984) wrote, "There is no such thing as a single-issue struggle because we do not live single-issue lives." One study found that only 10% of US-based student activists identify as single-issue activists, and 52% direct their activism at seven or more issues simultaneously (Conner, 2020).

In 2015, prior to the election of Donald Trump, the most prominent issues for US-based student activists were racial justice, gender equity, LGBTQ rights, human rights, and ending sexual assault. The top three issues have remained stable into 2021 (Greytak et al., 2021).¹ For a time, the mass shooting at Marjory Stoneman Douglas High School and the advent of March for Our Lives in 2018 vaulted gun violence prevention into the top issue areas of concern, but organizing in this area has since subsided. Climate activism has been on the ascendancy since Greta Thunberg emerged on the international stage, inspiring mass school strikes in 2019 and renewing calls for divestment from fossil fuels. Environmental and climate justice was the fourth most common issue area cited by a sample of 624 student activists in the summer of 2021 (Greytak et al., 2021). Meanwhile, racial justice has remained at the fore, especially after Trump referred to white supremacists as "very fine people," hate crimes surged on college campuses, and the police killings of George Floyd, Breonna Taylor, Walter Wallace, Jr., and other unarmed Black individuals drew hundreds of thousands to the streets to demand justice. Trump's Muslim ban, his fervor for building "the wall" between the US and Mexico, and his attempts to end the Deferred Action for Childhood Arrivals program also inspired considerable activism and advocacy among students, fighting for immigrant rights and protections for undocumented students. Finally, the Coronavirus pandemic and the economic precarity it exacerbated and exposed have contributed renewed energy to fights for tuition-free colleges and student debt relief. Outside the US, students remain engaged in pro-democracy struggles, in countries such as Argentina, Chile, Kenya, Mexico, Thailand, Uganda, and Venezuela, and in efforts to contest neoliberal education policy (Brooks, 2017).

Although both historically (Altbach, 1989) and currently, the majority of US college student activists tend to be leftist or progressive in their politics, conservative student activism is on the rise (Ince et al., 2018). The growing conservative student movement both reflects and feeds broader societal trends, such as that of Trumpism. Conservative student activists tend to be whiter and more economically privileged than their liberal counterparts (Ince et al., 2018). On some US campuses, conservative student activism is supported by deep-pocketed national organizations, such as the Young America's Foundation and the Leadership Institute, which issue "battle plans" for confrontational activism (Binder, 2013; Binder and Wood, 2013). In recent years, Turning Point USA has become particularly influential, growing from \$4.3 million in revenues in 2016 to 39.8 million in 2020 (Stone, 2021). In addition to staging events on and off campuses, Turning Point USA has encouraged students to record liberal professors and post the recordings online, and it maintains a "watch-list" of left-leaning faculty. On some campuses, as conservative student activists continue to stage provocative events and invite inflammatory speakers, they will incite clashes with liberal students and force administrators to articulate how they weigh competing interests, such as protecting freedom of speech and intellectual diversity on the one hand, while honoring institutional mission statements that call for respecting the dignity of all people, including the most marginalized, on the other.

Beyond campus, conservative students' complaints of a liberal bias and lack of free speech on campus are helping fuel efforts in some Republican-led states to defund higher education and to write legislation that erodes tenure protections for professors who teach critical perspectives on race and gender (Gaudiano, 2022). On their own, conservative student activists may have been less successful at changing their campuses than progressive student activists (Ince et al., 2018); however, in concert with policymakers, they are positioned to have a profound effect on higher education policy and practice in Republican controlled states.

Effects on students

The literature on US-based student activists suggests that students become activists for a variety of reasons. Some are moved to action by moments of moral outrage or "shocks to the conscience," that is, events or experiences that compel them to act (Holeman, 2007). Others journey into activism more gradually, in fits and starts, or through circuitous routes (Conner, 2020), and still others feel they have no choice but to be an activist—that their very existence is a form of activism (Logan et al., 2017; Stewart and Williams, 2019).

Regardless of how they come to identify as activists, scholars have framed this process in terms of sociopolitical development. Sociopolitical development involves aligning one's actions with one's social analysis and world view. According to the theory, activists evince both critical consciousness (or a critical perspective on social structures, systems, policies, and practices) and critical social action (or behaviors) intended to "influence or transform communities and societal systems to be more equitable and just" (Anyiwo et al., 2020, p. 86).

¹The American Civil Liberties Union (ACLU), a non-partisan organization with 1.6 million members and affiliates in every state, holds an annual ACLU National Advocacy Institute in late July (usually in Washington, DC, but virtually in 2020 and 2021). In its sixth year in 2021, the program has brought together more than 5000 high school and college-aged students (ages 15–24) from across the country. During the ACLU Advocacy Institute, students work with lawyers, organizers, and other experts to learn about the ACLU's issues, advocacy tools (judicial, legislative, and grassroots), and build networks to support each other as they return to their campuses and communities. For more information visit: www.aclu.org/institute.

Scholars have examined the benefits as well as the costs associated with engagement in activism and critical social action. These effects occur in various domains, most concentrated in academic/learning, emotional and physical, and relational domains. In each of these areas, the evidence for US-based student activists is mixed. It is, of course, important to recognize that these effects are shaped by the geopolitical context. In a relatively stable democracy, like the US, where freedom of speech protections exist, student activists may face less serious risks of harm than student activists in more autocratic and repressive regimes, like Russia, Sudan, or Nigeria, where protestors may be killed or disappeared or have their passports and assets seized by government officials.

Although some research with high school aged students has found associations between engagement in activism and positive academic outcomes, including higher GPA (Conner and Slattery, 2014; Seider et al., 2019) and SAT scores (Seider et al., 2019), the research with college student activists is contradictory. In a study of 25 student activists from 14 institutions, Linder et al. (2019) found students reported decreased academic performance and learning as a result of their activism, consistent with findings from other work (Vaccaro and Mena, 2011); however, in a study with 40 student activists from 32 different campuses, Conner (2020) found that 57% believe their activism enhanced their performance as students, while only 12% perceived an exclusively negative impact on their grades or learning, and 15% reported a mixed effect. Beyond indicators of academic achievement and self-reports, a large body of scholarship has identified a variety of learning outcomes associated with engagement in activism, including critical social analysis and critical thinking (Keeton et al., 2021; Kezar, 2010; Tsui, 2000; Quinn and Nguyen, 2017), time management and communication skills (Ollis, 2011), a better understanding of how to navigate institutional bureaucracies and systems (Keezar and Maxey, 2014), and leadership skills (Hoffman and Mitchell, 2016; Rhoads, 2016; Stewart and Quayle, 2019).

With regard to mental and physical health, the research is again inconclusive. Some studies highlight activism as a tool that allows students to cope with or even heal from the forms of oppression or violence they have experienced or continue to experience (Ginwright, 2010; Hope and Spencer, 2017). Some studies link activism to enhanced feelings of empowerment or agency (Christens and Peterson, 2012; Stake and Hoffman, 2001). Other work draw attention to the costs of engaging in activism, including exhaustion, fatigue, and burnout (Cho, 2020; Linder et al., 2019; Ortega-Williams, 2018; Smith et al., 2006; Stewart and Quayle, 2019; Stokes and Miller, 2019; Vaccaro and Mena, 2011). One study found that activism may decrease the stress and anxiety associated with experiencing microaggressions on campus for Latinx students, but increase it for Black students (Hope et al., 2018). Conner et al. (2021a) found that 60% of the student activists in their study reported that activism had adverse consequences for their psychological well-being. Nevertheless, the students remained committed to continuing their activism because the connection to others and sense of purpose they derived from their activism, in conjunction with the opportunities to effect change and practice self-care and collective care, helped to offset the stress, exhaustion, and guilt they experienced as a result of activism.

In the relational domain, results continue to vary. Some work finds that activism benefits students as they experience “meaningful interactions” and find “kinship” (Keeton et al., 2021). A recent quantitative study found youth organizers experienced greater gains in the areas of social capital and social skills than their matched peers, who were engaged in non-activist activities (Flores, 2020). Stake and Hoffman (2001) found that after participating in demonstrations, student activists felt less isolated. By contrast, other work has found that student activists experience increasing isolation from peers and family and their personal relationships suffer because of their activism (Jones, 2017; Linder et al., 2019). Results from a survey of 624 student activists demonstrated that on average, activists perceived the effect of their activism on their familial and peer relationships as more positive (4.36 and 5.88) than negative (3.28 and 3.47) (Greytak et al., 2021).

Clearly, activism is a complex phenomenon, with a wide range of possible effects. Questions remain about how the benefits and costs of activism may be experienced differently by students with different intersectional identities, practicing different forms of activism, for different reasons, under different institutional conditions, and in different regional contexts. The next generation of research on student activism can begin to tease out some of this variation.

Effects on institutions

Compared to research focused on student activists, considerably less contemporary scholarship examines the effects of student activism on institutions. A small but burgeoning literature documents institutional responses to activism (Cho, 2020; Cole and Harper, 2017; Fisher, 2018; Pryor, 2020). This research identifies a range of institutional tactics, including issuing statements, establishing task forces, launching surveys, holding town halls, and passing resolutions (Cho, 2020). Institutional responses to student activism may also include strategies such as gaslighting students; waiting out, placating or appeasing students; coopting student campaigns; or silencing students (Cho, 2020; Conner, 2020; Linder et al., 2019), as well as acceding to or accommodating students’ demands (Fisher, 2018; Williams and McGreevey, 2004). This literature has revealed that campus leaders’ responses to activists are motivated primarily by an imperative to minimize risk and protect the institution’s reputation. Nonetheless, questions of when, how, and why institutions change in response to student activists’ pressure campaigns have not been fully explored.

We know from earlier research that institutions do make changes because of pressure from student activists. For example, student demands in the 1960s led to the establishment of ethnic studies departments (Acuna, 2011; Claybrook, 2013; Pulido, 2006), the hiring of more faculty and administrators of color (Lee, 2011), and the creation of cultural centers on campus (Patton, 2010; Yosso and Benavides López, 2010). More recently, Bovill et al. (2020) credited student activism—enacted through protests, sit-ins, legal actions, federal complaints, and traditional and social media campaigns—in the US and the UK with challenging and changing federal policy regarding how cases of sexual assault are handled on college campuses, driving shifts in institutional

practices. Fisher (2018) analyzed four colleges' responses to student activists, finding that in all cases, institutions responded in "unconditionally constructive" ways to radical student activism, eventually accommodating student demands; however, this study stopped short of identifying "the specific factors determining the institutions' stances" (Fisher, 2018, p. 26).

Some work has found that even when institutions make changes in response to student activists' demands, these reforms do not translate into meaningful changes to the everyday experiences of students, leaving student activists feeling frustrated and demoralized (Cho, 2020; Hoffman and Mitchell, 2016; Pryor, 2020). As Kezar (2018) has argued, second-order changes that seek to fundamentally restructure systems are challenging to effectuate in higher education. Incremental change tends to be the norm. Furthermore, determining whether and how student activism changes institutional culture, influences public opinion, or foments social transformation is exceedingly difficult to measure. Innovative methods, including longitudinal approaches, are needed to better understand the short and long-term effects of student activism on their institutions and society.

Conclusion

Student activism has had a long and storied history. Around the world, students have catalyzed institutional changes, ignited social movements, and even helped topple governments (Altbach, 1989; Boren, 2001). Of course, not everyone has the privilege of being able to claim an identity as a student, and those who do tend to remain students only for a relatively short period of time. But in that fleeting moment, the small subset of students who don the mantle of activist, continue a rich tradition of envisioning a better future and pressing for progress.

References

- Acuña, R.F., 2011. *The Making of Chicana/o Studies: In the Trenches of Academe*. Rutgers University Press. <https://doi.org/10.36019/9780813550701>.
- Altbach, P.G., 1968. Student revolt in Europe. *Econ. Polit. Wkly.* 3 (19), 755–756.
- Altbach, P., 1989. Perspectives on student political activism. In: Altbach, P.G. (Ed.), *Student Political Activism: An International Reference Handbook*. Greenwood, pp. 1–17.
- Altbach, P.G., Klemenčič, M., 2014. Student activism remains a potent force worldwide. *Int. High. Educ.* 76, 2–3.
- Anyiwo, N., Palmer, G., Garrett, J., Starck, J., Hope, E., 2020. Racial and political resistance: an examination of the sociopolitical action of marginalized youth. *Curr. Opin. Psychol.* 35, 86–91.
- Barnhardt, C., 2014. Campus-based organizing: tactical repertoires of contemporary student movements. *N. Dir. High. Educ.* 167, 43–58. <https://doi.org/10.1002/he.20104>.
- Binder, A., 2013. How National Advocates Encourage Conservative Activism on America's College Campuses. Scholars Strategy Network. Available at: <https://scholars.org/contribution/how-national-advocates-encourage-conservative-activism-americas-college-campuses>.
- Binder, A., Wood, K., 2013. *Becoming Right: How Campuses Shape Young Conservatives*. Princeton University Press.
- Boren, M., 2001. *Student Resistance: A History of the Unruly Subject*. Routledge.
- Bovill, H., McMahon, S., Demers, J., Banyard, V., Carrasco, V., Keep, L., 2020. How does student activism drive cultural campus change in the UK and US regarding sexual violence on campus? *Crit. Soc. Pol.* 0, 1–23. <https://doi.org/10.1177/0261018320913967>.
- Braxton, E., 2016. Youth leadership for social justice: past and present. In: Conner, J., Rosen, S. (Eds.), *Contemporary Youth Activism: Advancing Social Justice in the United States*. ABC-CLIO, Santa Barbara, CA, pp. 25–38.
- Brodsky, A., 2016. Working for change, learning from work: student empowerment and challenges in the movement to end campus gender violence. In: Conner, J., Rosen, S.M. (Eds.), *Contemporary Youth Activism*. ABC-CLIO, pp. 245–261.
- Brooks, R., 2017. Student politics and protest: an introduction. In: Brooks, R. (Ed.), *Student Politics and Protest: International Perspectives*. Routledge, pp. 1–11.
- Brown, J., 2016. *The Road to Soweto*. James Currey.
- Brown, J., 2021. *June Forth: The Tiananmen Protests and Beijing Massacre of 1989*. Cambridge University Press.
- Cho, K.S., 2020. Responding to Campus Racism: Analyzing Student Activism and Institutional Responses. Cho_ucla_0031D_19082 [Doctoral Dissertation, University of California Los Angeles. ProQuest LLC].
- Christens, B.D., Peterson, N.A., 2012. The role of empowerment in youth development: a study of sociopolitical control as mediator of ecological systems' influence on developmental outcomes. *J. Youth Adolesc.* 41 (5), 623–635. <https://doi.org/10.1007/s10964-011-9724-9>.
- Claybrook Jr., M.K., 2013. Black power, black students, and the institutionalizing of change: Loyola Marymount University, 1968–1978. *J. Pan Afr. Stud.* 5 (10), 1–19.
- Cohen, R., 1993. *When the Old Left Was Young: Student Radicals and America's First Mass Student Movement, 1929–1941*. Oxford University Press, New York, NY.
- Cole, E.R., Harper, S.R., 2017. Race and rhetoric: an analysis of college presidents' statements on campus racial incidents. *J. Divers. High. Educ.* 10 (4), 318–333.
- Cole, R.M., Heinecke, W.F., 2020. Higher education after neoliberalism: student activism as a guiding light. *Pol. Futures Educ.* 18 (1), 90–116. <https://doi.org/10.1177/1478210318767459>.
- Conner, J., 2020. *The New Student Activists: The Rise of Neoactivism on College and University Campuses*. Johns Hopkins University Press, Baltimore, MD.
- Conner, J., Rosen, S., 2016. Introduction. In: Conner, J., Rosen, S. (Eds.), *Contemporary Youth Activism*. Praeger, pp. 1–18.
- Conner, J., Slattery, A., 2014. New media and the power of youth organizing: minding the gaps. *Equity Excell. Educ.* 47, 14–30. <https://doi.org/10.1080/10665684.2014.866868>.
- Conner, J., Galioto, M., Crawford, E., 2021a. The mental health effects of student activism: persisting despite psychological costs. *J. Adolesc. Res.* <https://doi.org/10.1177/07435584211006789>.
- Conner, J., Lotesta, J., Wang, T., Foley, A., Oshyn, K., Febus, M., 2021b. *Youth Movements and Voting*. CIRCLE, Boston, MA.
- Cowen, G.H., Hemer, K.M., Reason, R.D., 2019. Understanding American conservatism at its role in higher education. In: Morgan, D., Davis III, C.H.F. (Eds.), *Student Activism, Politics, and Campus Climate in Higher Education*. Routledge, New York, pp. 45–59.
- Cummings, P.M.M., 2015. Democracy and student discontent. *J. Polit. Lat. Am.* 7 (3), 49–84. <https://doi.org/10.1177/1866802X1500700302>.
- Danish students to demonstrate in five cities, 2017. *The Local* [online]. Available at: <https://www.thelocal.dk/20171005/danish-students-to-demonstrate-in-five-cities/>. (Accessed 7 July 2021).
- Davis III, C.H.F., 2019. Student activism, resource mobilization, and new tactical repertoires in the digital age. In: Morgan, D., Davis III, C.H.F. (Eds.), *Student Activism, Politics, and Campus Climate in Higher Education*. Routledge, New York, pp. 112–124.
- Fisher, M., 2018. Still fighting the good fight: an analysis of student activism and institutional response. *Publ. Relat. Rev.* 44 (1), 22–27. <https://doi.org/10.1016/j.pubrev.2017.11.005>.
- Flores, K.S., 2020. *Transforming Positive Youth Development: A Case for Youth Organizing*. Funders Collaborative on youth organizing, Brooklyn.

- Gallagher, M., 2002. The birth and death of apartheid. BBC News [online]. Available at: <http://news.bbc.co.uk/2/hi/africa/575204.stm>. (Accessed 7 July 2021).
- Giamporcaro, A., Dhlamini, X., 2017. South African Universities Need to Rethink How They Invest Their Millions. The Conversation [online]. Available at: <https://theconversation.com/south-african-universities-need-to-rethink-how-they-invest-their-millions-81306>. (Accessed 7 July 2021).
- Gill, J., de Fronzo, J., 2009. A comparative framework for the analysis of international student movements. *Soc. Mov. Stud.* 8 (3), 203–224. <https://doi.org/10.1080/14742830903024309>.
- Ginwright, S., 2010. *Black Youth Rising: Activism and Healing in Urban America*. Teachers College Press, New York.
- Green, J.N., Skidmore, T., 2021. *Brazil: Five Centuries of Change*. Oxford University Press.
- Greytak, E., Evich, C., Viveros, M., 2021. Evaluation of ACLU's National Advocacy Institute. Unpublished Report. American Civil Liberties Union, New York.
- Guidiano, N., 2022. Colleges are Fighting Conservative Bills in States Like Florida and Texas That They Say Throw Red Meat to Trump Supporters and Muzzle Educators. *Business Insider*. Available at: <https://www.businessinsider.com/colleges-fighting-conservative-woke-act-bills-critical-race-theory-gender-2022-3>.
- Guzmán-Valenzuela, C., 2017. Neoliberal discourse and the emergence of an agentic field: the Chilean student movement. In: Brooks, R. (Ed.), *Student Politics and Protest: International Perspectives*. Routledge, pp. 47–62.
- Harper, J., Dizon, J.P.M., Kezar, A., 2022. Using strategies elites understand: divestment as an economic approach to social change. *Peabody J. Educ.*
- Hoffman, G., Mitchell, T., 2016. Making diversity “everyone’s business”: a discourse analysis of institutional responses to student activism for equity and inclusion. *J. Divers. High. Educ.* 9 (3), 277–289. <https://doi.org/10.1037/dhe0000037>.
- Holeman, H., 2007. *Awakening a Social Conscience* (Unpublished Doctoral Dissertation). University of Oklahoma, Norman, Oklahoma.
- Hope, E.C., Velez, G., Offidani-Bertrand, C., Keels, M., Durkee, M.I., 2018. Political activism and mental health among Black and Latinx college students. *Cult. Divers. Ethn. Minor. Psychol.* 24 (1), 26–39. <https://doi.org/10.1037/cdp0000144>.
- Hope, E.C., Spencer, M.B., 2017. Civic engagement as an adaptive coping response to conditions of inequality: an application of phenomenological variant of ecological systems theory (PVEST). In: Cabrera, N.J., Leyendecker, B. (Eds.), *Handbook on Positive Development of Minority Children and Youth*. Springer International Publishing, Netherlands, pp. 421–435. https://doi.org/10.1007/978-3-319-43645-6_25.
- Hunter-Gault, C., 2013. Fifty years after the Birmingham children’s crusade. *The New Yorker*. <https://www.newyorker.com/news/news-desk/fifty-years-after-the-birmingham-childrens-crusade>.
- Ince, J., Finlay, B., Rojas, F., 2018. College campus activism: distinguishing between liberal reformers and conservative crusaders. *Sociol. Compass* 12, 1–14. <https://doi.org/10.1111/soc4.12603>.
- Jones, R., 2017. The heterogeneity of resistance: how Black students utilize engagement and activism to challenge PWI inequalities. *J. Negro Educ.* 86, 204–219. <https://doi.org/10.7709/jnegroeducation.86.3.0204>.
- Keeton, R.G., López, C.B., Aguilar-Hernández, J.M., 2021. “It shaped who I am”: reframing identities for justice through student activism. *Assoc. Mex. Am. Educ. J.* 15 (1), 1–28. <https://doi.org/10.24974/amae.15.1.414>.
- Kezar, A., 2010. Faculty and staff partnering with student activists: unexplored terrains of interaction and development. *J. Coll. Student Dev.* 51 (5), 451–480. <https://doi.org/10.1353/csd.2010.0001>.
- Kezar, A., 2018. *How Colleges Change*, second ed. Routledge, New York.
- Kezar, A., Maxey, D., 2014. Collective action on campus toward student development and democratic engagement. *N. Dir. High. Educ.* 167, 31–41. <https://doi.org/10.1002/he.20103>.
- Kusmer, A., 2021. Fossil fuel divestment activists inspired by anti-apartheid struggle. *The World* [online]. Available at: <https://www.pri.org/file/2021-03-18/fossil-fuel-divestment-activists-inspired-anti-apartheid-struggle>. (Accessed 7 July 2021).
- Lee, P., 2011. The Griswold 9 and student activism for faculty diversity at Harvard Law School in the early 1990s. *Harv. J. Racial Ethn. Justice* 27, 49–96.
- Linder, C., Quaye, S., Lange, A., Roberts, R., Lacy, M., Okello, W., 2019. “A student should have the privilege of just being a student”: student activism as labor. *Rev. High. Educ.* 42 (5), 37–62. <https://doi.org/10.1353/rhe.2019.0044>.
- Linder, C., Quaye, S., Lange, A., Evans, M.E., Stewart, T.J., 2020. *Identity-Based Student Activism: Power and Oppression on College Campuses*. Routledge, New York.
- Logan, G., Lightfoot, B., Contreras, A., 2017. Black and Brown millennial activism on a PWI campus in the era of Trump. *J. Negro Educ.* 86 (3), 252–268. <https://doi.org/10.7709/jnegroeducation.86.3.0252>.
- Lorde, A., 1984. *Sister Outsider: Essays and Speeches*. Crossing Press, Toronto.
- Luescher, T., Klemenčič, M., 2017. Student power in 21st-century Africa: the character and role of student organizing. In: Brooks, R. (Ed.), *Student Politics and Protest: International Perspectives*. Routledge, pp. 113–128.
- Macfarlane, B., 2017. “If not now, then when? If not us, who?” Understanding the student protest movement in Hong Kong. In: Brooks, R. (Ed.), *Student Politics and Protest: International Perspectives*. Routledge, pp. 143–156.
- Matthews, D., 2019. New Danish government ends budget squeeze. *Times Higher Education* [online]. Available at: <https://www.timeshighereducation.com/news/new-danish-government-ends-budget-squeeze-except-humanities>. (Accessed 7 July 2021).
- Mirza, M., Pinckney, J., Rivers, M., 2022. Young People are Overcoming Political Polarization. *The Forge*. Available at: <https://forgeorganizing.org/article/young-people-are-overcoming-political-polarization>.
- Mustaffa, J.B., Mayorga, O.J., 2019. The coalitional factors of student activism. In: Dache, A., Quaye, S.J., Linder, C., McGuire, K. (Eds.), *Rise Up!: Activism as Education*. Michigan State University Press, East Lansing, MI, pp. 177–199.
- Myklebust, J.P., 2015. Students Hold Mass Protests Against Education Cuts. *University World News* [online]. Available at: <https://www.universityworldnews.com/post.php?story=20151029205344289>. (Accessed 7 July 2021).
- Navarro, S., 2020. A digital movement: creating a safe space for black trojans seeking institutional change. *Daily Trojan* [online]. Available at: <https://dailytrojan.com/2020/07/06/a-digital-movement-creating-a-safe-space-for-black-trojans-seeking-institutional-change/>. (Accessed 7 July 2021).
- Ntuli, M.E., 2019. Student Activism and Its Role in Achieving the SDGs. *University World News* [online]. Available at: <https://www.universityworldnews.com/post.php?story=20191211132312407>. (Accessed 7 July 2021).
- Ollis, T., 2011. Learning in social action: the informal and social learning dimensions of circumstantial and lifelong activists. *Aust. J. Adult Learn.* 51, 248–268.
- Ordorika, I., 2021. Student movements and politics in Latin America: a historical reconceptualization. *High. Educ.* 83, 297–315. <https://doi.org/10.1007/s10734-020-00656-6>.
- Ortega-Williams, A., Wernick, L.J., DeBower, J., Brathwaite, B., 2018. Finding relief in action: the intersection of youth-led community organizing and mental health in Brooklyn, New York City. *Youth Soc.* 52 (4), 618–638. <https://doi.org/10.1177/0044118X18758542>.
- O’Driscoll, C., Fröhlich, M., Gehrke, E., Isoski, T., Maoláin, A.O., Meister, M., Nordal, E., Palomares, F.M.G., Pietkiewicz, K., Sanchez, I., Todorovsk, B., 2015. *Bologna With Students’ Eyes 2015. The European Students’ Union*, Brussels.
- O’Regan, V., 2021. Green power: UCT put in crosshairs by climate crisis activists taking a stand against fossil fuels. *Daily Maverick* [online]. Available at: <https://www.dailymaverick.co.za/article/2021-05-11-green-power-uct-put-in-crosshairs-by-climate-crisis-activists-taking-a-stand-against-fossil-fuels/>. (Accessed 7 July 2021).
- Patton, L.D., 2010. A call to action: historical and contemporary reflections on the relevance of campus culture centers in higher education. In: Patton, L.D. (Ed.), *Culture Centers in Higher Education: Perspectives on Identity, Theory, and Practice*. Stylus, Sterling, VA, pp. xiii–xvii.
- Pryor, D., 2020. *Institutional Responses to Student Activism Related to Racial Injustice at Two Predominately White Institutions: Administrator Actions and Student Perceptions* (Publication No. 27829843) [Doctoral Dissertation, New York University]. ProQuest LLC.
- Pulido, L., 2006. *Black, Brown, Yellow, and Left: Radical Activism in Los Angeles*. University of California Press. <https://doi.org/10.1525/9780520938892>.
- Quinn, R., Nguyen, C., 2017. Immigrant youth organizing as civic preparation. *Am. Educ. Res. J.* 54 (5), 972–1005.

- Rhoads, R.A., 1998. *Freedom's Web: Student Activism in an Age of Cultural Diversity*. Johns Hopkins University Press, Baltimore, MD.
- Rhoads, R., 2016. Introduction: student activism, diversity, and the struggle for social justice. *J. Divers. High. Educ.* 9 (3), 189–202. <https://doi.org/10.1037/dhe0000039>.
- Richman, J., Diaz-Cortez, A., 2008. Mexico's 1968 massacre: what really happened? *All Things Considered*. Available at. <https://www.npr.org/templates/story/story.php?storyId=97546687>.
- Robertson, P., 1969. Students on the barricades: Germany and Austria 1848. *Polit. Sci. Q.* 84 (2), 367–379. <https://www.jstor.org/stable/2147265>.
- Schmidt, C., 2018. *The Sit-Ins: Protests and Legal Change in the Civil Rights Era*. University of Chicago Press.
- Seider, S., Clark, S., Graves, D., 2019. The development of critical consciousness and its relation to academic achievement in adolescents of color. *Child Dev.* 91 (2), 451–454. <https://doi.org/10.1111/cdev.13262>.
- Shin, J., Kim, H., Choi, H., 2014. The evolution of student activism and its influence on tuition fees in South Korean universities. *Stud. High. Educ.* 39 (3), 441–454.
- Smith, W.A., Yosso, T.J., Solorzano, D.G., 2006. Challenging Racial Battle Fatigue on Historically White Campuses: A Critical Race Examination of Race-Related Stress.
- Stake, J., Hoffman, F., 2001. Changes in student social attitudes, activism, and personal confidence in higher education: the role of women's studies. *Am. Educ. Res. J.* 38 (2), 411–436. <https://doi.org/10.3102/00028312038002411>.
- Stewart, T.J., Quaye, S.J., 2019. Building bridges: rethinking student activist leadership. *N. Dir. Sch. Leader.* 161, 51–63.
- Stewart, T.J., Williams, B.M., 2019. Nuanced activism: the matrix of resistance. In: Dache, A., Quaye, S.J., Linder, C., McGuire, K. (Eds.), *Rise Up!: Activism as Education*. Michigan State University Press, East Lansing, MI, pp. 201–223.
- Stokes, S., Miller, D., 2019. Remembering "the black bruises:" a case study of supporting student activists at UCLA. In: Morgan, D., Davis III, C.H.F. (Eds.), *Student Activism, Politics, and Campus Climate in Higher Education*. Routledge, New York, pp. 143–163.
- Stone, P., 2021. Money and misinformation: how turning point USA became a formidable Trump force. *The Guardian*. Available at. <https://www.theguardian.com/us-news/2021/oct/23/turning-point-rightwing-youth-group-critics-tactics>.
- Tsui, L., 2000. Effects of campus culture on students' critical thinking. *Rev. High. Educ.* 23 (4), 421–441. <https://doi.org/10.1353/rhe.2000.0020>.
- Vaccaro, A., Mena, J.A., 2011. It's not burnout, it's more: Queer college activists of color and mental health. *J. Gay Lesb. Ment. Health* 15 (4), 339–367. <https://doi.org/10.1080/19359705.2011.600656>.
- Van de Velde, C., 2021. "Different struggles, the same fight?" A comparative analysis of student movements in Chile (2011), Quebec (2012), and Hong Kong (2014). In: Bessant, J., Mesinas, A.M., Pickard, S. (Eds.), *When Students Protest: Universities in the Global North*. Rowman & Littlefield, pp. 33–50.
- Walter, R., 1969. The intellectual background of the 1918 university reform in Argentina. *Hisp. Am. Hist. Rev.* 49 (2), 233–253. <https://doi.org/10.1215/00182168-49.2.233>.
- Warnock, D., Hurst, A., 2016. The poor kids' table: organizing around an invisible and stigmatized identity in flux. *J. Divers. High. Educ.* 9 (3), 261–276. <https://doi.org/10.1037/dhe0000029>.
- Weipert, A., 2020. When workers councils defeated the far-right coup in Germany. *Jacobin Magazine*. <https://jacobinmag.com/2020/03/november-revolution-kapp-putsch-1920>.
- Williams, J., 2013. *Consuming Higher Education. Why Learning Can't Be Bought*. Bloomsbury.
- Williams, P.R., McGreevey, M.R., 2004. Student activism at Ithaca college: reflections on management and leadership. *N. Dir. High. Educ.* 128, 61–72.
- Yosso, T.J., Benavides López, C., 2010. Counterspaces in a hostile place. In: Patton, L.D. (Ed.), *Culture Centers in Higher Education: Perspectives on Identity, Theory, and Practice*. Stylus, Sterling, VA, pp. 83–104.

FOURTH EDITION

International Encyclopedia of Education

Ways of Knowing
and Indigenous Education

VOLUME
3

Editors-In-Chief

Robert J. Tierney

Fazal Rizvi

Kadriye Ercikan

Volume Editors

Graham Hingangaroa Smith

Huia Jahnke

Margaret Maaka



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 3

Ways of Knowing and Indigenous Education

VOLUME 3 EDITORS

**Graham Hingangaroa Smith,
Huia Jahnke and Margaret Maaka**



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

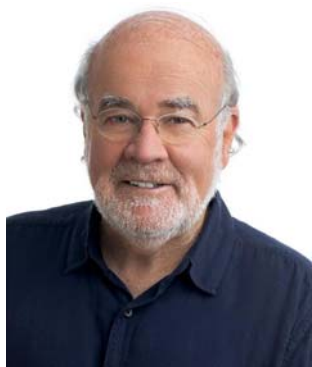
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence–based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



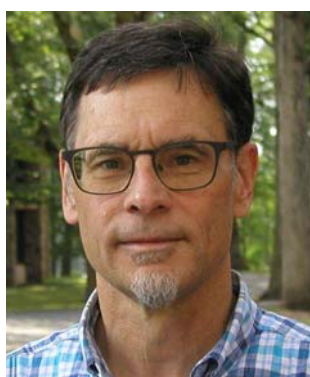
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 3: Ways of Knowing and Indigenous Education

Graham Hingangaroa Smith, Huia Jahnke, and Margaret Maaka

© 2023 Elsevier Ltd. All rights reserved.

Introduction: Transforming Indigenous Education

This volume of the Elsevier *Encyclopedia of Education* examines issues related to the education and schooling of Indigenous peoples. In recent years, indigenous education has become a focus due to persisting inequities in and as result of the educational opportunities experienced by most Indigenous peoples compared with dominant colonizing populations. The phrase “indigenous education and schooling,” as employed here, refers, on the one hand, to the experiences of Indigenous students both within and through conventional schooling developed mostly by and in the interests of dominant settler populations. On the other hand, this term references forms of education and schooling developed by indigenous groups, usually around their well-being, relating to language, cultural knowledge, cultural interests, and aspirations. A significant factor is that indigenous groups often struggle to assert their cultural recognition and survival within contexts that try to assimilate them into the culture and learning expectations of the dominant population.

In this section, we employ the definition of “indigenous” as defined by four conditions that predate the arrival of settler populations with their imposing societal norms:

- Historical and genealogical connections
- Strong cultural links to territories and surrounding natural resources
- Distinctive social, economic, and/or political systems
- Distinct language, culture, values, and beliefs

The reference “indigenous struggle” connotes indigenous resistance to the erosion of the four conditions named above. The destructive forces that eroded these conditions are called “colonization.” The term “struggle” describes indigenous resistance to these colonizing forces that are given impetus by unequal power and social relations. Settler populations have often employed various colonizing strategies thinking they were acting in the best interests of indigenous communities and that they were being helpful by assimilating the indigenous population as quickly as possible into the “mainstream” settler language, knowledge, and culture. However, in too many cases, the consequences of this strategy caused irreparable harm, such as the loss of indigenous languages and knowledge systems, including cultural beliefs, customs, and laws. Colonization has had other harmful consequences. Among them, are the forced and unjust treatment to deny Indigenous peoples their cultural rights through colonial laws and policies; economic exploitation through the loss of traditional resources such as forests and lands as a result of war, occupation, legal manipulation, and individualizing collective rights; religious persecution through the banning of traditional practices and instituting religious practices, ideals, and values associated with Judeo-Christianity; and differential access to democracy through the denial of voting rights. These strategies and many others served to reproduce settler interests. Colonization processes have been developed overtly and covertly by, and in the interests of, dominant settler populations, usually through a nexus of dominant power, culture, and economic interests. For Indigenous peoples, even with “enlightened” research and scholarship, the depressing news is that despite identifying and challenging these forms of colonization, colonization has not gone away. More often, colonization has simply changed shape and is being developed in new contemporary formations. In the present day, many of these new shapes of colonization are formed and legitimated within the neoliberal ideology of free-market economics, for example, where cultural oppression links to economic exploitation and where possessive individualism supersedes considerations of collective responsibility.

New forms of colonization are developed in multiple shapes, sites, and sizes comes to mind. The image of the Dutch boy standing in front of the leaking dike faced with the dilemma of numerous leaks of different shapes and sizes. He needs all sorts of stoppers to stem the leaking water, not just his fingers. The analogy here is that multiple forms of colonization require multiple “fit-for-purpose” responses. Our struggle in response to colonization is not a single struggle. This is the point made by Gramsci (1971) in his discussion of the concepts of “war of position” and “war of maneuver” in the form of struggles over ideas and beliefs. Thus, indigenous counternarratives that advance alternatives to dominant ideas of what is normal and legitimate need to be similarly positioned. They need to be different, nuanced, and fit for purpose. To reiterate, the indigenous struggle is not one struggle but many struggles across many different sites, shapes, and sizes.

The reason to position a volume of the Elsevier *Encyclopedia of Education* on the indigenous struggle over education and schooling is because

We will not have a sustainable social and economic revolution against the persisting inequities that accrue to indigenous populations without a prior or simultaneous education and learning revolution.

Smith (1997).

In this volume, we have assembled the work of several indigenous educators who contribute insights into their struggles within their own jurisdictions. The important point is that not all Indigenous people are in the same place regarding their struggle against colonization. Therefore, it is important not to make comparative judgments that create hierarchical value judgments about where different indigenous communities are at with their struggle. Instead, these contributions have been assembled simply to share and learn ways by which we might be able to improve our own education and schooling aspirations. We need to counter the colonizing imperative of divide and rule, where dominant interest groups often use a comparative strategy that uses one indigenous group to denigrate another. Comments such as, “Why can’t you be like ‘them’? They have high levels of success with their students.” or “What is wrong with your native community?” As an editing group, we understand that each of our indigenous contexts is very diverse, and our experiences of colonization are also different. This indigenous volume allows us to share our stories and to put all of our knowledge contributions into a central pool so that each of us can take what we need to reflect on and if appropriate, enhance our own contexts.

Transforming Indigenous Education

A key theme in this volume is transforming indigenous education. We deliberately use the term “transforming” instead of “transformation.” This is because transforming calls us to an ongoing enactment, whereas transformation suggests that we have arrived at some solution. The point here is that the indigenous struggle must be a continuous improvement cycle. [Smith \(1997\)](#) refers to this as a cumulative cycle of praxis. For most Indigenous peoples, the status quo is not working very well, and transforming is an important emphasis. Moreover, the need to focus on indigenous experience in and through education and schooling is driven by two imperatives. Firstly, there is the persistent problem of social, economic, and cultural inequalities suffered by indigenous communities. Secondly, the relatively recent focus that has been brought to bear on these issues by the release of the United Nations Declaration on the Rights of Indigenous Peoples (2007). Beyond this, significant issues focus on Education within different indigenous contexts. For example, the Truth and Reconciliation Commission inquiry into the Residential Schooling system in Canada and the subsequent list of 94 Calls to Action that impact the education sector; the major reform of Curriculum provision in Aotearoa New Zealand Education, and its focus on Māori learning success; also in Aotearoa New Zealand the new Government emphasis on the provision of Immersion Māori language schooling; the Education Compact discussions between the US Federal Government (DoE) and Native Alaskans which have resulted in five tribal state Schools to be established.

The phrase “transforming indigenous education” is ambiguous because it embodies two meanings. Both are important. The first meaning relates to the need to transform the process of how we educate and school indigenous learners. This calls attention to the need to scrutinize the material to be taught, how it is taught, and in whose interests the material has been selected. The second meaning relates to the outcomes of indigenous education and schooling, which also need to be transforming in order to provide schooling outcomes to realise success for indigenous students and to thereby make a difference.

The other connotation of the phrase “transforming indigenous education” emphasizes the need for indigenous education and schooling outcomes to realize successes for Indigenous students who also make a difference. Therefore, there is a need to focus on transforming both processes and outcomes to overcome using the same old strategies and practices that have been largely ineffective and are not making a significant difference in the lives of indigenous learners.

We bring together an eclectic array of contributions that are drawn from a range of indigenous jurisdictions. These describe multiple sites of struggle to overcome the constraints of dominant colonial education and schooling, which have historically reproduced the interests of the dominant settler populations and, simultaneously, the underdevelopment of the colonized indigenous peoples. There is an ongoing need to rearticulate the shape of the collective struggle as Indigenous peoples and a need to renew the “struggle” to maintain the impetus and effectiveness of the self-development revolution. Struggle in this sense is to be seen positively. The struggle is also important as it can help us clarify what it is that indigenous communities stand for as well as what it is that they stand against. The fact that most Indigenous people remain socially and economically marginalized is evidence that selected policy formation has limited effect. This reinforces the point that no silver bullet, magic pill, or single fix-it policy exists. Instead, there is a need to invest in change on a broad range of fronts.

It is also critical to recognize that, as [Freire \(1972\)](#) noted, we need to free ourselves before we can free others. This is an important point about hegemony and colonizing ourselves by uncritically believing in and taking aboard as common sense, the colonizing rhetoric of the settler population. Transforming indigenous experience rests not just on the individual but also on families, on communities, on tribes, and on Indigenous peoples collectively. An important factor in transforming is understanding how individual conscience can join with the collective consciousness of others.

In this sense, the indigenous struggle is also for our minds. We need to understand our participation in forming our own subordination, exploitation, and oppression. We must educate ourselves about false consciousness and free our minds from hegemony. Education has been a significant factor in embedding indigenous inferiority. We must reclaim the power of education to act in our interests. An important decolonizing act, therefore, is to struggle over the meaning and intention of education and schooling. It needs to serve all peoples and not simply be a means to reproduce dominant cultural expectations at the expense of indigenous interests.

References

- Freire, P., 1972. *Pedagogy of the Oppressed*. Penguin, Harmondsworth.
- Gramsci, A., 1971. *Selections From the Prison Notebooks*. International Publishers, New York.
- Smith, G.H., 1997. *The Development of Kaupapa Māori: Theory and Praxis*. The University of Auckland. Unpublished PhD Thesis.
- United Nations, 2007. Declaration on the Rights of Indigenous Peoples (UNDRIP). <https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html>.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 3

Emily Anson

*McMaster University, Hamilton, ON, Canada; and
Ohneganos Ohnegahdę:gyo, Hamilton, ON, Canada*

Jo-ann Archibald Q'um Q'um Xiem

University of British Columbia, Vancouver, BC, Canada

Sean Asikłuk Topkok

*Center for Cross-Cultural Studies, Indigenous Studies
Graduate Programs, Fairbanks, Alaska*

Katelyn Barney

The University of Queensland, Brisbane, QLD, Australia

Julie Blair

*Ontario Institute for Studies in Education, Toronto, ON,
Canada*

Abraham Bradfield

The University of Queensland, Brisbane, QLD, Australia

Tracey Bunda

The University of Queensland, Brisbane, QLD, Australia

Fred Dervin

University of Helsinki, Helsinki, Finland

Gabriel B Egbe

*Department of English and Literary Studies, Veritas
University, Abuja, Nigeria*

Bronwyn Fredericks

The University of Queensland, Brisbane, QLD, Australia

Danielle Gendron

*University of British Columbia, Vancouver, BC,
Canada; and Ohneganos Ohnegahdę:gyo, Hamilton,
ON, Canada*

Colin M Gibson

*University of Guelph, Guelph, ON, Canada; and
Ohneganos Ohnegahdę:gyo, Hamilton, ON, Canada*

Huia T Jahnke

*Te Pūtahi a Toi School of Māori Studies at Massey
University, Palmerston North, New Zealand*

Derek Jennings

*University of Washington, School of Public Health,
Seattle, WA, United States; and Bezruchka Family
Endowed Professor School of Public Health University of
Washington*

Theresa Arevgaq John

*Center for Cross-Cultural Studies, Indigenous Studies
Graduate Programs, Fairbanks, Alaska*

Michelle Johnson-Jennings

*Division of Environmental Health and Land-based
Healing, IWRI, University of Washington, School of
Social Work, School of Public Health, Seattle, WA,
United States; University of Washington, School of
Public Health, Seattle, WA, United States; and Full
Professor, Director of Indigenous Environmental Health
and Land-based Healing Division, Indigenous Wellness
Research Institute/IWRI, University of Washington,
School of Public Health and Social Work*

Patricia MG Johnston-Ak

*Te Whare Wananga o Awanuiaranig, Ngai te Rangi,
Ngati Ranginui, New Zealand*

Katrina-Ann R Kapā'anaokalāokeola Nākoa Oliveira

University of Hawai'i at Mānoa, Honolulu, Hawai'i

Michael Koskey

*Center for Cross-Cultural Studies, Indigenous Studies
Graduate Programs, Fairbanks, Alaska*

Tiffany S Lee

*Native American Studies, University of New Mexico,
Albuquerque, NM, United States*

Daniel Lipe

*Cal Poly Humboldt University, Environmental Science
Management, Arcata, CA, USA*

Makasa Looking Horse

*McMaster University, Hamilton, ON, Canada; and
Ohneganos Ohnegahdę:gyo, Hamilton, ON,
Canada*

Nēpia Mahuika
University of Waikato, Aotearoa, New Zealand

Dawn Martin-Hill
*McMaster University, Hamilton, ON, Canada; and
Ohneganos Ohnegahde:gyo, Hamilton, ON, Canada*

Summer Puanani Maunakea
*Curriculum Studies Department, University of Hawai'i
at Mānoa, Hawai'i, USA*

James McKenzie
*College of Education, University of Arizona, Tucson, AZ,
United States*

Kahontiyoha Cynthia Denise McQueen
*McMaster University, Hamilton, ON, Canada; and
Ohneganos Ohnegahde:gyo, Hamilton, ON, Canada*

Bob Morgan
University of Newcastle, Callaghan, NSW, Australia

Sereana Naepi
*Waipapa Taumata Rau, The University of Auckland,
Auckland, New Zealand*

Ryan Neepin
*Ontario Institute for Studies in Education, Toronto, ON,
Canada*

Sam L No'eau Warner
University of Hawai'i at Mānoa, Honolulu, Hawai'i

Kekailoa Perry
*College of Education—Educational Foundations,
University of Hawai'i at Mānoa, Honolulu, HI, United
States*

Michelle Pidgeon
*Faculty of Education, Simon Fraser University, Surrey,
BC, Canada*

Ngāti Porou
*University of Waikato, Aotearoa,
New Zealand*

Jean-Paul Restoule
*Department of Indigenous Education, Faculty of
Education, University of Victoria, Victoria, BC, Canada*

Martyn Reynolds
*School of Education, Victoria University of Wellington,
Wellington, New Zealand*

Kabini Sanga
*School of Education, Victoria University of Wellington,
Wellington, New Zealand*

Hinekura Lisa Smith
*Ngā Wai a Te Tūi, Māori and Indigenous Research
Center, UNITEC, Auckland, New Zealand*

Sandra Styres
*Ontario Institute for Studies in Education, Toronto, ON,
Canada*

Tehahenteh
*Indigenous Knowledge Exchange; and Ohneganos
Ohnegahde:gyo, Hamilton, ON, Canada*

Aleesha Towns-Bain (Alutiiq)
*Bristol Bay Native Corporation Education Foundation,
Anchorage, AK, United States*

Te Rina Warren
*Te Pūtahi a Toi School of Māori Studies at Massey
University, Palmerston North, New Zealand*

Tanya Wendt Samu
*School of Critical Studies, Faculty of Education and
Social Work, University of Auckland, Auckland, New
Zealand*

Maria Shaa Tláa Williams (Tlingit)
*University of Alaska Anchorage, Anchorage, AK, United
States*

Sweeney Windchief
*Department of Education, Montana State University,
Bozeman, MT, United States*

K Laiana Wong
University of Hawai'i at Mānoa, Honolulu, Hawai'i

Mei Yuan
Minzu University of China, Beijing, China

CONTENTS OF VOLUME 3

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 3: Ways of Knowing and Indigenous Education</i>	<i>xix</i>
<i>Contributors to Volume 3</i>	<i>xxiii</i>

HISTORICAL TRAUMA

"The mis-education" of Aotearoa: the struggle for Indigenous history in New Zealand <i>Nēpia Mahuika and Ngāti Porou</i>	1
Toward a rights-based and culturally responsive model of Aboriginal Education <i>Bob Morgan</i>	7
Impact of the 1971 Alaska Native Claims Settlement Act on education: 1971–2021 <i>Maria Shaa Tláa Williams (Tlingit) and Aleesha Towns-Bain (Alutiiq)</i>	22
Weaving together mind, body and spirit: envisioning pathways to trauma-informed classrooms from Indigenous perspectives <i>Sandra Styres, Ryan Neepin, and Julie Blair</i>	32

LANGUAGE REVITALIZATION

Ke Ala o ka No'eau: an infusion of Hawaiian into the Hawaiian language <i>K Laiana Wong, Katrina-Ann R Kapā'anaokaIāokeola Nākoa Oliveira, and Sam L No'eau Warner</i>	42
Indigenous language revitalization in the United States and Canada <i>Tiffany S Lee and James McKenzie</i>	50
Englishization and the language of education in Nigeria: literacy issues, challenges and hopes <i>Gabriel B Egbe</i>	61

HIGHER EDUCATION

- The salient truthfulness of Aloha 'Āina Theory and Guerrilla Praxis 69
Kekailoa Perry
- A public good? The commodification of Māori language and culture in higher education 81
Huia T Jahnke and Te Rina Warren
- Can we transform the academy? Indigenous-centered programming at the Department of
Indigenous Education at University of Victoria 89
Jean-Paul Restoule
- Being in good relations: higher education's responsibilities to Indigenous resurgence 99
Michelle Pidgeon
- The importance of "seeing oneself" in Australian universities: representations and reflections
of Indigeneity in higher education 109
Bronwyn Fredericks, Tracey Bunda, Katelyn Barney, and Abraham Bradfield
- Capturing storms: Pacific women's approaches to change in New Zealand universities 119
Sereana Naepi
- Is there a Dr in the house: building an indigenous doctoral program 129
Patricia MG Johnston-Ak
- Teaching graduate courses; indigenous students, course co-construction and bicultural accountability 136
Sweeney Windchief
- Developing transforming research—The University of Alaska Fairbanks Center for
Cross-Cultural Studies as case study 141
Sean Asikhuk Topkok, Michael Koskey, and Theresa Arevgaq John

LEADERSHIP

- Journeying to the common good: exploring Indigenous Solomon Islands leadership
through five moments 151
Kabini Sanga and Martyn Reynolds
- New frontiers in Pacific education, leadership and decolonized practice within Oceania: an initial
exploration 159
Tanya Wendt Samu

WAYS OF KNOWING

- Decolonizing psychology: weaving thrivance to disrupt colonial eugenics and shift
educational practices within Indigenous psychology 171
Michelle Johnson-Jennings and Derek Jennings
- Re-weaving thinking, re-thinking weaving: theorizing Indigenous arts practice as research
methodology 181
Hinekura Lisa Smith
- Indigenous storywork for education and research 189
Jo-ann Archibald Q'um Q'um Xi'em
- Seeing through the smoke: creating space for Indigenous Knowledge in science education 198
Daniel Lipe

<i>Kayaní:yo</i> (a good path)—Warrior Science Ohneganos—Indigenous ecological knowledge <i>Dawn Martin-Hill, Colin M Gibson, Makasa Looking Horse, Danielle Gendron, Emily Anson,</i> <i>Kahontiyoha Cynthia Denise McQueen, and Tehahenteh</i>	207
“Indigenous” ways of knowing in Chinese Minzu education: teachers reflecting on their engagement with STEM in inner Mongolia <i>Fred Dervin and Mei Yuan</i>	217
‘Āina-based education: Huli Ka Lima I Lalo Transformative praxis <i>Summer Puanani Maunakea</i>	228

“The mis-education” of Aotearoa: the struggle for Indigenous history in New Zealand

Nēpia Mahuika and Ngāti Porou, University of Waikato, Aotearoa, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
The indigenous past was never the colonizers pre-history	1
The myths we lived by	2
The history curriculum reset 2023	4
Teaching Māori and iwi histories—a struggle without end	5
References	6

Introduction

Oppressed peoples around the world are acutely aware of the power that history education has in shaping their cultural and political identities. In his influential text, *The Mis-Education of the Negro* (1933), for instance, the renowned African American Historian Carter G. Woodson wrote of the indoctrination of Black communities led to believe they were inferior in every class they entered and every book they read (p. 2). Mis-education, for him, was much more than a contemporary critique of the system but a cycle that results from the mis-educated continuing to teach and mis-educate others. For Indigenous Māori in Aotearoa New Zealand, this deliberate mis-education—what colonizers called “civilizing missions”, “nation building”, and “assimilation”—disconnected generations from their own history, identity and language. Like other colonized and indoctrinated groups, Māori could not see themselves or their histories in schools and textbooks, and when they were included, it was through racist stereotypes and patronizing colonial lenses that dismissed their past as prehistory and superstition.

The fight to have Māori histories taught in schools and taken seriously in the field has been a long and enduring struggle. For more than a century, Māori have sought to reclaim their past “site by site under Western eyes” (Smith, 1999, 34) from the hegemony and racism of an ongoing colonizer education system and the control and cannibalism of Western-centric academic disciplines and their gatekeepers. Since the arrival of Europeans, Māori have consistently embraced and experimented with new ideas and approaches to history but have always maintained their own pedagogies and key concepts about the nature, form, practice and ethics of historical knowledge transmission and research. Now in the 21st century Māori are beginning to see small incremental changes, albeit carefully controlled by those fearful of losing power. Nevertheless, together with non-Māori allies there is a developing public discussion about the need to address our shared and unresolved colonial trauma by teaching all New Zealanders the Indigenous histories of these lands, seascapes and peoples—to remind them where they are and that the Indigenous owners of these territories are still here. In 2019, the New Zealand government after years of petitions and protests finally announced plans to reset the National History Curriculum (yrs 1–10) for 2023. It was met with a mixture of cautious optimism and the expected racist backlash, but the struggle to reach this moment has taken generations of activism, conversations, contests, and a major rethinking of what history is, who it belongs to, how to research and teach it, and how it might yet heal our present and future.

The indigenous past was never the colonizers pre-history

History did not arrive in Aotearoa New Zealand with Europeans. The hundreds of years of prior Indigenous experiences are legitimate histories, and not the prelude or prehistory of a settler nation-making story. For centuries Māori taught their own history at home, out gardening, at funerals, while fishing, and in schools of learning called wānanga. These schools often held an explicit disciplinary focus on strategic warfare, agriculture, medicine, carving, and genealogy, and all transmitted local history and cultural politics. Through each generation, wānanga carried the “historical curricula of the tribe” and remains a practice in higher learning today across the country. Wānanga were more than simply curricula or content, but have been described as “a meditative practice, a pedagogy and discussion that encourages debate, dissent and critical analysis” (Mahuika and Mahuika, 2020). Māori kept wānanga as a practice and form of instruction, and it was a subject of some interest among early European ethnographers. Pākehā author Stephenson Percy Smith, for instance, wrote two volumes of *The Lore of the Whare Wānanga* in 1915 which became “the popular and state approved history of Pre-European Māori and Moriori” in the early 20th century (Anderson, 2014, 52). Smith’s Euro-centric interpretation of Māori history as “lore”, however, did not stop most non-Māori from dismissing local histories as merely myths and the puerile imaginings of the natives. Nevertheless, wānanga survived the establishment of Native Schools and the advent of the national public education system in New Zealand and is still practiced today by both Māori and now non-Māori. Indeed, most present-day tertiary institutions in Aotearoa commonly use the name wānanga in their titles. It is important to note that history was not only taught in wānanga, but transmitted within families in multiple ways through songs, stories, proverbs, and various artforms. Māori maintained an oral history where memorization was key to the learning of whakapapa (genealogy), the

stories of ancestors, practices, places and key turning points for each tribe. The past was woven, carved, performed, practiced, recited and lived by a chorus of storytellers who placed a high value on a persons’ ability to remember and narrate the collective experience of the tribe. To be a historian—a keeper, storyteller, and educator of the past—was a crucial and tapu role in Māori society held by women, men, weavers, singers and experts of various ages and backgrounds.

Māori schooling and history teaching looked very different from the European educational institutions that arrived with Pākehā in the early 1800s. Mātauranga Māori (Māori knowledge), and historical knowledge especially, were taught through games and hui (meetings), usually by the students’ whānau and hapū (family and extended kin). Children were taught culture, customs and history using key concepts in the Māori world such as tikanga, whakapapa and whānaungatanga (Hemara, 2000). In 2015, a group commissioned by the Ministry of Education wrote briefly about some of these concepts in *Te Takanga o te Wā* (Tamua, 2015) a booklet meant to assist teachers with an introduction to teaching Māori history.

When missionaries and books arrived in the early 19th century, Māori were highly receptive to reading and writing, and this new technique of remembering and collecting historical information developed, with many families keeping genealogy books and often treating them as sacred objects and treasures. Thus, the teaching of history in Aotearoa changed with the arrival of Europeans, and particularly Christianity, but it did not eradicate past practices and beliefs attached to the keeping of the past. Instead, Māori sought to “renegotiate” (Kaa, 2020) their mātauranga and embraced new ways to transmit history, enhancing their ability to share knowledge on their own terms.

Mission schools set about “civilizing” the local Indigenous population, hoping to facilitate the “Europeanizing of Māori” and their history (Simon and Smith, 2001, 3). The first mission school in New Zealand was established by Thomas Kendall at Rangihoua in 1816. Other schools soon followed, and during that period Māori embraced literacy and were avid readers of the Bible (McRae, 1997). Many of those who had attended mission schools returned to their villages and established kura (schools) in their own communities (Barrington and Beaglehole, 1974). Mission schools were eventually replaced by a Native Schooling system introduced by the Native Schools Act in 1867. According to James Belich the official policy toward Māori during this period (1814–1860) “was to convert them into Brown Britons” (cited in Simon and Smith, 1998, ix). By the 1860s, violent British colonial invasions in the Waikato region and wide-spread raupatu (land confiscations) emptied out mission schools around the country and left the government with very low educational expectations for Māori (not that they held high expectations at the time). In an 1862 report, Māori were viewed as better suited to “manual” rather than “mental labor” (Taylor Report 1862), and while the Education Act 1877 made provision for the first national curriculum, Section 81 (4) of the Act specified that “if history was taught, only English history was to be transmitted.” This remained the norm for history teaching up until 1904 when a new primary schools syllabus added New Zealand history content that was often ignored by teachers anyway (Manning, 2017, 121).

The myths we lived by

From the 1880s, newly appointed Native School teachers were expected to be “familiar with the discovery of New Zealand” and “the history of the land wars” which many taught as a triumphant and romantic account inconsistent with Māori histories of the same conflicts (Barrington & Beaglehole, 7; Manning, 121). This settler myth of peaceful settlement endured for so long it was alarmingly regurgitated by National prime-minister John Key in the early 2000s. The erasure, retelling, and mis-education of Māori and New Zealand history impacted all New Zealanders since the 19th century who were led to believe a range of myths about Māori history and the country’s colonial legacy. One longstanding myth described Māori navigation to Aotearoa as accidental, claiming Māori ancestors drifted haplessly and made landfall emaciated and starving. This was challenged decades later by scholars like Ranginui Walker who noted how the maritime achievements of Polynesian peoples in the Pacific had been “denigrated and romanticized” for too long (Walker, 25–26). A number of studies in the second half of the 20th century clearly demonstrated that traditional double hulled vessels were not only expertly sailed by Polynesian seafarers but that these skilled navigators traveled back and forth across the Pacific in multiple voyages.

Māori were also said to have arrived in one great fleet, a view popularized by Stephenson Percy Smith at the end of the 19th century. In one school text from 1960 it was suggested that Kupe discovered New Zealand about A.D. 950 and was followed by later explorers like Toi and a major migration about 1350 (Anderson, 52). The great fleet myth survived for some time until it was disproved by David Simmons extensive research in the 1960s and widely rejected by historians (King, 2004, 45). Yet another myth claimed that Māori were not the indigenous peoples of New Zealand at all and were in-fact responsible for the genocide of an earlier group, Moriori. This narrative was published in school journals in the early 20th century by Percy Smith and Elsdon Best and made its way into a 1934 school reader, *The Coming of the Māori to Ao-tea-roa*, published by A. W. Reed. It was later refuted by numerous scholars including H. D. Skinner, Roger Duff, and Michael King. King writes that “despite this overwhelming volume and weight of testimony some late 19th and early 20th century scholars convinced themselves and a gullible New Zealand public that Moriori were not Polynesian” (King, 2004, 55–56).

The undermining of Māori indigeneity and history has taken various mis-educated forms, and perhaps none more subtle than the reduction of the Māori past to prehistory. A range of history texts, including Keith Sinclair’s mid-20th century book, *A Destiny Apart* (1986), chaptered pre-European contact experiences as “Prehistory”, a precursor to the more important story of the nation. Māori have strongly rejected that idea as have most New Zealand historians today (King, 2001, p. 7). This mis-education of basic New Zealand history was notable in recent debates regarding the TUIA250 celebrations in 2015 (<https://mch.govt.nz/tuia250>),

a commemoration to the country’s “discovery”, that found its’ date in the arrivals of European explorers James Cook and Abel Tasman, unsurprisingly at the expense of earlier Māori navigators and discoverers.

The reduction of the Māori past to prehistory reflected the underlying tenets of a discipline grounded in Western historical methods and concepts. The professional History that arrived in New Zealand during the 19th century carried with it specific standards developed by Leopold Von Ranke and other historians between 1824 and 1871. Legitimate history, according to these standards must be based in verifiable objective empirical sources that rejected Indigenous oral histories as untrustworthy. This focus on narrow empirical evidence remained dominant in history teaching for the entire 20th century, despite arguments from scholars such as Jan Vansina, who argued for the viability of oral traditions as history (Vansina, 1985). Despite his optimism, Vansina maintained that oral histories should still be verified by Western historical methodology. But for Māori historians and scholars, like Ruka Broughton, tribal oral history and traditions were robust and contained “much that can be regarded as factual whether biological history” or genealogical (Broughton, 1979, 6).

Māori consistently challenged the displacement of their past throughout the 20th and early 21st centuries, with iwi and Māori activists, leaders, and historians opposing the insidious nature of this colonial mis-education (Mahuika, 2009). Writing on this issue to his friend Apirana Ngata in 1931, Māori leader Te Rangihira (Peter Buck), for instance, insisted:

Kua mutu te wā ki a Te Peehi ma, kua riro ma tāua, ma te Māori, tāua kōrero / The time for Best is over, we as Māori should take responsibility for researching our world for ourselves. It is left to us to straighten up what has been written by our Pākehā pioneers.

Cited in Sorrenson (1987).

Increasingly, Māori and non-Māori scholars challenged not only the content of popularized colonizer histories, but the underlying tenets of the discipline that delineated the field in explicitly non-Māori ways. Like Ngata and Buck, Hirini Moko Mead, several decades later in 1977, pointed out how Pākehā had for too long been “reaching into Māori culture and pulling out the features with which they can identify, taking hold of quite generous portions which they then try to fit into a Pākehā cultural world” (Mead, 1977). Māori organizations, like the Māori Women’s Welfare League, also challenged Government policy regarding the place of te reo and the teaching of Māori history in schools during the 20th century (Harris, 2004, 44). This activism, connected to protests of the 1960s and 1970s, disrupted the idea of New Zealand’s world leading race-relations mythology and reframed discussions around significant histories such as the Treaty of Waitangi and the 19th century Land Wars.

In reclaiming the past, Māori were keen to correct historical myths that perpetuated colonizer-centric stories, like the idea that Māori simply ceded sovereignty to the Crown and welcomed their own assimilation (Jackson, 1998, 73). This also meant challenging the power and place of Pākehā historians, like Michael King, who had set themselves up as experts on Māori history much the same as Elsdon Best and Percy Smith had decades earlier (Mahuika, 2011). This ownership of the past remained an ongoing discussion, with tribal scholars like Keri Kaa arguing that “for years we have provided academic ethnic fodder for research and researchers. Perhaps it is time we set things straight by getting down to the enormous task of writing about ourselves” (cited in King, 1999, 184). This led to questions about who can do Māori history? On this issue of ownership in the late nineteen eighties, Ngai Tahu leader, Sir Tipene O’Regan argued that “the past belongs to all New Zealanders, but first it is ours” (O’Regan, 1987). For Māori and iwi people, it was always a matter of ethics. Māori have pointed out that non-Indigenous peoples can do this work, but not by themselves; and would need to position themselves as non-Māori and be accountable to local communities (Mahuika, 2015). For many Māori, non-Indigenous researchers who built careers off the back of native peoples did more damage than good, despite self-serving arguments that their work rescued Native history from extinction. This question of ownership and access remains an ethical problem, evident in the critiques of Pākehā historian Paul Moon by Māori scholars today who have pointed out his Eurocentric mis-interpretations of their history and lack of appropriate Indigenous research ethics (Tau, 2009; Karere, 2018).

In 1992, Māori formed a national collective only for Māori and iwi historians called Te Pouhere Kōrero. It serves as a space for Māori to discuss the past away from Pākehā dominated forums like the New Zealand Historical Association (NZHA). For the last thirty years, several historians within this cohort have produced important work about the practice and nature of Māori history. These included important insights about the nature, form, ethics and practice of Māori and iwi histories. One of the ongoing issues has been the reduction of Māori history to overly-simplistic generalizations of Māori people. By the end of the 20th century, Māori scholars had adamantly argued that their histories should not be seen through a singular homogenous frame but are better understood from tribal perspectives. Tuhoe scholar, John Rangihau famously asserted, for instance, that “[m]y being Māori is absolutely dependent on my history as a Tuhoe person There are so many aspects about every tribal person. Each tribe has its own history. And it is not a history that can be shared among others.” (Rangihau, 1992, 190). This was reiterated by Tipene O Regan, who argued that the “psychological base of a Māori view of the past is both personal and tribal. It is intimately related to one’s notion of self” (O’Regan, 142).

Māori historians have consistently argued that their past should not be taught as merely colonial history, ignoring hundreds of years of nuanced historical experiences beyond the binaries of the colonized and the colonizer. Tribal scholars, for instance, have debunked the notion that 19th century Māori can be defined as either loyalists or rebels. Loyalist or “Kūpapa” were terms imposed on Māori historical subjects by non-Māori historians, effectively locking them into colonial nation-making binaries. In his biography of Rapata Wahawaha (Soutar, 2000), Ngāti Porou historian Monty Soutar convincingly showed that Wahawaha was by no means a loyalist and had been wrongly identified as Kūpapa. Likewise, in a recent study (2020), Denize Ewe also showed

that Ngāti Tipa leader, Waata Kukutai was inappropriately considered a loyalist Kūpapa as well, when like Wahawaha, Kukutai made decisions for the benefit of his people, not the Crown. This reclaiming by Māori historians disrupts longstanding myths that portrayed Māori resistors as violent rebels, blood thirsty warriors, and in more recent times as “wreckers and haters” or “terrorists” while colonial invaders were cast as the nation’s discoverers, settlers, pioneers, leaders, and citizens.

In challenging the mis-education of history in Aotearoa, Māori have tirelessly asserted the value and crucial importance of local Indigenous knowledge or “mātauranga Māori” (Ruru and Nikora, 2021). Mātauranga Māori, according to Rawiri Te Maire Tau “is simply the epistemology of Māori—what it is that gives point and meaning to Māori knowledge”. Tau questioned whether the Western dominated discipline of history is the right place for that knowledge, as it has often “twisted” mātauranga beyond anything our experts ever anticipated (Tau, 2001, 67–68). This is a common theme for Māori activists and scholars, with explicit concerns about the so-called rigors of empirical science which allowed Māori history and identity “to become the plaything of historiographers” (O’Regan, 141).

In more recent years, Māori have argued that those who research and teach history in Aotearoa must develop a stronger understanding of the way Māori conceptualize and practice the past. Māori, for instance, see viable historical sources as much more than rigidly empirical written documents. Objectivity in research is also not the driving goal for Māori historians or researchers who highly value the nuance and ethics in subjectivity. Like oral historians across the world, Māori have argued for the value and validity of oral sources, and indeed multi-sensory sources that are used to story the past (Mahuika, 2009, 2017). For Māori, genealogy or whakapapa is a key component in the way they conceptualize and structure history. Thus, while mātauranga Māori is seen by some as the appropriate place for history, whakapapa is seen by many as “the skeletal structure to Māori epistemology” (Te Maire Tau, 67–68). When non-Māori researchers arrived in Aotearoa, they gathered genealogical data, much of it through the Native Land Court system, and this became the basis of a number of written tribal histories. W. E. Gudgeon, a Native Land Court Judge and part time historian, for example, advocated arranging tribal histories and genealogies into a single history of Māori (Anderson, 52), while Sir Apirana Ngata experimented with the use of Western genealogical methods in the teaching of East Coast histories (Mahuika and Mahuika, 2020).

Like other Indigenous peoples, Māori also see time in a different way to the normative linear approaches introduced and normalized in schools and Western history (Marker, 2011). Māori histories are not simply linear, but can be circular, and understood through different conceptions of wā (time) and maramataka (lunar calendars). Some have argued that “Māori culture” is more “event time” driven rather than clock time oriented (Lo and Houkamau, 2012, 109). Māori have also widely rejected the argument made by those who suggest good objective history should avoid a presentist outlook—and not be impacted by present day perspectives and ideas (Williams, 2014). For Māori, however, the past is regularly seen as part of the present and collapsed as the narrator “walks backward into the future” (Rangiwai, 2021) and carries with them their ancestors’ stories. History is not a foreign country for Māori. It is a living practice deeply connected to the now.

Māori have also been outspoken about the importance of te reo Māori (the Māori language) in teaching and researching their past. In the late nineties, Monty Soutar highlighted the crucial place of language proficiency in undertaking tribal histories and the appropriate analysis of Māori historical sources (Soutar, 1996). Māori historians have been clear that New Zealand historians cannot rely on English to define and make sense of Māori and iwi histories, and that educators must grow their knowledge of te reo Māori to correctly pronounce and interpret the country’s past (Mahuika, 2015). English, as many other Native scholars across the world have argued, is a deeply problematic medium to teach Indigenous histories (Battiste, 2008, 504).

In challenging the colonial mis-education of their past, Māori have continued to illustrate how they see, teach and practice history. This includes writing about Māori history concepts (Tamua, 2015), sources and ethics (Mahuika, 2015). While these are perhaps more the interests of researchers and historians, it has been vitally important for Māori to challenge the mis-education of New Zealand’s past taught to the general public and generations of school goers. Māori have been explicitly outspoken in repeating the view that New Zealand was not settled peacefully. For Māori, settlers were invaders or at the least un-settlers, and history here should not revolve around their story of nation-building that has relegated Māori history to myths, legends, and prehistory. In reclaiming the past from colonizing educators, Māori experts have been adamant that the first law in New Zealand was Māori law not European law (Mikaere, 2005). Similarly, the first oral historians in New Zealand were not 20th century European journalists but Māori historians (Mahuika, 2017). And perhaps most importantly, the Treaty of Waitangi and its colonial past is not the focal point of iwi histories that are more interested in the storied experiences of their own ancestors, lands, rivers, and oceans.

Māori historians have been forthright in their assertions that there is no New Zealand history without Māori, and that the New Zealand narrative is only a very small chapter in a much longer and still ongoing Indigenous story (Mahuika, 2015). Unlike dominant Pākehā histories that focused on inevitable national progression, a Paradise Reforged (Belich, 2001), and a Frontier of Dreams (Dalley and McClean, 2005), Māori historians have produced much different narratives that emphasize an ongoing struggle without end (Walker, 1990), protest (Harris, 2004), and their position as the first and still native peoples or tangata whenua (Anderson, 2014).

The history curriculum reset 2023

Since the introduction of a national syllabus in 1904, Māori and iwi histories have been excluded in favor of both British and New Zealand content. History was optional in the syllabus until 1915, and for much of the 20th century the focus remained on English and settler colonial perspectives. The syllabus was revisited again in 1928 and influenced by a special committee formed in 1924.

The 1928 syllabus devoted more space to New Zealand topics but did not guarantee the incorporation of Māori perspectives (Manning, 2017). When Māori were included in textbooks, their histories reflected many of the colonial myths that Māori had sought to critique and correct for generations. From 1944, history as a subject fell under the newly introduced and compulsory social studies curriculum that added yet another layer through which Māori history would need to traverse to be seen and valued. Despite a further curriculum update in the late eighties, scholars note that New Zealand history was largely ignored by educators who preferred teaching the much less controversial 16th- and 17th-century Tudor England topic right through to 2010 (Manning, 2017). When Labor Prime Minister Jacinda Ardern announced plans for an overhaul of the New Zealand History curriculum (yrs 1–10) in mid-September of 2019, there was considerable skepticism from Māori (Loader et al., 2019).

The reset had been driven by a 2014 petition for the compulsory inclusion of the 19th century New Zealand Wars and increasing calls for New Zealanders to better understand their own colonial history. While Māori had been arguing for this for a century, there was little support from Pākehā, with “risk averse bureaucrats” and politicians too conservative or afraid to push for meaningful change despite the major shifts in bicultural education in the mid-eighties (Manning, 2017). When Māori had called for inclusion decades earlier, Pākehā Historians bemoaned the political correctness of it all, and complained that such a move would erode the autonomy of school boards to teach what they wanted. Following the 2019 announcement, Pākehā Historians like Michael Bassett complained that the curriculum advisory group was made up of “a collection of Māori radicals” with “no serious Pakeha historian in sight” (Bassett, 2021). In truth, the majority of that team were non-Māori. This gate-keeping of history education in Aotearoa New Zealand has long been dominated by “a number of high-profile academics,” none of them Māori (Sheehan, 680). And despite the energy and work around the yr 1–10 junior history curriculum, that unabashed gate-keeping is still evident in the current senior history writing group (yrs 11–13) which is still devoid of Māori historians despite calls from Te Pouhere Kōrero to Minster Hipkins and New Zealand History Association president Michael Belgrave to include Māori experts.

The struggle to have Māori history taught in New Zealand schools is as much a fight within the education system as it has been within the field of history globally. Māori historians share this similar struggle with indigenous peoples in Canada, Australia, Central, North and South America, and many other parts of the Indigenous world. In the United States, Native scholars have recently asserted that you cannot teach US history without Native Americans and have insisted that when “teachers embed Native American history more fully in the American story, students are challenged to think in new ways about larger themes in American history such as nation building, economic empowerment, citizenship and multiculturalism” (Sleeper-Smith et al., 2015). Similar work in Canada, led by First Nations scholars like Marie Battiste are also pushing for the inclusion and development of their history in local curricula.

Since the 2014 petition, there have been small advances in the inclusion of Māori history and a desire to address and decolonize the nations unresolved trauma. In 2018, for instance, the New Zealand History Teachers Association decided to change their constitution to reflect a more explicit decolonial focus and supported calls for inclusion of the New Zealand Wars in schools. The 2014 petition, and years of struggle, also led to the first Rā maumahara (commemoration day) hosted by Te Tai Tokerau iwi at Kororāreka in March 2018. It took another year for the government to officially announce that inclusion in the school curriculum. The 2022 History curriculum reset, then, has a long backstory of struggle and protest, renewed in recent petitions and largely embraced by a population more aware of the need to finally address the mis-education of their own past and not leave it simply to Māori.

Teaching Māori and iwi histories—a struggle without end

Over the past two centuries, Māori histories were deliberately excluded or misrepresented in various iterations of the history curriculum in New Zealand schools. From early mission schools (1814–1860s) to the advent of the Native School system (1867–1969), boarding schools and public schools (1877–). At all levels, History curricula and teaching favored powerfully embedded English and colonial-centric New Zealand nation-making perspectives and agendas. Māori struggled through more than a century to be in the conversation today, and to put right generations of racist mis-education about their past. The struggle was never just a matter of returning Māori history to Māori peoples, but ensuring the native history of these lands, oceans, and peoples are taught correctly to all New Zealanders today. It is a history far deeper than New Zealand nationalism and is still ongoing. European experiences and nation-state history is an only very small part of a much longer narrative that explicitly tells all those in Aotearoa how their local areas were named, and what it means to belong here—what it means to truly be native.

Like Māori, all New Zealanders were mis-educated across the 20th and early 21st centuries in ways that made it much harder for Māori to claim their historical knowledge as valid and reliable. That mis-education also proved a significant barrier for Indigenous people who became disconnected from their culture, language and history because they were unable to see themselves beyond the stereotypes and myths taught in schools. This led to generations of depression, poor health, resentment, anger, and unresolved trauma.

There is still significant concern about a new curriculum for 2022 and whether it will simply progress the state objective to assimilate Māori history into the well-funded and ever-present overarching power narrative of national becoming and progress. Despite the inclusion of Māori perspectives and content, the curriculum is still wedded to the idea of nationhood and is still controlled by a government that continues to assimilate local Indigenous peoples within their own historical articulation of one people. The struggle to re-educate the general population and fight for the right to define themselves, therefore continues for Māori and now for all New Zealanders. It is, much as Ranginui Walker once observed in his famous history, *Ka Whawhai Tonu Matou*, a struggle seemingly without end.

References

- Anderson, A. (Ed.), 2014. *Tangata Whenua: An Illustrated History*. Bridget Williams, Wellington.
- Barrington, M., Beaglehole, T.H., 1974. Māori Schools in a Changing Society: An Historical Review. New Zealand Council for Educational Research, Wellington.
- Bassett, M., 2021. New Zealand's Modern Culture Cringe, 18 Jan. <https://www.bassettrashandhide.com/post/new-zealand-s-modern-cultural-crige>.
- Battiste, M., 2008. Research ethics for protecting Indigenous knowledge and heritage: institutional and researcher responsibilities. In: Denzin, N.K., Lincoln, Y.S., Smith, L.T. (Eds.), *Handbook of Critical and Indigenous Methodologies*. London.
- Belich, J., 2001. *Paradise Reforged: A History of the New Zealanders From the 1800s to the Year 2000*. Allen Lane; Penguin, Auckland.
- Broughton, R., 1979. *Ko Ngaa Paiaa o Ngaa Rauru Kiitahi/the Origins of Nga Rauru Kiitahi*. Victoria University, MA thesis.
- Dalley, B., McClean, G. (Eds.), 2005. *Frontier of Dreams: The Story of New Zealand*. Hodder Moa Beckett, Auckland.
- Ewe, D., 2020. *Waata Kukutai: Rangitira o Ngaati Tiipa 1822–1867*. University of Waikato, MA Thesis.
- Harris, A., 2004. *Hikoi: Forty Years of Māori Protest*. Huia Publishers, Wellington.
- Hemara, W., 2000. *Māori Pedagogies: A View From the Literature*. New Zealand Council for Educational Research, Wellington.
- Jackson, M., 1998. Research and colonisation of Māori knowledge. *He Pukenga Kōrero* 4 (1), 69–76.
- Kaa, H., 2020. *Te Haahi Mihinare: The Māori Anglican Church*. Bridget Williams, Wellington.
- Karere, T., 2018. Reo Experts Say Professor Paul Moon's Claims Māori Language is in Trouble is "Ludicrous" and "White Noise". Jan 16th. <https://www.1news.co.nz/2018/01/16/te-reo-experts-say-professor-paul-moons-claims-maori-language-is-in-trouble-are-ludicrous-and-white-noise/>.
- King, M., 1999. *Being Pākehā Now: Reflections and Recollections of a White Native*. Penguin, Auckland.
- King, M., 2001. *Ngā iwi o Te Motu: One Thousand Years of Māori History*. Reed, Auckland.
- King, M., 2004. *The Penguin History of New Zealand*. Penguin, Auckland.
- Lo, K., Houkamau, C., 2012. Exploring the cultural origins of differences in time orientation between European New Zealanders and Māori. *NZJHRM* 12 (3), 105–123.
- Loader, A., Keene, B., Mullany, M., Stevens, M.J., Williams, M.M., 2019. Forum: new Zealand history in schools. In: *Te Pouhere Kōrero* 9, Māori History, Māori Peoples, p. 99. Auckland.
- Mahuika, N., 2009. Migration and the nation: revitalizing Te Ika a Maui. *N. Z. J. Hist.* 43 (2), 133–149.
- Mahuika, N., 2011. Closing the gaps: from post-colonialism to Kaupapa Māori and Beyond. *N. Z. J. Hist.* 45 (1), 15–32.
- Mahuika, N., 2015. New Zealand history is Māori history: tikanga as the ethical foundations of historical scholarship in Aotearoa New Zealand. *N. Z. J. Hist.* 49 (1), 5–30.
- Mahuika, N., 2017. An outsiders guide to public oral history in New Zealand. *Publ. His. N. Z.* 5, 3–18.
- Mahuika, N., Mahuika, R., 2020. Wānanga as a research methodology. *Alternative* 16 (4), 369–377.
- Manning, R.F., 2017. The New Zealand (school curriculum) "History Wars": the New Zealand land wars petition and the status of Māori histories in New Zealand schools (1877–2016). *Aust. J. Indig. Educ.* 47 (2), 120–130.
- Marker, M., 2011. Teaching history from an Indigenous perspective: four winding paths up the mountain. In: Clark, P. (Ed.), *New Possibilities for the Past*. UBC Press, Vancouver.
- McRae, J., 1997. *Māori Use of Writing and Print to 1900, Book & Print in New Zealand: A Guide to Print Culture in Aotearoa*. Victoria University Press, Wellington.
- Mead, H.M., 1977. Māoritanga, should it be shared? *Listener*, 10th Dec, 24.
- Mikaere, A., 2005. The treaty of Waitangi and recognition of Tikanga Māori. In: Belgrave, M., Kawharu, M., Williams, D., Waitangi, D. (Eds.), *Revisited: Perspectives on the Treaty of Waitangi*, p. 330.
- O'Regan, T., 1987. Who owns the past? Change in Māori perceptions of the past. In: Wilson, J. (Ed.), *From the Beginning: The Archaeology of the Māori*. Auckland: Penguin in Association With New Zealand Historic Places Trust, pp. 141–145.
- Rangihau, J., 1992. Being Māori. In: King, M. (Ed.), *Te Ao Hurihuri: Aspects of Māoritanga*. Reed, Auckland, pp. 183–190.
- Rangiwai, B., 2021. Walking backwards into the future: prophecy as an approach for embedding Indigenous values in tertiary education. *Te Kaharoa* 17.
- Reed, A.W., 1934. *The Coming of the Māori to Ao-Tea-Roa*. Reed, Place.
- Ruru, J., Nikora, L.W., 2021. *Ngā Kete Mātauranga: Māori Scholars at the Research interface*. Otago University press, Dunedin.
- Simon, J., Smith, L.T., 1998. *The Native Schools' System 1867–1969—Ngā Kura Māori*. Auckland University Press, Auckland.
- Simon, J., Smith, L.T., 2001. *A Civilising Mission? Perceptions and Representations of the New Zealand Native Schools System*. Auckland University Press, Auckland.
- Sinclair, K., 1986. *A Destiny Apart: New Zealand's Search for National identity*. Allen and Unwin/Port Nicholson Press, Wellington.
- Sleeper-Smith, S., Barr, J., O'Brien, J.M., Shoemaker, N., Manning Stevens, S. (Eds.), 2015. *Why You Can't Teach United States History Without American Indians*. North Carolina University Press and Chapel Hill, North Carolina.
- Smith, S.P., 1915. *The Lore of the Whare Wānanga or Teaching S of the Māori on Religion, Cosmogony and History*. T. Avery, New Plymouth.
- Smith, L.T., 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. University of Otago Press/Zed Books, Dunedin, London and New York.
- Sorrenson, M.P.K., 1987. *Na to Hoa Aroha: From Your Dear Friend, the Correspondence between Sir Apirana Ngata and Peter Buck, 1925–1950, vol. 2*. Auckland University Press, Auckland.
- Soutar, M., 1996. A framework for analyzing written iwi histories. *He Pukenga Kōrero* 2 (1), 43–57.
- Soutar, M., 2000. *Ngāti Porou Leadership; Rāpata Wahawaha: kei Te Ora Nei Hoki Tātou Mo to Tātou Whenua*. PhD Thesis. Massey University.
- Tamua, M., 2015. *Te Takanga o Te Wā*. Ministry of Education, Wellington.
- Tau, R.T.M., 2001. *Mātauranga Māori as an epistemology*. In: Andrew Sharp, A., McHugh, P. (Eds.), *Histories, Power and Loss: Uses of the Past- a New Zealand Commentary*. Bridget Williams Books, Wellington, pp. 61–73.
- Tau, R.T.M., 2009. Review of Paul Moon's *This Horrid Practice*. *Te Pouhere Kōrero* 3, Māori History, Māori Peoples, pp. 119–124. Wellington.
- Vansina, J., 1985. *Oral Tradition as History*. London, James Currey. University of Wisconsin Press, Wisconsin.
- Walker, R., 1990. *Ka Whāwhai Tonu Mātou, Struggle Without End*. Penguin, Auckland.
- Williams, D.V., 2014. Historians' context and lawyers' presentism: debating historiography or agreeing to differ. *N. Z. J. Hist.* 48 (2), 136–160.
- Woodson, C.G., 1933. *The Mis-education of the Negro*. The Associated Publishers, USA.

Toward a rights-based and culturally responsive model of Aboriginal Education

Bob Morgan, University of Newcastle, Callaghan, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	7
Dedication	9
Ethnic fraud	10
The Australian Government Productivity Commission	11
Moving from Aboriginal Education self-determination to education as sovereign right	12
Elements of a aboriginal rights-based education model	12
Instruments of possible change and transformation	12
Eurocentric education as a tool of assimilation and oppression	13
Aboriginal Education as social reproduction	13
Global Indigenous Education sites of struggles	14
Canada	14
United States of America	15
Aotearoa (New Zealand)	16
Connecting the eurocentric education and aboriginal incarceration dots	17
Meaninglessness, alienation and loss of culture and purpose (MAL) & Aboriginal Education	17
Positive signs of the promise and transformation	18
Culturally grounded education warriors	19
Final words	19
References	20

Introduction

The paper engages narrative storytelling, a process that privilege Aboriginal knowledges and worldviews while celebrating how Indigenous peoples create and share knowledge. My interpretation of narrative storytelling is influenced and shaped by the work of Jo-ann Archibald, who, when writing about Coyote the Trickster posits that the Trickster learns lessons “the hard way.” Archibald writes: “The Trickster stories remind us about the good power of interconnections within family, community, nation, culture and land” (Archibald, 2008, p. i). Archibald adds: “If we become disconnected, we lose the ability to make meaning from Indigenous stories.”

Tensions emerge in my narrative storytelling due to my use of the language of the coloniser to create and share meaning. This is compounded with the loss of my Gumilaroi language with the use of dominant cultures’ academic and literary standards creating a type of conformity and hegemony, within which I risk the loss of my authentic Aboriginal voice.

Thinking Gumilaroi and writing in the colonisers language is both a constant, and at times debilitating challenge. Why do I write and for whom? I often feel as if I’m losing something or surrendering my voice when I write using a style that is entrenched in and privileges the dominant culture.

I find myself wondering, is this why Indigenous movie makers and those involved with the creative and performing arts are so successful at their craft? In these fields they are less bound, or at least enjoy a certain freedom, enabling them to tell their story unencumbered by western dictates and conventions.

The haunting lyrics of Aboriginal storyteller, songwriter and entertainer, Kev Carmody reflect this freedom in his seminal song “Can Not Buy Our Soul.” Carmody’s lyrics remind us that oppression and hegemony is ever present but that we must never sell our soul nor should we ever surrender to injustice.

For 200 years us blacks are beaten down here too long on the dole

My dignity I’m losing here and mentally I’m old

There’s a system here that nails us and we’re left out in the cold

They took our life and liberty friend but they couldn't buy our soul

Joe Hill died, Che Guevara fought and Pemulwuy lay down dead

If a person speaks out critically here you can get loaded down with lead

How long can the majority wait for their story to unfold

They took life and liberty friend but they could not buy our soul

The cleverman spoke precisely, humanity he said was done

It's creed of greed could not proceed if our struggle's to be won

For humanity's more important here than that constant quest for gold

You may take life and liberty friend but you cannot buy our soul

Jesus woke one morning, rode a donkey up through the gate

He could see quite clearly he was going to face his fate

And the powers that be, could see that he, could not be bought or sold

They took his life and liberty friend but they could not buy his soul. (2007)

Addressing the tension of writing across cultures and ontologies, Nunga academic Irene Watson pondered: "Lately I have reflected a lot on why and how I speak, and write and that is because I feel a pressure to perform. Why do I feel this pressure to perform, and in what context do I feel it? In this one. The pressure is to locate myself within a space where the muldarbi has been working for centuries in dismantling my Nunga being.¹ The risk of entering this space is to become assimilated by the muldarbi, the challenge is to survive and remain a Nunga" (Watson, 1998, p. 28).

Aboriginal educator, researcher and scholar, Dr. Nerida Blair also writes of this tension in her book, *Privileging Australian Indigenous Knowledge: Sweet Patatoes, Spiders, Waterlilies and Brick Walls* (Blair, 2015). Blair shares:

Much of our energy as Indigenous peoples has been trying to 'fit in' to the Western Eurocentric system or fight being 'fitted in' by Eurocentric scholarship (Kovach, 2009, p. 31) Wilson (2008, p. 127), Arbon (2008, p. 27). I liken this to one of those boxes a lot of young children had in schools, at home to learn shapes, colours and numbers. The box was rectangular. The box has a number of different shapes cut into it; triangle, circle, square. The child has

¹Muldarbi: Demon Spirit.

a corresponding shaped block to fit into the shape in the box. If we as Indigenous peoples are diamond shape you can eventually fit us into every shape in the box. At what cost? How much of our exterior shape is rubbed raw in an effort to make us 'fit'; hammering, twisting, manipulating? Another word for this is assimilation.

(Blair, 2015, p. xix).

Similarly, Arbon (2008) argues that:

Aboriginal existence continues in a desperate struggle as empire manifests locally through a colonialism that demands homogeneity in most all undertaken in Indigenous lives.

(Arbon, 2008, p. 134).

Like, Watson, Blair, Arbon and others, I choose to speak and write my truth, albeit through the medium of the language of the coloniser, because Aboriginal voices have been silenced for far too long. I also choose in this paper to use my own narrative style within the boundaries and constraints of writing through the medium of the coloniser. When Indigenous peoples share our stories they are often misinterpreted, relegated and subjugated to a place where meaning is viewed by settler societies and defined as mere folklore and mythology. Voices that are steeped in Aboriginal culture and experience must be liberated and asserted if real and sustainable change is to occur in education and other areas of Aboriginal need. Aboriginal voices are critical in that place where truth and truth making reside so that a more equitable and respectful world of diversity and difference is realised.

Through my narrative storytelling I invoke 5 decades of experience in visioning, policy writing, managing teams and structures, facilitating community-led processes, empowering community-engagement in what is referred to as Aboriginal education. The valued voices of others are present in this journey and they are celebrated and woven into my narrative storytelling.

A central theme of the paper is that education failure levels currently experienced by a significant number of Aboriginal learners, particularly at the schooling level of education, should be viewed essentially as a rejection by Aboriginal learners of the assimilationist imperatives embedded in Eurocentric education systems.

Addressing assimilation and hegemonic forces embedded in Eurocentric education, Battiste (2000) argues ... "that Aboriginal people have been subjugated by the military, political, and economic systems, but the most effective oppressive attack on First Nations cultures has been the education system." Battiste adds: "the subtle influence of cognitive imperialism, modern educational theory and practice ... destroyed or distorted the ways of life, histories, identities, cultures, and languages of Aboriginal people" (p. 193).

This rejection by Aboriginal learners of the assimilationist imperatives, rooted as they are in Eurocentric cultures and ontology, are especially evident among Aboriginal learners who are strongly connected to their culture, their traditions, know their country and secure in their identity. Understanding what the existential essence of connection to country means for Aboriginal people is complex but it should never simply be interpreted as involving geography. Aboriginal people's connection to country, including lands, waterways, seas and sky involves symbiotic principals of custodianship, law obligation and accountability, spirituality, and kinships forged over thousands of years. Country is the library where hundreds of generations of knowledge and Aboriginal wisdom resides.

It is argued that achievement levels will not improve at the schooling level of education for the vast numbers of Aboriginal learners until they see themselves and their people positively and affirmingly in the curriculum they engage. This will require systems and the teachers who work within them, to be trained, untrained and retrained to ensure that this concept of country, and all that it involves, informs curriculum content and a culturally affirming and rights based pedagogy.

Providing a more nuanced definition of this assertion, Professor Lester - Irabinna Rigney uses the term "Indigenist epistemologies" to identify and define what he calls "provocations." These provocations are: privileging Indigenous voices, upholding integrity of cultural knowledges to resist hegemony, and empowering self-determination (Rigney, 2001).

The provocations cited by Rigney and a rights-based Aboriginal education pedagogy, are viewed as pivotal to constructing a more culturally affirming and responsive pedagogy and epistemology of Aboriginal education.

This paper builds on and supports the work of others who are challenge the assimilationist and hegemonic forces of Eurocentric education systems and their persistent failure to accommodate Aboriginal cultural and education rights and aspirations. This failure to accommodate is compounded by the compulsory nature of schooling in Australia. The compulsory requirement attached to schooling in Australia has been something that I've grappled with over the years because if it is compulsory for Aboriginal kids to attend school then surely there is a duty of care and obligation for schooling to accommodate and pay respectful attention to the culture and life circumstance of the kids who are compelled to attend government and other schools.

Dedication

The paper is dedicated to the tireless efforts of Aboriginal education warriors, parents, community members and the non-Aboriginal supporters who stand with Aboriginal people against oppression and who demand that historical wrongs are righted and that current inequities are corrected.

Some of these young warriors are gifted and culturally grounded but many lack the experience to withstand the seductive and hegemonic forces of the systems they are employed within, including education. In this context, the young ones are all too often expected to comply with Eurocentric education objectives and are advised to focus primarily on student attendance and behavioral matters thereby leaving curriculum content, assessment methods and teaching methods and other matters to others. Of course, the young ones work with others to address all these matters but they also understand that they have a duty of care to ensure that systems respect and accommodate their cultures, experiences and worldview. Navigating this terrain involves walking a tense tight-rope between systemic expectations and community and cultural obligations.

Illustrating the tension inherent in this terrain, I was recently invited to meet with a group of young Aboriginal educators employed with the NSW Department of Education and some of them shared that they constantly struggle to guard against implied and often explicit assimilation and social reproduction imperatives. These are young people who simply want to serve their community, live life with dignity and to provide for their family and kin without having to constantly fight against the assimilationist and hegemonic forces of the education system they are employed within. Eurocentric and assimilationist systems can be seductive, especially for isolated Aboriginal employees and those disconnected from country, kin and community.

Ironically, as is the case with many Aboriginal people in similar positions based in majority cultures and systems, members of this group of Aboriginal people were employed in “designated” or “identified” positions meaning that their cultural background and experience, it is expected, would position them to provide an Aboriginal voice and advocacy across the scope of their identified duties. However, some of these young ones reported that they were increasingly expected to attend to duties in a way that totally ignores and confounds the purpose for which they were employed. To paraphrase Dr. Eber Hampton, past president of the First Nations University in Regina, Saskatchewan, Canada, many of these young ones were expected to employ “Anglo methods to satisfy Anglo education purposes” (Hampton, 1995).

The young ones reported that they are increasingly forced to withstand the pressures to embrace and conform to non-Aboriginal assimilationist objectives, often resulting in burn-out or what I refer to as spiritual fatigue (Morgan, 2020). The young ones reported that they have no one to turn to for cultural and professional advice and support within the system, and if it does exist, it is often in the form of non-Aboriginal senior staff with limited knowledge of the young ones culture, community and country.

The young ones were advised to never surrender to or to be silent in the face of injustice. That they should strive to honor those on whose shoulders they stand; because to surrender, or to be silent, simply empowers the system.

This workplace tension has been there for many years and it has never been adequately addressed, forcing many of the young ones to surrender or seek employment in more culturally safe working environments. Of course, they have other obligations beyond their connection to culture, kin and country including rents or house mortgage and rental payments, car payments, and other living expenses so employment is critical to meeting these obligations, but at what a cost?

Culturally grounded Aboriginal people are better equipped to understand and therefore better able to function in this space compared to Aboriginal people whose identity is simply defined and informed by heritage.

“Heritage Aboriginal people” are those who I argue have little understanding of their country, kinship structures and obligations, they lack knowledge of language and other lessons derived from timeless song-lines, oral traditions and cultures that have sustained Aboriginal people for countless generations.

Heritage Aboriginal people merely connect through often ill defined or ambiguous heritage and are often unable to identify, precisely, the Aboriginal nation they reportedly draw their heritage from. Heritage Aboriginal people have rarely sat on country to engage with the old ones so that a more profound level of understanding is developed of critical matters that would define and inform their identity. Conflating heritage with genuine community, kinship and country based Aboriginality is highly problematic and insulting.

Ethnic fraud

Heritage Aboriginal people are different from culturally grounded Aboriginal people but there is another rapidly increasing group who fraudulently claim Aboriginality for the purpose of a perceived or real benefit, a benefit that is designed to address Aboriginal disadvantage and marginalization. This form of “ethnic fraud” is of considerable concern within Aboriginal communities and recently led to two NSW Local Aboriginal Land Councils (LALCs) submitting a formal complaint and request for an investigation by the NSW Independent Commission Against Corruption (ICAC) of perceived practices of ethnic fraud involving staff and students at one of Australia’s most prestigious higher education institution, Sydney University. (See “The Australian”, Thursday, November 25, 2021.)

While ethnic fraud might be considered, by some, to be extraneous to the right of Aboriginal people to access a rights-based education, ethnic fraud, left unchallenged, distorts and belittles genuine attempts of Aboriginal people to overcome long denied access to services and programs that are generally enjoyed by mainstream Australians. If left unchallenged, ethnic fraud also provides the opportunity for individuals who have used ethnic fraud to gain academic qualifications, and other advantages, to obtain positions where they can influence and shape public, institutional and corporate policy.

The Australian Government Productivity Commission

The report of the Australian Government Productivity Commission “Overcoming Indigenous Disadvantage: Key Indicators 2020” was released on December 3, 2020 (SCRGSP, 2020). The report measures the wellbeing of Aboriginal and Torres Strait Islander people and it is illuminating to note some of the statistics relative to Aboriginal and Torres Strait Islander education. The Productivity Commission reports:

This comprehensive report card measures where things have improved (or not) against 52 indicators across a range of areas including governance, leadership and culture, early childhood, education, economic participation, health, home environment and safe and supportive communities. The report includes case studies on governance, with a specific focus on identifying arrangements that support shared decision-making between Aboriginal and Torres Strait Islander people and Australian governments.

(Australian Productivity Commission, 2020).

Reading, writing and numeracy performance for Aboriginal and Torres Strait Islander students was reported thus:

Results for this indicator have fluctuated over time with no clear trend. Comparing 2008 and 2019, the proportion of Aboriginal and Torres Strait Islander students achieving national minimum standards increased for Years 3 and 5 for reading, and for Years 5 and 9 for numeracy. However, volatility in the data means conclusions drawn from comparisons should not be interpreted as consistent improvement over time.

(Australian Productivity Commission, 2020).

Addressing Year 1–10 attendance pattern data the report states:

In 2019, the overall attendance rate for Aboriginal and Torres Strait Islander students was 82 per cent, a decrease from 84 per cent in 2014.

The attendance rate for non-Indigenous students was 92 per cent in 2019.

The proportion of Aboriginal and Torres Strait Islander 20–24 years olds completing year 12 or equivalent or above increased from 45 per cent in 2008 to 66 per cent in 2018–19. For non-Indigenous people, the proportion remained relatively unchanged (90 per cent in 2017–18).

(Australian Productivity Commission, 2020).

In relation to Year 12 attainment, it is reported that:

The proportion of Aboriginal and Torres Strait Islander 20–24-year olds completing year 12 or equivalent or above increased from 45% in 2008 to 66% in 2018–19. For non-Indigenous people, the proportion remained relatively unchanged (90% in 2017–18) (Australian Productivity Commission, 2020).

These findings are not surprising when viewed through an Aboriginal rights based lens and reflects the ongoing failure of Euro-centric systems of schooling to effectively engage Aboriginal students in learning.

In light of the findings of Productivity Commission, and numerous other factors, it is advocated that education for Aboriginal learners must be seen as a right, provided in a culturally safe and responsive learning environment, utilizing a pedagogy that emanates from and is defined by and privileges the rights and freedoms of Aboriginal peoples expressed and recognized in instruments such as the UN Declaration on the Rights of Indigenous Peoples (UN, 2007) and asserted in the Coolangatta Statement.²

This paper does not offer any easy fix or instant remedies except to stress that the existing “guest paradigm” (Morgan, 2019) of assimilation and social reproduction model of Aboriginal education is flawed. The non-Aboriginal – Aboriginal power relationship operates creating a “partnership” in which Aboriginal people have all the need but systems retain and control all the resources. This “partnership” has generated irrelevant, and in many cases alienating curricula, ill-informed teaching methods and assessment procedures, poor parent and carer engagement processes and a host of other failings.

The positioning of Aboriginal learners as alienated guests in the Eurocentric education domain must end if the learning experiences and outcomes of Aboriginal learners are to improve.

²The Coolangatta Statement (WIPCE, 2006) was initiated during the 1993 World Indigenous People’s Conference on Education (WIPCE) and ratified by the 1996 WIPCE conference held in Hilo, Hawaii in 1999. The Coolangatta Statement was one of the documents that informed the education component of the UN Declaration on the Rights of Indigenous Peoples (UNDRIP).

Moving from Aboriginal Education self-determination to education as sovereign right

Aboriginal sovereignty is not recognized in the Australian Constitution or Australian law although calls are being made for constitutional amendments that would give effect to the recognition of a form of Aboriginal sovereignty including an Aboriginal and Torres Strait Islander voice to parliament. These matters and others are contained in the Uluru Statement from the Heart that emanated from the First Nations Constitutional Convention held in Central Australia in 2017. ([First Nations Constitutional Convention, 2017](#)).

Of course, the fact that Aboriginal sovereignty is not legally recognized in Australian law does not necessarily negate or invalidate its authenticity. The recognition of Aboriginal and Torres Strait Islander sovereignty is an enabler for treaty making, truth telling and sociopolitical change. Aboriginal children who are linked to culture and country intuitively know and are reminded that they belong to the oldest continuous culture on the planet and this sovereign reality must be part of their schooling journey.

Aboriginal sovereignty involves more than a claim to Western notions of prior-existence or land ownership; at its core Aboriginal sovereignty declares and defines the symbiotic and custodial relationship that Aboriginal people have with the land, rather than over it, and this relationship continues today and will persist into the future.

Hopefully a model of Aboriginal rights based education will be part of the commitment of the newly elected Albanese federal government to the Uluru Statement from the Heart. The new Australian Prime Minister, in claiming victory in the May 2022 election declared: “On behalf of the Australian Labor Party, I commit to the Uluru Statement from the Heart in full.” (See: <https://ulurustatement.org/the-statement/>).

Elements of a aboriginal rights-based education model

Discussing a rights-based education for Aboriginal students, Professor Rigney in his paper titled “Aboriginal Child as Knowledge Producer” ([Rigney, 2020](#)) offers what he terms “provocations on the rights of the knowledge producing Aboriginal child.” These provocations include:

1. The right to Indigenous subjectivity and languages
2. The right to produce knowledge for the dimensions of sovereignty, equality, democracy,
3. Solidarity, and freedom
4. The right to produce knowledge that privilege voice, integrity of Indigenous epistemologies, and intelligences, to resist hegemony and to strengthen self-determination
5. The right to global literacies, numeracies, and enriching ideas of humanity
6. The right to have rights and dignity as a human being
7. The right to belong, become, and be different
8. The right to produce digital, relational, existential, creative, embodied, scientific, political, and economic knowledge to move fluently across cultures without compromising those of self
9. The right to ecological sustainability where all things are interconnected
10. The right to love, peace, and treaty free from racism and colonialism
11. The right to not inherit a dead planet (see [Rigney, 2020](#)).

These “provocations”, and others, must form the basis for a new rights based and informed Aboriginal education policy and pedagogy.

Rigney’s work and the work of others, offer a fundamental lesson advocating that education for Aboriginal peoples will be better and more effectively addressed when Aboriginal people are directly engaged with and control its nature, scope and purpose. This “pedagogy of engagement” is a clear manifestation of a rights informed theory and praxis of “Aboriginal education.”

Instruments of possible change and transformation

Various instruments have been drafted to protect and advance the rights and freedom of Indigenous peoples including the WIPCE initiated Coolangatta Statement, that was used as a key document to inform the drafting of the education section of the United Nations Declaration on the Rights of Indigenous Peoples ([UN, 2007](#)). UNDRIP took over 20 years to negotiate and was finally adopted by the UN General Assembly in September 2007. Australia, the USA, Canada and New Zealand were the only member states that refused to initially sign and ratify the Declaration citing that, in their view, the instrument was “unfair and unworkable” particularly in respect to Article 26, which states that “Indigenous people have a right to own and control lands they traditionally used.”

Australia finally ratified and endorsed the declaration in 2009. It is important to understand however, that UNDRIP is not legally binding and has no recognition in international law. Its validity and efficacy is left to member states to decide how, and in what form, Indigenous rights and freedoms, including self-determination and the right to education, are acknowledged, respected and accommodated.

UNDRIP addresses the principle of self-determination at article 3 and assimilation at article 8 while education is specifically addressed at article 14.

With respect to the principle of self-determination, Article 3 reads:

Indigenous peoples have the right of self-determination. By virtue of that right they freely determine their political status and freely pursue their economic, social and cultural development.

Assimilation and the destruction of Indigenous culture are addressed at Article 8 which reads.

1. Indigenous peoples and individuals have the right not to be subjected to forced assimilation or destruction of their culture.
2. States shall provide effective mechanisms for prevention of, and redress for:
 - a. Any action which has the aim or effect of depriving them of their integrity as distinct peoples, or of their cultural values or ethnic identities;
 - b. Any action which has the aim or effect of dispossessing them of their lands, territories or resources;
 - c. Any form of forced population transfer which has the aim or effect of violating or undermining any of their rights;
 - d. Any form of forced assimilation or integration;
 - e. Any form of propaganda designed to promote or incite racial or ethnic discrimination directed against them.

Article 14 addresses the right of Indigenous peoples to education and states:

1. Indigenous peoples have the right to establish and control their educational systems and institutions providing education in their own languages, in a manner appropriate to their cultural methods of teaching and learning.
2. Indigenous individuals, particularly children, have the right to all levels and forms of education of the State without discrimination.
3. States shall, in conjunction with indigenous peoples, take effective measures, in order for indigenous individuals, particularly children, including those living outside their communities, to have access, when possible, to an education in their own culture and provided in their own language.

As previously argued, the effectiveness of these types of instruments are only as good as the respect and adherence they engender from nation states and dominant cultures. Sadly, history is riddled with examples of how nation states and dominant cultures have dishonoured and completely ignored such instruments as well as treaties.

A rights based education for Aboriginal people is a concept that many policy makers and non-Aboriginal education theorist struggle to comprehend. Perhaps it is because the despair and powerlessness that are born of invasion, dispossession and forced assimilation are absent from their lived experience; a virulent form of white privilege and an all too pervasive presence of wide-spread cognitive dissonance.

Eurocentric education as a tool of assimilation and oppression

The invasion of Aboriginal lands in Australia and the ancestral lands of Indigenous peoples around the globe was brutal, genocidal and relentless. With Darwinism and imperialism as their compass and conquest and domination as their social and political rationale, invading colonial forces used whatever means they considered necessary to usurp Aboriginal lands and subjugate its people. In Australia, massacres, poisoned waterholes and infected blankets were commonly used as weapons of oppression and as a means to aid in the spread of colonialism. The objective, though rarely stated, was clearly one of genocide.

Education was a tool weaponized by colonizing forces and was used as a mean to pacify and disconnect Aboriginal people from country and strip them of their cultural beliefs, rituals, customs, traditions and language. Whether it was Mission schools and the Stolen Generation in Australia, Indian Residential schools in Canada, the Indian schools in the US, the Native Schools in Aotearoa or similar experiences of other Indigenous peoples in other countries, each institution was established and operated clearly with a singular objective – forced assimilation. This is best captured in a statement made to American reformers in 1892 by Richard Pratt, head of the Carlisle Indian School when he argued: “Kill the Indian ... and save the man” (Pratt, 1892).

Aboriginal Education as social reproduction

The Council of Australian Government’s (COAG) National Education Agreement states that an agreed goal of schooling is that:

... all Australian school students acquire the knowledge and skills to participate effectively in society and employment in a globalised economy. (SCRCSP, 2010).

Tensions arise from this goal for Aboriginal students and their parents and communities when issues of racism, poverty, social inequalities and other marginalizing factors are understood as being present in the education journey of Aboriginal students.

Existing Aboriginal education interpretations are primarily structured so that they are fused onto white, Eurocentric knowledge systems that are fundamentally mired in and privileges what French sociologist and cultural theorist Pierre Bourdieu refers to as “education as social reproduction.” In a Bourdieuan view, education system, and more importantly, the curriculum that reproduces the dominant paradigm of inequalities ensures the continuation of conformity and disempowerment of socioeconomically and politically marginalized groups (Bourdieu and Passeron, 1977).

A balanced and synergized pedagogical model incorporating predominantly Aboriginal and a selection of aspects of non-Aboriginal knowledge from the majority culture domain, absent notions of social reproductions, is critical to providing Aboriginal learners the opportunity to develop the knowledge and skills necessary for living “a good life.”³

Unfortunately, there are no visible signs that this will happen anytime in the foreseeable future especially in light of what appears to be a shift to a corporate model of education administration in NSW and perhaps across Australia.

Global Indigenous Education sites of struggles

Indigenous peoples around the world have steadfastly held to their sovereign rights and freedoms, including the right to determine the nature and scope of the education systems that Indigenous people access and participate in. But colonial and settler forces have also held tenaciously to their power and privilege to control “Aboriginal education.” A brief and cursory analysis of “Aboriginal education” experiences in Australia, exhibits a journey of denied potential and what noted Maori scholar, Graham Hingangaroa Smith (2015) refers to as “sites of struggle.” Sites of struggle are manifest in the rejection of current assimilationist education in Australia and in most other global Indigenous contexts.

Canada

In Canada, the principle of Aboriginal Self Determination is manifest in the public policy on Indian education “Indian Control of Indian Education.” The policy, developed in 1972 by the National Indian Brotherhood (now the Assembly of First Nations), was a direct repudiation of the Canadian government’s White Paper on Indian Issues that included plans to abolish the Indian Act and promoted the notion that there should be no special status for Canadian First Nations peoples.

Discussing Indian rejection of the White Paper, Barman, Hebert and McCaskill (eds.) in *Indian Education in Canada: The Legacy* (1986) explain:

Indian reaction to the White Paper was overwhelmingly negative, Indian groups coalesced in opposition, the proposal being seen as a mechanism for the federal government to escape from its historical responsibility for Indian affairs, including the obligation to right past injustices in such areas as land claims and treaties.

In his book; “*The Unjust Society: The Tragedy of Canada’s Indians*” the late celebrated Indian leader Dr. Harold Cardinal, wrote:

The history of Canada is a shameful chronicle of the white man’s disinterest, his deliberate trampling of Indian rights and his repeated betrayal of our trust. Generations of Indians have grown up behind a buckskin curtain of indifference, ignorance and, all too often, plain bigotry. Now, at a time when our fellow Canadians consider the promise of a Just Society, once more the Indians are betrayed by a programme which offers nothing better than cultural genocide. (Cardinal, 1969, cited in Petten, 2017)

Cardinal described the Canadian Trudeau liberal government’s plan for Indian affairs as nothing more than a “thinly disguised program of extermination through assimilation.”

In his paper, titled “*Toward a Redefinition of Indian Education*” Hampton writes of Indian education experiences that are reflected in other Indigenous people’s experiences around the globe. This history and current realities indicate that the erasure of systemic education failure for Aboriginal peoples will only occur when Aboriginal people are in the position to shape the educational journey of Aboriginal learners. Aboriginal learners need to see themselves in the curriculum they engage and teachers must be trained, untrained and retrained to ensure that this is central to the learning journey of Aboriginal learners.

A rights-based system of education parallels what past president of the First Nations University in Saskatchewan, Dr. Eber Hampton describes as “*Schooling for Self Determination*” (Hampton, 1995).

Hampton contends that the term “Indian education” has five interpretative meanings, (1) Traditional Indian education, (2) Schooling for Self Determination, (3) Schooling for assimilation, (4) Education by Indians, and (5) Indian Education: *Sui Generis*.

³The concept of living “a good life” was shared by noted Yup’ik anthropologist and educator the late Dr. Oscar Kawagley during a private conversation during the 2010 American Education Research Association conference held in San Diego, California, USA. Dr. Kawagley was responding to the question: what is the main purpose of Indigenous life?.

Hampton's interpretative meaning of, "schooling for assimilation" feature in this paper. I argue that the assimilation of Aboriginal peoples was and remains the main objective of "Aboriginal education" in Australia.

Discussing Indian "Schooling for Self Determination", in his five "interpretative meanings" of Native education, Hampton advances that:

The second phase of Native education was the establishment of schools for Native children. Although schools as institutions were non-Native in origin and character, there have been vast differences in the goals, methods, and outcomes that it is possible to distinguish two kinds of schooling for Native: schooling for self-determination and schooling for assimilation.

Hampton (1995) further adds:

Although neglected in standard histories of Native education, there have been many examples of highly successful Native oriented schools. For example, schools established and controlled by the Chickasaw, Choctaw, and Cherokee nations, as well as Russian mission schools for the Yup'ik people, characterised by the use of Native languages, positive attitudes toward Native cultures, good school-community relations, and emphasise on self-determination rather than assimilation. These schools had high success rates in literacy and education attainment.

(Oleska and Dauenhauer, 1982).

Hampton shares, that these school were unfortunately closed by the unilateral actions of government.

Discussing "Schooling for Assimilation" Hampton characterizes this experience for Indian learners in the following manner:

Historically and in most contemporary situations, the education of Indians is carried out by Anglos using Anglo methods to satisfy Anglo purposes (American Indian Policy Review Commission, 1976). In contrast to schooling for self-determination, these schools for assimilation have been characterised by high failure rates in literacy and educational attainment, having assimilation rather than self-determination as goals, poor school – community relations, negative attitudes toward Native cultures, and prohibition or non-use of Native languages.

(Oleska and Dauenhauer, 1982).

Five decades after the introduction of the policy of Indian Control of Indian Education in Canada, evidence flowing from the seminal work of the Royal Commission on Aboriginal Peoples (1991–1996); and the Truth and Reconciliation Commission (2008–2015), Indian communities are beginning to slowly reap the benefits of their vision and determination. Education facilities of education excellence, as define by Indian people, such as the First Nations University (FNU) in Saskatchewan, the First Nations House of Learning (FNHL) at the University of British Columbia (UBC), The Nicola Valley Institute of Technology (NVIT), Old Sun Community College, Siksika First Nation, in Alberta and numerous other Indian controlled education institutions herald the true promise of academic success and cultural affirmation for Canadian First Nations peoples.

In Canada, as in other global Indigenous contexts, schooling has always been a site of struggle and trauma and we need to look no further than the tragedy of Residential Schools for evidence of the assimilation objective of colonizing and settler societies. What served as internment camps and masqueraded as schools, were in fact sources of pain and trauma that have scarred the lives of Indian learners and their families for generations. The recent discovery of 215 unmarked graves on the site of the Kamloops Indian Residential School (KIRS) in the Canadian province of British Columbia (BC) is distressing not only in terms of the secrecy that surrounded the circumstances of the burials but also the scale of the number of children who died in isolation and forcibly disconnected from their family and community (see Austen, 2021).

The tragic discovery of another 751 unmarked graves in a cemetery near the former Marieval Indian Residential School, north of the Saskatchewan provincial capital of Regina deepens the tragedy and signals that many more discoveries are likely to be made.

The families and loved ones of these children, some reportedly as young as 4, were denied the right to farewell their children with honor and within the traditions and protocols that govern how such ceremonies are conducted. No amount of financial compensation or apology can ever ease the painful and shameful history of the evil crimes that were committed against Indigenous peoples in the name of colonization and the settler society that have evolved.

United States of America

Indigenous peoples around the globe have suffered and seek freedom from the impact of colonialism. American Indian Nations have always asserted their right to culturally sustainable education, resolutely rejecting the assimilationist and social reproduction objectives embedded in Eurocentric education. The American Indian rejection of Eurocentric education is graphically captured in the classic statement made by an Indian leader as reported by Benjamin Franklin in 1794. Commenting upon the irrelevance and threat of Eurocentric education the Indian leader declared:

But you, who are wise, must know that different Nations have different Conceptions of things: and you will therefore not take it amiss, if our ideas of this kind of Education happen not to be the same as yours. We have had some Experience of it; Several of our young people were formally brought up at the College of the Northern Provinces; they were instructed in your Sciences; but, when they came back to us, they were bad Runners, ignorant of every means of living in the Woods, unable to bear either Cold or Hunger, knew neither how to build a Cabin, take a Deer, or kill an Enemy, spoke our language imperfectly, were therefore neither fit for Hunters, Warriors, nor Counsellors, they were totally good for nothing. We are however not the less oblig'd by your kind Offer, tho' we decline accepting it; and, to show our grateful Sense of it, if the Gentlemen of Virginia will send us a Dozen of their Sons, we will take great care of their Education, instruct them in all we know, and make Men of them.

This eloquent commentary has survived the test of time and its currency is as valid today as it was in 1794.

In more contemporary times, Native Americans continue to reject Eurocentric systems of education by asserting what is tantamount to an expression of sovereignty. This assertion of sovereignty can be found in the establishment and growth of Indian Tribal Colleges, the first of which was the Navajo Community College established in 1968. Reporting on the Tribal College movement, Ernest L. Boyer in his foreword to *Tribal Colleges-Shaping the Future of Native America*, (1989) made the following observation.

For many decades American Indian reservations have been demoralized by the seemingly permanent condition of extreme poverty. In some tribal communities, unemployment reaches 80 percent. The disease of alcoholism has taken a severe toll on Indians since its introduction into their culture, and today requires the most serious immediate attention. Health care is of the most critical concern. Infant mortality rates on some of the reservations are at shocking levels-as high as double the national average. The tragedy of early death from illness or suicide has touched all members of these families-oriented communities. These crises require urgent attention, and tribal colleges are working to provide the leadership, programs, and resources to meet the challenges.

(Boyer, 1989, p. xii).

Many of the circumstances Boyer comments upon are reflected in the Aboriginal experience throughout Australia. Arguing that American society can learn a great deal from Indian peoples Boyer further adds:

We believe that a good place to start learning from American Indians is the Tribal College, where we can learn about survival, about hope and determination in the face of extreme adversity, about renewal of community, about reclaiming the individual and the society from dependencies of all sorts, and about creatively connecting education to the larger world. Clearly American society has a great deal to gain by supporting the tribal college movement and learning from the first Americans.

(Boyer, 1989, p. xiv).

The Tribal Colleges are at once a rejection of the Eurocentric education systems and a beacon to the incredible resilience of Indigenous people across the world.

Aotearoa (New Zealand)

Closer to Australian shores, the failure of Eurocentric education systems to deliver meaningful and culturally affirming and academically empowering education experiences and outcomes for Maori led the Waitangi Tribunal in 1986 to assert:

The education system in New Zealand is operating unsuccessfully because too many Maori children are not reaching an acceptable standard of education. For some reason they do not or cannot take full advantage of it. Their language is not adequately protected and scholastic achievements full far short of what they should be. The promises of the Treaty of Waitangi of equality in education as well as in all other human rights are undeniable. Judged by the system's own standards Maori children are not being successfully taught, and for that reason alone, quite from a duty to protect the Maori language, the education system is being operated in breach of the Treaty.

But change is underway in Aotearoa with New Zealand Stats (2020) reporting that data from the 2018 NZ census showed: "...that in both secondary and tertiary studies, younger Māori and Pacific peoples are achieving outcomes closer to the national average than their counterparts in older age groups."

New Zealand Stats added: "In the 2018 Census, 80.6% of Māori and 83.0% of Pacific 15- to 24-year-olds had at least a level 1 qualification or equivalent (such as School Certificate), compared with 85.8% of 15- to 24-year-olds nationally."

While New Zealand Stats, Susan Hollows, senior manager census data delivery, observed: "There has been a focus over recent decades on improving cultural understanding and teaching approaches within New Zealand's education sector, which may have had an effect on the numbers of students achieving formal qualifications" (See New Zealand Stats 2018).

Unquestionably Maori are determined to control their educational destiny; a determination matched only by their tenacious and highly successful effort to save their language from extinction. The development of the innovative and creative language re-

vitalization program Te Kohanga Reo language immersion, that are tied to Maori education objectives, commenced in 1982 and is recognized and celebrated across the world as best practice, especially by Indigenous peoples involved with language protection and revitalization. The destruction of Indigenous languages was, and in certain circumstances remains, a fate common to many Indigenous peoples impacted by invading colonial forces and settler societies.

Some attempts are being made to save and revitalize Aboriginal languages in NSW and this is acknowledged and applauded but most of these efforts are school based and therefore susceptible to government rules and regulations, thereby relegating the voice of Aboriginal people as, at best, incidental or functionally inconsequential.

Even with this perfunctory examination of the international Indigenous education experience, there are clear signs that systemic education failure is manifest at alarming levels in multiple Indigenous contexts. The data also illustrates that success is possible when education systems are designed and controlled by Indigenous peoples. In the Canadian, New Zealand and US context, a rights-based education is reflected in their respective negotiated treaties. No such treaty exists in Australia.

Connecting the eurocentric education and aboriginal incarceration dots

Aboriginal adult and juvenile detention rates are at ever increasing and alarming levels in Australia, the extent of which is evidenced in an article published in the Guardian newspaper in 2019 that reported that 100% of juveniles in youth detention in the Northern Territory (NT) were Aboriginal.

The pain and trauma of incarceration and the ever increasing incidents of death in custody led to the establishment of The Royal Commission into Aboriginal Deaths in Custody (RCIADIC) in 1987. The Royal Commission delivered its final report in 1991 after examining the circumstances of 99 Aboriginal deaths in custody. At Volume 2 - Chapter 16 (1991) headed "Schooling" the RCIADIC reported:

The failure of schooling to provide a meaningful and useful experience for many Aboriginal people interacts with, and is a reflection of, their failure to achieve desired levels of participation in Australian society generally, and to command a level of services in respect of education, health and social justice which is commensurate with the rest of the Australian population. In many of the cases investigated, in hearings conducted into underlying issues, and in numerous submissions, this Commission has heard of the inextricable links between the formal education system, child welfare practices, juvenile justice, health and employment opportunity as factors contributing to the disproportionate representation of Aboriginal people in police and custodial facilities.

(RCIADIC, 1991).

The report of the Royal Commission included 339 recommendations each aimed at reducing Aboriginal incarceration rates, deaths in custody and other associated matters, but some thirty years after the tabling of the Royal Commission report very few of the recommendations have been operationalized by any Australian state or territory and to compound matters, a reported 474 extra number of Aboriginal deaths in custody have occurred since the tabling of the Commission's final report.

Aboriginal deaths in custody are traumatic for the families and loved ones and I was impacted by this when I read a 2020 coroner's report dealing with an Aboriginal death in custody case. Reading such a report is challenging under any circumstance but it is made even more so when the person the report relates to is a relative, in this case, a nephew. My nephew was not the first of our extended family who became yet another statistic in the ongoing crisis of Aboriginal deaths in custody. Therefore, I continue to struggle to understand a world within which the lives of Aboriginal men and women are so little valued by the authorities who seemingly act with callous indifference and impunity to the findings of the 1991 Royal Commission into Aboriginal Deaths in Custody.

My nephew was a special young man, proud of his Gumilaroi identity and of being the father of 4 loving children. He struggled with the education system but was, and will always be, a much-loved member of our family, the grandson of my eldest sister Mary.

Like so many Aboriginal people, the nephew felt alienated from the world into which he was born and was failed by the Eurocentric education system he was compelled to attend. He sought refuge in illicit substances to ease the pain of meaninglessness, alienation and historical trauma. The treatment of Aboriginal people at the hands of authorities continues to torment families and unfortunately lead some of our people into life choices that often end in tragedy. But no matter what the circumstances, the people who are driven to such despair are human and deserve better than to take their last breath alone and in custody.

Meaninglessness, alienation and loss of culture and purpose (MAL) & Aboriginal Education

Meaninglessness, alienation and the loss of culture and purpose emerged as a concept from a project commissioned by an Aboriginal Medical Service (AMS) in NSW. The project was to examine why Aboriginal men were not engaging with AMS clinical services and programs and to recommend possible new approaches.

Focus group consultations with Aboriginal men were conducted during the project and the men reported a sense of what the project team identified as meaninglessness, alienation and a loss of culture and purpose (MAL). The concept of MAL is informed

by and aligns with Viktor Frankl's theory of logotherapy (Frankl, 1938). The men had a strong belief that they were losing their culture, their role as men and their Aboriginal identity.

The men were considered to be suffering from what celebrated psychiatrist, philosopher and author of "Man's Search for Meaning" (1938), Dr. Viktor Frankl, calls an "existential vacuum." Writing about the mental health of some of his patients Frankl shares the reported "feeling of the total and ultimate meaninglessness of their lives." His patients, he argued, suffered a:

... lack of awareness of a meaning worth living for. They are haunted by the experience of their inner emptiness, a void within themselves; they are caught in that situation which I have called the "Existential Vacuum."

(Frankl, 1938, p. 111).

Inspired by Frankl's work I contend that an increasing number of Aboriginal peoples also suffer from what I refer to as "Spiritual Fatigue" (2020), a debilitating consequence of being forced to live and withstand trauma while either constantly struggling for human rights and freedoms or being forced to constantly defend them. There is no room for celebration and meaning in this paradigm - it is simply a choice between struggle or defense.

Aboriginal learners carry this existential vacuum and spiritual fatigue into classrooms and unless school curriculum, teaching methods and assessments policies and practices recognize and accommodate this reality the learning experience of Aboriginal learners, will continue to be one of failure, for both learner and teacher.

Although there is, as yet, no empirical evidence to support it, I am firmly convinced that MAL and spiritual fatigue create an "existential vacuum" for Aboriginal children and is present in the education journey of many Aboriginal school learners, especially those who are culturally grounded, connected to country, kinship systems and the obligations and responsibilities that stem from such relationships. Though untested, it is believed that MAL and spiritual fatigue form an "existential vacuum" which may also be present as a determinant in the alarming rates of Aboriginal youth incarceration and the distressing rates of Aboriginal youth alienation and suicide.

I contend that meaning and purpose were present in traditional Aboriginal knowledge settings and manifested in rites of passage such as initiation ceremonies for both Aboriginal boys and girls.

As Aboriginal people we must take greater control of the education in which our children are enrolled if the debilitating experience of education failure, incarceration rates and youth suicide and other life limitations are to be effectively addressed. A rights-based education pedagogy incorporating the provocations identified by Professor Rigney above, and other matters, must be the basis upon which greater Aboriginal control of education is exercised.

Positive signs of the promise and transformation

Two exceptional Aboriginal education leaders, Fiona Kelly and Tammy Anderson are principals of schools with a high proportion of Aboriginal students and they are developing and applying an Aboriginal informed and aligned pedagogy that, if the system does not interfere, will undoubtedly result in improved outcomes for Aboriginal and other learners enrolled in their schools. These improved learning experiences and outcomes will take time.

Both these exceptional education leaders are tethered in their culture and community and this groundedness has helped them create learning environments where teachers are valued and where positive learning achievements for students are beginning to take hold. Community engagement and involvement is a strong and celebrated feature of the learning environment created and present in each of these schools.

Added to this is a growing list of non-Aboriginal teachers who are also striving to create culturally affirming and responsive learning environments for Aboriginal learners. One example of this is the work of a newly appointed teacher with one of the schools that I work with as a critical friend. The school is small with a majority of Aboriginal student enrolments and is located in a rural and isolated region of NSW. These are the types of schools that first year out or new teachers are usually assigned, as was the case with this teacher at the beginning of 2021.

The teacher is an Australian who had been working as an academic in overseas universities. I initially met the teacher during an orientation program I was invited to attend and participate in. I was intrigued as to why the teacher had decided to move away from university teaching and be involved with teaching Aboriginal kids. After getting to know him I witnessed a profound level of commitment to social justice and equity and given his academic background I advised that he keep a journal of his teaching experience and of his engagement with the community.

The teacher took this advice onboard and is in the process of developing and recording what he refers to as "Emergent Discourses", a process that he uses to more effectively engage the students he shares the learning journey with. The teacher's work is exciting and a synopsis of one of his journal texts is provided to illustrate the nature and potential of his work.

The teacher describes the philosophical underpinnings of his work in the following manner.

Donald Winnicott, in *The Maturation Process and the Facilitating Environment* (1965), postulated a “potential space” between the significant carers and the child, in which the carers “held” the child and thus fostered her development in a safe and constructive context. In this space, the “good enough” carers knew somewhere in themselves that healthy development was affective as well as cognitive, that the child needed to be stimulated but also held and nurtured, and that they themselves needed to be present in that space just enough, and not too much, so as not to stifle the child’s spontaneity and creativity. This “negative capability” is a necessary quality, in carers of children at home as well as in other contexts, such as in school, to foster the capacity to learn on an autonomous basis. This assumption that schooling should foster the autonomy of the learner, even as it provides a safe and stimulating environment to nurture growth, is at the centre of my teaching philosophy, and implies the involvement not only of the teacher in the classroom, but also in other significant social contexts of which the learner is a part: the school outside the classroom, the family, and other community groups.

The social context is central to learning because it is through discourse with others that our actions, emotions, and ideas are socially validated. A child will feel included in a social context in which he can recognize his culture: ways of being, acting, and expressing himself. A person continually redefines her identity in response to social contexts, and so a continued sense of inclusion, caring, and belief in the person’s capabilities is important for the development of a strong validation of the self, including a belief in one’s own capacities. If the aim of mass schooling is to prepare children for a fulfilling life beyond school, then providing a vibrant and inclusive social context at school, in association with the local community, in which children feel held and secure, and in which their cognitive and affective needs are recognized and met, is of the highest importance. As for teacher qualities, along with negative capability, a nuanced sense of children’s affective needs, so that even the restless children can be included wherever possible, needs to be developed. Academic excellence is built upon a professional and caring relationship with the teacher, in which the teacher communicates to the learner that the teacher not only has high expectations of the learner and believes that the learner can meet them, but also that the teacher has her in mind, is thinking about the next steps in the learning process, and is thinking about her educative well-being generally. In this context, a person can continually construct and re-construct meaning from the teaching materials, discussion with classmates, past experience, and inner dialogues from past mentors, as they progress through the schooling experience. Moreover, a child who is developing their autonomy in a secure and validating context that is inclusive and supportive of academic achievement will bring these values to impact on the future of his society (teacher, personal communication).

Of course, the teacher’s work is at an embryonic stage of development but it shows exciting possibilities and promise. The teacher has reached out to parents and community members to better understand the concept of what is essentially a rights - based Aboriginal education pedagogy demonstrating that it is possible to affirm students culturally while engaging in other school-based learning. The ex-Principal of the school where the teacher is based, shared with me after a visit he made to the school and community that the teacher has his students “eating out of his hand.”

This is further evidence of the supportive role that non-Aboriginal people can play in Aboriginal education especially when they not only acknowledge and value the culture and identity of the students but that they also view this as a vital, valued and respected aspect of knowledge creation to enhance Aboriginal student learning experiences and outcomes.

Culturally grounded education warriors

The late 1970s and especially the seminal years of the 1980s provide the foundations and impetus for the emergence of a burgeoning number of Aboriginal thinkers; education warriors who have strong and well-established connection with community, culture and country who build on the work of earlier generations to inform and define their advocacy for educational justice and change. Culturally grounded warriors are strongly steeped in their culture and identity and are challenging and reforming systemic Aboriginal education failure. The individual and collective success of these warriors, and others, will undoubtedly lay the foundations for future educational change and transformation for generations to come.

In early 1980, the National Aboriginal Education Committee (NAEC) developed its 1000 by 1990 strategy (1982). Underpinning the strategy was the belief that having more Aboriginal teachers in classrooms would improve Aboriginal student learning outcomes. There are now, reportedly, more than 1000 Aboriginal teachers in the classrooms in NSW but improved Aboriginal student outcomes, with some marginal exceptions, remain largely aspirational.

Acknowledged as “the first of its kind,” the NSW Aboriginal Education policy was developed and released in 1982. The policy was largely the result of the tireless advocacy and work of the NSW Aboriginal Education Consultative Group (NSWAECG). The policy addressed many of the matters raised in this paper but 40 years after the release of the policy Aboriginal students are still being failed by the Eurocentric education system even though it is acknowledged that the little progress that is being made occurs at torturous and glacial pace for both student and parents.

Final words

Performance data emanating from the Closing the Gap (CTG) national agenda reports illustrate that after more than a decade since the introduction of the CTG, despite the strategy being underpinned through the allocation of a significant level of support funding, the evidence points to a situation where little meaningful progress is being made against identified CTG targets, including education.

Aboriginal people have always viewed education as a means of overcoming poverty, alienation and oppression, but education has never been viewed as a site, nor a process to surrender Aboriginal identity and culture.

The education experiences and outcomes for the majority of Aboriginal learners remain basically devoid of culturally affirming and responsive content and dominant culture Aboriginal education policies and programs, including instruction and assessment methods, with some minor exceptions, remain primarily aligned to pernicious assimilation objectives. These assimilation objectives are destructive when considered individually but their potency grows exponentially when considered collectively, especially when they are infused by racism and notions of deficit, whiteness and white fragility.

Continuing systemic Aboriginal education failure is evidence of an urgent need to develop a model of a rights-based education, one that is based on the principles contained in the UN Declaration on the Rights of Indigenous Peoples and the Coolangatta Statement. The provocations outlined by Rigney align with the voices of Aboriginal parents and communities and must inform and shape the development of this rights-based model of Aboriginal education.

Many years ago I was asked the question; "What do Aboriginal People want from Education?" I responded by saying:

Knowledge and skills to live life in safety and with dignity. To embrace, through language and knowledge systems, the teachings embedded in Country. To heed the lessons derived from timeless song-lines, oral traditions and cultures. To know, appreciate and celebrate the interconnected symmetry of all creation and the teachings that have sustained Aboriginal peoples since the beginning of time.

I remained firmly connected to this view because, as I see things, these values and principles were embedded in traditional Aboriginal knowledge systems and form the basis of a more empowering rights-based curriculum and pedagogy. Such a model will positively engage Aboriginal students affording them opportunities to achieve progressive learning experiences and outcomes while staying culturally grounded and connected to country and community.

The more I ponder the question of culturally affirming and intellectually enriching education for Aboriginal learners, the more I am convinced that the fullness of its promise will only be realized when a new pedagogy of engagement is designed and developed by Aboriginal people and their allies.

Cultural affirmation and survival are fundamental to the education journey of Aboriginal students but this journey should not be interpreted as a total rejection of the knowledge and lessons that resides in the non-Indigenous world. Rather it is an argument that cultural affirmation and survival should be viewed and embraced as foundational and upon which academic achievement should be built.

At its core, Aboriginal education should be viewed both as a means of cultural survival and an expression of a long denied right to exist and celebrate life as culturally distinct peoples.

It would be easy for Aboriginal people to simply surrender to the assimilationist and social reproduction imperatives embedded in Eurocentric education but history and experience clearly shows that even in the face of overwhelming odds, there has been one constant, the survival of Aboriginal peoples as culturally distinct, adaptive and resilient people. Surrender is not an option.

Perhaps the scars of history are too great and the degree of spiritual fatigue is too immense making the struggle for Aboriginal education reform arduous and elusive and possibly the best that can be hoped for is a benign form of assimilation and an accommodation of difference wrapped up in a not too subtle veil of hegemonic conformity.

To withstand hegemonic conformity and the debilitating experience of education failure, crippling juvenile incarceration and devastating youth suicide rates, and other life limitations, we must, as Aboriginal people take greater control of the education in which our children are enrolled. A rights-based education pedagogy incorporating the provocations identified by Professor Rigney above, and other matters, must be the basis upon which greater Aboriginal control of education is exercised.

References

- American Indian Policy Review Commission. United States, 1976. American Indian Policy Review Commission: Final Report Submitted to Congress. G.P.O. Washington, U.S.
- Arbon, V., 2008. Knowing from where?. In: *History, Politics and Knowledge : Essays in Australian Indigenous Studies*. Australian Scholarly Publishing, North Melbourne, Vic, pp. 134–146.
- Archibald, J.A., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. UBC Press.
- Austen, I., 2021. 'Horrible History': Mass Grave of Indigenous Children Reported in Canada. New York Times. Available at: <https://www.nytimes.com/2021/05/28/world/canada/kamloops-mass-grave-residential-schools.html>.
- Australian Productivity Commission, 2020. *Overcoming Indigenous Disadvantage: Key Indicators 2020*. Available at: <https://www.pc.gov.au/research/ongoing/overcoming-indigenous-disadvantage/2020>.
- Battiste, M., 2000. Maintaining aboriginal identity, language and culture in modern society. J. Barman, Y. Hebert, & D. McCaskill. In: Barman, J., Hebert, Y., McCaskill, D. (Eds.), *The Legacy of the Past: An Overview, Indian Education in Canada, Volume 1: The Legacy*. UBC Press, Vancouver, pp. 1–22.
- Blair, N., 2015. *Privileging Australian Indigenous Knowledge: Sweet Potatoes, Spiders, Waterlilies & Brick Walls*. Common Ground Publishing.
- Bourdieu, P., Passeron, J.-C., 1977. *Reproduction*. R. Nice, Trans. In: *Education, Society and Culture*. Sage, London and Beverly Hills.
- Bourdieu, P., and Passeron, J.-C., 1977. *Reproduction in education, society, and culture*. Sage, Beverly Hills, CA.
- Boyer, E.L., 1989. *Tribal Colleges: Shaping the Future of Native America*. Carnegie Foundation, New York.
- Cardinal, H., 1969. *The Unjust Society: The Tragedy of Canada's Indians*. M. G. Hurtig.
- First Nations National Constitutional Convention, 2017. *Uluru : Statement from the Heart*. Available at: <https://ulurustatement.org/>.
- Frankl, V.E., 1938/1992. *Man's Search for Meaning*. Beacon Press.
- Hampton, E., 1995. Towards a redefinition of Indian Education. In: Battiste, M.A., Barman, J. (Eds.), *First Nations Education in Canada: The Circle Unfolds*. UBC Press, Vancouver, pp. 5–46.
- Kovach, M., 2009. *Indigenous Methodologies: Characteristics, Conversations and Contexts*. University of Toronto Press.

- Morgan, B., 2019. Beyond the guest paradigm: eurocentric education and aboriginal peoples in NSW. In: McKinley, E., Smith, L. (Eds.), *Handbook of Indigenous Education*. Springer, Singapore. Available at: https://doi.org/10.1007/978-981-10-3899-0_60.
- Morgan, B., 2020. In NorDocs, northern rivers medical network. *Spiritual Fatigue* 10–11.
- Oleska, S., Dauenhauer, R., 1982. Education in Russian America. In: Stein, G. (Ed.), *Education in Alaskas Past*. Anchorage. Alaska Historical Society.
- Petten, C., 2017. Harold Cardinal Exposed White Paper's "thinly Disguised" Program of Extermination. *Windspeaker.com Archives*.
- Pratt, R.H., 1892. The advantages of mingling Indians with whites. In: Paper Presented at the National Conference of Charities and Correction, at the Paper Presented at the Nineteenth Annual Session of the National Conference of Charities and Correction.
- Rigney, I., 2020. Aboriginal child as knowledge producer: bringing into dialogue Indigenist epistemologies and culturally responsive pedagogies for schooling. In: Hokowhitu, B., Moreton-Robinson, A., Smith, L.T., Andersen, C., Larkin, S. (Eds.), *Routledge Handbook of Critical Indigenous Studies*. Routledge, Abingdon. Routledge Handbooks Online. (Accessed 10 June 2022).
- Rigney, L.I., 2001. A first perspective of Indigenous Australian participation in science: framing Indigenous research towards Indigenous Australian intellectual sovereignty. *Kaurna Higher Education Journal*. 7.
- Royal Commission into Aboriginal Deaths in Custody, 1991. *Royal Commission into Aboriginal Deaths in Custody National Report*. Australian Government, Adelaide.
- Smith, G.H., 2015. Chapter three: equity as critical praxis: the self-development of Te Whare Wānanga Awanuiārangi. *Counterpoints* 500, 55–77. Available at: <http://www.jstor.org/stable/45178204>.
- Steering Committee for the Review of Government Service Provision (SCRGSP), 2009. *National Agreement Performance Information 2008: National Education Agreement*. Productivity Commission, Canberra.
- Steering Committee for the Review of Government Service Provision (SCRGSP), 2010. *National Agreement Performance Information 2009: National Education Agreement*. Productivity Commission, Canberra.
- United Nations, 2007. *United Nations Declaration on the Rights of Indigenous Peoples*. United Nations, New York.
- Watson, I., 1998. Power of the Muldarbi, the road to its demise. *Aust. Fem. Law J.* 11 (1), 28–45.
- World Indigenous People's Conference on Education (WIPCE), 2006. The Coolangatta statement on indigenous peoples' rights in education. In: Read, P., Meyers, G., Reece, B. (Eds.), *What Good Condition?* (Vol. Aboriginal History Monograph 13. ANU Press, pp. 229–236.

Impact of the 1971 Alaska Native Claims Settlement Act on education: 1971–2021

Maria Shaa Tláa Williams (Tlingit)^a and Aleesha Towns-Bain (Alutiiq)^b, ^a University of Alaska Anchorage, Anchorage, AK, United States; and ^b Bristol Bay Native Corporation Education Foundation, Anchorage, AK, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	22
Overview of the geography and diversity of Alaska's Indigenous peoples	22
Colonial periods	23
Traditional education	24
Westernized education 1867–1971	24
Modern American period: 1959–present	24
Formation of Native Association	25
1971 Alaska Native Claims Settlement Act	25
Native sovereignty and Western education	26
Who is served—eligibility	27
ANSCA Education Consortium	27
Impact	28
Case study: Bristol Bay Native Corporation education foundation	28
A coalition of funders	29
Challenges and opportunities—looking ahead	29
Conclusion	30
References	31
Relevant websites	31

Introduction

The topic of Indigenous education is critically important in the 21st century. Alaska has a large Indigenous population and the Alaska Native regional corporations are the state's number one employer. The 2020 US Census data shows that the Alaska Native/American Indian population makes up 21.9% of the overall population of Alaska. This is the highest Indigenous demographic in the entire United States.¹ The Alaska Native population is also a very young population, with 28,620 Pre-K-12 AN/AI enrollment in 2020–21, which is 22% of the entire PreK-12 demographic; the highest in the nation.² The University of Alaska system, which includes three separately accredited universities (University of Alaska Anchorage, University of Alaska Fairbanks, and University of Alaska Southeast), Alaska Native students make up the highest “minority” population in their undergraduate programs, averaging 10–17% of the overall student body.

This article provides a brief overview of the colonial history of Alaska, with a specific focus on Western education, and recent initiatives that illustrate the amount of resources that the Alaska Native corporations and partner non-profits, as well as other tribal entities invest into educating our young population.³

Overview of the geography and diversity of Alaska's Indigenous peoples

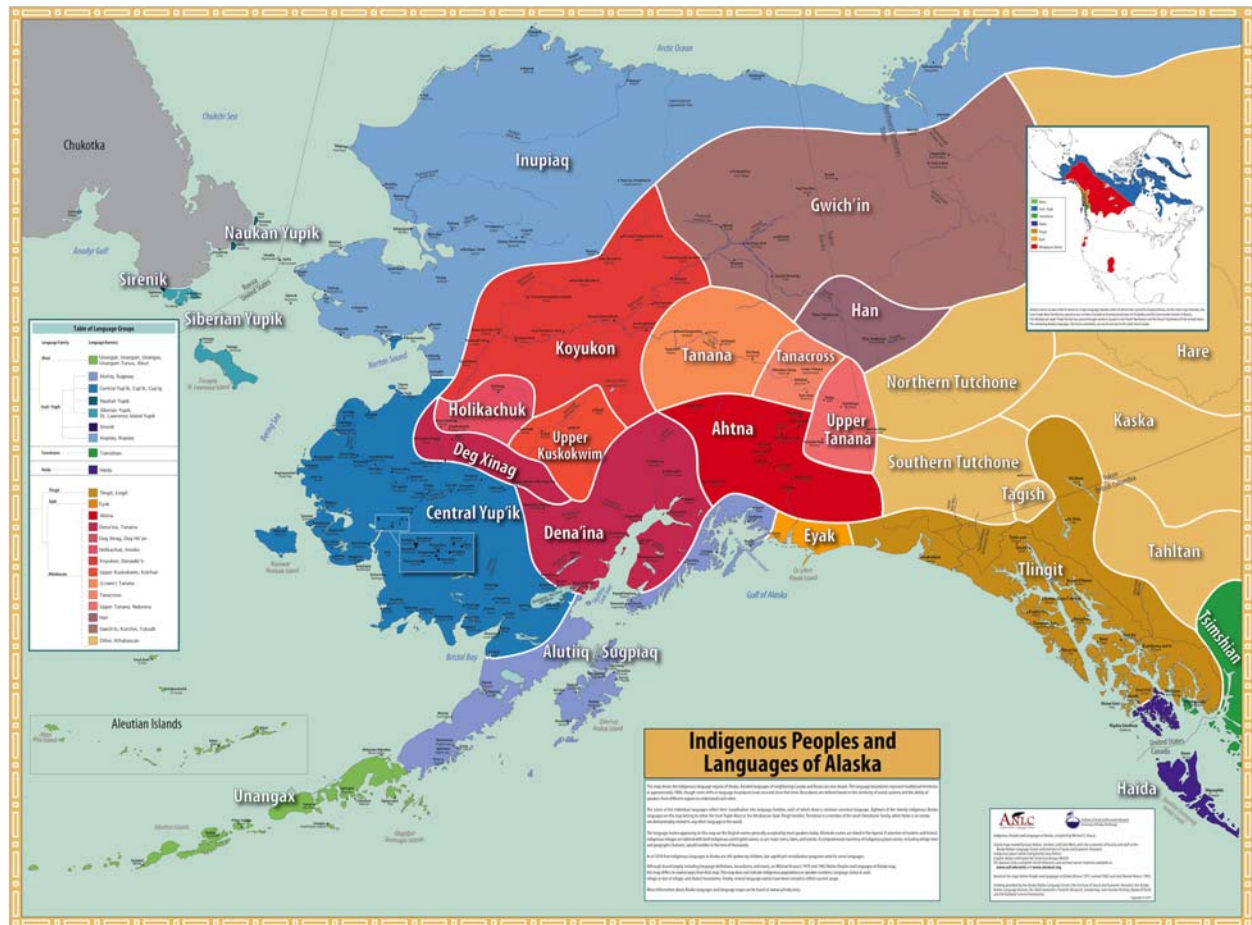
Alaska is a subcontinent of North America and encompasses over 570,00 square miles (1,717,856 km), has over 33,000 miles (53,108.352 km) of coastline and is in the Arctic and sub-Arctic regions of our planet. Alaska has six major geographic zones (North Slope, Interior, Aleutian Islands, Southwest, Southeast, and Southcentral). Each of these broad geographic zones include a wide array of flora and fauna and unique ecosystems, and geographic wonders. With over 17 mountain ranges, more than 3 million lakes, and major rivers such as the Yukon, Kuskokwim, Kobuk, Tanana, Susitna, and Copper Rivers—in addition to hundreds more, the diverse landscape of Alaska is matched only by the diversity of the Indigenous peoples who have inhabited it for over 20,000 years.

¹The total American Indian/Alaska Native (AI/AN) alone or in combination population is 160,287. The Census allows citizens to self identify as AI/AN alone, or AI/AN in combination with one or more races. (National Congress of American Indians Policy Research Center, Research Policy Update: A First Look at the 2020 Census American Indian/Alaska Native Redistricting Data, August 13, 2021).

²State of Alaska Department of Education & Early Development Statistics & Reports (<https://education.alaska.gov/data-center>).

³The Alaska Native Corporations are mandated by Congress to provide economic benefit to their shareholders. The corporations have succeeded this arena, however it is important to note that ANSCA did not mandate the funding of post-secondary education or cultural preservation, but the corporations have invested heavily in education and cultural programs.

The Indigenous peoples of Alaska live in all parts of Alaska and developed unique lifeways that reflect the environment they live in. There are over 20 different Alaska Native languages, most of which are still spoken, although they are considered endangered due to colonial practices, including boarding schools.



Alaska Native Language Map. <https://www.uaf.edu/anla/collections/map/anlmap.png>.

Colonial periods

The name of Alaska is derived from the Unangam Tunuu (Aleut) word *Alax̂sxâx̂*. The Aleutian Islands were the first part of Alaska that experienced colonial invasion beginning in the mid-18th century from the Russian Imperial Government. The global commodity during this period of history was fur, and Alaska had abundant fur animals, including the sea otter. The Russian Fur hunting companies arrived in droves once they discovered sea otters, and they enslaved the Unangax̂ (Aleut) and Suqpiq (Alutiiq) peoples so they could obtain fur. Russia claimed this region in 1754 and expanded their empire into southeastern Alaska and parts of the interior. Other European explorers (Cook, La Perouse, Arteaga and Bodega y Quadra, etc.) began to map out Alaskan coastal areas, claim them for their respective countries, and re-name sites. The United States began to grow by leaps and bounds in the 19th century and Russia was bankrupt by the end of the Crimean War, and the two nations negotiated the Treaty of Cession.⁴ In 1867 the Russian Imperial government ceded the right to trade in the Alaskan territories to the United States and the American colonial period began.

The American colonial period was marked by continued resource exploitation (Fur, gold, coal, salmon, timber, whale oil, etc.) and the Indigenous peoples were subject to new forms of colonization, which included a strong Christian Missionary presence (Williams, 2009), establishment of boarding schools, which forcibly removed children from their homes and families in order to facilitate their “americanization” and assimilation into the United States. Additionally, terrible epidemic diseases were introduced, beginning in the Russian era and continuing to the mid-20th century (smallpox, influenza, STD’s, tuberculosis, etc.) which had

⁴Major expansion of the continental United States occurred in the 19th century with the following purchases: Louisiana Purchase in 1803, Florida Treaty in 1819, Oregon Treaty in 1846, Texas Annexation in 1845, Treaty of Guadalupe/Hidalgo in 1848.

resulted in devastating loss of life and effects on the primarily hunting and fishing based economies. The population of Alaska Native people by 1950 was a little over 33,000 (alaskool.org) and most lived in over 200 villages across the entire region.

Traditional education

Education and knowledge have historically been highly valued by the indigenous peoples of Alaska, and the education of children and youth was integrated into life through oral storytelling, cultural values and behaviors, and daily practices. For example, in the Yup'ik tradition, the goal of this education was *Yuyyaraq*, the way of the human being, as described by [Napoleon \(1996\)](#). Education determined how individuals should interact with each other, their environment, the spirit world, and outlined protocols for all situations that a person would experience in their lives. Technical skills were taught in an experiential manner through apprenticeships in which a young person would observe and emulate an expert. Some skills taught were gender specific to create a functioning household and community ([Drabek, 2012](#)).

Westernized education 1867–1971

Non-Indigenous education was first introduced during the Russian colonial period and the first schools were established by the Russian Orthodoxy in the early 1800s. These schools were primarily in the coastal areas where the Russian American Company had control. The Alaska Native people who were enslaved by the Russian Fur hunters during the most brutal period (1754–1819) converted to Russian Orthodoxy and learned to read and speak Russian. The Russian Orthodox Priests attempted to protect the Indigenous people from the horrible treatment of the Russian Fur Companies. The kindness perhaps made conversion a protection against the more evil actions of the companies. After 1819, the American colonies came under the control of the Russian Navy and a more humane and structured approach was adopted. Slavery disappeared along with the more brutal aspects of conquest. Education was very important and there are examples of Indigenous students furthering their education in Russia. Specialized training such as navigation, map making were viewed as critical to making Alaska a viable and sustainable Russian colony.

In 1867 the Russians sold their outposts and the territory they claimed they owned to the United States under the Treaty of Cession. The Americans had a very different approach to colonization than the Russians. The government allied itself with various Christian denominations to convert all Alaska Natives to Christianity and to assimilate them. Education was a colonial tool and this led to the establishment of boarding schools for Indigenous youth all over the US and Canada, and Australia.

The western schools attempted to abolish traditional knowledge systems, including language(s). In the western school environment, children were forbidden to speak their Native languages and were punished for expressing any elements of their culture. Children were instructed about the necessity to become like white people as the essential element of their success. In addition to formalization of education, in some parts of Alaska, separate schools were created to educate White and Native children, predominantly in the Southeast. "Through a combination of laws and informal practice, segregated schools even within small communities became the norm. By 1929–30, segregated schools existed in at least 19 communities—white schools and native schools in the same tiny villages. In most of the communities, the native school stopped at the eighth grade while the white school went through grade 12," ([Cotton, 1984](#)).

By the 1960s 50% of the Alaska Native population had no more than a 6th grade education. 21% completed 7th and/or 8th grades, 14% went into 9–12th grade, but only 8% of that group graduated with a high school diploma. 2% attended a college, but less than 1% completed four or more years of college ([Arnold et al., 1968:17](#)). There were no high schools in most of the 200 villages in Alaska. Western education came in the form of Bureau of Indian Affairs (BIA) schools, state run boarding schools, which removed children from their rural villages to larger urban areas (Kodiak, Nome, Anchorage, Fairbanks) from 6th to 12th grade. The state also created regional boarding schools in Nome and Kodiak. These were disastrous and destructive to the young Indigenous students who were forced to leave their homes. This changed due to student activism and a major lawsuit. Often called the "Molly Hootch" case, *Tobeluk vs. Lind* became a landmark in education litigation and changed middle and high school education throughout the entire state for Indigenous students. 27 high school students in rural Alaska sued the State Commissioner of Education because they were forced to leave their home communities and attend boarding schools far from their homes. In 1976, the case of *Tobeluk vs. Lind* was settled by providing the establishment of high schools in 126 villages. Anna Tobeluk and the other 26 students' lawsuit revolutionized the delivery of secondary education in rural Alaska.

Modern American period: 1959–present

Alaska became the 49th state in the union in 1958/59 primarily due to the discovery of the massive oil fields in the North slope. The newly formed State of Alaska and the US Federal government began doing land selections and did not include Alaska Native people. The lands of the Indigenous peoples were being claimed by the State of Alaska and Federal Government. Alaska Native people in over 200 villages throughout the entire region were faced with losing their ancestral lands—access to their hunting and fishing and gathering areas. Additionally, the newly formed State of Alaska and Federal governments initiated major projects that would have detrimental effects on numerous Alaska Native communities. These included the Rampart Dam project, which proposed to create

a recreation lake larger than Lake Erie by damming the Yukon river and flooding the Rampart flats—which is a high yield subsistence area and home to several Dene/Athabaskan villages. The U.S. Defense Department created an Atomic Energy Commission that was planning on detonating atomic bombs in the coastal Arctic region of Alaska near several Inupiaq villages. Both of these projects would have destroyed hunting and fishing areas, as well as villages. No Alaska Native people were consulted and this led to what is now called the Land Claims movement, or Native Solidarity Movement (Williams, 2009:202–216). This led to a period of intense political activism, and the establishment of the *Tundra Times* newspaper in 1962. The *Tundra Times* became a vital conduit of information that was shared throughout all the various cultural groups and villages in Alaska and brought many groups together to fight and organize in order to protect their lands.

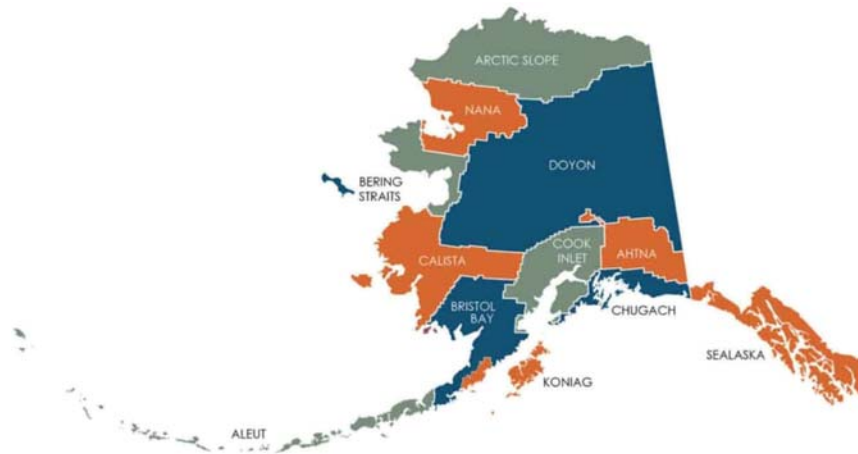
Formation of Native Association

The formation of Native Associations in the 1960s was a way in which each broad region could begin to patent their lands, and initiate their own land claims. This was a highly effective vehicle and each Association was roughly formed around pre-existing culture areas. These early associations began to patent land so they could protect their land base. In 1966 they formed an umbrella organization called the Alaska Federation of Natives. This became the political and lobbying arm and led to a land freeze. The State could not select any lands until there was a land settlement.

1971 Alaska Native Claims Settlement Act

On December 17th, 1971 President Richard Nixon signed the Alaska Native Claims Settlement Act (P.L.92-203). This was the largest land settlement in the history of the US Congress and represented sweeping new changes to previous treaties and agreements with American Indians. ANCSA had both a land settlement and an agreed upon sum of lands lost—so it included a land settlement of 44 million acres and over \$900 million land settlement for lands lost. The other unique feature of ANCSA was the establishment of 12 regional corporations and almost 200 village corporations. It was a social engineering approach designed to create a capitalist economy and a more western-based economic structure.

The Native associations formed in the 1960s were empowered to enroll shareholders, locate office space, elect boards of directors, and create a sustainable economic base. ANCSA, Section 7(a) identified these associations. “For purposes of this Act, the State of Alaska shall be divided by the Secretary within 1 year after the date of enactment at this Act into twelve geographic regions, with each region composed as far as practicable of Natives having a common heritage and sharing common interests....”



The Alaska Native Corporations and their respective regions. Courtesy of ANCSA Regional Association.

Within a 2 year period, the various corporations created enrollment mechanisms for their constituencies, and became incorporated. Each of the main regional corporations, and the village corporations had unique pre-existing histories and cultural practices. Also, each of the geographic areas had various resources that could be developed (timber in the rainforests of Southeast Alaska; the oil and gas in the North Slope and Cook Inlet area; mining such as zinc and copper, fisheries, etc.) in order to make profits and pay dividends to their shareholders. Each region also had shareholder enrollments that varied from a little over 1000 to almost 17,000; additionally the villages in each region varied from a handful to over 56.

<i>Regional corporation</i>	<i>Shareholders in 1974</i>	<i>Shareholders 2020</i>
Ahtna, Inc.	1092	2000+
Aleut Corporation	3353	4000+
Arctic Slope Regional Corporation	3906	13,000+
Bering Straits Native corporation	6916	8000+
Bristol Bay Native Corporation	5517	10,000+
Calista Corporation	13,441	33,700
Chugach Natives, Inc.	2099	2200+
Cook Inlet Region, Inc.	6243	9100
Doyon, Limited	9221	20,000+
Koniag	3340	3800+
NANA	4905	14,500+
Sealaska Corporation	16,493	23,000+

Source: Arnold (1976) & ANC websites in 2021.

Native sovereignty and Western education

Many of the architects of ANSCA were among the few Alaska Native people who had the opportunity to achieve Western higher education degrees, and/or were in military service in World War II and the Korean conflict (The military provided an avenue for providing knowledge on western practices, as did higher education). Early leaders such as Eben Hopson of the Arctic Slope region focused on the importance of education. He had an activist approach in reclaiming Inupiat control of the Arctic Slope area, and was one of the founders of the Arctic Slope Native Association, which later became the Arctic Slope Regional Corporation after the 1971 Alaska Native Claims Settlement Act. He identified traditional education vs. what he termed “coercive assimilationist educational era” in his Mayor’s address on December 19, 1975 (<http://ebenhopson.com/eben-hopson-on-inupiat-education/>). He and other leaders in the early 1970s focused on gaining more power and control in the arena of education, which was squarely in the hands of state run boarding schools.

The newly formed Alaska corporations, and the broader Native community felt that education and/or training beyond the secondary level was key to building a foundation for a solid financial livelihood, community stability and capacity for positive societal impact. Put simply, they recognized Native sovereignty in this time-period required Western education. The leadership also recognized that the western educational system could be a tool that could promote a more robust Indigenous community.

By the mid to late 1980s, the number of Alaska Native students graduating from high school began to increase and put post-secondary options in their sights. At the same time, with the creation and growing success of the Alaska Native corporations, a new suite of professional positions began to open. Suddenly, highly skilled professional positions were in demand—from accountants to attorneys, human relations specialists to C-suite executives, along with a wide variety of technical positions in industries as varied as forestry and oil. Locked out of professional fields and trades for so long, Alaska Native corporate board leaders believed first and foremost that those positions should be held by Alaska Native people—and if they did not have the training to hold those roles that it would pay for them to receive education.

Starting as early as 1977, each of the 12 regional corporations established entities to support activities to preserve its Alaska Native cultural heritage and/or provide higher education support for Alaska Native students. The corporations each provided significant financial resources and operational support to these newly formed nonprofit entities, in the form of endowments or annual cash donations.

Known colloquially as the “Alaska Native education foundations”, the 12 entities were organized differently, at separate times, and remain highly fluid. Some are private foundations, while others are nonprofit organizations, still others are organized as programs that are administered by the corporate offices. While the structures are different, remarkably, the overall purposes and missions of the organizations are similar. Most entities adopted missions to specifically provide scholarship funding and other educational support for students to pursue post-secondary credentials and degrees, and to promote and preserve the Alaska Native cultural heritage of their respective region.

Establishment of Foundations⁵

Arctic Education Foundation (1977)

Sealaska Heritage Institute (1980)

The CIRI Foundation (1982)

The Ahtna Heritage Foundation (1988) -Scholarship program now operated by corporation

⁵When data was unavailable from the entity, authors relied on listings available on public charities at [propublica.org/nonprofits/](https://projects.propublica.org/nonprofits/). In some cases, dates may reflect the state of IRS determination, and not the entity’s creation by founders.

Aleut Foundation (1996)
 The Koniag Education Foundation (1993)
 Robert Aqqaq Luk Newlin Sr. Memorial Trust (2014)
 The Beringia Settlement Trust (2018)—(formerly the Bering Straits Foundation—1991)—scholarship program now operated by Bering Straits Native Corporation
 Bristol Bay Native Corporation Education Foundation (1992)
 Calista Education and Culture, Inc. (2014, formerly Calista Heritage Foundation, 1980)
 The Chugach Heritage Foundation (1985)
 Doyon Foundation (1989)

Since inception, each organization has taken on specific programs and activities in response to community needs. The first private foundation was established by Arctic Slope Regional Corporation in 1977. The Arctic Education Foundation was created to prepare Iñupiaq students to enter the professional workforce in any job, particularly those required in the North Slope region ([Arctic Education Foundation, n.d.](#)). Sealaska Heritage Institute was created specifically to focus on perpetuating Tlingit, Haida and Tsimshian cultures of Southeast Alaska, with an additional focus on scholarships and Western education ([Sealaska Heritage Institute, n.d.](#)). The CIRI Foundation, based in Alaska's largest city of Anchorage, adopted a two-fold mission to promote individual self-development and economic self-sufficiency through education and to maintain pride in Alaska Native culture ([The CIRI Foundation, n.d.](#)).

As the organizations have matured, they have expanded to offer a variety of programs. The foundations have an impressive array of language and cultural materials, sponsor major dance festivals, and other types of educational initiatives, including publications. For example, the Chugach Heritage Foundation runs a very successful culture camp for youth and elders in Prince Williams Sound, hosts an annual Russian New Year celebration and maintains a robust scholarship program for students seeking college and university degrees, as well as vocational certificates. Calista Heritage Foundation was established in 1980 to fund scholarships for Calista Corporation shareholders and descendants. In 2014, it merged with the Calista Elders Council to create Calista Education and Culture, a nonprofit organization ([Calista Education and Culture, Inc., n.d.](#)). The organization is now active in cultural heritage work, including impressive language and cultural documentation of Yup'ik, Cup'ik and Dene/Athabascan ways of life, and continues its scholarship work to benefit students seeking college and vocational education. Similarly, Doyon Foundation has worked extensively in preservation and revitalization of the Dene/Athabascan languages and Iñupiaq, in addition to scholarships for students ([Doyon Foundation, n.d.](#)).

Who is served—eligibility

Most of the entities created by the Alaska Native Corporations initially limited their eligibility to serve a specific group of people—Alaska Native shareholders of each of the corporations. In order to become a shareholder of an Alaska Native Corporation, individuals had to meet a requirement of one-fourth blood quantum of Alaska Indian, Eskimo or Aleut blood, and were born on or before 11:59 p.m. on December 18, 1971. As a result, original shareholders are now 50 years old. The US Congress passed major amendments to ANSCA, which allowed corporations to amend their articles of incorporation to expand shareholder eligibility and allow for the enrollment of shareholder descendants and the gifting of shares to subsequent generations. It was up to each corporation to decide how and when to enroll descendants and offer services to this large, and ever growing, group. Similarly, the education entities also have been tasked with determining eligibility. Some entities have opted to serve descendants, and some still require students to become shareholders before they become eligible. Most entities require some tangible tie to an original shareholder, such as proving a student is a direct lineal descendant.

ANSCA Education Consortium

In the year 2000, the education entities began to meet regularly and formed a loose-knit association to be able to share information and ideas across the regions. The ANSCA Education Consortium (AEC) exists with a purpose of Alaska Native Education and Cultural advocacy and operates under the values of sharing for success, unity and the potential of our Alaska Native people. Led by annually elected co-chairs, the group operates under the mission that every Alaska Native person wishing to pursue postsecondary education should have the means to obtain it. Originally, its membership included the 12 Education nonprofit, foundation and programs created by 12 regional corporations. The Consortium now also includes scholarship entities created by village corporations and those organizations/entities that are invested in Alaska Native student success. In addition to the education and career-development of Alaska Natives, members of the ANSCA Education Consortium also focus on language revitalization, culture and tradition, early education, teacher training, and preservation and perpetuation of Alaska Native art forms.

Impact

To date, a comprehensive evaluation of the overall historical impact of the AEC organizations has not been completed, nor has one that more narrowly focuses on their significant contributions to supporting scholarships of Alaska Native students. This is an under-researched area. However, a number of professional reports offer a glimpse of their impact, and a larger scale report is planned.

According to the ANCSA Regional Association's website "Between 1977 and 1993, each of the twelve regional corporations chose to provide the initial financial investment to create separate non-profit foundations to support the advancement of the educational endeavors of their shareholders and descendants. To date, the twelve education foundations have awarded more than 54,000 individual scholarships to Alaska Native shareholders and descendants, totaling more than \$100 million for Alaska Native people to pursue various educational endeavors." ([ANCSA Regional Association, n.d.](#)) In all likelihood, the number of scholarships awarded and individual students impacted is under-reported as the independent nature, long lifespan, and capacity of the entities has impacted the ability to collect data from each. The private foundations report significant contributions on their tax filings with the IRS each year, for both scholarships, grants and related operating costs. In 2019, AEC organizations designated as private foundations by the IRS reported significant charitable disbursements totaling more than \$15.3 M ([Propublica, 2021](#)).

On websites and in correspondence, the AEC organizations offer a glimpse into their investments over time. The Arctic Education Foundation reports it has awarded \$39.2 M in scholarships since 1977. The CIRI Foundation reports that since 1982 it has contributed more than \$39.6 million for post-secondary scholarships (totaling 17,859 awards), and an additional \$3.7 million for its cultural heritage and education program ([Morris, 2021](#)). Sealaska Heritage Institute and Sealaska Corporation indicate that they have awarded approximately 11,000 scholarships totaling \$20 million in their 40-year history ([Sealaska Heritage Institute n.d.](#)). Since 1991, Bering Straits Native Corporation says it provided more than \$4 million in funding to support the postsecondary educational goals of BSNC shareholders and descendants ([Bering Straits Native Corporation, 2021](#)). Calista Education and Culture similarly reports it paid 2706 scholarships worth \$5.85 million to its shareholders and descendants ([Bavilla, 2021](#)). The Chugach Heritage Foundation indicates it has paid \$10.6 million to 2200 scholarship recipients in its history. Koniag Education Foundation reports that it has awarded over \$7 million in scholarships to Alutiiq students via 3000 individual scholarship awards ([Kreger, 2021](#)).

Case study: Bristol Bay Native Corporation education foundation

After more than a decade of providing scholarships to students via its corporate philanthropy, the Bristol Bay Native Corporation Education Foundation (BBNCEF) was created by Bristol Bay Native Corporation (BBNC) in 1992 as a private foundation with the mission to support BBNC shareholders with higher education and vocational scholarships ([McDowell Group, 2019](#)). In 2019, BBNCEF launched a comprehensive evaluation of its scholarship programs that serve as an illustrative snapshot of the overall impact of AEC investments. It was a significant undertaking that included unearthing board meeting minutes, notes, scholarship files starting in the early 1980s, and collating them for the first time in an electronic database. The evaluation was conducted by the McDowell Group (now called McKinley Research), a respected Juneau, Alaska-based research group. The evaluation found that the Bristol Bay Native Corporation began awarding scholarships as early as 1984, 8 years prior to formally creating its private foundation (McDowell Group, August 2019). From 1984 to 2019, BBNCEF had awarded 3787 scholarships worth nearly \$6.4 million to its shareholders for both higher education and short-term vocational training.

The same evaluation examined outcomes for 942 individual recipients who were tracked in the National Student Clearinghouse Database (77.8% of the total number of identified recipients). The report showed that scholarship recipients enrolled on average for 12 semesters and would often attend more than one institution in multiple states. The research indicated that 39.3% of BBNCEF scholarship recipients had graduated with a degree and were not currently enrolled, 12.6% were currently enrolled, 9.4 had previously graduated and had re-enrolled, and 38.6% had no degree and were not currently enrolled.

By adding both graduate categories together, the study showed that 48.7% of Alaska Native students who had received a scholarship had achieved a post-secondary credential. When put against the backdrop of overall postsecondary attainment for Alaska Native students, it represents the significance of both the funding and importance of the connectivity to their regional corporation, community and cultural identity. For BBNCEF, it is also important to note that the vast majority of Alaska Native students who apply for scholarship funding were awarded during this time period— in some years, it's 100%. Students are not cherry picked for outcomes but represent all eligible students seeking to complete secondary education.

In addition, according to the NSC data, many recipients were awarded more than one degree or credential. Forty three percent of degrees awarded were Bachelors degrees, 25% were Associate, 13% Advanced.

Of the major fields—the most common were business, management, marketing, and related support services (16%), health professions and related programs (11%), liberal arts and sciences, general studies and humanities (11%), and education (10%). Most student recipients attended post-secondary school in Alaska, followed by Washington State and Colorado. Scholarship recipients attended the University of Alaska Anchorage (54%) and University of Alaska Fairbanks (48%), followed by University of Alaska Southeast (9%) and Fort Lewis College (7%).

As a result of the evaluation, BBNCEF has altered some of its offerings to become more student-centric to address needs identified by the data as well as key respondents in a series of qualitative interviews with regional leaders, elders and students. BBNCEF recently increased its award levels for vocational scholarships interested in pursuing career focused vocational training such as welding, mechanics, commercial driving, and aviation. BBNCEF has also opened an emergency aid scholarship as a way to encourage students to stay enrolled. Funds are available to pay for emergencies such as travel to home communities, vehicle fixes, child care, or medical procedures.

Increasingly, BBNCEF is understanding that a strong sense of cultural identity, established early, contributes to academic success. In addition, it recognizes that cultural wisdom and knowledge is a complete and intact knowledge system that is as significant and worthy as Western knowledge systems. Therefore, BBNCEF has opened a cultural heritage portfolio to offer grants to communities interested in teaching native languages, dance, crafts, tools and a variety of other critical skills. Most grants fund activities that support activities that develop self identity and awareness for students Pre-K to grades 12, in a variety of settings from traditional village-based cultural camps to web-based classes for young adults.

A coalition of funders

The focus on providing educational support for higher education for Alaska Native students has been shared by organizations across the State of Alaska. Many village corporations also created scholarship programs to serve Alaska Native village corporation shareholder students and descendants. Other entities, such as regional economic development corporations, regional nonprofit associations, and other businesses, also offer scholarship support to Alaska Native students from their respective regions. This article does not explore the significant impacts of those partner funding entities, particularly village corporations, and would be a rich area for further data collection and research.

Challenges and opportunities—looking ahead

While funding exists for Alaska Native students to achieve a post-secondary degree or certification, significant challenges exist. A large number of Alaska Native students choose to attend one of the three University of Alaska Universities (University of Alaska Anchorage; University of Alaska Fairbanks; University of Alaska Southeast), and affiliated community campuses. Additionally there is one tribal college, Ilisagvik in Utqiagvik, Alaska Pacific University (Anchorage), Outer Coast College (Sitka), and Alaska Christian College (Kenai). Each of the universities have established various Alaska Native student support centers, as well as academic degree programs in Alaska Native and Indigenous studies. There are student centered programs that provide advisement, guidance, and programming to ensure the students feel a sense of belonging. As previously mentioned, the non-profit organizations provide tuition support for language and arts courses, as well as involvement in establishing new curricular programs.

Within the University of Alaska system academic and student supports exist, but they have been chronically underfunded by the university and have not kept pace with the growing Alaska Native student population. The University has historically also not supported the retention and development of Alaska Native faculty or staff. As a result, Alaska Native faculty make up 3% of the overall faculty, yet Alaska Native people are 20% of the state's overall population. In recent years, the numbers of Alaska Native faculty has actually decreased.

<i>Alaska Native faculty demographics</i>			
	2019	2020	2021
UAA	21	11	10
UAF	18	15	12
UAS	4	4	3

<i>Alaska Native undergraduate graduation rates</i>	
UAA	10% Alaska Native graduation rate within a 6–10 year window ⁶
UAF	9.7% Alaska Native graduation rate within a 6–10 year window ⁷

These facts are disturbing and require attention of the Alaska Native community to address. The Alaska Native Studies Council, an active group of Alaska Native Faculty and Faculty that teach in Alaska Native studies met with community members and the University administration. This led to what is called the Alaska Native Success Initiative or ANSI. The UA board of Regents endorsed recommendations of the ANSI, which included rigorous recruitment and retention of Alaska Native Faculty. This was passed in June of 2021. It is too soon to say if this will yield any changes.

⁶UAA Institutional Research. <https://app.powerbi.com/view?r=eyJrIjoibDE3MjMzODk0YUlmMi00Zjc2LTgzY2YtODU3OGJjNTk3OWI1IiwidCI6IjVhNmRjMDU5LTUwNjctNDE2ZS1hZjgxLTRiNTc5ZDkwMDMyMyIsImMiOiZ9.>

⁷[https://www.univstats.com/colleges/university-of-alaska-fairbanks/student-population/.](https://www.univstats.com/colleges/university-of-alaska-fairbanks/student-population/)

Conclusion

Alaska Native people are strong, resilient and practical people who have always valued, and continue to value, the importance of education and knowledge. These communities and leaders created a unique social infrastructure in Alaska Native corporations in order to protect their Native way of life and create a basis for a capitalistic economy. By utilizing the profits of the corporations, leaders subsequently formed nonprofit organizations to further focus on missions of providing financial support for students for attending Western education institutions and protecting cultural knowledge. Despite suffering significant physical and emotional abuse and racism in Western education environments historically and in more recent times, Alaska Native leaders continue to see value in learning and providing opportunities for young people to be exposed to worldly knowledge and how skills can create better communities.

While the entities continue to grow and evolve, the next stage of work is to truly document and understand the significant impact of funneling millions of dollars of investment into Alaska's greatest resource—its people. The impact and support of the Alaska Native regional corporations is undeniable, although significantly under researched and largely unrecognized outside the State of Alaska. The existence of a coalition of entities focused on the tasks of providing higher education scholarship support for indigenous students and supporting cultural heritage efforts is likely unique. It may be impossible to document historical investments in education, but systems could be put into place now to gather data on an ongoing basis and utilize that information in order to become greater advocates for Alaska Native people at the local, state, and federal levels.

The Alaska Native corporations and AEC may seek to leverage their respective political and economic clout to improve the K-12 educational system in Alaska, and also for the University of Alaska system. The amount of dollars that the Corporations and their non-profits have invested in education is significant, and yielded thousands of graduates. How can the investments that the Corporations and foundations have leverage to insure more success? Perhaps in a shared 5–10 year plan to increase high school and college graduation rates. High School graduation rates for Alaska Native/Indigenous students remain lower than the non-Indigenous students. If they were able to exert pressure on the State Department of Education, allow tribal compacting of schools, and use their clout to pressure the University system with an approach that incorporates benchmarks and goals, and repercussions if the institutions do not reach these goals. Each of the individual foundations have distinct profiles. This is a great strength, but is also challenging when identifying a metric for evaluation on the effectiveness of their educational programs, and/or support for education via scholarships or having a shared approach in dealing with the State Educational System. However, it must be pointed out that one of the strengths of the 12 Alaska Native corporations and their non-profits is distinct approaches and unique profiles and connection to their specific land base and language and cultural practices.

What will Alaska Native education preK-12 and post secondary look like in the next 50 years? 2021 was the 50th anniversary of the Alaska Native Claims Settlement Act. 2071 will be the 100th anniversary—what will the corporations and their associated non-profits look like? What will a decolonized education look like in Alaska? How do we get to a decolonized educational system?

The need to improve educational opportunities at K-12 and higher education level is a topic of broad discussions within the Alaska Native community. We highlight possible pathways forward:

- (1) Educational institutions in Alaska would benefit from incorporating Alaska Native worldviews, cultures, histories and languages broadly into curriculum content, not limited to posters or special programs. The results would be a more effective, rich and place-based educational system for everybody, not just Indigenous students. The Alaska State Board of Education needs to officially adopt the Alaska Standards for Culturally Responsive schools developed by the University of Alaska Fairbanks in concert with Alaska Native educational leaders and statewide groups such as the Alaska Federation of Natives. These are comprehensive and serve as a blueprint for making successful changes (<https://uaf.edu/ankn/publications/guides/alaska-standards-for-cult/>).
- (2) Higher education institutions should provide free or reduced tuition for Indigenous/Alaska Native/Native American Students. If Alaska Native students could attend universities with free or reduced tuition, it would allow the current funds that the Alaska Native Foundations and non-profits spend on higher education be re-directed on establishing think-tanks and publishing curriculum that incorporate Alaska Native worldview, languages, cultures. The funds could also be directed toward the Pre-K-12 institutions, which are in great need of being infused with Indigenous Alaska Native history, language, culture, and science. Content standards need to be further developed and implemented.
- (3) Tribal Compacting: The Alaska Federation of Natives is advocating for tribal compacting of education. If this moves forward it is a step in the right direction of supporting an Alaska Native-centric educational system in the PreK-12 and higher education by allowing tribes and regional associations greater input into their local school districts. Joel Isaak (Dena'ina) is the Project Coordinator for the State of Alaska's Tribal Compacting program, which is in development and ready to be launched (<https://education.alaska.gov/compacting>). The State of Alaska awarded the Alaska Federation of Natives a \$1 million dollar grant to develop and design how tribal compacting will be instituted.⁸ Under the award, the Alaska Federation of Natives plans to:
 - Identify relevant federal and state laws and regulations;
 - Compile relevant material related to educating Alaska Native students;
 - Engage with Alaska Native leaders and subject matter experts on educational compacting options;
 - Design of a comprehensive path forward;
 - Design seminars for educators and administrators on the why and what of compacting;

⁸The award period ended in July 2022 and was not available by the time this article was complete.

- Participate in discussions with the State of Alaska and various committees of the Alaska State Legislature to respond to questions on the final pathway report.
- (4) The State of Alaska should establish a College of Alaska Native Studies, which could create more opportunities for Alaska Native students. Similar to a tribal college, but with a specific focus on the current needs of Alaska Native communities, such as creating more preK-12 teachers, and more teachers of Alaska Native languages, and include Native Nation building as well.

Alaska Native people and communities are resilient. Through concerted efforts in the area of education we have collectively moved in a positive direction. We still have a ways to go and are confident that we will see more local and tribal control of education in the 2020s via the Tribal compacting initiative. There is optimism in the leadership of the younger Alaska Native community, who are agile thinkers and highly strategic. The cadre of Alaska Native university faculty, staff and students are a resource and voice that must be listened to. The Alaska Native Claims Settlement Act has had a positive impact on education and their continued involvement and financial support is necessary as we move forward.

References

- ANCSA Regional Association, (n.d.). Economic Impacts: Alaska Native Regional Corporations Investing in Alaska. Viewed November 29, 2021, <https://ancsaregional.com/economic-impacts/#investing>.
- ANCSA Regional Association, (n.d.). The Twelve Regional Corporations. <https://ancsaregional.com/the-twelve-regions/>.
- Arctic Education Foundation, (n.d.). About Us, Arctic Education Foundation. Viewed November 29, 2021, <https://arcticed.com/about-us/>.
- Arnold, R.D., 1976. Alaska Native Land Claims. Alaska Native Foundation, Anchorage.
- Arnold, R.D., Hickok, David M., Wunnicke, E.C., 1968. Alaska Natives and the Land. Federal Field Committee for Development Planning in Alaska. US Government Printing Office.
- Bavilla, R., 2021. Personal Communication. November 8, 2021.
- Bering Straits Native Corporation, 2021. Shareholder Benefits—Scholarship, Bering Straits Native Corporation. Viewed November 29, 2021. <https://beringstraits.com/shareholder-benefits/>.
- Calista Education and Culture, Inc., (n.d.). About Us, Calista Education and Culture, Inc., Viewed November 29, 2021, <http://www.calistaeducation.org/>.
- Cotton, S.E., 1984. Alaska's molly hootch case: high schools and the village voice. Educ. Res. Q. 8 (4).
- Doyon Foundation, (n.d.). About, Doyon Foundation. Viewed November 1, 2021, <https://doyonfoundation.com/about/>.
- Drabek, A., 2012. Liitukut Sugpiat'stun (we are learning how to be real people): exploring Kodiak Alutiiq literature through core values. A thesis in partial fulfillment of for the Degree Doctor of Philosophy. University of Alaska, Fairbanks.
- Kreger, J., 2021. Personal Communication. November 29, 2021.
- McDowell Group, August 2019. BBNC Education Foundation Scholarship Program: Database Analysis and Stakeholder Interview Findings. Bristol Bay Native Corporation Education Foundation, Anchorage.
- Morris, K., 2021. Personal communication. November 4, 2021.
- Napoleon, H., 1996. Yuuyaraq: The Way of the Human Being. Alaska Native Knowledge Network, Fairbanks.
- National Congress of American Indians Policy Research Center, August 13, 2021. Research Policy Update: A First Look at the 2020 Census American Indian/Alaska Native Redistricting Data.
- Propublica, November 15, 2021. Nonprofit Explorer: Research Tax-Exempt Organizations. Propublica viewed November 29, 2021. <https://projects.propublica.org/nonprofits/>.
- Sealaska Heritage Institute, (n.d.). About, Sealaska Heritage Institute. Viewed October 4, 202, <https://www.sealaskaheritage.org/>.
- State of Alaska Department of Education & Early Development Statistics & Reports. <https://education.alaska.gov/data-center>.
- The CIRI Foundation, (n.d.). About, the CIRI Foundation. Viewed October 16, 2021, <http://thecirifoundation.org/about/>.
- Williams, M., 2009. "The Comity Agreement: Missionization of Alaska Native People" the Alaska Native Reader. Duke University Press.

Relevant websites

- http://www.alaskool.org/resources/statistics/pop_isegr1971/aknative_population_trends.htm.
- <https://ancsaregional.com/the-twelve-regions/>.
- https://www.uaa.alaska.edu/academics/office-of-academic-affairs/institutional-effectiveness/departments/institutional-research/_documents/fact-book-2018-19.pdf.
- <https://www.ilisagvik.edu>.
- <https://uaf.edu/ankn/publications/guides/alaska-standards-for-cult/>.
- <https://education.alaska.gov/compacting>.

Weaving together mind, body and spirit: envisioning pathways to trauma-informed classrooms from Indigenous perspectives

Sandra Styres, Ryan Neeppin, and Julie Blair, Ontario Institute for Studies in Education, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Opening the circle with our “fallen feathers”	32
Trauma, development and classrooms	34
A journey of un-learning to re-learn	36
Closing the circle with feathers of hope	38
References	40

Weaving Together Mind, Body and Spirit: Envisioning Pathways to Trauma-informed Classrooms from Indigenous Perspectives.

Sandra Styres, Ryan Neeppin and Julie Blair.

Opening the circle with our “fallen feathers”¹

Who will sing for us? Who will dance for us?

Who will remind us of our relationship to the earth?

Who will Tell our Stories?²

The subject of trauma-informed classrooms from Indigenous³ perspectives in Canada, which is the focus of this chapter, began with community. Linda Tuhiwai Smith (2012) defines community as a multi-faceted construct that encompasses the “physical, political, social psychological, historical, linguistic, economic, cultural, and spiritual spaces” (p. 128) of community. An Elder from the urban Indigenous community in Tkaronto⁴ approached two of the authors with concerns about his children who are enrolled in classrooms within one of Tkaronto’s largest school boards. His concerns centered around several key issues: (a) literacy achievement rates for Indigenous learners in K-12⁵ classrooms; (b) affordable access to literacy tutoring having implications across diverse subject areas; and (c) low high school graduation rates for Indigenous learners. These conversations resulted in a literacy pilot study⁶ that collected baseline information at Kapapamahchakwew,⁷ an Indigenous-centric school in Tkaronto.

¹Tanya Talaga’s Seven Fallen Feathers.

²King, T., 2003, The Truth About Stories, p. 151.

³In the context of this chapter Indigenous refers to First Nations, Métis and Inuit peoples across Canada. According to Statistics Canada (2021) there are more than 630 First Nation communities in Canada, which represent more than 50 Nations and 50 Indigenous languages. More than half of all Indigenous peoples across Canada live in urban city centers with Ontario having the highest Indigenous population of any other province in Canada. Ontario is home to 23% of all Indigenous peoples in Canada. In Ontario there are 133 First Nations communities representing 7 overarching cultural and linguistic groups. Urban centers in Ontario with considerably high Indigenous populations are: Thunder Bay, Sudbury, Sault Ste. Marie, Timmins, Ottawa, and Toronto. <https://www.rcaanc-cimac.gc.ca/eng/1100100013791/1535470872302>

<https://www.sac-isc.gc.ca/eng/1603371542837/1603371807037>

⁴Tkaronto is the Mohawk word for Toronto. It means “where there are trees in the water.” (John Steckley, cited in Court, P. (2004). How Toronto Got Its Names. Sightlines XLIV(7). United States Institute for Theater Technology, Inc. © 2004, vol. XLIV, Number 7. Retrieved from <https://web.archive.org/web/20080720102358/http://www.usitt.org/sightlines/v44/n07/stories/Toronto.html>). Tkaronto is the capital city of Ontario, Canada. The extensive territories that encompass Tkaronto and the vast lands surrounding it have a long and tangled (hi)story that carries the storied foot prints of the Wendat, Haudenosaunee and Anishinaabek nations. Tkaronto has a treaty with the Mississaugas of the Credit River First Nation, and is part of the Dish With One Spoon Treaty which is a First Nation to First Nation agreement to share the hunting grounds of this territory in peaceful co-existence. The treaty was made between the Haudenosaunee League of the original Five Nations (Mohawk, Onondaga, Oneida, Cayuga, Seneca) their allies and a confederacy of Anishinaabek and allied nations after decades of inter-tribal strife.

⁵In Canada, elementary public schools have classrooms that begin in kindergarten (ages 4–5) and continue through to grade 12 (ages 17–18).

⁶The overarching aim of the pilot study was to provide baseline information in grades 1–3 at Kapapamahchakwew - Wandering Spirit school to support the development of a tutoring literacy program for Indigenous learners that would have implications across a wide range of subject areas. Utilizing OISE’s Master of Teaching student teachers as tutors and mentors, our aim was to increase Indigenous student success, achievement and well-being in K-12 classrooms.

⁷Kapapamahchakwew - Wandering Spirit School is an Indigenous-centric school in the Toronto District School Board. It was established in 1977 by Pauline Shirt and Vern Harper who were looking for a culturally safe and Indigenous identity focused school for their own child. In 1983, the school was recognized by the Toronto District Board as a Cultural Survival/Native Way program instead of an alternative school. In 1989, the program was renamed First Nations School of Toronto (FNST). Then in 2018/2019, the school returned to its original roots and after a renaming ceremony, is now called Kapapamahchakwew - Wandering Spirit School. The school is home to diverse urban Indigenous students from K-12 and is an ethically safe place where Indigenous children are taught in culturally relevant ways and Indigenous philosophies are centered. (<https://www.tdsb.on.ca/Community/Indigenous-Education/Schools/Kapapamahchakwew-Wandering-Spirit-School>).

One of the outcomes of this pilot study was that the community and researchers recognized that there was an undeniable need for training on Indigenous trauma-informed practices of care. This led us to develop a graduate course designed for both student teachers and educators⁸ entitled *Weaving Together Mind, Body and Spirit: Envisioning Pathways to Trauma-informed Classrooms from Indigenous Perspectives* (the course).

It is well documented that many Indigenous students remain underserved by mainstream models of education and continue to struggle in K-12 classrooms (Battiste, 2013; Cote-Meek, 2014; Cote-Meek and Moeke-Pickering, 2020; Moore, 2017; Tomlins-Jahnke et al., 2019; Styres, 2017, 2020; Styres and Neepin, forthcoming; Styres et al., 2020). The history of violence against Indigenous peoples perpetuated by the residential schools,⁹ the sixties scoop¹⁰, the millennial scoop,¹¹ missing and murdered Indigenous women and girls,¹² numerous and continuing social injustices and Indigenous rights violations, along with the ongoing discoveries of mass/unmarked graves in Canada at former residential schools¹³ have left a legacy of extensive inter-generational and compounded trauma with Indigenous peoples. Chief Willie Sellars said in an interview regarding the unmarked graves uncovered at the former St. Joseph's Mission Residential School located on T'exelc territory in British Columbia, that "the horrors that occurred inside walls of the [residential schools] are still very real for those who lived them. And the legacy of these atrocities is still readily apparent in the numerous ways that intergenerational trauma manifests in First Nation communities" (Fenton et al., 2022, n.p.).

Tanya Talaga (2017) recounts the story of seven Indigenous high school students from remote northern First Nations communities who died and/or went missing over a period of 11 years while at school in Thunder Bay. These children were forced to move hundreds of kilometers from their families and communities to attend high schools in Thunder Bay because there were no adequate high schools in their communities. Stan Beardy, former Grand Chief of the Nishnawbe Aski Nation, said that "going to high school is the right of every Canadian child. But these children have been treated differently, their needs forgotten in a country that prides itself on having one of the best education systems in the world" (cited in Talaga, pp. 20–21). One of those seven students who died while attending high school in Thunder Bay was Kyle Morriseau, grandson of the late world-renowned artist Norval Morriseau and son of artist Christian Morriseau who created the painting entitled *Seven Fallen Feathers*, which is also the title of Talaga's book. Christian, after the inquest into the seven deaths was complete, felt mired in his grief not only over his own son's death but the death of all seven youths. Out from that grief the image of the youths as seven individual but interconnected fallen feathers was gifted to him by the Great Spirit (pp. 300–301). The painting represents the ways the hearts and spirits of those fallen feathers are placed in the hands of their ancestors "who wait for them on the other side" (p. 301).

⁸In the context of this chapter, student teachers refers to students within the Ontario Institute for Studies in Education's (OISE), Master of Teaching program. The understanding of educator relates to students within the discipline and/or profession of education in its broadest context. Education occurs across a variety of informal and formal contexts. The course *Weaving Together Mind, Body and Spirit: Envisioning Pathways to Trauma-informed Classrooms from Indigenous Perspectives* includes students from a wide range of backgrounds, cultural traditions and spiritual beliefs. In addition, students in the course are positioned within different educational/employment backgrounds and trajectories as well as academic/professional programming and streams (i.e. MEd, MA, MT & PhD). As such, unless otherwise specified, when referring to all of the students within the course, we will be referring to them by the collective term "students".

⁹Over 150,000 children were forcefully and violently removed from their families by various policing agencies, clergy and Indian agents, and sent to residential schools. Many times these children would be sent to residential schools at great distances from their home communities in order to discourage run aways and familial visits. These children were isolated from the influence of their homes, families, traditions, language, and cultures. The 2015 Truth and Reconciliation Commission Report documents stories of neglect, disease, malnutrition, and many other horrendous abuses. Residential schools were a leg-islated and institutionalized agenda of mass assimilation and cultural genocide. In 1920, amendments to the 1876 Indian Act made it legally mandatory for every Indigenous child to attend residential school. Parents who refused faced heavy legal penalties and were denied legal representation. There are an estimated 80,000 former residential school survivors living today. The residential school era lasted for 160 years. The last residential school in Canada closed in 1996.

¹⁰Patrick Johnston (1981, 1983) coined the term "sixties scoop" in a 1983 report on Indigenous child welfare commissioned by the Canadian Council on Social Development. The sixties scoop, which overlapped with the residential school era, lasted from the mid-1950's into the 1980's. It was a period of time where policies and practices were enacted in Canada that gave Child Welfare Services the authority to remove children from their families in First Nations Communities and place them in non-Indigenous foster homes and adopt them out to white families. It is believed that approximately 20,000 children (some newly born) were forcibly and/or coercively taken during this period of time and adopted out in closed adoptions. Their Indigenous identity, history, and familial ties were purposefully kept from these children. The Canadian government's policies and practices purposefully ensured that there was no traceable connection between the children and their biological families. As well, their birth records were closed.

¹¹This "epilogue to the systematic removal of children in the sixties scoop" (Nelson, 2019, n.p.) has been called by critics of the system as the millennial scoop. Ridgen and Andrews (2019) reports that the forced removal of Indigenous children from their families continues to happen today in Canada, often with little to no evidence that they were at risk of abuse or neglect. Further, that the continued government-sanctioned apprehension of Indigenous children from their kinship ties is a human rights violation (n.p.). "Cindy Blackstock, Executive Director of the caring society, says she sees case after case of First Nations children in trouble being sent routinely into care because that's where the funding is instead of trying to help families deal with their problems" (The Canadian Press, 2011, n.p.).

¹²Colonial violence against Indigenous women and girls is not a new phenomenon but has its historical roots in colonization and the entrenchment of settler futurities; the dispossession of Indigenous peoples and lands; and the systemic cultural genocide of Indigenous peoples through the implementation and perpetuation of Indigenous residential schools and the forceful removal of Indigenous children from their families. After the release of the 2015 Truth and Reconciliation Report, then Justice Murray Sinclair, Chair for the Truth and Reconciliation Commission, stated at the release of the report that an inquiry "for murdered and missing aboriginal women is a must, and further, that the government has flat out rejected the idea for years" (Rabson, 2015, n.p.). For additional information please read the 2019 *Reclaiming Power and Place: The Final Report of the National Inquiry Into Missing and Murdered Indigenous Women and Girls* (2019) written by the National Inquiry into Missing and Murdered Indigenous Women and Girls (Canada).

¹³<https://www.npr.org/2021/07/01/1012100926/graves-found-at-new-site-canadian-indigenous-group-says>
<https://www.cnn.com/2021/07/13/americas/canada-unmarked-indigenous-graves/index.html>
<https://globalnews.ca/news/8074453/indigenous-residential-schools-canada-graves-map/>
<https://www.bbc.com/news/world-us-canada-57592243>
<https://www.ctvnews.ca/canada/b-c-first-nation-discovers-93-potential-graves-at-former-residential-school-1.5754168>
<https://www.wlfrb.com/news/live-wlfrb-reveal-st-josephs-mission-preliminary-findings/>

Indigenous peoples, their children, and communities who have been and continue to be impacted by historical and contemporary traumas and violences are the fallen feathers that situate our work. The Canadian [Truth and Reconciliation Commission's \(TRC\) 2015](#) report identified 94 calls to action with education as a critical site for reconciliation, healing and well-being. Reconciliation requires that educators identify the implications of settler colonialism on their educational and classroom practices, actively engage decolonial approaches to those practices ([Styres, 2020](#)), and appropriately integrate Indigenous perspectives and trauma-informed practices of care into classrooms.

Drawing from multiple conversations with community, students, and the experiences of the course instructors, this chapter reflects on some of the ways we engage with trauma-informed practices of care. We also examine themes of culturally and spiritually safe care in relation to the diversity of Indigenous peoples across Canada and the ways educators are implicated in colonial relations of power and privilege. We conclude with insights and lessons learned, and a consideration of the ways we can move forward to create more inclusive trauma-informed practices in professional and educational environments across diverse contexts.

Trauma, development and classrooms

We have changed our fundamental question from

"What's wrong with you?" to "What happened to you?"¹⁴

Trauma can be understood as experiencing/witnessing one, multiple and/or repeated threatening or harmful events resulting in long-lasting impacts on one's entire being ([Guarino and Chagnon, 2018](#)). Generally, traumatic experiences can include: abuse, violence, accidents, natural disasters, war, and ongoing systemic racism/oppression where one has "a close personal encounter with violence or death" ([Herman, 1997](#), p. 82). Although traumatic events may end at a point in time, witnessing/experiencing one or multiple traumatic events has the potential to change a person's physical and psychological responses to future real or perceived threats, can negatively impact the "psychological structures of the self", and "the systems of attachment and meaning that link individual and community" ([Herman, 1997](#), p. 121). Not everyone who has experienced a traumatic event is impacted by it in exactly the same way. The support one receives immediately following a traumatic incident, personal, socio-cultural and environmental factors all influence the degree to which trauma impacts an individual ([DuBois and Mistretta, 2019](#); [Guarino and Chagnon, 2018](#); [Linklater, 2014](#); [Siegel, 2010](#); [Van der Kolk, 2015](#)).

Because the course we are writing about in this chapter focuses on trauma-informed classrooms for children and youth specifically, it is important to understand the impacts of childhood trauma. Repeated childhood trauma such as sexual, physical and emotional abuse and neglect is defined as psychic abuse or complex trauma, which can result in long-lasting "damage on the psyche" ([APA, 2015](#), n.p.) This can include Complex Post-traumatic Stress Disorder (cPTSD)¹⁵ symptoms of emotional dysregulation, relationship difficulties, and a "negative self-concept" ([Hershler and Lee-Evov, 2021](#), p. 24), that includes feelings of worthlessness, guilt and shame. The diagnostic criteria for cPTSD also includes "re-experiencing" trauma through intrusive thoughts or mental images, "avoiding" reminders of trauma and feelings of being under threat in the present, where one may feel a heightened sense of being on alert for potential threats in one's environment (p. 24). To cope with and manage these symptoms, trauma survivors may use strategies such as substance misuse, self-harm, and risk-taking behaviors that may lead to additional trauma, guilt and shame ([Maté, 2008](#)).

Indigenous understandings of trauma are holistic in nature and recognize the spiritual impacts of trauma, where trauma and its resultant symptoms have resulted in an imbalance of one's sense of self; a "a soul wound" and "an injury where blood does not flow" ([Duran and Firehammer, 2017](#), p. 107; [Duran et al., 1998](#), p. 341). Generally, Western medical understandings of mental health focus on diagnosis and treatment of an illness, where "Indigenous healing begins with the Spirit" and connecting with self, others and the environment with an aim to restore balance and connection within these relationships ([Linklater, 2014](#), p. 125). In addition to individual experiences of trauma, Indigenous peoples' have experienced communal trauma in the form of multiple historical and ongoing traumas in present day, in the form of colonial violence, ongoing racism and oppression ([Burrage et al., 2021](#), p. 2; [Duran and Firehammer, 2017](#), p. 107; [Linklater, 2014](#), p. 14). Repeated instances of trauma, which has been experienced by multiple generations of Indigenous families and communities can result in "historical" "intergenerational" and "multi-generational" trauma where trauma symptoms and ways of coping are passed down through generations and "institutional patterning" via "blood memory" (Younging cited in [Linklater, 2014](#), p. 15). Colonial violence, racism, and oppression perpetrated against Indigenous peoples has been ongoing and continues to present day. [Linklater \(2014\)](#) asserts that "a vast majority of Indigenous people have lived, or are living in trauma; in most cases this trauma is multigenerational" (p. 20).

¹⁴Bloom, cited in [Perry and Winfrey \(2021\)](#), p. 19.

¹⁵The diagnosis of cPTSD has been included in the 11th version of the World Health Organization's (WHO) International Classification of Diseases (ICD-11) as of 2018.

Perhaps most relevant to learning about colonial and multigenerational violence perpetrated against Indigenous communities, families and children in relation to trauma-informed classrooms is the residential school system in Canada. Burrage et al. (2021) note that themes of loss - of identity and culture, of connection to self and others, of family violence, of "difficulties with parenting", and of destructive behaviors are all linked to the "intergenerational effects of residential schools" (p. 11). Given the likelihood of intergenerational transmission of trauma symptoms in Indigenous families and communities, it is important for educators to understand how Indigenous peoples may be impacted by attending educational institutions, when schools were places where generations of their families and relatives experienced violence, abuse and neglect.

Research shows that traumatic memories are stored in the brain and body. Further, that symptoms arising following traumatic experiences can have long lasting impacts on one's perception of the world and of themselves (Burrage et al., 2021; Hershler and Lee-Evoy, 2021; Siegel, 2010; Van der Kolk, 2015). Van der Kolk (2015) notes that "[a]fter trauma, the world is experienced with a different nervous system. The survivor's energy now becomes focused on suppressing inner chaos, at the expense of spontaneous involvement in their lives" (p. 110). In order to conceptualize how the brain is impacted by trauma, we can think of the brain having three main areas: brainstem, limbic region and cortex (Persaud, 2021; Siegel, 2010). The brainstem controls automatic body functions - breathing, heartbeat, digestion - and is the first part of the brain to develop. Our limbic region is involved in emotion creation, memory, and evaluative functions. The cortex is responsible for executive functioning and nervous system regulation is the last part of the brain to develop (DuBois and Mistretta, 2019; Siegel, 2010; Siegel and Payne Bryson, 2011; Van der Kolk, 2015).

When we are faced with a threat our natural and very necessary survival responses (fight, flight, freeze, collapse) kick in automatically. The amygdala, located in the limbic region, senses danger and quickly sends a signal to the brainstem to initiate a survival response, which "allows us to act before we think" (Siegel and Payne Bryson, 2011, p. 93). When these survival responses are activated, the prefrontal cortex goes offline¹⁶ - one is literally unable to think, reason, or learn. In trauma survivors, these responses can become primed over time to engage at even the hint of a real or perceived threat (Siegel, 2010). The overactivation of these survival responses puts enormous strain on one's system over time, and can have significant impact on the overall health of trauma survivors. Felitti's 1995 Adverse Childhood Experiences Study (ACE)¹⁷ found that there is a co-relation between experiences of trauma and neglect in childhood with increased depression, suicidal ideation, and addiction (Van der Kolk, 2015). The ACE study also found that people with 6+ incidents of adverse experiences in childhood had a 15% higher likelihood of developing major health problems such as heart disease and cancer (p. 289).

In addition to the findings from the ACE Study, experiences of trauma in childhood impact secure attachment, sense of safety, and brain function (DuBois and Mistretta, 2019; OFIFC, 2016), and children who experience neglect, abuse and poverty are "more at risk for delays in social, emotional and behavioral areas" (DuBois and Mistretta, 2019, p. 14). Some signs of trauma responses in children include: anxiety, difficulty concentrating, impulsivity, aggression, perfectionism, defiance and withdrawal (Cole et al., 2013; Guarino and Chagnon, 2018). Adolescents who experience trauma responses have lower academic achievement, social withdrawal, and engage in self-harming behaviors to cope with unwanted emotional and physical sensations associated with traumatic memories and frequent activation of survival responses (Guarino and Chagnon, 2018). Essentially, trauma experienced in childhood disrupts healthy development across a variety of markers. Strength-based approaches to understanding child development include Bruce Perry's Six Core Strengths for Healthy Child Development¹⁸ and traditional Indigenous Life Cycle Teachings.¹⁹ Children's brain and social development progresses in a scaffolded manner, where developmental milestones are dependent on mastery of the previously acquired skills. Perry describes the six core strengths of child development that follow the sequence of child brain growth as: attachment, self-regulation, affiliation, attunement, tolerance, and respect, with each core strength building on the previous one (Cole et al., 2013). Indigenous understandings of developmental stages can be illustrated using the Life Cycle Teachings model (OFIFC, 2016, p. 5) to understand the different roles and responsibilities we have at different stages of our lives, the strengths of being in each life stage, and the function of specific relationships between members of different life stages²⁰ (see also Wabano Health Center, 2017).

The ACE study also showed that childhood trauma is more prevalent in society than previously thought and, given the historical and intergenerational traumas Indigenous peoples' have experienced, it is likely that Indigenous students will experience trauma symptoms in the classroom (Cole et al., 2013; OFIFC, 2016). Teachers who have experienced or witnessed trauma (in or outside the classroom) also experience symptoms associated with burnout vicarious trauma and compassion fatigue²¹ (DuBois and Mistretta, 2019; Guarino and Chagnon, 2018). The good news is that human brains have the capacity to change especially during childhood - we have the capacity to form new neural pathways via environmental influences. Experiences of trauma can change our

¹⁶Siegel's hand model of the brain video link- <https://www.youtube.com/watch?v=f-m2YcdMdFw>.

¹⁷<https://developingchild.harvard.edu/resources/aces-and-toxic-stress-frequently-asked-questions/> (ACE STUDY).

¹⁸<https://youtu.be/skaYWK6iD4> (The Child Trauma Academy Channel, 2014).

¹⁹<https://youtu.be/ZotAL3BZqj4> (Wabano Health Center, 2017).

²⁰According to OFIFC's Trauma Informed Schools Report, "the lifecycle wheel teaches us that we come into this world with a number of gifts from creation including our clans, songs, colors, and our unique gifts. When we are born we have a responsibility at each of our life stages though out the lifecycle. With each of these stages of life it is the responsibility of the opposite stage directly across from them in the lifecycle to support them in fulfilling their responsibilities. We see that when one of the stages is not acknowledged in fulfilling their role that their growth may be stunted, such as if an infant is not able to provide joy to a family because the parents are impacted from trauma, they may not be able to enjoy their infant the infant may fail to thrive or return to the creator. We can notice this with each of the life stages. It is a natural process to move through each of these stages if they are supported and allowed to fulfill their responsibility. If there is no support, we return to these stages of the life cycle again and again to complete our responsibility/journey" (Ontario Federation of Indigenous Friendship Centers (OFIFC), 2016, p. 48).

²¹See DuBois and Mistretta (2019), Chapters 4 & 5, pages. 41-76 for additional information on burnout, vicarious trauma and compassion fatigue. Symptoms include but are not limited to anxiety, low energy, decreased ability to focus, emotional dysregulation, and absenteeism.

brains and so can “good experiences”, therefore supportive interventions in childhood can serve as protective factors that foster resiliency (Cole et al., 2013, p. 19). Since it is likely trauma will show up in the classroom, trauma-informed teachers need to be able to recognize what the flight, fight, freeze and collapse responses look like and to identify the coping strategies that children and youth engage in to cope with states of dysregulation associated with these survival responses (Cole et al., 2013). Educators need to be equipped with this knowledge to support students, create safer classroom spaces, and reduce the likelihood of responding in ways that cause additional harm (Guarino and Chagnon, 2018).

When trauma symptoms and survival responses arise in classrooms, educators may initially view these symptoms as disruptive/disengaged behaviors. When viewed through a trauma-informed lens we begin to understand that these may be survival responses - activated reactions to perceived or anticipated threats (Cole et al., 2013; Dubois and Mistretta, 2019). In trauma-informed classrooms we aim to shift perspectives, to cultivate curiosity and compassion where, instead of asking “What is wrong with you?”, we ask “What happened to you?”. This fosters an understanding that ultimately these trauma responses are highly attuned self-protective mechanisms (Bloom, cited in Perry and Winfrey, 2021, p. 19; OFIFC, 2016). This requires that educators embark on a journey of un-learning taken-for-granted biases and assumptions in order to re-learn trauma-informed care within their own classrooms. Using this trauma-informed lens offers an opportunity for educators to respond mindfully and to support students who are in the midst of a survival response or using unhealthy coping strategies to alleviate their emotional pain (Cole et al., 2013; Maté, 2008).

A journey of un-learning to re-learn²²

We make the road by walking²³

Without a doubt our communities are still trying to heal from having survived historical and ongoing colonial traumas and continued attacks on our lands, rights, nations, communities, families and kinships. Styres (2017) tells us that every journey “is an expression of movement from one place to another by making a conscious and deliberate decision to move into unfamiliar territory while maintaining an observing and reflective frame of mind ... and that ... every journey begins with a single movement forward” (pp. 8, 178). As already said, the course was born out of working with settler student teachers who, as part of a research project,²⁴ were completing their teaching practicum placements in primary classrooms in Kapapamahchakwew. Our children are still shouldering the burden of these various traumas which undoubtedly manifest in classrooms and schools and it became very clear to us that student teachers are ill prepared to complete their practicum placements at Kapapamahchakwew. The student teachers themselves stated they did not feel equipped to understand the historical and ongoing trauma or the lived realities of Indigenous peoples. They expressed the need for relevant training to work with learners who have and are currently experiencing trauma and for the need to create culturally safe places of learning that do not continue to perpetuate harm. Our course was developed in response to these conversations.

The importance of Indigenous students seeing themselves and their ways of knowing and being reflected in educational spaces (schools, curriculum, classroom practices) is well-documented in Indigenous scholarship. Often, educators believe that being trauma-informed means avoiding difficult topics and conversations. This is a position that serves educators to the detriment of Indigenous students. Indigenous students need and want to learn about their histories and to be able to safely discuss their current lived realities. Instead of avoidance, educators should focus on how to teach difficult topics in culturally safe trauma-informed ways. It is far more harmful for Indigenous students to have their histories, current lived realities, and ways of knowing and being erased in schools and the curriculum – these erasures, regardless of intentions, are acts of violence (Styres, 2020). Educators often find teaching and learning about Canada’s colonial relationship with Indigenous nations disruptive and uncomfortable because they become enmeshed in relations of power, privilege, and complicity. This leads to entangled cycles of resistances – both their own and their students. One way that educators can move past these resistances and begin teaching Indigenous content is through a journey of un-learning to re-learn. This un-learning to re-learn is often messy and uncomfortable involving a critical examination of Canada’s colonial history with Indigenous peoples and how they, as educators and responsible citizens, are implicated in the ongoing oppression and dispossession of Indigenous peoples (Styres and Neepin, forthcoming).

Students in our course are often surprised to find that it is only after learning about trauma, its impacts on brain development, and how it is embodied, that they begin learning about Canada’s colonial history. Students tell us that much of what they learn in

²²You will note the use of hyphenation which is our stylistic way for emphasizing the prefix un- and re-. Un-as a verb refers to a shift, a turning-around, a change in the direction of an action, positionality or state of being. Re- as a verb means again, afresh, and anew – the repeating or redoing of an action or process.

²³Horton and Freire (1990), pp. 6, 7.

²⁴The research was a pilot study entitled Tutoring for Indigenous Children and Youth in Toronto. The pilot research sought to gain base-line information that would determine where primary students in their literacy achievement rates, what interventions were currently in place, how and in what ways are they/are they not working to increase literacies, what does the community see as key priorities in increasing literacies. There were four very clear outcomes from this study: (1) the need for an Indigenous focused trauma-informed graduate education course; (2) ethics of relationality are key in any literacy project with the schools/children/families; (3) the need for a literacy program that encompasses the unique needs of the diverse urban Indigenous learners in Tkaronto; and (4) the need for the development and implement of Indigenous focused readers and diagnostic tools. The data from the pilot study as resulted in the development of a larger research project focused on literacy development.

this second part of the course is new to them - it is not a history they learned in their own educational journeys. Students in the course report undergoing the uncomfortable journey of un-learning to re-learn in class discussions/assignments. With a significant portion of the course focused on colonial history and its traumatic effects on Indigenous peoples, midway through the course it becomes clear that many students begin struggling with the course content and become entangled in and overwhelmed by their discomfort. Due to their privileged positionality as settler educators,²⁵ this discomfort alarms them and they begin forming counter-resistances against their un-learning to re-learn journeys. Anecdotally, many students read the course title and believe they are going to take a course where they can process their own experiences with trauma. Students have stated expectations about learning “Indigenous healing methods” or “ceremonies”. When they realize the course is not group therapy, a personal wellness journey, or participation in “healing” ceremonies, there is a shift in attitudes and expectations that lead to additional forms of resistance stemming from predetermined unmet expectations. Many students bring into the learning environment their preconceived performative notions about the ways Indigenous instructors should act and how they should look. Students are often surprised and disappointed that the course does not include beads, feathers, medicines and ceremonies (King, 2003, pp. 61–68). Educators in general, have very specific ideas of how Indigenous peoples are supposed to portray themselves and be characterized by others. This is evidenced by a recent case of a teacher in California who dressed up in a home-made paper headdress as an “Indian” and performed “ancient legends,” dances and chants in her classroom (Salcedo, 2021). Similarly, educators have very specific expectations of how Elders are supposed to “perform” for them when invited into their classrooms. Students in our class also hold very specific expectations of how we were supposed to “perform” as Indigenous instructors – essentially, we were not the “Indian [they] had in mind” (King, 2003, p. 31).

Our course is a tightly woven tapestry of resistances and counter-resistances. The introduction of this course into a colonial institution along with our mindful pedagogical approaches to the provocative course content are the first resistances. Zinga and Styres (2018) tell us that:

Decolonizing and anti-oppressive pedagogies are, by their very nature, resisting mainstream approaches to teaching and learning, as well as challenging taken-for-granted assumptions embedded in the status quo. Resistance implies a very particular resistance to colonial and oppressive structures, whereas counter-resistance implies resistance to learning about or being taught decolonizing and anti-oppressive pedagogy. In the classroom, these pedagogies challenge the networks and relations of power and privilege that are systemically embedded in academia (p. 3).

Teaching a course that flips the narrative concerning Canada’s historical and contemporary relationships with Indigenous people is resistance that pushes back against the dominant colonial institutional, political and social (hi)story. Our instructor team and pedagogical choice of privileging Indigenous scholarship strengthens that resistance. The course is co-taught by two Indigenous instructors that continuously engage with and resist the imposed hierarchies of the academy (and by extension the students) and make the deliberate choice to team-teach collaboratively and model the “ethics of relationality” (Styres and Kempf, 2022) which initiates other resistances in the course. Immediately students begin to counter-resist our pedagogies and course content. They become increasingly uncomfortable with the un-learning/re-learning work that is required to move forward. As instructors we respond by deliberately creating safe and supportive spaces for students to work through their counter-resistances. For example, as students begin learning about Canada’s colonial history and ongoing relationships with Indigenous peoples, they write critical reflection papers based on course readings. We prompt students to critically examine their “mainstream taken-for-granted assumptions” (Zinga and Styres, 2018, p. 4), to sit with their discomfort, and to engage in deeper and more meaningful ways. The aim is to get students to begin thinking critically about what they are learning, to describe what they are feeling, to name those emotions and what they are attached to (i.e. guilt, shame, defensive) which, if unexamined, can lead to “settler moves to innocence” (Tuck and Yang, 2012, p. 1). As instructors, we too, much like Zinga and Styres (2018) experience an unwillingness from students to critically engage as they are “frantically trying to reinstate the illusion of the meritocracy, which provides them with a level of comfort not found in critical reflection” (p. 7). Students are required to participate in small weekly reading groups where each group member presents a deeply thoughtful reflection on their assigned readings. The students are also expected to keep a journal so they can reflect on the tensions, challenges, and pivotal learning moments each week. While the weekly reflections are focused on analyzing their experiences with the readings, the journal serves to document the students’ learning journey throughout the course. Together these weekly exercises scaffold students’ experiences in preparation for their final paper. Students are immediately resistant to and actively complain about the amount work there is in the course, while simultaneously saying that what they are learning is new and shocking information. In order to do the work of un-learning to re-learn, critical reflection is crucial and documenting this process is key because it demonstrates those pivotal moments where the students’ thinking was challenged, critically engaged, emotively processed, and eventually changed. This type of reflective thinking and writing is intense because it is not merely surface-level regurgitation of what was read but rather critically engaging with what is being read and learned.

Educators have told us they are fearful of making mistakes when it comes to teaching Indigenous content. They are afraid of saying or doing the wrong thing and they are afraid of being accused of cultural appropriation. Quite often educators look for ways to “tick the check-box” of teaching Indigenous content by inviting in Indigenous Elders into their classrooms for a one-off

²⁵To date, a large majority of the students registered in our course identify as white-settler: student teachers, teachers in K-12 classrooms, and educational practitioners.

visit to teach their class. The problem with this is the expectation placed upon the Elder to do the heavy lifting of teaching their learners – essentially doing their job for them. Asking Elders to come into the classroom/school to “perform” ceremonies, to drum and sing, or worse, to perform a rain dance is inappropriate and disrespectful. It is an extractive, exploitive and voyeuristic practice because Elders are literally being asked to perform “being Indian” for the class. Our Elders are very precious to our communities and are exceedingly busy doing the important work of healing and language retention/revitalization efforts. We do, however, encourage educators to cultivate meaningful relationships with local Indigenous community members/organizations. Instead of one-time invitations, this relationship must be ongoing and reciprocal. This could include inviting parents, grandparents, aunts and uncles, as well as local artists, musicians and community leaders into schools (Styres and Neepin, forthcoming) – it is also critical that schools obtain funding to pay honouraria to the community members for their valuable time. Educators must actively participate and co-learn along with their learners from the community member. Another aspect of the co-learning method of teaching and learning is for educators to bring Indigenous-authored teaching resources into their classrooms (Styres and Neepin, forthcoming). This method is ongoing and emergent promoting a continuing journey of un-learning to re-learn.

Due to the students’ own long-held and reinforced hierarchical assumptions, expressions of resistances were also found in the ways expected of us to “perform” Indigeneity in our roles as instructors. There were three simultaneous performance roles being unconsciously enacted by students in our classes (Karpman, 2016).²⁶ One instructor was often seen by students as the “rescuer” and represented the safe place for students to land. The other instructor was cast in the role of “perpetrator” and represented the difficult content and challenging work. Students would often shift in and out of the “victim” role whenever they perceived they were being victimized by learning about the traumas Indigenous peoples have and continue to endure. What the students were counter-resisting continuously throughout their performance as “victim” was that they (as a collective) could not settle their expectations - the feel-good part they craved without the painful stories - with the difficult, unsettling, and often uncomfortable journey of un-learning to re-learn.

Students clearly went through various cycles of resistances and it was somewhere within these cycles where the authentic learning occurred. It was through iterative circularity (Styres, 2017) that began with the introduction of the course into an educational institute and our pedagogical approach to the course content as the initial resistance to colonialism. The students’ engagement with the content triggered a counter-resistance which, in turn, we as instructors also continuously mindfully and purposely resisted through engagement with our pedagogical approaches to the course content. To help guide students through this, we held space and offered support as they strived to work through these difficult challenges. It is at this point, when the students look back at their learning journeys and are continuing to do the difficult work of engaging with course content, that re-learning through critical self-reflexivity becomes a reality. The (new) tapestry that we have been weaving in this course created opportunities for us to bring light into the shadow places. The very things that created resistances and counter-resistances are the same things that support interventions, mindfulness, purposefulness, and being present while also speaking back to institutional violence, oppression, and colonialism through education.

Closing the circle with feathers of hope

Of course, you don’t have to pay attention to any of these stories ...

Do with it what you will. Cry over it, get angry.

Make it a topic of discussion group. Forget it.

But don’t say in the years to come that you would have lived your life differently if only you had heard this story. You’ve heard it now.²⁷

Our story returns full circle to community by drawing upon some of the insights gained concerning the critical ways a graduate course could have implications for the overall health and well-being of Indigenous learners within K-12 classrooms. Instructors aimed to both model a trauma-informed classroom while also teaching students ways to incorporate trauma-informed interventions into their future practices. This approach facilitated both experiential and theoretical knowledge transmission.

²⁶Karpman’s Drama Triangle (KDT), is based on theatrical roles that contribute to a good drama and in the context of this work, provides a framework for understanding ways resistances showed up in our classroom. KDT was developed by Stephan Karpman in order to be able to understand different roles that people enact, and patterns we can find ourselves in everyday interactions with others. The main roles in Karpman’s Triangle are the perpetrator, the victim, the rescuer and the neglectful or innocent bystander. (Hughes, 2021). Each actor in the triangle may be in a re-enactment of a situation they’ve experienced in the past and each of these roles serves a conscious or unconscious purpose for the people occupying them (Hughes, 2021). People can shift between roles, or be pushed into different roles by each actor in a scenario.

²⁷King, T., 2003, *The Truth About Stories*, pp. 60, 119.

The trauma-informed interventions taught in this course included gaining an understanding of what trauma and survival responses are generally, how trauma impacts Indigenous people specifically, and culturally appropriate ways to teach Indigenous content using a co-learning/-facilitation model. Cole et al. (2013), tell us that trauma-informed teaching practices include: assisting students with emotional regulation skills-building, setting high academic standards for all students, creating safety (including recognizing and responding to bullying and harassment), behavior management and limit setting, using strengths-based approaches while cultivating a sense of agency by creating opportunities for students to have choices.

As we know that survival responses are activated when one's sense of safety feels threatened, maintaining consistency in class structure and building a connected classroom community are ways of creating safety in classrooms to support students' nervous system and emotional regulation (Cole et al., 2013). Course instructors sought to teach about and demonstrate consistency, emotional regulation, open communication and accessibility in supporting learners, especially learners who have survived traumatic events. As a means to establish consistency and safety amongst our community of learners each class began with a sharing circle and agenda review, followed by small group reading conferences, an emotional regulation exercise, course content, experiential exercises, and ended with a sharing circle. Sharing circles were aimed at creating community, highlighting out common humanity, and were guided by prompts to promote emotional awareness and to name key learning themes at the end of class. To promote wellness and create connections students engaged in experiential exercises that included: creating personal wellness plans; arts-based strengths exploration; Indigenous education and wellness tools; online sharing forums and reflections on healthy endings; and ceremonies that mark transitions in student's own cultural backgrounds, families, or personal practices.

As Indigenous instructors teaching about Indigenous trauma to a class of mostly non-Indigenous students, we were aware that we ourselves were susceptible to experiencing vicarious trauma and compassion fatigue in the classroom. Due to the interconnected nature of Indigenous community relationships, the topics we taught also impacted our lives personally. To support one another and our own emotional well-being, we held weekly pre-brief and debrief sessions. During these sessions we also discussed how best to support students who appeared to be distressed by course content or who were dealing with challenges in their personal lives. These sessions supported our capacity to be intentional about providing tools to support students' wellness as they learned about multiple traumas Indigenous peoples' have been subjected to. Students were invited to participate in grounding exercises prior to each class as a way to ground themselves during lectures and discussions. Remaining attuned to the needs of others while remaining emotionally regulated is a skill that educators can cultivate through grounding²⁸ and mindfulness techniques.²⁹ Learning ways to stay grounded is an essential skill for educators in trauma-informed classrooms and professional practices because educators need to be emotionally regulated in order to co-regulate with students in dysregulated states of hyper- or hypo-arousal³⁰ (Hershler, 2021; Siegel and Payne Bryson, 2011).³¹

We want to conclude with the following insights and stories that we have drawn from weaving together a tapestry of mind, body and spirit which is trimmed in feathers of hope. These insights and learnings from the course go beyond the "walls" of our classroom spaces and reflection papers into the practices of educators and other helping professionals. Many students reported at the end of the course that they underwent a significant un-learning and re-learning journey. Further, that they had gathered many knowledges and strategies to take with them and integrate into their own professional practices and lives more widely. Some students named their past resistances to acknowledging their own complicity in Canada's colonial history and the continued and ongoing oppression of Indigenous peoples and communities. These same students engaged in the deeply emotional work of critical self-reflection and shared their ongoing learning with their own students and colleagues. Practicing educators and student teachers shared that they took what they learned and created lesson plans and professional development sessions based on course readings/content and taught them to their own students or shared them with colleagues.

One story of hope and healing is from a student in our course who self-identified as a white-settler educator and who had worked in Indigenous community schools for a majority of their career. In addition to weekly critical reading reflections and journaling, they expressed their thoughts and feelings about what they were un-learning to re-learn through art. The student worked on a "living" painting that continuously emerged and evolved as they added to it. In the end, the painting showcased the student's own healing and knowledge journey. The student gifted us as the instructors with the painting before moving away to resume their work as secondary school teacher in an Indigenous community. This story is similar to another settler-educator who also dedicated their decades-long career to working within northern and remote Indigenous communities. This educator also wanted to share what they were learning in the course with her colleagues and was working on their graduate degree to become principal. This student started a book club utilizing core course texts and materials as discussion points. Another settler-educator who told us about their transformative experience with the course worked at an east Tkaronto middle school with a large population of urban Indigenous students. They said that much of the course content was new and very shocking and they had to work through un-learning what they

²⁸Grounding strategies are outward focusing practices that involve a variety of physical, mental or soothing techniques that can be practiced at any time during triggered states when the thinking part of the brain is offline (Peckham, 2021).

²⁹See Nicholson et al. (2020). Chapter 3: Building a Culturally Responsive Self-Care Toolbox. (pp. 99–138).

³⁰Hershler (2021), p. 26, notes that fight, flight and freeze are the survival responses often seen in states of hyperarousal, and examples of hyperarousal include: anxiety, fear, anger, physical tension and urges to flee, yell or fight. Hershler also notes that the collapse survival response is often seen in states of hypoarousal and signs of hypoarousal include: feeling sad, lonely, shame, numb, disconnected and tired, and an urge to "sleep or hide".

³¹Human beings are wired for connection through mirror neurons in the brain and, when we engage with someone in survival response mode, we may feel a similar response rise within us. This means that our own ability to think clearly may become diminished in the moment (Katz et al., 2018; Siegel, 2010; Siegel and Payne Bryson, 2011). This also means that, using the same process of relating through mirror neurons, when we are within our "window of tolerance" (Siegel, 2010, p. 138) in an emotionally regulated state, and we connect with a student who is in a triggered state, we create the conditions for the student's nervous systems to re-regulate in response to ours (Siegel and Payne Bryson, 2011).

thought they knew, re-learning history, while untangling their own complicity as both a settler and teacher. They developed and implemented a series of professional development sessions based on their learning which featured readings and materials from the course. They felt very strongly that other teachers at her school, especially those with Indigenous students in their classes, ought to learn what they had in the course. This kind of knowledge transfer and sharing is another form of hope and healing. Finally, some students were so moved that they sought to move into “action” based on what they learned in the class and reported “wanting to do something tangible to make positive changes”. One student developed a relationship with an urban Indigenous agency in the city that supports under-housed and homeless Indigenous people in Tkaronto. The student learned that the agency required personal supplies so they coordinated a class donation drive to give to the agency.

These are a few examples of the feathers of hope and storied conversations that were woven into our tapestry informing the journey of un-learning to re-learn in our course. We see that Indigenous practices of trauma-informed care can disrupt and challenge the ways educational practitioners remain entrenched in colonial relations of power privilege that continue to perpetuate harm. A critical insight gained is that, Indigenous understandings of trauma-informed practices of care are an important step in addressing historical and contemporary socio-political issues for Indigenous learners and what it means to be Indigenous in K-12 classrooms across a wide-range of contexts.

References

- American Psychological Association (APA), April 2015. APA Dictionary of Psychology. American Psychological Association. Retrieved March 6, 2022, from: <https://dictionary.apa.org/psychic-trauma>.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. Purich Publishing Ltd, Saskatoon, SK.
- Burrage, R.L., Momper, S.L., Gone, J.P., 2021. Beyond trauma: decolonizing understandings of loss and healing in the Indian residential school system of Canada. *J. Soc. Issues*. <https://doi.org/10.1111/josi.12455>.
- Cole, S.F., et al., 2013. Helping traumatized children learn volume 1. Chapter 1: the impact of trauma on learning. In: Trauma Learning Policy Initiative, Massachusetts Advocates for Children and Harvard Law, pp. 14–41. Retrieved from: <https://traumasensitiveschools.org/tlpi-publications/download-a-free-copy-of-helping-traumatized-children-learn/>.
- Cote-Meek, S., 2014. *Colonized Classrooms: Racism, Trauma and Resistance in Post-Secondary Education*. Fernwood Publishing, Winnipeg, MA.
- Cote-Meek, S., Moeke-Pickering, T. (Eds.), 2020. *Decolonizing and Indigenizing Education in Canada*. Canadian Scholars Press, Toronto, ON.
- DuBois, A.L., Mistretta, M.A., 2019. *Overcoming Burnout and Compassion Fatigue in Schools: A Guide for Counselors, Administrators, and Educators*, first ed. Routledge. Retrieved from: <https://doi-org.myaccess.library.utoronto.ca/10.4324/9781351030021>.
- Duran, E., Firehammer, J., 2017. Injury where blood does not flow. In: Stewart, S., Moodley, R., Hyatt, A. (Eds.), *Indigenous Cultures and Mental Health Counselling: Four Directions for Integration with Counselling Psychology*, first ed. Routledge, New York, NY, pp. 107–124. <https://doi.org/10.4324/9781315681467>.
- Duran, E., Duran, B., Heart, Brave Heart, M.Y., Horse-Davis, S.Y., 1998. Healing the American Indian soul wound. In: Danieli (Ed.), *International Handbook of Multigenerational Legacies of Trauma*. Springer US, New York, NY, pp. 341–354.
- Fenton, C., Steacy, L., Wong, D., Martins, N., 2022, January 25. Williams Lake First Nation finds 93 possible burial sites around former residential school. Next City. Retrieved from: <https://vancouver.citynews.ca/2022/01/25/bc-residential-school-williams-lake/>.
- Guarino, K., Chagnon, E., 2018. *Trauma-Sensitive Schools Training Package*. National Center on Safe Supportive Learning Environments, Washington, DC. Retrieved from: <https://trauma.airprojects.org>.
- Herman, J.L., 1997. *Trauma and Recovery*. BasicBooks, New York, NY.
- Hershler, A., 2021. Window of tolerance. In: Hershler, A., Hughes, L., Nguyen, P., Wall, S. (Eds.), *Looking at Trauma: A Tool Kit for Clinicians*. The Pennsylvania State University Press, Essay, pp. 25–28.
- Hershler, A., Lee-Evov, J., 2021. Complex Posttraumatic Stress Disorder (cPTSD). In: Hershler, A., Hughes, L., Nguyen, P., Wall, S. (Eds.), *Looking at Trauma: A Tool Kit for Clinicians*. The Pennsylvania State University Press, Essay, pp. 21–24.
- Horton, M., Freire, P., 1990. *We Make the Road by Walking: Conversations on Education and Social Change*. Temple University Press, Philadelphia, PA.
- Hughes, L., 2021. Karpman's triangle. In: Hershler, A., Hughes, L., Nguyen, P., Wall, S. (Eds.), *Looking at Trauma: A Tool Kit for Clinicians*. The Pennsylvania State University Press, Essay, pp. 55–59.
- Johnston, P., 1981. Indigenous Children at Risk: Child Welfare Services Need Radical Changes. *Policy Options* (November–December 1981).
- Johnston, P., 1983. *Native Children and the Child Welfare System*. James Lorimer Limited, Toronto, ON.
- Karpman, S., 2016. A game free life. The definitive book on the drama triangle and compassion triangle by the originator and author. In: *The New Transactional Analysis of Intimacy, Openness, and Happiness*. Drama Triangle Publications, San Francisco, CA.
- Katz, J., Lamoureux, K., Moran, R., 2018. *Ensouling Our Schools: A Universally Designed Framework for Mental Health, Well-Being, and Reconciliation*. Portage & Main Press, Winnipeg, MB.
- King, T., 2003. *The Truth About Stories: A Native*. House of Anansi Press Inc, Toronto, ON.
- Linklater, R., 2014. *Decolonizing Trauma Work: Indigenous Stories and Strategies*. Fernwood, Winnipeg, pp. 11–50.
- Maté, D.G., 2008. *In the Realm of Hungry Ghosts*. Vermilion.
- Moore, S., 2017. *Trickster Chases the Tale of Education*. McGill-Queen's Press, Montreal, QU.
- National Inquiry into Missing and Murdered Indigenous Women and Girls Canada, 2019. *Reclaiming Power and Place: The Final Report of the National Inquiry into Missing and Murdered Indigenous Women and Girls*. Retrieved from: <https://www.mmiwg-ffada.ca/final-report/>.
- Nelson, K., 2019. The Millennium Scoop: Indigenous Youth Say Care System Repeats Horrors of the Past. CBC Radio. Retrieved from: <https://www.cbc.ca/radio/thecurrent/a-special-edition-of-the-current-for-january-25-2018-1.4503172/the-millennium-scoop-indigenous-youth-say-care-system-repeats-horrors-of-the-past-1.4503179>.
- Nicholson, J., Driscoll, P.S., Kurtz, J., Márquez, D., Wesley, L., 2020. Chapter 3: building a culturally responsive self-care toolbox. In: *Culturally Responsive Self-Care Practices for Early Childhood Educators*. Routledge, New York, NY, pp. 99–138.
- Ontario Federation of Indigenous Friendship Centres (OFIFC), 2016. “Ask Me About Trauma and I Will Show You How We are Trauma-Informed”: A Study on the Shift Toward Trauma-Informed Practices in Schools. In: OFIFC Research Series, vol. 4. Ontario Federation of Indigenous Friendship Centres. Retrieved from: <https://ofifc.org/wp-content/uploads/2020/03/Trauma-Informed-Schools-Report-2016.pdf>.
- Peckham, M., 2021. Self-care and grounding. In: Hershler, A., Hughes, L., Nguyen, P., Wall, S. (Eds.), *Looking at Trauma: A Tool Kit for Clinicians*. The Pennsylvania State University Press, Essay, pp. 13–20.
- Perry, B.D., Winfrey, O., 2021. *What Happened to You? Conversations on Trauma, Resilience, and Healing* [E-Book] Flatiron Books, Vancouver, BC.
- Persaud, S., 2021. The brain-three parts. In: Hershler, A., Hughes, L., Nguyen, P., Wall, S. (Eds.), *Looking at Trauma: A Tool Kit for Clinicians*. The Pennsylvania State University Press, Essay, pp. 39–43.

- Rabson, M., 2015. TRC Report Calls for Better Education, Health; Inquiry into Missing, Murdered Women. Winnipeg Free Press. Retrieved from. <https://www.winnipegfreepress.com/special/trc/TRC-report-calls-for-better-education-and-health-inquiry-into-missing-murdered-indigenous-women-305822541.html>.
- Ridgen, M., Andrews, B., 2019. Series: the child welfare industry. APTN National News. Retrieved from. <https://www.aptnnews.ca/child-welfare-industry/>.
- Salcedo, A., 2021. California teacher placed on leave after video shows her mocking Native American dance in headdress. Wash. Post. Retrieved from. <https://www.washingtonpost.com/nation/2021/10/22/california-teacher-suspended-video-native-american-headaddress/>.
- Siegel, D.J., 2010. *Mindsight: The New Science of Personal Transformation*. Bantam Books, New York, NY.
- Siegel, D.J., Payne Bryson, T., 2011. *The Whole-Brain Child: 12 Revolutionary Strategies to Nurture Your Child's Developing Mind*. Bantam Books, New York, NY.
- Smith, L.T., 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*, second ed. Zed Books, New York, NY.
- Statistics Canada, 2021. Government of Canada: Indigenous Communities in Ontario. Retrieved from. <https://www.sac-isc.gc.ca/eng/1603371542837/1603371807037>.
- Styres, S., 2017. *Pathways for Remembering and Recognizing Indigenous Thought in Education: Philosophies of lethi'nihténha Ohwentsia'kékha (Land)*. University of Toronto Press, Toronto, ON.
- Styres, S., 2020. Reconciliation: reconciling contestation in the academy. *Power and Educ.* 12 (2), 157–172.
- Styres, S., Kempf, A., 2022. Indigenous settler reconciliation and relational ethics. In: Styres, S., Kempf, A. (Eds.), *Troubling Truth & Reconciliation in Canadian Education: Critical Perspectives*. University of Alberta Press, Edmonton, AB.
- Styres, S., Neepin, R. (Forthcoming). Pedagogies of reconciliation. In: Berryman, M. (Ed.), *Relational and Responsive Pedagogy*. Encyclopedia of Teacher Education. Beijing, CN: Springer Singapore.
- Styres, S., Zinga, D., Douglas, V., Purton, P., 2020. Contested places: addressing the TRC in higher education. *Res. Educ.* 003452372093732.
- Talaga, T., 2017. *Seven Fallen Feathers: Racism, Death, and Hard Truths in a Northern City*. House of Anansi Press Inc, Toronto, ON.
- The Canadian Press, 2011. First Nations Children Still Taken From Parents. CBC News. Retrieved from. <https://www.cbc.ca/news/politics/first-nations-children-still-taken-from-parents-1.1065255>.
- The Child Trauma [sic] Academy Channel, January 14, 2014. Six core strengths for healthy child development: an overview [Video] YouTube. Retrieved from. <https://youtu.be/skaYWK6iD4>.
- Tomlin-Jahnke, H., Styres, S., Lilley, S., Zinga, D. (Eds.), 2019. *Indigenous Education: New Directions in Theory and Practice*. University of Alberta Press, Edmonton, AB.
- Truth and Reconciliation Commission of Canada, 2015. *Canada's Residential Schools: The Final Report of the Truth and Reconciliation Commission of Canada*. McGill-Queen's University Press, Montreal, QC.
- Tuck, E., Yang, W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Van der Kolk, B.A., 2015. *The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma*. Viking E-books, New York, NY.
- Wabano Health Centre, June 23, 2017. Life cycle teachings [Video] YouTube. Retrieved from. <https://youtu.be/ZotAL3BZqj4>.
- Zinga, D., Styres, S., 2018. Decolonizing curriculum: student resistances to anti-oppressive pedagogy. *Power and Educ.* 11 (1), 30–50. <https://doi.org/10.1177/1757743818810565>. Retrieved from.

Ke Ala o ka No'eau: an infusion of Hawaiian into the Hawaiian language

K. Laiana Wong, Katrina-Ann R. Kapā'anaokalaōkeola Nākoa Oliveira, and Sam L. No'eau Warner, University of Hawai'i at Mānoa, Honolulu, Hawai'i

© 2023 Elsevier Ltd. All rights reserved.

Introduction	42
An infusion of Hawaiian into Hawaiian language	43
The linguistic calque: a sticking point for Ka Lau 'Ulu	44
'Oko'a ka palaoa, 'oko'a ke koholā	44
Pupuka 'o hope o ka pīkake!	45
Pakalaki	45
The continuation of Ka Lau 'Ulu	47
Aia 'o Māmā i ke kaula'i lōle	47
He aha ka inoa o ka holoholona e umō mai nei?	47
Next steps	48
References	49

Introduction

The theme of the 2022 American Educational Research Association's annual meeting "Cultivating Equitable Education Systems for the 21st Century" obviously suggests that current systems are inadequate and that new, more equitable, systems must be established in order to support the advancement of all students. As we endeavor to improve on last year's direction recognizing that "Education researchers are not merely scholars; we are also citizens of the places in which our scholarship is produced, disseminated, and implemented. Equity and justice in these places depend as much on our deep thinking as they do on what we do with what we know. Racism, xenophobia, transphobia, Islamophobia, anti-Semitism, homophobia, and other manifestations of hate continually poison these places,"¹ we must recognize that these issues, while easily recognizable in others, might be operating surreptitiously within our own efforts to effect educational change. In an effort to create materials that promote "new" Hawaiian language literacies that reflect Hawaiian ways of speaking and knowing, we are unaware of the degree to which those efforts are guided by English worldviews, thus obscuring Hawaiian worldviews and Hawaiian ways of speaking that would distinguish it from English, and shield it from linguistic colonization, a patently inequitable process.

Obviously, in order to survive, Hawaiian, along with numerous other endangered indigenous languages, must be able to accommodate a 21st century world that has been inundated with English. As such, it is not feasible (let alone possible) to utilize some idealized version of the language that is completely pure and unadulterated by foreign influence, to include that imposed upon Hawaiians by English over the past two and a half centuries. Even if such a version had been hermetically sealed over that span of time and then suddenly made available to a community shielded from the effects of colonization, it would not accommodate our current linguistic needs. Moreover, having no actual shared experiential contexts with our forbearers, it would be quite challenging for us to understand their version of the language as they interacted with each other in communicative contexts. Given the heavy influence from English, and the number of generations that have acquired it, either along with Hawaiian or monolingually, its complete erasure from an emergent Hawaiian would be virtually impossible. Vigorous "updating" would be unavoidable. These efforts have been successful insofar as the overall number of new speakers is concerned, i.e., the number seems to have grown slowly along with levels of fluency, however, the range of topics that can be verbalized is not equally distributed across the speech community. Moreover, the "ways of speaking" that are reflective of a separate worldview, and which are not affected by advancements in technology, are fading into oblivion. We are so far removed that we don't even know what we don't know.

It is not our purpose here to avoid any particular ways of thinking or knowing merely because they derive from an English worldview, and neither did our ancestors. We recognize the value of foreign technologies and aspire to utilize them to our advantage. In fact, we see much of Western education as the vehicle for reshaping the English worldview that has possessed the community, and we have decided to confront these issues through the creation of materials that support a Hawaiian worldview and which will provide models for the use of Hawaiian ways of speaking that differ from English and are, therefore, not readily available to the vast majority of current Hawaiian language speakers and learners. This does not apply only to nascent speakers, but also to those who have been utilizing the language for years and even those who teach (and write about) the language. As we believe, opening pathways to discovery is critical, because the more you know, the more you know you don't know. However, it is important to maximize the use of all that has been bequeathed to us, finding ways to repurpose this inheritance for current contexts.

¹<https://www.aera.net/Events-Meetings/Annual-Meeting/2021-Annual-Meeting-Theme>.

An infusion of Hawaiian into Hawaiian language

In an effort to infuse more Hawaiian into Hawaiian, our late colleague and co-author, Dr. Sam L. No'eau Warner (henceforth No'eau), developed some 31 books designed to supplement a severe paucity of reading material written in Hawaiian and available to school children, as well as the broader community. These were illustrated, published, and distributed free of charge to all families in the Hawaiian Language Immersion Program. The project was funded by a federal grant, called I Mua Nō Ka 'Ulu, that No'eau had secured in order to offer well-written, well-illustrated, children's books to the Hawaiian learning community. These books were designed both to increase interest in reading Hawaiian and to correct a number of English calques that can be attributed to the prominence of English in the wider community. Such calques were, and are still, ubiquitous throughout this new Hawaiian language speaking community. The name of the project itself was Ka Lau 'Ulu.

No'eau's books represent a shift in thinking about immersion education in Hawai'i. The fact that these books were "real" in that they were hard-covered and not merely cut-and-pasted translations of English texts was not lost on the students. For one thing, English words did not bleed through the cut-and-pasted Hawaiian translations as seen in previous efforts to create curriculum materials, and for another, they were not dependent on the English ideas presented in the originals. That is, the Hawaiian was written in Hawaiian without relying on thoughts and expressions translated from English originals. Such translations clearly led to a stilted version of Hawaiian among second language learners. No'eau's books were both written in Hawaiian and designed to address certain grammatical issues that represented calques of English thought. Also, the recognition that these were "real" books was, and continues to be, of great psychological benefit, reinforcing the legitimacy of Hawaiian language immersion as a viable avenue for education, i.e., it is a "real" program and not just an English program that has been translated into Hawaiian.

In order to reach the children, and to correct the calques from English, No'eau felt that it was necessary to "go where the kids are at", as he would often describe it. This became a guiding philosophy that pervaded his efforts. He felt that Hawaiian would be relatively easy for them to learn if it was appropriately contextualized. It could perhaps be said that this subliminal nature of conveying Hawaiian to children would result in a natural form of learning language that does not generally occur in the second language classroom. That is, they would internalize these patterns despite the heavy influence from the ambient English in the community.

No'eau harbored concerns about the status quo of a community in which "someone who knows very little can easily impress someone who knows nothing". He would lament the fact that a person could merely string together a few Hawaiian street names and, to the novice learner, or to the uninitiated, pass as fluent. He wanted the community to achieve a state of full proficiency, wherein a tiny bit of knowledge was no longer worthy of remark. That is, it would be common knowledge possessed by all, not unlike the common ability to evacuate one's bowels. He devoted much effort to helping the families of immersion children, many of whom had little or no Hawaiian language background, who might otherwise place the responsibility for revitalizing the language on the next generation. No'eau recognized that, for revitalization to occur, it would be necessary for the whole family to be engaged in the effort, and not just the children. He endeavored to promote the intergenerational use of Hawaiian as normal, in an attempt to secure its future existence. He wrote:

Thus, for Hawaiian language regensis to successfully occur, there is a primary need for

the intergenerational use of the language among children who have learned Hawaiian in

the immersion schools and their parents.

Warner (1999, p. 315).

To this end, he worked to convince parents to learn as much as they could in order to bring real success to the cause. As such, he focused on families and the normal verbal interactions in which they were apt to engage. He did a lot of work outside the classroom and face to face with parents, creating opportunities to learn the language in real language use situations. He wanted them to gain a real understanding of Hawaiian and not one that is confined to a single domain, i.e., one marked by academic jargon.² Language cannot live in the classroom. Its speakers must be able to participate in a wide range of communicative events. After all, language revitalization is a community wide project, not one that is confined to a single context. And to reiterate, it should constitute a real gain, and not one that can only be appreciated by the easily impressed. That is, language revitalization goals can be deemed successful when a person's ability to speak Hawaiian is considered normal, and worthy of remark.

In sum, we intend for this work to offer some valuable improvements to the current body of literature written in or on Hawaiian. In this case, that includes the creation of illustrated books that are ostensibly meant for children, but are also of benefit to adults who aspire to know more than a little, and who wish to support the goals they have set for their children. It is important to imagine the trajectory that Hawaiian would have taken absent the overwhelming influence of English. This would help maintain a Hawaiian

²See Warner, 2001.

worldview while incorporating current ideas initially conceived in English. It would also lead to the creation of new ideas in Hawaiian.

Throughout this paper, we will present examples from No'eau's books that illustrate some of the common calques made by English speaking learners of Hawaiian. We imagine that this problem is found in many indigenous language revitalization situations throughout the world, particularly when second language learners are interacting and validating each other's calques. This paper is not intended to be critical with regard to the progress that has been made. It is meant, instead, to promote an unwavering desire to improve on the status quo, particularly under adverse circumstances, and to transcend the inclination to rest on one's laurels.

There is a lot more that can be said about the quantum leap that No'eau's books created in Hawaiian language education, but the focus of this paper is on the targeting of particular ways of speaking common to English that have come to transplant commonly used traditional Hawaiian expressions. In the present context, unless supported through education, these traditions will certainly fall into disuse and obscurity. We borrow on the idea that we should not speak one language in another, i.e., speak English in Hawaiian,³ but aspire instead to speak Hawaiian in Hawaiian. Although there is no unanimity of thought, even in the Hawaiian language speaking community, as to what constitutes real Hawaiian, it makes sense that No'eau would try to bolster those areas in which traditional forms of Hawaiian could be identified and used in order to promote a traditional way of expressing thought within a current context. This requires a different way of thinking, one that is foreign to that found in our current community. With regard to linguistic constructions, vigilance in this regard will yield a Hawaiian that is ultimately less dependent on English, but that will provide a model of language use that influences English. For now, the concern is that the closer the Hawaiian is to English, the more suspect it becomes. We aspire to maintain separation.

The linguistic calque: a sticking point for Ka Lau 'Ulu

The following are examples of No'eau's attempt to deal with the problem of speaking English in Hawaiian, which we refer to here as calquing. He chose to address the problem via his Ka Lau 'Ulu project wherein he used children's books to attack the problem. He recognized that English speaking learners of Hawaiian assume that the two languages handle ideas in the same way. This assumption of crosslinguistic equivalence is a well-known problem in language learning, and most people know (or are apprised of the fact) that, adherence to such an assumption⁴ will increase the possibility that meanings will be lost in translation. We view such loss as a sign of success in language revitalization. It perpetuates the difference in worldviews. The examples below are taken from a selection of No'eau's books:

'Okō'a ka palaoa, 'oko'a ke koholā

There are numerous examples of calquing in the Hawaiian language revitalization context. As stated above, this was one of the primary issues No'eau's books were meant to target. The following example, the title of which is shown above, is intended to help clarify one such issue. Among a number of ways of categorizing concepts, it is possible in Hawaiian to show either a relationship, or a lack thereof, between two ideas that are being compared. This example, as taken from one of No'eau's books will focus on the latter. In comparing the *palaoa* (sperm whale) with the *koholā* (humpback whale), No'eau targeted a common calque from English. He recognizes that English was set up to utilize a preposition to express this idea. That is, the *palaoa* could either be different from the *koholā*

from 4. (to express discrimination or distinction): *to be excluded from membership; to differ from one's father.*⁵

or there could exist a difference between the *palaoa* and the *koholā*.

between 6. distinguishing one from the other: *He couldn't see the difference between good and bad.*⁶

In both of these examples the meaning of one is dependent on its relationship to the other. There is some kind of connection without which there would be no point of comparison. These are common in English. In Hawaiian, however, as No'eau illustrates by way of his book, the two items are independent entities. We say: *'Okō'a ka palaoa, 'oko'a ke koholā* (the sperm whale is independent, the humpback whale is independent). As our colleague Margaret Maaka⁷ had pointed out, one does not depend on the other to derive its meaning. Just imagine the traditional Hawaiian speaker learning English and instead of saying that one differs from the other, says instead that one is independent and the other is independent. That would be both strange and amusing. Unfortunately, the linguistic power imbalance between English and Hawaiian always seems to favor the English, ultimately leading to the production of Hawaiian forms for "from" or "between".

'Okō'a ka palaoa, 'oko'a ke koholā represents a different way of understanding the world. It is no more or no less valid than the English understanding. The fact that emergent Hawaiian language speakers are saying that one differs from the other, or that there is

³An obituary article appeared in the New York Times (July 22, 1982) after the death of the well-known linguist Roman Jakobson in which a fellow linguist stated, "Jakobson is a peculiar man ... He speaks Russian fluently in six languages."

⁴Grace (1981), pointed out the existence of what he called the "postulate of intertranslatability", a belief that a phrase in one language can be accounted for by its corresponding translation in another language.

⁵1989. Webster's Encyclopedic Unabridged Dictionary of the English Language. Portland House. 1989.

⁶Ibid.

⁷Personal communication, 2009.

a difference between the two, reflects the ongoing dominance of English in the community—even in the Hawaiian speaking community. Current Hawaiian language speakers are intent on conveying their ideas to an interlocutor, even if those ideas are founded on English thoughts. Although there may be several Hawaiian ways to do so that are more in line with traditional Hawaiian thinking, they are either unaware of their existence, or couldn't be bothered to make the effort. It seems, in such cases, that the English message, by infiltrating the Hawaiian medium, acts as a sort of double agent pretending to be Hawaiian. In this case, the medium has been corrupted, and the message effectively altered. Moreover, this apparently “unnatural” construction in Hawaiian is obviously much more difficult to master, or “normalize”, than simply using the calque. After all, everyone knows the calque. We submit that, in a language revitalization context, fortifying the medium and its use should be the primary goal. Whether or not the medium is really the medium here perhaps extends Marshall McLuhan's theory⁸ in a new direction. We believe, instead, that it is worth putting in the effort to learn the Hawaiian way of speaking Hawaiian.

Pupuka 'o hope o ka pīkake!

There are a class of words in Hawaiian that indicate spatial relationships between nouns. These are called ha'iinoa honua and they represent a closed class, i.e., the class is constituted of a finite number of members, and does not accept new members. The normal pattern in which they appear is as follows: ma mua o (in front of), ma hope o (in back of/behind), ma luna o (on top of/above), ma lalo o (under/below), etc. In these cases, mua, hope, luna, and lalo are examples of ha'iinoa honua. The preposition “ma” preceding these phrases can be glossed as “in, at, on, etc.” Other prepositions such as “i” (in, at, on, to, etc.), “mai” (from), etc., also appear in this pattern. This represents the strong pattern in the current Hawaiian speaking community. But these words can also be used as nouns, and it seems that No'eau wanted to show examples of that, in parallel constructions, but with noun markers instead of prepositions. In the title above, for example, we see the subject marker 'o preceding the ha'iinoa honua (hope) indicating that it is being used as a subject. Thus, the phrase *Pupuka 'o hope o ka pīkake* can be glossed in English as *the backside of the peacock is ugly!*

No'eau noticed that young and nascent learners were following the strong English pattern when constructing this type of thought. They would say: *Pupuka ma hope o ka pīkake!* ([Something] is ugly in back of the peacock!) which is awkward in Hawaiian, or: *Pupuka ka hope o ka pīkake* (The backside of the peacock is ugly), which, despite yielding an intelligible translation in English, is also awkward in Hawaiian. He realized that these learners had not been exposed to the entire range of possible uses for these words. As such, this book is ostensibly designed to offer the common Hawaiian way of expressing the ha'iinoa honua as subjects simply by using the subject marker 'o, a marker that does not have a gloss in English. The use of 'o here as a subject marker is a sign of sophisticated language use not generally exhibited by second language learners. No'eau recognized that the language of wider currency in the community, i.e., English, would seep into areas where a pattern is less commonly used, and second language speakers would be inclined, at least subconsciously, to reproduce it.

Finally, the word *hope*, as indicated above, can refer to the buttocks. Although this might be considered inappropriate by some, particularly for younger children, we believe that such ideas of appropriateness are also imported from English and are not applicable to Hawaiian mores. The use of such terms and the raising of such topics are both appropriate in Hawaiian, even for children. Words relating to body parts and bodily functions were traditionally not censored in Hawaiian.⁹ We believe that English values have supplanted the traditional Hawaiian values with regard to the use of these words. Having said that, it is quite possible that their censorship has contributed substantially to the level of interest in, as well as the humorous nature of their use in No'eau's books.

Pakalaki

This book is ostensibly designed to normalize several “normal” Hawaiian language grammatical patterns for second language learners whose first language is English, and for whom these patterns are not acquired as part of their normal first language experience, i.e., via exposure to the ambient language of a speech community. The differences between first and second language speakers, with regard to how they handle particular speech situations, can present problems for the second language learner. In other words, the constructions required in the second language can be significantly, and unexpectedly more complex. Being constructed based on a different worldview, they are likely to be handled differently. Due to this inherent complexity, their introduction to the learner is often delayed until later in the learning process, wherein they are explained via a comparison to English. The inherent difference between the worldviews of each results in the solidification of English versions of Hawaiian in those contexts. That is, novel Hawaiian forms of speech would ultimately be constructed and understood via English ways of thinking and speaking.

No'eau saw an opportunity to bypass this dependence on an English thought-processes when divining meaning from Hawaiian texts. By utilizing context and repetition, he sought to reduce the powerful influence from English on our understanding of Hawaiian. He utilized several patterns common to Hawaiian speech. The three patterns of focus in Pakalaki are listed below:

1. Akā, he keu 'o ia a ka pakalaki!
However, he/she/it was extremely unlucky!

⁸see McLuhan, 1964.

⁹See Wong and Warner, 2019.

2. I ka manawa a Pakalaki i pā'ani ai ma kahi o ka 'ilio ...
When Pakalaki played near the dog ...
3. ua nahu 'ia maila kona wāwae e ka 'ilio nunui.
His/Her/Its foot was bitten by the large dog.

The first pattern (1) utilizes the words "keu ... a" in order to convey a superlative sense of the level of bad luck in this context. This was ostensibly designed to promote a Hawaiian way of expressing the superlative that might prevent the spread of English options not traditionally utilized, such as "Ua pakalaki loa, loa, loa, 'o ia!" (He/She/It was really, really, really unlucky!). There are a number of patterns that can express the superlative, many of which are idiomatic such as "keu a", with "keu" representing an excessiveness that transcends normal experience, and "a" is a type of preposition that assigns that excessiveness to a particular trait. This is not an uncommon phrase in Hawaiian. It is heard in a number of popular songs. However, because it is idiomatic, there is no appropriate English equivalent upon which to construct the Hawaiian. Learners would either have to learn to think in Hawaiian in order to produce it, or experience its use until it is internalized at the subconscious level. The amount of bad luck experienced by Pakalaki in this story is a great way to internalize this pattern.

The second pattern (2) contains both a relative clause to express time and a locative phrase that indicates proximity. The relative clause *ka manawa a Pakalaki i pā'ani ai* (the time that Pakalaki played) is one of the more difficult patterns for students to acquire because there are numerous permutations of the pattern with the choice of one or the other likely based on the grammatical context. This particular relative clause is built on the phrase *ka manawa* (the time) providing more information about the particular time in question, i.e., *a Pakalaki i pā'ani ai ma kahi o ka 'ilio* (when Pakalaki played near the dog). Due to its inherent complexity, teaching the various permutations of the Hawaiian relative clause has been the bane of many Hawaiian language teachers. We do not flesh out all the various permutations here, but suffice it to say that it is considerably less complex in English. In our opinion, the nature of its complexity suggests that it was likely less commonly used in Hawaiian than in English. As such, the motivation of English speakers learning Hawaiian as a second language is that they will be able to express their English thoughts in Hawaiian, as suggested by Grace's "postulate of intertranslatability" (see footnote 4). No'eau recognized that consistently using the pattern in appropriate contexts would be the model that learners would internalize when reading his books or hearing them read aloud. This would make sorting out the grammatical complexities moot.

Also in the second example, we find a prepositional phrase that locates where Pakalaki played, i.e., *ma kahi o ka 'ilio* (near the dog). The Hawaiian construction is comparatively more complex than its English counterpart. For one thing, it contains the contraction *kahi* which is made up of two words, *ka* (the) and *wahi* (place). Instead of a single word, i.e., "near", Hawaiian employs the phrase *ma kahi o* (at the place of). No'eau recognized that students were trying to say "near" by using the word "kokoke" yielding constructions such as "ua pā'ani 'o Pakalaki kokoke i ka 'ilio" which does not conform to normal grammatical structures in Hawaiian. In other words, there should be a preposition to introduce the phrase, e.g., *ma kahi kokoke i ka 'ilio* (at the place near the dog). Resorting to English constructions clearly leads to unusual expressions in Hawaiian. This particular example has taken hold of a large portion of the current Hawaiian speaking community, and No'eau tried to effect change by influencing speech through his books.

In example (3), No'eau utilizes another contraction *kona* or *ka + ona* (the + of him/her/it). As we can see, the word *ona* here is also a contraction, i.e., *o + ia* (of + he/she/it). One of the lessons to be taken from this contraction is that Hawaiian does not have gendered pronouns. This offers a major benefit to the speaker of Hawaiian, that is, there is no reason to disambiguate the referent by utilizing pronouns appropriate to the gender, and all the current confusion that an ever-increasing array of English pronominal identities has created is simplified. Hawaiian utilizes the word *ia* to denote the third person singular pronoun. This accounts for he/she/it/or any other identity that is not included therein. Whereas identification via gendered pronouns is not an issue for Hawaiian speakers, it might not have been No'eau's intention to utilize this contraction as a lesson. English speaking learners of Hawaiian do not seem to struggle when the Hawaiian system is less complex than that of English. It is only when the Hawaiian is more complex that English speaking learners tend to struggle.

Also, in example (3), No'eau utilizes a passive construction: *ua nahu 'ia maila kona wāwae* (its foot had been bitten). English speakers, and writers, are more inclined to use the active voice. Thus, the English-speaking learner of Hawaiian is likely to favor the active construction that points more directly at the agent, which appears as the subject in the initial position. That is, it is much more accusatory than the passive construction that is common in Hawaiian, and in which case the agent is relegated to the third position: *e ka 'ilio nunui* (by the large dog), constituting a less direct, and less accusatory way of indicating the agent¹⁰. This is much more reflective of a Hawaiian world view that recognizes the power of words to visit tangible harm upon others. The importance of avoiding such harm is a way for a community to maintain social harmony.¹¹

Of all of the English influenced ways of speaking that No'eau's books were designed to address by offering Hawaiian alternatives, there are many still constantly heard. The net result is that such English constructions, due to the current lack of counter evidence, i.e., traditional alternatives, become normalized, and ultimately entrenched in the psyche of the second language speaker. That is the power of old habits. Even when we are aware of the appropriate alternatives, we are unable to make the change. The habitual form has worked its way into the subconscious level, not unlike "I mean", "you know", and countless others, to the point that we are unaware that we are actually producing them, even as we speak Hawaiian. We believe that No'eau recognized this and tried to

¹⁰See Wong (2006).

¹¹See Pukui et al. (1972).

do something about it for new learners before the English forms became fossilized. The counter evidence memorialized in his books offers an opportunity to make such a change surreptitiously, i.e., without confrontation, without risking social discord or embarrassment, and without casting blame!

The continuation of Ka Lau 'Ulu

After publishing and distributing the first set of 31 books, No'eau fell victim to a stroke and passed away after a long battle. This represented a tremendous loss for the Hawaiian language community. Although many people recognized the adverse impact English was having on Hawaiian, No'eau was really the prime catalyst for change as evidenced by the popularity of his books. The mere exposure to Hawaiian alternatives for expressing contemporary thoughts, as presented in an attractive and understandable context, supports the continuing viability of Hawaiian ways of speaking within current contexts.

As it turns out, Kapā'anaokaōle Oliveira, one of the co-authors of this paper and one of No'eau's mentees, found another 21 books that No'eau had developed until near publication ready, but had not published. She took it upon herself to seek funding to have the books published and distributed in exactly the same way that No'eau had done with the first set. Again, all the books were well-received, and because they target grammatical issues that are primarily calques of English, they are valuable lessons for a program that is the primary vehicle for creating new Hawaiian language speakers. The following are a couple examples of this effort.

Aia 'o Māmā i ke kaula'i lolo

This book represents the repetitive pattern of question and response in a basic children's story. The younger brother repeatedly asks his older brother about the current location of a certain member of the family. The older brother responds in a way that is appropriate in Hawaiian but strange in terms of its English translation. Here is the first example:

E Kua'ana, aia lā i hea 'o Māmā? (Hey, older sibling of the same gender, where is Mommy?)
'O Māmā? Aia 'o Māmā i ke kaula'i lolo. (Mommy? Mommy is located at the laundry hanging.)
Mai ho'oluhi aku 'oe iā ia. (Don't you bother her.)

No'eau has taken the opportunity here to teach several lessons. The first involves the appropriate use of the dubitative marker *lā* which is not commonly used by second language learners, and has caused confusion with regard to its syntax. That is, many people think the syntax is as follows: *Aia i hea lā 'o Māmā?* The question seeks to ascertain the location of Māmā which, from an English perspective, would be the information that is unknown. However, this marker in Hawaiian is located after the predicate *aia* denoting the existence of a subject in space or time, and not after the prepositional phrase *i hea* (at where). Obviously, children do not need to know the reason for the grammatical order, they just need to hear it occurring in the normal position on a regular basis. We believe that this lesson, and its repetition in each scene, was designed to reinforce the normal use of this word.

There is another lesson that occurs in this first scene wherein the word *Kua'ana* or *older sibling of the same gender* is used as the appellation for the older brother, instead of the first name as we would expect in English. It is spelled with an upper case "K" because it is a name that is akin to a title. Moreover, it represents the inherent responsibility relationship an older sibling has with a younger sibling of the same gender. In English, by comparison, it is not uncommon for the younger brother to call an older brother by his first name. In many cultures, this might be considered disrespectful.

Another feature that was important for No'eau to reinforce was the appropriate response to a question about where someone is. If it is a location that is being sought, that is easy enough. One merely names the place in a prepositional phrase. However, Hawaiian can also mark an activity as a location, as is done here and throughout the story. This is clearly the primary lesson. When asked where Māmā is, the older brother says literally that *she is at the laundry hanging*. Again, this seems awkward from an English perspective, but not any more so than an English phrase does when translated into Hawaiian, e.g., *How old are you?* *Pehea kahiko'oe?* Hawaiian speakers, like those of some other cultures, are interested in how many years one has, not how "old" one is. Using the Hawaiian word for old, *kahiko*, in this question would make little or no sense. There are numerous other examples that illustrate the perils of direct translation. These can even be humorous, but when the learner tries to apply the range of meanings of a particular phrase in the first language cross-linguistically, it does not match up with the range of its phrasal counterparts in the second language. Unfortunately, this is something of which second language speakers of Hawaiian (as well as other languages) are generally unaware. Left to their own devices, they are liable to speak English in Hawaiian.

He aha ka inoa o ka holoholona e umō mai nei?

In this book, the flap technique is utilized wherein the first page asks a question about which animal makes the sound written on the next page. Upon opening the flap on that next page, the reader discovers the answer, both in the text and in picture form. Throughout the book, each animal makes a unique sound. The first up is the cow, and it is the one that makes the *umō* sound that is presented in the title. The first page asks: *He aha ka inoa o ka holoholona e umō mai nei?* (What is the name of the animal that is *umō*-ing?) Upon opening the flap, the answer appears along with the picture of a cow: *He pipi ka inoa o ka holoholona e umō mai nei!* (The animal that is *umō*-ing is a cow!).

Another lesson that should be noted here is that, unlike humans, cows are not considered to speak different languages. Nonetheless, the sounds they make are represented differently in various human languages. Instead of the common English representation,

"moo", the Hawaiian is employed. For language revitalization efforts throughout the world, various sounds are likely represented in ways unique to the language spoken in particular areas.

No'eau recognized that the expression "*O wai ka inoa o ... ?*" (Who is the name of ... ?) is the form most learners use for articulating this question, which is prevalent in inquiries about the personal name of a third party not involved in the conversation. One of the first lessons second language learners encounter in beginning level Hawaiian language classes is the direct question to an interlocutor "*O wai kou inoa?*" (Who is your name?).¹² Note that the word "wai" is translated as "who" when asking about a personal name. No'eau realized that, at a number of levels (e.g., spiritual, taxonomical, and grammatical), the word "inoa" does not translate perfectly to "name" in English. In this case, the type of animal has a taxonomical name and the particular animal itself might have a personal name. That is, the animal is referred to at a taxonomical level by the categorical name "pipi" (cow), but it might have a personal name such as "Lani Moo".¹³ Thus, a question about its name is constructed differently based on the type of name being sought. "*O wai ka inoa o ka pipi?*" would yield a response such as "*O Lani Moo kona inoa*" (Her name is Lani Moo). Instead, "*He aha ka inoa o ka holoholona?*" would yield "*He pipi ka inoa o ka holoholona.*"

The difference between taxonomical names and personal names has been leveled for many current speakers of Hawaiian. The interrogative "wai" (who) has displaced its colleague "aha" (what) to become the only avenue of interrogation when seeking either type of name. As such, No'eau made this problem the primary target of this book. Interestingly, the English gloss for "what" is, in fact, the appropriate English interrogative for seeking either the personal or taxonomical name, e.g., What is the name of the animal? Its name is Lani Moo. It is a cow. Interestingly, that being the case, one might expect that "*He aha ka inoa o ... ?*" would have been calqued before "*O wai ka inoa o ... ?*"

Next steps

No'eau has left a valuable legacy providing a map for future efforts to develop curriculum materials that effectively support the continued efforts to revitalize the Hawaiian language, and he did so in a way that maintains Hawaiian ways of speaking and knowing. As Fishman has stated, this is not an exercise in "backwards looking"¹⁴ which would leave us stuck in the past as we strive to maintain authenticity. Instead, this is about maintaining those ways of speaking and knowing that are constructed upon past models as we move forward into the future. This would necessitate the inclusion of technological changes in current curriculum materials and curriculum development. That is, this must necessarily apply both to subject matter and medium of delivery. As such, it is important, as No'eau has said, to go to "where the kids are at". In the current context, this means leaving a digital footprint on Hawaiian education.

No'eau's plan inspired Kapā'anaokalāokeola to expand access to the books by making them available online as audiobooks. While ebooks and audiobooks are commonplace in English 21st century early literacy, this marked a milestone for Hawaiian immersion education. For the first time, children could access Hawaiian audiobooks online. No longer was the breadth of distribution bounded by the limited quantity of physical books produced on grant funding; rather, the entire collection of 52 books was digitized and uploaded to www.noeau.org. Overnight, the books were available to anyone with internet connectivity. To catapult the movement further, a native speaker of Hawaiian was employed to read the texts aloud so that children would be able to mimic the voice of a native speaker of Hawaiian. This was a game changer because for many children, this was the first time that they were hearing the voice of a native speaker of Hawaiian.

While new technologies expand access, we still believe in the continued production of physical books that promote the tactile learning process. The model presented by No'eau clearly does this, and more funding should be sought and secured in order to produce a wider range of genres while also targeting the ever-present calques from English. No language revitalization effort can really be considered successful if it is primarily beholden to English for the language that is ultimately generated. As such, No'eau's model would be extremely valuable to other language revitalization efforts throughout the world, particularly those wherein a dominant language has hijacked native thought. It does not make sense merely to express thoughts from the dominant language by utilizing vocabulary and grammatical structures from the target language. Attention must be paid to world view as well, and a world view that is independent from English can be found in both the grammar and vocabulary of traditional Hawaiian. The goal of No'eau's work was NOT to speak English in Hawaiian. He did not see that as a worthwhile cause. One could simply speak English in English if that were the case. It was, instead, to speak Hawaiian in Hawaiian, maximizing the use of the knowledge that has been bequeathed to us by our ancestors. He left a valuable model. It is up to us to step in and continue the movement on *ke ala o ka no'eau*, i.e., *the path of wisdom*.

It is up to us to sense the *hō'ailona* that our *kūpuna* have bequeathed to us so that we too can add new *'ike* to our knowledge base for subsequent generations of *Kānaka*.

Oliveira (2014).

¹²It should be noted here that the authors consider asking an interlocutor (second person) his or her name without establishing a foundation to do so is somewhat presumptuous from a Hawaiian perspective. In this case, however, it is not the focus of analysis.

¹³Lani Moo was the name of the mascot representing Meadow Gold Dairies in Honolulu. She was a cow who had been suffering from cancer and was recently euthanized during the second week of April 2022.

¹⁴See Fishman, 1991.

References

- Fishman, J., 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Endangered Languages*. Multicultural Matters, Clevedon, England.
- Grace, G., 1981. *An Essay on Language*. Hornbeam Press, Colombia SC.
- McLuhan, M., 1964. *Understanding Media: The Extension of Man*. Signet Books, New York.
- Oliveira, K.R.K.N., 2014. *Ancestral Places: Understanding Kanaka Geographies*. Oregon State University Press.
- Pukui, M.K., Haertig, E.W., Lee, C.A., 1972. *Nānā i ke kumu*, vol. II. Hui Hānāi, Honolulu.
- Warner, S.L.N., 1999. Hawaiian language renaissance: planning for intergenerational use of Hawaiian beyond the school. In: Huebner, T., Davis, K.A. (Eds.), *Sociopolitical Perspectives on Language Policy and Planning in the USA*. John Benjamins Publishing Company, pp. 313–332.
- Warner, S.L.N., 2001. The movement to revitalize Hawaiian language and culture. In: Hinton, L.H., Hale, K. (Eds.), *The Green Book of Language Revitalization in Practice*. Academic Press, pp. 133–144.
- Warner, S.L.N., 2009a. Pupuka 'o hope o ka pīkake! Ka Lau 'Ulu.
- Warner, S.L.N., 2009b. 'Ōko'a ka palaoa, 'oko'a ke koholā. Ka Lau 'Ulu.
- Warner, S.L.N., 2009c. Pakalaki. Ka Lau 'Ulu.
- Warner, S.L.N., 2019a. Aia 'o Māmā i ke kaula'i lōle. Ka Lau 'Ulu.
- Warner, S.L.N., 2019b. He aha ka inoa o ka holoholona e umō mai nei? Ka Lau 'Ulu.
- Webster's Dictionary, 1989. *Webster's Encyclopedic Unabridged Dictionary of the English Language*. Portland House.
- Wong, K.L., 2006. *Kuhi Aku, kuhi mai, kuhi hewa ē: He mau ioina kuhikuhi'ākena no ka 'Ōlelo Hawai'i*. PhD Dissertation. University of Hawai'i at Mānoa, Hawai'i.
- Wong, K.L., Warner, S.L.N., 2019. He pelapela anei ka 'ōlelo a ka Hawai'i? In: Tomlins-Jahnke, H., Styres, S., Lilley, S., Zinga, D. (Eds.), *University of Alberta Press*.

Indigenous language revitalization in the United States and Canada

Tiffany S. Lee^a and James McKenzie^{b, a} Native American Studies, University of New Mexico, Albuquerque, NM, United States; and
^b College of Education, University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	50
Context for language change	51
Indigenous language change	52
Basis of need for Indigenous language revitalization	53
Indigenous language revitalization and Indigenous immersion programs	54
Planning for Indigenous language revitalization	54
Indigenous immersion education	54
Structural and contextual differences	56
Non-school-based ILR programs	57
Challenges and supports to Indigenous language revitalization	58
Conclusion	58
References	59

Glossary

Native American, First Nations, Indigenous, Native, Aboriginal: terms used interchangeably to identify the first peoples of the land in what is now the United States and Canada.

Introduction

This article will first cover the context of Indigenous language change and shift in the United States and Canada to provide the basis for understanding the need for language revitalization efforts. The campaign to eradicate Indigenous languages and thereby, their cultures, knowledges and traditional educational practices explains the overarching context for language change. This campaign created contexts of language endangerment where few to no speakers are left in some Indigenous communities. By some estimates, 90% of Native American languages are vulnerable reaching a point with no child speakers (Moseley, 2010). This first half of the article will also address the basis of the need for Indigenous language revitalization in terms of the important relationship of language to identity, wellbeing, and sovereignty.

Recognizing language change, Native American and First Nations communities have been involved in many types of language maintenance and revitalization efforts (Hinton and Hale, 2013; McIvor, 2009; Reyhner and Lockard, 2009). In the second half of this article, we move toward highlighting some of the diverse efforts for Indigenous language revitalization in the United States and Canada. While the article will privilege research on language immersion programs, schools, and efforts, it will also include community-based approaches with less research publication. Throughout the article, we use terminology in line with current research and framing of language change with limited use of damage-oriented terms (Tuck, 2009) that designate “loss” or “death.” Instead we use terms and framing that honor Indigenous languages’ important connections to identity, heritage, and living communities. In this sense, it is more respectful to describe languages as “sleeping” and with a focus on revitalization and wellbeing and as part of Indigenous “futures” (Smith et al., 2018). Smith et al. (2018) provide an eloquent description of Indigenous futures when they state,

We can use the word future to describe a time that comes after now, a time that we will come to inevitably. We can also use the word futurity, a word that imbues the future with what we are doing now to bring about different futures. The authors in this book attend to what we are doing now to bring about the futures we can’t even fully imagine yet. (p. 23)

We also use the terms Indigenous and Native as synonymous with Native American and First Nations. First Nations is a common term used in Canada, while Native American is common in the United States. We utilize both the terms language revitalization and language reclamation throughout the article, drawing on conceptual frameworks that do not limit such efforts to language-as-code and proficiency driven focus, but instead those that are implemented and experienced in Indigenous communities as guided by community needs, values, and goals, and grounded in Indigenous ways of being (Leonard, 2012, 2017; McKenzie, 2020). Finally, as we approach Indigenous language revitalization we offer critical questions with regard to some existing literature and its framing

of language change among Indigenous communities. The research we review here is only a sample of the extensive literature base that we wish we could include. While a comprehensive review would be beyond the scope of this article, we think you will find that the examples below provide an illuminating overview of current and emerging work.

Context for language change

Lomawaima and McCarty (2006) provide a thorough overview of the historical context for the campaign to eradicate Indigenous languages. The authors critically evaluate the historical educational landscape for understanding US educational policies and practices, which were aimed at the assimilation of Native students into the dominant American society through the suppression of Native students linguistic and cultural heritage. Interestingly, while those policies shifted over time between more and less oppressive, Lomawaima and McCarty argue they never leave the “safety zone.” The safety zone is a theoretical understanding where the authors argue that efforts to integrate Native American cultural and linguistic practices in schools have been allowed or disallowed based on whether such practices were deemed “safe.” That is, they did not threaten the ultimate goal of assimilation to an English language American identity. During some historical periods certain educational policies and practices have been deemed safe by federal agents because they posed no threat to American identity or economic superiority, while at other times similar policies and practices were considered a threat to that identity or American dominance and were therefore perceived as unsafe. In the end, Native American sovereign expression of linguistic and cultural heritage is held hostage through the power dynamics of federal policies. Lomawaima and McCarty raise important questions on the role of schools for creating American citizens, while at the same time fore fronting Native American peoples’ desire to control and oversee the education of their own people—a desire that eventually materializes in language revitalization efforts, among others.

In terms of learning about the trends in Native American language policy, Warhol (2011) provides a concise overview of federal, state, and tribal policies to demonstrate the shift in language policy to protect Indigenous languages. Warhol begins with an historical account of language policies as they were enacted in government boarding schools and moves to a discussion of policies in support of bilingual education. However, those policies were not necessarily focused on Indigenous language revitalization due to an emphasis on English-only. She then examines the Native American Languages Act and additional legislation that has helped or hindered Indigenous communities’ efforts at language revitalization. She ends her overview with attention to specific state and tribal policies as they have supported teacher certification, curriculum development and dual-language and language immersion education.

Similarly to Lomawaima and McCarty, Hinton (2016) argues that the state of endangerment of Indigenous languages today in the US and Canada is the result of state government suppression. Through schooling, policies, and laws, Indigenous languages were forbidden and repressed. However, schooling practices, policies, and laws starting in the 1970s shifted toward supporting their revival and sustainability. She also discusses the role Native communities have played to revitalize their own languages through immersion schools and innovation in pedagogical techniques for educators who learned their Indigenous languages as second languages.

In a more focused study on the impact of the campaign to eradicate Indigenous languages on Native youth, Romero-Little et al. (2007) examine the personal and learning experiences of Native youth as a result of language shift. Romero-Little et al. conducted 230 interviews, hundreds of hours of observations, and collected responses on 600 questionnaires from youth and teachers to understand their perspectives and experiences in this federally funded study called the Native Language Shift and Retention Study. Overall, they found that students and teachers alike valued English language but not at the expense of their Native language. Teachers also had a different perception than the students themselves of how much students valued their Native language. While 85% of Native students rated their Indigenous language as “very important” to learn and speak, only 29% of educators believed that Native students felt it was “very important.” Grandparents also had a stronger desire for their grandchildren to learn their Native language than did parents. Further, youth reported feeling linguistic shame, which meant feeling judged as inferior in social and intellectual status if they have an accent or if they demonstrated in front of their peers that they could speak the language. These findings are important for their implications on educational policy and practice. Federal and state policies at the time prioritized adequate yearly progress, which was measured by standardized English-based tests. These policies ignore the value of Native languages and center English-based curriculum, often through direct instruction methods that were lacking in substance and critical thinking, in order to improve test scores.

In response to such findings, McCarty and Nicholas (2014) offer insight into the roles and responsibilities of schools in reclaiming Indigenous languages. They explain the difference between natural language change and language shift, the latter being when the social structures supporting intergenerational transmission of language break down as a result of colonial or dominant control and suppression of Native languages. In this chapter, the authors review school-based language reclamation efforts followed by a thorough review of the language contexts, colonial policies, and Indigenous counter-movements and reclamation efforts across the US and Canada. They ask if schools, originally a site of Indigenous language destruction, can become effective sites for language reclamation. They argue it is the responsibility of schools to reverse the damage that they created before going on to offer a comprehensive overview of well-established, school-based (and community-based) language reclamation and immersion programs in several Native communities across the US and Canada.

In an example that highlights a wider range of potential factors influencing language shift, Spolsky (2002), who conducted a study on Diné (Navajo) language shift in the early 1970s, returns to the question of the influences on language shift to firmly

hold schools accountable for creating that change. But, he argues, larger, complex social, economic, and political contexts and influences also account for language shift among Diné people. He provides an overview of the state of Diné language and the state of Diné life across religious, political, economic, and educational settings, and across personal ideological perspectives, noting that language shift must be examined under this fuller sociohistorical context to understand contemporary conditions and to determine how to intervene.

Indigenous language change

This next section provides an overview of research that attempts to understand the stage at which language shift is taking hold and a critique of the perspectives under which Native communities and their languages are portrayed and understood.

McIvor and McCarty (2016) share that census data report that 169 Native languages are still spoken by 370,000 people in the US, and 240,815 First Nations people in Canada report speaking an Indigenous language at a conversational level in Canadian census data. While Krauss's research is more dated (1996), he provides a more nuanced understanding of the state of Native languages in the US and Canada by categorizing languages based on the age of the remaining speakers. Krauss claims 210 Indigenous languages across the US and Canada were still being spoken at the time of his research. He categorizes them into Class A – those with child age speakers and older, Class B – those with parental age speakers and older, Class C – those with grandparent age speakers and older, Class D – those with only elderly age speakers, and Class E – those language communities with no speakers left. At the time of his research, Class A had the smallest number of speakers, followed by Class B. Class C and D had the largest number of speakers. Like Sallabank (2012), he predicted that by the 2000s, between 20% and 50% of the world's 7000 languages would be at Class B or weaker. Sallabank predicted half or more would not be spoken by the end of this century. In her chapter Sallabank examines reasons for loss of linguistic diversity and what language policies can do to intervene. She discusses definitions of "language death" and notes how it is problematic to identify and assess when a language reaches "death." More so, languages exist through people and to suggest a language has died is to suggest the people have died. She provides examples of the many revitalization efforts of "dead" languages suggesting such use of the term is inaccurate and insulting. She continues to address the causes and influences that contribute to language shift and endangerment and deliberates measurements of linguistic diversity and health of languages. She reflects on the importance of linguistic diversity for the world and ends the chapter with a comprehensive overview of policies and practices that support languages facing more pronounced shift. What is unique is her assertion of policy makers as inclusive of government officials, language planners, community members, activists, members of language revitalization movements, as well as linguists or researchers.

Understanding language shift is complex, and while many scholars like Krauss provide much needed insight into that shift, other scholars like Sallabank convey the importance of discussing such change within a more holistic framework and less of a deficit-based portrayal, for the health and wellbeing of Indigenous communities involved in research. Fitzgerald (2017) argues that any discussion and predictions of Indigenous language vitality must include efforts at language revitalization in an effort to understand language change and shift through a framework based on Indigenous resilience, rather than a solely deficit-oriented focus on language shift, alone. In her commentary and literature review, Fitzgerald provides a thorough overview of efforts made by linguists and language communities to respond to language change, shift, endangerment, and loss. She synthesizes research for two decades of work in education, language revitalization, Indigenous studies, and health and well-being to inform conceptions and theories of language vitality and resilience. Her review includes key aspects of national policies, definitions of "language," metrics of vitality and success in language revitalization, and understanding the benefits of language revitalization. She asserts that defining the success of language revitalization in a holistic way accounts for better outcomes in health and education and supports findings that language maintenance and revitalization have benefits, while also acknowledging the negative effects of language shift within Indigenous communities in the US and Canada.

Benally and Viri (2005) and Lee and McLaughlin (2001) provide focused studies on Diné language of the US. Lee and McLaughlin present an analysis of social, economic, political, religious, and educational conditions that shape language and culture shift, and efforts to reverse it, within the Navajo Nation. They explain the status of Diné language in relation to each of the eight stages of Fishman's Graded Intergenerational Disruption Scale (GIDS) (1991), providing commentary on the appropriateness of Fishman's model for conceptualizing and working to reverse language shift. Benally and Viri assert that the Diné language is at a crossroads, from which it can move toward further decline, or toward a strengthened state through revitalization. They suggest the direction the language moves depends on Diné people and stress the importance of understanding the colonial (and internalized) influences to oppress Diné language along with the natural changes in adaptation to maintain a living language. It is important to note that Benally and Viri also explore the presence of "Navajo English" among Diné people, which is a form of English that has been influenced by the structure, accent, and cultural context of Diné language and people. They recognize Navajo English as a resource for strengthening Diné language.

Focused on another regional context, in a case study on the Indigenous languages of Oklahoma, Chiocca (2020) draws on Fishman's GIDS scale noted earlier to offer a critique of non-Native linguists and anthropologists' stances and viewpoints on language revitalization, arguing that Native communities must decide for themselves what constitutes success within the context of the demographic linguistic landscape of their communities. She shares examples of Native communities' language renewal efforts within Oklahoma and discusses their comparative growth and achievements. Importantly, she points out that it is the diverse contexts of each community's sociolinguistic environment that is important for understanding their linguistic goals and measures of success.

Basis of need for Indigenous language revitalization

This section will examine the relationship of Indigenous languages to Indigenous identity, wellbeing, and sovereignty. We position these relationships as the basis of the need for language revitalization in Indigenous communities. Grounded in the earlier described assertions that language change and revitalization must be examined more holistically in relationship to each community's social, historical, political, and other conditions, this section highlights scholarship and revitalization efforts within Indigenous communities who demonstrate the value and significance of their languages to identity, wellbeing, and sovereignty as aspects of Indigenous lifeways.

As will be shown in a later section of this article, Indigenous language immersion programs and models provide the most effective means for Indigenous language revitalization, encompassing approaches that enhance the holistically positive effects on learners. [Hermes and Kawai'ae'a \(2014\)](#) highlight immersion models in Ojibwe, Māori, and Hawaiian contexts. They illustrate how the nature of each program is shaped around building connections to family and community and result in influencing a strong language learner identity. Learners' language identities are strengthened and supported through immersion programs' revitalization of cultural traditions and practices and through reinforcing connections between learners' identities, languages and homelands.

The topic of the relationship of language to Indigenous young adult identity is taken up by [Lee \(2009\)](#) in her research with Diné and Pueblo youth. Her research investigated how young adults negotiate messages they receive about the importance of their Indigenous language for cultural continuity against messages of the need for English for success in the dominant society. The young adults in her study provided counter-narratives that demonstrated their perspectives of language and its relationship to their identity. Their perspectives were conveyed across five themes: respect for their language, experiences with stigmatization and shame regarding their language competencies, feelings and experiences of marginalization, these experiences' impact on their sense of their Native identity, and expressions of agency and ways to intervene to strengthen language revitalization within themselves, their families, and communities. Overall, regardless of their language fluency levels, the youth's sense of their Indigenous identity was not diminished. However, most of the young adults in the study expressed strong desires to improve their language fluency because of their value for their language and to increase their cultural knowledge through the language.

Related to identity is the issue of Indigenous wellbeing and its important connections to Indigenous language. [McCarty \(2020\)](#) addressed Indigenous individual and collective wellbeing by examining several Indigenous educational and communal program features that result in education for language revitalization and reclamation (ELR2) while at the same time recovering Indigenous practices of mutuality, relationality, respect, and equality. These features and practices lend to an environment that strengthens individual and collective wellbeing. McCarty cites examples from across North America and the Pacific to demonstrate these communal responses for revitalization and reclamation, approaches that take place in and out of school. She emphasizes the holistic nature of each community's approach in ELR2 within their unique sociolinguistic contexts. She concludes with how ELR2 influences wellbeing and holds beneficial implications for Indigenous self-determination, sustainable diversity, and social justice.

Pathways on a Language Landscape: A Planning Guide for Native Language Revitalization (n.d.) is a publication that was produced from a two-year planning process, a collection of Native community perspectives, and a conference in 2016 which gathered over 200 language educators, activists, scholars, community members, and leaders to share their visions for language revitalization. The guide is a framework of planning practices to aid Native communities in developing Native-led language programs. While providing an historical overview of language change in Native communities and a critique of the role of settler colonialism as the root cause for language shift and endangerment, the guide positions Native identity, wellness, knowledge, survival, sovereignty and liberation as the primary purposes for language revitalization and reclamation. It demonstrates the integrated relationships between these elements. The language landscape guide then provides nine common language practices, goals and strategies for achieving each practice, and shares resources for supporting Native communities' language planning. The guide is meant to engage all members of the community as a way to realize community renewal and building Native nations.

With profound implications for wellbeing, one Canadian study on the relationship between language use, other cultural continuity factors (such as land claims, education, and healthcare), and Indigenous youth suicide, pointed out that language use alone held a significant relationship with a drop in youth suicide rates ([Hallett et al., 2007](#)). Those communities where at least half of the population held a conversational level of knowledge of their Native language correlated with a suicide rate in their community that dropped to zero. Such results demonstrate that the more the Native language is present in a community, suicide rates are less. This suggests a direct, important connection between Native languages and wellbeing.

The research cited up to this point provides a holistic view of the connections between Indigenous languages and a variety of social and environmental factors that impact Indigenous lifeways. However, most unique, and perhaps most important, is the connection between language and Native sovereignty. [Blum-Martinez \(2000\)](#) argued that Pueblo languages should not be treated in similar ways to languages like Spanish in bilingual programs in schools. Pueblo cultures, histories, and communal and political structures view their languages and how they are taught or transmitted as to be controlled by their peoples for their communities. They are exercising their sovereignty through decision making about where, when, and how their languages are taught to their community members, and by whom. Blum-Martinez's work illuminates the fact that educational sovereignty is an act of community building and change to reflect community needs, resources, and goals.

Similarly, [Wilson and Kamanā \(2006\)](#) and [McCarty and Lee \(2015\)](#) discuss the significance of sovereignty and the role of schools in modeling the values and knowledge of the Indigenous communities they serve. Wilson and Kamanā raise the issue of a ban on Hawaiian language education as an attack on Hawaiian people as a distinct people with rights to learn and strengthen their language. They describe how Hawaiian-medium education has benefits that outweigh English-medium education and that do

not hold back students' progress in English language. [McCarty and Lee \(2015\)](#) also assert that Indigenous language education, and Indigenous education in general, must be advanced by Indigenous peoples. Native American language and culture programs operate in and out of schools in every state and US territory. In every case, these initiatives are embedded in the cultural context of local communities, highlighting the significance of context and place for creating a socially, linguistically, and culturally responsive learning environment. Thus, Indigenous experiences with language and cultural renewal are not homogeneous; while similarities and parallels can be drawn, what is appropriate in one context cannot be simply transported or translated to another.

Additionally, [Wong \(2004\)](#) critiques the academy, specifically Western institutions of higher education, with attempting to control the space for Hawaiian language. He asserts this is space that is still occupied by English, and explains the struggles Hawaiian language advocates still face to attain a space free of constraints and limitations.

In the next section, we share specific Indigenous language immersion programs and other language learning models across the US and Canada. This is not a comprehensive list, but it is a diverse list. [Hinton and Meek \(2018\)](#) and [Hinton and Hale \(2013\)](#) provide a comprehensive overview of the language revitalization movement, immersion models, mentor-apprentice, and other unique language educational approaches across Indigenous North America. The Green Book ([Hinton and Hale, 2013](#)) in particular moves away from deficit-oriented portrayals of the state of Native languages by sharing the optimistic and encouraging models showing success, while acknowledging the complex and challenging work of language revitalization.

Indigenous language revitalization and Indigenous immersion programs

Planning for Indigenous language revitalization

In order to see their languages thrive into future generations, Indigenous peoples of North America are implementing a diverse array of language revitalization efforts. Key to finding the approaches that work best for each community and language is the process of Indigenous language revitalization planning. Any and all community members and stakeholders can contribute to the process of planning for language revitalization ([Hinton, 2013](#); [McCarty, 2018](#)) and through collective input and decision-making Indigenous communities increase their potential for success, as they define it ([Gessner and Parker, 2020](#)). Because it is inextricably intertwined with socio-political aspects of life in Indigenous communities ([McCarty, 2018](#)), and because it involves multiple societal domains and constituent groups within them ([Gessner and Parker, 2020](#)), planning for effective Indigenous language revitalization must be holistic and comprehensive and must go beyond the short term to encompass a "Life-span approach" ([McIvor and Anisman, 2018](#), p. 101).

[McCarty \(2018\)](#) shares examples of the variety of ways that revitalization planning can manifest. Ranging from one person choosing to revive a sleeping language (not spoken for many years) from textual resources (see [Greendeer and Weston, 2021](#)), to large regional movements for cultural and linguistic revival (such as in Hawai'i), language revitalization planning takes many forms. Recognizing the complexities involved in such planning, and summarizing important lessons from practitioner-language planners, McCarty shares seven principles to aid in community-based language planning (CBLP):

1. "cannot be done to one or for one by others" ([Dauenhauer and Dauenhauer 1998](#), p. 97);
2. There are many CBLP pathways and no single vision or formula for "success;"
3. Language planning is community planning. CBLP is holistic and systemic;
4. CBLP requires a long view of the revitalization process;
5. CBLP requires ideological clarification and commitment;
6. "There is no language for which nothing at all can be done" ([Fishman 1991](#), p. 12);
- and 7. "Language issues [are] always people issues" ([Warner 1999](#), p. 89). ([McCarty, 2018](#), p. 30)

[Hinton \(2013\)](#) further provides nine stages for language planning for revitalization: (1) The Introductory Stage; (2) Goal Setting; (3) Preplanning and Research; (4) Needs Assessment; (5) Policy Formulation; (6) Goal Reassessment, and Developing Strategies and Methods to Reach Your Goals; (7) Implementation; (8) Evaluation; and (9) Replanning.

Indigenous immersion education

Recognizing that previous language education methods, including other forms of bilingual education, were not leading to the outcomes that they desired for their children ([Hinton and Meek, 2018](#)), one important approach to revitalization that Indigenous communities have turned to has been immersion education. Implemented in a multitude of ways, Indigenous immersion education is commonly understood as schooling in which the majority of instruction is carried out in the Indigenous language, often rooted in Indigenous ways of knowing and being ([Hermes and Kawai'ae'a, 2014](#); [Wilson and Kamanā, 2011](#)). Although they may have some similarities to world language immersion programs, Indigenous immersion programs are typically founded with the clear objectives of revitalizing Indigenous language, culture and lifeways. Representing Indigenous forms of resistance and refusal of hegemonic and assimilative US and Canadian systems of education, by "repurposing classrooms and schools," the work of such programs "is a de/anticolonial and restorative practice" ([McCarty et al., 2021](#), p. 350).

While other forms of Indigenous bilingual schooling emerged earlier, such as maintenance and transitional programs that were geared toward Indigenous learners who arrived to school with their Indigenous language as a first language, Indigenous immersion programs for learners beginning schooling with a dominant language (such as English or French) as their first language did not start in the US and Canada until the 1970s and 1980s. Early programs included a Mohawk immersion school in Kahnawà:ke ([McIvor](#)

and McCarty, 2016), and in 1984 the founding of the first Pūnana Leo language nest on the Hawaiian island of Kauaʻi (Wilson and Kamanā, 2011). Since that time, Indigenous immersion schooling has been initiated by numerous communities, with new programs founded each year across a wide range of settings and contexts. While it is common for a language community to have one immersion school, some languages such as Hawaiian and Ojibwe are now offered in multiple immersion schools. Often looked at as a leader in the Indigenous immersion movement, Hawaiʻi currently supports thirteen language nests (preschools) and twenty-seven K-12 immersion school programs (McCarty et al., 2021).

Among potential options for immersion education, growing numbers of Indigenous communities choose to implement immersion schooling with their youngest learners. Often drawing from the Māori programs that inspired such an approach, these programs are commonly referred to as language nests. Chambers and Saddleman (2020) note that, “A language nest is an early language learning program for young children from infancy to five years of age” (p. 27). These programs often have a goal of cultivating language in learners in ways that resemble natural forms of language acquisition, as closely as possible, often intergenerationally (McIvor and Parker, 2016). In an interview-based study focused on nests for languages such as Alutiiq, Hawaiian, Lakota, Mohawk, Ojibwe, SENĆOŦEN, and Tahltan, Okura (2017) examines topics such as what is necessary to start and maintain such programs, what resources they need, the challenges they face, advice that existing programs give, and the logistical and other differences and similarities among programs. Along with seeing positive language outcomes as children spent more time participating in the nests, Okura also notes family involvement, support and commitment as a “primary factor affecting success” in such programs (p. 156).

In line with McIvor and Parker (2016) she notes that there is great diversity among nests, some starting as grassroots initiatives, like those resembling “granny’s house,” while others might start with more formal institutional support – and that there is no one right way to create a nest.

Anchored in their experiences with a language nest in n̓kmapl̓qs in the Lake Okanagan Indian Band community in Vernon, British Columbia, Chambers and Saddleman (2020) document the history and trajectory of ns̓ylxc̓n language nests. Reflecting on their journey since 2007, they describe how the n̓kmapl̓qs program is “modeled on the extended family and intergenerational sharing of language and culture,” noting the positive effects on relationships between children and elders from the nest that “looks and feels like a regular home” (p. 37). They explain how the nest’s structure has evolved to incorporate split days - serving infants and toddlers early in the day, and preschool and kindergarten children later in the day. While the well-known example of the Pūnana Leo programs of Hawaiʻi continue to offer invaluable insights (see Wilson and Kamanā, 2011), as the number of language nests continue to grow, so too does scholarship on programs for languages like Keres, Xaad Kil (Haida language) and Tāltān (Moquino and Blum-Martinez, 2023; Young, 2019).

In addition to preschools or nests, a wide range of Indigenous communities in the US and Canada have founded K-6, and to a lesser extent, grade 7–12 immersion schools. In addition to some already mentioned, K-6 immersion schools for Cherokee, Cree, Salish, SENĆOŦEN, Blackfeet, Diné, Wōpanāak, Ojibwe, and a multitude of other languages now exist. As they nurture children beyond earliest years of development, these schools take a variety of approaches to curriculum and pedagogy, while continuing with the Indigenous language for all or a majority of instruction, and usually refraining from introduction of English through grade 3 or later (Hermes and Kawaiʻaeʻa, 2014). White (2015) provides a detailed account of one of the longest running Indigenous immersion programs, Akwesasne Freedom School (AFS). Founded in 1979, the curriculum and pedagogy of AFS is grounded in the Kanienʻkéha (Mohawk) language and traditional teachings, and arose out of a community movement for self-determination. Exercising significant control of its program, another Kanienʻkéha school that has enjoyed success is the Karonhianónhna Tsi Ionterihwaienstáhkwa of the Kahnawà:ke community in Quebec (Gomashie, 2019).

On the Navajo Nation the longtime immersion program Tséhootsooʻ Diné BíʻÓlta (TDB) was founded in 1986 to address the decline in speakers in Diné communities. Operating from grades K-6, and rooted in Diné philosophy of education, the school maintains high levels of Diné language instruction, consistently showing that students in the program perform academically as well or better than students in nearby English medium instruction programs (Arviso and Holm, 2013; Johnson and Legatz, 2006). Guided by the Diné Philosophy of Education, TDB operates according to the notion that all Diné relatives (and by that all staff at the school) are responsible for children’s Diné language development, which is facilitated by having a school space of their own (Johnson and Legatz, 2006).

Further emphasizing the need for these unique programs to have their own space, Blackfeet language immersion innovator Darrel Kipp warned, “that immersion schools demand a strong protective form of sanctuary” (Kipp, 2002, p. 4). Along with others, Kipp founded the private, non-profit Piegan Institute in 1987. The institute created the Nizi Puh Wah Sin Blackfoot Language Schools Immersion Program, which now encompasses three immersion schools for learners from grades K-8 (The Piegan Institute, 2022). The schools emphasize collectivism and cooperation over grades and competition in the language and identity development of their students (Kipp, 2002).

Seeking to establish immersion schooling that aligns with their own values, some communities have turned to educational approaches such as Montessori and Waldorf education in creating their K-6 schools. Two schools that have followed this approach have been the Weetumuw Katnuhtóhtákamuq Wōpanāak language immersion school and the Keres Children’s Learning Center of Cochiti Pueblo. Immersing their learners in Wōpanāak language and culture through decolonial learning that draws on community members’ knowledge and lessons learned at home in a caring, healthy environment that promotes self-efficacy and direction, the Weetumuw Katnuhtóhtákamuq school was founded in 2016 using Montessori pedagogy and now extends from pre-K through grade 4 (Greendeer and Weston, 2021; Wōpanāak Language Reclamation Project, n.d.). Similarly, the Keres Children’s Learning

Center (KCLC) of Cochiti Pueblo was founded in 2012. Recognizing parallels between Montessori approaches and Cochiti philosophy of holistic child development according to children's readiness, as well as goals of preparing children for culturally and spiritually grounded life guided by responsibility, respect and familial and community contribution, KCLC now hosts Keres Primary (ages 3–6 years) and Elementary Dual Language (ages 6–12) classes (Moquino and Blum-Martinez, 2023; Keres Children's Learning Center, 2022). KCLC has also led to the founding of the Indigenous Montessori Institute, which provides trainings in Indigenous philosophy of education and Montessori methods (Keres Children's Learning Center, 2022).

While the number of Indigenous immersion programs in preschool to grade 6 have increased significantly in the past two decades, both in the US and Canada, there continue to be few programs for grades 7–12. One program in British Columbia that reports programming through high school is the Chief Atahm School. Beginning as an early childhood program in the late 1980s the program serves Secwepemc members of Adams Lake Band, Neskonlith and Little Shuswap, and has now grown to offer schooling through grade 10 (Chief Atahm School, 2017; Michel, 2012). Another example is found in the highly successful program on the Lac Courte Oreilles reservation, Waadookodaading Ojibwe Immersion School. Founded in 2001, the school now serves children from grades K-8, with plans to expand into high school (Koran, 2021; Waadookodaading Ojibwe Immersion School, n.d.). Implementing culturally rich ways of learning driven by the natural cycles of their local environment, extending into secondary levels allows students at the school to build on the already fluid Ojibwe language they develop by elementary grades (Hermes and Haskins, 2018; Sullivan, 2019).

Perhaps the most well-known of US secondary Indigenous immersion programs are found in Hawai'i. Emerging with a pilot program in 1987 the Ke Kula Kaiapuni Hawai'i, or Hawaiian Immersion schools, provide instruction through the medium of Hawaiian in all subjects, usually introducing English language as a subject in grade 5. These programs now exist in a network across the islands of O'ahu, Maui, Hawai'i, and Moloka'i, with ten of the schools serving students at the grade 7–8 level and thirteen schools serving students in grades 9–12 (Hawai'i Department of Education; n.d.; Warner, 1999). One school on the island of Hawai'i.

Ke Kula 'O Nāwahīokalani'ōpu'u (Nāwahī), provides insights into the amazing potential of Indigenous immersion education. Offering K-12 programming, the school graduates high school students with a high sense of Native Hawaiian identity, who often return to Nāwahī or other immersion programs to contribute to Hawaiian language revitalization. At the same time the school has maintained a 100% graduation rate since 1999, and students graduate highly proficient in Hawaiian language (Wilson, 2021; Wilson and Kamana, 2011).

Hawai'i also provides an example for postsecondary immersion programs. College level Hawaiian language programming initially gained traction in 1982 as a four-year Bachelor of Arts program within the University of Hawai'i in response to calls from youth and elders. Later, along with programs at other sites, Ka Haka 'Ula O Ke'elikōlani College of Hawaiian Language was founded in Hilo, HI. Built upon the Kumu Honua Mauli Ola Hawaiian Philosophy of education, through the B.A. in Hawaiian Studies students complete rigorous Hawaiian language courses so that they can take all coursework through Hawaiian beginning in their second year (Wilson, 2020). Programs at the college have now extended to master's and doctoral levels, all of which can be completed in Hawaiian.

Elsewhere tribal colleges and other programs are creating unique college level programming. On the Navajo Nation, Diné College offers a bachelor's degree in Diné studies that is grounded in immersion and "aims to produce graduates highly proficient in the spoken Diné language, cultural teachings, and ways-of-thought, who embody these ways-of-being" (Diné College, 2020; McKenzie, 2020). Using another approach, with a goal of graduating teachers certified to contribute to language education and revitalization in their communities, the University of Victoria in British Columbia offers a sequence of programs that lead to a Bachelor of Education in Indigenous Language Revitalization, with extensive work in students' own languages. Another program in Canada, Kwi Awt Stelmexw, illustrates how community-based non-profit organizations can partner with postsecondary institutions to offer immersion programming that leads to college credit. They partner with Simon Fraser University to offer the two-year Squamish Language Proficiency Diploma through a full-time program of Skwxwu7mesh immersion learning (Chew et al., 2021).

Structural and contextual differences

Indigenous immersion schools in the US and Canada have successfully carried out their efforts within a number of structures and contexts. Contributing to the diverse ways in which these programs arise, one important characteristic that affects the schools significantly is their relationship to broader educational and governmental systems. Some immersion schools carve out space to operate under prescriptive regulation of state or federal school systems in the US Others, such as Nāwahī and Waadookodaading have operated successfully as charter schools that exercise increased control over curricular, pedagogical and other aspects. Still others, such as the Blackfeet schools founded by the Piegan Institute, and Akwesasne Freedom School, have chosen to operate as private, non-profit schools, independent of state and federal systems. In some cases, participation in broader systems can provide advantages like more stability, funding and facilities. On the other hand, participation as part of government regulated educational systems also comes with accountability to those systems, which often prioritize high stakes testing, and curriculum and pedagogy that do not reflect the learning sought by Indigenous communities, especially as they privilege English language (Hermes and Dyke, 2017; Kipp, 2002; White, 2015).

Related to their affiliation with larger educational systems, one prevalent difference in the ways that Indigenous immersion programs in schools can emerge is as either strand or stand-alone programs. While some schools such as the Blackfeet schools and KCLC have noted the importance of having a learning space dedicated solely to the Indigenous language (Kipp, 2002; Moquino and Blum-Martinez, 2023), it is sometimes more feasible for programs to begin and/or continue as a strand, or immersion track,

within an existing English medium school. This has been the case for a significant number of programs, including some Hawaiian medium programs and the Diné school Tséhootsooí Diné Bi'ólta' (May, 2013; Holm and Holm, 1995).

Another significant aspect in the diversity of schools is the places in which they are emerging. The majority of Indigenous immersion schools are found on the reserves and reservations of the peoples for whose languages they are founded. Yet, given an increase in the number of North American Indigenous peoples in urban areas (McIvor and McCarty, 2016), options for Indigenous immersion education are beginning to arise in cities that now occupy the ancestral lands of Indigenous peoples. One such school is Bdote Learning Center, a grade K-8 program named after the confluence of the Mississippi and Missouri rivers in the homelands of the Dakota people and present-day Minneapolis, MN. Although located in a large metropolitan area, the school, which has immersion pathways for both Dakota and Ojibwe languages, emphasizes interdisciplinary learning beyond the classroom, "tied to the geographical surroundings of the school, the local history, landscape, ecology, and Native languages and cultures" (Bdote Learning Center, 2021). Another urban-based program is Puente de Hózhq (PDH) in Flagstaff, AZ (Fillerup, 2011). A trilingual magnet school founded in a city with significant Native American and Latino populations, PDH offers both one-way Diné and two-way Spanish language immersion programs. With emphasis on academic rigor, and bilingualism, the school's infusion of culture, and presence of three languages, promotes their goal of multicultural competence. At the same time students regularly perform better academically than students who attend English medium schools in the area (Fillerup, 2011; McCarty and Lee, 2015).

Non-school-based ILR programs

While school-based programs continue to be common among ILR efforts in the US and Canada, a growing number of approaches outside of schools are also being implemented by Indigenous communities to support revitalization. Recognizing the need to increase the number of adult speakers who can foster language revitalization with all ages, a number of Indigenous communities have begun adult language immersion programs. Multiple programs of this kind have been developed for Kanien'kéha (Mohawk language). Programs have been established in the communities of Kahnawake Territory, Tyendinaga Territory, Wahta Territory, Six Nations Territory and Kanatsiowhareke Territory (Maracle, 2002). These communities reach across both the US and Canada. The community-based Onkwawenna Kentyohkwa (Our Language Society) program, for example, operating as full-time adult immersion program since 1999, has seen 140 students complete its program, many of whom have become teachers of the language, themselves (Gomashie, 2019; Onkwawenna Kentyohkwa, n.d.). Another program, Ojibwemotaadidaa Omaa Gidakiiminaang (OOG), offers immersion language learning through winter weekend and multi-week summer immersion academies aimed at helping learners build advanced language skills in Ojibwe. Initiated by Fond du Lac Tribal and Community College, and other Ojibwe revitalizers, OOG incorporates a variety of first language speakers and focuses on diverse contexts to prepare learners to eventually live everyday life in the language and contribute to language revitalization themselves (Erdrich, 2013; Ojibwemotaadidaa Omaa Gidakiiminaang, 2020).

Another approach to ILR that has become prevalent in recent decades is the Master-Apprentice Language Learning Program (or Mentor-Apprentice Language Learning Program) (MAP). Within these programs fluent speakers (often elders) are paired with adults who are committed to learning the language in one-on-one or small group settings. Based on immersive language interaction between the mentor(s) and apprentice(s) in the Indigenous language over extended periods of time, such programs promote language use and development on a regular basis and in a wide variety of life situations. In these programs the goal is often not just for the apprentice to learn, but for them to be able to pass the language on to others, as well (Hinton et al., 2018). Although Mentor-Apprentice Programs may be particularly helpful in situations in which there are very few remaining speakers of a language, or in which immersion schooling is not immediately viable or desired, they are not limited to such contexts, and can be implemented in a variety of situations and settings.

Since the 1990s Advocates for Indigenous California Language Survival (AICLS), the group who developed the MAP program, has trained more 100 Mentor-Apprentice teams in California alone, and the First Peoples' Cultural Council (FPCC) in British Columbia has supported 97 apprentices in 28 different languages in British Columbia (Hinton et al., 2018). Jenni et al. (2017) find that the programs not only cultivate language, but that they also have positive effects on cultural and spiritual health, as well as participants' engagement in their communities. In addition to experiencing success with many California languages the programs have been implemented for languages as far and wide as Chickasaw, Cherokee, Dakota, Euchee, Karuk, Ojibwe, Potawatomi, Sauk, Tlingit, Wompanaak, and Yakama, among others (Chew, 2015; Chew et al., 2021; Gordon and Huggins, 2020; Hinton et al., 2018; Morrison, 2020; Twitchell, 2016).

Other innovative approaches being created by language revitalizers place increased focus on the immersion environment. Based on her work with the Nsyilxcn language in British Columbia, Johnson (2013, 2017) describes a language house approach in which a group of committed language learners and facilitators come together to live in a house for five months, experiencing and supporting each other through significantly more language learning and use than conventional classes can offer. Supported by elders and focusing on curriculum development and assessment methods, along with the personal journeys of participants, Johnson demonstrates how such an intensive effort can lead to multiple resources for revitalization extending beyond individual participants and into the community. Focusing on another setting, some ILR efforts are drawing on the power of Indigenous connections to lands to facilitate language learning. Land-based approaches to immersion illustrate the importance of learning experiences on traditional homelands in creating the conditions through which learners can experience the relational and spiritual connections in which Indigenous languages are grounded (McKenzie, 2020; Rorick, 2018). One initiative has even turned land-based immersion language efforts into a comprehensive lifestyle. Reclaiming traditional homelands to found an off-grid community, the Ekvñ-Yefolecv

Maskoke ecovillage surrounds learners in the Maskoke language through prioritizing focus on traditional ecological and agricultural knowledge (Ekvn-Yefolecv, 2021).

While the number and diversity of Indigenous immersion programs continues to grow in the US and Canada, there exists a growing need to understand the similarities and unique aspects of these programs while also building a better understanding of how they affect learners, families and communities. To that end a multi-year multi-site, mixed methods study is currently underway to document and shed light on a diverse array of such schools. Sponsored by the Spencer Foundation, the Indigenous-Language Immersion and Native American Student Achievement Study will draw from quantitative work, as well as rich qualitative analysis to highlight both key commonalities and unique features of such programs (McCarty et al., 2021).

Challenges and supports to Indigenous language revitalization

One of the most prevalent challenges facing Indigenous language revitalization efforts is the declining numbers of first language speakers. Even languages with the largest number of speakers are seeing very few young people raised with the ability to speak their Indigenous languages. This situation is exacerbated by the ongoing assimilative environment fostered by hegemonic policies that prioritize English language and Euro-western knowledge as the basis of education in the US and Canada (Hermes and Dyke, 2017). Without adequate financial and other forms of support for Indigenous language medium education, Indigenous communities find themselves constantly lacking facilities, curricula, and training for teachers and program leaders, among other crucial resources (White, 2015). This often leads to situations in which insufficient time will be devoted to learning through the Indigenous languages (much less multi-year, full-immersion programs). Relatedly, those school-based programs that do exist have difficulty extending beyond grade 3 or 6, limiting the amount of language development that would be attainable if programs could run through grade 8 or even grade 12. Further, ideologically, in the Euro-western capitalist societies of the US and Canada, lack of support from government and educational leaders likely fuels what some have recognized as false perceptions that immersion education through a child's Indigenous language will be detrimental to their success in academic pursuits, and life in general (Kipp, 2002; Sumida Huaman, 2020).

Despite the challenges facing ILR initiatives, Indigenous communities are finding ways to contribute to revitalization of their languages. The rapid growth in programs that are being initiated and the ongoing success of programs across the continent for more than two decades, such as those described in this article, illustrate the unwavering resiliency and increasing momentum for revitalization efforts. Indigenous language revitalizers are working to build and sustain networks, both within and outside of institutional spaces, in grassroots community initiatives, and as individuals, to share resources through conferences, trainings and other forms of knowledge sharing and encouragement. ILR teacher training programs and scholarship around successful programs and approaches to ILR are increasing. Practitioners and communities are innovating with apps, programs, social media and other forms of technological resources that extend and expand learning opportunities beyond anything seen before. Perhaps most significantly, growing numbers of Indigenous peoples, from all walks of life, elders, youth, educators, researchers, speakers, learners, are all building awareness of and joining in Indigenous language revitalization initiatives for their languages.

Conclusion

Although Indigenous peoples in the US and Canada have endured assimilative oppression of their languages through settler states and their educational systems, ILR movements represent Indigenous communities' unwavering commitment to their sovereign abilities to maintain their languages and identities. Indigenous peoples are demonstrating their abilities to reclaim and nurture their languages and identities for generations to come through a diverse range of ILR approaches. On one hand these approaches are characterized by return to traditional conceptions of the purposes, meanings, uses and settings for Indigenous languages that differ significantly from predominant notions of education in mainstream society. In other instances, ILR approaches are embracing ways to innovate using contemporary systems and tools, repurposing them for the decolonial and re-Indigenizing goals of Indigenous language revitalization as defined by Indigenous communities, themselves.

With this in mind, the ILR movement can benefit from further research to address both its ongoing and recent successes, as well as current obstacles and challenges facing ILR approaches across North America. For example, the initiatives of previously discussed land based language immersion to influence comprehensive lifestyle change are promising and exciting. Learning about the holistic impact of these types of initiatives is significant for other communities to learn from and emulate. Similarly, further research on uniquely Indigenous immersion pedagogies, as well as the role(s) that Indigenous educational philosophies play in ILR will contribute to greater understanding of the holistic connections in the ILR movement. Conversely, further research on the continuing imposition of colonial structures on ILR, such as through state policies and funding that privilege English, is necessary to improve the ILR movement's understanding of how to intervene and confront mechanisms that continue to constrain and undermine ILR initiatives. Finally, further research on the challenges from within our own Indigenous communities is vital for ensuring community investment and the most potential success in the ILR movement. Addressing these challenges, including the internalized oppression of our minds toward our languages and lifeways, require what Smith calls for in "conscientization" (2003). Further research in these areas (and more) can contribute to the ongoing development and increasingly positive outcomes of ILR programming, and ultimately to the momentum of the ILR movement.

Across North America Indigenous communities are intentionally mobilizing in strategic and creative ways to re-conceptualize education for all ages of community members according to Indigenous aspirations for life, with Indigenous languages at the center. While similarities exist among the growing number of ILR initiatives, communities are increasingly focused on identifying the approach(es) that best fit their context and situation to promote well-being for their people, now and in the future.

References

- Arviso, M., Holm, W., 2013. Tséhootsooíi ólta'gí diné bizaad bihoo'aah A Navajo immersion program at Fort Defiance, Arizona. In: Hinton, L., Hale, K. (Eds.), *The Green Book of Language Revitalization in Practice*. Brill Publishers, Leiden, Netherlands, pp. 203–216.
- Bdote Learning Center, 2021. <https://bdote.org/>.
- Benally, A., Viri, D., 2005. Diné bizaad [Navajo language] at a crossroads: extinction or renewal? *Biling. Res. J.* 29 (1), 85–108. <https://doi.org/10.1080/15235882.2005.10162825>.
- Chambers, N.A., Saddleman, D., 2020. Moving towards a language nest: stories and insights from nKmaplqs. *First Peoples Child Fam. Rev.* 15 (1), 27. <https://doi.org/10.7202/1068361ar>.
- Chew, K.A.B., 2015. Family at the heart of Chickasaw language reclamation. *Am. Indian Q.* 39 (2), 154. <https://doi.org/10.5250/amerindiquar.39.2.0154>.
- Chew, K.A.B., Manatowa-Bailey, J., Lukaniec, M., McIvor, O., Linn, M., 2021. Growing the Fire within: Exploring Innovative and Successful Adult Language Learning Methods in Indigenous communities in Canada and the US. <https://shareok.org/handle/11244/331230>.
- Chief Atahm School. (2017). <http://www.chiefatahm.com/>.
- Chiocca, E.S., 2020. Language endangerment: diversity and specificities of Native American languages of Oklahoma. In: Brunn, S.D., Kehrein, R. (Eds.), *Handbook of the Changing World Language Map*. Springer International Publishing, pp. 1607–1626. https://doi.org/10.1007/978-3-030-02438-3_19.
- Diné College, 2020. <https://www.dinecollege.edu/>.
- Eknv-Yefolecv, 2021. <https://www.eknv-yefolecv.org/>.
- Erdrich, P., 2013. Ojibwemotaadidaa: preparing a new generation of fluent speakers. *Trib. Coll.* 24 (4), 18–21.
- Fillerup, M., 2011. Building a 'bridge of beauty': a preliminary report on promising practices in Native language and culture teaching at Puente de Hózhq Trilingual Magnet School. In: Romero-Little, M.E., Ortiz, S.J., McCarty, T.L., Chen, R. (Eds.), *Indigenous Languages Across the Generations—Strengthening Families and Communities*. Arizona State University Center for Indian Education, Tempe, pp. 145–164.
- Fishman, J., 1991. Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages, 76. *Multilingual Matters, LTD*, Clevedon, UK.
- Fitzgerald, C.M., 2017. Understanding language vitality and reclamation as resilience: a framework for language endangerment and "loss" (Commentary on Mufwene). *Language* 93 (4), 280–297. <https://doi.org/10.1353/lan.2017.0072>.
- Gessner, S., Parker, A., 2020. Language planning: planning a future for your language. In: Crowshoe, L., Genée, I., Peddle, M., Smith, J., Snoek, C. (Eds.), *Sustaining Indigenous Languages Connecting Communities, Teachers, and Scholars*. Northern Arizona University, Flagstaff, AZ.
- Gomashie, G.A., 2019. Kanien'keha/Mohawk Indigenous language revitalisation efforts in Canada. *McGill J. Educ.* 54 (1), 1060864ar. <https://doi.org/10.7202/1060864ar>.
- Gordon, H.S.J.G., Huggins, B., November 2020. Master/mentor-apprentice language learning model for Native language teacher training. Presented at. In: *The 2020 Administration for Native Americans Native Language Summit*. Online. https://www.anawestern.org/wp-content/uploads/2020/11/Master-Apprentice-Language-Learning-Model_Day2-530.pdf.
- Greendeer, N.H., Weston, J., 2021. Mother tongues and community well-being: survivance, decolonization, and the Wópanāak language reclamation project. In: Chisolm-Straker, M., Chon, K. (Eds.), *The Historical Roots of Human Trafficking*. Springer International Publishing, pp. 291–303. https://doi.org/10.1007/978-3-030-70675-3_16.
- Hallett, D., Chandler, M.J., Lalonde, C.E., 2007. Aboriginal language knowledge and youth suicide. *Cognit. Dev.* 22 (3), 392–399. <https://doi.org/10.1016/j.cogdev.2007.02.001>.
- Hawai'i Department of Education (n.d.). <https://www.hawaiipublicschools.org/TeachingAndLearning/StudentLearning/HawaiianEducation/Pages/Hawaiian-language-immersion-schools.aspx>.
- Hermes, M., Dyke, E., 2017. Still flourishing: enacting indigenizing language immersion pedagogies in the era of US common core state standards. In: McKinley, E.A., Smith, L.T. (Eds.), *Handbook of Indigenous Education*. Springer Singapore, pp. 1–23. https://doi.org/10.1007/978-981-10-1839-8_21-1.
- Hermes, M., Haskins, M., 2018. Unbecoming standards through Ojibwe immersion: the wolf meets Ma'íngan. In: Wigglesworth, G., Simpson, J., Vaughan, J. (Eds.), *Language Practices of Indigenous Children and Youth*. Palgrave Macmillan UK, pp. 99–116. https://doi.org/10.1057/978-1-137-60120-9_5.
- Hermes, M., Kawai'ae'a, K., 2014. Revitalizing Indigenous languages through Indigenous immersion education. *J. Immers. Content Based Lang. Educ.* 2 (2), 303–322. <https://doi.org/10.1075/jicb.2.2.10her>.
- Hinton, L., 2013. Language planning. In: Hinton, L., Hale, K. (Eds.), *The Green Book of Language Revitalization in Practice*. Brill Publishers, Leiden, Netherlands, pp. 51–59.
- Hinton, L., 2016. Language endangerment and revitalization. In: McCarty, T.L., May, S. (Eds.), *Language Policy and Political Issues in Education*. Springer International Publishing, Cham, Switzerland, pp. 1–16. https://doi.org/10.1007/978-3-319-02320-5_21-1.
- Hinton, L., Hale, K.L. (Eds.), 2013. *The Green Book of Language Revitalization in Practice*. Brill Publishers, Leiden, Netherlands.
- Hinton, L., Meek, B., 2018. Revitalization in Indigenous North America. In: Hinton, L., Huss, L., Roche, G., Huss, L., Roche, G. (Eds.), *The Routledge Handbook of Language Revitalization*. Routledge, Milton Park, Oxfordshire. <https://doi.org/10.4324/9781315561271>.
- Hinton, L., Florey, M., Gessner, S., Manatowa-Bailey, J., 2018. The master-apprentice language learning program. In: Hinton, L., Huss, L., Roche, G., Huss, L., Roche, G. (Eds.), *The Routledge Handbook of Language Revitalization*. Routledge, Milton Park, Oxfordshire. <https://doi.org/10.4324/9781315561271>.
- Holm, A., Holm, W., 1995. Navajo language education: retrospect and prospects. *Biling. Res. J.* 19 (1), 141–167. <https://doi.org/10.1080/15235882.1995.10668595>.
- Jenni, B., Anisman, A., McIvor, O., Jacobs, P., 2017. An exploration of the effects of mentor-apprentice programs on mentors and apprentices' wellbeing. *Int. J. Indigenous Health* 12 (2), 25–42. <https://doi.org/10.18357/ijih122201717783>.
- Johnson, M.K., 2013. n'áqwcín (Clear Speech): 1,000 Hours to Mid-intermediate N'syilxcn Proficiency (Indigenous Language, Syilx, Okanagan-Colville, N'qilxcn, Interior Salish) (Unpublished Doctoral Dissertation). University of British Columbia, Vancouver, BC. <https://doi.org/10.14288/1.0074314>.
- Johnson, M.K., 2017. Syilx language house: how and why we are delivering 2,000 decolonizing hours in N'syilxcn. *Can. Mod. Lang. Rev.* 73 (4), 509–537. <https://doi.org/10.3138/cmlr.4040>.
- Johnson, T., Legatz, J., 2006. Tséhootsooíi Diné bi'ólta' [The Navajo school at the meadow between the rocks]. *J. Am. Indian Educ.* 46 (2), 26–33.
- Keres Children's Learning Center (2022). <https://kclcmontessori.org/>.
- Kipp, D., 2002. Encouragement, guidance and lessons learned: 21 Years in the trenches of Indigenous language revitalization. In: Reyhner, J., Lockard, L. (Eds.), *Indigenous Language Revitalization: Encouragement, Guidance and Lessons Learned*. Northern Arizona University, Flagstaff.
- Koran, P., 2021. Race against the clock: the school fighting to save the Ojibwe language before its elders pass away. *Guardian* 16, 7 April.
- Krauss, M., 1996. Status of Native American language endangerment. In: Cantoni, G. (Ed.), *Stabilizing Indigenous Languages*. NAU Center for Excellence in Education, Flagstaff, AZ, pp. 16–21.
- Lee, T.S., 2009. Language, identity, and power: Navajo and Pueblo young adults' perspectives and experiences with competing language ideologies. *J. Lang. Ident. Educ.* 8 (5), 307–320. <http://www.tandfonline.com/doi/abs/10.1080/15348450903305106>.
- Lee, T., McLaughlin, D., 2001. Reversing Navajo language shift, revisited. In: Fishman, J.A. (Ed.), *Can Threatened Languages Be Saved? Multilingual Matters*. Bristol, U.K., pp. 23–43. <https://doi.org/10.21832/9781853597060-004>.
- Leonard, W., 2012. Framing language reclamation programmes for everybody's empowerment. *Gend. Lang.* 6, 339–367.
- Leonard, W., 2017. Producing language reclamation by decolonising 'language.'. *Lang. Doc. Des.* 14, 15–36.

- Lomawaima, K.T., McCarty, T.L., 2006. "To Remain an Indian": Lessons in Democracy from a Century of Native American Education. Teachers College Press, New York, NY.
- Maracle, B.J., 2002. Adult Mohawk language immersion programming. *McGill J. Educ.* 37 (3), 387.
- Martinez, R.B., 2000. Languages and tribal sovereignty: whose language is it anyway? *Theor. Pract.* 39 (4), 211–219.
- May, S., 2013. Indigenous immersion education: international developments. *J. Immers. Content Based Lang. Educ.* 1 (1), 34–69. <https://doi.org/10.1075/jicb.1.1.03may>.
- McCarty, T.L., 2018. Community-based language planning: perspectives from Indigenous language revitalization. In: Hinton, L., Huss, L., Roche, G. (Eds.), *The Routledge Handbook of Language Revitalization*. Routledge, pp. 22–35.
- McCarty, T.L., 2020. The holistic benefits of education for Indigenous language revitalisation and reclamation (ELR2). *J. Multiling. Multicult. Dev.* 0 (0), 1–14. <https://doi.org/10.1080/01434632.2020.1827647>.
- McCarty, T.L., Lee, T.S., 2015. The role of schools in Native American language and culture revitalization: a vision of linguistic and educational sovereignty. In: Jacob, W.J., Cheng, S.Y., Porter, M.K. (Eds.), *Indigenous Education: Language, Culture and Identity*. Springer, New York, NY, pp. 341–360. https://doi.org/10.1007/978-94-017-9355-1_17.
- McCarty, T.L., Nicholas, S.E., 2014. Reclaiming Indigenous languages: a reconsideration of the roles and responsibilities of schools. *Rev. Res. Educ.* 38 (1), 106–136. <https://doi.org/10.3102/0091732X13507894>.
- McCarty, T.L., Noguera, J., Lee, T.S., Nicholas, S.E., 2021. "A viable path for education"—Indigenous-language immersion and sustainable self-determination. *J. Lang. Ident. Educ.* 20 (5), 340–354. <https://doi.org/10.1080/15348458.2021.1957681>.
- McIvor, O., 2009. Strategies for Indigenous language revitalization and maintenance. In: *Encyclopedia of Language and Literacy Development*. Canadian Language and Literacy Research Network, Ottawa, Ontario, pp. 1–12.
- McIvor, O., Anisman, A., 2018. Keeping our languages alive: strategies for Indigenous language revitalization and maintenance. In: Watanabe, Y. (Ed.), *Handbook of Cultural Security*. Edward Elgar Publishing, Northampton, MA, pp. 90–109. <https://doi.org/10.4337/9781786437747>.
- McIvor, O., McCarty, T.L., 2016. Indigenous bilingual and revitalization- immersion education in Canada and the USA. In: García, O., Lin, A., May, S. (Eds.), *Bilingual and Multilingual Education*. *Encyclopedia of Language and Education*, third ed. Springer, Switzerland. https://doi.org/10.1007/978-3-319-02258-1_34.
- McIvor, O., Parker, A., 2016. Back to the future: recreating natural Indigenous language learning environments through language nest early childhood immersion programs. *Int. J. Holistic Early Learn. Dev.* 3 (0), 21–35.
- McKenzie, J., 2020. Approaching from many angles: seeing the connections for our languages to live. *Mod. Lang. J.* 104 (2), 501–506. <https://doi.org/10.1111/modl.12655>.
- Michel, K., 2012. Trickster's Path to Language Transformation: Stories of Secwépemc Immersion from Chief Atahm School (Unpublished Doctoral Dissertation). University of British Columbia, Vancouver, BC.
- Moquino, T., Blum Martinez, R., 2023. Keres Children's learning center: the search for a linguistically and culturally appropriate education. In: Peele-Eady, T.B. (Ed.), *Integrating Home and Heritage Languages in Schools: Pedagogy and Promise*. Routledge, New York. Retrieved from: https://static1.squarespace.com/static/5dd21f53f6b713e17aa9d19t/5f06218ebd6cfa7c6b832d8a/1594237327014/Moquino_Martinez+article.+TPE.pdf.
- Morrison, J.A., 2020. *CWY 501h00J T J0R0ET: J0ZPl 010100J JLTV00J0 We Are Cultivating the Cherokee Language: Stories of Language Renewal from the Cherokee Nation's Cherokee Language Master/Apprentice Program* (Unpublished Doctoral Dissertation). Texas Tech University, Lubbock, TX. <https://tu-ir.tdl.org/bitstream/handle/2346/85797/MORRISON-DISSERTATION-2020.pdf?sequence=1&isAllowed=y>.
- Moseley, C. (Ed.), 2010. *Atlas of the World's Languages in Danger: Context and Process*, third ed. UNESCO Publishing, Paris, France.
- Ojibwemotaadidaa Omaa Gidakiminaang. (Producer). (2020). Wadisman Workshop: How to Start an Indigenous language nest. [Online Workshop Series].
- Okura, E.K., 2017. *Language Nests and Language Acquisition: An Empirical Analysis* (Unpublished Doctoral Dissertation). University of Hawai'i at Mānoa, Honolulu, HI. <http://scholarspace.manoa.hawaii.edu/handle/10125/62500>.
- Onkwawenna Kentyohkwa (n.d.). <https://onkwawenna.info/>.
- Reyhner, J., Lockard, L. (Eds.), 2009. *Indigenous Language Revitalization: Encouragement, Guidance, and Lessons Learned*. Northern Arizona University.
- Romero-Little, M.E., McCarty, T.L., Warhol, L., Zepeda, O., 2007. Language policies in practice: preliminary findings from a large-scale national study of Native American language shift. *Tesol Q.* 41 (3), 607–618. <https://doi.org/10.1002/j.1545-7249.2007.tb00093.x>.
- Rorick, C.L., 2018. Wahaŋasuk?i Naananisqakqin: at the home of our ancestors: ancestral continuity in Indigenous land-based language immersion. In: Smith, L.T., Tuck, E., Yang, K.W. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, pp. 224–237.
- Sallabank, J., 2012. Diversity and language policy for endangered languages. In: Spolsky, B. (Ed.), *The Cambridge Handbook of Language Policy*. Cambridge University Press, Cambridge, UK, pp. 100–123. <https://doi.org/10.1017/CBO9780511979026.008>.
- Smith, G., 2003. Indigenous Struggle for the Transformation of Education and Schooling. *Transforming Institutions: Reclaiming Education and Schooling for Indigenous Peoples*. Keynote Address to the Alaskan Federation of Natives Convention, pp. 1–14 (October).
- Smith, L.T., Tuck, E., Yang, K.W. (Eds.), 2018. *Indigenous and Decolonizing Studies in Education*. Routledge, New York, NY.
- Spolsky, B., 2002. Prospects for the survival of the Navajo language: a reconsideration. *Anthropol. Educ. Q.* 33 (2), 139–162. <https://doi.org/10.1525/aeq.2002.33.2.139>.
- Sullivan, M., October 2019. Anooj Ikidowag Abinoojinyag: remarks on child language acquisition and assessment at Waadookodaading. Presented at: In: The 2019 National Indian Education Association Convention and Trade Show, Minneapolis, MN.
- Sumida Huanan, E., 2020. Small indigenous schools: indigenous resurgence and education in the Americas. *Anthropol. Educ. Q.* 51 (3), 262–281. <https://doi.org/10.1111/aeq.12335>.
- The Piegan Institute. (2022). <https://www.pieganinstitute.org/>.
- Tuck, E., 2009. Suspending damage: a letter to communities. *Harv. Educ. Rev.* 79 (3), 409–428. <https://doi.org/10.17763/haer.79.3.n0016675661t3n15>.
- Twitchell, X.L., 2016. P-20: Building a Tlingit Language Nest, K-12 Medium School, and Adult Immersion Program ([Unpublished manuscript]).
- Waadookodaading Ojibwe Immersion School (n.d.). www.waadookodaading.org.
- Warhol, L., 2011. Native American Language Policy in the United States. *Heritage Briefs*. The Center for Applied Linguistics. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.422.7543>. (Accessed 22 January 2022).
- Warner, S.N., 1999. Hawaiian language genesis: planning for intergenerational use of Hawaiian beyond the school. In: Huebner, T., Davis, K.A. (Eds.), *Studies in Bilingualism*, vol. 16. John Benjamins Publishing Company, p. 313. <https://doi.org/10.1075/sibil.16.20war>.
- White, L., 2015. Free to Be Mohawk: Indigenous Education at the Akwesasne Freedom School. University of Oklahoma Press.
- Wilson, W.H., 2020. The honua of the Hawaiian language college. In: Fornasiero, J., Reed, S.M.A., Amery, R., Bouvet, E., Enomoto, K., Xu, H.L. (Eds.), *Intersections in Language Planning and Policy: Establishing Connections in Languages and Cultures*. Springer International Publishing, pp. 421–437. https://doi.org/10.1007/978-3-030-50925-5_26.
- Wilson, W.H., 2021. Hawaiian History Month: The History of Ke Kula 'O Nāwahīkalanī'ōpu'u, UH Hilo's Hawaiian Language Medium Laboratory School – UH Hilo Stories. Retrieved January 23, 2022, from: <https://hilo.hawaii.edu/chancellor/stories/2021/09/21/history-of-nawahi-by-pila-wilson/>.
- Wilson, W.H., Kamanā, K., 2006. For the interest of the Hawaiians themselves: reclaiming the benefits of Hawaiian-medium education. *Hūlili* 3 (1), 153–181.
- Wilson, W.H., Kamanā, K., 2011. Insights from Indigenous language immersion in Hawai'i. In: Tedick, D.J., Christian, D., Fortune, T.W. (Eds.), *Immersion Education*. Multilingual Matters, pp. 36–57. <https://doi.org/10.21832/9781847694041-006>.
- Wong, K.L., 2004. He Hawa'e Kai Nui a Kau ma Kula. *Educ. Perspect.* 37 (1), 31–39.
- Wōpanāak Language Reclamation Project. (n.d.). <https://www.wlrp.org/>.
- Young, B., 2019. K'waa Gyáak'id Gáagaay Hl Kinsálang: Emerging Xaad Kúl (Haida Language) Use in a Preschool Setting (Unpublished Master's Thesis). Simon Fraser University, Burnaby, BC. <http://summit.sfu.ca/item/19663>.

Englishization and the language of education in Nigeria: literacy issues, challenges and hopes

Gabriel B. Egbe, Department of English and Literary Studies, Veritas University, Abuja, Nigeria

© 2023 Elsevier Ltd. All rights reserved.

Introduction	61
Conceptualizing Englishization	61
Englishization as restlessness	61
Englishization as domestication, hybridization and Anglicization	62
Anglicization of names	63
Englishization of indigenous languages and indigenization of English	63
Educational policies and English in Nigeria: some historical backdrop	63
English in the language policy provisions	64
English, language of education and literacy issues	65
Literacy education in Nigeria: prospects and hopes for Nigeria	65
Formulating and implementing a workable language of education and literacy policy	66
Producing and accessing adequate literacy materials	66
Adopting innovative instructional strategies	66
Deploying digital literacy technologies and approaches	66
Conclusion	66
References	67

Introduction

Portuguese was the first European language to be spoken in what is present day Nigeria in the 15th century. This period marked the first European entry into the West African coast. The English language came later at about 1553 when English men visited parts of the Ancient Benin Kingdom and Old Calabar areas of present day Nigeria. In spite of the fact that English came much later after Portuguese, the language has assumed a place of prominence which is not challenged by any other international language nor any of the over 500 indigenous languages spoken in Nigeria. This dominance is understandable following Nigeria's colonial history and the entrenchment of English as an official language even after independence in 1960. The fact that English is used as the medium of instruction and a subject of prestige in the school system is another factor. Not surprisingly, the place of English, especially its role in the educational system, has been a recurring decimal, that is, an issue that keeps occurring (Bamgbose, 2020) in the various educational enactments, policies, pronouncements and planning in relation to language and the language of education from the 19th century to the present. From the advent of English in Nigeria through the era of missionary and colonial activities to independence and post-independence periods (Adetugbo, 1979; Awonusi, 2004), English has been so entrenched that it is accepted as the language of integration and social cohesion within the Nigerian linguistic complexity. For some young Nigerians, especially those born and raised in some Nigerian urban centers, English, and Pidgin English in some cases, is not just a second language, but indeed, a first language. It is a home language and for some it is the only medium of communication and the preferred language of choice for this class of Nigerians (Egbe, 2014). This group of young Nigerians might choose to learn a second or an additional language for instrumental reasons.

The place and role of English in Nigeria's educational system vis-a-vis literacy learning is so central that when concerns are raised regarding the quality of education, people rather point at the performance of students in both their spoken and written English. There is therefore the tendency to measure educational standards in terms of how students perform in English both as a medium of instruction and as a school subject. The linguistic plurality of Nigeria appears a fertile ground to harvest the linguistic resources of English in Nigeria. The Englishization process is not challenged by any other foreign language in Nigeria nor any of the over 500 indigenous languages.

The focus of this paper is to examine the dominance of English in Nigeria conceived as Englishization in relation to the language of education in Nigeria, explore the concept of Englishization and identify the literacy issues and challenges in Nigeria given the dominance of English as well as highlight the prospects for Nigeria as a member of the global community contributing and shaping world Englishes.

Conceptualizing Englishization

Englishization as restlessness

Historically, English has been a restless language looking and searching for lands to invade and conquer. Since English itself is a product of conquest dating back to the invasion by some Germanic tribes (Angles, Saxons and Jutes) in AD 499, the behavior

of English should not come as a surprise. From inception therefore, English has remained “a language of conquest often and always seeking newer territories and peoples to dwell with ...” (Egbe, 2019: 4).

The dispersal of English from its ancestral birthplace to the present can be summarized in four waves. The first transportation of English started from within the British Isle when the language was taken to Scotland, Wales and Ireland (King, 2020; Douglas, 2020). This is regarded as the First Diaspora dispersal which was essentially internal. This was followed by the external Diaspora transcontinental re-rooting of English to United States (1584), Canada (1497), The Caribbean (early 1660s), Australia (1770), New Zealand (1769), and South Africa (1795). See Schneider (2020), Dollinger (2020), Kiesling (2020) and Aceto (2020). This is regarded as the Second Diaspora Dispersal. The Third Diaspora Dispersal accounts for the dispersal and planting of English to the South and South East Asia (1600s), former African British colonies (late 1600s), and South East Asia (1770s). See Gargesh (2020), Low (2020), Kamwangamalu (2020) and Bokamba (2020). The Fourth Diaspora Dispersal accounts for English in the South Pacific countries (1770s), English in South America, English in other parts of Europe including Russia, and parts of East Asia as well as English in the People’s Republic of China (Friedrich, 2020; Hilgendorf, 2020; Proshina, 2020; Honna, 2020; Zhang et al., 2020).

The journey of English from a local language of about four million speakers in the 15th century when “... English was the language of Englishmen” Strang (1977:156) to an international language with about 1.5 billion speakers representing 20% of the 7.5 billion of the world’s population (Lyons, 2017), the language has taken up diverse communicative, cultural, economic, educational and political and administrative responsibilities across the globe. In fact, Kachru (2020:454) listed the following as the functional domains of English in its status as the world’s lingua franca: the language is used as an access code, for advertising, corporate trade, development, government, linguistic impact, literary creativity, and literary renaissance. Others are its use in news broadcasting, newspapers, scientific higher education, scientific research and social interaction.

The dominance and centrality of English evoke fears and concerns. The global murderous march of English has earned English charges of “linguicide”. Schneider (2003:233) notes that “English ... has been damned as a ‘killer language’, responsible for the extinction of innumerable indigenous languages, dialects, and cultures around the globe”. The same fear is echoed by Crystal (2001) when he recalled Jim Erickson’s big question: “Will the English-dominated Internet spell the end of other tongues?” in Erickson’s article, “Cyberspeak: the death of diversity” (*Asianweek*, July 3, 1998, p. 15).

Whatever the apprehensions are for English, some scholars think English has come to the rescue of many. In fact, Kings (2020) regard English as “a gift to the globe, to millions of people, often to people who would not be able to express themselves to a wider audience if not for English”. Essentially because as he reasons, for good or for bad, ... English has become a world language in both senses of the term, international and global: international, as a medium of literary and other forms of cultural life in (mainly) countries of the former British Empire; global, as the cogenitor of the new technological age, the age of information (Kings, 2020:9).

While there might be aversions for the domineering role of English globally, the reality is that “English is too deeply entrenched, and if people are deprived of the chance of learning it they are the ones who suffer” and ironically “if you want to resist the exploitative power of English, you have to use English to do it” (Kings, 2020:7). That is the situation of some African writers and thinkers like Ngũgĩ wa Thiong’o who eloquently writes about decolonizing the African mind and African literacy education using English to do so regrettably. It is the same scenario with an African professor of Yoruba language for instance writing and delivery his or her inaugural lecture in English not in Yoruba, the language of the profession. To sum up the sentiments of English as both a global and international language, and the language one can hardly do without, Kachru (2020: 447) evokes the metaphor of English as “the Speaking Tree”, which “represents both fear and celebration, aversion and esteem, and, indeed, agony and ecstasy” rolled into one.

The re-rooting of English is a subject of many English discourses. For instance, Kachru’s concentric and overlapping circles of inner, outer and expanding circles categorization (Kachru, 1985, 1997). There is Graddol’s (1997) classification of (native) speakers, English as Second Language (ESL) speakers and English as a Foreign Language (EFL) speakers. There are also Yoneoka’s (2001) English Umbrella Tree and the revised English Umbrella Tree of Wenfang (2011) including the Dynamic Model espoused by Schneider (2003). What these models and paradigms point at is the spread of English globally and the conditions which favor and facilitate the spread of English and its domestication in the course of its transportation. In particular Schneider’s Dynamic Model shows how identity, contact and accommodation account for the conditions for the easy spread and domestication of English on the one hand and the emergence of “new Englishes” within the context of “World Englishes” on the other.

Englishization as domestication, hybridization and Anglicization

Following the transportation of English across oceans and continents, English found itself in the hands of new owners and as it is with languages in contact, a new English is and was bound to emerge. This is the basis for the birth of “Englishes” and “New Englishes” and the popularization of a new subdiscipline of English studies and scholarship known as “World Englishes”. Discourses on world Englishes have developed frameworks, paradigms and models to describe new Englishes dating back to the early 1980s (Strang, 1977) and more recently to Braj B. Kachru with the publication of the article on the Indianization of English (Kachru, 1983).

One of Africa’s finest writers, Chinua Achebe makes the point that while writing in English is inevitable, that English must be a “new English” able to express and carry the burden of his peculiar African experiences but still in full communion with international intelligibility. This is the English we find in use in his novels. This kind of claim is echoed by Germino Abad, the Philippine poet in another way when he declared, “English is now ours; we have colonized it” (Egbe, 2019:12). This sense of colonialization of English is important in understanding the domestication and hybridization of English within Kachru’s “outer” and “expanding”

circles whose population (over a billion) actually outstrips the “inner” circle population (over three hundred million) who are regarded as “native” speakers of English.

To account for membership of those who have domesticated and hybridized English, Platt et al. (1984) as well as Wenfang (2011) identify the following four features:

- (i) The English is developed through the education system. This may not necessarily be so because in some Nigerian families, English is acquired as a home language not a language learned or encountered in the formal school setting.
- (ii) The English is developed where majority of the people do not speak a native variety of English. This point is taken but raises the question of who is a “native” speaker of English or who qualifies to be called a “native speaker”.
- (iii) English is used for a range of functions by those who speak and write it.
- (iv) The English is localized and nativized.

It is the feature of localization and nativization (in terms of indigenizing) that we find the domestication and hybridization of English in Nigeria beyond the obvious differences in phonology, syntax, vocabulary, semantic, discourse and pragmatic resources from other varieties of world Englishes.

Anglicization of names

Whether as personal names, place names or product names, there is a growing tendency to Englishize everything. Although this tendency started with missionary activities in the early 20th century in Nigeria, to the present, some Nigerians still opt to give English names to their children; sometimes children opt for the Americanized versions and fancy coinages such as “Dooshima” shortened to “Dosh”, “Torkwase” shortened to “Tk”, “Iveren” shortened to “Iv”, and “Ter” which is a prefix to many names shortened to “Terry” or “Teddy” as found in Tiv language spoken in North Central Nigeria. You find some Nigerians prefer to bear such names as “Blessing”, “Faith”, “Gift”, “Glory”, “Goodness”, “Grace”, “Joy”, “Love”, “Mercy”, “Miracle”, “Peace”, and the like whereas there are local or indigenous rendering of such names. This is different from the Anglicization of African personal, object and place names such as “Bassey” for “Awase” or “Obasse”, “Ekureku” for “Ekwelekwu” as spoken in Leggbo, a language spoken in Cross River State, Nigeria. Even in the name of places, products and advertisement, one sees that the general linguistic landscape of Nigeria, is replete with Englishization mostly and other times the localization of English. In Cross River State, Nigeria for instance, the government has created some new economic destinations and established some new economic ventures. For example, “Cally Air” (for a new airline in the state) and “Calavegas” (a new economic and entertainment city near Calabar, the capital of Cross River State). One can understand the logic behind the coinages with no reference to any indigenous icons or symbols.

Englishization of indigenous languages and indigenization of English

Englishization is at work when we look at the orthographies of African languages. What we see generally are local languages wrapped in English through the use of Roman script. Some languages have extended Englishization to creating standardized structures to resemble English utterances. For example, in Igbo, a language spoken in South East Nigeria, “good morning” is rendered as “*utuṭu ọma*” instead of “*! bọla chi*” or “*! teala*” or “*! saala chi*” which translates to “have you woken up”. In a good number of cases there is a preponderant use of code switching and mixing of English and local languages in the speech of many educated Nigerians whether in conversations or even formal religious discourses. There are instances of English words taking into local languages which one might regard as the indigenization of English. For instance, we have “*tabul*” for “table”, “*mato*” for “motor”, “*beredi*” for “bread” among the Tivs in Nigeria which arguably represent an instance of the Tivinization of English. This feature is quite common in many other Nigerian languages.

Educational policies and English in Nigeria: some historical backdrop

Historically, the first formal pronouncement on education in Nigeria dates back to the Education Ordinance of 1882 promulgated for the British West African territories of Nigeria and Gold Coast (present day Ghana). The Ordinance was to regulate school’s curriculum and conditions under which schools were to obtain grants from the colonial government. In that Ordinance, the teaching of English was made compulsory. That means, if English were not taught, such a school was not qualified for any grants. The status which was given to English as a compulsory school subject has remained from 1882 to 2022, some 140 years later. The status and the role of English has not been reversed nor has it diminished in the educational sector. Yet, there have been over ten different education ordinances, codes, acts and edits (Wey-Amaewhule, 2018) in Nigeria. This is in addition to the revision of several versions of the National Policy on Education (NPE) first published in 1977 with its sixth edition published in 2013.

What has remained constant is the place of English in Nigeria’s educational policy and planning and the inconsistencies with regards to the language provisions in the various editions of the NPE (Bamgbose, 2020; Egbe and Ayeni, 2018; Trudell, 2018). In fact, Trudell (2018) compared the provisions of two versions of the NPE (2004 and 2013) and highlighted the inconsistencies as shown in the Table 1 below.

A closer look at the two versions of the NPE will reveal some of the following. The 2013 version admits the concerns raised by those who question the undue reference to the three “major” Nigerian languages which tend to diminish the so called “minority” Nigerian languages. There is also no reference to a major language. Perhaps, any Nigerian language will suffice which gives room for linguistic egalitarianism and the right of language. While the 2013 version rightly says Nigerian languages will be taught not learned, there are no clear provisions on how government will ensure that the mother tongue or the language of the immediate community

Table 1 NPE 2004 and 2013 compared.

NPE 2004	NPE 2013
Section 1.10: Every child is required to learn “one of the three Nigerian languages: Hausa, Igbo and Yoruba”	Section 1.10 has been deleted.
Section 5.24 and 5.25: “One major Nigerian language” listed as a core subject for junior and senior secondary school	No reference to Hausa, Igbo, Yoruba or “major Nigerian languages” is made in the document. <i>Taliced mine.</i> Junior and secondary school subject lists include “one Nigerian language”
Section 1.10: “Every child shall learn the language of the immediate environment”	Section 1.8: “Every child shall be taught in the mother tongue or the language of the immediate community for the first four years of basic education” (including pre-primary and P1-P3)
Section 2.14: Pre-primary education is to be carried out in the mother tongue or “the language of the immediate community”	Section 2.16: For early childhood care development and education, “Government will ensure that the medium of instruction is principally the mother tongue or the language of the immediate community”.
Section 4.19: “The medium of instruction in the primary school shall be the language of the environment for the first three years. During this period, English shall be taught as a subject”	Section 2.20: “The medium of instruction in the primary school shall be the language of the environment for the first three years in monolingual communities. During this period, English shall be taught as a subject”.
From P4 to P6, “language of the immediate environment” is a subject (as is French), with English as medium of instruction	Arabic added to subject languages in P4-P6.
Section 5.24: As a core subject in JSS, “the language of the environment shall be taught as L1 where it has orthography and literature. Where it does not have, it shall be taught with emphasis on oracy as L2”. French is also a core subject; Arabic is an elective.	Section 2.23: Junior secondary education curriculum subject list includes “one Nigerian language”. French and Arabic are optional.
Section 5.25: Senior secondary school non-vocational elective includes “any Nigerian language that has orthography and literature, etc.” French and Arabic are also elective subjects	Section 3.38: Senior secondary education curriculum includes an optional subject of “any Nigerian language that has curriculum”. French and Arabic are also elective subjects.

Source: Adapted from Language and Education in Nigeria (Trudell, 2018).

will be the principal medium of instruction. It is striking that the 2013 version alludes to monolingual communities. This is a far cry in Nigeria except in some rural and hard to reach communities. As if Nigeria is not already plagued with linguistic plurality, the 2013 version makes provision for more foreign languages (French and Arabic). This simply compounds the language in education situation in Nigeria. In fact, some states in Nigeria have added Chinese to the linguistic mix.

English in the language policy provisions

Apart from the inconsistencies highlighted above, what is clear is the status of English. The reality therefore is that English appears central in discourses on language policy globally (Pennycook, 1994; Spolsky, 2004; Sonntag, 2000; Ricento, 2000). This is because the logic of postcolonial policy is maintenance rather than change. While post-independence governments appear to be making language policies, most of the time, they are only perpetuating colonial language policy. This inheritance situation has meant a futile struggle between change and continuity, with the latter usually gaining the upper hand. In practically all African countries colonized by Britain, English remains an official or co-official language. Attempts to promote the use of any other language as national or official have resulted either in failure or limited success (Bamgbose, 2020:662). The case of South Africa where nine African languages are recognized as co-official languages with English and Afrikaans in the country’s Constitution may be an exception.

Even in countries where it seems the national language issue has been resolved, English has appeared inevitable. In Rwanda for instance, there are four official languages with Kinyarwanda as the national language spoken by 99.4% of the population. Yet in 2008, English was appointed a co-official language as the language of administration as well as the medium of instruction throughout the educational system from grade 1. The legitimization of English is positioned as “an inevitable, pragmatic and rational measure for economic development” (Rosendal and Ngabonziza, 2022:1). So the choice of English does not appear as free choice but a constrained choice which undermines the inclusiveness and equity conditions that should inform language policy and planning in its broad sense. It is easy to see why the spread of English is often viewed as “natural, neutral and beneficial” (Pennycook, 1994:9) a claim that is increasingly being disputed (Phillipson, 1992; Skuttnabb-Kangas, 2000; Tollefson, 2000).

Within the Nigerian context however, Bamgbose (2020) identifies prestige, status, and functionality as critical factors which favor English in the language policy discourse. Nevertheless, Bamgbose (2020) posits that further issues abound with regard to the role of English thus:

... what role English has as an official language, whether any other language can share this role at the national level, what further allocation of functions (particularly in education), when English should be introduced in schools, and what models of English should be aimed at in English language teaching (p. 660).

In the contexts of Outer Circle countries, where English is largely acquired through formal education such as Nigeria, the detrimental effects of the hegemony of English can be seen in social stratification, exclusion, and problems associated with education and literacy, the status of languages other than English, and language rights (Bamgbose, 2020:664). Fortunately, or unfortunately, Fishman (1996: 639) is of the view that “the socioeconomic factors that are behind the spread of English are now indigenous in most countries of the world and part and parcel of daily life and social stratification.” The issues this situation throw up will be addressed in the next subsections.

English, language of education and literacy issues

With regard to the use of English as the medium of instruction and as a preeminent school subject, Tollefson (2000: 14–15) describes the situation of virtually all Outer Circle countries using Philippines as a case study which is also true of Nigeria.

Access to higher education and to jobs is predicated on English proficiency. To gain admission into tertiary institutions in Nigeria, candidates are expected to write and pass a tertiary qualifying examination with English as a compulsory subject and a credit pass at the secondary school external examination.

There is a dual system of education in which children of the elite attend well-funded and well-staffed private schools, while children from poor families attend overcrowded and poorly-staffed public schools. This may not be the case for all the private schools in Nigeria. Some private schools do not have adequate staff yet parents would willingly prefer to send their children and wards to such private institutions as a social and class marker.

Although the medium of education is English in both public and private schools, the quality of instruction in the latter is superior and the products of the private schools end up with greater proficiency in English. In fact, there are some private schools that offer and opt for a completely British or American or Australian or Canadian curriculum for students in Nigeria. Besides the quality of instruction, the disruptive nature of public tertiary institutions in Nigeria due to frequent labor disputes open up the private schools, especially for higher education, to more access and quality of instruction.

The differential in English proficiency means a better chance of the elite getting more lucrative jobs and thus being able to send their children to private schools even when the private schools charge higher tuition.

Tollefson (2000) further argues that since the policy makers are drawn from the ranks of the elite, self-reinforcing and self-perpetuation of the social class (Dua, 1996; Rubagumya, 1991) appear as the guiding principles in the enactment of educational policies in favor of English in order to secure privileged positions and “an overwhelming advantage in terms of access to jobs, social status, and power” (Tollefson, 2000:15).

One of the challenges arising from the dominance of English as the medium of instruction and as a school subject is the increasing concern for educational failure and the role of English in Nigeria’s language policy and planning. In the Nigerian school system, for instance, in spite of the duration in which English has been in Nigeria, Bamgbose (2020) observes the following problems: teachers who are expected to teach English are themselves poor models of English language teaching; children are made to learn through English medium from the fourth year of primary education when their English proficiency is still low; concern for dropout rates which range from 37.4 percent to 52.7 percent at primary school level; lack of readiness for English language tasks which are expected at the beginning of secondary education. Others are: carryover of poor performance in English to end-of-secondary-school examination, which is marked by high failure rates of over 66 percent in English and between 56 percent and 73 percent in other subjects; mass failure in University Matriculation Examination compulsory English paper, as well as in other subjects; and poor standard of English among university undergraduates necessitating the requirement of a compulsory use of English course in tertiary institutions (Bamgbose, 2020:668).

Literacy education in Nigeria: prospects and hopes for Nigeria

Although it is obvious that no matter how desirable language policies may be, if the will to implement them are not strong and strategic, the impact of such policies will remain skeletal. The situation in Nigeria is that while the policy is being formulated, there is no effort to articulate the implementation mechanisms at any level. Sadly, it is at the point of implementation, especially with the language provisions, that we realize that we are engaged in an impossible mission and what teachers then do is to resort to the “go-for-English” solution (starting with English even when learners have no appropriate proficiency in English to learn in English). So, while the problems identified by Bamgbose (2020) are realities, and since the use of English as the medium of instruction in most schools in Nigeria is non negotiable at the moment, those problems have been attributed to English literacy failure (Egbe, 2021). That is, the inability of the students to meet the language and literacy demands the content areas make on them. This is not essentially the fault of the students but rather the students not being equipped with the content area literacy needs through systematic instruction. The solutions lie in formulating and implementing a workable language of education and literacy policy, producing and accessing adequate literacy materials, adopting innovative instructional strategies, and deploying digital literacy technologies and approaches across the various educational levels in Nigeria.

Formulating and implementing a workable language of education and literacy policy

There is no gainsaying that Nigeria is still struggling with formulating a coherent and implementable language policy. There is no literacy policy at the moment. I have pointed out the inconsistencies in the language provisions in the NPE and the difficulty of implementing the language stipulations in the 1999 Constitution of the Federal Government of Nigeria. In that document, English is the official language for the business of the national assembly and provides that Hausa, Igbo and Yoruba will be used after adequate arrangements have been made. Of course, no “adequate arrangements” have been made and may not likely to be made in the near future. This is understandable because of the seeming volatility that accompanies the discourse of national language(s) in Nigeria. English is still an integrating language. Even where some state houses of assembly are in what may be regarded as reasonably linguistically homogeneous, none of the indigenous languages have been adopted in the business of the house or in governance. Besides the provisions in the Constitution of Nigeria and the National Policy on Education (NPE), Nigeria needs a national language policy complimented with a comprehensive literacy policy. Such a policy must fundamentally pay attention to the multicultural and multilingual complexity of Nigeria and root for local, regional or state languages at appropriate domains and levels of use. In doing so, the enormity of developing such local languages (in terms of human and financial resources) cannot be glossed over. Whatever local languages adopted will complement English not the other way round because Nigeria cannot afford to do away with English as a national language no matter the volume of patriotic and nationalistic inclination voiced. English is a strong hope for Nigeria.

Producing and accessing adequate literacy materials

One of the failures of the current language provisions in the NPE is the dearth of literacy materials for educational purposes in both the local languages and English but more on the former. How are children supposed to be taught in their mother tongue or language of the immediate community when some of those languages have no orthography and where one exists, there is paucity of educational and literacy materials to assist and support teaching and learning. The huge financial and personnel cost required to provide for the numerous local languages defeats such a proposal and instead favors the continuous use of English. There are situations where children in over crowded classrooms cluster around the few reading and workbooks in English available in the classroom. How much children can learn in such a class setting is better imagined! Therefore, providing adequate literacy materials in English and making such materials easily accessible make more economic, educational and pragmatic sense to reduce the number of educational failure across Nigeria’s educational levels.

Adopting innovative instructional strategies

Educational failure in Nigeria is not just about the language of instruction which is English. Research and field observations have shown that some teachers are still depending on instructional strategies of the pre-independence era such as teaching reading by asking students to read a passage and answer the questions which follow when the students don’t know how to read. This shows that there is no distinction between learning to read and reading to learn. Also of great concern is the fact that most Nigerian primary schools are yet to include reading as a school subject and when reading is taught, the approach to reading instruction is a far cry from what contemporary reading research and instruction advocate. This is also manifest in the teaching of language whether English or any other language for that matter. So, to navigate the English medium instruction, the need to emphasize innovative instruction is important for desirable educational outcomes in Nigeria. There is great hope with the support of local and international development agencies working hard to boost the capacity of teachers especially at the basic education level in Nigeria through capacity building programmes both in some local languages and in English.

Deploying digital literacy technologies and approaches

As a corollary, to adopting innovative instructional strategies, is the compelling need to harness the resources of digital technologies and affordances across all educational levels. This is an issue which a national literacy policy will incorporate when such a policy is formulated. The astronomical transformation of how people learn and live in the 21st century through digital technologies has tremendous impact on educational delivery. How to use the knowledge and competencies of digital literacy as social practice within the context of the 4Cs of creativity, critical thinking, communication, and collaboration are essential ingredients of a digitally literate citizen. It goes without saying that as global citizens, who must remain relevant and competitive, any educational system which ignores the deployment and utilization of digital technologies in its educational and literacy development and promotion, does so at a very high and regrettable cost. Luckily, Nigeria has a vibrant information technology agency and policy which are positioned to transform Nigeria’s ICT landscape if adequately harnessed.

Conclusion

I have explored the concept of Englishization as spread and domestication, the influence of Englishization in the linguistic landscape of Nigeria as well as the critical issues of English as the language of instruction in the Nigerian educational context. I have also acknowledged the literacy concerns of Englishization in Nigeria and canvassed for the formulation of a language and literacy policy for Nigeria which pays homage to the multilingual and cultural plurality of Nigeria without undermining the place of English in Nigeria. There is hope for Nigeria if we fashion out an egalitarian language and literacy policy which accords the right to language and language rights to citizens in addition to promoting innovative instructional strategies, providing adequate literacy materials

and deploying digital technologies to maximize educational outcomes in line with 21st century global demands. While researchers in world Englishes cannot turn a blind eye to the problems of educational failure or unfavorable language policy outcomes, which sometimes is placed on the door steps of Englishization, English remains one major, global and powerful resource in the world today (Bamgbose, 2020) and the place of such a language, especially in a multilingual and multicultural context like Nigeria will be hard to replace for a long time to come.

References

- Aceto, M., 2020. Caribbean Englishes. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 87–103.
- Adetugbo, A., 1979. The development of English in Nigeria up to 1914: a sociolinguistic appraisal. *J. Hist. Soc. Nigeria* 9 (2), 89–105.
- Awonusi, V.O., 2004. Cycles of linguistic history: the development of English in Nigeria. In: Dadzie, A.B.K., Awonusi, V.O. (Eds.), *Nigerian English: Influence and Characteristic*. Concept Publications Limited, Lagos.
- Bamgbose, A., 2020. A recurring decimal: English in language policy and planning. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 659–673.
- Bokamba, E.G., 2020. African Englishes and creative writing. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 173–197.
- Crystal, D., 2001. *Language and the Internet*. Cambridge University Press, Cambridge.
- Dollinger, S., 2020. English in Canada. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 52–69.
- Douglas, F., 2020. English in Scotland. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 17–29.
- Dua, H.J., 1996. The politics of language conflict: implications for language planning and political theory. *Lang. Probl. Lang. Plan.* 20 (1), 1–17.
- Egbe, G.B., 2014. Language preference among Nigerian undergraduates and the future of English. *Int. J. English Lang. Transl. Stud.* 2 (4), 52–65. Retrieval at: <http://www.eltsjournals.org>.
- Egbe, G.B., Ayeni, A., 2018. Learning in English: the dilemma of Nigerian children. *J. English Teach. Assoc. Nigeria* 8, 1–13.
- Egbe, G.B., 2019. Of New Englishes and New Literacies. First Inaugural Lecture, Veritas University. Abuja delivered on Thursday, November 14, 2019.
- Egbe, G.B., 2021. Advanced and academic literacies: academic English proficiency for higher education success. In: Ushie, G.O., Inyabri, I.T., Ebim, A.M. (Eds.), *Language and Literary Studies in Society. A Festschrift for Professor Eno Grace Nta*. University of Lagos Press and Bookshop Limited, Lagos, pp. 639–658.
- Erickson, J., 1998. Cyberspeak: The Death of Diversity. *Asianweek*, p. 15. July 3, 1998.
- Fishman, J.A., 1996. Summary and interpretation: post-imperial English 1940–1990. In: Fishman, J.A., Conrad, A.W., Rubal-Lopez, A. (Eds.), *Post-imperial English*. Mouton de Gruyter, Berlin, pp. 623–641.
- Friedrich, P., 2020. South American Englishes and Englishes in South America. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 202–214.
- Gargesh, R., 2020. South Asian Englishes. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 107–134.
- Graddol, D., 1997. *The Future of English*. British Council, London.
- Hilgendorf, S.K., 2020. Euro-englishes. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 216–231.
- Honna, N., 2020. East Asian Englishes. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 248–265.
- Kiesling, S.F., 2020. English in Australia and New Zealand. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 70–86.
- Kachru, B.B., 1983. *The Indianization of English: The English Language in India*. Oxford University Press, Oxford.
- Kachru, B.B., 1985. Standards, codification and linguistic realism: the English language in the outer circle. In: Quirk, R., Widdowson, H. (Eds.), *English in the World: Teaching and Learning Language and Literature*. Cambridge University Press, Cambridge.
- Kachru, B.B., 1997. World Englishes and English-using communities. *Annu. Rev. Appl. Ling.* 17, 66–87.
- Kachru, B.B., 2020. World Englishes and culture wars. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 447–472.
- Kamwagamalu, N.M., 2020. South African Englishes: form and functions. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 160–172.
- Kings, R.D., 2020. Beginnings. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 5–16.
- Low, E.L., 2020. English in South East Asia. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 135–158.
- Lyons, D., 2017. How Many People Speak English, and Where is It Spoken? Retrieved from: <https://www.babbel.com/en/magazine/how-many-people-speak-english-and-where-is-it-spoken/>.
- Pennycook, A., 1994. *The Cultural Politics of English as an International Language*. Longman, Harlow.
- Phillipson, R., 1992. *Linguistic Imperialism*. Oxford University Press, Oxford.
- Platt, J., Webe, H., Ho, M.L., 1984. *The New Englishes*. Routledge and Kegan Paul, New York.
- Proshina, Z.G., 2020. Russian Englishes. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 232–247.
- Ricento, T., 2000. Introduction. In: Ricento, T. (Ed.), *Ideology, Politics and Language Policies: Focus on English*. John Benjamins, Amsterdam, pp. 1–8.
- Rosendal, T., Ngabonziza, J.A., 2022. Amid Signs of Change: Language Policy, Ideology and Power in the Linguistic Landscape of Urban Rwanda. <https://doi.org/10.1007/s10993-022-09624-5>. *Language Policy*.
- Rubagumya, C.M., 1991. Language promotion for educational purposes: the example of Tanzania. *Int. Rev. Educ.* 37 (1), 67–85.
- Schneider, E.W., 2003. The dynamics of new Englishes: from identity construction to dialect birth. *Language* 9 (2), 233–281.
- Schneider, E.W., 2020. English in the United States. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 31–51.
- Skutnabb-Kangas, T., 2000. Linguistic human rights and teachers of English. In: Hall, J.K., Eggington, W.G. (Eds.), *The Sociopolitics of English Language Teaching*. Clevedon: Multilingual Matters, pp. 24–44.
- Sonntag, S.K., 2000. Ideology and policy in the politics of the English language in north India. In: Ricento, T. (Ed.), *Ideology, Politics and Language Policies: Focus on English*. John Benjamins, Amsterdam, pp. 133–150.
- Spolsky, B., 2004. *Language Policy*. Cambridge University Press, Cambridge.
- Strang, B.M.H., 1977. *A History of English*. Methuen and Co. Ltd., London.
- Tollefson, J.W., 2000. Policy and ideology in the spread of English. In: Hall, J.K., Eggington, W.G. (Eds.), *The Sociopolitics of English Language Teaching*. Clevedon: Multilingual Matters, pp. 7–21.
- Trudell, B., 2018. *Language and Education in Nigeria: A Review of Policy and Practice*. British Council and UNICEF.

- Wenfang, C., 2011. Different models of English as an international language and the implications for teaching non-English majors. *Chin. J. Appl. Ling.* 34 (2), 5–17.
- Wey-Amaewhule, B., 2018. Historical development of education law and implications for educational policy and planning in Nigeria. *Int. J. Innov. Educ. Res.* 6 (2), 49–53.
- Yoneoka, J.S., 2001. The English umbrella model of a multicultural language system. *Asian English Stud.* 7 (2), 69–83.
- Zhang, W., Bolton, K., Botha, W., 2020. English in the People's Republic of China. In: Nelson, C.L., Proshina, Z.G., Davies, D.R. (Eds.), *The Handbook of World Englishes*, second ed. John Wiley & Sons, Inc., pp. 266–280

The salient truthfulness of Aloha 'Āina Theory and Guerrilla Praxis

Kekailoa Perry, College of Education—Educational Foundations, University of Hawai'i at Mānoa, Honolulu, HI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	69
Higher education, neoliberalism, and indigenous knowledge	70
Indigenizing the university—the new neoliberal construct	70
Academic gatekeeping and the pedagogy of refusal	71
Indigenizing the university of Hawai'i	72
Establishing native theory & praxis	73
Application of Kaupapa Māori Theory	73
The development of Aloha 'Āina as theory	73
Toward an Aloha 'Āina Theory & Guerrilla Praxis	74
Aloha 'Āina Theory and Guerrilla Praxis applications	74
The principle of sovereignty and de-occupation	76
The principle of centering genealogy and nationality	77
The principle of kuleana and moral obligations	77
The Principle of truth telling, radicalism and protest (community & liberation driven)	77
Conclusion	78
References	79

Introduction

21st century higher education is a treacherous endeavor. It is a dangerous tool of settler-colonial hegemony but can also create opportunity for emancipatory action for Indigenous and marginalized peoples (Tuhiwai-Smith, 2021: p. 38). Settler-colonial higher education is represented in the status quo or mainstream policies and pedagogies of contemporary Western academic institutions. Emancipatory pedagogies exist in the margins of the academy where Indigenous and liberation scholarship commonly reside. Indigenous scholars in particular face a sophisticated array of neoliberal and neocolonial policies that appropriate their ways of knowing and undermine the value of traditional and cultural knowledge with tokenism and a dollar. This chapter challenges the prevailing status quo of the Western academy through the application of Indigenous theory and praxis. Each section of this chapter will address salient issues concerning the neoliberal and neocolonial university system, the growing settler-colonial phenomenon of "Indigenizing the university," and the Indigenous theory and praxis that serve as pedagogies for liberation. Each salient point (Brayboy, 2005, p. 427) will provide space to introduce a developing Indigenous pedagogy called Aloha 'Āina Theory and Guerrilla Praxis as a counter hegemonic challenge to mainstream higher education.

However, before addressing the Indigenous theory and praxis of this chapter, several important foundational ideas that guide this research must be mentioned. First, this chapter is framed in the scholarly space of Indigenous knowledge and liberation studies. Those familiar with this academic area often apply anticolonial, postcolonial, settlercolonialism, critical race or Indigenous theory, and liberation methodologies as part of the analysis. I do the same here with one critical exception. That is, Hawaiians—the native or aboriginal people of the Hawaiian archipelago who are born from a genealogical relationship to the lands and gods—are not Indigenous like the native scholars of the US, Canada, or Oceania referenced in this chapter.

Sai (2008) rightly argues that in the context of international law Hawaiians are not Indigenous to the United States. Instead, he establishes a powerful historical and legal analysis that the Hawaiian nation state is occupied, not colonized. As a result, "the history of the Hawaiian Kingdom has fallen victim to the misuse of this term [colonization and indigeneity] by contemporary scholars" (2008: 111, bracket added). Sai also cautions that "when Hawaiian scholars and sovereignty activists, in particular, consistently employ the terms and theories associated with colonization and indigeneity, they are reinforcing the very control they seek to oppose" (2008: 113, brackets added). Young (2006) tempers Sai's rigid application of international law by asserting that "There is no inherent conflict between an identity that is Indigenous before 1795 or after 1843 [the year the Kingdom became a recognized nation state] ... it merely presumes a parallel coexistence with what is aboriginal from 1843 to this very second" (2006: 5, brackets added). Young's analysis suggests that even in an occupied state, Hawaiian aboriginal people's social and political experiences reflect those of other Indigenous peoples because of the "colonial-like" pattern of oppression imposed by the United States. Thus, I assert that Indigenous political and scholarly analysis can be applied here without spoiling the arguments made by Sai.

Second, the Aloha 'Āina Theory and Guerrilla Praxis analyzed in this chapter is under a state of continued development. Readers must keep that in mind as several aspects of the theory and praxis may require adjustment to maintain an effective and active resistance to the sophisticated hegemonic policy approaches utilized by mainstream academia.

Third, reference to “radical,” “activist,” “liberation,” and “sovereignty” are purposefully confrontational. From the perspective of academics like myself, promoting the status quo furthers the marginalization of natives and serves no purpose here. Thus, a more pro-activist voice is applied throughout. E mau.

Higher education, neoliberalism, and indigenous knowledge

The truth about contemporary Western higher education in the 21st century is that it reinforces the status quo at the expense of meaningful transformative change. For Indigenous peoples in the academy, status quo is revealed in a university's appropriation of native culture and identity to advance settler-colonial ideologies. In the past two decades, universities have feverishly shifted toward neoliberal economic and social policies that “protect the practices of the powerful from being publicly scrutinized as they appropriate resources to serve the interests of the few at the expense of the many” (McLaren, 2015: pp. 1–10). As such, the adoption of neoliberal policies in US educational institutions have “led to the defunding of schools and colleges” and further pushing universities to “conform to a corporate model of higher education” (Sudbury and Okazawa-Rey, 2009: p. 4). Apple (2019) succinctly outlines the hegemonic conditions of US education as a free-market labor process that encourages racial and economic domination of marginalized groups whose non-compliance is considered “somehow not fully American” thus rendering them invisible (2019: p. viii).

Smith (2004) describes the expansive reach of Western neoliberal policies on an international scale in his discussion of economic and educational reform and its impact on Māori in Aotearoa, New Zealand. G. Smith explains that a key function of the hegemonic strategy is to use white Western values and thinking “to marginalize, demean, derogate, and subjugate Māori people and their cultural preferences” (2004: p. 47). Similarly, Smith (2009) asserts that the university is an academic-industrial complex where Indigenous knowledge is treated as a commodity operating in a space that is governed by “multinational capitalism, US empire, and white supremacy” (2009: p. 40).

Indigenous educational spaces have been rendered almost negligible as a valued form of intellectual development in the context of US higher education. Trask's (1999) assessment two decades ago regarding the flagship Research 1 institution of the University of Hawai'i (UH) at Mānoa, highlights the prolonged struggle for Hawaiian academics working within the American academic-industrial complex. Trask's observation that universities function to “legitimate and entrench the power of the colonizing culture” by implementing a bureaucratic white male-dominated institution explains the structural powers driving the indigenizing university process today (1999: pp. 151–152).

Indigenizing the university—the new neoliberal construct

Since the second publication of Trask's influential book, *From a Native Daughter: Colonialism and Sovereignty in Hawai'i*, there have been several attempts by university administrators to “indigenize” the university. But, very little has been done to move beyond the neatly packaged, copious sprinkling of misappropriated native words and phrases that serve phatic university goals for faux diversity and accreditation (de Leeuw et al., 2013: p. 382). University efforts to appropriate Indigenous terminology for insincere claims of equality and inclusion only cheapens complex Indigenous ideas and covers up the dominant group's preservation of power (Apple, 2019: p. 166).

The process of indigenizing universities is common among many contemporary academic institutions. And while there are newly developed policies that acknowledge Indigenous lands and peoples, those seemingly well-intended policies have failed in creating meaningful transformative changes to the institution (Brayboy et al., 2015; de Leeuw et al., 2013; St. James, 2022; Stewart-Ambo and Yang, 2021). In fact, many Indigenous and non-Indigenous scholars have effectively argued that the “indigenizing the university” process supports neoliberal and neocolonial ideologies instead (Hill, 2012; de Leeuw et al., 2013; Gorrie et al., 2017; Antoine, 2017).

One way to understand the faulty logic behind an indigenizing university policy is to consider a system that has been designed by Indigenous scholars to address settler-colonial assimilationism and domination (Stein, 2003). American Indian Tribal Colleges (Tribal Colleges) developed as a response to the failures of the US higher education systems to assist the intellectual and cultural needs for native self-determination and sovereignty among other things (Benham and Stein, 2003; Shanley, 2003). According to Stein, the guiding principle for the development of Tribal Colleges was that “American Indian people must control the university” and that such a principle “is very much at the center of all successful movements of Indigenous, poor, or minority people” (Stein, 2003: p. 31). Stein argues that the rights and privilege to define and determine Indigenous education is a sovereign prerogative that resides with the Indigenous communities, not outsiders. Outsiders, like university administrations, set false presuppositions and goals that distract from native empowerment (2003: p. 31).

While the Tribal College systems are not without struggle and issues of settler-colonial influences (Shanley, 2003), the point here is to highlight the obvious limitations (and flaws) related to the academy's indigenizing policies. To state the obvious, Indigenous peoples create Indigenous knowledge and have the authority to decide (right or wrong) how to apply that knowledge. Contemporary universities may claim to have programs that support Indigenous initiatives, but unlike Tribal Colleges, Indigenous communities clearly do not control the power, direction, pedagogy, and delivery of knowledge in those systems (Brayboy et al., 2015). Instead, universities are executing nominal representation and ambiguous value systems in the name of indigenizing the universities.

Antoine (2017) asserts that contemporary universities are appropriating Indigenous knowledge to reinforce negative political tropes that tokenizes the native. Accordingly, native communities and knowledge are exploited and treated as window dressing (2017: p. 117). Indigenizing the university is an approach that may produce glossy marketing highlights to placate the well-intended administrator while simultaneously reinforcing rhetorical tropes (Keawe, 2008: p. 4; de Leeuw et al., 2013: p. 389; Mihe-suah and Wilson, 2004: p. 6) that undermine Indigenous peoples' knowledge and methodologies. In fact, administrative "good intentions" are the new neoliberal or neocolonial approaches employed by contemporary universities.

Authors de Leeuw, Greenwood and Lindsay suggest, that administrative good intentions create deep colonizing consequences that reinforce historical inequities and help to distract from more meaningful ways to establish anti-colonial or Indigenous world views in the academy (2013: p. 382). The authors provide a powerful critique of the indigenizing process used by universities in Canada that apply remarkably well to other institutions in the US and Oceania. Such indigenizing policies highlight the impulse of settler-colonial administrations to re-write history in a way that distances the administrators and the colonizing nation from the oppressive legacy of colonization and occupation. As a result, institutions are manifesting "new kinds of policies and practices ostensibly aimed toward addressing and ameliorating past wrongs" which the authors argue "demand critical attention because they function to obscure ongoing harms and injustices of colonial practice, subverting the very good intentions they purport to represent" (de Leeuw et al., 2013: p. 384, 391). Hill (2012) also critiques the "indigenizing the university" approach and argues that good intentions will not properly address Indigenous issues in the academy because the academy ignores "the profoundly colonial nature" of higher education (2012: p. 3). Unless there is "an explicit goal of decolonization" embedded in the policy, engagement with Indigenous thought might strengthen processes of colonization (2012: p. 3).

Aside from the occasional scattering of a few table scraps for Indigenous faculty positions, increased Indigenous enrollment, and targeted Indigenous program funds, the benefits of an indigenizing university principally favors the university's neoliberal "banking concept" and consumer economy (Freire, 2005: p. 72). In such a process, native academic efforts aimed at improving Western universities often come with unhealthy compromises that harm Indigenous knowledge and disempowers their communities (Hill, 2012: p. 3).

The truth about the academy is that it is rife with neoliberal and neocolonial ideologies utilizing policies like "indigenizing the university" to feign equality and recognition for marginalized peoples. However kind the indigenizing policies may appear, the lack of commitment, consequential sharing of power and control, and respect for native sovereignty by the academy, furthers the erasure of Indigenous communities.

Academic gatekeeping and the pedagogy of refusal

The Western university system serves as the gatekeeper of all knowledge. The system extracts Indigenous knowledge to reinforce Western biases while allowing various non-threatening strands of Indigenous and critical pedagogies limited spaces of recognition and functionality. Goodyear-Kaōpua and Kuwada (2018) assert that US political hegemony uses "safety zones" to negotiate acceptable forms of native knowledge in a piecemeal basis so that the settler state can sanction "spaces in which some aspects of Indigenous culture can be practiced as long as they do not disrupt or threaten settler society" (2018: p. 52). Thomas (2021) describes this form of native subjugation as "extractivism." For Thomas, extractivism is "a way of understanding forms of exploitation and subjection of people in capitalist systems" (2021: p. 1). She maintains that state-led engagement strategies with Indigenous peoples are not reciprocal arrangements and only those Indigenous concepts that most aligns with Western ways are cherry-picked out of context leading to the further oppression and the undermining of Indigenous communities (2021: p. 2).

The academy's idea of inclusivity is more reflective of the neoliberal status quo and less interested in meaningful applications of Indigenous knowledge. The result is that Indigenous scholars are incentivized to suppress the conscientization and transformative needs of the community. Thomas notes that well-intentioned collaboration between universities and Indigenous scholars will extract Indigenous knowledge by replacing it with "state language, systems and processes" (2021: p. 2). Even when Indigenous grassroots organizing successfully adapts counter-hegemonic voices to Western institutional learning, Indigenous energies and political capital are overextended on neoliberal equality and fairness debates rather than aggressively critiquing and transforming the institution (Hill, 2012: p. 1). Thus, Indigenous concepts are effectively extracted by university administrations to help make the academy's hegemony appear less overt and more palatable but operational nonetheless.

The practice of using Lorde's (2012) "master's tools" has shown some strategic usefulness but has its limitations. As Lorde predicted, those uses are proving far less transformative and unrelated to more salient issues like Indigenous sovereignty. Instead, neoliberal policies are used to skillfully divide and silence critical native voices allowing university administrators to selectively partner with willing and obedient native compradors "who want reform but don't really want to upset the status quo" (Bobo et al., 2001, pp. 19–21).

Grande (2018) offers sage guidance when challenging the neoliberal and neocolonial recognition policies of the academy. She cautions that settler-colonial policies like indigenizing the university, supplants more transformative Indigenous critiques within the "political project of liberal pluralism" and the "politics of recognition" that maintain the status quo (2018: p. 56). The politics of recognition superficially acknowledges the "respect for cultural differences" while asserting settler power over Indigenous peoples and knowledge (2018: pp. 49–50). Moreover, the politics of recognition encourages Indigenous collusion and complicity with neoliberal policies that will likely derail efforts to establish meaningful transformative change. The indigenizing the university approach is essentially a prescribed settler process that drives native academics to consent to their own oppression. Grande argues that Indigenous academics must engage in a pedagogy of Indigenous refusal. Indigenous refusal adopts radical alternatives that

“rejects the false promise of inclusion” found in the academy’s settler-colonial policies and that “asserts Indigenous sovereignty and peoplehood” (2018: p. 59). Indigenous refusal is an overt sovereign act that clarifies the positionality of the native with the neocolonial policies of the academy, and that encourages pedagogies that are culturally grounded, sovereignty based, and guerrilla in praxis.

Indigenizing the university of Hawai'i

Settler-colonialism has morphed in institutions across the US, Canada, and Oceania into more sinister and complex forms of oppression. Strategic planning at my institution, the UH, is an example of how the university employs neoliberal educational policies. Though this example appears space specific, native communities will find the UH policies strikingly similar to the negative policies being implemented in their own institutions.

The University of Hawai'i administration recently appropriated the traditional and cultural Hawaiian concept of “aloha 'āina,” falsely branding and redefining it with the UH's settler-colonial voice (University of Hawai'i, 2021: p. 8). The settler-colonial meaning of aloha 'āina ignores over a century of traditional stories and documented evidence on the topic including Sai-Dudoit and Tolentino's (2022) comprehensive research on the traditional, cultural, and political understanding of aloha 'āina that convincingly determined that it is, without question, a concept connected to the patriots or patriotism of the Hawaiian nation (2022: p. 27). The UH administration purposefully ignored and/or misread the available Hawaiian authoritative texts choosing instead to remake and reimagine aloha 'āina devoid of well grounded Hawaiian Indigenous knowledge.

The UH strategic plan *Manoa 2025: Our Kuleana to Hawai'i & the World, Strategic Plan 2015–2025* (“UH strategic plan” or “plan”), begins by appropriating a Hawaiian title then immediately asserting authority over Hawaiian knowledge and falsely presenting the university as the Hawaiian and worldly voice. The plan proclaims UH as a Hawaiian Place of Learning (University of Hawai'i, 2021: pp. 7–10) assigning itself the power over aloha 'āina, while ignoring the legitimate rights, duties, and claims of the Hawaiian community. The plan then fabricates a settler-colonial meaning for aloha 'āina:

Aloha 'āina is a relationship and worldview deeply rooted in Hawai'i. It is a recognition, commitment, and practice sustaining the ea—or life breath—between people and our natural environments that resulted in nearly 100 generations of sustainable care for Hawai'i. We recognize it is because of the aloha 'āina practiced by Native Hawaiians over those many centuries that we can enjoy the Hawai'i we have today. Therefore, as we continue to learn about our collective kuleana as a university truly of Hawai'i, we realize that we too have a role in aloha 'āina and we take guidance from Native Hawaiian ancestral knowledge and wisdom on this journey together (2021: p. 8).

The plan refers to Hawaiian people in the past tense, signaling that over time, UH can promote Hawaiian knowledge without including the Hawaiian community. The UH plan endorses erasure thus fulfilling Patrick Wolfe's warning that “settler-colonization is at base a winner-take-all project whose dominant feature is not exploitation but replacement” (1999: p. 163).

The UH plan also weaves in key neoliberal economic triggers for future UARC and private foundation grant funding. In particular, the phrase “sustainability” is purposely attached to aloha 'āina signaling to a specific segment of scholars and funders that Hawaiian culture and knowledge is for sale and UH is its pimp. Yet, nowhere in the authoritative Hawaiian Language dictionaries (Pukui and Elbert, 1986; Andrews, 2003) or in grounded Hawaiian epistemology is the term sustainability equal to the traditional, cultural, and political understanding of aloha 'āina. So, while aloha 'āina does have some association to cultural practices that might be considered sustainable, “sustainability,” as an ambiguous Western construct, lacks any cultural relevance (Salazar, 2014).

The most distressing part of the UH plan states “We recognize it is because of the aloha 'āina practiced by Native Hawaiians over those many centuries that we can enjoy the Hawai'i we have today” (University of Hawai'i, 2021: pp. 7–10). The statement is a blatant attack on Hawaiian knowledge that reinforces US hegemony and furthers native erasure. Indeed, claiming that Hawaiians can “enjoy the Hawai'i we have today” while US occupation maintains its assault on Hawaiian identity, culture, lands, and sovereignty is despicable. According to Wolfe, UH's plan to extract Hawaiian knowledge can never be innocent because “It is too deeply enmeshed in a historical relationship through which one's power is the other's disempowerment” and that such a process will surely support the settler-colonial goal of elimination (1999: pp. 163, 212–213). Today's settler-colonial Hawai'i maintains the Disney-like imagination where tourists and the military can frolic, exploit, and destroy native and cultural resources without accountability. Apple explains that “the ability of a group to make its knowledge into ‘knowledge for all’ is related to that group's power in the larger political and economic arena” (Apple, 2019: p. 65). Freire describes such political opera operations as indispensable and divisive mechanisms used by the oppressor to control consciousness and preserve the status quo (Freire, 2005: 139). Thus, the UH extractivist policy seems to fulfill Sai-Dudoit and Tolentino's prophetic statement that “Popular iterations [of aloha 'āina will] fall victim to the purposeful erasure of knowledge that was borne of and stewarded by Hawaiians through documentation in print, mele, and mo'olelo (stories) told over generations” (2022: p. 9, brackets added).

This example illustrates how effortless it is for universities to successfully deploy its settler-colonial power and embed fabricated cultural definitions into the mainstream, despite an abundance of contrary evidence from Indigenous academic and community scholars and primary resources. Like other so-called well-intentioned neoliberal universities, UH policies are securely anchored to US hegemony. Tuhiwai-Smith (2021) explains that no real progress has been made to decolonize the academy and even the most radical gestures of recognition “can end up as mealy mouthed, watered-down liberal devices that end up reinforcing the status quo” (2021: p. xii). UH purposely suppresses Hawaiian concerns relating to the misappropriation of culture, racist and

discriminatory policies, and desecration of the sacred cultural sites that they manage. Surely, such actions reflect “mealy mouthed” policies that reinforce the status quo.

The UH administration has followed the example of other Western higher education institutions by adhering to neoliberal and neocolonial economic and policy approaches that support the continued degradation of native knowledge. While limited improvements do occur through vigilant and purposeful struggle, they are not in and of themselves the liberation or sovereignty pedagogies that Indigenous peoples seek. The policy example here tells us that Indigenous knowledge in the academy is not an equal rights issue, it is an issue of liberation and sovereignty for the native nations and the awareness to accept nothing less (Smith, 2009: pp. 52–53).

Establishing native theory & praxis

For Indigenous peoples working within Western, neoliberal academies, advancing the wellbeing of native and marginalized communities often demands active resistance to—and refusal of—settler-colonial policies. Alfred (2004) explains that universities are “adamantly and aggressively opposed to Indigenous ways” and therefore, Indigenous academics have the responsibility “to work to defeat the operation of colonialism within the university” (2004: pp. 88–89). Similarly, author Kelly (2014) asserts that “Hawaiians can’t indigenize state institutions, but the state can certainly institutionalize Hawaiians” and cautions native academics from “academizing sovereignty” until it is “pressed like a flower between the pages of a book” written by the colonizer (2014: p. 46).

In this section, a developing pedagogy will be introduced as a response to the threat of settler-colonial policies expressed by Kelly and Alfred. It will emphasize the need for radical, activist-oriented approaches to empower the native community. Smith’s (1997) paradigm shifting research on Kaupapa Māori Theory will be a guiding methodology in the development of what I am calling “Aloha 'Āina Theory and Guerrilla Praxis.”

Application of Kaupapa Māori Theory

Kaupapa Māori Theory serves “as an intervention praxis into educational, cultural, and social crises which impacts disproportionately on large sections of the Māori population” (Smith, 1997: p. 26). Developed in large part because of the Māori cultural and political revolution of the 1980s, Graham Smith, the developer of Kaupapa Māori Theory, recounts the key to the revolution came by way of “Māori critical thinking and the realization by Māori that they could make changes themselves” (2004: p. 49). Kaupapa Māori theory functions as a “complex arrangement of conscientization, resistance and transformative praxis, which collectively seek to transform these twin crises related to education and culture” (2008: p. 27, citations omitted). The goal of Kaupapa Māori Theory is to “challenge and disrupt the commonly accepted forms of research” in order to privilege Indigenous ways of knowing and being (Mahuika, 2008: p. 4).

There are three major assumptions influencing Kaupapa Māori Theory. First, that Western institutions are culturally antagonistic to Māori/Indigenous peoples which entails greater pressure toward assimilation and thus conformity to settler-colonial ideologies. Second, that Western institutions are inherently biased and disadvantage those who are non-conforming. And finally, Western structures must be resisted to create healthy spaces for Māori and Indigenous peoples to fulfill their cultural aspirations. Recognizing the fluidity and shifting strategies of settler-colonial domination, Smith notes the need for Kaupapa Māori Theory to be equally flexible in its resistance strategies (Smith, 2004: p. 49).

Therefore, G. Smith provides space for other Indigenous and marginalized peoples, like Hawaiians, to apply Kaupapa Māori theory at various levels of consciousness and praxis in their own institutions. He explains that Indigenous educators “must develop radical pedagogy that is informed by their cultural preferences and by their own critical circumstances” and that they “must ensure ‘buy in’ from the communities that are purported to be served” (2004: p. 51).

Kaupapa Māori Theory demands a critique of both the settler-colonial structures of education on one hand, and the engaged native scholars who obtain positions of privilege on the other. The second area of critique is crucial because it requires self-reflexivity that can provide important intelligence for improving native consciousness and action in the academy (Mahuika, 2008: pp. 6–12). Thus, for Indigenous academics, the academy becomes a “crucial site for critical intervention” that requires a commitment to creating space for “the production of diverse perspectives, new ideas and different styles of thinking and writing” that contradict the status quo (hooks, 1990: p. 6-7).

Kaupapa Māori Theory has in fact become a widely accepted and powerful tool in developing critical and liberation pedagogies for Indigenous peoples. Mahuika argues that Kaupapa Māori’s greatest potential “may lie in its ability to challenge Māori to develop a greater awareness of who we are, what it is we really want, and how we want to go about achieving that” (2008: pp. 11–12). In a similar vein, I argue that the development of an Aloha 'Āina Theory combined with a Guerrilla Praxis are critical devices for interrogation and resistance against the neoliberal educational apparatus of the academy.

The development of Aloha 'Āina as theory

19th century Hawaiian educator and revolutionary Joseph Kaho'oluhi Nāwahīokalani'ōpu'u (Nāwahī) said that aloha 'āina was “ka ume mageneti iloko o ka puuwai o ka Lahui, e kaohi ana i ka noho Kuokoa Lanakila ana o kona one hanau” (Nāwahīokalaniopu'u,

1895). Editors of the Kamehameha Schools Ka'iwakīlōumoku Hawaiian Cultural Center awkwardly translate the phrase as “the magnetic pull in the heart of the patriot which compels the sovereign existence of the land of his birth” (Ke aloha, 2021). The nationalist context of Nāwahī's writing during the early stages of US occupation in 1895 is clear and unambiguous. Nāwahī implores Hawaiian patriots to understand aloha 'āina as the connection between the land, the Hawaiian nation, and the people that are clearly not of America.

Silva (2002) analyzed 19th century Hawaiian language articles relating to Hawaiian resistance to US annexation. She says Hawaiian people “united in unprecedented numbers under the banner of “aloha 'āina”—love for the land, their nation and their people—to draw on the strength such feelings had given them collectively” (2002: p. 119). Like Nāwahī before, Silva identifies aloha 'āina as a “central ideology for our ancestors” that is “a complex concept that includes recognizing that we are an integral part of the 'āina and the 'āina is an integral part of us” (2002: p. 4). Additionally, scholars Sai-Dudoit and Tolentino (2022) produced a chronological analysis of the Hawaiian term aloha 'āina by utilizing 19th century Hawaiian language primary resources (2022: p. 6). The authors conclude their research by stating unequivocally that, “There is no question that aloha 'āina refers to patriots or patriotism as it had come to be understood by the subjects of the Hawaiian Kingdom and its leaders” (2022: p. 27).

Aloha 'āina has a theoretical foundation based on customary and traditional practices. Goodyear-Kaōpua and Kuwada (2018) describe Hawaiian futurities through the traditional senet or 'aha cord metaphor, where the 'aha cord represents powerful links to the past that speak to future possibilities for Hawaiian nation building. The authors state, “Hawaiian independent futurities assemble practices of aloha 'āina (love for the land) that assert that our ability to sustain generation after generation requires healthier relationships with our lands and waters.” The authors go on to say that “one of the most enduring strands of our 'aha is aloha 'āina” (2018: p. 53).

Other Hawaiian scholars continue to reinforce the traditional and cultural understanding of aloha 'āina. For example, Tuti Baker's (2018) research demonstrates that aloha 'āina is a “political strategy of resistance that is anchored in an individual's spiritual and kinship relationship to place” (2018: p. 55). Likewise, Salazar (2014) identifies aloha 'āina as a grounding political philosophy that fuels Hawaiian struggles for sovereignty and human rights (2014: p. 205). He then advises future advocates that aloha 'āina is a Hawaiian political practice that “does not reduce important cultural differences to an idealized multiculturalism that sustains inherited, systemic, and institutional power” (2014: p. 223).

Aloha 'āina is a powerful theoretical foundation for today's liberation and activist pedagogies. It is a form of community conscientization that demands social justice, sovereignty, and the refusal of settler-state domination. G. Smith's Kaupapa Māori Theory and the cultural ideology of aloha 'āina are combined in the development of the Aloha 'Āina Theory and Guerrilla Praxis discussed below.

Toward an Aloha 'Āina Theory & Guerrilla Praxis

Aloha 'āina is a cultural concept that ties consciousness and praxis to Indigenous liberation and sovereignty. This section will cover several grounded assumptions of the theory. Then, four guiding principles of Aloha 'Āina Theory and Guerrilla Praxis will be introduced. To begin, I review the totality of George Helm's January 27, 1977, conscientized and transformative letter protesting the US military bombing of Kaho'olawe island (Ritte and Sawyer, 1978: p. 27; Helm, 1984; personal communication, 1977). The aloha 'āina protests over Kaho'olawe lasted more than 20 years resulting in an end to the bombing, a partial restoration of the island, and legislation requiring the island's return to the sovereign Hawaiian government (HI Rev Stat §6K). The letter is formatted in Table 1 for analytical purposes.

Helm's letter is a commanding narrative of emancipation, resistance, and refusal that recognizes the inseparable link between cultural integrity, God(s) and ancestral mana (spiritual power), and national sovereignty. These elements have guided the Hawaiian nation through 127 years of US occupation. More importantly, Helm offers an analytical framework for Aloha 'Āina Theory and Guerrilla Praxis.

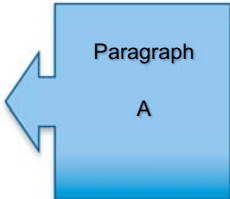
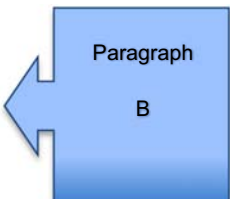
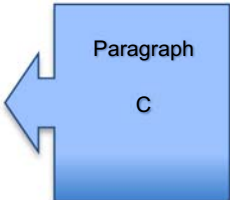
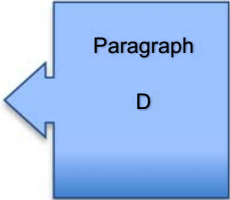
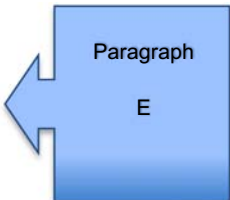
The analytical framework is derived from the major sections of Helm's letter and can be summarized in five developmental parts as outlined in Table 2 below.

Aloha 'Āina Theory and Guerrilla Praxis applications

Helm's letter provides several key assumptions for Aloha 'Āina Theory and Guerrilla Praxis. Helm begins with a broad conceptual framing that shows the complexity of aloha 'āina. In Paragraph A, he says, “I have my thoughts, you have your thoughts, simple for me difficult for you. Simply ... the reason is ... I am a Hawaiian and I've inherited the soul of my kupuna.” In Paragraphs B and C, Helm invokes the genealogical relationship with the maternal god Papa, describes the Hawaiian Islands as “sacred,” and recenters the Christian concept of “church” to Hawaiian land-based religions. Thus, he is engaging deep cultural and spiritual elements in his activism.

In Paragraph A, Helm identifies the first major tenet of Aloha 'Āina Theory—his positionality and grounding in Hawaiian (Indigenous) knowledge and spirituality. Helm's tenet aligns seamlessly with Tuhiwai-Smith's assessment that native spirituality is an important space of resistance because it privileges the Indigenous scholar and serves as “one of the few parts of ourselves which

Table 1 January 27, 1977, letter by George Helm.

I have my thoughts, you have your thoughts, simple for me, difficult for you. Simple ... the reason is ... I am a Hawaiian and I've inherited the soul of my kupuna. It is my moral responsibility to attempt an ending to this desecration of our sacred aina.; Kohe Malamalama o Kanaloa, for each bomb dropped adds further injury to my already wounded soul.	
The truth is, there is man and there is environment. One does not supersede [sic] the other. The breath in man is the breath of Papa (the earth). Man is merely caretaker of the land that maintains his life, and nourishes his soul. Therefore aina is sacred. The church of life is not in a building, it is the open sky, the surrounding ocean, the beautiful soil. My duty is to protect mother earth, who gives me life. And to give thanks with humility as well as ask forgiveness for the arrogance and insensitivity of man.	
What is national defense [sic] when what is being destroyed is the very thing the military is entrusted to defend, the sacred land of (Hawaii) America. The spirit of pride is left uncultivated, without truth and without meaning for keiki o ka aina, cut off from the land as a fetus is cut off from his mother. National defense is indefensible in terms of the loss of pride for many of the citizens of Hawaii nei.	
Call me radical for I refuse to remain idle. I will not have the foreigner prostitute the soul of my being and I will not make a whore out of my soul (my culture). All the archaeological discoveries, incredibly, are not enough it seems for the federal government to respect the sacredness of history. This continued disregard of our seriousness, this refusal to give credibility to the Hawaiian culture based on aloha 'āina, forces me to protest Aloha 'Āina	
E mau ko kakou lahui, e hoomau Hooulu ka pono o ka 'āina e Hooulu Hoola ka nani o ka 'āina e Hoola E MAU KE EA O KA 'āina I KA PONO /s/George	

the West cannot decipher, cannot understand and cannot control" (2021: p. 84). Similarly, Antonia Darder (2016: p. 8) argues that spiritual resistance creates conditions that "allow us to rethink spirituality, in our quest for freedom" (2016: p. 2).

In Paragraph A, Helm underscores the diversity of ideas and ideologies necessary for resistance by stating, "I have my thoughts ... you have yours." But in Paragraphs C and D he also notes the inequity and dominance of the settler-state ideology of the West. In fact, in Paragraph B he shows the conflict between competing Indigenous and Western values of land ("āina is sacred") and explains that Western ideology is an obstacle to justice because of its "refusal to give credibility to the Hawaiian culture." Helm understood that Western knowledge had become a default approach to mainstream knowing. Apple describes this dominating process as

Table 2 Aloha 'āina analytical framework.

<i>Developmental parts</i>	<i>Connection to Helm's letter by paragraph (from Table 1 above)</i>
Foundational or Grounded Assumptions	• Expressed strongly in Paragraph A • Expressed generally in full letter
Principle 1—Sovereignty & Deoccupation (decolonization)	• Expressed in Paragraph A • Expressed in Paragraph C • Expressed in Paragraph D • Expressed in Paragraph E
Principle 2—Centering Genealogy & Nationality	• Expressed in Paragraph A • Expressed in Paragraph C • Expressed in Paragraph D
Principle 3—Kuleana, Accountability & Moral Obligation	• Expressed in Paragraph A • Expressed in Paragraph B • Expressed in Paragraph D
Principle 4—Truth Telling, Radicalism & Protest (community and liberation focused)	• Expressed in Paragraph A • Expressed in Paragraph C • Expressed in Paragraph D

“modes of determination” that help to normalize hegemony by reinforcing Western norms, rules and routines in everyday life ([Apple, 2019](#): p. 65). For Helm, the best way to shine light on Western domination was through native resistance and refusal.

Helm understood that Hawaiian knowledge is peripheral in the context of US hegemony. For example, in Paragraph D he critiques the discipline of anthropology and identifies the hypocrisy of the academy when evidence-based research defers to Western settler-colonial demands. Helm established his own form of guerrilla resistance by radically and spiritually speaking truth to power. Helm's actions challenged the US military industrial complex and exposed the damage caused by US occupation but also exposed himself to retaliation. The US and state governments considered Helm a threat because of his fearless truth telling and resistance that awakened the people to their humanity and encouraged future acts of liberation ([Freire, 2005](#): p. 141).

Thus, another key understanding of Aloha 'Āina Theory is the courage to face settler-colonial retaliation, gas lighting, and hostility whenever state power is challenged. Tuhiwai-Smith describes the cruel responses commonly deployed by settler-colonial institutions arguing that decolonizing practices in the academy are extremely messy and met with violence thus resigning many Indigenous academics to mimic the corrupted ideologies they once resisted (2021: p. 288). Resistance will no doubt require courage and the willingness to remain uncompromising to activate meaningful change. Therefore, Aloha 'Āina Theory is intimately linked to, and guided by, a deep and abiding spiritual and cultural form of resistance and refusal.

The principle of sovereignty and de-occupation

At the core of Helm's letter is an underlying argument that natives must control their lands, culture, and politics. Specifically, Helm uses key phrases and terms of native sovereignty. Perhaps the clearest indication is found at Paragraph E where he quotes Alvan Issacs Sr.'s 'āina song, *E Mau ko Kakou Lahui* ([Ritte and Sawyer, 1978](#)). The verse starts with the line “E mau ko kakou lahui e ho'o-mau” and ends with “E mau ke ea o ka 'āina i ka pono.” Both phrases exalt the life of the Hawaiian nation and prays for its continued existence. Helm is asserting a position of anti-hegemony and identifying the US as a nation bent on maintaining its economic and political power over Hawaiian lands and people through the threat of violent military force and equally oppressive civil and criminal laws. In Paragraph C Helm argues “What is national defense [sic] when what is being destroyed is the very thing the military is entrusted to defend, the sacred land of (Hawaii) America.”

Helm also identifies the academy as a problematic state apparatus empowered by the US occupation to reify settler-colonial dominant ideologies including determining and redefining Hawaiian values of sacredness and culture. Linda Tuhiwai-Smith has made this point abundantly clear: “To acquiesce is to lose ourselves entirely and implicitly agree with all that has been said about us. To resist is to retrench in the margins, retrieve “what we were and remake ourselves.” The past, our stories local and global, the present, our communities, cultures, languages, and social practices—all may be spaces of marginalization, but they have also become spaces of resistance and hope” ([Tuhiwai-Smith, 2021](#): p. 4).

G. Smith notes that the principle of self-determination is essential for Indigenous academics to gain greater autonomy over key decision-making in schooling. Similarly, Helm is imploring Hawaiians (and other Indigenous and marginalized peoples) to encounter the state system to ensure the growth of more creative spaces that can reinforce native values and rights. However, Helm is also arguing that the settler-occupier's refusal to yield to Hawaiian concerns will require a more radical Indigenous

approach. Aloha 'āina demands more than seats on decision making bodies controlled by a hegemonic apparatus. Thus, one important principle of Aloha 'Āina Theory is that sovereignty and de-occupation are necessary and non-negotiable.

The principle of centering genealogy and nationality

Helm immediately centers Hawaiian knowledge and community as a core tenet. His positionality is stated directly through various cultural concepts such as, "I am a Hawaiian and I've inherited the soul of my kupuna" (Paragraph A), "the breath in man is the breath of Papa" (Paragraph B), "keiki o ka 'āina" (Paragraph C), "I will not prostitute the soul of my being and I will not make a whore out of my soul (my culture)" (Paragraph D), and "this refusal to give credibility to the Hawaiian culture based on aloha aina" (Paragraph D). Helm is clearly identifying a key principle—his genealogical relationship to the Hawaiian nation, land, gods, and people (community).

Noenoe Silva outlines a powerful methodology that ties the Hawaiian academic to their ancestors and gods. Silva's development of "mo'okū'auhau consciousness" describes an ethic and orientation to Hawaiian knowledge and world views. She critically asserts that recalling ancestral knowledge is part of a traditional and customary ethical duty to empower present and future generations (Silva, 2017: p. 7). Helm anchors his principle of aloha 'āina to ancestral knowledge ensuring that his own transformative strategies are accountable to his kupuna and the unborn generations who are yet to join the struggle.

The genealogy of Hawaiian nationalism and the Hawaiian state is a crucial tenet that identifies the unique political struggle of Hawaiians under US occupation. Sai provides a thorough examination of Hawaiian state sovereignty under international law and proves that "the Hawaiian Kingdom was not claiming to be a new State but rather exists as an independent State since the 19th century" (Sai, 2020: p. 25). He further establishes that the Hawaiian Kingdom is a government-less sovereign nation under prolonged occupation by the United States and that "in the absence of any evidence to the contrary, the United States has continuously violated Hawaiian sovereignty and its neutrality in many major conflicts to date" (Sai, 2020: p. 120).

Indigenous genealogical and national identity help native academics to establish clearer focus on the actions needed to educate, transform, resist, and refuse settler-colonial policies. Aloha 'Āina Theory creates a specific set of obligations that motivate direct action via Indigenous genealogy and sovereignty.

The principle of kuleana and moral obligations

Helm succinctly outlines the moral obligation and humanity that dictates movement, action and decision making for Aloha 'Āina Theory and Guerrilla Praxis. He asserts that wherever there is "desecration" (Paragraph A), "arrogance" (Paragraph B), hewa, or risk of life (Paragraph D), aloha 'āina imposes a duty of justice-oriented acts of resistance and refusal (Paragraphs A, B and D). In the movement to end the US military desecration of Kahoolawe island, Helm proclaimed: "it is my moral responsibility to attempt an ending to this desecration of our sacred aina, Kohe Malamalama o Kanaloa, for each bomb dropped adds further injury to an already wounded soul" (Paragraph A). Thus, a major understanding of this aloha 'āina principle is that all cultural practices, lands, people, and sovereignty are presumed sacred and therefore any act of hostility or threat against the Hawaiian culture, lands, people, and nation is a form of desecration. The duty of those engaging in Aloha 'Āina Theory requires an unflinching commitment to protecting and perpetuating the sanctity of the nation and community in total.

Kuleana or responsibility is a well-developed traditional concept. Goodyear-Ka'ōpua (2011) has articulated an approach that combines liberation pedagogy and anarcho-indigeneity with traditional and customary Hawaiian practices to support Hawaiian political movements. She describes kuleana lahui as "Indigenous concepts for thinking about and practicing collective autonomy" (2011: p. 1). Goodyear-Kaopua argues that patriotic lessons of Hawaiian activism strengthen future obligations to sovereignty and facilitates more innovative approaches that will improve Hawaiian identity and politics (2011: p. 4).

Hawaiian language scholar Warner (1999) expressed kuleana as a critical customary practice of empowerment and self-determination. His work demands that Indigenous people not compromise their sovereign responsibility for nominal improvements that do not challenge the status quo (Warner, 1999: p. 78). Kanalu Young further teaches that kuleana requires the "fortitude and resolve to do absolutely everything that is within our power and it all begins with determining for yourself the kuleana you are destined to fill on behalf of your country" (2006: p. 7).

The practice of kuleana is not fickle and is extremely action oriented. However, it requires an understanding of the moral compass that guides the native cultural and political wellbeing. Kuleana also requires a level of truthfulness that will allow critical critiques of power and position when concerns are raised regarding duty and responsibility of natives or allies.

The Principle of truth telling, radicalism and protest (community & liberation driven)

The fourth aloha 'āina principle discussed in this chapter focuses on Helm's methodical form of truth telling expressed through a promise of radicalism and protest in defiance of state power (Paragraphs A and D). The promise of direct action comes with a humble recognition of our collective humanity (Paragraph A and B), and of the existential threat that US hegemony poses to Hawaiian culture and nationhood (Paragraph C and D). Helm points to the US's state sanctioned settler-colonial forms of violence and erasure (Paragraph A and D). Helm is posing a radical, community driven solution to help people overcome their "fear of freedom" (Freire, 2005: p. 46); a fear that has been "prescribed" over time to produce native assimilation, colonization, and prolonged settler-state occupation. Helm's truth telling is grounded in cultural knowledge and expressed powerfully in Paragraph D: "Call me radical for I refuse to remain idle. The spirit of pride is left uncultivated, without truth and without meaning for keiki

o ka 'āina, cut off from the land as a fetus is cut off from its mother ... This continued disregard of our seriousness, this refusal to give credibility to the Hawaiian culture based on aloha 'āina, forces me to protest."

Freire notes that the oppressed have the kuleana to liberate themselves and their oppressor as well. Helms truth telling and refusal effectively reached state and federal decision makers and led to emancipatory action that ended the bombing. Helm appeals to the occupier's humanity but does so with a biting rendition of the truth. For the Kahoolawe movement, the truth telling, unrelenting radical protest and activism, and appeals to human rights, weakened the militaristic, violent position of the state apparatus over time. The presidential executive order that enabled the seizure and bombing of Kahoolawe was rescinded and the island was returned to the state in trust for a Hawaiian nation (HI Rev Stat §6K). Aloha 'āina activists who fought for Kaho'olawe, understood that their radical truth telling had very real human costs. Yet they willingly sacrificed their lives and livelihoods in pursuit of aloha 'āina led justice.

Activist scholarship and liberation pedagogies in Hawai'i grew because of the Kahoolawe issue and other Hawaiian movement actions that occurred over the years. Valuable lessons were learned on both sides. As Smith (1997, 2004) warns, the liberation work of today must adapt to the clever shift in hegemonic formations of oppression to ensure a lasting presence of native resistance and refusal. And, in that adjustment, Aloha 'Āina Theory and Guerrilla Praxis implores the activist scholar to reach for tools that challenge, rather than compromise, the settler-colonial educational systems.

Freire describes emancipatory actions through the analogy of "working" human hands that can transform the world. His analogy also denounces the dehumanizing "supplicating" hands that lead to further domination (2005: p. 45). Summer Maunakea (2019) explains how her kupuna Katherine Maunakea's "ne huli ka lima i lalo" Hawaiian culture lessons influence 'āina based pedagogies that lead to greater consciousness and transformative actions in the Hawaiian community (2019: p. 9). Maunakea notes that "When your hands are turned up you have an empty stomach, when your hands are turned down you have a full stomach" which conveys the Hawaiian concept of agency and transformative change (2019: p. 12). According to Freire, transformative action must come "from the oppressed themselves and from those who are truly solidary [sic] with them" (2005: p. 45). The oppressed, says Freire, understand the necessity of liberation but he also warns that they "will not gain this liberation by chance but through the praxis of their quest for it, through their recognition of the necessity to fight for it. And this fight, because of the purpose given it by the oppressed, will constitute an act of love opposing the lovelessness which lies at the heart of the oppressors' violence" (Freire, 2005: p. 45). Like Nāwahī's 19th century statement on aloha 'āina, Freire's idea of praxis is unvarnished and confrontational. Fanon (1963), sums up the necessity of a guerrilla-oriented praxis by stating that, "whatever may be the headings used or the new formulas introduced decolonization is always a violent phenomenon" (1963: p. 35). Thus, current efforts to indigenize the university through manipulative policies, native slogans or mainstreaming definitions should be resisted and refused as strategies of domination, not pedagogies of liberation.

In Indigenous/native nations and communities, an Aloha 'Āina Theory and Guerrilla Praxis is critically needed to unravel the evolving, complex tentacles of hegemony that instigate varying degrees of fear and ambiguity among the oppressed. In many ways, an Aloha 'Āina Theory and Guerrilla Praxis creates space for humanity and love to be fully applied as a legitimate political force for conscientization and transformation (Darder, 2016: p. 6). Darder declares that "it is precisely when the spiritual solidarity of a people is asserted organically in response to collective suffering and in the name of revolutionary action ... that political grace unfolds to support acts of resistance and dissent" (2016: p. 5).

The Principle of Truth Telling, Radicalism and Protest is a pedagogical element of aloha 'āina. The theory demands direct, radical actions of resistance and love by and for the community. Helm's endearing love for his nation, people and 'āina compelled him to act decisively against US hegemony. As G. Smith notes, a transformative process is not linear but fluid, allowing for those who are conscientized, like Helm, to organize direct action while simultaneously conscientizing others by exposing them to radical protest and activism. Meaning that a truly revolutionary praxis requires "action" and "critical reflection" occurring simultaneously (Freire, 2005: p. 128). The Guerrilla Praxis of Aloha 'Āina Theory is influenced by Helm's commitment to ending desecration through radical acts of love, resistance and protest all culminating in the refusal of US occupation.

Conclusion

Academia's sanctioning of neoliberal policies to "indigenize the university" is a direct assault on Indigenous knowledge and threatens efforts for native communities to pursue emancipatory pedagogies. The academic struggles described in the UH Indigenizing policy example above is an Indigenous call to action. Neoliberal and neocolonial policies can be hidden in plain sight and will be the kind of battles Indigenous, native, and marginalized peoples will experience in the coming years as more universities push to become "indigenized" despite lacking the community authority to do so.

For Indigenous and marginalized scholars, a pedagogy of Aloha 'Āina Theory and Guerrilla Praxis may prove useful in the struggle to save Indigenous knowledge from becoming a tokenized cultural fixture in the university's hegemonic status quo. Such a pedagogy may prove effective in challenging current university policies and weaken the deployment of tokenized academics. A Guerrilla Praxis can also empower outspoken native scholars to better manage the retaliation and hostility likely to occur during the struggle. Academics and community members must therefore muster the courage to speak a radical truth anchored in mo'oku'auhau and sovereignty in the face of the academy's settler-colonial aggression.

Academics must be proactive in truth telling that will respect a university administration's humanity while publicly exposing their extraction of Indigenous knowledge. Engagement with native communities and direct action in the form of radical resistance and refusal are inseparable elements to Aloha 'Āina Theory and Guerrilla Praxis. While change is always a possibility, it does not come without struggle. Many in the academy have attempted to use settler-colonial extraction devices, like "Indigenizing the university," to claim that change is coming soon, its arrival, however, is like "Waiting for Godot" (Beckett, 1982). Therefore, incremental successes celebrated by the academy are not the functional equivalent to actual emancipatory and transformative change.

Eager to build a stronger theoretical and praxis-oriented approach to reengage the academy on native terms, this chapter recommends development of an Aloha 'Āina Theory and Guerrilla Praxis. Aloha 'āina as an ideology is already well developed and soundly established through traditional and customary practices. Aloha 'āina also has a continuity that reinforces the sovereignty and cultural integrity of Indigenous and native communities in the present century. Most importantly, aloha 'āina is and has always been influenced, motivated and driven by the sovereignty of the nation and the native community. In the academy, aloha 'āina as a pedagogy requires an approach that can only be described as guerrilla-oriented because it is expected to undergo multiple changes to meet the morphing of settler-colonialism and to avoid being absorbed into the dominant, mainstream structures of the academy. The salient facts of this chapter tell us that, until the university is truly divested of its hegemonic ideology and structure—until there is sovereignty and de-occupation for native nations—resistance to settler-state hegemony in the academy will be a constant and will therefore require a deliberate approach that is aloha 'āina in theory and guerrilla in application. As Helm has taught us, Indigenous academics must always be radical and refuse to remain idle.

References

- Alfred, T., 2004. Warrior scholarship: seeing the university as a ground of contention. In: *Indigenizing the Academy: Transforming Scholarship and Empowering Communities*, p. 88.
- Andrews, L., 2003. A Dictionary of the Hawaiian Language. Island Heritage Publishing.
- Antoine, D., 2017. Pushing the academy: the need for decolonizing research. *Can. J. Commun.* 42 (1), 113–120.
- Apple, M.W., 2019. *Ideology and Curriculum*, fourth ed. Routledge.
- Baker, M.L., 2018. Ho'oulu 'Āina: Embodied Aloha 'Āina Enacting Indigenous Futurities. PhD Dissertation ProQuest No. 10992785.
- Beckett, S., 1982. *Waiting for Godot: A Tragicomedy in 2 Acts*. Grove Press.
- Benham, M.K.P., Stein, W.J. (Eds.), 2003. *The Renaissance of American Indian Higher Education: Capturing the Dream*. Lawrence Elbaum Association, Inc.
- Bobo, K., Kendall, J., Max, S., 2001. *Organizing for Social Change: Midwest Academy Manual for Social Activists (Third)*. Seven Locks Press.
- Brayboy, B.M.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37 (5), 425–446.
- Brayboy, B.M.J., Solyom, J.A., Castagno, A.E., 2015. Indigenous peoples in higher education. *J. Am. Indian Educ.* 54 (1), 154–186.
- Darder, A., 2016. Political grace and revolutionary critical pedagogy. *Rhizome Freirean* 21, 1–9.
- de Leeuw, S., Greenwood, M., Lindsay, N., 2013. Troubling good intentions. *Settl. Colon. Stud.* 3 (3–4), 381–394. <https://doi.org/10.1080/2201473X.2013.810694>.
- Fanon, F., 1963. *The Wretched of the Earth*. Grove Press.
- Freire, P., 2005. *Pedagogy of the Oppressed*. Routledge.
- Goodyear-Ka'ōpua, N., 2011. Kuleana lāhui: collective responsibility for Hawaiian nationhood in activists' praxis. *Affinities: A Journal of Radical Theory, Culture, and Action* 5 (1), 130–163.
- Goodyear-Ka'ōpua, N., Kuwada, B.K., 2018. Making 'āha: independent Hawaiian pasts, presents & futures. *Daedalus* 147 (2), 49–59.
- Gorrie, R., Mahood, A., Ashley, S., 2017. *Indigenizing the Academy*. Capilano University.
- Grande, S., 2018. Refusing the university. In: Tuck, E., Yang, W.K. (Eds.), *Toward What Justice? Describing Diverse Dreams of Justice in Education*. Routledge.
- Helm, G., January 30, 1977. Reasons for Fourth Occupation of Kahoolawe. Personal communication.
- Helm, G., 1984. Reasons for fourth occupation of Kahoolawe. In: *Ho'ihō'i Hou: A Tribute to George Helm and Kimo Mitchell*. Bamboo Ridge Press.
- Hill, E., 2012. A critique of the call to "Always Indigenize!". In: Mitchell, L. (Ed.), *Peninsula A Journal of Relational Politics* 2 (1).
- St. James, E., 2022. The Rise of Land Acknowledgements—and Their Limitations. *Vox*.
- hooks, b., 1990. *Yearning: Race, Gender, and Cultural Politics*. South End Press.
- Ke aloha 'āina; he aha ia?—Ka'iwakiloumoku—Hawaiian Cultural Center, 2021, July 7. <https://kaiwakiloumoku.ksbe.edu/article/essays-ke-aloha-aina-heaha-ia>.
- Keawe, L.O., 2008. *ki'i papalua: Imagery and Colonialism in Hawai'i*. Doctoral Dissertation University of Hawai'i at Mānoa.
- Kelly, A.K., 2014. Portrait. marie beltran and annie pau: resistance to empire, erasure, and selling out. In: *A Nation Rising*. Duke University Press, pp. 35–47.
- Lorde, A., 2012. *Sister Outsider: Essays and Speeches*. Crossing Press. Cross Press.
- Mahuika, R., 2008. Kaupapa Māori theory is critical and anti-colonial. *MAI Review* 3 (4), 1–16.
- Maunakea, S.A., 2019. Nē Huli ka Lima i Lalo Piha ka 'Ōpū: 'Ōiwi Agency and Outcomes of 'Āina-Based Education. PhD Dissertation.
- McLaren, P., 2015. *Pedagogy of Insurrection: From Resurrection to Revolution*. Peter Lang.
- Mihesuah, D.A., Wilson, A.C. (Eds.), 2004. *Indigenizing the Academy: Transforming Scholarship and Empowering Communities*. University of Nebraska Press.
- Native Hawaiian Place of Learning Advancement Office, 2019. *Native Hawaiian Student Data Report 2019*. University of Hawai'i at Mānoa.
- Nāwahioalaniopu'u, J., May 25, 1895. Ke aloha 'āina: heaha ia? Ke Aloha Aina 8.
- Pukui, M.K., Elbert, S.H., 1986. *Hawaiian Dictionary: English-hawaiian, hawaiian-English (Revised and Enlarged Edition)*. University of Hawai'i Press.
- Ritte Jr., W., Sawyer, R., 1978. Ka Mana'o Aloha o Kaho'olawe. Aloha 'Āina o Na Kūpuna, Inc., Honolulu.
- A slippery path towards Hawaiian indigeneity: an analysis and comparison between Hawaiian state sovereignty and Hawaiian indigeneity and its use and practice in Hawai'i today. In: Sai, D.K. (Ed.), *J. Law Soc. Challenges* 10, 68.
- Sai, D.K., 2020. The Royal Commission of Inquiry: Investigating War Crimes and Human Rights Violations Committed in the Hawaiian Kingdom. Royal Commission of Inquiry.
- Sai-Dudoit, K., Tolentino, B.N., 2022. Aloha 'āina: from the historical record. *Hawaiian J. Law Polit.* 4, 5–27.
- Salazar, J.A., 2014. *Multicultural Settler Colonialism and Indigenous Struggle in Hawai'i: The Politics of Astronomy on Mauna A Wākea*. PhD Dissertation. UMI No. 3691724.
- Shanley, J., 2003. Limitations and alternatives to developing a tribally controlled college. In: Benham, M.K.P., Stein, W.J. (Eds.), *The Renaissance of American Indian Higher Education: Capturing a Dream*. Lawrence Elbaum Associates, Inc.
- Silva, N.K., 2002. Translations of articles from the Hawaiian nationalist newspaper ke aloha aina. 'Ōiwi 2, 119–138.
- Silva, N.K., 2017. *The Power of the Steel-Tipped Pen: Reconstructing Native Hawaiian Intellectual History*. Duke University Press.
- Smith, G.H., 1997. *The Development of kaupapa Māori: Theory and Praxis* [Doctoral Dissertation], ResearchSpace@Auckland.
- Smith, G.H., 2004. Mai i te maramatanga, ki te putanga mai o te tauritanga: from conscientization to transformation. *Indig. Educ.* 37 (1), 46–52.

- Smith, A., 2009. Native studies and critical pedagogy: beyond the academic-industrial complex. In: Sudbury, J., Okazawa-Rey, M. (Eds.), *Activist Scholarship: Antiracism, Feminism, and Social Change*. Paradigm Publishers.
- Stein, W.J., 2003. Developmental action for implementing an Indigenous college: philosophical foundations and pragmatic steps. In: Benham, M.K.P., Stein, W.J. (Eds.), *The Renaissance of American Indian Higher Education: Capturing a Dream*. Lawrence Earlbaur Associates, Inc.
- Stewart-Ambo, T., Yang, K.W., 2021. Beyond land acknowledgement in settler institutions. *Soc. Text* 39 (1), 21–46.
- Sudbury, J., Okazawa-Rey, M. (Eds.), 2009. *Activist Scholarship: Antiracism, Feminism, and Social Change*. Paradigm Publishers.
- Thomas, A., 2021. Indigenous knowledge is not an extractable resource. *Academia Lett.* 2, 3832.
- Trask, H.-K., 1999. *From a Native Daughter: Colonialism and Sovereignty in hawaii*, revised ed. University of Hawaii Press.
- Tuhiwai-Smith, L., 2021. *Decolonizing Methodologies: Research and Indigenous Peoples*, third ed. Zed Books Limited.
- University of Hawai'i at Mānoa. 2021. *Mānoa 2025: Our Kuleana to Hawai'i & The World Strategic Plan 2015-2025*. Hawai'i: UH Mānoa.
- Warner, S.L.N.E, 1999. Kuleana: the right, responsibility, and authority of Indigenous Peoples to speak and make decisions for themselves in language and cultural revitalization. *Anthropol. Educ. Q.* 30 (1), 68–93.
- Young, K., 2006. Your Kuleana: toward a historiography of Hawaiian natinoal consciousness and the important question "so now what?". *Hawaiian Society of Law and Politics*.

A public good? The commodification of Māori language and culture in higher education

Huia T. Jahnke and Te Rina Warren, Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	81
Just whose public good is being served?	81
Who gets to speak te reo Māori - the Māori language?	83
Who gets to appropriate and commodify: the case of the academy	85
Who gets to speak freely	85
References	87

Introduction

"Māori have a long history and experience of schooling in a colonial and neo-colonial era and of dealing with and responding to the hegemonic consequences of the will of the State" (Tomlins-Jahnke, 2006, p. 3). The idea that tertiary education in New Zealand is a public good has been undermined by extreme policies that treat education as a commodity and the education sector as a market. Education is about producing highly skilled, knowledgeable and useful citizens and the notion of a "profit" is in "creating educated citizens" (Tertiary Education Union, 2018). While governments have promoted tertiary education as a core public good (The Ministry of Education, 2020), it is not what these same governments have emphasized. The government ignored the fact that most Māori enter tertiary as adult students because they are more likely to leave compulsory schooling without the necessary entry qualifications for university (Schulze and Green, 2017). As predicted, Māori access and participation has declined in recent years. So on the one hand the state promotes its commitment to Māori student success (The Ministry of Education, 2013, 2020) while implementing policies guaranteed to ensure barriers to Māori access and retention (Tomlins-Jahnke and Warren, 2011). Tertiary education is touted as a public good but ironically follows the trend established in the 19th century that perpetuates systemic Māori education failure within a neo-colonial education system (Penetito, 2010; Tomlins-Jahnke and Warren, 2011; Warren, 2013).

This chapter examines the notion of public good within Māori historical experiences of colonization and their effects on indigenous institutions and social systems through perpetuating ignorance as truth, myth creation (Hetaraka, 2022; Mikaere, 2011) or as Moana Jackson refers to as "myth-takes." The chapter interrogates the way in which tertiary institutions, supported by the State, exert their power and privilege to appropriate and commodify indigenous languages, culture and knowledge. This includes the power to justify and define Māori cultural terms and concepts in te reo Māori within institutional discourses that are superficial, cosmetic and are invariably aimed at maintaining the status quo.

Just whose public good is being served?

The notion of a public good is not foreign to Māori thinking. Customary Māori societies were founded upon the principle of collective wellbeing. In the early years of the 19th century and first contact with Pākehā (European) colonizers, Māori were intent on establishing new constitutional arrangements in the form of the Declaration of Independence (1835) 5 years earlier than the signing of the Treaty of Waitangi (1840) with the British Crown. The fact that Māori were considering such arrangements "... indicates an awareness of matters of national importance, transcending individual or even regional needs, and an expression of a form of public good on a national scale" (Piripi, 2011, p. 230).

Haami Piripi argues that within Māori social organization the notion of a public good is reflected in institutions like marae and tikanga (societal systems of ethics) which existed in the Pacific for thousands of years. The concept "mahī mō te katoa" meaning "for the benefit of all," was practiced customarily, and reflected in contemporary institutions of marae and tangihanga (ceremony of mourning) and in the gifting of land for public use (Piripi, 2011). Subsequently, early Māori business practice was founded on the principle of collective benefit, especially during the 1830s when Māori were major exporters in the southern hemisphere (Warren, 2009). This sort of "communism" as a public good contributed to settler opposition of Māori communal good. The individualization of benefits, including land title, became the foundation of the settler catch cry against perceived "Māori privilege" (Meihana, 2015) while ignoring the criminal and licentious activities of their fellow settlers. The perception that Māori calls for social justice, of righting past wrongs through equitable policies and law changes in order to mitigate the devastating impacts of colonization, somehow contributes to "Māori privilege" continues to persist within the national psyche.

When we consider the oppressive colonizing of indigenous worlds, invariably the colonizer asserts the superiority of their own political and constitutional structures as universal and normal thereby disregarding and replacing indigenous structures. This sentiment was openly stated in 1862 by colonial Inspector of Native Schools Henry Taylor in his report to parliament:

Corporal punishments and an over-rigid discipline have done much to drive away many children from the schools. A punishment, which to us would appear by no means harsh, would to a Native seem cruel and excessive. As Native parents never inflict chastisement upon an offending child, our summary mode of dealing with young delinquents must seem strange and tyrannical ... In carrying on the work of civilization among the aboriginal Native race, through the medium of schools, some impediments to progress, which may be gradually overcome by a diligent course of training... and the first and most serious of all is that state of communism in which all kinds of property are held amongst them. Their present social condition bears testimony to the ill-effects of such a system. Tribal rights destroyed personal ownership... few attempts have been made by the Natives to individualise property. In the school-room, by a careful and persevering system of appropriation we may gradually train them to a proper perception of and regard for the meum and tuum ... The Native language itself is also another obstacle in the way of civilization, so long as it exists there is a barrier to the free and unrestrained intercourse ... it shuts out the less civilized portion of the population from the benefits which intercourse with the more enlightened would confer.

Taylor (1862, p. 35).

In other words, and as Jackson asserts, "... what had been forged in the specific cultural and historical contexts of England and France suddenly became so 'culture free' and 'race-less' that a colonial governor could confidently assert that the English way of doing things simply existed in an 'independently supreme' condition" (Jackson, 2018). The Latin words *muem* (what is mine) and *tuum* (what is thine) re-enforces the colonial impulse to individualize ontological perspectives. The colonial settlers, supported by escalating immigrant numbers (Warren, 2017), simply exerted their European power and privilege over a declining native population in areas where it had not existed before. The racism inherent in that privileging was a prejudice pretending to be neutral.

Thus, self-serving fabrications about indigenous peoples buttressed by embedded racism are a cornerstone of colonial oppression, expansion and settlement and which remain a "normalized" presence within the bedrock of western societies (Jackson, 2019; Mutu, 2019). Untruths perpetuated about indigenous incapacity, about indigenous peoples as uncivilized, unintelligent and of a lower order of humankind, color coded and color charted as subspecies, are evident. Early government discourse on Māori children includes statements that reflect this very sentiment about Native Schooling when School Inspector Henry Taylor informed the Settler Parliament:

I do not advocate for the Natives under present circumstances a refined education or high mental culture: it would be inconsistent if we take account of the position they are likely to hold for many years to come in the social scale, and inappropriate if we remember that they are better calculated by nature to get their living by manual than by mental labour.

Taylor (1862, p. 38).

This mindset articulated by Taylor helped set the government's socially engineered education trajectory for Māori learners based on a differential curriculum whereby Māori boys were educated in agriculture to be farmers and Māori girls trained in domestic subjects as future wives of Māori farmers. Forty years later the same differential curriculum was reinforced in Parliament in reference to Hukarere Māori girls College and Te Aute College for Māori boys:

If ... it becomes necessary to choose between a course including Latin, geometry, and algebra in the higher forms, and scientific and technical subjects, preference be given to the latter.

The Commission on the Te Aute and Wanganui School Trusts (1906, p. v.).

Since the 1800's racist ideology embedded in Māori education policy is manifest for example in the profiling of indigenous people who "look like" criminals and the low education expectations of indigenous children who are considered less competent than the children of European settlers. As Moana Jackson says of racism "[a]t its most perverse, it corrupted reason to seek proof of the inferiority—and at its most dangerous, it stripped away human dignity with no thought for the hurt that it caused" (Jackson, 2018).

Being honest about the past and the present should be the first steps to settle injustice. Ignorance is a privilege that rarely extends to indigenous peoples working in higher education. Within the academy epistemologies of ignorance are repeated whereby non-indigenous try to be honest about the dishonest injustice of a racist history. But remedying the racism that continues to sustain the privileging of colonizing power requires even more honesty, as well as a certain courage (Jackson, 2018; Shaw, 2021). Pākehā educator Anne Milne, advocates for teaching pakeha/non-Māori children about their privilege;

Not to make them feel guilt—it is to make sure they do not grow up to perpetuate the situation. Knowing that our undeserved sense of entitlement and our unacknowledged privilege impact negatively on other people, and can be changed once we understand them, is a lesson our Pākehā children have the right to learn so they can contribute to a more equitable future and do better than we have done in the past. We are selling our Pākehā children short if we don't engage them in this learning

Milne (2017).

The history of Westminster education for Māori began with several Maori youth attending Rev Henry Williams classes in New South Wales (Jarvis, 1935), but the establishment of mission schools for Maori among Northern tribes confirms the racist and

oppressive nature of the colonial education system. Mission schools were established to encourage trade, agriculture and literacy in order to benefit the settler population in of Aotearoa. However, this imposed system

... aligned with international thinking and colonial practice in the nineteenth century ... aimed to assimilate indigenous people into European society. This assimilative ideology was clearly set out in the Native Trust Ordinance 1844, which provided mission schools with financial assistance and access to Māori land. The Education Ordinance 1847 established a set of education standards in religious instruction, industrial training, and instruction in the English language

Tomlins-Jahnke and Warren (2011, p. 22).

Following the establishment of a settler government in New Zealand, legislation was passed that embedded colonial ideals aimed at subjugating native children.¹ These laws worked in conjunction with a growing European immigrant population, and a dwindling Māori population, to disempower Māori through the education system where "... schools serve to act as key agents of cultural and ideological hegemony and of selective traditions ... what counts as school knowledge, the way school knowledge is organized ... the underlying codes that structure such knowledge access to and legitimation of school knowledge is determined by the dominant culture" (Tomlins-Jahnke, 2006, p. 4). Education as a "public good" was supported by Māori who willingly donated land and financed the establishment of schools in their districts. However within a short time Māori goodwill was undermined by the machinations of colonial power. Since then, the education system has maintained its original strategy of creating a Māori labor/underclass, where early strategies in education policy was aimed at social engineering in terms of a differential curriculum for Māori (Strong, 1931). As a public good Māori education was largely focused on a "civilizing mission" premised on the idea that it was necessary to remove the Māori child from the "demoralizing influence of the Māori villages" (Barrington, 1970, p. 28) in order to quickly assimilate Māori children into European ways (Barrington and Beaglehole, 1974). The long-term impacts on the inter-generational outcomes of Māori education remain extant. As Graham Smith reminds us:

The importance of examining Māori struggle within education and schooling, is that the effects of the educational schooling crises have a correlation within the wider social, economic, political and cultural positioning of Māori in New Zealand society as a whole. In this sense there is a need to understand why Māori remain marginal despite the apparent 'neutrality' of a schooling and education system that supposedly acts in the best interests of all, despite the 'fair and just' society created by the state.

Smith (2015, p. 25).

Who gets to speak te reo Māori - the Māori language?

The use of te reo Māori (the Māori language) in education faces a similar challenge. The 19th century mission schools were conducted in te reo Māori until colonial immigration numbers and legislation impacted te reo Māori usage. Māori children who spoke Māori within school precincts faced institutionally sanctioned violence when corporal punishment was introduced (Selby, 1999). The impending demise of te reo Māori and growing objections to the suppression of te reo was brought to a head in the 1970s when a group of Māori university students called Ngā Tama Toa, who were themselves denied opportunities to learn te reo Māori, petitioned government to reintroduce te reo into schools. In Wellington, the Te Reo Māori Society based in Victoria University led a national consciousness raising exercise supported by Ngā Tama Toa. According to Te Rito these students:

were first generation city-dwellers, their parents having migrated from their rural villages to the cities after World War II, as part of the mass urbanisation of the time. For some of these migrant families, physical detachment often led to a detachment from the culture and language of their rural beginnings. Many people of this generation of Māori bore an unwarranted shame that was not of their own making – that of not knowing their own native tongue.

Te Rito (2008, p. 2).

The raising of a national awareness on the dire state of te reo Māori led to a grass roots revolution and the emergence of Te Kohanga Reo (immersion pre-school) and the kura kaupapa Māori schooling system (immersion primary schools) where te reo Māori is the language of instruction and Māori knowledge, cultural values, philosophical traditions and tribal pedagogy are central. However, Te Kohanga Reo and Kura Kaupapa Māori are more than language revitalization projects. They are community spaces that Dr. Cathy Dewes describes as sites to rejuvenate the Māori spirit (Husband, 2018), and as sites of resistance² (Rawiri Wright, 2018). Te Kohanga reo and Kura Kaupapa Māori were established in the firm belief that with their language intact children will enjoy education success and flourish (Nepe, 1991; Smith, 1997). Forty years on, the assumptions of those early leaders proved correct when students of kaupapa Māori schools began outperforming students in English medium schools in National Certificate Education Achievement (NCEA) (Warren, 2013). As grassroots movements of collective action Te Kohanga Reo and Kura Kaupapa Māori spread rapidly across the country to effect change locally, regionally, and nationally despite the government denying these schools

¹For example: 1858 Native Schools Act, 1877 Education Act, 1880 Native Schols Act, 1894 School Attendance Act.

²Rawiri Wright, personal communication, January 29, 2018.

state funding and acknowledgment (Tomlins-Jahnke and Warren, 2011). It was only after the findings of the Waitangi Tribunal Te Reo claim of 1984 was issued did the government rescind and begin funding kohanga and kura (The Waitangi Tribunal, 1986).

Following the 1972 Te Reo Māori petition, Māori language week was instituted by government and is celebrated each year. More recently the focus on Māori language has extended to include Māori musicians whose productions have helped propel te reo Māori music into the public sphere. This newfound popularity in te reo inspired music, has led to several non-Māori “newcomers” to the te reo Māori music scene, including internationally acclaimed singer Lorde, receiving rave reviews for their contributions to te reo Māori. Support from renowned Māori language and music experts such as Sir Timoti Karetu and Dame Hinewehi Mohi (Hayden, 2021) served to add credibility to the newcomers and their efforts. In fact, the profile of te reo Māori music has come a long way since Dame Hinewehi was vilified in 1998 by the New Zealand public for performing the New Zealand national anthem in te reo Māori at an All-Black rugby game at Twickenham in England. By contrast, the recent foray into te reo Māori by non-Māori musicians did not go unnoticed by Māori associated with the industry (Dell, 2021; Findlay, 2021). For decades Māori artists struggled with few resources to bring te reo Māori into the public eye, and in this moment of international recognition of non-Māori performing in te reo, has served to undermine Māori struggle and reinforce white measures of acceptability and privileged airtime.

The government’s recent focus on te reo Māori as a public good, has resulted in unprecedented levels of state funding poured into professional development for all teachers and government employees (Maxwell, 2020; Te Puni Kōkiri, 2019) in a bid to grow the number of speakers of te reo Māori. This has included developing specific Māori language programmes for teaching a predominantly non-Māori workforce with allocated resources of over \$100 million for teachers (Maxwell, 2020) and \$33 million for government workers (Forrester, 2019) including promotions and additional remuneration based on language capability. While a mere \$6 million was allocated to Te Mātāwai the te reo Māori language board for language revitalization efforts among Māori. Meanwhile the Kura Kaupapa Māori sector remains seriously under resourced at every level such as capital works, employment wages, curriculum resources, professional development to name a few. A stark example of state negligence of the sector are the many Māori children in kura kaupapa experiencing their education in leaky buildings (O’Callaghan, 2020).

While the Government’s te reo Māori strategy is aimed at increasing the number of speakers of te reo Māori among New Zealanders generally, given the level of resourcing which benefits the dominant population, the plan is likely to widen the gap between the language *haves* (increasingly non-Māori) and the *have nots* (including many Māori) under the guise of a public good. The extraordinary uptake of learning te reo Māori is contributing to the growth of te reo Māori speakers. However, the growing number of Māori on waiting lists for te reo Māori courses (Tahana, 2021) offered by wānanga Māori³ raises the question about which “public” benefits?

The situation for Māori who are non-speakers of te reo was recently brought into sharp focus on social media that highlighted the shame and language trauma experienced by many generations of Māori who do not speak te reo (The Hui, 2022). The oft-repeated phrase that Māori language acquisition is up to each individual to commit to learning or not, overlooks the colonial history of the subjugation of te reo Māori where the choice by Māori to freely speak the language was removed and sanctioned by the state’s education policy until the latter years of the 20th century. In the main non-Māori learners are privileged beneficiaries of Māori loss and deprivation. Precious Clark cautions non-Māori learning about Māori language and aspects of culture:

... step onto the bridge, but remember, it’s our backbones you are walking across. Be guided by Māori in the use of our models such as pepeha to explore your identity and develop a deep connection of respect with this land and tangata whenua. Learn te reo Māori – after all, it’s beautiful! Move into te ao marama, and once you have learnt our culture, be an ally, a champion, an ambassador of Māori – being sure to avoid ‘weaponizing’ your knowledge of Māori culture against Māori.

Clark (2020).

Herein lies the perpetual challenge for many non-Māori to relinquish the colonial propensity to dominate, to show off their newfound knowledge at every opportunity and to consider themselves an expert in all things Māori.

This extends to the popular tendency by government, institutions and the academy to appropriate Māori terms often where such names bear no relation to Māori knowledge, custom or values. For example the name “kura auraki” delineates english medium schools from kura kaupapa Māori. What is problematic about appropriation is how epistemologies of ignorance of deep cultural knowledge is manifest and revealed by shallow application. The adoption of kura auraki for english medium schools is a case in point. According to Rawiri Wright “auraki” is a sacred ancestral ocean current on which Māori navigators traversed the pacific which highlights the flagrant contradiction of such naming. It is simply inappropriate to refer to a colonial system that continues to fail and disadvantage Māori children as *sacred*. In other words, the idea that a failing colonial education system is *sacred* for Māori children denies the reality of indigenous schooling where “... the dominant cultural and power inequalities remain: the mainstream system is well-funded, and the alternative sites are relatively under-resourced and marginalized within state policy” (Smith, 2017).

³Te Wananga o Raukawa, Te Wananga o Aotearoa, and Te Wananga o Awanui-a-Rangi.

Who gets to appropriate and commodify: the case of the academy

Commodification and appropriation of indigenous knowledge and culture by the academy has become popular and common practise worldwide. Such practices in Aotearoa New Zealand are subsumed within claims of upholding Treaty of Waitangi or Te Tiriti o Waitangi “principles” with a misguided assumption that the Māori version of the Treaty of Waitangi is therefore recognized. However applying the promises embedded in Te Tiriti (the te reo Māori text) such as sovereignty over lands and resources, customs, language, and self-governance are conveniently ignored, disregarded, and overlooked. Instead, appropriating indigenous names for all manner of things – buildings, university courses, groups, committees and so on at a deeply superficial level that when combined, commodifies the language as a branding exercise. The problem is intensified through the domestication of the indigenous individual whose consent is gained in the appropriation and commodification of cultural heritage within the academy.

In the main indigenous academics belong to disempowered communities both inside and outside of the academy. For example, Whare Wānanga⁴ are institutions of higher learning in the tradition of ancient schools of learning where customary traditions, knowledge, philosophy and theory draw from Te Ao Māori, the Māori world. While Whare Wānanga are denied the right to use the descriptor “university,” the government reserved the right for western universities to use “whare wānanga” where unlike whare wānanga, the teaching of disciplines of Māori language, knowledge and culture is largely limited to Māori studies departments (Wai-tangi Tribunal, 1999). For Māori, this is ironic as Linda Tuhiwai-Smith (2012) states:

By ‘naming the world’ people name their realities. For communities there are realities which can only be found, as self-evident concepts, in the indigenous language; they can never be captured by another language.

Tuhiwai-Smith (2012, p. 159).

Universities that co-opt whare wānanga as a descriptor invariably undermine the integrity embedded in the very institution of whare wānanga. For example charging exorbitant student fees thereby contributing to Māori student debt while employees receive free Māori language tuition. Similarly, attempts to insert Treaty obligations in strategic documents is most often conditioned by the full range of human expressions of prejudice, racism and ignorance laid out in all its naked contempt when indigenous priorities are privileged (Potter and Cooper, 2016). This is a battle that many Māori academics face when advocating for strategies to advance Māori education success, while simultaneously attempting to hold ground for all Māori within the academy.

Commodification of te reo Māori and tikanga Māori (Māori ethical protocols) is apparent in university policy for example when institutions (and Ministries) define Māori concepts such as “tikanga Māori,” how it will be considered and what funding will be allocated to whom. This process is reductionist and predicates further educational failure for Māori. The cosmetic application of te reo Māori is also evident in the vast translations undertaken by tertiary institutions of webpages, staff profiles, staff job descriptions, policies, and strategies in a bid to “normalize” te reo Māori. This process often forges ahead with set targets in a bid to “tick the box” without consultation with Māori or full consideration of the broader implications of such policies. For example, one tertiary institution has over 240 policies and the first and only policy chosen for translation into te reo Māori is the student disciplinary policy. The assumption being that the only university policy te reo Māori speakers need to be aware of is the punitive policy for student misconduct, and that of all the policies it is this one which is mostly applicable to Māori, and particular te reo Māori speakers. The actions of the translator employed to translate the policy could be interpreted as having carried out lateral violence against te reo Māori students while fiscally benefitting from the action.

The academy has long appropriated Māori cultural practises of pōwhiri (welcome events), hui (meetings) and karakia (ceremonial prayer) for events such as graduations, or as token ritual practises like the opening of hui with karakia. In these moments, Māori exercise limited authority adding performative value to maintain the illusion of inclusivity for the public good. Traditionally, universities have the responsibility to be the conscience of society, and therefore a foremost focus should be on the authentic application of Māori cultural practises that is informed and respectful. This is the preferred position rather than as cosmetic illusions and/or weaponizing the language and culture for the benefit of the institution especially when it is from a position of wilful ignorance.

Who gets to speak freely

The issue of “free speech” is often employed to enable hate speech and mistruth about Māori. In recent years, this has been played out against a background of escalation of racist discourses manipulated by mainly right-wing advocates of free speech, and therefore justified as a basis for the provision of public platforms.

The denial of two Canadian white supremacists, Stefan Molyneux and Lauren Southern, to speak at a public venue in Auckland because of their extreme racist views resulted in the emergence of The Free Speech Coalition led by ex-politician and former leader of the New Zealand National Party Don Brash. The right-wing members “... are more about securing their right to reproduce hate speech against Māori and other minority groups or to defend those who do so ... than they are about ‘free speech’” (Pihamā, 2018). Central to the modus operandi of the Coalition are litigious threats of court action and defamation suits against their critics.

⁴Te Wānanga o Raukawa, Te Wānanga o Aotearoa, and Te Whare Wānanga o Awanui-a-Rangi.

Don Brash, a long-time advocate of racist views about Māori that border on hate-speech, called for publicly funded radio bosses to remove Māori language from the airwaves deeming the language “pointless” (Pihama, 2018).

Brash is also involved with the Hobsons Pledge group that mobilized a reasonably successful nation-wide tactic to veto Māori wards in city council elections which would guarantee Māori representation on local governments. The group takes its name from a statement the first Governor of New Zealand Lieutenant William Hobson is purported to have uttered. After signing the Treaty of Waitangi Hobson pledged, “he iwi tātou” meaning “we are one people” (Keenan, 2017). But for Don Brash and his supporters, one people simply means we the white people.

The venue on the premises of one university that was booked for Don Brash’s invitation to address students was canceled after someone threatened to show up with a firearm. As a result university management came in for widespread condemnation and claims, including from academics within the university, that Don Brash’s rights to Freedom of Speech were breached and, in so doing threatened Academic Freedom.⁵ Attempts to censure the Vice Chancellor were mounted against claims they had brought the university into disrepute. Management felt conflicted about Brash’s pending visit due in large part to what had transpired when the Hobson’s Pledge representatives spoke on campus and disrespected a local elder, a group of Māori faculty and students. Among issues that conflated the situation was an assertion of the denial of free speech. In fact, Brash received masses of media time and attention which subsequently led to an offer of an alternative venue in another institution where he quite freely vented his anti-Māori views. The members of Hobsons Pledge exercised their right to “free speech” while displaying overtly racist behavior. What this truth demonstrates is the power of white privilege and who gets to speak freely. There was a deafening silence emanating from some senior professors and the press when it emerged the local elder was disrespected and abused on university premises while formally welcoming Hobson’s Pledge in the Māori language. Using the notion of “free speech” as a vehicle for abuse to prevent Māori from speaking freely in the official language of our country is an aggressive and violent act of oppression.

Claims that heinous acts of terrorism visited New Zealand in Christchurch can be described as “this is not us” is simply not true. The history of colonization in New Zealand suggests otherwise, as does its enduring neo-colonial operations within the academy. As legal scholar Moana Jackson maintains, the idea that “this is not us” is about “... the telling of a past that has been misremembered, and expressions of the ‘us’ that ought to be, rather than the reality that the legacy of colonization has left” (Jackson, 2018). A reality that includes colonial military assaults on any Māori community who defended their lands and resources (O’Malley, 2019; O’Malley and Kidman, 2018). Jackson reminds us that it is important to acknowledge the past and the present because the massacres in Christchurch underpinned by ideologies of racism and white supremacy are manifestations of colonization and a

... founding presumption that the so-called white people in Europe were inherently superior to everyone else. Some of Europe’s greatest thinkers contributed to the development of this presumption, and it eventually encompassed everything from the superiority of their form of government to the greater reason of their minds and even the beauty of their bodies. They were merely warped fantasies posing as fact, but they were eventually learned as the “truths” that enabled Europeans to assert that they had the right to take over the lands, lives, and power of those they had decided were the “lesser breeds”.

Jackson (2018).

Māori intellectual thinking is not new or an outcome of colonization. Our rich traditions of research and knowledge creation has always existed long before white explorers and settlers set foot on our shores. In Aotearoa New Zealand “whare wānanga” are systems of education underpinned by ritualized and formalized forms of pedagogy that extended into all aspects of everyday life (Pihama, 2019). We agree with Leonie Pihama’s contention that ...

The active suppression of Māori knowledge in colonial legislation and through ideological warfare meant that much Māori knowledge has been actively attacked or forced “underground” (Pihama 2001). This suppression was instrumental in the development of the colonial education system that sought to take the place of Māori systems of knowledge transmission.⁶ The marginalisation of Māori knowledge and Māori pedagogy has meant that our learning and teaching processes have been denied to generations of Māori people. This has without doubt been the situation in the University system, both in academic teaching and research. Researching within institutional frameworks often means having to deal in a daily way with critical issues related to Māori research, being a Māori researcher, being a Māori academic, a Māori woman academic.

Pihama (2019, p. 64).

In conclusion, our experiences tend to mirror those of indigenous academics across the world – education, indigenous language struggles, appropriation, commodification, and free speech remain sites of significant struggle for indigenous populations. While we create and regain cultural space and autonomy it is imperative that “... at the same time we must deconstruct the colonial power structures that show no love for our language, our culture or us” (Warren, 2017, p. 361). As Moana Jackson reminds us, we understand that the very idea of racism has become an abstraction described in various ways as unconscious bias, casual racism or unintended prejudice. Such terms can have some value, but they can also excuse the perpetrator and fail to address the unarticulated

⁵<https://www.donbrash.com/elocal/free-speech-again/>.

⁶“In the Native Schooling system there was a clear intention that te reo Māori be removed as a medium of instruction and that in order to receive support for a Native School that there be a commitment made to English as the medium of instruction with te reo Māori being used purely as a means of facilitating the learning of English” (footnote in Pihama, 2019).

distress and anguish that racism causes. The challenge for the academy is how brave and honest it can be in realizing that the just exercise of power can never be based on essentially racist assumptions that are themselves unjust (Jackson, 2018).

The Treaty of Waitangi was signed in 1840 between representatives of the independent tribes of Aotearoa and the British Crown. It is the foundation upon which Crown Māori relationships were based, that gave permission for British citizens to enter and when New Zealand was established, first as a British colony and later as an independent nation. Signing the Treaty in 1840 was a response by Māori tribes to, among others, the promises that modernity, in the form of new technology, new knowledge, resources and life-style, offered. However the failure of the Crown to live up to its Treaty promises almost from the outset provides the root of racism in Aotearoa/New Zealand. Racist behavior and attitudes must be addressed as a critical foundation of “public good” for the nation - for Māori as tangata whenua (First peoples of the land), for Pākehā (settler descendants) and for all other manuhiri (visitors and residents) as tangata tiriti (settlers to New Zealand under the jurisdiction of Te Tiriti o Waitangi). As Mikaere maintains of the settler/Pākehā society,

[t]here is nowhere else in the world that one can be Pākehā. Whether the term remains forever linked to the shameful role of the oppressor or whether it can become a positive source of identity and pride is up to Pākehā themselves. All that is required from them is a leap of faith.

Mikaere (2004).

For indigenous peoples, public good often comes at a cost. It is a price we pay in the exercise of our right to exist as Māori and as citizens of Aotearoa New Zealand. We continue to speak truth to power simply because our truth matters as first peoples of the land, and in the interests of the public good.

References

- Barrington, J.M., 1970. A historical review of policies and provisions. In: Ewing, J., Shallcross, J. (Eds.), *An Introduction to Māori Education* (np). N.Z. University Press.
- Barrington, J.M., Beaglehole, T.H., 1974. *Māori Schools in a Changing Society: An Historical Review*. New Zealand Council for Educational Research, Wellington, NZ.
- Clark, P., July 6, 2020. Opinion: Explore Tikanga Māori but Remember, It's Our Backbones You Walk Across. Newshub. <https://www.newshub.co.nz/home/new-zealand/2020/07/opinion-explore-tikanga-m-ori-but-remember-it-s-our-backbones-you-walk-across.html>.
- Collins, C.S., Jun, A., 2017. *White Out: Understanding White Privilege and Dominance in the Modern Age*. Peter Lang, New York.
- Dell, K., September 10, 2021. Why I Don't Get Warm Fuzzies over Lorde's Te Reo Māori Album. Kupumamae [Blog]. <https://kupumamae.com/2021/09/10/why-i-dont-get-warm-fuzzies-over-lordes-te-reo-maori-album/>.
- Findlay, S., September 28, 2021. A double-edged sword. Why Everyone's Talking about Lorde's Te Reo Māori Album. MamaMia. <https://www.mamamia.com.au/lorde-Māori/>.
- Forrester, G., 2019. One Million Kiwis to Speak Te Reo: The Plan behind Revitalising the Māori Language. Stuff. <https://www.stuff.co.nz/national/114928434/one-million-kiwis-to-speak-te-reo-the-plan-behind-revitalising-the-mori-language>.
- Hayden, L., 2021. "I'm Beginning a Journey": The Inside Story of Lorde's Surprise Mini-album in Te Reo Māori. The Spinoff. <https://thespinoff.co.nz/atea/09-09-2021/lorde-interview-Māori-lyrics-solar-power>.
- Hetarakā, M., 2022. Myth-making: on-going impacts of historical education policy on beliefs about Māori in education. N. Z. J. Educ. Stud. <https://doi.org/10.1007/s40841-022-00257-0>.
- Husband, D., November 25, 2018. Cathy Dewes: We Have a Right to Be Māori in Our Own Land. E-Tangata. <https://e-tangata.co.nz/korero/cathy-dewes-we-have-a-right-to-be-Māori-in-our-own-land/>.
- Jackson, M., February 25, 2018. Understanding Racism in This Country. E-Tangata. <https://e-tangata.co.nz/comment-and-analysis/maoana-jackson-understanding-racism-in-this-country/>.
- Jackson, M., March 24, 2019. The Connection between White Supremacy and Colonisation. E-Tangata. https://e-tangata.co.nz/comment-and-analysis/the-connection-between-white-supremacy/?fbclid=IwAR2S4wOen3UQkPjsNu-AJeN_jRI18SQre23HSX6G7QOSBji6hkay3W3TeE.
- Jackson, M., 2020. Where to next? decolonisation and the stories in the land. In: Kiddle, R., Elkington, B., Jackson, M., Mercier, O.R., Ross, M., Smeaton, J., Thomas, A. (Eds.), *Imagining Decolonisation*. Bridget Williams Books, pp. 55–65.
- Jarvis, J., 1935. The development and the settlement of Parramatta. Parramatta Hist. Soc., J Proc. 78.
- Kawiti-Bishara, T.K., July 27, 2022. Academic Questions Why Te Matatini Isn't Receiving Equitable Funding in Performing Arts Sector. Newshub.
- Keenan, D., November 30, 2017. Debunking the “one people” Myth: A Historian on the Invention of Hobson's Pledge. The Spinoff. <https://thespinoff.co.nz/atea/30-11-2017/debunking-the-one-people-myth-a-historian-on-the-invention-of-hobsons-pledge/>.
- Maxwell, J., December 8, 2020. Government Launches \$108m Te Reo in Schools Plan, Calling for 40,000 Staff to Upskill. Stuff. <https://www.stuff.co.nz/national/politics/123628606/government-launches-108m-te-reo-in-schools-plan-calling-for-40000-staff-to-upskill>.
- Meihana, P.N., 2015. *The Paradox of Māori Privilege: Historical Constructions of Māori Privilege circa 1769 to 1940*. Massey University. PhD thesis.
- Miakiere, A., 2011. *Colonising Myths, Māori Realities: He Rukuruku Whakaaro*. Huia Publishers.
- Mikaere, A., October 31, 2004. Are We All New Zealanders Now? A Māori Response to the Pākehā Quest for Indigeneity. Bruce Jessen Memorial Lecture. <https://www.brucejessen.com/ani-mikaere-2004-are-we-all-new-zealanders-now-a-Māori-response-to-the-pakeha-quest-for-indigeneity/>.
- Milne, A., October 23, 2017. Who Should Learn Most about White Privilege – Māori Children or Pākehā Children? NZARE Ipu Kererū. <https://nzareblog.wordpress.com/2017/10/23/white-privilege/>.
- Mutu, M., Feb 7, 2019. Guest Blog: Margaret Mutu - Māori Constitutional Transformations. The Daily Blog.
- Nepe, T., 1991. *E hao nei e tenei reanga, te toi huarewa tipuna : Kaupapa Māori, an educational intervention system*. Masterate thesis. Auckland University.
- O'Callaghan, J., October 5, 2020. Ministry Does “about-turn” Promises to Rebuild Mouldy Leaky Christchurch Kura. Stuff. <https://www.stuff.co.nz/national/education/122980390/ministry-does-aboutturn-promises-to-rebuild-mouldy-leaky-christchurch-kura>.
- O'Malley, V., 2019. *The New Zealand Wars = Ngā Pakanga o Aotearoa*. Bridget Williams Books.
- O'Malley, V., Kidman, J., 2018. Settler colonial history, commemoration and white backlash: remembering the New Zealand Wars. *Settler Colonial Stud.* 8 (3), 298–313. <https://doi.org/10.1080/2201473X.2017.1279831>.
- Penetito, W., 2010. *What's Māori about Māori Education? the Struggle for a Meaningful Context*. Victoria University Press.
- Pihama, L., 2018. Kaupapa Māori as Transformative Indigenous Analysis [Blog] Hate speech is not Free speech. kaupapamaori.com.

- Pihama, L., 2019. Kaupapa Māori within the academy. In: Tomlins-Jahnke, H., Styres, S., Lilley, S., Zinga, D. (Eds.), *Indigenous Education*. University of Alberta Press, pp. 63–82.
- Piripi, H., 2011. Te Tiriti o Waitangi and the New Zealand public sector. In: Tawhai, V., Sharp, A. (Eds.), *Always Speaking. The Treaty of Waitangi and Public Policy*. Huia Publications, pp. 201–214.
- Potter, H., Cooper, L., 2016. Project Whitestreaming: A Report on the Generalising of Māori Specialist Staff Positions in the Tertiary Education Sector. Tertiary Education Union.
- Schulze, H., Green, S., 2017. Change Agenda: Income Equity for Māori. Māori Futures Collective, Tokona Te Raki.
- Selby, R., 1999. *Still Being Punished*. Huia Publishers.
- Shaw, R., 2021. A tale of two stories: unsettling a settler family's history in Aotearoa New Zealand. *Genealogy* 5 (26).
- Smith, G.H., 1997. The Development of Kaupapa Māori Theory and Praxis. Doctoral dissertation. Auckland University.
- Smith, G.H., 2015. The Dialectic Relation of Theory and Practice in the Development of Kaupapa Māori Praxis. *Kaupapa rangahau: a reader*. Te Kotahi Research Institute, pp. 17–28.
- Smith, G.H., 2017. Kaupapa Māori theory: indigenous transforming of education. In: Hoskins, T.K., Jones, A. (Eds.), *Critical Conversation in Kaupapa Māori*. Huia Publishers, pp. 70–81.
- Strong, T.B., 1931. The education of South sea Island natives. In: Jackson, P.M. (Ed.), *Māori and Education: Or the Education of Natives in New Zealand and its Dependencies*. Fergusson and Osborn Ltd., p. 192.
- Tahana, J., September 13, 2021. Te Wiki o Te Reo Māori, Māori Language Week Begins. Morning Report. <https://www.nz.co.nz/national/programmes/morningreport/audio/2018812072/te-wiki-o-te-reo-Māori-Māori-language-week-begins>.
- Taylor, H., 1862. Further report on native schools in the province of Auckland. Appendix J. House Representatives 33–39. E-04. <https://paperspast.natlib.govt.nz/parliamentary/AJHR1862-I.2.1.6.4/1>.
- Te Puni Kōkiri, 2019. *Maihi Karauna: The Crown's Strategy for Māori Language Revitalisation 2019-2023*. Author.
- Te Rito, J., 2008. Struggle for the Māori language: he whawhai mō te reo Māori. *MAI Rev.* 2 (6), 1–8. <http://ojs.review.mai.ac.nz/index.php/MR/article/view/164/169>.
- Tertiary Education Union, 2018. *Voices from Tertiary Education. Response to Korero Mātauranga*. Author.
- The Commission on the Te Aute and Wanganui School Trusts, 1906. *reportTe Aute and Wanganui School Trusts (Report and Evidence of the Royal Commission on the)*. Appendix to the Journals of the House of Representatives. Session II, G-05.
- The Hui, July 26, 2022. A Comedian Has Been Shocked by the Mixed Response to a Te Reo Meme He Posted on Social Media. In: *Comedian Kajun Brooking Shocked by Reaction to His Te Reo Meme*. Newshub.
- The Ministry of Education, 2013. *Ka Hikitia: Accelerating Success, 2013-2017*. Author.
- The Ministry of Education, 2020. *Tertiary Education Strategy (TES)*. Author.
- The Waitangi Tribunal, 1986. *Report of the Waitangi Tribunal on the Te Reo Māori Claim (Wai 11)*. Brookers.
- The Waitangi Tribunal, 1999. *The Wānanga Capital Establishment Report* (G. P. Print).
- Tomlins-Jahnke, H., 2006. The place of cultural standards in indigenous education. *MAI Rev.* 1 (1), 1–11. <http://review.mai.ac.nz/MR/article/download/100/100-540-1-PB.pdf>.
- Tomlins-Jahnke, H., Warren, K.T., 2011. Full, exclusive and undisturbed possession: Māori education and the Treaty. In: Tawhai, V.M.H., Gray-Sharp, K. (Eds.), *Always Speaking: The Treaty of Waitangi and Public Policy*. Huia Publishers, Wellington, NZ, pp. 21–33.
- Tuhiwai-Smith, L., 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books.
- Warren, K.T., 2009. *Once upon a Tikanga: A Literature Review of Early Māori Business Practise*. Te Au Rangahau, Māori Business Research Centre.
- Warren, K.F., 2013. The legacy of Māori education: a view in 2013. In: Black, T., Cunningham, C., Jahnke, B., Tomlins-Jahnke, H., Kingi, T.K. (Eds.), *Matariki: A Monograph Prepared by Te Mata o Te Tau the Academy for Māori Research and Scholarship. Te Mata o te Tau (The Academy for Māori Research and Scholarship)*, Palmerston North, pp. 15–31.
- Warren, K.T.R., 2017. Te reo māori: analysing the colonial history of the native Māori language of Aotearoa (New Zealand). *Revista Lingüística/Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro* 13 (1), 349–366. ISSN 2238-975X 1. <https://revistas.ufrj.br/index.php/rl>.

Can we transform the academy? Indigenous-centered programming at the Department of Indigenous Education at University of Victoria

Jean-Paul Restoule, Department of Indigenous Education, Faculty of Education, University of Victoria, Victoria, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

CELÄNEN I SNEPENEK ŁTE	90
Respecting honoring and living Indigenous ways of knowing and being	90
Letting the community lead	92
Indigenous language revitalization and partnering with community	93
Laddering programs: how it works in practice	93
Reconciliation and resurgence	94
Challenges	96
Looking forward	97
References	98

Universities, as socially progressive as they imagine themselves, remain conservative in how they maintain systems and notions of colonialism, racial and cultural superiority, and patriarchy. The publication and take-up of the Truth and Reconciliation Commission (TRC) of Canada's Calls to Action in 2015 have led to a years-long rush among Canadian universities to demonstrate how they are contributing to reconciliation with Indigenous peoples. Claims of decolonizing and indigenizing are made yet many question whether the actions match the rhetoric (e.g., Gaudry and Lorenz, 2018). While I believe the gains for Indigenous people that can be made at most universities are modest, it remains important to attempt to transform the ways in which learning is carried out (Andreotti et al., 2015). Each small decolonizing step, each small indigenizing step, is a demonstration of Indigenous resilience.

So the question remains, can colonial institutions like universities be transformed in ways that are meaningful for Indigenous people and communities? In this entry, I will reflect on my first 5 years (2017–22) as chair of one of the only Indigenous Education departments at a university in Canada. I hope this reflection provides some inspiration for other Indigenous academics to find ways to support Indigenous aspirations and futures in their own institutions.

First, it's important to interrogate what is meant when I say decolonization, indigenization, reconciliation and resurgence. Decolonization is the divesting of colonial power from social, cultural and political institutions (Smith, 1999). While decolonization formerly referred specifically to the handover of government, it's increasingly realized that decolonization must extend beyond the political to the psychological. Indigenous individuals themselves have a decolonizing process required to divest our own internalization of colonial thinking, being and doing inculcated through education and cognitive assimilation (Battiste, 1998). Decolonizing is a process, perhaps never an arrival, and this notion has been described as a cycle involving five stages by Laenui, 2000 consisting of the recovery, mourning, dreaming, commitment and action phases. In this article, I will show how we at the Indigenous Education Department (IED) at the University of Victoria have engaged in each of these processes at various times.

Next, the word Indigenization is sometimes used uncritically (and carelessly) interchangeably with decolonization. At other times, it's critically assessed as a move that appears to be taking action while leaving current power structures comfortably unchanged or maintaining status quo. Indigenization means to make something Indigenous or to become influenced by Indigenous worldviews, perspectives and ways of doing things so as to change fundamentally. Another word that has been associated with Indigenization both in celebratory and desultory ways is infusion. To infuse a practice with Indigenization could mean that it becomes so Indigenous it's altered entirely so its entire being is transformed. This would be the ideal. But others have looked at infusion as "add and stir" approaches that don't fundamentally change the base. Like flavoring water, it's still water, but it tastes different. In my thinking, I've often assumed we can't fully Indigenize an institution until it's been decolonized first. But I also think we need to be working on both processes concurrently. If we wait for a place to decolonize we might not have the knowledge or capacity to Indigenize unless we've been working from that end too. Another way to look at this dilemma is internal and external. I feel like decolonizing work is aimed at colonial institutions and practices like the universities that are outside of our full control and our being. But Indigenization is bringing ourselves, our bodies, our practices and ways of teaching and learning into all the spaces we occupy, including universities. When we bring our knowledges with us into these spaces we are Indigenizing.

Another word that gets tied up in Canadian university policies of decolonizing and indigenizing is reconciliation. The TRC defined reconciliation as "establishing and maintaining a mutually respectful relationship between Aboriginal and non-Aboriginal peoples in Canada. For that to happen, there has to be awareness of the past, an acknowledgment of the harm that has been inflicted, atonement for the causes, and action to change behavior" (pp. 6–7). I've often made the distinction between reconciliation and resurgence as the work non-Indigenous and Indigenous people have to focus on respectively. Reconciliation is the work non-Indigenous people have to do, particularly around learning the truths, acknowledging them, atoning and taking action to change those root causes that led to assimilatory and genocidal policies. When you look at what's involved according to the TRC, that work is decolonial. Getting at the root causes and changing them is decolonizing. If this work is carried out, it

will free up space and place for resurgence activities, which is more properly the work of Indigenous people and communities. What is resurgence? Simply put it's a bringing forward from dormancy traditional Indigenous practices of relationship to land and community (Corn tassel, 2012). It's a renewal of culture through daily practice of language, prayer, ceremony, and responsibility to land and life.

One more term requires defining before we continue: Indigenous education. One useful definition (May and Aikman, 2003) focusses on the sources and approaches and how it's part of a resurgence process:

Indigenous education specifically focuses on teaching Indigenous knowledge, models, methods, and content within formal or non-formal educational systems. The growing recognition and use of Indigenous education methods can be a response to the erosion and loss of Indigenous knowledge through the processes of colonialism, globalization, and modernity.

Archibald (2020) adds, "Canadian Indigenous education includes education for Indigenous learners at all levels and ages and learning about Indigenous peoples' history, cultures/knowledges, and languages for all learners in educational systems." In IED, this dual approach of education for Indigenous learners and education about Indigenous ways for all learners is reflected in our two main areas of work: Indigenous language revitalization and teacher education for all learners.

Kirkness and Barnhardt (1991) identify the 4 Rs of Indigenous higher education as respect, relevance, reciprocity, and responsibility. As I progress through examples of the work we are doing in IED, I will demonstrate how our approaches are embedded in the 4 Rs. I hope to show that our work is, at times, decolonizing, at times focused on reconciliation, often Indigenizing and always resurgent. It is an attempt to, as Andreotti et al. (2015) have described it, hack the system, so that Indigenous communities can flourish, even if the institution's reforms are described as decolonial but are in practice far less radical.

There are a number of features to highlight in our work that I believe are worth striving toward including partnering directly with Indigenous communities, privileging community voices, centering Indigenous knowledges, revitalizing Indigenous languages, sitting in the space of dissonance, respecting the collective over the individual, reciprocity and educating for reconciliation. There are many lessons we have learned along the way and by sharing these we can help others be wary of overload and burnout, tokenization and ghettoization, funding issues, and avoiding barriers from standardization.

ĆELÁÑEN I SNEPENEK ŁTE

Respecting honoring and living Indigenous ways of knowing and being

Smith (2003) has written about the need to center our own ways of knowing, doing and being so that we, rather than the colonizers, are setting the agenda. One of the ways IED ensures we are centering Indigenous ways is by having an Advisory Board of Indigenous community members. Further, we enter into partnerships with Indigenous communities to develop programming with them in their territories. We put our peoples' writings and recordings in our courses and teach from Indigenous approaches about Indigenous topics. In our strategic plan, our very first set of goals fall under a heading called ĆELÁÑEN I SNEPENEK ŁTE, which is a SENĆOŦEN teaching meaning respecting, honoring, and living Indigenous ways of knowing and being. Our mandate in IED as it appears in our strategic plan is:

We want students to see the Department of Indigenous Education as supportive partners in life-long learning and a caring, safe Indigenous-led environment that respects, honours and lives Indigenous ways of knowing and being. We care deeply about supporting students on their learning journey and respect the protocols and practices of local Indigenous peoples.

We define success as a wholistic student experience respecting Indigenous knowledge, responsible relationships, community reciprocity and cultural relevance. Success is striving for balance among the spiritual, emotional, physical and intellectual aspects of being and participating in activities whose Indigeneity is not only valued but honoured.

In McGregor's (2014) evaluation of IED's B.Ed. in Indigenous Language Revitalization (BEDILR), she noted that the content and approach of courses created safe[r] spaces for the Indigenous students taking the program.

"The Bachelor of Education in Indigenous Language Revitalization (BEDILR) at the University of Victoria was launched in 2011 with two main goals: (1) to address the need to revitalize highly endangered Indigenous languages, and (2) to train a new generation of language speakers to teach the languages to other members of their communities. The four-year ladder program provides a foundation in Indigenous language, and skills in language revitalization strategies, and elementary school teaching" (McGregor, 2014). "Graduates from the BEDILR program are eligible to be certified teachers in the Province of BC."

Because the students enter the BEDILR from the Diploma program, there is an effort to support them in acquiring credits that may have been missed in their previous education. Indeed, one of the successes made possible by our Indigenous Language Revitalization (ILR) programming, is the acquisition of postsecondary education and credentials for students who were marginalized in their education journeys before encountering ILR at UVIC. They may have felt pushed out of western schools before, left school early for a host of reasons, or are healing from a legacy of residential school as a direct survivor or intergenerationally. A safer space for learning is created within the ILR courses where the histories are understood, and where the knowledge and approaches of Indigenous communities are respected, honored and employed in the programming. Some of the prerequisites for the BEDILR are requirements of the Ministry of Education and IED will arrange to put on courses meeting these conditions for the students that need them.

Respect for Indigenous ways and the centering of local knowledge means that students who were not given a chance to succeed in other western schooling find themselves going much further than they expected. In an evaluation of the first two BEDILR cohorts, [McGregor \(2014\)](#) found,

Both BEDILR communities cited as a central success the number of students who have graduated with diplomas and degrees. Many students succeeded in their programs despite entering the program with little formal education and with many family and community responsibilities. For example, one student began the program with a grade 5 education and graduated with a Diploma in Indigenous language revitalization. A well-known community Elder graduated with a Bachelor's degree. Many students were parents, grandparents, aunts, uncles, and important figures in their community who took lead roles in cultural events and celebrations. In different ways, all these students inspired others in their student cohort and community with their desire and motivation to learn.

More recently, a student reported

I took the language revitalization program through UVic and I didn't expect to make it this far. I come from a construction background ... I felt a big accomplishment in myself and doing my part in educating the greater community of the real history of our people and the real history of Canada really because it was all taken away from us through the residential school and day school experiences.

IED and UATÉ (2022).

Many graduates of the BEDILR have become teachers of Indigenous languages in school settings. A significant number entered immersion schools and are actively contributing to the development of curriculum and policies to ensure a new generation of speakers. Some have even become instructors for IED of the next cohort in their community.

We also chose to center Indigenous ways in our departmental governance by introducing policies for decision making and tenure and promotion policies that stem from Indigenous traditions. One of the first policies I shepherded through the newly created Department was a consensus model for decision making. IED does not use the standard Roberts rules or similar guides for running our meetings. We discuss topics or issues, often in circle and when a decision needs to be made we try to arrive at a consensus. We have never pushed an item to a vote. Consensus is different from unanimity and fortunately our department decisions have often been unanimous.

IED used to follow its former home department's tenure and promotion policy. While there were some progressive elements such as recognition of community-based research activities, there remained limits to the understanding of Indigenous scholarship, particularly around collaborativity and the difference between service and scholarship. In my first years as Chair, I sought to make the promotion and tenure review process more inclusive of Indigenous scholarship that is community-based, focused on language learning, and represents the often hidden Indigenous labor taken on to support Indigenous students and do unacknowledged Equity, Diversity and Inclusion (EDI) work. In 2019, IED moved toward its own unit standard, defining what is considered the bar for teaching research and service. It laid out clear standards for promotion to associate and full professor and tied these to a list of recognizable Indigenous scholarly activities. Included among the scholarly activities are Indigenous language learning, support and mentorship of Indigenous students, contribution to EDI efforts at the university, and a much longer list of community-based activities. This excerpt from our policy demonstrates the significance of the undertaking:

While the Department of Indigenous Education (IED) acknowledges the important role of rewarding performance in research, teaching and service, we also seek to reward performance based on the strategic goals and mission of IED: to guide transformative relationships in teaching and learning and extend a circle of welcome and support that enables Indigenous cultures and languages to flourish, while working to ensure that Indigenous ways of knowing, learning, being and doing thrive.

... Thereby, the Department of Indigenous Education recognizes that our mission is enhanced by our faculty members participating in Indigenous community and grounding their work in Indigenous ways of knowing, being and doing. By engaging and participating in Indigenous community, faculty enhance the viability, reputation, experiential opportunities, knowledge and depth of the Department's research, teaching and service.

IED continues to exist within a Eurocentric institution of learning and knowledge production. Our scholarly activities are necessarily a syncretic blend of Indigenous Knowledge ways with conventional academic approaches.

IED acknowledges that in Indigenous contexts, teaching, research and service activities can overlap and must be viewed from a more wholistic lens. This reality is addressed in Appendix B: List of accepted Research, Scholarly and Creative Activities.

Another step our department has taken to contribute to Indigenous community development is advance a complement plan for faculty and staff that prioritizes the hiring of Indigenous people. We have engaged in limited and preferential hiring practices (limited meaning only Indigenous people are considered, and preferential meaning the pool is composed of two lists, a first list of only Indigenous candidates that is reviewed first and a second list with non-Indigenous people consulted only if the first list is exhausted and no one is successful). To further support Indigenous people, our unit uses Indigenous community spaces for events like our Advisory Board meetings and planning retreats, and hires Indigenous caterers when we can. In our community partnerships we hire local First Nations coordinators, and instructors.

We try to live our Indigenous teachings as much as we can within the confines of the colonial institution we find ourselves in. When we request Elder support for events we pay an appropriate honorarium and we take care of their needs such as food, parking and travel support. We are committed to engaging in traditional protocols as much as possible. For example, when one of our beloved faculty members passed away, we organized a memorial on campus following her Kwakiutl protocols but because we were on Lekwungen and W'SÁNEĆ territory, we opened following local protocols with a practice to then welcome Kwakiutl to hold their ceremony within the house. After, the space was returned to W'SÁNEĆ so they could close in their way. When one of our students injured themselves on campus we followed ceremonial protocol. We went to the big house on their territory with their families, colleagues, instructors and IED present and held a feast and gifted her a collection to help her through the inconveniences she suffered. Another example is the moving of our Grandmother XAXE SIÁM SILA, the woven artwork depicting 13 moons. We followed protocol when she needed to be moved from one location on campus to another. Staff bathed grandmother and repaired fibers, smudged her, and ceremonially carried her from the library to her new home with drumming, singing, and regalia. The artist who facilitated the course where she was born came to advise us on the move saying, "you wouldn't just take your grandmother out of her home without giving her a bath and making her look her best, would you?" Our Indigenous Epistemologies courses are held on the land with involvement of revered knowledge keepers. These are all efforts to center Indigenous knowledge, worldview, being and ways of doing, teaching and learning.

Letting the community lead

Centering Indigenous ways also means prioritizing community needs and being accountable to the lands and the people in long relationship to the land. Within IED we have a few ways of ensuring accountability to community relations and privileging community voices. One of these is the Advisory Board that is Indigenous led and composed primarily of Indigenous members, and a vital one is our community partnership process that lays the ground for all of our Indigenous language revitalization work.

Laenui, 2000 described the dreaming as one of the critical decolonization processes necessary to effect change in our communities. He describes coming to a commitment through dreaming and dreaming about the dreaming: "the full panorama of possibilities are expressed, considered through debate, consultation, and building dreams on further dreams which eventually become the flooring for the creation of a new social order" (p. 145). In IED's work I hope that we are part of the dreaming and the dreaming about the dreaming, and also the commitment and action that Laenui described. Our programming is about fulfilling a vision that was forged by many community voices coming together and imagining a future where our people spoke their languages and carried out their ceremonies following the seasonal calendar. To enable the language dreams of the community, programs were created that were designed in such a way that they would be inclusive of all learners and would contribute to the building of ever greater language and teaching capacity in the communities.

IED's external Advisory Board is made up of members from local First Nations, and Indigenous education partners and has been in existence since 1999 when IED was a fledgling program inside of another department. The original co-chairs, Janice Simcoe, Director of Eye? Sq̓'lewen of Camosun College, and Nella Nelson of Aboriginal Education, School District 61, while recently retired, continue to co-chair, and their long experience and wisdom is much valued. Three times a year the board meets with members of the Faculty of Education (most in attendance are from IED plus the Dean of Education). The community members in attendance provide updates on activities in their respective settings and how IED can support their work. IED provides updates on activities in our unit and in our programming and also brings questions to the board when some direction is sought. It is here that the board often speaks strongly and forcefully about the need and ways to increase the number of Indigenous teachers in the province, or how to build understanding among non-Indigenous teachers in the local school districts. This is an example of how IED takes direction from community on how to engage in programming that is relevant, reciprocal respectful and responsible (Kirkness and Barnhardt).

Indigenous language revitalization and partnering with community

UVic is known internationally for its Indigenous Language Revitalization work, thanks in no small part, to IED's Director of many years, Dr. Onowa McIvor whose work has been foundational. IED's ILR programming is most closely tied to community partnership, hosts Indigenous student cohorts and directly contributes to Indigenous community aspirations. There are many features about our ILR programs that put community first. We strive to serve Indigenous students, peoples and communities well, helping to turn around some of the barriers that prevent Indigenous people from accessing university, staying in programs and completing them. So many aspects of our ILR work directly benefit Indigenous community, such as being able to change policies to facilitate student access and participation, including learners who wouldn't normally get a chance to go to university, and contributing to community needs by teaching the teachers who will be certified language instructors in the local schools. When we enter into a partnership with a First Nation we rely on Elders, local language mentors and knowledge keepers and we base our learning in community as much as possible. In a recent video made to illuminate the impacts of our Indigenous language programming beyond language learning, one IED student said: "Our classes were always in the community which made us more comfortable. I had babies while I was in those courses ... they were on my hip, wrapped and sleeping on the table. When I had to present my classmates were holding my babies ... We've lost so many of our fluent speakers in the last 10 years ... but the beautiful parts that have blossomed out of this whole experience in the last 10 years ... so many of our cohort that have graduated are still out there teaching" (IED and UATĒ, 2022). Being in community and respecting community ways has helped to build capacity among current speakers to teach the next generation in immersion schools that did not exist when they were younger.

Another Diploma in ILR graduate observed, "These Elders that were previously advocating for this language became the teachers of this course and the professors that came from UVic were more like the facilitators and it made this a natural space for learning [the language]" (IED and UATĒ, 2022). Indeed, we know the language resides in the community and we rely on that knowledge as the basis for the programming we do. While we have built the programming to be responsive to the latest research and to address community needs and desires and we share that in our courses, we also step back and give the knowledge keepers and language mentors the space and time needed to prepare the next generation. Dr. Wanost's'a7 Lorna Williams, professor emeritus and the founder of many IED courses and language certificates, said, "Wherever this program has been I know it has made a huge difference. It just lifted the people. I see those young people now and they're the key movers and leaders in their communities. That's what language does" (IED and UATĒ, 2022).

All IED language programming (with the exception of an elective course, SENĆOŦEN for beginners, that is open to anyone on campus) is a partnership with community, whether it is a First Nation or a Language society controlled by Indigenous members. Instead of individuals applying to a program with a list of required and elective courses, a community inquires with IED about the possibility of a partnership to deliver a program. IED and the community partner converse and explore timelines and feasibility. After coming to an agreement, we enter into a partnership with the community selecting the students who will enter the program.

UVIC requires the learners to register as students. Early on it was decided that community members should not be denied access to learning their language because they lacked credentials. There is an arrangement within the university to allow exceptional entry for our certificate students who may not have a high school diploma or specific pre-requisite courses. This means no one is denied opportunity to learn their language because of a lack of academic preparation. The IED staff and the local community coordinators work together to encourage students to register assisting them with the forms and providing plenty of support. Furthermore, the Indigenous language programming is laddered to provide learners with confidence as they move through each year of credits. The ladder approach to programs responds to a concern Indigenous teachers and students have voiced for years about retention and completion. Laddering has the advantage of providing successful exit opportunities at multiple points so that students who need to stop are recognized for their achievement. Those who continue receive encouragement to keep going.

Laddering programs: how it works in practice

Year 1 is either a Certificate in ILR (CILR) offered in association with UVIC Continuing Studies or the Certificate in Indigenous Language Proficiency Building (CILP) through IED. IED's language programming is a partnership with Linguistics. In the earliest days Linguistics had the professorial capacity and offered several courses that could be offered as Linguistics or IED. As Linguistics professors retired or moved away from the unit, the pendulum has swayed so that the majority of the teaching and administration takes place within IED. Communities choose whether to enroll in the CILR or the CILP. The CILP was a pathway that existed when I started as Chair but was a modified CILR. It was created in partnership with Indigenous communities who noted that students wanted to go deeper into the language with the certificate and were less interested in a couple of the courses that were more theoretical or academic. With some tweaks to the courses included as requirements and the overall number of credits required, the CILP was proposed and approved as a credential in 2019 that strongly contributed to one's proficiency development. The courses were designed to be taken in a certain sequence that promoted ever deepening knowledge and skill.

At the completion of either certificate, graduates are eligible to enter into the Diploma in ILR. When they complete the Diploma after 1 year, they are likely eligible to enter the B.Ed. in ILR (more on the BEDILR in the next section). This is a 4 year B.Ed. that leads to a certificate to teach in BC. The first 2 years are the Certificate and Diploma. The following 2 years provide all the education courses to make one eligible to apply for a teaching credential. The idea, borne out of consultation with communities and advisory

board members, was that a program was needed to prepare teachers to teach the language and in the language in their communities upon completion.

The Certificate and the Diploma courses are taught entirely in community or close to community (not on campus). They are usually all one language dialect and there is considerable attempt to be on the land and learn from place and community as much as possible. There are courses that employ the Mentor-Apprentice Program or MAP where mentors are identified and speak with apprentices in settings where the language vocabulary and application is relevant. Mentors may meet on the land or waters or in someone's kitchen making a meal. The goal is to speak in the language the entire time without relying on English to fall back on. More hours and practice in the language lead to more proficiency and increasing fluency (Hinton et al., 2002, 2001; See also FPCC, 2022).

IED's graduate programming includes the Masters in Indigenous Language Revitalization. (MILR). This 2-year program for ILR leaders and practitioners is designed to build leadership skills in ILR and training in research that furthers language revitalization work. The MILR begins with four courses that packaged together serve as the Graduate Certificate in ILR. Students who move on complete additional courses culminating in a research project fulfilling the requirements for an M.Ed. or M.A. Students are encouraged to complete a project that will assist their community with some aspect of a language initiative. It could be a curriculum in language, or the establishment of a MAP handbook for community language leaders, or a study of silent speakers, or creating an Indigenous language immersion home just to name a few of the projects from the first 5 years of the department.

The MILR, like the undergraduate ILR programming, takes a cohort-based approach. The cohort provides a built-in support network as the whole group moves through the courses together building a community of learners who become strong support for one another emotionally, and culturally. Cohorts have been known to be effective in Indigenous student retention (Landertinger et al., 2021).

In addition, ILR programming is scheduled in intensive course offerings with 39 h over a week to 10 days. One of our CILP offerings was offered over the term 1 weekend per month, rather than 3 h a week for 12–13 weeks like most UVIC undergraduate courses. Intensives allow learners to balance their education with their work, family and community responsibilities. For the MILR students, rather than move to Victoria for 2 years, they can visit for short periods of time. The 2022 MILR will try something new, picked up after the 2020 cohort was forced by the pandemic to move online. The 2020 cohort reported that they preferred some of the courses to meet more frequently for shorter periods as they did when they met online weekly. For 2022, some of the MILR courses will be offered in a blended format where the cohort will take them online for several weeks, then come together in an intensive in-person session. It will allow them more time to digest material and longer time to complete some assignments compared to an all in-person intensive. This flexibility in course scheduling is another feature of ILR programming that is responsive to learner lifestyles and puts community and learner/family priorities first. MILR course sequencing and content is reviewed every 2–4 years in response to student and instructor feedback to ensure that the material is relevant and the pedagogical approaches are engaging, flexible and suited to student lives. Like all the ILR programming, its flexibility, relevance and power stems from strong partnership, trust, respect and commitment to collaboration that is fostered between IED and the community.

Reconciliation and resurgence

While the students taking our teacher education courses outside the BEDILR are primarily non-Indigenous, altering the way all teachers are trained will change how generations of Canadians are educated about Indigenous histories, perspectives and futures. Our advisory board has prioritized our contributions to changing teacher education for the betterment of our people (and all people). In these ways our efforts to transform teacher education with an Indigenous lens is an attempt to foster a cadre of professionals committed to reconciliation.

The first innovation to note is course IED 373, EL TINIWT. This is the required course for all teacher candidates at UVIC and IED provides the instruction and curriculum oversight for the course. This requirement was established in UVIC before the province moved to make it a requirement for all teacher education programs in BC, and after the TRC Calls to action brought similar moves across Canada. The course provides

[a]n exploration of learning and teaching approaches to better address the educational requirements of Indigenous children, youth, and adults through understanding Indigenous peoples' relationship with land, language and community. Topics include: study of who Indigenous peoples are, diversity amongst Indigenous peoples, ways to Indigenize the learning and educational environment, and of the systems that impact teachers and learners.

Students get an overview of residential schools, the 60s scoop, and Indian Control of Indian Education. They also get an introduction to local ways of teaching and learning and the First Peoples Principles of Learning (FNESC, n.d.). One of the encouraging things that is happening right now is that we are starting to see students arrive with a greater understanding from which we can scaffold. We are not starting from zero knowledge in our 373 any longer. In 2022, IED will overhaul the IED 373 curriculum and resources to go deeper with our students, as they already have some understanding and exposure to Indigenous education histories, partly as a result of high schools and universities responding to the Calls to Action of the TRC by covering residential schools in curriculum and offering more Indigenous courses. These moves are complemented by the introduction of new provincial policies

that require ever more content and approaches that emerge from Indigenous cultures and knowledge. One example is the new professional standard for educators introduced in 2019 requiring knowledge of Indigenous perspectives, histories and pedagogies (BCTC, 2019), and the recently announced Indigenous course requirement for all secondary graduates starting in 2024 (BC, 2022).

Another way we are weaving Indigenous perspectives into the curriculum is through our Summer Institute. Every June, the IED Summer Institute provides the UVIC Secondary program students with an option to take 4 IED courses in a single month. The courses usually consist of an Indigenous Education History course, one in Indigenous Pedagogies, one in Local Knowledges and another course that is so unique, a book has been written about it (Tanaka, 2016). The IED 473 course, (CENENITEL TW TOLNEW: Helping each other to learn), had its genesis as a way for Indigenous presence to be established in the Education building at UVIC. A guest artist is identified and co-teaches the course with a UVIC faculty person (largely in the role of administrative facilitator). An art project is chosen and the whole class work together to make the piece that will have a home in the building. Past projects, still visible in the education building, include the Old Man pole Schalay'nung Sxwey'ga, representing Thunderbird and whale as protection, and XAXE SIÁM SILA, the Grandmother Mural that used to be the focal piece of the curriculum library, since moved to the main lobby of the Education building in a ceremony in 2018, as well as a button blanket representing numerous Indigenous values. The course has also been used to gift songs to the School District for teachers to use with young students (the Earth Songs course), and a canoe project (the canoe was moved to T'SAWOUT when it started to dry out—it needed to be used in the water or it would split). The past few years have seen students in the Summer Institute guided by a renowned Coast Salish artist to make hand drums with a Coast Salish art design painted on the skin. When they learn that it is custom to gift their first drum to someone, the emotions are high and it is an excellent teaching opportunity. The 473 course has also been used for making paddles, learning the metaphor of paddling together and paddling forward, with the knowledge gained from the course and the whole set (it was taken by the working teachers to close off a 13 month learning period in Indigenous education).

Tanaka (2016) looked at how pre-service teachers were transformed by the 473 course *Earth Fibers* which immersed them in Indigenous ways of learning and teaching by working alongside Indigenous wisdom keepers. "Together, they transformed cedar bark, buckskin, and wool into a mural that tells stories about the land upon which the course took place. In the process, they discovered new ways of learning that support not only intellectual but also tactile, emotional, and spiritual forms of knowledge" p. 22. Tanaka (2016), identified a number of teachings that emerge from the way in which the 473 course is done that represent fundamental Indigenous education values. Included among her findings were: working/walking alongside; putting aside the familiar/connecting to the unknown; letting go of perfection; fear of engaging in emotional territory; learning is uncertain and requires an inquiry stance; appreciating spirit; knowing nature and place; reflexivity requires safety; showing is effective; learning in families through story; opening up to other ways of knowing; faith in the learner; getting out of head and into hearts and hands. Our 473 course embodies these pedagogies, as does our whole Summer Institute.

In addition to providing in-depth experiences for future teachers, IED runs a 10 course set for in-service teachers to help them more fully and confidently integrate Indigenous perspectives in their work. Studies with teachers have shown there are many reasons they need to build confidence to move through their hesitancy (Carroll et al., 2020; Bascuñan et al., 2022; Nardozi and Restoule, 2019). Completing the 10 courses gives teachers the eligibility for a pay upgrade in BC and greater depth in Indigenous approaches and values. The teachers move as a cohort through the 10 courses. We have also allowed learners to register in one or two courses as electives for their programs or to acquire credits they're missing to be eligible for the upgrade. The courses consist of 2 offerings of the 473 course with 2 different projects. In addition to courses on pedagogies, teaching worldview through media, literature, and environment, there are two courses that are taken back-to-back going deeply into the self in relation to Indigenous education. Called Allies and Activists and Foundations of Decolonization, the two courses are meant to change one's orientation toward land and people, understanding responsibilities and obligations to these relationships and determining how to recognize colonial and racist structures and to advocate and work toward their dismantling. The courses begin with self and seeing one's location as a source of power and agency in order to disrupt racism and colonialism inside structures, policies and practices. The ultimate goal is to make changes that better life for Indigenous learners and communities. These courses were newly created and added (in 2021) and it's early still to know what impact they will have on the teachers moving through them in 2023 but we remain hopeful that some will find this learning transformative.

In the special topics versions of these two courses that were offered online during the pandemic, students reported being completely transformed. It caused them to question assumptions they had made about Canadian land settlement, and what they knew about the Indian Act, reserves, treaties and Aboriginal title. It fostered strong connections among the learners in the cohort and commitment to calling out problematic practices in their school districts. It encouraged teachers to find ways to make relationships with local Indigenous communities and to let that relationship guide their practices and learning journeys. The Integrated Course Set is designed to give teachers an opportunity to learn from and about Indigenous knowledge, and to make valuable connections to local Indigenous communities to strengthen the learners' understanding of Indigenous peoples and societies. The 10 courses are offered in concert with each other resulting in an immersive experience in Indigenous learning, teaching and knowledge where in-service teachers gain valuable experience and make meaningful connections.

These courses are scheduled to be tailored to teachers' schedules. The intensives are in July when they are not teaching so they can travel to Victoria and take the 6 summer courses over 6 weeks. The online courses are scheduled after school hours so that teachers can attend synchronous portions one night a week and do assignments the rest of the week at their own pace. Flexibility and relevance to teachers are the motivations for the scheduling.

In 2015, a new position was developed within the Faculty of Education (FoE) called the Indigenous Resurgence Coordinator (IRC). This position was created with the FoE commitment toward implementing the TRC Calls to Action pertaining to Education

and toward the goal of Indigenization and Indigenous resurgence in the FoE. The position has been able to advocate for greater inclusion of Indigenous content and resources in courses taught within the faculty and has provided support to faculty and instructors attempting to do more and need help with changing their approach or reviewing their course outlines. The IRC has also organized workshops on decolonizing practice, direct action support and trauma-informed learning. A challenge with the IRC role is that sometimes they are understood and used in token ways or have little power to enact the change that they see is needed in the unit. For IED, the IRC role has provided much needed program coordination support, especially in organizing the summer institute for secondary teacher candidates and the integrated course set for in-service teacher professional development. They have also been instrumental in forging relationships across campus and in the community to bring Indigenous resurgence activities forward in both spaces. It is hoped that in the future the training the IRC position provides to faculty members will build capacity for non-Indigenous teachers to more confidently take up reconciliation work instead of relying on support from the IRC or guest speakers.

Challenges

Centering Indigenous ways in a Western hegemonic institution means there are struggles. We are still in a university, not as nimble as local colleges in moving to assert Indigenous presence and programming, despite UVIC being recognized in Macleans' annual university rankings as number 1 in Indigenous visibility every year since the category was created in 2018. We are subject to collective agreements for staff and faculty that override Indigenous initiatives that might impact work contracts. We are still subject to BC rules and regulations governing academic programming like teacher education which is sometimes at odds with our own ways.

In some ways we are tokenized. As [Smith and Yang \(2019\)](#) observed,

The critical questions in terms of the call to "Indigenize" are, Who is making the call? Who is controlling the way that call is articulated? What Indigenous capacity is being developed and how is that being sustained over the long term? Sometimes these new ideas are opportunities to advance Indigenous thought and scholarship, Indigenous engagement and participation, but sometimes they simply add more work on to the shoulders of the few Indigenous scholars and other staff available in an institution with little attention being given to growing capacity, developing careers, improving relationships, or indeed transforming institutional practices" (p. 6/7).

There are still faculty that feel they don't need to include Indigenous programming or readings because "don't they have IED for that"? There is an assumption that students will have Indigenous material covered by us relieving them of the responsibility to cover it themselves or know it themselves. Some go as far as inviting IED to do the work in their classes. We have been asked to do land acknowledgments for them or to teach the Indigenous material as guest lectures.

There is also confusion within the institution about what it is Indigenous faculty do. UVIC has an Indigenous Studies program, an Indigenous Governance program and Indigenous Education (as well as many other great Indigenous programs), and we receive each other's mail and phone calls. Somehow still, with some in the university, anything Indigenous is treated as interchangeable.

Perhaps the biggest challenge is sustained funding for IED programs and students. ILR classes typically have fewer students than a regular class which is good for instruction, learning and practice but does not fit the standard university funding formula. In this time of neoliberalism, university programs often run on a cost recovery basis and are driven by market demand. Our ILR programming requires external funding to be able to run, be it from the communities themselves or through external grants. Each community delivery requires our partners to have their own (partial) funding in place as the added costs of ILR programs are not covered by university base funding. Many communities prioritize the language recovery and education among their members, but there are many others who are forced to cut language and culture funding when faced with housing, employment, safety and other needs that are not adequately addressed. Strong support for the added expense of ILR work would be good to see among the governments that were involved in the depletion and denigration of Indigenous languages during the residential schools era.

In addition to generating less revenue due to fewer students, ILR programs face additional costs. We often have to hire a coordinator that serves as a translator of university policy needs and deadlines for the Elders and language instructors who (rightly so) want to focus exclusively on sharing the language. Language mentors are often hired in addition to the facilitators who are the named appointees to courses. University guidelines don't accommodate this extra expense. Coordinating with community takes a lot of meeting time and labor not typical to most other programs. To remove barriers to students, much of the university interfacing is done for them. This is one of the reasons we need the staffing we have that is high for the number of faculty and students we have in our unit, but is required for the kind of programming that we do.

If universities are to truly work toward reconciliation, ILR has to be looked at as a long term investment rather than as a revenue stream. There are added costs for doing community-based programming such as travel, payment for community people, elders, knowledge keepers and language mentors. The MAP approach is costly because we pay our mentors as they are expected to do one on one language work for 75 up to 100 h per term per student (over and above appointed instructor costs). The added expense of renting space in community is not included in cost recovery and, it must be noted, even when we host courses on campus, running them in intensives means we often can't find suitable space on campus. Renting space for blocks of time is an added cost, as is food and gift giving which are culturally appropriate. Some courses require hiring a facilitator who deals with all of the administrative pieces of course instruction such as entering grades, processing academic concession requests and working

with university staff on issues in addition to the language speaker who is given the opportunity to focus exclusively on language teaching (they may find the whole workings of universities foreign and confusing).

While our admissions processes and partnership approach values the collective over the individual, as is common and expected in traditional Indigenous ways, we have experienced an institutional barrier in the funding of graduate students. The Faculty of Graduate Studies has a formula for providing support to units with a graduate program. The funding provided is meant to be transferred to the students in a way that the unit prefers. Perhaps one student receives a large grant or an incoming student is given a substantive entry scholarship. In IED, we like to distribute it as widely as possible to the entire group. However, this ends up being a small amount for each student (sometimes a \$1000 each—the minimum that can be disbursed). While every penny helps, it's also clearly not a system that has in mind supporting the collective but rather values competition and individualism. More Indigenous graduate student support is needed to truly reflect Indigenous values of community and collective. "Indigenous students said funding played a critical role in the success of their post-secondary experience. If the TRC's Calls are to be fulfilled in post-secondary spaces, it is important to increase funding for both Indigenous students and Indigenous services in post-secondary settings. Students said both play a critical role in their ability to be ready to learn and succeed in their programs" (Indspire, 2018, p. 7).

Sometimes our attempts to Indigenize practice have encountered colonial barriers that require us to challenge institutional practices and policies. For example, in our BEDILR, we have found the field placement guidelines of the provincial body overseeing teacher certification at odds with the realities of Indigenous community and language learning. The provincial conflict of interest guidelines require teacher candidates to be placed in schools where they have not worked before and where they do not have any close relations. When you are learning an endangered Indigenous language, the number of schools and communities where one can have a placement is incredibly low to begin with. It is highly likely that a candidate will have relatives who work there. IED succeeded in having our 2020 BEDILR cohort granted a one-time exemption for this condition but longer-term removal of these restrictions is a direction the province should be considering. As McIvor et al. (2018) noted, "the time to devote to language proficiency development continues to be limited by teacher education regulation leading to certification. An obvious new direction is the need to negotiate an alternative structure for degrees which are dually focused on Indigenous language proficiency building and creating professionally certified teachers" (p. 201).

One of the biggest challenges to our sustained success is Indigenous faculty and staff burnout. We are in a faculty of education as the smallest of four administrative units. We now have our autonomy as a department but being a faculty of just six (less with leaves) means administrative and service loads are significantly higher per faculty person than in larger departments. This is especially true as Indigenous faculty are called on by other units across the University to respond to the TRC calls to action for reconciliation, and demands for equity diversity and inclusion work. Indigenous faculty members are being pulled into more and more labor, some unaccounted for and undervalued. Indigenous faculty members get asked to support Indigenous students, mentor junior Indigenous faculty, be the Indigenous rep on graduate committees, and be an equity rep on service, representation and hiring committees. Indigenous faculty also commonly get asked to serve on research projects so that colleagues can access Indigenous funding envelopes. Public lectures, guest class visits and review of curriculum and policy with an Indigenous lens or perspective are just more demands on Indigenous faculty time that are all important but take a toll on one's own career progress, community commitments, and especially one's family and individual health.

Sessional instructors and limited term appointments have helped IED through some difficult periods with unexpected faculty sick leaves and the passing of one of our beloved colleagues but they should not be seen as a long-term solution to the issue of understaffing. In addition to the ethical issues surrounding the chronic underpaying of sessional instructors, term appointees are unrealistic supervisor choices for graduate students which is part of why faculty burnout occurs. Universities need to hire more Indigenous tenure stream professors in all areas of study if they want to be part of sustained and meaningful reconciliation efforts.

Looking forward

While IED remains embedded in an institution not of our own making that continues to have standardized requirements that restrict the full freedom and self-determination of our Indigenous knowledge systems and ways of being, there is still the possibility for positive social change as a result of our programming. I feel that in many ways our little unit is contributing to Indigenous futures in the ways that Smith and Yang (2019) wrote about:

My observation of earlier Maori or Indigenous scholars was that the academy consumed them, changed them, distracted them, and isolated them from the Indigenous world and that we, as a community, could not trust all of them to work for us. My critique of colonialism was not just about looking at the colonizer but also looking at what colonial hegemony was doing within our own Indigenous minds, spirits, and behaviors. I saw the need for a decolonizing agenda that dealt with the whole of the dialectic of colonizer-colonized and recognized the role of education as a means to transform colonialism at deep levels of knowledge, pedagogy, the shaping of minds and discourses.

Now, I feel that I am writing as part of a community of Indigenous scholars who have deepened understandings of the work of decolonizing education and, importantly, created new approaches to education that theorize, revitalize, enhance, and produce Indigenous educational experiences that support Indigenous futures" (Smith and Yang, 2019, p. 6).

Our work with Indigenous language revitalization, in partnership with Indigenous communities, taking guidance from our partners and our community-led Advisory Board, are a direct support for Indigenous futures.

Indirectly, if we ever want to see change in the relationship between Canada and Indigenous people, the next generation of Canadian school students has to learn some harsh truths about our shared history. They also need to learn about our current contributions, and aspirations and dispel stereotypes and fear about Indigenous people. One of the best ways to ensure this result is to give future teachers the skills they need and foster an orientation toward doing the right and responsible thing. Our teacher professional development activities and Summer Institute are contributing to teachers who will take the steps toward ever greater decolonization and inclusion of Indigenous perspectives and content in their own classrooms. Further, our BEDILR program helps ensure that Indigenous languages will be taught in schools, day cares, language nests and many other spaces.

It is an honor to be a part of the dreaming and the commitment and to action on our visions (Laenui, 2000). Perhaps we are seeking radical reform within a soft-reform space (Andreotti et al., 2015). Perhaps we are finding ways to hack a system (the university) to provide for flourishing Indigenous futures when universities have been opposed to Indigenous knowledge and being from their founding (Andreotti et al., 2015). By centering Indigenous knowledge, being accountable to community and always striving to be a place that is respectful of Indigenous people and practices, relevant to Indigenous community approaches, reciprocal to Indigenous communities and responsible to our relations to land and people, I believe IED is making space for decolonial, Indigenized spaces that support community aspirations and futures.

The Department of Indigenous education at UVIC has only been here since 2017. We are just getting started. I have been grateful to contribute these past years and can't wait to see what will happen next.

References

- Andreotti, V., Stein, S., Ahenakew, C., Hunt, D., 2015. Mapping interpretations of decolonization in the context of higher education. *Decolonization* 4 (1), 21–40.
- Archibald, J., 2020. Indigenous Education in Canada. In: *Oxford Research Encyclopedia of Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.818>.
- Bascuñán, D., Carroll, S., Restoule, J.P., Sinke, M., 2022. Teaching as trespass: avoiding places of innocence. *Equity Excell. Educ.* <https://doi.org/10.1080/10665684.2021.1993112>.
- Battiste, M., 1998. Enabling the autumn seed: toward a decolonized approach to aboriginal knowledge, language and education. *Can. J. Native Educ.* 22 (1), 16–27.
- British Columbia Ministry of Education, 2022. New Indigenous-Focused Graduation Requirement. https://www2.gov.bc.ca/assets/gov/education/kindergarten-to-grade-12/support/graduation/indigenous_focused_grad_requirement_summary.pdf.
- British Columbia Teachers Council, 2019. Professional Standards for BC Educators. https://www2.gov.bc.ca/assets/gov/education/kindergarten-to-grade-12/teach/teacher-regulation/standards-for-educators/edu_standards.pdf.
- Carroll, S., Bascuñán, D., Sinke, M., Restoule, J.P., 2020. How discomfort reproduces settler structures: moving beyond fear and becoming imperfect accomplices. *J. Curric. Teach.* 9 (2), 9–19. <https://doi.org/10.5430/jctv9n2p9>.
- Comtassell, J., 2012. Re-envisioning resurgence: indigenous pathways to decolonization and sustainable self-determination. *Decolonization* 1 (1), 86–101.
- First Nations Education Steering Committee, (n.d). First Peoples Principles of Learning. <http://www.fnesc.ca/wp/wp-content/uploads/2020/09/FNESC-Learning-First-Peoples-poster-11x17-hi-res-v2.pdf>.
- First Peoples Cultural Council, 2022. Language Resource Library. <https://fpcc.ca/resource/area/language/>.
- Gaudry, A., Lorenz, D., 2018. Indigenization as inclusion, reconciliation and decolonization: navigating the different visions for indigenizing the Canadian University. *Alternative* 14 (3), 218–227.
- Hinton, L., 2001. The master-apprentice language learning program. In: Hinton, L., Hale, K. (Eds.), *The Green Book of Language Revitalization in Practice*. Academic Press, San Diego, CA, pp. 217–226.
- Hinton, L., Vera, M., Steele, N., 2002. *How to Keep Your Language Alive: A Commonsense Approach to One-On-One Language Learning*. Heyday Books, Berkeley, CA.
- IED and UATÉ Stories, 2022. *Illuminating Impacts*. Video Produced by Indigenous Education UVIC and UATÉ Stories. Available. <https://vimeo.com/661300410>.
- Indspire, 2018. Truth and Reconciliation in Post-secondary Settings: Student Experience. https://indspire.ca/wp-content/uploads/2019/10/JMGD_003_IND_TR_REPORT_FINAL_V3_NOV15_V3.pdf.
- Kirkness, V.J., Barnhardt, R., 1991. First nations and higher education: the four R's—respect, relevance, reciprocity, responsibility. *J. Am. Indian Educ.* 30 (3), 1–15.
- Laenui, P., 2000. Processes of decolonization. In: Battiste, M. (Ed.), *Reclaiming Indigenous Voice and Vision*. UBC Press, Vancouver, BC, pp. 150–160.
- Landertinger, L., Tessaro, D., Restoule, J.P., 2021. We have to get more teachers to help our kids: recruitment and retention strategies for teacher education programs to increase the number of Indigenous teachers in Canada and abroad. *J. Glob. Educ. Res.* 5 (1), 36–53. <https://www.doi.org/10.5038/2577-509X.5.1.1066>.
- May, S., Aikman, S., 2003. Indigenous education: addressing current issues and developments. *Comp. Educ.* 39 (2), 139–145.
- McGregor, C., 2014. Bachelor of education in Indigenous Language Revitalization: honoring, respecting, and celebrating Indigenous Language Learning. In: *An Evaluation Study: Summary and Response*. Faculty of Education, Victoria. UVIC. Available. https://www.uvic.ca/education/indigenous/assets/docs/EDUC_5000_BEDILRevalReportV6OUT_web.pdf.
- Mclvor, O., Rosborough, T., McGregor, C., Marinakis, A., 2018. Lighting a fire: community-based delivery of a University Indigenous Language Teacher Education Program. In: Paul, W., del, M., de France, C.R., Mclvor, O. (Eds.), *Promising Practices in Indigenous Teacher Education*, pp. 189–204.
- Nardozi, A., Restoule, J.P., 2019. Exploring teacher candidate resistance to indigenous content in a teacher education program. In: Zinga, D., Styres, S., Jahnke, H., Lilley, S. (Eds.), *Indigenous Education: New Directions in Theory and Practice*. University of Alberta Press, Edmonton, pp. 311–337 (Chapter 14).
- Smith, G.H., October 2003. Indigenous Struggle for the Transformation of Education and Schooling (Keynote Address to the Alaskan Federation of Natives (AFN) Convention. Anchorage, Alaska, US October 2003).
- Smith, L., Yang, K.W. (Eds.), 2019. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge.
- Smith, L., 1999. *Decolonizing Methodologies: Research and Indigenous People*. Zed.
- Tanaka, M., 2016. *Learning and Teaching Together: Weaving Indigenous Ways of Knowing Into Education*. UBC Press, Vancouver.
- Truth and Reconciliation Commission of Canada, 2015. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. www.trc.ca.

Being in good relations: higher education's responsibilities to Indigenous resurgence

Michelle Pidgeon, Faculty of Education, Simon Fraser University, Surrey, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Unsettling colonialism	100
What it means to be in "good relations"	101
Tensions navigating transformation	102
Seeing the possible through individual and collective actions	103
Policies as good relations	104
Programs in good relations	104
People and practices in good relations	105
Transforming Indigenous education through good relations	106
References	107

Unceded and Occupied Shared Territories of the Coast Salish Peoples.

The aim of this paper is to explore what it means for higher education to be in good relations with Indigenous peoples' self-determination and resurgence in Kanada,¹ also known as Canada. "Indigenous resurgence means having the courage and imagination to envision life beyond the state" as defined through colonially imposed government structures (e.g., policies, programming) (Corntassel, 2012, p. 89). By focusing on "everyday" acts of resurgence, one disrupts the colonial physical, social and political boundaries designed to impede our actions to restore Indigenous nationhood (Corntassel, 2012). The specific context for this disruption of settler colonial narratives is Canadian higher education. I share understandings of what it means to wholistically support Indigenous resurgence that enacts Indigenous visions of transformation and change at the individual, community, and organizational levels. It is my intention to weave these lessons of how to be in good relations by sharing a few research stories. It is my intention to weave these lessons of how to be in good relations by sharing a few research stories. As I begin to think which stories to share, I reflect on the words of Kirkness and Barnhardt (1991), as their seminal work has been a guide for my own scholarship in Indigenous and higher education.

What First Nations people are seeking is not a lesser education, and not even an equal education, but rather a better education — an education that respects them for who they are, that is relevant to their view of the world, that offers reciprocity in their relationships with others, and that helps them exercise responsibility over their own lives. It is not enough for universities to focus their attention on "attrition" and "retention" as an excuse to intensify efforts at cultural assimilation. Such approaches in themselves have not made a significant difference, and often have resulted in further alienation. Instead, the very nature and purpose of higher education for First Nations people must be reconsidered, and when we do, we will find that the entire institution, as well as society as a whole, will be strengthened and everyone will benefit.

Kirkness and Barnhardt (1991, p. 14).

I intend to exemplify what it means for Canadian higher education to be in good relationships with Indigenous people's self-determining vision of education through an examination of system transformation. Ultimately, these stories are told to share of system transformation of education, by and for Indigenous peoples has occurred in Canada.

A clear vision of what Indigenous people's had for education of their people was first articulated in *Indigenous control over Indigenous education* (National Indian Brotherhood, 1972). As the authors of the report stated "It is a statement of the: philosophy, goals, principles, and directions which must form the foundation of any school program for Indian children" (National Indian Brotherhood, 1972, p. iii). This document has been guiding Indigenous resurgence through system change by and for Indigenous peoples across this country over the last 50+ years. Furthermore, this document provided an Indigenous re-visioning and re-awakening of the purpose of higher education for Indigenous peoples.

Considering the great need there is for professional people in Indian communities, every effort should be made to encourage and assist Indian students to succeed in post-secondary studies

National Indian Brotherhood (1972, p. 13).

¹Canada's name, is intentionally spelled here with a K to acknowledge the origin of the name is taken from Huron-Iroquois word "kanata," meaning "village" or "settlement."

Within their vision for post-secondary education, this report speaks to the recruitment processes and programs, entrance requirements to focus on ability, aptitude, intelligence, diligence, and maturity. There was also direct mention to federal financial support “those who have the motivation and talent to do post-graduate studies, should receive total financial assistance” ([National Indian Brotherhood, 1972](#), p. 13).

My goal is to help readers understand how transformation of Indigenous education in Canadian higher education has occurred through programs, policies, people, and practices. As the reader—whether settler or Indigenous—one is being asked to actively reflect and consider their own actions and commitments to transforming Indigenous education. This responsibility is both an individual and collective community effort deepen understandings, and ultimately one's enactment, of what it means to be in good relations in these efforts.

Unsettling colonialism

Most Indigenous communities living within colonized circumstances still suffer high and disproportionate levels of educational underdevelopment. [...] We intentionally use the term “underdevelopment” to connote the understanding that the “problematic” of Indigenous education is not just about the Indigenous learner deficit but also the recognition that there are often systemic and structural impediments that results from colonizations and associated approaches.

Jahnke and Smith (personal communication June 17, 2021).

Many of us within higher education systems critique the very structures that we are aiming to decolonize and leverage the resources available to us as faculty, administrators, and staff to support Indigenous resurgence through programs, policies, and practices. My thinking has been influenced by those Indigenous scholars who have been working to transform Indigenous post-secondary education (e.g., [Archibald, 2008](#); [Archibald et al., 1995](#); [Battiste, 2013, 1998](#); [Brayboy, 2005](#); [Brayboy et al., 2012](#); [Cote-Meek, 2014](#); [Cote-Meek and Moeke-Pickering, 2020](#); [Kirkness, 1998](#); [Kuokkanen, 2007](#); [Parent, 2017](#); [Restoule, 2011, 2005](#); [Restoule et al., 2013](#); [Shotton et al., 2013](#); [Stonechild, 2006](#); [Styres, 2017](#); [Waterman et al., 2018](#)). Such researchers are sometimes explicit but many times implicit in the ways that they individually and collectively support transforming the academy for Indigenous peoples. In fact, like [Comtassel \(2012\)](#), I see decolonization and resurgence as “interrelated actions and strategies that inform our pathways to resistance and freedom” (p. 89).

As an author in this collection, I recognize that the educational systems we are trying to transform, from within and from the outside, are colonial—even Indigenous K-12 schools and/or post-secondary institutions in Canada are required to engage in relationships with settler-organizations through transfer agreements, accreditation, and other fiscal reporting structures at provincial and federal government levels—while they simultaneously operate Indigenously through their governance system, academic programs, and support services. While Indigenous peoples work to decolonize and Indigenize alongside allies—we are simultaneously engaging, pushing back, and resisting imposed policies and practices that are deeply entrenched in settler colonialism. I am reminded of Ruana [Kuokkanen's \(2007\)](#) book “Reshaping the University: Responsibility, Indigenous Epistemes, and the Logic of the Gift”, in which she urges readers to consider how emmeshed our lives are with colonialism, and cautions us to be ever mindful that this is the context that we are in:

When we ignore the ways in which we are implicated in academic structures, and when we assume that we can remain uncontaminated by outside influences despite our daily interactions with the academy, we are only contributing to our own marginalization and to the construction of a monolithic understanding of our projects (p. xv)

This requires one to be conscious of the ongoing tensions of new forms of colonialism being enacted. [Kuokkanen \(2007\)](#) further suggests

that it is up to the academy to do its homework and address its own ignorance so that'll be able to recognize and give unconditional welcome to Indigenous peoples, worldviews and philosophies. But before the academy can recognize the gift of indigenous epistemes, it will have to profoundly transform itself; it will not be enough to merely include indigenous epistemologies (i.e., Indigenous systems of knowledge, or ways of knowing) in pedagogy and curricula. First and foremost, the academy will have to acknowledge that it is founded on the very limited conceptions of knowledge in the world. Because this constraint perception the academy cannot grasp or even hear views that are grounded in other epistemic conventions (pp. 2–3)

Indigenous ways of knowing and being encounters ongoing resistance and ridicule in a colonial western society where neoliberalism, economic exploitation, cultural oppression and individual gains in essence are valued over and run counter-to Indigenous values of community, collectivism, and interconnectedness (Jahnke & Smith, personal communication June 17, 2021). In recognizing the reality of colonization within our lives, “what we need instead is a more nuanced language as well as a careful understanding of our various circumstances and situations. Having established these will be able to examine more closely certain common indigenous philosophical premises and principles” ([Kuokkanen, 2007](#), p. 25). We cannot be complicit in assuming

that with transformation we will all of a sudden be in our ideal utopia nor that what works in one context will be universally transformative for all—yet, in reflecting on all the changes that have been advanced here in Canada—a beacon of possible hope is provided to help guide us through this ongoing resistance, reclamation, and resurgence. “But no matter how the educational landscape has been transformed, the transmission and continuity of Indigenous knowledges have been the foundation for learning that Indigenous families, communities, and nations have always maintained for their children” (Hare and Davidson, 2020, p. 203).

What it means to be in “good relations”

Central to Indigenous resurgence is the understanding of what it means to be in good relations; it is a cultural value and philosophy shared among many Indigenous peoples grounded in one's relationship with the animate and inanimate. It is a concept that is not bound by human-to-human relationships but our interconnected relationships to the physical and metaphysical worlds (Aluli-Meyer, 2006; Deloria, 2001; Kuokkanen, 2007; Mika, 2017). It is a responsibility to carrying one's self in good relationship on this planet. Being in good relations translates wholistically, in seeing oneself connected to family, community, and nation within a balance of the physical, emotional, spiritual, and intellectual (See Fig. 1—Indigenous Wholistic Framework). From their seminal article “First Nations and Higher Education: The Four R's—Respect, Relevance, Reciprocity, Responsibility,” Kirkness and Barnhardt (1991) describe good relationships as “reciprocal relationships,” which they deem important aspect of transformation, as they explain

One of the most frustrating aspects of the university experience for First Nations students is the role dichotomy between the producers and the consumers of knowledge in university settings. The conventional institutionalized roles of a university faculty member as the creator and dispenser of knowledge and expertise and the student as the passive recipient of that knowledge and expertise have a tendency to interfere with the establishment of the kinds of personalized “human” relationships to which First Nations students are most likely to respond (p. 9)

In naming the tensions of the dichotomy, they not only articulate the ongoing tension of being Indigenous within the academy but also provide insight to what is truly of value within Indigenous education—it is a journey to being in relationships with each other learning from each other, in essence—reciprocal learning. Respectful and reciprocal learning sets the foundation of being in good relationships.

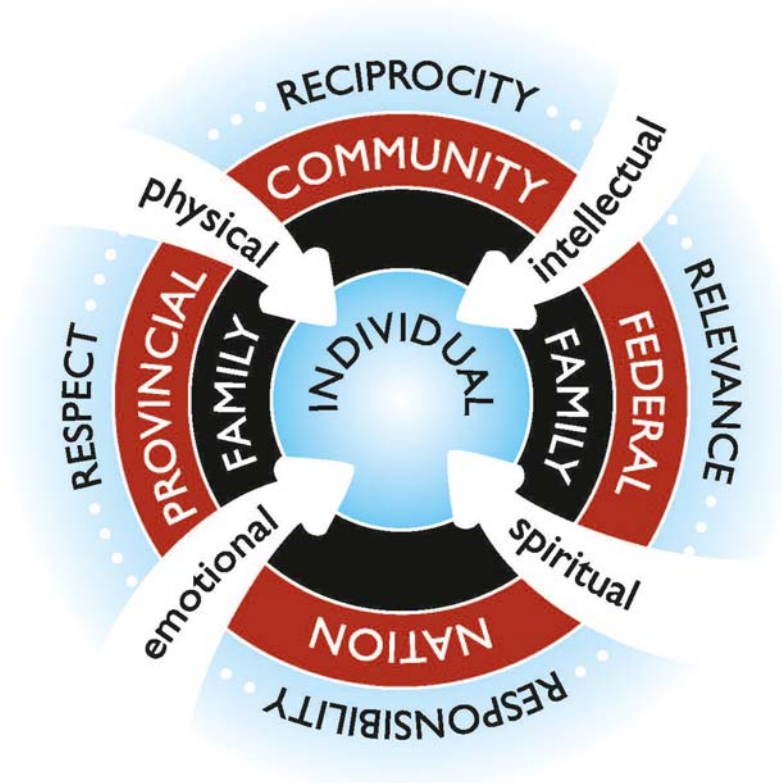


Fig. 1 Indigenous wholistic framework (Pidgeon, 2005, 2014, 2016a).

The Indigenous Wholistic Framework both theoretically and methodologically has guided and informed all aspects of my scholarship as it centers place and Indigenous ways of knowing and being in a wholistic sense. It connects the physical, emotional, cultural, and intellectual needs of the individual with their family, communities, and beyond. Around the framework, I have taken up the 4Rs (Kirkness and Barnhardt, 1991) to reflect the values that guide meaning making across the educational system—from the individual to community levels.

In thinking what transformation may mean for Indigenous education within this Framework, I think about various organizational and systemic structures that are involved through the interconnections and interrelatedness of policies, programs, people, and practices (Pidgeon and Rogerson, 2017). These 4Ps build upon the 4Rs embedded in the Indigenous Wholistic Framework moving this theoretical framework into a system change model for Indigenous education, that centers and privileges Indigenous ways of knowing, being, and doing in relationship to the lands and waters in which we do this work. I attest this is not the only way to do this work, it is just one way, one of the challenges of working within the academy is the ongoing homogenization of knowledges, practices, and ideas to be the “one” solution. This framework and what is shared through this writing is an offering of gifts of knowledges, experiences, and stories shared with me through research—many experiences inform what is shared—I raise my hands in thanks to all those who have helped inform, guide, and influence this work. In this acknowledgment, I honor the responsibility I have to those who have shared their stories with me; this responsibility lasts well beyond the conclusion of a particular research project. It is an example of how being in good relationships with Indigenous peoples is lived in practice.

Tensions navigating transformation

In centering how many Indigenous peoples articulate being in good relationships; we can now turn our attention to the tensions, pushes-pulls, and intentional manipulation of what it means to be in good relationships Indigenously while in a predominately Western educational system. In Canada, the ongoing colonial project is evident through power, the privileging of Western-valued ways of doing and being, and the reproduction of these systems over time. This educational system is rooted in settler colonialism and consequently, this privileging of whose knowledge counts is deeply connected to who maintains power over the lands and peoples connected to these lands. This power then influences the communities and individual actions that further enact colonialism through lateral violence, racism, and commodification of Indigenous knowledges.

Leigh Patel (2021) writes in her book *No Study without Struggle Confronting Settler Colonialism in Higher Education*, “barriers to higher education may take many forms, and although different populations experience different forms of discrimination, they are connected through the larger structure of settler colonialism” (p. 6). She then eloquently outlines how settler colonialism is interwoven into every aspect of higher education from the lands upon which institutions reside, to the policies that dictate who has access, who can afford access, and who can truly benefit from higher education (e.g., attainment of individual wealth and property due to their education). Kuokken (2007) and Patel (2021) make evident the ways the ongoing settler colonialism is deeply intertwined in any calls for transformation and change (e.g., Indigenization, decolonization, Land back).

The cautionary tale in this work is that in some ways, the more the Western canon comes to understand and embrace Indigenous ways of knowing and being—the more that canon appropriates Indigeneity and uses it against Indigenous peoples. Let me provide an example. More and more there is an understanding of the central importance and value of relationships within Indigenous communities; yet, frequently Indigenous scholars, students, and Elders are asked to do extra emotional and/or cultural labor for the institution to help “build good relationships” only to have it later used against them. This occurs when an Indigenous undergraduate student, who has been a leader in student-led initiatives is then penalized because they are taking too long to graduate, losing access to funding, housing, and other resources. It happens when Indigenous scholars who are under review for tenure and promotion are told they have done too much service even though this was the service they were asked to do on behalf of their institution. It also happens when Elders who speak from a place of wisdom and truth at public institutional events are then all of a sudden not invited back because they challenged the conception of what it means to be a “gentle Elder”. These examples are subversive acts of ongoing settler colonialism-infringing on what might be seen as cultural responsibilities that are then used against Indigenous peoples because we fail the Western colonially imposed and defined metrics of what it means to be “a good Indian.”

We can also see the ongoing colonial attempts at genocide and assimilation in the perpetuation of the view of homogeneity of Indigenous peoples that sees “sameness” to breakdown cultural and linguistic differences that have been deeply shaped by the lands and waters of their traditional territories. This is another example of the politics of distraction (Smith, 1997) as such homogenizing rhetoric requires Indigenous peoples alongside their allies to be hyper-vigilant of the colonial gaze and reminding the “others” (i.e., those who impose colonial views/biases) of the true diversity of Indigenous people. Being in good relations means coming to a deep understanding of the richness and diversity of Canada’s first peoples—even with the terms First Nations, Métis, and Inuit—we lose the uniqueness of the over 60 different Nations across this country—each of whom are defined by cultural protocols and understandings of relationship with lands and waters of their homelands. The lessons emerging from the lands, water, and air is what sets us in good relationships with the animate and inanimate. The relationship with our environment—whether mountain ranges, rugged coast lines, or the expanse of the tundra—inform our cultural knowledge, languages, and practices. As such, many Indigenous peoples believe everything is interconnected (Battiste, 2000; Mika, 2017). This interconnection is what colonization has attempted to erase.

It is such deep beliefs in the interconnection of all things that is offered through an Indigenous transformation of the educational system. The embodied Western knowledge and values of individualism and competitiveness becomes challenged by Indigenous

values of collectivism and community. The tensions between bringing Indigenous ways of knowing and being into the academy have been well discussed by others. However, in making visible the invisible means that we are also then able to also enact change. As I write this, I pause and think of black queer feminist, Audre Lorde, whose voice still rings true with her essay “the Master’s tools will never dismantle the Master’s house” (Lorde, 1984). The tools of colonization and genocide built the system we have today and calls for decolonization intentionally identify those tools for their true purpose—assimilation, genocide, erasure. What I attempt to offer here is another way of thinking about dismantling the masters house (e.g., education) through using Indigenous “tools” that empower an ongoing and iterative anti-colonial and transformational process. Can we dismantle the masters house with different tools, that is, Indigenous ways of knowing and being? Within the context of higher education—this means honoring the diversity of Indigenous epistemologies, ontologies, and axiologies and ensuring these deep knowledge systems are guiding the transformation through values and cultural principles may influence systemic transformation through decolonization means shifting institutional values from individual achievement and attainment to collective achievement and attainment. The short answer is “yes”—but it’s a long path to such changes, therefore, we must see this system change as iterative as the visions of Indigenous education futures are still in the dreaming, story making, and becoming phases. We are in the middle of the seventh generation preparing the system change needed to support the next 7 generations. This futurity of Indigenous education is the transformation process—the enacting and doing things Indigenously.

Seeing the possible through individual and collective actions

How do we “undo” colonization? It’s a deep energy that has tentacles across so many aspects of our lives. Yet, we also have a deepened consciousness of these tentacles and an understanding that colonialism has not gone away—these tentacles are much more visible, identifiable, and can be decolonized while we enact more anti-colonial practices. In these moments, we see how Indigenous peoples reclaim and enact Indigenous resurgence by taking back land, waters, and air to ensure Indigenous ways of being, knowing, and doing not only exist, but thrive. The question we need to be asking of ourselves as human beings on this fragile planet is “in whose interests does this work”? In many ways the other contributors to this volume, supported by research, attest that the educational system is failing many of us while further perpetuating broader social inequities. As Marie Battiste (2013) reminds us taking on this work

any attempt to decolonize education and actively resist colonial paradigms is a complex and daunting task. Educators must reject colonial curricula that offer students a fragmented and distorted picture of Indigenous peoples, and offer students a critical perspective of the historical context that created that fragmentation. In order to effect change, educators must help students understand the Eurocentric assumptions of superiority within the context of history and to recognize the continued dominance of these assumptions in all forms of contemporary knowledge” (p. 186).

Battiste (2013) was speaking to K-12 education system, I firmly believe it also applicable to all of us within post-secondary education—whether administrator, staff, faculty, instructor, or student. Consequently, the transformation of Indigenous higher education is about unveiling the hidden aspects of settler colonialism that go beyond the discussions and naming of anti-Indigenous racism—to ones that speak to ongoing enactment of settler colonization through systemic exclusion. As Gaudry and Lorde (2018) speak to the continuum of Indigenization, they acknowledge while institutional administrators are focusing on Indigenization inclusion (e.g., increasing Indigenous student, faculty, and staff) they acknowledge that Indigenous peoples are aiming for decolonial Indigenization which “envision[s] a wholesale overhaul of the academy to fundamentally reorient knowledge production based on balanced power relations between Indigenous peoples and Canadians, transforming the academy into something dynamic and new” (p. 226). System transformation is possible, but it is a long game thinking forward to the next seven generations. Stories of those who created the pathways in the 1960s and 1970s demonstrate that institutional change is possible, and with institutional change there is system change. From your own context you can probably identify who those Indigenous trail blazers were. In Canada, as part of Native Indian Brotherhood (now known as Assembly of First Nations), Verna Kirkness and others set the path for Indigenous programs (e.g., Native Teacher Education), and ultimately Indigenous student services in public post-secondary education (Pidgeon, 2005, 2014, 2016b; Stonechild, 2006).

Such individuals formed community around a shared vision (e.g., Indian Control over Indian Education, 1972), which then influenced educational systems across the provinces and territories and directly impacted national policies related to education (Carr-Stewart, 2019). Indigenous institutes of adult and higher education (e.g., Nicola Valley Institute of Technology (NVIT) (BC), Blue Quills College (SK)) were created to center and uphold Indigenous ways of knowing, being, and doing –by setting their own priorities and developing critical pathways of access to and for Indigenous learners. Furthermore, professional organizations, such as Indigenous Adult and Higher Learning Association (IAHLA <https://iahla.ca>) in BC and the National Associate of Indigenous Institutes of Higher Learning (https://www.afn.ca/uploads/files/2016_education_forum/session_5_-_post-secondary_education.pdf), are key advocates and voice for Indigenous community-based education.

Policies as good relations

In Canada, higher education is a provincial responsibility; however, Indigenous K-12 education is a legislated federal responsibility. The federal government sees post-secondary as a social obligation not a legal one (Stonechild, 2006). This relationship is set out in the Indian Act, which is a clear example of how well policy designed to eradicate Indigenous peoples from their lands, languages, and cultures can continue to influence society today (Battiste and Barman, 1995; Carr-Stewart, 2019, 2001; Fisher et al., 2006; Royal Commission on Aboriginal Peoples, 1996). For example, under the Indian Act, First Nations who pursued post-secondary education were enfranchised (i.e., lost their status as Indigenous peoples) until 1955 (Stonechild, 2006). It is no surprise that post-secondary education has not always been hospitable to Indigenous peoples—and the disparity of participation between Settler and Indigenous peoples in Canada continues to demonstrate the inequities across educational attainments.

While colonial policies have disrupted Indigenous education; it is also important to recognize the impact and influence how position papers and policies (e.g., National Indian Brotherhood, 1972; Royal Commission on Aboriginal Peoples, 1996; Truth and Reconciliation Commission of Canada (TRC), 2015), along with the international United Nations Declaration of the Rights of Indigenous Peoples (UNDRIP) (United Nations, 2007) are now shifting Canadian higher education to step into the work of living better relationships. What is evident from reading these documents is the consistent vision and calls for systemic change of Indigenous education by and for Indigenous peoples that connects lands, languages, cultural practices, and Indigenous knowledges as the foundation for Indigenous education (Hare and Davidson, 2020). The impact of the TRC (2015) report in particular has taken up calling on Canadians to embrace their responsibilities for learning the truth and being agents of change at the individual, community, and national levels regarding reconciliation.

Policy therefore can be a tool of decolonization and transformation, particularly at an institutional-level where local context and actors can enact change within their college or university, and become “good practice”, influencing others to follow suit. Within higher education, policies that support access, transition, and persistence through the PSE are a critical enactment of transformation. This is particularly relevant to supporting Indigenous student success—wholistically supporting their success means going beyond academic programs (i.e., individual intellectual needs) but also ensuring that the community supports (?) their emotional, cultural, and physical needs in culturally-relevant ways. In 2017–2019, my research team and I as part of a larger project to support Indigenous pathways to and through university undertook a national examination of institutional policies in this regard as a way of documenting how public universities in Canada were supporting Indigenous access, transition, and persistence to postsecondary education (Pidgeon et al., 2019). In this policy review, we noted that 80% of universities (N = 74) had institutional-specific awards and bursaries for Indigenous students. Additionally, many institutions connected Indigenous students to external awards (e.g., Indspire). These policies, however, remain couched in a colonial narrative. While over 50% of universities had Indigenous admissions policies, fewer institutions (13%) had specific housing policies for Indigenous students, or offered targeted on-campus family housing. In a deeper review of these policies—the tensions of how students self-identified (particularly for financial support) varied, from proof of status, to garnering letters of support from their communities or others who could verify who they were, to open self-declaration. The housing policies were of interest given that many First Nations, Métis, and Inuit students have to leave their home communities to attend university, yet policies focused on single occupancy and few offered family housing specific to Indigenous students. This means that those Indigenous students who are also parents—whether at the undergraduate or graduate—are left to find accommodations off-campus on their own.

Programs in good relations

The province of British Columbia committed to upholding the principles of UNDRIP through formal policy of Declaration on the Rights of Indigenous Peoples Act (DRIPA) on November 28, 2019 (DRIPA: <https://www2.gov.bc.ca/gov/content/governments/indigenous-people/new-relationship/united-nations-declaration-on-the-rights-of-indigenous-peoples>). While many public BC post-secondary institutions were already working to respond to the TRC (2015) calls to action, they now have further impetus to further their commitments to reconciliation via institutional plans and strategic documents, which ultimately guides (and shifts) institutional policies and practices. Such institutional responses to the TRC Calls to Action are evident across the country. However, it is also important to remember that transformation in the academy began well before 2015.

Academic change is a long process and began with those trail-blazers of the 1950s and continues with those who followed them. We can certainly identify where growth is happening— and yet, we also see the huge gaps and pitfalls that still lie in the way of advancing Indigenous control over Indigenous education, or even how the 4Rs of respect, relevance, responsibility, and reciprocity have been taken up (Kirkness and Barnhardt, 1991; Marker, 2004). A period of growth in the 1970s occurred with the establishment of Native Teacher Education and Indigenous Studies (e.g., formerly Native Studies) programs. This expansion of Indigenous specific programs, courses, and specializations continued in the late 1990s to present day (e.g., Geography, Linguistics, Forestry, Social Work, Health Sciences, STEAM), while reserved seating and culturally relevant programming (e.g., Law, Medicine, Nursing) have expanded Indigenous access to relevant academic programming and services. These changes are the result of the visions of our ancestors being enacted into being—yet there is still much to be done within the classrooms if we truly want to be more inclusive of Indigenous ways of knowing and being. Changes are still needed, from the curriculum and pedagogical practices to the classroom experiences of Indigenous learners.

However, the post-secondary experience is not just about the in-class experiences for Indigenous students, it is also about how their student experience outside of the classroom is supported and the systemic racism that exists on our campuses both within and outside of the classroom, is addressed. Early on in the 1970s it was recognized that due to the systemic and educational barriers facing them, Indigenous students needed to have wholistic supports in place when entering into post-secondary education. This awareness led to the establishment of Indigenous student services during the time of the first cohorts of Indigenous teacher education students (Pidgeon, 2016b). Between 2018 and 2019, my research team conducted a national study on Indigenous access, transition, and persistence programming—in which we identified that 95% of the 74 public universities in Canada offer culturally-relevant support services through Indigenous student services. This is notable growth—in 2001 I reviewed these same institutions and found only 49% provided some form of Indigenous student services (Pidgeon et al., 2019). Indigenous student services is the “home away from home” for many First Nations, Métis, and Inuit students along with Indigenous students from the United States, Australia, and elsewhere who choose to study in Canada. These culturally-relevant support services offered to meet the needs of the whole student tend to students’ physical, emotional, cultural, and spiritual needs (Shotton et al., 2013; Waterman et al., 2018). The Centers provide in essence sanctuary of being “Indigenous” and offer students a break from being “othered” across the campus.

In reviewing how Canadian public universities supported Indigenous students, we found that most (69%) focused on support programming for the transition of Indigenous students into their institutions (e.g., programming focusing on the first-year experience). About half provided some form of access programming (e.g., summer camps for K-12 students), but only 27% offer retention programs that support the persistence of Indigenous students beyond the second year of study (Pidgeon et al., 2019, p. 37). Even fewer universities (N = 8) supported Indigenous-focused co-op or experiential programs. For example, the University of Calgary provides an Indigenous Co-op Program that administers one-on-one dedicated support with an Indigenous career specialist. Carleton University promises its graduates the opportunity to be equipped to work in their band councils, community organizations, and the federal civil service in Ottawa (Pidgeon et al., 2019, p. 37). Having programs—whether academic or student services—is an important step to supporting Indigenous student success and transforming their university experiences both within and outside the classroom. However, it is also evident that programming alone or financial supports alone is not enough—there is an interconnection between wholistically supporting students and system transformation that is critical- and that directly relates to the people (and their practices).

The challenge to also note here is that these culturally-relevant services must continue to operate within an ongoing systemic colonial policies and procedures that continue to push back on how Indigenous programming and services are provided within the institution. For example, central financial office pushing-back on processing payments to knowledge holders due to lack of information (despite forms being fully completed). Another is students failing academic papers when citing Elder’s teachings in their paper. The stories I’ve heard continue to attest that while more programming and services exist there is an ongoing surveillance over these initiatives constraining how “Indigenous” ways of doing the work is limited. Within postsecondary institutions, the enactment of policies and development of culturally relevant academic programs and support services would not be possible without people practicing good relations.

People and practices in good relations

In Canada, as elsewhere, there is a recognition that Indigenous education transformation means having Indigenous peoples at every level of the institution. Being the seen face matters both in role modeling good practices, and establishing good relationships; but Indigenous ways of being and doing are going to be different and challenge Western-construed understandings of leadership. Elsewhere I’ve written about this in more detail (Pidgeon, 2008, 2014, 2016a) but reflecting on the importance of having the good people walking in this work, those who are focused on the collective vision of what is needed to support Indigenous higher education, I fully recognize that the policies and programs in place were built by people—Indigenous and others who walked alongside them—to do something meaningful, to enact change for the better for Indigenous students and their communities. Indigenous programming, policies, and services were built by those committed to calling out injustices, systemic racism, and pushing boundaries of what counted as knowledge, learning, and being in the academy. This relational way of being within the academy reminds me of the vision set out within *Indian Control over Indian Education*, where Indigenous peoples have meaningful representation, presence, voice, and leadership across the institution.

Indian people should seek representation on the governing bodies of institutions of higher learning. This includes university senates and boards of governors, as well as the governing councils of colleges, community colleges and technical schools.

National Indian Brotherhood (1972, pp. 13–14).

I recognize that this relationship between Indigenous peoples and post-secondary institutions is still largely one-way—grassroots change implemented by a core group of committed individuals pushing against the system. The question I am left with is how much of this collective change will lead to institutional transformation—the power structures that continue to privilege the western canon give way to value Indigenous ways of doing and being. Therefore, being in good relations means that we must also recognize the hidden tentacles of colonization that exists in our institutions—the overt and covert forms of anti-Indigenous racism that continues

to negatively impact, impede, or intentionally disrupt and distract those committed to institutional transformation. Some examples of that I see enacted are the questioning of “why are Indigenous peoples being treated special” and the systematic exclusion of Indigenous students from programs and opportunities because of mis-held beliefs that Indigenous students are “not academically ready” for post-secondary education. Within policy, it’s exemplified with the changes of tenure and promotion criteria to be inclusive of Indigenous research approaches and knowledge mobilization but the people interpreting these policies continue to value individual peer-reviewed publications in top tier journals and individual grant attainment.

Patel (2021) points to how Indigenous students and faculty, along with other BIPOC, are coveted for institutional promotional materials and communication. The optics of visual diversity, and the instances of “visible” Indigenous culture (i.e., wearing of traditional regalia during graduation, holding drums or feathers) and how these students “are disproportionately asked to provide input and even leadership on committees and task forces results in their identity being “... treated in a sense as commodities” (p. 39). Of course, having Indigenous students at various committees and decision-making tables is important, Patel (2021) doesn’t dispute that. However, what she does point out and I have witnessed over and over within the Canadian post-secondary system, is the extractive nature of these engagements, with institutions simply taking knowledge and wisdom from students as uncompensated labor. Such leadership opportunities. While potentially valuable, more often than not tend to distract students from the opportunity to focus on being a university or college student—to study and learn. Albeit they are definitely getting an applied education in systemic settler colonialism and the ways their presence can be tokenizing and homogenizing all at the same time. The costs are emotional, physical, intellectual, and spiritual—and we see more Indigenous students taking a break from their studies to recover from burn out due to their leadership activities on campus. This is particularly evident when their voices are raised—the colonists around the room may nod in agreement to the “articulate way” Indigenous students passionately speak to a concern, yet continue to respond with silence and, most hurtfully, a firm lack of action.

Transforming Indigenous education through good relations

Systemic and societal change is possible—there are examples of program and policy change that show how systemic and structural change has been occurring within the vision of Indigenous education. However, being mindful of whose story is being told (or not told) requires Indigenous peoples and those allies working alongside us in this work to sustain systemic change, in addition to advancing the vision through sheer determination and persistence. As broader societal movements and calls for change (e.g., #metoo, #BLM, #LandBack, #reconciliation, #decolonization) resound throughout our post-secondary campuses, hallways, and classrooms, we must see the opportunity to disrupt settler colonialism. In these movements and calls for change, I think about the Indigenous students, staff, and faculty who are instigating change—calling on transformation so the next generation of Indigenous peoples who come to our institutions do not have to encounter resistance and racism to who they are as Indigenous peoples. As Leigh Patel (2021) also recounts how settler colonialism is enacted in her work, I also want to clearly name and identify the politics of distraction that exists and the necessity of being hypervigilant to what is happening both institutionally and within society, responding to incidents of racism, ensuring Indigenous voices are at decision making tables, on hiring committees, supporting students, advancing policies, interrupting colonial practices, and the list goes on. This largely less-valued work continues on-top of my day-to-day contractual (?) duties and responsibilities. These are just a few colonial tactics at play that result in real burnout and lateral violence against Indigenous peoples and their allies who work to disrupt and transform systems within ongoing colonialism.

At times, one may even question if one can be in good relations in ongoing settler colonialism? That is the crux of the issue at hand in Canadian higher education, a system built upon the values of settler colonialism-displacement of Indigenous peoples from their lands and ways of being in the name of individuality, profit, land ownership, and resource extraction. At times, I feel overwhelmed by the responsibilities placed on Indigenous peoples in the academy to transform an unengaged and resistant colonial institution. It is relentless work, and even as a tenured professor, the costs of the work weigh on my physical, emotional, and spiritual well-being which leaves little energy to do the intellectual work that is my heart work. Yet, in these moments of frustration, anger, and resignation, I think of those who have gone before me to chart a path of Indigenous education by and for Indigenous peoples and then I think of those who are yet to come—and wanting them to have an educational experience that enriches their entire being and supports the development of their gifts in the ways that honor who they are as Indigenous peoples. I think of the inherent obligations and responsibilities that are part of our work of being in these places and the intergenerational mentoring that occurs when the next generation of students is supported by those who have come before them. I see the hope along with the pain of system change—it is the hope and the willingness to preserve knowing it isn’t about just one person—but a peoples, generations of families and communities that need us to do better by and for them. This passion ignites the words you read here, and my passions ultimately guide my thinking to the practicality of what we need to do to enact transformation through the interconnectedness of relationships and the whole being (emotional, physical, intellectual, and spiritual).

If we were to see the interconnectedness of all the changes needed for transformation, a few aspects that matter within post-secondary education become apparent: *Representation, Authority, Voice, and Resources* to have Indigenous representation across all aspects of the institution and *Indigenous-informed policies and practices* that align with Indigenous ways of knowing, being and doing, so whether it’s processing financial payment, whether it’s admission criteria, whether it’s graduation criteria, language requirements, all those parts, you need to have wholistic practices that align with supporting Indigeneity across the institution. This is across our governance structures, academic programs and student services, curriculum and pedagogical practices, and research, including

knowledge mobilization. From these elements of transformation we need to ask ourselves: How do we shift the way we do it to actually impact the mindsets, some would call their hearts, of the people and then their practices? Fundamentally-it comes down to building and sustaining good relations. Some of this system change has to begin with the people because you have to be able to shift their mindsets, practices, and what is valued to be able to then have them take up the work of not only interpreting policies but also rewriting policies. Being in good relations with people also connects to representation- "nothing about us without us" means ensuring Indigenous representation across the institution from senior leadership to the students and at every governance table within the institution. Post-secondary institutions need individuals on the ground willing to do the program development and delivery and they also need those who are on the ground doing the curriculum development along with those who are addressing system wide policies that support Indigenous education and relationships with the Indigenous peoples whose traditional territories the institutions reside upon.

In finishing these words, I remain hopeful that my being in good relations with post-secondary institutions will eventually develop a reciprocity where the institution in turn accepts the responsibility to be in good relations with me. That recognizing the work of changing the whole system involves all of us and then figuring out which pieces we connect to, and where we can help support, change, transform, or even dismantle. The 4Rs, respect, relevance, reciprocity, and responsibility, are then supported by the actionable 4Ps—people, policies, programs and practices. Being in good relations within your place, within your context, means knowing who your people are and seeing systemic change as intentional capacity building supporting intergenerational mentoring to fully realize good relations (e.g., who's coming up behind you that you need to be mentoring and encouraging and supporting).

I noted at the outset that as the reader- whether settler or Indigenous—you were being asked to actively reflect and consider your own actions and commitments to transforming Indigenous education both individually and as part of a collective community. I cannot tell you what it means to be in good relations in these efforts because it is up to you to live your own relations. Yet I leave you here with the questions to guide your own learning: What have you done to decolonize the institutions where you live, work and learn? What can you do to live in good relations with your Indigenous colleagues and the Indigenous communities within and outside your institution?

References

- Aluli-Meyer, M., 2006. Changing the culture of research: an introduction to the triangulation of meaning. *Hülli* 3, 263–279.
- Archibald, J., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. UBC Press, Vancouver.
- Archibald, J., Selkirk Bowman, S., Pepper, F., Urion, C., Mirenhouse, G., Shortt, R., 1995. Honoring what they say: post-secondary experiences of First Nations graduates. *Can. J. Native Educ.* 21, 1–247.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. Purich, Saskatoon, SK.
- Battiste, M. (Ed.), 2000. *Reclaiming Indigenous Voice and Vision*. UBC Press, British Columbia.
- Battiste, M., 1998. Enabling the autumn seed: toward a decolonized approach to aboriginal knowledge, language, and education. *Can. J. Native Educ.* 22, 16–27.
- Battiste, M., Barman, J. (Eds.), 1995. *First Nations Education in Canada: The Circle Unfolds*. University of British Columbia, British Columbia.
- Brayboy, B.M.J., 2005. Toward a tribally critical race theory in education. In: *The Urban Review Issues and Ideas in Public Education*, vol. 37, pp. 425–446. <https://doi.org/10.1007/s11256-005-0018-y>.
- Brayboy, B.M.J., Fann, A.J., Castagno, A.E., Solyom, J.A., 2012. Postsecondary education for American Indian and Alaska natives. *ASHE High. Educ. Rep.* 37, 1–154.
- Carr-Stewart, S., 2019. *Knowing the Past, Facing the Future: Indigenous Education in Canada*. Purich Books, Vancouver.
- Carr-Stewart, S., 2001. *Perceptions and Parameters of Education as a Treaty Right within the Context of Treaty 7* (PhD). University of Alberta, Edmonton, AB.
- Comtassel, J., 2012. Re-envisioning resurgence: Indigenous pathways to decolonization and sustainable self-determination. *Decolonization: Indigeneity, Education & Society* 1 (1), 86–101. <https://jps.library.utoronto.ca/index.php/des/article/view/18627/15550>.
- Cote-Meek, S., 2014. *Colonized Classrooms: Racism, Trauma and Resistance in Post-secondary Education*/Sheila Cote-Meek. Fernwood Publishing, Halifax; Winnipeg.
- Cote-Meek, S., Moeke-Pickering, T. (Eds.), 2020. *Decolonizing and Indigenizing Education in Canada*. Canadian Scholars, Toronto.
- Deloria, V., 2001. American Indian metaphysics. In: Deloria, V., Wildcat, D.R. (Eds.), *Power and Place: Indian Education in America*. Fulcrum Resources, Golden, CO, pp. 1–6.
- Fisher, D., Rubenson, K., Bernatchez, J., Clift, R., Jones, G., Lee, J., MacIvor, M., Meredith, J., Shanahan, T., Trotter, C., 2006. *Canadian Federal Policy and Postsecondary Education*. The Centre for Policy Studies in Higher Education and Training (CHET), Vancouver, BC.
- Gaudry, A., Lorenz, D., 2018. Indigenization as inclusion, reconciliation, and decolonization: navigating the different visions for indigenizing the Canadian Academy. *Alternative* 14, 218–227. <https://doi.org/10.1177/11771177180118785382>.
- Hare, J., Davidson, S.F., 2020. Learning from indigenous knowledge in education. In: *Visions of the Heart: Issues Involving Indigenous Peoples in Canada*. Oxford University Press, Don Mills, ON, pp. 203–219.
- Kirkness, V.J., 1998. Our peoples' education: cut the shackles; cut the crap; cut the mustard. *Can. J. Native Educ.* 22, 10–15.
- Kirkness, V.J., Barnhardt, R., 1991. First Nations and higher education: the four R's- respect, relevance, reciprocity, responsibility. *J. Am. Indian Educ.* 30, 1–15.
- Kuokkanen, R., 2007. *Reshaping the University: Responsibility, Indigenous Epistemologies and the Logic of the Gift*. UBC Press, Vancouver.
- The Master's tools will never dismantle the Master's house. In: Lorde, A. (Ed.), 1984. *Sister Outsider: Essays and Speeches*. Crossing Press, Berkeley, CA, pp. 110–114.
- Marker, M., 2004. The four Rs revisited: some reflections on first nations and higher education. In: Andres, L., Finlay, F. (Eds.), *Student Affairs: Experiences in and Through Canadian Post-Secondary Education*. UBC Press, Vancouver, pp. 171–188.
- Mika, C., 2017. Indigenous education and the metaphysics of presence. In: *A Worlded Philosophy*, first ed. Routledge, London.
- National Indian Brotherhood, 1972. *Indian Control of Indian Education*. National Indian Brotherhood, Ottawa.
- Parent, A., 2017. Visioning as an integral element to understanding indigenous learners' transition to university. *Can. J. High Educ.* 47, 153–170.
- Patel, L., 2021. *No Study Without Struggle: Confronting Settler Colonialism in Higher Education*. Beacon Press, Boston.
- Pidgeon, M., 2016a. More than a checklist: meaningful indigenous inclusion in higher education. *Soc. Incl.* 4, 77–91.
- Pidgeon, M., 2016b. Aboriginal student success & aboriginal student services. In: Hardy Cox, D., Strange, C. (Eds.), *Serving Diverse Students in Canadian Higher Education: Models and Practices for Success*. McGill University Press, Montreal, pp. 25–39.

- Pidgeon, M., 2014. Moving beyond good intentions: indigenizing higher education in British Columbia universities through Institutional responsibility and accountability. *J. Am. Indian Educ.* 53, 7–28.
- Pidgeon, M., 2008. Pushing against the margins: indigenous theorizing of “success” and retention in higher education. *J. Coll. Stud. Retent.* 10, 339–360. <https://doi.org/10.2190/CS.10.3>.
- Pidgeon, M., 2005. Weaving the story of Aboriginal student services in Canadian universities. *Communique* 5, 27–29.
- Pidgeon, M., Rogerson, C., 2017. Lessons learned from Aboriginal students' housing experiences: supporting Aboriginal student success. *J. Coll. Univ. Student Hous.* 44, 48–73.
- Pidgeon, M., Tobin, J., Setah, T., Leveille, A., Dunn, D., Johnson, K., Bubela, T., 2019. Looking Forward. Indigenous Pathways to and through Simon Fraser University. *Wholistic Understandings of Access, Transition, and Persistence*. Simon Fraser University, Burnaby.
- Restoule, J.-P., 2011. Looking for a way in: Aboriginal youth talk about access to university in Ontario. *Can. J. Native Stud.* 31, 47–62.
- Restoule, J.-P., 2005. Education as healing: how urban Aboriginal men described post-secondary schooling as decolonising. *Aust. J. Indig. Educ.* 34, 123–131.
- Restoule, J.-P., Mashford-Pringle, A., Chacaby, M., Smillie, C., Brunette, C., Russel, G., 2013. Supporting successful transitions to post-secondary education for indigenous students: lessons from an institutional ethnography in Ontario, Canada. *Int. Indigenous Policy J.* 4, 4.
- Royal Commission on Aboriginal Peoples, 1996. Report of the Royal Commission on Aboriginal Peoples. Vol. 3, *Gathering Strength*. Gathering strength. Ottawa, Ont. Indian and Northern Affairs Canada, Ottawa, Ont.
- Shotton, H., Lowe, S.C., Waterman, S.J. (Eds.), 2013. *Beyond the Asterisk: Understanding Native Students in Higher Education*. Stylus, Sterling, VA.
- Smith, G., 1997. *The Development of Kaupapa Maori: Theory and Praxis* (PhD). University of Auckland, Auckland, NZ.
- Stonechild, B., 2006. *The New Buffalo: The Struggle for Aboriginal Post-secondary Education in Canada*. University of Manitoba Press, Winnipeg.
- Styres, S.D., 2017. *Pathways for Remembering and Recognizing Indigenous Thought in Education: Philosophies of lethi'nihstenha ohwentsia'kekha* (Land). University of Toronto Press, Toronto, Buffalo, & London.
- Truth and Reconciliation Commission of Canada (TRC), 2015. *Honoring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. TRC, Ottawa.
- United Nations, 2007. *United Nations Declaration on Rights of Indigenous Peoples (UNDRIP)* [WWW Document]. United Nations. URL: https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf. (Accessed 5 April 2020).
- Waterman, S.J., Lowe, S.C., Shotton, H.J. (Eds.), 2018. *Beyond Access: Indigenizing Programs for Native American Student Success*. Stylus Publishing, LLC, Virginia.

The importance of “seeing oneself” in Australian universities: representations and reflections of Indigeneity in higher education

Bronwyn Fredericks, Tracey Bunda, Katelyn Barney, and Abraham Bradfield, The University of Queensland, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	109
Universities, whiteness and racism	110
Relationality and creating a sense of belonging	112
Designing frameworks that embrace Indigenous visibility	114
Conclusion	115
Acknowledgments	116
References	116

Introduction

In Bennett's 1988 artwork, *Echo and Narcissus* (McLean and Bennett, 1996), Narcissus, a white fetal-like figure devoid of defining features, peers into a reflective pool of water out of which Echo, an Aboriginal man emerges caressing Narcissus' face. *Echo and Narcissus* provides an insightful metaphor for some of the challenges relating to reconciliation and truth-telling in colonized Australia; asking, what is reflected when the white colonizer inspects themselves while on Indigenous sovereign land?

In this chapter, we take the university as akin to Narcissus in that it is an institution that mirrors white Anglo-centric ontologies, epistemologies, and power structures. When Aboriginal and Torres Strait Islander people gaze into the mirror that reflects the university, they rarely see echoes of themselves. Negative and racist experiences within the university system—or a failure to provide adequate support networks—can harbor feelings of alienation or dispossession, making universities unwelcoming places for many Aboriginal and Torres Strait Islander staff, students, and visitors (Fredericks, 2020). Here, we discuss how some Indigenous people have difficulty in seeing themselves in the social and built environment in Australian universities impacting enrollment, retention, and completion rates. We argue that to improve educational outcomes for Indigenous peoples; the staff, policies, and resources that govern behaviors and strategic decisions must reflect Aboriginal and Torres Strait Islander perspectives.

Enrollment of Aboriginal and Torres Strait Islander students within Australian universities have doubled since 2008 (Fredericks et al., 2022). This, however, remains varied with some universities achieving better results compared to others. Enrollment numbers remain disparate vis-a-vis the non-Indigenous population with enrollments of Aboriginal and Torres Straits Islander peoples in domestic bachelor courses comprising just 0.74% of the entire cohort (Fredericks et al., 2022). The Review of Higher Education Access and Outcomes for Aboriginal and Torres Strait Islander People: Final Report (Behrendt et al., 2012) highlights the need to shift away from a focus on access toward successful educational outcomes for Indigenous students and to identify what strategies universities with consistently high success rates employ. Further, a report funded by the National Center for Student Equity in Higher Education (NCSEHE) and led by scholars at the University of Queensland (Fredericks et al., 2022), has demonstrated that while completion rates of Indigenous students in higher education continues to lag national trends, institutions with Indigenous-led support centers, which are targeted toward and led by Indigenous peoples, are outperforming those that do not have them.

Universities Australia's *Indigenous Strategy for 2022–2025* (Universities Australia, 2022) highlights that access is not as a major priority as it once was—although access, especially in regional areas, persists as a barrier in need of redress—but rather argues that focus should turn toward retention and completion rates. As of 2019, completion rates among Indigenous students who graduated within 9 years of their undergraduate degree stands at 47% (Universities Australia, 2022: 24). This is a slight improvement from previous years but falls significantly short of reaching parity with the national cohort who have a completion rate of 74%. This raises questions such as, (1) what factors are contributing toward Indigenous students' decision to discontinue their studies; and (2) would completion rates improve if these barriers are identified and redressed within the university system?

To unpack these questions, we drawn on a recent report (Fredericks et al., 2022) which involved 66 qualitative interviews conducted across five Australian universities with Indigenous staff, enrolled students, and students who had discontinued their studies. The study's aim was to quantify data to provide evidence upon which greater understandings of the factors that impact completion and retention rates could be gained. It presents data that contributes to the development of effective strategies aiming to help Aboriginal and Torres Strait Islander peoples achieve greater academic success.

To inform our discussion of the direction higher education policy is currently heading, we also consider the policy framework outlined in Universities Australia's *Indigenous Strategy for 2022–2025*. Universities Australia is the peak body for Australian universities, representing 39 institutions nationwide. Their Indigenous Strategy outlines existing trends, identifies the areas in need of improvement, and reinforces a commitment to making the structural changes necessary to improve educational outcomes. We

finally draw on research conducted during the development of The University of Queensland’s *Aboriginal and Torres Strait Islander Design Framework Engagement Report* (Go-Sam et al., 2021). This qualitative research documents the views Aboriginal and Torres Strait Islander students and staff have toward the built-environment and infrastructure on the “sandstone” university campus.

These documents—along with a growing body of literature on critical race theory and place—are collectively encouraging a dialog around creating a university environment that is equally representative and welcoming of Indigenous peoples.

Universities, whiteness and racism

Universities are socially constructed institutions that reflect and produce social realities and hierarchies that are predominantly based on western Anglo-centric outlooks (Foley et al., 2015; Fredericks, 2020; Fredericks et al., 2019b; Nakata et al., 2019). As sites of knowledge production, universities play critical roles in the reclamation and re-territorialization of place as a “white possession” (Fredericks, 2009; Moreton-Robinson, 2007, 2021; Stewart et al., 2021). In other words, universities are sites where notions of *terra nullius*—the legal fiction that claimed the land we now know as Australia as “uninhabited”—are reinforced via its buildings, landscaping, systems of governance, placenames, and monuments that attempt to erase Indigenous cultures and presence (Butler, 2000; Fredericks, 2020; Watson, 2014).

In colonized spaces where white sovereignty is continuously reiterated as a means to take and retain power, Indigenous peoples are often characterized as an “unsettling presence” (Fredericks, 2020; Fredericks et al., 2019a; Kinnane et al., 2014; Williamson and Dalal, 2007). When Aboriginal and Torres Strait Islander people make their presence known by occupying senior positions or working in ways that align with Indigenous cultural protocols, it can expose the pervasiveness of white possession, the falsity of white sovereignty, and expose the ongoing impact of colonialism (Hokowhitu et al., 2020; Moreton-Robinson, 2015, 2021; Gunstone, 2009). The disregard of Indigenous cultures and perspectives—as well as decisions to challenge and resist to the status quo (Nakata, 2013)—can make the university a culturally unsafe environment for many Aboriginal and Torres Strait Islander people. In some regards, the university is a place where Indigenous peoples are treated as anomalies while on their own Country/Land or the countries/lands of other Indigenous peoples (Best, 2018; Fredericks, 2020; Moreton-Robinson, 2015; Newton, 2021). Fredericks (2020: para 7) has written on how,

Everything linked to the university, including its iconography, selection and display of artworks, photographic imagery used in marketing materials, and events held in and around the sandstone buildings, are evidence of their Anglo-centrality. The very buildings communicate cultural dominance. Cultural dominance is also communicated through the university’s physical infrastructure, systems, structures, budget allocations, and everyday interactions.

Indigenous peoples are unable to see themselves in the surrounding environment as their perspectives are rarely considered first or taken as the point of departure. Indigenous outlooks therefore need to be built into and flow through the strategies, policies, curricula, and built environment that make up a university. They must become places that reflect the Indigenous cultures and lands upon which they stand (Bunda et al., 2012). To achieve this however, the structural and institutional racism that is reflected and produced by universities needs to be confronted head-on.

Racism and discrimination toward Aboriginal and Torres Strait Islander peoples within the university significantly hinders Indigenous people’s ability to see themselves and develop a sense of belonging within the institution (Bunda, 2017a,b; Bunda et al., 2019; Uink et al., 2022). Acts that deny, disregard, or dismiss Indigenous presence communicate that they are spaces in which Indigenous peoples are neither welcomed nor seen to naturally belong (Leroy-Dyer and Heckenberg, 2022). Within higher education, 75% of Indigenous staff have experienced racism in the workplace (Universities Australia, 2022: 46) while experiences of racism by students are highly reported (Brown and Shay, 2021; Plater et al., 2020; Taylor et al., 2019). This suggests that racism is a systemic problem and deeply engrained.

Indigenous students and staffs’ ability to navigate, succeed, or be promoted within a university is often predicated on their capacity to *fit in* with the expectations and pedagogies of the overwhelmingly white institution (Bargallie, 2020; Bunda, 2014; Moreton-Robinson, 2020). Aboriginal and Torres Strait Islander peoples may not always see the university as a reflection of themselves, but they are expected to act as if it were. Racist experiences with staff, lecturers, tutors, and feelings of being unable to address concerns within a safe environment contribute to retention and completion rates (Fredericks et al., 2022). Dominant western pedagogies and epistemologies, for example, often conflict with Indigenous outlooks. Lecturers and/or tutors fail to consider how their positionality may be contested by others (Moreton-Robinson, 2020; Rigney, 1999, 2006; Fredericks et al., 2022; Nakata, 2011).

Creating Indigenous visibility within the university, however, entails more than tokenistic gestures or lumping Indigenous content into a single week. In some cases, Indigenous students can feel uncomfortable with being singled out or treated as “experts” who are expected to speak for all Indigenous peoples (Fredericks et al., 2022). These classes fail to recognize the diversity of Indigenous knowledge and cultures and the many ways in which people engage their Indigeneity (Sarra, 2014; Nakata, 2012; Waters, 2017). The lecture theater or tutorial room can quickly turn into an accusatory, voyeuristic, and culturally unsafe space in which micro-aggressions and blatant racism play out, while Anglo-centric dominance is reinforced and goes uncontested (Brown, 2010).

Decisions to make one’s Indigenous presence visible by pushing back or “talkin-up” to the system can be a daunting prospect, especially when an Indigenous person is in the minority or are the only Aboriginal or Torres Strait Islander person in the room

(Moreton-Robinson, 2020; Bargallie, 2020). Contesting the status quo within the university or challenging dominant structures can have significant ramifications on education, career pathways, and the prospect of attaining promotions and senior positions (Bargallie, 2020; Cassells et al., 2022; Coates et al., 2022a; Tamer, 2022; Trudgett et al., 2017). The risk and toll of raising concerns of structural racism to personnel who are understanding and responsive to colonial contexts is perhaps a product of a lack of Indigenous representation at senior and executive levels.

Having senior leadership will encourage top-down reform, rather than placing the burden on Indigenous staff and support centers. University Australia’s Indigenous Strategy has committed to addressing racism within the university sector, advising universities to take a zero-tolerance stance. Strategies include creating more leadership and senior positions for Indigenous peoples while confronting institutions’ complicity in maintaining systemic barriers and racial prejudice. This means that institutions and the executive teams and boards that determine universities’ strategies and policies need to remain receptive to the prospect of being challenged and recognize the history and ongoing impact of colonization (Foley et al., 2015, 2019). When Indigenous staff and students raise concerns of racism or propose pedagogies or methodologies that may not align with western outlooks, the visibility of senior staff is essential as it communicates that the institution values and takes Indigenous perspectives seriously.

Senior Indigenous positions also need to flow into other roles, either sideways, or upwards. Indigenous people within Indigenous positions can become stuck within a systemic racialized structure that blocks them from any further career development or promotion. Universities which do not address these structural biases are thus ensuring they will never have an Indigenous person who is a Senior Executive or Vice-Chancellor who has progressed via the Indigenous leadership positions (see Coates et al., 2022b for more discussion).

In addition to experiencing overt forms of racism within universities, Aboriginal and Torres Strait Islander people often have the burden having to explain and justify their actions or socio-cultural circumstances to non-Indigenous staff. For example, asking for time off or extensions for cultural reasons like mortuary rites, known as “sorry business”. This is not only emotionally draining but can cause discomfort, alienation, and in some circumstances resurface trauma (Grieve-Williams, 2008). The inability to feel safe, welcomed, or represented within governing structures means that some students avoid accessing support from non-Indigenous staff due to the lack of understanding they receive and the discomfort of having to explain and teach non-Indigenous staff on topics relating to Indigenous cultures. This can hinder a student’s education, affect their grades, or contribute to a student’s decision to discontinue their studies.

It is widely argued that cultural competency training and increasing cultural awareness among non-Indigenous peoples can improve outcomes for Indigenous staff and students (Fredericks and Bargallie, 2016, 2020a,b). More universities have cultural competency training and have taken steps to alleviate some of burdens placed on Indigenous staff to correct or withstand culturally inappropriate behavior within the institution (Universities Australia, 2020). It is unrealistic, however, to believe that the policies and structures that inform the negative experiences of Aboriginal and Torres Strait Islander peoples will be undone via the implementation of cultural training or Reconciliation Action Plans alone. We recognize that if not done correctly, cultural competency can become a colonizing tool through which non-Indigenous peoples aim to “manage” Indigenous peoples within the institution. It can have the adverse effect of empowering non-Indigenous peoples with the misguided perception that they can create change through understanding alone, while not addressing systemic racism or opening new opportunities and positions for Aboriginal and Torres Strait Islander staff.

Having said that, initiating difficult conversations can improve racial literacy and generate greater awareness and accountability (Bargallie, 2020; Brown et al., 2021; Fredericks and Bradfield, 2021a). Effective cultural competency should be informed by “critical race theory” which is a framework that seeks to unpack race as a determining factor of the ways in which the world is organized and responded (Bargallie, 2020; Delgado and Stefancic, 2017; McLaughlin and Whatman, 2011). Becoming “competent” in another’s culture, requires more than simply learning about a so-called “exotic other” within a de-politicized space (Kowal, 2010). It necessitates a deeper engagement with ongoing power structures and asks participants to reflect on the behaviors and policies that prevent them from understanding and recognizing Indigenous perspectives and sovereignties (Fredericks and Bargallie, 2016; Morrison et al., 2019). Where Indigenous people struggle to be seen within the university system, non-Indigenous peoples (and the structures they govern) often fail to see themselves in an open and honest way, veiling how their actions are the products and beneficiaries of colonial structures that privilege and reward Eurocentric outlooks (Bargallie, 2020; Gillborn, 2015; Moreton-Robinson, 1998).

In their evaluation of a cultural competency pilot program in central Queensland, Fredericks and Bargallie (2020a) challenge the “myth of safe space” within cultural competency frameworks. They argue that while courses should not be accusatory or aggressive, they must recognize that “safety” often correlates with dominant white groups avoiding uncomfortable topics and refusing to do the “hard work” of dismantling racism. Exposing what DiAngelo (2022) terms “white fragility”, dialogs about race that are facilitated and dominated by white voices can result in defensiveness that reinforces Eurocentric supremacy. Topics are avoided or watered down under the guise of creating a “safe space” in which the discomfort of racism and one’s complicity and profiting from it are negated (Anderson and Riley, 2021). Cultural competency training with non-Indigenous peoples should recognize that comfort within settler spaces cannot be assured; that discomfort and a lack of safety is something Indigenous peoples live with daily; and that discomfort is part of the learning process (Bond, 2014).

Cultural competency therefore means holding a mirror that exposes the systemic and institutional racism that impacts retention and completion rates. It calls for greater consideration of how structures and systems within the university can shift in ways that enable Indigenous staff and students to see themselves in the surrounding environment and develop a sense of belonging. As we discuss in further detail below, relationality, lies central to such recognition.

Relationality and creating a sense of belonging

Creating a university environment in which Aboriginal and Torres Strait Islander people can see themselves requires frameworks that are flexible, informed, and responsive to Indigenous peoples’ lived experiences. Studies have shown that increasing the visibility of Indigenous peoples and cultures can help improve completion rates, the retention of staff, and lead to greater advancement/promotion to senior leadership roles. Acting as cultural hubs, Indigenous support centers have the staff and resources needed to help student’s liaise with university faculties and departments—assisting in the negotiation of deadlines, extensions, systems, and time management. In faculties, schools, and educational settings where Indigeneity is often shrouded, Indigenous support centers are creating visibility that challenges their denial.

The support, dedication, and services provided by these centers, however, must flow throughout all faculties and schools. Cultural competency training, as discussed above, can assist in educating non-Indigenous staff, tutors, and lecturers of some of the factors that shape how and the circumstances under which Indigenous peoples learn. This can create a sense of shared understanding that will become the basis of collectively working through challenges that may arise. Competency, however, must be backed up by capacity building and increasing the numbers of Indigenous staff.

A sense of belonging is also created by having systems that are flexible and understanding of the sociocultural and socioeconomic circumstances of Indigenous students. Recent research, for example, found that Indigenous peoples tend to take longer to complete their degree compared to non-Indigenous students (Fredericks et al., 2022; Universities Australia, 2020). There are numerous explanations for this, including balancing commitments to family, employment, or having to take time off for mental health, cultural business, or other sociocultural reasons. The prolonged time it takes to complete a degree, however, should not be seen as a negative reflection of a student’s academic progress.

While the greater time it takes to complete a degree may be an indicator of increased risk of dropout among non-Indigenous students, for Indigenous peoples it can be conducive to their educational journey providing it is backed with support, infrastructure, and resources. Completion rates for Indigenous students who completed their undergraduate degree within 5–6 years at universities that have Indigenous support centers, for example, exceed the rates of non-Indigenous students. Therefore, more needs to be done to assist and support students over the course of their entire study—regardless of how long it takes.

Family and relationality also remain significant drivers of completion rates for Indigenous peoples (Oliver et al., 2013; Roche-couste et al., 2017). Family can be sources of motivation but can also potentially impact retention rates, as family and cultural commitments takes precedence over study. Indigenous peoples are community-orientated and motivated by an awareness that one’s knowledge and skills are best placed when they are for the benefit and advancement of their families. Building strong relationships with staff, peers, and other communities can help students navigate the university while also improve completion rates (Milne et al., 2016).

The ability to see one’s family, alongside creating an environment that fosters relationality, can contribute to a sense of belonging that improves education outcomes and experiences for Aboriginal and Torres Strait Islander people. The Fredericks et al. (2022) report demonstrates that growing numbers of Aboriginal and Torres Strait Islander peoples are completing their degrees and encouraging members of their family to do the same. This is despite past trends the demonstrate that Indigenous students are often the first in their family to attend university (Barney, 2013; Oliver et al., 2013). Many of the students interviewed as part of the research drew inspiration from family and friends who studied before them (Fredericks et al., 2022).

In some cases, family members of different generations were studying at the same institution simultaneously. Indigenous peoples position themselves through their relationships with kin and ancestors who are embedded in place and within the memory of place. Universities are Indigenous places where memory and lived experiences reside. Having multiple generations of Aboriginal and Torres Strait Islander people study at an institution and building a strong cohort of alumni can assist in making the space familiar and representative of Indigenous knowledges, cultures, languages, and histories. This emphasizes the importance of memory, place, and the significance of sharing stories of achievements and challenges relating to university life.

The *Indigenous Strategy 2017–2020* (Universities Australia, 2017) highlights “building effective community partnerships” as one of its key priorities to improving outcomes within the university sector. This is vital as Aboriginal and Torres Strait Islander people see themselves through their relationships with and within community. The University of Queensland, as many other universities, has recognized this, building “Community Engagement and Involvement” into its 2021–2025 *Aboriginal and Torres Strait Islander Research and Innovation Strategy* (The University of Queensland, 2021). This component of the strategy seeks to establish a cross-disciplinary hub to create community partnerships and opportunities to fund and broaden innovative research projects.

Community engagement however must flow through the entire university experience from *outreach programs* where students are encouraged to enroll; within the university via *acknowledgments* to Country, cultural literacy training, and cultural events; within *curriculum* that positions education from Indigenous perspectives; and in *pathways* toward further education and careers. The University of Melbourne, for example, has created an innovative Indigenous pedagogy as part of its Indigenous Studies program, integrating community and place-based knowledges within its curriculum. Each lecture seeks to incorporate a traditional owner who conducts a masterclass on a place-paced topic. In addition to building links and connections to community, initiatives such as these are creating and archiving knowledges that are accessible to future students and communities—making Indigenous visibility a long-term commitment that will span generations.

Integrating traditional knowledges within curriculum does not only assure academic rigor—providing innovative insights across all disciplines—but rather places tertiary education within the learner’s lived experience (Shay and Oliver, 2021). Education becomes less removed from place or approached as a discipline where one casts their objective gaze upon a removed “other”

and is recognized as being inseparable to place and culture. Indigenous students are more likely to see themselves in the research and be motivated by how education can benefit themselves, their families, communities, and Country (Barney, 2013; Bunda, 2014; Fredericks et al., 2015; Pagnini et al., 2014).

Indigenous support centers and units create spaces in which staff and peers act as a “de-facto” family, helping each other navigate the challenges associated with higher education (Andersen et al., 2008; Pechenkina et al., 2011; Rochecouste et al., 2017; Uink et al., 2022). Uink et al. (2022, p. 40), notes that a number of the students they interviewed felt “that there was a community of people on campus who were there for support”. Indigenous peoples are already impacted by higher rates of socio-economic disadvantages compared to the non-Indigenous population, particularly in relation to finance, income, and affordability of food, rent, electricity, transportation, and computers and technology (Fredericks and Bradfield, 2021b; Lowitja Institute, 2021). This is further exacerbated for students who balance study with work, education, and family responsibilities.

Indigenous support centers alleviate some of these pressures. For example, to curb food insecurity, they often stock food that students may access while also providing spaces, resources, and kitchens. This provides students with a space of their own with designated areas where they can concentrate on their education; do not have to worry about accessing computers or photocopiers; have the additional support of tutors, peers, and mentors; and can socialize and unwind (Fredericks et al., 2022; Kinnane et al., 2014; Shalley et al., 2019).

Creating a welcoming environment is especially important for 1st year enrollments. Indigenous students who attend universities with established Indigenous support centers were more likely to return after their 1st year—going on to complete their degree within the given timeframe. While negative experiences may be a motivating factor for regional students to return home and discontinue their studies generally, universities with Indigenous support centers were able to overcome such barriers achieving greater retention rates due to the support it provided Indigenous students.

Scholarships and financial support for students also opens doors and creates opportunities that would otherwise obstruct Aboriginal and Torres Strait Islander people from accessing tertiary education. Offering emergency financial support, food baskets, and providing access to computers and internet during the pandemic were particularly crucial to supporting students, many of whom experienced heightened financial strain and other social pressures (Fredericks et al., 2021). The pandemic, along with the necessity to limit face-to-face interaction, significantly impacted relationships between and among Indigenous staff and students, placing retention at risk due to disruptions to support networks. Many Indigenous support staff made conscious efforts to bridge the challenges posed by the pandemic, continuing social events online to maintain connection and wellbeing, while increasing access to online tutoring and academic support. Creating visibility and a sense of belonging within universities, therefore, refers to having the support networks in place that assist students and have their interests at heart.

Although financial support alleviates some of the financial pressures associated with studying, Fredericks et al. (2022) found no direct correlation between the attainment of Commonwealth Scholarships and completion rates. This suggests that retention is more likely determined by the social and cultural make-up of a universities’ environment, rather than only financial constraints. While greater investment in scholarships opens the door to tertiary education, Indigenous support centers play critical roles in retaining students and providing a safe and conducive educational environment. Both are therefore vital and necessary.

Universities and funding bodies, however, must be willing to invest in the staff and resources needed to increase the reach of Indigenous support centers and infrastructure. As reported by Universities Australia, in some cases, government funding intended as supplementary top-ups for pre-existing Indigenous support centers and networks have been used to fund universities’ entire Indigenous engagement strategies (Universities Australia, 2022: 23). In some universities, the facilities which house Indigenous support centers are in complete contrast to the facilities which house the chancellor, or the engineering, business, or anthropology programs. This is despite university leadership, and councils/senates saying how much they value Indigenous peoples and cultures and the Reconciliation Action Plans they have developed. Welcoming and celebrating Indigenous visibility must translate to meaningful outcomes and be more than tokenistic lip-service.

Universities must be willing to invest adequate funds for appropriate facilities, and to build strong support networks run by Indigenous staff and students. More staff and resources are needed to meet increased demand and to avoid burnout (Barney, 2018). Although the number of Aboriginal and Torres Strait Islander staff has nationally doubled since 2005, it remains disproportionate to the entire Indigenous population. Where Universities Australia’s previous Indigenous Strategy (2017–2021) focused on establishing frameworks to initiate change, the new iteration centers on delivery and execution across the entire university system. A whole of university approach, with transparent and measurable targets, is needed so that processes can be evaluated, monitored and successful initiatives and strategies shared. Indigenous perspectives and targets should be built into the foundation of a universities’ ethos, included in its corporate plans, research strategies, procurement policies, design protocols, teaching and learning, and employment strategies.

Indigenous centers and units create culturally safe spaces for students that help build a sense of community and belonging among staff and peers which extend beyond notions of access and individual success. This commonality is shared with other Indigenous peoples in Canada (Cote-Meek, 2014; Battiste, 2013; Tanaka, 2016), the United States (Brayboy et al., 2012; Huaman and Brayboy, 2017; Waterman et al., 2018), and New Zealand (Bishop et al., 2010). It’s important that universities continue to provide or establish designated spaces for Aboriginal and Torres Strait Islander peoples. Their place within the university, or a student’s motivations for accessing them, however, should not come because of negative or racist experiences encountered on campus. In the Fredericks et al. report (2022), some students expressed how Indigenous centers provide a welcoming and non-judgmental environment that was separated from the ignorant and often racist encounters experienced from some staff (both academic and administrative)

and peers on campus. While Indigenous support centers will always function as safe spaces for Indigenous peoples to be themselves without judgment, this must also flow throughout the entire university system—in both its physical and ideological make-up.

Designing frameworks that embrace Indigenous visibility

In 2021, the University of Queensland released its *Aboriginal and Torres Strait Islander Design Framework* (Go-Sam et al., 2021). During the consultations that lead to the drafting of this seminal document many Indigenous participants expressed how they felt inadequately represented in the built and social environment—if represented at all. University campuses and the buildings within them are not passive receptacles but living representations and manifestations of place. Indigenous staff and students believed that the reciprocal relationships between place and Country and the cultural obligation to care for Country needed to be reflected in the built environment (McKnight, 2016), creating a greater sense of belonging. After all, university campuses remain on Indigenous Country that is filled with stories, cultural sites, and memories. Such knowledge however is either unknown or publicly hidden from view. Buildings need to have a connection to the landscape, avoid homogeneity, and reflect the diversity of place.

Indigenous knowledge derives from Country where one learns and responds to the connections between changing patterns and seasonal cycles. Knowledge is relational, emerging, and based in memory and lived experience. University campuses do not merely facilitate knowledge translation but are sources of knowledge unto themselves. Some participants of University of Queensland’s *Design Framework* spoke of how buildings “belong” to Country and that they should maintain physical connections and relationships to it e.g., via integrating elements of the landscape within its design. The landscape of a campus, inclusive of the buildings and infrastructure, has agency which is overseen and mediated by the university. With this comes a responsibility and obligation to care and nurture for Country. Indigenous staff and students therefore wanted to see the university actively pursue and include Indigenous people in design initiatives to create spaces in which Country could be cared for.

Universities are also places where Aboriginal and Torres Strait Islander peoples have excelled in both “traditional” and “western” knowledge systems. In many cases Indigenous scholars have bridging the perceived divide between the two knowledge systems (Gilbert and Bunda, 2017; Gilbert and Tillman, 2017; Synot et al., 2021). While the success of these individuals may be known within smaller networks, it is important to celebrate their achievements and contributions more widely. Pre- and postcolonial Aboriginal and Torres Strait Islander scholars and leaders should also be acknowledged and celebrated. Griffith University, for example, is considering renaming the university after Dundalli, the Dalla warrior who led the resistance in southeast Queensland (Hinchliffe, 2020). Participants of the *Design Framework* suggested that each faculty and school within the university needed to do more to showcase Indigenous excellence, highlighting how Aboriginal and Torres Strait Islander scholarship has shaped how we come to understand and respond to the world. This not only creates visibility and reassurance to Indigenous’ students, but also generates awareness among non-Indigenous peoples who, whether they are aware of it or not, are building, adding, and benefiting from Indigenous scholarship.

While we previously discussed the importance of establishing Indigenous support networks through Indigenous-led centers, consideration must also be given to the place in which these centers are located within the university. If place is taken as a communicator of social value and purpose, the space allocated to Indigenous peoples can provide rich insight into an institutions’ priorities. One staff member from the University of Queensland, for example, spoke on their objection to the space allocated to the Indigenous support center stating that they were “offended by the fact that we have been put into this building. I think it is an insult to us as First Nations people” (in Go-Sam et al., 2021: 22).

Other students spoke of how the lack of acknowledgment of Indigenous cultures and the fact that Indigenous spaces were often hidden, difficult to find, or unmaintained, communicated messages that Aboriginal and Torres Strait Islander peoples were unworthy of investment or public exposure. Some highlighted the role of threshold spaces (such as entries/exits to campus or buildings) in reiterating that universities remain on Indigenous Country. Reminding visitors of this does not only generate greater awareness of Indigenous sovereignty among non-Indigenous peoples but follows protocols of welcoming newcomers to Country not theirs.

The notion that universities are contested spaces arising out of colonialism is recognized and experienced by Indigenous peoples, daily. Many attest to the discomfort they feel toward the placenames, statues, buildings, and monuments that glorify and celebrate colonial exploitation—often at the expense and silencing of Aboriginal and Torres Strait Islander people (Bunda et al., 2012; Foley et al., 2019; Daley, 2018). Despite recent debates over the so-called “statue wars” and the place of colonial monuments in public spaces, symbols of colonialism will most likely continue to dominate sandstone universities (Gregory, 2021; Moreton-Robinson, 2022). In many cases, buildings and facades are protected by heritage laws.

Creating increased visibility of Indigenous cultures and standpoints, however, is not predicated on the erasure or “canceling” of western heritage (Nakata, 2007). Like the cultural competence training discussed earlier, Indigenous histories and truths may be exposed via engaging in transparent, open, and honest discussion about Australia’s history, including the history of the university. Recognizing the history, context, and contention of place as understood by Indigenous peoples (whether through plaques, artwork, statues, or digital media) provides an opportunity to proclaim their survivance and sovereignty while generating greater awareness among the wider public.

The practice of making Acknowledgments of Country during public addresses, events, or attached to signatures in emails and correspondence has increasingly been used to recognize Indigenous people (Pelizzon and Kennedy, 2019). While some see it as

a positive sign of reconciliation many have warned that it is problematic due its symbolic and tokenistic nature (Kowal, 2010; McKenna, 2014; Pelizzon and Kennedy, 2019). If not backed by commitments to make meaningful change, Acknowledgments of and Welcomes to Country can obscure the history and context of colonialism, providing a false sense of legitimacy or authority through which white sovereignty and the exploits that maintain its possession go unchallenged. A member of staff at the University of Queensland, for example, expressed (in Go-Sam et al., 2021: 72) that,

You cannot provide an acknowledgment of this country without deeply thinking about the arrangements that got the brick and mortars here. The theft of this land from Aboriginal people—I do want people to think about that. I want people to think about that quite deeply. I want them to be troubled by that. Because my family has been troubled by that because they have been dispossessed ... I am desiring a more honest engagement with us as Indigenous peoples, and without that troubling [there is no understanding] and there is always a superficial relationship.

Requests for Indigenous peoples to perform Acknowledgments and Welcomes to Country—particularly by schools, centers, and faculties that remain underrepresented by Aboriginal and Torres Strait Islander staff and students—can be time consuming and exhausting. As the above quote attests, to acknowledge Country is to also confront the context which led to the theft of Aboriginal land and Indigenous people’s ability to prosper, profit, and be promoted within the systems imposed by white colonizers. An acknowledgment is more than a mere verbal or written statement but must be backed by meaningful action and commitments to change. Creating Indigenous visibility within white universities requires multiple approaches that celebrates Indigenous cultures and excellence; maintains connections and responsibilities to Country; and confronts colonial histories.

Conclusion

Throughout this chapter we have drawn on research led by University of Queensland that provides foundational data on the factors affecting completion and retention rates of Indigenous students in higher education. This seminal research can be used to inform the implementation of evidence-based policies to improve completion rates and educational outcomes for Aboriginal and Torres Strait Islander peoples. The report, alongside initiatives such as the University of Queensland’s *Indigenous Design Framework* (Go-Sam et al., 2021), the University Australia’s *Indigenous Strategy 2022–2025*, and other academic literature, consistently demonstrates that Indigenous staff and students’ inability see oneself in the university environment contributes to their sense of belonging. This can ultimately impact academic outcomes; contribute to decisions to continue or discontinue their studies; or hinder career pathways.

We have highlighted the importance of universities establishing Indigenous-led centers of support, speaking to how they counter negative and often racist interactions experienced on campus by establishing safe, welcoming, and culturally aware spaces. Evidence demonstrates that Indigenous support centers have positive outcomes in improving completion rates among Aboriginal and Torres Strait Islander students. The reason behind this lies in the dedication of its mostly Indigenous staff who work tirelessly to ensure that it is a uniquely “Indigenous” space where students are made to feel that they belong; can raise grievances; and have a support network which has its students back.

The burden of closing the gaps in disparate educational outcomes, however, can only be realized through wider systemic reform that flows throughout the entire university system. Indigenous perspectives, knowledges, histories, cultures, and those who have achieved, and continue to achieve feats of excellence, must be visible. While the recognition of Indigenous sovereignty does not equate to the dismantlement of western frameworks—which continue to govern universities—it does require an open and honest confrontation of colonialism, inclusive of the opportunities it has created for non-Indigenous peoples and the systemic disadvantages for Aboriginal and Torres Strait Islander peoples.

As we discussed in relation to cultural competency training and Acknowledgments of Country, symbolic recognition is not enough. If universities remain committed to improving outcomes for Aboriginal and Torres Strait Islander peoples and the KPIs set by their own Reconciliation Action Plans, there must be compromise and a willingness to listen and avail oneself to the guidance and wisdoms of Indigenous peoples. This needs to be realized in the physical appearance of the university; within curriculum and pedagogies; in the numbers of Indigenous staff employed; and in Indigenous representation at senior positions where policies and strategies are conceived and voted on. Based on the findings of our research the following seven points should be taken on board when considering how universities might better reflect Indigenous standpoints and become welcoming spaces:

1. Universities must recognize that they exist on Indigenous Country and therefore have responsibilities toward its care.
2. Recognizing Indigenous peoples and cultures must translate to meaningful structural reforms that incorporate Indigenous perspectives and personae across all levels, including senior and executive roles.
3. Cultural competency and racial literacy must be improved so that systemic racism is recognized as a tool of colonialism that silences Indigenous outlooks.
4. Indigenous support centers are crucial to helping build a sense of community and belonging among Indigenous staff and students.
5. Strategies to support Indigenous students must be flexible, accommodating to their lived experiences, ongoing, and flow throughout the entire university.

6. Relationality is central to engaging Indigenous communities and building strong relationships. Education must be recognized as a component of community engagement.
7. The physical environment must incorporate Indigenous outlooks and perspectives.

To return to Gordon Bennett’s artwork *Echo and Narcissus* (1988) discussed in the introduction. In this work, Narcissus appears a white naked fetal-like figure, his lack of definition and crawling posture provokes a feeling of anguish and discomfort, as if his presence is alien and unsettled. As Narcissus gazes into the pool, it is Echo, an Aboriginal man, who provides comfort and reassurance through his gentle caress of Narcissus’ face. Until Aboriginal and Torres Strait peoples are recognized as the sovereign custodians of Country and are adequately represented within the social and physical environment, universities will remain places that are in a perpetual state of infancy and discomfort. It is through recognizing and celebrating Indigenous sovereignty, and the fact that Indigenous cultures continue to echo throughout the land upon which Europeans invaded and built their monuments to western civilization that healing, truth-telling, and reconciliation may begin.

Acknowledgments

We acknowledge the generous support of the National Center for Student Equity in Higher Education who have funded this research. We wish to thank the members of the reference group for their advice and guidance through the project. For assistance with organizing interviews, we thank the staff in the Indigenous centers/units at the five university sites. Thanks also to all of the Aboriginal and Torres Strait Islander graduates and university staff who were interviewed as part of the research. The Aboriginal and Torres Strait Islander Studies Unit at The University of Queensland provided a supportive environment to undertake the project and some resourcing to complete the project.

References

- Anderson, L., Riley, L., 2021. Crafting safer spaces for teaching about race and intersectionality in Australian Indigenous Studies. *Aust. J. Indig. Educ.* 50, 229–236. <https://doi.org/10.1017/jie.2020.8>.
- Andersen, C., Bunda, T., Walter, M., 2008. Indigenous higher education: the role of universities in releasing the potential. *Aust. J. Indig. Educ.* 37, 1–8. <https://doi.org/10.1017/S1326011100016033>.
- Bargallie, D., 2020. Unmasking the Racial Contract: Indigenous Voices on Racism in the Australian Public Service. Aboriginal Studies Press.
- Barney, K., 2013. “Taking your mob with you”: giving voice to the experiences of Indigenous Australian postgraduate students. *High Educ. Res. Dev.* 32, 515–528. <https://doi.org/10.1080/07294360.2012.696186>.
- Barney, K., 2018. “We need more mob doing research”: developing university strategies to facilitate successful pathways for Indigenous students into Higher Degrees by Research. *High Educ. Res. Dev.* 37, 908–922. <https://doi.org/10.1080/07294360.2018.1467382>.
- Battiste, M., 2013. Decolonizing Education Nourishing the Learning Spirit. Purich Publishing Ltd., Saskatoon, Canada.
- Best, O., 2018. The cultural safety journey: an aboriginal Australian nursing and midwifery context. In: Yatdjuligin Aboriginal and Torres Strait Islander Nursing and Midwifery Care, pp. 46–64. <https://doi.org/10.1017/9781108123754.005>.
- Behrendt, L., Larkin, S., Griew, R., Kelly, P., 2012. Review of Higher Education Access and Outcomes for Aboriginal and Torres Strait Islander People: Final Report. Department of Industry, Innovation, Science, Research and Tertiary Education, Australian Government.
- Bishop, R., O’Sullivan, D., Berryman, M., 2010. Scaling Up Education Reform Addressing the Politics of Disparity. NZCER Press, Wellington, New Zealand.
- Bond, C., 2014. When the Object Teaches: Indigenous Academics in Australian Universities. Right Now, 14 November, Available. <https://rightnow.org.au/opinion-3/when-the-object-teaches-indigenous-academics-in-australian-universities/>. (Accessed 18 June 2022).
- Brayboy, B.M.J., Fann, A.J., Castagno, A.E., Solyom, J.A., 2012. Postsecondary education for American Indian and Alaskan Natives: higher education for nation-building and self-determination. *ASHE High. Educ. Rep.* 37 (5), 1–154.
- Brown, L., 2010. Nurturing relationships within a space created by “Indigenous Ways of Knowing”: a case study. *Aust. J. Indig. Educ.* 39, 15–22. <https://doi.org/10.1375/S1326011100001095>.
- Brown, C., Shay, M., 2021. From resilience to wellbeing: identity-building as an alternative framework for schools’ role in promoting children’s mental health. *Rev. Educ.* 9, 599–634. <https://doi.org/10.1002/rev3.3264>.
- Brown, L., Kelada, O., Jones, D., 2021. “While I knew I was raced, I didn’t think much of it”: the need for racial literacy in decolonising classrooms. *Postcolonial Stud.* 24, 82–103. <https://doi.org/10.1080/13688790.2020.1759755>.
- Bunda, T., 2014. The Relationship Between Indigenous Peoples and the University: Solid or What! PhD Thesis. University of South Australia.
- Bunda, T., 2017a. The Raced Space of Learning and Teaching in the Relationality of Race and Racism in Educational Research. In: Vass, G., Maxwell, J., Rudolph, S., Gulson, K.N. (Eds.). Routledge, London, pp. 85–96.
- Bunda, T., 2017b. Special issue: indigenous educational research. *Aust. Educ. Res.* 44, 1–4. <https://doi.org/10.1007/s13384-017-0227-x>.
- Bunda, T., Zipin, L., Brennan, M., 2012. Negotiating university “equity” from Indigenous standpoints: a shaky bridge. *Int. J. Incl. Educ.* 16, 941–957. <https://doi.org/10.1080/13603116.2010.523907>.
- Bunda, T., Heckenberg, R., Sneyvangers, K., Phillips, L.G., Lasczik, A., Black, A.L., 2019. Storymaking belonging. *Art Res. Int.* 4, 153–179.
- Butler, K., 2000. Overcoming Terra Nullius: aboriginal perspectives in schools as a site of philosophical struggle. *Educ. Philos. Theor.* 32, 93–101. <https://doi.org/10.1111/j.1469-5812.2000.tb00436.x>.
- Cassells, R., Duncan, A., Dockery, M., Kalsi, J., Loan Vu, L., Mavisakalyan, A., Salazar, S., 2022. Woort Koorliny: Australian Indigenous Employment Index 2022. Minderoo Foundation. <https://bcec.edu.au/publications/woort-koorliny-australian-indigenous-employment-index-2022/>.
- Coates, S.K., Trudgett, M., Page, S., 2022a. Ain’t no mountain high enough: aspirations of Indigenous academics within the academy. *Int. J. Leader. Educ.* 1–15. <https://doi.org/10.1080/13603124.2022.2068186>.
- Coates, S.K., Trudgett, M., Page, S., 2022b. Islands in the stream: Indigenous academic perceptions of Indigenous senior leadership roles. *High Educ. Res. Dev.* 41 (5), 1451–1467. <https://doi.org/10.1080/07294360.2021.1920894>.
- Cote-Meek, S., 2014. Colonized Classrooms: Racism, Trauma and Resistance in Post-secondary Education. Fernwood Publishing, Winnepeg, Manitoba.
- Daley, P., 2018. Heroes, monuments and history. *Meanjin* 77, 86–97. <https://meanjin.com.au/essays/heroes-monuments-and-history/>.
- Delgado, R., Stefancic, J., 2017. Critical Race Theory: An Introduction. NYU press. <https://doi.org/10.2307/j.ctt1ggjin3>.

- DiAngelo, R., 2022. *White Fragility Why Understanding Can Be So Hard For White People Adapted for Young Adults*. Beacon Press, New York.
- Foley, F., Martin-Chew, L., Nicoll, F., 2015. *Courting Blakness: Recalibrating Knowledge in the Sandstone University*. University of Queensland Press.
- Foley, F., Nicoll, F., Volcic, Z., O'donnell, D., 2019. Creative frictions in the neoliberal university: courting blakness at the university of Queensland. In: Bottrell, D., Manathunga, C. (Eds.), *Resisting Neoliberalism in Higher Education*, vol. 1. Springer. https://doi.org/10.1007/978-3-319-95942-9_9.
- Fredericks, B., 2009. The epistemology that maintains white race privilege, power and control of Indigenous studies and Indigenous peoples' participation in universities. *Crit. Race Whiteness Stud.* 5, 1–12.
- Fredericks, B., 2020. Collaborative creative processes that challenge us as “anomaly”, and affirm our indigeneity and enact our sovereignty. *M/C J.* 23 (5). <https://doi.org/10.5204/mcj.1674>.
- Fredericks, B., Bargallie, D., 2016. “Which way? Talking culture, talking race”: unpacking an Indigenous cultural competency course. *Int. J. Crit. Indig. Stud.* 9, 3–16. <https://doi.org/10.5204/ijcis.v9i1.141>.
- Fredericks, B., Bargallie, D., 2020a. An Indigenous Australian cultural competency course: talking culture, care and power. In: Frawley, J., Russell, G., Sherwood, J. (Eds.), *Cultural Competence and the Higher Education Sector: Perspectives, Policies and Practice*. Springer Publications, pp. 295–308. https://doi.org/10.1007/978-981-15-5362-2_16.
- Fredericks, B., Bargallie, D., 2020b. Situating race in cultural competency training: a site of self-revelation. *M/C J.* 23 (4). <https://doi.org/10.5204/mcj.1660>.
- Fredericks, B., Bradfield, A., 2021a. “I’m not afraid of the dark”: white colonial fears, anxieties, and racism in Australia and beyond. *M/C J.* 24 (2). <https://doi.org/10.5204/mcj.2761>.
- Fredericks, B., Bradfield, A., 2021b. Indigenous Australians and COVID-19: highlighting ongoing food security issues. *Int. J. Home Econ.* 14, 53–65.
- Fredericks, B., Kinnear, S., Daniels, C., Croftwarcon, P., Mann, J., 2015. *Path+ Ways. Towards Best Practice Bridging and Indigenous Participation Through Regional Dual-Sector Universities*. Report Submitted to the National Centre for Student Equity in Higher Education (NCSEHE). Curtin University, Perth. <https://www.ncsehe.edu.au/publications/pathways-towards-best-practice-bridging-indigenous-participation-regional-dual-sector-universities/>.
- Fredericks, B., White, N., Phillips, S., Bunda, T., Longbottom, M., Bargallie, D., 2019a. Being ourselves, naming ourselves, writing ourselves: Indigenous Australian women disrupting what it is to be academic within the academy. In: Thomas, M., Reinertsen, A. (Eds.), *Academic Writing and Identity Constructions*. Springer.
- Fredericks, B., White, N., Phillips, S., Bunda, T., Longbottom, M., Bargallie, D., 2019b. Being ourselves, naming ourselves, writing ourselves: Indigenous Australian women disrupting what it is to be academic within the academy. In: Thomas, L.M., Reinertsen, A.B. (Eds.), *Academic Writing and Identity Constructions: Performativity, Space and Territory in Academic Workplaces*. Palgrave Macmillan, Cham, Switzerland, pp. 75–96. https://doi.org/10.1007/978-3-030-01674-6_5.
- Fredericks, B., Bunda, T., Gilbert, S., 2021. Teaching and researching before and during# COVID19: the indigenous engagement division mob. *J. Glob. Indig. S.* 5, 1–13. <https://www.journalofglobalindigeneity.com/article/19428-teaching-and-researching-before-and-during-covid19-the-indigenous-engagement-division-mob>.
- Fredericks, B., Barney, K., Bunda, T., Hausia, K., Martin, A., Elston, J., Bernardino, B., Griffiths, D., 2022. Building the Evidence to Improve Completion Rates for Indigenous Students. NCSEHE Project Final Report. National Centre for Student Equity in Higher Education. <https://www.ncsehe.edu.au/publications/improve-completion-rates-indigenous-students/>.
- Gilbert, K., Bunda, T., 2017. The Pleasure and Pain of Aboriginal Being in the University in Producing Pleasure in Academia. In: Denaher, P., Harmes, M., Riddle, S. (Eds.), *Sense Publishers, Rotterdam. The Netherlands*, pp. 185–199.
- Gilbert, S., Tillman, G., 2017. Teaching practise utilising embedded Indigenous cultural standards. *Aust. J. Indig. Educ.* 46, 173–181. <https://doi.org/10.1017/jie.2017.4>.
- Gillborn, D., 2015. Intersectionality, critical race theory, and the primacy of racism: race, class, gender, and disability in education. *Qual. Inq.* 21, 277–287. <https://doi.org/10.1177/1077800414557827>.
- Go-Sam, C., Greenop, K., Marnane, K., Bower, T., 2021. *Campuses on Countries: Aboriginal and Torres Strait Islander Design Framework Engagement Report at The University of Queensland*. The University of Queensland, Brisbane. <https://doi.org/10.14264/955791e>.
- Gregory, J., 2021. Statue wars: collective memory reshaping the past. *Hist. Aust.* 18, 564–587. <https://doi.org/10.1080/14490854.2021.1956333>.
- Grieve-Williams, V., 2008. The “battlefields”: identity, authenticity and aboriginal knowledges in Australia. In: Minde, H. (Ed.), *Indigenous Peoples: Self-Determination, Knowledge and Indigeneity*. Eburon, Delft, pp. 287–311.
- Gunstone, A., 2009. Whiteness, Indigenous peoples and Australian universities. *ACRAWSA e J.* 5 (1), 1–8.
- Hinchliffe, J., 2020. “Enabler” of Massacres: The Push to Reexamine the Legacy of Founding Father Samuel Griffith. *The Guardian*, 5 June, Available. <https://www.theguardian.com/australia-news/2022/jun/05/enabler-of-massacres-the-push-to-reexamine-the-legacy-of-founding-father-samuel-griffith>. (Accessed 5 June 2020).
- Hokowhitu, B., Moreton-Robinson, A., Tuhiwai-Smith, L., Andersen, C., Larkin, S., 2020. *Routledge Handbook of Critical Indigenous Studies*. Routledge.
- Huaman, E.S., Brayboy, B.M.J. (Eds.), 2017. *Indigenous Innovations in Higher Education Local Knowledge and Critical Research*. Sense Publishers, Rotterdam, The Netherlands.
- Kinnear, S., Wilks, J., Wilson, K., Hughes, T., Thomas, S., 2014. “Can’t Be What You Cant See”: The Transition of Aboriginal and Torres Strait Islander Students Into Higher Education. The University of Notre Dame. The Australian Government Office of Learning & Teaching, Western Australia.
- Kowal, E., 2010. Welcome to country? *Meanjin* 69, 15–173.
- Leroy-Dyer, S., Heckenberg, S., 2022. The Gap Will Never Close if Aboriginal and Torres Strait Islander Students Don't Feel Safe on University Campuses. *The Conversation*, 31 March, Available. <https://theconversation.com/the-gap-will-never-close-if-aboriginal-and-torres-strait-islander-students-dont-feel-safe-on-university-campuses-180234>. (Accessed 17 April 2022).
- Lowitja Institute, 2021. *Close the Gap Report—2021—Leadership and Legacy Through Crises: Keeping Our Mob Safe*. <https://www.lowitja.org.au/page/services/resources/Cultural-and-social-determinants/culture-for-health-and-wellbeing/close-the-gap-report-2021>.
- Mckenna, M., 2014. Tokenism or belated recognition? Welcome to Country and the emergence of Indigenous protocol in Australia, 1991–2014. *J. Aust. Stud.* 38, 476–489. <https://doi.org/10.1080/14443058.2014.952765>.
- Mcknight, A., 2016. Meeting country and self to initiate an embodiment of knowledge: embedding a process for aboriginal perspectives. *Aust. J. Indig. Educ.* 45, 11–22. <https://doi.org/10.1017/jie.2016.10>.
- McLaughlin, J., Whatman, S., 2011. The potential of critical race theory in decolonizing university curricula. *Asia Pac. J. Educ.* 31, 365–377. <https://doi.org/10.1080/02188791.2011.621243>.
- McLean, I., Bennett, G., 1996. *The Art of Gordon Bennett*. Craftsman House, Roseville East.
- Milne, T., Creedy, D.K., West, R., 2016. Integrated systematic review on educational strategies that promote academic success and resilience in undergraduate indigenous students. *Nurse Educ. Today* 36, 387–394. <https://doi.org/10.1016/j.nedt.2015.10.008>.
- Moreton-Robinson, A.M., 1998. White race privilege: nullifying native title. In: *Bringing Australia Together: The Structure and Experience of Racism in Australia*, pp. 30–44.
- Moreton-Robinson, A., 2007. *Sovereign Subjects: Indigenous Sovereignty Matters*. Allen & Unwin, Crows Nest, N.S.W.
- Moreton-Robinson, A., 2015. *White Possessive: Property, Power, and Indigenous Sovereignty*. University of Minnesota Press, Minnesota. <https://doi.org/10.5749/minnesota/9780816692149.001.0001>.
- Moreton-Robinson, A., 2020. *Talkin' Up to the White Woman: Indigenous Women and Feminism*. University of Queensland Press, St. Lucia, QLD.
- Moreton-Robinson, A., 2021. *Place Names, Monuments and Cannibalism: James Cook and the Possessive Logics of Patriarchal White Sovereignty*. AILS Sinker Series, Cornell University. Youtube. https://www.youtube.com/watch?v=5L_QMLWcE5k.
- Moreton-Robinson, A., 2022. Monuments, place names and black lives matter: memorialising captain James cook. *Legalities* 2, 67–81. <https://doi.org/10.3366/legal.2022.0029>.
- Morrison, A., Rigney, L.-I., Hattam, R., Diplock, A., 2019. *Toward an Australian Culturally Responsive Pedagogy: A Narrative Review of the Literature*. University of South Australia.
- Nakata, M., 2007. The cultural interface. *Aust. J. Indig. Educ.* 36, 7–14. <https://doi.org/10.1017/s1326011100004646>.
- Nakata, M., 2011. Pathways for Indigenous education in the Australian curriculum framework. *Aust. J. Indig. Educ.* 40, 1–8. <https://doi.org/10.1375/ajie.40.1>.

- Nakata, M., 2012. A Torres Strait Islander's story of the struggle for a better education. In: Price, K. (Ed.), *Aboriginal and Torres Strait Islander Education: An Introduction for the Teaching Profession*. Cambridge University Press, Victoria.
- Nakata, M., 2013. The rights and blights of the politics in indigenous higher education. *Anthropol. Forum* 23, 289–303. <https://doi.org/10.1080/00664677.2013.803457>.
- Nakata, M., Nakata, V., Day, A., Peachey, M., 2019. Closing gaps in Indigenous undergraduate higher education outcomes: repositioning the role of student support services to improve retention and completion rates. *Aust. J. Indig. Educ.* 48, 1–11. <https://doi.org/10.1017/jie.2017.36>.
- Newton, B.J., 2021. Creating cultural safety as an aboriginal teacher in a class of non-aboriginal university students. *Aust. Soc. Work* 74, 4–12. <https://doi.org/10.1080/0312407X.2020.1799422>.
- Oliver, R., Rochecouste, J., Bennell, D., Anderson, R., Cooper, I., Forrest, S., Exell, M., 2013. Understanding Australian aboriginal tertiary student needs. *Int. J. High. Educ.* 2, 52–64. <https://doi.org/10.5430/ijhe.v2n4p52>.
- Pagnini, D., Stylianou, M., Kok, B., Johnson, D., 2014. Towards a Performance Measurement Framework for Equity in Higher Education. Cat. No. IHW 129. Australian Institute of Health and Welfare. <https://www.aihw.gov.au/reports/indigenous-australians/equity-in-higher-education/summary>.
- Pechenkina, E., Kowal, E., Paradies, Y., 2011. Indigenous Australian students' participation rates in higher education: exploring the role of universities. *Aust. J. Indig. Educ.* 40, 59–68. <https://doi.org/10.1375/ajie.40.59>.
- Pelizzon, A., Kennedy, J., 2019. “Welcome to country” and “acknowledgment of country”: (re) conciliatory protest. *Contention* 7, 13–28. <https://doi.org/10.3167/cont.2019.070103>.
- Plater, S., Mooney-Somers, J., Barclay, L., Boulton, J., 2020. Hitting the white ceiling: structural racism and aboriginal and Torres Strait Islander university graduates. *J. Sociol.* 56, 487–504. <https://doi.org/10.1177/1440783319859656>.
- Rigney, L.-I., 1999. Internationalization of an indigenous anticolonial cultural critique of research methodologies: a guide to indigenist research methodology and its principles. *Wicazo Sa Rev.* 14, 109–121. <https://doi.org/10.2307/1409555>.
- Rigney, L.-I., 2006. Indigenist research and aboriginal Australia. In: Kunnie, J.E., Godukka, N.I. (Eds.), *Indigenous Peoples' Wisdom and Power*. Routledge, London. <https://doi.org/10.4324/9781315252414-3>.
- Rochecouste, J., Oliver, R., Bennell, D., Anderson, R., Cooper, I., Forrest, S., 2017. Teaching Australian Aboriginal higher education students: what should universities do? *Stud. High. Educ.* 42, 2080–2098. <https://doi.org/10.1080/03075079.2015.1134474>.
- Sarra, C., 2014. *Strong and Smart-Towards a Pedagogy for Emancipation: Education for First Peoples*. Routledge.
- Shalley, F., Smith, J., Wood, D., Fredericks, B., Robertson, K., Larkin, S., 2019. Understanding Completion Rates of Indigenous Higher Education Students From Two Regional Universities, A Cohort Analysis. Final Report for Student Equity in Higher Education Research Grants Program, 2017. National Centre for Student Equity and Higher Education (NCSEHE), Perth, Western Australia. <https://www.ncsehe.edu.au/publications/completion-indigenous-higher-education-regional-universities/>.
- Shay, M., Oliver, R. (Eds.), 2021. *Indigenous Education in Australia: Learning and Teaching for Deadly Futures*. Routledge, Abingdon, Oxon United Kingdom.
- Stewart, G.T., Macdonald, L., Matapo, J., Fa'avae, D.T.M., Watson, B.K.I., Akiu, R.K., Martin, B., Mika, C., Sturm, S., 2021. Surviving academic whiteness: perspectives from the pacific. *Educ. Philos. Theor.* 1–12. <https://doi.org/10.1080/00131857.2021.2010542>.
- Synot, E., Graham, M., Graham, J., Valencia-Forrester, F., Longworth, C., Backhaus, B., 2021. Weaving First Peoples' knowledge into a university course. *Aust. J. Indig. Educ.* 50, 222–228. <https://doi.org/10.1017/jie.2019.29>.
- Tamer, R., 2022. “Hugely Disappointing”: Indigenous Employees “Almost Entirely Absent” From Senior Leadership, Report Finds. SBS News, 24 May, Available. <https://www.sbs.com.au/news/article/hugely-disappointing-indigenous-employees-almost-entirely-absent-from-senior-leadership-report-finds/extraft6rew>. (Accessed 23 June 2022).
- Tanaka, M.T.D., 2016. *Learning & Teaching Together Weaving Indigenous Ways of Knowing Into Education*. UBC Press, Vancouver, Canada.
- Taylor, E.V., Lalovic, A., Thompson, S.C., 2019. Beyond enrolments: a systematic review exploring the factors affecting the retention of Aboriginal and Torres Strait Islander health students in the tertiary education system. *Int. J. Equity Health* 18, 1–19. <https://doi.org/10.1186/s12939-019-1038-7>.
- The University of Queensland, 2021. The University of Queensland Aboriginal and Torres Strait Islander Research and Innovation Strategy 2021–2025. <https://www.uq.edu.au/research/files/39599/UQ%20Aboriginal%20and%20Torres%20Strait%20Islander%20Research%20and%20Innovation%20Strategy%202021-2025.pdf>.
- Trudgett, M., Page, S., Sullivan, C., 2017. Past, present and future: acknowledging Indigenous achievement and aspiration in higher education. *HERDSA Rev. High. Educ.* 4, 29–51. <http://www.herdsa.org.au/herdsa-review-higher-education-vol-4/29-51>.
- Uink, B., Dodd, J., Bennett, J., Bonson, D., Eades, A-M., Hill, B., 2022. Confidence, practices and training needs of people working with Aboriginal and Torres Strait Islander LGBTIQ+ clients. *Int. J. Res. Interv. Care*. <https://doi.org/10.1080/13691058.2022.2031298>.
- Universities Australia, 2017. *Indigenous Strategy 2017–20*. <https://www.universitiesaustralia.edu.au/wp-content/uploads/2019/06/Indigenous-Strategy-v16-1.pdf>.
- Universities Australia, 2020. *Indigenous Strategy Annual Report*. <https://apo.org.au/node/277226>.
- Universities Australia, 2022. *Indigenous Strategy 2022–25*. <https://www.universitiesaustralia.edu.au/policy-submissions/diversity-equity/universities-australias-indigenous-strategy-2022-2025/>.
- Waterman, S.J., Lowe, S.C., Shotton, H.J., 2018. *Beyond Access Indigenizing Program for Native American Students Success*. Stylus Publishing, LLC, Virginia.
- Waters, M.W., 2017. *Indigenous Knowledge Production Navigating Humanity Within the Western World*. Routledge, New York.
- Watson, I., 2014. Re-centring First Nations knowledge and places in a terra nullius space. *AlterNative* 10, 508–520. <https://doi.org/10.1177/117718011401000506>.
- Williamson, J., Dalal, P., 2007. Indigenising the curriculum or negotiating the tensions at the cultural interface? Embedding indigenous perspectives and pedagogies in a university curriculum. *Aust. J. Indig. Educ.* 36, 51–58. <https://doi.org/10.1017/s1326011100004701>.

Capturing storms: Pacific women's approaches to change in New Zealand universities

Sereana Naepi, Waipapa Taumata Rau, The University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	119
Mapping our journeys	121
Masi methodology and talanoa	121
Slow change	122
Utilizing the university to make change	123
Working with allies to make change	123
Collective action for change	124
Choosing change	125
Mentoring for change	125
Conclusion	126
Post-script	127
Acknowledgment	127
References	127

Introduction

Some of that change is generational. We're looking at generational change, and bringing our students through, and building our research literature, that takes decades.- Community talanoa one

Change is urgent in New Zealand universities. We have collected evidence for this for a significant amount of time, yet Pacific people continue to be under-served and excluded by our universities. Pacific academics make up less than 1% of Aotearoa–New Zealand professors and 1.7% of academics (Naepi, 2019a) and recently it was recognized that in order to achieve equity we needed 450 more Pacific academics and that at our current rate of change this would take 130 years (Tuhiwai Smith et al., 2020, p. 33). Pacific students have lower completion and retention rates (Naepi et al., 2021) and while over the last decade the overall all Pacific completion rate is increasing, closer inspection such as looking at the fields of society and culture, and the natural and physical sciences we can see that this increase is mirrored by the entire student body so the achievement gap persists (Naepi et al., 2021). UniversitiesNZ has previously identified that if parity of participation for Pacific students was achieved, an extra 1100 Pacific students would be completing degrees each year (UniversitiesNZ, 2018, p. 2). Significantly, solutions for this problem have traditionally focused on pedagogical approaches (Leenen-Young, 2020; Leenen-Young et al., 2021; Naepi et al., 2017) and knowledge inclusion (Nabobo-Baba, 2008; Naepi, 2021a; Suaalii-Sauni, 2008). Pacific peoples are a target group in higher education New Zealand for historical and contemporary reasons. Pacific peoples have had multiple migrations to Aotearoa New Zealand, the first wave of migration was that of settlement, when Eastern Pacific people explored and settled in Aotearoa and became tangata whenua. This first wave is important to the relationship between Māori (Tangata Whenua/Indigenous peoples of Aotearoa) and Pacific peoples as it created bonds and relationships through culture and genealogy in Te Moana Nui Kiwa (greater Oceania kinship connections) (Health Research Council, 2014). Importantly this relationship further cements Pacific peoples recognize the deep commitment needed to Article 2 of Te Tiriti o Waitangi – Tino rangatiratanga (self-determination) and the importance of tangata whenua authority and autonomy in Aotearoa. One hundred and fifty years ago, Pacific peoples arrived in Aotearoa as trainee teachers, missionaries, sailors, and whalers in a second wave of migration.¹ The third wave followed 70 years later when Pacific peoples who had served the colonial government as civil servants within the Pacific “territories” or in the colonial armed forces were able to move to Aotearoa (Macpherson et al., 2001). The fourth migration, which occurred 50 years ago, was perhaps the most significant and is the migration story with which most people are familiar today. For the fourth migration, Pacific people migrated for economic reasons and found work in the manufacturing and service sectors in post-war Aotearoa (Fleras and Spoonley, 1999; Macpherson et al., 2001; Macpherson, 2004; Te Punga Sommerville, 2012). This history has meant that Pacific peoples are more than a “diversity group” within New Zealand education systems.

¹New Zealand officially colonized Samoa from 1920 to 62, the Cook Islands from 1901 to 65, Niue from 1901 to 74, and Tokelau from 1926 to the present, with Niue and the Cook Islands remaining protectorates of New Zealand (Te Punga Sommerville, 2012).

For us, if you love your students, then if you see something's not working, then you change it! Talanoa Nine

The shift to parity/population matching for academics and students provides another avenue for change. However, as I have previously argued this approach to change should be considered through the lenses of racial capital (Leong, 2012) and specimen collecting (Musser, 2015) in order to ensure that Pacific people within institutions do not just become another desirable diversity body (Ahenakew and Naepi, 2015) that serves the purpose of the university but not of our own communities (Naepi, 2021a).

We can't wait for policies to change. We cannot wait for the big money to come through. So the service is integral to growth of community, otherwise we're going to go backwards. And in some regards, somewhere we have in terms of education. Government throws out these glamorous statistics, but they're all not as rosy as we know. So that is urgent. That's our requirement as part of the academy. And often we do it for no money. It's never about money. It's never about our position. It's just about support and encouragement of our peoples. - Talanoa Seven

The call for change has not gone unheard. In the 2013–17 Pacific Education Plan aimed for at least parity in all aspects of education for Pacific peoples (New Zealand Ministry of Education, 2013). The most recent Pacific Education Action Plan calls for addressing structural bias that prevents Pacific success (Ministry of Education, 2021). Importantly, the report that informs and provides explanation for the plan outlines that “in spite of universities’ public commitment to supporting Pacific peoples in tertiary study, a substantial amount of work is needed to address the ‘white masculine imprint’ on university structures” (Ministry of Education, 2021, p. 14). We have also seen a shift in how our universities base research grant funding is delivered; the Performance Based Research fund (PBRF) now weights research Evidence Portfolios from Pacific researchers at 2 and research Evidence Portfolios submitted to the Pacific research panel at 2.5 (Tertiary Education Commission, 2021a). This fund is worth \$321 million a year (Te Tai Ōhanga The Treasury, 2020) and unsurprisingly has resulted in a significant shift in university recruitment and research priorities as some universities faced a significant loss in revenue (Naepi, 2021a). For Pacific students the Tertiary Education Commission provides additional funding per Pacific EFT ranging from \$325 to \$450 to support higher education institutions to provide the support needed for learner success (Tertiary Education Commission, 2021b). Again, this is not a small amount of money as in “2020, the University of Auckland – which reported 2843 Pacific EFTs, of which 86.3% were below level 7 qualifications and 4% level 7 or above – would have received in the realm of \$747,000 in additional Tertiary Education Equity Funding to support Pacific student success” (Naepi, 2021a, p. 24). The government expectations of equity have resulted in universities across Aotearoa New Zealand making commitments to Pacific communities that are primarily focused on the politics of inclusion.

The plans are only as good as the implementation, and the energy of the people behind it. - Talanoa Twelve

At some point we must ask how much qualitative and quantitative (Ahenakew and Naepi, 2015; Airini et al., 2011; Barber and Naepi, 2020; Leenen-Young, 2020; Leenen-Young et al., 2021; McAllister et al., 2020a,b; Naepi, 2019a, 2020, 2021b; Naepi et al., 2020, 2021; Suaalii-Sauni, 2008; Theodore et al., 2018a,b; Thomsen et al., 2021) evidence is needed for change. Recording the different ways that we are excluded or underserved can support and drive structural change like the shift in the PBRF, however, that change can be slow and often the generation of students or staff in front of us need change today. This chapter attempts to complement the work that is being done to record the different ways that Pacific peoples are excluded and under-served by offering some insight into how Pacific women go about the everyday work of changing the university. Ultimately, what drives this chapter is the belief that the academy can be redeemed, and it is with that in mind that exploring how Pacific peoples engage in change becomes urgent; as we simply do not have a 130 years to wait for equity (Tuhiwai Smith et al., 2020, p. 134). Instead, we must engage with how we can accelerate change because business as usual or the polite politics of inclusion means that we will not experience the change necessary for another four generations.

Ideally change is supposed to be revolutionary - Community Talanoa Four

Utilizing Masi methodology to frame twenty-seven talanoa with Pacific women who worked in Aotearoa New Zealand's universities this chapter explores how Pacific women engage in change. In writing this Chapter I have fought the tension to make it completely comprehensible and to also reflect the voices of the women that I worked with. So, instead this chapter is something of a middle ground; a clear narrative with moments of interruption to enable what Ahenakew referred to as a shift from sense-making to sense-sensing (2016, p. 333). These interruptions provide an opportunity to see that while I write about the university, Pacific women experience the university; all these interruptions come from the talanoa.

So maybe our secrets will be shared, but we wouldn't have partaken in your research if we had fear of that being used against us at a later stage. I think that our voices need to be heard more broadly and more widely, and regionally, but internationally as well. - Talanoa Seven

What this chapter offers is an exploration of how Pacific women engage in change making in our universities with a slight bracketing that not all of our secrets to change will be recorded because some of them we need to continue sharing in our whisper networks so that institutional habit (Ahmed, 2012) doesn't catch us unawares and find ways to block the change we desire.

Mapping our journeys

It is common in Pacific education circles to use wayfaring as an analogy for our journeys through the education system. Just as our ancestors were able to articulate and travel to what was beyond the horizon, we can imagine a better educational system before we arrive there. Pacific peoples' have been scientists since time immemorial; our ancestors navigated the world's largest oceans repeatedly and intentionally by reading the stars, watching the birds and listening to the ocean as we navigated from one island to the next then beyond to the horizon. Pacific peoples in the academy continue this tradition by intentionally engaging in career mapping; by understanding what we need to travel to the next island and being aware of what slows us. Fig. 1 provides an image of the necessary navigation for Pacific people through the academy. In Fig. 1 the storm that slows us are institutional habits (Ahmed, 2012) that are foundational to our institution's existence. The numbers above don't just happen, instead the foundations of colonialism, neo-colonialism, neoliberalism, sexism and racism in the academy (Naepi, 2021a,b, 2022; Kidman, 2020; Kidman and Chu, 2017, 2019) combine to create a storm that slows our progress. Importantly, Pacific women recognized that these storms existed and that they did slows us (Naepi, 2021b). This is where the analogy of storms as opposed to walls is useful. Pacific navigators understood that storms could be weathered, avoided or harnessed. But before we could do any of that we had to capture them, to know they existed and so this chapter captures storms by recording how we respond to them; by enabling us to see how Pacific women have practiced change in our universities.

Masi methodology and talanoa

Masi methodology is an Indigenous Pacific methodology that centers Pacific women in the research process (Naepi, 2019b, p. 237). Masi methodology ensures that throughout the research process Pacific women's voices are included by "recognizing that the work and knowledge that Pacific women do and hold is valuable throughout our lives" (Naepi, 2019b, p. 238) and that the research is "intentional in viewing Pacific women as knowledge holders who are experts in their field. In practice, this means centering Pacific women's voices as authoritative and having something to offer beyond the limits of the academy's understanding." (Naepi, 2019b, p. 238). Masi methodology was the methodological framework that resulted in the decision to use talanoa as the method. Talanoa is a Pacific narrative inquiry method grounded in the understanding that knowledge is created in relationality (Naepi, 2019c, p. 11). Talanoa is an established Indigenous Pacific research method (Fa'avae et al., 2016; Farrelly and Nabobo-Baba, 2012; Otunuku, 2011; Prescott, 2008; Reynolds, 2017; Suaalii-Sauni and Fulu-Aiolupoeta, 2014; Vaiolotti, 2006) that is based on the Pacific cultural practice of talanoa (Manueli, 2012). Importantly, Prescott's (2008) in depth exploration of talanoa in comparison to the modern interview technique noted that an interview is primarily focused on gathering knowledge, whereas talanoa is about creating or strengthening relationships and coming to reach a state of understanding between those involved. In 2017, in a study on sport management and talanoa, Stewart-Withers, Sewabu, and Richardson summarized talanoa as:

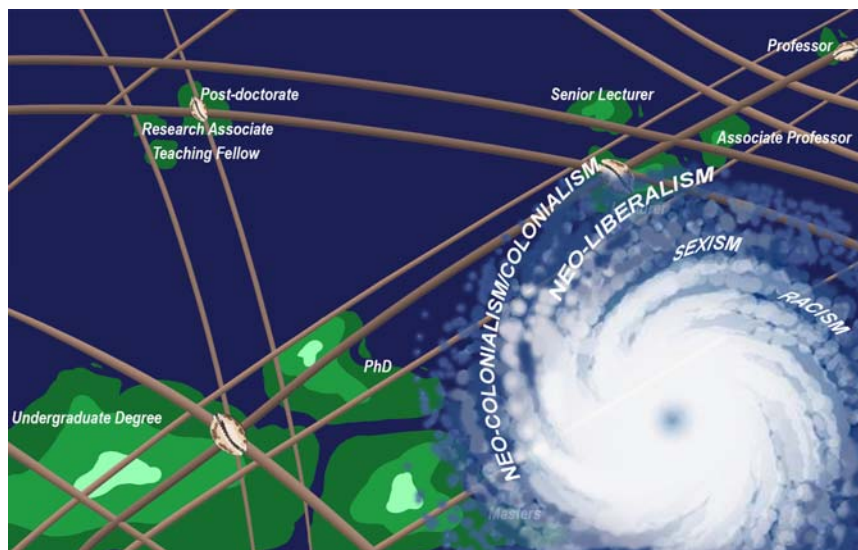


Fig. 1 Pacific navigation of academic pathways (Naepi et al., 2020, p. 153).

a process which affords value to, for example, kinship, land, tradition/custom, relationships, ancestors, ceremony, cosmology, space, language, ethics, the chiefly system, systems of faith, and protocols. Talanoa encompasses behaviors/manners such as openness, patience, tolerance, flexibility, silence, humility, generosity, gifting, reciprocity, humour, empowerment, listening, sharing, forgiveness, and subjectivity. (p. 59)

There are two types of talanoa represented below. The first are the initial one on one talanoa and the second are community talanoa. In alignment with masi methodology this research recognized that Pacific women should have the right to confirm the research reliability and validity as authoritative voices of their own stories. As such the community talanoa were in relation to the one-on-one findings where Pacific women were asked to confirm and expand on what had been shared so far. So, it is through a strengthening and building of relationships that these stories of change for Pacific women have been gathered. These talanoa represent 216 years of collective knowledge as Pacific women attempting to change our universities.

I was surprised how very angry I was, and I thought about that for a long time afterwards, the necessity to have platforms to talk about stuff, it keeps coming back to me a lot, between that one and this one. Having had a chance to process some of that stuff and going, okay, I can put it aside and continue on, but now at least I know that that is there and was there, and was a part of my experience, and it's better to know that something is living with you, then to have it suddenly flare up. Something that you weren't aware of, explode in the committee meeting somewhere, so thank you for allowing that part of my academic experience to breathe, to see the light of day, and I think move to a healthier place. - Community Talanoa One

Slow change

Before beginning a discussion on Pacific women's approaches to change in our universities we need to recognize that change is slow and how Pacific women have articulated this slowness helps us to further understand the different approaches that they take to change.

Change in universities is slow because so much of what universities do are steeped in tradition or as Ahmed would say in institutional habit (Ahmed, 2012):

a slow, it's ever slow, slow-moving, because it's so steeped in its tradition – Talanoa Nine

Change as something to be chipped away at, and Pacific women recognize that the revolutionary change we desire may not be possible:

It's a bit more of an indirect way, but challenging [the discipline] as whole, it's never really been a very revolutionary way. Well, at least I don't think so. I think it's just clawing or trying to bash your way or itch your way into where it's just a little bit at a time.....That's why you just have to chip away and put a little bits here and there - Talanoa Four

Part of the slowness of change is that some people simply don't want change and that the bureaucracy of change is a source of frustration:

It's there, it's just that some people don't want the office space to change. People in the office don't want to change, and then, say, if we wanted to bust down a wall that's like property services going nuts at us because we just take down one wall to make a student space. It's there. There's just too much red tape to get through for different people, and so that's really annoying - Talanoa One

However, change is a driver to keep Pacific women in our roles but also placed a time limit on how long we were willing to wait for change:

The quest to set things in change is what keeps you there. The fact that maybe you begin to see some change happening is very uplifting. But I think in order for me to stay in this position, something will have to change. I'm assessing the position, I'm in it for two years, so at the end of those two years I will certainly have a think about, am I able to manage this responsibly? - Talanoa Four

Pacific women recognize that change takes times, particularly in universities and that this can be a source of frustration. However, we also recognize how seeing change can be a motivating force for staying in a role. Importantly, for us while we are aware that change takes time we are also willing to put a limit on how long we are willing to wait for change.

Utilizing the university to make change

Pacific women know the importance of understanding the university and utilizing policy for change. The university is a tool to be wielded and the university can be used to support wider plans for change:

I think about being us. I think I recognize that the university does its thing. I see the university as being able to be co-opted into some other bigger plan, and that what I need from a university is the space to be me. - Talanoa Fourteen

We often understand the importance of understanding where power lies and who wields it:

It's understanding how policies get made, and who makes it. It's understanding how power is ... Where power resides, and who wields it. It's understanding how the system works, then that's kind of really as bland as that. When you come into the space you got to be a very good observer of who has the power, who has the power over what, and then who makes the decision over what. If there is something that I need to change, I always make sure that I know who is the last person down at the end of the chain that needs to make that decision, and then basically go straight to them. - Talanoa Twelve

Importantly, we also understand that in order to drive change sometimes it is Pacific women who need to need to wield that power:

"I've had someone, our Associate Dean Pacific saying things like, 'We need you guys to be sitting at those tables where those conversations happen', and for me, it was like that's never going to happen, and then you think, but then you have to, you have to, because it's not just you." - Community Talanoa One

Sitting at the table often also means understanding the people at the table and being aware of how they think:

I appreciate the, don't change who you are, your values is what got you here, and I absolutely believe that, but I'm learning that to be able to get to that table and sit around with those key people, you've got to be able to think like them too. - Community talanoa one

Driving change can also mean sharing with others how the system works and how to navigate around and through it:

also one way I do change is I tell the students how to get around the system... ... if you've been here long enough, you kind of know how the system works. but it's just knowing these little things that can make a difference, and that's how I see change - Community Talanoa Four

It is in the day to day details and being aware of finer details within the system that we can disrupt business as usual:

pragmatic things that we have to deal with every day, and we must read the details. We have to do the paperwork. We must attend to those finer details in the system, otherwise, that's where the system engulfs a more white privileged, white mainstream approach to understanding students. And those are major impacts on students. - Talanoa seven

Pacific women see the university as a tool to be utilized in change but also recognize that in order to be able to do this we must be aware of where power lies and have a seat at the table. The seat at the table means ensuring that we understand how others around the table think and what motivates them while simultaneously not changing who we are and our own values system. Often this can also mean reading the details and utilizing these details for change. Pacific women who do all these things also believe that for change to occur we need to share this knowledge so that others can navigate around and through the system.

Working with allies to make change

Pacific women see a role for allies in change making in our universities, particularly when it came to pedagogical practice. Sometimes these allies had established relationships:

I'm going to be working with a palangi friend to help us. She's trying to bring about introducing Pasifika methodologies and practices and theories into the way people teach Pasifika students in science. And you know, she's going to be asking them how are you going to transform your teaching practices to be inclusive of Pasifika world views. - Community Talanoa Four

It is helpful when allies are keen and want to engage as then they can be pushed to think about wider changes:

So what's also good about that, is that he is so open and, genuinely keen. And so he said "Oh this has been a great start." And I actually emailed and said, "Where do we go from here so why don't we have a meeting and start to think about how we could audit, not audit, but how we could provide some kind of mechanism to look at what's being taught and where there needs to be some changes in what's being taught. Looking at where our Pacific learners are struggling and what could be modified. - Talanoa Thirteen

Positive reinforcement when working with allies was also identified as a key change mechanism, with Pacific women reflecting that sometimes too much critique led to defensive responses:

And then we never get change. You know? So again that's very much from what my parents taught me, it's about working with people to effect change rather than continually criticizing because then they won't, they'll just get defensive. Talanoa Thirteen

One particularly insightful moment was when a Pacific women shared that she recognizes that authority that having an ally can bring to a program and that she chooses to protect her community in the short-term to enable long-term gains:

So, I share one of my main leadership roles, but sometimes people have difficulty seeing me as a leader. So, I choose to share this with a man- older man. It just provides a level of safety for my team. If it was just me, I would be okay to take that risk and just see if I can hold the land. But there's 10 other people whose jobs rely on us to be trusted, so I have my colleague there for those people who can't see me. They see him and they're reassured that there's someone in charge who looks like someone that they want to be in charge. - Talanoa Ten

Pacific women understand that allies are needed for change. Often this is occurring in relation to pedagogy, where Pacific women can support professional development for allies in how and what they teach. Importantly Pacific women recognized that allies need to be encouraged and that too much critique can lead to disengagement.

Collective action for change

Unsurprisingly, given the relational ontology of Pacific peoples collective action was seen as a key method for change in our universities.

Pacific women noted that despite critique of inclusion politics these things were needed for change and that everybody was playing their part increasing change:

Yeah, everybody's playing their part, though. We actually need that profiling in the choir. We also need the money, and then we need the people who are part of the advisory group, including [senior pacific male], to help set that up and to validate it as well. So, it's not going to be any one person. - Talanoa fourteen

Coming together is a way of bringing together a collective energy for change that can produce the momentum needed:

together as a collective we have decided that we would put all of our energies together and hopefully, the accumulation of all that energy is going to give us the momentum. - Talanoa Twelve

Collective action as provides strength:

It's strength in numbers, which you've got to have the support of those around you because, you can beat something when there's more of you behind it. It's hard to do by yourself - Talanoa Three

Collective action is an extension of how Pacific families operate and that when we do this, we will all thrive:

Whereas, if we have a collective, we're all moving forward together and we're all helping each other to get to that next level. Whatever it happens to be, it might be cultural skills, it might be language, it might be being comfortable going home to your home country with other professionals here. It might be teaching. It might be doing research. Whatever it takes. Whatever that next step is, and the next step. Just like we'd do in a family. And treat all kids the same, you know. Build on the things that they have and try to see a space for them that will thrive. – Talanoa Ten

Importantly, Pacific women identified that one of the reasons change maybe so slow currently is because there are individuals who drive their own agenda as opposed to working collectively:

Because, oftentimes, I think one of the reasons why change is slow is it's just based on one person's agenda, or what they think people should benefit, rather than this is what people are actually saying. They just maybe go on with their own agenda. It's all about you. – Talanoa Nine

Collective action enables change. Pacific women see collective action as a continuation of community practices and understand that collective action gives the energy, momentum and numbers needed for change. Collective action also enables us to recognize the different contributions that everybody makes to change whether is polite inclusive politics or radical calls for change.

Choosing change

Pacific women see change making as an active choice that can be a decision to go against the university or your colleagues. Pacific women recognize that they have a choice in change, and that sometimes the choice to disengage from change reflects the many battles they have had to engage in over their career, month, week or day:

And it comes from different places, that lack of safety, as well. And I think that sometimes we have more or less resilience to that, as well. So, there will be some times where I'm like, "Right, I'm gonna start a war. I'm gonna go to town on something here, and I have to say this." And then there's other times where I think, "Actually, right now I don't have the resilience to deal with what will come. Is this a stand on the big one, or can I let this battle go today just to save my own sanity?" Talanoa Ten

There is a risk to choosing change, but we engage anyway as we don't want to change who we are in order to be rewarded:

And people who tow the line get promoted, or they sit on committees, or they become the professors in some respects. The people who antagonize the systems like ourselves, and push the boundaries, we're seen as renegades, or we're seen as the problems ... For me, I can't be someone I'm not, and it makes me feel really uncomfortable, for example, that when decisions are made in the system or the institution that don't fit Pacific students' needs, or Maori students' needs, or indigenous students' needs, that we have to really be loud and scream about in some respects, and I hate feeling uncomfortable like that because I don't want to tow that line. I don't want to go against what our students are about, or our pupils are about, because in some respects, that's why our people are down, because people have been towing the line. We haven't been pushing enough boundaries. We've been giving in. We've been colonized, you know? It's like coming back to decolonizing our mindsets, and standing up for who we are and being who we are, but in a positive and a strength-based way, and affirming our identities - Talanoa Seven

For Pacific women change making is a choice that needs to be actively engaged with. Unfortunately, it is a choice that comes with risk and that often those who choose the status quo are rewarded. However, sometimes choosing not to engage with change is not because Pacific women are choosing the status quo but are instead choosing their own wellness.

Mentoring for change

Pacific women recognize the role that mentoring plays in changing their own trajectories. Pacific women often spoke of the roles that mentors played in supporting their wellness or career progression; mentoring was a key factor in change making.

Mentors did not always have to be Pacific, Pacific women noted that non-Pacific mentors who "clicked" were helpful and able to advocate for them around decision making tables in order to support more Pacific to come through the university:

You can have some really good mentors who are not Pacific. It's kind of getting a click with somebody who gets you, you know? There's that, and then there's the people who have opportunity to open up the minds and the awareness of those who sit around decision-making tables about the issues that are involved, and so allow for a lot more of Pacific people coming through - Talanoa two

There is a desire to interrupt any narrative that suggested we progressed without mentors. Instead, there was recognition that it is a collective effort to get the individual to the spaces and places they enjoy:

So that's mentor stuff. I think for anyone in academic career, anyone who manages to get into an academic position would always wanna identify mentors and stuff. It's rubbish that we're doing this on our own, though, really. And there's so many little things along the way to bump you and push you in the right direction. So those two people were really influential... And I sure there were many many mentors and support people throughout my experience that kept me going. - Talanoa four

Peer mentoring and the role of reciprocity and relationality is important in mentoring for Pacific women:

One of them was my writing buddy mentor. She refuses to be known as my mentor, cause she says in that our writing sessions, she gets as much out of it from me as I do from her. So there's that reciprocity that seems to work really well, as well as having really in common some of the basic principles of centrality of family, you know, joy of life, and living big lives. - Talanoa Five

Universities also provide opportunities for mentoring that are not always culturally bounded:

Well, the North Island University put on some great leadership development opportunities. It was a matter of stepping into them, really. So, they have a leadership development program. They have a women in leadership program. They had a chair mentoring program. All fantastic, and they're not culturally bounded. - Talanoa Fourteen

Mentors were the people who encouraged Pacific women to progress with study and the importance of having core qualifications to move through the university

it wasn't until I talked with a senior Pacific woman about how we're going to get to where we need to go? Well, we need some associate professors, we need some professors, because they're the people who will be in the room. That there's certain jobs that a university won't give you until you have that tick box and a PhD is one of them, but certainly the associate professor/professor stuff is one of them too. - Community Talanoa One

Mentoring plays a key role in providing critical advice about how to do our jobs:

And the one enduring practice that all our mentors have shared with us, whatever you do, you do it now. Make sure that you have the resources. Make sure that you share with others, and tell others what you want to teach and how you want to teach it. And also make sure that you have the time off. So it's about maintaining equilibrium within the teaching ratios and all that, and also the research and that - Talanoa Six

Mentoring was identified by Pacific women as a key practice for change. Mentors came in many different forms as peers, institutional, Pacific and non-Pacific. Mentors were seen as key informants for explaining institutional norms and progression advice. Mentoring for change reflects that Pacific women understand there is an aspect of personal development needed to advocate for change, to ensure that Pacific women can sit "at the table."

Conclusion

So you take the training that life, and universities, and research, and your own reflections on who you are and what you've got, and you put that into the policy government environment. All of those tools are needed, along with a good dose of relations, relationships. Then you just set about making the change that's needed. - Talanoa Fourteen

It is unsurprising that Pacific peoples have multiple approaches to change. Universities are complex institutions that require interventions at multiple sites in order to drive the long-term structural change that is needed to advance Pacific peoples. What these talanoa share is a recognition of both the capacity of higher education to be a positive force while also recognizing the behaviors and structures within higher education that need to change. While change is slow it needs to be recognized that Pacific women push this change in their everyday work in the university through sitting at tables, operating as a collective, mentoring each other, working with allies and actively choosing change.

A future question is, I think it's about potentiality. Maintaining that there are always possibilities. Anybody can do it. You're going to come across a lot of changes, but for Pasifika woman, it's about just getting it done. And always maintaining where you've come from, and celebrate that and celebrate the present, but also try and leave something behind. - Talanoa 6

Post-script

I completed this project before getting my first permanent academic job. Reflecting on these Pacific women's words has been a powerful and inspiring moment. At times as an early career Pacific researcher I can get lost in the weeds of the everyday. Revisiting these words have reminded me that I am not alone in this fight, that Pacific women have agitated for change in many ways and at many different levels for some time now. Importantly though, we have done so together. That is the takeaway from this work – that for anything to truly change we must advocate together and resist the temptation to embrace the neoliberal individualistic academy that would turn us against each other as we labor in the everyday work of change making.

The reason we can't pinpoint those moments of change-making is because our every moment in that institution is change-making. Our presence in those offices, our presence in those conversations is a moment of change. So how do you pinpoint that one moment, when every day what we do is change-making? - Community Talanoa Two – Sereana Naepi

Acknowledgment

I would like to acknowledge the University of British Columbia Public Scholars program for supporting this research. I would also like to thank the Pacific women who joined me in talanoa – I often reflect on what it meant to sit with you all so early on in my research career and how that has shaped my pathway forward Vinaka vaka levu.

References

- Ahenakew, C., 2016. Grafting indigenous ways of knowing onto non-indigenous ways of being: the (underestimated) challenges of a decolonial imagination. *Int. Rev. Qual. Res.* 9 (3), 323–340.
- Ahenakew, C., Naepi, S., 2015. The difficult task of turning walls into tables. In: Webber, M., Macfarlane, S., Macfarlane, A. (Eds.), *Sociocultural Realities: Exploring New Horizons*. University of Canterbury Press, Christchurch, pp. 181–194.
- Ahmed, S., 2012. *On Being Included*. Duke University Press.
- Airini, Curtis, E., Townsend, S., Rakena, T.O., Brown, D., Sauni, P., Smith, A., Luatua, F., Reynolds, G., Johnson, O., 2011. Teaching for student success: promising practices in university teaching. *Pacific-Asian Educ. J.* 23 (1).
- Barber, S., Naepi, S., 2020. Sociology in a crisis: Covid-19 and the colonial politics of knowledge production in Aotearoa New Zealand. *J. Sociol.* 56 (4), 693–703.
- Barber, S., Naepi, S., 2022. Introduction. *N. Z. Sociol.* 37 (1), 1–19.
- Fa'avae, D., Jones, A., Manu'atu, L., 2016. Talanoa'i 'A e Talanoa—talking about Talanoa: some dilemmas of a novice researcher. *AlterNative* 12 (2), 138–150.
- Farrelly, T., Nabobo-Baba, U., December 2012. Talanoa as empathic research. In: *International Development Conference*, p. 12.
- Fleras, A., Spoonley, P., 1999. *Recalling Aotearoa: Indigenous Politics and Ethnic Relations in New Zealand*. Oxford University Press, USA.
- Health Research Council, 2014. Health Research Council Pacific Guidelines. Retrieved December 29, 2017, from New Zealand Health Research Council: <http://www.hrc.govt.nz/sites/default/files/Pacific%20Health%20Research%20Guidelines%202014.pdf>.
- Kidman, J., 2020. Whither decolonisation? indigenous scholars and the problem of inclusion in the neoliberal university. *J. Sociol.* 56 (2), 247–262.
- Kidman, J., Chu, C., 2017. Scholar outsiders in the neoliberal university: transgressive academic labour in the whitestream. *N. Z. J. Educ. Stud.* 52 (1), 7–19.
- Kidman, J., Chu, C., 2019. 'We're not the hottest ethnicity': pacific scholars and the cultural politics of New Zealand universities. *Glob. Soc. Educ.* 17 (4), 489–499.
- Leenen-Young, M., 2020. Pasifika students and learning to learn at University. *MAI J.* 9 (1).
- Leenen-Young, M., Naepi, S., Thomsen, P.S., Fa'avae, D.T.M., Keil, M., Matapo, J., 2021. "Pillars of the colonial institution are like a knowledge prison": the significance of decolonizing knowledge and pedagogical practice for Pacific early career academics in higher education. *Teach. High. Educ.* 26 (7–8), 986–1001.
- Leong, N., 2012. Racial capitalism. *Harv. Law Rev.* 126, 2151.
- Macpherson, C., 2004. From Pacific islanders to Pacific people and beyond. *Tangata tangata: The changing ethnic contours of New Zealand* 135–155.
- Macpherson, C., Anae, M., Spoonley, P., 2001. *Evolving Identities of Pacific Peoples in Aotearoa/New Zealand*. Dunmore Press.
- Manueli, K.T., 2012. *Pasifika Tertiary Students' Use of Information and Communication Technologies*. Doctoral Thesis. Victoria University of Wellington. <http://researcharchive.vuw.ac.nz/handle/10063/2480>.
- McAllister, T.G., Kokaua, J., Naepi, S., Kidman, J., Theodore, R., 2020. Glass ceilings in New Zealand universities. *MAI J.* 9 (3), 272–285.
- McAllister, T.G., Naepi, S., Wilson, E., Hikuroa, D., Walker, L.A., 2022. Under-represented and overlooked: Māori and Pasifika scientists in Aotearoa New Zealand's universities and crown-research institutes. *J. Roy. Soc. N. Z.* 52 (1), 38–53.
- Ministry of Education, 2013. *Pasifika Education Plan 2013–2017*. New Zealand Ministry of Education. <https://assets.education.govt.nz/public/Documents/Ministry/Strategies-and-policies/PasifikaEdPlan2013To2017V2.pdf>.
- Ministry of Education, 2021. *Action Plan for Pacific Education 2020–2030*, New Zealand Ministry of Education. <https://www.education.govt.nz/our-work/overall-strategies-and-policies/action-plan-for-pacific-education-2020-2030/>.
- Musser, A.J., 2015. Specimen days: diversity, labor, and the university. *Femin. Format.* 1–20.
- Nabobo-Baba, U., 2008. Decolonising framings in Pacific research: indigenous Fijian Vanua research framework as an organic response. *AlterNative* 4 (2), 140–154.
- Naepi, S., 2019a. Why isn't my professor Pasifika. *Maine Bar J.* 8 (2), 219–234.
- Naepi, S., 2019b. Masi methodology: centring Pacific women's voices in research. *AlterNative* 15 (3), 234–242.
- Naepi, S., 2019c. Pacific research methodologies. In: *Oxford Research Encyclopedia of Education*.
- Naepi, S., 2020. "I didn't come to play": Pasifika women in the academy. In: Moeke-Pickering, T., Cote-Meek, S., Pegoraro, A. (Eds.), *Critical Reflections and Politics on Advancing Women in the Academy*. IGI Global, pp. 52–69.

- Naepi, S., 2021a. Shifting tides: the potential and limitations of critical mass to change discipline research habits. *J. N. Z. Stud.* (NS33).
- Naepi, S., 2021b. Pacific women's experiences working in universities in Aotearoa New Zealand. *High Educ. Res. Dev.* 40 (1), 63–74.
- Naepi, S., McAllister, T.G., Thomsen, P., Leenen-Young, M., Walker, L.A., McAllister, A.L., Theodore, R., Kidman, J., Suaalila, T., 2020. The Pakaru "pipeline": Māori and Pasifika pathways within the academy. *N. Z. Annu. Rev. Educ.* 24, 142–159.
- Naepi, S., Stein, S., Ahenakew, C., Andreotti, V.D.O., 2017. A cartography of higher education: attempts at inclusion and insights from Pasifika scholarship in Aotearoa New Zealand. In: Reid, C., Major, J. (Eds.), *Global Teaching*. Palgrave Macmillan, New York, pp. 81–99.
- Naepi, S., Wilson, E., Lagos, S., Manuela, S., McAllister, T.G., Kidman, J., Theodore, R., Kokaua, J., 2021. Where are we now? Patterns of Māori and Pasifika enrolment in the natural and physical science and society and culture fields in Aotearoa New Zealand. *High Educ. Res. Dev.* 40 (1), 90–103.
- Otunuku, M.A., 2011. How can talanoa be used effectively an indigenous research methodology with Tongan people? *Pacific-Asian Educ.* 23 (2), 43–52.
- Prescott, S.M., 2008. Using talanoa in Pacific business research in New Zealand: experiences with Tongan entrepreneurs. *AlterNative* 4 (1), 127–148.
- Stewart-Withers, R., Sewabu, K., Richardson, S., 2017. Talanoa: a contemporary qualitative methodology for sport management. *Sport Manag. Rev.* 20 (1), 55–68.
- Suaalii-Sauni, T.M., 2008. Critiquing Pasifika education at university. *AlterNative* 4 (1), 14–23.
- Suaalii-Sauni, T., Fulu-Aiolupotea, S.M., 2014. Decolonising Pacific research, building Pacific research communities and developing Pacific research tools: the case of the talanoa and the faafaletui in Samoa. *Asia Pac. Viewp.* 55 (3), 331–344.
- Te Punga Somerville, A., 2012. *Once Were Pacific: Māori Connections to Oceania*. University of Minnesota Press, Minneapolis.
- Te Tai Ōhanga The Treasury, 2020. Estimates of Appropriations: Vote Tertiary Education—Education and Workforce Sector—Estimates 2020/21, New Zealand Te Tai Ōhanga the Treasury. <https://www.treasury.govt.nz/sites/default/files/2020-05/est20-v2-tered.pdf>.
- Tertiary Education Commission, 2021a. Performance-Based Research Fund improvements. New Zealand Tertiary Education Commission. <https://www.tec.govt.nz/news-and-consultations/performance-based-research-fund-improvements/>.
- Tertiary Education Commission, 2021b. Equity Funding for 2022. New Zealand Tertiary Education Commission. <https://www.tec.govt.nz/funding/funding-and-performance/funding-fund-finder/equity-funding/equity-funding-for-2022/>.
- Theodore, R., Taumoepeau, M., Kokaua, J., Tustin, K., Gollop, M., Taylor, N., Hunter, J., Kiro, C., Poulton, R., 2018a. Equity in New Zealand university graduate outcomes: Māori and Pacific graduates. *High Educ. Res. Dev.* 37 (1), 206–221.
- Theodore, R., Taumoepeau, M., Tustin, K., Gollop, M., Unasa, C., Kokaua, J., Taylor, N., Ramrakha, S., Hunter, J., Poulton, R., 2018b. Pacific university graduates in New Zealand: what helps and hinders completion. *AlterNative* 14 (2), 138–146.
- Thomsen, P., Leenen-Young, M., Naepi, S., Müller, K., Manuela, S., Sisifa, S., Baice, T., 2021. In our own words: Pacific Early Career Academics (PECA) and Pacific knowledges in higher education pedagogical praxis. *High Educ. Res. Dev.* 40 (1), 49–62.
- Tuhiwai Smith, L., et al., 2020. E koekoe te tūi, e ketekete te kākā, e kūkū te 27ereru: The Report of the PBRF Review Panel. <https://www.education.govt.nz/assets/Documents/Further-education/PBRF-Review/The-Report-of-the-PBRF-Review-panel-E-koekoe-te-tuie-ketekete-te-kaka.-.pdf>.
- UniversitiesNZ, 2018. Achieving Parity for Māori and Pasifika—The University Sector View. UniversitiesNZ. <https://www.universitiesnz.ac.nz/sites/default/files/UNZ%20Parity%20Discussion%20Paper%20One%2028Aug%202018%29.pdf>.
- Vaioliti, T.M., 2006. Talanoa research methodology: a developing position on Pacific research. *Waikato J. Educ.* 12.

Is there a Dr in the house: building an indigenous doctoral program

Patricia MG. Johnston-Ak, Te Whare Wananga o Awanuiaranig, Ngai te Rangi, Ngati Ranginui, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Setting the context for why we need an indigenous doctorate	129
Setting the context for why we need an indigenous institution (step 1)	130
Developing a school of Indigenous Graduate Studies (step 2)	131
Developing an indigenous doctoral program (step 3)	132
New Zealand accreditation and approval of the program (step 4)	133
Acknowledgment	134
References	134

This article will examine the development of an indigenous doctoral program from Aotearoa, New Zealand. The program is located within an Indigenous institution, and is influenced by philosophies, methodologies and practices that are embedded within indigenous ways of knowing and doing.

The establishment of the program involved community ideas and expectations in the development phase that has translated into a program wherein students “give back” to their communities through their research and the solutions evident in that research. The focus is one of indigenous development and advancement that requires a commitment by students to create solutions to existing issues and matters deemed important by their communities or indigenous organizations.

Setting the context for why we need an indigenous doctorate

The main doctorate offered by institutions within the New Zealand tertiary (higher education) context is the Doctor of Philosophy (PhD), which is carried out over a minimum of two to four years of study. Middleton (2001, p. 4) states that the origin of the New Zealand PhD was in “the soil of British colonialism and the early provincial university ... were colonial sites for the assimilation of British national culture or civilization.” Māori engagement with these doctoral programmes has been through these very colonial sites, which have been hostile to Māori knowledge, practices and belief systems because the assimilatory agenda to civilize Māori had resulted in the banning of Māori language, restrictions placed to suppress Māori knowledge and the general overarching desire to introduce Māori to the ways and thinking of the British.

Additionally, Māori were also actively dissuaded from attending universities through practices like those at secondary schools, which limited their options and choices. For example, restricting curriculum choices for Māori resulted in them not meeting the entry criteria to attend university, options that were driven by policies and practices based on racial categorizations, which argued that Māori were more suited to menial/manual work rather than academic (Barrington, 1988).

While access to universities was theoretically achievable for those who could put aside their “Māori ness”, being able to enter the university system had its own set of challenges. Doctoral enrollments required individuals to have a foreign language requirement, English being the main language with Māori language not considered viable. Doctoral study for many Māori was seen as intimidating, foreign and unobtainable (Johnston, 2010) and with the earlier practices of undertaking doctoral study at a foreign university, restricting further the chances of Māori to engage in doctoral work.

In 2001, Sue Middleton completed research on “New Zealand Education PhD’s 1948–1998” in which she explored several factors that further contributed to and/or hindered the PhD process for students. Middleton (2001) examined students family backgrounds, the role those tertiary institutions played in supporting students, supervision practices, examination practices etc. Some Māori were involved in that study as participants.

Likewise, my own research in this area (Johnston, 2010) highlighted a few factors that dissuaded Māori from seeking doctoral qualifications which concurred with those of Middleton (2001) but additionally highlighted tribal and whanau (family) commitments often taking precedence over doctoral study, which saw many PhD students taking longer to complete their qualifications.

Between 1952 and the end of the year 2000, the number of Māori to graduate with PhD’s in New Zealand was small (approximately 55) but over the past decade there have been increasing enrollments of Māori into PhD programs (Ministry of Education, 2001). In 1999, history was made when 3 Māori women graduated PhD from the University of Auckland (Garner, 2000) and at the end of 2000, 3 Māori men from Massey University, the first time that more than one person of Māori descent graduated PhD in any given ceremony at any specific university.

The first person of Indigenous descent (Māori) to graduate with PhD was Maharaia Winiata (Tribes: Ngai te Rangi, Ngati Ranginui) from Edinburgh University in Scotland 1952. Between then and 1990, several more Māori received their doctorates from overseas universities with the first group of consolidated New Zealand based PhDs being produced in the 1990s (Pohatu, 2001). The latter is perceivable because of a shift in university practices regarding the use of Māori language and the incorporation of Māori language as either a linguistics subject or a Studies positioned under anthropology. This space however was still contested.

The difficulty of precisely defining the exact number, is a difficulty of working within a colonial climate whereby identifying as Māori was not necessarily a positive action because of ongoing racial beliefs and views that presided within New Zealand about the ability of Māori. The data on actual Māori graduates with PhD is therefore inconclusive. However, the 2007 National Māori Academic Excellence Awards noted and celebrated 183 graduates between 2002 and 2007 ([Prov-Vice Chancellor Māori Office, 2007](#)).

There is also evidence that Māori are gaining PhD in areas like Chemistry and Science, where historically they have not been located before. The increase of graduates means that universities can employ doctoral qualified Māori staff where previously, none were to be found. While the number of Māori graduating with PhD is on the increase, the numbers are still disproportionately lower than those for Pakeha students (members of the dominant group) ([Tertiary Education Commission, 2009](#)).

Research focusing explicitly on Māori with PhD is minimal although a three-year project investigating specifically the supervision of Māori Doctoral Students ([McKinley et al., 2007](#)) has been completed. Given the growing numbers of doctoral candidates ([Tertiary Education Commission, 2009](#)), this is an area within education that needs further investigation to enable us to identify and understand what contributes to success for Māori students, and what does not.

The focus on producing more Māori graduates has also been supported through the activities of organisations who are developing programmes to support Māori Doctoral students, for example Nga Pae o te Maramatanga—a National Institute of Research Excellence for Māori based at the University of Auckland and its national Program for Māori and Indigenous Post-graduate Advancement (MAI). Further incentive is perceivable in funding, like for example an equity component that rewards institutions for graduating Māori PhD (Performance Based Research Fund Research Degree Component) and policy strategies bedded into the [Māori Tertiary Education Framework \(2003\)](#) and [Tertiary Education Strategy 2007–2012 \(2007\)](#). Given this type of focus the likelihood of continual graduates, growth and development of Māori with PhD looks inevitable.

Setting the context for why we need an indigenous institution (step 1)

In 2005, Te Whare Wānanga o Awānuiārangi (Indigenous Tribal Institution based in Whakatane, New Zealand) enrolled their first students into their own PhD program. The program had been approved the previous year by the New Zealand Qualification Authority (Government Accreditation Organization) after several years of application that included substantial challenges leveled at a non-University institutions desire to offer a PhD program. While education policies at that time did not declare that Wānanga could offer doctoral programmes, neither did they state that they “could not” and it was in this latter space that Awānuiārangi contested the right to offer a PhD program.

Traditional Whare Wānanga existed in New Zealand prior to colonization and were considered the highest places of learning where the teaching of information was subjected to exceptionally rigorous and demanding processes, accessed by “experts” and superior families and not available to all. Whare Wānanga were clearly places of excellence, focused on quality that insisted on only the best. Indeed, the mission statement of Te Whare Wānanga o Awanuiārangi reads:

Rukuhia te mātauranga ki tōna hōhunutanga me tōna whānuitanga

Pursue knowledge to its greatest depth and broadest horizons.

Te Whare Wānanga o Awanuiārangi has qualities that make its contribution to the tertiary sector unique and necessary, unique in terms of āhuatanga Māori (Māori ways of knowing and doing) and necessary in terms of Māori decision-making by Māori for Māori. It offers programmes that are “niche” targeted, (meaning designed for and with Māori communities), programmes that are complementary to university and other tertiary operations, while resisting the assimilation of things Māori to be “described, labeled and categorized in a Western manner” ([Johnston, 1998](#)).

In terms of offering a doctoral program, the distinctive characteristics of Te Whare Wānanga o Awanuiārangi and its contributions should not be underestimated. The general perception of the Wānanga is one of being a “Māori institution”, interested in offering only “Māori things” to a largely “Māori audience”. However, section 162, 4(b) iv of the 1989 Education Act provides a broader definition.

A wānanga is characterized by teaching and research that maintains, advances and disseminates knowledge and develops intellectual independence, and assists the application of knowledge regarding āhuatanga Māori (Māori tradition) according to tikanga Māori (Māori custom).

[Mead \(2003\)](#) states that the definition of āhuatanga and tikanga Māori in the act is not clear and is slightly restrictive as the English equivalents do not do justice for the concepts of āhuatanga and tikanga. In viewing the first section of the clause however, there are several key responsibilities that a wānanga is charged with over and above āhuatanga and tikanga Māori, which are often not highlighted. The key words are “teaching” and “research”, “maintain”, “advance” and “disseminate”, “knowledge”, “intellectual independence” and “application”, the same sorts of activities one would expect from any tertiary institution—activities of quality, excellence and rigor. Thus, Wānanga are expected to perform according to national criteria (for example, research and teaching),

internationally set benchmarks (for example, the PhD program), and are subjected to the same forms of evaluation, monitoring and moderation as other tertiary institutions. And there the similarity with other institutions ends.

Wānanga are also expected to perform to tribal criteria, tribal benchmarks and are subjected to intense tribal evaluation, monitoring and moderation—which other tertiary institutions are not. While other tertiary institutions enjoy measures of autonomy and decision-making, Awanuiārangi must always consider the desires, visions and views of those tribes whom it works with and the Māori/indigenous communities that it serves—to build educational capability, dissemination of language and knowledge, to teach and research. This is one of the many unique features of this institution.

As a result of our mandate from the 1989 Education Act, Awānuiārangi is charged to consider how the development of new degrees and programmes are responsive to community needs, goals and interests and it is those needs etc., which drive what Te Whare Wānanga o Awanuiārangi produces. Those developments are different from what is offered through other tertiary institutions because of our dual national/tribal requirements.

A primary vision of the institution is to promote, grow and sustain all aspects of Māori language, knowledge and culture and have regard to tikanga Māori practice. Te Whare Wānanga o Awanuiārangi has also developed a range of Māori educational opportunities and advancements designed to positively impact the social, economic and cultural growth of Māori communities and iwi. Being able to offer a doctoral program was key to meeting those aspirations, thus incorporating the visions and aspirations of indigenous peoples more generically. The breadth of our learning provision is therefore expansive (based on need and aspirations) and extends from marae (traditional Māori community places) and community-based learning programmes up to graduate programmes at Master's and Doctoral levels. These programmes focus on the theme of Māori development and advancement and benchmark to a wider body of international indigenous research activity.

Developing a school of Indigenous Graduate Studies (step 2)

Te Whare Wānanga o Awānuiārangi launched the School of Indigenous Graduate Studies in June 2008, which was charged with providing a graduate program that incorporated quality learning/teaching experiences for students. Awanuiārangi was accredited two Master's Degrees in 1994 and as the inaugural Head of Graduate Studies (2004) and then the School of Graduate Studies (2008) one of my roles was to pursue accreditation of a Doctoral Program. This led as mentioned previously, to offering of a Doctor of Philosophy in 2004 in four areas: Indigenous Studies, Māori Studies, Education and Environmental Studies.

To support the fledgling doctoral program, a number of aspects in the School were created that drew from indigenous networks and ways of operating. For example, the creation of a teaching adjunct faculty of indigenous/Māori PhD, (many of whom are furthering iwi/hapū/tribal developments in their own spaces) was enabled, linking students into community networks or research centers. The School also drew from this pool of faculty members to offer students co-supervisors based in other areas and institutions, who could add to a student's topic by providing networking support and the latest up to date research information in their fields. The adjunct faculty members complimented Te Whare Wānanga o Awanuiārangi current supervision pool.

The growing demand for Doctoral Study was met through the growing number of PhD credentialed staff via contractual secondary supervision arrangements and a Graduate Adjunct Teaching Faculty comprising Indigenous PhD holders from within New Zealand (and from around the world), who were available for thesis supervision. These developments meant that academic programmes like our PhD were drawing from the experience, knowledge and vision of Māori at national and international levels and indigenous scholars internationally.

Other measures for providing support included inviting colleagues and friends in other institutions and communities to support students as advisors, mentorships for supervision. Additionally, four annual "block courses" were run (this is not usually done for doctoral students in other institutions) that saw persons with expertise invited to work with students, and for students themselves to work and be connected to each other, thus creating collegiality and support networks. Within the block courses, students stayed on site at the Wānanga having freedom and space to work independently or to further conversations and develop networks for and with each other, as their research and study necessitated.

Other processes included an orientation process for all new intending students, several noho (staying on site) throughout the year for supporting doctoral writing, attendances at conferences, seminars and writing workshops to support students' needs. A key element of the doctoral program was the high level of interaction with other indigenous graduate students and scholars internationally. In 2010, doctoral students attended and presented at international conferences as part of an international outreach and the School of Indigenous Graduate Studies also hosted and supported students to attend the national MAI conference in 2011 where two groups of indigenous students from Australia attended. Such attendances continued until Covid 19 intervened in 2020.

However, there was a problem with the PhD program one of control exerted over the structure and underlying philosophies of the program, which had its roots in the "colonial sites for the assimilation of British national culture", the very sites that the Wānanga was contesting and struggling against. So, while Te Whare Wānanga o Awānuiārangi gained approval to operate the PhD program, that very program sat within a certain mindset that did not allow Māori students to be indigenous in their ways of thinking and knowing, and which expected conformity to University set agenda. While as an institution we struggled to be as good as the universities when it came to the doctoral program, the truth of the matter was, that was never a "level playing field" that we could engage on or could hope to change to be conducive to indigenous ways of thinking: tinkering with the system did not change it!

Developing an indigenous doctoral program (step 3)

It was time to rethink the whole nature of what came to count as an Indigenous Doctorate. That rethinking was escalated in 2011, when we were approached by Professor Alan Parker of Chippewa Cree (Rocky Boy's Reservation in Montana), the co-coordinator for the Master of Public Administration, Evergreen College in Olympia, Washington State who retired July 2012. While visiting the Wānanga, Professor Parker and I had an opportune moment to discuss Graduate programmes and he proposed that Te Whare Wānanga o Awanuiārangī should meet with graduates from the Master of Public Administration Degree to establish if the students could be considered for enrollment into the Doctor of Philosophy.

Initial discussions with potential students occurred at the end of 2011. In those discussions (and in considering some of the shortcomings of our PhD program), I began to think that our PhD program was not going to be a good match for these potential students for several reasons. Firstly, most of them had a Master's that prepared them for entry into a "taught" doctorate (which is the model in the United States) and that the base for their Masters (particularly in the research thesis/dissertation area) was not as strong as in the New Zealand program.

Secondly, New Zealand PhD programmes relies on self-direction and independent study, which I thought might end up potentially isolating the students (which had occurred to a certain degree with our own Māori students), made worse by the fact that they were in Washington State (and our team was based in New Zealand).

Thirdly, I was still not convinced that the PhD program could enable transformation for indigenous communities. Developments for change were reliant on the commitment of the student to create change, an aspect of research and doctoral studies that I was becoming more aware of and committed to. It was at this point that I started to think about a practical type of doctorate (a professional doctorate) that was quite different from anything offered in New Zealand and which had the potential to meet the requirements for transformation and change for communities.

My own doctorate for example (PhD), sits on a shelf, is referred to by students and colleagues but for the most part, gathers dust. That is not the type of situation I envisaged for our students. This professional doctorate would need to be niche driven, to achieve a number of objectives that the PhD program could not. In discussion with Dr. Parker, it was time then to get out into the communities to see what they needed and wanted.

Over the next year, Dr. Parker and I visited several tribal groups (in Washington State and later in Hawaii), met with their councils and various members to discuss the possibility of offering a doctoral program into Washington State and what that might look like. The tribe's involved in that discussion were Lummi, Swinomish, Tulalip, Puyallup and Chehalis. In those discussions, there were some extremely forthright conversations about the nature of doctoral studies in Washington State and two main points were made that highlighted some of the concerns tribal members had about the doctoral programmes available to their community members.

Predominantly they spoke of their young people going off to university, gaining a doctorate and not returning because they were "seduced" by the academy. They noted that often these graduates were the only tribal person employed in a department (placing them at risk) with opportunities to engage in other activities and at other levels which the tribes could not compete against. The chances of them returning to work with and for their tribe was limited. Discussants asked if our doctoral program could keep their community members in the community?

A second area of discussion that required thinking through what a professional doctoral program might look like, related to communities highlighting that some tribal graduates returned to work with their communities having learned "strange ways of knowing, and doing strange things". Some tribal members were quite blunt about pointing out that these types of graduates were not helpful for community and tribal developments.

As discussions progressed with tribal communities, it was confirmed in my thinking, that a straight-out research doctorate (PhD) was not where we would be going and that a Professional "taught" doctorate (with a significant research component) was the way that we should be heading.

As tribal communities aired their concerns and raised issues, we responded (always "face to face") which required many meetings, discussions and planning. The role of Dr. Parker in this process was integral: his relationship with many of the Chairs of the tribal councils stemmed from his previously having served on the Washington State Gambling Commission for five years and he was known to many of the tribes throughout Washington State through that role. He was highly respected and his support of our endeavors to offer an Indigenous doctoral program was an endorsement of what we were attempting to achieve. I always admired that whenever we turned up at any of the tribal offices, his ability to meet with the Chairs (or educational representative) in those Tribes, occurred very quickly.

To seek further support for a doctoral program, we also had conversations with Evergreen College and North-West Indian College (Lummi territory). This was to gauge possibilities for relationships moving forward, that later extended into conversations (and a memorandum of understanding) with the University of Washington (Tacoma).

We were also joined by Professor Margie Maaka (University of Hawaii) who was interested in the development of a specific indigenous type of doctoral program for Hawaiian students. Her contacts with other indigenous educators (throughout Hawaii and Washington State) were to prove invaluable in the thinking phases of a new doctoral program. Our times together meeting with tribes, potential students and communities enabled the three of us to talk and think through this intended new development. We were also fortunate to later draw on the knowledge and expertise of Dr. Sharon Parker who worked at the University of Washington in Tacoma, who was able to provide much of the background information to the workings of doctoral studies in Washington State.

New Zealand accreditation and approval of the program (step 4)

Mid 2013, I began in earnest to develop and write up the Professional Doctorate taking into consideration the discussions and points outlined by the Washington State Tribes, and taking into consideration evaluations, comments and feedback to our current PhD Program within Aotearoa New Zealand and iwi (tribal) groups. My experiences with the PhD program (and what did not work well for our students) became an important part of thinking through what the new doctoral program would look like. This new doctorate had to be “niche” specific, created and written with a precise purpose to assist desired outcomes of communities and tribes, unlike the PhD which focuses on individual students and relies on their own personal commitment to implementing their research findings.

What evolved over four months, was a comprehensive program that took into account many of the issues, aspirations and desires of those Indigenous Communities who had indicated their support for such an endeavor. The program also met the academic criteria required for a Doctoral Program in New Zealand and was accredited (approved for delivery) by our National Accreditation organization (New Zealand Qualification Authority) in December 2013. It was called The Doctor of Indigenous Development and Advancement (with a second program available for Māori students called The Doctor of Māori Development and Advancement). The name was important because it signaled the intent of the Degree. Both programmes were first offered in 2014 with the Indigenous focus for Washington State, and both the Indigenous and Māori focus within New Zealand. There are several unique philosophical and operational aspects of the program outlined here that were put into place as a result of the discussions with tribal communities.

For example, students’ intending to enroll into the program were required to be supported and endorsed by their tribe/iwi in their endeavor to undertake their research. This was to such an extent that community members and tribal councils could (in negotiation with students) be involved in various stages of the students’ enrollments and/or study. In Washington State, this negotiation culminated in scholarships for students’, included advice on topic selection and (once research proposals were completed) ethical approval. The result is that students are tied to their research and communities because they literally must seek permission and support to do their research from tribal members and the council before enrollments are accepted. This also keeps them in the community to do their research.

The doctoral topics that students selected, had to also be relevant to the tribe that the student affiliated to. Students were asked to consider developing models, structures and frameworks to assist their communities to change. As a result, topics focused on what could work, (rather than what does not work - there is enough of this type of research in existence), thus enabling the potential for their research to be transformative and to create change.

Significantly, the first two years of the program focused on coursework. In that time and through those courses (Literature Review, Methodology, Methods and Research Proposals), students were exposed to indigenous theories, practices, research, other indigenous academics and community which addressed concerns about students learning “strange ways of knowing and doing things”. In the first two years, students were assisted to rethink how they had come to think about and know about research. For many students this represented a significant breaking away from that colonizing gaze that previous study had trained into them: a western perspective about “what counted as research” (Johnston, 1998). This decolonization process was an important aspect of the first two years in the program.

For example, one student who was writing a revisionist history of her tribe initially began by wanting to engage with the non-native anthropologists and historians (western perspectives), which would have set her whole research along a very different pathway. Her work would have acknowledged those accounts as the valid accounts (even if she were just critiquing them). As her supervisor, my conversations with her included questions like why would you do that, why would you privilege their voices? Instead, she began with the narratives and stories of her elders placing their knowledge at the forefront of her research, which created a much more powerful account for her tribe’s account of their history, then what an analysis of non-native accounts would have produced. Her whole dissertation (Dupuis, 2018) is her tribe’s story of how they came to be in the place they are located, their relationship with other tribes and an in-depth acknowledgment of their own cultural practices. I advised her to either place the non-native accounts at the end of her work and pick them up there (if she felt the need to) or let them drop off altogether. Such conversations enabled a re-centering away from western and colonial frameworks for the research, to clearly acknowledging and centering indigenous one’s.

Students were encouraged to think in terms of indigenous methodologies and practices in their work. This has potentially been the most exciting part of their program to date, seeing how they integrate their own cultural ways of knowing and doing into the very core of undertaking the research, which influences the research framework and how they collect and interpret their data and information.

Another student for example (Cultee, 2020), spoke to a specific practice around oral traditions and how in her working with elders, she was not allowed to tape or write down what she was told. Instead, she had to memorize the information and give it back (correctly) before she was given new information. As a result, she spoke of how her ability to “remember” increased over the time she was engaging in the research and this practice enabled her to challenge biased ideas about the simplicity associated with oral traditions and remembering. Her research framework is from her own tribe’s knowledge framework.

In their final term of coursework, students prepare a research proposal and deliver their first oral presentation about their research. Whanau (family) and tribal members are encouraged to attend those presentations and ask questions/pass comments. The presentation is also assessed by a panel of three academics that can include (in the case of Washington State students) indigenous academics from Washington State. The role of this oral, is at one level to gain support (and critical feedback) for the research

that students are undertaking, but at another level for them to be accountable to those whom the research intends to include by talking through what the research is about.

The second oral presentation is in “defense” of their research however, we see this more as a formality with the expectation that because of the processes leading up to that oral, students would pass through this stage as a matter of process. This second oral also enabled students to disseminate their research findings back to their communities and participants who once again, could attend the oral presentation. Communities get to “witness” the research and students are enabled to feed-back their findings as a part of their reciprocity and accountability to their communities for what they have produced.

Our accountability as in indigenous institution (Te Whare Wānanga o Awanuiārangi) came at other levels that included operational. In Washington State for example, students attended classes for the coursework and our teaching/learning spaces accommodation etc., were hosted by their tribes. Students would ask that their next class be held in their territory and this helped to build up a rapport with tribes, while also enabling us to give back (financially) to those tribes who supported our program. The importance also of doing this, was that often we would be asked to give “a talk” on a whole range of different issues to the broader community or education groups and so reciprocity (and accountability to those tribes) came in many different forms.

From November 2011, we met with 47 potential candidates however, the interest in the program extended to include students from Washington State University, Hawaii and University of British Columbia. Word spread about the program but there were difficulties that had to be worked through. Most significantly the availability of funding for our international students precluded many from participating because of federal rules about funding to external (New Zealand) organizations. Some tribes found ways to provide scholarships and there was many a tribal policy change to enable students to enroll into our doctoral program.

Currently, Covid 19 and the impact on our students both within New Zealand and Washington State, has to a certain extent, halted the Professional Doctorate because of the impact on students’ communities. Although we moved to an online delivery in 2020, this is not in keeping with some of the philosophy of meeting “face” to “face” and the support that is generated by being together in the same room. We continue to work through those concerns.

The program is still evolving as has my role in it. I shifted from being the developer of the program, negotiator with communities and tribes, administrator of the students and institutional requirements, teacher/learner with the students (and their communities) coordinator (with overall oversight) to now being that of a supervisor, colleague and mentor. Others now guide and shape the program.

I have supervised 3 students to completion in Washington State and several in New Zealand and I acknowledge in closing their willingness to be involved in the Professional Doctoral Program and for persevering while we worked through the details. As a part of that acknowledgment, I include here the words on our website from Dr. Marla Conwell, our first international indigenous graduate who graduated in 2018 (Chehalis Confederated Tribes) and who also received Te Onehou Eliza Phillis Award for Outstanding Iwi (Tribal) Research.

She states: I completed my Doctorate as an international doctoral candidate based in my home city of Olympia in Washington, USA. After my Masters, I had not intended to pursue a doctorate initially but was compelled to join because of the Awanuiārangi Indigenous Studies program. I have been blessed with the opportunity to further my education by my Chehalis Tribe. This doctorate was the perfect way to repay that faith in me. I focused on indigenous language revitalization—a personal passion of mine ... I wanted to contribute to the indigenous language revitalization field in the hope of assisting indigenous groups who want to revitalize their own languages.

My experience was simply amazing! As an international student based in the United States, I was fortunate that we met for week-long classes each quarter in addition to the online component, and my cohort maintained frequent contact outside of class for study sessions and writing. One of the benefits of studying with Awanuiārangi was the focus on decolonization and indigenous ways of knowing. This was life affirming. It caused me to re-evaluate my life and experiences as well as develop my critical thinking skills and approach to the colonized education and life systems in which we are often mired.

Would I recommend this program to other international scholars? Absolutely—it is one of the greatest accomplishments of my life. I am one of the very first doctors in both my family and my Tribe. This program is amazing, and I wholeheartedly recommend it to all indigenous people and their allies (<https://www.Wānanga.ac.nz/experience/our-students/marla-conwell-ida/>).

Acknowledgment

The following year Dr. Conwell’s twin sister Dr. Mary Dupuis graduated and received the Emeritus Professor Roger Green, ONZM Award for Top Doctoral Thesis.

References

- Barrington, J., 1988. Learning the dignity of labour: secondary education policy for Māori’s. *N. Z. J. Educ. Stud.* 23 (1), 45–58.
- Conwell, M., 2018. Doctor of Indigenous Development and Advancement Graduate. [Online] Available. <https://www.wananga.ac.nz-marla-conwell-ida>.
- Cultee, C., 2020. Xwleml Sche’lang’en: Life Stories Constitute an Indigenous Research Methodology. Thesis submitted in partial fulfillment for the Degree of Doctor of Indigenous Studies Development and Advancement. Te Whare Wānanga o Awanuiārangi, New Zealand.
- Dupuis, M., 2018. A Long Time Ago When the Earth Was Young: A History of the Confederated Tribes of the Chehalis Reservations. Thesis submitted in partial fulfillment for the Degree of Doctor of Indigenous Studies Development and Advancement. Te Whare Wānanga o Awanuiārangi, New Zealand.

- Garner, T., 2000. New PhDs Join an Elite Group. Available. <https://nzherald.co.nz/new-phds-join-an-elite-group>.
- Johnston, P., 1998. He Ao Rereke. Education Policy and Māori Educational Under-achievement: Mechanisms of Power and Difference Unpublished PhD Thesis in Education. University of Auckland, New Zealand.
- Johnston, P., 2010. Graduating Māori PhD: Some Firsts, Some Lessons Some Mistakes. American Educational Research Association, Denver. Colorado 30th April–May 4th.
- Māori Tertiary Education Framework, 2003. Māori Tertiary Education Framework: A Report. Ministry of Education, Wellington, NZ.
- McKinley, E., Grant, B., Middleton, S., Irwin, K., Williams, L.T., 2007. Teaching and learning in the supervision of Māori doctoral students. *Mai Rev.* 2, 6.
- Mead, H., 2003. *Tikanga Māori*. Huia Publishers, Wellington.
- Middleton, S., 2001. Educating Researchers: New Zealand Education PhDs 1948–1998. New Zealand Association for Research in Education Conference. Massey University, Palmerston North.
- Ministry of Education 2001. Available <http://www.educationcounts.govt.nz/statistics/tertiary-education/research>.
- Pohatu, G., 2001. Doctorate Degrees in New Zealand (As at 2001). Unpublished Notes.
- Pro-Vice Chancellor Māori Office, 2007. Te Amorangi: National Māori Academic Excellence Awards 2007. Waikato University. Available. <https://www.waikato.ac.nz/provcmaori/nmaea/archive/2007/eng>.
- Tertiary Education Commission, 2009. Available. http://www.educationcounts.govt.nz/statistics/tertiary_education/participation.
- Tertiary Education Strategy 2007–12: A Framework for Monitoring. Tertiary Sector Performance Analysis and Reporting Strategy and System Performance Ministry of Education Wellington: New Zealand.

Teaching graduate courses; indigenous students, course co-construction and bicultural accountability

Sweeney Windchief, Department of Education, Montana State University, Bozeman, MT, United States

© 2023 Elsevier Ltd. All rights reserved.

Indian education	136
Bicultural accountability	137
Graduate school as a professionalization process	137
Course co-construction	137
The final paper	138
Presentations/posters	138
Challenges	139
Accessibility	139
Conclusion	140
References	140

In following contemporary Indigenous ways of being, I share the following self-location statement to illuminate where my perspective comes from, my communities of accountability, and to be mindful that Shawn Wilson's concept of relationality (2008) is established. Though there are critiques of lengthy introductions (Gone, 2014), I hope that it is received as a presentation of my predispositions and perspective relevant to the passing of indigenous knowledge, in this case within the context of teaching at the graduate level at a predominately White institution (PWI).

Given by my grandfather Hokšína Oyágambisa.¹ I grew up with the Nakóna name Tataga Togáhe. I am a member of the Napéši² (Personal Communication Minerva Allen), Húdešana³ (Personal communication, Sweeney Windchief Sr.) and Wadópana⁴ (Personal communication, Larry Wetsit) clans of the Nakóna. In full disclosure, I was raised primarily off reservation, but feel privileged in that my parents spent the time and money to travel back to the Fort Peck Reservation in Northern Montana, in the summers. The highlights of my upbringing were within the context of family gatherings. As an enrolled Assiniboiné member of the Fort Peck tribes who grew up primarily off reservation, I now serve as an associate professor at a PWI located relatively close to, but ontologically distant from, the Fort Peck Indian Reservation. My intention is to share how this upbringing influences teaching that is inclusive of an Indigenous paradigm, even though it is at a PWI, where often students are not Indigenous but may find themselves engaging Indigenous community contexts, particularly through Montana constitutional law known as Indian Education For All (IEFA).

Indian education

In 1972, the Montana constitutional convention met to revise the state's constitution to include tribal peoples in the state. Though made official in 1999 it was not until 2005 that there was funding committed to the initiative. As a result of this funding there was a commitment, a structure for curriculum development, and dedicated personnel to train educators on the implementation of IEFA (Carjuzaa et al., 2010; Starnes, 2006). Though the majority of work has been at the primary and secondary levels, it is important to recognize that IEFA is not limited to the K-12 arena as stated by Carjuzaa:

OPI has developed several multimedia and on-line IEFA resources. Regional professional development allows many Montana teachers to access training tailored to local context. OPI Indian Education staff present on-site workshops for public schools, colleges, universities and educational organizations statewide (p. 196).

Therefor it is admissible to take the intent of IEFA and implement it into graduate education According to IEFA is focused on seven essential understandings. These include:

¹Assiniboiné name for Joe Day, Translation; "The boy everybody talks about".

²"Stakes themselves out" people through my paternal grandfather and father.

³"Red Bottom People" band of Assiniboiné on my grandmother's side.

⁴Assiniboiné band that settled on the Fort Peck Reservation, called canoe paddlers, Canoe band; Paddlers, band of Assiniboiné living around Wolf Point, MT.

- A recognition of tribal diversity
- A recognition of indigenous identity variability
- Understanding that there is an often missed discourse on pre-“discovery”
- Treaty processes are different for different tribal communities
- Federal policy is enacted, and the result extends beyond a discourse of tribal peoples being seen as wards of the state
- Indigenous curriculum implementation is to be considered using the most local tribes perspective
- The concept of tribal sovereignty is implemented differently among Montana’s tribal communities

In an effort to bring about culturally sustaining teaching (Paris, 2012) at the graduate school level, the purpose of this chapter is to share a teaching methodology that incorporates contemporary Indigenous knowledge while simultaneously incorporating a professionalization process that has the potential to inform what students might expect as they move from graduate programs into the field of academia and or academic administration. As students move from being consumers of knowledge and information in the undergraduate and professional realms, to providers of knowledge and information as graduate students (Neary and Winn, 2009), it is essential to foster development toward scholarship that aligns with local laws which subsequently provides the opportunity to include Indigenous values. (Badwound and Tierney, 1988; Harris and Wasilewski, 2004; Jackson, 2006; Kirkness and Barnhardt, 1991; Verbos et al., 2011), that subsequently confound typical teaching styles at PWIs. In reflecting on the courses that I teach; the students are all graduate students in a graduate classroom setting, and at this stage, the attainment of knowledge and understanding necessitates responsibility and complexity and eventually provides a unique position to influence others.

Bicultural accountability

Typically, graduate students are charged to become the expert in their field. How this expertise is communicated is often not aligned with an Indigenous community value of being humble (Zimmerman, 2005). In addition to a strange immersion into the field they study while balancing cultural integrity, they are in the process of contributing to the theories they ascribe to and taking their teaching responsibilities seriously. Furthermore, many Indigenous graduate students also feel an internal responsibility of contributing to their own communities and the broader indigenous community. Brayboy states “Consistent with findings regarding undergraduate AI/AN students, many graduate students report pursuing higher education out of a desire to give back to or to serve their communities” (Brayboy et al., 2015, p. 265). This is complicated by the fact that aside from Indigenous studies and its many iterations, the scholarship that they engage in their intellectual development is almost certainly centered in a non-Indigenous paradigm. As a university teacher and Indigenous researcher, and it is my charge to facilitate learning in alignment with Indigenous values as opposed to teaching in way that is preformed and predetermined by non-Indigenous principles, that often removes students from the process of innovation. In the majority of the courses I teach, I have to understand that bicultural accountability is paramount to successful classroom innovation. This occurs in three specific ways; (1) by measuring “Proficiency” and “Growth” paradigms which does not assume students have the same kinds of cultural capital, (2) by using consensus, co-construction and co-ownership through an indigenously cognizant model in the development of the course, and (3) implementing a professionalization process that is accountable to the institution as students move into academia and administration.

Graduate school as a professionalization process

Because students arrive to different topic areas with various levels of knowledge, students’ first assignment is typically an “Initial Reflection”; a 1–2 page paper asking them to reflect on their current knowledge directly related to the approved student learning outcomes. Often the responses are based on the student’s experiential knowledge serving to establish microaffirmations (Davison and Rolón-Dow, 2018). For example, students are explicitly asked *not* to Google, use Wikipedia, or review the literature in an effort to assess where each student is located in relation to their knowledge of the course content. The final assignment for the class mirrors this first assignment though at this point, students are encouraged to look at their initial reflection and their assignments as they respond. This teaching approach is conscious of two primary measures, (1) topic proficiency and (2) a knowledge growth paradigm. Topic proficiency is measured by the students’ knowledge of the learning objectives in class assignments while knowledge growth is measured by comparing students’ “initial reflection” with their “final reflection”, showing the amount of growth that each individual student made during their experience in the course, in terms of (1) knowledge, (2) skills, and importantly (3) dispositions (Rosebrook and Larkin, 2003).

Course co-construction

The central pedagogical innovation I use in the facilitation of learning combines using andragogical principles (Knowels, 1996; Merriam, 2001) which are techniques used to teach adults that include self-directed leaning, self-evaluation, and the development of scholarly independence. These principles are paired with a modified version of the Indigenous Leaders Interactive System (ILIS™), “an issues

management process that affirms the value of diverse opinions, clarifies a group vision, and fosters ownership in the collective outcome” AIO. (n.d.) to co-construct students’ collective experiences in the courses I teach, and eventually an agreed upon syllabus for each class.

To begin this process students are presented with the approved student learning outcomes that are stated on the syllabus. The adjusted⁵ ILISTTM prompting question asked of students during the first class is “How will we prove that at the end of the semester that we have met these student learning outcomes?” All students participate in providing responses to this question until as a collective, we have exhausted student contributions. This process provides rigor, by giving students ownership of their course; develops strong working relationships through active listening and considering classmates’ opinions; and provides the contemporary relevance required for a pedagogically sound doctoral level course by including both everchanging learning tools that are available online as well as making Indigenous inquiry a viable option.

As a result of the list of contributions, students are asked to cluster responses that overlap, and consider the contributions in terms of a “Grade-able” product. Students are then asked to connect the responses to the specific learning outcomes of the course, and finally they are asked to vote for the assignments they feel would be the most beneficial to their learning. In order to avoid students’ irritation related to assignments, deadlines, and expectations, the course assignments that we agree upon on are then given point values and due dates, based on guided conversation making room for traditional and family responsibilities that are too often left out. This technique provides the students with the opportunity to actively invest in their own learning (and learning styles) to the course material while feeling empowered to develop how they will learn in a systematic and calculated way. Finally, it allows for students to consider the fact that Indigenous community calendars’ and the academic calendar are not the same, making room for traditional community gatherings allowing students to be “whole” people.

The final paper

In implementing this approach, I have learned that students need some guidance regarding a final product that reflects the professional expectations of academia. The solution is to provide a realistic situation that emulates the postgraduate reality of the dissemination of knowledge. The students are asked to develop academic products that align with institutional and professional structures that are currently in place, to include indigenous research associations such as the American Indigenous Research Association or the He Manawa Whenua/He Au Honua Indigenous Research Conference.

In order to complete their degree, students are at some point required to use analytical and critical thinking to compose a final dissertation, thesis, or some sort of capstone document. The “Final Paper” and related “Poster/Presentation” assignment gives students experience in writing a well-organized paper that is pertinent to their own professional and/or research interests connected to the course. Part of the process is their responsibility to conceptualize and organize the paper during the semester. These papers are completed in successive steps similar to those used in proposing research papers at professional conferences in the field of higher education.

The required format for final papers is reflective of those accepted at regional and national educational researchers’ conferences. Specifically included are those conferences the students are likely to attend; the Northern Rocky Mountain Educational Researchers Association (NRMERA), the Association for the Study of Higher Education (ASHE), and the American Educational Researchers Association (AERA), or the above-mentioned Indigenous research conferences. Papers typically require six components (1) Objectives or purpose; (2) Theoretical or conceptual framework; (3) Methods; (4) Data sources (5) Conclusions or Preliminary results; and (6) The significance of the study.

Connected to the final paper, I require students to prepare a presentation or poster highlighting the elements of their final papers. This opportunity supports professional development in terms of dynamic teaching and presentation skills. Ideally students in the classes I teach will be able to adjust and refine the papers and presentations from my courses and deliver them at regional or national conferences, to their own students, or back home within their own respective Indigenous communities. If indeed students have a paper accepted, we then must develop the presentations that they will be expected to give and find travel funding.

Presentations/posters

The presentation is typically 10–15 min in length to include extra time for questions and feedback in order to prepare students for the academic critique they may experience in the future while simultaneously developing critical thinking skills and collegiality. Students are encouraged to be dynamic in their presentations and are directed to avoid only reading from the text or other materials in their presentation slides and are encouraged to use the provided IT equipment we have available in our classroom which mirrors what might be available at conferences. In addition to assessing the student, the practical significance of this is to test presentations if indeed they are to make a public presentation.

⁵Adjusted primarily because we are not allotted the time to be as thoroughly as practiced by AIO.

Challenges

One of the challenges that I have been confronted with in using this approach is that the course is often experienced as lacking in organization. This has been reflected in student evaluations in the courses I teach. It has been my experience that very rarely are students given the opportunity to co-construct their own course and are often unaccustomed to taking responsibility for their own learning. Nonetheless, by illuminating the realities of the profession, and a gentle reminder in that they will one day be leaders in the field as either professionals/administrators or faculty members, a canned curriculum does little to provide accountability for reflective innovation related to Indigenous community values or ontologies. My goals in the future are to better illuminate those realities while providing opportunities for students to learn about themselves as aspiring scholars and professionals. In order to accomplish these goals, I want to require an out of class presentation or poster session as opposed to “suggesting” or “encouraging” it. It is my hope that this requirement will bring even more investment in their own scholarship. Furthermore, a community accountability component could be added so that students’ home communities are aware of what they are doing in their time away at University. This serves to communicate what the students’ potential community contribution might be upon returning home.

Accessibility

Over the years I have had students who were unable to attend classes for multiple reasons. Students with documented physical disabilities have been unable to attend my courses in person. Others lived out of town and consequently were unable drive to campus because of winter road closures. Moreover, some were in the later stages of pregnancy and had difficulties getting to the classroom. Rather than exclude these students from classes, I have instituted a personal policy that all of my classes will be broadcast via internet video technology using our institutions video technology platform, allowing for participation when otherwise they would not be able to attend. This practice facilitated the “Covid-19 transition”.

In a review of student feedback, the strongest area of my teaching (Impartiality on Grades and exams) can be credited to the professionalization process found within the pedagogy that I implement. Students papers are treated as if they were submitted to an academic journal. I review the papers in terms of commendations and recommendations related to (1) APA seventh Ed. formatting (2) Content (3) Context/Relevance (4) Readability, and (5) Plagiarism. Student presentations are reviewed in a “mock conference presentation” where I act as a panel discussant and the other students are either co-panelists, or audience members.

The second strongest component as articulated by the students’ ratings (Stimulation of interest) can be credited to being conscious of avoiding a “predicament of relevance” for students. This is accomplished by letting them choose the topics they work on over the course of the semester, inviting student work that connects to Indigenous community. The only caveat is that they are germane to the learning outcomes of the course. In short, they are able to complete papers and projects they are interested in and carry community relevance.

This pedagogy opens the door for Indigenous students to do community centered work. As a result, students are aligning their work with IEFA. This is not without some concern. One question that I grapple with is; In a graduate course at a PWI, the students’ work is within an Indigenous paradigm, what is appropriate to share in a classroom, and what is not? Often In-class disclosures are often related to Indigenous cultural and spiritual values to include community specific custom. I am challenged by discourse that connects to students’ community held knowledges that are very close to the students but are conceptualized within specific community contexts and not considered the same way in an academic setting. To be clear, I do not see the knowledge as being wrong, rather, Indigenous knowledge is typically held within Indigenous spaces that are held to an indigenous standard that some consider “not academic”. I am unsure, as of now, how to proceed when it feels like the student is disclosing and sharing community-held information outside the Indigenous community. While their stories, iterations and perspectives are (in my view) absolutely appropriate within the auspices of Indigenous community life, or even the tribal college setting, the fact remains; most of the time depending on the course, at PWIs the majority of students are not Indigenous students. This can be the case even in such courses as Indigenous Methodologies in Research, and Native American Studies courses. To share community held knowledge seems to give non-Indigenous students access to information that is potentially at risk of being mis-used and misappropriated.

Furthermore, I am often challenged by students who wonder why I am such a stickler to APA sixth Ed. formatting, particularly when the knowledge that they are sharing is community held and protected Indigenous knowledge. I see the reason for students’ frustration. Why would an Indigenous student, in an Indigenous methodologies class, taught by a professor who is from an Indigenous community, be required to share in a way that is bound by non-Indigenous rules? This question, which speaks to a well-articulated argument against Western/colonized paradigms is to be taken seriously. My response to these students’ frustrations reflects that their voice is important, needs to be heard, and may empower others. The significance of their knowledge may not impact as many people if it is not written in a way that can be made public.

Reflectively, I have wondered, both in my own scholarship, and in the courses that I teach, the very same sentiments. To take it one step further I have become critical of the commodification of knowledge so apparent in the knowledge economy. In short, the Indigenous scholar, at any level, has the same burdens as their non-Indigenous counterparts/colleagues, but they also have a simultaneously unique source of power and knowledge and additional related responsibilities to protect that knowledge and power.

Conclusion

In conclusion, this chapter seeks to honor Indigenous scholarship by first locating myself for the purpose of establishing perspective. Please consider that this is one way to facilitate learning within an Indigenous paradigm, though performed at a PWI. It is not the only way. In this process I make use of the IEFA initiative in the state of Montana in order to bring Indigenous methods to the classroom while considering bicultural accountability, developing professionalization, self-directed learning, and access for all. Lastly the challenges of over disclosure of Indigenous community knowledge, and the fact that we are operating in colonized systems does not align well with graduate education as we know it. Further research in these areas is welcome, as we move forward providing spaces that help students keep their Indigenous concepts of self, intact. It is my intention to improve student learning in a way that broadens and deepens, professionalism, an investment in one's own learning, and prepare all students for the challenges they face as future scholars and educational leaders with particular emphasis on safeguarding community knowledge while promoting and particularizing what it means to be an Indigenous learner. This can be accomplished through innovative teaching methods which include by concentrating on both proficiency and growth paradigms for each student, facilitating course consensus and co-construction, and finally implementing a process by which students are able to experience what they will encounter in their professional lives, regardless of physical ability or Indigenous self-location and related complexities.

References

- AIO, n.d. Aio projects. Americans for Indian Opportunity:: Indigenous Leaders Interactive System (ILISTM). Retrieved April 8, 2020, from <https://aio.brownrice.com/projects/ilis>
- Badwound, E., Tierney, W.G., 1988. Leadership and American Indian values: the tribal college dilemma. *J. Am. Indian Educ.* 9–15.
- Brayboy, B.M.J., Solyom, J.A., Castagno, A.E., 2015. Indigenous peoples in higher education. *J. Am. Indian Educ.* 54 (1), 154–186.
- Carjuzaa, J., Jetty, M., Munson, M., Veltkamp, T., 2010. Montana's Indian education for all: applying multicultural education theory. *Multicult. Perspect.* 12 (4), 192–198.
- Davison, A., Rolón-Dow, R., 2018. Racial Microaffirmations: Learning From Student Stories of Moments That Matter.
- Gone, J.P., October 2014. Considering indigenous research methodologies: critical reflections by an indigenous knower. In: Symposium Conducted at the Annual Conference of the American Indigenous Research Association Held at the Salish Kootenai College in Pablo MT.
- Harris, L.D., Wasilewski, J., 2004. Indigeneity, an alternative worldview: four R's (relationship, responsibility, reciprocity, redistribution) vs. two P's (power and profit). *Sharing the journey towards conscious evolution. Syst. Res. Behav. Sci.* 21 (5), 489–503.
- Jackson, S., 2006. Compartmentalising culture: the articulation and consideration of Indigenous values in water resource management. *Aust. Geogr.* 37 (1), 19–31.
- Kirkness, V.J., Barnhardt, R., 1991. First nations and higher education: the four R's—respect, relevance, reciprocity, responsibility. *J. Am. Indian Educ.* 1–15.
- Knowles, M., 1996. Andragogy: An emerging technology for adult learning. *Boundaries of adult learning*, 82–98.
- Merriam, S.B., 2001. Andragogy and self-directed learning: Pillars of adult learning theory. *New directions for adult and continuing education*, 89, 3–13.
- Neary, M., Winn, J., 2009. Student as producer: reinventing the undergraduate curriculum. In: Neary, M., Stevenson, H., Bell, L. (Eds.), *The Future of Higher Education: Policy, Pedagogy and the Student Experience*. Continuum, London, pp. 192–210.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Rosebrook, V., Larkin, E., 2003. Introducing standards and guidelines: a rationale for defining the knowledge, skills, and dispositions of intergenerational practice. *J. Intergener. Relat.* 1 (1), 133–144.
- Starnes, B.A., 2006. Montana's Indian education for all: toward an education worthy of American ideals. *Phi Delta Kappan* 88 (3), 184–192.
- Verbos, A.K., Gladstone, J.S., Kennedy, D.M., 2011. Native American values and management education: envisioning an inclusive virtuous circle. *J. Manag. Educ.* 35 (1), 10–26.
- Zimmerman, L.J., 2005. *First, Be Humble: Working With Indigenous Peoples and Other Descendant Communities*.

Developing transforming research—The University of Alaska Fairbanks Center for Cross-Cultural Studies as case study

Sean Asikluk Topkok, Michael Koskey, and Theresa Arevgaq John, Center for Cross-Cultural Studies, Indigenous Studies Graduate Programs, Fairbanks, Alaska

© 2023 Elsevier Ltd. All rights reserved.

Student and faculty videos	145
CXCS faculty	146
Mission statement	146
Values	146
Respect	147
Relevance	147
Reciprocity	147
Responsibility	147
Shared Alaska Native values	147
Overall work ethics	148
How do these philosophies support an Indigenous/Native model of well-being?	149
References	149

The Center for Cross-Cultural Studies (CXCS) at the University of Alaska Fairbanks (UAF) was created in 1977 to address research, development, and instructional issues associated with educational policies, programs, and practices in culturally diverse contexts, with an emphasis on Indigenous Peoples worldwide, rural, and distance education.

CXCS was originally established in 1971 as the Center for Northern Educational Research (CNER) by action of the UA Board of Regents and with initial funding from the Ford Foundation. By 1977 the University had assumed full funding and CNER was merged with the Cross-Cultural Education Development Program to form the Center for Cross-Cultural Studies. CXCS was housed in the UAF College of Liberal Arts until Spring 2021. As of May 2021, CXCS is housed in the UAF College of Rural and Community Development (CRCD).

The research agenda for the center is established in cooperation with Native organizations, school districts, and state/federal agencies, with a focus on applied research that will benefit the people of Alaska. CXCS is structured to provide technical support and information to school districts, social service agencies, Native corporations, tribal governments, community colleges, and state and federal agencies in rural Alaska. This provides direction for the improvement of educational and professional development opportunities for rural Alaskans and a forum for the examination of Indigenous education and community development issues. The center also sponsors various state, national, and international seminars, conferences, and exchanges to bring people together around issues of concern to Alaska, the circumpolar North, and Indigenous people throughout the world. The Center houses the Alaska Native Knowledge Network (ANKN) and offers an M.A. and Ph.D. in Indigenous Studies with an emphasis on Indigenous knowledge and Indigenous ways of knowing. All faculty and students interested in issues associated with Indigenous knowledge systems, cultural diversity, and education are encouraged to participate in and contribute to the efforts of the center.

The Alaska Native Knowledge Network was established in 1995 through the Alaska Rural Systemic Initiative (AKRSI) (Hill et al., 2006). AKRSI was founded 1995–2005 to implement Native ways of learning and teaching in rural Alaska through various initiatives with about twenty school districts statewide. “The underlying purpose of these efforts has been to implement a set of research-based initiatives to systematically document the Indigenous knowledge systems of Alaska Native people and to develop pedagogical practices and school curricula that appropriately incorporate Indigenous knowledge and ways of knowing into the formal education system” (Barnhardt and Kawagley, 2005).

ANKN was established to become “a culturally based curriculum resources clearinghouse established to disseminate the information and materials that are being developed and accumulated as the AKRSI initiatives are implemented throughout rural Alaska” (Barnhardt and Kawagley, 1999). Native ways of knowing resources have been implemented in a variety of ways throughout Alaska. ANKN was developed to become a repository and distribution of various ways of teaching and learning:

In 1995 the University of Alaska Fairbanks (UAF) in collaboration with the Alaska Federation of Natives established the Alaska Native Knowledge Network (ANKN) to share and promote the exchange of cultural resources and knowledge among Indigenous communities throughout Alaska and beyond.

Topkok (2014, p. 143).

ANKN is a website and a clearinghouse that constitutes a part of the Center for Cross-Cultural Studies. ANKN serves as a resource for compiling and exchanging information related to Alaska Native knowledge systems and ways of knowing. It has been established to assist Native people, government agencies, educators, and the general public in gaining access to the knowledge base that Alaska Natives have acquired through cumulative experience over millennia.

While CXCS was founded in 1971 to serve Alaska Native research, development, and instructional issues; its two graduate programs include Indigenous collaboration worldwide. The cross-cultural studies programs prepare graduates capable of conducting basic and applied research on social, political, educational, economic, and cultural issues of concern to people and Indigenous communities.

The Master of Arts degree in Cross-Cultural Studies with an emphasis on Indigenous knowledge systems was approved by the UAF Board of Regents on March 9, 2001. The degree's title was changed in 2019 to a Master of Arts degree in Indigenous Studies to better reflect the nature of the curriculum. The program is designed to provide graduate students from various fields of interest an opportunity to pursue in-depth study focusing on the role and contributions of Indigenous knowledge in the contemporary world.

The M.A. degree program provides a means to expand our knowledge base in areas that have received only limited attention in the past, as well as to document and pass that knowledge on to future generations in a culturally sensitive way. The intent of the program is to incorporate and contribute to newly emerging bodies of scholarship that have much to offer in addressing critical needs of the state.

Graduates of the program are expected to bring greater depth and breadth of cultural understanding to many of the complex social issues and fields of endeavor that shape Alaska today, especially those involving cross-cultural considerations and utilizing Indigenous knowledge systems (e.g., education, ecological studies, natural resources, health care, community development, social services, justice, Native studies). Students are required to demonstrate their ability to work effectively with Indigenous people in their studies and to complete a final cultural documentation project in collaboration with knowledgeable Elders. The master's degree program in Cross-Cultural Studies/Indigenous Studies is housed in CXCS, and as of May 2021, thirty people had earned their master's degree with eight self-identifying as Alaska Native and four self-identifying as Indigenous outside Alaska.

In 2009, the doctoral degree program in Indigenous Studies was established and continues to be housed in CXCS. The Ph.D. in Indigenous Studies draws upon long-standing academic and research capabilities at UAF to offer an integrated, cross-disciplinary course of advanced graduate study. Native peoples in Alaska have usually been the subjects of research rather than the ones responsible for conducting it. However, the role of Alaska Natives in research is changing due to a concerted effort on the part of the University of Alaska and Native people themselves to develop new programs aimed at recruiting and preparing Native scholars in all academic fields who can take on leadership roles and bring an Indigenous perspective to the policy arenas at the local, state, national, and international levels.

The interdisciplinary Ph.D. program in Indigenous Studies is sponsored by the UAF Graduate School in collaboration with the Department of Alaska Native Studies and Rural Development (DANSRD), the Alaska Native Language Center (ANLC), the College of Liberal Arts (CLA), the School of Education (SOE), and the College of Rural and Community Development (CRCDD). The program consists of a common core curriculum that all students complete, coupled with six thematic specialty areas from which students choose a concentration:

- Indigenous Studies and Research
- Indigenous Knowledge Systems
- Indigenous Education and Pedagogy
- Indigenous Languages
- Indigenous Leadership
- Indigenous Sustainability

The Indigenous Studies and Research concentration provides the programmatic infrastructure for advanced, in-depth, interdisciplinary graduate studies and research in academic fields related to the role of Indigenous knowledge and ways of knowing in the contemporary world. This will provide students focusing on this concentration area to provide a venue to sponsor state, national and international seminars, conferences, exchanges, and comparative research programs that bring people together around issues of concern to Alaska, the circumpolar North, and Indigenous people throughout the world. Below is an abstract from one of our graduates from the Indigenous Studies Ph.D. program with a concentration in Indigenous Studies and Research. Dr. Norma Shorty's title of her dissertation is, "Inland Tlingit of Teslin, Yukon: Gaanax.ádi and Kookhittaán Clan Origin Stories for the Immediate and Clan Family of Emma Joanne Shorty (Nee Sidney)."

The purpose of my research is to learn the story of Mother's clan, and to document the processes of gathering knowledge about the clan connections between the Gaanax.ádi and Kookhittaán from Teslin, Yukon, Canada. The objective of this thesis is to document the stories and the story gathering processes of published and private holdings on my Mother's clan stories. The study includes published literature and non-Indigenous historians and oral history reviews, especially on those who have knowledge about the Kookhittaán and Gaanax.ádi clans and have connections to the Inland Tlingit from Teslin, Yukon. This Indigenous-led research focuses on my mother and her clan stories. I am an insider and an outsider to my culture. From an insider perspective I am privileged to hear, to learn, and to retell Mother's maternal clan stories. As a result of this research, Tlingit ways of documenting history are discovered and Tlingit research (literacy) frameworks are revealed. I learned that the Kookhittaán and Gaanax.ádi clans are one. Our oral history is validated by face paint designs, petroglyphs and clan shirt designs. In their published work some non-Indigenous ethnographers made changes to words and designs which distorted the Indigenous record. This dissertation compares all possible information sources showing the heavier

weight of evidence is provided by available Indigenous sources. Colonization has greatly impacted the perpetuation of Indigenous knowledge systems by referring to Indigenous knowledge as “traditional” because the term tradition conjures up images of living in the past.

Shorty (2015).

The Indigenous Knowledge Systems concentration contributes to and taps into newly emerging bodies of academic scholarship that address the role of Indigenous knowledge systems in fields such as ecological studies, natural resources management, health care, education, language revitalization, community development, social services, justice, and Native studies. Students focusing on this concentration area will compare and analyze the epistemological properties, worldviews, and modes of transmission associated with various Indigenous knowledge systems. Below is an abstract from the first graduate from the Indigenous Studies Ph.D. program with a concentration in Indigenous Knowledge Systems. Dr. Theresa Arevgaq John’s title of her dissertation is, “Yuraryarput Kangiit-llu: Our Ways of Dance and Their Meanings.”

The first purpose of this study is to describe the categories of dance. The second purpose is to describe how Yup’ik music and dance has played a functional role in organizing and maintaining various societal infrastructures (kinship, social, political, subsistence/economic, and spiritual) within the Yup’ik culture. This study seeks to further understand this role and how it has evolved over time.

The study utilizes an ethnographic methodology that includes historical and contemporary perspectives to describe Yup’ik music and dance categories and to explain how dance serves to organize various aspects of Yup’ik culture and societal infrastructure. Data includes interviews from Yup’ik elders and adults, fieldnotes, research journal entries, digital recordings, photographs, and observations of Yup’ik immersion school performers and rural community cultural events such as the Cama-i Festival.

The study suggests that Yup’ik dance and categories are important elements of the multiple cyclic rituals. It adds to the present literature revealing that there are twenty different dance types and categories, and many of the rituals are lost except for the ciuqitet (common dances), nangereciyaraq (the first dance), and iluriuracraq (teasing dance) dances.

The study also suggests that dancing is an essential part of the Yup’ik social infrastructure and that dancing is integral to the social system. This is demonstrated through six themes: Kinship, Physical/Mental Health, Form of Prayer, Spiritual Enlightenment, Leadership, and Teasing.

I also argue that there is connectedness in dance, music, and stories that are part of our yuuyaraq (epistemic worldview). Yuuyaraq is defined as a way of being a human or an absolute unified social web. This web is represented in our social infrastructures of kinship, health/physical and mental, form of prayer/rituals, spiritual enlightenment, leadership, and teasing. There is a relationship in storytelling genres in dance and oral stories that represent people’s historical and contemporary accounts, describing their social, cultural, and subsistence lifestyle. Interview participant data suggest these connections still exist in our society today.

John (2010).

The Indigenous Education and Pedagogy concentration provides an opportunity for students to engage in the comparative study of issues associated with the education of Indigenous peoples and communities with an emphasis on the role of language and culture through a review of educational practices, pedagogies, and theories that have emerged from a variety of sources. Students focusing on this concentration will demonstrate a greater understanding about the role of Indigenous culture and language in cognitive, social, psychological, spiritual, and physical well-being. Below is an abstract from a graduate in the Indigenous Studies program with a concentration in Indigenous Education and Pedagogy. Dr. Charleen Fisher’s title of her dissertation is, “Diideets’ii In Our Pathway (In Our Future): Gwich’in Educational Philosophy and transformative Praxis in K-12 Education.”

Gwich’in pedagogy is largely undocumented in Western academia. Gwich’in epistemology includes holistic perspectives on all Western content areas and crosses the usual segmented knowledge genres. Intergenerational transmission of Gwich’in knowledge occurs in many places including the natural environment, with long-standing cultural ties to place. Gwich’in pedagogy is relational, place-based, holistic, cooperative, purposeful, and subjective. Gwich’in gaagwidandaii, or communal knowledge, predates the inception of many world societies. Gwich’in concepts presented in this paper will include the introduction of a framework called Kheegwadadhaak’a, translated to mean, “We just keep the fire going.” This framework is a visualization. Important concepts of Gwich’in pedagogy include traditional ideas of assessments or standards using the phrases nil’ee t’ah’in and ch’aadaii, both meaning that someone has a natural talent or is adept at something, for girls and boys, respectively. Learning, or gik’yanjii in Gwich’in, also means “to find out, notice or sense.” This comprehension includes a deep, contextual understanding of traditional Gwich’in knowledge. The three types of Gwich’in knowledge are gaagwidandaii, gihk’agwagwaanjik, and gaatr’oahitan. These translate as “collectively known, individually learned, and taught knowledge,” respectively. Gwich’in have a complex and relational pedagogy. This pedagogy attempts to achieve contextuality, or duulee ginlii, which translates as “proficiency, agility, ability to do almost anything, being extremely good at anything they do, or overall ‘sharpness’ in life.” This process is importantly both a communal and personal journey.

Fisher (2018).

The Indigenous Language concentration strives to reinforce an Indigenous identity that is dependent on the language and culture, prepares the student for success in the world, and leads to acceptance at home. The concentration is based on the philosophy that a strong command of an Indigenous language leads to an understanding of Indigenous ways of life, the world around us, and our place in it. Below is an abstract from a graduate of the Indigenous Studies Ph.D. program with a concentration in Indigenous Language. Dr. Alisha Drabek's title of her dissertation is, "Liitukut Sugpiat'stun (We Are Learning How To Be Real People): Exploring Kodiak Alutiiq Literature Through Core Values."

The decline of Kodiak Alutiiq oral tradition practices and limited awareness or understanding of archived stories has kept them from being integrated into school curriculum. This study catalogs an anthology of archived Alutiiq literature documented since 1804 and provides an historical and values-based analysis of Alutiiq literature, focused on the educational significance of stories as tools for individual and community wellbeing. The study offers an exploration of values, worldview and knowledge embedded in Alutiiq stories. It also provides a history of colonial impacts on Alutiiq education and an in-depth study of the early colonial observers and ethnographers who collected Alutiiq oral literature, clarifying the context in which the stories have been retold or framed. Collections of traditional Indigenous literatures are valuable on many levels. This collection is of historical and personal significance for local Kodiak Alutiiq tribal members' identity as it makes these resources more accessible for community members and educators, and therefore accessible to younger and future generations. The conclusion also provides recommendations for next steps for developing curriculum and revitalizing Alutiiq oral traditions. The book is intended to contribute to an understanding of the evolution of cultural traditions in Alaska, and to serve as a model for similar cultural reclamation and education efforts.

Drabek (2012).

The Indigenous Leadership concentration prepares graduates who are capable of conducting basic and applied research on social, political, educational, economic and cultural issues of concern to people and communities in Indigenous communities. Students focusing on this concentration area will expand the pool of knowledgeable and highly skilled Alaskans who can assume leadership and technical positions with public and private sector organizations, including universities, school districts, social service agencies, Native corporations, tribal governments, and state and federal agencies in Alaska and beyond. Below is an abstract from a graduate from the Indigenous Studies Ph.D. program with a concentration in Indigenous Leadership. Dr. Pearl Kiyawn Brower's title of her dissertation is, "Tumitchiat: Iñuqqaat Aullarrisiatun Ilisaġviit—A New Pathway: Indigenous Leadership in Higher Education."

After centuries of colonization and assimilation policies in education, Indigenous people across our world are making commitments to nurture the next generation of Indigenous leaders. One area of focus is on Indigenous leadership through various forms of Higher Education. This dissertation supports the need for Indigenous leadership programs that have a foundation in Indigenous ways of knowing and learning and suggests a definition of what Indigenous leadership embodies. Specific Indigenous leadership programs are described.

Within the United States, tribal colleges are unique entities, but they share the same goals. These institutions create opportunities for hope and sustained Indigenous self-determination through their students, who are the next generation of Indigenous leaders. A broad review of tribal colleges is presented here. Then a particular tribal college, Ilisaġvik College, located in Barrow, Alaska, is discussed in more detail. The research also illuminates circumstances at University of Hawai'i (UH), a settler-colonial institution, that has recently decided to become a Native Hawaiian (NH) serving institution. How is this new responsibility viewed by UH's Indigenous leaders? I conclude with an overview and syllabus for an Arctic Indigenous leadership program to be implemented at Ilisaġvik College. The intent is to provide a path that others may use to create their own programs to meet their Indigenous communities' needs.

Brower (2016).

The Indigenous Sustainability concentration is concerned with the social, cultural, and physical well-being of Indigenous communities at the present and into the future, in part by understanding the changes of the past. Students focusing on this concentration area will explore interdisciplinary topics from a systems perspective that holistically examine the interactions between food, health, education, economic, and political subsystems, within a historical context. Knowledge sources for sustainability are not limited to a single cultural or ideological perspective, and new ideas and experiences concerning issues of sustainability are taken into consideration. The concept of sustainability is of growing concern to many academic disciplines; new course offerings are regularly being developed across the University of Alaska system, and students interested in this concentration area will have the opportunity to work closely with their graduate committee to find and select the courses most relevant to their individual studies. Below is an abstract from a graduate of the Indigenous Studies Ph.D. program with a concentration in Indigenous Sustainability. Dr. Heather Sauyaq Gordon's title of her dissertation is, "Self-Determination, Sustainability, and Wellbeing in the Alaska Native Community of Ninilchik."

Alaska Natives are a diverse group of people with different language groups and over 200 tribes. We have a history of colonization and are still a colonized people, but through all this, we strive for wellness for our people. This paper begins with an explanation of historical trauma, development, and the lack of fate control Alaska Native people experience. The literature review explains how colonization can negatively impact the colonized and details international, federal, and Alaska state law and court cases having to do with Indigenous sovereignty and self-determination. In this project the researcher works with the Ninilchik Village Tribe of Ninilchik, Alaska, to explore how community members utilize self-determination, either

individually and/or as a group, to achieve individual, community, and tribal sustainability and wellbeing. This project uses the method of ethnographic futures research to conduct scenarios about the future. The researcher conducted 30 interviews about three possible futures: the optimistic, pessimistic, and most likely, and followed the interviews with four focus groups to discuss the interview results. The results were coded through grounded theory in NVivo analysis software and compared with: (a) the Capabilities Approach, (b) Self-Determination Theory, (c) social science development theories of Dependency and World Systems, and (d) the Elements of Development Model. The Capabilities Approach and Self-Development Theory explain the links between self-determination and wellbeing. Dependency and World Systems Theories explain the importance of local self-determination for development. Finally, the Elements of Development Model provides an outline for different types of self-determining actions. The project analyzes Arctic wellbeing indicators and developed indicators of sustainability and wellbeing. The project results demonstrate what community members think that individuals, the community, and the tribe can do to improve sustainability and wellbeing in Ninilchik, and how to achieve those goals through self-determining actions. The dissemination document serves as the start to a 20-year strategic plan. This type of research demonstrates how tribes can address the results of historical trauma and take control of their fate through self-determination. The next steps in research would be asset mapping and capacity-building projects to work with the data and benefit the community.

Gordon (2019).

Ph.D. students participate in research activities across a variety of academic disciplines and applied fields at UAF. They are encouraged to engage in comparative studies with other Indigenous peoples around the world and to focus their dissertation research on issues of relevance to Alaska and the Arctic. Using the Interdisciplinary Ph.D. model of academic assignment, students' home-base is in the school or college of their major advisor or chair of their committee.

In collaboration with the graduate committee, each student develops a program of coursework and research that produces a unique intellectual contribution to the applied fields associated with Indigenous Studies. Students elect to focus on one of the six thematic specialty areas, or they may choose in collaboration with their graduate committee to draw on multiple themes to develop their own areas of study and dissertation research.

The Ph.D. Program in Indigenous Studies integrates the tools and approaches of the natural and social sciences in a cross-cultural and interdisciplinary framework for analysis to better understand the emerging dynamic between Indigenous knowledge systems, Western science, and higher education. We focus on the interface between Indigenous knowledge and science on an international scale, with opportunities for collaboration among Indigenous peoples statewide, nationally, and internationally. It also draws and builds upon past and current initiatives that seek to utilize Indigenous knowledge to strengthen the curriculum and pedagogical practices in education.

With numerous research initiatives currently in various stages of development and implementation around the world, there is an unprecedented window of opportunity to open new channels of communication between scientists, policymakers, and Indigenous communities, particularly as they relate to those research activities that are of the most consequence to Indigenous peoples (e.g., effects of climate change, environmental degradation, contaminants and subsistence resources, health and nutrition, bio/cultural diversity, Arctic observation networks, natural resource management, economic development, resilience and adaptation, community viability, cultural sustainability, language and education).

To the extent that there are potentially competing bodies of knowledge (Indigenous and Western) that have bearing on a comprehensive understanding of particular research initiatives associated with Indigenous-related themes, we offer opportunities for Indigenous Studies Ph.D. students to be embedded with ongoing research initiatives to contribute to and learn from the research process. In addition to conducting research on the inner dynamics of Indigenous knowledge systems, the Ph.D. students also examine the interplay between Indigenous and Western knowledge systems, particularly as it relates to scientific processes of knowledge construction and utilization. As of May 2021, there were 24 people who earned a doctoral degree in Indigenous Studies with 13 self-identifying as Alaska Native and 5 self-identifying as Indigenous outside of Alaska.

Alaska has at least 21 distinct ethnolinguistic Indigenous groups, and hundreds of village-based Indigenous communities. Both the master's and doctoral Indigenous Studies programs include Indigenous students statewide, nationally, and internationally. While a majority of the Indigenous Studies students are from Alaska, there are/have been students from the continental United States, Hawai'i, Guam, Canada, Mexico, (South) Korea, and other areas; and students from Japan, Ghana, Zimbabwe, Macedonia, Pakistan, Azerbaijan, Greece, Germany, Norway, the Netherlands, Denmark, Finland, and Switzerland have been enrolled in our classes.

Student and faculty videos

Dr. Charlene Aqpiq Apok <https://drive.google.com/file/d/1jzXihZYZRnRQbwE3EOqSw0ESR7dpMxZp/view?usp=sharing>.

Margie Mete https://drive.google.com/file/d/1u10lVpcW_2K9O5dfDhwhf9ddpZyfkllsC/view?usp=sharing.

Eldri Westmoreland https://drive.google.com/file/d/11S9ayKAiZ3xN_KDfDp2TbUKP4-LusF4N/view?usp=sharing.

Denise Newman <https://drive.google.com/file/d/1J-Qu7gYk7cbQW3SNd-h7v7YGWQP8nwt-/view?usp=sharing>.

Debra O'Gara https://drive.google.com/file/d/1adlw_ChPn4YfjNXHsPRIEFIOLDzdKfj1/view?usp=sharing.

Dr. Theresa Arevgaq John <https://drive.google.com/file/d/16mvsIXEKHYuPEEJYRetw2RBFNeiWv3ss/view?usp=sharing>.

Dr. Mike Koskey https://drive.google.com/file/d/1wBJmWgCPfHwQnM0RMzuPNvUhpI_5qDUx/view?usp=sharing.

Dr. Sean Asikluk Topkok <https://drive.google.com/file/d/1AxjtAkVQeUVConS2bsRY5KMB81PKwEQi/view?usp=sharing>.

CXCS faculty

Dr. Sean Asikluk Topkok is Iñupiaq, Sámi, Kven, Irish, and Norwegian. Asikluk is a tenured Associate Professor and Chair for the Center for Cross-Cultural Studies. He earned his doctorate in Indigenous Studies with a concentration in Indigenous Knowledge Systems. His research interests include multicultural and Indigenous education, decolonization and Indigenist methods and methodologies, working with communities to help them document their cultural heritages, and community well-being. His most recent research includes working with Nome residents (including Iñupiaq, St. Lawrence Island Yupik, and Central Yup'ik) about Arctic cruise tourism, implementing geoscience utilizing cultural knowledge systems in a Northwest Arctic Borough School District (mostly Iñupiaq residents), and interviewing Tlingit and Haida residents in Juneau utilizing Southeast Traditional Values dealing with the COVID pandemic. Asikluk is one of twenty-nine Iñupiaq scholars who earned a doctorate throughout history. He currently serves on the WINHEC Executive Board, World Indigenous Nations University (WINU) Pro-Vice Chancellor, University of the Arctic Avatitsinni (On Our Homelands—formerly known as the Indigenous Issues Committee), and other Indigenous committees. Asikluk started as a staff person for Alaska Native Knowledge Network and the Center for Cross-Cultural Studies in 1997. He created and taught Cultural Atlases as a Pedagogical Strategy for CXCS in 2006. Asikluk's institutional knowledge and being mentored by Drs. Ray Barnhardt and Angayuq Oscar Kawagley provide students through his and others' experiences.

Dr. Michael Koskey brings to the Center of Cross-Cultural Studies a diversity of life experiences and education that have focused on intercultural relations and understanding knowledge from local worldview perspectives. Born in Germany, raised in the American South, a Marine veteran of the Gulf War, and with research experience in Alaska, Siberia, and Belize, a lifelong interest in learning from local, especially Indigenous people, forms the basis of Koskey's cross-cultural scholarship. With an educational and professional background in anthropology, history, political economy, education, archeology, and international relations; Koskey brings a broad yet integrated perspective to his classes and research that is always rooted in local experiences and community interests. With this background, Koskey works to bridge cultures, by the exchange of knowledge, through mutual respect and understanding of one another's worldviews.

Dr. Theresa Arevgaq John is Yup'ik from Toksook Bay. Dr. John is a tenured Associate Professor in the Center for Cross-Cultural Studies at the University of Alaska Fairbanks. She received her Ph.D. in May 2010 from the UAF CXCS Indigenous Studies Program entitled Yuraryarput Kangiit-Illu: Our Ways of Dance and Their Meanings. Theresa is the first graduate of the Indigenous Studies Ph.D. program. President Obama selected Theresa to serve on the National Advisory Council on Indian Education in 2011. She co-authored a book entitled Yupiit Yuraryarit: Yup'ik Ways of Dancing that received a prestigious 2011 book of the year award from the Alaska State Library Association. Theresa has published academic articles in various academic agencies as well as local newspapers. Her professional interest areas include Alaska Native ceremonies and studies, Indigenous epistemology, ontology, ecology, cosmology, and worldview.

Mission statement

Our mission at the Center for Cross-Cultural Studies is the improvement of educational and professional development opportunities for Indigenous Peoples. We also serve as a forum for the examination of cross-cultural and Indigenous education and community development issues across the state, as well as the globe.

Through the interests of our students, our goal is to provide technical support and information to school districts, Native corporations, tribal governments, community organizations, and state and federal agencies serving Indigenous communities. Our Indigenous Ph.D. program is the highest Native Studies degree granting program in the state and combined with our MA program, we work hard to increase the recognition of Indigenous scholarship across a range of academic disciplines to best serve our students.

Values

As a reminder, the Center for Cross-Cultural Studies houses the Alaska Native Knowledge Network and the graduate programs in Indigenous Studies. We acknowledge that each 229 federally recognized Alaska tribes and Indigenous-recognized tribes have their own unique cultural values. [Roderick \(2008\)](#) states:

What these cultural groups share in common, however, are deeply-ingrained values, such as honoring the land and waters upon which life depends, having respect and reverence for fish and wildlife, valuing community over individuality, sharing with others, and respecting and learning survival skills and wisdom from Elders. Alaska Native cultural worldviews are holistic. Native cultures accept that everything in creation is connected, complex, dynamic, and in a constant state of flux. Alaska Native peoples have a deep and sophisticated qualitative understanding of the environment in which they live. This understanding comes from stories passed down for generations; it also comes from life experiences, learning from mentors beginning at a young age, observations of others in the community, and the guidance of Elders (p. 2).

CXCS serves Indigenous peoples internationally. In the *Alaska Standards for Culturally Responsive School* (also known as the *Alaska Cultural Standards*), under the Cultural Standards for Schools A(3), a school, "includes explicit statements regarding the cultural

values that are fostered in the community and integrates those values in all aspects of the school program and operation” ([Assembly of Alaska Native Educators \[AANE\], 1998](#), p. 17). As an academic and research center at UAF, we acknowledge CXCS as a school. Hence, CXCS’ Core Values include shared international Indigenous values of Respect, Relevance, Reciprocity, and Responsibility ([Kirkness and Barnhardt, 2001](#)).

Respect

CXCS recognizes Indigenous Peoples are inherently sovereign. Each student, each community, and each research project have its own cultural protocols and knowledge system unique to their respective areas. CXCS represents a personal, welcoming, and belonging environment, in which students, communities, and research bring cultural knowledge, traditions, and Indigenous values that are respected.

Relevance

[Kirkness and Barnhardt \(2001\)](#) state, “If universities are to respect the cultural integrity of First Nations students and communities, they must adopt a posture that goes beyond the usual generation and conveyance of literate knowledge, to include the institutional legitimization of Indigenous knowledge and skills.” One of CXCS’ foundational courses is Indigenous Knowledge Systems, created by the late Dr. Angayuqaq Oscar Kawagley. The course description states this course is, “a comparative survey and analysis of the epistemological properties, worldviews and modes of transmission associated with various Indigenous knowledge systems.” Each student and community have a unique worldview which should not only be accommodating but also structural throughout one’s research and education.

Reciprocity

[Kirkness and Barnhardt \(2001\)](#) share, “the emphasis is on making teaching and learning two-way processes, in which the give-and-take between faculty and students opens up new levels of understanding for everyone ... Faculty members and students in such a reciprocal relationship are in a position to create a new kind of education, to formulate new paradigms or explanatory frameworks that help us establish a greater equilibrium and congruence.” In the Alaska Cultural Standards, institutions are encouraged to nurture students’ experiences in schooling and to apply what they have gained in the real-world context.

Responsibility

[Kirkness and Barnhardt \(2001\)](#) state, “Gaining access to the university means more than gaining an education—it also means gaining access to power, authority, and an opportunity to exercise control over the affairs of everyday life, affairs that are usually taken for granted by most non-Native people.” Students who are enrolled in the Indigenous Studies programs are encouraged by their Elders to continue schooling to benefit their respective communities. Many students have expressed their motivation to earn a post-baccalaureate degree to benefit their respective communities.

Shared Alaska Native values

Alaska Native peoples have recognized shared cultural values. Though CXCS acknowledges and respects the international Indigenous cultural values of Respect, Relevance, Reciprocity, and Responsibility; we include these shared Alaska Native cultural and core values as an Indigenous program housed in an Alaska academic institution:

Show Respect to Others—Each Person Has a Special Gift

Share What You Have—Giving Makes You Richer

Know Who You Are—You Are a Reflection on Your Family

Accept What Life Brings—You Cannot Control Many Things

Have Patience—Some Things Cannot Be Rushed

Live Carefully—What You Do Will Come Back to You

Take Care of Others—You Cannot Live without Them

Honor Your Elders—They Show You the Way in Life

Pray for Guidance—Many Things Are Not Known

See Connections—All Things Are Related

ANKN (n.d.).

Each Alaska Native group has their own unique set of cultural values identified by Elders in the communities, and in some groups the cultural values may differ from village to village. However, all cultural values have been passed down from generation to generation for thousands of years. Some groups have created posters identifying their cultural values. Elders, community members, and other cultural bearers worked together to identify and create their own cultural values; they did not come from one individual. To include each one would be extensive. Hence, the shared cultural values listed above continue to ground our students, staff, faculty, and communities as we work together.

Overall work ethics

CXCS faculty, staff, and students work closely with communities and their partners. The First Alaskans Institute (FAI) is one of the many organizations whose focus includes Indigenous-led initiatives which CXCS emulates in research and instruction. FAI formalized a set of agreements (<https://firstalaskans.org/wp-content/uploads/2016/10/FAI-AGREEMENTS-POSTER.pdf>) which CXCS have adopted for all our meetings:

In Every Chair, a Leader

Speak to be Understood; Listen to Understand

Be Present; Be Engaged

Value Our Time Together

Safe Space for Meaningful Conversation

Challenges → Solutions

Takest Thou Hats Off

Our Value of Humor Helps Us

We are Responsible for Our Experience

Take Care of Yourself; Take Care of Each Other

Research is conducted in alignment with Indigenous guidelines and protocols including the “Guidelines for Respecting Cultural Knowledge,” “Coolangatta Statement on Indigenous Rights in Education,” and the “United Nations Declaration on the Rights of Indigenous Peoples” (ANKN, 2000; UN, 2007; WIPCE, 1993). In the “Guidelines for Respecting Cultural Knowledge,” there is a list of actions for cultural responsiveness for researchers. They include, “(a) Effectively identify and utilize the expertise in participating communities to enhance the quality of data gathering as well as the data itself and use caution in applying external frames of reference in its analysis and interpretation (p. 15).” Many students and faculty are actively involved with cultural activities. “(b) Ensure controlled access for sensitive cultural information that has not been explicitly authorized for general distribution, as determined by members of the local community (p. 15).” Students and faculty approach tribal councils to involve them with research at the very beginning. “(c) Submit research plans as well as results for review by a locally-knowledgeable group and abide by its recommendations to the maximum extent possible (p. 15).” Students and faculty share drafts and results to review before any publishing. “(d) Provide full disclosure of funding sources, sponsors, institutional affiliations and reviewers (p. 15).” Communities are well-informed of all sponsorships. “(e) Include explicit recognition of all research contributors in the final report (p. 16).” Participants are recognized for their contributions in dissertations and other publications, and in some cases, they are co-authors. “(f) Abide by the research principles and guidelines established by the Alaska Federation of Natives and other state, national and international organizations representing Indigenous peoples (p. 16).” These principles and guidelines also include the “Coolangatta Statement on Indigenous Rights in Education” and the “United Nations Declaration on the Rights of Indigenous Peoples.” Research by students and faculty are written and developed by Indigenous people, provide guidelines for conducting research ethically and respectfully; and to do so *with, by, and for* Indigenous people.

How do these philosophies support an Indigenous/Native model of well-being?

The CXCS Well-Being Model applies to students, faculty, and communities. While the faculty are involved with various research, some of the themes might overlap with each other. Additionally, students have a choice of at least one area of concentration, but they all have an option to include other areas of concentration.

The Center for Cross-Cultural Studies is a research center which offers master’s and doctorate academic degrees. Faculty are involved with research with communities. Students conduct research for their graduate degrees. As a research center, we encourage community members actively involved with Indigenous research. For all three (students, faculty, and communities), well-being is essential to begin with individuals and communities. Each individual and community define well-being themselves utilizing their own cultural values and epistemology. For ways of being, they are guided by their cultural values which remain constant throughout time immemorial. For ways of thinking, they are encouraged to utilize their epistemology and Indigenous worldviews. Ways of understanding their cultural values and epistemologies are deeply embedded in their ancestral and living heritage(s).

CXCS is an internationally Indigenous Studies program. There is no one Indigenous philosophy which our program draws for its foundation. We recognize students, faculty, and communities bring with them their own rich culture and knowledge systems. From these various philosophies, CXCS continues to grow and provide an Indigenous platform in a Western institution.

References

- Alaska Native Knowledge Network, (n.d.). Alaska Native Values for Curriculum. <http://ankn.uaf.edu/ANCR/Values/>.
- Alaska Native Knowledge Network, 2000. Guidelines for Respecting Cultural Knowledge. University of Alaska Fairbanks, Fairbanks, AK.
- Assembly of Alaska Native Educators, 1998. Alaska Standards for Culturally Responsive School. Alaska Native Knowledge Network, Fairbanks, Alaska.
- Barnhardt, R., Kawagley, A.O., 1999. Culture, Chaos and Complexity: Catalysts for Change in Indigenous Education. Alaska Native Knowledge Network, Fairbanks, Alaska.
- Barnhardt, R., Kawagley, A.O., 2005. Indigenous knowledge systems and Alaska Native ways of knowing. *Anthropol. Educ. Q.* 36 (1), 8–23.
- Brower, P.K., 2016. Tumitchiat: Iñuqqaat aullarrisiatun ilisagviit—A New Pathway: Indigenous Leadership in Higher Education. Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Drabek, A.S., 2012. Liitukut sugpiat’stun (We are Learning How to Be Real People): Exploring Kodiak Alutiq Literature Through Core Values. Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Fisher, C., 2018. Diideets’ii in Our Pathway (in Our Future): Gwich’in Educational Philosophy and Transformative Praxis in K-12 Education. Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Gordon, H.S.J., 2019. Self-Determination, Sustainability, and Wellbeing in the Alaska Native Community of Ninilchik. Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Hill, F., Barnhardt, R., Kawagley, A.O., 2006. 2005 AKRSI Final Report. Alaska Federation of Native, Anchorage, Alaska.

- John, T.A., 2010. Yuraryarput Kangiit-Ilu: Our Ways of Dance and Their Meanings. Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Kirkness, V.J., Barnhardt, R., 2001. First Nations and higher education: the four R's—respect, relevance, reciprocity, responsibility. In: *Knowledge Across Cultures: A Contribution to Dialogue Among Civilizations*. Hong Kong: Comparative Education Research Centre, The University of Hong Kong.
- Roderick, L., 2008. Do Alaska Native People Get Free Medical Care?: And Other Frequently Asked Questions About Alaska Native Issues and Cultures. University of Alaska Anchorage, Anchorage, Alaska.
- Shorty, N., 2015. Inland Tlingit of Teslin, Yukon: Gaanax.ádi and Kookhitta'an Clan Origin Stories for the immediate and Clan Family of Emma Joanne Shorty (Nee Sidney). Doctoral Dissertation, University of Alaska Fairbanks. ProQuest.
- Topkok, S., 2014. Native ways of networking. In: Alvares, C.A. (Ed.), *Multicultural Knowledge and the University*. Multiversity, Alor Setar, Malaysia, pp. 143–149.
- United Nations, 2007. United Nations Declaration on the Rights of Indigenous Peoples. United Nations, New York, NY.
- World Indigenous Peoples' Conference: Education, 1993. The Coolangatta statement on Indigenous rights in education. In: New South Wales, Australia: Task Force of the World Indigenous Peoples' Conference on Education.

Journeying to the common good: exploring Indigenous Solomon Islands leadership through five moments

Kabini Sanga and Martyn Reynolds, School of Education, Victoria University of Wellington, Wellington, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Indigenous leadership	151
Solomon Islands	151
Five moments	152
The labu: leadership as worship	152
Rara'aba: leadership as knowledge guardianship	153
School and kastom leadership: negotiation across domains	154
Climate change: leadership as hope	156
Reflection: leadership as consistent and contiguous	157
Closure	158
References	158

Indigenous leadership

All leadership, including Indigenous leadership, takes place in context. [Dei \(2011\)](#) helpfully suggests that “[f]undamentally, the Indigenous should be perceived as mostly about place-based knowing, and understanding of traditional sacred relationships between peoples, their cultures and their cosmologies. These relationships offer a holistic knowledge base” (p. 23). [Blakesley \(2010\)](#) adds, drawing on [Waters \(2005\)](#), that indigeneity is multifaceted, an attribute, always referenced against a specific place, but also an element in personal self-identification. Indigeneity is “a complex and nuanced construct, carrying with it attributes specific to place or geographic location, legal status, and identity” ([Blakesley, 2010](#)). In offering this chapter, we acknowledge complexities of space, time and belonging as aspects of context in the Indigenous leadership that Solomon Islanders recognize and practice. We utilize a lens that celebrates how people relate through leadership socially, culturally, cosmologically, geographically and temporally. Our intent when undertaking this enterprise is to celebrate and honor what Solomon Islanders know and practice daily, informed as they are by kastom (customary wisdom and practice) from the past, and visions for sustainable futures.

Much has been written about Indigenous leadership ([Long, 2017](#); [Khalifa et al., 2019](#); [Edwards, 2015](#); [Henry and Wolfram, 2018](#); [Von Der Porten, 2012](#); [McCalman et al., 2012](#) and others). Writers have stressed strengths-based approaches that honor Indigenous wisdom as complete and sufficient in itself; the significance of the knowledge on which leadership is based; communal approaches to group well-being; spirituality, often revealed in altruistic and service-framed leadership; collectivist communication practices; the social self; the ontological and epistemological foundations of leadership; and the need to embrace a timescale of millennia. These and other relational aspects of leadership resonate with our approach which is to take the “direction of naming” ([Sanga and Reynolds, 2017](#), p. 199) of leadership from the moments we present, paying attention to how leadership is understood in space and time, and thereby adding to the wisdom of leadership literature through specificity rather than generalization. We acknowledge that the methodology of using five moments upon which to base discussion may produce an incomplete view of Indigenous Solomon Islands leadership. However, the diversity of our examples affords significance to our claim that consistency in leadership illustrates contiguousness across space and time.

Solomon Islands

Geographically, Solomon Islands is an archipelago of over 900 islands situated east of Papua New Guinea (PNG) and west of Vanuatu. Culturally, Solomon Islands is diverse—around 80 Indigenous languages are spoken, and multiple Indigenous cultural groups live within the modern nation state ([Dinnen, 2002](#)). A Melanesian creole ([Jourdan and Angeli, 2014](#)), Pijin, provides a lingua franca for most Solomon Islanders. In conjunction with Tok Pisin (PNG) and Bislama (Vanuatu), mutual communication is also possible across other Melanesian populations.

Settled since the time of the Lapita peoples ([Irwin, 2005](#); [Kirch, 1997](#)), Solomon Islands has had a complex recent political history. In 1978, its status as a British Protectorate switched to that of an independent state and member of the British Commonwealth. Relationships that feature donor aid continue with several bodies including the governments of Australia, Aotearoa New Zealand, The United States, and, latterly, China ([Kabutaulaka, 2010](#)). The capital of Solomon Islands is Honiara on the island of Guadalcanal. Auki, the provincial capital of Mala'ita Province is another small township. Much of the country is rural and many people follow a subsistence lifestyle, often aquatically based ([Basel et al., 2020](#)).

Solomon Islanders generally recognize three domains of influence (Sanga and Reynolds, 2019). Kastom is the domain that encompasses customary understandings and practices; the church domain centers on various Christian denominations; the domain of formal institutional life includes government, education, and other statutory bodies. These domains compete for influence over individuals and groups, but their power is generally ordered as given. In accounting for matters such as leadership, one must consider in what domain leadership is being practiced, through what domain the legitimacy of a leader is established, and how leadership in one domain negotiates with leadership from others (Sanga et al., 2018, 2020; Sanga and Walker, 2012). In this account of Indigenous leadership, although much attention is paid to the kastom domain, we also show the strength of kastom-originated leadership as it influences other domains, particularly through the example of education.

Our approach of presenting five diverse moments entails a need to examine what Indigeneity might mean in the Solomon Islands context. The linguistic diversity of Solomon Islands reflects multiplicity so that “Indigenous” can be understood at village, clan and/or tribal levels. However, there is much population movement in the nation state of Solomon Islands and beyond which requires a nuanced appreciation of Indigeneity in Indigenous Solomon Islands leadership. Gegeo (2001) separates space and place in ways that are helpful here. Place is the site, location and root of Indigeneity. It is the genesis of the various Indigenous ideas of leadership understood by a specific group. Space, however, extends beyond the physical into the social, spiritual and moral. It expands so that Indigeneity travels with the voyager to new places, a relevant factor when, for instance, teachers are relocated by government around the archipelago. Thus, when presenting the moments, we embrace the contextual fluidity that affords significance to place-specific Indigenous leadership but also acknowledges space-relevant expressions of Indigenous leadership. This duality makes relevant the notion of Indigenous Solomon Islands leadership in ways that extend beyond the individual’s physical location in village and genealogical location in their clan.

Five moments

In order to investigate Indigenous Solomon Islands leadership, we have selected four moments as examples. These are moments in the sense of their transience in time, especially within a millennial timeframe. They are also “of moment,” significant in their effect on our storied account of leadership. The four range in temporal terms from a historical past through a contested present and into a speculative future. They also extend from being situated solely in the kastom domain to negotiation between kastom and other domains. Through their diversity, the moments offer opportunities to trace links that attest to the strength of Indigenous Solomon Islands leadership and its consistency as a journey to the common good. A fifth moment of reflection is also added.

The labu: leadership as worship

This moment is informed by Kabini as the principal guardian of the Ngongosila labu of the Gula’alā. The storied account draws from “The Mala’ita Labu: Rescripting the politics of living well together” (Sanga, *in press*) to be published by the Pacific Theological College in Re-Storying the Pacific Household Volume 1, edited by Upolu Va’ai and Aisake Casimira.

A labu is a fort-like village of refuge. Labu were common throughout central, east, west, and north Mala’ita, Solomon Islands, in tribal settings during the pre-Christian era. The Mala’ita labu was a sanctuary, a safe space for all where everyone was given their place, including an opportunity to start afresh in life. The labu was of great value to those being hunted. People who were cursed or accused would be listened to by the labu guardians. Those fleeing temporarily from abuse or other forms of harm would find care and welcome. Leadership ensured that, for visitors and refugees alike, the labu was a nurturing and healing space.

The labu guardians vowed to ensure sanctity by remaining loyal to the mission of the labu and by protecting its purpose and purity. Asylum seekers promised to surrender their authority as independent agents to the labu leaders in exchange for guarantees of safety, protection, justice and custodianship. In a pre-Christian multi-tribal, multi-lingual, multi-cultural environment, the labu was a spiritually charged oasis of rest, a place of hope, of new beginnings, new creations and of co-creations of new futures.

The structure of the labu involved a physical compound protected by hedges, walls, and strong fences. It was secured by sentries. This was a sacred space, closely protected in theocratic Mala’ita through the leadership of tribal priests via a stringent moral code and punishment system. For those inside, the labu was a place of promise. Against those outside who wished to cause harm, the labu had human and spirit guardians, feared by foes and hated by labu enemies.

Leadership in the Mala’ita labu is best understood as worship. The labu was a place of worship where residents’ relationships to and worship of their deities were central pillars of a theocratic community. In Indigenous Mala’ita, *rō lā* (obedience) was a primary societal value. According to Sanga (2014b) for the Gula’alā people, “*rō lā* or obedience is an act of worship, and a means of submission to and demonstration of loyalty to the other, particularly to authority” (p. 156). This understanding of worship as relational remains in today’s Christian Mala’ita.

Conceptualized as worship, labu leadership was manifested in the tribal priests’ regular offerings of sacrifices, both on their own behalf and for their people. Worship showed in the ceremonies which priests would perform, in the offerings made to ancestor spirits, in the purification rites they would administer, in the scheduled and contextually called sacred feasts, and in the gift-giving in which people would participate. To be effective in such leadership, strict statutes and codes governed labu communal and private living.

Through their abua or spiritual rootedness and abu lā (holy living), labu leaders exerted influence, exercised stewardship, and played their roles as hosts of their tribal communities. Coupled with an effective support system and cohesive communitarian

capacities, labu leadership provided safety and space for members to individually and communally worship, thereby allowing purposive growth and flourishing for labu members and the community as a whole.

One labu was the offshore island of Ngongosila in east Mala'ita, in operation until the early 1950s. In early 1900s, Christian missionaries from the South Sea Evangelical Mission leased a portion of Ngongosila as their base for Christianization along the northern, eastern, and southern regions of Mala'ita. The northern part of Ngongosila remained a stronghold for the principal tribal group, the Gwailao, whose labu was the island and its surrounding waterways.

The sanctity, mission, and effect of a labu relies on Indigenous Solomon Islands leadership. Ngongosila embodied a well-cared for labu for an extended period of time, stretching back into the period where kastom was the only lore, and continuing until quite recently. Sadly, the community has not remained the example it had been. Historical, national, and local politics created rifts so that community cohesion waned, and consensus was eroded. Leadership acts that were once applauded became derided. Like other Solomon Islands communities today, Ngongosila is experiencing acute vulnerability and needs revisiting and reimagination. More than restoration, a vision of a new created community is needed.

However, like the stones of Ngongosila, Indigenous Solomon Island leadership endures to offer a pathway forward to a sustainable future and a renewed labu vision. The Gula'alā people mourn the dug-up coral reefs and loss of marine life around the island; weep over the loss of effective tribal authority; regret the dismantled community cohesion; and deplore the inward-looking confrontational leadership and resultant contentious splintering of relationships in the village. However, there is hope centered on Indigenous Gula'alā ethical teachings, leadership such as is contained in the joyfully welcomed book, *Fānanau lā 'I Gula'alā* (Sanga, 2014a). Hope in this instance is borne out of contextual public grief and cultural ontological loss.

The mission of the labu is being revived and augmented. Through the book, elders can now teach their young ones again using Gula'alā language and their own Indigenous knowledge. In tandem, assurance comes in the form of generous Gula'alā landholding clan leaders who have gifted free lands on the eastern coast of the main island of Mala'ita to displaced and diasporic Gula'alā, those who have lost homes due to the damage of sea-level rise on Ngongosila atoll. As a result, new Gula'alā-speaking communities have been built and settled, new farms and lands cultivated on mainland Mala'ita.

Spiritual leadership is also being rejuvenated. Through prophetic teachings by Gula'alā pastors and church leaders, the Gula'alā are assured of God's sovereignty and care; spurring the people on to give of themselves as Gula'alā as they had done through the labu. The Gula'alā Volunteers Ministry Band (a men's team of builders) was formed and travels to different parts of Mala'ita and the Solomon Islands enacting the spirit of the labu by building without payment churches and community halls for other groups.

In labu understanding, departure is an essential part of alidea or homeward journeying. The Gula'alā have had to learn to develop capacities for adaption, responsiveness and change. Labu leadership has required the Gula'alā to depart in haste or as planned, in sorrow and/or in joy. After all, their historical settlement on Ngongosila atoll was a result of an intentional departure from South Mala'ita in a previous century.

For the Gula'alā, any departing must be joyous. Even if sad or hard (in physical and/or experiential ways), departure is liturgical and imaginative. The theocratic Mala'ita mind believes this. Labu understanding urges Gula'alā leaders to a departure frame of reference because it is counter to self-serving and enslavement ideologies and/or loyalties.

Rara'aba: leadership as knowledge guardianship

This moment is informed by Kabini as alafa or male tribal leader of Gwailao. The storied account draws on Sanga and Reynolds (2020).

A rara'aba is a calming of nerves meeting. The moment we describe took place in a rara'aba of the Gwailao clan conducted in the Gula'alā language. Material was translated by Kabini, member-checked and then cleared for publication by the council of the principal alafa of Gwailao. The subject of the meeting was debate of, and approval for, the tribal secret knowledge contained in *Fānanau lā 'I Gula'alā* (Sanga, 2014a) to be publicly gifted for education purposes.

Knowledge guardianship of private material is a key area of leadership in Solomon Islands. In Mala'ita, Solomon Islands, there are three kinds of knowledge: public, specialized and sacred. Secret knowledge can be located in the second and third of these categories, although in an integrated Mala'ita worldview, both knowing and not knowing affect the whole of life (Sanga and Reynolds, 2020). As an example, Gegeo and Watson-Gegeo (2001) observe that for the Kwara'ae of Mala'ita, knowledge associated with certain rituals, sorcery and healing persists in secret domains, notwithstanding the influence of Christianity. Similarly, Sanga (2014b) describes moral education in Gula'alā, Mala'ita, as existing in a private knowledge domain guarded by sub-groups in clans. Because guardianship assumes responsibility for sharing or withholding what is known, leadership, privacy and knowledge are inextricably linked in these Indigenous Solomon Islands contexts.

Examining authority to maintain and/or negotiate the placement of knowledge in certain domains offers an appreciation of the dynamics of Indigenous Solomon Islands leadership. Guardianship over knowledge involves position, but this is not solely an individual matter. In communitarian Mala'ita, a person's identity can be both individual ("I") and communal ("We"). When an individual plays the role of a knowledge guardian they lead as an individual and as a family or clan representative. Two principles exercised in the rara'aba show how authority in Indigenous knowledge guardianship is underpinned.

First, the principle of 'ado lā ana rau lea fainia tōa or stewardship articulates the value of sharing something good with other people. Knowledge is assumed to be a good thing, worth sharing now and into the future. Knowledge must have guardians to ensure that the knowledge is well cared for. Second, the 'ado lā or stewardship principle is operationalized through the notion of nao-nao lā or seniority. Thus, the question of leadership authority in the matter the publication of *Fānanau lā 'I Gula'alā* (Sanga, 2014a)

involves appreciating guardianship in relation to the knowledge, and nuanced sensitivity regarding position in relation to others. Developing these understandings in context was a purpose of the *rara'aba*.

In the context of the *rara'aba*, seven dimensions of seniority are given in Sanga and Reynolds (2020). Seniority can involve matters such as *garangi*, guardianship vested by virtue of a direct line to the first custodian of the knowledge; gender and age such as through *futa lā* and *wane me geni*; generational closeness to the principal knowledge guardian expressed as *fuli wane*; blood relations as in *futa lā ana tē 'abu*; wider relationships including blood, marriage and adoption as in *wane futa-wane kwaitā*; and *wane goumouri* which involves agency, credibility and esteem. Social change, such as population movement away from the village and the availability of publication technologies, is a source of dynamism that, when framed through principled thought, can result in changes in the way knowledge is positioned. Dimensions of seniority may appear fixed, but actually require fluid negotiation when faced with innovative circumstances such as book production.

Two examples of how the *rara'aba* interrogated secrecy are offered. These were addressed in a context where secrecy itself was not debated, but the qualities and extent of the secrecy were at stake. First, a consideration of time offered opportunities to sift the relationship between the origin of the knowledge and its secret status. A well-known person, *kōkō Ruth*, was identified as the origin of the knowledge to be published. However, genealogical questions were asked to probe where and when she obtained the knowledge. This shifted attention to the knowledge itself which pre-existed the identified individual and, consequently, drew attention to clan structures as means of transmission. The discussion of secrecy was thus extended to prior generations and the communal parameters of secrecy were widened. The *rara'aba* served to erode any unequivocal claims regarding the level and type of secrecy relevant in this case. In effect, definitive claims of secrecy were revealed to rest on less secure historical footing than might have been assumed.

Second, a consideration of the movement of knowledge through space enhanced participants' perspectives regarding the secrecy. One speaker pointed out the way that women bring knowledge with them when they enter a clan by marriage. As a result, the origin and distribution patterns of secret knowledge are unclear. When knowledge comes through this route, is it held in expressions of wisdom or the teaching that the wisdom embodies? Embodied wisdom is visible and therefore the boundaries of secrecy are more inclusive than exclusive and further transfer is possible by learning through action. If knowledge is understood in this way, the sense of secrecy may not be intense, and guardianship deserves reconsideration. In addition, some people who should be "inside" the borders of secrecy may behave in ways that suggest they have not acquired the wisdom embedded in the knowledge and are therefore "outside." This kind of interrogation of knowledge transfer through space questions the relevance of certain principles of seniority because people's relationships to knowledge deemed secret is not always predictable.

The application of principles of knowledge guardianship in the *rara'aba* reveals much about Indigenous Solomon Islands leadership.

- First, leadership involves the application of known and agreed principles, but in ways that sift a complex situation to arrive at a conclusion. Principled leadership, therefore, is not absolute but contextual. The intersection of many contributions to the *rara'aba* from clan members resulted in a decision by and on behalf of the clan. The decision was arrived at by peeling back layers of the context to reveal the quality, intensity and significance of secrecy.
- Second, leadership responds to wider context. In this case, an aspect of context is the movement away from the village of clan members who occupy certain genealogical and guardianship positions. In this circumstance, knowledge transfer cannot be guaranteed. Gifting of secret knowledge for publication is a way to support transfer that has implications for secrecy. Thus, leadership must deal with the relative value to the clan of transfer when compared to secrecy. Balance must be achieved in ways that honor all those who generated, hold, and will benefit in the future from the wisdom contained in knowledge. Thus, leadership is never static.
- Third, leadership honors unity. The process of the *rara'aba* allows what might seem to be conflicting and divisive principles to promote group unity. The discursive structure promotes language that wraps principles in dignity and steers discussion in indirect ways to prioritize harmony over a swift resolution. This takes place in a context where food is shared and socializing is central so that the coming together, form of the *rara'aba*, and discussion itself all support the enhancement of warmth between clan members as the most significant outcome.
- Fourth, leadership values action. Through the *rara'aba* the guardianship of knowledge came to rest with those who practice the wisdom it contains. Since secret knowledge as discussed here is centered on sustainability, the form and purpose of the *rara'aba* are united as expressions of the wisdom of unity. As a result, the *kastom* tenet of generosity actioned through gifting became the focus of the solution rather than breaking *kastom* for gain or overstepping *kastom*-derived protocol boundaries and principles.

Underpinning these four aspects of Indigenous Solomon Islands leadership is its relational core, in this case exercised dialogically. The *rara'aba* frames leadership primarily in ontological rather than political ways in a context. Clan members are both "I" and "We," and leadership decisions are made by, on behalf of, and for the benefit of the collective. Leadership takes place by acknowledging complexity as an aspect of context and prioritizing "We" as the touchstone of Indigenous Solomon Islands life.

School and *kastom* leadership: negotiation across domains

This moment is informed by research conducted through the auspices of Victoria University of Wellington and the Institute of Education, University of the South Pacific, funded by the Development Leadership Project and supported by the Australian Department of Foreign Affairs and Trade. The data were collected by research partners Grace Rohoana and Graham Hiele of Solomon Islands National University from school and community members in various Solomon Islands locations using a *tok stori* approach

(Sanga and Reynolds, 2019, 2021; Sanga et al., 2018). Tok stori is a Melanesian kastom relational orality used for many purposes including conversation, pedagogy (Sanga et al., 2021), evaluation (Paulsen and Spratt, 2020) and research (de la Torre Parra, 2021). To illustrate Indigenous Solomon Islands leadership negotiations with the institutional domain, we select data from tok stori sessions involving two school leaders, one female (F) and one male (M).

Through tok stori, F offered two metaphors for leadership, one potentially ancient and therefore kastom based and inherited, and the other comparatively recent and perhaps church derived. "Leadership is explained in the canoe concept—its paddlers are expected to paddle in both sides of the canoe so that the canoe will go on the straight direction. If the paddlers only paddle on one side of the canoe, the canoe will not go straight and may sink in the storm. Another metaphor ... is light. Light needs power. Supposing that there is no power, the light will not be able to shine on everyone so that they can see the leader. Like so, a leader has to shine on everyone, but this can only be possible through possessing important qualities. A leader must be transparent so that people know what they are doing rather than hiding in the darkness." These metaphors draw attention to the collective responsibilities of leaders and the shared outcomes of leadership. Through them, F suggests "transparency, trust and respect" are important aspects of Indigenous Solomon Islands leadership.

F's explanation of the two metaphors draws attention to the visionary and relational nature of leadership. In the kastom context of the canoe, navigation is important; the fate of paddlers is interlinked by the shared vessel. Trust is key when contributing to collective movement and the fate of the vessel through paddling. F applies this kastom thinking in school: "In the school, the head teacher and the school principal are leaders. The school board chair and school board look after the school. The roles of these leaders are to direct and lead people in order to carry out what is needed to build the community. From my experiences, if trust is not developed, even teachers will start to murmur—and they will choose not to come to school anytime they choose to."

Similarly, the metaphor of light shows the way institutional leadership is informed by church example. Describing her school context, F says "I learned from the catechist and my uncle the importance of being transparent. I admire some of the leadership attitudes and practices they demonstrated in the church circle ... The catechist is so transparent in their work, so I try to do likewise in my own work. Yes, if I do not practice transparency, I may not develop the trust from my teachers and parents."

While negotiation between domains leads to positive outcomes when key leadership elements travel into school leadership, the divisive consequence of self-interest can also be disruptive in multiple domains. F narrates an example. "In the community, I learned that leaders do not always trust each other ... One thing that pulled leadership back is logging ... Before logging arrived in the community, the people at home used a metaphor and said: "Do not open the parcel" and "Leave it closed" because if the parcel is opened, we will be separated—and that happened." She has seen similar divisions in school life. "One thing I learned is that [some in] the community see the school as their private school. According to the history of the school, a woman started the school and today her grandchildren are the ones that expect the school to give them money every time the grant money is received. So, they come and swear at us teachers and asked for school grant money. They regard this school as their private business."

The focus of F's school leadership is not her managerial function, although this is present. The heart of her practice is relational, underpinned by values that originate in kastom. The application of the canoe and light metaphors by F as a school leader suggest the relevance in institutional life of kastom ideas and divisive consequences regardless of domain when key aspects of kastom leadership are absent.

M inherited a leadership metaphor. "I always remember a teaching from my parents in a form of a metaphor—a house. When you build a house, the foundation of the house must be built properly till you reach the roof so that when the house is completed it is firm and strong. My parents always reminded us of this metaphor. Since then, I became aware of my father's ideas as a leader in the family. It has helped me a lot because I often reflect back on my father's leadership approaches and methods when dealing with difficult issues that arise."

M unpacks the metaphor through dispute resolution. "When there is an issue involving two persons in the community, it is not proper for me to make other members in the community become aware of their conflict. Such matters I must keep confidential. To settle such disagreements, and to be fair for both conflicting parties, I have a private discussion with each one to seek their views. I found that this approach is effective in settling disagreements among community members." This process promotes alignment between individuals to support a strengthened sense of the collective.

However, in diverse Solomon Islands, kastom leadership is not always identical. M recognizes the complex origins of parents and staff, and the varied ideas from Indigenous Solomon Islands leadership they carry. "In schools we know that we work with students, teachers and parents of different ethnic groups ... Because these people are from different backgrounds, it is impossible to just show up and say, "Let us do things this way" and expect them to adhere to my plans." M's inspiration when molding community from diversity is his parents' example. As with dispute resolution, this involves small scale relational activity within general transparency. To gain community traction in a classroom building project, "I collected the views from members of the community whom I could easily approach and then began designing the best options. I began to engage in networking. While some members at first ignored my request to help ..., the progress of the work done slowly got them to get on board. Today I am pleased with community members' level of support for the school."

M provides specific aspects of kastom practice. "I learned from the community ... to prepare a meal for anyone which I have invited to meet and discuss with me. I applied this approach as well in school ... Sometimes some parents complained that they are tired of doing work for the school. Because the reality is many parents have limited or no formal education at all and they do not earn money from engaging in schoolwork ... So, providing meals for parents that participate is a gesture of appreciating their commitment. In this way they are willing to participate in other school projects or work." When the gifts of time and labor are

paired with the gifts of food and the effort of cooking, the balance achieved through generosity lends support to leadership legitimacy. As a result, community is built through collective activity.

Like F, M has experienced the way leadership in the kastom domain can be eroded when the interests of individuals undermine the good of the collective. "This type of leadership causes some disgruntled members of community to not respect decisions or directives from community leaders. For example ... finances provided by some external sources for the benefit of the whole community are used by the leader, or only shared to some members of the community. Even our cultural practices have changed. For example, gifts and food during weddings are not given to every member of the community as practiced in the past." The institutional domain is not immune from such effects. "Some relatives sent three of their children to live with me and attend classes in ECE. But my relatives never attended schoolwork during Parents' Day. One evening I visited their houses and explained to them that the school does not belong to you and me. This is a school and if you are concerned with your child's education then you need to participate in schoolwork on Parents' Day."

These two school leaders' tok stori sessions reveal much about Indigenous Solomon Islands leadership.

- First, Indigenous Solomon Islands leadership values are influential beyond the kastom domain. In some cases, kastom leaders are also school leaders who carry values from their kastom socialization in their families. However, the values are also present in the relational bedrock on which kastom leadership practice is built. The degree to which the school and community are integrated is a predictor of the success of the school as an institution, at least as far as the provision of buildings is concerned. Kastom values such as transparency, trust, and respect bind people together through mutual appreciation and integrity, regardless of domain.
- Second, kastom leadership transcends domains through metaphor. Building and paddling are kastom activities undertaken as a collective that support survival. Light is a benefit shared by all. Each metaphor speaks in its own way of the significance to leadership of collective existence and the importance of shared fate when judging leaders' success. Taken together, the metaphors offered in this leadership moment relegate individualism to the status of an unproductive path at best, and a catastrophe such as drowning or blindness at worst.
- Third, leadership values social cohesion. Cohesion is represented in the values discussed above and is an aspect of all three offered metaphors. However, cohesion requires enactment. F's tok stori shows the deliberateness of a school leader seeking to enact trust and transparency in the school context. M's tok stori shows how kastom protocols such as house to house visits and the provision of food support social cohesion. For F and M, a cohesive relationship between school and community is prized as a way forward and is an aim of their leadership.

Climate change: leadership as hope

The Indigenous Gula'alā of east Mala'ita Solomon Islands, have been battling five decades of sea level rise and the daily erosion of their atoll home. Since the 1960s, the Islanders have fought against the destructive effects of higher water levels, cyclones and ongoing current erosion. A 1991 report, SOPAC Technical Report 121 (Rearic, 1991), noted that the coastal retreat for Ngongosila and Kwai atolls was 2–3 m per year, and that the process had been happening since the 1960s. The atolls are 300–500 m in length and 100 m wide. While the highest point on each atoll is 2 m above the mean sea level, much of the islands are around 1 m above mean sea level. Moreover, both atolls are exposed to 1000 km of open ocean.

These factors have meant that for several decades the Gula'alā have been exposed to environmental challenges on an everyday basis. Moreover, while densely forested and sparsely populated in the 1920s–1930s, the Gula'alā atolls of Ngongosila and Kwai have seen considerable overpopulation for their sizes, thereby adding to the environmental stresses which were and are present. It is little wonder that people have referred to the Gula'alā atolls as the "sinking islands" (van der Ploeg et al., 2020), portraying the Islanders as victims of climate change.

Beyond physical changes for Mala'ita people, the manifestations of community loss from the climate crisis are multiple. Loss is seen in the littered unplanned brown sago-leaved houses and unkempt smoky cooking huts. Loss is shown by the captive parrot flapping its wings all day on a house laundry line. Loss is present when feared ancestor skulls and aged bones are eaten by the Ocean deep. Loss is experienced in once gendered repositories and sites of knowledge being eroded, their purity watered away, thereby rendering them ordinary and secular.

As a leadership response, the sili (Indigenous Mala'ita lament) which follows, captures the grief of the Gula'alā as they respond to the changes of their environment. In it, the Gula'alā express grief over a diminished life and abandonment:

'Oi, ote wela 'ae (O you of youthful understanding).

Molu fā angī nau laona asoā 'I tara'ena (You have caused my heartache today).

Ku manata 'agu to'ona (Inciting memories of bygone days).

Na fataia na kō nau kini, na abua babatō gera kini, na bae abu gera kini, ma na ramoa gera kini laona tolo ma na asi. (The exploits of my ancestors, the sacred utensils which rooted my worlds, the sacred shrines of power and the life-giving spirit of my very existence).

Kō nau kini wane gera rangonia toa, gera adafafia toa, gera ogua toa, gera ka aofia fana toa (The generosity of my ancestors, their stewardship over an integrated world, their cohesive leadership as guardians and hosts).

Kō nau kini gera rigita ana 'aba gera (The autonomy, strength of hand and authority of my ancestors).

Molu ngalia nō fāsi nau. Molu fā inomae nau 'I tara'ena. Molu ada siofa agu (You have robbed me of life today. You have made me an orphan, a destitute in my own world).

'Oi, keleirurūfia lea nau 'ae (It is in grief that I lament uncontrollably now).

In Gula'alā leadership, grief is publicly voiced. Doing so produces hope as it is only in the context of loss and grief that hope is credible and authentic. Comfort arises out of the grief over Ngongosila's ontological and theological loss. Subsequently, hope comes to the people in the form of the occurrence of something new. In this instance, the newness comes in the form of the Indigenous Gula'alā teachings on ethics as contained in the book, *Fānanau lā 'I Gula'alā* (Sanga, 2014a). In introducing the compilation of the Indigenous Gula'alā ethics teachings, Sanga (2014a) explains, "These teachings are gifted to and to be generously shared with all Gula'alā children (p. 15)." He further urges: "Fathers and mothers are to sit with their children and teach them diligently and daily about these precepts" (p. 15).

Reflection: leadership as consistent and contiguous

As we reach the end of our chapter, we pause in a fifth moment, a moment of reflection. We reflect on in the way that consistency in Indigenous Solomon Islands leadership provides contiguousness across space and time. Across moments one to four, Indigenous Solomon Islands leadership can be observed seeking social cohesion in the form of unity. From the ancient labu to the modern school, the one-ness of Solomon Islands societies is an aim of leadership fostered by a communal approach. That approach is relational, a second contiguous element.

Relational leadership connects the leader to the group, and the effectiveness of leadership to group-based outcomes. Indigenous Solomon Islands leadership is concerned with how the actions of the leader benefit everyone. In the rara'aba, concern was shown through leadership for those who had originated secret knowledge, those holding it currently, and those yet to come. The lament references the past, present and the future. These temporal conceptions that extend beyond the here and now intensify the responsibility of present relationships because honor is significant, and time is not linear. A society in which honor is degraded is not well led and a society that denigrates its past can have no future.

Indigenous Solomon Islands relational leadership extends beyond the human. Leadership as worship in the labu shows how the geographical and cosmological are inextricably linked. Similarly, the metaphors supplied by educational leaders show how leadership delivers wisdom from the kastom domain to the institutional. In these cases, leadership is transcendent, a principled matter in a holistic world where the division of time and space or spirit and matter can lead to unethical behavior, but where relationality is the key to sense making.

The core of the common good to which Indigenous Solomon Islands leadership is oriented is sustainability. This is linked to cohesion, well-configured relationality, and holism. However, sustainability is not a synonym for a static approach to leadership. This is because the world is changing as it always has. To serve the common good, Indigenous Solomon Islands leadership is dynamic. It deals with leadership at the intersection of domains, the advent of new technologies and social settings, and increasingly with climate change.

The dialogic nature of Indigenous Solomon Islands leadership is consistent across moments. Principled leadership must enter into dialog with circumstance and those affected. Hope extends from worship, honoring the stones of Ngongosila labu in response to change. It provides a route to a future where the strengths of Solomon Islands communities adapt while being ontologically grounded. Dialog embodied in the rara'aba is leadership in action. The wisdom of the past structures practice in the present to enable hope for the future. Leadership aspects such as transparency, trust and respect cement together kastom and institutional domains as aspects of a single whole. In these ways, Indigenous Solomon Islands leadership ensures the temporal remains

integrated with, not separated from, the social, cultural, cosmological, and geographic in a connected reality. Thus, Indigenous Solomon Islands leadership provides forward looking vision in contiguous dialog with the past.

Closure

This chapter has made a case for Solomon Islands Indigenous leadership as springing from, and embedded in, Solomon Islands societies. It is leadership that seeks sustainability as the common good in socially, culturally, cosmologically, geographically, and temporally appropriate ways. Our storied exploration described Indigenous Solomon Islands leadership through five moments: the labu or sacred island space; the rara'aba or calming of nerves meeting; school leaders' negotiations of the intersections between kastom (customary) leadership and school leadership; lament as hopeful attention to climate change; and a moment of reflection. Together, the five moments represent diverse but contiguous aspects of life in Solomon Islands. Their assemblage encourages appreciation of the strength of Indigenous Solomon Islands leadership to respond constructively to changing circumstances on a journey that is framed around community cohesion in pursuit of a sustainable future.

References

- Basel, B., Goby, G., Johnson, J., 2020. Community-based adaptation to climate change in villages of Western Province, Solomon Islands. *Mar. Pollut. Bull.* 156, 111266.
- Blakesley, S., 2010. Storytelling as an insightful tool for understanding educational leadership in Indigenous Yukon contexts. *Can. J. Educ. Adm. Pol.* 111 (October 14), 1–31.
- de la Torre Parra, L., 2021. *Storying Place: A Tok Stori About Relationalities in Oceanic Education and Development*. Ph.D. Victoria University of Wellington.
- Dei, G.J.S., 2011. Revisiting the question of the "Indigenous". In: Dei, G.J.S. (Ed.), *Indigenous Philosophies and Critical Education: A Reader*. Peter Lang, New York, NY, pp. 21–33.
- Dinnen, S., 2002. Winners and losers: politics and disorder in the Solomon Islands 2000–2002. *J. Pac. Hist.* 37 (3), 285–298.
- Edwards, G., 2015. Anthropological accounts of leadership: historical and geographical interpretations from indigenous cultures. *Leadership* 11 (3), 335–350.
- Gege, D.W., 2001. Cultural rupture and indigeneity: the challenge of (re) visioning "place" in the Pacific. *Contemp. Pac.* 13 (2), 491–507.
- Gege, D.W., Watson-Gege, K.A., 2001. How we know: Kwara'ae rural villagers doing indigenous epistemology. *The Contemporary Pacific* 13 (1), 55–88.
- Henry, E., Wolfram, R., 2018. Relational leadership—an Indigenous Māori perspective. *Leadership* 14 (2), 203–219.
- Irvin, G., 2005. Pacific Migrations—Into Remote Oceania: Lapita People. *Te Ara, the New Zealand Encyclopedia*. <http://www.TeAra.govt.nz/en/pacific-migrations/page-3>.
- Jourdan, C., Angeli, J., 2014. Pijin and shifting language ideologies in urban Solomon Islands. *Lang. Soc.* 4 (3), 265–285.
- Kabutaulaka, T., 2010. Milking the dragon in Solomon Islands. In: Wesley-Smith, T., Porte, E.A. (Eds.), *China in Oceania: Reshaping the Pacific*. Berghahn Books, New York, NY, pp. 136–150.
- Khalifa, M.A., Khalil, D., Marsh, T.E., Halloran, C., 2019. Toward an indigenous, decolonizing school leadership: a literature review. *Educ. Adm. Q.* 55 (4), 571–614.
- Kirch, P.V., 1997. The Lapita peoples: ancestors of the oceanic world. In: *The Peoples of South-East Asia and the Pacific*. Blackwell, Cambridge, England.
- Long, J.E., 2017. Framing indigenous leadership. *Adv. Soc. Sci. Res. J.* 4 (6), 248–257.
- McCalman, J., Tsey, K., Kitau, R., McGinty, S., 2012. "Bringing us back to our origin": adapting and transferring an Indigenous Australian values-based leadership capacity-building course for community development in Papua New Guinea. *Community Dev.* 43 (3), 393–408.
- Paulsen, I., Spratt, R., 2020. When evaluation and learning are the intervention. In: Johansson-Fua, S., Jesson, R., Spratt, R., Coxon, E. (Eds.), *Relationality and Learning in Oceania*. Brill | Sense, Singapore, pp. 135–153.
- Rearic, D.M., 1991. Coastal Environment of Kwai and Ngongosila Islands, Malaita Province, Solomon Islands | SOPAC Technical Report 121. SOPAC, Suva, Fiji.
- Sanga, (in press). *The Mala'ita Labu: Rescripting the politics of living well together*. In: Va'ai, U., Casimira, A. (eds.), *Re-Storying the Pacific Household Volume 1*, Pacific Theological College, Suva, Fiji.
- Sanga, K., 2014a. *Fananau lā i Gula'alā*. TNR Literatures, Wellington, New Zealand.
- Sanga, K., 2014b. A first look at an Indigenous Pacific ethical system and its implications for research. In: Otunuku, M., Nabobo-Baba, U., Johansson-Fua, S. (Eds.), *Of Waves, Winds and Wonderful Things: A Decade of Rethinking Pacific Education*. University of the South Pacific Press, Suva, Fiji, pp. 148–160.
- Sanga, K., Reynolds, M., 2017. To know more of what it is and what it is not: Pacific research on the move. *Pac. Dyn.* 1 (2), 199–204.
- Sanga, K., Reynolds, M., 2019. Melanesian tok stori in leadership development: ontological and relational implications for donor-funded programmes in the Western Pacific. *Int. Educ. J.* 17 (4), 11–26.
- Sanga, K., Reynolds, M., 2020. Knowledge guardianship, custodianship and ethics: a Melanesian perspective. *AlterNative* 2 (16), 99–107.
- Sanga, K., Reynolds, M., 2021. Bringing research back home: exploring Indigenous Melanesian tok stori as ontology. *AlterNative* 17 (4), 532–542.
- Sanga, K., Walker, K., 2012. The Malaitan mind and teamship: implications of Indigenous knowledge for team development and performance. *Int. J. Knowl. Cult. Change Manag.* 11 (6), 223–235.
- Sanga, K., Reynolds, M., Paulsen, I., Spratt, R., Maneipuri, J., 2018. A tok stori about tok stori: Melanesian relationality in action as research, leadership and scholarship. *Global Comp. Educ.* 2 (1), 3–19.
- Sanga, K., Maebuta, J., Johansson-Fua, S., Reynolds, M., 2020. Re-thinking contextualisation in Solomon Islands school leadership professional learning and development. *Pac. Dyn.* 4 (1), 17–29.
- Sanga, K., Reynolds, M., Houma, S., Maebuta, J., 2021. Tok stori as pedagogy: an approach to school leadership education in Solomon Islands. *Aust. J. Indig. Educ.* 50 (2), 377–384.
- van der Ploeg, J., Sukulu, M., Govan, H., Minter, T., Eriksson, H., 2020. Sinking islands, drowned logic: climate change and community-based adaptation discourses in Solomon Islands. *Sustainability* 12 (17), 1–23.
- Von Der Porten, S., 2012. Canadian indigenous governance literature: a review. *AlterNative* 8 (1), 1–14.
- Waters, A., 2005. Indigeneity, self-determination and sovereignty. In: Hocking, B. (Ed.), *Unfinished Constitutional Business?: Rethinking Indigenous Self-Determination*. Aboriginal Studies Press, Canberra, Australia, pp. 190–209.

New frontiers in Pacific education, leadership and decolonized practice within Oceania: an initial exploration

Tanya Wendt Samu*, School of Critical Studies, Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	159
The overall context and positionality	160
Positioning myself	160
Locating Pacific/Pasifika/Pacific-heritage populations	160
Aotearoa New Zealand	161
United States	161
Conceptual underpinnings	162
Conceptualizing Oceania and taking an oceanic perspective	162
Conceptualizing an oceanic identity	162
A conceptual challenge from the past	163
The inquiry and its approach	163
Critical Exploratory (CE) questions	163
The selected education initiatives	164
Mana Academy (West Valley, Utah, USA)	164
Rise Up Academy, Rise Up Trust (Mangere East, South Auckland, New Zealand)	165
Mana Moana (National Leadership Program, Aotearoa New Zealand)	166
The National Pacific American Leadership Institute or NAPALI (National Leadership Program, Hawaii, USA)	167
Key insights, preliminary conclusions	167
References	170

Introduction

Between 2014 through to 2019, I had a series of seemingly unrelated experiences in the professional, personal and academic parts of my life. I made certain observations, and drew unexpected connections and overlaps across the (seemingly) disparate parts. I have attempted to carry out an initial, exploratory analysis of these experiences and observations in order to render them as educative and sources of learning and insight. The global pandemic that began in 2019 paused the process. The opportunity to write this chapter gently reactivated and accelerated this exploratory study.

In 2014, my attention was drawn beyond Aotearoa New Zealand (where I live and work); beyond the island nations and countries of the Pacific Region (which includes Samoa, my birthplace) to several education initiatives led by Pacific-heritage peoples, for Pacific-heritage learners within pockets of Pacific-heritage communities in Australia, the United States and New Zealand. I noticed that each of these education initiatives fore-grounded the valued knowledges and languages of specific Pacific-heritage home islands. I noticed that these initiatives were led by professional, urban-based Pacific-heritage Gen Xers (and a sprinkling of millennials) espousing strong commitment to transformation and change for their communities. Most of all, I noted that none of these initiatives were instigated by the state or local authorities. Rather they were grass roots in origin and subsequent development.

I began to tentatively theorize the possibility of a new conversation, one that could weave such initiatives together, focusing on the education, development and engagement of Pacific-heritage peoples, and leadership within the Pacific diaspora. Conversations about education and leadership within Pacific-heritage communities within metropolitan cities in Australia, the United States and New Zealand are not new—however, connecting such conversations together, and theorizing them across and within an existing network of relations IS (arguably) new. One of the earliest conceptualizations of this Pacific transnational space is Oceania, developed by Pacific elder-scholars Albert Wendt and Epeli Hau'ofa (Wendt, 1976; Hau'ofa, 1993). It conceptually brings together Pacific peoples located in within the Pacific and Pacific peoples with migrant histories in Pacific Rim nations such as Australia, New Zealand and the United States. There is no binary with Oceania. Samu (2010, p. 4) explained Hau'ofa's (1993, p. 6) conception of "world enlargement", a process carried out by Pacific families, as follows:

*Tanya Wendt Samu is a Senior Lecturer in Education at Waipapa Taumata Rau, the University of Auckland, New Zealand. Tanya's research and scholarship has focused on Pacific/Pasifika education and Pacific research in education, in addition to curriculum development (in social studies, social sciences), for over twenty years.

Pacific people map their worlds in terms of the location of their extended families. These maps frequently include New Zealand, Australia and the west coast of the United States of America, where parents, aunts and uncles, brothers, sisters and cousins reside. Modern technological advances in telecommunications and air travel have helped strengthen these familial networks. Traditional socio-cultural obligations are maintained by the flow of remittances, generally from the Pacific rim homelands of Pacific migrant communities to their Pacific country of origin, via the convenience of electronic transfers. There is also a counter-flow of cultural wealth such as tapa (traditional cloth made from mulberry tree bark), tivaevae (traditional quilts) and ie toga (fine mats) which are an intrinsic part of life events, when family members travel to visit one another for weddings, funerals, and church openings.

In framing this initial exploratory study, I have selected two education initiatives (one in the United States, one in New Zealand) and two leadership initiatives (again, one in the United State and one in New Zealand).

The rest of this chapter is organized into five sections. The first describes the context of the study and my positionality. The second presents the conceptual underpinnings. The third details the inquiry that was undertaken and the four key questions that inform and guide the exploratory study, while the fourth section discusses key insights that emerged.

The overall context and positionality

Positioning myself

I live in Auckland, Aotearoa New Zealand. My father is Samoan, my mother is Maori, Ngati Kahungunu to be precise. From one side, I am a migrant from Samoa. My children, nieces and nephews were born and raised in Aotearoa New Zealand. From the other side, I can trace my whakapapa or genealogy in Aotearoa back two generations before the waka that brought my ancestors to its shores. My connections to Aotearoa are deep. My legal connections to New Zealand are more recent in comparison. I was awarded citizenship when I was 11 years old, and I hold a current New Zealand passport.

Much of my academic career (spanning twenty five years) has focused on the education and development of Pacific-heritage peoples residing within Aotearoa New Zealand. I liken my teaching, research and service in the field that I have referred to as Pasifika education (Samu, 2010) as an all-encompassing personal, professional and intellectual fusion of educative experiences. These experiences have not necessarily been part of a systematic, orderly, sequential, intentional career program. I doubt such career trajectories are possible for many of us who were teachers and social workers before moving into the academy to be teacher educators, social work and counseling educators, academics and scholars.

Locating Pacific/Pasifika/Pacific-heritage populations

Oceania as a conception “sees” no international political boundaries. However,

please note that I will be using Pacific-heritage, Pasifika and Pacific interchangeably to refer to those who identify with one or more cultures indigenous to the island countries of the Pacific and are located within a Pacific Rim nation. I do not include New Zealand and Australia as Pacific countries or nations. Rather, I position them as Pacific Rim nations, along with the United States. This is an admittedly blunt typology, but Pacific Rim nation reflects the relative economic and political position Pacific Rim nations have compared to Pacific island nations. Pacific-heritage populations are minority populations within Pacific Rim nations. For example, a key feature of the Pacific population of New Zealand is its size relative to the total population. In 2018, the Pacific population in New Zealand was 8.1%, with 381,642 peoples. This is projected to increase significantly (Ministry of Pacific Peoples or MPP, 2020, p. 10).

However, the Pacific-heritage populations of the United States and Australia are considerably smaller as a proportion of the national population compared to New Zealand. For example, in the United States, “roughly 690,000 people identified as Native Hawaiian or other Pacific Islander alone but almost 900,000 identified as NHPI in combination with another race” according to the US Census Bureau,¹ drawing on 2020 census data. The total population of the United States was 331.4 million. According to Parliament of Australia,² the total population in the 2021 census was 25.4 million. According to Bately (2017), those claiming “Pacific Islands ancestry (excluding Maori)” were 0.88% of the total population in the 2016 census. Information about the proportion of the Australian population that identified as such for the 2021 census was not available at the time of writing this chapter.

From a state level perspective (for example, Queensland, Australia, or California in the United States) Pacific peoples are proportionally, very small minority in terms of numbers. From the perspective of metropolitan areas or cities, such as Brisbane, Queensland; or Los Angeles, California, Pacific peoples become more visible, but are still small, proportionally. However, it is in terms of suburbs or cities within mega-cities, where we find “pockets of Pacific Islanders” (the term Pacific Islanders appears to be used by institutions in the United States and Australia). It is at this level, where Pacific communities are much more visible and known to local government agencies and services such as schools and housing concentration. Examples include the city of Logan, part of the Brisbane metropolitan area. Logan has been described as “the heartland of Australia’s Pacific Islander population” (Chenoweth,

¹<https://www.census.gov/library/stories/2022/05/aanhpi-population-diverse-geographically-dispersed.html#:~:text=There%20are%2020.6%20million%20people,according%20to%20the%202020%20Census.>

²https://www.apb.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/FlagPost/2022/June/2021_Census#:~:text=On%20Census%20night%202021%20Australia,Australians%20is%20still%2038%20years.

2014). West Valley in Utah, the United States has been described as “the first large Utah city to be ‘minority majority’ where minorities combine to make up a majority of the population” (Davidson, 2018). 2.7% of the almost 137,000 population of West Valley identified as Pacific Islander. Given that the four selected education initiatives for this study are located in New Zealand and the United States, more contextual details will now be presented.

Aotearoa New Zealand

The formal, institutional name for Pacific heritage peoples in Aotearoa New Zealand has changed several times over the last 70 years. From Pacific Islanders, Pacific Polynesians, Pasifika, to simply Pacific (the current term of reference for many government departments), there has been the tendency to group together those who identified culturally with one or more Pacific nations (Samu, 2015). Prior to the mid-1990s, the underlying assumption of such terms was that the group was homogenous (Samu, 2020), and therefore could be responded to as one simplistic, unified block of sameness. Later in the 1990s, such umbrella terms acknowledged the diversity of languages, cultures, worldviews and histories within. Recognition of more nuanced and contemporary expressions of being “Pacific”, related to being urban, and homegrown in Aotearoa New Zealand, also began to emerge at the same time (Samu, 2015).

In a recent report (MPP, 2020), the Ministry of Pacific Peoples identified four key characteristics of Pacific population of New Zealand: it is young (with a median age of 23.4 years); growing (it grew by 29% between Census 2011 and Census 2018); diverse (in terms of cultural or ethnic identity; birthplace; settlement histories and regional distribution within New Zealand) and culturally rich (it includes people who identify with one or more of seventeen cultural groups, the main groups being: Samoan, Cook Island, Tongan, Niuean, Fijian, Tokelau, Tuvalu, I-Kiribati and Rotuman).

There are certain implications for Aotearoa New Zealand, as a consequence of this population profile. Most of the growing Pacific population is concentrated in specific suburbs or electorates. They represent voting power. The combined Pacific and Maori population represent almost a quarter of the total population of the nation. Both populations are youthful, with median ages in early twenties, and higher birthrates. Given the aging profile of the dominant European heritage population, these stats have implications for national level planning and provision in the employment; education; housing and health.

The cultural wealth of Pacific peoples is considered by many government agencies to be integral to developing appropriate, contextualized services for Pacific peoples in health, education, and social development, as evidenced by national level strategic plans with Pacific-specific targeting. Examples include the *Ministry of Education’s Action Plan in Pacific Education 2020–2030* (2020) and the *Ministry of Health’s Ola Manuia: The Pacific Health and Wellbeing Action Plan 2020–2025* (2020). The current position of Pacific peoples within New Zealand (developed over the past five decades) is best captured by the following somewhat aspirational statement within the Ministry of Pacific Peoples’ Pacific Aotearoa Lalanga Fou Report (2018, p. 5):

Our story as Pacific peoples in Aotearoa is one that is both familiar and ever evolving, in response to the changing currents of the modern world. Waves of migration have laid a foundation for our Pacific story; a story that involves aspirations to be a confident, thriving, prosperous and resilient Pacific Aotearoa. Our strong ties with tangata whenua¹ through our whakapapa², and shared moana³ help strengthen our identity in Aotearoa. Today, Pacific communities are recognised for the diversity they bring, the knowledge they impart and the contribution they make to the uniqueness of Aotearoa.

United States

The United States Census Bureau³ uses the term “Other Pacific Islander” with sub-groups including: Polynesian (Samoan, Tongan, Other Polynesian); Micronesian (Chamorro, Marshallese, Other Micronesian) and Melanesian (Fijian, Other Melanesian). It seems that usual categorization is Native Hawaiian and Other Pacific Islander or NHPI. The Bureau gathers data on NHPI and subgroups as well as NHPI “in combination with other race”. In the 2020 census, the NHPI population was 1.586 million with 689,966 identifying as NHPI alone, and 896,497 NHPI in combination.

Historically, Asians and Pacific Islanders have been grouped together by government classifications, and at some levels of government and some institutions, this usual practice. So it is important to bear this in mind when looking for and encountering data relating to the Pacific-heritage population within the United States. In terms of location, just over half of NHPI are located in Hawaii and California. The states with the next highest Native Hawaiian/Pacific Islander populations are Washington, Texas, Florida, and Utah. At the level of individual states, this population is still quite a small minority group. In the 2010 census, for example, there were approximately 37,000 NHPI in the state of Utah. The total population of Utah was 3.2 million. Smaller population sizes nation-wide and state-wide positions Pacific Islander populations to vie and compete for additional funding and resourcing, if needed by specific Pacific communities. One can assume that much more lobbying and community organization and activism is required in such settings and contexts (compared to the Pacific population within New Zealand).

³<https://www.census.gov/library/stories/2022/05/aanhpi-population-diverse-geographically-dispersed.html#:~:text=There%20are%2020.6%20million%20people,according%20to%20the%202020%20Census.>

Conceptual underpinnings

When radical post-colonial writers and artists from the Pacific were challenging cultural canons, and providing counter-hegemonic conceptions in the 1970s (for example, Albert Wendt, 1976) and early 1990s (for example, Epeli Hau'ofa), they focused on Pacific peoples living *within* the Pacific—and not on its *margins*, as in the Pacific Rim (Coxon, 2011). It is here that comparatively large communities of Pacific-heritage peoples have become established in certain suburbs or cities of metropolitan cities such as Auckland, Sydney, Los Angeles, Honolulu. These communities are a consequence of immigration from island nations such as Samoa, American Samoa and Tonga which began in the late 1960s and continued across the 1970s. Vibrant first, second and even third generation Pacific communities have since emerged within these metropolitan societies. This chapter collectively refers to them as the Pacific diaspora, recognizing global processes (such as labor migration) and inherently Oceanic processes (such as Hau'ofa's world enlargement) that have shaped contemporary Pacific peoples socio-economic positioning, relational connections and cultural identities across time and space.

Conceptualizing Oceania and taking an oceanic perspective

In his seminal essay "The New Oceania", published in 1976, Samoan writer and scholar Albert Wendt, described the provocative body of writing and artwork that was emerging within the Pacific Region at the time as ".... a revolt against the hypocritical/exploitative aspects of our traditional/commercial and religious hierarchies, colonialism and neo-colonialism and the degrading values being imposed from outside and by some elements in our own societies" (p. 59). His conceptualization of a "New Oceania" included counter-hegemonic features such as the importance of individual dissent. He warned that "without it our cultures would drown in self-love" (Wendt, 1976, p. 52). He advocated, not for the revival of past cultures but "for the creation of new cultures which are free of that taint of colonialism and based firmly on our own pasts" (p. 53). In exploring ideas such as Wendt's and expanding them and reframing them to education contexts and settings, it would seem that in taking an Oceanic perspective of Pacific or Pasifika education (Samu, 2010), facilitates new forms of expression and identity, and encompass and absorb creative 21st century approaches to resolving issues and concerns of Pacific peoples and their education in the 21st century—such as in Aotearoa New Zealand, and the United States.

Conceptualizing an oceanic identity

A poignant piece of prose that provides the concept of Oceania with life, is a quote by elder-scholar Teresia Teaiwa. According to her close friends (Teaiwa et al., 2021, p. xv)

As the late Professor Epeli Hau'ofa was finishing his essay 'The Ocean in Us' (1998), he decided he needed an epigraph, something wise and poetic to underscore his argument for grounding regional identities in the Pacific Ocean ... Hau'ofa asked a young poet, scholar, activist, and well-known wordsmith for a line he could use. She immediately offered her now most widely circulated quote:

We sweat and cry salt water, so we know the ocean is really in our blood

Hau'ofa and Teaiwa presented and promoted the value of developing a strong regional identity for Pacific peoples. This was via the shared inheritance of the Pacific Ocean, or the Moana. A strong shared identity as Pacific Islanders/Pacific peoples/Pasifika/Moana Peoples/Oceania involves ancestral connections to the actual Pacific Ocean—and connects people in Pacific island nations, and Pacific peoples within Pacific Rim nations such as New Zealand, Australia, the United States and Canada. Scholars of Pacific Studies as a multi-disciplinary field have been strengthening such connections, and such identities across the diaspora and into the Pacific region via symposia and conferences organized and hosted by universities for many years. For example the Pacific History Association⁴ has held its biennial conference for more than 24 years. The University of Hawaii at Manoa's Center for Pacific Island Studies⁵ was established more than 70 years ago. The University of Utah⁶ organized a two-day virtual symposium in 2021 entitled "Pedagogies for Indigeneity and Diaspora: Pacific Studies at Home and Abroad". Comprehensive, in-depth transnational research studies of the Pacific and Pacific diaspora and the sharing of new knowledge and perspectives demonstrate that critically important conversations are occurring at that level. I argue that something similar should be nurtured and developed in terms of specific conversations across the Pacific diaspora about education and leadership, particularly education initiatives for and by Pacific peoples.

⁴<https://hawaii.edu/cpis/pha-24th-biennial-conference/>.

⁵<https://hawaii.edu/cpis/2010-celebrating-connections-60-years-of-pacific-studies/>.

⁶<https://transform.utah.edu/pi-studies/symposium/>.

A conceptual challenge from the past

In 1985, Tongan scholar and philosopher, Epeli Hau'ofa issued a thought-provoking challenge to the “educated, relatively wealthy urban-based local elites ... locally instrumental in deciding the directions of change” (p. 163). He considered them paradoxically, as both a potential threat to the survival of “the indigenous elements of traditions” and potential change agents drawing on “the rich heritage from the past” (p. 166) to produce new images, new voices for the Pacific. The spatial focus for his writing was undoubtedly nations of Pacific the Pacific region. However, his essay also reflected the growing sense of political, economic and cultural connection as Pacific peoples “toward” the rest of the world, including recognition of negative external influences, particularly those from powerful and influential colonial/former colonial nations of the Pacific Rim such as New Zealand, the USA and Australia. That was then (the mid 1980s). What about now?

I believe his challenge is pertinent for very different 21st century metropolitan contexts and settings—within which “new”, quite different Pacific peoples reside, where many endeavor to survive, and not as many, regrettably, thrive. State provided education is one area which is problematic. For example, according to American Tongan activist and educator ‘Anapesi Kaili, speaking in 2012 about Pacific Islander students within the United States:

We have criticized them for not knowing their language and culture, yet we don't take the time to teach it. I have been in numerous meetings where administrators and teachers clearly do not see a need for multiple histories or epistemologies or anything that is not in par with mainstream culture and values. Yet, the minute our young people show any sign of resistance in this mainstream classroom they are labelled as having a behaviour disorder, tracked into ESL and special education courses and their Pacific cultures blamed for their so-called failures. The same culture that they are not allowed to practice or even emulate. The same culture that they hardly even know!

The inquiry and its approach

Elder -scholar Seu'ula Johansson-Fua has written simply and clearly about the value and importance of drawing on indigenous knowledges and perspectives (Johansson-Fua, 2014, p. 66) when undertaking research. She stated:

For me, as the apprentice, it was an exciting process of re-thinking about education, and about being Tongan. It was my own personal re-education as I worked with these two academics in search of a research framework that would capture the reality and truth about our world view. The Kakala research framework was a space that would allow us to be who we are, with all our insights, knowledge, experiences and inherited gifts and to position ourselves where we belong without shame or pretence. It is from this position as a researcher that we can do justice to our world view and to present an accurate account of what and who we really are.

Her wisdom resonates with me. Drawing on the concepts and ideas of Pacific elder-scholars in my search for meaning of the experiences I briefly described at in the Introduction of this chapter, and the subsequent thinking and writing involved has had a similar effect on me. I describe the personal re-education, during the process of sense-making, as the need to capture reality and “to be who we really are”.

Critical Exploratory (CE) questions

Four questions were applied to the exploration of the websites and public FaceBook pages of the four education initiatives that I selected. Three of these questions came from Hawaiian scholar Benham (2006, p. 45) who posed them when she was speaking over a decade ago to Native Hawaiian and Pacific Islander scholars about their research. I refer to them as Critical Exploratory Questions or CEQ.

Critical Exploratory Question 1:

Does it meet a high standard of social justice that ensures freedom to self-determine and plan for future progress?

Critical Exploratory Question 2:

Does it ask important questions that move us and others to transform our thinking and generate new ways of viewing learning and teaching that make a difference?

Critical Exploratory Question 3:

Does it invite to the conversation voices of cultural experts, elders, families and communities?

I argue that these questions reflect Wendt's views (1976) that our efforts must be counter-hegemonic; must recognize multiple Pacific worldviews; and must draw on indigenous epistemologies.

I add a fourth question, one posed by Haig-Brown (2012, p 74). She argued that such a question is necessary for the long process of decolonizing our countries and our lives, even when “our country” is a nation in which we or our parents or even grandparents settled. Her argument is about the need to educate ourselves, as diaspora, about the lands, histories and cultures of First Nations peoples, or Tangata Whenua.

Critical Exploratory Question 4:

“Whose traditional land are you on?”

The selected education initiatives

A distinguishing feature of these initiatives is that they are led by professional, urban-based, well-educated, bicultural, bilingual, relatively young members of their respective “local” Pacific communities (I define “young” in this instance in a culturally subjective sense—it means younger than I am by at least a decade and a half). Examples of local communities include West Valley, Utah in which Mana Academy is located; and Mangere East, south Auckland in which Rise Up Academy is located. Auckland is the main site for the Mana Moana Leadership program, which has national scope. Hawaii is the site for the National Pacific American Leadership Institute’s national leadership program.

Mana Academy (West Valley, Utah, USA)

Mana Academy is the first Pacific focused Charter school to be established in the world. It offers full primary schooling, and secondary schooling up to Grade 11, and is located in West Valley, in the state of Utah, the United States. Established in 2013, plans are currently underway to mark its ten year anniversary in April, 2023. In the academic year that ended in June of 2022, Mana had a student enrollment of 313 students. 49% of student enrollment identified as Native Hawaiian or Other Pacific Islander (NHPI) and 43.4% identified as Hispanic/Latino. Half the student population is defined by the state as being economically disadvantaged, and 97.1% as minority enrollment.⁷

According to its website, Mana Academy’s Mission is to “develop scholars who are rooted in their native cultures and are prepared to be the leaders of tomorrow”. Mana Academy’s Vision⁸ is as follows:

Mana Academy Charter School is a *community and family centred* school committed to *educational excellence* and *cultural sustainability*. Our administration, teachers and parents *work together as partners* to enhance the academic achievement and learning of all students. We provide a learning environment that is *culturally relevant and inclusive*. We strive to meet the needs of all learners by designing academically challenging work that is meaningful, engaging and inspire innovation, exploration, discovery and creativity. We empower all students to access, enrol and graduate from college.

All students who graduate from Mana Academy will be equipped with the necessary knowledge and skills to be successful leaders and positive agents of change in their local, regional and global communities (emphasis added).

Mana Academy Charter School’s Facebook public page features school activities and celebrations for secondary students, marking the end of the 2022 academic year (emphasis added).

From June 13, 2022:	• Our Mana Academy class of 2022 Salutatorian ... performing her final <i>Tongan tau’olunga</i> ... at our graduation ceremony • Our secondary campus scholars performing a <i>Samoaan sāsā</i> choreographed by Mana alumna and teacher ... This sāsā tells of the different legends of Polynesian Demi-God Maui when he fished up the islands, lifted the sky, snared the sun and brought fire to his people
From 11/6:	• Our class of 2022 have all been accepted to the college/university of their choice, <i>received full ride academic scholarships</i> and will begin their college career this coming fall.
From 1/6:	• Our elementary scholars <i>singing Nga Iwi e, composed by The late Hirini Melbourne, a Maori songwriter</i> and poet joined by a performance by some of our 3rd and 4th grade scholars. • Our high school scholars <i>performing a Norteno tribal</i> choreographed by our 10th grade scholar ... Norteno Tribal is native to the Mexican states of Sonora, Chihuahua, Coahuila and Nuevo Leon. The Norteno was developed from blending of the Mexican and Spanish musical traditions, military band instrumentation and Germanic musical styles such as polka and the waltz

I have selected these recent posts in order to provide insight into what “cultural sustainability”, “culturally relevant and inclusive” as well as “academic excellence” seems to look like for Mana Academic Charter School.

In addition, I share a post from the principal’s FB page. In 2018, I asked for her permission to use a specific photograph and her description of it for a presentation. The photo was from a visit to Mana Academy by His Highness Tui Atua Tupua Tamasese Ta’isi Tupuola Tupuola Efi and his wife, Masiofo Filifilia Imo, of Samoa. Tui Atua is a highly esteemed Samoan indigenous knowledge holder, with an international reputation as an academic scholar. The principal wrote the following (emphasis added):

Today our elementary scholars presented traditional gifts to His Highness Tui Atua Tupua Tamasese Efi and Her Highness Masiofo Filifilia Tamasese. We presented our gifts in the order of our migration to Utah.

⁷<https://www.usnews.com/education/k12/utah/mana-academy-charter-school-147712>.

⁸<https://www.themanaacademy.org/about>.

Our first presenters were our Navajo scholars representing the indigenous people of the land, followed by our scholars from Mexico, representing our Latinx community, then our Hawaiian scholars, Samoan scholars, Tongan scholars, Maori scholars and finishing up with our Marshallese scholars.

They were so moved by the fact that all gift presentations by our scholars were done in their native languages.

I took the Critical Exploratory Questions (CEQ) and applied them broadly.

CEQ1: Cultural graduation celebrations signal “freedom to self-determine and plan”, given the highly charged issue of high school bans on cultural attire during graduation ceremonies in other parts of Utah state. It made news via Maori media⁹ in New Zealand in June of 2022, when a young Samoan student made a formal appeal in the Jordan District of Utah, to overturn a by-law passed in December of 2021, forbidding high school graduates from wearing cultural items on top of their graduation regalia.

CEQ2 & CEQ3: Mana Academy offers a school curriculum in which all students are exposed to and learn from and about, the cultural groups represented within the school (Tongan, Hawaiian, Samoan, Marshallese, Hispanic and Navajo). Its website profiles staff, including a Culture Team of cultural teachers and mentors.

CEQ4: Learning about the indigenous peoples of their location is an explicit part of their school curriculum regardless of whether or not any students identify with the specific tribes whose lands the school and surrounding community are located on.

Rise Up Academy,¹⁰ Rise Up Trust¹¹ (Mangere East, South Auckland, New Zealand)

Rise Up Academy began in 2014 as a full primary Charter school via education policy changes instituted by a National government. It was founded by Rise Up Trust, a not-for-profit organization established in 2006. Rise Up Academy transitioned into a Designated Character School in 2019 (under a Labor government). Rise Up Academy’s website states that their special character consists of “A Hearts and Minds based Inquiry Learning model, which reflects Maori and Pasifika values. When we enroll your child, we enroll your whanau. Culturally responsive Whanau mentoring programmes.” Rise Up describes its vision and mission as the development of “SHARP minds, STRONG bodies and GOOD hearts by RAISING Leaders, one child at a time”. It states:

This acknowledges and reflects our shared Pasifika values – Mark 10:27. “Love the Lord your God with all your heart and with all your soul and with all your strength and with all your mind and, Love your neighbour as yourself”.

The whanau or family mentoring programmes are an important feature of Rise Up Academy. These programmes are also a requirement of parents/caregivers who enroll their children. Take, for example, the eight week Hearts and Mind course. It equips parents with the keys to unlock their child’s potential and engage with their learning. It is intended to enable parents to “feel empowered as they understand the connection between their child’s learning style, love languages and personality type and the best way to develop the child’s thinking and problem solving skills”. Another example of a key whanau program is Synergy, a “family program involving three way learning, where Whanau Educators take parents and their children through real life problem solving strategies. Through the Inquiry Learning Model and real life scenario the whanau plan a presentation about their child, which is presented by the child on the last evening.” Culture and identity is one of the seven stated objectives of the Synergy program. This program runs for seven weeks, in weekly 1.5 h sessions that occur in the evenings after work.

In her welcome statement on their website, the CEO of Rise Up Trust (who is also the principal of Rise Up Academy) stated the following:

For 15 years Rise UP Trust has partnered with our community to celebrate Pasifika uniqueness. With collective strength we have challenged, changed and lifted educational achievement and well-being in our community.

As a premium provider of evidenced based whanau engagement we have grown from a garage home-school, established The Rise UP Academy – Designated Character School and continue to address intergenerational disadvantage through the Building Learning Communities project.

In partnership with The Foundation North we are committed to a 10 year longitudinal study to track the learning journeys of a focus group within our community.

⁹<https://www.teaomaori.news/utah-polynesian-community-wins-appeal-wear-cultural-attire-graduation>.

¹⁰<https://www.riseupacademy.school.nz/building-learning-communities/>.

¹¹<https://www.riseuptrust.org.nz/>.

There is a Learning revolution going on in South Auckland and we welcome your support or expressions of interest in our programmes.

Rise Up Academy is the only one education initiative in this exploratory study that has been involved in sustained evidence based research studies of its services and practices. At least three quality assured academic publications have been produced from research activities to date (the most recent being [Sharp et al., 2020](#)).

After a broad application of the Critical Exploratory (or CE) questions:

CEQ1: Rise Up's well-documented history from an informal Saturday school in a family garage run by two Pacific primary school teachers, to a community trust to a school reflects the determination of its leadership to achieve a specific vision for Pacific learners and their families in south Auckland, in local board areas of Mangere-Otahuhu; Otara-Papatoetoe and Manurewa. Rise Up's development has been astute and strategic, drawing on one specific philanthropic source that shared their vision, collaborated with them to build their capacities and capabilities in alignment with the series of goals they set for themselves.

CEQ2 & CEQ3: The primary focus for Rise Up was on building up parents, to empower them in their efforts to effectively support their children's education, and strengthen their families. When the Trust established the school, parents were required to commit specific hours to serve and support the school (reciprocity) and to take the courses. Home-school partnership takes on different meaning when comparing Rise Up Academy to most "mainstream" schools in the New Zealand education system.

CEQ4: On the basis of the website scan or brushstroke, this CE question is not strongly evident. This does not mean it is absent, or goes unaddressed. Foregrounding matauranga Maori in relationship to Rise Up as a school is not evident within its online presence.

Mana Moana¹² (National Leadership Program, Aotearoa New Zealand)

Leadership New Zealand¹³ was established in 2003, and its overall purpose is "to create a connected community of leaders working together for a more just, equitable and sustainable Aotearoa". The Mana Moana Experience was launched in 2019 as one of four leadership development programmes offered by Leadership New Zealand. The website describes the Mana Moana Experience as an integrated program for Pasifika leaders. It "brings together the experience, expertise and networks of Leadership New Zealand and a research-based indigenous knowledge-derived specialist program created out of the post-doctoral research" of the Mana Moana Experience program director.

Mana Moana is designed for mid-career Pasifika leaders. It involves six three day retreats over a period of eight months. The program involves peer coaching and small group contact between whole-cohort retreats. Each retreat focuses on "a significant broad issue/theme that is critical and relevant to key issues facing the Pasifika community and New Zealand". Other program features include:

- the vitalization and mobilization of indigenous Pasifika languages, knowledge, values, culture, ways of knowing and viewing the world.
- the understanding that Pasifika ancestral cultural resources are fundamental and essential for optimal contemporary living and leadership.
- to succeed in Western contexts, Pasifika cultural resources are seen to be a source of advantage, pride, competitive edge, x-factor, complexity, cognitive flexibility, neuroplasticity, polycultural capital and a way of reaping the diversity dividend.

"This is leadership for a new era, drawing on ancient indigenous resources".

In order to capture a glimpse of the possible nature and tenor of aspects of the Mana Moana Experience, I invite readers to peruse "Our Ancestors Speak", described as a powerful short film work that is a call to arms for people across the Pacific and globally. Mana Moana -Pacific Voices¹⁴ is not directly associated with Leadership New Zealand's Mana Moana Experience. However its program director has curated this "... collection of emotive and artistic video works that have been created to amplify and support the Pacific drive to global action to reduce greenhouse emissions".

Based on a broad application of the Critical Exploratory (or CE) questions:

CEQ1: The focus on the development of an explicitly Pacific leadership, in which Pasifika cultural resources are more than sources of self-esteem and pride, but are identified and used as capital for good leadership and success seems to explicitly contribute to self-determination and planning for "future progress".

CEQ2: From the description of the program, and the participants' immersion within an indigenous framework for learning, it seems the potential for new learnings that "make a difference" is high.

CEQ3: This appears to be a program that includes cultural knowledge holders and experts and elders—but perhaps not families, given the residential retreat nature of the activities.

¹²<https://www.leadershipnz.co.nz/mana-moana>.

¹³<https://www.leadershipnz.co.nz/about/about-us>.

¹⁴<https://www.manamoana.co.nz/pacificvoices/>.

CEQ4: Learning about the relationship between Maori and Pacific peoples in Aotearoa New Zealand is deeply embedded in the program design and is an explicit part of the learning experience.

The National Pacific American Leadership Institute¹⁵ or NAPALI (National Leadership Program, Hawaii, USA)

NAPALI is a “non-profit tax exempt organization” that has been running since 1998. It provides Native Hawaiian and Pacific Islanders the opportunity to learn about Servant Leadership and to harness their own cultural heritage to be a better leaders. Their Mission statement states the following:

The Mission of the National Pacific American Leadership Institute (NAPALI) is to identify, train, and nurture Pacific American Emerging Leaders. NAPALI is unique in that it conducts the only leadership program in the nation that integrates the leadership values of Pacific American cultures and Western Society.

It is committed to making a difference in the quality of life of Americans who trace their ancestry to the indigenous peoples of the Pacific islands - American Samoans, Chamorros, Fijians, Maoris, Native Hawaiians, Tongans, Tahitians, and others connected with Pacific American cultures.

Over the 24 years since it began, 310 Fellows have graduated from the program. Here are extracts I have selected from the website to provide insight to the nature of the program that NAPALI offers (emphasis added).

- NAPALI ... educates and trains emerging Pacific American leaders to build upon *traditional cultural values* in understanding, harmonizing and applying leadership theories and skills necessary for achieving personal and professional growth, greater responsibilities, and a *higher level of service to our communities and country*.
- This course is based upon the belief that indigenous Pacific Americans have an established historical basis for their core values and norms, and that they will act in accordance with them as long as these values and norms are relevant in today's society.
- The course curriculum focuses on empowering participants by helping them *understand key cultural individual and collective* Pacific American values; how these values shape attitudes and perceptions; how they were expressed in leadership skills and practices; how they *fit into the dominant society*; analyze universal leadership attributes, skills and actions which increase potential for success; and *build bridges between cultural values and leadership skills for Pacific Americans and the dominant Western culture*.

The course consists of a one day multi-site session prior to a 78 h seminar conducted over a consecutive 8-day period in a setting conducive to learning, thinking and sharing.

After a broad application of the Critical Exploratory (or CE) questions:

CEQ1: The information on the NAPALI website is succinct, to the point. The broad, brushstroke application of this particular question did not yield anything particularly definitive other than an indicator that over the eight day course, career goal setting and personal development planning and presentation are intended activities.

CEQ2: A focus on specific cultural values occurs between Day 2 through to Day 6 of the residential program. It is interesting that while no specific details are provided, indigenous cultural activities are scheduled alongside activities that at first glance could be assumed to be the expected curricula for generic leadership development programmes (e.g., “public policy agendas” and “ethics and compliance”). The available photographs of previous groups show learning taking place outdoors in Hawaii—in taro patches, on board traditional Polynesian double hulled canoes with participants wearing clothing suited for active outdoor activities rather than corporate wear within conference centers.

CEQ3: Indigenous cultural elders or knowledge holders are described (and appear in photographs) as being an integral part of the program in addition to university based academics and scholars. Given that participants are selected from around the United States, and the residential or retreat nature of the eight day course, it not a program that can easily involve families.

CEQ4: I have inferred, based on the available text and photographs of the website, that participants experience indigenous Hawaiian culture, values and beliefs, which makes sense given the location of the retreat.

Key insights, preliminary conclusions

In some ways, the so-called new conversation I am calling for leveraged off an internal one that I began to have when I began to notice, observe certain education initiatives. Like Johansson-Fua (2014), I have been involved in an intriguing and at times exciting process of re-thinking about the education, development and engagement of Pacific peoples within the diaspora. The observations followed by aspects of this exploratory study has been a personal re-education for me.

I chose four to explore—two charter schools, two national Pacific leadership programmes, two different geographic settings, both Pacific diaspora. The others that caught my attention included a school to university pipeline project at Griffiths University

¹⁵<https://www.thenapali.org/>.

called the Griffith Maori and Pasifika LEAD¹⁶ (Legacy, Education, Achievement and Dream) Program, in Queensland Australia; Le Fetuao¹⁷ Samoan Language Center, a non-profit community organization in Honolulu, supporting young Samoan children to learn their heritage language and culture. It has strong partnerships of the University of Hawaii-Manoa's Samoan Language Program, churches and businesses. In 2015, I also came across an education program supporting Pacific and Aboriginal professional rugby league players prepare for life after professional sporting careers. A key partner with the Australian Rugby League Union was Australia National University or ANU (Canberra).

So, what have been key insights for me? I identify and share my reflections of three key learnings or sights.

(i) *Revise the Critical Exploratory Questions*

Perhaps the set of Critical Exploratory Questions that I "found" in my reading and reflections merely confirm the features that caught my attention, and their application confirms what I see as being of value—in terms of initiatives that are innovative; designed to respond to specific, contextualized contemporary needs identified by savvy, strategic minded Pacific community leaders and educators rather than state authorities; informed by indigenous Pacific values, cultures, ways of knowing and being; and implemented by Pacific leaders who have formed collaborations with universities, philanthropic organizations and other relevant entities.

If I were to explore further (in a subsequent study) I would reshape or adapt the CEQs as follows:

CEQ 1: This question might need to be reversed to "Does this initiative ensure freedom to self-determine and plan for future progress thereby ensuring a high standard of social justice?"

Or perhaps it should read "Does this initiative ensure the exercise of oceanic *agency*, thereby ensuring progression toward *equity* and *social justice*?"

CEQ 2: This question could also be reversed, to become "Does this initiative move those involved to transform their thinking; generate new ways of learning and teaching that make a measurable difference; and promote new questions?"

CEQ 3: I think this question could remain as it is currently.

CEQ 4: This question is far more important than I realized—so much so, that I have reflected on it in the next sub-section.

(ii) *The Importance of Decolonizing Pacific Perspectives of Place*

If a Pacific education or leadership initiative is explicit and purposeful about decolonizing leaders must take cognizance of the relationship between themselves as Pacific, and the indigenous peoples of the land. According to Haig-Brown:

I consider the possibility of decolonizing discourses of diaspora, by asking the central question not only where people of the diaspora come from, but *where have they come to?* (2012, p. 74, emphasis added).

In the context of Aotearoa New Zealand, according to Teaiwa and Mallon (2005), it is not uncommon in formal ceremonies for Pacific visitors to be welcomed and greeted by Maori as *tuakana*, or elder sibling. However, in the realities of contextualized encounters, relations can be one of "ambivalent kinship" (2005, p. 208), where depending on the specifics of the encounter or situation, there may be "a denial of kinship; or an assertion of seniority within a kinship model; or it may produce moments where the representation of a 'united happy family'" appears best. It is this kinship position that seems to be an appropriate inclusion for a reflective response to Haig-Brown's question. The Health Research Council (2014, p. 2) in New Zealand has what I consider to be a robust, astute answer:

Pacific peoples acknowledge, support and respect Te Tiriti o Waitangi as the foundation for their relationship with Tangata Whenua and recognise the Tangata Whenua status of New Zealand Maori and their right to exercise tino rangatiratanga. Pacific peoples also recognise the mandate accorded under citizenship rights and the entitlements that comply with international declarations and other legal instruments regarding the rights of human beings. Pacific peoples and Tangata Whenua are related and bonded culturally and geologically through the Moana Nui a Kiwa (greater Oceania kinship connections). This relationship is embedded in history, mythology, and cosmological beliefs, and transcends state boundaries and legal jurisdictions.

It means relevant, deeply thoughtful and authentic perspectives of the lands and histories of the people of the lands in which we as Pacific reside should be integral in the learning experiences that Pacific education initiatives design because they are part of our Pacific diasporic identities.

(iii) *Oceania as Relational Space*

According to Pacific elder-scholar Eve Coxon (2011), Epeli Hau'ofa's "influential vision" promoted "a revitalization of the pre-colonial historical interconnectedness" of Pacific peoples and "the development of extensive and expansive new connections with the countries of the Pacific Rim". She stated:

By reclaiming the term 'Oceania', he re-conceptualised the embodied life experiences of Pacific peoples by focusing not just on their small islands (their *places*) within the vast ocean, but on the varied *relationships* maintained over many centuries throughout the Pacific Ocean, their shared *space*, which

¹⁶<https://www.griffith.edu.au/student-support/outreach-and-engagement/pacific-islander-and-maori-student-outreach>.

¹⁷<http://www.lefetua.org/>.

more recently have given rise to new relationships with places on the Rim. Oceania is thus a *relational space*, a “much enlarged world of social networks that crisscross the ocean all the way from Australia and New Zealand in the southwest to the United States and Canada in the northeast” (p. 6)

It makes sense that the new connections she refers to are inclusive of places *within* the Rim nations, in which Pacific-heritage communities have settled, become established, and endeavor to become valued strands of the fabric of those societies.

It is specifically Pacific social networks, facilitated by 21st century technologies and modes of connection (social media platforms such as Facebook), that drew me in different ways to notice, observe grassroots initiatives in these outer reaches of Oceania. For example, when my younger sister and successful author of Pacific young adult fiction was invited to carry out a speaking tour in West Valley and Salt Lake City Utah in 2014, I joined her to provide big-sister moral support; to observe and meet community leaders and educators in this surprisingly sizable “pocket” of Pacific peoples that I had not been aware of until then.

This was how I first came to visit Mana Academy and meet the principal, deputy principal. It was through their social media posts that I became aware of the National Pacific American Leadership Institute (the Mana Academy principal is a Board member) and held initial discussions about it in subsequent interactions with Mana Academy leadership when they came on an education study tour to Auckland in 2016, and when I visited Mana Academy during my research and study leave in 2018. From 2014 through to the present, I became more involved with the American Research in Education Association (AERA) Indigenous Peoples of the Pacific Special Interest Group (IPP-SIG). Through IPP-SIG, I met and reached out to young emerging United States based Pacific doctoral scholars who turned out to be either NAPALI Fellows or were intending to apply to the program.

In 2016, I was formally approached to be the Pacific research advisor for a longitudinal project, Ngā Tau Nguahuru “Looking Beyond for Ten More Years”, a study into “success” as Maori and Pacific. It grew out of Foundation North’s Maori and Pacific Education Initiative 2009–2014. Foundation North is a major (in terms of scope and funding capacity) philanthropic organization, serving Auckland and Northland for over thirty years. This was how I became re-acquainted with Rise Up Academy. And then there’s my cousin who as then CEO of Leadership New Zealand, facilitated the establishment of the Mana Moana Leadership Experience. I began to follow its progress through her social media posts after its’ launch in 2019. Interestingly, the main funder for the Mana Moana Leadership program has been Foundation North.

What are the preliminary conclusions of this exploratory study? I identify three.

(i) *“When the answers are within us”*

Each initiative attests yet again to Pacific peoples capacities to draw on their resources and in collaboration with others, develop innovative strategies which hold the potential for meaningful change—particularly in the contexts, settings (such as the United States) where the initiatives rely far more heavily on Pacific peoples and their communities. We are fortunate, in Aotearoa New Zealand, that given the demographic profile of Pacific peoples, national education and social policies are in place that specifically target Pacific peoples because this population is a strategic priority. This has influenced both the private sector and the philanthropic sector, in terms of funding decisions.

(ii) *Counter-Hegemonic*

These initiatives are examples of education collaborations intended to be counter-hegemonic. They also represent efforts to decolonize the education of Pasifika learners. This is evident in the development and use of conceptual frameworks informed by indigenous knowledge and values; the foregrounding of heritage languages as assets for learning success and pedagogical responsiveness to the unique identities of the Pacific learners, their families and communities.

(iii) *The need to learn more*

To reiterate, this study was part of an exploratory, contemplative process of inquiry, one holding significant potential for deeper theorizing, more intensive examination. There is a case for the design and implementation of an over-arching, more comprehensive research project, one that examines rather than explores the strengths of these innovations and the outcomes for Pacific learners, participants and their families and communities.

Overall, this exploratory study draws attention to Pacific leadership, particularly the type of leadership that Hau’ofa believed was crucial for Pacific futures. I believe the leadership of the two quite different Pacific charter schools reflected that form of stewardship. The two different Pacific leadership programmes appear to focused on development of ‘that type’ of Pacific leadership. By virtue of their nature and structure, the two Pacific charter schools have developed incredible distinctive curriculum and learning programmes. They have responded to their respective visions for education of Pacific learners through deliberate purposeful design, the extent of which would be very challenging to achieve in “mainstream” state schooling.

I end this chapter with this quote by Benham (2006), which captures the essence of these education initiatives:

Keep this in mind: If it is not us, then who will do this and to what end?

Let us be the ones to answer the question:

¹⁸Benham (2006, p. 45).

'Where can we collectively be that is greater than where we are now?'¹⁸

References

- Benham, M.K.P., 2006. A challenge to Native Hawaiian and Pacific Islander scholars: what the research literature teaches us about our work. *Race Ethn. Educ.* 9 (1), 29–50.
- Bately, J., 2017. What does the 2016 Census reveal about Pacific Islands communities in Australia?. In: *State, Society and Governance in Melanesia Brief 2017/23*. https://bellschool.anu.edu.au/sites/default/files/publications/attachments/2017-09/ib_2017_23_bately_revised_final_0.pdf.
- Coxon, E., 2011. Inside and around the Pacific circle: educational places, spaces, and relationships. *Pac. Asian Educ.* 23 (2), 5–10.
- Chenoweth, L., 2014. Pacific Islanders and Education – Is Australia an “Unlucky Country”? *Oz Kiwi*. <https://ozkiwi2001.org/2014/02/pacific-islanders-and-education-is-australia-an-unlucky-country/>.
- Davidson, L., 2018. Census: West Valley City Is Now “Minority Majority” Where Minority Groups Outnumber White Population. *Salt Lake Tribune*, Sept 13th. <https://www.sltrib.com/news/politics/2018/09/13/census-west-valley-city/>.
- Haig-Brown, C., 2012. Decolonizing diaspora. In: *Decolonizing Philosophies of Education*. Sense Publishers, pp. 73–90.
- Hau'ofa, E., 1985. The future of our past. In: Kiste, R.C., Herr, R.A. (Eds.), *The Pacific Islands in the Year 2000*, pp. 151–169. Working Paper series [1985]. Honolulu, Hawaii: Pacific Islands Studies Program, Centre for Asian and Pacific Studies, University of Hawaii at Manoa.
- Hau'ofa, E., 1993. Our sea of islands. In: Waddell, N., Hau'ofa (Eds.), *A New Oceania. Rediscovering Our Sea of Islands*. School of Social and Economic Development, University of the South Pacific, Suva.
- Hau'ofa, E., 1998. The ocean in us. *Contemp. Pac.* 10 (2), 392–410. <http://www.jstor.org/stable/23706895>.
- Health Research Council, 2014. Pacific Research Guidelines. https://www.hrc.govt.nz/sites/default/files/2019-05/Resource%20Library%20PDF%20-%20Pacific%20Health%20Research%20Guidelines%202014_0.pdf.
- Johansson-Fua, Seu'ula, 2014. Kakala research framework: a Garland in celebration of a decade of rethinking education. In: Otunuku, M., Nabobo-Baba, U., Johansson-Fua, S.F. (Eds.), *Of Waves, Winds and Wonderful Things: A Decade of Rethinking Pacific Education*. https://www.usp.ac.fj/institute-of-education/wp-content/uploads/sites/132/2021/11/Of-Waves-Winds-Wonderful-Things_FINAL51.pdf.
- Kaili, A., 2012. I am in no way optimistic, but I remain a prisoner of hope. In: *From a Tongan Daughter: The World as I see It*. <http://anapesi.blogspot.co.nz/>.
- Ministry of Education, 2020. Action Plan in Pacific Education 2020–2030. <https://conversation-live-storage.s3.amazonaws.com/public/PACIFIC/Action-Plan-for-Pacific-Education/Pacific-Education-Plan-Summary.pdf>.
- Ministry of Health, 2020. Ola Manuia: Pacific Health and Well-Being Plan 2020–2030. https://www.health.govt.nz/system/files/documents/publications/ola_manuia-phwap-22june.pdf.
- Ministry of Pacific Peoples, 2018. Pacific Aotearoa: Lalanga Fou. <https://www.mpp.govt.nz/assets/Reports/Pacific-Aotearoa-Lalanga-Fou-Report.pdf>.
- Ministry of Pacific Peoples, 2020. Pacific Aotearoa Status Report: A Snapshot. <https://www.mpp.govt.nz/assets/Reports/Pacific-Peoples-in-Aotearoa-Report.pdf>.
- Samu, T.W., 1998. Social Studies: The Nebulous Cinderella Subject of the New Zealand School Curriculum. Unpublished Masters thesis, University of Auckland.
- Samu, T.W., 2010. Pasifika Education: an oceanic perspective. In: *Pacific Research in Education: New Directions*. MAI Review, Issue 10. www.review.mai.ac.nz/index.php/MR/issue/view/15.
- Samu, T.W., 2015. The “Pasifika Umbrella” and quality teaching: understanding and responding to the diverse realities within. *Waikato J. Educ.* 12 (1).
- Samu, T.W., 2020. Charting the origins, current status and new directions within Pacific/Pasifika education in Aotearoa New Zealand. *N. Z. Annu. Rev. Educ.* 26, 197–207. <https://doi.org/10.26686/nzaoe.v26.7138>.
- Sharp, T., Samu, T., Trotman, R., Cram, F., Theodore, R., 2020. Pacific educational success in Aotearoa New Zealand: comparing Ngā Tau Tuangahuru study findings with policy directions and research. *Pac. Asian Educ.* 59.
- Teaiwa, K., Henderson, A., Wesley-Smith, T., August 31, 2021. *Sweat and Salt Water* (Pacific Islands Monograph Series). University of Hawaii Press.
- Teaiwa, T., Mallon, S., 2005. Ambivalent kinships? Pacific people in New Zealand. In: Liu, J.H., McCreanor, T., McIntosh, T., Teaiwa, T. (Eds.), *New Zealand Identities: Departures and Destinations*. Victoria University Press, Wellington, pp. 207–229.
- Wendt, A., 1976. Towards a new Oceania. *Manag. Rev.* 1 (1), 49–60.

Decolonizing psychology: weaving thrivance to disrupt colonial eugenics and shift educational practices within Indigenous psychology

Michelle Johnson-Jennings^{a,b,c} and Derek Jennings^{b,d}, ^a Division of Environmental Health and Land-based Healing, IWRI, University of Washington, School of Social Work, School of Public Health, Seattle, WA, United States; ^b University of Washington, School of Public Health, Seattle, WA, United States; ^c Full Professor, Director of Indigenous Environmental Health and Land-based Healing Division, Indigenous Wellness Research Institute/IWRI, University of Washington, School of Public Health and Social Work; and ^d Bezruchka Family Endowed Professor School of Public Health University of Washington

© 2023 Elsevier Ltd. All rights reserved.

Introduction	171
Decolonizing psychology	172
The colonial tool of western psychology and eugenics	173
Cultural appropriation within western psychology	174
Plastic medicine wheel	175
Weaving thrivance: indigenous psychology	177
Conclusion	178
References	178

Introduction

Western psychology is often conceptualized as the study of the mind and behavior. Yet, western psychology as a field of study and practice is young-only developing since the mid to late 1800s (Asthana, 2015). According to the Merriam-Webster dictionary (Merriam-Webster, 2004), the word “psychology” originates from the Greek language with *psychē* meaning “breath, principle of life, life, soul,” and *logia* meaning “speech, word, reason.” Though great variation exists, there are commonalities between this western definition of psychology and the worldviews among over 476 million global Indigenous peoples speaking more than 400 distinct languages, including 5.2 million residing in North America (Worldbank, 2022). Similar to the Greek definition of psychology, many Indigenous cultures incorporate elements of the soul/spirit and words/reasoning into views of health (Struthers et al., 2004). As the first author was culturally instructed, our breath flows through us from our ancestors, guiding us with their loving touch. We carry this ancestral breath toward the future through our sacred bones (see Fig. 1). As we stand upon the ground connected to *iholitopa ishki*/mother earth, being embraced by the skies, we are unified and can be in a state of health under our ancestors’ watchful gaze and direction. To us, the concepts of *breath* and *life* are sacred, and words are seen as powerful in extending knowledge from our bodies to the world around us and beyond. These words and teaching are not distinct from our mind or body

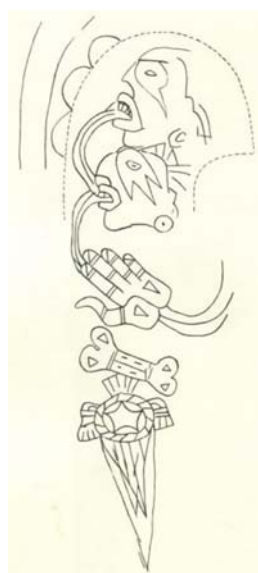


Fig. 1 Ancestral Mississippian shell engraving. <https://peachstatearchaeologicalsociety.org>.

and can be embodied during knowledge exchange, as seen within Indigenous worldviews across the globe (Batacharya and Wong, 2018).

Despite these initial overlaps in terminology, western psychology is shrouded in a cloak of oppression within Indigenous communities. As a colonial tool, it has objectified Indigenous peoples as “primitive” and “inferior” to whites in order to justify further trauma (Garth, 1931; Walker, 2018). With both Indigenous authors being raised in Oklahoma, they both often witnessed the need for psychological services. For instance, American Indian children have the highest depression rates than any other racial group in the US, as well as disparate rates of substance use disorders, post traumatic stress disorder, and suicide (American Psychiatric Association, 2017). Many are in dire need of emotional and social support to cope with historical and current traumas. Yet, there continues to be difficulty in accessing behavioral health services including long waits, few providers, and limited access to Indigenous traditional healers- if preferred (American Psychiatric Association, 2017; Aronson et al., 2016; Gone and Trimble, 2012). Furthermore, Indigenous communities often perceive western psychology as having high risks. Recent research has found that Indigenous communities can view western psychology as belonging to a foreign culture, or as being “white” (Gone, 2007). This may be due to differences in function and treatment. For instance, most western healing practices relegate spiritual health to spiritual healers, who are seen as distinct from health professionals; while Indigenous healers often address physical, mental and spiritual ailments at once viewing them as interconnected (Johnson-Jennings et al., 2018). Therefore, western psychology may not always be appropriate or effective for Indigenous patients. In addition, the history of western psychology continues to influence Indigenous community perceptions. As an Indigenous psychologist, the first author has heard Indigenous community narratives of psychologists, and their patient records, ending up in court to testify against Indigenous peoples, even for minor infractions. Many of these narratives culminate with stories of imprisonment of the community member(s) and, or removal of Indigenous children. In other words, the community, as whole, can experience large gaping wounds stemming from using western psychology. However, its not only the recent trauma that echoes across Indigenous communities and builds distrust, but also western psychology’s history of systemic oppressions. In particular, western psychology has been plagued by its systematic contribution to the genocide of Indigenous peoples using eugenic tactics, since its inception (Ilyes, 2020).

Given access issues to both western and traditional healing methods and varying cultural beliefs among Indigenous peoples, western psychology might be preferred and, or be the only avenue for Indigenous populations given the disparate mental health needs, as related to historical and, or ongoing systemic racism. Therefore, both western psychology and Indigenous psychology continue to be needed for Indigenous wellness (Aronson et al., 2016; Gone and Trimble, 2012). For instance, despite western medicine fragmenting healing by relying upon healing methods that compartmentalize- the body, mind, and spirit, it may still prove useful to Indigenous populations if employing appropriate care, cultural empathy, and overall respect (Goodkind et al., 2015). Yet, a need persists for western psychology to be decolonized in practice, education, and research. Despite western psychology having dominated medical approaches to healing beyond the physical body, many Indigenous psychologists have, remarkably, indigenized their psychological practices, research, and education. However, instead of developing a field of “Indigenous psychology” as grounded in Indigenous worldviews, international Indigenous scholars have argued that many have only imposed Indigenous elements onto the western psychology framework (Marsella, 2006). Some may call this, colloquially, as adding a feather (Walters et al., 2020). Though this may be the case for some, other revolutionary psychologists, such as Dr. Mason Durie indeed grounded his Te Whare Tapa Whā model in Indigenous worldviews (Durie, 1994). Though focused on a specific Indigenous group, Māori, this model inspired generations of Indigenous students across the globe to pursue psychology, such as the first author. Within this chapter, the authors (i.e., being Indigenous scientists, one psychologist and one health educator) will briefly discuss western psychology drawing from a decolonial and indigenist lens with the hope of supporting social transformation in practice, research, and education. To do so, western psychology’s foundation in racism and eugenics, as well as the appropriation of Indigenous psychology will be discussed. However, one cannot discuss these atrocities without uplifting Indigenous psychology and the thriving of Indigenous peoples today, which includes revitalizing and reinventing Indigenous healing methods, which will also be introduced.

Decolonizing psychology

As represented in Fig. 2, academic elder, Dr. Tuhiwai Smith (Smith, 2021) argues that decolonizing is a risky but necessary endeavor when working with Indigenous knowledges and ways of being. She provides a framework that has advised numerous Indigenous



Fig. 2 Tuhiwai Smith's (2021) decolonizing framework: Risky Business.

academic descendants, as well as allies, who seek to decolonize. She proposed that one must simultaneously unmask Indigenous knowledges, such as Indigenous psychologies and healing practices that have been subjugated by western psychology. As well, Tuhiwai Smith's encourages academics and others to actively create space for Indigenous voices-while challenging ethnocentric racism, both historically and present. To seize Tuhiwai Smith's call for activism within the field of psychology, for the psychologists, scientists, and educators must recognize the many social injustices that western psychology has committed against Indigenous peoples and challenge these ongoing colonial notions. In particular the historical and ongoing use of psychology to other Indigenous peoples and frame them as the deficit (Hartmann et al., 2019; Walker, 2018) must be addressed. This deficit approach in practice and research can be seen with the reference point, or norm, of "psychology" being European and, or western, white settler culture. Though found ethnocentric, western psychologist often advocate that their theories have accounted for scientific unbiases through "rigorous western methods" (Sherif, 1979). They further frequently ignore biased cultural influences on leading theorists. Therefore, a decolonial approach to psychology is necessary to move the field forward, to rebuild, and to effectively train psychologists and other mental health providers, educators, and scientists.

The colonial tool of western psychology and eugenics

"If there is to be reconciliation, there must first be truth.- Timothy Tyson remarks regards to the Truth and Reconciliation Commission of Canada (*If There Is To Be Reconciliation, There Must First Be Truth* – KapitalFM 92.9 Abuja, n.d.)

Western psychology has served as a colonial tool of genocide from early on in its development, as aforementioned. During the 19th and 20th century, western psychologists sought to relegate Indigenous persons as inferior. The Eugenics movement, which argued for "the practice or advocacy of controlled selective breeding of human populations (as by sterilization) to improve the population's genetic composition" (Merriam-Webster, 2004), took root in western psychology, forming the backbone of many practices. In particular, Indigenous culture, behaviors, and thoughts were deemed inferior and of lower intelligence than that of western, white settler families. The IQ test gave "evidence" to the racist notion of white settler superiority, while framing Indigenous adults and children as "feeble minded" (Walker, 2018). Scholars have since exposed the intertwinings of the Eugenics movements with western psychology's inception (Buss, 1976, Levine, 2016, Yakushko, 2019). However the impact on Indigenous communities has not often been addressed.

Eugenics psychology and intelligence testing were particularly problematic in increasing systemic racism. The results of these biased tests on Indigenous peoples informed policies and law (Levine, 2016; Yakushko, 2019). Terman, a Stanford University Eugenics proponent, standardized Alfred Binet's IQ test for the US. Unfortunately, he drew from his biased views of genetic and race based intelligence to retain culturally biased tests, and, or modify scoring scales to support views of inferiority. In doing so he reinforced that American Indian children, as well as Mexican American and Black, or African American, children were the deficit, or exhibited a presumed "feeble mindedness" (Terman, 1916). In doing so, he and his team tested, validated, and promoted inherently biased IQ tests that persist today. These racist biases can be seen within Terman's writings:

It is interesting to note that M. P. and C. P. represent the level of intelligence which is very, very common among Spanish-Indian and Mexican families of the Southwest and also among negroes. Their dullness seems to be racial, or at least inherent in the family stocks from which they come. The fact that one meets this type with such extraordinary frequency among Indians, Mexicans, and negroes suggests quite forcibly that the whole question of racial differences in mental traits will have to be taken up anew and by experimental methods.

Terman (1912/2007, p. 456).

To reify this "evidence," white settler children's scores were compared to Indigenous children, who were most often those who were forcibly and imprisoned within a boarding school and likely had no interest in completing these foreign tests. At the same time, Indigenous children were often experiencing human rights violations including rampant physical, mental, and sexual abuse, and forced child labor. As Dr. Dorothy Leighton found in one boarding school alone, over 40% of the children were malnourished as well as other untreated ill conditions (Schuck et al., 1964). However, Terman's researchers who helped norm and test the initial Stanford-Binet IQ test also held attitudes such as Rowe (1964) below:

The striking difference between the Indians and the two groups of white children cannot be explained by hygienic, social and educational differences ... Moreover, many of the children examined had spent three or more years at the government school where good food is plentiful and the school conditions at least average. The children go to school one half day and work the other. The half day of work is also educational in that each child is taught some trade or occupation. It seems therefore that the only satisfactory explanation of their inferiority in terms of the tests is to be found in an inferiority of native ability.

P. 456 Rowe (1914, p. 456).

With such a biased view towards the 268 Indigenous children who were made to take this test, the researcher continued to rationalize racist views of inferiority. Notions of second language acquisition, foreign cultural concepts, and, or not wishing to take such the test did not seem to occur to the researcher.

Western psychology continued to promote intelligence and aptitude tests to justify stereotypical notions of feeble mindedness toward Indigenous children. The IQ test, and a subsequent feeble-mindedness diagnosis, was used to track Indigenous children into subpar “special education.” Even in the late 1980s, the IQ tests were argued to *not* be culturally biased, as both teacher subjective assessments and scores were similar, despite the obvious cultural biases that may be shared between the two (Reschly, 1980). Today, inherently biased IQ tests continue to be utilized to track students in education and limit their opportunities (Morrier and Gallagher, 2012). Testing has further been used as a colonial tool to take children from “feeble-minded” Indigenous mothers. Approximately 25%–50% of all American Indian children had been removed to foster care or boarding schools by the 1940s (Abourezk, 1978). Sometimes, these tests would result in families having their land, money, and rights taken away due to their incompetence, or feeble-mindedness diagnosis (Flynn, 1988). Thereby this western psychological intelligence testing, and resulting stereotypes of Indigenous persons as inferior and unworthy of human rights, disrupted families across generations and led to continued subjugation, unjust policies, and violence. For instance, a US national Indigenous insane asylum operated from the early 1900s until 1933 with children and adults from over 52 tribes being forcibly admitted and held-most without a diagnosed mental health disorder. Children, as young as age 2 years, were admitted for causing problems at school and for their families not giving up “primitive” Indigenous cultural practices. Each inmate received unethical and deplorable treatment (Dilenschneider, 2013). This asylum, and related psychological “treatment,” have had long lasting and profound effects on Indigenous communal views on mental health practices and the legitimacy of such.

Western psychology’s eugenic roots and testing further promoted the genocide of Indigenous peoples by justifying forced sterilization. Prior to the 1910 US national law legalizing forced sterilizations, several states enacted laws regarding the ability to forcibly sterilize those deemed “feeble-minded,” often targeting American Indians (Cackler et al., 2016; Lawrence, 2000; Stern, 2005). Persons would be tested and if determined undesirable, as presumed by their Indigenous race and test scores, they could then be forcibly sterilized. Over 25% of all American Indian women were sterilized by 1976, and cases continue to occur into the present 21st century (Lawrence, 2000). This has affected not only the future generations of Indigenous peoples, but violently impacted the Indigenous mothers who were still caring for the next generation. For instance, in the early 2000s, Indigenous Women who were sterilized had 2.5 times the odds of poor mental health (Cackler et al., 2016). Hence western psychology, the very field that professed to increase mental health, has been a prime contributor to inducing ill mental health and genocide against Indigenous peoples.

Another test that has greatly impacted Indigenous populations is the ADHD testing that was birthed from the 1919, *continuous mental performance attention task*, forerunner to the modern Continuous Performance Task/CPT used in the diagnostic ADHD test. While early in the 20th century, psychologist Russell Garth (Garth, 1931) argued for the statistically significant superiority of the “primitive” American Indians on this test. Some argue that the CPT test was originally utilized to demonstrate the superiority of white children over other racial groups (Garth, 1931). The test was later modified and then generated inferior scores for American Indians. Recently, on one western reservation, up to 75% of all children were likely to be diagnosed with ADHD, and Indigenous children continue to be diagnosed at much higher rates today (Walker, 2006). Given these disparate rates, CPT’s history and use in today’s ADHD diagnostic testing has been argued to support a negative racial stereotype of the “ADHD Indigenous child” (Walker, 2006). This has further caused increased intergenerational distrust and tracking of Indigenous children.

Though “standardized” intelligence or aptitude tests have been blatantly developed on racist principles, as spread from a history of eugenics, they are still in full use within educational and health settings. Furthermore, they continue to influence policies, laws, and discrimination against Indigenous persons. Therefore, to truly decolonize western psychology, our society must reconsider the use of biased testing and how it impacts Indigenous groups.

Cultural appropriation within western psychology

Though western psychology is riddled with racism, there are several foundational theories within western psychology that were influenced and, or appropriated from Indigenous cultures. One of the more famous and contested theories is Maslow’s Hierarchy of Needs. Maslow was well documented to have traveled to the Blackfoot nation and to have studied under their elders- gathering much information about their teachings, or intellectual property, of self actualization and which basic needs are necessary for functioning within communities (Blood et al., 2011). Though the tribe has always held a well-developed framework for wellbeing, the tribal elders recognize that Maslow had not completely stolen their health theory, as several aspects differed. However, they were upset that their substantial influences were not referenced (Blackstock, 2011). As Blackstock (2011) cites, “Maslow did not fully incorporate Blood First Nation understandings of ancestral, knowledge, spirituality, and multiple dimensions of reality, nor did he fully situate the individual within the context of community ([personal communication] Wadsworth, 2008). Therefore, the Blackfoot Nation and Indigenous scholars have sought to reclaim their knowledge and to voice Blackfoot theoretical views, which have been silenced and long overlooked.

Richard Katz, a psychologist and personal friend of Maslow’s surmises, “Maslow may not have elevated Siksika [Blackfoot] teachings and influence on his hierarchy of needs as he may have believed that doing so would diminish the validity of the ideas” (Kaufman, 2019). In essence Katz recognizes that western psychology and research may have quickly rejected elevating and publishing any Indigenous theories of health and wellbeing within its existing racist framework. The resistance from white settler researchers,

even today, do support this notion. For example, any online search for “Maslow” and “Blackfoot” will yield more settler paper’s decrying his connection with Indigenous knowledges than Indigenous papers discussing such. Some settlers note their disbelief that Blackfoot elders knew about Maslow’s agenda, given that they had not met him prior. Through a settler view of Indigenous inferiority, one may not be able to comprehend that Indigenous peoples have known about white settlers, and their colonial agendas for centuries and have established strategies for coping with them, far before the individual arrives on the reservation. Given these reactionary settler papers disrupting the colonial narrative will take much activism, i.e., defined as “direct vigorous action especially in support of or opposition to one side of a controversial issues” (Merriam-Webster, 2004). Indeed, we are reminded by Tuhivai Smith that decolonial “attempts to break free can involve enormous violence: physical, social, economic, cultural and psychological” (Smith, 2021, p. 27).

In the mid 1930s, Erik Erikson, another notable psychologist who helped build developmental psychology, informed his area of psychological study under the tutelage of Indigenous knowledge keepers from the Lakota and the Yurok Nations (Erikson, 1945; 1993). From the Lakota, he learned their theories and practices on personality, identity development, and child development, and nurturance. In doing so, he noted what was obvious to most Indigenous peoples, i.e., culture and external factors shape behavior for adolescents. Afterwards Erikson used their teachings to make critical links and inform his developmental theories, especially around adolescence. Though the word “adolescence” appeared in the 15th century and came from the Latin word “adolescere,” which meant “to grow up or to grow into maturity” (Lerner and Steinberg, 2009, p. 1), it wasn’t until 1904 that the first president of the American Psychological Association, G. Stanley Hall, was credited with “discovering” adolescence (Henig, 2010, p. 4). Meanwhile, Indigenous cultures have long held the notion of adolescence, as we understand it today. Prearrival of European settlers, Indigenous pre-adulthood and exploration was well practiced within Indigenous communities. For instance among Choctaws, it is common knowledge that the ancestors did not promote marrying a young adolescent child nor having children until the age of 20 years or more. Yet at the same time, tribes supported healthy sexual development and as early as the 1700s, Europeans documented American Indians having invented and widely used an oral contraception effectively for birth control. This was 200 years before western science was to create the birth control pill (Keoke and Porterfield, 2009), thus demonstrating complex biological and pharmaceutical knowledges. Ironically, Margaret Sanger who developed the birth control pill for western medicine was also a large proponent of eugenics and forced sterilization (Franks, 2005). Though Erikson recognized Indigenous people’s influences on his work, racist notions of inferiority prevented the recognition of Indigenous knowledges as theories, and as intellectuals who could have readily co-authored. However, Indigenous academics, have been trained in western education and earning doctorate degrees for more than a century. For instance, the first Indigenous lawyer, and Indigenous rights proponent, was Choctaw tribal member, James McDonald, who graduated from Georgetown Law School in 1831 (Hoxie, 2013). Furthermore since the 1900s to 1930s, through the Society of American Indians, over 230 American Indian intellectuals, professors, and doctors were organized, actively met at conferences, and were indigenizing academic theories. They influenced US policies for Indigenous rights and impacted laws affecting Indigenous peoples (Allen, 2013). Meanwhile, many in western psychology were still calling Indigenous peoples “primitive” and claiming that they had “inferior” intellect.

Similarly esteemed Carl Jung developed the field of analytic psychology as guided from Indigenous knowledge and theories. However, he differed from Maslow in that he readily discussed the influence of tribal knowledge upon his work; albeit, through a paternalistic and ethnocentric lens. Indigenous scientists have argued that Jung objectified and romanticized Indigenous peoples using a colonial gaze that skewed his interpretations (see Durie, 1994; Deloria, 2009). This can be seen in Jung’s writing “it is not only primitive man whose psychology is archaic. It is the psychology also of modern civilized man. ... every civilized human being, however high his conscious development, is still an archaic man at the deeper levels of his psyche” (Jung, 1964, pp. 50–51). Jung further recognized that while “whites believe that they thought with her head, every one of the Pueblo knew that thinking comes from the heart” pp. 247–248. Unfortunately, Jung did not cite the Indigenous knowledge keepers from whom he learned as co-authors; instead, he simply integrated their intellectual property into his work as his own. Furthermore, he was often seen to twist Indigenous intellectual theories presenting them through his colonial lens. For instance, “In evoking the figures of the devil and the divine, Jung interpreted the trickster figure in comparative terms that made sense to European psychologists and scholars, but which had little to do with American Indians” (Deloria, 2009). A strong Indigenous co-author and intellectual could have readily assisted Jung in writing a unified and accurate view of Indigenous analytical psychology, which could have moved the field forward even more. One Indigenous co-author could have been the highly esteemed, Dr. Marigold Linton, who graduated with a PhD in Psychology in 1964 and was the first doctorate level American Indian psychologist (Trimble and Clearing-Sky, 2009). Despite many other Indigenous intellectuals with doctorate degrees, missed opportunities for advancing the field of western psychology were likely marred by racist and ethnocentric notions that did not respectfully recognize and integrate Indigenous intellectualism.

Plastic medicine wheel

-When the hen nurtures and raises the egg as her own, even her offspring doth not protest its intrusion- author’s Choctaw Family proverb

As aforementioned, non-Indigenous western scholars/researchers have taken Indigenous concepts, repackaged them within their western framework and then applied them to healing. Indigenous psychology has also been infiltrated by western concepts, which

are not properly cited or recognized as noted in the above Choctaw family proverb. One of the most prominent is the colonial view of the “medicine wheel” as related to behavioral health, which continues to plague Indigenous psychology. A brief online search will yield many papers by Indigenous, and non Indigenous allied academics, who write about “decolonizing” and using “indigenist/Indigenous research methods;” yet, they still utilize the compartmentalized westernized “medicine wheel” to guide their health interventions. This remains highly problematic for several reasons.

First, the *medicine wheel* term was created using a colonial settler gaze upon an authentically sacred symbol, which is shared across Turtle Island, or North America. The mere terminology has been argued as a form of cultural appropriation and misuse (Robertson, 2021). At the turn of the century, several white settlers set their eyes upon the Bighorn Wheel in Sheridan Wyoming and decided to term this symbol a “medicine wheel”, based on the circular symbol with four segments having a circle in the middle. These stones and shapes obviously depicted a distinctly sacred site. However, by asserting power and dominion over Indigenous symbolism, white settlers began to identify and categorize the naming of all similarly shaped icons both on ground and in drawings, as the “medicine wheel” (Brumley, 1988, 2006; Eddy, 1977; Kehoe and Thomas, 1979). Additionally, many tribes utilized similar iconography as found within carvings, writings, etc. For example the first author’s Choctaw culture utilizes a circular symbol, sometimes with a central ring, that often refers to our sacred fire and simultaneous connection to multiple worlds (Johnson-Jennings et al., 2022). In essence, the medicine wheel was named by settlers as “medicine,” but doing so reduces the fluidity and sacredness of the symbol to be about healing alone. From an Indigenous view, naming is ceremony, and for this object, settlers have done the naming, a “wheel.” Given one western definition of a wheel *being to publicly produce someone or something and use them to help you achieve something* (Merriam-Webster, 2004), settlers viewed these symbols as assisting humans alone, reflecting the colonial view of ownership and taking. Whereas Indigenous frameworks often view healing as bidirectional with the land/rocks/sky being active participants and kin in the healing process-rather than mere objects (Johnson-Jennings et al., 2020, 2022). Thus, this naming has eroded the original meaning behind these sacred sites and symbols. For some, this symbol is known to represent the sacred directions, movements of the cosmos, and align with star constellations (Robertson, 2021). However, over time the colonial lens and promotion of the medicine wheel has overridden the actual ancestral use of these symbols and sacred sites.

Secondly, the medicine wheel has been appropriated by settlers and then infiltrated western medicine and psychology to day. Like a cancer, it has proliferated and taken over. Beginning in the 20th century, the medicine wheel was invented as a potential health tool for American Indians, as created by what some may call a *plastic*, or fake, white settler, medicine man. In 1972 Charles/Arthur Storm, or as he called himself *Hyemeyohsts*, a reportedly invented name, falsely claimed to be a Cheyenne tribal member, even forging a tribal enrollment card, which was later exposed by the tribe (Velie and Mclinton-Temple, 2007). Storm then published what he called “Cheyenne” knowledges with the medicine wheel as a teaching tool for wellness. In doing so, he researched some elements of Indigenous healing and proceeded to insert yellow, black, white, red, and blue/black colors on his wheel, as well as assign animals and personality traits. He then devised the separation of “mind, body, spirit, emotions” for non-Natives. However, dissection of the self as separate mind, body, spirit and emotions has not been found in Indigenous traditional views pre 21st century (Robertson, 2021). In fact, Robertson (2021) argues that no tribal language in Canada and, to the authors’ knowledge, no tribes in the US have a word for *mental health* (i.e., referring to the health of the *mind* as distinct and separate from emotions). Hence, using this medicine wheel symbol to refer to the individual and separation of self into four quadrants are colonial constructs (Robertson, 2021). To make this model passible, Storm integrated concepts of balance and relationships that he appropriated from Indigenous culture. His cunning cultivation of the plastic medicine wheel took root, not only within new age movements that he said to have led, but clandestinely in our Indigenous communities and among Indigenous practitioners and non-Indigenous alike. Subsequently, Indigenous psychology began to emphasize the individual as self, fractured and guiding treatment with a facade of medicine wheel culture.

Third, today the medicine wheel continues to be used in psychological training, education, and practice, which unduly fragments our healing approaches and could cause harm. Some have argued that the medicine wheel remains a “mirage” that arbitrarily tries to present complex Indigenous ways of knowing into four quadrants through a colonial lens for practice or education (Widdowson and Howard, 2013). As Indigenous psychologists and educators, we must ask ourselves if what we practice actually aligns with Indigenous ancestral teachings and instruction, or if it has been filtered through a colonial lens to be palatable but not necessarily effective. One can then begin questioning if a treatment is incongruent with the community’s, family’s or individual’s ancestral teachings. One common incongruity is the western colonial framing of the medicine wheel tending to the *mind*, or as often referred to cognition, as separate from *emotions*. If utilizing this wheel and compartmentalized view of health, many Indigenous health interventions may not be as successful in addressing cultural health beliefs around wellness and balance, as they see the body as whole and inseparable. Believing its worth as an Indigenous tool, numerous Indigenous researchers base their research or health interventions upon this fractured model. However, they are often the offspring of those academic elders who have accepted this plastic medicine wheel as part of Indigenous culture. Many Indigenous psychologists may not know the medicine wheel’s history and, or feel pressure to follow precedent set by mentors who used this colonial symbol of self. Yet, Robertson (2021) argues not only does the medicine wheel do a disservice for Indigenous clients, often perpetuated by those who practice Indigenous psychology, but it limits culturally appropriate, holistic healing that could occur.

Overall the medicine wheel encourages viewing a dissolution of self, void of context. To teach this to a client or patient or student can be argued to further western colonial notions in psychological practice. This fragmentation of the person void of context can encourage ignoring the role of ignore their maltreatment; i.e., genocide, historical trauma, present trauma, and ongoing oppression on their current symptoms today. Instead of seeing self as intertwined and an interconnected whole within an intergenerational

environment, the Indigenous person and their actions are highlighted and blamed, either directly or indirectly. Though the *medicine wheel* symbol has been argued to affect one's emotion and communicate information (Ortner, 1979: 92), this conceptualization, nonetheless, could promote the blaming of the individual for not tending to their "distinct" four quadrants effectively to maintain balance and wellbeing. Hence, Indigenous psychologists, researchers, educators alike must carefully consider decolonizing the medicine wheel and practices attributed to such.

Weaving thriving: indigenous psychology

The majority of Indigenous persons are healthy and well. Yet, the vast systematic attempts at genocide, ethnocide, and assimilation have shattered Indigenous wellbeing, which reverberates throughout each generation. Indigenous people's ancestral/intergenerational connections, languages, family systems, original instructions and ceremonies have not been *lost*, they have been systematically *stolen* in an attempt to eradicate, subjugate, and dominate, as scholar Dr. Leonie Pihama has noted about Maori culture (Pihama, 2012). Beginning in the late 1400s to present day, Indigenous peoples in the Americas had their pillars of wellbeing stolen. This has spurred vast and ongoing health disparities with Indigenous peoples. For instance in the US, Indigenous peoples experience mortality rates higher than all US rates combined; including 3 times the diabetes rates and 4.6 times the rate of chronic liver diseases (mainly due to fatty liver diseases as related to obesity) (IHS, 2019). These disparities cannot be noted outside of the present day colonial settler context, which has nurtured and maintained this oppression (Czyzewski, 2011). Colonialism and ongoing structural racism continue to influence Indigenous persons lives in ways in which scientists are just beginning to understand. For instance, chronic pain rates increase for Indigenous persons who experience historical trauma, early childhood trauma, and racism (Chae and Walters, 2009), and Indigenous persons experience the highest rates of pain as compared to other racial ethnic groups (Johnson-Jennings et al., 2020). These high pains rates, coupled with racial discrimination and trauma, can also increase risks for smoking, which then increases perception of pain (Johnson-Jennings et al., 2014). Pain can become a vicious cycle and if untreated can result in individuals using substances to cope (Johnson-Jennings et al., 2020). In addition, more than 80% of Indigenous persons in the US have experienced violence, with over 90% of cases being non-Indigenous perpetrators (Department of justice, 2016). They are also more likely to be killed by the police, experiencing the highest rate (i.e., 2.9) as compared to Blacks (i.e., 2.6), Hispanic/Latinos (i.e., 1.7) and Whites (i.e., 0.9) (Hansen, 2017). This ongoing violence continues to feed mental health disparities in Indigenous communities today (Warne and Lajimodiere, 2015); however, most psychological interventions focus on the self as opposed to the structural system causing harm. Practitioners, educators, and scientists alike must consider the individual and their ability to cope and draw from resources within these colonial environments. In fact, Indigenous communities have, or are revitalizing, their traditional medicines, their kinship supports, their original instructions from their ancestors, etc. in order to promote resilience and wellness. And, Indigenous peoples are doing more than surviving, they are continuing to reclaim and thrive using their health approaches. Some describe this as *thriving*, which refers to drawing from the ancestor's (i.e., their ancestral knowledge and practices) while focusing on living in a good way for the future ancestors (Johnson-Jennings et al., 2020; Billiot et al., 2020). This thriving is seen in the many community engaged and grassroots projects that emanated during COVID (e.g., Jennings, 2020; Tuason et al., 2021).

Today Indigenous populations are continually revitalizing and reinventing Indigenous approaches to healing. They are strong and carrying forward their cultural beliefs and teachings for the next generations. They are influencing psychological education, practice, and building healing from the grassroots level. For instance, Canada's educational system is indigenizing as mandated by their psychological association. The Canadian Psychological Association's did so as a response to the Truth and Reconciliation Committee report asserts that all "undergraduate psychology students should have access to a course on Indigenous cultural literacy" that covers (a) an introduction to Indigenous knowledge and worldviews (b) the impacts of Canada's colonial policies (e.g., the Indian Act, Residential Schools, the Sixties Scoop, intergenerational trauma), on Indigenous peoples and communities, and (c) Indigenous psychology and cultural allyship" (Schmidt, 2019). However, while considering decolonizing, one may ask why western European based, settler psychology is termed "psychology", and practices related to Indigenous culture be considered "Indigenous psychology." This reference point normalizes western settler psychology as the baseline. However, Indigenous knowledges and health practices continue to be quite effective not only for Indigenous peoples, but as seen through western psychology's cultural appropriation of Indigenous healing ceremonies around psychedelics for non-Indigenous persons too (Gukasyan and Nayak, 2021). However, Indigenous scholars and knowledge keepers should be guiding these efforts, as they maintain ownership of their intellectual property.

Furthermore, Indigenous psychology has promoted psychological practices based on Indigenous culture as opposed to the eugenic-based, western psychology. Though this does prove difficult given the vast cultural differences that exist between tribes, commonly held beliefs that are emic, or shared. Though *Indigenous psychology* could simply be referred to as *Indigenous healing*, Indigenous psychologist scholars Dr. Mason Durie (see Durie, 1994), Dr. Teresa LaFromboise (see LaFromboise and Lewis, 2008) and Indigenous psychology philosophers such as Vine Deloria (Deloria, 2009), only to name a few, have paved the way for developing an Indigenous psychology field of practice and research that is decolonial. The authors working definition for *Indigenous psychology*—as distinct from western psychology, includes the following: (1) consideration of historical and present day traumas on Indigenous health and coping with a collective self vs. an individual. (2) including the study, pursuit, and practice of Indigenous healing, which is emically holistic and complex and wise. (3) viewing healing as interconnecting self with human and non-human kin and through multi-temporal (past, present, and future) and spatial dimensions of health, as guided by ancestral direction. Indigenous healing

practices have not ceased. They have continued across the millennia. Indigenous ancestors have been excellent scientists and researchers who intelligently drew from the spiritual world to guide their research, study, and practice. Presently, Indigenous academics, leaders, elders, and intellectuals continue to guide these initiatives to refine Indigenous psychology.

Through Indigenous psychology, western psychology's colonial history, narratives, and biases can be exposed and addressed to effectively serve Indigenous populations who are in need. At the same time, further space for Indigenous psychology must be created and maintained. Hence, a decolonial framework for western psychology is needed. As our academic elder, Professor Linda Tuhiwai Smith notes, "Decolonizing Methodologies is not a method for revolution in a political sense but provokes some revolutionary thinking about the roles that knowledge, knowledge production, knowledge hierarchies and knowledge institutions play in decolonization and social transformation" (Smith, 2021). Indigenous psychology is an ever-growing field that if led by Indigenous peoples can have vast impacts on improving health. There also exists a need for future research to determine how Indigenous healing methods could inform individual and group treatment for Indigenous peoples in innovative ways. This may include Indigenous practitioners, educators, and scientists appropriating some western psychology tools for our own use, on our own terms. Our ancestors have often done this since European invasion, with beadwork being a prime example. If grounded within Indigenous ways of knowing and science, Indigenous psychology can adapt and integrate as needed.

Finally, Indigenous communities are engaging in a plethora of strength's based research to build their health and wellbeing. Some are grounding their work in Indigenous healing paradigms of land and culture and are conducting land-based healing to improve health and prevent chronic diseases (e.g., Jennings, 2020; Johnson-Jennings et al., 2020, 2022; Redvers, 2020; Walters et al., 2020). They then augment their approaches with western psychology as appropriate. For instance, Johnson-Jennings has founded her research and approach in an Indigenous cultural framework and then weaves in her narrative therapy training as appropriate. For instance, through her blending, she has co-developed land-based narrative reactive therapy, which focuses on physically visiting land sites for historical trauma and re-narrating these as focused on the strength, resilience, and love of those ancestors who experienced such (Johnson-Jennings et al., 2022). However, the individual is not the only focus of this health intervention approach. The land is viewed as an active participant who is also healed through reconnection and ceremony (Johnson-Jennings et al., 2020). Still, other Indigenous communities are engaged in returning to traditional foods and practices and involving the entire community in feasts, activities, social dances as a method to restore mental health and healing (Johnson-Jennings et al., 2021). The communities are often utilizing their Indigenous healing frameworks and tending to the whole person, the community, and their environment, further spurring thriving. During COVID-19, Indigenous communities quickly rose to the challenge of maintaining wellness and developed much community activism and support (e.g., Akuhata-Huntington et al., 2020; Bratina, 2017; Tuason et al., 2021). Indigenous persons are no longer merely surviving, but are active leaders and intellectuals who are now in a state of promoting thriving.

Conclusion

Humankind has not woven the web of life. We are but one thread within it. Whatever we do to the web, we do to ourselves. All things are bound together. All things connect.

Chief Seattle, Duwamish.

This chapter is by no means exhaustive or a complete decolonial take on western psychology. It is however, one initial step in exposing the subjugation of Indigenous knowledges and intellectual property. It further has attempted to shed some light on the advances that Indigenous knowledges have gifted to western psychology. At the same time, this chapter has proposed to illuminate how colonial views of health can be masked within Indigenous cultures, such as the medicine wheel, and how this may influence treatment outcomes. Hence, a need exists for more Indigenous psychologists and allies who can guide western psychology in creating space for Indigenous psychological training, education, and practice.

References

- Abourezk, J., 1978. S.1214 - 95th Congress (1977-1978): Indian Child Welfare Act (1977/1978) [Legislation]. <http://www.congress.gov/>.
- Akuhata-Huntington, Z., Foster, S., Gillon, A., Merito, M., Oliver, L., Parata, N., Ualesi, Y., Naepi, S., 2020. COVID-19 and Indigenous resilience. *Higher Education Research & Development* 39 (7), 1377–1383. <https://doi.org/10.1080/07294360.2020.1823327>.
- Allen, C., 2013. Introduction: Locating the Society of American Indians. *Studies in American Indian Literatures* 25 (2), 1–22. <https://muse.jhu.edu/article/515529>.
- American Psychiatric Association, December 19, 2017. Mental Health Disparities: American Indians and Alaska Natives. Mental Health Disparities: American Indians and Alaska Natives. <https://www.psychiatry.org/File%20Library/Psychiatrists/Cultural-Competency/Mental-Health-Disparities/Mental-Health-Facts-for-American-Indian-Alaska-Natives.pdf>.
- Aronson, B.D., Johnson-Jennings, M., Kading, M.L., Smith, R.C., Walls, M.L., 2016. Mental health service and provider preferences among American Indians with type 2 diabetes. *Am. Indian Alaska Native Ment. Health Res* 23 (1), 1.
- Asthana, H.S., 2015. Wilhelm Wundt. *Psychol. Stud.* 60 (2), 244–248. <https://doi.org/10.1007/s12646-014-0295-1>.
- Batacharya, S., Wong, Y.-L.R., 2018. *Sharing breath: Embodied learning and decolonization*. AU Press, Athabasca University.
- Blackstock, C., 2021. The Emergence of the Breath of Life Theory 8 (1), 16.
- Blood, N., Head, R. Heavy, Microtraining Associates (Directors), 2011. *Naamitapiikoan: Blackfoot influences on Abraham Maslow's Developmental and organizational psychology*. Microtraining Associates.

- Bratina, B., 2017. Covid-19 and Indigenous Peoples: From Crisis Towards Meaningful Change, 107.
- Brumley, J.H., 1988. Medicine Wheels on the Northern Plains: A Summary and Appraisal, 12. Archaeological Survey of Alberta. <https://www.thecanadianencyclopedia.ca/en/article/medicine-wheels>.
- Brumley, J.H., 2006. Medicine Wheels. In: The Canadian Encyclopedia. <https://www.thecanadianencyclopedia.ca/en/article/medicine-wheels>.
- Buss, A.R., 1976. Galton and the birth of differential psychology and eugenics: Social, political, and economic forces. *Journal of the History of the Behavioral Sciences* 12 (1), 47–58.
- Cackler, C.J., Shapiro, V.B., Lahiff, M., 2016. Female Sterilization and Poor Mental Health: Rates and Relatedness among American Indian and Alaska Native Women. *Women's Health Issues* 26 (2), 168–175. <https://doi.org/10.1016/j.whi.2015.10.002>.
- Chae, D., Walters, K., 2009. Racial Discrimination and Racial Identity Attitudes in Relation to Self-Rated Health and Physical Pain and Impairment Among Two-Spirit. *American Journal of Public Health* 99 (S1), S144–S151. <https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2007.126003>.
- Czyzewski, K., 2011. Colonialism as a Broader Social Determinant of Health—ProQuest. *International Indigenous Policy Journal, Suppl. Health and Well-Being* 2 (1). <https://doi.org/10.18584/ijpi.2011.2.1.5>.
- Deloria, V., 2009. C.G. Jung and the Sioux Traditions: Dreams, Visions, Nature and the Primitive. Spring Journal Books.
- Dilenschneider, A., 2013. An Invitation to Restorative Justice: The Canton Asylum for Insane Indians. *Northern Plains Ethics Journal* 105–128.
- Durie, M., 1994. Whaiora: Maori Health Development. Oxford University Press.
- Eddy, J.A., 1977. Probing the Mystery of the Medicine Wheels. *National Geographic* 1 (151).
- Erikson, E.H., 1945. Childhood and Tradition in Two American Indian Tribes. *The Psychoanalytic Study of the Child* 1 (1), 319–350. <https://doi.org/10.1080/00797308.1945.11823141>.
- Erikson, E.H., 1993. *Childhood and Society*. W. W. Norton & Company.
- Flynn, S.J., 1988. Western assimilationist: Charles H. Burke and the Burke Act [Thesis. Texas Tech University. <https://ttu-ir.tdl.org/handle/2346/9747>].
- Franks, A., 2005. The Eugenically Driven Female Spirit: Rhetorical Continuity from Margaret Sanger to Today. *Association for Interdisciplinary Research in Values and Social Change* 18 (1). https://www.lifeissues.net/writers/air/air_vol18no1_2005.html#a23.
- Garth, T.R., 1931. *Race Psychology: A Study of Racial Mental Differences*. Whittlesey House, McGraw-Hill book Company Incorporated.
- Gone, J.P., 2007. We never was happy living like a Whiteman": mental health disparities and the postcolonial predicament in American Indian communities. *Am. J. Community Psychol.* 40 (3–4), 290–300. <https://doi.org/10.1007/s10464-007-9136-x>.
- Gone, J.P., Trimble, J.E., 2012. American Indian and Alaska native mental health: diverse perspectives on enduring disparities. *Annu. Rev. Clin. Psychol.* 8 (1), 131–160. <https://doi.org/10.1146/annurev-clinpsy-032511-143127>.
- Goodkind, J.R., Gorman, B., Hess, J.M., Parker, D.P., Hough, R.L., 2015. Reconsidering culturally competent approaches to American Indian healing and well-being. *Qual. Health Res.* 25 (4), 486–499. <https://doi.org/10.1177/1049732314551056>.
- Gukasyan, N., Nayak, S.M., 2021. Psychedelics, placebo effects, and set and setting: Insights from common factors theory of psychotherapy. *Transcultural Psychiatry*. <https://doi.org/10.1177/1363461520983684>, 1363461520983684.
- Hansen, E., 2017, November 10. The forgotten minority in police shootings. CNN. <https://www.cnn.com/2017/11/10/us/native-lives-matter/index.html>.
- Hartmann, W.E., Wendt, D.C., Burrage, R.L., Pomerville, A., Gone, J.P., 2019. American Indian Historical Trauma: Anticolonial Prescriptions for Healing, Resilience, and Survivance. *The American Psychologist* 74 (1), 6–19. <https://doi.org/10.1037/amp0000326>.
- Henig, B., 2010. Science, conscience, consciousness. *History of the Human Sciences* 23 (3), 15–28. <https://doi.org/10.1177/0952695110363353>.
- Hoxie, F.E., 2013. This Indian country: American Indian activists and the place they made.
- IHS, 2019. Disparities: Fact Sheets. Indian Health Service Newsroom. <https://www.ihs.gov/newsroom/factsheets/disparities/>.
- Ilyes, E., 2020. Psychology's eugenic history and the invention of intellectual disability. *Soc. Personality Psychol. Compass* 14 (7). <https://doi.org/10.1111/spc3.12537> e12537.
- Johnson-Jennings, M., Billiot, S., Walters, K., 2020. Returning to our roots: Tribal health and wellness through land-based healing. *Genealogy* 4 (3), 91.
- Johnson-Jennings, M.D., Tarraf, W., Xavier Hill, K., González, H.M., 2014. United States colorectal cancer screening practices among American Indians/Alaska Natives, blacks, and non-Hispanic whites in the new millennium (2001 to 2010). *Cancer* 120 (20), 3192–3299.
- Johnson-Jennings, M., Duran, B., Hakes, J., Paffrath, A., Little, M.M., 2020. The influence of undertreated chronic pain in a national survey: Prescription medication misuse among American Indians, Asian Pacific Islanders, Blacks, Hispanics and whites. *SSM-Population Health* 11, 100563.
- Johnson-Jennings, M., Paul, K., Olson, D., LaBeau, M., Jennings, D., 2020. Ode'imín Giizis: Proposing and piloting gardening as an indigenous childhood health intervention. *Journal of Health Care for the Poor and Underserved* 31 (2), 871–888.
- Johnson-Jennings, M., Punjabi, A., Paul, K., Jones, J., Jennings, D., 2021. Little Earth Strong: A community-level, culturally appropriate diabetes prevention pilot targeting urban American Indians. *Progress in Community Health Partnerships: Research, Education, and Action* 15 (1), 3.
- Johnson-Jennings, M., Walters, K., Little, M., 2018. And [they] even followed her into the hospital: primary care providers' attitudes toward referral for traditional healing practices and integrating care for indigenous patients. *J. Transcult. Nurs.* 29 (4), 354–362.
- Jung, C.G., Franz, M.-L. von, Henderson, J.L., Jaff, A., Jacobi, J., 1964. *Man and his symbols*.
- Kaufman, S., Director, 2019, January 10. Honoring the Wisdom of Indigenous Peoples with Richard Katz. <https://scottbarrykaufman.com/podcast/honoring-the-wisdom-of-indigenous-peoples-with-richard-katz/>.
- Kehoe, A., Thomas, F., 1979. Solstice-Aligned Boulder, Configurations in Saskatchewan. In: National Museum of Man Mercury Series, 48. Canadian Ethnology Service Paper.
- Keoke, E.D., Porterfield, K.M., 2009. *Encyclopedia of American Indian Contributions to the World: 15,000 Years of Inventions and Innovations*. Infobase Publishing.
- LaFromboise, T.D., Lewis, H.A., 2008. The Zuni life skills development program: a school/community-based suicide prevention intervention. *Suicide Life-Threatening Behav* 38 (3), 343–353. <https://doi.org/10.1521/suli.2008.38.3.343>.
- Lawrence, J., 2000. The Indian Health Service and the Sterilization of Native American Women. *American Indian Quarterly* 24 (3), 400–419. <https://www.jstor.org/stable/1185911>.
- Lerner, R.M., Steinberg, L., 2009. The scientific study of adolescent development: Historical and contemporary perspectives. In: *Handbook of adolescent psychology: Individual bases of adolescent development*, 3rd ed, 1. John Wiley & Sons Inc, pp. 3–14. <https://doi.org/10.1002/9780470479193.adlpsy001002>.
- Levine, P., 2016. *Eugenics*. University Press.
- Marsella, A.J., 2006. Indigenous and cultural psychology: understanding people in context: 2006, XXI, 518 p., 40 illus., Hardcover, ISBN-10: 0-387-28661-6, ISBN-13: 978-0-387-28661-7 (Kim, U., Yang, K.S. Hwang, K.K., Eds) *Int. J. Ment. Health Addiction* 4 (4): 377–378. doi:10.1007/s11469-006-9044-6.
- McNally, M.D., 2020. *Defend the Sacred: Native American Religious Freedom Beyond the First Amendment*—University of Washington. Princeton University Press.
- Merriam-Webster, 2004. *Merriam-Webster's Collegiate Dictionary*, eleventh ed. Merriam-Webster.
- Morrier, M.J., Gallagher, P.A., 2012. Racial Disparities in Preschool Special Education Eligibility for Five Southern States. *The Journal of Special Education* 46 (3), 152–169. <https://doi.org/10.1177/0022466910380465>.
- Ortner, S., 1979. On key symbols. In: Lessa, W.A., Vogt, E.Z. (Eds.), *Reader in Comparative Religion*. Harper, pp. 92–98.
- Redvers, J., 2020. The land is a healer": Perspectives on land-based healing from Indigenous practitioners in northern Canada. *International Journal of Indigenous Health* 15 (1), 90–107. <https://doi.org/10.32799/ijih.v15i1.34046>.
- Reschly, D.J., 1980. School psychologists and assessment in the future. *Professional Psychology* 11 (6), 841–848. <https://doi.org/10.1037/0735-7028.11.6.841>.
- Robertson, L.H., 2021. The Medicine Wheel Revisited: Reflections on Indigenization in Counseling and Education. *SAGE Open* 11 (2). <https://doi.org/10.1177/21582440211015202>, 21582440211015200.

- Rowe, E., 1914. Five Hundred Forty-seven White and Two Hundred Sixty-eight Indian Children Tested by the Binet-Simon Tests—ProQuest. *Pedagogical Seminary* 21, 454–468. <https://www.proquest.com/openview/b68512afe5fc539d1fdca730ace4a8ec/1?pq-origsite=gscholar&cbl=1818970>.
- Schmidt, H., 2019. Indigenizing and Decolonizing the Teaching of Psychology: Reflections on the Role of the Non-Indigenous Ally. *American Journal of Community Psychology* 64 (1–2), 59–71. <https://doi.org/10.1002/ajcp.12365>.
- Schuck, C., Wenberg, B.G., Boedeker, M.T., 1964. Evaluation of the Boarding School Diets of Indian Children of the Dakotas and Observations on the Growth and Development of Adolescent Indian Girls, 27.
- Shaw, C., 1995. A Theft of Spirit? *New Age Journal*. <https://thetrackingproject.org/a-theft-of-spirit/>.
- Sherif, C.W., 1979. Bias in Psychology. In: Sherman, J., Torton, E. (Eds.), *The Prism of Sex: Essays in Sociology of Knowledge*. University of Wisconsin, pp. 58–75.
- Smith, L.T., 2021. *Decolonizing Methodologies: Research and Indigenous Peoples*. Bloomsbury Academic & Professional.
- Stern, A.M., 2005. Sterilized in the name of public health: Race, immigration, and reproductive control in modern California. *American Journal of Public Health* 95 (7), 1128–1138. <https://doi.org/10.2105/AJPH.2004.041608>.
- Struthers, R., Eschiti, V., Patchell, B., 2004. Traditional indigenous healing, part 1. *Compl. Ther. Nurs. Midwifery* 10 (3), 141–149.
- Terman, L., 1916. *The Measurement of Intelligence by Lewis M. Terman*. Project Gutenberg eBook. <https://www.gutenberg.org/files/20662/20662-h/20662-h.htm>.
- Trimble, J.E., Clearing-Sky, M., 2009. An historical profile of American Indians and Alaska Natives in psychology. *Cultural Diversity and Ethnic Minority Psychology* 15 (4), 338–351. <https://doi.org/10.1037/a0015112>.
- Tuason, M.T., Güss, C.D., Boyd, L., 2021. Thriving during COVID-19: Predictors of psychological well-being and ways of coping. *PLOS ONE* 16 (3). <https://doi.org/10.1371/journal.pone.0248591> e0248591.
- Velie, A., McClinton-Temple, J., 2007. *Encyclopedia of American Indian Literature*. Facts On File, Incorporated. <https://www.biblio.com/book/encyclopedia-american-indian-literature-velie-alan/d/1405328328>.
- Walker, D., 2006. ADHD as the new ‘feble-mindedness’ of American Indian children. In: *Critical New Perspectives on ADHD*. Routledge.
- Walker, D.E., 2018. Betrayal by labels: the feble-minded, ADHD native child. *Indian Country Today*. <https://indiancountrytoday.com/archive/betrayal-by-labels-the-feble-minded-adhd-native-child>.
- Walters, K.L., Johnson-Jennings, M., Stroud, S., Rasmus, S., Charles, B., John, S., Allen, J., Kaholokula, J.K., Look, M.A., de Silva, M., Lowe, J., Baldwin, J.A., Lawrence, G., Brooks, J., Noonan, C.W., Belcourt, A., Quintana, E., Semmens, E.O., Boulafentis, J., 2020. Growing from our roots: strategies for developing culturally grounded health promotion interventions in American Indian, Alaska native, and native Hawaiian communities. *Prev. Sci.* 21 (1), 54–64. <https://doi.org/10.1007/s11121-018-0952-z>.
- Warne, D., Lajimodiere, D., 2015. American Indian health disparities: Psychosocial influences. *Social and Personality Psychology Compass* 9 (10), 567–579. <https://doi.org/10.1111/spc3.12198>.
- Widdowson, F., Howard, A., 2013. *Approaches to Aboriginal Education in Canada: Searching for Solutions*. Brush Education Inc. http://www.renaud-bray.com/Livres_Produit.aspx.
- Worldbank, 2022. *Indigenous Peoples*. <https://www.worldbank.org/en/topic/indigenouspeoples#>.
- Yakushko, O., 2019. Eugenics and its evolution in the history of western psychology: A critical archival review. *Psychotherapy and Politics International* 17 (2). <https://doi.org/10.1002/ppi.1495>.

Re-weaving thinking, re-thinking weaving: theorizing Indigenous arts practice as research methodology

Hinekura Lisa Smith, Ngā Wai a Te Tūi, Māori and Indigenous Research Center, UNITEC, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	181
Revealing my threads – who I am comes before what I do	182
Part one: re-weaving thinking about methodology	183
We have always been theorists	183
Claiming space to theorize art practice as methodology	183
Our theoreticians walk amongst us	184
Part two: re-thinking weaving – theorizing an arts-based methodology	185
An idea is awoken	185
Introducing Whatuora as methodology	186
Whatuora as a radical decolonizing methodology	186
Whatu as teacher - what I learnt about theorizing Indigenous methodology	186
Summary: tying off some threads and picking up others	187
References	187

Introduction

Indigenous weaving practices, and the storied artforms they produce, hold sophisticated systems of knowledge that guide how we move through and relate to the world, how we learn and are taught, how we experience life and pass down knowledge intergenerationally. Historically, these weaving practices were practical and purposeful - such as the knowledge required to weave clothing and create shelter. More recently, Māori (the Indigenous people of Aotearoa New Zealand) and other Indigenous researchers have looked to theorize our artistic practices as research methodology to tell our stories and explain our lived experiences in our words, on our terms, and through our own Indigenous artistic lens.

Methodology as an element of research, is an immutable aspect of academic scholarship. It refers to the concepts and theories that frame the way research is conducted, that is, the knowledges and influences that underpin research as a process for creating “new” knowledge. Over the last forty years, Indigenous research methodologies within academia have taken up critical Indigenous theories (Chilisa, 2012; Kovach, 2005, 2009; Smith, 1999) that seek to re-claim, re-frame and re-present the lived realities of Indigenous peoples. Rather than accept a Western academic definition of methodology, this chapter enthusiastically takes up threads woven by other Indigenous researchers that encourage us to look to our own “ways of being” in the world, re-weaving our thinking to theorize education research methodologies that will serve us in our diverse lived realities. By re-claiming methodology as an aspect of being Indigenous that has, for centuries, served our people and allowed them to flourish, I consciously seek to re-frame how methodology is viewed both within academic research paradigms and beyond.

While the approach I forward here is grounded in Kaupapa Māori (Māori centric) theory (Pihama, 2001; Smith, 1997, 2003), this chapter reaches beyond an Aotearoa New Zealand context to connect with other Indigenous decolonizing approaches. By weaving together our different and similar threads of theory and practice we simultaneously contribute to both our local decolonizing agenda and broader global Indigenous aspirations for self-determination. Like other Indigenous scholars I lean away from a convenient “pan-indigenous one size fits” approach to thinking about methodology, instead leaping in to research approaches which are critical, culturally based, decolonizing Indigenous approaches to research such as Kaupapa Māori research that is by Māori, for Māori with our language, values and aspirations at its center.

Kaupapa Māori theory emerged in the 1990’s as a radical response from Māori academics to create space in the academy for Māori-centric research methodologies to emerge (Henry and Pene, 2001; Irwin, 1994; Lee, 2008; Nepe, 1991; Pihama, 2001; Smith, 2003; Smith, 1999) or what Pihama (2001) calls “culturally defined theoretical space” (p. 77). Another prominent Māori academic, Kathy Irwin says that Kaupapa Māori research is that which is:

[C]ulturally safe, that involves the mentorship of kaumātua, that is culturally relevant and appropriate while satisfying the rigours of research, and that is undertaken by a Māori researcher, not a researcher who happens to be Māori.

(Irwin, 1994, p. 225).

Irwin’s description of Kaupapa Māori research, while written over twenty years ago remains relevant today, as Māori continue to claim theoretical space that centers our ways of being and doing in academic research.

This chapter is in two parts. The first part sets out to re-weave thinking about research methodology and how it shapes and informs education research. By reclaiming methodology from the academy as an exclusively academic activity, and re-framing it as an inherently Indigenous practice of enquiry, we reinstate ourselves as Indigenous methodologists, theorists and experts on ourselves and, as such, the people best placed to research “us”. When we claim space to theorize Indigenous ways of being and doing as methodology, particularly Indigenous arts-based methodology, I argue that we enact a decolonizing praxis of thinking and doing (Kovach, 2021), theory and practice, that holds within its strands a transforming potential to create positive change for Indigenous people. Practicing art is enjoyable. Theorizing arts practice is interesting – but when the theory is practiced, and the practice is theorized, a powerful praxis is potentialized that elevates “craft” and the artifacts it produces to robust and valid theoretical praxis.

The second part of this chapter offers an example of how an “old” set of arts practices was theorized as a “new” arts methodology. I share an approach I developed in a Kaupapa Māori research project about eight Māori women and the aspirations they hold for their children and grandchildren to live “culturally whole” lives as Māori, expressed through the storied heirloom cloaks they wove for their families.

Overall, this chapter has two simple yet complex aims. The first is to encourage Indigenous education researchers, particularly those in arts-based education, to look to their own practices and deeply methodological ways of being and doing to theorize methodology. The second is to celebrate the Indigenous theorists that walk amongst us and to draw on their ways of being and doing in the world to theorize the research praxis and methodologies best capable of researching our lived realities from our perspective. Finally, I share the lessons I learnt from re-thinking weaving as methodology and encourage other Indigenous creative artists, scholars, researchers and teachers to contribute culturally located ways of approaching research in education. To ensure the weaving of this narrative begins correctly, I claim a positionality to this topic by introducing myself.

Revealing my threads – who I am comes before what I do

An important, and indeed expected, element of Indigenous research is to locate oneself in relationship to people, place and purpose. In Aotearoa New Zealand, often the first question asked in a Māori language exchange is not WHO are you or WHAT do you do but more importantly WHERE are you from. The response identifies one's relationship to significant geographical landmarks, eponymous ancestors and tribal canoes that connect the person to the land and its people. A right to claim this cultural relationship accrues an immediate and almost greater cultural responsibility to act in “right” ways that seek to honor those people, places and stories.

In contrast, some Western research approaches forward researchers as objectively detached observers to the people, places and stories they encounter. Botswanan Indigenous scholar Bagele Chilisa (2012) argues that some Western research approaches stubbornly refuse to interrogate the impact of researcher positionality, instead championing an acultural objectivity as better capable of generating valid and robust research “results”. Hawaiian scholar Kū Kahakalau (2019) says the expectation of Hawaiian researchers to “remain neutral and unbiased, removing his/her personal opinion from the research process” (p. 13) is incomprehensible (p. 13) given that Hawaiians bring with their mana, encompassing strengths that are simultaneously physical and emotional, intellectual and spiritual, including skills, knowledge, experiences and aspirations. I argue that researcher positionality, their intentions, and the theoretical underpinnings of the research is as important, if not more, as articulating the methods, data collection tools, analysis and conclusions. As such, I locate my reflexive positionality (Kovach, 2021) here as a Māori woman, mother, and daughter, artist and researcher, storyteller and teacher, to declare, as I disclose (Grierson, 2007), the woven stories and experiences through which I view research.

I am the daughter of a Māori mother who, in the context of 1970's feminism and cultural renaissance, sought to reclaim and restore a positive Māori identity in and for her family – ways of being that had been systematically stripped away by a colonizing education system intent on cultural assimilation (Hohepa et al, 1996; Webber, 2008). My mother's deliberate actions to reclaim language, identity and artistic practices in, and for, her immediate family, served to disrupt our family's experience of cultural disconnection so that her children, and now her grandchildren, may experience a positive cultural identity as Māori.

As a child, I reveled in the creative act of making. Whether it was sewing, cooking, or crafting, I recall a deep sense of fulfilment in creating something “new” from something “old”. Scraps of fabric were given new life as blankets or clothing. Cardboard was decorated and fashioned into tissue box covers and match sticks were glued, sanded and varnished to make lighter holders (as abhorrent as it sounds, it was the 1980's). The meditative space of making was where I could sit with my thoughts and upon completion, I experienced an immense joy in gifting what I had made to others – what we name in Māori language as koha.

As a teenager I learnt raranga (the Māori art of weaving harakeke or flax leaves), an art form I picked up very quickly. Weaving was the space where my creativity and desire to make, gift and connect to others, could be located within Māori practices, language, and knowledge. Through raranga I felt connected to an old set of knowledge and practices that enabled me to express my Māori identity in storied and creative ways. I wove backpacks and bags that were gifted to family and friends. I wove mats that were given to a local school and flower arrangements for a relations wedding. At one point in time my ability to weave enabled me to pay my bills and buy food.

After more than two decades developing my weaving and teaching practice, I undertook a PhD that explored the practice of whatu (a traditional cloak weaving technique) beyond its tangible arts practice, to include a theorization of whatu as methodology (Smith, 2017). I named this methodology Whatuora and introduce it in the second part of this chapter as an example of how re-weaving thinking “about” methodology led to re-thinking weaving “as” methodology. First though, let's begin unpicking then re-weaving some theoretical threads.

Part one: re-weaving thinking about methodology

This part of the chapter begins by re-complexifying notions of methodology within the context of arts-based practice, discussing the possibilities created when we find the space, and the courage, to re-weave our thinking about methodology and theory. I introduce you to a key theoretician in my research – my then five-year-old daughter Kahukura - who, through her sharp, bi-literate insight, provided me with a moment of theoretical clarity that led me to developing Whatuora as a Māori arts-based methodology. By sharing Kahukura's story, I encourage Māori and Indigenous researchers to reinstate ourselves as Indigenous methodologists and theorists whilst reclaiming knowledge fragmented through colonization to effect positive change for our communities. By re-weaving how we think about methodology and theory, we honor the embedded and embodied ways of knowing contained within a wide range of Indigenous language, knowledge and narratives as an active and radical form of Indigenous scholarly resistance.

We have always been theorists

Māori, and Indigenous people globally, have always engaged in research, testing methods and methodologies to solve problems, advance technologies and develop theory. Activated by decolonizing politics, re-weaving how we think about research methodology in education seeks to re-claim and re-complexify methodology from its academic constraints to instead “see” methodology through Indigenous eyes and within our creative Indigenous hands. I argue that theorization and methodology does not reside exclusively in academic literature or with distant, sometimes dead, frequently male and often Western theoreticians. As Kaupapa Māori scholar Leonie Pihama (2001) reminds us, theory and methodology did not arrive with the coloniser. Rather, Indigenous peoples have for centuries engaged in their own forms of methodological research to test theories and advance thinking:

As Māori we have a history of investigation. It is an ancient history of exploration, of navigation, not solely in the physical domain, but in ways that reach throughout the many dimensions of Te Ao Māori. These are all forms of research, they are all ways within which our people have developed knowledge and have located ourselves in the wider world. (p. 47)

Indigenous people have always been theoretical, developing our own methodologies and methods to tackle problems. As Hawaiian academic Lisa Kahaleole Hall (2019) says however, that our rich theorisations have been “overshadowed at best, or delegitimized at worst” (p. 111) by dominant Western thought. Indigenous sets of knowledge have for centuries supported great advances enabling Indigenous communities around the world to explore, adapt and advance our technologies (Chillisa, 2012; Hau'ofa, 2008; Kovach, 2009; Pihama, 2001; Smith, 1999). Yet the notion of methodology in our more recent history has been captured and claimed by the academy nudging our ancient ways of research to the academic margins.

Indigenous methodologies have evolved to fit a contemporary reality shaped by the struggle to resist the assault of colonization through projects of cultural reclamation. Eminent Māori scholar Linda Tuhiwai Smith's seminal work *Decolonizing Methodologies* (1999) for example, strongly advocates for reframing and reclaiming methodologies as a critical element of a strategic Indigenous research agenda. Smith discusses the tides or states of survival, recovery, development and self-determination that intersect with those of decolonization, healing, transformation and mobilization, none of which are linear nor are they goals or ends in themselves. Instead they are “processes which connect, inform and clarify the tensions between the local, the regional and the global ... processes which can be incorporated into practices and methodologies” (p. 116). Similarly, Indigenous educationalist and academic Leah Abayao (2006) discusses Indigenous methodology as that which is:

[A]cquired over generations by communities as they interact with the environment ... exploring indigenous technological knowledge and knowledge transmission systems, and recasting the potentialities they represent. (2006, p. 180)

Consider for example the sophisticated navigational systems developed by Pacific voyagers that enabled them to criss-cross the Pacific Ocean on deliberate and purposeful voyages - technologies later regarded as superior to those of the Western world at the time (Evans, 2015). Navigational methods and methodologies such as these required high-level theorization, research, development and testing in order for return voyages across vast expanses of ocean to succeed. Importantly here I argue that methodology is a lived experience, a praxis if you will, of weaving together the theory and practice, the thinking and doing, to create solutions.

Claiming space to theorize art practice as methodology

Theorizing is not new to us. Indigenous methodologies that enable us to think about and understand our lived experiences have existed in various forms for thousands of years. Margaret Kovach (2009), a Professor in Education of Plains Cree and Saulteaux ancestry explains Indigenous methodologies as ways of knowing that are both practical and capable of transmitting values and

knowledge. In the Aotearoa New Zealand context, Māori across a wide range of fields are working to awaken “new” knowledge from “old” knowledge by theorizing mātauranga Māori or Māori knowledge in new and useful ways.

Māori arts researchers, Donna Campbell (2019) and Kahutoi Te Kanawa (2008), both recognized internationally as experts in traditional Māori fiber arts, theorize the traditional Māori weaving practices of tāniko and rāranga as research methodology. Campbell and Te Kanawa’s in-depth practical expertise in their respective fields, bound together with their theoretical scholarship in teaching and research, sees a re-weaving of “old” know how with innovative “new” ideas grounded in an active Māori political consciousness.

Yet we do not theorize for theory’s sake. As Māori arts practitioner, Charles Royal (2011), says the goal of theorization is not to create new knowledge for its own sake but rather “to develop a way of encountering the world that upholds life, deepens our relationships with the natural world and with each other (p. 48). Similarly, Hawaiian scholar Ho‘manawanui says that Indigenous researchers are rediscovering, recognizing and utilizing long standing ancestral analytical and intellectual tools (Ho‘omanawanui, 2019) drawing from the past before us (Wilson-Hokowhitu and Meyer, 2019), weaving it with the to find solutions to our problem.

When theorizing the artistic practice of whatu as the methodology most appropriate for my research, I did not invent something “new” as is often the imperative in research from a Western perspective. Instead I gathered existing strands of thinking around, language and identity, weaving them together in such a way as to create “new” fabric from “old” threads or “new” theory from “old” ideas. While Western theory can be a useful theoretical ally and, like critical and emancipatory theory, has contributed to Māori and Indigenous scholarly thinking, astute theoreticians walk both physically and spiritually amongst our own people - if we only have the eyes to see and the hearts to hear.

Our theoreticians walk amongst us

When I began my research, there was not an established methodology that immediately appeared to me as appropriate for a project that explored living as Māori women and weaving cloaks. I noted that some research projects explain methodology in a small section within a chapter where “established” methodological approaches are assumed then “applied” to the project. Whatuora methodology is a more diffuse, interwoven idea. Theorizing the artistic practice of whatu as methodology is concerned with creating methodology, and in turn methods, from within the embodied knowledge of whatu as an arts practice. When theorizing whatu as the methodology most appropriate for my research, I did not invent something “new.” Instead, I wove together several existing sets of ideas around Māori weaving practices, language, identity in a “new” way – in a sense creating “new” fabric from “old” threads.

While holding firm to the belief that the methodology for my research would be found within its practice, it was a gift from a key theoretician in my family – my then five- year-old daughter, who prompted me to theorize the practice of whatu beyond its practical application. I began to think about the interwoven meanings embodied in the complex term whatu – something my bi-literate daughter Kahukura did instinctively - and how these interwoven meanings could be bound together as methodology. This story was written several years ago about this moment, I share with you Kahukura’s story.

I sit deep in thought, working on a kākahu (cloak). Fingers moving rhythmically. Aho (weft strands) tangling and untangling. Whenu (warp strands) coming together. Building the kākahu whenu by whenu, one by one as the feathers are added. Some in pattern some not. Some selected, others discarded. I am in the zone. Thinking, thinking. Thinking about my research. Thinking about who this kākahu is for. Thinking about what to cook for dinner.

Thinking that it is getting dark and the girls need to come inside. Thinking and weaving. Kahukura is 5 years old. She often sits beside me and does her version of “whatu” – twisting and plaiting and twirling of whenu into curious knots. I think to myself, “I love it when she sits here with me,” and I tell her so. She smiles and I can tell that she enjoys the time we sit together. I think to myself, “I wish I had had someone to sit next to, to learn whatu. Yet, I do not “teach” Kahukura to whatu. She watches. She talks. She sings. The pedagogy at work here is not teaching. It is education. It is the deliberate act of whatu pedagogy – reclaiming knowledge and learning and passing this on to our children. My aspirations, handed down to her through the time we spend together, talking, telling stories, weaving. I am cloaking her in my aspirations that, in turn, she will continue these traditions with her daughters, my future mokopuna.

She comes in from outside and asks me in Māori language “Kei te aha koe Māmā? [what are you doing mum?] E whatu ana ahau bubba.” [I am weaving.] “Oh,” she says, hums to herself, plays with the whenu and then replies. “Whatu eh Mama is the same as seeing, like with these eyes (she points to her eyes) and your other eyes (she points to her heart)? I look up.

She smiles and walks away. I stare after her as a connection is made, blinding me like a bolt of lightning. Whatu means finger twining. Whatu also means eyes, to see or a way of seeing both the physical world and through your emotional soul. In this moment Kahukura helped me to “see” that the physical practice of whatu is simultaneously a way of being and being seen in the world through the stories we weave into its fabric. I watch her go and thank the universe for the simple yet complex brain of a 5-year-old and wish I could think with her clarity more often.

(Smith, 2017, p. 92).

Kahukura's (in)sight reminded me of the richness that exists within our own language and knowledge systems. Sometimes it takes a different "lens", with which to view that knowledge and weave it together in new ways. As Māori scholar Taina Pohatu (2011) says,

...an unwillingness to properly value the potential held within Māori bodies of knowledge, deposited by generations past into the "library of te reo" [language], will have the effect of limiting our visions, horizons, and so expectations (p. 3).

Kahukura's story is an example of what Kovach might name as abductive knowledge, that is a moment of creative insight that grabs your attention, an "a-ha moment" or a "flash of knowing" (Kovach, 2021, p. 34). Abduction weaves together observation, instinct and experience offering an "inclusive, holistic consideration of knowledge" (p. 35) in an effort to establish an hypothesis (Kovach, 2021) and aligns with an Indigenous approach to theorization. Kahukura's ability to re-think weaving, its language and practice, offered me a way to re-weave my thinking of whatu, beyond its practical sense, and instead see its potential in new, theoretical and methodological ways.

I am, unsurprisingly, challenged by scholars in naming Kahukura as a theorist, after all she is not published, in fact at the time she was just learning to write. Furthermore, abductive reasoning is labeled problematic in scientific research as intangible and unverifiable. Perhaps by viewing theory such as that which Kahukura imparted as only accessible to those equipped with the creative eyes to see and ears to hear, such theory can be elevated from hearsay and anecdote. I hold firmly to the notion that knowledge, wisdom and knowing comes from many forms, animate and inanimate, human and non-human. As Indigenous people we understand that there is story in the wind and wisdom in our mountains. It is not enough, however, to simply hear or receive a story and name it theory.

Kahukura's story could have appeared as an anecdote or cute vignette within the research. Instead, I chose to position her contribution as containing important theoretical knowledge and thus it was appropriate to acknowledge her as a key contributor to developing Whatuora methodology. Having been handed the theoretical threads of an idea through Kahukura's insight, I took up these threads, weaving and synthesizing literature from a wide range of Māori, Indigenous and Western thought in order to forward whatu as a theorized research methodology most capable of underpinning my research *Whatuora: Whatu Kākahu and Living as Māori Women*. Having foregrounded here the theoretical space that supports how Indigenous scholars can re-weave our thinking about methodology, the following section offers Whatuora as an example of theorizing an Indigenous arts practice as methodology.

Part two: re-thinking weaving – theorizing an arts-based methodology

Māori and Indigenous artists maintain an ancestral legacy of innovation and creativity (Lander, 2011), holding firmly to "old" cultural traditions and arts practices and weaving those threads in "new" contemporary ways. It is this innovative creativity that I drew upon to re-think the traditional Māori art of cloak weaving as decolonizing methodology. Whatu kākahu or traditionally woven cloaks were once viewed as prized art(efacts) by early colonial collectors' (Roth, 1924) intent on cataloging the "exotic South Seas" for European consumption (Te Rito, 2006). Within a short time however, what were once priceless storied treasures, became viewed as impractical "clothing" and labour-intensive women's "craft" drawing whatu kākahu as an embodied art form and creative practice to the brink of extinction (Tamarapa, 2011; Te Kanawa, 2008), as the intergenerational transmission of skill and practice was interrupted (Te Awakotuku, 1993). After many decades of active revitalization, traditional cloak making is once again flourishing as Māori women (and some men) weave storied cloaks for the next generation (Smith, 2017). My scholarly contribution to this legacy of arts innovation and Indigenous methodologies is theorizing the art form of whatu as research methodology.

Whatuora: Whatu kākahu and living as Māori women (Smith, 2017) explored the lived experiences of eight Māori women – including myself as reflexive researcher (Kovach, 2021). Over one year, as I taught the women to weave traditionally made Māori cloaks, they spoke and journaled their storied memories of reclaiming, restor(y)ing and re-visioning what it meant to live holistically well and culturally connected lives as Māori women. The act of weaving together in women's space, of being still and centered, provided a culturally grounded safe space for the women to discuss how their experiences of growing up and "being Māori" now inform the aspirations they have for their children and grandchildren to "live as Māori". To live "as Māori" is an education discourse popularized by renowned Māori scholar Mason Durie (2001) as a way to describe Māori-held aspirations to live well, fulfilled, holistically balanced lives with access to the Māori world including language, culture, community. The women's aspirations to develop for their children, a positive Māori identity, to speak Māori language and connect to the land, its history, people, and arts practices, do not fall out of thin air. Instead, they emerged from the women's, sometimes painful, stories of their own cultural assimilation and the work they actively engaged in to heal from cultural and identity trauma – for themselves and their families.

An idea is awoken

As my practice as a cloak weaver grew, so did my interest in theorizing this knowledge at a deeper level. I began to think about the embodied nature of whatu practice, how it made me feel about myself as a Māori woman and the maternal language associated with the practice. Several years ago, I was asked by a few Māori women friends to teach them what I had learnt about cloak weaving.

As we sat and wove, I noticed similarities in the experiences of the women who attended. We would talk and tell stories, share memories, laugh and sometimes cry as we reflected on knowledge lost and gained, memories of growing up and being Māori, being women, being mothers, grandmothers, aunts and sisters. I could see the transforming and healing power of whatu at work. As well as creating beautiful, tangible garments for their families, the women were weaving together their stories and reflections, creating an intangible fabric to cloak themselves in. Our conversations and my observations of the workshops, intertwined with my interest in education research prompted me think about whatu – both its practice and its theory as an intertwined praxis – as a site to reclaim and self-determine how we understand ourselves living as Māori women.

Introducing Whatuora as methodology

Whatuora, the new term I created to name my methodology, amalgamates two not commonly associated Maori words - whatu and ora. Whatu, the term for traditional Māori finger twining, is an ancient practice developed over centuries and used to create traditional clothing ranging from everyday rough rain capes through to ornately woven and highly prized cloaks adorned with native bird feathers (Tamarapa, 2011). Whatu is also the Māori word for eyes - the lenses through which we view the world, as introduced through Kahukura's story. By bringing these two words together I wanted to embody the intertwined nature of whatu practice by twining two distinct ideas into one "new" concept. It was interesting to note that, as our research progressed, some of the women began to use "Whatuora" to sign off email communications, for example "Whatuora ki a tātou" (may your eyes/vision be well), signaling perhaps that the women had claimed what was essentially a new word as having meaning beyond the constructs of our research.

At its simplest form, the practice of whatu requires the twining together of two sets of aho or horizontal weft threads, around a set of whenu or vertical warp threads as seen in the diagram below. As one set of aho threads come to the fore, it is exchanged for another that is then foregrounded, both sets of aho interacting interdependantly with the other. Consider these two sets of interacting and interlocking threads as theory and practice. One set of threads, for example theory, is interesting on its own but becomes most potent and useful when it interacts with its own set of practices as a transforming praxis and as such builds the "fabric" of an idea line by line.

Interdependant binding threads – whatu aho rua. Line drawing by Tim Galloway, reproduced with permission from the Te Papa Collection: © Te Papa (MA_I.278976).

Ora is a common word in the Māori language that means to be well, to survive, to be healthy, fit and vital (Williams, 1997). The ubiquitous greeting in Aotearoa New Zealand "kia ora" for example literally means "be well". Of particular relevance to a Whatuora methodology is ora as a stative - to be safe, cured, recovered and healed (Williams, 1997). Ora, in the context of the Whatuora research, is understood as a journey to wellness and recovery from the impacts of colonization, for Māori women who are reclaiming, restor(y)ing and self-determining what it means to them to live well lives as Māori women. When the complex notions of whatu as eyes, ways of seeing, ways to be seen, and the practice of finger twining are woven together with notions of ora as being well, cured, healed and able to participate culturally, a new bound together idea - Whatuora - comes into being as a methodology through which Māori women's stories and experiences can be viewed (Smith, 2017, 2019, 2020).

Whatuora as a radical decolonizing methodology

Radical ideas, methodologies and approaches are necessary to decolonize our current education conditions. If "radical" – another politicized word that deserves some re-complexifying attention – means to change from the accepted, to go beyond the norm, or to advocate for fundamental or revolutionary change, then Whatuora is forwarded here as a radical Indigenous methodology. Whatuora methodology seeks to push the boundaries of whatu beyond its established understanding as "eye" or "finger twining" to a theorized approach to research that expounds a way of "seeing" the world through decolonizing eyes.

Whatuora asserts that the ability to "see clearly" or to "see well" requires an active and critical reflection of our past experiences to better understand what and how we view our present and envision our future. Applied to this research, Whatuora is a lens with which to "see" the stories and experiences of Māori women living as Māori. While Whatuora was developed within a particular research project it is not intended to remain in this Māori arts and education space. Like all decolonizing methodologies, drawing on Indigenous thought has the potential to transcend academic fields, educational settings and Indigenous contexts. While the concept emerges from within the Māori practice of whatu, the political decolonizing lens of Whatuora is both fluid and flexible.

Activated by Kaupapa Māori theory, I looked to the knowledge, language and practices of whatu to guide this research and in the process gained some valuable insights about the challenges and rewards of theorizing a "new" Kaupapa Māori methodology.

Whatu as teacher - what I learnt about theorizing Indigenous methodology

I learnt that developing theory from an Indigenous practice-based art is necessarily difficult. To explain and capture the embodied, tactile and tangible relationship between the creative practice of whatu and Māori women in a linear written form is to attempt the impossible. The transformative nature of Whatuora as a research approach occurred in its fullest sense when the wāhine sat and talked, wove and shared stories; when they connected and related to each other as Māori women; presented their completed kākahu and cried; as they described the feeling of seeing it worn by their grandchild. Transforming moments were also experienced in our

own private and quiet spaces when we wove, wrote in our reflective journals and sat with our own thoughts in the quiet solitude that whatu weaving offers. I recognized early on the impossibility of trying to capture the essence of every embodied interaction in a linear and logical written.

I learnt that methodology does not belong solely to the academic realm. Māori and Indigenous people have always engaged in research, employing methods and methodologies to solve problems, advance technologies and test theories. By re-complexifying how methodology has been constructed in an academic context, I saw an opportunity to theorize the embedded knowledge of whatu as a research approach that emerged from within the research topic - that is whatu and Māori women. The opportunity to theorize whatu as methodology required me to look with deeply theoretical and decolonizing eyes in order to “see” the theory that existed there.

I learnt that theory is everywhere if we only have the eyes to see it – and the ears to hear it. It does not reside exclusively in Western academic literature or with distant and disconnected theoreticians. I was reminded that astute theoreticians walk both physically and spiritually amongst our own people. As Goodyear-Ka’ōpua reminds us “the theorists of my academic genealogy include the women of our family who formulated their ideas through love and daily struggle” (2007, p. 53). By positioning my daughter as a theorist I acknowledged her key contribution that enabled me to understand aspects of whatu weaving that I had not previously. In doing so she prompted me to “see” more clearly the relationship between Māori women and whatu kākahu, and to then theorize an appropriate research approach from within this relationship.

I learnt that theorizing Indigenous knowledge, in this case viewing our world through Whatuora - well and healed eyes - is not a final destination. Instead, viewing our lived realities through Whatuora is a constantly engaged process of learning, listening consciously and critically, decolonizing using the tools and materials at hand in the present – with an eye to the future. For me, enacting a Whatuora approach therefore did not end with the research. Instead, Whatuora has become an internalized way of thinking about *being* and *living* in the world as a Māori woman and as such both the theory and practice of Whatuora is woven methodically and consciously into the way I live and research.

Summary: tying off some threads and picking up others

Re-weaving how we think about methodology is a necessary and political research activity. This chapter draws on Indigenous decolonizing tools and arts-based theory to unpick the fabric of Western research methodology and then re-weave it as a transforming praxis of Indigenous thinking and doing – theory and practice. When we re-complexify accepted notions of methodology we create space for “new” methodological knowledge to be awakened from “old” arts practice that seeks to activate a decolonizing politics useful to address our current lived realities. By theorising our own methodologies, Māori and Indigenous researchers are re-weaving knowledge fragmented through colonization, then re-presenting this knowledge in new and useful scholarly ways. It is imperative that Indigenous arts researchers continue to decolonize our own research whatu (lens) in order to “see” not only our research topics through Whatuora (well eyes), but to simultaneously develop our own approaches to tell our stories in ways that create positive transformation for our diverse Indigenous communities.

Indigenous research methodologies drawn from artistic practices create space in the Western academy to view and conduct research through an Indigenous artistic lens. Kovach (2009) says that, in her early academic career, there was a paucity of theoretical literature that centered Indigenous conceptual frameworks and tribal knowledge. I join Kovach, Pihama, Lee and many other Māori and Indigenous scholars, to encourage, almost implore Indigenous researchers look to our own systems of knowledge to develop methodologies to investigate our problems, and make visible the ways in which we see the world. Often activated by decolonizing politics, Indigenous researchers continue to develop and use Indigenous approaches to speak back to Western research practices that marginalize Indigenous knowledge. In doing so Indigenous methodologies continue to re-weave and re-present Indigenous methodologies as valid and legitimate academic approaches to research and, more importantly, as capable of contributing to positive transformations for Indigenous people.

I share Whatuora as an example of a theorized Indigenous arts methodology that emerged from within whatu the Māori practice of cloak weaving. Theorizing a methodology from within its very practice and language, enabled a way to talk with Māori women about how they work in active ways to decolonize their lives, finding strength and reassurance that their children and grandchildren will live self-determined and positive identities “as Māori”. Whatuora is forwarded as an original contribution to countless other decolonizing ideas and actions, played out here in Aotearoa New Zealand and in global Indigenous research, that theorize our ways of being and doing to address challenges now and for the future.

Much like weaving a cloak, the artist makes the most of the materials available to her and when a new idea or resource is presented, she employs her adaptive expertise and innovation, twined together with her traditional knowledge, to make something new. This chapter is a call to other Indigenous arts researchers to take up the threads of innovation and creativity, activism and theorization to create “new” theoretical ideas from “old” sets of arts knowledge.

References

- Abayo, L.E., 2006. Articulating indigenous people's culture in education. In: Abu-Saad, I., Champagne, D. (Eds.), *Indigenous Education and Empowerment: International Perspectives*. AltaMira Press, United Kingdom, pp. 179–187.
- Campbell, D., 2019. Ngā kura a Hineteiwaiwa: The Embodiment of Mana Wahine in Māori Fibre Arts (PhD Thesis). The University of Waikato, Hamilton, New Zealand. Retrieved from: <https://hdl.handle.net/10289/12583>.
- Chilisa, B., 2012. *Indigenous Research Methodologies*. SAGE Publications, Thousand Oaks, CA.
- Durie, M., 2001. A Framework for Considering Maori Educational Advancement. Hui Taumata Mātauranga. Conference paper.

- Evans, J., 2015. Heke-nuku-mai-ngā-iwi Busby: Not Here by Chance. Huia Publishers, Wellington.
- Goodyear-Ka'opua, N., Ka'opua, L.S., 2007. Dialoguing across the Pacific: Kūkākūka and the cultivation of Wahine Maoli identities. *Pac. Stud.* 30 (1).
- Grierson, E.M., 2007. Ways of researching: creative arts methodologies and practices. In: Park, J.A. (Ed.), *Researching Art Education as Critical Cultural Inquiry*. Mijinsa, Seoul, pp. 90–101.
- Hall, L.K., 2019. All Our Relations: Mo'okūāhau and mo'olelo. In: Wilson-Hokowhitu, N. (Ed.), *The Past Before Us: Mo'okūāhau as Methodology*. University of Hawai'i Press, Honolulu, pp. 106–119.
- Hau'ofa, E., 2008. *We Are the Ocean: Selected Works*. University of Hawaii Press, Honolulu.
- Henry, E., Pene, H., 2001. Kaupapa Maori: locating indigenous ontology, epistemology and methodology in the academy. *Organization* 8 (2), 234–242.
- Hohepa, M., McNaughton, S., Jenkins, K., 1996. Māori pedagogies and the roles of the individual. *N. Z. J. Educ. Stud.* 31 (1).
- Ho'omanawanui, K., 2019. E Ho'i ka piko (returning to the center): theorizing Mo'okūāhau as methodology in and Indigenous literary context. In: *The Past Before Us: Mo'okūāhau as Methodology*. University of Hawai'i Press, Honolulu, pp. 50–68.
- Irwin, K., 1994. Mātauranga: Kathie Irwin. In: Brown, A. (Ed.), *Mana Wāhine: Women Who Show the Way*. Reed Publishing Ltd., Auckland, pp. 74–81.
- Kahakalau, K., 2019. Mā'awe Pono: treading on the trail of honor and responsibility. In: Wilson-Hokowhitu, N. (Ed.), *The Past Before Us: Mo'okūāhau as Methodology*. University of Hawai'i Press, Honolulu, pp. 9–27.
- Kovach, M., 2005. Emerging from the margins: indigenous methodologies. In: Brown, L., Strega, S. (Eds.), *Research as Resistance: Critical, Indigenous and Anti-oppressive Approaches*. Canadian Scholars Press; Women's Press, Toronto, ON, pp. 19–36.
- Kovach, M., 2009. *Indigenous Methodologies: Characteristics, Conversations and Contexts*. University of Toronto Press, Toronto, ON.
- Kovach, M., 2021. Second Edition. *Indigenous Methodologies: Characteristics, Conversations and Contexts*. University of Toronto Press, Toronto, ON.
- Lander, M., 2011. Te ao tāwhito/Te ao hou: entwined threads of tradition and innovation. In: Tamarapa, A. (Ed.), *Whatu Kākahu: Māori Cloaks*. Te Papa Press, Wellington, pp. 61–74.
- Lee, J., 2008. *Ako: Purakau of Maori Teachers' Work in Secondary Schools*. The University of Auckland, Auckland.
- Nepe, T., 1991. *Te Toi Huarewa Tipuna : Kaupapa Maori, an Educational Intervention System*. University of Auckland, Auckland.
- Pihama, L., 2001. *Tihei Mauri ora. Honouring our Voices: Mana Wāhine as a kaupapa Māori Theoretical Framework*. The University of Auckland, Auckland.
- Pohatu, T.W., 2011. Mauri – rethinking human wellbeing. *MAI Review* (3), 1–12.
- Roth, L., 1924. *The Maori Mantle*. The Scholar Press, Yorkshire.
- Royal, C.T.A., 2011. *Wānanga: The Creative Potential of Mātauranga Māori*. Mauriora-ki-te-Ao/Living Universe, Wellington.
- Smith, G.H., 1997. *The Development of Kaupapa Maori: Theory and Praxis*. The University of Auckland, Auckland.
- Smith, G.H., 2003. Kaupapa Maori theory: theorizing indigenous transformation of education and schooling. In: *Kaupapa Maori Symposium – NZARE/AARE Joint Conference*.
- Smith, H.J., 2017. *Whatuora: Whatu kakahu and Living as Māori Women* [Unpublished doctoral dissertation]. Auckland, New Zealand.
- Smith, H.J., 2019. Whatuora: theorizing “new” indigenous research methodology from “old” indigenous weaving practice. *Art/Res. Int.* 4 (1), 1–27. <https://doi.org/10.18432/ari29393>.
- Smith, H., 2020. “Cloaked in our aspirations”: Māori mothers and grandmothers weaving stories from their past for the future. *J. Indig. Wellbeing* 5 (1), 79–89. Retrieved from. https://journalindigenuswellbeing.com/journal_articles/cloaked-in-our-aspirations-maori-mothers-and-grandmothers-weaving-stories-from-their-past-for-the-future/.
- Smith, L.T., 1999. *Decolonizing Methodologies—Research and Indigenous Peoples*. University of Otago Press, Dunedin.
- Tamarapa, A., 2011. *Whatu kākahu*. Te Papa Press, Wellington.
- Te Awēkotuku, N., 1993. *Nga nui o Te Ra: Teaching, Nurturing, Developments of Māori Women's Weaving*. Te Waka Toi Council for Māori & South Pacific Arts, Auckland.
- Te Kanawa, K., 2008. *Mai i te ao kohatu. Weaving: an art form derived from mātauranga Māori as a gift from the ancestors*. *Te Kaharoa* 1, 138–151.
- Te Rito, J., 2006. *From the Margins of Te Whare Pora: Embracing Traditions of Innovation in Māori Textile Legacies*, Mātauranga Taketake Traditional Knowledge Conference, Indicators of Well-Being: Perspectives, Practices and Solutions.
- Webber, M., 2008. *Walking the Space Between: Identity and Māori/Pākehā*. NZCER Press, Wellington.
- Williams, H.W., 1997. *Dictionary of the Māori Language*, seventh ed. G. P. Publications, Wellington.
- Wilson-Hokowhitu, N., Meyer, M., 2019. Introduction: I ka Wā Mamua ed., *The Past Before Us: Mo'okūāhau as Methodology*. University of Hawai'i Press, Honolulu, pp. 1–8.

Indigenous storywork for education and research

Jo-ann Archibald Q'um Q'um Xiiem, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction & self-location	189
Story memory: making a cedar basket	190
Story memory: The Elders' fish camp	190
Story memory: becoming storywork ready	191
Reflections and discussion	192
Old Man Coyote's story: searching for the bone needle	192
Becoming storywork ready for education	193
Becoming storywork ready for research	194
Conclusion: my dear ones, the work is about to begin	196
References	197
Relevant websites	197

Introduction & self-location

Ey swayel, in the Halq'emeylem language means, good day. I write this chapter on the traditional and unceded lands of the X'məθk'əyəm (Musqueam), Skwxwú7mesh Úxwumixw (Squamish Nation) and Səlilwətaʔt (Tsleil-Waututh) Indigenous peoples of the Vancouver area of British Columbia (BC), Canada. Q'um Q'um Xiiem tel skwix- my Indigenous name is Q'um Q'um Xiiem. It means strong clear water. I am also known as Jo-ann Archibald. I am a member of the Stó:lō First Nation, whose traditional and unceded lands are in southwestern British Columbia, in the Fraser Valley area. I also have ancestry in St'át'imc First Nation, in the community of Xaxli'p, near Lillooet, located in the interior of BC. Stó:lō in the Halq'emeylem language, means, river. The term, unceded, means that First Nations have not "legally signed away their lands to the Crown or to Canada" (Wilson, 2018, p. 27).

This article begins with an introduction to who I am, as an Indigenous person. Indigenous scholarship often begins with the standpoint of identity, place, and community so that the reader has some contextual background about the author as a way to know where one comes from, culturally, geographically, politically, and more (Kovach, 2021). This form of self-location is part of an Indigenous methodology that is intimately linked to personal story as a foundational aspect of Indigenous knowledge. For those new to Indigenous people in Canada and British Columbia, I will provide basic background information to aid in understanding my self-location.

According to the 2016 Canadian Census, there were approximately, 1,673,785 Indigenous people who self-identified as either First Nations/Indian, Métis, or Inuit (FNMI), constituting 4.9% of the total Canadian population. FNMI peoples and their rights are acknowledged in Section 35 of the Canadian Constitution. The fastest growing segment of the Canadian population were young people under the age of 25, at 44% compared to 28% of the non-Indigenous population (Statistics Canada, 2017). In British Columbia, there were 270,585 FNMI, constituting 6% of the BC's total population (Statistics Canada, 2017). There are 34 First Nation languages spoken in BC in 203 First Nations communities (Dunlop et al., 2018). There is a rich diversity of Indigenous languages and cultural communities in British Columbia. I speak from the perspective of a First Nations person who grew up in a reserve community, attended public schools and university, and then became a school teacher and university professor. I did not have the opportunity to learn my Indigenous language, Halq'emeylem, in my home community or in educational systems that I attended due to colonial laws and policies, which will be discussed later. In this article, my personal stories exemplify aspects of Indigenous education, Indigenous ways of knowing, and an Indigenous methodology. I share three of my lived stories and one traditional story. Story characters such as Old Man Coyote and Owl also share their perspectives. Now to start with my personal story.

I grew up in Stó:lō territory, so I identify, culturally, as Stó:lō, which means people of the river. Soowahlie, is the name of the reserve community, where - as as child I had lots of freedom to play in the fields, woods, and streams with my cousins. I was fortunate to have spent time with my paternal grandparents, who lived beside us. As I reflect on my childhood, I feel fortunate to have had some place-based learning experiences with my immediate and extended family members. Fast forward to my adult years, as a beginning school teacher, I returned to the Stó:lō area, to teach primary level grades (2–3) in the public schools in 1974.

In the Stó:lō area, during the 1970s, a cultural revitalization movement was gaining momentum as a resistance response to a century of federal government laws, policies, and practices that focused on the assimilation, cultural genocide, and usurpation of lands of Indigenous peoples in Canada (Carlson, 1997). Some Stó:lō community leaders and Elders began to meet in each other's homes to revitalize and document, in written form, their Indigenous Halq'emeylem language, to discuss history from their perspectives, and to reclaim cultural values, stories, and practices (Archibald, 2008). In 1973, the Coqualeetza Cultural Education Center (Coqualeetza) was established as a non-profit charitable organization to serve Stó:lō communities and the wider Indigenous population with cultural programs and services (Coqualeetza Cultural Education Centre, 2010). The Coqualeetza Elders' Group became

a program of this center and those who were meeting previously, continued to meet and to revitalize Stó:lō culture and knowing through this program (Archibald, 2008).

In 1975, I began to attend some of Coqualeetza's cultural programs such as the Halq'emeylem language class, hide-tanning workshops, and a basket-making class. I was grateful to have these cultural-learning programs available because I had not participated in some of these cultural activities during my childhood. These types of cultural life-ways were forbidden through Canadian laws such as the Potlatch Ban where people could not gather for ceremonies or cultural activities and through education systems such as the Indian Residential Schools (Royal Commission on Aboriginal Peoples, RCAP, 1996) where children were separated from their families from the ages of 6 (sometimes younger) to 16, often to attend a school far from their home community for the school year. In these schools, numerous children experienced physical, sexual, and emotional abuse. They were forbidden to speak their Indigenous language, fed sparingly, were not allowed to practice their culture, and received a poor academic education (Truth and Reconciliation Commission of Canada, TRC, 2015). Many died while trying to run away from the school, to run back to their homes, or they died from disease such as tuberculosis. My mother and some of my grandparents attended Indian Residential Schools. At the time of this writing, many former Indian Residential School sites are using ground-penetrating technology to identify children's unmarked grave sites, after 215 unmarked graves were identified at the former Kamloops Indian Residential School in May 2021 (Snucins, 2021). I share this difficult impact of colonial history because all citizens of Canada have inherited this terrible legacy, but it is the generations of Indigenous people and their families, who have suffered and continue to experience inter-generational trauma and effects that are personal and at the same time, impact one's family and community. In the 1970s, I knew about some of this colonial history, but I had not thought critically about the prominent role of government and religions/churches in perpetuating colonization. As an adult and teacher, I was keen to learn more about Stó:lō cultural ways to strengthen my cultural identity.

Story memory: making a cedar basket

In 1975, I attended a summer basket-making class, sponsored by Coqualeetza. I was excited to attend because one of the master basket-makers, Teresa Michell, was teaching the course. She was also a very well-respected member of the Coqualeetza Elders' Group. I remember walking into the room where the class was held: a large room with lots of windows and tables and chairs. There were other women in attendance. I saw Elder Teresa Michell sitting at a table with a plastic tub of water filled with strips of cedar root material, and a small basket that she was weaving. Some of the women were seated at the tables with their cedar pieces. Some were weaving and others were sitting and visiting with them. I wondered if I had missed some of the instructions of the class because Elder Teresa was working away on her basket and not giving any whole or small group instructions. I just sat, watched, and waited, hoping that she would start telling us what to do. I kept watching as some women would sit with Elder Teresa and she would show them what to do, using her cedar pieces. Then they tried to copy her process. Eventually, I picked up the cedar pieces and began to follow what they were doing. When I needed help, I sat beside Elder Theresa to show her my work and to get her feedback. Sometimes, other women helped me.

This memory of experiential or "watch-then-do" learning has stayed with me for more than four decades. Even though I may not remember the details, what I remember, is the different teaching approach, where the Stó:lō teacher expected the learner to do lots of observation, then try the activity, and to seek help when needed. This approach is an Indigenous way of knowing (teaching and learning) that brings together, inter and intra-generational, experiential, observational, and mentoring approaches that are vastly different from direct and didactic instruction. I had expected the latter. My experiences with public school and university learning were more of the later. I had expected to sit down and that Elder Teresa would show us, the learners, a step-by-step approach where we would do each step with her as the teacher and where we as the learners, might interrupt her instruction, especially when we did not understand the directions. In the reflection and discussion section, I will return to this story and other stories that follow.

Story memory: The Elders' fish camp

Another story memory that I recall is about learning from Stó:lō Elders at their annual summer fish camp held near Yale, BC, where fishers donated sockeye salmon to the Elders for wind-drying and canning purposes. A few of us "young" people who were Coqualeetza staff and family members stayed at the fish camp to help with the fish preserving processes. I learned how to clean and cut the salmon so that it could be put on long thin poles, which were hung under small wooden wind-drying structures that had open sides and a roof. In the summer, the Yale area is windy and often has fairly hot temperatures; perfect conditions to let the salmon dry naturally. By that time, I had learned to watch the Elder do his/her work and then I tried to do it, with their watchful guidance. I already knew how to cut the fish for canning purposes, so I helped as needed. The Elders and the helpers would stay at the fish camp in tents or campers for a week or more, with some going home and returning the next day. I had a truck/camper, so I stayed with the Elders. I remember the laughter and teasing while we worked, and the stories that were told around the camp fire during meal times and in the evening. Mostly, I recall the feelings of being loved by the Elders, similar to the feelings of love from family members. I think that the Elders valued the interest and caring of the younger helpers, so that the Elders could continue the tradition of passing on their cultural values, knowledge and skills, albeit with newer tools. There were various Elders with particular talents and skills and all participated in the food preserving work accordingly. The Elders shared the teaching

responsibilities by using each other's 'gifts' or talents; no one person was considered the main teacher. The Elders' pedagogy at this fish camp continued to include inter and intra-generational, experiential, watch-then-do one-one or small group mentoring approaches. The Elders' love and caring made the young helpers feel valued and the inclusive approach ensured that everyone participated and contributed to the learning activity.

Story memory: becoming storywork ready

In 1976, in my role as a school district Indian Education Coordinator for elementary schools, I was invited by the Coqualeetza Cultural Education Center to assist with the development of an elementary school curriculum about Stó:lō culture (past and present). Coqualeetza was sponsoring this curriculum project, which was being led by the Coqualeetza Elders' Group. An advisory group, subsequently named the Stó:lō Sitel Advisory Committee, for which I was a member, was established that included educators, community members and Coqualeetza staff. This advisory committee provided policy guidance for this curriculum. Stó:lō *sxwoxwiyam* or traditional stories, and associated cultural knowledge became a core part of this curriculum that was prepared for the subject of social studies. At that time, the only subject area that provided an opening for using Indigenous topics as a locally developed curriculum was social studies.

The Elders met every Wednesday. A member cooked lunch for everyone and another member said some prayerful/thankful words before the food was shared. Then the Elders got to work. They worked with a linguist, the Coqualeetza staff, and other educators, such as myself. They would document the Halq'emeylem language, moving from the oral into a written form. This same approach was used with their stories where many remembered the stories told orally in Halq'emeylem during their daily life activities, before going to sleep at night, and during cultural events. During the documentation sessions, the Elders were translating the stories from Halq'emeylem to English and from the oral to the written form, which was in English. Some of that material was eventually tape-recorded, which was the main type of technology in use at that time. Those working on the curriculum ensured that the Elders held decision-making roles about all aspects of the curriculum process from the content (knowledge, stories) to learning approaches and published format.

With the Elders' permission, the meeting documentation and audio-tapes were subsequently used for developing the curriculum materials. This material has remained with Coqualeetza. A core purpose of this curriculum was to have it available and used in both public and band/Stó:lō schools. The Elders wanted to ensure that children, especially Stó:lō children, were able to learn about Stó:lō history and culture and ways of knowing from the perspective of Stó:lō people. The Elders named this curriculum, the Stó:lō Sitel. In Halq'emeylem, *sitel*, is a basket that holds treasured materials. The Elders believed that the Stó:lō curriculum contained treasured knowledge, values, and perspectives. Teresa Michell, a member of the Coqualeetza Elders and who is featured in the first story, wove a sitel basket to hold an original copy of the curriculum materials. At that time and still today, the majority of Stó:lō children attend public schools. The Elders felt that all children needed to learn about Stó:lō people so that a better intercultural understanding was developed amongst all students, for the long term goal of educating future citizens. The Elders' inclusive educational approach that began in the 1970s is a prime example of reconciliation as advocated 45 years later by the TRC: that students develop intercultural understandings through education (TRC, Call to Action 62, 2015).

The Elders also approved both the text and the artist's images for each story. Those of us working on the curriculum felt that it was our responsibility to ensure that the oral story was transferred to a new format, such as a published story book and that the story pedagogy, through teaching units and lesson plans were all satisfactory to the Elders. Today the term, authentic, is used when recommending types of Indigenous learning materials to use for educational purposes ([First Nations Education Steering Committee and First Nations Schools Association, 2016, 2021](#)). The Stó:lō Sitel curriculum is an example of authenticity. The Elders also fulfilled their cultural responsibility to pass on cultural knowledge and ways of knowing as they took the necessary time to remember, think about, talk about, or engage with particular stories. Each week when they would meet, the curriculum staff would show them what had been done with their story. They would either approve it or tell the staff what changes needed to be made.

This verification process also demonstrated respect because the curriculum staff wanted to ensure that the Elders felt comfortable and agreed with how their knowledge and stories were being used. In order to achieve this goal, we as the working group, took time to develop a respectful relationship with the Elders so that they could trust that we were carrying out our educational role, in good/ethical ways. Respect and responsibility often go hand-in-hand.

I left the Stó:lō territory in 1985 to work at the University of British Columbia's Vancouver campus with the Faculty of Education's Indigenous Teacher Education Program (formerly known as the Native Indian Teacher Education Program-NITEP) and then enrolled in a PhD program in 1989. During my PhD research, I was guided to learn more about the role and pedagogy of Indigenous stories, with a focus on Stó:lō and Coast Salish traditional and life experience stories. During the Stó:lō Sitel curriculum work, I learned to value the care and attention that the Elders demonstrated with documenting and using some stories for learning purposes. However, in my graduate student role, I felt that I needed to know more about Indigenous oral traditions, the role of storytelling, the relationship between stories and learning and teaching, and how colonization changed traditional Indigenous storied ways. To learn about these matters, I knew that it was necessary for me to return to my original teachers, the Stó:lō Elders, who were still working with Coqualeetza, to ask for their guidance in this next phase of my learning. This time, I focused on learning through the work of research.

My doctoral research learning journey spanned four years, learning from the Coqualeetza Elders, other Coast Salish Elders and Indigenous storytellers; learning from the small but growing literature written by Indigenous scholars and non-Indigenous allies

who worked respectfully and responsibly with Indigenous stories; and learning by doing/working with educators on other school curriculum that included stories as a core component and using stories in my university teaching practice. I was able to use some of the Elders' regularly scheduled meeting times where we discussed their experiences of listening to and living with stories throughout their life span; training methods for becoming a good speaker and storyteller; cultural contexts and protocols for stories, and their concerns about the use of and sustaining stories. We had many good talks over a one year period. As a result of these research and educational experiences, I developed an Indigenous storywork (ISW) framework that included seven principles of respect, responsibility, reverence, reciprocity, holism, inter-relatedness, and synergy. Indigenous storywork can be used for theoretical, methodological, and pedagogical purposes (Archibald, 1997). The origin of the name, storywork, comes from words that are said by a spokesperson at Stó:lō cultural gatherings, who knows about and facilitates the gathering's cultural work. To signal the start of the gathering, the spokesperson will say something like, "My dear ones, the work is about to begin." When the guests hear those words, they give their full attention to the work that is about to take place. I brought the concept of story together with work, so that when the name, Indigenous storywork, is heard or read, it signals that it is time to give serious attention to learning from and working with Indigenous stories.

Before I finished my doctoral dissertation, I returned to the Coqualeetza Elders to share what I had learned about the power and beauty of Indigenous stories; how stories could become like teachers; and how research could be carried out in ethical ways with Indigenous peoples' stories. I thanked them for their valuable guidance and later provided copies of my dissertation to Coqualeetza. My family also hosted a ceremony and feast to acknowledge the completion of my PhD and to thank those who had helped me on this journey.

It took many years before I felt comfortable and ready to turn my doctoral dissertation into a book. A demanding UBC work schedule, personal family responsibilities, and anxious feelings about how others would perceive Indigenous storywork kept me from proceeding with the book publication. I had continued to use the Indigenous storywork principles in my teaching practice and some research projects, so I believed that they were useful for educational and research purposes and could be used by others. Issues about disrespectful research were prevalent in the literature such as appropriation of Indigenous knowledge (Armstrong, 1993; Keeshig-Tobias, 1990; Maracle, 1992) and lack of Indigenous engagement in research decision-making and methodology (Smith, 1999). Despite these issues, my colleagues kept encouraging me to publish this work because they believed that Indigenous storywork contributed in positive ways to both research and education. They gave me useful suggestions for revising the manuscript that I produced. I kept revising the manuscript, until one colleague said that it didn't need any more revision and that it was time to submit it to a publisher. And it was sent. Finally in 2008, *Indigenous Storywork: Educating the Heart, Mind, Body and Spirit* was published by UBC Press.

Reflections and discussion

After sharing the three sqwelqwel-true stories, I had intended to write a reflection about what I had learned about Indigenous ways of knowing and Indigenous research from living and telling these stories. I will include reflective understandings of becoming storywork ready for both education and research purposes, through the use of the sqwelqwel stories, but there is an interruption that needs space here. Until now, I had resisted including a sxwoxwiyam-traditional story that has facilitated my understandings about education and research. I resisted because I have told and written about a particular Coyote story many times and wondered if instead, I should try another storied approach, which became the three sqwelqwel stories noted above. However, I was recently reminded that Elders and storytellers have often said that stories can be told over and over again. The Elders heard the same stories throughout their lives, so that the stories became a part of them, embedded in their heart, mind, body, and spirit. They also said that new ideas could surface from a story, depending on what the storyteller or listeners are open to and ready to experience. The Indigenous Trickster character of Coyote has become my critical story friend and helper throughout my doctoral journey and subsequent scholarly work. So, it is time to give Coyote space and voice in this part of the article. Before Coyote's story is told, I acknowledge that Eber Hampton of the Chickasaw Nation and university educational leader in Canada, told this story, which does not originate in Stó:lō territory. He eventually gave me permission to use this story and encouraged me to adapt it for my cultural context, so I renamed the Trickster, Old Man Coyote. Here is the story.

Old Man Coyote's story: searching for the bone needle

Old Man Coyote (OMC) had just finished a long, hard day of hunting. He had walked a long, long way through fields, up mountain trails, and along valley streams. But, he did not catch anything. He found a nice clearing to set up his camp for the night and started a fire for his meal. The fire burned brightly and after supper, OMC sat by the cozy warm fire and rubbed his tired feet from the long day's walk. He took his favourite moccasins out of his bag and noticed that there was a hole in the toe of one of them. He looked for his special bone needle to mend the moccasin but couldn't feel it in the bag.

Then Old Man Coyote started to crawl on his hands and knees around the fire to see if he could see or feel the needle. He went around and around the fire many times. Just then Owl came flying by and landed next to OMC. Owl asked him what he was looking for. Old Man Coyote told Owl his problem.

Owl said that she would help her friend look for the bone needle. After she made one swoop around the area of the fire, she told OMC that she didn't spot the needle. Owl asked OMC where he last used the needle. Old Man Coyote thought for a while then said that he used it quite far away, in the northern direction, to mend his jacket. Then Owl asked OMC why he kept going around and around the campfire when the needle clearly was not there. Old Man Coyote replied, "Well, it's easier to look for the needle here because the fire gives off such good light and it is warm and cozy." (adapted slightly, Archibald, 2008, pp. 35–36).

Many times, I have observed people acting like Old Man Coyote in educational systems, who get used to a way of thinking and doing and who do not really want to change their pedagogy or policies to create space for Indigeneity, which includes Indigenous ways of knowing, increasing the enrollment and retention of Indigenous students, teachers, and staff, and engaging Indigenous communities in meaningful decision-making roles. Calls for including Indigenous knowledges, values, history, languages, pedagogies and Indigenous peoples' involvement began through national policy documents in Canada such as Indian Control of Indian Education (National Indian Brotherhood, now known as the Assembly of First Nations, 1972), and continued with research resulting from the Royal Commission on Aboriginal Peoples (1996), and is still prevalent in the historical and current impact of Indian Residential Schools as portrayed in the Truth and Reconciliation Commission of Canada's Calls to Action (2015). At provincial and territorial levels, similar recommendations are also found (Archibald and Hare, 2021). Yet, some progress has been made to include Indigeneity in educational systems, but the overall progress is fairly slow and sporadic (RCAP, 1996; Archibald and Hare, 2021). At the same time, I have observed and worked with many Indigenous and non-Indigenous allies, at all levels of education whose work has included Indigenous perspectives and ways of knowing, especially through the use of Indigenous storywork (Archibald et al., 2019; Nicol et al., 2020). But Coyote figures it is time to highlight his role and impact with Indigenous storywork (Coyote's perspectives will be shown in italics).

Old Man Coyote. Finally, Q'um Q'um Xiiem (I will use her Indigenous name, which I prefer) has given me an opportunity to tell what I think about her sqwelqwel stories. It was really hard for me to be patient, to listen carefully, and to give my full attention to her stories. Besides, there was nothing about me in those stories, so it was hard for me to pay attention, although I am sure there were Coyote-like moments when people did not listen to each other, or tried to out-story one another, or tried to hurry and finish the stories. At least Q'um Q'um Xiiem found some stories about me and other Trickster characters that pointed her in the right direction for developing the Indigenous storywork principles- but it took her such a long time.

However, I was happy to travel with her on learning journeys about Indigenous storywork, especially, watching her learn to hear what the Elders were saying about the importance of stories to provide teachings if one is open and ready to engage with the story by using one's senses to address the heart, mind, body, and spirit. Usually, I am pretty good at the body/physical, but I don't like to think about what I have learned or what I could have done better, or how to recognize my own or others' feelings. Reverence still has me stumped! However, I am fortunate that the Owl continues to be my friend. She helps me in many ways. After staying near that cozy fire long enough, I knew that I would have to move away from the fire, to find the bone needle. Owl reminded me that I need to care for that needle so that I can use it when it is needed. Maybe that bone needle is like Indigenous stories, if we take care of them, they will sustain us. Or is it the other way around? Stories will take care of us, if we are open and storywork ready. Now, it is time to return to the sqwelqwel stories for educational purposes.

Becoming storywork ready for education

The basket-making and fish camp stories featured Stó:lō ways of knowing and Elders' pedagogy during cultural activities. The sqwelqwel stories are also a form of knowing and an example of an Indigenous pedagogy. When I attended various Elders' meetings and cultural activities, their talk was often in the form of sqwelqwel or true life-experience stories and sometimes included sxwox-wiyam or traditional stories. My own family ways of teaching and learning included similar approaches of watching and then doing. An action of respect is to listen and to observe, without interrupting the teacher or mentor. During the learning experience there is time to ask questions and opportunities to ask for help, when the time seems right. Knowing when the time is right, depends on the rhythm and difficulty of the activity. When someone interrupted or did not follow this pedagogy, I used to hear older people say to each other, "(that person) has no teachings," which I believed to mean that person did not know or practice the principles of respect, or that their needs were more important than others, or that they did not have patience, which were forms of disrespect. I wondered why Indigenous people would be disrespectful. During my graduate education I began to understand why. For decades, the predominance of Euro Western values and pedagogies in public educational systems, from kindergarten to post-secondary, have disallowed or marginalized Indigenous knowledges and pedagogies (Battiste, 2013).

My own experiences in public schools were such that I never felt proud of being Indigenous and when Indigenous topics were brought up, I could not identify with the Indigenous people or images being portrayed, since they were negative, fraught with problems, and often stereotypical. Yet, I lived an Indigenous life in an Indigenous reserve community. It was as if the Indigenous people being portrayed in the curriculum were not real. Even though there were social and other problems that people in my reserve community experienced, they were real people and family members cared for them. In relation to the time period that I attended public school, the stereotypical and racist portrayals of Indigenous people or omission of Indigenous contributions to Canada in public school curricula, particularly in the subject area of social studies were identified in evaluation research as pressing issues from the 1960s to 1970s (McDiarmid and Pratt, 1971; Manitoba Indian Brotherhood, 1974; Werner et al., 1977). A similar recommendation stemming from these studies emphasized the need for Indigenous people to be involved in developing and implementing Indigenous themed-curricula. As the story about the Stó:lō Sitel curriculum shows, the Coqualeetza Elders were living this

recommendation. I think that my commitment to improving education through systemic change traces back to my negative experiences in public school.

I persevered through the grades and did well academically to graduate high school. Upon reflection, I was fortunate to have had a positive home life and experienced many community-oriented activities, which I attribute to facilitating my educational success in public schooling. However, my schooling did not contribute to my Indigenous identity development or knowledge of Indigenous matters. Often, I felt like I lived in two very different worlds: one was the school life and the other was my family/community life. Now, I understand why some Indigenous people may act like they “have no teachings” because their public schooling did not accommodate Indigenous epistemology and pedagogy. Colonial laws and policies disrupted the family and community continuity of Indigenous knowledge and ways of knowing for many generations. They may have been adopted, fostered, and put in care with families that were not Indigenous and did not value or appreciate Indigeneity (RCAP, 1996; TRC, 2015).

The inter-generational impact of Indian Residential Schools in Canada has become more widely known in Canadian society since the release of the national Truth and Reconciliation Commission of Canada’s report and Calls to Action in 2015 and then the recent deluge of media reports about the growing numbers of unmarked graves being found at former residential school sites during 2021 (Malone, 2021). Another national inquiry, this time, the *National Inquiry into Missing and Murdered Indigenous Women and Girls* (2019) adds to the growing Canadian public’s interest, including post-secondary students, to know the “truth” of Indigenous peoples’ experiences, not only of residential schooling, but of on-going forms of colonization occurring through various sectors of society. I believe that the heightened interest and moral need to do better in relation to Indigenous peoples stems from the powerful personal lived stories that residential school survivors and Indigenous women and their families have shared. They have enacted their responsibility to tell their stories so that others will know and understand the intergenerational impact of colonization and that those of us who listen and understand the stories will contribute to making positive changes in the systems in which we are located. Old Man Coyote wonders how he can make systemic change.

Old Man Coyote. So much to consider in becoming storywork ready for educational purposes! I was quite happy staying by that cozy fire. Then Owl said that I have to work on myself before using those ISW principles, especially respect, responsibility, reverence, and reciprocity. In response, I figured that my knowledge of Indigenous ways of knowing and Indigenous pedagogy could be strengthened. My first step was to read books, especially ones Indigenous Elders had written, alone and with others, that focused on their stories. I was hoping that they included Tricksters like me. I sure liked “Khot-La-Cha: The Autobiography of Chief Simon Baker,” which was compiled by Verna J. Kirkness (1984). He learned a lot from his grandmother and used humor throughout his life story, even though there were troubling times that he experienced. He was a well-respected and resilient community leader. Khot-La-Cha respected his grandmother’s ways of teaching him and then carried out his responsibility to share this knowledge and to mentor the younger generation. Q’um Q’um Xiēm has said that Khot-La-Cha was always willing to help the younger generation because he knew that by mentoring them, he was continuing Indigenous ways of knowing. Since I spend lots of time on the unceded lands of the Skwxw7mesh Úxwumixw-Squamish Nation, where Khot-La-Cha was from, I thought that I should learn about Skwxw7mesh Úxwumixw history and perspectives as an action of respect and responsibility.

I was amazed by Kwulasulwut, Ellen White’s book, “Legends and Teachings of Xeel’s, The Creator” in which she tells four traditional stories that were passed from her ancestors, to older family members, to her. She worked with her daughter Vicki White, thereby, continuing an inter-generational pedagogy. In this book, first published in 2006 and reprinted in 2018, Kwulasulwut and Vicki talk about the story to provide contextual and cultural information for the reader and to share some ideas about the story to spark meaning making for the reader. Kwulasulwut from the Snuneymuxw First Nation of Nanaimo, BC, was a wonderful storyteller, healer, and educator. She demonstrated all of the 4R principles: respect for her ancestors’ stories and teachings by learning and authentically presenting them; responsibility of mentoring and guiding others to learn from the stories; reverence for the spiritual nature of the stories; and reciprocity by continuing to ensure that the stories are told and used, thereby, continuing the cycle of learning through stories. Q’um Q’um Xiēm also said that Kwulasulwut helped her appreciate the ‘core’ of a story (White and Archibald, 1992): somehow that core knowledge makes the story timeless, maybe it has to do with knowing the values, philosophy, or teachings of the story?

In their books, these two Elders shared traditional ways of knowing and learning, which are similar to the first two stories told by Q’um Q’um Xiēm about basket-making and fish camp, although she also calls them examples of Indigenous pedagogy. But Khot-La-Cha also talked about his difficult experiences at Indian Residential School, and what he and other kids did to survive. There are many more new books and films where Indigenous people tell their tragic stories, yet, they also demonstrate their resistance and resilience to this form of colonial education (Sellars, 2013; Webstad, 2019). These are the ‘truths’ that all Canadians must know and that must be placed within school curricula, so that the next generation of Canadians has a better understanding of the impact of colonization.

Then, I started reading more critically about the power of EuroWestern educational systems and a term that Indigenous scholar Marie Battiste used to describe this type of domination: cognitive imperialism (2013, p. 158). She also made me think about the need to be engaged in decolonization of education. Wow, that perspective made me think hard. Now, I am tired and need to rest. But, I am getting off track here, and Owl reminds me that I need to stay focused on Indigenous stories. But I tell her that all this reading has added to my basic knowledge about the struggles of decolonizing my thinking and actions so that I can be more open to letting the stories sustain me. Owl wants to take me on a journey to see what other Indigenous storyworkers are doing. I agree because I want the stories to look after me. But first, Q’um Q’um Xiēm has more to say about how research sparked the notion of becoming storywork ready for this purpose. I hope that she does not expect me to do research too!

Becoming storywork ready for research

In this section, I discuss the story memory of becoming storywork ready. This lengthy story includes my two journeys of learning that I experienced with the Coqualeetza Elders. Even though the first journey focused on revitalizing Stó:lō sxwoxwiyam stories for

educational purposes, I discuss that experience in this section to illustrate notions of research readiness. In fact, the Elders and others involved in the curriculum project undertook activities that involve methods of qualitative research, such as convening discussion/focus groups, asking open-ended and semi-structured questions, analysing the responses for common themes/topics, verifying and portraying these themes/findings in written and other material, and then disseminating the findings more broadly (Mishler, 1986; Hammersley and Atkinson, 1983). However, the group also carried out activities that exemplify what we now call, Indigenous methodologies, such as relationship building, ensuring that Indigenous people have decision-making roles throughout the research process, using Indigenous knowledge to structure the methods, analysis, and dissemination, and to ensure that research is beneficial to Indigenous people and their communities (Battiste, 2013; Kovach, 2021; RCAP, 1996; Smith, 1999, 2012; Wilson, 2008).

I developed a working relationship with the Coqualeetza Elders and staff members. My grandfather was a member of this group, so the Elders knew me through my family kinship and community location, but they did not know me personally. I recall feeling intimidated meeting with the Elders, at first. There seemed to be a large group of about 25. Many I did not know. Throughout my childhood and early adult years, I had learned to show respect and reverence to Elders who carried out their responsibilities in good, caring ways. Not all older people are Elders who take on a teaching/mentoring role and who live good teachings. Young people were expected to help and serve Elders food, if they indicated that they wanted help. Many preferred to get their own food and refreshments. But, they showed appreciation for the care shown by the younger people. During the more formal part of the Elders' session, I had the role of asking questions for the curriculum work. I recall some anxious moments when there was silence, or when the Elders did not respond to my questions, and sometimes they would speak quietly to the person next to them. One of the curriculum staff, who had very positive relationships with the Elders would help me by talking more about the intent of the question. Yet, there were silences. I learned to observe these silences and the Elders' reactions to the questions. Eventually, I became comfortable with silence because I knew that they were thinking about the question and forming their response or that they did not have anything they wanted to contribute at that moment. Sometimes, a week or more later, an Elder would bring up the question that was asked earlier and would share thoughts about it. Elders would mention that they had put certain knowledges to sleep and that it took a while to awaken their memory of it. I attributed their perspective of putting knowledge to sleep in response to colonial methods of oppression and denial of Indigenous knowledges. Perhaps they hoped for a future time to re-awaken their ways of knowing in a safe environment. Revitalization of Indigenous ways of knowing need time, additional effort and patience, cooperative engagement, and questioning through a decolonizing process (Archibald, 2008; Kovach, 2021; Smith, 2012).

On my second journey of returning to the Coqualeetza Elders for my PhD research, I first focused on re-establishing a responsible learning relationship with them. It had been over five years since I had worked with the Elders on the curriculum project. I followed ethical procedures for seeking their permission for me to carry out this research project with them, via their Elders' Council, as well as acquiring university ethics approval. The Elders' Council included a few members of the larger Elders' group who made decisions of this nature. I met with them to discuss my motivation and purpose for learning more about Stó:lō storytelling and ways of knowing through story and to seek their guidance and direction on how to proceed with these topics. I recall feelings of relief and excitement when they agreed to work with me, and this time, I felt more comfortable because of the previous relationship I had with them and that they voiced how important Stó:lō stories and the oral tradition were to them and their families and communities. This project was a way to continue passing on their teachings. Here is what I wrote in my research journal:

[The Elders Council] ... liked my work and said it was important to do; so many Elders have left us. We talked about the importance of stories for teaching and how they help us to live right-to be good human beings. Such a relief not to be doing this important work alone, but to have guides. (Archibald, 2008, p. 61).

During the second research period that I spent with the Coqualeetza Elders, I was better prepared and more confident to talk to them about my ponderings and questions, and to listen to their storied responses with my "three-ears." Various Elders have used this expression to indicate a deep listening with the ears to hear with and to bring that listening together with the heart/emotions/spirit. The Elders often told sqwelqwel stories in response to my questions. After our discussion sessions, I would go away, transcribe the talk/stories, and listen to it many times, to discern both common and different ideas, issues, and teachings. This term, teaching, is often used to mean Indigenous values, beliefs, knowledge, and way of knowing-all wrapped up in one concept. Two Indigenous storywork principles became evident in my meaning-making with the Elders' lived stories: holism and inter-relatedness. I identified parts of their stories that touched my heart (emotions), mind (intellect), body (actions), and spirit (spiritual) for a holistic beginning. I also thought about the stories in relation to their historical and current day experiences, and the cultural and education contexts for working with stories as an inter-related understanding.

The range of stories and perspectives related to what makes a good storyteller, becoming storytellers, cultural and place-based contexts for story learning, and issues such as breaking cultural rules, reclaiming responsibility, training new storytellers, and keeping the stories alive (Archibald, 2008, pp. 65–82). After I drafted a chapter about the aforementioned topics, I met with the Elders' Council to present the ideas and to receive their feedback and approval of what I had written. I also had given those who were quoted the transcript of their words and received their approval for using their words as quotes. The verification process took a few meetings over a one-year period. This second learning journey took about four years to complete. As mentioned previously, I also engaged in other research activities with others, apart from the Coqualeetza Elders. I also worked at UBC while completing this PhD research, which also added time for completion. I mention the time period and other factors for completion

so that graduate students who work while undertaking their graduate research will know that doing both is possible, but it does take longer to complete the graduate research. For graduate students thinking about undertaking storied aspects of research, I share this research reflection to indicate that researchers must expend more time and effort to become research ready. Living the ISW principles of respect, responsibility, reverence, and reciprocity is one way to become storywork ready.

Conclusion: my dear ones, the work is about to begin

Earlier, the words, *my dear ones, the work is about to begin*, signaled that it is time to give full attention to the work about to take place at a gathering. This time, these words are used to signal that it is time to get to work. The sqwelqwel stories, ponderings by Old Man Coyote, and my reflections focused on becoming storywork ready for engaging Indigenous traditional and life experience stories for educational and research purposes. The Indigenous storywork principles of respect, responsibility, reverence, and reciprocity need to be lived to understand and use them. The sqwelqwel stories of learning from Stó:lō Elders contained these principles when they used Indigenous pedagogies during cultural lifeway activities. For example, respect was demonstrated toward both the Elder teacher and younger learner. In the basket-making story, I as the learner, was given respect and agency to participate in the making of the basket, when I was ready. I could watch as long as I needed to and then begin when I was ready. The other women learners also used their agency and enacted responsibility to help me as needed. Reverence for the cedar root material was not discussed in the story, but I know from the Coqualeetza Elders' teachings that special words of thanks were said to the tree before taking its roots and bark for making baskets and other things (Stó:lō Site Curriculum, 1983). The roots and bark were carefully taken so that the tree continued to live: all human actions were carried out with reverential feelings. Elder Teresa's way of teaching, the women learners, and my sqwelqwel story are all forms of reciprocity, a way of giving back what we were fortunate to learn, and ways of continuing Indigenous knowledge and ways of knowing for current and future generations.

Recently, I shared this basket-making story memory with a colleague and friend, a non-Indigenous educator-ally. For many years, she and I have worked together on culturally responsive educational approaches that bring together, Indigenous storywork and math education. She commented on the agency dynamic of the watch-then-do approach as the "tremendous respect for the learner and trust in the teacher seen in the space for learners to pick up the materials when ready to start." She contrasted this approach with a common pedagogy of "I do, we do, you do ... [which] assumes everyone watches then everyone does together, then everyone on their own. There is little room for learner agency with learners and the teacher [has the major responsibility of] trying to keep everybody together" (personal communication, Dec 3, 2021). I appreciated her ideas, which sparked further contemplation from me. I had not considered learner agency in that story. *Old Man Coyote says that Owl is like Q'um Q'um Xiiem's friend who gave her new ideas to consider.*

Old Man Coyote. I understand that a concluding section in an article often summarizes and wraps up/closes the article, but I wonder why Q'um Q'um Xiiem says it is time to begin the work, after I have tried to pay attention to all that has been said so far. I even used my three ears! Didn't she say that we had to give our full attention to what was to take place? Owl says she wants to finally voice her opinion.

Owl. Yes, Old Man Coyote, Q'um Q'um Xiiem said both those things. I have observed your efforts in starting to live some of those Indigenous storywork principles by learning what Indigenous Elders and others have written about such as their life stories, the impact of colonization on generations of Indigenous people and what that has done to their families and communities, and even engaging in critical thoughts about the dominance of Euro Western knowledge and pedagogy in educational systems. You have shown responsibility and used your volition to find, read, and learn from this scholarship. I especially liked that you began with scholarship originating from the traditional and unceded territory on which you spend lots of time. I believe that you have shown respect by being open to learning and thinking. Understanding and practicing reverence will be a challenge, but I know that you can do it. I look forward to seeing how you put reciprocity into practice in the future. Now, I think that Q'um Q'um Xiiem is encouraging us both to travel on Indigenous storywork pathways either alone or together. But first, I want to share my perspective about the sqwelqwel stories.

I have appreciated learning more about Indigenous ways of knowing that can also be called Indigenous pedagogies, characterized as: inter and intra-generational, experiential, watch-then-do, and storied. These pedagogies were in all of Q'um Q'um Xiiem's stories. She has told me that these are not the only forms of Indigenous pedagogy and that it is necessary to learn culture-specific Indigenous pedagogy from communities. A part of me is very sad that the Elders who lived and taught these ways of knowing have passed to the Spirit World. Another part of me is happy, that they continued to teach these ways to the younger generation, who are now passing that knowledge to others. But, it is important to recognize that it is a continuous struggle to revitalize Indigenous knowledge systems that include stories as a core component after generations of colonial oppression and dominance of Euro Western ways in educational systems. However, I know that many Indigenous educators and non-Indigenous educator allies have worked for many years to decolonize these educational systems to create good learning spaces for Indigenous knowledge systems so that truth and reconciliation as advocated by the TRC Calls to Action can take place (2015). However, this work has to include Indigenous people in meaningful ways as decision-makers and teachers for both education and research. I often hear the phrase, "nothing about us, without us" being said across Canada.

I hope that those who are interested in using Indigenous storywork for educational and research purposes will try to live those 4Rs of respect, responsibility, reverence, and reciprocity in their own lives and practices as a beginning step to becoming storywork ready. The other ISW principles of holism, inter-relatedness, and synergy were present in the stories and reflections, although they were not emphasized as much as the 4Rs. I hope that Old Man Coyote understands why Q'um Q'um Xiiem used this approach.

Old Man Coyote. Yes, I know that I need to get myself ready to do Indigenous storywork in a good way, as I have heard that term used by lots of Indigenous people. I am sure there are various meanings of that term. Right now, for me that term means having good thoughts about what I am doing and what I am learning so that those good thoughts will seep into the work I am doing with stories and then result in beneficial learning. Maybe I will do a better job of looking after that bone needle in the future. Hey, Owl, didn't you say that we were going to see what other storyworkers have been doing? Maybe they have a fancy kit to share with us. Oh-oh, Owl is giving me the look, that means I should know better.

Thanks to Old Man Coyote and Owl for sharing their story and perspectives. I will have to tell OMC that the growing cadre of educators involved in Indigenous storywork resist the approach of preparing ready-to-use kits. The Coqualeetza Elders would smile at OMC and then continue to do their work, while OMC watches.

References

- Archibald, J., 1997. Coyote Learns to Make a Storybasket: The Place of First Nation Stories in Education. PhD thesis. Simon Fraser University, Canada.
- Archibald, J., 2008. Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit. UBC Press, Vancouver.
- Archibald, J., Hare, J., 2021. Thunderbird is rising: Indigenizing education in Canada. In: Graham, K., Newhouse, D. (Eds.), *Sharing the Land, Sharing a Future: The Legacy of the Royal Commission on Aboriginal Peoples*. University of Manitoba Press, Winnipeg, pp. 133–188.
- Archibald, J.Q.Q.X., Lee-Morgan, J.B.J., De Santolo, J. (Eds.), 2019. *Decolonizing Research: Indigenous Storywork as Methodology*. Zed Books, London.
- Armstrong, J. (Ed.), 1993. *Looking at the Words of Our Peoples: First Nations Analysis of Literature*. Theytus Books, Penticton, BC.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. Purich Publishing, Saskatoon, SK.
- Carlson, K., 1997. *You Are Asked to Witness: The Sto:lo in Canada's Pacific Coast History*. Sto:lo Heritage Trust, Chilliwack, BC.
- Dunlop, B., Gessner, S., Herbert, T., Parker, A., 2018. *Report on the Status of B.C. First Nations Languages*, third ed. First Peoples' Cultural Council, Brentwood Bay, BC.
- First Nations Education Steering Committee (FNESC) and First Nations Schools Association (FNSA), 2016. *Authentic First Peoples' Resources K-9*. FNESC & FNSA, West Vancouver, BC.
- First Nations Education Steering Committee (FNESC) and First Nations Schools Association (FNSA), 2021. *Authentic First Peoples Resources for Grades 10-12 and Adult Learning*. FNESC & FNSA, West Vancouver, BC.
- Hammersley, M., Atkinson, P., 1983. *Ethnography: Principles in Practice*. Tavistock, London.
- Keeshig-Tobias, L., 1990. Stop stealing Native stories. *Globe and Mail*, 26 January.
- Kirkness, V.J. (Ed.), 1984. *Khot-La-Cha: The Autobiography of Chief Simon Baker*. Douglas & McIntyre, Vancouver.
- Kovach, M., 2021. *Indigenous Methodologies: Characteristics, Conversations, and Contexts*, second ed. University of Toronto Press, Toronto.
- Malone, K.G., 2021. Children Who Never Returned from Residential Schools Newsmaker of the Year: CP Poll. *The Canadian Press*, 15 December, 2021.
- Manitoba Indian Brotherhood, 1974. *The Shocking Truth about Indians in Textbooks: Textbook Evaluation*. Manitoba Indian Brotherhood, Winnipeg.
- Maracle, L. (Ed.), 1992. *Give Back: First Nations Perspectives on Cultural Practice*. Gallerie, North Vancouver, BC.
- McDiarmid, G., Pratt, D., 1971. *Teaching Prejudice*. Ontario Institute for Studies in Education, Toronto.
- Mishler, E., 1986. *Research Interviewing: Context and Narrative*. Harvard University Press, Cambridge, MA, and London.
- National Indian Brotherhood, 1972. *Indian Control of Indian Education: Policy Paper Presented to the Minister of Indian Affairs and Northern Development*. National Indian Brotherhood, Ottawa, ON.
- National Inquiry into Missing and Murdered Indigenous Women and Girls, 2019. Executive summary of the final report : national inquiry into missing and murdered Indigenous women and girls. Available at: https://www.mmiwg-ffada.ca/wp-content/uploads/2019/06/Executive_Summary.pdf. (Accessed 13 December 2021).
- Nicol, C., Archibald, J.Q.Q.X., Gianfield, F., Dawson, A.J. (Eds.), 2020. *Living Culturally Responsive Mathematics Education With/in Indigenous Communities*. Brill/Sense, Boston.
- Royal Commission on Aboriginal Peoples, 1996. *Report of the Royal Commission on Aboriginal Peoples*. Vol. 3, *Gathering Strength*. Canada Communications Group, Ottawa.
- Sellers, B., 2013. *They Called Me Number One: Secrets and Survival at an Indian Residential School*. TalonBooks, Vancouver, BC.
- Smith, L.T., 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books, London.
- Smith, L.T., 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*, second ed. Otago University Press, New Zealand.
- Snucins, A., 2021. *Remains of 215 Children Found at Former Kamloops, B.C. Residential School*. The Canadian Press, 27 May 2021.
- Statistics Canada, 2017, October 25. *Aboriginal peoples in Canada: key results from the 2016 Census*. Daily. Available at: <https://www150.statcan.gc.ca/n1/en/daily-quotidien/171025/dq171025a-eng.pdf?st=zUdme2XF>. (Accessed 18 December 2021).
- Sto:lo Sitel Curriculum, 1983. *Upper Sto:lo interaction: teachings from our Elders*. Coqualeetza Education Training Centre, Sardis, BC.
- Truth and Reconciliation Commission of Canada, 2015. *Honouring the truth, reconciling the future: summary of the final report of the Truth and Reconciliation Commission of Canada*. Available at: http://www.trc.ca/assets/pdf/Honouring_the_Truth_Reconciling_for_the_Future_July_23_2015.pdf. (Accessed 13 December 2021).
- Werner, W., Connors, B., Aoki, T., Dahlie, J., 1977. *Whose Culture? Whose Heritage? Ethnicity with Canadian Social Studies Curricula*. Centre for the Study of Curriculum and Instruction, University of British Columbia, Vancouver.
- Webstad, P., 2019. *Phyllis's Orange Shirt*. Medicine Wheel Education, Victoria, BC.
- White, E., 2006. *Legends and Teachings of Xeel's, the Creator*. Pacific Educational Press, Vancouver.
- White, E., 2018. *Legends and Teachings of Xeel's, the Creator*. Theytus Books, Penticton, BC.
- White, E., Archibald, J., 1992. Kwulasulwut s yuth: Ellen White's teachings. *Can. J. Native Educ.* 19 (2), 150–164.
- Wilson, K., 2018. *Pulling Together: Foundations Guide*. BC campus, Victoria, BC. Available at: <https://opentextbc.ca/indigenizationfoundations/>. (Accessed 17 December 2021).
- Wilson, S., 2008. *Research Is Ceremony: Indigenous Research Methods*. Fernwood, Halifax.

Relevant websites

Coqualeetza Cultural Education Centre, 2010. *About Us* [online]. Available at: <http://www.coqualeetza.com>. (Accessed 16 December 2021).

Seeing through the smoke: creating space for Indigenous Knowledge in science education

Daniel Lipe, Cal Poly Humboldt University, Environmental Science Management, Arcata, CA, USA

© 2023 Elsevier Ltd. All rights reserved.

Introduction	198
Diverse frames of reference	198
Indigenous Knowledge Systems	198
Western science knowledge systems	199
Western settler colonialism	199
Impact of diseases	199
Settler greed	200
Shifting land management	200
Exclusion of indigenous land management	200
Protection versus sustainable resource utilization	200
Different perspectives for fire management	201
Indigenous fire and environmental management	201
Native American land management today	203
References	205

Introduction

When what counts as science and as ecological knowledge is narrowed to only include Western settler perspectives, scientific understandings, conversations, education, and access to environmental knowledge are limited to only one view. This paper discusses Indigenous land management practices outside of the contexts of settler colonialism in the United States. In addition, it explains the importance of utilizing diverse knowledge systems and perspectives within environmental education, on the ground land management practices, and land management policy decisions.

This article begins by describing why there is a need for diversity in science and natural resource education. I define Indigenous Knowledge Systems (IKS) and include the impacts that settler colonialism has had on Indigenous peoples of the United States. Utilizing fire management as an example, I illustrate the limitations and consequences of utilizing one scientific perspective in environmental land management. I also discuss settler colonialism from a land management perspective and how such an approach removes the Indigenous peoples from place. Removal includes not only the people but also their worldviews including those relating to land management practices. I end with a discussion the Sustainable Community Food Systems Bachelors of Applied Science (SCFS-BAS) degree program at the University of Hawaii, as an example of how curricula and programs can be developed in order to change higher educational science education through the inclusion of Indigenous perspectives.

Diverse frames of reference

Over the years, research has shown that groups working on complex multifaceted questions and issues benefited by having intellectual and cultural diversity (Adler, 2002; Mazur and Biatostocka, 2010; Page, 2007). According to Kincheloe and Steinberg (2008), "The more we understand about the world, the more complex it appears to be ... In this sense, different frames of reference produce multiple interpretations and multiple realities" (p. 138). As issues increase in both size and complexity, so does the need for diverse culturally intellectual frames of reference for identifying solutions to problems. Diverse frames of reference allow individuals to see issues through multiple perspectives, increasing their ability to come up with more complete and encompassing answers. In describing the importance of multiple, diverse perspectives, Kincheloe and Steinberg (2008) state, "In a sense, the single photograph of Cartesian thinking is replaced by the multiple angles of the holographic photograph" (p. 139). Diversity within sciences is likely to be as important as environmental and species diversity is to ecosystems.

Indigenous Knowledge Systems

Through generations of careful observation and hands-on scientific management, Indigenous peoples have and continue to impact local environments. Indigenous peoples of the United States understood and maintained their environmental knowledge according to the specific environments where they lived. In this way, IKS are as diverse as the environments and people who have developed them.

Indigenous peoples around the world have been deeply rooted in their homelands for thousands of years. Over time Indigenous peoples have created their own ontological understandings and knowledge bases, explaining their homelands and worldviews through oral traditions passed down from generation to generation (Barnhardt and Kawagley, 2005; Berkes, 1999; Cajete, 2000; Happynook, 2004; Lake, 2007; Rigney, 2001). Barnhardt and Kawagley (2005) explain, "The Depth of Indigenous knowledge rooted in the long inhabitation of a particular place offers lessons that can benefit everyone, from the educator to scientist, as we search for a more satisfying and sustainable way to live on this planet" (p. 9). Research suggests that integrating IKS alongside Western science practices is essential for maintaining global cultural diversity and the biological diversity connected to traditional management systems (Bohensky and Maru, 2011). Western scientific researchers are discovering and are growing more accepting with combining these two unique systems of knowledge that were once thought to be incompatible. For example, Bohensky and Maru (2011) state, "Co-management arrangements that allow knowledge to be integrated through collaboration can build social as well as ecological resilience" (p. 2). Integrating IKS with Western science can lead to more complete answers, because the solutions come from a variety of perspectives and also serve a greater range of invested stakeholders. Within the past 15 years there has been some progress made through participatory action research (PAR) projects that seek to include community knowledge into on-the-ground environmental research and educational projects. However, looking at a historical level the process of making change through collaborative efforts has been a very slow process.

Western science knowledge systems

Western science can and has included as diverse a set of definitions as Indigenous knowledge, incorporating everything within the natural world including social and cultural understandings. Western science information has been obtained primarily through the isolation of individual components. Webster's New Collegiate online dictionary defines science as knowledge attained through study or practice, covering truths about nature's general laws and has been tested through the scientific method. Since its conception western science has become the dominating perspective excluding many other ways of knowing as untruthful. The notion that all science knowledge needs to be tested through the scientific method has allowed western science to dominate science understandings. As western science increased in power it has also continued to narrow what can be categorized as scientific knowledge to only include information that is testable, repeatable, observable, and falsifiable (Shuttleworth, 2009). These same perspectives of western science knowledge systems being superior and allowing for the elimination of all other ways of knowing is the same notion that European colonizers utilized as a means to try to remove Indigenous peoples North America.

Western settler colonialism

Settler colonialism is a form of colonialism in which the colonizers come to stay. As noted by Tuck and Yang (2012) "settler colonialism is different from other forms of colonialism, in that the settlers come with the intention of making a new home on the land, a homemaking that insists on settler sovereignty over all things in their new domain" (p. 6). By design settler knowledge becomes the dominant form of knowledge, replacing all other forms of knowledge, including Indigenous Knowledge Systems.

Impact of diseases

While settler colonial intent and actions of destruction upon Indigenous peoples was by no means an accident (Tuck and Yang, 2012) it is important to understand how diseases at the point of impact affected most Indigenous peoples and their knowledge systems. Utilizing North America as an example one will learn that as population numbers of Europeans increased in America, so did the number of European-born pathogens (Carlos and Lewis, 2012). Diseases devastated many Native American populations causing loss of life and cultural knowledge, including land management practices. European diseases were the first devastating and immediate impacts on Indigenous peoples and their environmental management (White and Cronon, 1988).

Indigenous population counts and deaths associated with disease are debatable. Indigenous populations at the time of contact to the United States have been estimated at numbers as low as one million and as high as 18 million and every number in between (Carlos and Lewis, 2012). However, by the early 20th century Indigenous populations in the US and Canada have been reported at around 500,000 (Ubelaker, 1988). In general, tribes least affected lost up to 50% of their populations while others suffered much more devastating effects. Communities held a considerable amount of intellectual knowledge understood by all. In addition, there was family knowledge held within families. As diseases spread through communities it killed entire families and often communities. Consequently, unique knowledge held within each family was lost.

Some of the most unfathomable acts of settler colonialists involved utilizing many of the initial diseases that had already been proven to be destructive. It has been proven that settler colonialists' purposefully utilized smallpox blankets as a means to kill off Indigenous populations. According to Valencia-Weber (2003),

Smallpox blankets infected American Indians as the result of intentional acts where the donor (Euro-American colonial settlers) knew of the deadly microbes ... the concern here is for the intentional in the relationships mandated when Euro-Americans encountered the Natives and had to deal with mutual, as well as conflicting interests (p. 406).

Maliciously conceived settler colonialist's acts such as the use of germ warfare to remove Indigenous peoples from their homelands illustrate settler colonialist intentions to do anything possible to take over the Americas.

Settler greed

As settler populations continued to increase so too does their greed and desire for land and resources. Increased want for land within what is now the US came at the cost of the Indigenous peoples who continued to be removed, relocated, and decimated (Pevar, 2012). Policies such as the 1830 Indian Removal Act and the 1887 Dawes Act supplied additional lands through the removal of many Indigenous peoples from their traditional homelands (Pevar, 2012). In 1830 President Andrew Jackson put the Indian Removal Act into effect. President Jackson believed that Native Americans were a threat to American settlers and society and needed to be removed in order for American settlers to take their place in the newly forming United States (Bradey, 2011). In 1838 over 17,000 Cherokee men, women, and children in the Carolinas, Georgia, and Tennessee were forcefully taken and placed in stockades to await removal to Oklahoma country (Mooney, 1992). Provided with 645 wagons, the removal of the more than 13,000 remaining Cherokee people began in October 1838 and ended in March 1839. It is estimated by Mooney (1992) that more than 4000 Cherokee died as a direct result of removal. Acts like the one described above were not limited to only the Cherokee tribe. It is estimated that over 15,000 Indigenous peoples died on the Trail of Tears (Pevar, 2012).

In 1887 the US passed the General Allotment Act, also known as the Dawes Act (Pevar, 1992). Pevar (1992) describes the purpose of the Dawes Act, "... to break up tribal governments, abolish Indian reservations, and force Indians to assimilate into white society" (p. 5). The Dawes Act changed ownership from communal lands to individualized parcels with specific settler colonial management rules and regulations associated with each piece of land. If Indigenous peoples did not comply with regulations their parcels of land were subject to be taken and given or sold to European settlers, who understood how to fulfill regulations. Of the 140 million acres owned by Indigenous peoples in 1887, only 50 million acres remained under Indigenous ownership by the end of the Allotment Act period in 1934 (Pevar, 1992).

Shifting land management

With Indigenous removal by settler diseases and settler land policies came a shift in land management practices based upon settler colonialist philosophies of scientific management. Settler land philosophies were based upon the understanding of having dominion over the land and its resources. In addition, successful management was measured through one's ability to produce excess goods (Tuck and Yang, 2012). By design, settler colonizing science systems have been constructed to show that all other ways of knowing outside of the colonizers are inferior and therefore irrational (Rigney, 2001). As a means to colonize, Indigenous peoples were not allowed to maintain or practice their social systems, knowledge, traditional, and cultural sciences. Consequently, colonial settler-based science knowledge systems have designed an Indigenous reality through their own scientific discussions with no Indigenous input (Foley, 2003). When Indigenous Knowledge and peoples are left out of scientific discussions and education society as a whole loses out.

Exclusion of indigenous land management

Since the arrival of Euro-American colonizers to the United States, the Indigenous model of culture and science-based knowledge systems has been researched, described, and disseminated through a strictly Western-based lens (see Bohensky and Maru, 2011; Rigney, 2001). By design, the Western model forcefully removed perspectives that did not follow Western methodologies. This scientific model is described best through the title of Colin Scott's (1996) article "Science for the west, myth for the rest." In many cases exclusion of Indigenous knowledge systems has led to management and policy issues that affect not only Indigenous peoples but also the environments that sustain all peoples.

Protection versus sustainable resource utilization

Although many Indigenous peoples have maintained and expanded their knowledge systems and connections to the environment, Western laws and policies have often times overruled Indigenous knowledge and practices (Blackburn and Anderson, 1993). Western policies, in the form of national parks and federally protected environments, aim to protect environments, species, and lands. However, to ensure protection, these policies often prevent humans from interacting with those environments. Human removal goes directly against Indigenous cultural teachings, values, and responsibilities. As integrated members of the natural world Indigenous communities continue to shape environments through sustainable cultural management practices (Happybrook, 2004; Martinez et al., 2008). Therefore, policies and laws that exclude Indigenous cultural obligations and knowledge systems lead Native Americans to either be criminals in the Western sense or to assimilate and cease their cultural practices completely. This leads to devastating results for both the Indigenous communities as well as their environments. The concept of protection versus sustainable utilization remains an issue between Indigenous and Western land managers. One way to solve these conflicts is through educational discussions and on-the-ground participatory action research (PAR) projects that aim to manage and teach through the collective interested parties. In addition, there is a need for bi-cultural education programs in the natural resource and other sciences that can teach future resource managers about the foundational values found within Indigenous and Western worldviews (Lipe, 2013).

It is through the understanding of these two unique systems that a dialogue can start to take place on how and where IKS and Western science can be used to work together. Bi-cultural education programs can help future Indigenous and non-Indigenous scientists to learn how to communicate and work within Indigenous and Western science worlds. Bi-cultural education programs can also be used to help bridge Indigenous and western science worlds and create avenues for cross-cultural educational and on the ground land management partnerships. As discussed earlier collaborative cross-cultural education efforts create more comprehensive and encompassing natural resource management decisions by teaching individuals to see the world through diverse frames of reference. One example illustrating the need for diverse frames of reference is fire management within the United States.

Different perspectives for fire management

One Western policy example that has greatly impacted the Indigenous peoples and environments around the world, especially within the Pacific Northwest of the US, is the Weeks Act of 1911, also known as “fire suppression.” Fire, in its natural form, has transformed large portions of land throughout The United States and the world (Boyd, 1999). “Euro-American civilization, challenged by this uncontrollable force of nature, had a cultural aversion to fire. Since 1900 we (Euro-Americans) have tried to eliminate most forest fires with varying success. But, in just one century, settlement has changed the landscape, altering biodiversity, forest growth, wildlife habitat, and a host of other factors” (Barrett and Arno, 1999, p. 62). When the Indigenous peoples’ understanding, knowledge, and on-the-ground management were removed from their homelands of the Pacific Northwest and beyond, there were immediate repercussions to the environments they managed. For example, Ross (1999) states,

It has been documented that Native Americans were cognizant of benefits derived from selective burning. In contrast, most of the areas Euro-American settlers’ interpretation and policies suggest that they did not understand the ecological benefits of prescribed burning. By opposing Indian prescribed burning, Euro-Americans created an environmental situation that actually encouraged the heavy accumulation of dead vegetation, windfall, and deadfall; as well as change in certain species types, seed propagation, and alteration of fire resistance of some plants; increasing disease caused by bacteria, fungi, parasitic plants and viruses (p. 277).

Today society is dealing with the forest issues caused by the shift of management regimes from that of Indigenous peoples into the hands of Western policymakers.

Western fire suppression policies. The Weeks Act of 1911 was established as a land protection policy that included fire suppression of natural as well as human caused fires from the United States landscape (Otoole, 2007). Fire suppression policies and laws were made with no concern for Native American tribes and their knowledge systems. Fire suppression has created catastrophic changes in many forest ecosystems across the country (Stephens et al., 2007). The US government gave the Forest Service the right to suppress fires at any cost (Otoole, 2007). Since the establishment of the Weeks Act, fuel loads have increased in most US forests to a state of emergency. Helms (1998) identified fuel loads as the amount of combustible fuels that include ladder fuels identified as vegetation that provides vertical continuum between vegetation strata and can allow fires to reach crowns of trees. Increased fuel loads allow fires to grow to catastrophic crown levels. Crown level fires lead to the destruction of entire forest systems (Jain et al., 2012). With the removal of both natural and anthropogenic (man-made) fires by the Weeks Act, there has been a shift in fire regimes causing the very conditions that Indigenous peoples avoided through controlled burns.

Indigenous knowledge and burning practices. Indigenous peoples understood that fire, in its natural form, cannot be controlled or eliminated in ecosystems. Barrett and Arno (1999) state, “After a millennia of subsistence in the Northern Rockies, Indians were well accustomed to interacting with fire in the environment” (p. 62). Through careful observations and experimentation over multiple generations, Indigenous peoples within the Pacific Northwest gained the knowledge to respect and work with fire both in its natural and anthropogenic forms as a management tool. As noted by Turner (1999), “Controlled landscape burning required a detailed, in-depth knowledge of natural systems, especially ecological characteristics of vegetation and the successional nature of plant communities, as well as of various geographical and climate features of the landscape” (p. 201). For example, Native Americans demonstrated their in-depth knowledge by burning in a variety of areas during the late summer and early fall, coinciding with the natural fire regime (Blackburn and Anderson, 1993; Boyd, 1999). Furthermore, in observing how natural fires affected plant succession, fire was used as a tool to maintain desirable early successional stages with easily accessible nutrients (Boyd, 1999). Through time and use of fire the Indigenous peoples of the United States built a relationship with fire in order to maintain biodiversity and the health of their populations.

Indigenous fire and environmental management

In addition to establishing a relationship with fire for the benefit of humans, Indigenous peoples also used fire to enhance and manage ecosystems. Boyd (1999) describes Indigenous use of controlled burns in the form of spot fires. First, spot fires were used to reduce the fuel loads, decreasing chances of catastrophic crown fires. If a crown fire broke out, an entire ecosystem could be at high risk of being eliminated. Second, after a spot fire, new, highly-nutritious growth emerged that was utilized by animal species in the area. Third, spot fires were used to keep oak savannas open in order to make acorns more accessible to both animals and humans. Therefore, Indigenous peoples were able to establish a complex relationship between the environment, fire, and themselves. It is important to understand that while the Indigenous peoples of the United States utilized fire at different environmental

spatial and temporal scales. In this way, not all Indigenous peoples of the United States utilized fire to the same degree. The next several paragraphs provide specific ways Indigenous peoples of the Pacific Northwest utilized fire as land management tool in fishing, crop and range management, and for creating basketry materials.

Connecting fire to fisheries. With six different species of salmon returning to different river systems at different times throughout the Pacific Northwest, salmon is a main staple of food for many Indigenous peoples of the Pacific Northwest. Salmon returns for river systems like the Columbia River watershed were capable of having runs of 13–16 million fish annually (Taylor, 1999). Indigenous peoples are found from the mouth of the Columbia River along the Oregon/Washington Coast to the headwaters and tributaries that cover a span of over 260,000 square miles and seven states (“Columbia River,” 2012). A few reasons fire was used in fishing was to increase water levels, create salmon harvesting materials, baskets for salmon storage and trade materials, and for cooking salmon.

Fire was also used for creating materials utilized for making fishing nets and other fishing materials. Different techniques of fishing are utilized to capture salmon depending upon river conditions and where along the Columbia River continuum peoples live (Taylor, 1999). Different ways of storing food for wintertime and for trading was also very important. Most tribes along the Columbia River system dried salmon and pounded it into pemmican as a way to store the highly nutritious food in smaller packable sizes for winter storage and for trading (Taylor, 1999). Pemmican was stored in large baskets that were often traded for goods not found within the area. There was no place larger or better known to Native Americans of the Pacific Northwest for fishing and trading networks than Celilo Falls, found at the Dalles in Oregon (Taylor, 1999). Burning for materials varied depending upon the environment and the plant species found there. The fishing techniques most utilized by Indigenous peoples included but were not limited to spears, dipnets, poisons, seines, gillnets, traps, baskets, and weirs (Taylor, 1999). Each fishing technique and apparatus used required quality materials that came from maintained environments that produced the required resources. Most of the plant materials used in netting were from local areas and included iris, cedar bark, and silk grass. Through trading of harvested salmon, some tribes were able to obtain Indian hemp and bear grass from tribes found on the plateau (Taylor, 1999). Creating fishing and food storage materials required management of local environments. Management techniques and scale varied depending upon how much and what kinds of materials were needed.

Most tribes burned areas as a means to produce high quality materials for making cordage, netting, baskets, clothing, tools, and ceremonial items (Blackburn and Anderson, 1993). One such item was seine nets. Seine nets on the Columbia River have been reported as being up to 600 feet long and 16 feet tall (Taylor, 1999). According to Blackburn and Anderson (1993), “Craig Bates of the Yosemite Museum has estimated that approximately five stalks of Indian hemp or milkweed would have been required to manufacture one foot of cordage” (p. 23). Therefore, it would take a colossal number of high-quality materials to build and maintain a net the length of two football fields. Fire was the tool used to produce the majority of the quality materials used in making these goods. In addition, in many places in California, fire was used to burn large areas creating smoke filled environments at different times of the year. Heavy smoke areas from fire could have caused a change in lowering water temperature (Lake, forthcoming). Research in the Northern California area is being conducted to determine what effect dense smoke conditions may have had on rivers with salmon runs on them. Smoky covered environments in areas with heavy fire use such as northern California may have caused a decrease in water temperatures helping salmon species during late fall spawning times.

In addition, fire was utilized in fishing to increase water availability and levels. Burning areas with vegetation decreased water evapotranspiration rates, which increases underground water and replenishes aquifers (Lake, 2007). Replenished aquifers in turn increase the water flows and availability of water in salmon-producing rivers and tributaries. Depending upon species and area burned, fire could make significant increases in water availability, especially in later fall.

Enhancing crop production through burning. As a high context knowledge system incorporating multiple environmental variables and factors, Indigenous fire management practices were used at different temporal scales for multiple purposes. According to Boyd (1999), “Willamette Valley Indians must have had some understanding of the long-range affects that periodic firing had on brush and refuse to make acorn gathering easier, reduce competition from other plants, kill parasites, and encourage a higher level of acorn production” (p. 120). In some areas fire was used to facilitate growth of medicines (Lake, 2007). Keeping areas open was another primary reason used to increase crop yields while decreasing unwanted encroaching species (Lake, 2007). Burning also maintained and controlled pest species such as acorn weevils that could infest crops if not taken care of. Furthermore, Schenk and Gifford (1952) note, “Karuk women, for example, reported that the trees are better if they are scorched by fire each year. This kills disease and pests” (p. 382). Therefore, fire was used to manage crops by reducing weeds, killing parasites, and eliminating other encroaching species.

Fire and range management to increase ungulate populations. According to Boyd (1999), in certain areas fire was used “to maintain open, park-like conditions, which were conducive to ungulate species, fire removal of windfall and deadfall and understory fire management was necessary” (p. 279). Depending upon seasonality, many areas were burned as a means to both increase range and forage availability for deer and elk species in higher elevations, nearer to winter. Boyd continues, “Firing areas to shoot deer more easily resulted in increased browse and pasturage for larger numbers of deer and elk” (p. 120). Other areas were burned as a means to reduce palatable vegetation for ungulate species, forcing animals to move lower into valley hunting areas. Ross (1999) describes, “Burning destroyed tree moss which the deer feed on in winter season, and by burning the moss, the deer are obliged to descend into the valleys for food, making them more available for Indian predation” (p. 281). Other areas were burned as a means to drive ungulate species, primarily deer, into snares or hunting parties. Native American burning was not just for short term hunting goals, however. LaLande and Pullen (1999) point out, “Fire had longer-term wildlife goals as well; goals that might be years in the future ... Indians in South Western Oregon employed fire to improve and maintain wildlife habitat” (p. 261).

Fire environments and basket weavers' expertise. For most Indigenous groups, baskets and basket materials were among the most important and heavily utilized tools in everyday life. Controlled burning by Indigenous basket weavers has been utilized for centuries as a means to enhance the growth of certain basket materials (Ortiz, 1993). Basket weavers are plant biologists, botanists, and environmental managers all rolled into one. Weavers understand how plants work and how to create optimal conditions for growing materials best suited for making the best baskets. Basket weavers burn for a host of reasons. In describing the benefits of burning, "It [fire] returns important nutrients to the soil, eliminates unhealthy plants, reduces pests, and keeps understory vegetation sufficiently spaced, assuring that each plant will get the optimum light, water, and nutrients" (Ortiz, 1993, p. 204). Oftentimes basket material tracts of land are managed in much the same way as a garden. According to Ortiz (1993), "Basket weavers utilize the same horticultural methods used by gardeners in their yards: judicious harvesting, pruning, cultivation, and debris clearing. They also utilize controlled burning" (p. 199). Due to the changes in land management policies and practices many of these garden-like environments have been or are in jeopardy of being lost.

The example of fire illustrates how different cultures can bring different experiences and values to environmental land management discussions. In this case, one group views fire as damaging to the health of the ecosystem, therefore wanting to control and eliminate it, while another group views fire as a tool to maintain the health of ecosystems. However, because the US government and Western scientists have control over fire policies, the results have allowed for increased fuel loads and damages caused by increased crown fires. The wildfire crisis in the United States is a direct result of a longstanding land management policy failure (Busenberg, 2004). This crisis may have been avoided had the Indigenous peoples of the United States and their knowledge base been part of the educational and land management decision making processes since the arrival of colonizing settlers. Therefore, the scientific community needs to continue to expand their understandings of science in order to allow diverse perspectives to be heard and utilized.

Native American land management today

Through treaties that created Government-to-Government relationships between the US and Native American tribes, the US Government has a legal obligation to work with Indigenous tribes on land management issues (Pevar, 1992, 2012). Indigenous peoples of the United States own and manage millions of acres of both privately owned and US governmental ceded lands "that are entrusted to the US government. Trust lands are lands in collective tribal holding for which the secretary of the Interior has a continuing trust responsibility to manage in a manner to benefit the respective tribe" ("Appendix B," n.d., p. B-2). Tribal organizations maintain over 15,990,000 acres of forestlands, including more than 5,700,000 acres of commercial forestland and 8,700,000 acres of woodlands ("National Indian Forest," 2012). Found alongside these nearly six million acres of forestlands belonging to Indigenous peoples are over 2500 shared miles of common border and long-term stewardship responsibilities for the forests entrusted to their care ("Proposal for the Tribal Forest", 2011).

In many cases tribes continue to manage lands through both traditional and modern technology. While tribes do understand the economic value of timber, they also see the forest as much more than just trees and economic wealth. As mentioned by Wells (as cited in Simmons, 2012), in an interview with Chief Tanner of the Coquille tribe,

We get many things from the forest—canoes, baskets, clothes, shelter, fir, cedar, spruce, bear grass, camas, and we use all of these things, but they also have a spiritual value which we honor as our ancestors. What we take, we try to give back; what we don't need, we try not to take (p. 2).

In this way, tribes offer valuable insights and ethical standards that are useful to current land management practices. According to Prober et al. (2011),

There has been increasing global interest in accounting for different knowledge frameworks in the sustainable management of natural resources ... it is assumed that this will result in more resilient social-ecological systems through the incorporation of a broader information base, consideration of a wider diversity of culturally derived ideals, and elucidation of conflicts and synergies (p. 1).

In order to work together more effort needs to be made to create avenues for cross-cultural education between Indigenous and Western science.

Educating through diversity for the Future? Increasing the ability and number of individuals who can transit between Indigenous and Western ways of knowing and interacting with natural resources can prove beneficial for all. Collaborative involvement of tribes and tribal partners working with the Forest Service is essential to what the Department of Agriculture calls the "All Lands, All Hands" approach to managing ecosystems (National Association of Foresters, 2010). Creating a new land management regime through natural resource education curricula based upon sustainability in which all stakeholders work together can improve forest health, promote collaborative relationships, and increase landscape-scale management objectives and health of environments for all people.

While much of this paper focuses upon the United States, it is essential that we begin to identify and work with all Indigenous peoples around the world. It was recently reported that while Indigenous peoples around the world only control 20% of the lands surface they hold 80% of its diversity (<https://doi.org/10.4060/cb5131en>). The knowledge, understandings and knowhow can be found in and with Indigenous knowledge and the people places and relationships they hold to their local places. It is also through these same relationships that the earth herself can be healed. However, we must incorporate Indigenous peoples and their knowledge systems to help us understand how to create these relationships and help provide foundational understandings for how to move and be in place. It is through these understandings and connections that we be shown how to care for our earth. In a recent report by the Indigenous peoples currently only hold 20% of the earth's land but maintain over 80% of the world's diversity (<https://doi.org/10.4060/cb5131en>). This is by no means a coincident.

It is essential that we find a way to bring diverse perspectives of knowing and understanding together to be taught side by side. Natural resource education and curricula that includes cultural diverse perspectives is one way to provide future land managers and policy makers with the tools required for making better and more encompassing land management decisions. Bi-cultural education programs that teach the foundational values of both Western and Indigenous sciences can be helpful for teaching students from different cultural backgrounds about the similarities and differences between Western and Indigenous scientific perspectives (Lipe, 2013). In turn, these understandings can be beneficial for establishing a baseline point for dialogue used in bridging the gap between these two unique scientific perspectives. Whether policymaking or on-the-ground management, we as a larger society can no longer rely on Western science to come up with all the answers. Indigenous peoples and knowledge systems must be included into environmental education and management strategies if we are going to have any chance of a sustainable future.

Shifting paradigms of science education through Indigenous Knowledge courses and programmatic changes. The Sustainable Community Food Systems (SCFS) Bachelors of Applied Science Degree program at the University of Hawaii West Oahu, is a multidisciplinary experiential-based program that teaches the next generation of foods systems professionals the skills required in solving real world problems. In 2017, the SCFS program began developing a line of Indigenous Knowledge courses as a means to diversify its course offerings. These courses utilized both Indigenous and western science-based knowledge in order to expand students' knowledge bases and provide students with exposure into how western and Indigenous scientific based knowledge system can work together to provide answers to some of issues facing the state of Hawaii and beyond. The following is a list of courses that were developed along with a short description of what each course covered.

- **SCFS 315-Wild Game Management-** This 3-unit field course that provides an analysis of sustainable wild game management and processing techniques from both western science and Indigenous Knowledge perspectives. Through assigned readings, guest presentations from practitioners, and real-life first-hand experiences, this course seeks to teach the importance of wild game and food management as it relates to ecological sustainability and food production. This course describes Indigenous relationships and responsibilities involved in sustainable bio-cultural utilization of wild game management. It will also demonstrate how managing wild game can be another relevant component to helping fix today's global issues of natural resource management and food security.
- **SCFS 370-Ahupua'a Perspectives-** This course focuses on Hawaiian Natural Resource Management. It will provide an in-depth understanding of Hawaiian natural resource management through the study of connections found from mountain top to sea. Exploring historic as well as current documentation and case studies students will learn natural resource perspectives of Hawaiians pre-European contact. In addition, students will learn how Hawaiian based environmental management practices are still alive today and can provide valuable insights as we as a greater society look to living more sustainably.
- **SCFS 385-Fishponds and Near Shore Fisheries-** This course focuses on Hawaiian fisheries management past present and future. Centered around readings, case studies, guest presenters and field trips students will learn about Hawaiian based understandings of historical fisheries management. Included in this course will be in depth discussions about how and why Hawaiians designed, built and managed fishponds before Euro-American contact. In addition, the course will focus on teaching how many of these cultural practices are still relevant today. This course provides experiential based site visits to current fishponds that are being restored as a means to be a part of and learn current strategies for rebuilding and restoring Hawaiian fishpond for a more sustainable future.
- **SCFS 415-Mahi ai kalo-** This course provides and in depth understanding of Hawaiian lo'i agricultural practices both before and after Euro American contact. Through readings, guest presenters, and field trips students will learn and experience first-hand the intricacies within Hawaiian cultural and scientific understandings within lo'i kalo. In addition, this course will examine how these understandings are as relevant today and, in the future, as they were in the past.
- **SCFS 455-Ethics in Worldviews-** This course provides an in-depth understanding of how worldview shapes education, on the ground management, and policies of natural resource management. Students will learn about Native American, Hawaiian, and Western European based environmental world and how these beliefs that are shaped by the foundational values and ethics within them affect natural resource management. In addition, students will learn about how understanding diverse perspectives can help provide more encompassing answers to the complex multilevel environmental issues we are facing today, leading to a more sustainable future.
- **SCFS 485-Traditional Ecological Knowledge (TEK)-** provides an introduction to understanding of how Traditional Ecological Knowledge (TEK) can shape education, on the ground management, and policies of natural resource management. Students will learn about Native American, Hawaiian, and Western European based environmental knowledge and how these beliefs that are shaped by the foundational values affect natural resource management. In addition, students will learn about how

understanding diverse perspectives can help provide more encompassing answers to the complex multilevel environmental issues we are facing today, leading to a more sustainable future.

Including Indigenous courses as degree requirements—While the SCFS staff believed these courses were a valuable addition to the program they quickly realized that as electives it was not enough. So, in 2021, the program staff wrote a United States Department of Agriculture grant called, The Undergraduate Pathways Bridging Excellence in Agriculture and Tradition (UPBEAT), its main focus is to help Indigenize the Sustainable Community Food Systems Bachelors of Applied Science (SCFS-BAS) undergraduate degree program at the University of Hawaii West Oahu. The UP-BEAT project seeks to expand the systematic integration of Indigenous Knowledge and community-based experiential learning. One aspect of this grant was to make programmatic changes to include the courses listed above as graduating requirements instead of elective courses. By making these courses requirements both current and future SCFS students will take both Indigenous and western science-based courses in order to graduate. These courses will provide a detailed perspective along with hands on practical skills and tools required to utilize both Indigenous and Western science knowledge systems to solve food systems challenges of Hawaii and beyond. The SCFS program is using these experiential based courses to reconnect students to their local environments. In the end the program believes that by connecting Indigenous and western science knowledge it will provide the essential skills required to transition current food systems into a more equitable and sustainable system for all.

References

- Adler, N.J., 2002. International Dimensions of Organizational Behavior. South-Western, Cincinnati, Ohio.
- Barnhardt, R., Kawagley, A., 2005. Indigenous knowledge systems and Alaska native ways of knowing. *Anthropol. Educ. Q.* 36 (1), 8–23.
- Barrett, S., Arno, S., 1999. Indians Fire and the Northern Rockies: Ethnohistory and Ecology. *Indians Fire and the Land in the Pacific Northwest*. Oregon State University Press, pp. 50–64.
- Berkes, F., 1999. *Sacred Ecology: Traditional Ecological Knowledge and Resource Management*. Taylor & Francis, Philadelphia, PA.
- Blackburn, T., Anderson, K., 1993. *Before the Wilderness: Environmental Management by Native Californians*. Ballena Press, Menlo Park CA.
- Bohensky, E.L., Maru, Y., 2011. Indigenous knowledge, science, and resilience: what have we learned from a decade of international literature on “integration”? *Ecol. Soc.* 16 (4). <https://doi.org/10.5751/ES-04342-160406>.
- Boyd, R. (Ed.), 1999. *Indians, Fire, and the Land in the Pacific Northwest*, first ed. Oregon State University Press, Corvallis Or.
- Bradey, K.M., 2011. Indian removal act. In: Manning, M.J., Wyatt, C.R. (Eds.), *Encyclopedia of Media and Propaganda in Wartime America*. ABC-CLIO, Santa Barbara, Calif., pp. 203–204.
- Busenberg, G., 2004. Wildfire management in the United States: the evolution of a policy failure. *J. Rev. Policy Res.* 21 (2), 145–156.
- Cajete, G., 2000. *Native Science: Natural Laws of Interdependence*. Clear Light Publishers, Santa Fe, N.M.
- Carlos, A.M., Lewis, F.D., 2012. Small pox and Native American mortality: the 1780s epidemic in the Hudson Bay region. *J. Explor. Econ. Hist.* 49, 277–290.
- Columbia River, 2012. United States Environmental Protection Agency. July 13, 2012. Available at: <http://www.epa.gov/columbiariver/>.
- Foley, D., 2003. Indigenous epistemology and Indigenous standpoint theory. *Social Alternatives* 22 (1), 44–52.
- Happynook, T., 2004, May. *Securing Food, Health and Traditional Values through the Sustainable Use of Marine Resources*. Unpublished document.
- Helms, J.A. (Ed.), 1998. *The Dictionary of Forestry*. Society of American Foresters, Bethesda, MD.
- Jain, T.B., Battaglia, M.A., Han, H.-S., Graham, R.T., Keyes, C.R., Fried, J.S., Sandquist, J.E., 2012. *A Comprehensive Guide to Fuel Management Practices for Dry Mixed Conifer Forests in the Northwestern United States* (General Technical Report No. RMRS-GTR-292). United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, p. 331.
- Kincheloe, J.L., Steinberg, S.R., 2008. Indigenous Knowledge in Education. In: *Handbook of North American Indians; History of Indian-white Relations*. Sage, pp. 135–156.
- Lake, F., 2007. *Traditional Ecological Knowledge to Develop and Maintain Fire Regimes in Northwestern California, Klamath-Siskiyou Bioregion: Management and Restoration of Culturally Significant Habitats*. Doctoral dissertation. Oregon State University, Corvallis OR.
- Lalande, J., Pullen, R., 1999. Burning for a fine and beautiful open country. In: Boyd, R. (Ed.), *Indians Fire and the Land in the Pacific Northwest*. Oregon State University Press, pp. 255–276.
- Lipe, D., 2013. *Science as a Worldview: A Discussion on Indigenous Ways of Knowing and Doing*. Doctoral dissertation. University of Hawai'i at Manoa, Honolulu HI.
- Martinez, D., Salmon, E., Nelson, M.K., 2008. Restoring Indigenous History and Culture to Nature. In: *Original Instructions: Indigenous Teachings for a Sustainable Future*. Bear and company, pp. 88–115.
- Mazur, B., Biatostocka, P., 2010. Cultural diversity in organizational theory and practice. *J. Intercult. Manag.* 2 (2), 5–15.
- Mooney, J., 1992. *James Mooney's History, Myths, and Sacred Formulas of the Cherokees : Containing the Full Texts of Myths of the Cherokee (1900) and the Sacred Formulas of the Cherokees (1891) as Published by the Bureau of American Ethnology : With a New Biographical Introduction, James Mooney and the Eastern Cherokees*. Historical Images, Asheville, N.C.
- National Association of Foresters and the USDA Forest Service, 2010. *All Lands, All Hands: Sustaining America's State and Private Forests*. State and Private Forestry Annual Report, pp. 1–31. Available at: <http://www.stateforesters.org/sites/default/files/publicationdocuments/FY09-SPF-Annual-Report.pdf>.
- National Indian Forest Resources Management Act, 2012. US Forest Service. July 26, 2012. Available at: <http://www.fs.fed.us/spf/tribalrelations/documents/policy/statutes/PL101-630NIFRMA.pdf>.
- O'toole, R., 2007. *The Perfect Firestorm: Bringing Forest Service Wildfire Costs under Control*. Policy Analysis, executive summary No. 591, p. 16.
- Ortiz, B., 1993. Contemporary California Indian basketweavers and the environment. In: Blackburn, T., Anderson, K. (Eds.), *Before the Wilderness: Environmental Management by Native Californians*. A Ballena Press, pp. 195–212.
- Page, S.E., 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press, Princeton.
- Pevar, S.L., 1992. *The Rights of Indians and Tribes: The Basic ACLU Guide to Indian and Tribal Rights*. Southern Illinois University Press, Carbondale, IL.
- Pevar, S.L., 2012. *The Rights of Indians and Tribes*. Oxford University Press, Oxford; New York.
- Prober, S.M., O'Connor, M.H., Walsh, F.J., 2011. Australian Aboriginal peoples' seasonal knowledge: a potential basis for shared understanding in environmental management. *Journal of Ecology and Society* 16 (2), 12.
- Rigney, L., 2001. A first perspective of indigenous Australian participation in science: framing indigenous research towards indigenous Australian intellectual sovereignty. *Kauma High. Educ. J.* 7.
- Ross, J., 1999. Proto-historical and historical Spokane prescribed burning and stewardship of resource areas. In: *Indians Fire and the Land in the Pacific Northwest*. Oregon State University Press, pp. 277–291.

- Schenk, S.M., Gifford, E.W., 1952. Karuk ethnobotany. *Anthropol. Rec.* 13 (6), 377–392.
- Scott, C., 1996. Science for the west myth for the rest. In: Nader, L. (Ed.), *Naked Science: Anthropological Inquiry into Boundaries, Power, and Knowledge*. Rutledge, New York.
- Simmons, E., 2012. Non-timber forest resources-what is in this term? *BC j. ecosyst. manag.* 13 (1).
- Stephens, S.L., Martin, R.E., Clinton, N.E., 2007. Prehistoric fire area and emissions from California's forests, woodlands, shrublands, and grasslands. *For. Ecol. Manag.* 251 (3), 205–216. <https://doi.org/10.1016/j.foreco.2007.06.005>.
- Shuttleworth, M., 2009. The History of the Scientific Method Is a Fascinating and Long One, Covering Thousands of Years of History. Explorable. Available at: <http://explorable.com/history-of-the-scientific-method.html>.
- Taylor III, J., 1999. *Making Salmon: An Environmental History of the Northwest Fisheries Crisis*. University of Washington Press.
- Tuck, E., Yang, K.W., 2012. Decolonization Is Not a Metaphor. Unpublished manuscript.
- Turner, N.J., 1999. Time to burn: traditional use of fire to enhance resource productions by aboriginal peoples in British Columbia. In: *Indians Fire and the Land in the Pacific Northwest*. Oregon State University Press, pp. 185–218.
- Ubelaker, D.H., 1988. North American Indian population size, A.D. 1500 to 1985. *Am. J. Phys. Anthropol.* 77, 289–294.
- Valencia-Weber, G., 2003. Symposium: Native American and the Constitution: The Supreme Courts Indian Law Decisions: Deviations from Constitutional Principles and the Crafting of Judicial Smallpox Blankets. *University of Pennsylvania Journal of Constitutional Law*.
- White, R., Cronon, W., 1988. Ecological change and Indian-white relations. In: *Handbook of North American Indians; History of Indian-white Relations*, vol. 4. Smithsonian Institution, pp. 417–429.

Kayani:yo (a good path)—Warrior Science Ohneganos—Indigenous ecological knowledge

Dawn Martin-Hill^{a,e}, Colin M. Gibson^{b,e}, Makasa Looking Horse^{a,e}, Danielle Gendron^{c,e}, Emily Anson^{a,e}, Kahontiyoha Cynthia Denise McQueen^{a,e}, and Tehahenteh^{d,e}, ^aMcMaster University, Hamilton, ON, Canada; ^bUniversity of Guelph, Guelph, ON, Canada; ^cUniversity of British Columbia, Vancouver, BC, Canada; ^dIndigenous Knowledge Exchange; and ^eOhneganos Ohnegahde:gyo, Hamilton, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Ohneganos Ohnegahde:gyo	207
Why water?	208
Who we are	209
Kayani:yo (A good path)	212
Equity in scientific inquiry: Indigenous Knowledge dismantling structural barriers	213
Reconnecting youth to waterways and tradition	214
Knowledge mobilization and community engagement	214
Public dissemination matters	215
References	215

Ohneganos Ohnegahde:gyo

From the perspective of the traditional Haudenosaunee, we speak in terms of responsibilities with respect to water, not in terms of water rights. From time immemorial, we have held the view that the "law of the land" is not man-made law, but a greater natural law, the Great Law of Peacethe root words for "rain" in Mohawk means expensive, or precious or holy. Culturally, we would not abuse this resource.

King (2007, p. 452).

Indigenous communities in Canada are experiencing severe water and health crises caused by climate change and colonialism, but are also uniquely positioned to contribute to the solutions to these problems by accessing Indigenous Knowledge (IK). IK pedagogy acknowledges diverse ways of knowing and respects the pluralism of knowledge. Described by Daes (1994), IK is "... a complete knowledge system with its own concepts of epistemology, philosophy, and scientific and logical validity ... [which] can only be fully learned or understood by means of the pedagogy traditionally employed by [the] peoples themselves ..." (pp. 2–3). Further, Marlene Brant Castellano (2000) adds, "... knowledge valued in aboriginal societies derives from multiple sources, including traditional teachings, empirical observation, and revelation (p. 23) ... Aboriginal knowledge is said to be personal, oral, experiential, holistic, and conveyed in narrative or metaphorical language (p. 25)".

The philosophical foundation of IK prioritizes local geographies and relationships with the ecosystem over thousands of years. This is, however, not to say that the breadth of IK is made up of pockets of disconnected local or niche epistemologies, nor that Indigenous Nations stayed in one region and were static. Indeed, there is strong evidence that Māori wayfarers arrived on Salish territory (what is now the west coast of North America) long before Europeans. Recently, in a large genomic study conducted through an international collaboration (see Ioannidis et al., 2020), researchers found that "... Polynesians and Native Americans met at one point in history, and during that time people from the two cultures produced children with both Native American and Polynesian DNA" (Armitage, 2020). In a science communications article about the study, a co-lead researcher from Stanford University explained that the data shows, "... identical-by-descent segments of Native American ancestry across several Polynesian islands ... [serving as] conclusive evidence that there was a single shared contact event" (Armitage, 2020). The mounting DNA linkages of Salish (Native American) to Māori (Polynesian) ancestry is surprising to western scientists, but not to the Salish or Māori who have retained and preserved their oral histories; stories which now seemingly function to corroborate more formally accepted forms of "scientific" evidence.

The study referenced above is one example of how western academia is beginning to shatter notions that Indigenous people weren't global travelers pre-European contact; from the Bering Strait theory to broader human history, western science (WS) is learning that Indigenous oral histories are legitimate records. Indigenous histories and knowledge are encoded in songs, ceremonies and oral stories passed down from one generation to the next. The Haudenosaunee "social dances" are borrowed from other Nations; they are records of travel, allies, and observations kept alive and transmitted through song and dance. For example, the alligator dance carries accompanying oral histories of Haudenosaunee friendships with other tribes and ecosystems.

Indigenous Knowledge systems draw from within and from outside of neighboring tribes. This type of meeting and exchange between Indigenous communities across the globe has informed IK since time immemorial. Today, the resurgence of IK systems – including philosophies, epistemologies, identities, political governance, land-based practices, cultures, languages, and so on – weaves together multiple aspects of the fibers of Indigenous Knowledge methods and pedagogic practices of research. Whether

from health, environmental, cultural, or technological perspectives, IK research is intrinsically interdisciplinary, intersectional, and often inter-tribal. All of the natural world is interconnected, so too is IK. Indigenous Knowledge is "... a body of information about the interconnected elements of the natural environment which traditional Indigenous people have been taught, from generation to generation, to respect and give thanks for" (citing Jameson Brant in [National Aboriginal Forestry Association, 1996](#), p. 6). Ecocentric principles often bind Indigenous worldviews together regardless of region or Nation, whether from the Americas, Australia, Asia, or Africa. The real scientific question to pose is, how do Indigenous people from disparate parts of the globe hold common ecocentric worldviews, and why have "others" not held similar ontological truths throughout the ages?

Globally, IK has been shown to strengthen socio-ecological community resilience to the multiple stressors affiliated with large-scale environmental change. Following extensive engagement and discussions with community, researchers working with Six Nations of the Grand River (Six Nations) identified three primary areas of interest:

1. Bridging IK and WS [training];
2. Building youth resilience related to water security [wellness]; and
3. Exploring Indigenous laws through culture, language, and ways of knowing as essential to ecosystem management [governance].

Our project, termed *Ohneganos Ohnegahde:gyo* (Ohneganos) by the advising community knowledge holders, consists of leading experts in IK methodology who are working alongside western-trained scientists to co-create a holistic model of water research. The three Indigenous-led teams involved in the project (training, wellness, and governance) focus on the issues of addressing stewardship over time and crafting bilingual, relevant educational resources to foster resilience. Key deliverables of the project include developing mixed-method materials that are relevant for both a local and global audience (for instance, digital stories uploaded to YouTube¹) as well as establishing pathways for Indigenous secondary & post-secondary students to be trained in the braiding of Indigenous and western sciences. The overarching aim of our project is to holistically address cross-cutting and interconnected issues to create a "double dividend" that builds capacity and improves the resilience of the Six Nations community – acknowledging that Indigenous communities across Canada and beyond will stand to benefit from the knowledge, methodologies, technologies, and tools co-developed from the work completed.

Why water?

Indigenous communities in Canada are experiencing a water crisis that is exacerbated by climate change and has resulted in environmental, social, and health concerns. Water supplies are at high risk over a lack of access to water quality and quantity, water monitoring technologies (i.e., real time data and clear legal standards), and skilled management system operators. Inadequate infrastructure further increases the health burden in Indigenous communities.

Water crisis is widely experienced in Indigenous communities due to the "... ongoing struggle to have Indigenous voices heard in the decision-making processes that affect their lives, lands, and waters" ([McGregor, 2012](#), p. *abstract*). Research, capacity-building, and other supports are needed to advance a range of water-related topics such as governance, health, and expanding capabilities – including the development of Indigenous sustainability. The lattermost relies on IK, which is required to understand the water-health-legal nexus associated with water governance and is foundational to the mitigation of issues arising from inadequate water management. Such efforts and advancements necessitate thoughtful inclusion of diverse interests reflecting commonly held values in order to co-create and enhance innovative, sustainable, smart environmental management and training within communities that draws upon both IK and WS.

Indigenous Knowledge is increasingly recognized as a valuable source for adaptation and resilience-building in the face of impending environmental change and water insecurity – locally, regionally in North America (see [Finn et al., 2017](#)), as well as globally (see [UNFCCC, 2016](#); [Nakashima et al., 2018](#); [IPBES, 2019](#)). The Government of Canada's Federal Budget for 2017 underscored the priority of integrating IK in the country's pathway for sustainable growth to "... build a better understanding of climate change and to guide adaptation measures; enhance Indigenous community resilience through infrastructure planning ..." ([Government of Canada, 2017](#), p. 129). Similarly, the United Nations (UN) Sustainable Development Goals (SDGs) call on the governments of nation-states to "... [Goal 6.b] support and strengthen the participation of local communities in improving water and sanitation management" ([United Nations, 2015](#), p. 21).

With these goals in mind, Six Nations asked Ohneganos to develop specialized and accredited water management training, produce a bilingual virtual reality experience to disseminate co-created results from WS and IK, and create bilingual texts and educational resources for community knowledge mobilization. Knowledge mobilization on these issues involves listening, learning, educating, and training youth in the community in water management and governance to achieve self-determination and autonomy. In its entirety, the project consists of team leaders supported by dozens of community and university partners working in tandem to create strategic outcomes in Indigenous-led sustainable water monitoring and management.

Several challenges are inherent in these efforts. Establishing new, Indigenous-led pathways to post-secondary science and environmental streams is needed to build capacity in Indigenous communities, but at the moment this integration into postsecondary Indigenous Studies is struggling. This highlights the importance of capacity development, which includes the delivery of a cohesive

¹See [here](#) to view a YouTube playlist of digital stories published by Ohneganos.

and strategic stream of secondary school-level training that provides youth and community leaders with skills to respond to regional, national, and international developments impacting Indigenous water management. IK indicators of climate change, traditional governance principles, and determinants of wellness challenge the siloed approach of both community services and hyper-specialized university disciplines. Capacity development within university and community settings has – and will continue to – ensure long-term sustainable water stewardship, monitoring, and management.

Who we are

As Haudenosaunee peoples, Six Nations has an inherent responsibility to ensure that future generations can sustain their needs, and water is central to this ethos.

... For First Nations, water is a sacred gift, the life blood of Mother Earth, and all water, not just water for human use, needs protection. Through Indigenous ceremonies, laws, and protocols, First Nations have exercised inherent responsibilities to fulfill obligations to the Creator to ensure clean water for all living things since time immemorial ...

Longboat (2013, pp. 6–7).

Developing a Haudenosaunee-informed planning process to identify and manage surface waters and aquifers in Six Nations territories concerns land and waters outside of the government-imposed reserve boundaries of Six Nations. The area of concern extends to the Haldimand Tract (see Fig. 1) – which straddles the Grand River “... six miles deep from each side ... beginning at Lake Erie ... to the head of said river” (Johnston, 1964, pp. 50–51) – as well as beyond the tract, spanning numerous other disputed areas under “claim” by Six Nations, across the broader traditional territory of the Haudenosaunee.

Consumptive water takings, such as proposed commercial permits to remove groundwater² from the source with no return flow, are of significant concern to the Haudenosaunee leadership. Ohneganos consulted with the Haudenosaunee Confederacy Chiefs

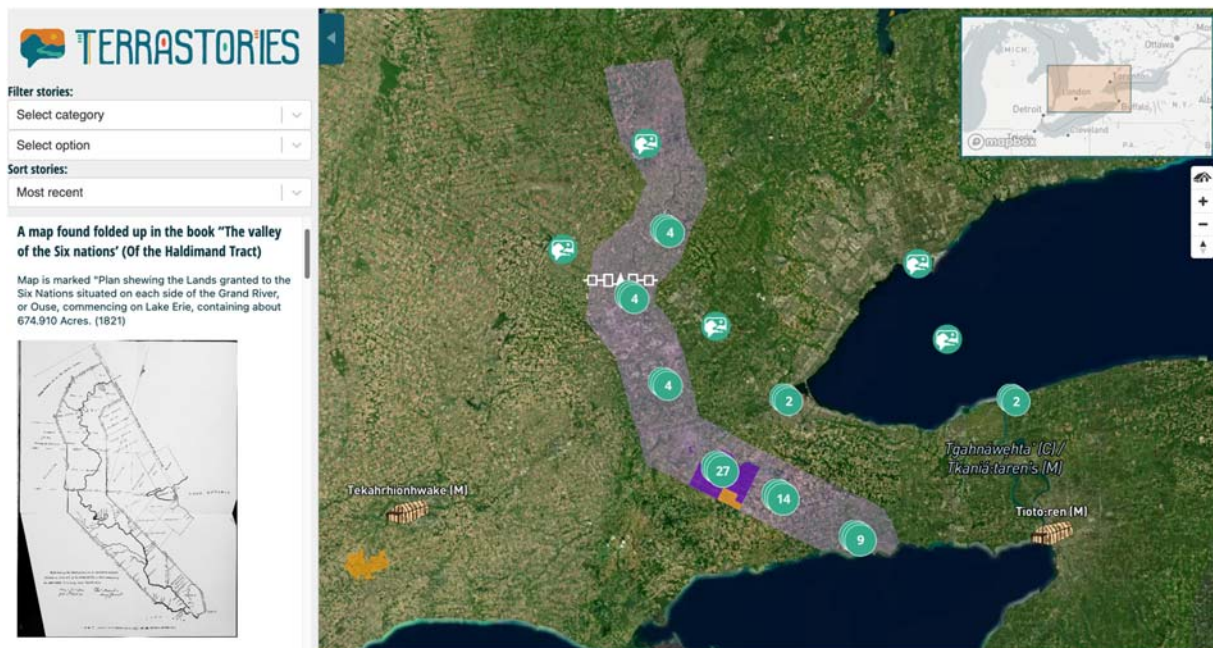


Fig. 1 Six Nations Indian Reserve No. 40 (purple block) and the Haldimand Tract (shaded strip). Image created using Terrastories by Dawn Martin-Hill and Rudo Kemper.

Council (the traditional governing body of the Haudenosaunee) as well as Six Nations Elected Council and community members. Through meetings, roundtables, workshops, information sessions, climate change actions, and youth “No Nestlé” protest marches,

²See <https://www.wellingtonwaterwatchers.ca/nestle-faq>.

Ohneganos has shared research findings with stakeholders and established a broader online and social media presence.³ This has helped to create formal networks with multiple stakeholders from community, surrounding organizations, governments, and grassroots groups all concerned with local water management and regional management of the Grand River watershed. Six Nations departments (e.g., Health Services, Social Services Public Works, Wildlife Management Office, Environment, and Birthing Center), Haudenosaunee organizations (e.g., the Haudenosaunee “Protect the Tract”⁴ and Haudenosaunee Environmental Task Force), along with local primary and secondary schools all contributed in direct and indirect ways to Ohneganos research and knowledge mobilization. The Ohneganos Water Governance subteam built off feedback from community consultation and intensive engagement with all partners, identifying the emphasis on knowledge mobilization through focusing on the youth. When combined, such networks in service, education, and governance institutions act to secure the holistic maintenance and ecological integrity of water. This in-depth, interdisciplinary, and multilateral response to a community call to action has led to greater awareness of the issues, leading to pragmatic and impactful change within the Six Nations. For example, during the COVID-19 lockdown in Summer 2022, Six Nations Elected Council and Public Works department offered septic removal free of charge as part of relief to water access. Generally, the fee for removal is \$90.00 and funds must be paid in advance for removal of waste from septic and holding tanks. Another example is the removal of all Nestlé water bottles from Six Nations stores and services, given the company’s illegal removal of millions of liters per day from the local aquifer.

A key component to fulfilling Indigenous ethical research requires answering the question: “who will benefit?”. This is why developing capacity for youth to monitor and govern their *own* waters is built into all co-creation research proposals. Thus far, Ohneganos has secured micro-credentials for secondary students working on the research program in science, language, and technology with the goal of ultimately combining them to form a university credit through McMaster University (a nearby postsecondary research institute). Working toward the goals of accrediting and training Indigenous youth required collaboration with science and language teachers from local secondary schools. Collaborative efforts have involved embedding GIS mapping techniques, water testing, and ecological assessments into digital Indigenous maps as well as soliciting students to help create digital stories through on-the-water technology training, contributing to valuable data collection for Ohneganos. Further, youth from Six Nations in post-secondary were particularly keen to develop governance and leadership skills, so international-level opportunities were created that resulted in several youth attending the United Nations Permanent Forum on Indigenous Issues (UNPFII) and the Expert Mechanism on the Rights of Indigenous Peoples (EMRIP). At the former event, youth presented at a side-table workshop which has led to numerous invitations throughout the past several years. The opportunity to build youth networking and sustaining those relationships has elevated Six Nations voices in water related UN-affiliated engagements (Fig. 2).



Fig. 2 Ohneganos youth training in a holistic Indigenous-led approach to water management. *Left to right:* student mapping workshop (A), protest against Nestlé (B), and youth lead opening at UN (C).

³Website (<https://www.ohneganos.com/>), Facebook (<https://www.facebook.com/ohneganos>), and YouTube (<https://www.youtube.com/c/OhneganosOhnegahd%C4%99gyo>).

⁴See <https://www.protectthetract.com/>.

A primary area of concern identified by Six Nations' youth were anxieties and grief over climate change and water insecurity. After retaining additional funding to expand Ohneganos in 2018, a community-led mental health research team was developed. A critically important determinant and indicator of health is access to clean water. Water insecurity negatively impacts quality of life and wellness, particularly mental health. Through surveys and interviews with Six Nations Health Services, Birthing Center, and youth focus groups, lack of access to clean water was found to be a significant stressor in day-to-day living. Further, youth experienced multiple barriers in harvesting the food and water they require for cultural continuity, thus impacting well-being (Chandler and Lalonde, 1998). If cultural ties to hunting, fishing, and ritual are impeded through environmental degradation, it further adds to stressors that affect overall health and wellness (Kirmayer and Valaskakis, 2009; Majeed and Lee, 2017).

Recent literature on environmental grief proposes that climate change creates environmental uncertainty, which can make youth more susceptible to depression and anxiety in general, while also specifically impacting other stressor sources, such as poor water quality and food insecurity. A recent study using data spanning several decades found that suicide rates are rising along with increasing temperatures because of global warming in both the United States and Mexico (Burke et al., 2018). Indigenous youth mental wellness must be centered in land, language, community, cultural identity, and empowerment (Martin-Hill, 2009; Mussell et al., 2004).

With this knowledge mind, Ohneganos youth lead, Makasa Looking Horse, spearheaded several youth-led initiatives for the project, including: organized climate change community actions, water protector teach-ins, No Nestlé marches, Youth Rise community leadership events, school art contests (see Fig. 3), and a tree giveaway for Earth Day; facilitated youth participation in the Two Row Paddle on the Grand River⁵; and created a renowned video podcast series called *Ohneganos: Let's Talk Water*.

Ohneganos supported outreach activities with youth designed to empower and elevate their voices. Through a partnership with Six Nations Social Services, Ohneganos provided schoolchildren with art supplies and asked them to express their views about water. This activity was highly successful and insightful, demonstrated by the submissions illustrated in Fig. 3. All of the activities designed by Ohneganos prioritize the empowerment of youth voices. By showcasing their work and written ideas and sharing their experiences in water protection, knowledge mobilization was advanced in a meaningful way, positively impacting youth-led water actions. This activity, which engaged youth as well as frontline health workers and Elders, served to foster community resilience in youth mental health.



Fig. 3 A collage of submissions from the “Water is Life” Ohneganos art contest in six nations.

⁵See <http://tworowonthegrandriver.com/>.

Bunch et al. (2011) suggest “... actions that address both biophysical and social environments have the potential to create a ‘double dividend’ that improves human health ... while also promoting sustainable development” (p. 2). From youth and community consultation to date, specific research tools have been identified that aim to create such a double dividend, including:

1. The need to develop a coordinated survey tool that includes information on water anxiety, perceptions of water quality, access to perceived healthy water and understanding of burden of cost, positive water actions, positive experiences with social justice, and resilience in youth.
2. The need for the documentation of water stories, emphasizing youth-led positive action as an open-access knowledge mobilization tool.
3. The need for a youth-informed resilience communication tool, such as a mobile application, that incorporates access to archives of youth stories, and serves a role for further water advocacy, youth agency, and a possible intervention in coping with water-related stressors.

In addition to developing a mental health mobile application, the ultimate goal is to create a multifunctional and “living” web platform that presents Ohneganos research while also facilitating digital access to all of the project’s co-created findings and tools.

Kayaní:yo (A good path)

Dissemination of science through the arts is in step with ancestral transmission of knowledge, encoded through song, dance, ceremony, and stories in our language. True to Haudenosaunee ways of knowing, the Ohneganos research project sought to co-create community stories that weaved traditional knowledge and western scientific data through visual, sensory, and emotive mediums. Communicating the research rationale, methods, and findings required a continual dialogue with numerous sectors of the community, engaging and communicating through various modes, such as by developing an immersive and interactive virtual reality (VR) experience and by 3D-modeling and data visualization. Another key aspect of developing a tangible and impactful data environment was integrating real-time water sensing data within a decolonial place-based mapping program called *Terrastories*, creating a holistic interactive map that houses both Haudenosaunee knowledge and technical water data. In this vein, the community also participated in the linguistic and cultural mapping of waterways as a method of decolonizing the local geography and creating an Indigenous map.

Yothá:te is a figurative “path” that implies a warrior’s intentions to be of aid – to be helpful; something which is considered as *kayaní:yo* (“a good path”) as it protects what our ancestors valued the most. In this way, co-creation approaches to science is putting our bodies back on the land and waterways, rediscovering our ancestral knowledge which has accumulated over thousands of years, and translating it to solve modern colonial-induced ecological problems. The elevation of IK, enhanced with WS and modern technologies, helps advance the protection of water, animals, and land ecosystems. The creation of ally networks and the application of scientific pluralism – when deemed appropriate – is, in fact, *Warrior Science*; it is clearing the *yothá:te* through rebuilding and nurturing passions, spiritual connections, and peaceful relationships to mitigate the climate crisis and rehabilitate biodiversity loss.

The climate crisis is leading to a period understood in scientific circles as the sixth mass extinction. The previous mass extinction was defined by the Cretaceous-Paleogene extinction event, during which the dinosaurs, along with thousands of other species, perished. A notable difference is that the current period of extinction we are in is manmade; it has not been caused by a natural event. Accordingly, anthropologists argue and geologists debate over if we have moved out of the Holocene and into a third Quaternary epoch, the *Anthropocene* (Zalasiewicz et al., 2011). However, it could be further posited that this would be a misnomer for the following reason: the United Nations Department of Economic and Social Affairs is clear in their analysis that Indigenous peoples, who are also human beings and part of the *anthropos*,

... account for only around 5 percent of the world’s population, [yet] they effectively manage an estimated 20–25 per cent of the Earth’s land surface. This land coincides with areas that hold 80 per cent of the planet’s biodiversity and about 40 per cent of all terrestrial protected areas and ecologically intact landscapes. Indigenous peoples therefore play a key role in efforts to protect the planet and biodiversity.

UN DESA (2021, p. 1).

Hence, perhaps this proposed epoch should be renamed the *Colonialcene*, for it is the imperialism of the past five centuries that has led to the great extinctions of the Holocene. Colonialism and western industrialism have driven climate change and the introduction of invasive species, accelerating the loss of life as plant and animal habitats disappear or are disrupted by new species. In fact, Michael Novacek, Senior Vice President of Science at the American Museum of Natural History, has stated, “... some biologists think that the current rate of species loss is probably a thousand times what the normal rate is ...” (AMNH, 2015). What mainstream scientific literature fails to acknowledge is identifying the socio-political driver of “human activity” that has led to the global crisis of biodiversity loss.

Co-creation has its challenges, which is often nested in the arrogance of western colonial epistemological constructions of the natural world and its peoples. Indigenous ecological knowledge (also commonly referred to as Traditional Ecological Knowledge or TEK) has been, and continues to be, dismissed by data-driven western scientists – even when the data is speaking clearly in global studies that are emphasizing the validity of Indigenous ways of knowing,

... this report recognises the value of diverse forms of knowledge such as scientific, as well as Indigenous knowledge and local knowledge in understanding and evaluating climate adaptation processes and actions to reduce risks from human-induced climate change.

IPCC (2022, p. 9).

Our oral histories tell us that we have experienced planetary climatic changes and cataclysmic events throughout history, and we have survived. We have stories about the Ice Age, the Great Fire caused by thunderers that chased the scary animals underground, and the Great Floods. In fact, Elders have recited and warned for nearly a century that settlers are creating an unsurvivable world for humans (Akwasasne Notes, 2005). IK draws upon this intellectual legacy while reinvigorating our language, stories, and ancestral wisdoms. By creating alliances with activists, artists, scientists, and leadership we catapult our voices into the consciousness of mainstream society. We carry out *kayaní:yo* by utilizing every tool in the toolbox, from social media to water monitoring sensors, VR technology, and scientific data circles. In doing so, we do more than decolonize, we lead the way. The overarching intention of Ohneganos was to demonstrate the successful harmonization of IK and WS by co-creating an enduring legacy of “tools” that target a holistic assessment of the environment, underpinned by Indigenous conceptions of human wellness and quality of life.

Equity in scientific inquiry: Indigenous Knowledge dismantling structural barriers

In an article titled *Why Indigenous Studies Matter for Natural Sciences Students* Hughes (2018) summarizes her work by stating,

... This article encourages students of the natural sciences to remember the importance of indigenous knowledge. Indigenous peoples work with scientists to share traditional ecological knowledge that has been honed over generations and allows for increased stewardship of natural resources. Indigenous knowledge should be integrated into university programs and Western scientists need to collaborate with indigenous groups towards conservation efforts.

The calls to include Indigenous ecological knowledge in the natural sciences appear to be a good *yothá:te*, but is it a *kayaní:yo*? There is the matter of colonialism framing dialogue between westerners and Indigenous people and, as a result, the potential for exploitation, appropriation, co-opting, and whitewashing IK is extraordinarily high. The hierarchical nature of academia and science is oppositional to IK, with structural racism in place that has been acknowledged by numerous Canadian inquiries (see RCAP (1996), TRC (2015), MMIWG (2019)). Yet, despite a chorus of calls to action to remove such structural barriers, nearly every recommendation on this matter has been ignored by the Canadian government and the general public.

Researchers from community, the local immersion school, and Indigenous undergraduate students – all who are part of the *TEK team* – have exhibited the Haudenosaunee ecocentric mixed-methods pedagogy through the recruitment of local stakeholders, such as service department representatives and secondary school teachers. Students and teachers working with the Ohneganos team are funded to create resources and programming, provide sustainable capacity and knowledge mobilization, and carry out research. Importantly, these students are also being credited by partners from McMaster University and Mohawk College (and potentially more local institutions; the opportunity is open to all who express enthusiasm and willingness to create and engage in the delivery of mixed-method programs). There have been preliminary agreements to formally accredit existing community researchers and develop a pathway for secondary students to pursue post-secondary environmental studies and water management programs that integrate experiential learning. An example is McMaster University Continuing Education agreeing to work on a pathway to credit the local language immersion school in Six Nations (Kawenní:io/Gawení:yo Private School) and students from the Six Nations Polytechnic the STEAM academy; teachers and educational assistants from which are assisting in current and ongoing knowledge co-creation and program development. These efforts include developing water-related digital stories, mapping traditional harvesting activity on water, and co-creating IK-led material in a bilingual text on Haudenosaunee Science. Mohawk College, McMaster University, and Six Nations education leaders have agreed to form an ad hoc committee to work together on multi-level accreditation pathways and programming to sustain forward solutions.

The goal of Indigenous-led research moves beyond single questions and answers; we must use research resources, expertise, and networks to solve community problems *while also* answering “the questions”. One way to develop capacity is by establishing sustainable ways of teaching IK and WS to train environmental researchers in ecological mixed-method approaches prior to engaging in research projects. Currently, the Indigenous scholars working on Ohneganos are burdened with on-the-job training of graduate students and faculty (e.g., professors from engineering, biology, health sciences, etc.) which imposes stress on all parties.

There is also a lack of equity in knowledge production in ecological research. Research contributions by Six Nations secondary students, undergraduate students, and knowledge holders are original and highly valued, but only one party receives scholarships and research funding resulting in Doctorates or Master’s accreditation. Indigenous Knowledge has no such support, scholarships, or funding equal to their partners and minimal accreditation; yet their contributions are imperative to the overall project design, implementation, and dissemination. Future research engaging Indigenous ecological knowledge should be strategic and thoughtful (with cumulative accreditation) to ensure an outcome that formalizes an IK/WS program that is operated by Indigenous Studies departments within institutions, as well as the people and community. This would address some of the current issues in academic institutions. Currently, the research funding, academic achievement, and unequal status of non-Indigenous researchers working on

Indigenous research projects gains awareness by community partners; this is discouraging, and community partners become disengaged overtime. As described, the IPCC, UN DESA, and other UN agencies value Indigenous ecological knowledge, but provide little means of structural support to elevate or advance IK-led research. Another significant barrier in equity is lack of practical bilingual teaching tools – from secondary school resources through to university level texts. As described, Ohneganos has been working to fill such gaps by developing resources and organizing them into a series of learning materials targeting both secondary and post-secondary students.

A primary example of the incredible contributions of Six Nations researchers is the Indigenous mapping of Haudenosaunee lands and waterways using *Terrastories*. The TEK team is carrying out GIS digital and decolonial mapping with the help of consultants from *Digital Democracy*⁶ as a way of storying our territory. Students and community researchers are assisting in developing Haudenosaunee placename maps that include digital stories, local oral histories, and traditional ecosystem inventories. The legacy of such an expansive Indigenous map will build on the seminal work of the *Decolonial Atlas*⁷ by focusing on Haudenosaunee heritage; the placenames of waterways, lakes, and aquifers; and by attaching both IK and WS data (e.g., historical cultural information about those bodies of water and water monitoring and climate change data). The digital archive accomplished through mapping will provide a teaching tool with language accessibility in both Mohawk and Cayuga. The digital maps will assert Haudenosaunee sovereignty and contest colonial boundaries. The information populating the map has been collected by secondary students with support from the Ohneganos Community Mapping Facilitator, Chris Martin, and would easily merit, at minimum, a Master's or Ph.D.; therein lie the inequity of knowledge production and accreditation. Aside from the programming and accreditation being co-led by Ohneganos, thus far we have only been able to produce ad hoc Certificates of Completion for student contributions with the support of McMaster University Continuing Education. This is a start, but this shows how community members receive no such recognition by institutions in community-research partnerships and demonstrates the need for institutional advancement in accreditation.

Key principles exemplified by the Ohneganos project include: a) defining research goals that are directed by Indigenous communities, and b) designing and implementing the co-creation of research, content, materials, and tools by observing ethical guidelines such as the *Tri-Council Policy Statement 2 Chapter 9: Research Involving the First Nations, Inuit, and Métis Peoples of Canada* (see [Government of Canada, 2018](#)) and the *Principles of Ownership, Control, Access, and Possession* (see [Schnarch, 2004](#)). Together, these guiding principles work to uphold the inherent rights of Indigenous Peoples (see [UNDRIP, 2007](#)) and Indigenous data sovereignty, helping the research unambiguously exemplify the phrase “nothing about us, without us”. Further, ecologically-informed community-developed capacity to monitor environmental health will build on foundational mixed-method pedagogy and outcomes by facilitating targeted dialogue between Western-trained scientists and Indigenous experts in all fields. Developing pathways to empower Indigenous communities to manage control over their water as a primary indicator of wellbeing is a crucial mechanism for achieving self-determination. Creating such pathways for Indigenous youth has had positive impacts on their careers and aspirations and created community pride in the work of the youth.

Reconnecting youth to waterways and tradition

Well-being among Indigenous people is firmly rooted in the connections to the land, the sea, and fresh waters, “... as a subsistence society, living off the natural resources of the land was fundamental to the social identities [of Natives] ... [an] intimate relationship in the reciprocal nature of caring for the land ... as it cares for the people, much like a family bond” ([Kanaŋiaupuni and Malone, 2006](#), pp. 289–290). Recognizing this, the TEK team supports the Two Row on the Grand River, a nine-day canoe paddle and symbolic enactment of the original Two Row treaty (see [Parmenter, 2013](#)) down the entire length of the Grand River in Ontario – fully equipped with educational events, traditional teachings, interactive dancing, storytelling, and sharing circles. Through involvement with the Two Row paddle, Ohneganos supports youth capacity-building by facilitating experiential mixed-method learning – on the water, about the water.

Knowledge mobilization and community engagement

The coronavirus pandemic interrupted many aspects of the Ohneganos project, from collecting water samples in homes to youth outreach and training. However, it also forced our team to think creatively and embark on a new journey of discovery. The *Ohneganos: Let's Talk Water* (Let's Talk Water) video podcast is an ongoing project hosted by Ohneganos youth lead, Makasa Looking Horse, with the help of Ohneganos full-time student staff. The production – which has been immensely successful in data sharing, awareness building, and co-creation – was a response to the COVID-19 lockdown and acted as a means to continue dialogue with the community, share research results, and elevate youth voices. Leveraging the power of social media, the show was hosted live on Facebook and YouTube with an average number of 600 views over the first 19 episodes (with a maximum of over 1000 views!). The Ohneganos Facebook page also skyrocketed from 54 to 1100 likes/follows. Importantly, Let's Talk Water also served as a safe,

⁶See <https://www.digital-democracy.org/>.

⁷See <https://decolonialatlas.wordpress.com/>.

innovative, and real-time methodology for collecting data, providing training, and disseminating research results. It required training of our all-female student team – being taught by consultants on cinematography, live productions, and film and sound engineering – as well as required sensitivity in recruiting and hosting a diverse set of guest speakers and participants. Let's Talk Water relied heavily on Makasa's network of well-known Indigenous singers, activists, Indigenous Knowledge holders, change-makers, and scientists to provide fresh and exciting content. The "science" was weaved into these conversations in a way that made sense to Indigenous audiences, though it is recognized that it may have been a bit confusing to the western-trained team members. The tensions of colonialism are never held far at bay in "co-creating" and can be, at times, draining for all parties. As such, co-creation requires an ongoing level of patience, understanding, and perseverance.

Public dissemination matters

The attention Let's Talk Water garnered in raising Indigenous voices poured in outreach from Vogue, Chatelaine Magazine, Yes! Magazine, Aboriginal Peoples Television Network, Vice News, CBC News and many others. The visibility then generated interest from multiple Indigenous communities across Canada, the United States and elsewhere because of Let's Talk Water's exploration of youth experience and the numerous consequences associated with water insecurity in Six Nations and other First Nations communities.

Youth represent the leading demographic in Indigenous communities and, as such, they are the driving force. In an era of climate change, Indigenous youth carry an added stressful burden – the burden of safeguarding the planet. Let's Talk Water is youth-led program giving a voice and platform to Haudenosaunee youth, as well as inspirational leaders from across the globe on the front lines of environmental justice, resistance to oil & tar sands, MMIW, health, and healing. Witnessing overt destruction to Indigenous lands and waters is an ongoing assault on Indigenous being. A recurring theme of Let's Talk Water repeated routinely by Makasa is that Indigenous youth need to connect more than ever while experiencing overwhelming ecological degradation and exploitation by corporations. In an upcoming film on ecological grief, Makasa describes the continued assault on local waters and land by corporations as "a million little cuts on my soul",

... whether it's Nestlé taking our precious water from the Grand River watershed, or my father's waters and lands marred by the Dakota Access Pipeline [DAPL], corporate greed and the extractive resource industry is devastating our planet ... [yet] we have found out that Nestlé is leaving Canada, the DAPL has been ordered to stop production until a review is complete and other pipeline projects halted, all due to youth activism and allies. The efforts of young water protectors have made impacts, but we have no definitive platform to share our experiences – such as those fighting at Standing Rock.
Looking Horse (Publication Pending).

In 2021, *Ohneganos: Let's Talk Water* won the David Suzuki Foundation's Future Ground Prize, People's Choice Award. Ohneganos was able to market the project's research to an audience that cares about Indigenous youth, climate change, and water security. At the same time, it also engaged people who aren't necessarily involved in academic science projects or who aren't Indigenous youth. The video podcast was a tremendous accomplishment while the community was reeling from the pandemic lockdown and demonstrated the strength and determination of a community-led research project, illustrating how incredibly resilient Indigenous communities are in the face of adversity.

All co-creators, Indigenous and non-Indigenous alike, rose to the occasion in accessing co-workers, allies, scientific colleagues, and community to build the Let's Talk Water social media knowledge mobilization platform. In each episode of Let's Talk Water, the viewers were routinely reminded of the era we are operating within; the sixth mass extinction – the *Colonialcene*. In battle, the meaning of *kayaní:yo* (a good path) can also mean "things are running along peacefully". Similarly, *yonkwarihwani:ron niyonkwarihò:ten* translates as, "we are steadfast in our ways of being". When our ways of being are not steadfast, we remove ourselves from that *yothá:te*, requiring time to reflect on ways to achieve peaceful relations without damaging our ways of being. The warrior is simply a skilled worker for the people to benefit their health and well-being; a worker with courage who is highly valued, paying close attention to the messaging in an effort to balance positives and negatives. Co-creation requires warriors who are willing to make some sacrifices so the people and creation may live, working through the tensions, discomfort, unequal power relations and social capital to remain on our *yothá:te yonkwarihwani:ron niyonkwarihò:ten*.

References

- Akwesasne Notes, 2005. Basic Call to Consciousness. Native Voices.
(AMNH) American Museum of Natural History, 2015. Episode twelve: six extinctions in six minutes. Shelf Life. Available at: <https://www.amnh.org/shelf-life/six-extinctions>.
Armitage, H., 2020. Polynesians, native Americans made contact before European arrival, genetic study finds. Stanford Medicine News Center. Available at: <https://med.stanford.edu>.
Bunch, M.J., et al., 2011. Promoting health and well-being by managing for social–ecological resilience: the potential of integrating Ecohealth and water resources management approaches. *Ecol. Soc.* 16 (1), 6. Available at: <http://www.ecologyandsociety.org>.
Burke, M., et al., 2018. Higher temperatures increase suicide rates in the United States and Mexico. *Nat. Clim. Change* 8 (8), 723–729. <https://doi.org/10.1038/s41558-018-0222-x>.

- Castellano, M.B., 2000. Updating aboriginal traditions of knowledge. In: Hall, B.L., Jerry Sefa Dei, G., Goldin Rosenberg, D. (Eds.), *Indigenous Knowledges in Global Contexts*, pp. 21–36.
- Chandler, M.J., Lalonde, C., 1998. Cultural continuity as a hedge against suicide in Canada's first nations. *Transcult. Psychiatr.* 35 (2), 191–219. <https://doi.org/10.1177/136346159803500202>.
- Daes, E.-I., 1994. Special Rapporteur: Protection of the Heritage of Indigenous People. United Nations Economic and Social Council.
- Finn, S., Herne, M., Castille, D., 2017. The value of traditional ecological knowledge for the environmental health sciences and biomedical research. *Environ. Health Perspect.* 125 (8), 085006. <https://doi.org/10.1289/EHP858>.
- Government of Canada, 2017. Budget 2017: Building a Strong Middle Class. Available at: <https://www.budget.gc.ca>.
- Government of Canada, 2018. Research involving the first nations, Inuit and Métis peoples of Canada. In: *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2*, pp. 107–132. Available at: <https://ethics.gc.ca>.
- Hughes, J., 2018. Why Indigenous Studies Matter for Natural Sciences Students. *Keystone Masterstudies*, 16 August. Available at: <https://www.nps.gov>.
- (IPCC) Intergovernmental Panel on Climate Change, 2022. *Climate Change 2022: Impacts, Adaptation and Vulnerability (Summary for Policymakers)*. Available at: <https://www.ipcc.ch>.
- Intergovernmental science-policy platform on biodiversity and ecosystem services, 2019. *The Global Assessment Report on Biodiversity and Ecosystem Services: Summary for Policymakers*. Available at: <https://ipbes.net>.
- Ioannidis, A.G., et al., 2020. Native American gene flow into Polynesia predating Easter Island settlement. *Nature* 583 (7817), 572–577. <https://doi.org/10.1038/s41586-020-2487-2>.
- Johnston, C.M., 1964. *The Valley of the Six Nations: A Collection of Documents on the Indian Lands of the Grand River*. University of Toronto Press. <https://doi.org/10.3138/9781442618510>.
- Kanaŋiaupuni, S.M., Malone, N., 2006. This land is my land: the role of place in Native Hawaiian identity. *Hūlili* 3 (1), 281–307. Available at: <https://kamehamehapublishing.org>.
- King, J.T., 2007. The value of water and the meaning of water law for the native Americans known as the Haudenosaunee. *Cornell J. Law Publ. Pol.* 16 (3), 448–472. Available at: <http://scholarship.law.cornell.edu>.
- Kirmayer, L.J., Valaskakis, G.G., 2009. *Healing Traditions: The Mental Health of Aboriginal Peoples in Canada*. UBC Press. Available at: <https://www.ubcpres.ca>.
- Longboat, S., 2013. First nations water security – security for Mother Earth. *Can. Woman Stud.* 30 (2,3), 6–13.
- Looking Horse, M. (Publication Pending). *Ecological grief film. Indigenous Youth-Led Documentary*.
- Majeed, H., Lee, J., 2017. The impact of climate change on youth depression and mental health, 1. Elsevier, pp. e94–e95. [https://doi.org/10.1016/S2542-5196\(17\)30045-1](https://doi.org/10.1016/S2542-5196(17)30045-1).
- Martin-Hill, D., 2009. Traditional medicine and restoration of wellness strategies. *J. Aboriginal Health* 5 (1), 26–42. Available at: <https://jps.library.utoronto.ca>.
- McGregor, D., 2012. Traditional knowledge: considerations for protecting water in Ontario. *Int. Indigenous Policy J.* 3 (3), 1–21. <https://doi.org/10.18584/ijpj.2012.3.3.11>.
- (MMIWG) National Inquiry into Missing and Murdered Indigenous Women and Girls, 2019. *Reclaiming Power and Place: The Final Report of the National Inquiry into Missing and Murdered Indigenous Women and Girls*, vol. 1a. Available at: <https://www.mmiwg-ffada.ca>.
- Mussell, B., Cardiff, K., White, J., 2004. *The Mental Health and Well-Being of Aboriginal Children and Youth: Guidance for New Approaches and Services*. Available at: <https://childhealthpolicy.ca>.
- Nakashima, D., Krupnik, I., Rubis, J.T. (Eds.), 2018. *Indigenous Knowledge for Climate Change Assessment and Adaptation*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781316481066>.
- National Aboriginal Forestry Association, 1996. *Aboriginal Forest-Based Ecological Knowledge in Canada (Discussion Paper)*. Available at: <http://www.nafaforestry.org>.
- Parmenter, J., 2013. The meaning of Kaswentha and the two Row Wampum Belt in Haudenosaunee (Iroquois) history: can indigenous oral tradition be reconciled with the documentary record? *J. Early Am. Hist.* 82–109. <https://doi.org/10.1163/18770703-00301005>.
- (RCAP) Royal Commission on Aboriginal Peoples, 1996. Volume 1: Looking Forward, Looking Back. Available at: <https://www.bac-lac.gc.ca>.
- Schnarch, B., January 2004. Ownership, control, access, and possession (OCAP) or self-determination applied to research. *J. Aboriginal Health* 80–95. Available at: <https://jps.library.utoronto.ca>.
- (TRC) Truth and Reconciliation Commission of Canada, 2015. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. Available at: http://nctr.ca/assets/reports/Final_Reports/Executive_Summary_English_Web.pdf.
- (UNFCCC) United Nations Framework Convention on Climate Change Assessment and Adaptation, 2016. *A Compilation of Good Practices, Tools and Available Data Collection Initiatives for the Use of Local, Indigenous and Traditional Knowledge and Practices for Adaptation*. Available at: <http://unfccc.int>.
- United Nations, 2007. *United Nations Declaration on the Rights of Indigenous Peoples*. Available at: <https://www.un.org>.
- United Nations, 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development*. <https://doi.org/10.1007/s13398-014-0173-7.2>.
- (UN DESA) United Nations Department of Economic and Social Affairs, 2021. *Challenges and Opportunities for Indigenous Peoples' Sustainability. Policy Brief 101*. Available at: <https://www.un.org>.
- Zalasiewicz, J., et al., 2011. The Anthropocene: a new epoch of geological time?, 369 *Royal Society*, pp. 835–841. <https://doi.org/10.1098/rsta.2010.0339>.

"Indigenous" ways of knowing in Chinese Minzu education: teachers reflecting on their engagement with STEM in inner Mongolia

Fred Dervin^a and Mei Yuan^b, ^a University of Helsinki, Helsinki, Finland; and ^b Minzu University of China, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	217
Problematising <i>indigeneity</i> for the context of Mainland China: a critical language review of the word <i>indigenous</i>	217
Minzu as a companion to indigeneity?	219
Minzu knowledge in Chinese education	220
How to identify Minzu knowledge?	221
Mongolian STEM teachers reflecting on the use of Minzu knowledges in their teaching	222
Discussion and conclusion	226
References	227

Introduction

In the current era of the *(critical) social justice turn* in education research and practice, epistemic diversity has been increasingly promoted, discussed, and problematized (e.g., Xu, 2022; Piller et al., 2022). China, the context of this article and one of the world's most diverse countries (Yuan, 2018), has also been active in trying to push for her own epistemic inclusion globally, as can be seen in the use of discourse instruments such as "Chinese cultural confidence" (文化自信) or "world-class education with Chinese characteristics" (具有中国特色的世界一流教育) in Chinese education (Dervin, 2022). Often, the ideas of "confidence" and "Chinese characteristics" remain unproblematised in relation to Mainland China and its 56 official Minzu (民族) "ethnic"¹ groups, especially in relation to indigenous knowledges.² This article aims to complement these discussions by focusing on such knowledges, from teachers' perspectives, in Chinese secondary education.

In Chinese academic circles, there is a special field attached to Minzu groups called Minzu education (中国民族教育), which, for at least the past 50 years or so, has endeavored to do research on and e.g., prepare teachers for working with different Minzu groups (Sude et al., 2020). Although research on some aspects of Chinese Minzu education has been discussed internationally (often under "Western" guises such as American multicultural education, see e.g., Wang, 2021), to our knowledge, none of these studies have looked at it from the lens of epistemic diversity and indigeneity.

This paper, which is embedded in the two authors' long-term cooperative work on Minzu education and interculturality (e.g., Dervin and Yuan, 2021), first problematizes the notion of indigeneity in relation to the Chinese Mainland by exploring how it is (not) used and offers a linguistic discussion of words related to the notion in Chinese. The notion of Minzu is then presented and offered as an epistemic complement for the English word *indigenous* in relation to knowledges, alongside with a short review of previous research on the topic. The chapter ends on an analysis of interview data collected among Chinese middle school teachers in a Minzu area (Inner Mongolia, Northwestern China) about their engagement with Minzu (indigenous) knowledges in their teaching. This serves as a case study and is not meant to be generalized to other contexts of Minzu education within Inner Mongolia and in the Middle Kingdom as a whole.

Problematising *indigeneity* for the context of Mainland China: a critical language review of the word *indigenous*

In this first section we discuss word use in Chinese and English. We note that both media and academic discussions of general Minzu affairs have been tense in the West, especially since the beginning of the 2019 pandemic in the midst of an increasing economic-political polarization of the world. These discussions appear to be systematically based on "Western" and especially US-led ideological takes on "race" and "ethnic" relations, which do not leave enough space to explore how the Chinese themselves engage with the notion of Minzu in academia, education and society at large (see Dervin and Yuan, 2022). Word use in different languages and the connotations of specific terms to discuss e.g., *diversity*, *majority/minority*, *indigeneity*, must be examined here to ensure that we are not simply "copy-pasting" ideological takes onto another context. We believe that this also helps us avoid universalizing the way we see self-other in specific societies. What follows is not meant to be read as "in defense of" China. As such, like all

¹The word "ethnic" is between inverted commas to indicate that the connotations of the English word, which are unstable worldwide, considering the very different economic-political contexts where it is used, may not reflect the specificities of the Chinese word Minzu. As a reminder, the English word "ethnic" was used first to refer to individuals who were not followers of either the Christian or the Jewish faiths.

²Since other parts of China such as the Special Administrative Region of Hong Kong do not follow similar patterns of classifications of "ethnic" groups, we do not discuss their specific cases here.

other countries, China is struggling to ensure equality, equity and social justice among its 1.4 billion inhabitants and many “Western” and Chinese scholars have attempted to describe such struggles in their own terms (e.g., Hao, 2020; Wang and Jiang, 2021; see Pasternak, 2020 about the context of this study, Inner Mongolia). Our goal here is to send a reminder concerning *multi-lingual* and *ideological* ways of speaking and interpreting other contexts to those of us working on indigeneity from international (comparative) perspectives in education.

Let us start with etymology. The English word *indigenous* derives from Latin *indigena* for *sprung from the land, native, in-born in a place*. It is based on Old Latin *indu* for *in* and *within* as well as *gignere* for *to produce, to beget*—with the latter derived itself from Proto-Indo-European for *to give birth*, and the idea of *familial* and *tribal* groups. In Chinese *indigenous* translates as 土著 (tǔzhù), with *native* and *aboriginal* as synonyms in English. Each of the two characters have the following meanings: 土 is *earth, dust* and 著 can translate as *writing* but also *receive, suffer, and make known* (among others). The definition of the word seems to include the features of being a “native”, having lived in a specific area for generations, and having faced invaders and colonizers (Zhang, 2002). It is thus very close to the United Nations definition of indigenous. This definition often includes at least three groups in the Chinese definition: *Aboriginal Australians, Native Americans* and *African Indigenous groups*.

Our next interest is in how the word *indigenous* has been used in recent international research in relation to China, with many different uses having been noted in English. The main strand relates to different fields of research and applies the word *indigenous* to China as a whole (in chronological order):

- Indigenous roots of modern Chinese psychology (e.g., Jing and Fu, 2001);
- Chinese indigenous swine (Kang et al., 2015) and goat (Yue et al., 2017) breeding;
- Theorizing China’s indigenous innovation capabilities beyond a traditional Western-centric theory of innovation (e.g., Vinig and Bossink, 2015). The Chinese term used in the global literature is 自主创新, which means *autonomous innovation*—*autonomous* referring here *to act for oneself*;
- Indigenous traditions of Chinese religions (Zhuo, 2017);
- Cultural roots in indigenous Chinese communication theories, including visualizing and modeling Chinese cultural factors like wind, grass, and water (Tian and Yu, 2022).

Some of the literature does use the word *indigenous* to refer to the knowledges of specific (“ethnic”) groups and communities within Mainland China (in chronological order):

- Indigenous knowledge of dye-yielding plants by Bai Minzu communities in Yunnan (Southwest) (Fan et al., 2018);
- Indigenous psychology among the Yi Minzu group in Southwestern China (Ting et al., 2019);
- Indigenous knowledge in integrating scientific and indigenous knowledge for community-based disaster risk reduction, e.g., taught by oral form through storytelling, in the Ningxia Hui Minzu Autonomous Region in Northwest China (Wang et al., 2019);
- Ethno-ecological knowledge of indigenous people regarding the use of ecological cues to conduct seasonal activities in relation to climate change in Tashkurgan, Xinjiang Province, near the Pakistani, Tajik, and Afghani borders (Yang et al., 2019);
- Indigenous knowledge in forest and wildlife conservation by Buland People in Yunnan Province (Su et al., 2020).

One notes that the fields of research include communication studies, forestry, psychology, (among others) but that, for example, no research on education has made use of the word *indigenous* in the English language, preferring another problematic term: “ethnic minorities”. Before we offer more details as to why in the next section, we comment on one paper that offers further elements to the linguistic discussions of Chinese equivalents to *indigenous* in Mainland China.

In *An Alternative to the “Indigenous” in Early Twenty-First-Century China: Guizhou’s Branding of Yuanshengtai* Luo (2017) examines the emergence and use of the Chinese concept of 原生态 (yuánshēngtài) through cultural industries and marketization at the turn of the 21st century. The concept translates as *original ecology* in English (原 means *original, former* and is used in the Chinese word for *aborigines*; 生态 means *ecology*). Luo (2017) argues that this rich term is difficult to translate properly in English and that some of its translations are problematic (e.g., *pristine* or *primitive*). The term refers to distinctive “native” and “original” cultural elements that have survived for hundreds of years, including dance, singing, “freehand” landscape painting but also housing, food and ecological protection. For example, folk songs such as *Manhan tunes* (a mix of Mongolian and Han songs), *the Twelve Muqams* (ancient Uigur music, combining music and dance) are considered as 原生态. Academia, the media and the cultural industry have made ample use of this term over the past decades in China. Assisting the development of the local economy through 原生态 is also often noted (Luo, 2017). Both ethnic and rural peoples from certain parts of China (e.g., highly multi-Minzu provinces like Yunnan on the country’s southwestern frontier, with over 40 million inhabitants) seem to be included in its use. We also note that Luo (2017: 74) is critical of the use of 原生态, calling it “romanticized symbolism (...) which conflates ethnicity, rurality, and nature, has come to be expressed through consumer desire”. She adds that the use of the concept could disfranchise discussions of historical and political issues, by emphasizing political economy and intellectual production (Luo, 2017: 90). The important issue of who represents who in the use of the concept is also put forward by Luo (2017). However, in her article (2017), she also shows that many of the discussions about environmental sustainability and heritage around the concept are very much reminiscent of international discussions of conservation around indigeneity. Finally, she maintains (2017) that the positive tone of 原生态 has contributed to e.g., valorize Chinese Minzu groups for their environmentally based knowledge and practices in China today. The notion of Minzu (民族), a central “companion” term to *indigenous* in the Chinese context, is discussed in the next section. We propose to use the pinyin romanized version of the notion in English, instead of *indigenous*, in our chapter.

Minzu as a companion to indigeneity?

In 2016 Hathaway wrote a paper describing how global environmental organizations had “unintentionally smuggled” the idea of *indigeneity* into China, although the Middle Kingdom does not use this notion to describe its diversity of peoples. As a reminder, there are many Chinese equivalents to the word and none of them seem to fit the polysemic English term. It is important to note at this stage that China *does* support indigenous rights internationally as stipulated by the United Nations (China signed the United Nations Declaration on the Rights of Indigenous Peoples in 2008) but, as Hathaway notes (2016), the country does not apply the notion to itself, using the word *Minzu* (民族) instead to describe itself as a land of “ethnic groups”. In his paper Hathaway shares his views on the question (2016: 2):

In the course of my research over the last twenty years, I have been asked many times: are there indigenous people in China or not? My answer has always been that it depends. The term is almost always understood as a natural category and the expected answer is either affirmative or negative: do any groups in China meet the criteria commonly expected of indigenous peoples? I argue that indigenous status is a political category. Identifying with the category of indigenous is a political strategy as much as it is an identity.

In a similar vein, in *The definition of indigenous peoples and its applicability in China*, Wang (2015) questions the applicability of the concept from an international law perspective. Commenting on the uncertainty of the definitions of indigeneity around the world, the complexity and specificities of the *Minzu* situation and the rejection of the term *indigenous* by Chinese authorities to talk about China, Wang urges us to clarify the Chinese terms in relation to Chinese *Minzu* (which he himself calls “ethnic minorities” in English...).

In her discussions of 原生态 (“original ecology”) in the aforementioned article, a term in Chinese used to refer to “indigeneity” from within, Luo (2017) also introduces the notion of *Minzu*. She reminds us that the term is polysemic in English and that it has been translated as *nations*, *nationalities*, *ethnicities*, and even *races* (Thomas, 2022; Dervin and Yuan, 2021; Dikotter, 2015). According to Luo (2017), 原生态 is often used to promote the ideas of “harmonious multiculturalism” and “unity in diversity”, which are also central discourse components of Chinese *Minzu*.

Equivalents to the word *indigenous* are thus many and varied in Chinese. We have also noticed that some Chinese might use it in English, without being aware of its differing connotations and ideological bases. They might use it to refer to all Chinese and to specific *Minzu* groups. In a similar vein, the same people might use the word *nationalities* in English to refer to *Minzu*. The fact that the authorities do not use the polysemic English term of indigenous to refer to the diverse situation of China does not mean that the Middle Kingdom disregards *Minzu* “minorities”. We now turn to the very word *Minzu*.

From a research and political perspective *Minzu* (民族) appears to be the closest to *indigenous*. Let us share a few more pieces of information about the notion. The Chinese word is composed of characters that mean “folk or common people” (*min* 民) and “consanguinity or lineage” (*zu* 族); in other words: *community of folks with a common origin*. China is officially a unified country with a diverse *Minzu* composition, where rights for *Minzu* groups are protected by laws and inscribed in the *Constitution of the People’s Republic of China* (1982). Current *Minzu* policies favor fostering a strong sense of community for the Chinese nation to ensure *Minzu* unity in diversity. The Chinese government recognizes 56 peoples of *Minzu* groups, with one group representing the *Minzu* “majority” (Han). It is important to note that about 800,000 people are still not recognized as being members of specific *Minzu* groups. The largest non-Han *Minzu* groups include e.g., Uyghurs, Mongols and Tibetans. The 55 *Minzu* “minority” groups represented around 8.9% of the population in 2020 (National Bureau of Statistics of China, 2021). Since China shares borders with 14 countries, a vast majority of *Minzu* groups are located on borders. Regional *Minzu* autonomous areas include for instance the Inner Mongolia Autonomous Region and the Xinjiang Uyghur Autonomous Region. Although Chinese citizens may have different mother tongues (e.g., Uyghur, Tibetan), most speak and study Mandarin Chinese, China’s common lingua franca—as much as in “Western” societies different languages and dialects might be spoken but one or two languages are used for official and educational purposes by all. It is important to say that economic development is central to *Minzu* affairs, with e.g., many economic stimulus programs for different *Minzu* groups having been established by the government.

Finally, we note that the notion of *Minzu* is included in the (2022) book *Changing Theory* (Menon, 2022), where the editor and authors chose linguistic concepts from 16 languages from Africa, the Arab world, Asia, and South America to open up global and multilingual conversations on social theory, beyond Euronormality (Kaviraj, 2009) and away from Euro-American formations (Menon, 2022). In the introduction to the book, Menon (2022: 2) makes an important point: “It is not about rendering visible words and ways of thinking across the Global South through mere translation within a monolingual space, which entrenches the politics of English as a universal language of rendition (...). A true conversation must engage with the nuances and hardness of multilingualism as much as the possible quiddity of concepts”. In using the notion of *Minzu* (instead of the polysemic English word *indigenous*) in the rest of this chapter, we wish to contribute to render the specificities and complexities of this notion in the English language, as a complement to discussions of indigeneity in education.

Minzu knowledge in Chinese education

The focus of this chapter is on the potential use of “alternative” Minzu knowledge in STEM education in a specific area of China (Inner Mongolia). This subsection problematizes the idea of Minzu knowledge, considering both discussions of epistemic diversity and indigenous and traditional knowledge. Before we do so, we provide a short review of how these issues have already been considered in research on Chinese (teacher) education.

Let us start with the field of *Minzu education* in Chinese Mainland (Sude et al., 2020), which is mostly unknown to international audiences. Based on anthropologist Fei’s (1989) theory of unity in diversity, Minzu education aims to acknowledge the diversity of different groups while promoting their togetherness and Minzu solidarity. For Wang (2021: 112), in China, “celebrations of cultural diversity must be built on support for unity”. A variety of preferential policies have been established to ensure educational opportunities and the socio-economic mobility of members of all Minzu groups. For example, students from certain Minzu groups, such as the Mongolian Minzu group under review in the next section, get bonus points in high school and college entrance examinations. Kong and Yu (2019) summarize the characteristics of Minzu education as follows: Equality and national solidarity as fundamental principles, as well as mutual assistance between Minzu groups and unfailing commitment to bilingual education and teacher quality. They note however the following problems in Minzu education (Kong and Yu, 2019): An unequal power relationship between certain Minzu groups, e.g., patriarchal relationship between Han (the so-called “majority”) and some Minzu groups; one-way assistance; but also disconnect between language teaching and cultural competence; support for Minzu minority language, but not Minzu minority culture.

A few review articles have been published about Minzu education in English between 2013 and 2022 detailing what has been done in this field. For example, in 2018, Wei Wang published a review paper on education and ethnicity in China covering the period 1990–2014. Wang’s focus is on how these topics were researched in Chinese in top journals published in Mainland China (Wang, 2018). The scholar identified six themes in 431 journal articles: *educational policy*, *multicultural education*, *minority cultural heritage in school*, *ideological-political education*, *multilingual education*, and *teachers and teacher education*. He notes that the inclusion of cultural heritage in curriculum settings seems to dominate such studies while teachers and teacher education are limited (Wang, 2018). He adds that the dominating theoretical trend in the papers is “multicultural education” (its “American” branch). Another example of a review is from 2020 with Liu et al. (2020) review of research from 2000 to 2018 about what they refer to as “Chinese multicultural education”. The data used for the review consists of 179 journal articles published in Chinese (157) and in English (22). Their results are categorized into:

- Theoretical discourses on Chinese multicultural education;
- Multicultural curricula and instruction;
- Students’ perspectives in multicultural contexts;
- Multilingual education and language policy;
- Citizenship education;
- Multicultural teaching and teacher education.

The authors discuss for example how some authors such as Teng (2003) have tried to propose a form of multicultural education model for working with Minzu. They also show how the topic *multicultural curricula and instruction* tends to dominate the scholarship but remains “at the stage of calling for action” (Liu et al., 2020: 290). Like Wang (2018) they also note a lack of empirical research on Minzu teacher education. All in all, they confirm that the discourse of national unity and integration is omnipresent in Chinese Minzu education.

In 2021, Wang (2021) published a research paper about the inclusion of Minzu cultural content in school practices by means of e.g., bilingual education and cultural courses and activities (see Zhang and Chen, 2014 as an example of a previous study on the same topic). Based on data collected in two Minzu areas with representatives of different Minzu groups, the author notes what he refers to as “superficial modalities” in the way Minzu cultural content is integrated in schools or their mere additive nature (e.g., Minzu issues being proposed as extracurricular activities and “performed” solely during traditional festivals). He also notes that “Underlying cultural attitudes, values and beliefs are not reflected in school practices” (Wang, 2021: 122). Teachers appear to be interested in promoting Minzu cultures but, for instance, the fact that the inclusion of Minzu knowledges in their teaching does not play a role in their yearly appraisal seems to not motivate them to include them systematically (Wang, 2021). Wang explains (2021) this lack of inclusion by the centralized nature of educational governance and, especially, by means of the concept of “Chinese internal orientalism”, whereby Minzu cultures are represented in static ways. In his conclusions, Wang (2021: 123), influenced by American (critical) multicultural education, argues that:

At present, the restrictive and highly centralised curriculum framework makes it rather difficult for teachers, whatever their own ethnicity [i.e., Minzu], to break through the curriculum structure and use advanced approaches to multicultural education, such as transformative and social actions which would encourage students to think about concepts, issues and events from the perspective of diverse ethnic groups, and empower them to make decisions on critical social issues.

What the reviews we have consulted in Minzu education (and our knowledge of the literature) show is a lack of discussions around knowledges from the 56 Minzu groups for education, especially beyond “Western” critiques of multiculturalism.

There is another strand of research which could be of interest to interrogate Minzu “indigenous” knowledges: *Chinese multilingual education* (see e.g., Rehamo and Harrell, 2017). This other broad field of research covers national, Minzu and foreign languages and

often includes discussions of epistemic diversity. In their (2021) edited volume Adamson and Feng note threats to the sustainability of weak languages and dialects in China where over 100 languages are spoken on top of Chinese, with many Minzu groups having their own languages (e.g., Mongolian, Uyghur, Yi, Zhuang). In one of the chapters, Wang et al. (2021) write about what they call “Local Knowledge Instruction” (LKI) for Minzu learners in Yunnan, inspired by the American concept of *Culturally Responsive Teaching*. LKI is described as daily life and native literacy and numeracy knowledge linked up with school literacy development (Wang et al., 2021: 171). This appears to be the closest to the idea of indigenous knowledges. In their chapter, Wang et al. (2021) describe an initiative dating back to 2009 whereby Minzu University of China (a university focusing on Minzu studies located in Beijing) and some NGOs published two LKI textbooks for primary schools. Their evaluation of the initiative reports teachers’ mixed feelings about introducing such knowledge in Yunnan. As such local Minzu teachers felt that the Minzu knowledge found in the textbooks was too localized, which could create tensions among representatives of the Minzu in other parts of the country. Wang et al. (2021) also note that the fact that there is no pre- and in-service teacher training around LKI represents a major inconvenient. In general, reading the chapter, one is not too clear about what LKI includes, who decides what to include and how to deal with the potential epistemological clashes between (and among) local knowledges and traditions and school knowledge. All these essential questions remain at the center of interrogating indigenous knowledges in relation to Minzu education.

How to identify Minzu knowledge?

In this section, we take a closer look at *if* and *how* Chinese teachers of Mongolian Minzu working in the region of Inner Mongolia both justify and engage with the use of Minzu knowledge in STEM education.

In a 2022 paper, Xu discusses the idea of epistemic diversity and cross-cultural comparative research. In the paper she problematizes and criticizes the very term of epistemic diversity and especially what she refers to as “long-standing normative appeals for epistemic diversity, decolonization, democratization and justice” (Xu, 2022: 37). She notes that questions such as “who has the right to create and own knowledge, and to be respected, heard, seen in their capacity as a knower?” (Xu, 2022: 37) are being addressed rightly today, through different perspectives such as decolonization and postcolonialism. However, she also maintains that the answers provided to the aforementioned questions are often very much framed within Western critical discourses, which still evacuate alternative voices. Xu maintains that “academic voices from the non-dominating spaces are often again marginalized, ignored, appropriated or excluded, demonstrating yet another layer of epistemic injustice” (2022: 38). For research and education around indigenous knowledge, these issues are very much relevant: *how to get out of the vicious circle of “still” making use of westernized and dominant ideologies? Who decides what knowledge is and how it should be imparted? And who has the right to judge if imparted knowledges are valuable, “essentialistic” and/or “superficial”* (see Wang, 2018)?

Although we the authors work constantly with people of other Chinese Minzu groups, none of us belongs to a Minzu “minority” group. Fred is from Europe and Mei is a member of the Han Minzu group. This does not make us feel too comfortable “judging” e.g., Mongolian Minzu teachers for their use (or not) of Minzu knowledges. Therefore, for us, identifying Minzu knowledges in teachers’ practices in the next section will not be about *evaluating* what they do (or not) but about providing a first snapshot of the topics of indigeneity, Minzu and knowledges in Chinese education. We always need to remember that teachers are educated and trained in specific ways, that they have their own beliefs and attitudes toward political issues such as the inclusion of specific knowledge and that they may not want to reveal their “inner” thoughts about the topic of Minzu knowledges during a research interview. Finally, since the data we collected was based on interviews (rather than e.g., observations), we need to remember that what people say during interviews may not always correspond to what they do in complex teaching situations.

Deciding upon how to identify Minzu knowledges in the research interviews at our disposal required careful consideration. After consulting the literature, we have decided to make us of Bruchac’s (2014) paper entitled “Indigenous Knowledge and Traditional Knowledge”, which appears to be a good starting point. Anchored within the field of global archeology, the paper defines traditional indigenous knowledge as follows (2014: 3816):

a network of knowledges, beliefs, and traditions intended to preserve, communicate, and contextualize Indigenous relationships with culture and landscape over time. One might distinguish “knowledge” as factual data, “belief” as religious concepts, and “tradition” as practice, but these terms are often used imprecisely and interchangeably to describe Indigenous epistemologies.

A series of keywords found in the definition can help us adopt an open approach to Minzu knowledges when combing through our data: *knowledges* (understood as “factual data”), *beliefs* (“religious concepts”, we add worldviews here) and *traditions* (“practice”). These three elements are used as “gauges” to examine our data in what follows. Bruchac (2014: 3817) also provides a concrete list of such knowledges:

oral narratives that recount human histories; cosmological observations and modes of reckoning time; symbolic and decorative modes of communication; techniques for planting and harvesting; hunting and gathering skills; specialized understandings of local ecosystems; and the manufacture of specialized tools and technologies (e.g., flint-knapping, hide tanning, pottery-making, and concocting medicinal remedies).

Traditional ecological knowledge or “Native science” (Cajete, 2000) is also presented as an important element to take into account by the author (Bruchac, 2014). These include: “philosophies and practical measures that are intended to preserve cultural heritage and protect ancestral landscapes and lifeways” (Bruchac, 2014: 3818). Finally, the importance of oral traditions (e.g., mythical stories) is also emphasized by Bruchac (2014: 3820):

Indigenous oral traditions often contain insightful explanations that focus on details: origin stories referencing natural and constructed features of the landscape; descriptions of the beings that inhabit this landscape; articulations of the reciprocal relations among these beings; and traditional beliefs that guide human interactions with place.

Bruchac (2014: 3821) reminds us that indigenous knowledges are often viewed with suspicion both by common people and scholars, especially when they are contrasted to “scientific knowledge”. However, she also notes that indigeneity can enhance scientific knowledges as many anthropological and indigenous scholars have demonstrated (Bruchac, 2014: 3821).

Mongolian STEM teachers reflecting on the use of Minzu knowledges in their teaching

In the rest of the chapter, we focus on a case study based on interview data with 9 teachers collected in different schools in the Inner Mongolia Autonomous Region (Northern China, on the border to Mongolia and Russia). The region has approximately 25 million inhabitants for a total area of about 1.1 million km². The vast majority of the population is constituted by the Han Minzu. The second largest Minzu group is that of the Mongols (about 17%). Other Minzu groups living in the region include e.g., Hui people, Manchus and Koreans. Some Mongols are followers of Tibetan Buddhism while e.g., some members of the Hui group follow Islam (Sude et al., 2020). According to the 2020 census, 67.48% of the population lives in urban areas while 32.52% in rural areas (National Bureau of Statistics of China, 2021). Worlddata.info (2022) claims that 5.64 million people speak Mongolian in China. Many speakers of the language are also located in provinces other than Inner Mongolia such as Qinghai, Xinjiang and Gansu. Mongolian has its own writing system adapted from an old Uyghur script. Although there were rumors in the global media in 2021 about the Mongolian language being “phased out” from education in Inner Mongolia (e.g., in the Xilingol League in the central region), Mongolian schools do teach subjects in Mongolian and offer bilingual education too. Chinese language and literature (the common language of China), Ethics and Law, and History, are taught in Chinese. Students can choose to go to Minzu schools where the Mongolian language is the medium of instruction and Chinese a subject or a Chinese-medium school where Chinese is the language of instruction and their Minzu language is taught as a subject. We collected the data from both kinds of schools (T: 5) in the city of Erdos (鄂尔多斯), a prefecture-level city located in the southwest of Inner Mongolia. Together with Hohhot and Baotou, Ordos constitute the economic triangle of the region. A cluster of different cities, Erdos has a population of about 2 million people (Statista, 2022).

In what follows we listen to 9 teachers of Science Technology Engineering and Mathematics (STEM). 7 of the teachers were of Mongolian Minzu, who all spoke Mongolian and Chinese and 2 Han teachers who also spoke both languages. 5 of the teachers were teaching in a Minzu school (Mongolian as the dominant language and Chinese as a subject) and 4 in Chinese-medium schools. All the schools were 初中 (middle school), i.e., the first three years of secondary education in China. Some teachers had just one year of experience teaching STEM while others 15. Finally, all the teachers were working with groups of both Han and Mongolian Minzu students. Table 1 provides basic information about the teachers.

The interviews were done in Chinese and Mongolian, translated and negotiated in English between us. Like other parts of the world, STEM is an important subject in China and according to Li (2018: 2): “Efforts are being made in China to stimulate student inquiry, innovation, and critical thinking in and through STEM education”. Many scholars have worked on STEM education worldwide within the subfields of e.g., culturally responsive STEM or Multicultural STEM education, arguing e.g., that it can support both teachers and students in revising their beliefs about e.g., mathematics being “neutral”; testing out culturally responsive instructional

Table 1 Basic information about the STEM teachers.

<i>Teacher</i>	<i>School subject</i>	<i>Minzu</i>	<i>Type of school</i>
Teacher 1	Technology	Mongolian	Mongolian school
Teacher 2	Science	Mongolian	Mongolian school
Teacher 3	Chemistry	Han	Mongolian school
Teacher 4	Technology	Mongolian	Chinese-medium school
Teacher 5	Science	Mongolian	Mongolian school
Teacher 6	Chemistry	Mongolian	Chinese-medium school
Teacher 7	Maths	Han	Mongolian school
Teacher 8	Science	Han	Chinese-medium school
Teacher 9	Engineering	Mongolian	Chinese-medium school

strategies and being opened to students’ thinking and problem solving (Ukpokodu, 2011). By including indigenous knowledges in STEM, one might also promote more interest in them. In what follows we are interested in seeing if and how Mongolian Minzu knowledge is included in STEM education in the schools under review.

Let us start with some comments on training and education. Officially, STEM education with Minzu characteristics is not discussed at school or in teacher education and training. All the teachers agree that there is no formal teacher education and training related to Minzu issues for STEM. When she details her preservice training, teacher 3 provides information about how she has learned to use e.g., PPTs for showing “real examples from life”. When asked how much she took Minzu examples into account, she replied that she did not. Teacher 2 also explains that he has taken a lot of extra training courses during his years as a STEM teacher. None of these courses were specifically targeted at Minzu knowledges but he argues that the imparted knowledge “is applicable to all Minzu groups” and that everyone can learn from “the excellent culture”, meaning from general international STEM knowledge. The phrase “excellent culture” (优秀的文化, Yōuxiù de wénhuà) to refer to “superior” and “more advanced” knowledge is commonly used in Chinese.

Teacher 1, who is herself Mongolian, also confirms that she received no specific training since “we teach all students the same way”. However, further in the interview, she mentions that her Mongolian family background has urged her to be interested in extracurricular activities such as physical education and technology and to learn more by herself. She claims for example that she has also taught music, painting, science and technology, and physical education in primary school. Contrasting Han and Mongolian families, she argues:

Excerpt 1

Many Han parents do not support students to participate in extracurricular activities. They support them to participate in activities that are beneficial to them, such as piano, English, etc. But they don’t support students to participate in subjects that they find useless such as science and technology and sports. But our Mongolian parents never do like that. As long as students like it, parents will support them. They believe that students should be widely involved in all kinds of knowledge and develop in an all-round way. We grew up with such an educational concept.

Although this statement might appear somewhat essentialist (Han families vs. Mongolian families), we are here within *beliefs* about education rather than *Minzu knowledges*. At this stage we should note that contrasts between Han and Mongolian, teachers, parents and students are omnipresent in the interviews, with Mongolian students being systematically constructed as “freer”, “more gifted for craft, singing and sports” and “less hardworking”/“passive”. In terms of STEM the teachers seem to agree that Mongolian students are more curious, better at experiments and manipulations, and eager to take part in activities such as competitions. Although some teachers might state that there is no difference between them, at different moments in the interviews, they often contradict themselves. For example, teacher 4 starts the interview by claiming that there are no differences between Han and Mongolian students and teachers. However, later on in the interview, when asked to compare them, he notes that, although Mongolian students are very much interested in STEM, their inventions are not new as such and that their lack of interest in innovations from other Minzu groups creates a gap in their STEM knowledge and creativity. About Mongolian teachers—as a reminder, the interviewee is himself a Mongolian Minzu—he argues that they tend to be too proud of “their own Minzu things” and that they “should pay more attention to other Minzu groups”. This argument was uttered by two other Mongolian teachers in the data, drawing systematically the same conclusion that one should pay attention to this issue in the future. To finish on the topic of comparing Han and Mongolian Minzu group members, we note that teacher 5 is the only one to divide Mongolian students into subgroups, on the basis of geography: “For example, we say that our children in the west are not working very hard. It may be because their families are very wealthy. Children in the northeast side are more hardworking.”

Finally, about training, a Han teacher working in a Mongolian school (teacher 7) describes how she was trained when she moved to Inner Mongolia:

Excerpt 2

When we came to Inner Mongolia, we took a special training, where they gave us some precautions. In fact, I think that in this school, many Mongolians have become more sinicized, and many living habits are not much different from those of the Han people. I think we should pay attention to it a little bit. But there are not as many taboos as some ethnic groups like the Hui, so it’s okay to be in a school with Mongolians.

This is the only teacher who mentions having received this kind of specific training, where she was given “some precautions” about working with Mongolians. Although she does not give much information about the content, at the end of the excerpt the keyword of “taboos” seems to indicate that this “special training” was about “do’s and don’ts” and cultural aspects of the Mongolian Minzu—see the use of “living habits” in the excerpt. The argument that Mongolians have become more “sinicized”—meaning more “similar” to the Han majority Minzu—is recurrent in some research on Chinese Minzu (e.g., Yang, 2017), however, what *sinicized* means considering the diversity of the Han Minzu too, is not clear-cut (on a critical review of the idea of “sinicization” see Cheng,

2021; on the Han “China’s diverse majority”, see [Joniak-Luthi, 2015](#)). Here again, the training did not seem to be concerned about including Minzu knowledges but more about beliefs and (maybe) traditions ([Bruchac, 2014](#)).

Before looking more specifically into Minzu knowledges that might be introduced into STEM education in the schools under review, let us say a few words about language issues. As said earlier, some teachers work in Mongolian-language schools with some Chinese used for specific courses and others in Chinese-medium school, with Mongolian taught as a subject. In the Chinese-medium school, where both Han and Mongolian teachers work, the majority language is Chinese and teachers have a duty to use the language in school (except in Mongolian language lessons). Teacher 2 explains:

Excerpt 3

If I teach in Mongolian in class, I would violate the education law, but it can be done after class. If some Mongolian students ask me questions after class, I will interact with them in Mongolian.

Both the Han and Mongolian teachers who work in the schools under review agree that Mongolian teachers are usually closer to Mongolian students. Teacher 7 maintains that “students will have a sense of identity in terms of culture and language, and this sense of identity may make them interested in this teacher and thus in this course.” In the interviews, the teachers who work at Mongolian schools all agree that the students should be taught and spoken to in Mongolian since some of them, especially from the pastoral areas, are not proficient enough in Chinese. However, some teachers argue that STEM should occur in Chinese because “some knowledge cannot be understood by students in Chinese” (Teacher 1). Teacher 2 is the only one who believes that Mongolian and Chinese should be both used in teaching, as “a complement to each other”. He gives the example of scientific and technological words which do not exist in Mongolian such as *tractor*, *airplane* or even the verb *to dissolve*, maintaining that the use of Chinese and/or English can support learning.

The following paragraphs focus on (Mongolian) Minzu knowledge in STEM education. When asked directly to mention how Minzu knowledges are included in the lessons, all teachers except one, provide some examples. Teacher 8, a Han Minzu science teacher working in a Chinese-medium school, explains:

Excerpt 4

If the topic at hand is close to Mongolian culture, I will adapt different methods, but the science subject is more theoretical and rational. Science just has one truth, so we don’t have any special ways, there is no difference between Minzu groups.

The argument of science being “neutral”, and thus that no other knowledge is required in class, has been noted in most research on e.g., culturally responsive STEM education around the world (e.g., [Ukpokodu, 2011](#): 48).

Like the other teachers, Teacher 7 seems eager to include Minzu aspects in their teaching but admits that their lack of knowledge prevents them from doing so. The only concrete example that they put forward is a reference to the largest open-pit coal mine in Inner Mongolia—which they admit contains no “Minzu aspect”. At the end of the interview, the same teacher replies to a final question about recommendations they would make for Minzu STEM education:

Excerpt 5

First, we have to learn Mongolian culture, whether it is language or some traditional customs. If we have such a traditional cultural heritage, we will have better communication with students, and they will have a sense of identity. A sense of identity may make the relationship between teachers and students more cordial and make them identify more with the teacher; second, in a class where Mongolian and Han students teach together, it’s best not to individually propose what should be done to Mongolian students. Let everyone blend together and don’t deliberately make this distinction. Because if there are very few Mongolian students in a class, it may make them bear some different perspectives, so I don’t think it is necessary to deliberately emphasize these differences.

Teacher 7 starts by insisting on the importance for a teacher to know the Mongolian language and “traditional cultural heritage”—without giving details, except the generic “traditional customs”—to stimulate Mongolian Minzu students and to create “bonds” with them. However, when she mentions both Han and Mongolian students, in order not to create some sort of disadvantage, she suggests treating them the same way (“let everyone blend together”) and thus put aside e.g., Mongolian culture and language (see similar results in [Wang, 2021](#)).

In what follows, based on [Bruchac’s \(2014\)](#) categories of *knowledges*, *beliefs* and *traditions*, we introduce some of the examples presented by the interviewees. Reading through what they say and the entire interviews (which focused on the very topic of including Minzu knowledges in STEM education), one notices that these could often appear to be “add-ons”, “improvised” and on the verge of stereotyping at times, and that there is no systematicity in their inclusion. Besides, we note that “dividing” the examples into the three categories was not always easy as some examples seemed to fit into different categories.

We start with knowledges, which is understood here as facts, theories but also consciousness and awareness derived from the Mongolian Minzu group. Three teachers provided the following examples:

Knowledge relating to living experiences (pastoral areas)

For example, for measurement in mathematics, especially for the children living in pastoral areas, the issue of distance, they can understand more easily if I mention that a few kilometres equal to a certain number of meters. Because in their real lives, they often have to do something a few kilometres away, they have already grasped this in practice (teacher 4).

Knowledge about nutrition (Huoliao porridge)

In terms of food, have you ever eaten Huoliao food (Mongolian porridge)? The ingredients are beef or mutton, rice, water and salt. For Mongolian people, this meal is a common home cooked meal. But I say to the students: in fact, relevant experts believe that this meal is the best meal one can get. Then the students ask *why? Isn't this the most common meal we eat every day?* I ask them in return: *why is your body so strong?* Although this is a kind of Mongolian food, its nutritional value is very high. The human body needs sugar, fat, protein, inorganic salts, vitamins and water, which are all included in this special porridge, so it is the best food from the perspective of nutrition (teacher 2).

Knowledge about place and space

For example, when teaching Han Minzu students, most teachers will use Tiananmen Square [in Beijing, opposite the Forbidden City] for discussing distance. When they talk about the size of 100 square meters, teachers will use Tiananmen Square or the school playground as an example. But the Mongolian teachers don't use these examples, they might refer to how big a sheep pen is and how big a cattle pen is. First, I thought in my lessons, as a Chinese, we should all know Tiananmen Square, but children living in pastoral areas have never been there (teacher 6).

For Bruchac (2014) beliefs refer mostly to “religion”. However, here we understand the word to mean principles, values, representations, philosophies, and views. Although examples from the two other categories also seem to fit in this one, we have identified two teachers making specific references to Mongolian beliefs in the examples they give:

Belief about water and ecological frugality

Let me think about it ... Mongolian ... Um ... About the Mongols ... (Long pause) Well, it could be a frugal aspect of Mongolian culture. Because we have water in the fourth chapter of the textbook, I just tell them that this is about frugality, we don't waste water. For Mongolians water is an essential part of their culture (teacher 2).

Belief about animals (the horse)

This is a horse's belly bag, which is put on the saddle. Ordinary teachers would simply perfunctorily explain that this horse's belly bag is made of a horse mane. When I explain to the students, I would say: This is the belly bag of a horse. It was made of a horse mane. I would ask: *Why did ancient people use horse hair to make this?* The students have a lot of ideas on this issue. I will then explain to them that the Mongols are nomadic people, so they have great feelings for horses. If you use leather or other things to make horse harness, it is not good for the horse's health. If you make clothes for the horse with the things that grow on the horse, it will do little or no harm to the horse. This is a manifestation of life skills (teacher 5).

Traditions is the category where most examples were identified in the interviews. They correspond to practices, interests and artifacts of the Mongolian Minzu:

Traditions related to manual work

I'd like to add some new contents in my teaching, such as enhancing handicraft. Because I pay attention to handwork, I found that Minzu students have a strong practical ability. They can make desks, chairs and daily necessities by themselves. They can also create some new things by hand, such as making some things with plenty of pencils. Mongolian tools and hand-made equipment are very developed (teacher 1).

Traditions related to architecture

Recently, we have specially organized training courses for teachers in this area. There is also laser cutting. Our division is more interested in laser cutting, so we developed and designed all the necessary components of the yurt [Mongolian tent] on the computer, and then cut it out to make something like a product. It can be sold, and students can also experience assembling this yurt for free. Yurts and other Mongolian elements are also decorated (teacher 9, Mongolian, engineering, Chinese-medium school).

Traditions related to music

We have a special class, where the students work on and design Mongolian musical instruments such as the matouqin [a two-stringed instrument with a trapezoidal body and handle shaped like a horse head] (teacher 3).

Behaviors

In two interviews, both a Han and a Mongolian teacher (teachers 8 and 3) share their surprise about having witnessed some Mongolian students kowtow to them.

Discussion and conclusion

To our knowledge this is the first attempt to link global questions of indigenous knowledges in education and the specific context of Chinese Minzu. We started by interrogating the words used in Chinese to refer to the (polysemic) notion of indigeneity in comparison to English, and noted that the more direct equivalent word is not used within China to discuss Chinese indigenous diversity. Other words were critically reviewed based on previous research and the notion of Minzu, used politically, in education and research in Mainland China, was introduced as a “companion” to issues of indigeneity in China. Reviewing some studies about how it has been implemented, the problems that both researchers and educators face when working on Minzu (e.g., an overreliance on “Western” multicultural education paradigms), we reflected on what identifying Minzu knowledges in Chinese education, especially in diverse Minzu areas, would mean. We explained that we felt uncomfortable about “judging” e.g., teachers about the kind of Minzu knowledges that they make use of. Using the “simple” model of indigenous knowledges proposed by Bruchac (2014), from the field of archeology, we combed through interview data with different kinds of teachers of STEM from Inner Mongolia.

Our study confirms first and foremost Wang’s (2021) study that there is a lack of formal training for teachers to include Minzu knowledges and that the kind of education and training that teachers seem to receive has to do with “general international knowledge” about STEM or with culturalist information about Mongolian Minzu students. What the analysis of the interviews reveals is that teachers do include some Minzu knowledges in their lessons, as well as aspects of Minzu beliefs and traditions—the latter being the most common one in the examples given by the teachers. This inclusion is clearly not systematic and one can see that most teachers are “juggling” between the key Chinese terms of “unity” (i.e., togetherness) and “diversity” in discussing if and how they include Minzu knowledges. In the interviews, some of the teachers maintain that they do not wish to disadvantage any student by including knowledges and/or multilingual elements while arguing that for some students (e.g., Mongolian students from pastoral areas who do not have proficiency in Chinese) such practices are highly recommended. Looking at the examples again, one can feel that the elements contained in the categories of knowledges and beliefs are potentially the most interesting ones since they relate what students learn to principles, values, representations, views and philosophies derived directly from their Minzu group. These seem to go beyond the “content integration” described by Wang (2021).

The discussions around indigeneity, Minzu, and epistemic diversity presented in this article leave us with (at least) two important sets of questions:

- Inner Mongolia is a complex ensemble of people from several Minzu groups (Han, Mongolian but also Hui, Manchu and Korean) who attend all kinds of schools (e.g., Mongolian and Chinese-medium), often together. Who decides then what Minzu *knowledges, beliefs and traditions* should be included and for whom? If the focus is only on Mongolian Minzu knowledges, is there a risk for both teachers and students to close themselves up in one “knowledgescape”? One could also ask the following questions: Should Minzu knowledges be introduced separately and/or concurrently? Should they be problematized through the concepts of *melange* or *mixing* (Pieterse, 1996) since Chinese Minzu groups have influenced each other throughout the centuries, instead of presenting them as separate blocks? For example, the Mongolian Minzu group has been highly influenced by both Tibetan and Uighur Minzu traditions and beliefs. For teachers this all relates to the important question of how to maintain the central ideology of *unity in diversity* while opening up students’ Minzu “knowledgescape”. Since teachers do not receive training or preparation on these issues, we can only suggest that individual teachers reflect on these questions for themselves and with others if possible. We also believe that these questions are relevant for many other contexts in the world.
- The question of languages is central to the case of Inner Mongolia since the Mongolian Minzu has its own language—like many other Minzu groups. Many examples found in the data refer to a subgroup from the Mongolian Minzu: students from pastoral (grassland) areas. These students are often described as having specific knowledges related to their upbringing and a lack of skills in the common Chinese language, using Mongolian as their main (and often only) language. In previous (international) research on Minzu education, the issue of language use in and outside education is rarely addressed in relation to Minzu knowledges. How could one link them up more systematically? In other words, how could language(s) provide more in-depth access to knowledges, beliefs and traditions? How could e.g., discussions around the translation of specific terms or etymologies deepen our understanding of indigenous knowledges? We also feel that these questions are central to global approaches to indigeneity in education.

References

- Adamson, B., Feng, A. (Eds.), 2021. *Multilingual China: National, Minority and Foreign Languages*. Routledge, London.
- Bruchac, M., 2014. Indigenous knowledge and traditional knowledge. In: Smith, C. (Ed.), *Encyclopedia of Global Archaeology*. Springer, New York, pp. 3814–3824.
- Cajete, G., 2000. *Native Science: Natural Laws of Interdependence*. Clear Light Publishers, Santa Fe, NM.
- Cheng, F., 2021. The evolution of "sinicisation". *J. R. Asiatic Soc.* 31 (2), 321–342.
- Constitution of the People's Republic of China, 1982. <http://english.mofcom.gov.cn/article/lawsdata/chineselaw/200211/20021100050385.html>.
- Dervin, F., 2022. Why "Chinese" stories of interculturality? In: Yuan, M., Dervin, F., Sude, Chen, N. (Eds.), *Change and Exchange in Global Education—Learning With Chinese Stories of Interculturality*. Palgrave Macmillan, London, pp. 1–16.
- Dervin, F., Yuan, M., 2021. *Revitalizing Interculturality in Education: Chinese Minzu as a Companion*. Routledge, London.
- Dervin, F., Yuan, M., 2022. Political ideology and atonality in language and intercultural education: a rejoinder to "between professionalism and political engagement in foreign language teaching practice" by Claire Kramsch. *J. Appl. Ling. Prof. Pract.* 16 (3), 31–45.
- Dikötter, F., 2015. *The Discourse of Race in Modern China*, Fully Revised and Expanded second ed. Oxford University Press, Oxford.
- Fan, Y., Zhao, Y., Liu, A., Hamilton, A., Wang, C., Li, L., Yang, Y., Yang, L., 2018. Indigenous knowledge of dye-yielding plants among Bai communities in Dali, Northwest Yunnan, China. *J. Ethnobiol. Ethnomed.* 14 (1), 1–12.
- Fei, X.T., 1989. *Zhong hu min zu duo yuan yi ti de ge ju* [The format of ethnic diversity in unified Chinese nation]. Central Minzu University Press, Beijing, China.
- Hao, S., 2020. *China's Solution to Its Ethno-national Issues*. Springer, Singapore.
- Hathaway, M.J., 2016. China's indigenous peoples? How Global environmentalism unintentionally smuggled the notion of indigeneity into China. *Humanities* 5, 1–17.
- Jing, Q., Fu, X., 2001. Modern Chinese psychology: its indigenous roots and international influences. *Int. J. Psychol.* 36 (6), 408–418.
- Joniak-Luthi, A., 2015. *The Han: China's Diverse Majority*. University of Washington Press, Washington.
- Kang, H., Wang, H., Fan, Z., Zhao, P., Khan, A., Yin, Z., Wanf, J., Bao, W., Wang, A., Zhang, Q., Liu, J.-F., 2015. Resequencing diverse Chinese indigenous breeds to enrich the map of genomic variations in swine. *Genomics* 106 (5), 206–294.
- Kaviraj, S., 2009. Marxism in translation: critical reflections on Indian political thought. In: Geuss, R., Bourke, R. (Eds.), *Political Judgement: Essays for John Dunn*. Cambridge University Press, Cambridge, pp. 172–200.
- Kong, P.A., Yu, X., 2019. Bilingual education for a harmonious multiculturalism: the importance of policy discourse for students of ethnic minority groups in China. *Multicult. Educ. Rev.* 11 (3), 189–215.
- Li, Y., 2018. Journal for STEM education research—promoting the development of interdisciplinary research in STEM education. *J. STEM Educ. Res.* 1, 1–6.
- Liu, Q., Zehra Colak, F., Agirdag, O., 2020. Characteristics, issues, and future directions in Chinese multicultural education: a review of selected research 2000–2018. *Asia Pac. Educ. Rev.* 21, 279–294.
- Luo, Y., 2017. An alternative to the "indigenous" in early twenty-first-century China: Guizhou's branding of yuanshengtai. *Mod. China* 44 (1), 68–102.
- Menon, D.M., 2022. Changing theory. Thinking concepts from the Global South. In: Menon, D.M. (Ed.), *Changing Theory*. Routledge, London, pp. 1–29.
- National Bureau of Statistics of China, 2021. Main Data of the Seventh National Population Census. http://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817185.html.
- Pasternak, B., 2020. *Cowboys and Cultivators*. Routledge, London.
- Pieterse, J.N., 1996. Globalisation and culture: three paradigms. *Econ. Polit. Wkly.* 31 (23), 1389–1393.
- Piller, I., Zhang, J., Li, J., 2022. Peripheral multilingual scholars confronting epistemic exclusion in global academic knowledge production: a positive case study. *Multilingua*. <https://doi.org/10.1515/multi-2022-0034>.
- Rehamo, A., Harrell, S., 2017. Theory and practice of bilingual education in China: lessons from Liangshan Yi autonomous prefecture. *Int. J. Biling. Educ. Biling.* 23 (10), 1–16.
- Statista, 2022. Population in China in 2020 by province or region. <https://www.statista.com/statistics/279013/population-in-china-by-region/>.
- Su, K., Ren, J., Qin, Y., Hou, Y., Wen, Y., 2020. Efforts of indigenous knowledge in forest and wildlife conservation: a case study on Bulang people in Mangba village in Yunnan Province, China. *Forests* 11 (11), 1–16.
- Sude, Yuan, M., Dervin, F., 2020. *An Introduction to Ethnic Minority Education in China*. Springer, London.
- Teng, X., 2003. *Zuqun, wenhua chayi yu xuexiao kecheng duoyang-hua* [Ethnic groups, cultural differences and school curriculum diversity]. *Jiangshu shehui kexue* 4, 24–29.
- Thomas, S., 2022. Minzu. In: Menon, D.M. (Ed.), *Changing Theory*. Routledge, London, pp. 127–141.
- Tian, D., Yu, H., 2022. On the construction of indigenous Chinese communication theories: an analysis of the cultural roots. *Commun. Theor.* 32 (1), 116–141.
- Ting, R.S.-K., Zhang, K., Huang, Q., 2019. An inclusive indigenous psychology for all Chinese: heeding the mind and spirit of ethnic minorities in China. In: Yeh, K.-H. (Ed.), *Asian Indigenous Psychologies in Global Context*. Palgrave Macmillan, London, pp. 249–276.
- Ukpokodu, O.U., 2011. How do I teach mathematics in a culturally responsive way? Identifying empowering teaching practices. *Multicult. Educ.* 19 (3), 47–56.
- Vinig, T., Bossink, B., 2015. China's indigenous innovation approach: the emergence of Chinese innovation theory? *Technol. Anal. Strat. Manag.* 27 (6), 621–627.
- Wang, L., 2015. The definition of Indigenous Peoples and its applicability in China. *Int. J. Minor. Group Rights* 22 (2), 232–258.
- Wang, W., 2018. Researching education and ethnicity in China: a critical review of the literature between 1990 and 2014. *Front. Educ. China* 13, 216–244.
- Wang, W., 2021. Ethnic minority cultures in Chinese schooling: manifestations, implementation pathways and teachers' practices. *Race Ethn. Educ.* 25 (1), 110–127.
- Wang, D.W., Jiang, L., 2021. *Ethnic Identity of the Kam People in Contemporary China*. Routledge, London.
- Wang, Z., Liu, J., Xu, N., Fan, C., Fan, Y., He, S., Jiao, L., Ma, N., 2019. The role of indigenous knowledge in integrating scientific and indigenous knowledge for community-based disaster risk reduction: a case of Haikou Village in Ningxia, China. *Int. J. Disaster Risk Reduc.* 41.
- Wang, G., Teng, X., Zhou, Z., 2021. Local knowledge instruction for young ethnic minority learners in Yunnan Province. In: Adamson, B., Feng, A. (Eds.), *Multilingual China: National, Minority and Foreign Languages*. Routledge, London, pp. 161–172.
- Worlddata.info, 2022. China. <https://www.worlddata.info/asia/china/index.php>.
- Xu, X., 2022. Epistemic diversity and cross-cultural comparative research: ontology, challenges, and outcomes. *Glob. Soc. Educ.* 20 (1), 36–48.
- Yang, M., 2017. *Learning to Be Tibetan: The Construction of Ethnic Identity at Minzu University of China*. Lexington Books, London.
- Yang, H., Ranjitar, S., Zhai, D., Zhong, M., Goldberg, S.D., Salim, M.A., Wang, Z., Jiang, Y., Xu, J., 2019. Role of traditional ecological knowledge and seasonal calendars in the context of climate change: a case study from China. *Sustainability* 2019 (11), 32–43.
- Yuan, H., 2018. Educating culturally responsive han teachers: case study of a teacher education program in China. *Int. J. Multicult. Educ.* 20 (2), 42–57.
- Yue, Y.-J., Yang, B.-H., Li, Y.J., Zhang, W., Zhang, H.-P., Wang, J.-M., Hong, Q.-H., 2017. Chinese indigenous goat breeds. In: Simões, J., Gutiérrez, C. (Eds.), *Sustainable Goat Production in Adverse Environments: Volume II*. Springer, London, pp. 41–54.
- Zhang, D., 2002. *Night Boat* (Published in Chinese as 夜航船). Sichuan Publishing Group, Chengdu.
- Zhang, D., Chen, L., 2014. Creating a multicultural curriculum in Han-dominant schools: the policy and practice of ethnic solidarity education in China. *Comp. Educ.* 50 (4), 400–416.
- Zhuo, X., 2017. Indigenous traditions of Chinese religions. In: Zhuo, X. (Ed.), *Religious Faith of the Chinese*. Springer, London, pp. 31–122.

'Āina-based education: Huli Ka Lima I Lalo Transformative praxis

Summer Puanani Maunakea, Curriculum Studies Department, University of Hawai'i at Mānoa, Hawai'i, USA

© 2023 Elsevier Ltd. All rights reserved.

Introduction	228
Opening with a story	228
Centering Kanaka 'Ōiwi perspectives of place and decolonizing place-based education	229
Revisiting the term "'Āina-based"	230
Addressing settler traditions of place	230
The need to claim and advance Indigenous peoples' processes of learning and teaching	230
The focus on transformation	231
Mo'olelo as methodology	232
Knowledge sources	232
Weaving a collaborative story	232
Findings	233
The naming of the framework	233
Four projects of 'Āina-based education	234
Huli Ka Lima I lalo	234
Structural issues of access	235
Kuleana	235
'Āina Momona and Ea	235
Ho'omana	236
Summary and discussion	236
Articulating what transformation looks like	236
He Leo Mahalo: closing with a voice of gratitude	237
References	237

Introduction

The story I would like to tell extends conversations on the intricacies of a pedagogy enacted by the agency of Kānaka 'Ōiwi¹ and grounded in 'āina, called 'āina-based education. Physically, 'āina includes the land, earth, and all that nourishes life. Holistically, "The 'āina sustains our identity, continuity, and well-being as a people. It embodies the tangible and intangible values of our culture that have developed and evolved over generations of experiences of our ancestors" (Kikiloi, 2010, p. 75). Kanaka educator and cultural practitioner, Kamoā'e Walk adds that the 'āina not only *sustains* our identity, it *is* our identity, means for continuity, and well-being as a people (Personal Communication, K. Walk, May 2021). As the Indigenous people of Hawai'i, our genealogies teach us that we descend from the primordial forces that birthed the islands we call home, documented in the many mele ko'ihonua (cosmogonic accounts) that inform Kānaka 'Ōiwi in the present day. In the 'ōlelo no'ēau (proverb) he ali'i ka 'āina, he kauwā ke kanaka (the land is a chief, man is a servant) (Pukui, 1983), the order and relationship that humanity and the earth coexist, is revealed .

Opening with a story

The first kalo (*colocasia esculenta*) harvest from my family's garden was very special. On a cool morning in Waipahū on the island of O'ahu, my father and I harvested the first row of kalo we planted in the red soil of Waipahū about eight months prior. We cut off the kalo corm, cut off the lū'au (leaves), and saved the huli (middle of the plant; stalk) to replant. We steamed the corm until the skin started to peel, scraped the skin off until it was clean, and I pounded it with a wooden board and stone until it was 'ūlika (soft and gummy). I carefully added water and mixed until it was the smooth consistency we call poi. My family ate the poi with the cooked lū'au and it physically and spiritually nourished us that day. This was the first time we ate kalo that we planted, cared for, and prepared with our own hands. My family had previously only eaten poi out of a plastic bag purchased from the grocery store. Here in Hawai'i, the price of poi ranges anywhere from five to twelve dollars per pound due to the rising cost of land and the meticulous process of growing and processing—making it inaccessible to many families whose ancestors may have eaten it daily.

¹Kānaka 'Ōiwi, Kānaka (plural), Kanaka (singular), Hawaiian are terms used to describe the indigenous people of Hawai'i. Kanaka and 'Ōiwi are also used as adjectives. When spelled kānaka (plural), this refers to all of humanity.

I share this mo’olelo (story, history, account) to begin this paper because as Kanaka ‘Ōiwi, kalo is my ancestor. But like many Kānaka ‘Ōiwi today, I did not grow up knowing the genealogies of my ancestors due to settler traditions of Western domination, assimilation, and hegemony. It was not until later in my life as young adult that I learned the genealogy of Papa and Wakea—one of many mele ko’ihonua that “identify and connect Kānaka to our kūpuna (ancestors; elders) and our ancestral homelands” (Oliveira, 2014, p. 24). *Malo (1903/1951)* provides this account of Papa and Wakea:

O Wakea noho ia Papa-hanau-moku, Hanau Hawaii, he moku, Hanau Maui, he moku. Hoi hou o Wakea noho ia Hoo-hoku-ka-lani. Hanau Molokai he moku, Hanau o Lanai ka ula, he moku. Lili-opu-punalua o Papa ia	Wakea lived with Papa, begetter of islands Begotten was Hawaii, an island Begotten was Maui, an island. Wakea made a new departure and Lived with Hoo-hoku-ka-lani. Begotten was Molokai, an island, Begotten was red Lanai, an island. The womb of Papa became jealous Hoo-hoku-ka-lani. Hoo-hoku-ka-lani. At its partnership with
Hoi hou o Papa noho ia Wakea. Hanau o Oahu, he moku, Hanau o Kauai, he moku, Hanau o Niihau, he moku, He ula a o Kahoolawe.	Papa returned and lived with Wakea. Begotten was Oahu, an island, Begotten was Kauai, an island, Begotten was Niihau, an island, A red rock was Kahoolawe (p. 243).

Following the creation of the islands, Wākea and Ho’ohōkūkalani’s firstborn child was of premature birth named Hāloa-naka. Hāloa-naka was buried in the earth and from the body grew the first kalo plant. The leaf was named lau-kapa-lili and the stem was known as Hāloa. Soon after, another child was born named Hāloa. Hāloanakalaukapalili is a common name that refers to the first kalo plant and Hāloa is known to be the first Hawaiian chief and primordial ancestor to all Kānaka ‘Ōiwi. While different accounts of Hāloa are told today, we are taught that this mo’olelo is an account of order and place in the universe. Older than humanity is Hāloa, the first Hawaiian chief. Older than Hāloa is Hāloanakalaukapalili, the first kalo plant and staple food of the Hawaiian people. Older than kalo are the islands, the land that Hāloanakalaukapalili grew from, and older than the islands are the elemental forces that created them. This genealogy reminds me of my place in the world. As a young adult, it also reminded me and my family of a way to connect with ‘āina in pono relationship.

Centering Kanaka ‘Ōiwi perspectives of place and decolonizing place-based education

As trending educational movements such as STEM and sustainability education take “place-based” approaches to address issues of the 21st century, it is critical that Indigenous Peoples center and advance the knowledge systems that have connected people and the land as life for generations. The limitations of place-based education in Indigenous communities are widely researched (Calderón, 2014; Corbett, 2020; Saffery, 2019; Seawright, 2014; Webber, 2017). Corbett (2020) posits that some versions of place-based education “might be described as place-limited education in the sense that they appear to begin and end in tightly bounded space” (p. 280). Extending from this idea, Calderón (2014) explains the importance of centering the relationship between land and settler colonialism:

Such a positioning is important, particularly from an Indigenous perspective, because it makes ideologies and structures of settler colonialism explicit. It demands that in articulating place-based models for forming relationships with place, the legal landscapes (Delaney, 1998) of settler colonialism be delineated and deconstructed (p. 26).

Without this delineation, place-based education may perpetuate settler notions of place, exemplified in the Western concept of “sense of place”, often ignorant of Indigenous processes of relating to land, time, and space (Brookfield, 1973; Fox and Cumberland, 1962; Lilomaiva-Doktor, 2020). Grande (2015) explains, “The centrality of place in the Indigenous thought-world is explicitly conveyed through tradition and language and implicitly through the relationship between human beings and the rest of nature” (p. 240). ‘Āina education (Saffery, 2019), aloha ‘āina-placed education (Oliveira, 2017), and for this paper I utilize the widely referenced term ‘āina-based education, seek to connect learners in appropriate ways to how Kānaka ‘Ōiwi relate to ‘āina through familial, spiritual relationship and kuleana (responsibility, privilege, obligation).

Revisiting the term “'Āina-based”

It is important to note that the term “'āina-based education” is contentious. Oliveira (2017) claims that “'āina-based” merely Hawaiianizes “place-based” in reference to place-based pedagogy, as “place” is simply translated to “'āina.” Instead, she uses the term “aloha 'āina-placed education” and emphasizes that aloha 'āina-placed education is:

... more than simply learning that occurs outside of the classroom. Aloha 'āina-placed education is a recognition that Kanaka are genealogically related to the 'āina and the 'āina is our ancestor. The 'āina is not only the source of our physical nourishment, but it also feeds us spiritually and mentally. Aloha 'āina-placed education reinforces the notion that Kanaka have a birthright to reside on the 'āina and by virtue of this birthright we also have a kuleana (responsibility, burden) to care for and protect the 'āina (Oliveira, 2017, p. 6).

I agree with Oliveira that the term 'āina-based education has become mainstream. Her justification that more emphasis is needed to articulate genealogical relationship and kuleana to 'āina resonates with me. Yet it is unclear where the term “'āina-based” originates or whether or not the term was intended to be a Hawaiianized version of “place-based education.” This uncertainty highlights the need to clarify the presence of Hawaiian terms and concepts in educational frameworks. Claiming the use of the term “'āina-based education” puts Kānaka 'Ōiwi in the place of power to define what 'āina-based education truly requires, especially as academic institutions find value in “place-based” educational approaches and begin to institutionalize 'āina-based pedagogies.

Research studies that detail the nature and impact of 'āina-based education have only emerged in the past two decades, whereas the foundational principles that undergird 'āina-based education are firmly rooted in ancestral Hawaiian epistemologies. The Hawaiian-focused Charter School Ko'olau 'Āina-based Hui defines 'āina-based education as:

Learning experiences that allow learners to create authentic, reciprocal relationships with the land, people, and resources of a specific community. This becomes an integral part of the learner's identity and a catalyst for a personal commitment to the health and well-being of that community. It is situated in learning that honors the past, present and future in order to ensure the sustainability and productivity of all forms of life. More specifically, 'āina-based education embodies the wisdom of the 'ōlelo no'ēau: 'O ka hā o ka 'āina, ke ola o ka po'e (The health of the land is the health of the people) (Kanuiakapono, 2020, para. 3).

'Āina-based learning approaches include dynamic and interdisciplinary processes of learning and teaching that hail from the natural landscapes and oceanscapes of Hawai'i's environment. These processes, which emphasize reciprocal interconnected relationships between Kānaka 'Ōiwi and the 'āina, draw upon place-specific intergenerational knowledge systems, language, and customary practices to frame curricula for all learners (Maunakea et al., 2020, 2022). This research explores *how* and *to what end* 'āina-based learning is being offered across diverse educational settings. I draw from Kānaka 'Ōiwi and Indigenous methodologies (Oliveira and Wright, 2016; Smith, 2021) to weave the voices of Kānaka 'Ōiwi educators, researchers, and cultural practitioners into a collective framework for 'āina-based education—further contributing to Kānaka 'Ōiwi directions for 'āina-based educational models and “place-based” approaches to learning.

Addressing settler traditions of place

'Āina-based education has been in large part operationalized in resistance to settler colonialism (Goodyear-Ka'ōpua, 2013; Vaughan, 2018). Seawright (2014) employs the term “settler traditions of place” to describe modes of domination built on “racialized, anthropocentric, and capitalistic understandings of place” (p. 554) that originate in Western epistemological foundations. Domination manifests itself as settler colonialism, white supremacy, anthropocentrism, heteropatriarchy, and other forms. Seawright (2014) explains, “healthy, sustainable, and reciprocal relationships to the social geography we are emplaced in have been bastardized, corrupted, and perverted by the same systems that let capital, competition, and private property define life” (p. 571). In Hawai'i, the commodification of the natural landscape due to development, industrialized agriculture, militarization, over-tourism, and urbanization has amounted to unforeseen levels of environmental degradation and separation of Kānaka 'Ōiwi from 'āina (Osorio, 2014). Considering Hawai'i's unique political status as a sovereign nation occupied by the United States of America government (Silva, 2004) and the history of American indoctrination on the education system (Kaomea, 2000; Benham and Heck, 1998), a critical gaze into place-based education literature is fitting.

The need to claim and advance Indigenous peoples' processes of learning and teaching

Once a space has been cleared of native culture and attachment, its voice becomes that of the detached scientist and the citizen-of-the-world environmentalist (Beamer, 2014, p. 41).

Kanaka ‘Ōiwi and Indigenous scholars assert the need to claim and advance Indigenous ways of knowing and being in research, language, schooling, and systems to make positive differences in their lives (ho‘omanawanui, 2008; Kana‘iaupuni et al., 2017; Maaka, 2004; Smith, 2021). Culturally sustaining and revitalizing pedagogies such as those that center land-based literacies provide counter-narratives to settler legacies in education (McCarty and Lee, 2014; Saffery, 2016). These learning approaches have been described as land education (Tuck et al., 2014); Indigenous land-based education (Wildcat et al., 2014); and Indigenous place-based education (Cajete, 1994; Kawagley and Barnhardt, 1999; Penetito, 2009). Land education as discussed by Tuck et al. (2014) has the power to

Remind people to place Indigenous understandings of land and life at the center of environmental issues and other (educational) issues; provide an explicit critique and rendering of settler colonialism, treaties, and sovereignty; invite and inspire acts of refusal, reclamation, and reimagination; and theorize pathways to living as separate sovereignties on shared territory (Simpson, 2013) (p. 19).

Similarly, Indigenous land-based education operates on the premise that if colonization is about dispossessing Indigenous Peoples from their lands, then decolonization must include “forms of education that reconnect Indigenous peoples to land and the social relations, knowledges and languages that arise from the land” (Wildcat et al., 2014, p. 1). Aligned with the values of land education and Indigenous land-based education, Indigenous place-based education emphasizes the responsibility of humanity in living harmoniously with the natural world (Cajete, 1994; Kawagley and Barnhardt, 1999; Seawright, 2014). Discussions around the questions of who we are, where we are, how we fit in, what the place means to us, and what we mean to the place comprise Indigenous approaches to place-based education (Penetito, 2009). The characteristics of these models possess the potential to “offer epistemic resistance to domination” (p. 554) and push “settler society away from ecological damage and racial and gender supremacy and toward the cultivation of a healthy reciprocal relationship with the natural world” (Seawright, 2014, p. 571).

Bowers (2013) advocates for an earth democracy, “the right of other species to play their role in the world’s diverse ecosystems without being reduced to having only economic value” (p. 228). This concept is fundamental in Kanaka ‘Ōiwi and Indigenous worldviews—all life is sacred and all natural creation has a function (Kanahele, 2005). Eco-justice requires a paradigm shift from the dominant anthropocentric worldview (Hamilton et al., 2015) to an ecocentric worldview, which places value on all animate and inanimate entities regardless of their practicality to human beings (Kortenkamp and Moore, 2001; Leopold, 1986). In this way, ‘āina-based efforts can evolve beyond cultivating relationships to ‘āina toward challenging humanity to shape consumption patterns relative to the limitations and prospects of the local environment (Bowers, 2001; Morishige et al., 2018) and directly challenging settler power structures (Calderón, 2014; Goodyear-Ka‘ōpua, 2013). Along these lines, ‘āina-based education and other land-centered pedagogies work to reclaim and advance the languages and cultures of Indigenous learners while at the same time pursuing Indigenous educational sovereignty (McCarty and Lee, 2014).

The focus on transformation

For this study, I applied the lens of transformative praxis (Smith, 2004) that aligns with Kanaka ‘Ōiwi movements to “rebuild the structures and regenerate the systems that sustain us as ‘Ōiwi” (Peralto, 2018, p. 361). In Graham Smith’s critical analysis of a common linear representation of transformative action, based predominantly on a Western type of thinking, Smith (2004) presents a cyclical model of transformation including the components of conscientization, resistance, and transformative action to illustrate how each element may happen in any order and simultaneously at one time. The use of this transformative praxis model moves beyond identifying and critically analyzing the problems Indigenous Peoples face toward focusing on the process of transformation. Smith articulates that Indigenous educators must understand more accurately, “(a) how and why communities buy in to a transformative process, (b) what counts as transformation, (c) how we know that transformation has actually taken place, and (d) whether or not the transformation is truly beneficial” he continues to explain that the significant need is “to move beyond conscientization to change through transformative praxis” (p. 46). When transformative praxis is applied as a lens for this research, transformative action may occur as a result of approaches that center land-based literacies. The collective action people take to resist further societal and ecological damage may also actualize transformation. Driven by the research and theory around ‘āina-based pedagogies and transformative praxis, with the intent of centering Kanaka ‘Ōiwi epistemologies, this research asks the following questions: *How do we offer education to people on the ‘āina in ways that are grounded in the Indigenous knowledge systems of Hawai‘i?* And, *why implement ‘āina-based education? To what end?* While the use of ‘āina-based approaches to learning and teaching needs no justification, asking the question *to what end* envisions Kanaka ‘Ōiwi futures while explicitly identifying what constitutes transformation.

Mo'olelo as methodology

The design of this study focuses on knowledge sharing (Smith, 2021)—guided by Indigenous research methods that emphasize relationality, agency, and healing (Maunakea, 2016; Smith, 2021; Wilson, 2008). I utilize mo'olelo (Oliveira and Wright, 2016) and collaborative storytelling (Bishop, 1999, 2020) as research methods to build a collective framework for 'āina-based education. Lipe and Lipe (2017) describe mo'olelo as a methodology in this way, “Mo'olelo aku refers to the sharing of mo'olelo with another. Mo'olelo mai refers to receiving a story. Hence, the methodology of mo'olelo aku mo'olelo mai is the sharing—giving and receiving—of stories” (p. 35). Additionally, collaborative storytelling requires the researcher to engage in the research process as a participant—to bring the experiences and stories as someone positioned in the collective to weave “a mutually constructed story created out of the lived experiences of the research participants” (Bishop, 1999, p. 6). I humbly asked three Kanaka 'Ōiwi leaders if they would share their experiences in education, their approaches to 'āina-based teaching, and what they see as future implications of 'āina-based learning in Hawai'i. Rather than having a preset list of formal questions, the conversations were organic in nature, and “driven by the course of the conversation” (Oliveira, 2016, p. 79).

Each kumu has over 30 years of experience as a cultural practitioner in the Hawaiian community. They are all well known and respected as leaders in the field of education because of the educational spaces they have opened for people to learn from and through a grounded and active relationship with 'āina. All talk-story sessions occurred on the island of O'ahu and what follows are brief introductions of each kumu and a description of the setting of each conversation.

Knowledge sources

Kaleo Hanohano is a high school teacher at Kahuku High School. Kaleo has been implementing 'āina-based learning with her students since she became a teacher in the public school system in the 1990s. Kaleo is the creator of the PIKO Rubric, a guide to help teachers in Hawai'i implement place-based learning.² On May 5, 2021, Kaleo and I met online via Zoom. Our conversation lasted 1 h and 30 min.

Kamoa'e Walk is a professor at Brigham Young University where he teaches Hawaiian language and Hawaiian studies courses. Kamoa'e cares for Kahuaola, a one-acre māla'ai (garden, cultivated field) where he, his students, and the community grow a variety of native and Polynesian introduced plants like kalo, mai'a (banana), and niu (coconut). On May 12, 2021, Kamoa'e and I met at Kahuaola to share mo'olelo. Our conversation lasted 3 h.

Herb Lee is the executive director of the Pacific American Foundation of Waikalua Loko I'a, a 400-year-old fishpond in the Ko'olaupoko district. Herb has been responsible for leading the restoration of Waikalua Loko I'a since 1997. He has served as a leader in Hawaiian education, training over 3000 educators in Hawaiian culture-based, place-based, and 'āina-based teaching. On May 18, 2021, Herb and I met at Waikalua Loko I'a. Our conversation lasted 2 h.

Weaving a collaborative story

Traditionally, mo'olelo were handed down orally and the continuity of stories relied upon the listener paying keen attention, making connections, and retelling the story with accuracy (Lipe and Lipe, 2017). Oliveira (2016) explains, “Because we as researchers may learn a great deal by listening to others, especially elders and master practitioners, the sense ability of listening is commonly used in indigenous communities as a methodology” (p. 79). To honor the methodology of mo'olelo, I chose not to utilize a recorder or take notes during talk story sessions for this project. Instead, I relied on my sense ability of listening to receive the mo'olelo and my sense ability of na'au to internalize and make connections with the knowledge being shared with me (Oliveira, 2016). Ultimately, mo'olelo as a methodology in this format requires the earnest responsibility of the researcher to accurately recall the information shared. It also requires the researcher to remain accountable to the participants for how their stories will be internalized and represented.

After each talk story session, I identified and documented the lessons that I learned from each kumu and had informal discussions with each of them to share what I internalized from our conversations. The lessons learned were then woven into the form of a conceptual model. After I completed the first draft of the findings section below, I sent it to each kumu by email to read over and provide feedback. As I continued to internalize their mo'olelo and the lessons I gleaned from all three conversations, the collective story below organically emerged—meant to inform, challenge, ground, encourage, and transform all who engage in 'āina-based education.

²For more information on the PIKO Rubric, visit <https://www.nawaiekolu.org/piko>.

Findings

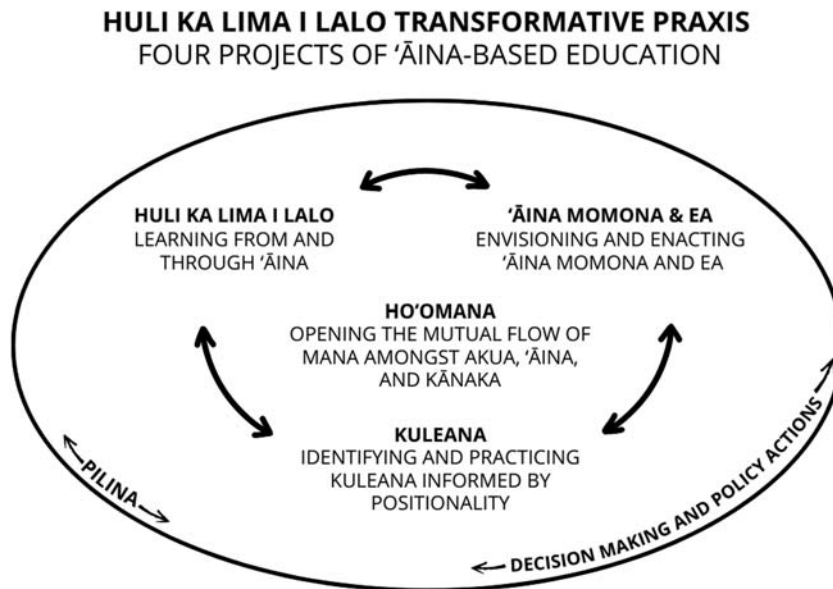


Fig. 1 Huli Ka Lima I Lalo transformative praxis.

The naming of the framework

As a graduate student at the University of Hawai'i at Mānoa (UHM), I was committed to learning about the relationship between 'āina, education, and well-being. I also had a strong desire to continue learning 'ōlelo Hawai'i. My father insisted I learn from one of his classmates, Dr. No'eau Warner, a professor at the UHM Kawaihuelani Center for Hawaiian Language. After a few classes, Kumu No'eau asked me if I had ever listened to Kupuna Katherine Maunakea's Ka Leo Hawai'i recordings. He instructed me to make my way over to Moore Hall on the University of Hawai'i at Mānoa campus and download her interviews. He wanted me to listen to Kupuna Katherine's voice and dialect to get a sense of what a first language speaker of my 'ohana sounded like. He wanted me to pay close attention to how she spoke and the topics she brought in her interviews because there was something in her recordings I needed to hear.

Kupuna Katherine was an educator, composer, lauhala weaver, and an inspiration to the lāhui (people, nation). In the early 1970s, Kupuna Katherine taught Hawaiian language and culture out of her Nānākuli home to many of my aunts, uncles, cousins, and keiki of the community. In one of the recordings, the youths led a sharing of the songs and teachings she taught them. One of her haumāna spoke, "Nē huli ka lima i luna pōloli ka 'ōpū, nē huli ka lima i lalo piha ka 'ōpū. When your hands are turned up you have an empty stomach, when your hands are turned down you have a full stomach."³ At that moment, I knew in my na'au that the agency of turning my hands down into the soil was how I would begin to address the questions I had about my health and the well-being of my family and the disconnection I felt from the 'āina as a source of nourishment. The action of "huli ka lima i lalo" became my guide to cultivating a strong relationship and kuleana to what 'āina we did have—a few hundred square feet behind my family's home in Waipahū where my father and I, soon after hearing Kupuna Katherine's teachings, planted our first māla kalo (kalo garden).

Each of the kumu I spoke with were familiar with Kupuna Katherine and recalled dearly their time spent with Kupuna Katherine and the legacy she has left us today. Thus, I thought it was fitting to name this framework in her honor. Literally, the saying describes a state of hunger "pōloli ka 'ōpū" when hands are turned up. But if one decides to turn the hands down—to the soil, in the 'āina—fullness, "piha ka 'ōpū", is the outcome. Figuratively, this saying may guide people into the future. If Kānaka 'Ōiwi enact agency, emptiness can be transformed to abundance. Because transformative praxis was used as a lens for this research, I adapted the model to serve as a structure to explore the dynamic functions and purposes of 'āina-based education.

³"Nē huli ka lima i luna pōloli ka 'ōpū, nē huli ka lima i lalo piha ka 'ōpū" is an 'ōlelo no'eau that was taught by Kupuna Katherine Maunakea. Because this 'ōlelo no'eau was audio recorded, the word "nē" has been transcribed to match the pronunciation on the recording. However, the word "nē" is a colloquial variation of the word "ina" and although "nē" is commonly heard in spoken Hawaiian, it is uncommon in written form. This 'ōlelo no'eau can be found on the Ka Leo Hawai'i tape HV24_106A. Ka Leo Hawai'i was a Hawaiian language radio show from 1972 to 1988 with interviews and phone calls from native speakers. For more information about Ka Leo Hawai'i and to access a select number of recordings, see: <http://ulukau.org/kaniaina/>.

Four projects of 'Āina-based education

A framework of four projects that ground 'āina-based education in Hawaiian epistemology surfaced from my conversations with the kumu. I call them projects because they are meant to be meaningfully experienced and not only theorized about. At the same time, the projects are grounded in ancestral principles embedded in Kānaka 'Ōiwi epistemology. Carl Urion in [Stewart-Harawira \(2005\)](#) explains:

The traditional principles of traditional knowledge ... remain fixed and provide the framework within which new experiences and situations are understood and given meaning. As such, these principles are the means by which cultural knowledge becomes remade and given meaning in our time (p. 155).

The principles I unpack below are critical factors in 'āina-based education because they strengthen familial and reciprocal relationships between Kānaka 'Ōiwi and 'āina. By identifying these principles, Kānaka 'Ōiwi are able to frame their cultural practices and teachings within the context of ancestral values while at the same time bring the relevance of place-specific knowledge and practice to the present day.

Similar to Graham Smith's model of transformative praxis, the arrows represent how participation in each project mutually reinforces the others. Herb Lee adds that the arrows signify pilina (relationships) because relationship building is constant, iterative, and flowing. *Huli ka lima i lalo* addresses the research question of how education is offered on the 'āina. Learning from and through 'āina acts as a catalyst for the project of *kuleana* where people become conscious of and practice their responsibilities to whom they are accountable to in the context of their own positionalities. The identification and practice of *kuleana* cultivates 'āina *momona* (productive, healthy, and resilient lands, waters, and people) (Morishige et al., 2018) and the enactment of *ea* (sovereignty). Simultaneously, 'āina *momona* and the enactment of *ea* further influences the approaches used in 'āina-based learning and how people understand and carry out their *kuleana*. At the center is *ho'omana* which must occur during each project and at the same time, it is a practice that guides the learner through each project. *Ho'omana* consists of *pule* (prayer), ceremony, and opening up the mutual flow of *mana* (power) that permeates physical and metaphysical realms.

The pilina label is to recognize that transformation occurs in the context of relationships—relationships among people, within self, with natural and built environments, between Indigenous and settler traditions, etc. The pilina that connect us are innumerable whether we are conscious of them or not—as Tewa teacher and scholar Gregory Cajete puts it, “we are all related” (Cajete, 1994, p. 74). Finally, the function of decision making and policy actions is included in the framework because policies whether they are good or bad affect relationships and participation in all four projects. At the same time, engagement in the four projects transform the nature of relationships, and the way we choose to conduct our relationships may transform and inform policy actions and decision-making processes. In the following sections, the projects are described with lessons learned from each kumu I spoke with.

Huli Ka Lima I lalo

Huli ka lima i lalo is the principle and project of directly immersing oneself in the 'āina—turning one's hands down into the *lepo* (soil), *wai* (freshwater), and *kai* (ocean). The learning environments in which 'āina-based education occurs are cultural *kīpuka* where Hawaiian culture can be regenerated and revitalized in contemporary settings (McGregor, 2006). 'Āina-based education has been observed across diverse learning environments which include, but are not limited to natural ecosystems such as native forests and ocean/reef systems; restorative community food systems such as *lo'i kalo*, *loko i'a*, 'āina *malo'o*: dryland and rain-fed field systems, organic farms, *māla'ai*, community gardens, medicine gardens, food forests, and school gardens; home and community life including *wahi pana*, home gardens, public health centers, and community non-profit organizations; and academic institutions including Hawai'i's private, public, and charter schools; post-secondary institutions such as the University of Hawai'i System; and private colleges and universities (Maunakea, 2021).

Students who engage in 'āina-based learning at Waikalua Loko I'a are taught what Herb calls the 'O IA principle. It is an acronym that guides students in “O” observation of what is happening on the 'āina, “I” interpretation of what is occurring through the lenses and experiences that each learner brings with them to the 'āina, and “A” adaptation—the ability for the learner to take what is learned and adapt the teachings to solve the issues in their communities. A fundamental instructional strategy of 'āina-based education involves *kilo*, the practice of keen place-based observation. Morishige et al. (2018) explains, “Native Hawaiian knowledge systems of *kilo* support multigenerational communities to build a collective and intimate understanding of biocultural landscapes and seascapes” (p. 9). Further, they emphasize the importance of “training our memories” to be able to identify the changes in the environment; the observed changes then inform human behavior.

There are many more processes of learning and teaching from 'āina-based settings to be explored in future studies such as, *lau-lima* (many hands working together to achieve a goal), *ma ka hana ka 'ike* (learning by doing), intergenerational learning, the sharing of *mo'olelo* and *mo'okū'auhau* (genealogy), and the blending of 'Ōiwi and foreign tools to cultivate abundance (Beamer, 2014; Kealiikanakaolehaililani and Giardina, 2016; Lincoln et al., 2018). While the methods of 'āina-based education discussed above inform educators in their pedagogical practice, it is also important to bring to light the challenges faced when trying to learn from and through 'āina.

Structural issues of access

Access to ‘āina-based education is an ongoing challenge. In the Hawai‘i State Department of Education, power is concentrated in one school district that spans all islands. Bus transportation to and from schools and sites that offer ‘āina-based learning is costly and the bell schedule often prevents adequate time for students to learn at ‘āina sites. Adequate professional development opportunities for teachers on how to meaningfully incorporate ‘āina-based experiences into their classroom curriculum is also lacking (Hawai‘i Leadership Forum, 2022). The same issue resonates with Indigenous Peoples globally. In the words of Ktunaxa scholar Sam (2019), “Access to place-based language and culture, which is key to the healthy identity development of any society, is limited, and as a result, cultural continuity is at risk” (p. 10). More power and resources must be available in community settings for inter-generational transfer of knowledge systems to occur, as we see that schooling—by way of its colonial nature—often gets in the way of learning.

While Kamo‘a‘e Walk and I talked about what ‘āina-based learning looks like at Kahuaola, the story he emphasized was the barrier of Kanaka ‘Ōiwi access to ‘āina. Kamo‘a‘e challenged me to consider that while culturally sustaining and place-based approaches to education are important, if Kānaka do not have access to ‘āina to fully implement the culture-based learning they receive in educational settings, then what is the point? Hence, a framework for ‘āina-based education must include aspects of decision making and policy actions that increase Kānaka ‘Ōiwi rematriation of connection to and kuleana for ‘āina. Nevertheless, ‘āina-based learning presents a powerful tool for people to strengthen their relationships with their place and community by learning from and through ‘āina. The strengthening of relationships acts as a catalyst for the next project of kuleana.

Kuleana

What can this generation leave for the next generation? Herb Lee posed this question as a catalyst for me to think about the potential of ‘āina-based education. Herb shared a mo‘olelo specific to the restoration of Waikalua Loko I‘a. His kuleana for the past 25 years has been to restore the structure of the fishpond. The new kuleana the next generation will take up is to restore the fishpond’s ability to feed the community. Kaleo shared a story of her students almost 30 years later about the ‘āina’s ability to catalyze one’s identification and practice of kuleana. In the summer of 2019, hundreds of kia‘i (guard, protector, caretaker) stood in protection of the sacred mountain on Hawai‘i Island, Mauna Kea against proposed development of the Thirty Meter Telescope. I recount Kaleo’s story here:

Nineteen ninety-three was the year our people had our 100-year anniversary of the overthrow of Queen Lili‘uokalani. I took 100 students, we made feather leis and we marched to ‘Iolani Palace. Almost 30 years later on Mauna Kea they ran to me. “Kumu, I’m here. Do you see me? I’m staying on the mauna (mountain), kumu. We’re all here. We’re all on the mauna. Remember when you took us to Wai‘anae and you told us that we had to get into the lo‘i?” I said yes, they’re like, “Yeah, that did it for me. I was done. I was all in, that did it kumu.” You just don’t know the impact the ‘āina can make until the students have an opportunity to tell you.

At the start of her teaching career, Kaleo took her students to a lo‘i which is a traditional agricultural structure used to grow kalo. A lo‘i system utilizes gravity-fed freshwater from mountain streams or springs that flow through human made ‘auwai (ditches, canals). Freshwater flows in and out of fields feeding all the kalo before joining natural flowing streams and exiting out to the ocean. Working in a lo‘i may entail clearing debris out of the ‘auwai so freshwater can freely flow to all the kalo plants, pulling weeds so the maximum amount of nutrients is available to the kalo, and sometimes harvesting or planting.

Corntassel (2008) uses the term sustainable self-determination—a concept that advocates for the teaching of sustainable indigenous food production within educational initiatives as a foundation for self-determination and well-being. According to Goodyear-Ka‘ōpua (2013), ‘āina-based pedagogies work toward sustainable self-determination because they disrupt “the hegemonic notion that only a globalized market economy dominated by transnational corporations can provide for our needs” (p. 128). Kaleo’s students expressed how the experience in the lo‘i, with their hands turned down caring for their ancestor Hāloanakalaukapilili, was a moment of transformation for them. Not only were they cultivating kalo to feed themselves and their community, almost 30 years later, they joined hundreds of kia‘i on Mauna Kea to uphold their kuleana to protect the mountain from further development—asserting ea—their right and kuleana as sovereign Kānaka in their homeland.

‘Āina Momona and Ea

‘Āina momona refers to productive and abundant land, watersheds, and oceans. Momona means fat, fertile, or rich as in soil (Pukui and Elbert, 1986). ‘Āina momona has been articulated as the “long-term goal for biocultural restoration in Hawai‘i that speaks to the productive, healthy, and resilient lands and oceans, including the intimate reciprocal relationships our ancestors had with ‘āina ...” (Morishige et al., 2018, p. 5). Herb told a story about 3 years into the restoration of the fishpond; a teacher at a local high school called him after seeing an article in the newspaper about the restoration project. As a science teacher utilizing conventional schooling approaches, she told him that she was not connecting with her students, most of them Kanaka ‘ōpio (youth) labeled “high risk”. She wanted to bring them to learn and aid in the restoration efforts at the fishpond. The teacher, Herb, and scientists

worked together for the school year and the fishpond became an outdoor classroom. At the end of the school year, students taught community members about the restoration of the fishpond and traditional fish gathering practices. The pride they exhibited and the transformation in the students could not be denied. At the beginning of the school year, they were unsure about their purpose at the fishpond, asking the teacher, why are we here? By the end, they were proud of working together to help restore a 400-year-old fishpond and their abilities to teach others how to “throw net”. In the cyclical process of 'āina-based education, lands and oceans are restored and can continue to provide, resulting in an interdependent dynamic of 'āina momona. At the same time, 'āina-based learning presents an opportunity for Kānaka 'Ōiwi to “reconnect to 'āina in holistic ways that unify the physical, emotional, and spiritual aspects of being 'Ōiwi” (Beamer, 2013, p. 60).

Conversation with Kamo'e revealed to me that 'āina-based education can only take us so far. It is up to the individual to decide then what to do with that experience and knowledge gained. In applying 'āina-based learning in a Western system of education or trying to live as a Kanaka in the Western structure that dominates society today, we fail because of the political and structural barriers in place. In addition, the extent to which we value nohona Hawai'i—Kanaka language, culture, traditions, lifeways, determines the extent to which we choose to live as Kānaka each day. Kamo'e encouraged me to first envision in detail where we want to be as Kānaka and to let that inform what “āina-based” education entails, its processes, and what it is proposed to do.

Ho'omana

Ho'omana is the central principle and project in the framework because the mana inherent in 'āina and spiritual realms can be accessed by Kānaka 'Ōiwi through the deepening of cultural skills and knowledge, ceremony, and exercise of kuleana. Ho'omana represents power, spirituality, reverence, and the ability for Kānaka 'Ōiwi to “open the mutual flow of mana within a relationship” (Goodyear-Ka'ōpua, 2013, p. 208). This inherited and acquired mana can be drawn on to develop innovative socio-cultural structures that improve the well-being of Kānaka 'Ōiwi communities and rebuild community understandings of abundance (Enos, 2015; Maunakea-Forth and Abbott, 2015). For Herb, what started as a fishpond restoration project in 1997, guided him on a spiritual journey that has illuminated how the healing of 'āina simultaneously heals kānaka. Herb learned that 'āina-based education included four steps. First, opening oneself to akua (god, spirit, divine). Second, allowing akua to guide one's life. Third, aloha kekahi i kekahi—expressing aloha to one another and joining together to restore the 'āina. Fourth and finally, aloha 'āina—the love one cultivates for their 'āina in the continuous return, protection, and care for it. Ultimately, if we take care of the 'āina it will take care of us. Transformation occurs when one is grounded in the 'āina and open to the guidance of akua.

In our conversation, Kaleo described the kupuna teaching of ho'i i ka piko (return to the center) as having the power to connect Kānaka to the essence of Hawaiian culture and lead Kānaka into the future. Kānaka believe that people have three piko. *Piko i* at the fontanel connects them to akua and the things of the spiritual realm; *piko o* at the navel area connects them to their ancestors; and *piko a* at pelvic area connects them to the future of their people. The teaching of ho'i i ka piko taught her that her kūpuna are always with her in everything she does. Kaleo emphasized that Kānaka do not need to wait for anyone to lead them into the future, they themselves are who they have been waiting for. When Kānaka enact agency to ho'i i ka piko, they holomua—go forward with the confidence and bravery that they can capture the teachings of their ancestors, live it in the present day, and share it with future generations (Young, 1998).

Summary and discussion

Articulating what transformation looks like

I set out to address the questions of *How do we offer education to people on the 'āina in ways that are grounded in the Indigenous knowledges of Hawai'i? And, why implement 'āina-based education? To what end?* Participation in the four projects is meant to actualize transformation across diverse realms—transformation within self, community, island life, decision making, and global well-being. 'Āina-based education is a means for people to learn through and in relation with 'āina. Huli ka lima i lalo is a powerful tool, but it is only one tool that can take us only so far. Learning through and in relation with 'āina must help people identify and practice their kuleana according to their roles and responsibilities in their communities. The experiences and relationships encountered through the projects of 'āina-based education are purposed to guide decision making on personal and policy levels that ultimately enact 'āina momona and ea. Through the dynamism of the Huli Ka Lima I Lalo Transformative Praxis, settler traditions of anthropocentrism, domination, and capitalism must be identified. In the identification process, we can strategize as one community with diverse strengths potential pathways for dismantling them.

More storytelling may help share lessons learned from individuals and communities involved in 'āina-based programs on how they work toward community healing and agency, become less reliant on the global market economy, and organize to protect their places from further ecological and societal degradation. In this regard, reflecting on the following questions may help amplify Kānaka 'Ōiwi voices and beliefs pertaining to what 'āina-based education truly requires:

- What comprises the goals of Kānaka 'Ōiwi communities? What potential does 'āina-based education have in the actualization of those goals?
- What can we learn from the communities utilizing 'āina-based pedagogies to disrupt settler traditions of place and challenge dominant structures of power? Such as MA'O Organic Farms on the Wai'anāe coast of O'ahu fostering youth leadership, feeding

their community, healing the land through regenerative farming practices, and becoming stewards of more ‘āina on the Wai‘anae coast.

- Without a pedagogical focus on settler traditions of place, is ‘āina-based education at risk of being watered down or appropriated as academic institutions find extractive value in its “place-based” educational approaches?
- What work in ‘āina-based education must be done in our generation for the next generation? And by whom? Who benefits?

We have much to do for the longevity of Hawai‘i and the coming generations of ‘Ōiwi who will continue to actualize ‘āina momona, the “depth and breadth” of our ‘ōlelo makuahine, and ea in its fullness. ‘Āina-based education as a pedagogical tool can serve as a means for these actualizations; but we must be maka‘ala (vigilant, alert). Indigenous leaders remind us of the importance of *reinventing the struggle* and not to get distracting in the “sprinkling” or “cherry-picking” of Hawaiian knowledge, language, and customs which are too often appropriated by settler logic. Educators, scholars, and entities interested in ‘āina-based education must critically reflect on their instructional strategies and outcomes to gauge the appropriateness of using an ‘āina-based approach. For example, taking a college-level class outside to a māla to look at a tree is not ‘āina-based education and should not be designated as such. Reflecting on the following critical question may help avoid appropriation of ‘āina-based pedagogies:

- What does the ‘āina-based learning experience(s) entail? Why is ‘āina-based education being used in the current context?
- What outcomes do you hope to achieve? Do the outcomes advance ‘āina momona and ea? How so? Who benefits from these outcomes?
- What languages and worldviews are being advanced in your ‘āina-based learning approach? What languages and worldviews are being silenced?
- How are people being educated to sustain that which has sustained Hawai‘i prior to foreign influence?
- Are learners able to identify the settler traditions at play that perpetuate dominance and degradation in your setting?

Finally, who gets to know what? Acknowledging the extractive tendencies of settler colonial institutions, when do we engage in ‘āina-based approaches and when must we practice academic refusal? (Grande, 2015). If responses to these questions further institutional or transactional outcomes (settler power) rather than cultivate healthy relations and advance Kanaka ‘Ōiwi goals toward ‘āina momona and ea, then the use of ‘āina-based education must be reconsidered and the outcomes redefined.

He Leo Mahalo: closing with a voice of gratitude

It is people who actualize their kuleana and do the liberatory work they are doing in their communities that we join when we engage in ‘āina-based education. Kānaka who inspire me like Kaleo Hanohano, Herb Lee, and Kamoa‘e Walk. Kukui and Gary Maunakea-Forth, Kauai Sana, and Chelsie Onaga of MA‘O Organic Farms. Puni Jackson and the Ho‘oulu ‘Āina crew, the Wilhelm ‘Ohana, Dani Espiritu, and Zach Pilien of Ho‘okua‘āina. The kumu and haumāna of Hālau Kū Mana. The Kahakalau ‘ohana. Haley Kailiehu and No‘eau Peralto of HuiMAU, Maya Saffery and Kaleo Wong of Kauluakalana. Leina‘ala Bright and Keoki Baclayon of Waimānalo Health Center, Aliah Irvine, Ku‘uleilani Samson, Tiana Te Rongopatahi Mo‘iha, Uncle Eric Enos, Uncle Calvin Hoe, Ikaika Lum, Kekailoa and Makuali‘i Perry to name a few.

In closing, this story is told from a personal lens illuminated by my conversations with Kaleo Hanohano, Kamoa‘e Walk, and Herb Lee. It is specific to this time in Hawai‘i, and where I am at right now in my understanding of the purpose of ‘āina-based education—which will continue to deepen and expand as time goes on. I hope that readers are able to glean lessons from this mo‘olelo and enact their own agency to transform their current state into abundance, relation, and transformation. “Nē huli ka lima i luna, pōloli ka ‘ōpū, nē huli ka lima i lalo, piha ka ‘ōpū”.

References

- Beamer, K., 2013. ‘Ōiwi leadership and ‘āina. In: Osorio, J.K. (Ed.), *I ulu i ka ‘āina = Land*. University of Hawai‘i Press, pp. 55–61.
- Beamer, K., 2014. *No mākou ka mana: Liberating the Nation*. Kamehameha Publishing.
- Benham, M.K.P.A.N., Heck, R.H., 1998. Culture and Educational Policy in Hawai‘i: The Silencing of Native Voices. Lawrence Erlbaum Associates, Inc.
- Bishop, R., 1999. Collaborative Storytelling: Meeting Indigenous Peoples’ Desires for Self-Determination in Research. <http://eric.ed.gov/?id=ED467396>.
- Bishop, R., 2020. Indigenous research methods. In: Ward, M.R.M., Delamont, S. (Eds.), *Handbook of Qualitative Research in Education*. Edward Elgar Publishing, pp. 81–92.
- Bowers, C.A., 2001. *Educating for eco-justice and community*. University of Georgia Press.
- Bowers, C.A., 2013. The role of environmental education in resisting the global forces undermining what remains of indigenous traditions of self sufficiency and mutual support. In: Kulnieks, A., Longboat, D.R., Young, K. (Eds.), *Contemporary Studies in Environmental and Indigenous Pedagogies: A Curricula of Stories and Place*. Sense Publishers, pp. 225–240.
- Brookfield, H., 1973. *The Pacific in Transition: Geographical Perspectives on Adaptation and Change*. Edward Arnold.
- Cajete, G.A., 1994. *Look to the Mountain: An Ecology of Indigenous Education*. Kivaki Press.
- Calderón, D., 2014. Speaking back to manifest destinies: a land education-based approach to critical curriculum inquiry. *Environ. Educ. Res.* 20, 24–36.
- Corbett, M., 2020. Place-based education: a critical appraisal from a rural perspective. In: *Rural Teacher Education*, pp. 279–298.
- Comtassel, J., 2008. Toward sustainable self-determination: rethinking the contemporary Indigenous-rights discourse. *Altern. Glob. Local Polit.* 33 (1), 105–132. <https://journals.sagepub.com/doi/10.1177/030437540803300106>.
- Delaney, D., 1998. *Race, place, and the law, 1836–1948*. University of Texas Press.
- Enos, K., May 2015. Using Ancestral Frameworks to Create Contemporary Abundance. Paper presented at The Indigenous Education Symposium. Honolulu Convention Center, Honolulu, Hawai‘i.
- Fox, J.W., Cumberland, K.B., 1962. *Western Samoa: Land, Life, and Agriculture in Tropical Polynesia*. Whitcombe and Tombs.
- Goodyear-Ka‘ōpua, N., 2013. *The Seeds We Planted: Portraits of a Native Hawaiian Charter School*. University of Minnesota Press.

- Grande, S., 2015. *Red Pedagogy: Native American Social and Political Thought*. Rowman & Littlefield.
- Hamilton, C., Gemenne, F., Bonneuil, C., 2015. The anthropocene and the global environmental crisis: Rethinking modernity in a new epoch. Routledge Publishing.
- Hawai'i Leadership Forum, 2022. 'Āina-Based Education Systems Map: Mapping What Enables and Inhibits 'āina-Based Education in Hawai'i. <https://kumu.io/hlf/aina-based-education>.
- ho'omanawanui, k., 2008. 'Ike 'āina: native Hawaiian culturally based indigenous literacy. *Hūlili* 5, 203–244.
- Kanahele, P.K., 2005. I am this land, and this land is me. *Hūlili* 2, 21–30.
- Kana'iaupuni, S.M., Ledward, B., Malone, N., 2017. Mohala i ka wai: cultural advantage as a framework for indigenous culture-based education and studecomes. *Am. Educ. Res. J.* 54 (Suppl. 1), 311s–334s. https://www.ksbe.edu/assets/pdfs/Mohala_i_ka_wai_Cultural_Advantage.pdf.
- Kanuikapono, 2020. Kanuikapono Curriculum Framework. <https://kanuikapono.org/wp-content/uploads/Website-extract-for-Kanuikapono-Curriculum-Framework-V5-FINAL-1.pdf>.
- Kaomea, J., 2000. A curriculum of aloha? Colonialism and tourism in Hawai'i's elementary textbooks. *Curric. Inq.* 30 (3), 319–344.
- Kawagley, A.O., Barnhardt, R., 1999. Education indigenous to place: Western science meets native reality. In: Smith, G.A., Williams, D.R. (Eds.), *Ecological Education in Action: On Weaving Education, Culture, and the Environment*. State University of New York, pp. 117–140.
- Kealiikanakaoleohailani, K., Giardina, C.P., 2016. Embracing the sacred: an Indigenous framework for tomorrow's sustainability science. *Sustain. Sci.* 11 (1), 57–67.
- Kikiloi, K., 2010. Rebirth of an archipelago: sustaining a Hawaiian cultural identity for people and homeland. *Hūlili* 6, 73–115.
- Kortenkamp, K.V., Moore, C.F., 2001. Ecocentrism and anthropocentrism: moral reasoning about ecological commons dilemmas. *J. Environ. Psychol.* 21 (3), 261–272.
- Leopold, A., 1986. *A Sand County Almanac*, reprint ed. Ballantine Books.
- Lilomaiaava-Doktor, S.I., 2020. Oral traditions, cultural significance of storytelling, and Samoan understandings of place or fanua. *Nativ. Am. Indig. Stud.* 7 (1), 121–151.
- Lincoln, N., Rossen, J., Vitousek, P., Kahooni, J., Shapiro, D., Kalawe, K., et al., 2018. Restoration of 'āina malo'o on Hawai'i island: expanding biocultural relationships. *Sustainability* 10 (11), 3985.
- Lipe, K., Lipe, D., 2017. Living the consciousness: navigating the academic pathway for our children and communities. *Int. J. Qual. Stud. Educ.* 30 (1), 32–47. <https://doi.org/10.1080/09518398.2016.1243270>.
- Maaka, M.J., 2004. E kua takoto te mānuka tūtahi: indigenous decolonization, self-determination, and education. *Educ. Perspect.* 37 (1), 3–13.
- Malo, D., 1951. *Hawaiian Antiquities (Mo'olelo Hawai'i)*, second ed. (N. B. Emerson, Trans.). Bishop Museum. (Original work published 1903).
- Maunakea, S.P., 2016. Arriving at an 'āina aloha research framework: what is our kuleana as the next generation of 'Ōiwi scholars? In: Oliveira, K.-A.R.K.N., Wright, E.K. (Eds.), *Kanaka 'Ōiwi Methodologies: Mo'olelo and Metaphor*. University of Hawai'i Press, pp. 142–159.
- Maunakea, S.P., 2021. Toward a framework for 'āina-based pedagogies: a Hawai'i approach to Indigenous land-based education. *J. High. Educ. Theor. Pract.* 21 (10), 277–285.
- Maunakea, S.P., Espiritu, D., Lum, I., Sana, C.K., Samson, K., 2020. Stories of 'āina-based learning, healing, and transformation: I ola kākou i ka ho'olōkahi. In: Cajete, G. (Ed.), *Sacred Journeys: Personal Visions of Indigenous Transformation*. John Charlton Ltd. Publishing, pp. 144–171.
- Maunakea, S.P., Simanu, C., Suluai-Mahuka, S., 2022. Culturally sustaining web 2.0 technology use in the Pacific: innovations from American Samoa and Hawai'i. *Middle Sch. J.* 53 (1), 16–25. <https://doi.org/10.1080/00940771.2021.1999747>.
- Maunakea-Forth, K., Abbott, M., May 2015. MA'O: A Community Endeavor to Restore Our Food & Education Sovereignty. Paper presented at the Indigenous Education Symposium, Honolulu Convention Center, Honolulu, Hawai'i.
- McCarty, T., Lee, T., 2014. Critical culturally sustaining/revitalizing pedagogy and Indigenous education sovereignty. *Harv. Educ. Rev.* 84 (1), 101–124.
- McGregor, D.P., 2006. *Nā kua'āina: Living Hawaiian Culture*. University of Hawai'i Press.
- Morishige, K., Andrade, P., Pascua, P., Steward, K., Cadiz, E., Kapono, L., Chong, U., 2018. Nā kilo 'āina: visions of biocultural restoration through indigenous relationships between people and place. *Sustainability* 10 (10), 3368.
- Oliveira, K.-A.R.K.N., 2014. *Ancestral places: Understanding Kanaka geographies*. Oregon State University Press.
- Oliveira, K.-A.R.K.N., 2016. Ka wai ola: The life-sustaining water of Kanaka knowledge. In: Oliveira, K.-A.R.K.N., Wright, E.K. (Eds.), *Kanaka 'Ōiwi Methodologies: Mo'olelo and Metaphor*. University of Hawai'i Press, pp. 72–85.
- Oliveira, K.-A.R.K.N., Wright, E.K. (Eds.), 2016. *Kanaka 'Ōiwi Methodologies: Mo'olelo and Metaphor*. University of Hawai'i Press.
- Oliveira, K.-A.R.K.N., 2017. Aloha 'āina-placed ho'omoana 'ōlelo Hawai'i: A path to language revitalization. In: McKinley, E., Smith, L. (Eds.), *Handbook of Indigenous Education* (reference work entry, pp. 1–18. https://link.springer.com/referenceworkentry/10.1007%2F978-981-10-1839-8_15-1.
- Osorio, J.K. (Ed.), 2014. *I ulu i ka 'āina = Land*. University of Hawai'i Press.
- Penitito, W., 2009. Place-based education: catering for curriculum, culture and community. *N. Z. Annu. Rev. Educ.* 18 (2008), 5–29.
- Peralto, L.N., 2018. *Kokolo mai ka mole uaua o 'i: The Resilience & Resurgence of Aloha 'āina in Hāmākua Hikina, Hawai'i* (Unpublished Doctoral Dissertation). University of Hawai'i, Mānoa.
- Pukui, M.K., 1983. *'Ōlelo no'eau: Hawaiian Proverbs & Poetical Sayings*. Bishop Museum Press.
- Pukui, M.K., Elbert, S.H., 1986. *Hawaiian Dictionary: Hawaiian-English, English-Hawaiian, revised and enlarged ed.* University of Hawai'i Press.
- Saffery, M.L.K., 2016. He ala nihinihi ia a hiki i ka mole: a precarious yet worthwhile path to kuleana through Hawaiian place-based education. In: Oliveira, K.-A.R.K.N., Wright, E.K. (Eds.), *Kanaka 'Ōiwi Methodologies: Mo'olelo and Metaphor*. University of Hawai'i Press, pp. 109–135.
- Saffery, M.L.K., 2019. *Mai ka piko a ke mole: Clearing Paths and Inspiring Journeys to Fulfill kuleana Through 'āina Education* (Doctoral dissertation, University of Hawai'i at Mānoa).
- Sam, M., 2019. Contextualizing approaches to Indigenous peoples' experiences of intractable conflict. *N. Engl. J. Publ. Pol.* 31 (1) article 5. <https://scholarworks.umb.edu/nejpp/vol31/iss1/5>.
- Seawright, G., 2014. Settler traditions of place: making explicit the epistemological legacy of white supremacy and settler colonialism for place-based education. *Educ. Stud.* 50 (6), 554–572.
- Silva, N.K., 2004. *Aloha Betrayed*. Duke University Press, Durham & London.
- Simpson, L., 2013. Politics based on justice, diplomacy based on love: What Indigenous diplomatic traditions can teach us. *Briarpatch Magazine Online*. <https://briarpatchmagazine.com/articles/view/politics-based-on-justice-diplomacy-based-on-love>.
- Smith, G.H., 2004. Mai i te Maramatanga, ki te Putanga mai o te Tahirutanga: from conscientization to transformation. *Educ. Perspect.* 37 (1), 46–52.
- Smith, L.T., 2021. *Decolonizing Methodologies: Research and Indigenous Peoples*, third ed. Zed Books.
- Stewart-Harawira, M., 2005. Cultural studies, indigenous knowledge and pedagogies of hope. *Pol. Futures Educ. Internet* 3 (2), 153–163. <https://journals.sagepub.com/doi/pdf/10.2304/pfie.2005.3.2.4>.
- Tuck, E., McKenzie, M., McCoy, K., 2014. Land education: indigenous, post-colonial, and decolonizing perspectives on place and environmental education research. *Environ. Educ. Res.* 20 (1), 1–23. <http://www.tandfonline.com/doi/abs/10.1080/13504622.2013.877708>.
- Vaughan, M.B., 2018. *Kaiāulu: Gathering Tides*. Oregon State University Press.
- Webber, G., 2017. *Intricate Waters: A Critical Literature Review of Place-Based Education* (Doctoral dissertation, University of Saskatchewan).
- Wildcat, M., McDonald, M., Irlbacher-Fox, S., Coulthard, G., 2014. Learning from the land: indigenous land based pedagogy and decolonization. *Decolonization* 3 (3), i–xv.
- Wilson, S., 2008. *Research Is Ceremony: Indigenous Research Methods*. Fernwood Publishing.
- Young, K.G.T., 1998. *Rethinking the Native Hawaiian Past*. Routledge Publishing.

FOURTH EDITION

International Encyclopedia of Education

Administration, Governance
and Leadership

VOLUME
4

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Kenneth Leithwood
Qing Gu



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 4

Administration, Governance, and Leadership

VOLUME 4 EDITORS

Kenneth Leithwood and Qing Gu



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

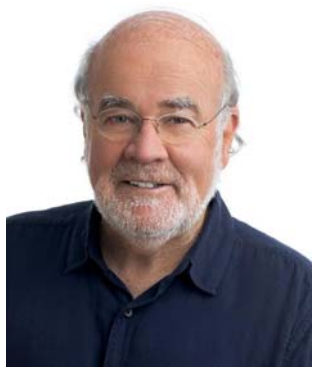
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati Mamoe Ngati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



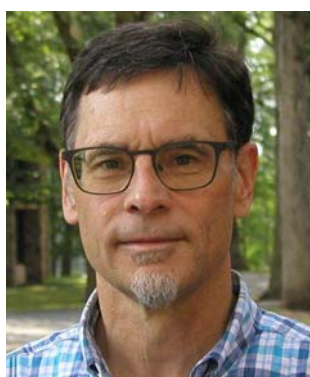
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 4: Administration, Governance, and Leadership

Kenneth Leithwood and Qing Gu

© 2023 Elsevier Ltd. All rights reserved.

Introduction

While the title of this section of the Encyclopedia suggests a balanced attention to three related concepts—administration, governance, and leadership—most chapters in the section are primarily focused on some aspect of leadership. This emphasis reflects a significant trend in the field over the past several decades. It is based on a growing body of evidence indicating that the practices, values, and dispositions of those exercising a positive influence on the work of their colleagues matter more to the success of the students they serve than the structures, policies, and regulations constituting the administrative and governance systems of educational organizations. This is not to say that these systems are unimportant, just that they should aid, rather than restrain or inhibit, leaders' efforts.

This understanding, however, brings with it a substantial—some might say existential—challenge to be addressed. Administrative and governance systems are typically designed to support, to lock into place, and to reinforce widely endorsed organizational behavior; that is, organizational behavior widely endorsed perhaps today although more likely yesterday. But as preferences and expectations for what organizations should accomplish change, the value of those formerly endorsed organizational behaviors diminishes creating the need to redesign existing administrative and governance systems. Such redesign work is a big part of leadership as well, of course, as helping to determine the new preferences and expectations for the organization; these are among the most fundamental responsibilities of leadership. All chapters in this volume of the Encyclopedia have something important to say about one or both of these leadership responsibilities. These chapters also reflect prominent themes in the field's most influential scholarly journals, themes that have been widely studied over some time, but still the focus of considerable ongoing research, as well as themes that have emerged more recently but considered to be quite promising.

As a whole, chapters in this volume reflect significant, closely related changes over about four decades¹ in our understandings about how leaders successfully address the two leadership responsibilities. To illustrate, one such change concerns overall approaches to leadership—from “great man,” “heroic” leadership exercised by individuals to leadership distributed across many members of the organization, as well as other stakeholders. Another closely related change over this period is the erosion of support for “charismatic” approaches to how leaders identify and build commitment to future directions and much greater support for more inclusive, member-engaged direction-setting processes.

This decades-long period also witnessed a shift in our understandings about the nature of “instructional” leadership: from a set of practices aimed at exercising direct influence on teaching and learning, to more indirect conceptions of instructional leadership that take account of the many mediators through which the influence of leadership flows to students. This period has also witnessed a change from understanding “transformational” approaches to “leadership” as an alternative to instructional leadership to transformational leadership as a mediator of effective instructional leadership.

Some of the trends in our understandings about leadership over the past four decades have been driven by changes in our understandings about the nature of learning. Cognitive science has shifted such understandings away from relatively behavioristic conceptions toward more constructivist orientations and an emphasis on learning as a social process. This emphasis on learning as a social process has encouraged leaders' attention to the development of learning cultures, collaborative forms of work (including the formation of professional learning communities, for example), and building into the design of schools' central features of learning organizations. One key intention of such organizational redesign more recently has been to support the evolution of schools' efforts from meeting the needs of “all students” to a more intentional, targeted focus on meeting the needs of historically underserved and diverse students.

Since at least the late 1990s, educational researchers have increased their attention to better understanding how leaders' contexts influence their practices. This attention has included, among others, contexts not well reflected in the mainstream literature on leadership and administration largely dominated by research conducted in Western contexts. This attention has also included how school leaders respond to the political, professional, and accountability challenges created by incoherent, disjointed, and at times contradictory external policy reforms. The political and sociological concerns about the pedagogical challenges, professional tensions, and ethical dilemmas that education reforms have produced for school leaders and teachers are well-documented. However, an increasing body of research knowledge on successful school leadership, policy enactment, and policy coherence offers more promising, more productive and in some ways, more hopeful evidence illustrating how and why some schools have the capacity not only to survive, but also to thrive in the same context of considerable and continuing external policy demands, whilst others do not.

Chapters in this volume of the Encyclopedia are organized around seven themes. The remainder of this introduction briefly summarizes chapters within each theme.

¹The beginning of this period coincides with the “effective schools movement” which significantly increased attention to, and empirical evidence about, school leadership.

Understanding Educational Leadership

There are four chapters in this section. In the first of these chapters, Robinson argues for the distinctiveness of leadership in education as compared with leadership in other organizational sectors. This argument rests on the unique purposes of educational institutions, the consequent role-related obligations of educational leaders, and the largely unique core technologies required to achieve them. Based on a science of learning and teaching, Robinson describes some of those core technologies and the capacities needed by leader to achieve the purposes of the organizations. Myran's chapter reinforces the importance of adopting a science of learning perspective on the meaning of instructional leadership. This chapter argues that reframing instructional leadership around the science of learning is needed in response to challenges facing leaders and schools which defy simple linear, hierarchical, and cause and effect conceptualizations.

Two additional chapters explore the potential of venerable as well as more recent theories to extend or reinforce Robinson's basic argument. Ylimaki highlights tensions faced by today's schools between the culturally diverse groups of students they serve and the common standards for learning proposed for all students by the curriculum policies in many education jurisdictions.

One approach to resolving these tensions, Ylimaki believes, can be found in Dewey's long-standing philosophical approach to education which advocates for individual growth, reflection, and social development (democracy) as a way to hold the needs of individual students at tension with needs for a common societal aim toward democracy. Wang's chapter reviews four categories of theories used over the past 20 years to frame empirical studies about educational leadership recommending that such theories be both "pruned" and refined to better guide future research about educational leadership.

Leadership Practices and Effects

Chapters in this section offer a deeper analysis of the approaches to leadership—and others—briefly described in Wang's earlier chapter. Chapters by Bush, as well as by Sun, Day, and Grice, offer comprehensive accounts of those approaches and the evidence justifying their claims to be successful in furthering the interests of students in schools. Sun reviews empirical evidence about the nature and effects of transformational school leadership while Leithwood describes what is known to date about the nature and effects of "integrated leadership," an approach that combines many of the leadership practices included in both instructional and transformational leadership theories.

Chapters by Spillane and colleagues and by Harris and colleagues describe the mechanisms explaining the positive effects of distributed approaches to leadership while Lumby offers a critique of this approach to leadership and typical claims about its value. Mujis reports unclear and weak effects of both instructional and distributed approaches to leadership on international student test results.

Two chapters are centered on the position of people providing leadership. Based on the distillation of an extensive body of evidence about "middle leaders" (teachers with additional organizational responsibilities), Gurr identifies "five trustworthy statements" about the work of these leaders. Hopkins and Harris describe the position of "system leader," outline its origins, review its current status internationally, and give us a glimpse of future challenges facing those promoting this approach to leadership.

In the final contribution to this section, Branson points to the recent escalation in attention to ethical leadership practices across many organizational sectors. This chapter describes three prominent approaches to ethical educational leadership practice and identifies some of their most serious limitations and how they might be overcome.

Leadership and School Improvement

Themes explored by chapters in this section inevitably overlap themes also discussed in other sections. So, this overview is limited to identifying only those variables or strategies giving rise to school improvement found in these chapters. According to this evidence, school improvement efforts are increasingly successful when schools are encouraged to become research engaged learning organizations (Godfrey, Hubbard et al., Wong et al.) building, for example, on evidence about turnaround school processes (Liu) and making good data central to their decision-making (Datnow and Park). School leaders contribute to the school improvement process when they ensure efficient and well-targeted gathering and assimilation of information on changes occurring both within the school and in the community, complemented by effective means of the knowledge's analysis, storage, and dissemination in an effort to advance the school's interests (Qadach et al.). Decision-making should be widely distributed (Klar, Eckert, Luschei), not just with those in the school but also with the many other groups, organizations, and institutions in its surrounding environment (Hatch et al.). Improvement success is also more likely when the overall improvement process is viewed from a system-wide perspective one which surfaces the structural, cultural, and other dimensions accounting for the complexities of continuous improvement (Slegers, Godfrey, Wong).

School Leadership and Organizational Factors: Pathways to Student Learning

Most chapters in this section acknowledge the indirect nature of leaders' influence on student learning, provide evidence about possible mediating variables, and propose how leaders might influence those mediating variables. These chapters explore evidence

about five such mediating variables—learning climate, trust, alterable family characteristics, collaborative decision-making, and teachers’ sense of internal accountability.

Sebastian and Park’s chapter uses survey and administrative data to compare the effects of a handful of theoretically promising mediators on student learning. Results indicate that, among such mediators, only a school’s learning climate has statistically significant effects on student learning. Learning climate is defined in this research as an environment fostering high academic expectations and a safe and orderly climate. Building on a considerable body of research in many contexts, reported over many years, Tschannen-Moran claims that leader/teacher trust is also strongly associated with student learning. Trusting relationships between and among leaders and their colleagues helps foster commitment to a shared mission and goals. Leaders earn the trust of their colleagues when they are perceived to be benevolent, honest, competent, reliable, and open. This chapter provides guidance to school leaders about how to build trusting relations with their staffs.

Two chapters in this section explore evidence about family characteristics as important mediators of leaders’ influence on students. Jaynes’s chapter, distinguishing between alterable and unalterable family variables, asserts that educational leaders make significant contributions to student success when they: help parents learn how to express their involvement in their children’s school; adopt an “authoritative” (as distinct from “authoritarian”) parental style; demonstrate high but reasonable expectations to their children for their success at school; and acquire the knowledge they need about schooling in order to successfully advocate for their children at school. Also arguing for school leaders to become more engaged in helping parents support their children’s schooling, FitzGerald’s chapter delineates the evolution of parental involvement in a Western context leading to a broader discussion of the work of school leaders with families. This discussion includes an examination of how standards for school leadership both signal the importance of families and communities and are borrowed in international contexts. The “community school” model is described in this chapter as an illustration of a comprehensive place-based reform strategy which locates parental involvement and community engagement at the heart of equity-minded school leadership.

Schools, Ainscow’s chapter argues, have untapped potential to improve their capacity for improving the achievement of all students, particularly those from disadvantaged backgrounds. This chapter points to conditions likely to foster collaboration among all school stakeholders in the interests of using their collective capacities to foster more equitable conditions for these children. School leaders mobilize these collective capacities, the chapter asserts, by nurturing processes of social learning and knowledge creation in settings that stretch from classrooms to the committee structures of senior policymakers. Research suggests that engagement with evidence is an effective way of making this happen.

Finally, Gilbert’s chapter argues that when leaders build a sense of internal professional accountability on the part of staff, they provide staff with a source of professional aspiration, pride, and support for teacher development and their schools. Among other things, this chapter explains how such an internal sense of professional accountability can be built by establishing a collaborative culture, as Ainscow also suggests, by aligning individual responsibility, as well as by developing shared expectations and action. The chapter concludes that public and professional accountability systems should complement each other but that public accountability systems should be judged on how well they support and improve school systems for professional accountability.

The Contexts in Which Leadership Is Enacted

Chapters in this section acknowledge that what educational leaders do is significantly shaped by the circumstances in which they find themselves. There are many such consequential circumstances, for example, the nature of the students and families served by leaders’ schools, both national and organizational cultural norms, educational policies, as well as interpersonal relationships and leaders’ personal values, beliefs, and skills.

Two chapters in this section review evidence about the effects of recent policy initiatives and new expectations about school leaders’ responsibilities on their work intensification (Pollock). Such intensification, as Bauer’s chapter points out, impacts leaders’ job satisfaction, sense of efficacy, as well as feelings of stress and isolation along with knock-on effects related to principal recruitment and retention. Both chapters offer strategies for significantly reducing the negative effects of work intensification. A third chapter by Thessin and Reyes offers an extended account of one such strategy—an approach to school leader supervision by district leaders that build the capacities needed by school leaders to meet the expectation of their districts and address the ambitious policy goals most are accountable for achieving. By conceptualising policy enactment as a leadership and organisational behaviour, the chapter by Gu shows how successful secondary principals in their research have managed to broaden and deepen their organizational, social, and intellectual capacities for the improvement of quality and standards in teaching and learning, *despite* rather than *because of* externally generated reforms.

Finally, Johnson and colleague’s chapter reviews evidence about the contextual factors influencing student outcomes found specifically in Indigenous and other low-socioeconomic contexts. This evidence is used to develop guidelines for empowering Indigenous families and for positively reshaping children’s futures.

Leadership Learning and Development

This section includes five chapters primarily concerned with the nature of, and challenges associated with, formal programs for leadership development. Two additional chapters explore processes entailed in the development of a leader identity and a further two chapters review evidence about the causes and effects of two issues influencing the quality of system-wide leadership.

Chapters in this section are broadly concerned with individual leader development. Several of these chapters focus on formal program characteristics. Chapters by both Bryant and colleagues and Young, rooted in the US context, provide insights into the field's evolving understanding of educational leadership, provide efforts to professionalize the field through the development of a knowledge base and standards, and describe the work needed to build a research-based understanding of quality leadership development. Leithwood's chapter aims to foster improvements in such quality through an approach to pedagogical decision-making better aligned to learning theory; these theories should guide the pedagogical choices made by program developers. Leadership educators are advised to be highly intentional about the capacities to be developed when implementing most of the more than dozen strategies described in the chapter and not to rely on just one or a few of these strategies. Carrington describes a sample of programs and other interventions aimed at helping leaders to develop the capacities they need to foster greater equity in their schools.

Both Earley and Bubb and Sugrue and O'Sullivan advocate for an expansion of the content of formal leadership development opportunities. Earley and Bubb's chapter argues that such opportunities should devote greater attention to individual leaders' personal development, including their well-being, and develop the personal qualities and support required to function successfully in the demanding conditions now facing school leaders. Extending this position, Sugrue and O'Sullivan's chapter claims that most leadership programs do not pay sufficient attention to the intimate relationship between the personal and professional lives of leaders. Ameliorating this neglect would enhance leaders' capacities for pursuing their own values and beliefs "while seeking out legitimate compromises as they grapple with aspects of the policy environment" resulting in more professionally responsible leaders.

While most of these chapters are concerned with the development of school-level leaders, Dobson reports the results of an evaluation of the Lloyds Banking Groups school governance programs concluding that the business sector has a role to play in school governance and understanding professional development is key.

Chapters by Moorosi and Showunmi, as well as Oplatka, describe processes through which leaders come to identify and be identified as leaders. The Moorosi and Showunmi chapter devotes special attention to the role of gender in such leader formation and how the construction of leadership as a male arena influences women's opportunities for advancement and their development as leaders. Oplatka's chapter provides a framework for better understanding how leaders come to view themselves as leaders of social justice.

Chapters by Yan, Cortez Ochoa, and Thomas address three issues that have significant effects on the development of leadership across education systems, as distinct from the quality of individual leaders. Yan's chapter acknowledges the often-negative effects of excessive principal turnover across school systems and reviews the relative effects of three factors potentially accounting for such turnover—leader characteristics, school context (including low student socioeconomic status), and leaders' working conditions. Implications for future research, policy, and practice are derived from these reviews. The Cortez Ochoa and Thomas chapter explores the extent to which the setting of leadership standards and the evaluation of principals guided by such standards is a suitable strategy for encouraging their professional growth and the improvement of schools.

Perspective on Educational Leadership in Non-Western Countries

The five chapters in this section examine a wide range of leadership research interest within Latin America, China, Malaysia, Thailand, and Oman. Latin America is the broad region of interest in the chapter by Flessa and his colleagues. This chapter reports the results of two reviews of published research, the latest conducted between 2016 and 2021. Results of these reviews highlight two issues in leadership research in Latin America, namely that the knowledge base is thin and uneven across countries and there is growing interest in school leadership among researchers and policymakers in who leaders are, what they do and with what impact, and how systems recruit, support, and evaluate them.

Based on individual empirical studies and international comparative reports, Qian and Walker's chapter describes the complexity of educational leadership practices in China. This complexity is viewed as an explanation for some of the discrepancies reported in research between key features of educational leadership in China and Western education systems. Research carried out in Malaysia and described in the Adams and Periasamy chapter was conducted with more than a dozen primary school headmistresses. Complementing and extending prior research on women leadership, the headmistresses in this study stressed the importance of women leaders become visible and vocal in their systems. These female leaders had extensive teaching experience and greater knowledge than most men in their institutions about instructional matters thus equipping them to lead instructional and curriculum changes in their schools.

Kulophas and Hallinger's chapter is based on a state-of-the-art review of research about school leadership in Thailand. Results indicate that successful school principalship in Thailand is defined as general excellence in school management processes, professional recognition, and/or attainment of highly regarded professional rank. This contrasts with Western nations where demonstrated impact on school outcomes is typically identified as the dominant criterion for judging principal success. This difference is explained by Thai culture in which norms of collectivism and power distance give precedence to process over performance. This chapter describes how political and societal factors in the Thai context shape principals' values, practices, and decision-making processes.

Oman is the final non-Western context in this section examining school leadership practices. AL-Mahadi and Mohamed's chapter reports evidence of an emerging interest in "authentic" approaches to leadership in that country. More specifically, evidence

is reported about the relationship between this approach to leadership and teacher turnover intentions mediated by teacher engagement. The chapter offers directions for the improvement of school leadership practices.

Selected Conceptual and Methodological Directions for Future Research

Of the four chapters in this section, two offer conceptual directions and two offer methodological guidance for future research. Louis and Murphy's chapter summarizes research supporting the importance of school leadership behaviors that focus on developing positive relationships with members of the school community. Arguing that relationships are a foundation for student learning and development, the chapter frames the ways in which positive relationships influence the conditions giving rise to such learning and development. School leaders are central to the quality of such relationships and their efforts to build such relationships throughout their school communities are central to the contributions schools make to their students. Relationships are also a central theme in the chapter by Liou and Daly focused on networks. This chapter introduces social network theory and methods as a way to conceptualize educational leadership within interdependent systems and offers several promising directions for future research in educational leadership from a network perspective.

Two chapters in this section make the case for uses, in the future, of more sophisticated methods for conducting research (Heck) and for providing insight into patterns of knowledge production and accumulation. Heck notes that over the past several decades research in the field has become more theoretically informed, but the quality of its research methods provides relatively weak evidence for its knowledge claims. Heck's chapter makes a case for increased uses of longitudinal designs, data collection, and methods of analysis in school leadership and school improvement research in the coming decade. Hallinger's chapter provides an introduction to bibliometric methods for conducting reviews of research. A rationale is provided for this systematic review method, some of the most frequently used analytical techniques are described, and illustrations of results described. The chapter also highlights several criteria that should be considered when evaluating the quality of bibliometric reviews.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 4

Donnie Adams

Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia

Mel Ainscow

University of Glasgow, Glasgow, United Kingdom; and University of Manchester, Manchester, United Kingdom

Asmaa Al-Fadala

World Innovation Summit for Education (WISE), Qatar Foundation, Doha, Qatar

Scott C Bauer

University of Colorado Denver, School of Education & Human Development, Denver, CO, United States

Yalong Bo

University of Hong Kong, Hong Kong, China

Daniela Bramwell

Toronto, Canada

Christopher M Branson

Australian Catholic University, Virginia, QLD, Australia

Darren A Bryant

Department of Educational Policy and Leadership, Asia Pacific Centre for Leadership and Change, The Education University of Hong Kong, Hong Kong SAR, PR China

Sara Bubb

UCL Institute of Education, University College London, London, United Kingdom

Tony Bush

Educational Leadership, University of Nottingham, Nottingham, United Kingdom

Suzanne Carrington

School of Early Childhood and Inclusive Education, Creative Industries, Education and Social Justice, QUT, Kelvin Grove, QLD, Australia

Cynthia E Coburn

Northwestern University, Evanston, IL, United States

Jordan Corson

School of Education, Stockton University, NJ, United States

Artemio Arturo Cortez Ochoa

University of Cambridge, United Kingdom

Alan J Daly

University of California, San Diego, Department of Education Studies, La Jolla, CA, United States

Rima'a Da'as

School of Education, The Hebrew University of Jerusalem, Jerusalem, Israel

Amanda Datnow

University of California San Diego, La Jolla, CA, United States; and Division of Social Sciences, Department of Education Studies, University of California, San Diego, CA, United States

Christopher Day

University of Nottingham, Nottingham, United Kingdom

Neil Dempster

Griffith University, Brisbane, QLD, Australia

Tom Dobson

Leeds Beckett University, Leeds, United Kingdom

Peter Earley

UCL Institute of Education, University College London, London, United Kingdom

Jonathan Eckert

Baylor University, Waco, TX, United States

Idit Finkelstein

Ono Academic College, Kiryat Ono, Israel

Anne Marie FitzGerald

University of Prince Edward Island, Charlottetown, Canada

Joseph Flessa

Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Bev Flückiger
Griffith University, Brisbane, QLD, Australia

Christine Gilbert
UCL Institute of Education, London, United Kingdom

David Godfrey
UCL Institute of Education, London, United Kingdom

Christine Grice
Sydney School of Education and Social Work, The University of Sydney, NSW, Sydney, Australia

Qing Gu
UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

David Gurr
The University of Melbourne, VIC, Australia

Philip Hallinger
College of Management, Mahidol University, Bangkok, Thailand; and Department of Educational Leadership and Management, University of Johannesburg, Johannesburg, South Africa

Alma Harris
Cardiff Metropolitan University, Cardiff, United Kingdom

Christopher W Harrison
Madison Metropolitan School District, Madison, WI, United States

Thomas Hatch
Teachers College, Columbia University, New York, NY, United States

Ronald H Heck
University of Hawaii at Manoa, Honolulu, HI, United States

Yasser F Hendawy Al-Mahdy
College of Education, Sultan Qaboos University, Muscat, Oman

David Hopkins
University of Bolton, Hope Valley, Derbyshire, United Kingdom

Lea Hubbard
Department of Leadership Studies, School of Leadership and Education Sciences, University of San Diego, San Diego, CA, United States

Alice Huguet
RAND Corporation, Santa Monica, CA, United States

Nashwa Ismail

Dong Wook Jeong
Department of Education, Seoul National University, Seoul, South Korea

William H Jeynes
Witherspoon Institute, Princeton, NJ, United States; and California State University, Long Beach, Long Beach, CA, United States

Greer Johnson
Griffith University, Brisbane, QLD, Australia

Michelle Jones

Hans W Klar
Clemson University, Clemson, SC, United States

Jasna Kovačević
Department of Management, School of Economics and Business, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

Dhirapat Kulophas
Department of Educational Policy, Management, and Leadership, Chulalongkorn University, Bangkok, Thailand

Moosung Lee
University of Canberra, Canberra, ACT, Australia

Kenneth Leithwood
University of Toronto, Toronto, ON, Canada; University College London, London, United Kingdom; and University of Nottingham, Nottingham, United Kingdom

Yi-Hwa Liou
National Taipei University of Education, Department of Educational Management, Taipei City, Taiwan

Peng Liu
University of Manitoba, Winnipeg, MB, Canada

Karen Seashore Louis
University of Minnesota-Twin Cities, Minneapolis, MN, United States

Jacky Lumby
Southampton Education School, University of Southampton, Southampton, United Kingdom

Thomas F Lushei
School of Educational Studies, Claremont Graduate University, Claremont, CA, United States

Hugh Mehan
Emeritus of Sociology and Education Studies, University of California, San Diego, CA, United States

Mahmoud Mohamed Emam
College of Education, Sultan Qaboos University, Muscat, Oman

Pontso Moorosi
Department of Education Studies, University of Warwick, Coventry, United Kingdom

Richard Paquin Morel
Research Scientist, San Francisco, CA, United States

Daniel Muijs
School of Social Sciences, Education and Social Work, Queen's University Belfast, Belfast, Northern Ireland

Joseph F Murphy
Vanderbilt University, Nashville, TN, United States

Steve Myran
Old Dominion University, Darden College of Education and Professional Studies, Department of Educational Foundations and Leadership, Norfolk, VA, United States

Izhar Oplatka
School of Education, Tel Aviv University, Tel Aviv, Israel

John O'Sullivan
School of Education, UCD, Dublin, Ireland

Seijoon Park
Educational Leadership and Policy Analysis, University of Missouri, Columbia, MO, United States

Vicki Park
Department of Educational Leadership, San Diego State University, San Diego, CA, United States

Ravadhi Periasamy
Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia

Katina Pollock
University of Western Ontario, Critical Policy, Equity and Leadership Studies, Faculty of Education, Western University, London, ON, Canada

Mowafaq Qadach
Al-Qasemi Academic College of Education, Department of Management and Organization of Educational Institutions and Ono Academic College, Israel

Haiyan Qian
Education University of Hong Kong, Tai Po, Hong Kong

Sandra Quiñones
University of Connecticut, Storrs, CT, United States

Coleen Reyes
The George Washington University, Baltimore Urban Debate League, Baltimore, MD, United States

Viviane Robinson
Distinguished Professor Emeritus, The University of Auckland, Auckland, New Zealand

Anthea Rose
Lincoln Higher Education Research Institute, University of Lincoln, Lincoln, United Kingdom

James Sebastian
Educational Leadership and Policy Analysis, University of Missouri, Columbia, MO, United States

Victoria Showunmi
UCL Institute of Education, University College London, London, United Kingdom

Peter Sleegers
BMC Consultancy, Amersfoort, The Netherlands

James P Spillane
Northwestern University, Evanston, IL, United States

Elizabeth Leisy Stosich
Fordham University, New York, NY, United States

Ciaran Sugrue
School of Education, UCD, Dublin, Ireland

Jingping Sun
Department of Educational Leadership, Policy and Technology Studies, The College of Education, University of Alabama, Tuscaloosa, AL, United States

Rebecca A Thessin
Department of Educational Administration, The George Washington University, Washington, DC, United States

Sally M Thomas
University of Bristol, United Kingdom

Megan Tschannen-Moran
William and Mary School of Education, Williamsburg, VA, United States

Allan Walker
Education University of Hong Kong, Tai Po, Hong Kong; and International Educational Leadership, Department of Educational Policy and Leadership, Asia Pacific Centre for Leadership and Change, The Education University of Hong Kong, Hong Kong SAR, PR China

Yinying Wang
Georgia State University, Atlanta, GA, United States

José Weinstein
Diego Portales University, Santiago, Chile

Lok-Sze Wong
University of North Texas, Denton, TX, United States

Rui Yan

*Office of Applied Research, Evaluation, and Data
Analytics, The City University of New York, New York,
NY, United States*

Rose M Ylimaki

*Education and Leadership, Northern Arizona University,
Flagstaff, AZ, United States*

Michelle D Young

*Loyola Marymount University, Los Angeles, CA,
United States; and University Council for
Educational Administration, Los Angeles, CA,
United States*

CONTENTS OF VOLUME 4

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vi</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 4: Administration, Governance, and Leadership</i>	<i>xix</i>
<i>Contributors to Volume 4</i>	<i>xxv</i>

UNDERSTANDING EDUCATIONAL LEADERSHIP

What is distinctive about educational leadership? <i>Viviane Robinson</i>	1
Re-conceptualizing instructional leadership through a science of learning lens <i>Steve Myran</i>	11
Leadership for education and learning in changing contexts <i>Rose M Ylimaki</i>	24
Future research on educational leadership: theoretical groundings of educational leadership research <i>Yinying Wang</i>	34

LEADERSHIP PRACTICES AND EFFECTS

Leadership practices and why they matter: evolving leadership models <i>Tony Bush</i>	44
Transformational school leadership <i>Jingping Sun</i>	53
Research on successful school leadership <i>Christopher Day, Jingping Sun, and Christine Grice</i>	62
Integrated educational leadership <i>Kenneth Leithwood</i>	73

Framing educational leadership as a <i>multilevel distributed</i> practice: a systemwide perspective <i>James P Spillane, Richard Paquin Morel, and Asmaa Al-Fadala</i>	82
Distributed leadership: taking a retrospective and contemporary view of the evidence base <i>Alma Harris, Nashwa Ismail, and Michelle Jones</i>	91
Distributed leadership: critical perspectives <i>Jacky Lumby</i>	101
Leadership in large-scale international assessments—diminishing returns? <i>Daniel Muijs</i>	109
A review of research on middle leaders in schools <i>David Gurr</i>	115
System leadership – emergence and maturity <i>David Hopkins and Alma Harris</i>	123
The ethical imperative in educational leadership <i>Christopher M Branson</i>	137

LEADERSHIP AND SCHOOL IMPROVEMENT

Schools as research-informed professional learning organizations <i>David Godfrey</i>	146
Understanding turnaround leadership research: continuity and change (2017–2021) <i>Peng Liu and Yalong Bo</i>	158
Evidence-informed decision making in schools and districts: the role of equity <i>Amanda Datnow and Vicki Park</i>	169
Principal’s information-processing mechanisms (PIPMs) as an effective tool for learning-centered leadership that reduces teacher’s intent to leave <i>Mowafaq Qadach, Rima’a Da’as, and Idit Finkelstein</i>	177
Successful leadership in high needs schools <i>Hans W Klar</i>	186
Leading together: developing teacher leadership <i>Jonathan Eckert</i>	196
Organizing schools for improvement: the role of school decision making and authority <i>Thomas F Luschei and Dong Wook Jeong</i>	203
Leadership and governance for improvement: how schools manage external demands <i>Thomas Hatch, Elizabeth Leisy Stosich, and Jordan Corson</i>	209
Leadership and school capacity for continuous improvement: toward a systemic approach <i>Peter Slegers</i>	216
School reform from a constructivist perspective <i>Lea Hubbard, Hugh Mehan, and Amanda Datnow</i>	227
School leaders’ use of research: viewing research use in decision making through an organizational lens <i>Lok-Sze Wong, Alice Huguet, Christopher W Harrison, Cynthia E Coburn, and James P Spillane</i>	235

SCHOOL LEADERSHIP AND ORGANIZATIONAL FACTORS: PATHWAYS TO STUDENT LEARNING

School leadership and organizational factors: evaluating pathways to student learning <i>James Sebastian and Seijoon Park</i>	243
Leadership practices and why they matter: trust in leaders <i>Megan Tschannen-Moran</i>	255
Effects of family educational cultures on student success at school: directions for leadership <i>William H Jeynes</i>	263
Leadership and family factors: critical perspectives on international leadership standards, models, and practices <i>Anne Marie FitzGerald and Sandra Quiñones</i>	272
Collaboration as a strategy for promoting equity in education <i>Mel Ainscow</i>	281
Reclaiming accountability <i>Christine Gilbert</i>	293

THE CONTEXTS IN WHICH LEADERSHIP IS ENACTED

Principals' work in public schools <i>Katina Pollock</i>	304
Leadership contexts and how they matter: leaders' working conditions <i>Scott C Bauer</i>	315
Leading in urban contexts: how principal supervisors coach principals to lead improvement <i>Rebecca A Thessin and Coleen Reyes</i>	325
Changing the context of Australian Indigenous education through leadership for learning "both ways" <i>Greer Johnson, Bev Flückiger, and Neil Dempster</i>	337
Leadership and policy: how principals of successful schools enact education policy for improvement <i>Qing Gu</i>	347

LEADERSHIP LEARNING AND DEVELOPMENT

International patterns and trajectories in principal preparation <i>Darren A Bryant, Allan Walker, and Moosung Lee</i>	356
The development of educational leaders: history and contemporary practice in the United States <i>Michelle D Young</i>	370
Pedagogical strategies for leadership development <i>Kenneth Leithwood</i>	381
Understanding what makes school leaders effective: the importance of personal development and support <i>Peter Earley and Sara Bubb</i>	393
Leadership for equity in schools <i>Suzanne Carrington</i>	401

Leadership formation: in the smithy of the personal and professional <i>Ciaran Sugrue and John O'Sullivan</i>	410
Diverse landscapes of school governing body practices <i>Tom Dobson and Anthea Rose</i>	420
Understanding leadership identity construction: a gendered analysis <i>Pontso Moorosi and Victoria Showunmi</i>	429
The emergence of social justice leaders: identity formation of new principals <i>Izhar Oplatka</i>	438
An overview of literature on factors that influence principal turnover <i>Rui Yan</i>	445
Review of evaluation approaches for school principals <i>Artemio Arturo Cortez Ochoa and Sally M Thomas</i>	453

PERSPECTIVES ON EDUCATIONAL LEADERSHIP IN NON-WESTERN COUNTRIES

School leadership in Latin America: research and policy trends <i>Joseph Flessa, Daniela Bramwell, and José Weinstein</i>	469
Leadership in non-western countries: China <i>Haiyan Qian and Allan Walker</i>	476
Women school leadership in Malaysia: perspectives and enactments <i>Donnie Adams and Ravadhi Periasamy</i>	484
Successful school principalship in Thailand: a literature review <i>Dhirapat Kulophas and Philip Hallinger</i>	490
School leadership practice in Oman: an investigation of authentic leadership, teacher engagement and turnover intention <i>Yasser F Hendawy Al-Mahdy and Mahmoud Mohamed Emam</i>	501

SELECTED CONCEPTUAL AND METHODOLOGICAL DIRECTIONS FOR FUTURE RESEARCH

The contributions of positive organizational studies to educational leadership and school improvement <i>Karen Seashore Louis and Joseph F Murphy</i>	511
Networked systems leadership: the potential of social network theory and analysis <i>Yi-Hwa Liou and Alan J Daly</i>	524
Longitudinal designs for school leadership research <i>Ronald H Heck</i>	536
Applying bibliometric review methods in education: rationale, definitions, analytical techniques, and illustrations <i>Philip Hallinger and Jasna Kovačević</i>	546

What is distinctive about educational leadership?

Viviane Robinson, Distinguished Professor Emeritus, The University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
The proper functions of educational institutions	1
The distinctiveness of the proper functions	3
Deriving the distinctive work of educational leadership	4
How students learn	4
Implications for classroom teaching	5
Implications for the leadership of learning and teaching	6
Leading by setting goals and expectations for deeper learning	6
Leading by resourcing strategically	6
Leading by ensuring quality teaching for deeper learning	7
Leading and participating in teacher professional learning	7
Leading by ensuring a safe and orderly learning environment	8
Concluding remarks	9
References	9

Introduction

The question posed by the title of this article goes to the heart of the theory and practice of educational leadership, for, before one can address the issue of distinctiveness, one must wrestle with the prior question “What is educational leadership?” How one answers that question will shape the degree to which educational leadership is seen as distinctive from leadership in other domains. If, for example, educational leadership is conceived as primarily concerned with forging productive relationships with staff, and the processes for doing so in schools and corporations are seen as very similar, then such similarity weakens the claim for distinctiveness of educational leadership. If, on the other hand, the work of educational leaders is conceptualized as primarily concerned with creating the conditions for the improvement of teaching and learning (Elmore, 2004), then the argument for distinctiveness is strengthened.

A great deal turns on the answer to the question posed in the title of this article. If leadership skills and knowledge are seen as generic, that is, as readily transferable from one type of organization to another (Shamir, 2013), then it makes sense to recruit and appoint leaders from non-educational organizations, waive requirements that leaders be registered teachers, and to include substantial business and management courses in leadership preparation and development programs. If, on the other hand, the skills and knowledge required to lead educational organizations are seen as particular to those organizations, then generic approaches will be found wanting, and more emphasis will be given to such things as how adult relationships are conditioned by the requirement to achieve excellence in learning and teaching (Robinson, 2006).

The position I take is that, although there are commonalities across educational and non-educational leadership, the unique purposes of the former, and the distinctive knowledge and skills required to pursue them, have been given too little emphasis. My argument for distinctiveness rests on the unique purposes of educational institutions, the consequent unique role-related obligations of educational leaders and the unique core technologies required to achieve them.

I begin with a brief review of philosophical writing on the proper functions, or purposes, of educational institutions. I then check the extent to which the three broad purposes derived from this analysis are reflected in the statements of purpose found in the policy and curriculum documents of seven jurisdictions. I argue that at least one of the three broad purposes are unique to educational institutions, and that, since the role obligations of educational leaders are partly derived from institutional purpose, educational leadership itself is distinctive from leadership in other domains. In the remainder of the chapter, I flesh out this somewhat abstract argument with an empirically based account of how the distinctive purposes shape aspects of the role of education leaders. In drawing out implications for the role, I show how successful pursuit of the purpose requires considerable understanding of the core technology—the science of teaching and learning in classrooms.

Throughout this article I treat educational leadership as distributed, in that it can be exercised by anyone regardless of whether or not they hold a formal leadership position (Robinson, 2008). My argument for the distinctiveness of educational leadership has implications for the roles of all educational leaders whether their responsibilities lie at the level of system, school, or team leadership, and whether or not those responsibilities are recognized in a formal leadership position.

The proper functions of educational institutions

Over many decades, philosophers of education have debated the purpose or proper function of educational institutions (Hand, 2014). Such debates, such as the relative importance of liberal and vocational purposes (Winch, 2002), are never finally resolved,

and a thorough discussion of them is beyond the scope of this article. The stance I take with respect to these debates is that educational institutions have several purposes and that a crucial role of leaders is to pursue them in ways that lead to their integration rather than polarization. Such integration requires not only the right philosophical commitments, but knowledge and skill in formulating curricula and pedagogies that enable integration of purposes that are too readily set in opposition to one another. After reviewing philosophical accounts of the purpose of educational institutions, I derived three purposes, which, in order to avoid using the taxonomy preferred by any particular writer, I call preparation, socialization and autonomy.

Purposes related to preparation are focused on the acquisition of valued knowledge and skills which enable children and young persons to lead satisfying and productive lives, including enabling them to make choices about the type of paid employment they desire (MacAllister, 2016; Winch, 2002). The valued knowledge and skills range from various forms of literacy and numeracy, to the specialist skills required for particular trades and professions. The importance of preparation is justified in terms of individual fulfillment (Winch, 2002), the relationship between a strong national economy and a skilled workforce (Biesta, 2009; Winch, 2002), and by the strong expectation of young people themselves that their schooling prepares them for the pursuit of fulfilling paid employment (Winch, 2002).

Purposes related to socialization are concerned with the initiation of students into particular cultures and communities, which may be associated with subject disciplines, professional groups, or political, social, religious, and ethnic traditions (Biesta, 2009). While the pursuit of such purposes is most obvious in, for example, religious schools, socialization is involved in the way in which any knowledge and skills are represented. In teaching math, for example, the subject can be taught in ways that, wittingly or unwittingly, communicate to students that the purpose of their activities is mathematical fluency i.e., to get the right answer as reliably and quickly as possible. Alternatively, math can be taught in ways that induct students into a mathematical community of practice through explicit teaching of mathematical ways of thinking and problem solving. In making these choices, educators are socializing students into very different ways of doing math and being a mathematician.

This second purpose needs little justification as it is an inevitable concomitant of pursuing the first “preparation” purpose. It is the worth of what students are being socialized into and the mode of that socialization that needs explicit debate. For MacIntyre and Dunne (2002), it is critical that students are exposed over extended periods of time to members of different traditions, so they are equipped to engage in the type of debate that characterizes an educated public. The choices educators make about socialization contribute to the formation of student identity and to the historical continuity and renewal of cultures and traditions.

My third set of purposes, which I call autonomy, has attracted the most debate among philosophers of education. For Biesta (2009), it is the counterpoint to socialization, for it involves the freedom to make choices about how to live in the world—choices that might question or even reject the cultures and traditions into which one has been socialized. According to Winch, the pursuit of autonomy in education involves developing three key personal attributes. First, as students grow older, they begin to get a sense of the kind of life they would like to lead, and so the development of autonomy involves exercising choice about how they can achieve the fulfillment they envisage. Second, students need sufficient knowledge, including self-knowledge, to make informed choices. Autonomy, in this educational sense, requires the development of a critical capacity to examine one’s life so that judgment is not surrendered to the will of others or to uncontrolled inner drives. Third, students “need to be equipped with the self-mastery to pursue projects to a successful conclusion in the face of doubts and difficulties” (Winch, 2002, p. 103).

To what extent are each of these three sets of purpose (preparation, socialization, and autonomy) found in the official statements of educational purpose of western democracies? Such statements are powerful because they signal what is valued by governing authorities and, in democratic societies, by those who elect them. In addition, such statements shape the frameworks against which curricula are developed, schools are evaluated, and their leaders are selected and evaluated.

References to preparation are found in the official documents of all seven jurisdictions reviewed (New Zealand, Victoria, California, Ontario, England, Scotland, Norway). While knowledge and skills are important, increasing emphasis is given to the integration of knowledge, skills, and attitudes into key competencies.

A competency is more than just knowledge and skills. It involves the ability to meet complex demands, by drawing on and mobilizing psychosocial resources (including skills and attitudes) in a particular context. For example, the ability to communicate effectively is a competency that may draw on an individual’s knowledge of language, practical IT skills and attitudes towards those with whom he or she is communicating

OECD (2005, p. 4).

The official documents reviewed refer to such competencies as collaboration, creativity, critical thinking, adaptability, taking responsibility and a strong sense of agency—competencies which are thought essential to living a fulfilling and productive life in the complex and uncertain world of the 21st century.

The more specific curriculum objectives that follow from the high level purposes describe the knowledge, skills, attitudes, and values that learners require to understand their world, analyze and solve problems in collaboration with others, transfer their learning to new contexts and manage their own learning. Doing history, for example, no longer involves learning screeds of facts, but understanding the structure of a discipline in which facts are organized into bodies of knowledge and engaging in the forms of historical inquiry that are particular to the discipline (Bransford et al., 2000; OECD, 2019a). Following the Committee on Defining Deeper Learning and 21st Century Skills, I call the learning involved in achieving these curriculum objectives deeper learning and use this term throughout this article. The committee defines deeper learning as “the process through which an individual becomes capable of taking what was learned in one situation and applying it to new situations (i.e., transfer)” (National Research Council,

2012, p. 5). The individual becomes expert in a particular knowledge or skill domain and knows when, why and how to apply that knowledge in solving newly encountered problems within that domain.

Far less emphasis is given in the reviewed documents to socialization and autonomy than to preparation. For example, California, Victoria, and Ontario all refer to acquisition of knowledge and skills needed to contribute to society and for life in the 21st century but make little reference to the role of educational institutions in socialization. In contrast, the English national curriculum document states that the English curriculum introduces students to “the best of what has been thought and said” (Department for Education, 2014, p. 6), and the New Zealand Ministry of Education makes the socialization purpose explicit in its vision that every New Zealander “is strong in their national and cultural identity” (Ministry of Education, 2018).

The Norwegian Ministry of Education and Research goes further by making explicit the values into which students shall be socialized in such statements as

Education and training shall be based on fundamental values in Christian and humanist heritage and traditions, such as respect for human dignity and nature, and on intellectual freedom, charity, forgiveness, equality and solidarity, values that also appear in different religions and beliefs and are rooted in human rights

Ministry of Education and Research (2019, p. 5).

Norway also stood out in its explicit reference to autonomy in such statements as “pupils and apprentices shall develop knowledge skills and attitudes so they can master their lives ...” (Ministry of Education and Research, 2019, p. 5), and “[educators] should stimulate pupils and apprentices/trainees to develop their own learning strategies and critical-thinking abilities” (Utdanningsdirektoratet, n.d., p. 2).

In summary, philosophical writings and official statements both suggest that educational institutions are charged with multiple complex purposes. While it is tempting to assume from these abstract formulations that there is tension between them (for example between socialization and autonomy), the actual degree of tension will depend on how more precise purposes are specified and on how they are pursued. For example, the pursuit of knowledge may be done in ways that indoctrinate, in which case there is considerable tension between preparation and the third autonomy purpose, or there may be little tension, if preparation is pursued through a pedagogy that is educational in the sense of cultivating reason and independent thought.

The distinctiveness of the proper functions

As Biesta (2020, p. 3) writes:

What is special and most likely unique about education is that it is not orientated to *one* purpose ... but actually is orientated to three purposes or, as I prefer to call it, three domains of purpose

Every decision about what is taught and how it is taught has implications for the quality of students’ preparation, socialization, and opportunity to think and act autonomously. For Biesta, the complexity and distinctiveness of education lies in the need to think in this three-dimensional way, because students not only learn from content, and how it is represented, but also from the quality of their interactions with their teachers. That is why he and other philosophers of education are calling for more attention to the values that are espoused and enacted in pursuit of each of the three purposes.

In my view, it is the substance of those purposes, rather than their number, that provides the stronger ground for claiming that the purposes of education are distinctive. While it could be argued that educational, religious, and cultural institutions all share a socialization purpose, only education has preparation as a major purpose. Preparation involves developing the distinctively human qualities and abilities, such as reasoning, wisdom, and understanding that enable people to manage their lives more intelligently and to appreciate the world in which they live (Pring, 2014). Only education is charged with developing students’ competencies by teaching the knowledge, skills and dispositions that are valued by a particular society, and which can, at least at secondary level, only be accessed at school.

Some may object that I have exaggerated the distinctiveness of educational institutions by overlooking the ways in which running a school is like running a business. That similarity is seen in the increasing intrusion of the language of business (e.g., inputs, outputs, targets, performance indicators, and chief executives) into education, and in calls for principals and other senior leaders to be trained in business management (Pring, 2014). Such intrusion is undeniable, particularly in those jurisdictions where decisions about budgets, staffing, property, and health and safety are devolved to the local or school level.

The necessity for some managerial work in any organization, including schools, does not undermine my claim about the distinctiveness of the purposes of education. The proper function of business is, subject to relevant ethical constraints, to produce “goods and services in such a way as to increase owner value ...” (Swanton, 2016, p. 689). The proper function of education is to educate students in ways that achieve the three purposes. The danger is that the management aspect of schools not only diverts leaders’

attention from pursuit of its educational purposes, but that its language and activities so infect the educational activities that they undermine or displace the pursuit of the educational purposes. Such displacement occurs, for example, when in pursuit of exam targets, teachers are pressured to concentrate on those students who are achieving just below the required standard, and to reduce their attention to those who are well below it (Pring, 2014). In such examples, the educational purpose of preparing all students in valued knowledge and skills is displaced by an improper purpose—increasing the school's ranking and reputation. One of the contributions of the philosophers of education discussed earlier, is their call for constant vigilance about the need to ensure that the necessary management aspects of schooling serve rather than displace educational purposes (Pring, 2014; MacIntyre and Dunne, 2002).

Deriving the distinctive work of educational leadership

In the remainder of this article, I discuss the leadership practices required to achieve the purposes of educational institutions and the extent to which they are distinctive. Ideally, this discussion would draw on empirical research on the links between leadership styles, practices, and valid indicators of achievement of the purposes. Unfortunately, this empirical strategy is limited by the fact that research on the impact of leadership on student outcomes has typically focused only on preparation, and the indicators of that purpose which have been employed to date, are far narrower than the key competencies that are so prominent in contemporary curriculum frameworks (Marzano et al., 2005; Robinson et al., 2008; Sun and Leithwood, 2012, 2015; Witziers et al., 2003.) While a program of research and development for the assessment of key competencies is underway (OECD, 2019b), it will be a long time before quantitative research will provide reliable evidence about the relationships between various types of leadership practice and the development of those competencies.

In the interim, I adopt an alternative backward mapping strategy, which starts from educational purpose, then asks how students learn to achieve those purposes, how teachers foster such learning, and how leaders can foster such teaching. The purpose on which I focus is preparation, understood as the acquisition of deeper learning. I use the science of how students learn to understand, problem solve and acquire adaptive expertise, to derive implications for teaching (Bransford et al., 2000). Once effective teaching for deeper learning is explained, I discuss some of its numerous implications for educational leadership and its distinctiveness.

This analytic strategy is very different from much of the discourse on educational leadership and management which takes theories of leader-follower relations as its starting point (Bush et al., 2019; Elmore, 2004). As Spillane (2013, p. 60) claims:

... with some exceptions, many analyses dwell on leading the schoolhouse rather than the core work of the schoolhouse. As a result, descriptions, and prescriptions for leading are only weakly related to the actual work of teaching and leading its improvement.

How students learn

There is now considerable consensus about how students learn (Bransford et al., 2000; Darling Hammond et al., 2019; Deans for Impact, 2015). Whether or not deeper learning occurs, depends, in part, on the type of academic task in which students are mostly engaged, because different tasks evoke and develop different cognitive operations (Doyle, 1983). My focus will be on comprehension tasks because growth in student comprehension and problem solving, is what is required to develop the deeper learning that is a necessary condition for building the competencies.

Comprehension and problem-solving tasks require students to actively construct a cognitive representation of the relevant concepts, events, or situations. As students gain experience in a subject domain, they develop more and better organized representations (schema) of task relevant material. Since schema are stored in long term memory, they free up limited short-term memory for the processing of new information. For students with little relevant prior knowledge—that is with few or no relevant schema—comprehension tasks create considerable cognitive overload as they struggle to bring meaning to apparently disconnected bits of information.

In short, competence in a domain requires a deep foundation of factual knowledge which is organized into a network of schema that aligns with the structure of the discipline or subject. There should be no opposition, therefore, between the learning of facts and deeper learning. The challenge for learners and teachers is to avoid memorizing disconnected facts and focus instead on their structure and interrelationships. This requires deepening learners' knowledge of the subject matter while developing their grasp of its conceptual structure (Bransford et al., 2000).

The development of expertise in a subject domain requires the encoding of valid information (Doyle, 1983). Students can develop schemata which encode faulty algorithms and concepts (Nuthall, 2007). When students have little relevant prior knowledge, or misunderstand what they have previously encountered, they will be unable to engage with tasks which presume accurate prior knowledge.

Deeper learning includes the development of the metacognitive skills required to regulate one's own learning. This involves setting goals, planning ahead, activating relevant prior knowledge, keeping success criteria in mind, and using them to monitor and adjust progress toward the goal (Bransford et al., 2000). Not all learners independently develop the internal talk that

self-regulation requires. Modern curricula require teachers to scaffold metacognitive strategies by explicit teaching and practice so that, over time, students will prompt themselves and monitor their own comprehension and problem solving. Research on metacognition shows that, as for critical thinking and problem solving, metacognitive strategies learned in one context do not readily transfer to other subject matter (Bransford et al., 2000). This means that attempts to teach generic problem solving or self-regulatory strategies are unlikely to be effective because such strategies need to be learned in conjunction with particular subject matter (National Research Council, 2012).

Academic work is conducted in classrooms in which teachers teach, not single students, but one or more groups of increasingly diverse students. Furthermore, students complete academic tasks in contexts where they are either recipients of, or witnesses to, the frequent evaluations of their teacher. For older students, these judgments contribute to a high stakes accountability system of credits or points, which determine report cards, academic pathways and, ultimately, their qualifications. These features of classrooms can create teacher-student interactions that profoundly alter students' opportunities to learn the cognitive processes required for deeper learning (Doyle, 1983). Low achieving students are likely to feel vulnerable in a highly evaluative climate and respond by withdrawal, so as to minimize the risk of public exposure and embarrassment (Peeters et al., 2020). Alternatively, they may attempt to increase their chances of success by using a variety of strategies, such as copying, and requesting detailed guides and model answers, that are designed to reduce the cognitive demands of the task. Teachers who give overly detailed instructions, prompts and advice, can unwittingly turn a comprehension task into a procedural one in which students happily follow their detailed guidance, without increasing their expertise in the subject domain (Doyle, 1986).

In short, academic work, particularly that required to achieve deep learning is complex. As Graesser et al. (2018, 9. 170), conclude:

Studies of the cognitive processes underlying academic work have revealed the enormously complex character of the operations and decisions that academic competence entails, a complexity that is often overlooked when the goals of schools are discussed.

Implications for classroom teaching

There are several excellent texts available which explain the implications for classroom teaching of how students learn (Bransford et al., 2000; Darling Hammond et al., 2019; Deans for Impact, 2015). Here, I concentrate on four implications that have particular relevance to teaching for deeper learning.

First, since what students learn is a function of the type of task they engage in, it is critical that teachers are skilled in the selection, design and assessment of tasks that are tightly aligned to the cognitive processes they intend students to develop (Tekkumru-Kisa et al., 2020). If the intended learning outcome is comprehension, then all task components, including explanations and assessments, should foster that outcome. For example, if a math teacher wants students to understand why an algebra formula works, then setting students 20 practice examples in which they apply the formula will foster procedural accuracy rather than comprehension. Comprehension is more likely to be fostered by assessing whether students can decompose the formula, apply its various steps to concrete materials, and provide a written or oral explanation of why it works.

Second, since deep learning requires the cumulative integration of prior knowledge with new information and concepts, effective teaching requires inquiry into students' prior knowledge. The purpose of such inquiry is not just to interest students in a topic, but to assess whether students have the prior knowledge assumed by the selected task and whether that prior knowledge includes misunderstandings that need to be challenged.

Third, teachers need well developed skills in managing groups of diverse learners while they grapple with intellectually challenging tasks. As already discussed, students attempt to shape tasks in ways that reduce cognitive load and, without high level skills in managing diverse learners, teachers may rely too much on memory and procedural tasks for which they are more likely to gain and retain students' cooperation (Doyle, 1983; Fulmer and Turner, 2014).

Fourth, teaching of the metacognitive skills that are required for the development of critical thinking, independence, adaptability, and other competencies, should be integrated into the subject matter students are learning, rather than presented as separate content-free processes (Bransford et al., 2000). The reason for this advice is that different metacognitive strategies are used in different subjects. For example, the types of question students ask about a word problem in math ("What pattern or principle might be relevant here?") are quite different from those they should ask in their analysis of a historical document ("Who was the intended audience?" "Why was this written in this way?"). Teachers should also be mindful that explicit teaching of metacognitive skills, even when embedded in content, adds cognitive load (Van Gog et al., 2011).

In summary, teaching for deep learning involves careful inquiry into students' relevant prior knowledge, and the provision of multiple opportunities for students to develop rich and well organized domain-specific knowledge. There should be an integration of, rather than opposition between, the learning of facts and higher order thinking, for the:

ability to plan a task, to notice patterns, to generate reasonable arguments and explanations, and to draw analogies to other problems are all more closely intertwined with factual knowledge than was once believed

Bransford et al. (2000, p. 16).

Implications for the leadership of learning and teaching

My discussion of the leadership implications of the science of learning and teaching for deeper learning is organized under five broad dimensions of leadership practice, (setting goals and expectations; resourcing strategically; ensuring quality teaching; leading teacher professional learning and ensuring a safe and orderly environment), which have been employed in empirical investigations of the effects of leadership on student outcomes. Since several reviews of the evidence on these effect effects are already available (Marzano et al., 2005; Robinson et al., 2008; Sun and Leithwood, 2012, 2015; Witziers et al., 2003), my focus will be on some of the leadership practices within each of these dimensions that are critical to supporting teachers in their pursuit of deeper learning, and on the distinctiveness of those practices.

Leading by setting goals and expectations for deeper learning

If the distinctive purposes of education are to be anything more than philosophical or policy abstractions, leaders will need to translate them into specific goals and expectations and incorporate them in relevant vision, mission, and goal statements. This goal setting activity requires a principled understanding of and commitment to the three purposes, and skill in engaging the relevant communities, so that officials, teachers, parents, and students are committed to the goals.

Success in this leadership work is certainly helped by an understanding of generic goal theory, which describes three conditions that need to be met if goal setting in any organization is to be effective. Goals need to be clearly specified, and those who are to achieve the goals must be committed to them, and have, or acquire, the capability needed to achieve the goals (Latham and Locke, 2006). But establishing these three conditions in any particular context requires much more than knowledge of goal theory. In achieving goal clarity, leaders need to clearly articulate and exemplify what they mean by deeper learning and how they believe teaching for deeper learning differs from current practice. They also need to be able to navigate the likely tensions between their own commitments and the goals that their superiors may require them to pursue (Louis and Robinson, 2012).

Teachers may value deeper learning in principle, but still not be committed to it, because their experience convinces them that success with their students is unlikely. It is critical that leaders recognize, that by setting ambitious goals, they are likely to create a gap between their teachers' current and required levels of capability. They should then initiate respectful and non-blaming discussion of the capability issues so that teachers trust that they will get the help they need (Robinson, 2018).

An important distinction is made in goal theory between performance goals and learning goals (Latham and Locke, 2007). When teachers' and leaders' current capability is not sufficient for achievement of a performance goal (e.g., student achievement in math problem solving will improve by 25% in a given year), leaders should set prior learning goals (e.g., all teachers at a particular grade level meet criteria for effective teaching of the vocabulary and strategies relevant to a type of math word problem). Learning goals create an expectation of, and accountability for, teacher and leader acquisition of the knowledge and skills required to achieve the performance goal. If leaders set performance goals when capability is insufficient, they are likely to create stress, lower commitment and invite goal displacement. Pressure to achieve performance goals that are beyond teachers' current capability, combined with high stakes accountability, invites gaming and other forms of unethical behavior (Welsh et al., 2019).

The three principles of goal theory are generic, in that they are applicable to the leadership of any type of organization. However, as I have illustrated, their successful application in education requires leaders who can integrate their knowledge of goal theory with distinctive educational knowledge and skills. For example, leaders need to be able to articulate and defend their choice of goal (deeper learning), evaluate the match between the required and current pedagogies, and provide a compelling account of how any capability gap will be closed. The knowledge and skills involved are distinctive to the work of educational leadership.

Leading by resourcing strategically

Enabling all students to succeed in deeper learning requires leaders to be very strategic in their management of the key resource of students' time. My focus on this resource is not intended to belittle the importance of the more typically discussed management of human, material, and financial resources (Mestry, 2019). Rather, it is to emphasize that students need multiple and varied opportunities to learn challenging material and, for those students whose homes do not provide the cultural and educational capital required to engage with such material, school provides their only opportunity to do so. Rushing through an over-crowded curriculum prevents the pedagogical approaches, such as respectful inquiry into students' emerging understandings and misunderstandings, that are essential to the goal of deeper learning.

One way to understand the use of students' time is via the construct of opportunity to learn (Berliner, 1987). Leaders can monitor students' opportunities to learn by inquiring into the quantity and quality of the opportunities that are planned and provided (Miles and Frank, 2008). Students' opportunities are eroded when time is lost through suspension of regular classes, pull-out programs, extracurricular activities conducted in class time, disciplinary removals from class and other disruptions. The quantity of opportunities is also reduced when teachers do not teach the planned curriculum because they believe that their students are not ready for such learning or because they lack confidence in teaching the scheduled lesson.

More subtle erosions of opportunities for deeper learning occur when there is misalignment between lesson activities and resources and the intended learning outcomes. Even when carefully aligned, students' opportunities are eroded if they lack the prior knowledge to engage with cognitively demanding tasks or they are bored by material they already know. Finally, opportunities to learn are eroded when, despite cognitive engagement, students experience repeated lack of success.

The management of students' time should be informed by rich feedback from students themselves about the pace and intensity of their learning. Such feedback, combined with opportunities to take control of portions of their own timetable, builds the meta-cognitive skills involved in becoming a self-regulated learner, and the autonomy that is one of the three purposes of educational institutions.

The strategic management of students' time so they have the quantity and quality of opportunities needed to succeed on complex tasks is a critical and much neglected aspect of educational leadership. The generic knowledge and skills involved in time-tabling, and the management of materials and money, need to be conditioned by distinctive educational knowledge and skills about how to allocate those resources in ways that ensure deeper learning (Miles and Frank, 2008).

Leading by ensuring quality teaching for deeper learning

Improving the overall quality of teaching for deeper learning, including reducing undesirable variability in such teaching, requires leaders at all levels in a school to be skillful, consistent, and persistent in monitoring and improving curriculum and pedagogy. Improvement should be based on a shared understanding of what teaching for understanding would look like in a given curriculum area and of how current curricula and pedagogy—as planned, delivered, and experienced by students—measure up to that standard. A key question for leaders is “Does the curriculum, as experienced by students, provide multiple opportunities for them to develop well organized bodies of knowledge, rather than discrete facts and isolated skills?”

A good starting point for addressing this question is an analysis of the types of academic task teachers are setting and assessing. The objective of such analysis is not only to assess the extent of focus on higher order skills such as comprehension, problem solving and critical thinking, but, more importantly, to understand the barriers to teachers setting, and students completing, tasks that require such skills. Since it is the assessments that students experience as the real tasks, alignment of intended and assessed learning outcomes is vital (Doyle, 1983; City et al., 2009). Light sampling of planned and completed assessment tasks provides an evidence base from which teachers and their curriculum leaders can discuss the cognitive processes being assessed, their alignment with deeper learning goals and the reasons for any misalignment.

A second focus for an evaluative review could be the depth and breadth of the current curriculum. Depth is needed because students need multiple and varied exposures to concepts, in order to develop schema that encode underlying principles and inter-relationships. Curricula that are too broad prevent learning for understanding because such learning takes more time than rote learning or the application of procedures. Leaders need to directly and collaboratively address the dilemma teachers often experience between fostering the in-depth learning that they espouse and covering the curriculum, for it is this dilemma that often stops them from narrowing and deepening their curriculum focus (OECD, 2019a).

A third focus for an evaluative review is the match between current teaching approaches (pedagogy) and those required to teach for understanding and critical thinking. It is critical that leaders have enough knowledge of and access to evaluative research on pedagogy to challenge the pedagogical fads and fashions that bedevil education and contribute to inequitable student outcomes (Bransford et al., 2000). For example, 50 years of research in cognitive and educational psychology has provided compelling evidence that novices are more likely to learn and transfer their learning if their teachers provide them with fully guided instruction rather than leave them to discover critical information for themselves, as happens in the various forms of discovery, inquiry, experiential and problem-based teaching (Clark et al., 2012). These discovery-based pedagogical strategies have minimal or detrimental effects on learning outcomes, because students lack the prior knowledge needed to make sense of the activities through which they are supposed to discover the intended understandings (Bransford et al., 2000; Van Gog et al., 2011).

A fourth focus for an evaluative review could be the instructional infrastructure that leaders employ in supporting the work of their teachers. The key question leaders should ask about the school's instructional framework, teacher induction handbook, teacher evaluation and classroom observation procedures, protocols for team meetings and sharing of teaching tips and best practice, is whether these tools and routines are fit for the purpose of improving teaching for deeper learning. For example, do classroom observation protocols require information about teacher inquiry in general, or is a distinction made in the protocol between teacher inquiry for an answer, for the right answer or for students' reasoning? It is only the latter type of inquiry that serves as an indicator of teaching for understanding and deeper learning.

The leadership knowledge and skills required to ensure teaching for deeper learning are deeply embedded in and distinctive to educational theory and practice. Because they are more extensive than can be reasonably expected of any one leader, the goal should be to enrich the knowledge and skills that are distributed across, and available to, all leaders of teaching and learning.

Leading and participating in teacher professional learning

By embracing the goal of deeper learning for all students, leaders are likely to create a considerable gap between the current and required capability of some of their teachers. Closing this gap is essential if leaders are to address inequities in students' understanding and problem solving. Doing so requires sustained, targeted, and high-quality opportunities for teacher professional learning. Such opportunities include both formally scheduled and resourced teacher professional development and the more informal on-the-job teacher discussion and feedback that is part of a learning-focused teacher culture.

A recent meta-analysis of the evidence on the effectiveness of formal teacher professional learning suggests that it brings little payoff in terms of improved outcomes for the students of the participating teachers (Filges et al., 2019). A likely explanation for these dismal results is that few professional development interventions are designed in accordance with our knowledge of how people learn. Since the science of learning is the same for students and adults, the applicable principles are those summarized in

the earlier section on student learning. Like their students, teachers need multiple opportunities to learn, transfer and apply the complex ideas and skills that are needed to teach diverse groups of students how to problem-solve in particular subject domains. This means that a formal professional learning program should be tightly focused, rather than the mile-wide and inch-deep provision that is still seen in some schools.

The teacher learning agenda should be highly responsive to the evidence about what students have yet to learn (Elmore, 2004). It should also be informed by diagnostic evidence about the links between what students learned and how they were taught (Bryk et al., 2015). Well-designed teacher professional learning includes respectful inquiry into the knowledge and beliefs that inform teachers' practice, because, just as for students, teachers' prior knowledge will interact with and profoundly affect how they understand and implement any new practices (Robinson, 2018).

Leaders need to be skilled and knowledgeable in explaining why professional learning is not voluntary, why meeting some student learning needs are given priority at this time over others, and why the leaders of each professional learning opportunity are chosen based on task-relevant expertise rather than on years of experience, nomination or formal position (Elmore, 2004). Leaders with deep educational expertise will defend such controversial decisions, by providing educational grounds for them. Leaders who lack such expertise cannot draw on this important source of leadership influence, and must rely instead, on the force of their personality and positional authority (Fay, 1987). When leaders rely on generic influence processes their leadership becomes more managerial and less educational.

Leading by ensuring a safe and orderly learning environment

Viewed from an educational rather than a narrow behavior management perspective, the work of creating a safe and orderly school is about increasing the behavioral, cognitive, and emotional engagement of students by meeting their needs for personal competence, a sense of belonging and control over and success in their learning (Wang and Holcombe, 2010). There is a considerable body of research on how school and classroom structures, routines and norms are perceived by students in ways that differentially impact their engagement. For example, if the preparation purpose of education is pursued through performance goals, the resulting high stakes competitive environment is likely to threaten students' need for personal competence and reduce their cognitive and emotional engagement with school. If, on the other hand, leaders ensure an emphasis on self-improvement through mastery or learning goals, then students are likely to make greater effort, including more use of self-regulation strategies (McInerney, 2011).

Students' need for autonomy and control over their learning is fostered by such things as student involvement in formative assessment processes, meaningful participation in school decision-making and choice about how they participate in school sports, cultural and arts activities. Their sense of belonging is fostered by confidence that their teachers understand and care about them (Bryk et al., 2010).

The academic tasks that are required for deep learning pose particular challenges to maintaining safe and orderly learning environments, because the answers to such tasks cannot be predicted and rehearsed in advance. Students may attempt to reduce their probability of failure by resisting, dissembling, or negotiating a reduction in the difficulty level of the task or its assessment. In an attempt to establish or restore order, teachers may simplify the task or select those which are more familiar and easier for students. It is important that leaders understand that their teachers cannot increase opportunities for deeper learning until they are able to integrate the provision of such opportunities with maintenance of an orderly class. As Doyle (2006, p. 111) writes:

... academic work can be swamped by the management function in teaching, and teachers can become preoccupied with getting work accomplished rather than promoting student achievement.

Since classroom order is a function of the interactions between teachers, students and academic tasks, leaders should help their teachers to diagnose the source of the difficulties before suggesting solutions. Is the cognitive load associated with the task excessive because the students lack the relevant prior knowledge (Center for Education Statistics and Evaluation, 2018)? Has there been too little explicit instruction in the relevant concepts with students expected to discover them on their own (Clark et al., 2012)? Or is the problem one of classroom management (e.g., poor classroom routines or unclear and inconsistent teacher expectations (Doyle, 2006)?

Leaders can help teachers reduce students' feelings of vulnerability and threat by showing them how to reduce cognitive load through explicit teaching of relevant prior knowledge. Similarly, they can help teachers to reduce evaluative threat through non-evaluative inquiry into several students' thinking and using the resulting information to give formative feedback to the group rather than to each individual student. They can model and develop teacher inquiry that assesses the impact of teaching rather than the correctness of the students' answers. Such strategies, if used consistently, will reduce the psychological risk students attribute to engaging with challenging academic tasks and increase their behavioral and cognitive engagement.

In pursuing a safe and orderly environment, leaders' focus should not be on classroom misbehavior per se, for such disruptions are usually symptomatic of teachers' and students' difficulties in establishing and maintaining the routines required to sustain attention on challenging academic work. The focus of leaders should be on inquiring into the source of student disengagement and assisting their teachers to create those classroom norms and routines that foster student engagement and success.

Concluding remarks

In this article I have argued that the preparation purpose of educational institutions is distinctive and that it is this purpose, expressed in contemporary policy and curricula as the promotion of deeper learning along with its associated competencies, that confers the distinctiveness of educational leadership. Given this purpose, I have argued that much of the work of educational leadership should be derived from the science of how students learn in classroom contexts, and of how teachers teach for deeper learning. That knowledge shapes the distinctive features of educational leadership and creates an extensive learning agenda for current and future educational leaders.

While the current emphasis on instructional leadership has moved the field from its reliance on more generic leadership theories, my argument suggests that more attention is needed, within the instructional leadership tradition, to the implications of pursuing ambitious deeper learning goals. Achieving those goals (e.g., high levels of understanding, complex problem-solving, metacognitive and collaborative competencies) is hugely challenging for learners and teachers, and unlikely to be achieved in settings where leaders do not have access to and utilize the science of learning and teaching for understanding.

Some might argue that, since most leaders are experienced teachers, they will have sufficient knowledge of the science of learning and teaching on appointment, and what they need is the more generic leadership skills involved in such things as setting goals, recruiting staff, and financial management. I have two responses. First, the extent to which leaders have the relevant knowledge on appointment is an important empirical question, answered by research on such things as the extent to which they can lead a discussion on the implications of cognitive load theory for pedagogical choice. Second, my argument for the distinctiveness of educational leadership does not imply that it has nothing in common with leadership in non-educational institutions. In my discussion of the five dimensions of leadership, I outlined how generic leadership skills such as goal setting, were important, but needed to be integrated with the particular educational constraints required to achieve the overarching purposes of today's educational institutions. Such integration transforms the decontextualized abstraction of generic leadership into a situated, distinctive, and content-rich process. If my account of educational leadership is accepted, then it opens up rich possibilities for research and development programmes on the knowledge, skills, attitudes and values which will enable educational leaders to succeed in the distinctive and highly challenging work that their roles entail.

References

- Berliner, D.C., 1987. Simple views of effective teaching and a simple theory of classroom instruction. In: Berliner, D., Rosenshine, B. (Eds.), *Talks to Teachers*. Random House, New York, NY, pp. 93–110.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21 (1), 33–46. <https://doi.org/10.1007/s11092-008-9064-9>.
- Biesta, G.J., 2020. Risking ourselves in education: qualification, socialization and subjectification revisited. *Educ. Theor.* 70 (1), 89–104.
- Bransford, J.D., Brown, A.L., Cocking, R.R. (Eds.), 2000. *How People Learn: Brain, Mind, Experience, and School*, expanded ed. National Academy Press, Washington, DC.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement*. University of Chicago Press, Chicago.
- Bryk, A.S., Gomez, L.M., Grunow, A., Le Mahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press, Cambridge, MA.
- Bush, T., Bell, L., Middlewood, D. (Eds.), 2019. *Principles of Educational Leadership and Management*, third ed. Sage, London.
- Centre for Education Statistics and Evaluation, 2018. *Cognitive Load Theory: Research that Teachers Really Need to Understand*. New South Wales Department of Education, Sydney.
- City, E., Elmore, R.F., Fiarman, S., Teitel, L., 2009. *Instructional Rounds in Education*. Harvard Education Press, Cambridge, MA.
- Clark, R.E., Kirschner, P.A., Sweller, J., 2012. Putting students on the path to learning: the case for fully guided instruction. *Am. Educat.* 36 (1), 6–11.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D., 2019. Implications for educational practice of the science of learning and development. *Appl. Dev. Sci.* 24 (2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>.
- Deans for Impact, 2015. *The Science of Learning*. Deans for Impact, Austin, TX.
- Department for Education, 2014. *The National Curriculum in England: Framework Document*. Department for Education. <https://www.gov.uk/government/organisations/department-for-education>.
- Doyle, W., 1983. Academic work. *Rev. Educ. Res.* 53 (2), 159–199.
- Doyle, W., 1986. Classroom organization and management. In: Wittrock, M.C. (Ed.), *Handbook of Research on Teaching*, 3rd ed. American Educational Research Association, New York, pp. 392–431.
- Doyle, W., 2006. Ecological approaches to classroom management. In: Weinstein, C.S., Evertson, C.M. (Eds.), *Handbook of Classroom Management: Research, Practice and Contemporary Issues*. Routledge, New York, NY, pp. 97–126.
- Elmore, R.F., 2004. *School Reform from the Inside Out: Policy, Practice, and Performance*. Harvard Education Press, Cambridge, MA.
- Fay, B., 1987. *Critical Social Science: Liberation and its Limits*. Cornell University Press, Ithaca, New York, NY.
- Filges, T., Torgerson, C., Gascoine, L., Dietrichson, J., Nielsen, C., Vinholt, B.A., 2019. Effectiveness of continuing professional development training of welfare professionals on outcomes for children and young people: a systematic review. *Campbell Syst. Rev.* <https://doi.org/10.1002/cl2.1060>. Available from:
- Fulmer, S.M., Turner, J.C., 2014. The perception and implementation of challenging instruction by middle school teachers. *Elem. Sch. J.* 114 (3), 303–326. <https://doi.org/10.1086/674053>.
- Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W., Hesse, F., 2018. Advancing the science of collaborative problem-solving. *Psychol. Sci. Publ. Interest* 19 (2), 59–92.
- Hand, M., 2014. Aims, concept of. In: Phillips, D.C. (Ed.), *Encyclopedia of Educational Theory and Philosophy*, vol. 1. Sage, Thousand Oaks, CA, pp. 30–32.
- Latham, G.P., Locke, E.A., 2006. Enhancing the benefits and overcoming the pitfalls of goal setting. *Organ. Dynam.* 35 (4), 332–340.
- Latham, G.P., Locke, E.A., 2007. New developments and directions for goal setting research. *Eur. Psychol.* 12 (4), 290–300.
- Louis, K.S., Robinson, V.M.J., 2012. External mandates and instructional leadership: school leaders as mediating agents. *J. Educ. Adm.* 50 (5), 629–665.
- MacAllister, J., 2016. What should educational institutions be for? *Br. J. Educ. Stud.* 64 (3), 375–391. <https://doi.org/10.1080/00071005.2015.1131811>.

- MacIntyre, A., Dunne, J., 2002. Alasdair MacIntyre on education: in dialogue with Joseph Dunne. *J. Philos. Educ.* 36 (1), 1–19. <https://doi.org/10.1111/1467-9752.00256>.
- Marzano, R.J., Waters, T., McNulty, B., 2005. *School Leadership that Works: From Research to Results*. ASCD and McREL, Aurora, CO.
- McInerney, D.M., 2011. Culture and self-regulation in educational contexts. In: Zimmerman, B.J., Schunk, D.H. (Eds.), *Handbook of Self-Regulation of Learning and Performance*. Routledge, pp. 442–464.
- Mestry, R., 2019. Resource management. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *Principles of Educational Leadership and Management*, third ed. Sage, London, pp. 149–164.
- Miles, K.H., Frank, S., 2008. *The Strategic School: Making the Most of People Time and Money*. Corwin Press, Thousand Oaks: California.
- Ministry of Education and Research, 2019. *Core Curriculum—Values and Principles for Primary and Secondary Education*. Ministry of Education and Research. Available from: <https://www.regjeringen.no>.
- Ministry of Education, 2018. *Statement of Intent 2018–2023*. Available from: <https://www.education.govt.nz>.
- National Research Council, 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. The National Academies Press, Washington DC.
- Nuthall, G., 2007. *The Hidden Lives of Learners*. NZCER Press, Wellington, New Zealand.
- OECD, 2005. *The Definition and Selection of Key Competencies: Executive Summary*. Available from: <http://www.oecd.org>.
- OECD, 2019a. *OECD Future of Education and Skills 2030 (Concept Note: Core Foundations for 2030)*. Available from: <http://www.oecd.org>.
- OECD, 2019b. *OECD Future of Education and Skills 2030: Transformative Competencies for 2030*. Available from: <https://www.oecd.org>.
- Peeters, A., Robinson, V., Rubie-Davies, C., 2020. Theories in use that explain adolescent help seeking and avoidance in mathematics. *J. Educ. Psychol.* 112 (3), 533–550. <https://doi.org/10.1037/edu0000423>.
- Pring, R., 2014. Leadership: skilled manager or virtuous professional?. In: *Investing in Our Education: Leading, Learning, Researching and the Doctorate (International Perspectives on Higher Education Research)*, vol. 13. Emerald Group Publishing Limited, pp. 59–73.
- Robinson, V.M.J., Lloyd, C., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership type. *Educ. Adm. Q.* 44 (5), 635–674. <https://doi.org/10.1177/0013161X08321509>.
- Robinson, V.M.J., 2006. Putting education back into educational leadership. *Lead. Manag.* 12 (1), 62–75.
- Robinson, V.M.J., 2008. Forging the links between distributed leadership and educational outcomes. *J. Educ. Adm.* 46 (2), 241–256. <https://doi.org/10.1108/09578230810863299>.
- Robinson, V., 2018. *Reduce Change to Increase Improvement*. Corwin, Thousand Oaks, CA.
- Shamir, B., 2013. Leadership in context and context in leadership studies. In: Rumsey, M.G. (Ed.), *The Oxford Handbook of Leadership*. Oxford University Press, New York, NY, pp. 343–355.
- Spillane, J.P., 2013. The practice of leading and managing teaching in educational organizations. In: OECD (Ed.), *Leadership for 21st Century Learning*. OECD Publishing, pp. 59–82.
- Sun, J., Leithwood, K., 2012. Transformational school leadership effects on student achievement. *Leader. Pol. Sch.* 11 (4), 418–451.
- Sun, J., Leithwood, K., 2015. Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *Sch. Effect. Sch. Improv.* 26 (4), 499–523.
- Swanton, C., 2016. A virtue ethical theory of role ethics. *J. Value Inq.* 50 (4), 687–702.
- Tekkumru-Kisa, M., Stein, M.K., Doyle, W., 2020. Theory and research on tasks revisited: task as a context for students' thinking in the era of ambitious reforms in mathematics and science. *Educ. Res.* <https://doi.org/10.3102/0013189X20932480>.
- Utdanningsdirektoratet (n.d.). *National Curriculum for Knowledge Promotion in Primary and Secondary Education and Training: The Quality Framework*. Available from: <https://www.udir.no>.
- Van Gog, T., Kester, L., Paas, F., 2011. Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Appl. Cogn. Psychol.* 25 (4), 584–587. <https://doi.org/10.1002/acp.1726>.
- Wang, M.-T., Holcombe, R., 2010. Adolescents' perceptions of school environment, engagement and academic achievement in middle school. *Am. Educ. Res. J.* 47 (3), 633–662. <https://doi.org/10.3102/0002831209361209>.
- Welsh, D., Bush, J., Thiel, C., Bonner, J., 2019. Reconceptualizing goal setting's dark side: the ethical consequences of learning versus outcome goals. *Organ. Behav. Hum. Decis. Process.* 150, 14–27.
- Winch, C., 2002. The economic aims of education. *J. Philos. Educ.* 36 (1), 101–117. <https://doi.org/10.1111/1467-9752.00262>.
- Witziers, B., Bosker, R.J., Krüger, M.L., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 39 (3), 398–425.

Re-conceptualizing instructional leadership through a science of learning lens

Steve Myran, Old Dominion University, Darden College of Education and Professional Studies, Department of Educational Foundations and Leadership, Norfolk, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Re-conceptualizing instructional leadership through a science of learning lens	11
Leadership content knowledge – bringing specificity to instructional leadership	12
The crossroad between scientific management and facilitating learning	12
Purpose statement	13
The missing paradigm	13
Methods: problem posing, thinking with theory, and theory building	13
Thinking with theory: using the Myran and Sutherland's synthesis of the science of learning as an analytic lens	14
A gestalt synthesis of the science of learning	14
Incongruent paradigms	14
Analysis	14
Findings	15
From management to learning: the evolution of instructional leadership and leadership content knowledge	15
From management to learning	16
Leadership content knowledge: refining the focus on learning	16
An emergent focus on learning	17
Hinting at theories of learning	17
Ambiguity, folk epistemologies and social consensus	18
Framing instructional leadership and learning through the lens of management	18
Linear, hierarchical, and cause and effect assumptions about learning	19
Hinting at reciprocal causation	19
The legacy of scientific management: incongruent theories of action	20
Discussion and implications	20
Navigating the learning turn	20
References	21

Re-conceptualizing instructional leadership through a science of learning lens

Over the past 30–40 years the emphasis on the impact of instructional leadership on student learning has gained increasing attention in both the scholarly and applied literatures (e.g., [Glickman et al., 2001](#); [Hallinger et al., 2020](#); [Hornig and Loeb, 2010](#); [O'Donnell and White, 2005](#)), becoming a nearly ubiquitous component of the school leadership field ([Hallinger et al., 2020](#)). Born in part out of the theory movement ([Halpin, 1958](#)), what some called the *New Movement* ([Greenfield and Ribbins, 2005](#); [Hughes, 2004](#)) in the middle of the twentieth century, and more pointedly, the effective schools movement ([Edmonds, 1979](#)) that followed, this literature highlights that school leaders who develop strong instructional leadership skills can have a marked impact on student learning. Specifically, school leaders who develop the capabilities to observe, assess, and identify needed instructional supports for students and teachers, to give feedback that helps teachers to iteratively improve their practice, evaluate curricular materials, identify and provide needed professional development opportunities, and lead systemic change are said to positively impact student performance (e.g., [Graczewski et al., 2009](#); [Leithwood and Seashore-Louis, 2011](#); [Olsen, 2006](#); [Robinson, 2010](#); [Steele et al., 2015](#)). Instructional leadership is concerned with what happens in the classroom ([Brazier and Bauer, 2013](#)) and is focused on improving the classroom practice of teachers ([Leithwood et al., 2004](#)) and as such, is focused on the organizational conditions that support student learning.

Nonetheless, while the literature provides bold language about instructional leadership, emphasizing a robust vision of research based teaching and learning ([Brazier and Bauer, 2013](#)), and a combination of philosophy, knowledge, and action ([McGhee and Lew, 2007](#)), the literature offers limited examples of precisely what this looks like ([Brazier and Bauer, 2013](#); [Neumer-ski, 2013](#)), and has been poorly defined ([Louis et al., 2010](#)). Much of this research has focused on what effective leaders “do,” not on how they think about what they do ([Stein and Nelson, 2003](#)), and the epistemic view of the student in most instructional leadership models overlook or ignore the reciprocal role of the students themselves. While it is well excepted in the broader educational administration literature that leadership has an impact on student learning, scholars and practitioners have struggled to define instructional leadership, allowing the concept to be interpreted and used in expansive ways ([McGhee and Lew, 2007](#)). Moreover, instructional leadership has been conceptualized, in large measure, through studies of school principals' experiences and perspectives and reified through current leadership roles and practices ([English, 2012](#)). Many of the most cited works on instructional leadership are consensus building articles that review the current models to develop a unified framework of existing research ([Hitt and Tucker, 2016](#)).

Despite broad intuitive and research based connections between instructional leadership and student learning, the knowledge of how instructional leadership impacts student learning is limited (Neumerski, 2013; Shaked, 2021). Similarly, because it has been operationalized so broadly it can be ambiguous and unobtainable (Brazner and Bauer, 2013; Rigby, 2014; Salo et al., 2015). Perhaps most importantly much of this research has tended to conflate student performance outcomes with learning, which is a an overarching challenge for the broader field, with notable practical and theoretical implications (Soderstrom and Bjork, 2015). This conflation of learning for performance risks allowing “surface level compliance oriented performance on standardize test to suffice for engaged, deliberative, meaningful, enduring, and transferable learning” (Myran and Sutherland, 2019a). In addition, the conceptualizations and definitions of instructional leadership don’t address how these leadership competencies and actions impact the actual cognitive and affective behaviors of students and as such, understanding the nature and quality of the learning experience and the learning “outcomes” is difficult to ascertain.

This conflation of learning and performance reveals the incongruence between the normative views on learning found in this literature and the empirical study about learning and motivation found in the learning sciences (e.g., educational, cognitive, and developmental psychology). This lack of a clearly defined operational definition of learning and the nature of the influence that instructional leadership has on students’ cognition and learning outcomes allows these gaps in knowledge to be filled with the assumptions of the prevailing logic, folk epistemologies, and social consensus (Anderson, 2001; English, 2000; Myran and Sutherland, 2019b), reinforcing “the normatively desirable role for principals who wished to be effective” (Hallinger, 2010, p. 5).

Leadership content knowledge – bringing specificity to instructional leadership

Bringing greater context and specificity to theories of instructional leadership, Stein and Nelson (2003), pointed out that all instructional leadership skills and knowledge are situated in the subject matter being taught and therefore, the leader’s knowledge of content will sharpen the lens through which they look at, understand, and support instruction and learning. Building from the growing interest in teacher pedagogical content knowledge (Shulman, 1986, 1987), which emphasizes a confluence of strong content knowledge and appropriate pedagogical knowledge needed to teach that specific content, Stein and Nelson (2003) introduced the concept of *leadership content knowledge* (LCK). Leadership content knowledge emphasizes that principals and other school leaders not only need to be conversant in the skills and knowledge associated with instructional leadership, but to developed fluency with the content being taught (Steele et al., 2015). However, little is known about how leaders’ LCK intersects with other instructionally and learning focused efforts to support teaching and learning (Quebec Fuentes and Jimerson, 2020), and the empirical evidence of its impact on teaching and learning is limited.

Moreover, like the broader work in instructional leadership, LCK also conflates performance outcomes with learning and does not conceptualize how this construct impacts students cognitive and affective learning behaviors. In this way, the construct, while compelling on a number of levels appears to sidestep the issue of how LCK influences the actual cognitive and affective behaviors of the students themselves. The impact is indirect, leaving important nuances about learning found in educational psychology and the related learning sciences, underexplored.

The crossroad between scientific management and facilitating learning

Myran and Sutherland’s (2019b) identification of the incongruence between educational leadership’s historic grounding in scientific management and managerial principles and the robust science of learning found in educational psychology and the related science of learning, may offer a means for identifying an underexplored antecedent to these challenges.

They point out that these differing theories of action represent competing conceptualizations about the nature of learning. On the one hand, the prominent managerial paradigm tends to frame learning as a linear, one-way delivery of content to students who in turn passively receive this content. On the other hand, the science of learning has painted an increasingly clear picture of the complex, dynamic and reciprocal interaction between the learners active and deliberate agency, their introspective capabilities, and the school’s social, ecological, and epistemological conditions. The first framing students as receiving an education, and the second framing student’s agency at the center, and the school and social conditions as reciprocally and bidirectionally interacting to foster learning. This differentiates structures that *produce* learning with those that *facilitate* learning (Myran and Sutherland, 2019b).

In this way, I argue that instructional leadership sits at the crossroads between long standing scientific management influenced habituated norms and assumptions about the goals and purpose of educational leadership and the growing push for school leaders to be managers *and* instructional leaders. Salo et al. (2015), described this as being at the “center of professional crossfire, between, on one hand, management practices built on efficiency, accountability and consumer orientation and, on the other, a collegial culture of trust and professional traditions built on ideals of democratic citizenship” (p. 494), what Moos et al. (2004) refer to as the “cross pressures” that school leaders face in their daily work between management and facilitating learning.

In short, if the prominent paradigm is grounded in principles of efficiency, uniformity, and productivity, then so too are the language, theories, metaphors, and strategies used by the field. How can we define and operationalize instructional leadership and leadership content knowledge with the language, theories, metaphors, and strategies of scientific management?

Even as the field has continued to evolve the specific understandings of how school leaders may influence the instructional climate and impact student learning, beyond lists of leadership traits and general behaviors (Neumerski, 2013), largely derived from professional consensus (English, 2000), we lack understanding of the process of how school leaders impact student learning

in their daily work and how these processes scaffold students cognitive and affective behaviors. If we are to address the challenges of refining our understanding of how school leadership shapes the instructional climate in ways that promote student learning, we need to reframe the narrative, metaphors, and strategies around a more appropriate science, that is the science of learning.

Purpose statement

In this chapter I identify the incongruence between our field's grounding in scientific management and the science of learning and how these disjunctures have shaped our conceptualizations of the school leaders' role in supporting instruction and student learning. Drawing from the related traditions of problem identification, theory building and thinking with theory, I unpack the instructional leadership and leadership content knowledge literatures and examine it through [Myran and Sutherland's \(2019b\)](#) framework to reveal the potential antecedents to the cross-pressures experienced by school leadership seeking to both manage and facilitate learning. The emergent focus on learning found broadly in the educational leadership field, and more specifically in the instructional leadership and leadership content knowledge literature appears to hold great promise for not only facilitating learning in general, but as a means of further addressing long standing structural and historic learning inequities. However, I build an argument that unless we address the various incongruences identified by Myran and Sutherland, we will continue to struggle to evolve these constructs to be better aligned to our field's core purpose and fail to achieve their potential.

Reinforcing this perspective, as [Bandura \(1985\)](#) stated, "what theorists believe people to be determines which aspects of human functioning they explore most thoroughly and which they leave unexamined. Conceptions of human nature thus focus inquiry on selected processes and are in turn strengthened by findings of paradigms embodying the particular view" (p. 81). In this way the more deterministic means by which our field has framed the role of school leaders in shaping student learning reifies these instrumental and functionalist perspectives and in turn constrain the epistemological, ontological, and methodological approaches to exploring these relationships. This chapter offers a means of unpacking the conceptualizations of leadership's role in shaping student learning, identifying the constraining incongruences, and reconceptualizing it through a science of learning lens.

The missing paradigm

As [Shulman \(1986\)](#) highlighted regarding the development of teacher pedagogical content knowledge, so neglected in the scholarly and applied fields was teachers' subject matter knowledge that they called it the "missing paradigm" in research on teaching. This seems to me a sort of counter-managerial pronouncement, an assertion that teaching is more than the delivery of standardized and codified content to passive students. I build from this notion of the missing paradigm, suggesting that the science of learning is the missing paradigm in educational leadership broadly, but specifically for instructional leadership.

In previous research, my colleague and I suggested that there is a fundamental paradox between the historically and conceptually rooted managerial leadership paradigm and our growing understanding about the principles and theories of learning found in the learning and psychological sciences ([Myran and Sutherland, 2019b](#)). We argued that even with the increased interest in more student centered outlooks about leadership and schooling, there is a tendency to replicate the linear cause and effect, hierarchical, and "decidedly functionalist and positivistic", if not "exclusively functionalist and instrumental view of the principal's role" ([Hallinger and Heck, 1996](#)). Offering a gestalt synthesis of the rich research on learning and motivation, we suggested that these managerial proclivities represent substantive departures from theories of learning and without a robust understanding of the theories and science of learning, we will continue to be tethered to a frame of reference that is antithetical to the core purpose of schooling – that is learning.

Methods: problem posing, thinking with theory, and theory building

As [Sarton \(1929\)](#) emphasized, "the most difficult thing in science, as in other fields, is to shake off accepted views" (p. 88). Framed within this challenge, I draw from [Myran and Sutherland's \(2019b\)](#) call to reframe the leadership narrative around the science of learning. Using theories of learning as a counter-managerial lens for understanding the reciprocal nature of human interaction, learning and knowledge creation, I engaged in deep problem identification utilizing the concept of *thinking with theory* ([Jackson and Mazzei, 2011](#)) to explore areas where instructional leadership and leadership content knowledge might continue to advance beyond the confines of the managerial project and its functionalist and instrumental assumptions about learning.

Specifically, I draw from problem posing/identification strategies ([Brown and Walter, 2012; Freire, 2013, 2021](#)), that have a strong focus on criticism and problematizing unrecognized or under-recognized assumptions and their implications and consequences ([AERA, 2009](#)), thinking with theory ([Jackson and Mazzei, 2011](#)), as a means of offering strong theoretical basis to problematizing challenging problems of practice and theory building methods ([Lynham, 2002; Lewis and Grimes, 1999; Gioia and Pitre, 1990](#)), as a means of offering new or refined theory that is better conceptualized to address the identified theoretical problems and problems of practice. I conceptualize the use of these related constructs as reciprocal; building on each and offering iterations of improved problem identification and the development of theory that is better adapted to the primary goals of organizational improvement and student learning.

Thinking with theory: using the Myran and Sutherland's synthesis of the science of learning as an analytic lens

First, in framing my use of theory as central to my analysis, I emphasize that theory shapes how we see, experience, and think about the world, whether we are cognizant of the specific theory or theories at work or if they are implicit to the cultural or organizational norms that shape our understanding of the world. In this way, thinking with theory allows us to be explicit about the various theories that are both knowingly and unknowingly at work in shaping our understanding of a given phenomenon. Moreover, thinking with theory provides a pathway for rigorously exploring new, and potentially more adaptive ways of understanding a given phenomenon and to “shake us out of the complacency of seeing/hearing/thinking/feeling as we always have ...” (Jackson and Mazzei, 2011), to make plain the epistemological and ontological assumptions that undergird and constrain how we see and understand the world (Jackson and Mazzei, 2011), and as a form of sensemaking (Weick, 1995). Similarly, theory offers a means of scaffolding for testing, adapting and refining theories in use and theories in action (Argyris and Schon, 1974).

A gestalt synthesis of the science of learning

In short, Myran and Sutherland's gestalt synthesis of the science of learning held that since the cognitive revolution some 70 years ago, compelling counter-managerial evidence describes learning and motivation as driven by the active, deliberate, and introspective agency of the individual which is reciprocally situated within complex and dynamic social contexts that serve to mediate, shape, or constrain the construction and co-construction of knowledge (Myran and Sutherland, 2019b). In this way, the learner is not the product of what schools and educators do. Instead, “the individual's behavioral, cognitive, and introspective capabilities interact bidirectionally with the environment and function as mediating factors that shape and constrain how one interacts within social structures. In this way, “we are both products and producers of our own learning environments” (Myran and Sutherland, 2019b). As Bandura (1989) highlights, “persons are neither autonomous agents nor simply mechanical conveyers of animating environmental influences. Rather, they make causal contribution to their own motivation and action within a system of triadic reciprocal causation. In this model of reciprocal causation, action, cognitive, affective, and other personal factors, and environmental events all operate as interacting determinants. Any account of the determinants of human action must, therefore, include self-generated influences as a contributing factor” (p. 1175).

Incongruent paradigms

Similarly, Myran and Sutherland (2019b) highlighted the disjunctures between the above gestalt synthesis of learning and the prominent managerial paradigm. They emphasized that the impacts of the science of management can be seen in its primary focus on the efficient and uniform operation of schools (Tyack and Cuban, 1995; Tyack, 1974), where all aspects of schooling were explicitly designed through their structures, schedules, and regiments to be analogous to the industrial-age factory (Callahan, 1962). While the direct commitments to scientific management and managerialism have been replaced by theories that increasingly recognize schools as complex socio-ecological organizations, the field's foundations in scientific management are part of our DNA, representing a “set of perverse incentives that shape these roles and steer daily activities toward the managerial” (Cuban, 1988).

Given this historic foundation, Myran and Sutherland (2019b) argued that scientific management continues to dominate our field's outlooks, what some have called Neo-Taylorism (Au, 2011; Gronn, 1982), Managerialism (Gunter, 1997, 2009), or Neo-Managerialism (Terry, 1998). As such, the managerial design philosophy reduces the complex and reciprocal interactions between and among the learner and the socio-educational ecology to mere “strategism”, what Morozov (2013) called “technological solutionism”, that is the formalized, routine, and procedural dissemination of knowledge. This dominant cultural perspective has in turn shaped our thinking about the nature of learning and teaching, how we conceptualize and design schools, as well as our roles and responsibilities within them. These scientific management influenced organizational practices assume (1) knowledge is a collection of facts and procedures, (2) the goal of education is to disseminate these facts and procedures, (3) teachers' job is to transmit this knowledge to passive students, (4) that simpler facts proceed more complex ideas, and (5) success is achieved when students can demonstrate they have mastered these facts and procedures through testing (Sawyer, 2006b).

Analysis

Drawing from the above, I used the peer-reviewed and seminal literature on instructional leadership and leadership content knowledge as my source of data and conducted a systematic review of these literatures from a thinking with theory and theory building perspective. That is, I interrogated this literature looking for congruence and incongruence with the science of learning, identifying emergent perspectives on human agency and the social and reciprocal dynamics of learning, as well as identifying aspects of the literature that were theoretically and practically tied to managerial perspectives. I identified articles on these topics using key word searches in both Google Scholar and the Education Resources Information Center (ERIC) with permutations of *instructional leadership* and *leadership content knowledge*, followed references and citations as indicators of impact on the field, followed lines of reasoning to new articles, and actively sought out counter evidence to curtail a priori biases.

Once identified, I created an EndNote bibliography including PDFs of all articles. Next, I imported this bibliography along with all the PDFs into NVivo qualitative research software. Using Myran and Sutherland's (2019b) gestalt synthesis of the science of learning literature and managerial critique discussed above, through an iterative and reciprocal process, I read, coded, and annotated all of the articles, first through a rounds of open coding, followed by subsequent rounds of axial and selective coding to

develop a thematic structure. As my reading and coding revealed possible additional articles, I added those to my databased and integrated these new data into my workflow. The process of writing the manuscript itself proved to be part of the iterative and reciprocal nature of this work, offering a means of sense-making, and refinement of the final thematic structure, which is outlined below.

Findings

The resulting analysis highlighted four interdependent themes. They are, (1) an evolving transformation from a dominant focus on the management and supervision of learning to an increasingly student centered and instructional leadership orientation, (2) the emergent, yet indirect referencing of the principals of learning, (3) given the identified disjunctures between management and the science of learning, the leader's role in supporting learning continues to be somewhat ambiguous and as a result, a host of folk epistemologies have evolved that are antithetical to the science of learning. Finally, (4) there is a strong theme of the legacy of scientific management found throughout the literature reviewed which reveals important, yet under-identified incongruent theories of action that must be examined, understood, and overcome in order to meet the potential of this evolving focus on learning (see Fig. 1).

I conceptualized this thematic structure as interrelated, irreducible, and cyclical. Each theme provides the next the basis for understanding the larger model and together they work as a self-perpetuating and self-referential system. The overall model presents a struggle between an emergent or nascent turn from a prominent managerial theory of action to a learning focused theory of action grounded in an understanding of the reciprocal causation among the various actors within complex social-ecological systems. Within this struggle we see this nascent focus on learning influenced by, if not misshapen by the managerial traditions and associated historically habituated structures, norms, and assumptions that are fundamentally incongruent with the science of learning. Given the incongruent nature of the theories of action found in the last theme in the model, these disjunctures between the science of learning and the legacy of scientific management perpetuate the struggle between the potential of instructional and content leadership and the managerial "conduits of pressure" (Mintrop and Trujillo, 2005; Mintrop, 2004) experienced by school leaders. In turn, this dynamic helps to maintain the ambiguity of the principal's role in supporting learning, reifies folk epistemologies, perpetuates the legacy of scientific management, and reinforces the incongruent theories of action. Each of these interrelated themes, along with examples from the literature reviewed are discussed in greater detail in the subsequent sections of the chapter.

From management to learning: the evolution of instructional leadership and leadership content knowledge

A prominent theme that emerged through my analysis was how educational leadership has been in a period of transition from a primary focus on the management and supervision of schooling to a more learner-centered and pedagogical orientation. This

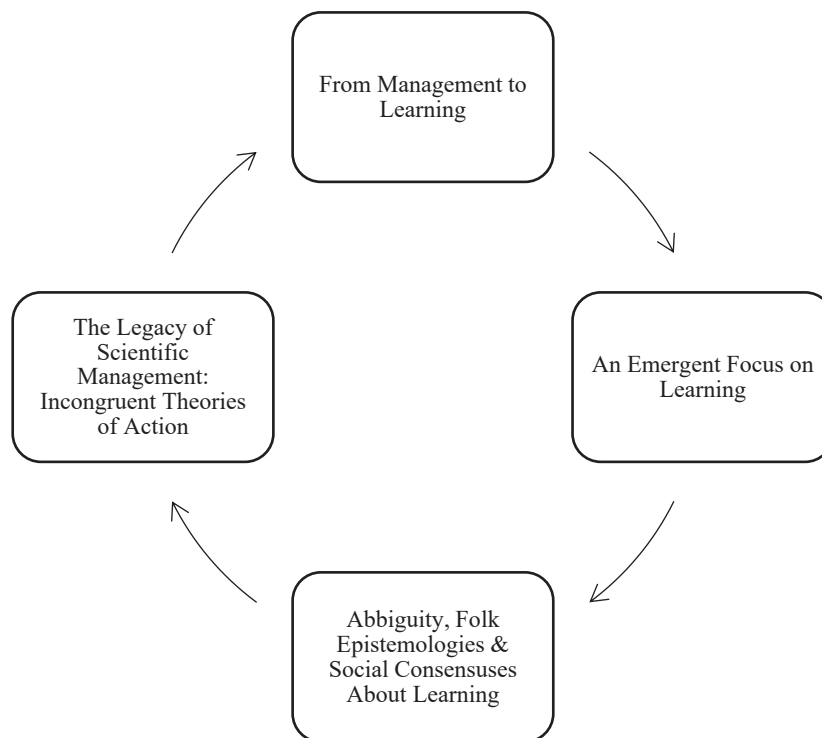


Fig. 1 Self-replicating incongruent theories of action.

literature paints a picture of a growing interest in the role school leaders play in developing, supporting, and maintaining a climate of high-quality teaching and learning. Central to this theme was a clear thread that ran through this literature that contrasted the leadership characteristics and capabilities of supporting teacher quality and student learning from those associated with the management and oversight of a school or school system. Relatedly, these management responsibilities tended to be framed as a fundamental barrier to the implementation and fidelity of instructional leadership. As scholars and practitioners have grappled with the transitions between management and instruction there has been a tendency to bifurcate the management of a school or school system from instructional leadership, what [Shen \(2020\)](#) called the bifurcated fault line between the school/principal and the teacher/classroom level.

From management to learning

Most scholars attribute the development of the construct of instructional leadership to effective schools research, generally attributing this to scholars such as [Edmonds \(1979\)](#), [Purkey and Smith \(1983\)](#), and [Weber \(1971\)](#), and the increasing interest in the role that principals and other school leaders play in impacting student learning. [Weber's \(1971\)](#) influential research actively differentiated the focus from the prior work of Coleman, Jensen and other researchers that made assumptions that the low achievement of poor children was primarily the result of "inherent disabilities characterizing the poor" ([Edmonds, 1979](#)). Building from Weber's work, [Edmonds' \(1979\)](#) seminal research, *Effective Schools for the Urban Poor*, which summarized a number of other studies, emphasized that effective schools have strong administrative leadership, are instructionally effective, promote a school's atmosphere that is orderly and safe without being too rigid or oppressive, prioritize learning over other school activities, and frequent monitoring. Just as importantly, Edmonds was emphatic that student's background was not a determinant in their ability to learn and emphasized that student failure was in part a school failure.

At the same time, this emergent shift from a primary focus on management to one of learning was nonetheless notably shaped by the prevailing scientific management design philosophy. Even with the increasing evidence of the limitations of scientific management as an appropriate frame of reference for education, because the massive bureaucracies that support public education had been well established prior to a science of learning ([Sawyer, 2006a](#)), the underlying organizational norms, values, assumptions, and practices were well entrenched ([Myran and Sutherland, 2019b](#)). It could be argued that the tenets of the effective schools movement were shaped by the same positivist frameworks that well may have been a key source of the dissatisfaction with traditional managerial models, and as such they ignored or overlooked the multidimensional social and cultural intricacies that managerial and managerial design philosophies are not effective at addressing ([Dantley, 1990](#)). The lists of factors of essential schools were "grafted" onto schools, a graft that rarely took hold due to competing ideologies ([Cuban, 1998](#)) and the well-established managerial perspectives.

In the more contemporary instructional leadership literature, the field continues to struggle with the transition from management to learning, what some have referred to as a shift from behaviorism to constructivism ([Nelson and Sassi, 2000](#)), or what [Robinson \(2006\)](#) referred to as putting education back into educational leadership. This literature offers a rich view on how scholars have grappled with the tensions between the dissatisfaction with the prominent managerial paradigm, the need to shift more effectively to learning centered models of leadership and the limitations of the historic grounding in managerialism. Here we can see an increasing emphasis that shifts from broad school level factors such as promoting high expectations, facilitating a positive school climate, decentralizing decision-making, to more teacher, classroom and student focused skills such as observing and assessing instruction, identifying instructional supports for struggling students, giving feedback to teachers, evaluating curricular materials, and engaging in meaningful, sustained instructionally focused conversations with teachers, and building and leveraging their knowledge of the content being taught to sharpen their focus on all of the above ([Carver, 2012](#); [Quebec Fuentes and Jimerson, 2020](#); [Steele et al., 2015](#); [Stein and Nelson, 2003](#); [Gracjewski et al., 2009](#); [Hallinger, 2010](#); [Robinson, 2010](#)).

Leadership content knowledge: refining the focus on learning

As discussed above, the instructional leadership literature highlights an evolving shift from managerial to learning perspectives and an increased focus on those elements that are closer to the source of learning, that is the classroom interactions between and among students and teachers. Within this increasingly more nuanced and learning-centered exploration, leadership content knowledge offers a refining focus on learning, helping to fill in gaps between the broad construct of instructional leadership and the nuanced skills, knowledge, and behaviors that may serve to scaffold and support learning at the actual source of learning. In this way, leadership content knowledge may serve as a means of sharpening the lens through which school leaders look at, understand, and support instruction and learning ([Stein and Nelson, 2003](#); [Steele et al., 2015](#)).

Just as importantly, as the field has moved from broad school level factors, to more teacher and classroom factors, there has also been a growing recognition of the importance of school leaders knowledge of the nature of student engagement ([Nelson and Sassi, 2000](#); [Steele et al., 2015](#)), a further indication of the push for more fully developed learning focused theories of action that can replace traditional managerial orientations. Just as the development of pedagogical content knowledge prompted new insights about teaching, leadership content knowledge offers a similar means of pushing the field further toward a more learning focused construct ([Stein and Nelson, 2003](#)). Taken together, leadership content knowledge is considered a co-defined construct, that is the blending of instructional leadership, content knowledge, and pedagogy ([Stein and Nelson, 2003](#)).

However, outside of the knowledge of the content itself, and the anecdotal and craft knowledge that comes with experience, there is a notable challenge in identifying the source of this pedagogical understanding. How do school leaders or teachers identify modes of inquiry best suited for a particular discipline, effective curriculum, and evidence-based practice in the absence of a robust

foundation in, and understanding of the science of learning? As Myran and Sutherland (2019b) point out, under the design principles of scientific management and managerialism, there is an assumption of a linear, cause and effect impact of leadership on the school climate, teachers and consequentially, student learning. The assumption is that the influence of leaders is primarily at the organizational and school level, and the evidence about the impacts on student learning begins at the school level and is traced back to through the levels of this hierarchy. As a result, the distance between school level factors and the more nuanced and bi-directional interactions between students, teachers, and the school climate, abstracts the students' active and reciprocal role and may risk defaulting to assumptions about students as the products of these school-level mechanisms and leadership practices. As Stein and Nelson (2003) note, "Without knowledge that connects subject matter, learning, and teaching to acts of leadership, leadership floats disconnected from the very processes it is designed to govern" (p. 446). This is a critical point, and I would push this logic further and suggest that without a foundation in the science of learning, leadership will continue to be disconnected from that which it seeks to have influence over.

An emergent focus on learning

Closely related to the theme of the evolving transformation from management to learning, I noted an emergent, yet indirect referencing of the principals of learning. Within this emergent shift from management to instructional leadership, we can also see a pattern that begins to hint at science of learning perspectives, moving from broad school level factors to classroom and teacher factors, and finally to more serious considerations for student level cognitive and motivational factors. Within this theme I identified a pattern of evidence that hinted at theories of learning, a push to move the field beyond generic conceptions of learning focused leadership, as well as a push to ground the field more fully in the core business of education, that is learning. Taken together these first two themes could be considered as an indication within the field of a *learning turn*, a paradigm shift from a managerial outlook about learning to one informed by the science of learning.

Hinting at theories of learning

Beyond the general shift from managerial to learning foci, I also observed an emergent emphasis on a more learner and person-centered framing of teaching and learning. Here the literature emphasized that instructional leadership should be fundamentally grounded in the core purpose of schooling (Mullen and Schunk, 2010), establishing a greater understanding of student and teacher thinking as foregrounding to the work of educational leadership (Steele et al., 2015), school leaders seeing themselves as learners (McGhee and Lew, 2007) and recognizing the complex and dynamic nature of learning and school organizations (Brazier and Bauer, 2013). More specifically, these person and learning-centered frameworks highlighted a participatory (Mullen and Schunk, 2010), bidirectional and reciprocal outlook about the nature of learning (Steele et al., 2015; Stein and Nelson, 2003; Mullen and Schunk, 2010), the importance of collaboration, discourse, and dialogue (Robinson et al., 2008; Quebec Fuentes and Jimerson, 2020; Nelson and Sassi, 2000; Brazier and Bauer, 2013; Mullen and Schunk, 2010), student, teacher, and leader engagement, (Graczewski et al., 2009; Nelson and Sassi, 2000), and intellectual risk taking (Nelson and Sassi, 2000).

In this way, leadership content knowledge provides the grounding for the core function of schooling, moving beyond what effective leaders "do," to how they think about, internalize, and understand what they do (Stein and Nelson, 2003) in reciprocal interactions with others in the learning context and in relationship to the core purpose of schooling. Stein and Nelson emphasized that learning is facilitated by what they described as "asymmetries in expertise" and expertise within the larger school organization shifts among different individuals or groups based on the nature of the work at hand. As such, they highlight that the job of school leaders is to organize and facilitate such that these knowledge resources can interact in productive ways; that is not transmitting knowledge, but discovering the learning needs of the organization, creating spaces for engagement, incentivizing learning, and assuring there are appropriate learning resources (Stein and Nelson, 2003). Instead of focusing on how students internalize content, school leadership should be more focused on how students think and engage with others, a shift that Brazier and Bauer's (2013) view as "seismic", that is, thinking about learning as the consumption of knowledge "misses the point. The action is in how students know in relation to the problems at hand and to their fellow students" (p. 3).

Emphasizing building leadership orientations around the core purpose of learning, Mullen and Schunk (2010) have emphasized that when school leaders engage in instructional leadership that is participatory in nature, collaborative, and reciprocal, authentic, involves co-learning, deep cultural exchange, shared power and control, and active engagement among critically reflexive participants, they are able to foster meaningful improvements in the instructional capacity of the school. Leadership content knowledge helps to connect the core function of schooling, learning, and teaching to leadership – helping to move beyond generic perspectives on school leadership. Without this, the work of school leadership is disconnected from the very processes it is designed to govern (Carver, 2012).

Similarly, strong content knowledge has been shown to help school leaders foreground student thinking, and student/teacher interactions, shifting their frame of reference from surface-level characteristics of effective teaching to the more complex and dynamic aspects of student thinking and student/teacher engagement (Steele et al., 2015), suggesting a nascent understanding of the reciprocal and bidirectional nature of learning. Representing a related shift in orientation, as McGhee and Lew (2007) have pointed out, school leaders who model intellectual curiosity, set learning goals, and see themselves as lifelong learners, proactively build a culture of continuous learning, work at fostering and sustaining a shared vision of commitment to student success, and more effectively foster a community of learners. However, as powerful as this potential seismic learning turn may be, few of these observations directly referenced the learning and psychological sciences, which offers significant theoretical substantiation for this turn.

Ambiguity, folk epistemologies and social consensus

Even with the evolution of this learning turn, my review of these related literatures suggests that there remain ambiguities about the principal's role in supporting learning. Given the tendency within the larger field to default to scientific management/managerial frames of reference (Au, 2011; Gronn, 1982; Gunter, 1997; Myran and Sutherland, 2019b; Donaldson and Allen-Handy, 2020) and the absence of well-defined applied and theoretical definitions of learning (Myran and Sutherland, 2019b), educators have been left to fill in these gaps based on their own experiences as learners, and their training and professional experience (Printy et al., 2008). This in turn has the effect of amplifying well-established and historically rooted assumptions about leadership, teaching, and learning, and privileging social consensus over a science of learning in the framing of instructionally focused leadership.

Through my analysis, I found that the ambiguity of the school leaders' role in supporting learning may tend to reify *folk epistemologies*. That is, our commonsense, everyday, or, intuitive conceptions of knowledge and knowing. (Gerken, 2020; Hardy-Vallée and Dubreuil, 2010), which are derived from experience, the contextual utility of a knowledge claim, and the subsequent social consensus. The absence of an empirical grounding in the science of learning and the implicit folk epistemologies tend to misrepresent, misunderstand, overlook, or disregard the student's active and deliberate agency and their reciprocal engagement in the socio-ecological learning context, sometimes treating students as the product of educational management.

Pushing against the constraining features of the prominent managerial paradigm, leadership content knowledge scholars (Stein and Nelson, 2003) sought to emphasize a type of connoisseurship and educational criticism (Brazier and Bauer, 2013) that would allow an administrator to know strong instruction when they see it; leveraging what they know about the content of a particular subject to build these capabilities. In this way, subject matter knowledge, pedagogy, and leadership are considered the basis of leadership content knowledge (Stein and D'Amico, 2000; Stein and Nelson, 2003). While this pushes against the limitations of the prominent managerial norms and begins to bring expertise of teaching, content knowledge and pedagogy to the center of leadership theory, it nonetheless leaves an empirical understanding about learning, the very point of these leadership constructs, undefined.

Framing instructional leadership and learning through the lens of management

In the absence of clearly defined theoretical and operational definitions of learning, I further noted a tendency for this literature to conceptualize learning through the prominent managerial lens. Across the reviewed literature I noted an underlying theme of viewing instructional leadership as an additional responsibility added on top of the foundational managerial responsibilities. Here the literature represents the management of the school as the floor, with new school leaders learning the skills of management that allows them to fulfill the basic needs of the school, upon which instructional leadership is grounded (Brazier and Bauer, 2013; Grissom and Loeb, 2011). Similarly, school leaders' management and instructional responsibilities have often been framed in contradictory ways (Brazier and Bauer, 2013; Bridges, 1967; Hallinger, 2011).

Grissom and Loeb (2011) note that one of the challenges for the broad and ambiguous construct of instructional leadership is "distilling which behaviors count as instructional leadership and which do not" (p. 433). This highlights the assumptions of the hierarchical, linear, and recursive impact of leadership on each subsequent and subordinate level of the organization. Here we see language such as, how some leadership actions "yield positive outcomes", which masks the reciprocal role of not only the students themselves, but those in roles subordinate to leadership. Robinson (2006) highlights that educational leaders have been historically assessed on their effectiveness as managers of the fiscal, organizational, and political climate of their building, that is setting the stage or creating the conditions in which teachers can carry out the educational mission. More recently however, there is a growing expectation that school leaders will take charge of the instructional climate as well, directly influencing student learning through their own enhanced knowledge of content and pedagogy and instructional leadership behaviors (Robinson, 2006). The underlying assumption being that through such linear and recursive structures leadership would have a direct and in-direct impact on student learning. Within this framing, many scholars have sought to understand the impact of leadership on learning and how the two are linked, generally assuming a hierarchical causal chain between leadership and student outcomes (e.g., Graczewski et al., 2009; Leithwood and Seashore-Louis, 2011; Olsen, 2006; Robinson, 2010; Steele et al., 2015).

Analysing this observation through a science of learning lens I noted a concerning disjuncture between the assumption of this causal chain and the gestalt synthesis of the science of learning (e.g., Alexander et al., 2009; Myran and Sutherland, 2019b). Some 70 years of post-cognitive revolution research makes it abundantly clear that learning is not a one-way causal dynamic between educators and learners; such a view significantly misrepresents that agency of the student, their introspective capabilities, motivations, and just as importantly, the bi-directional impact of their interactions within the complex social-ecological context of schools and classrooms. Because, "we are both products and producers of our own learning environments" (Myran and Sutherland, 2019b), the determinants of student learning can't be understood outside of what Bandura (1989) described as the causal contribution learners make to their own learning within a system of triadic reciprocal causation. Critically, as discussed earlier in the chapter, Bandura (1989), emphasized that any account of the factors that shape learning must include these reciprocal factors – that is the learners' own agency, their cognitive, affective, and other personal factors in bidirectional interactions with the learning environment, which collectively operate as interacting determinants of learning. As such, the traditional conception of a causal chain between leadership, teachers, and student learning disregards or overlooks the student's own self-generating influences on both the process and outcome of engagement in the learning environment.

Illustrating these disjunctures, one of the quandaries in the leadership content knowledge literature is how a principal or other instructional leaders could practically have substantive enough knowledge about every content area taught in the school to be an effective instructional leader with all teachers and all classrooms (Robinson, 2006; Stein and Nelson, 2003). Two solutions to this

challenge are offered in the literature. The first is through the delegation or distribution of instructional leadership (e.g., Spillane et al., 2003, 2004), and the second is through the concept Stein and Nelson (2003) refer to as *post-holing*, that is grounding a broader understanding of effective instruction in a depth of knowledge in a single subject area. Interestingly both solutions appear to be anchored to a management foundation and as such may overlook or disregard the theories and science of learning that I would argue should undergird the leadership content knowledge construct. This may represent a disjuncture between the foundation in management and the growing turn from this foundation to learning. Without an understanding about the nature of learning itself, in-depth knowledge of a given content area and the pedagogy unique to a given content area is insufficient to build broad and flexible understandings about what effective instruction looks like and how to provide meaningful instructional support to teachers. Here we can see the limitations of managerial outlooks as the foundation that leadership content knowledge is built on.

Linear, hierarchical, and cause and effect assumptions about learning

Across this literature I found a thematic pattern that framed learning within a largely linear, hierarchical, and cause and effect frame of reference. That is, discourse that frames the pressures of accountability as a management task that requires instructional leadership to help improve teaching, which in turn results in producing higher scores on benchmark tests and summative assessments. The assumption here is that principals are responsible for ensuring that teachers increase their knowledge and that this increased teacher knowledge will lead to changes in instruction and produce higher student achievement (Graczewski et al., 2009). In essence, there is a clearly defined linear, hierarchical, and cause and effect assumption about how improvements in learning are achieved, entirely leaving out the agency of the individual actors and the reciprocal role they play in what is in fact a complex and dynamic interplay between the individuals, their agency, motivations, introspective capabilities, and the complex socio-cultural school context.

The language used across the literature often suggested a strong leaning toward this linear and power-centric conceptualization about how improvements in learning take place. Here I found statements to the effect of, the role of schools in achieving student gains, identifying the determinants of student performance, linking improvements in teacher performance to student achievement, how a principals' modeling behaviors can impact teacher quality and student learning, principals as managers that assure that teachers are "delivering" mandated curriculum, and how support is delivered to struggling students. These types of statements seem to reveal a prevailing undertone of a recursive conceptualization of the role school leaders play in student learning, that is a *telescoping* (Printy et al., 2008) phenomenon where at each subsequent level of the organization the responsibility encompasses the learning needs of everyone at the each of the subordinate levels, requiring a superintendent, for example to understand, manage and control the learning needs of principals, teachers, and students. Here leaders become the teachers of the other educators who are subordinate to them, supervising, orchestrating the learning environment, implementing incentives for learning, assuring adequate resources to support learning, and assuming the responsibility for their learning (Printy et al., 2008). These conceptualizations are often represented with student learning outcomes as the dependent variable of the recursive impact of a linear and hierarchical set of independent variables (e.g., Bossert et al., 1982; Dumay et al., 2013).

Hinting at reciprocal causation

Even within the prevailing recursive framing, I also found counterexamples which are closely related to the theme of an emergent focus on learning discussed earlier. While much of the language across the reviewed literature reveals the linear, hierarchical, and cause and effect impact of leadership on student learning, that is the recursive and accumulative impact of leadership at each subordinate level of the organization, there were also counterexamples that present a more reciprocal and human agency supportive framing of the role of instructional leaders. Here I found language to the effect of honoring the professionalism of teachers, engaging in mutually beneficial discussions around instruction, principals and other school leadership learning alongside teachers, collaboration, student attitudes, self-concept, behavior, attitudes of peers, and participative leadership (e.g., Bell et al., 2003; Quebec Fuentes and Jimerson, 2020; Robinson, 2006; Robinson et al., 2008).

More pointedly, a few scholars have challenged the direct and mediated effects models of school leaders' impact on student learning, arguing these models do not capture the complex and dynamic interactions between and among the variables involved in these models (e.g., Hallinger and Heck, 2010, 2011; Oude Groote Beverborg et al., 2020; Shen et al., 2019; Shen, 2020). For example, Hallinger and Heck (2010, 2011) offer a reciprocal-effects model, that is a non-recursive conceptualization of the dynamic and changing relationships among the relevant variables over time. Here leadership is viewed not as a unitary and independent influence on learning, but one that both gives shape to, and is in return shaped by the other variables in the model. In other words, leadership and the other variables that may influence learning have both a cause and an effect on each other. This echoes aspects of Bandura's social learning theory (1979, 1985) and reciprocal determinism model (1978, 1984) from some 40 years ago. This may represent an evolution of how knowledge and learning are conceptualized; from a recursive and reductionist direct effect model, to mediated effects models that recognize that simple, direct, cause and effect framing oversimplifies the confluence of complex factors that come together in the learning setting, to a reciprocal model that is better supported within the science of learning (e.g., Bandura, 1978, 1979, 1984, 1985) systems thinking (e.g., Bronfenbrenner, 1992; Mingers, 1997) and complexity theories (e.g., Kaplan and Garner, 2017).

However, while these models begin to reconceptualize the complex and reciprocal interactions among the variables that may influence learning, they do not consider the active and reciprocal role of the students themselves in shaping the learning environment and their own learning outcomes. The reciprocal dynamics between the learner's prior knowledge, introspective capabilities, self-concept, and the socio-ecological educational setting are only hinted at. As such, these models do not capture the fuller reciprocal dynamics the science of learning has made an increasingly clear argument for (Myran and Sutherland, 2019b). This may

suggest an undercurrent within the field of a nascent intuition about the reciprocal dynamics among all the actors within the school setting, including the students themselves. However, without a robust discourse on these disjunctures, and an appropriate reconceptualization of instructional leadership and other leadership models, this emergent focus on learning may not realize its potential.

The legacy of scientific management: incongruent theories of action

As we can see in the discussion of the previous themes, even with this ongoing transformation there remain notable concerns about these disjunctures. The aspiration of instructional leadership, the sedimented histories of the field that tend to frame learning as linear, recursive, and the act of delivering and receiving content, and the science of learning represent incongruent theories of action and design philosophies. Donaldson and Allen-Handy (2020) described these managerial traditions as drawing from a *transfer/acquisition metaphor*, conceptualizing education as something that is *provided*, or *delivered*.

Despite a long history of school reform efforts in the United States and beyond, sustained implementation of research-based practices with fidelity is limited and sporadic (Brazier and Bauer, 2013; Cuban, 1984, 1990, 2013; Deschenes et al., 2001; Tyack and Cuban, 1995). Instructional leadership and leadership content knowledge models have not escaped this phenomenon. I would argue, that an under identified aspect of the lack of fidelity and sustained implementation is the legacy of scientific management and the incongruent theories of action between this foundation and the science of learning.

Discussion and implications

Taken together, I agree with Brazier and Bauer's (2013) view that educational leadership's shift in conceptualizing how students internalize content to how they think and engage with others is "seismic". The question remains however, how we make this seismic shift. The research on instructional leadership and leadership content knowledge offers both optimism and caution. Some have framed leadership as second only to classroom teaching as an influence on student learning (Leithwood et al., 2004; Leithwood and Seashore-Louis, 2011; Louis et al., 2010) and that this influence is achieved through leaderships' impact on the school organization and culture and teachers' classroom practices (Day et al., 2016; Witziers et al., 2003). Others have expressed caution, voicing concerns about the lack of precise definitions of instructional leadership (McGhee and Lew, 2007; Brazier and Bauer, 2013; Neumerski, 2013), which in turn make its implementation unobtainable (Brazier and Bauer, 2013; Rigby, 2014; Salo et al., 2015), and the contradictory nature of school management and instructional leadership responsibilities (McGhee and Lew, 2007).

Given the exploration of these interrelated themes, I suggest that what is missing is a grounding in the science of learning, which would help to replace assumptions rooted in theories of action more closely related to efficiency, uniformity, and productivity than theories of learning. On the one hand, it is difficult to argue against the appeal of efficiency and productivity; given limited financial, time, and human resources, why wouldn't educators be drawn to the appeal of relatively simple linear, hierarchical, cause and effect solutions. However, as a primary theory of action these managerial norms rest on shaky ground, potentially misrepresenting and oversimplifying the nature of learning itself, and therefore misguiding what forms of instruction and organizational support best facilitate substantive, lasting, and transferable learning. In this way, one of the implications of this analysis is that beyond the co-defined construct of instructional leadership, content knowledge and pedagogy, the field needs to continue following this evolution of learner and student-centered constructs and reframe instructional leadership around the science of learning.

I make the argument that what is needed is a reconceptualization of instructional leadership through a science of learning lens that disrupts the self-replicating incongruent theories of action outlined in Fig. 1. By introducing the science of learning as the foundational theory of action, the emergent focus on learning present in the literature to date can overcome the ambiguity, folk epistemologies, and social consensus, that may serve to constrain the emergent learning turn evident in the literature. With the introduction of this element to the model, the legacy of scientific management is disrupted, ambiguity and folk epistemologies are replaced with empirical evidence about learning, which in turn help to reveal the hidden, obscured, or taken for granted scientific management influenced perspectives on learning; supporting new insights that offers alternative explanations and counter theories that can build toward theories of leadership congruent with the science of learning (see Fig. 2).

Navigating the learning turn

The *learning turn* discussed here invites us to engage in discourse and action that is more substantively grounded in the science of learning. Such a reconceptualization of instructional leadership builds on the nascent focus on learning and offers a means to "escape the gravitational pull of the recalcitrant effects of scientific management in its past and current forms" (Myran and Sutherland, 2019b).

However, the sedimented histories and habituated norms established prior to the systematic study of learning presents a notable barrier. There is something seductive about hierarchical solutions, they offer uniform strategies that can be easily written into policy, implemented, overseen, and managed. However, given the complex nature of learning, such *solutionism* (Morozov, 2013) is incongruent with what we have known about the complex, reciprocal, dynamic and social-cultural nature of learning since the cognitive revolution some 70 years ago. Because our field does not offer any kind of theoretical or working definition of learning; often

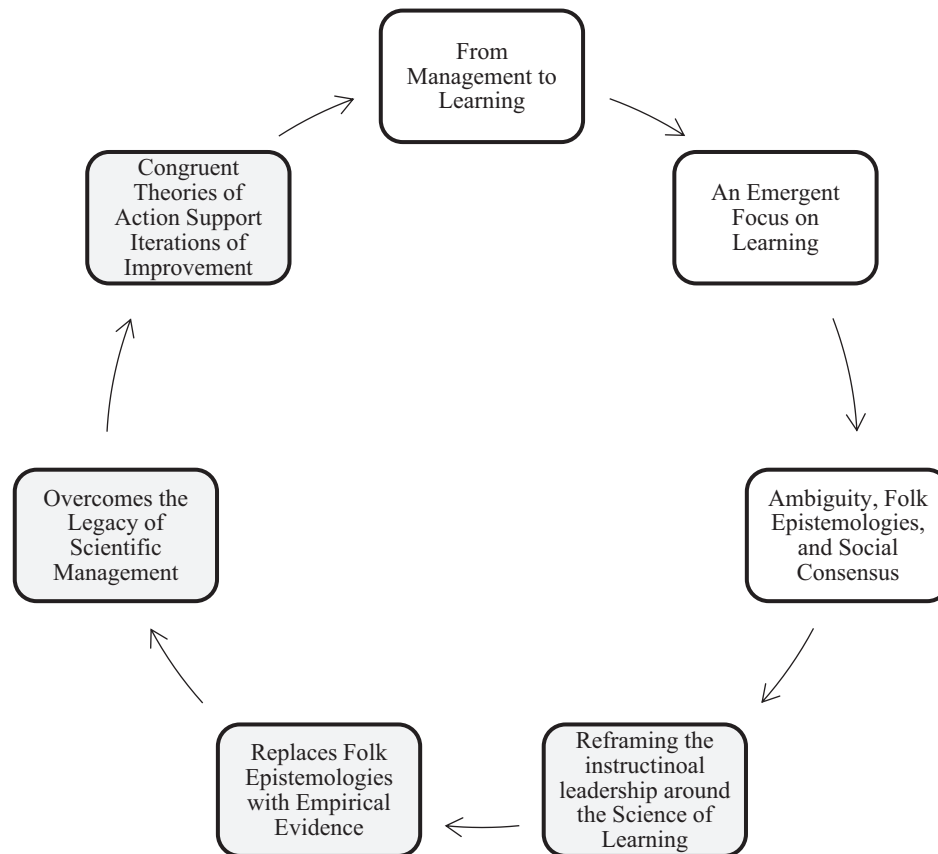


Fig. 2 Congruent theories of action which support iterations of improvement.

substituting performance as a proxy for learning, our ability to continue refining and strengthening the potential of instructional leadership is significantly limited, if not unattainable. It may be that we've reach the practical and functional limitations of a field that has yet to define its primary objective. In this respect, without reframing instructional leadership around the science of learning, the learning turn may merely leave us right where we've been for the past 100 years, attempting to cause and predict something that defies simple linear, hierarchical, cause and effect conceptualizations.

References

- AERA, 2009. Standards for reporting on humanities-oriented research in AERA publications. *Educ. Res.* 38, 481–486.
- Alexander, P.A., Schallert, D.L., Reynolds, R.E., 2009. What is learning anyway? A topographical perspective considered. *Educ. Psychol.* 44, 176–192.
- Anderson, G.L., 2001. Disciplining leaders: a critical discourse analysis of the ISLLC national examination and performance standards in educational administration. *Int. J. Leader. Educ.* 4, 199–216.
- Argyris, C., Schon, D.A., 1974. *Theory in Practice: Increasing Professional Effectiveness*. Jossey-bass.
- Au, W., 2011. Teaching under the new Taylorism: high-stakes testing and the standardization of the 21st century curriculum. *J. Curric. Stud.* 43, 25–45.
- Bandura, A., 1978. The self system in reciprocal determinism. *Am. Psychol.* 33, 344.
- Bandura, A., 1979. Self-referent Mechanisms in Social Learning Theory.
- Bandura, A., 1984. Representing Personal Determinants in Causal Structures.
- Bandura, A., 1985. Model of Causality in Social Learning Theory. *Cognition and Psychotherapy*. Springer.
- Bandura, A., 1989. Human agency in social cognitive theory. *Am. Psychol.* 44, 1175.
- Bell, L., Bolam, R., Cubillo, L., 2003. A Systematic Review of the Impact of School Headteachers and Principals on Student Outcomes. EPPI-Centre, Social Science Research Unit, Institute of Education, London.
- Bossert, S.T., Dwyer, D.C., Rowan, B., Lee, G.V., 1982. The instructional management role of the principal. *Educ. Adm. Q.* 18, 34–64.
- Brazer, S.D., Bauer, S.C., 2013. Preparing instructional leaders: a model. *Educ. Adm. Q.* 49, 645–684.
- Bridges, E.M., 19. Instructional leadership: a concept re-examined. *J. Educ. Adm.* 5 (2), 136–147.
- Bronfenbrenner, U., 1992. *Ecological Systems Theory*. Jessica Kingsley Publishers.
- Brown, S.I., Walter, M.I., 2012. *The Art of Problem Posing*. Routledge.
- Callahan, R.E., 1962. *Education and the Cult of Efficiency*. University of Chicago Press.
- Carver, C.L., 2012. Developing leadership content knowledge during school leader preparation. *Int. J. Educ. Leadersh. Preparation* 7, n3.
- Cuban, L., 1984. *How Teachers Taught: Constancy and Change in American Classrooms, 1890-1980*. Research on Teaching Monograph Series, ERIC.
- Cuban, L., 1988. *Managerial Imperative and the Practice of Leadership in Schools*. the, Suny Press.

- Cuban, L., 1990. Reforming again, again, and again. *Educ. Res.* 19, 3–13.
- Cuban, L., 1998. How schools change reforms: redefining reform success and failure. *Teach. Coll. Rec.* 99, 453–477.
- Cuban, L., 2013. Why so many structural changes in schools and so little reform in teaching practice? *J. Educ. Adm.* 51, 109–125.
- Dantley, M.E., 1990. The ineffectiveness of effective schools leadership: an analysis of the effective schools movement from a critical perspective. *J. Negro Educ.* 59, 585–598.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52, 221–258.
- Deschenes, S., Cuban, L., Tyack, D., 2001. Mismatch: historical perspectives on schools and students who don't fit them. *Teach. Coll. Rec.* 103, 525–547.
- Donaldson, J.P., Allen-Handy, A., 2020. The nature and power of conceptualizations of learning. *Educ. Psychol. Rev.* 32, 545–570.
- Dumay, X., Boonen, T., van Damme, J., 2013. Principal leadership long-term indirect effects on learning growth in mathematics. *Elem. Sch. J.* 114, 225–251.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader* 37, 15–24.
- English, F.W., 2000. Psssst! what does one call a set of non-empirical beliefs required to be accepted on faith and enforced by authority? [Answer: a religion, aka the ISLLC standards]. *Int. J. Leader. Educ.* 3, 159–167.
- English, F.W., 2012. Bourdieu's misrecognition: why educational leadership standards will not reform schools or leadership. *J. Educ. Adm. Hist.* 44, 155–170.
- Freire, P., 2013. *Pedagogy of the Oppressed*. Routledge.
- Freire, P., 2021. *Education for Critical Consciousness*. Bloomsbury Publishing.
- Gerken, M., 2020. Truth-sensitivity and folk epistemology. *Philos. Phenomenol. Res.* 100, 3–25.
- Gioia, D.A., Pitre, E., 1990. Multiparadigm perspectives on theory building. *Acad. Manag. Rev.* 15, 584–602.
- Glickman, C.D., Gordon, S.P., Ross-Gordon, J.M., 2001. *Supervision and Instructional Leadership: A Developmental Approach*. ERIC.
- Graczewski, C., Knudson, J., Holtzman, D.J., 2009. Instructional leadership in practice: what does it look like, and what influence does it have? *J. Educ. Stud. Placed A. T. Risk* 14, 72–96.
- Greenfield, T., Ribbins, P., 2005. *Greenfield on Educational Administration: Towards a Humane Craft*. Routledge.
- Grissom, J.A., Loeb, S., 2011. Triangulating principal effectiveness: how perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *Am. Educ. Res. J.* 48, 1091–1123.
- Gronn, P.C., 1982. Neo-Taylorism in educational administration? *Educ. Adm. Q.* 18, 17–35.
- Gunter, H., 1997. *Rethinking Education: The Consequences of Jurassic Management*. Burns & Oates.
- Gunter, H., 2009. Critical approaches to leadership in education. *J. Educ. Enq.* 2.
- Hallinger, P., 2010. *Developing Instructional Leadership. Developing Successful Leadership*. Springer.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49, 125–142.
- Hallinger, P., Gümüş, S., Bellibaş, M.Ş., 2020. 'Are principals instructional leaders yet?' A science map of the knowledge base on instructional leadership, 1940–2018. *Sci-entometrics* 122, 1629–1650.
- Hallinger, P., Heck, R.H., 1996. Reassessing the principal's role in school effectiveness: a review of empirical research, 1980–1995. *Educ. Adm. Q.* 32, 5–44.
- Hallinger, P., Heck, R.H., 2010. Leadership for learning: does collaborative leadership make a difference in school improvement? *Educ. Manag. Adm. Leader* 38, 654–678.
- Hallinger, P., Heck, R.H., 2011. Conceptual and methodological issues in studying school leadership effects as a reciprocal process. *Sch. Effect. Sch. Improv.* 22, 149–173.
- Halpin, A.W., 1958. *Administrative Theory in Education*. Midwest Administration Center, University of Chicago.
- Hardy-Vallée, B., Dubreuil, B., 2010. Folk epistemology as normative social cognition. *Rev. Philos. Psychol.* 1, 483–498.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86, 531–569.
- Hornig, E., Loeb, S., 2010. New thinking about instructional leadership. *Phi Delta Kappan* 92, 66–69.
- Hughes, M., 2004. Theory and practice in educational management. *Educ. Manag.: Major Themes in Education* 2, 117–160.
- Jackson, A.Y., Mazzei, L., 2011. *Thinking with Theory in Qualitative Research: Viewing Data across Multiple Perspectives*. Routledge.
- Kaplan, A., Garner, J.K., 2017. A complex dynamic systems perspective on identity and its development: the dynamic systems model of role identity. *Dev. Psychol.* 53, 2036.
- Leithwood, K., Seashore, K., Anderson, S., Wahlstrom, K., 2004. *Review of Research: How Leadership Influences Student Learning*.
- Leithwood, K., SEASHORE-Louis, K., 2011. *Linking Leadership to Student Learning*. Jossey-Bass, An Imprint of Wiley.
- Lewis, M.W., Grimes, A.I., 1999. Metatriangulation: building theory from multiple paradigms. *Acad. Manag. Rev.* 24, 672–690.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., Anderson, S.E., Michlin, M., Mascall, B., 2010. *Learning from Leadership: Investigating the Links to Improved Student Learning*. Wallace Foundation, New York, NY.
- Lynham, S.A., 2002. The general method of theory-building research in applied disciplines. *Adv. Develop. Hum. Resour.* 4, 221–241.
- McGhee, M.W., Lew, C., 2007. Leadership and writing: how principals' knowledge, beliefs, and interventions affect writing instruction in elementary and secondary schools. *Educ. Adm. Q.* 43, 358–380.
- Mingers, J., 1997. Systems typologies in the light of autopoiesis: a reconceptualization of Boulding's hierarchy, and a typology of self-referential systems. *Syst. Res. Behav. Sci.: The Official Journal of the International Federation for Systems Research* 14, 303–313.
- Mintrop, H., 2004. *Schools on Probation: How Accountability Works (And Doesn't Work)*. Teachers College Press.
- Mintrop, H., Trujillo, T., 2005. Corrective action in low performing schools: lessons for NCLB implementation from first-generation accountability systems. *Educ. Policy Anal. Arch./Archivos Analíticos de Políticas Educativas* 13, 1–27.
- Moos, L., Möller, J., Johansson, O., 2004. A Scandinavian perspective on educational leadership. *Educ. Forum* 200–210. Taylor & Francis.
- Morozov, E., 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. Public Affairs.
- Mullen, C., Schunk, D., 2010. A view of professional learning communities through three frames: leadership, organization, and culture. *McGill J. Educ./Revue des sciences de l'éducation de McGill* 45, 185–203.
- Myran, S., Sutherland, I., 2019a. Academic and epistemic drift and the re-production of leadership preparation norms. In: Danzig, A., Black, W. (Eds.), *Who Controls the Preparation of Education Administrators?* Information Age Publishing, Charlotte, NC.
- Myran, S., Sutherland, I., 2019b. Defining learning in educational leadership: reframing the narrative. *Educ. Adm. Q.* 55, 657–696.
- Nelson, B.S., Sassi, A., 2000. Shifting approaches to supervision: the case of mathematics supervision. *Educ. Adm. Q.* 36, 553–584.
- Neumerski, C.M., 2013. Rethinking instructional leadership, a review: what do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educ. Adm. Q.* 49, 310–347.
- O'donnell, R.J., White, G.P., 2005. Within the accountability era: principals' instructional leadership behaviors and student achievement. *NASSP Bull.* 89, 56–71.
- Olsen, H., 2006. *A Review of Literature on Principal Instructional Leadership*.
- Oude Groote Beverborg, A., Slegers, P.J., Moolenaar, N.M., van Veen, K., 2020. Fostering sustained teacher learning: a longitudinal assessment of the influence of vision building and goal interdependence on information sharing. *Sch. Effect. Sch. Improv.* 31, 576–604.
- Printy, S., Otten, L., Krantz-Selleck, K., Usiak, M., 2008. *Leadership Content Knowledge: What Do We Know*. CiteSeer.
- Purkey, S.C., Smith, M.S., 1983. Effective schools: a review. *Elem. Sch. J.* 83, 427–452.
- Quebec Fuentes, S., Jimerson, J.B., 2020. Role enactment and types of feedback: the influence of leadership content knowledge on instructional leadership efforts. *J. Educ. Superv.* 3, 6.
- Rigby, J.G., 2014. Three logics of instructional leadership. *Educ. Adm. Q.* 50, 610–644.
- Robinson, V.M., 2006. Putting education back into educational leadership. *Leading & Managing* 12, 62–75.

- Robinson, V.M., 2010. From instructional leadership to leadership capabilities: empirical findings and methodological challenges. *Leader. Pol. Sch.* 9, 1–26.
- Robinson, V.M., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44, 635–674.
- Salo, P., Nylund, J., Stjernström, E., 2015. On the practice architectures of instructional leadership. *Educ. Manag. Adm. Leader* 43, 490–506.
- Sarton, G., 1929. Science in the renaissance. In: Thompson, J.W., Rowley, G., Schevill, F., Sarton, G. (Eds.), *The Civilisation of the Renaissance*.
- Sawyer, R.K., 2006a. Educating for innovation. *Think. Skills Creativ.* 1, 41–48.
- Sawyer, R.K., 2006b. The new science of learning. *The Cambridge handbook of the learning sciences* 1, 18.
- Shaked, H., 2021. Perceptions of Israeli School Principals Regarding the Knowledge Needed for Instructional Leadership. *Educational Management Administration & Leadership*, 17411432211006092.
- Shen, J., 2020. The theory of bifurcated educational system and its implications for school improvement. *Int. J. Leader. Educ.* 1–21.
- Shen, J., Ma, X., Gao, X., Bierlien Palmer, L., Poppink, S., Burt, W., Leneway, R., Mccrumb, D., Pearson, C., Rainey, M., 2019. Developing and validating an instrument measuring school leadership. *Educ. Stud.* 45, 402–421.
- Shulman, L., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14.
- Shulman, L., 1987. Knowledge and teaching: foundations of the new reform. *Harv. Educ. Rev.* 57, 1–23.
- Soderstrom, N.C., Bjork, R.A., 2015. Learning versus performance: an integrative review. *Perspect. Psychol. Sci.* 10, 176–199.
- Spillane, J.P., Diamond, J.B., Jita, L., 2003. Leading instruction: the distribution of leadership for instruction. *J. Curric. Stud.* 35, 533–543.
- Spillane, J.P., Halverson, R., Diamond, J.B., 2004. Towards a theory of leadership practice: a distributed perspective. *J. Curric. Stud.* 36, 3–34.
- Steele, M.D., Johnson, K.R., Otten, S., Herbel-Eisenmann, B.A., Carver, C.L., 2015. Improving instructional leadership through the development of leadership content knowledge: the case of principal learning in algebra. *J. Res. Leadersh. Educ.* 10, 127–150.
- Stein, M.K., D'amico, L., 2000. How Subjects Matter in School Leadership. annual meeting of the American Educational Research Association, New Orleans.
- Stein, M.K., Nelson, B.S., 2003. Leadership content knowledge. *Educ. Eval. Pol. Anal.* 25, 423–448.
- Terry, L.D., 1998. Administrative Leadership, Neo-Managerialism, and the Public Management Movement. *Public Administration Review*, pp. 194–200.
- Tyack, D.B., 1974. *The One Best System: A History of American Urban Education*. Harvard University Press.
- Tyack, D.B., Cuban, L., 1995. *Tinkering toward Utopia: A Century of Public School Reform*. Harvard University Press.
- Weber, G., 1971. Inner-City Children Can Be Taught to Read: Four Successful Schools. *CBE Occasional Papers*, Number 18.
- Weick, K.E., 1995. *Sensemaking in Organizations*. Sage.
- Witziers, B., Bosker, R.J., Krüger, M.L., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 39, 398–425.

Leadership for education and learning in changing contexts

Rose M. Ylimaki, Education and Leadership, Northern Arizona University, Flagstaff, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	24
Changing student demographics	24
Policy trends	24
Studies of leadership for learning: strengths and limitations for changing contexts	26
Instructional leadership and the role of the principal	26
Transformational leadership	27
Shared or distributed leadership	27
Instructional leadership for academic, social justice, and equity	28
Education theorizing and concepts	29
Implications and conclusions: theorizing leadership for education and learning	30
References	31

Introduction

Across the United States and other nation states (e.g. UK, Canada, Australia, New Zealand), we can observe similar policy trends toward externalized evaluations, curriculum centralization, and evidence-based programs as well as demographic trends toward increasingly diverse student groups due to global population migrations and internal demographic shifts. Indeed, the United States is projected to be a majority minority nation state by 2050 (Hoy and Hoy, 2012). Thus, the educational leadership literature focused on curriculum, instruction, and learning is now complicated with tensions between diversity and commonality amid the Covid 19 pandemic and pressures for equity. The purpose of this chapter is to review extant literature on leadership for learning in relation to education theory and changing contexts. Consonant with this purpose, the chapter is organized into sections that consider changing demographics and policies that impact educational leadership, the strengths, and limitations of extant leadership literature as well as literature on education theorizing to enhance understandings of leadership for education and learning, and implications for leadership practice and preparation in light of changing contexts.

Changing student demographics

Demographics are changing due to global population migrations and internal demographic shifts. Currently, White people constitute the majority of the US population (62%); however, the percentage is expected to fall below 50% by 2050 with Hispanic populations to experience the largest increase (Colby and Ortman, 2015; Hoy and Hoy, 2012). As of the writing of this chapter, the combined nonwhite population—blacks, Hispanics, Asians, persons identifying as multiracial, and other races—comprized the dominant share of the under-age populations (50.1% compared with 49.8% in 2017), with Hispanics accounting for more than a quarter of this youth demographic. Diversity ranges by state, but the trend is spreading across the United States. These population changes are due to global population migrations as well as internal demographic shifts, including a decline in young whites. The decline in the number of white children under age 15 has declined over the 2010–2018 period by 2.2 million, continuing a trend toward diversity already observed in the first decade of the century (Brookings Institute, 2020).

Cultural diversity brings multiple experiences, values, and views of knowledge to educational institutions and the academics within them to schools. Since the 1960s, there has been a significant increase in the number and diversity of immigrants coming to the US, and the number has steadily increased over decades. The U.S. Census Bureau notes that the increasing number of international migrants will make the country “a ‘plurality’ of racial and ethnic groups” (Colby and Ortman, 2015, p. 9). For example, the Hispanic population has grown from 6.5% of the US population in 1980 to 17.6% in 2015 (Pew Research Center, 2015) while the Two or More Races population is projected to grow by 225% between 2014 and 2060 (Colby and Ortman, 2015). At the same time, policies are changing with trends toward centralization and an emphasis on evidence-based programs that “work” for all students (Ylimaki et al., 2021; Ylimaki and Brunderman, 2022).

Policy trends

Historically, in the US, policies under both Republican and Democratic administrations have supported the need for externalized evaluations (standardized assessment data) to hold schools and leaders accountable. Additionally, since 2002, there has been a drastic rise in the call for education research funded by the Office of Educational Research and Improvement (OERI) that addresses causal questions using random assignment designs [prior funding for randomized controlled trials represented 5% of federal funding for education research compared to 75% by 2002 (Morrison, 2012)]. Currently, funding applications for

U.S. Department of Education grants must include research designs that are based upon prior studies with “strong evidence” (explicitly defined by large-scale quantitative studies with randomized controlled trials or quasi-experimental designs). In the international arena, multinational organizations (the Organization for Economic Co-operation and Development (OECD) and the World Bank) have also made evidence-based policymaking a priority, both in their own work as well as for their members (Wiseman, 2010).

Globally, we can also observe a policy convergence toward such compliance requirements, curriculum centralization and rankings for school quality and performance (Pilton, 2009; Steiner-Khamsi, 2006). As Pilton (2009) explained, there are several ways that policy convergence contributes to evidence-based policies, curriculum centralization, and comparisons or rankings of schools, states, and nations. One way is through imposition. Imposition is typically a result of political demand or pressure (e.g. Organization for Economic Cooperation and Development/OECD agenda, World Bank) and requires compliance with international policy or institutional arrangements. There is no legal obligation to participate in the Program for International Student Assessment (PISA) with OECD; however, at OECD urging, member nations comply with evidence-based policies. Further, as Meyer (1977) and other institutional theorists would argue, when OECD identifies nation states with “high performing” schools, others emulate these practices (e.g. centralized curriculum) as they are legitimized. Such evidence-based policy convergence also contributed to research trends in funding applications, particularly research methods that are considered “scientific” and measure what works to improve schools.

New bureaucracies have also emerged in relation to neoliberal policies, all of which demand strong leaders who can mediate among many different (and sometimes conflicting) policies, diverse student needs, and democratic education values. Further, school leaders now must work collaboratively within and between governance systems or levels, balancing evidence-based values, and the humanistic values of education for increasingly diverse students.

In many nations, including the United States, education is primarily the responsibility of individual states; there are a variety of external evaluations (state tests) given in different states and sometimes in different districts. For example, in the state of Arizona, students currently take the AZMerit test of knowledge and skills. There is no longer a graduation exam in Arizona, but some school districts additionally require all students to take the ACT exam. In another US state of South Carolina, beginning in grade 3, students take the State Palmetto Assessment of State Standards (SCPASS) in English/Language Arts, Mathematics, Science and Social Studies. The High School Assessment Program (HSAP) is first administered in grade 10, and South Carolina high school students must pass this test in order to receive a high school diploma. National tests, such as the National Assessment of Educational Progress (NAEP), provide nationally representative achievement information in core subjects. International assessments that are widely taken in many countries (e.g. PISA, TIMSS) have often driven national policy decisions toward an increased use of tests as evidence of school quality or success. Although there is a range in the extent to which the results of these assessments directly affect individual students, the average student scores for schools, districts, states, and nations as a whole are widely accepted indicators of educational quality and have a major impact on educational policymaking (LeTendre et al., 2001).

The degree of equity and equality attained in an educational system is another reason frequently cited for assessing the performance of students and schools through externalized evaluations. By seeing who performs highly and who does not, it is possible for teachers, administrators, and policymakers to determine disparities among individual students, classrooms, and schools (Heilig and Darling-Hammond, 2008; Darling-Hammond, 2010). The logic behind using evidence to demonstrate equality suggests that policymakers can design policies and laws to close achievement gaps. For example, Arizona, like many other states, has a wide achievement gap between students from various social and economic groups. The state test (currently AZMerit) shows that this gap exists and the extent to which it exists in particular content areas. Arizona politicians have used this evidence to reward high performing schools and to punish low performing schools. Analysis of standardized achievement scores clearly falls along socioeconomic lines across the nation (Perry et al., 2010; Baker and Johnston, 2010; Cunningham and Sanzo, 2002) and has contributed to the rise of charter schools, vouchers, and public mechanisms to ensure greater equality among schools. Thus, politicians—with evidence from state tests—have passed legislation and policy targeting underperforming schools, requiring shifts in the curricular and classroom focus in order to close those achievement gaps. Under NCLB, if disaggregated data for each student subgroup (i.e. race, gender, language proficiency, etc.) do not demonstrate adequate yearly progress, then the U.S. Department of Education may withhold federal funding from the state or district within the state (McDermott and Jensen, 2005). In 2015, the reauthorization of NCLB resulted in the Every Student Succeeds Act (ESSA), which returned some control to states to determine their goals. Students are still required to be tested annually in grades 3 through high school, but the federal government removed many of the prescriptive pieces of the legislation (U.S. Department of Education, 2016). However, Arizona (as in many states) has not changed state legislation to reflect the greater flexibility of ESSA. It should also be noted that in Arizona, the state test was designed to have both criterion-references components and others that served as norm-referenced items by testing standards at, above, and below the grade level for which the test was designed.

Historically, federal and many state policy makers have attempted to address inequities through legislation. For example, the Coleman Report from the 1960s suggested that family and peer influences, not school resources, are the most important determinants of student performance (Baker and LeTendre, 2005). The Coleman Report was conducted in the wake of the court case *Brown vs. Board of Education of Topeka* that ruled that racial segregation of children in US public schools was unconstitutional. “Evidence” from the landmark court decision and the latter Coleman Report changed the course of equity-driven policy making in the United States and in many other countries. The Coleman Report was part of President Johnson’s strategy in the Great Society program to level the playing field by identifying evidence of inequality in schooling. Coleman surveyed over 560,000 students across the

US, a methodology partly due to the fact that the system was highly decentralized. And while recent externalized evaluation and curriculum policies have shifted toward centralization (e.g. Common Core curriculum), administration of education in the US remains state based. During the time of the Coleman Report, Johnson and other policymakers thought they could address inequality with additional resources for low-socioeconomic schools as a strategy for improvement. However, while they found unequal resources available according to class and race, Coleman's team found evidence that some equity indicators like resources did not predict achievement. Variations in family background and outside-of-school environment affected achievement at least as much as variations in school resources and quality. In other words, some policy impacts cannot be controlled by policymakers. Closely related, Mexican American parents successfully sued school districts in California (Roberto Alvarez vs. the Board of Trustees of the Lemon Grove School District, 1931, and Mendez vs. [Aguirre et al., 2014](#)), both of which attempted to segregate Spanish-speaking Mexican American students in inferior "Americanization" schools.

The 1968 passage of the Bilingual Education Act (Title VII of the Elementary and Secondary Education Act) provided federal protections for culturally and linguistically diverse students. The Bilingual Education Act built from the *Brown v. Board of Education* case and the civil rights movement and promoted instruction that was both in a student's native language and culturally responsive. The act was far reaching, but it (and related state policies) remains at tension with the need to educate students in a common language and other common understandings and civic dispositions that create a nation state. Thus, today's school leadership teams must balance tensions between recent shifts toward curriculum centralization in the Common Core (and state versions thereof), related externalized evaluation policies designed to measure student performance, pressures to use evidence-based instructional programs and the needs of increasingly culturally diverse student demographics.

US policy makers control school funding; at state and national levels, test results are increasingly tied to funding, with the most recent policy wave beginning with the No Child Left Behind Act and its requirement for schools to make adequate yearly progress. Because the U.S. Department of Education has the responsibility for evaluating schools' and states' performance, it can and does use test performance as evidence and a means to control the curriculum schools adopt, the content teachers teach and other components of schooling ([Wiseman, 2010](#)). Since the No Child Left Behind Act, schools identified as not making Adequate Yearly Progress (AYP) over a series of years have been in danger of losing federal funding, and as such this has become vital policymaking evidence for school systems in every state. Some would argue that policymakers have opted toward evidence-based approaches over which they have control. Thus, leaders must implement policies aimed at commonality and also lead curriculum and instruction for increasingly diverse students in schools.

Studies of leadership for learning: strengths and limitations for changing contexts

In this section, I consider mainstream studies of leadership focused on learning in relation to the changing contexts reviewed above. More specifically, I primarily analyze literature on instructional leadership and leadership for social justice with aims for school and societal change respectively. In so doing, I argue that a philosophical approach grounded in the dignity of practice in addition to an analytical approach from empirical research may support leaders to balance and mediate tensions as they educate for an increasingly diverse humanity.

Mainstream leadership literature frequently either focuses on effective organizational theory-grounded approaches, programs and strategies that contribute indirectly to student learning outcomes (Day, 2005, 2007; [Hallinger, 2005](#); [Leithwood and Jantzi, 2005](#); [Marzano et al., 2005](#)). Or critical theory informed approaches that emphasize equity and social justice through equitable student learning outcomes, an increased social consciousness, culturally responsive practices, and community relationships (e.g.). In the following two sub-sections, I consider strengths and limitations of these leadership strands and suggest a further grounding in education and pedagogy as well as organizational and critical social theories to inform future directions for leadership amid the contemporary context and beyond.

Instructional leadership and the role of the principal

In empirical literature, the term *instructional leadership* emerged from the effective schools' research on "outlier schools," that is, those schools that effectively educated children from low socioeconomic (SES) backgrounds (e.g. [Berman and McLaughlin, 1976](#); [Edmonds, 1979](#); [Hallinger and Murphy, 1986](#); [Louis and Miles, 1990](#)). Following this same tradition, conducted a large-scale survey of principals and other educators and identified four leadership tasks or functions that distinguish high-achieving from low-achieving schools. [Hallinger and Murphy \(1986\)](#) constructed an "instructional" leadership model from earlier effective schools research ([Edmonds, 1979](#)) and then tested that model with mixed methods on so-called outlier schools that exceeded expectations in neighborhoods with few resources. [Edmonds \(1979\)](#) conceived of instructional leadership as a role carried out by the school principal, a definition that was quickly adopted and utilized for successive studies of effective schools research (e.g. [Bossert et al., 1982](#); [Hallinger and Murphy, 1985](#); [Leithwood and Montgomery, 1982](#)).

According to Murphy's findings, effective instructional leaders (a) develop mission and goals; (b) promote quality instruction through supervision; (c) promote a positive learning culture through rewards, high expectations, and professional development; and (d) develop a safe and orderly school environment with appropriate economic and community resources. Murphy drew upon organizational perspectives to consider organizational processes and structures that support classroom instruction. Similarly, [Hallinger \(1984\)](#) identified several "instructional leadership behaviors" that yielded improved student outcomes—namely, setting high expectations and goals, supervising, and evaluating instruction, coordinating the curriculum, and monitoring

student progress. Hallinger and Murphy (1984) did not, however, find any connections between curriculum decision-making structures or learning culture and student achievement. Essentially, early instructional leadership models emerged from an empirical base that quantitatively measured principals' tasks and functions as variables that contributed to student outcomes in effective low-SES schools.

In this literature of the 1980s, it is important to note that relatively little reference was made to teachers, department heads or assistant principals as instructional leaders; there was little emphasis on instructional leadership as a distributed construct or a function to be shared. Instructional leaders were described as strong, directive leaders who had been successful in "turning their schools around" (Edmonds, 1979; Hallinger and Murphy, 1985); relatively few case examples of instructional leaders were conducted in schools with average or high performance in the early studies of the 1970s and 1980s. Drawing on results across this literature as well as their own empirical findings using survey methods, Hallinger and Murphy (1985) identified three dimensions of instructional leadership: (1) defining the school's mission, (2) managing the instructional program, and (3) promoting a positive school learning climate delineated into ten instructional leadership functions in educational organizations, including framing the school goals, supervising and evaluating instruction, coordinating curriculum, monitoring student progress, protecting instructional time, providing professional development, maintaining high visibility, providing incentives for teachers, and providing incentives for learning. Hallinger utilized his research over time to develop a rating scale for instructional leadership.

Shephard (2014) replicated Hallinger's study and identified the additional school factors of teacher commitment, professional involvement, and innovativeness as essential influences on classroom practice (with *practice* defined as instructional processes and student behavior). Similarly, other studies assumed a focus on the principal's leadership practices, including most often supervision and staff development, that contributed to improvements in teacher instruction (e.g. Blasé and Blasé, 1999; Heck et al., 1990; Leithwood and Duke, 1998). These studies also included a focus on motivation variables defined in psychology and learning theories rather than curriculum theory or education theory that might have illuminated the leader's work through sociopolitical and historical lenses essential for understanding diverse experiences (Pinar, 2004). At the same time, while early instructional leadership literature focused on structures and processes that support learning in classrooms, they did not explicitly account for change in ways that transformational leadership theories articulated.

Transformational leadership

In probably the most frequently cited example of transformational leadership, Leithwood (1994) posited a framework for transformational leadership elements that explicitly acknowledged change in educational organizations. He later conducted a meta-analysis (Leithwood and Riehl, 2003) of additional effective leadership studies and a North American research project (Leithwood and Seashore, 2011), indicating that four leadership practices were necessary but not sufficient for effectiveness in any context: setting directions, developing people, redesigning the organization, and managing the instructional program. Leithwood and Seashore (2011) concluded, based upon a mixed methods study with a large North American sample, the effects of leadership influence on student outcomes is indirect. Teaching or instruction and related classroom practices have a direct causal effect. Robinson (2010) conducted a synthesis of literature on instructional leadership and concluded that "while there is considerable evidence about the impact of instructional leadership on student outcomes, there is far less known about the leadership capabilities that are required to confidently engage in the practices involved" (p. 1–2). Her research suggested the importance of three interrelated capabilities: (a) using deep leadership content knowledge to (b) solve complex school-based problems, while (c) building relational trust with staff, parents, and students. Robinson (2010) argued that there is considerable interdependence between these three capabilities. Moreover, in an earlier synthesis of research, Robinson et al. (2008) concluded that the effect of instructional leadership on student outcomes was three to four times as great as that of transformational leadership. Here instructional leadership refers to those sets of leadership practices that involve the planning, evaluation, coordination, and improvement of teaching and learning. Going further with an emphasis on indirect relationships among leadership, instruction, and learning outcomes, Murphy et al. (2006) referred to instructional leadership with these capabilities as learning-centered leadership. Robinson et al.'s secondary analysis identified the relative effects of five sets of leadership dimensions: leading through promoting and participating in teacher learning and development; establishing goals and expectations; planning, coordinating, and evaluating teaching and the curriculum; strategic resourcing and ensuring an orderly and supportive environment. Robinson and colleagues concluded from these two meta-analyses that the more leaders focus their relationships, their work, and their learning on the core business of teaching and learning, the greater their influence on student outcomes (Robinson et al., 2008).

Shared or distributed leadership

Other recent studies have continued a focus on "instructional" leadership beyond the exclusive domain of the principal since teachers and other support professionals play a vital leadership role in the improvement of classroom practices (e.g. Hallinger, 2004; Jackson, 2000; Marks and Printy, 2003). Such a shared or integrated model conceptualizes instructional leadership as a capacity for school improvement in which the principal models appropriate instructional leadership behaviors (Bossert et al., 1982; Hallinger and Murphy, 1984; Purkey and Smith, 1983; Shephard, 2014) and invites teachers and others to join their efforts to change and improve their teaching practices. For example, Jackson (2000) used a mixed-methods approach to examine schools that consistently performed well on various assessment measures; he found that an "interactive leadership model"—whereby principals invited teachers to lead instructional improvement efforts and then worked with them in a shared instructional leadership

capacity to develop instructional innovations—could improve student learning. Like earlier instructional leadership studies that focused on individual principals, Jackson emphasized the distributed leader's role in pedagogical improvements more than curriculum content decisions. Spillane et al. (2001) come closer to an interdisciplinary perspective when they define distributed instructional leadership in terms of the various forms of capital that followers value (e.g. social, cultural, human). Here Spillane et al.'s definition of *human capital* encompasses academic content knowledge and implies the importance of leaders' understandings about students and their backgrounds. In other words, traditional (principal-focused) and distributed instructional leadership models account for instructional leaders' abilities to recognize excellent instruction but not necessarily leaders' abilities to foster rigorous instruction that recognizes diverse students' assets or funds of knowledge or to engage teachers and others in complicated conversations about equitable access to excellent curriculum content and instruction for *all* students. Other scholars (e.g. Kelley and Halverson, 2012; Klar, 2012; Moos et al., 2004; Møller, 2006) have considered instructional leadership as a distributed construct, drawing on Spillane's (2012) research to consider how social networks affect such objects as curriculum and instruction. According to Harris (2007), this view of distributed instructional leadership also rests on an idea of leadership as cognition as well as an idea of leadership as mediating social networks and the capital needed to develop democratic features. At the same time, these distributed leadership approaches to leadership for learning do not always explicitly address social justice. A social justice orientation toward instructional leadership is considered below.

Instructional leadership for academic, social justice, and equity

More recently, a growing number of scholars have expanded instructional leadership to focus on social justice with regard to curriculum and instruction (e.g. Oakes et al., 2000; Scheurich, 1998; Theoharis, 2007; Touchton and Acker-Hocevar, 2001). For instance, Touchton and Acker-Hocevar (2001) examined schools that were restructured to improve academic performance for students from traditionally marginalized groups. In this study, the school was the unit of analysis; however, Touchton and Acker-Hocevar focused on the role of the leader in the development of inclusive decision-making processes and classroom instructional practices. Likewise, Dantley and Tillman (2006) examined school restructuring and policy change efforts aimed at transforming procedures that perpetuate social inequalities and marginalization attributable to race, class, gender, and other markers of otherness.

Focusing more specifically on the leader, Theoharis (2007) studied principals' social justice orientations and argued that leaders must create inclusive instructional programs that account for race, class, gender, disability, sexual orientation, and other historically and currently marginalizing conditions (p. 223). Furthermore, Theoharis argued that to challenge the status quo of currently marginalizing instructional practices, such as segregating special education students, social justice leaders must have ideological clarity, self-awareness, passion, and courage. Frattura and Capper (2007) agree, arguing that the primary leadership characteristic of social justice leaders is that "leaders must believe in their core that students learn best when they are educated in heterogeneous educational settings" (p. 19). These and other social justice leadership studies move the needs of traditionally marginalized groups to the center of instructional leaders' efforts and push instructional leadership to incorporate dispositions and practices associated with critical theory (e.g. ideological clarity, critical self-awareness, aims for radical social change).

Marshall (2004, p. 7) argues that for leaders to develop critical self-awareness and ideological clarity, educational leadership preparation programs must have "more skills beyond scientific management or quick fixes" (p. 43) to address complex educational dilemmas steeped in an array of social, cultural, and political contexts. More specifically, McKenzie et al. (2008) posit three goals for educational leaders who seek educational equity and excellence: (a) They must believe that high test scores matter and raise the academic achievement for *all* students in the school, (b) they must prepare their students to live as critical citizens in society, and (c) they must recognize that both of these goals can be achieved only when leaders assign students to inclusive, heterogeneous classrooms that provide all students with access to a rich and engaging curriculum. According to McKenzie and colleagues, a critical consciousness about power, privilege, and social inequities motivates leaders to ensure that schools are safe places for all children and that academic achievement is explicitly linked with activism. To date, the literature on social justice and social equity leadership has primarily documented how principals have raised the achievement of typically marginalized students and, in some cases, done so in inclusive ways (e.g. Capper and Young, 2014; Dantley and Tillman, 2006; Theoharis, 2007; Touchton and Acker-Hocevar, 2001; Moral et al., 2020). As a result, instructional leadership preparation has expanded to include a clear focus on social justice and equity as well as instruction and learning.

Essentially, across this literature, leaders who create equity, which is different from equality, move beyond making sure that all children are treated the same, to ensuring that all students are succeeding academically the same (Larson and Murtdha, 2002). In other words, according to empirical research inspired by critical education theorizing, school leaders with this orientation are activist leaders who work to create justice in schools for all students (Bogotch, 2000; Capper, 2014; Capper et al., 2013; Dantley and Tillman, 2006; Jean-Marie, 2008; Marshall and Oliva, 2012; Shields, 2004). In contrast to earlier instructional leadership focused on planning, evaluation, coordination, and improvement of teaching and learning for individual student achievement and school improvement, this literature strand focuses on equity and social justice through curriculum and instruction. Such aims of social change require leaders to focus attention beyond schools and classrooms to community in ways that support students with culturally diverse backgrounds.

A number of scholars have recently developed culturally responsive leadership approaches to support social change aims (e.g. Horsford et al., 2011; Johnson, 2006, 2014; Scanlan and López, 2014). Johnson (2006) was one of the first to extend the asset-based approach of culturally responsive pedagogy to leadership. Johnson (2006) defined culturally responsive leadership as leadership

that involves philosophies, practices, and policies that create inclusive schooling environments for students and families from ethnically and culturally diverse backgrounds. Common practices include (1) emphasizing high expectations for student achievement; (2) incorporating the history, values, and cultural knowledge of students' home communities in the school curriculum; (3) working to develop a critical consciousness among both students and faculty to challenge inequities in the larger society; and (4) creating organizational structures at the school and district level that empower students and parents from diverse racial and ethnic communities (Johnson, 2007). Other researchers (e.g. Scheurich, 1998; Johnson, 2006, 2007). Riehl (2009) identified similar characteristics and practices of culturally responsive leadership.

Further, Horsford et al. (2011) incorporated antiracist pedagogy and leadership in diverse contexts to create a framework for culturally relevant leadership, including four dimensions: (1) knowledge of the political context; (2) inclusion of a culturally relevant and antiracist pedagogical approach; (3) a personal knowledge of cultural proficiency and challenges to it; and (4) the professional duty to work for educational equity. As Johnson (2014) and others have noted, culturally responsive leadership often overlaps with leadership for social justice approaches, a term that has been prevalent in the US educational leadership literature and focuses on improving the educational experiences and outcomes for all students, particularly those who have been traditionally marginalized in schools. Most recently, Johnson (2014) applied culturally responsive leadership to practices that bridge school and community concerns, advocate for cultural recognition and revitalization, and position educational leaders as advocates for racial equity and community development in diverse neighborhoods (p. 7).

Scanlan and López (2014) agree with the need to bridge school and community concerns and add explicit attention to cultural and linguistic diversity. Scanlan and López draw on culturally responsive leadership practices that reduce marginalization and successfully educate what they term the new mainstream of students. Based upon their analysis of the literature, Scanlan and López identified three essential dimensions to effectively educate culturally and linguistically diverse students, including promoting socio-cultural integration, cultivating language proficiency, and ensuring academic achievement. They proposed a theory of action that school leadership most effectively creates the learning architecture for successfully educating culturally and linguistically diverse students through an integrated service delivery.

In a recent review of literature on equity-oriented leadership, Leithwood (2021) identified practices and dispositions of equity-oriented leaders that contribute to inclusive, socially just educational experiences for diverse students in their schools. Drawing on his synthesis of literature on social justice, equity, and leadership, Leithwood (2021) identified common practices and dispositions of equity-oriented leaders. Findings indicated that successful equity-oriented leaders stress the importance of authentic school, family and community partnerships, recruit and retain a teaching staff reflecting the diversity of the school's students and families and develop a proactive disposition within the school. Within the classroom, leaders help teachers advance equity when they encourage and support ambitious forms of instruction and implement a curriculum perceived by students to be meaningful (e.g. culturally responsive). Importantly, Leithwood (2001) noted that the conception of leadership identified across the literature is a distributed one.

While culturally responsive and social justice-oriented leadership has contributed to the knowledge base about leadership for learning as well as social equity, like the instructional leadership studies, these studies do not explicitly consider education theory concepts (Dewey, 1887, 1916) like growth, pedagogical relations, experience, and the environment. Further, in the first strand, there is a preponderance of quantitative or mixed methods, and in the second strand aimed at equity, there is a preponderance of qualitative studies (Briggs et al., 2012).

Education theorizing and concepts

In the late 1800s and early 1900s, during another time of demographic changes due to immigration changes and industrialization, John Dewey formulated his education theorizing, considering both individual and social growth as well as the tensions between these. Dewey's education theorizing drew heavily on a continental tradition with its emphasis on a dialectic or tensions between competing forces but an ultimate aim of education for humanity (1877, 1938). For Dewey, education required a humanistic approach that supported both individual development and social aims. In *My Pedagogic Creed*, Dewey (1897/1972) told us:

I believe that education is a regulation of the process of coming to share in the social consciousness; and that the adjustment of individual activity on the basis of this social consciousness is the only sure method of social reconstruction I believe that in the ideal school we have the reconciliation of the individualistic and the institutional ideals (p. 93).

Here Dewey's (1897/1972) education theorizing aimed at unifying the individual journey and the societal goal of education: "education must be conceived as a continuing reconstruction of experience ... the process and the goal of education are one and the same thing" (p. 91). As he argued in *My Pedagogic Creed*,

I believe that the only true education comes through the stimulation of the child's powers by the demands of the social situations in which he finds himself. Through these demands he is stimulated to act as a member of a unity, to emerge from his original narrowness of action and feeling and to conceive of himself from the standpoint of the welfare of the group to which he belongs. Through the responses which others make to his own activities

he comes to know what these mean in social terms. The value which they have is reflected back into them. For instance, through the response which is made to the child's instinctive babbling the child comes to know what those babblings mean; they are transformed into articulate language and thus the child is introduced into the consolidated wealth of ideas and emotions which are now summed up in language.

Dewey (1916, p. 493) summarized his views about the child and the curriculum in *Democracy and Education*:

In sum, I believe that the individual who is to be educated is a social individual and that society is an organic union of individuals. If we eliminate the social factor from the child we are left only with an abstraction; if we eliminate the individual factor from society, we are left only with an inert and lifeless mass.

Dewey went on to argue, "We never educate directly, but indirectly by means of the environment." The influence of the environment operates through the habits of language, manners, good taste, and esthetic appreciation. These define the demands of social situations or the environment that contextualizes children's experiences in the world (Henderson et al., 2021). As Dewey put it, "In order to know what a power really is, we must know what its end, use, or function is; and this we cannot know save as we conceive of the individual as active in social relationships. But, on the other hand, the only possible adjustment which we can give to the child under existing conditions, is that which arises through putting him in complete possession of all his powers" (Dewey, 1897, p. 6). For Dewey, the individual followed his/her own journey of growth yet also connected and developed in social relationships as part of a democratic way of living.

Importantly, Dewey also reminded us that education must recognize that the future is an open question. He continued later in *Democracy and Education* (Dewey (1916).

With the advent of democracy and modern industrial conditions, it is impossible to foretell definitely just what civilization will be twenty years from now. Hence it is impossible to prepare the child for any precise set of conditions. To prepare him for the future life means to give him command of himself; it means so to train him that he will have the full and ready use of all his capacities; that his eye and ear and hand may be tools ready to command, that his judgment may be capable of grasping the conditions under which it has to work, and the executive forces be trained to act economically and efficiently.

In other words, Dewey (1897, 1916, 1938) encouraged continuous growth throughout life, and to avoid following a path to a point where learning could stop. In *Democracy and Education*, Dewey (1916) argued, in particular, that an education which only emphasizes the achievement of "external aims" (e.g. evidence from standardized test scores, grades, school letter grades, etc.) hinders students' capacity for continuous growth and leads students toward viewing learning as an overly burdensome activity which they should seek to end as quickly as possible. Rather, in a Deweyan notion of democratic society, there is a reliance on ordinary individuals increasing their wisdom through experience. Reality for Dewey is constantly changing and there is no one with a monopoly on the knowledge of reality which means that democracy requires that everyone continuously grow and adapt to changing conditions.

It should be noted that Dewey (1916) did not ignore the significance, nature and utility of subject areas, or outcomes. In fact, chapters of *Democracy and Education* are dedicated to deliberation on "geography and history" and "physical and social studies" as well as "play and work" and "vocational aspects of education." Dewey communicates the aims of education for social direction, preparation, formation, and reconstruction. Consequently, education is "based upon a consideration of what is already going on; upon the resources and difficulties of existing conditions," all of which require reflection and development or growth (democracy) (p. 493).

In Dewey's (e.g. 1897, 1916, 1938) writings, we gain a philosophical approach to education that requires individual growth, reflection, and social development (democracy). We see a way to hold the individual's needs at tension with needs for a common societal aim toward democracy. Dewey's education theorizing was developed during an earlier time when the US was developing as a nation state out of immigration amid wars and a pandemic. Here individual growth includes academic learning as part of an individual journey of understanding in relation to needs of society. During his career, Dewey opened a lab school aimed at the individual and society, but he did not explicitly consider the role of the organization or leadership support for instruction and pedagogy in ways that leadership scholars have emphasized. Dewey's work was pragmatic in nature; however, he did not conduct empirical studies drawing on his education theorizing.

Implications and conclusions: theorizing leadership for education and learning

Drawing on the extant empirical leadership literature aimed at effective instruction, learning, and schools, and Dewey's education theorizing aimed at the individual in relation to society, the purpose of this chapter was to consider leadership for

learning in relation to changing demographic, education, and policy contexts. By 2015, the US is projected to become a majority minority nation state due to immigration and internal demographic shifts. Such increasing cultural diversity is, in some ways, at tension with policy trends toward evidence-based programs, more centralized curriculum, and externalized evaluations. Today's leaders must mediate and balance these tensions in order to support education and learning for individuals and a democratic society.

Across the extant leadership literature, we can find two dominant strands with some distinctive aims. First, literature on instructional leadership focused on the principal and his/her support for instructional improvements aimed at student learning outcomes. Here, leadership support tends to be organizational, structural, or programmatic with learning outcomes from these measured with empirical research and externalized evaluations. This pattern makes sense because much of the literature on "instructional" leadership or leadership for learning is frequently grounded in organizational theory and policy trends support evidence-based programs. We can observe similar trends in other dominant leadership approaches (e.g. transformational leadership) with the additional focus on change.

In the other literature strand reviewed here, the aim is much more clearly grounded in critical theories with a focus on societal change and equity, and this literature strand provides important understandings about leadership for learning and equity. In John Dewey's (e.g., 1897, 1916, 1938) education theorizing, we can see an example of how to balance tensions between diversity and commonality in a dialectic approach; we recall from Dewey about longstanding conceptual understandings of education, growth and pedagogy that link the development of individuals and society. At the same times, Dewey's pedagogy does not include explicit attention to culturally responsive pedagogy or leadership.

Future *educational* leadership research and preparation grounded in education theorizing with understandings about tensions in a dialectic and implications for pedagogy as well as leadership grounded in organizational theory and change may offer a new framework for study and practice (Uljen and Ylimaki, 2017; Ylimaki and Uljen, 2017). In my view, implications for leadership preparation involve a greater emphasis on curriculum, philosophy, and foundations of education, including background in culture, as well as empirical evidence (Ylimaki, 2012; Ylimaki and Uljen, 2017). To date, much of the educational leadership literature is empirical in nature. Knowledge from empirical study has made significant contributions, but empirical study alone may not be insufficient to guide educational leaders in an increasingly complex policy and demographic context. Today's leaders who serve increasingly culturally diverse students at tension with policy requirements for common curriculum standards and externalized evaluations amid other complexities would also benefit from study of dialectics and education concepts from Dewey and the roots of his work as well as theories of organizations and culture, recognizing the future as an open question (Kraus and Ylimaki, *in press*).

References

- Aguirre, F.P., Bowman, K.L., Mendez, G., Mendez, S., Robbie, S., Strum, P., 2014. Mendez v. Westminster: A living history. *Mich. St. L. Rev.* p. 401.
- Baker, M., Johnston, P., 2010. The impact of socioeconomic status on high stakes testing reexamined. *J. Instr. Psychol.* 37 (3).
- Baker, D., LeTendre, G.K., 2005. *National Differences, Global Similarities: World Culture and the Future of Schooling*. Stanford University Press.
- Berman, P., McLaughlin, M.W., 1976. Implementation of educational innovation. In: *The Educational Forum*, vol. 40. Taylor & Francis Group, pp. 345–370.
- Blasé, J., Blasé, J., 1999. *Handbook of Instructional Leadership: How Successful Principals Promote Teaching and Learning*. Corwin Press, Thousand Oaks, CA.
- Bogotch, I.E., 2000. *Educational Leadership and Social Justice: Theory Into Practice*.
- Bossert, S.T., Dwyer, D.C., Rowan, B., Lee, G.V., 1982. The instructional management role of the principal. *Educ. Adm. Q.* 18 (3), 34–64.
- Briggs, A.R., Morrison, M., Coleman, M., 2012. *Research Methods in Educational Leadership and Management*. Sage Publications.
- Brookings Institute, 2020. Annual Report. <https://www.brookings.edu/wp-content/uploads/2020/12/2020-annual-report.pdf>. (Accessed 3 February 2022).
- Capper, C.A., Young, M.D., 2014. Ironies and limitations of educational leadership for social justice: a call to social justice educators. *Theor. Pract.* 53 (2), 158–164.
- Capper, C.A., Green, T.L., Tillman, L., Scheurich, J.J., 2013. Organizational theories and the development of leadership capacity for integrated, socially just schools. In: *Handbook of Educational Leadership for Equity and Diversity*, pp. 62–82.
- Colby, S.L., Ortman, J.M., 2015. *Projections of the Size and Composition of the US Population: 2014 to 2060, Current Population Reports, P25-1143*. US Census Bureau, Washington, DC.
- Cunningham, W.G., Sanzo, T.D., 2002. Is high-stakes testing harming lower socioeconomic status schools? *NASSP Bull.* 86 (631), 62–75.
- Dantley, M., Tillman, L., 2006. *Social Justice and Moral Transformative Leadership*. Pearson Education, Boston, MA.
- Darling-Hammond, L., 2010. Teacher education and the American future. *J. Teach. Educ.* 61 (1–2), 35–47.
- Day, C., 2005. Principals who sustain success: making a difference in schools in challenging circumstances. *Int. J. Leader. Educ.* 8 (4), 273–290.
- Day, C., 2007. Sustaining success in challenging contexts: Leadership in English schools. In: *Successful principal leadership in times of change*. Springer, Dordrecht, pp. 59–70.
- Dewey, J., 1887. Knowledge as idealisation. *Mind* 12 (47), 382–396.
- Dewey, J., 1897/1972. The psychological aspect of the school curriculum. In: Boydston, J.A. (Ed.), In: Bowers, F. (Ed.), *The Early Works of John Dewey 1882–1898*, vol. 5. Southern Illinois University Press, Carbondale, pp. 1895–1898 (Original work published in 1897).
- Dewey, J., 1897. My pedagogic creed (No. 25). *EL Kellogg & Company*.
- Dewey, J., 1916. *Democracy in Education*. The Free Press, New York, NY.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York, NY.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader.* 37, 15–24.
- Frattura, E., Capper, C., 2007. *Leading for Social Justice: Transforming Schools for All Learners*. Corwin Press, Thousand Oaks, CA.
- Hallinger, P., Murphy, J.F., 1984. School leadership studies: where do we go from here? *Am. Second. Educ.* 18–21.
- Hallinger, P., Murphy, J., 1985. Assessing the instructional management behavior of principals. *Elem. Sch. J.* 86 (2), 217–247.
- Hallinger, P., Murphy, J., 1986. The social context of effective schools. *Am. J. Educ.* 94 (3), 328–355.

- Hallinger, P., 2004. Reflections on the Practice of Instructional and Transformational Leadership. Paper presented at the Annual International Congress for School Effectiveness and School Improvement, Nottingham, UK.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leader. Pol. Sch.* 4 (3), 221–239.
- Harris, A., 2007. Distributed leadership: conceptual confusion and empirical reticence. *Int. J. Leader. Educ.* 10 (3), 315–325.
- Heck, R.H., Larsen, T.J., Marcoulides, G.A., 1990. Instructional leadership and school achievement: validation of a causal model. *Educ. Adm. Q.* 26 (2), 94–125.
- Hellig, J.V., Darling-Hammond, L., 2008. Accountability Texas-style: the progress and learning of urban minority students in a high-stakes testing context. *Educ. Eval. Pol. Anal.* 30 (2), 75–110.
- Henderson, J.G., Castner, D.J., Yilmaki, R.M., 2021. Terminology on Growth. A paper presented at workshop one on Growth and Bildung.
- Horsford, S.D., Grosland, T., Gunn, K.M., 2011. Pedagogy of the personal and professional: toward a framework for culturally relevant leadership. *J. Sch. Leader.* 21 (4), 582–606.
- Hoy, A., Hoy, W., 2012. *Instructional leadership: A research-based guide to learning in schools*. Pearson, New York, NY.
- Jackson, D.S., 2000. The school improvement journey: Perspectives on leadership. *Sch. Leadersh. Manag.* 20 (1), 61–78.
- Jean-Marie, G., 2008. Leadership for social justice: an agenda for 21st century schools. In: *The Educational Forum*, vol. 72. Taylor & Francis Group, pp. 340–354.
- Johnson, L., 2006. Making her community a better place to live": Culturally responsive urban school leadership in historical context. *Leadersh. Policy Sch.* 5 (1), 19–36.
- Johnson, L., 2007. Rethinking successful school leadership in challenging US schools: culturally responsive practices in school-community relationships. *Int. Stud. Educ. Adm.* 35 (3).
- Johnson, L., 2014. Culturally responsive leadership for community empowerment. *Multicult. Educ. Rev.* 6 (2), 145–170.
- Kraus, A., Yilmaki, R., (in progress). An historical to continental pedagogics from a North American perspective. *Educ. Philos. Theor.*
- Kelley, C., Halverson, R., 2012. The Comprehensive Assessment of Leadership for Learning: A Next Generation Formative Evaluation and Feedback System. *J. Appl. Res. Child.* 3 (2), 4.
- Klar, H.W., 2012. Fostering distributed instructional leadership: A sociocultural perspective of leadership development in urban high schools. *Leadersh. Policy Sch.* 11 (4), 365–390.
- Larson, C.L., Murtadha, K., 2002. Leadership for social justice. *Yearbk. Natl. Soc. Stud. Educ.* 101 (1), 134–161.
- Leithwood, K., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30 (4), 498–518.
- Leithwood, K., 2001. School leadership in the context of accountability policies. *Int. J. Leadersh. Educ.* 4 (3), 217–235.
- Leithwood, K., 2021. A review of evidence about equitable school leadership. *Educ. Sci* 11 (8), 377.
- Leithwood, K., Duke, D.L., 1998. Mapping the conceptual terrain of leadership: a critical point of departure for cross-cultural studies. *Peabody J. Educ.* 73 (2), 31–50.
- Leithwood, K., Jantzi, D., 2005. Transformational leadership. In: *The Essentials of School Leadership*, pp. 31–43.
- Leithwood, K.A., Montgomery, D.J., 1982. The role of the elementary school principal in program improvement. *Rev. Educ. Res.* 52 (3), 309–339.
- Leithwood, K., Riehl, C., 2003. What do we already know about successful school leadership. In: *annual meeting of the, (Vol. 22)*. American Educational Research Association, Chicago, IL.
- Leithwood, K., Seashore-Louis, K., 2011. *Linking leadership to student learning*. John Wiley & Sons.
- LeTendre, G.K., Baker, D.P., Akiba, M., Goesling, B., Wiseman, A., 2001. Teachers' work: institutional isomorphism and cultural variation in the US, Germany, and Japan. *Educ. Res.* 30 (6), 3–15.
- Louis, D., Miles, M.B., 1990. *Improving the Urban. High School*. Teachers College Press, New York.
- Marks, H., Printy, S., 2003. Principal leadership and school performance: an integration of transformation and instructional leadership. *Educ. Adm. Q.* 4 (4), 293–331.
- Marshall, C., Oliva, M., 2012. Building the capacities of social justice leaders. *Leader. Soc. Justice* 1–15.
- Marshall, C., 2004. Social justice challenges to educational administration: introduction to a special issue. *Educ. Adm. Q.* 40 (1), 3–13.
- Marzano, R.J., Waters, T., McNulty, B.A., 2005. Some theories and theorists on leadership. In: *School Leadership That Works*, pp. 13–27.
- McDermott, K.A., Jensen, L.S., 2005. Dubious sovereignty: federal conditions of aid and the no child left behind act. *Peabody J. Educ.* 80 (2), 39–56.
- McKenzie, K.B., Christman, D., Hernandez, F., Fierro, F., Capper, C., Dantley, D., et al., 2008. From the field: a proposal for educating leaders for social justice. *Educ. Adm. Q.* 44 (1), 111–138.
- Meyer, J.W., 1977. The effects of education as an institution. *Am. J. Sociol.* 83 (1), 55–77.
- Moos, L., Möller, J., Johansson, O., 2004, September. A Scandinavian perspective on educational leadership. In: *The Educational Forum*, Vol. 68, No. 3. Taylor & Francis Group, pp. 200–210.
- Møller, J., 2006. Democratic schooling in Norway: Implications for leadership in practice. *Leadersh. Policy Sch* 5 (1), 53–69.
- Moral, C., Higuera-Rodríguez, L., Martín-Romera, A., Valdivia, E.M., Morales-Ocaña, A., 2020. Effective practices in leadership for social justice. Evolution of successful secondary school principalship in disadvantaged contexts. *Int. J. Leader. Educ.* 23 (2), 107–130.
- Morrison, K.A., 2012. *Free school teaching: A journey into radical progressive education*. SUNY Press.
- Murphy, J., Elliott, S.N., Goldring, E., Porter, A.C., 2006. *Learning-Centered Leadership: A Conceptual Foundation*. Learning Sciences Institute, Vanderbilt University (NJ1).
- Oakes, J., Quartz, K.H., Ryan, S., Lipton, M., 2000. Becoming good American schools: the struggle for civic virtue in education reform. *Phi Delta Kappan* 81 (8), 568–575.
- Perry, T., Moses, R.P., Cortes Jr., E., Delpit, L., Wynne, J.T., 2010. *Quality Education as a Constitutional Right: Creating a Grassroots Movement to Transform Public Schools*. Beacon Press.
- Pew Research Center, 2015. Population Projections. Retrieved at. <https://www.pewresearch.org/topics/population-projections/2015/>.
- Pitton, J., 2009. *Policy Convergence in Education*. Unpublished manuscript. Lehigh University, Bethlehem, PA.
- Pinar, W., 2004. *What is Curriculum Theory?* Routledge, New York, NY.
- Purkey, S.C., Smith, M.S., 1983. Effective schools: a review. *Elem. Sch. J.* 83 (4), 427–452.
- Riehl, C.J., 2009. The principal's role in creating inclusive schools for diverse students: a review of normative, empirical, and critical literature on the practice of educational administration. *J. Educ.* 189 (1–2), 183–197.
- Robinson, V.M., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Robinson, V.M., 2010. From instructional leadership to leadership capabilities: empirical findings and methodological challenges. *Leader. Pol. Sch.* 9 (1), 1–26.
- Scanlan, M., López, F.A., 2014. *Leadership for Culturally and Linguistically Responsive Schools*. Routledge.
- Scheurich, J., 1998. Highly successful and loving, public elementary schools populated mainly by low-SES children of color: core beliefs and cultural characteristics. *Urban Educ.* 33 (4), 451–491.
- Shepherd, R.J., 2014. *Reflections of School Staff on Headteacher Leadership and How It is Practiced in Their State, Boys' Selective School*. Doctoral dissertation, University of Birmingham.
- Shields, C.M., 2004. Dialogic leadership for social justice: overcoming pathologies of silence. *Educ. Adm. Q.* 40 (1), 109–132.
- Spillane, J.P., 2012. *Distributed leadership*, Vol. 4. John Wiley & Sons.
- Spillane, J.P., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- Steiner-Khamisi, G., 2006. The economics of policy borrowing and lending: a study of late adopters. *Oxf. Rev. Educ.* 32 (5), 665–678.
- Theoharis, G., 2007. *The Schools are Children Deserve: Seven Keys to Equity, Social Justice, and School Reform*. Teachers College Press, New York, NY.
- Touchton, D., Acker-Hocevar, M., 2001. *Using a Lens of Social Justice to Reframe Principals' Interviews From High Poverty, Low Performing Schools*.

- Uljen, M., Yilmaki, R.M., 2017. Bridging Educational Leadership, Curriculum Theory and Didaktik. Springer International Publishing.
- Wiseman, A.W., 2010. The uses of evidence for educational policymaking: global contexts and international trends. *Rev. Res. Educ.* 34, 1–24.
- Yilmaki, R.M., Brunderman, L.A., 2022. Evidence-Based School Development in Changing Demographic Contexts. Springer Nature, p. 186.
- Yilmaki, R.M., Uljen, M., 2017. Theorizing educational leadership studies, curriculum, and didaktik: nonaffirmative education theory in bridging disparate fields. *Leader. Pol. Sch.* 16 (2), 175–227.
- Yilmaki, R., Brunderman, L., Moyl, P., 2021. Balancing evidence-based and humanistic education in school development for underperforming Arizona and South Carolina schools. *Leader. Pol. Sch.* 20 (3), 436–456.
- Yilmaki, R.M., 2012. Curriculum leadership in a conservative era. *Educ. Adm. Q.* 48 (2), 304–346.

Future research on educational leadership: theoretical groundings of educational leadership research

Yinying Wang, Georgia State University, Atlanta, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	34
What is theory?	34
What is theory for?	35
Where have we been?	36
Theories about leadership styles	36
Theories in organizational context	38
Sociological, psychological, and socio-psychological theories	39
Theories about social justice in education	39
What does the future hold?	39
Recommendation 1: refining definitions of theories	39
Recommendation 2: pruning theories	40
Conclusion	41
References	41

Introduction

What is theory? What is theory for? What are the theories that make up the current theoretical groundings in educational leadership research? How to contribute to theory development in educational leadership? This chapter aims to answer these questions. To peek into the future of theoretical groundings of educational leadership, we need a grip on the past and present. With an eye toward the future, I first chart advances over the past decades about what theories have been frequently used to frame empirical studies in educational leadership research. After taking stock of many past studies, I provide readers with two recommendations for theory development in educational leadership.

What is theory?

What is theory? This has been a long-standing question in the field of educational leadership. Asking educational leadership researchers, you may receive different answers. Some researchers use many terms interchangeably, including “theory,” “concept,” “construct,” “theoretical,” “conceptual,” “theorize,” and “conceptualize.” Others believe that nuances exist to differentiate those terms. In a recent article on the absence of student-centered learning in educational leadership, the authors stated, “all functions of leadership must stem conceptually, theoretically and practically from our understanding of the principles of learning” (Myran and Sutherland, 2019, p. 661). It seems the authors attached different meanings to “conceptually” and “theoretically.” Yet it is unclear what differentiated the two terms. The definition of theory has been quite confusing for many educational leadership researchers and practitioners.

What “theory” means can vary from one paper to the next (Abend, 2008). To set the stage for a discussion on the future theoretical groundings in this chapter, I introduce two major definitions of theory in educational leadership research. First, a theory can be logical propositions that establish testable hypotheses about the relationships among constructs (Leavitt et al., 2010). In this vein, a theory has two elements: (1) the constructs, “to which a particular meaning has been attached” (Griffiths, 1983, p. 38), as the building blocks of a theory, and (2) the assumptions of construct relationships (Wang, 2018a, b). To understand something, we have to start somewhere by labeling it. Social capital is a construct, meaning the sum of resources an individual or a group can access. A theory of social capital needs to propose some relationships among constructs. Take Burt’s (2005) social capital of structural holes as an example. This proposition established the relationship between social capital and structural holes. To build social capital, according to Burt, it is insufficient to simply build social ties. Rather, it is uniquely advantageous to build a specific type of ties—the ties that function as a bridge connecting holes in network structures. In Fig. 1, both Abby and Bill have four social ties, but they do not have the same amount of social capital. There is no structural hole in Abby’s social network, because everyone in Abby’s social network is connected to one another. By contrast, each tie connected to Bill functions as a bridge to a different pool of resources. The ties between Bill and Individual 2, 5, 6, and 8 enable Bill to access to resources in Group A, B, C, and D, respectively. Using the social capital of structural holes, researchers can conduct empirical research to expand existing theory—such as by adding mediators or moderators—or present new theories (Leavitt et al., 2010). In this sense, much of the research in educational leadership is “better labeled theoretically oriented than theory” (Willower, 1975, p. 4). This definition of theory is quite stringent.

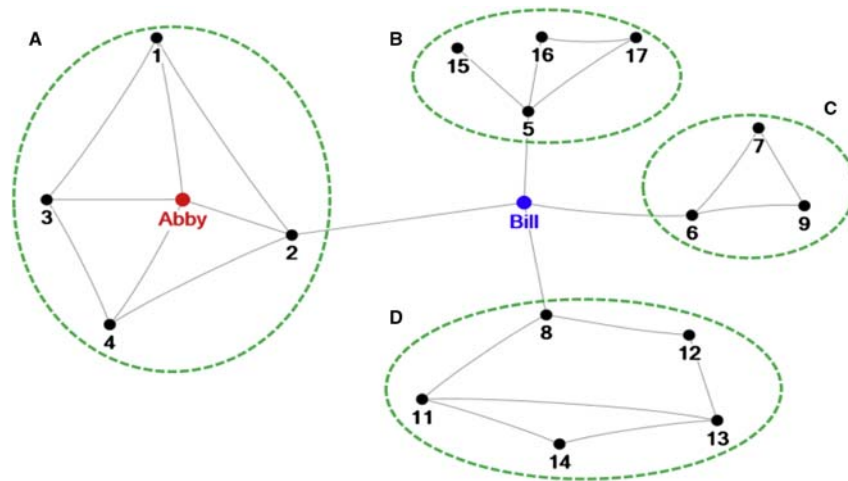


Fig. 1 Do Abby and Bill have the same amount of social capital?

A theory can also be an explanation of a particular social phenomenon. When we say that we have a theory about leadership, what we mean is that we can offer an explanation of it. This explanation should identify a number of factors or conditions (Abend, 2008). Many theories about leadership styles fall into this category. Leadership can be explained by leader behaviors. A pattern of leader behaviors is then considered as a leadership style (Wang, 2019). For example, transformational leadership presents an explanation of leadership, and this explanation has four dimensions: idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Bass, 1998). Ethical leadership can be explained by leader behaviors in seven dimensions: fairness, integrity, ethical guidance, people orientation, power sharing, role clarification, and concern for sustainability (Kalshoven et al., 2011). Culturally responsive school leadership presents another explanation of leadership, which encompasses four dimensions: critically self-reflecting on leadership behavior, developing culturally responsive teachers, promoting culturally inclusive school environment, and engaging students, parents, and indigenous contexts (Khalifa et al., 2016). Another example is organizational justice, which is closely related to organizational outcomes (e.g., employee job satisfaction, organizational commitment, organizational citizenship behavior, and organizational performance). Organizational justice has four dimensions: distributive, procedural, interpersonal, and informational justice (Colquitt et al., 2001). Building on these theories, researchers can use empirical research to evaluate and integrate the theories into other theories, thereby developing new theories. For example, shared instructional leadership was proposed after integrating instructional leadership and transformational leadership (Marks and Printy, 2003). Latino Critical Theory (LatCrit) was proposed after applying critical race theory and adding dimensions of language, immigration, ethnicity, culture, identity, phenotype, and sexuality to explain Latinas/Latinos' multidimensional identities and address the intersectionality of racism, sexism, classism, and other forms of oppression (Alemán, 2009; Bernal, 2002).

The above two definitions of theory are consistent with the saying, "empirical research should make a *theoretical* contribution." Granted, theory can also mean "an overall perspective from which one sees and interprets the world" (Abend, 2008, p. 179), such as postmodern theory, poststructuralist theory, feminist theory, queer theory, and critical theory. They are more about philosophical viewpoints on how to see and interpret the world, but less about the specific constructs and the relationships among them (Willower, 1980). A theory explains the essential elements of *what* (e.g., constructs and variables), *how* (e.g., How are the constructs and variables related?), and *why* (e.g., the underlying assumptions that justify the selection of the constructs and variables and the proposed relationships). A good theory is not a list of variables or the addition of a new variable to an existing list. Empirical evidence, according to (Poincare, 1902), is analogous to stones used to build a house. A pile of stones is not a house, and a pile of constructs and variables does not constitute a theory. By contrast, a good theory constitutes description and explanation—"a plausible, cogent explanation for why we should expect certain relationships in our data" (Whetten, 1989, p. 491). Kuhn's (2012) description of theory was even more detailed, "theories should be accurate in their predictions, consistent, broad in scope, present phenomena in an orderly and coherent way, and be fruitful in suggesting new phenomena or relationships between phenomena" (p. xxxi). Given that many educational leadership studies were only theoretically oriented (Willower, 1975), the term "theory" in this article is inclusive, encompassing both definitions introduced above, including a set of constructs, assumptions, and generalizations serving to explain a phenomenon.

What is theory for?

What is theory for? In other words, what are the functions of theory? With ever-increasing data—along with sophisticated measurements and data analyses, why not letting data speak for themselves—without a theoretical framework? To answer these questions, it is important to first look into three strands of a discipline or field: theoretical, methodological, and evaluative strands (Kuhn, 2012). Each strand represents an essential dimension of the disciplinary matrix: theories guide research by bringing clarity in explaining

phenomena; research methods allow empirical work to produce results that eliminate the ambiguities of theories; evaluative standards serve as a means “to scrutinize theories in great empirical detail” (Kuhn, 2012, p. 42). Taken together, the three strands are entangled, interwoven in research to advance a field. The development in one strand influence other stands, and vice versa.

A primary function of theory is to allow generalization, which derives from empirical evidence, to explain a phenomenon. The phenomena of interest are “too complex and varied to be explored at random” (Kuhn, 2012, p. 109). To study the phenomena of interest, it is too mentally demanding, if not cognitively impossible, to probe in all directions. Without theory, how can we decide what should be observed? It is theory that guides researchers as they explore the entities involved in the phenomena and how the entities behave. If you are a researcher or aspire to be one, the key is to ask good research questions. Theories help you do that. How? Here is an example. In crime TV shows, detectives examine existing evidence, generate their theory to explain what happened at a crime scene, and then find further evidence to either support or refute their theories, and maybe come up with new theories to explain new evidence. In their cases, the detectives’ theories are constantly evolving with the newly uncovered evidence. This process is very similar to the functions of scientific theories, as detectives always say, “Follow the evidence!” Theory in scientific research functions as an engine to generate concrete explanations of phenomena. Researchers use theories to derive assumptions and constructs in intellectual inquiry. Without theory, research methods are devoid of reasoning. For example, listing variables and their relationships without explanatory text is not theory-driven research. Theory forces researchers to dig into theoretical explanations and use logical reasoning.

In a field as interdisciplinary as educational leadership, theory is essential to empirical research (Van Maanen et al., 2007). This is why every empirical article in the field is required to include a section of theoretical framework. In the 1950s when educational leadership was still at its inception, Griffiths (1959) described a theory of administration as (1) a guide to action, (2) a guide to the collection of facts, (3) a guide to new knowledge, and (4) a guide to explain the nature of administration. After decades of development, educational leadership research addresses an ever-broader set of issues in an ever-increasing diversity of settings and more deeply embedded in its interdisciplinary knowledge base, including psychology, sociology, sociopsychology, and economics (Wang and Bowers, 2016). Given the interdisciplinary nature of educational leadership, theory is more important than ever.

Where have we been?

Before peeking into the future, it seems sensible to look back and take stock of theories that have been used to frame empirical research in educational leadership. Theories about educational leadership abound. In the 1950s, the theory movement was launched to respond to the mounting dissatisfaction with the “poverty of theory within our own field” and the criticism that “our own research has been anchored in naked empiricism” (Halpin, 1970, p. 156). Heavily influenced by logical positivism, the theory movement favored a cumulative, generalizable knowledge base built from rigorous, hypothetical-deductive empirical research and aimed at building a unique theory for educational leadership (Culbertson, 1988; Getzels, 1977; Griffiths, 1983; Halpin, 1970). Walton’s (1955) lamented that educational leadership lacked “a well-defined, highly organized body of subject matter; it has no elegant simple theoretical structure; and as literature it is singularly devoid of esthetic qualities” (p. 169). To build a unique theory for educational leadership, researchers ventured into the theories in other fields.

From the outset in the 1950s and the 1960s, theories in educational leadership were rooted in the theories in the social and behavioral sciences (Bates, 1980; Boyan, 1981; Hoy, 1982; Oplatka, 2009; Walton, 1955; Willower, 1975). This laid the foundation for educational leadership as an interdisciplinary field. The theoretical groundings of educational leadership have grown rapidly since the 1950s, particularly over the last two decades of the 2000s and 2010s. Fig. 2 illustrates the growth of the theoretical groundings of educational leadership since 1999. In each network, a node represents a theory that was used to theoretically frame empirical studies published in four leading educational leadership research journals: *Educational Administration Quarterly*, *Journal of Educational Administration*, *Educational Management Administration & Leadership*, and *Journal of School Leadership*. A tie connects a pair of theories if they were used to frame the same article. As seen in Fig. 2, the theory co-occurrence networks have grown rapidly in terms of both network size and complexity. By the end of 2018, a total of 483 theories were used to frame empirical articles theoretically. If we zoom in onto the center of the network in 2018, we can see the theories at the center of the network were the ones that were frequently used to frame empirical articles (See Fig. 3). Those theories fall into four categories: (1) theories about leadership styles (i.e., theories explain patterns of leader behaviors), (2) theories in organizational context (i.e., theories explain collective behaviors and phenomena in organizations), (3) sociological, psychological, and socio-psychological theories (i.e., theories explain mental processes which are manifested through behaviors), and (4) theories about social justice (i.e., theories explain justice, equity, diversity, and inclusion in schools as organizations). This categorization is not mutually exclusive, as they are all closely interconnected to one another in Fig. 3.

Theories about leadership styles

Many educational leadership studies took a behavioral approach, focusing on leader behaviors. When leader behaviors show a pattern, such a pattern is considered as a leadership style. Given the two definitions of theory in the earlier section, the theories about leadership styles do not fit the first definition which is very stringent, but fit the second definition which is offering explanations of a particular social phenomenon of leadership. Notably, other terms—such as leadership models and leadership approaches—bear similar, if not the same, meaning as leadership styles. Nevertheless, the leadership styles that have been frequently used to frame

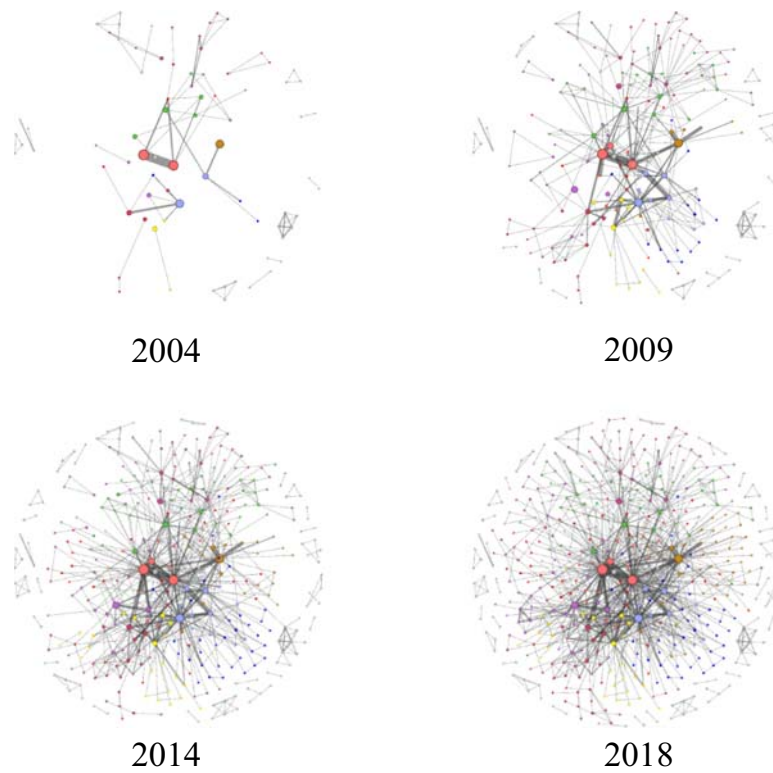


Fig. 2 Theory co-occurrence network of educational leadership from 1999 to 2018

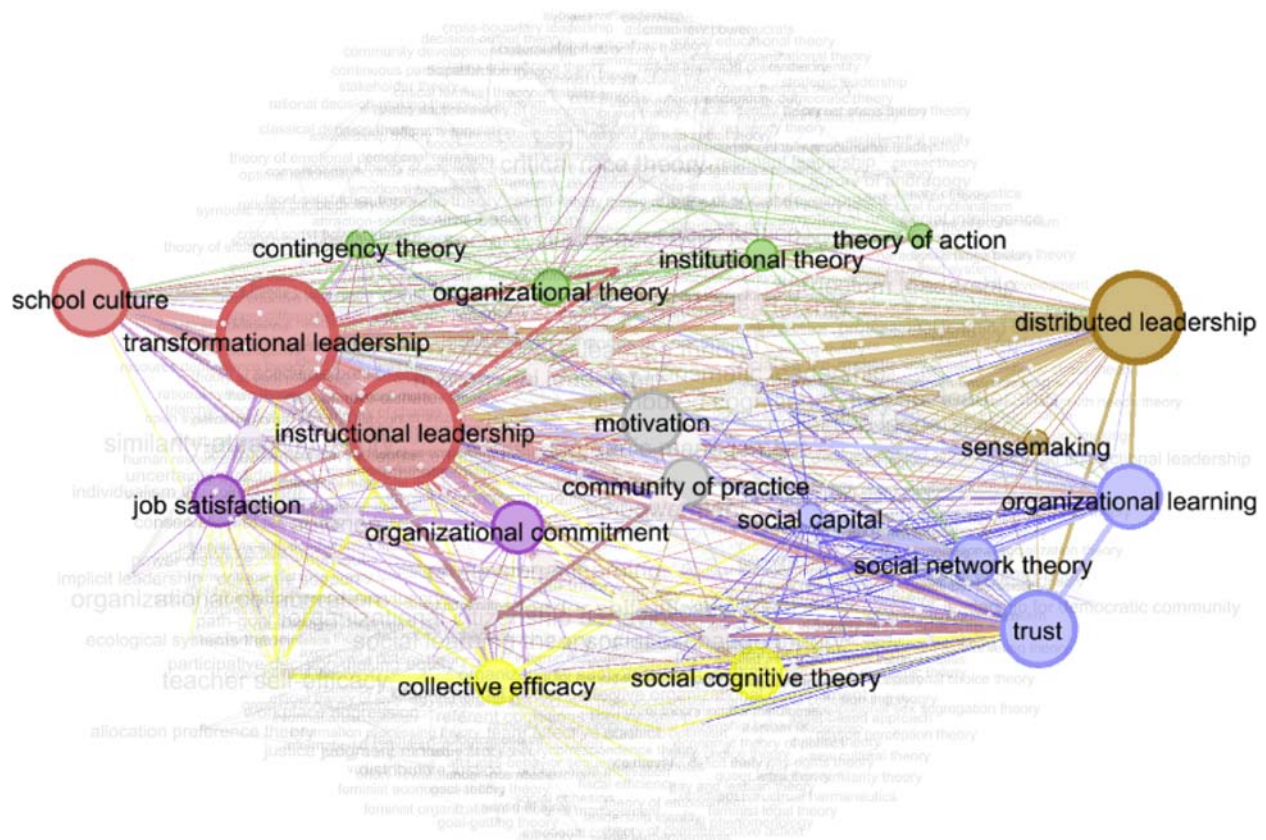


Fig. 3 Theories at the center of the theory co-occurrence network of educational leadership

Note. Node size indicates weighted degree centrality. The bigger the node size is, the more central a theory is in the network. Node color differentiates clusters detected by calculating the modularity score for each node.

empirical research include distributed leadership, instructional leadership, transformational leadership, social justice leadership, and teacher leadership (Wang, 2018a, b).

The theoretical roots of many leadership styles can be traced back to the literature in neighboring fields, including social behavioral science, psychology, organizational science, and administrative science (Boyan, 1981; Haller, 1968; Walton, 1955; Willower, 1975). For example, distributed leadership has its theoretical root in distributed cognition and activity theory (Spillane et al., 2001). Transformational leadership emerged in the body of literature that had a broad interest in leadership—not necessarily in schools but rather in all sorts of organizations and social groups (Bass, 1985), and then was applied to school settings (Leithwood, 1994).

After decades of theory development, some of the dominant leadership theories (e.g., distributed leadership and transformational leadership) remain as the mainstream in the field of educational leadership. Moreover, educational leadership scholars have developed theoretical explanations that are unique to educational organizations with a focus on teaching and learning, such as instructional leadership, teacher leadership, and leadership for learning. This appreciable progress spotlights the teachers and their instructional activities that are unique in school contexts in comparison with the leadership in non-school context (Greenfield, 1995).

Further, a wide range of theories has been used to explain educational leadership. In addition to the aforementioned theories, other theories include ethical leadership, democratic leadership, and participative leadership (Oplatka, 2010). Arguably, the latter two theories were very similar to distributed leadership which turned the focus from leaders at an individual level to a broader context. Instead of viewing leadership as extraordinary individuals with distinctive qualities and virtues, distributed leadership posited that leadership was the outcome of the interaction of leaders, followers, and situation (Spillane, 2006). The situation was mostly contextualized in schools as organizations (Bolden, 2011). Since 2000, in addition to distributed leadership, social justice leadership has gained traction in educational leadership, as evidenced by the rise of social justice-themed theories, such as social justice theory, critical race theory, and culturally responsive school leadership (Wang, 2018a, b)—a point I will return to shortly.

Theories in organizational context

In addition to leadership styles, another group of frequently used theories in educational leadership research is closely related to organizational context. The theories include organizational theory, organizational learning, organizational culture, and institutional theory. These theories were mostly borrowed from the neighboring field of organizational science (Oplatka, 2014).

In the 1950s and 1960s, social system and role theory were considered as major theories in educational leadership. These theories conceptualized a school and its people were in a closed-system organization, in which the absence of social, economic, and political impact was palpable (Hoy, 1982). In the 1970s, researchers acknowledged a “contingency-laden human context” of education leadership (Culbertson, 1981, p. 41). Instead of theorizing schools as a closed system marked with scientific management, bureaucracy, and social process theory, schools were considered as an open system. Universal theories of organizations were replaced by contingency theory (Hoy, 1982). Contingency theory considered the external environment (e.g., cultural, political, economic, informational and technical, and physical factors) as a moderator that influences the relationships between leadership and internal adjustments in organizations (e.g., decision-making process, structure, policies, practices, and organizational performance; Hanson, 1979). Despite theoretical advances, contingency theory was too inclusive because it suggests “it all depends.” Willower (1980) even questioned “whether contingency theory is a theory in the first place. Put another way, is contingency theory a set of related concepts and generalizations that serve to explain or is it a broad injunction to focus on the complexity of relationships and the contingency that affect them?” (p. 4).

In the 1980s the 1990s, as a response to the seminal report *A Nation at Risk: The Imperative for Educational Reform* (The National Commission on Excellence in Education, 1983), the rising interest in educational reform and school effectiveness (Oplatka, 2010; Wang et al., 2017) gave rise to a set of theories that address organizational change and reform. To study how to facilitate organizational change and reform in schools, many studies drew on organizational theories that were used in organizational science, including loosely coupled theory, organizational culture, organizational climate, bureaucracy, and empowerment.

From the 2000s to 2010s, there was an increasing number of theories that explained human behaviors in organizations and organizational outcomes. Among them, institutional theory, proposed by DiMaggio and Powell in the 1980s, was frequently used to frame research about organizational change. Institutional theory explained why organizations were so similar, regardless of who the leaders were. According to institutional theory, leaders have little impact, because organizations are exceedingly resistant to change, swept along by external forces, and constrained by norms (DiMaggio and Powell, 1983). However, there was a competing view derived from strategic leadership, which considered that leaders’ characteristics (e.g., personality, cognition, charisma, power, knowledge, skills, and abilities) greatly influence what happens to their organizations (Samimi et al., 2020). In educational leadership research, scholars investigated the relationship between school leadership and student achievement (Hitt and Tucker, 2016; ten Bruggencate et al., 2012; Witziers et al., 2003). In addition to student achievement as an organizational outcome, other theories about organizational outcomes were also frequently used in educational leadership research. Those theories included organizational citizenship behavior, organizational commitment, and employee job satisfaction. Moreover, as a response to the earlier criticism of contingency theory, researchers have worked to mitigate ambiguity and uncover intricacies in different contexts, including accountability policy implementation, teacher performance improvement in urban schools, and building organizational culture in turnaround schools. With student demographics changes, increasing attention has been devoted to how community values were manifest in school board elections, policymaking in schools as organizations, and organizational performance (Alsburry, 2003; Sampson, 2019).

Sociological, psychological, and socio-psychological theories

Given that educational leadership is a social process, many theories in sociology, psychology, and socio-psychology offer theoretical lenses for empirical research in educational leadership. These theories included trust, motivation, self-efficacy, collective efficacy, social capital, social exchange theory, and social network theory. These theories were not frequently applied to educational leadership research until the 2000s (Daly, 2010; Howard and Navarro, 2016).

Recently, those neighboring fields of educational leadership have ventured into neuroscience to investigate the brain mechanisms underlying social and psychological processes. In sociology, a subfield of social neuroscience has emerged to investigate human brains as social organs (Lieberman, 2013). In psychology, a subfield of cognitive neuroscience has emerged to investigate brain mechanisms underlying psychological processes (Gazzaniga et al., 2013). In economics, a subfield of neuroeconomics has emerged where behavioral economics, psychology, and neuroscience converge (Glimcher and Fehr, 2014). In organizational science—a field focusing on individual and collective behaviors in organizational context, a subfield of organizational neuroscience has emerged to “understand and incorporate the cognitive machinery behind our thoughts and actions into organizational theory” (Becker et al., 2011, p. 934). In educational leadership, Wang (2019) introduced brain mechanisms about some of the core constructs in leadership, such as trust, vision and envisioning, charisma, social identity, justice, feelings and emotions, and decision making.

Theories about social justice in education

Scholars have been wrestling with education inequities for decades. In 1966, sociologist James Coleman, commissioned by the United States Department of Health, Education, and Welfare, published the *Equality of Educational Opportunity Study*, also known as the Coleman Study (Coleman, 1995). This report uncovered unequal access to educational opportunities by children of different races, color, religions, and national origins. Since then, the Coleman Study has become the most cited report by educational leadership researchers (Wang and Bowers, 2016). However, the theories about education inequity and social justice did not emerge at the center of the theoretical groundings of educational leadership until the 2000s.

Among them, critical race theory and social justice leadership were frequently used to frame educational leadership studies. Critical race theory, first applied in education research in 1995 (Ladson-Billings and Tate, 1995), posits (1) permanence of racism, (2) whiteness as property, (3) counter storytelling and majoritarian narratives, (4) interest convergence, (5) critique of liberalism, and (6) intersectionality (Capper, 2015). Since then, critical race theory has been applied to study education politics (López, 2003), identify leadership practices for social justice and equity (Iverson, 2007), and analyze the language of the Interstate School Leaders Licensure Consortium (ISLLC) and the Educational Leaders Constituent Council (ELCC) standards (Davis et al., 2015).

Closely related to critical race theory is social justice leadership. With a strong emphasis on creating equal opportunities for students from marginalized groups (Furman and Gruenewald, 2004), socially just leaders are action oriented, transformative, committed, persistent, inclusive, democratic, relational, caring, reflective, and oriented toward a socially just pedagogy (Furman, 2012; Theoharis, 2008). They “translate democratic ideals and inclusive notions into concrete practices and redress injustices to meet the needs of diverse groups” (Wang, 2018a, p. 474). Social justice leadership offers an encompassing view on all forms of diversity, including race, class, gender, sexual orientation, disability, among others (Theoharis, 2007). To highlight students with disabilities, inclusion leadership was proposed (DeMatthews and Mawhinney, 2014). To highlight culture, culturally responsive school leadership was proposed (Khalifa et al., 2016). These theories about social justice have been gaining considerable traction over the last two decades, emerging as a core pillar undergirding the theoretical groundings of educational leadership.

What does the future hold?

Where should theory go? The theoretical groundings of educational leadership, as seen in Fig. 2, have grown more diverse than at any time in recent decades. The field has advanced from theories that focused on leadership at an individual level to leadership at the dyadic, group, and organizational levels. However, compared to a relatively large number of theoretically informed empirical studies, research on the theory development in educational leadership has been limited and has attracted scant attention in the research community. To promote theory development, I provide two recommendations for future researchers in educational leadership.

Recommendation 1: refining definitions of theories

The first recommendation is to refine loosely defined and underdeveloped theories. A problem in educational leadership is that many theories are widely used but poorly defined. Some definitions are largely overlapped with existing theories. Take teacher leadership as an example. In more than a decade of literature, teacher leadership was either rarely defined, or was defined ambiguously as mobilizing instructional expertise about teaching and learning to improve instruction and school culture (Wenner and Campbell, 2017). The latter was “closely aligned with both instructional and participative leadership” (York-Barr and Duke, 2004, p. 261). Moreover, in about 20% of teacher leadership studies, distributed leadership was used as the theoretical framework (Wenner and Campbell, 2017). To further complicating matters, when the definitions of two theories become too overlapped and indistinguishable, some scholars begin to use them interchangeably, such as organizational climate and organizational culture.

The definitions of some theories so elusive and overly inclusive that they offer little guidance in empirical research. Take sense-making as an example. Sensemaking, as the name suggests, refers to individuals or groups use prior experiences and knowledge to navigate ambiguous situations, make sense of information, and decide how to respond (Brown et al., 2015; Kiesler and Sproull, 1982). This definition of sensemaking is similar to social cognition which rests on two assumptions: (1) humans have a limited capacity to process all information; and (2) the cognitive process of making sense of information (i.e., perceiving, encoding, storing, retrieving, and inferring) is subject to individual capabilities (e.g., attention), social pressures, and limitations of information search. How school leaders made sense of race was subject to their own racial and role identity, as well as the history of their schools and communities, organizational norms, values, and beliefs (Evans, 2007). Sensemaking is also similar to upper echelons theory, which posits leaders' characteristics (e.g., knowledge, experience, values, personalities, and preferences) function as a filter and a processor of information (Hambrick and Mason, 1984). To do so, they go through a three-step process. Leaders' characteristics influences: (1) field of vision (i.e., the directions leaders look and listen), (2) selective perception (i.e., what leaders actually see and hear), and (3) interpretation (i.e., how leaders attach meaning to what they see and hear). While the content and labels for these constructs differ somewhat from one another, there is also a clear overlap. When it comes to what specific psychological and social processes that leaders' characteristics are converted into the sense they make in an ambiguous situation, it still remains largely a mystery (Hambrick, 2007a, b). Taken together, all these theories (i.e., sensemaking, social cognition, and upper echelons theory) were developed on the premise of bounded rationality—the idea that complex, uncertain situations are not objectively “knowable,” but rather are merely interpretation (March and Simon, 1958).

Neither indistinguishable nor nebulous definition is helpful for theory development in educational leadership. The purpose of defining a theory is to “set boundaries and limits,” and specify “what is included in the construct and what is not” (Yaniv, 2011, p. 590). Definitions are not just semantic clarifications. Rather, they are foundational to understand the scope and magnitude of a phenomenon. Defining and operationalizing a theory establish boundaries for empirical research. Without a clear definition, we lose grip on a phenomenon of interest. In this sense, I recommended that future researchers take up the challenge to refine definitions of theories in educational leadership. But how to do it?

To refine definitions of a theory, it is important to start with evaluating the utility of a theory or construct by applying a fundamental principle of Occam's razor (Banks et al., 2018). Occam's razor is a rule of thumb in science, suggesting that when two explanations of a phenomenon exist, all else equal, the simpler explanation should be employed (Blumer et al., 1987). Put differently, we choose a parsimonious explanation. The principle of parsimony has been recommended to address construct redundancy which was prevalent in leadership research. As a subfield of leadership in the context of education, the principle of parsimony is therefore recommended to evaluate the sheer volume of similar, overlapping, and indistinguishable theories in educational leadership.

To follow the principle of parsimony, the utility of a construct can be evaluated by (1) content validity, (2) convergent validity, and (3) discriminant validity. Content validity means a construct can be used interchangeably with others, and the construct definition is aligned with its measure(s). Convergent validity means different measures of the same construct need to have a degree of consistency. For example, four dimensions of transformational leadership (i.e., idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration) should be somewhat related, but not so highly related as to suggest redundancy. Discriminant validity refers to the distinctiveness of constructs. One approach to evaluate discriminant validity is to compare meta-analytic correlations across the leadership constructs, conduct confirmatory factor analysis, and calculate average variance extracted and composite reliability estimates (Bergh et al., 2016). If the average variance exceeds the squared factor correlation between any two constructs, then a construct is considered to have discriminant validity (Fornell and Larcker, 1981). After reviewing 57 meta-analytic leadership studies, Banks et al. (2018) reported that no evidence was found to indicate moral leadership behavior was distinct from transformational or charismatic leadership behavior.

Despite the wide variety of definitions and conceptualizations of educational leadership, many theories have been loosely defined, conceptually and operationally. To advance the understanding of educational leadership, it is therefore recommended that future researchers follow the principle of parsimony, evaluate theories by the content, convergent, and discriminant validity, better define and bring clarity to constructs and variables, and set theoretical boundaries in empirical research.

Recommendation 2: pruning theories

My second recommendation is to use empirical research to prune theories. Theory pruning is defined as “hypothesis specification and study design intended to bound and reduce theory” (Leavitt et al., 2010, p. 644). A problem with organizational science was “the focus is on the support or lack of support for a single theory, with little attention to an alternative theory that could also describe the results” (Leavitt et al., 2010, p. 646). The same problem exists in educational leadership research as well. Over the last two decades from 1999 to 2018, a total of 483 theories had been used to frame empirical studies in educational leadership. Most of the educational leadership empirical research was designed to expand existing theories (e.g., add mediators or moderators) or to present new theories (e.g., teacher leadership and culturally responsive leadership).

Despite the value of research that expands existing theories and present new theories, the scant attention to theory pruning can be problematic in educational leadership research (Hambrick, 2007a, b) criticized this focus on expand existing theories and presenting new theories as “too much of a good thing” (p. 1346) and characterized this insistence on new theory as a “fetish” (p. 1350). Other scholars echoed this criticism. After analyzing 10 years of publications in *Organizational Research Methods* and the top 25 papers in management, according to citations, it was reported that very few papers tested competing theories (Aguinis et al., 2009). Among them, there was a paper that reconciled two opposing views about the effects of leaders at the top of organizational

hierarchy on organizational outcomes by introducing the construct of managerial discretion as a moderator (Finkelstein and Hambrick, 1990). One view, derived from strategic leadership, was that leaders at the upper echelons greatly influenced their organizational outcomes. The competing view, derived from institutional theory (DiMaggio and Powell, 1983), was that executives have little effect because organizations are exceedingly inertial and constrained by external and internal forces. Finkelstein and Hambrick argued that both views are conditionally valid, depending on how much managerial discretion—or latitude of action exists. Discretion exists when there is a limited constraint (e.g., a weak, passive board) and when there is a great deal of means-ends ambiguity—that is, when there are multiple plausible alternatives. In educational leadership, the limited attention to testing competing theories is glaring. Among the top five most highly cited educational leadership studies in each decade since 1960 (Hallinger and Kovačević, 2019), none of those studies put competing theories to test. However, “the greatest scientific value emerges when at least two models are specified representing competing conceptualizations and one emerges the strongest, especially over several replications” (Vandenberg and Grelle, 2008, p. 170).

Corroborating existing theories—without pruning inaccurate or ill-fitting theories—may contribute limited insights, hindering theory development in educational leadership. In the light of new empirical evidence, we can eliminate false theories (e.g., the earth is the center of the universe), correct false assumptions of existing theories (e.g., prospect theory correcting the false assumption of rational decision making in classic economics), and replace existing theories with more accurate ones. Most educational leadership scholars have taken a “confirmatory” approach, adding propositions to an ever-growing and expansive body of knowledge. Using empirical evidence to verify theories is indeed important for any field. However, it is equally important to use empirical findings to prune theory (Shareef, 2007). Theories are not static entities. Rather, they “are tentatively held, waiting to be disapproved or simply discarded for a variety of reasons and replaced by new conjectures” (Willower, 1980, p. 1). Eric Scerri, a philosopher of science, further elaborated on the evolving nature of theories: “One strength of science is that it is often wrong. It proceeds by overthrowing preconception, perhaps replacing misconceptions by a more sophisticated misconception until that misconception is replaced by yet another, until ultimately (we optimists all hope) arriving at some version of the truth” (p. x). To advance a field, researchers first conduct empirical research driven by theories. With new evidence emerges, some theories would undergo revision as more discoveries are made. Theories are then undergoing patchwork: they are altered, refined, fine-tuned, changed, or refuted with adjustments, but the general model remains. Consider social justice leadership and culturally responsive leadership. Also consider critical race theory and critical Latino theory. Once new theories are proposed, some researchers conduct empirical research to test the new theories, and others may be skeptical and propose a counter-argument. Over time, empirical evidence accumulates. If the accumulating evidence can be explained by the theories, then the theories survive. Otherwise, they are left in the dust, and the exploration of new theories marches on. An example in organizational science is that we no longer think of an organization has been organized into a closed system but into multiple dynamic systems.

If educational leadership researchers always confirm their priors, we fall into the prey of confirmation bias. To prune theories, it is important to ask: What findings do not stand up to the rigors of science? What are the assumptions proven wrong? Heaps of evidence may show the earlier theories could be on shaky ground. The trial-and-error process is the essence of science.

Conclusion

This article seeks to peek into the future of theoretical groundings in educational leadership research. To do so, I first discussed what theory is, and what theory is for. I then outlined the current theoretical groundings to set the stage for a forward-looking discussion. To catalyze the theory development in educational leadership, I provided two recommendations: refining definitions of theories and pruning theories.

Theory development—along with methodological advances—make theories and constructs ripe for empirical exploration, conceptual elaboration, theory pruning, and evaluative scrutiny. No theory is perfect. Do not slavishly believe all theories, or submissively accept all assumptions of the theories are correct. Challenge them. Argue against them. Find the evidence. Come up with your own theories. Allow your theories to be challenged by others. Allow your theories to evolve as the new evidence emerge. It is up to you to ensure the theory development in educational leadership. Feel free to use your intellectual prowess to challenge preconceptions and misconceptions in our field. An analogy is to see theories and constructs as the ingredients in your kitchen. Be innovative, put different ingredients together, and see how they interact and what they offer in the end. They might end up with something that makes you salivate, or something you cannot wait to throw into your trashcan. Nevertheless, you are trying to create something new. It’s part of the trial-and-error process in science.

That is a future worth working hard.

References

- Abend, G., 2008. The meaning of “theory”. *Socio. Theor.* 26 (2), 173–199.
- Aguinis, H.H., Pierce, C.A., Bosco, F.A., Muslin, I.S., 2009. First decade of Organizational Research Methods: trends in design, measurement, and data-analysis topics. *Organ. Res. Methods* 12 (1), 69–112.
- Alémán Jr., E., 2009. LatCrit educational leadership and advocacy: struggling over whiteness as property in Texas school finance. *Equity Excell. Educ.* 42 (2), 183–201.

- Alsbury, T.L., 2003. Superintendent and school board member turnover: political versus apolitical turnover as a critical variable in the application of the dissatisfaction theory. *Educ. Adm. Q.* 39 (5), 667–698.
- Banks, G.C., Gooty, J., Ross, R.L., Williams, C., 2018. Construct redundancy in leader behaviors: a review and agenda for the future. *Leader. Q.* 29 (1), 236–251.
- Bass, B.M., 1985. *Leadership and Performance Beyond Expectations*. Collier Macmillan. Collier Macmillan, Springfield.
- Bass, B.M., 1998. *Transformational Leadership: Industrial, Military, and Educational Impact*. Erlbaum, Mahwah.
- Bates, R.J., 1980. Educational administration, the sociology of science, and the management of knowledge. *Educ. Adm. Q.* 16 (2), 1–20.
- Becker, W.J., Cropanzano, R., Sanfey, A.G., 2011. Organizational neuroscience: taking organizational theory inside the neural black box. *J. Manag.* 37, 933–961.
- Bergh, D.D., et al., 2016. Using meta-analytic structural equation modeling to advance strategic management research: guidelines and an empirical illustration via the strategic leadership-performance relationship. *Strat. Manag. J.* 37 (3), 477–497.
- Bernal, D.D., 2002. Critical race theory, Latino critical theory, and critical raced-gendered epistemologies: recognizing students of color as holders and creators of knowledge. *Qual. Inq.* 8 (1), 105–126.
- Blumer, A., Ehrenfeucht, A., Haussler, D., Warmuth, M.K., 1987. Occam's razor. *Inf. Process. Lett.* 24 (6), 377–380.
- Bolden, R., 2011. Distributed leadership in organizations: a review of theory and research. *Int. J. Manag. Rev.* 13 (3), 251–269.
- Boyan, N.J., 1981. Follow the leader: commentary on research in educational administration. *Educ. Res.* 10 (2), 6–21.
- Brown, A.D., Colville, I., Pye, A., 2015. Making sense of sensemaking in organizational studies. *Organ. Stud.* 36 (2), 265–277.
- Burt, R.S., 2005. *Brokerage and Closure: An Introduction to Social Capital*. Oxford university press, Cary.
- Capper, C.A., 2015. The 20th-year anniversary of critical race theory in education: implications for leading to eliminate racism. *Educ. Adm. Q.* 51 (5), 791–833.
- Coleman, J.S., 1995. *Equality of Educational Opportunity (Coleman) Study (EEO5)*, 1966, S.L. Inter-university Consortium for Political and Social Research.
- Colquitt, J.A., et al., 2001. Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *J. Appl. Psychol.* 86 (3), 425–445.
- Culbertson, J.A., 1981. Antecedents of the theory movement. *Educ. Adm. Q.* 17 (1), 25–47.
- Culbertson, J.A., 1988. A century's quest for a knowledge base. In: Boyan, N.J. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York, pp. 3–26.
- Daly, A.J., 2010. *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge.
- Davis, B.W., Gooden, M.A., Mischeaux, D.J., 2015. Color-blind leadership: a critical race theory analysis of the ISLLC and ELCC standards. *Educ. Adm. Q.* 51 (3), 335–371.
- DeMatthews, D., Mawhinney, H., 2014. Social justice leadership and inclusion: exploring challenges in an urban district struggling to address inequities. *Educ. Adm. Q.* 50 (5), 844–881.
- DiMaggio, P.J., Powell, W.W., 1983. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *Am. Socio. Rev.* 48, 147–160.
- Evans, A.E., 2007. School leaders and their sensemaking about race and demographic change. *Educ. Adm. Q.* 43 (2), 159–188.
- Finkelstein, S., Hambrick, D.C., 1990. Top-management-team tenure and organizational outcomes: the moderating role of managerial discretion. *Adm. Sci. Q.* 35 (3), 484–503.
- Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. *J. Market. Res.* 18 (1), 39–50.
- Furman, G.C., Gruenewald, D.A., 2004. Expanding the landscape of social justice: a critical ecological analysis. *Educ. Adm. Q.* 40 (1), 49–78.
- Furman, G., 2012. Social justice leadership as praxis: developing capacities through preparation programs. *Educ. Adm. Q.* 48 (2), 191–229.
- Gazzaniga, M.S., Ivry, R.B., Mangun, G.R., 2013. *Cognitive Neuroscience: The Biology of the Mind*, fourth ed. W.W. Norton, New York.
- Getzels, J.W., 1977. Educational administration twenty years later, 1954–1974. In: Cunningham, L.L., Hack, W.G. (Eds.), *Educational Administration: The Developing Decades*. McCutchan Publishing Corporation, Berkeley, pp. 3–24.
- Glimcher, P.W., Fehr, E., 2014. *Neuroeconomics: Decision Making and the Brain*. Elsevier, London.
- Greenfield Jr., W.D., 1995. Toward a theory of school administration: the centrality of leadership. *Educ. Adm. Q.* 31 (1), 61–85.
- Griffiths, D.E., 1959. *Administrative Theory*. Prentice-Hall, Englewood Cliffs.
- Griffiths, D.E., 1983. Evolution in research and theory: a study of prominent researchers. *Educ. Adm. Q.* 19 (3), 201–221.
- Haller, E.J., 1968. The interdisciplinary ideology in educational administration: some preliminary notes on the sociology of knowledge. *Educ. Adm. Q.* 4 (2), 61–77.
- Hallinger, P., Kovačević, J., 2019. *Science Mapping the Knowledge Base in Educational Leadership and Management: A Longitudinal Bibliometric Analysis, 1960–2018*. Educational Management Administration & Leadership.
- Halpin, A.W., 1970. Administrative theory: the fumbled torch. In: Kroll, A.M. (Ed.), *Issues in American Education: Commentary on the Current Sense*. Oxford University Press, New York, pp. 156–183.
- Hambrick, D.C., Mason, P.A., 1984. Upper echelons: the organization as a reflection of its top managers. *Acad. Manag. Rev.* 9 (2), 193–206.
- Hambrick, D.C., 2007a. The field of management's devotion to theory: too much of a good thing? *Acad. Manag. J.* 50 (6), 1346–1352.
- Hambrick, D.C., 2007b. Upper echelons theory: an update. *Acad. Manag. Rev.* 32 (2), 334–343.
- Hanson, E.M., 1979. School management and contingency theory: an emerging perspective. *Educ. Adm. Q.* 15 (2), 98–116.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86 (2), 531–569.
- Howard, T.C., Navarro, O., 2016. Critical race theory 20 years later: where do we go from here? *Urban Educ.* 51 (3), 253–273.
- Hoy, W.K., 1982. Recent developments in theory and research in educational administration. *Educ. Adm. Q.* 18 (3), 1–11.
- Iverson, S.V., 2007. Camouflaging power and privilege: a critical race analysis of university diversity policies. *Educ. Adm. Q.* 43 (5), 586–611.
- Kalshoven, K., Den Hartog, D.N., De Hoogh, A.H., 2011. Ethical leadership at work questionnaire (ELW): development and validation of a multidimensional measure. *Leader. Q.* 22 (1), 51–69.
- Khalifa, M.A., Gooden, M.A., Davis, J.E., 2016. Culturally responsive school leadership: a synthesis of the literature. *Rev. Educ. Res.* 86, 1272–1311.
- Kiesler, S., Sproull, L., 1982. Managerial response to changing environments: perspectives on problem sensing from social cognition. *Adm. Sci. Q.* 27, 548–570.
- Kuhn, T., 2012. *The Structure of Scientific Revolutions*, fiftieth anniversary ed. University of Chicago Press, Chicago.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Leavitt, K., Mitchell, T.R., Peterson, J., 2010. Theory pruning: strategies to reduce our dense theoretical landscape. *Organ. Res. Methods* 13 (4), 644–667.
- Leithwood, K.A., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30 (4), 498–518.
- Lieberman, M.D., 2013. *Social: Why Our Brains Are Wired to Connect*. Crown, New York.
- López, G.R., 2003. The (racially neutral) politics of education: a critical race theory perspective. *Educ. Adm. Q.* 39 (1), 68–94.
- March, J.G., Simon, H.A., 1958. *Organizations*. John Wiley & Sons, New York.
- Marks, H.M., Printy, S.M., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Myran, S., Sutherland, I., 2019. Defining learning in educational leadership: reframing the narrative. *Educ. Adm. Q.* 55 (4), 657–696.
- Oplatka, I., 2009. The field of educational administration: an historical overview of scholarly attempts to recognize epistemological identities, meanings and boundaries from 1960s onwards. *J. Educ. Adm.* 47 (1), 8–35.
- Oplatka, I., 2010. *The Legacy of Educational Administration: A Historical Analysis of an Academic Field*. Peter Lang, Berlin.
- Oplatka, I., 2014. Differentiating the scholarly identity of educational administration: an epistemological comparison of two neighbouring fields of study. *J. Educ. Adm.* 52 (1), 116–136.
- Poincaré, J.H., 1902. *La science et l'hypothèse [The science and the hypothesis]*. E. Flammarion, Paris.
- Samimi, M., Cortes, A.F., Anderson, M.H., Herrmann, P., 2020. What is strategic leadership? Developing a framework for future research. *Leader. Q.*
- Sampson, C., 2019. (Im) Possibilities of Latinx school board members' educational leadership toward equity. *Educ. Adm. Q.* 55 (2), 296–327.
- Shareef, R., 2007. Want better business theories? Maybe Karl Popper has the answer. *Acad. Manag. Learn. Educ.* 6, 272–280.

- Spihlane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco.
- Spihlane, J.P., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- ten Bruggencate, G., Luyten, H., Scheerens, J., Sleegers, P., 2012. Modeling the influence of school leaders on student achievement: how can school leaders make a difference? *Educ. Adm. Q.* 48 (4), 699–732.
- The National Commission on Excellence in Education, 1983. *A Nation at Risk: The Imperative for Educational Reform*. Available at: <https://www2.ed.gov/pubs/NatAtRisk/risk.html>.
- Theoharis, G., 2007. Social justice educational leaders and resistance: toward a theory of social justice leadership. *Educ. Adm. Q.* 43 (2), 221–258.
- Theoharis, G., 2008. Woven in deeply: identity and leadership of urban social justice principals. *Educ. Urban Soc.* 41 (1), 3–25.
- Van Maanen, J., Sørensen, J.B., Mitchell, T.R., 2007. The interplay between theory and method. *Acad. Manag. Rev.* 32 (4), 1145–1154.
- Vandenberg, R.J., Grelle, D.M., 2008. Alternative model specifications in structural equation modeling: fact, fictions and truth. In: Lance, C.E., Vandenberg, R.J. (Eds.), *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences*. Taylor & Francis Group, New York, pp. 165–192.
- Walton, J., 1955. The theoretical study of educational administration. *Harv. Educ. Rev.* 25 (3), 169–178.
- Wang, Y., Bowers, A.J., 2016. Mapping the field of educational administration research: a journal citation network analysis. *J. Educ. Adm.* 54 (3), 242–269.
- Wang, Y., Bowers, A.J., Fikis, D.J., 2017. Automated text data mining analysis of five decades of educational leadership research literature: probabilistic topic modeling of EAQ articles from 1965 to 2014. *Educ. Adm. Q.* 53 (2), 289–323.
- Wang, F., 2018a. Social justice leadership—theory and practice: a case of Ontario. *Educ. Adm. Q.* 54 (3), 470–498.
- Wang, Y., 2018b. The panorama of the last decade's theoretical groundings of educational leadership research: a concept co-occurrence network analysis. *Educ. Adm. Q.* 54 (3), 327–365.
- Wang, Y., 2019. Pulling at your heartstrings: examining four leadership approaches from the neuroscience perspective. *Educ. Adm. Q.* 55 (2), 328–359.
- Wenner, J.A., Campbell, T., 2017. The theoretical and empirical basis of teacher leadership: a review of the literature. *Rev. Educ. Res.* 87 (1), 134–171.
- Whetten, D.A., 1989. What constitutes a theoretical contribution? *Acad. Manag. Rev.* 14 (4), 490–495.
- Willower, D.J., 1975. Theory in educational administration. *J. Educ. Adm.* 13 (1), 77–91.
- Willower, D.J., 1980. Contemporary issues in theory in educational administration. *Educ. Adm. Q.* 16 (3), 1–25.
- Witziers, B., Bosker, R.J., Krüger, M.L., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 39 (3), 398–425.
- Yaniv, E., 2011. Construct clarity in theories of management and organization. *Acad. Manag. Rev.* 36 (3), 590–592.
- York-Barr, J., Duke, K., 2004. What do we know about teacher leadership? Findings from two decades of scholarship. *Rev. Educ. Res.* 74 (3), 255–316.

Leadership practices and why they matter: evolving leadership models

Tony Bush, Educational Leadership, University of Nottingham, Nottingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	44
The significance of school leadership	44
Definitions of school leadership	44
Leadership as influence	45
Leadership and values	45
Leadership and vision	45
A typology for leadership	45
Instructional leadership	45
Managerial leadership	46
Managerialism	46
Transactional leadership	46
Transformational leadership	47
Moral leadership	47
Distributed leadership	48
Allocative distribution	49
Teacher leadership	49
Contingent leadership	49
Conclusion: why leadership practices matter	50
References	51

Introduction

The growth in the importance of school leadership over the past 20 years has been accompanied by theory development, with new models emerging and established approaches being redefined and further developed. The purpose of this article is to review current and recent writing on school leadership models. The article reviews theoretical literature, to see how leadership is conceptualized, and empirical literature, to demonstrate whether and how the research evidence supports these concepts of school leadership. It draws on UK and international school leadership literature.

The significance of school leadership

It is widely recognized that leadership is second only to classroom teaching in its impact on student learning. [Leithwood et al. \(2006\)](#) widely cited report shows that “leadership acts as a catalyst” (p. 4) for beneficial effects, including pupil learning. The report also distinguishes between the impact of headteacher leadership (typically 5–7%) and total leadership (27%). This finding provides much of the empirical underpinning for the current interest in distributed leadership (see below) and for the concept of leadership “density”.

[Robinson \(2007\)](#) meta-analysis of published research shows that the nature of the leader’s role has a significant impact on learning outcomes. Direct leader involvement in curriculum planning and professional development is associated with moderate or large leadership effects. “This suggests that the closer leaders are to the core business of teaching and learning, the more likely they are to make a difference to students” ([Robinson, 2007: 21](#)). This finding indicates that instructional leadership effects are much greater than those of other leadership models (see below).

[Leithwood et al. \(2006: 5\)](#) conclusion, that “there is not a single documented case of a school successfully turning around its pupil achievement trajectory in the absence of talented leadership”, offers powerful support for the vital role of principals, senior and middle leaders in all types of school.

Definitions of school leadership

[Gunter \(2004\)](#) shows that the labels used to define this field have changed from “educational administration” to “educational management” and, more recently, to “educational leadership”. [Bush \(2008\)](#) discusses whether such changes are purely semantic or reflect substantive changes in the nature of the field. [Yukl \(2002: 4\)](#) argues that “the definition of leadership is arbitrary and very subjective”, but the following “working definition” includes its main features:

Leadership is a process of influence leading to the achievement of desired purposes. Successful leaders develop a vision for their schools based on their personal and professional values. They articulate this vision at every opportunity and influence their staff and other stakeholders to share the vision. The philosophy, structures and activities of the school are geared towards the achievement of this shared vision

Bush and Glover (2003: 5).

Three dimensions of leadership arise from this working definition.

Leadership as influence

Most definitions of leadership reflect the assumption that it involves a social influence process whereby intentional influence is exerted by one person [or group] over other people [or groups] to structure activities and relationships in a group or organization (Yukl, 2002: 3). Bush (2008: 277) refers to three key aspects of these definitions. First, the central concept is influence rather than authority. Both are dimensions of power, but the latter tends to reside in formal positions, such as principal or head teacher, while the former could be exercised by anyone in the school or college. Leadership is independent of positional authority while management is linked directly to it. Second, the process is intentional. The person seeking to exercise influence is doing so to achieve certain purposes. Third, influence may be exercised by groups as well as individuals. This notion provides support for the concept of distributed leadership and for constructs such as senior leadership teams.

This aspect of leadership portrays it as a fluid process, potentially emanating from any part of the school, independent of formal management positions and capable of residing with any member of the organization, including associate staff and students

Bush (2008: 277).

Leadership and values

The notion of “influence” is neutral in that it does not explain or recommend what goals or actions should be pursued. However, leadership is increasingly linked with values. Leaders are expected to ground their actions in clear personal and professional values. Hargreaves (2004), drawing on research in Canadian schools, finds that teachers report largely positive emotional experiences of self-initiated change but predominantly negative ones concerning mandated change. There is a tension here between the obligation to implement the policies of democratically elected governments and the need for teacher professionals to feel positive about new initiatives if they are to act on them successfully.

Leadership and vision

Vision has been regarded as an essential component of effective leadership for more than 20 years. Southworth (1993: 23–4) suggests that heads are motivated to work hard “because their leadership is the pursuit of their individual visions” (Southworth, 1993: 47). Murphy and Torre (2015: 177) claim that “vision routinely surfaces in studies of effective organizations”, adding that it is “a hallmark component of every important school leadership framework (Murphy and Torre, 2015: 78). They scaffold vision to comprise three levels—mission, goals and expectations.”

The articulation of a clear vision has the potential to develop schools, but the empirical evidence of its effectiveness remains mixed. Thoonen et al. (2011: 520) refer to the “adverse effects” of vision, which arise when principals do not involve teachers in the process of vision building. Hoyle and Wallace (2005: 139) add that visions are expected to conform to centralized expectations.

A typology for leadership

There are many alternative and competing models of school leadership. In this section, we review eight of these models, drawing on Leithwood et al. (1999), and Bush (2020).

Instructional leadership

The increasing emphasis on managing teaching and learning as the core activities of educational institutions has led to “instructional leadership”, or “learning-centred” leadership, being emphasized. The term “instructional leadership” derives from North America and Hallinger (2019) traces its “transformation” from an American to a global leadership construct. He concludes that “effective instructional leadership results when leaders align both instructional and managerial roles with their personal values” (Hallinger, 2019: 32).

Instructional leadership is the longest established concept linking leadership and learning. However, several other terms may be used to describe this relationship, including pedagogic leadership, curriculum leadership and leadership for learning. Despite its prominence and longevity, instructional leadership has been criticized on two grounds. First, it is perceived to be primarily concerned with teaching rather than learning (Bush, 2013). The second criticism is that it “focused too much on the principal as the

center of expertise, power and authority" (Hallinger, 2003: 330), ignoring or underplaying the role of other leaders such as deputy principals, middle managers, leadership teams, and teacher leaders. Lambert (2002: 37) claims that "the days of the lone instructional leader are over. We no longer believe that one administrator can serve as the instructional leader for the entire school without the substantial participation of other educators". Hallinger (2019) historical analysis shows that instructional leadership has moved from being regarded as a singular function, exercised by the principal, to being conceptualized as a shared role, including middle and teacher leadership.

Instructional leadership focuses primarily on the direction and purpose of leaders' influence; targeted at student learning via teachers. There is less emphasis on the influence process itself. However, most other leadership models focus strongly on leadership processes. The discussion below focuses on models which mainly address "how" to lead, with less emphasis on the purpose of leadership.

Managerial leadership

Managerial leadership assumes that the focus of leaders ought to be on procedures, functions, and activities. Influence accrues largely because of the formal authority of leaders and Leithwood et al. (1999) argue that influence is allocated in proportion to the status of those positions in the organizational hierarchy. Leithwood et al. (Leithwood et al., 1999: 15) add that "there is evidence of considerable support in the literature and among practicing leaders for managerial approaches to leadership". They add that "positional power, in combination with formal policies and procedures, is the source of influence exercised by managerial leadership" (Leithwood et al., 1999: 17).

Gumus et al. (2018) review of educational leadership models shows the enduring influence of managerial leadership, linked to scientific management theory. They add that this model "may be more visible in those countries in which the education system is centralized, since it considerably emphasizes the implementation of practices mandated by higher external management authorities within the bureaucratic hierarchy" (Gumus et al., 2018: 29). Semarco and Cho (2018: 911) show that what they describe as "task-oriented managerial leadership behaviors" are common in Ghana, a centralized and bureaucratic education system. As noted earlier, a managerial approach to leadership does not necessarily obviate principals from also being effective instructional leaders (Hallinger, 2019).

Managerialism

The shift in the language of school organization to favor "leadership" at the expense of "management" is partly semantic, as noted above, but also reflects anxiety about the dangers of value-free management, focusing on efficiency for its own sake, what Hoyle and Wallace (2005: 68) describe as "management to excess":

Effective leadership and management "take the strain" by creating structures and processes which allow teachers to engage as fully as possible in their key task. Managerialism, on the other hand, is leadership and management to excess. It transcends the support role of leadership and, in its extreme manifestation, becomes an end in itself.

Managerial leadership is the model which provides the greatest risk of a managerialist approach to school organization. By focusing on functions, the aims of education may be subordinated to the managerial aim of greater efficiency. Simkins (2005: 13–14) claims that, in England, managerialist values, such as rigid planning and target-setting regimes, were being set against traditional professional values. Managerial leadership is an essential component of successful schools, but it should complement, not supplant, values-based approaches. Effective management is essential but value-free managerialism is inappropriate and damaging (Bush, 2020).

Transactional leadership

Transactional leadership relates to relationships between leaders and teachers being based on the exchange of valued resources. In its simplest form, teachers provide educational services (teaching, pupil welfare, extracurricular activities) in exchange for salaries and other rewards. Anderson and Chang (2019: 87) say that transactional leadership is "sometimes called leader-member exchange theory". Miller and Miller (2001) definition refers to transactional leadership as an exchange process while Judge and Piccolo (2004: 755) add that transactional leaders "focus on the proper exchange of resources". Interaction between administrators and teachers is usually episodic, short-lived, and limited to the exchange transaction.

Boocock (2019) reports on what he regards as transactional leadership in the English further education sector. He argues that successive governments have operated in a transactional mode, since colleges were incorporated in 1993. This takes the form of financial incentives to increase student enrollment. In addition, he claims that the Ofsted, the English inspection agency, encourages transactional leadership, notably through benchmarking in respect of student retention, achievement, and success rates, linked to a top-down inspection framework.

The major limitation of transactional leadership is that it does not engage staff beyond the immediate gains arising from the transaction and does not produce long-term commitment to the values and vision being promoted by school leaders. However,

Bass (1998: 11) stresses that leaders often use both transformational and transactional approaches. “Consistent honoring of transactional agreements builds trust, dependability, and perceptions of consistency with leaders by followers, which are each a basis for transformational leadership”. Oterkil and Ertesvag (2014) also claim that the best leaders are both transactional and transformational.

Transformational leadership

Transformational leadership is often contrasted with transactional approaches (e.g., Miller and Miller, 2001). This form of leadership assumes that the central focus of leadership ought to be the commitments and capacities of organizational members. Higher levels of personal commitment to organizational goals, and greater capacities for accomplishing those goals, are assumed to result in extra effort and greater productivity (Leithwood et al., 1999: 9).

The contemporary policy climate within which schools operate also raises questions about the validity of the transformational model, despite its popularity in the literature. Transformational language is used by governments to encourage, or to require, practitioners to adopt and implement centrally determined policies. In South Africa, for example, the language of transformation is used to underpin a non-racist post-Apartheid education system. The policy is rich in symbolism but weak in practice because many school principals lack the capacity and the authority to implement change effectively (Bush and Glover, 2016).

In certain centralized systems, governments may be seen to require school leaders to adhere to government policies, which affect aims, curriculum content and pedagogy, as well as values. In this respect, transformation may be a unilateral process of implementation, not a context-specific assessment of the needs of individual schools and their communities. Particularly significant is the raft of research on this model in highly centralized Arab countries, including the United Arab Emirates (Litz and Scott, 2017; Stringer and Hourani, 2016), Jordan (Khasawneh et al., 2012), and Kuwait (Alsaedi and Male, 2013). Litz and Scott (2017) stress the need to adapt the model to make it suitable for Arab contexts, notably the expectations around hierarchy and collectivism. They propose a “cross-cultural transformational leadership model” to reflect such expectations.

International research supports the notion that transformational leadership may produce beneficial effects. Thomas et al. (2020), for example, show that the principal’s transformational leadership was related to positive teacher attitudes, in their study of first-year primary school teachers in Flanders (Belgium). Similarly, Ninkovic and Floric (2018) indicate that transformational leadership had positive effects on collective teacher efficacy in Serbia.

The transformational model is comprehensive in that it provides a normative approach to school leadership which focuses primarily on the process by which leaders seek to influence school outcomes rather than on the nature or direction of those outcomes. However, it may also be criticized for its potential to become “despotic” because of its strong, heroic, and charismatic features. Allix (2000) believes that the leader’s power ought to raise “moral qualms” and serious doubts about its appropriateness for democratic organizations.

The transformational model stresses the importance of values but the debate about its validity relates to the central question of “whose values”? Critics of this approach argue that the decisive values are often those of government or of the school principal, who may be acting on behalf of government. Educational values, as held and practiced by teachers, may be subjugated to externally imposed values.

Moral leadership

The moral leadership model assumes that the critical focus of leadership ought to be on leaders’ values, beliefs, and ethics. Sergiovanni (1984: 10) says that “excellent schools have central zones composed of values and beliefs that take on sacred or cultural characteristics”, adding that “administering” is a “moral craft” (Sergiovanni, 1991: 322):

The school must move beyond concern for goals and roles to the task of building purposes into its structure and embodying these purposes in everything that it does with the effect of transforming school members from neutral participants to committed followers. The embodiment of purpose and the development of followership are inescapably moral.

Sergiovanni (1991: 323).

Gumus et al. (2018) claim that moral leadership requires several qualities, including holding to values, following ethical decision rules, and becoming a role model through visible moral actions. Keddie (2016: 724) research in two Queensland (Australia) schools leads her to address issues of equity and to conclude that “morally-focused leadership [is] concerned with more than increasing attainment on external measures of performance”. This links to the notion of “moral purpose”, illustrated in Rintoul and Goulais (2010) study of Ontario (Canada) vice-principals and in Bezzina (2012) research in New South Wales (Australia).

Moral leadership may have a spiritual dimension, especially in religious schools. Woods (1997: 148) survey of headteachers in England found that 52% “were inspired or supported in their leadership by some kind of spiritual power”. Gibson (2014: 531) study of leadership in secular New Zealand primary schools shows that “spirituality is a complex and controversial human phenomenon, the meaning of which may be shaped and re-shaped by diverse perspectives and experiences.... These characteristics were woven through the lived reality of the principals’ lives, through their beliefs, moral values, attitudes, and the quality of their

actions'. The author concludes that "spirituality in principle leadership is not one form of leadership but invites principals to integrate personal meanings of spirituality into their preferred leadership practice" (Gibson, 2014: 533).

Hammad and Shah (2019) report on the challenges of leading faith schools in a secular society, with specific reference to head teachers of four English Muslim schools. They note that Muslim schools with an emphasis on Islamic values appear to provide a "safe haven", where young Muslims are expected to grow as worthy members of the community. The challenges facing such heads include issues in developing Muslim identity in a secular society, recruitment of qualified Muslim teachers, and the expectations of their local communities. The authors conclude that supporting faith schooling is consistent with approaches such as integration, cultural pluralism, widening participation and celebrating diversity.

Sergiovanni (1991: 329) argues for both moral and managerial leadership:

In the principalship, the challenge of leadership is to make peace with two competing imperatives, the managerial and the moral. The two imperatives are unavoidable and the neglect of either creates problems. Schools must be run effectively if they are to survive.... But for the school to transform itself into an institution, a learning community must emerge... [This] is the moral imperative that principals face.

Distributed leadership

The models discussed above are essentially about individual (usually principal) leadership. However, there have been several approaches which seek to widen the debate to include shared approaches to leadership. Collegial and participative leadership were popular shared approaches in the 1980s and 1990s, but distributed leadership has become the normatively preferred leadership model in the 21st century. Gronn (2010: 70) states that "there has been an accelerating amount of scholarly and practitioner attention accorded [to] the phenomenon of distributed leadership".

Lumby (2019: 10) argues that "distributed leadership is not underpinned by a clear definition", adding that its "unique selling point" is the potential for "emergent spontaneous leadership" (Lumby, 2019). Hairon and Goh (2015) add that the construct of distributed leadership is elusive. Tian et al. (2016) systematic review of the literature also concludes that there is no clear definition.

An important starting point for understanding this phenomenon is to uncouple it from positional authority. As Harris (2004: 13) indicates, "distributed leadership concentrates on engaging expertise wherever it exists within the organization rather than seeking this only through formal position or role". She claims that it "is characterized as a form of collective leadership" (Harris, 2004: 14). Crawford (2012: 611) notes that interest in distributed leadership is part of a wider shift away from leadership as a solo activity. "This solo stance has been mediated by more prominence on shared leadership, variously described as democratic leadership, collegiality, participative leadership and distributed leadership". Gumus et al. (2018: 31) add that distributed leadership is often used interchangeably with shared, collaborative, delegated, and dispersed leadership.

Gronn (2010: 70) refers to a normative switch "from heroics to distribution" but also cautions against a view that distributed leadership necessarily means any reduction in the scope of the principal's role. Indeed, Hartley (2010: 271) argues that "its popularity may be pragmatic: to ease the burden of overworked headteachers". Lumby (2009: 320) adds that distributed leadership "does not imply that school staff are necessarily enacting leadership any differently" to the time "when heroic, individual leadership was the focus of attention". Referring to the South African context, Du Plessis and Heystek (2019: 3) ask whether there is sufficient "space for principals to practice distributed leadership in a traditional hierarchical system?". This point highlights the dilemma about how to accommodate distribution alongside solo leadership. Gronn (2010: 77) "hybrid" model of leadership may offer the potential to harness the best of both individual and distributed approaches.

The interest in, and support for, distributed leadership is predicated on the assumption that it will bring about beneficial effects that would not occur with singular leadership. Leithwood et al. (2006) important study of the impact of school leadership led to the articulation of "seven strong claims" about successful school leadership. Two of these claims relate to distributed leadership and Leithwood et al. (2006: 12) also show that multiple leadership is much more effective than solo leadership, explaining as much as a 27% variation in student achievement across schools, much higher than is typically reported in studies of individual headteacher effects. Hallinger and Heck (2010: 670) also found that distributed leadership was significantly related to change in academic capacity and, thus, to growth in student learning.

The contemporary popularity of distributed leadership is illustrated by the many studies influenced by this model. These include research with schools in England (Bush and Glover, 2012), schools in South Africa (Du Plessis and Heystek, 2019), rural primary schools in Ireland (Brown et al., 2019), Singapore schools (Hairon and Goh, 2015; Ho et al., 2016), Malaysian schools (Bush and Ng, 2019), early childhood in Finland (Kangas et al., 2016), and in Australia (Heikka et al., 2013), learning organizations in Sweden (Liljenberg, 2015), higher education in Australia (Jones et al., 2014), and in Finland (Vuori, 2019), and vocational education in the Netherlands (Bouwman et al., 2019). This extensive body of research offers a mixed picture of the value and impact of distributed leadership. However, there is some evidence that this model is being identified as the "label" for activities that might have been badged differently without the current interest in distribution. For example, Bush and Glover (2012) research with high performing school leadership teams shows that many leaders claim that their success is attributable to distributed leadership, but this was at a time when the English NCSL was advocating this model.

Allocative distribution

As noted earlier, distributed leadership is conceptualized as an emergent and spontaneous phenomenon in western literature (Bennett et al., 2003; Harris, 2013; Tian et al., 2016). However, the support for this model in centralized contexts has led to a different mode of distribution, with principals allocating tasks and responsibilities in a manner often indistinguishable from the management concept of delegation.

Bush and Ng (2019) examine the extent and nature of distribution in 14 schools in two Malaysian states, Selangor, and Sarawak. Distributed leadership features prominently in the country's main reform document, the Malaysia Education Blueprint. The advocacy of distributed leadership by the Malaysian Ministry of Education is recognition of its potential to enhance leadership density, and thus potentially to contribute to improved student outcomes. However, the Blueprint links distribution to the hierarchy in two ways. First, the focus is firmly on middle and senior leaders holding formal roles in the structure. Second, the scope of distribution is circumscribed; leaders will be prepared to fully utilize the decision-making flexibilities "accorded to them" (MEB: E28) (present authors' emphasis).

The Blueprint's cautious approach to distribution is consistent with the notion of allocative distributed leadership (Bolden et al., 2009). This suggests an uneasy compromise between the free-flowing assumptions of distributed leadership theory and the rigid requirements of the hierarchy. There is evidence of allocative distribution, predominantly to senior leaders, in most of the case study schools. The evidence from the Malaysian schools is that distributed leadership is almost indistinguishable from delegation because principals remain in control and have firm reporting requirements (Bush and Ng, 2019).

Hairon and Goh (2015) research on distributed leadership in Singapore leads them to discuss the related notion of "bounded empowerment". "It was obvious that empowerment in the Singapore context has a certain degree of 'restriction', hence bounded empowerment... empowerment in Singapore suggests the necessity to only relinquish control of decisions with certain caveats" (Hairon and Goh, 2015: 707).

Such restricted modes of distribution are not confined to centralized contexts. Holloway et al. (2018) discuss "prescribed" distributed leadership in the US, in the context of high accountable demands. They conclude that "highly prescribed leadership roles might help schools respond to external policy demands, but they do not necessarily nurture opportunities for leadership growth" (Holloway et al., 2018: 551).

These studies indicate that the language of distribution has been captured to legitimize practices previously described as delegation, contributing to critiques of this leadership model. Tian et al. (2016) meta-analysis includes a section called "questioning the concept of distributed leadership". One aspect of this critique relates to micropolitics, with Johnson (2019) warning that distributed leadership might be camouflaged as a micropolitical strategy to rationalize top-down management. Similarly, Lumby (2013) argues that distributed leadership is a political phenomenon "replete with the uses and abuses of power".

Lumby (2019: 11) offers a criticism which links to the earlier discussion of allocative distribution. "What appears to be happening resembles... formal and informal delegation within a bureaucratic system". She concludes that "the discovery of persistent hierarchy in sites where distributed leadership has been researched suggests that at least elements of bureaucracy are a near-universal approach to leadership" (Lumby, 2019: 14).

Teacher leadership

There are clear links between teacher leadership and distributed leadership. Frost (2008: 337) characterizes the former as involving shared leadership, teachers' leadership of development work, teachers' knowledge building, and teachers' voice. Grant (2006) offers a four-part model of teacher leadership, based on research in South African schools: leadership within the classroom, leadership beyond the classroom, with other teachers, leadership for whole-school development, and leadership beyond the school and in the community. She (2006: 523) adds that there are three prerequisites for teacher leadership: First, a collaborative culture is required, with participatory decision-making, and vision sharing. Second, a set of values is needed, to assist in developing the collaborative culture. Third, the principal and the formal management team need to encourage a distributed leadership approach.

Lai and Cheung (2015) examine teacher leadership in the context of curriculum reform in Hong Kong. They argue that teacher leadership has three distinct strands; participation, learning and influence, and they all require both teacher agency and support from senior leadership.

Developing teacher leadership may be particularly challenging in centralized contexts, for example in Egypt (Emira, 2010), where the Ministry of Education restricts teacher leadership to the classroom, and in Iran (Aliakbari and Sadeghi, 2014), where the education system is "top-down". However, Bush et al. (2016) research on master teachers as teacher leaders in Malaysia and Philippines shows that master teachers were able to operate at all four levels of the Grant (2006) model.

Stevenson (2012) argues that the interpretation of teacher leadership is managerialist in nature and inherently conservative. Helderbran (2010: 363) notes that teacher leadership "remains largely an academic topic and, even though inroads have been made, teacher leadership remains more a concept than an actuality". A culture of trust and collaboration is essential, as is a shared vision of where the school needs to go, clear line management structures and strong leadership development programmes'.

Contingent leadership

The models of leadership examined above are all partial. They provide valid and helpful insights into one specific aspect of leadership. None of these models provide a complete picture of school leadership. As Lambert (1995: 2) notes, there is "no single best

type". The contingent model provides an alternative approach, recognizing the diverse nature of school contexts, and the advantages of adapting leadership styles to the situation, rather than adopting a "one size fits all" stance.

Yukl (2002: 234) adds that "the managerial job is too complex and unpredictable to rely on a set of standardized responses to events. Effective leaders are continuously reading the situation and evaluating how to adapt their behavior to it". Leadership requires effective diagnosis of problems, followed by adopting the most appropriate response to the issue or situation (Morgan, 1997). This reflexive approach is particularly important in periods of turbulence when leaders need to be able to assess the situation carefully and react as appropriate rather than relying on a standard leadership model.

Notman (2017) study of two urban school principals in New Zealand shows the importance of the context in deciding leadership approaches. These leaders "were highly responsive to the demands and challenges within and beyond their own school context" (Notman, 2017: 768). The author concludes by stressing the need for contingent or "adaptive" leadership, sensitive to the school context, rather than a single inflexible model.

A contingent approach also helps to counter the normative features of many leadership models and responds to the tendency to advocate one "right" approach to school leadership. By recognizing that a range of approaches can be valid, it provides a more complete picture of leadership practice. The contingent model is pragmatic and is not underpinned by a clear set of values.

Conclusion: why leadership practices matter

The concept of management has been joined, or superseded, by the language of leadership but the activities undertaken by principals and senior staff resist such labels. Successful leaders are increasingly focused on learning, the central and unique purpose focus of educational organizations. They also face unprecedented accountability pressures in many countries in what is clearly an increasingly "results driven" business. As these environmental pressures intensify, leaders and managers require greater understanding, skill, and resilience to sustain their institutions (Bush, 2020).

Theory is one of the four essential building blocks of school leadership. Alongside policy, research, and practice, it provides helpful insights into how schools are led and managed. The theory of leadership is important for two main reasons. First, it provides a way of understanding and interpreting the actions of leaders. The models outlined in this article provide lenses to view and comprehend the ways in which leadership is enacted in schools. Second, understanding theory provides a guide to leadership practice for principals and other leaders. It widens horizons and avoids drawing only on the inevitably limited individual or collective experience of any single leader.

Each of the leadership models discussed in this paper is partial. They provide distinct but uni-dimensional perspectives on school leadership. They are also artificial distinctions, or "ideal types", in that most successful leaders are likely to embody most or all of these approaches in their work. For example, heads may aspire to develop distributed instructional leadership.

The importance of context has been increasingly acknowledged (Hallinger, 2018). This has several dimensions and there is a distinction between national contexts, with political, economic and social aspects, and internal school contexts, often conceptualized as organizational culture (Walker and Haiyan, 2019). Contingent leadership acknowledges the diverse nature of school contexts, and the advantages of adapting leadership styles to the situation, rather than adopting a "one size fits all" stance. It is not a single model but represents a mode of responsiveness which requires effective diagnosis followed by careful selection of the most appropriate leadership style. It is pragmatic, rather than normative, unlike most of the other leadership models (Bush, 2020).

Leadership practices matter because of the evidence that effective leadership enhances student learning and school improvement (Leithwood et al., 2006). There is also a body of research showing that certain models may be particularly effective in raising student outcomes. The Leithwood et al. (2006) study, for example, shows the potential of distributed leadership. "School leadership has a greater influence of schools and pupils when it is widely distributed" (Leithwood et al., 2006: 12). In contrast, Robinson (2007) meta-analysis indicates that instructional leadership has more impact on student learning than other leadership models. Transformational leadership also has its advocates because of its potential to inspire greater commitment from teachers and other stakeholders.

The shifting popularity of different leadership models leads to consideration of the nature of theory in educational leadership. To what extent is it a representation of practice (description), and to what extent does it constitute advocacy; a normative perspective? Distributed leadership is the most recent model to be subject to a strongly normative approach; "the theory of choice for many" (Lumby, 2013: 581). Gunter et al. (2013) criticize the normative approach to distributed leadership. "Normative work tends to present the idea of distributed leadership as an imperative for practitioners as school improvers... the normative approach tends to have rationales that are based on exhorting how and why distributed leadership should be adopted" (Gunter et al., 2013: 563).

Lumby (2013: 582) comments that discussion of distributed leadership as a heuristic tool gave way to an evangelical approach, for example by the English National College for School Leadership (NCSL). This may explain, in part, the frequent references to distributed leadership by participants in a study of senior leadership teams at schools rated as "outstanding" by Ofsted (Bush and Glover, 2012). Unprompted, these leaders attributed their perceived success to adopting distributed approaches.

A normative approach to leadership theory, by academics, practitioners, or policy makers, is not new. Bureaucracy, collegiality, transformational leadership, and instructional leadership are also treated in this way. This suggests that leadership is over-theorized and under-researched.

The leadership typology discussed in this article provides many clues for heads, senior and middle leaders, and senior leadership teams. Managerial leadership, operating through the hierarchy, can mandate clearly targeted change, such as a stronger focus on examination and test scores. However, this often depends on a single leader and may not lead to sustainable change. Transformational

leadership approaches aim to widen commitment to school-wide objectives, through the development of shared vision, but the “vision” is often that of the head or principal with acquiescence, rather than genuine commitment, from teachers and other staff.

The limitations of the hierarchy have led to a plethora of alternative models, including distributed and teacher leadership, which are designed to broaden leadership and to stress lateral as well as vertical relationships. These are often manifested in team-based structures. [Bush and Glover \(2012\)](#) study of high performing senior leadership teams showed their value in providing coherence and leadership “density”.

While there are different approaches to leadership and management, recognition of the need for clear direction-setting, linked to moral purpose, is likely to be helpful in enhancing school leadership, for the benefit of pupils and students.

References

- Aliakbari, M., Sadeghi, A., 2014. Iranian teachers' perceptions of teacher leadership practices in schools. *Educ. Manag. Adm. Leader* 42 (4), 576–592.
- Allix, N.M., 2000. Transformational leadership: democratic or despotic? *Educ. Manag. Adm.* 28 (1), 7–20.
- Alsaedi, F., Male, T., 2013. Transformational leadership and globalization: attitudes of school principals in Kuwait. *Educ. Manag. Adm. Leader* 41 (5), 640–657.
- Anderson, G., Chang, E., 2019. Competing narratives of leadership in schools: the institutional and discursive turns in organizational theory. In: Connolly, M., Eddy-Spicer, D., James, C., Kruse, S. (Eds.), *The Sage Handbook of School Organization*. Sage, London.
- Bass, B.M., 1998. *Transformational Leadership: Industry, Military and Educational Impact*. Erlbaum, Mahwah, NJ.
- Bennett, N., Harvey, J., Wise, C., Woods, P., 2003. *Distributed Leadership: A Desk Study*. NCSL, Nottingham.
- Bezzina, M., 2012. Paying attention to moral purpose in leading learning: lessons from the leaders transforming learning and learners project. *Educ. Manag. Adm. Leader* 40 (2), 248–271.
- Bolden, R., Petrov, G., Gosling, J., 2009. Distributed leadership in higher education: rhetoric and reality. *Educ. Manag. Adm. Leader* 37 (2), 257–277.
- Boocock, A., 2019. Meeting the needs of local communities and businesses: from transactional to eco-leadership in the English further education sector. *Educ. Manag. Adm. Leader* 47 (3), 349–368.
- Bouwman, M., Runhaar, P., Wesseling, R., Mulder, M., 2019. Towards distributed leadership in vocational education and training schools: the interplay between formal leaders and team members. *Educ. Manag. Adm. Leader* 47 (4), 555–571.
- Brown, M., McNamara, G., O'Hara, J., Hood, S., Burns, D., Kurum, G., 2019. Evaluating the impact of distributed culturally responsive leadership in a disadvantaged rural primary school in Ireland. *Educ. Manag. Adm. Leader* 47 (3), 457–474.
- Bush, T., Glover, D., 2003. *School Leadership: Concepts and Evidence*. National College for School Leadership, Nottingham.
- Bush, T., Glover, D., 2012. Distributed leadership in action: leading high performing leadership teams in English schools. *Sch. Leader. Manag.* 32 (1), 21–36.
- Bush, T., Glover, D., 2016. School leadership and management in South Africa: findings from a systematic literature review. *Int. J. Educ. Manag.* 30 (2), 216–231.
- Bush, T., Ng, A.Y.M., 2019. Distributed leadership and the Malaysia Education Blueprint: from prescription to partial school-based enactment in a highly centralized context. *J. Educ. Adm.* 57 (3), 279–295.
- Bush, T., Ng, A., Glover, D., Romero, M.J., 2016. Master teachers as teacher leaders: evidence from Malaysia and the Philippines. *Int. Stud. Educ. Adm.* 43 (2), 81–102.
- Bush, T., 2008. From management to leadership: semantic or meaningful change? *Educ. Manag. Adm. Leader* 36 (2), 271–288.
- Bush, T., 2013. Instructional leadership and leadership for learning: Global and South African perspectives. *Education as Change* 17 (S1), 5–20.
- Bush, T., 2020. *Theories of Educational Management*, fifth ed. Sage, London.
- Crawford, M., 2012. Solo and distributed leadership: definitions and dilemmas. *Educ. Manag. Adm. Leader* 40 (5), 610–620.
- Du Plessis, A., Heystek, J., 2020. Possibilities for distributed leadership in South African schools: policy ambiguities and blind spots. *Educ. Manag. Adm. Leader* 48 (5), 840–860.
- Emira, M., 2010. Leading to decide or deciding to lead? Understanding the relationship between teacher leadership and decision-making. *Educ. Manag. Adm. Leader* 38 (5), 591–612.
- Frost, D., 2008. Teacher leadership: values and voice. *Sch. Leader. Manag.* 28 (4), 337–352.
- Gibson, A., 2014. Principals' and teachers' views of spirituality in principle leadership in three primary schools. *Educ. Manag. Adm. Leader* 42 (4), 520–535.
- Grant, C., 2006. Emerging voices on teacher leadership: some South African views. *Educ. Manag. Adm. Leader* 34 (4), 511–532.
- Gronn, P., 2010. Where to next for educational leadership? In: Bush, T., Bell, L., Middlewood, D. (Eds.), *The Principles of Educational Leadership and Management*. Sage, London.
- Gumus, S., Bellibas, M., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48.
- Gunter, H., Hall, D., Bragg, J., 2013. Distributed leadership: a study in knowledge production. *Educ. Manag. Adm. Leader* 41 (5), 555–580.
- Gunter, H., 2004. Labels and labeling in the field of educational leadership. *Discourse Stud. Cult. Polit. Educ.* 25 (1), 21–41.
- Hairon, S., Goh, J., 2015. Pursuing the elusive construct of distributed leadership: is the search over? *Educ. Manag. Adm. Leader* 43 (5), 693–718.
- Hallinger, P., Heck, R., 2010. Leadership for learning: does collaborative leadership make a difference in student learning? *Educ. Manag. Adm. Leader* 38 (6), 654–678.
- Hallinger, P., 2003. Leading educational change: reflections on the practice of instructional and transformational leadership. *Camb. J. Educ.* 33 (3), 329–350.
- Hallinger, P., 2018. Bringing context out of the shadows of leadership. *Educ. Manag. Adm. Leader* 46 (1), 5–24.
- Hallinger, P., 2019. Instructional leadership 1950–2018: transformation from an American to a global leadership construct. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *Principles of Educational Leadership and Management*, third ed. Sage, London.
- Hammad, W., Shah, S., 2019. Leading faith schools in a secular society: challenges facing head teachers of Muslim schools in the United Kingdom. *Educ. Manag. Adm. Leader* 47 (6), 943–959.
- Hargreaves, A., 2004. Inclusive and exclusive educational change: emotional responses of teachers and implications for leadership. *Sch. Leader. Manag.* 24 (3), 287–306.
- Harris, A., 2004. Distributed leadership in schools: leading or misleading? *Educ. Manag. Adm. Leader* 32 (1), 11–24.
- Harris, A., 2013. Distributed leadership: friend or foe? *Educ. Manag. Adm. Leader* 41 (5), 545.
- Hartley, D., 2010. Paradigms: how far does research in distributed leadership “stretch”. *Educ. Manag. Adm. Leader* 38 (3), 271–285.
- Heikka, J., Waniganayake, M., Hujala, E., 2013. Contextualizing distributed leadership within early childhood education: current understandings, research evidence and future challenges. *Educ. Manag. Adm. Leader* 41 (1), 30–44.
- Helterbran, V.R., 2010. Teacher leadership: overcoming “I am just a teacher” syndrome. *Education* 131 (2), 363–371.
- Ho, J., Chen, D.V., Ng, D., 2016. Distributed leadership through the lens of activity theory. *Educ. Manag. Adm. Leader* 44 (5), 814–836.
- Holloway, J., Nielsen, A., Saltmarsh, S., 2018. Prescribed distributed leadership in the era of accountability: the experiences of mentor teachers. *Educ. Manag. Adm. Leader* 46 (4), 538–555.
- Hoyle, E., Wallace, M., 2005. *Educational Leadership: Ambiguity, Professionals and Managerialism*. Sage, London.
- Johnson, B.L., 2019. Where have all the flowers gone? Reconnecting leadership preparation with the field of organization theory. In: Connolly, M., Eddy-Spicer, D., James, C., Kruse, S. (Eds.), *The Sage Handbook of School Organization*. Sage, Los Angeles.

- Jones, S., Harvey, M., Lefoe, G., Ryland, K., 2014. Synthesizing theory and practice: distributed leadership in higher education. *Educ. Manag. Adm. Leader* 42 (5), 603–619.
- Judge, T., Piccolo, R., 2004. Transformational and transactional leadership: a meta-analysis test of their relative validity. *J. Appl. Psychol.* 89 (5), 755–768.
- Kangas, J., Venninen, T., Ojala, M., 2016. Distributed leadership as administrative practice in Finnish early childhood education and care. *Educ. Manag. Adm. Leader* 44 (4), 617–631.
- Keddie, A., 2016. School autonomy as “the way of the future”: issues of equity, public purpose and moral leadership. *Educ. Manag. Adm. Leader* 44 (5), 713–727.
- Khasawneh, S., Omari, A., Abu-Tineh, A.M., 2012. The relationship between transformational leadership and organizational commitment: the case of vocational education in Jordan. *Educ. Manag. Adm. Leader* 40 (4), 494–508.
- Lai, E., Cheung, D., 2015. Enacting teacher leadership: the role of teachers in bringing about change. *Educ. Manag. Adm. Leader* 43 (5), 673–692.
- Lambert, L., 1995. New directions in the preparation of educational leaders. *Thrust Educ. Leadersh.* 24 (5), 6–10.
- Lambert, L., 2002. A framework for shared leadership. *Educ. Leader* 59 (8), 37–40.
- Leithwood, K., Jantzi, D., Steinbach, R., 1999. *Changing Leadership for Changing Times*. Open University Press, Buckingham.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven Strong Claims about Successful School Leadership*. DfES, London.
- Liljenberg, M., 2015. Distributing leadership to establish developing and learning school organizations in the Swedish context. *Educ. Manag. Adm. Leader* 43 (1), 152–170.
- Litz, D., Scott, S., 2017. Transformational leadership in the educational system of the United Arab Emirates. *Educ. Manag. Adm. Leader* 45 (4), 566–587.
- Lumby, J., 2009. Collective leadership of local school systems: power, autonomy and ethics. *Educ. Manag. Adm. Leader* 37 (3), 310–328.
- Lumby, J., 2013. Distributed leadership: the uses and abuses of power. *Educ. Manag. Adm. Leader* 41 (5), 581–597.
- Lumby, J., 2019. Distributed leadership and bureaucracy. *Educ. Manag. Adm. Leader* 47 (1), 5–19.
- Miller, T., Miller, J., 2001. Educational leadership in the new millennium: a vision for 2020. *Int. J. Leader. Educ.* 4 (2), 181–189.
- Morgan, G., 1997. *Images of Organization*. Sage, Newbury Park, CA.
- Murphy, J., Torre, D., 2015. Vision: essential scaffolding. *Educ. Manag. Adm. Leader* 43 (2), 177–197.
- Ninkovich, S.R., Floric, O.C.K., 2018. Transformational school leadership and teacher self-efficacy as predictors of perceived collective teacher efficacy. *Educ. Manag. Adm. Leader* 46 (1), 49–64.
- Notman, R., 2017. Professional identity, adaptation and self: cases of New Zealand principals during a time of change. *Educ. Manag. Adm. Leader* 45 (5), 759–773.
- Oterkil, C., Ertesvag, S., 2014. Development of a measurement for transformational and transactional leadership in schools taking on a school-based intervention. *Educ. Manag. Adm. Leader* 42 (4S), 5–27.
- Rintoul, H.M., Goulais, L., 2010. Vice principalship and moral literacy: developing a moral compass. *Educ. Manag. Adm. Leader* 38 (6), 745–757.
- Robinson, V., 2007. *School Leadership and Student Outcomes: Identifying What Works and Why*, Winmalee. Australian Council of Leaders, New South Wales.
- Semarco, S., Cho, S., 2018. The predictive influence of headteachers’ task-oriented managerial leadership behaviors on teachers’ retention intentions in Ghana. *Educ. Manag. Adm. Leader* 46 (6), 908–925.
- Sergiovanni, T., 1984. Leadership and excellence in schooling. *Leadership* 41 (5), 4–13.
- Sergiovanni, T.J., 1991. *The Principalship: A Reflective Practice Perspective*. Allyn Bacon, Needham Heights, MA.
- Simkins, T., 2005. Leadership in education: “What works” or “What make sense”. *Educ. Manag. Adm. Leader* 33 (1), 9–26.
- Southworth, G., 1993. School leadership and school development: reflections from research. *Sch. Organisat.* 12 (2), 73–87.
- Stevenson, H., 2012. Teacher leadership as intellectual leadership: creating spaces for alternative voices in the English school system. *Prof. Dev. Educ.* 38 (2), 345–360.
- Stringer, P., Hourani, R.B., 2016. Transformation of roles and responsibilities of principals in times of change. *Educ. Manag. Adm. Leader* 44 (2), 224–246.
- Thomas, L., Tuytens, M., Devos, G., Kelchtermans, G., Vanderlinde, R., 2020. Transformational school leadership as a key factor for teachers’ attitudes during their first year in the profession. *Educ. Manag. Adm. Leader* 48.
- Thoonen, E., Sleegers, P., Oort, F., Thea, T.D., Peetsma, T., Femke, P., Geijssels, P., 2011. How to improve teaching practices: the role of teacher motivation, organizational factors and leadership practices. *Educ. Adm. Q.* 47 (3), 496–536.
- Tian, M., Risku, M., Collin, K., 2016. A meta-analysis of distributed leadership from 2002 to 2013: theory development, empirical analysis and future research focus. *Educ. Manag. Adm. Leader* 44 (1), 146–164.
- Vuori, J., 2019. Distributed leadership in the construction of a new higher education campus and community. *Educ. Manag. Adm. Leader* 47 (2), 224–240.
- Walker, A., Haiyan, Q., 2019. Leadership and culture. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *Principles of Educational Leadership and Management*, third ed. Sage, London.
- Woods, G., 2007. The “bigger feeling”: the importance of spiritual experience in educational leadership. *Educ. Manag. Adm. Leader* 35 (1), 135–155.
- Yukl, G.A., 2002. *Leadership in Organizations*, fifth ed. Prentice-Hall, Upper Saddle River, NJ.

Transformational school leadership

Jingping Sun, Department of Educational Leadership, Policy and Technology Studies, The College of Education, University of Alabama, Tuscaloosa, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	53
The origin and major developments of transformational leadership	54
The birth of transformational school leadership (TSL)	54
Typical models and practices of transformational leadership	55
Setting directions	55
Developing a shared vision and building goal consensus	55
Holding high performance expectations	55
Developing people	55
Modeling behavior	55
Providing intellectual stimulation	55
Providing individualized support	56
Redesigning the organization	56
Strengthening school culture	56
Building collaborative structures	56
Engaging communities	56
Improving instruction	56
Focusing on instructional development	56
Managing and securing accountability	56
Contingent reward	56
Management by exception	56
Providing resources	56
Securing accountability	56
The impact of transformational leadership on school conditions	56
The impact of transformational leadership on teachers	57
The impact of transformational leadership on students	57
Direct effects of transformational school leadership on student achievement	57
Indirect effects of transformational leadership on student achievement	58
Conclusions	58
References	59

Introduction

This article synthesizes what is known about transformational leadership in school settings from previous reviews of such research (Chin, 2007; Leithwood et al., 1996; Leithwood and Jantzi, 2005; Leithwood and Sun, 2009, 2012; Sun and Leithwood, 2012), as well as from additional quantitative journal articles and dissertations reported in the past 10 years not included in previous reviews. The new evidence includes 99 dissertations from the ProQuest Dissertation database and 33 journal articles from the Eric-ProQuest Database, *Review of Educational Research* and the following eight major journals in the field of educational leadership: *Educational Administration Quarterly*, *Journal of School Leadership*, *Journal of Educational Administration*, *Leadership and Policy in Schools*, *International Journal of Leadership in Education*, *International Studies in Educational Administration*, *Educational Management, Administration and Leadership*, and *School Effectiveness and School Improvement*.

In order to be included in this review, a study must have been quantitative, empirically examined transformational leadership in school settings and reported a correlation between transformational leadership and at least one outcome variable. While this additional evidence may not be exhaustive, it does represent a sizable amount of research on transformational leadership in schools. Due to limited space, some of the new evidence providing background for the article is not included on the reference list.

The remainder of the article highlights the origins and development of the concept of transformational leadership (TL), identifies its typical practices, and reports the impact of TL on schools, teachers, and students. TL impact was estimated using vote-counting methods and the averaging of correlation coefficients or effect sizes. This article concludes with suggestions for future research.

The origin and major developments of transformational leadership

Transformational leadership was initially conceptualized by Burns (1978), in response to researcher and practitioner skepticism about inquiries into leadership effectiveness. Burns defined the purpose of leadership as exercising influence on followers to work toward transcendental goals instead of immediate self-interest and toward achievement and self-actualization instead of safety and security. Scholars from the traits, behavioral, and contingency schools of leadership had tried to find associations between leaders' traits and behaviors (especially contingent behaviors) in different contexts with varying levels of organizational effectiveness (Day and Antonakis, 2012), but these studies yielded mixed results.

Since the 1970s, the conception of TL has undergone major developments, including, for example, Kouzes and Posner's (1995) Leadership Practices Inventory (LPI) framework, Sashkin's (2004) conceptualization of transformational leadership as encompassing a leader's own behaviors, characteristics, and culture-building efforts, and Tichy and Devanna's (1990) three-act play of redesigning organizations. In non-educational contexts, the model of transformational leadership has been developed in its most mature form by Bass and his associates (e.g., Bass, 1985; Bass and Avolio, 1994). They developed a "two-factor theory" along with the Multifactor Leadership Questionnaire (MLQ) to measure transformational leadership. According to Bass and Avolio, transformational and transactional leadership are two separate ends of a continuum. Most leaders do some of both, but transformational leadership has augmented effects above transactional leadership.

While Burns saw an upward shift in the level of followers' needs as fundamental to the process of transformation, Bass (1985) suggested that the expansion of needs on the same level or even a downward shift could also be evidence of a transformation though he acknowledged that in the world of work, transformational processes usually involve the upgrading of needs. Bass also argues that in the process of transformation the change can be either good or bad, and that transactional leadership is a necessary component of transformational leadership. Bass and his colleagues placed more emphasis on the characteristics of leadership practices and their influential outcomes, rather than on the nature and the power of moral elevation of this type of leadership, when they formulated the two-factor theory of transformational leadership (Bass and Avolio, 1993).

The origin of the concept of transformational leadership is also associated with Weber (1922; 1947) coining of "charisma" from the Greek, meaning "divine gift." He defined charismatic leaders as those who use their considerable emotional appeal to direct their followers. Many scholars thereafter incorporated charisma as one of the most important components of transformational leadership. This component, as well as ideas about the "leader-hero" in some conceptualizations of transformational leadership led to critiques about the dark side of transformational leadership. Such critiques were about the ideological nature of "self-centered" "leader hero", excessive positivity, and emotional appeals to those who follow and make extra commitment without good reasons (Alvesson and Einola, 2019; Yukl, 1999). Proponents of transformational leadership (e.g., Bass and Steidlmeier, 1998), argue that true transformational leadership must be grounded in moral foundations to avoid the dark side of charismatic leaders.

The birth of transformational school leadership (TSL)

Transformational leadership was introduced into the education field in the early 90s (Hallinger, 1992; Leithwood, 1992, 1994). Leithwood and his colleagues modified and adapted earlier versions of transformational leadership (e.g., Podsakoff et al., 1990) to better suit the leadership demands found in schools and developed TSL to its fullest form over a number of years (e.g., Leithwood, 1992; Leithwood and Jantzi, 1990; 1999, 2005; Leithwood and Steinbach, 1991; Leithwood et al., 1996; Leithwood and Sun, 2009). Their most recent version (Leithwood, 2017; Leithwood and Jantzi, 2005; Leithwood and Sun, 2012; Leithwood et al., 2004) conceives of transformational leadership as including five broad categories of practices including a total of 21 specific practice: *Setting Directions*, *Developing People*, *Redesigning the Organization*, *Improving the instructional program*, and *Managing and Securing Accountability*. With the addition of *Improving the Instructional Program*, Leithwood's model of TSL became an integrated model (a combination of instructional and transformational leadership). While incorporating typical elements of other transformational leadership models developed for non-educational settings and integrating them in the first three domains, Leithwood's model of TSL was expanded to acknowledge the importance of attending to the technical core of schooling and incorporated practices central to instructional leadership. Further, unique leadership practices that support instruction in schools, such as those developing shared school goals and holding high expectations, improving school environment and instructionally helpful school culture, developing teachers, engaging communities, and building collaborative structures, were additions to each of the first three domains. None of these practices are included in TL models developed in non-school settings, or in most of the other TSL models.

The Ontario Leadership Framework (OLF) (Leithwood, 2012), built upon Leithwood's early versions of TSL, includes a full range of transformational school leadership elements. Hitt and Tucker (2016) claim that this is the most comprehensive effective school leadership framework currently available. This framework includes both transformational and instructional leadership practices, guidelines about how, where and when to enact those practices, as well as Personal Leadership Resources (cognitive, social, and psychological resources known to improve leader effects). A revision of the framework currently underway awards greater attention to leaders' ethics and values. Embedded in this integrated model are the most important elements of school leadership reported in the past three decades of research conducted from several major approaches, or schools of leadership theories, including, for example, trait theory (e.g., self-efficacy, emotional intelligence), motives and values, cognitive abilities, social and

problem-solving skills and expertise of leaders, information-processing theory, instructional leadership, transformational leadership, distributed leadership, inclusive leadership, and contextual school of leadership (see Antonakis et al., 2012 for short descriptions of these schools of leadership theories). Thus, OLF, the most fully developed TSL model developed to date, addresses the main critiques of earlier versions of transformational leadership such as the emphasis on the leader-hero, lack of consideration of the context the leader works in and with, and the ambiguity of leadership behaviors. While earlier studies in educational contexts tended to use Bass's MLQ to measure transformational leadership in schools, more recent studies in schools have tended to use Leithwood's survey.

Typical models and practices of transformational leadership

Over the past 35 years, twelve models have been used by researchers to examine and measure transformational leadership in schools (Leithwood and Sun, 2012; Sun, 2016). Three of these models were developed from evidence originally collected in non-educational settings including:

- Bass and his associates' two-factor theory measured using the Multifactor Leadership Questionnaire (MLQ) (Bass and Avolio, 1989, 1991, 1995, 2004, 2004; Avolio and Bass, 1999, 2004) and their adaptations (e.g., Edwards et al., 2010)
- Kouzes and Posner's (1993, 1995) Leadership Practices Inventory (LPI)
- Sashkin's visionary leadership measured by the Leadership Behavior Questionnaire (LBQ) (1990)

Nine transformational leadership models were developed specifically for educational settings, Leithwood's model is the most inclusive and most frequently used model, with the other eight only used in one or few studies each.

- Leithwood's TSL model measured by the Principal Leadership Questionnaire (PLQ) (e.g., Leithwood and Jantzi, 1991; 1997; 1999; 2006; Jantzi and Leithwood, 1996), early versions of the Nature of School Leadership survey (NSL) (Leithwood et al., 2001), more recent versions (Leithwood et al., 2006a,b; Leithwood, 2017), or adapted version (Day et al., 2016; Starks, 2018)
- Transformational leadership adapted to Dutch contexts (Geijsel et al., 2003; Thoonen et al., 2011; 2012 cited in Moolenaar and Sleegers, 2015; Moolenaar et al., 2010; Vermeulen et al., 2015)
- Three studies (Kuan, 2020; Huang and Hsieh, 2010; Urlick and Bowers, 2014) selected items from other school leadership surveys to measure transformational leadership
- Principal's Transformational Leadership Questionnaire (Chong-Hee No, 1994 in Ham 1999)
- Three author-constructed transformational leadership instruments (Perkins, 2018; Showalter, 2016; Wiley, 1998)

School leadership practices involved in these twelve models used in the original studies were synthesized into 13 practices (a practice is a set of leadership activities). These 13 practices, described further below, can be grouped into five categories: Setting Directions, Developing People, Redesigning the Organization, Improving Instruction, and Managing and Securing Accountability.

Setting directions

Developing a shared vision and building goal consensus

This dimension involves leadership practices that identify, develop, and articulate a shared vision that is appealing and inspiring to staff. It also includes practices related to achieving goal consensus among staff, motivating staff with challenging while achievable goals, giving staff an overall sense of purpose for their work, and monitoring and referring to school goals when staff are making decisions.

Holding high performance expectations

Holding high performance expectations means that leaders expect a high level of professionalism from staff, hold high expectations for students, and expect staff to be effective innovators (Leithwood et al., 2006a,b).

Developing people

Modeling behavior

The leadership practices involved in the various conceptualizations of modeling behavior include leaders walking the talk, providing a role model for highly ethical behavior, instilling pride, respect, and trust in staff, symbolizing success, and demonstrating a willingness to change their own practices as a result of new understandings.

Providing intellectual stimulation

The practices included in this dimension are leaders challenging staff's assumptions, stimulating and encouraging their creativity, and providing information to staff to help them evaluate their practices, refine them, and carry out their tasks more effectively.

Providing individualized support

This practice dimension involves leaders listening and attending to individual opinions and needs, acting as mentors or coaches to staff members, treating them as individuals with unique need and capacities, and supporting their professional development.

Redesigning the organization***Strengthening school culture***

The leadership practices related to strengthening school culture are leaders promoting an atmosphere of caring and trust among staff, building a collaborative school culture that reflects the school vision, and encouraging ongoing collaboration for program implementation.

Building collaborative structures

Building collaborative structures involves leadership practices that ensure that staff have adequate involvement in decisions about programs and instruction, establish working conditions that facilitate staff collaboration for planning and professional growth, and distribute leadership broadly among staff (Leithwood et al., 2006a,b). This dimension of leadership was mainly developed in Leithwood's model of TSL.

Engaging communities

Engaging communities was conceptualized and included only in Leithwood's model of TSL. The practices related to this dimension include leaders demonstrating sensitivity to community aspirations, diversity, and requests, incorporating community characteristics and values in the school, and actively encouraging parents and guardians to become involved in their children's education (Leithwood et al., 2006a,b).

Improving instruction***Focusing on instructional development***

Focusing on instructional development was conceptualized in both Leithwood's and Willey's models of TSL. It involves a broad range of leaders' practices related to planning and supervising instruction, providing instructional support, frequent and regular monitoring of school progress, and buffering staff from district or state initiatives that are potential distractions from school priorities. The development and inclusion of this dimension of leadership is a significant advancement or departure for transformational school leadership from the business sector where Bass' model was originally developed.

Managing and securing accountability***Contingent reward***

Contingent reward means that the leader rewards the follower for completing the agreed-upon work.

Management by exception

This leadership practice occurs when the leader monitors the performance of the subordinate and interferes when this performance deviates from the norm.

Providing resources

In this practice, the leader provides resources to teachers and schools to facilitate instruction, foster teacher professional development, and meet the needs of students.

Securing accountability

The OLF (Leithwood, 2012, 2017) expands the leadership model's attention to a broad policy context which holds most school leaders accountable for advancing their schools' achievement of widely valued student outcomes. This practice also includes leaders' efforts related to a supervisory and monitoring role in teaching and learning for quality assurance and accountability and managing staff (Kwan, 2020). A more detailed description of the early version of these TSL practices can be found in Leithwood (2010).

The impact of transformational leadership on school conditions

Evidence reviewed in this paper examined the effects of transformational leadership on 19 school organizational outcomes. There are significant numbers of analyses about the effects of transformational leadership on three of these school characteristics with smaller amounts of evidence about an additional 16 characteristics. Of those conditions for which there is significant evidence, Organizational Culture and Climate has been studied most (34 analyses), followed by Shared Goals and Mission (7 analyses) and Shared Decision making (6 analyses). Results indicate moderate to large effects of transformational leadership on these school

conditions. Transformational leadership appears to be most effective in achieving shared goals and nurturing cohesiveness and collaboration within schools.

This evidence demonstrates the strong contribution of transformational leadership to conditions that enable teaching and learning. Most of the individual transformational leadership practices have significant positive effects on some school conditions suggesting that transformational leaders' effects may depend on the enactment of as many of the TSL practices as possible.

The impact of transformational leadership on teachers

Evidence reviewed for this paper examined the effects of transformational leadership on 25 teacher characteristics, states or dispositions. There are significant numbers of analyses about the effects of transformational leadership on four of these characteristics or dispositions with smaller amounts of evidence about an additional 21. Of those teacher dispositions for which there is significant evidence, Teacher Commitment and Work Engagement has been studied most (31 analyses), followed by teachers' Job Satisfaction (25 analyses), Teachers-Perceived Leader Effectiveness (14 analyses), and teacher self-efficacy (11 analyses). Results indicate large effects of TSL on the first three teacher outcomes and moderate effects on teacher efficacy. TSL appears to be very influential in prompting positive individual perceptions of leaders and improving individual psychological states. This finding is generally consistent with the theory of transformational leadership conceptualized by Burns' (1978), in which the influence of transformational leadership is the elevation of motivation and morality in both the leader and followers, resulting in subordinates being motivated by the leader's vision and heightening their efforts to achieve that vision. This observation is also consistent with previous reviews (Leithwood and Sun, 2009, 2012; Leithwood and Jantzi, 2005; Chin, 2007). The large impacts of individual transformational leadership practices often arise from principals' building a shared vision, providing intellectual stimulation and individualized support, and modeling behaviors (Leithwood and Sun, 2012).

The impact of transformational leadership on students

A much smaller but growing number of studies have examined the impact of transformational leadership on students, compared with those on teachers and school conditions. Only one or few studies have examined the (mostly non-significant) effects of such leadership on attendance, college-going rates, drop-out rates, engagement, modernity, or graduation rates (e.g., Leithwood and Jantzi, 2000a,b; Wang, 2019). In contrast, 32 studies have examined the impact of transformational leadership on student achievement. While earlier studies directly measured the direct effects (relying on correlations) of transformational leadership on student achievement, more recent studies have examined the indirect effects of such leadership on student learning. This approach reflects the consensus that has been reached in the last 10 years in the field of educational leadership that school leaders' impact on student learning outcomes is largely indirect (Hallinger and Heck, 1996; Heck and Hallinger, 2014; Leithwood et al., 2004; Leithwood and Louis, 2012). A central challenge for this line of research is to locate potentially powerful mediators of leaders' effects on students and determine how leaders' can positively influence the status of those mediators. Studies addressing this challenge have used complicated statistical methods such as Structural Equation Modeling and regression mediation and/or moderation. The next section summarizes the results of research examining the direct and indirect effects of TSL on student achievement.

Direct effects of transformational school leadership on student achievement

Transformational leadership had small but significant, positive direct effects on student learning based on a meta-analysis of 24 dissertations (Leithwood and Sun, 2012). Separate meta-analyses of TSL's impact on student learning in reading and math yielded the same results. New evidence reviewed for this paper (published in eight major journals in educational leadership and *Review of Educational Research*) was consistent with this earlier evidence. These results also reflect the results of other studies and meta analyses examining the effects on students of diverse leadership models (Brown, 2001; Waters et al., 2010; Witziers et al., 2003). Chin's (2007) review of unpublished research on the effects of transformational leadership on student learning found an effect size (weighted mean r) of .49, higher than was found in our own meta-analysis of unpublished research (Leithwood and Sun, 2012).

While evidence supports the significant effect of transformational leadership on student achievement, Robinson and her associates' study (Robinson et al., 2008) report stronger effects for instructional as compared with transformational leadership. While their study reported a relationship of $ES = 0.11$ between transformational leadership and student outcomes, the average effect of instructional leadership on student outcomes was three to four times that size. It is difficult to make comparisons of the findings of our study with Robinson and colleagues for several reasons: the unknown nature of their study's mean effect sizes; the effect size calculation procedure they used, and the significance and homogeneity analysis of the effect sizes. That said, Robinson et al.'s (2008) study revealed strong average effects on student outcomes for the leadership dimensions involving promoting and participating in teacher learning and development and moderate effects for the dimensions concerned with goal setting and planning, coordinating, and evaluating teaching and curriculum, all of which are leadership practices included in Leithwood's model of transformational leadership and measured, for example, in the NSL (Leithwood et al., 2001). The impact of instructional and transformational types of leadership on student outcomes did not differ significantly in Brown's (2001) study ($d = .62, .74$).

Subsequent research about the effects on students of more fully conceptualized models of transformational leadership will no doubt clarify the relative and absolute effects of transformational leadership on students' learning. However, the practical answer to questions about relative effects is likely to reside in the development of leadership models that integrate practices associated with both instructional and transformational approaches. As Marks and Printy (2003) argue, it would be more fruitful for future research to develop such models and to identify individual leadership practices that are effective in improving student learning. Of the 13 transformational leadership practices summarized in the previous section (an integrated model of transformational leadership), Building Collaborative Structures and Providing Individualized Support are the most influential sets of practices; their direct effects are significant but small with weighted means r ranging from 0.15 to 0.17. (Leithwood and Sun, 2012).

Indirect effects of transformational leadership on student achievement

With school leadership effects now widely understood to be indirect, a growing number of studies have inquired about those variables which significantly mediate leadership effects. These are variables with more or less direct effects on students whose status can be influenced by leaders. These studies have approached indirect leadership effects by incorporating, in their research designs, (a) only mediating variables, (b) both mediating and moderating variables, or (c) only moderating variables.

The mediating variables that were included in the first group of studies were teacher collective efficacy, teacher commitment, professional community, teacher efficacy, teacher empowerment, shared school mission, and teacher job satisfaction. Among these mediators, most of these studies found that collective teacher efficacy and teacher commitment were positively and significantly related to student achievement. Studies conducted earlier using regression (e.g., Ross, 1998; Floyd, 1999) often reported a non-significant impact of transformational leadership on student achievement when mediators were put into the equation. These results suggest that the impact of transformational leadership on student learning is mostly indirect, and through mediators.

The several studies of transformational leadership effects on students that included both moderating and mediating variables reported generally larger effects than were reported in studies with simpler designs. Solomon's (2007) study, for example, reported a very large regression total R of .79 between transformational leadership, teacher commitment, teacher collective efficacy, SES, and student achievement. The effects of transformational leadership on student achievement seemed contingent upon which moderating and mediating variables were taken into account. As a second example, Nicholson's (2003) study in 146 Ohio elementary schools examined the influence of transformational leadership on student achievement as mediated by teacher collective efficacy. Teacher collective efficacy was positively related to student achievement even when controlling for SES and prior achievement. This study did not find a significant association between transformational leadership and collective teacher efficacy. However, structural equation modeling demonstrated that the transformational leadership practice of "intellectual stimulation" was directly related to math achievement; this relationship was not mediated by teacher collective efficacy. Future research that identifies the most important mediators that improve student learning effectively, and the transformational leadership practices and contextual factors or moderators that work best to improve the status of these mediators, would be very useful. Such combinations of TSL practices, mediating variables and contextual factors would uncover important contextualized paths through which transformational leadership improves student learning.

Conclusions

This synthesis of research on transformational school leadership demonstrates significant progress in clarifying the meaning and effects of this approach to leadership. Developed by many groups of scholars over decades, the various versions of transformational leadership, especially Leithwood and his colleagues' model (Leithwood et al., 2006a,b), subsume instructional and managerial forms of leadership (see the Improving the Instructional Program dimension of transformational leadership in Leithwood et al., 2006a,b for an example), ground its roots in moral foundations (see Bass, 1997; Bass and Steidlmeier, 1998 for detailed arguments), and can be participative (Bass, 1997; Channell, 2015; Thomas et al., 2020). About twelve versions of transformational leadership have guided research in school contexts. This research has examined transformational leadership effects on about 19 school outcomes, 25 teacher outcomes, and six student outcomes. Empirical research on transformational leadership in school settings has been gradually addressing the critiques of transformational leadership as promoting a "heroic" view of leadership, being excessively positive and paying insufficient attention to negative effects, being ambiguous about transformational behaviors and underlying influence processes, as well as overemphasizing dyadic processes, and (Alvesson and Einola, 2019; Yukl, 1999). That said, there are, however, at least five gaps in such progress that warrant future research. First the majority of studies were focused on a few school, teacher and student outcomes. Much more research is needed to clarify TSL's influence process on these and other outcomes.

Second, the development of transformational leadership in educational contexts has absorbed and integrated many other leadership models, a goal considered worth pursuing by a number of scholars (e.g., Antonakis et al., 2004; Hunt, 1999). The concept of TSL entails transforming teaching and learning in schools, as well as the school as an organization. Most studies of transformational leadership effects to date, however, have been based on narrow conceptions of transformational leadership (e.g., Bass and Avolio, 1995), conceptions that do not acknowledge leaders' organizational contexts and efforts to improve instruction in schools. For this reason, most studies have failed to identify or report the "real" impacts of TSL on school, teacher, and student outcomes. Especially

those instruction-related practices aimed at improving student learning need to be addressed with additional research. Future research should adopt versions of TSL that includes a broader range of practices suitable for school contexts and examine especially those TSL practices related to instruction and community building.

Third, the majority of studies have treated transformational leadership as a whole. It remains for future research to identify the relative contributions of individual TSL practices to improving school, teacher, and student outcomes. Three meta-analyses (Sun and Leithwood, 2012, 2015, 2017) illustrate the value of such research to date. For example, these reviews found that Direction-setting leadership practices (DSLPS), in aggregate, are effective in shaping a positive working environment and culture, achieving shared understanding of school goals among staff, and fostering shared decision-making in schools. Specifically, transformational school leaders' efforts to develop a shared vision in schools mostly influences teachers' individual internal states, especially teachers' judgment of leadership effectiveness and teachers' job satisfaction. School organizations with a shared vision motivate teachers to commit to and make extra efforts (Organizational Citizenship Behavior) toward that vision. DSLPs alone are not effective in building teachers' trust in school organizations. DSLPs have small and moderate impacts on improved instruction and teacher utilization of knowledge, and a close-to-large impact on various cultural outcomes.

Fourth, future research should aim to identify the critical paths or mediators through which individual transformational leadership practices have their effects on school conditions or teacher dispositions, which in turn improve student learning. Such an approach, adopting a broad range of TSL practices, would better estimate the overall impact of TSL on student learning. Using more sophisticated research methods, such as Structural Equation Modeling incorporating both important mediating and moderating variables, would uncover much more of the contextualized critical paths through which TSL improves student learning, and thus yield much more useful information to practitioners and policy makers.

Finally, more research should aim to understand what gives rise to TSL. Compared with inquiry into the nature and impact of transformational leadership, evidence about factors that give rise to TSL (the antecedents of TSL) is relatively sparser. Only two previous reviews (Leithwood and Jantzi, 2005; Sun et al., 2016) have touched on this problem. This very limited evidence identifies several predictors of TSL practices – principals' efficacy beliefs, values, formal training and emotional intelligence. Findings from this line of research can inform leadership preparation programs, the development of transformational school leaders and appointment of principals in different school contexts.

References

- Alvesson, M., Einola, K., 2019. Warning for excessive positivity: authentic leadership and other traps in leadership studies. *Leader. Q.* 30, 383–395.
- Antonakis, J., Cianciolo, A.T., Sternberg, R.J., 2012. Leadership: past, present, and future. In: Day, D., Antonakis, J. (Eds.), *The Nature of Leadership*. Sage, Thousand Oaks, CA, pp. 1–20.
- Avolio, B.J., Bass, B.M., 1999. Re-examining the components of transformational and transactional leadership using the multifactor leadership. *J. Occup. Organ. Psychol.* 72, 441–462.
- Avolio, B.J., Bass, B.M., 2004. Multifactor Leadership Questionnaire, third ed. Mindgarden, Redwood City, CA.
- Bass, B.M., 1985. Leadership and Performance beyond Expectations. The Free Press, New York.
- Bass, B.M., 1997. The ethics of transformational leadership. In: *KLSP: Transformational Leadership, Working Papers*. Academy of Leadership Press. Available at: http://www.academy.umd.edu/publications/klspdocs/bbass_p1.htm. (Accessed 26 December 2016).
- Bass, B.M., Avolio, B.J., 1989. Manual: The Multifactor Leadership Questionnaire. Consulting Psychologists Press, Palo Alto, CA.
- Bass, B.M., Avolio, B.J., 1991. The Multifactor Leadership Questionnaire: Form 5x. Center for Leadership Studies, State University of New York, Binghamton, NY.
- Bass, B.M., Avolio, B.J., 1993. Transformational leadership: a response to critiques. In: Chemers, M.M., Ayman, R. (Eds.), *Leadership Theory and Research*. Academic Press, Toronto, pp. 49–80.
- Bass, B.M., Avolio, B.J., 1994. Improving Organizational Effectiveness through Transformational Leadership. Sage, Thousand Oaks, CA.
- Bass, B.M., Avolio, B., 1995. Multifactor Leadership Questionnaire: Duplicate Set: Rate Form, for MLQ for 5x-Short. Mind Garden, California.
- Bass, B.M., Avolio, B.J., 2004. Multifactor Leadership Questionnaire: Manual and Sampler Set, third ed. Mind Garden, Palo Alto, CA.
- Bass, B.M., Steidlmeier, P., September 24, 1998. Ethics, Character, and Authentic Transformational Leadership. Available at: <http://cls.binghamton.edu/BassSteid.html>. (Accessed 3 June 2002).
- Brown, L.I., 2001. A Meta-Analysis of Research on the Influence of Leadership on Student Outcomes. Doctoral dissertation. Virginia Polytechnic Institute and State University, the United States.
- Burns, J.M., 1978. Leadership. Harper and Row, New York.
- Channell, J.G., 2015. The Transformational Leader in the Twenty-First Century: The Role of the Effective Classroom Leader in Student Achievement. Northcentral University.
- Chin, J.P., 2007. Meta-analysis of transformational school leadership effects on school outcomes in Taiwan and the USA. *Asia Pacific Education* 8 (2), 166–177.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes. *Educ. Adm. Q.* 52 (2), 221.
- Day, D., Antonakis, J., 2012. Leadership: past, present, and future. In: Day, D., Antonakis, J. (Eds.), *The Nature of Leadership*. Sage, Thousand Oaks, CA, pp. 1–20.
- Edwards, J.R., Knight, D.K., Broome, K.M., 2010. The development and validation of a transformational leadership survey for substance use treatment programs. *Subst. Use Misuse* 45, 1279–1302.
- Floyd, J.E., 1999. An Investigation of the Leadership Style of Principals and its Relation to Teachers' Perceptions of School Mission and Student Achievement. Dissertation Abstracts International, 6009A, AAI9946412.
- Geijsel, F.P., Sleegers, P.J.C., Leithwood, K., Jantzi, D., 2003. Transformational leadership effects on teachers' commitment and effort toward school reform. *J. Educ. Adm.* 41, 229–256.
- Hallinger, P., 1992. The evolving role of American principals: from managerial to instructional to transformational leaders. *J. Educ. Adm.* 30 (3), 35–50.
- Hallinger, P., Heck, R.H., 1996. Reassessing the principal's role in school effectiveness: a review of empirical research. *Educ. Adm. Q.* 32, 5–44.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681.
- Ham, S., 1999. A Study of the Relationship between Principal Leadership and School Climate in Korean Secondary Schools. Dissertation Abstracts International, 60(08)A, AAI9945411.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86 (2), 531–569.

- Huang, W.-Y., Hsieh, C.-L., 2010. A study on relationships between principal's transformational leadership and physical educational teachers' teaching performance in Taipei elementary schools. *Int. J. Oral Implant.* 3 (2), 80–96.
- Hunt, J.G., 1999. Transformational/Charismatic leadership's transformational of the field: an historical essay. *Leader. Q.* 102, 129–144.
- Jantzi, D., Leithwood, K., 1996. Toward an explanation of variation in teachers' perceptions of transformational school leadership. *Educ. Adm. Q.* 32 (4), 512–538. <https://doi.org/10.1177/0013161X9603200404>.
- Kouzes, J.M., Posner, B.Z., 1993. *Leadership Practices Inventory. A Self-Assessment and Analysis*, expanded ed. Jossey-Bass, San Francisco, CA.
- Kouzes, J.M., Posner, B.Z., 1995. *The Leadership Challenge: How to Keep Getting Extraordinary Things Done in Organizations*, Revised ed. Jossey-Bass, San Francisco, CA.
- Kwan, P., 2020. Is transformational leadership theory passé? revisiting the integrative effect of instructional leadership and transformational leadership on student outcomes. *Educ. Adm. Q.* 56 (2), 321–346.
- Leithwood, K., 1992. The move towards transformational leadership. *Educ. Leader* 495, 8–12.
- Leithwood, K., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30, 498–518.
- Leithwood, K., 2010. Transformational school leadership. In: *Encyclopedia*. Elsevier Ltd.
- Leithwood, K., 2012. Ontario Leadership Framework (Report to the Ontario Ministry of Education. Ontario, Canada).
- Leithwood, K., 2017. The Ontario leadership framework: successful school leadership practice and personal leadership resources. In: Leithwood, K., Sun, J.-P., Pollock, K. (Eds.), *How School Leadership Influences Student Learning: The Four Paths*. In: Leithwood, K. (Ed.), *Studies in Educational Leadership, Series Ed.*, vol. 22. Springer, New York, pp. 31–45.
- Leithwood, K., Aitken, R., Jantzi, D., 2001. *Making School Smarter: Leading with Evidence*. Corwin Press, Thousand Oaks, CA.
- Leithwood, K., Aitken, R., Jantzi, D., 2006a. *Making School Smarter: Leading with Evidence*, 3 ed. Corwin Press, Thousand Oaks, CA.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006b. *Seven Strong Claims about Successful School Leadership*. National College for School Leadership, Nottingham, England.
- Leithwood, K., Jantzi, D., 1990. Transformational leadership: how principals can help reform school cultures. *Sch. Effect. Sch. Improv.* 14, 249–280.
- Leithwood, K., Jantzi, D., 1991. The Primary Program School Survey. OISE Center for Leadership Development and Byron Dart, Ontario, Canada.
- Leithwood, K., Jantzi, D., 1997. Explaining variation in teachers' perceptions of principals' leadership: a replication. *J. Educ. Adm.* 35 (4), 312–331.
- Leithwood, K., Jantzi, D., 1999. Transformational school leadership effects: a replication. *Sch. Effect. Sch. Improv.* 10 (4), 451–479.
- Leithwood, K., Jantzi, D., 2000a. The relative effects of principal and teacher sources of leadership on student engagement with school. *Educ. Adm. Q.* 35, 679–706.
- Leithwood, K., Jantzi, D., 2000b. The effects of transformational leadership on organizational conditions and student engagement with school. *J. Educ. Adm.* 38, 113–129.
- Leithwood, K., Jantzi, D., 2005. A review of transformational school leadership research. In: Paper Presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Leithwood, K., Jantzi, D., 2006. Transformational school leadership for large-scale reform: effects on students, teachers, and their classroom practices. *Sch. Effect. Sch. Improv.* 17 (2), 201–227.
- Leithwood, K., Louis, K.S., 2012. *Linking Leadership to Student Learning*. John Wiley & Sons, San Francisco, CA.
- Leithwood, K., Louis, K., Anderson, S., Wahlstrom, K., 2004. *How Leadership Influences Student Learning: A Review of the Evidence Linking Leadership to Student Learning*. the Wallace Foundation, New York.
- Leithwood, K., Steinbach, R., 1991. Indicators of transformational leadership in the everyday problem solving of school administrators. *J. Person. Eval. Educ.* 4, 221–244.
- Leithwood, K., Sun, J.P., 2009. Transformational school leadership effects on schools, teachers and students. In: Hoy, W., Dipaola Series, M. (Eds.), *Research and Theory in Educational Administration: Vol. 8. Studies in School Improvement Programs*. Information Age, Charlotte, NC, pp. 1–22.
- Leithwood, K., Sun, J.-P., 2012. The nature and effects of transformational school leadership: a meta-analytic review of unpublished research. *Educ. Adm. Q.* 483, 387–423.
- Leithwood, K., Tomlinson, D., Genge, M., 1996. Transformational school leadership. In: Leithwood (Ed.), *International Handbook of Educational Leadership and Administration*. Kluwer Academic, Netherlands, pp. 785–840.
- Marks, H.M., Printy, S.M., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Moolenaar, N.M., Daly, A.J., Slegers, P., 2010. Occupying the principal position: examining relationships between transformational leadership, social network position, and schools' innovative climate. *Educ. Adm. Q.* 46 (5), 623–670.
- Moolenaar, N.M., Slegers, P.J.C., 2015. The networked principal. *J. Educ. Adm.* 53 (1), 8–33.
- Nicholson, M.R., 2003. Transformational Leadership and Collective Efficacy: A Model of School Achievement. *Dissertation Abstracts International*, 64(06)A, AAI3093682.
- Perkins, C.C., 2018. *Effective Leadership Styles as Perceived by Staff during the Recovery Phases after a School Shooting: A Comparative Study*. Northcentral University.
- Podsakoff, P., MacKenzie, S., Moorman, R., Fetter, R., 1990. Transformational leaders' behaviors and their effects on followers' trust in leader, satisfaction and organizational citizenship behavior. *Leader. Q.* 1 (2), 107–142.
- Robinson, V.M.J., Lloyd, C.A., Lloyd, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 445, 635–674.
- Ross, A.T., 1998. *Exploring Connections Among Teacher Empowerment, Teacher Efficacy, Transformational Leadership, and Student Achievement*. *Dissertation Abstracts International*, 5909A, AAI9906642.
- Sashkin, M., 1990. *The Visionary Leader: Leadership Behavior Questionnaire*. Human Resource Development Press, Amherst, MA.
- Sashkin, A., 2004. Transformational leadership approaches: a review and synthesis. In: Antonakis, J., Cianciolo, A.T., Sternberg, R.J. (Eds.), *The Nature of Leadership*. Sage Publications, Inc, pp. 171–196.
- Showalter, S.M., 2016. *Transformational Leadership Study of Catholic Inner-City Schools in the United States*. The Catholic University of America.
- Solomon, C., 2007. *The Relationships Among Middle Level Leadership, Teacher Commitment, Teacher Collective Efficacy, and Student Achievement*. Ph.D. dissertation. University of Missouri - Columbia, United States – Missouri. Retrieved July 24, 2010, from *Dissertations and Theses: Full Text*. Publication No. AAT 3322742.
- Starks, D.M., 2018. *A Study of the Relationships Among the Transformational Leadership Practices of the Michigan High School Principal, School Culture, and Student Achievement*. Eastern Michigan University.
- Sun, J.-P., 2016. The nature, impacts and antecedents of transformational school leadership: a review. In: Weinstein, J. (Ed.), *Educational Leadership in Schools: Nine Perspectives*. University Diego Portales Press, Santiago, Chile.
- Sun, J.-P., Chen, X.-J., Zang, S.-J., 2016. A review of research evidence on the antecedents of transformational leadership. *Educ. Sci.* 7 (1), 15.
- Sun, J.-P., Leithwood, K., 2012. Transformational school leadership effects on student achievement. *Leader. Pol. Sch.* 114, 418–451.
- Sun, J.-P., Leithwood, K., 2015. Direction-setting school leadership practices: a meta-analytical review of evidence about their influences. *Sch. Effect. Sch. Improv.* 26 (4), 499–523. <https://doi.org/10.1080/09243453.205.1005106>.
- Sun, J.-P., Leithwood, K., 2017. Calculating the power of alternative choices by school leaders for improving student achievement. *Sch. Leader. Manag.* 37 (1–2), 80–93.
- Thomas, L., Tuytens, M., Devos, G., Kelchtermans, G., Vanderlinde, R., 2020. Transformational school leadership as a key factor for teachers' job attitudes during their first year in the profession. *Educ. Manag. Adm. Leader* 48 (1), 106–132.
- Tichy, N.M., Devanna, M.A., 1990. *The Transformational Leader*. John Wiley and Sons, Toronto.
- Urlick, A., Bowers, A.J., 2014. What are the different types of principals across the United States? A latent class analysis of principal perception of leadership. *Educ. Adm. Q.* 50 (1), 96–123.
- Vermeulen, M., Acker, F.V., Kreijns, K., Buuren, H.V., 2015. Does transformational leadership encourage teachers' use of digital learning materials. *Educ. Manag. Adm. Leader* 43 (6), 1006–1025.

- Wang, S., 2019. School heads' transformational leadership and students' modernity: the multiple mediating effects of school climates. *Asia Pac. Educ. Rev.* 20 (3), 329–341.
- Waters, T., Marzano, R.J., McNulty, B., 2010. McREL's Balanced Leadership Framework: Developing the Science of Educational Leadership. Available at: <http://www.nga.org/cda/files/0404mcrel.pdf>. (Accessed 1 March 2010).
- Weber, M., 1922. 1947. *Theory of Social and Economic Organization*. Oxford University Press, New York (translated by Henderson, A. M. and Parsons, T. Originally published in 1922 in German under the title *Wirtschaft und Gesellschaft* (available online).
- Wiley, S.D., 1998. School Leadership and Professional Community: Effects on Student Achievement. *Dissertation Abstracts International*, 5910A, AA19910535.
- Witziers, B., Bosker, R.J., Krüger, M.L., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 393, 398–425.
- Yukl, G., 1999. An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *Leader. Q.* 10, 285–305.

Research on successful school leadership

Christopher Day^a, Jingping Sun^b, and Christine Grice^c, ^a University of Nottingham, Nottingham, United Kingdom; ^b The College of Education, University of Alabama, Tuscaloosa, AL, United States; and ^c Sydney School of Education and Social Work, The University of Sydney, NSW, Sydney, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	62
A brief history of leadership research	62
Major models of school leadership: strengths and limitations	63
Instructional leadership	64
Transformational leadership	64
Moral leadership	64
Inclusive and culturally responsive leadership	65
Distributed leadership	65
Five leadership frameworks	65
Leadership influence and impact	66
IMPACT project (England)	67
Australia	67
The international successful school principalship project: embracing complexity, acknowledging culture	67
Complexity theory	68
Cultural Differences	68
Future directions	69
References	70

Introduction

This article provides a critical synthesis of what selected research has shown internationally about who effective and successful school leaders are and what they do in order to achieve and sustain success in a range of school contexts, within and across national jurisdictions. It focuses on the most significant models, frameworks and pathways, and international research projects identified by researchers as furthering understandings of leadership, differentiating between those who are “effective” and “successful”. Alongside this, the article highlights the importance of acknowledging the influence of broader national cultures, and the need to understand schools as complex adaptive systems.

With few exceptions, research on school effectiveness and improvement has found that who school leaders are, what they do, and how and when they do it matters, and makes a significant difference to school culture, teacher motivation, quality, retention, and student learning, achievement and wellbeing trajectories. A range of research internationally has shown that high quality school leadership is second only to classroom teaching as a direct influence on pupil learning (Leithwood et al., 2006), and that it contributes, albeit mostly indirectly, through establishing key school conditions, and influencing the values, dispositions, aspirations, expectations, and practices of teachers (e.g., Day et al., 2000; Heck and Hallinger, 1999, 2014; Leithwood and Seashore-Louis, 2012; Marzano et al., 2005; Robinson et al., 2008). Because teachers are key contributors to student learning, achievement, and wellbeing, school leadership is most effective when it improves or sustains their curriculum knowledge, classroom skills, commitment, capabilities for change and development, active engagement, and capacity for resilience. However, just as research has shown that there are variations in teachers’ work, lives and effectiveness, so the same applies to the work, lives and effectiveness of school leaders. This article will focus on what research internationally has shown about who effective and successful school leaders are, what the similarities and differences between them are, and what they do in order to achieve success. It will also discuss strengths and limitations of research methodologies.

A brief history of leadership research

Research on school leadership in the mid 20th century was dominated by psychological paradigms and business leadership writers. Early research derived from these perspectives identified broad types of leadership, for example, “authoritative”, “democratic” and “laissez faire” leadership (Lewin et al., 1939). Later, leadership personality “traits”, and “styles” were identified. “Charismatic” and “heroic” leadership were identified as important (Parry and Bryman, 2006). Later, contrasting models comparing “transformational” with “transactional” leadership, derived from research on businesses by Burns (1978) also became popularized (e.g., Bass, 1985; Bass and Avolio, 1994). Over the last three decades especially, there has been an upsurge of researcher interest in school leadership. This coincided with the rise of new “neoliberal”, results-driven agendas in the public services in which teaching

and learning in schools have become more transparent, and school leaders have more accountability for raising the academic levels of attainment of all students, and alongside this, attending to their social and personal wellbeing. As “audit” cultures of external monitoring and evaluation became embedded (Power, 2019) a by-product of societies in which levels of trust in society and teachers in particular diminished, policy makers began to intervene in education more actively in order to solve challenges of international economic competition and social fragmentation. Researchers internationally began to conduct much more empirically grounded, nuanced portraits of “effective” and “successful” school leaders, and much of this found that previous research findings were less relevant to the increasingly complexities of school leadership and management in this century (Northouse, 2013).

Although this concentration of research has increased exponentially, educational researchers’ unprecedented interest in and pursuit of the “best,” most “effective” school leadership can be dated back to the Effective Schools movement in North America and the UK in the 1970s. In response to the assumption that schools have little effect on student achievement (e.g., James and Coleman, 1968; Jencks et al., 1972), researchers identified about a dozen essential qualities of effective schools, where students of all backgrounds, regardless of socio-economic status or family background, were successfully educated. They claimed that among those factors necessary to student success that schools can control is *effective* school leadership (Purkey and Smith, 1983; Downer, 1991; Sammons et al., 1995).

The school effectiveness movement emphasizes the importance of school leadership for the improvement of student outcomes. This ideal for those who seek to make a difference as educational leaders. However, school effectiveness research potentially ignores the complexity of school culture, narrowly conceptualizing pedagogy as a set of homogenous teaching techniques that can be applied in a range of school cultures (Hattie, 2009), rather than seeing pedagogy as a holistic educational act that is purpose driven and takes into account student diversity and different school cultures and values (Alexander, 2008). As Alexander (2008) explains:

Teaching is presented as value-neutral, content free and entirely devoid of the dilemmas of value and circumstance that confront real teachers daily. It conceals the technical deficiencies of the research and implies a degree of homogeneity in schools, classrooms and lessons that cannot be sustained empirically (p. 31).

Successful leadership literature has provided answers for educational leaders beyond the school effectiveness movement, in that it assumes that student outcomes represent more than measured academic attainment; that these are related to the quality of teachers and teaching and that the quality of school leadership is key (Pont et al., 2008). Research shows that while teacher effectiveness does not necessarily depend directly on the effectiveness of the principal, principal vision, values, structures, cultures and relationships make a difference. (Robinson et al., 2008; Day, 2017).

Stemming from the original work on effective schools, which pointed to the pivotal role of designated leaders, an increasing, if diverse, flow of research subsequently emerged, principally in Western nations, with school principals as the chief focus of interest. This now incorporates both psychologically and sociologically oriented research, as researchers recognized the importance of the social contexts as an influencing factor in schools which have become “policy mediators”, and, within this, the pivotal role of principals. So many “models” emerged that one Australian researcher remarked that the “myth” of individualism (typified by “charismatic” and “heroic” leadership descriptors) was being perpetuated:

Within the literature on school leadership there also remains an overwhelming endorsement of one-style-fits-all leadership, often accompanied with varying degrees of evangelical zeal. Promoted are a range of adjectival leaderships: “democratic” leadership, “strategic” leadership, “breakthrough” leadership, and so on.

Mulford (2008, p. 38).

There are many more adjectival models (for example “servant” leadership, “teacher” leadership, “collaborative” leadership, and “managerial” leadership), but for reasons of space these cannot all be included. The next part of this article will focus on those which most research reports have reported as being associated with successful student outcomes. In doing so, it will highlight their present strengths and limitations, finally pointing forward to the use of complexity theory perspectives as a means of transcending the “one size fits all” leadership “models”, “perspectives”, “pathways” and “frameworks”. It will also note that most of these are based on research in Western cultures, and point to the need for more research to take account of different cultural inheritances internationally.

Major models of school leadership: strengths and limitations

About seven major school leadership “adjectival” models have been developed since the Effective Schools movement. These include instructional, transformational, moral, participative, managerial, and contingent forms of school leadership (see more details about these models in Leithwood and Duke, 1999), as well as, more recently, distributed (Gronn, 2002; Leithwood and Jantzi, 2000; Spillane, 2017; Spillane et al., 2001) and inclusive or culturally responsive leadership (Ryan, 2006; Khalifa et al., 2016). These

last two may be more accurately represented as focusing upon social justice issues, and would be placed within the previous six. Among these, instructional and transformational models of leadership have been the most empirically studied, and have been advocated as promising approaches to understanding principals' contributions to improved student achievement and school reform.

Instructional leadership

Emerging from the research on "instructionally effective elementary schools" (e.g., Edmonds, 1979), this leadership model has undergone four decades of development and a variety of conceptualizations (see major developments in Andrews and Soder, 1987; Bossert et al., 1982; Hallinger and Heck, 1996; Hallinger and Murphy, 1985; Heck et al., 1990; Heck and Moriyama, 2010; Heck and Hallinger, 1999; Leithwood et al., 1990; Leithwood and Montgomery, 1982). To date, the model that has been clearly defined and reliably measured was developed by Hallinger and Murphy (Hallinger, 2005; Hallinger and Murphy, 1985; Murphy, 1990). Three dimensions with 10 leadership functions were conceptualized: *Defining the School's Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate*. Instructional leadership is claimed to influence the quality of school outcomes through the alignment of school structures and classrooms by means of the school's culture, and through modeling rather than through direct supervision or evaluation of teaching (Hallinger and Heck, 1996; Heck and Hallinger, 1999).

Whilst this "model" of school leadership is attractive and widely recognized amongst policy and practice communities, it is limited in its contribution to knowledge of what makes leadership successful. This is in part because it plays into the policy driven primary importance of functional outcomes, and in part because it assumes that better teaching will necessarily lead directly to better student outcomes in tests and examinations. In addition, even its advocates have found that, although emphasis on improving instruction is educationally meaningful, its direct impact is small.

Hallinger (2015), Robinson et al. (2008).

Transformational leadership

Transformational leadership (TSL) was introduced into the education field in the early 1990s to meet the demands of school restructuring (Hallinger, 1992; Leithwood, 1992; 1994; Podsakoff et al., 1990). Leithwood and his research colleagues developed transformational school leadership to its fullest form over a number of years (e.g., Leithwood and Jantzi, 1999, 2005; Leithwood and Steinbach, 1991; Leithwood et al., 1996; Leithwood and Sun, 2009). Their most recent version (Leithwood, 2017; Leithwood and Sun, 2012; Leithwood et al., 2004) expanded the concept of transformational school leadership to include five broad categories of practices with a total of thirteen specific sets of practice: *Setting Directions, Developing People, Redesigning the Organization, Improving the Instructional Program, and Managing and Securing Accountability*. The addition of leadership practices unique to a school setting that support instruction in schools, build an instructionally helpful school culture and collaborative structures, develop teachers, and engage communities is a major departure from the earlier models of transformational leadership developed in non-educational settings (e.g., Bass and Avolio, 1989), and from most other transformational school leadership models.

Transformational school leadership has been reported to be most effective in prompting positive individual perceptions of leaders, improving individual psychological states, achieving shared goals, nurturing cohesiveness and collaboration within schools, improving school culture and conditions, and having a small but significant impact on student learning (Leithwood and Sun, 2012; Sun, 2016; Sun and Leithwood, 2012). However, because the majority of findings on TSL impacts have been based on more narrowly conceptualized models, its actual impact on the many outcomes now required by schools, especially teaching and learning, is not clear yet.

Nevertheless, transformational leadership is a more inclusive model of leadership than Instructional. It subsumes Instructional leadership within its key strands of "Developing people" and "Improving the instructional programme", and goes beyond this, though it is important to note that this is disputed by adherents to Instructional leadership models.

Moral leadership

Implicit within transformational leadership is the importance attached to leadership values. Since the call from Greenfield (1976, 1978) for a value-laden inquiry into educational administration in the 1970s, a group of scholars has been developing research based on values theory (e.g., Hodgkinson, 1991; Sergiovanni, 1992; Starratt, 1994; Begley, 1996, 2003; Willower, 1997). Long ago, Hodgkinson (1991) argued, "Values constitute the essential problem of leadership—If there are no value conflicts then there is no need for leadership" (p. 11). Two central issues in moral leadership are "the nature of the values used by leaders in their decision-making, and how conflicts among values can be solved" (Leithwood et al., 1999, p. 10). In the last four decades, scholars have approached moral leadership in four ways: ethical reasoning; ethics as a way of informing perception and understanding; the purpose of education and the roles of educators; the establishment of caring schools; and naturalistic inquiry or valuation (Sun, 2011). The outcome of moral leadership is claimed to be morally justified decision-making, responsive, and democratic schools and classrooms (Leithwood and Duke, 1999; Noddings, 2013). Recently scholars have begun to develop measures of ethical leadership and its effects (e.g., Langlois et al., 2014).

It may be argued that moral leadership exists within the value systems of the successful application of both Instructional and Transformational leadership, rather than being a separate “model” in its own right.

Inclusive and culturally responsive leadership

Inclusive school leadership, arguably closely embedded in moral leadership values which include a belief in equity, was developed from the perspectives of critical theory and social justice. The first detailed inclusive leadership framework for schools was developed by Ryan (2006). According to him, inclusive leadership consists of a number of distinct practices, including advocating for inclusion, educating participants, developing critical consciousness, nurturing dialog, emphasizing student learning and classroom practice, adopting inclusive decision and policy-making strategies, and incorporating whole-school approaches. Expected outcomes of this kind of leadership include, increased awareness of injustice and inequity, increased emphasis on inclusive student learning, and enhanced instructional capacities and professional commitment (Ryan, 2006). A group of scholars (e.g., Gay, 2010; Johnson, 2006; Webb-Johnson and Carter, 2007) have extended this line of research by exploring *culturally responsive school leadership*, advocating culturally responsive classrooms and teaching, arguing that reforming and transforming all aspects of the educational enterprise, such as funding, policymaking and administration are needed, so that schools can be culturally responsive, especially with minority students. Culturally responsive leadership behaviors include: developing culturally responsive teachers; promoting a culturally responsive/inclusive school environment which engages students, parents, and indigenous contexts (Khalifa et al., 2016).

So far, inquiry into inclusive or culturally responsive school leadership is still at the conceptual stage. More empirical evidence is needed to understand its presence in successful leaders, and its influence on student or other school outcomes.

Distributed leadership

Although not a model in its own right, but rather, as with moral, and inclusive and culturally responsive, expressive of values held by successful transformational or instructional leaders, a distributed perspective of school leadership has been introduced and developed over the last two decades (Leithwood et al., 2009; Spillane et al., 2001, 2004; Spillane, 2017). Building upon work from other disciplines (e.g., Gronn, 2002), scholars have defined the concept of distributed leadership in different ways. Leithwood and his associates (Leithwood et al., 2009), for example, conceptualized distributed leadership as being exhibited in four forms based on leadership coordination: planful alignment, spontaneous alignment, spontaneous misalignment, and anarchic misalignment. Planful alignment is claimed to be more effective than the other three forms of leadership distribution, for example, in developing a safe and inclusive school environment (Anderson et al., 2009), and improving student learning (Leithwood et al., 2010). Day and his colleagues in a national mixed methods study found that planful distribution of leadership was progressive, associated with the building of trust by principals, took time, and is one of a number of factors which contribute, indirectly but powerfully to improvements in students’ attainment (Day et al., 2011).

Premised on activity and distributed cognition theories, Spillane and his associates (Spillane et al., 2001, 2004), were among the first scholars to apply distributed perspective as a frame for studying leadership practice. They focused on the productive and collegial interaction of all stakeholders, and argued that the conceptualization of the distributed leadership is not just about the actions of individual leaders, but rather, about *interactions* among leaders followers, and the situation. Further, he (Spillane, 2017) conceptualized instruction as being collective, situated, and distributed practice, not equivalent to a teacher’s behavior or a function of her/his knowledge and skill, but rather seen in the form of interactions between teachers and students related to particular intellectual material and to aspects of the situation. The role of teachers in leadership, and formal leadership roles in the spaces between school-wide senior leaders and classroom teaching colleagues have also been examined more in depth, as reflected from by two lines of research: teacher leadership (e.g., Muijs and Harris, 2003), and more recently, middle leadership (Day and Grice, 2019).

“Distributed” leadership focuses upon an important part of leadership which is likely to lead to success and thus is more finely delineated than the broader term of “effective”. Moreover, as with “moral” and “culturally responsive” leadership, distributed leadership is not a model but a value and a strategy; and would fit within the successful application of the “transformational” leadership model. As currently conceptualized, it does not take into account the influence of different educational histories or cultures, for example the prevailing “power-distance” culture in many Asian societies. Nevertheless, it does highlight the importance of recognizing that successful principals and the schools which they lead do not readily “fit” within any given model of successful leadership.

Five leadership frameworks

Several integrated, overarching school leadership frameworks have been developed in this century based on systematic reviews. We highlight five of these:

1. An extended version of Leithwood’s earlier work is the Ontario Leadership Framework (OLF) (Leithwood, 2012, 2017). Hitt and Tucker (2016) claimed that this is the most comprehensive school leadership framework to date. It not only includes *practices*

conceptualized in different leadership models such as transformational, instructional, and using distributed leadership practices, but also *personal resources* (cognitive, social, and psychological resources, along with values and ethics), and the *interaction* of leadership behaviors with contextual factors (i.e., information about how, where and when to enact those practices), the most important elements identified from different approaches to leadership in the past three decades. In this sense, it is the most holistic framework for capturing the complexities of successful school leadership to date.

2. Hitt and Tucker (2016) identified and compared three existing, noteworthy leadership frameworks: the OLF, the Learning-Centered Leadership Framework by Murphy et al. (2006), and the Essential Supports Framework derived by Sebring et al. (2006). They identified and classified 28 practices from these frameworks into five domains, and proposed a “United” Framework. These five domains are: establishing and conveying the vision; facilitating a high-quality learning experience for students; building professional capacity; creating a supportive organization for learning; and connecting with external partners.
3. A third major leadership framework is “Balanced” Leadership, consisting of 21 leadership responsibilities practices significantly correlated with student achievement identified by Waters and his colleagues (Waters et al., 2003) through a meta-analysis. It includes 21 leadership responsibilities, some of which are effective for first-order (direct) changes and the others are effective for second-order (indirect) change.
4. Robinson and her associates in New Zealand (Robinson et al., 2008) identified and proposed a school leadership framework based on a “best evidence synthesis” of selected quantitative studies on effective school leadership internationally. This framework consists of five domains of leadership practices found to influence student achievement effectively: establishing goals and expectations; resourcing strategically; planning, coordinating, and evaluating teaching and the curriculum; promoting and participating in teacher learning and development; and ensuring an orderly and supportive environment.

Leithwood and his colleagues (Leithwood et al., 2010, 2020) identified most promising empirically tested mediating school conditions from research evidence drawn from schools in Ontario, Canada. These were classified into four categories, and populated on Four Paths: Rational, Emotions, Organizational, and Family. School leaders, they claim, influence student learning through these four paths. Variables on the Rational Path are rooted in teachers’ knowledge and skills related to curriculum, teaching, and learning—the technical core of schooling. The Emotional Path includes teachers’ feelings, dispositions, or affective states, both individually and collectively. Variables on the Organizational Path include features of schools that structure the relationships and interactions among school members including, for example, cultures, policies, and standard operating procedures. Variables on the Family Path include those that reflect family expectations for their children, their culture and their support for their children, and community orientations toward school and general education. The variables on each of the four paths are potentially alterable by school leadership.

Leadership influence and impact

While these and hundreds of other, smaller scale studies, each applying a specific model of leadership, or interpreted framework that, have reported the influence of school leadership on a variety of school outcomes, the impact of school leadership on student learning has remained the most difficult to examine conceptually and methodologically due to its indirect nature. Research on school improvement and leadership over the last 30 years has identified significant school-level contributors to student learning, such as academic press and teacher collective efficacy (e.g., DiPaola and Hoy, 2005; Hoy and Miskel, 2013), and a small number of combinations of such variables (e.g., Bryk et al., 2010; Heck and Hallinger, 2014; Mitchell et al., 2016; Sebastian et al., 2017). These variables can be shaped significantly by school leadership. Very limited quantitative evidence shows important indirect impacts of school leadership (mostly transformational, and/or instructional models, or OLF) on student achievement through such mediating variables, moderated, or not. (e.g., Leithwood et al., 2010, 2017). As reported in the largest, five year mixed methods educational leadership project, funded by the Wallace Foundation, based on data from thousands of teachers in 180 schools in eight states in the USA, schools leaders have an impact on student achievement primarily through their influence on teachers’ motivation and working conditions. However, the study found that their influence on teachers’ knowledge and skills produced less direct impact on student achievement. Rather, it was collective leadership, particularly from teacher teams, parents, and students that, was claimed to have a stronger influence on student achievement than individual leadership (Leithwood and Louis, 2012).

While each of these models, perspectives and frameworks contributes to understandings of the qualities, values and capabilities (knowledge, skills) that leaders need, none of them alone capture how leaders assure student success over time.

Only a few studies have used both qualitative and quantitative longitudinal data to examine the link between school leadership, teacher and teaching effectiveness, and change in student learning (Day et al., 2011; Mulford and Silins, 2003). These two large scale, mixed methods researches, in Australia and England, and extensive case study research carried out by members of the International Successful School Principalship Project (ISSPP) over a twenty year period, have succeeded in providing substantive evidence of the “how” and “when”, of successful principalship, as well as the influence of cultures and contexts.

IMPACT project (England)

Based on student outcome data over at least a three year consecutive period showing improvement under the leadership of the same principal, national survey data including responses from 740 heads and 1775 key staff, and multi-perspective in-depth case studies of a subsample of 20 schools, Day and his associates (Day et al., 2016) investigated associations between principalship in effective and improving 1549 primary and 1140 secondary schools in England and student outcomes as defined (but not confined) by their national examination and assessment results over 3 years. This is most representative and comprehensive study using mixed methods and national longitudinal data in a single study. They found that successful principals directly and indirectly achieve and sustain improvement over time through values-led leadership which combines both transformational and instructional leadership strategies in ways which demonstrate consistency of values, possession and enactment of particular qualities, and strategies. Their research demonstrates that schools' abilities to improve and sustain effectiveness over the long term are not primarily the result of the principals' leadership "style" or "personality traits," but of their understanding, discernment and diagnosis of the school's needs and their application of clearly articulated, organizationally shared educational values. The research found that: (i) certain core leadership practices are effective in improving student learning across countries and across settings; (ii) successful principals are all values -led, possessing similar core attributes and qualities; (iii) they "layer" and enact multiple combinations and accumulations of key practices are strategically applied in different contexts for different periods of time, and progressively embedded in the school's work, culture, and achievements; (iv) the ways in which leaders apply leadership practices demonstrate responsiveness to, rather than dictation by, the contexts in which they work; and (v) they distribute leadership progressively according to levels of trust.

Australia

The Leadership for Organizational Learning and Student Outcomes (LOLSO) Research Project (Mulford and Silins, 2003) extended the understandings of school reform initiatives that aim to change school practices with the intention of supporting enhanced student learning (defined more broadly than academic achievement). A four year, longitudinal project, involving 3500 students and 2500 of their teachers and head teachers in 96 Australia secondary schools, it demonstrated that leadership that makes a difference to a high school has a community focus, sense of being valued among staff, and organizational leadership, and is transformational and distributive. It was one of the earliest research large scale projects outside of the USA and England that confirmed that school leadership influences students indirectly, by contributing to organizational learning, which in turn influences what happens in the core business of the school—the quality of teaching and learning.

The international successful school principalship project: embracing complexity, acknowledging culture

There are numerous global educational leadership projects that have focused on effective leadership. Of these, the ISSPP is the largest and longest running, spanning twenty countries over twenty years. Conducting multiple perspective, multi-level case studies using common research protocols, members of the ISSPP network have collectively produced the largest international, authoritative corpus of more than 150 case studies and more than 200 journal papers, book articles and books (Gurr and Moyi, forthcoming). It has been suggested that the ISSPP "is the most comprehensive and coherent international comparative study of the principalship ever undertaken." (Caldwell, 2014, xxii).

Determining what is success, who is successful, and how success is achieved is important in school leadership research. Success is a value, a process, an outcome, and an end in view, which is ongoing, as successful schools and their leaders do not stand still. Nor does success necessarily suggest a shared or transferable solution. In this on-going multi-national project, researchers across all participating countries have found that success does not have the same meaning as "effectiveness", or "improvement". *Successful principals, they found, are defined broadly across all jurisdictions and countries, cultures, and socio-economic contexts as those who attend, with their teachers, to moral, social, and ethical issues when educating students, as well as to their students' measurable achievements.*

Its members' work explores the qualities, characteristics, practices and identities of school principals who are successful over time in different countries and cultural settings. Its research focuses on how success is built and sustained over time in schools in rural and urban settings, and challenging socio-economic contexts; principal identities; and "visible" and "invisible" low performing schools, as well as educational career and principal preparation programs (Crow and Whiteman, 2016), and middle leadership (Farchi and Tubin, 2019; Gurr et al., 2018).

The studies in the member countries of the ISSPP project consistently indicate: (i) the importance of considering the part played by the combination of personal, professional, situational, policy and broader cultural contexts in achieving and sustaining success; (ii) the degree to which principals are respected and trusted by their school colleagues and external school communities; and (iii) the ways that relational trust, capacity, capability, shared professional practice, and individual and collective agency and efficacy, and raised levels of teaching, learning and student achievement are built successfully over time. These inform principals' diagnoses of need and their timely, contextually sensitive, consensual decision-making. Always in the ISSPP research findings, responding to "context" is defined as being the making of decisions which are taken in the broader as well as narrower interests of all interested stakeholders.

ISSPP research notes that transformational, instructional, distributed leadership all support successful principalship, depending on the degree of coordination and independence fostered, and the enabling of deep levels of trust and trustworthiness through

collaborative dialog; and that *values-led integrated leadership best represents the work of principals in successful schools as they tackle new contingencies within the complexity of their contexts*. How they express their values through management leadership systems and culture are key to their leadership success (Day, 2017).

Finally, the research in this project has found that school success is not achieved by the action—or inaction—of a single (heroic) person, but by the collective will and work of these and other leaders and teachers, and achieved through both the direct and indirect influence of principals.

The ISSPP studies have confirmed many of the earlier research studies and been able also to provide new insights and powerful images, particularly about the kinds of reflexivity, inner values, philosophies, and intra-as well as inter-professional qualities, skills, and strategies that successful principalship demands, regardless of culture, country, and school context; and they provide empirical affirmation of what was proposed by scholars 20 years ago (Hallinger, 1995; Hallinger and Heck, 1998; Hallinger and Leithwood, 1998; Heck, 1996; 1998; Leithwood and Duke, 1999; Walker and Dimmock, 2002), that certain parts of school leadership could be universal, i.e., culture-free; and other parts could be culture-specific (Day et al., 2020).

All three of the above studies have revealed more nuanced understandings of successful principals' work and lives in school contexts of complexity and change than those gained from research on particular "models," "frameworks" or "pathways" of leadership, or from "effective schools" and "school improvement" research discussed earlier in this article. However, the work of ISSPP members most closely embodies two key developments in the progressive nuancing of conception and conduct of school leadership research—the importance of acknowledging complexity, and of embracing cultural differences. This is in large measure because, either implicitly or explicitly, their research acknowledges that schools are complex adaptive systems, nested within and interactive with a number of other interconnected systems.

Complexity theory

What is unusual in school leadership research is that the ISSPP research is explicitly embedded in understandings that schools are complex adaptive systems, and that in these contexts, "rather than leading for efficiency and control, appropriate to manufacturing ... [successful]...organizations find themselves leading for adaptability, knowledge and learning" (Uhl-Bien et al., 2007, p. 301). In this conception, leadership has three main functions: adaptive (creative, that emerges from interactions); administrative (the formal leadership role); and enabling (combinations of bureaucratic (administrative) and emergent (adaptive)). Essentially, this notion arises from an understanding of schools as "living systems" in which:

There is no objective reality out there waiting to reveal its secrets. There are no recipes or formulae, no checklists or advice that describes "reality." There is only what we create through our engagement with others and events. Nothing really transfers; everything is always new and different and unique to each of us.

Wheatley (1992, p. 7).

The ISSPP research assumes that leaders are part of a profession, part of a school organization that is also part of the whole educational system, that is itself part of a country's cultural heritage. The system affects the environment and the environment affects the system (Morrison, 2002). Bronfenbrenner's (1995) "ecological systems theory" complements this, presenting a number of interactive system layers, each nested in and interactive with the others: micro (immediate environment of the classroom), meso (connections in the school), exo (indirect environment), macro (social/cultural values), and chrono (changes over time). The underlying rationale for the use of ecological systems theory, as found in much of the ISSPP research about successful leadership, is: (i) a view that schools are dynamic, policy influenced but not directed, task driven and relational in their nature; (ii) that success goes beyond the "functional", and (iii) that successful principals are agential, rather than compliant, influential through how they think and feel, who they are, what they do, and how they do it. The ISSPP research has found that successful principals *understand* that human behavior is likely to be complex, non-linear, interdependent, emergent, and uneven at and between different levels of this system, and that they are able to adapt to their ever-changing environments with values-based "systems thinking" (Day and Gurr, 2014; Day and Leithwood, 2007; Day and Gu, 2009 Day et al., 2011; Jacobson et al., 2005; Ylimaki and Jacobson, 2013).

Complexity theory in educational research suggests multiple causalities and multi-directional causes and effects (Cohen et al., 2011). This requires a paradigm shift for many, toward a "holistic, connectionist and integrationist view of the individual and the environment, rather than a fragmented, reductionist perspective" (Youngblood, 1997, p. 34). To date, there is relatively little research on system change which focuses, as the ISSPP research does, on the behaviors of individual agents and the relationships between these within a multi-level ecosystem (Jacobson et al., 2019, p. 113). In effect, these studies synthesize the new contextually-oriented, culturally informed, complexity-based knowledge about the nature of schools and successful school leadership, reaching beyond the previous successful school leadership research findings.

Cultural Differences

Educational leadership researchers cannot assume that all schooling is materially or functionally similar. Schooling and school resourcing do not operate in the same way in different countries, states and jurisdictions, or even within cities and sectors themselves. Within and between each of these are cultural assumptions about the purposes of schooling, what success means, and

who defines it. These assumptions may be unique to a school context, may be driven by national government education policies, and are increasingly determined by shared global assumptions about success, including the UNESCO sustainable development goals where education is fourth in priority, OECD research, and PISA test scores. Therefore, what determines success may be specific to a context, may travel between contexts, or be shared or borrowed by different contexts.

Country specific inherited cultures play an important part in principals' values, professional identities, the way they form relationships, and the way they lead. Culture can be defined as "the interactive aggregate of common characteristics that influence a human group's response to its environment" (Hofstede, 1980, p. 10). These visible manifestations of culture include values, symbols, heroes, and rituals. Culture shapes, but does not necessarily determine, the institutional and community context within which the school is situated, influences the predilections of individual leaders and the nature of their interactions with others in the school and its community, and strongly influences the particular educational emphasis or goals that prevail within a system of schooling (Hallinger and Leithwood, 1996). Culture can relate to the group of people at the school level (organizational culture) or at the national/regional level (societal culture) (Walker and Dimmock, 2002). Culture-common concepts can be found among people from different societies and cultures; culture-specific concepts represent different ways that people deal with culture-general demands, "both of which are necessary for an understanding of culture and cultural differences" (p. 71). Given the current convergence in understanding and development of school leadership globally, the study of school leadership across cultures can unveil distinctive features of leadership that are particularly effective in specific nations or cultures, and which should be celebrated (Walker and Dimmock, 2002).

The research of ISSPP found that, while the social, political, cultural, and educational histories of schools differ from country to country, all principals experienced change driven by the ever-increasing interest of central and municipal governments in measurable "performance" and "performativity", attended to new trends in instruction (e.g., personalization of the learning experience; teachers working collaboratively; greater use of feedback to improve performance at student, teacher, parent, school, and system levels), had a strong sense of moral and social/ethical purpose, shared the same cluster of values (e.g., caring for students, a love of learning, high expectations), increased their use of technology to support learning and qualities, strategies, actions, and relational connectedness, and enacted a set of core leadership practices (Day and Gu, 2009). According to this research, successful school principals focus on creating, establishing, sustaining, revisiting and where necessary renewing a set of curriculum, teaching, learning, achievement and relational principles and practices which are ethically, functionally and socially defensible, and relevant to the learning and achievement needs of every student in their school. They do this by understanding, acknowledging, and respecting the norms of external policy directives and the national and local cultures in which the school is located. They are not, however, always compliant with these, where they judge them not to be in the students' best interests. It will be evident from this, that successful leaders regard themselves as change leaders with strong ethical and moral purposes. Rather than acting as unquestioning implementors of externally defined policies, successful leaders interpret, modify, and incorporate these into preferred, values-led cultures and practices. In order to be successful, they know, also, that their priority must always be the quality of their students' learning, achievement and wellbeing, that they cannot achieve this directly or by their efforts alone, but that they can influence this indirectly through working in different ways over time to influence staff, students and their parents to work together for the longer-term broad as well as narrow educational benefits of the students. It is in choosing the right ways at the right times and with the right sets of internal and external relationships that they achieve success.

Such accumulations of multi-level case studies and longitudinal mixed-methods research designs as those above have the advantage of providing finer grained, more nuanced evidence-based understandings of the leadership roles and behaviors of principals who achieve and sustain educational outcomes in schools than single lens quantitative analyses, meta-analyses, or purely qualitative approaches. The findings provide support for more differentiated, context sensitive training and development for aspiring and serving principals. The importance of contextual influences identified by the Australian, English and ISSPP research were later affirmed by Hallinger's research, aptly entitled, "Bringing Context out of the Shadows" (Hallinger, 2016).

Future directions

As we move further into the 21st century, our educational contexts are likely to continue to be characterized by features such as accountability, globalization, and diversity. Research on successful school leadership needs to continue to be open to finding new elements to enhance its effectiveness and strength in new contexts. With the emergence of more empirical research and parallel evidence, comparative studies across continents can shed light on advanced leadership theories on an international platform. Findings about successful leadership exhibit certain commonalities across very different settings. The examination of successful leadership practices that are both culture-free and that are culture-specific can best inform leadership development within different contexts in a global context.

The research cited in this article unambiguously demonstrates the schools in this century are not static entities; they are complex adaptive systems, coping with uncertainties and complexity related to teachers, students, dynamic network structures, workload, school size, school level, policies, cultures, regions and countries. It follows that successful school leaders are adaptive, recognizing that classrooms at the micro level are nested within and interact with schools (at the meso level); that schools themselves are nested within and interact with policy, part of the indirect environment (the exo level), and society (at the macro level); and that all are subject to changes over time (the chrono level). It demonstrates, also, however, that working within and across these, principals who

succeed carry with them their resolute commitment to broad humanitarian educational values, equity *and* academic progress and achievement.

Thus, while the development of models, frameworks and pathways provides insights, these are unduly limiting. A further development of research within and across national jurisdictions using a combination of Integrated frameworks, or “pathways” and based on a view of schools as complex adaptive systems, would signal: (i) a growing awareness of the need to differentiate “effective” from “successful principalship”, the former being largely concerned with the functional, and the latter being concerned with both the functional and affective; (ii) a growing understanding of schools as complex systems which are nested in macro, exo and chrono systems and cultures, and which are influenced but not necessarily dictated by the vagaries of changing external policies and changes in societal demands; and (iii) an enhanced awareness of the importance of external, culturally inherited conceptions and practices of successful leadership. Identifying more indigenous elements of school leadership and its unique influence pathways, and more indigenous instruments developed in different cultural contexts, could help to develop indigenous leadership theories, which, combined with Western leadership theories, would eventually help us to understand and develop leadership with a global view and application.

Research now needs to examine more closely how leaders influence teachers and especially the quality of teachers’ work in classrooms in more detail, combining and complementing statistical structural equation modeling with qualitative methods in order to discover *how* change happens in classrooms as a consequence of principals’ influence and interventions.

While multi-level cultural perspectives need to be adopted to aid understanding of schools and their leaders, recent statistical advances (such as structural equation modeling and hierarchical linear modeling) also hold the promise of allowing researchers to investigate models across groups (e.g., schools or cultural settings) or between organizational levels (e.g., classes within schools within districts) or cultural layers (e.g., schools within communities within cultural settings). Research methods that incorporate the intuitive insights and speculative reflections of the observer from qualitative data and then submit them to systematic empirical testing offer the best hope for fulfilling the full potential of a comparative study. These will help to inform a more integrated, cross-cultural, sophisticated leadership theory that can meet the demands of globalization in the 21st century.

Two decades ago, Walker and Dimmock (2002) suggested that a cross-cultural comparative approach would allow administrative and leadership practices in different societies and cultures to be seen in relation to each other. As long as the bulk of theory and practice in the field is drawn almost exclusively from culturally conceived Anglo-American beliefs, philosophies, and understandings, our understandings regarding the field of educational administration and leadership are likely to remain partial and incomplete.

References

- Alexander, R., 2008. *Essays on Pedagogy*. Routledge, Oxon, England.
- Anderson, S., Moore, S., Sun, J.P., 2009. Positioning the principals in patterns of school leadership distribution. In: Leithwood, K., Mascal, B., Strauss, T. (Eds.), *Distributed Leadership According to the Evidence*. Routledge, New York, pp. 111–136.
- Andrews, R., Soder, R., 1987. Principal instructional leadership and school achievement. *Educ. Leader* 44, 9–11.
- Bass, B.M., Avolio, B.J., 1989. *Manual: The Multifactor Leadership Questionnaire*. Consulting Psychologists Press, Palo Alto, CA.
- Bass, B.M., Avolio, B.J., 1994. *Improving Organizational Effectiveness Through Transformational Leadership*. Sage, Thousand Oaks, CA.
- Bass, B.M., 1985. *Leadership and Performance Beyond Expectations*. The Free Press, New York.
- Begley, P.T., 1996. Cognitive perspectives on values in administration: a quest for coherence and relevance. *Educ. Adm. Q.* 32 (3), 403–426.
- Begley, P.T., 2003. In pursuit of authentic school leadership practices. In: Begley, P.T., Johansson, O. (Eds.), *The Ethical Dimensions of School Leadership*, pp. 1–12.
- Bossert, S., Dwyer, D., Rowan, B., Lee, G., 1982. The instructional management role of the principal. *Educ. Adm. Q.* 18, 34–64.
- Bronfenbrenner, U., 1995. Developmental ecology through space and time: a future perspective. In: Moen, P., Elder Jr., G.H. (Eds.), *Examining Lives in Context: Perspectives on the Ecology of Human Development*. American Psychological Association, Washington, DC, pp. 619–647.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement Lessons from Chicago*. University of Chicago Press, Chicago.
- Burns, J.M., 1978. *Leadership*. Harper and Row, New York.
- Caldwell, B.J., 2014. Forward. In: Day, C., Gurr, D. (Eds.), *Leading Schools Successfully: Stories from the Field*. Routledge, London, UK xxi–xxii.
- Cohen, L., Manion, L., Morrison, K., 2011. *Research Methods in Education*. Routledge, New York.
- Crow, G.M., Whiteman, R.S., 2016. Effective preparation program features: a literature review. *J. Res. Leadersh. Educ.* 11 (1), 120–148. <https://doi.org/10.1177/1942775116634694>.
- Day, C., Grice, C., 2019. *Investigating the Influence and Impact of Leading from the Middle: A School-Based Strategy for Middle Leaders in Schools*. The University of Sydney, Australia.
- Day, C., Gu, Q., 2009. Veteran teachers: commitment, resilience and quality retention. *Teach. Teach. Theor. Pract.* 15 (4), 441–457.
- Day, C., Gurr, D. (Eds.), 2014. *Leading Schools Successfully: Stories from the Field*. Routledge, London.
- Day, C., Leithwood, K., 2007. *Successful Principal Leadership in Times of Change: An International Perspective*. Springer, Dordrecht.
- Day, C., Harris, A., Hadfield, M., 2000. *Leading Schools in Times of Change: A Multiperspective Study of Effective Headteachers*. Open University Press, Buckingham.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *School Leadership and Student Outcomes: Linking Learning and Achievement*. Open University Press, Maidenhead.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Day, C., Sammons, P., Gorgen, K., 2020. *Successful School Leadership*. Education Development Trust, Reading, England.
- Day, C., 2017. *Teacher’s Worlds and Work: Understanding Complexity, Building Quality*. Routledge, London.
- DiPaola, M.F., Hoy, W.K., 2005. Organizational citizenship of faculty and achievement of high school students. *High Sch. J.* 88 (3), 35–44.
- Downer, D.F., 1991. Review of research on effective schools. *McGill J. Educ.* 26 (3), 323–332.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader* 37 (1), 15–24.
- Farchi, T., Tubin, D., 2019. Middle leaders in successful and less successful schools. *School Leadership and Management, Formerly School Organisation* 39 (3–4), 372–390.
- Gay, G., 2010. *Culturally Responsive Teaching: Theory, Research, and Practice*, second ed. Teacher College Press, New York, NY.
- Greenfield, T.B., 1976. Theory about what? Some more thoughts about theory in educational administration. *UCEA Rev.* 17, 4–9.
- Greenfield, T.B., 1978. Reflections on organization theory and the truths of irreconcilable realities. *Educ. Adm. Q.* 14 (2), 1–23.

- Gronn, P., 2002. Distributed leadership. In: Leithwood, K., Hallinger, P., Seashore-Louis, K., Furman-Brown, G., Gronn, P., Mulford, W., Riley, K. (Eds.), *Second International Handbook of Educational Leadership and Administration*. Kluwer, Dordrecht.
- Gurr, D and Mori, P., Editorial for the special issue, 'The International Successful School Principalship Project: Reflections and possibilities', *J. Educ. Admin.* forthcoming.
- Gurr, D., Drysdale, L., Longmuir, F., McCrohan, K., 2018. The leadership, culture and context nexus: lessons from the leadership of improving schools. *Int. Stud. Educ. Adm.* 46 (1), 22–44.
- Hallinger, P., Heck, R.H., 1996. Reassessing the principal's role in school effectiveness: a review of empirical research, 1980–1995. *Educ. Adm. Q.* 32 (1), 5–44.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9 (2), 157–191.
- Hallinger, P., Leithwood, K., 1996. Culture and educational administration. *J. Educ. Adm.* 34 (5), 98–116.
- Hallinger, P., Leithwood, K., 1998. Unseen forces: the impact of social cultural on school leadership. *Peabody J. Educ.* 73 (2), 126–151.
- Hallinger, P., Murphy, J., 1985. Assessing the instructional leadership behavior of principals. *Elem. Sch. J.* 86 (2), 217–248.
- Hallinger, P., 1992. The evolving role of American principals: from managerial to instructional to transformational leaders. *J. Educ. Adm.* 30 (3), 35–50.
- Hallinger, P., 1995. Culture and leadership: developing an international perspective of educational administration. *UCEA Rev.* 36 (1), 3–7.
- Hallinger, P., 2005. Instructional leadership and the school principals: a passing fancy that refuses to go away. *Leader. Pol. Sch.* 4, 221–239.
- Hallinger, P., 2015. *Assessing leadership for learning with the Principal Instructional Management Rating Scale*. Springer, Dordrecht Netherlands.
- Hallinger, P., 2016. Bringing context out of the shadows of leadership. *Educ. Manag. Adm. Leader* 46 (1), 5–24.
- Hattie, J., 2009. *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge, New York.
- Heck, R., Hallinger, P., 1999. Conceptual models, methodology, and methods for studying school leadership. In: Murphy, J., Seashore-Louis, K. (Eds.), *The 2nd Handbook of Research in Educational Administration*. McCutchan, San Francisco.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681.
- Heck, R.H., Moriyama, K., 2010. Examining Relationships Among Elementary Schools' Contexts, Leadership, Instructional Practices, and Added-Year Outcomes: A Regression Discontinuity Approach. *School Effectiveness and School Improvement*. Manuscript received from the author.2.
- Heck, R.H., Larson, T., Mercouliades, G., 1990. Principal instructional leadership and school achievement: validation of a causal model. *Elem. Sch. J.* 96, 527–549.
- Heck, R., 1996. Leadership and culture: conceptual and methodological issues in comparing models across cultural settings. *J. Educ. Adm.* 34 (5), 74–97.
- Heck, R., 1998. Conceptual and methodological issues in investigating principal leadership across cultures. *Peabody J. Educ.* 73 (2), 51–80.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86 (2), 531–569.
- Hodgkinson, C., 1991. *Educational Leadership: The Moral Art*. State University of New York Press, Albany.
- Hofstede, G., 1980. *Culture's Consequences Comparing Values, Behaviors, Institutions and Organizations Across Nations*, second ed. Sage Publications, Thousand Oaks, CA.
- Hoy, W.K., Miskel, C.G., 2013. *Educational Administration: Theory, Research, and Practice*, ninth ed. McGraw-Hill, New York.
- Jacobson, S., Johnson, L., Giles, C., Yilmaki, R., 2005. Successful leadership in U.S. schools: enabling principals, enabling schools. *J. Educ. Adm.* 43 (6), 607–618.
- Jacobson, M., Levin, J., Kapur, M., 2019. Education as a complex system: conceptual and methodological implications. *Educ. Res.* 48 (2), 112–119.
- James, S., Coleman, J.S., 1968. The concept of equality of educational opportunity. *Harv. Educ. Rev.* 38 (1), 14–16.
- Jencks, C.S., Smith, M., Ackland, H., Bane, M.1., Cohen, D., Gintis, H., Heyns, B., Michelson, S., 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. Basic Books, New York.
- Johnson, L.S., 2006. Making her community a better place to live: culturally responsive urban school leadership in historical context. *Leader. Pol. Sch.* 5, 19–36.
- Khalifa, M.A., Gooden, M.A., Davis, J.E., 2016. Culturally responsive school leadership: a synthesis of the literature. *Rev. Educ. Res.* 86 (4), 1272–1311.
- Langlois, L., Lapointe, C., Valois, P., de Leeuw, A., 2014. Development and validity of the ethical leadership questionnaire. *J. Educ. Adm.* 52 (3), 310–331.
- Leithwood, K., Duke, D., 1999. A century's quest to understand school leadership. In: Murphy, J., Louis, K.S. (Eds.), *Handbook of Research on Educational Administration*, vol. 2. Jossey-Bass, San Francisco, pp. 45–72.
- Leithwood, K., Jantzi, D., 1999. Transformational school leadership effects: a replication. *Sch. Effect. Sch. Improv.* 10 (4), 451–479.
- Leithwood, K., Jantzi, D., 2000. The effects of transformational leadership on organizational conditions and student engagement with school. *J. Educ. Adm.* 38, 113–129.
- Leithwood, K., Jantzi, D., 2005. A Review of Transformational School Leadership Research. Paper presented at the annual meeting of the American Educational Research Association. American Educational Research Association, Montreal, Canada.
- Leithwood, K., Louis, K.S., 2012. *Linking Leadership to Student Learning*. John Wiley & Sons, San Francisco, CA.
- Leithwood, K., Montgomery, D.J., 1982. The role of the elementary school principal in program improvement. *Rev. Educ. Res.* 52 (3), 309–339.
- Leithwood, K., Steinbach, R., 1991. Indicators of transformational leadership in the everyday problem solving of school administrators. *J. Person. Eval. Educ.* 4, 221–244.
- Leithwood, K., Sun, J.P., 2009. Transformational school leadership effects on schools, teachers and students. In: Hoy, W., Dipaola Series, M. (Eds.), *Research and Theory in Educational Administration*, vol. 8. Information Age, Charlotte, NC, pp. 1–22.
- Leithwood, K., Sun, J., 2012. The nature and effects of transformational school leadership: a meta-analytic review of unpublished research. *Educ. Adm. Q.* 48 (3), 387–423.
- Leithwood, K.A., Begley, P.T., Cousins, J.B., 1990. The nature, causes and consequences of principals' practices: an agenda for future research. *J. Educ. Adm.* 28 (4), 5–31.
- Leithwood, K., Tomlinson, D., Genge, M., 1996. Transformational school leadership. In: Leithwood (Ed.), *International Handbook of Educational Leadership and Administration*. Kluwer Academic, Netherlands, pp. 785–840.
- Leithwood, K., Jantzi, D., Steinbach, R., 1999. *Changing Leadership for Changing Times*. Open University Press, Philadelphia, PA.
- Leithwood, K., Louis, K., Anderson, S., Wahlstrom, K., 2004. *How Leadership Influences Student Learning: A Review of the Evidence Linking Leadership to Student Learning*. The Wallace Foundation, New York.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven Strong Claims About Successful School Leadership*. National College for School Leadership, Nottingham, England.
- Leithwood, K., Mascal, B., Strauss, D., 2009. *Distributed Leadership According to the Evidence*. Routledge, New York.
- Leithwood, K., Anderson, S., Mascal, B., Strauss, T., 2010. School leaders' influences on student learning: the four paths. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *The Principles of Educational Leadership and Management*, vol. 2. Sage, Los Angeles, pp. 13–30.
- Leithwood, K., Sun, J.P., Schumacker, R., 2020. How school leadership influences student learning: a test of "the four paths model". *Educ. Adm. Q.* 56 (4), 570–599.
- Leithwood, K., 1992. The move towards transformational leadership. *Educ. Leader* 495, 8–12.
- Leithwood, K., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30, 498–518.
- Leithwood, K.A., 2012. In: Louis, K.S., Anderson, S.E. (Eds.), *Linking Leadership to Student Learning*, first ed. Jossey-Bass, San Francisco, CA.
- Leithwood, K., 2017. The Ontario leadership framework: successful school leadership practice and personal leadership resources. In: Leithwood, K., Sun, J.P., Pollock, K. (Eds.), *How School Leadership Influences Student Learning: The Four Paths*. Springer, New York, pp. 31–45. In K. Leithwood (Series Ed.), *Studies in Educational Leadership*: 22.
- Lewin, K., Lippitt, R., White, K., 1939. Patterns of aggressive behavior in experimentally created "social climates". *J. Soc. Psychol.* 10 (2), 271–301.
- Marzano, R.J., Waters, T., McNulty, B., 2005. *School Leadership that Works: From Research to Results*. ASCD and McREL, Aurora, CO.
- Mitchell, R.M., Mendiola, B., Schumacker, R.E., Lowery, X.M., 2016. Creating a school context of success: the role of enabling school structure and academic optimism. *J. Educ. Adm.* 54 (6), 626–646.
- Morrison, K., 2002. *School Leadership and Complexity Theory*. Routledge, London.
- Muijs, D., Harris, A., 2003. Teacher leadership—improvement through empowerment?: an overview of the literature. *Educ. Manag. Adm.* 31 (4), 437–448. <https://doi.org/10.1177/0263211X030314007>.
- Mulford, B., Sillins, H., 2003. Leadership for organizational learning and improved student outcomes—what do we know? *Camb. J. Educ.* 33 (2), 175–195.

- Mulford, B., 2008. *The Leadership Challenge: Improving Learning in Schools*. Australian Council for Educational Research, Canberra.
- Murphy, J., Elliot, S.N., Goldring, E., Porter, A.C., 2006. *Learning-Centered Leadership: A Conceptual Foundation*. Wallace Foundation, New York, NY.
- Murphy, J., 1990. Principal instructional leadership. *Adv. Educ. Adm.* 1 (Part B), 163–200.
- Noddings, N., 2013. *Caring. A Relational Approach to Ethics and Moral Education*, second ed. University of California Press, Berkeley. Updated.
- Northouse, P.G., 2013. *Leadership Theory and Practices*. Sage Publications, Los Angeles.
- Parry, K., Bryman, A., 2006. I: 2.1 Leadership in Organizations. *The SAGE Handbook of Organization Studies*, p. 447.
- Podsakoff, P.M., MacKenzie, S.B., Moorman, R.H., Fetter, R., 1990. Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leader. Q.* 1 (2), 107–142.
- Pont, B., Nusche, D., Moorman, H., 2008. *Improving School Leadership Volume 1: Policy and Practice* OECD.
- Power, M., 2019. Modeling the Microfoundations of the Audit Society: Organizations and the Logic of the Audit Trail. *Academy of Management Review* (ja).
- Purkey, S.C., Smith, M.S., 1983. Effective schools: a review. *Elem. Sch. J.* 83 (4), 427–452.
- Robinson, V.M.J., Lloyd, C.A., Lloyd, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 445, 635–674.
- Ryan, J., 2006. Inclusive leadership and social justice for schools. *Leader. Pol. Sch.* 5 (1), 3–17.
- Sammons, P., Hillman, J., Mortimore, P., 1995. *Key Characteristics of Effective Schools: A Review of School Effectiveness Research*. Eric Database, London University, England. <https://files.eric.ed.gov/fulltext/ED389826.pdf>. (Accessed 19 September 2020).
- Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28 (3), 463–488.
- Sebring, P.B., Allensworth, E., Bryk, A.S., Easton, J.Q., Luppescu, S., 2006. *The Essential Supports for School Improvement*. Consortium on Chicago School Research, Chicago, IL.
- Sergiovanni, T.J., 1992. *Moral Leadership: Getting to the Heart of School Improvement*, first ed. Jossey-Bass, San Francisco.
- Spillane, J., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- Spillane, J.P., Halverson, R., Diamond, J., 2004. Towards a theory of school leadership practice: implications of a distributed perspective. *J. Curric. Stud.* 36 (1), 3–34.
- Spillane, J., 2017. Leadership and learning: conceptualizing relations between school administrative practice and instructional practice. In: Leithwood, K., Sun, J.P., Pollock, K. (Eds.), *How School Leadership Influences Student Learning: The Four Paths*. Springer, New York, pp. 49–67.
- Starratt, R.J., 1994. *Building on Ethical School: A Practical Response to the Moral Crises in Schools*. Falmer, London.
- Sun, J.P., Leithwood, K., 2012. Transformational school leadership effects on student achievement. *Leader. Pol. Sch.* 114, 418–451.
- Sun, J.P., 2011. Ethical decision-making and ethical responding: an analysis and critique of various approaches through case study. *Int. J. Leader. Educ.* 14 (1), 21–45.
- Sun, J.P., 2016. The nature, impacts and antecedents of transformational school leadership: a review. In: Weinstein, J. (Ed.), *Educational Leadership in Schools: Nine Perspectives*. University Diego Portales Press, Santiago, Chile.
- Uhl-Bien, R., Marion, R., McKelvey, B., 2007. Complexity Leadership Theory: shifting leadership from the industrial age to the knowledge era. *Leader. Q.* 18 (4), 298–318.
- Walker, A., Dimmock, C., 2002. Moving school leadership beyond its narrow boundaries: developing a cross-cultural approach. In: Leithwood, K., Hallinger, P. (Eds.), *Second International Handbook of Educational Leadership and Administration*. Springer Science + Business Media, Dordrecht, Netherlands, pp. 176–204.
- Waters, J.T., Marzano, R.J., McNulty, B., 2003. *Balanced Leadership: What 30 Years of Research Tells Us About the Effect of Leadership on Student Achievement*. Mid-continent Research for Education and Learning, Aurora, CO.
- Webb-Johnson, G.C., Carter, N., 2007. Culturally responsive urban school leadership: partnering to improve outcomes for African American learners. *Natl. J. Urban Educ. Pract.* 1, 77–99.
- Wheatley, M.J., 1992. *Leadership and the new science: learning about organization from an orderly universe*. Berrett-Koehler, San Francisco, CA, p. 7.
- Willower, D.J., 1997. Inquiry in educational administration: pitfalls and possibilities. *Educ. Manag. Adm.* 25 (4), 437–449.
- Ylimaki, R., Jacobson, S., 2013. School leadership practice and preparation: comparative perspectives on organizational learning (OL), instructional leadership (IL) and culturally responsive practices (CRP). *J. Educ. Adm.* 51 (1).
- Youngblood, M., 1997. *Life at the Edge of Chaos: Creating the Quantum Organization*. Perceval Publishing, Dallas, Texas.

Integrated educational leadership

Kenneth Leithwood^{a,b,c}, ^a University of Toronto, Toronto, ON, Canada; ^b University College London, London, United Kingdom; and ^c University of Nottingham, Nottingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	73
Integrated educational leadership	74
Instructional leadership (IL)	74
Transformational leadership (TL)	75
Illustrating the integration of IL and TL practices	75
Integrated educational leadership effects	75
Improving conceptions of integrated leadership: adding “personal leadership resources”	77
Cognitive resources	77
Problem solving expertise	78
Organizational improvement knowledge	78
Social resources	78
Perceiving emotions	78
Managing emotions	78
Acting in emotionally appropriate ways	79
Psychological resources	79
Optimism and resilience	79
Self-efficacy	79
Proactivity	79
Values and ethics	79
Conclusion	80
References	80

Introduction

This paper unpacks the meaning of integrated educational leadership (IEL), as it has typically been used in the educational leadership literature. Evidence is also summarized about the effects of such leadership on both school organizations and students. Throughout the paper, the work of my colleagues and me,^a especially as it has become embedded and extended in the *Ontario Leadership Framework* (OLF), is used to illustrate some features of IEL. A relatively comprehensive example of IEL, the OLF is also used to illustrate possible features of a major addition to most conceptions of IEL - leadership traits, dispositions or “personal leadership resources”.

Much of the research on leadership is focused on specific “models”. Typically, these are theoretical accounts of a selected set of leadership practices believed to be especially successful in accomplishing one or more valued organizational processes and/or outcomes. For example, *charismatic leadership* is about how to inspire followers to adapt and strive toward accomplishing a leader’s vision (Conger, 2015); *ethical leadership* stresses the importance of basing decisions on such widely held values as fairness, equity, transparency, honesty and the like (Brown and Trevino, 2006); *leader-member exchange theory* is about the importance of different types of relationships between leaders and followers (Gerstner and Day, 1987); and *servant leadership* encourages leaders to appreciate, and then fulfill, the needs and aspirations of followers (Greenleaf, 1996); *inclusive leadership* acknowledges the priority now awarded to leaders’ efforts at ensuring the educational needs all of their students are met (Khalifa et al., 2016a,b); and *distributed leadership* emphasizes making productive use of leadership resources potentially located in many people across the organization (Harris, 2013). Leadership models such as these typically excel in the coherence of their formulations and some have been the object of extensive empirical testing. Their attraction is due, in no small part, to the simplification of an otherwise highly complex role or function which demands a wide array of different practices, all contingent on the context in which leaders find themselves. Those practices are, themselves, shaped by a host of personal traits or dispositions. Such simplification is achieved by the focus of most leadership models on one or several undeniably important parts of what effective leaders actually do while ignoring the rest.

^aThis is work reflected in a four-decade long research program with colleagues and students. About a decade ago, this work was given considerable additional impetus through the formation of a partnership with the Canadian province of Ontario’s Ministry of Education and Ontario’s Institute for Educational Leadership. This partnership resulted in the *Ontario Leadership Framework* (Leithwood, 2012, 2020), now widely used in the province as a guide to leadership selection and development. As this paper is being written, the Framework is in the final stages of being updated and revised.

Adopting the position of test theorists, one might argue that those undeniably important parts of leadership can be treated as representative samples of the leadership domain as a whole. Leaders who enact those undeniably important parts of what effective leaders do, according to this position, can be expected to enact most of the other important parts, as well. There is no evidence to either confirm or disconfirm this argument. Alternatively, however, one might plausibly argue that the positive evidence about leadership effects, based on the conceptions of leadership represented by most leadership models, must radically underestimate those effects since such a small proportion of leaders' actual work is taken into account. Both of these positions, however, would encourage conceptions and measures of leaders' work that more fully capture what leaders actually do whenever feasible. This has been the basic warrant giving rise to the development of IEL models.

Integrated educational leadership

The concept of IEL became especially visible in 2003 with the publication of two papers. Hallinger (2003), reviewed relevant concepts and evidence, identified practices the literature had associated with both instructional leadership (IL) and transformational leadership (TL) to that point, and identified the similarities and differences between practices associated with the two approaches. Marks and Printy (2003) tested the relative effects of a selected sub-sample of IL and TL practices on classroom instruction and student achievement. A combination of the two sets of practices, they found, had significantly larger effects than either set of practices separately. Most subsequent research on IEL has assumed that it is the combination of practices associated with these two leadership models (Robinson et al., 2008).

While Hallinger (2003) and Marks and Printy (2003) provided much of the impetus for subsequent research on IEL, Dwyer (1985)^b anticipated important features of that work in his earlier study of IL. The results of that study identified what Dwyer referred to as two forms of principals' IL. One form was represented by a group of successful school leaders who focused primarily on setting the conditions of instruction in their classrooms ("working with their staffs to define and coordinate the what, when, where, and how of instruction" p. 14); this focus conforms closely to early views of IL that emerged from the effective schools movement. A second group of successful school leaders in Dwyer's study focused largely on setting the conditions for instruction in their schools "providing school environments that are supportive of teachers' work and students' learning" (p. 14); this focus reflects much of how TL is enacted in school contexts. So, Dwyer's (1985) study can be viewed as anticipating the IL/TL concept of integrated leadership as one which entails leaders setting the conditions both of and for instruction.

Instructional leadership (IL)

With a long history rooted in the effective schools movement (e.g., Edmonds, 1979), clear and testable models of IL have been proposed, for example, by Hallinger and Murphy (1985), Blasé and Blasé (2000) and others. Encompassing recent concepts such as "leadership for learning" (Hallinger, 2009; Knapp et al., 2012), the idea of IL still dominates the aspirations of a great many practicing school leaders even though the exact meaning of IL varies across models. Although sometimes accused of being too principal-centric and top down, shared or distributed conceptions of IL were advanced very early in research about IL (e.g., Dwyer, 1985; Rosenholtz, 1985). This shared conception of IL continues to receive considerable support (e.g., Bush, 2015; Supovitz et al., 2010).

While most models of instructional leadership place considerable weight on the relatively direct engagement of leaders in their school's instructional programs, the nature of that engagement varies. Three explicit models of IL illustrate such variation, as well as their "family resemblance". The most explicit, detailed and fully tested model of IL was developed by Hallinger and Murphy (Hallinger, 2005; Hallinger and Murphy, 1985; Murphy, 1990). This model includes three categories of practices - *Defines the school mission*, *Manages the instructional program* and *Develops a positive school learning climate*. A total of eleven sub-categories are included in this model. *Defining the school mission*, for example, encompasses the sub-dimensions *framing the school goals* and *communicating the school goals*.

A second example of an IL model, by Blasé and Blasé (2000), includes just two major dimensions of effective principal behavior. The first dimension, *Talking with teachers to promote reflection* includes *giving feedback*, *modeling*, *soliciting advice and opinions* and *giving praise*. The second dimension, *Making suggestions*, encompass nine behaviors, for example, *listening*, *sharing experience*, *giving teachers choice*, *encouraging risk taking*, and *recognizing teachers' strengths*. Bryk et al. (2010) is a third example. This IL model conceives of multiple principal leadership roles *developing the mission and goals of the school*, *setting high standards and communicating expectations to staff*, *developing trust and collaboration among staff*, and *supporting instruction*.

Based on a "broad reading of the literature on instructional leadership" since Dwyer's (1985) study, Hallinger (2005) indicates that conceptions of IL typically focus on some or all of the following dimensions:

- "Creating a shared sense of purpose in the school, including clear goals focused on student learning;
- Fostering the continuous improvement of the school through cyclical school development planning that involves a wide range of stakeholders;

^bAlthough now 35 years old, this remains one of the most methodologically impressive IL studies ever conducted. A considerable proportion of what subsequent studies of IL report do very little to advance Dwyer's findings even though leaders' contexts have changed dramatically.

- Developing a climate of high expectations and a school culture aimed at innovation and improvement of teaching and learning;
- Coordinating the curriculum and monitoring student learning outcomes;
- Shaping the reward structure of the school to reflect the school's mission;
- Organizing and monitoring a wide range of activities aimed at the continuous development of staff; and
- Being a visible presence in the school, modeling the desired values of the school's culture" (p. 233).

Transformational leadership (TL)

Accounts and evidence about transformational leadership, initially emerging from research in non-educational organizations, attracted significant attention by educators beginning in the late 1990s. In his seminal publication, Burns argued that transforming leadership "occurs when one or more persons *engage* with others in such a way that leaders and followers raise one another to higher levels of motivation and morality (1978, p. 20). Leaders do this by developing a vision for the organization, developing commitments and trust among workers, and facilitating organizational learning. Unlike many earlier theories of leadership which emphasized rational processes, transformational leadership theory emphasizes emotions and values, attributes of importance to symbolic behavior, and conceptualizes the role of the leader as helping make events meaningful for followers (Siangchokyo et al., 2020).

While closely aligned versions of transformational leadership were proposed by Burns (1978), Bass (1985) and Podsakoff et al. (1990), Bass launched an extensive program of empirical research (e.g. Bass and Avolio, 1994) which provided, among other things, evidence to justify the value of four categories of transformational leadership practices. These categories include *idealized influence*, *inspirational motivation*, *intellectual stimulation*, and *individualized consideration* (along with three transactional dimensions). To these dimensions, Podsakoff et al. (1990) added *articulating a vision*, *providing an appropriate model*, *fostering the acceptance of group goals*, and *high-performance expectations*.

Illustrating the integration of IL and TL practices

Leithwood and his colleagues initially developed a set of transformational leadership practices by adapting and extending those identified most directly by Podsakoff et al. (1990) with the aim of better capturing the consequences for leaders of working in school organizations (e.g., Leithwood et al., 1999; Leithwood and Jantzi, 2005). There were three broad categories of such practices in this early formulation including *setting directions*, *developing people*, and *redesigning the organization*.

Leadership practices aimed at *redesigning the organization* are unique to this conception of transformational leadership, although some accounts of Bass's conception refer to the importance of leaders influencing organizational cultures and structures (e.g., Avolio, 1994). This third domain of practices responds to the limited explicit attention devoted to leaders' organizational contexts by most models of transformational leadership. As part of developing the OLF, a fourth dimension was also added (*securing accountability*), expanding the models' attention to a broad policy context which holds most school leaders accountable for advancing their schools' achievement of widely valued student outcomes.

Finally, this model includes a fifth domain (*Improving the Instructional Program*) encompassing practices that are part of most (direct) IL models (e.g., Hallinger, 2005). This domain reflects the well documented claim that a considerable proportion of leaders' attention should be devoted to the technical core of schooling (e.g., Robinson et al., 2008).

A total of 21 specific practices are included in the five domains of this integrated model, as they appear in the OLF, including:

- Summaries of evidence justifying each of the domains and practices
- Explanations of when and how each of the practices might be enacted
- Explanations of the causal relationship between the domains of leader practices, follower responses and such distal outcomes as student learning.

Integrated educational leadership effects

There is a strong conceptual argument, touched on above, favoring an integrated model of school leadership over either IL or TL models alone. Integrated models potentially capture a larger proportion of what practicing leaders do and, as a result, ought to explain more of the variation in leadership effects than does either IL or TL alone. This section summarizes the results of nine quantitative studies testing IEL effects. Although small in number, these studies likely capture the bulk of such research to date.

Marks and Printy's (2003) mixed methods study tested integrated leadership effects, as well as the separate contribution of IL and TL practices to those effects. Their measure of instructional leadership were questions asking teachers to rate the priority which leaders attached to instruction, teachers' sharing in such leadership outside their own classrooms and the mutual engagement of principal and teacher leaders in efforts to improve instruction. The measures of transformation leadership in this study was based on a sub-set of the transformational practices included in Leithwood's early work on TL (Leithwood et al., 1999; Leithwood and Riehl, 2005). Results indicated that the quality of instruction and the achievement of students in schools led by principals enacting high levels of integrated leadership were significantly better than in schools led by principals enacting low or non-integrated approaches.

Day et al. (2016) study was part of a much larger three-year, mixed-methods research program about the role of primary and secondary school leadership in improving and effective schools in England. This study used an integrated model of leadership (after Leithwood et al., 2008) to examine how the practices of leaders changed over time as their schools' organizational features, student performance and staffs' capacities changed. The authors sum up one important set of findings as follows:

...Neither instructional leadership strategies nor transformational leadership strategies alone were sufficient to promote improvement.... Leadership strategies that built on change in organizational structures and conditions but that focused more closely on developing people to become innovative and more rigorous in their teaching practices and to learn to use data and observation to improve their teaching also played an important role in school improvement processes (p. 253).

Two parallel studies conducted in Chicago high and elementary schools (Sebastian and Allensworth, 2012; Sebastian et al., 2017), tested the extent to which an integrated model of school leadership influenced the quality of classroom instruction and student achievement. While the authors refer to the leadership model they tested as "instructional" (direct and indirect IL), it conforms closely (if incompletely) to the conception of integrated leadership common to some other studies reviewed here. Direct IL was measured with a set of questions about teachers' perceptions of their principal as an instructional leader, as well as the extent to which their principal set high standards for teaching and learning, set school vision and goals and supported effective classroom instruction. Transformational leadership (or indirect IL) was reflected in items measuring the extent to which teachers felt that their principals had confidence in them, had trusting relationships with them, encouraged collaboration, and looked out for their personal welfare.

Results indicated that leadership on the part of both elementary and high school principals had significant associations with classroom instruction and student achievement. These effects were mediated through the schools' learning climate alone, a climate strongly influenced by teacher leadership, and measured by student perceptions of "whether they got along, treated each other with respect, cared about each other ... and the extent to which students felt safe inside and around the school buildings" (Sebastian et al., 2016, p. 83).

ten Bruggencate et al. (2012) tested the effects of an integrated conception of leadership on student achievement and promotion rates using evidence from principals (103), teachers (998) and students (4336) in 97 Dutch schools serving students in five grades starting at age 12. Mediating variables in this study included student engagement and several features of teachers' work (support for students, class atmosphere, organization of instruction), along with the tendency for school cultures to reflect a performance orientation or a development orientation. School leadership was conceptualized according to the *Conflicting Values Framework* (CVR) (Quinn and Rohrbaugh, 1983), a framework broadly representing four different value positions underlying research on organizational effectiveness, each at least loosely reflecting one of the OLF leadership domains— a rational position (*setting directions*) a human relations position (*developing people*), an open systems position (*redesigning the organization*) and an internal processes position (*improving the instructional programs*). While these parallels seem conceptually imperfect, some of the actual items used to measure leadership were similar to those in Leithwood's survey (2017).

Results indicated that leader behavior, in aggregate, had significant total effects on both the performance orientation and development orientation of school cultures. Leader behavior associated with rational goals and open systems values had positive total effects on teachers' work. Smaller effects of leader behaviors were associated with internal process and human relations values. Leadership reflecting values associated with CVR's rational goals and open systems had negative effects on student promotion rates.

Three studies have tested the effects of leadership practices included in the OLF and all three studies used the same measure of leadership. Boberg and Bourgeois (2016), examined the direct and indirect effects of this model of IEL on student math and language achievement. Collective teacher efficacy (CTE), organizational citizenship behavior (teacher extra commitment and effort) and student engagement were hypothesized to mediate the effects of leadership on student achievement. Evidence for this study was provided by teachers (569) and students in 63 public charter middle and upper schools located in the south-central regions/areas of the United States. Results indicated that IEL had significant direct effects on teacher commitment and effort, as well as CTE, and that CTE had significant direct effects on both reading and math achievement as well as on students' emotional engagement. The authors concluded that:

This study's [IEL] model demonstrates combined indirect effect sizes on reading and mathematics that reveal considerable improvements over all earlier TL models evaluated by Marks and Printy (2003), Robinson et al. (2008), and Scheerens (2012) (p. 368).

Leithwood and his colleagues have conducted two tests of the OLF version of IEL, the second an approximate replication of the first. Both studies inquired not only about the effects of IEL on student achievement in math and language but also about the relative influence of a suite of four categories of conditions (12 in total) mediating such effects. Evidence for the first of these studies (Leithwood et al., 2010) were survey responses from 1445 teachers in 199 elementary schools in Ontario and government-collected achievement data for all students in those schools. Results indicated small but significant total effects of integrated leadership on both measures of achievement. These effects were significantly mediated, to about the same degree, by variables include in three of the four categories of mediators. Especially significant individual mediating conditions included CTE along with the school's academic press and disciplinary climate.

Evidence for the second of these studies (Leithwood, 2020) used data from 2017 teachers located in all elementary schools in six Texas school districts, along with school-level student achievement results on a multi-subject state test. The same four categories of mediators were measured in this study as in the first of these studies. Confirming results from the first study, results demonstrated small but significant effects of integrated leadership on student achievement. As in the first study, these effects were largely mediated by CTE, school academic press and disciplinary climate. Broadly similar results also have been reported by Boberg and Bourgeois (2016), Heck and Hallinger (2014), Sebastian and Allensworth (2012), and Day et al. (2016).

The ninth and final study in this review was by Kwan (2020) using responses from vice-principals in 177 Hong Kong schools to assess IEL effects on student outcomes and the interaction of TL and IL in producing those outcomes. The measure of TL in Kwan's study included items (15) classified as *strategic directions and policy environment, leader and teacher growth and development, and staff management* (broadly defined), while the measure of IL included items (10) classified as *Teaching, Learning, and Curriculum and Quality Assurance and Accountability*. Results indicate that the influence on student achievement of two of the three components of instructional leadership (quality assurance and accountability, teaching, learning and curriculum) were significantly moderated by transformational leadership. Kwan concluded that:

Instructional leadership by the principal will not lead to considerable improvement in student outcomes unless the principal has already made available a school environment in which teachers are competent and motivated (p. 344).

The nine studies reviewed in this section of the paper indicate, in sum, that IEL has generally positive effects on the quality of classroom instruction, as well the engagement and achievement of students in several different disciplines. These effects are mediated by a small number of school conditions, most notably the collective efficacy of teaching staffs, a strong academic school culture and a safe and orderly environment/disciplinary climate for both staff and students.

Nevertheless, this body of evidence places severe limitations on claims that can be made about IEL effects. First, it is a very small body of evidence. Second, only three of the studies compared the relative effects of IL and TL components and demonstrated that the two components together have larger effects than either component separately (Kwan, 2020; Marks and Printy, 2003; Day et al., 2016). Finally, conceptions and measures of IEL varied widely. Only three of the eight studies provided evidence about approximately the same model of IEL. The most defensible conclusion to be drawn from this evidence, as a whole, is that IEL has the potential to realize the values conceptualized for it but, to invoke that well-worn phrase "more research is needed".

Improving conceptions of integrated leadership: adding "personal leadership resources"^c

Most accounts of IL focus exclusively on the overt behaviors or practices of those exercising leadership, with only modest attention to the feelings, emotions or dispositions of either leaders or their colleagues. Accounts of TL acknowledge the central importance of leaders engaging with the dispositions of their colleagues but are largely silent on the dispositions of leaders themselves. Neglect of attention to leaders' dispositions has been influenced to a great extent by early leadership research casting doubts in the contribution of dispositions or traits to explanations of variation in leadership effectiveness (e.g., Stogdill, 1974). The warrant for excluding traits, however, began to be challenged in the 1980's (e.g., Kenny and Zaccaro, 1983) and by 2007 Zaccaro, for example. Noted that:

A number of studies have linked personality variables and other stable personal attributes to leader effectiveness, providing a substantial empirical foundation for the argument that traits do matter in the prediction of leader effectiveness (p. 6).

Considerable evidence now substantiates this argument (e.g., Judge et al., 2002; Peterson et al., 2003). Furthermore, as Zaccaro also claims, the "qualities that differentiate leaders from non-leaders are far ranging and include not only personality attributes but also motives, values, cognitive abilities, social and problem-solving skills, and expertise" (2007, p. 8). Taken as a whole, these personal qualities help to explain why leaders differ in their responses to the same situations or contexts. Personal resources heighten leaders' attention to selected features of the situations they face, sometimes triggering predictable response among those with similar personal resources (Gottfredson and Reina, 2020).

These strands of theory and evidence prompted the addition of four categories of personal leadership resources to the OLF - cognitive, social, and psychological resources, along with values and ethics - introduced here to exemplify the value added to conceptions of IEL by introducing leader dispositions.

Cognitive resources

Considerable evidence collected over many decades suggests that leaders' success or effectiveness is partly explained by intelligence and experience. This would only be surprising if it was not the case, although some early evidence indicates that stressful and hectic

^cThe term *Personal Leadership Resources* is used here (and in the OLF) to encompass traits, dispositions and other "internal" qualities of leaders.

environments (features of environments in which school leaders often find themselves) reduce the advantage of greater intelligence to near zero. Intelligence and experience, however, are “surface” traits of leaders offering little guidance to those selecting and developing leaders or to leaders and aspiring leaders themselves. Below the surface of what is typically referred to as leader’s intelligence are problem-solving capacities and below the surface of “experience” is the domain-specific knowledge useful for such problem solving.

Problem solving expertise

Problem solving is at the heart of leaders’ day-to-day work (Kelman et al., 2020) and effective IEL practices, whether of the IL or TL variety, can be viewed as the “solution processes” component included in most comprehensive problem-solving models. Research on expert problem solving is extensive and there are multiple possible models that could be included as part of an IEL model (e.g., Mumford et al., 2007).

The main components of the model included in the OLF, for example, include problem interpretation, identification of constraints, goal setting, clarification of relevant principles and values, identification of possible solutions, and considerations about how the mood of problem solvers influences their decisions. This problem-solving model is the outcome of a decade-long program of research on how expert school and district leaders solve unstructured problems (e.g., Leithwood and Steinbach, 1995).

“Systems thinking” is also a form of problem-solving writ large. Leaders able to engage in systems thinking understand the dense, complex and reciprocal connections among different elements of the organization; these Leaders also exhibit the foresight needed to engage their organizations in likely futures and consequences for action (Shaked and Schecter, 2017). Changing one organizational element typically has consequences for many others. These consequences may be positive or negative. So, leaders engaged in improvement efforts also need to anticipate and take into account possible second-order consequences of their planned improvements.

Organizational improvement knowledge

While is not possible to codify all of the knowledge likely to be useful to those exercising leadership, the well-documented indirect influence of leadership on students recommends priority to “organizational improvement knowledge”. Leaders’ contributions to the success of their students depends on them knowing about those malleable learning conditions in and around their schools with direct positive effects on student success, diagnosing the status of those conditions in their own schools; selecting those most likely to be constraining student success and improving the status of those conditions.

The studies of IEL effects reviewed earlier provided examples of malleable learning conditions with direct effects on student success – an academic culture of high expectations, collective teacher efficacy and a safe and orderly environment. The OLF, for example, identifies twelve of these malleable learning conditions recommended for leaders’ attention by the results of significant bodies of research.

Social resources

The ability to develop and sustain productive working relationships has long been acknowledged as fundamental for leaders in almost all organizational contexts. Socio-emotional skills are also fundamental capacities for nurturing creative leadership in an organization. (Mumford et al., 2007). However, the importance of this ability grows with the interpersonal intensity typically experienced within an organization and the demands such intensity places on its leadership. Schools typically experience a level of interpersonal intensity virtually unmatched in any other type of organization. The capacity of especially school leaders to do such emotionally laborious work depends on their “social intelligence”, “social appraisal skills” or “emotional intelligence” (Slaski and Cartwright, 2003; Schutte et al., 2013). Three types of these social resources, closely aligned with Mayer and Salovey (1997) theory of emotional intelligence, are especially important contributions to leaders’ work.

Perceiving emotions

This set of social resources includes the ability to detect, from a wide array of clues, one’s own emotions (self-awareness) and the emotions of others. People with this relational resource are able to recognize their own emotional responses and how those emotional responses shape their focus of attention and influence their actions. They are also able to discern the emotions being experienced by others from, for example, their tone of voice, facial expressions, body language and other verbal and non-verbal information.

Managing emotions

This set of social resources includes managing one’s own and others’ emotions, including the interaction of emotions on the part of different people in pairs and groups. People with this relational resource are able to understand the reasons for their own “intuitive” emotional responses and are able to reflect on the potential consequences of those responses. Leaders with this capacity are also able to persuade others to be more reflective about their own “intuitive” emotional responses and to reflect on the potential consequences of those responses (e.g., Little et al., 2016).

Acting in emotionally appropriate ways

This set of social resources entails the ability to respond to the emotions of others in ways that support the purposes for the interaction. Leaders do this by exercising a high level of cognitive control over which emotions are allowed to guide their actions and by assisting others to act on emotions most likely to best serve their interests (e.g., [Burch, 2013](#)).

Psychological resources

Three psychological traits (discussed below) make significant contributions to leaders' success individually. But when these resources act in synergy, that is, when one person possesses all these resources, they make an especially large contribution to such success.

Optimism and resilience

Optimism and resilience are closely associated traits. Optimism is the habitual expectation of success in one's efforts to address challenges and confront change now and in the future. At the core of resilience is the ability to "bounce back" from failure and even move beyond one's initial goals while doing so. Resilient leaders have the ability to thrive in the challenging circumstances commonly encountered by school leaders. Optimistic and resilient leaders expect their efforts to be successful in relation to those things over which they have direct influence or control but not necessarily to be powerful enough to overcome negative forces in their organizations over which they have little or no influence or control; they are realistic as well as optimistic. Optimistic leaders are likely to take initiative and responsible risks with positive expectations regardless of past problems or setbacks (e.g., [Avey et al., 2011](#); [Zeni et al., 2016](#)).

Self-efficacy

Sense of efficacy is a belief about one's own ability to perform a task or achieve a goal. It is a belief about ability, not actual ability. That is, efficacious leaders believe they have the ability to solve whatever challenges, hurdles or problems that might come their way in their efforts to help their organizations succeed. Self-efficacy beliefs contribute to leaders' success through their directive effects on leaders' choices of activities and settings and can affect coping efforts once those activities are begun ([Bandura, 1997](#)). Efficacy beliefs determine how much risk people will take, how much effort they will expend and how long they will persist in the face of failure or difficulty. The stronger the self-efficacy the longer the persistence.

The combined effect of Optimism, Resilience and Efficacy is characterized by confidence about succeeding, having positive expectations about succeeding now and in the future, persevering to meet those expectations and being resilient in the face of challenges that threaten success ([Luthans et al., 2004](#)).

Proactivity

Among the personal leadership resources in the *Ontario Leadership Framework*, proactivity alone reflects much of the evidence about the contribution that personality traits make to the explanation of variation in leadership practices; its effects are also an extension of those associated with optimism. Recent evidence about leaders' personality traits has been largely guided by the "five factor model of personality" including Extraversion, Openness to Experience, Agreeableness, Neuroticism and Conscientiousness. Results of this research point to Extraversion as a key contribution to both leader effectiveness and leader emergence (e.g., [Judge et al., 2002](#)). While the contribution of Extraversion to leadership is well established, the concept itself has negative connotations, especially in cultures which value collaboration and leadership distribution. Proactivity is a closely related but more socially palatable trait in such cultures. Emerging from positive psychology, [Crant and Bateman \(2000\)](#) describe Proactivity as follows:

People who are proactive effect environmental change; they identify opportunities and act on them, show initiative, and persevere until they bring about meaningful change. They transform their organization's mission, find and solve problems, and take it upon themselves to have an impact on the world around them. Less proactive people show little initiative, tending to passively adapt to their circumstances rather than change them (p. 65).

Values and ethics

Many decisions faced by leaders are ethical dilemmas ([Cranston et al., 2014](#)) and almost all decisions are influenced, one-way or another, by leaders' values. Considerable evidence associated with ethical leadership points to both personal and professional values influencing leaders' success. Personal values that evidence collected in Western contexts suggests are generally associated with ethical leadership ([Cranston et al., 2014](#)) include, for example: fairness, honesty, integrity, openness, altruism and commitment to the common good. Such values also include trustworthiness and respect for others.

Among the most consequential professional values held by successful educational leaders are belief that the best interests of all students should be at the center of decision making (e.g., NPBEA ISSLC standards, 2015). Other widely endorsed professional values include:

- Social justice and the equitable treatment of students no matter their economic conditions, sexual orientation, cultural and ethnic background or religious affiliation.

- Commitment to high standards of leadership practice.
- Commitment to act ethically and professionally in one's personal conduct and all aspects of one's leadership to the benefit of all stakeholders.
- The responsibility to safeguard and promote the values of democracy, individual freedom and responsibility, equity, social justice, community, and diversity.
- The responsibility to provide moral direction for all members of the organization, the district's staff and elected Board.

These professional values reflect, in education contexts, accounts of ethical leadership reviewed in Freeman et al. (2006). Most are also a direct reflection of values embedded in current educational policy priorities for which school and district leaders are responsible as, for example, a commitment to equity and social justice and the primacy of students' best interests.

Conclusion

This paper has explored the meaning of integrated educational leadership (IEL), reviewed a significant sample of evidence about the effects of IEL on student outcomes, and pointed to a small handful of conditions mediating the effects of such leadership on those outcomes. The *Ontario Leadership Framework* has been used to illustrate possible key features of IEL models, as well as to demonstrate the value of adding "personal leadership resources" or dispositions to models of IEL which are currently limited to the combination of instructional and transformation leadership practices. This addition extends the proportion of leaders' actual work significantly beyond the narrow representation of that work accounted for by most leadership models and begins to explain variation across leaders in their responses to the same context or situation.

References

- Avey, J., Avolio, B., Luthans, F., 2011. Experimentally analyzing the impact of leader positivity on follower positivity and performance. *Leader. Q.* 22 (2), 282–294.
- Avolio, B.J., 1994. The alliance of total quality and the full range of leadership. In: Bass, B.M., Avolio, B.J. (Eds.), *Improving Organizational Effectiveness Through Transformational Leadership*. Sage, Thousand Oaks, CA, pp. 121–145.
- Bandura, A., 1997. Self-efficacy: toward a unifying theory of behavior change. *Psychol. Rev.* 84, 191–215.
- Bass, B.M., Avolio, B.J., 1994. *Improving Organizational Effectiveness through Transformational Leadership*. Sage, Thousand Oaks, CA.
- Bass, B.M., 1985. *Leadership and Performance beyond Expectations*. The Free Press, New York.
- Blase, J., Blase, J., 2000. Effective instructional leadership. *J. Educ. Adm.* 38(2), 130–141.
- Boberg, J., Bourgeois, S., 2016. The effects of integrated transformational leadership on achievement. *J. Educ. Adm.* 56(4), 357–374.
- Brown, M., Trevino, L., 2006. Ethical leadership: a review and future directions. *Leader. Q.* 17, 595–596.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement Lessons From Chicago*. University of Chicago Press, Chicago.
- Burch, G., 2013. *Interactive Empathy and Leader Effectiveness: An Evaluation of How Sensing Emotion and Responding With Empathy Influence Corporate Leader Effectiveness*. Virginia Commonwealth University. PhD. Dissertation. <https://scholarscompass.vcu.edu/etd/540>.
- Burns, J.M., 1978. *Leadership*. Harper and Row, New York.
- Bush, T., 2015. Understanding instructional leadership. *Educ. Manag. Adm. Leader* 43 (4), 487–489.
- Conger, J., 2015. Charismatic Leadership. *Wiley Encyclopedia of Management*, pp. 1–2.
- Cranston, N., Ehrich, L., Kimber, M., 2014. Managing ethical dilemmas. In: Branson, C., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York.
- Crant, J.M., Bateman, T.S., 2000. Charismatic leadership viewed from above: the impact of proactive personality. *J. Organ. Behav.* 21 (1), 63–75.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 44 (5), 635–674.
- Dwyer, D., 1985. Understanding the principal's contribution to instruction. *Peabody J. Educ.* 63 (1), 3–18.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader* 37, 15–24.
- Freeman, R., Martin, K., Parmar, B., Cording, M., Werhane, P., 2006. Leading through values and ethical principles. In: Burke, R., Cooper, C. (Eds.), *Inspiring Leaders*. Routledge, New York (Chapter 8).
- Gerstner, C.R., Day, D.V., 1997. Meta-analytic review of leader-member exchange theory: correlates and construct issues. *J. Appl. Psychol.* 82 (6), 827–844.
- Gottfredson, R., Reina, C., 2020. Exploring why leaders do what they do: an integrative review of the situation-trait approach and situation-encoding schemas. *Leader. Q.* 39, 101373.
- Greenleaf, R.K., 1996. *On Becoming a Servant Leader*. Jossey-Bass, San Francisco.
- Hallinger, P., Murphy, J., 1985. Assessing the instructional leadership behavior of principals. *Elem. Sch. J.* 86 (2), 217–248.
- Hallinger, P., 2003. Leading educational change: reflections on the practices of instructional and transformational leadership. *Camb. J. Educ.* 33 (3), 329–351.
- Hallinger, P., 2005. Instructional leadership and the school principals: a passing fancy that refuses to go away. *Leader. Pol. Sch.* 4, 221–239.
- Hallinger, P., 2009. *Leadership for 21st Century Schools: From Instructional Leadership to Leadership for Learning*. Hong Kong Institute of Education, Hong Kong.
- Harris, A., 2013. *Distributed leadership matters: Perspectives, practicalities, and potential*. Corwin Press.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681.
- Judge, T.A., Bono, J.E., Ilies, R., Gerhardt, M.W., 2002. Personality and leadership: a qualitative and quantitative review. *J. Appl. Psychol.* 87 (4), 765.
- Keleman, T., Matthews, S., Breevaart, K., 2020. Leading day-to-day: a review of the daily causes and consequences of leadership behaviors. *Leader. Q.* 31, 101344.
- Kenny, D.A., Zaccaro, S.J., 1983. An Estimate of Variance Due to Traits in Leadership, p. 678.
- Khalifa, M., Gooden, M., Davis, J., 2016a. Culturally responsive school leadership: a synthesis of the literature. *Rev. Educ. Res.* 86. <https://doi.org/10.3102/0034654316630383>.
- Khalifa, M., Gooden, M., Davis, J., 2016b. Culturally responsive school leadership: a synthesis of the literature. *Rev. Educ. Res.* 86 (4), 1272–1311.
- Knapp, M., Honig, M., Plecki, M., Portin, B., Copeland, M., 2012. *Learning-Focused Leadership in Action: Improving Instruction in Schools and Districts*. Routledge, New York.
- Kwan, P., 2020. Is transformational leadership theory passe? Revisiting the integrative effect of instructional leadership and transformational leadership on student outcomes. *Educ. Adm. Q.* 56 (2), 321–349.
- Leithwood, K., Jantzi, D., 2005. A review of transformational leadership research 1996–2005. *Leader. Pol. Sch.* 4 (3), 177–200.

- Leithwood, K., Riehl, C., 2005. What we already know about successful school leadership. In: Firestone, W.A., Riehl, C. (Eds.), *A New Agenda: Directions for Research on Educational Leadership*. Teachers College Press, New York, NY.
- Leithwood, K., Steinbach, R., 1995. *Expert Problem-Solving Processes: Evidence From Principals and Superintendents*. State University of New York Press, Albany, NY.
- Leithwood, K., Jantzi, D., Steinbach, R., 1999. *Changing Leadership for Changing Times*. Open University Press, Buckingham, UK.
- Leithwood, K., Day, C., Sammons, P., Hopkins, D., Harris, A., 2008. *Successful School Leadership: What is It and How It Influences Pupil Learning*. Department for Education and Skills, Nottingham, England.
- Leithwood, K., Patten, S., Jantzi, D., 2010. Testing a conception of how school leadership influences student learning. *Educ. Adm. Q.* 46 (5), 6712–706.
- Leithwood, K., 2012, 2020. *The Ontario Leadership Framework*. The Institute for Educational Leadership, Toronto.
- Little, L., Gooty, J., Williams, M., 2016. The role of leader emotion management in leader–member exchange and follower outcomes. *Leader. Q.* 27 (1), 85–97.
- Luthans, F., Luthans, K., Luthans, B., 2004. Positive psychological capital: beyond human and social capital. *Bus. Horiz.* 47, 45–50.
- Marks, S., Printy, H., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Mayer, J., Salovey, P., 1997. What is emotional intelligence? In: Salovey, P., Sluyter, D. (Eds.), *Emotional Development and Emotional Intelligence: Implications for Educators*. Basic Books, New York, pp. 3–31.
- Mumford, M., Hunter, S., Eubanks, D., Bedell, K., Murphy, S., 2007. Developing leaders for creative efforts: a domain-based approach to leadership development. *Hum. Resour. Manag. Rev.* 17, 402–417.
- Murphy, J., 1990. Principal instructional leadership. *Adv. Educ. Adm.* 1 (Part B), 163–200.
- Peterson, R.S., Smith, D.B., Martorana, P.V., Owens, P.D., 2003. The impact of chief executive officer personality on top management team dynamics: one mechanism by which leadership affects organizational performance. *J. Appl. Psychol.* 88 (5), 795–808.
- Podsakoff, P., MacKenzie, S., Moorman, R., Fetter, R., 1990. Transformational leaders' behaviors and their effects on followers' trust in leader, satisfaction and organizational citizenship behavior. *Leader. Q.* 1 (2), 107–142.
- Quinn, R.E., Rohrbaugh, J., 1983. A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis. *Manag. Sci.* 29, 363–377.
- Robinson, V., Lloyd, C., Rowe, K., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Rosenholtz, S.J., 1985. Effective schools: interpreting the evidence. *Am. J. Educ.* 93 (3), 352–387.
- Scheerens, J., 2012. Summary and conclusion: instructional leadership in schools as loosely coupled organizations. In: Scheerens, J. (Ed.), *School Leadership Effects Revisited: Review and Meta-Analysis of Empirical Studies*. Springer, London, pp. 131–152.
- Schutte, N., Malouff, M., Thorsteinsson, E., 2013. Increasing emotional intelligence through training. *Int. J. Emot. Educ.* 5 (1), 56–72.
- Sebastian, J., Allensworth, E., 2012. The influence of principal leadership on classroom instruction and student learning: a study of mediated pathways to learning. *Educ. Adm. Q.* 48 (4), 626–663.
- Sebastian, J., Allenworth, E., Huang, H., 2016. The role of teacher leadership in how principals influence classroom instruction and student learning. *Am. J. Educ.* 123, 69–108.
- Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28 (3), 463–488.
- Shaked, H., Schechter, C., 2017. *Systems Thinking for School Leaders: Holistic Leadership for Excellence in Education*. Springer Publishers.
- Siangchokyo, N., Klinger, R., Campion, E., 2020. Follower transformation as the linchpin of transformational leadership theory: a systematic review and future research agenda. *Leader. Q.* 31, 101341.
- Slaski, M., Cartwright, S., 2003. Emotional intelligence training and its implications for stress, health and performance. *Stress Health* 19, 233–239.
- Stogdill, R.M., 1974. *Handbook of Leadership: A Survey of Theory and Research*. Free Press.
- Supovitz, J., Sirinides, P., May, H., 2010. How principals and peers influence teaching and learning. *Educ. Adm. Q.* 46 (1), 31–56.
- ten Bruggencage, G., Luyten, H., Scheerens, J., Sleegers, P., 2012. Modeling the influence of school leaders on student achievement: how can school leaders make a difference. *Educ. Adm. Q.* 48 (4), 699–732.
- Zeni, T., Buckley, M., Mumford, M., Griffith, J., 2016. Making “sense” of ethical decision making. *Leader. Q.* 27, 838–855.

Framing educational leadership as a *multilevel distributed practice*: a systemwide perspective

James P. Spillane^a, Richard Paquin Morel^b, and Asmaa Al-Fadala^c, ^aNorthwestern University, Evanston, IL, United States; ^bResearch Scientist, San Francisco, CA, United States; and ^cWorld Innovation Summit for Education (WISE), Qatar Foundation, Doha, Qatar

© 2023 Elsevier Ltd. All rights reserved.

Introduction	82
Educational leadership framed from a distributed perspective	83
Educational leadership framed from a multilevel distributed framework	83
Conclusion	88
References	89

Introduction

Deliberations about educational leadership tend to be person-based or place-based. We write about teacher leadership, principal or headteacher leadership, student leadership, and so on. We study department leadership, school leadership, and increasingly system leadership. *Yet*, educational leadership is neither place nor person bound. While what a principal or headteacher does is important to educational leadership as experienced by students, parents, and school staff, it rarely matters in isolation. It interacts with other actors and artifacts within and beyond the school, implicating a wide variety of people and organizations. Framing educational leaderships as a systemwide phenomenon focuses attention on how leadership practice is constituted in the interactions among actors (assistant principals, teacher leaders, system leaders, students, parents, teachers) and artifacts that stretch from the classroom to the principal's office and beyond to the system hub whether that is a Local Educational Authority, a provincial or ministry of education, or IB central. Our article sketches a multilevel distributed framework for doing diagnostic, design, and development work about educational leadership that embraces the *systemwide* nature of the phenomenon.^a

We outline and argue for a multilevel distributed framework in order to explore educational leadership *systemically*; that is, as a *systemwide* phenomenon. Our approach should not be confused with arguments for studying system leadership, though we agree with such arguments and incorporate systems into our discussion. Rather, system leadership is just one level of consideration in the multilevel distributed framing we develop in this article. Our multilevel distributed perspective also takes into account that actors and artifacts stretched across multiple levels of education systems may interact in different ways depending on the particular educational system and the national educational sector in which the system operates. Our framework enables comparative work on educational leadership.

Our article is organized like this: We begin by motivating and anchoring our work in research in the distributed leadership tradition. We then outline the elements of our multilevel distributed framework for thinking about educational leadership. We do so by anchoring our framework in teaching which is the core technology of education systems. We then map backwards from the classroom sketching a multilevel framework that takes the school, community, educational system, and educational sector into account. We conclude with a discussion of the entailments of the framework for diagnostic, design, and development work on educational leadership.

Some discussion of the construct of leadership is always wise at the outset, as there is no shortage of definitions. Definitions of leadership and attempts to operationalize them for diagnostic, design, and development work vary widely. Indeed, people often do not explicitly discuss their definitions of the construct. So, in an effort at clarity at the outset we offer a working definition of leadership and educational leadership in particular. By educational leadership we mean social relationships or interactions tied to the core work of schooling—teaching—that are intended or understood to influence the motivation, knowledge, affect, and practice of educational stakeholders including teachers, students, parents, school and system staff, and others who have a stake in the educational sector (Spillane, 2006; Bass, 1990). Defining leadership in this way anchors our discussion of the phenomena in supporting, improving, and maintaining the quality of teaching, the core technology of educational systems, something we return to in the next section.^b

^aWe include empirical research under diagnostic work. We use diagnostic rather than research in order acknowledge that an array of actors in the education sector do diagnostic work though for many it would not count as empirical research. Yet such work is critical to education and its improvement.

^bThere may be other important ways that school leaders can improve outcomes beyond teaching and learning, such as through building social relationships between the school and community institutions. However, we put aside such other functions to focus on the core enterprise of schooling—teaching and learning.

Educational leadership framed from a distributed perspective

One development in the field of educational administration and leadership over the past few decades involved attention to the practice of leadership; that is, how the work of leadership is actually accomplished inside school. Some work on practice centers on understanding the actions or behaviors of individual school leaders, especially head teachers or principals (Blase and Blase, 1998, 1999; Blase and Kirby, 2000). Others have argued for framing practice differently, not simply a matter of the actions or behaviors of individual leaders but centered on those interactions that are the essence of social influence relationships. Framed from a distributed perspective, educational leadership involves two essential components aspects; first the practice aspect from which the second aspect—the leader-plus—follows (Spillane, 2006; Spillane et al., 2008; Spillane and Ortiz, 2016; Spillane, 2020).

The practice aspect foregrounds the *practice* of leadership – *leading* – and frames our attention on what leaders and followers actually *do together* and how they do it. From a distributed perspective, practice is framed in a particular way: Drawing on activity theory, distributed cognition and socio-cultural activity theory (Hutchins, 1995a, 1995b), a distributed perspective frames leadership practice as unfolding in the interactions among people as these interactions are enabled and constrained by aspects of their situation. A distributed perspective sees practice as taking shape in the interactions among leaders and followers (acknowledging that people move in and out of these roles), and aspects of their situation such as tools, routines, norms, regulations, and so on (Gronn, 2002; Spillane et al., 2001, 2004; Spillane, 2006). To understand the practice of educational leadership from a distributed perspective we have to get beyond an exclusive focus on the actions of individual leaders and get to the *interactions* among them. From a distributed perspective, aggregating the actions individual leaders to understand leadership practice is insufficient because people (teachers, school administrators, parents, students) co-enact leadership practice in their everyday interactions with one another.

From a distributed perspective, the situation is afforded an especially prominent position in that interactions among people in schools are fundamentally defined – enabled and constrained – by aspects of their *situation*. Often, the situation is framed as the stage on which leaders' practice and when aspects of the situation are taken into consideration, they are typically cast as aids or accessories to practice that enable educational leaders to work more or less efficiently or effectively. For example, new tools (e.g., data analysis package, online tools for hosting meetings) or organizational routines (e.g., Learning Walks, Instructional Rounds) framed as enabling people to practice in more efficient or effective ways. From a distributed perspective, this conventional approach to framing the situation is inadequate for understanding practice because aspects of the situation are positioned differently in leadership practice as fundamentally defining practice. These often taken for granted aspects of the situation are an essential constituting element of leadership practice because they focus and frame interactions among people in particular ways – who interacts with whom, about what, and how they interact with one another. Viewed this way, aspects of the situation are not external to leadership practice, influencing it from the outside in, moderating or mediating practice, but rather are a core constituting element of that practice similar to leaders and followers. A distributed perspective frames practice as taking place in the interactions among human and nonhuman actors.

By way of example consider two ubiquitous aspects of any organization – organizational routines and tools or artifacts. Organizational routines such as teacher evaluations or supervisions, faculty and grade-level or high school department meetings, school improvement planning, student assemblies, among others, are staples in schools and essential to daily work practice in the schoolhouse. Similarly, tools as externalized representations of ideas that people use in everyday practice (Norman, 1988), such as classroom observation protocols, test score reports, and lesson plans, figure prominently in educational leadership practice. From a distributed perspective, these two aspects of the situation – organizational routines and tools – do not simply influence or accessorize educational leadership practice, but rather define that practice as they are used by school staff in their interactions with one another by shaping who interacts with whom and the substance of their interactions. Aspects of the situation such as tools and routines are never neutral as they privilege some ideas and ways of working and position some organizational members differently (e.g., as experts, with more authority) than others. These often taken for granted and innocuous appearing aspects of the situation reify and reinforce, sometimes challenge, authority and influence social interactions.

The leader-plus aspect follows from the practice aspect recognizing that there are multiple leaders, some of whom may have no formal leadership position, and understanding who they are and how they work is essential. Employing a distributed framework urges us to identify who does the work of educational leadership, that is, who takes responsibility for leading and managing teaching. A distributed framework presses us to move beyond an exclusive focus on the principal and explore how leadership might be distributed across other formal leaders such as assistant principals, teacher leaders, and so on. The framework also allows for individuals with no formal leadership position taking responsibility for educational leadership. From a distributed perspective then, understanding who takes responsibility for educational leadership and how that work is arranged and stretched over people is an important challenge (Spillane, 2006; Spillane and Coldren, 2011).

Educational leadership framed from a multilevel distributed framework

Much of the empirical work taking a distributed perspective has used the schoolhouse as the unit of analysis. While this work has managed to foreground educational leadership practice, it has mostly focused on that practice in the schoolhouse; one might say the horizontal distribution of leadership practice among school staff and stakeholders. As we argue above, however, educational leadership practice is neither place nor person bound. To begin with, schools serve communities that, more or less, organize to exercise influence over education. At the same time, schools typically are part of an education system (sometimes more than one system)

and, more broadly, operate in an educational sector that generates essential resources for schools and sometimes attempt to regulate work in schools. Education systems, some more than others and in distinctly different ways, attempt to guide and support teaching and learning in schools through various means including teacher preparation, student assessments, curricular materials, school inspections, and so on. Depending on the country, the organization of education systems differ as does the extent to which the educational sector is regulated by government or non-government agencies. The organization and workings of education systems and the broader national educational sector are, more or less depending on the system and country, consequential for efforts to lead and manage teaching and learning inside schools. System leaders such as inspectors from a national or provincial ministry of education or instructional coaches from Local Educational Authorities are a potential source of education leadership. Further, many aspects of the situation that are constitutive of leadership practice inside schools are a product of system level design efforts such as school review or improvement routines that are often mandated at the system level. Hence, it is necessary to think about the distributed perspective not just horizontally at the school level, but also vertically across levels of the education system in order to understand how the practice of educational leadership is stretched across classrooms, grade levels, schools, educational systems, and sectors.

Fortunately, research on educational leadership over the past quarter century has moved beyond an exclusive focus on the school-level and the school principal to explore other sources of leadership. Research on teacher leadership makes a case for going beyond the school principal and other school administrators. At the same time, several scholars have shrewdly drawn attention to leadership as a system-level phenomenon. Work in these traditions have contributed to broadening our understanding of educational leadership beyond the schoolhouse. A key challenge that remains, however, is figuring how these various sources of leadership, sometimes operating at different levels and typically studied independently of one another, work in tandem (or not) in an effort to influence teaching. The tendency to compartmentalize leadership by pigeonholing principal leadership, teacher leadership, system leadership and so on, contributes to a disjointed portrayal of educational leadership practice. After all, when it comes to the actual work of teaching and learning, the interrelationships among leadership at different levels is what will matter ultimately for teachers and students, not the independent influence of any one source of leadership.

A multilevel distributed framework seeks to conceptualize leadership as spanning classrooms, school, system, and the educational sector. A core challenge involves figuring how leadership operating at different levels, typically studied independent of one another, work in interaction with each other (or not) to support teaching. Moving beyond focusing on educational leadership at any one level of an education system, such as the grade level, school, district office or ministry levels, a multilevel distributed framework frames the practice of educational leadership as stretched over the interactions of actors and artifacts both within different levels and spanning different levels. It frames educational leadership systemically, as a systemwide phenomenon. A multilevel distributed perspective focuses on how people and aspects of their situation at different levels of the educational interact, not only within but also across levels, to define educational leadership practice. We begin by anchoring the multilevel framework in teaching, the core technology of educational systems, and then engage in a “backward mapping” exercise from classrooms to schools, communities, educational systems, and educational sectors (Elmore, 1979).

Teaching. Any attempt to frame educational leadership as a multilevel distributed practice must engage the challenge of maintaining consistency when working across different levels. We address this challenge by anchoring our multilevel distributed framing in teaching. Differentiating teaching from learning, we argue that teaching is the primary means that education systems have for enabling student learning.^c While student learning may be the ultimate object or goal of educational leadership, we cannot afford to sideline the means that schools and school systems have for getting there – teaching. Educational leadership can enable improvement in student learning in great part by cultivating the conditions that support and maintain the quality of teaching and enable its improvement. In this way, teaching is not only the object of educational leadership it is also its subject.

Teaching as a practice is understood in various ways from the behaviors of the individual teacher to a collective practice co-performed by teachers and students with particular materials. For some, teaching is what the teacher does and says – the actions of the classroom teacher. We embrace a different understanding of teaching, seeing it as a distributed practice, co-performed by teachers and their students as they interact with one another using particular materials (e.g., mathematics, textbooks) (Cohen and Ball, 1999; Cohen, 2011; Delpit, 1995; Freire, 2018). Though the teacher is critical, so are students by virtue of how they interact with the teacher and one another, the ideas they voice, and their ways of being in these interactions. At the same time, how the teacher engages students, both individually and collectively, and treats their ideas and ways of knowing is also constitutive of teaching practice. Teaching is fundamentally about teacher-student relationships that are framed and focused by materials of various sorts, including the school subject. Interactions between teachers and students as they co-perform teaching mathematics or writing are enabled and constrained by the materials they work with including texts, tests, manipulatives, and others.

Framed this way, teaching is a distributed practice stretched over teachers and students and the materials that enable and constrain their interactions with one another, including the cultural resources that students “bring” to the classroom. By participating in cultural practices such as dominoes, dice, basketball, story-telling, rapping, bartering, and dancing, children gather cultural resources that enable them to develop an identity and a sense of belonging (Lee, 1995; Nasir, 2002; Saxe, 1988; Taylor, 2009). But these cultural resources often go ignored, and are sometimes even shunned, in schools and classrooms, contributing to feelings of alienation and not being valued as students (Farkas et al., 1990; Nasir and Saxe, 2003). Students’ ideas and ways of knowing often are seen by teachers as “deficit” (Ainscow, 2005; Clycq et al., 2014; Garcia and Guerra, 2004; Valencia, 2012). When teachers fail to

^cSee Cohen, 1988.

notice and recognize the cultural resources that students have developed from their everyday experiences at homes and in their neighborhoods, resources that could support teaching practice go untapped and unused in teaching. Teaching then is an *embedded* practice in that relationships among teachers and students in the classroom are fundamentally shaped by relationships with others beyond the classroom such as children's relationships with peers in their neighborhoods or with siblings and parents; relationships that involve everyday experiences with numeracy, literacy, science, and nature that do not evaporate once students enter the classroom and co-perform teaching with their teacher.

Relationships and Resources for Teaching. While relationships among teachers and students are fundamental to teaching, we also see educational leadership as fundamentally about relationships; that is, cultivating and channeling relationships that access and activate resources for supporting teaching as a distributed practice. Whereas cultivating relationships refers to strategically initiating and developing them, channeling relationships denotes focusing and directing these relationships substantively. Accessing resources is chiefly about building, procuring, and distributing resources whereas activating resources refers to noticing and using resources in practice. Though the two are closely interrelated, distinguishing between access and activation draws attention to the fact that resources often go unrecognized as such and unused in practice (Cohen et al., 2003). As noted above, teachers sometimes fail to recognize their students' cultural resources and thereby fail to activate these resources in co-performing teaching with students.

Embracing the embeddedness of teaching as a distributed practice acknowledges that the relationships and resources critical for supporting teaching extend beyond the classroom and even the schoolhouse to relationships with, and among, an array of other stakeholders such as parents, community members, neighborhood organizations, educational system leaders, among others. These overlapping and interacting relationships are critical because they shape the resources that teachers and students have access to and condition what teachers and students notice (and thereby activate) and the choices they make in the classroom influencing how they co-perform teaching.

Educational Leadership in The Schoolhouse. The work of educational leadership involves strategically cultivating and channeling relationships that support accessing and activating those resources essential for teaching. The relationships that support accessing and activating resources essential for teaching practice don't just happen naturally (Bourdieu, 1986). Yes, teachers interact with one another, but these interactions often are only tangentially about teaching (Coburn and Russell, 2008). Consistent with a distributed perspective, we focus on the practice of educational leadership and acknowledge that practice is not confined to the principal's office or school faculty meetings, but occurs throughout the school as administrators, teachers, coaches, and many others engage in leadership practices. Though schools are relatively flat organizations, educational leadership unfolds within and across levels of the school organization; it is a multi-level phenomenon as formal and informal school leaders engage in cultivating and channeling relationships that are essential for accessing and activating resources for teaching.

Research suggests that the school's formal organizational arrangements, such as leadership positions and grade level assignments, shape social networks in schools (Moolenaar et al., 2011; Spillane et al., 2015). Teachers of the same primary schools grades, for example, are much more likely to interact with one another about their teaching (Spillane et al., 2015; Spillane et al., 2012). Similarly individuals, including regular classroom teachers, who have a formal leadership position, or teachers who report having more professional development on teaching a particular subject are more likely to be sought out by their colleagues for teaching advice (Spillane et al., 2017). Educational leadership can shape relationships among school staff about teaching in ways that contribute to accessing and activating resources for teaching. School leaders can strategically cultivate new relationships among staff by formally assigning new staff members a mentor or coach. A principal or lead teacher can cultivate relations informally by connecting a teacher who is struggling with teaching a particular subject or unit with a successful colleague in that subject or unit. In this way, school leaders strategically broker relations among their colleagues. Formal and informal brokering of relations among staff, however, must be strategic to contribute to building, accessing, and activating resources. Simply creating a new formal position does not mean that relationships will automatically form around the person in that position as intended (Daly et al., 2010). Some school leaders may be better positioned in their school's relationship network to cultivate particular sorts of relationships. Teacher leaders who teach full or part-time, for example, seem to play more prominent roles in brokering relations among teachers about teaching than full-time administrators such as principals and deputy principals (Spillane et al., 2010).

Cultivating supportive staff relationships about teaching, however, involves more than designing routines and creating positions that bring teachers together regularly. Several other conditions have to be met if these relationships are to genuinely contribute to accessing and activating resources for teaching. To work together productively teachers and staff need to develop norms such as trust (Bryk and Schneider, 2002). Without trust, teachers are less likely to engage openly and honestly with one another about their teaching, the sorts of interactions that contribute to the development of new knowledge about teaching that can contribute to improving teaching practice.

Cultivating and channeling relationships for accessing and activating resources to support teaching in the schoolhouse depend in important measure on a reasonably clear and well-developed shared vision for what and how students should learn (Anagnostopoulos and Rutledge, 2007; Cobb et al., 2018; Leithwood et al., 2004). If the curriculum and teaching strategies intended to support it are ambiguous and ill-defined, or if school staff hold radically divergent views, then cultivating and channeling relationships that support teaching is difficult as teachers and school leaders are pursuing different educational goals. Similarly using different curricular materials, texts, and student assessments among classrooms makes cultivating deep, productive relationships about teaching among teachers difficult. Teachers negotiate understandings about teaching developing new knowledge as they use shared materials and deliberate about their use in their classrooms (Lampert et al., 2011). Curricular materials are important for teachers' joint work outside the classroom as they work together to develop resources (Ball and Cohen, 1996; Davis and Krajcik, 2005; Davis et al., 2014).

Shared visions for instruction, shared materials, and shared expectations for student learning are essential because they help coordinate the work of cultivating and channeling relationships for accessing and activating resources to support teaching. The linchpin of coordination is a clear, shared, and mutually acknowledged vision for instructional practice. It entails well-articulated beliefs about high-quality instruction, norms of professionalism, and ambitious, but achievable, goals for improving teaching. Coordination supports two critical factors that enable high-quality instruction; coordination fosters coherence in how human, social, material, and cultural resources are used (Cobb et al., 2018; Newmann et al., 2001), and it provides direction, motivation, and morale by challenging process of instructional improvement (Bryk et al., 2010; Hallinger, 2011). Coordination is about much more than all school staff getting on the same page and saying, doing, and valuing the same things in an effort to send consistent messages to students, teachers, and other stakeholders. It is about getting the material, social, human, and cultural resources into a generative and interactive system support high-quality instruction (Lampert et al., 2011).

The coordination challenge is in great part about structuring interdependencies among relations and resources in ways that support teaching practice—and maintaining that structure over time. To help conceptualize this, we use the idea of an *infrastructure*, connoting a set of entities and the relationships among them required for the operation of an enterprise. An *educational infrastructure* consists the structures and resources that educational systems develop, intended to support and coordinate teaching, maintain its quality, and enable improvement (Cohen et al., 2013, 2018; Hopkins et al., 2013; Peurach and Neumerski, 2015; Spillane et al., 2019). Much of the work of creating an educational infrastructure occurs outside the schoolhouse, though it varies depending on the national educational sector and the educational systems. At the same time, school leaders' access to key resources for supporting teaching, such as skilled and knowledgeable teachers, depends on circumstances beyond the schoolhouse even in the same educational sector. As several studies document, access to quality teachers differs substantially for urban schools compared to suburban schools. Further, in some national educational sectors either a government agency or a national ministry of education exercises influence over who gets to teach and the content of their teacher preparation (e.g., Ireland) whereas in other situations such matters are left to universities with only minimal oversight (e.g., US).

For all these reasons educational leaders in schools depend on their environments for not only access to, but also the coordination of all sorts of relationships and resources that are essential for supporting teaching in schools (Scott and Davis, 2015). Hence, the importance of taking a multilevel distributed framework to research on educational leadership.

Schools and "National" Educational Sectors. Scholars frame the environments in which schools and other organizations operate as organizational fields or, more commonly, as educational sectors. A sector – such as the health sector or the education sector – consists of the set of actors and organizations operating within a given domain, supporting those who provide a product or service. We can think of schools as the focal organizations in the educational sector, but we also have to pay attention to the array of ancillary actors – organizations and individuals – that are critical to the work of schooling. The educational sector comprises a hodgepodge of actors and organizations in addition to schools and educational systems, including government agencies, professional development providers, community and professional organizations, unions, philanthropy, research firms and institutes, supplemental educational providers, and others. These organizations typically do not directly provide schooling but play a supportive role in providing resources (human, social, material), sometimes directly in school or classroom practice (Rowan, 2002). Other organizations serve regulatory functions regarding resources (e.g., curricular materials) and services (e.g., teacher preparation). In some educational sectors, the same organization performs more than one of these functions.

The in-school relationships that are essential for accessing and activating resources are embedded in relationships that extend beyond the schoolhouse to the educational sector. These sectoral organizations are critical to cultivating and channeling relationships that access and activate resources. Indeed, the work of designing, deploying, and maintaining an educational infrastructure involves various actors and organizations in an educational sector, not just schools. Depending on the particular national educational sector and the decisions of regulatory agencies, the presence, form, and role of these organization can vary. These ancillary organizations can influence the work of leadership in the schoolhouse and beyond by shaping the resources school leaders and teachers can access and support the use of resources in practice (Coburn, 2005).

Schools depend on the educational sector in which they operate for legitimacy, and for other essential resources to improve the quality of teaching, and address inequities in students' opportunities to learn. How schools are embedded within an educational sector is a critical consideration for educational leadership in the schoolhouse because it shapes which resources are available, how they can be accessed and their use supported in practice. Governmental organizations may regulate the training and certification of teachers and other professional staff—or this essential resource for teaching may be delegated to quasi-governmental organizations. Alternatively, the training and certification of teachers may go mostly unregulated or with minimal government oversight. Organizations in an educational sector produce and distribute various resources – some provide curricular materials, others develop, administer, and score student assessments, others provide pre- and in-service professional preparation and development. In some national sectors, government and/or non-governmental organizations work to regulate the efforts of these organizations and their provisions. These variations are consequential for leadership in the schoolhouse because they contribute to differences in the work and the challenges that school leaders face as they attempt to cultivate relationships with various organizations in the sector toward supporting teaching.

Some organizations in an educational sector are transnational, raising the question of how to define boundaries. Examples are publishers such as Pearson or organizations that design and administer international student assessments (e.g., PISA), the Organization for Economic Cooperation and Development (OECD). Some educational systems operate across various national educational sectors such as International Baccalaureate and AMI Educational. Although organizational fields are often not constrained by national borders, because of distinct differences among nations, we believe it is useful to think about them, for analytical

purposes, in terms of national educational sectors. A nation's regulatory agencies often have the power to shape the educational sector by authorizing who can provide particular services within national borders.

Schools, Sectors, and Educational Systems. Educational systems are involved, in various ways and to varying degrees, in coordinating the design, deployment, and support of educational infrastructure. Often the terms "school system" and "educational system" are used interchangeably; however, we see them as capturing an important analytical distinction: they refer to different ways of organizing (Peurach et al., 2019; Spillane et al., 2019a,b). An educational system is concerned with the educational function of schooling and schools; it is focused on supporting the day-to-day teaching and learning in schools. Educational systems include the set of central organizations that drive efforts to support instruction and its improvement; these could include district offices (or teams within the office), charter management organization, a diocese office, or even a national or provincial ministry of education (Peurach et al., 2019). System roles and responsibilities can overlap, spanning national, state and provincial governments; some schools belong to two or more educational systems, such as when a local public school district adopts International Baccalaureate. We use "school system" to refer to organizations at the national or regional levels with the legal authority to provide schooling by building and operating schools, sorting students into schools and classrooms, and resourcing classrooms with teachers and sometimes instructional materials. Analytically, the key distinction between school systems and educational systems is that the latter establish the vision for learning and improving instruction, build and coordinate resources to support the vision in school, and addressing inequalities in students' opportunities to learn (Peurach et al., 2019).

There are also religion or faith-based and independent systems which operate with various degrees of government regulation or quasi-government regulation depending on the national educational sector and the particular system. In some countries, national or provincial government ministries regulate all school systems receiving public funding. In others, there is limited government regulation and limited public funding for such systems. Finally, there are international educational systems, like Montessori and the International Baccalaureate, which operate schools with varying levels of involvement from governmental entities depending on where the national educational sector in which they operate. Though educational systems take different forms, they typically include schools and some central agency or hub that more or less manages the core educational functions.

One way in which educational systems matter to educational leadership is that they attempt (some more eagerly and comprehensively than others and in different ways) to coordinate the resources and relationships essential for supporting teaching. Educational systems coordinate the designing and building of educational infrastructures, work to support the use of that infrastructure in school and classroom practice, and manage educational infrastructure and its use in order to improve and maintain the quality of teaching and reduce inequities in opportunities to learn (Peurach et al., 2019; Spillane et al., 2019a,b). Educational infrastructure includes instruments and tools that are the material of teaching, such as curricula and student assessments. It also includes the formal positions, procedures, organizational routines, and tools that educational systems design, acquire, and deploy to support teaching and its improvement. Further, it includes the coordination of core educational functions, including building a technical culture anchored in consistency among the components of infrastructure, and representing a shared vision for teaching and learning. Other core functions include recruiting and hiring professional staff, teacher educational and ongoing professional learning for staff, monitoring teaching quality, and leadership and management and improvement of teaching.

Education systems vary in how the design and deployment of infrastructure is achieved. In some nationalized and centralized systems, a central or provincial ministry often takes a prominent role. In other educational systems, central offices may play a more modest role, leaving the building, maintaining and supporting of educational infrastructure to schools. The work of leadership in schools varies significantly.

In thinking about educational leadership then, we must also take the educational system into account. People in educational systems perform key leadership functions and take on important responsibilities, such as articulating a coherent vision for teaching and learning and for improving teaching. Systems leaders develop, provide access to, and support the use of critical resources, such as curricular materials or professional development (Cobb et al., 2018). The extent and quality of work by system leaders, and how they are organized (managerial, market, federated, networked), has important impact for leadership in the schoolhouse. Further, depending on the country several different educational systems may operate in any one national educational sector. As a result, simple cross-national comparisons of educational leadership are problematic as the work and challenges can differ depending on the educational system and how the national educational sector is organized.

The type of educational system(s) a school belongs to shapes the coordination among actors and organizations in the sector. Some schools are a part to more tightly coordinated centralized systems where much of the work of designing and deploying an infrastructure is done by central offices. Leaders in schools may therefore rely less on other ancillary organizations to access and activate support. In contrast, some schools exist in systems where school leaders shoulder much of the responsibility for reaching out to other organizations in the sector to secure instructional materials or professional training. Charter schools in the United States are an example.

Above, we frame educational leadership in the schoolhouse as a distributed practice; that is, educational leadership is constituted in the interactions among school staff as they are enabled – or constrained – by key aspects of their situation. Educational infrastructure is a critical aspect of that situation. In other words, the nature of an educational system's infrastructure defines the practice of educational leadership in the schoolhouse differently as it focuses and frames interactions among school staff, students, and other stakeholders.

Designing and deploying an educational infrastructure involves multiple actors and organizations across the sector; which actors and organizations and how they are arranged varies depending on the system and the national educational sector. Further, some education systems take a more central role than other systems in designing and deploying educational infrastructures and

supporting the use of those infrastructures in school and classroom practice which is critical for coordinating the resources and relationships essential for supporting teaching. This underscores the importance of taking a multi-level perspective by thinking about leadership as extending over the school and the educational system – and indeed over the educational sector more broadly. This is the subject we examine next.

The work of supporting and improving teaching is not solely the province of school leaders, teachers and other professional staff and stakeholders. Schools are typically part of an educational system or systems,^d and embedded in an educational sector that provides essential resources and support. We conceptualize leadership as a multilevel, distributed process, reflecting the collective nature of designing and deploying an infrastructure that involves several organizations. Moving beyond the schoolhouse, our account makes the case for thinking about leadership most broadly from a multilevel distributed perspective.

We expanded our view of educational leadership by recognizing that (1) schools are embedded in multiple critical relationships for improving teaching; and (2) leadership is distributed across the entire sector in order to best coordinate the design, deployment, and use of an educational infrastructure. Focusing narrowly on individual school responsibility misses the array of relationships that school leaders should cultivate to improve teaching. Education systems vary in how they do the work of coordinating resources and relationships fundamental to successful efforts.

Conclusion

A multi-level distributed framework eschews compartmentalizing leadership by pigeonholing principal leadership, teacher leadership, system leadership and so on because such framings contribute to a disjointed portrayal of educational leadership practice. After all, when it comes to the actual work of teaching and learning, the interrelationships among leadership at different levels is what will matter most, not the influence of any one source of leadership independent of the others.

Our article sketched an analytical framework to address two blind spots in current approaches to educational leadership by exploring the interrelationships among different sources of educational leadership operating at various levels, from classrooms, to schools and beyond, to educational systems and the educational sector *and* by elucidating how the practice of educational leadership is not only distributed within the schoolhouse, but also distributed across educational systems and sectors. Our multilevel distributed framework has entailments for work on educational leadership and educational improvement.

Entailments for Leadership Practice and Developing Practice. To engage effectively with the multilevel distributed framing to educational leadership, suggests that aspiring leaders acknowledge the following. First, with a few notable exceptions (DeFlaminis et al., 2016), efforts to develop educational leadership tend to be role bound and individual-centered. Educational leadership development programs, whether pre-service or in-service, rarely distinguish leadership practice from developing individual leaders. We believe programs need to focus more on developing leadership practice. Whether one is a school or system leader, understanding that it is not all about “me” (the individual) is critical; the practice of educational leadership is about the individual in interaction with others, from across the educational system, as enabled and constrained by aspects of their situation. For aspiring practitioners of an effective multilevel distributed process, this is difficult but critical. Teaching, as both the object and subject of educational leadership, provides an important anchor for working on educational leadership. Educational leadership development must also focus more on preparing leadership teams to engage in the twin processes of diagnosis and design thinking (Breakspear et al., 2017).

Entailments for policy. Educational policymakers in several countries are paying more attention to leadership, perhaps in part because of a growing concern about the quality of teaching and inequalities in students’ opportunities. Educational policymakers in some countries are introducing certification requirements for school leaders, in particular school principals and head teachers. Several governments have introduced legislation that hold the school principal or head teacher accountable for school performance, often as measured by a handful of metrics, typically student performance on standardized tests and attendance. Policymaking on educational leadership tends to target leadership in the schoolhouse, and mostly fixates on the school principal or head teacher. Our multilevel distributed framework suggests that caution is in order regarding these policy initiatives; if policymakers sincerely wish to legislate action to improve educational leadership (whether they should or not is a different issue), they should to adopt a more comprehensive approach as outlined and described in our text.

With respect to policymaking that addresses certification, preparation, and advancement of educational leaders, a systemic approach that focuses on both the educational system and schoolhouse leadership together is essential. A comprehensive policy approach to leadership advancement should attend to everyone from teacher leaders to school administrators and system leaders. Moreover, it should accommodate non-linear movement among positions so that, for example, a deputy head or assistant principal with an opportunity to take a system leadership position could also have the flexibility to move back to a school leadership position.

Entailments for research. Applying a multilevel distributed framework to research on educational leadership argues for attention to sources of leadership not only within but beyond the schoolhouse in the educational sector and seek to understand the interdependencies among these various sources for leadership practice. Research must recognize that most schools are part of an educational system(s), and that we should examine the work of educational leadership as part of this system. Adopting a multilevel

^dWe acknowledge that there are independent schools that do not belong to educational or school systems as defined here, but for the purposes of this brief, we do not attend to these schools.

mindset will necessitate examining leadership as a practice that stretches over the broader educational system, and exploring how leadership responsibilities are distributed and managed across different levels from classrooms to the central offices of the system. Such work involves recognizing how educational systems and educational sectors differ around the world and exploring how these sector and system differences contribute to differences in the practice of educational leadership.

References

- Ainscow, M., 2005. Developing inclusive educational systems: what are the levers for change? *J. Educ. Change* 6 (2), 109–124.
- Anagnostopoulos, D., Rutledge, S.A., 2007. Making sense of school sanctioning policies in urban high schools: charting the depth and drift of school and classroom change. *Teach. Coll. Rec.* 109 (5), 1261–1302.
- Ball, D.L., Cohen, D.K., 1996. Reform by the book: what is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educ. Res.* 25 (9), 6–14.
- Bass, B.M., 1990. From transactional to transformational leadership: learning to share the vision. *Organ. Dynam.* 18 (3), 19–31.
- Blase, J., Blase, J., 1998. *Handbook of Instructional Leadership: How Really Good Principals Promote Teaching and Learning*. Corwin Press, Thousand Oaks, CA.
- Blase, J., Blase, J., 1999. Principals' instructional leadership and teacher development: teachers' perspectives. *Educ. Adm. Q.* 35, 349–378.
- Blase, J., Kirby, P.C., 2000. *Bringing Out the Best in Teachers: What Effective Principals Do*, second ed. Corwin Press, Thousand Oaks, CA.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood, New York, pp. 241–258.
- Breakspear, S., Peterson, A., Alfadala, A., Khair, M.S., 2017. *Developing Agile Leadership of Learning: School Leadership Policy for Dynamic Times*. World Innovation Summit on Education (WISE) Research Report 7.
- Bryk, A., Schneider, B., 2002. *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation, New York.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Easton, J.Q., Luppescu, S., 2010. *Organizing Schools for Improvement: Lessons From Chicago*. University of Chicago Press, Chicago, IL.
- Clycq, N., Ward Nouwen, M.A., Vandenbroucke, A., 2014. Meritocracy, deficit thinking and the invisibility of the system: discourses on educational success and failure. *Br. Educ. Res. J.* 40 (5), 796–819.
- Cobb, P., Jackson, K., Henrick, E.C., Smith, T.M., MIST Team, 2018. *Systems for Instructional Improvement: Creating Coherence From the Classroom to the District Office*. Harvard Educational Press, Cambridge, MA.
- Coburn, C.E., Russell, J.L., 2008. District policy and teachers' social networks. *Educ. Eval. Pol. Anal.* 30 (3), 203–235.
- Coburn, C.E., 2005. Shaping teacher sensemaking: school leaders and the enactment of reading policy. *Educ. Pol.* 19 (3), 476–509.
- Cohen, D.K., Ball, D.L., 1999. Instruction, Capacity, and Improvement. CPRE Research Report No RR-043. Consortium for Policy Research in Education, Philadelphia, PA.
- Cohen, D.K., Raudenbush, S.W., Ball, D.L., 2003. Resources, instruction, and research. *Educ. Eval. Pol. Anal.* 25 (2), 119–142.
- Cohen, D.K., Peurach, D., Glazer, J.L., Gates, K.E., Goldin, S., 2013. *Improvement by Design: The Promise of Better Schools*. University of Chicago Press, Chicago, IL.
- Cohen, D.K., Spillane, J.P., Peurach, D.J., 2018. The dilemmas of educational reform. *Educ. Res.* 47 (3), 204–212.
- Cohen, D.K., 1988. *Teaching Practice: Plus ça Change*. National Center for Research on Teacher Education, East Lansing, MI.
- Cohen, D.K., 2011. *Teaching and Its Predicaments*. Harvard College, Cambridge, MA.
- Daly, A.J., Moolenaar, N.M., Bolivar, J.M., Burke, P., 2010. Relationships in reform: the role of teachers' social networks. *J. Educ. Adm.* 48 (3), 359–391.
- Davis, E.A., Krajcik, J.S., 2005. Designing educative curriculum materials to promote teacher learning. *Educ. Res.* 34 (3), 3–14.
- Davis, E., Palincsar, A.S., Arias, A.M., Bismack, A.S., Marulis, L., Iwashyna, S., 2014. Designing educative curriculum materials: a theoretically and empirically driven process. *Harv. Educ. Rev.* 84 (1), 24–52.
- DeFlaminis, J.A., Abdul-Jabbar, M., Yoak, E., 2016. *Distributed Leadership in Schools: A Practical Guide for Learning and Improvement*, 1st ed. Routledge, New York.
- Delpit, L., 1995. *Other People's Children: Cultural Conflict in the Classroom*. The New Press, New York.
- Elmore, R., 1979. Backward mapping: implementation research and policy decisions. *Polit. Sci. Q.* 94 (4), 601–616.
- Farkas, G., Grobe, R., Sheehan, D., Shaun, Y., 1990. Cultural resources and school success: gender, ethnicity, and poverty groups within an urban district. *Am. Socio. Rev.* 55, 127–142.
- Freire, P., 2018. *Pedagogy of the Oppressed*. Bloomsbury Publishing, New York.
- Garcia, S.B., Guerra, P.L., 2004. Deconstructing deficit thinking: working with educators to create more equitable learning environments. *Educ. Urban Soc.* 36 (2), 150–168.
- Gronn, P., 2002. Distributed leadership as a unit of analysis. *Leader. Q.* 13 (4), 423–451.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142.
- Hopkins, M., Spillane, J.P., Jakopovic, P., Heaton, R.M., 2013. Infrastructure redesign and instructional reform in mathematics: formal structure and teacher leadership. *Elem. Sch. J.* 114 (2), 200–224.
- Hutchins, E., 1995a. *Cognition in the Wild*. MIT Press, Cambridge, MA.
- Hutchins, E., 1995b. How a cockpit remembers its speeds. *Cogn. Sci.* 19 (3), 265–288.
- Lampert, M., Boerst, T.A., Graziani, F., 2011. Organizational resources in the service of school-wide ambitious teaching practice. *Teach. Coll. Rec.* 113 (7), 1361–1400.
- Lee, C.D., 1995. A culturally based cognitive apprenticeship: teaching African American high school students skills in literary interpretation. *Read. Res. Q.* 608–630.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K., 2004. How Leadership Influences Student Learning. Review of research. The Wallace Foundation, New York.
- Moolenaar, N.M., Daly, A.J., Sleegers, P.J., 2011. Ties with potential: social network structure and innovative climate in Dutch schools. *Teach. Coll. Rec.* 113 (9), 1983–2017.
- Nasir, N.I.S., Saxe, G.B., 2003. Ethnic and academic identities: a cultural practice perspective on emerging tensions and their management in the lives of minority students. *Educ. Res.* 32 (5), 14–18.
- Nasir, N.I.S., 2002. Identity, goals, and learning: mathematics in cultural practice. *Math. Think. Learn.* 4 (2–3), 213–247.
- Newmann, F.M., Smith, B., Allensworth, E., Bryk, A.S., 2001. Instructional program coherence: what it is and why it should guide school improvement policy. *Educ. Eval. Pol. Anal.* 23 (4), 297–321.
- Norman, D.A., 1988. *The Psychology of Everyday Things*. Basic Books, New York.
- Peurach, D.J., Neumerski, C.M., 2015. Mixing metaphors: building infrastructure for large scale school turnaround. *J. Educ. Change* 16 (4), 379–420.
- Peurach, D.J., Cohen, D.K., Yurkofsky, M., Spillane, J.P., 2019. From mass schooling to educational systems: changing patterns in the organization and management of instruction. *Rev. Res. Educ.* 43 (1), 32–67.
- Rowan, B., 2002. The ecology of school improvement: notes on the school improvement industry in the United States. *J. Educ. Change* 3 (3–4), 283–314.
- Saxe, G.B., 1988. Candy selling and math learning. *Educ. Res.* 17 (6), 14–21.
- Scott, W.R., Davis, G.F., 2015. *Organizations and Organizing: Rational, Natural and Open Systems Perspectives*. Routledge, New York.
- Spillane, J.P., Coldren, A.F., 2011. *Diagnosis and Design for School Improvement: Using a Distributed Perspective to Lead and Manage Change*. Teachers College Press, New York.
- Spillane, J.P., Ortiz, M., 2016. Perspectiva distribuida del liderazgo y la gestión escolar: implicancias cruciales. In: Weinstein, J. (Ed.), *Liderazgo educativo en la escuela: Nueve miradas*. Universidad Diego Portales, Santiago, Chile, pp. 155–176.
- Spillane, J.P., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- Spillane, J.P., Halverson, R., Diamond, J.B., 2004. Towards a theory of leadership practice: a distributed perspective. *J. Curric. Stud.* 36 (1), 3–34.

- Spillane, J.P., Halverson, R., Diamond, J.B., 2008. Théorisation du leadership en éducation : une analyse en termes de cognition située. *Educ. Soc.* 21 (1), 121–149.
- Spillane, J.P., Healey, K., Kim, C.M., 2010. Leading and managing instruction: using social network analysis to explore formal and informal aspects of the elementary school organization. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge, MA, pp. 129–156.
- Spillane, J.P., Kim, C.M., Frank, K.A., 2012. Instructional advice and information providing and receiving behavior in elementary schools: exploring tie formation as a building block in social capital development. *Am. Educ. Res. J.* 49 (6), 1112–1145.
- Spillane, J.P., Hopkins, M., Sweet, T.M., 2015. Intra-and interschool interactions about instruction: exploring the conditions for social capital development. *Am. J. Educ.* 122 (1), 71–110.
- Spillane, J.P., Shirrell, M., Sweet, T.M., 2017. The elephant in the schoolhouse: the role of propinquity in school staff interactions about teaching. *Sociol. Educ.* 90 (2), 149–171.
- Spillane, J.P., Peurach, D.J., Cohen, D.K., 2019a. Comparatively studying educational system (re)building cross-nationally: another agenda for cross-national educational research? *Educ. Pol.* 33 (6), 916–945.
- Spillane, J.P., Seelig, J.L., Blaushild, N., Cohen, D.K., Peurach, D., 2019b. Educational system building in a changing educational sector: environment, organization, and the technical core. *Educ. Pol.* 33 (6), 846–881.
- Spillane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco.
- Spillane, J.P., 2020. Educational leadership from a distributed and multilevel perspective. In: Strauss, N.C., Anderegg, N. (Eds.), *Teacher Leadership: Leading School as a Community*. Hep, der Bildungsverlag, Bern.
- Taylor, E.V., 2009. The purchasing practice of low-income students: the relationship to mathematical development. *J. Learn. Sci.* 18 (3), 370–415.
- Valencia, R.R., 2012. *The Evolution of Deficit Thinking: Educational Thought and Practice*. Routledge, Abingdon, UK.

Distributed leadership: taking a retrospective and contemporary view of the evidence base

Alma Harris, Nashwa Ismail, and Michelle Jones

© 2023 Elsevier Ltd. All rights reserved.

Introduction	91
Distributed leadership: a retrospective view	92
Putting distributed leadership to the empirical test	92
Distributed leadership and student outcomes	93
Distributed leadership- exploring the contemporary evidence	94
Theme 1—exploring relationships	95
Theme 2—different contexts	96
Theme 3—instrument development	96
Theme 4—critique	97
Theme 5—other disciplines	98
Commentary	98
References	99
2001–2010	99
2010–2021 (Review)	99
Other	100

Introduction

Distributed leadership represents one of the most influential ideas to emerge in the field of educational leadership, prompted by the foundational and seminal work of [Spillane et al. \(2001\)](#). The idea of distributed leadership has found favor with researchers, policy-makers practitioners and educational reformers around the globe and has captured the attention of many international researchers (e.g., [Spillane, 2006](#); [Harris, 2008](#); [Leithwood et al., 2009](#); [Gronn, 2002](#)). Few ideas, it seems, have provoked so much attention, debate, and controversy, in the field of educational leadership, as distributed leadership.

Unlike many other fashionable leadership ideas, labels, and theories, distributed leadership has stood the test of time and it could be argued has experienced a renaissance during COVID 19 times, where most leadership practice was virtual, remote, and inevitably distributed ([Harris and Jones, 2020](#)). This chapter looks at the two decades of research that followed the [Spillane et al. \(2001\)](#) article on distributed leadership, primarily in the education sector. It looks at the evidence base in two-halves.

Firstly, it draws upon the literature from 2001 to 2011 mapping out the key empirical terrain, taking a retrospective view. Secondly, it offers a new review of the literature from 2011 to 2021 giving a contemporary assessment of the more recent research evidence concerning distributed leadership. The chapter concludes with some thoughts about the next decade of empirical enquiry and considers whether distributed leadership remains a relevant concept for those working within the field of educational leadership.

In 2001, distributed leadership emerged as a *key* leadership idea of that moment along with ongoing interest in instructional and transformational leadership. While the notion of shared, collaborative, or participative leadership was far from new, distributed leadership theory provided a new lens on a familiar theme. The work of [Spillane et al. \(2001\)](#) sparked renewed interest in leadership as *practice* focusing particularly on the interactions between leaders, followers, and their situation. This seminal work on distributed leadership theory reinforced the multiple sources of influence that operate within any organization and focused empirical attention on the *leader plus* aspect of leadership work ([Spillane, 2006: 3](#)). From a distributed perspective, interactions are a critical part of leadership practice. Analyzing and understanding patterns of influence from a distributed perspective, it was proposed, provides a closer and more nuanced understanding the practice of leadership.

From the outset, however, distributed leadership was an idea not devoid of controversy or critique. Writes like [Fitzgerald and Gunter \(2006\)](#) and [Hargreaves and Fink \(2009\)](#) called into question the motivation of those espousing distributed leadership. In their view, distributed leadership was little more than a palatable way of encouraging teachers to do more work, a way of reinforcing standardization practices, simply old managerialism in a contemporary guise. Instead of being a more democratic form of leadership. [Hargreaves and Fink \(2009\)](#) proposed that distributed leadership could in fact be another, more attractive, mechanism for delivering government policy.

In their critique, [Fitzgerald and Gunter \(2008: 335\)](#) suggested that teacher leadership, which is closely associated with distributed leadership, “merely cements authority and hierarchy whereby leaders monitor teachers and their work to ensure set of predetermined standards are met.” They argued “that teacher leadership is deeply rooted in neo-liberal versions of the performing school and that it is a management strategy and not a radical alternative” ([Fitzgerald and Gunter, 2006: 335](#)).

It is important not to lose sight of these critical perspectives even though, in most cases, there was little empirical evidence to substantiate the claims made. Those opposed to distributed leadership largely came from an ideological position, rather than an empirically grounded one (Fitzgerald and Gunter, 2006). Those researching in the leadership field, however, have taken care to point out the shortcomings and limitations of distributed leadership in their writing (e.g., Harris, 2007).

The next section of this chapter offers a retrospective view of the evidence, focusing on the decade following the Spillane et al. (2001) article.

Distributed leadership: a retrospective view

In the decade from 2001 to 2011, it was consistently argued by those writing about distributed leadership that the practice was pivotally important in securing organizational improvement. Knowing if, how and in what way distributed leadership practice influenced organizational outcomes was at the heart of a great deal of contemporary scholarly enquiry during this period (Leithwood et al., 2009). The available evidence also suggested a positive relationship between distributed leadership, organizational improvement, and student achievement (Harris, 2008, 2009; Hallinger and Heck, 2009; Leithwood and Mascal, 2008). Many studies reinforced the importance of distributed leadership as a potential contributor, under the right conditions, to positive organizational change and improvement.

Even though the evidence base about distributed leadership was still emerging, at this time, the concept had already been enthusiastically adopted within educational policy circles in several countries including the UK, USA, Australia, parts of Europe, and New Zealand. Most of these policies took a normative stance on leadership distribution and emphasized extended forms of leadership at school, district, and system level. For example, in Wales, distributed leadership was a key part of the national school effectiveness framework in 2012. In England, distributed leadership featured heavily in workforce remodeling and reform along with the introduction of new models of schooling such as federations, partnerships, networks, and multi-agency working. In the USA, distributed leadership was a feature of many of the comprehensive school reform approaches.

Despite words of caution from those in the research community, distributed leadership was an idea that was widely advocated and endorsed in many education circles around the world. This proved to be both good news and bad news. On the positive side there was evidence to suggest that certain forms of collective leadership or forms of distributed influence “have a modest but significant indirect effect on student achievement” (Leithwood and Mascal, 2008, p. 546). It was also clear, from the available evidence, however, that *how* leadership was distributed mattered greatly. Early research underlined that distributed leadership was not intrinsically a positive or negative thing, rather it depended upon the purpose, form, and nature of the distribution.

As Leithwood et al. (2009) highlighted in their empirical analysis, purposeful or planned leadership distribution was more likely to impact positively on school development and change. Other researchers similarly reinforced that successful leadership distribution occurs by careful design rather than some random default (Day et al., 2009). The empirical work, at that time, highlighted that it was *how* leadership was distributed (i.e., the patterns of distribution) that explained its subsequent impact.

Putting distributed leadership to the empirical test

In the decade that followed Spillane et al. (2001), distributed leadership became a very slippery concept, often moving between theoretical, descriptive, analytical, normative, and applied interpretations. The term was used liberally to mean distinctively different things and such discrepancy in meaning meant that sometimes researchers simply talked past each other (Mayrowetz, 2008: 425). One common misuse of the term was to align distributed leadership with any form of collaboration thus blurring the precise meaning of the term even further.

It was clear from those writing about distributed leadership, at that time, that different conceptualizations and interpretations of the term persisted and prevailed (Harris, 2007). As Mayrowetz (2008) proposed, it was crucial to inventory the multiple usages of the term distributed leadership for two reasons. Firstly, because of the wide variation in meaning and secondly, to make clearer connections between such usages and the goal of organizational improvement. As Leithwood et al. (2009) underlined, there was little to be gained in continuing to debate the rights and wrongs of different positions or perspectives on distributed leadership. Instead, he proposed that the main challenge was to measure the degree and extent of any organizational impact and effect from distributed leadership.

Apart from wandering definitions, another early issue with distributed leadership was its popular positioning as the antithesis of top-down, hierarchical leadership. This misinterpretation proved to be deeply unhelpful. While distributed leadership is essentially an alternative way of understanding leadership practice, to view it simply as the opposite of formal, autocratic leadership is to misrepresent the concept even further. As researchers at the time underlined, distributed leadership encompassed both formal and informal forms of leadership practice within its framing, analysis, and interpretation. As Spillane (2006: 58) clearly noted, distributed leadership involved the “co-performance of leadership and the reciprocal interdependencies that shaped leadership practice.”

Researchers, at the time, continued to investigate two pressing questions. Firstly, what difference, if any, did patterns or configurations of distributed leadership make to classrooms, schools, and school systems? Secondly, how could we know? On the second of these issues, it was clear that the methodological challenges associated with research into distributed leadership were complex

and considerable. The proliferation of accounts of the single, often “heroic” leaders in the leadership literature was largely explained by the methodological ease of data gathering. If leadership equates with role and position, then it is relatively straightforward to collect evidence about the actions and responsibilities of those occupying leadership positions. It is easy to compile case studies based upon the actions, usually self-reported, of individual leaders. Indeed, the leadership literature is replete with such case studies. Alternatively, if leadership is construed as interaction and the aggregate of a complex set of social processes then the methodological challenges increase exponentially.

Without question, it more difficult to investigate distributed leadership empirically because of the multiple sources of influence. But it is not impossible. As Spillane (2006) showed in his extensive empirical work, data collection methods needed to be more sophisticated and nuanced to capture distributed leadership practice. His research work, with colleagues over this decade, demonstrated that it was perfectly possible to collect data about distributed leadership through new research techniques, tools and alternative methodological approaches.

Earlier work by Spillane et al. (2001) focused on the nature of interdependencies and the co-performance of leadership practices in schools. Implicit in the notion of “co-performance” was the possibility that those performing the practices might be pursuing different or even contradictory goals. “From a distributed perspective, leaders can interact in the co-performance of leadership routines even when they seek different or conflicting outcomes” (Spillane, 2006: 84).

But what do we know from the early empirical evidence about distributed leadership? Firstly, we know that the empirical base about distributed leadership could be found in different research fields and traditions. Early studies that offered research informed insights into distributed leadership were contained in literatures pertaining to school improvement; organizational change; teacher leadership and school leadership. This research terrain was diverse but despite the miscellaneous nature of the early evidential base, consistent messages about distributed leadership and organizational change were forthcoming. For example, there was evidence of a positive relationship between organizational change and wider leadership distribution (Harris, 2008; Leithwood and Mascal, 2008).

Secondly, there were a growing number of early empirical studies that focused attention explicitly on the impact of distributed leadership upon teaching and learning (Leithwood et al., 2009). Such studies provided evidence about the nature, form, and impact of distributed leadership practices in schools. Studies by Hallinger and Heck (2009) highlighted a positive relationship between distributed leadership and certain student outcomes. These studies also reinforced a positive impact on teachers’ level of self-efficacy and motivation. Leithwood et al. (2009) concluded that the “field was now much closer to developing impact studies of distributed leadership because of the theoretical, conceptual and empirical work undertaken.” They noted, however, that the major challenge to the field was to start designing such research studies.

Thirdly, early research studies tended to conceptualize distributed leadership as a form of work-redesign and focused on distributed leadership as a form of job re-design or work restructuring (Harris, 2008). Other researchers paid close attention to the different patterns of leadership distribution in schools (Leithwood et al., 2009; Spillane and Camburn, 2006). The main question pursued, at this time, concerned different patterns of distributed leadership that existed in schools and whether, and in what way, if any, they made a difference to organizational outcomes. As noted earlier, generally the empirical findings suggested that the configuration of leadership distribution was critically important and that certain patterns of distribution had a more positive effect upon organizational development and change (Leithwood et al., 2009).

The early evidence proposed that there were different formations of leadership distribution in schools, some were random, some emergent, and some were carefully orchestrated. Leithwood et al. (2009) suggested that there were different patterns of distribution reflecting differing degrees of active alignment and co-ordination. This insightful work reinforced the importance of planned, aligned distributed leadership practice that is purposeful and focused. Other evidence pointed toward the importance of “principal directed” approaches to distributed leadership (Harris, 2008; Day et al., 2009).

The fourth main finding from the early empirical evidence highlighted a relationship between carefully distributed leadership and organizational performance. As already noted, some writers advocated that distributed leadership was in fact a form of work-redesign. The implication arising from this position was that certain forms of work-redesign created key critical psychological states where employees felt more responsible for their work and were given greater autonomy.

The early evidence also reinforced that distributed leadership had a greater impact upon organizational development where certain structural and cultural conditions were in place (Louis et al., 2009). Within this early literature there was the strong implication that distributed leadership had the potential to positively influence organizational change and student outcomes, under the right conditions. The relationship between distributed leadership and student outcomes is considered next.

Distributed leadership and student outcomes

Without question, the relationship between distributed leadership and student outcomes, in the early evidence base, was the most controversial and contested issue. Positions on the relationship between distributed leadership and student outcomes varied. Some writers argued that seeking to explore this relationship was simply a futile exercise. They suggested that the “search for normative links between specific leadership distribution patterns and student achievement results is unlikely to yield clear guidelines for practice”. Conversely, others argued that distributing leadership is only desirable if the quality of leadership activities contributes to “assisting teachers to provide more effective instruction to their students” (Timperley, 2005: 220).

In essence, there were a limited number of studies that explicitly explored the relationship between distributed leadership and important organizational and student outcomes. Two pivotal studies, however, stand out, offering useful insights into the relationship between distributed leadership and student outcomes. The first study by [Leithwood and Jantzi \(2000\)](#) suggested that distributing a larger proportion of leadership activity to teachers had a positive influence on teacher effectiveness and student engagement. They also noted that teacher leadership had a significant effect on student engagement that far outweighs principal leadership effects after considering home family background.

The second study by [Silns and Mulford's \(2002\)](#) also provided empirical confirmation of the key processes through which more distributed kinds of leadership influence student outcomes. Their work concluded that student outcomes are more likely to improve when leadership sources are distributed throughout the school community and when teachers are empowered in areas of importance to them.

There was other early, empirical work that demonstrated a positive relationship between distributed leadership and student outcomes. A study of teacher leadership conducted in England found positive relationships between the degree of teachers' involvement in decision making and student motivation and self-efficacy ([Harris and Muijs, 2004](#)). This study explored the relationship between teacher involvement in decision making within the school and a range of student outcomes. These findings showed a positive relationship between distributed leadership and student engagement.

[Camburn and Han \(2009\)](#) explored the outcomes of distributed leadership by drawing upon extensive evidence from an investigation into the America's Choice CSR program. A core design feature of this program was the requirement to distribute leadership responsibilities to teacher leaders in schools to act as a key lever for instructional change. Their research work outlined the impact of this program in 30 urban elementary schools and investigated the association between the distribution of leadership to teachers and instructional change. This study concluded that distributing leadership to teachers can support positive instructional change under the right conditions.

[Hallinger and Heck \(2009\)](#) explored the impact of system policies on the development of distributed school leadership and school improvement. Their quantitative analysis and results supported a relationship between distributed leadership and the school's capacity to improve. They concluded that distributed leadership was an important co-effect of school improvement processes. Large scale, mixed methods research by Day et al. (2007: 17) found that substantial leadership distribution was very important to a school's success in improving pupil outcomes. The findings from this study showed that distributed leadership was positively correlated with the conditions within the organization, including staff morale, which in turn impacted positively upon student behavior and student outcomes.

In their work [Leithwood et al. \(2009\)](#) and [Louis et al. \(2009\)](#) highlighted the potentially positive influence of distributed leadership practices on student outcomes. In summary, the early empirical evidence about distributed leadership and student learning was encouraging but it was clear that more work was needed to consolidate and fully explain this relationship.

The next section of this chapter moves into the more contemporary literature on distributed leadership. It draws upon scholarly work produced between 2011 and 2021 to outline the contours of the most recent research about distributed leadership. This is not a systematic review of the literature but rather a review of selected, indicative, empirical work about distributed leadership undertaken in the last decade. Initially, the methodology for the review is outlined and then, the contemporary evidence base is presented thematically.

Distributed leadership- exploring the contemporary evidence

This review of the literature looks at the latest international, contemporary evidence about distributed leadership, largely in education. As noted already, this is not a systematic review of the literature but rather, it offers an overview, based on selected, relevant, and scholarly contemporary evidence. This review focused solely on peer reviewed articles in reputable journals, that were mainly but not exclusively within the field of education. It is fully accepted therefore that there are potentially many books, book chapters and reports on this subject that are not included. The review also discounted opinion pieces, blogs, and other media.

The review focused on peer refereed, academic journals only, with the prime aim of mapping trends in the more contemporary, empirical evidence base. It is fully acknowledged that this review may have missed key work published outside journals and accepts that it has omitted scholarly work in journals written in other languages.

In this review, the data retrieval and analysis process followed the review method guidelines outlined by [Kitchenham \(2004\)](#) that also corresponds with the PRISMA ([2008](#)) methodology. In summary, the stages were:

- (1) Setting Keywords
- (2) Selecting databases
- (3) Defining search criteria
- (4) Selected articles
- (5) Analysis framework
- (6) Extracting information from the articles into the analysis framework

In terms of searching the academic literature, it was decided to use a range of databases to identify relevant articles and to cross-check these databases to ensure that all the retrieved information was accurate, and no publication was missed. EBSCOHOST was used

and the search was subsequently extended to include Web of Science for cross checking. The databases explored within EBSCO-HOST included:

- British Education Index (BEI)
- ERIC
- Education Abstracts
- Education Administration
- Academic Research Complete

The inclusion criteria for this review were:

- Source: refereed journal articles only
- Search Dates 2011–2021
- Language: English
- Search in article: title against the keywords (Distributed Leadership, Distributing Leadership, Leadership Distribution)
- Sources: only peer reviewed journals indexed in SCOPUS/ISI

The data retrieval and analysis process followed four steps, outlined in [Table 1](#) below.

- *Identification*—all studies identified in the initial results of the electronic literature search were recorded.
- *Screening*—studies that did not fully meet the inclusion criteria were excluded and duplicate results were removed from the list.
- *Eligibility*—abstract screening for each output was conducted to ensure only directly relevant and appropriate articles were selected for the next stage.
- *Inclusion*—full PDFs of the selected articles were read to ensure they remained relevant and appropriate for inclusion in the final selection.

The initial results from the search were significantly higher than the final list of articles selected. This is because many of the articles that appeared in the identification stage were irrelevant on closer inspection (e.g., many articles contained the word “distributed” but did not focus explicitly on leadership). Also, many articles were duplicates and therefore were removed. Despite a significant number of articles identified initially in the search, the rigorous filtering process for quality, direct relevance and appropriateness resulted in 44 articles being included finally in this review.

The findings from this contemporary consideration of the literature are summarized under five key themes that feature heavily and consistently in the final list of articles included. It is noticeable that while earlier research indicated that the form or patterns of distributed leadership was a significant explanation for variation in distributed leadership outcomes, the empirical work, reviewed in the next section, rarely explored the issue. The implications for future research into distributed leadership is then assessed and outlined.

Theme 1—exploring relationships

Within the contemporary literature on distributed leadership, large scale, quantitative analyses of relationships feature heavily. There are many studies that have explored the relationship between distributed leadership and some other factor or variable (e.g., academic optimism, trust). In the articles included in this overview, there are a range of relationships explored. For example, [Amels et al. \(2020\)](#) examines the relationship between teachers’ perceptions of distributed leadership and inquiry-based work in primary schools and the resulting impact on those teachers’ capacities to contribute to educational change. The results indicate a direct, positive effect of distributed leadership on teachers’ collaboration and collegiality, as well as on their motivation to contribute to educational change.

Table 1

<i>Step</i>	<i>Process</i>	<i>Number of retrieved journals (EBSCOHOST and Web of science)</i>
1. Identification	This total number of hits from the databases.	669
2. Screening	Studies that did not fully meet the inclusion criteria were excluded and duplicate results were removed from the list.	265
3. Eligibility	Abstract screening for each output was conducted to ensure only directly relevant and appropriate articles were selected for the next stage.	190
4. Inclusion	Full PDFs of the selected articles were read to ensure they remained relevant and appropriate for inclusion in the final selection.	44

Another quantitative study by [Chang \(2011\)](#) explored the relationship between distributed leadership, teachers' academic optimism and student achievement in learning. The study involved 1500 teachers and a structural equation modeling approach was applied to the data. This study found that not only did distributed leadership have a positive influence on academic optimism, but also it indirectly affected student achievement. Other quantitative studies included in this review ([Chen, 2018](#); [Hulpia et al., 2011](#); [Aldaihani, 2019](#); [Tashi, 2015](#)) focused on the relationship between distributed leadership and instructional leadership, organizational commitment, teacher professionalism, applications in high schools and teacher engagement respectively. In all but one of these studies a positive relationship was established.

Distributed leadership within the sphere of digital technologies, and the relationships therein, are the focus of some articles within this review ([Rikkerink et al., 2015](#); [Harris et al., 2013](#)) while the larger body of evidence on distributed leadership continues to explore a wide range of organizational aspects. Since 2011 there has been a burgeoning of studies, in many different countries exploring the relationship between distributed leadership and a wide range of factors or variables (e.g., [Al-Harathi and Mahdy, 2017](#); [Bellibas and Liu, 2016](#); [Coban and Atasoy, 2020](#); [Joo, 2020](#); [Zheng et al., 2019](#)). Most of these studies focus on some aspect of organizational change and largely, but not exclusively, highlight a positive relationship. Such positive relationships would be in keeping with the earlier evidence about the potential of distributed leadership to contribute to organizational change.

Conversely, a review by [Tian et al. \(2015\)](#) concluded, that the positive impact of distributed leadership on organizational change and learner outcomes remains questionable. This review, however, was limited insofar that it did not fully consider many of the key pieces of research, including some in this contemporary overview, that offer empirical confirmation of a positive relationship between distributed leadership and key aspects of organizational change.

This is not to suggest that a relationship between distributed leadership and organizational change is unequivocally established but to propose that there is sufficient empirical evidence, particularly from the contemporary research, to conclude that there is a strong possibility of a positive relationship between distributed leadership and many key organizational aspects.

Theme 2—different contexts

While the early evidence on distributed leadership tended to be Anglo-centric, with studies coming predominantly from North America and Europe, the more recent literature reveals a much broader range of countries and settings for research into distributed leadership. The contemporary literature included in this review illustrates that many more researchers from a wide range of countries and contexts are exploring distributed leadership.

Of note is the fact that there has been an uptake in research into distributed leadership in middle eastern countries. For example, [Aldaihani \(2020\)](#) examined different applications of distributed leadership in high schools in the State of Kuwait from teachers' perspectives. As a quantitative study, the study sample consisted of 1210 high schools, and it concluded that the practice of distributed leadership in Kuwaiti schools was low.

[Al-Harathi et al. \(2017\)](#) similarly examined the prevalence of distributed leadership in Egypt and Oman, using a quantitative approach. This study similarly showed that the adoption of distributed leadership practices was low, and training therefore was required. [Hashem \(2020\)](#) focuses on Jordan and is the first of the articles, in this review, to focus on the way culture affects the possibility of distributed leadership in practice. [Hashem \(2020\)](#) talks about being culturally bound by certain norms and values that make the possibility of distributed leadership practices less likely.

Other studies looked further afield and encompass countries like Turkey where [Coban and Atasoy \(2020\)](#) found a positive relationship between distributed leadership, organizational innovativeness, and teacher collaboration. [García \(2017\)](#) reports on a study into distributed leadership practices in Spain while [Hairon and Goh \(2014\)](#) offer some empirical insights into distributed leadership practices in Singapore, again largely focused on examining collaborative practices and distributed leadership. [Kangas et al. \(2015\)](#) explore distributed leadership in Finland noting that it is generally perceived to be an important part of professional development and learning.

[Harris and DeFlaminis \(2016\)](#) report on a major distributed leadership research project in the United States. The article outlines evidence about the impact of distributed leadership in schools and districts engaged in a large-scale research and development project over many years. This is one of the few longitudinal research studies of distributed leadership practice and the findings suggest a positive relationship between distributed leadership and organizational change.

This review illustrates that research is being undertaken on the topic of distributed leadership in more diverse countries and settings than was the case in the first decade of work. This is very promising and important. Yet, the contextual and cultural contours of these countries is not always given sufficient attention in the analysis or reporting of findings.

In countries where more hierarchical structures are in place and where there are high degrees of accountability, it is unsurprising that levels of distributed leadership practice are low. Conversely, where more lateral, democratic processes are in place distributed leadership is reported to be flourishing. Possibly, cultural, and contextual influences require more attention in studies going forward and it is suggested that such influences have greater explanatory power than currently explored or acknowledged.

Theme 3—instrument development

One of the pleas from researchers in the early phase of research work on distributed leadership was the need for more sophisticated instruments to capture data about distributed leadership practices. There was also a call for more large-scale mixed methods studies

of a longitudinal nature. In the articles selected for this contemporary review, only one article, [Sun and Xia \(2018\)](#) used an existing data base to explore distributed leadership and selected variables. Other studies tended to design new instruments aimed at collecting data about specific aspects of distributed leadership.

[Baloglu \(2012\)](#) focuses on the relationship between value-based leadership and distributed leadership behaviors of school principals. A simple correlational design was employed in the study involving 225 primary teachers. Data were gathered by means of a specially developed “Values Based Leadership and Distributed Leadership Inventory.” Confirmatory and exploratory factor analyses were used to define the construct validity of the scale. The findings showed that there was a 0.74 positive correlation between value based and distributed leadership and that there were four positive dimensions of distributed leadership. These four dimensions were teamwork, support, vision creating and control.

[Blitz and Modeste \(2015\)](#) outline the development and application of a Comprehensive Assessment of Leadership for Learning (CALL) which they advocate is a multi-source assessment of distributed leadership. The authors outline the use of this instrument to examine differences between teacher and leader ratings in assessing distributed leadership practice.

[Carbone et al. \(2017\)](#) describe how they used a “Distributed Leadership Benchmarking Tool” to assess how effectively distributed leadership was enabled at five Australian institutions that were implementing a collaborative teaching quality development scheme premised on bringing teams of academics, co-ordinators etc. together with the prime goal of enhancing teaching and learning. The conclusion drawn is that the benchmarking tool is useful in assessing the alignment (or otherwise) of teaching and learning.

[Mullick et al. \(2013\)](#) report on a study of teachers’ perceptions about distributed leadership for inclusive education in Bangladesh. “A Distributed Leadership Practice for Inclusive Education Scale” was developed specifically for this study. The results indicate that teachers perceived that distributed leadership practices for IE were present in primary schools in Bangladesh. The findings show that teachers’ perceptions about distributed leadership practices for inclusive education have a positive correlation with their satisfaction about the implementation of inclusive education policy.

Of those articles selected as part of this contemporary overview, only [Ersozlu and Ulusoy \(2016\)](#) adapted an existing instrument for their research purposes. The “Distributed Leadership Scale” originally developed by Davis was used to explore aspects of distributed leadership in Turkey.

While the development of new instruments to capture distributed leadership is undoubtedly a positive aspect of the more contemporary literature it is noticeable that firstly, many of these instruments capture the *perceptions* of respondents, and secondly, many of the instruments, except for [Ersozlu and Ulusoy \(2016\)](#) are *single use* only, in other words they are only reported to be used in one research study. This inevitably calls into question the extent to which a cogent knowledge base about distributed leadership can be built without the consistent, repeated use of reliable and valid instruments to collect data in different context and settings.

Theme 4—critique

The early work on distributed leadership, as acknowledged in this chapter, was not without critique. Of the selected contemporary literature within this review, there are critical pieces that follow some of the arguments made about distributed leadership in the early years of empirical enquiry. Indeed, many of the critics of the early work on distributed leadership remain vocal about the topic, a decade on.

[Gunter et al. \(2013\)](#) present distributed leadership as a study in knowledge production. They outline the development of a mapping framework to present and analyze knowledge production and distributed leadership in schools. They propose that their analysis identified the prevalence and political dominance of functional approaches and provides explanations regarding the interplay between state, public policy, and the preferred types of knowledge. [Hall \(2013\)](#) follows a similar line of argument positing that distributed leadership is located within wider structural reform of education in England as part of the New Public management movement in public service delivery. [Hall \(2013\)](#) suggests that there are “officially authorized discourses” of distributed leadership that are imposed on teachers and schools.

[Corrigan \(2013\)](#) considers the rhetoric and reality of distributed leadership, arguing that there are contrasting oppositional messages in the literature and that this calls into question the resurgence of distributed leadership. [Lumby \(2017\)](#) considers distributed leadership in the context of post bureaucratic organizations and proposes that there is little evidence that distributed leadership offers an authentic re-distribution of power. [Woods \(2016\)](#) similarly looks at authority, power, and distributed leadership, as does [Bush \(2018\)](#) who explores leadership and bureaucracy. [Crawford \(2012\)](#) returns to the issue of definitional confusion concerning distributed leadership.

Torrance challenges five generally held assumptions about distributed leadership, while [Mifsud \(2016\)](#) looks at the tension between democracy and autocracy within a distributed leadership scenario. [Or and Berkovich \(2021\)](#) look at participative decision making in schools in individualistic and collectivist cultures which speaks to the point made earlier in this chapter about the importance of acknowledging contextual influences. Such critiques remain an important dimension of the contemporary literature on distributed leadership and reinforce how, a decade later, the concept of distributed leadership is still a source of debate and discussion.

Theme 5—other disciplines

A major difference between the contemporary evidence on distributed leadership and the early work is the introduction of research that emanates from other disciplines. Within the selected review articles, a number focused on distributed leadership within other sectors. [Currie and Lockett \(2011\)](#) examine distributed leadership in practice, drawing on [Spillane's \(2006\)](#) work within the context of health and social care. [Floyd and Fung \(2015\)](#) consider the idea of distributed leadership within universities, noting that there is relatively little written on this topic. The article reports on a funded study which explored how one institution had implemented a newly conceived distributed leadership model. Their findings suggest that the challenge of distributed leadership in universities is complex on several levels that might be antithetical to the strategic direction of the institution.

[Martin et al. \(2014\)](#) describe the utility and impact of a distributed leadership model to implement a National Health Service (NHS) national quality improvement program. They report on the inclusion different groups in decision making. [Jonasson et al. \(2018\)](#) consider the implementation of distributed leadership during a hospital merger suggesting the need to understand the complex dynamics of widened and restricted leadership distribution. [Boak et al. \(2015\)](#) focus on distributed leadership, team working and service improvement in healthcare. [Nayani et al. \(2017\)](#) provide a literature review that looks at occupational safety, health leadership and the management of distributed workers.

The selected contemporary studies, highlighted here, offer only some of the insights into the challenges of distributed leadership within different settings, contexts, and structures. These studies add considerable value to the literature and provide an important comparative element that can offer researchers in the field of education new insights and perspectives.

Commentary

An assessment of the contemporary evidence base confirms that there is still much scholarly interest in the topic of distributed leadership. The sheer scale of research activity suggests that empirical interest in distributed leadership has not waned in the last decade, but conversely, has increased. This conclusion is also supported by [Gumus et al. \(2018\)](#) who found distributed leadership to be the most researched model of leadership followed by instructional leadership and transformational leadership.

It is clear, however, that most of the contemporary studies undertaken of distributed leadership within the education field still tend to be quantitative in orientation and are focused on exploring specific relationships. In contrast, qualitative studies tend to be far fewer in number and smaller in scale. Within the contemporary selected literature, there were six articles that drew upon case study data ([Ng and Ho, 2012](#); [Or et al., 2021](#); [Piot and Kelchtermans, 2015](#); [Tahir et al., 2016](#); [Tam, 2018](#); [Torrance, 2013](#)). The largest sample in these selected articles was six schools. This raises important methodological considerations and highlights the need for more broad-based, mixed method studies that can generate data with the power to generalize beyond the individual school or group of schools.

To conclude, the 64-million-dollar question is does the contemporary evidence about distributed leadership offer any new conclusions? To answer this question, requires a return to the 2008, "Seven Strong Claims about Successful School Leadership" In 2008, it made two strong claims about distributed leadership. In 2020, the authors of this seminal work revisited the claims, based on the evidence that had emerged since 2008. They concluded that a considerable body of relevant evidence had been reported since 2008, significantly increasing certainty about the validity of their claims about distributed leadership. Namely, that school leadership can have an especially positive influence on school and student outcomes when it is distributed (Claim 5) and some patterns of distribution are more effective than others (Claim 6).

A decade ago, it would have been unwise to develop bespoke leadership programs based solely on the idea of distributed leadership. The definitional, operational, and empirical platform was simply not robust enough. Now there is a clearer conceptualization of distributed leadership practice in the contemporary literature and recent empirical work has started to map relationships and sources of influence more carefully and systematically. There is also more research activity, in a diverse range of countries, addressing the outcomes and effects of distributed leadership. After two decades of empirical work on distributed leadership the research base is unquestionably richer and the understanding of the organizational influences from distributed leadership is so much deeper.

So, what has the contemporary evidence about distributed leadership yielded? While the introduction of more large-scale quantitative studies is to be welcomed, the black box of distributed leadership practice remains only partially open, there is more focused and collective empirical work to do. There are strong rays of light, however, in the form of the work around networking and professional community that show the benefits of collaborative working and distributed leadership practices. These offer strong lines of future empirical enquiry ([Harris et al., 2021](#); [Azorin et al., 2020](#)).

To conclude, future research into distributed leadership would benefit greatly from large-scale, longitudinal mixed methods studies, if the knowledge base is to improve significantly. Such empirical work should record, capture, and illuminate the way in which distributed leadership is best nurtured, incubated and supported, so that organizational transformation is possible.

An assessment of the contemporary evidence base would also suggest that more co-ordinated, collaborative, connected sets of scholarly activities could contribute greatly to a better understanding of distributed leadership. Collaborative, comparative research activity, across different contexts or settings, would ensure that the future knowledge base on distributed leadership is grounded offering the real possibility of relevant policy advice and sound professional guidance.

References

2001–2010

- Day, C., Sammons, P., Leithwood, K., Harris, A., Hopkins, D., 2009. The Impact of Leadership on Pupil Outcomes. Final Report. DCSF, London.
- Camburn, E., Han, S.W., 2009. Investigating connections between distributed leadership and instructional change. In: Harris, A. (Ed.), *Distributed Leadership: Different Perspectives*. Netherlands Springer Press.
- Fitzgerald, T., Gunter, H., 2006. Teacher leadership? A new form of managerialism. *N. Z. J. Educ. Leader.* 21 (2), 44–57.
- Gronn, P., 2002. Distributed leadership as a unit of analysis. *Leader. Q.* 13 (4), 423–451.
- Hallinger, P., Heck, R., 2009. Distributed leadership in schools: does system policy make a difference? In: Harris, A. (Ed.), *Distributed Leadership: Different Perspectives*. Netherlands Springer Press.
- Hargreaves, A., Fink, D., 2009. Distributed leadership: democracy or delivery? In: Harris, A. (Ed.), *Distributed Leadership: Different Perspectives*. Springer Press, Netherlands.
- Harris, A., 2007. Distributed leadership: conceptual confusion and empirical reticence. *Int. J. Leader. Educ.* 10 (3), 1–11.
- Harris, A., 2008. Distributed Leadership: Developing Tomorrow's Leaders. Routledge, London.
- Harris, A., 2009. Distributed Leadership: Different Perspectives. Netherlands Springer Press.
- Harris, A., Jones, M., 2020. COVID 19—school leadership in disruptive times. *Sch. Leader. Manag.* 40 (4), 243–247. <https://doi.org/10.1080/13632434.2020.1811479>.
- Harris, A., Muijs, D., 2004. Improving Schools Through Teacher Leadership. Routledge, London.
- Kitchenham, B., 2004. Procedures for Performing Systematic Reviews, vol. 33. Keele University, pp. 1–26.
- Leithwood, K., Jantzi, D., 2000. The effects of transformational leadership on organizational conditions and student engagement with school. *J. Educ. Adm.* 38 (2), 112–129. <https://doi.org/10.1108/09578230010320064>.
- Leithwood, K., Mascal, B., 2008. Collective leadership effects on student achievement. *Educ. Adm. Q.* 44 (4), 529–561.
- Leithwood, K., Mascal, B., Strauss, T., 2009. Distributed Leadership According to the Evidence. Routledge, London.
- Mayrowetz, D., 2008. Making sense of distributed leadership: exploring the multiple usages of the concept in the field. *Educ. Adm. Q.* 44 (3), 424–435.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group, 2008. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2008 (6), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Sillins, H., Mulford, B., 2002. Leadership and school results. In: Leithwood, K., Hallinger, P. (Eds.), *Second International Handbook of Educational Leadership and Administration*. Springer, Dordrecht, pp. 561–612 (2012).
- Spillane, J.P., 2006. Distributed Leadership. Jossey-Bass, San Francisco CA.
- Spillane, J.P., Camburn, E., 2006. The Practice of Leading and Managing: The Distribution of Responsibility for Leadership and Management in the Schoolhouse. American Educational Research Association, San Francisco.
- Spillane, J., Halverson, R., Diamond, J., 2001. Towards a Theory of Leadership Practice: A Distributed Perspective. Northwestern University, Institute for Policy Research Working Article.
- Timperley, H.S., 2005. Distributed leadership: developing theory from practice. *J. Curric. Stud.* 37 (4), 395–420.

2010–2021 (Review)

- Aldaihani, S.G., 2019. Distributed leadership applications in high schools in the State of Kuwait from teachers' viewpoints. *Int. J. Leader. Educ.* 23 (3), 355–370. <https://doi.org/10.1080/13603124.2018.1562096>.
- Al-Harathi, A.S.A., Al-Mahdy, Y.F.H., 2017. Distributed leadership and school effectiveness in Egypt and Oman: an exploratory study. *Int. J. Educ. Manag.* 31 (6), 801–813. <https://doi.org/10.1108/ijem-05-2016-0132>.
- Amels, J., Krüger, M.L., Suhre, C.J.M., van Veen, K., 2020. The effects of distributed leadership and inquiry-based work on primary teachers' capacity to change: testing a model. *Sch. Effect. Sch. Improv.* 31 (3), 468–485. <https://doi.org/10.1080/09243453.2020.1746363>.
- Baloglu, N., 2012. Relations between value-based leadership and distributed leadership: a causal research on school principals behaviours. *Educ. Sci. Theor. Pract.* 12 (20), 1375–1378.
- Bellibas, M.S., Liu, Y., 2016. The effects of principals' perceived instructional and distributed leadership practices on their perceptions of school climate. *Int. J. Leader. Educ.* 21 (2), 226–244. <https://doi.org/10.1080/13603124.2016.1147608>.
- Boak, G., Dickens, V., Newson, A., Brown, L., 2015. Distributed leadership, team working and service improvement in healthcare. *Leader. Health Serv.* 28 (4), 332–344. <https://doi.org/10.1108/lhs-02-2015-0001>.
- Blitz, M.H., Modeste, M., 2015. The differences across distributed leadership practices by school position according to the comprehensive assessment of leadership for learning. *Leader. Pol. Sch.* 14 (3), 341–379.
- Bush, T., 2018. Distributed leadership and bureaucracy: changing fashions in educational leadership. *Educ. Manag. Adm. Leader.* 47 (1), 3–4. <https://doi.org/10.1177/1741143218806704>.
- Carbone, A., Evans, J., Ross, B., Drew, S., Phelan, L., Lindsay, K., Cottman, C., Stoney, S., Ye, J., 2017. Assessing distributed leadership for learning and teaching quality: a multi-institutional study. *J. High. Educ. Pol. Manag.* 39 (2), 183–196. <https://doi.org/10.1080/1360080x.2017.1276629>.
- Chang, I.H., 2011. A study of the relationships between distributed leadership, teacher academic optimism and student achievement in Taiwanese elementary schools. *Sch. Leader. Manag.* 31 (5), 491–515. <https://doi.org/10.1080/13632434.2011.614945>.
- Chen, S., 2018. Influence of functional and relational perspectives of distributed leadership on instructional change and school climate. *Elem. Sch. J.* 119 (2), 298–326. <https://doi.org/10.1086/700279>.
- Crawford, M., 2012. Solo and distributed leadership : definitions and dilemmas. *Educ. Manag. Adm. Leader.* 40 (5), 610–620.
- Coban, M., Atasoy, R., 2020. Relationship between distributed leadership, teacher collaboration and organizational innovativeness. *Int. J. Eval. Res. Educ.* 9 (4), 903. <https://doi.org/10.11591/ijere.v9i4.20679>.
- Corrigan, J., 2013. Distributed leadership: rhetoric or reality? *J. High. Educ. Pol. Manag.* 35 (1), 66–71. <https://doi.org/10.1080/1360080x.2013.748479>.
- Currie, G., Lockett, A., 2011. Distributing leadership in health and social care: concertive, conjoint or collective? *Int. J. Manag. Rev.* 13 (3), 286–300. <https://doi.org/10.1111/j.1468-2370.2011.00308.x>.
- Floyd, A., Fung, D., 2015. Focusing the kaleidoscope: exploring distributed leadership in an English university. *Stud. High. Educ.* 42 (8), 1488–1503. <https://doi.org/10.1080/03075079.2015.1110692>.
- García Torres, D., 2017. Distributed leadership and teacher job satisfaction in Singapore. *J. Educ. Adm.* 56 (1), 127–142. <https://doi.org/10.1108/jea-12-2016-0140>.
- Gunter, H., Hall, D., Bragg, J., 2013. Distributed leadership. *Educ. Manag. Adm. Leader.* 41 (5), 555–580. <https://doi.org/10.1177/1741143213488586>.
- Hairon, S., Goh, J.W., 2014. Pursuing the elusive construct of distributed leadership. *Educ. Manag. Adm. Leader.* 43 (5), 693–718. <https://doi.org/10.1177/1741143214535745>.
- Hall, D.J., 2013. The strange case of the emergence of distributed leadership in schools in England. *Educ. Rev.* 65 (4), 467–487. <https://doi.org/10.1080/00131911.2012.718257>.
- Harris, A., DeFlaminis, J., 2016. Distributed leadership in practice. *Manag. Educ.* 30 (4), 141–146. <https://doi.org/10.1177/0892020616656734>.
- Harris, A., Jones, M., Baba, S., 2013. Distributed leadership and digital collaborative learning: a synergistic relationship? *Br. J. Educ. Technol.* 44 (6), 926–939. <https://doi.org/10.1111/bjet.12107>.

- Hashem, R., 2020. "Al Faza'a" leadership: an implicit cultural barrier to distributed leadership in Jordanian public schools. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/1741143220932580>, 174114322093258.
- Hulpia, H., Devos, G., van Keer, H., 2011. The relation between school leadership from a distributed perspective and teachers' organizational commitment. *Educ. Adm. Q.* 47 (5), 728–771. <https://doi.org/10.1177/0013161x11402065>.
- Jonasson, C., Kjeldsen, A.M., Ovesen, M.S., 2018. Dynamics of distributed leadership during a hospital merger. *J. Health Organ. Manag.* 32 (5), 691–707. <https://doi.org/10.1108/jhom-08-2017-0225>.
- Joo, Y.H., 2020. The effects of distributed leadership on teacher professionalism: the case of Korean middle schools. *Int. J. Educ. Res.* 99, 101500. <https://doi.org/10.1016/j.ijer.2019.101500>.
- Kangas, J., Venninen, T., Ojala, M., 2015. Distributed leadership as administrative practice in Finnish early childhood education and care. *Educ. Manag. Adm. Leader.* 44 (4), 617–631. <https://doi.org/10.1177/1741143214559226>.
- Lumby, J., 2017. Distributed leadership and bureaucracy. *Educ. Manag. Adm. Leader.* 47 (1), 5–19. <https://doi.org/10.1177/1741143217711190>.
- Martin, G., Beech, N., MacIntosh, R., Bushfield, S., 2014. Potential challenges facing distributed leadership in health care: evidence from the UK National Health Service. *Sociol. Health Illness* 37 (1), 14–29. <https://doi.org/10.1111/1467-9566.12171>.
- Mifsud, D., 2016. Distribution dilemmas: exploring the presence of a tension between democracy and autocracy within a distributed leadership scenario. *Educ. Manag. Adm. Leader.* 45 (6), 978–1001. <https://doi.org/10.1177/1741143216653974>.
- Mullick, J., Sharma, U., Deppeler, J., 2013. School teachers' perception about distributed leadership practices for inclusive education in primary schools in Bangladesh. *Sch. Leader. Manag.* 33 (2), 151–168. <https://doi.org/10.1080/13632434.2012.723615>.
- Nayani, R.J., Nielsen, K., Daniels, K., Donaldson-Feilder, E.J., Lewis, R.C., 2017. Out of sight and out of mind? A literature review of occupational safety and health leadership and management of distributed workers. *Work Stress* 32 (2), 124–146. <https://doi.org/10.1080/02678373.2017.1390797>.
- Ng, D., Ho, J., 2012. Distributed leadership for ICT reform in Singapore. *Peabody J. Educ.* 87 (2), 235–252. <https://doi.org/10.1080/0161956x.2012.664478>.
- Or, M.H., Berkovich, I., 2021. Participative decision making in schools in individualist and collectivist cultures: the micro-politics behind distributed leadership. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/17411432211001364>, 174114322110013.
- Piot, L., Kelchtermans, G., 2015. The micropolitics of distributed leadership. *Educ. Manag. Adm. Leader.* 44 (4), 632–649. <https://doi.org/10.1177/1741143214559224>.
- Rikkerink, M., Verbeeten, H., Simons, R.J., Ritzen, H., 2015. A new model of educational innovation: exploring the nexus of organizational learning, distributed leadership, and digital technologies. *J. Educ. Change* 17 (2), 223–249. <https://doi.org/10.1007/s10833-015-9253-5>.
- Sun, A., Xia, J., 2018. Teacher-perceived distributed leadership, teacher self-efficacy and job satisfaction: a multilevel SEM approach using the 2013 TALIS data. *Int. J. Educ. Res.* 92, 86–97. <https://doi.org/10.1016/j.ijer.2018.09.006>.
- Tahir, L.M., Lee, S.L., Musah, M.B., Jaffri, H., Said, M.N.H.M., Yasin, M.H.M., 2016. Challenges in distributed leadership: evidence from the perspective of headteachers. *Int. J. Educ. Manag.* 30 (6), 848–863. <https://doi.org/10.1108/ijem-02-2015-0014>.
- Tam, A.C.F., 2018. Conceptualizing distributed leadership: diverse voices of positional leaders in early childhood education. *Leader. Pol. Sch.* 18 (4), 701–718. <https://doi.org/10.1080/15700763.2018.1513156>.
- Tashi, K., 2015. A quantitative analysis of distributed leadership in practice: teachers' perception of their engagement in four dimensions of distributed leadership in Bhutanese schools. *Asia Pac. Educ. Rev.* 16 (3), 353–366. <https://doi.org/10.1007/s12564-015-9387-4>.
- Tian, M., Risku, M., Collin, K., 2015. A meta-analysis of distributed leadership from 2002 to 2013. *Educ. Manag. Adm. Leader.* 44 (1), 146–164. <https://doi.org/10.1177/1741143214558576>.
- Torrance, D., 2013. Distributed leadership: challenging five generally held assumptions. *Sch. Leader. Manag.* 33 (4), 354–372. <https://doi.org/10.1080/13632434.2013.813463>.
- Woods, P.A., 2016. Authority, power and distributed leadership. *Manag. Educ.* 30 (4), 155–160. <https://doi.org/10.1177/0892020616665779>.
- Zheng, X., Yin, H., Liu, Y., 2019. The relationship between distributed leadership and teacher efficacy in China: the mediation of satisfaction and trust. *Asia Pac. Educ. Res.* 28 (6), 509–518. <https://doi.org/10.1007/s40299-019-00451-7>.

Other

- Azorín, C., Harris, A., Jones, M., 2020. Taking a distributed perspective on leading professional learning networks. *Sch. Leader. Manag.* 40 (2–3), 111–127.
- Gumus, S., Bellibas, M.S., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader.* 46 (1), 25–48.
- Harris, A., Azorin, C., Jones, M., 2021. Network leadership: a new educational imperative? *Int. J. Leader. Educ.* <https://doi.org/10.1080/13603124.2021.1919320>.

Distributed leadership: critical perspectives

Jacky Lumby, Southampton Education School, University of Southampton, Southampton, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	101
Critiques of distributed leadership	101
Distributing what?	101
Power and distributed leadership	102
Bureaucracy	103
Sustaining the weakness of the field	104
Poor historical awareness	104
Theoretical evangelism	104
Insufficient interrogation of research methodology	105
Colonization by research theories	105
Blindness to exclusion	106
The future of leadership theory	106
References	107

Introduction

The worldwide grip of distributed leadership as a favored theory of educational leadership is apparent in the academic and gray literature. Research articles flow from Africa, the Arab nations, Europe, Scandinavia and Anglophone countries. What appears evident is the one facet of distributed leadership on which the field of educational leadership can agree: that since the 1990s the theory has held a dominant role in the research and development of practice (Gumus et al., 2018). Much else is contested, including a definition, whether it is solely a heuristic lens or also a way of leading, and if its impact on learners and faculty has been positive or negative. This article adopts a critical perspective, in broad terms arguing that the theory is flawed and serves purposes other than those that are espoused. It explores some of the assertions, uncertainties, and contradictions of distributed leadership as both concept and practice. It suggests that bureaucracy is a more enduring and universal form in education and that its disparagement and disregard has been to the detriment of the field.

Critiques of distributed leadership

Distributing what?

It is not surprising that there is no agreement on a definition of distributed leadership, as it adds one opaque idea, distribution, to that of another, leadership, which itself is a highly ambiguous and contested concept. Some, such as Lakomski (2008), go so far as to consider the concept of leadership “a folk theory” and “therefore a contender for outright elimination” (p. 161). Historically, the extensive literature on theories of leadership has focused on who and what leaders are, their values, ethics, personality traits and so on, and their actions—that is, what they do, the two being interrelated. Distributed leadership has introduced into this scenario greater emphasis on leadership outcomes as unintended, emerging from the concatenation of intention and the material and social context: “leadership is the outcome both of people’s intentions (intentionality) and the complex flow of interactions in the daily life of schools (emergence)” (Woods and Roberts, 2019: 665).

Though distributed leadership reaches back at least as far as the 1950s as a concept that has much in common with shared leadership, systems leadership and other generic leadership theory, there is a more recent insistence on a theoretical distinctiveness, drawing on the developing science of cognition as a socially distributed process. Halverson (2002) uses Hutchins (1995) as a source of distributed cognition theory: “a model of the operation of a sociocultural system from which the human actor has been removed” (p. 363, original emphasis), in contrast to an earlier person-centered “model of an individual’s internal cognitive processes” (p. 249). Lakomski and Evers (2016) describe the process of socially distributed cognition as recognizing that individual intentional agency cannot account for how humans function and that “cognition, in this view, ‘leaks’ into the world, therefore is no longer bounded by skin and skull” (p. 9).

If leadership is reconceived as distributed cognition leaking from multiple individuals into the organization, it perpetuates long-term uncertainty over the concept of leadership, in particular over distinguishing leadership activity from other unintentional cognition, as potentially, everything that happens in an organization could be labeled leadership. Clarke calls this “cognitive bloat” (Clark, 2008: 80, 278). Gronn (2009: 200) translates this tendency into the context of education leadership as “a somewhat promiscuous inclination to think of virtually every initiative on the part of teachers and administrators as leadership” (p. 201).

This impulse can be illustrated by [Murphy et al. \(2009\)](#) longitudinal study of six schools in the United States over three years. This presents teachers meeting together to plan instruction more effectively as an example of distributed leadership in action. In Chinese schools, similar collaborative meetings to improve student learning are routine and are known as teacher learning or teacher research, not leadership ([Liu and Hallinger, 2018](#)). In Finland, teachers collaborate at district level to achieve curriculum reform, and this is conceived as shared sense-making, not leadership ([Pietarinen et al., 2019](#)). Collaboration by teachers to improve curriculum and learning can be traced back in the literature over half a century, variously referred to as cooperative teacher preparation, teacher development, networking, cooperative learning, sharing professional knowledge, and many other terms ([Brooks Smith, 1968](#)). Timetabling is another example of an activity undertaken by a teacher that is defined as leadership in [Murphy et al.](#)'s study on the grounds that it involves not just software manipulation but also value judgments ([Gronn, 2003](#)). The difficulty is that all educational activity involves value judgments and much of it entails working with other people. If most activities, even the most mundane, can be defined as leadership, then leadership as a distinct activity disappears.

Leadership also disappears in a second sense. [Woods and Roberts \(2019\)](#) assert that "Leadership is not reducible just to the intentions and actions of individuals" (p. 665). Such a concept of distributed leadership as an emerging, spontaneous, fluid and dispersed activity offers a solipsistic vision in which the embodied human leader disappears behind the disembodied concept of leadership. In using [Hutchins' \(1995\)](#) construction of cognition as a sociocultural system separate from human actors, [Halverson \(2002\)](#) acknowledges as much. A critical perspective insists, to the contrary, that educational organizations' service to learners is embodied, dependent on the socially (un)just intentions and actions of individuals, and that the impact of action relates to the differential power that they hold. Such power may draw on the authority of a particular role but, equally, may relate to knowledge, access to resources or networks, charisma, or a host of other factors. The critical factor in supporting leaders, for practical purposes, is that individual human beings use conscious intention. The guidance of practice engages the leader's intention and, in any attempt to bring about change, the consciousness of power is crucial.

Power and distributed leadership

Power is essentially the capacity to influence or direct others to act in a way that the powerholder wishes ([Lumby, 2019a](#)). This straightforward definition overlies a number of complex and contested theories on how such capacity is acquired and implemented. Direct orders may be used but, equally, indirect strategies may shape the playing field, and even the mindset of the players so that their actions are influenced, circumscribed, or shaped without them necessarily being aware of it ([Lukes, 1974](#)).

Competing ontologies underlie the research on leadership in education and, consequently, our concepts of power and organizations. On one hand are concepts in which faculty can be trusted to use their professional knowledge in the interests of learners. Here, cooperation functions effectively, based on "educational traditions and changes infused by different visions—those that prioritize collaboration, democratic values and humanistic aims" ([Woods and Roberts, 2019: 663](#)) that are embodied in practice; that is, "a shift away from a leadership paradigm based on power and control" ([Halpin, 2003: 85](#)). On the other hand, a much less promulgated vision sees leaders, teachers and organizations quite differently. Research from the 1980s ([Hoyle, 1982](#)) to more recently ([Brosky, 2011](#); [Samier, 2014](#); [Lochmiller and Pawlicki, 2018](#)) has shown that educators are perceived to work for what is best for their learners only in part ([Piot and Kelchtermans, 2016](#)). They also pursue individual and or partisan goals by means of relationships, forming coalitions and engaging in conflict. What is best for learners is contested, and action is arrived at through the temporary resolution of the ongoing competition between these interests. This view presents organizations as places where, regardless of which structural changes are made, hierarchies and coalitions emerge ([Courpasson and Clegg, 2006](#)).

Power in the field of education is generally viewed negatively, as compulsion or manipulation, both of which are assumed to undermine professionalism. Professionalism "gives central stage to and trusts the professional judgment of teacher educators to do what they deem is good, appropriate, or best" ([Vanassche et al., 2019: 485](#)). Faculty's freedom to do as they wish is argued to be essential to the effectiveness of education and the self-identity of educators, justified by their context-specific knowledge and skills. An erosion of professionalism has been strongly linked to analyses of managerialism since the 1990s and embodies resistance to the perceived growing power of senior authority figures. Distributed leadership uses the same argument as professionalism, linking greater faculty autonomy to effectiveness by means of lessening power differentials and opening up leadership.

A vision that rejects oligarchy and offers greater freedom and equity, as befits the professionalism of educators, is an attractive prospect for self-identity and, perhaps because of this, is widely evident in the literature on educational leadership in the last three decades. Micropolitically based research and analyses are less frequent: the view of educators pursuing self-interest as well as learners' interests is far less comfortable. However, many argue this it is neither a negative nor a cynical picture but simply a reflection of the reality of day-to-day life in schools, colleges, and universities ([Bolden, 2004](#); [Brosky, 2011](#)). It is perhaps a more clear-sighted assessment that both at national and organizational level: "Much of the time, education is not about what is best for children; it is about the adult issues of power and control" ([Owen, 2006: 103](#)).

Schools, colleges and universities are unlikely to be staffed by faculty who completely put aside their self-interest and engage in no conflict or micropolitical strategies. Nor are they devoid of those who wish to do the best for learners for much of the time and sincerely pursue what they deem to be a path of integrity. Most faculty are neither angels nor demons but a complex mix of self-awareness and self-deception, and who both collaborate and oppose. The absolutes to be relied upon are that power will be in play and that trust in faculty will inevitably be circumscribed by a realistic vision of human capacity rather than an idealization of professionalism.

Such a position can be traced in the critical literature on professionalism, which recognizes that educators' power is typically invisible. Where it is discernible, power is presented as necessary to the interests of the learner (Osgood, 2006). The classic features of professionalism, which include the acquisition and deployment of a specialized set of skills and vocabulary to provide appropriate ethical service to the client, shape it as an "apolitical and common-sense construct" (p. 5). Many commentators suggest that this "appearance of neutrality" (Bourdieu and Wacquant, 1992: 242) disguises a struggle for control, where the discourse is a tool for the acquisition of power. Its purpose as a means for the professional to gain greater autonomy and control must be hidden, even from the professionals in question, as pursuing power contradicts a self-identity based on devotion to service (Noordegraaf and Schinkel, 2011).

Until the 1990s, the functional literature on professionalism largely served as a means to resist perceived attempts by the state to control educators (Dillabough, 1999). From the 1990s onwards, the concept of distributed leadership has arguably taken over as the major discourse proposing greater autonomy for educators to lead in the interests of the learner (Leithwood et al., 2019). Just as the literature on professionalism insisted that improved service to the learner was the underlying justification for faculty's preference for autonomy, so has distributed leadership emphasized the learner as the primary beneficiary. For example, Leithwood et al. (2019) insist that distributed leadership's positive impact on learner outcomes has been proven empirically and those who question it are mistaken. Consequently, from professionalism to distributed leadership, an unbroken tradition can be discerned in the education leadership literature. Both discourses are used essentially to fight for greater faculty autonomy. In both cases, improving learning is the major justification and the power that shapes which faculty gain greater autonomy or otherwise is largely bleached out, as is faculty's resistance to the power of, or the control by those in authority roles and the state.

Distributed leadership has been frequently critiqued for not translating ideals about power into specific actions, structures, and processes (Hatcher, 2005; Flessa, 2009). For example, Murphy et al. (2009) suggest that "distributed leadership must grow from the soil of changed organizational and professional norms and be actively nurtured by administrative leaders in a variety of ways" (p. 189): these "ways" are not explained. Flessa (2009) points out that, while Spillane and Diamond (2007) instruct us to pay close attention to interaction, not just action, they neither explicitly engage with micropolitical understandings of how human beings interact, nor draw out how we are to interpret interactions and respond in terms of improving learners' experience. Woods and Roberts (2018: 677) provide a list of principles to support the development of collaborative leadership. Their suggestions, such as "a partnership between schools and external agencies" or "opportunities for open discussion", ignore the likely power plays inherent in such partnerships and communications (Griffiths et al., 2009). A detailed analysis of the literature on distributed leadership reveals either a general pursuit of idealized change, such as unspecified transformations of culture or structure, or, when more explicit, actions that for decades have been part of leaders' repertoire and fail to distinguish distributed leadership as different from what went before.

Osgood (2006) argues that communities locate themselves in relation to discourses that "reflect the socially sanctioned dominance of certain ideologies and subjugation of others" (Sinclair, 1996: 232) and that "alternative counter-discourses become pathologized and marginalized" (p. 6). This is the case with distributed leadership, which has largely supplanted other theories of leadership. Following Foucault's insistence that we consider alternative perspectives, the questions then are: If the guidance based on distributed leadership as a concept lacks engagement with the reality of day-to-day life in educational organizations, are any other theories more compelling? What theory or theories best support fallible human beings, working in contexts where ideals and micropolitical interests coexist? It might be a helpful exercise to consider the most pathologized theory of all, bureaucracy.

Bureaucracy

The features of bureaucracy as outlined by Weber (1947) are widely known: the appointment and reward of faculty on merit related to their education, training and experience; labor demarcated by role and expertise; a hierarchy within which each is subject to the instruction of superiors through agreed rules and procedures; and security of tenure. Bureaucracy is "founded on authority, that is, the belief in a legitimate, rational-legal political order" (Olsen, 2005: 2). Both in societal discourse and in education, the term has come to be connected with, at best, negative and restricting structures and processes and, at worst, malign intentions that dehumanize and undermine individual agency (Courpasson and Clegg, 2006; Fitzgerald, 2009). However, its origins relate strongly to values of justice and equality. Just as distributed leadership now claims to be a response to the changing context of society and organizations, and specifically increased complexity, so Weber's (1989a) conception of bureaucracy was in response to "the general growth in those demands on the administration, which result from the increasing complexity of civilization" (p. 348), and a shift from nineteenth-century models of leadership to one that he saw as more rational and so potentially more effective.

As initially conceived, bureaucracy offered an alternative to the untrammelled power of oligarchy, with its historic baggage of nepotism, simony, corruption and other forms of exclusion, such as "race" (Weber, 1989b). The distinctive, rule-based approach that is currently deplored by many was an attempt by Weber (1947) to conceive a method not to inhibit all human initiative but to prevent illegitimate behavior, thus protecting the individual and, by this means, to secure effective organizational functioning (Kallinikos, 2004). Just as many believe that society moves forward by offering all citizens the protection of being subject to the rule of law, so bureaucracy subordinates all members of an organization to the same rules, thereby attempting to circumvent arbitrary or self-seeking action by those with most power.

In its vision of greater equality and more open access to leadership, distributed leadership is based on a positive view of human nature and capacities. Similarly, bureaucracy implies a view of human nature, but it is a less optimistic view. It assumes that human

beings are not always fitted for particular leadership roles, nor will they always put the interests of the organization first. Selection and boundaries are required. Bureaucracy assumes that faculty will want a structure that promotes their career, rules about salary, security of tenure, and access to opportunities that are based on rational decisions: the “‘objective’ execution of business in this context means execution which has ‘no regard for persons’ and which follows calculable rules” (Weber, 1989a: 351).

Weber’s hope was that a bureaucratic organization might go some way toward eliminating the sometimes overt and sometimes subtle or covert means by which some people rise more than others in terms of status, salary and access to challenging work, to the detriment of those whom the organization serves. However, the notion that an organization with classic bureaucratic features would banish inappropriate behavior is as much an ideal type as distributed leadership. Weber (1947) suggested that sociologists need to conceive of pure forms of organization in order to be able to make distinctions and evaluate implications; however, “this ideal *pure* form, is perhaps as little likely to be found in the real world as is a physical reaction, calculated on the assumption of an absolute vacuum” (Weber, 1989b: 23, original emphasis). He reflected on the potential of bureaucratic organizations to support ethical and effective functioning that opened up opportunities to individuals based on training, experience and equality before the rules, yet he was also aware of the possibility that application of the model could tip into a rigid and dysfunctional application of systems. From its inception, the concept of bureaucracy accepted the inevitability that human beings would use organizational forms of leadership both to achieve positive outcomes and to subvert them with negative consequences.

Some critics of distributed leadership have similarly conceived of negative outcomes from distributed leadership in practice (Bolden et al., 2009; Hatcher, 2005), but many more have promoted it on the basis that deliberate encouragement of many into leadership roles and fluid self-selection of leaders represents a liberation that leads to greater organizational effectiveness. The potential for distributed leadership to deepen exclusion, and thereby reduce effectiveness, is less evident in the literature. The formal or informal appointment or the spontaneous emergence of a leader may be predicated not just on an individual’s choice but on the willingness of others to accept him or her as a leader. Consequently, stereotypical prototypes of what a leader looks like serve to shape the willingness to support. These may relate as much to unjustified prejudice as to the individual’s expertise. The flows into and out of leadership according to organizational need, which is suggested by some to be an identifier of distributed leadership, depict an organization as apolitical, free from sexism, racism, and other forms of prejudice. Simultaneously, numerous studies of distributed leadership in action repeatedly show both control over entry into leadership by those in senior positions and the exclusion of some individuals for reasons other than capability (Hairon et al., 2014; Murphy et al., 2009).

Qualifiers such as “bounded empowerment” (Bouwman et al., 2019: 568) or “pragmatic schools” (Hairon et al., 2014: 370) have emerged to limit the earlier optimistic explorations of distributed leadership. The latter study suggests that distributed leadership is useful in Singapore schools, while also making it clear that faculty are uncomfortable with the notion and do not wish to expend time on leadership which they see as the province of those appointed to such roles: “they do not wish to spend time for the self-organization work needed” (p. 379). This quotation raises the issue of the cultural inappropriateness of distributed leadership in parts of the world that do not prize a reduction in power distance to the same extent that is evident in the West, where the concept of distributed leadership evolved.

Sustaining the weakness of the field

The research on distributed leadership perpetuates several weaknesses in the field of education leadership: poor historical awareness; side-lining of inconvenient alternative perspectives; unacknowledged methodological limitations; cultural colonialism through global promotion of a concept based on Western values; and marginalization of the issue of exclusion.

Poor historical awareness

An example of poor historical awareness is the excoriation of bureaucracy and the presentation of distributed leadership as overturning what are characterized as the nineteenth-century values and structures that inhibit leadership by the many (Murphy et al., 2009; Day et al., 2010). Such assertions are based on “oversimplified and stylized images of the bureaucratic form ... marked by an astonishingly naïve functionalism devoid of any historical awareness” (Kallinikos, 2004: 2). The initial purpose of bureaucratic forms, to create more egalitarian, just, and effective organizations, is lost and replaced by a caricature that presents only its potential disadvantages, which Weber (1947) fully anticipated and acknowledged, and none of its advantages.

Theoretical evangelism

Much of distributed leadership research’s blindness to alternative perspectives can be illustrated by reference to bureaucracy. Some assert that distributed leadership is evident in all educational organizations, given that spontaneous leadership will emerge in a group of people whatever its intentions, structure, or culture (Gronn, 2016). Woods and Roberts (2013) insist that “leadership is a distributed phenomenon”; however, there is at least as much evidence that education embodies bureaucracy, given its widely evident framework of rules, standards and processes, sometimes state-mandated, matching the features of bureaucracy. The persistence of evidence showing that in education the dominant form is a formal hierarchy (Earley et al., 2012) is played down by assertions that, within a distributed leadership system, formal leaders are still necessary and that hierarchy runs alongside heterarchy (Zala-Mezö et al., 2019). What a hierarchy brings positively to leadership is either underexplored or categorized as

a component of distributed leadership. A key characteristic of bureaucracy, the allocation of roles to those with relevant expertise, is explained exclusively as a facet of distributed leadership: for example, “patterns of leadership distribution tend to be based on patterns of expertise” in “authentic distributed leadership” (Leithwood et al., 2019: 10) rather than being analyzed as a defining characteristic of bureaucracy. There is a tendency to characterize factors as illustrating or providing evidence of the presence of distributed leadership even though they could just as well be evidence of alternative forms of leadership using other theoretical frameworks.

Compare research in educational leadership with that in generic leadership and a different breadth becomes apparent. Bureaucracy rarely appears in educational leadership research except as something to be avoided, replaced, or deplored (Lumby, 2019b). By contrast, generic leadership research has engaged with challenging and developing bureaucracy, evolving new concepts such as representative bureaucracy (Kemp et al., 2019) and hybrid adaptive bureaucracy (Zhi and Pearson, 2017), just two illustrative examples of many attempts to evolve hybrid and alternative models in public services (Fotaki and Jingit, 2018). The narrow focus on distributed leadership in educational leadership research has deprived the field of necessary knowledge and insights into the bureaucratic form that is, to some degree, evident in all educational organizations.

Insufficient interrogation of research methodology

The primary justification for the introduction of distributed leadership as a practice is that it will improve learning. The literature on distributed leadership asserts its connection to improved student outcomes. For example, Harris (2008: 3) claims that “Distributed leadership is a common denominator of highly effective organizations across different sectors”. Many researchers make similar claims (Timperley, 2008; Day et al., 2010). Research has focused increasingly on identifying the connection definitively. For instance, Heck and Hallinger (2010) conducted a large-scale, longitudinal study and concluded that “the evidence therefore suggests that change in distributed leadership can be empirically linked to change in school improvement capacity and subsequent growth in student learning” (p. 881). Such claims continue (Supovitz et al., 2019).

The extent to which relevant research is considered trustworthy determines whether a connection is accepted between distributed leadership and improved student outcomes. Multiple issues have been raised. First is the widely expressed difficulty in reaching a clear definition of distributed leadership that is adequate for translation into distinctive and reliable research factors across differing contexts. However, increasingly, commentators insist that distributed leadership takes various forms (Bolden et al., 2009; Robinson, 2009; Leithwood et al., 2019), leading Robinson to suggest that the appropriate question is not whether there is a connection between distributed leadership and improved student outcomes but “What are the relative impacts of particular types of distributed leadership practice?” (p. 238). The evidence about any such connection would need to be sorted by its relation to the particular form of distributed leadership, were there a sufficiently clear definition of each and a way to identify which was under scrutiny. The evidence’s relationship to the culture of the research location would also need to be assessed. A form may appear positive in, say, Hong Kong but not South Africa. Only then could the data be interrogated to see if the evidence is sufficiently robust. Often these conditions have not been met, and some researchers have acknowledged weaknesses in the research (Heck and Hallinger, 2010) or suggested that, while the connection is unproven it remains “intuitively attractive” (Hairon and Goh, 2015: 694).

A further problem is that if the argument is accepted that leadership is always distributed (Woods and Roberts, 2013), then a counterfactual is impossible: it is not feasible to eliminate distributed leadership from a control group school(s) or parts of a school in order to assess its impact. Despite assertions that those who question the evidence of improved student outcomes have either consulted insufficient literature or misunderstood it (Harris and DeFlaminis, 2016; Leithwood et al., 2019), many research issues remain unresolved: inadequate definition (Lakowski, 2008); reliance on self-reported data, whether through interview or survey; insufficient attention to the impact of the researcher’s presence, understood by some as the “Hawthorne effect” (Stand, 2000); the difficulty in establishing a control group and researcher bias, given the extent of investment by both national policy and research systems in this form of leadership (Lumby, 2019b). All conceptual frameworks bias the results to some degree, but it would be useful if those researching distributed leadership could interrogate the weaknesses in this research more fully.

Colonization by research theories

Such interrogation is less likely in a study whose researchers and the faculty at the research site are proponents of distributed leadership, leading to potential bias in the study’s self-reported data and analysis. In particular, the cultural fit of distributed leadership appears at times to be set aside to demonstrate its presence, however it is defined by a study, even when its data suggest that this is doubtful. For example, Hairon et al.’s. (2014) data analysis shows leaders’ “reluctance to relinquish key operational decisions to subordinates or only encourage faculty to make decisions within their work scope”, relating this to Singapore’s culture. Yet no question is raised about whether the data indicate the presence of distributed leadership or, instead, another kind of leadership. In many parts of the world, propelled by the support of many researchers and the OECD (2008), in the field of education leadership the square pegs of distributed leadership are hammered into cultural round holes. This perpetuates the tradition of colonization by theory; that is, the eager export and adoption of theories originating in Anglophone countries by those whose culture and values may be quite different (Normore and Collard, 2007).

Blindness to exclusion

Education leadership research is generally bifurcated. Much includes, at best, a token reference to the forms of prejudice embedded in education, such as sexism and racism. Much less is critical research giving centrality to issues of inequity and exclusion among both faculty and learners. Distributed leadership generally falls into the first category.

One justification of distributed leadership is that it opens up leadership to more people; however, if opportunities were indeed being opened up, one would expect to see more people with stigmatized characteristics using informal opportunities as a lever to access formal paid roles or, at least, generally experiencing less exclusion. The literature from many parts of the world supports no such change. Those groups of people who are underrepresented in leadership continue to be so, and in some cases their experience in the workplace blights their professional life (Bush, 2019; Ho, 2015; deLeon and Cryss Brunner, 2012). Distributed leadership appears to have raised expectations for greater leadership opportunities that are not always fulfilled. For example, Holloway et al. (2018) examined one form of distributed leadership in the United States and found that teachers' "expectations were unrealistic" (p. 538). They concluded "that prescribed, incentive-driven forms of distributed leadership can place teacher leaders in precarious positions that demand more of their time, while limiting their capacities to participate in the leadership practices they deem most valuable" (Holloway et al., 2018). Far from improving their opportunities, distributed leadership had, in some cases, been a mechanism to extract more work from teachers. Although some critical literature attempts to point out that people's participation in the putative greater opportunities persists in being shaped by power structures (Lumby, 2013) or social and professional capitals (Woods, 2016), a great many articles set this understanding aside.

A second justification for distributed leadership is that it offers a lens through which to understand the totality of leadership. If this promise were met, it would be a valuable step forward. One might expect to see a range of research encompassing political and critical as well as functional perspectives, exploring in detail how people access leadership or not, their experiences while leading, and how these impact on learners. Instead, there are characteristically surveys, where teachers self-report using a range of factors, some of which could be elements from other forms of leadership, correlated to learner outcomes. Distributed leadership does not appear to have contributed to a fuller picture of leadership to lead us to understand the complex mesh of interactions that enable some, but not others, to take on leadership roles.

The future of leadership theory

Just as there are debates on whether schools reflect or lead society, so one might question whether theories of leadership improve the experience and outcomes of faculty and learners' or contribute to the mesh of phenomena that construct the current inequalities and failings. The very least one might expect of research in this field is self-interrogation, a critical awareness of the potentialities of any form of leadership to resist or sustain unequal systems; however, rather than rigorous self-interrogation, distributed leadership has proved a ready tool for smoothing over the complexities of leadership. Spillane and Diamond (2007) point out the irony whereby the very vagueness of its meaning contributes to distributed leadership's popularity. It has proved a useful label to attach to various leadership activities, carrying with it its positive spin of admirable intentions. Despite claims by some that it widens leadership opportunities, there is evidence that it extracts more work, so is part of the work-intensification project of globalization, and that it sidesteps any explicit engagement in creating opportunities for those groups most likely to be excluded. In an era of populism, the advent of highly visible leaders who flout the traditional public-sector values of honesty, transparency and inclusion demonstrates the dangers of leadership that has insufficient safeguards, balances, and boundaries. The field of education leadership has focused on distributed leadership and set aside the development of forms of leadership that may offer greater protections, such as bureaucracy.

The problem then, is not distributed leadership itself, whose inception was intended to interrogate the complexity of leadership more fully, but rather how it has become a mechanism for the mainstream field of educational leadership to persist in largely ignoring the political and discriminatory aspects of education organizations. This is not to argue that its adoption, in its very many configurations, has not had positive impacts: the Hawthorne effect alone would guarantee this. Nor is it to argue that other forms of leadership will not have their disadvantages as well as advantages. Rather, a critical perspective would demand a more rigorous exploration into distributed leadership's potential negatives by mainstream research instead of banishment to its margins. A deeper engagement with concepts of power and organization is needed, recognizing that, as Ahlberg and Nightingale (2018) argue, "acts of resistance to subjection ... simultaneously confirm and reiterate social hierarchies and discriminatory norms" (p. 385).

Distributed leadership as practice has sought, to use Judith Butler's (1990) term, to "trouble" the prevailing relations in the hierarchies of education organizations without reference to the long-acknowledged paradox whereby resistance both diminishes and heightens existing power relations. Often this is achieved by using three-dimensional power where a field of action serves to manipulate individuals to internalize beliefs that perpetuate the power differentials while they imagine that they are doing the opposite (Foucault, 1974). Distributed leadership has operated in this way. Alongside its apparent role in opening up opportunities and redressing power imbalance, distributed leadership has been a vehicle to perpetuate weaknesses in the field of education leadership research and practice, and so it has served leaders less well than is claimed by many.

References

- Ahlberg, H., Nightingale, A.J., 2018. Theorizing power in political ecology: the where of power in resource governance projects. *J. Polit. Ecol.* 25 (1), 382–401.
- Bolden, R., Petrov, G., Gosling, J., 2009. Distributed leadership in higher education: rhetoric and reality. *Educ. Manag. Adm. Leader* 37 (2), 257–277.
- Bolden, R., 2004. What is Leadership? Leadership South West Research Report 1. Exeter: Leadership South West. Exeter University.
- Bourdieu, P., Wacquant, L., 1992. *An Invitation to Reflexive Sociology*. Polity, Cambridge.
- Bouwman, M., Runhaar, P., Wesselink, R., Mulder, M., 2019. Towards distributed leadership in vocational education and training schools: the interplay between formal leaders and team members. *Educ. Manag. Adm. Leader* 47 (4), 555–571.
- Brooks Smith, E., 1968. Needed: A New Order in Student Teaching that Brings Joint Accountability for Professional Development. Available at: <https://files.eric.ed.gov/fulltext/ED023624.pdf>.
- Brosky, D., 2011. Micropolitics in the school: teacher leaders' use of political skill and influence tactics. *Int. J. Educ. Leader. Preparation* 6 (1), 1–11.
- Bush, T., 2019. Talent, determination and resilience: leadership accession and enactment for black, Asian and minority ethnic leaders. In: Miller, P., Callender, C. (Eds.), *Race, Education and Educational Leadership in England: An Integrated Analysis*. Bloomsbury Academic, London, pp. 81–98.
- Butler, J., 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, Abingdon, Oxon.
- Clark, A., 2008. *Supersizing the Mind*. Oxford University Press, New York.
- Courpasson, D., Clegg, S., 2006. Dissolving the iron cages? Tocqueville, Michels, bureaucracy and the perpetuation of elite power. *Organization* 13 (3), 319–343.
- Day, C., Sammons, P., Hopkins, D., et al., 2010. Ten Strong Claims about Leadership. NCSL, Nottingham. Available at: <https://dera.ioe.ac.uk>.
- deLeon, M.J., Cryss Brunner, C., 2012. Cycles of fear: a model of lesbian and gay educational leaders' lived experiences. *Educ. Adm. Q.* 49 (1), 161–203.
- Dillabough, J.A., 1999. Gender politics and conceptions of the modern teacher: women, identity and professionalism. *Br. J. Sociol. Educ.* 20 (3), 373–394.
- Earley, P., Higham, R., Allen, R., et al., 2012. Review of the School Leadership Landscape. NCSL, Nottingham.
- Fitzgerald, T., 2009. The tyranny of bureaucracy: continuing challenges of leading and managing from the middle. *Educ. Manag. Adm. Leader* 37 (1), 51–65.
- Flessa, J., 2009. Educational micropolitics and distributed leadership. *Peabody J. Educ.* 84 (3), 331–349.
- Fotaki, M., Jingjit, R., 2018. Humanizing bureaucracy: clan-oriented culture in the Thai civil service. In: Bice, S., Poole, A., Sullivan, H. (Eds.), *Public Policy in the "Asian Century"*, International Series on Public Policy. Palgrave Macmillan, London, pp. 151–183.
- Foucault, M., 1974. Human nature: justice versus power. In: Fols, E. (Ed.), *Reflexive Water: The Basic Concerns of Mankind*. Souvenir Press, London, pp. 135–197.
- Griffiths, J., Vidovich, L., Chapman, A., 2009. Policy "partnerships"? Power dynamics in curriculum reform. *J. Educ. Adm. Hist.* 41 (2), 193–208.
- Gronn, P., 2003. Leadership: who needs it? *Sch. Leader. Manag.* 23 (3), 267–291.
- Gronn, P., 2009. From distributed to hybrid leadership practice. In: Harris, A. (Ed.), *Distributed Leadership*. Springer, Dordrecht, pp. 197–217.
- Gronn, P., 2016. Fit for purpose no more? *Manag. Educ.* 30 (4), 168–172.
- Gumus, S., Bellibas, M.S., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48.
- Hairon, S., Goh, J.W., 2015. Pursuing the elusive construct of distributed leadership: is the search over? *Educ. Manag. Adm. Leader* 43 (5), 693–718.
- Hairon, S., Goh, J.W.P., Lin, T.B., 2014. Distributed leadership to support PLCs in Asian pragmatic Singapore schools. *Int. J. Leader. Educ.* 17 (3), 370–386.
- Halpin, D., 2003. *Hope and Education. The Role of the Utopian Imagination*. Routledge Falmer, London.
- Halverson, C.A., 2002. Activity theory and distributed cognition: or what does CSCW need to DO with theories? *Comput. Support. Coop. Work* 11 (1–2), 243–267.
- Harris, A., DeFlaminis, J., 2016. Distributed leadership in practice: evidence, misconceptions and possibilities. *Manag. Educ.* 30 (4), 141–146.
- Harris, A., 2008. Distributed leadership: according to the evidence. *J. Educ. Adm.* 46 (2), 172–188.
- Hatcher, R., 2005. The distribution of leadership and power in schools. *Br. J. Sociol. Educ.* 26 (2), 253–267.
- Heck, R.H., Hallinger, P., 2010. Testing a longitudinal model of distributed leadership effects on school improvement. *Leader. Q.* 21, 867–885.
- Ho, E.S.C., 2015. Confronting the barriers of women leaders around the world. In: Reilly, E.C., Bauer, Q. (Eds.), *Women Leading Education across the Continents: Overcoming the Barriers*. Rowman & Littlefield, Lanham, ON, p. 85.
- Holloway, J., Nielsen, A., Saltmarsh, S., 2018. Prescribed distributed leadership in the era of accountability: the experiences of mentor teachers. *Educ. Manag. Adm. Leader* 46 (4), 538–555.
- Hoyle, E., 1982. Micropolitics of educational organizations. *Educ. Manag. Adm.* 10 (2), 87–98.
- Hutchins, E.L., 1995. *Cognition in the Wild*. MIT Press, Cambridge, MA.
- Kallinikos, J., 2004. The social foundations of the bureaucratic order. *Organization* 11 (1), 13–36.
- Kemp, L.J., Mathias, M., Raji, M., 2019. Representative bureaucracy in the Arab Gulf states. *Int. J. Public Sect. Manag.* 32 (5), 392–412.
- Lakomski, G., Evers, C.W., 2016. Challenging leadership: the issues. In: Lakomski, G., Evers, C.W. (Eds.), *Questioning Leadership*. Routledge, Abingdon, Oxon, pp. 17–30.
- Lakomski, G., 2008. Functionally adequate but causally idle: w(h)ither distributed leadership? *J. Educ. Adm.* 46 (2), 159–171.
- Leithwood, K., Harris, A., Hopkins, D., 2019. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* <https://doi.org/10.1080/13632434.2019.1596077>.
- Liu, S., Hallinger, P., 2018. Principal instructional leadership, teacher self-efficacy, and teacher professional learning in China: testing a mediated-effects model. *Educ. Adm. Q.* 54 (4), 501–528.
- Lochmiller, C.R., Pawlicki, C.R., 2018. Leadership and resource allocation in schools: applying micropolitical perspectives. In: Touchton, D.J., Rodriguez, M.A., Ivory, G., Acker-Hocevar, M. (Eds.), *Quandaries of School Leadership*. Springer, Cham, pp. 181–206.
- Lukes, S., 1974. *Power: A Radical View*. Macmillan, London.
- Lumby, J., 2013. Distributed leadership: the uses and abuses of power. *Educ. Manag. Adm. Leader* 41 (5), 581–597.
- Lumby, J., 2019a. Leadership and power in higher education. *Stud. High Educ.* 44 (9), 1619–1629. <https://doi.org/10.1080/03075079.2018.1458221>.
- Lumby, J., 2019b. Distributed leadership and bureaucracy. *Educ. Manag. Adm. Leader* 47 (1), 5–19.
- Murphy, J., Smylie, M., Mayrowetz, D., Seashore Louis, K., 2009. The role of the principal in fostering the development of distributed leadership. *Sch. Leader. Manag.* 29 (2), 181–214.
- Noordegraaf, M., Schinkel, W., 2011. Professionalism as symbolic capital: materials for a Bourdieusian theory of professionalism. *Comp. Sociol.* 10 (1), 67–96.
- Normore, A., Collard, J., 2007. Constructing theory for leadership in intercultural contexts. *J. Educ. Adm.* 45 (6), 740–755.
- Organisation for Economic Cooperation and Development (OECD), 2008. *Improving School Leadership*. OECD, Paris.
- Olsen, J.P., 2005. Maybe it is time to rediscover bureaucracy. *J. Publ. Adm. Res. Theor.* 16 (1), 1–24.
- Osgood, J., 2006. Deconstructing professionalism in early childhood education: resisting the regulatory gaze. *Contemp. Issues Early Child.* 7 (1), 5–14.
- Owen, J., 2006. *The Impact of Politics in Local Education: Navigating White Water*. Rowman & Littlefield Education, Lanham, MD.
- Pietarinen, J., Pyhältö, K., Soini, T., 2019. Shared sense-making in curriculum reform: orchestrating the local curriculum work. *Scand. J. Educ. Res.* 63 (4), 491–505.
- Piot, L., Kelchtermans, G., 2016. The micropolitics of distributed leadership: four case studies of school federations. *Educ. Manag. Adm. Leader* 44 (4), 632–649.
- Robinson, V.M.J., 2009. Fit for purpose: an educationally relevant account of distributed leadership. In: Harris, A. (Ed.), *Distributed Leadership*. Springer, Dordrecht, pp. 219–240.
- Samier, E., 2014. *Secrecy and Tradecraft in Educational Administration: The Covert Side of Educational Life*. Routledge, Abingdon, Oxon.
- Sinclair, A., 1996. Leadership in administration: rediscovering a lost discourse. In: Weller, P., Davis, G. (Eds.), *New Ideas, Better Government*. Allen & Unwin, Sydney, pp. 229–244.
- Spillane, J.P., Diamond, J. (Eds.), 2007. *Distributed Leadership in Practice*. Teachers College Press, New York.

- Stand, J., 2000. The "Hawthorne effect"—what did the original Hawthorne studies actually show? *Scand. J. Work. Environ. Health* 26 (4), 363–367.
- Supovitz, J.A., D'Auria, J., Spillane, J.P., 2019. Meaningful and Sustainable School Improvement with Distributed Leadership. CPRE Research Reports. Available at: https://repository.upenn.edu/cpre_researchreports/112.
- Timperley, H., 2008. A distributed perspective on leadership and enhancing valued outcomes for students. *J. Curric. Stud.* 40 (6), 821–833.
- Vanassche, E., Kidd, W., Murray, J., 2019. Articulating, reclaiming and celebrating the professionalism of teacher educators in England. *Eur. J. Teach. Educ.* 42 (4), 478–491. <https://doi.org/10.1080/02619768.2019.1628211>.
- Weber, M., 1947. Legitimate authority and bureaucracy. In: Henderson, A.M., Parsons, T. (Eds.), *Theory of Social and Economic Organization*. Oxford University Press, New York, pp. 328–340.
- Weber, M., 1989a. Race relations. In: Runciman, W.G. (Ed.), *Weber: Selections in Translation*. Cambridge University Press, London trans. Matthews, E.
- Weber, M., 1989b. The development of bureaucracy. In: Runciman, W.G. (Ed.), *Weber: Selections in Translation*. Cambridge University Press, London trans. Matthews, E.
- Woods, P.A., Roberts, A., 2013. Leadership is Distributed. YouTube Video. Available at: <https://youtu.be/oQl8H7XN4l4>.
- Woods, P., Roberts, A., 2018. *Collaborative School Leadership: A Critical Guide*. Sage, London.
- Woods, P., Roberts, A., 2019. Collaborative school leadership in a global society: a critical perspective. *Educ. Manag. Adm. Leader* 47 (5), 663–677.
- Woods, P.A., 2016. Authority, power and distributed leadership. *Manag. Educ.* 30 (4), 155–160.
- Zala-Mezö, E., Bormann, I., Strauss, N., Müller-Kuhn, D., 2019. Distributed leadership practice in Swiss "eco-schools" and its influence on school improvement. *Leader. Pol. Sch.* <https://doi.org/10.1080/15700763.2019.1631855>.
- Zhi, Q., Pearson, M.M., 2017. China's hybrid adaptive bureaucracy: the case of the 863 program for science and technology. *Governance* 30 (3), 407–424.

Leadership in large-scale international assessments—diminishing returns?

Daniel Muijs, School of Social Sciences, Education and Social Work, Queen's University Belfast, Belfast, Northern Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	109
Defining leadership	109
What does leadership look like?	110
Leadership and attainment	111
Conclusion	112
References	113

Introduction

Leadership has long been acknowledged as a key factor in school effectiveness, with leadership being only second to teaching in its impact on pupil outcomes. Many studies have shown leadership to have a significant albeit modest and largely indirect effect on attainment. School leadership typically account for approximately a fifth to a quarter of school level variance in pupil attainment, once intake and pupil background have been factored in the indirect effect of leadership is even greater, as it influences school climate, disciplinary climate, teaching quality, curriculum, relations and therefore almost everything that happens in the school as an organization. Furthermore, the impact of school leadership is greatest in schools serving the most disadvantaged (Leithwood et al., 2004, 2006; Huber and Muijs, 2010; Robinson and Gray, 2019). It is therefore not surprising to find that large-scale international studies have also looked at this factor as an explanatory factor in pupil attainment (as in PISA), and as a factor of importance in and of itself (as in TALIS).

The broad conclusions on the importance and impact of leadership drawn in the literature on leadership and school effectiveness do however mask some differences in how leadership has been conceptualized. Over time, a consensus view has emerged, though, as well will see below, this is now being contested.

In this chapter we will therefore look at how leadership is conceptualized in LISA's, such as TALIS, TIMMSS/PIRLS and PISA, what the key findings from these studies are, and how the current conceptualization and use of leadership in the leadership corresponds with the newest insights on leadership in education.

In this review I will disregard studies that have simply used international study rankings to choose countries high (or low) and then separately study these on leadership and other factors, due to the methodological weaknesses of such designs in that they do not take sufficient account of the variety of cultural and other differences between countries, and therefore tend to more resemble a fishing expedition than scientific research.

Defining leadership

While long considered a key factor in organizational performance, leadership can be a somewhat nebulous and ill-defined construct. For some, leadership is essentially defined by organizational role, thus in an educational context equating leadership with the person(s) of the leader, typically the principal and the senior leadership team in the school. Leadership is then essentially what the leader(s) does (do). A second research tradition focuses more on the characteristics of the person in the role. Here we see an evolution from charismatic views of leadership, whereby innate traits are studied, to an emphasis on behavior and beliefs of leaders, leading to transformational theories of leadership. In education, these have often been supplemented or supplanted by instructional leadership, about which we will write more below. However, prevailing views tend to see leadership as much as a characteristic of the organization as of individuals, with distributed leadership models having become popular in education, not least in the form of teacher leadership. Increasingly, as the role of schools as part of a wider social and educational system have increasingly been acknowledged, system leadership has become to be seen as part of the school leader job description (Muijs, 2011).

Recently, there has been growing critique of these prevailing models of educational leadership, which are seen as too generic, and insufficiently focused on knowledge. In this alternative perspective, contextual knowledge about education, curriculum, didactics and pedagogy should form the basis for effective school leadership (Lock and Rees, 2020). This view has started to influence leadership development in a number of places, notably England, where it has helped inform the new national professional leadership qualifications (Jerrim and Sims, 2019).

This new perspective sees school leadership as:

- Complex. This recognizes the complexity of the role, and therefore that much of the work of leaders is solving complex problems while realizing their own limitations as they engage in this work.
- Domain-specific. Leadership is much more about specific knowledge in a particular field than generic typologies such as distributed or transformational leadership. It therefore requires in-depth knowledge of education in areas such as pedagogy and curriculum

- Knowledge-centric. Rather than a set of more-or-less generic skills, leadership is built on in-depth knowledge
- Problem-focused: key to leadership is the ability to deal with persistent (complex) problems
- Context-specific. Leadership start with understanding the specific school context.

In light of these differences, it is important to understand how leadership is defined in the major international studies.

TALIS, which looks at leadership most comprehensively of the international studies, provides a short theoretical summary in their conceptual framework for the 2018 survey of teachers and school leaders. While no overall definition of leadership is given, the framework draws on Day et al. (2016) summary of characteristics of effective school leadership. TALIS puts a lot of emphasis on instructional leadership, defined as “supporting and developing high-quality instructional practices, developing and implementing policies that support student achievement, developing learning communities, providing feedback on instruction, modeling effective instruction, and supporting the use of assessment data” (Ainley and Carstens, 2018). TALIS also points to a growth in the importance and attention paid to system leadership in the field since the previous survey. The third pillar of the framework is distributed leadership, again seen as an area that has received increasing prominence in the period before development of TALIS 2018. In the study itself TALIS looks at five dimensions through its questionnaires:

1. Characteristics of school leaders, e.g. qualifications, recruitment;
2. What school leaders do: role, function and work of the principal, instructional leadership, terms and conditions, workload, hours, autonomy, function, and actions;
3. Distributed leadership and teacher leadership;
4. Principal job satisfaction; and
5. System leadership and leadership in networks of schools.

PISA included items on school management/leadership in the school context survey up to 2015. In the subsequent two administrations these items were no longer included. The items in the earlier editions reflect instructional and distributed leadership models, as well as an emphasis on governance models (who at which level from national government to teacher is responsible for what) (OECD, 2017, 2019).

The latest iterations of TIMSS/PIRRLS focus primarily on principal qualifications and experience, though previously a “leadership activities scale” was included in the school questionnaire. In this scale principals were asked how much time they spent on a range of leadership activities, reflecting one of four aspects of school leadership: (1) school vision, (2) safe and orderly environment, (3) faculty trust, and (4) professional development. Principals were asked to respond using a rating scale with three categories (1 = No time, 2 = Some time, 3 = A lot of time) IEA (Mullins et al., 2012; Reynolds et al., 2023).

The LISA frameworks therefore reflect a view of leadership both as a role, with a clear focus on school principals, their background, their characteristics and what they do, supplemented by some attention (especially in TALIS and earlier iterations of PISA up to 2015) on instructional distributed leadership. These frameworks tend toward the traditional definitions of leadership around key concepts such as distributed or instructional leadership, rather than toward the new thinking that stresses knowledge, complexity and specificity, and primarily focus on the principal on the one hand and teacher input on the other, with little attention to intermediary levels of leadership in schools (such as middle leaders).

It is notable that the attention paid to leadership has diminished in recent versions of both PISA and TIMSS/PIRRLS. We will discuss more later in this chapter.

What does leadership look like?

The main evidence on what school leadership looks like internationally comes from the TALIS studies.

A range of analyses of the TALIS 2013 and 2018 datasets have been undertaken, looking at different aspects of leadership, such as instructional and distributed leadership behaviors, based on either ratings from the principal scales or from the teacher survey (e.g. teacher rating of principal instructional leadership and distributed leadership, Xia and O’Shea, 2022).

TALIS generally reports quite high levels of instructional leadership across countries, while the picture on distributed leadership is more mixed.

In terms of instructional leadership, almost 60% of principals across TALIS (2018) participating countries claim to be using direct forms of instructional leadership, such as collaborating with teachers on school discipline, leading professional development, observing teachers and providing feedback. An even greater proportion of principals (typically over two-thirds of respondents) were involved in activities related to indirect instructional leadership, such as actions to ensure that teachers feel responsible for their students’ learning outcomes, actions to ensure that teachers take responsibility for improving their teaching skills, and actions to support co-operation among teachers to develop new teaching practices (OECD, 2021b).

Looking then at distributed leadership we find that involvement of teachers in school management and decision making is variable, and quite limited in many schools and countries (OECD, 2021b). Overall, according to responses to the principal survey around 42% of schools have high levels of teacher participation in decision making. There are, however, clear differences between questions asked (e.g. when principals are asked whether teachers are involved in decision-making almost all say they are, when asked more specifically about teacher membership of the senior management team the proportion in some countries is as low as 25%). There are also large differences between teachers’ and principals’ responses to questions on their perceived influence and control, with the former, unsurprisingly, reporting much lower levels than the latter.

Principals don't collaborate with other principals and school leaders all that often (just over a third), suggesting that the system leadership elements of the role are less developed (OECD, 2021b).

These results from TALIS factors remain largely unchanged compared to TALIS 2013.

Of course, a key question is the extent to which the constructs used in TALIS are in fact invariant across countries. This is not entirely clear using the Rasch model to analyze the leadership activities scale from PIRLS and TIMSS, Zhang and Wind (2019) compared Hong Kong SAR, Chinese Taipei, the United States, and Canada, finding that perceived ordering of activities across countries differed to some extent. In particular, the two East-Asian countries differed significantly from the two North American countries on perceptions in two main areas. The main differences were on items relating to regulating student behavior, on which principals in the North American countries spent significantly more time, and items on visiting other schools and conferences, which was more common among principals in the East Asian countries. Differences in most of the other items were limited, however. Differences in factor structure between Germany and Chinese Taipei were found in a study comparing TIMSS (2011) data. In Germany, the researcher extracted four leadership styles using CFA and MLSEM: distributed leadership style (26%), the integrated instructional and transformational leadership style (27%), the transformational leadership style (7%), and the instructional leadership style (40%). In Chinese Taipei only two styles were extracted: the integrated instructional and transformational leadership style (36%), and the mixed integrated instructional and distributed leadership style (64%) (Chen, 2018). However, using the same data the research team found that the basic dimensions of vision/goal (S1), management (S2), maintaining school climate (S3), and principal leadership did fit the data in both countries (Chen et al., 2020). Looking at PISA 2015 data, Gao et al. (2018) claim factorial invariance for the construct of principal decision making across China and the US. However, this was only achieved by the statistically dubious practice of removing the constraints with the largest modification indices.

Not only are there cross-country differences in how leadership is conceptualized and understood, but there is also the methodological issue of differences in self-report measures from principals and teachers. This is evident in TALIS, for example, where principals' and teachers' views on distributed leadership in their schools differ substantially (with principals generally perceiving greater levels of DL than teachers) (Printy and Liu, 2021), and where teachers see instructional leadership as a unidimensional construct, whereas it is multi-faceted among principals (setting goals, professional development, and supervision of instruction) (Urick and Bowers, 2019).

Country differences are not always apparent, however. Analyzing TALIS 2018 data, Veletic and Olsen (2021) used cluster analysis to identify five clusters of practice relating to leadership for learning. They found that these clusters differ substantially more between schools within countries than between countries, and that country differences are unrelated to geographical/cultural groupings of countries.

Leadership and attainment

As mentioned in the introduction, the consensus among researchers into educational leadership is that there is a significant, albeit mainly indirect, impact of leadership on pupil attainment. The LISA studies paint a far less clear picture of this relationship, however.

Analysis of the 2018 TALIS-PISA link dataset does indeed show an indirect effect of principal leadership on pupil attainment. The findings also point to a number of key factors leaders need to focus on to enhance pupil attainment, especially ensuring a disciplined school climate focused on achievement, and supporting teachers by maximizing the time they can spend on teaching and learning and minimizing unnecessary workload and stress (OECD, 2021a,b,c).

An analysis of TALIS data for the English education ministry showed significant differences between countries in requirements for training and development of principal, salary, teaching requirements and available CPD. There was however, little evidence of a relationship between these factors and student outcomes. There was some limited evidence that principals in high performing countries, with principals in selected high performing countries valuing instructional and distributed leadership, but the methodology (selecting countries) makes such conclusions highly tentative (Greatbatch and Tate, 2018).

As well as the 2018 PISA-TALIS link dataset, there are also a range of analyses from PISA 2015 (in which the school questionnaire contained a number of scales and items relating to principal leadership). One study, for example, looked at four dimensions of leadership from the school questionnaire and correlated these to outcomes in science, maths and English. They found a positive relationship between instructional leadership and outcomes, while support for professional development was negatively related to outcomes. Goal setting and distributed leadership were not significant, and the overall relationship with leadership was negative (Wu et al., 2019). On the other hand, in a study using data from the teacher questionnaire, Wu et al. (2020) did find a significant relationship between principal leadership and attainment in science in the US 2015 PISA database. Using the German data from PISA 2015, a positive relationship was found between instructional leadership and exchange of materials between teachers, but this in turn was negatively related to attainment (Mora-Ruano et al., 2021). Using TIMSS 2011 data comparing Germany and Chinese Taipei, Chen et al. (2020) found that principal leadership in general had a negative relationship with students' mathematics achievement in Germany, but had a non-significant relationship in Chinese Taipei. The researchers also looked at the relationship between each category of principal leadership and students' mathematics achievement in both countries. The results showed a stronger focus on setting vision/goal and on maintaining school climate was negatively related to students' mathematics achievement in Germany, while a stronger focus on school management was positively related to achievement. In Chinese Taipei

maintaining school climate was positively associated with achievement, and school management negatively. Setting vision/goal was unrelated to students' mathematics achievement.

Using PISA 2009 data for a comparative study of Korea, Mexico and the US, [Shin et al. \(2013\)](#) also found very mixed results. Principals' perceptions of their instructional leadership were positively related to PISA scores in Korea but negatively in Mexico and the United States. Similarly, principals' perceptions of their leadership of teachers showed a positive relationship with PISA scores in Mexico but was negative in the other two countries; and principals' perceptions of their leadership for professional development were negatively correlated in Korea but positively in Mexico and the United States.

Looking at distributed leadership with PISA 2015 data, [Luschei and Jeong \(2021\)](#) report that increases in teacher decision-making responsibilities were positively related to student achievement in math, reading, and science. This was not the case for any other aspect of principal leadership, however.

The final set of studies on leadership and attainment use TIMSS and PIRLS data. Again, these tend to be from earlier studies as the emphasis on leadership has decreased in recent administrations.

Comparing the US and Flanders using TIMSS (2011) data, [Urlick et al. \(2022\)](#) found that principal instructional leadership was related to opportunity to learn in both countries, but that the relationship was stronger in Flanders. The pathways were also differentially mediated.

Using TIMSS (2007) data to look at eight countries (Australia, Colombia, Egypt, England, Hungary, Korea, Turkey and USA) [Shin and Slater \(2010\)](#) found that neither the time that principals spent on instructional leadership (e.g., developing curriculum and pedagogy) nor the time they spent supervising and evaluating teachers and other staff were related to student achievement.

Elements to relate other aspects of principal's work to attainment have also not been particularly successful. [Sealy et al. \(2016\)](#) used the 2012–13 PISA-TALIS link study data to look at the relationship between principal job satisfaction and attainment, but found a significant relationship in only 2 out of 8 participating countries.

While the relationship between leadership and attainment remains questionable from this data, there is some evidence from international studies of a significant relationship between leadership and other variables relating to school and teacher effectiveness through which leadership could have an indirect effect on pupil outcomes.

Using TALIS 2018 data, for example, [OECD \(2021b\)](#) report a positive correlation between teacher autonomy, innovation and collaboration. [O'Shea \(2021\)](#) found a positive relationship between distributed leadership and teachers' use of innovative practices in schools. [Bellibaş et al. \(2021\)](#) found a positive relationship between principals' instructional leadership and instructional quality, while the effect of distributed leadership was mainly indirect, mediated by teacher collaboration and job satisfaction. Using the same dataset, [Liu et al. \(2021\)](#) report that distributed leadership and instructional leadership were both positively and directly associated with teacher job satisfaction and teacher self-efficacy.

[Maslowski et al. \(2016\)](#) in their report on TALIS 2013 note the relationship between instructional leadership and teacher collaboration. They also found a relationship between instructional leadership and the extent to which principal feedback leads to positive change in instructional practices. Distributed leadership, in turn, was related to better staff-student relationships and a greater sense of shared purpose in the school. Similarly, [Bellibas and Liu \(2018\)](#) found a positive relationship between distributed leadership and staff mutual respect, but not with school safety, while [Sun and Xia \(2018\)](#) show a positive relationship between teacher self-efficacy and job satisfaction. In their report on TALIS 2013, [OECD \(2016\)](#) point to a positive relationship between distributed leadership and sense of purpose across the school. They also found that in schools where distributed leadership was practised there were more positive relations between students and staff. Schools in which leadership included elements of both distributed and instructional leadership were more likely to show the characteristics of professional learning communities. However, they also report that principals have only limited, indirect influence on the establishment of the learning climate in the school.

Extent of instructional leadership in schools is correlated with the extent to which principals are able to involve teachers, parents and stakeholders in school decisions ([OECD, 2021b](#)), suggesting that distributed and instructional leadership are correlated with one another.

Generally, then, the LISA studies present a mixed picture of the relationship between leadership and student attainment. It is by no means clear that there is a consistent relationship in these studies, calling into question the strong conclusions often drawn about leadership effects. On the other hand, there is converging evidence of a positive relationship between leadership practices and staff satisfaction and behaviors, which one would hypothesize as benefitting pupil outcomes.

Conclusion

There are a few conclusions to be drawn from these studies.

Firstly, the frameworks used reflect current orthodoxy in educational leadership research in their emphasis on core generic aspects such as distributed leadership, taken from the prevailing consensus in educational leadership research.

Secondly, we have over time seen a reduction in the attention spent on leadership in the LISA studies, this both in the OECD and IEA. Scales and items in PISA and TIMSS/PIRLS school questionnaires which were previously devoted to leadership have been replaced by items on other important concerns, such as diversity.

Thirdly (and probably not unrelatedly), results with regards to the impact of leadership have been disappointing. There is little in the studies summarized above that would support an emphasis on leadership as a factor that can help explain differences in pupil attainment. These findings therefore somewhat contradict the bulk of research conducted outwith the LISA studies.

It has for quite some time been a given in the literature on leadership and school effectiveness that leadership has a significant impact on attainment. Furthermore, there is a general consensus (at least until recently) on what factors matter, to the extent that Leithwood et al. (2020) claim to be able to make “strong statements on educational leadership”. The findings from LISAs are far less clear, however. Invariance across countries of what leadership means is not a given, and relationships with attainment are variable.

Why is this? We can of course not come to any definitive empirical conclusions on the basis of this review of the literature, but there are a number of possible explanations.

Firstly, it is possible that the correlational nature of LISA’s does not allow the full impact of leadership on outcomes to be measured, especially as the relationship is largely indirect. However, the majority of non-LISA studies on which the claim of significant impacts are based are also cross-sectional (Muijs, 2011).

Secondly, it is possible that due to the inevitable limitations of space within a wide-ranging survey of busy professionals means that the concepts are measured by fewer items than in studies focused specifically on leadership, potentially reducing reliability. This is certainly the case in comparison to some, if not all other studies. However, careful development and selection of items, as certainly happens in the LISA studies, should at least partially mitigate this effect.

Thirdly, due to the international nature of the data the constructs may be differently understood in different countries (the frequent finding of limited invariance suggests this may be the case). However, in-country studies using LISA data tend to converge around relatively weak effects, suggesting that this too is not the full answer.

Lastly, it is possible that the rigorous measures used in LISA impose constraints that the existing leadership studies cannot meet (for example in how attainment is measured), and that other studies may therefore overestimate effects. There is also potentially an issue of publication bias in existing studies that may have led to an overestimation of effect sizes.

The fact of these disappointing results makes the lower emphasis on leadership in recent iterations of PISA and TIMSS/PIRLS understandable. However, the level of evidence in other studies, and the theoretical basis for positing a leadership effect, mean that it is nonetheless unfortunate. An interesting exercise would be to see whether some of the newer hypotheses around leadership (see Lock and Rees, 2020), could be tested in future iterations. This would be valuable as while theoretically intriguing in light of similar developments in the study of teaching and learning, they are not currently supported by much empirical evidence.

Both theoretically and intuitively it would seem that leadership is bound to matter, and non-LISA studies tend to suggest this is the case. So let us not give up on studying this factor in relation to pupil attainment, but instead seek to refine our understanding of leadership, and develop instruments and scales that are able to measure accordingly.

References

- Ainley, J., Carstens, R., 2018. Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework. <https://doi.org/10.1787/19939019>. https://www.oecd-ilibrary.org/education/teaching-and-learning-international-survey-talis-2018-conceptual-framework_799337c2-en.
- Bellibas, M.S., Liu, Y., 2018. The effects of principals’ perceived instructional and distributed leadership practices on their perceptions of school climate. *Int. J. Leader. Educ.* 21 (2), 226–244. <https://doi.org/10.1080/13603124.2016.1147608>.
- Bellibas, M.S., Gümüş, S., Liu, Y., 2021. Does school leadership matter for teachers’ classroom practice? The influence of instructional leadership and distributed leadership on instructional quality. *Sch. Effect. Sch. Improv.* 32 (3), 387–412. <https://doi.org/10.1080/09243453.2020.1858119>.
- Chen, D., 2018. The Relationship Between School Leadership and Student Mathematics Achievement. A Comparative Study between Germany and Chinese Taipei Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.) Technische Universität Dortmund. Fakultät Erziehungswissenschaft, Psychologie und Soziologie.
- Chen, D., Ning, B., Bos, W., 2020. Relationship between principal leadership and students’ mathematics achievement: a comparative study between Germany and Chinese Taipei. *Asia Pac. J. Educ.* <https://doi.org/10.1080/02188791.2020.1816899>.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258. <https://doi.org/10.1177/0013161X15616863>.
- Gao, X., Xia, J., Shen, J., Ma, X., 2018. A comparison between U.S. And Chinese principal decision-making power: a measurement perspective based on PISA 2015. *Chin. Educ. Soc.* 51 (5), 410–425. <https://doi.org/10.1080/10611932.2018.1510691>.
- Greatbatch, D., Tate, S., 2018. Evidence on School Leadership in an International Context. Department for Education, London. https://www.lse.org.uk/uploads/datahub/4567ceb%5E03in-03/2018-08-23-School_Leadership_Review.pdf.
- Huber, S.G., Muijs, D., 2010. School leadership effectiveness: the growing insight in the importance of school leadership for the quality and development of schools and their pupils. In: Huber, S.G. (Ed.), *School Leadership: International Perspectives*. Springer, Berlin, pp. 57–77.
- Jerrim, J., Sims, S., 2019. The Teaching and Learning International Survey. Research Report. Department for Education, London. https://dera.ioe.ac.uk/33612/1/TALIS_2018_research.pdf.
- Leithwood, K., Seashore-Lewis, K., Anderson, S., Wahlstrom, K., 2004. How Leadership Influences Student Learning. The Wallace Foundation, New York. <http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Documents/How-Leadership-Influences-Student-Learning.pdf>.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. Successful School Leadership: What It Is and How It Influences Pupil Learning. DFES and Nottingham: NCSL, London.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22. <https://doi.org/10.1080/13632434.2019.1596077>.
- Liu, Y., Bellibas, M.S., Gümüş, S., 2021. The effect of instructional leadership and distributed leadership on teacher self-efficacy and job satisfaction: mediating roles of supportive school culture and teacher collaboration. *Educ. Manag. Adm. Leader* 49 (3), 430–453. <https://doi.org/10.1177/1741143220910438>.
- Lock, S., Rees, T., 2020. The ResearchEd Guide to School Leadership. John Catt Educational, Woodbridge.
- Luschei, T.F., Jeong, D.W., 2021. School governance and student achievement: cross-national evidence from the 2015 PISA. *Educ. Adm. Q.* 57 (3), 331–371. <https://doi.org/10.1177/0013161X20936346>.
- Maslowski, R., Rekers-Mombarg, L., Bosker, R., Veldman, M., 2016. School Leadership for Learning: Insights From TALIS 2013. OECD. <https://doi.org/10.1787/9789264258341-en>.
- Mora-Ruano, J.G., Schurig, M., Wittmann, E., 2021. Instructional leadership as a vehicle for teacher collaboration and student achievement. What the German PISA 2015 sample tells us. *Front. Educ.* <https://doi.org/10.3389/educ.2021.582773>. <https://www.frontiersin.org/articles/10.3389/educ.2021.582773/full>.

- Muijs, D., 2011. Leadership and organisational performance: from research to prescription? *Int. J. Educ. Manag.* 25 (1), 45–60. <https://doi.org/10.1108/09513541111100116>.
- Mullins, I., Martin, M., Foy, P., Drucker, K., 2012. PIRLS 2011 International Results in Reading. IEA.
- O'Shea, C., 2021. Distributed leadership and innovative teaching practices. *Int. J. Educ. Res. Open* 2. <https://doi.org/10.1016/j.ijedro.2021.100088>. <https://www.sciencedirect.com/science/article/pii/S2666374021000583?via%3Dihub><https://www.sciencedirect.com/science/article/pii/S2666374021000583?via%3Dihub>.
- OECD, 2016. School Leadership for Learning. <https://doi.org/10.1787/9789264258341-en>. Insights from TALIS 2013. https://www.oecd-ilibrary.org/education/school-leadership-for-learning/executive-summary_9789264258341-2-en.
- OECD, 2017. PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving. OECD Publishing, Paris. <https://doi.org/10.1787/9789264281820-9-en>.
- OECD, 2019. PISA 2018 Assessment and Analytical Framework. <https://www.oecd-ilibrary.org/docserver/b25efab8-en.pdf?expires=1655225506&id=id&accname=guest&checksum=348667E346C36D71384067F654E4D9FD>.
- OECD, 2021a. Positive, High-Achieving Students? What Schools and Teachers Can Do. <https://www.oecd-ilibrary.org/sites/9bc425ba-en/index.html?itemId=/content/component/9bc425ba-en>.
- OECD, 2021b. TALIS2018 Results Vol II: Teachers and School Leaders as Valued Professionals. https://www.oecd-ilibrary.org/sites/19cf08df-en/1/3/6/index.html?itemId=/content/publication/19cf08df-en&_csp_=67e65b72be0b468ed3dac915593716de&itemGO=oecd&itemContentType=book#s135.
- OECD, 2021c. PISA2021 Context Questionnaire Framework. <https://www.oecd.org/pisa/sitedocument/PISA-2021-questionnaire-framework.pdf>.
- Printy, S., Liu, Y., 2021. Distributed leadership globally: the interactive nature of principal and teacher leadership in 32 countries. *Educ. Adm. Q.* 57 (2), 290–325. <https://doi.org/10.1177/0013161X20926548>.
- Reynolds, K.A., Mullins, I.V.S., Martin, M.O., 2023. TIMSS 2023 Context Questionnaire Framework. https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks_Ch3_Context-Questionnaires.pdf.
- Robinson, V., Gray, E., 2019. What difference does school leadership make to student outcomes? *J. Roy. Soc. N. Z.* 49 (2), 171–187. <https://doi.org/10.1080/03036758.2019.1582075>.
- Sealy, K.M., Perry, S.M., DeNicola, T., 2016. Relationships and Predictors of Principal Job Satisfaction across Multiple Countries: A Study Using Talis 2013 and PISA 2012. <https://doi.org/10.2139/ssrn.2717056>. Available at SSRN. <https://ssrn.com/abstract=2717056>.
- Shin, S.-H., Slater, C.L., 2010. Principal leadership and mathematics achievement: an international comparative study. *Sch. Leader. Manag.* 30 (4), 317–334. <https://doi.org/10.1080/13632434.2010.498995>.
- Shin, S.-H., Slater, C.L., Backhoff, E., 2013. Principal perceptions and student achievement in reading in Korea, Mexico, and the United States: educational leadership, school autonomy, and use of test results. *Educ. Adm. Q.* 49 (3), 489–527. <https://doi.org/10.1177/0013161X12458796>.
- Sun, A., Xia, J., 2018. Teacher-perceived distributed leadership, teacher self-efficacy and job satisfaction: a multilevel SEM approach using the 2013 TALIS data. *Int. J. Educ. Res.* 92 (1), 86–97. <https://doi.org/10.1016/j.ijer.2018.09.006>.
- Urick, A., Bowers, A.J., 2019. Assessing international teacher and principal perceptions of instructional leadership: a multilevel factor analysis of TALIS 2008. *Leader. Pol. Sch.* 18 (3), 249–269. <https://doi.org/10.1080/15700763.2017.1384499>.
- Urick, A.M., Ford, T.G., Page Wilson, A.S., Consuegra, E., 2022. How does instructional leadership influence opportunity to learn in mathematics? A comparative study of pathways for grade 4 students in the U.S. and Belgium. *Res. Comp. Int. Educ.* <https://doi.org/10.1177/17454999221086360>.
- Veletić, J., Vegar Olsen, R., 2021. Exploring school leadership profiles across the world: a cluster analysis approach to TALIS 2018. *Int. J. Leader. Educ.* <https://doi.org/10.1080/13603124.2021.1953612>.
- Wu, H., Gao, X., Shen, J., 2019. Principal leadership effects on student achievement: a multilevel analysis using Programme for International Student Assessment 2015 data. *Educ. Stud.* 46 (3), 316–336. <https://doi.org/10.1080/03055698.2019.1584853>.
- Wu, H., Gao, X., Shen, J., 2020a. Principal leadership effects on student achievement: a multilevel analysis using Programme for International Student Assessment 2015 data. *Educ. Stud.* 46 (3), 316–336. <https://doi.org/10.1080/03055698.2019.1584853>.
- Wu, H., Shen, X., Zhang, Y., Zheng, Y., 2020b. Examining the effect of principal leadership on student science achievement. *Int. J. Sci. Educ.* 42 (6), 1017–1039. <https://doi.org/10.1080/09500693.2020.1747664>.
- Xia, J., O'Shea, C., 2022. To what extent does distributed leadership support principal instructional leadership? Evidence from TALIS 2013 data. *Leader. Pol. Sch.* <https://doi.org/10.1080/15700763.2022.2056059>.
- Zhang, S., Wind, S., 2019. Cross-cultural comparisons of school leadership using Rasch measurement. *J. Appl. Meas.* 20 (2), 1–17.

A review of research on middle leaders in schools

David Gurr, The University of Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	115
Middle leader definition	115
Middle leaders in schools	115
Reflection	119
References	121

Introduction

This paper provides a reflection on research exploring middle leaders in schools. It begins with definitional issues and arrives at a statement of what middle leaders are and are not. On the basis of this, it then explores reviews and research since 2006 that are helpful to understand the work of middle leaders. It concludes with five trustworthy statements about middle leaders and consideration of future research. It is not a systematic review of the literature and I do not claim completeness. Rather, it is a somewhat personal narrative, but one which will consider important papers that help to understand the phenomenon, highlight conceptual and methodological issues, and make suggestions for future research and practice directions.

Middle leader definition

It is always important to define clearly the focus of a review and in the middle leader area, this is problematic. For example, [De Nobile \(2017\)](#) described differences between middle managers, middle leaders and teacher leaders. He also noted that the phenomenon is bound by context and perspective. For De Nobile, middle leaders are those people who are positioned organizationally below the principal and other senior leaders (if present), but above teachers. They differ from teachers in that they will have a defined organizational responsibility, which will typically be part of a formal role, and they differ from senior leaders in that they will retain a substantial teaching responsibility that is likely to be their main work focus ([Grootenboer et al., 2015](#)). Defined this way, middle leaders are likely to have position titles such as director of teaching and learning, curriculum coordinator, subject coordinator, domain leader, professional learning team leader, head of department, student well-being coordinator, year level coordinator and so forth, depending on context (e.g., location, school type and level). Obviously principals are excluded, but so too are deputy principals, or those with similar overseeing roles such as a head of a campus or school section. These people are better placed in a senior leadership category although some researchers include them as middle leaders (e.g., [Cranston, 2009](#)) and others note that an administrative-focused deputy principal can be less like a leader than a middle leader with significant strategic responsibilities ([De Nobile and Ridden, 2014](#)). Teachers, or teacher leaders, are also excluded. Again, there is a frustration with the literature, because it is common for definitions of teacher leaders to not make distinctions between teachers and middle leaders, with this particularly evident in North American literature ([Wenner and Campbell, 2017](#); [York-Barr and Duke, 2004](#)). There are further confusions such as: the term “leading from the middle” which can be either a system view focused on principals (e.g., [Flessa, 2012](#); [Fullan, 2015](#); [Hargreaves and Ainscow, 2015](#)), or focused on schools (e.g., [Day and Grice, 2019](#); [Edwards-Groves et al., 2019](#); [Reid, 2020](#)); the inclusion of non-teaching staff as middle leaders ([De Nobile, 2019](#)).

For this paper, middle leaders are teachers who have an additional formal organizational responsibility. The research reviewed here will conform to this definition as much as possible, unless stated otherwise.

Middle leaders in schools

In 2017, Harris and Jones in an editorial for *School Leadership and Management*, made several important observations about middle leaders. They suggested that the interest in middle leadership had faded, and that the literature on distributed and teacher leadership “may just have overshadowed and supplanted any interest in middle leadership roles” ([Harris and Jones, 2017](#), p. 214). Their paper was a call to for a renewed research effort.

Since then, *School Leadership and Management* has published a seven-article special issue edited by [Grootenboer et al. \(2017, 2019\)](#), which included a review on middle leaders ([Harris et al., 2019](#)). [Harris et al. \(2019\)](#) took, as their point of departure, the middle leader reviews conducted by Bennett and colleagues a decade earlier ([Bennett, 2005, 2006](#); [Bennett et al., 2003, 2007](#)). Before considering the review of [Harris et al. \(2019\)](#), it is worth reflecting on these earlier reviews.

The reviews of Bennett and colleagues were focused on subject leaders, middle managers, heads of department and curriculum coordinators, mostly in secondary contexts, mostly in England or comparable countries (Australia, Canada, Hong Kong, New Zealand, the Netherlands and the USA; and additionally Singapore and South Africa for the 2003 review) reporting on mostly small-scale qualitative or survey-based empirical studies from English language journals from 1988 and books from 1995. The sources for the review included six major databases: EBSCO, SSCI, ERIC, Education Line, JSTOR and SwetsWise. The initial 2003 review included 101 articles, the 2005 review added another 51 articles, with the 2006 and 2007 papers providing further review data. The conclusions arising from these reports include:

- Middle leaders can be important to schools in terms of their influence on teacher work and student learning, but the nature of this is context dependent.
- They have a focus on their responsibility area and sometimes this means there is a tension between meeting expectations at the whole school level and supporting their area. Senior leaders want them to be involved more broadly in whole-school matters, but the middle leaders are more comfortable with a focus on their responsibility area.
- Relatedly, there is tension between line management demands and the collegiality these leaders seek to foster. This is shown by increasing supervisory and accountability requirements (often enacted through observation of colleague's work) in competition with the desire and need to work collaboratively with colleagues in improving teaching and learning.
- These teachers will often have doubts about whether they have sufficient teacher competence and subject knowledge to be advising others.
- There was a need to study the influence of this leadership work on teaching and learning, and to understand the effectiveness of the professional development of middle leaders.

From these reviews, it seemed that middle leaders were acknowledged as being potentially important to school success, but with some caveats related to context, expectations, support and tensions; that is, middle leaders can positively influence teacher work and student learning, but this depends on how their role is constructed and how they are selected and supported in the role to do this influence work. I will return to consider evidence about the impact of middle leaders later in the paper.

The review of [Harris et al. \(2019\)](#) builds upon this work by Bennett and colleagues and reflects the changing nature of academic publications. They relied on journal articles that were English language, published, peer-reviewed, empirical and available through searches on EBSCOHOST, Web of Science, SCOPUS and six further education databases. The journal scope was perhaps broader than the earlier reviews, but the review was narrower in that other "materials such as books, masters and Ph.D. theses, conference proceedings, policy papers and articles published in non-ISI/SCOPUS journals were excluded" ([Harris et al., 2019](#), p. 270). The review by [Harris et al. \(2019\)](#) departs from the earlier reviews in that it is largely a quantitative exercise identifying broad trends in regard to knowledge production on middle leadership in schools; it does not provide the more nuanced perspectives that Bennett and colleagues gained from reading their review papers closely. Their criteria for inclusion meant that the [Harris et al. \(2019\)](#) review used 49 articles. The review did not provide bibliographic details about all the papers reviewed. As is the trend of late with science mapping reviews (e.g., [Kovacević and Hallinger, 2020](#)), it is a bibliometric review that is absent of substantial content. Rather, it reports on patterns and connections. Findings included:

- The amount of research increased but not greatly; there were 2.8 publications per year between 2003 and 2007 (5 years), compared to 3.5 publications per year between 2008 and 2017 (10 years).
- Nearly 50% of papers were published in journals that are not within the field of educational leadership, although the main journal for publication was *School Leadership and Management* with 22% of the publication; this is a journal strongly connected with the UK.
- Most research was published in the UK, although increasing country diversity was evident (especially from Asian and Middle East countries post-2007)
- 43 of the 49 articles used qualitative or mixed-methods, with the most common methods being interview. Post 2007, there was an increase in use of more diverse methods with interviews being supplemented with observation and document analysis.
- More than half were focused on leadership practice and roles, with other areas of focus including professional development/agency, professional identity, culture and leadership distribution.

There are reviews and reports of research streams published after the reviews by Bennett and colleagues that provide some of the rich detail missing from the review of [Harris et al. \(2019\)](#), and it is these to which I now turn. In the following paragraphs, I describe a broadly focused middle leader review by [De Nobile \(2017\)](#), before considering examples of two research programs; the practice architecture research by [Grootenboer et al. \(2019\)](#), and a series of doctoral research from Australia ([Gurr, 2019](#); [Gurr and Drysdale, 2013](#)). I conclude by exploring leadership impact studies by [Dinham \(2005, 2007\)](#) and [Highfield \(2012\)](#), and [Leithwood's \(2016\)](#) review of secondary school heads of departments.

De Nobile has produced a middle leadership review ([De Nobile, 2018](#)), constructed a model informed by the review ([De Nobile, 2017, 2018](#)) and is developing a questionnaire ([De Nobile, 2019](#)). The 2018 review was modeled on the work of [Bennett et al. \(2003\)](#) and involved a search of databases from 2003 to 2016 that included ERIC, A+ Education, EBSCO, JSTOR and Informit. The search uncovered 252 journal articles, book articles and books, institutional reports, and school system reports. Search terms began with middle management and middle leadership but were expanded to include more publications by using typical middle leader positional terms such as department head, subject leader, year advisor. Teacher leader and teacher leadership were also included as search terms and this, unfortunately, leads to some doubts about the veracity of the findings in terms of middle leaders. Through

content analysis of definitions, descriptions, findings, recommendations and practices relating to middle leadership, and recording of publication demographics, a comprehensive review was provided. The review did not provide bibliographic details about all the papers reviewed, with only about half the reference details provided in [De Nobile \(2017, 2018\)](#). Findings from the review included:

- A decrease in publications between the periods 1999–2007 and 2008–2016. Note that the longer time frame compare to [Harris et al. \(2019\)](#) suggests a different pattern.
- Of the 234 publications with a single country focus, most were from the UK (42%), Australia (26%) and the USA (12%). Note, this review has not identified the increasing diversity found in [Harris et al. \(2019\)](#).
- Most publications were in academic journals (67%), books (10%) or reports (8%).
- Publications were research reports (73%), reviews (14%), conceptual papers (9%) or policy papers (4%).
- Publications were focused on primary and secondary schools (36%), secondary schools (34%) or primary schools (15%).

There were also findings in regard to terminology use. [De Nobile \(2018\)](#) noted both the overlap and distinctiveness between teacher leadership and middle leadership research, yet, as noted already, included teacher leadership in his search terms. [De Nobile \(2018\)](#) found that middle manager was the preferred term from 1999 to 2005 and was then supplanted with the emergence of middle leader and teacher leader such that by 2008 the term middle manager had largely vanished. [De Nobile \(2018\)](#) suggested that the terms middle leadership and teacher leadership were becoming interchangeable.

The review by [De Nobile \(2018\)](#) then focused on reporting evidence that informed the middle leader model. [De Nobile \(2017, 2018\)](#) describes an input/output role model that helps to clarify the influences on the work of middle leaders:

- **Inputs:** principal support, school/system culture, professional development, enthusiasm/drive, and knowledge of curriculum, pedagogy and assessment.
- **Roles:** student focused, administrative, organizational, supervisory, staff development and strategic. These are ordered from managerial to leadership.
- **Enactment of roles:** managing relationships, leading teams, communicating effectively, managing time and managing self.
- **Outputs:** teacher quality, teacher attitudes and student outcomes.

[De Nobile \(2018\)](#) provides some of the evidence base for each element of the model and there is further clarification in [De Nobile \(2019\)](#). Suggestions for further research centered on quantitative and qualitative empirical research to confirm and develop the model ([De Nobile, 2018, 2019](#)). Input/output middle leader models have been developed before, such as those focused on heads of departments by [Turner and Bolam \(1998\)](#) and [White \(2000\)](#).

A different approach to researching middle leaders is found in the work of Grootenboer and Edwards-Groves from Australia, and Rönnerman from Sweden. Over several years they have used practice architectures to explore the interactions between persons, practices and contexts (e.g., [Edwards-Groves et al., 2019](#); [Grootenboer, 2018](#); [Grootenboer et al., 2017](#)). In practice architectures, beliefs (sayings), practices (doings) and relationships (relatings) in various contexts (including temporal) need to be considered. This view is not dissimilar to Spillane's interaction approach to researching distributed leadership ([Spillane, 2006](#); also see [Grootenboer, 2018](#), p. 24), and, indeed, [Edwards-Groves et al. \(2019\)](#) view their research as extending Spillane's research. This work is mostly small scale and interview-based, often in primary settings, sometimes using action-research and sometimes focused on how middle leaders support professional learning of their peers. Surprisingly given the practice focus, while observation may be included in the research, sometimes the observations are not used to inform the findings (e.g., [Edwards-Groves et al., 2019](#)). [Grootenboer \(2018\)](#) provides a complex view of this research in a book about the practices of school middle leaders that uses a varied (conceptually, methodologically and in terms of the research and school focus) collection of five studies. He describes three broad practices of middle leaders related to leading teaching (termed leading-teaching to emphasize simultaneous work in both leading and teaching), managing and facilitating, and collaborating and communicating, with a focus on curriculum development and professional development of teachers. This is explained in [Grootenboer et al. \(2017, p. 248\)](#) as:

The practice of middle learning involves engaging in (simultaneous) *leading-teaching* by *managing and facilitating* educational development through *collaborating and communicating* to create communicative spaces. [italics from original]

Communicative spaces are where the practice of middle leading occurs. [Grootenboer \(2018\)](#) provided many examples of practice from the five studies that informed the book.

Doctoral research is often not represented well in reviews, yet they can make a significant contribution. As an example, I will refer here to a doctoral research series exploring middle leaders that Drysdale and I have supervised, and, later in the paper, to a thesis by [Highfield \(2012\)](#). [Gurr \(2019\)](#) reported on a sequence of six doctoral theses over a 20-year period that focused on the work of middle leaders (the five published theses are [Cotter, 2011](#); [Huerta Villalobos, 2018](#); [Keane, 2010](#); [Koh, 2018](#); [White, 2000](#)). This built upon an earlier report by [Gurr and Drysdale \(2013\)](#). All studies utilized multiple-perspective interviews. Four were from secondary schools in Melbourne, Australia, one researched primary schools in Singapore and one researched secondary schools in Chile. All were focused on teachers with middle leader roles who were leading curriculum or learning areas, and one also included teachers in pastoral roles. The review found that increasingly middle leaders were being viewed as key personnel in improving teaching and learning, and when they had well-defined roles and sufficient expectation and support could impact on

student outcomes through providing direction, leading professional support and focusing on improving teaching practice. Conclusions about student impact were based on perception and not through formal evidence of learning gains. Unfortunately, a consistent finding over the two decades covered was that too often the middle leaders had limited impact on improving teaching and learning, did not receive sufficient support from senior leaders, and worked in school structures that hindered their work. To support the development of middle leaders it was suggested that they needed to be career focused, supported in their development by principals and other senior leaders, with access to appropriate professional learning programs, supported in a student-focused school culture and structure, and with clearly defined and high expectation role descriptions.

What is surprising in the middle leader research is that there is often little said about the impact of middle leaders on student learning. Mainly the evidence is centered on their influence on colleagues, and then anecdotally on students. There is an assumption that this work then leads to positive impacts on student learning (Day and Grice, 2019). There is, however, some research that makes a more substantial case for impact on student learning and three examples are explored below (Dinham, 2005, 2007; Highfield, 2012; Leithwood, 2016).

One of the most compelling studies in terms of the impact of middle leaders has been the ÆSOP (An Exceptional Schooling Outcomes Project). ÆSOP involved the study of 50 subject departments and cross-school programs (e.g., student welfare) in years 7–10, across 38 secondary schools in New South Wales, Australia (Dinham, 2005, 2007). Initial selection of the 50 “sites” was based on a combination of reputation (nomination from parent groups, principals, system officers) and analysis of publicly available data on student achievement (public examination performance and value-adding data) that demonstrated outstanding educational outcomes over at least a 4-year period. The study involved a reflection on what influenced the achievement of these results. Multiple-perspective, individual and group interviews involved the principal, head teacher/leader of the outstanding department/program, staff group forum, two student forums (year 7/8 and 9/10 groups), parent forum, classroom observation and document analysis. Principal leadership (Dinham, 2005), and leadership of the Heads of Department/Programs (Dinham, 2007) were both important for student success. The middle leaders were found to promote success through a combination of core and contributing strategies (Dinham, 2007).

Core strategy

- A focus on students and their learning across academic, personal and social dimensions.

Contributing strategies

- High-level interpersonal skills, and generally being well-liked and trusted.
- High-level professional capacity and strategic resource allocation.
- Promotion and advocacy of their departments and maintaining good external relations with the school.
- Influencing department planning and organization.
- Developing common purpose, collaboration and sense of team within their department.
- Fostering teacher learning and developing a culture of shared responsibility and trust.
- Clear vision, high expectations of themselves and others, and developing a culture of success.

While the focus of the research was on the outstanding faculties and teams, it also found that principal, and other senior leadership support, were important contributors to site success. Principal leadership that supported middle leader success included (Dinham, 2005):

- Having a core focus on students and their learning across academic, personal and social dimensions.
- External awareness and engagement with the wider environment
- A bias toward innovation and action
- High level interpersonal skills, and generally being well-liked and trusted
- Having a clear vision, high expectations and fostering a culture of success
- Encouraging teacher learning and responsibility, and showing trust
- Promoting student support, developing common purpose and encouraging teacher collaboration

Dinham (2007) noted the significant overlap between the strategies of middle leaders and principals, with the main difference being that middle leader work was, understandably more focused on their area responsibility than the wider school, albeit that some middle roles have a whole-school focus (e.g., student welfare). This research highlighted both how middle leaders contribute to teacher and student success, and the importance of the principals and other senior leaders in “providing the conditions where teachers can operate effectively and students can learn” (Dinham, 2005, p. 355).

A study that provided estimates of the impact of middle leaders on student learning was a quantitative New Zealand study on secondary Heads of Department by Highfield (2012). This is another example of doctoral thesis research, and, in this case, provides findings that build upon the qualitative ÆSOP study (as noted by Highfield, 2012). Drawing upon literature on within-school variation and middle leaders, Highfield (2012) constructed a three-phase study. The first explored within-school variation in English, mathematics and science departments in 41 secondary schools. The second phase used 10 of the 41 schools that volunteered to be involved in a survey-based exploration of middle leadership practices; 95 (87%) middle leaders, and 232 (65%) teachers completed complementary surveys on middle leader practices. The third phase connected middle leadership practices with student achievement. Middle leadership was mostly found to contribute positively to student learning. For the last two years of secondary schooling, middle leadership practices were shown to explain an important amount of student achievement variation; for the third

last year of secondary school there was no significant relationship found. At Level 2 NCEA (equivalent to the second last year of secondary schooling), while school decile (a measure of school community socio-economic status and educational advantage) accounted for 62% of variance in student academic outcomes, middle leadership accounted for an additional 22%, while at Level 3 NCEA, school decile accounted for 46% and middle leadership 16% of variance. Highfield (2012, p. 155) concluded that,

a combination of school decile and middle leadership practices were strong overall predictors of student academic achievement at Level 2 and 3 NCEA. This is a critical finding of the study because it demonstrates that middle leaders and the teachers in their departments can make a difference to student academic outcomes despite the socioeconomic status or ethnicity of students.

The most important leadership practices impacting positively on student outcomes were management of resources, goals and expectations, and creating a positive learning environment for students and teachers. Negative impacts were, surprisingly, compared to other research (e.g., Bennett, 2005; Dinham, 2005; Tay et al., 2020), associated with practices that promote collegiality among teachers, and a focus on academic results.

Leithwood (2016) conducted a review of research focused on secondary department heads to address a lack of attention to this middle leadership role. The review overlaps considerably with other reviews and some of the research already noted; for example, the research of Dinham (2007) and Highfield (2012) are represented in the review, as is much of the earlier UK research on department heads such as the work of Siskin (1991, 1997) and Brown and Rutherford (1998), which were included in the review by Bennett et al. (2003). Leithwood (2016) searched five educational leadership journals (*Educational Administration Quarterly*, *International Journal of Leadership in Education*, *School Leadership and Management*, *Educational Management, Administration and Leadership*, *Management in Education*), three other education journals (*School Effectiveness and School Improvement*, *Journal of Educational Research*, *Educational Researcher*), and the ERIC, Proquest and Google Scholar databases. The period of review seems to have covered at least 30 years, with some studies from the mid-1980s included. The review uncovered 42 studies (29 qualitative, 7 quantitative and 6 mixed-method), with most of the studies from the UK (16) and the USA (16). Additionally, 15 literature reviews and 10 related studies were used to inform the findings. It was suggested that contextual differences across countries were less important than differences in teaching cultures, union regulations, and department head and principal expectations about the role. Effective leadership practices were described by matching research evidence from 32 relevant studies with the Ontario Leadership Framework (OLF) (Leithwood, 2012), with Leithwood (2016, p. 131) suggesting that,

... effective department-head leadership practices and personal leadership resources are closely aligned to those practices and resources included in the OLF. Most of the differences are a reflection of the "middle leader" position of the department head in the school. This position creates a unique need to build productive relationships with individual teacher colleagues and heads of other departments, and to forge a significant school-wide decision-making role for themselves in collaboration with school-level leaders who have considerable discretion in determining how much authority to award their department heads.

Leithwood (2016) concluded that

"department and department-head effects on students are consistently positive, practically meaningful and larger than school effects" (p. 117), and "that secondary-school principals and department heads, acting in concert, maybe especially well-situated to provide both instructional and transformational leadership practices and, as a consequence, make powerful contributions to secondary-school improvement" (p. 136).

Like Dinham's (2005, 2007) research, and findings from other studies (e.g., Bassett, 2016; Cotter, 2011; Keane, 2010; White, 2000), it seems that having leadership expectations of the role that are shared by middle leaders and principals (and other senior leaders), and operating in a supportive environment, are especially important if middle leaders are to impact positively on teaching and learning. Leithwood (2016) distinguished between middle leaders that evaluate and supervise teachers versus those that offer collegial support and guidance, with the former having a far greater impact on teaching and learning. Relatedly, Farchi and Tubin (2019) found that successful schools tended to have middle leaders who focused more on pedagogical behavior than administrative behavior.

Reflection

The call from Harris and Jones (2017) for more middle leader research remains worthwhile, but it is clear from this and other reviews that there is a substantial existing knowledge base. Considering the state of knowledge from the reviews by Bennett and colleagues, there is some development since, but perhaps not as much as would be expected over the 13 years since 2007. The nature and extent of the development in knowledge, however, is somewhat dependent on review methodology, especially in relation to time period comparisons, what type of publications are included, and, for journals, which journals are included. Nevertheless, there is more clarity about the influence of middle leaders on teacher work and more certainty about their impact on

student learning—although the extent of both influences remains context dependent. Compared to the earlier reviews, increasingly middle leaders are being expected to be involved in whole-school as well as area responsibilities. The tension between supervision *of* staff versus collaboration *with* staff remains, although increased use of collaborative structures and enhanced knowledge of teaching and learning are enhancing the work of middle leaders and possibly reducing this tension.

Without overstating the knowledge base, in 2020 there are several trustworthy statements that can be made about middle leaders (indicative references provided):

1. Middle leaders can impact positively on teacher work and student outcomes (Carter, 2016; Dinham, 2005, 2007; Highfield, 2012; Leithwood, 2016). Indeed, they are an important source of leadership impact on student outcomes in many school contexts, and especially so when principals and middle leaders act together (Leithwood, 2016).
2. The way middle leaders improve student outcomes is through establishing the conditions (e.g., area direction, resource acquisition and distribution, staff development, supportive culture) that lead to collective endeavor by a group of teachers to improve teaching and learning (Bryant et al., 2020; Dinham, 2005; Grootenboer, 2018; Grootenboer et al., 2017; De Nobile, 2017, 2018; Gurr, 2019; Gurr and Drysdale, 2013). Developing staff collegiality can also have many benefits, but to have a positive influence on student outcomes this needs to have a focus on improving teaching and learning (Highfield, 2012; Leithwood, 2016).
3. High expectations and role clarity are important for these people to have a leadership focus rather than a managerial or administrative focus (Farchi and Tubin, 2019; Gurr, 2019). High expectations come from both the middle leaders themselves in terms of their sense of agency, and from principals and senior leaders about what they want from the roles (Gurr and Drysdale, 2013). Principals working in concert with middle leaders enhances the success and impact of middle leaders (Bassett, 2016; Dinham, 2007; Leithwood, 2016). Role clarity serves to give balance to leadership and management functions (Drysdale et al., 2016; Farchi and Tubin, 2019).
4. Recruitment and professional development are important both in terms of the quantity and quality of middle leaders (Gurr and Drysdale, 2013; Rhodes and Brundrett, 2009). To ensure there are sufficient teachers who have the skills and characteristics to be successful middle leaders there is a need to encourage teacher aspirations, have appropriate selection and induction processes, and ensure there is on-going review and development (Bassett and Robson, 2017; Gurr and Drysdale, 2013; Ng and Chan, 2014; Rhodes and Brundrett, 2009; Thorpe and Bennett-Powell, 2014).
5. There are factors that diminish the effectiveness of middle leaders with these being the negative of many of the factors noted above. These factors include a lack of understanding and organizational support by senior leaders, lack of professional preparation and leadership development, underdeveloped professional knowledge and capability, and a focus on administrative practices.

While there are some differences across the reviews about where the research is originating, there is a consistent presence of research from the UK over the past three decades. There is also likely to be a substantial contribution from the USA, although this is obscured somewhat with the focus on teacher leaders in preference to distinguishing between teacher and middle leaders. There is a consistent contribution from Australian researchers, with a notable increase from about 2017. Also evident from 2007 is the expansion of the knowledge base to include research from Asia and the Middle East. Of course, this is reporting upon research written in English language journals. A more complete review is needed that includes a more diverse language background, such as research published in Spanish and Chinese.

The research is dominated by qualitative and mixed method research, especially the use of interviews. In recent years, more diversity is evident with observation and document analysis being used along with interviews. Mixed method research, while not as prevalent, continues and the careful work of De Nobile (2017, 2019) is an example. Nevertheless, future research may need more innovative methods as indicated below.

While an effort has been made to only include middle leader research in this paper, some ambiguity is evident with, for example, the review by De Nobile (2018) deliberately including teacher leadership literature. This highlights an ongoing issue for middle leader research. There is, no doubt, considerable overlap between teacher leader, distributed leadership and middle leader research. It is not easy to resolve the overlap and what it means for the knowledge of leadership in schools—the segmentation allows for detail to be explored but may be artificial compared to the reality of working in schools. It is understandable that De Nobile (2018) wanted to expand his review by including teacher leadership, as there is literature in this area, which is relevant to a study of middle leaders, especially from the USA where teacher leadership is a broad term (Murphy, 2005; Wenner and Campbell, 2017; York-Barr and Duke, 2004). Yet, it also means that his review was not entirely focused on middle leaders because he gave no sense of how he might have vetted the teacher leader papers for their suitability to inform knowledge about middle leaders.

It is also understandable to seek to include research about leadership distribution. For example, the distributed leadership literature, while also suffering from conceptual confusion has much to say about the how non-principal leadership impacts on schools (Tian et al., 2016). While there is evidence through social network analysis that many in schools can have substantial influence (Supovitz et al., 2019), there is also evidence that middle leaders are important (Bryant et al., 2020; Nicholas, 2019). For example, Nicholas (2019) conducted a network analysis across three successful secondary schools which involved a total of 186 staff completing an eight-question network instrument. He then interviewed 26 highly connected staff about their work. In terms of the 20 most highly connected people in each school, there were only two people out of the 60 that didn't have a senior (7 people)

or middle leader (51 people) role. Being connected allows for leadership opportunities, and so the idea that “true” teacher leaders are common is clearly not the case. It is also clear that from a distributed leadership perspective, middle leaders are likely to be an important source for study.

Perhaps the contemporary reality of many schools, with increased use of collegial/collaboration arrangements (e.g., Hargreaves and O’Conner, 2018), means that dividing leadership research into principal, senior leaders, middle leaders and teacher leaders, and focusing on the distribution of leadership, is no longer sufficient. More holistic research is called for and this may benefit from a practice focus such as the interaction perspective of distributed leadership researchers (e.g., Spillane and Zuberi, 2009) or the practice architecture approach (e.g., Grootenboer, 2018). This will likely require innovative research strategies that can, in unobtrusive ways, collect rich practice data. In terms of ways to conceptualize middle leaders, while general leadership models can be adapted and used to locate the work of middle leaders (e.g., Gurr and Drysdale, 2013), and there are middle leader specific models (De Nobile, 2017, 2019), non-positional and relationship focused leadership models (e.g., Angelle and DeHart, 2016) and practice community views (e.g., Hord and Summers, 2008) are likely to become important to gain a better understanding of the work of middle leaders.

The writing of this paper predates the publication of another review of middle leaders (Lipscombe et al., 2021). The conclusions of the current paper are not changed by this more recent review, but there are additional details about the impact, work and characteristics of middle leaders that readers will find useful.

References

- Angelle, P.S., DeHart, C.A., 2016. Four models of teacher leadership: comparison and evaluation. *Res. Educ. Adm. Leadersh.* 1 (1), 85–119.
- Bassett, M., Robson, J., 2017. The two towers: the quest for appraisal and leadership development of middle leaders online. *J. Open Flexible Distance Learn.* 21 (2), 20–30.
- Bassett, M., 2016. The role of middle leaders in New Zealand secondary schools: expectations and challenges. *Waikato J. Educ.* 21 (1), 97–108.
- Bennett, N., Newton, W., Wise, C., Woods, P.A., Economou, A., 2003. The Role and Purpose of Middle Leaders in Schools. National College for School Leadership, Nottingham, UK.
- Bennett, N., Woods, P., Wise, C., Newton, W., 2007. Understandings of middle leadership in secondary schools: a review of empirical research. *Sch. Leader. Manag.* 27 (5), 453–470.
- Bennett, N., 2005. The Role of Subject Leaders in Secondary Schools Facing Challenging Circumstances: A Report for the National College for School Leadership. National College for School Leadership, Nottingham, UK.
- Bennett, N., 2006. Making a Difference: A Study of Effective Middle Leadership in Schools Facing Challenging Circumstances. National College for School Leadership, Nottingham, UK.
- Brown, M., Rutherford, D., 1998. Changing roles and raising standards: new challenges for heads of department. *Sch. Leader. Manag.* 18 (1), 75–88.
- Bryant, D.A., Wong, Y.L., Adames, A., 2020. How middle leaders support in-service teachers’ on-site professional. *Int. J. Educ. Res.* 100, 17.
- Carter, A., 2016. Empowering middle leaders—trends in school leadership research on the principal’s impact on school effectiveness. *Aust. Educ. Leader* 38 (1), 37–41.
- Cotter, M., 2011. Examination of the Leadership Expectations of Curriculum Coordinators in the Archdiocese of Melbourne—A Case Study Approach. Doctoral thesis. The University of Melbourne.
- Cranston, N.C., 2009. Middle-level school leaders: understanding their roles and aspirations. In: Cranston, N.C., Ehrich, L.C. (Eds.), *Australian School Leadership Today*. Australian Academic Press, Bowen Hills, QLD, pp. 217–241.
- Day, C., Grice, C., 2019. Investigating the Influence and Impact of Leading From the Middle: A School-Based Strategy for Middle Leaders in Schools. Executive Summary. Report Commissioned by the Association of Independent Schools Leadership Centre. University of Sydney, New South Wales, Sydney, NSW.
- De Nobile, J., Ridden, P., 2014. Middle leaders in schools: who are they and what do they do? *Aust. Educ. Leader* 36 (2), 22–25.
- De Nobile, J., 2017. Towards a theoretical model of middle leadership in schools. *Sch. Leader. Manag.* 38 (4), 395–416.
- De Nobile, J., 2018. The “State of the Art” of Research Into Middle Leadership in Schools. Paper Presented at the Australian Association for Research in Education Annual Conference, Sydney, December 2–6.
- De Nobile, J., 2019. The roles of middle leaders in schools: developing a conceptual framework for research. *Lead. Manag.* 25 (1), 1–14.
- Dinham, S., 2005. Principal leadership for outstanding educational outcomes. *J. Educ. Adm.* 43 (4), 338–356.
- Dinham, S., 2007. The secondary head of department and the achievement of exceptional student outcomes. *J. Educ. Adm.* 45 (1), 62–79.
- Drysdale, L., Gurr, D., Goode, H., 2016. Dare to make a difference: successful principals who explore the potential of their role. *Int. Stud. Educ. Adm.* 44 (3), 37–54.
- Edwards-Groves, C., Grootenboer, P., Hardy, I., Rönnerman, K., 2019. Driving change from “the middle”: middle leading for site based educational development. *Sch. Leader. Manag.* 39 (3–4), 315–333.
- Farchi, T., Tubin, D., 2019. Middle leaders in successful and less successful schools. *Sch. Leader. Manag.* 39 (3–4), 372–390.
- Flessa, J.J., 2012. Principals as middle managers: school leadership during the implementation of primary class size reduction policy in Ontario. *Leader. Pol. Sch.* 11 (3), 325–343.
- Fullan, M., 2015. Leadership from the middle. *Educ. Can.* 55 (4), 22–26.
- Grootenboer, P., Edwards-Groves, C., Rönnerman, K., 2015. Leading practice development: voices from the middle. *Prof. Dev. Educ.* 41 (3), 508–526.
- Grootenboer, P., Rönnerman, K., Edwards-Groves, C., 2017. Leading from the middle. A praxis-oriented approach. In: Grootenboer, P., Edwards-Groves, C., Choy, S. (Eds.), *Practice Theory Perspectives on Pedagogy and Education: Praxis, Diversity and Contestation*. Springer Singapore, Singapore, pp. 243–264.
- Grootenboer, P., Edwards-Groves, C., Rönnerman, K., 2019. Understanding middle leadership: practices and policies. *Sch. Leader. Manag.* 39 (3–4), 251–254.
- Grootenboer, P.C., 2018. The Practices of School Middle Leadership. *Leading Professional Learning*. Springer Nature, Singapore.
- Gurr, D., Drysdale, L., 2013. Middle-level school leaders: potential, constraints and implications for leadership preparation. *J. Educ. Adm.* 51 (1), 55–71.
- Gurr, D., 2019. School middle leaders in Australia, Chile and Singapore. *Sch. Leader. Manag.* 39 (3–4), 278–296.
- Hargreaves, A., Ainscow, M., 2015. The top and bottom of leadership and change. *Phi Delta Kappan* 97 (3), 42–48.
- Hargreaves, A., O’Connor, M.T., 2018. *Collaborative Professionalism: When Teaching Together Means Learning for All*. Corwin Press, Thousand Oaks, CA.
- Harris, A., Jones, M., 2017. Middle leaders matter: reflections, recognition, and renaissance. *Sch. Leader. Manag.* 37 (3), 213–216.
- Harris, A., Jones, M., Ismail, N., Nguyen, D., 2019. Middle leaders and middle leadership in schools: exploring the knowledge base (2003–2017). *Sch. Leader. Manag.* 39 (3–4), 255–277.
- Highfield, C., 2012. The Impact of Middle Leadership Practices on Student Academic Outcomes in New Zealand Secondary Schools. Doctoral thesis. The University of Auckland.
- Hord, S.M., Sommers, W., 2008. *Leading Professional Learning Communities: Voices from Research and Practice*. Corwin Press, Thousand Oaks, CA.
- Huerta Villalobos, M., 2018. Educational Leadership and the Contribution of the Technical Pedagogical Head Across Three Types of Schools in One Chilean Region. Doctoral Thesis. The University of Melbourne.

- Keane, W., 2010. Case Studies in Learning Area Leadership in Catholic Secondary Schools in Melbourne. Doctoral thesis. The University of Melbourne.
- Koh, H.H., 2018. The Leadership Role of Middle Leaders in Six Selected Primary Schools in Singapore. Doctoral thesis. The University of Melbourne.
- Kovacević, J., Hallinger, P., 2020. Finding Europe's niche: science mapping the knowledge base on educational leadership and management in Europe, 1960–2018. *Sch. Effect. Sch. Improv.* 31 (3), 405–425.
- Leithwood, K., 2012. The Ontario Leadership Framework 2012 with a Discussion of the Research Foundations. Ontario Institute for Studies in Education, Toronto, Ontario.
- Leithwood, K., 2016. Department-head leadership for school improvement. *Leader. Pol. Sch.* 15 (2), 117–140.
- Lipscombe, K., Tindall-Ford, S., Lamanna, J., 2021. School middle leadership: a systematic review. *Educ. Manag. Adm. Leadersh.* 19.
- Murphy, J., 2005. *Connecting Teacher Leadership and School Improvement*. Sage, New York, NY.
- Ng, S., Chan, T.K., 2014. Continuing professional development for middle leaders in primary schools in Hong Kong. *J. Educ. Adm.* 52 (6), 869–886.
- Nicholas, D., 2019. Distributed Leadership in Successful Schools. Doctoral thesis. The University of Melbourne.
- Reid, D.B., 2020. School principals acting as middle leaders implementing new teacher evaluation systems. *Sch. Leader. Manag.* 40 (1), 88–104.
- Rhodes, C., Brundrett, M., 2009. Growing the leadership talent pool: perceptions of heads, middle leaders and classroom teachers about professional development and leadership succession planning within their own schools. *Prof. Dev. Educ.* 35 (3), 381–398.
- Siskin, L.S., 1991. Departments as different worlds: subject subcultures in secondary schools. *Educ. Adm. Q.* 27 (2), 134–160.
- Siskin, L.S., 1997. The challenge of leadership in comprehensive high schools: school vision and departmental divisions. *Educ. Adm. Q.* 33 (1), 604–623.
- Spillane, J., Zuberi, A., 2009. Designing and piloting a leadership daily practice log: using logs to study the practice of leadership. *Educ. Adm. Q.* 45 (3), 375–423.
- Spillane, J., 2006. *Distributed Leadership*. John Wiley and Sons, San Francisco, CA.
- Supovitz, J.A., D'Auria, J., Spillane, J.P., 2019. Meaningful and Sustainable School Improvement with Distributed Leadership. CPRE Research Reports. Retrieved from: https://repository.upenn.edu/cpre_researchreports/112.
- Tay, H.Y., Tan, K.H.K., Deneen, C.C., Leong, W.S., Fulmer, G.W., Brown, G.T., 2020. Middle leaders' perceptions and actions on assessment: the technical, tactical and ethical. *Sch. Leader. Manag.* 40 (1), 45–63.
- Thorpe, A., Bennett-Powell, G., 2014. The perceptions of secondary school middle leaders regarding their needs following a middle leadership development program. *Manag. Educ.* 28 (2), 52–57.
- Tian, M., Risku, M., Collin, K., 2016. A meta-analysis of distributed leadership from 2002 to 2013: theory development, empirical evidence and future research focus. *Educ. Manag. Adm. Leader* 44 (1), 146–164.
- Turner, C., Bolam, R., 1998. Analyzing the role of the subject head of department in secondary schools in England and Wales: towards a theoretical framework. *Sch. Leader. Manag.* 18 (3), 373–388.
- Wenner, J.A., Campbell, T., 2017. The theoretical and empirical basis of teacher leadership: a review of the literature. *Rev. Educ. Res.* 87 (1), 134–171.
- White, P., 2000. The Leadership of Curriculum Area Middle Managers in Selected Victorian Government Secondary Schools. Doctoral thesis. The University of Melbourne.
- York-Barr, J., Duke, K., 2004. What do we know about teacher leadership? Findings from two decades of scholarship. *Rev. Educ. Res.* 74 (3), 255–316.

System leadership – emergence and maturity

David Hopkins^a and Alma Harris^b, ^a University of Bolton, Hope Valley, Derbyshire, United Kingdom; and ^b Cardiff Metropolitan University, Cardiff, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	123
Part one	124
Defining system leadership	124
System leadership for system transformation	125
The emergence of system leadership – the case of England	127
Global perspectives on system leadership pre-2010	128
Part two	128
A macro and micro level framework for system leadership	128
Macro level	129
Micro level	130
System leader and system leadership – selected country examples	131
Australia	131
Singapore	132
An update on England	132
Ontario, Canada	132
Part three	133
Toward the future and a reflection	133
References	134

Introduction

The concept of ‘system leadership’ is one that over the past two decades has caught the educational imagination. System leaders are those Heads or Principals who are willing to shoulder system wide roles in order to support the improvement of other schools as well as their own. As such, system leadership is a new and emerging practice that embraces a variety of responsibilities that are developing either locally or within discrete Local/State/Regional networks or policy driven initiatives that when taken together have the potential to contribute to system transformation.

The educational policy direction in many developed countries is changing quite dramatically at the present time (Harris and Jones, 2017; Hopkins, 2013; Schleicher, 2018). There is currently a rapid shift away from the government managed educational changes of the 1990s and 2000s to far more decentralised systems based on the principle of ‘autonomy’. This reflects the genuine belief on the part of many politicians and policy makers that there is a need to unleash the power of the profession that has been harnessed in the recent past by too much control. It is within such a context that system leadership assumes increasing importance.

The education leadership field is subject to fads and fashions that quickly come and swiftly go. In sharp contrast, the leadership models that have a sound and extensive empirical base tend to converge around a few core leadership areas (Leithwood et al., 2019). A systematic review of the evidence conducted by Gumus et al. (2018) concluded that distributed leadership, instructional leadership, teacher leadership and transformational leadership continue to be the most researched models of leadership. The review also found that empirical interest in all these leadership models has increased significantly over the past decade further strengthening the collective evidence base.

While system leadership does not feature in the current list of most researched models of leadership, it certainly remains a very popular idea and is still reflected in the policy landscape of many countries. Its presence is still evident in much of the contemporary writing about educational and organisational change.

Since its emergence in the early 2000s, system leadership was viewed as an important system driver and as a means of building the system capacity required for positive change. In 2008, the OECD undertook a significant research program and produced two volumes about system leadership, one of these volumes included a range of detailed country cases (Pont et al., 2008a, 2008b). At this time, prominent academics and influential policy makers endorsed the idea by writing about the processes, practices, and impact of system leadership in action (e.g. Fullan, 2005; Higham et al., 2009; Hopkins, 2009). The term ‘system leadership’ also links to earlier work on systems thinking and has strong connections with systems theory (Senge, 1990).

This chapter outlines the development of the concept of System Leadership and reviews the current international evidence base. In making the case for the potential of system leadership as the key driver for systemic reform the argument in this paper is divided into three key sections:

- **Part 1** outlines the genesis of System leadership pre-2010, in particular:
 - Explores emerging definitions of system leadership
 - Conceptualises its contribution to systemic reform
 - Describes its emergence in a case study of England
 - Reviews its emerging practice by referring to the 2008 OECD School Leadership study.
- **Part 2** reviews the literature post 2010 as the concept of system achieves maturity and is better embedded in policy reforms in many countries
- **Part 3** identifies future challenges.

Part one

Defining system leadership

System leadership, in the sense it is used in this chapter, implies a significantly more substantive engagement with other schools in order to bring about system transformation. Specifically, we define it as a form of leadership where a head teacher or principal is willing and able to shoulder wider system roles and in so doing is almost as concerned with the success and attainment of students in other schools as they are with their own.

Although the concept of 'system leadership' began to catch the educational imagination in the early 2000s it was not initially defined with such clarity. At this time, it was seen more as an emerging role, an instinct or aspiration rather than an operational job description. Take for example these quotations from two significant opinion makers. The first from the General Secretary of the Association of School and College Leaders in England and the other from a leading educational commentator whose work has a global reach.

[John Dunford \(2005\)](#) in an address to the National Conference of the Specialist Schools and Academies Trust (SSAT) argued that:

The greatest challenge on our leadership journey is how we can bring about system improvement. How can we contribute to the raising of standards, not only in our own school, but in others and colleges too? What types of leaders are needed for this task? What style of leadership is required if we are to achieve the sea-change in performance that is demanded of us?

In *Systems Thinkers in Action* [Michael Fullan \(2004\)](#) argued that:

a new kind of leadership is necessary to break through the status quo. Systematic forces, sometimes called inertia, have the upper hand in preventing system shifts. Therefore, it will take powerful, proactive forces to change the existing system (to change context). This can be done directly and indirectly through systems thinking in action. These new theoreticians are leaders who work intensely in their own schools, or national agencies, and at the same time connect with and participate in the bigger picture. To change organizations and systems will require leaders to get experience in linking other parts of the system. These leaders in turn must help develop other leaders within similar characteristics.

These two quotes reflect the more tentative and emerging concept of the role in the early 2000s. The Dunford quote asks about future styles of leadership without proposing any answer to the question. The Fullan quote mentions the importance of leaders engaging in systems thinking but does not allude to the concept of system leaders developed in this chapter.

These quotations do however contain three implicit and important assumptions relevant to the emerging concept of system leadership. The first is that if we are ever to achieve sustainable education change it must be led by those close to the school [this is an explicit claim made by Fullan] the second is that this must have a systemic focus [also an explicit claim by Fullan]; the third is that "system leadership" is an emerging practice [that reflects the challenge implied by Dunford's questions].

Despite this tentativeness, system leadership as a concept has a rich theoretical and research context. The conceptual concerns of system theory for relationships, structures and interdependencies ([Katz and Kahn, 1966](#); [Senge, 1990](#); [Campbell et al., 1994](#)) underpin the contemporary work of system leaders in practice. The key insight here has been well summarised by [Kofman and Senge \(1995:27\)](#) when they state that the "... defining characteristic of a system is that it cannot be understood as a function of its isolated components. ... the system doesn't depend on what each part is doing but on how each part is interacting with the rest".

These demands are further illuminated in theory by [Peter Senge \(1990\)](#), who argues that for organisations to excel, they have to become "learning organisations", which he defines as "organisations where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together" (p. 3). To Senge, the key to becoming a learning organisation is for leaders

to tap into people's commitment and capacity to learn at all levels, to clarify broader systemic interdependencies and how to make them more effective (ibid, p. 4).

This theme, as seen in the quote above, underpins Fullan's (2005) exposition, in a later publication than the previous quote, of the role he believes school leaders will need to play as "system thinkers in action" if sustainable large-scale reform is to be achieved. This, Fullan argues, will necessarily involve adaptive challenges that "require the deep participation of the people with the problem; [and] that is why it is more complex and why it requires more sophisticated leadership" (p. 53). For Fullan, examples of this new work include:

- Leading and facilitating a revolution in pedagogy (p. 57)
- Understanding and changing the culture of a school for the better (p. 57)
- Relating to the broader community, in particular with parents, and integrating and co-ordinating the work of social service agencies into the school as a hub (p. 61)
- This will demand "... above all ... powerful strategies that enable people to question and alter certain values and beliefs as they create new forms of learning within and between schools, and across levels of the system" (ibid, p. 60).

The challenges that Fullan posits represent an ambitious agenda and, as will be seen later, not all "system leaders" take on all these roles. This list also begs the question - Are there enough Heads or Principals with the existing or potential dispositions and capacities, required to achieve such an ambitious set of expectations to lead the huge numbers of existing schools? This is not only an adaptive challenge, but also a logistical one. Our response, as is seen in Part Two of this chapter, is that there is no expectation that every Head or Principal will be a system leader, that would be both unrealistic and unnecessary. What however is both realistic and necessary is that a significant minority of such leaders adopt elements of this role if we are to move toward system transformation.

So, this discussion of system leadership is leading us to three initial conclusions. First the concept of system leadership flows from the general literature on systems theory and thinking. Second system leadership is a theory of action that embraces a range of disciplines in order to exert its power (see for example Elmore, 2004; Leithwood and Riehl, 2005). And third, system leadership will only exert any influence to the extent that it focuses on teaching and learning (i.e. is instructional), shares its authority with others (i.e. is distributed).

These assumptions lead to three other observations that are discussed in detail in this chapter. The first is that there is a tension between system leadership being a national policy or a professional movement. This dichotomy as we will see has profound implications for the prospect of sustainable educational reform. The second is that although system leadership is emerging as a professional practice, it is a concept that as we have seen is located in a rich theoretical and research context. The third observation is that while many school leaders in most jurisdictions are involved in some form of collaborative activity or networking, this is categorically not the same as system leadership. We continue to explore these observations in the following sections of the chapter.

System leadership for system transformation

As has been noted already, the concept of system leadership emerged in the early 2000s, at a time when the long-term utility of large-scale top-down change was being questioned (Fullan, 2011; Hopkins, 2007; Sahlberg, 2012).

There is a general lesson to be drawn for large scale/systemic reform from the global experience in the early 2000s. It is that top-down large-scale reform which has an early narrow focus on key measurable skills can produce a rapid increase in standards at least initially. However, this initial rise in standards will inevitably level off over time if top-down micro-management becomes the default change strategy.

To move beyond this inevitable plateau of achievement requires a system wide school improvement approach that can deliver continuous improvement beyond the early gains. In other words, large scale reform has characteristically focused on short term objectives, whereas systemic change envisages a multi phased process that ensures that early gains do not level off but continue to improve as a consequence of employing strategies that at the same time raise achievement and build capacity.

However, achieving this shift is not straight forward. As Michael Fullan (2003:7) has said, it takes capacity to build capacity, and if there is insufficient capacity to begin with it is folly to announce that a move to 'professionalism' provides the basis of a new approach. The key question is 'how do we get there?', because we cannot simply move from one era to the other without self-consciously building professional capacity throughout the system.

This analysis is informed by the English experiment with large scale reform in the late 90s/early 2000s. The argument goes something like this (Hopkins, 2007):

- Most agreed that standards were too low and too varied in the 1990s and that some form of direct state intervention was necessary.
- However, when progress inevitably plateaued one has to question whether prescription, or more micro-management still guaranteed sustained large-scale reform in the medium term.
- The answer of course is to re-balance top down and bottom up and as competence and confidence increases give more responsibility to schools and system leaders to lead improvement efforts.

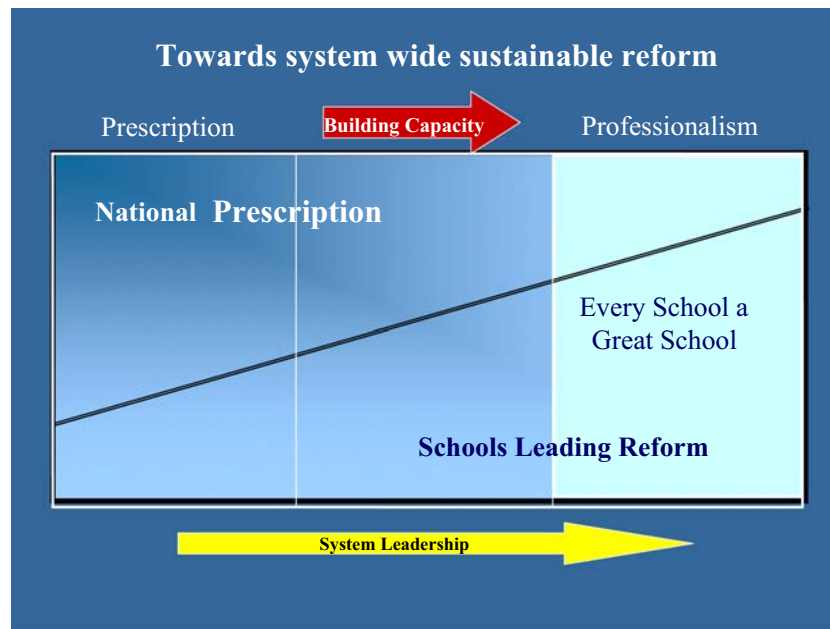


Fig. 1 Toward system wide sustainable reform.

It is this progression that is illustrated in Fig. 1, where system wide reform is driven by system leadership. The intent of the diagram is to indicate that when a school or system is under performing i.e. has low levels of local capacity, then national prescription is necessary. As local capacity increases however, national prescription needs to gradually be replaced with local discretion. This is the re-balancing that is implied as one moves from the left-hand to the right-hand side of the diagram.

It is worth taking a little more time unpacking the thinking underlying the diagram. Five points need making.

The first is to emphasise that this not an argument against 'top down' change. It is clear that neither 'top down' nor 'bottom-up change' work just by themselves, they have to be in balance – in creative tension. The balance between the two at any one time will of course depend on context.

Secondly, it must be realised that in England in 1997 it was obvious that more central direction was needed. This reflects the balance toward national prescription as seen in the left-hand segment of the diagram.

Third it should be no surprise to realise that the right-hand segment is relatively unknown territory. It implies horizontal and lateral ways of working with assumptions and governance arrangements very different from what we know now. The main difficulty in imagining this landscape is that the thinking of most people is constrained by their experiences within the power structure and norms of the left-hand segment of the diagram.

The fourth point is both complex and critical. In terms of the diagram, effective system wide sustainable reform requires a movement from the left to right with all that implies. The left to right movement is necessarily incremental as it builds on, rather than contradicts, the success of previous phases. Yet, and this is the crucial point, the achievement of creating the educational landscape implied by the right-hand segment represents a step change from what has gone before. Yes, the difference between left and right-hand segments represents a radical change or a transformation; but the process or journey from left to right will be incremental, building on past success and re-shaping in light of learning from experience. It is in this way that the language of school improvement (logical incremental steps building on past experience) and transformation (a qualitatively different state from what was known previously) is reconciled.

Finally, of course it is not being suggested that one always has to start from the left-hand side of the diagram and move in some sort of uniform way to the right. That is just how it was in England in 1997. Other systems may well be in the middle and need to move left briefly to firm up certain conditions before rapidly proceeding into the right-hand segment. Some may believe that they are already in the right-hand segment. If this diagram has any value, it is as a heuristic – its purpose is to help people think rather than tell them what to do.

One further point needs to be re-iterated. As noted earlier, the transition from 'prescription' to 'professionalism' is not straight forward. In order to move from one to the other, strategies are required that not only continue to raise standards but also build capacity within the system. This point is key, one cannot just drive to continue to raise standards in an instrumental way, one also needs to develop social, intellectual and organisational capital. This also explains why the concept of "school-based management" by itself is insufficient to affect the transition from the left hand to the right-hand side of the diagram. It is system leadership rather than management that is required.

The emergence of system leadership – the case of England

It was a policy context such as that described above that led to the emergence of System leaders in England in the early 2000s. At this time Ministers deliberately, strategically and boldly embraced the logic of [Fig. 1](#), re-balanced top-down/bottom-up change policy initiatives and embraced system leadership and entered into a new relationship with schools.

In *Every School a Great School* ([Hopkins, 2007](#)) it was suggested that the five striking characteristics of system leaders, those distinguishing them from broader collaborative activity, are that they deploy their experience, knowledge and skills to:

- actively lead improvements in others schools and measure their success in terms of student learning, achievement and welfare, both to raise the bar and narrow the gap(s).
- commit staff in their own and other schools to the improvement of teaching and learning, engaging them deeply in the organization of teaching, learning, curriculum and assessment so as to ensure learning is increasingly personalized for students.
- lead the development of schools as personal and professional learning communities, building relationships across and beyond each school to provide a range of learning experiences and professional development opportunities.
- lead work for equity and inclusion through acting on context and culture, not only in response to poverty but also to employ educational resources to help give communities a greater sense of aspiration and empowerment.
- manage strategically the impact of the classroom, school and system on one another, understanding that in order to change the larger system one has to both engage with it in a meaningful way and manage subsequent change at a school level.

This set of five purposes or goals make a powerful case for system leadership. Taken as a whole, they address most of the issues or problems identified with the challenges introduced in the earlier section on defining system leadership.

It is a compelling proposition that such leadership holds significant potential to contribute to systemic educational improvement. It is an idea that was positively advocated by Government Ministers. David Miliband MP, for example, when Minister of State for School Standards, saw that the development and deployment of cadre of system leaders could go a long way to responding to the key challenges he had identified for school leadership as part of a new relationship between schools and Government ([Miliband, 2004](#)). This included raising productivity in education, effecting greater social justice and ensuring sustainable improvement ([DfES, 2005](#); [Miliband, 2003](#)).

Further research identified a variety of system leader roles emerging in England and elsewhere ([Higham et al., 2009](#)). As it happens this taxonomy has received support from subsequent research as seen in Part 2 of this chapter. The role of the National Leader of Education (NLE) has developed considerably in response to these policy suggestions ([Matthews, 2008](#)); as has the concept of the system leader in a more global context ([Hargreaves, 2012](#)). The initial framework suggested the following roles ([Hopkins and Higham, 2007](#)):

- Developing and *leading a successful educational improvement partnership* between several schools, often focused on a set of specific themes that have significant and clear outcomes that reach beyond the capacity of any one single institution.
- Choosing to *lead and improve a school in extremely challenging circumstances* and change local contexts by building and sustaining a culture of success as a high valued added institution.
- *Partnering another school facing difficulties and improve it*, either as an Executive Head Principal of a Federation or as the leader of a more informal improvement arrangement.
- Acting as a *community leader* to broker and shape partnerships and/or networks of wider relationships across local communities to support children's welfare and potential, often through multi agency work.
- Working as a *change agent* or expert leader within the system, identifying best classroom practice and transferring it to support improvement in other's schools.

No doubt these roles will expand and mature over time. What is particularly interesting about them is how they have evolved in the recent past as a response to the adaptive challenge of system change. It is also important to note that the taxonomy includes system leaders working in either national or state programs as well as locally organised often *ad hoc* roles, a point that is picked up again below.

The underlying assumption of this chapter is that the real prize and potential of system leadership is the realisation of systemic improvement. It is important to appreciate a further insight that the aspiration of system transformation can be facilitated by the degree of segmentation existing in the system. Segmentation implies using the natural variation in school performance within the system as a means of improvement through collaboration ([Hopkins, 2007](#)).

There are two crucial aspects here. First, that there is increased clarity on the nature of intervention and support for schools at each phase of the performance cycle. Second, that schools at each phase are clear as to the most productive ways in which to collaborate in order to capitalise on the diversity within system. A summary of this 'segmentation' approach is set out in [Table 1](#).

In the left-hand column is a basic taxonomy of secondary schools in England produced by the DfES in the mid 2000s based on their phase in the performance cycle. It is this that is the focus of the second column, where a range of strategies for supporting schools at different phases of their development are briefly described.

The key point to consider is that these different forms of intervention and support are increasingly being provided by schools themselves, rather than being imposed and delivered by some external agency. This approach to system transformation relies fundamentally on school-to-school support as the basis of the improvement strategy.

Table 1 The ‘segmentation approach’ to school improvement.

<i>Type of school</i>	<i>Key strategies – responsive to context and need</i>
Leading Schools	Become leading practitioners
Succeeding, self-improving schools	Formal federation with lower-performing schools Regular local networking for school leaders Between-school curriculum development
Succeeding schools with internal variations	Consistency interventions: Such as assessment for learning Subject specialist support to particular curriculum areas
Underperforming schools	Linked school support for underperforming departments/year groups Underperforming pupil programs for catch-up
Low attaining schools	Formal support in Federation structure Consultancy in core subjects and best practice
Failing schools	Intensive support from system leader or Restart with new name and leadership

In this section of the chapter, the skills and potential roles of system leaders have been described as well as their strategic contribution to systemic improvement. This of course has been based in the English context, but the overall strategic approach is generic and can apply in a range of settings and jurisdictions. The skills themselves, the range of improvement categories and the terminology associated with intervention will vary from setting to setting and each system will have to develop its own language and framework suited to its own particular purposes and context. The crucial point being that not all schools are the same, each requires different forms of support and at each stage of development there will be a productive re-balancing between prescription and professionalism.

Global perspectives on system leadership pre-2010

In the mid 2000s the OECD launched a major research and development program into School Leadership in participating countries. The program culminated in the publication of two books *Improving School Leadership*, Volume 1: Policy and Practice (Pont et al., 2008a), and *Improving School Leadership*, Volume 2: Case Studies on System Leadership (Pont et al., 2008b). The authors state that throughout OECD countries, there is a great deal of school leadership co-operation and collaboration going on. Practitioners do not work alone, and many benefit from a variety of networking arrangements or new structures that share management across schools. Although the authors were not able to gauge the extent of their success in most countries because they had not pursued the issue systematically, they did say that most initiatives had explicit objectives that concentrated on the following: sharing and rationalising resources, improving coherence of educational provision, supporting wellbeing and improving educational opportunities and outcomes (Pont et al., 2008a).

In fact, in all countries participating in the OECD Improving School Leadership activity, there were some arrangements for co-operation between schools. School leaders were the key to these and are also strongly influenced by them. In order to give a flavor of this activity Table 2 lists different types of approaches and some of the reasons for co-operation across OECD countries.

To summarise, in this first part of the chapter we have discussed some of the theoretical underpinnings of System Leadership, described a framework for explaining its contribution to systemic reform, provided a case study of emerging practice in England, and reviewed briefly the adoption of collaborative ways of working in twelve countries. In Part 2, the discussion moves to a consideration of System leadership post 2010 as the practice begins to achieve maturity and is better embedded in the policy reforms in many countries.

Part two¹

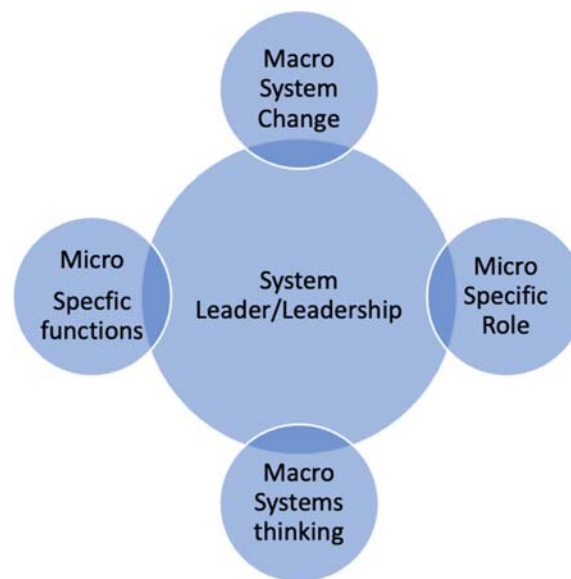
A macro and micro level framework for system leadership

The different interpretations and conceptualisations of system leaders/ship within the international evidence base post 2010 can be separated into *macro* and *micro* level explanations (Harris, 2020). These interpretations sometimes cross-over, making much of the writing on the subject complex to navigate (Boylan, 2016). To make sense of these different interpretations, and the inter-relationships between them, it is helpful to explore system leaders/ship through macro and micro level lenses, respectively. Fig. 2 illustrates the different interpretations and the inter-relationships.

¹Acknowledgment - this section is based upon an unpublished report, on the topic of system leaders and leadership, prepared by one of the authors for the National Academy for Educational Leadership in Wales (Harris, 2020).

Table 2 Different types of approaches to school leader collaboration across OECD countries (Pontet et al., 2008a).

Belgium	- School communities have been created as voluntary collaborative partnerships between schools. They aim to have common staffing, ICT and welfare resources management.
Denmark	- Co-operation in post-compulsory education has been promoted by the creation of administrative groups set up locally or regionally to optimise the joint resources of several self-governing institutions.
Finland	- 2003 legislative reform has enhanced school co-operation aiming to ensure integrity of students' study paths.
France	- "school basins" have been implemented to ensure collaborative partnerships between schools to work together in student orientation, educational coherence between different types of schools, common management of shared material and human resources.
Hungary	- "micro-regional partnerships" based on economic and professional rationalisation have resulted in the spreading of common school maintenance in almost all Hungarian micro regions. This network-type co-operation enables professional and organisational learning leading to new forms of education governance and efficient innovation.
Korea	- small schools cooperate to overcome problems of size in teacher exchange, curriculum organising, joint development activities, and integrated use of facilities.
Netherlands	- In primary education, "upper management" takes management function responsibility for several schools. About 80% of primary school boards have an upper school management bureau for central management, policy staff and support staff.
New Zealand	- school clusters are based around geographical communities and communities of interest.
Norway	- tendency is to merge several schools to form an administrative unit governed by a school principal. Three-level municipalities require networks between schools.
Portugal	- schools are commonly grouped together with a collective management structure; executive, pedagogical and administrative councils are responsible for their areas.
Scotland	- important political promotion of collaboration. "Heads together" is a nationwide online community for sharing leadership experience.
Sweden	- municipal directors of education steer principals. Most of them are members of director of education steering group where strategy, development and results are discussed.
England	- there are different approaches to co-operation stimulated by the government – federations of schools, national leaders of education, school improvement partners ...

**Fig. 2** A Macro and Micro level framework for system leadership.

Macro level

At the macro level, the evidence on system leaders/ship diverges into two major areas of study. Firstly, as was seen earlier, there is a literature on *systems thinking* and *systems theory* that focuses on ways of solving complex problems within and across organisations (e.g. Levin and Schrum, 2013). Secondly, there is a considerable evidence base around system leaders/ship as a policy driver for system change and as the mobilisation of system improvement in different countries and contexts (e.g. Harris et al., 2014; Cousin, 2016; Nasmyth, 2012; Hamilton et al., 2018).

As intimated earlier, systems-thinking, is an approach that is concerned with taking a holistic view of the system. It is a management approach to organisational change that has been applied to schools and connects with closely system science (Shaked and Schechter, 2020). Systems thinking reflects a theoretical orientation on organizational change rather than a practical one. It is viewed as an approach for understanding change within complex systems.

Systems thinking as Fullan (2005) noted places an emphasis upon system thinkers rather than system leaders, reinforcing the capacities of individual leaders within a specific eco-system (Shaked and Schechter, 2014; Toh et al., 2014). Systems thinking is also linked to idea of systemic leadership that appears within the business literature (e.g. Thompson and Zakhirova, 2019) which views organisations as open systems that are visible and inter-connected (Beehner, 2020).

In the second area of study, system leaders/ship is represented and understood as a *policy lever* for school and system improvement. In this respect, system leadership is presented as an adaptive leadership style or disposition that is based on an understanding of the complexity of educational contexts as involving multiple, interrelated systems. System leadership at the macro level focuses on how leadership is exercised for the greater good and how *capacity* is built to deliver system wide change and improvement.

The international literature reflects an extensive evidential base around system leaders/ship as a policy driver for system improvement. In Canada, Finland, Georgia, England, Hong Kong, New York, Latin America, Singapore, Scotland, Australia and New Zealand there are specific examples of policy interventions, around system leader/ship, directly aimed at building the capacity for school/system improvement (Mitchell and Sackney, 2016; Sentocnick et al., 2016; Elwick, 2017; Weinstein and Hernández, 2015; Hamilton et al., 2018; Hopkins 2016; Mowat, 2018; McNaughton et al., 2011; Watterston et al., 2011; Dimmock, 2016; Simon, 2015; Smeds-Nylund et al., 2016; Southworth and Quesnay 2005; Szeto, 2019; Ng, 2017).

System leadership is essentially defined, at the macro level, as having a systemic orientation where system leaders influence those beyond their immediate context and are deliberately deployed to bring about large-scale change. This interpretation has been seen in policy developments in England (Cousin, 2019; Close, 2016) and in many other countries (Pont et al., 2008c).

The implementation of a system model, it has been argued, happens when all stakeholders are active co-collaborators in the modeling of solution seeking processes (Wilkinson, 2016). Within the international evidence base, there are frequent references to system leaders as solution seekers and change agents, as levers for system level change and improvement.

The academic literature similarly reinforces the centrality of system leaders/ship as a macro lever for school and system wide change. For example, Cousin (2018) views system leaders/ship as a public service delivery mechanism within a narrative of global governance. Focusing on England, Cousin (2018) outlines findings from a longitudinal study that demonstrates a positive relationship between system leaders (National Leaders of Education) and school/system improvement. Brown and Greany (2018) also turn their attention to the English education system but offer a more critical view of the evidence about system leaders/ship and its relationship with the self-improving system. Other writers offer examples of system leaders/ship in various settings (e.g. Naicker, 2015; Bryant, 2018; Godfrey, 2017) each reinforcing that system leadership is a macro policy lever and a catalyst for system wide change (Hamilton et al., 2018).

In the macro sense, therefore, system leaders are collectively charged with the development of others and the improvement of the system (Sheppard and Dibbon, 2011; Hopkins et al., 2014) where their sphere of influence varies according to role, responsibility, and level.

Micro level

Moving to the micro-level, re-positions the idea of system leaders/ship away from the collective spheres of influence toward individual functions, roles, traits, or characteristics (Accardi, 2016; Dimmock, 2016). In his work, Hopkins (2009) suggests that system leaders operate at *three levels* –.

1. The level of the school i.e. helping other schools.
2. The local level i.e. working with networks of schools or at the national level.
3. The national/system level i.e. operating as a leader of system change (e.g a National Leader of Education) (Fig. 3).

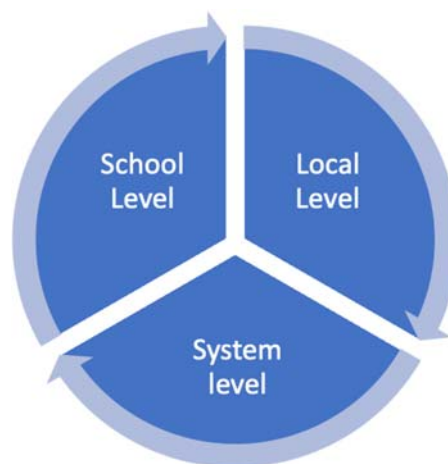


Fig. 3 The three levels of System Leadership.

Hopkins (2009) further suggests that all these system leader roles are necessary and operate interdependently to secure system level change. This tripartite distinction suggests that system leaders need to understand the micro and macro influences in the system to be most effective. Heifetz and Linsky (2002) suggest that system leaders must exist on ‘the balcony and the dancefloor’. In summary, system leaders not only need to see the bigger picture, but they also need to understand the detail of how to secure change.

The international evidence is replete with accounts of the work of different categories of system leaders e.g. headteacher, principal, district leader, network leader, middle leader or teacher leader (Robinson, 2012; Senge et al., 2015; Cousin, 2019; Boylan, 2016). This evidence base proposes that those occupying existing *leadership roles* within the system are, by default, system leaders with the potential to exercise leadership throughout the system. Simkins et al. (2019) focus on primary heads and argue that through their leadership within the system it is possible to have a vision of a school led system. Spillane et al. (2019) focuses on six school systems in the USA and argues that school leaders are system leaders who affect change and improvement through the processes of sense making.

In their work, Senge et al. (2015) talk about core capabilities for system leader/ship that include the ability to see the needs of the larger system, fostering reflection, building collective engagement, and not being reactive but shaping the future. Earley (2016) similarly offers a series of expectations of system leaders but these tend toward the aspirational rather than the practical.

To summarise, the contemporary evidence base views system leaders/ship as a macro level policy driver and a micro level role or responsibility. The literature reflects three distinctive but interrelated interpretations:

- system leaders/ship *as* system change – e.g. in Singapore, leaders within the system (school, district, and Ministry) collectively operate *as* system change agents (Ng, 2017).
- system leaders/ship *for* system change -e.g. in England, system leaders, (National Leaders of Education) are responsible *for* leading system change (Cousin, 2018)
- system leaders/ship *through* system change - e.g. in Ontario, system change produced literacy co-ordinators and champions who became system leaders *as a direct result of* system wide change (Gallagher et al., 2016)

The next section aims to illuminate these different interpretations further, by offering country specific examples of system leader/leadership.

System leader and system leadership – selected country examples

As seen previously, over a decade ago, the OECD (Pont et al., 2008) highlighted several country cases where system leader/ship was present and prevalent including those summarised here. The short examples that follow (Australia, Singapore, England, and Canada) offer contemporary insights into how system leadership is currently interpreted, understood, and enacted in different contexts. Each of the examples is informed by policy evidence, collected as part of this review, and recent communication with identified country experts.

Australia

To begin with, two experts from Australia offer some important insights into system leaders/leadership in their context:

In Australia, system leadership varies by school type. For government schools, the state or territory Department of Education will have the most system leadership influence and this will have central and regional components. Yet, principals and other school leaders have an equally important role in school success. For Catholic schools, the main system leadership will come from the local diocese. For Independent schools it will vary according to whether they are part of a system or not. As well as this, governance at the school level varies greatly and this will impact on the influence of systems and the work of principals (Gurr, 2020²).

In NSW, the term system leadership, is used to describe the role of senior educators who occupy formal positions that have the responsibility of supporting the effectiveness of schools and school leaders. In NSW, the term system leadership is also used to describe educators who voluntarily undertake roles, with the purpose of contributing to the development of other educators beyond those within their own school. These roles generally include, supporting the professional learning of others as an in-program facilitator, or leading a local network with a focus on an area of educational interest such as curriculum, ICT, or leadership development. Current research regarding in-program facilitators, indicates that the key drivers for educators to undertake these roles include-

- *a commitment to the profession and the well-being of the public education system,*
- *a desire to engage in professional learning through supporting the learning of others,*
- *a desire to network with, and expand their sphere of influence, beyond the school* (McIntyre, 2020³).

²Personal Communication.

³Personal Communication.

Singapore

In Singapore, system leadership is *the* central way that all system change is delivered. School leaders, at all levels, are prepared from day one to take up responsibilities anywhere in the education system, including the Ministry. The National Institute of Education (NIE) trains all teachers and leaders, throughout the country, to a consistently high standard. It prepares them to develop other leaders at all levels in the system. When a teacher is identified as being capable of a leadership responsibility, the final beneficiary is often another school. When the teacher is ready for a leadership position, they may be identified (by the Ministry of Education) to be posted to another school that needs them more or where a better opportunity for further development is available.

In Singapore, teachers and principals are allocated to their posts by the Ministry, it is not their personal choice. Many capable heads of department, who have been in one school, become vice-principals in another school. Similarly, many capable vice-principals, who have worked extensively in one school, will eventually leave, and lead another school. Many highly effective principals join the district or the Ministry level.

In this way leadership positions and responsibilities are rotated in the system. So, in Singapore, school leaders know that they are developing people for the system, rather than just for one school. This is a system high in trust and collaboration. School leaders also trust the system to provide the human resources the school needs (Ng (2017).

In Singapore, school leaders do not just lead a school. They are called to lead 'nationally'. That means that a school leader belongs to a community of leaders. School leaders care for the whole education system, not just the school each of them is currently leading. They must take a broader view and consider how their decisions in school will affect other schools or even the whole nation. Social capital in Singapore is 'nationalistic'. This is the spirit of competitive collaboration. Centralized decentralization can work because the spirit of competitive collaboration helps to drive improvement while keeping the system united (Ng, 2020⁴).

An update on England

In England, the concept of system leadership continues to mature and the 'self-improving school-led system' (SISS) is now the overarching narrative for education policy. This agenda has encompassed an ensemble of reforms since 2010 including academies, the promotion of multi-academy trusts (MATs), the roll back of local authorities (LAs) from school oversight, and the development of new school-to-school support models, such as Teaching School Alliances (TSAs). The government's rationale for this mix of reforms is, ostensibly, to move control to the frontline by giving greater autonomy to schools. To support the SISS, a tranche of National Leaders of Education (NLE's) currently operates to mobilise change across the system and to act as system leaders for school self-improvement. NLE's tend to be high performing leaders and their teams (e.g. as National Leaders of Education or Teaching Schools) who are designated and deployed to support other schools (Earley and Greany, 2017).

Prior to 2010, there was a broad interpretation of system leadership in England (Hopkins and Higham, 2007). Any school leader who played a role beyond their context could consider themselves as engaging in 'system leadership'. Latterly, however, system leaders are formally designated as NLE's with clear lines of accountability to government. The primary function of the NLEs is to deliver school improvement support on behalf of the Department for Education in England (DfE, 2021). The core purpose of the NLE is to deliver support to underperforming schools, helping them to identify and address areas in need of improvement.

Within England tensions exist, however, within the NLE system leadership role. On the one hand, system leaders (NLEs) are encouraged and expected by government to drive relentless and rapid improvement across multiple schools. On the other hand, they are expected simultaneously to address other 'school-led' policy priorities, such as to expand the numbers of trainees on school-led Initial Teacher Training routes. NLEs are charged with working on behalf of the self-improving system, engaging their peers in building collaborative alliances while operating in a competitive marketplace for school improvement services (Greany, 2018).

Greany and Higham (2018) in *Hierarchy, Markets and Networks* highlighted that NLEs in England were experiencing conflicting and unreasonable demands from central government while being perceived by their peers to be a 'co-opted elite', working as part of the managerial state and accruing a range of personal and organizational benefits as a result. Concerns about the NLE role have also been raised because of the importance placed on the Ofsted 'Outstanding grade' to designate those selected as system leaders.

System leaders all identified benefits from their roles, but there are also a series of challenges, including significant pressure to make short-term improvements in other schools and the fear that their own school might drop in performance as a result of their external work

Greany (2020)⁵.

Ontario, Canada

Since 2007, the Province of Ontario in Canada has successfully embarked on a major reform process, largely but not exclusively focusing on improving literacy and numeracy outcomes for young people, particularly those who are most disadvantaged

⁴Personal Communication.

⁵Personal Communication.

(Campbell, 2020). Ontario's educational reform agenda has concentrated consistently on raising the bar and narrowing the achievement gap and on increasing the rates of high school completion.

An important and prominent part of this reform agenda has been the focus on system leadership across the Province. A by-product of system level reform in Ontario has been the attention paid to building professional capacity and leadership capacity at all levels of the education system (province, district, school, and classroom). A central element of this focus has been developing leaders and leadership practices at all levels of the education system. In this respect, system leadership in Ontario has emerged through large-scale system level change. The following two quotes from *Our Call to Action* (National Academy for Educational Leadership Wales 2019, p.11) offer insights into the role of the system leader in Ontario:

One of the strengths of the system is a comprehensive, well-established Ontario Leadership Framework (OLF) providing valued support for professional learning, into roles as centrally assigned principals and beyond as superintendents. This provides a clear pathway for developing future leaders in the system. Principals and vice-principals are allocated from a central pool to schools by the Toronto Board of Education with the expectation that they remain in one school for approximately five years

Nevertheless, there is a compromise to be made, as the expectation to move schools can affect the well-being of principals. However, principals are well supported in their roles by centrally assigned principals who are all experienced leaders. This leads to a high level of trust and mutual respect, which in turn creates an ethos of support rather than challenge in the system

All the country examples reflect the fact that system leadership within and across educational organisations is still considered important to build the necessary professional capital and professional capacity for improvement. As this review of post 2010 practice has highlighted, there is evidence that when properly deployed, system leadership can be a positive force for system change and transformation.

Part three

Toward the future and a reflection

The substance of the previous two parts of this chapter suggests that the concept of system leadership is an idea whose time has come. One can characterise contemporary practice by saying that system leadership is increasingly being seen as:

- A wider resource for school improvement, making more of the most successful school leaders' by encouraging and enabling them to identify and transfer best practice, reduce the risk of innovation in other schools and lead partnerships that improve and diversify educational pathways for students within and across localities.
- A more authentic response to low attaining schools. Strong leadership is vital to turn these schools around. However, a central challenge is that these schools are often the least able to attract suitable leaders. The most successful heads hold the potential to impact on these schools, which need their expertise, by working to develop and mobilize leadership capacity in the pursuit of whole school improvement.
- A potential means to resolve, in the longer term, the emerging and related challenges of a declining demographic supply of well-qualified school leaders, falling student rolls and hence increasingly non-viable schools, and yet ongoing pressures to sustain educational provision in all localities. The evidence cited here suggests that system leadership solutions may include fewer head teachers across some groups of schools, new challenges and incentives for the retention of the most experienced head teachers, as well as new development opportunities for deputies and middle leaders to experience aspects of headship at first hand before taking on full head teacher responsibilities.

Before we get too carried away with such enthusiasm for the concept of system leadership, however, we need to admit that much of the well-intentioned advocacy is based on aspiration and individual cases rather than systemic evidence. Despite its attractiveness as a new catalyst for system change there are a range of challenges to be faced up to if we are to avoid a simplified and uncritical approach to system leadership. We have several concerns:

- First, there is currently no clear or systemic knowledge of how leaders undertake system leadership roles.
- Second, there is insufficient evidence of effectiveness including on the proportion of system leaders who are successful, their impact in different contexts and what might constitute best practice.
- Third, there is no certainty that sufficient leadership quality and crucially whole school capacity already exists to provide a platform on which to develop a wide range of effective and sustainable system leadership roles.
- Fourth, it is not always clear how capacity is renewed within the schools from which system leaders work, especially those in challenging circumstances.
- Fifth, there is no consensus about how best (or who is best placed) to deploy and develop a range of existing and aspirant system leaders to ensure that they gain the skills, experience and support needed to be effective.

- And sixth, there is a wider debate within the head teacher profession about how to reconcile on the one hand an impetus for collaboration between schools with on the other hand an accountability system focused on individual schools (which has the potential to act as an impetus for competition). This perceived tension certainly does not circumscribe nascent system leadership, but it is an important contextual dimension when wider systemic roles are being considered or developed.

In reflecting on these concerns however, there is an overarching challenge to face. It is the distinction between system leaders working in Governmental or State programmes and those working in locally organised often *ad hoc* roles.

As seen in **Part Two**, the majority of system leaders increasingly operate in Governmental or Regional programs that have incentivised activity through organisation, funding and professional development - this is the 'enabling state' at work. It is an important strategy for encouraging Heads and Principals to lead technical and adaptive solutions in a widening professional domain of cross-school and system improvement. It is a phenomenon that is increasingly being seen in those school systems that are accelerating up the PISA international benchmarking scales.

More freedom exists on the other side of the divide, in the roles that are locally developed, often *ad hoc* and contextually responsive. It is understandable why many conceive of these roles to be a more authentic form of system leadership. With no single framework or protocol, a range of models are developed in relation to specific needs and times. Furthermore, from this perspective, the role of an 'enabling state' becomes focused on reducing both barriers to collaboration and wider policy disincentives at all levels. On the other hand, it encourages agencies to provide specific support in networking and bespoke professional development to individual system leaders.

There are of course variations to this bottom up/top-down distinction that are beyond the scope of this chapter and will be dependent on a range of criteria. If, however, a shared criterion is to develop effective system leadership in a growing number of schools, then the following suggestions for action may prove instructive.

Suggestion One: Incentivise rather than legislate. The traditional response has been intervention and management from Governmental, State and Federal Agencies. The argument here is that this leadership now needs to come more from Heads and Principals themselves or from agencies committed to working with them in authentic ways. It is clear that the more bureaucratic the response, the less likely it will be to work. A more lateral approach may be to create the conditions within the system to promote system leadership and collaborative activity through for example, adjusting accountability requirements, and funding for capacity building. With the right incentives in place schools will naturally move toward these new ways of working and mold them to the context in which they operate and to the challenges they face.

Suggestion Two: Place the agency close to the school. There are now in many systems, system leadership roles whose remit is specifically school improvement. The intention that must be maintained is that instead of creating a new bureaucracy, their brief is increasingly focused on facilitating relationships between schools to maximise the potential of purposive collaboration. This approach to school transformation is made increasingly possible by the range of sophisticated data potentially available on school and student performance. It enables groups of schools to identify (a) issues where they shared both strengths and weaknesses i.e. their capacity for sharing, and (b) common issues where they are likely to need some external input.

Suggestion Three: Use school 'independence' collaboratively to tackle underperformance. The underlying assumption here is that autonomous schools working collaboratively is a particularly appropriate organisational format for contexts where rapid transformation of standards and support for students are most needed. The key point is that the freedoms associated with increased autonomy can be used to promote collaboration and inclusion to directly address the needs of students. The crucial condition is that all schools accept responsibility for the education of all the students within their geographic area.

The purpose of this chapter has been to demonstrate that system leadership represents a powerful combination of practices that give us a glimpse of the crucial importance of leadership in the new educational landscape. The collective sharing of skills, expertise and experience creates much richer and more sustainable opportunities for rigorous transformation than can ever be provided by isolated institutions. Realising this landscape, however, may also require a bigger shift within the broader education system, in particular, by giving school leaders more agency to take the lead – in short to light their own fires. The future is certainly theirs.

References

- Accardi, I., 2016. The Role of System Leaders in Promoting Teacher Leadership in Small Schools. PhD. Esteves School of Education, Russell Sage College.
- Beehner, C.G., 2020. System Leadership for Sustainability. Routledge, New York.
- Boylan, M., 2016. Deepening system leadership- teachers leading from below. *Educ. Manag. Adm. Leader* 44 (1), 57–72. <https://doi.org/10.1177/1741143213501314>.
- Brown, C., Greany, T., 2018. The evidence-informed school system in England: where should school leaders Be focusing their efforts? *Leader. Pol. Sch.* 17 (1), 115–137. <https://doi.org/10.1080/15700763.2016.1270330>.
- Bryant, D.A., 2018. Conditions that support middle leaders' work in organisational and system leadership: Hong Kong case studies. *Sch. Leader. Manag.* 39 (5), 415–433. <https://doi.org/10.1080/13632434.2018.1489790>.
- Campbell, C., 2020. Educational equity in Canada: the case of Ontario's strategies and actions to advance excellence and equity for students. *Sch. Leader. Manag.* 41 (4–5), 409–428. <https://doi.org/10.1080/13632434.2019.1709165>.
- Campbell, D., Coldicott, T., Kinsella, K., 1994. Systemic Work with Organisations. A New Model for Managers and Change Agents. Karnac Books, London.
- Close, P., 2016. 'System leader' consultancy development in English schools: a long-term agenda for a democratic future? *Sch. Leader. Manag.* 36 (1), 113–132. <https://doi.org/10.1080/13632434.2016.1160214>.
- Cousin, S., 2016. System Leadership: Are System Leaders in the English Education System Leaders of a Profession Or Instruments of Government? (PhD). UCL (University College London, London).

- Cousin, S., 2018. System leaders in England: leaders of a profession, or instruments of government? *Educ. Manag. Adm. Leader* 47 (4), 520–537. <https://doi.org/10.1177/1741143217751726>.
- Cousin, S., 2019. *System Leadership: Policy and Practice in the English Schools System*, first ed. Bloomsbury Academic.
- DfES, 2005. *Education Improvement Partnerships: Local Collaboration for School Improvement and Better Service Delivery*. DfES, London.
- Department for Education, 2021. *Get School-To-School Support from System Leaders*. DfE, London. <https://www.gov.uk/guidance/system-leaders-who-they-are-and-what-they-do>.
- Dimmock, C., 2016. System leadership for school improvement: a developing concept and set of practices. *Scot. Educ. Rev.* 48 (2), 60–79.
- Dunford, J., 2005. Watering the plants: leading schools and improving the system. In: *Address to the National Conference of the Specialist Schools and Academies Trust*.
- Earley, P., 2016. Global trends and challenges for school leaders: keeping the focus on learning. *J. Educ. Cultur. Psychol. Stud.* 14, 21–33. <https://doi.org/10.7358/ecps-2016-014-earl>.
- Earley, P., Greany, T., 2017. *School Leadership and Education System Reform*. Bloomsbury Publishing Plc, London.
- Elmore, R., 2004. *School Reform from the Inside Out: Policy, Practice, and Performance*. Harvard Educational Press, Cambridge.
- Elwick, A., 2017. Education reform in New York city (2002–2013). *Oxf. Rev. Educ.* 43 (6), 677–694. <https://doi.org/10.1080/03054985.2017.1296421>.
- Fullan, M., 2003. *The Moral Imperative of School Leadership*. Corwin Press, London.
- Fullan, M., 2004. *Systems Thinkers in Action: Moving beyond the Standards Plateau*. DfES Innovation Unit/NCSL, London/Nottingham.
- Fullan, M., 2005. *Leadership & Sustainability: System Thinkers in Action*. Corwin Press.
- Fullan, M., 2011. *Choosing the Wrong Drivers for Whole System Reform*. Seminar Series Paper no. 204. Centre for Strategic Education, Melbourne.
- Gallagher, M.J., Malloy, J., Ryerson, R., 2016. Achieving excellence- Bringing effective literacy pedagogy to scale in Ontario's publicly-funded education system. *J. Educ. Change* 17 (4), 477–504.
- Godfrey, D., 2017. What is the proposed role of research evidence in England's 'self-improving' school system? *Oxf. Rev. Educ.* 43 (4), 433–446. <https://doi.org/10.1080/03054985.2017.1329718>.
- Greany, T., 2018. Innovation is possible, it's just not easy: improvement, innovation and legitimacy in England's autonomous and accountable school system. *Educ. Manag. Adm. Leader* 46 (1), 65–85.
- Greany, Higham, 2018. *Hierarchy, Markets and Networks*. UCL Institute of Education Press, London. <https://www.ucl-ioe-press.com/ioe-content/uploads/2018/08/Hierarchy-Markets-and-Networks.pdf>.
- Gumus, S., Bellibas, M.S., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48.
- Hamilton, G., Forde, C., McMahon, M., 2018. Developing a coherent strategy to build leadership capacity in Scottish education. *Manag. Educ.* 32 (2), 72–78. <https://doi.org/10.1177/0892020618762715>.
- Hargreaves, D.H., 2012. *A Self-Improving School System in International Context*. National College for School Leadership, Nottingham.
- Harris, A., 2020. *System Leaders and System Leadership: Exploring the Contemporary Evidence Base* (Un-published report prepared for the National Academy for Educational Leadership in Wales).
- Harris, A., Jones, M.S., Adams, D., Perera, C.J., Sharma, S., 2014. High-performing education systems in Asia: leadership art meets implementation science. *Asia-Pacific Educ. Res.* 23 (4), 861–869. <https://doi.org/10.1007/s40299-014-0209-y>.
- Harris, A., Jones, M., 2017. Leading educational change and improvement at scale: some inconvenient truths about system performance. *Int. J. Leader. Educ.* 20 (5), 632–641.
- Heifetz, R.A., Linsky, M., 2002. A survival guide for leaders. *Harv. Bus. Rev.* 80 (6), 65–74.
- Higham, R., Hopkins, D., Matthews, P., 2009. *System Leadership in Practice*. Open University Press/McGraw Hill, Maidenhead, Berkshire.
- Hopkins, D., 2007. *Every School a Great School*. McGraw-Hill/Open University Press, Maidenhead.
- Hopkins, D., 2009. *The Emergence of System Leadership*. NCSL, Nottingham.
- Hopkins, D., 2013. *Exploding the Myths of School Reform*. Open University Press, McGraw Hill Education, Berkshire.
- Hopkins, D., 2016. School and system reform—an agenda for Wales. *Cylchgrawn Addysg Cymru/Wales J. Educ.* 18 (1), 87–110.
- Hopkins, D., Higham, R., 2007. System leadership: mapping the landscape. *Sch. Leader. Manag.* 27 (2), 147–166.
- Hopkins, D., Stringfield, S., Harris, A., Stoll, L., Mackay, T., 2014. School and system improvement: a narrative state-of-the-art review. *Sch. Effect. Sch. Improv.* 25 (2), 257–281. <https://doi.org/10.1080/09243453.2014.885452>.
- Katz, D., Kahn, R.L., 1966. *The Social Psychology of Organizations*. Wiley, New York.
- Kofman, F., Senge, P.M., 1995. Communities of commitment: the heart of learning organizations. In: Chawla, Renesch's (Eds.), *Learning Organizations: Developing Cultures for Tomorrow's Workplace*. Productivity Press, Oregon.
- Leithwood, K., Riehl, C., 2005. What we know about successful school leadership. In: Firestone, W., Riehl, C. (Eds.), *A New Agenda; Directions for Research on Educational Leadership*. Teachers College Press, New York.
- Leithwood, K., Harris, A., Hopkins, D., 2019. Seven Strong Claims about Successful School Leadership Revisited. *School Leadership & Management*. <https://doi.org/10.1080/13632434.2019.1596077>.
- Levin, B.B., Schrum, L., 2013. Using systems thinking to leverage technology for school improvement. *J. Res. Technol. Educ.* 46 (1), 29–51. <https://doi.org/10.1080/15391523.2013.10782612>.
- Matthews, P., 2008. Attributes of the first National Leaders of Education in England: what do they bring to the role? *Res. Rep. Commission. Nation. Coll. School Leader*.
- McNaughton, S., Robinson, V.M.J., Timperley, H., 2011. Building capacity in a self-managing schooling system- the New Zealand experience. *J. Educ. Adm.* 49 (6), 720–738.
- Miliband, D., 2003. School leadership: the productivity challenge. In: *Address at the National College for School Leadership*, October 2003.
- Miliband, D., 2004. A new relationship with schools. In: *Address at the North of England Conference*, January 2004.
- Mitchell, C., Sackney, L., 2016. School improvement in high-capacity schools. *Educ. Manag. Adm. Leader* 44 (5), 853–868. <https://doi.org/10.1177/1741143214564772>.
- Mowat, J.G., 2018. 'Closing the gap': systems leadership is no leadership at all without a moral compass – a Scottish perspective. *Sch. Leader. Manag.* 39 (1), 48–75. <https://doi.org/10.1080/13632434.2018.1447457>.
- Naicker, S.R., 2015. The role of the school district in promoting system leadership. In: *Paper Presented at the 8th International Conference of Education, Research and Innovation (ICERI)*, Seville, SPAIN, NOV 16-20, 2015.
- Nasmyth, G., 2012. *Leadership as a System-wide Construct: A Narrative Exploration of the Collaborative Process*. PhD. Fielding Graduate University.
- National Academy for Educational Leadership Wales, 2019. *Our Call to Action. Inspiring Leaders – Enriching Lives*. National Academy for Educational Leadership Wales, Swansea, Wales.
- Ng, P.T., 2017. *Learning from Singapore: The Power of Paradoxes*. Routledge, New York.
- Pont, B., Nusche, D., Moorman, H., 2008a. Improving school leadership. In: *Policy and Practice*, vol. 1. OECD Publications, Paris.
- Pont, B., Nusche, D., Hopkins, D., 2008b. Improving school leadership. In: *Case Studies on System Leadership*, vol. 1. OECD Publications, Paris.
- Pont, B., Nusche, D., Hopkins, D., 2008c. *Improving School Leadership*, Vol 2. OECD.
- Robinson, S., 2012. *School and System Leadership: Changing Roles for Primary Headteachers*, first ed. Bloomsbury publishing Plc, London.
- Sahlberg, P., 2012. *Global Reform Is Here!*. Available at: <https://parsisahlberg/global-educational-reform-movement-is-here!>.
- Schleicher, A., 2018. *World Class: How to Build a 21st-Century School System*. OECD Publications, Paris.
- Senge, P.M., 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency, New York.
- Senge, P., Hamilton, H., Winter, J.K., 2015. The Dawn of System Leadership. Retrieved 17/04/2020, from: https://ssir.org/articles/entry/the_dawn_of_system_leadership.

- Sentočnik, S., Sales, G.C., Richardson, J.W., 2016. Educational reform in Georgia: recommendations for building leadership capacity as a critical success factor for instructional change. *Int. J. Leader. Educ.* 1–18. <https://doi.org/10.1080/13603124.2016.1224931>.
- Shaked, H., Schechter, C., 2014. Systems school leadership: exploring an emerging construct. *J. Educ. Adm.* 52 (6), 792–811. <https://doi.org/10.1108/jea-07-2013-0081>.
- Shaked, H., Schechter, C., 2020. Systems thinking leadership: new explorations for school improvement. *Manag. Educ.* <https://doi.org/10.1177/0892020620907327>, 0892020620907327.
- Sheppard, B., Dibbon, D., 2011. Improving the capacity of school system leaders and teachers to design productive learning environments. *Leader. Pol. Sch.* 10 (2), 125–144. <https://doi.org/10.1080/15700763.2010.502610>.
- Simkins, T., Coldron, J., Crawford, M., Maxwell, B., 2019. Emerging schooling landscapes in England: how primary system leaders are responding to new school groupings. *Educ. Manag. Adm. Leader* 47 (3), 331–348.
- Simon, C.A., 2015. Are two heads better than one? System school leadership explained and critiqued. *Sch. Leader. Manag.* 35 (5), 544–558. <https://doi.org/10.1080/13632434.2015.1107035>.
- Smeds-Nylund, A.-S., Sundqvist, R., Uljens, M., 2016. Educational leadership for sustained multi-level school development in Finland: A non-affirmative approach. *Nordic Stud. Educ.* 36 (02), 103–124. <https://doi.org/10.18261/issn.1891-5949-2016-02-03>.
- Southworth, G., Quesnay, H.D., 2005. School leadership and system leadership. *Educ. Forum* 69 (2), 212–220. <https://doi.org/10.1080/00131720508984685>.
- Spillane, J.P., Seelig, J.L., Blaushild, N.L., Cohen, D.K., Peurach, D.J., 2019. Educational system building in a changing educational sector: environment, organization, and the technical core. *Educ. Pol.* 33 (6), 846–881. <https://doi.org/10.1177/0895904819866269>.
- Szeto, E., 2019. School Leadership in the Reforms of the Hong Kong Education System: Insights into School-Based Development in Policy Borrowing and Indigenising. *School Leadership & Management*, pp. 1–22. <https://doi.org/10.1080/13632434.2019.1616172>.
- Thompson, W.R., Zakhirova, L., 2019. *Racing to the Top: How Energy Fuels System Leadership in World Politics*. Oxford University Press, New York.
- Toh, Y., Jamaludin, A., Hung, W.L.D., Chua, P.M.-H., 2014. Ecological leadership: going beyond system leadership for diffusing school-based innovations in the crucible of change for 21st century learning. *Asia-Pacific Educ. Res.* 23 (4), 835–850. <https://doi.org/10.1007/s40299-014-0211-4>.
- Watterston, J., Dinham, S., Caldwell, B., 2011. System alignment as a key strategy in building capacity for school transformation. *J. Educ. Adm.* 49 (6), 637–652. <https://doi.org/10.1108/09578231111174794>.
- Weinstein, J., Hernández, M., 2015. Birth pains: emerging school leadership policies in eight school systems of Latin America. *Int. J. Leader. Educ.* 19 (3), 241–263. <https://doi.org/10.1080/13603124.2015.1020344>.
- Wilkinson, S., 2016. *The Qualities and Practices of Effective System Leaders: Primary Headteachers*. PhD. Institute of Education, UCL.

The ethical imperative in educational leadership

Christopher M. Branson, Australian Catholic University, Virginia, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	137
Why ethical educational leadership?	138
Ethical leadership defined	139
Ethical educational leadership practice	140
Why is the issue persisting?	141
Personal competence (self-knowledge)	142
Social competence (connectedness)	142
Organisational competence (transcendence)	142
Personal integrity (ethicalness)	143
Social integrity (trustworthiness)	143
Organisational integrity (responsibility)	143
Personal authenticity (autonomy)	143
Social authenticity (presence)	143
Organisational authenticity (wise judgment)	144
Summation - where to from here?	144
References	145

Introduction

"If it isn't ethical, it isn't leadership" was the title of the final chapter I wrote for the *Handbook of Ethical Educational Leadership* that I co-edited with Steven Gross (Branson and Gross, 2014). Here the term, ethical, implies that a leader's actions are founded upon recognisable principles or values that automatically embrace the common good. Essentially, this particular chapter challenged the widely held view that a person appointed to a position in which they are labeled a leader has the sole right to choose how they will lead. In the opinion of such individuals, leadership is incorrectly assumed to be any activity performed by a person in a leadership position. As a result, the world has witnessed brutally inhumane actions being justified or given credibility by being labeled as leadership practices by persons in influential positions. Arguably however, the wide-ranging influence of Burns' (1978) seminal and Pulitzer Prize-winning text, *Leadership*, in which he described the concept of transformational leadership, as distinct from transactional leadership, initiated a challenge to this potentially precarious autocratic perspective.

According to Burns (1978), the previous leadership theories were essentially transactional in nature because they afforded total control to the person in the leadership position. Such transactional leadership theories concentrated almost solely on what the leader had to do or provide to those they were supposedly leading in exchange for their loyalty and commitment toward maximizing their work quality and individual output so as to continually benefit the organisation's purpose. In contrast, Burns argued that leadership practice would be better served if it concentrated on transforming those being led. In order to transform the people, the leader should not only concentrate on making those being led more knowledgeable and skillful about their work but also that their work would become more meaningful. Burns' view is that once the person feels capable in doing work that is more meaningful to them, they will naturally perform their work more efficiently and effectively. Essentially, the application of a transformational approach to leadership introduced the need for the person in the leadership position to be far more considerate and inclusive of the needs of those being led.

Important within this discussion of the essential role of ethical practice in leadership is Burns' (1978) argument that a fundamental requirement in leadership is that

leaders address themselves to followers wants, needs, and other motivations, as well as their own, and thus they serve as an independent force in changing the make up of the followers' motive base through gratifying their motives (p. 20, italics in original text).

Later he urges "leaders must accommodate followers wants and needs without sacrificing basic principles" (p. 77) so that they can "help transform followers' needs into positive hopes and aspirations" (p. 117). Finally, in a very powerful manner, Burns warns that the

test of leadership in all its forms is the realization of purpose measured by popular needs manifested in social and human values. Leaders can operate off the “skin” of public opinion - off surges of trending opinion, the applause of idolizing spectators, the bubbling up of passing social and political fads, trumped up foreign crises, and exaggerated dangers to national security - without recognizing the persisting, widespread, and intensive needs and goals that motivate followers and are there for the evoking by leaders (p. 251).

The challenge for the leader, according to Burns, is not to be moved and influenced by the superficial, generalist, presumed values and needs of the followers but those that are at the very core of their thoughts, feelings and actions. Simply stated, this is a view of leadership whereby the needs of those being led, and not the leader's or the organization's interests, is the primary departure point for the leader's actions, which is the cornerstone of ethical leadership practice. This understanding is pivotal in denouncing the claims to leadership by those in positions of high office who use charismatic appeal or misleading information to manipulate the emotions and beliefs of others for their own self-interests.

Subsequent leadership theories have followed Burn's lead of adding an ethical expectation to leadership practice. For example, [Hodgkinson \(1978\)](#) called for the restoration of a moral foundation of theories on educational leadership. A few years later, [Greenfield \(1981\)](#) and [Foster \(1986\)](#) emphasized the need for further study into the ethical and moral aspects of educational leadership. [Begley \(2003\)](#) described the concept of authentic leadership as a “professionally effective, ethically sound, and consciously reflective” (p. 1) form of leadership. Also, when promoting moral leadership, both [Sergiovanni \(1992\)](#) and [Starratt \(1994\)](#) shared the view that the leader must closely consider what is significant, what is right and what is worthwhile in all of their decision-making activities and that commitment to such ethical and moral behavior elevates the leader above mere pragmatics or expediency so as to raise their moral reasoning in order to transform their self and those they lead. Also, [Day \(2000\)](#) argued that the fulfillment of a truly transformational approach to leadership was depended upon a far more “values-led” commitment than initially considered due to the ever-increasing influence of change and uncertainty upon the role of the leader. Thus, a substantial and growing body of research ensued aimed at developing an understanding regarding the ethical dimensions of leadership in schools ([Arar et al., 2016](#); [Langlois et al., 2014](#)).

However, despite more than 40 years of literature describing and promoting the ethical imperative in educational leadership, it is yet to gain profound understanding and therefore universal traction. Hence, this chapter seeks to overcome this deficiency in three ways. First, it will outline the diverse and compelling reasons for why 21st Century educational leadership must be unequivocally founded upon an ethical base. Then this chapter will briefly review the expansive ethical leadership literature in order to not only show what has been proposed to date regarding the practice of ethical educational leadership but also to highlight its apparent insufficiency. Finally, this chapter will describe how this insufficiency can be overcome and thereby providing an explanation for why true leadership is intrinsically ethical. The fundamental message of this chapter is that educational leadership is not about becoming an ethical leader but rather it is about becoming a leader in the true sense of the word because this demands a commitment to being ethical.

Why ethical educational leadership?

Clearly the 2008 global financial crisis brought the concept of ethical leadership into the international spotlight. The worldwide physical, financial and psychological devastation caused by the unethical practices of a relatively few individuals and institutions within the banking and investment world drew universal attention to the need for ethical leadership practices. However, in reality this incident was only one of an ever-growing number of very serious unethical activities by people in leadership positions in the political, industrial, religious, media, and education spheres. Within the political sphere there have been a growing number of coalition governments as voters became disillusioned and skeptical about whom to support as decisions by politicians regardless of persuasion appear to be founded on self-interest rather than common good reasons. In the industrial scene, decisions by those in leadership positions being based upon economic benefit rather than human safety were behind the catastrophic oil pollution incident in the Gulf of Mexico and the mining disaster in New Zealand that led to the death of 39 miners. Individuals in religious leadership positions have acted contrary to their religious principles and norms by becoming civil militia leaders in Ireland, the Middle East and North Africa or have concealed criminal behavior by some of their ministers. England witnessed phone hacking by media employees in leadership positions who were desperate to boost sales. Also, persons in educational leadership positions have been found to have acted fraudulently when overseeing formal student assessment procedures while others have misappropriated their institution's financial and/or physical resources.

As a result of seemingly burgeoning occurrences of unethical practices by people in leadership positions, having a commitment to being ethical is now being viewed as an inviolable expectation of leaders. It is expected that our leaders will act morally whereby they will not produce harm but rather will show the virtues of doing good, of honoring others, of taking positive stands, and of behaving in ways that clearly show that their own self-interests are not the driving motivation behind their actions ([Branson, 2014a](#)). People want to know that their leaders are there for them; that the leader is one of them. People will only fully support leaders who act fairly, justly and promote good rather than harm. They want their leaders to be accountable to them in what they do just as they accept they will be accountable to the leader and others in what they do. More so than ever before, people will only respect leaders who are ethical whereby they are guided by “an internal moral compass that directs them to take on tasks,

assume styles, and behave commensurate with organisational beliefs regarding right and wrong, virtue and vice, and social responsibility" (Normore and Doscher, 2007, p. 2).

In conjunction with this call for ethical leadership arising from a general concern about the prevalence of unethical leadership there now exists a very clear realization of its importance in maintaining global economic stability as our world transitions from the third to the fourth stage of the industrial revolution (Industry 4.0). From an Industry 3.0 world influenced by electronics, computers and automated manufacturing to one that will be influenced by cyber-physical systems in which real objects and virtual processes are interlinked. But getting to this fourth stage is not solely dependent upon technological developments because particular socio-cultural changes will be required, too. As artificial intelligence, the internet-of-things, and machine learning replace people in most routine and repetitive type tasks, it will be essential for leaders to be ever mindful of protecting working conditions, of helping to shape the extent of automation as applied to jobs, and of preventing the digital revolution from widening inequalities. Simply stated, as the world moves toward Industry 4.0 the perspective of leaders must switch from "could we" to "should we" (Deloitte, 2020), which implies a mandatory commitment to ethical decision-making. As the 2020 Deloitte *Global Human Capital Trends Report* argues, "how organizations combine people and machines, govern new human-machine work combinations, and operationalize the working relationship between humans, teams, and machines will be at the center of how ethical concerns can be managed for the broadest range of benefits." (p. 98).

For the educational leader, the growing general demand for ethical leadership in conjunction with this Industry 4.0 perspective adds considerable import for ethical educational leadership. Certainly, the need for educational leaders to abide by legal obligations and proper fiscal practices is unequivocal and it is only on very few occasions that individual principals have not met this mandatory expectation. But, as more and more are being asked of the educational leaders, choosing what to attend to and what to put off or overlook is an ethical decision. Each person can be influenced by personal preferences and our mind is very adept at defending these perspectives. Thus, any choice by an educational leader about where to concentrate her/his interests and energy must be based on common good and not self-interest – it must be an ethical choice and not for personal benefit.

Also, the learning environment within an educational setting must be in the best interests of each student's future and not what maintains, necessarily, the professional and personal interests and comfort of the teachers. This means that change is a constant companion for everyone working in an educational organization. However, we all know that leading change, which begins in the role of the leader, is fraught with difficult challenges. But to ignore the future needs of the students, or to ignore the need to change because of its challenges, are unethical leadership decisions.

Into these ethical dilemmas for the educational leader add issues associated with the need for an authentic commitment to properly attending to cultural needs, family needs, and local community needs within the educational setting. Also, there is the need by the educational leader to lead their educational community to appropriately respond to political initiatives that may or may not benefit the students and the school. Not least, of course, is the need for the educational leader to ensure they make proper choices about their own life balance whereby neither work nor family/leisure gain an inappropriate hold or prominence on their lifestyle.

Every choice an educational leader makes is based on values that are either known or unknown, acknowledged or unacknowledged, by them. The less known or acknowledged are the values that direct a choice, the more likely that an unethical decision will be made, particularly in very unusual, complex and challenging situations. More and more today, educational leaders are being regularly confronted by unusual, complex and challenging situations, which demand them to make a choice. Indeed, ethical leadership is part and parcel of educational leadership, today. Perhaps the most important choice that an educational leader has to make is in answer to the question – Do I want to be an ethical educational leader?

But, before looking at what is promoted as proper ethical educational leadership practice, it is important to, first, establish a common understanding of what constitutes ethical leadership.

Ethical leadership defined

According to the widely supported view of Brown et al. (2005), ethical leadership can be defined as "the demonstration of normatively appropriate conduct through personal actions and interpersonal relationships, and the promotion of such conduct to followers through two-way communication, reinforcement, and decision-making" (p. 120). The first part of this definition relates to the *ethical person* responsibility of ethical leadership and the second part to the *ethical developer* responsibility. Leaders embracing the ethical person perspective strive to be a role model for what constitutes an ethical person through their integrity, honesty, fairness, trustworthiness, and care for the well-being of others. The ethical developer proactively advocates and engenders ethical behaviors in others through their efforts to influence and guide the ethical behavior of those they are leading by regularly communicating and explaining ethical standards, acknowledging and affirming those who act ethically, and suitably disciplining employees who demonstrate unethical behaviors.

In regard to what ethical leadership might imply in an educational context, Nancy Tuana (2014) provides a very similar view to that of Brown and colleagues when arguing that "effective and responsible educational leaders must also be ethical leaders [whereby] their leadership decisions and practices be based on ethical values and principles [and] that they lead with ethics in the sense of being committed to, and take a leadership role, in ensuring the development of moral literacy throughout their communities" (p. 262). She goes on to argue that effective ethical educational leadership is grounded on the recognition that ethical leadership mandates putting ethics in the foreground of all activities. It is not possible for the actions of only one person, no matter how ethical or how effective a leader, to make an entire community ethical. An ethical community exists because the commitment to

ethical leadership permeates the entire community. Hence, the focus on ethical leadership cannot be solely limited to the practises of the school leader as the ultimate aim must surely be about educating the entire school community to act ethically. So, what might this mean in practice?

Ethical educational leadership practice

Although there is a substantial pool of resources available for those wishing to develop ethical educational leadership practices, these resources mostly concentrate on the *ethical developer* perspective, particularly ethical decision-making. This is because these resources treat being ethical more like a recipe than a manifestation of the leader's character. The inherent assumption here is that all the leader requires for being ethical is to be seen as acting ethically according to a prescribed process. Moreover, the critically important time for the leader to be seen as ethical is presumed to be in their decision-making. Arguably, this concentration upon ensuring that a leader's decision-making process is ethical is founded upon the realization that, in today's far less predictable, controllable and uniform world, leaders can no longer depend upon past solutions when solving current problems that often contain value contestations or ethical dilemmas. Ethical dilemmas are ill-defined problems that have high-stakes consequences. Navigating ethical dilemmas requires recognition and proper representation of multiple pieces of information. It also requires intuitive judgment about potential outcomes. Ethical decision making is about the leader doing what they ought to do for the good of all rather than just doing what they would like to do. Thus, the aim of ethical decision-making is to ensure that the leader's conscious or unconscious interests and preferences to not undermine the process so that it remains just, fair, thorough and defensible. Thus, for a leader's decision to be ethical the process must be seen to be inclusive of all necessary information, be free from the leader's self-interests, be generally applicable, and be trustworthy in producing understandable and explainable outcomes.

To this end, there are three different types of approaches to ethical decision-making. First, there is the simple logical-sequential approach whereby the author proposes an ordered series of steps for the leader to follow toward the achievement of an ethical decision. For example, the Markkula Center for Applied Ethics at Santa Clara University suggests the following five steps when making an ethical decision: Recognize an ethical issue; Get the facts; Evaluate alternative actions; Make a decision and test it; Act and reflect on the outcomes. This resource, like others that are similar, provides further clarifying questions and/or statements to be addressed within each step. A Google search of ethical decision-making models returns more than 200 million sites. Many of these sites provide a similar version of this logical-sequential approach. Although such an approach provides clarity of purpose during a time of ambiguity and uncertainty its significant limitation is in its assumption that the leader has the capacity to know how to recognize the presence of an ethical issue and to be able to suspend unconscious biases or ignorance when getting the facts and evaluating alternative actions.

An alternative approach to ethical decision-making strives to overcome these limitations by providing a far more holistic process often presented as a flowchart or model. For example, the ethical decision-making model proposed by [Cranston et al. \(2014, p. 237\)](#) provides the following nine lenses the leader can use to ensure they "get the facts" about the ethical dilemma; professional ethics, organisational culture, public interest, global context, economic and financial context, legal issues and policies, institutional context, society, and political framework. In addition, these authors acknowledge the potential influence of the leader's unconscious biases and propose that the leader should consult with a "significant other" (p. 238) in order to nullify any such limitation on the choices and judgements. A somewhat similar model is the *Leader Sensemaking Model* posited by [Thiel et al. \(2012\)](#). The uniqueness of this particular model is the view that, for most leaders, "ethical awareness is grossly misunderstood and under simplified" (p. 49). Hence, the inclusion of a sensemaking component in the model whereby sensemaking is defined as "the complex cognitive process engaged in when one is faced with complex and high-risk situations" (p. 50). Although this approach strives to comprehensively illustrate the decision-making process in order to offer a far more prescriptive procedure, it is precisely this characteristic that produces its limitation. Such an approach can appear as being far too cumbersome with too many issues to consider making the process even more complex and obscure.

The third approach to ethical decision-making is based on a philosophical rather than a logical sequential foundation. The research of [Shapiro and Stefkovich \(2016\)](#) illustrates this approach; their text describing the approach has been published in many different languages and is used widely around the world. Instead of a sequence of steps to follow, this approach guides the leader toward applying the four paradigms of the ethics of justice, critique, care, and the profession when gathering the facts and making a decision in regard to an ethical dilemma. The ethic of justice directs the leader's attention toward determining which laws, rights, and policies need to be considered. The ethic of critique stands in opposition to that of justice by requiring the leader to review any inherent problems, inequities, or deficiencies within the information gathered via the ethic of justice. The ethic of care calls upon the leader to listen to others, to be concerned for the well-being of all, and to attend to the needs of all affected by the outcomes generated by the decision-making process. Finally, the ethic of the profession recognizes that those who work in education must personify the highest moral ideals of the profession by presenting an ideal image of the moral character of both the profession and the professional. It is argued that the leader is far more able to arrive at an ethically authentic and defensible decision if they are able to sincerely and comprehensively apply each of these paradigms to a complex ethical dilemma. From the diverse information gathered by applying each of these paradigms, the leader will be more able to determine the best, the most ethically appropriate, decision.

However, despite the abundance and diversity of advice to leaders, particularly educational leaders, over the past 25 years on how to implement an ethical decision-making process it would seem that little progress has been made. The issue persists not

only in the negative form of unethical leadership behavior but also, and arguably more concerning, in the form of ongoing leadership doubt and uncertainty about how to confidently enact ethical decision-making processes. This suggests that, despite all our efforts to the contrary, there must be some factor that is often missing from each of these different approaches to ethical decision-making.

Why is the issue persisting?

The genesis of this missing factor is captured in [Nancy Tuana's \(2014\)](#) view that there is more to ethical leadership than simply ethical reasoning as encompassed in each of the aforementioned approaches to ethical decision-making. According to Tuana, leaders also require moral literacy whereby they must gain the essential skills and abilities to become moral agents. As a moral agent not only must "their leadership decisions and practices be based on ethical values and principles, but also that they 'lead with ethics' in the sense of being committed to, and taking a leadership role in, ensuring the development of moral literacy throughout their school community" (2014, p. 153). Moreover, Tuana posits that such moral literacy consists of the following three elements: Ethical sensitivity, Ethical decision-making skills, and Ethical motivation.

The first element, ethical sensitivity, is defined as consisting of three major components: 1. The ability to determine whether or not a situation involves ethical issues; 2. The ability to identify the moral virtues or values underlying an ethical issue; and 3. Awareness of the moral intensity of the ethical issue. The second element, ethical decision-making skills has already been well addressed in the previous section. Indeed, there is so much available literature describing different versions of the three ethical decision-making approaches described above because it has been mistakenly assumed that this is all that a leader required in order to enact ethical leadership. But, if the leader is not ethically sensitive, then any knowledge about ethical decision-making processes is likely to be rarely or inappropriately applied. However, even with sensitivity and process knowledge and skills, a leader is prone to avoiding making an ethical decision unless they are absolutely committed to doing so. In the heat of the moment, when there are heightened emotions and calls for quick action due to the seriousness of the situation, the tendency has been that leaders most often revert to past actions and quick fixes. Ethical processes are intentionally or unintentionally forgotten. In such circumstances, it takes the undeterred and wholehearted commitment by the leader to ensure ethical procedures are thoroughly adopted despite the pressures from many sources for this not to be the case. Sadly, research shows that, for many educational leaders, the desire to be ethical when combined with a strong predisposition for individual gain tends to produce moral masquerading ([Branson, 2014b](#)). This research suggests that many educational leaders are motivated more toward appearing ethical rather than actually being ethical. Moreover, it highlights the importance of exploring the concept of ethical motivation as a precursor to the successful development of ethical leadership practice. Without ethical motivation, ethical educational leadership is likely to remain more of an ideal than a regular reality.

While much has already been done to help leaders understand what constitutes ethical leadership and how to implement ethical decision-making practices, little has been directed toward exploring the nature and practice of ethical motivation. To this end, however, [Branson \(2014b\)](#) argues that ethical motivation is about instinctively and consistently doing what is right for the good of others in the absence of incentives or sanctions. This is the highest form of human motivation as it is founded, not on the leader's own physical needs or personal gains but, rather, on intersubjective relationships, interpersonal friendships, mutual appreciation, and a feeling of belonging with and for others. This level of motivation has been variously referred to as an innate desire to fulfill potentialities ([May, 1967](#)), to achieve self-actualization ([Maslow, 1968](#)), to reach for self-completion ([Hultman and Gellermann, 2002](#)), or to attain self-enhancement ([Osborne, 1996](#)). Furthermore, it is noted that people pursue this motive for personal betterment or self-enhancement through two secondary motives ([Branson, 2014b](#)). Firstly, people seek self-enhancement through the motive to achieve success, which focuses on competence. Secondly, people seek self-enhancement through the motive to do things properly, which focuses on integrity. To have worth, people must believe that they are both competent and have integrity in what they are doing and in what others see them doing. People are motivated to do things that not only show to their selves that they are competent and have integrity but also so that they can believe that others share this same positive view of them. This suggests that competence and integrity have both a personal and a social dimension. In other words, the fundamental human motive for self-enhancement causes people to strive toward personal and social competence as well as striving to have personal and social integrity.

However, this perspective requires some further development to make it more specific to the role of an educational leader. It is argued that educational leaders are not only motivated to do things that show to their selves and others that they are competent and have integrity but also that show expected leadership qualities within their educational organization. Leadership does not happen in a vacuum; leadership occurs within a particular designated social group or organization. Also, the credibility within the leader's own eyes, and the eyes of others, is very dependent upon the leader's performance in enabling the organization to achieve a meaningful and successful outcome. This means that not only is the organizational context a source of motivation for the leader but also the leader strives to display expected leadership behavior. Within our current leadership literature, such outcomes are described as the leader's pursuit of authenticity. Authenticity implies a genuine, personal kind of leadership that is hopeful, open-ended, visionary, and creative ([Gardner et al., 2005](#)). In addition, leadership based on authenticity has a social dimension, as it is described as being sophisticated, knowledge based, and skillful in its response to the social demands of people having to work together ([Begley, 2006](#)). Being authentic means that the leader is able to acknowledge and accommodate in an integrative way the legitimate needs of individuals, groups, organizations, communities, and cultures rather than being only concerned by the organizational perspectives that have been the usual preoccupation of much of the leadership literature.

In summary, then, this discussion argues that the fundamental human motive for personal betterment or self-enhancement causes an ethical educational leader to strive toward displaying:

- Personal, social, and organizational competence
- Personal, social, and organizational integrity
- Personal, social, and organizational authenticity

Based upon this understanding of the importance of the fundamental human motive for personal betterment or self-enhancement, the following personal characteristics matrix is used to further clarify and describe the values that can be nurtured as a means of intentionally developing ethical leadership practices (Table 1).

This matrix of values suggests that the development of ethical leadership can be more readily accomplished by encouraging and supporting the educational leaders' growth in applying the values of self-knowledge, connectedness, transcendence, ethicalness, trustworthiness, responsibility, autonomy, presence, and wise judgment in guiding all of their leadership activities. A brief outline of each of these values follows.

Personal competence (self-knowledge)

Self-knowledge is about educational leaders having the capacity and commitment to use self-reflection to form an accurate, veridical model of their selves and to be able to use this unique source of knowledge to operate more effectively in their role as a leader (Bran-son, 2009). Such self-knowledge enables the educational leader to clarify their thinking, raise their self-awareness, develop their relationships on an intimate level, and get in touch with their emotions, thus unleashing their intuition and putting their selves in immediate contact with the school and its community. Such leaders have the courage and commitment to work with the reality of their thoughts, feelings, and environment, going at their own pace on a journey of self-discovery, developing their inner selves, growing in their courage to make ethical decisions, and building supportive and sustainable relationships.

Social competence (connectedness)

Connectedness in ethical leadership promotes the belief that it is imperative for today's educational leader to know that, given their own bounded rationality, discussion, dialogue, and soliciting divergent points of view is the only way to develop better solutions to ethical problems (Berkovich, 2018). This acknowledges that no single person alone is likely to have the combined knowledge and capacities necessary to reach the most appropriate ethical decision. Furthermore, connectedness in leadership acknowledges that contemporary educational leaders need to be more sensitive and caring in their attitudes and relationships. Connected educational leaders transcend the ordinary level of social relations by listening to the burdens of others, caring for them, and putting their selves in the place of others. In this way, the connected leader is able to consider the impact of the ethical issue and the chosen response on all affected by it, and to provide the appropriate support accordingly.

Organisational competence (transcendence)

Transcendence in leadership means that the educational leader has moved beyond, or transcended, their personal ordinary self-interested or self-centred ways. Transcended leaders realize that there are far more important considerations than those that focus on their self and that their leadership role is not about what is right for them, but about the things that are important to the people and the educational organization that they are leading (Starratt, 2004). Moreover, the transcended leader's life is about knowing the right things to do and doing them well. They can see the bigger picture and are willing to become a part of the school community that is striving to make this a reality. Those being led can see and follow the leader's ethical dedication, commitment, and convictions, can discern and appreciate the leader's values and motives, and can be inspired by the leader's sense of purpose and sincerity. An important part of such a leader is in their ability to model ethical values, beliefs, and behaviors. Through their transcended approach, ethical educational leaders become capable of inspiring others by their truthfulness and virtuousness in what they say and do, and in the honesty and transparency in the way they make clear the ethical reasons upon which their decisions are made.

Table 1 A matrix of values for sustaining the self-enhancement motivation, which is fundamental in the development of ethical educational leadership.

		<i>Context</i>		
		<i>Personal</i>	<i>Social</i>	<i>Organizational</i>
Desired outcome	Competence	Self-knowledge	Connectedness	Transcendence
	Integrity	Ethicalness	Trustworthiness	Responsibility
	Authenticity	Autonomy	Presence	Wise-judgment

Personal integrity (ethicalness)

It is not enough for educational leaders just to make the right moves for any purpose that suits them or for any vision that they might have of what their educational organizations should be like — they must consider carefully the impact of all they do on the livelihood of others (Badaracco, 2006). Educational leaders must be fully committed to act ethically such that their purposes and visions are socially useful, serve the common good, meet the needs of their staff, and elevate their staff, students and school community to a higher moral level. This requires them to be a person of courage and integrity, who is passionately committed to doing the right thing regardless of the consequences. Hence, they must continually monitor their own behavior to ensure that it is being influenced by the right values. This commitment to an ethical approach acknowledges that an essential responsibility of educational leadership is in influencing all within their school community to act ethically.

Social integrity (trustworthiness)

People need to be able to trust their leader. Being trustworthy is about the educational leader willingly acting openly, honestly, ethically and consistently (Walumbwa et al., 2011). It is more than simply telling the truth. Trustworthiness means that the educational leader consistently displays total congruence between who they say they are and what they do. Such congruence builds not only trustworthiness but also acceptance. Trustworthy leaders do not seek to be treated more advantageously than those they lead. They do not seek more from others than they are willing to give of themselves. A truly trustworthy educational leader is seen and accepted by those in their school community as being one of them, someone who is truly a part of the community. This is a leader who willingly and readily lives the community's norms and principles and, more, advocates and champions the benefits to be gained for all who follow suit. Trustworthy leaders are consistent and fair in the way they treat each of their school community members. All are treated equitably, with no one gaining benefit or advantage from the leader through preferential treatment, unacknowledged actions, or exclusive opportunities.

Organisational integrity (responsibility)

Responsibility comes from self-reference and self-organization, compliance comes from regulations and policies; responsibility comes from a conscious commitment to what is meaningful and beneficial not only to one's self but also to others (Cloud, 2009). Educational leaders who willingly embrace their need to show organisational responsibility take ownership of the results achieved by the school community and do not try to make excuses for poor performances or blame someone else for them. They accept that they are accountable for the choices they make and take responsibility for what is happening in the school as well as for their own performance. Responsible educational leaders learn from errors or omissions in judgment regardless of who makes them and, rather than seeing faults, they see opportunities to learn and grow both in themselves and others. Moreover, responsible educational leaders have a deep conviction that they must help create the best learning environment for the students and they devote themselves to making this happen despite obstacles, frustrations, resistances, mistakes, and competing demands. But they do this in a way that shows that they accept their moral responsibility for the welfare of others. The responsible educational leader strives to infuse a morally appropriate school culture in which everyone respects the dignity, integrity, and worth of each of their school community members.

Personal authenticity (autonomy)

Acting with autonomy means invariably using one's freedom to act according to one's own consciousness as informed by the full consideration of one's objective and subjective reasoning. For the ethical educational leader to be acting with autonomy means that they use their informed consciousness and sense of what is right to stave off influences that would prevent them behaving or living in keeping with ethical principles. As such, the leader's autonomy does not depend on external authority; instead, it ultimately depends on how they understand and remain true to their selves. It is the freedom the leader gains by repelling interference, manipulation, temptation, and social pressure. It is the freedom that, though often hard won, is nevertheless there to be won. Furthermore, autonomous leaders are consistently authentic to their selves. They are not expected to copy someone else's way of being a leader. As authentic leaders, they find the style of leadership that brings out the best in them and offers the best help to those they lead. As described by Gardner et al. (2005), authentic leadership extends beyond the authenticity of the leader as a person to encompass authentic relations with followers and colleagues. Having personal autonomy enables the leader to be consistently true to their core beliefs and values, which not only ensures they exhibit authentic behavior but also positively fosters the development of ethical behavior in those being led.

Social authenticity (presence)

Effective educational leadership inherently necessitates an influencing relationship – influencing others to see how they can be continually growing personally and professionally. Here, the concept of “presence” has gained prominence. According to Senge et al. (2007), presence involves being fully conscious and aware in the present moment through deep listening, and of being open beyond one's preconceptions and historical ways of making sense. Being able to consciously feel, understand, analyze,

and employ one's immediate experience is an essential aspect of being present. This involves being present to one's self and being present to the others. Being present to one's self is about thinking differently whereby both objective and subjective data from the immediate experience are brought into awareness and, together, these inform the person's deliberations rather than simply repeating past habitual ways of thinking. While being present to the others does not only mean being there physically and hearing what others have to say but rather it necessitates the need to be there fully for the others. Presence in educational leadership implies that the leader willingly and consistently directs alert attention and empathic sensitivity to each of their school community members whenever necessary so that their presence activates not only their own authenticity but also the authenticity of these others.

Organisational authenticity (wise judgment)

Simpson and French (2006, p. 249) propose that "leadership is evoked precisely in moments or situations where we do not know - either what we are doing, or where we are going, or how to get to where we want to be". To be a leader is to be thrown into ambiguity (Endres et al., 2009) and situations that demand "swift and nimble action" (Ropo and Sauer, 2008, p. 565). Moreover, Lawler (2005, p. 227) concludes that leadership is "inherently a chaotic, irrational, emotional phenomenon" such that Pye (2005, p. 47) argues that leadership is "sensemaking in action." In other words, leadership in general, and ethical leadership in particular, calls for wise judgment in the face of uncertainty, unpredictability, and ethical implications. Thus Sternberg (2007) describes successful and effective leadership as the result of a synthesis of three key attributes — wisdom, intelligence, and creativity. He adds that a successful leader needs creativity to generate good ideas, academic intelligence to ascertain whether the ideas are good, practical intelligence to persuade other people, and wisdom to make the ideas work for the common good. Similarly, Bierly et al. (2000) contend that wisdom relates to the ability to effectively choose and apply appropriate knowledge in a given situation and define wisdom as the ability to best use knowledge to establish and achieve desired goals. Importantly, Small (2004) proposes that wisdom is having the power to apply experience and knowledge critically, ethically, and practically.

Summation - where to from here?

This chapter began by arguing that now more than ever before, it is time to challenge the superficial, erroneous and potentially harmful assumption that any person appointed or elected to a leadership position automatically becomes a leader. It must be acknowledged that a leader who is not deliberately and sincerely striving to be ethical is only simulating, if not fabricating, leadership. The actions of a leader, who is not deliberately and sincerely striving to be ethical, is management, at best, or something far more sinister and deceptive, at worst. It's time to stop calling this "leadership" because this only justifies or gives credibility to such actions. Leading ethically must not remain optional but rather become an imperative for anyone wishing to be called a leader especially for an educational leader.

Moreover, there is much more to actually being an ethical educational leader than simply knowing how to make an ethical decision. It is crucial that ethical educational leadership is seen as a personal attribute rather than a professional capability. Knowing how to enact ethical decision-making is insufficient in and of itself. If this were not so, then the growing concern for an apparent lack of ethical educational leadership would not be evident because there has been an abundant amount of literature during the past 30 years describing how leaders can implement ethical decisions. We cannot be satisfied if our educational leaders can demonstrate that they know how to solve theoretical ethical dilemmas or mount a moral argument but struggle to apply such knowledge and capability in their daily work. They need to go beyond *potential* to *habits* of behavior whereby they naturally live out a coherent and defensible values position, *choosing* to apply their ethical knowledge and capabilities daily in a rapidly evolving complex reality. An outside observer should be able to say, not that our educational leaders have a particular ethical capability, but that they live it. Being ethical is a personal attribute that one associates with them. Thus, our focus should be on developing the ethical person, not predominantly on developing their ethical capability.

This implies that the genesis of ethical educational leadership is in the leader's "being" and not just their "doing". It manifests from the educational leader's person, their character, and their behavior is but the tangible reflection of this. Simply having the desire to act ethically does not translate, necessarily, to actually being an ethical educational leader. Ethical educational leadership is an achievement and not a given. A leader has to strive, to be determinedly motivated, to become an authentically ethical person in order to be able to enact ethical leadership. As I argued previously (Branson, 2014b), if we want ethical leaders who naturally act with heart and soul, leaders with courage based on inner ethical strength, wisdom, and grace, if we want leaders with breadth of ethical vision and depth of character, then we will need to support them in developing themselves to their real depths, not by following some decision-making formula, but by supporting them in finding out who they are, what they are made of, who it is that they really want to become, and what values will help them to become this ethical person. If it is possible to help individuals to develop the skills and knowledge necessary to cater for their physical, safety, belonging, and esteem needs, then it must be similarly possible to help educational leaders to develop their skills, knowledge, and character so that they are able to reach their full ethical educational leadership potential.

References

- Arar, K., Haj, I., Abramovitz, R., Oplatka, I., 2016. Ethical leadership in education and its relation to ethical decision-making: the case of Arab school leaders in Israel. *Ethical leadership in education and its relation to ethical*. J. Educ. Adm. 54 (6), 647–660.
- Badaracco, J.L., 2006. *Questions of Character: Illuminating the Heart of Leadership through Literature*. Harvard Business School Press, Boston.
- Begley, P.T., 2003. In pursuit of authentic school leadership practices. In: P. T. Begley and O. Johansson (Eds.), *The ethical dimensions of school leadership*. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 1–12.
- Begley, P.T., 2006. Self-knowledge, capacity and sensitivity: prerequisites to authentic leadership by school principals. J. Educ. Adm. 44 (6), 570–589.
- Berkovich, I., 2018. Typology of trust relationships: profiles of teachers' trust in principal and their implications. *Teachers Teach*. 24 (7), 749–767.
- Bierly, P., Kessler, E., Christensen, E., 2000. Organizational learning knowledge and wisdom. J. Organiz. Change 13 (6), 595–618.
- Branson, C.M., Gross, S.J. (Eds.), 2014. *Handbook of Ethical Educational Leadership*. Routledge, New York.
- Branson, C.M., 2009. *Leadership for an Age of Wisdom*. Springer Educational Publishing, Dordrecht, Netherlands.
- Branson, C.M., 2014a. Conclusion: if it isn't ethical, it isn't leadership. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 439–454.
- Branson, C.M., 2014b. De-constructing ethical motivation. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 439–454.
- Brown, M.E., Trevino, L.K., Harrison, D., 2005. Ethical leadership: a social learning perspective for construct development and testing. *Organ. Behav. Hum. Decis. Process*. 97 (2), 117–134.
- Burns, J.M., 1978. *Leadership*. Harper & Row, New York.
- Cloud, H., 2009. *Integrity: The Courage to Meet the Demands of Reality: How Six Essential Qualities Determine Your Success in Business*. HarperCollins, New York.
- Cranston, N., Ehrich, L.C., Kimber, M., 2014. Managing ethical dilemmas. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 229–245.
- Day, C., 2000. Beyond transformational leadership. *Educ. Leader* 57 (7), 56–59.
- Deloitte, 2020. *The social enterprise at work: Paradox as a path forward*. Deloitte Global Human Capital Trends, London.
- Endres, M., Chowdhury, S., Milner, M., 2009. Ambiguity tolerance and accurate assessment of self-efficacy in a complex decision task. J. Manag. Organ. 15, 31–46.
- Foster, W., 1986. *Paradigms and Promises: New Approaches to Educational Administration*. Prometheus, Buffalo, NY.
- Gardner, W.L., Avolio, B.J., Luthans, F., May, D.R., Walumbwa, F., 2005. Can you see the real me? A self-based model of authentic leader and follower development. *Leader. Q.* 16 (3), 343–372.
- Greenfield, T.B., 1981. Can science guide the administrator's hand? In: Aoki, T.T. (Ed.), *Rethinking Education Modes of Inquiry in the Human Sciences*. University of Alberta Press, Edmonton, AL.
- Hodgkinson, C., 1978. *The Philosophy of Leadership*. Blackwell, Oxford.
- Hultman, K., Gellermann, B., 2002. *Balancing Individual and Organizational Values: Walking the Tightrope to Success*. Jossey-Bass, San Francisco.
- Langlois, L., Lapointe, C., Valois, P., de Leeuw, A., 2014. Development and validity of the ethical leadership questionnaire. J. Educ. Adm. 52 (3), 310–331.
- Lawler, J., 2005. The essence of leadership? Existentialism and leadership. *Leadership* 1, 215–231.
- Maslow, A., 1968. *Towards a Psychology of Being*, second ed. Van Nostrand, Princeton, NJ.
- May, R., 1967. *Psychology and the Human Dilemma*. Van Nostrand, Princeton, NJ.
- Normore, A.H., Doscher, S.P., 2007. Using media as the basis for a social issues approach to promoting moral literacy in university teaching. J. Educ. Adm. 45 (4), 427–450.
- Osborne, R.E., 1996. *Self: An Eclectic Approach*. Simon & Schuster, Needham Heights, MA.
- Pye, A., 2005. Leadership and organizing: sensemaking in action. *Leadership* 1, 31–50.
- Ropo, A., Sauer, E., 2008. Dances of leadership: bridging theory and practice through an aesthetic approach. J. Manag. Organ. 14, 560–572.
- Senge, P., Scharmer, C.O., Jaworski, J., Flowers, B.S., 2007. *Presence: Exploring Profound Change in People, Organizations and Society*. Nicholas Brealey, London.
- Sergiovanni, T., 1992. *Moral Leadership: Getting to the Heart of School Improvement*. Jossey Bass, San Francisco, CA.
- Shapiro, J.P., Stefkovich, J.A., 2016. *Ethical Leadership and Decision Making in Education: Applying Theoretical Perspectives to Complex Dilemmas*. Routledge, New York.
- Simpson, P., French, R., 2006. Negative capability and the capacity to think in the present moment: some implications for leadership practice. *Leadership* 2, 245–255.
- Small, M.W., 2004. Wisdom and now managerial wisdom: do they have a place in management development programs? J. Manag. Dev. 23, 751–764.
- Starratt, R.J., 1994. *Building an ethical school: A practical response to the moral crisis in school*. Falmer Press, London.
- Starratt, R.J., 2004. *Ethical Leadership*. Jossey-Bass, San Francisco.
- Sternberg, R.J., 2007. A system model of leadership. *Am. Psychol.* 62 (1), 34–42.
- Thiel, C.E., Bagdasarov, Z., Harkrider, L., Johnson, J.F., Mumford, M.D., 2012. Leader ethical decision-making in organizations: strategies for sensemaking. J. Bus. Ethics 107, 49–64.
- Tuana, N., 2014. An ethical leadership developmental framework. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 262–297.
- Walumbwa, F.O., Luthans, F., Avey, J.B., Oke, A., 2011. Authentically leading groups: the mediating role of collective psychological capital and trust. J. Organ. Behav. 32, 4–24.

Schools as research-informed professional learning organizations

David Godfrey, UCL Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	146
The theoretical landscape	146
The research-engaged school	150
Recent literature related to research-engaged schools	151
Research on evidence and research-informed practice in schools	151
System and practitioner (macro and micro level) patterns of research and evidence-informed practices in schools	151
The exo-system: models for collaborative research-informed school improvement	153
School (meso-level) factors affecting research engagement	154
The strength of the evidence	154
Conclusions	156
References	156

Introduction

Internationally there have been many calls for schools to be learning organizations and for leadership and teaching policies and practices to be more closely informed by research. For instance, in USA, since 2002, the Institute of Education Sciences has funded over 300 evaluations in schools (Tipton et al., 2021). In Australia, there have been renewed calls for schools and teachers to become research-rich (Shore et al., 2017) and for the school systems to use evidence for improvement (APC) 2016. Similar policy movements can be seen in several other jurisdictions.

This chapter has three overall aims: to describe this field of academic study, to bring the reader up to date with some relevant research and to assess the strength of this scholarship and identify gaps for further exploration. The chapter first describes the parameters used in describing schools as research-informed learning organizations vis a vis the identification of the academic fields that intersect this topic, key concepts, and their definitions. Some of the debates and discussions in this area will also be outlined, including the relationship between academic research and the practices of teaching and school leadership. The intention here is to outline the contested nature of some of the theoretical and normative positions adopted by scholars and researchers in this area, for instance, the notion of *what works* research and of teaching as a research-informed profession. This is important, since resolutions to these arguments result in different outcomes about school organizational configurations, the ways that teachers work with research and the kinds of research used by school leaders to inform decisions.

The author's and others' preferred conception of this area, that is, research-engaged schools, and the (eco)-system perspective on how such institutions serve as lynchpins in a dynamic and improving school system, is then outlined. The chapter then brings the reader up to date with a range of international publications, providing some conclusions from these studies, particularly the leadership conditions and factors involved in promoting research-engagement in schools. This is achieved by analyzing 17 publications from 2017 to 2021 covering a wide range of pertinent foci and empirical evidence from nine countries and examining the issues from all levels of the ecosystem. The chapter ends with a look at the strength of claims made in this academic field, first by looking at research that attempts to link research-engagement with school improvement and secondly by describing systematic reviews in areas highly pertinent to this topic. In summarizing this work, the intention is to gauge current thinking on schools as research-informed organizations. Areas that require more exploration are discussed and further research avenues are proposed to explore gaps and to bolster claims made, including the efficacy of approaches designed to promote research engagement.

The theoretical landscape

As will be discussed further in this chapter, the field concerning schools as research-informed learning organizations covers several overlapping areas of scholarship. A quick reflection on the elements involvement in making practice and policy within schools "research-informed" will make its multi-faceted nature clear; these include, among others: the nature of research-use by teachers and school leaders, the use of professional learning communities, the creation of a learning organization and an infrastructure that helps to mobilize evidence to enable practitioners to make sense of it and to translate this into effective teaching and learning strategies.

A simple Google Scholar key word search on some of these areas provides a picture of the shape of the field: (Table 1)

Table 1 Google scholar searches related to schools as research-informed learning organizations.

Search term	Number of returns (October 7, 2021)
"Professional learning communities" and schools	77,000
"Teachers as researchers"	11,500
"Schools as learning organizations"	1310
"Data use in schools"	930
"Research learning communities" and schools	472
"Research-engaged schools"	244
"Evidence-informed practice in schools"	39
"Practice-based research in schools"	32
"Research-informed schools"	13
"Research-informed school leadership"	4

The table above shows the maturity of some areas of scholarship, such as professional learning communities (PLCs), schools as learning organizations (LOs) and teachers as researchers. There has been a relatively recent uptick in publications in the areas of research-engaged schools (RESs) (e.g., [Godfrey and Brown, \(2019\)](#); [Dimmock, 2019](#); [Godfrey and Brown \(2019\)](#); [Godfrey and Handscomb, 2019](#); [Godfrey, 2016](#)) and research learning communities (RLCs) (e.g., [Godfrey and Brown, 2019](#)), albeit both of these are still limited to relatively few scholars. It also important to note some conceptual overlap between research on data use in schools and with scholarship on evidence-informed practice (EIP) ([Brown et al., 2017](#)). While recognizing the blunt instrument of a Google Scholar search, nevertheless, there appears to be a paucity of published work that looks specifically at research and evidence-informed leadership and EIP at school organizational level. Some recent additions to this work are explored later.

More systematic work by [Farley-Ripple et al. \(2020\)](#) on Use of Research Evidence (URE), points to the fragmented nature of scholarship in this area. Surveying a wide community of scholars in the USA and UK, conducting both a social network and bibliometric analysis, they identified the different strands of recognized, published work and key scholars across sectors and disciplines involved in URE research. The most influential scholars named were from: sociology (27%), political science (20%), organizational studies (14%), psychology (12%), and economics (2%). Multiple key words were used to describe this field too: policy studies (21%), knowledge utilization (18%), evidence-based or evidence-informed decision making, policy, practice (17%), research impact (15%), implementation science (11%), and knowledge mobilization (9%). In their network analysis they found low density and high fragmentation of scholars in this area; respondents nominated 185 references across a range of disciplines, policy issues, and publication formats spanning more than five decades, with articles from 57 journals, overwhelmingly from health and education fields. However, even the most referenced journal, *Implementation Science*, had only with five publications, and the most nominated author was Carol Weiss, but most were nominated by only one person.

Some of the complexity of this area is brought together via the concept of *research-engaged schools*, which will be outlined later. Before this, the two key terms within the chapter title will be explained, including some key distinctions and debates about them, that is: *learning organizations* and *research-informed practice in schools*.

Learning organizations (LOs) essentially do two things: they promote the learning of their members through the culture and structural conditions of the organization, especially through its leadership, and secondly, they feed this learning back into the organization itself, making it better adapted, effective or efficient in the performance of its key functions. In other words, the "learning" flows both to the members/staff in an organization and back to the organization itself in a relationship of mutual advantage. Based on extensive survey data from Wales and the use of component and reliability analysis, eight dimensions are proposed for schools as LOs, these are:

- (a) a shared vision centred on the learning of all students, (b) partners contributing to school vision, (c) continuous learning opportunities, (d) team learning and collaboration, (e) a culture of enquiry, innovation and exploration, (f) systems for collecting and exchanging knowledge and learning, (g) learning with and from the external environment, and (h) modelling learning leadership.

[Kools et al \(2020, p. 24\).](#)

Numerous and contested assumptions follow from the concept of LOs. These include the nature of a school "organization", the notion of "learning" itself and criteria for determining "improvement". To begin with, viewing schools as organizations may promote a business orientation to schools and to see them as isolated units instead of as part of a wider embedded system underpinned by values. For this reason, some have proposed schools as "institutions" as preferable to the term organization (e.g., [Glatter, 2015](#)). A further distinction relates to the disciplinary underpinnings taken by scholars. For instance, some work employs cognitive processing theories to explain information flows in organizations analogous to the workings of the human brain ([Friedman et al., 2005](#)). While these may be applicable to individuals, the charge is that, by anthropomorphizing organizations this ignores many of the institutional, historical and practical enablers and barriers that allow or prevent learning taking place in an organization. By

contrast, social learning perspectives see organizations as sites of social and historically located practices (Lave and Wenger, 1991). The latter provides the foundation for the voluminous work on professional learning communities in schools. While enjoying strong favor among many researchers and practitioners, PLCs can be critiqued for the tendency to assume a consensus or collectivist view of professional work that may belie tensions with the notion of professional autonomy as well as between individuals' different educational values and aims that may make collaboration difficult.

One of the critiques of LO scholarship is that it is replete with assertions and terms that are insufficiently underpinned by empirical evidence (Schechter and Mowafaq, 2013). While organizational learning (OL) as a field of academic study does not promote any particular model or ideal, the broad LO literature often does exactly this (Friedman et al., 2005). However, such ideals are hard to empirically test due to the complexity in ascertaining when and how LOs have been demonstrably better at exercising their functions than comparable organizations. In schooling, the issue also touches on concerns about the propensity for research to measure what is easier to measure rather than what is important to measure, especially by narrowing down improvement criteria to student attainment in standardized testing (Biesta, 2017).

Proponents of LOs nevertheless provide extensive theoretical rationales about the usefulness of promoting schools as LOs to serve the needs of society, students, the economy and so on. One of the arguments in favor of the view of schools as learning organizations is that they fit a 21st Century model of what is required of schools, to equip their students for the complexities of the (post-) modern world. With the advent of ubiquitous knowledge via the internet, the job of teachers is less about imparting information and more so than ever, about making students into expert life-long learners. To do so, it is argued that teachers need to be knowledge workers, skilled in the analysis of information, and creation of new knowledge through enquiry (Price and Weatherby, 2018). Teaching, as in other professional domains (medicine, law and so on), it is argued, should therefore be underpinned by a range of knowledge: that obtained during by and in practice, observation, from student data **and from research**, in other words, it should be "evidence-informed".

Part of the picture that leads to evidence or research-informed practice in schools is the "infrastructure" that enables knowledge to flow to and from the separate domains of research and practice. The debates about "useful" research in academia for instance, have a long history in education, for example, in the UK, with Lawrence Stenhouse's work in the 1970s. These will not be covered once again here in full but will be touched upon below in the positioning and critiques of the *what works* model. More recently, as part of efforts to bring research knowledge closer to practice, some countries have adopted country-wide knowledge mobilization strategies. Knowledge mobilization (KM) works at interpersonal, institutional and system level and can be defined as:

The active and dynamic process whereby stakeholders (e.g., research, practitioners, policy makers and community members) share, create and use research evidence, to inform programming, policy, decision-making and practice'

Malik (2016, p. 11), in Pollock et al. (2019, p. 27).

In the US, the Campbell Collaboration is a social science research network that produces and disseminates systematic reviews and other evidence to inform public policy and practice. In the UK, the EPPI-Centre has a similar mission and does this through developing methods for systematic reviews and research syntheses, conducts reviews, supports others to undertake them, and provides guidance and training in this area. They also study the use of research evidence in personal, practice and political decision-making, supporting those who wish to find and use research to help solve problems, and providing guidance and training in this area. In the UK, the Education Endowment Foundation (EEF) was established in 2011 to improve the educational attainment of the poorest pupils in English schools. The EEF aims to support teachers and senior leaders by providing evidence-based resources to improve teaching practice and students' learning. As well as funding research (particularly randomized control trials) they also have a widely used teaching and learning toolkit that synthesizes research and produces guidance on the effect sizes of numerous educational interventions. Since 2016, the EEF has funded a network of "Research-schools" in England whose remit is to mobilize knowledge from research to other schools in their area, with the aim of stimulating evidence-based school improvement (Gu et al., 2020).

Debates about the relationship between research and practice are manifold and complex, and affect the way that practitioners, especially school leaders and teachers, researchers and policymakers interact with research. A way to think about this, is to contrast the *what works/evidence-based practice* model with an improvement science approach. The "what works" model proposes the use of "rigorous" research to demonstrate the efficacy and effectiveness of educational strategies. The gold standard of this approach is to use randomized control trials (RCTs), where large scale experiments randomly assign students, teachers or schools to an "treatment" or a control group and then compare outcomes after an appropriate time span. The second predominant form of evidence is through research-syntheses that provide summaries of a wide range of research on issues relevant to the education sector. The idea is to summarize, using the best available research, which strategies practitioners, policymakers and school leaders should use. In the EEF teaching and learning toolkit, this is exemplified with reference to the effect sizes of various interventions, the strength of the evidence and the cost effectiveness of the approach.

Proponents of the *what works* model can point to the advances that these approaches have made in translating evidence into more useful, practical and understandable forms of output that teachers and school leaders are able to access more readily (e.g., Collins and Coleman, 2021). However, a critique of this model is the exclusionary view of evidence that favors experimental

approaches and undervalues other kinds of research (Wisby and Whitty, 2019). RCTs can address simple cause-effect relationships in relation to an educational strategies, that is, whether it works better compared to usual practice, however, they fail to address how and why this method works better. This is why most RCTs in education also have a strong qualitative component to them. However, this only partially addresses their shortcomings. For instance, it has been shown that random allocation to groups is frequently “subverted”, dosage and compliance in treatment groups can vary significantly - raising doubt about how much of the “treatment” actually reached this group. Finally, due to the long time span of research the local and political milieux can affect ongoing participation and motivation in an experimental trial (Camburn et al., 2016).

Further research throws doubt on the pre-eminence notion for RCTs in Education. Lortie-Forgues and Inglis (2019) conducted an analysis of 141 trials from the EEF in England and the National Center for Educational Evaluation and Regional Assistance (NCEE). They made two important observations: first that around 40% of findings from these trials were “uninformative”, remarking that, “Such trials allow us to conclude neither that an intervention should be implemented at scale nor that this should be avoided to prevent the waste of public money” (Lortie-Forgues and Inglis, 2019, p. 164). Secondly, they found very low effect sizes in the trials they examined – much lower than would be expected from the wider educational literature. While the authors offer variations potential reasons for this (including the laudible fact that both the EEF and NCEE publish all trials, including those with negative outcomes), the solutions they offer are themselves problematic, requiring extensive replication of basic research before conducting experiments, larger trial sizes, or tailored outcome measures. The last approach, while giving the potential to detect larger effect sizes, sacrifices comparability across trials, while the other two solutions require significant time and significant additional financial commitment.

Academic research is itself, best considered one of many forms of evidence and knowledge that “informs” practice. Therefore, the term EIP is preferred by many to evidence-based practice (EBP), since the latter fits the *what works* model outlined above (Godfrey, 2017a) and fails to capture the way that humans interpret and make sense of the world around them, reducing the educational process to a mechanistic ontology where the effects of interventions “go on behind teachers’ backs” (Biesta and Burbules, 2003). Not only does this create a disempowering model of professionalism, it also overlooks the practical uses of *theory* in educational practice, including its conceptual and enlightenment role, focusing as it does on applying only research outcomes (Godfrey, 2017a). Thus, for a school-based version of a learning organization connected to research, I propose a model of professional educational practice that is “research-informed”, that is:

An actively enquiry mode of professionalism that involves critical reflection and engagement in (“doing”) and with (“using”) academic and practitioner forms of research, taking into account both the findings and theories generated from them.

Godfrey (2017a, p. 438).

Some scholars go further and suggest the need for a new paradigm of improvement science, in which practitioners and researchers work together as improvers, instead of one providing the knowledge and the other “doing” the improvement. This contrast with the *what works* adoption of a technical-rational professionalism in which expert knowledge resides in the academy and closely guides the way practitioners should work (Schön, 1983). Two of the key pillars of improvement science are: (i) to understand why programmes are often implemented with wide disparities in effect sizes in different contexts; and (ii) to work efficiently, avoiding imposing new programmes for schools (and workload) on top of existing strategies (Bryk, 2015). The imposition of top down prescriptions for teaching does not address issues of consistency because this elides the judgment needed by teachers and school leaders to adapt strategies based on the demographics of their schools and students and to align policies efficiently and effectively alongside extant approaches.

Bryk (2015) suggests that to make this improvement science paradigm work, and to close disparities between what society expects of its education system and what it achieves, there is an urgent need for an effective Research and Development (R&D) infrastructure. Such an infrastructure should focus on improving teaching and learning in schools and allow for better analysis of the problems they face, and the alternative solutions to them – avoiding what he calls “solutionitis” in which one solution is imposed on top of another, without ever properly testing and evaluating the system and how it works to achieve success through its institutional routines, rules and processes. The aim of this R&D infrastructure would not just be to know *what can work* but also to develop the *know how to make it work at scale, in different contexts, and for different sub groups*. Later in this chapter, some recent research comparing models of R&D infrastructure will be discussed, drawing out the central principles from them.

Having outlined some of the definitions, assumptions and debates that underpin this topic, the chapter now turns to a concept of the school where research, along with other types of knowledge creation and use, forms part of an ecosystem that has the potential to enhance the quality of education. It is important to state that, it is possible to view research-led institutions akin to businesses that are looking to gain “market advantage” over their rivals. Arguments can be made about the benefits or otherwise of encouraging rivalry and competition between schools to raise the overall level. However, many have pointed out that at the system-level, while it may be “theoretically” possible for competition between schools to raise standards, in reality these conditions are rarely in place and tend only to lead to greater inequalities and inconsistencies in standards (e.g., Ladd and Fiske, 2003). More scholars, in education at least, have pointed to the advantages of collaboration at school level to raise standards (e.g., Ainscow et al., 2006). Therefore, the research-engaged school is envisaged as part of a collaborative, inter-connected education ecosystem aimed at producing policies and practices that are informed (and improved) by the integration of research knowledge.

The research-engaged school

The notion of an ecosystem of research-engaged schools (RESs) has been comprehensively outlined in an edited book which looks at various levels of the system and with a range of international contributions (Godfrey and Brown, 2019). In this book, the levels of the ecosystem are mapped onto dimensions employed originally by Bronfenbrenner (1979) in his research into child development. Applied to RESs, these are: the macro-system (educational values, the purpose of schooling etc); the exo-system (the institutional configurations, collaborations and policy outcomes of these values); the meso-system (organizational level, the inter-actions between those in the micro-system); and the micro-system (the direct interactions and influences of teachers, school leaders and practitioners on students). The “ideal” RES ecosystem requires alignment of numerous elements such as: teacher training, teacher professional bodies; the accountability framework, school networks, research-practice collaborations with universities and other bodies, the intermediary brokering of research to school practitioners and within school ongoing research use, encouraged by school leaders. Seen as an ecosystem, the relationship between elements is dynamic and mutually constitutive rather than a linear chain of cause and effect’ with (evidence-based) policy being “implemented” by research school leaders, teachers and other professionals. Instead, the ecosystem has levels of affordances from its state of equilibrium and feedback loops present in different elements and sub levels of the system. For example, the accountability system may encourage innovation and research-use, or dampen it through a process of institutional isomorphism, game playing and conservatism (Ehren, 2019). At the macro-level, values may encourage cooperation or competition, or suggest a narrow or wider view of schools in relation to society, communities and so on. The teacher training system may produce new teachers accustomed to engaging critically in published research or it may turn out teachers concerned more by short term exam-based targets and teaching to the test. Seen in this way, RESs discourage thinking about leadership of individual schools as the only unit of interest in encouraging research-informed practice, the ultimate goal.

The research-engaged school operates within the ecosystem as a sub level: the meso-system. This organizational level is crucial, as it concerns the relationships between micro-level actors – especially teachers but also other professionals and stakeholders – who have the most direct influence on educational outcomes and processes. The term “research-engaged school” was originally coined by Handscomb and MacBeath (2003) and was described as a school with a research-rich pedagogy, a research orientation, that promotes research communities within and beyond the school and with research at the heart of its policies and practice. Drawing on a range of empirical evidence and scholarship, the author (Godfrey, 2016 p. 306) describes the nutrients of a research-engaged school culture as consisting of: leadership that encourages research capacity building and distributed knowledge creation; a teaching profession that is informed by a critical engagement with research; a school that is set up consciously to be a learning organization; and one that is connected to the wider system, outward looking and belonging to various networks (Fig. 1).

The focus on the meso-system brings to the forefront the importance of leadership in research-engaged schools and of effective learning communities in particular. Dimmock (2016) describes RESs as lynchpins in a system that address three concerns:

1. How to bridge the research-policy-practice gap by mobilizing knowledge more effectively through knowledge producers and consumers working collaboratively
2. Valuing and integrating both tacit knowledge and academic coded (explicit) knowledge
3. Raising the professionalism and reflectivity of teachers and leaders

Godfrey and Brown (2019, p. 9).

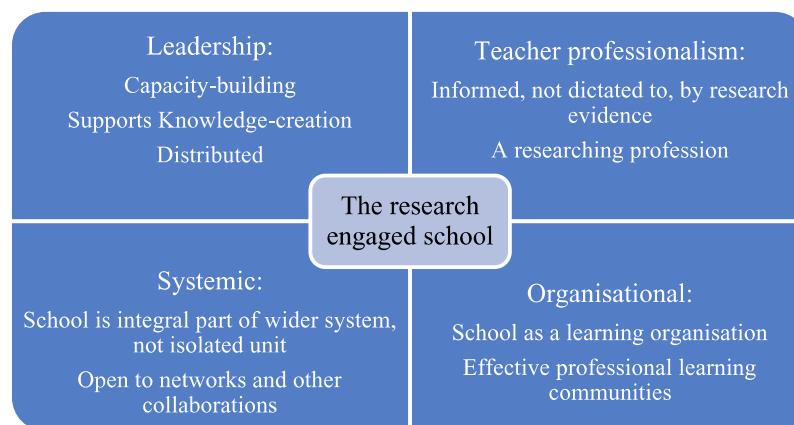


Fig. 1 Nutrients of the research - engaged school “culture”.

Dimmock sees one of the roles of leadership of schools as being to increase its organizational capital, that is, “to change and redesign the school by making better use of its intellectual and social capital to produce high-leverage strategies of teaching and learning” (Dimmock, 2019, p. 58). Research-engagement is seen as key to achieving this improvement in the innovative and learning capacities of schools. He sees three key leadership tasks in this vein:

- i. To encourage teachers to be reflective practitioners, valuing evidence and research-based approaches
- ii. To build a school culture with expectations that leadership and teaching can be improved through research-engagement
- iii. To establish processes and structures in the organization that embed such ways of working, including sustaining and scaling-up research into practice.

Godfrey and Brown (2019, p. 60).

One of the ways that this shifts traditional leadership practices is that school leaders need to be cognisant of effective methods of bringing research into practice. In Godfrey and Brown (2019), two such models are proposed, including research-informed peer review and research learning communities. More recently a model of collaborative peer enquiry has been proposed (Godfrey, 2020). These models require participants to learn procedures and principles that, especially in the early stages of use, involve modeling by external (usually university based) facilitators, intermediary brokering of relevant research, and effective partnerships with other schools. School leaders need to focus on process rather than outcome accountability for those involved (usually teachers). This is because by only by adhering to the central principles and practices of the approaches will practitioners achieve the deep level of sustained learning they are designed to engender. Judging participants by outcomes such as short-term improvement in student attainment endangers this process and leads to conservatism and a mentality of performativity. This kind of accountability is likely to be counterproductive because participants feel more inhibited, less innovative and disinclined to share their areas for development with peers, focusing rather on proving to their superiors that they have made gains. Nevertheless, for researchers, the challenge remains how to convince practitioners that research-practice models are worth the investment of time that they require. This, among other themes, is explored below in terms of the current evidence in the field.

Recent literature related to research-engaged schools

The chapter now turns to what insights have been gained on this topic from 17 articles published from 2017 to 2021 covering a range of research topics pertinent to the issue of schools as research-informed organizations. Two overall areas will be explored from these. The first are studies comparing current EIP by teachers and school leaders in various countries and how to promote EIP. The second area is the evidence-base supporting research-engagement at individual, school and system level. The research covers 11 jurisdictions: England, France, The Netherlands, Canada, USA, Finland, Turkey, Australia, Sweden, Germany, Ireland. The studies employ a wide variety of research methods: systematic reviews, surveys, interviews, focus groups, case studies, mixed methods, document analysis, participatory action research; data templates, engagement data and professional reflection. Below the analysis is divided into two overall areas: 1) research and EIPs in schools; 2) the current strength of the evidence base supporting various aspects of research-engagement.

Research on evidence and research-informed practice in schools

Research from the articles in focus covers seven areas: a comparison of EIP and policies across countries and the barriers and facilitators to EIP in these jurisdictions; approaches to improving teacher evaluative and research skills; how school leaders can encourage inquiry-based school practices; teachers as embedded researchers; how teacher education programmes can contribute to a school’s research culture; the use of research-champions; a research-brokering initiative and finally a study of research use by the Swedish school inspectorate. The intention is to examine research concerned with all levels from micro-macro level of RESs.

System and practitioner (macro and micro level) patterns of research and evidence-informed practices in schools

Before examining the ways in which school leaders can influence research-engagement at the school (meso) level, looking at EIP in different national school systems can identify ways that macro-level environments affect teacher (micro) level research-engagement. Some of these reflect ideas and values about the nature of teacher (and other educators’) professionalism and the extent to which research plays a part in this.

Malin et al. (2020), examined the use of EIP in four countries and the formal and informal institutional rules that constrain or enable these practices. In a comparative case studies’ approach, they employed a range of secondary data, looked at EIP in schools in Catalonia (Spain), England (UK), Massachusetts (MA, United States), and Rheinland-Pfalz (RP) (Germany). They classified each region employing a cohesion/regulation matrix developed by Hood (1998). This four-quadrant matrix looks at the overall macro/system level factors that define school systems, comparing countries that have high or low social cohesion (SCO) and high or low social regulation (SR). In the fatalist way, the context is defined by rule bound approaches but with little cooperation (high SR, low SC). In the hierarchist way, social cohesion and cooperation is high and ends are achieved through rule bound and bureaucratic approaches (high SC, high SR). The individualist way involves a fragmented system with informal, negotiated approaches to operating (low SR, low SC). Finally, the egalitarian way involves high participation in decisions and lots of peer to peer support (high SC, low SR).

Table 2 Four region's positions on the social cohesion-regulation matrix.

	<i>High social cohesion</i>	<i>Low social cohesion</i>
High regulation	<i>The hierarchist way</i> Catalonia	<i>The fatalist way</i> England, Massachusetts
Low regulation	<i>The egalitarian way</i> Germany (Rheinland-Pfalz)	<i>The individualist way</i>

In this classification, Catalonia was assessed as having a hierarchist approach; the English system as fatalist, as was Massachusetts (albeit with complexities and some evidence of collaboration); while in Germany the approach was considered to be egalitarian (See [Table 2](#) below).

The analysis showed how these overarching values affected the ways in which evidence was used to inform practice. In Catalonia there was a diffuse, fragmented picture of school system and research use. Spanish teachers tended to disconnect the conception of research from educational practice, in favor of experiential and peer knowledge. National education reforms were themselves not considered evidence informed, and culturally, institutions tended not to have a research culture. In England there was moderate evidence of some research use by teachers, and the EEF toolkit and research schools were given as examples of enablers. However, a lack of time for practitioners to engage in EIP and budget cuts were considered highly significant barriers. In Massachusetts, a strong top-down approach was used to push certain types of evidence into schools. Nevertheless, there were some isolated examples of strong EIP and untapped potential for cross-district collaboration. "Data" was more often used than research-evidence, although indirect access of research through brokering institutions was also seen. Lack of time was also considered a barrier to EIP in schools. In Rheinland-Pfalz (Germany), data from new evaluation and assessment practices was distrusted, and research was rarely used to inform teaching quality. There was partial uptake of the idea of the school as learning organization, but the professional autonomy of teachers made organizational level change and research-engagement a challenge, albeit not an insurmountable one, for school leaders.

Further evidence of how top-down values permeate EIP comes from research by [Gaussel et al. \(2021\)](#) in the context of the French school system. The authors describe a situation in France dominated by centralized top-down control and in which teachers see school leaders as sometimes threatening to their professional autonomy. Lack of time to collaborate with others was also seen as a barrier to French teachers. In a survey of 172 primary and secondary school teachers, the researchers wanted to find out to what extent EIP was associated with school improvement and to what extent leadership and issues of trust could encourage EIP. Among other variables, the survey measured perceptions of research use (PoRU), climates of trust (CoT), and innovation (CoI), and climates of research-use (CoRU). The researchers found that French respondents were less favourable toward the idea that theory from research could contribute to their practice but open to research findings contributing. Weak correlations were found between all 4 factors, the strongest were between PoRU and CoRU. School climate that encourages research use was seen as the most important determinant of teacher perceptions of research use and more important than a climate of trust. However, as leaders in France have little instructional influence, the authors recommended establishing norms of research-engagement centrally, in the teaching profession.

Three studies examined research-engagement in the Australian school system ([Mockler and Stacey, 2021](#); [Rickinson et al., 2021](#) and [Prendergast and Rickinson, 2019](#)). There has been an increasing call for the Australian school system to use evidence for improvement in recent years (e.g., [Australian Productivity Commission \(APC\) 2016](#)) and for Australian schools to become research-rich ([Shore et al., 2017](#)).

[Rickinson et al. \(2021\)](#) studied how Australian educators find and use research and evidence. In a survey of 492 practitioners from 414 school across four Australian states, they were interested in finding out how schools were using research evidence, what was involved in using research evidence well and how quality use of research evidence could be developed. They found that research is sourced and used less often than other types of evidence but that school leaders had positive attitudes toward research use, especially when this was contextually relevant. They also found that teachers tended to rely on social methods to engage with evidence; and were concerned with time and access issues. Collaboration with other schools was seen as a key factor in determining uptake of research engagement, along with the contextual specificity of the evidence.

[Prendergast and Rickinson \(2019\)](#) were interested in multiple aspects of research-engagement. In a survey and focus groups involving principals and deputies from Catholic schools in Melbourne, they wanted to know about school leaders' reasons for participating in research projects, how research was used by teachers, and engagement by teachers in research-enquiry. The three most frequent reasons for agreeing to a project were when:

- (i) It was identified as an area of need of development in the school improvement plan
- (ii) the topic was of interest to staff, students, and/or families and:
- (iii) school leaders believed research would produce tangible outcomes.

The principal reason for not agreeing to a project were when demands on the school were considered too great. School leaders surveyed were clear that the most beneficial research projects included a capacity building and professional learning dimension. Regarding research use, schools accessed research in indirect and informal ways including conferences, professional associations and professional publications. However, they conceded that there was infrequent daily use of research to inform practices.

Mockler and Stacey (2021) looked at how teachers understood and engaged with educational research and data with the aim of understanding teachers' professional learning and development needs in relation to research and data literacy. Drawing on data from across two phases of a research project known as Teachers, Educational Data and Evidence-based Practice (2018), they interviewed 21 teachers, recruited via social media, and surveyed 524 respondents in a questionnaire. Acknowledging the bias of this group of self-selected participants, their results found that teachers were generally more well disposed toward research than data, with 63% declaring themselves to be data enthusiasts and 76% research enthusiasts. Over a quarter of the participants identified as data sceptics, while only 12% identified as research sceptics. Ninety-four percent indicated that they regarded teachers' own systematic observation and reflection to be the best source of evidence. Student learning and assessment tasks, interviews/focus group interviews with students, classroom observation by a peer and senior colleague, student surveys and in-class tests were also highly regarded forms of evidence. Less trusted, however, were "walkthroughs" by senior school staff, Higher School Certificate or leaving credential results, other external assessment data and national student test data. Analysis of the values of this group of research enthusiasts showed that some forms of evidence, especially those which were external accountability-driven, could be seen as adversarial and therefore not contribute to improved practice. Teachers valued internally generated evidence that showed how their teaching had improved.

The macro-level considerations about how evidence from research can and should influence practice, comes into studies of other actors in the education system. Segerholm et al. (2021), researched how educational research is used by the Swedish School Inspectorate (SSI). Their research looked at evidence-based governing of the school system and the implications of involving educational research in this process. The research explored, through case studies, two activities performed by the SSI: (1) regular supervision and the establishment of a common knowledge base for that supervision, and (2) The planning, design and conduct of quality audits. With the intent of enhancing research-informed practice, the Swedish government employed staff with investigative backgrounds and listed scholars as research resources for the SSI. Through interviews with school leaders, inspectors and educational researchers they found some tensions in how decisions were arrived at and viewed, through interactions between the various agents (inspectors, teachers, school leaders and researchers). When it came to generating examples of "good practice" in the audits, the researchers felt that it was necessary to problematize this in terms of the context of the practice, while the inspectors did not. Paradoxically, the researchers expressed concerns about their involvement in the inspection process; first about losing their independence and criticality by working for a state agency, secondly, they wanted to emphasize the voice of teachers and protect them from unwarranted criticism.

The Swedish study, as well as the Australian research, shows misalignment in values between accountability dimensions of research-informed practices with those of practitioners. This suggests that need to consider intelligent forms of accountability that complement school improvement efforts. Other clear commonalities in the findings from the above studies include the need to make research relevant and specific to schools, to be presented in accessible and user-friendly formats, to find ways to increase teacher time and staff motivation for EIP and to improve their confidence and capability to translate research into classroom practice.

The exo-system: models for collaborative research-informed school improvement

In addition to the overarching values about professionalism, accountability and institutional regulation and cohesion, the infrastructure of research and development (R&D) also plays a part in influencing the effectiveness of research in improving practices in schools. Several collaborative approaches and models have emerged that attempt to engage multiple participants and stakeholders—and not just researchers, from different organizations, to address problems requiring systemic change. For school leaders, policy makers and researchers in this area, it has become important to make sense of these different forms of EIP "architecture".

Penuel et al. (2020), looked at the similarities and guiding principles of four R&D models in USA, namely: Strategic Education Research Partnerships (SERP), Design-Based Implementation Research (DBIR), Networked Improvement Communities (NICs) and Community-Based Design Research (CBDR). In a day and a half meeting, presentations, and pre-reading of select papers, researchers and experts in these approaches used a Venn diagram to identify principles in common or that were distinctive to each methodology. Eight core principles were identified:

- 1 Participants should collaborate in research
- 2 Roles and contributions of partners to the whole venture need to be clearly defined
- 3 Stakeholders unite around a shared problem
- 4 Research should be sensitive to context
- 5 Research is valuable in that it provides something of practical value to the stakeholders and users of it
- 6 The research should look at intended and actual outcomes (and gaps)
- 7 Should be logical, specific plans and iterations to test out, refine and evaluate the approaches
- 8 The research should be valued by people/bodies outside of the partnership too.

While more research is needed to evaluate the merits of these four models, this study is a useful step in teasing out principles that many have in common and that experts agree lead to successful school improvement. With several R&D approaches vying for attention, the validity of such claims and relative efficacy of the above models can be further tested in relation to these precepts.

School (meso-level) factors affecting research engagement

While the focus of the above research looks at how system level factors influence research and EIP, the influence of school level factors, especially leadership is clearly important. The following section thus focuses on leadership roles and processes that are involved in promoting school research-engagement.

While R&D models such as those described in the above section have the potential to lead to changes in practice informed by research evidence, involvement in such projects may also lead to changes in the way things are done at school, that is, culture. In the Netherlands, Geijssel et al. (2020) examined the extent to which participation in funded R&D projects encouraged the integration of inquiry-based practices. To this end, 28 school leaders of Dutch secondary schools were interviewed retrospectively about their involvement in a series of R&D projects designed to improve teachers research literacy, to use research for accountability and diagnose and analyze problems of practice. School leaders in several of the projects said inquiry-based practices served the development of their schools; and external researchers and school leaders were able to work with teachers to develop their research skills, helping to get answers to school-specific issues. Teachers communicated research in presentations and practice demonstrations to the professional community. Some school leaders planned to apply for further funding for practice-based research and to integrate research as yearlong activity in their school improvement plans. External researchers were able to add specific expertise on research areas such as effective student feedback and organizational capacity was improved through knowledge sharing. In a minority of projects, research did not develop a professional learning community and in one case the problem cited was that the results of the research came too late to feed into teachers' learning.

A leadership innovation to develop research-informed practice in schools, particularly in the English context, is to use teachers, middle or sometimes senior leaders to coordinate research-engagement within the organization and in collaboration with partners. These research champions (RCs) are sometimes funded by the school directly, by virtue of external research grants or in ongoing school university partnerships. In an example of the latter, Burn et al. (2020) gathered data from nine RCs from a collaboration called the Oxford Deanery, linked to Initial Teacher Education (ITE) and Continual Professional Development (CPD). They looked at the strategic intentions of school-based research champions and the activities they carried out and with whom they carried them out. The researchers carried out data collection over 18 months and analyzed it using a cultural historical analysis of motivations and contradictions between the introduction of a new system and the continuation of the old one. Research champions were found to have two broad intentions: creating a research culture and supporting CPD and promoting knowledge mobilization and collaboration. Relationships with senior management were considered crucial, as well as "upward management" of knowledge. Research champions worked to build their own expertise and targeted teachers and departments more so than broader audiences, with little two-way response from teachers to RCs. However, RCs encouraged teachers to use research informed discretion and they performed a gatekeeper role, selecting relevant research for practitioners but also informing the university of what was relevant. Overall, the relational aspects of knowledge-mobilization were apparent in the work of RCs, making them interesting intermediaries in the process and, given the expansion of RCs, an area for further research.

Another innovative approach to bridge the academic-practice gap is to use teachers as embedded practitioner-researchers. In Australia, Blannin et al. (2020) studied schools implementing innovative learning environments (ILEs). Their research formed part of a much larger government initiative to change layouts of schools to improve learning spaces, across Australia. In this study, researchers from universities worked to develop school-based research and solutions. In this three-year study at eight schools, researchers worked with the schools to develop a mixed methodology to find out the school's current knowledge and practice regarding ILE's, to work out how to assist teachers and students in using spatial competencies to improve their teaching and learning and to gather evidence from the project. The study found leadership of embedded research in schools was helped by the support of a senior leader or a research champion. Senior leaders particularly helped with modeling, culture, trust and support through mentoring and time. Although linking teachers' professional development to school improvement was overall seen as a positive, sometimes this faced challenges if teachers were less personally invested in the research.

The above studies show how research-engagement can be promoted and supported by leadership roles in the school, either senior leaders or dedicated research champions and that leadership of research-engagement is also a relational process. Growing research use and enquiry is about changing the school's culture, so that such practices are built into day-to-day dialogue, planning and professional learning. The research itself, as already mentioned above, must not only be context specific but also timely if it is to feed back into practice in and effect change.

The strength of the evidence

Below, some of the evidence base on school research-engagement is highlighted from recent publications. This looks firstly at empirical support for linking research-engagement to school improvement, and then two systematic reviews of the literature are examined, one on the role of teachers in evaluating their own teaching programmes and another looking at the strength of the evidence underpinning various strategies for research use.

Perhaps the most difficult question to answer with certainty is whether and to what extent research-engagement and school improvement are linked. Bernhard et al (2020) interviewed 22 Head teachers, in seven deprived Boroughs of London to examine whether any such link could be found. Using a conceptual framework informed by the dynamic approach to school improvement (DAIS) (Creemers and Kyriakides, 2006) and research-engaged schools (Godfrey and Brown, 2018), they found that the most improved schools exhibited many of these characteristics. Two-thirds of their sample were found to be research-engaged and most of the other schools put a strong emphasis on effective professional development, which was itself guided by academic

evidence. In these schools, teachers and school leaders directed their work toward improving students' learning, including using learning walks. Many were strongly inspired by research and aligned their practices according to what they had read from the academic literature.

In other recent research, [Schenke et al. \(2017\)](#) found, from case studies of 19 secondary schools in the Netherlands, several positive outcomes from involvement in practice-based research. They looked at new knowledge created by performing research, which fell into three domains: innovation, teachers' practice, and the quality of the school organization. In 13 of the 19 projects research results were used to inform school development; and in some cases, changes to teaching practice were observed, such as in the use of feedback. Their research suggested overall that research-engagement played a strong part in professional development and inspiring improvement in these schools. Both this and the above study by Bernhard and colleagues corroborate previous research showing a moderate link between research engagement and student outcomes ([Rose et al., 2017](#)).

One purpose of involving teachers in their own research is to evaluate the effectiveness of their teaching or programmes. Using [Patton's \(2008\)](#) definition of evaluation as the systematic investigation of the effectiveness, worth, or merit of a program, practice or policy, [McFadden and Williams \(2020, p. 1\)](#), systematically reviewed the current literature. One of their questions was how such evaluative capacity could be developed in teachers. They found very few quantitative studies, nine were qualitative, none were experimental, and one quasi-experimental; three used pre and post design; three used cross sectional quantitative designs and two used mixed methods. The most common form of capacity building was support for teachers to undertake small scale research, including forming inquiry communities with university mentors and mostly conducting action research. Eight studies showed pre- and in-service teacher coursework and one study looked at research mentorship. Only one took a whole school approach to research, and this was perhaps a special case, being a Professional Development School in the Netherlands. Teacher skills development included: sourcing and evaluating research findings (research literacy); attitudes, knowledge, skills, and confidence in research processes; professional competences or reflective practice skills; teacher identity and self-efficacy; and collaboration, professional conversations, and collegial relationships. In this systematic review, the authors show a logic model in which short term changes from these activities were the development of increased skills and positivity toward research use, then in the medium term there was an increased use of a systematic research base and an empowered teacher identity, and finally more effective teaching was evident through increase in evidence-based practices. Important enablers for in-service teachers included release time, opportunities for collaboration and minimizing workload issues. An environment of research engagement and support from universities also increases confidence. Overall, this review supports the idea that school research-engagement can be intentionally nurtured and grown, and that intermediate markers that can be monitored along the way by school leaders to check its progress; these conclusions confirm findings in previous work (e.g., [Godfrey, 2016](#) and [2019](#)).

Which approaches most effectively encourage (high-quality) research use in education has not up to now been thoroughly evaluated. [Gorard et al. \(2019\)](#) therefore conducted a systematic review to answer this question, examining 323 studies from multiple sectors, screening for quality on: design; scale; missing data; measurement quality; and threats. They found very few robust studies, and most came from outside the education sector, very few showed the counterfactual needed to show what would have happened if the approach had not been used. In their research they show nine variants of evidence use (a–i) on the two dimensions: the extent to which this was modified, that is, plain (e.g., open access to journals), modified (e.g., EEf toolkit) and engineered (lesson plans/scripts), and the extent to which it was actively transferred, that is, passive (reading research), engagement in transfer (e.g., research champions or brokers), active, in transfer (e.g., teacher research). Their findings are summarized below: ([Table 3](#))

The above table shows two things, first the paucity of evidence in many areas suggests much more research is needed on research use in education. Secondly, only approaches c, h and especially f look particularly promising in the educational realm. Generally speaking, and especially for complex changes, active approaches to evidence transfer appear the most effective, albeit sources of passive evidence may play a first useful stage in a later more active process involving engagement or active transfer. Research on using a research-brokering initiative in Ireland ([McGann et al., 2020](#)) is a case in point. The T-REX initiative (teachers research

Table 3 Summary of systematic review of evidence on research use.

	<i>Passive transfer</i>	<i>Engagement in transfer</i>	<i>(Inter) active transfer</i>
Plain evidence	a does not work	b no evidence of change or improvement to practice	c limited evidence that this helps target research at concerns and because staff learn more about research use
Modified evidence	d evidence of engagement from downloads but not of changing practice.	e no evidence using knowledge brokers helps research use. May only work where change is very simple and partly accepted already.	f the most promising evidence. An audit approach and feedback/coaching is effective way to get evidence into use.
Engineered evidence	g no good evidence in education but feasible could work. However, concerns teachers might feel they are losing personal autonomy, for example, if given scripts in the way that nurses are given clinical guidelines.	h some evidence behavioral nudges/reminders can be effective but mostly as quick fix solutions.	i no good evidence however possible to legislate for EIP to be used but must only be most rigorous studies.

exchange) has a membership made up of over 2500 educational professionals from early years to higher education, with the aim of promoting communities of practice, inter professional dialogue through discussion, interviews, clips, and research stories. An example of cell e above, this may lead to more active engagement or transfer of research over time, without showing such an effect in the early stages.

Some caution needs to be attached to Gorard et al.'s methodology however, in that they use a particularly narrow definition of "rigorous" research within a *what works* paradigm (see above debate about this issue). Furthermore, they are looking at demonstrable effects over often short time periods of time on student (or other equivalent end user) outcomes. Research on whole school approaches to research-engagement tends to indicate cultural and structural changes which may take many years (Godfrey, 2017b), so strategies may become more, or less effective or appropriate over time. For instance, if teachers become more research literate, then "plain" evidence may have a greater effect and legislating the use of highly engineered, evidence-based practices could be counter-productive as it would be seen as impinging on teachers' professional autonomy.

Conclusions

The research linking research-engagement to specific school improvement outcomes is a burgeoning one and warrants further attention albeit this will remain a difficult one to support conclusively. As this research tends to be correlational, conducting longitudinal research where schools intentionally develop research-engagement and tracking trajectory over time may help illustrate this further. This approach is more feasible than comparing matched schools given the complexities of controlling the myriad of variables involved. Studies on research use are still lacking in education and more should be conducted testing the efficacy of various approaches. This research should however, employ a wide range of outcome measures, including at teacher level, and not rely solely on standardized student attainment data. From the little research so far, active approaches to research use look the most promising. This will come as little surprise to many who have worked with schools to apply research in practice in collaboration with practitioners and to develop cultures of research-engagement. Nevertheless, the principles of encouraging active research use still warrant further attention and testing of for the validity of their claims. The above work by Farley-Ripple et al. (2020) suggests there would be much to gain by bringing together different communities to understand each other's research traditions, terminology and contributions to an interdisciplinary conversation about research use. As the authors state, this will require coordination, investment, and opportunities to share learning.

Throughout most of the studies above, there is a recurring and common theme too: leadership of research-engagement is important, it takes time and is about changing the school culture. There are also structural factors built into school life and the teaching profession that are both easy to understand and difficult to resolve – teachers need time to be research-engaged and these practices need to be valued, and supported across the system.

References

- Ainscow, M., Muijs, D., West, M., 2006. Collaboration as a strategy for improving schools in challenging circumstances. *Improv. Sch.* 9 (3), 192–202.
- Australian Productivity Commission (APC) (2016). National education evidence base: report no. 80. Accessed online (11th Oct 2021) at: <https://www.pc.gov.au/inquiries/completed/education-evidence/report>.
- Bernhard, R., Burn, K., Sammons, P., 2020. Research-based School Improvement: How Highly Effective and Improving Schools in England Apply Research. American Educational Research Association online paper repository.
- Biesta, G., 2017. Education, measurement and the professions: reclaiming a space for democratic professionalism in education. *Educ. Philos. Theor.* 49 (4), 315–330.
- Biesta, G.J., Burbules, N.C., 2003. *Pragmatism and Educational Research*. Rowman & Littlefield.
- Blannin, J., Mahat, M., Cleveland, B., Morris, J.E., Imms, W., 2020. Teachers as embedded practitioner-researchers in innovative learning environments. *CEPS J.* 10 (3), 99–116.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press, Cambridge, MA.
- Brown, C., Schildkamp, K., Hubers, M.D., 2017. Combining the best of two worlds: a conceptual proposal for evidence-informed school improvement. *Educ. Res.* 59 (2), 154–172.
- Bryk, A.S., 2015. 2014 AERA distinguished lecture: accelerating how we learn to improve. *Educ. Res.* 44 (9), 467–477.
- Burn, K., Conway, R., Edwards, A., Harries, E., 2020. The role of school-based research champions in a school–university partnership. *Br. Educ. Res. J.* 47 (3), 616–633.
- Camburn, E.M., Goldring, E., Sebastian, J., May, H., Huff, J., 2016. An examination of the benefits, limitations, and challenges of conducting randomized experiments with principals. *Educ. Adm. Q.* 52 (2), 187–220.
- Collins, K., Coleman, R., 2021. Evidence-informed policy and practice. In: *School Leadership and Education System Reform*, p. 19.
- Creemers, B.P.M., Kyriakides, L., 2006. Critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model. *Sch. Effect. Sch. Improv.* 17 (3), 347–366. <https://doi.org/10.1080/09243450600697242>.
- Dimmock, C., 2019. Leading research-informed practice in schools. In: *An Ecosystem for Research-Engaged Schools: Reforming Education Through Research*. Routledge, London, pp. 56–72.
- Dimmock, C., 2016. Conceptualising the research–practice–professional development nexus: mobilising schools as 'research-engaged' professional learning communities. *Prof. Dev. Educ.* 42 (1), 36–53.
- Ehren, M., 2019. Accountability structures that support school self-evaluation, enquiry and learning. In: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 41–55.
- Farley-Ripple, E.N., Oliver, K., Boaz, A., 2020. Mapping the community: use of research evidence in policy and practice. *Human. Social Sci. Commun.* 7 (1), 1–10.
- Friedman, V.J., Lipshitz, R., Popper, M., 2005. The mystification of organizational learning. *J. Manag. Inq.* 14 (1), 19–30.
- Gaussel, M., MacGregor, S., Brown, C., Piedfer-Quêney, L., 2021. Climates of trust, innovation, and research use in fostering evidence-informed practice in French schools. *Int. J. Educ. Res.* 109, 101810.
- Geijsel, F., Schenke, W., van Driel, J., Volman, M., 2020. Embedding inquiry-based practices in schools: the strategic role of school leaders. *Eur. J. Educ.* 55 (2), 233–247.
- Glatzer, R., 2015. Are schools and colleges institutions? *Manag. Educ.* 29 (3), 100–104.

- Godfrey, D., 2020. From peer review to collaborative peer enquiry: action research for school improvement and leadership development. *Lond. Rev. Educ.* 18 (3), 373–389. <https://doi.org/10.14324/LRE.18.3.04>.
- Godfrey, D., 2019. Moving forward—how to create and sustain an evidence-informed school ecosystem. In: *An Ecosystem for Research-Engaged Schools*. in: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 202–219.
- Godfrey, D., 2017a. What is the proposed role of research evidence in England's 'self-improving' school system? *Oxf. Rev. Educ.* 43 (4), 433–446.
- Godfrey, D., 2017b. Exploring Cultures of Research Engagement at Eight English Secondary Schools. Doctoral thesis. UCL (University College London).
- Godfrey, D., 2016. Leadership of schools as research-led organisations in the English educational environment: cultivating a research-engaged school culture. *Educ. Manag. Adm. Leader* 44 (2), 301–321.
- Godfrey, D., Brown, C., 2019. Innovative models that bridge the research—practice divide. In: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 91–107.
- Godfrey, D., Brown, C., 2018. How effective is the research and development ecosystem for England's schools? *Lond. Rev. Educ.* 16 (1), 136–151. <https://doi.org/10.18546/LRE.16.1.12>.
- Godfrey, D., Handscomb, G., 2019. Evidence use, research-engaged schools and the concept of an ecosystem. In: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 4–21.
- Gorard, S., Griffin, N., See, B.H., Siddiqui, N., 2019. How Can We Get Educators to Use Research Evidence? Lulu Press, US, 193 pages.
- Gu, Q., Rea, S., Seymour, K., Smethem, L., Bryant, B., Armstrong, P., Ahn, M., Hodgen, J., Knight, R., 2020. The research schools network: supporting schools to develop evidence-informed practice. Accessed online (10th October 2021) at: https://d2tic4wv01iusb.cloudfront.net/documents/projects/RS_Evaluation.pdf.
- Handscomb, G., MacBeath, J., 2003. *The Research-Engaged School*. Essex County Council.
- Hood, C., 1998. *The Art of the State, Culture Rhetoric and Public Management*. Oxford University Press.
- Kools, M., Stoll, L., George, B., Steijn, B., Bekkers, V., Gouëdard, P., 2020. The school as a learning organisation: the concept and its measurement. *Eur. J. Educ.* 55 (1), 24–42.
- Ladd, H.F., Fiske, E.B., 2003. Does competition improve teaching and learning? evidence from New Zealand. *Educ. Eval. Pol. Anal.* 25 (1), 97–112.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge university press.
- Lortie-Forgues, H., Inglis, M., 2019. Rigorous large-scale educational RCTs are often uninformative: should we be concerned? *Educ. Res.* 48 (3), 158–166.
- Malik, S., 2016. Knowledge Mobilization in Ontario: A Multi-Case Study of Education Organizations. Doctoral dissertation. University of Toronto, Canada.
- Malin, J.R., Brown, C., Ion, G., van Ackeren, I., Bremm, N., Luzmore, R., Flood, J., Rind, G.M., 2020. World-wide barriers and enablers to achieving evidence-informed practice in education: what can be learnt from Spain, England, the United States, and Germany? *Human. Soc. Sci. Commun.* 7 (1), 1–14.
- McFadden, A., Williams, K.E., 2020. Teachers as evaluators: results from a systematic literature review. *Stud. Educ. Eval.* 64, 100830.
- McGann, M., Ryan, M., McMahon, J., Hall, T., 2020. T-REX: the teachers' research exchange. overcoming the research-practice gap in education. *TechTrends* 1–14.
- Mockler, N., Stacey, M., 2021. Evidence of teaching practice in an age of accountability: when what can be counted isn't all that counts. *Oxf. Rev. Educ.* 47 (2), 170–188.
- Penuel, W.R., Riedy, R., Barber, M.S., Peurach, D.J., LeBouef, W.A., Clark, T., 2020. Principles of collaborative education research with stakeholders: toward requirements for a new research and development infrastructure. *Rev. Educ. Res.* 90 (5), 627–674.
- Patton, M., 2008. *Utilization-focused Evaluation*. SAGE, Thousand Oaks, CA.
- Pollock, K., Campbell, C., McWhorter, D., Bairos, K., van Roosmalen, E., 2019. Developing a system for knowledge mobilisation. In: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 22–40.
- Prendergast, S., Rickinson, M., 2019. Understanding school engagement in and with research. *Aust. Educ. Res.* 46 (1), 17–39.
- Price, H.E., Weatherby, K., 2018. The global teaching profession: how treating teachers as knowledge workers improves the esteem of the teaching profession. *Sch. Effect. Sch. Improv.* 29 (1), 113–149.
- Rickinson, M., Gleeson, J., Walsh, L., Cutler, B., Cirkony, C., Salisbury, M., 2021. Research and evidence use in Australian schools: early insights from educators. Accessed online (10th October 2021) at: <https://apo.org.au/sites/default/files/resource-files/2021-03/apo-nid311651.pdf>.
- Rose, J., Thomas, S., Zhang, L., Edwards, A., Augero, A., Roney, P., 2017. *Research Learning Communities: Evaluation Report and Executive Summary*. Education Endowment Foundation.
- Schechter, C., Mowafaq, Q., 2013. From illusion to reality: schools as learning organizations. *Int. J. Educ. Manag.* 27 (5), 505–516.
- Schenke, W., van Driel, J.H., Geijssels, F.P., Volman, M.L., 2017. Closing the feedback loop: a productive interplay between practice-based research and school development through cross-professional collaboration in secondary education. *Prof. Dev. Educ.* 43 (5), 860–880.
- Schön, D.A., 1983. *The Reflective Practitioner. How Professionals Think in Action*. Ashgate publishing limited, London.
- Segerholm, C., Lindgren, J., Novak, J., 2021. Evidence-based governing? educational research in the service of the Swedish schools inspectorate. *Scand. J. Educ. Res.* 1–16.
- Shore, S., White, S., Nuttall, J., Mills, M., Down, B., Woods, A., Bussey, K., 2017. A research-rich teaching profession: insights from the field. In: *Paper Presented at the Australian Association for Research in Education Conference*, Canberra.
- Tipton, E., Spybrook, J., Fitzgerald, K.G., Wang, Q., Davidson, C., 2021. Toward a system of evidence for all: current practices and future opportunities in 37 randomized trials. *Educ. Res.* 50 (3), 145–156.
- Wisby, E., Whitty, G., 2019. Maintaining (ecosystems for) a broad view of educational research and its relationship to practice. In: *An Ecosystem for Research-Engaged Schools: Reforming Education through Research*. Routledge, London, pp. 187–201.

Understanding turnaround leadership research: continuity and change (2017–2021)

Peng Liu^a and Yalong Bo^b, ^a University of Manitoba, Winnipeg, MB, Canada; and ^b University of Hong Kong, Hong Kong, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	158
Conceptual framework	158
Data sources	159
Method	159
Findings	159
Turnaround as a concept	159
Turnaround as a condition	160
Internal conditions	160
External conditions	161
Turnaround as a process	161
Leadership as a key to reintegration	161
Changing leadership as a key to turnaround	162
Rationale for leadership changes	162
Restrains on turnaround leaders	162
Political environment	162
Resistance from the school and the community	162
Inability to perform	163
Antecedents and factors affecting turnaround	163
Creating a collaborative school culture and structure	163
Principals' role in school turnaround	163
Leadership practices in school turnaround	164
Getting to know the terrain	164
Setting up a new vision and planning	164
Building trust and demonstrating respect	164
Improving communication	164
Turnaround leadership competence	165
District leadership in turnaround contexts	165
Types of leadership for successful school turnaround	165
Conclusions	166
References	167

Introduction

Turnaround leadership is an important research field, but not many scholars have examined the areas of continuity and change in this field since Liu (2020). Adopting Murphy's (2008) turnaround leadership review framework, this review fills this gap by exploring turnaround as a concept, a condition, a process, and a consequence in relevant research published between 2017 and 2021. This review also explores other aspects of turnaround leadership: leadership as a key to reintegration, why leadership change is necessary in low-performing schools, restraints on turnaround leaders, collaborative school culture and structure, antecedents of turnaround leadership, and the characteristics of effective turnaround leadership. In addition to documenting the ongoing efforts that have been made in these lines of research, this review also finds that there has been no significant progress in regard to school leaders as the cause of problems or the replacement of managers as signaling. New themes that have emerged in recent research include the role of principals in school turnaround, effective turnaround leadership practices, types of leadership, and the effects of district leadership in transforming turnaround schools.

Conceptual framework

Hallinger (2013) believed in the importance of limiting a review to a specific period for historical and pragmatic reasons. Liu (2020) concluded that more publications in turnaround leadership had emerged and prospered after Murphy's (2008) comprehensive review on the topic of turnaround leadership. Continuing Liu's (2020) efforts, this review focused on the relevant literature from 2017 to 2021.

Data sources

This review examined English-language book sections, peer-reviewed journal articles, and master's and doctoral students' theses on turnaround leadership in the education sector published between 2017 and 2021. In total, 44 publications were reviewed, including 28 peer-reviewed journal articles, sections or chapters from seven books, and eight doctoral theses. Multiple digital research databases such as ProQuest, ERIC, and EBSCOhost were searched for the aforementioned data sources. In addition, the researchers used Google Scholar to assist in the literature search. Guided by the literature collection strategy of [Murphy \(2008\)](#) and [Liu \(2020\)](#), key words used to identify the relevant publications included "turnaround leadership," "turnaround leaders," "principalship in school turnaround," "turnaround school leaders," "turnaround school principals," and "leaders/leadership in low-performing school."

The 28 journal articles come from the following international journals: *Journal of Educational Administration*, *Asia Pacific Journal of Education*, *Educational Administration Quarterly*, *Journal of School Leadership*, *Journal of Cases in Educational Leadership*, *AERA Open*, *Education Policy Analysis Archives*, *International Journal of Educational Management*, *Leadership and Policy in Schools*, *School Effectiveness and School Improvement*, *Journal of Education for Students Placed at Risk (JESPAR)*, *Journal of World Englishes and Educational Practices*, *Africa Education Review*, *NASSP Bulletin*, *Journal of Educational Administration*, *Journal of Educational Administration*, *International Perspectives on School Settings*, *Education Policy and Digital Strategies*, *School Leadership and Management*, *Educational Management Administration & Leadership*, *On the Horizon*, *International Perspectives on Leading Low-Performing Schools*, *School Leadership & Management*, and *Leadership & Policy in Schools*.

During the period between 2017 and 2021, multiple scholars published books on the topic of turnaround leadership in the educational context. [Hochbein and Mahone \(2017\)](#) published their first book in a series on school turnaround to help researchers and practitioners in the field to see through a few common and long-existing misconceptions (referred to by the authors as "myths") to better understand how certain schools decline or even fail in terms of performance, and, more importantly, how to more effectively reform low-performing schools. In the same year, [Hitt and Meyers \(2017\)](#) published their book and introduced some promising leadership practices that can produce school improvements. Another book, titled *International Perspectives on Leading Low-Performing Schools* and edited by [Meyers and Darwin \(2018\)](#), was one of the most significant publications during this period. Scholars from all over the world contributed chapters that shed light on the turnaround leadership in their respective countries or regions, such as Japan, China, Costa Rica, Chile, Germany, Malaysia, Indonesia, etc. Such literature helps researchers and practitioners learn from both successful and ineffective leadership cases from all over the world. 2019 saw the publication of two important books by [Murphy and Bleiberg \(2019\)](#) and [Petrich \(2019\)](#); the former reflected on a number of ineffective school turnaround policies and practices and identified some common barriers that may hinder or jeopardize school turnaround attempts, while the latter focused on the leadership practices of multiple elementary school principals who had succeeded in turning their schools around. Most recently, [Klein et al. \(2021\)](#) discussed, in a book chapter, effective leadership in schools that serve disadvantaged communities in countries such as the United States and Germany.

Method

The method used in this study was based on [Hallinger's \(2013\)](#) suggested considerations for a review:

1. What are the guiding questions?
2. What are the main conceptual perspectives that direct the selection, evaluation, and interpretation of studies?
3. What sources and types of data are adopted in the review?
4. How are data evaluated, analyzed, and synthesized in the review?
5. What are the major results, limitations, and implications of the review?

This study used [Murphy's \(2008\)](#) method to understand and synthesize the literature. The analysis was conducted using the following procedures. The initial analytical framework was derived from the main themes in [Murphy \(2008\)](#). These themes were coded using capital letters from A to K. Sub-themes were coded with numbers. For instance, "Setting direction" was initially coded as "4M," where the letter "M" represented the theme of "District leadership practices" used in transforming turnaround schools. The number 4 represented a specific district leadership practice. Then, the initial codes were grouped with other similar codes to form larger themes. For instance, "Setting direction" (4M) was grouped with other codes such as "Communication" to form the higher-level theme "District leadership practices." This particular example fell within [Murphy's \(2008\)](#) themes, but new themes were also added to the initial analytical framework as they emerged.

Findings

Turnaround as a concept

School "turnaround" is not a new or revolutionary notion. The scientific investigation of this topic can be traced back to organizational change and school effectiveness research about four decades ago ([Gerade, 2019](#)). Similar to many other sociological constructs, there is no universally agreed-upon definition for school turnaround ([Harris et al., 2018](#); [Hochbein and Mahone, 2017](#)). However, it is generally agreed to be different from regular school improvement or reform, in that it aims to enable drastic

change in schools with continuous underperformance in terms of students achievement and school development (Hochbein and Mahone, 2017; Meyers and Hitt, 2017). Hitt and Meyers (2018, p. 7) quantified turnaround as “a rapid and dramatic improvement in a school within the bottom 5% of schools that demonstrates at least 20% point gains in 2 years in both math and English Language Arts.” Mania-Singer (2018) pointed out that, according to the available literature, school turnaround denotes quick and dramatic change, but that in reality an increasing body of research has found that it generally takes a couple of years before meaningful and significant change (p. 139) is noticeable.

Citing Hochbein (2012) and other sources such as Collins (2009), Senge (1990), Pearson et al. (2015), and Gatewood and Gynelle (2020) summarized school decline into three scenarios: (1) a single event that sparks a significant decrease in performance, (2) a series of interconnected events that send a school into a gradual but noticeable downward spiral, and (3) an undetected slow descending trajectory (p. 46). Gatewood and Gynelle (2020) studied the decline of a school in the United States where the majority of the students were African American. They learned that, internally, a drop in the number of student admissions will have an impact on funding, which affects multiple areas such as the number of educational programs the school can offer and the number of teachers and managers the school can afford to hire (Gatewood and Gynelle, 2020, p. 105). Externally, changes that have occurred in the community can influence the performance of a school as well. Such influencing factors can include both inward and outward mobility of community members for reasons such as property values and taxes, especially when such mobility changes the racial or ethnic diversity of the neighborhood (Gatewood and Gynelle, 2020).

More often than not, school turnaround is a political or institutional decision or campaign rather than a voluntary endeavor undertaken by schools and their stakeholders. Turnaround has become a prevailing theme in the education policy of the United States as a response to the political and moral call to rescue continuously underperforming schools (Meyers and Smylie, 2017). In particular, in the United States, transforming turnaround schools has become an important political agenda (Mania-Singer, 2018; Meyers, 2020; Pham et al., 2020; Schueler, 2019). Although not as systematically developed, improving low-performing schools, especially those serving disadvantaged communities, has become one of the major foci of educational policy making in Germany (Klein et al., 2021). Furthermore, Tao and Liu (2020) reported that school turnaround in China has relied on school leaders’ developing an adequate understand the relationship dynamics between their schools and relevant governmental departments or institutions.

The continuity in this line of research has been that turnaround has mainly been viewed as a management concept and as representing a decline in the organization. Previous reviews (Liu, 2020; Murphy, 2008) have identified a relatively new theme in the area of turnaround as a concept: educational reform and turnaround as democratic processes. In addition, this review found a new trend: turnaround as a political concept. For instance, scholars have begun to describe decline as both a single event and an interconnected series of events, such as change in the demographic composition of a larger community triggering change in the student population of a school (Gatewood and Gynelle, 2020). In this approach to turnaround, school leaders need to have a good understanding of the political and policy situation in their country.

Turnaround as a condition

A large volume of the school turnaround literature has mainly focused on current low performance in a school and its potentially better state in the future, rather than on its troubled past (Meyers and Smylie, 2017). In comparison, much less literature is available that has discussed the factors that contribute to school decline and low performance (Gatewood and Gynelle, 2020; Harbatkin, 2020). To fill this gap, this section reviews literature that shines some light on the conditions that cause schools to be in need of turnaround.

Internal conditions

Internally, one of the essential factors that may cause a school to have performance challenges is related to the central focus of all educational activities—the students. Such complications may be the result of behavioral problems or high student turnover caused by a change in the racial or ethnical diversity of the student population (Harbatkin, 2020), or difficult conditions, such as socio-economic struggles or other life challenges, that students face outside the school walls (Coby, 2018). In addition to students’ personal challenges, risks factors at the school level can also have negative impacts on the performance status of a school, such as challenges related to school culture and climate, funding and resources, and the human side of school—factors related to teachers and administrators.

An organization’s culture and climate can greatly influence its performance and the effectiveness of its leaders, and schools are no exceptions (Harris et al., 2018). School culture and climate may contribute negatively to the development of a school and become the key factors resulting in its turnaround status (Whyte, 2018). Harbatkin (2020) discovered low or decreased trust among teachers and between teachers and students, the negative status of teacher leadership or school leadership, a low level of community support and involvement in school administration, and a high or increased level of violent behaviors among students may signal a drop in a school’s performance (pp. 86–87).

In addition to the intangible yet significantly influential factors that characterize school climate and culture, many low-performing schools are also economically disadvantaged with poor facilities and technological supports (Harris et al., 2018). However, Meyers and Smylie (2017) pointed out that superficial changes and improvement in the physical facilities of a school do not necessarily lead to fundamental and sustainable changes, and that schools need meaningful changes in their organizational behaviors if increased funding is to be effective.

Besides the environmental factors discussed above, the human side of the school organization plays a role too, i.e., teachers and school administrators. Such variables can include student-teacher ratio, the proportion of new teachers in a school, turnover of teachers as well as administrators, wages and other forms of compensation for teachers and principals, and so forth (Harbatkin, 2020). Furthermore, low teacher qualifications or the incompetency of school management or leadership can jeopardize students' learning outcomes either directly or indirectly through decreasing teachers' work motivation or commitment (Coby, 2018; Meyers and Hitt, 2017).

External conditions

Aside from the conditions and restraints within schools, there are also forces from the outside that, more often than not, are beyond the control of the school itself. Such external conditions can be poverty in a school's community, change in student composition especially in racially or ethnically diverse societies, and the imbalanced distribution of education resources (Harris et al., 2018). In their review of the relevant literature, Arar and Örüçü (2021) identified several other external unfavorable factors ranging from terrorism and high crime rates in the community, to students' diverse religious and cultural backgrounds, and rapid societal change.

The continuity in this line of research has focused on funding and resources (Harris et al., 2018; Meyers and Smylie, 2017). Research has confirmed that funding and resources are important for the successful transformation of turnaround schools, but that change in organizational and people-related factors may also significantly contribute to such transformation.

To sum up, as reviewed above, new developments in the internal factors related to transforming turnaround schools have centered on the human side of schools (Meyers and Sadler, 2018; Meyers and Hitt, 2017) and school culture and school climate (Harris et al., 2018). However, in the line of research, detailed strategies for developing school culture and school climate need to be further explored. Importantly, human-related factors like teacher motivation and well-being also require further exploration. Furthermore, extensive research on the roles of the external community in transforming turnaround schools is needed using diverse research methods (qualitative, quantitative, and mixed).

Turnaround as a process

There is a rich volume of literature describing the process of school turnaround. As mentioned frequently in the turnaround leadership literature, Leithwood et al. (2010) proposed breaking it down into three stages: stopping the decline and creating conditions for early improvements, ensuring survival and realizing early performance improvements, and reaching positive performance (Gerade, 2019; Liu, 2020).

Inspired and informed by such a framework, Yoon and Barton (2019) examined the notion of time in the process of school turnaround and proposed an innovative approach to school turnaround that the authors referred to metaphorically as *shifting gears*. In the first gear, turnaround leaders start by "repairing and rebuilding school climate and culture"; in the second gear, they "push for instructional initiatives and structures" to facilitate student learning; finally, they should try to imbed the instructional and cultural transformation into the school's DNA, so it can be sustained in the long run (Yoon and Barton, 2019, pp. 695–696). Similar to Yoon and Barton's (2019) final two steps, Hitt and Meyers (2018) emphasized two stages in the process of school turnaround: turnaround and sustainability. The "turnaround" phase demands school leaders' dramatic intervention, and the "sustainability" phase requires the continued use of both turnaround leadership practices and a gradual blending in of other effective leadership practices.

Similarly but more specifically, Harris et al. (2018) also introduced their summary of the turnaround process in Malaysia and Indonesia. At the beginning, principals attach importance to multiple aspects of reform such as institutional reorganization, the renewal of teaching staff, changing the school culture, and creating clear disciplines. During the more advanced phases of school turnaround, principals then concentrate on supporting teachers, increasing collaboration, and offering feedback about the improvement of instructional strategies. Furthermore, principals will also focus on monitoring students' engagement and learning outcomes.

The continuity in this line of research has been a focus on the conceptualization of the turnaround process in low-performing schools (Gerade, 2019; Hitt and Meyers, 2018; Liu, 2020; Yoon and Barton, 2019). The progress that has been made in this research has focused on the details of the change process. For instance, Harris et al. (2018) specified the precise strategies school leaders can use to transform turnaround schools. Additionally, it is notable that Hitt and Meyers (2018) described the sustainability stage, drawing the attention of turnaround leaders to combining varied leadership styles to improve their education practices. In future research, exploring the relationship between leadership styles and student achievement might deepen this line of research. In addition, further international studies on how to transform low-performing schools would solidify the knowledge base about effective change across cultures.

Leadership as a key to reintegration

Leadership plays an essential role in school turnaround and improvement (Liu, 2020). There has been a rich body of international literature with empirical evidence that has attested to the importance of leadership for improving students' achievement outcomes and the overall performance of a school in multiple ways (Leithwood et al., 2008; Lochmiller Chad and Chesnut Colleen, 2017).

Although there has been a large amount of interest in more democratic leadership styles such as distributed leadership, teacher leadership, and so forth, principals are still the central piece of educational leadership (Bush, 2020). Principals with proven talent and successful experience in school management and leadership play an undeniable role in turning around the performance of a school (Harbatkin, 2020; Meyers and Hitt, 2017).

Without a talented and experienced turnaround leader, struggling schools are not very likely to be turned around; therefore, a turnaround principal, together with the leadership styles and strategies that he or she chooses to adopt, serves as a catalyst that helps to stop a school's downward trajectory and to lift it up in a more positive direction (Wana, 2019).

Changing leadership as a key to turnaround

Strategic replacement of school leaders and sometimes part of the teaching faculty is a common measure in many school turnaround cases (Meyers and Smylie, 2017; Wana, 2019). This practice is believed to bring about the "necessary drastic disruptions" that are considered the "*sine qua non*" [italics in original] of school turnaround (Meyers and Smylie, 2017, p. 506). For example, Pham et al. (2020) evaluated the sustainability and maturation of school turnaround in Tennessee where replacement of the principals was a major component of the turnaround strategy. While agreeing with the necessity of such means, Meyers and Smylie (2017) also criticized that a change in leadership should not mistakenly be equated with the actual turning around of the suffering school. The continuity in this line of research has centered on the significance of replacing school principals in turnaround schools, but further research is required to explore the antecedents of successful turnaround leadership and how leaders can effectively empower teachers in the transformation process. Also, further exploration of the effects of replacing school principals would contribute to this field.

Rationale for leadership changes

As mentioned above, it is a common practice to replace the principal of a continuously struggling school, due to the frequent assumption that the original leaders in a struggling school may have become "too entrenched in the existing culture" of the school to bring about effective changes (Mania-Singer, 2018, p. 140). Instead, effective turnaround leaders "possess the knowledge and skills to stimulate quick and profound change that disrupts the original practices of the teachers and improves student learning outcomes" (Woulfin and Weiner, 2019). However, just like new CEOs do not always improve the performance of struggling companies according to management literature, a new principal hired from outside may not necessarily lift a poorly performing school out of its current problematic condition (Meyers and Smylie, 2017).

While short-term increases have been documented by researchers, Meyers and Smylie (2017) asserted that evidence of long-standing sustainable success is rare, unless the new principals deeply engage themselves in "the fundamental changes in school organization, operations and practices" that are essential to ensure that the new improvements last instead of withering away following the initial short-lived excitement of having a new leader (p. 510). However, Same (2017) also warned that we should not solely rely on certain leaders dropped in from the outside to magically save a school, but that all stakeholders of the school should collaborate with one another to ensure the success of the turnaround. The continuity in this line of research lies in the significance of the replacement of school principals in turnaround schools. The progress in this line of research has been recognition of the roles of different stakeholders in school organizations. This goes beyond the assumption that turnaround leadership only comes from the school principal.

Restraints on turnaround leaders

In both Western and Asian contexts, as mentioned above, prior successful leadership experiences cannot be mechanically transferred from one school to another with guaranteed results (Harbatkin, 2020; Paletta et al., 2017). In spite of the ambitious expectations and timelines in many school turnaround situations, turnaround leaders need to be aware of the restraints they face when diving in to turn around a low-performing school.

Political environment

The political environment from the national level to the local district level can have rather strong impacts on the turnaround of a low-performing school, in that it can influence both policy development on a macro level and the actual school district- or even school-level implementation of the turnaround measures as well as the principal's behaviors (Schueler, 2019). For example, China has adopted a highly centralized and top-down education governance model under which the government is in a hierarchically higher position and may enact certain restrictions over turnaround principals' power (Tao and Liu, 2020).

Resistance from the school and the community

It takes time for teachers and students in a school whose principal is reassigned, as well as the parents in the community, to understand and accept a new turnaround principal and/or his or her new visions and measures for the school (Hitt and Meyers, 2017; Same, 2017; Whyte, 2018). In a case study on turnaround leadership effectiveness conducted by (Same, 2017), the majority of the

participating turnaround leaders recalled resistance or resentment from both teachers and parents. Therefore, a turnaround leader cannot mechanically change everything, but “provide a synthesis of the past and the present to render sustainable change” (Whyte, 2018, p. 19).

Inability to perform

A turnaround principal’s effectiveness is to a large extent determined by his or her competence. This competence includes the principal’s own disposition as well as prior experience. A person’s experience can be the key to his or her success or the very reason that leads to setbacks; therefore, Hitt et al. (2018) regarded a turnaround principal’s prior experience as an essential contributing factor to the effectiveness of a turnaround principal. Such experience can come from previous firsthand work experience on the ground or from the preparation program he or she attended (Browne-Ferrigno and Muth, 2009; Hitt et al., 2018; Leithwood et al., 1996). However, based on empirical evidence, Adejumo (2017) asserted that principal replacement is not always necessary; if the principals of underperforming schools have open minds, so they will not be imprisoned by their previous knowledge and successful experience, they can evolve as capable turnaround school leaders, adapt in accordance to the current reality, and lift their schools out of their underperforming status.

Antecedents and factors affecting turnaround

Many factors can either provide aid for or hinder the result of school turnaround. For instance, as mentioned above, the original culture and structure of a school as well as teachers’ complacency about the status quo can create hurdles for turnaround attempts. Moreover, successfully turning around a struggling school takes time; while it is common for both the district and the turnaround leaders to see noticeable effects of the turnaround attempt, it is important for the stakeholders to allow enough time and patience, as well as necessary support during the process (Mania-Singer, 2018; Murphy and Bleiberg, 2019). It is also possible that some turnaround measures may also come across active resistance from teachers as well as the community, especially when the measures are perceived as loss (Murphy and Bleiberg, 2019; Pham et al., 2020). Therefore, it is essential for turnaround leaders to maintain active communication with these stakeholders and provide evidence of progress or success along the way. Other factors such as teachers’ experience and professional capacity, as well as teacher turnover, may also have significant impacts on the effectiveness of school turnaround (Pham et al., 2020). Moreover, as reviewed previously, many turnaround projects are part of the political campaigns or strategies of the local or even the state government (Mania-Singer, 2018; Meyers, 2020; Pham et al., 2020; Schueler, 2019). Therefore, another vital element in the success or failure of turnaround is the level of support a turnaround school receives from its superior governing body such as the district or relevant government division(s) (Meyers and Sadler, 2018).

Creating a collaborative school culture and structure

Just as Manla (2021) articulated, a school would not be likely to achieve any instructional or curriculum successes without an enabling culture and a climate that celebrates respect, clear expectations, recognition, personal responsibility, and so forth (p. 21). Turnaround leaders with a successful track record can be helpful for repairing and rebuilding school climate and culture, in that they are better at accommodating change—a crucial leadership quality turnaround schools badly need (Yoon and Barton, 2019).

Realizing and taking into consideration the importance of school culture and climate, a Southern US school district (Turner School District) explicitly included changing the school culture as an essential part of their strategy for the turnaround of the schools in the district (Meyers, 2020). Harbatkin (2020) also came to the conclusion from the literature that a change in school culture and climate is necessary for successful school turnaround.

If turnaround principals are to consistently apply the democratic and collaborative leadership styles that have been considered more effective in the school turnaround leadership literature, it is necessary for them to develop and maintain a school environment that allows for collaboration. In her doctoral research project on leadership reform in low-performing schools in the United States, Harbatkin (2020) identified that the amount of teachers’ time used for collaboration was positively related to school performance. In other words, when teachers spend less time on collaboration, there may be a decline in the school’s performance.

When turnaround leaders use supportive leadership strategies such as shared or distributed leadership, it is necessary to create an inclusive environment in the school where all teachers and staff, along with school administrators, can collaborate on a collective vision. In doing so, the school can mobilize all its manpower and resources to achieve better curriculum and instructional performance and change the downward trajectory of the school development in a positive direction (Huguet, 2017).

Principals’ role in school turnaround

However, not all well-performing principals can turn around a struggling school. Drawing on a substantial amount of literature, Meyers and Hitt (2017) pointed out a common belief that turnaround principals are fundamentally different than regular effective principals. For instance, they have higher levels of emotional intelligence and are usually better at making use of both support and accountability as well as catalyzing change in the chronically struggling school, while effective principals in already well-performing schools mainly need to maintain their good condition and practices (Meyers and Hitt, 2017).

Yoon and Barton (2019) also pointed out that turnaround leaders with a successful track record can be helpful in turning around a struggling school because they are more accommodating of change—a crucial leadership quality turnaround schools badly need. Reitzug and Hewitt (2017) asserted that turnaround principals need to be able to provide or instill a sense of vision or direction in the faculty, establish and strengthen connections with the staff, and at the same time help develop the capacity and efficacy of the teachers. These authors also argued that successful turnaround principals must be able to create “ongoing dynamic tensions” [italics in original] between elements such as direction, connection, and development in their schools and at the same maintain balance between them so as to disrupt complacency and the status quo and keep their faculty members “at the edge of their competence and capabilities” (Reitzug and Hewitt, 2017, p. 517).

Leadership practices in school turnaround

As mentioned above, each turnaround school has its unique conditions and challenges; therefore, the strategies and measures developed to turn them around should be tailor-made with consideration for the particular situation of each school rather than mechanically relying on the prior experiences of successful cases elsewhere (Harris et al., 2018). Wana (2019) argued that effective turnaround leaders ought to *maintain that focus on student learning, provide clear and consistent expectations, develop a clear sense of accountability, and adjust their instructional and leadership strategies based upon actual data from the turnaround process*. In addition, other effective leadership practices can include building and developing teachers’ professional capacity, create a collaborative and supportive school culture and climate, and connecting with external partners (Same, 2017).

Getting to know the terrain

Upon entering a failing school, it is unwise for turnaround principals to rush into looking for solutions without a thorough understanding of the root cause of the school’s struggles (Mania-Singer, 2018). Instead, it is essential for the turnaround principals to try, upon arrival at a low-performing school, to familiarize themselves with its history, its existing organizational culture and climate, and the strengths and weaknesses of the teachers and the students, so that they can develop a meaningful turnaround strategy that truly fits the school’s needs.

Setting up a new vision and planning

High-quality school improvement planning aligned with a resonating vision is an essential step toward improving the climate and student learning of underperforming schools (Meyers and Hitt, 2018). Without it, teachers and other supporting staff will not be able to find a clear purpose but will instead “muddle through activities” without knowing why (Huguet, 2017).

Redding and Searby (2020) extracted a fourfold school improvement plan from a rich volume of relevant literature: “(1) establish a shared vision, (2) set strategic goals that help achieve this vision, (3) identify areas of staff development, and (4) create measurable outcomes that can be routinely monitored by school, district and state personnel” (p. 69). However, to do so, turnaround leaders and policy makers ought to learn about and take into consideration the actual organizational context of the schools while planning (Meyers and Hitt, 2018).

It is a common assumption as well as practice for turnaround leaders to make short-cycled plans for the schools they are entrusted to save, but Meyers and Hitt (2018) discovered that, in reality, most of such attempts were unsuccessful, possible reasons for which include the following factors:

1. The plans and the goals or objectives are not clearly communicated within the school;
2. The components of these short-cycled plans are not aligned with one another or the priorities, goals, and vision of the turnaround mission;
3. These plans are not meaningful for or connected to the new vision and objectives.

Building trust and demonstrating respect

One important factor in a school’s culture and climate is mutual trust between the school administrators and teachers, as well as among the teachers. Principals are more likely to trust teachers who help in meeting school goals, while teachers are more likely to trust principals who display personal consideration and benevolence (Bush, 2020).

On the other hand, if teachers feel that they can trust the principals to provide them with support and encouragement, they inclined to be more enthusiastic and committed to their job and the organization (Huguet, 2017). Such a trusting and supportive school culture and climate will also have profound positive impacts on the behavior as well as the academic performance of the students (Wana, 2019).

Improving communication

Effective communication can help turnaround leaders better understand the current conditions and problems of the school. After learning the terrain, turnaround principals need to create or find a new vision with clearly defined goals and objectives, so that they can initiate and implement activities that truly pull the school in the intended directions of development (Huber, 2018). Successful turnaround principals use communication of ideals and beliefs to promote their turnaround campaign and also to influence and facilitate change in the mindsets of the teachers as well as parents. These changed mindsets then cause change in the teachers’ and parents’ behaviors and ultimately enable the process of turning the school around (Adejumo, 2017).

Turnaround leadership competence

Compared with past research, a new advancement in turnaround leadership research has focused on the effects of turnaround leadership on student achievement. Hitt et al. (2019) explored the internal states of principals who orchestrate school turnaround with improved student achievement. This research identified seven important competencies of a successful principal, including: Initiates and Persists, Inspires and Motivates Others, Elicits Intended Responses, Builds Capacity with Accountability and Support, Commits to Student Learning, Crystalizes Problems and Creates Solutions, and Uses Inquiry to Systematically Frame Problems and Solutions. The research indicated that only Inspires and Motivates Others and Builds Capacity with Accountability and Support had higher-level scores, while Commits to Student Learning had a low-level score.

District leadership in turnaround contexts

Meyers and Sadler (2018) specified 12 district leadership practices. For instance, district leadership promotes bold changes to prioritize turnaround work while involving multiple members of the leadership team to initiate change. District leadership not only demonstrates high expectations to principals and school leadership teams, but also provides suitable resource utilization and responds to key needs in a timely way. In addition, district leaders need to give sufficient authority to turnaround principals and prioritize the hiring of school leaders in high priority schools. Meanwhile, district leaders are required to ensure good recruitment, placement, and retention while holding teachers accountable for teaching performance. In this process, rigorous assessment is clearly aligned with fundamental standards. The authors also explained that district leaders need to have a curriculum strategy to guide teachers' work. Subsequently, a data-driven culture and student management data system are required for instruction monitoring.

Importantly, Meyers (2020) also specifies the district leadership practices that are necessary for transforming turnaround schools. This research found that Commitment to Equity is the primary responsibility for district leaders who wish to initiate change and improvement in turnaround schools. District leadership also needs to create necessary structures while prioritizing talent and sets of skills to lead the work in schools. In addition, Meyers (2020) suggested that district leaders need to plan intentionally and with urgency, aiming to motivate and engage teachers into the change process. Importantly, district leaders need to provide the supports that schools need. Furthermore, district leadership needs to cope with three tensions strategically: the tension between focusing on the mission or personal connections, the tension between high expectations and high trust and respect, and the tension between disturbing satisfaction and continuity.

Types of leadership for successful school turnaround

Not all schools are created equal. The leadership strategies that were successful in turning around one failing school are not guaranteed to be as effective in a different context; therefore, turnaround leaders have to be familiar with multiple leadership styles so that they can adapt to the particular context of the school they are trying to save (Gerade, 2019; Paletta et al., 2017; Yoon and Barton, 2019).

Leadership scholars have been attempting for a long time to make more systemic sense of different types of leadership styles in order to better inform both scholars and practitioners in the field. In as early as the 1950s, Tannenbaum and Schmidt (1958) brought forward their classic Continuum of Leadership Behavior that puts different leadership patterns on a spectrum with *boss-centered leadership* using managers' positional authority on one end and *subordinate-centered leadership* allowing more freedom for subordinates on the other. A more modern but similar conceptualization can be seen in the comparison between authoritative and democratic leadership styles. Another popular school of leadership thoughts pioneered by Bass and Avolio (1990) among other scholars divides leadership into a duality of transformational and transactional leadership. Building on top of the previously available leadership knowledge, upon the beginning of the 21st century, leadership scholars developed a new perspective in the robust leadership literature—distributed leadership (Grint, 2011). More recently, Makgato and Mudzanani (2019) divide leadership into task-oriented and relationship-oriented styles. According to the scholars, task-oriented leaders emphasize more on the completion of the tasks but such an approach cannot effectively promote the team members' creativity and innovation. In contrast, the authors posit that relationship-oriented leaders support, motivate, and develop their subordinates, and that they foster positive relationships in the workplace and try to create open channels of communication.

By and large, supported by the recent literature, many turnaround leadership scholars believe that more subordinate-centered, democratic, and transformational leadership styles are more effective in elevating underperforming schools. For instance, Makgato and Mudzanani (2019), who has been mentioned above, conclude that relationship-oriented leadership approach can have a positive influence on students' learning performance. Such arguments resonate with the leadership practices reviewed above—effective turnaround leaders try to create a workplace culture filled with trust and respect and empowered with effective communication. In the same article, the scholars also assert that relationship-oriented leadership approach is not any particular single leadership style; instead, it integrates democratic leadership, transformational leadership, and shared instructional leadership.

With a similar tone, referring to recent literature, [Lochmiller Chad and Chesnut Colleen \(2017\)](#) also point out that effective leadership in schools does not rely on any single style but is rather an integration of transformational, managerial and instructional leadership behaviors (p. 86). Drawing upon data collected from 15 turnaround school leaders in LA County, [Same \(2017\)](#) also found that leadership practices under transformational and situational leadership styles are more effective than other approaches. Such research findings are also in agreement with what has been mentioned earlier that effective turnaround leaders should get to know the situation of the school they aspire to help instead of imposing measures from their past experiences. After all, each under-performing school has their own unique challenges.

Informed by multiple years of leadership literature, [Meyers and Hitt \(2017\)](#) noticed that distributed leadership has been given a significant amount of attention and emphasis. [Huguet \(2017\)](#) also echoed that collaborative leadership styles such as shared leadership, distributed leadership, and teacher leadership are most effective in positively influencing school performance. [Miles \(2018\)](#) identified that distributed leadership and teacher leadership had positive effects on the turnaround of a case school by revitalizing the school culture and improving the mutual relationship between teachers and students.

Similarly, drawing on the literature on leadership styles such as instructional leadership, transactional, and transformational leadership, as well as contingency leadership, [Petrich \(2019\)](#) studied how principals who successfully turned around failing elementary schools adjusted their leadership behaviors in a mid-Atlantic school district in the United States and identified five effective leadership practices: “inspire a shared vision, model the way, challenge the process, encourage the heart and enable others to act” (p. 77). In his doctoral research project, [Wana \(2019\)](#) assessed successful turnaround principals in Minnesota in the United States, and identified empirical evidence of similar leadership behaviors.

However, [Harris et al. \(2018\)](#) point out that turnaround leadership literature is largely Western-centric, and therefore has paid much more attention to more democratic leadership styles or approaches, such as distributed leadership, transformational leadership, teacher leadership, and so forth. One major reason for such a preference is that culturally western societies have much lower hierarchical power distance than other countries or regions especially in Asia. For instance, according to the big-data calculation by [Hofstede Insights \(2021\)](#), the Power Distance Index in many Asian countries is twice or three times as high as Western societies in North America and Europe. Through a case study in 10 successful turnaround schools in Malaysia and Indonesia, the authors also identified that many effective turnaround principals in Asia adopt top-down leadership approaches, considering themselves to be moral leaders and role models for teachers to follow rather than placing teachers at the center of the stage. [Harris et al. \(2018\)](#) contributed greatly to this line of research by exploring the emergent context of turnaround leadership. The development of the understanding of turnaround leadership in the literature has been illustrated through further exploration of themes related to such leadership. However, future exploration of the effects of turnaround leadership and the paths to improve student learning and achievement is necessary.

Conclusions

In total, 44 publications were reviewed, including 28 peer-reviewed journal articles, sections or chapters from seven books, and eight doctoral theses. These publications provides a relatively clear picture of the research progress and future research directions in turnaround leadership. Some important findings includes the leadership roles, quality of principals, replacement of principals, restraints on school principals, and districted leadership. This review emphasizes leadership’s role in integrating various turnaround processes, but the specific ways in which this integration takes place needs to be further explored. For example, [Bush \(2020\)](#) confirmed that it is efficient to retain the school principal’s role as the ultimate leader in school management, even when using a democratic leadership style. In addition, [Harbatkin \(2020\)](#) and [Meyers and Hitt \(2017\)](#), confirmed the importance of principals with good qualities and successful leadership and management experience in the process of transforming turnaround schools. Replacement of the school principal has been a widely used method to conduct school turnaround in the field. It is worth further exploring the effects of replacement of the school principal on turnaround school transformation. Principal replacement might benefit the receiving school. Also, the sustaining turnaround is problematic when the original school principal was commissioned to another turnaround school. It will be meaningful to further explore how turnaround schools can further develop their school principals to be effective turnaround principals.

In terms of restraints on school principals in turnaround contexts, this review identified that the political environment, resistance from the school and its community, and the principal’s own lack of ability can restrain a principal’s work. An understanding of the political environment is important for school principals to transform turnaround schools effectively ([Schueler, 2019](#)). District leadership in turnaround contexts was another research trend that has been picking up its momentum in this field. [Meyers and Sadler \(2018\)](#) specified 12 district leadership practices, among which providing support is crucial. This emerging field needs to be further explored to understand effective transformation in different contexts. Overall, the important aspects of turnaround leadership research have made new progress, but substantiated evidences are required for further understanding this field.

References

- Adejumo, M., 2017. Beyond Socioeconomic Status: The Impact of Principal Leadership in Urban and High Poverty Turnaround Schools (Doctor of Education). Seton Hall University, ProQuest. Retrieved from: <https://www.proquest.com.eproxy.lib.hku.hk/docview/2009556096?accountid=14548>.
- Arar, K., Örüçü, D., 2021. Rethinking educational leadership and policy in schools in challenging circumstances: inter/intra-continental perspectives. *Leader. Pol. Sch.* 20 (1), 1–5. <https://doi.org/10.1080/15700763.2020.1838555>.
- Bass, B.M., Avolio, B.J., 1990. Developing transformational leadership: 1992 and beyond. *J. Eur. Ind. Train.* 14 (5).
- Browne-Ferrigno, T., Muth, R., 2009. Candidates in educational leadership graduate programs. In: *Handbook of Research on the Education of School Leaders*, pp. 195–224.
- Bush, T., 2020. Turnaround Leadership in Education: A Restatement of Heroic Leadership or a Flawed Concept. SAGE Publications Sage UK, London, England.
- Coby, V.M., 2018. *International Perspectives on Leading Low-Performing Schools*. Information Age Publishing, Charlotte, NC.
- Collins, J.C., 2009. *How the Mighty Fall: And Why Some Companies Never Give in*. Random House.
- Gatewood, R., Gynelle, T., 2020. A Case Study of School Decline at a Historic, Predominantly African American, Urban High School (Doctorate). University of Huston. Retrieved from: <https://uh-ir.tdl.org/handle/10657/7294>.
- Gerade, A.O., 2019. A New Order of Things: How Does a Secondary School Principal Lead a Community Out of School Turnaround (Doctorate). University of Massachusetts Lowell.
- Grint, K., 2011. A history of leadership. In: *The SAGE Handbook of Leadership*, vol. 14, pp. 3–14, 1.
- Hallinger, P., 2013. A conceptual framework for systematic reviews of research in educational leadership and management. *J. Educ. Adm.*
- Harbatkin, E., 2020. Low-Performing Schools and School Reform: Three Essays on School Turnaround, the Mechanisms of Low Performance, and Leadership for Reform (Doctorate). Vanderbilt University.
- Harris, A., Jones, M., Ismail, N., Adams, D., Sumintono, B., 2018. Leading turnaround and improvement in low-performing schools in Malaysia and Indonesia. In: *International Perspectives on Leading Low-Performing Schools. Contemporary Perspectives on School Turnaround and Reform*. Information Age Publishing, pp. 267–287.
- Hitt, D.H., Meyers, C.V., 2017. Promising Leadership Practices for Rapid School Improvement That Lasts. WestEd, San Francisco, CA.
- Hitt, D.H., Meyers, C.V., 2018. Beyond turnaround: a synthesis of relevant frameworks for leaders of sustained improvement in previously low-performing schools. *Sch. Leader. Manag.* 38 (1), 4–31.
- Hitt, D.H., Woodruff, D., Meyers, C.V., Zhu, G., 2018. Principal competencies that make a difference: identifying a model for leaders of school turnaround. *J. Sch. Leader.* 28 (1), 56–81.
- Hitt, D.M., Meyers, C.M., Woodruff, D., Zhu, G., et al., 2019. Investigating the relationship between turnaround principal competencies and student achievement. *NASSP Bulletin*.
- Hochbein, C., Mahone, A., 2017. The failure fallacy. In: Meyers, C.V., Darwin, M.J. (Eds.), *Enduring Myths That Inhibit School Turnaround*. Information Age Publishing, Charlotte, NC, pp. 11–28.
- Hochbein, C., 2012. Downward spirals, boiled frogs, and catastrophes: examining the rate of school decline. *Leader. Pol. Sch.* 11 (1), 66–91. <https://doi.org/10.1080/15700763.2011.577927>. Retrieved from.
- Hofstede Insights, 2021. Country Comparison. Retrieved from: <https://www.hofstede-insights.com/country-comparison/>.
- Huber, S.G., 2018. No simple fixes for schools in challenging circumstances: contextualization for Germany. In: *International Perspectives on Leading Low-Performing Schools*, pp. 243–266.
- Huguet, B.C.S., 2017. Effective leadership can positively impact school performance. *Horizon* 25 (2), 96–102. <https://doi.org/10.1108/OTH-07-2016-0044>.
- Klein, E.D., Young, M.D., Böse, S., 2021. Successful leadership in schools serving disadvantaged communities in Germany and the USA. In: *International Perspectives on School Settings, Education Policy and Digital Strategies*, p. 47.
- Leithwood, K., Jantzi, D., Coffin, G., Wilson, P., 1996. Preparing school leaders: what works? *J. Sch. Leader.* 6 (3), 316–341.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28 (1), 27–42.
- Leithwood, K., Harris, A., Strauss, T., 2010. *Leading School Turnaround: How Successful Leaders Transform Low-Performing Schools*. John Wiley & Sons.
- Liu, P., 2020. Understanding turnaround leadership research: continuity and change (2009–2016). *Educ. Manag. Adm. Leader.* 48 (1), 6–24.
- Lochmiller Chad, R., Chesnut Colleen, E., 2017. Preparing turnaround leaders for high needs urban schools. *J. Educ. Adm.* 55 (1), 85–102. <https://doi.org/10.1108/JEA-11-2015-0099>.
- Makgato, M., Mudzanani, N.N., 2019. Exploring school principals' leadership styles and learners' educational performance: a perspective from high- and low-performing schools. *Afr. Educ. Rev.* 16 (2), 90–108. <https://doi.org/10.1080/18146627.2017.1411201>.
- Mania-Singer, J., 2018. Making waves in riverside: a superintendent/principal partnership for dramatic school turnaround. *J. Cases Educ. Leader.* 21 (4), 131–145. <https://doi.org/10.1177/1555458918769119>.
- Manla, V.H., 2021. School climate: its impact on teachers' commitment and school performance. *J. World Engl. Educ. Pract.* 3 (2), 21–35.
- Meyers, C., Darwin, M., 2018. *International Perspectives on Leading Low-Performing Schools*. IAP.
- Meyers, C.V., Hitt, D.H., 2017. School turnaround principals: what does initial research literature suggest they are doing to be successful? *J. Educ. Stud. Placed Risk* 22 (1), 38–56.
- Meyers, C.V., Hitt, D.H., 2018. Planning for school turnaround in the United States: an analysis of the quality of principal-developed quick wins. *Sch. Effect. Sch. Improv.* 29 (3), 362–382. <https://doi.org/10.1080/09243453.2018.1428202>.
- Meyers, C., Sadler, J., 2018. District leaders engaged in school turnaround: identified challenges and espoused responses. *NASSP Bull.* 102 (2), 89–110.
- Meyers, C.V., Smylie, M.A., 2017. Five myths of school turnaround policy and practice. *Leader. Pol. Sch.* 16 (3), 502–523.
- Meyers, C.V., 2020. District-led school turnaround: a case study of one US district's turnaround launch for multiple schools. *Leader. Pol. Sch.* 1–20.
- Miles, K., 2018. *Shared Leadership Perceptions in a Turnaround Elementary School* (Doctorate). University of Portland. Retrieved from: <https://pilotsscholars.up.edu/etd/37>.
- Murphy, J., 2008. The place of leadership in turnaround schools: insights from organizational recovery in the public and private sectors. *J. Educ. Adm.*
- Murphy, J.F., Bleiberg, J.F., 2019. Barriers to turnaround. In: *School Turnaround Policies and Practices in the US: Learning From Failed School Reform*. Springer International Publishing, Cham, pp. 53–72.
- Paletta, A., Alivernini, F., Manganelli, S., 2017. Leadership for learning: the relationships between school context, principal leadership and mediating variables. *Int. J. Educ. Manag.* 31 (2).
- Pearson, T., Wolgemuth, J.R., Colomer, S.E., 2015. Spiral of decline or “beacon of hope”: stories of school choice in a dual language school. *Educ. Pol. Anal. Arch.* 23 (25). Retrieved from: <https://eric.ed.gov/?id=EJ1062317>.
- Petrich, E.R., 2019. *Leadership Practices of Successful Elementary Turnaround Principals*. Virginia Tech.
- Pham, L.D., Henry, G.T., Kho, A., Zimmer, R., 2020. Sustainability and maturation of school turnaround: a multiyear evaluation of Tennessee's achievement school district and local innovation zones. *AERA Open* 6 (2). <https://doi.org/10.1177/2332858420922841>.
- Redding, C., Searby, L., 2020. The map is not the territory: considering the role of school improvement plans in turnaround schools. *J. Cases Educ. Leader.* 23 (3), 63–75.
- Reitzug, U.C., Hewitt, K.K., 2017. “The click-clack of her heels and the jingle of her keys”: exploring the tensions in the leadership of a successful turnaround principal. *J. Sch. Leader.* 27 (4), 491–520.
- Same, E.H.R., 2017. *Best Leadership Practices of Turnaround K-12 Administrators* (Doctorate). Pepperdine University. Retrieved from: <https://digitalcommons.pepperdine.edu/etd/885>.
- Schueler, B.E., 2019. A third way: the politics of school district takeover and turnaround in Lawrence, Massachusetts. *Educ. Adm. Q.* 55 (1), 116–153.
- Senge, P.M., 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday.

- Tannenbaum, R., Schmidt, W.H., 1958. *How to Choose a Leadership Pattern*. Boston.
- Tao, Y., Liu, S., 2020. Network governance in education: the experiences and struggles of local governments in Chinese school turnaround. *Asia Pac. J. Educ.* 1–15.
- Wana, K.C., 2019. *Assessment of Leadership Practices of Principals of Turnaround Elementary Schools in Minnesota* (Doctorate). St. Cloud State University. Retrieved from. https://repository.stcloudstate.edu/edad_etds/56. Available from EBSCOhost OpenDissertations database.
- Whyte, P.A., 2018. *Turnaround as an Experience: Using School Culture and Climate as the Driver for School Turnaround* (Doctorate). Columbia University, Teachers College.
- Woulfin, S.L., Weiner, J., 2019. Triggering change: an investigation of the logics of turnaround leadership. *Educ. Urban Soc.* 51 (2), 222–246. <https://doi.org/10.1177/0013124517714865>.
- Yoon, I.H., Barton, A., 2019. Turnaround leaders' shifting gears in chronos and kairos time. *J. Educ. Adm.* 57 (6).

Evidence-informed decision making in schools and districts: the role of equity[☆]

Amanda Datnow^a and Vicki Park^b, ^a University of California San Diego, La Jolla, CA, United States; and ^b Department of Educational Leadership, San Diego State University, San Diego, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	169
Equity issues that arise with data use	170
Accountability-driven data use and data use for continuous improvement	170
Using data to confirm assumptions and using data to challenge beliefs	173
Using data for tracking and flexible grouping	173
Conclusion	174
References	175

Introduction

Ensuring equitable opportunities and outcomes for all students is an enduring vision for educators and policymakers across the globe. Although equity has been an abiding concern in various educational reform efforts, accountability policies have intensified the use of data to illuminate and address differences in achievement across racial, linguistic, and income groups. Data on student achievement are now ubiquitous in schools and school systems. Large-scale assessment data are widely available in a large number of countries, and many schools and systems are supporting teachers to engage in formative assessment as well. There is a heightened expectation that data will inform decision-making in all aspects of schooling.

A growing body of research has been published examining the supports and constraints for data use, the role of teachers and leaders in the data use process, professional development for data use, and the successes and challenges of engaging in data-driven decision making (e.g., Coburn and Turner, 2011; Lai and Schildkamp, 2016; Mandinach and Honey, 2008; Marsh, 2012; Means et al., 2009; Schildkamp and Poortman, 2015). The research on data use tends to be dominated by conceptual frameworks rooted in organizational learning (Christman et al., 2009). In an organizational learning framework, the focus is on identifying problems, creating action plans to solve them, and developing a collaborative culture of inquiry (Christman et al., 2009). Data use is conceptualized as part of a cycle of instructional improvement (Goertz et al., 2010) in which individuals engage in a process of defining a problem or setting goals, gathering data, analyzing data, followed by action planning and evaluating outcomes (Coburn and Turner, 2011; Schildkamp and Poortman, 2015).

The cycle of instructional improvement has its roots in the plan-do-study-act cycle of Total Quality Management, an approach that was prominent in the corporate sector in the 1980s (Goertz et al., 2010). These frameworks are not only common in data use research, they are also widely prevalent in school, district, and state data use practices (Goertz et al., 2010). Meanwhile, researchers have explained that the process of data use is more complex, contextual, and less rational than this linear model might suggest (e.g., Coburn and Turner, 2011; Schildkamp and Poortman, 2015; Vanlommel et al., 2017).

Power issues and other social justice concerns are features of the data use process, although many studies of data use have not focused on these elements. Power comes into play as individuals in higher level positions use data use processes as a form of managerial control, influence which data are prioritized, or influence interpretations of data (Coburn and Turner, 2011). Some problematic data use practices further marginalize low performing students, such as publicly displaying test scores (Neuman, 2016) or engaging in “educational triage” by focusing remediation efforts on students on the cusp of proficiency (Booher-Jennings, 2005; Gillborn and Youdell, 1999). Equity concerns related more broadly to high stakes accountability systems, such as a narrowing of the curriculum in schools serving low-income students of color, have also been raised and associated with data use (Diamond and Cooper, 2007; Lipman, 2004).

Large-scale accountability policies, while drawing attention to systemic inequities, are often narrowly focused on highlighting student achievement gaps at the expense of understanding and mitigating the effects of unequal educational conditions and processes. Reform efforts have consequently focused on ways to close these gaps. We concur with Ladson-Billings (2006) who argues that the term “achievement gap” unfairly constructs low-income and minority students as defective and lacking and places the burden on students to catch up. Instead, using the term “education debt,” moves us to a discourse that holds educators and policymakers accountable for providing students access to the educational resources that will allow them to achieve at high levels. Ladson-Billings further advocates for deploying our best research knowledge, skills, and expertise to alleviate the poor educational conditions and learning experiences that students receive in many schools. Ladson-Billings’ arguments dovetail with those of Valli et al. (1997) who explain that, “Equity has replace the older concept of equality of opportunity ... equity places more emphasis on

[☆] This chapter is an edited version of the following article: Datnow, A., and Park, V. Opening or closing doors for students? Equity and data use in schools. *Journal of Educational Change*, 19 (2), 2018, 131–152. It has been reprinted with permission.

notions of fairness and justice, even if that requires an unequal distribution of goods and services” (p. 254) for equitable outcomes to be achieved (cited in Gillborn and Youdell, 1999, p. 3).

Data use could be an important lever for achieving equity and ameliorating the educational debt, but how and under what circumstances this may occur has not been well documented. It is important to reflect upon the question, *how do data use practices open or close doors for students?* The aforementioned studies provide insight into some of the equity issues that may arise with data use, but a more thorough investigation is in order. In this article, we draw on prior research and data from our qualitative own studies to discuss the tensions that educators face in using data and the consequences for equity. We argue that an equity agenda needs to be at the forefront of the field’s understanding and study of data use in schools.

Equity issues that arise with data use

Data-informed decision making can help schools monitor their progress toward equity goals (Skrla et al., 2004). But we cannot simply promote data use and expect good things to happen. Educators play a critical role in shaping how and why data are used, what counts as data, and what people are aiming for when they push the use of data in schools (Coburn and Turner, 2011). Data do not drive decisions by themselves (Dowd, 2005). Data-informed decision making is thus a more appropriate term for educational practice (Knapp et al., 2007). Although the term data-driven decision making is commonly used in the field, from here onward we will refer to the practice as data-informed decision making to signal this important shift in thinking about data use.

Our interest here is uncovering the ways in which data use may serve to open or close doors for students. There are a variety of ways in which data can be used in the service of expanding students’ opportunities to learn. Data can be also used, sometimes inadvertently, to limit students’ opportunities. We bring both extant literature and our own research findings to bear on a set of data use practices in which some of these core tensions are present. These include tensions between:

- Accountability-driven data use and data use for continuous improvement
 - Data use meetings focused on instructional improvement and administrative compliance
 - Focusing on standardized test data and using multiple types of data
 - Focusing on narrow bands of students and carefully examining data on all students
- Using data to confirm assumptions and using data to challenge beliefs
- Using data for tracking and flexible grouping to promote student growth

In exploring these tensions, we are not suggesting that these are dichotomous practices, with each existing on one end of a spectrum. In reality, there are a range of practices undertaken within each realm, some of them existing simultaneously within the same school. For example, it is quite common to find districts that focus primarily on large scale assessment data, whereas teachers are much more inclined to look at students’ achievement holistically (Coburn and Talbert, 2006). Within the same school, we can find educators who use data to confirm assumptions about students, as well as those who try to use data to challenge beliefs, and educators who do both, depending on the context. So too, even in schools that use data to track students, we find educators using data in the service of flexible grouping. Our aim is to explore some of the equity issues that arise along these dimensions, while elucidating the complexity and variation in data use processes.

Accountability-driven data use and data use for continuous improvement

Accountability demands have figured strongly into data use in many countries including the Australia, Belgium, the Netherlands, the United Kingdom, and the United States, among many others (Pierce and Chick, 2011; Mandinach and Honey, 2008; Schildkamp et al., 2014). What is clear in the literature is the delineation between high-stakes, accountability-driven data use, which emphasizes complying with external pressures and bureaucratic demands, and data use for continuous school improvement and organizational learning. Across the globe, educators often have to balance these two competing agendas simultaneously (Lai and Schildkamp, 2016). As educators balance various demands, key decisions are made with respect to data use and equity.

Firestone and Gonzalez (2007) explain that an accountability-driven data use culture focuses primarily on raising student test scores and tends to have a short-term time frame. Data are used mainly to identify problems and monitor compliance. When the stakes are high, “data can slant the system toward trickery and treachery” (Hargreaves and Shirley, 2012, p. 39). These misuses of data and the overemphasis on test score data have led policymakers in some places, such as Wales and Alberta, to abolish assessments at various grade levels (Hargreaves and Shirley, 2012). Other countries, such as Finland and Germany, test only a subsample of students rather than engaging in high-stakes testing of all students and, as Shirley (2017) explains, “This has sheltered German educators from the curriculum narrowing that occurred in US schools” as a result of No Child Left Behind (NCLB) (p. 61).

In contrast to accountability-driven data use, Firestone and Gonzalez (2007) note that data use for continuous improvement focuses on student and organizational learning and instructional improvement and is long-term in scope. In this model, data are used for diagnostic purposes and with integrity (Shirley, 2017). Data are used as part of a reflective process aimed at enlightenment and guiding and mobilizing action to improve the school over time (Firestone and Gonzalez, 2007). Also, while

continuous improvement can be a useful frame for data use, equity issues may still go unexamined, unless problems are framed as such, as we will explain. However, a continuous improvement focus is more likely to cohere with equity than an accountability focus.

It is also important to note that even when data use is framed in terms of continuous improvement, educators still may experience data use as a form of accountability. In Goertz et al.'s (2010) study, while the district's implementation of benchmark assessment was ostensibly aimed at continuous improvement in mathematics, teachers felt pressured by principals. "If you don't do well, you get talked to," explained a teacher (Goertz et al., 2010, p. 76). School context also matters. As Diamond and Cooper's (2007) study of Chicago showed, higher performing schools could meet the vision of using data for instructional improvement, whereas low-performing schools who were under sanction looked for quick-fixes to get out of probationary status. Many of these probationary schools served large numbers of low-income students of color and were characterized by a narrow focus on particular forms of data, as well as targeted students, grade levels, subject areas (Diamond and Cooper, 2007).

Data use meetings focused on instruction and administrative compliance. Leaders help set the tone for data use among teachers, focusing them away from or toward accountability, continuous improvement, and equity concerns. Correspondingly, the tension between accountability and continuous improvement is reflected in whether teacher collaboration meetings are seen as a place for genuine inquiry around student learning, or whether they are administratively regulated around compliance demands related to data use. These different foci have other important implications for equity, particularly when we consider the types of schools in which these different patterns of data use take shape. The focus on data use solely for accountability purposes (or not) shapes the decisions about what data are used, as well as how discussions among teachers unfold. Because professional learning communities are often the primary vehicle by which teachers learn to use data, the character of them matters a great deal (Daly, 2018; Jimerison and Wayman, 2015; Marsh and Farrell, 2014). As Huguet et al. (2017) argue, if teachers have unequal opportunities to learn, so will their students.

Horn et al. (2015) study found that teacher workgroups used different logics when interpreting the same district math assessment data. One teacher work group was focused on instructional management, spending their time categorizing students according to achievement levels. This teacher work group's instructional management logic aligned with the principal's focus on accountability. They adopted his frame of data use as a monitoring activity, rather than as a vehicle for examining students' mathematical understanding. In contrast, another teacher work group's discussions focused on instructional improvement, including deeply investigating the source of student mistakes. This pattern has been documented in other studies as well. Gannon-Slater et al. (2017) found that an accountability-focused data culture at the school they studied led teachers to focus on the technical dimensions of teaching and "what works," displacing equity aims. Huguet et al. (2017) explain, "In some cases, the work in a PLC [Professional Learning Community] can lead to evidence-based conversations that lead to a re-thinking of instruction, while in other cases, teachers can experience PLCs where time is invested in more procedural matters rather than reflection and learning tied to data" (p. 378).

In our own research, we have also observed different orientations in the teacher team meetings we have observed. In our earlier studies, numerous schools utilized data analysis protocols and tools in teacher team meetings. Many teachers described productive work related to these data discussion protocols (Datnow et al., 2012). However, in other cases, the administrative regulation that accompanied the protocols led some groups of teachers to focus on the tasks (e.g., completing a form describing the outcome of their discussions of data), rather than on meaningful discussions around data and instructional improvement. For example, in one data team meeting we observed, the overall point of the meeting appeared to be making sure that all parts of the protocol were addressed. With the focus on form completion, teachers spent less time engaging in a deep discussion of instruction.

We have also found that teacher team meetings can have competing purposes, defined by teachers in terms of instructional improvement but in terms of accountability by the principal. Park et al. (2013) reported on schools and districts in which leaders strategically framed the use of data in an effort to produce more equitable learning opportunities and outcomes for all their students. These leaders constructed sensemaking frames that centered on the need to confront disparities in opportunities-to-learn, school improvement as a shared collective responsibility, and making incremental structural and cultural changes to ensure sustainability of reform goals. The school principal and a coach in a school we studied redirected dialog toward students' strengths rather than weaknesses and oriented the conversation around improving practice (Park, 2018). In both of these cases, we see a deliberate commitment on the part of leaders to frame data use for instructional improvement as an *equity-driven activity*.

Focusing on standardized test data and using multiple types of data. Educators' and policymakers' decisions about what counts as data play an important role as well. Standardized tests themselves raise questions from an equity standpoint as they have long been criticized for their orientation toward forms of knowledge that privilege white, middle-class students (Garner et al., 2017; McNeil, 2002; Santelices and Wilson, 2010). As such, they put students from marginalized communities at an unfair disadvantage. High stakes testing also "creates intense, unfair pressure for teachers and administrators to raise scores in a system that does not support success for students of color, emergent bilinguals, and students from low-income families" (Garner et al., 2017, p. 422). These are concerns with standardized tests not with data use per se, but it is important to note that these issues arise when data use is narrowly defined to large-scale accountability measures.

With these important pitfalls of data use in mind, it is essential to address the relationship between federal and state accountability policies and data-informed decision making. First, the emphasis on using evidence (broadly speaking) to inform educational decision making is not new and can be found in many earlier reforms in and outside education. Thus, we should not confuse data use and accountability, nor should we assume that the two must be linked. The problem is that the type of data use that is connected to accountability policies is one that is narrowly construed, both in terms of what data is used and how it is used. Accountability

policies have prompted many districts and schools to be more focused on data use (Ingram et al., 2004; Marsh et al., 2006). These accountability pressures also motivated many districts to adopt benchmark assessments aligned to state tests in order to examine on an interim basis how students were achieving relative to state standards (Braaten et al., 2017; Jimerson and Childs, 2017; Snodgrass Rangel et al., 2017). In some cases, examining data on benchmark assessments has motivated teachers to look at a range of other information that can inform them about student learning and guide instructional changes (Datnow and Park, 2014). In other instances, data use is organized *only* around these benchmark assessments and has led to problematic consequences, such as a district's strong accountability orientation inadvertently compromising teachers' attempts at meaningful instructional improvement (Braaten et al., 2017).

Educators focused on continuous improvement actively seek out a wide range of data and do not limit themselves to data linked to accountability mechanisms. In our own research, we found that the sources of data that teachers actually relied upon went well beyond external assessment measures and included teacher created assessments, curriculum embedded assessments, writing portfolios, results from student work with online instructional tools, and of course their own observations of student learning. Drawing on this wide range of data allows for a much fuller portrait of student learning, allowing for a greater possibility that the strengths of students will be evident (Datnow and Park, 2019). Using multiple forms of data to inform instructional decision making can help provide a more holistic picture of student achievement and provide more ways to see evidence of student growth. This allows for a greater possibility that the strengths of students who have historically been disadvantaged by standardized measures will be evident. Drawing on a wider range of data may also lead to more informed, differentiated instruction (Park and Datnow, 2017).

Conversely, focusing on limited data, especially data from high-stakes assessments that are themselves biased against low-income students of color, can work against equity goals, contributing to narrow instructional goals and higher dropout rates (McNeil, 2005; Valenzuela, 2005). As Valenzuela (2005) argues, "When the test is the sole or primary arbiter in decisions with such long-lasting consequences for children, we insist that students have the right to be assessed in a complete and fair manner, using as many criteria as may reasonably indicate children's cognitive abilities and potential" (p. 2). At the same time, while using multiple forms of data is necessary, it is not sufficient, as we will explain.

Focusing on narrow bands of students and carefully examining data on all students. Under increased public scrutiny, NCLB and related policies in the US provided rewards and sanctions for schools based on the number of students who are performing at or above grade level. As one unintended consequence of NCLB, educators in Texas engaged in "educational rationing and triage" by examining state assessment data to determine which students are on the "bubble" or cusp of grade level proficiency and then investing resources to push these students over the threshold of the proficiency marker (Booher-Jennings, 2005). The same practice was used by low performing schools in Chicago (Diamond and Cooper, 2007). These practices were not limited to the US, however, as Hargreaves and Shirley (2012) documented a school in Ontario that focused test preparation on the 20% of students just below the passing mark. We found a similar pattern in our own research (Datnow and Park, 2014, 2019). Students who are very low performing were considered "lost causes," and those performing above grade level were not focused on either. This "educational triage" approach leaves many students behind (Booher-Jennings, 2005). By focusing on the use of data to avoid sanctions, educators, intentionally or unintentionally, subvert the intended goal of these policies.

This focus on students on the bubble shows up not just in individual teacher planning, but in teacher conversations as well. Halverson et al. (2007) found that faculty data discussions overwhelmingly centered on helping students below proficiency levels with few discussions focused on raising students from proficient to advanced levels. While such practices may result in greater numbers of students achieving "above the mark," they inadvertently compromise the overarching purpose of data driven practice, which is to expand opportunities for all students.

Achieving goals of equity requires carefully examining data on each and every student. One school we studied has long used a process by which teams of teachers sit down with the principal, a counselor, and 2–3 intervention teachers three times a year (fall, winter, and spring) to discuss data and plan instructional interventions for every student (Datnow and Park, 2019). At the spring meeting, the group also discussed classroom placements for the following year. A notable feature of these meetings, however, was that while meeting participants all had numeric data on student achievement in front of them, their discussion was not restricted to numbers. Most discussion focused on providing a full picture of the student beyond academic achievement and involved discussions of students' home lives, social adjustment, classroom behaviors, work habits, completion of homework and so forth, none of which were reflected on the data sheets. Wardrip and Herman (2017) also found that "Teachers framed students' instructional needs with information that fell outside of the data the school collected" (p. 14). In their study, knowing the students' "backstories" (e.g., their friends, aspirations, interests) provided a context by which educators reasoned about student data.

Examining data on all students also promotes shared responsibility, which is a key component of data use for equity. District and school leaders play an important role in cultivating a belief that teachers must share responsibility for all of the students at the school, not just those they have contact with in their classroom. Equity cannot be achieved unless data use efforts focus on all students. When data use is narrowly focused on students "on the bubble," many students get left behind. A focus on all students allows data use efforts to address students' individual needs and for instructional changes to focus on lifting the class as whole, rather than small groups of students. It also helps to promote and is supported by a shared sense of responsibility for all students.

Using data to confirm assumptions and using data to challenge beliefs

A goal of data driven decision making is to bring evidence to light that will help educators think about student achievement in a new way. However, examining data does not always lead to new interpretations. In a study of data use in Belgium, [Vanlommel et al. \(2017\)](#) found that “teachers seek for data that confirms what they believe to be true and avoid data that questions their beliefs” (p. 81). Data can be used to validate existing understandings of students’ learning profiles ([Oláh et al., 2010](#)). As [Oláh et al. \(2010\)](#) explain, teachers interpret data in the context of their expectations for individual students or for the class as a whole and rely on their background knowledge about students when interpreting data. When data do not fit with teachers’ expectations of students this can send up a “red flag” as one teacher in this study noted, which may turn out to be useful in addressing students’ learning needs.

However, when educators use student characteristics as explanations for results, they can reinforce a culture of low expectations and stereotypes ([Bertrand and Marsh, 2015](#)). For example, a teacher in Bertrand and Marsh’s study discussed data on her students such that “the connection between ‘resource kids’ and the lack of proficiency was self-evident, suggesting an assumption that students in special education score poorly on tests by nature” (p. 19). Moreover, this teacher attributed the assessment results to her students, not to her teaching. The attribution of results to what were thought to be stable student characteristics was a pattern in this study, and English learners, special education students, and other struggling students were “blamed” for outcomes.

When educators use data to confirm assumptions, this often goes hand-in-hand with comments about students’ home lives being the primary explanation for student achievement. When such assumptions are made, it is very difficult for educators to see data as a useful tool for improving their own practices, as they locate achievement patterns in the students’ backgrounds.

Paying close attention to teacher talk in the context of data collaboration meetings can help researchers examine how teachers think about student ability, as teachers often discuss data in relation to individual student achievement. In our own research ([Datnow et al., 2018](#)) and in [Bertrand and Marsh’s \(2015\)](#) study, the No Child Left Behind era accountability language figured strongly into teachers’ talk. Terms used to describe students such as “low,” “high,” or “below basic” are reinforced by assessment data showing which students were below, at, or above grade level. These labels are imposed upon teachers by broader policies for accountability and have become part of teachers’ lexicon. This lexicon appears to reify a hierarchy of ability and can constrain the ways in which teachers talk about student achievement and learning, as well as their expectations of students.

At the same time, data use has been shown to be a powerful tool to push teachers to challenge existing assumptions about student learning and to reflect critically on instructional practices ([Lachat and Smith, 2005](#)). The process of closely examining data in the context of teacher team meetings can facilitate teachers’ focus on student growth, thereby shaping teachers’ beliefs about what they think their students are capable of. The team meeting can provide a space and a routine for teachers to point to various forms of “data” when making claims about student achievement. For example, in our own work ([Park, 2018](#)), these opportunities for dialog around data helped teachers move away from deficit framing of students and toward discussions of student growth and potential. Coaches or other instructional leaders played a critical role in reframing conversations around student assets.

In the end, data can play a very powerful role in challenging stereotypes and providing an opportunity for educators to examine the relationship between instructional practices and achievement. But simply deciding to build professional learning communities is not sufficient to bring about change. As [Pollock \(2017\)](#) argued, school talk must debunk myths about intelligence as easily measurable and explicitly challenge common comments about young people or families that are harmful. Framing conversations carefully and providing the opportunity for educators to bring multiple sources to bear on conversations about student achievement are critical.

Using data for tracking and flexible grouping

Educators’ conceptions of student ability inform their logics about how to organize instruction for students. In a study of ten detracking schools, [Oakes et al. \(1997\)](#) found that some teachers held a conventional view of intelligence in which ability was seen as an uni-dimensional, fixed, and innate characteristic. This led them to believe that intelligence could be easily assessed and to support tracking and ability grouping. Beliefs that overlapped race with ability were also prevalent among this group. In contrast, another group of teachers believed intelligence was multidimensional and plastic, leading them to support mixed ability grouped classes ([Oakes et al., 1997](#), p. 494). Data have long been an important part of tracking practices in schools. The use of data starts early in the educational process as students are often assessed for school readiness in kindergarten and then move through the system at different speeds ([Oakes, 2005](#)). By high school, disproportionate percentages of African American and Latino students are in courses that include less rigorous instruction, less qualified teachers, fewer resources, and lower expectations than their white or Asian counterparts in higher track classes ([Oakes et al., 1992](#)). Thus, tracking has significant consequences for students’ opportunity to learn.

Just as data have been used in the past to justify tracking decisions, in recent years, educators have turned to benchmark assessments as a placement tool, which is not their intended purpose. Instead, these assessments are designed to provide educators with interim feedback on student progress relative to curriculum standards. This has been documented in numerous studies ([Davidson and Frohbieter, 2011](#); [Heppen et al., 2012](#); [Shepard et al., 2011](#)). In addition to misusing the assessments for unintended purposes, the sole use of benchmark assessments places too much emphasis on one form of data to make such a high-stakes decision.

Uncritical and high-stakes uses of data can have long-term consequences for students, whose confidence and access to rigorous instruction is often tied to the labels that are applied to them.

Whereas the use of data for tracking purposes limits student opportunities, the use of data for flexible grouping of students can help to expand opportunities. As part of their comprehensive, three-times a year data analysis process, educators at an elementary school we described earlier created language arts and spelling groups that shifted often (Datnow and Park, 2019). In another elementary school, teachers used formative assessment data daily to differentiate instruction and to place students in flexible learning groups to address particular skill areas. Along these lines, Oláh et al.'s (2010) study found that elementary school teachers used small, pullout group instruction during a designated re-teaching week. Based on results on a benchmark assessment, teachers determined the students of most need. As a teacher explained, "If it's half the class ... I'll just reteach the whole thing. But if it's a few children ... then I would definitely pull them out and get some special homework for them to work on" (Oláh et al., 2010, p. 240). High scoring students received less direct instruction during the re-teaching week and instead were given enrichment worksheets, games, or more time on the computer. This approach to flexible grouping based solely around remediation also has the unintended consequence of stigmatizing the students who are repeatedly "called to the back to table" to work with the teacher.

Overall, while we have observed instances in which data use has inspired flexible grouping geared around students' individual needs and skills, there is certainly a strong cautionary tale to be told about the connection between data use and student grouping. Problematic practices of tracking and ability grouping with long-term consequences continue to abound in schools and are legitimated with data. In fact, tracking remains one of the most enduring practices in American high schools, in spite of a robust research base denouncing it.

Conclusion

This chapter drew upon extant studies to examine the ways in which the data use practices in schools may serve to open or close doors for students. We explored tensions between accountability-driven data use and data use for continuous improvement, data use to confirm assumptions and data use to challenge beliefs, and the use of data for tracking and flexible grouping. These tensions are examined and discussed across the scholarship on data use. By synthesizing these tensions as broad sets of decisions, we contribute to research on how data use practices can promote equity.

These sets of decisions are not necessarily dualisms with dichotomous sets of practices. Rather, along each of these dimensions, there are active decision makers, complex processes of data use at play, and a great deal of variation both within and across contexts. These decisions are also intertwined and may need to be considered in relationship to one another. For example, the decision to challenge deficit beliefs about student abilities using data will not likely occur without a broader definition of what counts as data, especially beyond standardized assessments. Conversely, the decision to use multiple forms of data alone will not lead to a goal of expanding students' opportunities to learn, without the value for using data to challenge assumptions or building on students' assets. Educators and policymakers are faced with a set of critical choices in their work with data that can profoundly affect students' daily educational experiences and trajectories. Some of these choices are more likely to expand opportunities for students, whereas others are more likely to limit them. This especially true when there is a pattern of choices within a particular school.

When accountability frameworks dominate teacher conversations and instructional interventions are focused on students for whom schools will get the most "bang for the buck" on standardized measures, students are more likely to suffer. The curriculum in these schools is likely to be narrower, with student achievement primarily focused on single measure, which often disadvantages low-income, minority students. The use of data for tracking, long-term ability grouping, and placement is of particular concern, as it can serve to reinforce hierarchies among students. In schools where data are more often used to actively challenge stereotypes, to examine student growth as well as weaknesses, and to differentiate instruction in dynamic ways, students likely have more opportunities for success. Students are also more likely to thrive when educators share responsibility for their success.

Just as equity needs to be an explicit goal in data use practices, research on data use also needs an equity lens. Educational improvement and policy lenses tend to prioritize organizational changes associated with reforms. Although many educational researchers list equity as an important goal of school reform, this concern may not be central to their work. Putting equity at the center of studies of data use—and indeed all educational reforms—may involve the use of different research questions, methodologies and/or theoretical frameworks. It requires explicit attention to questions and issues raised in this article and in others. In what ways do data use lead to inequitable outcomes for differing groups of students and what are the specific data use conditions and processes that contribute to unequal or equal opportunities to learn? How do educators' assumptions, often tied to beliefs about students' backgrounds, become challenged or reified by data use? What data use practices support shifts away from deficit thinking and fixed beliefs about intelligence or ability? How do data-informed leaders with an explicit equity-orientation build a school culture that expands opportunities to learn not just for students but also for teachers?

This equity-centered agenda on data use also may involve assembling diverse teams of researchers, a willingness to engage in difficult conversations, and carefully framing findings to do justice to the complexity of equity-driven change. With data use efforts occurring in a wide range of settings across the globe, researchers have an unprecedented opportunity to further document ways in which data use opens or closes doors for students. Studying data use through an equity lens provides unique affordances that can help to eliminate problematic practices and promote better schooling for all.

References

- Bertrand, M., Marsh, J.A., 2015. Teachers' sensemaking of data and implications for equity. *Am. Educ. Res. J.* 52 (5), 861–893.
- Booher-Jennings, J., 2005. Below the bubble: "Educational triage" and the Texas accountability system. *Am. Educ. Res. J.* 42 (2), 231–268.
- Braaten, M., Bradford, C., Kirchgassler, K.L., Baracos, S., 2017. How data use for accountability undermines equitable science education. *J. Educ. Adm.* 55 (4), 427–446.
- Christman, J.B., Neild, R.C., Bulkley, K., Blanc, S., Liu, R., Mitchell, C., Travers, E., 2009. Making the Most of Interim Assessment Data. Lessons From Philadelphia. Retrieved from http://www.researchforaction.org/wp-content/uploads/publication-photos/41/Christman_J_Making_the_Most_of_Interim_Assessment_Data.pdf.
- Coburn, C., Talbert, J., 2006. Conceptions of evidence use in school districts: mapping the terrain. *Am. J. Educ.* 112 (4), 469–495.
- Coburn, C.E., Turner, E.O., 2011. Research on data use: a framework and analysis. *Measurement* 9 (4), 173–206.
- Daly, A.J., 2018. Data, dyads, and dynamics: exploring data use and social networks in educational improvement. *Teach. Coll. Rec.* 114 (11), 1–38.
- Datnow, A., Park, V., 2014. *Data-Driven Leadership*. Jossey Bass, San Francisco.
- Datnow, A., Park, V., 2019. *Professional Collaboration With Purpose: Teacher Learning for Equitable and Excellent Schools*. Routledge, New York.
- Datnow, A., Park, V., Kennedy-Lewis, B., 2012. High school teachers' use of data to inform instruction. *J. Educ. Stud. Placed Risk* 17, 247–265.
- Datnow, A., Choi, B., Park, V., St John, E., 2018. Teacher talk about student ability and achievement in the era of data-driven decision making. *Teach. Coll. Rec.* 120 (4), 1–34.
- Davidson, K.L., Frohbieter, G., 2011. District Adoption and Implementation of Interim and Benchmark Assessments. Report No. 806. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA.
- Diamond, J.B., Cooper, K., 2007. The uses of testing data in urban elementary schools: some lessons from Chicago. *Yearbk. Natl. Soc. Stud. Educ.* 106 (1), 241–263.
- Dowd, A., 2005. *Data Don't Drive: Building a Practitioner-Driven Culture of Inquiry to Assess Community College Performance*. University of Massachusetts, Lumina Foundation for Education, Boston.
- Firestone, W.A., González, R.A., 2007. Culture and processes affecting data use in school districts. In: Moss, P.A. (Ed.), *Evidence and Decision Making*. Yearbook of the National Society for the Study of Education. Blackwell, Malden, MA, pp. 132–154.
- Gannon-Slater, N., La Londe, P.G., Crenshaw, H.L., Evans, M.E., Greene, J.C., Schwandt, T.A., 2017. Advancing equity in accountability and organizational cultures of data use. *J. Educ. Adm.* 55 (4), 361–375.
- Garner, B., Kahn, J., Horn, I., 2017. Teachers interpreting data for instructional decisions: where does equity come in? *J. Educ. Adm.* 55 (4), 407–426.
- Gillborn, D., Youdell, D., 1999. *Rationing Education: Policy, Practice, Reform, and Equity*. Open University Press, Buckingham.
- Goertz, M.E., Nabors Oláh, L., Riggan, M., 2010. From Testing to Teaching: The Use of Interim Assessments in Classroom Instruction. CPRE Research Report No. RR-65. Consortium for Policy Research in Education, Philadelphia, PA.
- Halverson, R., Grigg, J., Pritchett, R., Thomas, C., 2007. The new instructional leadership: creating data-driven instructional systems in schools. *J. Sch. Leader.* 17 (2), 159–193.
- Hargreaves, A., Shirley, D., 2012. *The Global Fourth Way: The Quest for Educational Excellence*. Corwin Press, Thousand Oaks, CA.
- Heppen, J., Jones, W., Faria, A., Sawyer, K., Lewis, S., Horwitz, A., Casserly, M., 2012. *Using Data to Improve Instruction in the Great City Schools: Documenting Current Practice*. American Institutes for Research and The Council of Great City Schools, Washington, DC.
- Horn, I., Kane, B., Wilson, B., 2015. Making sense of student performance data: data use logics and mathematics teachers' learning opportunities. *Am. Educ. Res. J.* 52 (2), 208–242.
- Huguet, A., Farrell, C.C., Marsh, J.A., 2017. Light touch, heavy hand: principals and data-use PLCs. *J. Educ. Adm.* 55 (4), 376–389.
- Ingram, D., Louis, K.S., Schroeder, R., 2004. Accountability policies and teacher decision making: barriers to the use of data to improve practice. *Teach. Coll. Rec.* 106 (6), 1258–1287.
- Jimerson, J.B., Childs, J., 2017. Signal and symbol: how state and local policies address data-informed practice. *Educ. Pol.* 31 (5), 584–614.
- Jimerson, J.B., Wayman, J.C., 2015. Professional learning for using data: examining teacher needs and supports. *Teach. Coll. Rec.* 117 (4), 1–36.
- Knapp, M.S., Copland, M.A., Swinnerton, J.A., 2007. Understanding the promise and dynamics of data-informed leadership. *Yearbk. Natl. Soc. Stud. Educ.* 106 (1), 74–104.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in U.S. schools. *Educ. Res.* 35 (7), 3–12.
- Lai, M.K., Schildkamp, K., 2016. In-service teacher professional learning: use of assessment in data-based decision-making. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York, pp. 77–94.
- Lachat, M.A., Smith, S., 2005. Practices that support data use in urban high schools. *J. Educ. Stud. Placed Risk* 10 (3), 333–339.
- Lipman, P., 2004. *High Stakes Education: Inequality, Globalization, and Urban School Reform*. Routledge Falmer Press, New York.
- Mandinach, E.B., Honey, M. (Eds.), 2008. *Data Driven School Improvement: Linking Data and Learning*. Teachers College Press, New York, NY.
- Marsh, J., Farrell, C., 2014. Supporting teachers with data-driven decision making: a framework for understanding capacity-building. *Educ. Manag. Adm. Leader* 1–21.
- Marsh, J.A., Pane, J.F., Hamilton, L.S., 2006. *Making Sense of Data-Driven Decision Making in Education*. RAND, Santa Monica, CA.
- Marsh, J.A., 2012. Interventions promoting educators' use of data: research insights and gaps. *Teach. Coll. Rec.* 114 (11), 1–48.
- McNeil, L., 2002. *Contradictions of School Reform: Educational Costs of Standardized Testing*. Routledge.
- McNeil, L.M., 2005. Faking equity: high-stakes testing and the education of Latino youth. Leaving children behind: how "Texas-style" accountability fails Latino youth. In: Valenzuela, A. (Ed.), *Leaving Children Behind: How "Texas-Style" Accountability Fails Latino Youth*. SUNY Press, Albany, NY, pp. 57–111.
- Means, B., Padilla, C., DeBarger, A., Bakia, M., 2009. *Implementing Data-Informed Decision Making in Schools—Teacher Access, Supports and Use*. U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Washington, DC.
- Neuman, S.B., 2016. Code red: the danger of data-driven instruction. *Educ. Leader* 74 (3), 24–29.
- Oakes, J., Gamoran, A., Page, R.N., 1992. Curriculum differentiation: opportunities, outcomes, and meanings. In: Jackson, P. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 570–608.
- Oakes, J., Wells, A.S., Jones, M., Datnow, A., 1997. Detracking: the social construction of ability, cultural politics, and resistance to reform. *Teach. Coll. Rec.* 98 (3), 482–510.
- Oakes, J., 2005. *Keeping Track: How Schools Structure Inequality*, second ed. Yale University Press, New Haven, Conn.
- Oláh, L.N., Lawrence, N.R., Riggan, M., 2010. Learning to learn from benchmark assessment data: how teachers analyze results. *Peabody J. Educ.* 85, 226–245.
- Park, V., Datnow, A., 2017. Ability grouping and differentiated instruction in an era of data-driven decision making. *Am. J. Educ.* 123 (2), 281–306.
- Park, V., Daly, A.J., Guerra, A.W., 2013. Strategic framing: how leaders craft the meaning of data use for equity and learning. *Educ. Pol.* 27 (4), 645–675.
- Park, V., 2018. Data conversation moves: towards data-informed leadership for equity and learning. *Educ. Adm. Q.* 1–34.
- Pierce, R., Chick, H., 2011. Teachers' intentions to use national literacy and numeracy assessment: a pilot study. *Aust. Educ. Res.* 38 (3), 433–477.
- Pollock, M., 2017. *Schooltalk: Rethinking What We Say to—and About—Students Every Day*. New Press, New York.
- Santelices, M.V., Wilson, M., 2010. Unfair treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harv. Educ. Rev.* 80 (1), 106–134.
- Schildkamp, K., Poortman, C., 2015. Factors influencing the functioning of data teams. *Teach. Coll. Rec.* 117 (4).
- Schildkamp, K., Karbautzki, L., Vanhoof, J., 2014. Exploring data use practices around Europe: identifying enablers and barriers. *Stud. Educ. Eval.* 42, 15–24.
- Shepard, L., Davidson, K., Bowman, R., 2011. *How Middle School Mathematics Teachers Use Interim and Benchmark Assessment Data*. CSE Technical Report. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA.
- Shirley, D., 2017. *The New Imperatives of Educational Change: Achievement With Integrity*. Routledge, New York.
- Skrlla, L., Scheurich, J.J., Garcia, J., Nolly, G., 2004. Equity audits: a practical leadership tool for developing equitable and excellent schools. *Educ. Adm. Q.* 40 (1), 133–161.

- Snodgrass Rangel, V., Bell, E., Monroy, C., 2017. Teachers' sensemaking and data use implementation in science classrooms. *Educ. Urban Soc.* 51 (4), 526–544.
- Valenzuela, A. (Ed.), 2005. *Leaving Children Behind: How "Texas-Style" Accountability Fails Latino Youth*. SUNY Press, Albany, NY.
- Valli, L., Cooper, D., and Franks, L., 1997. Professional development schools and equity: a critical analysis of rhetoric and research, In: M. W. Apple (Ed.), *Review of Research in Education*, Am. Educ. Res. Assoc. Washington, DC, 51–304.
- Vanlommel, K., Van Gasse, R., Vanhoof, J., Petegem, V., 2017. Teachers' decision making: data-based or intuition driven? *Int. J. Educ. Res.* 83, 75–83.
- Wardrip, P.S., Herman, P., 2017. We're keeping on top of the students': making sense of test data with more informal data in a grade-level instructional team. *Teach. Dev.* 22 (1), 31–50.

Principal's information-processing mechanisms (PIPMs) as an effective tool for learning-centered leadership that reduces teacher's intent to leave

Mowafaq Qadach^a, Rima'a Da'as^b, and Idit Finkelstein^c, ^a Al-Qasemi Academic College of Education, Department of Management and Organization of Educational Institutions and Ono Academic College, Israel; ^b School of Education, The Hebrew University of Jerusalem, Jerusalem, Israel; and ^c Ono Academic College, Kiryat Ono, Israel

© 2023 Elsevier Ltd. All rights reserved.

Introduction	177
The principal's role in teacher retention: teacher's intent to leave	178
Principals' support of school learning	179
School principals' information processing mechanisms (PIPMs)	179
School principal learning-centered leadership impact on teacher's intent to leave	180
Method	180
Participants and procedures	180
Measures	181
PIPMs	181
Teacher's intent to leave	181
Control variables	181
Results	181
Discussion	182
Practical implications	182
Appendix: PIPM questionnaire	183
Factor 1: storing and retrieving information	183
Factor 2: receiving information from students, parents, the community, and superiors	183
Factor 3: disseminating information to teachers	183
Factor 4: analyzing external and internal information	183
References	183

Purpose: Keeping experienced and competent teachers in schools is becoming an important challenge for school leadership. Hence, this research explored an innovative relationship between principal's information-processing mechanisms (PIPMs) (as a school level variable that reflects school principals' ability to gather, assimilated and disseminate information), and a teacher's intent to leave his or her school (as the individual level variable).

Design/methodology/approach: Data were collected through a survey of 101 principals and 1130 teachers from 101 Jewish and Arab elementary schools randomly selected from the database of the Israeli education system. To test the proposed multilevel model, we conducted a multi-level SEM by the M-plus program in which individual responses are nested within organizational units (schools).

Findings: With controlling for principals' seniority and teachers' tenure in teaching, the relationship between the quality of leadership learning, as indicated by PIPMs, and teacher's intent to leave is significant and negative. The analysis confirms that PIPMs negatively predicts teacher's intent to leave.

Practical implications: Our findings point to principals' information processing as an important factor that contributes to a teacher's intent to leave, thus highlighting the role principal's leadership learning plays in teacher retention and how it can aid in reducing the phenomenon of teacher's intent to leave.

Originality/value: This study adds to the body of research directed at identifying school principals' learning processes, which may decrease a teacher's intent to leave and are amenable to leadership intervention.

Introduction

Organizational learning is vital to school performance, with schools that incorporate this approach shown to perform better (for example, [Mulford and Silins, 2011](#); [Qadach et al., 2020](#)). Such an approach entails the development of learning processes, strategies and structures geared toward a more effective response to changing, uncertain or dynamic circumstances ([Fauske and Raybould, 2005](#); [Giles and Hargreaves, 2006](#); [Kruse, 2003](#)). Adopting an organizational learning approach necessitates the establishment and nurturing of a framework for the continuous collective processing, understanding and application of knowledge about teaching and learning ([Louis, 2006](#); [Schechter, 2008](#)). In this chapter, we focus on the impact of school principals' learning in the context of their implementation of school principal's information processing mechanisms (PIPMs) on teacher's intent to leave.

The principal's role in teacher retention: teacher's intent to leave

Withdrawal behavior has been studied by looking at employees' responses to items concerning contemplating leaving and the desire and likelihood of quitting (Hanisch and Hulin, 1991). A teacher's intent to leave can be viewed as a form of withdrawal, that is, a set of attitudes and behaviors characteristic of workers who become less participative (Kaplan et al., 2009) or leave the job altogether (Rahim and Afza, 1993). Withdrawal behavior in teachers manifests itself as increased burnout and decreased motivation and organizational commitment (Conley and You, 2009), and lower work effort and productivity (Fore et al., 2002; Tett and Meyer, 1993). Consequently, much research in recent years has focused on identifying administrative support and work-related factors that can reduce a teacher's desire to leave.

Intent to leave is viewed as the extent to which workers seek alternative employment elsewhere (Shapira-Lishchinsky and Rosenblatt, 2009). Intent to leave is considered a contributing factor to reduced school effectiveness (Ingersoll, 2012; Shaw et al., 2005), teacher productivity and motivation, and overall school morale (Tett and Meyer, 1993).

Schools worldwide find it difficult to retain effective teachers, which entails considerable human resource management costs (Harris et al., 2005). More importantly, a teacher's intent to leave has unique organizational and pedagogical implications. Those seeking to leave are typically the more qualified teachers who are more likely to secure employment elsewhere. This jeopardizes the school's teaching standards and negatively impacts the motivation and effort of the remaining teachers in the school (Ingersoll, 2012; Josephson et al., 2008; Parry, 2008; Rosenblatt and Shirom, 2006). A teacher's intent to leave also increases the pressure on the remaining faculty (Torrez and Kritsonis, 2008), leading to a decline in morale, with additional staff potentially following suit (Shaw et al., 2005). It can also damage a school's reputation and teaching staff cohesion, which inevitably reduces the school's effectiveness (Ingersoll, 2012; Schaefer et al., 2012; Tett and Meyer, 1993). It is, therefore, vital to elucidate the reasons behind a teacher's decision to leave a school or the educational system altogether to reduce the scope of this negative phenomenon (Lopez, 2018).

Teachers contemplate quitting when they have little work satisfaction and consider leaving a low-risk act. The lack of professional opportunities, limited professional autonomy, unsatisfactory wages and poor job satisfaction all play a dominant role in considering to leave (Fochsen et al., 2005; Morrell, 2004; Zembylas and Papanastasiou, 2004).

As discussed earlier, an ever-growing body of literature affirms the key role principals play in improving school effectiveness and promoting change (for example, Da'as, 2019; Sun and Leithwood, 2017). It is particularly important when it comes to teacher retention (Cha and Cohen-Vogel, 2011; Hambrick, 2007). Research has linked leaders' actions to teacher job satisfaction, turnover and retention (Brown and Wynn, 2009; McCann et al., 2009). Other studies have linked teacher's perception of their school leaders and their intent to leave (Hughes et al., 2015; Qadach et al., 2019). Teachers' intent to leave was also inversely correlated with leadership supportive of an environment that allows teachers to exercise their potential as educators (Markow and Pieteres, 2009). In line, highly rated principals succeeded in retaining good teachers, with principal support an important factor in new teacher retention (Grissom and Loeb, 2011).

A recent important study in an elementary school shone a light on the negative correlation between the organizational learning climate and a teacher's intent to leave at the individual level ($\gamma = -0.14$, $p < 0.05$; Da'as et al., 2020). It likewise found a positive relationship between the innovative approach of a principal and organizational learning atmosphere ($\gamma = 0.30$, $p < 0.001$). Further, the study confirmed the mediation effect: the relationship between principals' innovative behavior and a teacher's intent to leave is mediated by the perceptions of the organizational learning climate. This previous study asserted that school principal behaviors and the learning climate that he or she designed can influence teacher's intent to leave on the individual level.

Our recent studies identified two other school-level variables that exhibit a direct, negative relationship with teacher's desire to leave: collective teacher efficacy (CTE) and shared vision (Qadach et al., 2019). Conversely, instructional leadership showed no such correlation. School culture, principal leadership (i.e., is supportive and creates school environments conducive to learning), and the collegial relationships are interrelated key elements that were found to improve the social context of work and predict teacher retention (Johnson et al., 2012). Teachers' perceptions of the school administration have by far the greatest influence on teacher retention decisions (Boyd et al., 2011). Further, Studies have shown that the quality of school leadership is a strong indicator of teacher retention (e.g., Boyd et al., 2011; Grissom, 2011; Ladd, 2011). This finding suggests that principals that succeed in creating a more positive school climates are likewise better able to retain teachers. This is because teachers receive greater support, more opportunities for professional growth, and more favorable working conditions, the outcome of which is greater job satisfaction and loyalty (Ingersoll, 2001; Johnson et al., 2012; Kraft et al., 2016).

A new approach to teacher retention has been coined *strategic retention*, denoting a differential approach whereby efforts are made both to retain effective teachers and to not retain ineffective ones (Grissom and Bartanen, 2019). For example, principals may seek to reduce the overall teacher turnover rate, but direct greater efforts toward their top teachers. Another strategy is to actively encourage turnover among low performers, through, for example, opting not to renew a contract, "counseling out" or making the job less palatable (Drake et al., 2016).

Previous findings indicated that teachers who believed their principal took an inclusive approach to leadership exhibited greater enthusiasm and commitment to their school. Conversely, teachers who considered their principal's approach to be instrumental, i.e., demanding compliance rather than including them in school decisions and processes, experienced frustration and an increased intent to leave the school (Johnson, et al., 2012).

Principals' support of school learning

Considering schools' complex, dynamic environment, where interconnected components must work together to ensure the continued high-performance of teachers (Ingersoll, 2012; Schaefer et al., 2012), it is imperative to investigate the way by which principals affect organizational features and processes and employee perceptions and attitudes. Previous studies indicate that teacher performance and classroom practices are all influenced by leadership attributes and behavior (for example, Louis et al., 2010). Sebastian et al. (2017) found that leadership affects learning indirectly by shaping conditions that promote effective teaching and learning. Louis et al. (2010) pointed to shared and instructional leadership as an indirect contributor to student achievements. Moreover, Day et al. (2016) founded that both transformational and instructional school principal leadership directly and indirectly promote schools', teachers' and students' success. These outcomes seem to be a result of the way teachers organize themselves into professional communities with strongly held norms and values, reflective discussions about instruction, and a sense of collective responsibility for student learning.

As those tasked with improving school performance, school principals also play a key role in organizational learning processes (Da'as et al., 2019; Davis, 2005; Desjardins and Donaldson, 2008; Drago-Severson and Pinto, 2009; Leithwood, 2004; Waters and Grubb, 2004). Charismatic leadership is conducive to a positive climate of organizational learning in the school, namely, its employees' attitudes toward the work setting's positive or negative impact on learning, and to the formation of trust among the teaching staff, which translates into school outcomes (according to superintendent and parent evaluation) (Berson et al., 2015).

In fact, the last decade has seen organizational learning become a major focus of school leaders. Among the driving forces for principal's support of this focus on organizational learning (AKA learning-centered schools) is the intricate socio-cultural milieu and policy accountability reforms (for example, Darling-Hammond et al., 2007; Drago-Severson, 2009). This has given rise to approaches such as "Learning-focused leadership" (Knapp and Feldman, 2012) or "learning-centered leadership" (Murphy, 2006), which deal with leaders' roles in student, professional, and system learning. Strong instructional leaders have been shown to form a shared vision (Hallinger and Wang, 2015), create a school culture (Heck and Hallinger, 2014), establish a positive instructional climate (May and Supovitz, 2011), enhance collective teacher efficacy (Qadach et al., 2019), and participate in curricular and instructional issues with their teaching staff (Hornig and Loeb, 2010).

To bring about organizational learning, the leader is tasked with overseeing collective learning processes and formulating a school-wide learning agenda (Knapp and Copland, 2006). According to Ash and Persall (2000), the principal is the chief learning officer, who is required to ensure the existence of a reward-based environment conducive to teacher leadership, collaboration and professional development. This necessitates principals to ensure that relevant knowledge is gathered and then appropriately shared among the teaching faculty and staff. Information gathering can be achieved through experiential learning, vicarious learning, recruiting new members, and searching and observing the environment. Experiential learning denotes intra-organizational experiments and self-appraisal. Vicarious learning entails learning from other organizations' strategies and technologies. In practical terms, principals need to constantly provide new information and opportunities for collaborative planning and learning and to set in place an environment that promotes leadership throughout the school (see also McGough, 2003). Leaders are accountable for arranging the interactive social environment and ensuring that adequate resources are available to support teacher learning (for example, Halverson et al., 2005). They also need to be capable of building strong professional communities that can foster teacher collaboration, dialogue, and learning (Fullan, 2009).

To form such a community, principals must rely on methodological research concerning the development of school learning, but tailor it to suit their school's particular characteristics and needs. In other words, they must adopt information-processing mechanisms that ensure efficient and well-targeted gathering and assimilation of information on changes occurring both within the school and in the community, complemented by effective means of the knowledge's analysis, storage and dissemination in an effort to advance the school's interests. This includes addressing social and political pressures while shielding their staff from detrimental policies, initiating school-improvement plans that align with external reforms, and catering to the needs of the school's students and community. Effort should also be directed to ensure ongoing learning, even when faced with counter social, political and economic forces. Collaborating with school community members can aid in determining the school's particular attributes that cannot be assessed via quantitative measurements that will ensure the school's ongoing success (Kochan et al., 2002). Organizational learning entails both extending one's gaze beyond the school grounds in search of learning opportunities and working within the school to develop professional learning communities (Knapp and Feldman, 2012). Particularly given the growing emphasis on accountability and standardization, principals must actively strive to realize a school culture that features both individual and collective learning (Desjardins and Donaldson, 2008; Drago-Severson and Pinto, 2009; Schlechty, 2009).

School principals' information processing mechanisms (PIPMs)

The great promise of learning-centered school systems is an improvement in student scholastic performance (Browne-Ferrigno, 2007; Fullan, 2009). This has led to principal accountability for teaching and learning issues (Orr, 2006), with the expectation that principals establish collective learning opportunities in elementary schools for both faculty and students, collectively referred to as "principals' information processing mechanisms" (PIPMs), an approach explored by Schechter and Qadach (2016). In simple terms, PIPMs is a data-gathering and processing approach that enables school principals to collect information about school and relevant external processes and use the data as a guide to inform their leadership decisions and practices. In order to elucidate the ways by which principals can improve structural and procedural school learning arrangements for assembling, assessing, storing,

disseminating, retrieving and employing school-related information, the Authors adopted an information-processing perspective, which is based on assembling and analyzing information from internal (organizational) and external (environmental) sources. This approach, outlined in detail in [Schechter and Qadach \(2016\)](#), led the Authors to define the following four-factor model, which best suited their empirical results and conceptual formulation (see [Appendix A](#)), which together provide an indication of the quality of principals' learning mechanisms:

1. *Information storage and retrieval* – the manner by which principals store personal and organizational experiences, code them into school memory, and access them to guide decision making and actions (for example, “I document meetings with the superintendent, municipal authority and/or parents”).
2. *Stakeholder information retrieval* – the manner by which principals attain information from students, parents, community members and superiors (for example, “I hold meetings with my superiors in order to learn from external events that affect the school”).
3. *Teacher information dissemination* – the manner by which principals convey information to school personnel for the purpose of collective analysis (for example, “I disseminate teachers' success stories among the staff”).
4. *External and internal information analysis* – the manner by which the school principal and teaching faculty discuss and analyze important events and/or useable data that affect the school, such as the school atmosphere and student academic performance (for example, “I examine external data [such as national exams] with the teaching staff”).

PIPMs support a learning-centered school approach by forming a central social platform for the formulation, analysis and sharing of knowledge and new standard routines and procedures and their dissemination to all the school's stakeholders ([Ellis and Shpilberg, 2003](#); [Lipshitz et al., 2002](#)). As noted by [Schechter and Qadach \(2016\)](#), the PIPM perspective is supported by the growing realization that most principal-assessment instruments focus on general management (for example, implementing vision) or traits rather than on the leaders' critical behaviors and actions within the learning-centered leadership framework ([Hitt and Tucker, 2016](#); [Leithwood and Sun, 2012](#)).

[Schechter and Qadach, \(2016\)](#) validated the four-factor model via constructing a PIPM questionnaire, which they then used to assess the correlation of the developed scale with established work-related constructs. [Schechter and Qadach \(2016\)](#) found PIPMs to positively and significantly correlate with teachers' organizational commitment, as well as with the three factors of instructional leadership (IL), namely, determining the school mission, overseeing the instructional program, and working toward a school atmosphere conducive of learning. All the PIPMs factors, except for “external and internal information analysis”, showed positive (moderate) and significant correlation with principal IL factors.

We also assessed PIPMs' influence on teacher's organizational commitment ([Schechter and Qadach, 2016](#)). Here, too, the entire PIPMs questionnaire moderately correlated with affective commitment, but not with normative or continuance commitment. These findings are in line with earlier studies that found affective commitment to be the most dominant determinant among the three commitments in terms of staff behavior ([Cohen, 2003](#); [Shapira-Lishchinsky and Rosenblatt, 2009](#)). Further, a recent study found that PIPM's directly and positively predicts school principal instructional leadership (IL), and the IL mediates the relationship between PIPM's and school Organizational Learning Mechanisms (OLM's) ([Qadach et al., 2020](#)).

In summary, the PIPM questionnaire had a notable, positive, moderate correlation with instructional leadership and teachers' affective commitment, but not with normative commitment or continuance commitment. Thus, the PIPM questionnaire is an efficient tool for principal evaluation of the effectiveness of various strategies for promoting learning-centered schools. It could further offer the means to self-monitor the contribution of the principal to disseminating organizational learning approaches and subsequently modifying the strategies to reach a more optimal balance between the various components ([Schechter and Qadach, 2016](#)). Information analysis actions, for example, could be balanced with information collection ones to provide information that aids the implementation of effective school initiatives.

School principal learning-centered leadership impact on teacher's intent to leave

A yet to be explored aspect of the principal's learning-centered leadership is to delineate what are the strategies, as reflected in PIPMs, that directly influences teacher retention. To investigate this issue, we applied PIPMs in the context of a teacher's intent to leave (see [Fig. 1](#)) with the goal of identifying principal learning-centered leadership variable that counteract a teacher's intent to leave that can be employed as a means of leadership intervention.

Method

Participants and procedures

We randomly selected 101 elementary schools from a school list on the Israeli educational system website (i.e., from all of Israel). In this study, 101 principals and 1130 teachers participated; the teachers were randomly selected from each school. The principals filled a questionnaire on PIPMs and demographic variables and the teachers completed a questionnaire regarding the intent to leave and demographic variables. School size was determined based on the number of teachers in the school; the average was 33 (SD = 11.42). We ensured a random sampling minimum of 30% of the teaching staff at

each school. The approval of the Israeli Ministry of Education and of each school was obtained prior to administering the questionnaire.

The principals' average tenure in their job was 9.20 years ($SD = 7.01$). Ninety one percent of the teachers were women, and their average number of teaching years was 21.0 ($SD = 8.19$); 26% had a B.A. degree, and 67% an M.A. degree.

Measures

PIPMs

We used a questionnaire developed by [Schechter and Qadach \(2016\)](#), which included 22 items. Each principal was asked to complete one questionnaire. The principals rated the items on a 5-point Likert scale, ranging from *never* (1) to *always* (5). The questionnaire addressed four components ([Fig. 1](#)): (i) Receiving information (eight items), for example, "I meet with students in order to get exposed to difficulties, special requests, comments and insights," with an Alpha Cronbach of 0.84. (ii) Storing and retrieving information (five items), for example, "I document meetings with the superintendent, municipal authorities, and parents," with an Alpha reliability of 0.79. (iii) Analyzing external and internal information (four items), for example, "I examine external data (such as national exams) with the teaching staff," with an Alpha reliability of 0.80. (iv) Disseminating information to teachers (five items), for example, "I disseminate reports concerning professional changes and innovations among the staff," with an Alpha Cronbach of 0.81. Construct validity for PIPMs was measured by confirmatory factor analysis, using the AMOS program. The data fit the model, with a comparative fit index (CFI) of 0.90 and a root mean square error of approximation (RMSEA) of 0.06. We considered school principals' information processing mechanisms to be a unified construct affecting all members of a unit ([Kark et al., 2003](#)). Thus, an individual principals' PIPM score PIPMs was considered an organizational construct that reflects the school level.

Teacher's intent to leave

Teachers completed the 5-item intent-to-leave subscale of the [Shapira-Lishchinsky \(2012\)](#) survey. Each teacher was asked to complete one survey. Teachers rated items, such as "I'm considering contacting professional bodies about other work options," on a 5-point Likert scale, ranging from *strongly disagree* (1) to *strongly agree* (5), with an Alpha Cronbach of 0.88. The CFA results indicated good fit indices: $\chi^2 (5) = 29.47$, RMSEA = 0.05, CFI = 0.96, TLI = 0.94, IFI = 0.95.

Control variables

We controlled for the effects of principals' tenure in their job and teachers' years of teaching, variables that have been shown to be potentially relevant to school processes and outcomes (for example, [Qadach et al., 2020](#)).

Results

The data collected reflect a hierarchical structure in which individual responses are nested within organizational units (schools) ([Qadach et al., 2020](#), p. 14). Thus, we conducted multi-level structural equation modeling by the M-plus program ([Muthén and Muthén, 1998](#)). Our results indicate a good fit between the model ([Fig. 1](#)) and the data (CFI = 0.98; RMSEA = 0.05). As seen in

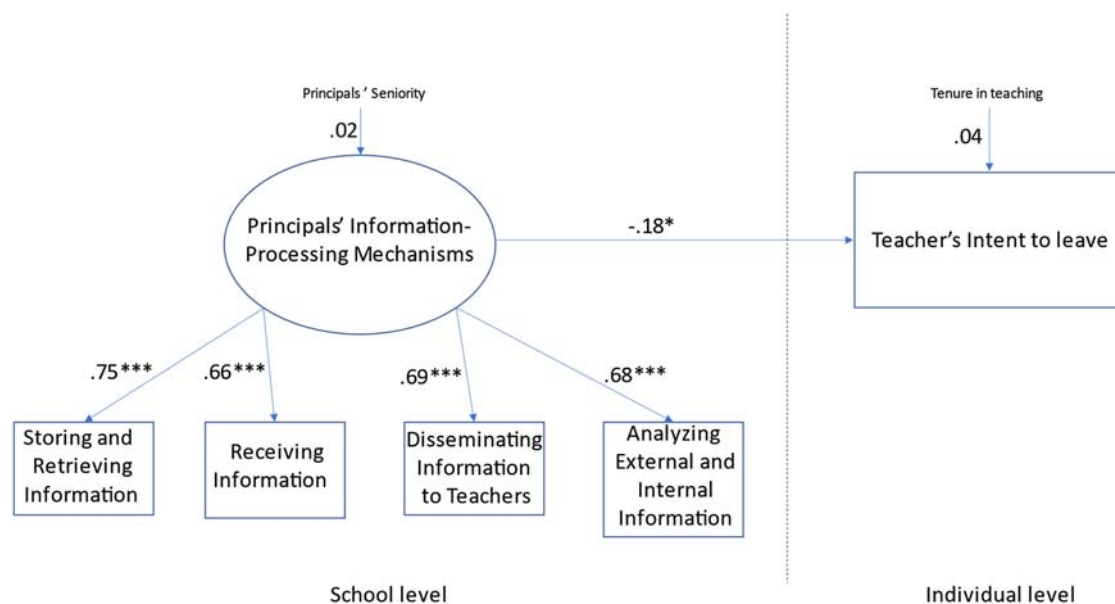


Fig. 1 Standardized path coefficients from principals to teacher's intent to leave ($N = 103$); $*** p < 0.001$; $* p < 0.05$

Fig. 1, with controlling for principals' seniority and teachers' tenure in teaching, the relationship between PIPMs and teacher's intent to leave is significant and negative ($\gamma = -0.18, p < 0.05$). In other words, once the effect of principal seniority as well as teacher seniority variables are controlled, there is a statistically significant negative relationship between principal learning-centered actions (PIPM'S) and the tendency of a teacher to leave the profession. A school principal learning-focused leadership through the assimilation of PIPMS mitigates this intent.

Discussion

This study adds to the limited existing research that found that teachers who are more likely to leave the profession are less likely to have a principal who creates a learning-centered environment that helps them exercise their potential as teachers (Markow and Pieteres, 2009). It further indicates that highly rated principals are successful in keeping high-performing teachers (Grissom and Loeb, 2011). Thus, as teachers who intend to leave often exhibit increased burnout and decreased motivation and organizational commitment (Conley and You, 2009), this study adds to the body of research directed at identifying the power and impact of PIPMs as a learning leader for the school teachers, which may decrease a teacher's intent to leave. These results are aligned with previous research that found that school principals' innovative behavior promotes an Organizational Learning Climate (OLC), which, in turn, negatively impacts teachers' intent to leave in the Israeli education system (Da'as et al., 2020).

School learning is critical to realizing school effectiveness. This is because schools that operate as learning organizations perform better than schools that do not (Da'as and Qadach, 2018). Furthermore, PIPMs, which offers school principals daily opportunities to share and evaluate information regarding their practices as a means to improve them further, was shown to significantly decrease a teacher's intent to leave. The PIPM learning processes, strategies and structures strengthen a principal's ability to respond in an effective manner to complex and dynamic situations, which inevitably affect teacher behaviors at the individual level, and, in line, a teacher's intention to leave. This study, therefore, adds to the ongoing efforts to find proxies and links between organizational-level and individual-level factors in the intricate workings of the school system (Da'as et al., 2019).

While limited research has been conducted on the relationship between PIPMs and a teacher's intent to leave, this study suggests a new theoretical contribution by confirming the negative correlations between the two variables. This emphasizes the significant role of PIPMs in schools, which according to our current results is a key predictor of future teacher behavior. In other words, when school principal PIPMs are implemented and the school organization enjoys the fruits of the learning added value, this will possibly decrease a teacher's intent to leave the school (Richardson et al., 2008). This study adds to the limited existing research directed at identifying school principal factors that decrease a teacher's intent to leave.

It is reasonable to assume that leaders provide their teachers with a supportive environment that allows them to expand their learning capacity, lead, and work together to improve their students' achievements (Seaton et al., 2008). As proposed by Printy (2008), leaders both set the agenda for teacher learning and school improvement and serve as brokers of knowledge and providers of resources in support of teacher learning and innovation initiatives. And setting the agenda for teacher learning (Printy, 2008), by way of modeling their learning (PIPMs) to teachers, most likely further promotes organizational learning and teacher cooperation (for example, information sharing) and work motivation. This, in turn, should reduce the likelihood that teachers will consider leaving their school, work or profession.

Principals who rely more heavily on PIPM processes may also elicit stronger appreciation from their teaching staff, which is probably based more on emotional attachment than a sense of obligation. The outcome may be a greater willingness of teachers to get involved and invest time in formal and informal activities in an effort to achieve the school's mission. Our findings indicate that principals that effectively operate PIPMs will be rewarded with heightened teacher affective commitment (Schechter and Qadach, 2016), and with it, a lower likelihood of teachers leaving the school.

In summary, in this complex educational leadership context, we argue that, in order to decrease a teacher's intent to leave, in line with the PIMP model, school leaders need to receive, store and retrieve information, think and plan strategically, analyze external and internal information, so as to make decisions effectively, disseminate information to teachers, and communicate effectively with teachers and promote their emotional well-being.

Practical implications

Our study's findings are important for policy-makers and for school superintendents, both of whom are responsible for training principals. Leadership programs should consider focusing on helping principals develop and promote PIPMs. In this regard, the nature of the interface between the principals' contribution to organizational learning in schools must be assessed, especially those exchanges that either foster or hamper the principals' efforts to adopt information processing mechanisms to decrease teachers' intent to leave.

Further, the results highlight the pivotal role of school principal learning in helping determine teacher's withdrawal behavior. Day-to-day principal concerns often preclude teams and organizations from investing in learning activities. We, therefore, propose to design organizational mechanisms or systems (OLMs) that continually address learning. For example, ongoing team meetings

could be devoted to analyzing, seeking, sharing and disseminating information as mechanisms for learning schools (Qadach et al., 2020).

This study also has important implications in the educational realm, specifically for policies concerning public school administration. Principals face the challenge of boosting school performance, and facilitating a climate of organizational learning may provide them with an effective means to do so. Our findings suggest that a principal's learning processes may be the key to effective processes in schools.

Appendix: PIPM questionnaire

Factor 1: storing and retrieving information

I direct the staff to catalog learning materials for the students.

I document meetings with the superintendent, municipal authority, or the parents.

For evaluation purposes, I encourage the use of previous reports dealing with learning and teaching.

I maintain meeting protocols.

I make use of databases of teacher data (evaluation reports, absences, achievements) in decision making.

Factor 2: receiving information from students, parents, the community, and superiors

I meet with community representatives to promote school needs.

I hold meetings with my superiors in order to learn from external events that affect the school.

I examine external evaluation data together with the local authority.

I meet with representatives of the PTA in order to receive feedback about school affairs.

I meet with students in order to get exposed to difficulties, special requests, comments and insights.

I participate in professional development programs for principals.

I meet with parents in order to get exposed to difficulties, special requests, comments and insights.

I examine external evaluation data with the superintendent.

Factor 3: disseminating information to teachers

I disseminate among the staff reports concerning professional changes and innovations.

I help teachers share their special ideas with other teachers in the school.

I disseminate teachers' successes among the staff.

I disseminate the district publications and ministry circulars among the staff.

Factor 4: analyzing external and internal information

I hold meetings with the teaching staff in order to examine information about external events that affect the school.

I examine external data (such as national exams) with the teaching staff.

I initiate and participate in teaching staff meetings in order to evaluate the school climate.

I initiate and participate in teaching staff meetings in order to evaluate student academic achievements.

I observe teachers' lessons.

Source: [Schechter and Qadach \(2016\)](#).

References

- Ash, R.C., Persall, J.M., 2000. The principal as chief learning officer: developing teacher leaders. *NASSP Bull.* 84 (16), 15–22.
- Berson, Y., Da'as, R., Waldman, D.A., 2015. How do leaders and their teams bring about organizational learning and outcomes? *Person. Psychol.* 68 (1), 79–108.
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., Wyckoff, J., 2011. The influence of school administrators on teacher retention decisions. *Am. Educ. Res. J.* 48, 303–333.
- Browne-Ferrigno, T., 2007. Developing school leaders: practitioner growth during an advanced leadership development program for principals and administrator-trained teachers. *J. Res. Leadersh. Educ.* 2 (3), 1–30.
- Brown, K.M., Wynn, S.R., 2009. Finding, supporting, and keeping: the role of the principal in teacher retention issues. *Leader. Pol. Sch.* 8 (1), 37–63.
- Cha, S., Cohen-Vogel, L., 2011. Why they quit: a focused look at teachers who leave for other occupations. *Sch. Effect. Sch. Improv.* 22 (4), 371–392.
- Cohen, A., 2003. *Multiple Commitments in the Workplace: An Integrative Approach*. Mahwah, Erlbaum, NJ.
- Conley, S., You, S., 2009. Teacher role stress, satisfaction, commitment, and intentions to leave: a structural model. *Psychol. Rep.* 105 (3), 771–786.
- Da'as, R., 2019. Teachers' skill flexibility: examining the impact of principals' skills and teachers' participation in decision making during educational reform. *Int. J. Educ. Manag.* 33 (2), 287–299.
- Da'as, R., Schechter, C., Qadach, M., 2019. From principal's cognitive complexity to teacher's intent to leave: exploring the mediating role of school's absorptive capacity and teacher's commitment. *J. Educ. Adm.* 58 (2), 227–245.
- Da'as, R., Qadach, M., 2018. Examining organizational absorptive capacity construct: a validation study in the school context. *Leader. Pol. Sch.* 19 (3), 327–345.

- Da'as, R., Watted, A., Barak, M., 2020. Teacher's withdrawal behavior: examining the impact of principals' innovative behavior and climate of organizational learning. *Int. J. Educ. Manag.* 34 (8), 1339–1355.
- Darling-Hammond, L., LaPointe, M., Meyerson, D., Orr, M., Cohen, C., 2007. *Preparing School Leaders for a Changing World: Lessons from Exemplary Leadership Development Program*. Stanford Educational Leadership Institute, Stanford, CA.
- Davis, S., Darling-Hammond, L., LaPointe, M., Meyerson, D., 2005. *School Leadership Study: Developing Successful Principals (Review of Research)*. Stanford Educational Leadership Institute, Stanford, CA.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Desjardins, F., Donaldson, G., 2008. High school achievement in Maine: where you come from matters more than school size and expenditures. *Me. Policy Rev.* 17 (1), 1–11.
- Drago-Severson, E., 2009. *Leading Adult Learning: Supporting Adult Development in Our Schools*. Corwin/Sage, Thousand Oaks, CA.
- Drago-Severson, E., Pinto, K., 2009. School leadership in support of teachers' learning: the dramatic differences that financial resources make. *Leader. Pol. Sch.* 8, 446–476.
- Drake, T.A., Goldring, E., Grissom, J.A., Cannata, M.A., Neumerski, C., Rubin, M., Schuermann, P., 2016. Development or dismissal? Exploring principals' use of teacher effectiveness data. In: Grissom, J.A., Youngs, P. (Eds.), *Improving Teacher Evaluation Systems: Making the Most of Multiple Measures*. Teachers College Press, New York, NY, pp. 169–183.
- Ellis, S., Shpilberg, N., 2003. Organizational learning mechanisms and managers' perceived uncertainty. *Hum. Relat.* 56 (10), 1233–1254.
- Fauske, J.R., Raybould, R., 2005. Organizational learning theory in schools. *J. Educ. Adm.* 43 (1), 22–40.
- Fochsen, G.K., Sjogren, K., Josephson, M., Lagerstrom, M., 2005. Factors contributing to the decision to leave nursing care: a study among Swedish nursing personnel. *J. Nurs. Manag.* 13 (4), 338–344.
- Fore, C., Martin, C., Bender, W.N., 2002. Teacher burnout in special education: the causes and the recommended solutions. *High Sch. J.* 86 (1), 36–44.
- Fullan, M., 2009. Large-scale reform comes of age. *J. Educ. Change* 10, 101–113.
- Giles, C., Hargreaves, A., 2006. The sustainability of schools as learning organizations and professional learning communities during standard-based reform. *Educ. Adm. Q.* 42 (1), 124–156.
- Grissom, J.A., 2011. Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record* 113, 2552–2585.
- Grissom, J.A., Brendan, B., 2019. Strategic retention: Principal effectiveness and teacher turnover in multiple-measure teacher evaluation systems. *American Educational Research Journal* 56 (2), 514–555.
- Grissom, J.A., Loeb, S., 2011. Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal* 48 (5), 1091–1123.
- Hallinger, P., Wang, W.C., 2015. *Assessing Instructional Leadership with the Principal Instructional Management Rating Scale*. Springer, Dordrecht.
- Halverson, R., Grigg, J., Prichett, R., Thomas, C., 2005. *The New Instructional Leadership: Creating Data-Driven Instructional Systems in Schools*. paper presented at the. In: *Annual Meeting of the National Council of Professors of Educational Administration*. Washington, DC.
- Hambrick, D.C., 2007. Upper echelons theory: an update. *Acad. Manag. Rev.* 32 (2), 334–343.
- Hanisch, K.A., Hulin, C., 1991. General attitudes and organizational withdrawal: an evaluation of causal model. *J. Vocat. Behav.* 39 (1), 110–128.
- Harris, K.J., James, M.L., Boonthanom, R., 2005. Perceptions of organizational politics and cooperation as moderators of the relationship between stress and intent to turnover. *J. Manag. Issues* 17 (1), 26–42.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of the school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement. *Rev. Educ. Res.* 86 (2), 531–569.
- Horng, E., Loeb, S., 2010. New thinking about instructional leadership. *Phi Delta Kappan* 92 (3), 66–69.
- Hughes, A.L., Matt, J.J., O'Reilly, F.L., 2015. Principal support is imperative to the retention of teachers in hard-to-staff schools. *J. Educ. Train. Stud.* 3 (1), 129–134.
- Ingersoll, R.M., 2001. Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal* 38, 499–534.
- Ingersoll, R.M., 2012. Beginning teacher induction: what the data tells us. *Phi Delta Kappa International* 93 (8), 47–51.
- Josephson, M., Lindberg, P., Voss, M., Alfredsson, L., Vingard, E., 2008. The same factors influence job turnover and long spell of sick leave: a 3 year follow-up of Swedish nurses. *Eur. J. Publ. Health* 18 (4), 380–385.
- Johnson, S.M., Kraft, M.A., Papay, J.P., 2012. How context matters in high-needs schools: the effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teach. Coll. Rec.* 114, 1–39.
- Kaplan, S., Bradley, J.C., Lachman, J.N., Hayness, D., 2009. On the role of positive and negative affectivity in job performance: a meta-analytic investigation. *J. Appl. Psychol.* 94 (1), 76–162.
- Kark, R., Shamir, B., Chen, G., 2003. The two faces of transformational leadership: dependence and empowerment. *J. Appl. Psychol.* 88 (2), 243–255.
- Knapp, M.S., Copland, M., 2006. *Connecting Leadership with Learning: A Framework for Reflection, Planning, and Action*. Association for Supervision & Curriculum Development, Alexandria, VA.
- Knapp, M.S., Feldman, S.B., 2012. Managing the intersection of internal and external accountability: challenge for urban school leadership in the US. *J. Educ. Adm.* 50 (5), 666–694.
- Kochan, F.K., Bredeson, P., Riehl, C., 2002. Rethinking the professional development of school leaders. *Yearbk. Natl. Soc. Stud. Educ.* 101, 289–306.
- Kraft, M.A., Marinell, W.H., Yee, D., 2016. School organizational contexts, teacher turnover, and student achievement: evidence from panel data. *Am. Educ. Res. J.* 53, 1411–1449.
- Kruse, S.D., 2003. Remembering as an organizational memory. *J. Educ. Adm.* 41 (4), 332–347.
- Ladd, H.F., 2011. Teachers' perceptions of their working conditions: how predictive of planned and actual teacher movement? *Educ. Eval. Pol. Anal.* 33, 235–261.
- Leithwood, K., Seashore Louis, K., Anderson, S., Wahlstrom, K., 2004. *Review of Research: How Leadership Influences Student Learning*. The Wallace Foundation, New York, NY.
- Leithwood, K., Sun, J., 2012. The nature and effects of transformational school leadership: a meta-analytic review of unpublished research. *Educ. Adm. Q.* 48 (3), 387–423.
- Lipshitz, R., Popper, M., Friedman, V.A., 2002. Multi-faceted model of organizational learning. *J. Appl. Behav. Sci.* 38, 78–98.
- Lopez, V., 2018. *Teachers' Job Satisfaction and Efficacy as Indicators of Intent-to-Leave Teaching*. doctoral thesis. Walden University, Minneapolis, MA. viewed 19 October, 2018. <http://scholarworks.waldenu.edu/dissertations>.
- Louis, K.S., 2006. Changing the culture of schools: professional community, organizational learning, and trust. *J. Sch. Leader.* 16 (5), 477–489.
- Louis, K.S., Dretzke, B., Wahlstrom, K., 2010. How does leadership affect student achievement? Results from a national US survey. *Sch. Effect. Sch. Improv.* 21 (3), 315–336.
- Markow, D., Pieteres, A., 2009. *The MetLife Survey of the American Teacher: Collaborating for Student Success*. MetLife, New York, NY.
- May, H., Supovitz, J.A., 2011. The scope of principal efforts to improve instruction. *Educ. Adm. Q.* 47 (2), 332–352.
- McCann, L., Hughes, C.M., Adair, C.G., 2009. Assessing job satisfaction and stress among pharmacists in Northern Ireland. *Pharm. World Sci.* 32 (2), 188–194.
- McGough, D.J., 2003. Leaders as learners: an inquiry into the formation and transformation of principals' professional perspectives. *Educ. Eval. Pol. Anal.* 25 (4), 449–471.
- Morrell, K., 2004. Towards a typology of nursing turnover: the role of shocks in nurses' decisions to leave. *J. Adv. Nurs.* 49 (3), 315–322.
- Mulford, B., Silins, H., 2011. Revised models and conceptualization of successful school principalship that improve student outcomes. *Int. J. Educ. Manag.* 25 (1), 61–82.
- Murphy, J., 2006. Some thoughts on rethinking the pre-service education of school leaders. *J. Res. Leadersh. Educ.* 1 (1), 1–3.
- Muthén, L.K., Muthén, B.O., 1998. *Mplus User's Guide*, sixth ed. Muthén And Muthén, Los Angeles, CA.
- Orr, M.T., 2006. Mapping innovation in leadership preparation in our nation's schools of education. *Phi Delta Kappan* 87 (7), 492–499.

- Parry, J., 2008. Intention to leave the profession: antecedents and role in nurse turnover. *J. Adv. Nurs.* 64 (2), 157–167.
- Printy, S., 2008. Leadership for teacher learning: a community of practice perspective. *Educ. Adm. Q.* 44 (2), 187–228.
- Qadach, M., Schechter, C., Da'as, R., 2019. Instructional leadership and teacher's intent to leave: the mediating role of collective teachers' efficacy and shared vision. *Educ. Manag. Adm. Leader* 48 (4), 617–634. <https://journals.sagepub.com/doi/abs/10.1177/1741143219836683>.
- Qadach, M., Schechter, C., Da'as, R., 2020. From principals to teachers to students: exploring an integrative model for predicting students' achievements. *Educ. Adm. Q.* 56 (5), 736–778. <https://doi.org/10.1177/0013161X20907133>.
- Rahim, M.A., Afza, M., 1993. Leader power, commitment, satisfaction, compliance and propensity to leave a job among U.S. accountants. *J. Soc. Psychol.* 133 (5), 611–625.
- Richardson, B.K., Alexander, A., Castleberry, T., 2008. Examining teacher turnover in low performing schools: relationships among emotional labor, communication symmetry, and intent to leave. *Commun. Res. Rep.* 25 (1), 10–22.
- Rosenblatt, Z., Shirom, A., 2006. School ethnicity and governance influence on work absence of teachers and school administrators. *Educ. Adm. Q.* 42 (3), 361–384.
- Schaefer, L., Long, J.S., Clandinin, D.J., 2012. Questioning the research on early career teacher attrition and retention. *Alberta J. Educ. Res.* 58 (1), 106–121.
- Schechter, C., 2008. Organizational learning mechanisms: its meaning, measure, and implications for school improvement. *Educ. Adm. Q.* 44 (2), 155–186.
- Schechter, C., Qadach, M., 2016. Principals' learning mechanisms: exploring an emerging construct. *Leader. Pol. Sch.* 15 (2), 141–167. <https://doi.org/10.1080/15700763.2015.1031256>.
- Schlechty, P.C., 2009. *Leading for Learning: How to Transform Schools into Learning Organizations*. Jossey-Bass.
- Seaton, M., Emmett, R.E., Welsh, K., Petrossian, A., 2008. Teaming up for teaching and learning. *Leadership* 37 (3), 26–29.
- Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28 (3), 463–488.
- Shapira-Lishchinsky, O., 2012. Teachers' withdrawal behaviors: integrating theory and findings. *J. Educ. Adm.* 50 (3), 307–326.
- Shapira-Lishchinsky, O., Rosenblatt, Z., 2009. Perceptions of organizational ethics as predictors of work absence: a test of alternative absence measures. *J. Bus. Ethics* 88 (4), 717–734.
- Shaw, J.D., Gupta, N., Delery, J.E., 2005. Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. *Acad. Manag. J.* 48 (1), 50–68.
- Sun, J., Leithwood, K., 2017. Calculating the power of alternative choices by school leaders for improving student achievement. *Sch. Leader. Manag.* 37 (1–2), 80–93.
- Tett, R.P., Meyer, J.P., 1993. Job satisfaction, organizational commitment, turnover intention, and turnover: path analysis based on meta-analytic findings. *Person. Psychol.* 46 (2), 259–293.
- Torrez, A., Kritsonis, W., 2008. *The Recruitment, Retention, and Development of Quality Teachers*. National Education Association viewed 28 June, 2011. www.idonline.org/article/The_Recruitment%2C_Retention%2C_and_Development_of_Quality_Teachers_in_Hard-to-Staff_Schools.
- Waters, T., Grubb, S., 2004. *Leading Schools*. Mid-continent Research for Education and Learning, Aurora, CO.
- Zembylas, M., Papanastasiou, E., 2004. Job satisfaction among school teachers in Cyprus. *J. Educ. Adm.* 42 (8), 357–374.

Successful leadership in high needs schools

Hans W. Klar, Clemson University, Clemson, SC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	186
High needs schools	186
Characterizing high needs schools	187
Resources for high needs schools	187
Effective school leadership	188
Effective leadership for high needs schools	189
Taking a contextually relevant approach	189
Increasing organizational capacity	190
Leading collaborative and continuous school-wide improvement efforts	190
Combining contextual relevance, capacity building, and continuous improvement	191
High needs school leader development	191
Conclusion	192
Acknowledgment	193
References	193
Relevant websites	195

Introduction

In addition to the moral obligation to prepare the youth under their care for success in post-secondary education or careers in increasingly competitive and highly skilled workforces, educational leaders and policymakers around the world face an intensifying demand for increased student academic outcomes. These pressures are exacerbated by factors such as shrinking global markets, the use of standardized tests for national comparisons, the proliferation of alternatives to traditional public schools and a lack of public trust in education. The overwhelming political and economic response to these pressures has been grounded in neoliberalism, the notion that economic and education policy can enable more efficient education systems through open market competition, standardization and high-stakes testing (Butler, 2014). Furthermore, this perspective has engendered an accountability context in which teachers, leaders and policymakers are held responsible for student success (Zepeda et al., 2017).

Meeting the demand for increased academic outcomes is challenging for all educators. Yet, realizing these outcomes is more challenging in some schools than in others. The climate of accountability for higher educational outcomes has led to a focus on schools with low levels of academic performance on standardized tests and other outcomes, such as graduation rates. The conditions that can negatively influence these indicators of success, not surprisingly, are those that undermine teaching and learning practices. These conditions include economic, political, historical, geographic, socio-economic and environmental factors, which can also exacerbate the disparities between schools' abilities to provide equitable educational opportunities for students.

Student academic outcomes are viewed as the collective responsibility of educators at multiple levels in Organization for Economic Cooperation and Development (OECD) countries (Pont, 2020). However, given the vast body of research on the influence of school leaders on student learning outcomes and the tendency to hold formally recognized leaders responsible for these outcomes, for the purpose of this article, I focus on the role of the singular school leader. Yet, I do so with the understanding that successful leaders enlist and enable other leaders in meeting the learning needs of their students. Furthermore, I use the term school leader throughout the article rather than principal as the term for the sole, formally recognized leader of a school varies in different regions of the world and can include headmaster, head teacher and director among others (see Pont, 2020).

Viewing schools where educators encounter greater than usual challenges meeting the social, emotional and academic needs of their students as high needs schools, and considering the key role of school leaders in meeting these challenges, in the following sections of this article I describe the characteristics of high needs schools, effective school leadership practices and leadership practices that are particularly effective in high needs schools. In describing the effective practices of leaders in high needs schools, I focus on three integrated and mutually reinforcing strategies: taking a contextually relevant approach, increasing organizational capacity and leading collaborative and continuous school-wide improvement efforts. Finally, given the limited attention to the development of leaders for high needs schools (Lochmiller and Chesnut, 2016), I conclude the article by discussing strategies for preparing and supporting leaders for such schools.

High needs schools

The phrase high needs is usually attributed to those schools in which educators find it more challenging to realize high levels of academic outcomes for their students. Indeed, in the previous edition of this encyclopedia, Harris (2010) referred to such schools

as “challenging” (p. 778). The phrase high needs schools often conjures up images of inner city or rural schools situated in economically depressed locations. Some organizations in the United States, in fact, use poverty to classify schools as high needs. This classification is typically based on the percentage of students eligible to receive free or reduced-priced meals at school under the National School Lunch Program (NLSP). For instance, the Teacher Quality Act (2002) defined high needs elementary schools as those in which 60% of students are eligible for free or reduced-price meals. Use of this proxy for poverty levels is consistent with a report by the [Mid-Continent Research for Education and Learning \(2005\)](#), which defined high needs elementary schools as those in which at least 50% of students were eligible for free or reduced-priced meals. According to the National Center for Educational Statistics (NCES), in 2016 at least 50% of students were eligible for FRPL in at least 50% of public schools in the United States ([NCES, 2018](#)).

Characterizing high needs schools

Though poverty is a factor in a large percentage of public schools in the United States, racial and ethnic groups are not equally distributed among high poverty schools. Schools that serve higher percentages of students from low income households also tend to have higher concentrations of students who are non-white and do not speak English as a first language ([Musu-Gillette et al., 2017](#); [NCES, 2018](#)). Given the well-documented and negative influences of poverty on student learning outcomes ([Berliner, 2009](#); [Lacour and Tissington, 2011](#)), the disparities in educational funding ([EdBuild, 2019](#)) and the persistent achievement gap between white and non-white students ([Musu-Gillette et al., 2017](#)), it is imperative that leaders in high needs schools understand the historical, political, social and cultural norms that create and maintain these conditions ([Berry et al., 2019](#)). In addition, these leaders need to meet the demands of accountability policies while developing a collective vision that transcends the immediate teaching and learning challenges to focus instead on marshalling a sustained commitment to meeting the learning needs of all students.

Despite poverty often serving as the defining characteristic of high needs schools, other characteristics also warrant consideration. These aspects can include those that relate specifically to physical, financial, political, human and environmental factors. ([Baran and Berry, 2015](#); [Murakami et al., 2019](#)). These factors can be seen in many rural schools, with their smaller population centers and long distances from developed areas. Remoteness, an aging population ([Argent, 2016](#)) and high levels of poverty ([Showalter et al., 2019](#)) can lead to fewer resources being generated by local tax bases to support school initiatives. This lack of resources can in turn hinder a school leader’s ability to fund adequate instructional programs and recruit effective teachers ([Wieczorek and Manard, 2018](#)).

Recruiting and retaining effective teachers has been identified as a particular challenge for rural schools in multiple studies. A study of the teaching force in the United States over the last three decades ([Ingersoll et al., 2018](#)), a study of rural Australian schools ([Downes and Roberts, 2018](#)) and a study of rural schools in Australia, Canada and New Zealand ([Sullivan et al., 2018](#)) all indicated that rural schools had the highest teacher turnover rates. The higher level of teacher turnover in rural schools is further exacerbated by higher levels of school leader turnover ([Hansen, 2018](#); [Pendola and Fuller, 2018](#)). Furthermore, the higher rate of teacher and leader turnover in rural and high poverty schools negatively impacts student learning outcomes ([Bêteille et al., 2012](#)) and leads to greater inequities between rural and non-rural students ([Sullivan et al., 2018](#)).

Additionally, the messages students in high needs schools receive about themselves, their schools and their generally lower levels of academic achievement can have a negative influence on their social and emotional development. Over time, these messages can become internalized and detrimentally impact students’ educational aspirations ([Byun et al., 2012](#)). Thus, leaders of high needs schools need to attend to students’ social and emotional needs as well as their academic ones.

Resources for high needs schools

Educational leaders and policymakers around the world are tasked with meeting the demand for increased academic outcomes in schools. Increasingly common approaches to meeting this demand being adopted in Australia, Canada, the US and some European countries are grounded in the neoliberal philosophies that emanated from the UK and the US in the 1980s ([Sahlberg, 2016](#)). Though some countries in Europe, East Asia and elsewhere have resisted such approaches to educational reform, the trend toward competition, standardization and high-stakes testing is expanding across the globe ([Sahlberg, 2016](#)). Policymakers in the United States have attempted to redress the conditions that contribute to lower levels of academic outcomes in high needs schools with a combination of accountability demands, resource allocations and, more recently, greater flexibility for how resources are employed. This balance of accountability, resources and flexibility can be seen in the [Every Student Succeeds Act \(ESSA\) \(2015\)](#), the most recent re-authorization of the [Elementary and Secondary Education Act \(ESEA\) \(1965\)](#). This balance is most apparent in ESSA’s aims “to require states to align their education programs with college and career ready standards and to extend the federal focus on equity by providing resources for poor students, students of color, English language learners, and students with disabilities” ([Young et al., 2017](#), p. 706).

The most prominent avenue for distributing federal money to local education agencies is through Title I, Part A of ESSA. Title I money, as it is commonly referred, is intended to fund schoolwide programs and supplemental instruction that can aid in raising the academic outcomes of low achieving students in schools where at least 40% of students are from low-income families. In the 2015–16 school year, more than 26 million students studied in nearly 56,000 Title I schools ([U.S. Department of Education, 2018](#)).

Under ESSA, states are required to submit plans for meeting the challenges of high needs schools. In the US state of South Carolina, the state department of education created a school improvement framework that includes three programs focused on

improving the educational outcomes of high poverty and low performing schools (South Carolina Department of Education, 2017). Through these programs, schools receive differentiated levels of comprehensive and targeted grant money and other supports to facilitate school improvement efforts. Another example of state level support for schools in the state of South Carolina is the Rural Recruitment Initiative (Center for Educator Recruitment, Retention and Advancement (CERRA), 2017). Through this incentive program, administered by CERRA, resources are provided to rural and underserved districts with an annual teacher turnover rate of more than 11%. The resources provided by the initiative include incentives such as financial assistance or supplements to attract and retain teachers.

In addition to resources provided by state and federal authorities to assist in meeting academic outcomes in high needs schools, school leaders can also take advantage of resources within their local communities. Bauch (2001) viewed businesses, parent groups and religious and other community-based organizations as potential assets with which schools can form partnerships. Through such partnerships, Bauch suggested, students can learn more about the historical, geographic and cultural assets in their communities. Bauch also stated that these interactions between students and community members can enhance mutual interests and levels of commitment. Similarly, Green described how school leaders can use geographic information systems (GIS) technology (2015) and community-based equity audits (2017) to identify and map community assets in what are often considered low-opportunity communities.

As described in this section, high needs schools are typically characterized on the basis of poverty levels and are associated with high concentrations of minoritized students and low levels of academic achievement. However, for the purpose of understanding how school leaders can meet the challenges of teaching and learning in these schools, I have chosen to more broadly conceptualize high needs schools as those in which teachers and leaders encounter greater than usual challenges meeting the social, emotional and academic needs of their students. I have chosen this approach to avoid a focus on deficits which often position students and their communities as the source of the problem. Rather, in the following sections, I focus on how school leaders can marshal the resources and will to provide students in high needs schools with equitable learning opportunities.

Effective school leadership

Successful leadership in high needs schools is enacted by leaders who create and sustain teaching and learning conditions that meet the needs of their students and satisfy the demands of their accountability contexts. Though the quality of classroom instruction is the most significant school level factor in realizing academic outcomes, school leadership practices comprise the second most significant factor (Louis et al., 2010). Indeed, that leadership is a key factor in successful schools is widely accepted as can be seen in education policy in the US (Herman et al., 2017; Young et al., 2017) and other OECD countries (Pont, 2020). School leaders' significant if indirect effects on student learning outcomes are primarily associated with their roles in exerting leadership that fosters and sustains teaching and learning environments (Louis et al., 2010; Zepeda et al., 2017).

The research basis for this focus on school leadership is often attributed to the effective schools movement (Edmonds, 1979; Levine and Lezotte, 1990). The studies conducted as part of this effort were largely based in high-poverty urban schools where students were performing at higher than expected levels. Seeking to discover the correlates of success in these schools, scholars identified a number of leadership practices, including establishing a clear mission, high expectations, a safe environment and a strong focus on instruction (Edmonds, 1979; Levine and Lezotte, 1990). This focus on instruction was further buoyed by the work of Hallinger and Murphy (1986) who identified three primary tasks of instructional leadership: define the mission, manage the instructional program and promote the school climate.

Since this period, both instructional leadership, with its focus on improving classroom instruction (Hallinger, 2005), and transformational leadership, with its broader focus on creating a positive school environment (Leithwood and Jantzi, 2005), have been widely studied as part of an effort to inform and enhance school leadership practice. After completing a meta-analysis of effective school and leadership studies conducted over three decades, Leithwood and Riehl (2005) identified a combination of core instructional and transformational leadership practices they described as necessary but not sufficient for school effectiveness. These practices-setting directions, developing people, redesigning the organization and managing the instructional program-have been widely cited and have served as the building block of numerous educational leadership studies.

In a more recent study on the influence of school leadership on student learning outcomes, Leithwood et al. (2019) found that school leaders can enact the core leadership practices in four different paths, or categories of mediators, to influence student learning outcomes. The four paths include the rational, emotional, organizational and family path. The authors found that school leaders could most significantly influence student learning outcomes by improving variables on the rational path, which is "rooted in the knowledge and skills of school staff members about curriculum, teaching, and learning" (p. 5) and includes classroom instruction, academic press, disciplinary climate and teachers' use of instructional time. The findings from this study not only reinforced the benefits of combining instructional and transformational leadership practices but also emphasized the critical nature of those practices related to teaching and learning and their influence on student learning outcomes.

Research in the area of school leader effectiveness has also emphasized the need to adapt leadership practices to suit immediate contexts (Hallinger, 2011; Leithwood et al., 2008; Louis et al., 2010). As Leithwood et al. (2019) noted in the conclusion of their study, school leaders have to make choices regarding which strategies to use when and how with little guidance from the research field. Given the dynamic challenges of high needs schools, the ability to enact practice in contextually relevant ways is essential (Klar and Huggins, 2020).

Effective leadership for high needs schools

The findings that emerged from the effective schools movement and many of the subsequent studies on effective school leadership were in fact generated in high needs schools. As such, it is important to acknowledge that what is today widely recognized as effective leadership is not only applicable to most schools but is in fact largely based on what is known about effective leadership in high needs schools. Building on this sound research basis, in the following section, I examine more closely three leadership practices that have been shown to be particularly effective in high needs schools: taking a contextually relevant approach, building organizational capacity and leading collaborative and continuous improvement efforts.

Taking a contextually relevant approach

As conceptualized in this article, educators in high needs schools face greater than usual challenges meeting the social, emotional and academic needs of their students. Yet, the needs of each high needs school are unique and require a unique set of strategies. That context influences teaching and learning conditions, and thus leadership practice, is widely recognized in the literature on effective school leadership. Some of the contextual factors that enable and constrain leadership practice include the community, cultural, economic, organizational and political contexts (Bredeson et al., 2011; Hallinger, 2018; Klar and Brewer, 2013). Also widely recognized, as noted previously, is the need to adapt leadership practices to suit these contextual circumstances (Clarke and O'Donoghue, 2017; Gurr and Drysdale, 2018; Hallinger, 2018; Okilwa and Barnett, 2018; Pashiardis et al., 2018). In a comparative study of successful school superintendents in Sweden and the United States, Bredeson et al. (2011) found that these leaders drew on a range of knowledge, skills and dispositions rather than relying on pre-determined leadership styles to lead in ways that were context-responsive, or contextually relevant (Klar and Huggins, 2020). Additionally, instead of merely addressing their unique contextual challenges reactively, these leaders also pushed back and shaped their contexts in ways that positively influenced the teaching and learning conditions in their schools.

The findings from these studies on district-level leaders' contextually relevant practices were echoed in a review of research studies on school leaders conducted as part of the International School Leadership Development Network (ISLDN). The ISLDN is a joint venture between the British Educational Leadership, Management and Administration Society and the University Council for Educational Administration (Baran and Berry, 2015). ISLDN members from over 20 countries have conducted international comparisons of effective leaders in high needs schools for nearly 10 years. As noted in Gur and Drysdale's (2018) review of this research, successful school leaders are not victims of their contextual circumstances, rather, as Bredeson et al. (2011) reported, they are able to push back and shape the influences of context on their schools.

A contextually relevant leadership approach requires a high level of understanding of the contextual factors positively and negatively influencing the educational environment. Bredeson and Klar (2008) described this understanding as contextual literacy, or the awareness of key aspects of a particular context, how those aspects influence and are influenced by each other and how they can be leveraged to achieve organizational goals. To better understand how leaders develop their contextual literacy, Bredeson and Klar developed a framework that includes five intersecting sources of their knowledge, skills and dispositions. This framework (people, personalization, place, professional knowledge and purpose) has since been used to study effective leaders in high needs rural schools (Klar et al., 2019b; Klar and Huggins, 2020).

In these studies, leaders demonstrated the need to work with and through people, develop trusting relationships and meet others' personal and professional needs while holding them to high expectations. The leaders personalized their roles by drawing on their own sets of unique skills, values, experiences and perspectives to initiate and model change. They displayed a sound understanding of the place and time in which their schools were situated, including the historical, geographic and economic influences and cultural norms of their settings. The leaders utilized their professional knowledge about teaching, and the resources, policies and practices that supported or limited the creation of effective learning environments to support and sustain change. Lastly, they were guided by a strong sense of purpose which positioned students at the center of decision making, was widely communicated and confronted deficit thinking wherever it occurred. Though this framework provides some insight into how school and district leaders develop the contextual literacy to lead in contextually relevant ways, as noted by Leithwood et al. (2019), more research is required to determine when and how school leaders focus their improvement efforts.

Though understanding and adapting leadership practices to suit a school's unique context is critical to successful student learning outcomes in all schools, it is particularly important in high needs schools. This distinction is due to the very challenges that characterize schools as being high needs. That is, the greater-than-average challenges of realizing high levels of student learning outcomes that accountability systems demand of all schools require unique approaches. Yet, the creativity, risk-taking and personalizing of the learning environment required to meet the needs of students in an environment characterized by higher levels of poverty and other political, socio-economic and historical stressors are challenging to develop (Berry et al., 2019), especially when educators are incentivized to narrow the curriculum to align with standardized accountability measures (Sahlberg, 2016). Thus, understanding how contextual factors positively and negatively influence their schools and how they can reshape the factors to ensure positive teaching and learning environments while also meeting their accountability demands requires leaders of high needs schools to possess a greater-than average focus on contextually relevant leadership practices.

Increasing organizational capacity

Despite the pressures for immediate changes in academic outcomes, successful leaders in high needs schools resist the temptation to grasp at quick fixes. Instead, they work to build the capacity of their schools to support and sustain appropriate teaching and learning activities. Given the indirect manner in which school leaders are able to enhance student learning outcomes through their support of teaching and learning activities (Leithwood et al., 2019; Robinson et al., 2008), it is essential that leaders in high needs schools focus on developing and sustaining effective teaching practices as a high priority and with a relentless focus (Harris, 2010).

The process of building capacity within any school involves increasing the motivation, resilience and other conditions necessary to sustain continuous learning efforts (Stoll and Bolam, 2005). The multiple aspects of building capacity can be seen in the core instructional and transformational leadership practices identified by Leithwood and Riehl (2005) and in Hallinger's (2011) review of decades of research on effective school leadership, from which he derived a model of leadership for learning. Like Leithwood and Riehl (2005), Hallinger (2011) recognized that leaders contribute to learning through multiple dimensions, which include having a vision for the future, developing processes that enhance academic outcomes and fostering the development of teachers and other leaders to assist in creating and maintaining effective policies and practices.

However, as Cohen et al. (2007) and Cohen et al. (2017) noted, there is a dilemma regarding education policies and practice. This dilemma is that policymakers rely on educators to solve problems, but the educators who have the problem are relying on the policymakers for solutions. Furthermore, Cohen et al. (2007) suggested that a school's level of extant capacity is relative to the amount of change required to meet the aims of a particular initiative. As this dilemma suggests, universal approaches to capacity building are unlikely to be successful. Rather, school leaders must work to draw upon and develop their contextual understanding to determine when and how to enact capacity building strategies in accordance with the scope of change required.

In addition, the work of capacity building requires a systemic focus on developing organizational learning rather than a singular focus on individuals or programs. Organizational learning is the school's collective ability to meet its teaching and learning challenges (Mulford and Silins, 2003) and provides an avenue to improving classroom instruction (Spillane and Louis, 2005). In high needs schools with high levels of organizational learning, members of the school community are able to assess and respond to challenges in ways that can lead to sustainable improvement. In their review of effective school leadership in rural high needs schools, Klar and Huggins (2020) developed a framework for understanding organizational learning. Building on the work of Mitchell and Sackney (2011) and Gurr and Drysdale (2007), the framework suggests that school leaders can foster organizational learning through a combination of developing the individual capacities of teachers and other staff, the abilities of people to collaborate, the organizational structures for these interactions to occur and the school's connections with families and other community based resources that could benefit students.

Building organizational capacity to respond to teaching and learning challenges is a key approach for any school leader. However, building the capacity of a school to meet these challenges is more difficult in high needs schools. One particularly important factor, given the importance of classroom instruction, is the higher level of teacher turnover in schools situated in low-income rural and urban communities (Ingersoll et al., 2018). This trend results in students in high needs schools receiving instruction from greater numbers of inexperienced teachers and illustrates the critical need for leaders of such schools to focus on building instructional capacity. Furthermore, given that the poorer working conditions often found in high needs schools are among the most common reasons teachers provide for leaving the profession (Ingersoll et al., 2018), leaders in high needs schools need to work to enhance the teaching and learning conditions. This can be done in part by developing or modifying organizational structures to create collaborative cultures (Louis et al., 2010), enable greater levels of individual and group learning (Mitchell and Sackney, 2011) and increase social interaction among teachers (Moolenaar, 2012).

Leading collaborative and continuous school-wide improvement efforts

In the United States, there has been a particularly intense focus on improving perennially low performing schools over the last two decades. Beginning with the *No Child Left Behind Act* (2001) and continuing through other initiatives, such as *Race to the Top* (2009), this focus has been on turning around low performing schools (Duke, 2012, 2014). Like its predecessors, ESSA has kept a focus on improving low performing schools. Though ESSA has provided more flexibility in achieving these aims, it also requires states and school districts to provide evidence of effectiveness through research-based interventions and school improvement planning processes (Sparks, 2018).

This focus on planning and data use to monitor progress in student learning outcomes runs counter to the urge to adopt quick fixes in the belief that long-struggling schools can be quickly improved simply by adopting new programs (Rohanna, 2017). Instead, there has been an increase in systemic and sustainable change efforts informed by the use of continuous school improvement processes (Sparks, 2018). The term continuous improvement describes the process of addressing problems of practice by using evidence and systematic approaches to improve systemic practices (Park et al., 2013; Sparks, 2018; Tichnor-Wagner et al., 2017).

Continuous improvement is closely aligned with improvement science, a disciplined approach to improving practice through inquiry (Bryk et al., 2015). Such inquiries often take the form of small-scale changes using models such as *Plan Do Study Act* (Langley et al., 2009; Park et al., 2013) and action research (Logan, 2014). In both cases, individuals or teams may investigate a problem of practice using disciplined inquiry, local practitioner knowledge and research to implement and study the effects of interventions. Based on the results of the intervention, the cycle may be continued, potentially resulting in deeper and broader levels of change. Bryk et al. (2015) described this approach to school improvement as accelerating "learning by doing" (p. 8).

Adopting an improvement science approach to school improvement allows leaders of high needs schools in particular to effectively and rapidly enhance their schools' individual and organizational capacities to address the misalignments within their systems. Studies of leaders in high needs schools implementing action research (Andreoli et al., 2019; Klar and Huggins, 2020) and PDSA cycles (Andreoli and Klar, 2020) highlight the benefits continuous improvement can bring to school improvement processes. These benefits include a deeper understanding of the assets and challenges related to problems, a more systematic and reflective approach to identifying and addressing challenges and an increase in the involvement, commitment and leadership of other teachers and administrators in meeting schools' teaching and learning needs.

In a recent study of a high needs, rural elementary school, Andreoli and Klar (2020) found that the principal and leadership team members collaboratively completing a continuous improvement inquiry cycle significantly enhanced their ability to learn by doing (Bryk et al., 2015). In this case, the school leaders use of the school improvement process led to more effectively utilizing data and prioritizing improvement strategies, incorporating multiple perspectives into problem-solving efforts and establishing a culture of risk-taking and ownership of teaching and learning practices. Combined, these strategies allowed these leaders to become the drivers of their own change processes and resulted in them significantly improving their ability to meet the needs of their students.

Combining contextual relevance, capacity building, and continuous improvement

Though described separately, the three practices explained in this section are both interconnected and mutually reinforcing. Combined, they can result in the development of schools and communities that are committed to and capable of meeting the social, emotional and academic needs of their students. This combination of approaches highlights the benefits of a systems perspective, which includes seeing both the parts and the whole rather than focusing on isolated components of systems (Shaked and Schechter, 2014).

As conceptualized in this article, educators in high needs schools have greater than usual challenges meeting the learning needs of their students. Effective leaders in high needs schools draw on their contextual awareness to focus their capacity building and continuous improvement efforts in ways that align with their unique contexts. Furthermore, as meeting the demands of accountability in organizations as dynamic as high needs schools requires the contributions of all stakeholders, a key benefit of a contextually-relevant approach is the opportunity to recognize and develop the capacity of other teachers, leaders and community members to participate in and enhance continuous improvement efforts. When leaders engage in continuous improvement processes, they also increase their contextual literacy. Andreoli et al. (2019) reported that 10 school leaders in a three-year leadership development initiative who learned to lead change through action research interacted closely with other school members and shared ownership of the school improvement process. As a result, other members of the school community took on greater levels of leadership responsibility and began to approach school improvement processes with a deeper and more reflective and systematic perspective.

High needs school leader development

High needs schools require leaders who are able to identify and challenge the conditions limiting the success of their students while simultaneously meeting the demands placed on them by the increased accountability climate. Furthermore, these leaders need to be able to draw on a variety of strategies to develop the capacity of their schools to lead sustainable, systematic and systemic change. In this regard, leaders of high needs schools face additional challenges that require ongoing support (Duke, 2014; Lochmiller and Chesnut, 2016). In some high needs schools, such as ones in rural areas, these challenges can result in school leaders taking on a larger number of duties than leaders in schools without high needs (Hansen, 2018; Preston et al., 2013; Wieczorek and Manard, 2018). As a result of these additional burdens, leaders in high needs schools can have additional stress levels (Starr and White, 2008; Wood et al., 2013) and higher rates of turnover (Béteille et al., 2012).

Given the influence school leaders can have on students' academic outcomes (Day et al., 2016; Heck and Hallinger 2014; Louis et al., 2010), the detrimental effects of school leader turnover (Béteille et al., 2012) and the potential of professional development to enhance school leader retention (Goldring et al., 2014), there has been an increased focus on developing school leaders to address the challenges of improving the teaching and learning conditions in high needs schools (Herman et al., 2017; Manna, 2015). In the United States, this emphasis can be seen in ESSA (2015), which allows states the flexibility to direct their allocated funds to support leadership development efforts. In a review of states' ESSA accountability plans, which included the District of Columbia and Puerto Rico, Reedy and Doiron (2018) found that 44 of 52 plans described a focus on leadership development activities rather than simply enhancing compensation or advancement opportunities. This shift to structured leader development is noteworthy, since more focus has traditionally been centered on teacher development, and leadership development activities are rarely initiated by leaders themselves (Zepeda et al., 2014). Studies of successful districts in Canada (Leithwood and Azah, 2017) and the United States (Zepeda et al., 2014) have demonstrated that school leader professional development is a factor in school improvement success. Other studies have illustrated how large urban districts can provide collaborative professional learning experiences for school leaders focused on instruction (Louis et al., 2010; Honig, 2012; Honig and Rainey, 2014).

Leading high needs schools in contextually relevant ways that build capacity and enable continuous improvement efforts requires learning opportunities embedded in the context in which the leading and learning is to take place. This type of personalized learning opportunity can be provided through leadership coaching. Though coaching has been commonly used as a strategy for

supporting classroom teachers for many years, it is increasingly being used to support new and experienced school leaders (Louis et al., 2010; Reiss, 2015; Riley and Meredith, 2017).

As opposed to mentoring, coaching leaders involves those with expertise in school leadership working to develop leaders' knowledge and skills to address targeted areas for improvement (Goff et al., 2014; Grissom and Harrington, 2010). Though coaches may use a combination of blended strategies to assist leaders' development (Bloom et al., 2005; Reiss, 2015), the primary coaching approach is the use of questioning to engage leaders in critical self-reflection (Robertson, 2016). In this way, coaching is more personalized and effective than traditional one-size-fits-all forms of professional development (Lochmiller, 2014; Reiss, 2015).

Though leadership coaching typically occurs in the form of coach-leader pairings, the benefits of coaching can be enhanced when leaders are also able to participate in professional communities of their peers (Robertson, 2016). Like coaching, professional communities, have typically been seen as a vehicle for teacher development. However, they are more recently being considered valuable learning opportunities for school leaders. Professional communities are those in which participants have established norms of shared values, reflective dialog, de-privatized practice, a focus on student learning, and collaboration (Kruse et al., 1995). When school leaders participate in professional communities, they are able to learn with and from their colleagues and receive just-in-time advice from trusted peers.

While examples of leadership coaching and professional communities can be found in some larger, better resourced districts, smaller and less well-resourced districts often do not have the capacity to provide leaders with these forms of professional learning (Lock et al., 2012; Preston et al., 2013). The inability to provide their leaders with adequate professional learning opportunities to meet the challenges of their high needs schools has led to an increasing number of districts in the United States forming research practice partnerships with universities and other organizations (Coburn and Penuel, 2016; Coburn et al., 2013). Through research practice partnerships, schools are able to develop and implement research-based strategies to address their problems of practice while also serving as sites for furthering research on the topic.

Though still relatively limited in number, some districts have been able to form successful research practice partnerships to provide leadership coaching and professional communities for pre-service and in-service leaders of high needs schools (Cosner et al., 2018; Fusarelli et al., 2018; Klar and Huggins, 2020). In the case of Klar and Huggins (2020), faculty and graduate students from two universities and a consortium of 12 predominantly rural and high-poverty school districts formed a three-year research practice partnership called the Leadership Learning Community (LLC). Through the LLC, 10 school leaders (principals, assistant principals and teacher leaders) and eight leadership coaches (experienced principals and district leaders) attended monthly professional development sessions facilitated by the university-based faculty members. The content of the sessions related to leading collaborative and continuous change through action research cycles. In addition, the leadership coaches received professional development related to coaching strategies from the facilitators and met with their coaching peers in a professional community monthly (see Huggins et al., 2020).

The LLC leaders received coaching twice a month from their leadership coaches. These coaching sessions took place both in the leaders' schools and at off-site locations (see Andreoli et al., 2019; Klar et al., 2019a). Importantly, many of the leaders' districts did not have experienced school leaders who could serve as their leadership coaches. Instead, the leaders were paired with coaches from other districts. This aspect of the LLC highlights the challenges that some districts have supporting leadership development. Though it also created some logistical challenges for the coaches, leaders described a clear preference for having coaches from outside their districts who they could share their leadership challenges with in a non-evaluative environment.

As a result of participating in the LLC, leaders increased their abilities to lead change in their school, gained a greater understanding of how to engage other school-based members in improvement efforts and enabled distributed leadership practices that increased their schools' capacities to respond to their students' needs. The coaches also increased their understandings of how to use systemic and systematic change processes. Furthermore, they were able to use their newly developed coaching perspectives to work with other leaders in their own districts. In this way, this initiative's focus on leadership development and school improvement not only impacted the schools of leaders who participated in it but also other schools and districts in the consortium.

Though the LLC is merely one example of a research practice partnership, it illustrates the potential affordances of high needs districts seeking new forms of collaboration to address their leadership development challenges. It also adds to the growing research base on the role of school leader supervisors (Thessin, 2019). Importantly, the LLC demonstrated that leaders in high needs schools and districts can be supported in ways that increase their abilities to lead in contextually relevant ways that focus on building their school's capacities for leading systematic and sustainable change geared toward enhancing their teaching and learning environments.

Conclusion

Leaders of all schools have a moral obligation to meet the social, emotional and academic needs of their students in the process of preparing them for future success. Leaders of all schools must also meet the demands for accountability imposed on them by policy-makers and the general public. While all leaders may find achieving this balance of student interests and accountability demands challenging, leaders in high needs schools experience greater challenges in doing so than leaders in less challenging schools. Indeed, in the previous edition of this encyclopedia, Harris (2010) asked rhetorically whether improvement in such schools is even possible. Harris went on to say that, though challenging, both increased student outcomes and sustainable improvement are possible. Yet, to be successful "requires a school improvement process that fits the individual school context and is underpinned by an unrelenting focus on improving conditions at the learning level" (p. 781).

Ten years on, it remains the case that school leaders can successfully respond to the combination of student needs and accountability demands by engaging in contextually-relevant leadership practices, enhancing organizational learning through capacity building efforts and leading collaborative and continuous improvement processes that focus on enhancing teaching and learning practices within schools. However, leaders in high needs schools require additional levels of preparation and support to successfully meet the various challenges presented by their school contexts. In meeting these challenges school leaders benefit from personalized and job embedded learning opportunities geared toward addressing their own problems of practice, such as receiving leadership coaching and joining professional communities. To provide greater access to these types of professional learning in the United States, an increasing number of high needs school districts are forming research practice partnerships with universities and other organizations. Given the critical role the leaders of high needs school can play in enhancing the academic and career opportunities for the millions of children studying in them, it is essential that more research and funding be dedicated to this process.

Acknowledgment

I would like to thank Parker M. Andreoli for her valuable assistance in developing this manuscript.

References

- Andreoli, P.M., Klar, H.W., 2020. Becoming the drivers of change: continuous improvement in a rural research-practice partnership. *J. Educ. Adm.* 59 (2), 161–176.
- Andreoli, P.M., Klar, H.W., Huggins, K.S., Buskey, F.C., 2019. Learning to lead school improvement: an analysis of rural school leadership development. *J. Educ. Change* 21, 517–542.
- Argent, N., 2016. Demographic change: beyond the urban-rural divide. In: Shucksmith, M., Brown, D.L. (Eds.), *Routledge International Handbook of Rural Studies*. Routledge, London, pp. 29–35.
- Baran, M.L., Berry, J.R., 2015. The International School Leadership Development Network (ISLDN) High Needs Schools Group Research Protocol and Members' Guide. British Educational Leadership, Management, and Administration and the University Council for Educational Administration.
- Bauch, P.A., 2001. School-community partnerships in rural schools: leadership, renewal, and a sense of place. *Peabody J. Educ.* 76 (2), 204–221.
- Berliner, D.C., 2009. Poverty and Potential: Out-of-School Factors and School Success. Education and the Public Interest Center & Education Policy Research Unit, Boulder, CO and Tempe, AZ.
- Berry, R.J., Moss, S.C., Gore, P., 2019. Leadership in high-needs/high performing schools: success stories from an urban school district. In: Murakami, E., Notman, R., Gurr, D. (Eds.), *Educational Leadership, Culture and Success in High-Need Schools*. Information Age Publishing, Charlotte, NC, pp. 131–148.
- Béteille, T., Kalogrides, D., Loeb, S., 2012. Stepping stones: principal career paths and school outcomes. *Soc. Sci. Res.* 41, 904–919.
- Bloom, G., Danilovich, D.L., Fogel, J., 2005. Passing the baton: a new program from ACSA and the new teacher center at UC Santa Cruz is improving the way a new generation of site leaders is prepared and supported. *Leadership* 35 (1), 30.
- Bredeson, P.V., Klar, H.W., 2008. Context matters: lessons learned from successful superintendents about preparation, practice, and professional development. In: Papa, R. (Ed.), *Leadership on the Frontlines: Changes in Preparation and Practice*, Yearbook of the National Council of Professors of Educational Administration. ProActive Publications, Lancaster, PA, pp. 274–282.
- Bredeson, P.V., Klar, H.W., Johansson, O., 2011. Context-responsive leadership: Examining superintendent leadership in context. *Educ. Pol. Anal. Arch.* 19.
- Bryk, A.S., Gomez, L.M., Grunow, A., LeMahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press, Cambridge, MA.
- Butler, T.A., 2014. School leadership in the 21st century: leading in the age of reform. *Peabody J. Educ.* 89 (5), 593–602.
- Byun, S., Meece, J.L., Irvin, M.J., Hutchins, B.C., 2012. The role of social capital in educational aspirations of rural youth. *Rural Sociol.* 77 (3), 355–379.
- Clarke, S., O'Donoghue, T., 2017. Educational leadership and context: a rendering of an inseparable relationship. *Br. J. Educ. Stud.* 65, 167–182.
- Coburn, C.E., Penuel, W.R., 2016. Research-practice partnerships in education: outcomes, dynamics, and open questions. *Educ. Res.* 45 (1), 48–54.
- Coburn, C.E., Penuel, W.R., Geil, K.E., 2013. *Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts*. William T. Grant Foundation, New York, NY.
- Cohen, D.K., Moffitt, S.L., Goldin, S., 2007. Policy and practice: the dilemma. *Am. J. Educ.* 113 (4), 515–548.
- Cohen, D.K., Spillane, J.P., Peurach, D.J., 2017. The dilemmas of educational reform. *Educ. Res.* 47 (3), 204–212.
- Cosner, S., Walker, L., Swanson, J., Hebert, M., Whalen, S.P., 2018. Examining the architecture of leadership coaching. *J. Educ. Adm.* 56 (3), 364–380.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52, 221–258.
- Downes, N., Roberts, P., 2018. Revisiting the schoolhouse: a literature review on staffing rural, remote and isolated schools in Australia 2004–2016. *Aust. Int. J. Rural Educ.* 28 (1), 31–54.
- Duke, D.L., 2012. Tinkering and turnarounds: understanding the contemporary campaign to improve low-performing schools. *J. Educ. Stud. Placed Risk* 17 (1–2), 9–24.
- Duke, D.L., 2014. A bold approach to developing leaders for low-performing schools. *Manag. Educ.* 28 (3), 80–85.
- EdBuild, 2019. 23 Billion. Jersey City, NJ.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader* 37, 15–24.
- Elementary and Secondary Education Act of 1965, 1965. H.R. 2362, 89th Cong. Pub. L. No. 89-10.
- Every Student Succeeds Act, 2015. Every Student Succeeds Act. Pub. L. No. 114-195.
- Fusarelli, B.C., Fusarelli, L.D., Wirt, L.G., 2018. Developing and sustaining school- university-community collaborative partnerships to develop school leaders in rural areas. In: Reardon, R.M., Leonard, J. (Eds.), *Innovation and Implementation in Rural Places: School-University-Community Collaboration in Education*. Information Age Publishing, Charlotte, NC, pp. 3–28.
- Goff, P., Guthrie, E., Goldring, E., Bickman, L., 2014. Changing principals' leadership through feedback and coaching. *J. Educ. Adm.* 52 (5), 682–704.
- Goldring, R., Taie, S., Riddles, M., 2014. Teacher Attrition and Mobility: Results From the 2012–13 Teacher Follow-Up Survey. National Center for Education Statistics, Washington, DC.
- Green, T.L., 2015. Places of inequality, places of possibility: mapping "opportunity in geography" across urban school-communities. *Urban Rev.* 47, 717–741.
- Green, T.L., 2017. Community-based equity audits: a practical approach for educational leaders to support equitable community-school improvements. *Educ. Adm. Q.* 53 (1), 3–39.
- Grissom, J.A., Harrington, J.R., 2010. Investing in administrator efficacy: an examination of professional development as a tool for enhancing principal effectiveness. *Amer. J. Educ.* 116, 583–612.

- Gurr, D., Drysdale, L., 2007. Models of successful principal leadership: Victorian case studies. In: Day, C., Leithwood, K. (Eds.), *Successful Principal Leadership in Times of Change. Studies in Educational Leadership*, vol. 5. Springer, Dordrecht.
- Gurr, D., Drysdale, L., 2018. Leading high needs schools: findings from the international school leadership development network. *Int. Stud. Educ. Adm.* 46 (1), 147–155.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leader. Pol. Sch.* 4 (3), 221–239.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49, 125–142. Retrieved from <https://doi.org/10.1108/09578231111116699>.
- Hallinger, P., 2018. Bringing context out of the shadows of leadership. *Educ. Manag. Adm. Leader* 46 (1), 5–24.
- Hallinger, P., Murphy, J.F., 1986. The social context of effective schools. *Am. J. Educ.* 94, 328–355.
- Hansen, C., 2018. Why rural principals leave. *Rural Educat.* 39 (1), 41–53.
- Harris, A., 2010. Effective leadership in challenging schools. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *The International Encyclopedia of Education*, third ed. Oxford University Press, Oxford, pp. 777–782.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52, 653–681.
- Herman, R., Gates, S.M., Arifkhanova, A., Bega, A., Chaves-Herrerias, E.R., Han, E., 2017. *School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review*. RAND Corporation, Santa Monica, CA.
- Honig, M.I., 2012. District central office leadership as teaching: how central office administrators support principals' development as instructional leaders. *Educ. Adm. Q.* 48 (4), 733–774.
- Honig, M.I., Rainey, L.R., 2014. Central office leadership in principle professional learning communities: the practice beneath the policy. *Teach. Coll. Rec.* 116, 1–48.
- Huggins, K.S., Klar, H.W., Andreoli, P.M., 2020. Facilitating leadership coach capacity for school leadership development: the intersection of structured community and experiential learning. *Educ. Adm. Q.* 1–31.
- Ingersoll, R.M., Merrill, E., Stuckey, D., Collins, G., 2018. Seven Trends: The Transformation of the Teaching Force – Updated October 2018. CPRE Research Reports.
- Klar, H.W., Brewer, C.A., 2013. Successful leadership in high-needs schools: an examination of core leadership practices enacted in challenging contexts. *Educ. Adm. Q.* 49 (5), 768–808.
- Klar, H.W., Huggins, K.S., 2020. *Developing Rural School Leaders: Building Capacity through Transformative Leadership Coaching*. Routledge, New York, NY.
- Klar, H.W., Huggins, K.S., Andreoli, P.M., Buskey, F.C., 2019a. Developing rural school leaders through leadership coaching: a transformative approach. *Leader. Pol. Sch.* 1–21.
- Klar, H.W., Moyi, P., Yimaki, R.M., Hardie, S., Andreoli, P.M., Dou, J., Harrington, K., Roper, C., Buskey, F.C., 2019b. Getting off the list: leadership, learning, and context in two rural, high-needs schools. *J. Sch. Leader.* 1–22.
- Kruse, S.D., Louis, K.S., Bryk, A.S., 1995. An emerging framework for analyzing school-based professional community. In: Louis, K.S., Kruse, S.D. (Eds.), *Professionalism and Community: Perspectives on Reforming Urban Schools*. Corwin Press, Thousand Oaks, CA, pp. 23–44.
- Lacour, M., Tislington, L.D., 2011. The effects of poverty on academic achievement. *Educ. Res. Rev.* 6 (7), 522–527.
- Langley, G.J., Moen, R.D., Nolan, K.M., Nolan, T.W., Norman, C.L., Provost, L.P., 2009. *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*, second ed. Jossey-Bass, San Francisco, CA.
- Leithwood, K., Azah, V.N., 2017. Characteristics of high-performing school districts. *Leader. Pol. Sch.* 16 (1), 27–53.
- Leithwood, K., Jantzi, D., 2005. A review of transformational school leadership research 1996–2005. *Leader. Pol. Sch.* 4, 177–199.
- Leithwood, K., Riehl, C., 2005. What we know about successful school leadership. In: Firestone, W., Riehl, C. (Eds.), *A New Agenda: Directions for Research on Educational Leadership*. Teachers College Press, New York, NY, pp. 12–27.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28 (1), 27–42.
- Leithwood, K., Sun, J., Schumacker, R., 2019. How school leadership influences student learning: a test of “the four paths model”. *Educ. Adm. Q.* 1–30.
- Levine, D.U., Lezotte, L.W., 1990. *Unusually Effective Schools: A Review and Analysis of Research and Practice*. National Center for Effective Schools Research and Development.
- Lochmiller, C.R., 2014. Leadership coaching in an induction program for novice principals: a 3-year study. *J. Res. Lead. Educ.* 9 (1), 59–84.
- Lochmiller, C.R., Chesnut, C.E., 2016. Preparing turnaround leaders for high needs urban schools. *J. Educ. Adm.* 55 (1), 85–102.
- Lock, G., Budgen, F., Lunay, R., Oakley, G., 2012. The loneliness of the long-distance principal: tales from remote western Australia. *Aust. Int. J. Rural Educ.* 22 (2), 65–77.
- Logan, J.P., 2014. *School Leadership Through Action Research*. Pearson, New York, NY.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., Anderson, S.E., 2010. *Investigating the Links to Improved Student Learning: Final Report of Research Findings*. Wallace Foundation, New York, NY.
- Manna, P., 2015. *Developing Excellent School Principals to Advance Teaching and Learning: Considerations for State Policy*. Wallace Foundation, New York, NY.
- Mid-Continent Research for Education and Learning, 2005. *Final Report: High-Needs Schools – What Does It Take to Beat the Odds?* Mid-Continent Research for Education.
- Mitchell, C., Sackney, L., 2011. *Profound Improvement: Building Learning-Community Capacity on Living System Principles*, second ed. Routledge, New York, NY.
- Moolenaar, N.M., 2012. A social network perspective on teacher collaboration in schools: theory, methodology, and applications. *Am. J. Educ.* 119 (1), 7–39.
- Mulford, B., Silins, H., 2003. Leadership for organisational learning and improved student outcomes – what do we know? *Camb. J. Educ.* 33 (2), 175–195.
- Murakami, E., Notman, R., Gurr, D., 2019. Educational leadership, culture, and success in high-need schools: Introduction from the editors. In: Murakami, E., Gurr, D., Notman, R. (Eds.), *Educational Leadership, Culture, and Success in High-Need Schools*. Information Age Publishing, Charlotte, NC, pp. xi–xix.
- Musu-Gillette, L., de Brey, C., McFarland, J., Hussar, W., Sonnenberg, W., Wilkinson-Flicker, S., 2017. Status and Trends in the Education of Racial and Ethnic Groups 2017 (No. NCES 2017-051). Retrieved from U.S. Department of Education, National Center for Education Statistics website. <http://nces.ed.gov/pubsearch>.
- National Center for Education Statistics (NCES), 2018. Concentration of public school students eligible for free or reduced-price lunch. https://nces.ed.gov/programs/coe/indicator_clb.asp.
- No Child Left Behind Act of 2001, 2001. No Child Left Behind Act of 2001. Pub. L. No. 107-110.
- Okiwa, N., Barnett, B.S., 2018. Four successive school leaders' response to a high needs urban elementary school context. *Int. Stud. Educ. Adm.* 46 (1), 45–85.
- Park, S., Hironaka, S., Carver, P., Nordstrum, L., 2013. *Continuous Improvement in Education*. Carnegie Foundation for the Advancement of Teaching.
- Pashiardis, P., Brauckmann, S., Kafa, A., 2018. Let the context become your ally: school principalship in two cases from low performing schools in Cyprus. *Sch. Leader. Manag.* 1–18.
- Pendola, A., Fuller, E.J., 2018. Principal stability and the rural divide. In E. McHenry-Sorber & D. Hall (Eds.), *The diversity of rural educational leadership [Special Issue]* *J. Res. Rural Educ.* 34 (1), 1–20.
- Pont, B., 2020. A literature review of school leadership policy reforms. *Eur. J. Educ.* 55, 154–168.
- Preston, J.P., Jakubiec, B.A., Koymans, R., 2013. Common challenges faced by rural principals: a review of the literature. *Rural Educat.* 35 (1).
- Reedy, M., Doiron, T.A., 2018. *Certification of Educational Leaders*. University Council for Educational Administration, Charlottesville, VA.
- Reiss, K.J., 2015. *Leadership coaching for educators: bringing out the best in school administrators*, Second ed. Thousand Oaks, CA.
- Riley, D.L., Meredith, J., 2017. State efforts to strengthen school leadership: insights from CCSSO action groups. Policy Studies Associates, Inc.
- Robertson, J., 2016. *Coaching leadership: building educational leadership capacity through partnership*, Second ed. Wellington, NZ.
- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Rohanna, K., 2017. Breaking the “adopt, attack, abandon” cycle: a case for improvement science in K-12 education. *N. Dir. Eval.* (153), 65–77.
- Sahlberg, P., 2016. The global educational reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley-Blackwell, West Sussex, pp. 128–144.

- Shaked, H., Schechter, C., 2014. Systems school leadership: exploring an emerging construct. *J. Educ. Adm.* 52 (6), 792–811.
- Showalter, D., Hartman, S.L., Johnson, J., Klein, R., 2019. *Why Rural Matters 2018–2019: The Time is Now*. Rural School and Community Trust.
- South Carolina Department of Education, 2017. South Carolina School Improvement Framework (version 2.0). Columbia, SC, Retrieved from South Carolina Department of Education website. <https://ed.sc.gov/districts-schools/school-improvement/school-improvement-programs/sc-school-improvement-framework/>.
- Sparks, S.D., 2018. A primer on continuous improvement. *Educ. Week* 37 (19).
- Spillane, J.P., Louis, K.S., 2005. School improvement processes and practices: professional learning for building instructional capacity. *Yearbk. Natl. Soc. Stud. Educ.* 101 (1), 83–104.
- Starr, K., White, S., 2008. The small rural school principalship: key challenges and cross-school responses. *J. Res. Rural Educ.* 23 (5), 1–11.
- Stoll, L., Bolam, R., 2005. Developing leadership for learning communities. In: Coles, M., Southworth, G. (Eds.), *Developing Leadership: Creating the Schools of Tomorrow*. Open University Press, Maidenhead, Berkshire, UK, pp. 50–64.
- Sullivan, K., McConney, A., Perry, L.B., 2018. A comparison of rural educational disadvantage in Australia, Canada, and New Zealand using OECD's PISA. *SAGE Open* 1–12.
- Thessin, R.A., 2019. Establishing productive principal/principal supervisor partnerships for instructional leadership. *J. Educ. Adm.* 57 (5), 463–483.
- Tichnor-Wagner, A., Wachen, J., Cannata, M., Cohen-Vagel, L., 2017. Continuous improvement in the public school context: understanding how educators respond to plan-do-study-act cycles. *J. Educ. Change* 18, 465–494.
- U.S. Department of Education, 2018. Non-Regulatory Guidance Title I, Part A: Improving Basic Programs Operated By Local Educational Agencies. Non-Regulatory Guidance Title I, Part A of the Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act of 2015. Author, Washington, DC.
- Wieczorek, D., Manard, C., 2018. Instructional leadership challenges and practices of novice principals in rural schools. *J. Res. Rural Educ.* 34 (2), 1–21.
- Wood, J.N., Finch, K., Mirecki, R.M., 2013. If we get you, how can we keep you? Problems with recruiting and retaining rural administrators. *Rural Educat.* 34 (2).
- Young, M.D., Winn, K.M., Reedy, M.A., 2017. The Every student Succeeds act: strengthening the focus on educational leadership. *Educ. Adm. Q.* 53 (5), 705–726.
- Zepeda, S.J., Parylo, O., Bengtson, E., 2014. Analyzing principal professional development practices through the lens of adult learning theory. *Prof. Dev. Educ.* 40 (2), 295–315.
- Zepeda, S.J., Parylo, O., Klar, H.W., 2017. Educational leadership for teaching and learning. In: Waite, D., Bogoch, I. (Eds.), *The Wiley International Handbook of Educational Leadership*. Wiley & Sons, Hoboken, NJ, pp. 227–252.

Relevant websites

- <https://nces.ed.gov/fastfacts/display.asp?id=898> – National Center for Educational Statistics (NCES) Education Fast Facts.
- <https://www2.ed.gov/programs/titleiparta/index.html> – United States Department of Education Improving Basic Programs Operated by Local Educational Agencies (Title I, Part A).
- <https://ed.sc.gov/districts-schools/school-improvement/school-improvement-programs/sc-school-improvement-framework/> – South Carolina School Improvement Framework.
- <https://www.cerra.org> – South Carolina Center for Educator Recruitment, Retention, and Advancement (CERRA).
- <https://www.cerra.org/rural-recruitment-initiative.html> – South Carolina Center for Educator Recruitment, Retention, and Advancement (CERRA) Rural Recruitment Initiative.
- <https://isldn.weebly.com> – International School Leadership Development Network (ISLDN).

Leading together: developing teacher leadership

Jonathan Eckert, Baylor University, Waco, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Teacher leadership	196
Teacher leadership development	197
Teacher leadership through collective leadership	197
A model of collective leadership development	198
Teachers and administrators as catalysts for school improvement	199
The future of teacher leadership for school improvement	200
References	201

Since at least the 1980s teacher leadership has been viewed as a panacea for many education ills. For decades, we have known that the two most significant school-level factor influencing student outcomes are teachers and school leaders (Chetty et al., 2011; Hanushek, 1992, 2010; Hanushek and Rivkin, 2012; Kane and Staiger, 2008; Leithwood et al., 2004; Rockoff, 2004). By combining “teacher” and “leadership,” policymakers, practitioners, and researchers hoped to identify an elegant solution to adaptive challenges related to unclear problems and solutions (Heifetz, 1994).

However, as is the case with most adaptive challenges, the solution has not been elegant or simple. Even the term, “teacher leadership” has remained somewhat ill-defined but widely used. Calls for teacher leadership increased (Curtis, 2013; Duncan, 2014; Pennington, 2013) and there were stronger arguments for why it was needed (Goldstein, 2014; Mangin and Stoelinga, 2008; Pounder, 2006). Advocacy for teacher leadership abounded. New models of and opportunities for teacher leadership proliferated (Alexandrou and Swaffield, 2014; Berry, 2010; Hargreaves and Fullan, 2012; Margolis, 2012). Substantial numbers of teachers have indicated they want opportunities to lead (MetLife, 2013).

Comprehensive reviews of the literature on teacher leadership have identified definitions, cultivated a theoretical grounding, advocated for enhanced teacher voice in policymaking, and indicated its role in school improvement with mostly distal connections to student outcomes (Murphy, 2005; Schott et al., 2020; Wenner and Campbell, 2017; York-Barr and Duke, 2004). Three literature reviews that examine teacher leadership literature from 1980 to 2018 and are particularly useful for framing the teacher leadership landscape (see Table 1).

Teacher leadership

York-Barr and Duke reviewed relevant literature from 1980 to 2002. According to their definition, teacher leadership is “the process by which teachers, individually or collectively, influence their colleagues, principals, and other members of school communities to improve teacher and learning practices with the aim of increased student learning and achievement” (2004, pp. 287–288). Others have attempted to expand this definition. For example, in the consultancy literature, Katzenmeyer and Moller (2009) broadened

Table 1 Teacher leadership literature since 1980.

	York-Barr and Duke (2004)	Wenner and Campbell (2017)	Schott et al. (2020) ^a
Years reviewed	1980–2002	2004–2013	2014–2018
Total studies included in review	34	83	84
Qualitative studies	19 (56%)	60 (74%)	59 (70%)
Quantitative studies	8 (23%)	3 (6%)	12 (14%)
Mixed methods studies	7 (21%)	20 (20%)	13 (16%)
Definition of teacher leadership	“The process by which teachers, individually or collectively, influence their colleagues, principals, and other members of school communities to improve teacher and learning practices with the aim of increased student learning and achievement” (p. 287–288).	“Teachers who maintain K-12 classroom-based teaching responsibilities, while also taking on leadership responsibilities outside of the classroom” (p. 5).	Authors recommend using York-Barr and Duke’s definition as they found that 39 (42%) of the studies included in their review lacked a definition of teacher leadership.

^aBased on analysis for Schott et al. (2020).

this definition, observing that “teacher leaders lead within and beyond the classroom; identify with and contribute to a community of teacher learners and leaders; influence others toward improved educational practice; and accept responsibility for achieving the outcomes of their leadership” (p. 6). York-Barr and Duke (2004) described a conceptual framework to guide inquiry into differentiated paths teachers take to organizational influence, professionalism, and teaching and learning. The framework has influenced subsequent teacher leadership research.

In their review of literature from 2004 to 2013, Wenner and Campbell (2017) found that teacher leadership research focused on roles beyond the classroom—roles that supported student outcomes by influencing policy outcomes and the professional learning of peers. Although the theoretical grounding improved since York-Barr and Duke’s review, the grounding remained limited and rarely focused on justice or equity issues. Wenner and Campbell highlighted the importance of administrators, school structures, and norms for supporting or constraining teacher leadership. Wenner and Campbell (2017) define teacher leadership as “teachers who maintain K-12 classroom-based teaching responsibilities, while also taking on leadership responsibilities outside of the classroom” (p. 5). The proliferation of various definitions or sometimes ill-defined teacher leadership has limited the theoretical grounding of the concept. However, their findings represented progress in relation to the grounding York-Barr and Duke described as “largely atheoretical” (2004, p. 291).

Schott et al. (2020) reviewed teacher leadership studies from 2014 to 2018. They recommended a return to York-Barr and Duke’s definition of teacher leadership to undergird the theoretical foundation of the research as 42% of the studies they included did not define teacher leadership. They also cited the methodological limitations of some of the studies and recommended greater attention to conceptual clarity, cross-country research designs, and attention to some of the negative aspects of teacher leadership as every study in their review demonstrated positive outcomes.

Teacher leadership development

There are several pathways by which teachers might come to leadership. Educational administrators may “appoint and anoint” teachers to leadership positions and leadership work outside the classroom (Smylie and Denny, 1990). Often, administrators are encouraged to share or “distribute” leadership work, involving teachers in decision-making and school improvement (e.g., Bryk et al., 2010; Louis et al., 2010; Spillane, 2006).

Another pathway to teacher leadership is initiated by teachers themselves. Teachers are encouraged to step forward and assume leadership at the school, district, and even state and federal policy levels. They are enticed with images of “superhero” teacher leaders and successful “teacherpreneurs” (Berry et al., 2013) and new influential roles at the national level (Eckert et al., 2016). They are exhorted to “speak truth to power” and to “bust” through the organizational, administrative, and institutions “cages” that constrain them (Hess, 2015).

While these paths can lead to productive leadership and some teachers may come to leadership by way of one or more of them, a more influential and sustainable pathway is teacher leadership development—the intentional, systematic, systemic development of teachers’ capacity for organizational-level leadership and its practice. The path of development is different from the paths of advocacy, exhortation, invitation, and self-initiation. None of those paths necessarily involve development and, indeed, few actually do. These other pathways become problematic as they assume that teachers who are tapped for or who self-initiate leadership possess the capacity and will receive adequate preparation and support to engage leadership work successfully. Advocacy for teacher leadership is not the same as the development of it. Building a new role or providing an opportunity for leadership does not mean that teachers will come forward or that they will be able to perform the work effectively (Smylie and Eckert, 2018).

Teacher leadership, even when well-supported, designed, and implemented is limited in its impact when performed in isolation from administrators. Therefore, educators, researchers, and policymakers have advocated for numerous forms of leadership – teacher (Murphy, 2005; Schott et al., 2020; Wenner and Campbell, 2017; York-Barr and Duke, 2004), distributed (Muijs and Harris, 2006; Spillane et al., 2001; Supovitz, 2008), transformational combined with instructional (Day et al., 2016) shared (MacBeath and Dempster, 2009), collaborative (DeWitt, 2017), or collective (Leithwood and Mascal, 2008; Louis et al., 2010). New models of leadership, particularly those involving teachers, have proliferated (Alexandrou and Swaffield, 2014; Berry et al., 2013; Hargreaves and Fullan, 2012; Margolis, 2012).

Teacher leadership through collective leadership

Regardless of the term or model used, education leadership requires teachers and administrators to work together toward shared organizational goals. Therefore, a more effective way to think about leadership for school improvement is to consider how to develop teachers and administrators together. “Collective leadership encompasses the practices through which teachers and administrators influence colleagues, policymakers, and others to improve teaching and learning” (Eckert, 2018, p. 5). Whether transformational, instructional, or both, collective leadership is a useful construct because it emphasizes joint goal setting and strategic implementation of those goals in the service of the primary purpose of schools – teaching and learning. This is an important stipulation because in some cases, terms such as distributed leadership have been reduced to task delegation from administrators to teachers. Moreover, collective leadership has been found to have greater influence on student outcomes than individual leadership (Louis et al., 2010). If this type of leadership can enhance collective teacher efficacy, the belief that faculty can positively affect

students, outcomes for students improve. Collective teacher efficacy is one of the most influential factors influencing student learning (Hattie, 2015).

Typically, schools and districts have focused on the practice of teacher and administrative leadership separately. This is inadequate for meaningful, sustainable school improvement that requires the coordinated leadership development of both teachers and administrators. There has been much less attention in scholarly and professional literature to teacher or administrator leadership development than to its purposes, design (e.g., attention to “hybrid” divisions of roles), its practice, and its efficacy. This lack of emphasis on development leads to practices in which teachers add work to the point of burnout or experience unhelpful delegation of work under the guise of leadership development (Eckert et al., 2016; Eckert, 2018). Emphasis on exceptional teachers simply doing more, delegating work differently, or focusing on individual leadership will do little to change complex education systems (Louis et al., 2010). Attention that has been paid to teacher leadership development has drawn primarily from the educational literature, notably research and writing from earlier efforts to implement teacher career ladders, master and mentor teacher programs, and participative decision making (Murphy, 2005; Schott et al., 2020; Wenner and Campbell, 2017; York-Barr and Duke, 2004).

Collective leadership development more accurately reflects the type of development necessary for school improvement as it addresses the mutual growth of teachers and administrators together (Eckert, 2019; Eckert and Butler, 2021). Literature on work design and redesign (e.g., Campion et al., 2005; Hackman and Oldham, 1980; Humphrey et al., 2007), leadership development generally and across organizational type and sector (e.g., Avolio, 2010; Conger, 1992; Day et al., 2004; Fulmer, 1997; Popper, 2005; Van Velsor et al., 2010; Yukl, 2013), and teacher leadership (e.g., Berg, 2018; Lieberman and Miller, 2004; Mangin and Stoelinga, 2008; Murphy, 2005; Wenner and Campbell, 2017; York-Barr and Duke, 2004) formed the foundation of a model of collective leadership development (see Fig. 1). Additionally, the early theoretical work design literature (Hackman and Oldham, 1980), form the conceptual foundation of collective leadership as a shift in the way teachers and administrators work and also has a way to examine collective leadership’s potential efficacy. The literature demonstrates two issues: there has not been enough focus on the “collective” or “development” of “collective leadership development.” Developing teacher and administrative leadership together have been overlooked in research and practice.

A model of collective leadership development

Fig. 1 represents a model of collective leadership development (Eckert and Butler, 2021).

The focus of the model is on leadership, not the leader, just as focusing on teaching instead of the teacher is beneficial (Hiebert and Stigler, 2017). This theory views leadership functionally (Firestone, 1996), as redesigned work to be performed rather than as a particular role to be assumed, considering this type of leadership as redesigned work (Mayrowetz et al., 2007; Smylie, 1994). The model captures the essential elements of a school, particularly as they relate to leadership for school improvement. The various elements influence leadership development. The theory supporting the analytic model posits that the development and performance of this new work is based on seven conditions:

- (1) A vision and strategy that is supported by dynamic, strategically “ambidextrous” leadership development (Fullan, 2005; Hargreaves and Fink, 2006; McCauley, 2008; Mumford et al., 2007; Mumford et al., 2007; Smylie, 2010; Tushman and O’Reilly, 1996) that is democratic, participative inclusive leadership (Bryk et al., 2010; Leithwood and Louis, 2012; Light, 1998);
- (2) Supportive administration that believes in this kind of leadership (Birky et al., 2006; Camburn et al., 2008; Murphy et al., 2009; Smylie and Brownlee-Conyers, 1992; Van Velsor et al., 2010) that recalibrates authority relationships (Murphy, 2005);
- (3) Resources and initial leadership capacity that provide the necessary talent, time, and space to lead together (Avolio and Hannah, 2008; Bond et al., 2008; Conger, 1992; Martineau, 2004);

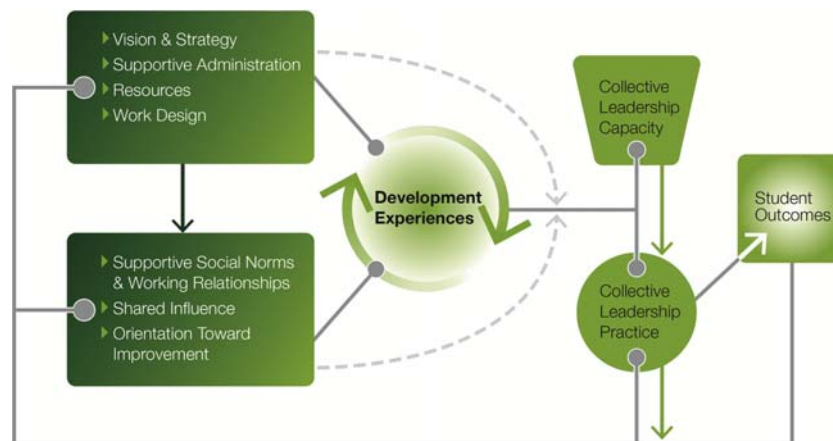


Fig. 1 Collective leadership development model.

- (4) Adequate work design to facilitate collaboration, improvement, and the spread of teaching expertise (Campion et al., 2005; Hackman and Oldham, 1980; Margolis, 2012; Smylie and Denny, 1990);
- (5) Supportive social norms and working relationships that foster relational trust (Bryk and Schneider, 2002; Hargreaves and O'Connor, 2018; Muijs and Harris, 2007; Van Velsor and McCauley, 2004);
- (6) Shared influence of administrators and teachers on themselves and others (Day et al., 2004; Lai and Cheung, 2015; Muijs and Harris, 2007); and
- (7) Orientation toward improvement that facilitates growth (Eckert, 2018, 2019; Eckert and Daughtrey, 2019; Smylie and Eckert, 2018).

Of the seven conditions, the first four are antecedent to the other three, but there are systemic interactive relationships among all seven (Bass, 1990; Firestone, 1996; Smylie, 2010; Yukl, 2013).

Constructed primarily from the literature on teacher leadership and factors that influenced its practice (Murphy, 2005; York-Barr and Duke, 2004), the model builds on Wenner and Campbell's definition of teacher leadership in which teachers lead in and beyond the classroom (2017). The development experiences that flow from the seven conditions are not artificial team-building experiences. Instead, they are the work of school improvement that is co-led by teachers and administrators building on the notion that leadership is work to be performed rather than a particular role to be assumed (Firestone, 1996). The model frames collective leadership as work redesign, a case of broadening teachers' work from an exclusive focus on students and classroom instruction to working with colleagues, administrators, and policymakers at various levels. Work redesign through leadership development experiences is essential for both teachers and administrators if efforts are to be coordinated, sustained, and amplified. Leadership development through this shared work leads to increased relational trust (Bryk and Schneider, 2002; Tschannen-Moran, 2004) and professional capital—the systematic development of human, social, and decisional capital (Hargreaves and Fullan, 2012)—which leads to improved leadership development experiences, increased leadership capacity, and enhanced leadership practice and outcomes (Eckert and Butler, 2021). Effective practitioners of the leadership work in the service of shared goals becoming *de facto* leaders thus eliminating concerns about who should or should not be leading. The shared work identifies the leaders and flattens the organizational hierarchy (Skrtic, 1991).

In the model, dotted lines extend from the seven conditions beyond the development experiences as either supports or constraints to leadership capacity. For example, if a principal continues to expand the role of teachers based on their expertise and leadership in shared work, their collective leadership capacity will increase with the principal serving as a support. If, however, social norms and working relationships are poor, this condition will limit the teacher leaders' development of capacity. The model is iterative in that school leaders must attend to the seven conditions collectively and repeatedly to maximize growth. If any condition is under-developed, readiness for school improvement through collective leadership will be compromised.

As the collective leadership capacity of teachers and administrators increases, collective leadership practice should improve commensurately. As the model is predicated on work toward shared goals, leadership capacity and leadership practice are tightly-coupled. Improved leadership practice results in enhanced student outcomes. Because collective leadership is typically aligned with instructional practices, student learning gains, while still distal, are more proximal than other school improvement efforts such as facilities enhancement, resource allocation, or organizational re-structuring (Eckert, 2019).

Student learning outcomes in turn influence collective leadership practice, development experiences, and the seven conditions. The model results in a virtuous cycle of improvement as conditions improve even as external factors such as economic, social, and political realities influence school improvement. Due to the broad participation required by collective leadership, schools are more likely to find multiple, context-specific paths to improvement (Smylie and Eckert, 2018; Eckert, 2019). This concept of "equifinality" (Burke, 2014) is an important consideration when policymakers and educational leaders attempt to leverage collective leadership for school improvement. Collective leadership is not a formula or a top-down solution, but it can support continuous school improvement.

Teachers and administrators as catalysts for school improvement

Context matters a great deal in the development and enactment of leadership work. Principal support is a catalyst for leadership development and is contingent on initial teacher capacity, school conditions, resources, and work design. These antecedent factors shape the development experiences that either enhance or diminish trust in pursuit of improved collective leadership practice for improved student outcomes. A catalyst is something that accelerates a reaction. The catalyst is not the focus of the reaction. In an organizational context, a catalyst speeds up the work of others.

In schools where strong collective leadership exists, there is trust between school administrators and teachers that develops through catalytic relationships (Smylie et al., 2007). Principals are essential for instructional leadership development (Leithwood et al., 2004; Mangin, 2007) especially as it relates to shared vision and coherence (Firestone et al., 2005). Although teacher leadership is often used as an umbrella term for many school leadership efforts (Neumerski, 2012), collective leadership is even broader due to its inclusion of administrators.

Catalytic administrators serve particular roles in collective leadership development. They bridge and buffer external initiatives (Aldrich and Herker, 1977; Durand et al., 2015; DiPaola and Tschannen-Moran, 2005) as they bring coherence to school improvement by building connections to external groups and limiting work that does not serve teachers and students well. Their ability to

reduce the impact of outside demands on teachers increases the likelihood that collective leadership will lead to school improvement (Eckert, 2019). Because being a catalyst requires humility, the principals and their leadership teams are primarily composed of humble educators who are highly relational with other educators and students (Schein and Schein, 2018). Humble leaders express vulnerability and empathy for others and do not take credit for improving student outcomes. Instead, they highlight the work that others are doing (Eckert, 2018).

Principals' antecedent support of teacher leadership combined with the initial capacity of teachers for leadership is predictive of schools' opportunities for improvement. The impact of principal support is not related to the teacher to administrator ratio. In rural, urban, and suburban contexts with teacher to administrator ratios that ranged from 9:1 to 14:1, the development of collective leadership capacity through teacher-led initiatives, formal Professional Learning Communities, and informal problem-solving teams influences principal support and increased capacity of teachers in iterative feedback loops (Eckert, 2019). The capacity of the teachers for leadership work was central to the transformation of a school. Capacity can be constrained or supported by many factors such as preparation (Darling-Hammond et al., 2005; Jensen et al., 2016), professional development (Darling-Hammond et al., 2009; Jensen et al., 2016), or colleagues (Hattie, 2015; Jackson and Bruegmann, 2009; Ronfeldt et al., 2015). If teachers are already at capacity with their instructional responsibilities, adding additional work—even if it is meaningful—will not necessarily improve outcomes for students (Eckert et al., 2016).

Capacity is something that is malleable and can grow through deliberate practice. Deliberate practice is a vehicle for collective leadership development and is built into the analytic model. Deliberate practice is not just experience, it is practice that includes feedback, reflection, and opportunities for improvement (Ericsson et al., 1993). In schools, this occurs best with other educators. Over a ten-year period, teachers working in schools with strong professional environments improved 38% more than peers in schools with weak environments (Kraft and Papay, 2016). Additionally, student outcomes are dependent on collective teacher efficacy (Hattie, 2015). Where greater leadership capacity is perceived, a greater sense of teacher efficacy follows.

The establishment of relational trust (Bryk and Schneider, 2002) and the expansion of professional capital for teachers (Hargreaves and Fullan, 2012) is evident at schools with higher degrees of collective leadership. Teachers, administrators, and students collaboratively develop School Improvement Plans, teachers are leading staff meetings, and relationships between administrators, teachers, and school communities improve through shared work (Eckert, 2018). Leadership teams report that relational trust is essential for collective leadership development (Bryk and Schneider, 2002; Bryk et al., 2010). Unlike contractual trust—trust based on agreed upon conditions and work products—relational trust develops as teachers, students, and administrators respect others' work, value their competence, demonstrate integrity, and communicate personal regard for others (Bryk and Schneider, 2002). The ability to speak openly, honestly, and regularly is necessary between teachers and administrators. Often, effective communication manifests itself in administrators' "open door" policies for teachers who need to discuss issues or suggest opportunities for improvement. This openness is also needed for teachers to welcome others to observe their teaching. This openness is especially important in the development of followership (Day et al., 2004). Especially in contexts where collective leadership is being practiced, followership is not guaranteed, and therefore, the relationships developed through the work determine leaders and followers.

Another way to build relational trust is through the development of professional capital. Professional capital is comprised of human capital, social capital, and decisional capital. Improvement, both organizational and individual, occurs through transparent practice, collection of evidence and personal judgment, and collective responsibility coupled with collective accountability. The professional capital is most evident in schools that elevate leaders who were effective educators. Teachers or administrators who do effective work with students or colleagues in support of school improvement become leaders. Their credibility is derived from their expertise and rarely from their positions. This kind of professional capital is evident in schools with strong collective leadership, although the capital sometimes creates friction. With this kind of leadership rising from the student and classroom level, there are competing interests that require the support of teachers and administrators for support and coherence. Trust and professional capital are essential for accelerating collective leadership development and are often the result of effective development experiences and activities (Hargreaves and Fullan, 2012).

The future of teacher leadership for school improvement

While researchers and practitioners have made progress since the 1980s (Schott et al., 2020; Wenner and Campbell, 2017; York-Barr and Duke, 2004), much work remains for the development of teacher leadership as a vehicle for school improvement. Leadership work and development are done simultaneously and require both administrators and teachers. The design of the work (Hackman and Oldham, 1980), resources, and conditions influence the development experiences available to teachers and administrators. The work design of most schools can be challenging for collective leadership development. Providing time for collaboration, reciprocal observation, co-teaching, and planning varies depending on countries' education systems and other professions (Hiebert and Stigler, 2017; Jensen et al., 2014; Jensen et al., 2016; Margolis, 2012). Variation in work design exists across schools attempting to elevate the leadership of teachers. Depending on school size, there might not be formal roles for teacher leadership, opportunities for planning with other teachers in the same discipline, or job-embedded professional learning opportunities. However, even and especially in small schools, there is leadership work that needs to be done. The freedom that teachers derive from not being concerned with whom has a formal title to lead the work, allows them to step into the work and become *de facto* leaders. School climate can improve as a goal of the leadership work and also as a by-product of working alongside others. This improved climate leads to

more development experiences and meaningful changes in work design (Bryk et al., 2010; Eckert, 2019; Smylie, 2010). Further cross-country studies could inform ways to improve work design and leadership development.

One additional component of work design as it relates to collective leadership development is the sharing of evidence between schools for systemic improvement. Tony Bryk and his colleagues at the Carnegie Foundation (Bryk et al., 2015) prescribe networked improvement communities (NICs). They suggest systematically collecting evidence of what does and does not work and sharing those data between organizations committed to improvement. The impetus for data collection is not accountability, but improvement. Due to the need for reciprocal leadership development, this orientation toward improvement is about systematic development of which collective leadership must necessarily be a part. No longer do schools have a lone instructional leader. In a comprehensive high school with departments ranging from performing arts to English to world languages to science, the notion of one instructional leader is unrealistic. A school's orientation should be toward improvement, not overly simplistic accountability (Murphy, 2005; Smylie, 2010). This occurs through work that is done together by teachers and administrators. Leadership teams develop through mutual respect. To expand and scale this work, evidence of impact should be shared across schools, which requires work design that affords the time, opportunity, and networks for collective improvement around shared goals.

References

- Aldrich, H., Herker, D., 1977. Boundary spanning roles and organizational structure. *Acad. Manag. Rev.* 2, 217–230.
- Alexandrou, A., Swaffield, S., 2014. *Teacher Leadership and Professional Development*. Routledge, New York, NY.
- Avolio, B.J., Hannah, S.T., 2008. Developmental readiness: accelerating leader development. *Psychol. J. Pract. Res.* 60, 331–347.
- Avolio, B.J., 2010. *Full Range Leadership Development*, second ed. Sage, Thousand Oaks, CA.
- Bass, B.M., 1990. *Handbook of Leadership: Theory, Research, and Managerial Applications*, third ed. Free Press, New York, NY.
- Berg, J.H., 2018. Leading in Sync: Teacher Leaders and Principals Working Together for Student Learning. ASCD, Alexandria, VA.
- Berry, B., Byrd, A., Wieder, A., 2013. *Teacherpreneurs: Innovative Teachers Who Led but Don't Leave*. Jossey-Bass, San Francisco, CA.
- Berry, B., 2010. *Teaching 2030: What We Must Do for Our Students and Public Schools*. Teachers College Press, New York, NY.
- Birky, V.D., Shelton, M., Headley, S., 2006. An administrator's challenge: encouraging teachers to be leaders. *NASSP Bull.* 90, 87–101.
- Bond, F.W., Flaxman, P.E., Bunce, D., 2008. The influence of psychological flexibility on work redesign: mediated moderation of a work reorganization intervention. *J. Appl. Psychol.* 93, 645–654.
- Bryk, A., Schneider, B., 2002. *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation, New York, NY.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press, Chicago.
- Bryk, A.S., Gomez, L.M., Grunow, A., LeMahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press, Cambridge, MA.
- Burke, W.W., 2014. *Organizational Change: Theory and Practice*, fourth ed. Sage, Los Angeles, CA.
- Camburn, E.M., Kimball, S.M., Lowenhaupt, R., 2008. Going to scale with teacher leadership: lessons learned from a districtwide literacy coach initiative. In: Mangin, M.M., Stoeilinga, S.R. (Eds.), *Effective Teacher Leadership: Using Research to Inform and Reform*. Teachers College Press, New York, NY, pp. 120–143.
- Campion, M.A., Mumford, T.V., Morgeson, F.P., Nahrgang, J.D., 2005. Work redesign: eight obstacles and opportunities. *Hum. Resour. Manag.* 44, 367–390.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2011. *The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood*. National Bureau of Economic Research, Washington, DC.
- Conger, J., 1992. *Learning to Lead: The Art of Transforming Managers into Leaders*. Jossey-Bass, San Francisco, CA.
- Curtis, R., 2013. *Finding a New Way: Leveraging Teacher Leadership to Meet Unprecedented Demands*. The Aspen Institute, Washington, DC.
- Darling-Hammond, L., Bransford, J., LePage, P., Hammerness, K., Duffy, H. (Eds.), 2005. *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, San Francisco, CA.
- Darling-Hammond, L., Wei, R.C., Andree, A., Richardson, N., Orphanos, S., 2009. *Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad*. National Staff Development Council, Washington, DC.
- Day, D.V., Zaccaro, S.J., Halpin, S.M. (Eds.), 2004. *Leader Development for Transforming Organizations: Growing Leaders for Tomorrow*. Erlbaum, Mahwah, NJ.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes. *Educ. Adm. Q.* 52, 221–258. <https://doi.org/10.1177/0013161X15616863>.
- DeWitt, P.M., 2017. *Collaborative Leadership: Six Influences that Matter Most*. Corwin Press, Thousand Oaks, CA.
- DiPaola, M.F., Tschannen-Moran, M., 2005. Bridging or buffering?: the impact of schools' adaptive strategies on student achievement. *J. Educ. Adm.* 43, 60–71. <https://doi.org/10.1108/09578230510577290>.
- Duncan, A., 2014. Teach to lead: advancing teacher leadership. In: *Teaching and Learning Conference*. Presented at the National Board on Professional Teaching, U.S. Department of Education, Washington, DC.
- Durand, F.T., Lawson, H.A., Wilcox, K.C., Schiller, K.S., 2015. The role of district office leaders in the adoption and implementation of the common core state standards in elementary schools. *Educ. Adm. Q.* 52, 45–74. <https://doi.org/10.1177/0013161X15615391>.
- Forthcoming. Eckert, J., Butler, J., 2021. Teaching and leading for exemplary STEM learning: a multiple-case study. *Elem. Sch. J.*
- Eckert, J., Daughtrey, A., 2019. Teacher leadership development: tracking one district's progress over three years. *Educ. Pol. Anal. Arch.* 27, 1–22. <https://doi.org/10.14507/epaa.27.4130>.
- Eckert, J., Ulmer, J., Khatchatryan, E., Ledesma, P., 2016. Career pathways of teacher leaders: adding and path-finding new professional roles. *Prof. Dev. Educ.* 42, 687–709.
- Eckert, J., 2018. *Leading Together: Teachers and Administrators Improving Student Outcomes*. Corwin Press, Thousand Oaks, CA.
- Eckert, J., 2019. Collective leadership development: emerging themes from urban, suburban, and rural high schools. *Educ. Adm. Q.* 55, 477–509. <https://doi.org/10.1177/0013161X18799435>.
- Ericsson, K.A., Krampe, R.T., Tesch-Romer, C., 1993. The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363–406.
- Firestone, W.A., Mangin, M.M., Martinez, M.C., Polovsky, T., 2005. Content and coherence in district professional development: three case studies. *Educ. Adm. Q.* 41, 413–448.
- Firestone, W.A., 1996. Leadership: roles or functions? In: Leithwood, K., Chapman, J., Corson, D., Hallinger, P., Hart, A. (Eds.), *International Handbook of Educational Leadership*. Sage, Thousand Oaks, CA.
- Fullan, M., 2005. *Leadership and Sustainability: System Thinkers in Action*. Corwin Press, Thousand Oaks, CA.
- Fulmer, R.M., 1997. The evolving paradigm of leadership development. *Organ. Dynam.* 25, 59–72.
- Goldstein, D., 2014. *The Teacher Wars: A History of America's Most Embattled Profession*. Doubleday, New York, NY.
- Hackman, J.R., Oldham, G.R., 1980. *Work Redesign*. Addison-Wesley, Reading, MA.
- Hanushek, E.A., Rivkin, S.G., 2012. The distribution of teacher quality and implications for policy. *Annu. Rev. Econ.* 4, 131–157.
- Hanushek, E.A., 1992. The trade-off between child quantity and quality. *J. Polit. Econ.* 100, 84–117.

- Hanushek, E.A., 2010. The Economic Value of Higher Teacher Quality. National Bureau of Economic Research, Cambridge, MA.
- Hargreaves, A., Fink, D., 2006. Sustainable Leadership. Jossey-Bass, San Francisco, CA.
- Hargreaves, A., Fullan, M., 2012. Professional Capital: Transforming Teaching in Every School. Teachers College Press, New York, NY.
- Hargreaves, A., O'Connor, M.T., 2018. Collaborative Professionalism: When Teaching Together Means Learning for All. Corwin Press, Thousand Oaks, CA.
- Hattie, J., 2015. What Works Best in Education: The Politics of Collaborative Expertise, Open Ideas. Pearson.
- Heifetz, R., 1994. Leadership without Easy Answers. Harvard University Press, Cambridge, MA.
- Hess, F.M., 2015. The cage-busting teacher. Harvard Education Press.
- Hiebert, J., Stigler, J.W., 2017. Teaching versus teachers as a lever for change: comparing a Japanese and a U.S. perspective on improving instruction. *Educ. Res.* 46, 169–176. <https://doi.org/10.3102/0013189X17711899>.
- Humphrey, S.E., Nahrgang, J.D., Morgeson, F.P., 2007. Integrating motivational, social, and contextual work design features: a meta-analytic summary and theoretical extension of the work design literature. *J. Appl. Psychol.* 92, 1332–1356.
- Jackson, C.K., Bruegmann, E., 2009. Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers (No. NBER Working Paper 15202). National Bureau of Economic Research, Cambridge, MA.
- Jensen, B., Hunter, J., Sonnemann, J., Cooper, S., 2014. Making Time for Great Teaching. Grattan Institute, Melbourne, Australia.
- Jensen, B., Roberts-Hull, K., Magee, J., Ginnivan, L., 2016. Not So Elementary: Primary School Teacher Quality in High-Performing Systems. National Center on Education and the Economy, Washington, DC.
- Kane, T.J., Staiger, D., 2008. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. NBER, Cambridge, MA.
- Katzenmeyer, K., Moller, G., 2009. Awakening the Sleeping Giant: Helping Teacher Leaders Develop as Leaders, third ed. Corwin Press, Thousand Oaks, CA.
- Kraft, M.A., Papay, J.P., 2016. Developing Workplaces where Teachers Stay, Improve, and Succeed. The Albert Shanker Institute.
- Lai, E., Cheung, D., 2015. Enacting teacher leadership: the role of teachers in bringing about change. *Educ. Manag. Adm. Leader.* 43, 673–692. <https://doi.org/10.1177/1741143214535742>.
- Leithwood, K., Louis, K.S., 2012. Linking Leadership to Student Learning. Jossey-Bass, San Francisco, CA.
- Leithwood, K., Mascal, B., 2008. Collective leadership effects on student achievement. *Educ. Adm. Q.* 44, 529–561.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K., 2004. How Leadership Influences Student Learning. The Wallace Foundation, Ontario.
- Lieberman, A., Miller, L., 2004. Teacher Leadership. Jossey-Bass, San Francisco, CA.
- Light, P.C., 1998. Sustaining Innovation: Creating Nonprofit and Government Organizations that Innovate Naturally. Jossey-Bass, San Francisco, CA.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., Anderson, S.E., 2010. Investigating the Links to Improved Student Learning: Final Report of Research Findings. The Wallace Foundation.
- McBeath, J., Dempster, N. (Eds.), 2009. Connecting Leadership and Learning: Principles for Practice. Routledge, Abingdon.
- Mangin, M.M., Stoelinga, S.R., 2008. Effective Teacher Leadership: Using Research to Inform and Reform. Teachers College Press, New York, NY.
- Mangin, M.M., 2007. Facilitating elementary principals' support for instructional teacher leadership. *Educ. Adm. Q.* 43, 319–357. <https://doi.org/10.1177/0013161X07299438>.
- Margolis, J., 2012. Hybrid teacher leaders and the new professional development ecology. *Prof. Dev. Educ.* 38, 291–315.
- Martineau, J.W., 2004. Evaluating the impact of leader development. In: McCauley, C.D., Van Velsor, E. (Eds.), *The Center for Creative Leadership Handbook of Leadership Development*. Jossey-Bass, San Francisco, CA, pp. 234–267.
- Mayrowetz, D., Murphy, J., Louis, K.S., Smylie, M.A., 2007. Distributed leadership as work redesign: retrofitting the job characteristics model. *Leader. Pol. Sch.* 6, 69–101.
- McCauley, C.D., 2008. Leader Development: A Review of Research. Center for Creative Leadership, Greensboro, NC.
- MetLife, 2013. The MetLife Survey of the American Teacher: Challenges for School Leadership.
- Muijs, D., Harris, A., 2006. Teacher-led school improvement: teacher leadership in the UK. *Teach. Teach. Educ.* 22, 961–972. <https://doi.org/10.1016/j.tate.2006.04.010>.
- Muijs, D., Harris, A., 2007. Teacher leadership in (in)action: three case studies of contrasting schools. *Educ. Manag. Adm. Leader.* 35, 111–134.
- Mumford, M.D., Hunter, S.T., Eubanks, D.L., Bedell, K.E., Murphy, S.T., 2007. Developing leaders for creative efforts: a domain-based approach to leadership development. *Hum. Resour. Manag. Rev.* 17, 402–417.
- Murphy, J., Smylie, M.A., Mayrowetz, D., Louis, K.S., 2009. The role of the principal in fostering the development of distributed leadership. *Sch. Leader. Manag.* 29, 181–214.
- Murphy, J., 2005. Connecting Teacher Leadership and School Improvement. Corwin Press, Thousand Oaks, CA.
- Neumerski, C.M., 2012. Rethinking instructional leadership, a review: what do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educ. Adm. Q.* 49, 310–347. <https://doi.org/10.1177/0013161X12456700>.
- Pennington, K., 2013. The Landscape of Today's Teachers Shaping Policy. The Center for American Progress, Washington, DC.
- Popper, M., 2005. Main principles and practices in leader development. *Leader. Organ. Dev. J.* 26, 62–75.
- Pounder, J.S., 2006. Transformational classroom leadership: the fourth wave of teacher leadership? *Educ. Manag. Adm. Leader.* 34, 533–545.
- Rockoff, J.E., 2004. The impact of individual teachers on student achievement: evidence from panel data. *Am. Econ. Rev. Pap. Proc.* 94, 247–252.
- Ronfeldt, M., Farmer, S.O., McQueen, K., Grissom, J.A., 2015. Teacher collaboration in instructional teams and student achievement. *Am. Educ. Res. J.* 52, 475–514.
- Schein, E.H., Schein, P.A., 2018. Humble Leadership: The Power of Relationships, Openness, and Trust. Berrett-Koehler Publishers, Inc., Oakland, CA.
- Schott, C., van Roekel, H., Tummers, L.G., 2020. Teacher leadership: a systematic review, methodological quality assessment and conceptual framework. *Educ. Res. Rev.* 31, 100352. <https://doi.org/10.1016/j.edurev.2020.100352>.
- Skrkt, T.M., 1991. The special education paradox: equity as the way to excellence. *Harv. Educ. Rev.* 61, 148–206.
- Smylie, M.A., Brownlee-Conyers, J., 1992. Teacher leaders and their principals: exploring the development of new working relationships. *Educ. Adm. Q.* 28, 150–184.
- Smylie, M.A., Denny, J.W., 1990. Teacher leadership: tensions and ambiguities in organizational perspective. *Educ. Adm. Q.* 26, 235–259.
- Smylie, M.A., Eckert, J., 2018. Beyond superheroes and advocacy: the pathway of teacher leadership development. *Educ. Manag. Adm. Leader.* 46, 556–577. <https://doi.org/10.1177/1741143217684893>.
- Smylie, M.A., Mayrowetz, D., Murphy, J., Louis, K.S., 2007. Trust and the development of distributed leadership. *J. Sch. Leader.* 17, 469–503.
- Smylie, M.A., 1994. Redesigning teachers' work: connections to the classroom. In: Darling-Hammond, L. (Ed.), *Review in Research in Education*. American Educational Research Association, Washington, D.C., pp. 129–177.
- Smylie, M.A., 2010. Continuous School Improvement. Corwin Press.
- Spillane, J.P., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30, 23–38.
- Spillane, J.P., 2006. Distributed Leadership. Jossey-Bass, San Francisco, CA.
- Supovitz, J.A., 2008. Instructional influence in American high schools. In: Mangin, M.M., Stoelinga, S.R. (Eds.), *Effective Teacher Leadership: Using Research to Inform and Reform*. Teachers College Press, New York, NY, pp. 142–162.
- Tschannen-Moran, M.T., 2004. Trust Matters: Leadership for Successful Schools. Jossey-Bass, San Francisco, CA.
- Tushman, M.L., O'Reilly, C.A., 1996. Ambidextrous organizations: managing evolutionary and revolutionary change. *Calif. Manag. Rev.* 38, 8–30.
- Van Velsor, E., McCauley, C.D., 2004. Our View of Leadership Development. Jossey-Bass, San Francisco, CA.
- Van Velsor, E., McCauley, C.D., Ruderman, M.N. (Eds.), 2010. *The Center for Creative Leadership Handbook of Leadership Development*, third ed. Jossey-Bass, San Francisco, CA.
- Wenner, J.A., Campbell, T., 2017. The theoretical and empirical basis of teacher leadership: a review of the literature. *Rev. Educ. Res.* 87, 134–171. <https://doi.org/10.3102/0034654316653478>.
- York-Barr, J., Duke, D., 2004. What do we know about teacher leadership? Findings from two decades of scholarship. *Rev. Educ. Res.* 74, 255–316.
- Yukl, G., 2013. *Leadership in Organizations*, eighth ed. Pearson, Boston, MA.

Organizing schools for improvement: the role of school decision making and authority

Thomas F. Luschei^a and Dong Wook Jeong^b, ^aSchool of Educational Studies, Claremont Graduate University, Claremont, CA, United States; and ^bDepartment of Education, Seoul National University, Seoul, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Introduction	203
Conceptual framework: a fixed or growing pie of power?	204
Power-sharing and trust in schools	204
Leadership style and distributed leadership	204
Empirical evidence	204
The need for an international perspective	205
Principal and teacher decision making: who should do what?	205
International trends in school decision making	205
School decision making and student achievement	206
Conclusion and recommendations	207
Summary	207
Recommendations for organizing decision making for school improvement	207
References	207

Introduction

School leaders are paramount in leading and sustaining school improvement efforts. In addition to managing day-to-day educational operations, school leaders represent their schools in their local communities, advocate for students, and supervise and support teachers. Most importantly, effective school leaders can support the creation of school environments that encourage excellent teaching, ensure consistent learning, and promote student engagement.

The importance of school principals is evident in the school reform and improvement literature. Drawing on research from 15 years of school improvement efforts in Chicago elementary schools, Bryk (2010) identified five “essential supports” for school improvement: (1) a coherent instruction guidance system, (2) professional capacity, (3) strong parent-community-school ties, (4) student-centered learning climate, and (5) leadership drives change. In addition to the key role that school leaders play in each of the first four essential supports, the fifth support exemplifies the central role of school principals in advancing instructional improvement efforts and uniting the entire school community around a common vision of improvement. Bryk (2010) argues that as they lead the process of school improvement, “principals cultivate a growing cadre of leaders (teachers, parents, and community members) who can help expand the reach of this work and share overall responsibility for improvement” (p. 25). In other words, school principals must simultaneously lead and cede, shouldering key responsibilities while surrendering some of their authority to other school stakeholders, especially teachers.

Although literature on “distributed leadership” illustrates the potential benefits of sharing leadership responsibilities among various school actors (e.g., Spillane, 2012), power sharing within schools also introduces the potential for conflict and misunderstanding about who should take on which roles. The sharing of responsibilities within schools also raises several important questions about school governance and leadership styles, such as whether the pie of school authority is fixed or capable of growing; who is better positioned to execute which roles; and how much supervision, support, or autonomy school principals should offer to teachers.

The importance of these questions has grown as many school systems across the globe decentralized in the 1990s and early 2000s, leading to an expansion of school autonomy and a greater role for school-level actors in school decision making (Jeong and Luschei, 2018). At the same time, some education systems (such as in the United States) have concentrated more educational decision making in the hands of the government, while other education systems have simultaneously devolved authority to schools while weakening decision making authority of teachers over their own instructional decisions (Jeong and Luschei, 2018). These diverging international trends illustrate the importance of considering the political and governance context to understand the manner and extent to which responsibilities are shared or contested within schools. For example, centralized education systems may concentrate decision making in supra-school entities like district or municipal offices, or even ministries of education. In contrast, decentralized systems are more likely to cede decision making to schools, in which case power may be concentrated in the hands of school principals or distributed across principals, teachers, or other actors. Yet since education systems generally do not vary within countries, it is difficult to examine the interaction between education system context and school-level decision making in a given country. As a result, an international comparative perspective is essential to understand how different configurations of school-level decision making relate to key outcomes like student achievement.

In this article, we apply an international comparative perspective to explore important questions about the role of principal and teacher authority in supporting school improvement and student achievement. By reviewing recent evidence—including our own published work—we aim to distill key lessons for how schools can organize and distribute authority appropriately and effectively across key school stakeholders, with a focus on school principals and teachers.

To explore the conceptual and empirical foundation for school-level power sharing, our conceptual framework lays out the competing perspectives of zero-sum versus variable-sum conceptions of school decision making authority. We then describe recent evidence regarding international trends in school governance and principal/teacher authority, distributed leadership, and ultimately, what the evidence tells us about how school governance may impact student achievement. We conclude with a set of concrete recommendations regarding the organization of school decision making and authority for school improvement.

Conceptual framework: a fixed or growing pie of power?

The question of whether and how authority is shared across school actors rests on competing ideas of power within schools as either fixed or variable. If power is fixed, then decision making authority cannot be allocated to one actor without being stripped from another. That is, power follows a “zero-sum” arrangement in which school principals cannot allocate authority to teachers without losing some of their own. By contrast, in variable-sum arrangements, school principals can cede decision making control without losing their own authority, resulting in a “win-win” relationship that increases the overall size of the “pie” of school-level power (Jeong and Luschei, 2018).

Power-sharing and trust in schools

Of course, the nature of power-sharing agreements undoubtedly varies across schools based on key conditions like principal leadership style and the capacity of school personnel. Perhaps most importantly, school leaders must trust their teachers to use their power to support the school’s vision of instructional improvement and raise student outcomes. Similarly, teachers must trust that principals will support their exercise of power and autonomy in the classroom and avoid undercutting new innovations and strategies.

Bryk (2010) describes the important role of school principals in establishing and nurturing trust in urban school communities, where power is often not shared across principals, teachers, and parents:

Principals establish both respect and personal regard when they acknowledge the vulnerabilities of others, actively listen to their concerns, and eschew arbitrary actions. If principals couple this empathy with a compelling school vision, and if teachers see their behavior as advancing this vision, their personal integrity is also affirmed. Then, assuming principals are competent at managing routine school affairs, an overall ethos conducive to building trust is likely to emerge (p. 27).

Leadership style and distributed leadership

In addition to competence, school principals’ leadership style will also influence how well they are able to engender trust among teachers to enable greater distribution of decision making authority across school actors. One key aspect of leadership style is whether school leaders ensure that their relationships with teachers are collaborative and supportive. A collaborative approach not only empowers teachers to take on greater decision making roles, but it can also promote greater effectiveness, creativity, and innovation on the part of teachers (Somech, 2010).

Collaborative leadership approaches can also lead to the sharing of responsibilities and roles through distributed approaches that empower teachers and other school actors to lead key aspects of school improvement (Spillane, 2012). From the perspective of the school leader, a central component of distributed leadership is their openness to sharing leadership functions with other school actors (Liu, 2020). Although distributed leadership implies the exercise of power by multiple actors, the approval and involvement of the school principal is an essential element in the presence and success of distributed leadership approaches (Printy and Liu, 2021). In other words, decision making authority cannot be shared across school actors if the school principal does not believe in or authorize collaborative approaches.

Empirical evidence

The positive potential of sharing decision making within schools is supported by a small but important set of empirical studies. In examining whether school power follows a zero- or variable-sum power-sharing arrangement, Shen and Xia’s (2012) analysis of the 1999–2000 Schools and Staffing Survey (SASS) found that principal and teacher authority are positively related across six of seven decision making dimensions, implying a variable-sum arrangement and an overall increase in school-level authority. However, Shen and Xia also found that such win-win relationships are more likely to occur at the elementary rather than secondary level. In a study of the 2004 SASS, Xia (2014) found positive relationships among principal and teacher decision making authority in five of seven decision making domains. Together, these results suggest that principal and teacher power sharing depends both on the context of the school and the decision making domain involved.

Context and decision making domain also appear to be important for the effectiveness of distributed leadership models, which by extending participation and decision making across multiple actors within a school site, assume the existence of variable-sum power-sharing arrangements. For example, [Printy and Liu's \(2021\)](#) analysis of data from the 2013 Teaching and Learning International Survey (TALIS) found that the distribution of school leadership is practiced and understood differently across countries. Printy and Liu also found that the relationship between principal and teacher leadership reports are positive and significant for some leadership functions, such as managing instruction, but negative for others, most notably organizational decision making. According to [Printy and Liu \(2021\)](#), these findings suggest that "principals think differently about who participates in leadership activity according to the situation." Specifically, whereas teachers are more likely to engage in activities and decisions related to instruction, teacher engagement is less likely in general organization decision making, where "principals most often retain hierarchical authority" (p. 316).

The need for an international perspective

Limited empirical evidence suggests that principals can share power with teachers without compromising their own decision making authority. One key avenue for doing so is the use of distributed leadership approaches that extend responsibilities and authority across many school actors. In fact, empirical evidence supports the exercise of distributed leadership for important school functions and outcomes. According to [Liu \(2020\)](#), distributed leadership is positively associated with school-level academic capacity, teachers' pedagogical skill and knowledge, teacher and principal motivation, perceptions of school cohesion, and teacher motivation. However, effective sharing or distribution of leadership responsibilities appears to be sensitive to the types and domains of decision making undertaken and the contexts under which leadership is practiced. Here the literature suggests the importance of cross-country differences, especially in terms of educational governance arrangements between governments and schools and within schools ([Printy and Liu, 2021](#)). In particular, the level of centralization or decentralization within an education system is likely to play an important role in how power is shared, and whether school principals act as agents of the government or as partners with teachers ([Jeong and Luschei, 2018](#)). Since educational governance generally varies little within a given country, a cross-country perspective is essential to understand relationships among educational governance, the sharing of authority within schools, and teachers' authority over their important classroom decisions. In the section that follows, we discuss evidence related to each of these importance pieces and how they relate to the central educational outcome of student achievement.

Principal and teacher decision making: who should do what?

International trends in school decision making

The 1990s and early 2000s brought a wave of educational decentralization across much of the globe, under the theory that local actors could solve educational problems more effectively than distant government bureaucrats. In many cases, this decentralizing trend brought new decision making authority to teachers, not just in the classroom, but also over issues such as school budgeting, staffing, and planning ([Eurydice, 2008](#)). This global trend was best exemplified by Finland, where teachers were empowered by a "culture of trust" to act as autonomous, well-trained professionals in their schools and classrooms ([Sahlberg, 2007](#)). Finland's stellar performance on the Program for International Student Assessment (PISA) and other international assessments provided evidence that investments in teacher preparation and autonomy represented a powerful way to raise student achievement. Finland's success was also cited to support growing calls for teacher professionalization, whose advocates argue that as well-trained professionals, teachers should be empowered to make key curricular and instructional decisions in their schools and classrooms ([Darling-Hammond, 2017](#); [Jeong and Luschei, 2018](#)).

Finland's performance notwithstanding, many other countries took a different approach toward school governance in the 1990s and 2000s, either centralizing overall school governance or decentralizing governance while concentrating school decision making in the hands of school principals and other non-teacher actors. In an analysis of educational decision making in 33 countries that participated in the 2000 and 2015 applications of PISA, [Jeong and Luschei \(2018\)](#) found that from 2000 to 2015, most of these countries actually centralized decision making in the hands of governments. Further, this centralization occurred not just in the domains of budgeting and staffing, but also in curriculum and instruction. In fact, 27 of 33 countries moved decision making authority over curriculum and instruction away from teachers and toward governments during this period. In contrast, school principals gained decision making authority over curriculum and instruction during the same period.

Surprisingly, [Jeong and Luschei \(2018\)](#) found that Finland, long identified with teacher professionalization and autonomy, was one of the 27 countries to centralize curricular and instructional decision making between 2000 and 2015. Not surprisingly, one of the countries with the largest moves toward centralization during this period was the United States. This period corresponded to the US No Child Left Behind (NCLB) reform, which increased the role of federal and state governments in standards-based accountability and oversight of school- and classroom-based decision making. Consequently, among the 33 countries in [Jeong and Luschei's \(2018\)](#) analysis, the United States was among the four countries with the highest level of overall centralization and was second only to Greece in terms of centralization of decisions related to curriculum and instruction.

[Jeong and Luschei's \(2018\)](#) study includes three important lessons for understanding the role of school decision making and the sharing of authority within schools:

1. Apparent global trends toward decentralization are actually much more nuanced than they appear, as decentralization can look very different within each country. For this reason, it is important to examine how decision making is shared *within schools*.
2. Despite global calls for teacher professionalization and autonomy, teachers in many countries have lost decision making authority over what happens in their own classrooms. This may be explained in part by the countervailing force of school accountability policies, best exemplified by NCLB in the United States.
3. Educational centralization and loss of decision making authority on the part of teachers does not imply a similar loss for school principals. In fact, during the same time that teachers lost decision making authority over curriculum and instruction, school principals gained authority in this domain.

School decision making and student achievement

Teachers' loss of authority over curriculum and instruction defies the logic that teachers know best how to address their students' needs and that they should have the autonomy to make key decisions regarding curriculum and instruction. In fact, research related to teacher autonomy indicates that teachers with greater autonomy have greater job satisfaction, motivation, and retention, which are all likely to be related to their effectiveness with students (Ingersoll, 2006; Ingersoll and May 2012; Jeong and Luschei, 2018). Consequently, the trends identified by Jeong and Luschei (2018) raise concern regarding the impact of teachers' loss of decision making authority on student achievement. More generally, the impact of teacher and principal authority on student achievement represents one of the most important illustrations of why policy makers and education officials should care about how decision making authority is shared within schools.

Although a fairly robust research literature has examined the impact of educational decentralization on student achievement, this research tends to look at school-level decision making as monolithic. That is, studies often examine the relative influence of government-versus school-level decision making across multiple dimensions or contexts on student achievement. For example, Fuchs and Woessmann (2007) found that greater school autonomy in the areas of textbook choice, hiring of teachers, and school budgeting is positively related to student achievement. Hanushek et al. (2013) found that decentralization has a more positive impact in more developed and higher capacity environments. However, this research does not examine how power is distributed across principals and teachers within schools. Yet as our conceptual framework emphasizes, power-sharing within schools can look very different across contexts, and such differences may be important to understand the impact of power sharing on student achievement.

In one of the few empirical studies to examine the relative importance of principal and teacher decision making authority for student achievement, Luschei and Jeong (2020) examined data from the 2015 PISA to assess whether perceived governance arrangements are related to students' PISA test scores. The authors found that increases in teacher decision making authority are consistently and positively related to student test scores in math, reading, and science. Further, this positive result tends to be greater in the domain of curriculum and instruction. In contrast to these positive results for teacher decision making, Luschei and Jeong (2020) found that decision making responsibilities of school principals are not significantly related to student achievement in any subject. When examining these relationships across different stakeholders and domains of decision making, the authors found that greater responsibility for school budgeting among school boards is consistently and positively related to student achievement in all subjects. However, greater decision making responsibility on the part of governments has a consistently negative or nonsignificant relationship with student achievement across all subject areas and decision making dimensions, suggesting a negative impact of educational centralization.

In addition to evidence supporting the positive benefits of sharing decision making authority with teachers, three other results from Luschei and Jeong's (2020) analysis deserve mention:

1. To examine the role of teacher capacity in decision making, the authors conducted additional analyses in which they categorized schools as high or low capacity according to the average level of teacher education in a school. They found that the positive relationship between teacher responsibilities and student achievement holds only in schools with high-capacity teachers. In schools with less educated teachers, the relationship between teacher responsibility and student achievement is not statistically significant. In other words, the positive results for teacher decision making appear to be largely driven by higher-capacity teachers. Additionally, the relationship between teacher decision making and student achievement is strongest for decision making related to curriculum and instruction. These results demonstrate once again that the impact of decentralizing school responsibilities or sharing power with teachers is not monolithic. Importantly, school leaders' decisions to share power with teachers must be informed by contextual considerations like the capacity of teachers and other school stakeholders, as higher-capacity schools are more likely to leverage greater autonomy to positively affect important outcomes like student achievement.
2. Luschei and Jeong (2020) also examined relationships among decision making authority and achievement across countries that were members of the OECD (representing wealthier countries) and those that were not OECD members. They found that the positive relationship between teacher decision making and student achievement is only significant in non-OECD countries. The impact of decision making by school boards also appears to be more positive in non-OECD countries. Although these results defy easy interpretation, they do suggest important cross-country differences in the impact of decision making arrangements. They also highlight a potentially important but understudied topic in educational governance, the role and impact of school boards on student achievement.

3. To understand the role of principals' leadership style, the authors also included models that controlled for various aspects of principal leadership style, as measured by PISA questionnaires that were filled out by school leaders. These variables include principals' responses related to (1) communication of school goals and curricular development, (2) instructional leadership, (3) promotion of instructional improvements and professional development, and (4) teacher participation in leadership. The analysis found only one of these variables to be significantly related to student achievement. This variable, which represents the degree to which principals are involved in teachers' classroom level decision making and problem solving, was negatively related to student achievement in all subjects. Additionally, the inclusion of principal leadership variables had little impact on the positive and significant relationship between teacher decision making and student achievement, which suggests that teacher decision making can have a positive impact on student learning regardless of principals' self-reported leadership style.

Conclusion and recommendations

Summary

In summarizing their review of literature related to the impact of school principals on students and schools, [Grissom et al. \(2021\)](#) observed, "it is difficult to envision an investment with a higher ceiling on its potential return than a successful effort to improve principal leadership" (p. 43). In this article, we argue that a key path to ensuring the effectiveness of principals is through the strategic sharing of decision making authority with teachers. We point to compelling evidence across education systems, subject matters, and domains of decision making that greater decision making responsibilities for teachers are consistently and positively related to student achievement. However, there are some caveats to this conclusion. First, we identified some evidence that the impact of sharing decision making with teachers varies according to national context, although this question needs further cross-national research. Second, the positive impact of teacher decision making is most apparent in the domain of curriculum and instruction, which is most directly aligned with the day-to-day "instructional core" of teachers' work. Finally, the positive impact of teacher decision making appears to be largely driven by more educated, or higher capacity, teachers. Consequently, principals must consider teacher capacity when deciding which responsibilities to devolve to whom.

Recommendations for organizing decision making for school improvement

Our review of conceptual and empirical literature leads us to make three tentative recommendations for organizing school-level decision making for school improvement.

1. Both school leaders and teachers must recognize that the sharing of decision making authority does not have to be a zero-sum game. The literature cited above offers evidence that decision making responsibilities for principals can grow along with those of teachers. The ample literature on distributed leadership also provides robust evidence for the positive potential of sharing leadership tasks within schools. However, the extent to which power can be shared appears to vary by context and according to decision making dimension, so principals and teachers must engage in power-sharing carefully, collaboratively, and with an eye toward how these arrangements may change as school contexts and conditions change.
2. To make collaborative decision making work, principals must build trust among teachers and other school stakeholders. School-level trust is the *sine qua non* of any collaborative school improvement effort. Teachers must trust that principals will recognize and support their professionalism, especially in the core areas of curriculum and instruction. Similarly, principals must trust that teachers will use their autonomy and authority to support a school-wide vision for improvement to make positive instructional changes for students.
3. Principals must share decision making authority strategically, considering both the type of decisions teachers are willing and able to make and the capacity of teachers to do so. Decisions related to curriculum and instruction are closest to teachers' training and day-to-day work, and they best encapsulate teachers' knowledge and skills as professionals. We have also presented evidence that among all decision making dimensions, teachers are most likely to use decisions related to curriculum and instruction to improve student achievement. On the other hand, inexperienced or less educated teachers may have less skill to leverage their authority or autonomy to improve instruction. In such cases, principals may need to supplement autonomy with support and professional development. As [Luschei and Jeong \(2020\)](#) argue, "The lesson for school leaders is to recognize which type of teaching staff they have and provide support or autonomy accordingly" (p. 363).

References

- Bryk, A.S., 2010. Organizing schools for improvement. *Phi Delta Kappan* 91 (7), 23–30.
- Darling-Hammond, L., 2017. Teacher education around the world: what can we learn from international practice? *Eur. J. Teach. Educ.* 40 (3), 291–309.
- Eurydice, 2008. Levels of Autonomy and Responsibilities of Teachers in Europe. Eurydice European Unit.
- Fuchs, T., Woessmann, L., 2007. What accounts for international differences in student performance? A re-examination using PISA data. *Empir. Econ.* 32 (2–3), 433–464.
- Grissom, J.A., Egalite, A.J., Lindsay, C.A., 2021. How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research. The Wallace Foundation, New York.

- Hanushek, E.A., Link, S., Woessmann, L., 2013. Does school autonomy make sense everywhere? Panel estimates from PISA. *J. Dev. Econ.* 104, 212–232. <https://doi.org/10.1016/j.jdeveco.2012.08.002>.
- Ingersoll, R., May, H., 2012. The magnitude, destinations, and determinants of mathematics and science teacher turnover. *Educ. Eval. Pol. Anal.* 34 (4), 435–464.
- Ingersoll, R., 2006. *Who Controls Teachers' Work? Power and Accountability in America's Schools*. Harvard University Press, Cambridge, MA.
- Jeong, D.W., Luschei, T.F., 2018. Are teachers losing control of the classroom? Global changes in school governance and teacher responsibilities, 2000–2015. *Int. J. Educ. Dev.* 62, 289–301.
- Liu, Y., 2020. Focusing on the practice of distributed leadership: the international evidence from the 2013 TALIS. *Educ. Adm. Q.* 56 (5), 779–818.
- Luschei, T.F., Jeong, D.W., 2020. School governance and student achievement: cross-national evidence from the 2015 PISA. *Educ. Adm. Q.* 57 (3), 331–371.
- Printy, S., Liu, Y., 2021. Distributed leadership globally: the interactive nature of principal and teacher leadership in 32 countries. *Educ. Adm. Q.* 57 (2), 290–325.
- Sahlberg, P., 2007. Education policies for raising student learning: the Finnish approach. *J. Educ. Pol.* 22 (2), 147–171.
- Shen, J., Xia, J., 2012. The relationship between teachers' and principals' decision-making power: is it a win-win situation or a zero-sum game? *Int. J. Leader. Educ.* 15 (2), 153–174.
- Somech, A., 2010. Participative decision making in schools: a mediating moderating analytical framework for understanding school and teacher outcomes. *Educ. Adm. Q.* 46 (2), 174–209.
- Spillane, J.P., 2012. *Distributed Leadership*, vol. 4. John Wiley & Sons.
- Xia, J., 2014. *Principal-Teacher's Decision-Making Power Relationship: A National Study Based on SASS 2003–04 Data* [Unpublished doctoral dissertation]. Western Michigan University.

Leadership and governance for improvement: how schools manage external demands

Thomas Hatch^a, Elizabeth Leisy Stosich^b, and Jordan Corson^c, ^aTeachers College, Columbia University, New York, NY, United States; ^bFordham University, New York, NY, United States; and ^cSchool of Education, Stockton University, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

School capacity: what does it take to improve student learning?	209
Strategically managing external demands	210
Illustrations: managing external demands in schools, networks, and systems	211
Managing external demands in schools: integration or adoption?	211
Leveraging networks to sustain attention to internal goals in the face of external change	212
Creating a context for coherence: managing external demands at the system level	212
Summary	213
References	214

The ability to make constant improvements and adapt to a changing environment enables organizations not only to survive but also to increase their effectiveness over time (Collins, 2001; March, 1991). This focus on how organizations adapt to and work with their environments highlights the critical role that managing external demands plays in organizational development. From this perspective, the extent to which organizations are able to improve their performance depends on more than their internal characteristics; improvement depends on the relationship between the organization and the many other groups, organizations and institutions in the surrounding environment. By strategically managing their relationship with these other entities, organizations like schools can get access to information, resources and expertise and develop networks of connections that help them create coherent approaches to improvement that advance their organizational goals (Honig and Hatch, 2004). In this article, we describe one theoretical framework on capacity building in schools that helps to explain why managing demands is so important for organizational development; outline some key strategies education leaders can use to manage external demands; and provide a series of examples to explain the role that managing demands can play in school improvement efforts.

School capacity: what does it take to improve student learning?

Increasing access to and quality of schooling remains a central concern of societies around the world. In many cases, the hope is that enabling schools to be more effective will fuel a whole variety of benefits, including individual success, economic development, civic participation, and quality of life. Correspondingly, improvement efforts often focus on providing a series of “inputs” to schools – funding, curricula, qualified teachers – with the expectation that this infusion of resources will provide schools with the capacity to reach their goals. In a sense, this approach conceptualizes schools as containers that need to be filled: once schools have adequate resources, they should be able to reach their goals. This is a linear conception that underlies policy interventions like the development of Title 1 Elementary and Secondary Schools Act in 1965 and other “compensatory” programs in the United States around that time. These programs sought to achieve goals like improving schools and eliminating poverty by providing additional funds, curricula, or other programs to schools, particularly for schools in communities with lower levels of income and limited access to other resources.

By the 1980’s, however, the concerns about the continuing challenges of improving schools contributed to calls in the United States to rethink what it might take to improve schools and to go beyond providing more resources and to “restructure” schools entirely so that they could use those resources more effectively (Carnegie Forum on Education and the Economy, 1986; Elmore, 2004). At the same time, attention began to expand beyond schools themselves to consider the ways in which changes in many different aspects of education systems needed to be made (see for example, Smith and O’Day, 1990). The subsequent efforts to promote “systemic reform” in the United States also began to draw heavily on comparisons with countries that were seen as “higher-performing” on international tests like TIMMS and eventually the PISA tests that came into wider use in the 1990’s.

These shifts in some of the most prominent school improvement efforts at the time reflected a new and emerging view of school capacity that went beyond the typical focus on providing schools with additional resources (Cohen and Ball, 1999). First, this emerging view highlighted that simply having resources is not sufficient (Corcoran and Goertz, 1995; Malen and Rice, 2004). As Malen and Rice express it, a schools’ capacity to meet its goals depends on turning resources into expected outcomes (p. 636). In turn, education reformers began to highlight that maximizing the use of resources depends on the relationships among technical capital (funding, facilities, materials), the human capital (knowledge and expertise) needed to use those materials effectively, and social capital (networks of relationships and trust) that provides access to and facilitates the use of materials and the development of expertise (Cohen and Ball, 1999; Corcoran and Goertz, 1995; O’Day et al., 1995; Newmann et al., 2000; Putnam, 2000; Spillane et al., 2003; Spillane and Thompson, 1997).

Second, instead of treating schools as containers waiting to be filled, this view defined capacity as the “potential of material, a product person, or group, to fulfill a function if it is used in a particular way” (Newmann et al., 2000, p. 261). In this case, a school’s capacity to reach a particular goal depends both on the resources available and the complexity and magnitude of the work needed to accomplish that function (Hatch, 2001). For example, schools have low capacity if they have to get access to new materials, develop more expertise, or make substantial changes in their operations in order to enable all their students to reach high standards of achievement; but they have high capacity if few changes or additional investments are required to meet that goal. At the same time, the evaluations of the many standards-based and systemic reform efforts at the time, also highlighted changes in the structures, practices, and people in other aspects of the system that also needed to be made in order for schools to develop the capacity to meet these ambitious goals (Goertz et al., 1996; Fullan, 1999; O’Day and Smith, 1993).

An analogy to the evolution of municipal water systems helps to illustrate this new relational view of capacity (Hatch, 2009). In the early 1900’s urban centers needed a much larger supply of water. Initially, efforts focused on increasing the supply of water – by building more and bigger dams and reservoirs and expanding the amount of water they could deliver. The economic and ecological costs, however, quickly demonstrated the limits of that approach. Consequently, municipalities began to turn their attention to managing the *demand* for water: developing appliances that used less water and educating the public about the benefits of conserving water. Controlling demand lessened the pressures to increase the supply.

This analogy reveals that the capacity to meet particular goals – whether to provide 24/7 access to water or to increase student learning outcomes – is a function of the resources available and the demands that exist. Thus, increasing supply and the amount of resources available is only one means of building capacity. Monitoring and managing the demands on the system plays a critical role in building capacity as well. Returning to the analogy also illustrates that the extent to which the resources available meet the current demands for water can also be affected by a host of contextual factors including weather patterns, technological developments, and demographic changes. Furthermore, managing the water supply not only involves factoring in the external conditions, it also requires acting systematically – considering the connections among different neighborhoods and different regions and the way their interactions affect supply and demand.

In terms of schools, several lessons stand out. First, building the capacity to improve student learning depends on the relationship between the school and the surrounding community: changing the demands placed on schools can change the resources and the amount of effort schools need to achieve their goals. Second, schools face all kinds of demands that they do not always have the capacity to meet (Hatch, 2002). These include demands to meet performance goals on conventional measures such as attendance, test scores, and graduation rates. However, schools are also expected to fulfill a variety of other implicit and explicit functions such as housing and looking after school-age children for a good portion of the work day. Schools also face a wide variety of demands for compliance and conformity from the district, the state, support providers, students, families and others who expect them to implement certain practices and operate in certain ways. In particular, all schools – and in fact all educational entities – have to contend with the general expectations about what “real school” and “real learning” look like (Hatch, 2021; Metz, 1988; Tyack and Cuban, 1995). As a consequence, schools face multiple, often conflicting demands that can sap time and energy and may undermine rather than enhance their ability to achieve their performance goals.

Strategically managing external demands

Despite the challenges, schools can strategically manage external demands. In a review of literature related to policy coherence and organizational learning, Honig and Hatch (2004) identify a number of approaches schools can utilize to recognize and deal with external demands. First, goal setting creates a framework that can help members of an organization to understand what they are supposed to do and to identify external resources and opportunities that can help them meet those goals. This framework serves as a “simplification system” that enables organization members to make sense of complex and conflicting information and to make complex problems manageable (March, 1994; Weick, 1995). Without some common understanding of their purposes and goals, organization members have little direction to help them make decisions or to decide what aspects of the external environment to pay attention to. In particular, educators benefit from having a shared understanding of goals that are both specific and open-ended: providing some direction but also allowing for diverse views and the development of new information and expertise that can lead to productive adaptations.

Second, organizations can take advantage of two general approaches to help them manage external demands and use them to advance their collective goals: bridging and buffering. Bridging involves engaging with external demands in order to gain information, resources, and assistance. Bridging strategies include “pulling the environment in” by working with and incorporating people and groups in their external environment into their own activities. These efforts blur the boundaries between the organization and the external environment, fostering the development of reciprocal relationships, facilitating the sharing of information, expertise, and resources, and supporting mutual adaptation. Developing these relationships also gives organizations opportunities to “shape the terms of compliance” by influencing the demands that they have to face. For example, organizations that develop connections to policymakers by including them on their boards and among their advisors can both gain information about policies in development and can provide information and advocacy that encourages policymakers to develop initiatives that reflect the organization’s goals and interests.

Conversely, buffering involves deliberate efforts to deflect or ignore external demands. Buffering strategies include attempting to “suspend ties to the environment” as much as possible. For example, schools may choose to apply for magnet or charter status or to operate as private schools in order to avoid having to respond to the demands of the public education system. Public schools may

also choose to avoid some external demands by seeking waivers that allow them to avoid demands (such as those associated with particular evaluation systems), choosing not participate in particular programs or refusing funds that bring additional demands. Notably, however, it is extremely difficult for schools to be able to eliminate all external demands. Even schools that operate outside the regulations of public education systems still have to contend with the expectations and requirements of students, parents, and other institutions with whom they interact.

Despite their differences, bridging and buffering are not mutually exclusive strategies. Organizations can use strategies that reflect both approaches at the same time – connecting to some external groups and some demands while attempting to ignore or protect themselves from others. Several strategies reflect a “hybrid” approach that involves both bridging and buffering simultaneously. For example, developing “peripheral structures” enables organizations to deal with some external demands while insulating other units and members of the organization from those demands (Elmore and McLaughlin, 1988) (in other words, partially suspending ties with their environment). Organizations can also symbolically respond to or adopt external demands (Cuban, 1984; DiMaggio and Powell, 1983; Meyer and Rowan, 1977; Zajac and Westphal, 1994). For example, a school might develop a strategic plan that incorporates goals and activities required by external authorities, but then may not implement that plan; or organizations may adopt language, terms and structures that make it sound and look like the organization is responding, without actually altering any of their core operations.

The literature on how organizations deal with external demands suggests several key lessons for education leaders. First, members of organizations respond to their environment by bridging and buffering all the time. However, educators are better able to strategically manage those demands when members of the organization have a shared understanding of their goals and work collectively to use bridging and buffering strategies explicitly to advance their goals. Second, members of organizations, the units within them, and the other organizations they are connected to, are also all attempting to manage external demands at the same time. Therefore, members of other organizations like staff of district and municipal offices, external support providers, and state agencies can help schools to manage external demands or they can contribute to the demands to which schools have to respond (or both). As a consequence, developing and maintaining relationships with key actors in the external environment remains a crucial ingredient for managing external demands. In addition, deliberate efforts by education leaders, policy makers, and funders to support common goals, shared understanding, coordinated activity, and collective impact across organizations can create an environment that makes managing external demands easier rather than harder (Kania and Kramer, 2011).

Illustrations: managing external demands in schools, networks, and systems

In the sections that follow, we review research on how members of organizations manage external demands at the level of the school, system, and network and the lessons for education.

Managing external demands in schools: integration or adoption?

In practice, capable school leaders are able to bridge to and integrate external demands in ways that reinforce and advance internal priorities. In the process, they help to craft a shared coherent, understanding of what school members are supposed to do when faced with different choices and conflicting demands (Honig and Hatch, 2004). By contrast, schools with weaker leaders are more likely to accept and adopt the goals of external policies as their internal priorities (Louis and Robinson, 2012). In the process, schools become focused on implementing the initiatives developed by others; and their members have to figure out how to deal with the competing expectations largely on their own. Given the number and short shelf life of many improvement policies and initiatives, adopting the goals of external initiatives opens schools up to considerable uncertainty and leaves them vulnerable to the constant changes in their environment.

Integrating external demands rather than adopting them wholesale depends on leaders who can foster a shared understanding of school priorities (Fullan and Quinn, 2015; Johnson et al., 2014) and can establish what Abelman et al. (1999) describe as “internal accountability” or what Forman et al. (2017) describe as a school’s level of “internal coherence,” or “the collective capability ... to connect and align resources to carry out an improvement strategy” (p. 2–3). Internal accountability and coherence develop as leaders put in place structures and processes that support the development of relationships, sharing of information and expertise, and a sense of collective responsibility (Hatch, 2020). This internal coherence then serves as a foundation for distributed leadership (Harris and Spillane, 2008) and enables many members of the school to advance the school’s interests. The effective integration of external demands is also more likely when schools have more capable instructional leaders – principals and teachers who are knowledgeable about standards, subject matter, and the leadership practices that can improve teaching and learning (Kaufman and Berglund, 2018; Stein and Nelson, 2003) – and have stronger organizational structures and resources for professional learning.

In addition to focusing on developing the structures and processes inside their schools, school leaders can manage external demands by drawing on several different strategies for working with people and groups outside their organization (Hatch, 2009). In particular, leaders can cultivate connections with key allies and share information with them about their school’s goals, needs, and operations. These connections enable leaders to “scan the environment” (Honig and Hatch, 2004) by giving them access to information about what’s happening in other organizations and what may happen in the future. This information can help leaders to anticipate changes in external demands and conditions and to prepare to respond. At the same time, these connections

enable leaders to “seed” the environment by giving them allies in the external environment who can help to develop new policies and initiatives that already reflect the needs and concerns of the school. In addition, these connections give schools access to people who can advocate on the schools’ behalf to help shape the terms of compliance and to help schools get waivers or other means of buffering from problematic demands.

In a study of the implementation of new academic standards for students and a new teacher evaluation policy in three high-poverty elementary schools in New York, for example, one principal leveraged these policies to strengthen existing curriculum and better address the needs of their students, particularly English learners (Stosich, 2018). Teachers in this school described engaging in professional development and receiving instructional feedback that supported them in learning how to use and adapt their existing curriculum to meet the new standards, the expectations for teacher evaluation, and their students’ needs. Furthermore, the school’s sustained work with a curriculum developer provided access to new materials and support for professional learning as they navigated the shifts required by the new standards while building on their ongoing efforts to strengthen instruction in English language arts.

By contrast, the other two schools in the study lacked a well-defined vision for curriculum and instruction. Notably, one school lacked a well-defined curriculum and the curriculum had changed frequently in the other school, leaving teachers in both schools without sustained opportunities for learning how to use curriculum to implement the new academic standards. In these instances, the principals adopted the practices described in the teacher evaluation framework as their school priorities. Given the limited attention to curriculum, educators in both schools focused on the pedagogical approaches outlined in the teacher evaluation framework as a way of simplifying the work of teaching to the new standards; however, this strategy failed to bring attention to how these instructional strategies would be used to address specific content and students’ learning needs. Teacher evaluation policies focus on teachers’ pedagogy, whereas academic standards focus on students’ learning; thus, responding to the demands of more student-focused policies may present a greater challenge for educators than teacher-focused policies. For these schools, the implementation of the teacher evaluation framework detracted attention from the new standards for students in the two schools with lower levels of organizational capacity.

Leveraging networks to sustain attention to internal goals in the face of external change

Beyond independent connections with different individuals and groups, school leaders can also form or join networks of organizations that share some goals and interests. Like schools, networks with shared goals and a history of success may be better able to buffer themselves from problematic policies and initiatives while maintaining their legitimacy. Research on the Internationals Network for Public Schools in the United States presents one promising example (Jaffe-Walter, 2008). The network of high schools serves newcomer youth who have been in the country for four years or less at their time of admission and are learning English and is organized based on common principles that prioritize project-based learning and collaborative inquiry for students and educators.

Since the network was founded in the 1990s, increased focus on standardization and high-stakes testing has presented a challenge for the network and member schools as they work to maintain their focus on interdisciplinary, project-based learning in the face of narrower, content-based external assessments. As Jaffe-Walter (2008) explains, “the network is critical because it provides a space for educators to explore ways of implementing core commitments in the context of external pressures” (p. 2053). This research suggests that the network is not only a place for collectively negotiating how to respond to conflicting demands while maintaining core commitments but can also be a site for reactivating a collective memory of resistance to inform responses to new demands. Notably, this network is strongly positioned to buffer external policies given the success of network schools in supporting recent immigrants and English learners in graduating from high school at much higher rates than other schools in their same school districts. Even in successful networks, however, negotiating the connections between external demands and internal priorities is an ongoing process. Thus, as policies increased the number of both formative and summative assessments required for high school graduation, educators in the study described an ongoing debate with colleagues in their school and across schools in the network about the extent to which they should allow external assessment policies to influence internal priorities. Networks can connect schools with the technical, human, and social capital that increases capacity for sustaining attention to internal goals even in the face of competing or conflicting external demands.

Creating a context for coherence: managing external demands at the system level

Reflecting the relationship between schools and the context with which they interact, school leaders’ ability to manage external demands depends as well on the extent to which district administrations, external support providers, and other organizations and government agencies contribute to or mitigate multiple, conflicting demands. Notably, district administrators can amplify conflicting demands if they respond to new policies and initiatives by adding layers of requirements to facilitate monitoring and to ensure compliance at the school level. Such amplification becomes even more problematic when municipalities and national governments create several new policies at once or when those policies are in conflict with others that are already in place. For example, in the US, the adoptions by many states of the Common Core Learning Standards and new teacher evaluation approaches required by the Federal Race to the Top legislation created separate demands for professional development and training that schools and districts had to figure out how to coordinate (Stosich et al., 2021).

Relatedly, policymakers can make it more difficult for schools to implement policies if they fail to acknowledge the need to build new capabilities to meet the new demands. According to [Elmore's \(2004\)](#) theory of reciprocal accountability: "If the formal authority of my role requires that I hold you accountable for some action or outcome, then I have an equal and complementary responsibility to ensure you have the capacity to do what I am asking you to do" (p. 68). In practice, however, schools and districts often complain that new policies include "unfunded" mandates and related demands that schools have not been given the funds and means to meet.

At the same time, district administrators and municipal and national policymakers can also work to align external demands and act in ways that help schools to coordinate those demands effectively. Alignment describes the process of lining up policy expectations, resources, and rewards or consequences so that they create a consistent and linked set of demands. In Singapore, for example, regular, systematic reviews of curriculum in each subject are centrally coordinated with work on professional development and assessment. In the US, the "systemic" reforms of the 1990's also sought explicitly to encourage states to develop higher standards for student learning and to align the development of assessments, teacher preparation, and other supports with those new standards.

Nonetheless, alignment is not the same thing as coherence, and even highly coordinated reforms can be experienced as incoherent and overwhelming if educators do not have sufficient support. Coherence is a "state of mind" that can result from educators' experience with external demands and the degree to which they interpret them as connected and reinforcing shared goals ([Hatch, 2015](#), p. 105). Thus, coherence is not a characteristic of these demands but rather a perception that is developed and renegotiated over time as educators are confronted with new information and experiences.

Intentional efforts from district leaders to bring together educators to make sense of and learn about new initiatives can contribute to the development of the close social ties necessary for sharing complex information and encouraging a sense of coherence and coordination ([Hodge and Stosich, 2022](#); [Stosich et al., 2021](#)). District leaders can use organizing routines, such as creating opportunities for shared observation of instructional practice or learning about shared tools, to encourage coordination across individuals and roles ([Hatch et al., 2016](#)). Coordination of activities around shared goals has been identified as a key attribute of effective organizations, both inside and outside of education ([Lawler, 1990](#); [Stosich, 2017](#); [Wohlstetter et al., 1994](#)).

In addition, the nature of particular policies may encourage coordination among actors in the same roles, across roles, or present unique challenges for coordination. For instance, studying the efforts of school and district leaders to implement common academic standards for students and a new teacher evaluation framework, we found that the teacher evaluation framework, which focused on aspects of teaching that were presumed to apply across levels and disciplines, allowed for shared learning experiences for educational leaders across roles. By contrast, the districts focused opportunities for learning about academic standards on teachers rather than school leaders, and these opportunities were organized based on content area and grade-level, limiting opportunities for teachers and leaders across roles to develop common understandings and approaches to support implementation of the new expectations for students. Thus, the design of academic standards presented a unique challenge for coordinating the work of administrators as reform leaders ([Stosich, 2018](#); [Stosich et al., 2021](#)).

Municipalities and national systems can also support coherence by putting in place mechanisms that bring together stakeholders to develop common understanding of educational goals and to foster the collective responsibility to meet them. For example, roughly every ten years Finland has engaged in a curriculum renewal process which supports collective reflections on the goals of the comprehensive education system since it was established in the 1970s ([Hatch et al., 2021](#)). The process has evolved over time, but in the last two renewal cycles there has been extensive involvement of key education stakeholders from the very beginning of the discussions. Numerous "curriculum groups" go to work developing the guidelines and objectives in each subject and aspect of the core curriculum. Teachers are at the center of these committees, though the committees also include school leaders, municipal administrators, teacher educators, and researchers among others (and many of those other representatives have themselves been teachers at some point in time). In addition, an advisory board overseeing the whole process includes a cross-section of representatives of teachers, school leaders, parents, students, textbook publishers, researchers, teacher educators, ethnic groups (for instance representatives of the Sami people), and municipalities. Recommendations for changes in the curriculum framework are then discussed and debated by committees of stakeholders at both the municipal and local school levels. The wide engagement of teachers, leaders, teacher educators, textbook publishers, researchers, parents, students and others in the process creates connections among stakeholders at different levels of the system. In turn, those "bridges" give those at the school level the opportunity to influence policy development at both the municipal and national levels and those at the local level also get information and knowledge about major changes *long before those changes are actually made*. As a consequence, schools all across the country are more likely to be ready to head in new directions in concert. More than a mere adjustment to ensure the system is "aligned", however, the curriculum renewal process can also be seen as an opportunity for many people to develop a sense of coherence and an understanding of why schools are doing what they are being asked to do.

Summary

The research on the role of external demands in organizational development shows that leaders have to look beyond the boundaries of their own authority. Effective governance and leadership depend on the abilities to work with and influence those individuals and groups who are both inside and outside schools and districts. In particular, to manage external demands, leaders have to make choices about when to bridge and buffer – when to engage with and when to protect their organizations from external forces.

At the same time, leaders have to learn how to balance their time and energy – focusing enough attention on understanding and dealing with external forces while still attending to and supporting those who report to them directly. To achieve that balance, leaders can foster shared understanding and common priorities with both internal colleagues and external allies, enabling both to advance their collective interests.

When considering which strategies might best achieve their desired goals, leaders of schools, networks, district offices, states and other organizations can also consider their context, including the historical investment in internal priorities, the current capacity of educators to achieve these goals, and the extent to which the policy environment is supportive. Sustained investment, high levels of educator capacity, and strong social networks that reinforce shared goals can support schools in advancing internal priorities while effectively managing external demands. Notably, at each level of the system educational leaders not only manage demands but also create them. The introduction of additional policies or programs, even when complementary, increases the complexity of policy implementation and, thus, may require increased and coordinated leadership support for successful integration in practice.

References

- Abelmann, C., Elmore, R., Even, J., Kenyon, S., Marshall, J., 1999. When Accountability Knocks, Will Anyone Answer? CPRE Research Report Series RR-42. Consortium for Policy Research in Education.
- Carnegie Forum on Education and the Economy, 1986. *A Nation Prepared: Teachers for the 21st Century*. The Report of the Task Force on Teaching as a Profession. Carnegie Corporation, New York.
- Cohen, D.K., Ball, D.L., 1999. Instruction, Capacity, and Improvement. CPRE Research Report Series RR-43. Consortium for Policy Research in Education, University of Pennsylvania. <https://doi.org/10.1037/e382692004-001>.
- Collins, D.B., 2001. Organizational performance: the future focus of leadership development programs. *J. Leader. Stud.* 7 (4), 43–54. <https://doi.org/10.1177/107179190100700404>.
- Corcoran, T., Goertz, M., 1995. Instructional capacity and high performance. *Educ. Res.* 24 (9), 27–31. <https://doi.org/10.3102/0013189X024009027>.
- Cuban, L., 1984. Transforming the frog into a prince: effective schools research, policy, and practice at the district level. *Harv. Educ. Rev.* 54 (2), 129–152. <https://doi.org/10.17763/haer.54.2.m07133p087795h38>.
- DiMaggio, P.J., Powell, W.W., 1983. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *Am. Socio. Rev.* 147–160. <https://doi.org/10.2307/2095101>.
- Elmore, R.F., McLaughlin, M.W., 1988. *Steady Work. Policy, Practice, and the Reform of American Education*. The RAND Corporation.
- Elmore, R.F., 2004. *School Reform from the Inside Out: Policy, Practice, and Performance*. Harvard Education Press. <https://doi.org/10.1037/e565212006-012>.
- Forman, M.L., Stosich, E.L., Bocala, C., 2017. *The Internal Coherence Framework: Creating the Conditions for Continuous Improvement in Schools*. Harvard Education Press.
- Fullan, M., Quinn, J., 2015. *Coherence: The Right Drivers in Action for Schools, Districts, and Systems*. Corwin, Thousand Oaks, CA.
- Fullan, M., 1999. On effecting change in arts education. *Arts Educ. Pol. Rev.* 100 (3), 17–18. <https://doi.org/10.1080/10632919909599458>.
- Goertz, M.E., Floden, R.E., O'Day, J.A., 1996. *The Bumpy Road to Education Reform*. Consortium for Policy Research in Education. University of Pennsylvania.
- Harris, A., Spillane, J., 2008. Distributed leadership through the looking glass. *Manag. Educ.* 22 (1), 31–34. <https://doi.org/10.1177/0892020607085623>.
- Hatch, T., Hill, K., Roegman, R., 2016. Investigating the role of instructional rounds in the development of social networks and district-wide improvement. *Am. Educ. Res. J.* 53 (4), 1022–1053. <https://doi.org/10.3102/0002831216653205>.
- Hatch, T., Corson, J., van den Berg, S.G., 2021. *The Education We Need for a Future We Can't Predict*. Corwin.
- Hatch, T., 2001. Incoherence in the system: three perspectives on the implementation of multiple improvement initiatives in one district. *Am. J. Educ.* 109 (4), 107–137. <https://www.jstor.org/stable/1209265>.
- Hatch, T., 2002. When improvement programs collide. *Phi Delta Kappan* 83 (8), 626–634. <https://doi.org/10.1177/003172170208300812>.
- Hatch, T., 2009. *Managing to change: How schools survive (and sometimes thrive) in turbulent times*. Teachers College Press.
- Hatch, T., 2015. Connections, coherence, and common understanding in the common core. In: Supovitz, J.A., Spillane, J.P. (Eds.), *Challenging Standards: Navigating Conflict and Building Capacity in the Era of the Common Core*. Rowman & Littlefield, pp. 103–112.
- Hatch, T., 2020. Building the capacity for collective responsibility in Norway. In: Jones, M., Harris, A. (Eds.), *Leading and Transforming Education Systems*. Springer, pp. 121–131.
- Hatch, T., 2021. Expanding the boundaries of learning. *Phi Delta Kappan* 102 (8), 8–13.
- Hodge, E.M., Stosich E.L., 2022. Accountability, alignment, and coherence: how educators made sense of complex policy environments in the Common Core era. *Educ. Eval. Policy Anal.* <https://doi.org/10.3102/01623737221079650>.
- Honig, M., Hatch, T., 2004. Crafting coherence: how schools strategically manage multiple, external demands. *Educ. Res.* 33 (8), 16–30. <https://doi.org/10.3102/0013189X033008016>.
- Jaffe-Walter, R., 2008. Negotiating mandates and memory: inside a small schools network for immigrant youth. *Teach. Coll. Rec.* 110 (9), 2040–2066.
- Johnson, S.M., Marietta, G., Higgins, M.C., Mapp, K.L., Grossman, A.S., 2014. *Achieving Coherence in District Improvement: Managing the Relationship Between the Central Office and Schools*. Harvard Education Press.
- Kania, J., Kramer, M., 2011. Collective impact. *Stanford Soc. Innovat. Rev.* 9 (1), 36–41.
- Kaufman, J.H., Berglund, T., 2018. School Supports for Teachers' Implementation of State Standards: Findings from the American School Leader Panel. RAND Corporation. <https://doi.org/10.7249/RR2318>.
- Lawler III, E.E., 1990. *Strategic Pay: Aligning Organizational Strategies and Pay Systems*. Jossey-Bass.
- Louis, K.S., Robinson, V.M., 2012. External mandates and instructional leadership: school leaders as mediating agents. *J. Educ. Adm.* 50 (5), 629–665. <https://doi.org/10.1108/09578231211249853>.
- Malen, B., Rice, J.K., 2004. A framework for assessing the impact of education reforms on school capacity: insights from studies of high-stakes accountability initiatives. *Educ. Pol.* 18 (5), 631–660. <https://doi.org/10.1177/0895904804268901>.
- March, J., 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* 2 (1), 71–87. <https://www.jstor.org/stable/2634940>.
- March, J., 1994. *Primer on Decision Making: How Decisions Happen*. The Free Press.
- Metz, M.H., 1988. Some missing elements in the educational reform movement. *Educ. Adm. Q.* 24 (4), 446–460. <https://doi.org/10.1177/0013161X88024004009>.
- Meyer, J.W., Rowan, B., 1977. Institutionalized organizations: formal structure as myth and ceremony. *Am. J. Sociol.* 83 (2), 340–363. <https://www.jstor.org/stable/2778293>.
- Newmann, F.M., King, M.B., Youngs, P., 2000. Professional development that addresses school capacity: lessons from urban elementary schools. *Am. J. Educ.* 108 (4), 259–299. <https://doi.org/10.1086/444249>.
- O'Day, J.A., Smith, M.S., 1993. Systemic reform and educational opportunity. In: Fuhrman, S. (Ed.), *Designing Coherent Education Policy: Improving the System*. Josey-Bass, pp. 250–312.

- O'Day, J., Floden, M.E., Goertz, R.E., 1995. Building Capacity for Education Reform. Consortium for Policy Research in Education. <https://doi.org/10.1037/e383992004-001>.
- Putnam, R., 2000. Bowling Alone: The Collapse and Revival of American Community. Simon and Shuster. <https://doi.org/10.1145/358916.361990>.
- Smith, M.S., O'Day, J., 1990. Systemic school reform. *J. Educ. Pol.* 5 (5), 223–267. <https://doi.org/10.1080/02680939008549074>.
- Spillane, J.P., Thompson, C.L., 1997. Reconstructing conceptions of local capacity: the local education agency's capacity for ambitious instructional reform. *Educ. Eval. Pol. Anal.* 19 (2), 185–203. <https://doi.org/10.3102/01623737019002185>.
- Spillane, J.P., Hallett, T., Diamond, J.B., 2003. Forms of capital and the construction of leadership: instructional leadership in urban elementary schools. *Sociol. Educ.* 76 (1), 1–17. <https://doi.org/10.2307/3090258>.
- Stein, M.K., Nelson, B.S., 2003. Leadership content knowledge. *Educ. Eval. Pol. Anal.* 25 (4), 423–448. <http://www.jstor.org/stable/3699585>.
- Stosich, E.L., Hatch, T., Hill, K., Roegman, R., Allen, T.D., 2021. Social networks and policy coherence: administrators' Common Core and teacher evaluation advice networks. *Educ. Pol. Anal. Arch.* 29 (60). <https://doi.org/10.14507/epaa.29.5990>.
- Stosich, E.L., 2017. Leading in a time of ambitious reform: principals in high-poverty urban elementary schools frame the challenge of the Common Core State Standards. *Elem. Sch. J.* 117 (4), 539–565. <https://doi.org/10.1086/691585>.
- Stosich, E.L., 2018. Principals and teachers 'craft coherence' among accountability policies. *J. Educ. Adm.* 56 (2), 203–219. <https://doi.org/10.1108/JEA-10-2016-0124>.
- Tyack, D.B., Cuban, L., 1995. *Tinkering Toward Utopia*. Harvard University Press.
- Weick, K.E., 1995. *Sensemaking in Organizations*, vol. 3. Sage.
- Wohlstetter, P., Smyer, R., Mohrman, S.A., 1994. New boundaries for school-based management: the high involvement model. *Educ. Eval. Pol. Anal.* 16 (3), 268–286. <https://doi.org/10.3102/01623737016003268>.
- Zajac, E.J., Westphal, J.D., 1994. The costs and benefits of managerial incentives and monitoring in large US corporations: when is more not better? *Strat. Manag. J.* 15 (S1), 121–142. <https://www.jstor.org/stable/2486814>.

Leadership and school capacity for continuous improvement: toward a systemic approach

Peter Slegers, BMC Consultancy, Amersfoort, The Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	216
Perspectives on school improvement	217
Planned change	217
Policy implementation	217
Effective schools	217
School culture	217
Social networks	218
School improvement journeys	218
Inside and outside views	218
School capacity for improvement: conditions that support teaching and learning	218
Leadership	218
Teacher collaboration	219
Participative decision making	220
Trust	220
Teacher motivation	220
Teacher learning	220
Building school capacity for improvement: changes and journeys	221
Changes and reciprocal effects	221
Journeys	222
School improvement as a system-wide phenomenon	222
Conclusion	223
References	223

Introduction

Over the past 30 years, schools have had to be responsive to different fragmented educational policies aimed at improving teaching and learning, including standard-based reforms, the increased demand for accountability, and the growing marketization in education. In addition to these pressures to coordinate the pursuit of excellence and equity from system actors (e.g. state agencies, district offices, school administration), different so-called non-system actors (e.g. voluntary and professional organizations, universities, and private firms) have taken a stronger role in the governance of education, the provision of educational services and the implementation of policy. These changes have altered the institutional landscape of schools worldwide (Cohen et al., 2018). Schools also have faced broader social environmental changes such as shifts in student demographics, increased teacher shortages, the call for customization, the growing social inequality in education, and more recently the rise and rapidly spreading of the COVID-19 pandemic. To effectively cope with these changes, schools have to develop their school-wide capacity to improve teaching and student learning. As a consequence, building the improvement capacity of schools has become a key challenge for practitioners, policymakers and educational researchers.

In their efforts to understand the nature and dynamics of school improvement as an organizational process, the vast majority of scholars conducted qualitative case studies and large-scale quantitative cross-sectional surveys. These studies have offered a “one-time snapshot” of leadership practices and school conditions that support teaching and learning (e.g. Hallinger and Heck, 1998; Slegers et al., 2002). More recently, scholars have begun to use longitudinal research designs to investigate the process of building school-wide capacity for improvement over time and patterns of change during schools’ improvement journeys (e.g. Duff, 2021; Hallinger and Heck, 2011; Heck and Hallinger, 2009, 2010a,b; Mulford and Silins, 2011; Thoonen et al., 2012; Slegers et al., 2014). Although substantial progress has been made, school improvement research faces some conceptual and methodological challenges (Feldhof and Radisch, 2021). The goal of this chapter is to summarize the key findings of school improvement research and discuss one of the key conceptual challenges facing school improvement research. I do so by arguing for using a systemic perspective to better understand the complexity and dynamics of school improvement as a system wide phenomenon.

The structure of the chapter mirrors this goal. I will start by discussing the different perspectives scholars have used in analyzing supportive conditions and processes associated with school improvement. I then sketch the key findings of the large body of school improvement research, making a distinction between studies using case study and cross-sectional designs on the one hand and longitudinal designs on the other. Following this overview, I elaborate on the potential of using a systemic perspective to investigate

how schools, that are embedded in a broader educational system, can build school capacity for improvement. I end the chapter with some suggestions for future research.

Perspectives on school improvement

In conceptualizing school improvement scholars have used a variety of theoretical perspectives (Slegers and Leitwood, 2010). Below, I briefly discuss these perspectives that are often grounded in different assumptions and research traditions.

Planned change

Since the late 1970s, scholars have started to frame school improvement as a form of planned change, while taking the school as the key unit of change. An important perspective on the process of implementing change in schools has been the modeling of school improvement as overlapping developmental stages, including adaptation, implementation and institutionalization (Fullan, 1991). This line of inquiry focused on processes of personal change (Evans, 1996; Rogers, 2003), organizational change (O'Toole, 1995) and educational change (Hall and Hord, 2006; Van den Berg and Slegers, 1996a,b), identifying the impact of mechanism, strategies and "change agents" at various stages in the change process. Findings from these studies indicate that reform programs have been often changed during implementation through "mutual adaptation" and that events and structures limit the impact of managing change successfully (Berman and McLaughlin, 1978; Firestone and Corbett, 1988; Louis and Miles, 1990; Van den Berg and Slegers, 1996a).

Policy implementation

Researchers studying the implementation of standards-based and accountability reforms reframed the notion of mutual adaption as a learning process and a lack of alignment between local, district and state policies as key mechanisms by which policy influences school and class instructional practice (Coburn et al., 2016). This research has demonstrated how educators (administrators, school leaders and teachers) respond to policy reforms and that their understanding of policy is shaped by their own beliefs about "good" instruction. This often resulted in the reinforcing of preexisting practices, the transformation of innovative curricula into traditional ones and the adoption of surface-level changes rather than deeper fundamental changes to instructional practice. Research in this domain has also identified that system leaders' sense making, organizational routines and artifacts in schools play an important role in how the work of organizing for, and supporting instruction is coupled with the broader institutional environment (Spillane and Burch, 2006; Spillane et al., 2011, 2019; Woulfin, 2015).

Effective schools

A great deal of evidence for school-wide interventions to improve teaching and student learning has emerged from the knowledge base on effective teaching and schools. This line of research has focused on "what works" to guarantee school success, leading to a variety of lists of educational effectiveness enhancing conditions (see for example Scheerens, 2016). Although many of the variables included are amenable to school-wide interventions, the issue of how to build school-wide capacity for less effective schools to become more effective has not been addressed (Louis et al., 1999; Duff, 2021). Scholars have, therefore, called for a synthesis of school effectiveness with school improvement research (Reynolds et al., 2000; Slegers et al., 2002). Comprehensive School Reform Programs can be seen as reform designs or models that are grounded in educational effectiveness research and offer solutions to schools' improvement problems. Research into the effects and implementation of these models have found modest to positive effects on student achievement scores examined (Borman et al., 2003). Moreover, most of the models were actively shaped and reshaped ("co-constructed") by teachers, school leaders and other local educators to accommodate the various aspects of the models to the local context, thereby underlining the important role "mutual adaptation" plays in implementing reform programs as founded in earlier research (Datnow et al., 2002).

School culture

Another perspective emerged from research that focused on how school culture shapes educational reforms and connects to school improvement (Rosenholtz, 1991; Sarason, 1982). Researchers have highlighted the importance of a strong school culture as the softer strategy in school improvement and believed that "reculturing" is crucial to entail deep change (Fullan, 2000). In exploring the link between school culture and school improvement, scholars have situated professional learning communities and organizational learning as the building blocks for school culture (Leithwood and Louis, 1998; Louis et al., 1996; Mitchell and Sackney, 2000; Lee and Louis, 2019). This work suggests that teachers' collaborations as organizational mechanism provide the habits of searching for ideas, sharing information with colleagues and of generating new ideas within and from outside the school. Findings have shown that dimensions of organizational learning and PLC promote teaching and learning (Bolam et al., 2005; Lomos et al., 2011; Vescio et al., 2008).

Social networks

More recently, scholars have started to investigate how social relations between teachers may support school improvement, using social network theory (Daly, 2010). The study of social networks in schools have granted more attention to the social embeddedness of teachers' social relationships in the informal patterns of interaction in schools that play a role in the exchange of knowledge, information and materials, and the facilitation of individual and collective action. This research views the school as an organization from a social interactional perspective in which teachers' interactions are assumed to create opportunities for professional learning in support of school-wide capacity for improvement (Sleegers et al., 2018). The concepts and tools offered by social network theory and research supply opportunities for understanding more fully the conditions, processes and outcomes of school improvement and educational change.

School improvement journeys

As school leaders must take the environmental and organizational conditions into account as they work toward continuous school improvement, researchers have started to investigate the relationship between schools' contexts, leadership practices, and improvement processes during change (Day et al., 2011; Duff, 2021; Hallinger and Heck, 2011). They try to analyze different patterns or stages in the journey of school improvement and relate these patterns to contextual conditions. Findings from the few studies conducted so far have started to increase our knowledge of how building school-wide capacity is embedded in the school improvement context.

Inside and outside views

Sleegers and Leithwood (2010) identified two views that underlie the different perspectives that dominated research on school improvement and educational change. The first view, which they term the "inside" view, focuses on the capacity of schools to transform themselves into supportive environments for teacher and student learning. By focusing on the professional growth of teachers and the problem-solving capacities of the system itself, change is considered as part of a larger process of making sense of situations through individual and collective learning. The second view, the "outside" view, concerns the implementation of policy reforms into schools to improve the quality of instruction and enhance student learning. This view focuses on the attempts of reform policies and those who are charged with their implementation to force or induce changes in teachers' practice via mandates, incentives and normative pressures. Although these two views use different assumptions about the nature and direction of school improvement and educational change, they inform each other. According to Sleegers and Leithwood (2010), findings from school improvement research suggest that schools with high improvement capacity are not only better able to transform themselves into supportive contexts for teaching and learning but also to effectively respond to the pressures of external change and align the various goals, materials and demands of policy programs to the local context. So, school capacity for improvement connects research on top-down (external view) and bottom-up approaches (internal view) to school improvement and is considered an important prerequisite for continuous and sustained school improvement (Sleegers et al., 2014).

In the next sections, I therefore will look at research on school improvement through the lens of school capacity for improvement and will organize a wide range of studies that have provided insights into conditions that foster continuous school improvement. I will start with a summary of the main benefits of correlational and case studies, followed by a discussion of the key findings from the few available longitudinal studies.

School capacity for improvement: conditions that support teaching and learning

School capacity for improvement has been defined as a set of conditions at school and teacher level that enable teachers' professional learning, support teaching and student learning and provide strategies aimed at continuous school improvement (Heck and Hallinger, 2009; Mulford and Silins, 2003; Sleegers et al., 2005; Stoll et al., 2006). In conceptualizing school capacity for improvement, scholars have used models in which variation in the improvement capacity of schools is assumed to be a function of the capacities and motivations of workplace personnel (*teacher conditions*), the characteristics of the organizational setting (*organizational conditions, including leadership*) in which they work, and the *external environment* (Leithwood et al., 2002; Sleegers et al., 2014). In these models, leadership practices and school organizational conditions are assumed to have indirect effects on teachers' instructional practices needed for gains in student outcomes through teacher motivation and learning.

Most research in this domain has consisted of case studies and cross-sectional surveys and focused on relating organizational and teacher conditions to instructional practices and student learning in a specific context or at one moment in time. Six sets of conditions at school and teacher level that support teaching and student learning appear to have robust empirical support: leadership, teacher collaboration, participative decision making, trust, teacher motivation and teacher learning. I briefly review these conditions to provide some context about how these conditions are addressed in school improvement research.

Leadership

Leadership is critical in school effectiveness and school improvement (Leithwood et al., 2008, 2020). To understand the role leadership plays in the promotion of school improvement and educational change, scholars investigated specific sets of leadership

practices covered by two dominant models, including transformational and instructional leadership (Hallinger, 2003). Researchers focusing on organizational learning have used transformational leadership models to investigate how leadership practices bring about change. Transformational leaders enhance the commitment, competencies and extra effort of teachers required for organizational change by vision building, providing individual support and stimulating intellectual stimulation. Research in this domain suggests that transformational leadership is related with various conditions for school improvement, including shared decision making, collaboration, teacher efficacy, extra effort, teacher learning and changes in teachers' instructional practices (Geijsel et al., 2003; Leithwood and Jantzi, 2005; Leithwood and Sun, 2012; Thoonen et al., 2011). Those scholars that are more concerned with the *outcome* than the *process* of school improvement, have emphasized instructional leadership as a key construct to study the impact of leadership on teaching and learning. They argue that the primary task of school leaders is to improve instruction as student learning is the bottom line of school improvement. Instructional leaders are goal-oriented, academically focused, coordinate, supervise and monitor teaching and learning and create "academic press" through defining the school's mission, managing the instructional program and promoting a positive school climate (Hallinger, 2005). The large amount of studies conducted in this field has yielded indirect effects of instructional leadership practices on student achievement as mediated by organizational conditions (e.g. school mission and goals, expectations, curriculum).

To better understand and unravel the paths through which school leaders have an impact on teaching and student learning, scholars have suggested to use a more integrated model in which transformational leadership and instructional leadership coexist as the generic functions of leadership practices are often obscured by the use of different styles of leadership (Bruggencate et al., 2012; Hallinger, 2003; Leithwood et al., 2004; Robinson et al., 2008). Research on the influence of integrated leadership on teaching and learning has shown that leadership for school improvement requires both instructional and transformational leadership (Marks and Printy, 2003; Printy et al., 2009; Urlick and Bowers, 2014). Successful school leaders make use of the same basic repertoire of practices, including: (1) developing a vision and giving direction, (2) understanding and developing people, (3) redesigning the organization and (4) managing the teaching and learning program (Leithwood et al., 2006a,b). In line with this, an approach called leadership for learning, also representing a blend of transformational and instructional leadership, has become popular and gained international currency (Hallinger, 2011; MacBeath and Cheng, 2008). Leadership for learning signals, among other things, the critical role that leadership plays in building school capacity for improvement by highlighting the importance of a school-wide focus on learning, both for teachers and students. Based on these multiple leadership frames and available evidence, Hitt and Tucker (2016) recently developed an overarching set of leadership practices for school improvement that accounts for transformational, instructional and leadership for learning. This unified model distinguishes the following five leadership practices: (1) establishing and conveying a vision, (2) facilitating a high-quality learning experience for students, (3) building professional capacity, (4) creating a supportive organization for learning, and (5) connecting with external partners.

The large body of research on the impact of leadership for school improvement has mainly focused on the role of the principal as formal leader. More and more scholars have started to shift the focus from formal leaders to school-wide leadership that is distributed (Spillane, 2006): leadership that is exercised by those in both management (e.g. the principal, assistant principals, department heads) and non-management roles (e.g. teachers, staff). Researchers have found that patterns of distributed leadership (e.g. more or less complex and coordinated patterns) appear to depend on goals and the expertise of those who provide leadership (Harris et al., 2007). Evidence from studies indicates that some of the above mentioned basic set of leadership practices, including developing people and managing the instructional program, are more distributed than the others (Day et al., 2010; Hulpia et al., 2009; Leithwood et al., 2010). Setting directions and redesigning organizations are practices that are mostly carried out by those in formal management roles. Moreover, findings suggest that various patterns of distributed leadership affect school capacity for school improvement and support teaching and learning (Leithwood et al., 2006a,b; Lu and Hallinger, 2018).

Teacher collaboration

Cooperative and collegial relationships and open communication provide emotional and psychological support for teachers' work and opportunities for free exchange of ideas and professional learning (Geijsel et al., 2001; Kwakman, 2003; Rosenholtz, 1991; Smylie et al., 1996). Several studies have shown that teacher collaboration is positively associated with teachers' self-efficacy, their organizational commitment, professional learning and can improve, if focused on student learning, classroom practices (Bryk et al., 1999; Louis and Marks, 1998; Rosenholtz, 1989; Smylie, 1988; Stoll et al., 2006). So, teacher collaboration as a school organizational condition can promote organizational and instructional improvement (Darling-Hammond and Richardson, 2009; Hargreaves and Fullan, 2012; Nguyen and Ng, 2020; Shirrell et al., 2019; Slegers and Leithwood, 2010; Stoll, 2009). Research on professional learning communities has underlined the important role of teachers collaborations and interactions for school improvement (Lee and Louis, 2019). In this line of research, teacher collaboration has been interpreted as one of the five dimensions, (together with shared values, deprivatized practice, reflective dialog and a focus on student learning) that constitute a professional community and more broadly as a feature of or embedded in school culture. Findings from studies into schools as professional learning communities indicate that schools with higher levels of teachers collaboration and norms of collegiality are associated with higher levels of improved instruction, data use and school performance (Leithwood et al., 2001; Earl and Katz, 2006; Smylie and Wenzel, 2003; Coburn and Turner, 2011). By focusing on the importance of social ties and social relationships among teachers, social network theory and methods have supplied some needed specificity to delve more fully in teachers' collaboration as an organizational condition for school improvement. Research on social networks in schools suggests that the patterns of social relationships among educators that result from teachers' interaction promote teachers' motivation (collective

efficacy) and matters for different outcomes (e.g. innovative climate and student learning) of school improvement (Moolenaar et al., 2011, 2012; Moolenaar and Sleegers, 2010; Sleegers et al., 2018).

Participative decision making

Participative decision making is assumed to foster teachers' commitment to change and reinforce their motivation. Participation in decision making refers to joint decision-making or shared influence in decision-making by school leaders and teachers. When teachers participate more in decision making they receive more feedback and create more opportunities to exchange new knowledge and skills and to experiment. Research has shown that teachers' participation in decision making is positively related to teacher motivation (Geijssel et al., 2009; Rosenholtz, 1989; Thoonen et al., 2011), improved teaching (Smylie et al., 1996), their acceptance of policy (Smylie, 1988), and is important for schools' policy-making capacity (Sleegers et al., 1994). As such, participation in decision-making may provide standards teachers can use to evaluate their own practice and clarify school goals. Despite the positive impact on teachers' work, participation in decision-making appears to be contingent on conditions in schools. For example, teachers working at schools with a high innovative capacity participate more in decision making than teachers working at schools with a low innovative capacity (Geijssel et al., 1999; Van den Berg and Sleegers, 1996b). Moreover, teachers are willing to participate most in decisions that are related to the curriculum and instruction, while they are least willing to participate in decisions about personnel and managerial issues (Bacharach et al., 1990). So, a lack of connection with the technical core of teachers' work appears to be an important hindrance for many teachers to become engaged in participative decision-making (Marks and Louis, 1997).

Trust

Trust among organizational members is considered an important emotional or normative resource for school improvement (Bryk and Schneider, 2002; Coburn and Russel, 2008; Forsyth et al., 2011). Trust emerges from social interactions among organizational members and can refer to an individual emotional state that may be shared by others (relational trust) or a norm that is socially constructed by group members (collective trust). Trust is often associated with organizational climate, group cohesion, collaboration and the school culture and facilitates collective action (Adams and Forsyth, 2013). Positive social interactions may create a safe environment and foster trust by reducing uncertainty about the engagement of the other party and establishing a tolerance for failure without the fear of being punished or ridiculed (Bryk and Schneider, 2002; Lipshitz et al., 2007). Research has indicated that densely connected social networks in schools increase teacher trust, which in turn facilitates a more innovative-supported school climate (Moolenaar and Sleegers, 2010). Trust is also positive related to teacher professionalism and teachers' willingness to contribute to organizational goals (Penuel et al., 2007; Smylie and Hart, 1999; Tschannen-Moran and Hoy, 1998; Tschannen-Moran, 2001, 2009). Moreover, trust contributes to improved classroom practice and the performance of schools (Bryk and Schneider, 2002; Goddard et al., 2001, 2009; Tschannen-Moran, 2014). By keeping the professional learning community together, trust facilitates the continuous improvement of schools which may lead to less rigid responses to high-stakes accountability pressures (Daly, 2009).

Teacher motivation

Among teacher conditions that support teaching and learning, teacher motivation plays an important role (Geijssel et al., 2009; Thoonen et al., 2011). Motivational processes are a function of one's personal goals and beliefs about one's capacities and one's context and are aimed at helping the person to evaluate the need for change or action (Ford, 1992). In research on teacher commitment to change a strong belief in and acceptance of the organization's goals and values is often considered as an element of teacher motivation (Geijssel et al., 2003; Leithwood et al., 1999). Teachers are more motivated if they have internalized school goals and values as their personal goals. In setting where there are challenging personal goals, individuals are like to develop a positive sense of self-efficacy (Bandura, 1986). Self-efficacy is an important motivational construct: it is a future-oriented belief about the competence a person expects to display in a given situation. Self-efficacy can strengthen teachers in pursuing action and may take away fear of failure (Bandura, 1986; Geijssel et al., 2009). For example, teachers with higher levels of self-efficacy challenge themselves to reach more challenging goals, will persist in pursuing their interest and actions when faced with obstacles, have less fear of failure, and will find satisfying solutions (Bandura, 1997; Schwarzer and Hallum, 2008). Research has shown that both teachers' internalization of school goals into personal goals and their self-efficacy beliefs affect instructional practices as mediated by teacher learning (Geijssel et al., 2009; Thoonen et al., 2011). Of these two motivational factors, teachers' sense of efficacy appears to be the most important for explaining teacher learning and instruction.

Teacher learning

Scholars have emphasized that professional learning embedded in schools is crucial for school improvement (Clarke and Hollingsworth, 2002; Kwakman, 2003; Putnam and Borko, 2000; Sleegers et al., 2005). Research on organizational learning in schools suggest that learning among staff as well as the development of the school as a whole depend on the opportunities schools create for teacher's professional learning (Sleegers and Leithwood, 2010). Scholars have distinguished different professional learning activities that can promote school improvement (Kwakman, 2003; Lohman and Woolf, 2001; Oude Groote Beverborg et al., 2015a).

These professional learning activities vary from individual activities aimed at acquiring new knowledge, skills, and information, where other colleagues are not necessarily needed, to social learning activities requiring others. Examples of individual learning activities are self-reflection, keeping up to date, and experimentation, while sharing information and asking for feedback are examples of social learning activities. Research has shown that both individual and social learning activities affect instruction (Garet et al., 2001; Geijsel et al., 2001; Supovitz, 2002; Thoonen et al., 2011). Teachers engagement in professional learning activities is facilitated by motivational factors of which self-efficacy is the most important one. Moreover, self-efficacy mediates the relationship between leadership practices, organizational conditions and teacher learning. These findings demonstrate how the interplay between leadership practices, organizational conditions and teacher's motivation affect teaching learning and instruction.

Building school capacity for improvement: changes and journeys

Knowing which set of conditions of school capacity for improvement matter for teaching and student learning is not the same as knowing *how* schools can build their capacity for continuous improvement. Because school improvement involves change in the state of different set of conditions supporting teaching and learning, longitudinal research is needed that take into account the changing relationships between these conditions (Hallinger and Heck, 2010; Thoonen et al., 2012; Slegers et al., 2014). Scholars have therefore started to conduct longitudinal research to examine the causal and reciprocal relationships between conditions that comprise school capacity for improvement, teaching and student learning over time. For a systematic review of the methods and designs used in this line of research, I refer to Feldhoff et al. (2016).

Changes and reciprocal effects

Longitudinal research indicates that school capacity for improvement can develop over time (Heck and Hallinger, 2010a,b; Thoonen et al., 2012). For example, in a longitudinal study of a large sample of Dutch elementary schools, Thoonen et al. (2012) investigated the increase or decrease of school capacity for improvement over time. They found that leadership practices, organizational conditions, motivational factors and teacher engagement in learning activities showed a development over time, both positive and negative. While leadership practices and organizational conditions gradually increased over time, motivational factors and teachers engagement in learning activities barely improved or even become less during the years. Moreover, the "momentum" on which the variables show improvement differs: leadership appears to improve at an earlier point in time than school organizational conditions, teacher motivational factors, and teachers' engagement in professional learning activities.

Building on this work, Slegers et al. (2014) examined the relationship between changes in the improvement capacity of schools and changes in teachers' classroom practices over a four years period of time. The results showed that teachers' classroom practices changed over time and that teachers succeed in sustaining their improved instruction. In addition, motivational factors and teacher learning had the greatest impact on changes in teachers' classroom practices, while leadership and organizational conditions were particularly important in enhancing teacher's motivations and learning.

Based on a large random sample of US elementary schools over four years, Heck and Hallinger (2009) investigated the effects of change in distributed leadership and school improvement capacity on student performance. They found that changes in distributed leadership and the improvement capacity of schools have significant impact on growth in student learning in reading and math. Furthermore, the findings indicate that the effects of change in distributed leadership on growth in student learning is mediated by change in the school improvement capacity.

These findings shed more light on the nature of the causal relations and the interaction over time between leadership practices and different conditions for improvement that support teaching and student learning. However, they do not provide evidence in support of possible reciprocal relationships between these variables and how they mutually shape each other over time. Although scholars have emphasized the need to use reciprocal effects models in school improvement research, little progress has been made so far. A notable exception is the work of Hallinger and Heck. They analyzed the reciprocal relationships between distributed (or what they call collaborative leadership), school improvement capacity and growth in student math and reading achievements (Hallinger and Heck, 2010; Heck and Hallinger 2010a,b). The findings support the reciprocity between collaborative leadership and school improvement capacity. The effects of changes in leadership on changes in school improvement capacity appeared to be smaller than the effects of school improvement capacity on leadership over time. Moreover, changes in collaborative leadership and school improvement capacity were related to growth in student achievement and vice versa. Regardless of the level of student achievement, schools can thus improve student learning by changes in leadership practices and improvement capacity.

Findings from a three-wave longitudinal study of a large sample of Vocational Educational Training Colleges in the Netherlands, showed that leadership practices and teacher collaboration mutual influence each other over time (Oude Groote Beverborg et al., 2015b, 2020). Moreover, some evidence was found for the reciprocity of teacher collaboration and teachers' engagement in learning activities, especially self-reflection. Self-efficacy appeared not to play a pivotal role as a psychological lever between leadership and teacher learning as found in cross-sectional studies. However, a reverse effect was found: self-reflection as an individual learning activity had a positive influence on self-efficacy over time.

Journeys

Rather than focusing on changes in the average improvement capacity among a sample of schools, researchers have recently started to investigate differences between schools' improvement trajectories (i.e. stable, improving, declining) over time. Analyzing these school improvement journeys helps to reveal how building school-wide capacity is situated in the school improvement context. Hallinger and Heck (2011) were able to identify three school trajectories based on student growth in math achievement over a period of 4 years. The results suggest that the positive relationships between change in leadership and capacity building is stronger in schools with the highest student growth over time than in schools with moderate and small growth in student achievement. In their 3-year mixed-methods national study, Day et al. (2016) found evidence for four broad improvement phases: foundational, developmental, enrichment, and renewal phases. These different phases contained different combinations of leadership practices, ranging from attention to "basics," making the school more student-centered to distributed leadership. Based on teachers' perceptions of the improvement capacity of 1200 elementary and middle schools in the New York City context between 2017 and 2019, during COVID-19, Duff (2021) found six types of schools through Latent Class Analyses: versatile, collaborative, developing, balkanized, controlled and demoralized (see also Duff and Bowers, 2021). The findings also showed that schools moved from one type to another between 2017 and 2019, reinforcing that school improvement capacity is not fixed. This typology and the increase and decrease of the improvement capacity was supported by the qualitative case studies conducted. Moreover, the case studies provided strong support for the critical role leadership plays in the school improvement journeys and that the improvement capacity helps schools to manage both internal and external change, particularly to overcome the pandemic's challenges.

School improvement as a system-wide phenomenon

Much of the empirical work on school improvement has used the schoolhouse as the unit of analysis and focused on how leadership practices, organizational conditions, motivational factors and teacher learning support teaching and student learning inside schools. The work of improving teaching and enhancing student learning is not solely the province of school leaders and teachers as schools are part of school systems and embedded in a broader educational environment that provides legitimacy, essential resources and support (Cohen et al., 2018). The organization and the working of school systems and the broader educational environment are consequential for building school-wide capacity for improvement. Many aspects of the school improvement context are shaped by all kinds of design efforts often mandated by political and policy initiatives and guided by an institutional logic aimed at the pursuit of excellence and equity. Focusing narrowly on school capacity for improvement without taking the broader educational environment in which schools are embedded into account misses the array of influences mutually reinforcing system components and actors operating at various system levels may have on building school-wide capacity for continuous improvement (Cohen and Moffitt, 2010). Some scholars have paid attention to how districts, superintendents or school boards may foster school improvement inside schools. There is some evidence that local education authorities and superintendents can support leadership and the improvement capacity of schools and, in turn, affect teaching and learning by district-wide, job-embedded professional development, using data for organizational learning and accountability, and a district-wide focus on educational goals and student achievement (Daly et al., 2017; Hooge et al., 2019; Leithwood and Seashore-Louis, 2011; Lee and Louis, 2019; Louis, 2015; Saatcioglu et al., 2011; Togneri and Anderson, 2003). More recently, a small but growing body of research on network-based educational improvement examines the organizational dynamics of exploration and exploitation as the core learning processes in school improvement networks—characterized by a general hub-member network form—as a strategy for improving education (Glazer and Peurach, 2013; Peurach, 2016; Peurach et al., 2018). Findings from this research suggest that, although some school improvement networks are successful in creating highly developed capabilities for managing and organizing instructionally-focused continuous improvement, success is not the rule. Furthermore, the findings indicate that a system-level innovation infrastructure which is strong in support for evaluating impact, but weak in support for continuous improvement provides uneven support for developing the learning dynamics in networks that have the potential to enhance school capacity for continuous improvement.

However, scholars have devoted little attention to the paths and mechanisms through which the institutional environment can incentivize and resource school systems to build and (re)design their educational infrastructures to coordinate and enable instructional improvement. One of the key challenges for school improvement research is to understand how the interrelationships among resources, processes and (system and non-system) actors, operating at various levels, from schools to school systems and the broader educational environment, work in tandem in an effort to support school wide capacity for continuous improvement (Cohen et al., 2018). Research on school improvement would profit from using a more systemic approach that focuses on the interaction between leadership, school improvement capacity, school systems and networks and how these linkages constitute instructionally-focused continuous improvement, while taking the broader institutional system in which schools are embedded as the relevant unit of analysis (Cohen et al., 2018). Hence, scholars should frame school improvement as a system-wide phenomenon and investigate the dynamics and complexity of school improvement from a systematic perspective by drawing on system thinking and complexity theory (Schechter and Shaked, 2018; Slegers et al., 2018).

Conclusion

In this chapter I reviewed the large amount of research on school improvement and argued for framing school improvement as a system-wide phenomenon. By using two views, the internal and external view, that underlie the different perspectives scholars have used to examine school improvement, I brought together key findings of school improvement research from the lens of school improvement capacity. Robust empirical support evidence has been found for a set of conditions comprising school capacity for improvement that support teaching and student learning: leadership, teacher collaboration, participative decision making, trust, teacher motivation and teacher learning. Findings from the available longitudinal studies also suggest that school capacity for improvement fluctuates over time and seems to be malleable. While leadership appeared to play a critical role as a driving force for school improvement, leadership and school improvement capacity mutual influence each other to support teaching and student learning. By revealing the variation in change trajectories, studies into school improvement journeys indicate that all schools and not only the poor performing school have to be engaged in building school-wide capacity to improve teaching and learning.

Although substantial progress has been made, I argued that scholars should look at school improvement from a systematic approach. Using a multi-dimensional systemic framework would help school improvement research to focus on the dynamic changes between conditions at school and teacher level that support teaching and learning and the social interactions, patterns and networks between actors and schools at different levels of the educational system that support school-wide capacity for continuous improvement. This may result in a fruitful attempt to synthesize the diverse theoretical perspectives on school improvement into a single overarching conceptual framework and help us to better understand the complexity and dynamics of school improvement.

Future directions for research can be derived from the weak spots of theoretical perspectives and empirical evidence. Key issues for research are to examine and explain how large-scale reforms at the macro-level and patterns of instructional organization and management within and between school systems at the meso-level interact to shape school wide capacity for continuous improvement. Future research should also focus on how change can be sustained over time and how reforms can be extended to more teachers, schools and districts. Little empirical evidence exists for differences between educational systems around the world and how these differences contribute to differences in the practice of leadership and the improvement capacity of schools to support teaching and learning. To unravel the complexity and dynamics of school improvement, researchers should use different designs (e.g. longitudinal designs, comparative research designs and design-based implementation research) and mixed methods to gather rich, dynamic, contextual and longitudinal data. Findings from these studies will help us to deepen our understanding of how schools that are embedded in an educational system can build school capacity for continuous improvement. Findings from these studies may also help us to develop a theory of design that can provide practitioners with tools to redesign and change system mechanisms that constrain and enable school capacity for continuous improvement.

References

- Adams, C.M., Forsyth, B.F., 2013. Revisiting the trust effect in urban elementary schools. *Elem. Sch. J.* 114, 1–21.
- Bacharach, S.B., Bamberger, P., Conley, S.C., Bauer, S., 1990. The dimensionality of decision participation in educational organizations: the value of a multi-domain evaluative approach. *Educ. Adm. Q.* 26 (2), 126–167.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, NJ.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. W. H. Freeman and Company, New York.
- Berman, P., McLaughlin, M., 1978. Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations. In: *The Findings in Review*. Rand, Santa Monica, CA.
- Bolam, R., McMahon, A., Stoll, L., Thomas, S., Wallace, M., Greenwood, A., Hawkey, K., Ingram, M., Atkinson, A., Smith, M., 2005. *Creating and Sustaining Effective Professional Learning Communities*. Research Report No. 637. DfES and University of Bristol, London.
- Borman, G., Hewes, G., Overman, L.T., Brown, S., 2003. Comprehensive school reform and achievement: a meta-analysis. *Rev. Educ. Res.* 73, 125–130.
- Bruggencate, G., Luyten, H., Scheerens, J., Sleegers, P., 2012. Modeling the influence of school leaders on student achievement: how can school leaders make a difference? *Educ. Adm. Q.* 48 (4), 699–732.
- Bryk, A., Schneider, B., 2002. *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation, New York.
- Bryk, A., Camburn, E., Louis, K.S., 1999. Professional community in Chicago elementary schools: facilitating factors and organizational consequences. *Educ. Adm. Q.* 35 (5), 751–781.
- Clarke, D., Hollingsworth, H., 2002. Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18 (8), 947–967.
- Coburn, C.E., Russell, J.L., 2008. District policy and teachers' social networks. *Educ. Eval. Pol. Anal.* 30 (3), 203–235.
- Coburn, C.E., Turner, E.O., 2011. Research on data use: a framework and analysis. *Measurement* 9 (4), 173–206.
- Coburn, C.E., Hill, H.C., Spillane, J.P., 2016. Alignment and accountability in policy design and implementation: the common core state standards and implementation research. *Educ. Res.* 45 (4), 243–251.
- Cohen, D.K., Moffitt, S.L., 2010. *The Ordeal of Equality: Did Federal Regulation Fix the Schools?* Harvard University Press, Cambridge.
- Cohen, D.K., Spillane, J.P., Peurach, D.J., 2018. The dilemmas of educational reform. *Educ. Res.* 47 (3), 204–212.
- Daly, A.D., Moolenaar, N.M., Carrier, N., Fresno, M., del, 2017. The school district as complex adaptive system: exploring a complexity theory perspectives on policy. *J. Res. Organ. Educ.* 1, 31–61.
- Daly, A.J., 2009. Rigid response in an age of accountability: the potential of leadership and trust. *Educ. Adm. Q.* 45 (2), 168–216.
- Daly, A.J. (Ed.), 2010. *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge, MA.
- Darling-Hammond, L., Richardson, N., 2009. Research review/teacher learning: what matters. *Educ. Leadersh.* 66 (5), 46–53.
- Datnow, A., Hubbard, L., Mehan, H., 2002. *Extending Educational Reform: From One School to Many*. Routledge Falmer Press, New York.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Harris, A., Gu, Q., Brown, E., 2010. *Ten Strong Claims About Successful School Leadership*. The National College for School Leadership, Nottingham.

- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *School Leadership and Student Outcomes: Building and Sustaining Success*. Open University Press, Maidenhead.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Duff, M., Bowers, A.J., 2021. Identifying a typology of New York city schools through teacher perceptions of organizational capacity: a latent class analysis. *Leadersh. Pol. Sch.* 1–25.
- Duff, M., 2021. *A Detour in School Improvement Journeys: A Mixed Methods Analysis of School Change During the COVID-19 Pandemic*. Dissertation. Columbia University, Columbia.
- Earl, L., Katz, S., 2006. Leading schools in a data-rich world. In: *Harnessing Data for School Improvement*. Corwin Press, Thousand Oaks, California.
- Evans, R., 1996. *The Human Side of School Change*. Jossey-Bass, San Francisco, CA.
- Feldhoff, T., Radisch, F., 2021. Why must everything be so complicated? Demands and challenges on methods for analyzing school improvement processes. In: Oude Groote Beverborg, A., Feldhoff, T., Maag Merki, K., Radisch, F. (Eds.), *Concept and Design Developments in School Improvement Research: Longitudinal, Multilevel and Mixed Methods and Their Relevance for Educational Accountability*. Springer, Dordrecht, The Netherlands, pp. 9–27.
- Feldhoff, T., Radisch, F., Bischof, L.M., 2016. Designs and methods in school improvement research: a systematic review. *J. Educ. Adm.* 54 (2), 209–240.
- Firestone, W.A., Corbett, H.D., 1988. Planned organizational change. In: Boyan, N.J. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York & London, pp. 321–341.
- Ford, M.E., 1992. *Motivating Humans: Goals, Emotions, and Personal Agency Beliefs*. Sage, Newbury Park, CA.
- Forsyth, P.B., Adams, C.M., Hoy, W.K., 2011. *Collective Trust. Why Schools Can't Improve Without It*. Teachers College Press, New York.
- Fullan, M., 1991. *The New Meaning of Educational Change*. Teachers College Press, New York.
- Fullan, M., 2000. The three stories of education reform. *Phi Delta Kappan* 81 (8), 581–584.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., Yoon, K.S., 2001. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38 (4), 915–945.
- Geijsel, F., Van Den Berg, R., Slegers, P., 1999. The innovative capacity of schools in primary education: a qualitative study. *Int. J. Qual. Stud. Educ.* 12 (2), 175–191.
- Geijsel, F., Slegers, P., van den Berg, R., Kelchtermans, G., 2001. Conditions fostering the implementation of large-scale innovation programs in schools: teachers' perspectives. *Educ. Adm. Q.* 37 (1), 130–166.
- Geijsel, F., Slegers, P., Leithwood, K., Jantzi, D., 2003. Transformational leadership effects on teachers' commitment and effort toward school reform. *J. Educ. Adm.* 41 (3), 228–256.
- Geijsel, F.P., Slegers, P.J.C., Stoel, R.D., Krüger, M.L., 2009. The effect of teacher psychological and school organizational and leadership factors on teachers' professional learning in Dutch schools. *Elem. Sch. J.* 109 (4), 406–427.
- Glazer, J.L., Peurach, D.P., 2013. School improvement networks as a strategy for large-scale education reform: the role of environments. *Educ. Pol.* 27 (4), 676–710.
- Goddard, R.D., Tschannen-Moran, M., Hoy, W.K., 2001. A multilevel examination of the distribution and effects of teacher trust in students and parents in urban elementary schools. *Elem. Sch. J.* 102 (1), 3–17.
- Goddard, R.D., Salloum, S.J., Berebitsky, D., 2009. Trust as a mediator of the relationships between poverty, racial composition, and academic achievement: evidence from Michigan's public elementary schools. *Educ. Adm. Q.* 45 (2), 292–311.
- Hall, G.E., Hord, S.M., 2006. *Implementing Changes: Patterns, Principles and Pitfalls*, second ed. Pearson/Allyn & Bacon, Boston.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9 (2), 157–191.
- Hallinger, P., Heck, R.H., 2010. Collaborative leadership and school improvement: understanding the impact on school capacity and student learning. *Sch. Leadersh. Manag.* 30 (2), 95–110.
- Hallinger, P., Heck, R.H., 2011. Exploring the journey of school improvement: classifying and analyzing patterns of change in school improvement processes and learning outcomes. *Sch. Effect. Sch. Improv.* 22 (1), 1–27.
- Hallinger, P., 2003. Leading educational change: reflections on the practice of instructional and transformational leadership. *Camb. J. Educ.* 33 (3), 329–351.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leadersh. Pol. Sch.* 4 (3), 221–239.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press/Ontario Principals' Council.
- Harris, A., Leithwood, K., Day, C., Sammons, P., Hopkins, D., 2007. Distributed leadership and organizational change: reviewing the evidence. *J. Educ. Change* 4, 337–347.
- Heck, R.H., Hallinger, P., 2009. Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *Am. Educ. Res. J.* 46 (3), 626–658.
- Heck, R.H., Hallinger, P., 2010a. Testing a longitudinal model of distributed leadership effects on school improvement. *Leadersh. Q.* 21 (5), 867–885.
- Heck, R.H., Hallinger, P., 2010b. Collaborative leadership effects on school improvement: integrating unidirectional-and reciprocal-effects models. *Elem. Sch. J.* 111 (2), 226–252.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86 (2), 531–569.
- Hooge, E.H., Moolenaar, N.M., van Look, C.J., Janssen, S.K., Slegers, P.J.C., 2019. The role of district leaders for organization social capital. *J. Educ. Adm.* 57 (3), 296–316.
- Hulpia, H., Devos, G., Rosseel, Y., 2009. The relationship between the perception of distributed leadership in secondary schools and teachers' and teacher leaders' job satisfaction and organizational commitment. *Sch. Effect. Sch. Improv.* 20 (3), 291–317.
- Kwakman, K., 2003. Factors affecting teachers' participation in professional learning activities. *Teach. Teach. Educ.* 19 (2), 149–170.
- Lee, M., Louis, K.S., 2019. Mapping a strong school culture and linking it to sustainable school improvement. *Teach. Teach. Educ.* 81, 84–96.
- Leithwood, K., Jantzi, D., 2005. A review of transformational leadership research 1996–2005. *Leadersh. Pol. Sch.* 4 (3), 177–200.
- Leithwood, K., Louis, K.S. (Eds.), 1998. *Organizational Learning in Schools*. Swets and Zeitlinger, Lisse, The Netherlands.
- Leithwood, K., Seashore-Louis, K., 2011. *Linking Leadership to Student Learning*. John Wiley & Sons Inc., New Jersey.
- Leithwood, K., Sun, J., 2012. The nature and effects of transformational leadership: a meta-analytic review of unpublished research. *Educ. Adm. Q.* 48 (3), 387–423.
- Leithwood, K.A., Jantzi, D., Steinbach, R., 1999. *Changing Leadership for Changing Times*. McGraw-Hill Education, Buckingham.
- Leithwood, K.A., Aiken, R., Jantzi, D., 2001. *Making Schools Smarter*. Corwin Press, Thousand Oaks, California.
- Leithwood, K., Jantzi, D., Mascall, B., 2002. A framework for research on large-scale reform. *J. Educ. Change* 3, 7–33.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K.T., 2004. *Review of Research: How Leadership Influences Student Learning*. Wallace Foundation, New York, NY.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006a. *Successful School Leadership: What It Is and How It Influences Student Learning*. Research Report 800. Department for Education, London, UK.
- Leithwood, K., Mascall, B., Strauss, T., Sacks, R., Memon, N., Yashkina, A., 2006b. Distributing leadership to make schools smarter: tacking the ego out of the system. *Leadersh. Pol. Sch.* 6 (1), 37–67.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leadersh. Manag.* 28 (1), 27–42.
- Leithwood, K., Louis, K.S., Wahlstrom, K., Anderson, S., Mascall, B., Gordon, M., 2010. How successful leadership influences student learning: the second installment of a longer story. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*. Springer, Dordrecht, The Netherlands, pp. 611–629.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leadersh. Manag.* 40 (1), 5–22.
- Lipshitz, R., Friedman, V.J., Popper, M., 2007. *Demystifying Organizational Learning*. Sage Publishing, London.

- Lohman, M.C., Woolf, N.H., 2001. Self-initiated learning activities of experienced public school teachers: methods, sources, and relevant organizational influences. *Teach. Teach.* 7 (1), 59–74.
- Lomos, C., Hofman, R.H., Bosker, R.J., 2011. Professional community and student achievement—a meta-analysis. *Sch. Effect. Sch. Improv.* 22, 121–148.
- Louis, K.S., Marks, H.M., 1998. Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *Am. J. Educ.* 106 (4), 532–575.
- Louis, D., Miles, M., 1990. *Improving the Urban High School*. Teachers College Press, New York.
- Louis, K.S., Marks, H.M., Kruse, S., 1996. Teachers' professional community in restructuring schools. *Am. Educ. Res. J.* 33 (4), 757–800.
- Louis, K.S., Toole, J., Hargreaves, A., 1999. Rethinking school improvement. In: Murphy, J., Louis, K.S. (Eds.), *Handbook of Research on Educational Administration*, second ed. Jossey-Bass, San Francisco, pp. 251–276.
- Louis, K.S., 2015. Linking leadership to learning: state, district and local effects. *Nord. J. Stud. Educ. Pol.* 2015 (3), 30321.
- Lu, J., Hallinger, P., 2018. A mirroring process: from school management team cooperation to teacher collaboration. *Leadersh. Pol. Sch.* 17 (2), 238–263.
- MacBeath, J., Cheng, Y.C., 2008. *Leadership for Learning: International Perspectives*. Sense Publishers, Rotterdam.
- Marks, H.M., Louis, K.S., 1997. Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic achievement. *Educ. Eval. Pol. Anal.* 19 (3), 245–275.
- Marks, H.M., Printy, S.M., 2003. Principal leadership and school performance: integrating transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Mitchell, C., Sackney, L., 2000. *Profound Improvement: Building Capacity for a Learning Community*. Swets & Zeitlinger, Lisse.
- Moolenaar, N.M., Sleegers, P.J.C., 2010. Social networks, trust, and innovation. How social relationships support trust and innovative climates in Dutch Schools. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard University Press, Cambridge, pp. 97–115.
- Moolenaar, N.M., Daly, A.J., Sleegers, P.J.C., 2011. Ties with potential: social network structure and innovative climate in Dutch schools. *Teach. Coll. Rec.* 113 (9), 1983–2017.
- Moolenaar, N.M., Sleegers, P.J., Daly, A.J., 2012. Teaming up: linking collaboration networks, collective efficacy, and student achievement. *Teach. Teach. Educ.* 28 (2), 251–262.
- Mulford, B., Silins, H., 2003. Leadership for organizational learning and improved student outcomes. *Camb. J. Educ.* 33 (2), 175–195.
- Mulford, B., Silins, H., 2011. Revised models and conceptualization of successful school principalship for improved student outcomes. *Int. J. Educ. Manag.* 25 (1), 61–82.
- Nguyen, D., Ng, D., 2020. Teacher collaboration for change: sharing, improving, and spreading. *Prof. Dev. Educ.* 46 (4), 638–651.
- O'Toole, J., 1995. *Leading Change*. Jossey-Bass, San Francisco, CA.
- Oude Groote Beverborg, A., Sleegers, P.J., Van Veen, K., 2015a. Promoting VET teachers' individual and social learning activities: the empowering and purposeful role of transformational leadership, interdependence, and self-efficacy. *Empir. Res. Vocat. Educ. Train.* 7 (1), 1–20.
- Oude Groote Beverborg, A., Sleegers, P.J., Endedijk, M.D., Veen, K.V., 2015b. Towards sustaining levels of reflective learning: how do transformational leadership, task interdependence, and self-efficacy shape teacher learning in schools? *Societies* 5 (1), 2075–4698.
- Oude Groote Beverborg, A., Sleegers, P.J., Moolenaar, N.M., van Veen, K., 2020. Fostering sustained teacher learning: a longitudinal assessment of the influence of vision building and goal interdependence on information sharing. *Sch. Effect. Sch. Improv.* 31 (4), 576–604.
- Penuel, W.R., Fishman, B.J., Yamaguchi, R., Gallagher, L.P., 2007. What makes professional development effective? Strategies that foster curriculum implementation. *Am. Educ. Res. J.* 44 (4), 921–958.
- Peurach, D.J., Penuel, W.R., Russell, J.L., 2018. Beyond ritualized rationality: organizational dynamics of instructionally-focused continuous improvement. In: Connolly, M., Eddy-Spicer, D.H., James, C., Kruse, S.D. (Eds.), *The SAGE Handbook of School Organization*. Sage Publishing, London, pp. 465–489.
- Peurach, D.J., 2016. Innovating at the nexus of impact and improvement: leading educational improvement networks. *Educ. Res.* 45, 421–429.
- Printy, S.M., Marks, H.M., Bowers, A.J., 2009. Integrated leadership: how principals and teachers share transformational and instructional influence. *J. Sch. Leadersh.* 19 (5), 504–532.
- Putnam, R.T., Borko, H., 2000. What do new views of knowledge and thinking have to say about research on teacher learning? *Educ. Res.* 29 (1), 4–15.
- Reynolds, D., Teddlie, C., Hopkins, D., Stringfield, S., 2000. Linking school effectiveness and school improvement. In: Teddlie, C., Reynolds, D. (Eds.), *The International Handbook of School Effectiveness Research*. Falmer Press, London, UK, pp. 206–231.
- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on school outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Rogers, E.M., 2003. *Diffusion of Innovations*. Simon and Schuster, New York, NY.
- Rosenholtz, S.J., 1989. Workplace conditions that affect teacher quality and commitment: implications for teacher induction programs. *Elem. Sch. J.* 89 (4), 421–439.
- Rosenholtz, S.J., 1991. *Teachers' Workplace. The Social Organization of Schools*. Teachers College Press, New York.
- Saatcioglu, A., Moore, S., Sargut, G., Bajaj, A., 2011. The role of school district social capital in district governance: effects on financial and academic outcomes. *Leadersh. Pol. Sch.* 10 (1), 1–42.
- Sarason, S.B., 1982. *The Culture of the School and the Problem of Change*. Allyn and Bacon, Boston.
- Schechter, C., Shaked, H., 2018. System thinking in school organization. In: Connolly, M., Eddy-Spicer, D.H., James, C., Kruse, S.D. (Eds.), *The SAGE Handbook of School Organization*. Sage, London, pp. 249–266.
- Scheerens, J., 2016. *Educational Effectiveness and Ineffectiveness. A Critical Review of the Knowledge Base*. Springer, Dordrecht, The Netherlands.
- Schwarzer, R., Hallum, S., 2008. Perceived teacher self-efficacy as a predictor of job stress and burnout: mediation analyses. *Appl. Psychol.* 57, 152–171.
- Shirrell, M., Hopkins, M., Spillane, J.P., 2019. Educational infrastructure, professional learning, and changes in teachers' instructional practices and beliefs. *Prof. Dev. Educ.* 45 (4), 599–613.
- Sleegers, P., Leithwood, K., 2010. School development for teacher learning and change. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, vol. 7. Elsevier, Oxford, pp. 557–562.
- Sleegers, P.J.C., Bergen, T.C.M., Giebers, J.H.G.I., 1994. The policy-making capacity of schools: result of a Dutch study. *Educ. Man. Adm.* 22 (3), 147–159.
- Sleegers, P., Geijsel, F., van den Berg, R., 2002. Conditions fostering educational change. In: Leithwood, K., Hallinger, P. (Eds.), *Second International Handbook of Educational Leadership and Administration*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 75–102.
- Sleegers, P., Bolhuis, S., Geijsel, F., 2005. School improvement within a knowledge economy: fostering professional learning from a multidimensional perspective. In: Bascia, N., Cumming, A., Datnow, A., Leithwood, K., Livingstone, D. (Eds.), *International Handbook of Educational Policy*. Kluwer, Dordrecht, the Netherlands, pp. 527–543.
- Sleegers, P.J.C., Thoonen, E.E.J., Oort, F.J., Peetsma, T.T.D., 2014. Changing classroom practices: the role of school-wide capacity for sustainable improvement. *J. Educ. Adm.* 52 (5), 617–652.
- Sleegers, P., Moolenaar, N., Daly, A., 2018. The interactional nature of schools as social organizations: three theoretical perspectives. In: Connolly, M., Eddy-Spicer, D.H., James, C., Kruse, S.D. (Eds.), *The SAGE Handbook of School Organization*. Sage Publishing, London, pp. 267–284.
- Smylie, M.A., Hart, A.W., 1999. School leadership for teacher learning and change: a human and social capital development perspective. In: Murphy, J., Louis, K.S. (Eds.), *Handbook of Research on Educational Administration*, second ed. Jossey-Bass, San Francisco, CA, pp. 421–443.
- Smylie, M.A., Wenzel, S.A., 2003. The Chicago Annenberg Challenge: successes, failures, and lessons for the future. In: *Final Technical Report of the Chicago Annenberg Research Project*. Consortium on Chicago School Research, Chicago, IL.
- Smylie, M.A., Lazarus, V., Brownlee-Conyers, J., 1996. Instructional outcomes of school-based participative decision making. *Educ. Eval. Pol. Anal.* 18 (3), 181–198.
- Smylie, M.A., 1988. The enhancement function of staff development: organizational and psychological antecedents to individual teacher change. *Am. Educ. Res. J.* 25 (1), 1–30.
- Spillane, J., Burch, P., 2006. The instructional environment and instructional practice: changing patterns of guidance and control in public education. In: Meyer, H.-D., Rowan, B. (Eds.), *The New Institutionalism in Education*. State University of New York Press, Albany, pp. 87–103.

- Spillane, J.P., Parise, L.M., Sherer, J.Z., 2011. Organizational routines as coupling mechanisms policy, school administration, and the technical core. *Am. Educ. Res. J.* 48 (3), 586–619.
- Spillane, J.P., Seelig, J.L., Blaushild, N.L., Cohen, D.K., Peurach, D.J., 2019. Educational system building in a changing educational sector: environment, organization, and the technical core. *Educ. Pol.* 33 (6), 846–881.
- Spillane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco, CA.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., Thomas, S., 2006. Professional learning communities: a review of the literature. *J. Educ. Change* 7 (4), 221–258.
- Stoll, L., 2009. Capacity building for school improvement or creating capacity for learning? A changing landscape. *J. Educ. Change* 10 (2), 115–127.
- Supovitz, J.A., 2002. Developing communities of instructional practice. *Teach. Coll. Rec.* 104 (8), 1591–1626.
- Thoonen, E.E.J., Slegers, P.J.C., Oort, F.J., Peetsma, T.T.D., Geijssel, F.P., 2011. How to improve teaching practices: the role of teacher motivation, organizational factors, and leadership practices. *Educ. Adm. Q.* 47 (3), 496–536.
- Thoonen, E.E.J., Slegers, P.J.C., Oort, F.J., Peetsma, T.T.D., 2012. Building school-wide capacity for improvement: the role of leadership, school organizational conditions and teacher factors. *Sch. Effect. Sch. Improv.* 23 (4), 441–460.
- Togneri, W., Anderson, S.E., 2003. *Beyond Islands of Excellence: What Districts Can Do to Improve Instruction and Achievement in all Schools*. Retrieved October 2011 from. <http://www.learningfirst.org/publications/>.
- Tschannen-Moran, M., Hoy, W., 1998. Trust in schools: a conceptual and empirical analysis. *J. Educ. Adm.* 36, 334–352.
- Tschannen-Moran, M., 2001. Collaboration and the need for trust. *J. Educ. Adm.* 39 (4), 308–331.
- Tschannen-Moran, M., 2009. Fostering teacher professionalism in schools: the role of leadership orientation and trust. *Educ. Adm. Q.* 45 (2), 217–247.
- Tschannen-Moran, M., 2014. *Trust Matters: Leadership for Successful Schools*, second ed. Jossey-Bass, San Francisco, CA.
- Urlick, A., Bowers, A.J., 2014. What are the different types of principals across the United States? A latent class analysis of principal perception of leadership. *Educ. Adm. Q.* 50 (1), 96–134.
- van den Berg, R., Slegers, P., 1996a. Building innovative capacity and leadership. In: Leithwood, K., Chapman, J., Corson, D., Hallinger, P., Hart, A. (Eds.), *International Handbook of Educational Leadership and Administration*. Kluwer, Dordrecht, pp. 653–699.
- van den Berg, R., Slegers, P., 1996b. The innovative capacity of schools in secondary education: a qualitative study. *Int. J. Qual. Stud. Educ.* 9, 201–223.
- Vescio, V., Ross, D., Adams, A., 2008. A review of research on the impact of professional learning communities on teaching practice and student learning. *Teach. Teach. Educ.* 24 (1), 80–91.
- Woulfin, S.L., 2015. Highway to reform: the coupling of district reading policy and instructional practice. *J. Educ. Change* 16 (4), 535–557.

School reform from a constructivist perspective

Lea Hubbard^a, Hugh Mehan^b, and Amanda Datnow^c, ^a Department of Leadership Studies, School of Leadership and Education Sciences, University of San Diego, San Diego, CA, United States; ^b Emeritus of Sociology and Education Studies, University of California, San Diego, CA, United States; and ^c Division of Social Sciences, Department of Education Studies, University of California, San Diego, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	227
Culture, structure, and agency: components of a constructivist theory of school reform	227
The value of “teacher buy-in” when introducing reform	228
Reculturing teaching practices	229
Political clashes between the district and valued constituencies	229
Summary	230
Schools and reform design teams co-construct reform	230
The value of coherence in restructuring and reculturing school reform	231
District coherence matters	231
Summary	232
Leadership, coherence, and the co-construction of reform	232
Summary	233
Conclusion	233
References	233

Introduction

Attempts to improve schooling and better prepare students have persisted as long as schools have existed. This ameliorative intention has translated into the development, adoption, and implementation of a wide variety of reform efforts, from those that focus on one or two aspects of classroom life to comprehensive efforts that affect classrooms, schools, districts, and even national policy.

We have been studying the school reform process primarily by engaging in qualitative research. Our studies have involved hundreds of interviews with district leaders, principals, teachers, and parents throughout the United States. We spent hours conducting observations of classrooms and professional development events organized to support educational leaders and teachers, and we have examined community factors external to the school to understand their contribution to the shaping of reform.

Our research has taken a constructivist orientation that focuses on the dynamic interplay between culture, structure, and agency. These studies have led us to examine the role coherence and leadership plays in reform. When cultural and structural coherence is present, reform efforts benefit from educators sharing a common language and beliefs about the reform. In contrast, when coherence is absent, the possibility of a reform being successful is often challenged. Each of the cases we discuss here also illustrate the important role leadership plays in affecting reform outcomes. Models of school leadership vary from top-down mandates to “distributive collaborative” models, which are models that offer a more responsive and inclusive set of conditions (Diamond and Spillane, 2016; Woods and Roberts, 2019). Each leadership model defines meaning and mediates educators’ action—either supporting or pushing against the proposed reform.

We begin with a discussion of our theoretical frame and then describe each of the cases of reform change through this lens. By depicting culture, structure, and agency as equal participants in a reflexive system of social action, we gain better insight into the processes that facilitate or constrain school reform efforts.

Culture, structure, and agency: components of a constructivist theory of school reform

Our constructivist position regards reform policy as an ongoing process of negotiating meaning in real-life educational settings. Politically, reform policy is a mechanism for powerful actors to manage contested perceptions by focusing attention on some conditions rather than on others and by promoting a particular interpretation of those conditions. It is also a means for powerful actors to legitimate particular meanings by granting them the authority derived from these actors’ official statuses. Finally, policy works to shore up institutional authority by communicating an institution’s commitment to particular values and ideas.

The constructivist perspective contributes to public policy theory by reinforcing the idea that public policy is both a process and a product of constitutive human activity (Datnow et al., 2002; Hall, 2017a,b; Hall and McGinty, 1997; Hubbard et al., 2006; Mehan, 2012). As a *process*, policy is a means by which statements of value and definitions of reality are constructed, asserted, validated, and negotiated. As *products*, policies are the material residues of these actions. They are cultural objects and artifacts that “embody the

authority to define goals and command means" (Levinson and Sutton, 2001, p. 5), legitimize and reinforce particular views of reality, and grant those definitions some form of "official" (i.e., institutionalized or publicly recognized) status.

Leaders often construct reform policy—the reform process and the products—but policy studies often assume leadership and decision making occur first at the top of a hierarchy. The implication is leaders at the top are more analytically important (Pressman and Wildavsky, 1973; Ravitch, 1995; Smith and Keith, 1971; Tucker and Coddling, 1998). By contrast, the constructivist position does not automatically assign an *a priori* importance to any one hierarchical context, including those at the top of a hierarchy. Instead, it recognizes leadership can be distributed across key agents (Diamond and Spillane, 2016). Woods and Roberts (2019) expanded notions of distributed leadership by emphasizing leadership can also be a collaborative process constructed both from the intentions, emergent actions, and ideas of leaders and other educators.

Our constructivist approach to understanding educational reform also takes into account the influence of reform coherence—the extent to which a reform has been created to ensure consistency within and across schools. According to Fullan and Quinn (2016), coherence can allow educators to share a "depth of understanding about the purpose and nature of the work" (p. 1). When teachers and leaders individually and collectively cohere, they are "focusing direction, cultivating collaborative cultures, deepening learning and securing accountability" (Fullan and Quinn, 2016, p. 12).

An earlier school improvement study of Chicago schools showed elementary schools with stronger program coherence had higher gains in academic achievement (Newmann et al., 2001). "Instructional program coherence," as defined in this study, included a common instructional framework supported by conditions and practices that avoided other competing improvement efforts. Coherence assists teacher effectiveness because teachers work with other teachers in an integrated manner (Newmann et al., 2001). As Cohen et al. (2017) pointed out, however, the pressure to build coherence pushes against a legacy of incoherence in schools and environments. Efforts to build consensus are challenged by those who support differentiated outcomes and methods. As open systems, schools must respond to parents and community members and navigate political pressures.

At the level of whole school and/or district-wide change, building coherence involves respecting and building upon teachers' professional wisdom (Datnow and Park, 2019). We know, for example, teachers are not passive adopters of reform. They may resist, comply, or adapt. Even when they believe in the reform, teachers often modify a reform by exerting their judgment, in many cases to respond to the diverse needs of their students. District-level coherence can facilitate school improvement as well, but, more often, districts keen on improving academic achievement adopt multiple initiatives. If these initiatives lack a coherent fit, teachers are confused and feel overburdened with the demands of each reform (Hubbard et al., 2014). Conflicting program goals can challenge implementation coherence and "conflict with those [programs] currently in place in local school communities" (Madda et al., 2007, p. 1959). Investigating coherence provides additional insight into factors that support or challenge reform success.

The value of "teacher buy-in" when introducing reform

In this section, we describe how a comprehensive reform played out in San Diego City Schools (SDCS) from its inception in July 1998 to August 2002. We examine the cultural and political clashes that erupted within the district, and between the district and valued constituencies, to shape reform outcomes. The reform was notable for its attention to the structural and cultural dimensions of the reform, while minimizing the role of agency within the organization. Although reform leaders attended to the actions of philanthropists and actors in positions of power, they paid less attention to actors "on the ground" (i.e., teachers, principals, and community members). This lack of attention to those most intimately involved in adopting and implementing the reform contributed significantly to the premature demise of this ambitious reform (Hubbard et al., 2006).

To achieve the reform's goal of improving student learning, district reform leaders installed *The Blueprint for Student Success*, a content-driven, centralized, comprehensive, and fast-paced reform starting in 1998. This plan was put forth to ensure there was coherence across all schools in the district. The most notable structural features emanated from centralized leadership in which major instructional, operational, and professional development decisions originated from the top of the district while administrators at school sites, staff developers, and resource teachers were tasked with implementing plans designed by the superintendent and "chancellor of instruction."

The district leadership attempted to "reculture" the SDCS by organizing educators into *communities of practice*, moving teaching from private to public practice, and transforming operational managers into "instructional leaders." This reculturing and push against the taken-for-granted grammar of schooling (Tyack and Cuban, 1995) involved developing common thoughts, beliefs, and a common language about skills, practices, and accountability among educators in every part of the school system to reach the district's goal, which was to improve student achievement by supporting teaching and learning in the classroom.

The district's centralized, fast-paced, and comprehensive reform was legitimated on moral and political grounds. Comprehensive, simultaneous, and centralized implementation was legitimated morally by saying significant numbers of students could not be asked to wait until an entire system was ready to provide them with the support they needed to learn. Their theory of action was legitimated politically by saying piecemeal modifications were all too often absorbed by the system and therefore did not lead to paradigmatic shifts. The SDCS' investments in instruction and professional development were legitimated by claiming these changes in professional practices rested on a well-established, empirically supported disciplinary knowledge base.

Reculturing teaching practices

Prior to the move to systemic reform in the mid-1990s, teachers were generally regarded as autonomous agents. Teachers were relatively free to assemble the curriculum and instruction they enacted in their classroom. The SDCS district leaders concluded this “private practice” seldom synchronized and calibrated teachers’ practices across classrooms at one school and across schools within the district, resulting in differences in performance across schools. Instead, SDCS leaders aimed for a coherent system of change designed to create a common set of beliefs and some shared and valued practices visible in all classrooms.

This attempt to “reculture” (Fullan and Hargreaves, 2012; Hargreaves, 1994) moved teaching from private practice to communities of public practice. Reculturing involved opening classroom doors to principals’ and coaches’ scrutiny, working collaboratively with peers to solve common problems, and developing common thoughts, beliefs, and language about skills, practices, and accountability. Many teachers—following the lead of their union—resisted these changes in their practice because they challenged their beliefs that they had earned autonomy, or the right to teach as they saw fit. Teachers and union leaders especially resented calls for changes in teaching practice because they were designed and implemented from the “top down” without consulting teachers. Resentments turned into clashes.

Another significant reform move that clashed with teachers’ beliefs about their instructional roles was generated when reform leaders directed high school principals and their teachers to teach reading strategies. High school principals and teachers insisted they knew how to teach biology, algebra, English literature, and other subjects, but they neither knew how nor wanted to teach reading strategies. Elementary school teachers should teach reading, they insisted, not high school teachers. Furthermore, whereas the elementary school calendar enabled extended learning time for literacy, the high school calendar that scheduled classes into discrete time periods worked against the extended learning time needed for improving students’ reading skills.

Many teachers and teachers’ union leaders who opposed the reform said it was their job to teach—to present the material; it was students’ responsibility to learn—to “get it.” If students did not get it—then it was students’ fault; it was not teachers’ responsibility to change instruction to meet students’ needs.

The SDCS reform leaders countered that all children can learn and that it is teachers’ responsibility to find the means to actualize this basic principle of the reform. This idea—that teachers’ high expectations and instruction can improve students’ academic performance—challenged the idea that social factors outside the school, such as students’ impoverished conditions, their lack of mastery over the English language, and their frequent movement between schools, are reasons why students may struggle to learn. Critics of the reform complained the district’s focus on instruction as the means to close the achievement gap did not give enough weight to the influence of students’ social circumstances and instead blamed teachers for students’ failures.

Another set of clashes based in institutional culture erupted over one of the most basic tenets of the reform: Area superintendents and principals should abandon their operational responsibilities and become “instructional leaders.” Instructional leaders became responsible to teach and support teachers in the successful enactment of the reform. This demand challenged some teachers’ and principals’ strongly held beliefs about their leadership role. Moreover, many principals rose through the ranks on the operational and administrative side of the district and were not prepared to coach all teachers in all subjects.

Political clashes between the district and valued constituencies

The district’s theory of action that placed instructional change on the leading edge of the reform agenda generated conflict within the district and between district leaders and valued constituencies. While these structural changes and commitments to raising student achievement won the district leadership the support of the powerful local business community, the national press, and granting agencies, it did not secure teachers’ buy-in or win significant union, community, or unanimous school board support. The resulting political and cultural conflicts, combined with a devastating state budget crisis, led the SDCS in new directions.

In promoting the progressive ideal of closing the achievement gap, the district leadership called attention to the poor performance of students in low-income neighborhoods who tend to have teachers with the least teaching experience and the least professional qualifications (e.g., degrees and credentials). In response, they allocated significant material resources to these “focus schools.” Exposing these “social facts” was intended to rally the community, especially parents of color, by admitting the public school system had not served their children well and by encouraging support for a more rigorous curriculum and higher quality instruction.

These rhetorical moves spawned an unfortunate side effect, especially among some members of the teaching corps, however. Teachers felt they had been unfairly blamed for the achievement gap, and some parents and students agreed with them. Adding to this perception of blame was a perception of disrespect—that local teachers’ knowledge did not count; only the expertise of people from New York City (the prior home of the “chancellor of instruction”) and New Zealand (the home of the literacy program adopted by SDCS) mattered.

While supporting the structural aspects of the reform through increased funding and new construction, parents in San Diego’s affluent neighborhoods lambasted the pedagogical dimensions of the reform. They said the new structure of 2- and 3- hour literacy blocks watered down high school instruction, forced more valued elective courses from students’ courses of study, and restricted teachers’ creativity. A coalition of wealthy parents and teachers exercised their social capital and negotiated concessions from the district over these complaints. A high school in an affluent neighborhood was granted “pilot” status, which exempted the school from conformity to the reform after a coalition of parents and teachers threatened to form an independent charter school. Their actions began to erode the reform leaders’ plan for a coherent comprehensive reform.

Parents and community activists in San Diego's less affluent and predominantly ethnic neighborhoods voiced different concerns than their more affluent contemporaries, albeit no less stridently. The "Latino Coalition," one particularly vocal group within the Latino community, supported the *Blueprint* reform for 18 months because of its promise to close the achievement gap, but the coalition withdrew its support in the fall of 2000 because their leaders said the reforms were not meeting the needs of English language learners, leadership profiles did not reflect the diversity of students in the district, and the district was diminishing due process and respect for parents (Latino, 2001).

In addition, a combination of community groups, supported by the board minority, accused the district of making administrative decisions "top down" without parent or community consultation. Whereas parents in affluent neighborhoods traded their social capital for concessions from the district, when parent and community groups in less affluent neighborhoods were not able to do so, they resorted to walkouts, lawsuits, and appeals to the US and California Secretaries of Education. None of these actions were successful, however.

Summary

The SDCS district's attempt to introduce this coherent, wide-ranging, and innovative reform produced unintended consequences. Introducing innovations into complex systems such as school districts is unpredictable because the existing system pushes back in sometimes predictable, sometimes mysterious ways. "Push back" rooted in organizational culture—including the agency of the educators within the organization and those embedded in community politics—contributed to the premature demise of this innovative reform.

The culturally based reason for resistance to proposed innovations concerned the conflicting meanings teachers, principals, instructional leaders, community members, and reform leaders attributed to new norms, beliefs, and standard operating procedures. The district leaders' desire to convert the area superintendents' and principals' roles from operational to instructional, transform all their teachers into literacy instructors, and make teachers' practice public not private collided with many teachers' and union leaders' long-standing and deeply held beliefs about leadership and teaching roles.

Findings on the underutilization of the expertise of actors on the ground are especially troubling in light of the literature on the importance of "teacher buy in" or "relations of trust" (Bryk and Schneider, 2003) for successful school reforms. Datnow (2000) concluded new programs are not successful without meaningful involvement of teachers in their adoption and implementation. The process of reform initiation, Datnow said, echoing Fullan (1991), is as important as the reform itself. There is a world of difference, between compliance with the parameters of a particular solution and genuine support for its principles (Datnow, 2000; Hubbard et al., 2006).

District leaders invested more heavily in building political alliances with powerful community and national groups than they did in cultivating "buy-in" with educators up and down the line. As a result, the local teachers' union, a vocal school board minority, and advocacy groups based in well-to-do and less-well-to-do communities voiced loud and consistent opposition to the district's content-driven, comprehensive, and centralized reforms. The reform design team took a calculated risk that a fast-paced, centralized reform would transform instruction that, in turn, would produce positive gains in student performance. This evidence would then convince all concerned about the efficacy of the reforms. The short unhappy life of the reform suggests this bet did not pay off.

Schools and reform design teams co-construct reform

A well-documented finding in studies of educational reform is that schools modify reform models as much as, if not more than, reform models change schools (Cuban, 1998). As the examples we present have shown, however, these accommodations are not necessarily the result of recalcitrance, resistance, or ignorance of educators in local school sites. Conditions in local contexts cause educators to modify reform models to better suit their needs. Some adaptations are rather benign, other adaptations enhance reform designs, whereas others threaten the intensity and potential success of the reform.

Achievement Via Individual Determination (AVID) is an educational reform program designed to narrow the achievement gap between socioeconomic and ethnic groups. AVID promises to improve the achievement of low-income and minority students and increase their chances of attending college by offering them a more rigorous college preparatory high school curriculum accompanied by a special college-prep class.

The implementation of the AVID college preparatory program in California, Kentucky, North Carolina, and Virginia was not only influenced "top down" by the actions of AVID's design team; the actions of the design team was also influenced "bottom up" by the actions of educators in schools, as well as by other forces existing in different layers of the system (Datnow et al., 2002).

An example of a design team changing in response to a combination of political actions initiated by the state and pressures from local constituents occurred when AVID redefined its student population. From 1980 until 1995, AVID said it recruited low-income students with high academic potential from ethnic groups historically underrepresented in colleges and universities. AVID changed its way of representing students, partially in response to the elimination of affirmative action, first by the UC Regents and then by California voters, and partially in response to sentiments expressed at new school sites, especially in Kentucky and Virginia. Although changes in California state law that eliminated affirmative action did not directly affect AVID, the leadership felt AVID would be on safer ground to no longer mention race or ethnicity in its program statements. Instead, AVID changed the representation of the student population it serves. AVID now says it provides college prep services to "students in the middle."

Local political pressures in Kentucky and Virginia reinforced the need for AVID to redefine its intended student population. Educators in high schools beginning to implement AVID recoiled from depicting AVID students in racialized terms. Doing so, they said, would offend White parents and community members who would see AVID as engaged in reverse discrimination. More fundamentally, AVID's reform model challenged entrenched cultural beliefs about academic success and the distribution of intelligence and ability along racial lines. AVID's policy to place underrepresented minority students in college prep courses, however, met resistance from some educators and community members who believed intelligence is fixed, natural, and racially based. They said Black and Latino students lacked these skills. These long-standing and deep-seated cultural beliefs in schools throughout the study sites had led to tracking practices that excluded Black students from honors and AP courses and left them with fewer opportunities for future academic success than their White peers.

To navigate the oppositional discourse about intelligence, race, and academic achievement, the AVID design team adopted a strategy of accommodation. Rather than retreating from serving schools in Kentucky and Virginia that had requested their services, AVID redefined their student population just as they had in California. Instead of explicitly taking race or ethnicity into account, they advertised recruiting "students in the middle." When teachers were unable to gain district approval for offering an AVID academic support class during the regular school day, they resorted to offering a "Zero Period" class before the regular school day, or a "seventh Period" after school.

This case reveals two dimensions of school reform efforts: (a) the conditions in local contexts, and contexts far away from local contexts, influence educators to modify reform models to better suit their needs; and (b) cultural beliefs about race and intelligence can become structural constraints impinging upon school reform.

AVID's decision to change the definition of students targeted for admission and to offer college prep courses before or after school was influenced by decisions made in California's government chambers—which are far removed from AVID's schools—and then by local educators in Virginia and Kentucky reluctant to devote class time to a program devoted primarily to underachieving students. Deep-seated and long-standing cultural beliefs about race-based intelligence became structural constraints that threatened to derail this reform effort. Local AVID educators offered support classes at alternative times—actions that salvaged their fledgling programs by working around structural constraints.

The value of coherence in restructuring and reculturing school reform

In 2015, AVID expanded its opportunities to narrow the achievement gap by moving from solely a middle and high school program to a program for elementary schools. AVID created "AVID Elementary" (AE), a schoolwide program that created a structure and culture intended to motivate students to begin to think and prepare for an academic path that would eventually qualify them for college admission.

In a recent study of AE in six schools across the United States, we found each school was expected to be structurally and culturally coherent across all grade levels within the school and coherent with all other AE schools within the larger AVID Program (Hubbard, 2019; Hubbard et al., 2022). AE educators are expected to adhere to the AVID Essentials by embedding a specific set of instructional practices.

Understanding reform as a co-constructed process—that is, taking structural and cultural factors into account—is clarified when considering the role coherence plays in shaping reform. The AE reform achieved coherence through a centralized and comprehensive professional development (PD) system, primarily through district trainings and national summer institutes. Each year, more than 40,000 educators gather at one of 12 institutes across the country for AVID's premiere training event. Districts also host their own AVID PD often with members from AVID Central.

Structural coherence in AE schools is achieved through the use of common educational tools. AE teachers cohere around instructional practice by implementing common pedagogical strategies such as collaborative groups, philosophical chairs, and Socratic seminars. For example, teachers in every classroom use writing, inquiry, collaboration, organization, and reading (WICOR) strategies across the curriculum. Every AE school has a site team with a lead teacher who organizes ongoing instructional training for all teachers across all grade levels (Hubbard, 2019; Hubbard et al., 2022).

Elementary schools traditionally have focused on a school curriculum that gives little, if any, attention to college. By contrast, AE attempts to *reculture* a school by shifting expectations and beliefs among teachers about the necessity of preparing students for college and career. School site teams help to support this goal by creating a common language shared by all teachers, empowering and creating opportunities for teacher collaboration. Teachers across grade levels work together routinely to reflect and revise their practice and create new strategies to push the work deeper. All principals and school counselors are expected to attend their own AVID trainings to enable them to support the structural and cultural changes required to support the goals of AE.

District coherence matters

The role of the district is always important in the success of school improvement. This axiom proved to be true for the schools in our study as we examined the success of AE's adoption, implementation, and the ongoing support they received to meet their immediate and long-term goals. District, school leaders, and teachers pointed to the important distinction they made between districts "*with AVID*" (an attitude that we have the AVID program) and "*an AVID district*" (an attitude that AVID is what our district is about). When AVID is districtwide and the district is all about AVID, there is a high level of support for AE. When district leaders offer

a greater sensitivity to making sure principals and teachers are clear on how AVID fits with the overall vision of the district and filter any other initiatives that might prevent them from deepening the work of AE, AE seems to run more smoothly. Teachers were taught how they could use their AVID strategies to respond to the goals of other district reform initiatives. Moreover, district leaders explained they would only adopt an initiative if it was consistent with the AVID mission. District leaders committed to AVID took actions to ensure all other reform initiatives created a consistent message and a consistent ideology that put AVID first.

Coherence broke down at the district level in some districts, however, when AE students attempted to transition into an AVID elective in middle school. Students found they were blocked from participation because too few sections of the AVID elective class were available to them. Realizing middle schools students want to have the opportunity to take a variety of electives, middle school principals opted instead to offer an elective “wheel,” an arrangement where AVID is only one among several electives rather than scheduling multiple sections of AVID.

In response to the limitations of the wheel approach to elective offerings at the middle school level, AVID Central adopted a new *schoolwide* approach. AVID leaders reasoned, by embedding AVID strategies throughout all classes in the middle school, all students would gain some exposure to AVID’s academic support strategies, even if they missed out on the in-depth intervention a single AVID class would provide. All middle school teachers, ideally, would receive some AVID training so they could implement AVID strategies in their classes.

This approach has had unintended consequences. Depending on the commitment of school educators at each school, AVID is more or less evident, and, as a result, the reform has had varying levels of impact on students’ academic outcomes. AVID calls this interruption in support an example of another “opportunity gap.” Without full participation in AVID in middle school, the likelihood that the typical AVID population of students will seek AVID in high school and continue on a pathway toward college is limited. AVID’s own research shows students need 3 years of participation in the AVID class to outperform their non-AVID peers, and this interruption may derail students.

Summary

The case of AE shows the profound affect schoolwide structural and cultural coherence can have on building and sustaining reform. Moreover, AE is functioning under a leadership model that privileges collaborative professionalism. Principals, teachers, and counselors work together to socialize students into a college-going culture. Common professional development provided by AVID Central helps to build consistent practice at the school and district level. When a reform is institutionalized through the development of common structures, and a schoolwide and district culture is created that motivates educators to cohere around a focused common goal, then reform goals are more likely to be successful. When coherence breaks down, however, at the systems level, due in part to AVID Central’s own middle school approach to participation in AVID, coupled with the unwillingness on the part of middle school principals to adequately support AVID schoolwide, AE students miss educational opportunities to continue with AVID.

Leadership, coherence, and the co-construction of reform

The case of AVID Elementary points to the importance of coherence in the co-construction of educational reform. The next case, a story of school turnaround, reinforces the need for coherence and also highlights the role of collaborative leadership (For more detail on this case, see [Hubbard and Datnow, 2020](#)). A previously underperforming, under-enrolled, urban US middle school underwent a dramatic change when new leadership took over and engaged teachers in a revisioning process. The school restructured into a magnet school that has gained international attention for their innovative practices. In the past 3 years, approximately 1000 educators from across the globe visit the school annually in an attempt to replicate their model. There are now 4 times as many applicants as there are open spots each year.

A new principal began by engaging staff in a reculturing process. Morale among teachers was low, as they had experienced a frequent turnover of principals and felt pressured by accountability measures that did not inspire them or their students. The principal saw an immediate need to include teachers in the visioning for the school and to shift from a more hierarchical model of leadership to an inclusive one. Many honest conversations were held among the then-new principal and the staff, addressing tough topics about student expectations and mindsets. This began the shift toward a school culture that embodied love, respect, and trust for students and teachers alike. With these foundational elements in place, the principal and teachers collaborated to chart plans for instructional improvement.

After a period of much discussion and joint learning, the group decided design thinking (DT) would be at the core of the school’s reform efforts. DT is an instructional approach to increase student engagement, motivate innovative and critical thinking, and teach students to problem solve. DT offers a departure from traditional schooling and emphasizes knowledge construction oriented around real-world problems ([Bereiter and Scardamalia, 2006](#); [Koh et al., 2015](#)). A focus on DT helped build coherence in the school, and staff were provided with an array of professional learning opportunities to learn about and become expert in DT. DT now serves as a focal point and is embodied across the curriculum and in how staff engage in joint problem solving.

A structural element that supported DT was the creation of Design Lab electives. The principal encouraged teachers to create Design Labs built around their own knowledge and passion as well as what would foster student engagement and creativity. Students choose from a menu of innovative offerings, including robotics, architecture, business entrepreneurship, forensic analysis, rocketry, wearable technologies, among others. In these student-centered learning environments, students are exposed to DT. They

identify a problem, develop empathy, ideate, explore potential solutions, and create a prototype or model that displays a solution to that problem. Creativity, innovation, and critical thinking are prioritized, and the memorization and rote learning are eschewed.

While the school staff was committed to offering Design Labs, finding time in the schedule was initially a challenge. Organizational structures like the master schedule often serve as a barrier to change in secondary schools. To make space for Design Labs, staff decided students would have a rotating block schedule, attending odd and even periods on alternating days. The traditional 90 min blocks are split into 50 min Design Lab and 40 min lunch periods every day. Every student participates in a semester-long Design Lab class every day, which is uncommon at the middle school level where students needing remediation often have to forgo electives. An additional elective period for classes such as music or student government still exists as well.

At this school, DT is deliberately not marginalized within the Design Lab electives. Educators realized DT had to be integrated into core content classes, including math, English, science, and social studies. The DT framework allowed teachers and students to share a common language and common practices, engaging students in DT as they interrogate texts, address social problems, or engage in science projects. There is a shared sense of the importance of this instructional model among teachers and principals that allowed DT to spread and deepen. Coherency allows teachers to talk to one another about teaching strategies, and it provides consistency across the curriculum for students.

In everyone's estimation, the school has undergone enormous changes in the past 5 years. Undergoing such a dramatic transformation, however, has taken significant effort and tenacity. First, some teachers initially remained somewhat wedded to their traditional ways of teaching, and this challenge had to be overcome. Teachers were given freedom in how they engaged in change, but the nonnegotiable was they had to better support student learning. Opportunities to build teacher capacity were essential. Many teacher preparation programs do not teach DT or other nontraditional ways of organizing schooling. Educating parents in the community about the school and how it functions is an ongoing challenge. Since nontraditional teaching looks different, educators and parents have also considered students' preparation for high school and more traditional ways of teaching that involve note taking, testing, and considerably less flexibility. Ultimately, staff defended the structure and culture of their school against these challenges, believing it is more important to chart a new path than to prepare students for traditional, outdated learning settings. Fortunately, the district supported the school's efforts as well.

Summary

This case of school reform reveals how educators exercise agency and leadership to transform the structure and culture of the school. The principal was mindful that reform had to be a collective endeavor and that teachers were central to the change effort. The decision to orient the school around DT brought coherence in the educational program for students and teachers. It also served as a way for educators to problem solve when challenges arose. Given that various school features departed from the traditional grammar of schooling (Tyack and Cuban, 1995), it is not surprising the school's structural features received some pushback from teachers and parents accustomed to the usual ways of "doing school." Sustaining change thus requires deep commitment among powerful agents who are willing to stay the course.

Conclusion

The case studies we reviewed here show the value of depicting the reform policy process as a constructed project involving structural, cultural, and agentic dimensions. We learned implementing structures that support change, alongside a process that involves reculturing in which educators begin to share common beliefs about teaching, can facilitate the success of reform goals. These cases also show the actions of educators and leaders matter. The street-level bureaucrats tend to be neither necessarily compliant nor passive actors. Likewise, actors in positions of power are not necessarily seen as rationally and intentionally calculating courses of action and then directing their underlings to implement these plans. Instead, those at the top, those distributed throughout schools, and those at the bottom of bureaucracies are understood to be making organizational choices in situations that involve uncertainty.

By deploying a constructivist perspective, we exposed the impact reform coherence has when it is within and between structure, culture, and agency at *all* levels of the system. It contributes to the success of the policymaking process, and that process is characterized by continuous interactions among social actors within and between levels of organizational systems, such as what happens within and between schools and with district administration.

We reported on cases of social policy, as it was generated in face-to-face interactions among real people who were confronting real problems, in concrete social settings. In the cases we presented, social actors responded to policy directives in a variety of ways—initiating alternatives, advancing or sustaining directives, and resisting or actively subverting them. Most importantly, the agency of and interaction between participants in the educational policymaking process is perceived as being part of a complex dynamic, one that *shapes* and *is shaped by* the structural and cultural features of school and society.

References

- Bereiter, C., Scardamalia, M., 2006. Education and mind in the knowledge age. In: Alexander, P.A., Winne, P.H. (Eds.), *Handbook in Educational Psychology*, second ed. Lawrence Erlbaum, Mahwah N.J., pp. 695–713.
- Bryk, A.S., Schneider, B., 2003. *Trust in Schools: A Core Resource of Improvement*. Russell Sage, New York.
- Cohen, D.K., Spillane, J.P., Peurach, D.J., 2017. The dilemmas of educational reform. *Educ. Res.* 47 (3), 204–212. <https://doi.org/10.3102/0013189X17743488>.

- Cuban, L., 1998. How schools change reforms: redefining school success and failure. *Teach. Coll. Rec.* 99 (3), 153–177.
- Datnow, A., Park, V., 2019. *Professional Collaboration with Purpose: Teacher Learning for Equitable and Excellent Schools*. Routledge, London.
- Datnow, A., Hubbard, L., Mehan, H., 2002. Extending Educational Reform: From One School to Many. Routledge Falmer Press, London.
- Datnow, A., 2000. Power and politics in the adoption of school reform models. *Educ. Eval. Pol. Anal.* 22 (4), 357–374.
- Diamond, J.B., Spillane, J.P., 2016. School leadership and management from a distributed perspective: a 2016 retrospective and prospective. *Manag. Educ.* 30 (4), 147–154. <https://doi.org/10.1177/0892020616665938>.
- Fullan, M., Hargreaves, A., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press, New York.
- Fullan, M., Quinn, J., 2016. *Coherence: The Right Drivers in Action for Schools, Districts, and Systems*. Sage Publications, Thousand Oaks.
- Fullan, M., 1991. *The New Meaning of Educational Change*, second ed. Teachers College Press, New York.
- Hall, P.M., McGinty, P.J.W., 1997. Policy as the transformation of intentions: producing program from statutes. *Socio. Q.* 38 (3), p439–467. <https://doi.org/10.1111/j.1533-8525.1997.tb00487.x>.
- Hall, P.M., 2017a. Interaction, and the study of social organization. In: Burns, T.R., Hall, P.M. (Eds.), *The Meta-power Paradigm: Impacts and Transformations of Agents, Institutions, and Social Systems—Capitalism, State, and Democracy in a Global Context*, pp. 107–128. Frankfurt am Main: Peter Lang.
- Hall, P.M., 2017b. Meta-power, social organization, and the shaping of social organization. In: Burns, T.R., Hall, P.M. (Eds.), *The Meta-power Paradigm: Impacts and Transformations of Agents, Institutions, and Social Systems—Capitalism, State, and Democracy in a Global Context*, pp. 237–258. Frankfurt am Main: Peter Lang.
- Hargreaves, A., 1994. *Changing Teachers, Changing Times*. Teachers College Press, New York.
- Hubbard, L., Datnow, A., 2020. Design thinking, leadership, and the grammar of schooling: implications for educational change. *Am. J. Educ.* 126 (4) <https://doi.org/10.1086/709510>.
- Hubbard, L., Metz-Matthews, K., Barbara, R., 2022. Building coherence through whole school reform: the case of the AVID elementary program, *Journal of Education for Students Placed at Risk (JESPAR)*, 27 (1), 80–101. <https://doi.org/10.1080/10824669.2021.1981907>.
- Hubbard, L., Datnow, A., Prunty, L., 2014, September. Multiple initiatives, multiple challenges: The promise and pitfalls of implementing data. In: Schildkamp, K., Vanhoof, J. (Eds.), *International Studies in Educational Evaluation*, 42, pp. 54–62. Special Issue.
- Hubbard, L., Mehan, H., Stein, M.K., 2006. *Reform as Learning: When School Reform Collided with School Culture and Community Politics in San Diego*. Routledge Falmer, New York.
- Hubbard, L., 2019. *AVID Elementary (Report Prepared for AVID Center)*.
- Koh, J.H.L., Chai, C.S., Wong, B., Hong, H.Y., 2015. *Design Thinking for Education*. Springer, New York.
- Latino, C., 2001. Presentation to the San Diego City Schools Board of Education. San Diego, CA.
- Levinson, B., Sutton, M., 2001. *Policy as Practice: A Sociocultural Approach to the Study of Educational Policy*. Ablex, Boston.
- Madda, C., Halverson, R., Gomez, L., 2007. Exploring coherence as an organizational resource for carrying out reform initiatives. *Teach. Coll. Rec.* 109 (8), 1957–1979.
- Mehan, H., 2012. In the Front Door: Constructing a College-Going Culture of Learning. Paradigm Publishers, Boulder CO.
- Newmann, F.M., Smith, B., Allensworth, E., Bryk, A.S., 2001. School Instructional Program Coherence: Benefits and Challenges. Annenberg Research Project, Chicago. <https://consortium.uchicago.edu/publications/school-instructional-program-coherence-benefits-and-challenges>.
- Pressman, J.L., Wildavsky, A., 1973. *Implementation*. University of California Press, Berkeley.
- Ravitch, D., 1995. *Developing National Standards in Education*. U.S. Department of Education, Washington DC.
- Smith, L., Keith, P., 1971. *Anatomy of an Educational Innovation*. John Wiley and Sons, New York.
- Tucker, M.S., Coddling, J.B., 1998. *Standards for Our Schools: How to Set Them, Measure Them, and Reach Them*. Jossey-Bass, San Francisco.
- Tyack, D., Cuban, L., 1995. *Tinkering Toward Utopia: A Century of Public School Reform*. Harvard University Press, Cambridge, MA.
- Woods, P., Roberts, A., 2019. *Collaborative School Leadership: A Critical Guide*. Sage Publications, Thousand Oaks.

School leaders' use of research: viewing research use in decision making through an organizational lens

Lok-Sze Wong^a, Alice Huguet^b, Christopher W. Harrison^c, Cynthia E. Coburn^d, and James P. Spillane^d, ^aUniversity of North Texas, Denton, TX, United States; ^bRAND Corporation, Santa Monica, CA, United States; ^cMadison Metropolitan School District, Madison, WI, United States; and ^dNorthwestern University, Evanston, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

The decision making process in schools	235
The use of research in decision making	236
School level factors that shape research use in decision making	237
Formal and informal authority	237
Status	238
Organizational routines	238
District-school arrangements that shape school level decision making	238
Jurisdiction	239
District-school organizational routines	239
District-school informal social networks	239
Animating our conceptual framework	239
Conclusion	240
Acknowledgment	241
References	241

Calls for evidence based decision making have drawn attention to whether and how research is used in school level decision making. Researchers have found that a significant gap exists between the worlds of research and practice (Coburn and Stein, 2010). Some suggest that, in order for research to meaningfully impact practice, characteristics of the research itself must change. Such critics assert that research oversimplifies complex problems (Lather, 2004) or is not written in a manner that is easily digestible for practitioners (Rickinson et al., 2004; Vanderlinde and van Braak, 2010). Others suggest research should be more pertinent to topics that teachers and other practitioners care about (Vanderlinde and van Braak, 2010), or disseminated in a fashion more suited to reaching teachers and leaders (Cook et al., 2013). At the same time, still other studies of research use tell us that it is the practitioners who must change—that they must find the time or the motivation to consume research or would benefit from graduate level courses to deepen their capacity to understand and use research (Ion and Iucu, 2016).

While these suggestions may serve to lessen the gap between research and school level actors, they miss some important contextual layers that mediate practitioners' ability to engage with research. Perhaps we should not isolate the research as the central object, nor the researchers or the practitioners.

We expand on these factors by conceptualizing research use in schools using an *organizational* perspective. By organizational, we mean that we consider the complex contexts within which teachers and school leaders work, and how these contexts facilitate and hinder opportunities to bridge research and practice. Organizational theory delineates how factors inherent to organizations such as schools—and the systems they exist within—influence their functioning. We synthesize findings from existing studies and draw upon organizational theory to build a contextualized conceptual framework that aims to contribute to our understanding of whether or not—and why, and how—educators interact with research. In so doing, we answer the question: what organizational factors and district-school arrangements may influence school actors' use of research? In designing a multi-district study of educators' evidence use in education, our research team developed an original conceptual framework to guide our investigation of school level research use. In this article, we describe this framework and the literature that undergirds it.

To orient you to our conceptual framework, we begin with discussing the existing literature on school leaders' decision making and the use of research within. We then describe key elements of the broader organizational context that we see as influencing educators' decision making process. Understanding whether and how research is used in school level decision making first requires an understanding of the process of decision making within school settings. We then describe the key organizational dimensions that shape the decision making process, and how these dimensions can broaden or narrow the opportunities for educators to engage with research in school settings. Finally, we illustrate how these dimensions dynamically interact in the decision making process to mediate research use in schools. We conclude by suggesting ways that this framework may prove useful to the research community moving forward.

The decision making process in schools

We view decision making as a complex social process (Cohen et al., 1972; Hannaway, 1989; Kennedy, 1982; Majone, 1989; Weiss, 1988; Weiss and Bucuvalas, 1980). In schools, it can involve administrators, coaches, teachers, and other staff members

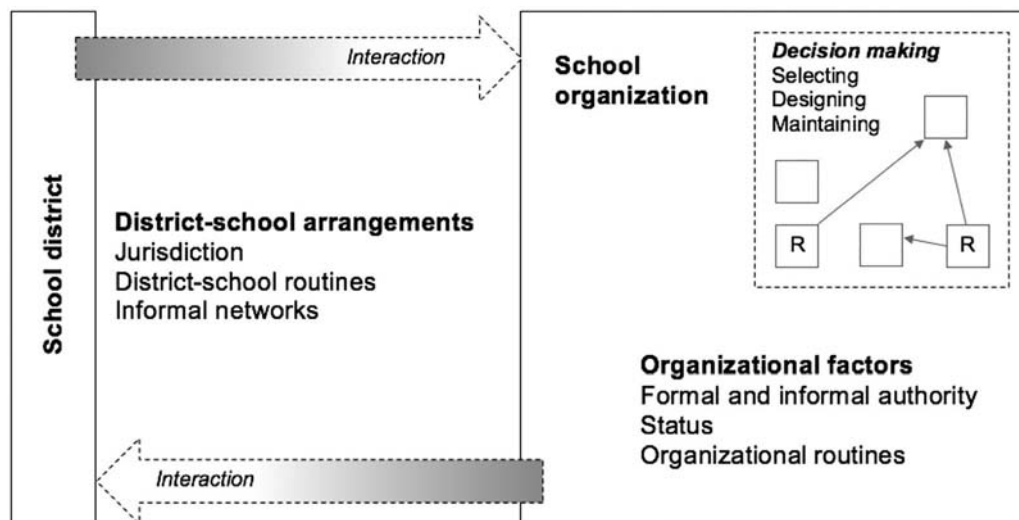


Fig. 1 Conceptual framework for research use in school-level decision making. Reprinted from Huguet et al. (2018).

(Spillane, 2006). Decision making activities include selecting, (re)designing, and maintaining policies and programs (Weiss and Bucuvalas, 1980). In the process of selecting, (re)designing, and maintaining, decision makers interact in a complex social dance, in which they engage with one another to “recommend, advise, confer, draw up budgets, testify, develop plans, write guidelines, report, supervise, propose legislation, assist, meet, argue, train, [and] consult” (Weiss, 1980, p. 163).

In this view of decision making, decisions are not singular momentary choices, but rather a process of accretion (Weiss, 1988). Decisions are a series of choices and actions connected over time across departments, grade levels, programs (e.g., English learners, special education, general education), and an organizational hierarchy (Hannaway, 1989; Kennedy, 1982; Majone, 1989). In our conceptual framework (see Fig. 1; Huguet et al., 2018, p. 626), the dashed box in the upper right corner of the school level represents the decision making process as it occurs within a school. Each smaller square within this box is a point of deliberation within the process, connected to one another through varied linkages, and not occurring in a predictable or linear manner. These squares represent both planned points of deliberation, such as meetings, and casual conversations in school hallways and classrooms.

Some decisions may seem to be made by an individual, such as a teacher making an instructional decision for their classroom. If we unpack the teacher’s decision making process, however, we see that even these decisions are influenced by interactions with other people inside and outside of their school. An individual teacher’s decision may be influenced by decisions made within the teacher’s team, school leaders’ instructional decisions, guidance provided by central office leaders, and events occurring in students’ communities (Gamoran et al., 2000).

The use of research in decision making

We understand research as an activity or product of an activity in which people employ systematic, empirical methods to answer a specific question. Research might be conducted by university researchers, but is also undertaken by others, such as district officials or practitioners within schools. Investigations may draw upon qualitative, quantitative, or mixed methods. Research can appear in books, academic journals, practitioner-oriented journals, and in policy and evaluation reports developed by researchers within a district (Penuel et al., 2016).

By focusing on research, we choose to exclude other forms of information often discussed in evidence use literature, such as data. The field of data driven decision making has a number of established frameworks from which researchers and practitioners can draw when designing inquiries (see Coburn and Turner, 2011; Marsh and Farrell, 2015). Yet there is no research use framework from which to do the same. There is a need for a conceptual framework on research use to structure studies in this area. Data use in school level decision making can be linked to research use. However, we chose to keep the categories separated to focus on research use in school level decision making.

We see research as a potentially central form of information that can contribute to decision making in schools. Research can inform or shape choices and actions among decision makers as decisions accrete over time. Indeed, studies of school leaders’ research use find that most claim to value and use research regularly (Levin, 2010; Saha, 2009), particularly when the research is perceived as relevant to their school (Vanderlinde and van Braak, 2010). In fact, some argue that research can be more relevant for decision making if educators are situated as knowledge creators and active partners in the research process, in addition to knowledge consumers (Hiebert et al., 2002; Stanovich and Stanovich, 2003; Vanderlinde and van Braak, 2010). Action research has, in particular, been positioned as a framework through which educators might approach systematic inquiry and reflective practice

(Cochran-Smith and Lytle, 1990, 2009; Hiebert et al., 2002). More recently, frameworks like design-based methods and improvement science have been forwarded as tools for guiding practitioners' efforts to engage in systematic inquiry (Anderson and Shattuck, 2012; Bryk, 2015; Cohen-Vogel et al., 2015).

In Fig. 1, we represent the role of research in decision making with an "R." At some points of deliberation, research may enter the conversation. The introduction of research may, but does not always, happen in a predictable way. When decisions are made to select, (re)design, or maintain educational policies and practices in schools, the process is most often a social, irregular one that occurs over a period of time as individuals deliberate in varied groupings.

Studies have found multiple factors that mediate research use in daily school level decision making. To begin, school leaders and teachers often lack easy access to research, as well as the time and support needed to read and reflect on the pertinence of the research for their schools (Hemsley-Brown and Oplatka, 2005; Sheard and Sharples, 2016; Williams and Coles, 2007). Further, while some leaders do have access to research through journals and books, they also rely on secondary sources such as colleagues and professional meetings to highlight worthwhile research (Biddle and Saha, 2006; Levin, 2010). Indeed, many school leaders report personal experience and the input of colleagues to be more powerful influences on their beliefs than research (Levin et al., 2011). Additionally, school leaders and teachers stated the nature and format of research can effect its use, such as its communication quality and its timeliness (Cousins and Leithwood, 1993; Levin et al., 2011; West and Rhoton, 1994). For instance, in Vanderlinde and van Braak's study (2010, p. 306), principals shared that "... many researchers speak and write in a language that is unintelligible to practitioners", and that the use of such technical language may impede principals' efforts to use research. Moreover, as shared above, educators' perception of the validity of research and its relevance to their professional practice mediate their use of research in their decision making (Bickel and Hattrup, 1995; Boardman et al., 2005; Cochran-Smith and Lytle, 1990; Cousins and Leithwood, 1993; Levin et al., 2011; West and Rhoton, 1994). These findings indicate that leaders prefer research with clear, practical suggestions about what works (Vanderlinde and van Braak, 2010). To add to this, Coburn and Talbert (2006) find that principals are more likely than teachers and central office leaders to be skeptical of the value of research, and to prefer multiple forms of evidence. The authors surmise that principals' responsibilities "place demands on them to bring coherence among multiple kinds of validity standards represented in various assessments and data" (Coburn and Talbert, 2006, p. 485).

Additionally, a few studies point to organizational characteristics that also mediate research use in decision making (Cousins and Leithwood, 1993; Levin et al., 2011; Sheard and Sharples, 2016). For instance, Cousins and Leithwood (1993) find that an organization's information needs, focus for improvement, political climate, and competing information shaped whether and how school leaders used research. Research is also more likely to be used when organizational structures, processes, and culture support it (Levin et al., 2011; Sharp et al., 2006), such as when school leaders are explicitly guided on how to apply research, or when they have knowledgeable mentors who help create research engaged environments. However, as we discuss next, research use is likely shaped by other organizational factors not yet examined in this literature.

School level factors that shape research use in decision making

Building on the limited existing literature on how organizational context shapes the use of research in school level decision making, we synthesize existing organizational literature to highlight several important factors that likely shape school level decisions. We understand schools to be complex social organizations. The core activities in a school are managed by both formal and informal leaders (Spillane and Zuberi, 2009; Sun et al., 2013) who jointly select, (re)design, and maintain a school's policies, internal structure, culture, and routines (Scherer and Spillane, 2011). Here, we focus on three organizational factors that shape that process—(1) authority, (2) status, and, (3) organizational routines.

Formal and informal authority

Decision making, and the role of research and other information therein, is influenced by formal and informal authority relations. For instance, school leaders and teachers with authority can intentionally or unintentionally shape the decision making process (Wong, 2019). This includes bringing information such as research to the decision making process and influencing how their colleagues use that information during their deliberations (Coburn, 2005).

Formal authority is power "coded into structural design" (McAdam and Scott, 2005, p. 10) as comes with hired positions. We often think of formal authority at the school level belonging to principals. Indeed, studies exploring research use at schools often reference this authoritative position. Levin (2010) review of the literature, for instance, found that across studies, school level research use was influenced by social setting, and, in his account, principals were a key influence on that setting. Sharp et al. (2006) also found that school staff recognized when their leaders valued and held knowledge about research, suggesting that leaders' knowledge may be a strong predictor of research use throughout the school. In a mixed method study of principals in Australia and the United States, Saha (2009) found that principals passed on their own research literacy skills to teachers at their schools. Another way to state this could be that school staff are exposed to research through their principals' knowledge—or lack thereof—emphasizing the importance of this formal position. One could imagine a principal being a gatekeeper, bringing research into the organization and influencing staff's exposure, and an agenda setter, calling for particular research use in meetings

or professional learning opportunities. In this way, the principal can influence the flow of research into the school as well as its use within the school. At the same time, they may also serve as a limiting factor, choosing only research that fits with their individual preferences.

However, administrators are not the only actors to exert influence on schools' decision making processes. Any effort to understand school level decision making must go beyond the principal to explore the role of informal leaders in the process. These individuals may have informal authority, or authority ascribed to an actor related to characteristics other than their officially recognized position (Scott and Davis, 2007). Informal authority belongs to those who actually take responsibility in practice, which may deviate from official arrangements. Individuals who have no formal leadership designation, but may have informal authority, may influence the role of research in decision making. While this has not yet, to our knowledge, been studied in relationship to research use at schools, we know that informal authority does carry influence in schools. Evidence suggests that instructional decision making at the school level, in fact, involves an array of actors (Gronn, 2003; Leithwood, 2007), indicating the need to look beyond formal authority as an influence in decision making and research use.

Status

Similarly, people with higher status in a group might find their perspectives privileged during decision making, creating imbalances in interpretive contexts (Coburn et al., 2008) and allowing them to hold sway in negotiations (Camburn et al., 2003; Leithwood and Mascall, 2008; Spillane et al., 2007). Educators with less experience, for instance, may defer to the judgment of their more experienced colleagues on how to use and interpret research. Status refers to power associated with "unobserved performance expectations" (Balkwell, 1994, p. 124) that often result in prestige and may influence interactions within groups. Status is typically negotiated in a social situation, and can be rooted in fundamental characteristics (i.e., race, gender, or age) or achieved ones (i.e., academic background or occupational history) (Coburn et al., 2008; Dornbusch and Scott, 1975). Though not explicit, there are echoes of status in existing school research use literature. In a study by Sheard and Sharples (2016), the authors partnered with three English primary schools on a process of research engagement. Among other factors, they found that "recognition of professional experience and expertise" (p. 682) among the involved school actors contributed to a climate more supportive of in-depth research use. In this example, status led to a culture more conducive to research use. One could imagine status working in other ways; an individual with a prestigious background, for instance, could influence which, when, and how research gets interpreted.

Organizational routines

Organizational routines are another school level factor that shapes decision making. Organizational routines are "repetitive, recognizable pattern[s] of interdependent actions, involving multiple actors" (Feldman and Pentland, 2003, p. 95). In schools, routines related to decision making might include such activities as school improvement planning, designing and facilitating professional development, leadership team meetings, or instructional rounds, among countless others. These routines are both mediums for and outcomes of decision making; they can emerge from everyday practice or can be intentionally designed by school and district decision makers (Cohen and Bacdayan, 1994; Pentland and Feldman, 2008). Some routines are developed at the school site, while others are district mandated.

Organizational routines are organizational structures that shape who collaborates with whom about what, determining which individuals participate in decision making and what research and other information can be brought to the table during a discussion. Investigations into school level research use commonly—though not always overtly—demonstrate that organizational structure has the potential to improve school actors' use of research (Datnow and Park, 2010; Sharp et al., 2006; Sheard and Sharples, 2016). According to Levin (2010, p. 312), leaders who want to build a culture of research use at their schools should consider organizational structure: "The very fact that research findings and implications have a place on the agendas of staff meetings or principal meetings changes the way people in the organization think about research and, if done in ways respectful of professional responsibility, is likely to increase receptivity." Yet, studies have not investigated research use explicitly in light of routines. To understand research's role in schools, it is important to pay attention to routines as they relate to decision making practices.

District-school arrangements that shape school level decision making

Schools are components of broader school systems, and system leaders can influence school leaders' and teachers' practice, including influencing the use of research in school level decision making.¹ School level decision making is a phenomenon that can include stakeholders and activities across the school system, involving people and policies from district central offices, other schools, community members, Board members, and other organizations and stakeholders. Here, we focus on how district-school arrangements can shape research use in school level decision making. District central offices can exert varying degrees of control over

¹The relationship between schools and districts has bidirectional influence (Honig and Venkateswaran, 2012). However, since our focus is on school level decision making, we do not specify in this framework ways that schools may influence districts.

the decisions that schools have a say in, the research that school leaders have access to, how school leaders interpret and utilize research, and the viable options school leaders can consider (Honig and Venkateswaran, 2012; Ikemoto and Marsh, 2007). Thus, we highlight three district-school arrangements—(1) jurisdiction, (2) routines, and (3) informal social networks—to understand how these arrangements can shape the role of research in school level decision making.

Jurisdiction

Districts differ in the domains over which central office administrators have authority, and which aspects are controlled by school administrators. That is, districts vary in patterns of jurisdiction; in one, schools may be responsible for the majority of their instructional decisions, while in another, the central office makes the preponderance of calls related to instruction. Jurisdiction over decision making may also vary within school districts, where different schools may be given varying levels of autonomy. For example, some districts grant school leaders more or less autonomy based on student achievement scores.

Jurisdiction may be formal or informal. For instance, while a district may officially have authority over one aspect of instruction, schools may informally make decisions about it themselves. We also found the opposite in a recent analysis (Wong et al., 2020). Even though school leaders in two site-based decision making districts had formal jurisdiction over most aspects of instruction, their central office leaders' successfully used persuasive strategies to encourage them to make decisions in line with central office preferences. Further, there may be ambiguity with respect to which level has jurisdiction over a given issue. In some districts, the central office might be responsible for selecting research, compiling summaries, and facilitating meetings in which research is discussed (Levin et al., 2011). Schools in such districts would potentially gain access to research that they may not otherwise have had, but the district would in this case be responsible for narrowing the content to which school actors were exposed. This stands in contrast to schools where professional development is almost exclusively managed at the school site. Jurisdiction may also be influential in less direct ways—for instance, who has the final word about curriculum selection, or who designs discipline policies, both cases in which research may arise.

District-school organizational routines

Similar to school level organizational routines, district-school organizational routines are opportunities for central office and school leaders to collaborate and share information such as research. District-school organizational routines influence which central office and school staff are involved in discussions, what research and other information are used, and in what ways (Coburn and Turner, 2011; Little, 2012). For instance, in districts where central office leaders do have jurisdiction over professional development, the planning and delivery of the training are examples of routines that connect the school with the district central office. If individuals from both of those levels work with one another at different steps of the process, there is further opportunity for access and exposure to research. Of course, training is only one illustration of a routine connecting the school and district. Other examples of connecting routines include teacher evaluation processes, district walkthroughs, cross-district meetings for teacher leaders or school based coaches, or principal professional development sessions.

District-school informal social networks

District-school connections may also occur through informal social networks, opportunities for central office and school leaders to reach out to each other as needed. For example, school leaders may reach out to central office leaders for advice on a decision (Wong et al., 2020). The research and other information central office leaders offer can then shape school level decision making processes. Ties that span boundaries allow for new information to enter each organization and minimize "groupthink" among organizational members (Burt, 2000; Hansen, 1999; Krackhardt and Stern, 1988). The greater the trust built within these networks, the more willing participants may be to share and exchange information in situations that require them to be vulnerable and take risks. When connections between central office and schools are weak, they tend to handicap schools' use of research (Daly and Finnigan, 2010). The nature of information passed through a network, and the likelihood that it will influence decision making at the school level, may depend upon the structure of social networks beyond the formal connections already discussed here. School actors may be exposed to research through a friend or past colleague, or a district actor with whom they share an informal association.

Animating our conceptual framework

In this section, we bring together the key concepts from our framework to illustrate how they collectively and dynamically interact as dimensions of the decision making process. Our conceptual framework utilizes an organizational lens to provide a contextualized approach for understanding research use in schools—an approach that we currently use in our own studies. We provide an extended example to illustrate how these organizational dimensions of schools and school systems interact to influence the role that research plays in school level decision making. While the anecdote is hypothetical, it is rooted in composite experiences our team has observed throughout various studies, combined to illustrate the application of this conceptual framework.

Consider this scenario: an elementary school leadership team must decide what math curriculum they should adopt in the upcoming school year. They have a number of options on the table and, given that this is a potentially expensive and impactful

decision, the topic is introduced and reintroduced in full staff meetings, at math department meetings, and in schoolwide leadership meetings. This curricular decision is also discussed during informal conversations. During some of these deliberations, research is introduced. At one point, the math department chair introduces research that she was given during an evening graduate level class, which appears to support one program over the others. A third grade teacher searches online for expert reviews on the potential options available, then mentions them to colleagues over lunch. Over the course of months, the conversation begins to take shape around a particular set of ideas, some informed by research. This seemingly messy process exists, in our diagram, in the dashed box in the upper right corner.

The process is influenced by a number of factors within the organization. These elements in our diagram are present at the school level—the larger box on the right. Formal and informal authority, status, and organizational routines each have the potential to influence the role research plays in these processes. For instance, in our hypothetical example, authority may determine to some degree which research is introduced into deliberations about mathematics curriculum. The principal, for instance, may include time on a staff meeting agenda to discuss opinions about new math curriculum, including the introduction of an article she believes to be relevant. Given her position as principal, one which is imbued with formal authority, the research she introduces is not questioned publicly, though a number of teachers may find flaws with its premise. Perhaps while discussing this particular piece of research in the staff meeting, the principal's interpretation of how its findings apply to their selection of curriculum is either the only interpretation discussed, or the one given the most attention. In these ways and myriad others, an individual in a position of authority may influence which research reaches deliberations, and how it is interpreted.

Other organizational factors—such as status and organizational routines—also play a role in how research may be used at the school level. For instance, in our hypothetical example, a teacher with a graduate degree may leverage her status to challenge the principal's interpretation of research, publicly or in private. Further, the principal's inclusion of research on the staff meeting agenda is one instantiation of how research use may be supported through the employment of organizational routines. But, as the organization is porous and influenced by elements outside its own boundaries, the school's use of research in decision making is also permeable, meaning that research can enter through other means.

Our conceptual framework focuses on three particularly generative ways in which research use may be influenced. These are arrangements between the school and the central district office: jurisdiction, district-school organizational routines, and district-school informal social networks. Jurisdiction influences research use at a number of points in our hypothetical example of school level decision making. If the school district, for instance, mandated a particular mathematics curriculum, the question of selection would be moot. Research regarding mathematics curricula may not be pulled into the school's decision making arena at all, in this case. In addition, it is possible that the district office could mandate a particular professional development series, therefore limiting the degree to which the principal could introduce her own selection of research to the staff at large. Organizational routines that connect the school and district, such as the planning and implementation of professional development, might favor a different set of research, based on the training and preferences of the district level team involved. Finally, informal social networks that connect school level actors with those in the district serve as another channel for research—and interpretations of research—to enter the deliberative field. These social networks are again a means for sharing information that may influence the types of research on the table and how they are leveraged in discussions. Again, the examples provided above are compiled from a variety of research experiences and are selected for illustrative purposes. It is not often that such clean cut instances of research use arise in reality. Although we focus in this article on the school level, the bidirectional arrows between the school and the district serve as recognition that interactions with the school can influence district level deliberations, as well.

Conclusion

To the conversation on research use in schools, we offer a view of decision making that is social and additive, an organizational lens, and most tangibly a framework through which to further explore this pressing topic. Our team is currently using this conceptual framework in a multi-city, multi-year study of research use in districts and schools. Though we anticipate new and improved iterations will follow in future investigations—ours as well as others'—our hope is to set the stage for deeper explorations of research use at the school level, remaining cognizant to factors that influence the processes of decision making in which research plays a role.

Our framework opens up school level decision making to include the impact that organizational context can have on whether and how educators use research within their decision making. District-school and school level organizational routines, informal social connections, jurisdiction, status, and authority all mediate whether educators learn about research and how, if at all, they use the research in their decision making. In addition to individual and access factors, we can examine these organizational factors to understand and better support research use in schools. Viewing research use through an organizational lens helps us situate and contextualize whether and how research is used in instructional decision making.

This organizational framework is a critical addition to the existing conversation on research use in schools, and a useful one for pressing forward in our understanding of school level decision making and the role of research therein. To separate the decision making process from school and district organizational influences misses elemental components of the story. In presenting this framework, we make the argument that consideration of organizational factors should not be incidental to any study of decision making, but should instead play a central, framing role from the inception of the investigation.

Further studies are necessary as current policies continue to press educational practitioners toward evidence-based decision making. We argue that the school level is becoming increasingly important in this conversation, as policy changes—such as the

Every Student Succeeds Act in the US—signal a move to more on-the-ground autonomy in instructional choices. If this policy is taken as an indicator, we are entering an era in which school level decision making will be increasingly taking center stage. We end with a call to our research colleagues to usher it in with the goal of understanding how organizational dimensions play a role in this important process.

Acknowledgment

This work has been supported by the Institute of Education Sciences, US Department of Education, through Grant RC305C140008. The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education.

References

- Anderson, T., Shattuck, J., 2012. Design based research: a decade of progress in education research? *Educ. Res.* 41 (1), 16–25.
- Balkwell, J.W., 1994. Status. *Group Processes: Sociological Analyses*. Wadsworth, Belmont, CA, pp. 119–148.
- Bickel, W.E., Hattrup, R.A., 1995. Teachers and researchers in collaboration: reflections on the process. *Am. Educ. Res. J.* 32 (1), 35–62.
- Biddle, B.J., Saha, L.J., 2006. How principals use research. *Educ. Leader* 63 (6), 72–77.
- Boardman, A.G., Argüelles, M.E., Vaughn, S., Hughes, M.T., Klingner, J., 2005. Special education teachers' views of research-based practices. *J. Spec. Educ.* 39 (3), 168–180.
- Bryk, A.S., 2015. 2014 AERA distinguished lecture: accelerating how we learn to improve. *Educ. Res.* 44 (9), 467–477.
- Burt, R.S., 2000. The network structure of social capital. *Res. Organ. Behav.* 22, 345–423.
- Camburn, E., Rowan, B., Taylor, J.E., 2003. Distributed leadership in schools: the case of elementary schools adopting comprehensive school reform models. *Educ. Eval. Pol. Anal.* 25 (4), 347–373.
- Coburn, C.E., Stein, M.K., 2010. *Research and Practice in Education: Building Alliances, Bridging the Divide*. Rowman & Littlefield, Lanham, MD.
- Coburn, C.E., Talbert, J.E., 2006. Conceptions of evidence use in school districts: mapping the terrain. *Am. J. Educ.* 112 (4), 469–495.
- Coburn, C.E., Turner, E.O., 2011. Research on data use: a framework and analysis. *Measurement* 9 (4), 173–206.
- Coburn, C.E., Bae, S., Turner, E.O., 2008. Authority, status, and the dynamics of insider–outsider partnerships at the district level. *Peabody J. Educ.* 83 (3), 364–399.
- Coburn, C.E., 2005. Shaping teacher sensemaking: school leaders and the enactment of reading policy. *Educ. Pol.* 19 (3), 476–509.
- Cochran-Smith, M., Lytle, S.L., 1990. Research on teaching and teacher research: the issues that divide. *Educ. Res.* 19 (2), 2–11.
- Cochran-Smith, M., Lytle, S.L., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York.
- Cohen, M.D., Bacdayan, P., 1994. Organizational routines are stored as procedural memory: evidence from a laboratory study. *Organ. Sci.* 5 (4), 554–568.
- Cohen, M.D., March, J.G., Olsen, J.P., 1972. A garbage can model of organizational choice. *Adm. Sci. Q.* 17 (1), 1–25.
- Cohen-Vogel, L., Tichnor-Wagner, A., Allen, D., Harrison, C., Kainz, K., Socol, A.R., Wang, Q., 2015. Implementing educational innovations at scale: transforming researchers into continuous improvement scientists. *Educ. Pol.* 29 (1), 257–277.
- Cook, B.G., Cook, L., Landrum, T.J., 2013. Moving research into practice: can we make dissemination stick? *Except. Child.* 79 (2), 163–180.
- Cousins, J.B., Leithwood, K.A., 1993. Enhancing knowledge utilization as a strategy for school improvement. *Sci. Commun.* 14 (3), 305–333.
- Daly, A.J., Finnigan, K.S., 2010. A bridge between worlds: understanding network structure to understand change strategy. *J. Educ. Change* 11 (2), 111–138.
- Datnow, A., Park, V., 2010. Large-scale reform in the era of accountability: the system role in supporting data-driven decision making. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*, Springer International Handbooks of Education, vol. 23. Springer Netherlands, Dordrecht.
- Dornbusch, S.M., Scott, W.R., 1975. *Evaluation and the Exercise of Authority*. Jossey-Bass, San Francisco, CA.
- Feldman, M.S., Pentland, B.T., 2003. Reconceptualizing organizational routines as a source of flexibility and change. *Adm. Sci. Q.* 48 (1), 94–118.
- Gamoran, A., Secada, W.G., Marrett, C.B., 2000. The organizational context of teaching and learning. In: Hallinan, M.T. (Ed.), *Handbook of the Sociology of Education*. Springer, New York, pp. 37–63.
- Gronn, P., 2003. Leadership: who needs it? *Sch. Leader. Manag.* 23 (3), 267–290.
- Hannaway, J., 1989. *Managers Managing: The Workings of an Administrative System*. Oxford University Press, Oxford.
- Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. *Adm. Sci. Q.* 44 (1), 82–111.
- Hemsley-Brown, J., Oplatka, I., 2005. Bridging the research-practice gap: barriers and facilitators to research use among school principals from England and Israel. *Int. J. Public Sect. Manag.* 18 (5), 424–446.
- Hiebert, J., Gallimore, R., Stigler, J.W., 2002. A knowledge base for the teaching profession: what would it look like and how can we get one? *Educ. Res.* 31 (5), 3–15.
- Honig, M.I., Venkateswaran, N., 2012. School–central office relationships in evidence use: understanding evidence use as a systems problem. *Am. J. Educ.* 118 (2), 199–222.
- Huguet, A., Wong, L., Harrison, C.W., Coburn, C.E., Spillane, J.P., 2018. Research use in schools: a framework for understanding research use in school-level decision making. In: Connolly, M., Eddy-Spicer, D.H., James, C., Kruse, S.D. (Eds.), *The SAGE Handbook of School Organization*. SAGE Publications, London, pp. 618–631.
- Ikemoto, G.S., Marsh, J.A., 2007. Chapter 5: cutting through the “data-driven” mantra: different conceptions of data-driven decision making. *Yearbk. Natl. Soc. Stud. Educ.* 106 (1), 105–131.
- Ion, G., Iucu, R., 2016. The impact of postgraduate studies on the teachers' practice. *Eur. J. Teach. Educ.* 39 (5), 602–615.
- Kennedy, M.M., 1982. Evidence and decision. In: Kennedy, M.M. (Ed.), *Working Knowledge and Other Essays*. Huron Institute, Cambridge, MA, pp. 59–103.
- Krackhardt, D., Stern, R.N., 1988. Informal networks and organizational crises: an experimental simulation. *Soc. Psychol. Q.* 5 (2), 123–140.
- Lather, P., 2004. Scientific research in education: a critical perspective. *Br. Educ. Res. J.* 30 (6), 759–772.
- Leithwood, K., Mascall, B., 2008. Collective leadership effects on student achievement. *Educ. Adm. Q.* 44 (4), 529–561.
- Leithwood, K., 2007. What we know about educational leadership. In: Burger, J.M., Webber, C.F., Klinck, P. (Eds.), *Intelligent Leadership: Studies in Educational Leadership*, vol. 6. Springer Netherlands, Dordrecht.
- Levin, B., Cooper, A., Arjomand, S., Thompson, K., 2011. *Research Use and its Impact in Secondary Schools: Exploring Knowledge Mobilization in Education*. Canadian Education Association, Toronto.
- Levin, B., 2010. Leadership for evidence informed education. *Sch. Leader. Manag.* 30 (4), 303–315.
- Little, J.W., 2012. Professional community and professional development in the learning centered school. In: Kooy, M., van Veen, K. (Eds.), *Teacher Learning that Matters: International Perspectives*. Routledge, London, pp. 22–46.
- Majone, G., 1989. *Evidence, Argument, and Persuasion in the Policy Process*. Yale University Press, New Haven, CT.
- Marsh, J.A., Farrell, C.C., 2015. How leaders can support teachers with data-driven decision making: a framework for understanding capacity building. *Educ. Manag. Adm. Leader* 43 (2), 269–289.

- McAdam, D., Scott, W.R., 2005. Organizations and movements. In: Davis, G.F., McAdam, D., Scott, W.R., Zald, M.N. (Eds.), *Social Movements and Organizational Theory*. Cambridge University Press, Cambridge, UK, pp. 4–40.
- Pentland, B.T., Feldman, M.S., 2008. Designing routines: on the folly of designing artifacts, while hoping for patterns of action. *Inf. Organ.* 18 (4), 235–250.
- Penuel, W.R., Briggs, D.C., Davidson, K.L., Herlihy, C., Sherer, D., Hill, H.C., Farrell, C.C., Allen, A.R., 2016. Findings from a National Study on Research Use Among School and District Leaders (Technical Report No. 1). National Center for Research in Policy and Practice, Boulder, CO.
- Rickinson, M., Clark, A., McLeod, S., Poulton, P., Sargent, J., 2004. What on earth has research got to do with me? *Teach. Dev.* 8 (2/3), 201–220.
- Saha, L.J., 2009. The dissemination of knowledge about research on teachers, to the teachers. In: Saha, L., Dworkin, A.G. (Eds.), *International Handbook of Research on Teachers and Teaching*. Springer, New York, pp. 71–79.
- Scott, W.R., Davis, G.F., 2007. *Organizations and Organization: Rational, Natural, and Open Systems Perspectives*. Pearson Prentice-Hall, Upper Saddle River, NJ.
- Sharp, C., Eames, A., Sanders, D., Tomlinson, K., 2006. *Leading a Research-Engaged School*. NCSL, Nottingham, UK.
- Sheard, M.K., Sharples, J., 2016. School leaders' engagement with the concept of evidence-based practice as a management tool for school improvement. *Educ. Manag. Adm. Leader* 44 (4), 668–687.
- Sherer, J.Z., Spillane, J.P., 2011. Constancy and change in work practice in schools: the role of organizational routines. *Teach. Coll. Rec.* 113 (3), 611–657.
- Spillane, J.P., Zuberi, A., 2009. Designing and piloting a leadership daily practice log: using logs to study the practice of leadership. *Educ. Adm. Q.* 45, 375–423.
- Spillane, J.P., Camburn, E.M., Stitzel Pareja, A., 2007. Taking a distributed perspective to the school principal's workday. *Leader. Pol. Sch.* 6 (1), 103–125.
- Spillane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco, CA.
- Stanovich, P.J., Stanovich, K.E., 2003. *Using Research and Reason in Education: How Teachers Can Use Scientifically Based Research to Make Curricular and Instructional Decisions*. National Institute for Literacy, Washington, DC.
- Sun, M., Frank, K., Penuel, W., Kim, C., 2013. How external institutions penetrate schools through formal and informal leaders. *Educ. Adm. Q.* 49, 610–644.
- Vanderlinde, R., van Braak, J., 2010. The gap between educational research and practice: views of teachers, school leaders, intermediaries and researchers. *Br. Educ. Res. J.* 36 (2), 299–316.
- Weiss, C.H., Bucuvalas, M.J., 1980. *Social Science Research and Decision-Making*. Columbia University Press, New York.
- Weiss, C.H., 1980. Knowledge creep and decision accretion. *Knowledge* 1 (3), 381–404.
- Weiss, C.H., 1988. Evaluation for decisions: is anybody there? Does anybody care? *Eval. Pract.* 9 (1), 5–19.
- West, R.F., Rhoton, C., 1994. School district administrators' perceptions of educational research and barriers to research utilization. *ERS Spectr.* 12 (1), 23–30.
- Williams, D., Coles, L., 2007. Teachers' approaches to finding and using research evidence: an information literacy perspective. *Educ. Res.* 49 (2), 185–206.
- Wong, L., Coburn, C.E., Kamel, A., 2020. How central office leaders influence school leaders' decision making: unpacking power dynamics in two school-based decision-making systems. *Peabody J. Educ.* 95 (4).
- Wong, L., 2019. Administrators' unintentional sensegiving and system reform outcomes. *Educ. Pol. Anal. Arch.* 27 (3).

School leadership and organizational factors: evaluating pathways to student learning

James Sebastian and Seijoon Park, Educational Leadership and Policy Analysis, University of Missouri, Columbia, MO, United States

© 2023 Elsevier Ltd. All rights reserved.

Essential supports theoretical framework	244
School leadership studies in Chicago Public Schools (CPS)	244
Results of the Chicago Consortium leadership studies	245
Disentangling pathways from leadership to student learning	246
Discussion and implications	246
Appendix A. Examination of mediational studies reviewed by	249
Acknowledgments	252
References	252

Principal leadership has consistently been regarded as a critical factor in influencing student learning and school improvement (Grissom et al., 2021; Hallinger et al., 1996; Hallinger and Heck, 1996b, 1998; Leithwood et al., 2004; Witziers et al., 2003). A review of research on leadership published by the Wallace Foundation in 2004 (Leithwood et al., 2004), claimed that “Leadership is second only to classroom instruction among all school-related factors that contribute to what students learn at school” (p. 5). A subsequent publication by the same research team (see Louis et al., 2010) strengthened this claim, noting that “After six additional years of research, we are even more confident about this claim. To date we have not found a single case of a school improving its student achievement record in the absence of talented leadership” (p. 9). A more recent study commissioned by the Wallace foundation and authored by Grissom et al. (2021), noted that “Effective principals are at least as important for student achievement as previous reports have concluded—and in fact, their importance may not have been stated strongly enough” (p. xviii). This review further noted that “it is difficult to envision an investment with a higher ceiling on its potential return than a successful effort to improve principal leadership” (p. xiv).

The initial Wallace report that reviewed prior leadership studies noted that leadership had small but statistically significant effects on student outcomes. Other reviews of school leadership studies have reached similar conclusions (see for example, Hallinger, 2005; Hallinger and Heck, 1998; Liebowitz and Porter, 2019; Neumerski, 2012). Hallinger (2005) notes that “The preponderance of evidence indicates that school principals contribute to school effectiveness and student achievement indirectly through actions they take to influence school and classroom conditions.” (p. 9), and that “The size of the effects that principals indirectly contribute toward student learning, though statistically significant, is also quite small.” (p. 9). More recently, Tan et al. (2020) conducted a second-order meta-analysis of primary studies that spanned four decades of research and found small overall mean effects size for school leadership. The description of leadership effects as small is based on Cohen’s criteria by which an effect size of 0.20 is characterized as small, 0.50 as medium, and 0.80 as large (Cohen, 1962). A different way to interpret these effects is to compare them with other organizational factors or educational interventions. Grissom et al. (2021) for example, compared leadership effects with teacher impacts and found that “the impact of having an effective principal on student achievement is nearly as large as the effect of having a similarly effective teacher” (p. xiii).

Besides the finding that leadership effects can be characterized as small (according to Cohen’s criteria), there is also consistent evidence that these effects mostly occur indirectly (Neumerski, 2012), through mediating processes such as goal setting and maintaining or changing the school’s mission and vision (Hallinger et al., 1996; Leithwood and Jantzi, 2008; Louis et al., 2010; Supovitz et al., 2010), teacher development and support (Leithwood, 1994; Marfan and Pascual, 2018; Shen et al., 2016), organizational management (Leithwood and Jantzi, 2008), instructional and program coherence (Sebastian and Allensworth, 2012), parental involvement (Sebastian et al., 2016; Shen et al., 2016), and school climate (Dutta and Sahney, 2016; Sebastian et al., 2014, 2017). Hallinger and Heck (1998) note that “we conclude that convincing empirical evidence suggests that principals do affect school outcomes. The impact is, however, achieved primarily through complex indirect relationships with other external environmental forces (e.g., community) and in-school processes” (p. 185). It is important to note here that a few studies have argued that leadership can also have direct effects (Hallinger and Heck, 1998; Silva et al., 2011; Wu, 2020; Wu et al., 2020).

Researchers interested in studying indirect effects of school leadership typically use Structural Equation Modeling (SEM) to examine the role of specific organizational features in mediating leadership effects (see for example, Hallinger and Heck, 1996a,b, 2010a; Heck and Hallinger, 2009; Sebastian and Allensworth, 2012; Sebastian et al., 2016, 2017). Researchers have examined a wide array of mediational factors to understand school leadership effects, depending on their conceptualizations of leadership and theoretical perspectives about school organization. For example, Leithwood et al. (2010) and (2020) classified four unique paths through which principals influence learning: (a) rational path, that includes foundational aspects of schooling such as instruction and classroom practice, academic press, and disciplinary climate; (b) emotional path, which contains job satisfaction, teacher

academic optimism and expectation, trust, teacher collaboration, and teacher self and collective efficacy; (c) organizational path, which includes safe and orderly environment, school organization aspects, and capacity for school improvement, and (d) family path, which includes variables such as parent and parent involvement. Wu (2020) used Leithwood's criteria to review leadership studies that have used the indirect-effects approach and noted that the most commonly examined set of organizational variables in mediational studies were classroom instruction, teacher collective efficacy, professional learning community, and school climate.

So far, in prior school leadership research, there is little discussion about which component of leadership mediational pathways results in overall small effects from leadership to student outcomes. In general, if we consider any path from school leadership (X) to learning (Y) via an organizational mediating process (M), the total indirect effect (X-M-Y), involves two components, school leadership to the mediator (X-M), and then mediator to outcome (M-Y). It is important to examine these two components to understand where mediational pathways are diminished. For example, the finding that a specific leadership pathway to any school outcome (e.g., teaching/learning) is not statistically significant, could be due to X-M or M-Y, or both. Similarly, if an overall indirect pathway via a specific mediator is significant but small, it could be that X-M is small, but M-Y is large, or vice-versa. Alternately, all components of the mediational pathway could be small. In the present chapter, we review evidence from a large urban school district—Chicago Public Schools (CPS), using longitudinal data and an array of organizational mediators to disentangle the relationships between school leadership, organizational processes, classroom instruction, and student learning outcomes. In the following sections, we first summarize the Essential Supports framework that we use as the conceptual basis for the study, we next review the main findings from a series of school leadership studies conducted using CPS data, and then examine specific components of the mediational pathways examined in these studies.

Essential supports theoretical framework

As noted earlier, school leadership researchers have examined a wide array of school organizational factors depending on their theoretical frameworks and definitions of leadership. Hitt and Tucker (2016) conducted a review of school leadership studies and highlighted three leadership frameworks as noteworthy—the Ontario Leadership Framework (OLF), Learning-Centered Leadership Framework (LCL), and the Essential Supports Framework. These were identified as influential frameworks that described effective leadership practices and had strong empirical support. The Essential supports framework is of relevant here because the CPS leadership studies reviewed in the present study used this specific framework to examine mediational leadership relationships.

The Essential Supports Framework proposes a multiple-mediational pathway that begins with school leadership and works through organizational processes of parent-community ties, staff professional capacity, student-centered learning environment, and ambitious instruction (Bryk et al., 2010; Sebring et al., 2006).

Based on longitudinal studies of elementary schools in Chicago, spanning from 1990 to 2005, Bryk et al. (2010) showed that schools that were strong on multiple dimensions of the five essentials framework were about 10 times more likely to improve during a period of education reform than schools that were weak on multiple dimensions. The focus of leadership in this framework is principal instructional leadership, but it also recognizes the importance of distributing leadership throughout the school, as evidenced by the involvement of teachers in school policy and work. As shown in Fig. 1, in this model, leadership works indirectly through multiple mediational processes to improve classroom instruction and student learning. The framework highlights how the main school organizational factors that influence student achievement are interconnected, and proposes that without a strong leader, school reform efforts are significantly less likely to be successful (Bryk et al., 2010).

School leadership studies in Chicago Public Schools (CPS)

In a series of studies funded through a research grant from the Institute of Education Sciences (IES) Department of Education, the University of Chicago Consortium on School Research utilized personnel survey data and student achievement data from CPS to study school organizational factors through which school leaders influence student learning. The results of these studies have been published (see Allensworth and Hart, 2018; Allensworth et al., 2020; Sebastian and Allensworth, 2012, 2013, 2019; Sebastian et al., 2016, 2014, 2019, 2017; Wiedermann and Sebastian, 2020a,b). The survey instruments for these studies were developed by the Chicago Consortium based on the Essential Supports framework. The Chicago Consortium collects longitudinal data on different aspects of the five essentials framework (including leadership, organizational factors, teaching, and learning), keeps rigorous documentation of instrument development and data collection, uses consistent survey items, and follows Rasch modeling for measure development. This allows for longitudinal modeling and analyses and valid comparisons across years (Bryk et al., 2010; Luppescu and Ehrlich, 2012; Sebring et al., 2006).

The Chicago Consortium leadership studies used Structural Equation Modeling (SEM) to link survey data on leadership and organizational processes with CPS administrative data on student test scores and grades. The longitudinal design of the studies allowed for studying the relationship of school leadership with improvement in student learning over time. Further, these studies controlled for student demographic and school contextual characteristics and examined elementary and secondary levels separately. A criticism of the SEM approach is that it evaluates pathways net of each other (Fiss, 2007, 2011; Ragin, 2008; Ragin and Fiss, 2008). For example, using SEM to study leadership allows us to evaluate the importance of professional development as a mediator, net of all other mediators. However, in actual practice, school leaders work on multiple organizational factors together and their

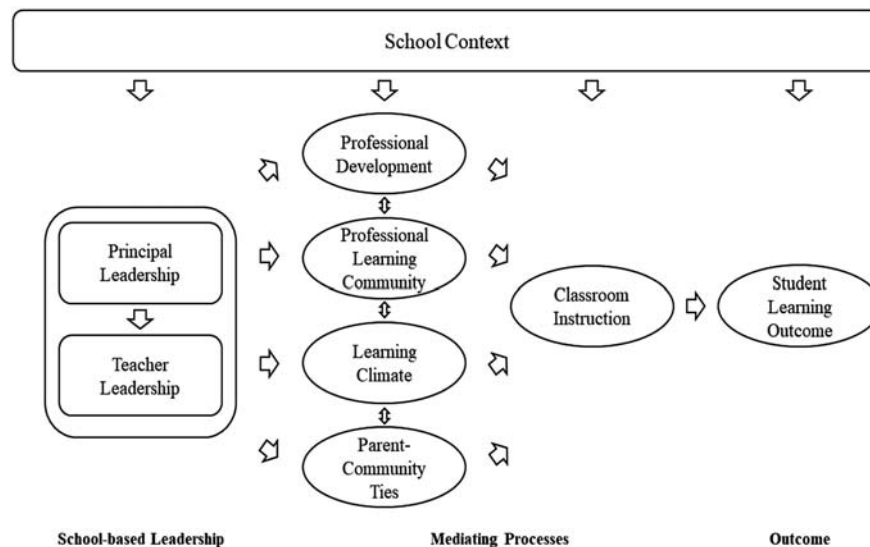


Fig. 1 Adaptation of the five essentials framework to examine mediational leadership pathways. Adapted from Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement Lessons From Chicago*. University of Chicago Press.

combined work on two or more mediators could result in improvements to student learning. In order to examine combinatory effects of multiple mediators acting together, the Chicago Consortium leadership studies also used a method called Qualitative Comparative Analysis (QCA) that helps to uncover combinatory patterns of the five essential supports as they relate to student outcomes (see Sebastian et al., 2014). As part of the IES grant, qualitative studies also sought to examine what differentiated the work of strong school leaders that ultimately translated into improvements in student learning. This was conducted using case studies of principals and teachers in schools which reported having a strong school principal but showed variation in their student achievement gains (see Allensworth et al., 2020). In the present manuscript, we will focus only on the quantitative studies to examine specific components of SEM models more closely.

Results of the Chicago Consortium leadership studies

The main results of the CPS leadership studies have been reported in refereed journal articles (see above). In brief, we found across multiple studies, examining elementary and secondary schools, and using cross-sectional and longitudinal data, that only the leadership pathway via school learning climate was statistically significant. We conceptualized school learning climate as an environment fostering high academic expectations and a safe and orderly climate. Our mediational studies found that none of the other pathways were statistically significant, net of learning climate. The remaining direct path between leadership and student learning was also not statistically significant. Further, the estimate of the only significant indirect pathway linking leadership to student learning (via learning climate) could be classified as small according to Cohen's criteria. Fig. 2 shows one example of the results and describes mediational pathways from high school data; these results were consistent across the other SEM studies using the Chicago Consortium data. Our results are also consistent with previous leadership mediational studies that found that the relationship of school leadership with student learning is largely indirect and overall effects tend to be small. However, prior studies typically had often included only one mediator in their analysis (see for example, Dumay et al., 2013; May and Supovitz, 2011). Other studies combined multiple school organizational factors into one overall global factor such as school capacity (Hallinger and Heck, 2010a,b; Heck and Hallinger, 2009). The CPS studies utilized the Essential Supports framework to examine multiple school mediators simultaneously. This approach helped in identifying which organizational processes school leaders could prioritize in their efforts to improve student learning.

The Consortium studies that examined the relationship of combinations of multiple organizational processes, including school leadership, also highlighted school learning climate as the most important component (Sebastian et al., 2014). In sum, across all our studies, we found that the strongest and only significant link between leadership and teaching/learning was via school climate, and that school climate mattered in all contexts. Even in schools that already had a strong school climate, improvements in school climate translated to further student achievement gains (Sebastian and Allensworth, 2019). We interpreted these findings in the following way:

Because all classrooms are affected by school climate, a principal's influence on climate may have the broadest reach across all the different classes in a school. As the school leader, the principal manages the work environment for everyone in the school. The more that students, teachers, and staff feel

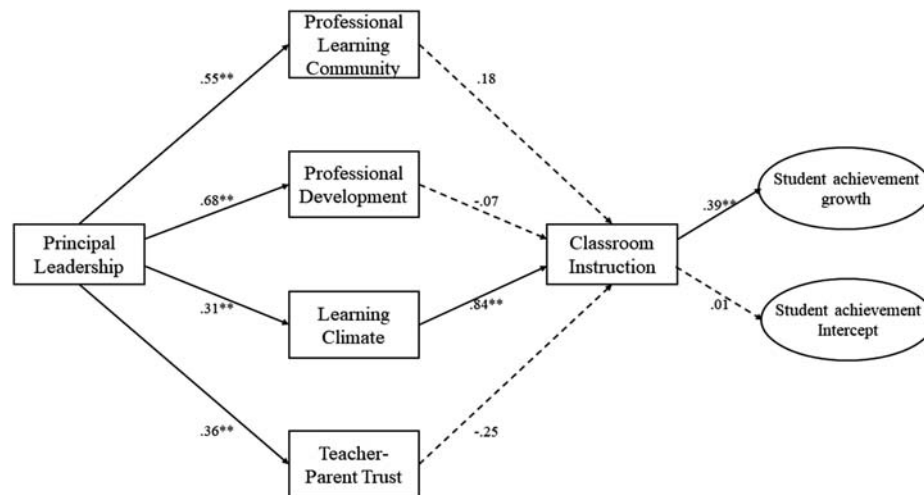


Fig. 2 Direct and indirect pathways from principal leadership to student achievement and student achievement gains. Adapted from Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28(3), 463–488.

safe and hold high expectations, the easier for everyone to be successful at teaching and learning. Efforts by principals to improve a school's climate do not require subject-specific knowledge or skills. While school leaders may not have content knowledge or expertise to be proficient in every subject, they can indirectly promote good instructional practices by creating a supportive school climate. This has some implications for how school leaders should organize their daily work. If school climate is the primary mechanism through which principals can influence classroom instruction and student learning, then principals need to prioritize climate in school improvement efforts (p. 9).

Disentangling pathways from leadership to student learning

In our previous CPS leadership studies, we reported the direct and indirect pathways from the SEM models, their coefficients, and standard errors. However, our discussion of results emphasized only the overall statistical significance of indirect pathways. In the present chapter, we consider the SEM results in closer light, to re-examine different sections of each leadership pathway.

Across all the Consortium leadership studies that used SEM models, the relationships between measures of leadership and all measures of mediating organizational processes were statistically significant. Fig. 2 shows that all paths from leadership to mediators (X-M)—parent/community ties, program quality, professional development, and learning climate were statistically significant. Using Cohen's criteria, the path estimates from leadership to professional learning community and professional development were moderate, whereas the path estimates for parent ties and learning climate could be classified as small. However, when examining the next stage of each mediational pathway, from the mediators to classroom instruction (M-Y), none of the relationships were statistically significant except for learning climate. Therefore, for most mediators, it is the second stage of leadership pathways that results in overall non-significant indirect effects. When considering the pathway via learning climate, net of the other, which is the only significant mediator, the path estimate for the first stage—leadership to climate (X-M), is weaker than the second stage (M-Y). For every other pathway, the correlation between leadership and the mediator (X-M) is stronger than the correlation between that mediator and the outcome (M-Y).

Taken together, these results show that for all school organizational processes that school leadership is theoretically expected to impact, empirical evidence suggests that leadership is indeed significantly related to those processes. When teachers in CPS reported higher levels of instructional leadership, they also reported higher levels of program quality, professional community, teacher professional development, parent/community involvement, higher academic expectations for students, and higher perceptions of safety. One limitation in using survey data is the potential for subjective bias that could arise from single-source survey reporting (Camburn et al., 2017). However, we utilized student, teacher, and principal survey reports for informing different measures used in the SEM models, and except for learning climate, the same patterns were consistently observed—strong correlations from leadership to mediating factors and non-significant relationship from the mediator to the outcome, except for learning climate.

Discussion and implications

The results of the present study show that when examining the second stage of leadership pathways to school outcomes (M-Y) using CPS data, we observe that in most cases relationships are not statistically significant. Although theory proposes that different school organizational processes should relate to the quality of classroom instruction and student learning, empirical evidence in the CPS

Table 1 Variables descriptive statistics.

<i>Name</i>	<i>Mean</i>	<i>SD</i>
Principal leadership	−0.02	0.60
Professional learning community	−0.01	0.68
Professional development	−0.02	0.47
Learning climate	0.03	1.49
Teacher parent trust	0.00	0.57

Adapted from Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28(3), 463–488.

context does not bear this out. This could be due to the unique context from where the data was collected, a large urban district where school learning climate is particularly important. Another reason could be that differences between schools on several mediators (organizational processes) are not large when compared to differences in school learning climate (see [Table 1](#)), or that those differences do not relate to school-level differences in classroom instruction and student learning, net of school climate.

Comparing our results to other mediational studies that have examined leadership pathways to student outcomes, many of them show results that are consistent with what we described here—stronger relationships from leadership to mediators and weaker (and often non-significant) relationships from the mediator to student learning. [Leithwood et al. \(2020\)](#) examined the links between leadership and student learning via four mediational paths—rational, emotional, organizational, and family; for all these paths, the path estimates from the leadership measure to each mediator was stronger than the corresponding path from that mediator to student learning. Similarly, other leadership studies (see for example, [Dumay et al., 2013](#); [Supovitz et al., 2010](#)) also show that links from leadership to various mediational factors are relatively stronger than the corresponding relationships between mediators and student outcomes.

[Wu \(2020\)](#) recently reviewed leadership studies that have utilized SEM models to examine direct and indirect pathways to student outcomes. [Wu \(2020\)](#) classified these studies based on [Leithwood et al. \(2010\)](#) and (2020)'s four-path model including rational, emotional, organizational, and family paths as mediators. We examined the mediational studies reviewed by [Wu \(2020\)](#) and examined specific components of mediational pathways from leadership (X) to student outcome (Y) via mediating processes (M). As shown in [Appendix A](#), we cataloged 95 mediational pathways from X to Y (via M) in 36 articles. Among 95 X-M-Y pathways, there were 62 cases in which X-M had a larger effect size than M-Y, 28 cases in which X-M had a smaller effect size than M-Y, and 5 cases for which estimates for specific pathway sections were not available. As with the Consortium IES leadership studies, for most pathways, the specific component of most leadership pathways where the overall indirect effect becomes non-significant or small (according to Cohen's criteria), is the section from the mediator to the outcome (M-Y) and not leadership to mediator (X-M). Taken together, these results can help the field clarify the reasons overall leadership effects are small. The weak indirect leadership effects are not due to weak links from leadership to organizational processes that leadership is expected to influence. Instead, the weak effects arise from weak correlations between school organizational processes and student outcomes.

The CPS studies show that leadership has an influence on all organizational processes that it is theoretically expected to influence. However, the links between these organizational processes and teaching/learning are not significant. As noted earlier, one of the reasons for this could be that school-level variation among organizational mediating factors is highest for learning climate (see [Table 1](#)). Therefore, schools vary widely in their learning climates and are relatively more similar on parent/community ties, professional development, and professional learning communities. This could be unique to urban school contexts such as CPS; however, the importance of learning climate has been emphasized in other studies using more generalizable data (see for example, [Leithwood and Sun, 2018](#); [Leithwood et al., 2020](#); [OECD, 2010, 2016](#)). [Leithwood and Sun \(2018\)](#) showed that academic culture—a latent variable comprised of academic press, disciplinary culture, and teachers' use of instructional time, significantly mediated the relationship between school leadership and student achievement. It is noteworthy that their academic culture construct is conceptually similar to the learning climate construct used in the CPS studies. [Leithwood et al. \(2020\)](#) compared multiple mediational pathways using the four-paths model and found that only the rational path mediated the relationship between leadership and achievement, again emphasizing the importance of academic culture.

Another reason could be the net-effects approach used in SEM studies, where each mediational pathway is evaluated net of other pathways ([Ragin and Fiss, 2008](#)). Our QCA study ([Sebastian et al., 2014](#)) sought to examine configurations of multiple organizational processes in relation to student outcomes. However, here again, the organizational process that was shown to be critical to effective school leadership was school learning climate alone. The use of QCA is only one method that has been proposed to examine the importance of multiple configurations and some critiques of the method have been noted (see for example, [Achen, 2005](#); [Seawright, 2005](#)). The research on examining combinations of multiple mediational pathways from leadership to student outcomes is still developing. Another possibility is a theoretical reconceptualization of school organization where school climate is seen as a primary mediator that in turn influences other organizational factors as secondary mediators. In the Consortium studies, if learning climate was not included in the SEM models, other mediators such as professional learning community were significant mediators of leadership and achievement ([Sebastian and Allensworth, 2012](#)). One possibility is that organizational factors could

matter to student outcomes to the extent that they influence school climate and instruction. [Leithwood et al. \(2020\)](#), whose work also emphasized the importance of “academic culture” (a construct similar to learning climate), similarly suggested the need for future research to examine the complex relationships among the mediators themselves, for e.g., other mediational pathways could influence achievement through the rational path.

Our results showed that learning climate has a strong correlation with classroom instruction and student learning, whereas the links between leadership and learning climate can be classified as small. These results could be influenced by how teachers rate principals. Specifically, the measure of principal leadership may have functioned as a proxy for principals’ strengths in other aspects of leadership including organizational management skills ([Sebastian et al., 2016](#)). Our results also imply that there are many factors that affect a school’s learning climate other than the school leadership such as a school’s contextual and demographic characteristics. Taken together, principals provide an organizing framework to improve the overall climate of the school by supporting and guiding the work of school personnel ([Sebastian et al., 2017](#)).

Overall, our results provide a useful empirical basis for understanding principal leadership effects on student outcomes and considering new directions for research and practice. Principals strongly influence student learning to the degree to which a school is set up to support student learning. However, instruction and learning are influenced by factors other than leadership and school organization. Of all the organizational conditions studied here, school climate has the largest relationship to instruction. Whether professional development or trusting relationships with families leads to stronger instruction depends on what professional development is about, and how relationships are structured to support student achievement in school. More empirical research is needed to verify these results and more theoretical work is needed to consider the implications for school organization theory and leadership preparation. Based on these results, it does not appear that further emphasizing leadership skills in parent-community ties, professional development, and professional learning community, through professional training or through leadership preparation programs can have direct benefits for learning outcomes. This is because there is little correlation between those organizational factors themselves and student learning when we also consider learning climate in analytical models. An important caveat to this claim is that our evidence is only based on large-scale quantitative data sets. Further, they may be relevant for other outcomes and for the effective functioning of schools. For example, [Leithwood \(2021\)](#) reviewed prior research on leadership practices that improve equitable school outcomes and pointed to three sets of practices (see page 33): (i) creating “authentic partnership between schools, families, and communities”, (ii) reflecting the “ethnic, cultural, and linguistic assets that diverse students and their families bring to the school” (p. 33) in the school curriculum, and (iii) assisting teachers “with the implementation of ambitious forms of instruction for traditionally underserved student” (p. 33).

Finally, the relatively weak correlation between leadership and school climate shows that there is room for strengthening leadership skills in improving climate and that other factors besides leadership are relevant for influencing climate. Identifying those factors through reviews of prior research and more empirical organizational research that prioritizes learning climate appears to be key for school improvement efforts.

Appendix A. Examination of mediational studies reviewed by Wu (2020)

Articles	X – M	M – Y	X – M > M – Y
Boberg and Bourgeois (2016)	X (Integrated transformational leadership) – M (Collective teacher efficacy)	M (Collective teacher efficacy) – Y (Reading achievement)	No
	X (Integrated transformational leadership) – M (Collective teacher efficacy)	M (Collective teacher efficacy)—via (Student emotional engagement) – Y (Mathematics achievement)	No
Bruggencate et al. (2012)	X (Rational goals) – M (Development orientation)	M (Development orientation)—via (Teacher's work)— via (Promotion rate) – Y (Average exam score)	Yes
	X (Internal process) – M (Development orientation)	M (Development orientation)—via (Teacher's work)— via (Promotion rate) – Y (Average exam score)	Yes
	X (Human relations) – M (Development orientation)	M (Development orientation)—via (Teacher's work)— via (Promotion rate) – Y (Average exam score)	Yes
	X (Open systems) – M (Development orientation)	M (Development orientation)—via (Teacher's work)— via (Promotion rate) – Y (Average exam score)	Yes
Chang (2011)	X (Distributed leadership) – M (Teacher academic optimism)	M (Teacher academic optimism) – Y (Student achievement)	No
Dumay et al. (2013)	X (Principal leadership) – M (Teacher collaboration)	M (Teacher collaboration)—via (Teachers' collective efficacy) – Y (Students' learning growth in mathematics)	Yes
	X (Principal leadership)—via (Teacher collaboration) – M (Teachers' collective efficacy)	M (Teachers' collective efficacy) – Y (Students' learning growth in mathematics)	Yes
Dutta and Sahney (2016)	X (Transformational leadership) – M (Social and affective environment)	M (Social and affective environment) – Y (Student achievement)	Yes
	X (Transformational leadership) – M (Job satisfaction)	M (Job satisfaction) – Y (Student achievement)	No
	X (Instructional leadership) – M (Social and affective environment)	M (Social and affective environment) – Y (Student achievement)	Yes
	X (Instructional leadership) – M (Physical environment)	M (Physical environment) – Y (Student achievement)	Yes
Fancera and Bliss (2011)	X (Instructional leadership) – M (Job satisfaction)	M (Job satisfaction) – Y (Student achievement)	No
	X (Protect instructional time) – M (Collective teacher efficacy)	M (Collective teacher efficacy) – Y (School achievement in SAT-CR)	No
Goddard et al. (2015)	X (Leadership) – M (Collaboration)	M (Collaboration)—via (Collective efficacy) – Y (2009 math)	Yes
	X (Leadership) – M (Collaboration)	M (Collaboration)—via (Collective efficacy) – Y (2009 reading)	Yes
	X (Leadership)—via (Collaboration) – M (Collective efficacy)	M (Collective efficacy) – Y (2009 math)	No
	X (Leadership)—via (Collaboration) – M (Collective efficacy)	M (Collective efficacy) – Y (2009 reading)	No
Goddard et al. (2010)	X (Shared instructional leadership) – M (Collaboration)	M (Collaboration) – Y (2008 mathematics achievement)	Yes
	X (Shared instructional leadership) – M (Collaboration)	M (Collaboration) – Y (2008 reading achievement)	Yes
Griffith (2004)	X (Principal transformational leadership) – M (Staff job satisfaction)	M (Staff job satisfaction) – Y (School achievement progress)	No
Hallinger and Heck (2010a)	X (Change in collaborative leadership) – M (Change in school capacity)	M (Change in school capacity) – Y (School growth)	Yes
Hallinger and Heck (2010b)	X (Change in leadership) – M (Change in capacity)	M (Change in capacity) – Y (Read growth rate)	Yes
Heck and Hallinger (2009)	X (Change in leadership) – M (Change in capacity)	M (Change in capacity) – Y (Math growth rate)	Yes
Heck and Hallinger (2010a)	X (Change in leadership) – M (Change in capacity)	M (Change in capacity) – Y (School growth)	Yes

(Continued)

Articles	$X - M$	$M - Y$	$X - M > M - Y$
Heck and Hallinger (2010b)	X (Change in leadership) – M (Change in capacity)	M (Change in capacity) – Y (Growth rate math)	Yes
	X (Change in leadership) – M (Change in capacity)	M (Change in capacity) – Y (Growth rate read)	Yes
Heck and Hallinger (2014)	X (Leadership) – M (Instructional environment)	M (Instructional environment) – Y (Growth)	Yes
Kwan (2016)	X (Transformational leadership) – M (Relational trust)	M (Relational trust) – Y (Student academic outcomes)	–
	X (Transformational leadership) – M (Relational trust)	M (Relational trust) – Y (Student non-academic outcomes)	–
	X (Instructional leadership) – M (Relational trust)	M (Relational trust) – Y (Student academic outcomes)	–
	X (Instructional leadership) – M (Relational trust)	M (Relational trust) – Y (Student non-academic outcomes)	–
Kythreotis et al. (2010)	X (Organizational commitment) – M (Academic emphasis)	M (Academic emphasis) – Y (Student achievement in Greek language)	Yes
Leithwood and Jantzi (2006)	X (Transformational leadership) – M (Motivation)	M (Motivation) – Y (Classroom practices/literacy)	Yes
	X (Transformational leadership) – M (Capacity)	M (Capacity) – Y (Classroom practices/literacy)	No
	X (Transformational leadership) – M (Situation)	M (Situation) – Y (Classroom practices/literacy)	Yes
	X (Transformational leadership) – M (Motivation)	M (Motivation) – Y (Classroom practices/numeracy)	Yes
	X (Transformational leadership) – M (Capacity)	M (Capacity) – Y (Classroom practices/numeracy)	No
	X (Transformational leadership) – M (Situation)	M (Situation) – Y (Classroom practices/numeracy)	Yes
Leithwood and Jantzi (2008)	X (School leader behavior) – M (School conditions)	M (School conditions) – Y (Student achievement)	Yes
Leithwood and Mascall (2008)	X (Collective leadership) – M (Capacity)	M (Capacity) – Y (Student achievement)	Yes
	X (Collective leadership) – M (Motivation)	M (Motivation) – Y (Student achievement)	No
	X (Collective leadership) – M (Setting)	M (Setting) – Y (Student achievement)	Yes
Leithwood et al. (2010)	X (Leadership) – M (Rational path)	M (Rational path) – Y (Combined achievement)	Yes
	X (Leadership) – M (Emotions path)	M (Emotions path) – Y (Combined achievement)	Yes
	X (Leadership) – M (Organizational path)	M (Organizational path) – Y (Combined achievement)	Yes
Louis et al. (2010)	X (Leadership) – M (Family path)	M (Family path) – Y (Combined achievement)	No
	X (Instructional leadership) – M (Professional community)	M (Professional community)—via (Focused instruction) – Y (2004–05 Math)	Yes
	X (Trust in principal) – M (Professional community)	M (Professional community)—via (Focused instruction) – Y (2004–05 Math)	Yes
	X (Shared leadership) – M (Professional community)	M (Professional community)—via (Focused instruction) – Y (2004–05 Math)	Yes
	X (Building level) – M (Professional community)	M (Professional community)—via (Focused instruction) – Y (2004–05 Math)	No
Marfan and Pascual (2018)	X (Principal's leadership practices) – M (Teachers' related aspects that affect school climate)	M (Teachers' related aspects that affect school climate) – Y (Plausible value in reading 1)	No
	X (Principal's leadership practices) – M (Teachers' stimulation of student engagement with reading)	M (Teachers' stimulation of student engagement with reading) – Y (Plausible value in reading 1)	Yes
McGuigan and Hoy (2006)	X (Enabling structure) – M (Academic optimism)	M (Academic optimism) – Y (Math school achievement)	No
	X (Enabling structure) – M (Academic optimism)	M (Academic optimism) – Y (Reading school achievement)	No
Miller et al. (2010)	X (Instructional leadership) – M (Collaboration)	M (Collaboration) – Y (Mathematics)	Yes
	X (Instructional leadership) – M (Collaboration)	M (Collaboration) – Y (Reading)	Yes

(Continued)

Articles	$X - M$	$M - Y$	$X - M > M - Y$
Mitchell et al. (2015)	X (Instructional leadership) – M (School academic press)	M (School academic press) – Y (Academic achievement)	No
Park et al. (2019)	X (Principal support) – M (Professional learning community)	M (Professional learning community)—via (Group-level teacher expectations) – Y (Student math achievement)	Yes
	X (Principal support) – M (Collective responsibility)	M (Collective responsibility)—via (Group-level teacher expectations) – Y (Student math achievement)	Yes
Ross and Gray (2006)	X (Transformational leadership) – M (Commitment to school mission)	M (Commitment to school mission) – Y (Grade 3 and 6 achievement)	Yes
	X (Transformational leadership) – M (Commitment to professional community)	M (Commitment to professional community) – Y (Grade 3 and 6 achievement)	Yes
	X (Transformational leadership) – M (Commitment to community partnerships)	M (Commitment to community partnerships) – Y (Grade 3 and 6 achievement)	No
Savvides and Pashiardis (2016)	X (School leadership) – M1 (School academic optimism)	M2 (Classroom as a learning environment—dealing with misbehaviour/positive aspects) – Y (Student cognitive outcomes in citizenship education)	–
Sebastian and Allensworth (2012)	X (Principal leadership) – M (Professional community)	M (Professional community) – Y (Classroom instruction)	Yes
	X (Principal leadership) – M (Program quality)	M (Program quality) – Y (Classroom instruction)	Yes
	X (Principal leadership) – M (Learning climate)	M (Learning climate) – Y (Classroom instruction)	No
	X (Principal leadership) – M (Parent-community ties)	M (Parent-community ties) – Y (Classroom instruction)	No
Sebastian and Allensworth (2019)	X (Change in principal leadership) – M (Change in program quality)	M (Change in program quality) – Y (Growth in ISAT)	Yes
	X (Change in principal leadership) – M (Change professional community)	M (Change professional community) – Y (Growth in ISAT)	Yes
	X (Change in principal leadership) – M (Change in teacher-parent trust)	M (Change in teacher-parent trust) – Y (Growth in ISAT)	Yes
	X (Change in principal leadership) – M (Change in school safety)	M (Change in school safety) – Y (Growth in ISAT)	Yes
Sebastian et al. (2016)	X (Principal leadership)—via (Teacher leadership) – M (Learning climate)	M (Learning climate) – Y (Classroom instruction)	No
	X (Principal leadership)—via (Teacher leadership) – M (Community & human resources)	M (Community & human resources) – Y (Classroom instruction)	Yes
	X (Principal leadership)—via (Teacher leadership) – M (Professional development program quality)	M (Professional development program quality) – Y (Classroom instruction)	Yes
	X (Principal leadership)—via (Teacher leadership) – M (Professional community)	M (Professional community) – Y (Classroom instruction)	Yes
Sebastian et al. (2017)	X (Principal leadership)—via (Teacher leadership) – M (Professional learning community)	M (Professional learning community) – Y (Classroom instruction)	Yes
	X (Principal leadership)—via (Teacher leadership) – M (Professional development)	M (Professional development) – Y (Classroom instruction)	Yes
	X (Principal leadership)—via (Teacher leadership) – M (Learning climate)	M (Learning climate) – Y (Classroom instruction)	No
	X (Principal leadership)—via (Teacher leadership) – M (Teacher-parent trust)	M (Teacher-parent trust) – Y (Classroom instruction)	Yes

(Continued)

Articles	$X - M$	$M - Y$	$X - M > M - Y$
Shen et al. (2016)	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (School culture)	M (School culture) – Y (Student performance)	Yes
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Academic press)	M (Academic press) – Y (Student performance)	Yes
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Instructional improvement)	M (Instructional improvement) – Y (Student performance)	Yes
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Professional development)	M (Professional development) – Y (Student performance)	Yes
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Disciplinary climate)	M (Disciplinary climate) – Y (Student performance)	No
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Teacher job satisfaction)	M (Teacher job satisfaction) – Y (Student performance)	No
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Parental involvement)	M (Parental involvement) – Y (Student performance)	No
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Principal leadership)	M (Principal leadership) – Y (Student performance)	Yes
	X (Principals' DDHS/data-informed decision-making on high-impact strategies) – M (Teacher influence)	M (Teacher influence) – Y (Student performance)	Yes
Supovitz et al. (2010)	X (Principal leadership) – M (Peer influence)	M (Peer influence) – Y (Change in instruction/English language arts)	Yes
	X (Principal leadership) – M (Peer influence)	M (Peer influence) – Y (Change in instruction/mathematics)	Yes
Wu et al. (2020)	X (Principal leadership) – M (Teacher self-efficacy)	M (Teacher self-efficacy) – Y (Student science achievement)	No
	X (Principal leadership) – M (Collaboration)	M (Collaboration) – Y (Student science achievement)	No
	X (Principal leadership) – M (Teacher job satisfaction)	M (Teacher job satisfaction) – Y (Student science achievement)	Yes

Acknowledgments

No potential conflict of interest was reported by the authors.

This work was supported by the Institute of Educational Sciences, U.S. Department of Education, under grant number R305A120706 to the University of Chicago.

References

- Achen, C.H., 2005. Two cheers for Charles Ragin. *Stud. Comp. Int. Dev.* 40 (1), 27–32.
- Allensworth, E., Hart, H., 2018. How Do Principals Influence Student Achievement? <https://consortium.uchicago.edu/publications/essential-supports-school-improvement>.
- Allensworth, E., Sebastian, J., Gordon, M.F., 2020. Principal leadership practices, organizational improvement, and student achievement. In: Peter, Y., Kim, J., Mavrogordato, M. (Eds.), *Exploring Principal Development and Teacher Outcomes: How School Leaders Can Strengthen Teacher Efficacy and Commitment*. Routledge, pp. 189–205.
- Boberg, J.E., Bourgeois, S.J., 2016. The effects of integrated transformational leadership on achievement. *J. Educ. Adm.* 54 (3), 357–374. <https://doi.org/10.1108/jea-07-2014-0086>.
- Bruggencate, G., Luyten, H., Scheerens, J., Slegers, P., 2012. Modeling the influence of school leaders on student achievement: how can school leaders make a difference? *Educ. Adm. Q.* 48 (4), 699–732. <https://doi.org/10.1177/0013161x11436272>.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement: Lessons From Chicago*. University of Chicago Press. <http://site.ebrary.com/lib/uchicago/Doc?id=10370350>.
- Camburn, E., Han, S.W., Sebastian, J., 2017. Assessing the validity of an annual survey for measuring the enacted literacy curriculum. *Educ. Pol.* 31 (1), 73–107. <https://doi.org/10.1177/0895904815586848>.

- Chang, I-Hua., 2011. A study of the relationships between distributed leadership, teacher academic optimism and student achievement in Taiwanese elementary schools. *Sch. Leadersh. Manag.* 31 (5), 491–515. <https://doi.org/10.1080/13632434.2011.614945>.
- Cohen, J., 1962. The statistical power of abnormal-social psychological research: a review. *J. Abnorm. Soc. Psychol.* 65 (3), 145–153.
- Dumay, X., Boonen, T., Van Damme, J., 2013. Principal leadership long-term indirect effects on learning growth in mathematics. *Elem. Sch. J.* 114 (2), 225–251.
- Dutta, V., Sahney, S., 2016. School leadership and its impact on student achievement. *Int. J. Educ. Manag.* 30 (6), 941–958. <https://doi.org/10.1108/ijem-12-2014-0170>.
- Fancera, S.F., Bliss, J.R., 2011. Instructional leadership influence on collective teacher efficacy to improve school achievement. *Leadersh. Policy Sch.* 10 (3), 349–370. <https://doi.org/10.1080/15700763.2011.585537>.
- Fiss, P.C., 2007. A set-theoretic approach to organizational configurations. *Acad. Manag. Rev.* 32 (4), 1180–1198.
- Fiss, P.C., 2011. Building better causal theories: a fuzzy set approach to typologies in organizational research. *Acad. Manag. J.* 54 (2), 393–420. <https://doi.org/10.5465/amj.2011.60263120>.
- Goddard, Y., Miller, R., Larson, R., Goddard, R., Madsen, J., Schroeder, P., 2010. Connecting Principal Leadership, Teacher Collaboration, and Student Achievement. Annual Meeting of the American Educational Research Association, Denver, CO, United States. Available at: <https://files.eric.ed.gov/fulltext/ED528704.pdf>.
- Goddard, R., Goddard, Y., Sook Kim, E., Miller, R., 2015. A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *Am. J. Educ.* 121 (4), 501–530. <https://doi.org/10.1086/681925>.
- Griffith, J., 2004. Relation of principal transformational leadership to school staff job satisfaction, staff turnover, and school performance. *J. Educ. Adm.* 42 (3), 333–356. <https://doi.org/10.1108/09578230410534667>.
- Grisson, J.A., Egalite, A.J., Lindsay, C.A., 2021. How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research. The Wallace Foundation, New York. <https://www.wallacefoundation.org/knowledge-center/Documents/How-Principals-Affect-Students-and-Schools.pdf>.
- Hallinger, P., Heck, R.H., 1996a. The principal's role in school effectiveness: an assessment of methodological progress, 1980–95. In: Leithwood, K., Chapman, J., Corson, D., Hallinger, P., Hart, A. (Eds.), *The International Handbook of Educational Leadership and Administration*. Kluwer, pp. 723–784.
- Hallinger, P., Heck, R.H., 1996b. Reassessing the principal's role in school effectiveness: a review of empirical research, 1980–1995. *Educ. Adm. Q.* 32 (1), 5–44.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9 (2), 157–191.
- Hallinger, P., Heck, R.H., 2010a. Collaborative leadership and school improvement: understanding the impact on school capacity and student learning. *Sch. Leader. Manag.* 30 (2), 95–110.
- Hallinger, P., Heck, R.H., 2010b. Leadership for learning: does collaborative leadership make a difference in school improvement? *Educ. Manag. Adm. Leader.* 38 (6), 654–678.
- Hallinger, P., Bickman, L., Davis, K., 1996. School context, principal leadership, and student reading achievement. *Elem. Sch. J.* 96 (5), 527–549.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leader. Pol. Sch.* 4 (3), 221–239.
- Heck, R.H., Hallinger, P., 2009. Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *Am. Educ. Res. J.* 46 (3), 659–689. <https://doi.org/10.3102/0002831209340042>.
- Heck, R.H., Hallinger, P., 2010a. Collaborative leadership effects on school improvement: Integrating unidirectional- and reciprocal-effects models. *Elem. Sch. J.* 111 (2), 226–252. <https://doi.org/10.1086/656299>.
- Heck, R.H., Hallinger, P., 2010b. Testing a longitudinal model of distributed leadership effects on school improvement. *Leadersh. Q.* 21 (5), 867–885. <https://doi.org/10.1016/j.leaqua.2010.07.013>.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681. <https://doi.org/10.1108/jea-08-2013-0097>.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement. *Rev. Educ. Res.* 86 (2), 531–569. <https://doi.org/10.3102/0034654315614911>.
- Kwan, P., 2016. The effect of trust on the relationship between instructional leadership and student outcomes in Hong Kong secondary schools. *Asia-Pac. Educ. Res.* 25 (1), 111–121. <https://doi.org/10.1007/s40299-015-0242-5>.
- Kythreotis, A., Pashiardis, P., Kyriakides, L., 2010. The influence of school leadership styles and culture on students' achievement in Cyprus primary schools. *J. Educ. Adm.* 48 (2), 218–240. <https://doi.org/10.1108/09578231011027860>.
- Leithwood, K., Jantzi, D., 2006. Transformational school leadership for large-scale reform: effects on students, teachers, and their classroom practices. *Sch. Effect. Sch. Improv.* 17 (2), 201–227.
- Leithwood, K., Jantzi, D., 2008. Linking leadership to student learning: the contributions of leader efficacy. *Educ. Adm. Q.* 44 (4), 496–528. <https://doi.org/10.1177/0013161x08321501>.
- Leithwood, K., Mascall, B., 2008. Collective leadership effects on student achievement. *Educ. Adm. Q.* 44 (4), 529–561. <https://doi.org/10.1177/0013161x08321221>.
- Leithwood, K., Sun, J., 2018. Academic culture: a promising mediator of school leaders' influence on student learning. *J. Educ. Adm.* 56 (3), 350–363. <https://doi.org/10.1108/JEA-01-2017-0009>.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K., 2004. Review of Research: How Leadership Influences Student Learning. The Wallace Foundation, New York. <https://www.wallacefoundation.org/knowledge-center/pages/how-leadership-influences-student-learning.aspx>.
- Leithwood, K., Patten, S., Jantzi, D., 2010. Testing a conception of how school leadership influences student learning. *Educ. Adm. Q.* 46 (5), 671–706. <https://doi.org/10.1177/0013161x10377347>.
- Leithwood, K., Sun, J., Schumacker, R., 2020. How school leadership influences student learning: a test of “The Four Paths Model”. *Educ. Adm. Q.* 56 (4), 570–599. <https://doi.org/10.1177/0013161x19878772>.
- Leithwood, K., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30 (4), 498–518. <https://doi.org/10.1177/0013161x94030004006>.
- Leithwood, K., 2021. A review of evidence about equitable school leadership. *Educ. Sci.* 11 (8), 377. <https://doi.org/10.3390/educsci11080377>.
- Liebowitz, D.D., Porter, L., 2019. The effect of principal behaviors on student, teacher, and school outcomes: a systematic review and meta-analysis of the empirical literature. *Rev. Educ. Res.* 89 (5), 785–827.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., Anderson, S.E., Michlin, M., Mascall, B., 2010. Learning From Leadership: Investigating the Links to Improved Student Learning. Wallace Foundation, New York. <https://www.wallacefoundation.org/knowledge-center/pages/investigating-the-links-to-improved-student-learning.aspx>.
- Lupescu, S., Ehrlich, S.B., 2012. UChicago Consortium Rasch Measurement Model Primer, 2012 UChicago Consortium Rasch Measurement Model Primer.
- Marfan, J., Pascual, J., 2018. Comparative study of school principals' leadership practices: lessons for Chile from a cross-country analysis. *Educ. Manag. Adm. Leader.* 46 (2), 279–300.
- May, H., Supovitz, J.A., 2011. The scope of principal efforts to improve instruction. *Educ. Adm. Q.* 47 (2), 332–352.
- McGuigan, L., Hoy, W.K., 2006. Principal leadership: creating a culture of academic optimism to improve achievement for all students. *Leadersh. Policy Sch.* 5 (3), 203–229. <https://doi.org/10.1080/15700760600805816>.
- Miller, R., Goddard, Y., Goddard, R., Larsen, R., Jacob, R., 2010. Instructional Leadership: A Pathway to Teacher Collaboration and Student Achievement. University Council for Educational Administration Convention, New Orleans, LA, United States. <https://files.eric.ed.gov/fulltext/ED528591.pdf>. (Accessed 10 July 2022).
- Mitchell, R.M., Kensler, L.A.W., Tschannen-Moran, M., 2015. Examining the effects of instructional leadership on school academic press and student achievement. *J. School Leadersh.* 25 (2), 223–251. <https://doi.org/10.1177/105268461502500202>.
- Neumerski, C.M., 2012. Rethinking instructional leadership, a review. *Educ. Adm. Q.* 49 (2), 310–347. <https://doi.org/10.1177/0013161x12456700>.
- OECD, 2010. PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices. Organization for Economic Co-operation and Development.
- OECD, 2016. PISA 2015 Results (Volume II): Policies and Practices for Successful Schools. OECD Publishing.

- Park, J.-H., Lee, I.H., Cooc, N., 2018. The role of school-level mechanisms: how principal support, professional learning communities, collective responsibility, and group-level teacher expectations affect student achievement. *Educ. Adm. Q.* 55 (5), 742–780. <https://doi.org/10.1177/0013161x18821355>.
- Ragin, C.C., Fiss, P.C., 2008. Net effects analysis versus configurational analysis. In: Ragin, C.C. (Ed.), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press, pp. 190–212.
- Ragin, C.C., 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press.
- Ross, J.A., Gray, P., 2006. School leadership and student achievement: the mediating effects of teacher beliefs. *Can. J. Edu. / Revue canadienne de l'éducation* 29 (3), 798. <https://doi.org/10.2307/20054196>.
- Savvides, V., Pashiardis, P., 2016. An exploration of relationships between leadership and student citizenship outcomes in Cyprus middle schools. *Educ. Adm. Q.* 52 (3), 497–526. <https://doi.org/10.1177/0013161x16638415>.
- Seawright, J., 2005. Qualitative comparative analysis vis-à-vis regression. *Stud. Comp. Int. Dev.* 40 (1), 3–26.
- Sebastian, J., Allensworth, E., 2012. The influence of principal leadership on classroom instruction and student learning: a study of mediated pathways to learning. *Educ. Adm. Q.* 48 (4), 626–663.
- Sebastian, J., Allensworth, E., 2013. How do secondary principals influence teaching and learning? *Princ. Res. Rev.* 8 (4), 1–5. <https://consortium.uchicago.edu/publications/essential-supports-school-improvement>.
- Sebastian, J., Allensworth, E., 2019. Linking principal leadership to organizational growth and student achievement: a moderation mediation analysis. *Teach. Coll. Rec.* 121 (9), 1–32.
- Sebastian, J., Allensworth, E., Stevens, D., 2014. The influence of school leadership on classroom participation: examining configurations of organizational supports. *Teach. Coll. Rec.* 116 (8), 1–36.
- Sebastian, J., Allensworth, E., Huang, H., 2016. The role of teacher leadership in how principals influence classroom instruction and student learning. *Am. J. Educ.* 123 (1), 69–108.
- Sebastian, J., Huang, H., Allensworth, E., 2017. Examining integrated leadership systems in high schools: connecting principal and teacher leadership to organizational processes and student outcomes. *Sch. Effect. Sch. Improv.* 28 (3), 463–488. <https://doi.org/10.1080/09243453.2017.1319392>.
- Sebastian, J., Allensworth, E., Wiedermann, W., Hochbein, C., Cunningham, M., 2019. Principal leadership and school performance: an examination of instructional leadership and organizational management. *Leader. Pol. Sch.* 18 (4), 591–613. <https://doi.org/10.1080/15700763.2018.1513151>.
- Sebring, P.B., Allensworth, E., Bryk, A.S., Easton, J.Q., Luppescu, S., 2006. *The Essential Supports for School Improvement*. Research report. Consortium on Chicago School Research. <https://consortium.uchicago.edu/publications/essential-supports-school-improvement>.
- Shen, J., Ma, X., Cooley, V.E., Burt, W.L., 2016. Mediating effects of school process on the relationship between principals' data-informed decision-making and student achievement. *Int. J. Leader. Educ.* 19 (4), 373–401. <https://doi.org/10.1080/13603124.2014.986208>.
- Silva, J.P., White, G.P., Yoshida, R.K., 2011. The direct effects of principal–student discussions on eighth grade students' gains in reading achievement. *Educ. Adm. Q.* 47 (5), 772–793. <https://doi.org/10.1177/0013161x11404219>.
- Supovitz, J., Sirinides, P., May, H., 2010. How principals and peers influence teaching and learning. *Educ. Adm. Q.* 46 (1), 31–56. <https://doi.org/10.1177/1094670509353043>.
- Tan, C.Y., Gao, L., Shi, M., 2020. Second-order meta-analysis synthesizing the evidence on associations between school leadership and different school outcomes. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/1741143220935456>.
- Wiedermann, W., Sebastian, J., 2020a. Direction dependence analysis in the presence of confounders: applications to linear mediation models using observational data. *Multivar. Behav. Res.* 55 (4), 495–515. <https://doi.org/10.1080/00273171.2018.1528542>.
- Wiedermann, W., Sebastian, J., 2020b. Sensitivity analysis and extensions of testing the causal direction of dependence: a rejoinder to Thoemmes (2019). *Multivar. Behav. Res.* 55 (4), 523–530. <https://doi.org/10.1080/00273171.2019.1659127>.
- Witziers, B., Bosker, R.J., Kruger, M.L., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 39 (3), 398–425.
- Wu, H., Shen, J., Zhang, Y., Zheng, Y., 2020. Examining the effect of principal leadership on student science achievement. *Int. J. Sci. Educ.* 42 (6), 1017–1039. <https://doi.org/10.1080/09500693.2020.1747664>.
- Wu, H., 2020. *The Effect of Principal Leadership on Student Achievement: A Multivariate Meta-analysis*. Doctoral dissertation. Western Michigan University.
- Zhu, H., Li, L., Li, H., 2020. How school leadership influences Chinese students' reading literacy: a test of the rational, emotions, and organizational paths in rural schools. *Child. Youth Serv. Rev.* 119. <https://doi.org/10.1016/j.childyouth.2020.105534>.

Leadership practices and why they matter: trust in leaders

Megan Tschannen-Moran, William and Mary School of Education, Williamsburg, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

What do school leaders need to know about trust?	255
Benevolence	256
Conveying caring, empathy, and respect	257
Cultivating a culture of care	257
Honesty	257
Be honest and authentic	258
Foster a culture of integrity	258
Competence	258
Mediate conflict	258
Nurture the competence of those you lead	259
Reliability	259
Foster a culture of professionalism	259
Openness	260
Be open	260
Extend discretion and shared decision making	260
Trusted and trusting	261
References	261

Schools are challenged to adapt to the needs of a rapidly changing world. Trust is a vital resource in times of crisis and change (Misra, 1996), thus school leaders would do well to invest in the cultivation of trust in their schools. Schools that foster high-trust environments are better positioned to accomplish the challenging task of educating a diverse group of students in a changing world.

What do school leaders need to know about trust?

It is imperative that school leaders are knowledgeable about and attend to issues of trust because trust is strongly associated with student outcomes in schools. A growing body of evidence attests to the powerful role that trust plays in student achievement (Adams, 2008; Forsyth et al., 2006; Goddard et al., 2001, 2009; Hoy, 2002; Hoy et al., 2006; Hoy and Tschannen-Moran, 1999; Tschannen-Moran, 2004). In a study of two districts, one urban and the other suburban, I found that 75% of the variance in student achievement was explained by a set of trust variables that included faculty trust in students, parents, colleagues, and principals, as well as student trust in teachers and parent trust in schools (Tschannen-Moran, 2014b). Moreover, in a decade-long study of Chicago public schools engaged in reform initiatives, Bryk and Schneider (2002) concluded that trust was a critical factor in predicting which schools would make the greatest gains in student achievement and which would sustain those gains over time. Further, comparative international research across seven countries demonstrated the key role that trust plays in student outcomes within these national systems of schooling (Fink, 2016). In the past five decades, educational researchers have searched diligently and largely in vain to identify variables that explain student achievement above and beyond the effects of SES. Trust, and specifically the reciprocal trust between teachers and students, has proven to be just such a powerful variable (Goddard et al., 2001, 2009; Hoy, 2002; Hoy et al., 2006).

Trust is essential for leaders in binding the hearts and minds of followers to a shared mission or goal. Faculty trust in the principal has been linked to student academic achievement in both direct and indirect ways (Chughtai and Buckley, 2009; Forsyth and Adams, 2014; Handford and Leithwood, 2013; Notman and Henry, 2011; Salfi, 2011; Tschannen-Moran, 2014b; Zeinabadi, 2014). Faculty trust in the principal has been related to both the collegial orientation of a principal's leadership as well as to instructional leadership (Tschannen-Moran and Gareis, 2015). To be productive and accomplish organizational goals, schools need cohesive and cooperative relationships, and trust is essential to fostering these relationships. Trust supports schools' effectiveness, improvement initiatives, and persistence in reform efforts (Bryk and Schneider, 2002; Cerna, 2014; Forsyth et al., 2011, 2006; Goddard et al., 2001; Louis, 2007; Tschannen-Moran, 2014a; Uline et al., 1998; Van Maele and Van Houtte, 2012). A culture of innovation and continuous learning, as well as the success of improvement initiatives, have all been tied to strong bonds of trust (Bryk and Schneider, 2002; Cerna, 2014; Cosner, 2009; Daly et al., 2014; Forsyth et al., 2011, 2006; Goddard et al., 2001; Hoy et al., 1996; Louis, 2007; Moolenaar et al., 2010).

Faculty trust has been related to the level and quality of collaboration in a school (Tschannen-Moran, 2001), teachers' perceptions of the professionalism of their colleagues (Tschannen-Moran, 2009), and the quality of professional learning communities (Hallam et al., 2014). When trust has been damaged or has failed to develop in the first place, teachers and staff members spend

their energy in hypervigilance, perennially on guard to protect themselves and their interests, which inevitably comes at the expense of collective goals. Students who do not trust their teachers or each other divert energy in self-protection and away from engagement with the learning task. Furthermore, students who do not feel trusted by their teachers and administrators may psychologically distance themselves from schools and build an alienated, rebellious youth culture that creates barriers to participation in the learning task. Students may, in fact, live down to the low expectations of a distrustful school environment. Moreover, the degree to which students trust their teachers has also been found to be strongly related to student achievement across a variety of contexts (Adams, 2010; Mitchell et al., 2008, 2018; Tschannen-Moran et al., 2013).

Because trust is so important to the productive functioning of schools, defining just what trust is and how it works in a school is important. The need for trust emerges in a situation of interdependence, in which valued outcomes cannot be achieved without the involvement and contribution of others (Tschannen-Moran, 2014a). A core aspect of trust is the deliberate acceptance of vulnerability at the hand of a trustee, despite the absence of guarantees that the other party will not violate that trust by acting opportunistically (Heyns and Rothmann, 2018). In schools, principals, teachers, staff, students, and parents are interdependent because none can accomplish the complex task of educating young people alone, but the shape and meaning of trust may differ based on the nature of that interdependence. Managing a diversity of trust relationships makes the work of educators particularly complex. Adding to the challenge of defining and studying trust is that fact that it is a complex and dynamic construct that can change over the course of a relationship, as the nature of the interdependence between parties changes and as expectations are either fulfilled or disappointed.

People make judgments as to whom to trust and to what extent based on several criteria or facets of trust. Much of the research on trust in schools has relied upon the following definition.

Trust is a feeling at ease in a situation of interdependence based on the confidence that the other person or group is benevolent, honest, open, reliable, and competent

Tschannen-Moran (2014a).

In the relationships between teachers and principals, this model in which the five facets form a unitary and coherent concept of trust has consistently held up when examined through exploratory and confirmatory factor analyses as well as qualitative investigations (Hoy, 2002; Hoy and Tschannen-Moran, 1999; Tschannen-Moran, 2014b). However, qualitative studies have also uncovered the potential for unevenness in trust relationships, based on differing expectations. For example, a teacher in an underperforming urban school reported that while he trusted his colleagues as friends, he did not trust them as professionals (Tschannen-Moran, 2014a). Thus, trust and distrust can both be present in a single relationship. Although various facets of trust may be salient, their relative weight will depend on the referent of trust (who is being trusted) and the nature of the interdependence between the parties. In a study of 205 schools (including survey responses from 205 principals and 1150 teachers) in Chile, Weinstein et al. (2020) found that in the relationships between teachers and their principal, benevolence was the facet of trust that carried the most weight in teachers' trust assessments. Furthermore, the facets most tied to character, that is benevolence, openness and honesty, as a group, carried more weight than those related to competence (i.e., competence and reliability). When principals were assessing their trust in teachers, however, the weightings were reversed. Reliability and competence played a stronger role, whereas those facets related to character were somewhat less potent.

To nurture high trust environments and to reap the benefits of trust, school leaders must earn the trust of their followers as well as cultivate a culture of trust throughout the organization. Although the constructs of trustworthiness and trust are related, Heyns and Rothmann (2018) highlight the distinction. While trustworthiness is a quality possessed by the person being trusted, trust itself is something that the trusting party does. A culture of trust is a set of tacit beliefs and assumptions that a group of people in an organization hold in common. Weinstein et al. (2020) observed that principals are the main drivers of a culture of trust in schools, setting the tone, as well as influencing the intensity and scope of the trust relationship in their school. In the sections that follow, we will discuss leadership behaviors associated with being trustworthy, as well as actions they can take to cultivate a high trust culture in their schools, exploring each of the five facets of trust in turn.

Benevolence

One of the most universally acknowledged foundations of trust is a sense of benevolence or positive regard expressed as a mutual spirit of good will and a willingness to extend oneself in support of the well-being of the other (Bryk and Schneider, 2002; Tschannen-Moran, 2014a). It is grounded in the confidence that one's well-being or something one cares about will be protected and not harmed by the person in whom they have placed their trust (Zand, 1997). Benevolence also entails a willingness to forgo personal gain if that gain would bring potential harm to the trusting party. The nature of benevolence may vary with the nature of the relationship. When there is a power differential in the relationship, as in the relationship between school leaders and teachers or between teachers and students, the person with greater power may characterize the sense of good will they are looking for in a subordinate as respect, whereas the person with less power in a relationship may characterize mutual good will as a sense of caring about one's wellbeing (Weinstein et al., 2020).

Conveying caring, empathy, and respect

For school leaders to be regarded as trustworthy, they must evidence an abiding sense of care for their constituents as persons, care for the relationship, and care for the shared project of schooling. Teachers expect principals to be aware of their particular situation, and to be concerned for their physical and psychological well-being. Teachers expect principals to play a role as protectors of their rights and to act as mediators between them and upper-level administrators to prevent teachers from being adversely affected by decisions made. Teachers trust in their principals grows when they note ways their principal takes initiative to improve the working conditions of teachers. Teachers expect the principal to view the staff not as mere workers but as real individuals with wants and needs. Benevolence involves having empathy for and being responsive to the needs of the other (Tschannen-Moran, 2014a). Weinstein et al. (2020) found that teachers were especially sensitive to the principals' knowledge about their specific health problems or major personal difficulties (for example, a relative passing away), and their willingness to be flexible regarding their duties while they were working through their personal difficulty. In a study of 116 schools in the United States, principal trust was directly related to teachers' perceptions of principal caring (Louis and Murphy, 2017). Supportive principal behaviors have been found to strongly predict teachers' trust in the principal; they also influence teachers' trust in colleagues and in students and parents (Demirtaş et al., 2017; Yilmaz, 2004; Yilmaz and Altinkurt, 2012). Various forms of support, such as emotional support, instrumental support, professional support, and appraisal support have all been found to be significant predictors of trust in the principal as well as to organizational commitment (Demirtaş et al., 2017).

In their role as change agents, a sense of benevolence, empathy, and caring is especially important for principals. If school leaders want their staff to engage in making changes in their practice, they must demonstrate that care in ways that are evident to their followers and be understanding of the challenges brought about by the change. Organizational change can be disruptive to a person's sense of wellbeing, as well as their sense of meaning and purpose at work (Bolman and Deal, 2017). It can create feelings of uncertainty as they encounter new processes and procedures for accomplishing their objectives. Communicating clearly about expectations of teachers and what they can expect through each stage of the change process helps to reduce the level of uncertainty that teachers have to cope with. Change also inevitably involves loss (Bolman and Deal, 2017). As they cope with the losses inherent in organizational change, teachers may exhibit any of the five stages of grief (denial, anger, bargaining, depression, and finally acceptance) in response to a change initiatives. It is important that school leaders recognize these reactions as normal responses to grief, and not evidence that those resisting are malevolent or incompetent if they are initially less than enthusiastic at the prospect of doing things differently. Inviting conversations about how teachers are feeling about changes in their school, listening skillfully to their concerns, and expressing empathy are potent means for assisting teachers and staff members to move through the grief process.

Cultivating a culture of care

To cultivate a school-wide culture of care, principals need to attend to the elements of culture, such as norms, stories, ceremonies, and the use of humor (Bolman and Deal, 2017). Establishing shared norms is an important part of developing a shared vision for the school. Changing norms regarding acceptable behavior and speech will likely require some explicit norm-setting conversations. Once these norms are agreed upon, it is a school leader's responsibility to respond to behaviors that fall outside these norms and to teach followers to intervene when these norms are violated. In calling people on their behavior, a leader has to ensure that they abide by the norms themselves, addressing the issues in private and avoiding embarrassing a teacher in front of others. While a school leader may communicate their desire to have a safe and risk-free environment for teaching and learning, when they fail to confront teachers and students on issues of bad behavior in the classroom in a helpful and supportive way, they appear to be giving tacit approval to these behaviors. The stories that are told about people whose actions are admired as well as those told as cautionary tales are another important element of culture that either foster a climate of care or undermine it. One prime area that often needs to be addressed is the tone of how teachers speak about students whose behavior they disapprove of. The ceremonies and rituals of the school also convey the collective values in the school, and warrant attention as to the messages that are being sent. What counts as acceptable forms of humor and play are also areas ripe for exploration in terms of how they can support a culture of care. Humor that involves sarcasm or humiliation is detrimental to a caring culture. Students are attentive to signs that their teachers care about them as individuals and not just about keeping order or moving efficiently through the curriculum (Cooper and Minness, 2014). When teachers speak about this aspect of trust in relation to students, they more often frame it as respect—respect for the shared project of schooling, for the rules, processes, and procedures of school, as well as respect for themselves, for fellow students, for school property, and for educators themselves (Tschannen-Moran, 2014a).

Honesty

Honesty is a fundamental facet of trust. Honesty is evidenced through behaviors that demonstrate integrity, authenticity, and accountability for one's actions (Hoy and Tschannen-Moran, 1999; Tschannen-Moran and Hoy, 1998). Trust requires having confidence that a person's word can be relied upon and can accurately predict future actions. Trust facilitates communication and contributes to greater efficiency when people have confidence in the integrity of other people's words and deeds.

Be honest and authentic

To earn the trust of their followers, school leaders need to earn a reputation for honesty. Integrity speaks to living according to a set of core values or principles, being sincere, and telling the truth. Integrity entails an alignment between a person's words and their deeds. When teachers begin to perceive a discrepancy between their leader's words and actions, trust is undermined. Once a leader's verbal statements are regarded with suspicion, it will be hard for them to regain trust because the spoken word is an essential tool for fostering trust. Having integrity also involves refraining from using one's power to play favorites or to improve one's personal outcomes. A sense of fairness and fair play or what is sometimes referred to as organizational justice is also an essential element of integrity (Yilmaz and Altinkurt, 2012).

Authenticity is expressed when a school leader reveals to followers their true personality and values, as opposed to the projection of an image that masks their nature and intentions. Authenticity involves a willingness to share one's heart, humanity, and foibles with others. School leaders who come across as too guarded in what they are willing to reveal about themselves can be perceived as simply playing a theatrical role, and thus their motivations may be regarded with suspicion. Accountability involves a willingness to take responsibility for one's mistakes and refraining from blaming others for personal failings. School leaders demonstrate vulnerability by acknowledging their mistakes publicly and sharing a plan for growth.

Foster a culture of integrity

If we hope to instill a sense of integrity in our students, then the adults must consistently model these behaviors. Even very young children are observant of instances in which teachers' actions do not align with their admonitions regarding their expectations of student behavior. Just as leaders are expected to tell the truth and keep their promises, teachers must be held to these same standards. One area of promise-keeping that may seem counterintuitive is in the area of threats. Failure to follow through on a threat can be as damaging to trust as failure to keep a promise, even if the misbehaving student is relieved to escape the promised punishment (Tschannen-Moran, 2014a). Thus, teachers need to be cautioned not to make empty threats, and supported to learn more effective ways of winning the cooperation and compliance of students. Cultivating a culture of integrity also means moving away from conveying a tacit image of teachers as all-knowing and infallible. Teachers need to model for students that it is acceptable to say "I don't know (let's see if we can figure out how to find out)" and "I made a mistake (here's what I learned and what I will do next time)."

Competence

Competence is the ability to perform a task as expected, according to appropriate standards. In work that is as strongly interpersonal in nature as education, there are two dimensions to the competence of school leaders: competence with the task and competence with relationships. Both are necessary for trust to develop and be sustained. Competence as an educational leader involves inspiring action, setting and maintaining high standards and pressing for results. School leaders inspire teachers in their commitment to students, but also challenge and support teachers who fall short in their duty to improve their instructional practice, holding teachers accountable in ways that seem fair and reasonable to their staff (Bryk and Schneider, 2002; Chugtai and Buckley, 2009; Forsyth and Adams, 2014; Handford and Leithwood, 2013; Weinstein et al., 2020; Zeinabadi, 2014). Teachers expect principals to efficiently manage their school community and create collaborative spaces, fostering positive relationships that promote cooperation, a spirit of teamwork, and mutual support among colleagues. In particular, principals are expected to play a key role in establishing harmonious working relationships between diverse groups of teachers, such as novice versus senior teachers or regular versus special education teachers (Weinstein et al., 2020).

Mediate conflict

Of the multitude of skills that effective principals evidence, one of the most potent when it comes to building and maintaining trust is to constructively manage conflict. The health of a school community is dependent upon a leader who can mediate the inevitable conflicts inherent in the vital and complex work of schools. Stakeholders rely on educators to be skillful problem-solvers and to demonstrate the requisite conflict-resolution skills necessary to navigate the competing needs and interests of a diverse educational community (Tschannen-Moran, 2014a). Weinstein et al. (2020) found that the principals in their study who were regarded as competent were those who made thoughtful decisions, taking the time to ponder all sides of the issue and the possible consequences of various courses of action, who consulted various stakeholders, and who acted with equanimity, as opposed to acting impulsively. These principals were seen as acting in ways that were fair and just, prioritizing the interests of the school over the interests of any particular group. The teachers interviewed especially emphasized the importance of the way principals made decisions about sensitive matters regarding teachers, such as the assignment of responsibilities or sanctions to address problems such as chronic absenteeism or tardiness.

Principals need to not only have the skills to manage conflict themselves, but to create the structures and provide the training necessary for teachers and students to learn constructive conflict-management strategies as well (Cosner, 2009). When teachers are supported by norms and processes that help them negotiate solutions that meet the needs of all parties, it can help restore trust that

has been broken and also to prevent further breakdowns of trust. Both administrators and teachers need to be aware of the dramatic costs of broken trust. Conflict, resistance, fear, and anxiety can erode trust when people perceive that the actions of others fall outside of established norms and expectations (Tschannen-Moran, 2014a). When trust is broken between administrators and teachers, suspicion and psychological withdrawal are likely to result (Tschannen-Moran, 2014a). When trust is broken between teachers and students, a cycle of punishment and withdrawal or rebellion may result, setting up a dynamic that is deleterious to cognitive and social-emotional development. This is important because the process of repairing broken trust is difficult and costly. Effective educators are on guard to prevent the abuse of power and encourage the cultivation of high-trust relationships.

Organizational change initiatives create conditions that are ripe for conflict. Change processes inevitably disrupt the power dynamics within a school, advantaging some and disadvantaging others (Bolman and Deal, 2017). Even where there is agreement on the goals of the change, there are often differences of opinion as to how to achieve those goals. Breakdowns in trust may emerge around issues of misunderstanding the purpose of the change process, what new behaviors and outcomes will be expected, and how and when the success of the change will be assessed. Honing the competence to engage rather than avoid controversy by skillfully debating ideas is a valuable asset during times of change. Strong norms of civil discourse that guide participants to challenge the strategies, ideas, or thinking of those with whom they disagree, while steadfastly avoiding personal attacks, are important to keeping the dialog constructive (Uline et al., 2003).

Nurture the competence of those you lead

A school leader cannot adequately fulfill their role unless they can lead their teachers and staff to high levels of performance (Tschannen-Moran, 2003, 2009, 2014a). For teachers to be considered competent, they are expected to provide engaging, accurate, and informative instruction, as well as act as diagnosticians who can assess the learning difficulties of particular students and construct effective interventions for struggling students. To support these competencies, a school leader must be consistent in providing clear, specific feedback to followers and in delivering the tools and support necessary for their steady improvement. One of the payoffs of high trust is the development of an organizational culture that supports innovation and professional learning, and a willingness to take risks (Bryk and Schneider, 2002; Cosner, 2009; Daly et al., 2014; Forsyth et al., 2011; Goddard et al., 2001; Louis, 2007). Engaging teachers in instructional rounds encourages a collaborative approach to instructional improvement and when managed carefully has the potential to elevate trust among the faculty.

School leaders must be insistent on and encouraging of improvement, serving as a support system for teachers who are lacking in the adaptive skills necessary to provide high quality instruction to all students. School leaders have a moral and ethical responsibility to safeguard the wellbeing of the students in their school. Therefore, they must respond swiftly and strategically to teachers who engage in teacher aggression that evidences a lack of classroom management skills that undermine trust, such as yelling, using sarcasm and humiliation, as well as group punishments (Riley, 2010; Tschannen-Moran, 2014a). At the same time, they must give people a fair chance to improve before moving toward dismissal. The evocative coaching model suggests building on teacher strengths and then inviting them to design experiments around ways they want to or need to improve their instruction (Tschannen-Moran and Tschannen-Moran, 2020).

Reliability

In a situation of interdependence, when something of importance is required from another person or group that impacts joint outcomes, trust is evidenced by a sense of confidence that the trusted person will come through with what is needed on a consistent basis, in a timely fashion, and that the work will be of sufficient quality to meet joint standards (Mishra, 1996). If the reliability of the trusted person or group is in question, the trustor must invest energy in making mental provisions for how they will manage in case the other person should fail to follow through on agreements. Reliability involves keeping commitments, keeping promises, keeping appointments, and keeping confidences. Trust may survive a broken promise if a plausible explanation is given along with an apology; however, a pattern of broken agreements will pose a threat to trust.

Foster a culture of professionalism

One of the reasons that trust is so powerfully related to school outcomes is because of the role trust plays in fostering a culture of professionalism (Tschannen-Moran, 2003, 2009, 2014a). The vision for a school as a vibrant professional learning community is characterized by a deep commitment to diagnosing and addressing the needs of individual students through collaboration and joint deliberation. This involves ongoing, rigorous professional inquiry, a collective focus on student learning, data to bolster decision making, de-privatized teaching practice, and reflective dialog (Fullan, 2003; Louis et al., 1996), all of which would require trust. Principals' work to support and enhance collaboration among teachers has been identified as important with respect to the cultivation of collegial trust (Cosner, 2009; Tschannen-Moran, 2001). Greater professionalism has been linked to stronger motivation, greater commitment to shared goals, and a willingness to voluntarily go beyond minimum expectations (Tschannen-Moran, 2003), as well as improved student achievement (Tschannen-Moran et al., 2006; Tschannen-Moran, 2009). If school leaders hope to inspire greater levels of professionalism among their faculty, they need to adopt a professional orientation in their exercise of authority because principal professional orientation has been found to be strongly related to the level of teacher professionalism

(Mitchell and Tarter, 2016; Tschannen-Moran, 2009). School leaders with a professional orientation hold in high regard the professional expertise of teachers and evidence faith in their capability; and because of this, they extend greater trust and give teachers greater autonomy and discretion to exercise their professional judgment in response to the diverse needs of clients (Hoy and Sweetland, 2001; Tschannen-Moran and Gareis, 2015).

In a sense of urgency to bring about improvement in their schools, school leaders may unwittingly fall into a trap that is not uncommon in organizations facing a competitive and changing external environment. Under these challenging circumstances, schools often develop a “rigid response,” becoming more centralized, doubling down on control over teachers, and enforcing rigid adherence to standardized procedures such as curricular scope and sequence plans or scripted curriculums (Daly, 2009). Trust plays a central role in a school’s ability to avoid these counterproductive tactics. High trust school leaders eschew rigid rules and micro-management as means to control teachers, and instead focus on strengthening joint norms of professionalism and a shared vision of a school in which all children thrive (Cosner, 2009). However, these leaders do not abdicate their responsibility for leadership. They engage in coaching and collaboration to bring underperforming teachers into alignment with professional standards, as well as to provide resources to continually extend the professional knowledge of all the teachers in their building (Tschannen-Moran, 2014a).

Openness

Openness as a facet of trust evidences the reciprocal nature of trust, as people garner trust by extending trust. In being open, people make themselves vulnerable to others by sharing information, influence, and control (Zand, 1997). Leaders who are guarded with the information they share provoke suspicion as people wonder what is being hidden and why. Withholding important information is a strategy that leaders may use to maintain power or manipulate employees, but it may cost them the trust of their followers (Mishra, 1996).

Be open

Strong principals inspire a shared vision by listening carefully to the hopes and needs of the people they lead. This process is a key opportunity for school leaders to ensure that underrepresented voices in the community are heard and included as part of the vision. A school leader may arrive with a powerful and positive vision, however fail to draw others in. Adopting a strengths-based process, such as appreciative inquiry, can help guide a principal and the school community to craft a vision that is truly shared.

Trust contributes to organizational effectiveness, in part, by creating the conditions for improved communication. In schools where trust is high, communication tends to flow more easily; and, the candor and open exchange of information allows problems to be disclosed, diagnosed, and corrected before they are compounded. In these high-trust environments, mistakes are viewed as opportunities for learning and refinement rather than as occasions for criticism and blame; this makes for greater openness and honesty in the face of disappointing results (Hoy and Sweetland, 2001; Tschannen-Moran, 2014a). Leaders help foster the flow of information to them by being open with communication that flows from them (Bryk and Schneider, 2002; Heyns and Rothmann, 2018). In doing so, school leaders make themselves vulnerable as an act of trust, modeling the courage it takes to confront and discuss difficult issues, especially those likely to evoke an emotional reaction (Kutsyruba et al., 2016; Meyer et al., 2017). When leaders reveal their ideas, thoughts and feelings freely, sharing facts, alternatives, intentions, and judgments, it not only enhances perceptions of trustworthiness but leads to greater openness on the part of subordinates as well.

Trust that involves disclosure of critical and sensitive information of either work-related or personal nature, such as communicating personal insights, task-related knowledge, thoughts and beliefs, problems, concerns, feelings, hunches, and emotional responses to work-related events openly and honestly requires a deeper level of trust than relying only on another’s competence and direction because it is a more active form of engagement. In these situations, trust is more complicated when there is a relationship of unequal power as, for example, between school leaders and teachers (Bryk and Schneider, 2002; Meyer et al., 2017). Those in power can diminish the tension experienced by those with less power through behaviors that signal their intention to foster trust (Bryk and Schneider, 2002). Teachers and staff can be invited to think about what types of obstacles they encounter that prevent them from sharing their views and how they could start to remove these barriers (Cosner, 2009).

Extend discretion and shared decision making

School leaders foster trust through extending trust, with openness in granting discretion and autonomy to teachers, in delegating tasks, and by engaging stakeholders in distributed leadership and democratic decision making. For teachers to work as professionals, they need to be granted discretion to make decisions in their classrooms, even as they are held accountable for the outcomes in student learning and wellbeing (Tschannen-Moran and Gareis, 2015). As professionals, teachers need to learn, grow, and express themselves authentically in their work roles. Autonomy can be promoted by giving teachers more freedom of choice in the way in which instructional strategies are executed, encouraging initiative, and re-designing jobs in ways that enhance meaningfulness and so that work becomes a form of internally motivated self-expression. At the same time, attention should be given to clarifying the expectations and responsibilities that go hand in hand with more freedom, such as the discussion of obstacles to the pursuit of excellence in a frank and fearless manner (Moye and Henkins, 2005).

Delegating important tasks without micromanaging is another way that leaders extend trust in ways that garner greater trust (Heyns and Rothmann, 2018; Tschannen-Moran, 2014a). Teachers who perceived that they were empowered in their work environments and who reported substantial influence in their work environments had higher levels of trust in their principals (Moye and Henkins, 2005). Distributed leadership has been found to be positively related to trust among colleagues and trust in the principals (Adiguzelli, 2016; Beycioglu et al., 2012; Cansoy and Parlar, 2018).

Leaders who engage in authentic forms of shared decision-making in which teachers, students, and parents have genuine influence over decisions help foster trust (Tschannen-Moran, 2001). This is in contrast to leaders who orchestrate contrived processes in which subordinates are asked to give their input to a decision that the subordinates are suspicious has already been made. Democratic principal behaviors in which teacher have the opportunity to express their opinions on issues that concern them and ensuring teacher participation in decisions over which they have expertise and a personal stake have been found to be significantly related to trust in the principal (Heyns and Rothmann, 2018; Kars and Inandi, 2018). One of the ways that trust serves to facilitate school because it expands the zone of acceptance in decisions (Bryk and Schneider, 2002; Hoy and Tarter, 2008). Where trust is low, teachers want a place at the decision-making table in order to guard their interests. When they trust that their interests will be well looked after, they may feel freed to invest their energies elsewhere.

Trusted and trusting

For school leaders to be successful, they must both earn the trust of their followers and foster high trust relationships between the members of their staff, as well as trusting relationships between teachers, students, and their families. Developing strategies for fostering deeper trust are crucial skills for those who would lead schools in our increasingly diverse society. Attention to the five facets of trust, including benevolence, honesty, competence, reliability, and openness provides potent guidance to school leaders for tackling these two challenging and critical objectives.

References

- Adams, C.M., 2008. Building trust in schools: a review of the empirical evidence. In: Hoy, W.K., DiPaola, M. (Eds.), *Improving Schools: Studies in Leadership and Culture*. Information Age Publishing, Charlotte, NC, pp. 29–54.
- Adams, C.M., 2010. Social determinants of student trust in high poverty elementary schools. In: Hoy, W.K., DiPaola, M. (Eds.), *Analyzing School Contexts: Influences of Principals and Teachers in the Service of Students*. Information Age, Charlotte, NC, pp. 255–280.
- Adiguzelli, Y., 2016. Examining the relationship between distributed leadership and organizational trust according to opinions of teachers. *Educ. Sci.* 41, 185.
- Beycioglu, K., Ozer, N., Ugurlu, C.T., 2012. Distributed leadership and organizational trust: the case of elementary schools. *Proc. Soc. Behav. Sci.* 46, 3316–3319.
- Bolman, L.G., Deal, T.E., 2017. *Reframing Organizations: Artistry, Choice, and Leadership*. John Wiley and Sons, New York.
- Bryk, A.S., Schneider, B., 2002. *Trust in Schools: A Core Resource for School Improvement*. Russell Sage Foundation, New York.
- Cansoy, R., Parlar, H., 2018. Examining the relationships among trust in administrator, distributed leadership and school academic optimism. *Educ. Manag. Theor. Pract.* 24, 1–44.
- Cerna, L., 2014. Trust: What it is and Why it Matters for Governance and Education, OECD Education Working Papers, No. 108. OECD Publishing, Paris. <https://doi.org/10.1787/5jxswcg0t6wl-en>.
- Chughtai, A.A., Buckley, F., 2009. Linking trust in the principal to school outcomes: the mediating role of organizational identification and work engagement. *Int. J. Educ. Manag.* 23, 574–589.
- Cooper, K.S., Miness, A., 2014. The co-creation of caring student/teacher relationships: does teacher understanding matter? *High Sch. J.* 97, 264–290.
- Cosner, S., 2009. Building organizational capacity through trust. *Educ. Adm. Q.* 45, 248–291.
- Daly, A.J., Liou, Y., Moolenaar, N., 2014. The principal connection: trust and innovative climate in a network of reform. In: Van Maele, D., Forsyth, P.B., Van Houtte, M. (Eds.), *Trust and School Life: The Role of Trust for Learning, Teaching, Leading, and Bridging*. Springer Publisher, Dordrecht, pp. 285–312.
- Daly, A., 2009. Rigid response in an age of accountability: the potential of leadership and trust. *Educ. Adm. Q.* 45, 168–216.
- Demirtaş, H., Özer, N., Demirebilek, N., Onur, B., 2017. Relationship between the perceived principal support, trust in principal and organizational commitment. *Int. Online J. Educ. Sci.* 9, 1075–1092.
- Fink, D., 2016. *Trust and Verify: The Real Keys to School Improvement*. Institute of Education Press, London, England.
- Forsyth, P.B., Adams, C.M., 2014. Organizational predictability, the school principal, and achievement. In: *Trust and School Life*. Springer, Dordrecht, pp. 83–98.
- Forsyth, P.B., Barnes, L.L., Adams, C.M., 2006. Trust-effectiveness patterns in schools. *J. Educ. Adm.* 44, 121–141.
- Forsyth, P.B., Adams, C.M., Hoy, W.K., 2011. *Collective Trust: Why Schools Can't Improve Without It*. Teachers College Press, New York.
- Fullan, M. (Ed.), 2003. *The Moral Imperative of School Leadership*. Corwin Press, Thousand Oaks, CA.
- Goddard, R.D., Tschannen-Moran, M., Hoy, W.K., 2001. A multilevel examination of the distribution and effects of teacher trust in students and parents in urban elementary schools. *Elem. Sch. J.* 102, 3–17.
- Goddard, R.D., Salloum, S.J., Berebitsky, D., 2009. Trust as a mediator of the relationship between poverty, racial composition, and academic achievement. *Educ. Adm. Q.* 45, 292–311.
- Hallam, P.R., Dulany, S.K., Hite, J.M., Smith, H.R., 2014. Trust at ground zero: trust and collaboration within professional learning community. In: Van Maele, D., Forsyth, P.B., Van Houtte, M. (Eds.), *Trust and School Life: The Role of Trust for Learning, Teaching, Leading, and Bridging*. Springer Publisher, Dordrecht, pp. 145–170.
- Handford, V., Leithwood, K., 2013. Why teachers trust school leaders. *J. Educ. Adm.* 51, 194–212.
- Heyns, M., Rothmann, S., 2018. Volitional trust, autonomy satisfaction, and engagement at work. *Psychol. Rep.* 121, 112–134.
- Hoy, W.K., Sweetland, S.R., 2001. Designing better schools: the meaning and measure of enabling school structures. *Educ. Adm. Q.* 37, 296–321.
- Hoy, W.K., Tarter, C.J., 2008. *Administrators Solving the Problems of Practice: Decision-Making, Concepts, Cases, and Consequences*, third ed. Allyn and Bacon, Boston.
- Hoy, W.K., Tschannen-Moran, M., 1999. Five faces of trust: an empirical confirmation in urban elementary schools. *J. Sch. Leader.* 9, 184–208.
- Hoy, W.K., Sabo, D., Barnes, K., 1996. Organizational health and faculty trust: a view from the middle level. *Res. Middle Level Educ. Q.* 19 (3), 21–39.
- Hoy, W.K., Tarter, C.J., Hoy, A.W., 2006. Academic optimism of schools: a force for student achievement. *Am. Educ. Res. J.* 43, 425–446.
- Hoy, W.K., 2002. Faculty trust: a key to student achievement. *J. Sch. Publ. Relat.* 23, 88–103.
- Kars, M., Inandi, Y., 2018. Relationship between school principals' leadership behaviors and teachers' organizational trust. *Eur. J. Educ. Res.* 74, 145–164.

- Kutsyuruba, B., Walker, K., Noonan, B., 2016. The trust imperative in the school principalship: the Canadian perspective. *Leader. Pol. Sch.* 15, 343–372.
- Louis, K.S., Murphy, J., 2017. Trust, caring and organizational learning: the leader's role. *J. Educ. Adm.* 55, 103–126.
- Louis, K.S., Marks, H.M., Kruse, S., 1996. Teachers' professional community in restructuring schools. *Am. Educ. Res. J.* 33, 757–798.
- Louis, K.S., 2007. Trust and improvement in schools. *J. Educ. Change* 8, 1–24.
- Meyer, F., Le Fevre, D.M., Robinson, V.M., 2017. How leaders communicate their vulnerability: implications for trust building. *Int. J. Educ. Manag.* 31, 221–235.
- Mishra, A.K., 1996. Organizational responses to crisis: the centrality of trust. In: Kramer, R., Tyler, T. (Eds.), *Trust in Organizations*. Sage, Thousand Oaks, CA, pp. 261–287.
- Mitchell, R.M., Tarter, C.J., 2016. A path analysis of the effects of principal professional orientation towards leadership, professional teacher behavior, and school academic optimism on school reading achievement. *Societies* 6 (5), 1–11.
- Mitchell, R.M., Forsyth, P.B., Robinson, U., 2008. Parent trust, student trust and identification with school. *J. Res. Educ.* 18, 116–124.
- Mitchell, R.M., Kensler, L.W., Tschannen-Moran, M., 2018. Student trust in teachers and student perceptions of safety: positive predictors of student identification with school. *Int. J. Leader. Educ.* 21 (2), 135–154.
- Moolenaar, N.M., Karsten, S., Sleegers, P.J.C., Zijlstra, B.J.H., 2010. Linking social networks and trust: a social capital perspective on professional learning communities. In: Moolenaar, N.M. (Ed.), *Ties with Potential: Nature, Antecedents, and Consequences of Social Networks in School Teams*. Organisation for Scientific Research, Netherlands.
- Moye, M.J., Henkin, A.B., 2005. Teacher-principal relationships: exploring linkages between empowerment and interpersonal trust. *J. Educ. Adm.* 43, 260–277.
- Notman, R., Henry, D.A., 2011. Building and sustaining successful school leadership in New Zealand. *Leader. Pol. Sch.* 10, 375–394. <https://doi.org/10.1080/15700763.2011.610555>.
- Riley, P., 2010. Attachment Theory and the Teacher-Student Relationship: A Practical Guide for Teachers, Teacher Educators and School Leaders. Routledge, London.
- Salfi, N.A., 2011. Successful leadership practices of head teachers for school improvement: some evidence from Pakistan. *J. Educ. Adm.* 49, 414–432. <https://doi.org/10.1108/09578231111146489>.
- Tschannen-Moran, M., Gareis, C.R., 2015. Faculty trust in the principal: an essential ingredient in high-performing schools. *J. Educ. Adm.* 53, 66–92.
- Tschannen-Moran, M., Hoy, W.K., 1998. Trust in schools: a conceptual and empirical analysis. *J. Educ. Adm.* 36, 334–352.
- Tschannen-Moran, M., Tschannen-Moran, B., 2020. *Evocative Coaching: Transforming Schools One Conversation at a Time*, second ed. Corwin, Thousand Oaks, CA.
- Tschannen-Moran, M., Parish, J., DiPaola, M.F., 2006. School climate and state standards: how interpersonal relationships influence student achievement. *J. Sch. Leader.* 16, 386–415.
- Tschannen-Moran, M., Bankole, R., Mitchell, R., Moore, D., 2013. Student academic optimism: a confirmatory factor analysis. *J. Educ. Adm.* 51, 150–175.
- Tschannen-Moran, M., 2001. Collaboration and the need for trust. *J. Educ. Adm.* 39, 308–331.
- Tschannen-Moran, M., 2003. Fostering organizational citizenship: transformational leadership and trust. In: Hoy, W.K., Miskel, C.G. (Eds.), *Studies in Leading and Organizing Schools*. Information Age Publishing, Charlotte, NC, pp. 157–179.
- Tschannen-Moran, M., 2004. What's Trust Got to Do With It? The Role of Faculty and Principal Trust in Fostering Student Achievement. Paper Presented at the Annual Meeting of the University Council for Educational Administration, Kansas City, MO.
- Tschannen-Moran, M., 2009. Fostering teacher professionalism: the role of professional orientation and trust. *Educ. Adm. Q.* 45, 217–247.
- Tschannen-Moran, M., 2014a. *Trust Matters: Leadership for Successful Schools*, second ed. Jossey-Bass.
- Tschannen-Moran, M., 2014b. The interconnectivity of trust in schools. In: Van Maele, D., Forsyth, P.B., Van Houtte, M. (Eds.), *Trust and School Life: The Role of Trust for Learning, Teaching, Leading, and Bridging*. Springer Publisher, Dordrecht, pp. 57–81. https://doi.org/10.1007/978-94-017-8014-8_3.
- Uline, C., Miller, D., Tschannen-Moran, M., 1998. School effectiveness: the underlying dimensions. *Educ. Adm. Q.* 34, 462–483.
- Uline, C., Tschannen-Moran, M., Perez, L., 2003. Constructive conflict: how controversy can contribute to school improvement. *Teach. Coll. Rec.* 105, 782–815.
- Van Maele, D., Van Houtte, M., 2012. The role of teacher and faculty trust in forming teachers' job satisfaction: do years of experience make a difference? *Teach. Teach. Educ.* 28, 879–889.
- Weinstein, J., Raczynski, D., Pena, J., 2020. Relational trust and positional power between school principals and teachers in Chile: a study of primary schools. *Educ. Manag. Adm. Leader* 48, 64–81.
- Yilmaz, K., Altinkurt, Y., 2012. Relationship between the leadership behaviors, organizational justice, and organizational trust. *Çukurova Univ. Fac. Educ. J.* 41 (1).
- Yilmaz, K., 2004. Primary school teachers' perception about the supportive leadership behavior of school administrators and trust in school. *İnönü Univ. J. Fac. Educ.* 5 (8), 117–131.
- Zand, D.E., 1997. *The Leadership Triad: Knowledge, Trust, and Power*. Oxford University Press, New York.
- Zeinabadi, H.R., 2014. Principal-teacher high-quality exchange indicators and student achievement: testing a model. *J. Educ. Adm.* 52, 404–420.

Effects of family educational cultures on student success at school: directions for leadership

William H. Jeynes^{a,b}, ^aWitherspoon Institute, Princeton, NJ, United States; and ^b California State University, Long Beach, Long Beach, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Alterable family variables	263
Alterable family variable #1: parental involvement including parental expectations	264
Elements of parental involvement	265
Having high, but reasonable parental expectations	265
Alterable family variable #2: parental style and the importance of love and structure in the home	265
Alterable family variable #3: social capital	265
How can schools and their leaders help parents further develop these features of family educational cultures?	266
Educators need to become more aware of how great the impacts are of parental family structure, fatherlessness, and parental involvement on students	266
School leaders need to help parents, students, and teachers rediscover the salience of social capital	267
Principals and other school leaders should train school personnel about the most beneficial kinds of parental -expectations, -style, and -involvement	267
Information and direction that school leaders can provide for parents	267
School leaders should make sure parents and students who have added hurdles are sufficiently informed about the advantages of parental involvement	268
Principals and other leaders should train school personnel about the most beneficial kinds of parental involvement for children of different ages	268
School leaders need to do more to involve fathers	269
Concluding thoughts	269
References	269

Increasingly over the past several decades, there has been a growing appreciation by educators, scholars, and those in a variety of leadership positions that family factors are even more important than school variables in terms of influencing academic outcomes of youth (Coleman et al., 1982; Kyriakides and Creemers, 2008; Leithwood et al., 2010). Leithwood and his colleagues, as a result of examining a good amount of data on the topic, have concluded that the family path is perhaps the most salient and explains a sizable portion of the variance in scholastic outcomes. Other prominent studies suggest that family factors may explain about 50% of the variance in pupil achievement (Coleman et al., 1982; Kyriakides and Creemers, 2008). In contrast, research indicates that school factors contribute about 20% of the variance (Coleman et al., 1982; Creemers and Reezigt, 1996). Consequently, an educational leader who recognizes the potential of the family path could emerge as a powerful human catalyst, inspiring students to reach high levels in achievement (Jeynes, 2006b).

Admittedly, in recent decades many principals and teachers have understated the role of the family in school outcomes because they opine that educators can have more of an impact on school variables rather than on home-based ones (Katz et al., 2015; Robinson, 2014). In addition, there is a natural proclivity for educators to study largely education journals and similar publications, sociologists to study sociological ones, psychologists to read psychological publications, economists examine economic works, and so forth (Jeynes, 2012b). If superintendents and principals are to be more efficacious, they need a more interdisciplinary educational purview (Flecha and Soler, 2013; Leithwood et al., 2010; National Academy of Sciences, 2005).

Alterable family variables

With all this in mind, what is key for school leaders to understand is that is that there is a difference between what are *alterable* family variables and those that are unalterable. Unalterable family-related variables are those over which the school has no influence (e.g., parental education and parental income). *Alterable* family variables, sometimes referred to as family educational culture, are potentially open to influence from the school and its leadership. The encouraging trend is that copious researchers believe that there are increasingly more family factors that are alterable than what was thought three decades ago (Hattie, 2009; Jeynes, 2015b).

Five of the family influences that have the most puissant effects on children are: (1) parental involvement in their children’s lives, (2) the degree to which children are loved (Levine, 2012), (3) parental family structure (Jeynes, 2011b; Shirakawa, 2010), (4) social capital, and (5) parental style. Of the five primary ways that family factors can influence pupil academic results, parental involvement, social capital, and parental style are patently the ones on which school leaders can have the greatest impact. It is these three

with that will be addressed in this article, which are the most *alterable*. It should be noted, however, that school leaders can also make a difference in instructing parents on how to lovingly support their children and having their knowledge about the effects of family structure help students and parents (McLanahan and Sandefur, 1994; Wallerstein et al., 2002). Given that it is these three aspects of caregiving that are the most alterable, in this article, the most emphasis will be placed on parental involvement. The reader should keep in mind that when the term “parental involvement” and “parental style” are used this does not refer to any particular family structure(s), but is based on the reality that all children have parents in some sense of the term.

Alterable family variable #1: parental involvement including parental expectations

Parental involvement has become one of the primary areas of research and practice in education today (Hornby, 2011). Parental engagement has been around for many centuries (Harris et al., 2009; Wartman and Savage, 2008). It was present in the teachings of Old Testament and the New Testament, as well as in the daily practices of the Pilgrims and Puritans (Jeynes, 2011b; Wartman and Savage, 2008).

Nevertheless, a series of societal changes in various countries around the world have caused scholars to rediscover parental involvement and this has helped lift it to a central place of scholarship. Beginning around 1963 the divorce rate surged and out of wedlock births also rose precipitously (Phillips, 1988; Weaver and Schofield, 2015). Over the 1963–1980 period, the divorce rate rose consistently for seventeen sequential years (Phillips, 1988; Weaver and Schofield, 2015). Concurrently, by the early 1990s out of wedlock births in the United States had inflated by seven times (U.S. Department of Health and Human Services, 2012; U.S. Department of Justice, 1999). The expressions “latch key children” and “deadbeat dads,” which had been almost nonexistent for the bulk of world history, were by the 1960s and 1970s, at the foreground of interactions among family scholars in many nations (Harris et al., 2009; Jeynes, 2019; Weaver and Schofield, 2015).

Moreover, innumerable children born out of wedlock were oblivious of who their fathers even were (Baskerville, 2007; Mincy et al., 2015; Turkat, 1995). Globally, mothers also entered the workforce in greater numbers and higher percentages than ever before, which translated into parents spending less time with their children, on average, than had been the case in past decades (Shirakawa, 2010; Weaver and Schofield, 2015).

Out of these concerns about the breakdown of the traditional family beginning in about 1963, arose a plethora of concerns about the stability of family. This spawned two branches of research, although one emerged before the other. The first branch focused on the effects of being in a non-traditional family, especially in single-parent families and blended ones (Hampden-Thompson and Galindo, 2015; Moor and Komter, 2012). Much of the purpose of these studies was to educate mothers and fathers about the typical consequences of their children potentially living in a single-parent or blended family. It was hoped, therefore, that if present- and future-parents are cognizant of these statistics, they will hesitate to engage in behavior that increases the likelihood that children will be raised in these family structures. (Moor and Komter, 2012).

The diligence with which family structure was studied was strongly related to the decline in achievement test scores that was transpiring concurrent to the surge in divorce rates (Carter and Glick, 2013; Raymo et al., 2013). Academics concluded that it was highly unusual that the exact period of time in which SAT scores in the US plummeted for seventeen consecutive years, 1963–1980, was identical to the timeframe during which the American divorce rate surged for 17 consecutive years (Carter and Glick, 2013; Stevenson and Wolfers, 2007). Moreover, also in 1963 juvenile -crime rates and -measures of anti-social behavior began to surge (Carter and Glick, 2013; Jeynes, 2007). Studies have shown that when there is an absent parent, children have a higher proclivity to become depressed, suicidal, take illegal drugs, consume large amounts of alcohol, and become sexual promiscuous (Jeynes, 2015a).

The second branch was less focused on results that might prevent children unnecessarily being born into a difficult family situation. Rather, this branch acknowledged that children were being born into single-parent families and wanted to see if high levels of involvement by the custodial parent could compensate for coming from a non-traditional family structure (Jeynes, 2005, 2011a; Tan, 2017). Researchers were also naturally convinced that parental involvement could benefit youth from all family structures, but they were especially concerned about children at risk, particularly as these risks emanated from family structure (Vaughans and Spielberg, 2014).

One of the academic dynamics that caused this second branch to develop is that a number of the most prominent social scientists in the field of parental involvement have also studied family structure, e.g., Joyce Epstein (Epstein, 1989). The developing body of literature on parental family structure’s influence on children therefore eventually produced a new discipline within family studies, i.e., the study of parental involvement (Jeynes, 2011b). This was also a salient advancement, because educators viewed parental involvement as more *alterable* than is parental family structure.

Meta-analytic research demonstrates that parental involvement tends to have a positive impact of about 0.60–0.75 of a grade point (Jeynes, 2003). Parental family structure tends to have an impact of a similar size, i.e., the scholastic difference between youth in intact families in which both the mother and father are present and children from the never married single parent family structure is usually about 0.60–0.75 of a grade point (Jeynes, 2011b). The influence of parental-involvement and family structure have a similar impact on children’s behavior (Shirakawa, 2010; Weaver and Schofield, 2015).

What is vital for parents, educators, and society’s leaders to understand is that parental involvement is not a series of tasks. In other words, it is not simply volunteering at school, attending school functions, checking homework, and establishing household rules. Rather, parental involvement is about love and relationships (Jeynes, 2006a, 2018). It is about bonding with the child. Meta-analytic research has revealed that the most salient components of caregiving, in this sense are largely relational. A meta-analysis is a utilized technique that statistically combines all the research that has been done on a particular topic (Hedges and Vevea, 1998).

That is, a meta-analysis gives a quantitative summary of the entire body of research. Government leaders and academics, in particular, love meta-analyses because rather than read all of the articles and books that have been published on a particular subject, they can just read a meta-analysis on the topic and it summarizes the existing body of research on that topic.

The meta-analyses that have been completed have given considerably greater insight into the workings of caregiving, particularly regarding what components of this engagement are the most efficacious.

Elements of parental involvement

Having high, but reasonable parental expectations

The findings of [Jeynes \(2010\)](#) and [Hattie \(2009\)](#) concur that parents' high, but reasonable, expectations is probably the foremost aspect of parental involvement. Other social scientists agree that having high expectations is the most salient component among all forms of parental involvement ([Klein, 2007](#); [Long and Pang, 2016](#); [Winter, 2006](#)).

Family members should not convey high expectations in a way that rigidly emphasizes, "You will go to Harvard, Oxford, Cambridge, Princeton, or Tokyo University" but rather in a loving and supportive way that subtly communicates that children should do their best ([Jeynes, 2011b, 2018](#); [Phillipson and Phillipson, 2017](#)). As a result of growing up in this kind of an atmosphere, young people understand that in this home it is expected that people give their best effort and aim high ([Klein, 2007](#); [Winter, 2006](#)). School leaders need to communicate to teachers, staff, and parents that one can do more academic and psychological damage to a child by calling them "stupid" than the good that emerges from attending every school function in the year ([Jeynes, 2010](#)). The high expectations that are communicated to youth and adolescents are not to be of the nature that causes undue stress. Instead, they are subtly expressed in parental examples that influence girls and boys on a daily basis.

Many mothers and fathers become involved in the traditional school-based ways, call their child "stupid," and then cannot understand why their child is not performing at his or her potential ([Honore, 2008](#)).

Via meta-analyses it is now evident that the home-based components of parental involvement, such as having high but reasonable expectations of children are, on average, more powerful than the school-based ones ([Harris et al., 2009](#); [Jeynes, 2010](#)). This fact is *extremely important* for school leaders to fully appreciate, because when most educators think of suggestions they can give to fathers and mothers to become more involved, they think of attending school assemblies and events, checking homework, volunteering in school, etc.

Alterable family variable #2: parental style and the importance of love and structure in the home

Research by [Baumrind \(1971\)](#) demonstrates that youngsters develop best if they are reared in a home atmosphere that has a balance of love and structure. Children progress best in an environment in which they feel loved ([Coleman, 2012](#); [Levine, 2012](#)). Youths thrive in circumstances which they feel loved and valued ([Coleman, 2012](#); [Levine, 2012](#)). When they are frequently criticized and put down, it hinders their ability to perform at their highest levels ([Stipek and Seal, 2001](#); [Sax, 2009](#)). As part of the expression of this parental love, mothers and fathers also need to raise their children to exercise self-discipline so that they complete their assignments on time, etc ([Hagelin, 2005](#); [Stipek and Seal, 2001](#)). Parental love also plays a prominent role in helping children to succeed later in life ([Jeynes, 2016b](#); [Levine, 2012](#); [Stipek and Seal, 2001](#)). Parental engagement and making it a priority to keep the family together are probably two of the major expressions of love that caregivers can show ([Hagelin, 2005](#)). Nevertheless, parental love can be manifested in numerous ways and can be a marvelous asset and source of strength to children as they develop.

Sadly, the word *love* is probably the most feared four-letter word among scholars and often in schools as well. However, studies suggest that if schools reach out to children and families in an attitude of love, caregivers are much more likely to become engaged in their children's schooling ([Mapp et al., 2008](#)). Studies utilizing nationwide data sets and meta-analyses indicate that part of the reason why parents become more involved- and students excel more-at religious schools than in public schools is because they feel loved more ([Jeynes, 1999, 2012a](#)). If school leaders create an atmosphere of both love and structure in their schools and instruct parents to become the same way, students and their families will tend to flourish.

Alterable family variable #3: social capital

It is important for educators to acknowledge the social capital to which students have access.

Social capital is defined in the Collins Dictionary as "the network of social connections that exist between people and their shared values and norms of behavior among people, which enable and encourage mutually advantageous social cooperation ... and a sense of belonging." There are many sources of social capital, but the two that are most often referred to are the family and houses of worship ([Coleman et al., 1982](#); [Jeynes, 1999](#)). James Coleman and his colleagues often pointed to the home and the houses of worship as sources of social capital. Technically, Coleman and his colleagues examined churches and synagogues, but one can argue it may well apply to other houses of worship as well. He also believed that one of the strengths of religious education was its ability to bring these two sources of social capital together in order to better educate children ([Coleman et al., 1982](#); [Halpern, 2005](#)). Coleman claimed that one of the reasons why schools of faith did so well in providing scholastic and behavioral tutelage for pupils is because these schools encourage parental participation and hence bring together family- and faith-based sources of social capital ([Coleman et al., 1982](#)).

The power of the family is manifested not only in the atmosphere in the home, but also in various practices of the family. *Communication* is essential if one is to maximize source of social capital that the family is. The communication between parents and their children needs to be both frequent and supportive in order for it to maximize its potential to do good (Chapman and Campbell, 2016; Meyer, 2009). In order to enhance parent-child communication school principals and teachers must not only convey to fathers and mothers just how important this is, but they must also demonstrate exemplary behavior to the children. School staff can talk about communication all they want, but if they do not communicate much or denigrate others when they do express themselves, their words will have little impact (Jansen, 2011). Caregivers can come to a comprehension of this fact, because one of the chief peeves of low-SES parents is that educators speak down to them (Jeynes and Littell, 2000). In loving communication, people view others as more important than themselves (Chapman and Campbell, 2016). If caregivers comprehend that they do not like educators speaking down to them, they will more easily fathom why similarly children do not enjoy their parents speaking down to them.

Aldrich (2012, p. 2) stated in his book *Building resilience: Social Capital in Post-Disaster Recovery*, "I argue that higher levels of social capital—more than such factors as greater economic resources, assistance from the government or outside agencies, and low levels of damage—facilitate recovery and help survivors coordinate for more effective reconstruction." Aldrich (2012, p. 9) continues by stating that, "Empirical studies of the effect of aid and government expenditure have failed to provide evidence of a causal link between more aid money and enhanced recovery."

Robert Putnam (1993) helped popularize the concept in his book, *Making Democracy Work*.

Putnam argued that social capital yielded a type of *civic virtue*. Putnam (1993) believed that there were three aspects of *social capital*: trustworthiness, norms, and networks. All of three of these aspects provide support for students. It helps them develop strong connections with their family, friends, and institutions that are a source of values and inner strength. Often families, friends, and houses of worship provide some of the strongest sources of social capital. In the case of the house of worship, for example, there are values and norms that are taught, networks and connections that are made, and a sense of trustworthiness that results. Similarly, in the case of the families there are certain commonalities that exist in the family, especially to the degree that it develops and unites over a sense of shared values.

Both faith and family often provide a real solid foundation of social capital. They both provide youths a sense of principles and guidelines regarding how the world works and, in contrast, how it should work. James Coleman highlighted the role of religious worship centers in providing social capital both through worship services and their many schools (Coleman et al., 1982). According to Coleman, churches, synagogues, and houses of worship provide children a foundation of an emphasis on love of God and love of neighbor as the most important qualities that one can have, in the midst of a world of largely self-serving people (Gatto, 2001). According to Coleman and his colleagues, churches, synagogues, and faith-based worship schools provide those houses of refuge that people, particularly children and adolescents, in order to maintain their strength to do well in school and in life (Coleman et al., 1982). Families also provide a source of strength and a close network that is a source of great support for youths (Bronfenbrenner, 1979). One of the most important expressions of mothers and fathers providing social capital to help their offspring in school is reading with them. Research evidence indicates that the sooner parents begin to read to their children the better. As a result, it is advisable for family reading to begin well before the child knows the alphabet (Risley and Hart, 2006). Some parents even initiate reading to and talking with their children at birth or even before they are born (Risley and Hart, 2006).

How can schools and their leaders help parents further develop these features of family educational cultures?

With the principles examined in mind, there are actions that teachers and school leaders can take to help shape family educational cultures in such a way that students will benefit.

Educators need to become more aware of how great the impacts are of parental family structure, fatherlessness, and parental involvement on students

School Superintendents, principals, and other school leaders need to understand that family factors can make children either stronger or more vulnerable. Boys from single-parent families are much more likely to end up in gangs than their counterparts in two-parent families (Gerdes, 2012; Popenoe, 2009). Generally speaking, about 90% of adolescents and pre-adolescents in gangs come from single-parent families (Jeynes, 2011a, 2012a; Popenoe, 2009). Psychologists assert that this situation exists, because these youths are looking for a surrogate-father (Popenoe, 2009). As Blankenhorn (1998, p. 88) observes, "Put simply, we have too many boys with guns primarily because we have too few fathers."

When educators are familiar with these realities, they can take action to provide some support that includes (1) helping support parents and the family, (2) taking on certain parental roles that will help the child; and (3) becoming aware of the fact that some people in the world will take advantage of vulnerable students and lead them into trouble. It is a calling to serve and minister to children (Jeynes, 2011b).

Educators need to more expeditiously and correctly identify the trials that children from single-parent-, blended-families, and homes with low levels of parental involvement encounter and how it can ultimately influence their schoolwork, ability to focus, penchant toward depression and even suicide, school behavior and rates of suspension, and likelihood of joining a gang and committing criminal acts of violence (Gerdes, 2012; Jeynes, 2007). Granted, these truths may be unpleasant to digest, but thwarting them or steering a youth away from such activities can often be the most fulfilling aspects of K-12 leadership (Gnezda, 2005; Shore, 2002).

There are a number of reasons why children flourish better in intact families rather than in single-parent families. First, family dissolution through divorce, the death of a parent, or a tenuous relationship that comes to an end are often very distressing experiences for children (Moor and Komter, 2012; Pew Research Center, 2007). Second, marital break-up nearly always causes a drop in family income (Barigozzi et al., 2019; Hampden-Thompson and Galindo, 2015; Pew Research Center, 2007). Third, youths often question the extent to which they are loved when parental relationships collapse (Roberts, 2014; Wallerstein et al., 2002). Fourth, as a result of family break-up, overall children nearly always spend less time with their two parents (Pew Research Center, 2007; Wallerstein et al., 2002).

One of the best ways that instructors can try to compensate for some of the challenges that these young people face is to encourage families to practice parental involvement. Even if a pupil is from a non-intact family, some of the disadvantages can be compensated for if parents are willing to become engaged in their children's schooling and overall education (Jeynes, 2005).

School leaders need to help parents, students, and teachers rediscover the salience of social capital

Forty years ago it was fairly popular for people to talk about the importance of social capital, in large part because of the work of John Coleman and his associates (Coleman et al., 1982). Coleman conducted his research at just the right time. From 1963 to 1980 divorce rates in the United States skyrocketed, the Bible and prayer were removed from the public schools in 1963, and the 1960s and early 1970s in many ways epitomized what it meant for a society to be in turmoil. The assassination of President Kennedy, Martin Luther King, and Robert Kennedy rocked the nation. President Lyndon Johnson sending well over half a million troops to Vietnam and Watergate divided the country (Hunt, 2011). This period was one in which it appeared that the country had lost its moral compass. Coleman's research of the mid-sixties through the early 1980s, in particular appeared to be a call for the nation to regain its bearings, as well as rediscover its foundations and the values that made it great (Coleman et al., 1982).

However, the days when social capital had a prominent place in intellectual discussions have passed. Many caregivers, students, and their families are no longer familiar with the term. Educators often lament the fact that it is easier for them to blame teachers, politicians, and societal factors at large for their children seemingly not living up to their potential than to recognize that the solution of social capital is often not much farther than their own breath (Aldrich, 2012). Therefore, there is a great need to reintroduce this concept and describe the ways that the families can help their children.

A substantial number of social scientists have asserted that they believe that the social capital available to children in the home can give them a decided advantage in their scholastic potential and performance (Coleman et al., 1982; Halpern, 2005). Social capital largely consists of culture, relationships, and moral foundations such as the work ethic that are a source of strength for youngsters (Coleman et al., 1982; Halpern, 2005).

Parents help set a moral tone and foundation for their children in the home. Over the last several decades, there has been a tendency for parents to view the instructor's responsibilities as increasingly greater and their own responsibilities as more and more narrow. In the midst of these trends, many parents do not realize the social capital that they can provide, their house of worship can provide, etc. These centers of worship can provide a sense of purpose in life, which can inspire a strong work ethic (Aldrich, 2012; Coleman et al., 1982). Many churches, for example, also have tutoring services that can help pupils thrive and achieve new heights in school. Many mothers and fathers are not aware of these opportunities (Jeynes, 2011a). Educational leaders can make them more aware of such opportunities and how to capitalize from them.

Principals and other school leaders should train school personnel about the most beneficial kinds of parental -expectations, -style, and -involvement

It was likely a misstep when scholars chose to examine school-based parental involvement well before they gave careful examination to its home-based expression. Having stated this, the error was an understandable one, because academics felt they could control classroom environments far more easily than they could home environments (Harris et al., 2009; Sax, 2009). As a result, teachers often highlight school-based components of family engagement. They also tend to underestimate the primacy of parental expectations and parental style, which usually have their greatest impact at home. However, meta-analytic data indicate that home-based expressions of parenting such as high, but reasonable, expectations and rearing style are actually more powerful and more influential than other components of family rearing. Hence, school leaders need to adjust their advice to parents and teachers, so that it matches the data. It is not only vital that leaders direct family members to understand that parenting is salient in a general sense, but also convey to parents that home-based aspects of parental participation are usually the most vital (Jeynes, 2017, 2018).

Information and direction that school leaders can provide for parents

There is copious amount of information and advice that educators can provide for families that can help them familiarize themselves with the full benefits of mother and father engagement. To be sure, in order for teachers to be efficacious in supplying this advice, they themselves need to be cognizant of parenting research. Unfortunately, the reality is that most school leaders and instructors are not familiar with the body of research in this discipline (Henderson and Mapp, 2002; Jeynes, 2011b). Their knowledge is often limited. As a result, they direct parents into the standard expressions of school-based parental involvement, with little or no emphasis on parental expectations and parental style (Jeynes, 2011a).

It is therefore absolutely necessary for educational leaders to become well versed in the most comprehensive studies examining the benefits of parenting, which usually means those that utilized meta-analyses or the examination of nationwide data sets (Hedges and Vevea, 1998). Once leaders familiarize themselves with the research and literature on parental engagement, they can offer guidance on the following points. First, it is important that educators communicate to mothers and fathers that both the family-based and subtle components of parental involvement are generally the most important (Jeynes, 2010). The vast majority of parents are not aware of this fact. Moreover, they really are not familiar with the aspects of engagement that have the greatest impact (Henderson and Mapp, 2002; Jeynes, 2011b). They are unaware of just how key it is to have high but reasonable expectations, a balanced parental style, and upbuilding communication.

Second, school leaders need to convey certain messages to teachers, for them to then share with parents. Among them are these. For example, mothers and fathers need to understand that the parenting done at home regarding having high but reasonable expectations and the style of parenting that they engage in is generally considerably more important than the parenting they exhibit at school. School leaders need to instill in their instructional staff the salience of love. Johann Pestalozzi (1898, 1916), an educator who had a major influence in the emphasis of early public schools and considered one of the finest educators ever, asked an important rhetorical question. He asked why was it that children learned best at home than in any other environment. Given that the question was a rhetorical one, Pestalozzi (1898, 1916), answered his own question and asserted that the reason was because children were loved at home. High and reasonable expectations and a parenting style that combines love and structure are some of the most important qualities that a parent can have.

School leaders should make sure parents and students who have added hurdles are sufficiently informed about the advantages of parental involvement

It is also important for school leaders to communicate to families of low socioeconomic status, people of color, immigrants, and other groups who might be more hesitant to believe that they can contribute as much as other parents, that *parental involvement works for all groups* (Jeynes, 2016a, 2017; Jeynes and Littell, 2000). Even if fathers or mothers work two jobs, because the most salient forms of parental involvement are home-based and subtle, they can make a huge difference. If the parents are immigrants it is not easy, particularly if their first language is not English (Klein, 2007). However, school leaders can instruct teachers to reach out in love so that especially if the language is common among students, the teacher can, via the parent, learn some of the language and in turn teach the parents more English.

Principals and other leaders should train school personnel about the most beneficial kinds of parental involvement for children of different ages

The results of the meta-analysis provide some excellent guidance for educators in *dealing with children of different ages*. These findings indicate that the type of parenting that is the most effectual during a student's elementary school years is often not the most suitable for the secondary school years (Hulbert, 2003; Steinberg, 2004). It is interesting to note that family activities such as checking homework, establishing household rules for when it is time for work and leisure, and attending student classes and other school functions are associated with higher scholastic outcomes to a statistically significant degree among elementary school students, but have little or no relationship with these outcomes at the secondary school level (Jeynes, 2011b). The reality is that younger children are often far more at ease with direct parental participation than are adolescents (Cherlin, 2008; Graber and Brooks-Gunn, 1996). Adolescents, on the other hand, generally feel more comfortable with less direct parental intervention (Downey and Eccles, 2005; Fuller, 2009).

Unfortunately, most parental involvement strategies pay little or no attention to the age of the students. The overwhelming percentage of parental engagement strategies are based on the pupil's elementary school years rather than the subtler aspects of that involvement that become more salient at the secondary school level.

Principals and superintendents, can also provide families beneficial advice when it comes to rearing practices that are best for different ages of children (Hulbert, 2003; Steinberg, 2004). One of the greatest lessons a mother or father can internalize is that just because you are a good parent with a five year old does not mean that you will be a good parent with a fifteen year old. Helpful parenting practices are not static, but change to some degree depending on the age of the child. Social science statistical analysis now supports this declaration. Over the course of the last three decades, a substantial body of research has accrued indicating that parental involvement has a considerable degree of impact at both the elementary and secondary school level (Green et al., 2007; Domina, 2005; Englund et al., 2004). Meta-analyses indicate that the most effective parental involvement components are not the same among young children versus adolescents (Jeynes, 2011b). That is, more overt expressions of parental involvement are associated with higher academic outcomes among elementary school students, but are no longer associated with these better results when students reach high school (Jeynes, 2005, 2011a). By the time students attain secondary school, virtually all the aspects of parental involvement that yield higher academic outcomes are subtle components (Jeynes, 2011b). School leaders need to adequately disseminate this information to parents.

School leaders need to do more to involve fathers

In reality, when education leaders use the term “parental involvement,” what they often mean is “mother involvement” (Nielsen, 2019, 2020). This is especially the case when there has been family dissolution (Nielsen, 2019). Often, in this case, educators virtually ignore the fathers and in this practice it is often the children that suffer the most (Nielsen, 2019, 2020). However, if educators are to be people who are filled with integrity and keep their word, they need to do all they can to draw in *both* parents (Nielsen, 2019).

A meta-analysis on the effects of father involvement indicates that nearly half of the benefits of parental engagement is generated from dads (Jeynes, 2015b). Great Britain may practice the most outstanding program to inspire father involvement. The UK’s *Sure Start* program appears superior to similar initiatives in the United States, such as *Head Start* (Potter and Carpenter, 2008).

The first way *Sure Start* is superior is that it demonstrates a sophisticated approach to welcoming father involvement in that it seeks to combat sexist stereotypes that portray males as not very nurturing. Potter and Carpenter (2008) reported that fathers often complained that one reason they tended to eschew school-parent gatherings is because they felt that the women stared at them in a judgmental way. *Sure Start* attempts to improve the situations for fathers by making teachers more aware of these stereotypes and the paternal perceptions of them. *Sure Start* advocates also point out that many school leaders subtly impede father participation by providing knitting classes and aerobics workouts, which are overwhelmingly attended by mothers, but not providing other activities that encourage fathers to come to the schools. *Sure Start* has seen a surge in fathers coming to the schools by providing weightlifting classes (Potter and Carpenter, 2008).

Second, *Sure Start* focuses more specifically on drawing in fathers into the schooling process. *Sure Start* advocates that special consideration be given to non-custodial fathers. Nevertheless, these advocates also assert that fathers in two-parent families need to understand their importance in the family. *Sure Start* is designed to communicate to both single parent and two-parent fathers how they can help lift the educational outcomes of boys (Potter and Carpenter, 2008).

Concluding thoughts

For centuries parents and educators have known that the family culture has a major impact on the scholastic outcomes of their children. Social scientists today, via their various studies and theories, are simply confirming this overall reality that the general population has known for centuries. When parents are determined, in both their actions and attitudes, to become more active in their children’s education, they can make a profound difference in how they their offspring do in their classes (Domina, 2005). If school leaders implement the recommendations given in this article, mothers and fathers have a tremendous opportunity to reach their full potential in helping their children excel in their studies. To the extent that school leaders identify what aspects of parenting can make a difference and implement strategies that will make a difference, the impact will probably be greater than what most school officials can currently appreciate. Moreover, it is particularly clear that if school leaders have a deeper appreciation of some of the components of educational culture that are especially underutilized, the difference it would make could be particularly outstanding. For example, social capital can be a powerful influence, but is currently not particularly in vogue. Second, it is extremely rare that educators at any level appreciate and discuss how the effectiveness of different aspects of parental involvement are especially salient at different age levels. Third, few educators have deep discussions of how best to raise children, especially when it comes to addressing high and realistic expectations for children and the importance of a balanced style of parenting. Fourth, few school leaders fully appreciate the important role of fathering.

Previously in this article, it was emphasized that love and relationships are at the heart of effective parenting. Children are smart and can usually tell if their parents love them and want to foster and strengthen a relationship with them (Jeynes, 2007). Consistent with this logic, families can tell if teachers and school administrators truly care about them (Mapp et al., 2008). A school can promote all the efficacious programs that they desire, but unless families perceive teachers and principals as loving and caring individuals who want to build relationships, such programs will have little impact (Mapp et al., 2008). Truly, sincere school leaders can make a difference.

References

- Aldrich, D.P., 2012. Building Resilience: Social Capital in Post-disaster Recovery. University of Chicago, Chicago, IL.
- Barigozzi, F., Cremer, H., Roeder, K., 2019. Till taxes do us *part*: tax penalties or bonuses and the marriage decision. *Eur. Econ. Rev.* 118, 37–50.
- Baskerville, S., 2007. Taken Into Custody: The War Against Fathers, Marriage, and the Family. Cumberland House, Nashville, TN.
- Baumrind, D., 1971. Current patterns of parental authority. *Dev. Psychol.* 4 (1), 1–103.
- Blankenhorn, D., 1998. Fatherlessness is to blame for poverty in single-parent families. In: Bruno Leone, N. (Ed.), *Inequality: Opposing Viewpoints in Sexual Problems*. Greenhaven, San Diego, CA, pp. 86–98.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press, Cambridge, MA.
- Carter, H., Glick, P.C., 2013. *Marriage and Divorce: A Social and Economic Study*. Harvard University Press, Cambridge, MA.
- Chapman, G.D., Campbell, R., 2016. *The Five Love Languages of Children: The Secret to Loving Children Effectively*. Northfield Publishing, Chicago.
- Cherlin, A.J., 2008. *Public and Private Families: An Introduction*. McGraw Hill, Boston.
- Coleman, J., Hoffer, T., Kilgore, S., 1982. *High School Achievement: Public, Catholic, and Private Schools Compared*. Basic Books, New York.
- Coleman, D., 2012. *The Thriving Family*. Hachette Books, Dublin, Ireland.

- Greemers, B.P.M., Reezigt, G.J., 1996. School level conditions affecting the effectiveness of instruction. *Sch. Effect. Sch. Improv.* 7, 197–228.
- Domina, T., 2005. Leveling the home advantage: asserting the effectiveness of parental involvement in elementary school. *Sociol. Educ.* 78 (3), 233–249.
- Downey, G., Eccles, J.S., 2005. *Navigating the Future: Social Identity, Coping, and Life Tasks*. Russell Sage, New York.
- Englund, M.M., Luckner, A.E., Whaley, G.J., Egeland, B., 2004. Children's achievement in early elementary school longitudinal effects of parental involvement, expectations, and quality of assistance. *J. Educ. Psychol.* 96 (4), 723–730.
- Epstein, J.L., 1989. Family structure and student motivation: a developmental perspective. In: Ames, C., Ames, R. (Eds.), *Research on Motivation in Education*. Academic Press, San Diego, CA, pp. 259–295.
- Flecha, R., Soler, M., 2013. Turning difficulties into possibilities: engaging Roma families and students in school through dialogic learning. *Camb. J. Educ.* 43 (4), 451–465.
- Fuller, J.R., 2009. *Juvenile Delinquency: Mainstream and Crosscurrents*. Pearson/Prentice Hall, Upper Saddle River, NJ.
- Gatto, J.T., 2001. *The Underground History of American Education*. Oxford Village Press, New York.
- Gerdes, L.I., 2012. *Juvenile Crime*. Greenhaven Press, Detroit.
- Gnezd, N.M., 2005. *Teaching Difficult Children: Blue Jays in the Classroom*. Scarecrow, Lanham, MD.
- Graber, J.A., Brooks-Gunn, J., 1996. *Transitions Through Adolescence: Interpersonal Domains and Context*. Erlbaum, Mahwah, NJ.
- Green, C.L., Walker, J.M.T., Hoover-Dempsey, K.V., 2007. Parents' motivation for involvement in children's education: an empirical test of a theoretical model of parental involvement. *J. Educ. Psychol.* 99 (3), 532–544.
- Hagelin, R., 2005. *Home Invasion*. Nelson, Nashville, TN.
- Halpern, D., 2005. *Social Capital*. Polity, Cambridge, UK.
- Hampden-Thompson, G., Galindo, C., 2015. Family structure instability and the educational persistence of young people in England. *Br. Educ. Res. J.* 41 (5), 749–766.
- Harris, A., Andrew-Power, K., Goodall, J., 2009. *Do Parents Know They Matter?: Raising Achievement Through Parental Engagement*. Continuum, London.
- Hattie, J., 2009. *Visible Learning: A Synthesis of Meta-analyses Relating to Achievement*. Routledge, New York, NY.
- Hedges, L.V., Vevea, J.L., 1998. Fixed and random effects models in meta-analysis. *Psychol. Methods* 3, 486–504.
- Henderson, A.T., Mapp, K.L., 2002. *A New Wave of Evidence: The Impact of School, Family, and Community Connections on Student Achievement*. Southwest Educational Development Laboratory, Austin, TX.
- Honore, C., 2008. *Upper Pressure: Rescuing Our Children from the Culture of Hyper-parenting*. HarperOne, New York.
- Hornby, G., 2011. *Parental Involvement in Childhood Education: Building Effective School-Family Partnerships*. Springer, New York.
- Hulbert, A., 2003. *Raising America: Experts, Parents, and a Century of Advice About Children*. Knopf, New York.
- Hunt, M.H., 2011. *Lyndon Johnson's War: America's Cold War Crusade in Vietnam*. Hill and Wang, New York.
- Jansen, J.D., 2011. *We Need to Talk*. Booksorm, Northcliff, South Africa.
- Jeynes, W., Littell, S., 2000. A meta-analysis of studies examining the effect of whole language instruction on the literacy of low-SES students. *Elem. Sch. J.* 101 (1), 21–33.
- Jeynes, W., 1999. The effects of religious commitment on the academic achievement of black and Hispanic children. *Urban Educ.* 34 (4), 458–479.
- Jeynes, W., 2003. A meta-analysis: the effects of parental involvement on minority children's academic achievement. *Educ. Urban Soc.* 35 (2), 202–218.
- Jeynes, W., 2005. Effects of parental involvement and family structure on the academic achievement of adolescents. *Marriage Fam. Rev.* 37 (3), 99–116.
- Jeynes, W., 2006a. Standardized tests and Froebel's original kindergarten model. *Teach. Coll. Rec.* 108 (10), 1937–1959.
- Jeynes, W., 2006b. The impact of parental remarriage on children: a meta-analysis. *Marriage Fam. Rev.* 40 (4), 75–102.
- Jeynes, W., 2007. *American Educational History: School, Society and the Common Good*. Sage, Thousand Oaks, CA.
- Jeynes, W., 2010. The salience of the subtle aspects of parental involvement and encouraging that involvement: implications for school-based programs. *Teach. Coll. Rec.* 112 (3), 747–774.
- Jeynes, W., 2011a. Help families by fostering parental involvement. *Phi Delta Kappan* 93 (3), 38–39.
- Jeynes, W., 2011b. *Parental Involvement and Academic Success*. Taylor and Francis/Routledge, New York.
- Jeynes, W., 2012a. A meta-analysis on the effects and contributions of public, public charter, and religious schools on student outcomes. *Peabody J. Educ.* 87 (3), 265–305.
- Jeynes, W., 2012b. A meta-analysis of the efficacy of different types of parental involvement programs for urban students. *Urban Educ.* 47 (4), 706–742.
- Jeynes, W., 2015a. A meta-analysis on the factors that best reduce the achievement gap. *Educ. Urban Soc.* 47 (5), 523–554.
- Jeynes, W., 2015b. A meta-analysis: the relationship between father involvement and student academic achievement. *Urban Educ.* 50 (4), 387–423.
- Jeynes, W., 2016a. A meta-analysis: the relationship between parental involvement and African American School and Behavioral Outcomes. *J. Black Stud.* 47 (3), 195–216.
- Jeynes, W., 2016b. Meta-analysis on the roles of fathers in parenting: are they unique? *Marriage Fam. Rev.* 52 (7), 665–688.
- Jeynes, W., 2017. A meta-analysis: the relationship between parental involvement and Latino student outcomes. *Educ. Urban Soc.* 49 (1), 1–28.
- Jeynes, W., 2018. A practical model for school leaders to encourage parental involvement and parental engagement. *Sch. Leader. Manag.* 38 (2), 147–163.
- Jeynes, W., 2019. A meta-analysis: the relationship between character education programs and student outcomes. *Educ. Urban Soc.* 51 (1), 33–71.
- Katz, R., Lowenstein, A., Halpern, D., Tur-Sinai, A., 2015. Generational solidarity in Europe and Israel. *Can. J. Aging* 34 (3), 342–355.
- Klein, A., 2007. *A Class Apart: Prodigies, Pressure, and Passion Inside One of America's Best High Schools*. Simon and Schuster, New York.
- Kyriakides, L., Creemers, B.P., 2008. Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: a study testing the validity of the dynamic model. *Sch. Effect. Sch. Improv.* 19, 183–205.
- Leithwood, K., Patten, S., Jantzi, D., 2010. Testing a conception of how school leadership influences student learning. *Educ. Adm. Q.* 46 (5), 671–706.
- Levine, M., 2012. *Teach Your Children Well: Parenting for Authentic Success*. Harper, New York.
- Long, H., Pang, W., 2016. Family socioeconomic status, parental expectations, and adolescents' academic achievements: a case of China. *Educ. Res. Eval.* 22 (5/6), 283–304.
- Mapp, K.L., Johnson, V.R., Strickland, C.S., Meza, C., 2008. High school family centers: transformative spaces linking schools and families in support of student learning. *Marriage Fam. Rev.* 43 (3/4), 338–368.
- McLanahan, S., Sandefur, G., 1994. *Growing Up With a Single Parent: What Hurts, What Helps*. Harvard University Press, Cambridge.
- Meyer, J., 2009. *The Love Revolution*. Faithwords, New York.
- Mincy, R.B., Jethwani, M., Klempin, S., 2015. *Failing Our Fathers: Confronting the Crisis of Economically Vulnerable, Nonresident Fathers*. Oxford University Press, Oxford, UK.
- Moor, N., Komter, A., 2012. The impact of family structure and disruption on intergenerational emotional exchange in Eastern Europe. *Eur. J. Aging* 9 (2), 155–167.
- National Academy of Sciences, 2005. *Facilitating Interdisciplinary Research*. National Academic Press, Washington, DC.
- Nielsen, L., 2019. *Father–Daughter Relationships: Contemporary Research and Issues*. Routledge, New York.
- Nielsen, L., 2020. *Improving Father–Daughter Relationships: A Guide for Women and Their Dads*. Taylor and Francis, Milton, UK.
- Pestalozzi, J., 1898. In: Holland, L., Turner, F. (Eds.), *How Gertrude Teaches Her Children: An Attempt to Help Mothers to Teach Their Own Children and an Account of the Method*. Bardeen, Syracuse.
- Pestalozzi, J., 1916. How a child is led to god through maternal love. In: Green, J.A. (Ed.), *Pestalozzi's Educational Writings*. Edward Arnold, London.
- Pew Research Center, 2007. *As Marriage and Parenthood Drift Apart, Public is Concerned About Social Impact*. Pew Research Center, Washington, DC.
- Phillips, R., 1988. *Untying the Knot: A Short History of Divorce*. Cambridge University of Press, New York.
- Phillipson, S., Phillipson, S., 2017. Generalizability in the mediation effects of parental expectations on children's cognitive ability and self-concept. *J. Child Fam. Stud.* 26 (12), 3388–3400.
- Popenoe, D., 2009. Cohabitation, marriage and child well-being: a cross-national perspective. *Society* 46 (5), 429–436.
- Potter, C., Carpenter, J., 2008. Something in it for dads. Getting fathers involved in Sure Start. *Early Child. Dev. Care* 18 (7/8), 761–772.

- Putnam, R., 1993. *Making Democracy Work*. Princeton University Press, Princeton, NJ.
- Raymo, J.M., Fukuda, S., Iwasawa, M., 2013. Educational differences in divorce in Japan. *Demogr. Res.* 28 (6), 177–206.
- Risley, T., Hart, B., 2006. Promoting early language development. In: Watt, N.F., Ayoub, C., Bradley, R.F., Puma, J.E., LeBeouf, W.A. (Eds.), *The Crisis in Youth Mental Health: Critical Issues and Effective Programs*, vol. 4. Praeger, Westport, CT, pp. 83–88.
- Roberts, R., 2014. *Parents with Price Tags: Heal Yourself of Family Dysfunction and Love Your Children Unconditionally*. BookBaby, Cork.
- Robinson, K., 2014. *The Broken Compass*. Harvard University Press, Cambridge, MA.
- Sax, L., 2009. *Boys Adrift: The Five Factors that Driving the Growing Epidemic of Unmotivated Boys and Underachieving Young Men*. Basic Books, New York.
- Shirakawa, T., 2010. Family structure and children's reading literacy: academic achievement of children from single-parent families in Japan and the United States. *Sociol. Theory Methods* 25 (2), 249–265.
- Shore, R., 2002. *What Kids Need*. Beacon Press, Boston.
- Steinberg, L.D., 2004. *Ten Basic Principles of Good Parenting*. Simon and Schuster, New York.
- Stevenson, B., Wolfers, J., 2007. *Marriage and Divorce: Changes and Their Driving Forces*. Bureau of Economic Research, Cambridge, MA.
- Stipek, D.J., Seal, K., 2001. *Motivated Minds: Raising Children to Love Learning*. Holt and Co, New York.
- Tan, C.Y., 2017. Do parent attitudes toward their children's education and future jobs matter for their children's school. *Br. Educ. Res. J.* 43 (6), 1111–1130.
- Turkat, I.D., 1995. Divorce-related malicious mother syndrome. *J. Fam. Violence* 10 (3), 253–264.
- United States Department of Health and Human Services, 2012. *Statistical Abstracts of the United States*. U.S. Department of Health and Human Services, Washington, DC.
- United States Department of Justice, 1999. *Age-Specific Arrest Rate and Race-Specific Arrest Rates for Selected Offenses 1965–1992*. U.S. Department of Justice, Washington, DC.
- Vaughans, K., Spielberg, W., 2014. *The Psychology of Black Boys and Adolescents*. Praeger, Santa Barbara, CA.
- Wallerstein, J.S., Lewis, J., Blakeslee, S., 2002. *The Unexpected Legacy of Divorce: A 25 Year Landmark Study*. Hyperion, New York.
- Wartman, K.L., Savage, M., 2008. *Parental Involvement in Higher Education: Understanding the Relationship Among Students, Parents, and the Institution*. Wiley/Jossey-Bass, San Francisco.
- Weaver, J.M., Schofield, T.J., 2015. Mediation and moderation of divorce effects on children's behavior problems. *J. Fam. Psychol.* 29 (1), 39–48.
- Winter, J., 2006. *Breakthrough Parenting for Children With Special Needs: Raising the Bar of Expectations*. Jossey-Bass, San Francisco.

Leadership and family factors: critical perspectives on international leadership standards, models, and practices

Anne Marie FitzGerald^a and Sandra Quiñones^b, ^a University of Prince Edward Island, Charlottetown, Canada; and ^b University of Connecticut, Storrs, CT, United States

© 2023 Elsevier Ltd. All rights reserved.

Impacts of family engagement and school leadership	272
Defining family engagement	273
Models of family engagement	273
Historical look at family engagement: pendulum swings of power and control	273
Leadership standards related to family and community engagement	274
The influence of Western leadership standards in policy borrowing	274
International leadership standards and family engagement	275
Family engagement around the world: an international view	276
Comprehensive reform efforts: community schools as a promising strategy	278
Conclusion	278
References	279

Three factors with especially significant effects on student achievement are teacher quality (Wayne and Youngs, 2003), school leadership (Leithwood et al., 2020), and family engagement (Smith et al., 2021). Here, we focus on the intersection of school leadership and family factors. In the past, school-centric foci and the practices of dominant families shaped definitions of parental involvement and family engagement (Ishimaru, 2014, 2020). Thus, we begin by using a critical lens to define and describe models of family engagement. Next, we take a historical look at the evolution of family engagement in a Western context and use an international perspective to discuss promising practices globally. Here, we take an international perspective and elucidate promising levers and pathways, such as targeted leadership standards for family engagement, equity-based theories that inform models of family engagement, and results-based system reform efforts such as community schools.

Impacts of family engagement and school leadership

Family engagement in children's education correlates positively with a range of positive outcomes for students, such as attendance, achievement, and social and emotional wellbeing (Jeynes, 2018, 2021; Ma et al., 2016; Jacobson and Blank, 2015a,b). Schools shape and are shaped by their communities. In other words, context matters. Specifically, social and economic conditions impact families, communities, and schools. Educational equity gaps based on culture, race, ethnicity, language, sexual orientation, and class (Ishimaru and Galloway, 2014; Leithwood, 2021) can be mitigated or worsened, depending upon the internal compass, values, knowledge, and skills of school leaders (Briscoe and Pollock, 2017; Leithwood, 2021). School leadership impacts students' lives both directly and indirectly (Leithwood et al., 2020). Evolving leadership standards, enriched preparation for aspiring principals, and more nuanced evaluation of practicing principals acknowledge the importance of partnering with families and communities to enhance equitable access to education (Alberta Teachers' Association and Canadian Association of Principals, 2014; Young and Perrone, 2016).

In a volatile, uncertain, complex, and ambiguous (VUCA) world, there is room for hope and opportunity for change. Multiple sources of injustice continue to impact the educational experiences of students and families around the world. Social justice literature identifies systemic racism, socioeconomic, digital divides, and colonialism as factors contributing to structural injustice.

We describe structural injustice as associational, distributive, and cultural. The term associational refers to the ability of all groups to govern, participate, and make decisions that affect their lives (Gerwitz and Cribb, 2002; DeMatthews et al., 2016). Thus, associational injustices include marginalized communities with limited (or no) participation and/or ongoing mistrust with social groups that engage with decision-making practices. Distributive injustices include inadequate housing, health care, education and economic opportunities. Cultural injustices include racism, nonrecognition or invisibility, deficit perspectives, othering, stigma, and exclusion.

We conceptualize colonialism as a structural injustice encompassing associational, distributive, and cultural factors (Gerwitz and Cribb, 2002; DeMatthews et al., 2016). Horvath (1972) defined colonialism as: "a form of domination—the control by individuals or groups over the territory and/or behavior of other individuals or groups"; as possessing "political" and "economic variables"; and as a "culture-change process" (p. 46). Horvath differentiated colonialism from imperialism by the presence of "large numbers of settlers" with aims of "assimilation" (p. 47). Battiste (2004) further defined colonialism to include "cognitive colonialism" (p. xvi) that described the centering of Western values and forms of knowledge in education. All of these factors represent structural

injustices that are aggravated further by global stressors, such as pandemics, climate change, political and economic instability, and violence.

These intersecting factors continue to shape present systems and in many cases oppress non-dominant families. We define nondominant families as “those impacted by systemic oppression, such as marginalization based on race, class, language, or immigration status” (Ishimaru, 2020, p. 8). In some instances, the perspective of parents is not sought or acted upon. However, some policy makers, systems, and schools are shifting the narrative and taking collective action. This view of nondominant families is applicable to a wide variety of contexts at an international level. Such an approach accentuates the importance of power differentials and context.

Defining family engagement

We conceptualize family engagement as full, equal, and equitable partnership among families, educators, and community partners to promote children’s learning and development from birth through college and career (Mapp and Bergman, 2021). Rather than presenting this definition as the most ideal or authentic form of family engagement, we view it as an aspirational definition located on one end of the continuum. This definition of family engagement was co-created with families, policy makers, educators, and community leaders as part of collective efforts to enact more liberatory (free of dominance), solidarity-driven (in union and fellowship), and equity-focused (fair and just) family engagement practices in schools (Mapp and Bergman, 2021). This approach views family engagement as a practice—not a program—that is interrelated with teaching and learning, health and wellbeing, neighborhood development, and overall community uplift, particularly in relation to nondominant families. With this equity-based view of family engagement in mind, we now turn to models of family engagement.

Models of family engagement

The above definition builds upon and expands previous models of family engagement, such as Joyce Epstein’s (1987), Epstein and Dauber (1991) influential model of parental involvement which outlines six categories by which educators can involve parents in their children’s education (parenting, communicating, volunteering, learning at home, decision-making, and collaborating with the community). Epstein et al. (2018) extended the term “parents” to include guardians, foster parents, and extended family members and argued for “overlapping” rather than separate spheres of influence in a child’s life. In other words, children learn at home, in school, and in the community; thus, schools and families must partner. By its’ very structure, early versions of the Epstein model reflected uni-directional ways of thinking about families and outlined school-centric avenues to involve parents. Later versions of the model elaborated examples of working *with* families and identified specific challenges to implementing the model. Terminology in the field shifted from the phrase “parental involvement” to “family engagement,” reflecting awareness of (a) the multiple configurations of family and (b) the need for family agency and autonomy, and (c) the need for bi-directional communication, decision-making, and action (Weiss et al., 2018).

In line with this shift in terminology, Hong (2011) described three phases or “a cord of three strands” (induction, integration, and investment) that characterized successful community-based family engagement. Hong (2011) contended that equitable change involved working across ecologies of home, school, and community to make a cultural shift that honored the ways in which parents contribute to their children’s education. Similarly, Auerbach’s (2010, 2012) model of family engagement provided guidance to school leaders by outlining a continuum toward authentic partnerships with families. Elements of the continuum include how leaders position families and community organizations and the extent to which leaders commit to co-constructing goals and reducing power differentials between educators and families. Auerbach’s model identified authentic family and community engagement as a matter of social justice.

At the same time, education scholars of color propose equity-based conceptualizations of family engagement. For example, Ishimaru (2019) calls for family engagement strategies to redress “asymmetrical power dynamics, and cultural, class, and language divides” (p. 351). Alameda-Lawson and Lawson (2019) proffer a relational model of collective parental engagement (CPE) that seeks to engage groups of parents in the community and school by cultivating change at individual, social, and institutional levels. Hernandez et al. (2014) emphasize the racial identity development of educational leaders, thus leading them to prioritize racial identity development in students and culturally responsive ways of engaging families.

In summary, these models illustrate the need for families to exercise agency in deciding how, when, and where they will engage in their children’s education. In other words, there is no “culture-less” or isolated family engagement. Thus, in what follows, we take a historical approach that takes into account power dynamics and contextual factors that shape family engagement and school leadership.

Historical look at family engagement: pendulum swings of power and control

Prior to discussing current issues in and future directions of family engagement and school leadership, it is important to consider historical accounts that reveal cyclical patterns and ongoing issues in this focus area. Therefore, in this section, we turn our attention

to parent involvement from a historical perspective and utilize the United States as a contested terrain with shifting power dynamics over time.

In a historical perspective of parent involvement in American public schools, [Hiatt-Michael \(1994/2000\)](#) highlights the “pendulum swings” from parent involvement in the home and parent control of schools, followed by local control of education, to the eventual bureaucratic factory model schools of the industrial revolution. Hiatt-Michael also describes how many schools in the colonial period were organized along social class, but later “the shift from parent education to public education occurred first in Europe and then was transmitted to America” (p. 248).

After the Revolutionary War era, American leaders promoted tax-supported universal education for all children (including the poor), with particular attention to teaching reading, writing, and rhetoric to America’s European immigrants. Subsequently, America’s public school system functioned as tools of socialization and assimilation toward a “common culture,” particularly in heavily populated urban cities ([Pulliam, 1987](#); [Spring, 1986](#); [Tyack, 1974](#)). Later, there was a rise in compulsory education and child labor laws, followed by the development of school bureaucratization and the professionalization of teachers ([Hiatt-Michael, 1994/2000](#)). This was the time of the single-room schoolhouse with children of multiple ages that was organized and operated by a locally hired teacher. Soon after, “children were classified by grade” (p. 252) as part of a graded school plan.

During this time period in America’s public schooling, “the increasing separation between parent control and public school was perceived by parents,” and mothers (mostly middle and upper class) sought intervention. This led to the National Congress of Mothers (NCM) in 1897, a convening which later turned into what is now known as the Parent Teacher Association (PTA). According to [Davies \(2000\)](#), the PTA served to Americanize newcomers to the country and to teach middle class parenting. As parents regained more control, a growing number of court cases ensued, including landmark court cases such as *Mendez v. Westminster* (segregation of Mexican Americans in California), *Brown v. Board of Education of Topeka* (segregation of Black and White children in US) and *Lau v. Nichols* (bilingual education in the US). Subsequently, “many parents became so enraged” with desegregation practices “that they removed children from public schools and enrolled them in private schools, created new schools, or began home schooling” ([Hiatt-Michael, 1994/2000](#), p. 256). Now, in the contemporary information age, the “pendulum appears to be swinging back again ... toward schooling which increasingly involves parents” (p. 256). There is an increasing emphasis on collaboration with the aim of better understanding and supporting families and schools as interdependent (rather than independent) entities (p. 257). Moreover, [Baquedano-López et al. \(2013\)](#) describe how inequities in parental and community involvement in US schools are longstanding and persisting issues. In the United States context, schools are essentially run by dominant populations and there is no cultureless family engagement; it is integral part of what anthropologists and academics refer to as “culture wars” in local and global contexts.

In summary, a historical look reveals that leadership and family engagement practices function within a complex and dynamic continuum: at times “at odds” with each other (high tension), at times in harmony (low tension), and most of the time in a state of productive struggle (somewhere in the middle). To this extent, we see how pendulum swings among multiple stakeholders with competing interests are an ongoing issue in school leadership and family engagement in education. Next, we turn our attention to the professionalism of school leaders as signaled through standards, the influence of Western standards in international contexts, and the extent to which standards highlight the importance of leaders’ work with families and communities.

Leadership standards related to family and community engagement

The influence of Western leadership standards in policy borrowing

We begin by discussing the purpose, evolution, and influence of leadership standards in Western and Anglophone settings (US, Canada, and UK), follow with a discussion on how select countries around the world have “borrowed” and adapted these standards, and conclude with a particular focus on standards guiding educational leaders’ work with families and communities.

Standards, both for in-practice educational leaders and for educational leadership preparation programs, create common expectations for the preparation and work of principals and, in many cases, form the basis of their evaluation. The [Center for Study of Policies and Practices in Education \(2013\)](#) performed a comparative study of standards in countries belonging to the Organization for Economic Cooperation and Development (OECD) and contend that standards “clarify and provide focus to the principals’ function” (p. 49). Leadership standards are not without critics. For example, [Riveros et al. \(2016\)](#) contend that standards create a false optimism that the agency of leaders can overcome systemic and structural obstacles. [Weinstein and Hernández \(2016\)](#) argue that standards created in Western context and culture cannot be transplanted to other contexts and cultures, such as those found in Latin American countries. [Carrasco et al. \(2021\)](#) contend such standardization decenters educators’ professional judgment and attempts to provide cookie cutter solutions to unique school environments. Furthermore, scholars [Moorosi \(2020\)](#) and [Davis et al. \(2015\)](#) critique the “colour-blind” language of standards documents. In a similar vein, [Antony-Newman \(2019a,b\)](#) analysis of educational policy documents in Ontario reveals narrow definitions of parent involvement laced with discourses of barriers and parental deficiency (p. 143). Moreover, issues of inequality are omitted rather than positioned as central in policies and practices. However, together with the intrinsic values school leaders bring to their work ([Gu et al., 2018](#)) leadership standards hold the potential to improve the lives of students and families. Here, we examine the signaling of educational policy makers through standards and the relative importance placed on engaging families and communities in children’s education.

In the 1990s, the standards movement emerged in Western and Anglophone countries and spread rapidly to non-Anglophone countries ([Berkovich and Bogler, 2020](#); [Flessa et al., 2018](#)). In 1996, the US based Interstate School Leaders Licensure Consortium

(ISSLC) created a set of six standards for school leaders. Standard 4 addressed principals' work with parents and families: "Collaborating with family and community members, responding to diverse community interests and needs and mobilizing community resources." Recently, the National Policy Board of Educational Administration (NPBEA) revised these professional standards and reworded the standard addressing family and community engagement (Young and Perrone, 2016). "Effective educational leaders engage families and community in meaningful, reciprocal, and mutually beneficial ways to promote each student's academic success and well-being" (National Policy Board for Educational Administration, 2015). In addition, cultural responsiveness with an emphasis on equity earned its own standard, highlighting the need for an assets-based and holistic approach to educational leadership (Murphy et al., 2017). Thus, in the US context, leadership standards promoted working *with* families and communities as integral to the work of school leaders.

In Canada, the Ontario Leadership Framework, introduced in 2006 and revised in 2013 in response to evolving leadership research (Leithwood, 2012), outlined five core capacities for school leaders (setting goals, aligning resources, promoting collaborative learning cultures, using data, and engaging in courageous conversations). In a more fulsome description, the School Level Leadership model detailed an inclusive approach to "Building productive relationships with families and communities" that included welcoming parents, viewing them as partners, building trust, and working with diverse families. Similarly, in the United Kingdom (UK), National Standards for Headteachers outline 10 broad areas of responsibility, and the ninth standard, Working in Partnership, states that headteachers must "forge constructive relationships beyond the school, working in partnership with parents, carers and the local community" (UK Government, 2021). In sum, leadership standards in three Western and Anglophone countries specifically signal the importance of partnering with families and communities.

International leadership standards and family engagement

Non-Anglophone education systems around the world quickly adopted or adapted standards to their own contexts or created their own educational leadership standards (Berkovich and Bogler, 2020) with a greater or lesser degree of guidance regarding the work of school leaders with families and communities. Emerging research on the professionalization of school leadership in non-Anglophone countries *and* space constraints limits our discussion to selective and illustrative examples. Next, we consider a cross-cultural analysis of leadership standards in China (Wei, 2017); an examination of standards' adoption in eight countries/subregion in Latin America (Weinstein and Hernández, 2016) with a particular emphasis on Chile, and a critical view of the standards document informing the work of South African school leaders (Moorosi, 2020).

The Chinese Ministry of Education adopted a set of professional standards for school leaders in 2013 that were influenced by the ISSLC (2008) standards. Wei (2017) performed a cross-cultural analysis of these standards and concluded that, while there were similarities between the US and Chinese standards, significant differences could be explained by unique cultural and national contexts. A crosswalk between the ISSLC (2008) standard regarding families and communities (Collaborating with faculty and community members, responding to diverse community interests and needs, and mobilizing community resources) and the corresponding Chinese standard reveals one of these differences: school leaders were to "adjust to external contexts."

In Latin America, the adoption or adaptation of leadership frameworks and standards has varied (Flessa et al., 2018). Weinstein and Hernandez, 2016) studied eight Latin American school systems, seven countries (Argentina, Chile, Colombia, Ecuador, Mexico, Dominican Republic, and Peru) and one Brazilian sub-region (Ceará). Results revealed a wide range: systems with centrally developed leadership frameworks and standards in place (Chile, Colombia, and Ecuador) to systems with no leadership standards in place (Argentina, Ceará/Brazil). However, Brazil's recent adoption of learning standards for students (Rivas and Sanchez, 2020) may indicate that leadership standards are not far behind. In Chile, for example, The Framework for Good Principals and School Leadership (Ministerio de Educación de Chile, 2015) delineates a number of areas (including vision, goal setting, and monitoring of instruction) that guide principals' behaviors. The framework includes direction on how principals should work with families to "encourage and facilitate the involvement of families in the education of their children ... for which they provide information and generate concrete and systematic spaces to involve them, in order to generate a collaborative relationship" (translated from Spanish, Sisto, 2018, p. 149). Independent of the extent to which standards align with professional development for school leaders or leaders' actual behaviors, Chilean educational policy points to the priority placed on principals working collaboratively with families.

In 2015, the South African government formally adopted "The South African Standard for Principals" which provides guidance for school leaders in eight key areas, including a particular emphasis on the processes of teaching and learning (Moorosi, 2020). The sixth area, working with and for the community, advises principals to be cognizant of "the socio-economic, political and cultural characteristics of the wider school community" and further directs principals to "build and maintain partnerships between the school, the home and the wider community" (Department of Basic Education, South Africa, 2015, p. 21). However, within the context of highly racialized educational and community environments, Moorosi (2020) critiques the "colour-blind" language used throughout the policy.

Standards for school leaders continue to be created, borrowed, and/or adapted in non-Western countries around the world. In many cases, professional development for principals has yet to be offered and changes in practice has yet to be documented in research. Nevertheless, as Table 1 below illustrates, a clear emphasis on working *with* families and community partners in a reciprocal and equity-minded manner is evident in an examination of policy documents.

Table 1 A cross-country comparison of school leaders' professional standards related to family and community engagement.

Country	Author and date	Leadership standards	Standard related to family and community engagement
United States	National Policy Board of Educational Administration, 2015	Professional Standards for Educational Leaders (PSEL)	"Effective educational leaders engage families and community in meaningful, reciprocal, and mutually beneficial ways to promote each student's academic success and well-being" (PSEL, 2015).
United Kingdom	Department for Education, UK Government, 2020	National Standards for Headteachers	Headteachers must "forge constructive relationships beyond the school, working in partnership with parents, carers and the local community" (UK Government, 2020).
Ontario, Canada	Ontario Ministry of Education, 2013	The Ontario Leadership Framework, School Level Leadership Model	School leaders must "Build productive relationships with families and communities" (Ontario Ministry of Education, 2013) that include welcoming parents, viewing them as partners, building trust, and working with diverse families.
China	The Ministry of Education for the People's Republic of China, 2013	Professional Standards for Compulsory Education School Principals	The public relations standard encourages, "the capacity to negotiate with public relations, and to establish collaborations between the school, government, families, communities, and institutions" (Ministry of Education for the People's Republic of China, 2013).
Chile	Ministerio de Educación de Chile, 2015	The Framework for Good Principalship and School Leadership	States that principals should work with families to "encourage and facilitate the involvement of families in the education of their children ... for which they provide information and generate concrete and systematic spaces to involve them, in order to generate a collaborative relationship" (translated, Sisto, 2018, p. 149).
South Africa	Department of Basic Education, South Africa, 2015	The South African Standard for Principalship	The sixth area, working with and for the community, advises principals to be cognizant of "the socio-economic, political and cultural characteristics of the wider school community" and further directs principals to "build and maintain partnerships between the school, the home and the wider community" (Department of Basic Education, South Africa, 2015, p. 21).

Family engagement around the world: an international view

We know that school leaders play an important role for establishing school-family partnerships (Jung and Sheldon, 2020) and positioning family engagement "both as an outcome *and* a means for providing a quality education for students" (emphasis added, DeMatthews et al., 2016, p. 756). However, as noted earlier, we live in a challenging and changing world that shapes the everyday context of schools, families, and communities. Thus, "research in highly demanding contexts across the globe is needed to further develop theory and practical applications" in the field (DeMatthews et al., 2016, p. 755). Accordingly, in what follows, we discuss research conducted in highly demanding contexts. We also consider international contexts related to Indigenous education.

In the context of the United Kingdom, contemporary social challenges that shape school leaders' work include security/safety, poverty, inequality, chronic absenteeism, and high school dropout rates. In a study examining the dynamics between primary school leadership and parental engagement in the United Kingdom, Mleczko and Kington (2013) demonstrate "the importance of the school principal in facilitating opportunities for increased parental engagement that promoted social inclusion and cohesion between/among parents, volunteers, and professionals, particularly in the context of highly disadvantaged areas" (p. 143). The schools that were studied as part of this qualitative research were "characterized by a high presence of ethnic minority children and were located in a neighborhood of low social and economic status in the Midlands region of the United Kingdom" (p. 129). The findings from the research brings evidence "that the role of the school principal is crucial in the introduction, implementation and sustainability of solutions focused on parental involvement, as well as bringing benefits to the social cohesion of the local community" (p. 143). In their own words, the scholars noted:

... [A]s part of establishing a culture in schools that encourages parental involvement, it is important that the leadership vision for the school is clearly communicated, and responsive to the school and community context. Furthermore, this vision needs to be positive and optimistic in tone if members of staff are to feel that they have the opportunity to innovate and be part of a team that can achieve the goal of continuing to foster improved parental involvement (p. 143).

In other words, school leaders play a key role in mitigating distributive and cultural factors that shape school-family dynamics and promote social cohesion within ethnically diverse communities.

In Mexico, DeMatthews et al. (2016) offer a qualitative case study exploring the enactment of social justice leadership in a private elementary school located in the *colonia* of Ciudad Juarez, a highly demanding context in the US-Mexico border. As noted by the authors,

Colonias are unincorporated and unregulated settlement with substandard housing and living conditions that present numerous obstacles to student well-being and achievement, including (a) high proportions of adult illiteracy; (b) lack of access to quality public schools; (c) severe poverty; (d) domestic violence, gang/cartel violence, and the persistence of violent and nonviolent crime; (e) government malaise/corruption; and (f) a lack of community cohesion and solidarity

Bejerano (2002), Hernandez and Grineski (2012), Heyman and Campbell (2004), Hill (2003), p. 756.

The research describes how the school principal prioritized meaningful family engagement as part of addressing *out-of-school* factors that influenced the social and academic outcomes of students. For instance, to address issues related to violence, extortion, governmental oppression, corruption, and precarious housing and everyday living conditions, the school principal engaged with family members to increase safety and accessibility, provide expanded learning opportunities (e.g., adult education, extended day, after school programs), and cultivate critical leadership skills. While this study used a year-long data collection process, it is important to note that the principal, the school's founder, had been working with families and the community for over a decade. With this sustained focus, the principal's aim was to disrupt status quo thinking, address local problems, and support long-term community capacity development. In other words, the explicit intention was to (re)build trust and increase family engagement and community solidarity over time. In so doing, the principal worked with families in an inclusive and democratic manner that leveraged school resources and supported marginalized families.

In the Australian context, cultural and linguistic diversity, especially with regards to Aboriginal children and families, has inspired school leaders to move toward a two-way street approach to school-family partnerships in early childhood and elementary schools. A one-way street approach is one where schools focus on educating families as part of pre-decided school-based efforts and goals (Sisson et al., 2021). However, a more democratic two-way and parents/families, teachers/staff, and school leadership collaborate and learn together. To this extent, Sisson and colleague's multiple case study research explored how two schools in Australia engaged with Reggio Emilia education principles to reimagine school-family partnerships (p. 2). As a result of engaging with theory-grounded family engagement practices, school leaders and teachers enacted a more collaborative "two-way street" approach when working with families. This approach included co-designing school spaces and co-planning school-based programming and community-wide events that were informed by shared goals. Here it is also important to point out that school leaders *and* teachers collaborated in reframing family engagement efforts at both school communities. In other words, school leaders cannot do this work this work alone. Teachers, staff, families, and neighborhood residents are all important stakeholders in the process of moving toward a two-way street approach to school-family partnerships in education.

In recent years, there has been a growing recognition of the ways in which a long and grievous history of colonization in countries around the world continues to impact negatively the school experiences of Indigenous students and families (Battiste, 2017; Bishop, 2021; Smith et al., 2018). In particular, in countries such as Canada, Australia, French Indochina, and the United States, government officials removed Indigenous and part Indigenous children from their homes and families and placed them in residential schools where children faced inhumane conditions and abuse (Firpo and Jacobs, 2018). Bang et al. (2018) write "systems of education for Indigenous children and youth often remain sites of trauma, assault, and aims of Indigenous erasures" (pp. 1–2) and advise instead, family engagement should honor "Indigenous families' and communities' ways of knowing and being" and engage "Indigenous families and communities as dreamers, nation-builders, and future elders" (Bang et al., 2018, p. 2). In a similar vein, Garcia (2014) queries how re-framing parental involvement from an Indigenous perspective might "disrupt notions of treating parents as uninvolved or deficient, and begin to privilege Indigenous families as a source of strength and commitment to education from an Indigenous *and* western perspective?" (p. 61). Corn tassel and Hardbarger's (2019) work elucidates the importance of everyday actions *in* families and communities and *on* the land that forms the basis for education and decision-making from an Indigenous perspective. Thus, from an Indigenous perspective, parents and families play a central, not peripheral, role in the perpetuation of knowledge and land-based, not school-based, education is prioritized.

While most of the aforementioned research accounts for single and/or multiple case studies conducted at the school site level, a growing number of scholars are examining family engagement from a broader, yet place-based perspective that accounts for school-family-community dynamics within a district or in a geographic area ((Alameda-Lawson, 2016) Hopson, 2014). For instance, a qualitative thematic synthesis (Alameda-Lawson, 2014; Lawson et al., 2016) of school-family-community partnerships in three low-income, ethnically concentrated urban school communities in the United States (one in South Florida and two in Northern California) utilized a *collective parent engagement approach* (CPE). Social workers engaged with parents to design and develop school-based and/or school-linked programs and services that could support the strengths, needs, and challenges of children and families in the neighborhood community (p. 1085). The authors noted, "when implemented as an isolated or 'stand-alone' service strategy, CPE generally does not influence school outcomes. But when tied to a broader system of reform efforts,

CPE can help transform the social-institutional landscape of low-income urban school communities" (p. 1085). Thus, the findings "serve as an important reminder of the limitations of any single intervention to affect change, especially interventions that are not well integrated with the work of school leaders, teachers, social and neighborhood providers, funders, and policy makers (Gutierrez and Penuel, 2014)" (p. 1114). Accordingly, this body of literature supports the idea that family engagement is not a program or a service, but rather a set of leadership practices and mediating conditions in schools and families that form part of broader improvement goals or reform efforts (Mapp and Bergman, 2021; Leithwood, 2021, in press).

Comprehensive reform efforts: community schools as a promising strategy

An international review of the scholarship about leadership factors and family engagement points to existing and thriving practices that are important to recognize, support, and study further. Thus, we highlight the **community school model** as an international approach to education that is grounded school-family-community partnerships (Anderson-Butcher and Ashton, 2004; Anderson-Butcher et al., 2018; Dryfoos, 1994; Dryfoos et al., 2005; Galindo et al., 2017; Galindo and Sanders, 2019a,b (Sanders, 2006, 2016, 2018)). As noted by Jenkins and Duffy (2016), community school is an umbrella term that refers to "schools that adopt a broad and varying range of services to address the comprehensive needs of students, families, and communities (p. 2)." Accordingly, "each community school should be unique to serve the specific needs of the local population" (p. 2). Moreover, the "composition of services and support might also vary due to funding constraints, funding considerations, or other factors" (p. 2). However, a key component of the community school model is the use of external partnerships "to transform the school into a neighborhood hub for social services and integrated student supports" (p. 2).

Since approaches to school-community partnerships vary widely in scope, purpose, and requirements, Valli and colleagues (Stefanski et al., 2016; Valli, 2016; Valli et al., 2014) offer a helpful typology of school-community partnerships. This analytical tool can be used to better understand the similarities and differences between family and interagency collaboration, full-service schools, full-service community schools, and a community development model. That being said, most approaches share the purpose of responding to community needs as an equity strategy (i.e., addressing education inequality more systemically, Task Force on Next Generation Community Schools, 2021).

For example, Community schools in Saskatchewan, Canada (Salm et al., 2016); in Ghent, Belgium (Blaton and Van Avermaet, 2016), and in the northern territory of Australia (Kamara, 2017) share a comprehensive approach yet vary in response to particular contexts and community needs (FitzGerald and Quiñones, 2018a, p. 1). For this reason, the *Coalition for Community Schools* (2021) notes that a community school is often referred to as "the hub of its neighborhood; it is a place and space that unites families, educators, and community partners to promote equity and excellence for children and families; it is an approach that strengthens families and communities" (p. 1).

Although implementation of the community school model vary-in-context, scholar-practitioners recommend the adoption of *four core pillars* to increase fidelity within the "scaling up" process (Task Force on Next Generation Community Schools, 2021, p. 24). According to Maier, Daniel, and Oakes from The Learning Policy Institute (2017), four features characterizing community schools that effectively support conditions for teaching and learning include: (a) integrated student supports; (b) expanded learning time and opportunities; (c) active family and community engagement; and (d) collaborative leadership and practice. Furthermore, effective community schools "fund full-time community school coordinators/directors to ensure quality implementation and increase return on investment by leveraging civic engagement and community resources toward desired outcomes" (*Coalition for Community Schools*, 2021, p. 13). It is important to note that community school coordinators, in tandem with school principals and community partners, lead efforts to integrate family-centered services and to expand learning opportunities for students (Bloodworth and Horner, 2019; *Coalition for Community Schools*, 2021; Mayger and Hochbein, 2019; FitzGerald and Quiñones, 2018b; Sanders et al., 2019).

In so doing, the community school model accounts for a promising strategy that rejects deficit views of families and calls upon families and community partners to implement educational practices grounded in shared goals, civic engagement, and systemic efforts (Fehrer and Leos-Urbel, 2016; Green, 2018; Heers et al., 2016 (McWayne et al., 2021; Min et al., 2017)).

Conclusion

As we close this scholarly journey through the intersection of leadership and family factors, we re-emphasize the importance of school leaders who co-design models of systemic engagement with multiple stakeholders. Family engagement is not a program, but rather a core element of effective and equitable educational practices. Historical views accentuate the evolution of parental involvement in a Western context and lead to broader discussions about the role and responsibilities of school leaders when working with families to support educational outcomes. Standards for school leadership signal the importance of families and communities in a local and global context, and the community school model illustrates a comprehensive place-based reform strategy that centers asset-based family engagement.

References

- Alameda-Lawson, T., Lawson, M.A., 2019. Ecologies of collective parent engagement in urban education. *Urban Educ.* 54, 1085–1120. <https://doi.org/10.1177/0042085916636654>.
- Alameda-Lawson, T., 2014. A pilot study of collective parent engagement and children's academic achievement. *Child. Sch.* 36, 199–209.
- Alameda-Lawson, T., 2016. A university-assisted, place-based model for enhancing students' peer, family, and community ecologies. *Educ. Sci.* 6, 1–15. <https://doi.org/10.3390/educsci6020016>.
- Alberta Teachers' Association and Canadian Association of Principals, 2014. The Future of the Principalship in Canada: A National Research Study. Alberta Teachers' Association.
- Anderson-Butcher, D., Ashton, D., 2004. Innovative models of collaboration to serve children, youth, families, and communities. *Child. Sch.* 26, 39–53.
- Anderson-Butcher, D., Paluta, L., Sterling, K., Anderson, C., 2018. Ensuring healthy development through community schools: a case study. *Child. Sch.* 40, 7–16.
- Antony-Newman, M., 2019a. Parental involvement of immigrant parents: a meta-synthesis. *Educ. Rev.* 71, 362–381. <https://doi.org/10.1080/00131911.2017.1423278>.
- Antony-Newman, M., 2019b. Parental involvement policies in Ontario: a critical analysis. *Sch. Community J.* 29, 143–170. <http://www.schoolcommunitynetwork.org/SCJ.aspx>.
- Auerbach, S., 2010. Beyond coffee with the principal: toward leadership for authentic school-family partnerships. *J. Sch. Leader.* 20, 728–757.
- Auerbach, S., 2012. Conceptualizing leadership for authentic partnerships. In: Auerbach, S. (Ed.), *School Leadership for Authentic Family and Community Partnerships: Research Perspectives for Transforming Practice*. Routledge, New York, NY, pp. 29–51.
- Bang, M., Montañón Nolan, C., McDaid-Morgan, N., 2018. Indigenous family engagement: strong families, strong nations. In: *Handbook of Indigenous Education*, pp. 1–22.
- Baquezano-López, P., Alexander, R.A., Hernández, S.J., 2013. Equity issues in parental and community involvement in schools: what teacher educators need to know. *Rev. Res. Educ.* 37, 149–182. <https://doi.org/10.3102/0091732X12459718>.
- Battiste, M., 2004. Unfolding the lessons of colonization. In: Sugars, C. (Ed.), *Unhomely States: Theorizing English-Canadian Postcolonialism*, pp. 209–220.
- Battiste, M., 2017. *Decolonizing Education: Nourishing the Learning Spirit*. University of British Columbia Press.
- Bejarano, C.L., 2002. Las super madres de Latino America: transforming motherhood by challenging violence in Mexico, Argentina, and El Salvador. *Front. J. Women Stud.* 23, 126–150.
- Berkovich, I., Bogler, R., 2020. The relationship between school leadership standards and school administration imperatives: an international perspective. *Sch. Leader. Manag.* 40, 321–334.
- Bishop, M., 2021. A rationale for the urgency of Indigenous education sovereignty: enough's enough. *Aust. Educ. Res.* 48, 419–432.
- Blaton, L., Avermaet, P.V., 2016. Community schools in Ghent: strengthening neighbourhoods in Belgium (Flanders). In: *Developing Community Schools, Community Learning Centers, Extended-service Schools and Multi-service Schools*. Springer, Cham, pp. 173–204.
- Bloodworth, M.R., Horner, A.C., 2019. Return on Investment of a Community School Coordinator: A Case Study. Apex, Albuquerque, NM.
- Briscoe, P., Pollock, K., 2017. Principals' perceptions of difference and diversity in their student bodies. *Canadian Association of Principals Journal* 10–14.
- Carrasco, A., Fromm, G., Gunter, H., 2021. Critical perspectives in and approaches to educational leadership in Chile. In: *Understanding Educational Leadership: Critical Perspectives and Approaches*. Bloomsbury Publishing PLC, pp. 121–135.
- Centre of Study for Policies and Practices in Education, 2013. *Learning Standards, Teaching Standards and Standards for School Principals: A Comparative Study*. OECD.
- Coalition for Community Schools, 2021. *Community Schools: Supporting One Another as a Blueprint for the Future*. Institute for Educational Leadership, Washington, DC.
- Comtassel, J., Hardbarger, T., 2019. Educate to perpetuate: land-based pedagogies and community resurgence. *Int. Rev. Educ.* 65, 87–116.
- Davies, D., 2000. Interview With Author. Institute of Responsive Education.
- Davis, B.W., Gooden, M.A., Micheaux, D.J., 2015. Color-blind leadership: a critical race theory analysis of the ISLLC and ELCC standards. *Educ. Adm. Q.* 51, 335–371.
- DeMatthews, D.E., Edwards Jr., D.B., Rincones, R., 2016. Social justice leadership and family engagement: a successful case from Ciudad Juárez, Mexico. *Educ. Adm. Q.* 52, 754–792. <https://doi.org/10.1177/0013161X16664006>.
- Dryfoos, J.G., Quinn, J., Barkin, C. (Eds.), 2005. *Community Schools in Action: Lessons From a Decade of Practice*. Oxford University Press.
- Dryfoos, J.G., 1994. *Full-Service Schools: A Revolution in Health and Social Services for Children, Youth, and Families*. Jossey-Bass.
- Epstein, J.L., Dauber, S.L., 1991. School programs and teacher practices of parent involvement in inner-city elementary and middle schools. *Elem. Sch. J.* 91, 289–305.
- Epstein, J.L., Sanders, M.G., Sheldon, S.B., Simon, B.S., Salinas, K.C., Jansorn, N.R., et al., 2018. *School, Family, and Community Partnerships: Your Handbook for Action*. Corwin Press.
- Epstein, J.L., 1987. What Principals Should Know About Parent Involvement. *Principal*, pp. 6–9.
- Fehr, K., Leos-Urbel, J., 2016. "We're one team": examining community school implementation strategies in Oakland. *Educ. Sci.* 26, 1–24. <https://doi.org/10.3390/educsci6030026>.
- Firpo, C., Jacobs, M., 2018. Taking children, ruling colonies: child removal and colonial subjugation in Australia, Canada, French Indochina, and the United States, 1870–1950s. *J. World Hist.* 29, 529–562.
- FitzGerald, A.M., Quiñones, S., 2018a. Working in and with community: leading for partnerships in a community school. *Leader. Pol. Sch.* 18, 511–532. <https://doi.org/10.1080/15700763.2018.1453938>.
- FitzGerald, A.M., Quiñones, S., 2018b. The community school coordinator: leader and professional capital builder. *J. Prof. Cap. Community* 3, 272–286. <https://doi.org/10.1108/JPC-02-2018-0008>.
- Flessa, J., Bramwell, D., Fernandez, M., Weinstein, J., 2018. School leadership in Latin America 2000–2016. *Educ. Manag. Adm. Leader* 46, 182–206.
- Galindo, C., Sanders, M.G., 2019a. Achieving equity in education through full-service community schools. In: Sheldon, S., Turner-Vorbeck, T. (Eds.), *The Wiley Handbook of Family, School, and Community Relationships in Education*. Wiley, pp. 511–530.
- Galindo, C., Sanders, M.G., 2019b. Final thoughts: working towards the sustainability of full-service community schools. *J. Educ. Stud. Placed Risk* 24, 1–6. <https://doi.org/10.1080/10824669.2019.1615959>.
- Galindo, C., Sanders, M., Abel, Y., 2017. Transforming educational experiences in low-income communities: a qualitative case study of social capital in a full-service community school. *Am. Educ. Res. J.* 54, 140S–163S.
- García, J., 2014. Re-examining indigenous conceptualizations of family and community involvement. *J. Family Divers. Educ.* 1, 58–74. <https://doi.org/10.53956/jfde.2014.24>.
- Gewirtz, S., Cribb, A., 2002. Plural conceptions of social justice: implications for policy sociology. *J. Educ. Pol.* 17, 499–509.
- Green, T.L., 2018. School as community, community as school: examining principal leadership for urban school reform and community development. *Educ. Urban Soc.* 50, 111–135. <https://doi.org/10.1177/0013124516683997>.
- Gu, Q., Day, C., Walker, A., Leithwood, K., 2018. How successful secondary school principals enact policy. *Leader. Pol. Sch.* 17, 327–331.
- Gutiérrez, K.D., Penuel, W.R., 2014. Relevance to practice as a criterion for rigor. *Educ. Res.* 43, 19–23.
- Heers, M., VanKlaverin, C., Groot, W., van den Brink, H., 2016. Community schools: what we know and what we need to know. *Rev. Educ. Res.* 86, 1016–1051.
- Hernández, A.A., Grineski, S.E., 2012. Disrupted by violence: children's well-being and families' economic, social, and cultural capital in Ciudad Juárez, Mexico. *Rev. Panam. Salud Públ.* 31, 373–379.
- Hernandez, F., Murakami, E.T., Cerecer, P.Q., 2014. A Latina principal leading for social justice: influences of racial and gender identity. *J. Sch. Leader.* 24, 568–598.
- Heyman, J.M., Campbell, H., 2004. Recent research on the US-Mexico border. *Lat. Am. Res. Rev.* 39, 205–220.
- Hiatt-Michael, D.B., 1994/2000. Schools: an historical perspective 1642–2000. *Sch. Community J.* 4, 247–258.

- Hill, S., 2003. Metaphoric enrichment and material poverty: the making of "colonias." In: Vila, P. (Ed.), *Ethnography at the Border*. University of Minnesota Press, Minneapolis, pp. 141–165.
- Hong, S., 2011. *A Cord of Three Strands: A New Approach to Parent Engagement in Schools*. Harvard Education Press.
- Hopson, R., 2014. Why are studies of neighborhoods and communities central to educational policy and reform? *Urban Educ.* 49, 992–995.
- Horvath, R.J., 1972. A definition of colonialism. *Curr. Anthropol.* 13, 45–57.
- Ishimaru, A., 2014. When new relationships meet old narratives: the journey towards improving parent-school relations in a district-community organizing collaboration. *Teach. Coll. Rec.* 116, 1–49.
- Ishimaru, A.M., 2019. From family engagement to equitable collaboration. *Educ. Pol.* 33, 350–385. <https://doi.org/10.1177/0895904817691841>.
- Ishimaru, A.M., 2020. *Just Schools: Building Equitable Collaborations With Families and Communities*. Teachers College Press, New York.
- Ishimaru, A.M., Galloway, M.K., 2014. Beyond individual effectiveness: conceptualizing organizational leadership for equity. *Leadersh. Policy Sch.* 13 (1), 93–146.
- Jacobson, R., Blank, M.J., 2015a. Expanding the learning day: an essential component of the community schools strategy. *N. Dir. Youth Dev.* 131, 55–67.
- Jacobson, R., Blank, M.J., 2015b. A framework for more and better learning through community school partnerships. *Inst. Educ. Leader.* 1–36.
- Jenkins, D., Duffy, M., 2016. *Community Schools in Practice: Research on Implementation and Impact*. Clearinghouse for Educational Research, Research for Action, Pennsylvania.
- Jeynes, W.H., 2018. A practical model for school leaders to encourage parental involvement and parental engagement. *Sch. Leader. Manag.* 38, 147–163.
- Jeynes, W., 2021. Parent involvement for urban students and youth of color. In: *Handbook of Urban Education*. Routledge, New York, pp. 418–433.
- Jung, S.B., Sheldon, S., 2020. Connecting dimensions of school leadership for partnerships with school and teacher practices of family engagement. *Sch. Community J.* 30, 9–32. <http://www.schoolcommunitynetwork.org/SCJ.aspx>.
- Kamara, M., 2017. Remote and invisible: the voices of female indigenous educational leaders in Northern Territory remote community schools in Australia. *J. Educ. Adm. Hist.* 49 (2), 128–143.
- Lawson, M.A., Alameda-Lawson, T., Richards, K.A.R., 2016. A university-assisted, place-based model for enhancing students' peer, family, and community ecologies. *Educ. Sci.* 6 (2), 1–16.
- Leithwood, K., Sun, J., Schumacker, R., 2020. How school leadership influences student learning: a test of "The Four Paths Model". *Educ. Adm. Q.* 56, 570–599.
- Leithwood, K., 2012. Ontario Leadership Framework With a Discussion of the Leadership Foundations. Institute for Education Leadership, OISE, Ottawa, Ontario, Canada.
- Leithwood, K., 2021. A review of evidence about equitable school leadership. *Educ. Sci.* 11, 1–49.
- Leithwood, K., (in press). Building productive relationships with families and communities: a priority for leaders to improve equity in their schools. In: Jeynes, W. (Ed.), *Parental Involvement: Cultivating Relationships*. New York: Routledge.
- Ma, X., Shen, J., Krenn, H.Y., Hu, S., Yuan, J., 2016. A meta-analysis of the relationship between learning outcomes and parental involvement during early childhood education and early elementary education. *Educ. Psychol. Rev.* 28, 771–801. <https://doi.org/10.1007/s10648-015-9351-1>.
- Mapp, K.L., Bergman, E., 2021. *Embracing a New Normal: Toward a More Liberatory Approach to Family Engagement*. Carnegie Corporation of New York, NY.
- Mayger, L.K., Hochbein, C.D., 2019. Spanning boundaries and balancing tensions: a systems perspective on community school coordinators. *Sch. Community J.* 29, 225–254.
- McWayne, C., Hyun, S., Diez, V., Mistry, J., 2021. "We feel connected... and like we belong": a parent-led, staff-supported model of family engagement in early childhood. *Early Child. Educ. J.* 1–13. <https://doi.org/10.1007/s10643-021-01160-x>.
- Min, M., Anderson, J.A., Chen, M., 2017. What do we know about full-service community schools? Integrative research review with NVivo. *Sch. Community J.* 27, 29–54.
- Ministerio de Educación de Chile, 2015. *The School Management and Leadership Framework [Marco para la Buena Dirección y Liderazgo Educativo]*. Ministerio de Educación, Chile.
- Mleccko, A., Kingston, A., 2013. The impact of school leadership on parental engagement: a study of inclusion and cohesion. *Inter. Res. Educ.* 1 (1), 129–148.
- Moorosi, P., 2020. Colour-blind educational leadership policy: a critical race theory analysis of school principalship standards in South Africa. *Educ. Manag. Adm. Leader* 174.
- Murphy, J., Louis, K.S., Smylie, M., 2017. Positive school leadership: how the Professional Standards for Educational Leaders can be brought to life. *Phi Delta Kappan* 99, 21–24.
- National Policy Board for Educational Administration, 2015. *Professional standards for educational leaders 2015*. Author, Reston, VA.
- Oakes, J., Maier, A., Daniel, J., 2017. *Community Schools: An Evidence-Based Strategy for Equitable School Improvement*. Learning Policy Institute.
- Pulliam, J.D., 1987. *History of Education in America*, fourth ed. Merrill, Columbus.
- Rivas, A., Sanchez, B., 2020. Race to the classroom: the governance turn in Latin American education: the emerging era of accountability, control and prescribed curriculum. *Compare* 1–19.
- Riveros, A., Newton, P., Burgess, D., 2016. Leadership standards and the discursive repositioning of leadership, leaders and non-leaders: a critical examination. In: *Questioning Leadership*. Routledge, UK, pp. 165–177.
- Salm, T., Caswell, E., Storey, S.G., Nunn, A., 2016. Enhancing and extending full service community schools in Saskatchewan, Canada: educators becoming part of the hub. In: *Developing community schools, community learning centers, extended-service schools and multi-service schools*. Springer, Cham, pp. 149–171.
- Sanders, M.G., Galindo, C., DeTablan, D., 2019. Leadership for collaboration: exploring how community school coordinators advance goals of full-service community schools. *Child. Sch.* 41, 89–99.
- Sanders, M.G., 2006. *Building Community Partnerships: Collaboration for Student Success*. Skyhorse, NY.
- Sanders, M.G., 2016. Leadership, partnerships, and organizational development: exploring components of effectiveness in three full-service community schools, school effectiveness and school improvement. *Int. Res. J. Pol. Pract.* 27, 157–177.
- Sanders, M.G., 2018. Crossing boundaries: a qualitative exploration of relational leadership in three full-service community schools. *Teach. Coll. Rec.* 120, 1–36.
- Sisson, J.H., Shin, A.M., Whittington, V., 2021. Re-imagining family engagement as a two-way street. *Aust. Educ. Res.* 1–18. <https://doi.org/10.1007/s13384-020-00422-8>.
- Sisto, V., 2018. Managerialismo, autoritarismo y la lucha por el alma de la gestión: el caso de las últimas reformas en políticas de dirección escolar en Chile. *Educação e Contemporaneidade* 2, 141–156.
- Smith, L.T., Tuck, E., Yang, K.W. (Eds.), 2018. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, UK.
- Smith, T.E., Reinke, W.M., Herman, K.C., Sebastian, J., 2021. Exploring the link between principal leadership and family engagement across elementary and middle school. *J. Sch. Psychol.* 84, 49–62. <https://doi.org/10.1016/j.jsp.2020.12.006>.
- Spring, J., 1986. *The American School 1642–1985*. Longman, New York.
- Stefanski, A., Valli, L., Jacobson, R., 2016. Beyond involvement and engagement: the role of the family in school-community partnerships. *Sch. Community J.* 26, 135–160.
- Tyack, D.B., 1974. *The One Best System: A History of American Urban Education*. Harvard University Press, Cambridge.
- Valli, L., Stefanski, A., Jacobson, R., 2014. Leadership in school-community partnerships. *Soc. Behav. Sci.* 141, 110–114.
- Valli, L., 2016. Typologizing school-community partnerships: a framework for analysis and action. *Urban Educ.* 5, 719–747. <http://doi:10.1177/004205914549366>.
- Wayne, A.J., Youngs, P., 2003. Teacher characteristics and student achievement gains: a review. *Rev. Educ. Res.* 73, 89–122.
- Wei, W., 2017. Education policy borrowing: professional standards for school leaders in China. *Chin. Educ. Soc.* 50, 181–202. <https://doi.org/10.1080/10611932.2017.1331012>.
- Weinstein, J., Hernández, M., 2016. Birth pains: emerging school leadership policies in eight school systems of Latin America. *Int. J. Leader. Educ.* 19, 241–263.
- Weinstein, J., Muñoz, G., Hernández, M., 2014. *El liderazgo escolar en América Latina y el Caribe. Un estado del arte en base a ocho sistemas escolares de la región*. Orealc-Unesco, Santiago.
- Weiss, H.B., Lopez, M.E., Caspe, M., 2018. *Joining Together to Create a Bold Vision for Next Generation Family Engagement: Engaging Families to Transform Education*. Global Family Research Project.
- Young, M.D., Perrone, F., 2016. How are standards used, by whom, and to what end? *J. Res. Leader. Educ.* 11, 3–11. <https://doi.org/10.1177/1942775116647511>.

Collaboration as a strategy for promoting equity in education

Mel Ainscow^{a,b}, ^a University of Glasgow, Glasgow, United Kingdom; and ^b University of Manchester, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

A global agenda	281
A different way	282
Inquiry-based developments	282
A family of approaches	283
A framework for analysis	284
Making sense of the process	286
Anomalies and organizational learning	287
Social learning	288
The nature of practice	289
Final thoughts	289
Acknowledgments	290
References	290

Many educators take the view that working together with others is a major means of achieving improvements in schools. The danger is that this belief can distract attention from the complexities involved in achieving effective forms of collaboration. There is a related risk that policy makers, seeing collaboration as the silver bullet of educational reform, might decide to mandate it as a means of achieving success. In so doing, they may overlook how other policies make this difficult.

With these concerns in mind, in this paper I reflect on many years of working with colleagues on a programme of development and research projects that have been based around collaborative ways of working (see summaries in: Ainscow, 2015, 2020). In so doing, I point to the many possibilities that this approach offers whilst, at the same time, examining the barriers that can obstruct progress. This analysis leads me to identify the conditions that are needed in order to make collaboration work, focusing in particular on what I see as the major challenge facing education systems around the world, that of achieving inclusion and equity. In this context, my focus is on factors at all levels of an education system, from the office of the minister, to the district level and through to what teachers do in their classrooms.

A global agenda

Since 1990, international impetus regarding inclusion in schools has been provided by the United Nation's Education for All (EFA) movement, which has worked to make basic education available to all learners. However, despite the progress that has been made over that period, recent UNESCO data show that progress has slowed in recent years, leaving some 258 million children, adolescents and youth still not in school, and more than 617 million children and adolescents not achieving minimum proficiency levels in reading and mathematics (UIS, 2019). It is also important to recognize that inequality in education has increased, with the poorest and most disadvantaged shouldering the heaviest burden.

Whilst this situation is most acute in the developing world, there are similar concerns in many wealthier countries, as noted by the OECD (2012), which reports that across its member countries, almost one of five students does not reach a basic minimum level of skills to function in today's societies. It is also noted that students from low socio-economic background are twice as likely to be low performers, implying that personal or social circumstances are obstacles to achieving their educational potential.

In responding to these challenges there is growing interest internationally in the use of strategies that place an emphasis on the power of market forces (Lubienski, 2003). In particular, some national education policies are encouraging schools to become autonomous; for example, in parts of Australia, the independent public schools; the academies in England; charter schools in the USA; the voucher reforms in Chile; and free schools in Sweden. Alongside this emphasis on school autonomy is a focus on parental choice. This usually takes place within education systems where high-stakes testing systems are intended to inform decision making, whilst at the same time driving improvement efforts (Au, 2009). In addition, narrowly defined measures of effectiveness are used for purposes of accountability (Schildkamp et al., 2012).

This global policy trend is a matter of considerable debate and there are varied views as to the extent to which it is leading to the desired outcomes. In particular, there is a concern that the development of education systems based on autonomy, coupled with high-stakes accountability and increased competition between schools, will further disadvantage learners from low income and minority families (Salokangas and Ainscow, 2017).

Whilst such developments have the potential to open up possibilities to inject new energy into the improvement of education systems, there is growing evidence from a range of countries that they are leading to increased segregation that further disadvantage learners from economically poorer backgrounds (Pickett and Vanderbloemen, 2015). So, for example, talking about the development of charter schools in the USA, Kahlenberg and Potter (2014) suggest they have led to increased segregation in school systems across the country and similar patterns are reported from Sweden (Wiborg, 2010). In relation to Chile, Zancajo (2019) explains how the context of a competitive environment has led to processes of exclusion that affect students from low socioeconomic backgrounds, foster school segregation and increase educational inequalities. Meanwhile, in New Zealand, a recent government concluded that a key reason for the country's poor equity and achievement outcomes is that schools have predominantly operated as autonomous, self-managing entities, loosely connected to each other, and with a distant relationship with the center (Ministry of Education, 2019). This autonomy, it is argued, has left schools to operate largely on their own and without sufficient support.

Meanwhile, in my own country, England, we have reported how the creation of academy schools has reduced the level of funding that is available elsewhere throughout the education system and, in particular, diverted resources away from local authorities and the schools that remain within their control (Ainscow et al., 2016). Moreover, the manipulation of admissions suggests that these initiatives are contributing more directly to social segregation by effectively selecting "easy-to-teach students" into favorably funded schools. It seems, that, as with other reforms badged as movements towards greater school autonomy, any gain in this direction comes at a price.

A different way

There are, however, countries that have made progress by adopting a very different approach, one that involves an emphasis on collaboration and networking in order to combine quality with equity (OECD, 2012). In these contexts, the vast majority of students have the opportunity to attain high level skills, regardless of their own personal and socio-economic circumstances.

Encouragement for this view is provided by the "Report Card" prepared for Unicef by the Innocenti Center (2018). Focusing on high and middle-income countries, it concludes:

Tackling educational inequality does not mean sacrificing high standards. Countries with higher average achievement tend to have lower levels of inequality... Bringing the worst performing students up does not mean pulling the best-performing students down (p. 3).

Warning against the temptation to copy what more successful countries do well, the authors of the report argue that, whilst there are undoubted lessons that can be learned from these countries, they must be replicated with care. There are, it argues, many sources of inequality in education related to political, economic, social, cultural and institutional contexts, and these vary across countries.

In thinking about what is needed to use collaboration as a strategy for promoting equity within education system, there is also a need to rethink the way systems are managed. In particular, we need to reflect on the recommendations of the OECD in regard to developments in Scotland:

In our view there is need now for a bold approach that moves beyond system management... new dynamic and energy to be generated nearer the learning level... We call for a strengthened "middle" operating through networks... within and across local authorities... to create coherent and cohesive cultures of system-wide improvement.

OECD (2015, p. 15).

This suggests that some form of local coordination is needed in order to encourage collaborative activity.

Bringing the international policy context up to date, the year 2016 was particularly significant. It saw the publication of the Education 2030 Framework for Action (UNESCO, 2017). This emphasizes inclusion and equity as laying the foundations for quality education. It also stresses the need to address all forms of exclusion and marginalization, disparities and inequalities in access, participation, and learning processes and outcomes.

Taking account of this international policy agenda, in what follows I argue that effective change requires processes of social learning and knowledge creation within settings that stretch from multiple classrooms to the committee structures of senior policy makers. In particular, I reflect on evidence from a series of studies, carried out with my colleagues over the last three decades in a variety of national contexts, in order to consider how such approaches can be operationalized.

Inquiry-based developments

The development of the approach I am recommending began with the "Improving the Quality of Education for All" (IQEA) project in the early 1990s. Initially this involved a small group of researchers at the University of Cambridge working with schools in and around London (Ainscow and Hopkins, 1992). Subsequently, IQEA led to developments in other parts of the world (see Ainscow,

1999; Clarke et al., 2006; Hopkins, 2007; Hopkins et al., 1994; Hopkins et al., 1996; West and Ainscow, 2010 for more detailed accounts of some of these projects). All of these activities involved teams of researchers working in partnership with networks of schools to identify ways in which the learning of all members of a school community—students, parents and staff—could be enhanced.

Work with schools in the IQEA projects was based upon a contract that attempted to define the parameters for our involvement, and the obligations those involved owed to one another. It emphasized that all staff be consulted; that an in-school team of coordinators be appointed to carry the work forward; that a critical mass of staff were to be actively involved; and that sufficient time would be made available for necessary classroom and staff development activities. Meanwhile, we committed ourselves to supporting the school's developments, usually in the first place for 1 year. Often the arrangement continued, however, and, in some instances, we were involved for periods as long as 7 years. We provided training for the school coordinators, made regular school visits and contributed to school-based staff development activities. In addition, we attempted to work with the schools in recording and analyzing their experiences in a way that also provided data relevant to our own on-going research agendas.

As a result of such engagements with schools involved in the IQEA project, we evolved a style of collaboration that was to influence later initiatives, which we referred to as “working with, rather than working on” (Ainscow and Southworth, 1996). This phrase attempted to sum up an approach that deliberately allowed each project school considerable autonomy to determine its own priorities for development and, indeed, its methods for achieving these priorities. In attempting to work in this way, we found ourselves confronted with staggering complexity, and by a bewildering array of policy and strategy options. It was our belief that only through a regular engagement with these complexities could a greater understanding of school change be achieved.

These ideas were taken forward and developed through a series of further studies:

- *Understanding and Developing Inclusive Practices in Schools*. This collaborative action research study occurred between 2000 and 2004, and involved a network of 25 urban schools in three English local education authorities and three partner universities. Within the network schools, we saw how the use of evidence to study teaching can help to foster the development of more inclusive forms of teaching (see: Ainscow et al., 2004, 2006a; Dyson et al., 2003; Howes et al., 2004, 2005).
- *The Stockborough Equity Research Network*. Between 2006 and 2011 we had a chance to explore these ideas in more detail through our involvement in another network of schools England (see Ainscow et al., 2012a,b). This initiative was located in a local authority, which we called “Stockborough”, characterized by socio-economic disadvantage, and social and ethnic segregation. It involved staff inquiry groups in participating schools, usually consisting of five or six members representing different perspectives within their school communities.
- *Ethical Leadership: A collaborative investigation of equity-driven evidence-based school reform*. This study involved a research network of schools, spread across the geographically spread State of Queensland, Australia, and a team of eight university researchers who, together, explored how ethical leadership could promote ways of interpreting and using various forms of evidence to promote learning and equity (see: Harris et al., 2017, 2019; Spina et al., 2019).
- *Inclusive Inquiry*. Carried out between 2012 and 2020, this programme of research focused more specifically on how students can contribute to the development of inclusive teaching and learning. It involved the development of a strategy in which children becoming researchers who learn how to use research techniques to gather the views of their classmates, as well as observing lessons (see: Ainscow and Messiou, 2017; Messiou et al., 2016; Messiou and Ainscow, 2015, 2020).

All of these studies involved teams of university researchers in supporting, recording and analyzing collaborative action research as it occurred in project schools. As a result, we came to define this as a process of knowledge-generation that occurs when researcher and practitioner knowledge meet in particular sites and is aimed at producing new knowledge about ways in which broad values might better be realized in future practice.

This approach has significant implications for the relationship between researchers and practitioners. It requires new forms of partnership, in the way that is outlined helpfully by Hiebert et al. (2002) who suggest that fruitful forms of collaboration require a reorientation of values and goals amongst both groups. So, they argue, teachers need to move away from the dominant view that teaching is a “personal and private activity”. Rather, teachers have to adopt the “more risky view” that it is an activity that can be continuously improved, provided it is made public and examined openly. At the same time, they argue that researchers must stop undervaluing the knowledge teachers acquire in their own classrooms. In this way researchers are able to recognize the potential of personal knowledge as it becomes transformed into professional knowledge.

A family of approaches

These studies adopted different terms to describe their efforts, such as collaborative inquiry, practitioner action research, action learning and lesson study. Recently Yurkofsky (2020) have labeled approaches such as these collectively as “continuous improvement” methods, noting that they share four characteristics:

1. Grounding improvement efforts in local problems or needs;
2. Empowering practitioners to take an active role in research and improvement;
3. Engaging in iteration, which involves a cyclical process of action, assessment, reflection, and adjustment; and
4. Striving to encourage change across schools and systems, not just individual classrooms.

They can be seen as being part of a “family” of approaches within the overall tradition of action research (Reason and Bradbury, 2001) that developed out of the action research tradition of Kurt Lewin (1946) and the work of many other academics (e.g., Cochran-Smith and Lytle, 2009; Elliot, 1991; Kemmis and McTaggart, 1988; Schön, 1983; Stenhouse, 1975). They emphasize an engagement in inquiry to inform and improve practice, and the intentional combining of knowing and doing for achieving positive social change (Kemmis, 2010).

Since the focus of the inquiry is usually defined by the engagement of the participants, careful consideration is required in determining who is included, and how and who will speak for whom, and who sets the research agenda. Therefore, negotiating power issues and the relationships between collaborating practitioners, stakeholders and academics “requires ethical probity where each party recognizes, understands and respects mutual responsibilities” (Campbell and Groundwater-Smith, 2007, p. 2).

The preoccupation with equity means that it also requires a particular concern to giving voice to those who may be powerless or unheard in the decision-making processes (Groundwater-Smith, 2011). Moreover, actively engaging participants in inquiry “problematizes the question of who is researcher and who is researched, raising issues around anonymity, the ‘ownership’ of findings and dissemination” (Locke et al., 2013, p. 107).

Related considerations concern decision-making around how outcomes will be evaluated in order to determine whether the inquiry processes have led to positive change. Questions to be considered include:

- What forms of evidence will be used to determine the outcomes for those who have participated in the process and those who are the intended beneficiaries of the change efforts that were the focus of the project?
- How far did processes improve the social and professional practices in the contexts in which they were conducted?
- How will the outcomes from the inquiry be used?

Participatory and collaborative inquiry processes, although not set on achieving consensus, do provide opportunities for negotiation of multiple perspectives around shared agendas, rather than producing evidence from a single viewpoint or source. A helpful strategy in this respect is what Wasser and Bresler (1996) refer to as “group interpretive processes” as a means of analyzing and interpreting evidence. This involves an engagement with the different perspectives of practitioners, students and university researchers in ways that are intended to encourage critical reflection, collaborative learning and mutual critique. In these contexts, the varied theoretical perspectives of members of the research teams provide a valuable means of questioning taken-for-granted assumptions and helping practitioners to reconsider neglected possibilities for moving practice forward.

The issue of trustworthiness is a particular challenge to collaborative research of this sort, which involves such a high degree of participation among stakeholders. Commenting on this issue, Schon (1991) argues that without a serious effort to make clear what is meant by rigor, participatory research “becomes an open sesame to woolly-headedness, a never-never land where anything goes” (p. 10). He goes on to suggest that appropriate rigor in the study of practice should focus on validity (e.g., how do we know what we claim to know?) and utility (e.g., how useful is the research to practitioners?).

With this in mind, the studies I have described made use of three forms of triangulation, supporting observations and reports from a number of viewpoints. These involved comparing and contrasting evidence about the same actions and activities from different people (e.g., teachers, support staff and students); scrutinizing events from different angles by making use of a variety of methods for collecting information; and using “outsiders” as observers.

Influenced by the ideas of Karl Popper, Schon (1991) argues that the fundamental test for validity in participatory enquiry is through “competitive resistance to refutation”. This involves juxtaposing alternate plausible accounts of the phenomenon in question. Schon notes: “In the absence of an alternate hypothesis, one is likely to be overwhelmed by the obviousness of what one already knows” (p. 348).

Guided by this advice, members of our research teams usually discussed with their practitioner partners, and with one another, written accounts of the work carried out in the schools, including alternative explanations as to what lessons could be drawn from these experiences. In this way, the aim was to draw conclusions that were both valid and relevant.

A framework for analysis

The studies summarized so far were all guided by the principle of equity, which, following the lead of OECD (2012), is taken to involve notions of inclusion and fairness. Working with schools in different contexts over many years, those involved in the research became increasingly aware of the complexities this involves. This led to the development of an overall framework to guide the development of the collaborative action research, referred to as an “ecology of equity” (Ainscow et al, 2012a). This formulation involves three interlinked areas within which equity issues arise:

- *Within-schools*. These are issues that arise from school and teacher practices.
- *Between-schools*. These are issues that arise from the characteristics of local school systems.
- *Beyond-schools*. This far-reaching arena includes: the wider policy context within which schools operate; the family processes and resources which shape how children learn and develop; the interests and understandings of the professionals working in schools; and the demographics, economics, cultures and histories of the areas served by schools.

Drawing on the findings of the studies referred to above, in what follows possibilities for linking within-school, between-school and beyond-schools factors in order to develop effective collaborative improvement approaches are explored.

Within-school factors. My colleagues and I have argued that “schools know more than they use” (Ainscow et al., 2012a). In other words, schools tend to have overlooked expertise that can be mobilized to address the challenge of equity. This suggests that the starting point for strengthening the work of a school is with the sharing of existing practices through collaboration amongst staff, leading to experimentation with new practices that will reach out to all students. This approach draws on the work of the many scholars who have explored processes of collaborative inquiry as a strategy for professional learning and school improvement.

Much of what teachers do during the intensive encounters that occur is carried out at an automatic, intuitive level. Furthermore, there is little time to stop and think. This is why having the opportunity to see colleagues at work is so crucial to the success of attempts to develop practice. It is through shared experiences that colleagues can help one another to articulate what they currently do and define what they might like to do (Hiebert et al., 2002). It is also the means whereby space is created within which taken-for-granted assumptions about particular groups of learners and what they might achieve can be subjected to mutual critique, a factor that is crucial to the promotion of inclusive practices.

At the heart of the processes in schools where changes in practice do occur is the development of a common language with which colleagues can talk to one another and, indeed, to themselves, about detailed aspects of their practice (Huberman, 1993). Without such a language, teachers find it very difficult to experiment with new possibilities (Ainscow, 1999).

In a recent article, Lefstein et al. (2020) point to the importance of the conversations embedded in teachers’ day-to-day work, through which they learn from one another what it means to be a teacher and how to perform their duties. The authors explain how professional knowledge and skills, much of which are implicit, are learned on the job, through participation in work practices and informal interactions with colleagues. This means that professional knowledge is intimately related to the practices through which it is constructed and to which it is applied. Given the situated nature of knowledge, ideas constructed with colleagues at school are more likely to be used within a school, whereas ideas constructed within professional development workshops are only likely to be applied in the social practice of participating in such events.

An engagement with evidence of various kinds to study teaching within a school can help in generating school-based professional learning (Ainscow et al., 2004). This, in turn, can help to foster the development of practices that are more effective in engaging hard to reach learners. Specifically, it can create space for rethinking by interrupting existing discourses. The starting point for such processes is often with a consideration of statistical evidence regarding student progress. However, the need to dig deeper into factors that influence progress usually requires an engagement with qualitative forms of evidence. Particularly powerful techniques in this respect involve the use of mutual lesson observation, sometimes through video recordings, and evidence collected from students about teaching and learning arrangements within a school.

Under certain conditions such approaches provide “interruptions” that help to make the familiar unfamiliar in ways that stimulate self-questioning, creativity and action. In so doing they can sometimes lead to a reframing of perceived problems that, in turn, draws the teacher’s attention to overlooked possibilities for addressing barriers to participation and learning. In this way, differences amongst students, staff and schools can become a catalyst for improvement.

Here, the concern with the principle of equity means that there also has to be focus on the thinking that is *behind* actions and the impacts of such thinking on practices. In particular, there has to be a concern with the attitudes and assumptions that influence what teachers do, some of which may be unconscious (Sadker et al., 2009), and how these can be modified through dialogs with others, especially with learners themselves. All of this points to the importance of forms of leadership that encourages colleagues to challenge one another’s assumptions about particular students (Riehl, 2000).

Between-school factors. The approach outlined so far is based on the idea of those within schools collecting and engaging with various forms of evidence in order to stimulate moves to create more inclusive practices. The studies I have summarized provide encouraging evidence of the potential of this approach. However, they have also thrown light on the difficulties in putting such thinking into practice, particularly within policy contexts that put pressure on schools to compete. This points to some of the limitations of within-school strategies, suggesting that these should be complemented with efforts to encourage greater cooperation between schools, and between schools and their wider communities.

The projects have generated considerable evidence that school-to-school collaboration can strengthen improvement processes by adding to the range of expertise made available (see: Ainscow, 2010, 2012, 2015; Ainscow and Howes, 2007; Ainscow and West, 2006; Ainscow et al., 2004, 2006b; Muijs et al., 2010, 2011). Together with the work of others (e.g., Chapman and Hadfield, 2010; Fielding et al., 2005; Hill, 2008), these studies indicate that collaboration between schools has an enormous potential for fostering the capacity of education systems to respond to learner diversity. More specifically, they show how such partnerships can sometimes help to reduce the polarization of schools within an education system, to the particular benefit of those students who seem marginalized at the edges of the system, and whose performance and attitudes cause concern.

There is also evidence that when schools seek to develop more collaborative ways of working this can have an impact on how teachers perceive themselves and their work (Muijs et al., 2011). Specifically, comparisons of practices in different schools can lead teachers to view underachieving students in a new light. In this way, learners who cannot easily be educated within a school’s established routines come to be seen less as “having problems” but as challenging teachers to re-examine their practices in order to make them more responsive and flexible.

It is important to realize, however, that developing partnerships between schools is not a straightforward process. Too often they can lead to meetings without any significant action. For example, in the networks referred to earlier, schools were encouraged to visit one another in order to generate evidence regarding their shared focus on developing more inclusive practices. However, these visits were not always successful. This seemed to be particularly so when the host teachers interpreted the visits solely as opportunities for the visitors to learn. On these occasions, the hosts positioned themselves as teachers rather than learners. Typically, the visit then

consisted of a demonstration of various teaching strategies that had been judged to be successful, usually followed by a short question and answer session. On these occasions, those receiving the visit merely rehearsed what they already knew and responded to questions beyond the procedural as if they were challenges, rather than openings for debate. On the other hand, successful visits were usually characterized by a sense of mutual learning amongst hosts and visitors. It was noticeable, too, that the focus for these visits often took some time to identify and clarify. Indeed, the preliminary negotiations that took place were in themselves a key aspect of the process.

Further convincing evidence of the power of schools working together comes from the City Challenge programmes in London and Greater Manchester. The success of these initiatives has been widely reported, leading to extensive debates as to what were the key factors that led to their impact (e.g., [Barrs et al., 2014](#); [Claeys et al., 2014](#); [Greaves et al., 2014](#); [Hutchings et al., 2012](#); [Kidson and Norris, 2014](#)).

An important feature of City Challenge was the emphasis placed on contextual analysis as the starting point for developments. This led to the conclusion that plenty of good practice existed across the schools in the two city regions. Consequently, it was decided that collaboration and networking would be the key strategies for strengthening the overall capacity of the system to reach out to vulnerable groups of learners. More specifically this involved a series of inter-connected activities for “moving knowledge around” ([Ainscow, 2015](#)). Once again, an engagement with evidence proved to be important in making this happen.

The powerful impact of the collaborative strategies developed in City Challenge points to ways in which improvement processes used within individual schools can be deepened and, therefore, strengthened. This involves an emphasis on mutual critique, within schools and between schools, based on an engagement with shared data. This, in turn, requires strong collective commitment from senior school staff and a willingness to share responsibility for system reform.

Beyond-school factors. Research suggests that the development of education systems that are effective for all children will only happen when what happens outside as well as inside a school changes ([Kerr et al., 2014](#)). Indeed, there is encouraging evidence of what can happen when what schools do is aligned in a coherent strategy with the efforts of other community players—families, employers, community groups, universities and public services. This does not necessarily mean schools doing more, but it does imply partnerships beyond the school, where partners multiply the impacts of each other’s efforts. However, experience suggests that the success of such partnerships is dependent upon a common understanding of what they are trying to achieve and, once again, an engagement with various forms of evidence to stimulate collective effort.

This approach draws on the principles underpinning the highly acclaimed Harlem Children’s Zone in the USA ([Whitehurst and Croft, 2010](#)). These involve efforts to improve outcomes for children and young people in areas of disadvantage through an approach that they characterize as being “doubly holistic”. That is to say, they seek to develop coordinated efforts to tackle the factors that disadvantage children and enhance the factors which support them, across all aspects of their lives, and across their life spans, from conception through to adulthood. [Dobbie and Fryer \(2009\)](#) describe the Children’s Zone as “arguably the most ambitious social experiment to alleviate poverty of our time” (p. 1).

Another American initiative, StriveTogether, acts as a central backbone organization for sites using similar ideas across the country that are locally tailored ([Grossman et al., 2014](#)). These initiatives are guided by indicators that span young people’s lives “from cradle to career”, with progress determined using data at all stages. The challenges involved in applying this thinking have been explored in the Children’s Neighborhoods Scotland initiative ([Chapman and Ainscow, 2021](#)).

Developments such as these have implications for the various key stakeholders within education systems. In particular, teachers, especially those in senior positions, have to see themselves as having a wider responsibility for all children and young people, not just those that attend their own schools. They also have to develop patterns of working that enable them to have the flexibility to cooperate with other schools and their wider communities. It means, too, that those who administer area school systems have to adjust their priorities and ways of working in response to improvement efforts that are led from within schools.

Some form of local coordination is needed in order to encourage such forms of collaboration ([Armstrong and Ainscow, 2018](#)). Indeed, a recent report noted that four of the most successful national education systems—Singapore, Estonia, Finland, and Ontario—each has a coherent “middle tier”, regardless of their differing extents of school autonomy or devolution of decision-making ([Bubb et al., 2019](#)). In particular, they all have district level structures that offer a consistent view that, to maintain equity as well as excellence, there needs to be an authoritative co-ordinating influence, with local accountability.

Making sense of the process

Underlying these ways of using collaboration to promote equity within education systems is a common pattern. Most importantly, they involve an engagement with evidence collected by practitioners with support from university researchers. Usually, this begins with a consideration of an established set of practices that are largely taken for granted. An interruption occurs that problematizes these practices and provokes consideration of why current practice is the way it is and how it might be improved. This may then lead to actual changes in practice.

Given my focus on the possibilities of developing more equitable ways of working, this pattern begs two important questions:

- What is it that provokes the problematization of established practice?
- And why does this necessarily lead to more inclusive ways practices?

In addressing these questions, we have found it helpful to draw on a range of theoretical resources. One that has proved particularly powerful is the idea of “communities of practice”, as developed by Wenger (1998), focusing specifically on the way he sees learning as “a characteristic of practice”.

Wenger explains practice in terms of those things that individuals within a community do, drawing on available resources, to further a set of shared goals. This goes beyond how practitioners complete their tasks, to include, for example, how they make it through the day, commiserating about the pressures and constraints within which they have to operate. Practices are thus ways of negotiating meaning through social action, which underlines the importance of the conversations embedded in teachers’ day-to-day work, referred to earlier in this paper.

In explaining this process, Wenger argues that communities ‘reify’ their practices by producing concrete representations of them, such as tools, symbols, rules and documents (and even concepts and theories). However, these reifications have to be given meaning through a process of participation, which consists of the shared experiences and negotiations that result from social interaction within a purposive community.

Wenger offers some helpful guidelines for judging whether a particular social collective can be considered as a community of practice. Since such a community involves mutual engagement, a negotiated enterprise, and a repertoire of resources and practices, then its members should be expected to:

- Interact more intensively with, and know more about, others in the group than those outside the group;
- Hold their actions accountable (and be willing for others in the community to hold them accountable) more to the group’s joint enterprise than to some other enterprise;
- Be more able to evaluate the actions of other members of the group than the actions of those outside the group; and
- Draw on locally produced resources and artifacts to negotiate meaning, more so than those that are imported from outside the group.

By these criteria the staff teams in the studies reported earlier can, to varying degrees, be seen as communities of practice. Much of the evidence for this assertion rests on what was witnessed of the ongoing and informal interactions between groups of teachers. So, for example, amongst staff inquiry groups, hours of meetings, shared experiences and informal discussions over hurriedly taken lunches were observed. These sometimes involved the development of particular meanings of frequently used phrases, such as “raising standards”, “equity” and “inclusion”. These shared meanings help to define a teacher’s experience of being a teacher. In the same way it can be assumed that groups of colleagues doing similar work in another school have their own shared histories that give meaning to being a teacher in that particular context.

There is no reason to suppose that teachers are conscious of such processes for the most part, though occasionally they may be able to articulate their importance for developments in their schools. What was significant in some of the project schools, however, was not simply the high level of collaboration that was claimed by the teachers (such claims can, of course, be challenged), but the implication that “good practice” is defined through such collaborative processes. In these contexts, good practice is defined not by what researchers or policy-makers say, or others do elsewhere, but by what “we” think “works here”.

Once again, the notion of communities of practice is important here in that it views practice as being intimately bound up with the norms, values, beliefs and assumptions of a group of teachers, in a particular school context. The implication is that practices cannot be understood—or, more to the point, changed—without also understanding and changing those local patterns. What is needed is something that disturbs existing assumptions and provokes some sort of reformulation, both of practice and of the thinking on which it is based.

In the context of the collaborative research processes I have described, disturbances often come from an engagement with evidence generated within a school that helped to make the familiar unfamiliar. Sometimes, too, they arose from outside the school in the shape of a planned intervention—a report from the university group to the school inquiry team, or a visit to another school. However, the communities of practice idea suggests that meaning-making does not depend on such planned and structured events but is an ongoing process, as groups of teachers come to terms with new situations, new disturbances and new perspectives.

Anomalies and organizational learning

This relationship between practice and local meaning-making suggests that external agendas cannot simply be ‘imposed’ on communities of practice. Specifically, external proposals for change, however powerfully enforced, have to be endowed with meaning within local contexts before they can inform practice. This implies that schools (or, at least, the communities of practice within schools) may well negotiate local meanings for those agendas that are different from those of the formulators themselves or, indeed, of other schools.

The significance of communities of practice is usefully summed up by Wenger when he argues:

Communities of practice are not intrinsically beneficial or harmful.... Yet they are a force to be reckoned with, for better or for worse. As a locus of engagement in action, interpersonal relationships, shared knowledge, and negotiation of enterprises, such communities hold the key to real transformation - the kind that has real effect on people’s lives... The influence of other forces (e.g. the control of an institution or the authority of an individual) are no less important, but they are mediated by the communities in which their meanings are negotiated in practice.

Wenger (1998, p. 85).

There is, however, an important caveat here. Communities of practice “are not intrinsically beneficial or harmful” precisely because the values and assumptions to which they will subscribe are determined locally. In terms of the development of practices, there is nothing inherent in even the most dynamic community of practice which predisposes it towards generating more equitable outcomes. Indeed, a dynamic community that defends the status quo or moves rapidly in a non-inclusive direction is entirely conceivable. So, for example, Yurkofsky et al. (2020) argue:

Educators who believe in supporting equitable schools can still carry implicit biases that affect their practices, and teachers who aspire to improve their pedagogy may in practice have trouble giving up the belief that external factors (e.g., parental and neighbourhood influence) - as opposed to their own actions as teachers - are the primary determinants of students’ achievement (p. 415).

What, then, makes the difference between instances where collaborative meaning-making create a potential for the development of inclusive practices and those where such practices actually arise? For this, two sets of concepts from the literature on organizational development are helpful: Argyris and Schön’s (1996) idea of “single and double-loop learning”; and Skrtic’s (1991) distinction between bureaucracies and adhocracies, together with his notion of the recognition of “anomalies” as the catalyst for the transition from one to the other.

Argyris and Schön argue that organizations are capable of learning, but to different extents and, indeed, at different levels. What they refer to as single-loop learning takes the form of what in an educational context might be called the improvement of existing practices, but without any fundamental reconsideration of the assumptions upon which those practices are based. On the other hand, double-loop learning asks questions about the underlying aims of practice and about the implicit theories which underpin it.

To take an example, in the first year of the Stockborough Equity Research Network, described above, discussions involving the head teachers led to an agreement that equity was a central issue facing each of the partner schools. It soon became evident, however, that what this meant was different in each context, not least in respect to the groups of students who were seen as a cause for concern. As a result, it was decided that the work of the network would focus attention on those learners thought to be “missing out” within existing arrangements. This led to a concern amongst our university team that this might lead to narrowly focused efforts to “fix” students seen as being in some sense inadequate. However, collecting evidence about these groups eventually led to a process of double-loop learning that re-focused attention on an analysis of contextual factors that were acting as barriers to their participation and learning. In this way, most of the projects gradually became whole school improvement efforts that had the potential to benefit many students.

Skrtic also presupposes a fundamental distinction in the way organizations solve problems. He argues that bureaucratic organizations deal with such problems by creating different sub-units and specialisms to contain them, leaving practice elsewhere in the organization undisturbed. So, for example, a school may decide to establish a separate unit to deal with the problems of disruptive behavior, such that it avoids the need to examine ways in which its own practices may have helped to generate these problems. On the other hand, what Skrtic calls adhocratic organizations see such problems as an opportunity to rethink their existing practices in fundamental ways. Moreover, he argues that bureaucratic organizations can become adhocratic if enough of their members recognize “anomalies” in existing practice.

It seems, then, that faced with some form of disturbance—what I referred to earlier as an interruption—some schools will close the problem down and make largely technical responses. Others may open the problem up and use it as the basis for a critical interrogation, and reformulation of practice and the assumptions on which practice is based. However, as Skrtic argues, someone has to recognize a problem as an anomaly and convince others.

Although the notion of communities of practice would encourage an emphasis on the collaborative nature of such recognition, there is no doubt that individuals can also be powerful inhibitors or facilitators. Here, the leadership style and practices of senior staff are important. So, for example, Lambert et al. (2005) write about the importance of “constructivist leadership”, which can engage colleagues in what was characterized earlier in this chapter as shared meaning-making.

A further important factor appears to be access to evidence and alternative perspectives which lie outside current assumptions and constitute a basis on which they can be problematized. This is where the involvement of external research partners who provide critical interventions can be so important as those in schools come to terms with data they have collected. A similar process can occur through school-to-school collaboration.

Here, again, we are in the territory of communities of practice, where processes of meaning making in a particular community and in the local context of that community are at work. In other words, “what we see working here” is more important than what others tell us to do. However, in this case, established patterns of practice and meaning are disturbed sufficiently such that distinctly new patterns begin to emerge.

Social learning

This way of thinking is based on the idea that schools and their local communities have untapped potential to improve their capacity for improving the achievement of all of their students, not least those from economically poorer backgrounds. The

challenge therefore is to mobilize this potential. This reinforces the argument that school improvement is a social process that involves practitioners in learning from one another, from their students, and from others involved in the lives of the young people they teach.

In thinking about how this thinking might be used more widely it is essential to recognize that it does not offer a straight forward recipe that can be lifted and transferred from place to place. Rather, it defines an approach to improvement that uses processes of contextual analysis in order to create strategies that fit particular circumstances. This involves an engagement with various forms of evidence, leading to the development of locally determined improvement pathways. In this way, those involved probe beneath the surface of headline performance indicators to understand how local dynamics shape particular outcomes for students. In so doing, this helps to identify barriers to progress and resources that can inject pace into efforts to move things forward. What is distinctive in the approach is that it is mainly led from within schools, with senior school staff having a central role as “system leaders” (Hopkins, 2007). And, as argued earlier, this requires new thinking, practices and relationship across education systems.

It is predictable that such changes will lead to periods of organizational “turbulence” (Hopkins et al., 1994). The nature of this phenomenon will vary from place to place, but in general it arises as a result of the reactions of individuals to ideas and approaches that disrupt the status quo of their day-to-day lives. It is worth noting, however, that there is research evidence to suggest that without periods of turbulence, successful, long-lasting change is unlikely to occur (Fullan, 2007). In this sense turbulence can be seen as a useful indication that things are on the move.

The nature of practice

There is one further factor at work, which may be more fundamental. We are familiar with the notion that professionals operate in arenas characterized by complexity and uncertainty, such that their practice must constantly adjust to new and unique situations (Eraut, 1994; Schön, 1983/1991). As Eraut notes, this creates a significant problem, given the tendency of human actors to routinize their actions:

The development of routines is a natural process, essential for coping with the job and responsible for increased efficiency; but the combination of tacit knowledge and intuitive decision-making makes them difficult to monitor and to keep under critical control. As a result, routines tend to become progressively dysfunctional over time; not only do they fail to adjust to new circumstances but ‘shortcuts’ gradually intrude, some of which only help professionals to cope with pressure at the expense of their clients.

Eraut (1994, pp. 111–112).

The “clients” of teachers are, of course, their students. However, not only are these students active participants in the schooling process—making their views and feelings known, sometimes in no uncertain terms—but the aim of teaching is to elicit some sort of response from them, whether that be in the form of long-term outcomes or immediate classroom behaviors. If, as Eraut suggests, teachers’ routines become progressively dysfunctional, then the desired responses from students will fail to materialize. This situation will be exacerbated where students are less reluctant to make their dissatisfaction known to their teachers, and where teachers and schools are held tightly to account for delivering certain sorts of outcomes.

The implication of all this is that there is an in-built dynamic which brings teachers face-to-face with a mismatch between their routinized practice and the actual responses of their students. If inclusion is understood as a process of reducing barriers to learning and participation, then it seems reasonable to suggest that this in-built dynamic also brings teachers face-to-face with the non-inclusive nature of their practices. If we then go on to see school staffs as consisting of communities of practice, we can see how a core task of such groups is to develop shared meanings (and therefore practices) around the recurrent experience of mismatch between established practices and the realities of their students. This does not mean, of course, that every mismatch will be acknowledged as an “anomaly”, or that the shared meanings that are developed around them will lead to more inclusive practices. However, it does mean that fundamental questions about inclusiveness arise and can potentially be made explicit as a matter of course in *every* school, regardless of whether or not those involved espouse a rhetoric of inclusion.

Our work suggests that this is not a merely theoretical possibility. In some of our project schools we saw how a mismatch between established practices and the actual outcomes of that practice became apparent. This might relate to a group of students whose exclusion from classrooms ceases to be taken for granted, or those who “disrupt” established practices, or students who “fail” to reach desired targets. We have also worked with schools where shifting demographics produced intakes that responded differently to the schools’ traditional population, or classrooms in which individual students seemed to be doing less well than their teachers hoped, or experiences of new practices which suggested to teachers that their students could do better if they abandoned their existing ways of working.

Final thoughts

The research summarized in this chapter points to the sorts of conditions that are needed in order to use processes of collaboration to foster equity within education systems. As I have explained, this is based on the idea that schools have untapped potential to

improve their capacity for improving the achievement of all of their students, particularly those from poorer backgrounds. The challenge therefore is to mobilize this potential. This reinforces the argument that school improvement is a social process that involves practitioners in learning from one another, from their students, and from others involved in the lives of the young people they teach. As we have seen, an engagement with evidence can be a powerful catalyst for making this happen.

A helpful theoretical interpretation that can be made of these processes is that, together, they help to strengthen social capital. In other words, they create pathways through which expertise and lessons from innovations can spread. In recent years, the work of Robert Putnam has been influential in making the idea of social capital a focus for research and policy discussion. He explains:

Whereas physical capital refers to physical objects and human capital refers to the properties of individuals, social capital refers to connections among individuals – social networks and the norms of reciprocity and trustworthiness that arise from them.

Putnam (2000, p. 19).

Writing about the United States, Putnam states that “what many high-achieving school districts have in abundance is social capital, which is educationally more important than financial capital” (p. 306). He also suggests that this can help to mitigate the insidious effects of socioeconomic disadvantage.

It seems, then, that a major factor in determining success in promoting equity in education is our ability to strengthen social capital. Therefore, the task of those involved in leadership roles is to create the climate that will support such developments.

This means that equitable education systems must be connected systems, and that schools have to work together, and with others, both within and beyond their local areas. It also points to the vital role of local area coordination in promoting greater equity. It shows, too, that pursuing arrangements that are more equitable does not always mean schools individually doing more. However, as I have argued, it does imply partnerships beyond the school, where partners multiply the impacts of each other's efforts.

Finally, there are significant implications here for national policy makers. To support the promotion of greater equity, they need to develop policy that can learn from and nurture innovative developments. They also have to determine how they can foster the conditions needed to encourage the interpretation of policy on the ground in ways that promote the principle that all learners matter and matter equally.

Acknowledgments

Many colleagues have contributed to the programme of research described in this paper, too many to mention. However, the ideas of Alan Dyson were particularly important to the argument I develop.

References

- Ainscow, M., 1999. *Understanding the Development of Inclusive Schools*. Routledge, London.
- Ainscow, M., 2010. Achieving excellence and equity: reflections on the development of practices in one local district over 10 years. *Sch. Eff. Sch. Improv.* 21 (1), 75–91.
- Ainscow, M., 2012. Moving knowledge around: strategies for fostering equity within educational systems. *J. Educ. Change* 13 (3), 289–310.
- Ainscow, M., 2015. *Towards Self-Improving School Systems: Lessons From a City Challenge*. Routledge, London.
- Ainscow, M., Booth, T., Dyson, A., Farrell, P., Frankham, J., Gallannaugh, F., Howes, A., Smith, R., 2006a. *Improving Schools, Developing Inclusion*. Routledge, London.
- Ainscow, M., Chapman, C., Hadfield, M., 2020. *Changing Education Systems: A Research-Based Approach*. Routledge, London.
- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012a. *Developing Equitable Education Systems*. Routledge, London.
- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012b. Making schools effective for all: rethinking the task. *Sch. Leadersh. Manag.* 32 (3), 1–17.
- Ainscow, M., Dyson, A., Hopwood, L., Thomson, S., 2016. *Primary Schools Responding to Diversity: Barriers and Possibilities*. Cambridge Primary Review Trust, New York, NY.
- Ainscow, M., Hargreaves, D.H., Hopkins, D., 1995. Mapping the process of change in schools: the development of six new research techniques. *Eval. Res. Educ.* 9 (2), 75–89.
- Ainscow, M., Hopkins, D., 1992. Aboard the 'moving school'. *Educ. Leadersh.* 50 (3), 79–81.
- Ainscow, M., Howes, A., 2007. Working together to improve urban secondary schools: a study of practice in one city. *Sch. Leadersh. Manag.* 27, 285–300.
- Ainscow, M., Howes, A.J., Farrell, P., Frankham, J., 2003. Making sense of the development of inclusive practices. *Eur. J. Spec. Needs Educ.* 18 (2), 227–242.
- Ainscow, M., Muijs, D., West, M., 2006b. Collaboration as a strategy for improving schools in challenging circumstances. *Improv. Sch.* 9 (3), 1–11.
- Ainscow, M., Southworth, G., 1996. School improvement: a study of the roles of leaders and external consultants. *Sch. Eff. Sch. Improv.* 7, 229–251.
- Ainscow, M., West, M. (Eds.), 2006. *Improving Urban Schools: Leadership and Collaboration*. Open University Press.
- Ainscow, M., West, M., Nicolaidou, M., 2005. Putting our heads together: a study of headteacher collaboration as a strategy for school improvement. In: Clarke, C. (Ed.), *Improving Schools in Difficult Circumstances*. Continuum, London.
- Argyris, C., Schon, D., 1996. *Organisational Learning II: Theory, Method and Practice*. Addison Wesley, Reading, MA.
- Armstrong, P., Ainscow, M., 2018. School-to-school support within a competitive education system: views from the inside. *Sch. Eff. Sch. Improv.* 29 (4), 614–633.
- Au, W., 2009. *Unequal by Design: High-Stakes Testing and the Standardization of Inequality*. Routledge, London.
- Barrs, S., Bernardes, E., Elwick, A., Malortie, A., McAleavy, T., McInerney, L., Menzies, L., Rigall, A., 2014. *Lessons From London Schools: Investigating the Success*. CfBT Trust, Reading.
- Bubb, S., Crossley-Holland, J., Cordiner, J., Cousin, S., Earley, P., 2019. *Understanding the Middle Tier: Comparative Costs of Academy and LA-Maintained School Systems*. Sara Bubb Associates, London.
- Campbell, A., Groundwater-Smith, S. (Eds.), 2007. *An Ethical Approach to Practitioner Research: Dealing With Issues and Dilemmas in Action Research*. Routledge, London.
- Chapman, C., Ainscow, M. (Eds.), 2021. *Educational Equity: Pathways to Success*. Routledge, London.

- Chapman, C., Hadfield, M., 2010. School-based networking for educational change. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second Handbook of Educational Change*. Springer, London.
- Claeys, A., Kempton, J., Paterson, C., 2014. Regional Challenges: A Collaborative Approach to.
- Cochran-Smith, M., Lytle, S.L., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York, NY.
- Copland, M.A., 2003. Leadership of inquiry: building and sustaining capacity for school improvement. *Educ. Eval. Policy. Anal.* 25 (4), 375–395.
- Deppeler, J., 2013. Developing equitable practices in schools: Professional collaboration in research. In: Jones, P. (Ed.), *Bringing Insider Perspectives into Inclusive Teacher Learning: Potentials and Challenges for Educational Professionals*. Routledge, New York, NY, pp. 178–188.
- Dobbie, W., Fryer, R.G., 2009. Are High-Quality Schools Enough to Close the Achievement Gap? Evidence From a Bold Social Experiment in Harlem. Harvard University, Cambridge.
- Dyson, A., Gallannaugh, F., Millward, A., 2003. Making space in the standards agenda: Developing inclusive practices in schools. *Eur. J. Educ. Res.* 2, 228–244.
- Dyson, A., Howes, A., Roberts, B., 2004. What do we really know about inclusive schools? A systematic review of the research evidence. In: Mitchell, D. (Ed.), *Special Educational Needs and Inclusive Education: Major Themes in Education*. Routledge, London.
- Dyson, A., Kerr, K., 2013. *Developing Children's Zones for England: What's the Evidence? Save the Children*, London.
- Elliott, J., 1991. *Action Research for Educational Change*. Open University Press, Buckingham.
- Fielding, M., 2001. Students as radical agents of change. *J. Educ. Change* 2 (2), 123–141.
- Fielding, M., Bragg, S., Craig, J., Cunningham, I., Eraut, M., Gillinson, S., Horne, M., Robinson, C., Thorp, J., 2005. *Factors Influencing the Transfer of Good Practice*. DfES Publications, Nottingham.
- Fullan, M., 2007. *The New Meaning of Educational Change*, fourth ed. Teachers College Press, New York, NY.
- Fullan, M., Rincon-Gallardo, S., Hargreaves, A., 2015. Professional capital as accountability. *Educ. Policy Anal. Arch.* 23 (15). <https://doi.org/10.14507/epaa.v23.1998>.
- Greaves, E., Macmillan, L., Sibieta, L., 2014. *Lessons From London schools for Attainment Gaps and Social Mobility*. The Social Mobility and Child Poverty Commission, London.
- Grossman, A.S., Lombard, A., Fisher, N., 2014. *Strive Together: Reinventing the Local Education Ecosystem*. Harvard Business School.
- Groundwater-Smith, S., 2011. Concerning equity: the voice of young people. *Lead. Manag.* 17 (2), 52–65.
- Hargreaves, A., Ainscow, M., 2015. *The Top and Bottom of Leadership and Change*. Phi Delta Kappa. November, 2015.
- Hargreaves, A., Shirley, D., 2009. *The Fourth Way: The Inspiring Future for Educational Change*. Corwin Press, Thousand Oaks, CA.
- Hiebert, J., Gallimore, R., Stigler, J.W., 2002. A knowledge base for the teaching profession: what would it look like and how can we get one? *Res. Educ.* 31 (5), 3–15.
- Hill, R., 2008. *Achieving more together: adding value through partnership*. Leicester: ASCL.
- Hopkins, D., 2007. *Every School a Great School: Realizing the Potential of System Leadership*. Open University Press, Maidenhead.
- Hopkins, D., Ainscow, M., West, M., 1994. *School Improvement in an Era of Change*. Cassell, London.
- Howes, A., Booth, T., Dyson, A., Frankham, J., 2005. Teacher learning and the development of inclusive practices and policies: Framing and context. *Res. Pap. Educ.* 20, 133–148.
- Howes, A., Frankham, J., Ainscow, M., Farrell, P., 2004. The action in action research: Mediating and developing inclusive intentions. *Educ. Action Res.* 12, 239–258.
- Huberman, M., 1993. *The model of the independent artisan in teachers' professional relationships*. In: Little, J.W., McLaughlin, M.W. (Eds.), *Teachers' Work: Individuals, Colleagues and Contexts*. Teachers College Press, New York, NY.
- Hutchings, M., Hollingworth, S., Mansaray, A., Rose, R., Greenwood, C., 2012. *Research Report DFE-RR215: Evaluation of the City Challenge Programme*. Department for Education, London.
- Jensen, B., Sonnemann, J., 2014. *Turning Around Schools: It Can Be Done*. Grattan Institute, Melbourne.
- Kahlenberg, R.D., Potter, H., 2014. *A Smarter Charter: Finding What Works for Charter Schools and Public Education*. Teachers College Press, New York, NY.
- Kemmis, S., 2010. *Research for praxis: knowing doing*. *Pedagogy Cult. Soc.* 18, 9–27.
- Kemmis, S., McTaggart, R., 1988. *The Action Research Planner*, third ed. Deakin University Press, Geelong.
- Kerr, K., Dyson, A., Raffo, C., 2014. *Education, Disadvantage and Place: Making the Local Matter*. Policy Press, Bristol.
- Kershaw, I., 2014. *Investigation Report: Trojan Horse Letter (Prepared for Birmingham City Council)*. Eversheds, London.
- Kidson, M., Norris, E., 2014. *Implementing the London Challenge*. Joseph Rowntree Foundation, London.
- Lambert, L., 2005. Constructivist leadership. In: Davies, B. (Ed.), *The Essentials of School Leadership*. Paul Chapman Publishing and Corwin Press, London, pp. 93–109.
- Lewin, K., 1946. Action research and minority problems. *J. Soc. Issues* 2 (4), 34–46.
- Lewis, C., Perry, R., Murata, A., 2006. How should research contribute to instructional improvement? The case of lesson study. *Res. Educ.* 35 (3), 3–14.
- Locke, T., Alcorn, N., O'Neill, J., 2013. Ethical issues in collaborative action research. *Educ. Action Res.* 21, 107–123.
- Lubienski, C., 2003. Innovation in education markets: theory and evidence on the impact of competition and choice in Charter Schools. *Am. Educ. Res. J.* 40 (2), 395–443.
- McIntyre, D., Pedder, D., Rudduck, J., 2005. Pupil voice: comfortable and uncomfortable learnings for teachers. *Res. Pap. Educ.* 20 (2), 149–168.
- Messiou, K., Ainscow, M., 2015. Engaging with the views of pupils: a catalyst for powerful teacher development? *Teach. Teach. Educ.* 51, 246–255.
- Messiou, K., Ainscow, M., Echeita, G., Goldrick, S., Hope, M., Paes, I., et al., 2016. Learning from differences: a strategy for teacher development in respect to student diversity. *Sch. Eff. Sch. Improv.* 1, 27. <https://doi.org/10.1080/09243453.2014.966726>. January 2016.
- Miles, S., 2010. Leading under pressure: leadership for social inclusion. *Sch. Leadersh. Manag.* 143–157.
- Miles, S., Ainscow, M., 2011. *Responding to Diversity in Schools*. Routledge, London.
- Ministry of Education, 2019. *Supporting All Schools to Succeed: Reform of the Tomorrow's Schools System*. Ministry of Education, New Zealand.
- Mitra, D.L., 2004. The significance of students: can increasing 'student voice' in schools lead to gains in youth development? *Teach. Coll. Rec.* 106 (4), 651–688.
- Muijs, D., Ainscow, M., Chapman, C., West, M., 2011. *Collaboration and Networking in Education*. Springer, London.
- Muijs, D., Ainscow, M., Dyson, A., Raffo, C., Goldrick, S., Kerr, K., Lennie, C., Muijs, D., West, M., Ainscow, M., 2010. Why network? Theoretical perspectives on networking. *Sch. Eff. Sch. Improv.* 21 (1), 5–26.
- Mulford, B., 2007. Building social capital in professional learning communities: Importance, challenges and a way forward. In: Stoll, L., Seashore Louis, K. (Eds.), *Professional Learning Communities: Divergence, Depth and Dilemmas*. Open University Press, London.
- OECD, 2012. *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. OECD Publishing, Paris.
- OECD, 2015. *Improving Schools in Scotland: An OECD Perspective*. OECD Publishing, Paris.
- Payne, C.M., 2008. *So Much Reform, So Little Change: The Persistence of Failure in Urban Schools*. Harvard Education Press, Cambridge.
- Pickett, K., Vanderbloemen, L., 2015. *Mind the Gap: Tackling Social and Educational Inequality*. Cambridge Primary Review Trust, New York, NY.
- Putnam, R.D., 2000. *Bowling Alone*. Simon & Schuster, New York, NY.
- Reason, P., Bradbury, H. (Eds.), 2001. *Handbook of Action Research: Participative Inquiry and Practice*. Sage, London.
- Riehl, C.J., 2000. The principal's role in creating inclusive schools for diverse students: a review of normative, empirical, and critical literature on the practice of educational administration. *Rev. Educ. Res.* 70 (1), 55–81.
- Sabel, C., Saxenian, A., Miettinen, R., Kristenson, P.H., Hautamaki, J., 2011. *Individualized Service Provision in the New Welfare State: Lessons From Special Education in Finland*. SITRA, Helsinki.
- Sadker, D.M., Sadker, M., Zittleman, K.R., Sadker, M., 2009. *Still Failing at Fairness: How Gender Bias Cheats Girls and Boys in School and What We Can Do About It*. Scribner, New York, NY.
- Schildkamp, K., Ehren, M., Kuin Lai, M.K., 2012. Editorial article for the special issue on data-based decision making around the world: from policy to practice to results. *Sch. Eff. Sch. Improv.* 23 (2), 123–131.

- Schleicher, A., 2010. International comparisons of student learning outcomes. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second Handbook of Educational Change*. Springer, London.
- Schön, D., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York, NY.
- Skrlic, T.M., 1991. The special education paradox: equity as the way to excellence. *Harv. Educ. Rev.* 61 (2), 148–206.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development*. Heinemann, London.
- Talbert, J.E., Mileva, L., Chen, P., Cor, K., McLaughlin, M., 2010. *Developing School Capacity for Inquiry-Based Improvement: Progress, Challenges, and Resources*. Stanford University; Center for Research on the Context of Teaching, Stanford, CA.
- Takala, M., Pirttimaa, R., Tormane, M., 2009. Inclusive special education: the role of special education teachers in Finland. *Br. J. Spec. Educ.* 36 (3), 162–172.
- UNESCO, 2015. *Education for All 2000–2015: Achievements and Challenges*. UNESCO, Paris.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, Cambridge.
- West, M., Ainscow, M., 2010. Improving schools in Hong Kong: a description of the improvement model and some reflections on its impact on schools, teachers and school principals. In: Huber, S. (Ed.), *School Leadership - International Perspectives*. Springer, London.
- Whitehurst, G.J., Croft, M., 2010. *The Harlem Children's Zone, Promise Neighborhoods, and the Broader, Bolder Approach to Education*. The Brookings Institution, Washington, DC.
- Wiborg, S., 2010. *Swedish Free Schools: Do They Work?* Centre for Learning and Life Chances in Knowledge Economies and Societies. <http://www.llakes.org>.
- Yurkofsky, M.M., Peterson, A.J., Mehta, J.D., Horwitz-Willis, R., Frumin, K.M., 2020. Research on continuous improvement: exploring the complexities of managing educational change. *Educ. Res. Rev.* 44, 403–433.
- Zancajo, A., 2019. Education markets and schools' mechanisms of exclusion: the case of Chile. *Educ. Policy Anal. Arch.* 27 (130).

Reclaiming accountability

Christine Gilbert, UCL Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	293
Defining accountability	293
Accountability and professional practice	294
The impact of public accountability systems	294
Developing a system of professional accountability	296
The shift in culture and practice needed	296
Building professional ownership of accountability	297
Leadership for professional accountability	297
Schools as centers of professional learning	298
Developing collaborative and rigorous self-evaluation	299
Locality partnerships as support for professional accountability	300
Inspection as support for professional accountability	301
Conclusion	301
References	302

Introduction

It is hard to imagine any sustained debate about education reform in which the word “accountability” would not be used. Usually, it is interpreted negatively as unhelpful pressure on schools and teachers. This is because it is generally associated with public accountability: results, league tables and inspection. But accountability in its truest sense is much broader and more important than that. A system of professional accountability operating at the level of the individual teacher, the individual school and across networks of schools can provide a source of professional aspiration and improve teachers’ knowledge, skills and practice so that children are supported better in their learning. It provides important support for professionalism in schools, teacher development and school improvement. As such, it deserves championing as an essential and even energizing element of any school system.

Defining accountability

The word accountability has fallen into disrepute among professionals over the years because it is too often understood as something imposed on schools by external authorities. Many now suggest it should be replaced by the word “responsibility,” which does not carry the same pejorative connotations. Although the two words “responsibility” and “accountability” are often used interchangeably, there is a subtle difference in meaning which is important and should not be lost.

Responsibility is best explained as a duty to take care of something. Each individual’s sense of responsibility differs and stems from their experiences, their beliefs and personal values. The term “accountability” assumes responsibility but it does more than that. It describes a relationship whereby one party—such as an individual, an institution or another entity—has an obligation to account for their actions or performance to another (Brundrett and Rhodes, 2011). Put simply, they are answerable to someone and there are expectations—whether explicit or not—related to that. The obligation assumes that whoever is giving account, and being held to account, has some responsibility for the actions or performance. They might therefore be considered worthy of praise or blame. So, accountability involves responsibility, expectations and accounting, which may be evaluated against established or expected standards and action taken. Responsibility is the more personal concept, with people feeling an intrinsic sense of responsibility, for individual children, for example. Giving an account is seen as a less personal, more systemic concept (Mongon and Chapman, 2012).

The key question for schools is: who are they accountable to and for what? In their exploration of accountability in education, Earley and Weindling (2004) identify four relationships for schools. Each relationship signals some sort of stake in the outcomes of education and, therefore, expectations of schooling. Schools have responsibilities for and have to account to:

- Pupils—and, I would add to parents and the local community (*moral accountability*)
- Colleagues (*professional accountability*)
- Employers or government (*contractual accountability*)
- The market, where clients, such as parents, have a choice of institution (*market accountability*).

These four relationships encompass two key approaches to accountability. The first is a performance or productivity model which emphasizes outputs such as test and exam results (Elliott et al., 1981) or outcomes. The second is an improvement or process model

which focuses on school evaluation, opening practice to debate and critique. The purpose of the first approach, which would be characterized as a summative model, is to prove quality—it emphasizes what has been achieved. The purpose of the second model is more formative as it is designed to improve quality (Hopkins, 2007).

In England, the first model is well embedded in the current public accountability framework. The most hostile perception of this model is that it is based just on a centralized regime of data, tests and inspection, set by government and, invariably, negative, mechanistic and stressful. The model is generally now described as high stakes accountability, with sanctions or rewards for those being assessed or judged, whereas low stakes accountability provides information without direct actions linked to results.

Ten years ago, when I left Ofsted, England's inspection organization, where I had been Chief Inspector for 5 years, I presented the case (Gilbert, 2012) for a better balance between the two models. I continue to believe the second model needs to assume greater importance as it offers more leverage for change. However, in systems where accountability is as high stakes as the English system, balance is not possible as the first model is so deeply embedded. I am not arguing for the end of an external accountability system; we need one that not only gives an account at system level, without the current pressures, but also supports improvement to the system by contributing to teachers' and schools' professionalism. This should be more supportive of their work and result in greater ownership of the accountability system.

The best school leaders embrace both forms of accountability but with a working emphasis on the second. Elmore's research (2004: 134) indicates that schools with strong internal accountability, "a high level of agreement among members of the organization on the norms, values and expectations that shape their work," do well within the public accountability system. The word "high" is implicit here in front of "expectations." He (2004: 134) also highlights that schools with weak internal accountability, "low agreement and atomization," tend not to perform so well. Elmore's definition of a strong professional accountability system within schools is what I am using in this chapter and this is explored more fully in [Developing a system of professional accountability](#) section.

Accountability and professional practice

There is a wealth of research in health about responsibility and accountability at the individual, the organizational and the whole system level. This is far more extensive than for education, perhaps because the consequences of practice mistakes and failure can be so obviously life threatening. We can learn from the experience of other professions. Eraut's research (1994) focused on all professions and he described professional practice as autonomous and self-regulating, with individual professionals accountable firstly to themselves. He argued that professional accountability should be drawn from a commitment to service and should include:

- A moral commitment to serve the interests of clients
- A professional obligation to self-monitor and to periodically review the effectiveness of one's practice
- A professional obligation to expand one's repertoire, to reflect on one's experience and develop one's expertise
- An obligation that is professional as well as contractual to contribute to the quality of one's organization
- An obligation to reflect upon and contribute to discussions about the changing role of one's profession in wider society (Eraut, 1994: 236).

Written nearly 30 years ago, this provides a view of professional accountability that could not be more relevant in education today. It underlines the importance of the individual's accountability for the quality of their own work and to the people the profession serves. It points to the significance of school leaders and teachers as agents of their own accountability and the need to develop an internal system of professional accountability in schools, with clear roles for teachers and, indeed, for all those involved in the education community.

The impact of public accountability systems

Schools are publicly funded organizations and so governments are expected to account for the quality of education they provide. In establishing public accountability systems, governments want to ensure that public money is being well spent on high quality education. They—and the organizations that serve the systems—invariably have a belief too that accountability itself contributes to improvement in education. For example, when first established, Ofsted took on the strapline, "Improvement through Inspection" changing it in 2006 to "Raising standards, Improving Lives," when it extended its inspection remit and also assumed regulatory responsibilities.

The research evidence, however, is mixed, with many accountability systems showing both benefits and problems. Research is also restricted by the limited nature of the evidence base, particularly as the details of inspection frameworks change so frequently. Nevertheless, a search of the literature provides evidence of impact, both negative and positive, which is important contextually in shaping a different system. Consequently, I outline below significant findings about its impact on leaders and teachers, curriculum, pupils and groups of pupils.

The overriding message from an analysis of much of the evidence is that the impact of external accountability systems stems from the pressure they generate on the organizations being held to account and on the people within them. Pressure can be productive, but it can also have negative and unintended consequences.

Matthews (2014) argues that inspection raises standards by defining high expectations for effective schools and assessing practice against them. Ehren (2016), in summarizing the findings from a range of international empirical research, identifies ways in which inspection can impact on school improvement. Several of these research studies indicate that when the outcomes of inspections are negative, greater improvement results. Ehren also highlights the influence on schools' practice of the standards inspectors use to assess quality and the pressure on leaders from the prospect and consequences of either inspection failure or reward. There are considerable risks relating to such standardization, in particular that it can lead to conformity and compliance, as well as distracting schools from developing a culture of professional learning (Ehren, 2019).

Using online survey data from approximately 2300 principals in seven European countries, Altricher and Kemethofer (2015) also looked at whether the accountability pressure generated by school inspections promoted school improvement. School leaders in high stakes systems, such as England, reported more pressure than low stake systems such as Austria and Switzerland. They found that school leaders who feel more pressure to do well in inspections also took more account of the quality expectations established by inspections. These leaders also showed more sensitivity to stakeholders' reactions to inspection and undertook more development activities in their schools both pre and post inspection. However, pressure was also found to be detrimental to development; school leaders experiencing most pressure identified unintended consequences, such as narrowing of the curriculum and the discouragement of innovative teaching strategies.

A literature review on the impact of accountability (Brill et al., 2018) found it could stimulate the development of teachers' skills to engage well with assessment data in order to improve teaching and learning. Such professional development is positive. However, the accountability system (Lynch et al., 2016) is also experienced as a source of frustration and additional workload for teachers. It is seen as impacting on job satisfaction and lowering morale, including resulting in some teachers leaving the profession early. Wilkins (2011) reports a generational shift in attitudes to accountability with new teachers and those early in their careers seeing it as a core part of the job. Certainly, 10 years on, in the English system there are very few teachers who haven't grown up with a strong system of external accountability deeply embedded in the school system. This needs to be considered in any move to a system driven more strongly by professional accountability.

There are numerous studies reflecting widespread concern about the impact of external accountability systems on the curriculum. The literature review (2018), referred to above, identified concerns about a focus on literacy and numeracy at the expense of a broad and balanced curriculum and also a tendency to "teach to the test." The review found that this was particularly the case when the accountability measures concentrated on pupil performance in a limited range of subjects.

International benchmarking, for example that provided by survey results from PISA, TIMSS or PIRLS, can have an impact on system-wide accountability. It has influenced national discussions about curriculum and led to reforms in some countries. For example, it has been seen to have a direct impact on national education policy and provision. Japan, a consistently high performer in PISA, responded to its decline in the PISA rankings by revising "key competencies" within its national curriculum and placing greater emphasis on cross-disciplinary skills such as independent thinking (Nakayasu, 2016). These were more in line with the skills PISA aimed to measure.

A narrowing of the curriculum is problematic for all children but is likely to have greatest impact on low performing pupils. Often these consequences are unintended. For example, a report on the Welsh education system (OECD, 2014) highlights the danger of a performance target for pupil attainment, laudably aspirational, leading to schools focusing attention on pupils just below that required target level. This might well have the unintended consequence of very low attainers, often including many disadvantaged pupils, being among those left out. The performance target was at variance with the Welsh government's other reform objective of reducing the impact of deprivation on pupil performance.

In Australia too, a similarly negative effect on the curricular experiences of children was identified (APPA, 2010, cited in the 2018 literature review), with the suggestion that parents of low-performing children were discouraged from allowing them to take the test.

These examples point to ongoing concerns about unethical practices and "gaming the system."

Much research draws attention to the close relationship between public accountability and pupils' standards of achievement, with the latter often being used as a way of measuring educational quality. For example, research undertaken by Burgess et al. (2013) tested the hypothesis that the abolition of school performance tables, often described as league tables, negatively affects school effectiveness. Comparing outcomes in England and Wales before and after the abolition of tables by the Welsh government in 2001, they concluded that this change markedly reduced school effectiveness in Wales. Standards in secondary schools fell and the gap between Wales and England widened. They also found heterogeneity across schools, with those previously in the top quartile of the tables showing that abolition had no effect on results; no effect was seen either when analyzing by ability or by socio-economic status.

Brill et al. (2018) report the potential of the public accountability system for both increasing or reducing the attainment gap. For example, just as narrowing the curriculum can impact on low attaining pupils, often the most disadvantaged, so too will the tendency to focus extra support on those pupils who are most likely to contribute to the achievement of targets. There is evidence that pupils also become less engaged when emphasis is given to some groups at the expense of others.

Groups of pupils working below the expected achievement threshold are at risk here. The Bew Review (DfE, 2011), which looked at the English system of key stage 2 testing (for children aged 10–11 years), assessment and accountability, drew attention to the concern that the accountability system discouraged schools from admitting pupils with special educational needs and the difficulties of addressing this concern. On the positive side, there is evidence that data collected and used for accountability purposes can be used to target additional, funded support for lower performing pupils. In Japan, for example (OECD, 2012), additional teachers are allocated to those low-performing schools which have a greater number of disadvantaged pupils.

Research evidence has highlighted test anxiety as a potential consequence when pupil assessment features strongly in accountability regimes. Goodman and Burton (2012) suggest that disadvantaged children, in particular, may be affected by anxiety and pressure. Kuramoto and Koizumi (2016) argue that tests in the Japanese system are perceived to lead to a competitive classroom atmosphere and anxiety. However, given the many other factors that may influence pupils' experiences of assessment, the NFER Review concludes that it is difficult to assess the extent to which it is testing, carried out for the purposes of accountability, that causes anxiety.

While national accountability frameworks clearly have value in establishing transparency and giving public confidence about whether public money is being well-spent and schools are delivering high quality education, there is little evidence that external accountability systems are the major driver of improvements in pupil and school performance. Nevertheless, key stakeholders, such as parents, value the information provided and use it to inform choice. Governments might improve public accountability frameworks but they are unlikely to abandon them. It is important, therefore, for policy makers to consider the many ways in which transparency can be provided to the public without such unproductive pressure on teachers and schools and in ways that support professional accountability.

A range of different approaches, some of which are referenced in 5.7 below, could continue to provide public reassurance and also contribute to the development of professional accountability within and across schools. These systems could still monitor schools and identify those requiring intervention and intensive support. They could both prove quality but focus more on improving it too.

Developing a system of professional accountability

The shift in culture and practice needed

As indicated earlier, schools manage a range of accountability relationships. Particularly in England, the focus of the last 30 years has been on accountability to government or employers (*contractual accountability*) and to the market (*market accountability*) where parents have a choice of school. If we are to create a strong system of professional accountability which supports teachers and schools as they improve learning, greater leverage for improvement has to come more directly from accountability to pupils (*moral accountability*) and to colleagues (*professional accountability*). It has to come too from the external accountability framework.

Professional accountability requires an emphasis on both the individual teacher and the school. Schleicher (2018) describes jurisdictions such as Ontario, Finland, Japan and New Zealand as placing greater importance on the development of professionalism and collaboration. In some systems, such as Singapore, teachers' annual appraisal includes not only their contribution to the academic and character development of their pupils but also their collaboration with parents and community, and their impact on their colleagues and the school as a whole. In Shanghai, each teacher's appraisal focuses not only on how well children have been taught and what they have achieved but on the teacher's contribution to the profession or wider education system. These approaches to accountability recognize the importance of individual professional accountabilities in driving change and innovation at system level.

In any system, it is the teachers themselves—in particular through the quality of their teaching and their relationship and interactions with pupils—who make the most difference to pupils' learning. Teachers themselves have to be at the center of any internal accountability system. They have to own it and drive it. They must feel supported by it. As Matthews (2014) argues, "Ultimately, accountability of a profession rests with the professionals. The more they can manage their own quality control and continuing improvement, the less need there should be for external inspection" (2014: 28).

The best school leaders establish a culture that enables teachers to feel greater ownership of all forms of accountability and see it as something that supports their work. They encourage teachers to develop a strong sense of both individual and collective purpose and to feel part of a professional learning community where a continuing process of review and dialog about pedagogy and practice leads to change, innovation and better learning for pupils. An open and trusting analysis would include dialog about things that are not working well and using this to learn and develop professionally. So, a system of professional accountability, rooted in moral purpose and professionalism, requires us to move forward in a number of ways:

- Creating a shift in mind-set and culture so that accountability is professionally owned rather than being seen as externally imposed
- Using accountability as a force for learning to energize schools, motivate teachers and grow their self-efficacy, thereby building professional and social capital, and so system capability and capacity
- Using data, evidence and observation to support organizational improvement, particularly in relation to teaching and learning
- Ensuring school evaluation is a dynamic and inclusive process, which involves pupils, parents, staff and key players in the local and wider educational community, building professionalism and leading to better practice
- Establishing a culture of professional reflection, enquiry and learning within and across schools that increases teachers' aspirations and the development of better practice in teaching and pupils' learning
- Capturing the benefits more fully of within-school and across-school collaboration, in particular through peer learning, for challenging and developing the work of teachers and the learning of pupils
- Using school networks to develop capacity and ensure no school is isolated
- Focusing inspection to give greater support to professional accountability.

Building professional ownership of accountability

The key unit within any accountability framework remains the individual school. Each child and parent relates to the school and, indeed, to the teachers most concerned with that child's education. At this level, too, the teacher's moral and professional accountabilities are not only clear, but inextricably linked to and focused on real children and young people. A research study from the General Teaching Council (GTC, 2009) showed that teachers felt accountability most strongly in relation to their pupils. They felt particularly accountable for the quality of their own teaching and they acknowledged that maintaining public confidence in standards of teaching was an important part of this. Good leaders are able to build on these individual feelings of professional accountability so they become collective and an essential part of thinking and practice in the school. They build a culture so teachers feel accountable for all pupils in the school, not just those they teach.

At the beginning of 2012, I visited a number of first-tranche teaching schools in England (Gilbert, 2017). Each had been involved in system improvement for some time. All these schools:

- Expected teachers to be accountable for the quality of teaching and the impact it was having on the learning of individual pupils
- Were creating strong professional communities, where peer learning was central and focused on the detail of practice and pedagogy
- Used school-to-school collaboration to strengthen their review of practice and pedagogy
- Were engaging pupils in reviewing their own learning
- Targeted resources flexibly to meet both pupil and staff needs
- Were giving regular accounts to parents of their children's progress
- Gave regular reports to governing bodies about progress and performance at school level and, increasingly, about the progress and performance of the schools with which they were working.

The competition to become a teaching school in the first tranche had been considerable. These schools had all been judged outstanding by Ofsted; they had to demonstrate not only a strong tradition of professional development and learning but also of collaboration with other schools. This was therefore a skewed sample of schools but they had all embedded professional accountability. Many resented the pressure they felt from the public accountability framework but had the confidence and motivation not to be weighed down by it. Teachers talked with enthusiasm about working in schools with the best professional practices but they were also motivated by the impact of their work on children. It was a mark of their professionalism that they moved comfortably between formative and summative accountability; they saw the need for each, and several teachers described both approaches as "feeding each other." These schools had a confidence that is still too rare.

In their analysis of system leadership, Higham et al. (2009) point to the importance of strong internal accountability within schools, very similar to what I am describing as professional accountability. The authors looked at three schools in challenging circumstances where improvement had been sustained. Staff in these schools reported feeling very accountable. This sometimes stemmed from what they perceived to be managerial pressure but over time the pressure shifted to something more professional: "In particular, regular peer observation and collaborative planning helped to create a shared language about what was being found effective in engaging pupils in their learning" (Higham et al., 2009: 50).

In schools where professional accountability is embedded, teachers see themselves as responsible for the quality of their teaching and its impact on pupils' learning. They see learning from each other as routine. They welcome opportunities to engage pupils, peers and their own managers confidently and openly to generate greater intelligence, which they use to develop their knowledge and skills. They see such engagement as important not only in offering them practical insights into their work and pupils' learning but a good platform for developing better and sometimes innovative practice. This approach is central to their professional accountability and, in particular, their accountability for improving quality.

How can we find ways of enabling more schools to build strong, internal systems of professional accountability? How can we encourage them to ease their anxieties about a more vigorous approach to formative accountability and to help them see its potential as an aid to improvement as well as a support for test results and the outcomes of inspection?

Leadership for professional accountability

Leadership is the essential ingredient for change. It is leadership that establishes the vision, values, expectations, structures and processes for getting the work done. It is leadership which builds an internal, professional accountability system that is owned collectively. Whether they acknowledge it or not, all schools operate with various forms of accountability built into their day-to-day work. These stem from the leadership of the school and the culture that has developed—the expectations, behaviors and routines - as well as the different sense of responsibility and commitment that individuals bring. Accounting to whom, for what, and how therefore operates differently in different schools. The consequences of accounting are different too.

The four core domains of leadership practice identified through research (Leithwood et al., 2006) have stood the test of time (Leithwood et al., 2019):

- Building vision and setting clear directions
- Developing relationships and people

- Ensuring organizational design works in support of what's wanted
- Managing and giving active supporting to the teaching and learning program.

These are important but research also indicates that it is how these practices are enacted that remains crucial (Day et al., 2016, 2011). Leadership practice is what connects vision and policy to implementation and performance in schools.

Large scale statistical research (Louis et al., 2010) across several countries about effective school leaders found collective or shared leadership more powerful than individual leadership. The former is leadership that involves people at every level of the community in decision-making, whereas individual leadership is top-down leadership by the principal and a few senior leaders. The research concluded:

- Collective leadership has a stronger influence on student achievement than individual leadership
- Almost everyone in high-performing schools has greater influence on decisions than in low-performing schools
- Higher-performing schools award greater influence to teacher teams, parents and students
- Principals and district leaders have the most influence on decisions in all schools but they do not lose influence as others gain it
- School leaders have an impact on student achievement primarily through their influence on teachers' motivation and working conditions; their influence on teachers' knowledge and skills produces less impact on student achievement (2010: 19).

This sort of collective leadership has the potential for creating a culture where accountability is energizing and a force for improvement. It can create the conditions for individuals, teams and schools to own accountability and see it as an essential part of their work and development. They establish a culture of learning which feeds their internal accountability system to improve the quality of their teaching and their pupils' learning, enabling the school to work more effectively.

Based on research with a range of schools, Elmore (2004) also concluded that successful schools build a collaborative culture and internal accountability aligning individual responsibility and collective expectations. These in turn guided the motivation and actions of teachers. Staff share what I would describe as strong professional accountability. The schools in Elmore's research also performed well under the pressures of external accountability. Further research by Elmore (2008) confirms that schools with strong professional accountability seem to perform well on tests, regardless of whether their educational beliefs align with the tests.

Most research on leadership, for example in both the work of Elmore and Louis et al. cited above, emphasizes the importance of relationships and, linked to that, trust. Schools that have a strong moral purpose, captured in a collective vision and values, underpinned by a clearly articulated model for professional improvement have found it easier to establish a culture of openness and trust which makes people feel that change is possible. Trust is built on interactions and good relationships over time. If information-sharing, problem-solving and decision-making are genuinely collaborative, trust builds and improves the nature and quality of those relationships (Finnigan and Daly, 2012) so that individuals and schools are supported to learn better, take risks and innovate.

A good school accountability system must be able to identify and respond to specific problems in context. This can only work well if teachers feel safe to identify specific difficulties, to seek support and to innovate in a specific context. This includes permission to take risks and make mistakes, using these to learn. Greany and Earley (2017: 5) underline the importance of high trust in increasing leadership agency: "... where schools collaborate and share their expertise and capacity so that effective practice spreads, where leadership development and capacity building are prioritized and where leaders have a voice in shaping policy so they are committed to achieving shared goals ..."

Schools as centers of professional learning

School leaders who take collective professional accountability seriously know that it cannot be short term and is certainly not a quick fix. Culture takes time to develop and become embedded. A collaborative culture, which demands attention is paid to individual behavior, to teamwork and to organizational practices, needs to be supported by a focus on consistently building the school's capacity for learning. This requires organizational and resource investment, especially when it is rooted, as it should be, in a culture of practice analysis, professional development and peer learning.

The OECD (2016: 1) identified "seven action-oriented dimensions" as essential for what they describe as a model for the school as a learning organization:

- Developing and sharing a vision centered on the learning of all students
- Creating and supporting continuous learning opportunities for all staff
- Promoting team learning and collaboration among all staff
- Establishing a culture of inquiry, innovation and exploration
- Embedding systems for collecting and exchanging knowledge and learning
- Learning with and from the external environment and larger learning system
- Modeling and growing learning leadership.

Many successful school leaders operate this sort of model for learning. It can be used to help establish the conditions for an approach to accountability which energizes and supports staff. This requires an emphasis on locating the learning as close as possible to the internal work of the school, particularly in terms of practice relating to the detail of teaching and learning. This approach should include finding time for collective work on both individual and common problems. This builds professional and social capital. The negative connotations of the word mean that few leaders would use "accountability" to describe such

a learning environment. Placing “professional” in front of “accountability” gives back ownership of, and pride in, its moral purpose and professionalism.

Investment in learning is important. For example, time has to be found for teachers to work together, to observe, analyze, reflect on the detail of their teaching and pupils’ learning and then (which takes even longer) to shift deeply embedded practice through, for example, approaches such as instructional or incremental coaching. It is important that the focus is on visible evidence gained through observation of interactions, usually in the classroom. No single model of development guarantees success. Change needs to take account of organizational factors and specific contextual conditions of the school community (Gu et al., 2019).

Keeping learning as close as possible to the internal work of the school does not negate the need to look outwards. If the profession is to take greater ownership of accountability, evidence remains key. Collins and Coleman (2017) indicate evidence has become a feature of school system reforms over the last few decades. They argue persuasively that if its use is to become deeply embedded, it must be through an evidence-rich system and an approach focused on evidence for improvement. Professional accountability needs the sustenance gained through use of external research and evidence. It helps the school open its eyes as wide as possible, and as a consequence to learn and to thrive.

Developing collaborative and rigorous self-evaluation

The best schools know themselves well. This is the essential base for a professional and internal accountability system but also serves the demands of external accountability. The process of review and evaluation—at individual teacher level, at team level and whole school level—is fundamental to the way the school works. It is sensitive to the nuances of context and looks back as well as forward. With its focus on learners and learning, planned well, internal evaluation is an essential professional discipline. It can increase the sensitivity of teachers to the conditions for better practice, including the need for new knowledge and skills. It can serve the purposes of both internal and external accountability. As an extensive literature review on internal evaluation (Nelson et al., 2015) indicates, it is a powerful support for continuous improvement and professional development.

Data too, generated internally or from the public accountability framework, is important. This can raise questions, stimulate thinking and generate insights leading to action to improve practice. Data enables progress to be monitored. The role of quantitative data as a tool for school improvement is well rehearsed and widely accepted but data can emerge from a range of activities, principally observation, but also others, such as analysis of problems and case studies, or from interviews or focus groups. This sort of data is a strong support for professional accountability within and across schools.

The process of evaluation should operate at all levels within a school from interactions between teacher and pupil to whole school benchmarking. It may well be professionally led but it has to engage others. If moral accountability is to become more meaningful professionally, pupils, as well as their parents, have a key role to play in driving improvement and quality. They must have agency.

Pupils have a key role in evaluating the quality of support for learning and in helping to shape it too. Hattie (2009) undertook a synthesis of 134 meta-analyses of all possible influences on achievement. This showed clearly that pupil feedback was one of the most important influences. He captures the power of this approach in this description: “It was only when I discovered that feedback was most powerful when it is *from the student to the teacher* that I began to understand it better. When teachers seek, or at least are open to, feedback from students as to what students know, what they understand, where they make errors, where they have misconceptions, when they are not engaged—then teaching can be synchronized and powerful. Feedback to teachers helps make learning visible” (2009: 173).

As we saw during the pandemic, many schools have found new ways of interacting with parents about learning, including those who do not readily engage. There are examples where parents themselves have sometimes built their own virtual peer learning groups to support their children’s learning (Gilbert, 2021). There are signs too that schools’ increased confidence and skill in the use of technology post-pandemic is likely to lead to new and creative ways of engaging parents better.

The external accountability framework provides information for parents but test results, league tables and inspection reports give a narrow view about the school’s purposes. A global review (Cheng and Moses, 2016) considered 14 countries that used “school report cards” to inform the general public about school performance. No one model emerges from the review that would work for everyone but the concept of an annual school report card or record of achievement has the potential for providing parents and the local community with a broader and more balanced account of a school’s achievements. Indeed, the act of accounting could encourage an ethical, principled and professional stance by schools. For example, the school report card could recognize the achievements of pupils who do not often show up positively in tests or league tables or who arrive in school with a more challenging journey ahead of them, for example refugees. In giving an account of its work, the school would need to capture the voices of children, parents and employees. This sort of report to parents and community could prove a valuable part of the school’s professional accountability system.

At whole school level, self-evaluation is also a tool to help redress the balance between summative and formative accountabilitys. If it is to strengthen the professionalism of teachers, it has to be undertaken as a collaborative exercise including collective decision making. Evaluation should flow from the school’s vision and values, capturing its moral and professional accountabilitys, but also its contractual and market responsibilities. Good leaders establish a culture of supportive review, of disciplined enquiry leading to professional learning and action for improvement. This is what sustains formative accountability and, ultimately, produces good outcomes to evidence quality.

Increasingly, good leaders know that if this culture is to be driven by the best knowledge and understanding, lateral interaction beyond the school is vital. They are keen to seek professional excellence by extending knowledge and thinking to improve practice. This resonates with Mourshed et al. (2010: 81) who report leaders' identification of three key professional changes brought about by collaborative practice:

- Teaching practice is made public and the entire teaching profession shares responsibility for student learning
- There is a cultural shift from an emphasis on what teachers teach to what students learn
- A normative model of pedagogy develops, with teachers the custodians of that model, much as "doctors, lawyers and accountants establish norms of good practice and hold each other accountable for following them"

The use of professionals beyond the school to strengthen self-evaluation supports collaborative practice and, if followed through with sufficient rigor, could support the purposes of both internal and external accountability. Indeed, it has the potential for significant change in the way many external inspection systems currently work.

As Matthews and Ehren (2021) reference, there is evidence that both self-evaluation and peer evaluation provide powerful support for improvement as well as for meeting the demands of any external accountability system. The assumption is that reviewers will bring a trusted and fresh professional perspective, rooted in school-based evidence, that stimulates, challenges and supports. One of the uses of peer evaluation is as a means of validating self-evaluation.

Collins and Coleman (2017: 17) argue that incorporating evidence as central to decision-making will have a significant impact on pupil outcomes. They define evidence as "knowledge generated by external research that is transparent and rigorous." This has considerable value but I would go further and argue that a disciplined and collaborative approach to school evaluation, particularly involving colleagues external to the school, yields good evidence for decision-making and builds professional accountability.

Given the more recent growth of peer evaluation, there is less extensive research about its impact than there is about self-evaluation. The research that has taken place indicates that the process can lead to powerful professional development and mutual benefits for reviewers and their schools (Matthews and Headon, 2015; Matthews and Ehren, 2021). At its best, peer evaluation is formative lateral accountability in action and has the potential for enriching professional accountability.

In England, the use of various forms of peer review has been growing for several years with one study showing that 48% of school leaders reported that their school had been engaged in peer review in 2017 (Greany and Higham, 2018). It has emerged more slowly in other countries but peer review practices have been identified in Belgium, the Czech Republic, Finland and Sweden (OECD, 2013). Godfrey's study of peer review (2020) describes further instances of peer review in Australia, Bulgaria and Chile.

Godfrey's extensive analyses of different models characterizes peer review as either formative ("subjectivist") or summative ("objectivist"). He argues that the latter is more common and is sometimes used to prepare schools for an inspection. Nevertheless, he sees potential in further research and development of the model of peer review. Subjectivist models, he argues, focus more on deepening and sharing knowledge and learning about practice. Certainly, a model of collaborative enquiry, Research-informed Peer Review (RIPR), which Godfrey (2021) describes, has considerable potential for strengthening professional accountability through collaboration between leaders from schools and academic researchers. In RIPR, leaders and researchers identify a shared enquiry theme, which has impact on student outcomes. Working in clusters, they review the academic literature and evidence, then work through a cycle of iterative school review visits supported by workshops to develop practice. This approach is similar to action research and participatory evaluation.

Locality partnerships as support for professional accountability

In England, the growth of locality or place-based partnerships over recent years represents a response to the challenges resulting from an increase in the number of autonomous schools and the reduced power of local authorities to ensure a fair and coherent local education system (Gilbert, 2021). This situation is not unique to England; in many other countries, districts or local authorities have lost the power to coordinate local school systems.

Schools choose to join these partnerships because they share an overriding commitment to place and community. They see the partnership as a way beyond their school of helping them bring about the difference they want to make to the lives of children and young people.

In most instances, local authorities, who would previously have held considerable responsibilities for school performance and improvement, have driven the establishment of these partnerships. They are using them as engines of professional development, particularly to address local issues of quality and equity, which cannot easily be tackled by autonomous schools working alone. They commission a range of improvement services from the partnerships. This foundation of democratic and local accountability stemming from the role of the local authority but blended with the professional and moral accountabilities of schools, underpins the development of place-based partnerships. It gives powerful legitimacy to their work. It also gives them a strong platform to support the development of professional accountability within and across schools.

All partnerships have a strong focus on how professional learning can be shared or developed across schools. They have brokered and organized activities such as school-to-school support, peer review, teacher networks reflecting different communities of professional interest, practice improvement work, continuing professional development and action research. For many partnerships, this work involves using practitioners from schools as key source of expertise and creativity. By being involved in these sorts of initiatives, leaders and teachers start to feel professional and moral accountability for the children beyond their own schools.

Teachers in many areas have spoken with enthusiasm about the opportunities afforded by co-designing improvement initiatives created by their partnerships (Gilbert, 2017). They enjoy the creativity of working collaboratively to establish better or even innovative practice and they are developing their own knowledge and skills. When teachers work together to develop practice, not only are they benefiting professionally, they are also beginning to feel pride in the system by contributing to the greater good (Fullan and Boyle, 2014: 121). It is this sort of collaboration that is likely to build individual and collective capacity to sustain professional accountability.

At their best, place-based partnerships in England provide leadership for the strategic oversight of education in their area. They are bringing autonomous schools together, moving knowledge and learning around the system, increasing coherence and reducing the risks of fragmentation and isolation. Schools themselves take responsibility for their own collective performance, supporting each other and helping to improve outcomes for children and young people. This is lateral shared accountability, embedded in collective moral purpose and professional networks. Partnerships can help strengthen a model of accountability, professionally owned by teachers and schools. This can motivate, inspire teachers and incentivize system-wide collaboration to improve learning, thereby building greater professional capital and strengthening accountability.

Inspection as support for professional accountability

Governments need public accountability systems that ensure public money is being well spent on high quality education. Generally, they would expect systems to identify strengths as well as the improvements needed. Greater emphasis on schools developing more inclusive and rigorous professional accountability systems could support both internal and external accountability better. Inspection has a key role in accounting to government and the public about the quality of education but it could also have a stronger part to play in supporting schools to develop their approach to professional accountability. This would be particularly important in systems, such as England, which continue to operate systems of high stakes accountability.

Ehren (2019) argues for using inspection to help schools develop a culture of professional learning, within which schools are supported to become more enquiring and research informed. She presents a radically different model of inspection in which, “evaluation criteria are designed around how schools engage in research and use research evidence to inform their practice” (2019: 52) rather than a reliance on an inspection framework. Evaluations would be grounded in the needs of the school and its stakeholders. Certainly, this alternative model has considerable potential for supporting all aspects of school evaluation and professional accountability.

In England, if Ofsted were to take a different approach to the inspection of those schools that had undertaken a strong self-evaluation process, as described earlier, change would be dramatic. Ofsted might test the rigor of the school’s annual, summative self-evaluation and the professional accountability processes underpinning it. If inspectors found these to be strong, they would give formal and public endorsement of the school’s judgments and evaluation of itself. However, in undertaking scrutiny of the self-evaluation, if inspectors had concerns about judgments or process, they would extend their time in the school to carry out a full inspection and produce a full inspection report. For the best schools in the system, those with strong professional accountability embedded in their work, Ofsted would therefore be adopting a transparent, more cost-effective, quality assurance role which would not necessarily entail inspection.

At the same time, this process would enable Ofsted to identify samples of schools that might be usefully selected for in-depth studies of particular issues or aspects of education of interest to the government or indeed, the public. Examples of such studies might include current issues such as the impact of Covid-19 on children’s learning or the use of Artificial Intelligence for student learning in schools. In-depth reports of this nature have the potential not only to strengthen learning in schools by sharing professional capital across the system but also to inform government about where investment is needed in the system.

The British Educational Research Association (Moss et al., 2021) proposed reform of England’s high stakes accountability system that would also improve the quality of pupils’ and teachers’ experiences. Having surveyed other countries use of national sampling, their report points to the success of the National Education Panels (NEPs) in Germany. Data and research are used together to inform federal policy and practice. Moss et al. suggest replacing tests and school league tables with a nationally representative sample of pupils who would be tested over time on a broad range of skills and competences. The system would also involve gathering a range of in-depth, contextual information on schools and their communities using school, parent and pupil surveys. This combination of data could then be used to inform investment and other aspects of system improvement. Summary reports based on the national sample would be accessible to the public. Schools would be able to use the national data provided by the sample to support their self-evaluation as well as learning more widely from the summary reports. The report proposes too that inspection processes should be redesigned to support schools’ use of the research evidence and also peer learning. If enacted, all these proposals would have considerable potential for strengthening professional accountability.

Conclusion

The impact of external accountability practices on both equity and educational quality is hard to establish. Greater focus on using it to support both is crucial. This requires policymakers to design national accountability policy and practices that not only provide transparent information about education but also steer and support schools in their work. External accountability should be judged on how well it supports and improves schools’ systems for professional accountability as well as the quality and impact of its

information to government and the public. This would foster an emphasis on encouraging self-improvement and help build the professional capital of leaders and teachers across the system. The public accountability system could continue to prove quality but give greater support to improving it too.

However, real change will only occur when accountability has been reclaimed by schools and is seen as absolutely central to their work and teachers' professionalism. This means leaders and teachers taking pride in their role as agents of professional accountability and working collaboratively within and across schools to strengthen it. This culture of collective responsibility, trust and high expectations for performance and learning, underpinned by shared values and practices, has the potential for transforming systems. A strong approach to professional accountability within and across schools should feed effortlessly—and without as much pressure as evidenced currently—into public accountability frameworks. Indeed, the two approaches should become mutually reinforcing thereby strengthening learning, development and performance in all parts of the system.

References

- Altricher, H., Kemethofer, K., 2015. Does accountability pressure through school inspections promote school improvement? School effectiveness and school improvement. *Int. J. Res. Pol. Pract.* 26 (1), 32–56.
- Australian Primary Principals Association (APPA), 2010. The reporting and use of NAPLAN. Hindmarsh, S. Australia: APPA. In: Klenowski, V., Wyatt-Smith, C. (Eds.), *The Impact of High Stakes Testing: The Australian Story*, Assessment in Education: Principles, Policy & Practice, vol. 19, pp. 65–79.
- Bew Review, 2011. Independent Review of Key Stage 2 Testing, Assessment and Accountability. DfE, London.
- Brill, F., Grayson, H., Kuhn, L., O'Donnell, S., 2018. What Impact Does Accountability Have on Curriculum, Standards and Engagement in Education? A literature Review. NFER, Slough.
- Brundett, M., Rhodes, C., 2011. *Leadership for Quality and Accountability in Education*. Routledge, London & New York.
- Burgess, S., Wilson, D., Worth, J., 2013. A natural experiment in school accountability: the impact of school performance information on pupil progress. *J. Publ. Econ.* 106, 57–67.
- Cheng, X., Moses, K., 2016. *Promoting Transparency Through Information: A Global Review of School Report Cards*. IIEP-UNESCO, Paris.
- Collins, K., Coleman, R., 2017. Evidence-informed policy and practice ahead. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *Successful School Leadership: Linking With Learning*. Open University, Maidenhead.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2).
- Earley, P., Weindling, D., 2004. *Understanding School Leadership*. Paul Chapman Publishing, London.
- Ehren, M. (Ed.), 2016. *Methods and Modalities of School Inspections*. Springer, Switzerland.
- Ehren, M., 2019. Accountability structures that support school self-evaluation, enquiry and learning. In: Godfrey, D., Brown, C. (Eds.), *An Ecosystem for Research-Engaged Schools: Reforming Education Through Research*. Routledge, London & New York.
- Elliott, J., Bridges, D., Ebbutt, D., Gibson, R., Nias, J., 1981. *School Accountability: The SSRC Cambridge Accountability Project*. Grant McIntyre Limited, London.
- Elmore, R., 2004. *School Reform From the Inside Out: Policy, Practice and Performance*. Harvard Education Press, Cambridge, MA.
- Elmore, R., 2008. Leadership as the practice of improvement. In: Pont, B., Nusche, D., Hopkins, D. (Eds.), *Improving School Leadership, Case Studies on System Leadership*, vol. 2. OECD, Paris.
- Eraut, M., 1994. *Developing Professional Knowledge and Competence*. Falmer Press, London.
- Finnigan, K., Daly, A., 2012. Exploring the space between social networks, trust and urban school district leaders. *J. Sch. Leader.* 22, 493–530.
- Fullan, M., Boyle, A., 2014. *Big-City School Reforms: Lessons From New York, Toronto and London*. Teachers College Press, New York.
- General Teaching Council (GTC), 2009. *Accountability in Teaching: Two Messages From Two Research Studies*. General Teaching Council for England, London.
- Gilbert, C., 2012. *Towards a Self-Improving System: The Role of School Accountability*. NCSL, Nottingham.
- Gilbert, C., 2017. *Optimism of the Will: The Development of Local Area-Based Education Partnerships*. LCCL, London.
- Gilbert, C., 2021. Place-based leadership in autonomous school systems. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London.
- Godfrey, D., 2020. *School Peer Review for Educational Improvement and Accountability: Theory, Practice and Policy Implications*. Springer, Dordrecht.
- Godfrey, D., 2021. The growth of school peer review, its characteristics and the way ahead. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London.
- Goodman, R., Burton, D., 2012. What is the nature of the achievement gap, why does it persist and are government goals sufficient to create social justice in the education system? *Education* 3–13 (40), 500–514.
- Greany, T., Earley, P., 2017. In: Earley, P., Greany, T. (Eds.), *School Leadership and Education System Reform*. Bloomsbury, London, pp. 1–13.
- Greany, T., Higham, R., 2018. *Hierarchy, Markets and Networks: Analysing the "Self-Improving School-Led System" Agenda in England and the Implications for Schools*. UCL IOE Press, London.
- Gu, Q., Hodgen, J., Adkins, M., Armstrong, P., 2019. *Incentivising Schools to Take Up Evidence-Based Practice to Improve Teaching and Learning: Evidence From the Evaluation of the Suffolk Challenge Fund*. Education Endowment Foundation & UCL, London.
- Hattie, J., 2009. *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge, London and New York.
- Higham, R., Hopkins, D., Matthews, P., 2009. *System Leadership in Practice*. Open University Press & McGraw-Hill, Buckinghamshire.
- Hopkins, D., 2007. *Every School a Great School*. Open University Press: McGraw-Hill, Buckinghamshire.
- Kuramoto, N., Koizumi, R., 2016. Current issues in large-scale educational assessment in Japan: focus on national assessment of academic ability and university entrance examinations. *Assess. Educ. Princ. Pol. Pract.* <https://doi.org/10.1080/0969594X.2016.1225667> [online].
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven Strong Claims About Successful School Leadership*. National College for School Leadership, Nottingham, UK.
- Leithwood, K., Harris, A., Hopkins, D., 2019. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (4), 1–18.
- Louis, K., Leithwood, K., Wahlstrom, K., Anderson, S., 2010. *Investigating the Links to Improved Student Learning: Final Report of Research Findings*. Carei, Minnesota.
- Lynch, S., Worth, J., Bamford, S., Wespieser, K., 2016. *Engaging Teachers: NFER Analysis of Teacher Retention*. NFER, Slough.
- Matthews, P., Ehren, M., 2021. Accountability and improvement in self-improving school systems. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London.
- Matthews, P., Headon, M., 2015. *Multiple Gains—an Independent Evaluation of Challenge Partners Peer Reviews of Schools*. IOE Press, London.
- Matthews, P., 2014. *Professional Accountability—Redesigning Schooling 7*. SSAT, London.
- Mongon, D., Chapman, C., 2012. *High Level Leadership*. Routledge, London & New York.

- Moss, G., Goldstein, H., Hayes, S., Chereau, B.M., Sammons, P., Sinnott, G., Stobart, G., 2021. High Standards, Not High Stakes: An Alternative to SATs That Will Transform England's Testing & School Accountability System in Primary Education & Beyond. BERA, London.
- Mourshed, M., Chijioke, C., Barber, M., 2010. How the World's Most Improved Systems Keep Getting Better. McKinsey & Company, London.
- Nakayasu, C., 2016. School curriculum in Japan. *Curric. J.* 27 (1), 134–150.
- Nelson, R., Ehren, M., Godfrey, D., 2015. Literature Review on Internal Evaluation. Institute of Education, London.
- OECD, 2012. Lessons from PISA for Japan: Strong Performers and Successful Reformers in Education. OECD, Paris.
- OECD, 2013. Synergies for Better Learning: An International Perspective on Evaluation and Assessment. OECD, Pisa, Paris.
- OECD, 2014. Improving Schools in Wales: An OECD Perspective. OECD, Paris.
- OECD, 2016. What Makes a School a Learning Organization: A Guide for Policymakers, School Leaders and Teachers. OECD, Paris.
- Schleicher, A., 2018. World Class: How to Build a 21st-Century School System. OECD, Paris.
- Wilkins, C., 2011. Professionalism and the post-performative teacher: new teachers reflect on autonomy and accountability in the English school system. *Prof. Dev. Educ.* 37 (3), 389–409.

Principals' work in public schools

Katina Pollock, University of Western Ontario, Critical Policy, Equity and Leadership Studies, Faculty of Education, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	304
Role	304
Time on tasks	305
Work	305
What do we know about principals' work?	305
Leading curriculum	305
Managing the budget	306
Dealing with personnel	306
Maintaining the school site	306
Building and sustaining partnerships	306
Processing paperwork and forms	307
Changing nature of principals' work	307
Advances in information communication technology and social media	307
Equity, diversity, inclusion and decolonization (EDID)	307
Policies, programs and accountability initiatives	308
Student wellness	308
COVID-19 pandemic and occupational health and safety guidelines	308
Work intensification	308
Working long hours	308
Increased volume of similar work	309
Engaging in new work demands	309
Increased pace of work	309
Perceived limited resources	309
No downtime	310
Why this is important to know	310
Conclusion	311
References	311

Introduction

This article explores what principals do using the lens of *work*. It not only considers what principals do but also why this work is changing, and the consequences of this changing work on principals themselves, such as the impact of work intensification. Scholars in educational administration and leadership have explored what principals do using several lenses: (a) role, (b) time on task, and (c) work. This article is organized into several sections. The first section presents some of the existing literature on the aforementioned lenses. The second section discusses what is known about principals' work, including (a) what they do and (b) the changing nature of their work in contemporary times. The third section discusses why knowing about principals' work is important.

Role

Getzels and Guba (1957) were the first scholars to prominently consider the concept of *role* in organizational systems. Since then, scholars have employed the concept of role in a number of ways (e.g., Goodwin et al., 2003; Hadjiithoma-Garstka, 2011; Mitchell and Sackney, 2006; Whitaker, 2003). The problem with these studies, however, is that many of them tend to focus on a particular policy/issue or approach role through a narrow lens rather than recognizing that principals have multiple roles. A more serious shortcoming of using role to understand what principals do is that it is "unrealistic to assume, as most role theorists do, that the position that a person occupies will dictate what he or she does" (Ryan, 2007, p. 344). Examining what principals do solely through the lens of role erroneously assumes that principals exclusively engage in the actions associated with their assigned roles. Doing so fails to acknowledge that the work of principals may fall outside of their official roles, be shared with others, or be influenced by the unique contexts in which they work.

Time on tasks

Another lens explores what principals do using the concept of time spent on tasks. Time-on-task studies have existed over several decades (Lee, 2021). Although the majority of these studies avoid the issues of the role lens, they nevertheless have their own shortcomings. Perhaps the most cited shortcoming for some of the seminal studies is the lack of attention given in time-on-task studies to the nature of the task itself (Gronn, 1982). Specifically, time-on-tasks studies in the past have generally failed to interrogate the meaning of the tasks or why administrators are engaged in them. For example, although they may meticulously document the frequency and duration of telephone conversations, few inquired into the nature of the conversations (e.g., negotiating with a school board employee, talking with a parent, seeking advice from a colleague, etc.). Other efforts have been made to try to interrogate the meaning of tasks and why administrators engage in them; methodological approaches have included the use of structured observations and reflective interviews (Martinko and Gardner, 1990; Pollock and Hauseman, 2016) as well as self-report activity logs (Goldring et al., 2008; May and Supovitz, 2011). However, these methodological approaches have limitations as well. Rowan (2021) presents a list of assumptions that exist with this research, from conceptual, to scale construction, to generalizability. Lastly, there is an effort to move away from the use of clock time to help determine what principals do and why to consider a different theoretical approach instead, where rather than use clock time as the unit of analysis, to flip this notion on its head and analyze what it is that principals actually do in their work regardless of how much time it takes for the task to be completed (Eacott, 2018).

Work

The notion of work used in this article acknowledges the wide-ranging, diverse, and complex nature of what principals do. Work is defined as labor or effort expended to achieve a particular set of goals (Merriam Webster, n.d.). Work can be both paid and unpaid. It includes employment-related unpaid work but not other kinds of unpaid work such as volunteerism, household activities, or family responsibilities (Drago, 2007). This labor may be expended both within *and outside* position-related roles that principals enact. It is difficult in this day and age to erect clear boundaries around work efforts, just as it is to define organizational boundaries (Ryan, 1996). Work can take place on and off the school site. It can occur after the official opening and closing of the school day. Work also comprises particular experiential components such as physical, mental, and emotional aspects (Applebaum, 1992; Gamst, 1995). This article employs a lens that acknowledges the behavioral, cognitive, *and* emotional elements of work. In some existing research, the term workload is used interchangeably with *work* (See Oplatka, 2017).

What do we know about principals' work?

Interest in principals' work has existed for centuries (Ensign, 1923; Pierce, 1934; Rousmaniere, 2013). Scholars interested in this area have focused on topics such as:

- the growth of school systems in the 18th and 19th centuries and beyond (Kafka, 2009),
- the duties of principals (Davis, 1921; Bolton and Howard, 1925),
- the professionalization of school leadership (UNESCO, 2016),
- the roles of the principalship including the ongoing debate around leadership and management (Spillane, 2009),
- school improvement (Harris et al., 2013),
- recruitment and retention (Dite, 2020; Jackson, 2018),
- leadership frameworks (Gümüç et al., 2018; Leithwood et al., 2017),
- accountability (Andreyko, 2010),
- job satisfaction and motivation (Wang et al., 2018a,b),
- work and labor such as workload (Oplatka, 2017),
- intensification of work (Riley, 2014, 2019; Walker, 2021), and
- wellness (Richardson, 2020; Riley, 2019; Wang et al., 2018a,b).

Historically, time use studies have divided principals' work into four dimensions: (a) leadership, management, and administration; (b) external relations; (c) internal relations, and (d) principals' own professional development (Lee, 2021). However, these categories are not necessarily discrete, as managing both internal and external relations often falls under the umbrella of leading and managing, for example. Given that it is impossible to comprehensively cover all four dimensions within a single article, I collapsed these four dimensions into two foci: (a) what contemporary principals' work entails, and (b) the changing nature of their work. These foci were selected because they provided a way to structure a manageable overview of the work that principals do across these four categories and how this work influences their wellness.

Leading curriculum

There are numerous purposes of K–12 education (Gutek, 2014). Regardless of which purposes are espoused, they all include the co-creation of understanding, knowledge, and skills with children who attend these K–12 education systems. Depending on the

geographical context, school principals are referred to as educational/educative leaders (Duignan and MacPherson, 1992) leaders of instruction (Hallinger, 2005), curriculum leaders (Glatthorn et al., 2018), pedagogical leaders (Macneil et al., 2005), leaders of learning (Hallinger, 2011), and leaders of student learning (Robinson, 2011). Their overall work is to lead, support, and facilitate at their work site (Pollock and Hauseman, 2015).

Managing the budget

Schools have always had budgets and leaders of schools have always had some kind of interaction with these budgets, from merely knowing what the fiscal amount is and where the funds are going to site-based management with full discretion on how to spend the annual budget (Pollock and Hauseman, 2015). However, it appears that with the onset of reform efforts and school improvement initiatives, “budgeting” has become more complex than previously experienced. In 2013, the Organisation for Economic Co-operation and Development (OECD) Teaching and Learning International Survey (TALIS) report ranked “inadequate school budget and resources” as the top barrier to principal effectiveness; in Alberta, Canada it was ranked second (OECD, 2013). This lack of financial resources is becoming a global reality (Caribbean Policy Research Institute [CaPRI], 2014; DeAngelis et al., 2018; Lupton and Thrupp, 2013; Mahoney, 2013; Martin et al., 2018; Semuels, 2016; OECD, 2012, 2017; World Bank, 2019; Zimmer, 2020). It is not only the lack of resources that is changing but also the demand to find creative ways to either secure these resources (OECD, 2017; Walker, 2021) or manage with less. In Ontario, a recent study indicated that nearly one-third of the 70 principals interviewed required additional professional learning to manage the school’s budget or deal with other financial aspects of the position (Pollock et al., 2015). There is also growing interest in the work of fundraising for school principals and budgetary allocation (Poole and Fallon, 2015; Winton and Milani, 2017; Walker, 2021). For example, concerns around budgetary allocation have also heightened during the COVID-19 pandemic as educational policymakers and leaders face budgetary cuts (Al-Samarrai et al., 2020; Mitra, 2020) at a time when additional costs are incurred to not only keep students safe but also to address growing inequities (Allen et al., 2020; Daniela et al., 2021; Friedman et al., 2021; Sider, 2020).

Dealing with personnel

Managing personnel has always been a component of principals’ work (Norton, 2008). Whom they manage and for what purposes also slightly vary depending on the organizational structure of the public schooling. For example, some school principals who work in schools that engage in site-based management have more decision-making power about who works at that site (Patrinios and Fasih, 2009; Rose, 2007; Whitaker, 2003; White, 1989; Wohlstetter and Mohrman, 1993), whereas school principals who find themselves in school systems that are very bureaucratic and centralized have limited control over hiring practices (Hanson, 1976; Maynes and Hatt, 2015; Oplatka, 2004; UNESCO, 2016). Some argue that managing personnel and leading teachers cannot be separated; although they are closely connected, the focus here is on the technical aspect of working with people that requires principals to be knowledgeable about labor contracts and possess skills such as conflict resolution, the ability to supervise multiple employee groups, and to address grievances (Castro, 2020; Engel and Curran, 2016; Matubatuba, 2008; Sorenson and Goldsmith, 2008).

Maintaining the school site

Principals are responsible for maintaining school sites. It is, however, not an easy task. It can also take on many forms. In some jurisdictions, principals with limited financial freedom are only responsible for reporting building issues and providing information about buildings when asked by their system managers. Although this might seem to be a simple reporting procedure, the work can contribute to the growing paperwork and procedural work that school principals find themselves doing. Other principals with more decision-making power working in older schools are faced with difficult decisions about allocating resources; they must self-educate on what needs to be prioritized, and must oversee any renovations and/or repairs (Díaz-Vicario and Sallán, 2020; Federal Emergency Management Agency, 2017; Nhalpo, 2020; Preiss, 2014). However, it is not just older buildings that require principals’ time and attention—new school development projects can be just as taxing on a school principal’s time. New school buildings require an effective maintenance plan that includes meeting warranty deadlines and constant revisiting as the building ages (Chan, 2000). And yet, in decentralized schools in some regions, principals may find themselves literally having to source supplies, purchase them, and in some cases actually engage in the skilled labor work themselves. For example, in Jamaica, some school principals actually fundraise, purchase the lumber for additions, and then build the additions themselves (Walker, 2021). In more recent times, school principals now have the added layer of having to create safe learning and teaching spaces during the pandemic, where expectations of social distancing influence class size, structure of the school day, and so forth (World Health Organization, 2020).

Building and sustaining partnerships

Principals are often engaged in partnerships. In this article, a partnership is understood as a relationship with an outside organization or group (as opposed to relationships with individual parents or families). Many advocate that partnerships can be a way to support student achievement (Hands, 2010), and can be considered a part of instructional leadership. However, I have separated this category from instructional leadership as it does not facilitate principals’ direct involvement in curriculum instruction in their

daily work. Schools are being encouraged to engage in outside partnerships for all sorts of reasons: to generate revenue, such as by renting out building space after school hours (Clandfield, 2010); to foster closer connections to services that the school cannot effectively provide, such as supplemental educational services (Koyama, 2011); or as a way of engaging ethnic, racial, and religious communities with the school and with student learning (Auerbach, 2010; Rogers et al., 2012). Effective partnerships require trusting relationships, a central purpose, and leadership to name but a few components (Best and Holmes, 2010). For many principals, engaging in partnerships at the school level requires engaging in partnership-building and maintenance tasks that can at times take them away from direct school functions. Notably, during the COVID-19 pandemic there have been calls for schools and school systems to work more closely with outside agencies, especially those in public health as concerns over student mental health and wellness have increased (Hertz and Barrios, 2021).

Processing paperwork and forms

Paperwork has always been a part of principals' work (Georgiades and Jones, 1989; Wang et al., 2018a,b; Williams, 1979). Before the recent advances in information communication technology (ICT), paperwork would have arrived on a principal's desk through the general post or through intercourier district mail bags dropped off at specific times at the school site, then eventually through fax correspondence in the office; now, principals are likely filling out PDF forms received via email either by smartphone, tablet, or laptop or through an interactive web platform where they electronically respond to requests and provide information to those requesting it. Paperwork can be understood as the medium for accountability mechanisms that principals are required to respond to on a daily basis. The majority of paperwork that contemporary principals respond to is generated from the accountability mechanisms entrenched in jurisdictional policies that they are to follow (Fitzgerald, 2009).

Changing nature of principals' work

Principals' work is always changing and both researchers and practitioners have documented these changes over the centuries (Alvoid and Black, 2014; Bredeson, 1989; Darling-Hammond et al., 2007; Day et al., 2000; Hallinger, 1992; Huber, 2004; Klein and Ninio, 2019; Pollock et al., 2015). The nature of principals' work continues to evolve in response to local and global contexts. Long before the COVID-19 pandemic, principals were responding to several global changes. A few are reported here, such as working with new advances in ICT; evolving understandings of equity, diversity, inclusion, and decolonization; program, policies and accountability initiatives; student wellness; and COVID-19 and occupational health and safety guidelines.

Advances in information communication technology and social media

Because school principals spend a considerable amount of time communicating, they have always capitalized on the various modes of communication available to them beyond face-to-face interaction. In the past, this could have involved writing handwritten letters, making telephone calls, and sending faxes. With advances in ICT and now social media, principals are faced with an entire new suite of communication media. The Internet, email, texting, smartphones/mobile technology, and social media platforms and programs have profoundly impacted the work environments of a number of professions. The education sector, specifically the school principalship, has been and continues to be highly affected (Anderson and Dexter, 2005; Cho, 2016; Cho and Jimerson, 2017; Davies, 2010; Fancera, 2020; Fullan, 2014; Liu et al., 2018; Omar and Ismail, 2020; Pollock and Hauseman, 2018; Sheninger, 2014; Uğur and Koç, 2019). Many scholars have purported that technology makes work easier, increases efficiency, reduces paperwork, and promotes work-life balance (Fullan, 2014; Sheninger, 2014). My own research has indicated that the use of email can lead to fast and efficient communication with various stakeholders, assist school principals in managing their workload by allowing them to complete tasks faster, while emails leave an accountability trail through recording communication and an accurate trail of tasks and activities (Pollock and Hauseman, 2018). But there is also a flipside to the use of ICT as well. Emerging evidence has demonstrated that how school principals manage the changes resulting from increased use of email in the workplace is not uniformly positive. Principals also report high volumes of email, longer workdays, increased workload, increased expectations of shorter response times, and the blurring of boundaries between work and home—all of which contribute to the intensification of work (Pollock and Hauseman, 2018).

Equity, diversity, inclusion and decolonization (EDID)

Systemic inequity exists globally. These inequities—for example, social, racial, religious, socioeconomic, gender-based—prevent some groups of students from receiving the same quality education as students from other groups. Over time there has been a growing movement to reach and support *all* groups of students. Although not universal, most democratic countries are striving for this goal through policy changes and program modifications (OECD, 2020, 2012; UNESCO, 2015), embracing diversity with the populations they serve and the workforces they employ (Hewins-Maroney and Williams, 2013; Pollock and Briscoe, 2019), engaging in inclusive and equitable practices (DeMatthews et al., 2021; Gümüş et al., 2021; Ryan, 2014), and working

toward decolonizing the systems they work within (Hohepa, 2013; Munroe et al., 2013; Stewart and Warn, 2017). Efforts to enact EDID often require school principals to challenge their own biases, change their mindsets around their assumptions and beliefs, and engage in their work in a different manner.

Policies, programs and accountability initiatives

Principals' work is influenced by policies. Reform movements and accountability initiatives have used policies as tools to initiate and support educational change. For this article, policy is understood as a way to change practice through official formal legislation, mandates, government initiatives, and board regulations. Policies can act to promote, regulate, constrain, and/or monitor actions and outcomes. Recent studies in Canada indicate that policies are playing a more significant role in principals' work than in the past. Many school principals feel that their role is constrained by the number of initiatives imposed on a school (Alberta Teachers' Association and Canadian Association of Principals, 2014; Blakesley, 2012). Not only do principals in Ontario feel constrained by formal policy initiatives, but multiple policy initiatives within short time periods have also influenced how they do their work (Pinto, 2015; Wang et al., 2018a,b). The OECD Teaching and Learning International Survey (TALIS) report ranked "government regulation and policy" as third on its list of barriers to principal effectiveness (OECD, 2013).

Student wellness

Principals spend time promoting student wellness. In North America, student wellness has become a focus in public education. Some practitioners have indicated that they deal with increasing incidences of issues relating to the mental health and well-being of students and staff (Pollock, 2016). There appears to be no consensus on why these incidents are increasing (Olson et al., 2015; Whitaker, 2010). Some argue that the number of incidences is not increasing, but rather it is awareness of mental health issues that has increased (Whitaker, 2010). Principals are not ignoring the health and well-being of students and teachers. For this reason, they often find themselves seeking out additional community services and resources for students and families, engaging in informal counseling and information-sharing sessions, and providing support for teacher professional learning—activities in which they would not have been involved in the past.

COVID-19 pandemic and occupational health and safety guidelines

Principals play a pivotal role in creating a healthy school community. This work has been especially critical during the COVID-19 pandemic. Principals are legally responsible for this through various labor laws, public health laws, and education laws. Many of these responsibilities are connected to employee and public safety through labor laws or through public health and education legislation that focuses on students' well-being, health, and safety. These responsibilities and tasks have always existed and vary between jurisdictions but generally include being legally responsible for the health and safety of all students within the school. For example, in some jurisdictions principals are required to implement school-wide antibullying or antiviolen policies and plans. Other jurisdictions expect principals to connect with representatives from other government agencies such as public health in an effort to better meet student needs or provide student support services. Some principals are also explicitly expected to report any child welfare concerns to their superintendent and appropriate government officials (Pollock and Hauseman, 2015). Last but not least, principals have been charged with the responsibility to notify the appropriate authorities of any outbreaks of infectious or contagious diseases. It is precisely this last responsibility that has recently influenced the nature of public school principals' lives. Because of the pandemic, educators, including school principals, have been subjected to drastic changes in how public education has been delivered. All jurisdictions around the world have been impacted by the pandemic. The majority of school systems have engaged in a series of practices from full school system shutdown to full online learning to various hybrid versions of learning (Netolicky, 2020; Pont, 2020).

Work intensification

It is one thing to document the ever-changing nature of principals' work; it is another to also explore the influence of these changes on school principals. Some researchers explore the impact of student performance and system success, others consider the impact on the health and wellness of individuals in the school principal role and yet others consider the labor impact of school principals. The intensification of work can be found in several work sectors such as engineering, nursing, and medicine (Brannon, 1994; Willis et al., 2015; Wong et al., 2019). Principals also experience work intensification (Alberta Teachers' Association and Canadian Principals Association, 2014; Armstrong, 2015; Leithwood and Azah, 2014). The remainder of the article argues that the changing nature of principals' work is leading to work intensification.

Working long hours

Principals work long hours. There is a global trend of long work hours culture (Chatzitheochari and Arber, 2009; Kodz et al., 1998; Pencavel, 2018; Waheed, 2010). Principal workforces are no different. Time-on-task studies provide various quotes on hours worked per week. For example, in 2019, in a regular work week in Ontario, principals worked an average of 57.3 h. Almost all participating principals (97.5%) worked more than the normal 40 h work week; almost half of the participants reported

(47.3%) working between 50 and 60 h per week. In addition, 21.4% worked more than 60 h per week. This is approximately 16 h per week more than Canadian occupational managers (Statistics Canada, 2020) who work 40.9 h per week. The trend toward long work hours is not isolated to Canada, however. In Australia, it appears that school principals spend on average 51–60 h per week during regular school time (Riley et al., 2020). It would be a statistical feat to compare these time-on-task studies as not all studies are exact replicas of each other (Rowan, 2021). However, cumulative statistical findings and qualitative studies indicate that most school principals around the world work long hours. Although working long hours can reap benefits and/or pose disadvantages, most employed people experience periods of long work hours. Working long hours may be considered part of intense work situations, but by itself it does not constitute work intensification, especially given that we do not know whether principals have always worked these long hours. However, working long hours can be an intermediary factor that, combined with other factors, can lead to work intensification.

Increased volume of similar work

Work volume is often referred to as the amount of work or number of tasks that an employee is to do within a specified period of time. Recent studies have discovered that school principals are reporting an increased volume of similar work. Specifically, they describe doing more of the same types of work. One example is paperwork. The principalship has always required some amount of paperwork; over the past decade, principals have reported engaging in increased amounts of paperwork (Alberta Teachers' Association and Canadian Principals Association, 2014; Hancock et al., 2019). An increase in work volume means more work, which may not pose issues for employees if they are allowed to delegate other work tasks or are able to work additional hours until the additional work is completed. However, in most situations, the volume increases without any ability for principals to delegate other work and the expected time to complete this additional work has not changed but remained the same. For those in the principalship who experience work intensification, they work longer hours to attend to both an increased volume of work and to additional different kinds of work.

Engaging in new work demands

As mentioned previously, the work of principals is dictated by the local context and global trends. In many cases, this means that how they go about their work has changed. For example, principals have always been responsible for communicating with parents. How they go about doing that today has changed. In the past, principals may have met parents in face-to-face meetings; they most likely still do meet face-to-face, but they also communicate through paper notices, electronic newsletter, and by email. The end goal is the same but the mode and medium are quite different. In some cases, this also means that the very nature of their work has changed. Over the past decade, many countries have experienced a concentration on student mental health and wellness. As mentioned earlier, there has been a shift in supporting student mental health and wellness. This concern has also extended to teacher mental health and wellness, especially given that there is increasing evidence that teacher wellness may influence student wellness and school success (Harding et al., 2019). This means that principals are engaged in additional support-seeking for not just student wellness but also teacher wellness, something that in the past was not always a consideration. Moreover, since 2020, the COVID-19 pandemic has seen yet again additional work for school principals such as extending the notion of care to preventing school breakouts, containing breakouts, working toward continued student learning through physical school closures and multiple modes of curriculum delivery, and concentrating on being a technology leader (Osmond-Johnson et al., 2020; Pollock, 2020).

New work demands do not directly lead to work intensification. In many organizations, when an employee or team is provided with new tasks, other work is reassigned, put on pause, or additional staff are hired; however, in the public education system, this is not the case. School principals often find themselves continuing to juggle the increased work that they already have with the new expectations and tasks assigned to them while little is removed from their workload. Many principals are also finding that the pace of their work has also increased.

Increased pace of work

Principals must deal with an increased pace of work. This pace of work is sometimes determined by policies and procedures that stimulate timelines and deadlines; other times it is because of bureaucratic decisions made elsewhere. Principals may not experience work intensification through working long hours with additional work and new tasks if the expectations for completion are flexible and there can be extensions. In reality, however, not only is the time of return less flexible, principals have indicated over the past decade that the expectation for completion or responses to requests has shortened, which increases the pace of their work. The increased pace of work coupled with increased work and additional roles can result in principals finding their work challenging at times. Working in these conditions when there is constant pressure over perceived or limited resources contributes to work intensification.

Perceived limited resources

In some jurisdictions, there are limited resources and financing for public schools to successfully function (Walker, 2021); in other jurisdictions, there may be plenty of funding for public schooling but that funding might be ear-marked for specific initiatives and can only be spent in particular ways (Pollock et al., 2015). This is an issue for principals and can contribute to work intensification. For example, in other sectors, departments and organizations often have control over their budget and can hire more staff or professionals to share work volume and additional tasks to meet the quicker turn-around times. In public education, some principals may

be fortunate enough to reassign some of their work to other administrative staff such as the vice-principal and/or other administrative assistants but in some cases this then creates other consequences such as the downloading of work to vice-principals or assistant principals (Leithwood and Azah, 2014; Lim and Pollock, 2019; Pollock et al., 2017; Swain, 2021). In some site-based management arrangements, a principal may be able to hire additional administrative support but these situations are uncommon. Many principals not only have little control over their budget in terms of hiring staff but many also have found themselves with less administrative and staff support over the past few decades (Beusaert et al., 2016; Friedman, 2002). Limited resources alone does not mean work intensification; however, being asked to do more under increased pressure and pace with less increases the chances that a school principal will experience work intensification. The key however is whether principals working in this environment have the opportunity for any downtime.

No downtime

Principals have little downtime. Of all the key components of work intensification, downtime is the most influential. The notion of downtime was originally used in relation to innate networks and machinery systems in production plants during the mid-1900s. When a piece of machinery was out of service, either planned or unplanned, the production process was experiencing downtime (Online Etymology Dictionary n.d.). The use of the concept was not applied to social and production practices of individuals until the early 1980s; in these cases, downtime was an employee's time away from paid work or for individuals it can mean time away from paid work where the employee experiences a state of physical relaxation and psychological detachment from paid work (Dugan and Barnes-Farrell, 2017).

People who work in fast-paced environments with incredible amounts of work to be accomplished within a short time frame by working long hours with limited resources or work accommodations may experience work intensification; however, if these intense periods of work are followed by periods where employees or workers are rewarded with additional time off and/or the work environment converts to a slower pace with reduced workload expectations and reduced hours of work, then these employees are not necessarily experiencing work intensification. Work patterns where employees can recharge, regroup, and recover from the intense work are not experiencing work intensification. They may have experienced a period of intensified work but work intensification is the ongoing intense work environment that has no downtime. Downtime, or more accurately, little to no downtime, is usually the deciding factor on whether or not a workplace is experiencing work intensification. No downtime usually leads to all sorts of unintended consequences such as lost days of work for the employer, burnout, and other occupational wellness issues for the employee (Giurge et al., 2020; Lei et al., 2019; Shattell, 2018).

Some skeptics of work intensification in the public education workforce argue that educators and principals receive downtime on weekends, statutory holidays, and school system summer holidays. Although structurally this can appear to be the case, emerging time-on-task data and qualitative studies indicate that these scheduled downtimes are in reality the times during which many principals work additional hours to catch up on unfinished work or work that demands uninterrupted concentration on specific tasks. A recent Australian study, for example, reported school principals spend 25–30 h working per week during holiday periods (Riley et al., 2020). In many jurisdictions, school principals are responsible for continued professional learning, yet this expected learning is rarely scheduled in a principal's work day and therefore is fulfilled during time periods that would be considered downtime.

Why this is important to know

Work intensification, if left unchecked, can lead to all sorts of consequences for those individuals experiencing it. Work intensification places greater demand on workers' cognitive, psychological, and physical abilities (Fournier et al., 2011); it is associated with numerous psychological hazards such as "chronic stress, anxiety, depression, psychosomatic disorders, psychological decompensation mechanisms, work addiction (workaholism) and burnout" (Legault and Belarbi-Basbous, 2006, p. 26). A 2010 report by the World Health Organization highlighted these types of psychosocial hazards as an emerging trend in occupational health and safety: "Work overload would appear to constitute one of the main psychosocial risk factors leading to psychological distress at work" (Fournier et al., 2011, p. 4). Some of the consequences of work intensification include poor employee performance, increased job stress, low morale, burnout, poor occupational mental health and wellness, negative repercussions for leisure, and personal and familial issues (Burchielli et al., 2006; Caudroit et al., 2011; Canadian Center for Occupational Health and Safety [CCOHS], 2012).

Principals in public education appear to be exhibiting signs of unwellness. Concerns regarding stress, burnout, coping strategies, workload, and the general health and wellness of school principals and administrators have proliferated since the 1970s. According to Poirel et al. (2012), between 1966 and 1988, more than 1300 articles were published concerning principals and stress. Many of these studies point directly to school principals' work and workloads as directly associated with their health and wellness concerns (Whitaker, 1995). Presently, much academic research and practice has focused on the well-being of students and teachers, with limited focus on school principals (Hendriks and Scheerens, 2013; Pollock et al., 2015).

Some parents, policymakers, and researchers might ask why principal wellness is such an issue in the public education sector—especially when a body of research clearly demonstrates that in schools, teachers have the most influence on student learning (Leithwood and Louis, 2012). Even though teachers have direct influence on student learning and there are more teachers than principals, principals play an important part in student and school success and system change as well. It has been established from

decades of research that principals indirectly affect student success (Hallinger and Heck, 1998; Leithwood et al., 2004); what is less commonly known is that principals also directly impact student success as well (Lee, 2021). Emerging research indicates that school principals do interact with students more than originally thought and that these interactions are significant both in terms of time on task but also impactful for students and their school success (Lee et al., 2021; Silva et al., 2011). How are students expected to receive quality education from school principals when many are struggling with their wellness and work intensification?

Principals play a key role in supporting students and teachers; they are critical for school and school system success. However, work intensification must be addressed for three main reasons: (a) work intensification is contributing to principal unwellness (Earley, 2020; Heffernan and Pierpoint, 2020; Wang et al., 2018a,b; Whitaker, 2003; Xiao and Newton, 2020); (b) unwell principals cannot support student success as effectively as well principals (Whitaker, 2003; Wang et al., 2018a,b); and (c) the job is becoming less desirable and there are growing recruitment and retention issues for the principal workforce in many countries as a result (Earley, 2020; Heffernan and Pierpoint, 2020; Whitaker, 2003; Wang et al., 2018a,b; Xiao and Newton, 2020).

Conclusion

This article explored what principals do using a work lens. It provided examples of some of the work that school principals do and a snapshot of some of the local and external influences on the nature of principals' work. The article concluded that many principals around the world find themselves experiencing work intensification. Because principals are the middle managers of public education school systems (Flessa, 2012; Kaul et al., 2021), the work they do will always be crucial. They champion new initiatives and system changes, are key supporters of students and teachers and pivotal change agents. It is crucial that politicians, policymakers, administrators, educators, and parents find ways to support principals in the important work they do.

References

- Alberta Teachers' Association [ATA], Canadian Association of Principals, 2014. The Future of the Principalship in Canada. Alberta Teachers Association, Edmonton, AB.
- Allen, J., Mahamed, F., Williams, K., 2020. Disparities in education: e-learning and COVID-19, who matters? *Child. Youth Serv.* 41 (3), 208–210. <https://doi.org/10.1080/0145935X.2020.1834125>.
- Al-Samarrai, S., Gangwar, M., Gala, P., 2020. The Impact of the COVID-19 Pandemic on Education Financing. World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/33739>.
- Alvold, L., Black Jr., W.L., 2014. The Changing Role of the Principal: How High Achieving Districts Are Recalibrating School Leadership. Center for American Progress. <https://www.americanprogress.org/issues/education-k-12/reports/2014/07/01/93015/the-changing-role-of-the-principal/>.
- Anderson, R.E., Dexter, S., 2005. School technology leadership: an empirical investigation of prevalence and effect. *Educ. Adm. Q.* 41 (1), 49–82.
- Andreyko, T.A., 2010. Principal Leadership in the Accountability Era: Influence of Expanding Job Responsibilities on Functional Work Performance, Stress Management, and Overall Job Satisfaction. Unpublished doctoral dissertation. University of Pittsburgh.
- Applebaum, H.A., 1992. The Concept of Work: Ancient, Medieval, and Modern. SUNY Press.
- Armstrong, D., 2015. Transition to the Role of Principal and Vice-Principal: Final Report. Institute for Educational Leadership, Toronto, ON.
- Auerbach, S., 2010. Beyond coffee with the principal: toward leadership for authentic school-family partnerships. *J. Sch. Leader.* 20, 728–757.
- Beausaert, S., Froehlich, D.E., Devos, C., Riley, P., 2016. Effects of support on stress and burnout in school principals. *Educ. Res.* 58 (4), 347–365.
- Best, A., Holmes, B., 2010. Systems thinking, knowledge and action: towards better models and methods. *Evid. Policy* 6, 145–159.
- Blakesley, S., 2012. Juggling educational ends: non-Indigenous Yukon principals and the policy challenges that they face. *Int. J. Educ. Policy Leadersh.* 7, 1–14.
- Bolton, E.B., Howard, C., 1925. Duties of the elementary-school principal. *Peabody J. Educ.* 3 (3), 139–145. <https://doi.org/10.1080/01619562509534701>.
- Brannon, R.L., 1994. Professionalization and work intensification: nursing in the cost containment era. *Work Occup.* 21 (2), 157–178.
- Bredeson, P.V., 1989. Redefining leadership and the roles of school principals: response to changes in the professional work life of teachers. *High Sch. J.* 73 (1), 9–20.
- Burchielli, R., Pearson, P., Thanacoody, P.R., 2006. Exploring work intensification in teaching: a research agenda. *N. Z. J. Employ. Relat.* 31, 71–82.
- Canadian Centre for Occupational Health and Safety (CCOHS), 2007. Fatigue: The Foe You Don't Want to Know at Work. <https://www.ccohs.ca/newsletters/hsreport/issues/2007/11/ezine.html>.
- Caribbean Policy Research Institute (CaPRI), 2014. Jamaica: Failing the Education Test, vol. 5. Think Tank. https://issuu.com/capricaribbean/docs/capri_issue3.
- Castro, A.J., 2020. Tensions between what principals know and do: the role of labor market conditions on teacher recruitment and hiring. *Leader. Pol. Sch.* 1–24.
- Caudroit, J., Boiché, J., Stephan, Y., Scanff, C., Trouilloud, D., 2011. Predictors of work/family interference and leisure-time physical activity among teachers: the role of passion towards work. *Eur. J. Work. Organ. Psychol.* 20 (3), 326–344. <https://doi.org/10.1080/13594320903507124>.
- Chan, T.C., 2000. Beyond the status quo: creating a school maintenance program. *Princ. Leadersh.* 3, 64–67.
- Chatzitheochari, S., Arber, S., 2009. Lack of sleep, work and the long hours culture: evidence from the UK Time Use Survey. *Work. Employ. Soc.* 23 (1), 30–48.
- Cho, V., Jimerson, J.B., 2017. Managing digital identity on twitter: the case of school administrators. *Educ. Manag. Adm. Leader* 45 (5), 884–900.
- Cho, V., 2016. Administrators' professional learning via twitter. *J. Educ. Adm.* 54 (3), 340–356.
- Clandfield, D., 2010. The school as a community hub: a public alternative to the neoliberal threat to Ontario schools. *Our Sch. Our Selves* 19, 5–74.
- Daniela, L., Rubene, Z., Rudolfa, A., 2021. Parents' perspectives on remote learning in the pandemic context. *Sustainability* 13 (7). <https://doi.org/10.3390/su13073640>.
- Darling-Hammond, L., LaPointe, M., Meyerson, D., Orr, M.T., 2007. Preparing School Leaders for a Changing World: Lessons From Exemplary Leadership Development Programs. Stanford Educational Leadership Institute, Stanford.
- Davies, P.M., 2010. On school educational technology leadership. *Manag. Educ.* 24 (2), 55–61.
- Davis, C.O., 1921. The duties of high-school principals. *Sch. Rev.* 29 (5), 337–350.
- Day, C., Harris, A., Hadfield, M., Tolley, H., Beresford, J., 2000. Leading Schools in Times of Change. Open University Press, Buckingham.
- DeAngelis, C.A., Wolf, P.J., Maloney, L.D., May, J.F., 2018. Charter School Funding: More Inequity in the System. University of Arkansas, Fayetteville. <http://www.uaedreform.org/downloads/2018/11/charter-school-funding-more-inequity-in-the-city.pdf>.
- DeMatthews, D.E., Serafini, A., Watson, T.N., 2021. Leading inclusive schools: principal perceptions, practices, and challenges to meaningful change. *Educ. Adm. Q.* 57 (1), 3–48.
- Díaz-Vicario, A., Sallán, J., 2020. A comprehensive approach to management school safety: case studies in Catalonia Spain. *Educ. Res.* 59 (1), 89–106. <https://doi.org/10.1080/00131881.2016.1272430>.

- Dite, E.A., 2020. Recruitment and Retention of Catholic School Principals: The Importance of School Leadership in the Success of Catholic Schools. Doctoral Dissertation. Seton Hall University.
- Drago, R.W., 2007. Striking a Balance: Work, Family, Life. Dollars & Sense.
- Dugan, A.G., Barnes-Farrell, J.L., 2017. Time for self-care: downtime recovery as a buffer of work and home/family time pressures. *J. Occup. Environ. Med.* 59, 46–56.
- Duignan, P.A., MacPherson, R.J., 1992. *Educative Leadership: A Practical Theory for New Administrators and Managers*. Falmer Press, Bristol.
- Eacott, S., 2018. Theoretical notes on a relational approach to principals' time use. *J. Educ. Adm. Hist.* 50 (4), 284–298.
- Earley, P., 2020. Surviving, thriving and reviving in leadership: the personal and professional development needs of educational leaders. *Manag. Educ.* 34 (3), 117–121.
- Engel, M., Curran, F.C., 2016. Toward understanding principals' hiring practices. *J. Educ. Adm.* 54 (2), 173–190.
- Ensign, F.C., 1923. Evolution of the high-school principalship. *Sch. Rev.* 31 (3), 179–190.
- Fancera, S.F., 2020. School leadership for professional development: the role of social media and networks. *Prof. Dev. Educ.* 46 (4), 664–676. <https://doi.org/10.1080/19415257.2019.1629615>.
- Federal Emergency Management Agency, 2017. Safer, Stronger, Smarter: A Guide to Improving School Natural Hazard Safety. https://preparecenter.org/sites/default/files/natural_hazards_school_safety.pdf.
- Fitzgerald, T., 2009. The tyranny of bureaucracy: continuing challenges of leading and managing from the middle. *Educ. Manag. Adm. Leader* 37 (1), 51–65. <https://doi.org/10.1177/1741143208098164>.
- Flessa, J.J., 2012. Principals as middle managers: school leadership during the implementation of primary class size reduction policy in Ontario. *Leader. Pol. Sch.* 11 (3), 325–343. <https://doi.org/10.1080/15700763.2012.692429>.
- Fournier, P.-S., Montreuil, S., Brun, J.-P., Bilodeau, C., Villa, J., 2011. Exploratory Study to Identify Workload Factors that Have an Impact on Health and Safety: A Case Study in the Service Sector (Report R-107). Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSTT). <http://irsst.qc.ca/media/documents/PubIRSTT/R-701.pdf>.
- Friedman, J., York, H., Mokdad, A.H., Gakidou, E., 2021. U.S. children "learning online" during COVID-19 without the internet or a computer: visualizing the gradient by race/ethnicity and parental educational attainment. *Socius* 7. <https://doi.org/10.1177/2378023121992607>.
- Friedman, I.A., 2002. Burnout in school principals: role related antecedents. *Soc. Psychol. Educ.* 5 (3), 229–251.
- Fullan, M., 2014. *The Principal: Three Keys to Maximizing Impact*. Jossey-Bass, San Francisco.
- Gamst, F.C., 1995. *Meanings of Work: Considerations for the Twenty-First Century*. SUNY Press.
- Georgiades, W.D.H., Jones, H.L., 1989. A review of research on headmaster and school principalship in developing countries. In: Education and Employment Division in the Population and Human Resources Department of the World Bank, Document No. PHREE/89/11, PHREE Background Paper Series.
- Getzels, J.W., Guba, E.G., 1957. Social behavior and the administrative process. *Sch. Rev.* 65 (4), 423–441.
- Giurge, L.M., Whillans, A.V., West, C., 2020. Why time poverty matters for individuals, organizations and nations. *Nat. Hum. Behav.* 4 (10), 993–1003.
- Glatthorn, A.A., Boschee, F., Whitehead, B.M., Boschee, B.F., 2018. *Curriculum Leadership: Strategies for Development and Implementation*, fifth ed. SAGE, Thousand Oaks.
- Golding, E., Huff, J., May, H., Camburn, E., 2008. School context and individual characteristics: what influences principal practice? *J. Educ. Adm.* 46 (3), 332–352.
- Goodwin, R.H., Cunningham, M.L., Childress, R., 2003. The changing role of the secondary principal. *NASSP Bull.* 87 (634), 26–42.
- Gronn, P.C., 1982. Neo-Taylorism in educational administration? *Educ. Adm. Q.* 18 (4), 17–35.
- Gümüş, S., Bellibas, M.S., Esen, M., Gümüş, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48.
- Gümüş, S., Arar, K., Oplatka, I., 2021. Review of international research on school leadership for social justice, equity and diversity. *J. Educ. Adm. Hist.* 53 (1), 81–99.
- Gutek, G.L. (Ed.), 2014. *Philosophical, Ideological, and Theoretical Perspectives on Education*, second ed. Pearson, Boston.
- Hadjithoma-Garstka, C., 2011. The role of the principal's leadership style in the implementation of ICT policy. *Br. J. Educ. Technol.* 42 (2), 311–326.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9 (2), 157–191.
- Hallinger, P., 1992. Changing norms of principal leadership in the United States. *J. Educ. Adm.* 30 (3), 35–48.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leader. Pol. Sch.* 4 (3), 221–239.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142.
- Hancock, D.R., Müller, U., Stricker, T., Wang, C., Lee, S., Hachen, J., 2019. Causes of stress and strategies for managing stress among German and US principals. *Res. Comp. Int. Educ.* 14 (2), 201–214.
- Hands, C., 2010. Why collaborate? The differing reasons for secondary school educators' establishment of school-community partnerships. *Sch. Effect. Sch. Improv.* 21 (2), 189–207.
- Hanson, M., 1976. School Governance and the Professional/Bureaucratic Interface: A Case Study of Educational Decision-Making. Paper presented at the Annual Meeting of the American Educational Research Association. <https://files.eric.ed.gov/fulltext/ED125124.pdf>.
- Harding, S., Morris, R., Gunnell, D., Ford, T., Hollingworth, W., Tilling, K., Evans, R., Bell, S., Grey, J., Brockman, R., Campbell, R., Araya, R., Murphy, S., Kidger, J., 2019. Is teachers' mental health and wellbeing associated with students' mental health and wellbeing? *J. Affect. Disord.* 242 (1), 180–187. <https://doi.org/10.1016/j.jad.2018.08.080>.
- Harris, A., Day, C., Hopkins, D., Hadfield, M., Hargreaves, A., Chapman, C., 2013. *Effective Leadership for School Improvement*. Routledge, London.
- Heffernan, A., Pierpoint, A., 2020. *Autonomy, Accountability, and Principals' Work: An Australian Study*. Final report. Monash University. https://researchmgt.monash.edu/ws/portalfiles/portal/324416419/314450604_oa.pdf.
- Hendriks, M.A., Scheerens, J., 2013. School leadership effects revisited: a review of empirical studies guided by indirect-effect models. *Sch. Leader. Manag.* 33, 373–394.
- Hertz, M.F., Barrios, L.C., 2021. Adolescent mental health, COVID-19, and the value of school–community partnerships. *Inj. Prev.* 27 (1), 85–86. <https://doi.org/10.1136/injuryprev-2020-044050>.
- Hewins-Maroney, B., Williams, E., 2013. The role of public administrators in responding to changing workforce demographics: global challenges to preparing a diverse workforce. *Publ. Adm. Q.* 37 (3), 456–490.
- Ngāpuhi Hohepa, M.K., 2013. Educational leadership and indigeneity: doing things the same, differently. *Am. J. Educ.* 119 (4), 617–631.
- Huber, S.G., 2004. *Preparing School Leaders for the 21st Century*. RoutledgeFalmer, London.
- Jackson, S., 2018. *The Principal Pipeline: A Qualitative Exploration of Strategies for the Recruitment and Retention of Black Male Public School Principals*. Doctoral dissertation. St. John Fisher College.
- Kafka, J., 2009. The principalship in historical perspective. *Peabody J. Educ.* 84 (3), 318–330.
- Kaul, M., Comstock, M., Simon, N., 2021. Leading From the Middle: How Principals Rely on District Guidance and Organizational Conditions in Times of Crisis. CPRE. Working Papers. https://repository.upenn.edu/cpre_workingpapers/26.
- Klein, J., Ninio, R., 2019. School principals' key experiences and changing management patterns. *Int. J. Educ. Manag.* 33 (5), 903–918.
- Kodz, J., Kersley, B., Strebler, M.T., O'Regan, S., 1998. Breaking the Long Hours Culture. IES, Brighton. Institute for Employment Studies Report 352.
- Koyama, J., 2011. Principals, power, and policy: enacting "supplemental educational services". *Anthropol. Educ. Q.* 42, 20–36.
- Lee, M., Walker, A., Riordan, G., 2021. Principals' direct interaction with individual students. In: Lee, M., Pollock, K., Tulowitzki, P. (Eds.), *How School Principals Use Their Time: Implications for School Improvement, Administration, and Leadership*. Routledge, pp. 75–92.
- Lee, M., 2021. Principal time use as a research area: notes on theoretical perspectives, leadership domains, and future directions. In: Lee, M., Pollock, K., Tulowitzki, P. (Eds.), *How School Principals Use Their Time*. Routledge, Abingdon, pp. 9–21.
- Legault, M.-J., Belarbi-Basbous, H., 2006. Gestion par projets et risques pour la santé psychologique au travail dans la nouvelle économie. *Pistes* 8 (1).
- Lei, X., Kaplan, S.A., Dye, C.E., Wong, C.M., 2019. On the subjective experience and correlates of downtime at work: a mixed-method examination. *J. Organ. Behav.* 40, 360–381.

- Leithwood, K., Azah, V.N., 2014. Executive Summary: Elementary and Secondary Principals' and Vice Principals' Workload Studies. Catholic Principals' Council of Ontario, Toronto, ON.
- Leithwood, K., Louis, K.S., 2012. Linking Leadership to Student Learning. Jossey-Bass, San Francisco.
- Leithwood, K., Louis, K., Wahlstrom, K., Anderson, S., 2004. How Leadership Influences Student Learning: A Review of Research for the Learning From Leadership Project. Wallace Foundation, New York.
- Leithwood, K., Sun, J., Pollock, K. (Eds.), 2017. How School Leaders Contribute to Student Success. Springer Nature, Cham.
- Lim, L., Pollock, K., 2019. "...the workload has proliferated, absolutely proliferated": secondary principals' perspectives on the impact of work intensification on the secondary vice-principal role. *Lead. Manag.* 25 (2), 80–98.
- Liu, C., Ready, C., Roman, A., Van Wart, M., Wang, Q., McCarthy, A., Kim, S., 2018. E-leadership: an empirical study of organizational leaders' virtual communication adoption. *Leader. Organ. Dev. J.* 39 (7), 826–843. <https://doi.org/10.1108/LODJ-10-2017-0297>.
- Lupton, R., Thrupp, M., 2013. Headteachers' readings of and responses to disadvantaged contexts: evidence from English primary schools. *Br. Educ. Res. J.* 39 (4), 769–788.
- Macneill, N., Cavanagh, R.F., Silcox, S., 2005. Pedagogic leadership: refocusing on learning and teaching. *Int. Electron. J. Leader. Learn* 9 (2), 2.
- Mahoney, M.H., 2013. Inequity in American schools: a new perspective on distributional effects of school expenditures on economic well-being. In: *Review of Income and Wealth Series*, vol. 59, pp. 728–755, 4.
- Martin, C., Boser, U., Benner, M., Baffour, P., 2018. A Quality Approach to School Funding: Lessons Learned From School Finance Litigation. Center for American Progress. <https://www.americanprogress.org/issues/education-k-12/reports/2018/11/13/460397/quality-approach-school-funding/>.
- Martinko, M.J., Gardner, W.L., 1990. Structured observation of managerial work: a replication and synthesis. *J. Manag. Stud.* 27 (3), 329–357.
- Matubutuba, H.S., 2008. An Evaluation of the Outcomes of the Training and Developing School Principals in Labor Relations During the Ten Years of Democracy. Doctoral Dissertation. North-West University, South Africa.
- May, H., Supovitz, J.A., 2011. The scope of principal efforts to improve instruction. *Educ. Adm. Q.* 47 (2), 332–352.
- Maynes, N., Hatt, B.E., 2015. The Complexity of Hiring, Supporting, and Retaining New Teachers Across Canada. Canadian Association for Teacher Education (CATE).
- Merriam-Webster, (n.d.). Downtime. <https://www.merriam-webster.com/dictionary/downtime>.
- Mitchell, C., Sackney, L., 2006. Building schools, building people: the school principal's role in leading a learning community. *J. Sch. Leader.* 16 (5), 627–640.
- Mitra, R., 2020. COVID-19 is killing education budgets: are educational public-private partnerships an answer? *J. Prof. Cap. Community* 5 (3/4), 255–264. <https://doi.org/10.1108/JPC-07-2020-0056>.
- Munroe, E.A., Borden, L.L., Orr, A.M., Toney, D., Meader, J., 2013. Decolonizing aboriginal education in the 21st century. *McGill J. Educ.* 48 (2), 317–338.
- Netolicky, D., 2020. School leadership during a pandemic: navigating tensions. *J. Prof. Cap. Community* 5 (3/4), 391–395. <https://doi.org/10.1108/JPC-05-2020-0017>.
- Nhalpo, V.A., 2020. The leadership role of the principal in fostering sustainable maintenance of school facilities. *S. Afr. J. Educ.* 40 (2), 1–9. <https://doi.org/10.15700/saje.v40n2a1790>.
- Norton, M.S., 2008. Human Resources Administration for Educational Leaders. SAGE Publications.
- Olfson, M., Druss, B.G., Marcus, S.C., 2015. Trends in mental health care among children and adolescents. *N. Engl. J. Med.* 372, 2029–2038.
- Omar, M.N., Ismail, S.N., 2020. Mobile technology integration in the 2020s: the impact of technology leadership in the Malaysian context. *Univers. J. Educ. Res.* 8 (5), 1874–1883. <https://doi.org/10.13189/ujer.2020.080524>.
- Oplatka, I., 2004. The principalship in developing countries: context, characteristics and reality. *Comp. Educ.* 40 (3), 427–448.
- Oplatka, I., 2. Principal workload: components, determinants and coping strategies in an era of standardization and accountability. *J. Educ. Adm.* 55 (5), 552–568.
- Organization of Economic Co-operation and Development (OECD), 2013. Education at a Glance 2013: OECD Indicators. OECD Publishing, Paris. <https://doi.org/10.1787/eag-2013-en>.
- Organization of Economic Co-operation and Development (OECD), 2017. The Funding of School Education: Connecting Resources and Learning. OECD Publishing, Paris.
- Organization of Economic Co-operation and Development (OECD), 2020. Strength Through Diversity: Education for Inclusive Societies. OECD Publishing, Paris. <https://www.oecd.org/education/strength-through-diversity/Design-and-Implementation-Plan.pdf>.
- Organization of Economic Co-operation and Development (OECD), 2012. Equity and Quality in Education: Supporting Disadvantaged Students and Schools. OECD Publishing, Paris. <https://www.oecd.org/education/school/50293148.pdf>.
- Osmond-Johnson, P., Campbell, C., Pollock, K., 2020. Moving Forward in the COVID-19 Era: Reflections for Canadian Education. EdCan Magazine. <https://www.edcan.ca/articles/moving-forward-in-the-covid-19-era/>.
- Patrinos, H.A., Fasih, T., 2009. Decentralized Decision-Making in Schools: The Theory and Evidence on School-Based Management. World Bank Publications, Washington DC.
- Pencavel, J.H., 2018. Diminishing Returns at Work: The Consequences of Long Working Hours. Oxford University Press, New York.
- Pierce, P.R., 1934. The Origin and Development of the Public-School Principalship. Doctoral Dissertation. The University of Chicago.
- Pinto, L.E., 2015. Fear and loathing in neoliberalism: school administrator responses to policy layers. *J. Educ. Adm. Hist.* 46, 140–154.
- Poiré, E., Lapointe, P., Yvon, F., 2012. Coping with administrative constraints by Quebec school principals. *Can. J. Sch. Psychol.* 27 (4), 302–318.
- Pollock, K., Briscoe, P., 2019. School principals' understandings of difference and diversity and how these understandings influence their work. *Int. J. Educ. Manag.* 34 (3), 518–534. <https://doi.org/10.1108/IJEM-07-2019-0243>.
- Pollock, K., Hauseman, D.C., 2015. School leadership in Canada. In: Årlestig, H., Day, C., Johansson, O. (Eds.), *Cross Country Histories of School Leadership Research: Focus and Findings*. Springer, Dordrecht, pp. 202–232.
- Pollock, K., Hauseman, D.C., 2016. Observational research on school principals: to time or not or not to time. In: Ling, L., Ling, P. (Eds.), *Paradigms and Methods in Educational Research*. IGI Global, Hershey, PA, pp. 88–107. <https://doi.org/10.4018/978-1-5225-1738-2.ch006>.
- Pollock, K., Hauseman, D.C., 2018. The use of e-mail and principals' work: a double-edged sword. *Leader. Pol. Sch.* 18 (3), 382–393. <https://doi.org/10.1080/15700763.2017.1398338>.
- Pollock, K., Murakami, E., Swapp, D.H., 2015. The work of school leaders: north American similarities, local differences. *Int. Stud. Educ. Adm.* 43 (2), 5–20. http://www.cceam.org/fileadmin/user_upload/SEA_MEM/SEA43.2.pdf.
- Pollock, K., Wang, F., Hauseman, D.C., 2017. The Changing Nature of Vice Principals' Work: Final Report for the Ontario Principals' Council. Ontario Principals' Council, Toronto.
- Pollock, K., 2016. Principals' work in Ontario, Canada: changing demographics, advancements in informational communication technology and health and well-being. *J. Commonw. Council Educ. Adm. Manag.* 44 (3), 55–74.
- Pollock, K., 2020. School leaders' work during the COVID-19 pandemic: a two-pronged approach. *Int. Stud. Educ. Adm.* 48 (3), 38–45.
- Pont, B., 2020. Education Leadership in Times of Uncertainty: Rising to the Challenge. Qatar Foundation. <https://www.wise-qatar.org/education-leadership-in-times-of-uncertainty-rising-to-the-challenge/>.
- Poole, W., Fallon, G., 2015. The emerging fourth tier in K-12 education finance in British Columbia, Canada: increasing privatization and implications for social justice. *Glob. Soc. Educ.* 13 (3), 339–368. <https://doi.org/10.1080/14767724.2014.996857>.
- Preiss, B., 2014. State School Principals Face Impossible Choices on Building Maintenance. The Victoria Age. <https://www.theage.com.au/national/victoria/state-school-principals-face-impossible-choices-on-building-maintenance-20141015-116k7x.html>.
- Richardson, L.M., 2020. Principal Wellness: Principals' Managing Their Wellbeing. Doctoral Dissertation. The University of Nebraska-Lincoln.
- Riley, P., Sioau-Mae, S., Marsh, H., Dicke, T., 2020. The Australian Principals' Occupational Health, Safety, and Well-Being Survey (IPPE Report). Institute for Positive Psychology and Education, Australian Catholic University, Sydney.
- Riley, P., 2014. Australian Principal Occupational Health, Safety and Wellbeing Survey: 2011–2014 Data. Institute for Positive Psychology and Education (ACU).

- Riley, P., 2019. The Australian Principal Occupational Health, Safety and Wellbeing Survey: 2018 Data. Institute for Positive Psychology and Education (ACU).
- Robinson, V.M.J., 2011. *Student-Centered Leadership*. Jossey-Bass, San Francisco.
- Rogers, J., Freeton, R., Terriquez, V., 2012. Enlisting collective help: urban principals' encouragement of parent participation in school decision-making. In: Auerbach, S. (Ed.), *School Leadership for Authentic Family and Community Partnerships: Research Perspectives for Transforming Practice*. Routledge, New York City, pp. 55–57.
- Rose, G.C., 2007. The role of principal empowerment within a site-based management environment: empirical testing of a structural model. *J. Educ. Change* 8 (3), 207–233.
- Rousmaniere, K., 2013. The principal: the most misunderstood person in all of education. Atlantic. <https://www.theatlantic.com/education/archive/2013/11/the-principal-the-most-misunderstood-person-in-all-of-education/281223/>.
- Rowan, B., 2021. Measurement assumptions in principals' time-use research: past problems and needed studies. In: Lee, M., Pollock, K., Tulowitzki, P. (Eds.), *How School Principals Use Their Time*. Routledge, Abingdon, pp. 22–42.
- Ryan, J., 1996. Putting Humpty Dumpty together again: order, anxiety and systemic reform. *Int. J. Educ. Reform* 5 (4), 453–462.
- Ryan, J., 2007. Dialogue, identity and inclusion: administrators as mediators in diverse school contexts. *J. Sch. Leader.* 17 (3), 340–369.
- Ryan, J., 2014. Promoting inclusion in Ontario: principals' work in diverse settings. In: Bogotch, I., Shields, C.M. (Eds.), *International Handbook of Educational Leadership and Social (In)justice*. Springer Netherlands, Dordrecht, pp. 359–380.
- Samuels, A., 2016. Good schools, rich schools; bad schools, poor schools: the inequality at the heart of America's education system. Atlantic. <https://www.theatlantic.com/business/archive/2016/08/property-taxes-and-unequal-schools/497333/>.
- Shattell, M., 2018. Getting away from it all: the importance of vacation and downtime recovery from work. *J. Psychosoc. Nurs.* 56 (5), 1–4.
- Sheninger, E., 2014. *Digital Leadership: Changing Paradigms for Changing Times*. Corwin, Thousand Oaks.
- Sider, S.R., 2020. School principals and students with special education needs in a pandemic: emerging insights from Ontario, Canada. *Int. Stud. Educ. Adm.* 48 (2), 78–84.
- Silva, J.P., White, G.P., Yoshida, K.R., 2011. The direct effects of principal–student discussions on eighth grade students' gains in reading achievement: an experimental study. *Educ. Adm. Q.* 47 (5), 772–793.
- Sorenson, R.D., Goldsmith, L.M. (Eds.), 2008. *The Principal's Guide to Managing School Personnel*. Corwin, Thousand Oaks.
- Spillane, J.P., 2009. Managing to lead: reframing school leadership and management. *Phi Delta Kappan* 91 (3), 70–73.
- Statistics Canada, 2020. Usual Hours Worked by Occupational, Annual. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410029901>.
- Stewart, J., Warr, J., 2017. Between two worlds: indigenous leaders exercising influence and working across boundaries. *Aust. J. Publ. Adm.* 76 (1), 3–17. <https://doi.org/10.1111/1467-8500.12218>.
- Swain, G., 2021. Exploring the Principal and Vice Principal Relationship: More Vice than Principle? LinkedIn. <https://www.linkedin.com/pulse/exploring-principal-vice-relationship-more-than-principle-swain>.
- Uğur, N.G., Kog, T., 2019. Leading and teaching with technology: school principals' perspective. *Int. J. Educ. Leadersh. Manag.* 7 (1), 42–71. <https://doi.org/10.17583/ijelm.2018.3758>.
- UNESCO, 2015. Education 2030 Incheon Declaration and Framework for Action: Towards Inclusive and Equitable Quality Education and Lifelong Learning for All. http://uis.unesco.org/sites/default/files/documents/education-2030-incheon-framework-for-action-2016-en_0.pdf.
- UNESCO, 2016. Leading Better Learning: School Leadership and Quality in the Education 2030 Agenda. <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/pdf/leadership-report.pdf>.
- Waheed, Z., 2010. The long work hours culture: causes, consequences and choices. *Facilities* 28 (7/8), 395–396.
- Walker, A., 2021. *Jamaican Secondary School Principals' Occupational Mental Health and Well-Being*. Unpublished Doctoral Thesis. University of Western Ontario.
- Wang, F., Pollock, K.E., Hauseman, C., 2018a. School principals' job satisfaction: the effects of work intensification. *Can. J. Educ. Adm. Pol.* 185, 73–90.
- Wang, F., Pollock, K., Hauseman, D.C., 2018b. In: Cherkowski, S., Walker, K. (Eds.), *Ontario Principals' and Vice-Principals' Well-Being and Coping Strategies in the Context of Work Intensification*.
- Whitaker, K.S., 1995. Principal burnout: implications for professional development. *J. Pers. Eval. Educ.* 9, 287–296.
- Whitaker, K.S., 2003. Principal role changes and influence on principal recruitment and selection: an international perspective. *J. Educ. Adm.* 41 (1), 37–54.
- Whitaker, R., 2010. *Anatomy of an Epidemic: Magic Bullets, Psychiatric Drugs, and the Astonishing Rise of Mental Illness in America*, first ed. Crown Publishing Group, New York City.
- White, P.A., 1989. An overview of school-based management: what does the research say? *NASSP Bull.* 73 (518), 1–8.
- Williams, R.C., 1979. The impact of collective bargaining on the principal: what do we know? *Educ. Urban Soc.* 11 (2), 168–180.
- Willis, E., Henderson, J., Hamilton, P., Toffoli, L., Blackman, I., Couzner, L., Verrall, C., 2015. Work intensification as missed care. *Lab. Ind.* 18 (38), 1–19.
- Winton, S., Milani, M., 2017. Policy advocacy, inequity, and school fees and fundraising in Ontario, Canada. *Educ. Pol. Anal. Arch.* 25 (40), 1–33. <https://doi.org/10.14507/epaa.25.2679>.
- Wohlstetter, P., Mohrman, S.A., 1993. *School-Based Management: Strategies for Success*. Consortium for Policy Research in Education, Philadelphia.
- Wong, K., Chan, A.H.S., Ngan, S.C., 2019. The effect of long working hours and overtime on occupational health: a meta-analysis of evidence from 1998 to 2018. *Int. J. Environ. Res. Publ. Health* 16 (12), 1–22. <https://doi.org/10.3390/ijerph16122102>.
- World Bank, 2019. The Education Crisis. Being in School is Not the Same as Learning. <https://www.worldbank.org/en/news/immersive-story/2019/01/22/pass-or-fail-how-can-the-world-do-its-homework>.
- World Health Organization, 2020. Strategies for Protecting K-12 School Staff From COVID-19. <https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/ppcovidwho-2639?lang=en>.
- Xiao, J., Newton, P., 2020. School leadership challenges in Canada. In: Papa, R. (Ed.), *Oxford Research Encyclopedia of Education*. Bloomsbury, New York City.
- Zimmer, E., 2020. Australian Report Reveals Accelerating Social Inequality in Education. World Socialist Web Site. <https://www.wsws.org/en/articles/2020/08/31/educ-a31.html>.

Leadership contexts and how they matter: leaders' working conditions

Scott C. Bauer, University of Colorado Denver, School of Education & Human Development, Denver, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	315
Search parameters and opening considerations	316
Job satisfaction	316
Principal efficacy	317
Stress/burnout	318
Isolation	319
Discussion	320
References	321

Introduction

For well over half a century, scholars and practitioners alike have asserted the vital importance of school leadership to the success of schools and to the promotion of academic achievement for students. Initially, this claim seemed as much wishful thinking or the result of a chain of logical deduction as one supported by rigorous research. The more recent emergence of a considerable body of work based on well thought-out theoretical models which have been systematically examined empirically is an indicator of the field emerging from its adolescence. While it is unlikely researchers will discover a “unified theory” of effective school leadership, it is impossible to ignore the progress the field has made both in quantity and quality of attention to the questions: How does school leadership matter, and what do effective (impactful) school leaders do? (See, for example: Day et al., 2011, 2016; Gumus et al., 2018; Leithwood and Day, 2008; Leithwood and Mascal, 2008; Leithwood et al., 2010, 2017; Urlick and Bowers, 2014). Consequently, it is virtually obligatory in any scholarly work to reference the vital importance of school leadership to the success of schools worldwide—in fact, it is practically mandatory to mention that leadership is second only to quality teaching in importance among school-related factors that influence student achievement (Leithwood et al., 2008, 2017; Louis et al., 2010), a claim recently updated by Leithwood et al. (2020, p. 6) to read: “School leadership has a significant effect on features of the school organization which positively influences the quality of teaching and learning. While moderate in size, this leadership effect is vital to the success of most school improvement efforts.” This is not meant as a pejorative or sarcastic observation; quite the contrary, the progress scholars have made and the reasonableness and rigor of both evidence and resultant claims, especially when triangulated across settings and sources, is incredible.

An almost-equally compulsory requirement of scholarship on school principals, heads, or heads-of-school is to acknowledge that the role is horribly intricate and getting more so every year (Day, 2014; Horng et al., 2010; Huang et al., 2020; Oplatka, 2017; Robinson, 2011; Southworth, 2008; Wang et al., 2018). School principals’ work is described as “fragmented, relentless and stressful” (Bezzina et al., 2018, p. 842), increasingly diverse (Hauseman et al., 2017), and while leaders are continually admonished to focus on instructional leadership, evidence suggests that demands best described as bureaucratic and managerial tend to dominate their days (Grissom et al., 2013, 2015; Hauseman et al., 2017; Horng et al., 2010; Huang et al., 2020). In a recent analysis of a national sample of American middle school principals, Huang et al. (2020) characterized the principal’s day as unrelenting, hectic, largely reactive and unplanned; Earley and Bubb (2013) mention that just over 60% of headteachers in the UK characterize their work-life balance as poor. Their study showed that new headteachers felt challenged by isolation, external pressures, and a desire to distribute leadership without the requisite understanding of how to do this.

While leadership might be aptly described as varied and filled with ambiguities in any context, principal workload is extreme (Oplatka, 2017) and increasingly associated with negative outcomes such as burnout and turnover (Earley and Bubb, 2013; Hauseman et al., 2017; Maxwell and Riley, 2017; Türkoğlu and Cansoy, 2020; Wells, 2013; West et al., 2010). This leads to a vicious cycle—schools need high-quality leaders but the nature of the role and the toll it can take on incumbents results in potentially unsustainable turnover rates and fewer applicants for positions (Cierninski, 2018; Norton, 2002; Snodgrass Rangel, 2018; Teklelessie and Villarreal, 2011). This turnover is costly (Alenezi, 2020; Snodgrass Rangel, 2018), has an adverse impact on school performance and climate/culture (Mascal and Leithwood, 2010; Snodgrass Rangel, 2018), and particularly impacts the neediest schools (Alenezi, 2020; Beckett, 2018; Hansen, 2018) as well as charter schools (Ni et al., 2015; Sun and Ni, 2015).

Given the importance of school leadership to the mission and performance of schools, the challenging nature of the role, and the impact of consequences such as turnover, this paper focuses on the question: What does scholarship tell us about the quality of work life of school leaders? Specifically, the literature was probed to determine what is known about factors that are controllable by school and system leaders and policy makers that have been shown to have an impact on principal job satisfaction; efficacy; stress/burnout; and isolation. Taken together, these outcomes paint a picture of the quality of work life of school principals/heads, and understanding the malleable factors that could be used to enhance them may provide insight into how we might make the role more attractive and provide “space” for the continual improvement of school leaders.

Search parameters and opening considerations

To address the question posed above, a search was done across varied electronic databases for empirical research articles published in peer-reviewed journals, initially delimited to work published since 2010 and later extended back an additional ten years based on a review of papers first consulted in order to expand and triangulate the inquiry. The search was conducted using the keyword pairings depicted in [Table 1](#).

To further ensure that relevant empirical work was discovered, additional searches were done on ancillary keywords: "turnover/persistence/intention to leave," "workload," "social support," "working conditions," and "work environment/school climate/work setting."

Several initial considerations seem prudent to mention at the outset that may affect interpretation of the results, or at very least, are worth sharing with readers to augment trustworthiness. First, not surprisingly far more research has focused on teacher working conditions in schools rather than administrators or leaders. For the most part, this work was ignored—while conditions and many responsibilities between teachers and other school staff have some similarities, there are significant limits on the degree to which we might generalize from findings related to teachers and reasonably apply them to individuals who serve as principals, heads, or headteachers. Second, each of the factors investigated have a bipolar nature that complicates discussion. For example, while we will be seeking to understand what organizational factors might limit job stress, the absolute absence of stress is death and some degree of stress is absolutely essential to job performance. Similarly, while we hope to discover work-related factors that enhance job satisfaction, some degree of dissatisfaction is likely motivating and a benefit to individual leaders and the organization. Nevertheless, for simplicity, each was treated as unidimensional, e.g., that stress, burnout, isolation and dissatisfaction should be minimized.

Finally, all four of the constructs of interest in this paper are inherently subjective in nature and are most often operationalized using attitude surveys using well-established measures that have been subject to psychometric examination to establish validity and reliability. An issue that plagues investigation of all outcomes of interest related to emotions or quality of work life is that even though each has been widely measured, there are no universally accepted definitions and none have empirically-determined norms that would permit determining "levels" or a judgment of what "high" and "low" might be, even delimited by various contextual factors. Conceptually, even if norms existed, applying them to individuals or groups is complicated by the nature of the phenomena—individuals have varied experiences and abilities to cope with such things as stressors or experience with satisfiers and dissatisfiers. Consequently, studies tend to be relational in nature and seek to uncover constructs in the work world that could be manipulated to promote change in a direction rather than to discover an absolute level of "high" or "low." This suggests, though, that two basic kinds of leverage exist in relation to these factors: we might seek to lessen or minimize their incidence, or alternatively we can take steps to enhance individuals' ability to cope (or both).

Job satisfaction

Job satisfaction is an employee's favorable assessment of their job ([Chang et al., 2015](#); [De Nobile and McCormick, 2008](#)) and serves as a sort of omnibus outcome of interest. It is no exaggeration to suggest that job satisfaction is the most studied work outcome in organizational and management studies ([Locke, 1976](#)), dating back to Taylor's work on scientific management and the subsequent interest in human relations and leadership through the middle of the last century ([Bauer and Brazer, 2013](#)). [Hackman and Oldham's \(1980\)](#) work connects job satisfaction experienced by employees in different occupations to work design and factors such as role stress, and in a synthesis of research findings related to studies of job satisfaction in educational settings, [Thompson et al. \(1997\)](#) found factors such as role ambiguity and conflict to be among the most important predictors of satisfaction.

A variety of studies have been directed at understanding principal's job satisfaction (though considerably more work has been done related to teachers), and in contrast to the other outcomes featured in this paper, satisfaction is virtually always defined as a dependent variable predicted by personal, contextual, or job-related factors. Focusing on the latter, [Chang et al. \(2015\)](#) found that principals are more likely to be committed to their organization and more satisfied at work when they perceive their superintendents provide them with greater autonomy in their work; similarly, in a study of German principals [Dadaczynski et al. \(2020\)](#) found that decision latitude and sense of control is related to job satisfaction. In a study of Greek principals, [Saiti and Fassoulis \(2012\)](#) also found relations with supervisors to be important, including recognition for one's work and fair treatment.

In a study of principals in a US state, [Sodoma and Else \(2018\)](#) noted the general increase in duties and work pressure experienced by principals affecting stress and satisfaction, a finding echoed by [Wang et al. \(2018\)](#), who found "work intensification" to be related to job satisfaction of Canadian school leaders, and [Darmody and Smyth \(2016\)](#) who found workload-related variables to affect satisfaction of a sample of Irish headteachers. [Wang et al. \(2018\)](#) added, "Faced with increasing job expectations, contemporary principals are more likely to experience role ambiguity, overload, and conflict on a regular basis" (p. 75), thus exacerbating the situation.

Table 1 Keyword search parameters.

"Principal" or "head of school" or "headteacher"	and	"Satisfaction" or "job satisfaction" "Isolation" "Stress" or "job stress" or "burnout" "Efficacy" or "principal efficacy"
--	-----	--

Maforah (2015) expressed the workload issue a bit differently in noting that most principals find an imbalance between the effort they put into the job and the rewards—tangible and otherwise—that they receive from it. This may be especially true among principals in disadvantaged schools (Maforah and Schulze, 2012) and rural schools (Graham and Messner, 1998). Hu et al. (2016) found that salary and promotion satisfaction were related to perceived well-being, which were also related to job satisfaction, and Tran (2017) found pay satisfaction to be predictive of turnover intentions of principals, although Wang et al. (2019) found Korean principal's job satisfaction to be more related to symbolic prestige and authority than actual salary.

In a study using Teaching and Learning International Survey (TALIS) data, Garcia and Flores (2018) found that factors they describe as associated with bureaucracy (paperwork, excessive rules) most related to job dissatisfaction, whereas principals report greater satisfaction when they are able to involve others in decisions and promote leadership distribution. Hulpia and Devos (2009) also investigated the relationship between leadership distribution and principal satisfaction, finding that a cooperative and engaged leadership team was a strong influence on satisfaction, while the extent of distribution of leadership to teachers was inconsequential.

Conley et al. (2007) report that attachment to coworkers was an important predictor of job satisfaction for school principals and assistant principals, a finding supported by Chaplain (2001) in a study of primary headteachers who were particularly affected by support from mentors in the same role. In a study using the TALIS data, Liu and Bellibas (2018) found that while there were differences among countries studied, the most important predictor of principal job satisfaction—present for all 32 countries—related to social interactions and mutual respect among school staff. In a study of new principals in a single US state, Bauer and Brazer (2013) found that social support significantly reduced perceived isolation, which in turn enhanced job satisfaction, even when accounting for the negative effects of role ambiguity and overload. This affirms the central role of social relations and peer support as a satisfier.

A final point worth noting from the literature on job satisfaction is that among the four outcomes reported in this paper, satisfaction is most studied in relation to the others—that is, many of the studies reviewed involve appreciating the impact of efficacy, stress, burnout and isolation on job satisfaction, and as noted earlier, satisfaction is most often cast as a dependent variable. In rarer cases, however, satisfaction is a mediator—for example, Federici and Skaalvik (2012) first, and Bauer and Silver (2018) in a replication of their work, position job satisfaction as a mediator between efficacy and burnout, with satisfaction preceding burnout as a predictor of principal's intention to leave.

In sum, then, job satisfaction has been empirically associated with a variety of working conditions including workload; job attributes such as autonomy and scope of influence; and other work outcomes such as efficacy and stress. Schermuly et al. (2011) suggest that school systems might see dissatisfaction as a sort of “early alert” to the potential of leaders' impending burnout, and in a similar way, satisfaction is apt to be cast as an early warning of turnover. Let it suffice to say that the ways the four constructs discussed in this paper relate to one another is itself a worthy focus of empirical study.

Principal efficacy

Efficacy is an individual's belief about his or her ability to complete a task successfully (Bandura, 1997). In schools, the study of self-efficacy has focused mostly on classroom teachers; research shows that higher levels of student achievement are associated with higher degrees of teacher self-efficacy and collective efficacy (Goddard et al., 2000; Tschannen-Moran and Woolfolk-Hoy, 2001). While empirical examinations of leader efficacy in schools are somewhat few in number (Federici, 2013), principal efficacy has garnered more attention recently. For instance, Leithwood and Jantzi (2008) found that higher degrees of principals' self-efficacy and perceptions of their schools' collective efficacy are associated with enhanced school improvement, and that principals' self-efficacy is malleable and susceptible to district influences. Louis et al. (2010) reported that principals' sense of efficacy related to school improvement is impacted by the degree of clarity and consensus on important school goals, and that both leader self-efficacy and sense of collective efficacy are important outcomes related to leader effectiveness. Higher levels of principal efficacy may be related to effective decision making (Baroudi and Hojeij, 2020) and principal's ability to lead change (Daly et al., 2011; Tschannen-Moran and Gareis, 2004). Principals with higher levels of efficacy tend to be more engaged at work and report higher levels of job satisfaction and lower levels of burnout (Baroudi and Hojeij, 2020; Federici and Skaalvik, 2012; Tschannen-Moran and Gareis, 2004). Grissom and Loeb (2009) remind us, though, that the causal direction between principal self-efficacy and school outcomes may be complex—principal self-efficacy might lead to more effective schools, but also principals may feel more efficacious in schools that are higher performing.

Efficacy is often characterized as an individual resource that leaders bring to the role (Guglielmi et al., 2012); consistent with this perspective, it is cast as a predictor of other work outcomes. Federici and Skaalvik (2011, 2012; also Federici, 2013), for example, conducted a number of studies related to leader efficacy of Norwegian principals, and this work paints a picture of the importance of leader efficacy as a predictor of various consequences, including work engagement (Federici and Skaalvik, 2011), job satisfaction and burnout (Federici, 2013; Federici and Skaalvik, 2012). Skaalvik (2020) positioned efficacy for instructional leadership as a predictor and it emerged as statistically significant in relation to engagement, emotional exhaustion, and intention to leave. Hesbol (2019) found principal efficacy beliefs to be related to perceptions of the school as a learning organization, i.e., principals with a more robust sense of efficacy were more apt to pursue ambitious improvements and had stronger beliefs in the capabilities of teachers and others to engage in such change. In a study set in Italy, Guglielmi et al. (2012) found efficacy to be related to enhanced engagement, but it was unrelated to principal burnout. Leithwood et al. (2020) identify self-efficacy as among the psychological personal leadership resources highly effective principals possess and develop.

Daly et al. (2011) found that threat of sanctioning under-performing schools within existing accountability frameworks may lower principal self-esteem. They further speculated that districts would tend to react rigidly in the face of threat and “shut down” options for creativity and change, creating a downward and reinforcing spiral—principals of schools under sanction experience self-doubt and lower self-efficacy while their districts reduce degrees of freedom and schools’ capacity for change, leading to a greater probability of dire consequences for the already-sanctioned schools. While it is unlikely that sanctions will disappear in the current policy environment, this work suggests that districts might do well to explore their reactions to threat and consider ways they might help promote principal and school staff efficacy that could lead to greater capacity for improvement.

In terms of work-related factors that may relate to principal efficacy, Leithwood and Jantzi (2008) note that most of the research on efficacy tends to focus on its outcomes rather than its antecedents. They elaborate that of antecedents studied to that date, most were evident in only a single study and those most studied tended to involve demographics or attributes unrelated to organizational policies and practices. Their study focused on district leadership and organizational conditions, which they found to be more related to leaders’ collective efficacy beliefs than to their self-efficacy, although all factors studied were statistically meaningful predictors of leader self-efficacy. The strongest predictors were emphasis on teamwork; focus on achievement and quality of instruction; and district culture (e.g., clear, well understood goals). Their integrated model of the relationships among variables related to collective and self-efficacy beliefs showed that district conditions, as an aggregate, was highly predictive of leader self-efficacy. Leithwood and Jantzi also found that district investment in instructional leadership development was the weakest relationship observed, prompting them to speculate whether such investment was more a symbolic gesture by districts than an actual commitment to change and growth.

In contrast to these findings related to professional development, Grissom and Harrington (2010) found that principal professional development enhanced principal efficacy, particularly participating in a formal mentoring, coaching or peer observation process, a finding supported by Gümüş and Bellibaş (2020) in a study of Turkish principals. They found that self-efficacy was supported by participation in leadership development, and that it played a strong mediating role between participating in professional development and self-reported engagement in learning centered leadership practices. In a study of Iranian primary school principals, incidentally, Hallinger et al. (2018) also found that self-efficacy affected principals’ report of instructional leadership behaviors.

Federici (2013) studied several work-related conditions as predictors of efficacy. Specifically, this study found that respondents who report greater autonomy and fewer job-related constraints to autonomy in their work reported feeling more efficacious, and both autonomy and efficacy contributed to their perceived job satisfaction. Bauer and Brazer (2010) examined how role stressors (ambiguity, role conflict, role overload), social support, and involvement in a coaching program affected new principals’ isolation and self-efficacy beliefs, positioning isolation as a mediator and efficacy as the dependent variable. They showed, first, that role ambiguity, role overload, and social support were significant predictors of principal efficacy, and that the role stressors remained important predictors when introducing isolation into the model with isolation mediating the effects of social support (i.e., greater social support predicted reduced isolation, which predicted greater efficacy).

In contrast to the other outcomes presented in this review, efficacy appears most often in the literature as a precursor of other work-related outcomes, suggesting that principal efficacy may have a wide-ranging array of resulting effects. In general, efficacy seems to be cast as malleable, but evidence is hardly definitive regarding, for instance, whether leadership development and other organization-sponsored efforts are generally effective at promoting leader efficacy. We might speculate, based on the evidence presented, that if school systems could work to clarify goals, provide greater autonomy, adjust workload, and provide opportunities for networking and support from others, school leaders might be both less isolated and experience greater sense of efficacy, resulting in additional positive results such as greater job satisfaction.

Stress/burnout

Stress has to do with a person’s ability to respond to and cope with factors in the work environment (Beausaert et al., 2016; Friedman, 2002; Klocko and Wells, 2015), whereas Maslach (2003) describes burnout as a prolonged, negative response to persistent stressors in the workplace. A leader’s ability to cope with stressors is thus related to burnout, which may be mitigated, in turn, by such things as social support. The inability to cope, over time, may result in negative physical, emotional, and cognitive job-induced symptoms (Federici and Skaalvik, 2012). While all living creatures are constantly interacting with their environment which, in turn, stresses them, and individuals can and do learn new and better coping mechanisms all the time, consistent stress beyond one’s coping abilities can be debilitating and result in serious emotional, mental, and physical repercussions. Burnout, which might be construed as a form of stress that incessantly exceeds coping capacities, is most often associated with emotional exhaustion; it has been particularly associated with helping or human-service professions such as the principalship (Friedman, 2002; Oplatka, 2002; Stephenson and Bauer, 2010), perhaps accounting for the fact that the majority of empirical research found dealt specifically with burnout. DeMatthews et al. (2019) explain that individuals in helping professions often interact with people who are distressed or experiencing trauma themselves and may become distressed themselves by their limited abilities to remedy these situations. Stress and burnout have been associated with a wide variety of unfortunate outcomes, including reduced job satisfaction, turnover, and decreased ability to do one’s job (Darmody and Smyth, 2016; Friedman, 2002; Gursel et al., 2002; Maslach, 2003; Mitani, 2018; Steward, 2014).

Stressors may be personal or organizational in nature. Focusing on the latter, it is unsurprising that the most often-mentioned precipitator of principal stress is related to the overwhelming complexity of the role and overload (Beusaert et al., 2016; Darnoddy and Smyth, 2016; Grissom et al., 2015; Gursel et al., 2002; Klocko and Wells, 2015; Mahfouz, 2020; Somech and Miassy-Maljak, 2003; Tikkanen et al., 2017). Wells (2013) reminds us that "although specific pressures might be new, the call for principals to accomplish great things with little support, and to be all things to all people is certainly not" (p. 337). Nevertheless, recent accounts suggest that demands have expanded at an unsustainable rate (Darnoddy and Smyth, 2016). As Mahfouz (2020) put it, "Increasingly placed under scrutiny, school administrators are under unprecedented pressure. They must work in dynamic, rapidly changing environments and are expected to adapt while remaining composed and positive" (p. 440). Not only have the number of demands increased in number, but principals are under ever-increasing pressure to accomplish more in shorter timeframes, seldom with added resources. And as helping professionals, she adds, "They are expected to take full responsibility for any conflict, misunderstanding or mistake that takes place in their buildings and to fulfill all of their duties with considerable urgency, which can be emotionally exhausting" (p. 452).

A variety of additional studies from a wide variety of contexts reveal additional stressors that may well lead to burnout and other unwanted outcomes. In a study of Zimbabwean administrators, Nhundu (1999) found lack of autonomy as a potent stressor, as was lack of resources (a finding shared in many contexts) along with paperwork and workload-related issues. Yildirim and Dinc (2019), in a survey of Flemish leaders, found that in addition to workload, principals who had higher levels of role ambiguity and role conflict also reported higher levels of burnout symptoms. In a survey study of Quebec principals, Poirel et al. (2012) found that stressors were mainly administrative and bureaucratic constraints, especially at the elementary school level. Tikkanen et al. (2017) found relatively low degrees of stress and burnout among a sample of Finnish school leaders, but noted social relations playing a role in the incidence of burnout. Gursel et al. (2002) found Turkish principals affected by low salary and lack of access to information needed to do their jobs, and Friedman (2002), in a study of Israeli school leaders, found that more than overload, the dominant stressors leading to burnout were related to parental demands and weak teachers (often leading to discipline issues).

In the US, Klocko and Wells (2015) repeated a survey study in 2009 and 2012. They found that at the initial data collection, the most stressful work characteristics were managerial in nature and included concerns over loss of revenues, no doubt fueled by the temporal proximity to the global crisis of 2008. Paperwork, insufficient time to get work done, and constant interruptions were among the highest stressors both years. Mahfouz (2020) conducted a qualitative study of 13 principals and found an abundance of stressors she organized into three groups: work-related; relationship-related; and time-related. Among the work-related stressors were constant change, power dynamics, and bureaucracy; time-related issues included family/work balance, overload, and no time for self-care. Mitani (2018) finds strong evidence that sanctions associated with high-stakes policies in the US exacerbate principal stress and turnover. Finally, Wells and Klocko (2018) pose an intriguing question: does the current and expanding expectation that principals distribute leadership emerge as a stressor and source of angst for leaders or does it offer the possibility of reducing pressure associated with workload? Or perhaps, the corollary: Do many or most school leaders have the skills and abilities to distribute leadership effectively?

There seems to be less written about possible coping strategies or remedies for stress and burnout than its incidence and causes (Poirel et al., 2012), or as Mahfouz (2020) found in her qualitative study when she interviewed thirteen US principals, respondents reported far more stressors than coping strategies. Having a healthy work-life balance (Steward, 2014) and learning self-regulation strategies (Tikkanen et al., 2017) are recommended, and developing effective time-management skills was found to reduce job stress (Grissom et al., 2015). If there is a theme among the rather limited sources, it may be that increased quantity and quality of social support helps mitigate stress and burnout (Beusaert et al., 2016; Stephenson and Bauer, 2010). Chaplain (2001) found that the quality of relationships matter, and that headteachers found it especially helpful to receive support from other headteachers who could empathize and share their experience with their colleagues. Stephenson and Bauer (2010) found that increased social support and reduced isolation were among the most potent factors that mitigated or lessened emotional burnout, and Beusaert et al. (2016) found that certain types of social support decreased stress and burnout in their study of Australian school leaders. Specifically, they found that support from the broader community tended to mitigate burnout, which they speculate as related to the idea that feeling supported by the community makes principals feel more connected or a part of it.

Overall, the literature suggests that leaders' experience of stress is dependent on both the individual's ability to cope and adjust to environmental and organizational stressors, and factors controllable by the organization, and that repeated or persistent stress beyond this coping capacity may lead to burnout. If there is a prevalent theme in this literature, it is that the complexity associated with school leadership, work overload, and leaders' ability to control or influence what goes on in schools has a significant affect on the degree of stress and burnout experienced. Less seems to be known about coping strategies available to leaders, a deficiency that researchers and practitioners alike might do well to address.

Isolation

How is it possible, Cook et al. (2015) wonder, that individuals who are surrounded by literally hundreds of people and who are working in jobs that involve an unrelenting variety of tasks involving interaction with others might feel alone? "This is a conundrum of the modern world. In schools, we are in constant contact with teachers, parents, students, and community members, and yet we feel isolated" (p. 3)? And yet, isolation has affected the quality of work life for leaders and other professionals for decades (Ackerman and Maslin-Ostrowski, 2004; Bauer and Brazier, 2010; Bauer and Silver, 2018; Johnson et al., 2018; Stephenson and

Bauer, 2010), and may be an especially prevalent issue for new school leaders (Hobson and Sharp, 2005; Slater et al., 2018). The traditional egg-crate structure of schools is part of the explanation, but surely this affects teachers more than school leaders. Principals, however, are very much alone in terms of their role and responsibilities and have neither a peer group or true colleagues (Kelchtermans et al., 2011), even in larger and more complex schools with a variety of administrators. Ackerman and Maslin-Ostrowski (2004) identify isolation as "the endemic and chronic condition of leadership life" (p. 311); Kelchtermans et al. (2011) associates the structural loneliness of the role as linked to a struggle with belonging. This is exacerbated by the inherent role ambiguity and conflict associated with the position (Koromuz, 2016).

Cook et al. (2015) elaborate that in fact, education is not historically a collaborative industry; time for adults to work together is expanding with interest in distribution of leadership, communities of practice, and learning communities, but this is not historically valued. Johnson et al. (2018) found that even when teacher teams are employed in part to reduce isolation, they only do so tangentially; teams may be instrumental in promoting improvements in teaching, but such things as emotional support or collaboration are less central. Principals are held accountable and judged against criteria of effectiveness and efficiency, not either their or their subordinate's emotional well-being (Kelchtermans et al., 2011), and leaders may be reluctant to share problems, concerns, or other "weaknesses" with others.

Of the outcomes reported in this paper, isolation is the least studied in relation to the principalship (Bauer and Brazer, 2013; Koromuz, 2016; Stephenson and Bauer, 2010) and the literature on principal isolation is sparse (Beaudoin and Taylor, 2004; Cookson, 2005; Garmston, 2007; Hord, 2007; Schlechte et al., 2005). The concept of isolation has to do with the sense of feeling alone at work; it is less a structural reality than an emotional response. As schools have moved from being led by "principal teachers," to having principals that must take full responsibility for all of the administrative and instructional imperatives of complex organizations (Cuban, 1988), even today as we stress leadership capacity as a schoolwide attribute principals retain sole responsibility for school outcomes and make many of their key decisions in isolation.

A study conducted by Izgar (2009) shows that there is a statistical link between measures of principal loneliness and depression, and Cubitt and Burt (2002) found that isolation (loneliness) and work overload were the most significant predictors of stress among primary principals in a study set in New Zealand. Howard and Mallory's (2008) work supports the inclusion of isolation as a variable that operates in concert with the stress created by principal job expectations in the form of role ambiguity, role overload, and role conflict, and similarly, role ambiguity, conflict and overload were found to be statistically related to greater levels of isolation in a study of new principals in a US state (Bauer and Brazer, 2013; Bauer et al., 2019). In a study of Turkish school leaders Koromuz (2016) determined that communication and lack of time/overload were the most impactful work-related conditions associated with isolation.

In a study of recent graduates who became principals, Nichols and McBride (2016) found that social relationships, especially from former peers, contributed to resentment from those who may have been passed over for the position and isolation from former teacher peers. A study of new school leaders set in Malaysia found that while the incidence of loneliness was moderate, isolation was exacerbated by workload demands, autocratic leadership, and poor relations with teachers (Tahir et al., 2017). In contrast, studies of networking, mentoring or collaborative leadership learning found that isolation could be reduced through enhancement of collegial and mentoring relationships with professional colleagues (Hansford and Ehrich, 2006; Piggot-Irvine, 2004; Service et al., 2018). Trust et al. (2018) suggest that reduced isolation can be accomplished through social media and other on-demand learning and networking technologies.

To summarize, among the work outcomes reviewed in this paper, isolation represents a bit of a paradox; school leaders are in relationship with and regularly interact with many individuals during any workday, and yet they sometimes feel alone. Further, they are uniquely responsible and accountable for myriad consequences related to what schools do, including safety and learning. And in many school contexts, leaders are expected to have a depth and range of expertise that can be genuinely daunting. While the empirical work on isolation is relatively scant, one theme that appears well-supported from the few empirical studies on isolation is that the quality of collegial relationships at work are related to the degree of isolation experienced by school leaders (Bauer and Brazer, 2010; Bauer et al., 2019; Beusaert et al., 2016; Kelchtermans et al., 2011).

Discussion

A few things might be concluded from this review that can assist those who want to better understand the work and work lives of school leaders, improve those lives, and consequently improve the schools. There are several variables that emerge in relation to more than one of the work outcomes reviewed that may be relatively easily addressed that could improve the work lives of school leaders, among them: role stressors, particularly work overload and role ambiguity; communications, networking opportunities, mentoring and social support; and autonomy and decision-making control. The fact that these constructs are themselves related makes this even more important. We might speculate, for instance, that demands on principals to perform as instructional, transformational and managerial leaders makes their work more complex and leads to a much greater likelihood that they will experience role ambiguity, conflict, or overload on a regular basis; restrictions on their autonomy or lack of opportunities for social support might easily exacerbate or otherwise complicate their ability to lead and problem-solve effectively.

In addition, several questions recur throughout that might serve as suggestions for attention by researchers, aside from the obvious one that more research is needed in relation to each of these work outcomes for principals. Among them is the fascinating question of the impact of leadership distribution on each outcome and particularly, assuming school leaders have the skills and

vision to distribute leadership and followers have the capacity to successfully contribute as leaders, does this reduce isolation, stress, and work overload of principals and lead to greater job satisfaction and/or efficacy?

A significant limitation of this paper is the emphasis on the four constructs of interest as dependent variables; by stressing the work-related factors that predict each, we neglect the possible inter-relationships between and among these factors and how examining them as connected phenomena might promote greater understanding of the work experience of principals and thus provide only a limited picture of the ways principal quality of work life is affected and might be enhanced. As Conley and Woosley (2000) note, the relationships among variables associated with the design of leaders' work and outcomes like job satisfaction is unlikely to be "simple and direct" (p. 181), but rather mediated by factors associated with the work context. Throughout the discussions presented, of course, some literature dealing with each construct as a predictor rather than outcome was presented, and yet it seems quite likely more complex relationships exist between and among each AND among these variables and other outcomes such as persistence, job effectiveness, and perhaps even school effectiveness. Some few pieces were found that probe these relationships, for instance, Federici and Skaalvik (2012) and Bauer and Silver (2018) examine the relationships between principal efficacy, job satisfaction, and burnout in relation to their impact on principals' intention to leave their job (with Bauer and Silver adding the impact of isolation as a precursor). However, considerably more research needs to be done to tease out these more complex relationships and better understand the opportunities that might exist to help enhance the quality of work life and work outcomes for leaders.

The most important take-away from this review is expressed eloquently by Ackerman and Maslin-Ostrowski (2004): Much of what is studied and written about school leaders centers on aspects of leader behavior, what effective leaders do (or should do), and the impact leaders have on organizational outcomes rather than who leaders are or what they experience at work. "The leadership life," they write, "... is a complex balance of conflicting forces and tensions that manage to function most of the time. Leadership is 'practiced' at an extremely busy and sometimes hazardous intersection of the personal and professional where there is often a thundering flow of traffic" (p. 317). While it is certainly true that scholarship on the role of the principal has gained in both volume and quality over the past few decades, hindsight—and this review—suggests that we understand much more about effective school leadership behavior and its impact on teachers and students than we do about the quality of work life of school leaders themselves, their well-being, or how to enrich their work lives. We might wonder, consequently, whether we contribute what we should to leaders' abilities to perform their vitally important role or enhance, as we might, their motivation to do so.

References

- Ackerman, R.H., Maslin-Ostrowski, P., 2004. The wounded leader and emotional learning in the schoolhouse. *Sch. Leader. Manag.* 24 (3), 311–328. <https://doi.org/10.1080/1363243042000266945>.
- Alenezi, A.S., 2020. Risk of turnover among U.S. principals based on personal and school characteristics. *J. Res. Educ.* 26 (1), 1–19.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman, New York, NY.
- Baroudi, S., Hojeij, Z., 2020. The role of self-efficacy as an attribute of principals' leadership effectiveness in K-12 private and public institutions in Lebanon. *Int. J. Leader. Educ.* 23 (4), 457–471. <https://doi.org/10.1080/13603124.2018.1529822>.
- Bauer, S.C., Brazer, S.D., 2010. The impact of isolation on the self-efficacy of new principals'. In: Conley, S., Cooper, B. (Eds.), *Teacher Collaboration: Advancing Professionalism & School Quality*. Rowman and Littlefield, New York, pp. 89–112.
- Bauer, S.C., Brazer, S.D., 2013. The impact of isolation on the job satisfaction of new principals. *J. Sch. Leader.* 23 (1), 152–177. <https://doi.org/10.1177/105268461302300106>.
- Bauer, S.C., Silver, L., 2018. The impact of job isolation on new principals' sense of efficacy, job satisfaction, burnout and persistence. *J. Educ. Adm.* 56 (3), 315–331. <https://doi.org/10.1108/JEA-07-2017-0078>.
- Bauer, S.C., Silver, L., Schwartz, J., 2019. The impact of isolation on new principals' persistence: evidence from a southern US state. *Educ. Manag. Adm. Leader* 47 (3), 383–399. <https://doi.org/10.1177/1741143217739359>.
- Beaudoin, M., Taylor, M.E., 2004. *Creating a Positive School Culture: How Principals and Teachers Can Solve Problems Together*. Corwin Press, Thousand Oaks, CA.
- Beausaert, S., Froehlich, D., Devos, C., Riley, P., 2016. Effects of support on stress and burnout in school principals. *Educ. Res.* 58 (4), 347–365. <https://doi.org/10.1080/00131881.2016.1220810>.
- Beckett, L.O., 2018. Predictors of urban principal turnover. *Urban Educ.* <https://doi.org/10.1177/0042085918772629>.
- Bezzina, C., Paletta, A., Alimehmeti, G., 2018. What are school leaders in Italy doing? An observational study. *Educ. Manag. Adm. Leader* 46 (5), 841–863. <https://doi.org/10.1177/1741143217694896>.
- Chang, Y., Leach, N., Anderman, E.M., 2015. The role of perceived autonomy support in principals' affective organizational commitment and job satisfaction. *Soc. Psychol. Educ.* 18 (2), 315–336. <https://doi.org/10.1007/s11218-014-9289-z>.
- Chaplain, R.P., 2001. Stress and job satisfaction among primary headteachers: a question of balance? *Educ. Manag. Adm.* 29 (2), 197–215. <https://doi.org/10.1177/0263211X010292005>.
- Cieminski, A.B., 2018. Practices that support leadership succession and principal retention. *Educ. Leadersh. Rev.* 19 (1), 21–41.
- Conley, S., Shaw, S., Glasman, N., 2007. Correlates of job and growth satisfaction among secondary school administrators. *J. Sch. Leader.* 17 (1), 54–88.
- Conley, S., Woosley, S., 2000. Teacher role stress, higher order needs, and work outcomes. *J. Educ. Adm.* 38 (2), 179–201.
- Cook, S.C., Johnson, J.J., Stager, T.C., 2015. *Breaking Out of Isolation: Becoming a Connected School Leader*. SAGE Publications, Thousand Oaks, CA.
- Cookson, P.W., 2005. The challenge of isolation: professional development-your first year. *Teach. Pre K-8* 36, 14.
- Cuban, L., 1988. *The Managerial Imperative and the Practice of Leadership in Schools*. SUNY Press, Albany.
- Cubitt, S., Burt, C., 2002. Leadership style, loneliness and occupational stress in New Zealand primary school principals. *N. Z. J. Educ. Stud.* 37 (2), 159–169.
- Dadaczynski, K., Paulus, P., Horstmann, D., 2020. The predictive value of individual and work-related resources for the health and work satisfaction of German school principals. *Health Educ. J.* 79 (2), 225–236. <https://doi.org/10.1177/0017896919867118>.
- Daly, A.J., Der-Martirosian, C., Ong-Dean, C., Park, V., Wishard-Guerra, C., 2011. Leading under sanction: principals' perceptions of threat rigidity, efficacy, and leadership in underperforming schools. *Leader. Pol. Sch.* 10 (2), 171–206. <https://doi.org/10.1080/15700763.2011.557517>.

- Darmody, M., Smyth, E., 2016. Primary school principals' job satisfaction and occupational stress. *Int. J. Educ. Manag.* 30 (1), 115–128. <https://doi.org/10.1108/IJEM-12-2014-0162>.
- Day, C., 2014. Resilient principals in challenging schools: the courage and costs of conviction. *Teach. Teach.* 20 (5), 638–654. <https://doi.org/10.1080/13540602.2014.937959>.
- Day, C., Gu, Q., Sammons, P., 2016. School leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Day, C., Sammons, P., Leithwood, K., Harris, A., Hopkins, D., Gu, Q., Brown, E., Antariou, E., 2011. *Successful School Leadership: Linking with Learning and Achievement*. Open University Press, London.
- De Nobile, J.J., McCormick, J., 2008. Job satisfaction of Catholic primary school staff: a study of biographical differences. *Int. J. Educ. Manag.* 22 (2), 135–150. <https://doi.org/10.1108/09513540810853549>.
- DeMatthews, D.E., Carrola, P., Knight, D., Izquierdo, E., 2019. Principal burnout: how urban school leaders experience secondary trauma on the US-Mexico border. *Leader. Pol. Sch.* 18 (4), 681–700. <https://doi.org/10.1080/15700763.2018.1513153>.
- Earley, P., Bubb, S., 2013. A day in the life of new headteachers: learning from observation. *Educ. Manag. Adm. Leader* 41 (6), 782–799. <https://doi.org/10.1177/1741143213494189>.
- Federici, R.A., 2013. Principals' self-efficacy: relations with job autonomy, job satisfaction, and contextual constraints. *Eur. J. Psychol. Educ.* 28 (1), 73–86. <https://doi.org/10.1007/s10212-011-0102-5>.
- Federici, R.A., Skaalvik, E.M., 2011. Principal self-efficacy and work engagement: assessing a Norwegian principal self-efficacy scale. *Soc. Psychol. Educ.* 14 (4), 575–600. <https://doi.org/10.1007/s11218-011-9160-4>.
- Federici, R.A., Skaalvik, E.M., 2012. Principal self-efficacy: relations with burnout, job satisfaction and motivation to quit. *Soc. Psychol. Educ.* 15 (3), 295–320. <https://doi.org/10.1007/s11218-012-9183-5>.
- Friedman, I.A., 2002. Burnout in school principals: role related antecedents. *Soc. Psychol. Educ.* 5, 229–251.
- Garcia, I.M., Flores, J.G., 2018. An explanation of job satisfaction in school principals based on their performance of the position. *Educ. Knowl. Soc.* 19 (1), 77–95. <https://doi.org/10.14201/eks20181917795>.
- Garmston, R., 2007. Collaborative culture. *J. Staff Dev.* 28, 55–57.
- Goddard, R., Hoy, W.K., Woolfolk Hoy, A., 2000. Collective teacher efficacy: its meaning, measure, and impact on student achievement. *Am. Educ. Res. J.* 37 (2), 479–508.
- Graham, M.W., Messner, P.E., 1998. Principals and job satisfaction. *Int. J. Educ. Manag.* 12 (5), 196–202. <https://doi.org/10.1108/09513549810225925>.
- Grissom, J.A., Harrington, J.R., 2010. Investing in administrator efficacy: an examination of professional development as a tool for enhancing principal effectiveness. *Am. J. Educ.* 116 (4), 583–612. <https://doi.org/10.1086/653631>.
- Grissom, J.A., Loeb, S., 2009. *Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills*. School Leadership Research Report No. 09-1. Stanford University, Institute for Research on Education Policy and Practice, Stanford, CA.
- Grissom, J.A., Loeb, S., Master, B., 2013. Effective instructional time use for school leaders: longitudinal evidence from observations of principals. *Educ. Res.* 42 (8), 433–444. <https://doi.org/10.3102/0013189X13510020>.
- Grissom, J.A., Loeb, S., Mitani, H., 2015. Principal time management skills: explaining patterns in principals' time use, job stress, and perceived effectiveness. *J. Educ. Adm.* 53 (6), 773–793. <https://doi.org/10.1108/JEA-09-2014-0117>.
- Guglielmi, D., Simbula, S., Schaufeli, W., Depola, M., 2012. Self-efficacy and workaholism as initiators of the job demands-resources model. *Career Dev. Int.* 17 (4), 375–389. <https://doi.org/10.1108/13620431211255842>.
- Gümüş, S., Bellibas, M., 2020. The relationship between professional development and school principals' leadership practices: the mediating role of self-efficacy. *Int. J. Educ. Manag.* 34 (7), 1155–1170. <https://doi.org/10.1108/IJEM-10-2019-0380>.
- Gumus, S., Bellibas, M., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48. <https://doi.org/10.1177/1741143216659296>.
- Gursel, M., Sunbul, A.M., Sari, H., 2002. An analysis of burnout and job satisfaction between Turkish headteachers and teachers. *Eur. J. Psychol. Educ.* 17 (1), 35–45. <https://doi.org/10.1007/BF03173203>.
- Hackman, J., Oldham, G., 1980. *Work Redesign*. Addison-Wesley, Reading, MA.
- Hallinger, P., Hosseingholizadeh, R., Hashemi, N., Kouhsari, M., 2018. Do beliefs make a difference? Exploring how principal self-efficacy and instructional leadership impact teacher efficacy and commitment in Iran. *Educ. Manag. Adm. Leader* 46 (5), 800–819. <https://doi.org/10.1177/1741143217700283>.
- Hansen, C., 2018. Why rural principals leave. *Rural Educ.* 39 (1), 41–53. <https://doi.org/10.35608/ruraled.v39i1.214>.
- Hansford, B., Ehrich, L.C., 2006. The principalship: how significant is mentoring? *J. Educ. Adm.* 44 (1), 36–52. <https://doi.org/10.1108/09578230610642647>.
- Hauserman, D.C., Pollock, K., Wang, F., 2017. Inconvenient, but essential: impact and influence of school-community involvement on principals' work and workload. *Sch. Community J.* 27 (1), 83–105.
- Hesbol, K.A., 2019. Principal self-efficacy and learning organizations: influencing school improvement. *Int. J. Educ. Leadersh. Preparation* 14 (1), 33–51.
- Hobson, A.J., Sharp, C., 2005. Head to head: a systematic review of the research evidence on mentoring new head teachers. *Sch. Leader. Manag.* 25 (1), 25–42. <https://doi.org/10.1080/1363243052000317073>.
- Hord, S.M., 2007. Learn in community with others. *J. Staff Dev.* 28, 39–42.
- Horng, E.L., Klasik, D., Loeb, S., 2010. Principal's time use and school effectiveness. *Am. J. Educ.* 116 (4), 491–523. <https://doi.org/10.1086/653625>.
- Howard, M.P., Mallory, B.J., 2008. Perceptions of isolation among high school principals. *J. Women Educ. Leadersh.* 6, 7–27.
- Hu, C., Cui, S., Wang, L., 2016. Path analysis of work family conflict, job salary and promotion satisfaction, work engagement to subjective well-being of the primary and middle school principals. *J. Educ. Train. Stud.* 4 (9), 10–15. <https://doi.org/10.11114/jets.v4i9.1619>.
- Huang, T., Hochbein, C., Simons, J., 2020. The relationship among school contexts, principal time use, school climate, and student achievement. *Educ. Manag. Adm. Leader* 48 (2), 305–323. <https://doi.org/10.1177/1741143218802595>.
- Hulpia, H., Devos, G., 2009. Exploring the link between distributed leadership and job satisfaction of school leaders. *Educ. Stud.* 35 (2), 153–171. <https://doi.org/10.1080/03055690802648739>.
- Izgar, H., 2009. An investigation of depression and loneliness among school principals. *Educ. Sci. Theor. Pract.* 9, 247–258.
- Johnson, S.M., Reinhorn, S., Simon, N., 2018. Ending isolation: the payoff of teacher teams in successful high-poverty urban schools. *Teach. Coll. Rec.* 120 (5), 1–46.
- Kelchtermans, G., Piot, L., Ballet, K., 2011. The lucid loneliness of the gatekeeper: exploring the emotional dimension in principals' work lives. *Oxf. Rev. Educ.* 37 (1), 93–108. <https://doi.org/10.1080/03054985.2010.545192>.
- Klocko, B.A., Wells, C.M., 2015. Workload pressures of principals: a focus on renewal, support, and mindfulness. *NASSP Bull.* 99 (4), 332–355. <https://doi.org/10.1177/0192636515619727>.
- Koromuz, S., 2016. Invisible barriers: the loneliness of school principals at Turkish elementary schools. *S. Afr. J. Educ.* 36 (4) <https://doi.org/10.15700/saje.v36n4a1302>.
- Leithwood, K., Day, C., 2008. The impact of school leadership on pupil outcomes. *Sch. Leader. Manag.* 28 (1), 1–4. <https://doi.org/10.1080/13632430701799718>.
- Leithwood, K., Jantzi, D., 2008. Linking leadership to student learning: the contributions of leader efficacy. *Educ. Adm. Q.* 44 (4), 496–528. <https://doi.org/10.1177/0013161X08321501>.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28 (1), 27–42. <https://doi.org/10.1080/13632430701800060>.
- Leithwood, K., Mascal, B., 2008. Collective leadership effects on student achievement. *Educ. Adm. Q.* 44 (4), 529–561. <https://doi.org/10.1177/0013161X08321221>.

- Leithwood, K., Patten, S., Jantzi, D., 2010. Testing a conception of how school leadership influences student learning. *Educ. Adm. Q.* 46 (5), 671–706. <https://doi.org/10.1177/0013161X10377347>.
- Leithwood, K., Sun, J., Pollock, K., 2017. *How School Leadership Influences Student Learning: The Four Paths*. Springer, The Netherlands.
- Liu, Y., Bellibas, M.S., 2018. School factors that are related to school principals' job satisfaction and organizational commitment. *Int. J. Educ. Res.* 90, 1–19. <https://doi.org/10.1016/j.ijer.2018.04.002>.
- Locke, E., 1976. The nature and causes of job satisfaction. In: Dunnette, M. (Ed.), *Handbook of Industrial and Organizational Psychology*. Rand McNally, Chicago, pp. 1297–1349.
- Louis, K.S., et al., 2010. Final Report of Research to the Wallace Foundation. Wallace Foundation, New York.
- Maforah, T.P., 2015. Secondary schools principals and their job satisfaction: a test of process theories. *J. Int. Educ. Leadersh.* 5 (2).
- Maforah, T.P., Schulze, S., 2012. The job satisfaction of principals of previously disadvantaged schools: new light on an old issue. *S. Afr. J. Educ.* 32 (3), 227–239. <https://doi.org/10.15700/saje.v32n3a571>.
- Mahfouz, J., 2020. Principals and stress: few coping strategies for abundant stressors. *Educ. Manag. Adm. Leader* 48 (3), 440–458. <https://doi.org/10.1177/1741143218817562>.
- Mascall, B., Leithwood, K., 2010. Investing in leadership: the district's role in managing principal turnover. *Leader. Pol. Sch.* 9 (4), 367–383. <https://doi.org/10.1080/15700763.2010.493633>.
- Maslach, C., 2003. Job burnout: new directions in research and intervention. *Curr. Dir. Psychol. Sci.* 12 (5), 189–192.
- Maxwell, A., Riley, P., 2017. Emotional demands, emotional labor and occupational outcomes in school principals: modeling the relationships. *Educ. Manag. Adm. Leader* 45 (3), 484–502. <https://doi.org/10.1177/1741143215607878>.
- Mitani, H., 2018. Principals' working conditions, job stress, and turnover behaviors under NCLB accountability pressure. *Educ. Adm. Q.* 54 (5), 822–862. <https://doi.org/10.1177/0013161X18785874>.
- Nhundu, T.J., 1999. Determinants and prevalence of occupational stress among Zimbabwean school administrators. *J. Educ. Adm.* 37 (3), 256–272. <https://doi.org/10.1108/09578239910275490>.
- Ni, Y., Sun, M., Rorrer, A., 2015. Principal turnover: upheaval and uncertainty in charter schools? *Educ. Adm. Q.* 51 (3), 409–437. <https://doi.org/10.1177/0013161X14539808>.
- Nichols, J., McBride, J., 2016. Promoted from within: preparing beginning educational leaders for executive loneliness that occurs in their new position. *Coll. Student J.* 51 (1), 47–56.
- Norton, M.S., 2002. Let's keep our quality school principals on the job. *High Sch. J.* 86 (2), 50–56. <https://doi.org/10.1353/hsj.2002.0024>.
- Oplatka, I., 2002. Women principals and the concept of burnout: an alternative voice? *Int. J. Leader. Educ.* 5 (3), 211–226. <https://doi.org/10.1080/136031200112873>.
- Oplatka, I., 2017. Principal workload: components, determinants and coping strategies in an era of standardization and accountability. *J. Educ. Adm.* 55 (5), 552–568. <https://doi.org/10.1108/JEA-06-2016-0071>.
- Piggot-Irvine, E., 2004. Growth, development and a way out of principalship's isolation. *Manag. Educ.* 18 (1), 24–29. <https://doi.org/10.1177/08920206040180010501>.
- Poirrel, E., Lapointe, P., Yvon, F., 2012. Coping with administrative constraints by Quebec school principals. *Can. J. Sch. Psychol.* 27 (4), 302–318. <https://doi.org/10.1177/0829573512461131>.
- Robinson, S., 2011. Primary headteachers: new leadership roles inside and outside the school. *Educ. Manag. Adm. Leader* 39 (1), 63–83. <https://doi.org/10.1177/1741143210383899>.
- Saiti, A., Fassoulis, K., 2012. Job satisfaction: factor analysis of Greek primary school principals' perceptions. *Int. J. Educ. Manag.* 26 (4), 370–380. <https://doi.org/10.1108/09513541211227773>.
- Schermuly, C.C., Schermuly, R.A., Meyer, B., 2011. Effects of vice-principals' psychological empowerment on job satisfaction and burnout. *Int. J. Educ. Manag.* 25 (3), 252–264. <https://doi.org/10.1108/09513541111120097>.
- Schlechte, J., Yssel, N., Merbler, J., 2005. Case studies in teacher isolation and alienation. *Prev. Sch. Fail.* 50, 35.
- Service, B., Dalgic, G.E., Thornton, K., 2018. Benefits of a shadowing/mentoring intervention for New Zealand school principals. *Prof. Dev. Educ.* 44 (4), 507–520. <https://doi.org/10.1080/19415257.2017.1378705>.
- Skaalvik, C., 2020. School principal self-efficacy for instructional leadership: relations with engagement, emotional exhaustion and motivation to quit. *Soc. Psychol. Educ.* 23 (2), 479–498. <https://doi.org/10.1007/s11218-020-09544-4>.
- Slater, C.L., Garcia Garduno, J.M., Mentz, K., 2018. Frameworks for principal preparation and leadership development: contributions of the international study of principal preparation (ISPP). *Manag. Educ.* 32 (3), 126–134. <https://doi.org/10.1177/0892020617747611>.
- Snodgrass Rangel, V., 2018. A review of the literature on principal turnover. *Rev. Educ. Res.* 88 (1), 87–124. <https://doi.org/10.3102/0034654317743197>.
- Sodoma, B., Else, D., 2018. Job satisfaction of Iowa public school principals. *Rural Educ.* 31 (1) <https://doi.org/10.35608/ruraled.v31i1.438>.
- Somech, A., Miassy-Maljak, N., 2003. The relationship between religiosity and burnout of principals: the meaning of educational work and role variables as mediators. *Soc. Psychol. Educ.* 6, 61–90.
- Southworth, G., 2008. Primary school leadership today and tomorrow. *Sch. Leader. Manag.* 28 (5), 413–434. <https://doi.org/10.1080/13632430802499978>.
- Stephenson, L.E., Bauer, S.C., 2010. The role of isolation in predicting new principals' burnout. *Int. J. Educ. Policy Leadersh.* 5 (9) <https://doi.org/10.22230/ijepi.2010v5n9a275>.
- Steward, J., 2014. Sustaining emotional resilience for school leadership. *Sch. Leader. Manag.* 34 (1), 52–68. <https://doi.org/10.1080/13632434.2013.849686>.
- Sun, M., Ni, Y., 2015. Work environments and labor markets. *Educ. Adm. Q.* 52 (1), 144–183.
- Tahir, L., et al., 2017. Novice head teachers' isolation and loneliness experiences: a mixed-methods study. *Educ. Manag. Adm. Leader* 45 (1), 164–189. <https://doi.org/10.1177/1741143215587302>.
- Tekleselassie, A.A., Villarreal, P., 2011. Career mobility and departure intentions among school principals in the United States: incentives and disincentives. *Leader. Pol. Sch.* 10 (3), 251–293. <https://doi.org/10.1080/15700763.2011.585536>.
- Thompson, D., McNamara, J., Hoyle, J., 1997. Job satisfaction in educational organizations: a synthesis of research findings. *Educ. Adm. Q.* 33 (1), 7–37.
- Tikkanen, L., Pyhalto, K., Pietarinen, J., Soini, T., 2017. Interrelations between principals' risk of burnout profiles and proactive self-regulation strategies. *Soc. Psychol. Educ.* 20 (2), 259–274. <https://doi.org/10.1007/s11218-017-9379-9>.
- Tran, H., 2017. The impact of pay satisfaction and school achievement on high school principals' turnover intentions. *Educ. Manag. Adm. Leader* 45 (4), 621–638. <https://doi.org/10.1177/1741143216636115>.
- Trust, T., Carpenter, J.P., Krutka, D.G., 2018. Leading by learning: exploring the professional learning networks of instructional leaders. *Educ. Media Int.* 55 (2), 137–152. <https://doi.org/10.1080/09523987.2018.1484041>.
- Tschannen-Moran, M., Gareis, C.R., 2004. Principals' sense of efficacy: assessing a promising construct. *J. Educ. Adm.* 42 (5), 573–585. <https://doi.org/10.1108/09578230410554070>.
- Tschannen-Moran, M., Woolfolk Hoy, A., 2001. Teacher efficacy: capturing an elusive construct. *Teach. Teach. Educ.* 17 (7), 783–805.
- Türkoglu, M.E., Cansoy, R., 2020. School principals' opinions on the meaning and sources of workload. *Int. J. Contemp. Educ. Res.* <https://doi.org/10.33200/ijer.657994>.
- Urick, A., Bowers, A.J., 2014. What are the different types of principals across the United States? A latent class analysis of principal perception of leadership. *Educ. Adm. Q.* 50 (1), 96–134. <https://doi.org/10.1177/0013161X13489019>.
- Wang, F., Pollock, K., Hauseman, C., 2018. School principals' job satisfaction: the effects of work intensification. *Can. J. Educ. Adm. Pol.* Available at: <https://eric.ed.gov/?id=EJ1179195>. (Accessed 21 June 2020).
- Wang, C., et al., 2019. Perspectives of South Korean school principals on job satisfaction: in comparison to German and U.S. principals. *Leadersh. Policy Sch.* 1–16. <https://doi.org/10.1080/15700763.2019.1696372>.

- Wells, C.M., 2013. Principals responding to constant pressure: finding a source of stress management. *NASSP Bull.* 97 (4), 335–349. <https://doi.org/10.1177/0192636513504453>.
- Wells, C.M., Klocko, B.A., 2018. Principal well-being and resilience: mindfulness as a means to that end. *NASSP Bull.* 102 (2), 161–173. <https://doi.org/10.1177/0192636518777813>.
- West, D.L., Peck, C., Reitzug, U.C., 2010. Limited control and relentless accountability: examining historical changes in urban school principal pressure. *J. Sch. Leader.* 20 (2), 238–266. <https://doi.org/10.1177/105268461002000206>.
- Yildirim, F., Dinc, M.S., 2019. Factors influencing burnout of the principals: a pilot study in Flemish schools of Belgium. *Econ. Res.* 32 (1), 3538–3553. <https://doi.org/10.1080/1331677X.2019.1660200>.

Leading in urban contexts: how principal supervisors coach principals to lead improvement

Rebecca A. Thessin^a and Coleen Reyes^b, ^aDepartment of Educational Administration, The George Washington University, Washington, DC, United States; and ^bThe George Washington University, Baltimore Urban Debate League, Baltimore, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	325
Methodology	326
Data sources	326
Interviews	326
Observations	327
Documents	327
Surveys	327
Data analysis	327
Findings	328
Qualities of a productive principal supervisor–principal partnership	328
Background	328
Summary of findings	328
Engagement in joint work	330
Background	330
Summary of findings	330
Principal supervisor differentiation strategies	331
Background	331
Summary of findings	331
Developing partnerships	331
Use of context and principal needs	332
Principal supervisor interactions with leadership teams	332
Background	332
Summary of findings	333
Implications for practice and policy	333
Prioritize selection and preparation	334
Communicate clear expectations	334
Establish thoughtful partnerships	334
Support learning of school-based instructional leadership teams	334
Conclusion	334
Acknowledgment	335
References	335

Introduction

High-needs schools, or those with repeated low student achievement and other compounding challenges like poverty and lack of resources, demand more of principals than their lower-needs and higher-resourced counterparts (Finnegan and Stewart, 2009). Challenges associated with leading high-needs schools are typically greater, particularly when facing accountability pressures, requiring nuanced actions and considerations by school-based leaders and those who support improving teaching and learning in these schools (Cosner and Jones, 2016). Since the positive impact of the principal on student success is second only to that of teachers (Leithwood et al., 2020), the principal's role as an instructional leader who leads teaching and learning improvement is particularly crucial in high-needs settings. In order to best support high-needs schools, principals must secure additional and strategic resources, focus resources on the instructional program of the school (Klar and Brewer, 2013), engage in practices to increase continuous improvement such as data monitoring and analysis (Day et al., 2016), and facilitate ongoing aligned professional learning for leaders, teachers, and staff, among other responsibilities (Finnegan and Stewart, 2009). In other words, principals of high-needs schools must engage in focused instructional improvement efforts.

Yet, principals cannot effectively lead instructional improvement by themselves, a reality that is particularly apparent in high-needs schools (Thessin et al., 2020). The school district central office plays a critical role in supporting these improvement efforts by providing supervision and support for principals (Thessin and Louis, 2020). Within the central office, the role of the principal

supervisor has gained renewed attention in the last decade as it has been identified as a key lever for principals' instructional leadership learning and development (Goldring et al., 2018; Turnbull et al., 2015).

To address this need for ongoing, school-based, professional learning to grow principals' skills as instructional leaders, many of our nation's largest urban school systems have responded by changing expectations for those who support and supervise principals (Honig, 2012; Thessin and Louis, 2019). To accomplish this goal, many large city school systems in the US have allocated additional funding to reduce the ratio of principal supervisors to principals (Goldring et al., 2018). These central offices have reshaped the role of principal supervisors to focus their work on providing intensive, job-embedded coaching to principals (Corcoran et al., 2013; Honig, 2012) while also continuing their responsibilities as evaluators. Principal supervisors in these districts support principals' learning as coaches who provide feedback, model practices, broker resources, engage in joint classroom observations and development of tools, prioritize time on task to instructional leadership development, and convene groups of principals for collaborative learning, among other tasks (Corcoran et al., 2013; Goldring et al., 2018; Honig, 2008, 2012; Honig and Rainey, 2014; Thessin, 2021; Turnbull et al., 2015).

For 16 months, from 2017 to 2018, we followed the work of 12 principal supervisors and the principals they supported at high needs schools in one large, urban Mid-Atlantic school district, to be referred to by the pseudonym of Cityline Schools (pseudonyms are also used for all participants and schools in the study). In Cityline Schools, about 60% of students qualify for free and reduced price meals and approximately 55% of district students are African American and roughly 36% are Hispanic. This site was selected due to its early implementation of a new model of principal supervision. The district had reduced the span of control for principal supervisors to approximately 12–17 principals per supervisor to facilitate increased focus on developing principals' instructional leadership capacity. At the time of the study, there were 14 principal supervisors who had all been former principals in the district. This particular school district had also facilitated professional learning for most principal supervisors on learning-focused supervision (Lipton and Wellman, 2013) and had dedicated additional personnel and resources to supporting improvement in the district's schools with the greatest needs for improvement.

The purpose of this chapter is to look across the findings that emerged from a series of examinations of our data set to report on what we now know about principal supervisors' work with principals in this large urban district that may inform supervisors' work and district policies in other large urban centers that seek to reorient this role (Reyes, 2021; Thessin, 2019, 2021; Thessin et al., 2020). Broadly, our findings show that the goal of building the instructional leadership capacity of the school principal to lead improvement can be achieved by a skilled principal supervisor. Results and implications from our study can inform other urban districts striving to support principals in learning to lead instructional improvement in high-needs schools.

In the following sections, we will first provide a brief overview of the methodology of our study. We will then present and discuss each of the four findings we will articulate. Following our discussion of these findings, we will consider implications for educational leaders as they consider how to support principal learning and growth through the role of the principal supervisor. We will then end with a brief conclusion.

Methodology

We relied primarily on a qualitative multiple case study design for this project, as described in prior publications (Reyes, 2021; Thessin, 2019, 2021). For one study, we utilized quantitative methods to analyze results of a survey administered to instructional leadership team members (Thessin et al., 2020). All study procedures were reviewed and approved by the Institutional Review Board at the university where the research team members worked and by Cityline Schools' Office of Research. A brief overview of methods will be provided below.

Data sources

Principal supervisors and principals were purposefully recruited for participation in the study based on their work at schools designated by the district as *high needs* due to the district's focus on facilitating leadership development and school improvement at these schools. These included Title I schools, schools that had received School Improvement Grants from the state, and schools that had consistently low student test scores on second grade reading assessment indicators. We invited principal supervisors who were assigned to at least one of the high-needs schools in the district to join the study. After five principal supervisors consented to participate, we contacted and invited all principals whom they supervised who were leaders of high-needs schools. Twelve principals enrolled in the study, with each principal supervisor overseeing two or three principals who joined the study. A summary of the participants is included in Table 1.

Interviews

We conducted two 50–90 min interviews with each of the five principal supervisors (10 interviews total) and each of the 12 participating principals (absent one principal who elected not to complete the second interview, resulting in 23 principal interviews). The first interview took place when the participant joined the study, the second at the conclusion of the study. We designed and utilized semi-structured interview protocols that included questions on both the nature of principal supervisors' work with principals, such as the frequency and structures of their work, and the actual practices the principal supervisors utilized with principals. Principals

Table 1 Study participants: principal supervisor and principal pairs.

<i>Supervisor/principal</i>	<i>Years experience in education</i>	<i>Years experience in administration^a</i>	<i>Years in current position</i>	<i>School level</i>
Bill	33	17	4	(Supervisor)
Rory	25	11	1	Elementary
Tim	14	5	2	Elementary
Carmen	20	15	4	(Supervisor)
^b Jennifer	19	6	2	Secondary
^b Nancy	24	16	2	Secondary
^b Yora	20	13	2	Elementary
Jim	28.5	16	4	(Supervisor)
Mark	14	8	4	Secondary
Terri	20	11	7	Elementary
Sara	24	20	1.5	(Supervisor)
^b Samantha	19	9	1	Elementary
^b Paul	21	22	2	Elementary
Reagan	28	18	3	Elementary
Tammy	27	18	7	(Supervisor)
^b Carson	22	6	3	Elementary
^b Kara	21	10	3	Elementary

^aYears experience in administration may include years as an assistant principal, principal, principal supervisor, and/or another type of educational administrative position.

^bPrincipals who demonstrated evidence of engagement in joint work in this study.

and principal supervisors were asked about the practices they engaged in that contributed to principals' instructional leadership learning and development and about specific events and actions we observed.

Observations

We conducted 86 observations, ranging in length between 1 hour and a full day, totaling nearly 150 observational hours over the 16-month period of the study. We observed a variety of interactions between principal supervisors, principals, and school teams, including one-on-one meetings; team meetings; and central office network meetings that included members of the school-based instructional leadership team and central office administrators who supported their work.

Documents

To supplement our interviews and observations, we also collected documents related to supervisors' work with principals and school-based teams, including meeting agendas, feedback summaries, protocols, and tools used by principal supervisors and principals. The documents we gathered helped us to acquire a clearer understanding of the interactions between the supervisor and the principal at each school.

Surveys

To understand the relationship between principal supervisors' interactions with their instructional leadership teams (ILTs) and team members' perceptions of their schools' leadership and organizational conditions, we administered the Internal Coherence Survey (Stosich, 2014) in spring 2018 to ILT members in 10 schools who agreed to participate in this portion of our study (Thessin et al., 2020). The Internal Coherence Survey, which was previously piloted and validated (Elmore et al., 2014; Forman et al., 2017; Stosich, 2014), focuses on three domains of leadership practice that research strongly links to school improvement. The principals of the ten schools distributed a link to the online survey to their ILT members but had no other involvement in the survey data collection and were unaware of which ILT members completed the survey. The overall response rate for the survey was 74% (87 out of 117 possible participants).

Data analysis

Descriptive coding was appropriate for this analysis due to the numerous types of data being coded and the extended length of the study (Saldaña, 2013). Through each iteration of our data analysis process, new research questions arose for us that led us to review additional literature relevant to our new questions. We then reinitiated our codebook development by deriving deductive codes from our review of the body of literature relevant to our analysis. During our coding process, we paid particular attention to themes that emerged to identify new inductive codes and to instances in which codes overlapped. We strove to achieve intercoder agreement by having each member of the research team first independently code one-two observations or interviews at a time, after which we continually compared and revised the code definitions we had previously written (Saldaña, 2013).

We coded interview transcripts and all notes from observations using AtlasTI software, and all research team members wrote reflective memos following each observation and after coding each interview. We utilized documents we gathered as additional evidence to confirm, or negate, themes that emerged during data analysis.

To analyze our survey results, we examined the consistency of responses within the sections of the survey, each of which measured a separate construct of the Internal Coherence Framework (Stosich, 2014). We found a great deal of consistency within sections, with alphas ranging from 0.85 to 0.94 for the eight measured constructs; we therefore computed the mean score for each respondent for each section of the survey (Thessin et al., 2020). We then created tables and figures of means for each construct measured by our survey, both overall and based on categories of principal supervisor involvement with ILTs that emerged from our qualitative analysis. We tested for statistically significant differences between group means using simple linear regressions where the dependent variables were individual ILT members' ratings of the survey scales and the independent variables were indicator variables for these groups.

Findings

While interest in reshaping the principal supervisor role has grown in urban districts nationally in the last decade, support from outside foundations has also contributed to an increased interest in and attention to this role. The Wallace Foundation, for example, has been working to redefine the position of the principal supervisor to one focused on providing instructional leadership through the Principal Supervisor Initiative (PSI) since 2014. Goldring et al. (2018) articulated that in the third year of the PSI Initiative in six districts, principal supervisors commonly understood their role to be: "to support principals to improve schools" (p. 31). While the district in which this series of studies took place was not one of the six funded by the PSI, the intended outcome of refocusing principal supervisors' work on developing and supporting principals' instructional leadership across districts engaged in this work is ultimately to improve instruction in schools (Goldring et al., 2018).

From our 16 months of data collection and numerous analyses and reexaminations of our data, we concluded: (1) the qualities of the principal supervisor–principal partnership influenced the types of changes principals made to their instructional leadership practice through the support of the supervisor; (2) principal supervisors' and principals' engagement in joint work tasks provided the opportunity for principals' learning and practice development, and when these experiences were situated in a team context, other school-based leaders had the opportunity to benefit as well; (3) principal supervisors differentiated their support of principals through the development of partnerships and usage of contextual information; and (4) principal supervisors who engaged with school-based instructional leadership teams (ILTs) focused their interactions in specific areas, and the degree of these interactions related to ILT members' perceptions of the school's conditions for improvement. The work of building principals' instructional leadership capacity begins with the establishment of a productive principal supervisor–principal partnership.

Qualities of a productive principal supervisor–principal partnership

Background

Traditionally, the role of central office administrators has been to monitor the work of principals but not to engage in the work themselves (Honig, 2012). Principal supervisors are now being asked to facilitate learning about instructional leadership for principals as a coach, while still serving in a dual role as a supervisor, historically a role of overseeing, approving, and evaluating (Thessin, 2019). As the span of control for principal supervisors was reduced, Goldring et al. (2018) suggested that the change in the nature of interactions between the principal supervisor and the principal also changed the nature of their relationships.

A number of sources in the principal coaching literature and in the field of executive coaching highlight the need for an effective working relationship between the coach and coachee to facilitate changes in practice (Baron and Morin, 2009; de Haan et al., 2011; Huff et al., 2013; Natale and Diamante, 2005; Thessin, 2019). Robertson (2009) describes how leadership learning occurs through professional partnership in stating "Approaching leadership learning through the concept of partnership means entering the relationship as an equal learner" (p. 40). A professional partnership also requires trust to be productive (Baron and Morin, 2009; de Haan et al., 2011; Feldman and Lankau, 2005; Huff et al., 2013; Robertson, 2009; Swinnerton, 2007). In light of the expectations for both principals' instructional leadership learning and improved student outcomes, the new principal supervisor–principal relationship is best described as a partnership to which both the supervisor and the principal contribute.

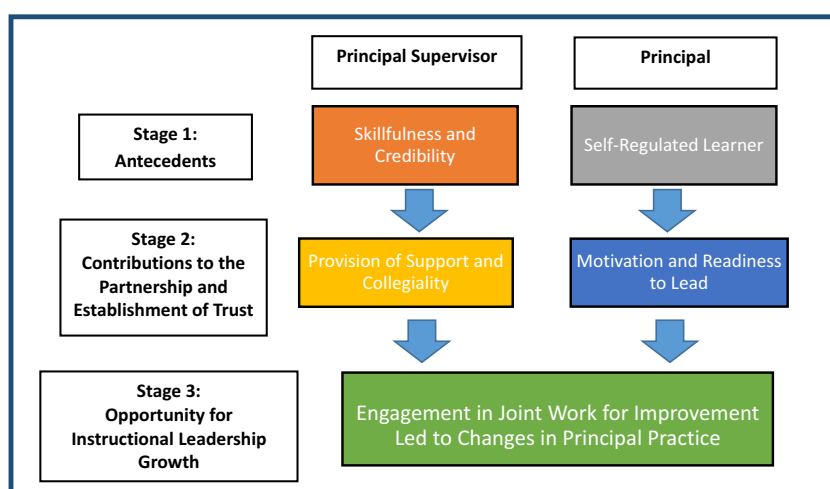
Summary of findings

As we engaged in our study, we noticed that some principals began to lead instructional improvement in new ways as result of the support of their principal supervisors, while others did not. We therefore examined how specific aspects of the principal supervisor–principal partnership affected the degree to which the principal grows as an instructional leader who can facilitate instructional improvement. We found that the types of changes principals made in their instructional leadership practice through the support of the principal supervisor varied from robust, to simple, to no change at all, with these degrees of change being influenced by qualities of the partnership (please see Table 2). We described productive partnerships as ones that generated *robust* changes in the principal's leadership practice, with robust being defined as changes to the way the school principal leads and builds instructional capacity in others (Thessin, 2019).

Table 2 Degrees of change by principals to their instructional leadership practice.

No change at all	Simple changes	Robust changes
Jennifer ^a	Paul	Yora ^a
Terri	Tim	Carson ^a
Samantha	Mark	Nancy ^a
Reagan	Kara ^a	
Rory		

^aThese principals participated as members of the principal supervisor–principal pairs in the study on principal supervisor differentiation practices.


Fig. 1 Antecedents and contributions to the principal/principal supervisor partnership that facilitate engagement in joint work for improvement.

Of the 12 participating principal supervisor–principal pairs, we identified three that met our definition of a productive partnership. For these three, the partnership developed in three stages: (1) recognition of antecedent characteristics; (2) contributions of support and collegiality, coupled with principals’ motivation to learn, led to the presence of trust in the relationship; and (3) engagement in joint work led to changes in principals’ instructional leadership practice (Thessin, 2019). Fig. 1 illustrates the stages of partnership development that occur.

Stage 1

As antecedents to the partnership, principal supervisors brought both skillfulness and credibility (Baron and Morin, 2009; Huff et al., 2013) as coaches of instructional leadership to their partnerships with principals. Principal supervisors also brought past experiences as effective principals to their partnerships and further established their own credibility by engaging in joint learning and joint work with principals at the inception of the partnership. Principals brought a readiness to be a learner to the partnership and generally aimed to improve their own competencies as self-regulated learners (Zimmerman, 2002). Principal supervisor skills, combined with the principal’s desire to learn, led to the continued development of a productive partnership in Stage 2 (Thessin, 2019).

Stage 2

Principal supervisors in productive partnerships dedicated time to providing support and establishing collegiality with their principals (Huff et al., 2013; Robertson, 2009; Wise and Hammack, 2011; Wise and Jacobo, 2010). Principals contributed motivation, a readiness to lead the work of improvement as communicated by clear goals, and a willingness to engage in critical reflection (Baron and Morin, 2009; Bozer et al., 2013). Principal supervisors supported the achievement of these goals, offered feedback, and established the pathway to build trust that, in turn, led to a shared commitment to improvement.

Stage 3

In this final stage, principal supervisors and principals engaged in joint work. We found that it was through the act of collaborating on joint work that principals learned how to refine their practice as instructional leaders. Principal supervisors in productive partnerships did not direct the work of improvement for the principal, but facilitated the work jointly with the principal, and sometimes with other school-based leaders. Moreover, these principal supervisors approached their work as equal partners who both had a stake in the school’s progress and in each student’s success.

In productive partnerships, principals made robust changes to their instructional leadership practice by directly engaging in the work of improvement collaboratively with the supervisor. These changes were not made in isolation, but were modeled by, planned collectively with, or facilitated with the supervisor's guidance and support at the school site. Our finding of the importance of joint work for principal's learning and practice development led us to investigate the forms of joint work that occurred across all 12 of the participating pairs in our study.

Engagement in joint work

Background

Our finding on the potential significance of engagement in joint work for principal learning and development is supported by a few other recent studies on joint work among educational leaders (Honig, 2012; Honig and Rainey, 2020; Thessin, 2019). Honig (2012) first identified joint work as a practice that supported principals' development as instructional leaders. Cosner et al. (2019) described joint work between aspiring principals and leadership coaches by articulating episodes of mutual engagement in practice that took place between leadership coach/aspirant pairs when conducting joint classroom observations. Most recently, Honig and Rainey (2020) included *joint work moves* as one of seven teaching and learning approaches in their framework on how principal supervisors work with principals to facilitate changes in principal leadership practices.

In our study, we found that principals made robust changes in their instructional leadership practice primarily through engagement in joint work (Thessin, 2019). We also noticed that principal supervisors engaged in joint work with principals in varied situations, with some supervisors engaging in joint work with principals in larger team-based settings. This finding is supported by a small body of literature on the benefits of situating leader coaching interactions in large-group settings (Cosner et al., 2018), particularly since instructional leadership is now thought of as a collective endeavor (Robertson, 2009; Spillane et al., 2001). We therefore reexamined our data to understand the types of key work tasks taken up jointly by the principal supervisor and the principal and the situational contexts in which these joint work experiences occurred (Thessin, 2021).

Summary of findings

Using Honig's (2012) work to inform our thinking and our analysis, we defined joint work as, "*engagement in the work of instructional leadership by both the principal supervisor and the principal*" and applied this definition to interview and observational data sources among the 12 principal supervisor–principal pairs in our study (Thessin, 2021). We identified evidence of engagement in joint work between principal supervisors and principals in 7 of the 12 participating pairs in our study, and therefore relied primarily on data from these pairs in developing our findings (please see Table 1). Five of these seven had been identified as making either robust or simple changes in their instructional leadership practices through the support of the principal supervisor. For the other two principals, we did not have sufficient evidence to conclude that they made changes to their practice, but they did have developmental opportunities through engagement in joint work (Thessin, 2021).

We found that principal supervisors and principals jointly engaged in specific instructional leadership tasks that provided opportunities for practice-based experiences and leader development, including: (1) development and use of tools; (2) classroom observations; and (3) planning, facilitating, and reflecting on leading teams. We also found that these three types of joint work tasks generally took place in four different situational contexts: (1) as joint work in a dyad; (2) as joint facilitation of learning in a team; (3) through the supervisor's leadership of learning in a team with the principal participating; and (4) as joint planning followed by the supervisor's leadership of team learning (Thessin, 2021).

Development and use of tools

We observed principal supervisors and principals develop or modify tools and protocols together in a dyad, and then jointly utilize them afterward in a team setting. Examples of these tools included a Below Grade Level Monitoring tool, the stop light protocol, and a data analysis protocol for upcoming PLC meetings, among others (Thessin, 2021).

Classroom visits

Principals described the value of engaging in classroom visits in a dyad with their supervisor, which contributed to their capacity to provide effective feedback to teachers and provided an opportunity for them to reflect on next steps in improving instruction with their supervisors. At another school, the principal supervisor observed and then coached the principal on learning-focused supervision practices following his provision of feedback to a teacher. Then the assistant principal (AP) joined the next classroom observation, and the principal supervisor supported the principal as he coached the AP following her provision of feedback to a teacher after the observation. Both the principal and the AP benefited from the learning opportunity facilitated by the supervisor in this instance (Thessin, 2021).

Planning, facilitating, and reflecting on leading teams

At four of the seven schools, we observed principal supervisors regularly engage in joint planning for and co-facilitating of team meetings with the principal for the instructional leadership team (ILT), administrative team, or collaborative grade level teams. At a number of these school sites, this cycle of learning and improvement included reflection between the principal supervisor and the principal, and sometimes a broader group of leaders, as well. Further, these learning opportunities were situated in the four different contexts we articulated earlier. Principal supervisors situated leader learning in the context of

school-based teams to grow not only the principal's capacity to lead, but also the capacity of other school leaders, allowing other leaders to try out new instructional leadership practices as well. As one example, one principal supervisor actually jointly planned upcoming school improvement meetings with a school's APs, which the principal recognized as a direct opportunity for her APs to benefit from her supervisor's knowledge as well. Principal supervisors and principals who engaged in this ongoing cycle of learning and development in team settings recognized that the difficult work of school improvement could not be led by the principal alone.

Our examination of principal supervisors' and principals' engagement in joint work demonstrated four key takeaways: the value of situated learning in building principals' instructional leadership capacity; the principal supervisor's role in supporting learning for school-based leaders other than the principal; the collective ownership for school-based improvement demonstrated by supervisors and principals engaging in joint work; and the opportunity for engagement in other coaching practices, such as modeling and facilitation of learning through experience, through engagement in joint work. Further, these outcomes would not be possible under the traditional conception of the principal supervisor's role in which the supervisor offered feedback after an observation of the principal and then did not return to the school for weeks, or even months. At four of the schools in this study, the frequent, ongoing sequence of planning, facilitating, and reflecting on leading teams led to opportunities for practice development for principals and for other school-based leaders (Thessin, 2021). Achievement of this goal, however, particularly in the context of high-needs schools, requires the principal supervisor to know each principal's and each school's needs well through the use of a differentiated approach.

Principal supervisor differentiation strategies

Background

As an individualized practice, coaching has recently grown in popularity as a potentially effective and popular approach to professional learning, and leadership coaching has an increasing presence in education. Effective coaching, by its nature, offers the opportunity to provide differentiated support to principals because coaches work individually with their coachees (Baron and Morin, 2009; Huff et al., 2013; Weathers and White, 2015) on real-life situations that they face (Beusaert et al., 2016). Principal supervisors now engage in job-embedded professional learning for principals through coaching (Grissom and Harrington, 2010; Reiss, 2012; Spanneut et al., 2012), offering skilled supervisors the opportunity to differentiate their support and supervision of principals to best meet principals' and schools' needs.

Honig (2012) first highlighted six teaching and learning strategies that central office-based principal supervisors used to support principal learning and growth. The strategies were: engaging in joint work (as described in an earlier section), modeling, developing and using tools, brokering through bridging (connecting) and buffering (blocking), and differentiation. The strategy of differentiation was described as one that was used to strategically combine other strategies to individualize work based on principals' strengths and needs. While Honig (2012) shared some examples of supervisor differentiation, no other literature has gone further to determine, classify, or describe how differentiation works in the principal supervisor's practice. Much like teachers differentiating for students in a classroom, principal supervisors become the agile connector of learning based on needs, the content, and the individual principal. Through the analysis of our data set, we sought to understand how principal supervisors differentiated their practice to effectively grow principal instructional leadership.

Summary of findings

We analyzed a subset of our data that included two principal supervisors and their work with five principals, allowing us to illuminate the myriad approaches utilized to support principal learning and growth. We found that these principal supervisors differentiated their support through developing partnerships and using context and principal needs (Reyes, 2021). In three of the five pairs, the principal demonstrated robust changes in instructional leadership practice, and in a fourth, the principal made simple changes based on our evidence (please see Table 2). In the last pair, however, there was no marked change in principal instructional leadership. In this pair with no evidence of change by the principal, though, the principal supervisor still provided learning opportunities for other members of the school ILT. The findings described below focus on the three partnerships where robust changes occurred.

Specifically, we found that principal supervisors in these three partnerships differentiated by developing partnerships through (1) building collegial relationships, (2) fostering trust, and (3) building relationships with others in each school. Additionally, we found that principal supervisors used context and principal needs when differentiating through (1) engaging in observations and feedback, (2) using school data and principal learning needs, and (3) implementing cycles of goal-setting, learning, and reflecting.

Developing partnerships

Principal and principal supervisors alike described and demonstrated partnership through collegial relationships. Three of the five principals described their supervisors as someone they trusted, a partner in the work, and part of their team. They acknowledged their supervisors were also responsible for their evaluation and yet described the relationship as feeling equitable and not "top-down." Principals noted that this was not a usual feeling among their fellow principals. Furthermore, they all described times when their supervisors stepped in to support them in a moment of need. The development of this collegial relationship acted as the foundation for principal supervisors to foster trust.

When building effective partnerships, principal supervisors fostered trust. Through trust, principals shared openly and honestly. A principal described it this way: “I would just say she’s been very open and transparent about her success and even some of her failures and I think that’s just made me very comfortable with having that kind of dialog and interaction with her.” In the three relationships where principals made robust or simple changes in their instructional leadership practices, the principals repeatedly described trust that their conversations were confidential with their supervisor. For example, two principals shared during interviews that they knew the supervisor would not use their communication of need for support during challenging situations punitively.

Principal supervisors also differentiated their support to other school-based leaders, thereby supporting principal learning and school improvement in areas where support was needed. For example, principal supervisors led professional learning sessions directly with assistant principals on teacher observation and feedback at one school, and with school-based instructional leadership teams on analyzing data at another site. The principal supervisors were not seen as “strangers” in the building, but were known by teachers and staff. This development of collegial relationships with other school-based leaders strengthened the principal supervisor’s partnership with the school as well as with the principal.

Use of context and principal needs

In order to differentiate their support to principals, principal supervisors utilized evidence gathered through observation. For example, principal supervisors used evidence they gathered through both planned and unplanned school visits to inform coaching conversations with principals. These visits allowed supervisors to see principals as they facilitated meetings, led learning walks, gave teachers feedback, and met with parents/guardians or community partners. After these observations, many supervisors provided both written and oral feedback to principals. In the feedback, they often praised the principal first and then offered constructive feedback on an area of growth that might guide their future collaborative work or support achievement of one of the principal’s expressed goals.

Additionally, principal supervisors used school data and principal learning needs to inform differentiated approaches to providing support. During observations, principals and principal supervisors analyzed or discussed various kinds of data. Data included: student reading and math achievement data through Partnership for Assessment of Readiness for College and Careers (PARCC) or Measures of Academic Performance-Reading (MAP-R), the perception of principal leadership by the community through Vanderbilt Assessment of Leadership in Education (VAL-ED), teacher evaluation data, and even climate surveys. By triangulating these various data points, principal supervisors determined learning needs and next steps in their future work together. Through use of formal and informal data, principal supervisors also brokered resources based on principal and school needs.

Finally, principal supervisors implemented cycles of goal-setting, learning, and reflecting. In order to engage in this process effectively, it must be based on evidence from the principal’s own practice, connected to principal standards, and relevant to the school context. To address all of these characteristics, a principal supervisor must differentiate. One supervisor described her approach: “I’m trying to make the connections. I’m trying to connect from the previous work and loop it around to what we did that particular day for it to make sense as to where we need to go.” Principal supervisors helped principals connect their goals to feedback, new learnings, and next steps as part of the coaching cycle.

Principal supervisors built the instructional leadership capacity of principals through a variety of differentiation strategies. Their ongoing use of observations, data analysis, and regular communication ensured that principal supervisors understood and utilized contextual information to tailor their work with each principal. Further, our findings that principal supervisors differentiated their support of ILTs in varied ways across schools raised new questions for us that we sought to examine more deeply.

Principal supervisor interactions with leadership teams

Background

While principals have a central role in leading school improvement, existing literature is also clear that principals cannot lead improvement on their own, particularly in high-needs schools. When principals and teachers share responsibility for instructional leadership, the quality of instruction improves, leading to gains in student learning (Leithwood and Jantzi, 2012; Marks and Printy, 2003; Printy et al., 2009). Principals who are effective instructional leaders not only improve teaching and learning in the classroom, but also build the capacity of the individuals they supervise and support to lead and sustain schoolwide instructional improvement efforts with them (Edwards and Gammell, 2016; Wenner and Campbell, 2017; York-Barr and Duke, 2004). Instructional leadership teams (ILTs), therefore, play an important role in school improvement efforts by engaging other school staff in improvement work (Portin et al., 2013).

In their new roles as partners with principals in school improvement efforts, principal supervisors also have the opportunity to contribute to principals’ growth as coaches for other school-based leaders, and to the growth and development of instructional leadership team (ILT) members directly (Thessin et al., 2020). Further, in the Model Principal Supervisor Standards (Council of Chief State School Officers, 2015), one of the central responsibilities of principal supervisors is to “help principals create distributed leadership systems and structures that support teaching and learning” (p. 19). Still, districts have not specified the principal supervisor’s role in supporting the work of school-based leadership teams. In light of the potential influence that principal supervisors could have on the work of ILTs and the principal’s own role in leading the ILT, a role of greatest consequence in high-needs schools, we sought to understand how principal supervisors interact with ILTs to establish the conditions for instructional improvement.

Summary of findings

Across one district, over our 16 months of data gathering, we observed principal supervisors take differing approaches to their interactions with school-based ILTs. We therefore analyzed our data sources, including a survey administered to ILT members, to explore the degree and focus of principal supervisors' interactions with ILTs. We also examined whether these interactions were related to the ILT members' perceptions of the leadership and organizational conditions that are conducive to school improvement (Thessin et al., 2020). As only ten school principals elected to administer the survey to their ILT members, data from these ten schools were analyzed for this portion of the study.

Degree of interactions

While some principal supervisors interacted directly with ILT members to support team members' learning and development, others discussed the role of the ILT in coaching conversations with the principal, but did not have direct interactions with ILT members. To analyze this variation in the degree of interactions, we grouped schools together in three categories: "minimal interaction" between the supervisor and ILT, "some interaction," and "a great deal of interaction." At four schools, the schools with "a great deal of interaction," the principal supervisor interacted directly with ILT members by collaborating with the school principal (or other school leaders) in planning ILT meetings, modeling facilitation of ILT meetings, and leading professional learning at ILT meetings, among other actions. At these schools where principal supervisors engaged in a great deal of interactions with the ILT, principals identified their supervisors as members of the team with whom they collectively owned the work of improvement (Thessin et al., 2020).

Focus of interactions

We found that principal supervisors' interactions with principals regarding the work of the ILT, and with ILT members directly, fell within three categories of the internal coherence framework (Elmore et al., 2014): (a) leadership for learning, (b) professional development, and (c) support for teams. (a) We observed leadership for learning with the ILT when principal supervisors visited classrooms with members of the ILT, discussed observations, and planned feedback to provide to the presenting teachers. At times, these learning walks were also joined by the district's Deputy Superintendent (Thessin et al., 2020). (b) In schools with a great deal of principal supervisor involvement in ILT work, principal supervisors often designed, facilitated, and co-facilitated professional learning opportunities for ILT and administrative team members. These professional learning experiences included how to utilize new data tools, analyze data for the purpose of differentiating instruction, engage in school improvement processes, form effective teams, observe classroom instruction, and provide feedback to teachers. (c) In some cases, supervisors provided support for teams by interacting with grade level or content area teams with the principal and with one or more members of the ILT. In four schools in which there was a great deal of interaction with the ILT, principal supervisors also engaged in a regular cycle of planning, facilitating, and reflecting on leading teams, as described previously (Thessin, 2021).

Relations of principal supervisor interactions to ILT members' perceptions

We found that the perceptions of ILT members of their schools' conditions for improvement differed depending on the degree of principal supervisor interactions with the ILT. ILT members' perceptions of their schools' leadership and organizational conditions for improvement at schools with "some interaction" and "a great deal of interaction" by the principal supervisor were more positive than ILT members' perceptions of the conditions for improvement in schools with "minimal" interactions by the principal supervisor. It is therefore possible that principal supervisors' interactions with ILT members contributed to their perceptions of the conditions for improvement, but further research is needed to understand this relationship.

Based on our findings, there are many potential benefits to the principal supervisor's interactions with school-based ILTs, including additional learning opportunities for principals and for ILT members and likely contributions to ILT members' perceptions of the school's readiness to engage in improvement practices. Still, principal supervisors interacted with ILTs in these potentially beneficial ways at only four school sites in our study, suggesting that there is a lack of understanding of whether, and when, the principal supervisor should be engaging in these ways with school-based ILTs.

Implications for practice and policy

Based on our study, there is clear evidence that principal supervisors who establish productive partnerships, engage in joint work, differentiate their support, and interact with school-based leadership teams contribute to changes in principals' instructional leadership practice, to the leadership growth of other school-based leaders, and to school improvement efforts (Reyes, 2021; Thessin, 2019, 2021; Thessin et al., 2020). To grow this impact, increasing numbers of urban districts are reexamining the role of the principal supervisor to ensure supervision is focused on building principals' instructional leadership capacity, as opposed to the traditional compliance and accountability-focused supervision model. Yet, this redesign process requires thoughtful planning and clarity of expectations at the district level, and additional guidance through the continued refinement of national standards. Even greater attention to the role of the principal supervisor is warranted for those principal supervisors who support high-needs schools, where students have not been served well by the educational system over time. Four specific implications for practice and policy at the school district level and at the national level are articulated below:

Prioritize selection and preparation

In order to change the traditional focus of the principal supervisor role, district leaders must first select principal supervisors who are ready to provide job-embedded professional learning to principals as a coach and partner in the work of school improvement. The selection process for principal supervisors should consider candidates' practices and abilities beyond their effectiveness in past roles as principals. Subsequently, individuals selected for the principal supervisor role should be given the professional learning and support they need to effectively build principals' instructional leadership capacities, beginning with those supervisors who support principals at high-need schools. Throughout the professional learning experience, focus should be paid to the principal supervisor standards, district expectations for the role, coaching strategies to provide ongoing learning for principals in their work, and the understanding that supervision is an ongoing process between the principal supervisor and principal. Further, those districts that have redesigned the principal supervisor role should ensure senior district leaders are prepared to model this type of support and supervision through their work with principal supervisors.

Communicate clear expectations

School districts need to provide not only ongoing professional learning to all principal supervisors who are selected for the position, but also clear guidance and expectations for their work. As we found in our study of 12 principal supervisor–principal pairs, the degree to which principal supervisors were successful in supporting principals' instructional leadership growth to facilitate improvement varied significantly across pairs. This wide variation in one district likely resulted at least partially from a lack of clear expectations for and supervision of the role of the supervisor. In this district, a key criterion for the selection of principal supervisors was a track record as a successful principal (as viewed by those directly involved in the principal supervisor selection process); yet, successful principals may not be effective principal supervisors, as we found in our study. Principal supervisors also shared that the district's evaluation rubric for their roles did not align with the district's new focus on coaching and instructional leadership support for principals. In the absence of clear expectations for changing their practice, leaders are likely to continue to lead as they have traditionally led in prior roles, roles which likely did not prioritize instructional leadership growth through the provision of support and job-embedded professional learning.

Establish thoughtful partnerships

Principal supervisor–principal partnerships should be established strategically and sustained over time. Partnering based on principals' needs and principal supervisors' abilities and strengths provides increased opportunity for robust changes in principal practice to occur. In order for principal supervisors to be able to work directly with the principal to meet learning needs and further support school teams to increase school improvement and effectiveness, there must be enough time and opportunity to know the principal and school context well. Through time and the building of partnerships, principal supervisors will be able to engage in more frequent, aligned, and strategic coaching cycles over time. Each cycle can enhance the principal's learning and further the school in its improvement efforts. This depth can increase only through continuous relationship- building and strategic coaching with the principal and members of the instructional leadership team.

Support learning of school-based instructional leadership teams

We assert that the title *principal supervisor* underestimates the power of this individual to facilitate learning that is consequential for practice development for principals and for school-based leaders more broadly (Thessin, 2021). Yet, principal supervisors have differing understandings of the degree to which they should interact with and support school-based ILTs, if at all. Clear guidelines from school district leaders to support not only the school principal, but also the increased instructional leadership capacity of ILT members, may help in this regard. By situating learning within the ILT, the principal supervisor can partner with the principal to establish the organizational conditions needed for improvement and build ILT members' skills to lead this challenging work. Distributing leadership of improvement also increases the possibility for instructional leadership practices to become embedded at the school and for improvements to continue even if principals change. Additionally, refining the Model Principal Supervisor Standards (Council of Chief State School Officers, 2015) to further specify the role of the principal supervisor in supporting the learning of school-based ILTs on a policy level may provide districts with additional leverage to make refinements at the local level (Thessin, 2021).

Conclusion

Research conducted over the last ten years has begun to show that the role of the principal supervisor is critical to principals' instructional leadership practice. Based on findings from our studies of one urban district's approach to a new model of principal supervision, supervisors have the potential to contribute to school improvement efforts, particularly in high-needs schools in which improvement is needed quickly, by facilitating both the principal's instructional leadership learning and the learning of other school-based leaders. As we found, however, principal supervisors bring a wide range of approaches to their role and to their

work to supervise and support school principals. Districts must therefore consider how to: align the hiring, professional learning, and evaluation of principal supervisors to the revised role; partner principals and supervisors strategically and sustain partnerships over time; and specify expectations for the revised principal supervisor role through policy at the district and national level.

As we have learned from urban districts engaged in this work, revising the principal supervisor role from a role focused on management and compliance to one focused on instructional leadership is a significant shift that the entire district must be ready to support. As more districts engage in this work, and as interest in shifting the principal supervisor role grows in suburban and rural districts as well, further research will be needed to understand how the supervisor's role can be most supportive to principals' growth as instructional leaders in other contexts.

Acknowledgment

Funding for this research project was generously provided by the Spencer Foundation. Findings presented in the article represent the author's work and may not represent the Foundation.

References

- Baron, L., Morin, L., 2009. The coach-coachee relationship in executive coaching: a field study. *Hum. Resour. Dev. Q.* 20 (1), 85–106.
- Beausaert, S., Froehlich, D., Devos, C., Riley, P., 2016. Effects of support on stress and burnout in school principals. *Educ. Res.* 58 (4), 347–365.
- Bozer, G., Sarros, J., Santora, J., 2013. The role of coachee characteristics in executive coaching for effective sustainability. *J. Manag. Dev.* 32 (3), 277–294.
- Corcoran, A., Casserly, M., Price-Baugh, R., Walston, D., Hall, R., Simon, C., 2013. Rethinking Leadership: The Changing Role of Principal Supervisors. <https://www.wallacefoundation.org/knowledge-center>.
- Cosner, S., Jones, M.F., 2016. Leading school-wide improvement in low-performing schools facing conditions of accountability: key actions and considerations. *J. Educ. Adm.* 54 (1), 41–57.
- Cosner, S., DeVoto, C., Rahman, A., 2018. Drawing in the school context as a learning resource in school leader development: application-oriented projects in active learning designs. *J. Res. Leadersh. Educ.* 13 (3), 238–255.
- Cosner, S., Walker, L., Swanson, J., Herbert, M., 2019. Elaborating Learning-Focused Leadership Coaching Activities Experienced by Aspiring Leaders During Principal Preparation [Unpublished Manuscript]. Department of Educational Policy Studies, University of Illinois.
- Council of Chief State School Officers, 2015. Model Principal Supervisor Professional Standards 2015. <https://ccsso.org>.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- de Haan, E., Culpin, V., Curd, J., 2011. Executive coaching in practice: what determines helpfulness for clients of coaching? *Person. Rev.* 40 (1), 24–44.
- Edwards, B., Gammell, J., 2016. Building strong school leadership teams to sustain reform. *Leadership* 45 (3), 20–22.
- Elmore, R., Forman, M., Stosich, E., Bocala, C., 2014. The Internal Coherence Assessment Protocol and Developmental Framework: Building the Organizational Capacity for Instructional Improvement in Schools. Research Paper. In: Strategic Education Research Partnerships.
- Feldman, D.C., Lankau, M.J., 2005. Executive coaching: a review and agenda for future research. *J. Manag.* 31 (6), 829–848.
- Finnegan, K., Stewart, T., 2009. Leading change under pressure: an examination of principal leadership in low-performing schools. *J. Sch. Leader.* 19.
- Forman, M., Stosich, E., Bocala, C., 2017. The Internal Coherence Framework: Creating the Conditions for Continuous Improvement in Schools. Harvard Education Press, Cambridge, MA.
- Goldring, E., Grissom, J., Rubin, M., Rogers, L., Neel, M., Clark, M., 2018. A New Role Emerges for Principal Supervisors: Evidence from Six Districts in the Principal Supervisor Initiative. Wallace Foundation, New York. <http://www.wallacefoundation.org/knowledge-center>.
- Grissom, J., Harrington, J., 2010. Investing in administrator efficacy: an examination of professional learning as a tool for enhancing principal effectiveness. *Am. J. Educ.* 116 (4), 583–612.
- Harris, A., Chapman, C., 2004. Improving schools in difficult contexts: towards a differentiated approach. *Br. J. Educ. Stud.* 52 (4), 417–431.
- Honig, M.I., Rainey, L.I., 2014. Central office leadership in principle professional learning communities: the practice beneath the policy. *Teach. Coll. Rec.* 116 (040304), 1–48.
- Honig, M., Rainey, L., 2020. Supervising Principals for Instructional Leadership: A Teaching and Learning Approach. Harvard Education Press, Cambridge, MA.
- Honig, M.I., 2008. District central offices as learning organizations: how sociocultural and organizational learning theories elaborate district central office administrators' participation in teaching and learning improvement efforts. *Am. J. Educ.* 114 (4), 627–664.
- Honig, M.I., 2012. District central office leadership as teaching: how central office administrators support principals' development as instructional leaders. *Educ. Adm. Q.* 48 (4), 733–774.
- Huff, J., Preston, C., Goldring, E., 2013. Implementation of a coaching program for school principals: evaluating coaches' strategies and the results. *Educ. Manag. Adm. Leader* 41 (4), 504–526.
- Klar, H.W., Brewer, C.A., 2013. Successful leadership in high-needs schools: an examination of core leadership practices enacted in challenging contexts. *Educ. Adm. Q.* 49 (5), 768–808.
- Leithwood, K., Jantzi, D., 2012. Collective leadership: the reality of leadership distribution within the school community. In: Leithwood, K., Louis, K. (Eds.), *Linking Leadership to Student Learning*. Jossey-Bass, Los Angeles, CA.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22.
- Lipton, L., Wellman, B., 2013. Learning-Focused Supervision: Developing Professional Expertise in Standards-Driven Systems. Charlotte, VT.
- Marks, H., Printy, S., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Natale, S., Diamante, T., 2005. The five stages of executive coaching: better process makes better practice. *J. Bus. Ethics* 59 (4), 361–374.
- Portin, B., Russell, F., Samuelson, C., Knapp, M., 2013. Leading learning-focused teacher leadership in urban high schools. *J. Sch. Leader.* 23 (2), 220–252.
- Printy, S., Marks, H., Bowers, A.J., 2009. Integrated leadership: how principals and teachers share transformational and instructional influence. *J. Sch. Leader.* 19 (5), 504–532.
- Reiss, K., 2012. Be a Changemaker: 12 Coaching Strategies for Leading Professional and Personal Change. Corwin.
- Reyes, C., 2021. Principal Supervisor Differentiation Practices that Support Principal Professional Learning. D. Thesis. The George Washington University.
- Robertson, J., 2009. Coaching leadership learning through partnership. *Sch. Leader. Manag.* 29 (1), 39–49.
- Saldaña, J., 2013. The Coding Manual for Qualitative Researchers, second ed. Sage Publications, Thousand Oaks, CA.
- Spanneut, G., Tobin, J., Ayers, S., 2012. Identifying the professional learning needs of public school principals based on the interstate school leader licensure consortium standards. *NASSP Bull.* 96 (1), 67–88.

- Spillane, J.P., Halverson, R., Diamond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- Stosich, E.L., 2014. Measuring school capacity for improvement: piloting the internal coherence survey. In: Bowers, A., Barnett, B., Shoho, A. (Eds.), *Using Data in Schools to Inform Leadership and Decision Making*. Information Age Publishing.
- Swinnerton, J., 2007. Brokers and boundary crossers in an urban school district: understanding central office coaches as instructional leaders. *J. Sch. Leader.* 17 (2), 195–221.
- Thessin, R.A., Louis, K.S., 2019. Supervising school leaders in a rapidly changing world. *J. Educ. Adm.* 57 (5), 434–444.
- Thessin, R.A., Louis, K.S., 2020. The principal's role has changed. Is professional learning keeping up? *Learn. Prof.* 41 (2), 40–46.
- Thessin, R.A., Shirrell, M., Richardson, T., 2020. How do principal supervisors interact with leadership teams in high needs schools? *Plan. Changing* 49 (3/4), 173–201.
- Thessin, R.A., 2019. Establishing productive principal/principal supervisor partnerships for instructional leadership. *J. Educ. Adm.* 57 (5), 463–483.
- Thessin, R.A., 2021. Engaging in joint work with principals: how principal supervisors' joint facilitation of teams contributes to principals' practice development. *Leader. Pol. Sch.*
- Turnbull, B., Riley, D., MacFarlane, J., 2015. *Districts Taking Charge of the Principal Pipeline*. Wallace Foundation, New York.
- Weathers, J., White, G., 2015. Executive coaching of schools leaders in a mid-sized urban school district: learning of a model of effective practice leading small and mid- sized urban districts. *Adv. Educ. Adm.* 22, 191–222.
- Wenner, J., Campbell, T., 2017. The theoretical and empirical basis of teacher leadership: a review of the literature. *Rev. Educ. Res.* 87 (1), 134–171.
- Wise, D., Hammack, M., 2011. Leadership coaching: coaching competencies and best practices. *J. Sch. Leader.* 21 (3), 449.
- Wise, D., Jacobo, A., 2010. Towards a framework for leadership coaching. *Sch. Leader. Manag.* 30 (2), 159–169.
- York-Barr, J., Duke, K., 2004. What do we know about teacher leadership? Findings from two decades of scholarship. *Rev. Educ. Res.* 74 (3), 255–316.
- Zimmerman, B.J., 2002. Becoming a self-regulated learner: an overview. *Theor. Pract.* 41 (2), 64–70.

Changing the context of Australian ¹Indigenous education through leadership for learning “both ways”

Greer Johnson, Bev Flückiger, and Neil Dempster, Griffith University, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	337
Methodology	338
Findings and discussion: Four central premises to reshape learning within Indigenous schools and communities	339
Premise 1: A strengths-based approach to Indigenous education has the potential to reframe deficit assumptions	339
Premise 2: Working two ways or both ways across governments, systems, schools, families and communities creates new ways of leading and learning	340
Premise 3: Enriching children’s development and learning from conception improves lifelong general health, well-being and learning outcomes	341
Premise 4: An integrated service delivery approach to health and education supports children’s overall development, well-being and learning	341
Implications of the review: From premises to practice—reshaping Indigenous contexts through leadership for learning	342
Leadership for learning	342
Curriculum	343
Limitations of the review	344
Concluding comments	344
References	344

Introduction

Improving Indigenous children’s lives such that their health and education outcomes compare favorably with more mainstream peers has been an ongoing challenge for Western governments. We take Australia as a case since it exhibits many of the issues that impact the education of Indigenous children in post-colonial societies. In spite of significant policy reform and investment over a lengthy period, targeted improvements have been largely disappointing. In the Australian setting, since the implementation of the Closing the Gap strategy in 2008 which was intent on reducing disadvantage in the Indigenous Australian populations and measuring improvement, achieving sustained results has proven equally elusive. After 12 annual reports documenting change in seven targets, just two are on target to reach their goals: early childhood education (By 2025, increase the proportion of Aboriginal and Torres Strait Islander children enrolled in Year Before Fulltime Schooling [YBFS] early childhood education to 95%) and Year 12 attainment (By 2031, increase the proportion of Aboriginal and Torres Strait Islander people [age 20–24] attaining Year 12 or equivalent qualification to 96%; see <https://www.closingthegap.gov.au/targets>).

The latest Close the Gap report, “We Nurture our Culture for our Future, and our Culture Nurtures Us” (The Lowitja Institute, 2020), designed by the Lowitja Institute for the Australian federal government’s Close the Gap Steering Committee, was released in March 2020. It recognizes that poverty and economic disengagement have been barriers to success. However, there are impediments of a more fundamental and ideological nature, specifically in terms of governance reform. The 2020 Closing the Gap strategy calls for equal decision-making between government and Aboriginal people at local, regional, state and national levels. Following the advice of community elders is all important in changing the context in which Indigenous Australian people live. Increasingly, implementation of such policies is being recognized as a value-laden process which not only identifies and measures social problems but also produces and reproduces them through discourse (Dawson et al., 2020). Understanding how deficit discourses are produced and reproduced is essential to challenging them (Fogarty et al., 2018).

The Australian federal government’s 2020 national Agreement on Closing the Gap strategy commits all Australian governments to 16 targets, three of which are related directly to early childhood and school education. These are:

- To increase the proportion of Aboriginal and Torres Strait Islander children enrolled in early childhood education to 95% by 2025;
- To increase the proportion of Aboriginal and Torres Strait Islander children assessed as developmentally on track in all five domains (physical health and well-being, social competence, emotional maturity, language and cognitive skills, communication skills and general knowledge) of the Australian Early Development Census (AEDC) to 55% by 2031; and
- To increase the proportion of Aboriginal and Torres Strait Islander people (aged 20–24) with a Year 12 or equivalent qualification to 96% by 2031.

¹The term “Indigenous Australian” has been used in this article to represent Aboriginal and Torres Strait Islander peoples of Australia, while “Indigenous” refers to first nations peoples more broadly.

Writing for *The Conversation*, [Graham et al. \(2020\)](#) argue that although these targets are warranted, they miss the mark, particularly in the way “success” is to be measured. Success needs to be measured in Indigenous as well as Western terms, especially with respect to meeting educational targets.

Our literature review aims to find out more about “what does (not) work and why” in a range of Indigenous and other low-socioeconomic contexts for improving educational outcomes for children from birth to school leaving age, inside and outside school. In our development and research agenda over the past 10 years, use of the theory of “leadership for learning ‘both ways’” has proven beneficial in making a difference in Indigenous education contexts. It is therefore fitting that we have again employed the concepts inherent in leadership for learning both ways to our collection, analysis and discussion of review materials leading to a set of interrelated premises that hold unprecedented possibilities for empowering Indigenous families and communities to work with schools in positively reshaping children’s futures.

Briefly, the leadership for learning both ways concept was borne and enacted during the Principals as Literacy Leaders with Indigenous Communities (PALLIC) project, through the mechanism of core participants: Indigenous leadership partners working alongside mostly white school leaders in the school, with a remit to lead engagement with their communities in children’s literacy learning ([Johnson et al., 2014](#)). The expression “both ways” refers to the value placed on both Western and Indigenous knowledges and was first articulated by Yolŋu educators in North-East Arnhem Land in the late 1980s. A metaphor was used to explain it which likens the meeting of knowledge systems to the meeting of two bodies of water in a lagoon where salt and fresh water come together ([van Gelderen and Guthadiaka, 2019](#)).

The PALLIC project was supported by an explicit position on school leadership in general and shared leadership in particular. Compelling research evidence has shown that quality leadership makes a difference to children’s learning and achievement, particularly in challenging school environments ([Bishop, 2011](#); [Day et al., 2010](#); [Hallinger, 2011](#); [Leithwood et al., 2006](#); [MacBeath and Dempster, 2009](#); [Masters, 2009](#); [OECD, 2008](#); [Robinson et al., 2009](#)). Added to this is the need for open “intercultural space” ([Taylor, 2003](#)) in which leadership both ways (Western and Indigenous) is accepted and valued in Indigenous schools and communities ([Ober and Bat, 2007](#); [Priest et al., 2008](#)).

Making improvements in learning and achievement requires both ways partnerships inside and outside the school to share the leading of reading ([Bishop et al., 2011](#); [DEEWR, 2009](#); [Hallinger and Heck, 2010](#); [McNaughton and Lai, 2009](#); [Ober and Bat, 2007](#); [PricewaterhouseCoopers, 2010](#); [Priest et al., 2008](#)). In practice, the Indigenous leadership partners and the school leaders worked collaboratively within the schools to enrich their personal and professional partnerships and the effects they might have on children learning to read. That this occurred is visible across the research findings: Indigenous leadership partners were highly respected by their principals, their teachers and their Indigenous communities, while they themselves valued the opportunities presented by the role to build stronger partnerships both ways, not only with the schools but also with the children, their families and community members. The long-term intention was to teach and learn literacy through a shared Indigenous and Western knowledge base.

Methodology

There is no consensus on the standard structure of a narrative or traditional literature review. Therefore, we explain the methodology for locating and evaluating the literature first and then present the findings and discussion and the limitations of the review. Since 2010, our long-term strategy has been to build and progressively update a narrative review of literature, including relevant authored and edited books in the field as well as policy documents on the broader topic of Australian Indigenous education and relevant work from Canada and New Zealand. We have updated the broader literature review for the purpose of this article to provide a concise understanding of “What works and why?”. Subsidiary questions were used to critique the notions of “What works?” and “What does success look like from a variety of viewpoints?”

As a first step we examined our current database using keywords to identify the most relevant texts: Indigenous children’s learning outcomes; success; teaching and learning in Indigenous schools; leadership for learning in Indigenous schools; remote Indigenous contexts; Indigenous family, school, community engagement; community empowerment. Similar keywords were used to guide scholarly electronic database searches in identifying further research papers published in English language peer-reviewed journals. Databases employed were Scopus, Science Direct, PsychInfo, ProQuest, Web of Knowledge, Sage, Google Scholar and Google. Additional information was generated from reference lists and citation records of papers found in database searches.

The most recent search has included work—including recent policy documents—published in Australia and internationally between 2015 and 2020, because this timeframe is rich with critique of the current concepts and practices driving Indigenous education policy and practice in the Australian context. In this instance we have not tried to review all of the relevant literature in this paper. Instead, we have drawn out specific concepts that support/substantiate our premises.

A close examination of these documents has highlighted the understanding that “what works” is not just a matter of listing a range of initiatives that have shown promise in strengthening Indigenous students’ learning outcomes and futures, but that it is also important to delve deeper to distil the philosophical and ideological drivers underpinning them. We found no publications that reviewed the literature on these terms.

The next section presents the literature review findings and discussion as a central set of interrelated premises that, when shared by stakeholders, hold unprecedented possibility for reshaping the contexts of leadership for learning in Indigenous contexts—for a more diverse and equitable notion of success.

Findings and discussion: Four central premises to reshape learning within Indigenous schools and communities

Based on a concise narrative review of contemporary research and evaluation reports of initiatives implemented in Australia, Canada and New Zealand, we found four interrelated premises that are integral for all school leaders and other stakeholders intent on enacting change in the context of Indigenous children’s learning outcomes. The premises, in combination, represent a positive set of beliefs, values and ideologies that systems, policymakers, principals, teachers, and Indigenous elders, families and children must share if Indigenous children are to gain greater access to educational opportunities and the outcomes that are forthcoming from achievement during the early years and across their years of schooling and into further study and work.

Premise 1: A strengths-based approach to Indigenous education has the potential to reframe deficit assumptions

A strengths-based approach to improving Indigenous children’s health and learning outcomes is an overarching premise that needs to drive policy reform, new initiatives and programs. It represents a significant shift in mindset that has the power to change the context in which Indigenous children live and learn. However, the notion of whose strengths are valued has been a longstanding point of contention that has impeded change to date. In the main, policies and practices that aim to improve Indigenous children’s learning outcomes have embodied a deficit mindset that assumes Indigenous communities lack strength to impact positive change (e.g., [Sarra, 2011](#)). Indigenous educators have repeatedly called for a reconceptualization that acknowledges strength on all sides. Such an approach respects Indigenous culture, acknowledges the inherent strengths of families, communities and children, and recognizes that Indigenous children’s experiences growing up are culturally informed and shape their behavior and attitudes to education, health and life more generally. Within the 2020 Close the Gap Report, the narratives of Aboriginal and Torres Strait Islander people provided a strengths-based analysis of Aboriginal and Torres Strait Islander health and well-being. They explain further:

... a strengths-based approach does not overlook the relationship of power and inequality on health outcomes and the responsibility of governments in leading systems reform. The work of communities must be matched by governments through shared decision-making, equitable (and sustained) funding and a cohesive policy approach.

The Lowitja Institute (2020, p. 5).

In Australia, [Priest et al. \(2008\)](#) warned that too many educational interventions and initiatives take a deficit approach to Aboriginal and Torres Strait Islander parents and communities: a key factor known to impede progress of schools and disenfranchise families partnering to improve children’s learning outcomes. In a deficit approach the potential learning opportunities and existing accomplishments of Indigenous Australian parents and communities are ignored in favor of imposing strategies and initiatives that aim to fix the perceived shortfalls with a top-down approach to governance. Others have similarly argued that a deficit model constrains a genuine working partnership between government, school leaders, educators and Indigenous communities ([Beresford and Partington, 2011](#); [Fogarty et al., 2018](#)).

At the heart of a strengths-based approach to improving learning outcomes for Indigenous children is the need for policymakers to be more sensitive to the cultural beliefs and child-rearing practices of Indigenous families and communities ([Lohoar et al., 2014](#)) and to build continuities as children transition across home, care and educational settings ([Kellard and Paddon, 2016](#)). The underlying cultural values and attitudes that inform child-rearing practices in Indigenous homes, however, may often be unseen and unexplained ([Geia et al., 2013](#)), so there is a need to find ways to build cross-cultural understandings as a prerequisite for policy reform. The lack of understanding about the potential incongruence between the culture of an early childhood education setting and the cultures of Indigenous families and groups may lead to non-Indigenous educators undermining Indigenous family and culture, and to general racism in early childhood settings ([Grace and Trudgett, 2012](#)). This signals the need for new approaches to Indigenous education that facilitate place-based understandings and reform.

It is broadly acknowledged that the way all young children grow up is culturally constructed ([Krakouer et al., 2018](#)). Cultural values and beliefs influence the attitudes, expectations and emotional orientation of parents and caregivers in relation to the child-rearing and development of their children ([Muir and Bohr, 2019](#); [Wise and Sanson, 2000](#)). In considering the values, expectations and orientations that influence Indigenous children growing up, it must be remembered that Indigenous Australians are diverse ([McTurk et al., 2011](#)), and also that the cultural beliefs and values that inform Indigenous child-rearing practices vary significantly from Western traditions ([Grant and Yerlo, 2011](#)). [Wise and Sanson \(2000\)](#) argued that “the early childhood developmental goals valued by parents and educators in many Western societies reflect an underlying ‘individualistic’ cultural script” (p. 3). This individualist focus is characterized by the acquisition of competence, independence and competition. In contrast, [Wise and Sanson](#) suggest that traditional cultural groups usually value more collectivist or interdependent cultural scripts. In these groups, collective goals such as living in harmony with one another, competent participation in social events, and obedience to authority are valued more highly. It is reasonable to argue that many Western-based reforms to Indigenous education are generated from a focus on individual competence and therefore fail to utilize the strength of community-based knowledge and competencies that Indigenous families and children have to offer the learning process in schools. The Longitudinal Study of Indigenous Children ([Penman, 2006](#)) provides further insight into children’s experiences and the child-rearing practices within Indigenous families and communities

across Australia. The study revealed that Indigenous Australians believe that culture, respect and family harmony are the fundamental elements of family life, child-rearing and kinship maintenance. Penman explained:

... culture consists of complex systems of obligations and responsibilities that bind all members of the kinship system through a strong oral tradition. Respect reflects the absolute belief that all people are worthy of respect and it is right that others will pay them due regard. Family harmony is premised on the principle of reciprocity—a mutuality of expectations. (p. 35)

In the same study, summaries of reports on Indigenous attitudes to children and child-rearing practices for people living traditional or semi-traditional lifestyles and those in urban environments included the following common themes: children are not viewed as helpless and in need of adult-imposed routines; children are kept close and their needs met as they arise; and children are cared for by the whole extended family. Taylor (2011) noted that the skills of autonomy, sibling and peer solidarity, motor skills, visual-spatial skills and the capacity to self-judge and take risks are encouraged in Indigenous children from an early age. McTurk et al. (2011) suggested that many of these distinctive characteristics of Australian Aboriginal family systems demonstrate early learning. Yet such an approach to learning is very seldom acknowledged in schools (Kellard and Paddon, 2016).

Many of the initiatives examined do not recognize the powerful insights, knowledge and cultural identities that Indigenous children bring to their schooling (Lohoar et al., 2014), nor do they recognize that learning for all children takes place in homes and communities from birth, as well as in schools. The differences in the goals and expectations of caregivers, which often mediate the daily experiences of Indigenous children and their interactions with people, objects and symbols in their immediate environment (Rosenthal, 2000; Santoro et al., 2011), must be considered in children’s learning if learning outcomes are to be profiled transparently with appreciation (Harrison et al., 2019; Parkinson and Jones, 2018).

School principals, educators and Indigenous leaders are in a position to build effective two-way communication practices and to create protocols for respectful interactions among educators, school officers and parents and carers as a means of understanding the learning support roles each is undertaking: to challenge the deficit context that often constrains Indigenous participation in schooling. However, while relationships that are mutually reinforcing, respectful and directed toward improving children’s learning are integral to the success of initiatives, they are often difficult and slow to establish in particular communities and in systems more widely (Kendall et al., 2008).

Indigenous children are dependent on informed, sensitive educators in all care and education settings to ensure experiences are socially inclusive. This will only be achieved through community engagement, respect for culture, and genuine working partnerships with communities and between schools, homes and the broader community: the hallmarks of a strengths-based approach (Kellard and Paddon, 2016).

Premise 2: Working two ways or both ways across governments, systems, schools, families and communities creates new ways of leading and learning

Working two ways or both ways involves genuine consultation between Western-styled governments, systems, schools, families and communities and Indigenous people. The terms “two ways” and “both ways” are used interchangeably (Disbray, 2016). Broadly, both terms indicate the possibility for Indigenous people to develop strong intercultural identities supported by Indigenous and Western knowledges. It is widely understood that the two-ways or both-ways Indigenous inspired methodology (Robertson et al., 2017) is essential for understanding how Indigenous and non-Indigenous people might learn from each other (Smith, 2011). As van Gelderen and Guthadiaka (2019) state,

“bothways”, by definition, represents the legitimacy and power of different theories and epistemologies working together, whilst retaining their distinctiveness. However, “bothways” also represents the position that in the process of negotiating knowledge practices, new ideas and indeed, worlds are born. (p. 4)

These authors liken both-ways methodologies to transculturation, an exchange between two cultures in which something is always given in return for something that is received. In these exchanges, both cultures are active, contributing and cooperating, resulting in knowledge practices being modified and new, transformed realities emerging. In the forming of such partnerships, Christie and Grotorex (2004) have pointed out that the positioning of “authority” in Indigenous communities must be taken into account. For example, unlike a top-down Western leadership style, Indigenous Australians work in the opposite direction, “with the group endowing authority on the basis of expertise” (Guilfoyle, 2016, p. 40).

In practice, a both-ways approach to educational reform is underpinned by an ideological alignment to a strengths-based approach to Western partnering with Indigenous people; that is, not assuming that Indigenous parents and children are educationally deficit and therefore needing to assimilate into the dominant education system in order to be “fixed”. D’Arbon et al. (2009, p. 19), in their project called “Linking Worlds: Strengthening the Leadership Capacity of Indigenous Educational Leaders in Remote Education Settings,” have explicated what is meant by the radically alternative educational ideology, first developed in the mid-1970s by Pincher Nyurmiyarri, a member of The Gurindji, who were concerned that their children were receiving a “one-way”

education. The Gurindji, from remote Northern Territory’s Victoria River region, proposed a two-way education system that involved “reciprocity and obligation, involving curriculum, knowledge, policies and power” (McConvell, 1982, p. 63). It also encompassed an approach to learning within The Gurindji which aimed to “re-establish relationships between the younger and older generations, healing rifts that had developed in the transmission of knowledge.” However, at this point “two-ways” became a “domain separation theory” which saw Indigenous people contributing knowledge and culture as an “add on” to dominant education, but with little opportunity to innovate and practice leadership.

There has been a recent move to a more socially critical intercultural model in which the conceptual and practical understandings of both ways is (re)presented as two overlapping circles. The intersection between the two circles demonstrates how the new methodology moves from a focus on systems and schools merely consulting Indigenous people about contributing Indigenous knowledge and content to educational curriculum (“both-ways”) to a new set of practices where there are viable possibilities for Indigenous people “to shape and develop the educational programs and to implement them” (Wunungmurra, 1989, cited in D’Arbon et al., 2009).

An authentic both-ways approach recognizes the strength of Indigenous people to work within systems alongside non-Indigenous people in school leadership roles, planning and implementing innovative curriculum, and increasing the possibilities for partnerships built on a platform of mutual respect, resolve and responsibility from the outset (van Gelderen and Guthadjaka, 2019).

Premise 3: Enriching children’s development and learning from conception improves lifelong general health, well-being and learning outcomes

Research in numerous disciplines has contributed to our understanding of the critical importance of education and care in early childhood and the role that day-to-day interactions with caregivers plays in brain development from conception (Lally and Mangione, 2017; Shonkoff and Phillips, 2000). Studies in neuroscience, for example, have identified that the quality of the early caregiving environment has a significant influence on the way in which the brain develops (Glaser, 2014; Morgan et al., 2018). Studies in psychology have confirmed the importance of sensitive parenting and attachment security for children’s social and emotional development (Bergström et al., 2020; van der Voort et al., 2014), and psychology and education studies have shown that a complex range of both distal (income, parent education) and proximal factors (quality of the early home learning environment) impact on the acquisition and development of literacy skills (Mathers et al., 2014; Webb and Williams, 2018). A child’s home and parents, therefore, are viewed as the main agents of influence during the early years.

Longitudinal studies in the USA have demonstrated the benefit of education and care programs in early childhood that support families and children during the early years. In particular, the Chicago Child Parent Centers (Reynolds et al., 2010), the Abecedarian intensive early childhood development project (Campbell et al., 2002; Page et al., 2019), and the Perry Preschool Project (Belfield et al., 2006) have found that high-quality early childhood education and care programs result in a higher rate of high school completion and lower rates of juvenile arrest and school dropout. In addition, the Perry Preschool Project has recorded lasting effects through to age 40 on employment rates and earnings (Belfield et al., 2006) and, more recently, has reported positive inter-generational effects lifting multiple generations out of poverty (Heckman and Karapakula, 2019).

In economic terms, the lasting effects of quality early childhood education and care produce a high cost-benefit. Jenkins (2014), along with Heckman and Karapakula (2019), explained that the marginal program cost for an individual child has been found to be significantly less than the collateral benefits to society. This is achieved by minimizing future costs to society in addressing key issues, for example, delinquency and dropout, while at the same time increasing individual resources in the form of education and earnings (including the earnings of future generations). In response, governments worldwide are investing resources during early childhood to promote these long-term individual and societal returns (Bishop and Berryman, 2010; Jenkins, 2014). While the provision of quality early childhood education and care is an important consideration for all children, investment in this area is seen as a key mechanism for closing the equity gap between Indigenous and non-Indigenous Australians (Council of Australian Governments [COAG], 2009; El Sayed, 2018).

A 2014 review of early childhood parenting, education and health initiatives for Indigenous children and families in Australia (Bowes and Grace, 2014) found that most Indigenous early childhood research in these areas was descriptive and tended to focus on identifying and describing “gaps.” As a result, the collective research story has been seen as one of deficits within the Indigenous community. Similar to Canada (Preston et al., 2012), it appears the societal norms and values of the dominant, white, middle-class culture determine what constitutes quality and best practice in relation to childrearing and educative practices. However, Bowes and Grace (2014) pointed out that “the Indigenous experience in Australia is one marked by strength and resilience as much as by inequity and disadvantage” (p. 2). It is by harnessing the strength and resilience of Indigenous Australian families and communities and providing opportunities for them to determine the co-design and implementation of early childhood programs based on contextualized realities that the early learning of Indigenous children can be best supported.

Premise 4: An integrated service delivery approach to health and education supports children’s overall development, well-being and learning

A report on the findings from the study *Obstacles to Success*, conducted for the Australian Primary Principals Association (APPA) by O’Keeffe et al. (2012), cited insufficient out-of-school support for Indigenous children as one of the key obstacles to children’s

learning outcomes. The report discussed the need to strengthen family and community practices to ensure children feel safe and are fed, sheltered and given relevant health services. These affordances are seen to work well when they are wrapped around parental and community valuing and support of education. O’Keefe et al. (2012) identified that “absence from school, including low rates of daily attendance, lateness and high mobility, was reported to be a prime obstacle to learning by 72% of participants” (p. 39). Masters (2013) reinforced the need to improve Indigenous health and well-being of families, carers and communities in parallel with strategies to improve Indigenous children’s learning outcomes.

An holistic approach to the health, well-being and education of children recognizes that early life outcomes are complex and influenced by a wide range of factors. It considers children and families in the context of their communities and cultures and takes into account physical and mental health, emotional well-being and development. Such an approach can be realized through integrated service delivery that attends simultaneously to health, well-being and education (Bowes and Grace, 2014; Emerson et al., 2015). An integrated service delivery model can improve service responsiveness to family needs, reduce service fragmentation, and improve service continuity (Robinson et al., 2009; Ware, 2012). In an analysis of the Vulnerable and Disadvantaged Client Access Strategies, Robinson et al. (2012) found collaboration between service providers and the delivery of “wrap-around” services were particularly successful in engaging with families and providing multilayered support in disadvantaged communities.

Holistic programs and services, delivered using a strengths-based approach that recognizes and builds on the expertise of parents, along with programs that offer learning opportunities for parents, have been found to be even more successful (Emerson et al., 2015; Sims, 2011). Sims (2011) explained that such an approach requires non-Indigenous policymakers and service deliverers to clearly understand their assumptions based on experiences with non-Indigenous early education and encouraged them to set about becoming competent in Indigenous culture.

In a relevant Australian project, Priest et al. (2012) developed a conceptual framework which privileges Aboriginal knowledge of child health and well-being in an urban context. The holistic framework consists of four main themes: strong culture; strong child; strong environment; and strengths and challenges. Critical elements of an Aboriginal cultural perspective identified by participants were identity, being proud and strong, kinship and family connection, respect for Elders, connection to Country, connection to family, connection to community, ceremony, art and artifacts, and they avowed that none of these elements should be tokenistic or superficial.

While the findings of the study cannot be generalized due to the urban context, Priest et al. (2012) stated that the findings “resonate with recommendations from research and discussions regarding Indigenous health” (p. 191). They highlighted the need for deeper consideration of culture in policy and practices related to Aboriginal child health and well-being and strongly supported an holistic, integrated and strengths-based approach to service delivery. Similarly, a recent scoping review of published literature in Canada, Australia, New Zealand and the United States assessed the quality of the evidence base on family-centered interventions for Indigenous early childhood care and well-being. The authors of this review (McCalman et al., 2017) found family-centered care with integrated service delivery to be a positive approach, generating diverse healthcare outcomes for Indigenous children and their parents/caregivers. It would seem, therefore, that a coordinated and integrated service-delivery approach among government, community-controlled agencies, and health and education sectors is needed. However, such a model must foreground capacity-building in Aboriginal communities that is developed through consultation with and co-designed with communities.

Implications of the review: From premises to practice—reshaping Indigenous contexts through leadership for learning

Collectively, the four premises discussed above provide an evidence-based platform for challenging the status quo of what is historically an assimilationist approach to Indigenous education. Contemporary approaches to equitable and respectful family-school-community relations, when based on the four premises, can help to reshape the context in which Indigenous children the world over learn in school and in their homes and communities. However, moving from premises to practice requires a dedicated practical attention over the longer term. We propose two key interconnected areas for reform: leadership for learning and curriculum.

Leadership for learning

In combination, the four premises foster multiple opportunities for leadership for learning beyond the formal positions available to school staff. In particular, we refer to recent work that is applicable to facilitating leadership and agency of non-dominant and Indigenous family and community, parents and carers in supporting children’s learning inside and outside schools. Many engagement efforts by schools never succeed in achieving true partnership with parents/carers (Ishimaru et al., 2016). To overcome these barriers, the engagement process must be co-designed with all parties from the outset. This involves collectively establishing respectful ground rules of engagement and then living them throughout the inevitable ups and downs. Once the journey toward partnership commences, the school can no longer unilaterally change these ground rules (Arnstein, 1969).

These partnerships are like any relationship—an ongoing, negotiated process. In the words of one educator, partnerships are “about sharing and co-constructing the school together ... and once they (parents) have a voice, and you have to listen to that voice” (Auerbach, 2010, p. 728). This can result in extensive debate, disagreement and conflict between parents, between the school and parents and with the community more broadly as the enactment of the partnership is (re)negotiated (Auerbach, 2009; Epstein,

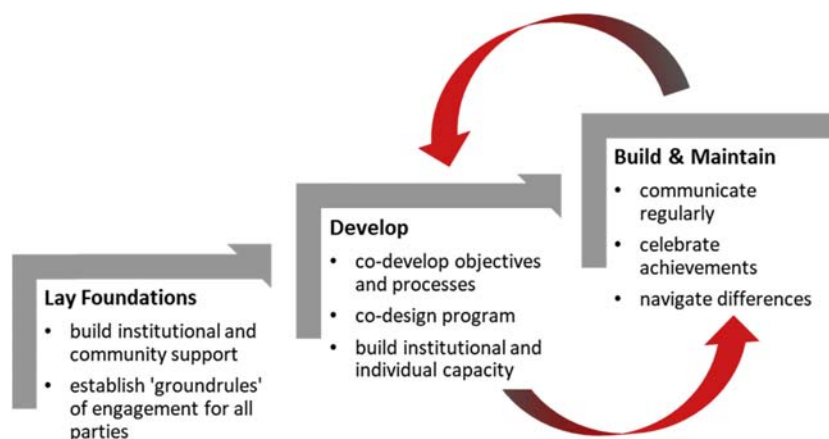


Fig. 1 Building blocks for authentic community-led engagement with schools.

2010). It is also important to recognize that engaging families and the community is a long-term undertaking involving a commitment to the mindset and process involved in shared leadership.

Fig. 1 breaks the engagement journey into broad stages, noting that individual communities and schools will progress at differing rates. The fundamental tasks involved in building such partnerships start with laying the foundations for engagement. The actions in this stage should be derived from first principles of participants and leaders—that the engagement of families, support for social justice and democratic participation are important because this is a moral imperative (Auerbach, 2009). Embedding these principles in symbolic and substantive ways will underpin all successive activities.

In the development stage, all parties co-develop the objectives and specifics of the engagement program. Building the participatory governance research, this stage will involve extensive discussion, deliberation and debate (Fung and Wright, 2001; Kania and Kramer, 2011). It is important at this stage to consider objectives of engagement that are broader than academic achievement for school children. For example, although improving children’s learning outcomes is the long-term aim, a focus on strengthening aspects of children’s health and well-being and Indigenous identity could well be the focus of a school-wide initiative in a remote Indigenous community. This broader perspective recognizes the school as part of the community and the potential for mutually beneficial engagement that extends beyond the classroom. In parallel to this, schools, Indigenous community groups and (grand)parents can explore where capacity to support greater engagement is needed and identify how this can be developed. It may, among many options, equally involve professional development for teachers to help them be “in relationship with and working alongside parents and families” (Pushor, 2012, p. 464) at childhood literacy workshops for parents (Riley and Webster, 2016; Saltmarsh et al., 2015).

The development stage and the subsequent implementation stage are recursive, reflecting the recognition that family-school-community engagement is iterative and ongoing. We argue that a more authentic measure of success involves a genuine attempt to value western and Indigenous ways of learning (Barr and Saltmarsh, 2014; Hanleybrown et al., 2012; Kania and Kramer, 2011; Yohani, 2013).

Curriculum

Closely linked to leadership for learning is the question of what counts as curriculum inside and outside Indigenous schools. A recent systematic literature review conducted by Harrison et al. (2019) found sharp contrasts between Indigenous and Western views of curriculum and learning. Whereas curriculum in the West is a body of knowledge that has been written and distributed by classroom teachers and by educational experts, Indigenous knowledge is learned through daily interaction with community “on country”.

In the main, Australian Indigenous children attend schools offering a Western national curriculum delivered in English which is not always the child’s first or perhaps even second language. Schools are led primarily by non-Indigenous principals and Indigenous staff are employed primarily in the role of teaching assistants to White qualified teachers. Although there are increasing efforts to incorporate Indigenous knowledges in the national Australian curriculum, the concepts of both-ways education are still in development.

Guenther et al. (2015) have critically examined what Aboriginal people from remote Australian cross-cultural communities thought about success in children’s education and how it might be achieved. Remote Indigenous Australian families reported three factors, linked to leadership for learning and curriculum, as integral to their children’s success in learning. First, high levels of importance are placed on family, school, and community engagement to support children’s learning. This factor emphasizes the empowerment of families to contribute to children’s learning inside school and at home: the importance of authentic partnerships where home knowledge is represented in the co-design of school curriculum. Second, is an urgent need for Indigenous teachers, especially local Aboriginal language teachers, to be equitably represented as permanent staff in community schools. Third, remote parents also

see the need for reflection and reform on the qualities of non-Indigenous teaching staff recruited from outside the community to work in remote schools. The qualities seen as key to successful teaching and learning included “flexibility, being friendly, kind, the teacher as a learner, being prepared for the environment, being respectful, patient, listening, passionate, having commitment and being dedicated to doing the best for the kids” (Guenther et al., 2015, p. 200). Underpinning these reforms is the implicit desire for a both-ways approach to learning that facilitates success in Indigenous and Western knowledge and identity.

Recent work by Bishop and Vass (2020) goes some way to explaining the disjunct between Western and Aboriginal thinking of “success” in terms of the cultural politics of Indigenous education. Culturally responsive approaches to curricula are not enough to make the difference that Indigenous people desire for their children. Bishop and Vass argue for the urgent need for systems change in the way teachers understand and work with Indigenous students and their families and communities. At the heart of the matter, pre-service and in-service teachers should be encouraged to question their ontology and axiology—that is, who they are and where they are from (Bishop and Durksen, 2020). In effect, no amount of curriculum reform using culturally responsive schooling approaches can make up for deficit approaches to pedagogy with Indigenous children and community.

In summary, changing the context in which Indigenous children learn at school and in their community is closely bound to the beliefs of school leaders, families and other stakeholders. Recognizing the need to deeply challenge deficit assumptions about the rights, responsibilities and capabilities of Indigenous school-families and communities as educators is paramount.

Limitations of the review

A significant limitation of the review is the high number of Australian studies utilized in it. However, we feel this is justified as the focus of this article relates to the Australian Indigenous context. Another limitation is the high proportion of qualitative studies included. We acknowledge that quantitative studies are important for gauging effect sizes; however, qualitative studies have helped us to understand the complex social issues from a variety of perspectives and to provide insight into these issues. As explained above, we have not tried to review all of the relevant literature in this paper. Instead, we have drawn out specific concepts that support/substantiate our premises. Yet, from the proportion of papers reviewed over a lengthy period, we are confident in the validity of the findings. While many of the findings may have implications for enhancing Indigenous children’s learning in other colonized countries, there is ready recognition that education policy and approaches to leadership and learning may differ across countries and regions.

Concluding comments

The momentum for changing Indigenous Australian policy and practice is growing (Dawson et al., 2020). Yet, responsibility for strengths-based change lies with us all through the sharing of new cultural frameworks and co-creating of equitable environments in which all forms of knowledge are valued, through a process of transculturation and shared leadership for learning. We argue that it is not just the effectiveness of Indigenous policy measures that needs to be refreshed: The way in which the problem is understood needs to be examined more critically.

The purpose of this article was to address the questions: “What works and why?” and “What does success look like from a variety of viewpoints?”. We focused this task on an Australian case study. However, a key message is that Australia’s Closing the Gap policy is a policy agenda that can be seen in other countries; therefore, the research and policy relevance of our four premises and two implications are applicable to the wider international education systems/context.

The four premises provide a means of critically examining some of the constraints engendered by Western colonial mindsets about Indigenous peoples’ rights and capabilities to be fully engaged in governance of their futures. Central to change for Indigenous children in particular is the need for a deep-seated shift by all Western systems, schools, families and communities from an entrenched deficit mindset about the rights and capabilities of Indigenous families and community to more positive territory. Strengths-based approaches and genuine engagement with Indigenous people may offer ways forward to provide policy measures that are based on Indigenous cultures and their people’s aspirations; how this will influence future Closing the Gap like strategies around the world is yet to be seen. It is clear that Indigenous Australian communities measure educational success as more than improvement in national literacy and numeracy scores sought at periodic times from primary to secondary schooling. We suspect this measure of success exists elsewhere.

References

- Arnstein, S.R., 1969. A ladder of citizen participation. *J. Am. Inst. Plan.* 35 (4), 216–224.
- Auerbach, S., 2009. Walking the walk: portraits in leadership for family engagement in urban schools. *Sch. Community J.* 19 (1), 9–31.
- Auerbach, S., 2010. Beyond coffee with the principal: toward a leadership for authentic school-family partnerships. *J. Sch. Leader.* 20, 728–757.
- Barr, J., Saltmarsh, S., 2014. “It all comes down to the leadership”: the role of the school principal in fostering parent-school engagement. *Educ. Manag. Adm. Leader* 42 (4), 491–505.
- Belfield, C., Nores, M., Barnett, W., Schweinhart, L., 2006. The High/Scope Perry Preschool Program: cost-benefit analysis using data from the age-40 followup. *J. Hum. Resour.* 42 (1), 162–190.

- Beresford, Q., Partington, G. (Eds.), 2011. *Reform and Resistance in Aboriginal Education: The Australian Experience*. University of Western Australia Press, Crawley.
- Bergström, M., Gebreslassie, M., Hedqvist, M., Lindberg, L., Sarkadi, A., Hjerm, A., 2020. Narrative review of interventions suitable for well-baby clinics to promote infant attachment security and parents' sensitivity. *Acta Paediatr.* 109 (9), 1745–1757. <https://doi.org/10.1111/apa.15212>.
- Bishop, R., Berryman, M., 2010. Te Kotahitanga: culturally responsive professional development for teachers. *Teach. Dev.* 14 (2), 173–187.
- Bishop, M., Durksen, T., 2020. What are the personal attributes a teacher needs to engage Indigenous students effectively in the learning process? Re-viewing the literature. *Educ. Res.* 62 (2), 181–198. <https://doi.org/10.1080/00131881.2020.1755334>.
- Bishop, M., Vass, G., 2020. Talking about culturally responsive approaches to education: teacher professional learning, Indigenous learners and the politics of schooling. *Aust. J. Indig. Educ.* 1–8. <https://doi.org/10.1017/jie.2020.30>.
- Bishop, R., Berryman, M., Wearmouth, J., Peter, M., 2011. Developing an effective education reform model for Indigenous and other minoritized students. *Sch. Effect. Sch. Improv.* 23 (1), 49–70.
- Bishop, R., 2011. Education leaders can reduce educational disparities. In: Townsend, T., MacBeath, J. (Eds.), *International Handbook of Leadership for Learning*. Springer, Dordrecht, NLD, pp. 1069–1081.
- Bowes, J., Grace, R., 2014. Review of early childhood parenting, education and health intervention programs for Indigenous children and families in Australia. Issues paper no. 8. In: *Produced for the Closing the Gap Clearinghouse*. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies.
- Campbell, F., Ramey, C., Pungello, E., Sparling, J., Miller-Johnson, S., 2002. Early childhood education: young adult outcomes from the abecedarian project. *Appl. Dev. Sci.* 6 (1), 42–57.
- Christie, M., Greateorex, J., 2004. Yolngu life in the Northern Territory of Australia: the significance of community and social capital. *Asia Pac. J. Public Adm.* 26 (1), 55–69.
- Council of Australian Governments (COAG), 2009. National Indigenous Reform Agreement: Intergovernmental Agreement on Federal Financial Relations. Department of Families, Housing Community Services and Indigenous Affairs, Canberra.
- D'Arbon, T., Fasoli, L., Frawley, J., Ober, R., 2009. *Linking Worlds: Strengthening the Leadership Capacity of Indigenous Educational Leaders in Remote Education Settings*. Batchelor Press, Batchelor, NT.
- Dawson, J., Augoustinos, M., Sjöberg, D., Canuto, K., Glover, K., Rumbold, A., 2020. Closing the gap: examining how the problem of Aboriginal and Torres Strait Islander disadvantage is represented in policy. *Aust. J. Soc. Issues.* <https://doi.org/10.1002/ajs4.125>.
- Day, L., Hanson, K., Maltby, J., Proctor, C., Wood, A., 2010. Hope uniquely predicts objective academic achievement above intelligence, personality, and previous academic achievement. *J. Res. Pers.* 47 (4), 550–553. <https://doi.org/10.1016/j.jrp.2010.05.009>.
- Department of Education, Employment and Workplace Relations (DEEWR), 2009. Early Years Learning Framework (Research Report). <http://deewr.gov.au/early-years-learning-framework>.
- Disbray, S., 2016. Spaces for learning: policy and practice for Indigenous languages in a remote context. *Lang. Educ.* 30 (4), 317–336.
- El Sayed, S., 2018. What will it take to close the gap in the Northern Territory? *Bedrock* 23 (2), 14–15.
- Emerson, L., Fox, S., Smith, C., 2015. *Good Beginnings: Getting it Right in the Early Years*. The Lowitja Institute, Carlton South, Melbourne.
- Epstein, J.L., 2010. School/family/community partnerships: caring for the children we share. *Phi Delta Kappan* 92 (3). <https://doi.org/10.1177/003172171009200326>.
- Fogarty, W., Lovell, M., Langenberg, J., Heron, M.-J., 2018. *Deficit Discourse and Strengths-Based Approaches: Changing the Narrative of Aboriginal and Torres Strait Islander Health and Wellbeing*. The Lowitja Institute, Melbourne.
- Fung, A., Wright, E.O., 2001. Deepening democracy: innovations in empowered participatory governance. *Polit. Soc.* 29 (1), 5–41.
- Geia, L.K., Hayes, B., Usher, K., 2013. Narrative or yarning/aboriginal storytelling: towards an understanding of an Indigenous perspective and its implications for research practice. *Contemp. Nurse* 46 (1), 13–17. <https://doi.org/10.5172/conu.2013.46.1.13>.
- Glaser, D., 2014. Neurodevelopment in the first three years—implications for child development, professional practice and policy. *J. Child. Serv.* 9 (2), 154–164.
- Grace, R., Trudgett, M., 2012. It's not rocket science: the perspectives of Indigenous early childhood workers on supporting the engagement of Indigenous families in early childhood settings. *Australas. J. Early Child.* 37 (2), 10.
- Graham, A., Nutton, G., He, V., 2020. There are 3 New Closing the Gap Education Targets: Here's What They Miss. *The Conversation*. <https://theconversation.com/au/education>.
- Grant, E., Yerlo, W., 2011. Whyalla Aboriginal Child and Family Centre: Indigenous Design Considerations. Report to the Department of Transport, Energy and Infrastructure, South Australia. http://digital.library.adelaide.edu.au/dspace/bitstream/2440/67062/2/hdl_67062.pdf.
- Guenther, J., Disbray, S., Osborne, S., 2015. Building on “Red Dirt” perspectives: what counts as important for remote education? *Aust. J. Indig. Educ.* 44 (2), 194–206.
- Guilfoyle, D.H., 2016. *Developing a Both-Ways Management Education Pedagogy from the Key Motivations of Aboriginal Australians*. Unpublished Thesis. Charles Sturt University, Australia.
- Hallinger, P., Heck, R.H., 2010. Leadership for learning: does collaborative leadership make a difference in school improvement? *Educ. Manag. Adm. Leader* 38 (6), 654–678.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142.
- Hanleybrown, F., Kania, J., Kramer, M., 2012. Channeling change: making collective impact work. *Stanford Soc. Innovat. Rev.* 1–8. ssir.org/tags/Collective+Impact.
- Harrison, N., Tennent, C., Vass, G., Guenther, J., Lowe, K., Moodie, N., 2019. Curriculum and learning in Aboriginal and Torres Strait Islander education: a systematic review. *Aust. Educ. Res.* 46, 233–251. <https://doi.org/10.1007/s13384-019-00311-9>.
- Heckman, J.J., Karapakula, G., 2019. Intergenerational and Intragenerational Externalities of the Perry Preschool Project. IZA Discussion Papers No. 12363. Institute of Labor Economics (IZA), Bonn.
- Ishimaru, A.M., Torres, K.E., Salvador, J.E., Joe Lott, I., Cameron Williams, D.M., Tran, C., 2016. Reinforcing deficit, journeying toward equity: cultural brokering in family engagement initiatives. *Am. Educ. Res. J.* 53 (4), 850–882.
- Jenkins, J.M., 2014. Early childhood development as economic development considerations for state-level policy innovation and experimentation. *Econ. Dev. Q.* 28 (2), 147–165.
- Johnson, G., Dempster, N., McKenzie, L., Klieve, H., Fluckiger, B., Lovett, S., et al., 2014. Principals as Literacy Leaders with Indigenous Communities: Leadership for Learning to Read—“Both Ways”. The Australian Primary Principals Association, Canberra.
- Kania, J., Kramer, M., 2011. Collective impact. *Stanford Soc. Innovat. Rev.* 9 (1), 36–41.
- Kellard, K., Paddon, H., 2016. Indigenous Participation in Early Childhood Education and Care—Qualitative Case Studies. Report Prepared for Department of Education and Training. The Social Research Centre, Canberra ACT, Melbourne.
- Kendall, L., Morrison, J., Yeshanew, T., Sharp, C., 2008. The Longer-Term Impact of Creative Partnerships on the Attainment of Young People: Report from 2005 and 2006. Research Report. http://www.nfer.ac.uk/publications/CPY01/CPY01_home.cfm.
- Krakouer, J., Wise, S., Connelly, M., 2018. “We live and breathe through culture”: conceptualizing cultural connection for Indigenous Australian children in out-of-home care. *Aust. Soc. Work* 71 (3), 1–12. <https://doi.org/10.1080/0312407X.2018.1454485>.
- Lally, J., Mangione, P., 2017. Caring relationships: the heart of early brain development. *Young Child.* 72 (2), 17–24. <https://www.jstor.org/stable/90004118>.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven Strong Claims About Successful Leadership*. National College for School Leadership, Nottingham.
- Lohar, S., Butera, N., Kennedy, E., 2014. Strengths of Australian Aboriginal Cultural Practices in Family Life and Child Rearing. CFCA Paper 25. <https://aifs.gov.au/cfca/publications/strengths-australian-aboriginal-cultural-practices-fam/summary-aboriginal-strengths>.
- MacBeath, J., Dempster, N. (Eds.), 2009. *Connecting Leadership and Learning*. Routledge, London.
- Masters, G., 2009. *A Shared Challenge: Improving Literacy, Numeracy and Science Learning in Queensland Primary Schools*. ACER, Melbourne.
- Masters, G., 2013. *Reforming Education Assessment: Imperatives, Principles and Challenges*, vol. 57. Australian Education Review. <http://research.acer.edu.au/aer/12/>.
- Mathers, S., Sylva, K., Eisenstadt, N., Soukakou, E., Ereky-Stevens, K., 2014. Supporting early learning for children under three: research and practice. *J. Child. Serv.* 9 (2).

- McCalman, J., Heyeres, M., Campbell, S., Bainbridge, R., Chamberlain, C., Strobel, N., Ruben, A., 2017. Family-centred interventions by primary healthcare services for Indigenous early childhood wellbeing in Australia, Canada, New Zealand and the United States: a systematic scoping review. *BMC Pregnancy Childbirth* 17, 71. <https://doi.org/10.1186/s12884-017-1247-2>.
- McConnell, P., 1982. Supporting the two-way school. In: Bell, J. (Ed.), *Language Planning for Australian Aboriginal Languages*. Alice Springs: ALA/AD, pp. 60–76.
- McNaughton, S., Lai, M.K., 2009. A model of school change for culturally and linguistically diverse students in New Zealand: a summary and evidence from systematic replication. *Teach. Educ.* 20 (1), 55–75.
- McTurk, N., Lea, T., Robinson, G., Nutton, G., Carapetis, J.R., 2011. Defining and assessing the school readiness of Indigenous Australian children. *Australas. J. Early Child.* 36 (1), 69–76.
- Morgan, S.E., White, S.R., Bullmore, E.T., Vértés, P.E., 2018. A network neuroscience approach to typical and atypical brain development. *Biol. Psychiatry* 3 (9), 754–766.
- Muir, N.M., Bohr, Y., 2019. Contemporary practice of traditional Aboriginal child rearing: a review. *First Peoples Child Fam. Rev.* 9 (1), 53–165.
- Ober, R., Bat, M., 2007. Paper 1: both ways: the philosophy. *Ngoonjook* 31, 64–86.
- Organisation for Economic Co-operation and Development (OECD), 2008. *Improving School Leadership*. Volume 1: Policy and Practice. OECD Publishing.
- O’Keefe, K., Olney, H., Angus, M., 2012. *Obstacles to Success: Indigenous Students in Primary Schools*. The Australian Primary Principals’ Association (APPA), Canberra.
- Page, J., Cock, L.M., Murray, L., Eadie, T., Niklas, F., Scull, J., Sparling, J., 2019. An abecedarian approach with Aboriginal families and their young children in Australia: playgroup participation and developmental outcomes. *Int. J. Early Child.* 51, 233–250. <https://doi.org/10.1007/s13158-019-00246-3>.
- Parkinson, C., Jones, T., 2018. Aboriginal people’s aspirations and the Australian curriculum: a critical analysis. *Educ. Res. Pol. Pract.* <https://doi.org/10.1007/s10671-018-9228-4>.
- Penman, R., 2006. The “Growing Up” of Aboriginal and Torres Strait Islander Children: A Literature Review. Occasional Paper No. 15. Department of Families, Community Services and Indigenous Affairs. http://www.dss.gov.au/sites/default/files/documents/05_2012/op15.pdf.
- Preston, J., Cottrell, M., Pelletier, T., Pearce, J., 2012. Aboriginal early childhood education in Canada: issues of context. *J. Early Child. Res.* 10 (3), 3–18.
- PricewaterhouseCoopers, 2010. *Creativity, Culture & Education: The Costs and Benefits of Creative Partnerships (Research Report)*. Retrieved from. <http://www.creativitycultureeducation.org/wp-content/uploads/PWC-report-the-costs-and-benefits-of-creative-partnerships.pdf>.
- Priest, K., King, S., Nangala, I., Nungurrayi Brown, W., Nangala, M., 2008. Warri Jarrinjaku “working together everyone and listening”: growing together as leaders for Aboriginal children in remote central Australia. *Eur. Early Child. Educ. Res.* 16 (1), 117–130.
- Priest, N., Mackean, T., Davis, E., Briggs, L., Waters, E., 2012. Aboriginal perspectives of child health and wellbeing in an urban setting: developing a conceptual framework. *Health Sociol. Rev.* 21 (2), 180–195. <https://doi.org/10.5172/hesr.2012.21.2.180>.
- Pushor, D., 2012. Tracing my research on parent engagement: working to interrupt the story of school as protectorate. *Action Teach. Educ.* 34 (5–6).
- Reynolds, A., Temple, J., Ou, S., 2010. Preschool education, educational attainment, and crime prevention: contributions of cognitive and non-cognitive skills. *Child. Youth Serv. Rev.* 32 (8), 1054–1063.
- Riley, T., Webster, A.A., 2016. Principals as Literacy Leaders with Indigenous Communities (PALLIC) building relationships: one school’s quest to raise Indigenous learners’ literacy. *Teach. Educ.* 27 (2), 136–155.
- Robertson, F., Barrow, J., Wajrak, M., Nannup, N., Bishop, C., Nannup, A., 2017. Participatory action and dual lens research. *Qual. Res. J.* 17 (4), 283–293.
- Robinson, V., Hohepa, M., Lloyd, C., 2009. *School Leadership and Student Outcomes: Identifying What Works and Why*. Best Evidence Synthesis Iteration. Ministry of Education, Wellington: New Zealand.
- Robinson, E., Scott, D., Meredith, V., Nair, L., Higgins, D., 2012. Good and Innovative Practice in Service Delivery to Vulnerable and Disadvantaged Families and Children. *Children Family Community Australia (CFCA) Paper No. 9*. <http://www.aifs.gov.au/cfca/pubs/papers/a142861/cfca09.pdf>.
- Rosenthal, M.K., 2000. Home to early childhood service: an ecological perspective. *Child. Issues* 4 (1), 7–15.
- Saltmarsh, S., Barr, J., Chapman, A., 2015. Preparing for parents: how Australian teacher education is addressing the question of parent-school engagement. *Asia Pac. J. Educ.* 35 (1), 69–84.
- Santoro, N., Reid, J.-A., Crawford, L., Simpson, L., 2011. Teaching Indigenous children: listening to and learning from Indigenous teachers. *Aust. J. Teach. Educ.* 36 (10). <https://doi.org/10.14221/ajte.2011v36n10.2>.
- Sarra, C., 2011. *Strong and Smart—Towards a Pedagogy for Emancipation: Education for First Peoples*. Routledge, New York.
- Shonkoff, J.P., Phillips, D.A., 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academy Press, Washington, DC.
- Sims, M., 2011. *Early Childhood and Education Services for Indigenous Children Prior to Starting School*. Produced for the Closing the Gap Clearinghouse. Australian Institute of Health and Welfare and Melbourne: Australian Institute of Family Studies, Canberra.
- Smith, B., 2011. Community involvement. In: Craven, R. (Ed.), *Teaching Aboriginal Studies: A Practical Resource for Primary and Secondary Teaching*, second ed. ProQuest Ebook Central, pp. 194–209. <https://ebookcentral-proquest-com.libraryproxy.griffith.edu.au>.
- Taylor, R., 2003. An indigenous perspective on evaluations in the inter-cultural context: how far can one throw a Moree boomerang? *Eval. J. Australas.* 3 (2), 44–52.
- Taylor, A., 2011. Coming, ready or not: Aboriginal children’s transition to school in urban Australia and the policy push. *Int. J. Early Years Educ.* 19 (2), 145–161.
- The Lowitja Institute, 2020. *We Nurture Our Culture for Our Future, and Our Culture Nurtures Us*. Close the Gap Campaign Steering Committee for Indigenous Health Equality.
- van der Voort, A., Juffer, F., Bakermans-Kranenburg, M., 2014. Sensitive parenting is the foundation for secure attachment relationships and positive social-emotional development of children. *J. Child. Serv.* 9 (2).
- van Gelderen, B., Guthadiaka, K., 2019. Renewing the Yolŋu “bothways” philosophy: Warramiri transculturation education at Gäwa. *Aust. J. Indig. Educ.* 1–11. <https://doi.org/10.1017/jie.2019.2>.
- Ware, V., 2012. Improving access to urban and regional early childhood services. Resource sheet no. 17. In: Produced for the Closing the Gap Clearinghouse. Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies, Canberra.
- Webb, G.L., Williams, C.J., 2018. Factors affecting language and literacy development in Australian Aboriginal children: considering dialect, culture and health. *J. Early Child. Res.* 16 (1), 104–116.
- Wise, S., Sanson, A., 2000. Child care in cultural context: issues for new research. *Aust. Inst. Fam. Stud.* 22, 1–24.
- Wunungmurra, W., 1989. Dhawurrpunaramirr: finding the common for a new Aboriginal curriculum. *Ngoonjook* (2), 12–16.
- Yohani, S., 2013. Educational cultural brokers and the school adaptation of refugee children and families: challenges and opportunities. *J. Int. Migrat. Integrat.* 14 (1), 61–79.

Leadership and policy: how principals of successful schools enact education policy for improvement

Qing Gu, UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	347
Connecting leadership with policy enactment	348
The research	349
Case studies of four secondary schools in England	350
Findings: how senior and middle leaders see their principals enact education policy for sustained improvement	
building capacity: using policy as impetus for change	350
The Shaw Lane girls academy: from good to outstanding	351
The contexts	351
Opportunity of purpose: anchoring core values	351
Opportunity of ownership: communication and consultation	352
Discussion and conclusion	353
References	354

Introduction

Schools in many countries across the world have undergone considerable reform over the past two decades and their principals have had to learn to manage increased volumes of government educational-policy initiatives designed to raise standards of teaching, learning, and academic outcomes for all students. Although these initiatives are seen by governments as a means of building human, economic, and social capital in increasingly competitive and socially turbulent global environments, there are continuing concerns over attempts by school leaders and teachers to implement them in ways that appear to be manifestly unlikely to work (e.g. [Payne, 2008](#); [Gu et al., 2021](#)).

It is perhaps then no surprise that many researchers and publications have explored the nature of educational reforms and their impact on schools and teachers over time. These, though rigorous in their approaches, tend to stress the negative consequences of reform on, for example, teacher morale, an increasing emphasis on the academic to the disadvantage of other humanistic areas of curriculum, and continuing problems of narrowing the achievement gap experienced by students from socio-economically disadvantaged communities. I, together with my fellow colleagues who have worked with and researched many passionate and inspiring teachers and school leaders over the last two decades, do not disagree with these generalities. They resonate with more general concerns about the ability of externally mandated reforms at system level ([Ainscow, 2015](#); [Elmore, 2004](#); [Fullan, 2016](#)) to provide a sustained impact upon the quality of teaching and learning and achievement without the active mediation of principals and teachers in the schools themselves (e.g. [Johnson, 2019](#); [Gu et al., 2020, 2021](#)).

However, an increasing body of research knowledge has identified improving and successful schools that do not compromise or sacrifice broader educational purposes, can beat the odds, and enable their teachers and students to achieve and thrive – both individually and collectively – in the face of considerable, continuing and at times, disruptive external policy demands ([Day et al., 2016](#); [Day and Leithwood, 2007](#); [Day et al., 2011](#); [Day and Gu, 2018](#); [Elmore, 2004, 2011](#); [Gu et al., 2016](#); [Leithwood, 2018](#); [Matthews et al., 2014](#)). The insight from such knowledge is that schools continuing to thrive in the reform process are “active agents” ([Hubbard et al., 2006](#), p. 14) who reshape policy initiatives into actions that are culturally, organizationally and educationally meaningful to their teachers and students in their daily realities. Put differently, how these schools respond to and do policy is “an act of co-construction” ([Hubbard et al., 2006](#), p. 14) and *enactment* in context, rather than *implementation* with fidelity.

The research upon which this chapter is based is about how successful principals lead their schools in times of intensive and pervasive policy reforms. The research does not, therefore, engage with the political and sociological debates about the pedagogical challenges, professional tensions, and ethical dilemmas that education policies and reforms have produced for school leaders and teachers (e.g., [Fuhrman, 1999](#); [Mitchell et al., 2011](#); [Payne, 2008](#)). Rather, we argue that the political, professional, and accountability pressures created by incoherent, disjointed, and at times contradictory external policy initiatives are part of the broad environments in which “schools and education policy subsist” ([Cohen et al., 2007](#), p. 526) in many systems including England, and within which schools in our research managed to continue to make a difference to students’ academic performance ([Gu et al., 2018](#)). Put simply, policy shifts have become unavoidable political realities of education in many systems. However, whilst some schools not only survive but also continue to thrive in the face of challenging and changing environments, others struggle and falter.

By using a longitudinal, mixed-methods design to investigate the interface of reform at macro (country), meso (school), and micro (classroom) levels, this UK research council funded three-year research investigated how government reforms (mandatory and non-mandatory) were received and mediated by principals and senior and middle leaders in improved and effective schools serving communities of contrasting socio-economic advantage. “Success” was defined as characterizing those schools that had shown sustained improvement in student academic outcomes over time (measured by pupil progress and attainment outcomes over a nine-year period), other key aspects of school improvement (measured by national inspection results), and reputation. In these successful schools, policies were conceptualized as “opportunities,” resources that leaders skilfully weaved into their processes of school improvement to create educationally equitable, and values-based “landscapes of success.”

Much of this research has been published in a special issue of *Leadership and Policy in Schools* (e.g. Day and Gu, 2018; Gu et al., 2018a,b). In this chapter I investigate, from the perspective of senior and middle leaders, the ways in which successful secondary principals in England had led their schools to achieve sustainable high performance in the face of intense reform efforts and policy shifts. Although the literature on successful principal leadership has become increasingly extensive over the last decade, knowledge about how key staff (i.e., senior and middle leaders in their schools) perceive the ways in which school principals make a difference to student outcomes tends to be limited. In particular, the literature has been relatively silent on how the key staff perceive the roles and practices of their principals in managing external policy demands and increased accountability pressures for sustained improvement over time.

Connecting leadership with policy enactment

Over the last 40 years, scholars have based their conceptualization of policy enactment and implementation on a range of critical sociological (Ball, 1994; Ball et al., 2012; Maguire et al., 2015), organizational (Elmore, 1995; Hubbard et al., 2006; Manning, 1982; Newmann, 1996), social psychological (Weick, 1979, 1995, 2005), and cognitive theories (e.g., Spillane et al., 2002; Spillane et al., 2002b). The critical sociological approach to examining how schools enact policy conceptualizes policy as text (Ball, 1994) in that it is “complexly encoded in sets of texts and various documents and it is also decoded in complex ways” (Braun et al., 2011, p. 586). Such a conceptual lens has allowed for an understanding of policy enactment as a creative, sophisticated, and complex process (Braun et al., 2010) in which “policies are interpreted and ‘translated’ by diverse policy actors in the school environment” (Ball, 1994, p. 19). Ball et al. (2012) argue that the enactment of policies is “an iterative process of making institutional texts and putting those texts into action” (2012, p. 45), and that enactment of policies is “always more than just implementation” because “they bring together contextual, historic and psychosocial dynamics into a relation with texts and imperatives to produce action and activities that are policy” (2012, p. 71). In their seminal research on how four “ordinary” coeducational, nondenominational, and nonselective secondary schools enact policy, Ball et al. (2012) observed that

At the center of policy enactment is the school—but the school is neither a simple nor a coherent entity, there is a need to understand schools as far more differentiated and loosely assembled than is often the case. Schools are not of a piece. They are precarious networks of different and overlapping groups of people, artefacts and practices. (2012, p. 144)

Indeed, the narrative accounts of how teachers and other adults from the four case-study schools interpreted, translated, and implemented various external policies in their own contexts of work showed that these policy actors are not only producers and consumer of policy, but also readers and writers of policy (Ball et al., 2011). This evidence highlights the localized and situated nature of policy actions (Braun et al., 2011). It also reveals that enacting policies in schools is a process of recontextualization in which policy actors work creatively in diverse ways to “fabricate and forge practices out of policy texts and policy ideas in the light of their situated realities” (Ball et al., 2012, p. 142).

Although writing from different theoretical perspectives to understand and explain how people in organizations and schools make sense of policies and implement them, Weick (1995, 2005) and Spillane (2004) have both emphasized the situated nature of sense making in the policy enactment process and the importance of considering how people make sense of their environments in this process. Thus, making sense of policies is not a passive process of decoding the information in the policy texts (Von Glasersfeld, 1989). Rather, in this process people as social agents “construct, rearrange, single out, demolish many objective features of their surroundings” (Weick, 1979, p. 164) and ultimately transform their environments (Spillane et al., 2002). Using the sense-making frame to examine how school leaders enact district-level accountability policies, Spillane et al. (2002) concluded that:

Managing in the middle in an era of accountability can also have advantages. Skillful school leaders can use accountability policies to augment their authority with respect to instruction... Hence, school leaders can interpret district accountability policies in ways that support their own reform agendas and use them to augment their influence over staff. Because the stakes are high, they can use district accountability measures to add considerable clout to their own efforts to transform practice in particular ways.

Spillane et al. (2002, p. 760).

The conceptual and empirical connections that Spillane and his colleagues have established between school leadership and policy enactment are important because schools’ responses to external policies are the result of the “function not only of leaders’

identities but also the multiple contexts in which their sense-making is situated" (2002, p. 755). They remind us that in schools, enacting policies is an organizational behavior which is crafted and shaped by school leaders, and principals especially, who set the directions of the school and can act to redesign the organization. How these leaders interpret and make sense, rationally and emotionally, of what a particular policy means to their schools and then decide "whether and how to ignore, adapt, or adopt" this policy locally (Spillane et al., 2002, p. 733) influences not only how the policy is interpreted by their teachers and how effectively it is implemented in the school, but importantly, the extent to which the actions of "enactment" are likely to disrupt, constrain, or advance further improvement of the school. In this sense, enactment links closely with the principal's role in their diagnosis of the school's needs and challenges and the focus on particular priorities in consequence. As our empirical research in this article shows, principals who do well know how to use policies as opportunities to create organizational conditions and regenerate school capacities for enhanced progress and performance which are not restricted only to academic attainment results.

In their analysis of how schools strategically manage multiple external demands, Honig and Hatch (2004) found that although some researchers argue that multiple policy demands in such environments strain schools' "ability to operate in coordinated and productive ways," others maintain that they may "add up to important new opportunities for school improvement" (Honig and Hatch, 2004, p. 16). By conceptualizing policy coherence as a dynamic and ongoing process, as opposed to an objective reality, they argue that schools are a central agent in crafting coherence between external demands and internal goals and strategies. In his earlier work, Hatch (2002) emphasized that schools are part of an ecosystem in which "many different entities are trying to co-exist" and in which "changes are constantly underway" (Hatch, 2002, p. 632). Honig and Hatch argue that "multiple external demands do not present a problem to be solved but an ongoing challenge to be managed, a potential opportunity for schools to increase necessary resources, and an important arena of organizational activity" (2004, p. 26–27). By extension, we argue that to create and embed coherence between policy and practice within the particular context of a school's organization requires strong leadership. We ascribe to the view of the superintendent in Hubbard et al.'s (2006) research on how schools learn from reforms that "coherence making at the end is what leadership is about" (2006, p. 157).

The research

The chapter draws upon empirical evidence from a UK Economic and Social Research Council (ESRC) and HK Research Grants Council (RGC) jointly funded bilateral project (ES/J017035/1) on how "successful" secondary schools in England and Hong Kong mediate government policies in furthering their own broad improvement agendas. The primary purpose of this research was to advance understandings of how mandated government reforms were mediated by principals, senior and middle leaders, and teachers in improved and effective schools which served communities of contrasting socioeconomic advantage. The research was guided by three broad questions:

- (1) How do leaders in successful secondary schools across different contexts respond to government systemic reforms?
- (2) What key challenges and issues do they face in sustaining academic standards for all whilst forwarding their broader educational success agendas?
- (3) How and to what extent do school leaders at all levels in these schools maintain a strategic and operational focus on the leadership of learning and teaching whilst managing wider structural and cultural changes?

The research began with a mapping exercise of the perceived challenges of current government reforms in Hong Kong and England and a secondary analysis of two existing national surveys on senior and middle leaders' perceptions of their principals' contribution to change and improvement in their schools. These surveys were part of two parallel three-year research projects led by the research teams which investigated the impact of school leadership on the improvement of students' academic and social outcomes in the wider educational reform contexts in Hong Kong and England, respectively (Day et al., 2011; Gu et al., 2008; Walker, 2011). By merging the datasets of the two surveys and taking into account contextual differences, this secondary analysis sought to develop an evidence-based model to analyze the influence of leadership at all levels on the implementation of systemic reform on the structures, processes, and outcomes of school and classroom improvement.

These analyses formed the basis for case studies of eight improved and effective secondary schools (four from the parallel study in each country) across diversified school populations in different socioeconomic contexts. A purposive sample of four secondary schools in each country was selected using the following criteria:

- (1) *School performance*: "improved" and "effective" schools. All schools were selected from the research samples of two parallel large-scale research projects on school leadership and change. The parallel studies involved ten secondary case-study schools across different contexts in England and Hong Kong respectively. Four from each country, which had shown significant improvement or higher than average improvement in their value-added scores in English (in England)/Chinese (in Hong Kong) and mathematics over the three-year period prior to the commencement of the projects, were invited to participate in the research.
- (2) *School contexts*: The indicator of free school meal bands in England and the indicators of the medium of instruction and school sponsoring body in Hong Kong were used to select four schools which serve communities of contrasting socioeconomic advantage/disadvantage.
- (3) *School size*: Ranging from small to large.

The main methodological advantage of studying these schools was that we were able to build rich, longitudinal accounts (2005–2014) of their values, strategies, and practices and thus ensure an in-depth investigation of this new research area. The case studies focused upon the ways in which reforms were mediated by principals and senior and middle leaders and teachers in order to assess the extent to which the primary intentions had been translated into practice and sustained and if not, why not. This helped us better understand the dynamics underpinning the effectiveness of the models of government reforms in England and Hong Kong and provide school-level, evidence-based analytical insights into the processes by which reforms were enacted.

Case studies of four secondary schools in England

The qualitative strand used four in-depth case studies of a subset of the original 10 secondary case study schools in the IMPACT study (Day et al., 2011). These four cases, together with the four in Hong Kong, were selected to enable in-depth, longitudinal analyses of leadership and change. All four schools showed significant improvement or higher than average in their value-added scores in English and mathematics at Key Stage 4 (age 16) in public GCSE examinations over the seven-year period prior to the commencement of this research (including the three years prior to the beginning of the IMPACT project). These case studies represented schools in different levels of socioeconomic advantage and disadvantage (as measured through the low-income indicator percentage of student intake eligible for Free School Meals) and ethnic diversity.

Data were collected through two visits over a two-year period with detailed interviews of principals and a range of key staff and stakeholders. The main methodological advantage of studying these schools was that we were able to build on and extend existing datasets from the IMPACT study (2005–2009) and thus achieve longitudinal narrative and reflective accounts of the challenges, processes, and outcomes of change and improvement that schools had experienced in mediating externally imposed structural and curriculum reforms across a period of nine years (2005–2014). By focusing on “success” (i.e., schools that had shown sustained improvement in academic performance over time and also were judged as Good or Outstanding by national inspections), we were able to identify effective principles and practices of leadership and educational practice for change and improvement in schools and classrooms. Through these, we extended our knowledge of how and to what extent some school leaders and teachers successfully mediate government models of systemic reform whilst maintaining the broad educational purposes of learning and teaching in their places of work.

Although this chapter focuses on case studies from England only, research findings from England and Hong Kong are broadly consistent in suggesting that school organizations are adaptive social systems in which school improvement must be understood both in contexts of policy enactment, contexts of use, and in connection with the values and capabilities of the principals and staff at all levels of the systems (Day and Gu, 2018). This is despite the differences in national cultures and education systems between England and Hong Kong in our research. In his postscript of our special issue, Leithwood (2018) concludes that ‘policy-driven school reform depends critically on school leaders’ responses to external policy initiatives’ (2018, p. 392):

The mediated effects on students of a set of practices commonly enacted by successful leaders in widely different contexts certainly count as an important finding but is not novel enough to count as an insight. What does count as an insight is that these same sets of practices may be successful in accomplishing a wide array of outcomes for students. ... Leaders sampled by the four articles were initially selected because they had been successful in helping their students achieve those “tool skills” measured by external government tests, among other things. But these principals and their staffs had more ambitious and varied goals for their students such as, for example, internationalism, entrepreneurialism, creativity, and more. (2018, p. 394)

Findings: how senior and middle leaders see their principals enact education policy for sustained improvement building capacity: using policy as impetus for change

The four case-study schools served communities of contrasting socioeconomic disadvantage and were led by principals with different years of experience in the school. However, despite differences in context and leadership history, what shone through the interviews with senior and middle leaders in all four schools was a strong, collective sense of positive leadership which embraced external policies and innovations as the catalyst for further growth and higher performance.

For the purpose of this chapter, I have selected a story of a secondary school that provides an example of how the principal led an already successful school to even higher performance. This principal had the shortest tenure of all the principals in this research, having only been in post for three years when the first data were collected. Her values and vision for excellence were perceived by senior and middle leadership as the key drivers for change. Shaw Lane was inspected within a term of her appointment and was judged to be a “Good” school. Since then, she had made considerable progress in addressing the areas of weakness the inspection had highlighted, most notably the teaching and learning agenda which had been radically overhauled via a range of strategies to foster more collaboration and creativity amongst teachers. We interviewed seven senior and middle leaders in the school and most were interviewed twice over this two-year project. This example is used to illustrate that for principals in our case studies, the key to success in enacting external policy demands is to use them as opportunities to develop and transform people—such that they share the same values and passion for further growth and development, and that they become change actors (as opposed to receptors) who possess enhanced knowledge, qualities, and capacity to regenerate the social and intellectual culture of the school, and importantly, to manage new changes, collectively and collaboratively, for unified goals and core purposes.

The Shaw Lane girls academy: from good to outstanding

The contexts

Shaw Lane, a single-sex school situated on the outskirts of a large urban area within the Midlands region of England, was a relatively small secondary school and sixth form that provided an education for approximately 800 students between the ages of 11–19. The school served a culturally homogenous population with the majority of students from a non-White heritage background and many speaking English as an additional language. The number of students eligible for free school meals was higher than the national average (46% vs. 13%), as was the proportion with a special educational need (75.9% vs. 16.2%) at the time the study was conducted.

The school was judged to be a Good school by the Office for Standards in Education (Ofsted) shortly after the principal's appointment in 2011. The inspection report drew attention to the good progress made by all students and the high-quality teaching and learning that took place across the school. It also highlighted the positive impact of the recently appointed principal within the context of the "legacy of underperformance" she inherited from her predecessor. By 2014, the school achieved an overall grade of Outstanding.

During this critical period of school improvement and renewal (2011–2014), four major new government initiatives were introduced to schools in England, reflecting a change of government in 2010—all of which demanded deep structural, financial, and cultural changes: performance-related pay and appraisal (Department for Education, 2013); a new Ofsted inspection framework; a review of the National Curriculum and change in the way school performance was measured; and pressure for the academization of all schools (which are funded by the Department for Education but independent of local-authority control). The principal's aspiration for Outstanding was proudly shared by her staff as a "powerful" aim and morally just vision for the school (Assistant Principal). Driven by such an aspiration, she used the enactment of these various external policies as an opportunity to raise expectations and anchor core values; to consolidate consultation and enhance ownership of change; and, above all, to build, broaden, and deepen the capacity required to lead further change and improvement.

Opportunity of purpose: anchoring core values

Deal and Peterson (2009) argue that "Central to successful schools is a powerful sense of purpose that is focused on students and learning" (2009, p. 250). Shaw Lane was no exception. All the senior and middle leaders reported that the school positioned itself ahead of educational policy and that they were cautious of being too reactive toward government reform. Managing change was seen pragmatically as the nature of the job: "When a policy comes in—if it's statutory—then it's got to be done" (Head of Mathematics). Because government policy initiatives were, more often than not, unpredictable and potentially transient in nature (seen as likely to shift with a change of government), it was believed that the priorities of the school had to be centered on moral purpose, upon doing:

...what is right for the students because if you do what is right for the students then you must be doing the right thing. ... At the end of the day, if you look at government policy it comes and it goes, doesn't it? But the needs of the students don't change that much; they might change in terms of the contexts they have but, fundamentally, what you're trying to do is to prepare them for a future in which they can be useful citizens.

Senior Deputy Principal.

Such moral compass defined a shared direction for the school and a strong commitment in the staff who wanted to move the school forward "to be more than Outstanding: making sure that whatever we are doing is for the benefit of the students" (Middle Leader—Humanities). What also came across consistently and powerfully here was how the shared purpose and direction had run deep to shape a confident attitude toward change and policy demands. Student needs rather than government priorities informed the way that the school was organized and operated:

I don't feel shackled by policy because I've got a lot of experience behind me and [the principal] wants Outstanding and so she is putting structures into place to achieve that; and that is why I don't feel shackled by it because if you are doing your job right then you have nothing to worry about. And we share a similar ethos which is about producing well-formed individuals and it's about us finding the best way for these girls in Shaw Lane to achieve and to be well-formed individuals. So [the principal] will do that her way.

Assistant Principal.

Associated with this positive mindset of change was a sense of positive, assertive, and responsible professionalism in the school. The ethical principles and standards that acted as drivers for improvement were to "do your best every time a student comes through" because "every student who comes to this school has only one chance" (Senior Deputy Principal). Importantly, these standards were defined, believed in, and pursued by the staff, manifesting themselves in the form of what Elmore (2003) calls "internal accountability":

At the end of the day, what you are trying to achieve for the students is the best possible exam results and qualifications that are going to enable them to go on and be successful in the future; and you want the students to have the qualities and the characteristics which will enable them to be productive citizens when they leave here. If you can do all of that then whatever accountability framework comes along it should meet it, shouldn't it? If you're doing what is right for the student, then it should meet whatever framework does come along.

Senior Deputy Principal.

Opportunity of ownership: communication and consultation

The principal and her senior leadership team created clear lines of communication to ensure that staff members were thoroughly consulted and well informed about the policymaking process at school. They were given ample opportunity to air their views on the ways in which the latest policy initiative would affect them and their work. This process, driven by the principal's insistence on communication and transparency, also ensured that staff members fully understood the policy in question and that they were given time to negotiate how it would be enacted on the ground in their departments and classrooms:

Mostly the senior team is very good and they will consult with middle management and we have a group set up where we will meet up at lunchtimes or after-school sessions where, if the senior team have got certain policies that they are going to implement, they will take our views on board before they implement the policy. So, there is that going on and we, as heads of department, meet regularly with the senior team and, of course, those issues are discussed there as well and heads of department will get a chance to have their say as well. So that level of communication will go on before that policy is actually put into practice, so it's not something that has just been sprung on us and there is a certain amount of discussion and communication that goes on beforehand.

Head of Mathematics Department.

There was also a forum for departmental heads to discuss the implications of policy changes through their monthly curriculum meetings before returning to their respective faculties to feedback these discussions to their teaching staff and consult them on any changes that might have been proposed. The following senior leader explained how the organizational structure had facilitated communication and the sharing of information, which helped them to manage change in practice:

ICT is no longer a topic in itself—it's changed to computer science—and we were looking at that yesterday and whether we can do an ICT audit across our curriculum because every subject is responsible for teaching ICT. This is what the government is saying and we need to adjust to these changes. So, it's in curriculum meetings and it's in faculty meetings as well because when we go back we discuss it as a faculty on the necessary changes that are taking place. Even the changes in the courses for GCSE now are fed down from senior management and then passed through so that everyone is aware of what is going on across the school.

Acting Deputy Principal.

It is perhaps, then, no surprise that a strong sense of ownership and collective loyalty was shared in the school. This was seen as a sign of an open, cohesive, and trusting culture that the principal and her senior leadership team had regenerated. Shaw Lane was described by middle leaders in particular as a "happy place" where staff were treated well (Head of Humanities) and "a listening school" where "I've got a voice": "You've got the freedom to say what you are happy with and what you feel might need to be improved" (Head of English). Making policies was regarded as "a whole-school thing" (Head of Humanities), as the decision-making process involved open discussions with staff members at every level in a very democratic way. As a result, there was a high level of consistency in behavior across the school: "Policies we have integrated are whole school policies. It's almost like nobody deviates from the norm and we are all doing the same thing: the marking policy is the same; the data collection is the same; behavior policy is the same" (Head of Humanities). Such consistency in behavior and vision helped to deepen the coherent and cohesive culture in the school, which made further growth and improvement possible.

Opportunity for capacity building: Focusing on the basics Improving teaching and learning was seen as "the bread and butter of what we do" (Assistant Principal), and was believed to be "a positive drive" (Head of Music) that had turned the school around:

That is all about capacity building and it's all about developing the staff to have the skills and knowledge that they need and that is obviously going to feed through to the students and lead to them getting a better experience.

Senior Deputy Principal.

Put simply, to embrace change effectively in the school required capacity building. Key in this regard was quality professional development. This was because, at least in part, the policy "is actually saying that this is the direction that we think you ought to go, and this is the sort of path we want you to take. But how you walk along the path—'how the garden grows'—is actually up to you" (Senior Deputy Principal). Thus, knowing how to make sense of external policy and recontextualize it in ways that were fit for purpose required sustained attention to improve the knowledge, skills, and practices of the staff on the ground.

The newly introduced Ofsted inspection framework, for example, was used as a vehicle to raise the standards of provision of teaching and learning in the school. This was achieved through the provision of a series of in-house training events designed to ensure that the staff had a thorough grasp of this framework. At the same time, there was an “Open Door Community” (Assistant Principal) in the school where the staff felt “safe” to share practice and discuss what outstanding teaching and learning looked like. After some initial worries, the Head of Humanities realized that the new framework had not really had much impact on her “because our lessons are well-planned and we are well-resourced anyway.”

Similarly, when enacting the new performance-related pay and appraisal policy in the school, efforts had been focused on how this policy could be used to join up with support for learning and development, and through this, foster a professional culture of high expectations.

We’ve also looked at the idea that performance management is something that underpins what happens throughout the whole school. So your performance management should be tied to your CPD and your school development plan, and so every department, having identified what the department needs are to fit in with the whole school development plan, is able to create a training plan for the department and, therefore, the CPD will actually support individual staff in specific CPD that will help them meet department needs, their own needs, and whole-school needs.

Senior Deputy Principal.

The leadership intention was well-received by their middle leaders: “they’ve [Senior Leadership Team] really pushed the idea that performance management is not a whip for us to be beaten with and it’s meant to be about self-development” (Assistant Principal). The quote from the Head of Mathematics below expressed a similar view, which represented not only a common voice from the interviews, but also, and importantly, a testimony to the success of the principal leadership in that she had shaped a high degree of consistency in values, expectations, and behavior across the school.

I think colleagues understand the need for accountability. You would think that there are very few colleagues who are in that position where they are bordering on being incompetent, if you like. The majority are hardworking and conscientious people who want to do right by the kids. So, the idea of the appraisal system is to support them to improve and it’s just about evidencing what you do day in and day out. That’s what needs to happen really: it’s finding out the people who are perhaps dragging their heels and making them more aware and more accountable and getting them to perform to their best.

Head of Mathematics.

Discussion and conclusion

A central concern of this research was to explore whether, and in what ways, various senior and middle leaders had a consistent understanding of how their school principals purposefully incorporated and embedded, rather than were passively led by, external policy demands in sustaining school improvement and growth in student learning over time. The findings of the research led to two observations.

The first observation is about principal leadership. Evidence from our research suggests that building internal school capacity for improvement is not a simple, linear process. It requires directions from inspiring and visionary school leadership to create, develop, and sustain coherent and fit-for-purpose structures, cultures, and conditions to grow the knowledge, skills, and commitment of individuals and harness them to become the collective capacity of the school. This observation confirms what we already know from the research literature on successful school leadership: shared directions and goals and consistency in understandings of the standards of teaching and learning are key characteristics of high-performing schools where fundamental principles and values unite and drive teachers and leaders to be intellectually and emotionally committed to make a positive difference to the lives of their students (Day et al., 2011; Drysdale and Gurr, 2011; Gu et al., 2014; Klar and Brewer, 2013; Leithwood et al., 2006, 2010; Sun and Leithwood, 2015).

Investigating how successful principals had managed external policy demands from the perspectives of their key staff enabled us to learn more about leadership and change. In high-performing schools a key defining aspect of principal leadership is an unrelenting focus on fostering consistent values, expectations, and standards, and through these, empowering and transforming staff capacities and organizational conditions to embrace change and improvement. A growing body of research on senior and middle-level leaders, albeit limited, has shown the importance of their roles in influencing decision making, supporting the professional development of their peers, promoting shared vision and high expectations, developing a culture of shared responsibility and trust, and fostering collaboration at different levels within and across the school (Cranston, 2009; Dinham, 2007; Gurr and Drysdale, 2013; Leithwood and Riehl, 2005; Shaked and Schechter, 2018; Wenner and Campbell, 2017). Put succinctly, senior and middle-level leaders hold a key position in schools’ efforts to manage change, embed innovation, and sustain improvement. However, the extent to which they are able to fulfill their leadership roles and responsibilities effectively and consistently is influenced, directly and profoundly, by the leadership values, qualities, and practices of their principals (Day et al., 2011; Gurr and Drysdale, 2013).

In this research, the marked similarities between the principal and key-staff SEM models (Gu et al., 2018a), together with the narrative accounts from the case-study examples, highlight—most powerfully—that establishing consistency and coherence in

school structures, cultures, and improvement processes holds the key to engaging the heads, hearts and hands of the school community to achieve sustained and sustainable performance over time. Such consistency and coherence were reflected in understandings—between those who led and those who were led—of how and why their schools were able to become successful and, as importantly, to stay successful. They helped to shape the day-to-day practice of school life and enabled schools in our research to remain focused on the core business of schools – quality provision of teaching and learning – irrespective of external pressures of change.

The second observation is about how successful principals respond to, manage, and enact external policy demands. We found that in successful schools the process of policy enactment and the process of school improvement are not two separate processes. They are intertwined to form one overall process in which external policy initiatives and internal school improvement practices are purposefully aligned by principals to serve their moral purposes, educational values, and goals for the school. At the heart of this intertwined process are continuous leadership efforts to support collaborative professional learning and development and, through this, to build the necessary whole-school capacity for sustainable personal, social, and academic improvement in student outcomes. Investigating how school principals shape their school's improvement efforts from the perspectives of middle and senior leaders provides us with more insightful evidence on how and why some schools have the capacity to manage change, despite facing the same external challenges and demands, whilst others do not.

In summary, policy enactment is, in essence, about change. External policy initiatives—whether they are foreground or background noises that schools can or cannot ignore—represent some, but not all, of the many demands, challenges, and opportunities that schools face in their everyday working worlds. Enacting policy successfully essentially relies on building and consolidating the capacity for further growth and development. Key in this regard are strong leaders who know how to design the social and intellectual conditions which engage the heart and mind of individuals in the school and, through this, harness their ideas, experiences, knowledge, and relationships to fulfill shared values and achieve shared goals. Kotter (1996) argues that although managing change is challenging, the much better challenge for most organizations is “leading change” (1996, p. 30). Evidence from our research shows that in successful schools, what the leaders appear to be doing exceptionally well is using policies and reforms as opportunities for change—purposefully, progressively, and strategically—to regenerate collaborative and coherent cultures and conditions which encourage and support the staff to learn, to reflect, and to renew their practice.

References

- Ainscow, M., 2015. *Towards a Self-Improving System: Lessons From a City Challenge*. Routledge, London.
- Ball, S.J., 1994. *Education Reform: A Critical and Post-Structural Approach*. Open University Press, Buckingham.
- Ball, S., Maguire, M., Braun, A., Hoskins, K., 2011. Policy subjects and policy actors in schools: some necessary but insufficient analyses. *Discourse* 32 (4), 611–624.
- Ball, S.J., Maguire, M., Braun, A., 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. Routledge, London.
- Braun, A., Maguire, M., Ball, S., 2010. Policy enactments in the UK secondary school: examining policy, practice and school positioning. *J. Educ. Pol.* 25 (4), 547–560.
- Braun, A., Ball, S., Maguire, M., Hoskins, K., 2011a. Taking context seriously: towards explaining policy enactments in the secondary school. *Discourse* 32 (4), 585–596.
- Braun, A., Ball, S., Maguire, M., Hoskins, K., 2011b. Taking context seriously: towards explaining policy enactments in the secondary school. *J. Educ. Pol.* 32 (4), 585–596.
- Cohen, D.K., Moffitt, S.L., Goldin, S., 2007. Policy and practice: the dilemma. *Am. J. Educ.* 113 (4), 515–548.
- Cranston, N., 2009. Middle-level school leaders: understanding their roles and aspirations. In: Cranston, N., Ehrich, L. (Eds.), *Australian Educational Leadership Today*. Australian Academic Press, Brisbane, pp. 217–241.
- Day, C., Gu, Q., 2018. How successful secondary school principals in England respond to policy reforms: the influence of biography. *Leader. Pol. Sch.* 17 (3), 332–344.
- Day, C., Leithwood, K. (Eds.), 2007. *Successful Principal Leadership in Times of Change: International Perspectives*. Springer, Dordrecht.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *Successful School Leadership: Linking with Learning and Achievement*. McGraw Hill Open University Press, Maidenhead.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Deal, T., Peterson, K., 2009. *Shaping School Culture: Pitfalls, Paradoxes, & Promises*, second ed. Jossey-Bass, San Francisco, CA.
- Department for Education, 2013. *New Advice to Help Schools Set Performance-Related Pay*. Press Release. Retrieved from: <https://www.gov.uk/government/news/new-advice-to-help-schools-set-performance-related-pay>.
- Dinham, S., 2007. The secondary head of department and the achievement of exceptional student outcomes. *J. Educ. Adm.* 45 (1), 62–79.
- Drysdale, L., Gurr, D., 2011. Theory and practice of successful school leadership in Australia. *Sch. Leader. Manag.* 31 (4), 355–368.
- Elmore, R.F., 1995. Structural reform and educational practice. *Educ. Res.* 24 (9), 23–26.
- Elmore, R.F., 2003. Accountability and capacity. In: Carnoy, M., Elmore, R., Siskin, L. (Eds.), *The New Accountability*. Routledge; Falmer, New York, NY and London, pp. 195–209.
- Elmore, R.F., 2004. *School Reform from Inside Out: Policy, Practice, and Performance*. Harvard Education Press, Cambridge, MA.
- Elmore, R.F., 2011. Policy is the problem, and other hard-won insights. In: Elmore, R.F. (Ed.), *I Used to Think... and Now I Think*. Harvard Education Press, Cambridge, MA, pp. 33–40.
- Fuhrman, S., 1999. *The new accountability*. RB-27-January. Consortium for Policy Research in Education, Philadelphia, PA.
- Fullan, M., 2016. *The New Meaning of Educational Change*, fifth ed. Teachers College Press, New York, NY.
- Gu, Q., Sammons, P., Mehta, P., 2008. Leadership characteristics and practices in schools with different effectiveness and improvement profiles. *Sch. Leader. Manag.* 28 (1), 43–63.
- Gu, Q., Rea, S., Hill, R., Smethem, L., Dunford, J., 2014. *Teaching Schools Evaluation: Emerging Issues from the Early Development of Case Study Teaching School Alliances*. Department for Education, London.
- Gu, Q., Rea, S., Smethem, L., Dunford, J., Varley, M., Sammons, P., Power, L., 2016. *Teaching Schools Evaluation: Final Report*. Department for Education, London.
- Gu, Q., Day, C., Walker, A., Leithwood, K., 2018a. Editorial: how successful secondary school principals enact policy. *Leader. Pol. Sch.* 17 (3), 327–331.
- Gu, Q., Sammons, P., Chen, J., 2018b. How principals of successful schools enact education policy: perceptions and accounts from senior and middle leaders. *Leader. Pol. Sch.* 17 (3), 373–390.

- Gu, Q., Rea, S., Seymour, K., Smethem, L., Bryant, B., Armstrong, P., Ahn, M., Hodgen, J., Knight, R., 2020. The research schools network: supporting schools to develop evidence-informed practice. Evaluation Report. Education Endowment Foundation, London.
- Gu, Q., Seymour, K., Rea, S., Knight, R., Ahn, M., Sammons, P., Kameshwara, K., Hodgen, J., 2021. The Research Schools Programme in Opportunity Areas: Investigating the Impact of Research Schools in Promoting Better Outcomes in Schools. Evaluation Report. Education Endowment Foundation, London.
- Gurr, D., Drysdale, L., 2013. Middle-level secondary school leaders: potential, constraints and implications for leadership preparation and development. *J. Educ. Adm.* 51 (1), 55–71.
- Hatch, T., 2002. When Improvement Programs Collide. *Phi Delta Kappan*, pp. 626–639.
- Honig, M., Hatch, T., 2004. Crafting coherence: how schools strategically manage multiple, external demands. *Educ. Res.* 33 (8), 16–30.
- Hubbard, L., Mehan, H., Stein, M., 2006. *Reform as Learning*. Routledge, New York, NY.
- Johnson, S., 2019. *Where Teachers Survive: Organising Schools for Success*. Harvard Education Press, Cambridge, MA.
- Klar, H.W., Brewer, C., 2013. Successful leadership in high-needs schools: an examination of core leadership practices enacted in challenging contexts. *Educ. Adm. Q.* 49 (5), 768–808.
- Kotter, J., 1996. *Leading Change*. Harvard Business School Press, Boston, MA.
- Leithwood, K., 2018. Postscript: five insights about school leaders' policy enactment. *Leader. Pol. Sch.* 17 (3), 391–395.
- Leithwood, K., Riehl, C., 2005. What do we already know about school leadership? In: Firestone, W., Riehl, C. (Eds.), *A New Agenda: Directions for Research on Educational Leadership*. Teachers College Press, New York, NY, pp. 22–47.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. Seven Strong Claims about Successful School Leadership. National College for School Leadership, Nottingham.
- Leithwood, K., Harris, A., Strauss, T., 2010. *Leading School Turnaround: How Successful Leaders Transform Low-Performing Schools*. Jossey-Bass, San Francisco, CA.
- Maguire, M., Braun, A., Ball, S.J., 2015. Where you stand depends on where you sit: the social construction of policy enactments in the (English) secondary school. *Discourse* 36 (4), 485–499.
- Manning, P.K., 1982. Organizational work: structuration of environments. *Br. J. Sociol.* 33 (1), 118–134.
- Matthews, P., Rea, S., Hill, R., Gu, Q., 2014. *Freedom to Lead: A Study of Outstanding Primary School Leadership in England*. Department for Education & National College for Teaching and Leadership, London. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/363794/RR374A_-_Outstanding_primaries_final_report.pdf.
- Mitchell, D., Crowson, R., Shippis, D., 2011. *Shaping Education Policy: Power and Process*. Routledge, New York, NY and London.
- Newmann, F.M., 1996. *Authentic Achievement: Restructuring Schools for Intellectual Quality*. Jossey-Bass, San Francisco, CA.
- Payne, C.M., 2008. *So Much Reform, So Little Change: The Persistence of Failure in Urban Schools*. Harvard Education Press, Cambridge, MA.
- Shaked, H., Schechter, C., 2018. School middle leaders' sense making of a generally outlined education reform. *Leader. Pol. Sch.* <https://doi.org/10.1080/15700763.2018.1450513>.
- Spillane, J.P., 2004. *Standards Deviation: How Schools Misunderstand Education Policy*. Harvard University Press, Cambridge, MA.
- Spillane, J., Diamond, J., Burch, P., Hallett, T., Jita, L., Zoltner, J., 2002a. Managing in the middle: school leaders and the enactment of accountability policy. *Educ. Pol.* 16 (5), 731–762.
- Spillane, J., Reiser, B., Reimer, T., 2002b. Policy implementation and cognition: reframing and refocusing implementation research. *Rev. Educ. Res.* 72 (3), 387–431.
- Sun, J., Leithwood, K., 2015. Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *Sch. Effect. Sch. Improv.* 26 (4), 499–523.
- von Glasersfeld, E., 1989. Cognition, construction of knowledge, and teaching. *Synthese* 80 (1), 121–140.
- Walker, A., 2011. *School Leadership as Connective Activity*. Australian Council for Educational Leaders, Australia.
- Weick, K.E., 1979. *The Social Psychology of Organizing*. McGraw-Hill, New York, NY.
- Weick, K., 1995. *Sensemaking in Organizations*. Sage, Thousand Oaks, CA.
- Weick, K.E., 2005. Organizing and the process of sense-making. *Organ. Sci.* 16 (4), 409–421.
- Wenner, J.A., Campbell, T., 2017. The theoretical and empirical basis of teacher leadership: a review of the literature. *Rev. Educ. Res.* 87 (1), 134–171.

International patterns and trajectories in principal preparation

Darren A. Bryant^a, Allan Walker^b, and Moosung Lee^c, ^a Department of Educational Policy and Leadership, Asia Pacific Centre for Leadership and Change, The Education University of Hong Kong, Hong Kong SAR, PR China; ^b International Educational Leadership, Department of Educational Policy and Leadership, Asia Pacific Centre for Leadership and Change, The Education University of Hong Kong, Hong Kong SAR, PR China; and ^c University of Canberra, Canberra, ACT, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	356
Literature	357
Further development	358
Methodology	358
Site selection	358
Sources and data collection	359
Analytic strategies	359
Developing trajectories	359
Program profiles, contexts and frameworks	359
Canada (Ontario): principals qualification program (PQP)	359
Further development	360
Australia (Victoria)	360
Further development	360
Singapore: leaders in education program (LEP)	361
Further development	361
Hong Kong: certification for principalship (CFP)	362
Further development	362
The United States (New York city): aspiring principals program (APP)	362
Further development	363
Commonalities, variations and trajectories in leadership preparation programs	363
Frameworks	363
Common content areas	364
Sources of content	364
National and local contexts	364
Student and community need	364
Professional and academic content	364
Operational features	364
The practitioner	364
Tripartite collaboration	365
Provider expertise	365
Formal credentials	365
Selection	365
Trajectories	365
Embedding leadership into school improvement agendas	365
Contextually responsive changes	365
New content emphases	366
Brokerage and tendering models of delivery	366
Systems perspectives	366
Conclusion	366
Summary	366
Implications for policy, practice and research	367
References	368

Introduction

Leadership remains a critical determinant of performance within and across organizations, including schools (Leithwood et al., 2020). While often indirect, the impact of principals' leadership on various indicators of organizational capacity and performance is significant (Kulophas and Hallinger, 2020). In this context, attention internationally has focused on developing successful principals through high quality leadership development programs (Barnett and Okilwa, 2020). Research has documented core

attributes of high-quality programs that best prepare aspiring principals for the job ahead (e.g., Darling-Hammond et al., 2010; Slater et al., 2018). Our previous research found that high-performing leader development programs, despite operating in different countries, seem to share certain core designs, structures and practices (Walker et al., 2013). Yet they can be differentiated jurisdictionally based on governments' policy intentions (Bryant et al., 2012). However research has not investigated how specific programs adapt or remain stagnant over time in response to changing contexts. Addressing this can provide clues as to how leader development programs retain currency to meet the needs of aspiring educational leaders and school systems. Therefore, the purpose of this article is to assess how five high-quality leader development programs in different Asian and Western jurisdictions compare in three key design domains and how they change over time in response to shifting system environments. We are guided by four research questions:

- Are there any common or different *frameworks* which bind leadership development programs with related leadership functions or local conditions?
- What major *content areas* are delivered through the programs?
- How are programs *operationalized* in terms of provider expertise, formal credentials, and selection?
- How and why have preparation programs evolved in the above design domains over the last decade?

Through considering these questions we suggest implications for program development and identify major directions that principal preparation programs have and are taking to meet the complex needs of principals (Normore and Lahera, 2019).

In this article we examine developments in principle preparation internationally over the last decade by first recounting our previously published analyses on the state of principal preparation programs across five countries in 2011 (Walker et al., 2011; also see Walker et al., 2013; Bryant et al., 2012). We do this in each section of the article by extracting segments of verbatim text from our original manuscript (Walker et al., 2013), followed by updates with new analyses of current data (predominantly policy papers and program documents) for the programs. As such, we include discussion of how and why the structure and practices of leader development have changed in each jurisdiction, track their trajectories over the last decade and suggest issues for consideration going forward.

In order to expand understanding of leadership preparation across the five different national contexts we first present relevant literature reviewed in our original paper (Walker et al., 2013) with a brief update in order to provide an analytic framework to compare programs. We then recount methodological decisions including site selection, data collection, and analysis. Next, we delineate the major features of the selected programs and then analyze program commonalities and variations. We also discuss in a new section emerging trajectories and their possible implications for the programs.

Literature

Over the last 20 years, school principals' work has increased in complexity in response to wide-ranging reforms such as the coupling of school-based management with accountability measures (Cheng, 2009; Lee et al., 2012a; Walker and Ko, 2011), standardized testing (Carnoy and Loeb, 2002), and the use of data related to student learning outcomes to drive school improvement (Lee et al., 2012b). The pressures of implementing such changes in schools, the skills needed to make them work and the time commitments required has resulted in many potential leaders electing not to enter the principalship (MacBeath, 2011; Walker and Kwan, 2010). In Asian societies, such as Hong Kong and Singapore, this problem is not evident (Walker and Dimmock, 2006). Nevertheless, the bulk of recent research on leadership preparation has been conducted in western societies.

Within this context, preparing effective school leaders has been placed at the core of many educational reform agendas (Bryant et al., 2012) and has been the subject of much research. Indeed, recent studies have focused on indicators of effectiveness of high-quality leadership programs. For example, Darling-Hammond et al. (2010) studied eight pre-service principal training programs in the United States. These were selected because of the strength of their outcomes. Despite varied approaches, the following features were embedded in each:

- The alignment of a coherent curriculum with externally developed state and/or professional standards;
- An emphasis on instructional leadership and school improvement to guide the program philosophy and curriculum;
- The use of active and student-centered learning strategies;
- A teaching faculty that includes practitioners and scholars who are experts in their fields;
- Diverse support structures such as formal mentoring systems and cohorts;
- A rigorous approach to participant selection;
- Supervised internships or school attachments.

Consistent with these indicators, the overall trajectory of leadership preparation programs has converged across different societies. Based on our previous work (Walker et al., 2011), we particularly note this convergence in terms of three interdependent components of leadership programs—i.e., framework, content, and operation.

With the increasing global emphasis on accountability, states or professional associations frequently mandate leadership frameworks that are centrally defined and aligned to district or state level educational policies (Bryant et al., 2012). Within jurisdictions, various providers may be endorsed to deliver leadership programs that meet the requirements stipulated in the frameworks.

Although conducted mostly in western countries, previous research suggests a high level of commonality in the content of contemporary leadership development programs. For example, Bush and Jackson's (2002) found a shared emphasis on

organizational, transformational and instructional leadership as well as traditional content areas such as “professional development, finance, curriculum, and external relations” (p. 421).

A body of research also shows that common features embedded in leadership programs include the engagement of participants in the field through experiences such as internships or university-district partnerships (e.g., Darling-Hammond et al., 2010). Both tactics aim to address the provision of authentic leadership experiences for program participants. The internship serves to contextualize skills that are otherwise explored through simulations, case studies, and problem-based learning, learning strategies that fall short of “the same urgency, sense of responsibility, and discomfort” (Perez et al., 2011: 218–9) encountered in the principalship. In such programs, field experiences form the “centerpiece” of authentic development.

Another notable common feature is participation in action-oriented projects and experiential project-based learning that allow application of learning in real-world contexts. Findings by Piggot-Irvine (2011) suggest that for such approaches to be effective they require careful delineation of scope, time constraints and support from academic staff and/or mentors. This suggests that leadership projects would benefit from being coupled with extended periods in the field and well-developed mentorship structures. In sum, programs previously conducted in university classrooms only now include school-based leadership development that emphasizes learning through practice in context (Piggot-Irvine, 2011).

Darling-Hammond et al. (2010) suggest that excellent leadership programs demonstrate positive relationships between universities and districts. More specifically, Sanzo et al. (2011) report that when courses are taught in an integrated fashion involving university faculty and district leadership, participants seemed more able to connect theory to practice.

Many of the above developments reflect a swing toward learning through practice. This development has been reinforced in many countries by the involvement of professional associations, unions and nongovernmental agencies as program providers—alone or in partnership with universities—and/or as contributors to state leadership frameworks (Bryant et al., 2012; MacBeath, 2011).

All these features illustrate a degree of convergence in terms of the operational features in current leadership preparation programs; however, differing district and national contexts impact on their operation (Leithwood and Levin, 2008). Operationalizing leadership programs seems to be inextricably intertwined with local needs. Thus, whereas contemporary leadership programs employ similar features in terms of framework and content, state goals and local needs significantly impact on enactment (Bryant et al., 2012). Given that most previous research targeted western-based leadership programs, we believe that a joint focus on both western and eastern societies may better help identify variations. For instance, systems such as Hong Kong and Singapore, which perform at the high end of international school outcome measures and retain examination-oriented cultures, may honor differing values to inform leadership preparation. We believe that capturing both commonality and variations in quality leadership programs will provide a fuller picture for school leadership program developers and policy makers. This perspective sets the scene of our inquiry.

Further development

More recently, preparing effective school leaders has been placed at the core of many educational reform agendas in both developing and developed countries. Leader development research has focused on reviewing preparation programs in specific regions or countries (e.g., Devi and Fernandes, 2019; Gurm, 2020), evaluating and suggesting frameworks for designing programs (e.g., Normore and Lahera, 2019; Okoko, 2020); and comparing preparation programs within or across continents and systems (Barnett and Okilwa, 2020; Eacott and Asuga, 2014). However, there remains limited probing of synergies around government policy, local context, service brokers/providers, and program structures.

Methodology

In this section, we first relate verbatim the major thrust of our original 2011 methodology (Walker et al., 2013) and then explain our approach to updating the analysis.

Site selection

For the purposes of this study, we identified five principal preparation programs from five different societies for analysis: Australia, Canada, Hong Kong, Singapore, and the United States. Given our aim to draw international comparisons in leadership preparation, selection criteria addressed congruence around indicators of excellence and diversity of context.

At the state level, each of the selected programs is located in jurisdictions reported as top performers in the Program for International Student Assessment (PISA). We recognize that PISA results do not evidence a direct relationship to leadership preparation. However, PISA results are closely followed by policymakers, stimulate international competition, and legitimate reform efforts (Afonso and Costa, 2009; Rautalin and Alasuutari, 2009). High performing societies become magnets of interest for researchers and policymakers looking for strategies to improve their own educational systems (Phillips and Ochs, 2003). Accordingly, performance on PISA is recognized in consultancy reports (c.f. Barber et al., 2010), policy papers (c.f. National Governors Association and Council of Chief State School Officers and Achieve Inc., 2008) and academic research (c.f. Matthews et al., 2008) on school leadership, and informs our selection. Each of the jurisdictions selected for study performs in the top 10 of PISA measures. Further, the selected programs are offered in jurisdictions identified for excellence in educational leadership (Barber et al., 2010). Given the diverse performance on PISA that is found within large federations such as the United States and Canada, this helped to narrow our selection (e.g., to specify New York and Ontario).

At the program level, selection criteria used recently delineated research-based indicators of exemplary leadership programs with high performance outcomes (Darling-Hammond et al., 2010). These include alignment to “state and professional standards”, an emphasis on “instructional leadership and school improvement”, student-centered learning, knowledgeable faculty, cohort structures, “formalized mentoring and advisement”, rigorous participant selection, and site-based internships (Darling-Hammond et al., 2010: 181–2). Although the selected programs share such features, commonalities and variations in their development, employment and emphases need to be further understood. Selecting programs with broadly shared features allow for a more fine-grained analysis of similarities and distinctives.

Each program leads to a credential accredited by regulatory agencies or universities. We exempted Victoria, Australia from this final selection criterion in order to meet selection criteria for the diversity necessary to obtain an international perspective on principal preparation in high performing educational systems. We note that of the top ranked PISA performers, most do not have required pre-services preparation programs, which eliminated countries such as Finland, New Zealand and the Netherlands from our selection.

Finally, we limited our study to systems that operate leadership programs in the medium of English. This served practically to assist in data collection and analysis as the selected systems had documents in English which were publicly available on the Internet or by request.

Sources and data collection

The data set comprises reviews of program evaluation reports, in-house materials such as syllabi, and Internet searches. Each of the researchers initially explored different programs, contacted key personnel, searched the Internet, and uncovered academic papers and internal documents related to the specific programs. While Internet data often provided the general structure of programs, the evaluation reports, program guidelines and interviews served to tease out key or unique features of each (See Walker et al., 2013 for sources in addition to those provided in the reference list).

Analytic strategies

We devised a comparative analytic framework based on our review of extant literature. Leithwood and Levin’s (2008) leadership development typology informed the initial draft of the framework. Leithwood and Levin (2008) observe a lack of a comprehensive leadership typology. Accordingly, through a literature review, we identified other facets of leadership preparation programs and wrote descriptors of each in order to build a broader analytic typology. [The full typology and details of its development can be found in the original article (see Appendix 2 in Walker et al., 2013)]. While we used these to code programs, given that our purpose aimed for comparison we used a constant comparative approach (Corbin and Strauss, 2008) to identify content common across the five programs.

Nearer to the end of our analysis we arranged the codes under five broad themes: purpose, framework, content, delivery and operational features. However, in this paper we limit our discussion to the three themes that emerged as most significant and inter-related¹: framework, content and operational features. We report updates to these design domains as applicable.

Developing trajectories

In order to assess developments over time, we decided to analyze the current state of the original five leader development programs while recognizing possible debate about whether or not they may still be classified as exemplary. To do this, we collected current policy documents, leadership frameworks, and program designs found on government and program provider websites. We assessed change as represented in these documents in the framework, content and operational features domains.

Program profiles, contexts and frameworks

In this section, we outline key features of the selected programs, including their evolution over the decade since our initial analysis (Walker et al., 2013).

Canada (Ontario): principals qualification program (PQP)

In Canada, Ontario has long mandated a pre-service qualification for entry into the principalship, termed the Principals Qualification Program (PQP). The PQP is regulated by the Ontario College of Teachers (OCT), which mandates program expectations, content requirements and learning outcomes. The OCT also accredits providers, which include universities and professional associations (Ontario College of Teachers, 2011). The OCT’s guidelines in turn cohere to the Ontario Leadership Framework, developed by the Institute for Education Leadership (IEL, 2008), which cites Leithwood et al. (2006) influential research synthesis as providing

¹Discussion of other categories may be found in Bryant et al. (2012), and Walker et al. (2011).

Table 1 Changes to the Ontario leadership frameworks.

2008 OLF (IEL, 2008)	2012 OLF (Leithwood, 2012)
Setting directions	Setting directions
Developing people	Building relationships and developing people
Redesigning the organization	Developing the organization to support desired practices
Managing the instructional program	Improving the instructional program
Securing accountability	Securing accountability

its theoretical foundation. These informed four of the framework's leadership domains with a fifth emerging from the province's accountability requirements (IEL, 2008) (see column one of [Table 1](#)).

The framework also informs various other facets of leader development, including mentoring ([Ontario, 2010a](#)), and performance appraisal ([Ontario, 2010b](#)) and is delineated for principals, district supervisors, and government and Catholic schools.

We selected the PQP delivered by the Ontario Principals Council (OPC) for analysis. The OPC, as the professional association for principals, offers a program distinctive from those typically run through universities. The OPC delivers the PQP throughout the province in urban and rural districts, making it one of the most widely distributed leadership programs in Ontario. OPC is the largest PQD provider in the province.

Further development

Recent changes to the OCT's PQP guidelines followed a 2012 revision of the Ontario Leadership Framework that incorporated research on high-performing systems. Accordingly, the OLF revision includes "a set of effective individual and small group leadership practices to acknowledge the unique contexts, challenges and opportunities ... for both school and system level leaders" ([Leithwood, 2012](#), p. 4).

Notably, the framework emphasizes the importance of situated, shared and relationally oriented leadership practices, as opposed to articulating a set of universal competencies that are applicable to principals in all contexts. Building on the revised OLF, the [OCT \(2017\)](#) articulated the need to address school and system leadership within Ontario's diverse cultural and educational context and updated the PQP framework in response. [Table 1](#) shows changes in the OLF, which informed the current PQP framework ([OCT 2017, 2020](#)).

In response to the framework, the OPC has articulated new modules that relate to new emphases on accountability and system leadership and OLF's emphasis on student learning and wellbeing. Modules include "Equitable and Inclusive Schools", "Inclusive Education", "Co-creating Safe Schools" and "Building Relationships with Parents and the Community" ([OPC, 2020](#)). In 2020, the OCT accredited nine program providers, which include six universities and three professional associations ([OCT, 2020](#)).

Australia (Victoria)

The Australian state of Victoria placed leadership development as key to the state's school improvement efforts ([Elmore, 2007](#)). This led to the creation of the government-operated Bastow Institute of Educational Leadership, which supported leadership development for current and aspiring school leaders. It was responsible for implementing Victoria's Developmental Learning Framework for School Leaders (DLFSL) ([Department of Education, 2007](#)). The framework was derived from [Sergiovanni's \(2009\)](#) five forces of leadership—Technical, Human, Educational, Symbolic and Cultural leadership. Competency indicators for each domain were provided to inform principal induction, development, selection, mentoring, succession planning, and professional learning ([Department of Education, 2007](#)).

No single credential was required of principals; however, Bastow developed a range of pathways for leader development. These ranged from weekend workshops to a 2-year Master of School Leadership (MSL) delivered by two providers: the University of Melbourne and Monash University. Bastow served to regulate and accredit the MSL.

Further development

Our original analysis focused on the MSL delivered by the University of Melbourne. However, structural changes quickly followed our 2012 analysis. Shifts in leader development emerged from 2015 when a new Labor government launched its "Education State" policy that aimed to make Victoria's school system the best performing in Australia. The policy aimed to bolster students' achievement, engagement and well-being. A newly established Department of Education and Training (DET) published an "evidence-based framework" for school improvement, the *Framework for Improving Student Outcomes* (FISO), which became compulsory for all Victorian government schools ([DET, 2020](#)). Professional Leadership, one of FISO's four priorities, identifies four dimensions for school leader development (see [Table 2](#)). Each is justified with reference to syntheses of research, (primarily [Day and Sammons \(2013\)](#), [Hattie \(2015\)](#), [Marzano et al. \(2005\)](#), [Robinson et al. \(2008\)](#), and [Sun and Leithwood \(2015\)](#)). Each priority is delineated on a four-point leader developmental continuum from emerging to excelling. As the governments' leadership development agency, Bastow

Table 2 Rough alignment of Victoria's DLFSL, FISO and APSP frameworks in Victoria.

<i>Developmental learning framework (DoE, 2007)</i>	<i>FISO (DET, 2020)</i>	<i>APSP (AITSL, 2015)</i>
Human	Building leadership teams	Developing self & others
Educational	Instructional and shared leadership	Leading teaching and learning
Symbolic	Vision values and culture	Leading improvement, innovation and change
Culture	Vision values and culture	Engaging and working with the community
Technical	Strategic resource management	Leading the management of the school

Institute of Educational Leadership (2020) adopted FISO as its central framework for developing teacher, principal and system leaders. Additionally, DET and Bastow adopted the Australian Professional Standard for Principals (APSP, Australian Institute for Teaching and School Leadership (AITSL), 2015) as the leadership capabilities and professional practices needed to enact FISO. Table 2 shows a rough alignment of the frameworks.

Bastow reoriented its operations toward a Brokerage model, linking a range of professional learning opportunities and providers with schools and educators. It arranged for credit transfer from some modules into postgraduate degrees at three universities. However, it positioned "Unlocking Potential" as its flagship certificate program to "increase the pool of high-quality aspirant principals". The program, launched in September 2016, uses FISO and APSP for its organizational frameworks.

Singapore: leaders in education program (LEP)

Since 1997, Singapore's educational reforms have been aligned with its vision of "Thinking Schools, Learning Nation" (TSLN) and further specified in 2005 through the state's policy initiative "Teach Less Learning More" (TLLM)" (Ng, 2008; Ministry of Education (MOE), 2005). Under TLLM, the role of school leaders stressed the development of creative and innovative transformational leaders, and instructional leaders capable of creating and implementing innovative programs for diverse learners (MOE, 2005; National Institute of Education (NIE), 2010). To achieve this goal, Singapore mandates a single pathway to the principalship through the Leaders in Education Program (LEP), a 7-month intensive full-time program, which in 2001 replaced the Diploma in Educational Administration. It is delivered only by the NIE.

In keeping with government policy, the LEP draws its framework from international models adapted to focus on promoting creativity, innovation, diversity, and distributed leadership. Singapore, like Victoria, draws on Sergiovanni's *Five Forces of Leadership* but also applies Howard Gardner's (2008) *Five minds for the future* to identify key competences and attitudes, termed mind-sets, that are "needed to perform the roles" (NIE, 2011:5–6) (see Table 3).

The framework, developed by faculty at the NIE and endorsed by the Ministry of Education is applied narrowly to the LEP, for which it was specifically created.

Further development

More recently, Singapore's MoE has adopted 21st Century Competencies (known as 21CC) as an outgrowth of TSLN and the current driver of the school system. 21CC aligns to Singapore's current economic policy aims around innovation and entrepreneurialism and, mediated by Desired Outcomes for Students, frames curriculum review and teacher professional development. Recent revisions to the desired outcomes emphasize values, emotions and interpersonal skills (MoE, 2020).

The NIE continues to offer the NEP for school and teacher leaders. The most recent program structure remains rooted in the blended Sergiovanni—Gardner framework but has added the dimension of "value-creating", a term used in Singapore's current economic policy agenda. This dimension serves to drive the Creative Action Project component of the LEP, in which participants anticipate a future-oriented need for their placement schools and design and implement innovative solutions.

Table 3 Framework guiding Singapore's LEP.

<i>Sergiovanni's framework (2009)</i>	<i>Gardner's framework (2008)</i>
Educational	The disciplined mind develops expertise in "at least one discipline" that entails a specific set of skills, knowledge and understanding
Technical	The synthesizing mind collects, analyses, evaluates, and synthesizes information from wide-ranging sources
Human	The creating mind moves knowledge ahead of technological advances through innovative questioning and thinking
Culture	The respectful mind, cognizant of the globalized world's interconnectivity, negotiates differences among individuals and societies
Symbolic	The ethical mind acts for the greater social good over self-interest

Hong Kong: certification for principalship (CFP)

Until 1999, leadership preparation in Hong Kong followed no clear, coherent pathway. A shift toward involving academics and practitioners in developing and delivering a mandatory leadership preparation credential occurred in response to two consultation reports in 1999 and 2002, resulting in establishing the Certification for Principalship (CFP) as a mandatory entry requirement for principals. The government licenses three providers, all tertiary institutes, to deliver the CFP. Providers must demonstrate an alignment of the program they designed to a required framework, *Key Qualities of the Principalship in Hong Kong* (Education Bureau, 2002; Walker et al., 2002).

The Hong Kong framework is located in a wide range of international research, which has been adapted in response to research conducted on leadership in the Hong Kong context (Walker et al., 2002). It comprises four leadership domains:

- Strategic leadership,
- Instructional leadership,
- Organizational leadership, and
- Community leadership.

These in turn inform six core areas of program foci:

- Strategic direction and policy environment,
- Learning, teaching and curriculum,
- Teacher professional growth and development,
- Staff and resources management,
- Quality assurance and accountability,
- External communication, and connection to the outside world.

The six core areas are further delineated by a structure of values, knowledge, skills and attributes that guide leadership development programs for aspiring principals, which the CFP targets, newly appointed principals, and serving principals (Education Bureau, 2002). CFP candidates complete a needs analysis, preparation for the principalship course and a portfolio.

Further development

No major changes have been made to the structure and content of the CFP over the past twenty years. However, two major policies are accounted for in program delivery. First is the introduction of the *Professional Standards for Principals of Hong Kong* in 2018 that identifies three standards that focus on student learning and personal growth:

- *Ethical enablers* of all-round growth and balanced advancement,
- *Versatile architects* of vibrant learning organizations, and
- *Visionary entrepreneurs* of educational transformation and continuous improvement.

These are articulated on a three-point continuum and are suggested as broad-based aspirational descriptors, rather than as a series of competencies. They serve as a reference point for professional development and learning, rather than as a mandate for program delivery or performance appraisal.

Secondly, following a review of professional development, the Education Bureau (2020) established a “professional ladder” from pre-service teachers through to the principalship that identifies the new teacher and principal standards as the goals for professional development, provides additional middle leader and vice-principalship positions in schools, and requires formal PD in leadership and professionalism as a mechanism for developing the leadership pool.

While the EdB has not changed any of the CFP leadership domains, program foci or delivery, it permits individual providers to include a limited number of additional topics in their programs to address the perceived needs of participants.

The United States (New York city): aspiring principals program (APP)

New York City comprises the largest public school district in the United States. Its reform efforts explicitly highlight leadership, empowerment, and accountability as its three key pillars. These aim to close the achievement gap by stressing high student standards and the use of data to drive improvement and shape accountability processes (Fryer, 2011). Within this context, the New York City Leadership Academy’s (NYCLA) Aspiring Principals Program (APP) has demonstrated its success in raising student outcomes (Corcoran et al., 2009).

The APP program framework was developed by the provider, the NYC Leadership Academy (NYCLA), and derived from multiple standards set at the national, state and city levels. Termed the Leadership Performance Standards Matrix, the framework emphasizes transformational and instructional leadership and drives content and assessment across twelve dimensions:

- Personal behavior;
- Resilience;
- Communication;

- Focus on student performance;
- Situational problem-solving;
- Learning;
- Accountability for professional practice;
- Supervision;
- Leadership development;
- Climate and culture;
- Time, task and project management, and;
- Technology. (New York City Leadership Academy (NYCLA), 2012).

This framework guides “participant selection, curricular scope, assignments, program advancement, interventions and the comprehensive evaluation of each participant” (NYCLA, 2012). Graduates must demonstrate competency in each dimension, further delineated by 55 behavioral criteria, which are in turn defined along a continuum of not meeting, progressing toward or meeting the standard. The Matrix has in turn been applied to the development of a Leadership Performance Planning Worksheet in conjunction with the states of Delaware and Kentucky and the Wallace Foundation. The Worksheet is used in the development of novice principals and is aligned to the Educational Leadership Policy Standards (ELPS, Council of Chief State School Officers, 2008), which are derived from a database of 83 international empirical studies and 47 other references.

Further development

The years 2017–2019 marked a pivotal transition for school leadership development in NYC. This period included New York City Department of Education’s (NYCDE) introduction of a new “Instructional Leadership Framework” (ILF) for school improvement that aims to build “system-wide” coherence and improve “achievement for every student” (NYCDE, 2019a, p. 1). The state formally adopted the National Policy Board for Educational Administration’s (NPBEA) 2015 Professional Standards for Educational Leadership (PSEL) to replace 2008s ELPS with effect from 2020. PSEL included new emphases on equity and cultural responsiveness, alignment from school to district while focused on school leadership, recognition that leadership is broadly based in schools, and aspirational in orientation by calling on principals to “move beyond established practices and systems” (NPBEA, 2015, p. 4) through situational awareness rather than prescribed actions.

NYCLA’s (2019) review of the APP emphasized its impact on the career pipeline of principal preparation and on the structure of principal development in other jurisdictions; but stated the program was discontinued in 2017. NYCDE (2019b) issued the Chancellor’s Regulation C-30 to regulate the requirement and procedure of principalship applications. C-30 required aspiring principals to hold School Building Leader (SBL) certification. SBL qualifications can be obtained by completing the NYCDE’s *Leaders in Education Apprenticeship Program* (LEAP)—an outgrowth of the APP, the Wallace Leadership Fellows Program offered with various university partners, or through independent coursework, such as a master’s degree or an Advanced certificate offered through different universities. The former two programs retain internships, cohort-based learning, and financial support. Holders of SBL must meet an array of criteria related to instructional leadership, school culture, school improvement, and pedagogic experience to enter the Principal Candidate Pool. Significantly, the NYCLA remains an influential partner through collaborating with universities to launch LEAP.

Commonalities, variations and trajectories in leadership preparation programs

We now look across the five programs to identify their commonalities, variations and trajectories. First, we outline our original analysis organized under the three design domains. Second, we analyze subsequent changes and their implications using the subheading trajectories. Finally, we update some conclusions around commonalities, variations and trajectories.

Frameworks

Network formation is heavily influenced by governments, central/district offices, and/or international pools of research, and mediated by regulating agencies and program providers. The frameworks are applied or adapted to guide a range of leadership development strategies and programs. Internationally respected theories and/or empirical research informed frameworks with adaptations accounting for the local context and new research findings. For example, Ontario added a fifth dimension to Leithwood et al.’s framework, Hong Kong adjusted international findings to account for local leadership research, and Singapore focused its the framework on national policy concerns of nation building and international competitiveness.

Variations across the sources of the frameworks and in their application were apparent. The Ontario and Hong Kong programs used research syntheses to design program standards and structures. Ontario applied a synthesis of Leithwood et al.’s (2006) work which was originally published by England’s National College of School Leaders. Hence, it drew partly on an established framework that had informed practices elsewhere. Hong Kong’s framework was derived from an international synthesis of literature filtered through research on leadership in the local context. The Singaporean, Australian and NYC programs provided interesting variations, all were inspired by Sergiovanni’s (2009) model. However, Singapore added Gardner’s (2008) framework as a way to focus leadership practices.

Common content areas

We noted that each program emphasizes facets of instructional and transformational leadership and addresses specific topics as determined by their respective frameworks. As presented by Walker et al. (2013), there was a great deal of commonality across a plurality of topics. The most salient common feature across the programs was that “content is to some extent determined in the form of centrally defined competency frameworks” (Ontario, Victoria, and Hong Kong) or “frameworks developed by the provider but which draw on a range of sources for their construction” (Singapore and New York). Where content expectations were delineated by regulators, providers shaped program content by allocating instructional resources, determining assessment strategies, and stressing areas of respective expertise. Coupled with this overall commonality across programs, we detail below that there are context-specific variations across programs. Variations in context and content sources shape program foci, at times narrowing the focus to the needs of communities and students.

Sources of content

In each of the programs, content was derived from leadership frameworks, however other factors shaped specific programs.

National and local contexts

By analyzing the programs comparatively, contextual features emerge as drivers of content selection. In some instances, statutory or national aims determined the key areas. The OPC’s PQP program drew particular attention to issues pertaining to inclusion and diversity, which reflected Canada’s constitutional status as a multicultural and bilingual country, and legislative requirements to address issues of access for aboriginal peoples as well as those with learning and developmental challenges (Ontario Principals’ Council, 2011a,b). The content of Singapore’s LEP aligned to aims of nation-building in an era marked by international competitiveness, globalization and societal change. This included an emphasis on *knowledge creation and innovation, the development of networks, understanding international and corporate leadership contexts, and social constructivist processes* (MOE, 2005; NIE, 2010, 2011). To provide instruction in these areas, the LEP comprised foundational coursework, a school attachment, a school based “Creative Action Project”, participation in a learning community called the “Learning Syndicate”, and “management dialog sessions” and industrial attachments (NIE, 2011). In Victoria, the state’s emphasis on leadership development informs MSL content through the government’s Developmental Learning Framework and the Bastow’s 13 core leadership modules. In contrast state defined content, New York City’s APP targets local school district concerns, which includes a large number of inner-city schools. Local needs drive its focus, emphasizing *instructional leadership* as the way to narrow the achievement gap. This specifically involves leadership preparation for instructional improvement in low-performing schools located in high poverty communities.

Student and community need

Hong Kong’s CFP addressed common content areas, but also incorporated an emphasis on a personal needs analysis that stressed the centrality of a strong and meaningful joint academic-practitioner partnership. The needs analysis formed a profile for individual participants based on multi-point feedback. Participants responded to the needs analysis by creating a learning portfolio to demonstrate their in-school learning. Instructors, mentors and syndicate group leaders, who were senior school principals, used individual needs-analysis results to provide feedback and guide participants’ development. The needs analyses were in turn used to create a cohort profile that allowed instructors to tailor the content of subsequent modules. In the OPC’s PQP, the involvement of practitioners, serving principals and district specialists, as primary instructional leaders, guest lecturers and mentors was intended to permit a focus on the particular needs of the wide-ranging communities which the program served.

Professional and academic content

Providers’ expertise, and that of their faculty members, focus program content. The OPC’s status as a professional organization gives its PQP curriculum writers access to its counseling and legal arms. This informed content pertaining to law and ethics, giving it genuine currency in these areas. In New York, the APP’s contextual aims and practitioner orientation, with instructors and mentors comprised of current and retired school leaders and district personnel, shaped the interpretation or emphasis of content. Melbourne’s MSL emphasizes leaders’ self-knowledge through “positive psychology” to ground effective leaders (Waters and Luck, 2011). The MSL program provides students with access to international researchers via the University of Melbourne faculty and visiting academics. The MSL differs from the other programs we examined given its exceptionally strong emphasis on developing leaders as researchers who contribute to theory through action research. Thus, while frameworks, often informed by state legislation, determine the primary content direction, content specification is mediated by identified student community needs and the particular expertise and organizational purposes of the providers.

Operational features

The practitioner

Practitioners taking on expanded roles in leadership preparation was apparent across all programs. Practitioners played crucial roles in providing guidance to aspiring leaders through mentorship or job shadowing and tended to be involved in all aspects of the program including on-site or in-class mentoring, formative and summative feedback. This was particularly so in the OPC’s PQP

and NYCLA's APP where practitioners acted as formal instructors and assessors, replacing positions conventionally held by university staff. Practitioners were seen to lend credibility by providing relevant tacit knowledge and the tailoring of programs to the needs and resources of their local communities. The involvement of professional mentors in Victoria or district leaders in Ontario helped contextualize the content and skills articulated in the framework. Whereas Hong Kong's CFP, the University of Melbourne's MSL, and the NIE's LEP emphasized content derived from research and academic networks to support programs.

Tripartite collaboration

Selected leadership programs leveraged tripartite collaboration among government/regulator, provider and schools. This collaboration informed the development of leadership frameworks, program regulation and content. For instance, Ontario included providers in planning the redevelopment of the PQP requirements. Tripartite collaboration permitted multiple providers to deliver the programs in a manner that ensured some continuity in philosophy and content, simultaneously catering for local needs by using local expertise. In most programs, this occurred at the school level where mentors guided aspiring leaders. The MSL exemplifies tripartite collaboration in that participants worked with three mentors reflecting professional, academic, and in-school program aims. This mirrored the tripartite aims of encouraging participants' progression along the developmental continuum, honing their skills as researchers, and meeting school needs.

Provider expertise

Across the systems, except for Singapore, candidates could choose from different providers. Providing some measure of choice to candidates appeared a significant operational feature of leadership training programs. However, once admitted to an individual program, candidates progress as cohorts through a prescribed curriculum. The providers in Ontario, New York, Victoria and Hong Kong are each among one of several in their respective jurisdictions. Programs were shaped by the providers' particular expertise. Programs offered by professional associations and not-for profit groups draw on networks and resources that differ from university-based programs. They tended to emphasize practitioner orientations, local networks, and content derived from professional sources. For example, the OPC drew on its professional networks to identify top school leaders to act as instructors.

Formal credentials

The programs delivered a range of credential options. All but the MSL led to the formal, required licensure needed to become a principal. In addition to formal licensure, some programs offered pathways to academic credentials, although in different ways. The MSL alone terminated in a postgraduate degree. However, in Ontario, the OPC was affiliated with universities in Canada, Australia and the UK which recognized its PQP as partial credit toward their master's degree programs. In Hong Kong, all providers are universities and the CFP was recognized as postgraduate credit at the discretion of the various universities. The Singaporean and New York programs led to the earning of a license but not a postgraduate degree.

Selection

All programs had rigorous selection criteria, which typically required the prior completion of other academic or professional development programs as well as the endorsement of a participants' supervisor. In Singapore, applicants compete for limited places. Singapore limited places to vice-principals and Ministry of Education officers, which also selected the participants. This limited selection may reflect the locally-gearred goals of the program. In contrast, the other jurisdictions admitted a broad range of qualified applicants.

Trajectories

Our assessment of trends since 2010 suggests that many of the above conclusions still hold. Leadership frameworks still drive most programs. Of the programs we analyzed, all have adapted the pre-existing framework (Ontario, NYC and Singapore) and/or introduced new complementary frameworks (Victoria, NYC and Hong Kong). However, we identified five adaptations over the decade.

Embedding leadership into school improvement agendas

In three instances frameworks complementary to leadership standards have been developed within the last five years that embed instructional leadership within wider school improvement goals (New York's Instructional Leadership, Victoria's FISO, and Hong Kong's Professional Standards). Across all systems except for Singapore, new or revised frameworks explicitly reference aspirational improvement-orientations and situational-contextual considerations, as opposed to essentialist core competencies applicable in all contexts.

Contextually responsive changes

Changes to frameworks and programs continue to reflect shifts in governments' economic and political agendas. This is seen most explicitly in Singapore's introduction of the new leadership domain and Victoria's new FISO framework in response to new State policy. Where available, jurisdictions have adopted or slightly adapted national standards (NYC, Victoria) and/or revised existing standards in light of new research and demands of practice (Ontario and NYC).

New content emphases

Across all societies, there were changes in the content included in frameworks and programs. These changes tracked directly to local contextual changes. The most notable inclusions are around leadership to address cultural diversity and equity, and to support student wellbeing. Hong Kong and Singapore both explicitly address the entrepreneurial side of educational leadership which calls on leaders to be adaptive, responsive and innovative. These are complementary to the aspirational nature of the frameworks as mentioned above.

Brokerage and tendering models of delivery

Similar to our previous findings, all jurisdictions except for Singapore allow for the potential of different providers. This may be through a tendering process (Hong Kong), or through alternative pathways such as state agencies, or partnerships with several universities, PD providers and NGOs (NYC, Victoria, Ontario) or through a brokerage role by which government agencies direct aspirants to different providers. Consistent across this is the notion of providing some choice in terms of the nature of the program. Choices ranged from focused modules to intensive certificate programs, to postgraduate degrees or mechanisms for postgraduate credit recognition.

Systems perspectives

In most jurisdictions, frameworks and programs are now positioned as part of a professional ladder or pipeline for leadership preparation that develops teacher leaders into senior school and district leaders. A systems perspective is also manifest in programs that see schools and the principalship as part of the broader school system. The systems view emphasizes the principal's role in forging stronger community networks and with other school and district leaders.

Conclusion

Summary

Similar trajectories of leadership preparation have emerged across the five high-performing education systems in terms of three interdependent program components—i.e., framework, content, and operation. Apart from these commonalities, we also note that differing local and national contexts have an impact on creating variations in terms of the program enactment. These are summarized as follows.

First, major commonalities in terms of framework show that these:

- are based in empirically grounded research or theory;
- are derived from research syntheses or widely respected international sources;
- provide the potential of a common leadership language across the respective jurisdiction;
- are adjusted to suit jurisdictional needs.

At the same time, however, major variations range from:

- Derived from a synthesis of research to derived from a singular theoretical construct;
- Mandated from the top (Ontario, Hong Kong and Victoria) or framed by program developers (Singapore and New York).

Observed trajectories include:

- A shift from universal competencies toward practices that are situationally conditional;
- An increasing emphasis on embedding school improvement and systems-orientations into leadership frameworks, or emphasizing these through complementary frameworks.

Second, with respect to content, major commonalities include:

- Established frameworks drive overarching content;
- Provider expertise determine specific foci;
- Practitioner involvement supports contextualization of content;
- Content tends to emphasize a change orientation and stresses instructional and transformational leadership.

Content variations include:

- Specific program foci are dictated by local concerns as mediated by the particular expertise of academics and/or practitioners;
- Content ranges from addressing “national” priorities such as globalization and the knowledge society (Hong Kong and Singapore) to more pragmatic concerns of the “local” context (e.g., NYC on closing the achievement gap).

Observed trajectories include:

- An enhanced emphasis on leadership for student wellbeing, diversity and equity;
- An emphasis on the innovative, entrepreneurial, and aspirational nature of school leadership.

Finally, major commonalities in terms of operational features include that programmes:

- Accept aspiring and potential principals;
- Have admission requirements (e.g., experience and professional qualification);
- Provide a formal licensure/certification—this is required by the system although some provide credit toward formal degrees;
- Are offered by multiple providers who are centrally selected/regulated (except Singapore);
- Are expanding practitioners' roles in line with the call for more contextualized knowledge and skills.

The variations of operational features include:

- Some programs (e.g., Singapore and NYC) offer a limited number of fully funded places to attract talent;
- Other programs (e.g., Hong Kong and Ontario) focus on wider access to potential leaders and so an expanded pool of future principals. (In Ontario, completion rates now exceed principal positions);
- Programs which encourage wide access are offered on a part-time basis (Ontario, Melbourne-Victoria, Hong Kong); highly selective programs (NYC and Singapore) are offered only on a full-time basis.
- While programs depend on a tripartite partnership, variation in operational features across programs is shaped by the key providers' particular expertise and emphasis.

Observed trajectories include:

- A tendency toward brokerage and modular delivery of training that enhances alternative pathways to principal certification;
- Locating the principalship as just one part of a professional ladder or pipeline by which prior leadership training (supported through the modular approach) serves to develop a pool of potential principals.

Implications for policy, practice and research

These findings point to several features that researchers and providers of leader development programs may consider.

First, leadership frameworks inform each program. How broadly these frameworks are applied varies. The benefits of aligning all leadership development and evaluation in a jurisdiction to a common framework allows for a common leadership language to be shared among government, regulators, providers and participants across a system. This may further promote opportunities for continued professional learning in networks across schools (e.g., networked learning communities). Such benefits should be considered in light maintaining flexibility to address local needs. In brief, program developers may consider the impact of broadly applied leadership frameworks when they further develop or revise programs for potential/aspiring principals. And indeed, official frameworks now tend to recognize the situated nature of leadership.

Second, greater openness to qualified candidates may impact healthily on a system's leadership language and networks for those who wish to build their leadership capacity without aspiring to the principalship immediately. Where programs allow all qualified candidates to participate, they broaden the pool of leaders from diverse backgrounds. As Ontario, Victoria and Hong Kong apply their frameworks to a wide range of programs, admitting participants who may not become principals serves to increase the capacity of middle leaders. Having multiple providers also broadens opportunities for more leaders to participate. Possible merits of this include that programs can train numbers beyond the positions immediately available for the principalship, embed a common leadership language (Walker et al., 2011), and even encourage teachers to see themselves as leaders (Townsend, 2011). This may boost a shared understanding of leadership among teacher-leaders at various levels, not just principal aspirants. From a practical standpoint, accepting larger numbers of candidates who meet admission requirements (e.g., Hong Kong and Ontario) could ease principal shortages in high-needs schools. Several jurisdictions now envision a pipeline of leadership preparation extending from teachers to systems leaders with modular training available at key career stages.

Third, at the same time, programs that are highly selective/competitive and provide a limited number of places (Singapore and New York City) may promote (1) smoother, more predictable program operation or management, (2) programs that can be organized and implemented around team-based learning units, (3) participants who can access more customized and intensive learning opportunities, and (4) allows those selecting to be more discerning about the talent and type of leaders they want to shape their system. Whether systems open or restrict access to programs may reflect the extent to which a jurisdiction wants to promote a common understanding of leadership for all potential leaders, regardless of their actual position or aspiration. It may also relate to funding availability, supply and demand, or desire for central quality control of programs, and ultimately schools.

Fourth, our analysis draws predominantly on documentary sources that emphasize program design. We have observed that program developers look elsewhere to inform program frameworks. However, as is common with many reform efforts, international borrowing may emphasize the needs of the state, but it is "on the ground" where adaptations to local context are made. Although the diverse aims of states and the expertise of providers may explain some variation in programs, how societal-cultural differences shape variation requires further research and the attentiveness of program designers. The documentary data used in this study did little to explain cultural impacts on program variation.

Finally, the study shows that across all programs, regardless of their national context, the role of the practitioner as mentor and trainer is gaining in currency, with alignment to frameworks often being the primary academic contribution. Where regulators delegate provision to accredited universities, professional and non-profit agencies, a measure of choice may be offered to candidates. This appears a significant option in larger jurisdictions. However, providers should carefully consider the differing expertise that

academics and practitioners can bring to leadership preparation programs. The increasing primacy of practitioners over academic staff in program implementation appears to lend credibility and relevance to leadership preparation programs. Although, the implications of this development are not fully understood, it behooves scholars to consider specifically how their role in the tripartite relationship can better inform leadership preparation programs and the potential of engaging stakeholders through diffused delivery options (Walker and Dimmock, 2005; Walker, 2015).

This article updates our previous publication: Walker, A., Bryant, D., Lee, M., 2013. International patterns in principle preparation: commonalities and variations in pre-service programs. *Educ. Manag. Adm. Leader.* 41(4), 405–434.

References

- Afonso, N., Costa, E., 2009. The influence of the program for international student assessment (PISA) on policy decision in Portugal: the education policies of the 17th Portuguese constitutional government. *Educ. Sci. J.* 10, 53–64.
- Australian Institute for Teaching and School Leadership, 2015. Australian Professional Standard for Principals and the Leadership Profiles. AITSL, Melbourne.
- Barber, M., Whelan, F., Clark, M., 2010. Capturing the Leadership Premium: How the World's Top School Systems are Building Leadership Capacity for the Future. McKinsey and Company. Retrieved from: http://www.mckinsey.com/client-service/Social_Sector/our_practices/Education/Knowledge_Highlights/~/_media/Reports/SSO/schoolleadership_final.ashx. (Accessed 15 February 2011).
- Barnett, B., Okilwa, N., 2020. Preparation programs for school leaders. In: Barnett, B.G., Okilwa, N.S.A. (Eds.), *Oxford Research Encyclopedia of Education*. Oxford University Press, Oxford.
- Bastow Institute of Educational Leadership, 2020. Browse Professional Learning. Retrieved from: https://www.bastow.vic.edu.au/professional-learning/browse?tid=19&field_course_mode_value=All&tid_1=All&tid_2=All&tid_3=All. (Accessed 24 August 2020).
- Bryant, D.A., Walker, A., Lee, M., 2012. The impact of international, national, and local forces on the enactment of quality leadership preparation programs. In: Brundrett, M. (Ed.), *Principles of School Leadership*. Sage, London, pp. 279–308.
- Bush, T., Jackson, D., 2002. A preparation for school leadership: international perspectives. *Educ. Manag. Adm. Leader* 30 (4), 417–429.
- Carnoy, M., Loeb, S., 2002. Does external accountability affect student outcomes? A cross-state analysis. *Educ. Eval. Policy Anal.* 24 (4), 305–331.
- Cheng, Y.C., 2009. Hong Kong educational reforms in the last decade: reform syndrome and new developments. *Int. J. Educ. Manag.* 23 (1), 65–86.
- Corbin, J., Strauss, A., 2008. *Basics of Qualitative Research*, third ed. Sage Publications, Thousand Oaks, CA.
- Corcoran, S.P., Schwartz, A.E., Weinstein, M., 2009. The New York City Aspiring Principals Program: A School-Level Evaluation. Institute for Education and Social Policy, New York University. Retrieved from: <http://steinhardt.nyu.edu/esp/featured/APP>. (Accessed 1 February 2011).
- Council of Chief State School Officers, 2008. Educational Leadership Policy Standards: ISLLC as Adopted by the National Policy Board for Educational Administration. Retrieved from: http://www.ccsso.org/Resources/Publications/Educational_Leadership_Policy_Standards_ISLLC_2008_as_Adopted_by_the_National_Policy_Board_for_Educational_Administration.html. (Accessed 15 January 2012).
- Darling-Hammond, L., Meyerson, D., LaPointe, M., Orr, M.T., 2010. *Preparing Principals for a Changing World: Lessons from Effective School Leadership Programs*. Jossey-Bass, San Francisco, CA.
- Day, C., Sammons, P., 2013. *Successful Leadership: A Review of the International Literature*. CfBT Education Trust, CfBT Education Trust, Reading, Berkshire.
- Department of Education and Training, 2020. Framework for Improving Student Outcomes (FISO). Retrieved from: <https://www2.education.vic.gov.au/pal/fiso/policy?Redirect=1>. (Accessed 24 August 2020).
- Department of Education, 2007. *The Developmental Learning Framework for School Leaders*. Department of Education, Melbourne.
- Devi, M., Fernandes, V., 2019. The preparation of Fijian school leaders: a framework for principal preparation in a South Pacific context. *J. Educ. Adm. Hist.* 51 (1), 53–65.
- Eacott, S., Asuga, G., 2014. School leadership preparation and development in Africa: a critical insight. *Educ. Manag. Adm. Leader* 42 (6), 919–934.
- Education Bureau, 2002. *Continuing Professional Development for School Excellence—Consultation Paper on Continuing Professional Development of Principals*. Printing Department, Hong Kong.
- Education Bureau, 2020. *Continuing Professional Development of Teachers*. Retrieved from: <https://www.edb.gov.hk/en/teacher/qualification-training-development/development/cpd-teachers/index.html>. (Accessed 20 August 2020).
- Elmore, R., 2007. *Educational Improvement in Victoria*. Department of Education, Office for Government School Education, Victoria.
- Fryer, R., 2011. Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. NBER Working Paper, 16850. Retrieved from: <http://www.nber.org/papers/w16850>. (Accessed 20 February 2012).
- Gardner, H., 2008. *5 Minds for the Future*. Harvard Business Press, Boston.
- Gurmu, T.G., 2020. Primary school principals in Ethiopia: selection and preparation. *Educ. Manag. Adm. Leader* 48 (4), 651–681.
- Hattie, J., 2015. High impact leadership. *Educ. Leader* 72 (5), 36–40.
- Institute for Education Leadership, 2008. *Putting Ontario's Leadership Framework Into Action: A Guide for School and System Leaders*. Retrieved from: www.education-leadership-ontario.ca. (Accessed 28 January 2011).
- Kulophas, D., Hallinger, P., 2020. Leadership that matters: creating cultures of academic optimism that support teacher learning in Thailand. *J. Educ. Adm.* 58 (6), 605–627.
- Lee, M., Walker, A., Chui, Y.L., 2012a. Contrasting effects of instructional leadership practices on student learning in a high accountability context. *J. Educ. Adm.* 50 (5), 586–611.
- Lee, M., Louis, K.S., Anderson, S., 2012b. Local education authorities and student learning: the effects of policies and practices. *Sch. Eff. Sch. Improv.* 23 (2), 133–158.
- Leithwood, K., Levin, B., 2008. Understanding and assessing the impact of leadership development. In: Lumby, J., Crow, G., Pashardis, P. (Eds.), *International Handbook on the Preparation and Development of School Leaders*. Routledge, New York, pp. 280–300.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Successful School Leadership: What it is and How it Influences Pupil Learning*. National College for School Leadership, Nottingham, UK.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven Strong Claims About Successful School Leadership Revisited. *School Leadership and Management*, pp. 1–18.
- Leithwood, K., 2012. *The Ontario Leadership Framework 2012: With a Discussion of the Research Foundations*. The Institute for Education Leadership, Ontario.
- MacBeath, J., 2011. No lack of principles: leadership development in England and Scotland. *Sch. Leader. Manag.* 31 (2), 105–121.
- Marzano, R.J., Waters, T., McNulty, B.A., 2005. *School Leadership that Works*. ASCD, Alexandria, VA.
- Matthews, P., Moorman, H., Nusche, D., 2008. Building leadership capacity for system improvement in Victoria, Australia. In: Point, B., Nusche, D., Hopkins, D. (Eds.), *Improving School Leadership, Case Studies on System Leadership*, vol. 2. Organization for Economic Co-operation and Development, pp. 179–213.
- Ministry of Education, 2005. *Framework for Teach Less, Learn More: Recommendations*. Retrieved from: <http://www3.moe.edu.sg/bluesky/tllm.htm>. (Accessed 28 February 2011).
- Ministry of Education, 2020. *Desired Outcomes of Education*. Retrieved from: <https://beta.moe.gov.sg/education-in-SG/desired-outcomes>. (Accessed 24 August 2020).
- National Governors Association, Council of Chief State School Officers and Achieve Inc., 2008. *Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education*. National Governors Association, Washington, DC.
- National Institute of Education, 2010. *Leaders in Education Program: Handbook for Participants*. National Institute of Education, Nanyang Technological University, Singapore.

- National Institute of Education, 2011. *Leaders in Education Program: Handbook for Participants*. National Institute of Education, Nanyang Technological University, Singapore.
- National Policy Board for Educational Administration, 2015. *Professional Standards for Educational Leaders*. National Policy Board for Educational Administration, Reston, VA.
- New York City Department of Education, 2019a. *Instructional Leadership Framework: Building System-Wide Coherence to Accelerate Learning for Every Student*. New York City Department of Education.
- New York City Department of Education, 2019b. *Regulation Governing the Selection, Assignment and Appointment of Principals and Assistant Principals*. New York City Department of Education.
- New York City Leadership Academy, 2012. *Aspiring Principals Program*. Retrieved from: <http://www.nycleadershipacademy.org/aspiring/aspiring>. (Accessed 18 February 2012).
- New York City Leadership Academy, 2019. *Aspiring Principals Program: Impacting the Pipeline*. NYC Leadership Academy.
- Ng, P.T., 2008. Developing forward-looking and innovative school leaders: the Singapore leaders in education program. *J. Serv. Educ.* 34 (2), 237–255.
- Normore, A., Lahera, A., 2019. The evolution of educational leadership preparation programs. *J. Educ. Adm. Hist.* 51 (1), 27–42. <https://doi.org/10.1080/00220620.2018.1513914>.
- Okoko, J.M., 2020. Framing school leadership preparation and development for Kenya: context matters. *Educ. Manag. Adm. Leader* 48 (2), 396–413.
- Ontario College of Teachers, 2011. *Course Providers*. Retrieved from: http://www.oct.ca/additional_qualifications/providers/?lang=en-CA. (Accessed 20 January 2012).
- Ontario College of Teachers, 2017. *Principal's Qualification Program Guideline*. Ontario College of Teachers, Toronto.
- Ontario College of Teachers, 2020. *Principal's Development Course*. Retrieved from: https://www.oct.ca/members/additional-qualifications/schedules-and-guidelines/principals-development?sc_lang=en. (Accessed 24 August 2020).
- Ontario Principals Council, 2011a. *Principal's Qualification Program Part I: Candidate Acceptance Package*.
- Ontario Principals Council, 2011b. *Principal's Qualification Program Part II: Candidate Acceptance Package*.
- Ontario Principals Council, 2020. *Principal's Qualification Program*. Retrieved from: <https://www.principals.ca/en/professional-learning/principals-qualification-program-pqp.aspx>. (Accessed 24 August 2020).
- Ontario, 2010a. *Ontario Leadership Strategy: Principal/Vice-Principal Performance Appraisal, Technical Requirements Manual*. Queen's Printer for Ontario.
- Ontario, 2010b. *Ontario Leadership Strategy: Mentoring for Newly Appointed School Leaders, Requirements Manual*. Queen's Printer for Ontario.
- Perez, L.G., Ulline, C.L., Johnson Jr., J.F., James-Ward, C., 2011. Foregrounding fieldwork in leadership preparation: the transformative capacity of authentic inquiry. *Educ. Adm. Q.* 47 (1), 217–257.
- Phillips, D., Ochs, K., 2003. Processes of Policy Borrowing in Education: some explanatory and analytical devices. *Comp. Educ.* 39 (4), 451–461.
- Piggot-Irvine, E., 2011. Principal development: self-directed project efficacy. *Educ. Manag. Adm. Leader* 39 (3), 283–295.
- Rautalin, M., Alasuutari, P., 2009. The uses of the national PISA results by Finnish officials in central government. *J. Educ. Pol.* 24 (5), 539–556.
- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Sanzo, K.L., Myran, S., Clayton, J.K., 2011. Building bridges between knowledge and practice. *J. Educ. Adm.* 49 (3), 292–312.
- Sergiovanni, T.J., 2009. *The Principalship: A Reflective Practice Perspective*. Pearson, Boston.
- Slater, C.L., Garduno, G.J.M., Mentz, K., 2018. Frameworks for principal preparation and leadership development: contributions of the International Study of Principal Preparation (ISPP). *Manag. Educ.* 32 (3), 126–134.
- Sun, J., Leithwood, K., 2015. Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *Sch. Effect. Sch. Improv.* 26 (4), 499–523.
- Townsend, T., 2011. Changing times: new issues for school leaders. *Sch. Leader. Manag.* 31 (2), 91–92.
- Walker, A., Dimmock, C., 2005. Developing leadership in context. In: Coles, M., Southworth, G. (Eds.), *Developing Leadership: Creating the Schools of Tomorrow*. Open University Press, Milton Keynes, pp. 80–94.
- Walker, A., Dimmock, C., 2006. Preparing leaders, preparing learners: the Hong Kong experience. *Sch. Leader. Manag.* 26 (2), 125–147.
- Walker, A., Ko, J., 2011. Principal leadership in an era of accountability: a perspective from the Hong Kong context. *Sch. Leadersh. Manag.* 31 (4), 369–392.
- Walker, A., Kwan, P., 2010. Leadership in diverse cultures. In: Peterson, P., Baker, E., McGraw, B. (Eds.), *International encyclopedia of education*, Third ed. Elsevier, Oxford, pp. 824–831.
- Walker, A., Dimmock, C., Chan, A., Chan, W.K., Cheung, M.B., Wong, Y.H., 2002. *Key Qualities of the Principalship in Hong Kong*. Centre for the Development of Educational Leadership, Hong Kong.
- Walker, A., Lee, M., Bryant, D.A., Qian, H., 2011. *Principal Preparation Programs: A Synthesis of Best Practice Internationally*. National College for Leadership of Schools and Children's Services, UK [Confidential]. Submitted to.
- Walker, A., Bryant, D., Lee, M., 2013. International patterns in principle preparation: commonalities and variations in pre-service programs. *Educ. Manag. Adm. Leader* 41 (4), 405–434.
- Walker, A., 2015. Clones, drones and dragons: ongoing uncertainties around school leader development. *Sch. Leader. Manag.* 35 (3), 300–320.
- Waters, L., Luck, F., 2011. *Master of School Leadership: 2011 Program Information Handbook*. Melbourne Graduate School of Education.

The development of educational leaders: history and contemporary practice in the United States

Michelle D. Young^{a,b}, ^a Loyola Marymount University, Los Angeles, CA, United States; and ^b University Council for Educational Administration, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

The history of leadership development in the US	370
Educational leadership: an evolving understanding	372
Defining educational leadership	372
Researching quality preparation	373
Leadership development	374
Student selection	374
Cohorts	375
Curriculum	375
Teaching	375
Field-based experiences	377
Partnerships	377
Program faculty	378
Conclusion	378
References	379

Quality leadership matters. A growing body of knowledge highlights the importance of educational leadership to school improvement, teacher effectiveness and student learning (Leithwood and Jantzi, 2000; Robinson et al., 2008). In fact, school leadership is second only to teacher effects in its impact on student learning (Leithwood and Seashore, 2004). Likewise, there is growing understanding of the importance of effective leadership development to high quality practice (Young et al., 2009). Given the importance of educational leadership to student learning and the importance of preparation to leadership practice, effectively developing educational leaders could not be more important.

Research accumulated over the last two decades has enhanced our understanding of the features and characteristics of leadership development that positively influence the knowledge, experiences, and skills needed for successful leadership practice (Young et al., 2009). Leadership development also shapes leadership identity, assumptions about leadership, professional aspirations and career paths.

Importantly, many educational leadership faculty members have capitalized on research and development work in the field to redesign the content and delivery of their leadership development programs to have a stronger and more positive influence on their candidates' learning, their subsequent leadership practices, and their ability to foster school effectiveness (Cosner et al., 2015). Over the past 2 decades, significant attention has been directed to innovation and best practice in leadership preparation programs, both within and outside of universities (Jackson and Kelley, 2002; Young et al., 2009; Young and Crow, 2017).

This article has several goals. In its design, it describes the broad contours as well as the core structures of leadership development in the United States (US). According to Kandel (1970), to truly understand and appreciate an education system, it is essential to know something of its history and traditions. Thus, the article opens with brief history of educational leadership development in the United States, highlighting major markers in the development of the field. The three subsequent sections build on three key elements of the historical overview: (1) the field's evolving understanding of educational leadership, (2) efforts to professionalize the field through the development of a knowledge-base and standards, and (3) work to build a research-based understanding of quality leadership preparation. This is followed by a review of the key features of leadership preparation in the United States.

While the diffusion of research and concepts on preparation from the US and other countries has gone on for many years, and while some issues and trends may be relevant to scholars and educators in other national contexts, it is unlikely that work is universally applicable. This article offers an analysis of US-based educational leadership development research, rather than proposing a definitive statement about leadership preparation, writ large.

The history of leadership development in the US

As in many countries, the position of school leader in the US was originally bestowed upon head teachers. Over time, as school systems in the US have become more complex and definitions of educational leadership have evolved, the requirements of taking a leadership position have increased. In the majority of states, in order to lead a school, one must hold a special license or

certification that affirms their expertise to hold an educational leadership position. Typically, certification is tied to the completion of a graduate level educational leadership program, and in some cases, earning a master's degree.

The preparation of educational leaders in the US can trace its origins to the early part of the twentieth century, though it came into its own in the 1950s. Since that time, leadership development in the US has consisted of both pre-service for aspiring administrators and in-service preparation, both of which are designed to develop leaders in specific knowledge and skill domains. For the most part, pre-service preparation is provided by universities and in-service preparation is provided by professional organizations or through collaborative efforts between universities and school districts following initial preparation.

In the early years, the preparation of educational leaders borrowed concepts from private sector management, reflecting a concern with efficiency and an emphasis on rationalizing educational organizations (Black and Murtadha, 2007; Murphy, 2003). Drawing from the behavioral sciences, leadership research sought to identify key best practices or models of effective leadership that could be generalized across contexts and taught to future educational leaders. Subsequently, educational leadership scholars began borrowing from fields like human relations, sociology, political science, etc. Ideas and frameworks borrowed from these fields were used to guide leadership development and to explore what leaders did and how they thought and felt.

Beginning in the mid-1980s, both the preparation and professional development of educational leaders became the focus of much criticism and reform (Young et al., 2015). Several national commissions, task forces, and individual scholars examined the nature of leadership and university-based leadership preparation programs. These efforts were instrumental in shedding light on how demands for reform were altering the nature and direction of leadership development programs. The first major effort to address criticisms of the field was undertaken in the mid-1980s when the National Commission on Excellence in Educational Administration (NCEE) was developed. In 1987, the NCEE issued a set of eight recommendations:

Calling for the redefinition of educational leadership; the establishment of a national board to shape policy related to school administration; the modeling of administrator preparation programs after other professional schools; a reduction in the number of preparation programs; the increased recruitment, preparation, and placement of ethnic minorities and women; the establishment of partnerships with the public schools in the preparation of school administrators; increased emphasis on professional development for practicing administrators; and reform of licensure and certification.

Forsyth (1999, p. 75).

These recommendations influenced thinking about school leadership development and were widely endorsed across the 50 states in the US. Following the release of NCEE's report, efforts focused on supporting the quality of leadership preparation intensified. For example, the Danforth Foundation launched a major initiative to support preparation program redesign and innovation, the National Policy Board for Educational Administration (NPBEA) was established, and a multi-state initiative was launched to develop standards to guide the preparation and practice of educational leaders (Young et al., 2002).

Although many of NCEE's recommendations were implemented and made significant and positive impacts on the field, satisfaction with their impact was short lived. By the early 2000s, a new wave of critique was in full swing and they were coming from many directions. In 2006, the Southern Regional Education Board (SREB) report, *Schools Can't Wait: Accelerating the redesign of university educational leadership preparation programs* argued:

There is a lack of urgency for refocusing the design, content, process, and outcomes of principal preparation programs based on the needs of schools and student achievement and little will happen until there are committed leaders of change at every level—state, university, and local school district.

Fry et al. (2006, p. 4).

The report also criticized faculty independence in course content development and resistance to changing current coursework. Similarly, the report, *Educating School Leaders*, launched a high-profile critique of the quality of educational leadership programs, arguing that leadership preparation programs lacked purpose, curricular coherence, adequate clinical instruction, appropriate faculty, and high admission standards (Levine, 2005). The report also questioned district and state policies that based salary schedules on the acquisition of graduate degrees and credit hours and claimed that the field would be better served by requiring a Masters of Business Administration.

Critiques also focused on program assessment. Here, critics argued that preparation programs had not systematically examined the efficacy of their own practice. Research on practicing school leaders indicated that preparation programs did not prepare them sufficiently for the myriad demands of their job (Fry et al., 2006; Schulman et al., 2006). Due to the lack of focus on program assessment, programs had little evidence from which to respond to such critiques. Soon thereafter, the University Council for Educational Administration (UCEA) took several important steps to rectify this situation. First, in 2000, UCEA launched a leadership preparation program evaluation initiative. Subsequently, in 2001, it spearheaded the National Commission for the Advancement of Educational Leadership Preparation (NCAELP), and in 2004, it formed a Joint Research Taskforce on educational leadership preparation. These efforts will be described in greater detail later in this article.

Educational leadership: an evolving understanding

Grogan and Andrews (2002), described early university leadership preparation programs as preparing aspiring administrators for the role of a top-down manager, emphasizing courses on management, administration, planning, supervising, budgeting, scheduling, etc. In their historical overview of the preparation of school leaders, Murphy and Vriesenga (2004) concurred with this characterization and charted several distinct waves of educational leadership scholarship, each with its own impact on leadership development. The periods included the ideological era (1820–1900), the prescriptive era (1900–1915), the scientific era (1947–1985) and the dialectic era (1986–2006). Over time, they argued, the content and character of leadership preparation in the US changed, with the most recent major shifts occurring in the 1980s shifting from a more conventional emphasis on school management to an emphasis on leadership, shared governance, participatory decision making, and school-based councils; in the late 1990s shifting to learning focused leadership; and in the early 2000s adding distributed and equity orientations to learning focused leadership.

In the 1990s, several ideas were particularly influential on the field of educational leadership in the US: learning organizations and leadership for learning. The notion of a learning organization, introduced by Senge (1990) shifted the way many scholars thought about schools from static organizations that required technical management to dynamic, non-stable, and self-generating organizations that required leadership and management (Senge, 1990). Additionally, a focus on leading learning, which emerged from the effective schools' literature shifted emphasis to specific ways educational leaders exercise a powerful influence on student achievement (Leithwood and Jantzi, 2000). Finally, a focus on working for equity emerged in the 2000s as a counterforce to the accountability movement, placing emphasis on the current and future needs and aspirations of each and every child. In response, leadership preparation programs began emphasizing transformational and instructional leadership (Firestone et al., 2004).

In many ways, these adjustments in focus reflect larger changes in theory, policy and general society. They also reflect efforts within the field to define educational leadership, to identify a knowledge base and to develop a set of national standards for leadership preparation and practice. Furthermore, they reflect critiques of such efforts (Dantley and Tillman, 2006; Donmoyer et al., 1995; Marshall and Oliva, 2006). Critical scholars were particularly active during the late 1990s and early 2000s, exposing the limitations of leadership models that downplayed context and equity, while overemphasizing heroic notions of leadership and leadership characteristics (Bennett and Anderson, 2003). Many of these same scholars, argued that traditional notions of leadership were "inherently biased toward maintaining stability, and thus served to reproduce inequity in education" (Black and Murtadha, 2007, p. 2).

Defining educational leadership

Starting in the mid-1980s, several efforts were undertaken to define the knowledge base in educational leadership and to use that knowledge base to develop a set of standards for the field (Donmoyer et al., 1995). In the early years, efforts to define standards were undertaken chiefly by organizations involved in the NPBEA with university faculty and practitioners playing a limited role; however, in more recent years the scope of engagement has enlarged significantly.

In the mid-1990s, UCEA engaged scholars from across the country in defining the knowledge base in educational leadership. Considered a controversial project by some, the UCEA Knowledge Base project resulted in the identification of seven knowledge domains: (1) societal and cultural influences on schooling, (2) teaching and learning processes, (3) organizational studies, (4) leadership and management processes, (5) policy and political studies, (6) legal and ethical dimensions of schooling, and (7) economic and financial dimensions of schooling. In support of these seven areas, which "currently recognized for comprising the field" and serve as organizers for mapping educational administration," UCEA scholars developed a document warehouse of what they considered to be essential readings on educational administration (Bredeson, 1995, p. 52). The document warehouse, titled *Primis* and published for UCEA by McGraw Hill, enabled faculty to develop course readers from the identified knowledge base.

Soon thereafter, the NPBEA initiated the development of the Interstate School Leadership Licensure Consortium (ISLLC) standards for the professional practice of school leaders. The purpose of the ISLLC standards, according to consortium's sponsoring organization—the Council of Chief State School Officers (CCSSO)—was to provide a clear, organized set of criteria that could be used to drive the preparation, professional development, and licensure of school leaders. Murphy (2003), noted that two questions were central to the standard setting work: (1) "What do we know about schools where all youngsters achieve at high levels?" and (2) "What do we know about the actions and values of the women and men who led effective schools and productive school systems?" (pp. 7 and 9). Extensive research reviews were conducted on each of these questions to identify the key responsibilities and practices of effective educational leaders. The resulting ISSLC standards provided a foundation for the field of educational leadership that was significantly different and more comprehensive than previous notions of educational leadership (Young et al., 2009). According to Murphy (2003), the standards pushed the field from a focus on administration and management to leadership and learning.

However, the ISSLC standards have not been immune to criticism (Canole and Young, 2013). Some of the more significant and recurring concerns included: (1) a lack of direct connection between the leadership standards and supporting student achievement gains, (2) the omission of specific areas like school technology leadership and data use, (3) the under-specification of criteria to be met, (4) the lack of consideration given to the role of context in leadership practices, (5) an assumption that leadership is provided

by a single person, (6) the lack of focus on issues of diversity, equity or inclusion, and (7) the failure to identify the empirical knowledge/research base upon which the standards are based.

Nonetheless, the ISSLC standards substantially influenced preparation and professional development in the US. By 2001, 35 states had adopted or adapted the standards for use in guiding the practice, preparation and professional development of educational leaders. Complementing this effort, the Education Testing Service (ETS) used the ISSLC standards to develop a performance-based assessment, the School Leaders Licensure Assessment (SLLA), to assess proficiency in each of the standard areas and thus the effectiveness of leadership preparation and professional development. By 2008, when the ISSLC standards were last revised, almost 40 states had adopted or adapted the standards as platforms for licensure, evaluation and professional development, and many also required universities to align their leadership preparation programs to the standards.

The ISSLC standards were moderately revised in 2008 and then overhauled in 2015, resulting in the adoption of the Professional Standards for Educational Leaders (PSEL). The PSEL standards include the following ten domains:

1. Mission, Vision, and Core Values
2. Ethics and Professional Norms
3. Equity and Cultural Responsiveness
4. Curriculum, Instruction and Assessment
5. Community of Care and Support for Students
6. Professional Capacity of School Personnel
7. Professional Community for Teachers and Staff
8. Meaningful Engagement of Families and Community
9. Operations and Management
10. School Improvement (NPBEA, 2015, p. 3)

Notably, the standards emphasize the importance of both academics as well as the support and care required for students to excel. “The 2015 Standards have been recast with a stronger, clearer emphasis on students and student learning, outlining foundational principles of leadership to help ensure that each child is well-educated and prepared for the 21st century” (NPBEA, 2015, p. 3). Collectively, the ISSLC 1996, ISSLC 2008 and PSEL 2015 standards have provided guidance for leadership preparation, practice and evaluation at the state level for over 20 years.

According to Crow et al. (2008), similar sets of competencies have been adopted in other countries. Despite widespread acknowledgment that cultures vary considerably and that leadership preparation and development could be adjusted in relation to culture, the content, method of delivery and assessment of preparation and development shows relatively little variation throughout the world (Bush and Jackson, 2002).

Researching quality preparation

In 2000, the NPBEA established a working group with representatives from AACTE, ASCD, AASA, NASSP, NCPEA, NAESP, NASB and UCEA to design program accreditation standards for the National Council for the Accreditation of Teacher Education’s (NCATE) Specialized Professional Association (SPA) review of educational leadership programs. These standards, which were adopted by NCATE in the spring of 2002, have undergone several revisions and were recently revised and renamed the National Educational Leadership Preparation (NELP) standards (Young et al., 2018). The new version, which is aligned to the PSEL standards is strongly anchored to the research based on leadership practice and effective preparation (Young et al., 2017). The NELP standards are not only used in national accreditation, but they are also used by states for program review and approval (National Policy Board for Educational Administration, 2018a, b).

In addition to these national efforts, researchers have worked to develop an empirical base of knowledge to guide preparation program improvement efforts. Both individual researchers and organizations have participated in these efforts. For example, in 1999 UCEA conducted a study of school principals and superintendents, their jobs, preparation and the problems they faced. In 2000, UCEA affiliated with Division A of the American Education Research Association and the Learning and Teaching in Educational Leadership (LTEL) special interest group to develop a joint research taskforce to design evaluation tools and support research on educational leadership preparation. The field needed to have useful data for program improvement and accreditation purposes.

In 2003, following the recommendations of the National Commission on the Advancement of Educational Leadership Preparation (NCAELP), UCEA spearheaded a Joint Taskforce on Research on Leadership Education under the sponsorship of UCEA, NCPEA (National Council of Professors of Educational Administration), Division A of AERA (American Educational Research Association), and the TEA-SIG of AERA. Despite methodological challenges, the taskforce developed multiple research strands, including studies focused on specific program attributes, longitudinal studies of preparation programs and their impact, and backward mapping studies that examined effective schools, the nature of their leadership, and the source of their leadership skill development.

These two taskforces engaged in individual and joint research on leadership preparation, leading to significant increases in the knowledge base. Their work, along with expert opinion, provided important insight into a set of essential core program attributes, including a well-defined, leadership-for-learning focus; coherence; challenging and reflective content; student-centered instructional

practices; competent faculty; positive student relationships; a cohort structure; supportive organizational structures; and substantive and lengthy internships (Darling-Hammond et al., 2007; Jackson and Kelley, 2002; Young et al., 2009). Research continues to evolve around the above-mentioned preparation program features as well as their effects on graduates' knowledge, movement into leadership positions, and leadership practices.

To support this work, UCEA established new outlets for publishing and disseminating research, including the first peer reviewed journal focused on school leadership development, the *Journal of Research on Leadership Education* (JRLE) as well as two editions of the *Handbook of Research on Leadership Education* (Young et al., 2009; Young and Crow, 2017). In addition to fostering the development of a base of research on leadership education that provides insight into the development and delivery of leadership education, these publications also provide information, case descriptions, and data that enable larger conversations about leadership education.

As noted above, UCEA has played a pivotal role in developing a process for documenting the practices of educational leadership preparation programs. In 2011, a group of researchers from UCEA and the Utah Education Policy Center developed and validated the INSPIRE Leadership Survey Suite, which examines the programmatic elements of educational leadership preparation programs and the experiences of program participants both while they are in the program and once they have taken leadership positions. The INSPIRE Leadership Surveys are aligned to PSEL and NELP standards and provide data to help programs examine relationships between their program features, graduate student experiences and their practice once they assume leadership roles (Ni et al., 2019). Today, the INSPIRE leadership survey suite, which includes a program, graduate, leader in practice and 360 surveys, is used by many programs in the US to document their programs and gather data from graduates and their employing organizations. The critical friends review process, which involves having external site visitors review and provide feedback on program quality, is also commonly used by programs in the US.

The above described efforts have had an important impact on preparation programs in the United States. For example, an increasing number of programs have adopted research-based program features, and many have increased opportunities for candidates to deepen their understanding of critical leadership practices, such as building a collaborative vision, supporting and advocating for student learning, fostering and engaging in data-informed decision making, developing a supportive organizational and instructional culture, seeking out and incorporating diverse perspectives, engaging stakeholders, acquiring and effectively using resources, and reflecting on how their knowledge bases interact with each other and how to apply them (Leithwood et al., 1996; Young et al., 2009).

Leadership development

What has resulted from the increased emphasis on quality preparation? What does preparation entail in the US? Research on leadership preparation has identified a number of program features as essential to high quality preparation. These include: (a) selective admissions, (b) cohort structures, (c) curricular coherence, (e) powerful learning experiences, (d) field-based experiences, (f) mentoring and coaching, and (g) collaboration between universities and school districts (Darling-Hammond et al., 2007; Jackson and Kelley, 2002; Young and Crow, 2017; Young et al., 2021). In the following sub-sections, I review each of these leadership preparation program features. In doing so, I highlight a number of US leadership preparation programs that have been identified by experts in the field as exemplary or innovative (Jacobson et al., 2015; Young, 2015; Young et al., 2021) or which have been recognized by the UCEA Exemplary Educational Leadership Preparation (EELP) award.

Student selection

Selective admissions is a key characteristic of exemplary preparation programming (Fry et al., 2006). In the typical program, applicants are identified, selected, and screened to reflect leadership potential. How these functions are conducted and why differentiates the average program from the exemplary. In most programs, applicants are asked to submit their academic credentials, an essay and a set of recommendations from individuals who they feel can speak to their leadership potential. Based on this information, the applicant's grade point average, and their score on the Graduate Record Exam (GRE), program faculty select candidates who they believe will gain from and contribute to the program in meaningful ways.

Student selection practices in exemplary programs may use some of these same resources for the first phase of the screening process, but not to determine who they will ultimately invite to participate in the program. Rather, selection processes in exemplary programs usually develop from the program's purpose and theory of action (Young et al., 2021). That is, how the program goes about selecting candidates for the program will reflect the program's purpose and assumptions about leadership and its development. For example, the process used by the University of Texas at Austin assumes that school leaders must be able to transform schools into educational environments that work for each and every child, and, as a result, program faculty select candidates who are excellent teachers and who can provide evidence of teacher leadership. In addition to evaluating each application portfolio, a team of observers visits each applicant's classroom to assess teaching quality and to personally interview the applicant and his or her supervisor.

At the University of Illinois-Chicago, the selection process begins with the review of candidate application materials (Cosner et al., 2015). Potential candidates must have a master's degree and a grade point average (GPA) score of 3.5 or better. The faculty also review the candidates' GRE scores, personal statements, and letters of reference. In their review, they look for evidence of

successful teaching and leadership of adults to achieve shared goals. The next phase of the selection process involves a campus interview experience. Applicants who are selected for campus visits are forwarded case material on a struggling school in advance of the interview. During the campus interview, candidates: (1) observe and orally respond to questions about an instructional video, (2) present an analysis of a struggling school, (3) respond to an oral examination of his/her portfolio, and (4) respond to a writing prompt.

Cohorts

In many exemplary programs applicants are selected as a cohort (VanGronigen et al., 2018; Young et al., 2021). This is often based on the assumption that part of the learning that takes place in leadership development is peer to peer as well as the desire to foster a sense of community, social capital, and enhanced academic depth and to develop a mix of students that will support one another's learning and development (Barnett et al., 2000; Young et al., 2009). Ninety percent of UCEA programs use a full (55%), partial (13%), or de facto (13%) cohort model.

A full cohort model is a defined group of individuals who begin the program together and stay together throughout the entire program, a partial model is where there is defined group of individuals for some of their courses, and the more informal de facto model is where students enter at the same time and take the majority of the same courses and in the same sequence but are not deemed a cohort.

Anderson et al. (2017, p. 17).

For example, Loyola Marymount and Brigham Young Universities use cohorts with the intention of developing professional learning communities and career-long networks. Similarly, the University of Washington, the University of Texas, and the University of Illinois-Chicago organize students into cohorts for the duration of their program offering highly specified and sequenced courses that enhance cohort-based learning.

Curriculum

In effective programs, there is a clear, well-defined curriculum that reflects agreement on the relevant knowledge and skills needed for educational leaders in their first year, or first few years in the profession (Young et al., 2021). As such, contemporary leadership preparation curriculum reflects the heightened emphasis on school improvement (e.g., courses focused on change, applied research and data analysis), equity and social justice (e.g., courses on diversity, culturally responsiveness, advocacy and engagement), and talent management (e.g., courses on hiring, building a collaborative learning culture, and teacher supervision and development). For example, in the University of Illinois-Chicago program, the curriculum is designed to support the program's logic model which focuses on the processes through which leadership supports student learning and school transformation. The program curriculum has three strands: (a) instructional diagnosis and development, (b) organizational/leadership diagnosis and development, and (c) cycles of inquiry for school-wide improvement (Cosner et al., 2015).

Although program faculty ultimately determine course content, national, state and, in some cases, district leadership standards are used to frame preparation program focus and alignment. Subsequent to the adoption of the ELCC or NELP standards for the accreditation review of leadership preparation programs, their impact was significant (Canole and Young, 2013; Young et al., 2009). Over a third of all leadership preparation programs in the US had received national recognition from NCATE between 2002 and 2012, and that number grew under the Council for the Accreditation of Educator Preparation (CAEP). According to Young et al. (2017) that number is now close to 1000.

Like accreditation, state licensure policy has also supported the adoption of a default "national curriculum" in educational leadership preparation. Only a few of the states in the US have resisted adopting the ISSLC standards since their development in the late 1990s, and many are in the process of adopting the latest version, the PSEL standards (Young et al., 2017). As noted in a previous section, state adoption of standards not only provides a common vision of leadership preparation, it tends to impact the content and focus of preparation, practice, evaluation and professional development.

Importantly, a curriculum based on standards, such as the NELP standards, conveys a hidden curriculum (Osterman and Hafner, 2009). That is, what is included in the standards conveys a larger conception of what counts as educational leadership, what knowledge and skills novice leaders should accrue before taking a position, as well as what quality educational leaders do. As such, other thoughts and practices that do appear within the standards and components are excluded and by their exclusion can be interpreted as not essential to the practice of beginning leaders (Young et al., 2017). The standardization of what counts as leadership (and what does not), as well as what knowledge is considered an essential part of the leadership knowledge base, has drawn significant criticism over the years.

Teaching

The short history of preparation for educational leaders is punctuated by research and development activity around the issue of pedagogy. UCEA archives are full of simulation and case study materials dating back to the late 1950s. Simulations and cases were considered to be the state of the art in the 1950s, just as problem-based learning and the use of district data-sets were considered cutting-edge in the late 1990s (Bridges and Hallinger, 1997; Copland, 2003). Today, evidence suggests the continued use of

Table 1 Powerful learning experiences framework.

Features	Descriptions
Authentic	The learning experience reflects an authentic, meaningful, and relevant aspect of leadership practice.
Active engagement	The learning experience requires active engagement (e.g., examining, diagnosing, and addressing problems of leadership practice, linking theory and practice).
Sense-making	The learning experience fosters sense-making around critical problems of leadership practice.
Centers equity	The learning experience requires that candidates explore, critique, and deconstruct problems, beliefs, practices, and policies from an equity perspective.
Reflection	The learning experience requires reflection.
Collaboration and interdependence	The learning experience requires collaboration and interdependence.
Responsibility for learning	The learning experience empowers learners to take responsibility for their own learning.
Learner and knower	The learning experience positions both professor and students as both knowers and learners.
Broadens perspective	The learning experience broadens and shifts perspective from the classroom to school, district, or state level.
Confidence building	The learning experience develops confidence in leadership.

problem-based learning, case studies and the encouragement of individual and group reflection (Young and Crow, 2017). In a study of UCEA programs, Anderson et al. (2017) found that “UCEA programs create learning experiences that reflect the job of a practicing administrator by assigning field-based, course-related projects (95%), using problem-based learning (83%), requiring action research or inquiry projects (75%), and analyzing case studies (78%)” (p. 12).

Problem-based and case-based teaching methods are increasingly the primary mode of teaching in educational leadership preparation programs. These pedagogical tools offer situated learning and an opportunity to experiment with multiple perspectives within a traditional classroom setting. Representing real world situations, case studies and simulations can provide powerful learning opportunities that draw attention to meaning making, patterns of behaviors, beliefs and values, and economic and political structures. Similarly, problem-based learning, an approach adapted for educational leadership by Bridges (1992), provides an opportunity to ground aspiring leaders in the problems of their context, expanding their problem-framing and problem-solving capabilities, and working on problems collaboratively with others.

Building on some of the work in adult learning, in the early 2000s UCEA fostered conversations around the development of a signature pedagogy in school leadership development, one that would include problem-focused learning, action research and other transformative learning experiences (Mezirow, 1991; Black and Murtadha, 2007; Cunningham et al., 2019). According to Schulman (2005), a signature pedagogy is “... a set of assumptions about how best to impart a certain body of knowledge and know-how. And it has an implicit structure, a moral dimension, that comprises a set of beliefs about professional attitudes, values, and dispositions” (p. 55).

Drawing from developments in adult learning theory and principles, such as Mezirow (1991), the signature pedagogies promoted by UCEA scholars (e.g., Black and Murtadha, 2007; Young et al., 2017; Cunningham et al., 2019) incorporate experiential learning, reflective practice, structured dialog, problem-based learning and learning communities in their courses. For example, Cunningham et al. (2019), VanGronigen et al. (2018) and Young et al. (2021) promote the Powerful Learning Experiences (PLE) framework captured in Table 1.

Pedagogical features, like those included in the PLE framework address the concern that “school leaders too often remain overly localized with little of the peripheral vision that data, theory, and processes of inquiry might foster” (Black and Murtadha, 2007, p. 12).

Along these same lines, Young et al. (2017) suggest that aspiring leaders develop three core propensities: critical reflection, critical consciousness, and multiple perspective taking. The first propensity, critical reflection, is a process that enables a leader to examine long-held, socially constructed ideas, beliefs and assumptions about their experiences (Mezirow, 1991). Reflective practice is increasingly structured into and throughout programs as a means of provoking and challenging students’ presuppositions, helping them to surface theories-in-use that shape behavior, and enabling them to try on other frames for leadership development (Young et al., 2017). The second, critical consciousness, is a tool for engaging in “broad social analysis, to take a macro level perspective, which is essential to understanding education and its place in society, to thinking about it strategically and to potentially change it” (Young et al., 2017, p. 251). The third propensity, multiple perspective taking, involves the “ability and desire to consider issues from a variety of theoretical, experiential and epistemological perspectives” (p. 251). Multiple perspective taking enables leaders to better understand problems, issues, and relationships, to check their understanding or tenets, which were formerly accepted as given (Young, 1999). Together, critical reflection, critical consciousness and multiple perspective taking, enable a leader to see and think in new and fresh ways.

Field-based experiences

Although internships only began to appear in educational leadership training in the 1950s, field-based experiences are now considered a cornerstone of educational leadership preparation programs. [Reyes-Guerra and Barnett \(2016\)](#) report that around 90% of leadership preparation programs require the internship. While research on internships, practica, apprenticeships and other field-based experiences is less robust than other areas, available research points to its importance and state policy and national accreditation bolster its wide-spread use. According to [Reyes-Guerra and Barnett \(2016\)](#), two goals provide a foundation for the wide-spread use of internships:

On one hand, these experiences are intended to have general effects on participant learning, which include encouraging independent learning, improving self-confidence, promoting social justice and change, and increasing the commitment to school leadership as a career. On the other hand, internships are structured to strengthen learners' specific leadership and administrative skills, such as communication, personnel evaluation, collaboration, and resource management. By directly experiencing and practicing administrative tasks and responsibilities, aspirants learn about the role, how to perform the job, and why to take certain courses of action (p. 8).

The majority of states in the US require an internship experience as part of educational leadership preparation. However, requirements differ from state to state. For example, the state of North Carolina requires a yearlong internship for all administrator candidates, while Illinois requires a yearlong part-time internship. The state of Colorado does not specify whether the internship needs to be full-time or part-time; instead it requires a minimum of 300 h of supervised field-based experiences.

In some states, funding is provided to off-set the cost of these mandates. For example, the state of Washington offers a state-funded school leadership internship program wherein funds are provided directly to school districts to cover up to 45 days of release time for participating school leadership candidates. The state of Mississippi provides a limited number of year-long sabbaticals with full pay and benefits to teachers enrolled in educational leadership programs to support their participation in a full-time internship.

To achieve national recognition through the CAEP's NELP SPA, programs must have a high-quality internship. CAEP and NELP standards include a specific focus on clinical experiences for advanced programs in educational leadership. For example:

The provider works with partners to design clinical experiences of sufficient depth, breadth, diversity, coherence, and duration to ensure that advanced program candidates demonstrate their developing effectiveness in creating environments that support all students' learning and development. Clinical experiences, including technology-enhanced learning opportunities, are structured to have multiple performance-based assessments at key points within the program to demonstrate advanced program candidates' development of the knowledge, skills, and professional dispositions, as delineated in Standard 1, that are associated with creating a supportive school environment that results in a positive impact on the learning and development of all P-12 students.

CAEP (2015, p. 6).

Most internships involve over 300 contact hours and span at least two semesters. In some programs, the internship extends through a program's 2-year duration. In a handful of programs, like that of the University of Illinois-Chicago, the internship is full-time; however, in most programs the internship is a part time experience. In the latter situation, field-based experiences are typically integrated into course work. In these programs, the curriculum is sequenced and mapped against the annual cycle of regular work responsibilities of educational leaders and leadership standards and the program's vision of leadership for learning ([Young et al., 2021](#)). Furthermore, reflection, portfolio development and coaching and mentoring have been incorporated into internships to encourage theory-to-practice connections and generally make them more meaningful learning experiences.

A sizable amount of research has been devoted to the importance of mentoring and coaching as an important component of aspiring leaders' learning about leadership ([Reyes-Guerra and Barnett, 2016](#)). During clinical experiences, coaching and mentors serve a key role in developing aspiring leaders' capabilities to improve instructional practices, provide meaningful professional development for teachers, and lead school improvement initiatives.

Partnerships

Less than 20 years ago, district-university partnerships were small-scale initiatives involving districts and regional colleges and universities. Today, however, district and university collaborations around preparation are increasingly common and involve colleges and universities of many sizes. Moreover, the nature of the collaborations tends to be much more substantive and reciprocal, reflecting leadership needs for local contexts (e.g., high teacher turnover, turnaround schools). District leaders are no longer relegated to university program advisory boards; rather they participate as partners in program design and delivery.

District partnerships often involve a number of activities, include partnering with universities to identify and screen candidates, carefully prepare mentors, shape the curriculum content and delivery, and assist in assessing candidates' leadership capabilities. According to research on UCEA programs, the most frequent form of collaboration was for districts to partner with the university to support the organization and supervision of the internship (92%). Those programs with more extensive formal relationships had official memorandums of understanding (47%) and structured advisory committees (44%) ([Anderson et al., 2017](#)).

According to research on leadership preparation, learning is advanced through university–district partnerships that tailor preparation to the needs of the districts, more closely aligning theory to practice, and use internships that provide candidates with the opportunity to practice in core leadership responsibilities (Anderson et al., 2017). In a study of high-quality programs described as exceptional, Jackson and Kelley (2002) found that programs partnered in a variety of ways, including: (a) the selection and recruitment process, (b) the continued support of program elements, and (c) the organization and development of the clinical experience.

Additionally, partnerships often develop beyond the provision of pre-service experiences, to the development of leadership development pipelines, which include components of recruitment, preparation, selection, induction, development, and evaluation of school leaders. Hitt et al. (2012) conceptual model of the leadership professional pipeline divides the pipeline by preservice and in-service leadership. Research presented with this model demonstrates the growing knowledge base about effective leadership practices, important relationships between preservice and in-service, and the need for “strategic and intentional coordination of efforts by multiple stakeholders in leadership preparation and practice” (Hitt et al., 2012, p. 1). This model provides an outline for building, monitoring, and evaluating a professional pipeline from preparation to career-long learning.

Program faculty

The personnel staffing a program are fundamental to program quality (Young and Crow, 2017; Young et al., 2009). Thanks to a series of national surveys beginning in 1973 (Campbell and Newell, 1973) and continuing under the leadership of Martha McCarthy (Hackmann and McCarthy, 2011) we have a longitudinal knowledge base on faculty characteristics and trends over time. Most pertinent to this article are the following findings: (1) an increasing number of educational leadership faculty today have prior administrative experience, and (2) there are increasing numbers of part-time clinical and adjunct faculty in preparation programs and a decreasing number of full-time clinical faculty members (Hackmann and McCarthy, 2011).

According to Darling-Hammond et al. (2007), faculty knowledge is most closely associated with program quality. In their review of exemplary programs Jacobson et al. (2015) found that having a critical mass of faculty with strong content knowledge and experience was very important, as was the stability of program leadership and staff support.

At UCEA institutions, full-time faculty members provide the majority of instruction. UCEA programs use tenured or tenure-track faculty to teach over half (55%) of the courses, and clinical instructors teach a third (33%) of courses (Anderson et al., 2017). Part time instructors, such as adjuncts (often current practitioners) also contribute approximately 30% of instruction in UCEA programs. Furthermore, 99% of programs call upon practitioners to present to classes and other program events as guest speakers or panelists. In addition to teaching and presenting, practitioners are often utilized in supporting or supervisory roles, such as the supervision of clinical fieldwork as well as to help with programmatic decision making (89%) and design work. Thus, it appears that few UCEA programs are working in isolation in the proverbial “ivory tower;” rather, the majority work with practitioners to select candidates, design coursework and delivery of instruction, and to support the clinical component.

In addition to having a strong faculty, research also indicates that faculty community is an important condition of program quality and a focus on continuous improvement. The universities identified as exemplary by the UCEA EELP award all have very strong and engaged faculty communities (Jacobson et al., 2015; Young et al., 2021). For example, the UIC faculty culture focuses on continuous improvement which mirrors the approach taught within their program to aspiring leaders. The faculty have established work routines and cycles of inquiry that enable them to ensure they are continuously adapting the program to meet the needs of their student population.

Due to major breakthroughs in research as well as significant investments in program improvement and policy alignment to this research, preparation in the US, is becoming increasingly aligned with the research on exemplary preparation programs. How well this reflects preparation in other areas of the globe, remains to be seen.

Conclusion

Never has it been more important to have learning focused leaders who understand teaching and learning. “There’s nothing new or especially controversial about that idea,” yet how to ensure consistently high-quality leadership for all schools continues to be a challenge (Leithwood and Seashore, 2004, p. 1).

One of the challenges and opportunities of designing, delivering and researching educational leadership development is the dynamic nature of the field. Educational leadership is affected by rapid and evolving knowledge on learning, teaching and leading, the demands of society, the persistent demographic changes of schools, and the increasing complexity of the job of school leader. As a result, leadership preparation and development should be considered an evolving and dynamic area of work, wherein curricula, teaching methods and modes of delivery evolve with developing needs and understanding.

Furthermore, context matters. As noted in the introduction, this article offers a context-bound treatment of leadership development, firmly anchored in the US experience. What worked in the US, may work less well in other contexts. History, culture, language and the role and structure of education and policy systems differ from nation to nation, making the transfer or translation of practices from one context to another complicated. Nonetheless, several lessons learned through the US experience may prove important to a spectrum of contexts. For example, while the importance of leadership in education is increasingly considered common

knowledge across the globe, the idea that leaders require high quality preparation and development is less widely accepted. Yet, quality preparation and development has been proven to make a significant difference in the practice of school leaders in the US and could potentially make an equally important impact on the quality of education in other countries.

The US journey to understand what counts as effective leadership and leadership preparation provides a road map that prove useful to other nations. Key elements of the US journey included broad participation and collaboration, the development of leadership standards that reflected a wide-spread understanding of effective leadership practice and preparation, and the creation of and dependence on a research infrastructure to inform standards for preparation and practice. The importance of research cannot be overstated. It is through research that we have been able to identify the key areas of leadership to leverage for increased student learning and success, and it is through research that we continue to build our understanding of the kinds of development experiences that lead to effective leadership practice.

In the US and elsewhere, dedicating resources to understanding effective leadership preparation and practice is essential to building a quality education system. Although the body of research on educational leadership preparation and development in the US is growing, much more work is needed. The same is true in other countries. To date there has been little systematic or longitudinal work. Furthermore, what we know is based on a limited set of methodological approaches: descriptive, survey and cross-sectional methods. While there are important insights that can be gleaned from such work, particularly in the development of program innovations, they reveal less about what works and why. This should be viewed as an opportunity to develop a serious international research agenda that can contribute to the understanding of effective leadership in a variety of contexts, the design of leadership preparation and development, and the formation of policies that support them.

References

- Anderson, E., Winn, K., Young, M.D., Groth, C., Korach, S., Pounder, D., Rorrer, A., 2017. Examining university leadership preparation: an analysis of program attributes and practices. *J. Res. Leadersh. Educ.* 1–23. <https://doi.org/10.1177/1942775117735873>.
- Barnett, B.G., Basom, M.R., Yerkes, D.M., Norris, C.J., 2000. Cohorts in educational leadership programs: benefits, difficulties, and the potential for developing school leaders. *Educ. Adm. Q.* 36 (2), 255–282.
- Bennett, N., Anderson, L., 2003. In: *Rethinking Educational Leadership: Challenging the Conventions*. Sage, London.
- Black, W., Murtadha, K., 2007. Toward a signature pedagogy in educational leadership preparation and program assessment. *J. Res. Leadersh. Educ.* 2 (1), 1–29.
- Bredeson, P., 1995. Building a professional knowledge base in educational administration: opportunities and obstacles. In: Donmoyer, R., Imber, M., Scheurich, J. (Eds.), *The Knowledge Base in Educational Administration*. SUNY Press, Albany, pp. 47–73.
- Bridges, E., 1992. *Problem-Based Learning for Administrators*. ERIC Clearinghouse on Educational Management, Eugene, OR.
- Bridges, E.M., Hallinger, P., 1997. Using problem-based learning to prepare educational administrators. *Peabody J. Educ.* 72 (2), 131–146.
- Bush, T., Jackson, D., 2002. A preparation for school leadership: international perspectives. *Educ. Manag. Adm.* 30 (4), 417–430.
- CAEP., 2015. Council for the Accreditation of Education Preparation (CAEP) Program Standards. Author, Washington, DC.
- Campbell, R.F., Newell, L.J., 1973. A Study of Professors of Educational Administration. University Council for Educational Administration, Columbus, OH.
- Canole, M., Young, M.D., 2013. Standards for Educational Leaders: An Analysis. Council of Chief State School Officers, Washington, DC.
- Copland, M.A., 2003. Leadership of inquiry: building and sustaining capacity for school improvement. *Educ. Eval. Policy Anal.* 25 (4), 375–395. <https://doi.org/10.3102/01623737025004375>
- Cosner, S., Tozer, S., Zavitskovsky, P., Whalen, S., 2015. Cultivating exemplary school leadership preparation at a research intensive university. *J. Res. Leadersh. Educ.* 10 (1), 11–38.
- Crow, G., Lumby, J., Patriardis, P., 2008. *International Handbook on the Preparation and Development of School Leaders*. Taylor & Francis, New York.
- Cunningham, K.M., VanGronigen, B.A., Tucker, P.D., Young, M.D., 2019. Using powerful learning experiences to prepare school leaders. *J. Res. Leadersh. Educ.* 14 (1), 74–97. <https://doi.org/10.1177/1942775118819672>.
- Dantley, M., Tillman, L., 2006. Social justice and moral transformative leadership. In: Marshall, C., Oliva, M. (Eds.), *Leadership for Social Justice: Making Revolutions in Education*. Pearson Publishing, Boston.
- Darling-Hammond, L., LaPointe, M., Meyerson, D., Orr, M.T., Cohen, C., 2007. *Preparing School Leaders for a Changing World: Lessons From Exemplary Leadership Development Programs*, vol. 6. Stanford Educational Leadership Institute, Stanford, CA.
- Donmoyer, R., Imber, M., Scheurich, J.J., 1995. *The Knowledge Base in Educational Administration: Multiple Perspectives*. State University of New York Press, Albany, NY.
- Firestone, W., Monfils, L., Hayes, M., Polovsky, T., Martinez, C., Hicks, J., 2004. The principal, test preparation, and educational reform. In: Firestone, W., Schorr, R., Monfils, L. (Eds.), *The Ambiguity of Teaching to the Test*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Forsyth, P., 1999. Redesigning the preparation of school administrators: toward consensus. In: Thompson, S. (Ed.), *School Leadership: A Blueprint for Change*. Corwin Press, Newbury Park, CA, pp. 23–33.
- Fry, B., Bottoms, G., O'Neill, K., 2006. *Schools Can't Wait: Accelerating the Redesign of University Principal Preparation Programs*. Southern Regional Educational Board, Atlanta, GA.
- Grogan, M., Andrews, R., 2002. Defining preparation and professional development for the future. *Educ. Adm. Q.* 38 (2), 233–256.
- Hackmann, D.G., McCarthy, M.M., 2011. At a Crossroads: The Educational Leadership Professoriate in the 21st Century. Information Age, Charlotte, NC.
- Hitt, D.H., Tucker, P.D., Young, M.D., 2012. *The Professional Pipeline for Educational Leadership: A White Paper Developed to Inform the Work of the National Policy Board for Educational Administration*. University Council for Educational Administration, Charlottesville, VA.
- Jackson, B., Kelley, C., 2002. Exceptional and innovative programs in educational leadership. *Educ. Adm. Q.* 38 (2), 192–212.
- Jacobson, S., McCarthy, M., Pounder, D.G., 2015. What makes a leadership preparation program exemplary? *J. Res. Leadersh. Educ.* 10, 63–76. <https://doi.org/10.1177/1942775115569420>.
- Kandel, I.L., 1970. *Comparative Education*. Greenwood Press, Westport, CT.
- Leithwood, K., Jantzi, D., 2000. The effects of transformational leadership on organizational conditions and student engagement. *J. Educ. Adm.* 38 (2), 112–129.
- Leithwood, K., Seashore, K., 2004. *How Leadership Influences Student Learning*. Wallace Foundation, New York.
- Leithwood, K., Jantzi, D., Coffin, G., Wilson, P., 1996. Preparing school leaders: what works? *J. Sch. Leader.* 6 (3), 316–342.
- Levine, A., 2005. *Educating School Leaders*. Education Schools Project, New York.
- Marshall, C., Oliva, M. (Eds.), 2006. *Leadership for Social Justice: Making Revolutions in Education*. Pearson Publishing, Boston.
- Mezirow, J., 1991. *Fostering Critical Reflection in Adulthood: A Guide to Transformative and Emancipatory Learning*. Jossey, Bass, San Francisco.

- Murphy, J., Vriesenga, M., 2004. Research on Preparation Program in Educational Administration: An Analysis. University Council for Educational Administration, Columbia, MO.
- Murphy, J., 2003. Reculturing Educational Leadership: The ISLLC Standards Ten Years Out. Paper prepared for the National Policy Board for Educational Administration.
- National Policy Board for Educational Administration, 2015. Professional Standards for Educational Leaders, 2015. National Policy Board for Educational Administration, Reston, VA.
- National Policy Board for Educational Administration, 2018a. National Educational Leadership Preparation Standards for Building Level Leaders. National Policy Board for Educational Administration, Reston, VA.
- National Policy Board for Educational Administration, 2018b. National Educational Leadership Preparation Standards for District Level Leaders. National Policy Board for Educational Administration, Reston, VA.
- Ni, Y., Rorrer, A., Pounder, D., Young, M., Korach, S., 2019. Leadership matters: Preparation program quality and learning outcomes. *J. Educ. Adm.* 57 (2), 185–206.
- Osterman, K., Hafner, M., 2009. Research on the educational leadership preparation curriculum. In: Young, M.D., Crow, G., Ogawa, R., Murphy J. (Eds), *The handbook of research on leadership preparation*. Routledge, New York.
- Reyes-Guerra, D., Barnett, B., 2016. Clinical practice in educational leadership. In: Young, M.D., Crow, G. (Eds.), *The Handbook of Research on the Education of School Leaders*, second ed. Routledge, New York.
- Robinson, V.M., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44, 635–674.
- Schulman, L., Golde, C., Bueschel, A., Garabedian, K., 2006. Reclaiming education's doctorates: a critique and a proposal. *Educ. Res.* 35 (3), 25–32.
- Schulman, L., 2005. Signature pedagogies in the professions. *Daedalus* 134 (3), 52–59.
- Senge, P., 1990. The leader's new work: building learning organizations. *Sloan Manag. Rev.* 32 (1), 7–23.
- VanGronigen, B.A., Cunningham, K.M.W., Young, M.D., 2018. How exemplary educational leadership preparation programs hone the interpersonal-intrapersonal (i^2) skills of future leaders. *J. Transform. Leadersh. Policy Stud.* 7 (2). <https://doi.org/10.36851/jtlps.v7i2.503>.
- Young, M.D., Crow, G., 2017. *The Handbook of Research on the Education of School Leaders*, second ed. Routledge, New York.
- Young, M.D., Petersen, G.J., Short, P.M., 2002. The complexity of substantive reform: a call for interdependence among key stakeholders. *Educ. Adm. Q.* 38 (2), 137–175.
- Young, M.D., Crow, G., Murphy, J., Ogawa, R., 2009. *The Handbook of Research on the Education of School Leaders*. Routledge, New York.
- Young, M.D., Gooden, M.A., O'Doherty, A.P., 2015. Preparing leaders to support diverse learners: a curriculum development initiative by UCEA. In: Witherspoon-Arnold, N., Grant-Overton, C., Khalifa, M. (Eds.), *Handbook of Urban Educational Leadership*. Rowman & Littlefield, Lanham, MD.
- Young, M.D., Anderson, E., Nash, A., 2017. Preparing school leaders: standards-based curriculum in the United States. *J. Leadersh. Policy Sch.* 16 (2), 228–271.
- Young, M.D., Auchter, J., Shiffman, C., Richardson, I., 2018. Preparing for the National Educational Leadership Preparation (NELP) Program Review: A Companion Guide. The National Policy Board for Educational Administration, Washington, DC.
- Young, M.D., O'Doherty, A.P., Cunningham, K.M., 2021. *Leadership Preparation for Equity: Strategies for Redesign, Innovation, and Improvement*. Routledge, New York.
- Young, M.D., 1999. Multifocal educational policy research: toward a method for enhancing traditional educational policy studies. *Am. Educ. Res. J.* 36 (4), 677–714.
- Young, M.D., 2015. Effective leadership preparation: we know what it looks like and what it can do. *J. Res. Leadersh. Educ.* 10 (1), 3–10.

Pedagogical strategies for leadership development

Kenneth Leithwood, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Framework	381
Leadership skills	382
Underlying processes	382
Pedagogical strategies for enhancing leaders' skills	383
Case study	383
Problem-based learning (PBL)	383
Coaching and mentoring	384
Leadership dispositions	384
Social emotional dispositions	384
Underlying processes	384
Pedagogical strategies	385
Psychological dispositions	385
Underlying processes	385
Pedagogical strategies	386
Leadership knowledge and understanding	386
Underlying processes	387
Pedagogical strategies	387
Lectures/discussions and other forms of information presentation	387
Individual reading	387
Multi-use pedagogical strategies	388
Needs analysis	388
Digital technology	388
Cohort and network structures	389
University-district/school partnerships	389
Transfer of learning	389
Conclusion	390
References	391

This article reviews promising pedagogical strategies for the initial and ongoing development of school leaders in such widely different settings as universities, schools, districts, networks and regional development centers. While [Cosner \(2020\)](#) has recently lamented “the dearth of existing literature that attends to pedagogical approaches for educational leadership” ([Cosner, 2020](#), p. 167), an exclusive focus on the educational leadership literature for such approaches is unnecessary, especially since so much of it is descriptive and normative ([Crow and Whiteman, 2016](#); [Hallinger and Bridges, 2017](#)). This review points to an extensive and quite relevant literature about pedagogies for leadership development carried out in both school and non-school sectors. Emphasis in the review is placed on evaluative (rather than descriptive and normative) literature, especially systematic syntheses when they are available.

Framework

The objectives for leadership education are the skills (behaviors, practices), dispositions and specialized professional knowledge needed to be successful as a leader. As [Table 1](#) indicates, underlying processes explaining the acquisition of leadership skills, dispositions and knowledge are described in the paper, along with evidence about promising pedagogical strategies fostering such acquisition. Theories of expertise and how it develops are invoked to explain the achievement of skills objectives (e.g., [Chi et al., 2014](#)). Elements of socio-psychological theory are used to explain the development of several types of dispositions (e.g., [Mayer et al., 2004](#)) and social construction cognitive theory provides the perspective selected for explaining the achievement of knowledge objectives (e.g., [Schunk, 2012](#)). While some pedagogical strategies are much better suited to developing one of the three types of capacities, others can be useful in accomplishing several types of such capacities depending on how the strategies are implemented.

No matter the category of objective for leaders' learning, a large and persistent challenge faced by leadership development initiatives is to increase the likelihood that participants will apply what they learn in their own work settings ([Moldoveanu and Narayandas, 2019](#)). This “transfer of learning” challenge is addressed in the final section of the paper. Throughout the paper, examples of leadership skills, knowledge and dispositions are used to illustrate the focus of pedagogical strategies being described. These

Table 1 Summary of framework.

<i>Objectives for leader education</i>	<i>Underlying processes</i>	<i>Pedagogical strategies</i>
Skills (procedural knowledge)	Development of expertise	• Case study • Problem based learning • Coaching/Mentoring
Dispositions		
Social emotional	Skill development (as above)	• Strategies for skill development (as above)
	Self-regulation theory	• Active, concrete experiential methods
Psychological	Self-regulation theory	• Mastery experience • Vicarious experience • Verbal persuasion • Emotional arousal
Knowledge (propositional knowledge)	Social Constructivist learning theory	• Information presentation (e.g. lectures, discussion) • Individual reading • Needs analysis • Digital technology • Cohort and network structures • Partnerships
Multi-use Pedagogical strategies		
<i>Objectives for leader education</i>	<i>Underlying processes</i>	<i>Pedagogical strategies</i>
Skills (procedural knowledge)	Development of expertise	• Case study • Problem based learning • Coaching/Mentoring
Dispositions		
Social emotional	Self-regulation theory	• Strategies for skill development (as above)
		• Active, concrete experiential methods
Psychological		• Mastery experience • Vicarious experience • Verbal persuasion • Emotional arousal
Knowledge (propositional knowledge)	Social Constructivist learning theory	• Information presentation (e.g. lectures, discussion) • Individual reading • Needs analysis • Digital technology • Cohort and network structures • Partnerships
Multi-use Pedagogical strategies		

illustrations are drawn from a well-developed leadership framework (e.g., [Leithwood, 2012](#)) and several recent extensions of that framework (e.g., [Leithwood et al., 2017](#); [Leithwood, in 2021a](#)) including a version aimed at providing more specific guidance to school leaders about how to significantly improve equity in their schools ([Leithwood, 2021 b,c](#)). Although used only for illustrative purposes in this paper, adherence to a consistent framework or model of leadership has been identified as a key feature of especially effective leadership education initiatives ([Darling-Hammond et al., 2020](#)).

Leadership skills

The leadership framework from which illustrations are drawn includes a total of 22 relatively specific skills or practices organized into five domains including setting directions, developing people, designing the organization, improving the instructional program and securing accountability. One recent study ([Leithwood, Sun & Schumacker., 2020](#)) found that, after controlling for SES, this set of school leadership practices explained up to 35.8% of the variance in student achievement depending on the choice of mediator included in the analysis.

Underlying processes

The foundation of a person's performance of a skill (e.g., execution of a leadership practice) is "procedural" knowledge. Such knowledge is stored in memory as, for example, a list of things to do and a sequence about when to do them. By way of illustration, consider the leadership skill or practice "Creating high performance expectations among staff", a specific skill or practice included as part of the domain "Setting Directions". While high performance expectations do not define the substance of organizational goals

or directions, they signal the degree of effort expected to accomplish those goals, and the level of performance associated with the goals (e.g., achievement of the goals by all students not just some).

Now imagine a principal new to a school serving a high proportion of students from economically disadvantaged families struggling to make ends meet. After a few months in the school, the principal comes to appreciate that most staff members care a great deal about their students, spend considerable time offering students social support and know the family circumstances of their student very well. Most staff members are also defensive about the relatively low academic achievement of these students but believe their paramount responsibility is for their students' social and emotional well-being. The principal believes, however, that the long-term well-being of these students is closely linked to their academic success and sets out to increase staff expectations for such achievement.

Among several of the elements of the skill included for this principal in "creating high performance expectations" are helping staff understand and eventually agree on the long-term value of their students' academic success, helping staff identify instructional strategies with the potential to improve their students' academic success and providing constructive feedback to staff about how well those practices are being implemented. Each of these elements contains its own set of more specific skills and each can be enacted with more or less expertise.

So, improving a novice learners' skill (e.g., DeKeyser, 2015) requires pedagogical processes that include opportunities for the learner to acquire initial rudimentary understandings of the elements entailed in performing the skill, as well as the order of performance of those elements when order might be important (procedural knowledge). Learners then need opportunities to perform the skill, guided by those initial, rudimentary understandings of what is involved. Repeated cycles of feedback and performance gradually improve learners' understandings of what is entailed in performing the skill, as well as their performance of the skill, until such understandings and performances approximate those who are at least competent at the skill.

Moving especially from competent to expert levels of skill depends on time spent in deliberate practice (Ericsson, 2006) focused on those aspect of a skill one is weakest at performing. Deliberate practice requires concentration and practice beyond the point of being able to concentrate does not do much to advance one's expertise. Furthermore, as expertise increases over time there is a tendency for skills to become automated and consciously unavailable to the learner for further improvement. So, achieving high levels of expertise also depends on the motivation to continue achieving ever higher levels of expertise (Ryan and Deci, 2012) and the ability (or access to help) to retrieve and monitor one's performance. High levels of expertise also depend on the learners' self-regulation of their own responses to that performance (Leithwood and Steinbach, 1995), that is their ability to "systematically focus on their thoughts, feelings and action on the attainment of their goals" (Zimmerman, 2000).

Pedagogical strategies for enhancing leaders' skills

Expertise enhancement strategies aim to provide learners with positive examples of expert performance and opportunities for practice and refinement of both procedural knowledge and its enactment. There is considerable evidence about the nature and effects of three types of expertise enhancement strategies. One of these types, mentoring/coaching, enable individual practice and feedback in the learner's own context. Two of the strategies, case study and problem-based learning (PBL), can be used with groups of learners and address the significant challenge, in real school settings, of providing learners with opportunities for repeated cycles of practice and feedback. Dexter et al. (2020) argue that the uses of cases and simulations like PBL:

can serve as imagined contexts for practice onto which students can project what they are learning in courses. Rich, complex contexts allow instructors to guide experiential learning, but in a fail-safe way, and to contrive critical or even rare emergency situations ... Feedback about hypothetical contexts can be authentic and constructive, in addition to being research based, without being traumatizing (p. 176)

Case study

Evidence indicates that the structured use of case studies can be an excellent method for building skills or "procedural" knowledge (e.g., how to solve a problem). Many effective leadership programs use these types of techniques to foster their participants' deep engagement with program content of several types (Davis and Darling-Hammond, 2012). The use of case studies is a staple of most business programs and is now common in many educational leadership programs, as well. The US educational leadership community has created the *Journal of Cases in Educational Leadership* devoted to the provision of cases that can be used for such purposes with educational leaders.

Problem-based learning (PBL)

Problem-based learning is also increasingly used for leadership education and evidence suggests that its effects are at least moderately positive (Cosner et al., 2015; Dochy et al., 2003; Hallinger and Bridges, 2017). Most approaches to PBL depend on the prior development of carefully structured cases and include contextual and other forms of information key to problem solution. At their best, those structured cases closely approximate the circumstances faced by leaders engaging with them as a means of prompting transfer to participants' own work settings.

Hallinger and Bridges have provided the most comprehensive account to date about the uses of PBL in the preparation of educational leaders. This account also provides a model, including six fundamental characteristics, of problem-based learning that have been used in many of the applications of PBL for educational leadership development (Hallinger and Bridges (2017), p. 257).

1. The starting point for learning is a problem, not a theory.
2. Learning is stimulated by and organized around a high impact problem that students are apt to face in the workplace.
3. Curriculum units are organized around problems rather than subject disciplines.
4. Students assume major responsibility for managing their own learning in teams.
5. Most learning occurs in the context of student-directed small groups rather than teacher-directed lectures.
6. To the extent practical, students “demonstrate” simulated solutions to the problem (rather than just analyses) and receive formative feed-back on their efforts

Coaching and mentoring

Coaching and mentoring are typically considered overlapping strategies. Evidence suggests that some forms of individual coaching and mentoring significantly influence individual learning, especially feedback provided by coaches about leaders’ work in their schools (Bickman et al., 2012). In spite of their value, Damore and Stacy Rieckoff (2021) report few job-embedded coaching opportunities available especially for experienced US school leaders. A recent comprehensive review of cross-sector executive coaching literature (Athanasopoulou and Dopson, 2018) reported such positive outcomes for those being coached as better personal management skills and self-control, improved leadership skills, more productive interpersonal relationships as well as improved organization-level outcomes. Positive feedback about one’s leadership potential has also been associated with subsequent ambition to become a leader as well as enhanced performance (Steffens et al., 2018).

Coaches are also valuable resources for learning that takes place in the context of teams and networks (Rapp et al., 2016; Ford et al., 2013). One purpose for a coach in these contexts is to act as an advisor to network members specifically aimed at improving the quality of network interactions. A second purpose is to work with individual program or network participants on their own problems of practice within a networked arrangement (Ely et al., 2010).

Evidence reported by Hamlin and Sage (2011) indicates that mentors (and likely coaches) value mentees who are open, friendly and receptive, take responsibility for their own learning and for the success of the mentoring relationship and reveal their concerns and provide feedback on the impact the mentoring is having. Mentors also value mentees who help their mentors to understand quickly both the individual and organizational contextual factors that impact positively or negatively on the mentoring situation.

Mentees appreciate their mentors when they build an open and mutually trusting dyadic relationship that leads to friendship and convey positive regard for mentees and induce in them self-reliance, self-esteem and confirmation” of self-worth. A “coaching” style of mentoring is appreciated by mentees as are mentoring processes that take mentors outside their comfort zone so they can learn to cope effectively with more challenging work assignments.

Feedback can be provided to individual learners other than through coaching and mentoring, however. Individuals receive feedback when they compare their actual performance with their understanding (or schema) of ideal performance. This feedback is always available to the individual, although not always taken advantage of. But the value of this feedback is limited by the sophistication of their own understanding and the extent to which they are able to monitor their own performance. Feedback from a relatively expert coach or mentor overcomes both of these limitations. At its core, the coach or mentor’s feedback is based on his or her assessment of the gap between the participant’s performance and the coach or mentor’s more refined understanding of what an expert performance should look like.

Leadership dispositions

The framework from which illustrations are drawn about leadership dispositions includes two categories of such dispositions. Socio-emotional and psychological dispositions.¹ Underlying processes giving rise to these two sets of dispositions are partly different from one another, and some of those processes depend on the development of expertise (described in the previous section) as well as propositional knowledge (described in the next section).

Many dispositions, in addition to the two sets included in the illustrative framework, have been included in both leadership programs and the standards on which some are based. However, there is little consensus about either the meaning of “dispositions” or which dispositions matter most.

Social emotional dispositions

Underlying processes

Social-emotional dispositions, conceived of as “abilities” (Daus and Ashkanasy, 2005), include perceiving and managing emotions, as well as acting in emotionally appropriate ways. Reflecting the dimensions of emotional intelligence included in Mayer et al.,

¹Recent refinements to the framework include “Character” as a third category of dispositions, one encompassing leaders’ values and the use of those values in ethical decision making. Underlying processes and pedagogical strategies for developing leaders’ Character will be addressed in a subsequent paper.

(2004) influential framework, these social emotional dispositions are strongly associated with such leadership-relevant outcomes as job satisfaction, improved management and coaching skills, teamwork and conflict management ability, along with positive attitudes toward work (Little et al., 2016; Kotsou et al., 2019). Relationships such as these are especially significant for those in jobs with relatively high levels of emotional labor such as school leadership. There is broad consensus about the improvability of at least some of these abilities (Dulewicz and Higgs, 2004), although interventions are generally believed to be most effective during childhood. Underlying processes accounting for improvements are adaptations of the skill development processes leading to greater expertise.

Pedagogical strategies

Adopting skill development as the process underlying social emotional learning means that case study, PBL, coaching and feedback are suitable pedagogical strategies for the socio-emotional development of leaders (Green et al., 2011). Cherniss et al. (1998) point to elements of the skill development process especially significant for developing this set of dispositions. Pedagogical strategies should be especially clear about the goals to be achieved:

Specific and challenging goals help support social and emotional learning because they maximize self-efficacy, mastery, and motivation. (p. 17)

These strategies should also maximize practice opportunities:

In social and emotional learning, there often must be more practice than in other types of learning because old, ineffective neural connections need to be weakened and new, more effective ones established. Such a process requires repetition over a prolonged period of time. And learners need to practice on the job, not just in the training situation, for transfer to occur." (p. 17)

Especially frequent feedback on practice should be provided:

Feedback is especially useful in social and emotional learning because the learners often have trouble recognizing how their social and emotional behavior manifests itself. In fact, because self-awareness is a core competence, those who need the most help in emotional competence programs may be particularly weak in this area. (p. 17).

And experiential methods are much to be preferred in developing socio emotional skills:

More active, concrete, experiential methods, such as role plays, group discussions, and simulations, usually work better than lecturing or assigned reading for social and emotional learning. In order to re-program neural circuits connecting the amygdala and neocortex, people need to actually engage in the desired pattern of thought, feeling, and action. (p. 18)

Evidence suggests that strategies for increasing leaders' sensitivity to their own and others' emotions include coaching and feedback for recognizing the emotions of self and others from facial and voice cues, as well as awareness of one's own body states relating to emotions. Also recommended is coaching about how to adjust the overt expression of leaders' own emotions so they are appropriate to the circumstances, and how to draw on positive moods to enhance creative thought (e.g., Schutte et al., 2013).

Psychological dispositions

Underlying processes

Four psychological dispositions, included in the illustrative leadership framework, are associated especially strongly with more efficacious leadership including optimism (Kluemper et al., 2009), resilience (Foerster and Duchek, 2018) self-efficacy (Leithwood and Jantzi, 2008) and proactivity (Crant and Bateman, 2000). Underlying processes giving rise to these psychological dispositions are explained by a line of theorizing in social psychology about the self-regulation of behavior (e.g., Bandura, 1991).

Self-regulation theory explains high levels performance as a virtuous circle of self-beliefs influenced by the setting of ambitious goals, attention to feedback about progress in achieving those goals (Schier and Carver, 1992) and belief that most capacities can be learned, Dweck's "growth mindset" (2007). People who believe that achieving ambitious outcomes is possible and likely (optimism), as well as something they have or can acquire the ability to achieve (role-specific self-efficacy), continue to pursue those outcomes even in the face of adversity (resilience). Closely related psychological dispositions, optimism, self-efficacy and resilience, have their most powerful effect on a person's behaviors acting in unison to promote "proactivity". Proactivity is a motivational state

predisposing a person toward initiating future-oriented action to improve organizational functioning (Parker et al., 2006; Crant and Bateman, 2000).

Pedagogical strategies

Consistent with self-regulation theory, Bandura et al. (1999) identified four types of learning experiences on which to base leadership education strategies aimed at enhancing psychological dispositions. Although not pedagogical strategies themselves, these four types of experiences mediate the influence of pedagogical strategies on the learning of leaders. The four types of experiences are described in this section along with examples of pedagogical strategies likely to stimulate each type of experience.

Mastery experiences

Leaders' past successes and failures with a task will have strong effects on their feelings of optimism and self-efficacy about accomplishing similar future tasks. These mastery experiences might arise for example, from participation in some form of especially relevant district professional development or from opportunities to solve manageable problems in their schools, as when a principal asks a teacher leader to chair a school committee and supports their efforts to do so. Evidence reported by Rafferty and Griffin (2006) found that "mastery experiences have a stronger impact on efficacy beliefs than other factors ... such as vicarious experiences, cognitive simulations, verbal persuasion or physiological and affective state. (p. 56). Self-regulation theory also points to strong efficacy beliefs as a key point of departure for the for eventual achievement of ambitious goals.

Vicarious experiences

Visualization of successful performance may serve as vicarious experience and be prompted by observing a skilled model (e.g., another leader working in a similar context) or by hearing about how other leaders have successfully accomplished relevant tasks in their own contexts. Self-regulation theory endorses exposure of learners to models who demonstrate the importance of resilience in accomplishing ambitious outcomes. Well-functioning peer networks offer considerable potential for vicarious experiences such as these, as well as opportunities for the information-rich feedback so essential for skill development (Leithwood, 2018).

Verbal persuasion

This source of stimulation includes encouragement and feedback from a credible, trusted, and respected colleague (perhaps a coach or mentor). Self-regulation theory points to verbal persuasion as among the most obvious and promising ways of persuading learners that they are able to learn novel (to them) and complex leadership skills – a "growth mindset", to use Dweck (2007) term. Transformational models of leadership encompass the provision of such persuasion in a practice labeled "individualized support" (Rafferty and Griffin, 2006). Evidence from non-education sectors also associate individualized support with increases in self-efficacy and employee optimism, as well (Rich, 1999).

Emotional arousal

This is a form of enhanced motivation that may be triggered by an inspirational other, a charismatic senior leader for example, who helps elevate other learners' standards and aspirations and helps them see the relationship between their organization's goals and larger social and moral purposes. A compelling speech about, or model of, those elevated standards might provide such inspiration. Some evidence (Rodríguez-Sánchez et al., 2019) indicates that clarifying an organization's commitment to social responsibility makes significant contributions to employee resilience. Elevated aspirations and more fundamental purposes are compelling reasons for "staying the course".

Leadership knowledge and understanding

The framework from which illustrations are drawn does not aspire to capture all that a leader needs to know and understand. Rather, the focus is primarily on the "mediators" of leadership influence on students, that is, those in-school and out-of-school conditions that are known to have relatively direct positive effects on students *and* that are amenable to intervention by school leaders (Sun and Leithwood, 2017). Knowledge about such mediating conditions allows leaders to focus their improvement efforts on diagnosing and enhancing the status of those conditions that matter most to student success. Acting on this knowledge is the practical consequence of understanding that the majority of a school leader's influence on students is indirect (e.g., Sebastian and Allensworth, 2012).

The illustrative framework identifies four categories of mediating conditions, a dozen in total, all justified by considerable previous research (Leithwood, Sun & Pollock). One recent study (Leithwood et al., 2020) found especially strong influences on student achievement of one of the four categories of mediators ("rational") including academic press, disciplinary climate and teachers' uses of instructional time.

Underlying processes

An earlier section of this paper described “procedural” knowledge (knowledge about how to do something) as the foundation of skill development. Leaders’ understanding, on the other hand, depends on a type of knowledge often referred to as “propositional” or “declarative”; this is knowledge about something² – it is factual and conceptual in nature.

Social construction theory, especially dialectical approaches to it, is used here to explain how propositional knowledge is developed. While rote learning will sometimes be sufficient to acquire new facts, acquiring more complex forms of new propositional knowledge depends on learners making connections between what they already know and the new ideas or concepts to be learned (Bransford et al., 2004). This process entails learners accommodating what they already know by adapting or changing it in a way that makes sense to them in the context of the settings in which they find themselves (Schunk, 2012). A personal level of understanding of new concepts or ideas is acquired through such processes: this is a social constructivist interpretation of “true beliefs”.

Piagetian notions about intellectual development suggest that, in the face of potentially new ideas, the mind’s default response is to attempt to assimilate them into one’s existing understandings avoiding the hard work of restructuring or adding to one’s existing understandings (Block, 1982). The hard work of accommodating new knowledge is prompted by challenges to existing ways of understanding, clear definitions and explanations, as well as by experiences that distinguish new concepts and ideas from already understood concepts and ideas.

Pedagogical strategies

Social construction learning theory itself *rules out* very few specific forms of pedagogy. The main focus of the theory is not on pedagogical strategies but on learners’ responses to those strategies. Any strategy that results in learners’ actively processing concepts and ideas and actively working to make sense of them will do. Two routinely but mistakenly criticized strategies exemplify this key point.

Lectures/discussions and other forms of information presentation

Information-based strategies such as lectures, lectures combined with discussion and a wide array of digital presentation modes are often considered passive forms of information transmission ineffective in promoting learners’ active processing (e.g., Dexter et al., 2020). Meta-analytic reviews of evidence in the non-education sector, however, indicate that lectures or lectures followed by discussion make substantial contributions to participants’ propositional learning (Lacerenza et al., 2017; Collins and Holton, 2004). The bad press about lectures, for example, is really bad press about bad lectures. As most of us know through our own experience, a lecture from an engaging person about new ideas that, for example, challenge our existing understandings or offer new perspectives on familiar issues can prompt a great deal of sense-making on our parts. We have all witnessed very engaging TED talks, for example. The “good” lecture is actually a very efficient way of instilling propositional knowledge as well as encouraging transfer of knowledge use to other settings.

Individual reading

Some quite recent evidence has indicated that many school leaders rank their personal reading as the most valuable source of their own professional knowledge (Leithwood, 2018). Individual reading also offers the potential to infuse new knowledge into network interactions as a means of stimulating collective knowledge creation. Leadership education initiatives can capitalize on the value of their participants’ individual reading by providing them with a carefully curated set of readings designed for two purposes. One purpose of these readings is to present basic information related to the content of a leadership education initiative. Readings may also serve as points of reference for helping participants to become more sophisticated and habitual consumers of research, an important long-term goal that an effective leadership education initiative might aim to accomplish.

Program readings can fall short of their potential for several reasons. Johnson (2014) provides a list of ways this can happen, as well as how to ensure that they do not. Many participants are reluctant to read original research reports, often the most authentic sources of relevant ideas (Darling-Hammond et al., 2010) because of difficulties in understanding complex data analysis techniques. Ensuring participants that all readings of this sort are based on credible techniques that they need not spend time on is an easy way to overcome some of this reluctance (admittedly this minimizes the development of some of the skills ideally needed to be a sophisticated long-term consumer of research). Avoiding readings that provide very small returns of ideas or stimulation for the time it costs to read them is another method for encouraging personal reading of individual program participants. Spending significant time in networks and other program forums discussing the assigned readings helps to signify and unpack their value to participants.³

²Philosophers such as Moser (1987), for example, argue that the conditions for claiming to have propositional knowledge about something include “justified true belief” and spend a lot of their time arguing about whether those conditions are sufficient. “Propositional” knowledge, as it is used in this paper, does not depend on wading into these arguments. Suffice to say that the propositional knowledge included in the illustrative framework includes an understanding of the concepts representing the mediators of leadership influence. Robust empirical research about each of the mediators justifies the argument that they deserve the attention of leaders’ school improvement efforts. The leadership education challenge is to help leaders understand the concepts and the empirical justification of claims that they matter to students.

³Some of the claims in this paragraph reflect my own experiences delivering leadership development programs over many years; it is not based on other more formal evidence unless cited.

Multi-use pedagogical strategies

The description, to this point, about how to improve leaders' skills, dispositions and knowledge indicates considerable overlap among categories of pedagogies useful for improving a wide range of leaders' capacities. For such pedagogies to be so broadly useful, however, those responsible for leadership education need to be highly intentional in the application of the pedagogies. In addition to such overlap, some pedagogical strategies can be useful no matter the objectives or content to be learned. Four approaches not yet touched on are described in this section of the paper.

Needs analysis

Because much of the leadership development research in the education sector is concerned with the impact of whole programs, it has provided very little evidence about the value of needs analysis, at least at the individual level. Reviews of non-education sector research (e.g., [Collins and Holton, 2004](#)), however, suggest that needs analysis is a useful tool for personalizing leadership development; avoiding wasting participants' time on skills and knowledge they either do not need or already possess. Absent such an analysis, programs are more likely to be generic in nature and not suitable for at least a portion of individuals and their organizations. Needs analysis, however, needs to guard against relying solely on participants' existing understandings of what they should learn ([Ford et al., 2013](#)). There is an important role for some form of coaching in this process.

Digital technology

Digital technology now permits the ready presentation of new information online, along with interactions among participants about the meaning and use of that information. One recent study found that 73% of the leadership preparation programs sampled by the study used some form of distance learning ([Crow and Whiteman, 2016](#)). Webinars have become a core strategy of some organizations (likely most organizations during the pandemic) for introducing program content to participants and for engaging participants in sense making about such content. Online strategies such as these are replacing the heavy use of seminar and lecture formats traditionally used for the presentation of new ideas, although a small number of face-to-face sessions are still likely to be useful. While face-to-face experiences provide immediate flexibility in response to participants' concerns, they can also be a forum in which the voices and views of some members dominate those of others. Online provision is noted for leveling the opportunities to be heard by all members ([Barakat et al., 2019](#)) but for this to happen, discussions need to be carefully moderated.

Evidence and theory from research on e-leadership, especially the leadership of virtual teams, has some implications for the use of digital technologies as tools for leadership education. Two lines of theory seem especially useful to consider. One line of theory, media synchronicity theory, distinguishes synchronous from asynchronous forms of digital communication. Face-to-face, phone conversation and videoconferencing are examples of synchronous communication, while text and email are asynchronous examples. The flow of conversation is controlled by the degree to which a digital medium allows for synchronous communication, as in the case in face-to-face communication.

A second line of theory, media richness theory, is explained as follows:

A rich medium allows for: (1) transmitting multiple verbal and nonverbal cues, (2) using natural language, (3) providing immediate feedback, and (4) conveying personal feelings and emotions ... The richest medium is [face-to-face] communication, followed by telephone, chat, e-mail, and print communications. Newer technologies, such as videoconference, are thought to rank above the telephone, but below FTF communication in terms of media richness

Hambley et al., 2007, p. 3.

While some comparative evidence indicates that participants rate the learning that occurs through online and face-to-face program experiences about the same (e.g., [Ritter et al., 2010](#)), evidence and theory about e-leadership suggest more nuanced leadership education applications; that is, as the complexity of what is to be learned increases, so should the synchronicity and richness of the media used.

Writing about corporate sector leadership development, [Moldoveanu and Narayandas \(2019\)](#) argue for corporations to conceptualize the huge range of forms and providers available to them as a *Professional Learning Cloud* (PLC). As technology becomes increasingly used by providers and increasingly sophisticated (e.g., media rich and synchronous) in how it engages learners, both corporations and individual learners will be able construct professional learning experiences that are increasingly personalized, social and contextual in nature. These experiences will also potentially be much more cost effective to the extent that the content of many traditional programs becomes accessible in a piecemeal fashion. The potential for Artificial Intelligence, now being actively applied to assist learning for K-12 students ([CoSN and Microsoft, 2020](#)) will, no doubt, begin to emerge as a more prominent tool for leadership development in the near future.

Corporate sector professional learning opportunities would seem to be significantly greater and more varied than are learning opportunities for educational leaders. But developing a *Professional Learning Cloud* for educational leadership development would be a promising initiative.

Cohort and network structures

Many highly regarded leadership programs promote collegial communities or networks that have a sustained focus and clear expectations for improved student engagement and learning outcomes. The motivation of members in collegial communities to engage actively in their own learning is typically very high. Many effective leadership development programs now create collegial communities by developing peer networks among participants sometimes by organizing participants in cohorts which continue together throughout their program experiences (e.g., Ford et al., 2013).

Whether through cohort organization or in some other manner, effective peer networks can vary in size and members may interact both in person and online. Principals included in two recent studies (Leithwood and Azah, 2016; Leithwood, 2018) ranked participation in their district-developed peer leadership networks as second only to their own professional reading as a source of their professional learning among a dozen possible sources.

A well-functioning network builds on strong interpersonal relationships (e.g., Gallego et al., 2001) and contributes to the capacity of individual members by exposing them to the practices, dispositions and ideas of others faced with similar tasks and responsibilities (e.g., Lieberman and Grolnick, 1996). This is the diffusion goal of networks, one that exploits knowledge already developed by ensuring that such knowledge goes to scale. The diffusion process often also encourages network members to make explicit, for themselves and their colleagues, the large amounts of tacit knowledge (Polanyi, 1983) on which a large proportion of their expertise depends.

A network is also a structure which, under the right conditions, stimulates potentially rich interactions among members resulting in new and creative ideas or practices not initially part of the repertoire of any individual network member (e.g., Lieberman and Grolnick, 1996). The whole becomes more than the sum of its parts. This is the knowledge creation goal of networks. Knowledge creation depends on meeting at least four closely related conditions:

- willingness on the part of participants to collaborate in the solving of some shared problem or the meeting of a shared purpose
- willingness, during collaborative interactions, to genuinely listen to the ideas of one's network colleagues. Such listening entails a conscious change in mindset, from one that is primarily about assimilation of new ideas to one that is focused on the adaptation of one's existing understandings
- emergence of solutions or ideas that are not simply personally creative (individual learning) but are creative or novel and useful to the collaborating groups, as a whole
- infusion of relevant ideas located in sources outside the network.

University-district/school partnerships

Anderson et al.'s (2018) study of leadership programs affiliated with the *University Council for Educational Administration* (UCEA) found that among the 97 sampled programs, a very large proportion were actively engaged in formalized partnerships with districts. Examples of the advantages ascribed to such university-district/school partnerships include bridging theory and practice, creating more options for delivering programs, emphasizing collaborative leadership, aligning district and program goals, and sharing of program costs. Some evidence also points to improved program quality and increased opportunities for districts to develop leaders with the capacities needed to help districts achieve their strategic priorities. University-district/school partnerships, however, bring with them their own challenges. Perhaps the largest of those challenges is obtaining and maintaining sustained participation and commitment from all stakeholders, often a challenge in the face of rapid district leadership mobility and succession (Crow and Whiteman, 2016; Darling-Hammond et al., 2010).

A recent, large-scale UK study (Day et al., 2021) unpacked how university-school partnerships may contribute to fundamental and sustained large-scale educational reform and the conditions needed for those contributions to be realized. Insights from this study also include how and under what conditions such partnerships are likely to contribute to leadership development. Partnerships offer most of the advantages to leaders' learning described in the previous account of peer network participation. However, as compared with peer networks for example, university-school partnerships are likely to encompass much greater diversity of knowledge, culture, perspectives on challenges for improvement and resources. As the authors claim:

The joining of resources, knowledge, experiences and expertise creates a larger entity that can potentially expand the capacity of each individual partner and offer them the promise of deep change and improvement through trustful collaboration. (p. 127)

Learning through participating in university-school partnerships would seem to depend on leaders: treating the diversities to be found in partnerships as opportunities for improving their own capacities; welcoming the opportunities for collaboration with different "others"; contributing to the development of trusting, productive working relationships among members of the partnership and: readily engaging with others in the partnership to build new knowledge.

Transfer of learning

Transfer of learning occurs when learning in one context has a positive or negative influence on a person's performance in another context. The transfer process is generally considered to be challenging and unlikely to occur unless specific strategies are used to

encourage it. Failure to transfer learning from a leadership education experience to a learner's own work setting is a major threat to the impact of leadership development initiatives (Moldoveanu and Narayandas, 2019). The often-significant investments in those initiatives can easily result in little or no return. For this reason, no matter the category of objective being addressed by a leadership education initiative, including strategies to help ensure transfer of learning to participants' own work settings should be considered a basic requirement.

The seminal work of Perkins and Salomon initiated three decades ago (Perkins and Salomon, 1992; Salomon and Perkins, 1989; Perkins and Salomon, 1989) helps clarify the transfer of learning process and identify basic strategies for helping to ensure it. A key distinction introduced by these researchers is between near or "low road" transfer and far or "high road" transfer. An example of low road transfer would be a school leader applying a goal-setting procedure she developed in one school to a new school to which she had been transferred. The basic procedure might remain the same but the explanation of it to the different group of teachers in the new school might have to be modified. This low road transfer process is reflexive, that is, it involves the triggering of well-practiced routines under conditions that are similar to those in the original learning context.

The best case for leaders to learn from to ensure low road transfer will often be their own school or district, the challenges it faces and the many features that make it different from other schools or districts. Using the example of one leadership development program in Texas (Rice University's Educational Entrepreneurial program), immersing participants in their own cases entailed program providers collecting evidence in the early stages of participants' engagement in the program about key aspects of each participant's school. This evidence, summarized into readily useable form by program staff, was fed back to each participant in a report used as a major resource for school improvement work. Knowledge and skills about leadership for school improvement were important elements of the program's content. Participants' guided efforts to improve their own schools were a central strategy used by this program to foster development of such knowledge and skills (also see Salisbury and Irby, 2020).

High road transfer entails the application of learning that occurred in one context to inform performance in a very different context. An example of high road transfer would be an approach to stimulating staff innovation described in a case study of a high-tech company included in a training program attended by a superintendent of education. For the superintendent to apply this case to encourage innovation among school leaders in his or her district the superintendent would need to work quite deliberately at understanding the fundamental ideas used by the high-tech company to prompt more staff innovation and to have sorted out the differences in contexts that would need to be addressed for successful application in his or her own district. Perkins and Solomon refer to this process as "deliberate effortful abstraction and a search for connections" (Perkins and Salomon, 1992; p. 8). Examples of strategies for improving the chances of successful high road transfer include:

- explicit abstractions of principles from the original learning situation
- metacognitive reflection on the part of learners to their own thinking
- use of metaphors and analogies during the original learning experience.

While evidence indicates a much greater likelihood of low road, as compared with high road transfer, the value of successful high road transfer is its flexible use across multiple contexts. Low road transfer of learning, in contrast, is limited to applications in very similar circumstances. Successful low road transfer cannot be taken for granted, however. It requires the learning, which is to be transferred, to be well automated, something typically requiring significant practice. It also requires the context for transfer to be very similar to the context of the original learning.

Conclusion

The article has provided evidence to justify three basic claims central to the goals for leadership development and their achievement. First, almost all leadership capacities can be conceptualized as either skills, dispositions, or knowledge and understanding. Second, development of each type of capacity can be explained by at least partly distinctly bodies of theory, for example, theories of expertise in the case of skills, self-regulation theory in the case of many dispositions and social constructivism in the case of knowledge and understanding. Each of these bodies of theory, thirdly, provide useful guidance when selecting and implementing pedagogical strategies for use in leadership education initiatives. Some pedagogical strategies have the potential for helping achieve multiple types of leadership capacities depending on *how* they are used. Leadership educators are advised to be highly intentional about the capacities to be developed when implementing most of the more than dozen strategies described in the paper and not to rely on just one strategy (Lacerenza et al., 2017).

Some common understandings of the term "active learning strategies" (e.g., Dexter, et al., 2020; Salisbury and Irby, 2020), used in more recent leadership education programs, have been challenged in the article. In descriptions of these programs, the term "active" does seem to describe the physical and social circumstances of participants; for example, they are working on problems with others and engaged in networking activity. But active *instructional* strategies would be a more accurate label for them. Whether or not participants are actively *learning* depends entirely on whether they are cognitively and/or emotionally engaged in making sense of their experiences. While such engagement does seem likely with many of these active instructional strategies, participants can be engaged in making sense of their experiences by well implemented learning strategies (a lecture, for example) often viewed as passive or simply designed to transmit information.

The article also acknowledges the waste of resources that occur when leaders acquire new skills, dispositions, and knowledge through participation in a learning initiative but subsequently fail to use those new capacities in their actual work. Transfer of

learning theory has been invoked as a source of guidance for leader educators about how to design their education initiatives to better ensure use of new learning in practice. The important distinction in transfer of learning theory between “high” and “low-road” transfer also helps to both explain and challenge the types of educational initiatives preferred by many practicing leaders. These preferred initiatives typically stress low-road transfer because they offer relatively immediate “practical” solutions to current problems and require relatively small amounts of cognitive accommodation on the part of learners. But many of the challenges facing today’s educational leaders in the future will be predictably unpredictable, guaranteed only to be significantly different than the challenges they currently face. Preparing school leaders for such unpredictable futures is the job of “high road” transfer – arming today’s leaders with powerful theories, principles and concepts that can be used as tools for addressing many different future circumstances should be an important feature of educational leader education today.

References

- Anderson, E., Winn, K.M., Young, M.D., Groth, C., Korach, S., Pounder, D., Rorrer, A.K., 2018. Examining university leadership preparation: an analysis of program attributes and practices. *J. Res. Leadersh. Educ.* 13 (4), 375–397.
- Athanasopoulou, A., Dopson, S., 2018. A systematic review of executive coaching outcomes: is it the journey or the destination that matters the most? *Leader. Q.* 29 (1), 70–88.
- Bandura, A., 1991. Social cognitive theory of self-regulation. *Organ. Behav. Hum. Decis. Process.* 50, 248–287.
- Bandura, A., Freeman, W.H., Lightsey, R., 1999. Self-efficacy: The Exercise of Control, pp. 158–166.
- Barakat, M., Reames, E., Kensler, L., 2019. Leadership preparation programs: preparing culturally competent educational leaders. *J. Res. Leadersh. Educ.* 14 (3), 212–235.
- Bickman, L., Goldring, E., De Andrade, A., Breda, C., Goff, P., 2012. Improving principal leadership through feedback and coaching. *Rep. Res. ED530123*.
- Block, J., 1982. Assimilation, accommodation and the dynamics of personality development. *Child Dev.* 53, 281–295.
- Bransford, J.D., Brown, A.L., Cocking, R.R., Donovan, M.S., Pellegrino, J.W., 2004. *How People Learn*. National Academy Press, Washington, D.C.
- Cherniss, C., Goleman, D., Emmerling, R., Cowan, K., Adler, M., 1998. *Bringing Emotional Intelligence to the Workplace*, A Technical Report Issued by the Consortium for Research on Emotional Intelligence in Organizations. Rutgers University, New Brunswick, NJ.
- Chi, M.T., Glaser, R., Farr, M.J. (Eds.), 2014. *The Nature of Expertise*. Psychology Press, New York.
- Collins, D., Holton, E., 2004. The effectiveness of managerial leadership development programs: a meta-analysis of studies from 1982 to 2001. *Hum. Resour. Dev. Q.* 15 (2), 217–248.
- CoSN & Microsoft, 2020. Artificial Intelligence in K-12. Prepared for. CoSN in partnership with Microsoft. (Spring).
- Cosner, S., 2020. A deeper look into next generation active learning designs for educational leader preparation. *J. Res. Leadersh. Educ.* 15 (3), 167–172.
- Crant, J.M., Bateman, T.S., 2000. Charismatic leadership viewed from above: the impact of proactive personality. *J. Organ. Behav.* 21 (1), 63–75.
- Crow, G., Whiteman, R., 2016. Effective preparation program features: a literature review. *J. Res. Leadersh. Educ.* 11 (1), 120–148.
- Damore, S., Stacy Rieckoff, B., 2021. School leader perceptions: coaching tools and processes. *J. Res. Leadersh. Educ.* 16 (1), 57–80.
- Darling-Hammond, L., Meyerson, D., LaPointe, M., Orr, M.T., 2010. *Preparing Principals for a Changing World: Lessons from Effective School Leadership Programs*. Jossey-Bass, San Francisco, CA.
- Darling-Hammond, L., Wechsler, M., Levin, S., Campoli, A., Leung, M., Tozer, S., 2020. *Developing Principals: What we Know, Why it Matters, and What can be Done*. The Wallace Foundation (Draft), New York.
- Daus, C., Ashkanasy, N., 2005. The case for the ability-based model of emotional intelligence in organizational behavior. *J. Organ. Behav.* 26, 453–466.
- Davis, S., Darling-Hammond, L., 2012. Innovative principal preparation programs: what works and how we know. *Plan. Changing* 43 (1–2), 25–45.
- Day, C., Gu, Q., Townsend, A., Holdich, C., 2021. *School-University Partnerships in Action*. Routledge, London.
- DeKeyser, R., 2015. Skill acquisition theory. In: Bill VanPatten, B., Williams, J. (Eds.), *Theories in Second Language Acquisition*, second ed. Routledge, New York, pp. 94–112.
- Dexter, S., Clement, G., Moraguez, D., Watson, G., 2020. (Inter) active learning tools and pedagogical strategies in educational leadership preparation. *J. Res. Leadersh. Educ.* 15 (3), 173–191.
- Dochy, F., Segers, M., Van den Bossche, P., Gijbels, D., 2003. Effects of problem-based learning: a meta-analysis. *Learn. Instruct.* 13 (5), 533–568.
- Dulewicz, V., Higgs, M., 2004. Can emotional intelligence be developed? *Int. J. Hum. Resour. Manag.* 15 (1), 95–111.
- Dweck, C., 2007. *Mindset: The New Psychology of Success*. Random House, New York.
- Ely, B., Boyce, L., Nelson, J., Zaccaro, J., Hernez-Broome, G., Whyman, W., 2010. Evaluating leadership coaching: a review and integrated framework. *Leader. Q.* 21, 585–599.
- Ericsson, A., 2006. The influence on expertise and deliberate practice on the development of superior expert performance. In: Anders Ericsson, A. (Ed.), *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*. Cambridge University Press, Cambridge, pp. 637–706.
- Foerster, C., Ducheck, S., 2018. Leaders’ resilience-A systematic literature review and future research agenda. *Acad. Manag. Proc.* 2018 (1), 13879. Briarcliff Manor, NY 10510: Academy of Management.
- Ford, C., McMahon, M., Gronn, P., 2013. Designing individualised leadership development programmes. *Sch. Leader. Manag.* 33 (5), 440–456.
- Gallego, S., Hollingsworth, S., Whitenack, D., 2001. Relational knowing in the reform of educational cultures. *Teach. Coll. Rec.* 103 (2), 240–266.
- Green, J., Chirichello, M., Mallory, B., Melton, T., Lindahl, R., 2011. Assessing leadership dispositions: issues, challenges, and promising practices. *Inte. J. Edu. Leadersh. Prep.* 6 (4), 1–10.
- Hallinger, P., Bridges, E., 2017. A systematic review of research on the use of problem-based learning in the preparation and development of school leaders. *Educ. Adm. Q.* 53 (2), 255–288.
- Hamlin, R., Sage, L., 2011. Behavioral criteria of perceived mentoring effectiveness. *J. Eur. Ind. Train.* 35 (8), 752–778.
- Johnson, E., 2014. Expertise and decision under uncertainty: performance and process. In: Chi, M., Glaser, R., Farr, M. (Eds.), *The Nature of Expertise*. Psychology Press, New York, pp. 209–260.
- Kluemper, D., Little, L., DeGroot, T., 2009. State or trait: effects of state optimism on job-related outcomes. *J. Organ. Behav.* 30, 209–231.
- Kotsou, I., Mikolajcak, M., Heeren, A., Gregoire, J., 2019. Improving emotional intelligence: a systematic review of existing work and future challenges. *Emotion Rev.* 11 (2), 151–165.
- Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L., Salas, E., 2017. Leadership training design, delivery, and implementation: a meta-analysis. *J. Appl. Psychol.* 102 (12), 1686–1718.
- Leiberman, A., Grolnick, M., 1996. Networks and reform in American education. *Teach. Coll. Rec.* 98 (1), 7–45.
- Leithwood, K., 2012. *The Ontario Leadership Framework with a Discussion of Research Foundations*. Institute for Educational Leadership and the Ontario Ministry of Education, Toronto.
- Leithwood, K., 2018. Characteristics of effective leadership networks: a replication and extension. *Sch. Leader. Manag.* 39 (2), 145–17.
- Leithwood, K., Azah, V., 2016. Characteristics of effective leadership networks. *J. Educ. Adm.* 54 (4), 409–433.
- Leithwood, K., Jantzi, D., 2008. Linking leadership to student learning: the contributions of leader efficacy. *Educ. Adm. Q.* 44 (4), 496–528.

- Leithwood, K., Steinbach, R., 1995. *Expert Problem-Solving Processes: Evidence from Principals and Superintendents*. State University of New York Press, Albany, NY.
- Leithwood, K. (2012). *The Ontario Leadership Framework with a discussion of its Research Foundations*. Toronto: Institute for Educational Leadership and Ontario Ministry of Education 6. p0180:
- Leithwood, K., Sun, J., Pollock, K. (Eds.), 2017. *How School Leaders Contribute to Student Success: The Four Paths Model*. Springer International Publishers, New York.
- Leithwood, K., Sun, J.-P., Schumacker, R., 2020. How school leadership influences student learning: a test of "the four paths model". *Educ. Adm. Q.* 56 (4), 570–599.
- Leithwood, K., 2021a. Integrated educational leadership. In: Tierney, R., Rizvi, F., Ericikan, K., Smith, G. (Eds.), *Encyclopedia of Education*, fourth ed. Elsevier Publishers, Amsterdam, NL.
- Leithwood, K., 2021b. A review of evidence about equitable school leadership. *Educ. Sci.* 11, 377 doi:10.3390.
- Leithwood, K., 2021c. Building productive relationships with families and communities: a priority for leaders to improve equity in their schools. In: Jaynes, W. (Ed.), *Parental Involvement: Cultivating Relationships*. Routledge, New York.
- Little, L., Gooty, J., Williams, M., 2016. The role of leader emotion management in leader–member exchange and follower outcomes. *Leader. Q.* 27 (1), 85–97.
- Mayer, J.D., Salovey, P., Caruso, D.R., 2004. Emotional intelligence: theory, findings, and implications. *Psychol. Inq.* 15 (3), 197–215.
- Moldoveanu, M., Narayandas, D., 2019. Gaps in traditional executive education are creating room for approaches that are more tailored and democratic. *Harv. Bus. Rev.* (March–April).
- Moser, P.K., 1987. Propositional knowledge. *Phil. Stud.* 52 (1), 91–114.
- Parker, S.K., Williams, H.M., Turner, N., 2006. Modeling the antecedents of proactive behavior at work. *J. Appl. Psychol.* 91 (3), 636–652.
- Perkins, D.N., Salomon, G., 1989. Are cognitive skills context bound? *Educ. Res.* 18 (1), 16–25.
- Perkins, D., Salomon, G., 1992. Transfer of learning. In: *International Encyclopedia of Education*, second ed. Pergamon Press, Oxford, England.
- Polanyi, M., 1983. *The Tacit Dimension*. Peter Smith, Gloucester, MA.
- Rafferty, A.E., Griffin, M.A., 2006. Refining individualized consideration: distinguishing developmental leadership and supportive leadership. *J. Occup. Organ. Psychol.* 79 (1), 37–61.
- Rich, G.A., 1999. Salesperson optimism: can sales managers enhance it and so what if they do? *J. Market. Theor. Pract.* 7 (1), 5–63.
- Ritter, C., Polnick, B., Fink II, R., Oescher, J., 2010. Classroom learning communities in educational leadership: a comparison study of three delivery options. *Internet High Educ.* 13 (1–2), 96–100.
- Rodríguez-Sánchez, A., Guinot, J., Chiva, R., López-Cabral, Á., 2019. How to emerge stronger: antecedents and consequences of organizational resilience. *J. Manag. Organ.* 25 (1), 1–18.
- Ryan, R., Deci, E., 2012. An overview of self-determination theory: an organismic dialectical perspective. In: Deci, E., Ryan, R. (Eds.), *Handbook of Self-Determination Research*. University of Rochester Press, Rochester, pp. 3–33.
- Salisbury, J., Irby, D., 2020. Leveraging active learning pedagogy in a scaffolded approach: reconceptualizing instructional leadership learning. *J. Res. Leadersh. Edu.* 15 (3), 210–226.
- Salomon, G., Perkins, D., 1989. Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Educ. Psychol.* 24 (2), 113–142.
- Schier, M., Carver, C., 1992. Effects of optimism on psychological and physical well-being-theoretical review and update. *Cognit. Ther. Res.* 16 (2), 201–228.
- Schunk, D., 2012. *Learning Theories: An Educational Perspective*, sixth ed. Allyn & Bacon Publishers, Boston.
- Schutte, N.S., Malouff, J.M., Thorsteinsson, E.B., 2013. Increasing emotional intelligence through training: current status and future directions. *Int. J. Emotion. Edu.* 5 (1), 56–72.
- Sebastian, J., Allensworth, E., 2012. The influence of principal leadership on classroom instruction and student learning: a study of mediated pathways to learning. *Educ. Adm. Q.* 48 (4), 626–663.
- Steffens, N.K., Fonseca, M.A., Ryan, M.K., Rink, F.A., Stoker, J.L., Pieterse, A.N., 2018. How feedback about leadership potential impacts ambition, organizational commitment, and performance. *Leader. Q.* 29 (6), 637–647.
- Sun, J.-P., Leithwood, K., 2017. Calculating the power of alternative choices by school leaders for improving student achievement. *Sch. Leader. Manag.* 37 (1–2), 80–93.
- Zimmerman, B.J., 2000. Attaining self-regulation: a social cognitive perspective. In: Boekaerts, M., Pintrich, P.R., Zeidner, M. (Eds.), *Handbook of Self-Regulation*. Academic Press, San Diego, pp. 13–39.

Understanding what makes school leaders effective: the importance of personal development and support^a

Peter Earley and Sara Bubb, UCL Institute of Education, University College London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	393
Leadership development and leader development	393
The challenges of headship	394
Developing leaders	396
Opportunities to thrive, survive and revive	396
Conclusion	398
References	399

Introduction

International research (e.g., [Huber, 2010](#); [Day and Sammons, 2013](#); [Day and Gurr, 2016](#)) consistently shows that highly successful principals and headteachers demonstrate a number of personal and professional characteristics. They are optimistic, enthusiastic and curious, committed to social justice, equity and excellence, and they show respect and empathy for others. Self-aware, resilient and persistent, they show drive and determination, enthusiasm, courage, conviction and integrity in pursuing excellence and putting students first. Only the best will do for such professionals. Similarly, [Lovett et al., \(2015, p. 138\)](#) note that there is general agreement that “school leaders need high levels of self-efficacy, resilience, self-awareness and judgment in order to cope with the emotional demands and complexities inherent in school-based ethical decision-making”.

Research evidence also shows that given the intensity of the job, the emotional demands, the accompanying workload and the high-stakes accountability cultures in which they work, principals suffer from high levels of occupational stress ([Bakker, 2015](#); [OECD, 2013](#); [Phillips and Sen, 2011](#)). It is perhaps unsurprising therefore that many education systems are experiencing difficulties in recruiting and retaining school principals ([Cheng and Szeto, 2016](#); [Fink, 2010](#); [MacBeath, 2009](#); [MacBeath et al., 2012](#); [OECD, 2009](#); [Pollock, 2016](#); [Thomson, 2009](#)). In England, “The School Leadership Challenge” ([Future Leaders Trust, 2016](#)) provided an analysis of the supply and demand for school leaders until 2022. It noted that schools in England face real challenges in attracting, developing and retaining leaders. The research found 40% of school governing bodies, who are responsible for the appointment of heads, face challenges in finding good candidates for senior roles.

Within this context of recruitment and retention challenges, growing emotional demands, work intensification, stress and overload, this paper considers the personal development needs of school leaders. It argues that while the provision of leadership development opportunities continues to be essential, greater attention needs to be given to individuals and their personal development - their wellbeing and developing the personal qualities and traits required to operate in such demanding working conditions. For as [Mahfouz \(2020, p. 442\)](#) notes, “when school administrators are emotionally exhausted, overworked and fail to prioritize their own wellbeing, they typically become frustrated, and their performance is hindered”.

Consideration is also given to the forms of support needed to ensure leaders’ continuing effective performance, all the more important while working within a high stakes accountability culture. In other words what is needed for headteachers to “survive”, “thrive” and also “revive” (or rejuvenate) in the current educational landscape? Drawing upon evidence from a variety of sources, it is suggested that there is a need for leadership development programmes to give as much if not more prominence to personal development over professional development - what [Waters \(1998a\)](#) calls the “pro-personal” and [Leithwood \(2012\)](#) terms personal leadership resources. Also a number of suggestions are made which will help make the work of principals - both the work itself and the way it is approached - more sustainable and to ensure they continue to operate effectively. If implemented, these suggestions would help principals not only to survive but thrive and/or revive, thus bringing many benefits to the individuals themselves, the schools they lead and the students within them.

Leadership development and leader development

Leadership development refers to the activities involved in strengthening a leader’s ability to establish clear vision and achievable goals, and to motivate others to subscribe to the same vision and goals. Leader development, however, focuses more on the individual school leader and their personal needs, what is sometimes referred to as personal effectiveness or personal efficacy. Broadly

^aA version of this article was published in 2020 as “Surviving, thriving and reviving in leadership: the personal and professional development needs of educational leaders”, *Opinion Piece, Manag. Educ.* 34 (3), pp. 117–121.

speaking, leadership development encompasses all formal and informal learning that enables leaders to improve their own practice while performing the role. It is often focused on developing competences or skill acquisition (Bush, 2018). Leader development is an aspect of personal (leader) development and, wherever possible, the two - leader and leadership development - should interact and complement each other: it is not “an either/or” but “both”. The former is mainly about occupational role development, whereas personal or leader development is about the growth of the individual (the leader), often the “whole” person, and it almost always involves changes in self-awareness and understanding.

Personal development which almost always involves changes in self-awareness (Waters, 1998a) is therefore seen as a fundamental part of the job not an optional extra. Leader development is “above all the development of the individual and only secondarily the development of the role. The premise is that the more we learn about ourselves the better we will perform in our job – a direct consequence the one of the other” (Gray, 1987 cited in Waters, 1998b, p. 32). However, the more common approach to personal development is perhaps best described by Waters’ term “propersonal development”. This is development that is not genuinely “personal” in its focus but is personal development for professional development purposes (Waters, 1998a, p. 35). Tomlinson (2004, p. 1) states, “Personal effectiveness is a precondition of professional excellence” and notes the importance of self-management and self-awareness thus:

Improved self-management increases an education leader’s ability to cope with stress, resolve conflict, manage change and manage to change, achieve sustainable peak performance, build and lead effective teams, and influence organizational cultures (p.1).

In a similar vein, Nicholas and West-Burnham (2016) have argued that although leadership development and training opportunities are more available for leaders than ever before, the bulk of them focus on the technical and the organizational, yet:

There are other dimensions to (leader) sustainability that we believe are being neglected – that is the personal aspects of well-being and wellness. This might be best understood as personal efficacy (i.e., the development of the whole person) – the recognition that leadership is more than an aggregation of technical skills and that it requires the engagement of all aspects of the person (pp. 203–204).

School leaders need to acquire and develop their “intelligences” (e.g., emotional, contextual, strategic) as well as their knowledge and skills (Day and Sammons, 2013). Research, especially into successful and highly effective leaders (e.g., Day and Gurr, 2016) points, *inter alia*, to the importance of whole person development and it reinforces Waters’ exhortation for development which links improved personal management with increased professional efficacy. In 2012 Leithwood introduced the concept of “personal leadership resources” (PLRs) which included cognitive (e.g., problem-solving), social (e.g., acting in emotionally appropriate ways) and psychological resources (e.g., self-efficacy), seeing them as showing “promise of explaining a high proportion of variation in the practices enacted by school leaders” (Leithwood et al., 2020, p. 10).

One of the psychological resources noted by Leithwood was resilience and its importance has been noted in the literature as crucial to individual teacher and leader effectiveness (Beltman et al., 2011; Day and Hong, 2016; Day and Gurr, 2016; Sardar and Galdames, 2017). Although writing only about teachers, Beltman et al. (2011, p. 197) note how resilience is essential to “ensure that those who stay in the profession do not just survive, but thrive as confident and healthy professionals”.

Professional and personal spheres are mutually supportive and beneficial, however, we wish to propose that there is a need for development programmes to give greater prominence to elements of personal development and PLRs, including developing self-awareness, self-efficacy, reflection, managing emotions, wellbeing and resilience, which are even more important as the role of the principal changes and their work becomes ever more demanding and all-consuming.

The challenges of headship

It is known that the work of principals is complex and demanding. Research using observation, diaries and logs completed by headteachers show a relentless, complex and emotionally-demanding workload. An observational study in England (Earley and Bubb, 2013) showed principals spending long hours of demanding, varied and complex work, confirming the commonly-held view of headship as pressured, highly accountable and stressful (Phillips and Sen, 2011). For MacBeath and colleagues a good descriptor of what it means to be a headteacher in the 21st century is “intensification”: “increasing pressure to do more in less time, to be responsive to a greater range of demands from external sources and to meet a greater range of targets, accompanied by impatient deadlines to be met” (2012, p. 422). Fink summed it up well when he said principals were reporting being overburdened, overworked and overwhelmed (2010, p. 35).

The work of school leaders is more complex than ever with many challenges to be met. Examples of these challenges include accountability pressures associated with high-stakes testing; school inspections; league tables and exam performance; curriculum developments; changes in assessment; and human resources issues, such as recruiting quality teachers or maintaining staff morale. The main difficulty, however, is that at any point in time there are always numerous challenges and initiatives that have to be dealt with, often concurrently. For example, as part of a research project conducted for the English independent school sector, when heads

were asked to identify their three most significant challenges from a list of 16 provided by the researchers, one said, “The hard part is that there are never three things. It’s usually about a dozen!” (Earley et al., 2016, p. 10).

The role of school principals has always been demanding but it has changed considerably over recent decades, in particular with regard to the levels and patterns of accountability, the nature of their responsibilities and the extent of institutional autonomy (OECD, 2009; Schleicher, 2012). School autonomy and decentralization policies, found in an increasing number of national education systems, have given new and enhanced powers to school leaders who:

Sit at the fulcrum of high autonomy-high accountability systems and are expected to resolve the policy paradoxes of both competition and co-operation they should: exercise their autonomy to innovate in response to parental needs, whilst at the same time meeting centrally prescribed targets and requirements; improve literacy and numeracy scores every year, whilst maintaining a broad and balanced curriculum; close attainment gaps, while pushing the brightest and the best; and collaborate with their peers to develop skills and capacity, whilst competing to ensure that (their schools) move up the local hierarchy.

(Greany and Earley, 2021a,b, p. 5)

The challenges and demands are many and although they may change they are unlikely to disappear. School leaders, regardless of phase or sector, will need considerable support and on-going development, both professional and personal, if their schools are to provide the kind of education needed for students to live and work in the 21st century.

Across the world headteachers have had to come to terms with heightened expectations and increasing public accountabilities (MacBeath et al., 2012). This is a process that Gronn (2003, p. 147) has called “greedy work”, where school leaders work increasingly in a “high stakes” accountability culture where their job may well depend on the results of their last external school inspection. He comments on the “greedy work” of organizations for leadership in the current educational climate, reflecting the intense and potentially overwhelming nature of modern headship. The all-encompassing nature of headship is also noted by Bottery (2016) when discussing “the threats to leadership sustainability”:

Much of this extra work is hidden from public view in the additional hours at home, not just on weekday evenings but also at weekends and on “holiday”. If leaders do not do this then the job begins to run away from them. So a Catch 22 situation can develop, where work requires unsustainable hours if individuals are to keep on top of it, but where the job becomes unsustainable if they do not. Either route is likely to lead to extra worry and stress, to the four-in-the-morning syndrome, and to less personal sustainability (p. 98).

Similarly, Lee (2016) notes that while the use of technology may be desirable for faster and more efficient organizational communication, “using ICT outside of school and before/after office hours for school work seems to be a sign that principals’ work is adopting the characteristics of a 24/7 ubiquitous service, which commonly occurs in business sectors” (2016, p. 133).

All-consuming workloads and challenges must be seen as surmountable and the leadership task perceived as possible for “ordinary mortals” for organizations to ensure a continuing succession of effective leaders. As the punishing pace of their work increases, school leaders therefore need to be helped to develop confidence in their capabilities and to acquire appropriate knowledge, understanding and skills. They also need to know that they have or can develop the self-awareness, personal resolve and resilience to maintain their wellbeing and work-life balance and, importantly, that they have the support of others.

Research from the OECD (OECD, 2020; Schleicher, 2012) alerts us to the variations in the role and workload of principals in different countries. In England nearly two-thirds (61%) of headteachers described their work-life balance as poor or very poor, due “to long working hours or to deficiencies in working practices, such as school headteachers not knowing how to prioritize or delegate their work” (ibid, p. 18). The Teacher Workload Survey (Walker et al., 2019) found that headteachers on average reported working 56.9 hours in a week in March 2019. Not only is work-life balance desirable, it is increasingly clear that working long hours can affect wellbeing and damage health. A meta-review analysis of 25 studies in Europe, the USA and Australia, involving over half a million participants, found that those spending 55+ hours working per week were at greater risk of strokes than those working a standard week of 35–40 hours (Kivimäki, 2015). It is important for principals to manage their workload and wellbeing, deciding what *not* to do, to prioritize time and energies. In a study of Generation X school leaders (i.e., those under 40) it was noted that many principals were unintentionally discouraging the next generation to want to take on senior roles (Edge et al., 2021). Many people were being put off aspiring to headship by the responsibilities and risks they associated with the role and by the poor examples and role models they witnessed. Paradoxically perhaps, a large proportion of school leaders find the job enormously satisfying and would not want to do anything else (e.g., Earley and Weindling, 2004; Earley et al., 2016).

The loneliness of the principal’s role is considered in many studies and the emotional demands in the face of possible feelings of professional isolation are seen to be considerable and have not lessened over the years (Earley, 2013). A recent research study for the independent sector, which involved data collection from a variety of sources including a questionnaire survey, of the 155 respondents, a fifth said they did not manage professional isolation well, with experienced heads feeling it slightly more. One said “the isolation can be overwhelming and it sneaks up on you at funny times” (Earley et al., 2016).

This can be challenging in its impact on the health and wellbeing of leaders. The above survey also showed that heads’ wellbeing and work-life balance often needed consideration. For example, fewer than ten per cent stated that they were dealing with work-life balance “very well”, and fewer than 40 per cent were prepared to claim that they were coping even “quite well”. Nearly two-thirds

(64%) of respondents believed that they handled it “not well” or “badly”. Lifestyle advice was important for heads for as one experienced school leader noted “my mantra is ‘look after yourself, look after yourself’ – and what are you doing about it?” Annual health checks were considered valuable because: “heads are often bad at recognizing their own problems. They need to be more honest with themselves and actually require a bit of bullying to do something about their own wellbeing” (Earley et al., 2016).

Support for principals to deal with matters of professional isolation, workload and wellbeing may come from a variety of sources, both professional and personal. For example, research studies have pointed to the importance of fellow heads, senior colleagues, the chair of governors, local district officers, school advisers, networks (real and virtual), a coach or mentor, a spouse or partner. However, this support needs to be ongoing, flexible and responsive to changing needs over time (Bush, 2018; Earley et al., 2011).

Leaders need intellectual breadth and agility, self-awareness and self-confidence, good interpersonal skills, emotional management and high degrees of resilience if they are to deal successfully with complexity, resolve paradoxes and make informed decisions. They also need opportunities to think and reflect which is becoming increasingly difficult to find in the hurly-burly life of schools and the “busyness” of headship. The provision of support and opportunities to develop as highly effective and successful school principals is therefore crucial and given the earlier described context, more important than ever.

Surveying the leadership landscape, we find that the skills and capabilities required of school leaders have changed substantially, especially in England since the Education Reform Act of 1988 and the introduction of decentralization and school-based management (Earley and Greany, 2021). Simkins (2012) offers a detailed account of these changes and maps three distinct eras of school leadership training and development in England: the era of Administration; the era of Management; and the era of Leadership. To this we wish to argue the need for a fourth era of Leadership Development – or more accurately Leader Development – which in recognition of the current context in which principals work necessitates a greater focus on the personal needs of individual school leaders and the development of their personal leadership resources.

Developing leaders

Leader and leadership development are both critical at almost any level in an organization. We know that the most effective leadership development is that which provides a range of formal and informal learning and development activities and that the best programmes make good use of the workplace as a site for leadership learning. High levels of learning about educational leadership can occur as a result of attending an off-site training programme or because of the learning opportunities created within the workplace. Indeed the most effective leadership development is that which makes use of several methods in a complementary and reciprocal manner (Earley and Jones, 2010).

Preparation for leadership and on-going development programmes are common in both the private and public sectors. The trends are for organizations to have an approach to development which is wider than any programme and integrated in the work leaders do. Information that might previously have been delivered face-to-face is now often accessed online or covered in pre-course assignments. Face-to-face time is more oriented toward practice, discussion, reflection and feedback. Combinations of learning strategies are used, including lectures, case studies, experiential exercises, and simulations underpinned by feedback and coaching.

Approaches to leader and leadership development should be bespoke or personalized. The most successful training and development programmes are adapted to the needs of the individual and elements of customized programmes are now more commonly found (Pillans, 2015). It appears, as with teacher development more generally, it is not so much the case of “one size fits all” but rather that “one size fits no one” (Bubb and Earley, 2013). As such school leaders need to take greater responsibility for planning and implementing their own learning experiences to meet their personal and professional needs. In a similar vein, Lovett et al., (2015, p. 128) in their international review of school leadership development strategies, note “the limited role individuals seem to take in shaping their own professional learning, seemingly preferring to rely on others to determine what, why and how they should learn”. For them, leaders need to be able to reflect systematically and across time on their knowledge and understanding of leadership and that “knowledge of personal shortcomings and knowing what to do to overcome them is a particularly desirable self-related learning goal for leaders, because it helps them guard against arrogance, complacency, pretentiousness and narcissism” (ibid, p. 137).

In our view it is desirable to strive toward a healthy balance between what an individual “wants” and what they might “need”. Customized programmes are therefore essential and need a greater focus on the development of the individual leader, especially the development of the person – their personal effectiveness or personal efficacy – to enable them to be able to survive and thrive; and when necessary to revive. School leaders should be encouraged to take greater responsibility for their own personal and professional learning.

Opportunities to thrive, survive and revive

Principals and other school leaders’ development is most effective when time is built in for reflection but as noted above it is often hard to find such time given high workloads and “greedy” organizations. Structured opportunities to reflect with others (such as with a coach or in a tutorial or network) can lead to powerful learning. Reflection has long been identified as an important

leadership skill but it is in danger of getting lost due to work intensification. Adult learning theory or andragogy tells us that people need time and space to reflect on what they have learned and plan for how they will put such learning into action. But reflection does not come easily to busy people. The frenetic pace of life in schools makes it difficult to create space for leaders to reflect on what they have learned, and receive feedback and support, such as through coaching, all of which have been found to be critically important to sustaining adult learning (Pillans, 2015). Former headteacher and coach Viv Grant says:

What is needed is emotional support and a space for headteachers to reflect on how well they are doing the job and what they could do better. Social workers have supervision to help them process their toughest cases and corporate executives have space for lessons learned in between projects. School leaders need something similar.

(Grant, 2015 cited in Berkeley, 2021, p. 38).

Similarly, Nicholas and West-Burnham (2016) note that “in practical terms, the most powerful basis for (leadership) learning is supported reflection – support being provided through coaching and mentoring, the use of a reflective journal, structured reading to inform review and, perhaps most importantly, peer review and feedback on actual practice” (p. 22). It is increasingly recognized that learning is a key component of the job itself – not an adjunct or a bonus but a crucial element in the definition of the role – and therefore there is a need to schedule time and space for regular reflection along with networks (virtual and actual) to nourish, support and challenge. Layen (2015), using autobiography and life histories in her own exploratory development work with early childhood leaders, also notes the value of reflection for increasing self-awareness and self-concept. Furthermore, she sees insights so gained “may have potential to increase resilience, motivation and commitment and develop leadership maturity” (p. 286).

Experiences are important yet, “if people haven’t reflected on and learned from them, just having the experience in itself doesn’t necessarily mean people perform any better as a result” (Bowerman cited in Pillans, 2015, p. 31). Experience is not everything; people do not learn solely from experience but from reflecting on that experience – and that requires time. Leaders’ development is most effective when time is built in or created for reflection. Learning intensifies when we reflect on what we have experienced; and the more recent the experience we reflect on, the greater and more intense the learning. The key question, then, is how can opportunities for reflection and critical thinking be created for school leaders which will help them survive and thrive?

The role of coaching and mentoring in supporting reflection is discussed by Barnett and O’Mahony (2008) who argue that the coach’s ability to pose reflective questions to stimulate leaders’ thinking and action is perhaps the most important reflective tool. Coaches can stimulate thoughtful reflection through active listening and posing questions in non-threatening ways (Robertson, 2008) and there is no shortage of evidence that coaches and mentors influence the reflective abilities of the individual leaders they work with, helping them to examine various viewpoints, consider options and determine the best ways forward (Hobson, 2003). Coaching, especially executive coaching, can increase leaders’ self-awareness, improve their skills and decision-making, and help them become more reflective (O’Neal and Glasson, 2019; Jackson and Berkeley 2021a,b; Woods et al., 2020). Indeed, “in the quest for principled autonomy in leadership, nurturing the practice of critical reflexivity is essential” (Woods et al., 2020). Significantly, coaching can also help reduce professional isolation.

As well as coaching, opportunities for reflection can be created in other ways, e.g., forums, work discussion groups, action-learning sets, university programmes or courses, systematic study or short sabbaticals. They all have a role in encouraging reflection, research and critical thinking. In some parts of the corporate sector, leaders may have lots of peers around them, and bosses and team members who are increasingly used to giving support and feedback. These advantages are not routinely available to school principals, which may suggest that an emphasis on peer learning methods such as structured shadowing, coaching, mentoring and action learning are important, and more generally the creation of communities and networks where principals can learn and develop together. Some of this could be conducted virtually. Networks, both real and virtual, can be both supportive and reassuring to group members helping to combat the feelings of professional and personal isolation, while also encouraging reflection and ideas for strategic development.

Effective school leaders demonstrate certain qualities such as optimism, self-efficacy, resilience, proactivity and perseverance and they possess a range of “intelligences” – contextual, emotional, strategic, and of course cognitive intelligence. They need to draw upon all of these to be able to make the best possible decisions given the context and circumstances they face. Long working days are the norm but due to the nature of their job, heads cannot have an “off-day” as they constantly need to “perform”. As earlier noted, the importance of headship to effective school performance is now well established but ensuring that heads remain “at the top of their game” and do not “plateau out” (after 5–7 years) has long been recognized as a problem (Earley and Weindling, 2007). Headteachers need to be enabled to continue to thrive – or given opportunities to revive – if they are to continue to perform at the high levels demanded of them.

Many years ago, the James Report (DES, 1972) advocated a sabbatical period or study leave of one term for educational professionals after 7 years of service. This idea has been taken up in various forms in other parts of the world (Downing et al., 2004) and in some organizations employees (usually senior staff) with over 5 years’ service may be entitled to apply for a period of sabbatical leave. James’ idea was never taken up in England but in 2001 a sabbatical scheme for experienced teachers working in areas of high challenge was introduced in England (DfEE, 2001). This was part of a wider, ambitious government CPD policy, and although short lived, the NFER evaluation showed many benefits for both individuals and their schools (Downing et al., 2004) of this 6-week funded period of absence. For example, improvements in confidence, refreshment and self-esteem, along with raised skills and

knowledge were reported in 87% of sabbaticals, while in 83% of cases there were reports of the school developing as a professional learning community as a result of the sabbatical.

Crucially, a sabbatical, even if only for a relatively short-period, should help principals to survive (e.g., by knowing that there is a period of time during one's career to rest and recharge), to thrive (e.g., by undertaking study visits or a high quality course) and, if necessary, to revive (e.g., by being given time to rethink, to recharge, to be rejuvenated or refreshed). In the USA, sabbaticals or "time out" have traditionally been treated as "R and R" (rest and recuperation) rather than as providing development opportunities as has been the case in most other education systems where such schemes are found (Downing et al., 2004, p. 3). In some case the time away from school has been devoted to undertaking a course of study (study leave). Serious and systematic study, which is what undertaking a higher degree involves, provides opportunities for reflection and updating, research and networking. It helps to develop the "intelligences" earlier identified and an awareness of the qualities possessed by highly effective and successful principals. Academic studies are the place to ponder, study and grow, with experience-based debates scaffolded by theory and deeper conceptualization. For participants they provide "rich opportunities for the kind of learning that reflected personal goals in education" (Hargreaves and Preece, 2014, p. 144), including the spaces to interact with other leaders, to explore existing published research and to undertake research into areas they personally found important.

Some sets of leadership competences include "thinking" e.g., analytical, systems and conceptual thinking and the eagerness to learn – all are found in university accredited Masters and Doctoral programmes, along with the all-important critical approach to reading research and policy. Undertaking higher degrees and professional qualifications in leadership can be helpful and could be offered at various career stages when people are ready to gain maximum benefit from them. These and similar opportunities would enable school leaders to know themselves, "their ethics, values, strengths, weaknesses and motivations – and what it might take for them to make informed transitions from aspirant, to novice, to experienced leader. Knowing self is also fundamental to sustaining purpose, interest and commitment throughout a career (Lovett et al., 2015).

Secondments too, can provide opportunities for development, offer access to expert personnel and provide a safe way of exploring career options, although for some the return to school can prove difficult (Tuohy and Lodge, 2003).

In our view, following James (1972), the opportunity to take sabbaticals or study leave for principals, indeed all teachers, should be an entitlement. The idea needs consideration once again, despite existing budgetary constraints and teacher shortages. However, the manner in which a period of leave is used should not be prescribed but more the result of a discussion of needs with a coach, fellow head or line manager such as the governing body or board of trustees. Sabbaticals and study leave provide the opportunity for both leader development and leadership development to be much more personalized. They also provide opportunities to survive, thrive and revive as school leaders.

The school governing body or board of trustees can assist in providing services and offering support for school leaders. In English schools the governing body is specifically responsible for the wellbeing of the principal and there is greater recognition of the part "healthy" organizations can play in issues around staff wellbeing and welfare (Bubb and Earley, 2004; Bingham and Bubb, 2021). Steward (2014, p. 67) with specific reference to resilience, recommends that governors should be involved in supporting heads to undertake a regular risk-analysis with a resulting action plan. Regardless of its source, psychological support is needed since many school leaders work under considerable pressure, both physically and emotionally, feel exhausted, and often experience feelings of isolation and loneliness. Regular meetings of principals, both face-to-face and virtual, with a trusted group of others where problems can be shared and solutions found, along with sabbaticals or study leave for a period (a term, a month or even only a week) to undertake research, read or observe and learn from an outstanding leader, are examples of valuable opportunities for both surviving, thriving and, where needed, reviving.

Another benefit of sabbaticals, study leave and secondments is that other members of staff will have to "act-up" temporarily. Highly effective leaders are recognized as developing other leaders and participating in such activities will help develop leadership capacity and capability in others, as well as assisting with succession planning. James' (DES, 1972) recommendation of a term's secondment every 7 years is an important means by which exhaustion and burn out can be prevented and opportunities given for personal and professional refreshment and rejuvenation.

Conclusion

The OECD's leadership reports (Pont et al., 2008; OECD, 2009; Schleicher, 2012) identified four policy levers which, taken together, can improve school leadership practice, namely:

1. (Re)defining school leadership responsibilities
2. Distributing school leadership
3. Making school leadership an attractive profession
4. Developing skills for effective school leadership

To the last OECD policy lever we would wish to add the need for a greater focus on the individual leader and their personal and professional needs, to personalize development and to offer the support necessary to survive and thrive in what is generally perceived to be an ever-demanding role in a challenging and constantly changing environment. Following Simkins' (2012) classification, we wish therefore to argue for a fourth era of "leader development" where there is a greater focus on personal development and personal leadership resources.

Learning to be a leader at any level in an organization is not a one-time or one-off event but a continuing or career-long process. Leaders' skills can be developed when their role and responsibilities are clearly defined and specific training in those skills is made available to them (Schleicher, 2012). An emphasis on peer learning methods such as structured shadowing, coaching, mentoring and action learning are going to be important in any leadership development programme and more generally the creation of communities and networks of leaders, both virtual and real, who can learn and develop together (Darling Hammond et al., 2009; Fluckiger et al., 2014).

In this article we have argued for the need to ensure we do not lose sight of the individual and their personal development, or what Water's (1998a,b) refers to as "pro-personal" development. A greater emphasis needs to be given to Leithwood's (2012) personal leadership resources. The importance of developing self-awareness, emotions management, being reflective and resilient while managing workload and wellbeing should not be underplayed. Individual or personal effectiveness or efficacy is key for continuing success as a leader yet is not given the prominence it deserves. We have suggested there is a need for a combination of customized or bespoke school-based and externally-provided developmental opportunities and experience to be provided in a range of settings. These should be based on careful analysis of an individual's needs (Crawford and Earley, 2011) and include strategies to enhance personal efficacy.

There needs to be a growing recognition that schools should be seen as "healthy organizations" where concern is shown not only for professional development but also for individuals' personal development, their workload, welfare and wellbeing. Unsurprisingly, when educational leaders are overworked, emotionally strained and forgetful about their own wellbeing, their performance is affected negatively. Mahfouz (2020) when discussing stress management argues that if we wish to retain school leaders and help them lead their schools effectively then there is an urgent need to focus on social emotional skills based or mindfulness-based programmes that include self-awareness and self-management. Similarly, O'Neal and Glasson (2019) in an evaluation of an innovative, multi-dimensional development programme for experienced heads in Australia, point to the need to focus on health and wellbeing (underpinned by an executive health assessment) as this was found by far to be the most successful aspect of the program.

This article has argued that greater consideration be given to the role of sabbaticals and study leave to help principals – and other educational professionals – to survive and thrive. Coaching too is an important part of the personalized route and where it exists it has been highly valued, helping to develop leaders' thinking and reflection (e.g., Blackman, 2010; Crawford and Earley, 2011; O'Neal and Glasson, 2019; Jackson and Berkeley 2021a,b) and resilience (e.g., Sardar and Galdames, 2017; Steward, 2014). For Jackson and Berkeley (2021b) "coaching should be seen as a *basic duty of care* rather than an individual indulgence—especially in the context of national concerns regarding headteacher wellbeing, recruitment and retention" (p. 98 emphasis in original). However, as they further note:

Rather than viewing regular professional development, such as coaching for themselves, as a fundamental duty of care owed to them—as they might with any other member of staff—they are often inclined to think of it as a *luxury* or even an *indulgence* (p. 114).

Capacities necessary for on-going effective leadership include the capacity to collaborate with others, the capacity for resilience, and the capacity to survive and thrive in complex contexts. As Kellerman (2012) has remarked "leading has become a high wire act that only the most skilled are able to perform successfully over a protracted period of time" (2012, p. 263). We must ask whether it is reasonable to expect principals to continue to undertake such a high powered, challenging and demanding role for a number of years and continue to do it well without the provision of opportunities to survive and thrive, and where necessary, revive. If the role of the principal is to be seen as attractive *and* manageable, then consideration must be given to providing more assistance, greater support and development – especially personal (leader) development – with the recognition of the need for opportunities for professional refreshment and rejuvenation. In this way the chances of schools being well-led and managed and high achieving will be considerably improved.

References

- Bakker, A., 2015. Towards a multilevel approach of employee well-being. *Eur. J. Work. Organ. Psychol.* 24 (6), 839–843.
- Barnett, B., O'Mahony, G., 2008. Mentoring and coaching programmes for the professional development of school leaders. In: Lumby, J., Crow, G., Pashiardis, P. (Eds.), *International Handbook on the Preparation and Development of School Leaders*. Routledge, London pp. x-y.
- Beltman, S., Mansfield, C., Price, A., 2011. Thriving not just surviving: a review of research on teacher resilience. *Educ. Res. Rev.* 6, 185–207.
- Berkeley, A., 2021. Keeping your head: the unspoken realities of headship. In: Jackson, E., Berkeley, A. (Eds.), *Sustaining Depth and Meaning in School Leadership: Keeping Your Head*. Routledge, London, pp. 17–40.
- Bingham, D., Bubb, S., 2021. Leadership for wellbeing. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London. pp. 143–152.
- Blackman, A., 2010. Coaching as a leadership development tool for teachers. *Prof. Dev. Educ.* 36 (3), 421–441.
- Bottery, M., 2016. Not so simple: the threats to leadership sustainability. *Manag. Educ.* 30 (3), 97–101.
- Bubb, S., Earley, P., 2004. *Managing Teacher Workload: Work-Life Balance and Wellbeing*. Sage, London.
- Bubb, S., Earley, P., 2013. The use of training days: finding time for teachers' professional development. *Educ. Res.* 55 (3), 235–247.
- Bush, T., 2018. Preparation and induction for school principals: global perspectives. *Manag. Educ.* 32 (2), 66–71.
- Cheng, A., Szeto, E., 2016. Principals' changing work in a time of Hong Kong education reform: challenges and opportunities. *Int. Stud. Educ. Adm.* 44 (3), 21–36.

- Crawford, M., Earley, P., 2011. Personalised leadership development? Lessons from the pilot NPQH in England. *Educ. Rev.* 63 (1), 105–115.
- Darling Hammond, L., Meyerson, D., LaPointe, M., Orr, M., 2009. *Preparing Principals for a Changing World: Lessons from Effective School Leadership Programs*. John Wiley, New York.
- Day, C., Gurr, D. (Eds.), 2016. *Leading Schools Successfully: Stories from the Field*. Routledge, London.
- Day, C., Hong, J., 2016. Influences on the capacities for emotional resilience of teachers in schools serving disadvantaged urban communities: challenges of living on the edge. *Teach. Teach. Educ.* 59, 115–125.
- Day, C., Sammons, P., 2013. *Successful Leadership: A Review of the International Literature*. CfBT Education Trust, Reading.
- Department for Education and Employment, 2001. *A Sabbatical Scheme for Experienced Teachers in Challenging Schools*. DfEE, London.
- Department for Education and Science, 1972. *The James Report: Teacher Education and Training*. HMSO, London.
- Downing, D., Watson, R., Johnson, F., Lord, P., Jones, M., Ashworth, M., 2004. *Sabbaticals for Teachers: An Evaluation of a Scheme Offering Sabbaticals for Experienced Teachers Working in Challenging Schools*. NFER, Slough.
- Earley, P., Bubbs, S., 2013. A day in the life of new headteachers: learning from observation. *Educ. Manag. Adm. Leader* 41 (6), 782–799.
- Earley, P., Greany, T., 2021. Postscript. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London. pp. 273–279.
- Earley, P., Jones, J., 2010. *Accelerated Leadership Development: Fast-Tracking School Leaders*. Institute of Education, London. Bedford Way series.
- Earley, P., Weindling, D., 2004. *Understanding School Leadership*. Paul Chapman/Sage, London.
- Earley, P., Weindling, D., 2007. Do school leaders have a shelf life? Career stages and headteacher performance. *Educ. Manag. Adm. Leader* 35 (1), 73–88.
- Earley, P., Nelson, R., Higham, R., Bubbs, S., Porritt, V., Coates, M., 2011. *Experiences of New Headteachers in Cities*. NCSL, Nottingham.
- Earley, P., Bubbs, S., Berry, J., 2016. *The Professional Development Needs of Heads of HMC Schools: Final Report*. UCL/HMC, London.
- Earley, P., 2013. *Exploring the School Leadership Landscape: Changing Demands, Changing Realities*. Bloomsbury, London.
- Edge, K., Galdames, S., Horton, J., 2021. Diversity: new leaders and new leadership. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London. pp. 231–242.
- Fink, D., 2010. *The Succession Challenge: Building and Sustaining Leadership Capacity through Succession Management*. Sage, London.
- Fluckiger, B., Lovett, S., Dempster, N., 2014. Judging the quality of school leadership learning programmes: an international search. *Prof. Dev. Educ.* 40 (4), 561–575.
- Future Leaders Trust, 2016. *The School Leadership Challenge: 2022*. FLT, Teach First and Teaching Leaders, London.
- Greany, T., Earley, P. (Eds.), 2021a. *School Leadership and Education System Reform*, second ed. Bloomsbury, London.
- Greany, T., Earley, P., 2021b. Introduction. In: Greany, T., Earley, P. (Eds.), *School Leadership and Education System Reform*, second ed. Bloomsbury, London. pp. 1–15.
- Gronn, P., 2003. *The New Work of Educational Leaders: Changing Leadership in an Era of School Reform*. Sage, London.
- Hargreaves, E., Preece, P., 2014. The value of the personal in teachers' professional learning: a case study. *Prof. Dev. Educ.* 40 (1), 130–146.
- Hobson, A., 2003. *Mentoring and Coaching for New Leaders*. NCSL, Nottingham.
- Huber, S. (Ed.), 2010. *School Leadership: International Perspectives*. Springer, Utrecht.
- Jackson, E., Berkeley, A. (Eds.), 2021a. *Sustaining Depth and Meaning in School Leadership: Keeping Your Head*. Routledge, London.
- Jackson, E., Berkeley, A., 2021b. Executive coaching for the hidden realities of life as a school leader. In: Jackson, E., Berkeley, A. (Eds.), *Sustaining Depth and Meaning in School Leadership: Keeping Your Head*. Routledge, London. pp. 97–115.
- Kellerman, B., 2012. *The End of Leadership*. Harper Collins, New York.
- Kivimäki, M., 2015. Long working hours and risk of coronary heart disease and stroke: a systematic review and meta-analysis of published and unpublished data for 603,838 individuals. *Lancet* 386 (10005), 1739–1746.
- Layen, S., 2015. Do reflections on personal autobiography as capture in narrated life-stories illuminate leadership development in the field of early childhood? *Prof. Dev. Educ.* 41 (2), 273–289.
- Lee, M., 2016. The changing nature of school principals' work: lessons and future directions for school leadership research. *Int. Stud. Educ. Adm.* 44 (3), 129–135.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* <https://doi.org/10.1080/13632434.2019.1596077>.
- Leithwood, K., 2012. *Strong Districts and Their Leadership*. Council of Ontario Directors of Education, Toronto, ON.
- Lovett, S., Dempster, N., Fluckiger, B., 2015. Personal agency in leadership learning using an Australian heuristic. *Prof. Dev. Educ.* 41 (1), 127–143.
- MacBeath, J., O'Brien, J., Gronn, P., 2012. Drowning or waving? Coping strategies among Scottish head teachers. *Sch. Leader. Manag.* 32 (5), 421–437.
- MacBeath, J., 2009. Recruitment and retention of senior school leaders: meeting the challenge. *Eur. Educ. Res. J.* 8 (3), 1–12.
- Mahfouz, J., 2020. Principals and stress: few coping strategies for abundant stressors. *Educ. Manag. Adm. Leader* 48 (3), 440–458.
- Nicholas, L., West-Burnham, J., 2016. *Understanding Leadership: Challenges and Reflections*. Crown House, Carmarthen.
- O'Neal, M., Glasson, S., 2019. Revitalising professional learning for experienced principals: energy v ennui. *Educ. Manag. Adm. Leader* 47 (6), 887–908.
- OECD, 2009. *Improving School Leadership: The Toolkit*. OECD Publishing, Paris.
- OECD, 2013. *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. OECD Publishing, Paris.
- OECD, 2020. How much time do teachers and school heads spend teaching and working?. In: *Education at a Glance 2020: OECD Indicators*. OECD Publishing, Paris.
- Phillips, S., Sen, D., 2011. Stress in headteachers. In: Langan-Fox, J., Cooper, C. (Eds.), *Handbook of Stress in the Occupations*. Edward Elgar Press, Cheltenham, pp. 177–200.
- Pillans, G., 2015. *Leadership Development: Is It Fit for Purpose?* Corporate Research Forum, London.
- Pollock, K., 2016. Principals' work in Ontario, Canada: changing demographics, advancements in information communication technology and health and wellbeing. *Int. Stud. Educ. Adm.* 44 (3), 55–74.
- Pont, B., Moorman, H., Nusche, D., 2008. *Improving School Leadership: Policy and Practice*. OECD Publishing, Geneva, Switzerland.
- Robertson, J., 2008. *Coaching Educational Leadership: Building Leadership Capacity through Partnership*. Sage, London.
- Sardar, H., Galdames, S., 2017. School leaders' resilience; does coaching help in supporting headteachers and deputies? Coaching. On-line access.
- Schleicher, A., 2012. *Preparing Teachers and Developing School Leaders for the 21st Century*. OECD, Paris.
- Simkins, T., 2012. Understanding school leadership and management development in England: retrospect and prospect. *Educ. Manag. Adm. Leader* 40 (5), 621–640.
- Steward, J., 2014. Sustaining emotional resilience for school leadership. *Sch. Leader. Manag.* 34 (1), 52–68.
- Thomson, P., 2009. *School Leadership: Heads on the Block?* Routledge, London.
- Tomlinson, H., 2004. *Educational Leadership – Personal Growth for Professional Development*. Paul Chapman publishing, London.
- Tuohy, D., Lodge, A., 2003. *The Secondment of Teachers: How Secondment Impacts on Career Dynamics in Teaching and the Costs within the Education System*. DES/Clare: Clare Education Centre Publication.
- Walker, M., Worth, J., Van den Brande, J., 2019. *Teacher Workload Survey 2019*. DfE, London.
- Waters, M., 1998a. Personal development for teachers – part 1. *Prof. Dev. Today* 1 (2), 29–38.
- Waters, M., 1998b. Personal development for teachers – part 2. *Prof. Dev. Today* 1 (3), 29–36.
- Woods, P., Roberts, A., Jarvis, J., Culshaw, S., 2020. Autonomy, leadership and leadership development in England's school system. *Sch. Leader. Manag.* <https://doi.org/10.1080/13632434.2020.1811661>, 1–20.

Leadership for equity in schools

Suzanne Carrington, School of Early Childhood and Inclusive Education, Creative Industries, Education and Social Justice, QUT, Kelvin Grove, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	401
International trends	401
An inclusive approach to education	402
A way forward: leadership for equity	402
Leadership for equity	403
Challenges in leading for equity	403
Leadership frameworks and taxonomies	404
Yukl (2012)	404
Hitt and Tucker (2016)	405
Establishing and conveying vision	405
Facilitating a high-quality learning experience for students	405
Building professional capacity	405
Creating a supportive organization for learning	405
Connecting with external partners	405
Starratt (2014)	405
Duignan (2012)	406
Brown (2004)	406
Theoharis (2007)	406
School leaders enacting change: Australian studies	406
Study one: promoting equity in schools	406
Study two: using the index for inclusion to promote equity	407
Conclusion	407
References	408

Introduction

Education systems around the world are developing policy frameworks to support greater equity for all within a context of global uncertainty and challenge, increasing school autonomy, accountability, and competition (Ainscow et al., 2018; Brown, 2004; Riddle and Apple, 2019). A focus on equity ensures that people are treated fairly according to need rather than everyone treated equally in the same way. Equal treatment perpetuates existing inequalities (Ryan and Rottmann, 2007). Leadership for equity is also aligned with social justice leadership (Blackmore, 2002) that aims to advance equity and address marginalization in education. Effective school leadership is instrumental in education reform and change in an education system and at school level (Fullan, 2005).

School leaders working to enact socially just change need support to “confront socially difficult topics with respect, dialog, and a continuous expansion of awareness, acknowledgment, and action” (Brown, 2004: 80) and this requires leaders to understand their own beliefs and values that inform their ethical and moral behavior. This article will firstly consider a range of international trends that influence the challenges and priorities for school leaders. Secondly, theoretical approaches are considered that guide a way forward for school leaders. Thirdly, the article addresses the challenges and ethical dilemmas that school leaders face when mobilizing equity in schools. Fourthly, a range of leadership taxonomies and frameworks are described that could inform professional learning for aspiring and current school leaders. Finally, two Australian studies describe ways that school leaders have enacted change for equity within policy contexts that emphasize competition and top-down accountability.

International trends

Addressing equity in education is a challenge when “exclusion resides deep in the bones of education” (Slee, 2018: 11) where there is a long history of connection between schooling and divisions in social class due to race, ethnicity, gender, geography, sexuality, ability and disability, and religion. In current times, there is also a focus on standardization, testing and increased accountability, and the quality of school leadership and teaching. This trend has been described as the Global Education Reform Movement (GERM) (Sahlberg, 2012). These international trends have led to increased rather than decreased disparities in education quality and opportunity between advantaged and marginalized groups because it may be perceived that supporting students who need additional help, such as students with disability, poses a threat to success in the competitive school market

(Ainscow et al., 2012). Patterns of disadvantage may be so deeply ingrained in education and may not be recognized, particularly by school leaders and teachers. There is a need for school leaders to address unexamined everyday practices in schools including patterns of interaction between students, staff, and parents. As Slee (2018: 16) suggests:

Schools are forged within the furnace of competitive individualism, and students are reduced to the bearers of results. They become individually measurable units that when aggregated reveal the performance of a teacher, a school, a school district, state jurisdiction or nation-state. As individual units, students manifest risk or opportunity.

Despite the impact of economic, political, and cultural factors on inequality in education and society, there has been a steady movement toward a more inclusive approach to schooling that promotes equity, influenced by a human rights perspective.

Governments often call for school leaders to address new policy initiatives such as equity without consideration of the complexity required for long-term sustained change (Niesche and Gowlett, 2019). There may also be a lack of understanding about how system and school organizations perpetuate educational inequity despite school leaders trying to lead organizational change in their schools (Rottmann, 2007). Change in education requires challenging dominant norms in ideas and values that impact on how people are viewed and included or excluded in schools and school leaders may not have the knowledge and skills to help people in their schools to critically reflect and consider alternative approaches to schooling. Preparing school leaders to drive change for equity and social justice requires allocating time for leaders to learn about how schooling perpetuates inequality so that they can consider different points of view and develop a values base that aligns action with vision. The frameworks and taxonomies (see “**Leadership frameworks and taxonomies**” section) will be useful for school leaders to support self-reflection and guide understanding of knowledge, skills, and mindsets that are required to lead schools in increasingly complex times.

An inclusive approach to education

A focus on inclusion in education is based on a commitment to key values such as equity, justice, moral and ethical integrity, an appreciation of diversity, and the importance of contributing to and benefiting from a deeper understanding of other cultures. Inclusive education is an approach where all people are valued and treated with respect (Booth and Ainscow, 2011), and is endorsed by the widely ratified United Nations (UN) Convention on the Rights of Persons with Disabilities (CRPD) (UN, 2006) and Sustainable Development Goal Four, which is to “Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all” (UN, 2015). These international conventions, combined with anti-discrimination legislation, have influenced national policy around the world to support inclusion for all students in schools. We now have a conclusive body of evidence that shows that an inclusive approach to education leads to positive academic and social emotional outcomes for all students (Cologon, 2019; Hehir et al., 2016; Szumski et al., 2017).

Achieving an inclusive approach to education is challenging because it requires school leaders and teachers to challenge their understanding of difference that draws on a social-cultural perspective of diversity in society (Carrington and Saggars, 2021). The values and beliefs embedded in these perspectives create a new set of possibilities, expectations, and commitments in education (Skrtic, 1991) that could be achieved in a “moving school” which is continually seeking to develop and refine its responses to the challenges it meets (Rosenholtz, 1989). Moving a school forward to support equity for all students requires a more reflective and critical approach to school leadership at both the systems and school levels, as well as collaborative teamwork (AuCoin et al., 2020).

A way forward: leadership for equity

Effective school leadership is an important factor in bringing about any significant change in the education system (Day et al., 2009) and a critical approach to school leadership is consistent with progressing the ideals of equity and inclusion (Ryan, 2007). School leaders who take a critical approach drive change in the education system, the school, and for individuals with a focus on addressing inequity for the less fortunate and marginalized students. These leaders recognize that they have a leadership role to raise awareness of oppression and injustice so they can act to change them by working collaboratively with school stakeholders (Ryan, 2007).

A critical approach to school leadership links leadership and change (Rottmann, 2007) and acknowledges the importance of school context (Niesche and Gowlett, 2019). School leaders would consider opportunities for change in all areas of school culture, policy, and practice and would be “open to challenging systemic oppression” (Rottmann, 2007: 77). There is a need for school leaders to make choices between competing forces between the political and bureaucratic requirements and the needs of their teaching staff and students (Mulford et al., 2009).

Shields (2014) describes transformative leadership as a critical approach that is linked to promoting equity and inclusion with a focus on a sense of values and beliefs that inform how leaders think and lead. School leaders have a transformational role that can impact individuals and the collective of staff, students, and parents in their school community by developing relationships that are equitable, caring, and open. A transformative leadership approach requires collaboration and inclusive governing structures to address inequity and fosters an inclusive school culture (Cooper, 2009). Leaders who are committed to enact equity need to be activist oriented and morally transformative (Dantley and Tillman, 2006).

This type of approach is required for school leaders wanting to establish an inclusive school culture (Carrington, 1999).

It is acknowledged that taking a critical and transformative approach to leadership has a risk of focusing too much on critique without considering practical ways forward (Niesche and Gowlett, 2019). These authors suggest that “education leaders want straightforward answers and solutions to help make their jobs easier and their leadership more effective and that socially critical perspectives hinder this work” (Niesche and Gowlett, 2019: 18). Leading socially just change requires school leaders to “disrupt rather than reinforce the current social order” (Rottmann, 2007: 60) and this takes time at an individual, collective, and theory level. With these thoughts in mind, the educational change required for our schools to be equitable and inclusive for all students will need to be slow and incremental and require critical dialog and planning over time with people at school and system levels of education (Rottmann, 2007). This is because there is a need for multiple levels of change. While it is necessary to acknowledge this complexity, school leaders have an important role in their own school communities to recognize inequities and injustices and interrupt and challenge the status quo. It is clear that long-term change requires school leaders and staff to work together over a sustained period of time.

Leadership for equity

School leaders who wish to promote an inclusive, equitable, and transformative agenda for their schools are often confronted with questions about how to ensure that their leadership is ethical (Shields, 2014). “Ethical leadership is the attempt to act from the principles, beliefs, assumptions, and values in the leader’s espoused system of ethics” (Starratt, 2004: 5). A school leader wanting to progress equity needs to have the courage to examine, challenge, and, as necessary, correct situations and practices that promote inequity. This will involve drawing on ethical principles that guide decision making to drive change, action, and advocacy to affect the change. An ethical school leader adopts an analytical stance when examining their actions and resolving problems (Yasir and Mohamad, 2016).

Recently there has been a stronger emphasis on understanding how ethics and morals guide school leaders to recognize that leadership is values-based (Cranston et al., 2006; Fullan, 2003; Sergiovanni, 1992; Starratt, 2012). “Values are fundamental guides and prompts to action. They spur us forward, give us a sense of direction and define a destination” (Booth, 2011: 33). School leaders can guide their school community to develop a shared set of inclusive values that supports a clear vision for the school (Cologon, 2019). The leader’s ethical decision making and actions need to be consistent with that vision and shared values. It is acknowledged that school leaders operate “within established cultures and existing contexts that influence the effect of actions, and—consciously or unconsciously—change the culture of their schools through their actions” (Swaffield and Major, 2019).

Challenges in leading for equity

School leaders have challenges and face complex ethical dilemmas when making decisions that will impact on people’s lives and this is what they say often keeps them awake at night (Cranston et al., 2006). They are often grappling with situations that concern individuals, collectives in the organization, systemic policy, and the social-cultural and political context. With a focus on progressing an inclusive approach to schooling that promotes equity, there is a need for a school leader to be brave in leading dialog where there will be different views and values that need to be heard. Compassion and respect for all, good communication, and a calm and steady approach will be required when discussing challenging topics such as: Who gets what?; How to distribute resources between individuals and groups?; Who is treated in what ways? (are learners valued equally?); Is there a hierarchy of valuing some characteristics and cultures more than others?; Who can do what? Drawing on the work of Kidder (1995), Duignan (2012) described the following categories of tensions that impact on making decisions:

- **Common good and individual good.** This challenge is about how a decision will promote the good of a group such as a group of students or parents, or the rights of an individual such as a teacher or a student.
- **Care and rules.** Decisions are influenced by the compassion of the school leader for the individual circumstances and consideration of how that might override policy or rules.
- **Long-term and short-term considerations.** Any decision made by a school leader will have long-term or short-term consequences. This is particularly relevant as leading change to support equity and inclusion requires a long-term plan to allow for dialog, reflection, and planning with a school community.
- **Loyalty and honesty.** School leaders are torn between being committed to an organization and colleagues, for example, and this may influence if they can speak truthfully about a situation.
- **Service and economic rationalism.** Promoting equity for all in schools is likely to clash with the imperatives of economic rationalism associated with increasing accountability and improving school performance.
- **Status quo and development.** School leaders progressing equity in schools often need to challenge the status quo and enact, develop, and embrace change.

School leaders usually have many opportunities to learn administration and managerial skills but often have not had the opportunity to learn and embrace the ethical dimensions of leadership that impact on leading schools for equity (Starratt, 2004). The impact of critical theory has led to approaches to school leadership that are critical and transformative and this requires school leaders to have a greater self-awareness and ideological clarity (Dantley and Tillman, 2006). If we expect school leaders to challenge

others' thinking and practice then we need to support school leaders to have a better understanding of leadership roles and attributes that may influence their work. This is particularly the case for school leaders working to support equity and social justice as they face many pressures and dilemmas as discussed previously. School leaders need to not only have the knowledge and expertise to drive students' performance higher but also have the leadership attributes to support equitable decision making to equalize outcomes for students (Gross et al., 2003). A range of leadership frameworks and taxonomies can support professional learning for school leaders who want to work in critical and transformative ways.

Leadership frameworks and taxonomies

A range of leadership taxonomies and frameworks have been developed to support a better understanding of the roles and attributes of school leaders (Yukl, 2012). Leadership attributes could include the values that influence reasoning that might influence how a school leader enacts socially just change. A school leader will have values that influence approaches to decisions about staff, students, resources, assessment, curriculum, and pedagogy. Attributes of school leadership include whether or not leaders adopt a collaborative approach and the extent to which leaders develop trust among staff (Karami-Akkary et al., 2019). Taxonomies and frameworks of behaviors and attributes can support understanding about how they interact in complementary ways in context and could inform professional learning programs for school leaders (Yukl, 2012).

Following is an overview and comparison of the taxonomies and frameworks that can be used to support the work of school leaders working for socially just change.

Yukl (2012)

Yukl (2012) reviewed the literature and developed a four-level hierarchical behavior taxonomy with 15 specific behaviors for effective school leadership. The particular focus of Yukl's work was on school leadership to improve performance. The four elements of this taxonomy (Yukl, 2012: 68) are: (1) task-oriented ("to accomplish work in an efficient and reliable way"); (2) relations-oriented ("to increase the quality of human resources and relations"); (3) change-oriented ("to increase innovation, collective learning, and adaptation to the external environment"); and (4) external ("to acquire necessary information and resources, and to promote and defend the interests of the team or organization"). The 15 specific behaviors associated with each level are listed in Table 1.

Yukl (2012) concluded that more research is needed to explore how leadership behaviors interact in different situations or contexts. This was because he noted that aspects of the situation or context, such as the socio-economic and cultural backgrounds of families and incidence of students with disability or learning difficulties in the school, determine which leadership behaviors are relevant in the context. He emphasized that it is the flexibility of behaviors that is important for effective school leadership (Yukl and Mahsud, 2010) and this will be particularly relevant for school leaders working to achieve equity. Yukl actually acknowledged that future research could establish a fifth focus on ethical and socially responsible behaviors that could include a values approach to behavior (2012). This would be a welcome addition to this taxonomy and support professional learning for leadership for equity. This might involve behaviors like communicating ethical standards, encouraging ethical conduct based on a set of inclusive values, modeling ethical behavior, and opposing non-ethical conduct. The idea about another area of consideration that reflects ethics and values is also supported by Brown and Treviño (2006) review of ethical leadership and theoretical influences. It is clear that research on understanding leadership ethical behaviors and the effects of ethical leadership is limited.

Table 1 Comparison of Hitt and Tucker (2016) and Yukl (2012) taxonomies.

Hitt and Tucker (2016) taxonomy	Yukl (2012) taxonomy	
Establishing and conveying vision	Relations-oriented	Supporting
Facilitating a high-quality learning experience for students		Developing
Building professional capacity	Change-oriented	Recognizing
		Empowering
		Advocating change
		Envisioning change
Creating a supportive organization for learning	Task-oriented	Encouraging innovation
		Facilitating collective learning
		Clarifying
		Planning
Connecting with external partners	External	Monitoring operations
		Problem solving
		Networking
		External monitoring
		Representing

Hitt and Tucker (2016)

Hitt and Tucker (2016) synthesized the research about leadership frameworks from 100 studies published between 2000 and 2014 to develop a unified framework that organizes leadership practice and behavior into five overarching domains: (1) establishing and conveying vision; (2) facilitating a high-quality learning experience for students; (3) building professional capacity; (4) creating a supportive organization for learning; and (5) connecting with external partners. Each of these domains will now be described.

Establishing and conveying vision

This domain includes creating and articulating the vision for the school and setting goals and performance standards that are linked to the vision. The key point is that school leaders need to work with staff over time in collaboratively developing the vision to ensure that their values are reflected in the everyday work. Hitt and Tucker (2016) emphasize that it is the methods by which the direction is decided and the associated activities that are important.

Facilitating a high-quality learning experience for students

Hitt and Tucker (2016) acknowledge the need to personalize the learning environment to meet the diverse needs of students. This is supported by DeMatthews et al. (2020) who suggest that in the context of inclusive education, school principals need to lead teachers in discussing and planning learning experiences for students and challenge the status quo and deficit-based beliefs associated with teaching students with disability.

Building professional capacity

This area of the framework focuses on providing opportunities for school staff to learn together and create communities of practice. It is generally understood that education staff need to be confident, have a positive attitude, and have the knowledge and skills to teach all students in an inclusive approach (Cologon, 2019) and it is clear that professional capacity building and support can influence professional growth for teachers who are trying to be inclusive (Saggers et al., 2019). School principals provide leadership for the development of professional learning communities to support teacher learning (Hoppey and McLeskey, 2013).

Creating a supportive organization for learning

This domain is focused on the organization and considers how decisions are made about access and use of resources in the school. There is an expectation of collaboration and shared leadership that values diversity, supports high expectations, and builds a strong school culture. Behaviors that could be considered in this domain reflect relationships and how people in the school community work together.

Connecting with external partners

This domain includes three interrelated dimensions that are important for progressing inclusion: building positive relationships with families and community; engaging families and community in collaborative processes to support student learning; and ensuring that schools are the base of a community and that leaders and teachers can connect families with support when required.

DeMatthews et al. (2020) aligned the Hitt and Tucker framework with the work of school leaders of inclusive schools in the US and suggest this framework represents the field's most current understanding of effective leadership practices for inclusive education despite a lack of targeted focus on inclusion. The DeMatthews review (2020) uses the five domains from Hitt and Tucker (2016) to organize and present findings about inclusive school leadership literature in the US. The authors suggest that the five domains, which are interrelated, are consistent with the leadership practices used by principals in inclusive schools as they influence teacher capacity, school culture, and organizational conditions.

Neither Yukl (2012) nor Hitt and Tucker (2016) leadership taxonomies include leadership relation-oriented behaviors such as collaboration, support, and encouragement based on values and ethics (Daniëls et al., 2020).

Starratt (2014)

Starratt (2014) ethics of leadership—which incorporates care, justice, critique, and community—is a useful framework to support leaders to progress inclusive education (Harris et al., 2020). School leaders can pursue an ethic of care by valuing diversity, confronting stereotypes, and having high expectations for every child. The ethic of justice relates to authority and management within an education system, particularly in a school that considers equity and fairness. The ethic of critique is where school leaders look at their school in terms of structural justice and injustice to promote common good, such as ensuring high expectations of every member of the school community. Using the ethic of critique is essential to ensure an education system or school supports an ethical and inclusive culture.

Duignan (2012)

Duignan (2012) framework for analyzing ethical tensions supports school leaders to make choices in situations where there may be opposing values and ethical positions. The framework encourages a both/and mindset rather than an either/or mindset. Taking a both/and mindset encourages leaders to “look for the positives and the complementaries in these tensions and try to choose options that reflect a dynamic balance between considerations for both rules and caring” (Duignan, 2012: 82) by opening up situations to dialog through collaborative analysis, and multiple and mutually dependent relationships. This results in informed decisions and choices that consider the complexity of the dilemma. This framework is appropriate for leaders taking a critical and ethical stand because they are encouraged to look for multiple perspectives, perhaps challenge the status quo, and contest historical ways of thinking and practice. Duignan explains that this type of approach supports systems thinking (Senge, 1992) which helps the school leader to see the whole situation and the interrelationships and patterns of change.

Brown (2004)

Brown (2004) developed a framework to support a critical transformative approach for school leadership for social justice and equity. “A Pedagogy of Transformative Leaders” is a framework for professional learning for school leaders that highlights the context, theoretical understandings, and pedagogical strategies to support knowledge, skills, and desire to support equity (Brown, 2004: 79). Key theories such as critical theory (Giroux, 1983) and transformative learning (Mezirow, 1991) underlie the importance of facilitating critical reflection through dialog with consideration of local context required for a school leader to enact socially just change. Critical social theory drawing on the work of Freire (1994) supports a commitment to constant reflection and social change for a more equitable world. In the framework, these vertical theoretical threads are overlaid with horizontal threads focused on critical, transformative pedagogy to support learning. A transformative learning approach (Mezirow, 1991) drives critical examination of world views, values, and beliefs that are influenced by social and cultural contexts that drive goals and actions for enacting change.

Theoharis (2007)

Drawing on a qualitative study with seven school principals in the Midwestern states of America, Theoharis (2007) developed a three-pronged framework of resistance: (1) the resistance principals enact against historic marginalization of particular students; (2) the resistance principals face as a result of their social justice agenda; and (3) the resistance principals develop to sustain their social justice agenda in the face of resistance (Theoharis, 2007: 248). This framework can inform professional learning for aspiring school leaders. Theoharis (2007) suggests that the strategies described by the principals in his study and organized in his framework can be used to prepare leaders “to both enact and develop resistance in the face of significant barriers” (Theoharis, 2007: 249). He describes how the combination of knowledge and skills combined with the ability to facilitate critical reflection can guide future planning. He also provides some guidance for future leaders to develop their own resistance and resilience as they lead the change. These strategies acknowledge the ethical dilemmas faced by school leaders and support the need for a professional network to reduce the stress and burnout involved in this important leadership work.

School leaders who are driving socially just change will challenge and re-invent existing structures, practices, and relationships; however, there are few studies that report on ways that school leaders enact change for social justice and equity (Theoharis, 2007). The following section of this article reports on two studies that highlight how school leaders in Australia have enacted change.

School leaders enacting change: Australian studies**Study one: promoting equity in schools**

A team of academic researchers (Harris et al., 2018) worked with a network of six schools over three years to explore how school leaders worked with staff and students to develop more inclusive and equitable education. The school leaders were challenged in many ways to focus on equity and inclusion in a context of high-stakes testing and school choice which often focused on increased segregation and particular disadvantage for young people from minority backgrounds. This study built on international research that has established that inclusive schools are most successful in improving all students’ learning by establishing an inclusive school culture where staff, students, and parents’ voices are valued (Harris et al., 2018). The outcomes of this Australian study highlight how school leaders work in ethical ways, drawing on Starratt’s Framework of Ethical Leadership (2014). The resulting book—“Promoting Equity in Schools” (Harris et al., 2018)—provides detailed accounts of how school leadership facilitates the ethical, collaborative, and critical work with senior school leaders, teachers, and students.

One school project considered how using school performance data from the national testing system stimulated new ways of thinking to support greater attention to equitable student support. Critical discussions led by school leaders focused on questions such as: Why are things as they are? What type of support needs to be actioned to support more equitable outcomes for all students? School leaders in these schools interrogated the data with a collaborative team of school stakeholders that informed new ways of thinking about promoting change in relation to equity. Other inquiries in schools focused on teaching and learning projects within and across classrooms. Some school projects highlighted how students played active roles in collaborative inquiries and offered new

perspectives on teaching and learning. School leaders, who valued student perspectives, were able to interrupt existing assumptions that were not aligned with an equitable approach.

Further examples of a whole school approach to supporting equity and collaborations with other schools and the wider community also highlighted how collaborative inquiries can support school leaders to facilitate new ways of thinking and practice to support equity in schools (Harris et al., 2018). A critical and transformative approach is evident in the accounts of school leaders who developed an understanding of ethical challenges by gathering evidence to inform their decision-making using cycles of inquiry. This critical and transformative approach could facilitate school leaders to become more aware of how their leadership practice and behavior reflect the five domains in the Hitt and Tucker (2016) leadership framework and this could be important future research.

Study two: using the index for inclusion to promote equity

This study reports on the phases and outcomes of a whole school approach that utilized the Index for Inclusion (Booth and Ainscow, 2011) to promote a sense of belonging and connectedness for all students (Carrington et al., 2021). The Index for Inclusion (Booth and Ainscow, 2011) is informed by a values framework that guides “how we should live together” (Booth, 2011: 38) and supports school leaders, teachers, students, and families to reflect and have dialogs about how they can make their school a better place for all with a focus on inclusion, belonging, and connectedness. School leaders who want to promote equity and socially just change in their schools will find the Index for Inclusion provides a set of tools and resources to support a critical and transformative approach (Carrington et al., 2012) that aligns with Yukl’s fifth taxonomy missing element on ethical and socially responsible behaviors (2012). The process of using the Index for Inclusion in transformative leadership work enables school leaders to develop a values framework to support continuous school renewal that can lead to deeper and committed change and reform. The processes of engagement with the Index for Inclusion supports the development of a collective vision and action for change such as explained in the Hitt and Tucker framework (2016). Through questioning current values and practice or “the way we do things around here”, school leaders can guide teachers, parents, and students to create a more equitable and inclusive school.

The recent study reporting a whole-school approach in an Australian case study school (Carrington et al., 2021) highlights how school leadership can support a process of critical self-review and development by inviting stakeholders to have dialog and explore “other’s opinions and assumptions on the basis of equality, honesty and trust” (Booth and Ainscow, 2011: 19). The approach described in this case study explains how a school committee led cycles of planning, data collection, collaborative analysis and setting new school priorities about how students could be better supported in equitable ways. The Index provides a framework for sharing of perspectives and issues of concern between parent/community members, staff, school leaders, and students in a way that can facilitate deep learning and understanding about the values of inclusive education. In order to support use of the Index, a committee made up of school community stakeholders was established in the school to oversee the five phases of the Index: (i) getting started with the Index, (ii) finding out about the school, (iii) producing an inclusive school development plan, (iv) implementing priorities, and (v) reviewing actions and progress. This study highlights how leadership of a shared commitment to values that guides a process of critical review and development, using the Index for Inclusion, can support the development of greater awareness of equity and inclusive education.

Conclusion

If education systems expect school leaders to implement policy and create school cultures to support greater equity, then it is appropriate to consider current trends in education and the knowledge and skills needed to mobilize change in schools. Education reforms to improve equity in schools in the US have been unsuccessful at changing embedded practices that result in inequitable education. Reasons for this lack of success can be attributed to little understanding of the knowledge, skills, and attributes required for school leaders who are working in complex contexts with people who have multiple values and perspectives (Kozleski et al., 2020). Change for equity in schools takes time due to the critical reflection, dialog, and commitment to shared values of equity and inclusion required to support long-term actions for the greater good in schools. A critical and transformative approach to school leadership, as described in the Australian studies, will support school leaders to build relationships and work in partnership with students, parents, and educators in schools to initiate transformative reform and support equity in education. Future research should include a greater focus on the leadership taxonomies and frameworks described in this article. This future research could support a better understanding of how school leaders enact socially just change and support professional learning for school leaders.

Apple (2013) provides us with hope when he suggests that education can change society to supporting equity. We know that schools are communities made up of school leaders, teachers, students, and families and they all have obligations and roles to play in the school to ensure they are respected, included, and belong. However, it is the school leaders that lead the visioning and action to reframe school culture, policy, and practice. Cooper (2009) acknowledged that critical and transformative leadership to support equity is challenging but suggests that research indicates that it can be achieved with a collaborative approach with education stakeholders. The two Australian studies illustrate how school leaders can work in critical and transformative ways,

engaging in cycles of inquiry and collaboration with staff, students, and parents. The theoretical approaches and the taxonomies and frameworks described in this article can guide and support future and current school leaders to learn about leadership behaviors, attributes and values to support an appreciation of diversity and greater equity in schools.

References

- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012. *Developing Equitable Education Systems*. Routledge, New York.
- Ainscow, M., Harris, J., Carrington, S., 2018. Addressing the challenge of equity. In: Harris, J., Carrington, S., Ainscow, M. (Eds.), *Promoting Equity in Schools: Collaboration, Inquiry and Ethical Leadership*. Routledge, London, pp. 1–16.
- Apple, M., 2013. *Can Education Change Society?* Routledge, New York.
- AuCoin, A., Porter, G., Baker-Korotkov, K., 2020. New Brunswick's journey to inclusive education. *Prospects* 49, 313–328.
- Blackmore, J., 2002. Leadership for socially just schooling: more substance and less style in high risk, low trust times? *J. Sch. Leader.* 12, 198–222.
- Booth, T., Ainscow, M., 2011. *Index for Inclusion: Developing Learning and Participation in Schools*. Centre for Studies on Inclusive Education, Bristol.
- Booth, T., 2011. Curricula for the common school: what shall we tell our children? *Forum* 53, 31–48.
- Brown, M.E., Treviño, L.K., 2006. Ethical leadership: a review and future directions. *Leader. Q.* 17, 595–616.
- Brown, K.M., 2004. Leadership for social justice and equity: weaving a transformative framework and pedagogy. *Educ. Adm. Q.* 40, 77–108.
- Carrington, S., Saggars, B., 2021. An introduction to research in inclusive education: empirical evidence for supporting an inclusive approach for students on the autism spectrum. In: Carrington, S., Saggars, B., Harper-Hill, K., Whelan, M. (Eds.), *Research Approaches to Supporting Students on the Autism Spectrum in Inclusive Schools: Outcomes, Challenges and Impact*. Routledge, London.
- Carrington, S., Bourke, R., Dharan, V., 2012. Using the index for inclusion to develop inclusive school communities. In: Carrington, S., MacArthur, J. (Eds.), *Teaching in Inclusive School Communities*. John Wiley & Sons Australia Ltd, Brisbane, pp. 341–366.
- Carrington, S., Saggars, B., Shochet, I., et al., 2021. Researching a whole school approach to school connectedness. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2021.1878298>.
- Carrington, S., 1999. Inclusion needs a different school culture. *Int. J. Incl. Educ.* 3, 257–268.
- Cologon, K., 2019. *Towards Inclusive Education: A Necessary Process of Transformation*. Children and Young People with Disability Australia, Melbourne.
- Cooper, C.W., 2009. Performing cultural work in demographically changing schools: implications for expanding transformative leadership frameworks. *Educ. Adm. Q.* 45, 694–724.
- Cranston, N., Ehrich, L.C., Kimber, M.P., 2006. Ethical dilemmas: the bread and butter of educational leaders' lives. *J. Educ. Adm.* 44, 106–121.
- Daniëls, E., Hondégheem, A., Heystek, J., 2020. School leaders' and teachers' leadership perceptions: differences and similarities. *J. Educ. Adm.* 58, 645–660.
- Dantley, M., Tillman, L.C., 2006. Social justice and moral transformative leadership. In: Marshall, C., Oliva, M. (Eds.), *Leadership for Social Justice: Making Revolutions in Education*. Pearson Education, Boston, pp. 16–30.
- Day, C., Sammons, P., Hopkins, D., et al., 2009. *The Impact of School Leadership on Pupil Outcomes*. University of Nottingham, Nottingham.
- DeMatthews, D., Billingsley, B., McLeskey, J., Sharma, U., 2020. Principal leadership for students with disabilities in effective inclusive schools. *J. Educ. Adm.* 58, 539–554.
- Duignan, P., 2012. *Educational Leadership*, second ed. Cambridge University Press, New York.
- Freire, P., 1994. *Pedagogy of Hope*. Continuum, New York.
- Fullan, M., 2003. *The Moral Imperative of School Leadership*. Corwin Press, Thousand Oaks.
- Fullan, M., 2005. *Leadership and Sustainability: System Thinkers in Action*. Corwin Press, Thousand Oaks.
- Giroux, H., 1983. *Theory and Resistance in Education: A Pedagogy for the Opposition*. Bergin & Harvey, Westport.
- Gross, S.J., Shaw, K., Shapiro, J.P., 2003. Deconstructing accountability through the lens of democratic philosophies: toward a new analytic framework. *J. Res. Educ. Leadersh.* 1, 5–27.
- Harris, J., Carrington, S., Ainscow, M., 2018. *Promoting Equity in Schools*. Routledge, London.
- Harris, J., Ainscow, M., Carrington, S., Kimber, M., 2020. Developing inclusive school cultures through ethical practices. In: Graham, L. (Ed.), *Inclusive Education for the 21st Century: Theory, Policy and Practice*. Allen & Unwin, Sydney, pp. 247–265.
- Hehir, T., Grindal, T., Freeman, B., et al., 2016. *A Summary of the Evidence on Inclusive Education*. Alana Institute, São Paulo.
- Hitt, D., Tucker, P., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86, 531–569.
- Hoppey, D., McLeskey, J., 2013. A case study of principal leadership in an effective inclusive school. *J. Spec. Educ.* 46, 245–264.
- Karami-Akkary, R., Mahfouz, J., Mansour, S., 2019. Sustaining school-based improvement: considering emotional responses to change. *J. Educ. Adm.* 57, 50–67.
- Kidder, R.M., 1995. *How Good People Make Tough Choices: Resolving the Dilemmas of Ethical Living*. William Morrow, New York.
- Kozleski, E.B., Stepaniuk, I., Proffitt, W., 2020. Leading through a critical lens: the application of DisCrit in framing, implementing, and improving equity-driven, educational systems for all students. *J. Educ. Adm.* 58, 489–505.
- Mezirow, J., 1991. *Transformative Dimensions of Adult Learning*. Jossey-Bass, San Francisco.
- Mulford, B., Cranston, N., Ehrich, L., 2009. Australian school leadership today: conclusions. In: Cranston, N., Ehrich, L. (Eds.), *Australian School Leadership Today*. Australian Academic Press, Brisbane, pp. 417–425.
- Niesche, R., Gowllett, C., 2019. *Social, Critical and Political Theories for Educational Leadership*. Springer, Singapore.
- Riddle, S., Apple, M., 2019. Education and democracy in dangerous times. In: Riddle, S., Apple, M. (Eds.), *Reimagining Education for Democracy*. Routledge, New York, pp. 1–9.
- Rosenholtz, S., 1989. *Teachers' Workplace: The Social Organization of Schools*. Longman, New York.
- Rottmann, C., 2007. Leadership and change for social justice: mapping the conceptual terrain. *J. Educ. Adm. Found.* 18, 52–91.
- Ryan, J., Rottmann, C., 2007. Educational leadership and policy approaches to critical social justice. *J. Educ. Adm. Found.* 18, 9–23.
- Ryan, J., 2007. Inclusive leadership: a review. *J. Educ. Adm. Found.* 18, 92–125.
- Saggars, B., Tones, M., Dunne, J., Aberdein, R., 2019. Tele-classroom consultation: promoting an inclusive approach to supporting the needs of educators, families and early years learners on the autism spectrum in rural and remote areas in contextually responsive ways. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2019.1609103>.
- Sahlberg, P., 2012. How GERM is Infecting Schools Around the World? <https://pasisahlberg.com/text-test/>.
- Senge, P., 1992. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Random House, New York.
- Sergiovanni, T.J., 1992. *Moral Leadership: Getting to the Heart of School Improvement*. Jossey-Bass, San Francisco.
- Shields, C., 2014. Ethical leadership: a critical transformative approach. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 24–42.
- Skrtic, T.M., 1991. The special education paradox: equity as the way to excellence. *Harv. Educ. Rev.* 61, 148–206.
- Slee, R., 2018. *Inclusive Education isn't Dead, It Just Smells Funny*. Routledge, London.
- Starratt, R.J., 2004. *Ethical Leadership*. Jossey-Bass, San Francisco.
- Starratt, R.J., 2012. *Cultivating an Ethical School*. Routledge, New York.
- Starratt, R., 2014. The purpose of education. In: Branson, C.M., Gross, S.J. (Eds.), *Handbook of Ethical Educational Leadership*. Routledge, New York, pp. 43–69.

- Swaffield, S., Major, L., 2019. Inclusive educational leadership to establish a co-operative school cluster trust? Exploring perspectives and making links with leadership for learning. *Int. J. Incl. Educ.* 23, 1149–1163.
- Szumski, G., Smogorzewska, J., Karwowski, M., 2017. Academic achievement of students without special educational needs in inclusive classrooms: a meta-analysis. *Educ. Res. Rev.* 21, 33–54.
- Theoharis, G., 2007. Social justice educational leaders and resistance: toward a theory of social justice leadership. *Educ. Adm. Q.* 43, 221–258.
- United Nations (UN), 2006. Convention on the Rights of Persons with Disabilities. United Nations, New York.
- United Nations (UN), 2015. Sustainable Development Goals. United Nations, New York.
- Yasir, Y., Mohamad, N.A., 2016. Ethics and morality: comparing ethical leadership with servant, authentic and transformational leadership styles. *Int. Rev. Manag. Market.* 6, 210–216.
- Yukl, G., Mahsud, R., 2010. Why flexible, adaptive leadership is important. *Consult. Psychol. J.* 62, 81–93.
- Yukl, G., 2012. Effective leadership behavior: what we know and what questions need more attention. *Acad. Manag. Perspect.* 26, 66–85.

Leadership formation: in the smithy of the personal and professional

Ciaran Sugrue and John O'Sullivan, School of Education, UCD, Dublin, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	410
Leader formation: finding the “sweet spot” between fragmentation and coherence?	411
School leadership and identity politics?	411
The “character” of leadership: leader formation	412
Foucault’s concept of ethics	413
Moral codes and acts	413
Foucault’s key aspects of ethics	413
The character of leadership: ethical work as a basis for praxis	414
Standing for the common good	414
Values and accountability: professionally responsible leadership-formation	415
The crucible of leadership character formation	416
Concluding discussion	417
References	418

Introduction

According to John’s gospel (14:2), “in my father’s house are many mansions,” rather resonant with contemporary accounts of the multifarious sub-fields of leadership, its conceptions and practices. Nonetheless, in recent decades there has been a sustained effort, through various aspects of new public management (NPM), to reduce what counts in leading schools to being accountable for learning outcomes, frequently narrowly conceived. Increasing external policy prescriptions have serious consequences for the formation of school leaders, and, importantly also, for learners. Our focus in this chapter is on school leaders’ formation, and how this process, shaped significantly by external policy environments, has enormous consequences for leaders, who they have been and continue to become, while emphasizing also that, as research indicates, school leaders’ impact on the quality of teaching and learning, though indirect, is nonetheless, considerable (Day et al., 2016; Leithwood et al., 2008, 2020; Pont et al., 2008; Robinson et al., 2008). Thus, by paying focused attention to leader formation, particularly the dynamic interplay between the personal and professional, space to grow and flourish is extended, with positive consequences for all. In the Irish context, as schools reopened after Summer vacation, there have been calls for schools to cater for “messers, dreamers and misfits” (O’Toole, 2021, pp. 9–10) in addition to the mainstream, a provision that allegedly exists under the mantle of “Inclusion.” The author in question readily recognizes the necessity for rigor and standards but appeals also for space to accommodate and educate “the misfits and the dreamers and the people who don’t know what they’re doing until they stumble into it” (O’Toole, 2021). However, the “formation” of students cannot be divorced from the formation of teachers, principals, curricula and pedagogies (Leithwood et al., 2004). Thus, schools are the smithies wherein the formation of school leaders, teachers, and students alike, are crafted on the anvil where the personal and professional collide, subjected to the heat of the policy moment, immediate context, teachers and principal’s values and priorities, as well as the vicissitudes of our time.

What is formation? According to Sutphen and de Lange (2015, p. 412) “formation [is] ... the process of change that students undergo ... as well as ... faculty members ... as they teach” and as students learn. From a student perspective, formation is a broad church that encompasses: “the changes that students undergo in their outlook, values, behaviors, ideals, goals and interpretations of experiences in the course of being a ... student” (Sutphen and de Lange, 2015), while the term may be extended to embrace the life course, including the personal and professional lives of teachers and school leaders. Such a formative education extends well beyond performance as measured on a narrow range of learning outcomes, and embraces a moral dimension. It is legitimate to ask when focusing on contemporary education: to what extent do current external policy environments, curricula, pedagogies, and leadership practices, promote or constrain student, teacher and school leader formation as envisaged here?

In the increasingly fast paced globalized world, there are constant appeals for “leadership,” needed now more than ever, these clarion calls suggest, while others lament both the lack and apparent loss of leadership (Iacocca, 2008). However, such calls, no matter how well founded, are rendered somewhat impotent in a post-truth society, where: facts, alternative facts, emotion, reason, fueled by a pervasive social media and a news cycle that never sleeps become a miasma of confusion rather than clarity; an erosion of shared senses of “common good” where, it is argued: “The climate crisis, the global pandemic and its inevitable economic aftermath have revealed the extent to which our existing social contract is no longer working” (Shafik, 2021, p. xiv). Despite the “achievement gap” (Wagner, 2008), and recognition of its existence and increase (Piketty, 2014), there is general support for “the value of education” since it “reflects both the benefits to the individual and the gains to society as a whole of having an educated population”

(Shafik, 2021, p. 49). What is the nature of the social contract that cements the relationships between school leaders, teachers, their learners, and their communities? Arguably, all four volumes of this Encyclopedia are devoted in one way or another to responding to this enormous challenge. More modestly then, it is appropriate to ask: what aspect of this monumental task is to be addressed here?

The next section takes a narrative approach to extant literature, exploring the significance of values in leadership formation, with particular emphasis on the interplay between the personal and the professional. As a “tool” for unpacking this, the subsequent section presents a synthesis of Foucault’s notion of ethics, subsequently employed in the empirical section to explain and elaborate upon examples of leadership formation from the perspective of the lived realities of school leaders. The concluding discussion recapitulates the importance of the moral dimension of leadership, the values that shape it, while acknowledging the necessity for greater attention to this aspect of leader formation, and, when necessary, recognize the place of resistance in leadership when values conflicts arise.

Leader formation: finding the “sweet spot” between fragmentation and coherence?

In our pursuit of an elusive “spot” in an increasingly fragmented field of leadership literature, and the more mundane recognition that school principals and leaders have a responsibility to make decisions, and to act, often in challenging circumstances, we readily recognize that the contours of our journey are not terra incognita, but rather have been mapped in various ways by a variety of scholars. For example, the following clearly asserts the significance of values in the lives and work of school leaders:

In Day et al.’s (2001) study, they stress the importance of school leaders’ personal values informing their professional values and practices. Woods’ (2000) research reveals the extent to which school leaders’ values and reflexivity influence their actions and perceptions.

Campbell et al. (2003, p. 205).

What this evidence suggests is that recognizing the importance of values is significant, though not an end in itself. Rather, from a leadership formation perspective, values combined with reflexivity, have potential to contribute to the ongoing formation of leaders and teachers, working toward and contributing to a leadership praxis, that:

...implies a combination of reflection and action; not one or the other, but both. It seeks to hold in productive tension thought and action through critical reflection. ... it includes a moral dimension, informed by and imbued with one’s values, and beliefs that constitute a world-view.

Solbrekke and Sugrue (2020, p. 8).

Arising from this particular confluence of considerations there is recognition that values permeate principal’s “perception of their leadership role, their relationships with students, staff and the local community” (Campbell et al., 2003, p. 217). Consequently, there is sustained clamor in the literature that more attention be paid to values in principle preparation programmes (Cruz-Gonzalez et al., 2019), but also, and importantly, “throughout the professional life of principals” (Nordholm et al., 2020, p. 347) (see also, Angelle, 2017; Bezzina and Bufalino, 2019; Bird et al., 2013). However, it is important to add that attention to values is not merely a matter of articulating mission and vision statements that frequently reside in the limbo or no-man’s-land of school websites, but to recognize also that principals “are expected to be ‘moral leaders’” and since “some principals may not have sufficient access to skills to confront the ethical and critical challenges of their profession” appropriate pedagogical and leadership preparation is warranted (Cruz-Gonzalez et al., 2019, p. 325).

School leadership and identity politics?

Such a focus on values, however, and how they influence thought, deliberation and action, has become fraught and fragmented due to the exponential focus on “identity politics,” fueled in recent times by a combination of #Me Too, Black Lives Matter, and “group affinity” that is fired and facilitated by a variety of social media platforms, thus also a shift in emphasis to “group membership rather than individual rights and responsibilities” (D’Ancona, 2020, pp. 9–10). The reality of identity politics has seeped into leadership research, simultaneously providing both anchoring as well as complicating influences, the latter more disempowering rather than fueling a sense of agency. For example, Tricia, a black principal in Toronto, recognizes the significance of her own biography in the following:

I must be interested in social justice and equity because of who I am ... sometimes you’ll hear people say, “But that’s not important to me, I’m black, but I’m not affected.” But I think that’s nonsense (Tricia).

Armstrong and Mitchell (2017, p. 829).

For both principals in this Canadian study, there is recognition that there are inevitable tensions between jousting with and juggling different identities, simultaneously “authoring and effacing identity” but crucially also “standing up and standing tall for personal and professional values” (Armstrong and Mitchell, 2017). In the context of this chapter, these principals recognize the dynamic between what is identified as the “public persona” (professional) and the “shadow persona” (personal), an ongoing tension between the desire to “fit in” while simultaneously wanting to be allowed the space and opportunity to “be myself” (Armstrong and Mitchell, 2017, p. 835). In similar vein, others too have claimed “it is important simultaneously to recognize the impossibility of escaping the personal and the situational” (Sugrue, 2015a, p. 15). Consequently, recognition of the increasing complexity of the principal’s role has spawned “multiple identities” (embracing “problem solver, guardian of culture, change agent, advocate, or mediator”) (Notman, 2017, p. 760), while such shattering of the self has encouraged others to argue that in such circumstances, there is a desire if not a craving on the part of school leaders to “create in their relationships with others ... a sense of coherence, worth and belonging” (Lumby and English, 2009, p. 95). Added to this complexity, a vein of research on leader emotions, while important, has potential also to create further fragmentation and vulnerability through greater exposure due to social media (Gavish and Oplatka, 2012). In recognition of the demands of the role creating an emotional roller coaster, efforts have been made to find some “situated certainty” (Hargreaves, 1994) in such fractured circumstances—resulting in “the PIVOT framework—Purpose, Identity, Values, Options and Trust” as a means of explaining “current leaders’ practice and development” (Cain and Gunter, 2012, p. 187). At the risk of being accused of a reductionist perspective, or an overly simplistic approach to leader formation, we argue that the fundamental pivot between the professional and the personal results from a manifestation of “moral integration” (Coles, 2000) through which a degree of continuity and coherence is found; finding “legitimate compromise” between competing priorities and values (May 1992, 1996).

When the discrepancy between personal values grate and grind against the values embedded in policy prescriptions that fail to recognize or accommodate particular school and community contexts, there is increasing evidence that due to the changing nature and increasing complexity of the principal’s role, there is growing attrition, and, depending on its frequency, with negative consequence for community, continuity and stability, as well as student learning and achievement (Mascall and Leithwood, 2010). Similarly, when framed within a milieu of “school improvement” and “policy enactment”, as concluded in the following, these concerns cannot be separated from “connection with the values and biographies of the principals and actors at all levels of the system” (Day and Gu, 2018, p. 342). Additionally, further research confirms this intimate relationship, when it states: “school principals, ...as educators” not only require a “strong moral purpose” but that such a set of values is both “a critical antecedent and co-requisite” of their capacity for effective practices” (Crow et al., 2017, p. 265).

The “character” of leadership: leader formation

In order to advance research, as well as the argument of this chapter, and in the interest of being inclusive and comprehensive in our approach to leader formation, we understand that multiple identities, the responsibility to care, trust and challenge, often simultaneously, is the ongoing “crucible” (Bennis and Thomas, 2002) in which leaders find themselves, not merely as an entré or initiation into the principal’s role; closer to Dante’s *Inferno*, when there are days that seem like real “purgatory,” while others harbor significant psychic rewards, and though not entirely “paradise,” as Dante would have it, render the role worthwhile and rewarding. It is such experiences that re-form (repeatedly) the character of leaders. When as academics we are invited to provide references in support of a colleague’s application, typically for a promotion, in addition to their capacities, and accomplishments, almost invariably there is a request to comment on their “character,” their communication skills, disposition, and evidence of collaboration, given the increasing significance of the latter for leadership (Leithwood et al., 2019). While cumulative evidence over time regarding “successful school leadership” has identified significant characteristics of “good” leaders (Leithwood et al., 2006) what is not predictable is how such qualities become embedded in the leader; that requires the ongoing work in the crucible where personal and professional contexts continuously encounter one another. This work extends over a rather wide spectrum of activities and engagements for teachers and leaders, embracing:

... professional and personal–professional activities beyond the immediate pedagogical or administrative. From formal and sustained programmes of higher degree study and certified participation to decidedly informal and oftentimes fleeting undertakings increasingly pursued in the unregulated environment online.

O’Brien (2018, p. 229).

Across this range of activities there are varying degrees of an “ethical relation” as well as a “relation of power or government” often articulated in policy texts (O’Brien, 2018, p. 230). In informal settings, individuals have the courage to speak freely and uninhibitedly, while in formal gatherings, the requirement for self-government is felt more keenly. When the external environment, the immediate context of the school and/or national policy, is experienced as demanding or even coercive, governance of the self becomes considerably more circumspect. In such circumstances, the “enfolding” of self-governance may induce compliance and conformity rather than, the courage to “stand tall” as articulated above. The notion of “the fold” or the process of enfolding, where self (personal) and governance (professional context) continually meet is what O’Brien (2018, p. 231) identifies as “The pocket of ‘interiority’ [that] ... produces ... the zone of the self, simultaneously ‘inside’ and ‘outside’”. In this zone, the self, and what it stands for is continuously remade as it encounters the realities of professional life within the school community, the dynamics of which he

captures in the following: “practices of the self and practices of government might be better regarded as hybrid ones and might, more accurately, be described as ethical-governmental practices” (O’Brien, 2018, p. 232). However, maintaining the ethical element of this tension, necessitates moving beyond the language and logic of accountability, and to reach for the language and logic of professional responsibility which necessitates a moral dimension (Solbrekke and Sugrue, 2011). Thus, in order to vindicate one’s personal values, as they become enfolded in governmental requirements, as part of the process of re-constituting the personal and professional, it is necessary to seek out legitimate compromise (May, 1992). A note of caution however is warranted since research demonstrates the autonomous ‘new professional’ is subjectified to embody a neoliberal subject in practices of self-formation (Anderson and Cohen, 2015; Cohen, 2014; Gobby, 2013). Consequently, to vindicate one’s professional responsibility, resistance rather than compliance to governance requirements may feature prominently in the process of enfolding, thereby enabling the evolving self to embody moral considerations. Accordingly, it is timely to turn attention to a notion of “ethical work” (Foucault, 2000b) that is particularly apposite to explicate such a processes.

Foucault’s concept of ethics

Foucault’s work has proved beneficial in the area of educational leadership since it provides insight into ways in which school leaders are subjected to (Mifsud, 2016, 2017; Niesche, 2011), and to a lesser extent resist (Dolan, 2020; Niesche, 2011), the influence of discourses emanating from NPM. Consequently, the use of his oeuvre in research pertaining to educational leadership has become “an established approach” (Heffernan, 2018, p. 511). Generally speaking, the term ethics is used to refer “to well-founded standards of right and wrong that prescribe what humans ought to do” (Velasquez et al., 2013). Foucault, however, advanced a concept of ethics that centers upon the relationship that a person has with oneself to behave in a moral manner. This latter notion of ethics provides a useful means of illuminating the interplay between the personal and professional in the enfolding that underpins the character of leadership. To explicate this further, we focus on Foucault’s notion of ethics, its relationship with moral codes and acts.

Moral codes and acts

Foucault uses the term moral codes to refer to sets of prescriptions that determine which behaviors of people are permitted and forbidden and ascertain the positive or negative value of different possible behaviors (Foucault, 1977, 2000b). For example, the idea that people are not allowed to commit adultery is an element (prescription) of a moral code. In effect, moral codes are sets “of values and rules of action that are prescribed by agencies such as the family, educational institutions and churches” (Niesche, 2011, p. 37). The Bible, Koran and educational leadership policy documents typify texts that set out moral prescriptions. In view of this, ethics can be seen as originating from an individual’s interpretation of such moral codes.

Ethics is concerned with this relationship an individual has with him/herself in terms of interpreting how one should behave to consider themselves as acting in a moral fashion (Foucault, 1992, 2000b). Additionally, Foucault uses the word acts to denote, for example, the way a principal determines how the code concerning distributed leadership should apply to himself or herself culminating in his/her actual behavior in cultivating it, or not. Hence, ethics can be viewed as the process through which an individual determines the physical enactment of their philosophical code.

Foucault’s key aspects of ethics

Foucault suggests that this process of ethical conduct has four major aspects (Foucault, 1992; 2000b). One of these he terms “ethical substance” which can refer to, for example, a person’s acts, desire or feelings (Foucault, 2000b, p. 263). It denotes the part of an individual that they themselves must refine as the “prime material” of their moral conduct (Foucault, 1992). For example, with the ancient Greeks it was moral for a man to be in love with a boy so long as he did not touch him, whereas for early Christians, it was immoral to even desire a boy in such a manner (Foucault, 2000b). Hence, the ethical substance that informed the ethical judgment of the ancient Greeks concerned acts while for early Christians it centered upon desire. In a complex, relativist plurality of values faced by contemporary school leaders, the character of their leadership is forged and burnished through finding legitimate compromise between their values and those inherent in policy mandates, in the absence of which resistance becomes a moral obligation.

Foucault also draws attention to what he calls “mode of subjectivation ... the way in which people are invited or incited to recognize their moral obligations” (Foucault, 2000b, p. 264). It concerns the way in which an individual establishes their relation to a rule and recognizes himself or herself as obliged to implement it (Foucault, 1992). It alludes to the factors that influence individuals to accept and comply with their moral obligations. It could refer to divine law in a text such as the Bible or Koran (Niesche, 2011; O’Farrell, 2005), policy documents by an education authority (Niesche and Haase, 2012), to make one’s life an example to others, to be a role model (Foucault, 1992; O’Farrell, 2005) or criteria inspectors use to judge schools (Gillies, 2013). For example, a principal might advance a policy reform initiative to enhance the educational experiences of students, while another might do so to receive favorable rank from the Inspectorate; the latter being a less worthy value espousal, although Foucault avers that one’s mode of subjectivation is premised upon “a personal choice” (Foucault, 2000b, p. 266).

An additional aspect which Foucault refers to as “self-forming activity” relates to the means by which a person attempts to change themselves in order to behave in a moral fashion to achieve their ethical aspiration (Foucault, 2000b, p. 265). Such steps could include physical and mental techniques such as fasting or only eating certain types of food in the area of diet (O’Farrell, 2005), principals undertaking administrative work outside of school hours (Niesche, 2011) or school leaders participating in Continuing Professional Development (CPD) activities (Gillies, 2013).

Yet another dimension of this ongoing character formation or subjectification/exercise of agency for Foucault he terms “telos”, which is “the kind of being to which we aspire when we behave in a moral way;” an individual’s idealized notion of themselves as a moral person (Foucault, 2000b, p. 265). It relates to achieving “a certain mode of being, a mode of being characteristic of the ethical subject” (Foucault, 1992, p. 28).

The empirical section of this chapter indicates and illustrates how senior school leaders bring together the public and shadow persona, forging legitimate compromises as integral to the ongoing development of their ethical character.

The character of leadership: ethical work as a basis for praxis

The empirical data reported here draws selectively on a multi-site case study conducted across 12 second level schools in Ireland. While this research embraced interviews with senior school leaders - principals and deputy principals - and teachers at each of these schools, illustrative data are presented to showcase different dimensions of the leadership character formation of three school leaders in specific professional contexts. Notably, the associated national context, emulating those of a variety of jurisdictions across the globe (Apple, 2006, 2013, 2014; Mac Ruairc, 2010), is characterized in significant ways by the influence NPM agendas (Ball, 2016; Sugrue, 2015b). What follows is an exploration of how values of senior school leaders and the professional contexts in which they are situated meet, are juxtaposed, yet coalesce as constituents of their emerging “ethical substance”, that is their leadership formation (Foucault, 2000b).

Standing for the common good

The first example of enfolded transpired at a fee charging school, one at which the focus on student attainment in the government Leaving Certificate examination (end of secondary education) was particularly pronounced. It centered upon a letter from parents that the deputy principal there was asked to endorse, which sought an “accommodation”¹ for their son for said examination. To give a “flavor” of the kind of things he was “dealing with on the hoof” (Deputy principal, School A), this deputy principal mentioned that while he was on the corridor earlier that morning, the career guidance teacher had shown him a letter from the parents in question seeking the aforementioned accommodation. Regarding this letter the deputy principal stated:

... they wrote things about [their child] being a victim of bullying for six years and so on which are categorically untrue... To be quite blunt they’re lying in the hope of getting an accommodation for him... I understand that any parent will kind of try anything to get what they think will help their child, but we can’t allow them to say these things. I mean if they’re true, fine, let’s deal with them... To my mind it is clearly an attempt just to use sympathy and so on but we can’t stand over that. We certainly can’t endorse what’s in our opinion a complete lie. You know it’s wrong... I wouldn’t turn a blind eye and say, ah he’s a good kid and if this helps him get it.

Deputy principal, School A.

This citation illustrates the complexity of the, frequently fleeting (O’Brien, 2018), competing and conflicting personal and professional demands that are imposed on school leaders by both themselves and the professional contexts in which they enact their roles. It confirms that the parents in question had submitted a dubious letter for the school to endorse possible enhancement of their son’s leaving certificate results and in so doing were “playing a sympathy card” to influence the deputy principal to achieve this. Moreover, it demonstrates that he was able to empathize with the parents’ desire to support their child, whilst also being aware that a person in his position could be swayed in the matter to help a “good” student. Hence, the decision-making process for the deputy principal did not just center on whether, or not, the assertions being made in the letter were legitimate, as other potentially mitigating factors were at play.

In the above instance the deputy principal resisted the temptation of succumbing to the pressure from the parents, which bodes well for the character of his leadership. To understand how and why this enfolded as it did, Foucault’s notion of ethics is most useful. The citation below is pertinent as it clarifies associated elements of the deputy principal’s “telos” (Foucault, 2000b) of himself as a good leader and his “mode of subjectivation” (Foucault, 2000b) - which in general terms alludes to

¹An “accommodation” refers to additional support that may be granted by the Ministry of Education to facilitate access to the certificate examinations by students who would have difficulty in accessing the examination and/or communicating what they know and/or can do due to an examiner consequent to a physical and/or visual and/or learning difficulty.

the factors that influence individuals to accept and honor or live up to their moral obligations – in choosing the leadership response he did, over possible alternatives. In referring to the possibility of endorsing the aforementioned letter from parents, the deputy principal stated:

Legally we could even be vulnerable to a case later; ... you endorsed it [the letter] when you sent it off to the Department. You basically accepted it's true... If we sent that off now in the envelope today and pretend we didn't see it or collude with it, the chances are it would never come back and bite us. The boy might get his "accommodation" or it might be to his benefit, but it's wrong and also it would disadvantage let's say a girl in another school who didn't get a course because he got an accommodation he didn't deserve under false pretences. So, you know, morally I couldn't stand over it. I'm not making myself out to be the most moral person in the world but there has to be some kind of sense that this is acceptable and that's not acceptable.

Deputy principal, School A.

When this quotation is considered in conjunction with this deputy principal's previous one, it is evident that he knew that if the school were to endorse the letter it would be validating its assertions, including one that the student in question was a victim of bullying at the school for six years, and could therefore have a legal case taken against it. The above citation also confirms that the deputy principal was aware that he could endorse the letter to the benefit of the student and that this course of action would be unlikely to culminate in adverse repercussions for the school. However, while he was aware of these scenarios, they were not crucial in guiding his decision. On the contrary, it can be deduced from his statement, "there has to be some kind of sense that this is acceptable and that's not acceptable" that it was premised on his *telos* that good leadership must reflect certain moral standards. Specifically, it was his view that, simply put, "it's wrong" to endorse false claims for an accommodation and that to do so could result in inequalities in access to further education for students elsewhere in the country.

The deputy principal's leadership formation (ethical substance), which resulted in his decision not to endorse the letter, was contingent upon his personal values that a good educational leader (*telos*) must embody certain standards of right as opposed to wrong and be fair to students across schools in terms of progression to further education (mode of subjectivation). As such, in engaging in "ethical work" (Foucault, 2000b) his personal values, consistent with extant research (Campbell et al., 2003, p. 205), played a crucial role in his leadership formation, evidence of professionally responsible leadership praxis combining values and reflexivity (Solbrekke and Sugrue, 2020). In the face of conflict between his personal values and professional context, he stood up and stood tall (Armstrong and Mitchell, 2017) for the common good.

Values and accountability: professionally responsible leadership-formation

A second example of discord between the values of a senior school leader and professional context manifest itself at School B, where parents, by virtue of the privilege they ascribe to having their children taught by excellent teachers, sought to encroach on the potential appointment of a teacher for secondment to a non-teaching position at the school. This conflicted with the principal's moral aspirations regarding such appointments. Foucault's key aspects of ethics are useful in illuminating the key factors that influenced his associated leadership formation.

A key feature of the principal at School B's "telos" (Foucault, 2000b) of a good principal incorporated the notion that teachers should be appointed to positions on merit. This can be deduced from the rationale he articulated for selecting a teacher for appointment to a post of responsibility (a middle management role) to promote academic excellence at the school. He stated that the teacher in question

... is very enthusiastic about ... promoting academic achievement... and ... I know it's something she's passionate about. She got a prize relating to her subject a number of years ago for the top leaving cert result in the country in a particular year. The top A1 was one of her students ... and she's the ideal person. She's also a very senior member of staff, well respected! I certainly feel that staff members would take her in that role very seriously.

Principal, School B.

This indicates clearly that the principal had selected the teacher in question to undertake the role of promoting academic excellence as she had exemplified characteristics that suggest she would undertake it competently. Thus, it can be appreciated that appointing teachers to positions on merit formed part of his *telos* of a good principal. However, the professional context that arose consequent to the perspectives of parents regarding another appointment at his school was at variance with his meritocratic values.

The issue with this particular appointment arose when several teachers at School B applied for secondment to this non-teaching position, including one whom the principal described as "a particularly strong teacher ... The teacher would probably be the best teacher in the school. Absolutely, and ... we have some very talented ones, so absolutely excellent" (Principal, School B). He recognized that if this teacher were to be appointed to the seconded position it would be, as he put it, "a huge loss from the teaching staff ... and [I] knew then of course there was going to be a backlash from parents" (Principal, School B). This citation indicates that parents at this school would have preferred if the excellent teacher in question was not appointed to the seconded position, a perspective, as evidenced by the interview data, underpinned by their desire to maximize their children's leaving certificate results. The quotation also confirms that if the excellent teacher was appointed to the position, the specter of reverberations from parents

loomed large. Thus, the principal's meritocratic values in respect of appointments were in direct conflict with pressure emanating from parents to privilege their children's interests.

In connection with the above dilemma, the principal clarified his associated "mode of subjectivation" (Foucault, 2000b) - the factors that incited him to "stand by" his moral obligations - in stating:

I really had to make a decision that I couldn't allow, and I was tempted, but I couldn't allow that teacher to be penalised almost for being exceptional and that's what it would have amounted to if that teacher hadn't got that role. She would have been penalised because she was just so good and so committed ... so she was appointed to the position.

Principal, School B.

This highlights the gravity of the situation for, not just the excellent teacher and the principal's leadership formation as the latter "was tempted" not to appoint the former to the position but, also, regarding the positive impact she could have as an excellent teacher in cultivating excellence among colleagues, thus raising standards across the school. Additionally, it reveals the principal's mode of subjectivation centered on resisting the influence of parents as this would have involved short-changing a teacher for excelling, a values position he was unwilling to be compromised on. Crucially, evoking his mode of subjectivation in this way was premised upon a "a personal choice" (Foucault, 2000b, p. 266) to honor his meritocratic values while resisting being "disciplined" (Niesche, 2011) by the influence of parents. His shadow persona influenced his professional decision.

Acting in a professionally responsible manner played a significant role in enabling the principal subsequently to defend the moral integrity of the appointment process as there was a backlash from parents, whom in his words were "very concerned and questioned me" (Principal, School B). However, he had engaged in a "self-forming activity" (Foucault, 2000b) - steps taken to conduct himself in a moral fashion - which enabled him to respond to them and achieve his ethical aspiration, as illustrated:

My response to them at the Parents' Association meeting was ... that I advertised, I interviewed, and I sourced an excellent ... very highly qualified teacher to fill that vacancy and the teacher in question didn't relinquish her teaching duties until the newly appointed teacher started and there was an appropriate hand over. And I feel I am very well covered ... but that's an element of accountability.

Principal, School B.

It suggests that the appointment process complied with relevant policy requirements, an accomplished teacher had been appointed to replace the one being seconded and there had been an appropriate "hand over", thereby alleviating potential concerns parents might have in respect of their children's academic attainment. Thus, through his "ethical work" (Foucault, 2000b) premised upon upholding meritocratic values (telos) in a professional context that privileges measurable outcomes of students, conducting the appointment process, not merely in an accountable but a professionally responsible manner (self-forming activity), consistent with his personal values-based choice (mode of subjectivation), the principal's leadership character (ethical substance) was enfolded (O'Brien, 2018).

Consistent with extant research (Crow et al., 2017), personal values-based choices are pertinent to leadership formation in the contemporary era, whereby senior school leaders must choose between their moral values and values inherent to educational environments characterized by NPM influences. Moreover, it is evident that in such predicaments they can exercise a "degree of freedom" in determining their acts (Foucault, 2000a), even fusing personal values and accountability, thus acting in a professionally responsible manner while simultaneously cultivating the moral character of leadership.

The crucible of leadership character formation

The values of senior school leaders pertaining to inclusive education are concerned with both the academic and social development of all students and enable us to shine additional light on leader's character formation (Ainscow et al., 2019). Inclusion policies validate the plurality of difference within and across students (Booth and Ainscow, 2016) but contexts vary in their receptivity to such policy prescriptions, as the following example illustrates. As indicated by the deputy principal in School C, the progress of first year students was of concern due to student behavior in the associated "mixed ability classes", with the result that "teachers were overwhelmed by this situation" (Deputy principal, School C). Consequently, to address the situation for future years she led an initiative that culminated in the decision that, thereafter, for first year classes the "top two classes will be streamed and the bottom two classes will have mixed ability" (Deputy principal, School C). Further, she indicated that for a group of incoming first year students whose conduct reflected the misbehavior of their current first year cohort, a "transition program" had been developed whereby such students will have their own classroom and experience an alternative curriculum tailored specifically for them. Hence, the deputy principal was responsible for leading change from allocating students, upon entering the school, to "mixed ability" class groupings to employing hierarchical arrangements in the form of "banding" (Ireson and Hallam, 1999). Research suggests the use of such hierarchical approaches have adverse effects for some students thus run counter to the philosophy of inclusive education (Boaler et al., 2000; Gillborn and Youdell, 2001; Ireson and Hallam, 2009; Ireson et al., 2005; McGillicuddy and Devine, 2020; Smyth, 2018). The application of Foucault's concept of ethics that follows reveals that, even if it fell somewhat short of an inferno, the

associated enfoldings of the self (O'Brien, 2018) placed the deputy principal's values at odds with her leadership formation in a no-win situation in an era of austerity.

Empirical data indicate that the Deputy Principal at School C's "telos" (Foucault, 2000b) of herself as a good deputy principal incorporated the notion she should promote inclusive education. Indeed, she accorded this facet of her telos a high priority as evidenced by her statement, "one of the things that I'm most proud of" is the school is catering for students "from all sorts of diverse backgrounds" (Deputy principal, School C). Consequently, her "ethical substance" (Foucault, 2000b), the prime part of herself that she must refine regarding fulfilling the above facet of her telos, is her leadership formation. Thus, at a cursory glance it appears that she ought to have promoted the use of mixed ability as opposed to banding as the basis for forming first year classes to consider herself to be acting in a moral way. However, further examination of her engagement with ethics verifies this as an overly simplistic interpretation of the situation, as her leadership formation was complicated by other factors.

In respect of the Ministry's policy promoting inclusive education, she stated:

I'm one hundred percent supportive of it but sometimes it's not achievable because we don't have the resources to reach those ideals... I think, for example, our mixed ability might be a much more pleasant experience for students and for teachers ... if we had more resources for our students in terms of attaining their educational potential, the best that they can do, with extra trained people in a classroom. Then, I think that would really help us live the idea of an inclusive school.

Deputy principal, School C.

While the above confirms the deputy principal's endorsement of inclusive education policy and her commitment to having it reflected in the lived reality at her school, it also asserts that from an organizational perspective, the formation of classes based on mixed ability is, in and of itself, insufficient for students to experience inclusive education within such classes, as this can fall short in terms of the pleasantness of the associated educational experiences for both students and teachers. Indeed, the educational outcomes of students would undoubtedly suffer too when, as previously cited, teachers are "overwhelmed" in such settings. Accordingly, the realization of inclusive education within mixed ability classes for first year students at School C would have been contingent upon the allocation of additional resources to the school for team teaching within such classes. Notably, at the time the school had lost significant personnel and middle management resources arising from cutbacks in education at national level (Drudy, 2009). Thus, in the professional context in question, realizing the deputy principal's ethical ambition was not as simplistic as choosing between operationalizing mixed ability that would enable students to experience inclusive education or employing banding that would militate against its actualisation.

On the contrary, on the one hand the character of the deputy principal's leadership could have embraced the continued use of mixed ability as the basis for class formation. The associated "mode of subjectivation" (Foucault, 2000b) would be consistent with the deputy principal's ethical aspirations in respect of inclusive education from an organizational perspective but would, due to a lack of resources, fall short of enabling students to achieve appropriate levels of educational attainment in a suitable environment. On the other hand, the deputy principal's leadership formation could have been characterized, as it was, by the introduction of banding. Thus, her "mode of subjectivation" (Foucault, 2000b) conflicted with her ethical ideals from a class formation perspective due to concomitant adverse social (Smyth, 2018) and psychological effects for students (Ireson and Hallam, 2009; McGillicuddy and Devine, 2020), while it would be more likely to yield higher levels of academic attainment for some students (Ireson et al., 2005; McGillicuddy and Devine, 2020). The decision taken in the circumstances might not be regarded by all readers as a "legitimate compromise" (May 1992) between the deputy principal's endorsement of inclusion as a policy and the possible negative consequences of her decision for some learners, but in the crucible between choosing between the wellbeing of her teaching colleagues, the progress of some learners or adverse effects for both teachers and the entire student cohort, resistance to the policy in the circumstances seemed to her legitimate. In a time of austerity, securing additional resources was not an option. However, values conflicts sustained over time contribute to "corrosion of character" and burnout (Sennett, 1998).

Concluding discussion

The character of leadership is not just something "out there" to be aspired to by leaders, described in mission statements but otherwise largely ignored. What this chapter has endeavored to demonstrate is that the character of leadership is embodied, it is enfolded into the public and shadow persona of leaders, their ongoing formation. The particular examples we provide as illustrative of this ethical process reinforces the view that leading has an inescapable moral dimension. This indicates the extent to which the moral fabric of education systems is contingent upon the "ethical work" (Foucault, 2000b) of senior leaders in particular, but extends to all members of the school community. This ethical imperative is supported by extant literature that highlights the importance of ensuring that ethical leadership formation receives due recognition in leadership preparation programmes (Cruz-Gonzalez et al., 2019).

Moreover, ethical decision-making requires a degree of professional autonomy to choose one action over another since "freedom is the ontological condition of ethics" (Foucault, 2000a, p. 284). However, due to the complexity of the principal's role, the policy constraints and contexts in which they are obliged to lead, frequently necessitates "legitimate compromise" (May 1992) that includes reflexive consideration of the way s/he has decided to act. The ethical considerations indicated in this chapter

are deployed by senior school leaders as they enact their responsibilities. They are part of their deliberations, and reflections as integral to their character formation, while seeking to act in an ethically appropriate manner. The enfolding of values into character formation in the everyday lives of school leaders evident above, lends legitimacy to resistance to those policies that violate the personal and professional sense of self, and where legitimate compromise remains out of reach. Rather, it calls for a more “deliberative leadership” that recognizes both the shadow and substance of leadership practice, while, through reflexivity seeks to elevate it to a sustainable praxis (Sugrue and Solbrekke, 2020).

References

- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676. <https://doi.org/10.1080/13603116.2019.1622800>.
- Anderson, G., Cohen, M.I., 2015. Redesigning the identities of teachers and leaders: a framework for studying new professionalism and educator resistance. *Educ. Pol. Anal. Arch.* 23 (85).
- Angelle, P.S., 2017. Leading authentically: a new principal in challenging circumstances. *Research in Educational Administration & Leadership* 2 (1), 10–27.
- Apple, M.W., 2006. *Educating the “right” Way: Markets, Standards, God and Inequality*, 2nd ed. Routledge Taylor & Francis Group, New York, NY.
- Apple, M.W., 2013. *Can Education Change Society?* Routledge, New York, NY.
- Apple, M.W., 2014. Official knowledge: democratic education in a conservative age. In: *Official Knowledge: Democratic Education in a Conservative Age*, 3rd ed., pp. 1–233. <https://doi.org/10.4324/9780203814383>.
- Armstrong, D., Mitchell, C., 2017. Shifting identities: negotiating intersections of race and gender in Canadian administrative contexts. *Educ. Manag. Adm. Leader* 45 (5), 825–841. <https://doi.org/10.1177/1741143217712721>.
- Ball, S.J., 2016. Neoliberal education? Confronting the slouching beast. *Pol. Futures Educ. Internet* 14 (8), 1046–1059. <https://doi.org/10.1177/1478210316664259>.
- Bennis, W.G., Thomas, R.J., 2002. *Geeks & Geezers How Era, Values, and Defining Moments Shape Leaders*. Harvard Business School Press, Cambridge, MA.
- Bezzina, C., Bufalino, G., 2019. Nurturing authentic leadership for teacher leaders: the challenges ahead. *Kappa Delta Pi Rec.* 55 (1), 18–23.
- Bird, J.J., Dunaway, D.M., Hancock, D.R., Wang, C., 2013. The superintendent’s leadership role in school improvement: relationships between authenticity and best practices. *Leader. Pol. Sch.* 12 (1), 37–59.
- Boaler, J., William, D., Brown, M., 2000. Students’ Experiences of Ability Grouping – disaffection, polarisation and the construction of failure. *Br. Educ. Res. J.* 26 (5), 631–648. <https://doi.org/10.1080/01411920020007832>.
- Booth, T., Ainscow, M., 2016. *Index for Inclusion: Guide to School Development Led by Inclusive Values*. Centre for Studies in Inclusive Education, Bristol.
- Cain, M.E., Gunter, H., 2012. An investigation into primary school leadership practice: introducing the PIVOT framework of leadership. *Manag. Educ.* 26 (4), 187–191. <https://doi.org/10.1177/0892020612447247>.
- Campbell, C., Gold, A., Lunt, I., 2003. Articulating leadership values in action: conversations with school leaders. *Int. J. Lead. Educ.* 6 (3), 203–221. <https://doi.org/10.1080/136031203200090064>.
- Cohen, M.I., 2014. “In the back of our minds always”: reflexivity as resistance for the performing principal. *Int. J. Leader. Educ.* 17 (1), 1–22. <https://doi.org/10.1080/13603124.2013.804208>.
- Coles, R., 2000. *Lives of Moral Leadership Men and Women Who Have Made a Difference*. New York Random House.
- Crow, G., Day, C., Moller, J., 2017. Framing research on school principals’ identity. *Int. J. Leader. Educ.* 20 (3), 265–277.
- Cruz-Gonzalez, C., Segovia, S.D., Rodriguez, C.L., 2019. School principals and leadership identity: a thematic exploration of the literature. *Educ. Res.* 61 (3), 319–336. <https://doi.org/10.1080/00131881.2019.1633941>.
- D’Ancona, M., 2020. *Identity, Ignorance, Innovation*. Hodder & Stoughton, London.
- Day, C., Gu, Q., 2018. How successful secondary school principals in England respond to policy reforms: the influence of biography. *Leader. Pol. Sch.* 17 (3), 332–344.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258. <https://doi.org/10.1177/0013161X15616863>.
- Day, C., Harris, A., Hadfield, M., 2001. Challenging the orthodoxy of effective school leadership. *Int. J. Leader. Educ.* 4 (1), 39–56. <https://doi.org/10.1080/13603120010010615>.
- Dolan, C., 2020. *Paradox and the school leader: The struggle for the soul of the principal in neoliberal times*. Springer.
- Drudy, S., 2009. Education and the knowledge economy: a challenge for Ireland in changing times. In: Drudy, S. (Ed.), *Education in Ireland: Challenge and Change*. Gill and Macmillan, Dublin.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison* (A. Sheridan, Trans.). Penguin, London.
- Foucault, M., 1992 (R. Hurley, Trans.). *The Use of Pleasure: The History of Sexuality*, vol. 2. Penguin, London.
- Foucault, M., 2000a. The ethics of the concern for self as a practice of freedom. In: Rabinow, P. (Ed.), *Ethics, subjectivity and truth: Essential works of Foucault 1954–1984*, Vol 1. Penguin, London, pp. 281–301.
- Foucault, M., 2000b. On the genealogy of ethics: an overview of work in progress. In: Rabinow, P. (Ed.), *Ethics, subjectivity and truth: Essential works of Foucault 1954–1984*, Vol 1. Penguin, London, pp. 253–280.
- Gavish, T., Oplatka, O., 2012. Tali Gavish & Izhar Oplatka (2012) Educational leadership in the era of mass media: state, consequences and repercussions. *Sch. Leader. Manag.* 32 (1), 73–89. <https://doi.org/10.1080/13632434.2011.642351>.
- Gillborn, D., Youdell, D., 2001. The new IQism: ‘Ability’ and the rationing of education. In: Demaine, J. (Ed.), *Sociology of Education Today*. Palgrave Publishers Ltd, Hampshire, pp. 65–99.
- Gillies, D., 2013. Educational leadership and Michel Foucault. *Educational Leadership and Michel Foucault* 1–141. <https://doi.org/10.4324/9780203495353>.
- Gobby, B., 2013. Principal self-government and subjectification: the exercise of principal autonomy in the Western Australian independent public schools programme. *Crit. Stud. Educ.* 54 (3), 273–285.
- Hargreaves, A., 1994. *Changing Teachers, Changing Times*. Cassell, London.
- Heffernan, A., 2018. The accountability generation: exploring an emerging leadership paradigm for beginning principals. *Discourse* 39 (4), 509–520. <https://doi.org/10.1080/01596306.2017.1280001>.
- Iacocca, L., (with C. Whitney). (2008). *Where have all the leaders gone?* New York, NY, London, Toronto & Sydney: Scribner.
- Ireson, J., Hallam, S., 1999. Raising standards: is ability grouping the answer? *Oxf. Rev. Educ.* 25 (3), 344. <https://doi.org/10.1080/030549899104026>.
- Ireson, J., Hallam, S., 2009. Academic self-concepts in adolescence: relations with achievement and ability grouping in schools. *Learn. Instruct.* 19 (3), 201–213. <https://doi.org/10.1016/j.learninstruc.2008.04.001>.
- Ireson, J., Hallam, S., Hurley, C., 2005. What are the effects of ability grouping on GCSE attainment? *Br. Educ. Res. J.* 31 (4), 443–458. <https://doi.org/10.1080/01411920500148663>.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven strong claims about successful school leadership*. National College for School Leadership, Nottingham.

- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28, 27–42. <https://doi.org/10.1080/13632430701800060>.
- Leithwood, K., Harris, A., Hopkins, D., 2019. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* <https://doi.org/10.1080/13632434.2019.1596077>.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22. <https://doi.org/10.1080/13632434.2019.1596077>.
- Leithwood, K., Seashore Louis, K., Anderson, S., Wahlstrom, K., 2004. How Leadership Influences Student Learning. Retrieved from New York: www.wallacefoundation.org.
- Lumby, J., English, F., 2009. From simplicism to complexity in leadership identity and preparation: exploring the lineage and dark secrets. *Int. J. Leader. Educ.* 12 (2), 95–114.
- Mac Ruairc, G., 2010. This way please! Improving school leadership: the OECD way. *J. Educ. Adm. Hist.* 42 (3), 223–246. <https://doi.org/10.1080/00220620.2010.492961>.
- Mascall, B., Leithwood, K., 2010. Investing in leadership: the district's role in managing principal turnover. *Leadersh. Policy Sch* 9 (4), 367–383. <https://doi.org/10.1080/15700763.2010.493633>.
- May, L., 1992. *Sharing Responsibility*. University of Chicago Press, Chicago, IL.
- May, L., 1996. *The Socially Responsive Self. Social Theory and Professional Ethics*. Chicago University Press, Chicago, IL.
- McGillicuddy, D., Devine, D., 2020. "You feel ashamed that you are not in the higher group"—children's psychosocial response to ability grouping in primary school. *Br. Educ. Res. J.* 46 (3), 553–573. <https://doi.org/10.1002/berj.3595>.
- Mifsud, D., 2016. The policy discourse of networking and its effect on school autonomy: a Foucauldian interpretation. *J. Educ. Adm. Hist.* 48 (1), 89–112. <https://doi.org/10.1080/00220620.2016.1092427>.
- Mifsud, D., 2017. Distribution dilemmas: exploring the presence of a tension between democracy and autocracy within a distributed leadership scenario. *Educ. Manag. Adm. Leader* 45 (6), 978–1001. <https://doi.org/10.1177/1741143216653974>.
- Niesche, R., 2011. *Foucault and Educational Leadership: Disciplining the Principal*, 1 ed. Routledge, New York, NY.
- Niesche, R., Haase, M., 2012. Emotions and ethics: a foucauldian framework for becoming an ethical educator. *Educ. Philos. Theor.* 44 (3), 276–288. <https://doi.org/10.1111/j.1469-5812.2010.00655.x>.
- Nordholm, D., Arnqvist, A., Nihlfors, E., 2020. Principals' emotional identity – the Swedish case. *Sch. Leader. Manag.* 40 (4), 335–351. <https://doi.org/10.1080/13632434.2020.1716326>.
- Notman, R., 2017. Professional identity, adaptation and the self: cases of New Zealand school principals during a time of change. *Educ. Manag. Adm. Leader* 45 (5), 759–773. <https://doi.org/10.1177/1741143216670650>.
- O'Brien, P.C., 2018. The adaptive professional: teachers, school leaders and ethical-governmental practices of (self-) formation. *Educ. Philos. Theor.* 50 (3), 229–243. <https://doi.org/10.1080/00131857.2017.1339585>.
- O'Toole, F., 2021. We need schools for messers, dreamers and misfits, opinion. *The Irish Times* 14.
- O'Farrell, C., 2005. *Michel Foucault*. Sage Publications, London.
- Piketty, T., 2014. *Capital in the Twenty-First Century* (Translated by Arthur Goldhammer). Cambridge Massachusetts & London The Bellnap Press of Harvard University Press.
- Pont, B., Nusche, D., Moorman, H., 2008. *Improving School Leadership, Volume 1: Policy and Practice*. OECD Publishing, Paris.
- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Sennett, R., 1998. *The Corrosion of Character*. W. W. Norton and Co, London & New York, NY.
- Shafik, M., 2021. *What We Owe Each Other A New Social Contract*. The Bodley Head, London.
- Smyth, E., 2018. Working at a different level? Curriculum differentiation in Irish lower secondary education. *Oxf. Rev. Educ.* 44 (1), 37–55. <https://doi.org/10.1080/03054985.2018.1409967>.
- Solbrekke, T.D., Sugrue, C., 2011. Professional responsibility – back to the future. In: Dyrda Solbrekke, T., Sugrue, C. (Eds.), *Professional Responsibility: New Horizons of Praxis*. Routledge, London & New York, NY, pp. 10–28.
- Solbrekke, T.D., Sugrue, C., 2020. Leading higher education as and for public good. In: Solbrekke, T.D., Sugrue, C. (Eds.), *Leading Higher Education as and for Public Good: Rekindling Education as Praxis*. Routledge, London & New York, NY, pp. 3–15.
- Sugrue, C., 2015a. *School Leadership in Context: National Policies, International Influences*, pp. 25–45.
- Sugrue, C., 2015b. *Unmasking School Leadership: A Longitudinal Life History of School Leaders*. Springer.
- Sugrue, C., Solbrekke, T.D., 2020. Rekindling education as praxis: the promise of deliberative leadership. In: Solbrekke, T.D., Sugrue, C. (Eds.), *Leading Higher Education as and for Public Good Rekindling Education as Praxis*. Routledge, London & New York, NY, pp. 159–176.
- Sutphen, M., de Lange, T., 2015. What is formation? A conceptual discussion. *High. Educ. Res. Dev.* 34 (2), 411–419. <https://doi.org/10.1080/07294360.2014.956690>.
- Velasquez, M., Andre, C., Shanks, T., Meyer, M., Meyer, S., 2013. What is ethics? Retrieved from. <http://www.scu.edu/ethics/practicing/decision/whatisethics.html>.
- Wagner, T., 2008. *The Global Achievement Gap Why Even Our Best Schools Don't Teach the New Survival Skills Our Children Need – and what We Can Do about it*. Basic Books, New York, NY.
- Woods, P.A., 2000. Varieties and Themes in Producer Engagement: structure and agency in the schools public-market. *Br. J. Sociol. Educ.* 21 (2), 219–242. <https://doi.org/10.1080/713655342>.

Diverse landscapes of school governing body practices

Tom Dobson^a and Anthea Rose^b, ^a Leeds Beckett University, Leeds, United Kingdom; and ^b Lincoln Higher Education Research Institute, University of Lincoln, Lincoln, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	420
Policy context	421
Business sector involvement in school governance	421
Project design	423
IEB governors acquiring educational knowledge through formal and informal training to cross boundaries	423
IEB governors embodying lay knowledge	424
E-govs' experiences of barriers to acquiring educational knowledge and embodying lay knowledge	425
The impact of the hybridization of IEBs' practices on GBs and the business sector	425
Conclusions	426
References	427

Glossary

IEB governor (Independent, External, Business-based governors) An individual from the business sector who sits on school governing body

MAT (Multi Academy Trust) An academy trust that is directly funded by the government and that more than one academy school

NED (Non-Executive Director) An independent director who is a member of a MAT's governing body

Ofsted (The Office for Standards in Education, Children's Services and Skills) A non-ministerial department of the UK government that is responsible for inspecting educational institutions

Introduction

In this article, we conceptualize the governing body (GB) of a school as a “diverse landscape of practice” (Wenger-Trayner and Wenger Trayner, 2014), bringing together individuals with different knowledge sets to improve outcomes for children and young people. This diversity has been accelerated within the global context of neoliberalism, with government policy at national level decentralizing school governance in a wide range of countries, including Japan (Lin and Wu, 2013) and Portugal (Viseu and Carvalho, 2018). Schools are required to run themselves as businesses and those that are successful against performance criteria are rewarded financially (Ball, 2003), meaning that the skills of the business sector have become key part of the diverse landscape of practice of GBs globally.

This global move toward the involvement of the business sector in school governance is explored in this article through research which took place in England. Here a 20-year policy shift toward encouraging schools to become financially independent has led to an emphasis upon appointing individuals from the business sector to school GBs (Balarin et al., 2008, p. 58). This article brings together previously analyzed data (Dobson et al., 2020; Dobson and Rose, 2019), which was collected as part of an 18-month evaluation of a banking group's school governance volunteering program. The program involved Lloyd's Banking Group (LBG) recruiting their staff to become independent, external, business-based governors (IEBs) in the North of England where schools were either underperforming, experiencing governor shortages, or both. The evaluation focused on the early experiences of 18 IEBs as school governors, highlighting both the barriers and the facilitators that shaped their participation in a new and diverse landscape of practice. In order to capture their development, the data collected includes four interviews with each of the 18 IEBs at different stages of their first 18 months of tenure, as well as interviews with key stakeholders (Headteachers, Chair of Governors, Business Managers) in each school at the end of the evaluation.

Set against research which claims that governors from the business sector marginalize the contributions of other GB members (Young, 2016, 2017), the analysis of this interview data identifies a different type of participation from these IEBs—rather than marginalizing the contributions of others, IEBs were often tentative in their participation, aware of the need to learn from the experiences and knowledge sets of the other governors in this new and diverse landscape of practice.

Given our dependency upon online communication as a result of the COVID-19 pandemic, the analysis of the interview data is timely because 5 of the 18 IEBs were recruited as “e-governors”. This meant that they were expected to attend GB

meetings remotely, allowing them to be part of a GB in schools which traditionally struggled to recruit IEBs due to their remote location. The analysis of the interviews of the 5 e-governors and their schools' stakeholders highlights the affordances and barriers of IEBs participating remotely. While all e-governors were able to contribute to their respective GBs, notably in relation to improving the schools' business practices, the barriers they experienced were significant. These included operational barriers, for example accessing remote training and remote meetings, as well as relational barriers, namely building social relationships with other GB members in order to participate in a different and diverse landscape of practice. In a post-pandemic context, these operational and relational barriers can now be thought about more widely as schools (and businesses) engage more readily with technology.

Policy context

As indicated above, the decentralization of the running of schools can be seen as part of global neoliberalism. In England, where this research took place, the process of decentralization of schools began in 2001 when New Labor introduced the first academization of schools program for failing schools and it is reflected in more recent legislation to ensure that schools have a professionalized GB that offers technical expertise and secures performance evaluation to improve public service delivery (Wilkins, 2015: 182). In England, therefore, a GB is no longer positioned simply as a "critical friend" but is responsible for the three core functions: ensuring clarity of vision, ethos and strategic direction; holding the Headteacher to account for the educational performance of the school, its pupils and staff; and overseeing the financial performance of the school, making sure money is well spent (DfE, 2015: 7).

In line with this and in line with the technologies of neoliberalism, in 2013 the government inspectorate (The Office for Standards in Education, Children's Services and Skills—Ofsted) significantly raised demands in terms of GBs (James et al., 2013), and the latest Ofsted inspection handbook (2018: 29) emphasizes the importance of a high performing GB: "the contribution of governors to the school's performance is evaluated as part of the judgment on the effectiveness of leadership and management". In England, a school cannot be rated "outstanding" if its GB falls short on any of criteria against which it is assessed. School GBs should include individuals with business expertise and the following skill-sets: knowledge of the legal sector in a number of different areas; the ability to manage processes; the experience and knowledge of running a business; and finance and leadership. Such skill-sets are often associated with the private sector. Running a school is now akin to running a business (Wilkins, 2015: 188).

In the North of England where the research took place, the government launched the "Northern Powerhouse" agenda in 2014 to challenge economic underperformance in the region (Clifton et al., 2016: 7). Initially the focus was transport and connectivity but this later widened as a result of a report from the Institute for Public Policy Research North (Clifton et al., 2016), which stated that if the North is to succeed in reaching and sustaining its economic potential, educational disadvantage must be "at the heart" of transformational objectives. This was based on data which indicates the workforce in the North of England is lower qualified than the national average (Clifton et al., 2016). According to Ofsted (2015: 11), 75% of all failing secondary schools are in the North and the Midlands.

In relation to school governance, the challenges facing GBs are greater in areas of socio-economic disadvantage (James et al., 2011: 415) like Yorkshire and Humberside, the specific region in the North of England where this research took place. One of the ways in which improved educational standards for all pupils can occur is by GBs appointing individual members with the appropriate skills, knowledge and experience to help drive schools forward. Yet the National Governance Association's most recent report (NGA, 2020) demonstrates that in England schools are increasingly struggling to recruit governors and that this is most pronounced in Yorkshire and Humberside, where 65% of schools have struggled to recruit to their GB.

LBG's school governance volunteer program, therefore, should be viewed as part of the wider Northern Powerhouse strategy and is particularly focused on Yorkshire and Humber, with 80 out of their 400 IEBs recruited in this area. Given that 25% of new governors claim to have had no training in their first year of governance (NGA, 2020), it should be emphasized that LBG's offer encompasses recruitment and training through the establishment of their National Governors' Network, which seeks to provide IEBs with the tools they needed to support their schools. This includes dial-in webinars during working hours to address particular topics relevant to school governance.

LBG's offer in Yorkshire and Humber also targeted "cold spots" in the region where rural or coastal schools were finding it difficult to recruit governors due to their remote locations. Here LBG piloted IEB governors who would dial into meetings rather than attend in person. LBG provided schools with sound and vision technology to facilitate this. At the start of the evaluation, ten remotely operating IEBs were recruited, five of whom took part in our interviews.

Business sector involvement in school governance

Despite recent policy changes to school governance in England, there is a lack of research into how business volunteers develop into their roles as school governors. This lack of research is surprising, however, when the results of an extensive evaluation of GBs undertaken prior to recent policy changes is taken into account (Balarin et al., 2008). In their evaluation report, the second aim of which was to "review the business contribution of governors", Balarin et al. (2008: 58) undertook a survey with over 3000 governors from

different backgrounds, 80% of whom agreed that the expertise of governors primarily recruited for their business expertise was “crucial to the running of the schools”. Given the quantitative nature of the evaluation, however, what is unclear is why and how the business community’s contribution to school governance was held to be crucial and whether it was particular stakeholders who tended to hold this view.

Instead of exploring the reasons behind the positive contribution of the business sector to school GBs, subsequent research into governance has focused on the negative effects of policy change. Connolly et al. (2017) literature review evaluates the effectiveness of the new skills-based model of governance as opposed to the traditional, stakeholder model where members of the GB are elected to reflect the composition of the school’s local community. While acknowledging the need for GBs to be more skills-based in terms of their composition in light of policy changes, Connolly et al. (2017: 17) conclude that “the stakeholder approach to school governance continues to have much merit as it affords a degree of authenticity to those representatives as well as a degree of independence not easily replicated in the skills based approach”. By authenticity, Connolly et al. (2017: 6) mean that the GB representatives have a genuine “interest” in their schools.

The idea of a skills-based model of school governance as a threat to the stakeholder model is articulated further by Young in research articles (2016, 2017) based on her Ph.D. thesis (data collected in 2011/12). Drawing upon semi-structured interviews and observations in four schools as well as wider policy critique, Young identifies how the rise of a skills discourse in policy serves to create a false “skills/representation binary” (Young, 2016: 170). This binary is further explored in a later paper (Young, 2017: 42), which interestingly draws upon theories of “deliberative democracy” to conceptualize GBs. Here the focus is on the nature of “expert” knowledge and how this knowledge interacts with “lay” knowledge throughout decision-making processes in a GB. While emphasizing the “slippery” nature of knowledge types, here Young (2017: 42) differentiates between two forms of “expert” knowledge at play within a GB—“educational” and “managerial”.

Young’s key finding is that “lay” knowledge is both marginalized and co-opted by “managerial” knowledge and she identifies three key reasons for this. Firstly, within a wider context of neoliberalism where education becomes an auditable product (Ball, 2003), managerial knowledge is privileged as the business of the GB becomes primarily associated with “compliance-checking” and “accountability” (Young, 2017: 52)—a point further developed by Wilkins (2016: 112), who views the “performative accountability” of the neoliberal education system as promoting “calculative technologies” which play to the skills and discourses of the business sector. Secondly, in recognizing the value of educational knowledge as expert knowledge, managerial knowledge “lays claim” to educational knowledge (Young, 2017: 53). And finally, because lay knowledge tends to be associated with “vague conceptions of common sense” (Young, 2017: 42), managerial knowledge co-opts conceptions of “common sense” leaving little space for lay knowledge to be articulated.

This article and the analysis of the data generated for the evaluation directly respond to Young’s findings. In doing so, this article also addresses a gap in the research literature to capture the experiences of IEBs as they begin in their roles on school GBs. In order to think about how the IEB governors developed in these roles, I draw upon the theoretical framework of Wenger-Trayner and Wenger-Trayner (2014) who view professional development as occupying specific “landscapes of practice”. For Wenger-Trayner and Wenger-Trayner (2014: 19), professional identity and associated knowledge results from participation in landscapes of practice which have three dimensions—“flat”, “diverse” and “political”. Here, they see “flat” as the everyday experiences of a group of professionals who experience their landscape as a predictable and taken for granted normality. GBs, however, are comprised of a “diverse” group of individuals who, depending upon the landscape of practice they experience outside of their governance roles, will find themselves participating by drawing upon knowledge which, according to Young (2017), is either educational, managerial or lay. The lived experience of individuals within GBs is, therefore, more likely to be “diverse” than “flat”. Linked to this, the diverse landscape’s “political” dimension is potentially more apparent than other landscapes of practice as, according to Young, expert knowledge is valued and lay knowledge is marginalized.

What a landscape of practice like a GB brings to the surface, therefore, are boundaries between different practices where individuals have a “lack of a shared history”, which means that these “boundaries are places of potential misunderstanding and confusion arising from different regimes of competences and commitments, values, repertoires, and perspectives” (Wenger-Trayner and Wenger-Trayner, 2014: 9). Wenger-Trayner and Wenger-Trayner’s (2014) key idea is that learning in the form of professional development occurs through “crossing boundaries”—an often uncomfortable maneuver that involves shifting away from the safety of a “flat” landscape. For Clark et al. (2017), in order for professional development to occur, it is the nature of the relationships between participants as they cross boundaries which needs articulating and exploring. When this occurs, boundaries between practices are not merely identified through difference but are crossed as learning takes place through a “hybridization” of practice (Clark et al., 2017: 245). Young’s research where “lay” knowledge is seen as marginalized by managerial knowledge suggests hybridization which involves lay knowledge could be difficult to achieve within GBs, especially as educational and managerial knowledge form their own exclusive hybrid practice within a wider neoliberal context.

With relation to IEBs who were participating remotely, while there is no literature relating directly to remote governance of schools, the concept of e-governing, where achieving governor responsibilities through “the use of IT” (Mundy and Musa, 2010: 148), is relevant. Here e-governing is defined as using IT to allow governors to attend meetings remotely rather than using IT to facilitate paperless meetings which governors attend in person. As indicated at the start of this article, the pandemic has forced the education sector to engage with online communication and in England in 2020 86% of GBs met remotely (NGA, 2020). What is of interest here is what can be learned from how technology was used by schools to support the e-governor IEBs to cross boundaries in the new and diverse landscape of a school GB.

Project design

All new IEBs recruited by LBG in the Yorkshire and Humber region were invited by LBG to be part of the evaluation. Eighteen volunteered, representing 23% of those recruited by LBG in this region at the start of the evaluation. As indicated earlier, the evaluation tracked IEB governors from the start of their journey to approximately 18 months into their tenure. The IEB governors fell into three distinct groups: nine school governors (FGs); five non executive directors (NEDs); and five governors who attended GBs remotely (e-govs). FGs were traditional in nature—part of an individual school's GB, they attended meetings in person and often held roles on other school committees. The NEDs were senior leaders within LBG, and their governance role was strategic across a number of schools at Multi Academy Trust (MAT) level. E-govs were appointed to schools that struggle to recruit governors, largely due to their rural location, the idea being that they could participate in GB meetings remotely.

Each IEB governor was interviewed four times by phone over an 18-month period. The semi structured nature of the interviews provided the evaluation team with longitudinal and in-depth data on each LBG governor and their experience of being a new school governor. Each interview lasted approximately 30 min and the interviews were spread approximately 4 months apart. The interviews aimed to capture the IEB governors' experience of school governance at different stages and covered issues such as: the recruitment and induction process; training and support offered; and how well they settled into their role.

At the end of evaluation, key school stakeholders were also interviewed in order to provide their perspective on the development and impact of the IEB governors. These stakeholders included: School Executives (13); Chairs of Governors (14); School Business Managers (5); IEBs' line managers (4); and other key stakeholders (4). Teachers and parents were not included in this process, although the value of including their perspectives in future research is recognized.

All interviews were audio recorded and fully transcribed. The evaluation process followed the University's ethics procedures, in line with the British Education Research Association (BERA) guidelines. All participants were given information and consent forms, informing them of their right to withdraw at any time and that all data would be anonymized. A potential problem of evaluation is bias, with participants and researchers wanting to provide a purely positive outlook on the project under evaluation. In order to reduce the potential for bias, each IEB governor was interviewed by the same researcher who was, therefore, able to build a sense of rapport and openness. Perceptions of the IEB governors were also triangulated with perceptions of key stakeholders in order to help verify any claims made.

Using NVivo, a thematic approach to analyzing the data across participants through coding was taken (Miles and Huberman, 2014). This involved breaking down the evaluation's key foci on recruitment, training and development into themes. For this article, the thematic coding was then further interrogated in relation to Young's (2017) heuristic of knowledge as either educational, managerial or lay. Using the idea of professional development on a GB as boundary crossing, the themes were also looked at across time to explore how IEB governors experienced the educational knowledge with which they were initially unfamiliar over an 18-month period. This involved thinking about how their claims to "lay" knowledge manifested itself, particularly in relation to their "authenticity" as governors with a genuine interest (Connolly et al., 2017: 9) in their schools.

As a result of this theoretical analysis, four key themes emerged, which are explored below:

- IEB governors acquiring educational knowledge through formal and informal training to cross boundaries;
- IEB governors embodying lay knowledge;
- E-govs' experiences of barriers to acquiring educational knowledge and embodying lay knowledge;
- The impact of the hybridization of IEBs' practices on GBs and the business sector.

IEB governors acquiring educational knowledge through formal and informal training to cross boundaries

Initially, the evaluation was set to run for one academic year. However, following our second round of interviews with IEB governors in the spring term, our analysis of the interview transcripts indicated that it was likely to take more than a year for IEBs to settle into their roles and participate more fully in the diverse landscape of practice of their GB. This was due to the initial "peripheral participation" (Lave and Wenger, 1991) of IEB governors as they realized their need to acquire educational knowledge. In light of this and in light of the aims of the evaluation, the funders agreed to extend the evaluation by one school term.

Some IEB governors had school-aged children but this tended to be the extent of their educational knowledge prior to becoming a school governor. All were aware of their lack of educational knowledge and some exhibited the "confusion" caused by a "lack of a shared history", which is symptomatic of going through a process of boundary crossing (Wenger-Trayner and Wenger-Trayner, 2014: 19). Most talked about the educational "jargon" which they had to learn in order to fully participate and understand governor business. Some also commented on the length of documents circulated prior to meetings.

While the educational discourse of the GB's landscape of practice initially served as a barrier to participation, on the whole, and in line with Young (2017), the IEB governors "valued" educational knowledge as expert knowledge and were quick to acquire, understand and use it. This meant that rather than being marginalized by the need to cross boundaries, all IEB governors began to make some kind of contribution to meetings from the outset. One FG, for example, was surprised by the speed with

which she began to participate: “At the first meeting I thought I’d probably just sit there, just listening because I’m not an extrovert person really. But I did actually make quite a few comments, suggestions and raise a few things, ask questions, queries about things.”

According to another FG’s Chair of Governors, the relative speed of participation in the GB meetings was because, “if he didn’t understand a particular piece of jargon, he certainly asked, ‘what does that mean?’ and ‘what’s the implication of that?’ Whereas some other people would just let it wash over them and hope to pick things up; but FG was keen to learn more quickly, which was good.” Indeed, not being afraid to ask questions or ask for clarification was seen by some school stakeholders as one of the key strengths of a governor primarily recruited for their business expertise as opposed to a lay governor.

An awareness of the need to acquire educational knowledge in order to participate was linked to the different kinds of governor training accessed by different IEB governors, which included local authority training, school training, and LBG training. Some IEB governors also valued filling in a reflective log. Initially, the log was part of the research team’s data collection tools, but as the majority of the IEB governors found they did not have time to fill them in, the logs were ultimately not used for this purpose. However, two IEB governors did continue using the logs as a “boundary object” (Clark et al., 2017), which helped them think about the nature of their participation in the new landscape. For one FG, the log helped her to place her development as a school governor acquiring educational knowledge in a landscape of practice: “it reminds you of the incremental steps that you’ve taken along the way because otherwise it’s very easy, 16 months on, to think well actually now you’re asking me, what did I do?”

By comparing the three governor types, some generalizations about the ways in which these different groups acquired educational knowledge and what this meant for their participation in the landscape of their GB were identified. The barriers experienced by e-govs are discussed below, but it was the NEDs who, as a result of the skill-sets acquired in their jobs as senior leaders, were least likely to access formal training and most likely to learn through professional conversations with key school stakeholders. One NED, for example, said, “I’ve not tapped into any training. I felt like I didn’t need it, I’ve done leadership days before, I’ve facilitated before, I’ve set strategies for teams and areas before”; and his CEO testified how this NED had “over-performed in terms of his uptake of complicated sector specific knowledge.”

IEB governors embodying lay knowledge

As well as the importance of acquiring educational knowledge, the interviews with the IEB governors emphasized the importance of personal relationships in terms of participating in the diverse landscape of the GB. Indeed, the building of these relationships with other GB members was evident with most IEB governors and was symptomatic of both the ideological values which had initially motivated their participation as well as the ways in which they endeavored to gain a contextual understanding of their school or schools. Taken together, these motivations and context-specific actions meant that IEB governors were developing “authenticity” associated with the stakeholder model (Connolly et al., 2017) through acquiring contextual, local knowledge, which is similar to Young’s lay knowledge (2017).

From a broad ideological standpoint, all IEB governors spoke about wanting to take on a governing role in order to “make a difference.” Often this was linked to pupil employability and how their organization could help ensure pupils had the right skills and understanding to enter the world of work, but more generally it was about volunteers working with a school which was part of, or close to, their communities in order to “give something back.” One FG, for example, proclaimed a “passion” for schools and a belief that education can change lives: “Education for me is one of those things that makes a big difference. So my family background, I am the first to go to university. That seriously changed the kind of work I do compared to what the rest of my family do ... I wish I’d become a governor years ago.”

For IEB governors, their purpose went beyond “compliance-checking” where education is viewed by professionals as an audit-able commodity (Wilkins, 2016; Young, 2017). Another FG, while acknowledging external measures of school competence were important, had a clear sense of the need to provide a holistic education bespoke to the pupils in the school: “It’s that constructive challenge to help the school move forward with its aims, one of which is of course to become outstanding. But also, I suppose on the softer side, retaining the character of the school while at the same time clearly ensuring that the right education is provided to sectors of the pupil population.”

For the majority of IEB governors, understanding the local community as well as the teachers was key in terms of ensuring their effective participation on their GB. This enabled IEB governors to understand and share the values of their schools, and in order to do this, most IEB governors visited classes, spoke to teachers and, in some instances where they were unfamiliar with the community, accompanied Community Officers on visits. The NED’s School Executive recounted: “He spent a morning in school and went out on home visits with our New to English Team who support Roma students who have very poor attendance records and he went out with them in the minibus and he watched them. In fact, he did a home visit with one of them and knocking a kid up who hasn’t got up and out to school but trying to get them in for the rest of day.”

Equally, there was an awareness and understanding of the importance of the GB in leading the school and how the IEB governors had an obligation to commit their time and energies to ensuring the school was successful. One FG, for example, was acutely aware of the history of her GB and what this meant for her participation: “One of the things they’re hoping to achieve, which I would like to support them with, is a little bit of stability within their GB as they have had a lot of change for the last couple of years and someone who is hopefully going to stick around.”

E-govs' experiences of barriers to acquiring educational knowledge and embodying lay knowledge

From the perspective of the e-govs themselves, there was a clear perception of barriers to them participating fully in the new and diverse landscape of a school GB. These barriers were apparent at each stage of the settling-in process, from initial training and induction to throughout the academic year. In relation to training, while all IEBs had access to the same range of online resources and training programs, in some instances there was an expectation placed by schools upon these IEBs to attend face-to-face induction training. For three of the e-govs, the travel time was impractical, with one commenting: "because of the distance that's involved it's physically impossible to attend those."

Training aside, all five e-govs and their schools experienced different degrees of problems with the technology designed to engage them in GB meetings. These were quickly resolved in the case of one e-gov, but the other four experienced ongoing technical problems that, in some cases, were never resolved. For one e-gov this barrier was so extreme as the school were never able to work any of the technology meaning that the e-gov had to correspond directly and only with the Chair outside of meetings. As the e-gov identified, this was in part down to the culture of the GB: "they just don't operate in that way, they work on a face-to-face basis—that you're in the room for the meetings." More than this, the school's physical environment was seen by the Chair as not accommodating the remote governance concept: "it's a very old school. The technology is about 30 years behind the times and we just really had problems in getting an audio contact with our e-gov during meetings."

The barriers in relation to induction, training and virtual attendance at GB meetings meant that these IEBs were less likely than other IEBs to acquire "educational" knowledge in order to take on specific roles on sub-committees within their schools. Indeed, while other FGs and NEDs had at least one specific role in addition to being on the main GB, this was the case for only two of the e-govs. The barriers to engagement experienced by the IEBs meant, as a whole, e-govs did not settle into their roles as quickly or effectively as the other IEBs.

For one e-gov the barriers to engagement became so great that by the end of her first year in post she felt that she had no other option but to resign. However, according to the School Executive, this Multi Academy Trust (MAT) made two similar appointments, who were not part of this evaluation but who took to their roles more effectively than this e-gov as they lived near enough their schools to make a visit and attend GB meetings in person. Accordingly, one of these IEBs was appointed as an advisor to the MAT's Audit and Finance committee; the other has physically attended most meetings and is now Chair of Governors at her allocated school. The School Executive explained: "She was able to physically attend meetings more. She came in through the remote governance program but she's able to be more than a remote governor but if we'd not engaged in the program we would not have ended up with her being a physical presence around our governing body table and now Chair."

What this emphasizes is the importance of embodying lay values and knowledge by forging relationships with other GB members. Given that GBs are together perhaps once or twice a term, this was even a challenge for those IEBs who did attend meetings, but when e-govs attended virtually with technological issues, barriers to building relationships become exacerbated. One Chair commented that this was a key issue which required further thought: "I think that one of the areas that [needs] addressing is how you build up the social side of things, which is important."

One e-gov ultimately felt so compelled to try to build a relationship with his GB that he undertook a 200-mile round trip in order to do so. He commented: "I understand the actual principles behind the project in the first place to try to reach these remote locations, but having that face-to-face contact, you know, the hand shake and the sitting down with governors, I think helps cement the relationship that much stronger." In line with this, where IEBs did settle into their roles more quickly, a key variable was the relative proximity between their homes and schools. For two e-govs, the distance from their homes to their schools was not prohibitive of them attending occasional GB meetings in person and undertaking school visits, and this meant that they both felt part of their respective GBs and acquired the "lay" knowledge relating to the specific context of their schools, its staff and pupils. As one of these e-govs commented, attending meetings in person at the start of their appointment helped her develop a working relationship with the rest of the GB and enabled her to then participate more fully in meetings by dialing in: "I'd been to a few of the meetings so I'd met them all face-to-face and they were a really good group of people. So being on the phone worked really well because I never felt like they'd forgotten that I was there and they kept me included as and when if anything was happening."

The impact of the hybridization of IEBs' practices on GBs and the business sector

Despite the barriers experienced by e-govs, the perspectives of IEB governors and their school stakeholders demonstrated that all IEB governors had ultimately been able to participate in the diverse landscape of their GB in order to make some kind of impact upon different aspects of the running of their school. This was because all IEB governors brought with them knowledge and skills from their business-based landscape of practice which were shared in order to help GBs run their schools as a business. This, in turn, impacted positively upon a range of aspects including: outcomes for pupils; the confidence, knowledge and skills of the GB; and school business practices. In relation to pupil outcomes, school stakeholders felt that their IEB governors had, on the whole, indirectly contributed to improving school performance and the educational outcomes of pupils. As one

Headteacher stated, this tended to be associated with ensuring the “financial and business state of the school is intact.” Indeed, all school stakeholders felt their IEB governors had increased the confidence, knowledge and skills of their GB as a result of the complementary skill-set they brought with them. In relation to school business practices, depending upon the existing knowledge of the GBs, some schools were in greater need than others of the IEB governors’ professional skill-sets. One Chair, who had a background in purely educational rather than managerial knowledge, spoke about how their e-gov had been able to “act as a mentor to me when I was starting to move things forward with the Headteacher.” This meant that the IEB governor was informally employing mentoring techniques to develop the Chair’s managerial knowledge in the specific context of the school’s finances.

However, as indicated above, the impact of IEB governors upon their GBs and schools was by no means immediate and tended to become more apparent to the research team at the fourth and final interview. What this demonstrates is that the application of the IEB governor’s existing skill-set to a GB was contingent upon them crossing boundaries to acquire educational knowledge rather than tenuously laying “claim” to educational knowledge (Young, 2017). Perhaps more fundamentally, this acquiring of educational knowledge seemed to go hand in hand with the IEB governors holding idealistic views about the purpose of education for all children in relation to their own participation in their GB. This meant that the acquiring of educational knowledge was inextricable from the acquiring of lay knowledge for IEB governors—a stark counterpoint to earlier research which held managerial knowledge as a direct threat to lay knowledge (Young, 2017).

That being said, even once some IEB governors had acquired new knowledge, on occasion they found it hard to bring their business practices into the landscape of the GB. Often these difficulties related to the ways in which the GB meetings were run, with IEB governors expecting short, outcome-driven meetings. One FG, for example, had recommended to his GB that they streamline their documentation prior to meetings and adopt a more corporate approach within meetings: “What I found is from the governance meeting is that they just have stuff to talk about and it’s not actually clear whether a decision needs to be made or not. I have raised [it] at the last governor meeting and ... I think they were more comfortable with how it’s currently set out.” As discussed above, these operational barriers were even more acutely experienced by e-governors who experienced insurmountable technological problems.

This indicates that the managerial knowledge of IEB governors, even when transformed by the acquisition of educational and lay knowledge, was at times subject to being marginalized within the GB. In contrast to Young, therefore, what our study shows is that managerial knowledge was not only tentative in developing agency in GB meetings but that managerial knowledge also appreciated the need to be transformed and hybridized by educational and lay knowledge—a process which meant that managerial knowledge itself was sometimes marginalized by what had become the “flat, common-sense” practices of the GBs (Wenger-Trayner and Wenger-Trayner, 2014).

What is also evident is that the acquisition of new knowledge in order to participate in GBs resulted in the transformation of practice for IEB governors in different ways. These new hybrid practices, which resulted from crossing boundaries (Clark et al., 2017), were not simply enacted in GBs but were transforming in relation to the workplace. Indeed, most IEB governors reported that their governance role had positively impacted on their professional work and benefitted them in terms of their continuing professional development. The professional benefit most commonly cited by IEB governors and their line managers was not the acquiring of new skills per se, but rather a growth in their “self-confidence” which enabled them to carry out their professional role more effectively. This was explicitly linked to a “broadening of horizons”, “working in a new environment” and “working with people from different backgrounds, cultures and communities”. As one e-gov commented, “I think because I’ve gone into something that I’ve not done before, it’s given me more confidence because I’ve taken the skills that I’ve got in my current role, from a finance background, to use them in another environment and with people that I didn’t know and who didn’t know me. And I feel like they appreciated my input and appreciate the support they got from us, so it’s just given me a bit more confidence in my role.”

In line with the theoretical framework of this article, the confidence gained by IEB governors is seen as arising from the changes undertaken through acquiring the knowledge to participate effectively and with agency in a landscape other than their usual workplace. As one FG’s line manager stated, gaining confidence seemed to have a more profound effect upon his participation in the businesses landscape of practice: “I feel that he’s now got that drive and passion within him. It will end up seeping through into his work. I feel that it has had a positive impact on giving him that purpose that he has been looking for. That spark.”

Conclusions

By capturing the perspectives of new business governors as they settled into their roles as well as the perspectives of key school stakeholders on this process, this article provides an insight into the ways in which IEB governors cross boundaries in the acquisition of educational and lay knowledge to transform their practices within the diverse landscape of practice of a GB. In contrast to earlier research, which tends to see the business community as a threat to both the stakeholder model (Connolly et al., 2017; Wilkins, 2016) and lay knowledge (Young, 2017), our longitudinal study captures some of the complexities of practices in GBs involving new business governors.

Seen from this longitudinal perspective, we arrive at a new conceptual understanding of the participation of managerial knowledge in school GBs. Rather than participating through exerting power which claims educational knowledge and co-opts and marginalizes lay knowledge (Young, 2017), managerial knowledge is often tentative in its initial participation in this diverse landscape. Furthermore, managerial knowledge's tentative participation is symptomatic of an awareness of the need to acquire rather than lay claim to both educational and lay knowledge in order to be able to contextualize skills and practices. More fundamentally, managerial knowledge seems to be underpinned by a value set which is altruistic and authentic and which has otherwise been seen as only belonging to lay governors within a stakeholder model (Connolly et al., 2017). This tentative participation, involving those with managerial knowledge seeking to acquire both expert and lay knowledge, means that while becoming more diverse, the GB becomes less hierarchical as a landscape of practice (Wenger-Trayner and Wenger-Trayner, 2014), which in turn facilitates boundary crossing and the transformation of practice through "hybridization" (Clark et al., 2017). This suggests that Connolly's (2017) argument about the lack of "authenticity" of a GB with members from the business sector does not always hold true.

Given the global context of neoliberalism and the global trend toward the decentralization of school governance, the implications of this new conceptual understanding of the role of the business sector in the running of school GBs are manifold. In relation to England, where recruitment of skilled governors is difficult (NGA, 2020) and where this is especially the case in areas of economic disadvantage (James et al., 2013), operationalizing these implications will help address governor shortages and, albeit indirectly, improve outcomes for children.

Firstly, this case study indicates how that other businesses should follow suit as part of their civic duty. This would include businesses allowing their staff time to undertake school governance. The resulting hybridization of practice, we have shown, is likely to have benefits for schools and businesses alike, especially within the wider context of performativity. Secondly, given the various ways in which different IEB governors acquired educational knowledge, an understanding of the bespoke nature of professional development is needed by those who provide induction and training (local authorities, private companies, schools). This includes allowing for development to be bespoke and informal as appropriate, with an emphasis also upon providing high quality online training for IEBs operating remotely. Thirdly, in relation to the barriers to participation experienced by e-govs and given the way that the pandemic has since forced school GBs to change their cultures and practices, IEB governors operating remotely could in this current context offer a more significant contribution to recruitment issues faced by GBs in both disadvantaged and remote areas. This could be especially impactful were GBs to think about ways of building relationships with remote IEBs in order to overcome social barriers that have been apparent in this study. And finally, further research needs to be undertaken by the academic community and shared with school leaders in order to understand the ways in which different GB members utilize their knowledge sets and interact with other members. This could involve researchers working alongside or within GB meetings in order to develop a deep understanding of the nature of different stakeholder participations in a diverse landscape of practice.

References

- Balarin, M., Brammer, S., James, C., McCormack, M., 2008. The school governance study. In: *Business in the Community*. Business in the Community, London.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Clark, J., Laing, K., Leat, D., 2017. Transformation in interdisciplinary research methodology: the importance of shared experiences in landscapes of practice. *Int. J. Res. Method Educ.* 40 (3), 243–256.
- Clifton, J., Round, A., Raikes, L., 2016. *Northern Schools: Putting Education at the Heart of the Northern Powerhouse*. Institute for Public Policy Research, Manchester.
- Connolly, M., Farrell, C., James, C., 2017. An analysis of the stakeholder model of public boards and the case of school GBs in England and Wales. *Educ. Manag. Adm. Leader* 45 (1), 5–19.
- DFE, 2015. *The Governance Handbook: For Trustees of Academies and Multi-academy Trusts and Governors of Maintained Schools*. Available at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/481147/Governance_handbook_November_2015.pdf. (Accessed 20 May 2021).
- Dobson, T., Rose, A., 2019. The use of technology to enable school governors to participate remotely in meetings: a solution to skilled governor shortages in schools? *Manag. Educ.* 33 (3), 123–139.
- Dobson, T., Rose, A., Parton, G., Hart, M., 2020. Boundary crossing in school governing bodies: perspectives from the business community. *Educ. Manag. Adm. Leader* 48 (3), 532–548.
- James, C., Brammer, S., Connolly, M., Fertig, M., James, J., Jones, J., 2011. "The Hidden Givers": A Study of School GBs in England. CFBT Education Trust, Reading.
- James, C., Brammer, S., Connolly, M., Spicer, D.E., James, J., Jones, J., 2013. The challenges facing school GBs in England: a "perfect storm"? *Manag. Educ.* 27 (3), 84–90.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. CUP, Cambridge.
- Lin, Y., Wu, C., 2013. Analyses and changing models of school governance in Japan: school council cases. *J. Res. Educ. Sci.* 58 (4), 103–131.
- Miles, M., Huberman, M., 2014. *Qualitative Data Analysis*, third ed. SAGE, London.
- Mundy, D., Musa, M., 2010. Towards a framework for eGovernment development in Nigeria. *Electron. J. eGovernment* 8 (2), 148–161.
- NGA, 2020. *School Governance in 2020*. Available at: <https://www.nga.org.uk/getmedia/3b313a00-2e13-4ae5-8d9f-7445de07f395/School-Governance-Report-2020-WEB2-AW.pdf>. (Accessed 24 March 2021).
- OFSTED, 2015. *The Annual Report 2014–15*. OFSTED, Manchester.
- OFSTED, 2018. *School Inspection Handbook: Handbook for Inspecting Schools Under Section 5 of the Education Act 2005*. OFSTED, Manchester.
- Viseu, S., Carvalho, L.M., 2018. Changes in school governance and the reshaping of head teachers' roles and identities in Portugal. In: Normand, R., Min, L., Carvalho, L.M. (Eds.), *Education Policies and the Restructuring of the Educational Profession: Global and Comparative Perspectives*. Springer, Singapore, pp. 57–70.
- Wenger-Trayner, E., Wenger-Trayner, B., 2014. Learning in a landscape of practice. In: Wenger-Trayner, E., Fenton O'Creedy, M., Hutchinson, S., Kubiak, C., Wenger-Trayner, B. (Eds.), *Learning in Landscapes of Practice: Boundaries, Identities and Knowledgeability in Practice-Based Learning*. Routledge, Oxford, pp. 12–19.

- Wilkins, A., 2015. Professionalizing school governance: the disciplinary effects of school autonomy and inspection on the changing role of school governors. *J. Educ. Pol.* 30 (2), 182–200.
- Wilkins, A., 2016. *Modernizing School Governance: Corporate Planning and Expert Handling in State Education*. Routledge, Oxford.
- Young, H., 2016. Asking the “right” questions: the constitution of school GBs as apolitical. *J. Educ. Pol.* 31 (2), 161–177.
- Young, H., 2017. Knowledge, experts and accountability in school GBs. *Educ. Manag. Adm. Leader* 45 (1), 40–56.

Understanding leadership identity construction: a gendered analysis

Pontso Moorosi^a and Victoria Showunmi^b, ^a Department of Education Studies, University of Warwick, Coventry, United Kingdom; and ^b UCL Institute of Education, University College London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introducing leadership	429
Leadership development	430
Women and leadership	430
Intersecting gender and ethnic identities	431
Leadership identity construction: a gendered analysis	431
Individual internalization (personal)	432
Relational recognition	433
Collective endorsement	433
Toward inclusive leadership identity theories	433
Implications for leadership development	435
Conclusion	435
References	435

Introducing leadership

Leadership has a long history as a theme in management studies and, more recently, trumping management as a more preferred term and permeating educational landscape and research. As a concept, leadership has been riddled with controversies since its inception. The first controversy concerns the lack of agreement on what leadership means. Stogdill (1974) had earlier argued that “there are as many definitions of leadership as those who have tried to define it” (p. 1). Carroll and Levy (2010) have labeled leadership a “mysterious and complex concept” (p. 212), with others, including Billsberry (2009: 3) and Iwowo (2015: 114) contending that leadership is “in the eye of the beholder”. This social constructivist perspective suggests the significance of defining the contextual use of the term leadership. As Iwowo perceives, the lack of conceptual clarity surrounding leadership presents a dilemma for leadership development as it is difficult to develop people for something that is not known. Equally, identifying effective leadership practices becomes well-nigh impossible if the nature of what is being sought is not known.

In western conception, originally, leadership research focused on the “who” of leadership, seeking to distil the personality traits that separated leaders from non-leaders. These early models of leadership were influenced by the notion that [great] leaders possessed particular traits, and were based on research conducted in business, military and government settings, and the White men who led these institutions (Middlehurst, 2008). Trait leadership theory thus, suppressed and neutralized “difference”, including considerations of how gender and race/ethnic dimensions may impact leadership (Parker, 2005). Leadership stereotypes tended to be linked to whiteness, maleness, heterosexuality, and middle-class attributes (Coleman, 2012). Although leadership research has moved away from trait approaches and into more behavioristic and situational approaches, these criticisms remain true of much contemporary leadership theory.

Secondly, leadership theorizing has generally shifted from individual to relational dimensions. Fletcher (2003) argued that modern conceptions of leadership are “post heroic” and rely on a network of interdependencies at different levels within organizations. Approaches such as distributed leadership (Lumby, 2013) and transformational leadership (Dvir et al., 2002), emphasize the significance of leaders’ engagement and leaders’ impact on followers. Despite this shift, researchers have tended to minimize how such relationships may be influenced by social identity and status inequality. Leadership scholars (e.g., Ayman and Korabik, 2010; Chin, 2010) have argued for closer integration between leadership research and diversity, challenging assumptions that conception of leadership is gender, race and culture neutral. Ayman and Korabik (2010) argued that in actuality, “gender and culture affect leaders’ style, behavior, emergence and effectiveness in many complex ways’ (p. 166).

The third controversy is that the phenomenon of leadership has traditionally been associated with the conception of formality of hierarchical structures and positional power within organizations as opposed to “informal and emergent leadership” (Middlehurst, 2008, p. 327). In this conception, leaders are typically those who get assigned to managerial roles, usually at the top of organizational structures. In educational institutions, such as schools, those at the top of organizational structures, who by default are regarded as leaders whether they do indeed provide organizational leadership or not, are school principals or head-teachers and those who follow them in the hierarchy. It is an expectation that those charged with the responsibility of headship will be leaders irrespective of whether they exhibit leadership qualities or not (Setlhodi, 2020). This hierarchical and top-down conception of leadership is increasingly problematic as it neither explains ineffective principalship nor does it explain the locus of influence that does not come from the top. This notion of hierarchical leadership is found unhelpful in contexts that are indigenously collaborative and interdependent, such as the majority of African societies, which, despite their own hierarchies, have traditionally sought to lead by

consensus (Iwowo, 2015; Setlhodi, 2020). Indeed, scholars have intimated the need to understand leader/follower relationships (Learnmonth and Morell, 2017) and the relationship between leader and the process of leading (Grint, 1997; Setlhodi, 2020) if we are to fully understand leadership as a social process of interaction.

In trying to understand this relationship, leadership scholars have viewed leadership in terms of a joint effort of “mutual influence” that is independent of any formal position of responsibility (De Rue and Ashford, 2010, p. 627). Furthermore, to understand the relationship between leader and leading, it is important to understand how those who end up with the badge of leadership develop this identity. Grint (1997) argued that leadership is a product of social influence that emerges overtime. Understanding the power dynamics in this social influence is central to understanding of gender inequity in traditional and contemporary notions of leadership and can equip us with skills to develop more relevant and responsive ways of developing leaders and bridging the gender inequality in leadership.

Leadership development

Leadership development is itself viewed as an essentially longitudinal process that happens over time (Day et al., 2009; Day, 2011; Lord and Hall, 2005; Miscenko and Day, 2016). From this lifespan perspective, leadership development is a process which begins in childhood and continues through adulthood, occurring through life experiences as leader identities are forged (Day et al., 2009). Moorosi and Grant (2018) illustrated this lifespan construction through the leadership socialization of African school leaders, suggesting a “relationship between early formations of leadership identity and a life-time commitment to providing leadership for schools and communities” (p. 655). While a distinction between leader development and leadership development has become necessary in interrogating processes that build organizational leadership versus individual leadership (Day, 2001; Dalakoura, 2010), there has also been an argument toward understanding these processes as complementary sides of a coin. In this regard, the most helpful distinction is that made by Day (2001) and Dalakoura (2010) that leader development focuses on human capital while leadership development builds social capital. Having made that distinction, the authors agree that both are necessary for leadership effectiveness within an organization. In fact, leader identity development or leader identity construction denotes the process by which one comes to know oneself and to be known as a leader, and central to that is the notion of leadership development. By definition, leader identity construction is an aspect of leadership development that entails an individual process of self-concept, which van Knippenberg et al. (2004) perceive to be a “mediator of the relationship between leadership and follower attitudes and behavior” (p. 829). This self-concept gets affirmed by the collective followership within the organization through the reciprocal process of claiming and granting (DeRue and Ashford, 2010). Indeed, DeRue and Ashford (2010) contend that the process of claiming and granting is a social construction of identities that make identities “inherently related” (p. 628). It is through this social construction process that leader identities can shift between group members from leader to follower and vice versa (DeRue and Ashford, 2010).

As a body of research, leadership identity research has involved differing terminology including identity work (Carroll and Levy, 2010; Petriglieri, 2011), identity formation (Skinner, 2014; Komives et al., 2005); and identity construction (De Rue and Ashford, 2010; Moorosi, 2014). Although we do not have scope to deal with these different aspects to the same degree, we contend in this chapter that all aspects of identity are important to understanding leadership and how it remains a contested terrain of power and inequality. Thus, we align our thinking with those who see leadership as a process that entails more than just inhabiting a role, but a process of influence and direction-giving wherein an individual should identify as a leader and be sanctioned as such by others (Grint, 1997; Miscenko and Day, 2016).

Although leadership identity is well-established as an area of study within broader leadership and management studies, it is not widely explored in school leadership, particularly with regard to women leaders, who are still in danger of internalizing male perceptions of leadership within their own identities, as the field has been developed using male-generated theories (Murakami and Torsen, 2017; Parker, 2005). These male conceptions of leadership have implications for leadership emergence, leadership effectiveness and leadership development as aspects of leadership identity development (Ibarra et al., 2014). Indeed, men and women have been found to experience different impediments to leader identity development (Ely et al., 2011). When women prioritize the male perceptions of leadership, they inadvertently suppress their own feminine identities and when one identity is suppressed, this can lead to identity conflict (Debebe and Reinert, 2014). Karelaia and Guillén (2014) have intimated that an “identity conflict” arises when women leaders perceive an “incongruity between being a leader and being a woman” (p. 204), or at the disjuncture between how others perceive an individual and how an individual perceives herself (Debebe and Reinert, 2014, 284). Debebe and Reinert add that, “identity conflict can activate a learning process that will take an individual through a path to constructing a personal identity” (p. 284). Indeed, social identity theories of leadership view gender and other social identities as central to the construction of leadership identity (van Knippenberg, 2011).

Women and leadership

The focus on the gendered perspective of leadership identity construction is necessitated by the traditional association of leadership with masculinity, that make it difficult for women to assert their presence in leadership despite breaking the glass-ceiling and dominating the workforce in schools and other educational institutions (Eagly and Karau, 2002). Eagly and Karau argued that leadership

holds prescriptive norms that render women less favorably evaluated than men in their leadership potential and actual leadership behavior. Ely et al. (2011) argued that subtle, institutionalized forms of gender bias—stemming from workplace structures, cultures, and patterns of interaction that inadvertently favor men—shape, and often interfere with, the identity work of women leaders. This leads to the scarcity of women in leadership despite a plethora of equal opportunity legislation across the globe, including some developing contexts (Mestry and Schmidt, 2012). In many regions across the globe, women have yet to attain leadership positions within their chosen professions in any comparable numbers to their male counterparts. Leader identity construction is thus fraught with difficulties as it is more challenging for women to find the most powerful mentors and sponsors. This is due to lack of role models for women identify with, gendered career paths that still favor men and gendered organizations that reflect men's lives and men's situations (Ely et al., 2011).

As a social group, women's experiences have not been used to develop current leadership theory and experiences of women of color continue to be underrepresented in the leadership discourse. Being white, male and middle class is much more of a guarantee of career advancement than possessing attributes which do not conform to the accepted cultural stereotype of what leadership looks like and how leadership is executed (Alston, 2012). Organizational cultures and practices ensure that the gatekeepers reproduce the status quo (Eagly and Carli, 2007), disadvantaging women who are as committed to their professional goals as men, and who "pour out of the top schools with great credentials and flood into the workforce" (Slaughter, 2015, p. 81) in abundance. Yet their decreasing visibility on the upper ranks of the career ladder is testament to the fact that women's participation in positions of power is still a phenomenon less ordinary, and even less ordinary for women of color (Alston, 2012). While the "what" of leadership, i.e., the complicated and multifaceted nature of leadership and recognition that being in a senior management role is a demanding and onerous responsibility, is widely accepted, and that those who take up the mantle of leadership forfeit much in terms of their personal lives, the "who" presents as a much more problematic issue, lacking agreement and acceptance of its representation in a multiplicity of jurisdictions.

Intersecting gender and ethnic identities

Ethnic and gender identities are both central to self-concept (Showunmi et al., 2015), yet in research they are often considered independently as the determining feature of individuals' experiences, isolated from other identity strands. Leadership studies on diversity often focus on single categories, limiting understanding of the complexity of individuals' experiences (Coleman, 2012). However, it can be argued that asking people to account for experiences based on one category (e.g., ethnicity) to the exclusion of another (e.g., gender) is an invalid conceptualization of reality when membership of both is confounded in individuals (Cole, 2009). Such "essentialism" categorizes individuals' cognitive processes, personalities and experiences in relation to single identity categories (such as gender) without considering how other factors (such as other identities, historical and current social context) influence these (Debebe and Reinert, 2014).

Intersectionality as a framework for understanding the interconnectedness of social identities disrupts the notion of essentialism in leadership research. Intersectionality is the "idea that social identities such as race, class, and gender interact to form qualitatively different meanings and experiences" (Warner, 2008, p. 454). The term is often attributed to Crenshaw (1991), who drew attention to African American women's invisibility resulting from their minority status in gender and race scholarship. Consequently, ethnic minority women became subsumed into categories that did not accurately reflect their perspectives and experiences. In the twenty years since Crenshaw's work, researchers have adopted the concept of intersectionality to examine how multiple identity dimensions such as sexual orientation, social class, and nationality influence women's leadership and work experiences across the globe (e.g., Hite, 2007; Pompper, 2007; Moorosi, 2014). Moorosi (2014) used intersectional analysis to show how women leaders' identity strengthened from a leadership development program. Intersectional perspectives integrate feminist and multiculturalist perspectives to understand women's experiences from a more nuanced perspective. They highlight how all identity categories depend on each other for meaning and significance (Richardson and Loubier, 2008; Debebe and Reinert, 2014).

Leadership identity construction: a gendered analysis

Leadership identity is defined as an individual's self-definition as a leader, based on a relatively stable set of meanings associated with being a leader. In simple terms, leadership identity means the extent to which one identifies and is identified as a leader. Shollen (2018) suggests that leader identity involves journeys that are related to what the leaders do as leaders and how they experience leadership and how they are identified by others as leaders. The processes by which leaders are identified and affirmed by others indicate reciprocation which entails leader identity development as necessarily a relational and social process that essentially involves other people (Ely et al., 2011; van Knippenberg et al., 2004). In this sense, leader identity is hardly an "individual pursuit" but one that actually needs a team of people, making it an ultimate "social-relational process" (Shollen, 2018, p. 162). DeRue and Ashford (2010) argue that the extent to which someone views themselves as a leader has three benefits: it enhances motivation to lead, it improves engagement in leadership processes, and it promotes the drive to actively seek leadership development opportunities (p. 628). Consequently, internalizing this leader identity becomes an essential ingredient for effective leadership. However, women leaders have been found to experience more challenges than men in developing leader identity, due to the bias inherent in social dynamics (Ely et al., 2011).

In a school context, this internalized leader identity would inform the school leaders' ability to lead their schools toward the achievement of set goals (Tubin, 2017). Tubin suggests that a leadership identity influences the clarity of the vision for the school, the extent to which the school principal sees themselves as a leader and the extent to which their authority is accepted by followers. This latter part is particularly important for women, as a significant body of literature shows women principals struggling to have their authority accepted by their subordinates (Davids, 2018; Moorosi, 2010; Shapira et al., 2011). This acceptance or lack thereof can affect women's self-efficacy and professional identity as leaders. Indeed, Shollen (2018) shows how building relationships within a team of people is particularly essential for women developing a leadership identity, as women represent a group that has traditionally not been viewed in leadership terms.

Scholars of leadership identity have established numerous ways of researching identities in leadership, many of which are hardly explored in educational leadership research. These researchers have also established that identity in leadership can be conceptualized at three different levels: personal, relational and collective identities (see, for example: Brewer and Gardner, 1996; Debebe and Reinert, 2014; DeRue and Ashford, 2010; Ibarra et al., 2014; Lord and Hall, 2005). Other scholars have focused on social identity theory and have brought to the debate the influence of social identities that people bring to leadership (Karelaia and Guillén, 2014; Showunmi et al., 2015; van Knippenberg et al., 2004). While these conceptions differ on their focus, they do overlap, demonstrating that identity can be thought of in multiple terms (Debebe and Reinert, 2014) and that these multiple terms vary depending on the subjects and their cultural orientation. The section below examines these different levels of identity development as constructed by DeRue and Ashford (2010) and others. An attempt is made to provide a gendered analysis of these levels within educational leadership.

Individual internalization (personal)

According to DeRue and Ashford (2010), the personal level of leader identity development involves individual internalization wherein personal attributes of a leader are reconciled with one's personal identity forming a persona that marks one's individual identity (Sethodi, 2020). Ely et al. (2011) found identity construction at this level to be a developmental process that indicates acquisition of some basic skills and a clear development of self-concept or "self-definition". This self-concept assumes that to be in a position to send the right messages about who they are, leaders must first understand themselves. Self-concept is inextricably gendered as women have not traditionally perceived themselves as leaders (Moller, 2003; Moorosi, 2020).

Indeed, as Lord and Hall (2005) earlier established, this self-understanding occurs before a certain level of expertise in leadership can be achieved. At this level, leader identity may change as people go through personal development experiences such as a leadership development program, as found by Richardson and Loubier (2008) and Moorosi (2014). Moorosi found that women leaders' identity developed as they went through a leadership development program, learning new skills and dispositions that gave them the confidence to lead. However, other studies established that, since personal attributes precede leadership development opportunities, they become central aspects on which leaders build on the process of knowing themselves (see, for example, Jacobson et al., 2007; Moller, 2003). In both studies, Jacobson et al. (2007) and Moller (2003) found that principals integrated their personal identities with being school principals, thereby enhancing their leadership identities.

Bennis (2009) asserted that knowing oneself means "separating who you are, and who you want to be, from what the world thinks you are and wants you to be" (p. 50). He perceives knowing oneself as a lifetime self-invention process that happens at different stages for different people. Day (2001) identified three capabilities (self-awareness, self-regulation and self-motivation) that he argued contribute to the enhancement of one's "individual knowledge, trust, and personal power" (p. 584). Self-awareness helps leaders to: (a) clarify their own values, strengths, and limitations; (b) recognize the influence of self on others; (c) be in tune with how others perceive them; and (d) understand the contextual factors that have shaped both self-concept and others' perceptions (Day, 2001; Hall, 2004). In a study with women leaders on a leadership development program, Moorosi (2014) found that the studied women leaders' "self-awareness, self-confidence and self-regulation" (p. 11) were more pronounced at the end of program which suggested that the leadership development program acted as a workspace for identity development, while others (e.g., Hall, 2004) have stated that leaders' self-confidence and personal authority grows as leaders progress in hierarchy and experience. Working with a different group of African women leaders, researchers found that women's self-confidence grew as they climbed the leadership ladder and as they occupied more than one principalship (Komiti and Moorosi, 2020; Moorosi, 2020).

Ibarra et al. (2014) assert that one way in which leader identity is internalized is by "developing a sense of purpose" or "taking a purposeful action", which would be affirmed or resisted by others. This they found to be an iterative process, which entails one's view of oneself as a leader being reciprocated by affirmations and confirmed by others' views, reactions and interactions. Indeed, women were found to "led by example" doing things that they want others to do (Moorosi, 2020). One such participant in Moorosi's study established that leading by example minimizes resistance and buys her reciprocation of her leadership. Affirmation or resistance to such initiatives become ways in which others communicate their views about one's fitness for leadership and can respectfully encourage or discourage subsequent assertions, which may affect one's sense of self as a leader. For example, resistance may diminish one's self-confidence, discouraging further attempts to take leaderful actions, which would ultimately affect how one is viewed as a potential leader (Ibarra et al., 2014; Moorosi, 2020). Thus, individual internalization is shaped by a variety of factors which include personal attributes, leadership development training and experiencing different leadership roles.

Relational recognition

The relational level of leader identity development suggests that, in addition to leader identity internalization, the leader's identity will be stronger if it is relationally recognized (DeRue and Ashford, 2010; Moorosi, 2014). This means that leadership is expressed through an interactive process of roles that are recognized and reciprocated by others. Indeed Moller (2003) affirmed that "through interactions and what the school principals tell they are doing, or not doing, they are making visible who they are or are becoming" (p. 23). Thus, as one's leadership is affirmed or disaffirmed by others within a particular context, leading to a more relational aspect of leadership identity construction (Ely et al., 2011), one's self-identity as a leader may be strengthened (DeRue and Ashford, 2010) encouraging one to seek further opportunities for leadership development. As Moller (2003) and later Shollen (2018) established, this affects men and women differently, given that traditionally people have held different perceptions about men and women as leaders.

DeRue and Ashford (2010) argued that the construction of leadership, and that of a leadership identity, are essentially about constructing a reciprocal relationship that must be affirmed by the parties concerned. Within the school leadership context, this would involve school principals and their staff working together and affirming each other through their relationship. Leithwood et al. (2008) reported that, when school principals' identities are sound, staff reported improved motivation and commitment, while staff in Jacobsen et al. (2007) study reported their contributions being acknowledged. Cunliffe and Eriksen (2011) argued that, in this relational process, organizations are understood as "communities of people and conversations" (p. 1431), to be engaged and related to. It is the "dialogical relationships with others—both in terms of how one is acknowledged or not acknowledged by others" (Davids, 2018, p. 171) within these organizations that shape identity. Fletcher (2003) has argued that relational forms of leadership, which include empathy, community and collaboration, are "post heroic" and associated with femininity and are often invisible and unrewarded. Indeed, the South African women principals in Davids (2018) study resorted to aggressive masculine styles of leadership to gain respect from colleagues. Moorosi (2020) also found that women were socialized in masculine ways. It is within this relational process of leader identity development that transformational leadership emerges from the interaction between leaders and followers (Yeager and Callahan, 2016). These post heroic theories of leadership are relational and foreground the interaction between leaders and followers but arguably, they ignore the inherent power issues that lead women to constructing their leader identities in masculine terms.

Collective endorsement

The third level is that of "collective endorsement" (DeRue and Ashford, 2010, p. 629), wherein one is seen as part of a larger social group. It is at the collective level that social identities such as gender, race/ethnicity and culture interact with the persona identified at the individual level, shaping and strengthening their leader identity. Karelaia and Guillén (2014) research demonstrates that women's positive social identities play a significant role in psychologically adjusting to the requirements of both personal and professional lives. In the school context, a school principal "has to gain the endorsement of the broader social environment as the leader of a particular school" (Tubin, 2017, p. 791). This involves being part of the broader social groups that shape social identity. In Lumby and Azaola (2014) study, women principals identified "mothering" as a contributor to their professional identities. A mothering identity in this context made women principals "belong" to a social group that offered more than just principalship.

DeRue and Ashford (2010) argue that collective endorsement may reinforce related identities, strengthening leader identity. In a study of both male and female principals in South Africa, Mpungose (2010) concluded that both male and female principals were inspired by role models whose involvement helped them in constructing their identities as school leaders. This literature suggests that individuals with a strong and integrated leader identity would be more motivated to engage in leader development and to exercise leadership than individuals not holding a leader identity, or those whose leader identity is less important to them. In another South African study, Davids (2018) found that women principals did not identify as leaders despite their success in leadership. Their sense of self-identity did not fit in with the norm of what principalship was in that context. Indeed, Ibarra et al. (2014) stated that leader identity should be more closely linked to career development particularly for women leaders. Women who view themselves as leaders may actually bring themselves to think like leaders and seek opportunities for leadership development, thereby strengthening their leader identity. Davids (2018) suggested that women leaders must be prepared to disrupt the situated frameworks of their leadership so that they are able to perceive and be perceived as leaders.

Toward inclusive leadership identity theories

What this presentation so far illustrates is that leadership and socio-demographic identities are closely entwined (Sanchez-Hucles and Davis, 2010), despite their limited attention in the literature, and particularly in the literature on women leaders in educational leadership. Leaders' social identities relative to their followers, influence their effectiveness and the extent to which they feel able to enact their leader identity. However, the extent to which this affects women leaders in various school contexts is questionable. In fact, most of the evidence from the literature suggests that women may not be able to bring their whole selves into leadership, arguably because of the dominant masculine cultures in which they work. Fletcher (2003) argues that the shift from self to "self-in-relation to" (p. 6) is useful as it is aligned with current notions of leadership. However, it is at same time complex as it is related to systemic gender dynamics in the workplace. van Knippenberg (2011) work suggests that prototypical leaders (typical or exemplary

representatives of a group) are liked more and have more positive relationships with subordinates and colleagues. This increases the likelihood that people will comply with requests, further reinforcing leader-follower relationships and strengthening the reciprocal claiming and granting process that strengthens leadership identity. However, such compliance is attributed to masculinity and not femininity. Van Knippenberg (2011) argued that the gendered group norms “render leader group prototypicality contingent on leader gender”, (p. 1081) which is more likely to be male because of male domination in organizational leadership.

The experience of many women leaders through the review of the literature suggests that women find themselves “disowning” their other identities because they just do not work to their advantage. Indeed, what has been labeled the feminine advantage to leadership, which makes women better transformational leaders for organizational effectiveness (Eagly and Karau, 2002; Fletcher, 2003), is in this sense not becoming an advantage for women. In this case, an assumed prototype of leadership is matched by men (and not women) and this is evident across many cultures in both western and non-western contexts (see Showunmi et al., 2015). Thus, the leadership perceptions that colleagues hold of minority group members may be weaker as they do not so easily fit into the mental representations held by people of these roles. For example, incongruence between female stereotypical attributes and the generally masculine environment at the top of many organizations may contribute to prevailing “glass ceiling” or “concrete ceiling” effects faced by women in the workplace (Eagly and Karau, 2002).

Showunmi et al. (2015) work illustrated to this notion of social identity focusing on gender, working with women in the western context. Their study sought to explore whether one’s own identity impacts the way in which women lead. They established that women hid their identity as they climbed the ladder to success, in order to fit into the organizations. They concluded that placing culture at the center of leadership development enables us to fully acknowledge what we bring to the table as leaders. Debebe and Reinert (2014) argued that the impact of not being able to bring one’s whole self into leadership gives a one-sided story of a leader’s identity and does not help develop leadership potential to its full capacity. Suppressing their [invisible] identities, which women were doing by not acknowledging who they are argue can happen consciously or subconsciously (Debebe and Reinert, 2014). Debebe and Reinert problematized the inability to bring one’s whole identity self as identity “miniaturization of self” (p. 290) that constrains leadership potential. Indeed, as the authors argued, “identity miniaturization poses a problem for effective leadership development as it constrains potential leadership qualities that do not fit with the miniaturized self” (p. 290). To further illustrate the occurrence of miniaturized identity, Lumby (2008) study of South African women found that women did not use gender as part of their collective identity. She concluded that this act of “disappearing gender” diminishes women’s ways of defining their leadership. In another study, Lumby and Azaola (2014) contended that “women taking up a school principal role may therefore face persistent and prescriptive stereotypes which mean, whether competent or not, nurturing or not, they will be transgressing one prescription or another, as woman or leader” (p. 16). Showunmi et al. study involved participants who identified as “women of color” in England; they were also found to “disappear” their culture and ethnicity because these identities did not fit in with the white masculine ways of leading within their organizations. Here, intersectionality is at play, revealing the disadvantage of multiple identities that are miniaturized.

Thus, this multidimensional approach to leader identity development helps reveal the shifting views of self from personal to collective, suggesting that identity work can be an iterative and co-constructed process that is fluid and non-linear (Ibarra et al., 2014). Sinclair (2011) suggests that identity work is active construction which involves leaders as “both authors of and objects in identity production” (p. 512), as leaders and followers are both involved [albeit to varying degrees] in identity work. Collinson (2006) found that followers are often a neglected “fantasy” in the process of identity construction. However, Sinclair (2011) asserts that central to identity work is the need to go back into backgrounds, histories and childhood, which is usually intended to help individuals begin to “unpack and work consciously with their beliefs, practices and assumptions about authority and leadership” (p. 512). It is indeed in the unpacking of women leaders’ stories that followers’ roles come to life, especially through the reciprocation of women’s leadership. Moorosi (2020) work reveals the significance of values in these childhood histories that the women leaders use to construct their leader identities. As a social and relational process of identity development, gender affects perceptions that people hold about women leaders and is a characteristic that is visibly noticeable and is thus central to leadership identity development (Moller, 2003; Shollen, 2018). Leadership is more typically associated with males, who are most often presumed to be competent leaders who are confident—women are more likely to question their ability to identify as a leader and may be less likely to be viewed by others as a leader (Shollen, 2018).

The literature on leadership identity development suggests that there are different experiences between men and women, which ultimately make the leadership development process a gendered career path (Ely et al., 2011; Moorosi, 2010, 2014). Given this difference, Boatwright and Egidio (2003) argued that women tend to view their leadership efficacy differently, necessitating an understanding of women’s leadership identity development processes. The question is how women leaders position this identity development and what identities are brought in. Debebe and Reinert (2014) reference to identity miniaturization becomes relevant in taking this work forward. Leading by example and acknowledging their values, women are showing signs of bringing their whole selves to leadership, which needs to be consciously encouraged as part of leadership development. The benefit of the latter would not only be on our ability to design effective leadership development programs, but also on the improvement of our ability to provide appropriate support for men and women, providing them opportunities to develop an identity of true self that exploits their full leadership potential.

The question of whether leadership identity development straddles at least three theoretical explanations of leadership behavior also remains and needs to be fully explored, as Brungardt (1996) suggested. To attempt an explanation, the personal level of identity development does seem to draw support from in the trait theory due to its emphasis on personal attributes of leaders. Here, there is an apparent interaction between the leaders’ personality as a precursor for how they internalize leadership identity. Behavioral

theories on the other hand explain leaders' relational ability, which is akin to group performance, while relationship theories such as transformational leadership emphasizes the "leader's role in the creation of culture" (Brungardt, 1996, p. 83). Indeed, this seems to present the integration of all three theoretical explanations as a good platform for leadership identity. Fletcher (2003) argued that the post heroic dilemma needs more scrutiny as to the extent of its benefit to women, which remains questionable. What implications does this hold for leadership development? Brungardt (1996) warned that leadership development should not be narrowly construed to imply only programs of training. He suggested that leadership development should include development activities throughout the lifetime of potential leaders, leadership education, and leadership training. The next section explores what this might mean in an educational context.

Implications for leadership development

The literature on leadership identity development programs (e.g., Komives et al., 2005; McKenzie, 2018) suggests that leadership identity is usually developed from those programs. Eagly and Karau (2002) suggest that because leadership is typically cognitively associated with male, women more likely question their ability to identify as a leader and may be less likely to be viewed by others as a leader. Recent studies show that when women have a more prominent and central leader identity, they are more likely to intentionally seek out opportunities to develop as leaders (Komiti and Moorosi, 2020). These women were active in networks and had informal mentoring relationships that supported their career development. This suggests that leadership development must be strategically placed within the career path and entail different activities.

In this sense, leadership development activities would be most effective if they enabled individuals to understand themselves within their contexts so that they are able to address issues that are unique to them and bring their whole identities, cultures and values into leadership. Ibarra et al. (2014) observed that when they plan for leadership development, most organizations tend to focus exclusively on the acquisition of new skills, which is not sufficient in and of itself. They argue that learning new leadership skills must be accompanied by a growing sense of one's identity as a leader. This, they argue, can be achieved by a greater understanding of second-generation bias, creating safe spaces for leadership identity development, where women can learn, experiment and create communities, and by encouraging women to anchor their leadership purpose—leading by example. Without these features, skills are developed within a vacuum, and, within a context that is less supportive of women as leaders and one that does not encourage others to recognize and accept women's efforts, women leaders are not likely to succeed as leaders in their leadership positions. Shah (2010) stated that,

people from diverse philosophical, ideological and faith backgrounds conceive and perceive educational leadership differently, particularly across the gender divide, drawing upon their beliefs, values and knowledge sources (p. 28).

Thus, as a multiple identity approach suggests, women can bring their holistic selves into the construction of identity in leadership. The women leaders can see their leadership in multiple identities shaped by their familial values, communities and socio-political contexts (Moorosi, 2020; Moorosi and Grant, 2018). They can bring mothering (Lumby and Azaola, 2014); role modeling (Moorosi et al., 2018), spirituality (Nkomo and Ngambi, 2009), religiosity (Shah, 2010) into leadership and leadership identity development. A successful approach to leadership development would thus be one that acknowledges the multiplicity of identities, values, cultures and roles. It is this expansive role of leadership identity development that needs to be considered in leadership development research so that it informs leadership understanding in context.

Conclusion

Through this conceptual analysis, we have argued that leadership is best understood as a social construction through which an identity of leadership is constantly negotiated. The analysis shows that the current understanding of leadership is influenced by hierarchical, masculine and relational notions, that despite the contemporary nature of relational leadership, leadership identity is a gendered space that so far does not appear to benefit women. The literature shows evidence of gender as a disappearing social identity and women leaders not enjoying the benefit of the feminine advantage. This calls for the expansive understanding of the context of leadership that might help facilitate the eradication of the masculine stereotypical image and identity of a leader.

References

- Alston, J.A., 2012. Standing on the promises: a new generation of Black women scholars in educational leadership and beyond. *Int. J. Qual. Stud. Educ.* 25 (1), 127–129.
- Ayman, R., Korabik, K., 2010. Leadership: why gender and culture matter. *Am. Psychol.* 65 (3), 157–170.
- Bennis, W., 2009. *On becoming a leader*. Basic Books, New York.
- Billsberry, J., 2009. The social construction of leadership education. *J. Leadersh. Educ.* 8 (2), 1–10.
- Boatwright, K.J., Egidio, R.K., 2003. Psychological predictors of college women's leadership aspirations. *J. Coll. Stud. Dev.* 44, 653–669.

- Brewer, M.B., Gardner, W., 1996. Who is this "we"? Levels of collective identity and self-representations. *J. Pers. Soc. Psychol.* 71 (1), 83–93.
- Brungardt, C., 1996. The making of leaders: a review of the research in leadership development and education. *J. Leadersh. Stud.* 3 (3), 81–95.
- Carroll, B., Levy, L., 2010. Leadership development as identity construction. *Manag. Commun. Q.* 24 (2), 211–231.
- Chin, J.L., 2010. Introduction to the special issue on diversity and leadership. *Am. Psychol.* 65 (3), 150–156.
- Cole, B.A., 2009. Gender, narratives and intersectionality: can personal experience approaches to research contribute to "undoing gender"? *Int. Rev. Educ.* 55 (5), 561–578.
- Coleman, M., 2012. Leadership and diversity. *Educ. Manag. Adm. Leadersh.* 40 (5), 592–609.
- Collinson, D., 2006. Rethinking followership: a post-structuralist analysis of follower identities. *Leadersh. Q.* 17, 179–189.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics and violence against women of color. *Stanford Law Rev.* 43, 1241–1299.
- Cunliffe, A.L., Eriksen, M., 2011. Relational leadership. *Hum. Relat.* 64 (11), 1425–1449.
- Dalakoura, A., 2010. Differentiating leader and leadership development: a collective framework for leadership development. *J. Manag. Dev.* 29 (5), 432–441.
- Davids, N., 2018. When identity and leadership intersect: the experiences of six female principals in South Africa. *Afr. Educ. Rev.* 15, 157–174.
- Day, D.V., Harrison, M.M., Halpin, S.M., 2009. An Integrative Approach to Leader Development: Connecting Adult Development, Identity, and Expertise. Routledge, New York, NY.
- Day, D.V., 2001. Leadership development: a review in context. *Leadersh. Q.* 11, 581–613.
- Day, D.V., 2011. Integrative perspectives on longitudinal investigations of leader development: from childhood to adulthood. *Leadersh. Q.* 22, 561–571.
- Debebe, G., Reinert, K.A., 2014. Leading with our whole selves: a multiple identity approach to leadership development. In: Miville, M.L., Ferguson, A.D. (Eds.), *The Handbook of Race-Ethnicity and Gender in Psychology*. Springer, New York, pp. 271–293.
- DeRue, D.S., Ashford, S.J., 2010. Who will lead and who will follow: a social process of leadership identity construction in organizations? *Acad. Manag. Rev.* 35 (4), 627–647.
- Dvir, T., Eden, D., Avolio, B.J., Shamir, B., 2002. Impact of transformational leadership on follower development and performance: a field experiment. *Acad. Manag. J.* 45 (4), 735–744.
- Eagly, A.H., Carli, L.L., 2007. *Through the Labyrinth: The Truth About How Women Become Leaders*. Harvard Business School, Boston, MA.
- Eagly, A.H., Karau, S.J., 2002. Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* 109, 573–598.
- Ely, R.J., Ibarra, H., Kolb, D., 2011. Taking gender into account: theory and design for women's leadership development programs. *Acad. Manag. Learn. Educ.* 10, 374–393.
- Fletcher, J.K., 2003. The Paradox of Post Heroic Leadership: Gender Matters. CGO Working Paper No. 17. Center for Gender in Organizations, Boston, MA.
- Grint, K., 1997. *Leadership: Classical, Contemporary, and Critical Approaches*. Oxford University Press, Oxford.
- Hall, D.T., 2004. Self-awareness, identity, and leader development. In: Day, D.V., Zaccaro, S.J., Halpin, S.M. (Eds.), *Leader Development for Transforming Organizations: Growing Leaders for Tomorrow*. Erlbaum, Mahwah, NJ, pp. 153–176.
- Hite, L.M., 2007. Hispanic women managers and professionals: reflections on life and work. *Gen. Work. Organ.* 14 (1), 20–36.
- Ibarra, H., Wittman, S., Petriglieri, G., Day, D.V., 2014. Leadership and identity: an examination of three theories and new research directions. In: Day, D.V. (Ed.), *The Oxford Handbook of Leadership and Organizations*. Oxford University Press, New York, NY, pp. 289–305.
- Iwowo, V.A., 2015. Leadership in Africa: rethinking development. *Person. Rev.* 44 (3), 408–429.
- Jacobson, S., Brooks, S., Giles, C., Johnson, L., Ylimaki, R., 2007. Successful leadership in three high-poverty urban elementary schools. *Leadersh. Pol. Sch.* 6 (4), 1–27.
- Karelaia, N., Guillén, L., 2014. Me, a woman and a leader: positive social identity and identity conflict. *Organ. Behav. Hum. Decis. Process.* 125, 204–219.
- Komiti, M.C., Moorosi, P., 2020. Career development of women principals in Lesotho: influences, opportunities and challenges. *J. Educ.* 79, 95–114.
- Komives, S.R., Owen, J.E., Longenecker, S.D., Mainella, F.C., Osteen, L., 2005. Developing a leadership identity: a grounded theory. *J. Coll. Student Dev.* 46 (6), 593–611.
- Learmonth, M., Morell, K., 2017. Is critical leadership studies 'critical'? *Leadership* 13 (3), 257–271.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leadersh. Manag.* 28 (1), 27–42.
- Lord, R.G., Hall, R.J., 2005. Identity, deep structure and the development of leadership skill. *Leader. Q.* 16, 591–615.
- Lumby, J., Azaola, M.C., 2014. Women principals in South Africa: gender, mothering and leadership. *Br. Educ. Res. J.* 40 (1), 30–44.
- Lumby, J., 2008. Disappearing gender: choices in identity. In: Soberhart, H. (Ed.), *Women Leading Education Across the Continents: Sharing the Spirit, Fanning the Flame*. Rowman and Littlefield, Lanham, pp. 29–38.
- Lumby, J., 2013. Distributed leadership: the uses and abuses of power. *Educ. Manag. Adm. Leadersh.* 41 (5), 581–597.
- McKenzie, B.L., 2018. Am I a leader? Female students' leader identity development. *J. Leadersh. Educ.* 17, 1–18.
- Mestry, R., Schmidt, M., 2012. A feminist postcolonial examination of female principals' experiences in South African secondary schools. *Gen. Educ.* 24 (5), 535–551.
- Middlehurst, R., 2008. Not enough science or not enough learning? Exploring the gaps between leadership theory and practice. *High Educ. Q.* 62 (4), 322–339.
- Miscenko, D., Day, D.V., 2016. Identity and identification at work. *Organ. Psychol. Rev.* 6, 215–247.
- Moller, J., 2003. Gender and leadership identities—negotiated realities for women as school principals. *Improv. Sch.* 6 (3), 23–44.
- Moorosi, P., Fuller, K., Reilly, E., 2018. Leadership and intersectionality: constructions of successful leadership among black women school principals in three different contexts. *Manag. Educ.* 32 (4), 152–159.
- Moorosi, P., Grant, C., 2018. The socialization and leader identity development of school leaders in Southern African countries. *J. Educ. Adm.* 56, 643–658.
- Moorosi, P., 2010. South African female principals' career paths: understanding the gender gap in secondary school management. *Educ. Manag. Adm. Leadersh.* 38, 547–562.
- Moorosi, P., 2014. Constructing a leader's identity development through a leadership development programme: an intersectional analysis. *Educ. Manag. Adm. Leadersh.* 42, 792–807.
- Moorosi, P., 2020. Constructions of leadership identities via narratives of African women leaders. *Front. Educ.* 5 (86). <https://doi.org/10.3389/educ.2020.00086>.
- Mpungose, J., 2010. Constructing principals' professional identities through life stories: an exploration. *S. Afr. J. Educ.* 30, 527–537.
- Murakami, E.T., Torsen, M., 2017. Female secondary school principals: equity in the development of professional identities. *Educ. Manag. Adm. Leadersh.* 45 (5), 806–824.
- Nkomo, S.M., Ngambi, H., 2009. African women in leadership: current knowledge and a framework for future studies. *Int. J. Afr. Renaiss. Stud.* 4 (1), 49–68.
- Parker, P.S., 2005. *Race, Gender and Leadership: Re-envisioning Organizational Leadership Form Perspectives of American Women Executives*. Routledge, Mahwah.
- Petriglieri, G., 2011. Identity workspaces for leadership development. In: Snook, S., Nohria, N., Khurana, R. (Eds.), *The Handbook for Teaching Leadership*. SAGE, Thousand Oaks, CA, pp. 295–312.
- Pompper, D., 2007. The gender-ethnicity construct in public relations organizations: using feminist standpoint theory to discover Latinas' realities. *Howard J. Commun.* 18 (4), 291–311.
- Richardson, A., Loubier, C., 2008. Intersectionality and leadership. *Int. J. Leadersh. Stud.* 3 (2), 142–161.
- Sanchez-Hucles, J.V., Davis, D.D., 2010. Women and women of color in leadership: complexity, identity, and intersectionality. *Am. Psychol.* 65 (3), 171–181.
- Sethodi, I.I., 2020. Deconstructing Leadership: Engaging Leading Through a Socially Constructed Process. <https://doi.org/10.5772/intechopen.90630>.
- Shah, S.J.A., 2010. Re-thinking educational leadership: exploring the impact of cultural and belief systems. *Int. J. Leadersh. Educ.* 13 (1), 27–44.
- Shapira, T., Arar, K., Azaiza, F., 2011. They didn't consider me and no-one even took me into account: female school principals in the Arab education system in Israel. *Educ. Manag. Adm. Leadersh.* 39 (1), 25–43.
- Shollen, S.L., 2018. Women's leader identity development: building a team for the journey. In: Tan, S.J., DeFrank-Cole, L. (Eds.), *Women's Leadership Journeys: Stories, Research and Novel Perspectives*. Routledge, New York, NY, pp. 162–177.
- Showunmi, V., Atewologun, D., Bebbington, D., 2015. Ethnic, gender and class intersections in British women's leadership experiences. *Educ. Manag. Adm. Leadersh.* 44 (6), 917–935.
- Sinclair, A., 2011. Being leaders: identities and identity work in leadership. In: Bryman, A., Collinson, D., Grint, K., Jackson, B., Uhl-Bien, M. (Eds.), *The SAGE Handbook of Leadership*. Sage, Los Angeles, CA, pp. 508–517.

- Skinner, S., 2014. Understanding the importance of gender and leader identity formation in executive coaching for senior women. *Coaching* 7 (2), 102–114.
- Slaughter, A., 2015. *Unfinished Business: Women Men Work Family*. Oneworld Publications, London.
- Stogdill, R.M., 1974. *Handbook of Leadership: A Survey of Theory and Research*. Free Press, New York.
- Tubin, D., 2017. Leadership identity construction practices: the case of successful Israeli school principals. *Educ. Manag. Adm. Leadersh.* 45 (5), 790–805.
- van Knippenberg, D., van Knippenberg, B., De Cremer, D., Hogg, M.A., 2004. Leadership, self, and identity: a review and research agenda. *Leadersh. Q.* 15, 825–856.
- Van Knippenberg, D., 2011. Embodying who we are: leader group prototypicality and leadership effectiveness. *Leadersh. Q.* 22 (6), 1078–1091.
- Warner, L.R., 2008. A best practices guide to intersectional approaches in psychological research. *Sex. Roles* 59 (5), 454–463.
- Yeager, K.L., Callahan, J.L., 2016. Learning to lead: foundations of emerging leader identity development. *Adv. Dev. Hum. Resour.* 18 (3), 286–300.

The emergence of social justice leaders: identity formation of new principals

Izhar Oplatka, School of Education, Tel Aviv University, Tel Aviv, Israel

© 2023 Elsevier Ltd. All rights reserved.

Introduction	438
Leadership identity formation for SJ	439
The importance of leadership identity building	439
Forms of leadership identity building	440
The major factors of leadership identity formation for SJ	441
Pre-service and in-service training of principals	441
Mentors/role models	442
Belonging to a professional group	442
Education policy	442
Personal characteristics	443
Barriers to leadership identity construction for SJ	443
Conclusion	443
References	444

Introduction

The concept of social justice (SJ) has a host of meanings, and therefore scholars and researchers tend to analyze different dimensions of leadership for SJ such as inclusion, caring relationships, moral engagement, and so forth (Oplatka, 2013; Furman, 2012). Yet, although there is no one dominant, holistic theory of SJ, it is widely accepted that this term refers to a society which values diversity by providing all people and groups with respect, dignity, fairness, basic rights, and opportunities (Arar, 2015; Flores et al., 2014).

In light of this epistemological nature of SJ concept, Lipman (2004) framed four imperatives that create a vision of SJ education: “(1). *Equity*: all children should have equal opportunities and rights, but special efforts must be made to overcome past injustice and inequalities of race, gender, and class; (2). *Agency*: education should allow students to take action to change their personal situations and social injustice ...; (3). *Cultural Relevance*: teachers should use students’ cultures to support academic success. They should work to develop sociopolitical consciousness and challenge unjust conditions; (4). *Critical Literacy*: schools should be a place in which students can examine knowledge and their own life experiences critically. The curriculum should be grounded in students’ experiences and challenge official knowledge that distorts the histories of marginalized groups (p. 17)”.

This vision is very similar to common conceptualizations of leadership for SJ that construct this model of leadership as action-oriented and transformative, committed and persistent, inclusive and democratic, relational and caring, reflective, and oriented toward a socially just pedagogy (Furman, 2012). In practice, according to Theocharis (2007, p. 222), leaders for SJ “advocate, lead, and keep at the center of their practice and vision issues of race, class, gender, disability, sexual orientation, and other historically and currently marginalizing conditions in the United States”. They also identify inequities in school policies, procedures, and practice and promote inclusive practice and equitable access to the curriculum. To this end, they assist teachers in planning curriculum and instruction, which is inclusive of students and families’ perspectives and experiences and challenge. They also modify staff’s negative conjectures and misperceptions about diverse students, families, and communities and thereby alter values (Guerra et al., 2013).

In their review of the scholarship about leadership for SJ, Nicholson et al. (2020, p. 96) indicated that this leadership focuses on the experiences of marginalized groups and inequality in schools. They sums up the major themes that appear in k-12 social justice leadership:

(a) the need for maintaining persistence, commitment and courage in one’s work ...; (b) building inclusive and participatory processes that engage students, families and educators and widen the voices contributing to decision-making in schools ...; (c) valuing caring relationships and continuously investing time and effort towards building and strengthening relationships ...; (d) engaging in ongoing critical self-reflection and personal work to gain clarity on one’s own perspectives, motives and biases and to expand and challenge them ...; and (e) continuously analyzing the evidence of student learning to determine where inequities exist that require immediate and responsive intervention

This comprehensive analysis of leadership for SJ provides us with the major concepts and practices leaders for SJ are characterized with in the literature and, consistent with Furman (2012), highlights the importance of this type of leadership in our neoliberal world. New leaders may need to develop the capacities required for individuals to engage in the complex, manifold tasks of leaders for SJ who strive to promote issues of social justice, equity, and equality in their school.

The aim of this article is to propose a research-informed analytical framework of the process through which new principals may undergo when constructing, explicitly and implicitly, an identity of leadership for SJ. Two points merit highlighting in the framework: (1) the emergence of awareness to SJ and the need to lead schools more ethically. (2) The factors affecting the formation of leadership identity for SJ in early career. After all, hierarchical status alone is insufficient to explain how and when the followers (or subordinate) perceive others as leaders (Marchiondo et al., 2015). Underlying my conjecture, then, is that new educational leaders for SJ should not only claim their role as “the leader” but also as the persons who led the school toward SJ.

According to the literature about early career in management outlined below, new leaders are likely to construct a leadership identity. There are several terms for this process of identity development in the literature, including: identity construction—the iterative and reciprocal process of identity shaping (DeRue and Ashford, 2010); identity formation—the product of interrelatedness between the context and the individual’s developmental and learning processes (Flum and Kaplan, 2012); and identity work—the process of engagement in forming, restoring, maintaining, strengthening or reversing the identity (Sveningsson and Alvesson, 2003). In broad terms, identity construction theory centers on the dynamics by which leadership emerges, as well as on leadership processes as they appear among all organizational members, rather than only among formal leaders (DeRue and Ashford, 2010). Accordingly, individuals have an active role in defining, redefining, and modifying their identities. In fact, they play an active role in molding the varied components of their identity, but certainly also in shaping social identity both in the subjective and supra-individual sense of the term (Postmes et al., 2005).

The formation of leadership identity is not limited to leadership development programs. Rather, it is formed and constructed during the first years in leadership positions. An understanding of the process of leadership formation and its factors during early career in a more holistic manner may help us better understand how leadership for SJ might be developed and molded. This in turn may not only enrich our knowledge of early career in educational leadership but also help design leadership preparation program contents that set the stage for the emergence of leadership for SJ in our schools.

Note, however, that in no way do I claim that every new principal is likely to construct an identity of leadership for SJ. It depends on many personal and organizational elements, some of them are discussed in this paper. In this sense, the current article presents an analytical framework of leadership identity building for SJ that some new principals may experience. In the next part, I analyze different aspects of leadership identity construction, in general, and of that of leaders for SJ, in particular.

Leadership identity formation for SJ

In early career, new educational leaders undergo distinctive career tasks and experiences (Oplatka, 2012). Thus, they may feel a “reality shock”, surprise, and a lack of full understanding of what an educational leader should do in practice. Their staff may be hostile toward them in the first several months after their arrival, and the new leaders are likely to face enthusiasm vs. stress, loneliness, professional insecurity and fear of failure. Likewise, they allocate most of their time to engage in technical rather than instructional work aspects, resulting in some sense of frustration and a lack of professional fulfillment. Above all, they are required to deal with multiple complex tasks and unanticipated negative events before acquiring sufficient practical managerial knowledge and skills.

Concurrently, new educational leaders face several unique career tasks as follows:

1. Mastering the basic skills and competencies needed to perform effectively and efficiently in principalship.
2. Developing an understanding of the school structure, culture and environment in order to stabilize one’s position in the school.
3. Becoming accepted as a competent educational leader whose staff and stakeholders trust his/her expertise and ability to run their schools.
4. Devising a new school vision with staff and constituencies leading to a new change program for the school.
5. Coping successfully with stress, frustration and negative emotions that accompany the establishment of a new role (Oplatka, 2012, p. 147).

The performance of these particular tasks in early career is intended to enable new leaders to establish themselves successfully in their role before moving into the next career stage and becoming experts in educational administration. One may see these tasks as part of the process of leadership identity formation in early career.

The importance of leadership identity building

It is widely accepted that an educational leader is not granted as a leader by school members and stakeholders only because s/he has been officially appointed to a leadership position in the school. In fact, the leader has to undergo a process of leadership identity formation in order to be accepted by school staff and by him/herself as a “leader” (DeRue and Ashford, 2010). In their leadership identity construction model, DeRue and Ashford (p. 631) suggest that “individuals ‘claim’ an identity and others affirm or ‘grant’ that identity as the underlying process by which leader and follower identity become socially constructed and form the basis of leader-follower relationship”. This claiming-granting process includes verbal/non-verbal and direct/indirect actions. For instance, a direct verbal act is making a statement that one is a leader (e.g., I am the principal here!), whereas an indirect verbal tactic is name-dropping with regard to knowing other notable leaders and having relationships with them (e.g., sitting in the principal’s higher chair, running the management meeting).

Based on this model, leadership can be defined as a process of mutual influence that change across time and situations as individuals claim and grant leader and follower roles. Leadership in organizations emerges, then, as a result of individuals claiming (for themselves) and granting (to others) leader identities, which we frame as a negotiated process of relational identity construction (Marchiondo et al., 2015, p. 903). Thus, when a teacher claim to be a leader for SJ in the school, s/he needs to negotiate with other teachers verbally and indirectly before they could affirm him/her as such a leader and follow his/her vision of SJ, honesty, and so on. For example, as part of the negotiation, the new leader might convince the teachers of his/her commitment to diversity, inclusion, and equality and in his/her capacity to promote justice in the school.

Through leadership claiming, then an emergent educational leader negotiates, explicitly or implicitly, a leadership role for him/herself (Kopelman, 2014). The emergent leader may claim this role by explicitly negotiating during a conversation—proclaiming that s/he will take the lead in a meeting or project about ethical consideration in staff meetings, for example. Another novice leader might instead rely on implicit negotiations, communicating indirectly to accomplish similar claiming, for example by sitting at the head of the table in a meeting with parents that handles issues of social injustice in the school (note, in many cases the head of the parent association could head the table).

Thus, claims serve to establish oneself as a leader within a particular group at a particular time (Marchiondo et al., 2015). For the purpose of this article, though, the emergent educational leader for SJ should not only claim his/her role as the leader but also as the person who leads the school toward SJ. For example, and consistent with McKenzie et al. (2006), it is not enough for the emergent leader to have developed a consciousness of SJ, equity, equality and the like, but rather s/he should encourage and support teachers in their attempts to develop it as well. Leading the staff toward SJ-oriented teaching strengthens the emergent leader's request to be acknowledged as a leader for SJ, at least in part because the teachers who consequently improved their skills to teach diverse students effectively may grant him/her as their leader.

More specifically, the act of granting a leader identity can precede a claim (e.g., “would you like to take the lead on this project?”) or may arise in response to another's claim (e.g., “yes, I agree that you should take the lead on this project”) (Marchiondo et al., 2015). However, claiming to be the leader is insufficient; the followers should accept the claims based on their perceptions of the “right” leader and their judgments of the ability of that person to be their leader. Note, followers' responses to another's leadership claiming and granting are important in molding leadership perceptions of both parties involved in the leadership identity construction process.

Based on Conger (1992) typology of leadership development, it is likely that granting a leader identity to the person who claims to be a leader for SJ depends, largely, on his/her ability to display skills and competencies that are expected from such a leader (e.g., the ability to run a reflective debate about the influence of the curriculum on unprivileged students). The granting aspect in this case is composed of many feedbacks that the emergent leader for SJ receive from school members and stakeholders who are committed to equity, equality, and just education. A support to the importance of feedbacks we gain from the next quote:

His self-perception of teacher leadership was reinforced by school administrators, who recognized his skills and gave him the responsibility to lead school science department activities. Feedback from his colleagues also made him realize that he had “good ideas” and could be a credible source of information whenever “they are introducing something new”.

Somnath and Hnuscin (2017, p. 362).

To further this issue, let us look at the components of equity consciousness as McKenzie et al. (2006) constructed:

By equity consciousness, we mean that teachers are aware of, accept, and act on four central beliefs: that all children—except only a very small percentage, that is, those with profound disabilities—are capable of high levels of academic success; that this academic success equitably includes all student groups, regardless of race, social class, gender, sexual orientation, learning differences, culture, language, religion, and so forth; that the adults in schools are primarily responsible for seeing that all children reach this success; and that traditional school practices result in inequity for individual students and groups of students and that these must be changed to ensure success for every child (p. 160).

Thus, teachers who have already developed SJ consciousness (and probably work in disadvantaged communities) might expect their new leader to develop such consciousness before they grant him/her as a leader for SJ.

Forms of leadership identity building

Yet, to be granted by the school members and stakeholders as an educational leader for SJ, the educational leader has to undergo a process of identity formation. From the career literature we learn that new managers' central concern is to become established in their new position, just as other employees (Hall, 2002). Indeed, they usually experience a kind of identity transformation as they learn how to think, feel and value as managers rather than as subordinates (Hill, 2003). This includes “taking charge,” i.e., the process of learning and taking action that a manager goes through until he or she has mastered a new assignment in sufficient depth to run the organization, and “taking hold,” i.e., developing an understanding of the new situation and an initial set of priorities and expectations with key subordinates as a basis for effective working relationships (Gabarro, 2007).

Hill (2003) work is especially insightful in illustrating the life of new managers in organizations. Briefly, she found that the transition to managerial positions is not constrained to acquiring competencies and building relationships, but chiefly includes a transition from specialist and doer who directly performs specific technical tasks done mostly by one's own efforts to a generalist and agenda-setter who initiate diverse tasks and get them done through others. During this identity construction new managers make the psychological adjustment required to make sense of their new responsibilities, learn what it means to be a manager, develop interpersonal judgment, gain self-knowledge, and cope with stress and intense emotions associated with their new position and the evolution of a new professional identity. Additionally, new managers in her study worked hard to build effective relationships with their subordinates and were eager to receive feedback from superiors and peers.

But, identity construction is slower moving and shaped by multiple roles taken in various settings (Carden and Callahan, 2007). "Leadership identity" is established when leadership is exhibited across both personal and professional contexts. When leadership identity is formed personally and professionally, it is easier to become comfortable as "leader" in an organizational role (Regan, 2016). For example, creating educational leadership roles seems to be an identity formation process for Emirati women. They are embedded in a multifaceted identity formation and maintenance dynamic which consists of being Muslim, being Emirati, being a woman, and achieving a managerial or leadership role organizationally (Samier, 2015).

Regan (2016) found that leadership identity formation in nonprofit human service organizations includes (a) encountering the possible clash in values between the person and the organizational setting, (b) renegotiating "belonging" with client, staff groups or executive teams that one may be leaving or joining, (c) becoming comfortable in an organizational role through the development of a "provisional self" in which power and authority can be comfortably exercised, and (d) exhibiting leadership behaviors across personal and professional roles (p. 448–449).

In light of Regan's components of identity formation, one may view the process of leadership identity formation for SJ as follows. Firstly, the new leader should balance between his/her values and those in the school and develop consciousness and passion for SJ. Secondly, to become comfortable with his/her ability to lead the schools toward SJ s/he has to develop a managerial self in which power and authority are legitimate and constitute an integral part of becoming a leader for SJ. Thirdly, to fulfill his/her passion for SJ, the new leader ought to be seen by staff and stakeholders as an expert in leading just education that promotes, among other things, equity, equality, respect, fairness, critical thinking, and honesty. Following Flores et al. (2014), I believe the new leader has to develop competencies such as building awareness of one's own biases, prejudice, and stereotypes, building awareness of how one internalizes notions of superiority and inferiority, developing the ability to appreciate other peoples' ways of thinking, and so on. In addition, the new leader for SJ has to acquire knowledge of social inequality, interpersonal and structural forms of oppression, and the impact of societal inequality on peoples' experience of privilege and oppression.

Finally, the new leader has to exhibit leadership behaviors across personal and professional roles that increase his/her image of leader for SJ and promote this issue in the school. For example, s/he has to develop skills to work with diversity and promote equity, engage in difficult dialogs, foster inclusion, and encourage the staff to participate in courses and workshops intended to help them promote SJ in their classroom. The question of what influences the identity construction of leaders for SJ is discussed in the next section.

The major factors of leadership identity formation for SJ

Now let us probe into five factors that may facilitate the emergence of leadership identity for SJ among new educational leaders. Some attention is also given to potential barriers to the emergence of leadership for SJ.

Pre-service and in-service training of principals

The seeds of SJ in the soul of the new leader may be planted in leadership development program that emphasize equity and social justice (prior to the entrance of the leader into his/her office). After all, training for future principals is fundamental to the development of leadership identity, as Cruz-Gonzalez et al. (2020) noted. In this sense, Pounder et al. (2002) indicated that leaders ought to be provided with new analytical skills, knowledge, and dispositions to promote SJ in schools. Therefore, they suggested including field-based inquiry focused on oppression and discrimination in leadership development programs. In addition, they recommended analyzing empirical data regarding racism in schools, examining stereotypes related to oppression, and developing socially just practices in the courses the consist this program. One may assume that these kinds of contents may increase prospective educational leaders' awareness of injustice and inculcate a desire (or even passion) to challenge unjust conditions in the school.

In other words, leadership preparation programs have to encourage future school leaders to think very differently about organizational structures and leadership roles, i.e., to develop a leadership identity that is oriented not only toward neoliberal policies but also toward issues of SJ. Thus, instead of adopting incremental reforms that simply add more layers to existing structures, the construction of leadership identity for SJ is influenced by programs that encourage prospective educational leaders to reconstruct roles and relationships at the school level around a vibrant core purpose focused on SJ and directed at improving student learning. But, according to Cambron-McCabe and McCarthy (2005), this cannot be accomplished without concentrating on teaching and learning. In their view, educational leaders have to possess high quality instructional skills and to be able to raise critical issues concerning equity, race, diversity, and privilege. When the prospective leaders enter their new school, well-equipped with skills and

competencies necessary to lead their school toward SJ, they are more likely to construct leadership identity for SJ. After all, without proper skills and competencies required to be considered by school members and environmental constituencies as an “effective leader”, it is unlikely that new leaders will be able to develop leadership identity effectively.

Mentors/role models

Every new leader needs some practical guidance and support to face successfully with the need to acquire new leadership competencies and skills.

Higgins and Kram (2001) identified developmental networks of mentors and defined them as “a set of people named by a protégé as taking an active interest and action to advance their career by providing developmental assistance”. Thus, when the mentor is also a leader for SJ, it is likely that the new leader will be able to form an identity of such a leader more effectively, because the mentor can help him/her in increasing awareness to SJ and implementing new policies and projects intended to face issues of race, class, gender, disability and so forth in the school.

However, when it comes to identity formation, role models can be seen as a network providing both near and active as well as distant and passive career guidance. Young women in Singh et al. (2006) study perceive senior women and men as remote role models, as well as closer role models such as managers and mentors. Likewise, role models are often considered to be key to the successful development of young aspiring managers. According to Shapiro et al. (1978), role models are individuals whose behaviors, styles and attributes are emulated by others. Note, role models are not familiar with the role model user, nor do they give their permission to be used in this way (Singh et al., 2006).

Role models may have strong impact about leadership identity construction for SJ. Thus, new leaders whose awareness of SJ is high (due to, for example, their principal preparation program) may use a selection of role models who have promoted SJ in their communities. Let's take for example leaders such as Martin Luther King or Maria Teresa. A new leader may be inspired by their life story to build appropriate leadership identities such as the articulation of messages that sharpen social injustice, the ways to harness the efforts of teachers in favor of disadvantaged communities, or the value of inclusion in the staffroom. It is important to note that the new leader has to learn how to lead the staff and convince stakeholders about his/her ability to be an educational leader. When great leaders for SJ constitute a source of inspiration, they “help” the new leader in forming a leadership identity that goes beyond the routine and the known. It is an identity that contrasts common and accepted neoliberal views of education and convince others to adopt and more social-emotional views of education.

Belonging to a professional group

Group belonging is the basis of leadership identity construction. Every group establishes standards and norms and if one excels at the norms and values created by that particular group, s/he acquired “rank” (power, authority, privilege) and is poised to take on leadership (Regan, 2016). Thus, a new educational leader who follows the dominant educational norms and values in the society and in the staffroom demonstrates his/her commitment to the school and, in turn, receive the legitimacy to be the leader. One should bear in mind that the leader is expected more than others to follow entrenched common norms and values and promote them in the group. When it comes to SJ, the new leader should become familiar with the dominant values as regard to justice, equity, equality and so on in the school and in the community and develop managerial strategies to promote them in the classrooms and the community. Furthermore, s/he should value diversity and encourage teachers to adopt socially just pedagogies.

Likewise, leadership identity is formed as a result of social interactions as much as based on personal abilities and traits. Individuals constantly strive to shape their personal identity in organizations, and are being shaped by discursive forces (Saarukka, 2014). Here again, the new leader's identity formation depends largely on those s/he is going to meet in the teacher's lounge, the parent association, the education district and the community. Leadership identity formation is strongly shaped by social interactions with other teachers and stakeholders. For example, if the interactions focus on social injustice and a lack of equal opportunities in the school, they are likely to mold leadership identity for SJ among the new leaders, and vice versa. If the staff ignores issues of SJ and “chooses” to center on the advantaged communities, the process of leadership identity formation for SJ will be slower and include many obstacles.

Education policy

Evidence supports the strong impact of educational policies on the construction of a professional identity in educational leadership (Cruz-Gonzalez et al., 2020). For example, accountability in a form of standards of performance, testing and student achievement as a measure of successful education, characterizing many western countries nowadays, may limit the new leader's autonomy and make it very difficult for him/her to construct leadership identity for SJ.

In contrast, when the policies encourage equity, equality, respect to diversity, and SJ and provide schools with adequate curricula and projects, it is more likely that new leaders will construct leadership identity for SJ. Knowing that their superiors (e.g., superintendents, supervisor, and school boards) support varied notions of SJ in education enhances new leaders' awareness of the importance of SJ in education. It may also drives teachers to perceive a new leader as a leader for SJ because they have been encouraged by the district/local education authority to engage in SJ issue prior to the arrival of the new leader.

This, in turn, assists the new leader develop interpersonal judgment as regard to SJ, gain self-knowledge about SJ procedures in the school, and cope with stress and intense emotions associated with his/her attempts to promote, for example, equal educational opportunities.

Personal characteristics

Leadership identity is very much affected by personal features. According to the principals' narratives, awareness of individual abilities was expressed and related to leadership issues (Saarukka, 2014). In the case of leadership for SJ, new leaders need to have personal features such as flexibility, sensitivity to others, high professional efficacy, respect to others and so on. These characteristics may facilitate their process of leadership identity formation, because they will be capable of being sensitive to disadvantaged students and believe they have the knowledge and skills to increase their achievements and well-being. As leaders, though, they need to believe they can inspire the staff and sweep them to follow the principles of SJ and identify inequalities in education. This belief is being developed a day after day in the process of leadership identity construction.

Likewise, new educational leaders who have experienced social injustice, poverty, inequality and so on in a certain period in their life may be aware of the need to promote SJ in education.

Barriers to leadership identity construction for SJ

From the factors discussed thus far we can conclude that leadership identity construction for SJ depends not only on the new leader's traits and experiences but also on contextual and environmental elements and constituencies. Theoharis (2007) identifies a number of barriers that can obviate SJ work in schools. These include the deficit thinking about marginalized groups that characterizes some schools, the preference of "technical" leadership over moral leadership in many educational systems, national and local policies that work against equality and SJ, and the burden to the individual of engaging in transformative leadership.

Based on Theoharis's analysis, leadership identity formation for SJ may be inhibited when school staff tend to marginalize the importance of equity, caring, and equality in the society or ignore the special needs of students from disadvantaged communities. Similarly, education districts or local education authorities that glorify high achievements and accountability over any other educational goals and needs may in turn decrease the significance of SJ in education. Thus, new leaders may interpret the educational policy as promoting high ranks and therefore concentrate most of their time on the improvement of teaching, learning, and external exams regardless of issues of class, ethnicity, equality and justice.

From the new leaders' perspectives, some may focus to a large extent on technical issues of leadership in their process of leadership identity construction over issues of moral education and SJ in education. This may be caused by their interpretation of current neoliberal ideologies of education as measurement-oriented, i.e., their performance is measured by technical standards rather than moral ones.

Conclusion

This article has attempted to address the issue of SJ leadership identity formation by proposing a research-informed analytical framework for the process and factors of this process and arguing that the central purpose of new leaders for SJ is to develop the capacities for leading teachers toward identifying unjust in the schooling process and promoting SJ. This framework captures the dynamic interplay between leadership identity construction and its personal and contextual factors in the school, and thus has the potential to clarify the process through which educational leaders become leaders for SJ.

New educational leaders need to play an active role in forming the central elements in their emergent leadership identity and in modifying their previous professional identity as teachers. More particularly, they may re-define the elements of SJ in their current identity to become not only advocates of SJ who face social injustice as teachers, but also leaders who led other to do it. Put another way, in the process of leadership identity formation the new leader becomes capable of making teachers aware of SJ and promote it in their classes.

To this end, it is important to emphasize issues of SJ and social injustice in leadership development programs and provide prospective educational leaders with skills required to lead the school toward SJ. But, we shall bear in mind that the practice of leadership for SJ (or the praxis) can be developed only after the new leader entered his/her role. The concept of identity construction sheds light on the process through which a teacher who has adopted critical, inclusive pedagogies and curricula becomes a leader who led others to do the same (and more than that). Practically speaking, the new educational leader needs to learn how to think, feel, and act as a leader rather than as teacher as well as to learn how to lead others particularly toward SJ in the school. This is not an easy task indeed.

While culture and society play a key role in our subjective conceptualizations of SJ (Oplatka and Arar, 2016), one should see the framework I proposed here as influenced mostly by the culture of Anglo-American nations. In no way do I claim for it to be universal, but encourage researchers from different societies to explore SJ leadership identity formation in their educational systems to fill the gap in our knowledge about leadership for SJ in developing countries.

References

- Arar, K., 2015. Leadership for equity and SJ in Arab and Jewish schools in Israel: leadership trajectories and pedagogical praxis. *Int. J. Multicult. Educ.* 17 (1), 162–187.
- Cambron-McCabe, N., McCarthy, M., 2005. Educating school leaders for SJ. *Educ. Pol.* 19 (1), 201–222.
- Carden, L., Callahan, J., 2007. Creating leaders or loyalists: conflicting identities in a leadership development program. *Hum. Resour. Dev. Int.* 10 (2), 169–186.
- Conger, J., 1992. *Learning to Lead*. Jossey-Bass, San Francisco.
- Cruz-Gonzalez, C., Rodriguez, C.L., Segovia, J.D., 2020. A systematic review of principals' leadership identity from 1993 to 2019. *Educ. Manag. Adm. Leader* 1–23.
- DeRue, D.S., Ashford, S.J., 2010. Who will lead and who will follow? A social process of leadership identity construction in organizations. *Acad. Manag. Rev.* 35 (4), 627–647.
- Flores, M.P., De La Rue, L., Neville, H.A., Santiago, S., Ben Rakemayahu, K., Garite, K., Brawn, B., Valgoi, M., Brooks, J., Lee, E.S., Ginsburg, R., 2014. Developing SJ competencies: a consultation training approach. *Counsel. Psychol.* 42 (7), 998–1020.
- Flum, H., Kaplan, A., 2012. Identity formation in educational settings: a contextualized view of theory and research in practice. *Contemp. Educ. Psychol.* 37, 240–245.
- Furman, G., 2012. Social justice leadership as praxis: developing capacities through preparation programs. *Educ. Adm. Q.* 48 (2), 191–229.
- Gabarro, J.J., 2007. When a new manager takes charge. *Harv. Bus. Rev.* 1–15.
- Guerra, P.L., Nelson, S.W., Jacobs, J., Yamamura, E., 2013. Developing educational leaders for SJ: programmatic elements that work or need improvement. *Educ. Res. Perspect.* 40, 124–149.
- Hall, D.T., 2002. *Careers in and Out of Organizations*. Sage Publications, Thousand Oaks, CA.
- Higgins, M.C., Kram, K.E., 2001. Reconceptualizing mentoring at work: a developmental network perspective. *Acad. Manag. Rev.* 26 (2), 264–288.
- Hill, L.A., 2003. *Becoming a Manager: How New Managers Master the Challenges of Leadership*. Harvard Business School Press, Boston, MA.
- Kopelman, S., 2014. *Negotiating Genuinely: Being Yourself in Business*. Stanford University Press, Stanford, California.
- Lipman, P., 2004. *High Stakes Education: Inequality, Globalization, and Urban School Reform*. Routledge Falmer, New York.
- Marchiondo, L.A., Myers, C.G., Kopelman, S., 2015. The relational nature of leadership identity construction: how and when it influences perceived leadership and decision-making. *Leader. Q.* 26 (5), 892–908.
- McKenzie, K.B., Skrla, L., Scheuich, J.J., 2006. Preparing instructional leaders for SJ. *J. Sch. Leader.* 16, 158–170.
- Nicholson, J., Kuhl, K., Maniates, H., Lin, B., Bonetti, S., 2020. A review of the literature on leadership in early childhood: examining epistemological foundations and considerations of social justice. *Early Child. Dev. Care* 190 (2), 91–122.
- Oplatka, I., Arar, K., 2016. Leadership for social justice and the characteristics of traditional society: ponderings on the application of Western-grounded models. *Int. J. Leader. Educ.* 19 (3), 352–369.
- Oplatka, I., 2012. Towards a conceptualization of the early career stage of principalship: current research, idiosyncrasies and future directions. *Int. J. Leader. Educ.* 15 (2), 129–151.
- Oplatka, I., 2013. The place of “social justice” in the field of educational administration: an historical overview of emergent area of study. In: Bogoch, I., Shields, C. (Eds.), *International Handbook of Social [In]Justice and Educational Leadership*. Springer Publishing, Dordrecht, pp. 15–35.
- Postmes, T., Haslam, S.A., Swaab, R.I., 2005. Social influence in small groups: an interactive model of social identity formation. *Eur. Rev. Soc. Psychol.* 16 (1), 1–42.
- Pounder, D., Reitzug, U., Young, M., 2002. Preparing school leaders for school improvement, SJ, and community. In: Murphy, J. (Ed.), *The Educational Leadership Challenge: Redefining Leadership for the 21st Century*. University of Chicago Press, Chicago, pp. 261–288.
- Regan, K., 2016. Leadership identity formation in nonprofit human service organizations. *Hum. Serv. Organ.* 40 (5), 435–440.
- Saarukka, Z., 2014. *How School Principals Form Their Leadership Identity*, vol. 12. Vodenze, pp. 5–18.
- Samier, E., 2015. Emirati women's higher educational leadership formation under globalization: culture, religion, politics, and the dialectics of modernization. *Gend. Educ.* 27 (3), 239–254.
- Shapiro, E.C., Haseltine, F.P., Rowe, M.P., 1978. Moving up: role models, mentors, and the patron system. *Sloan Manag. Rev.* 19, 51–58.
- Singh, V., Vinnicombe, S., James, K., 2006. Constructing a professional identity: how young female managers use role models. *Women Manag. Rev.* 21 (1), 67–81.
- Somnath, S., Hanuscin, D.L., 2017. Development of teacher leadership identity: a multiple case study. *Teach. Teach. Educ.* 63, 356–371.
- Sveningsson, S.F., Alvesson, M., 2003. Managing managerial identities: organizational fragmentation, discourse and identity struggle. *Hum. Relat.* 56, 1163–1193.
- Theoharis, G., 2007. Navigating rough waters: a synthesis of the countervailing pressures against leading for SJ. *J. Sch. Leader.* 17, 4–27.

An overview of literature on factors that influence principal turnover

Rui Yan, Office of Applied Research, Evaluation, and Data Analytics, The City University of New York, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	445
Theoretical framework	446
Microeconomic Labor Market Theory	446
Job Choice Theory	446
Factors influencing principal turnover	447
Principal characteristics	447
School context	448
Working conditions	449
Economic benefits	449
Nonpecuniary working conditions	450
Summary	451
References	451

Introduction

Principal leadership is the second most influential school-based factor that influence student performance after classroom instruction, accounting for one-quarter of all school effects on student achievement (Leithwood et al., 2004; Marzano et al., 2005; Robinson et al., 2008). Given the strong influence of principal leadership on school success, leadership stability is also a critical concern of well-run schools.

Principal turnover typically includes principals transferring to another school, becoming a superintendent or other district staff, becoming a teacher, retiring, or pursuing another career inside or outside of the education system (Farley-Ripple et al., 2012b; Parlow, 2007). According to the National Center for Education Statistics (NCES), in the United States, 18% of public school principals in 2015–16 didn't stay in the same position at the same school in 2016–17, including 6% moved to a position as principal at another school and 10% stopped working as a principal. Principal turnover rates are relatively higher in schools with high concentrations of low-income, low-performing students, and/or students of color (Béteille et al., 2012; DeAngelis and White, 2011).

Principal turnover could have negative or positive effects on school success, depending on turnover frequency, the quality of principals that leave or move, and whether the school has a collaborative school culture and distributed leadership structure (Hargreaves and Goodson, 2006; Mascall and Leithwood, 2010). Generally it takes at least 5 years for principals to change their identity from peripheral to insider so as to exert effective influence on school success (Hargreaves and Goodson, 2006; Mascall and Leithwood, 2010). That is, for principals to be known, trusted, and accepted by the staff, they need to become an "insider" who has an influential power in the school community of practice, rather than staying peripheral (Fink and Brayman, 2006; Hargreaves and Goodson, 2006). These insider principals are able to build trusting relationships with people that could support a reform and deal with some complex cultural issues, such as teachers' resistance to change; they are thus able to acculturate themselves into new environments (Mascall and Leithwood, 2010). Insider principals also exert strong influence on a school's community of practice and are able to sustain structural changes as well as attitudinal and social changes (Fink and Brayman, 2006).

From this perspective, frequent rotation or turnover of principals is harmful if it results in "revolving door syndrome," in which teachers perceive that their principals come and go frequently as if through revolving doors (Fink and Brayman, 2006; Mascall and Leithwood, 2010). Teachers may stop investing in change efforts and making progress in learning and instruction and "wait principals out" (Hargreaves and Goodson, 2006), and may become more cynical and resistant to change and less-committed, thus resulting in a lack of sustainability and development for teacher and student learning, and the interruption of trusting working relationships between principals and teachers (Macmillan, 2000). Empirical research suggested that frequent principal turnover is often associated with lower teacher morale, higher teacher turnover, and a decline in student achievement gains (Béteille et al., 2012; Copland, 2003; Farley-Ripple et al., 2012a; Fink and Brayman, 2006; Leithwood et al., 2004; Miller, 2013).

Moreover, turbulent change in school leadership can be highly disruptive to long-term school development and school culture (Béteille et al., 2012; Fink and Brayman, 2006; Gates et al., 2006; MacMillan, 2000). Frequent principal turnover may interrupt the implementation or even the termination of long-term instructional programs, thus causing loss of institutional memory, inconsistency of instructional goals and reform agenda, as well as a loss of educational resources (Béteille et al., 2012; Ni et al., 2015).

From an economic perspective, a great deal of human and financial resources are lost if principals leave the education system within a short time or are frequently transferred or rotated without necessary causes (Mascall and Leithwood, 2010). According to a national report conducted by the School Leaders Network in 2014 (School Leaders Network, 2014), conservative estimates of the cost to develop, hire, and transition each principal is \$75,000. Increasing principal retention rates in high-poverty schools to match those in affluent schools could save school districts in the United States \$163 million annually (Sable et al., 2010).

Principal turnover could have a positive impact on school management under certain conditions. For instance, principal turnover could be beneficial to school management and student performance depending on the effectiveness of a principal that leaves and the successor who comes in (Hallinger and Heck, 1998; Leithwood et al., 2004). Therefore, losing a high-quality and effective principal could be detrimental to a school, but it could be beneficial for a school when an ineffective principal leaves and a more effective principal takes over (DeAngelis and White, 2011; Gates et al., 2006). Furthermore, a certain amount of principal turnover could be beneficial to schools if it results in better matches between principals and schools, infuses innovative and productive policies into schools, and dismisses ineffective employees from the school (Abelson and Baysinger, 1984; Baker et al., 2010; Fullan, 1993).

Theoretical framework

Microeconomic Labor Market Theory

Principal labor markets have the common characteristics of the broader labor markets. Based on Microeconomic Labor Market Theory, principal labor markets can be comprised of three dimensions: the supply side, demand side, and policy environments. The supply side of the principal labor market includes principals and principal candidates. They have their own preferences and benefits regarding their entry into, mobility in, and exit from the principalship. The demand side of the principal labor market includes schools and district central offices. Schools need principals to lead, manage, and maintain the functioning of the school system. District central offices define the roles and responsibilities of principals in school management, create principal working conditions, as well as determine the hiring, rotation, and dismissal of principals (Ornstein et al., 2016).

School districts, also serving as state policy extensions, assume major responsibilities for setting school policies and ensuring that local schools are properly managed and meet requirements and regulations from the federal and state governments to enhance student achievement (Ornstein et al., 2016). In addition to school districts, the federal and state governments as policy makers can influence the principal labor markets through establishing and enacting statutes to raise student academic performance, overseeing policy implementation, administering grant funds, determining school taxes and financial aid to local school districts, engaging in educational innovation and research, and setting minimum standards for training and recruiting personnel, as well as providing curriculum guidelines and assessment requirements (Ornstein et al., 2016).

In summary, policy makers at the federal and state levels as well as local school districts create policy environments for principals to work in, which influence both the supply and demand sides and determine the equilibrium of the principal labor markets. During the past three decades, through the widespread introduction of accountability policies, state governments, local school boards, superintendents, principals, and teachers all have been assuming growing responsibilities and pressures to enhance student academic performance. These accountability policies exert a great deal of influence on principal labor markets. For instance, on the demand side of the principal labor markets, in order to meet student academic standards mandated by the state governments, school districts aim at searching for higher-quality principals. They often raise hiring standards and qualifications such as requiring a license/certification in administration, advanced education levels, and more working experience from principals, thus influencing the supply side of the labor markets, as well (Loeb et al., 2010; Whitaker and Vogel, 2005). Meanwhile, given pressures from state accountability policies, on the supply side of the principal labor markets, principals and principal candidates may want to work at schools with a higher proportion of high-performing students and supportive parents that could result in less pressures to achieve state academic standards (Farley-Ripple et al., 2012a).

Job Choice Theory

In addition to labor market theory, Job Choice Theory can be used to interpret principals' career decisions. Job Choice Theory, first introduced to the educational field by Young et al. (1989), can provide objective and subjective perspectives to help understand the career choices made by principals and principal candidates. Both personal and external factors are often weighed based on their relative importance to each individual as they make their career decisions (Behling et al., 1968).

From the objective perspective of Job Choice Theory, principals are "economic beings" who seek to maximize their economic status by joining the organization that is most economically competitive for them. Economic benefits such as salary/benefit packages and prospects for job advancement are all important factors in principals' career choices (Pounder and Merrill, 2001). Empirically, researchers have found that principals' salaries and other job benefits such as stipends and retirement benefits, are important factors in both drawing teachers into leadership and in retaining principals (Akiba and Reichardt, 2004; Fraser and Brock, 2006; Pounder and Merrill, 2001). For example, Fraser and Brock (2006) conducted a qualitative study on principal retention among Catholic school principals in New South Wales and Australia and found that financial security is a critical factor for principal retention.

From the subjective perspective of Job Choice Theory, principals are "psychological beings" who choose jobs to meet their psychological needs and emotional expectations (Young et al., 1989). Principals seek principalships in specific districts, schools, and community environments that are more satisfying and compatible with their own psychological needs (Behling et al., 1968; Pounder and Merrill, 2001). For instance, if a principal prefers to work in a democratic-leadership rather than a bureaucratic-leadership environment, then he is more likely to work in a district or school with this kind of organizational climate (Pounder and Young, 1996). As another example, principals often need adequate autonomy and time to work with their school communities

to achieve meaningful school improvement goals and to cultivate a positive school culture (Fink and Brayman, 2006). A lack of authority on school decision making could undermine principal job satisfaction, which could cause a principal to transfer to another school/district or even leave the education system (Vroom, 1964).

It is worth noting that, as an independent human being, a principal's individual preferences rather than school district assignments often play a more important role in his principal turnover decisions (Loeb et al., 2010). Although, in cases such as emergencies or reassignments within region or district offices, a principal could be directly appointed to a position without a formal selection process. More often, interested principals apply for and are then chosen for the open positions (Loeb et al., 2010).

Factors influencing principal turnover

With the guidance of the Microeconomic Labor Market Theory and Job Choice Theory, the following paragraphs review literature on factors that influence principal turnover, which can be categorized into principal characteristics, school context, and working conditions.

Principal characteristics

As the supply side of the principal labor market, principals' individual characteristics can influence principal turnover, since their characteristics are interactively associated with their objective/subjective perspectives regarding values and benefits of their job choice and transition (Baker et al., 2010; Fuller et al., 2007). Principal demographic and professional characteristics that are associated with principal turnover include principals' age (Fuller et al., 2007; Fuller and Young, 2009), gender (Akiba and Reichardt, 2004), race (Akiba and Reichardt, 2004; Baker et al., 2010; Chapman and International Institute for Educational Planning, 2005; Fuller et al., 2007; Gates et al., 2006), as well as professional experience, including years of administrative and teaching experience (Gates et al., 2006; Papa, 2007), and educational level (Gates et al., 2006).

Research findings regarding principal demographics have shown few consistent results. In terms of principals' age, Fuller and Young (2009) examined the tenure and retention rates of newly hired principals in Texas public schools from 1996 through 2008. Their findings suggested that newly hired principals between 40 and 44 years old have the longest tenure, and principals who are under 30 years old have the shortest tenure. Additionally, principals between the ages of 35 and 49 tend to have greater school retention than principals of other ages, especially for middle and high school principals. Examining interactions with race and gender, Akiba and Reichardt (2004) performed logistic regressions and concluded that, in Colorado, female administrators and administrators of color (both principals and assistant principals), younger than 40 or older than 56, had statistically significant higher attrition rates than male and nonminority groups.

Further categorizing different facets of principal turnover, Gates et al. (2006) performed multinomial logit regression on administrative data from Illinois and North Carolina from the years 1987–2001, and found that women are more likely to leave the education system and change positions than male principals. In contrast to the above findings, Baker et al. (2010) found that minority or male principals were more likely to be unstable in their positions. The inconsistent findings across the studies could result from varied principal individual preferences and from different methodology and datasets utilized in these studies. For example, Baker et al. (2010) categorized different types of principal turnover (in Missouri) and created different indicators of the time period at which a cohort member: (a) left the principalship altogether, (b) made a first move to another school, or (c) made a second move to another school. They then constructed a "stability" ratio to identify the amount of time a principal spent in any given school as a percentage of the total time that principal was in the dataset, and utilized truncated regression models and Cox Proportional Hazard models.

The professional characteristics that influence principal turnover include principals' education level, working experience, and whether having attended preservice professional training programs. In terms of education level, Gates et al. (2006) found that principals who had a Master's degree or higher were less likely to change positions within the state system than those without an advanced degree, but educational level had no significant effect on the probability of leaving the school system or transferring schools. They also applied Barron's ranking of the undergraduate college that a principal graduated from as a proxy for unobserved principal quality, but revealed no statistically significant relationship between principal quality and principal turnover.

For principals' years of experience, Gates et al. (2006) found negative coefficients on both linear and quadratic terms in regression for principals' years of experience, which indicated that very experienced and very inexperienced principals are least likely to change schools or positions. Similarly, Baker et al. (2010) found that more experienced principals were more likely to stay in one place and have higher stability ratios. However, they also indicated the possibility that ineffective but experienced principals were staying in their positions, which could reduce the reliability of using principals' years of experience as a signal of principal quality.

Additionally, principals' preservice professional training programs could influence principal turnover. These programs are mostly facilitated by universities that provide course work to aspiring principals in curriculum leadership, assessment, as well as school improvement and management (Grissom and Harrington, 2010). McKibben (2013) examined the influence of principal preservice training programs on principal turnover, but found no significant influence of district-level preservice training programs on the probability of principals moving to another school or leaving the principal profession, based on multinomial logistic regression on SASS and PFS in 2008–2009. She speculated that gaining access to preservice training programs did not necessarily imply the value of these programs, and that the quality of training should also be considered.

School context

From the demand side of the principal labor markets, school contextual factors reflect the circumstances that principals are working under, thus influencing principals' mobility and attrition decisions. School contextual characteristics include the proportion of students of color and low-income students (Baker et al., 2010; Gates et al., 2006; Loeb et al., 2010), prior student performance (Akiba and Reichardt, 2004; DeAngelis and White, 2011; Loeb et al., 2010; Partlow, 2007), school size/enrollment (Akiba and Reichardt, 2004; Baker et al., 2010; Gates et al., 2003), school urbanicity (Gates et al., 2006; Papa et al., 2002), school level (Baker et al., 2010; Gates et al., 2006), school type (Ni et al., 2015), number of students per full-time equivalent teachers (DeAngelis and White, 2011), and percentage of students or teachers in the school who are racial/ethnic minority (Gates et al., 2006; Papa, 2007).

Compared with the findings of the influence of principal characteristics on principal turnover, the literature on the effects of school context on principal turnover has shown more consistent conclusions. Generally, schools with higher proportions of low-income, lower-performing students, students of color, and less-qualified teachers are associated with higher possibility of principal turnover (Akiba and Reichardt, 2004; Baker et al., 2010; Gates et al., 2006; Loeb et al., 2010; Papa, 2007; Partlow, 2007). The following paragraphs elaborate on these findings.

First, principals in schools with large concentrations of students of color often have higher principal turnover rates (Baker et al., 2010; Gates et al., 2006; Loeb et al., 2010). With administrative data from Illinois and North Carolina from 1987 to 2001, Gates et al. (2006) found that principals in schools with larger proportions of students of color are more likely to transfer or leave the principalship, but principals who are the same race as a plurality of students have higher career stability. Baker et al. (2010) found that a higher proportion of African American students in school was associated with higher principal turnover with truncated regression models and Cox Proportional Hazard models. By dividing schools into different quartile groups based on proportion of students of color, Loeb et al. (2010) found that the odds of principals leaving were 60% higher in schools in the top quartile of minority students than the bottom quartile, which corresponded to a 5%-point difference in the probability.

Additionally, a larger proportion of low-income students are often associated with higher probability of principal turnover. Loeb et al. (2010) found that the odds of principals leaving the principalship were about 30% higher in schools in the middle quartile of free or reduced-price lunch than those in the bottom quartile, which is possibly caused by many principals' preference to work in schools with a lower proportion of minority, low-income, and lower-performing students (Loeb et al., 2010).

Moreover, higher student performance in a school is often associated with lower probability of principal turnover. Due to the formal and informal state and district sanctions and accountability pressures placed on schools, many principals prefer to work in schools with a larger proportion of high-performing students (Loeb et al., 2010). Higher-performing students are more likely to achieve higher test scores that could meet the state academic standards, and more involvement from affluent parents and communities could bring in more support and resources, thus attracting many principals to work in these types of schools (DeAngelis and White, 2011). This hypothesis is not universally agreed upon: Akiba and Reichardt (2004) found that lower school achievement only predicted higher attrition among female leaders. Further, Partlow (2007) found student math test scores to be the only factor that significantly influences principal turnover among eight school contextual factors, including superintendent turnover rate, building enrollment, student attendance, student mobility, pupil-teacher ratio, teacher attendance. DeAngelis and White (2011) found that schools with lower average test scores or those that fail to make Adequate Yearly Progress (AYP) often have higher odds of principal transitions. Moreover, principals who moved across districts often moved to schools with lower percentages of low-income and lower-performing students, but those who moved within districts moved to schools with similar student demographic characteristics (DeAngelis and White, 2011). They speculated that principal mobility within districts in Illinois may be determined more by district needs and decisions than by principals' preferences.

School enrollment has shown mixed results on the influence of the probability of principal turnover. Specifically, principals who transferred to another school in the same city often moved to schools with smaller enrollments, which indicates principals' preferences for smaller schools (Papa et al., 2002). Consistent with these findings, Akiba and Reichardt (2004) found that larger schools were more likely to have higher rates of principal turnover than middle-size and small schools. They speculated that smaller schools provide leaders with more opportunities to build close relationships with faculty and students. On the contrary, Gates et al. (2006) found that principals in larger schools were less likely to transfer to another school or leave the principalship to pursue another position in the school system. Although larger schools tend to have more problems, the higher principal salary in these schools could compensate for the more complex tasks and higher pressures that they need to tackle to some extent (Gates et al., 2003).

For urbanicity, Gates et al. (2006) found that the probability of changing schools was higher for principals in urban areas of Chicago and North Carolina relative to principals in rural areas, although the magnitude of these differences were small. In terms of school type, Ni et al. (2015) found that principals in charter schools were much more likely to leave the education system (including retirement) compared with traditional public schools with a longitudinal dataset in Utah from 2004 to 2011.

School level is another school contextual factor that influences principal turnover. Gates et al. (2006) concluded that elementary school principals had the least principal turnover, while high school principals were nearly twice as likely to change positions compared with elementary school principals, and middle school principals were about 25% more likely to change positions than elementary school principals. Again, other studies have different results; for example, Baker et al. (2010) found that middle school principals had higher turnover than those in high schools. They speculated that many principals regard middle school principalship as a stepping stone to obtain high school principalship, partly because high school principals receive a salary premium compared with principals in elementary and middle schools.

Finally, in addition to student demographics, schools with lower percentages of highly qualified teachers, as defined by No Child Left Behind (NCLB),¹ have significantly higher odds of principals moving to schools outside of districts and leaving the education system altogether (DeAngelis and White, 2011; Papa, 2007). Since NCLB mandates a basic level of qualifications, including higher expectations and requirements for teachers of core subjects, principals are under more pressures to recruit and hire highly qualified teachers, and to help teachers meet professional standards and enhance student performance (DeAngelis and White, 2011).

Working conditions

According to the International Labor Organization, working conditions cover a broad range of topics and issues, from working time (hours of work, rest periods, and work schedules) and remuneration, to physical conditions and mental demands that exist in the workplace (International Labor Organization, 2015). Compared with principal characteristics and school contextual factors that influence principal turnover, there has been far less research on the influence of principal working conditions on turnover.

Principal working conditions is often conflated with school context in the field of principal turnover research. The major reason is that many principals prefer to work in schools with fewer low-income, low-performing students, and students of color; and these types of schools often have better working conditions, including more resources, fewer safety and disciplinary problems, more parental involvement, fewer teacher vacancies, and less teacher and student turnover (Loeb et al., 2010). For instance, Akiba and Reichardt (2004) constructed principal working conditions as school contexts, including poverty level, proportion of minority students, school size, school location, and student achievement level, as well as instructional and administrative expenditure per student. These studies are significant explorations of the impact of working conditions on principal turnover, however, the equating of principal working conditions and school context can cause some difficulties for researchers to distinguish between the effects of school context and working conditions on principal turnover. Loeb et al. (2010) was one of the first studies that responded to this issue in research. They separated the influences of student demographics and school climate on principal turnover. Their findings suggested that high principal turnover in schools with more low-income, minority, and low student achievement could be driven by undesirable working conditions rather than student demographics in the schools.

From the Job Choice Theory Perspective, working conditions reflect principals' economic and psychological needs as they are making career choices, thus affecting principal turnover behaviors. Principal working conditions include economic benefits and nonpecuniary working conditions. Economic benefits, including principal salary and job benefits such as stipends and retirement benefits, are important factors to draw teachers into leadership positions and for principals to remain in the principalship (Akiba and Reichardt, 2004; Fraser and Brock, 2006; Pounder and Merrill, 2001). Nonpecuniary working conditions include all the other factors that are not related to salary and financial benefits (e.g., workload, autonomy, and advancement opportunities), are also critical concerns for many principals upon entering into, transitioning, or leaving schools (Farley-Ripple et al., 2012a; Loeb et al., 2010; Pijanowski et al., 2009).

Economic benefits

Principals are "economic beings" who seek to maximize their economic status by joining the organization that is most economically competitive (Young et al., 1989). Therefore, salary and benefit packages are important factors in principals' career choices (Pounder and Merrill, 2001). On the one hand, principals' salaries and job benefits are competitive compared with those of teachers (Pounder and Merrill, 2001). For instance, the average base salary of a high school principal in a medium-sized district in Utah exceeded that of a teacher with the same degree and years of experience in the system by 24%–55% depending on administrative experience (Pounder and Merrill, 2001). Moreover, in Arkansas' largest high schools, the difference between a midcareer teacher with a master's degree and a principal was a 97.5% increase in annual salary, or \$46,640 more pay per year for the principal in 2006–2007. Although most principals are on 12-month contracts rather than 9-month contracts like most teachers, after adjusting contract day difference, principals still showed a 67% increase in daily pay over teachers with Master's degrees and 15 years' experience (Pijanowski and Brady, 2009). The gap between principal and teacher salaries could incentivize teachers to move to various leadership positions (Pounder and Merrill, 2001). The relative salaries, which reflect pecuniary advantages of school principalship over many other educator positions, are perceived by principals as the strongest objective factor of job desirability of the principalship, which explained 38% of total variance in rating of job attraction and intentions (Composite Job Desirability Index) (Pounder and Merrill, 2001).

On the other hand, many principals perceive the gap between teacher and principal salary to be too small to reflect the much greater responsibility that principals assume compared with teachers (Pounder and Merrill, 2001). Thus, unsatisfactory principal salaries could be one of the reasons why principals leave for other positions (Pounder and Merrill, 2001). In fact, the gap in the compensation between principals and teachers has narrowed to the point that it may engender a discouraging impact on principal candidates or current principals who are considering entering or remaining in the principalship (Pijanowski and Brandy, 2009). Furthermore, the compensation gap between principals in different locations also varies dramatically. For example, principals' salaries are much lower in smaller, poorer, and more rural districts (Pijanowski and Brady, 2009).

¹NCLB requires states to (1) measure the extent to which all students have highly qualified teachers, particularly minority and disadvantaged students, (2) adopt goals and plans to ensure all teachers are highly qualified and, (3) publicly report plans and progress in meeting teacher quality goals. <http://www2.ed.gov/nclb.html>.

Given the strong influence of salary on principals' job desirability and satisfaction, empirical evidence has indicated that higher principal salaries are often associated with lower principal turnover rates (Akiba and Reichardt, 2004; Baker et al., 2010; Papa et al., 2002; Papa, 2007; Pijanowski and Brandy, 2009; Yan, 2020). Akiba and Reichardt (2004) concluded that salary was significantly associated with principal turnover decisions, and principals were more likely to leave when they expected an increase in compensation if transferring to another school. Additionally, Papa (2007) found that schools paying one standard deviation below the mean salary were 9.5 times more likely to lose their principals compared to schools paying one standard deviation above the mean salary. Baker et al. (2010) compares principal's relative salary to their peers in the same labor market, and found that principals are able to leverage school-to-school moves for an average change in relative salary of about 5% (Baker et al., 2010). They regard principals' relative salary as the most "consistent and potential policy lever for principal retention" (p. 551). Further, Yan (2020) found that higher salary lowers the odds of principals moving to another school relative to staying, while holding principal characteristics and school context constant.

Nonpecuniary working conditions

With state and federal accountability systems placing increasing pressures on school-level leaders to improve student performance, principals assume expanding roles, challenges, and a growing workload from multiple stakeholders (Loeb et al., 2010; Pijanowski et al., 2009). Under these circumstances, oftentimes economic benefits alone are not adequate to compensate for the stress and adverse working conditions principals face (Pijanowski et al., 2009). Therefore, principals often seek the principalship in schools and districts with the community environments that are more satisfying and compatible with their own psychological needs (Behling et al., 1968; Pounder and Merrill, 2001). Thus, many nonpecuniary working conditions, such as workload, autonomy, district/teacher/parent supports, and professional development and career advancement opportunities, are also important concerns for many principals when considering the entry, mobility, or exit of the principalship (Farley-Ripple et al., 2012a; Loeb et al., 2010; Pijanowski et al., 2009).

Given the broad range of nonpecuniary working conditions and the lack of measurements of many possible causal factors such as principal autonomy and district support, very limited empirical research has examined the relationship between nonpecuniary working conditions and principal turnover (Fuller et al., 2015; Tekleselassie and Villarreal, 2011; Yan, 2020). For instance, Fraser and Brock (2006) applied narrative surveys and structured interviews on 47 random principals in elementary and secondary schools in New South Wales, and identified incentives and disincentives of working conditions that influence principal turnover. The incentives included district support, professional development opportunities, union, support from teachers and parents, available resources, as well as principal autonomy to make decisions. The disincentives of principal working conditions included a sense of isolation, stress from work, insufficient remuneration, staff issues, demanding and disgruntled parents, and unrealistic expectations from employing authorities.

These trends are also similar for principals in the United States. In a case study with more than 100 administrators' career transitions within Delaware education system, and semi-structured interview data from 48 principals and assistant principals, Farley-Ripple et al. (2012a) found that economic benefits, working relations with multiple stakeholders (teachers, parents, school districts, and school boards), as well as the availability of opportunities into administration or the district central office are all critical incentives for principal retention.

Fuller et al. (2015) summarized four indirect yet powerful working conditions that influence principal turnover in small and mid-sized urban districts. These four factors include state policy, leadership preparation programs, principal autonomy, and district policy. They found that intrinsic rewards, overall workload, and a feeling of effectiveness were important factors influencing principal retention. Moreover, principals in small and mid-sized urban districts perceived state accountability policies and statewide testing systems, as well as principal workload, to be more influential than those in other districts. More recently, Yan (2020) utilized multinomial logistic regression on national survey data from National Center for Educational Statistics (NCES), and examined how contract type, school disciplinary environments, workload, professional development, and principal influences in school are associated with different types of principal turnover.

The non-pecuniary working conditions that influence principal turnover include principal workload, school disciplinary environments, and principal influences on school decision making. In terms of workload, the balance of professional and personal lives reflects principals' concerns over their life quality and health (Pounder and Merrill, 2001). Therefore, a heavy workload and extra responsibilities are regarded by principals as unattractive job attributes that affect the job desirability. Analyzing the influence of principal workload on principal turnover intentions, Tekleselassie and Villarreal (2011) conducted a three-level Generalized Multi-level Model based on Schools and Staffing Survey in 2003–2004, and found that work overload was unrelated to principals' departure intentions, but increased the intentions of principals changing schools. They speculated that a heavier workload could cause principals to consider transferring to another school that has a lighter workload and shorter working hours. Moreover, Fuller et al. (2015)'s descriptive analysis indicates that workload is among the most important concerns for principal retention.

For school disciplinary environment, empirical studies indicate the positive relationship between disciplinary environment and principal retention (Loeb et al., 2010; Sun and Ni, 2015; Tekleselassie and Villarreal, 2011; Yan, 2020). For instance, Yan (2020) found that one standard deviation improvement in school disciplinary environment rating lowers the odds of principals moving to another school by 36.4% relative to staying. By further separating school disciplinary environment from student characteristics regarding the effects on principal turnover, this study indicates that as school disciplinary environment rating increases by one standard deviation, the odds of moving to another school for principals in schools serving the highest quartile of students of color are

about 1/3 lower than those in other schools. It indicates that improving school disciplinary environment plays a more important role in lowering principal turnover in schools serving high concentrations of students of color.

Another important factor that influence principal turnover is principal influence in school decision making. Principals need adequate autonomy to work with school communities to establish and achieve meaningful school improvement goals and cultivate positive school culture (Fink and Brayman, 2006). Providing principals with adequate autonomy allows them to develop shared vision and efforts within the school communities, thus facilitating school improvement goals and enhancing student learning (Fink and Brayman, 2006; Hargreaves and Goodson, 2006). However, if states and local school districts exercise excessive influence on setting directions, establishing policies, and hiring teachers, principals may feel themselves as “puppets” without influential power on school reforms and have little leeway to influence school decision making (Tekleselassie and Villarreal, 2011).

Among all the areas upon which principals exert their influence, the areas of recruiting effective teachers, transferring teachers, and discharging unsuitable teachers are where principals experience the greatest gap between the authority they need in order to make changes in school and the actual influence they possess that is delegated from the district central office (Adamowski et al., 2007). This “autonomy gap” is especially true for principals in urban and low-performing schools (Adamowski et al., 2007; Papa and Baxter, 2008). Lack of influence over major decisions on school matters could hurt principals’ emotions in school reforms and job satisfaction, which causes principals transfer to another school, or even leave the education system (Adamowski et al., 2007).

Therefore, the lack of influence over supervision (e.g., principals’ perceived influence over spending, teacher evaluations, hiring teachers, and disciplinary policies) are often significantly associated with principals’ departure and mobility intentions (Tekleselassie and Villarreal, 2011). Furthermore, Yan (2020) found that principals having more influence on determining teacher professional development and budget spending and less influence on setting performance standards are associated with a lower likelihood of principals leaving the education system.

Summary

Principal leadership plays a significant role in school improvement and student success, therefore, the retention and sustainability of principals has a substantial impact on student success and the education system. Frequent turnover of school principals often has negative impacts on school success, thus it is critical to examine how various factors are associated with principal turnover. This article applies Microeconomic Labor Market Theory and categorizes the influencing factors and their interactions of principal labor markets from the supply side, demand side, and policy environment perspectives. Additionally, utilizing Job Choice Theory, this article examines how principals’ career choices are affected by objective and subjective factors.

With the guidance of the two frameworks, this article has reviewed factors that are associated with principal turnover from three dimensions: principal characteristics, school context, and working conditions. Overall, most principals’ demographic and professional characteristics have not shown a consistent impact on principal turnover. These findings could be supported by Job Choice Theory, as each principal holds his/her own objective and subjective perspectives as well as individual preferences regarding the relative importance of multiple aspects of factors in their job decision making.

In terms of school contextual factors, many schools with more low-income, low-achieving students, students of color, and less-qualified teachers are disadvantaged in terms of the ability to retain quality principals. These schools often have principals who have less experience, less education, and have attended less selective colleges. Furthermore, these disadvantaged schools are often used as “stepping stones” for principals to more desirable schools, which may partly explain the higher principal turnover rates in these schools.

Compared with principal characteristics and school contextual factors, research on how on the relationships between principal working conditions and principal turnover is rather limited. With the expanding roles and responsibilities of principals to improve student success, economic and non-pecuniary working conditions are playing an increasingly significant role in principals’ job satisfaction and retention. Factors such as salary, contract type, school disciplinary environments, workload, and principal influences in school have a significant impact on principal turnover.

Given the limited research on principal turnover, more empirical research remains to be conducted to unravel the patterns and dynamics of principal career transitions and factors that influence different types of principal turnover. For instance, longitudinal studies could be examined to study the dynamic transition of principals over time. Additionally, qualitative research methodology such as case study or focus groups could be applied in order to obtain an in-depth understanding regarding voluntary vs. involuntary reasons behind different principal transitions and various organizational factors that influence principal retention. Moreover, researchers could conduct more inter-disciplinary research on principal career transitions that apply theories of labor economics, psychology, and leadership in order to reveal the economic and psychological reasons that drive principals’ career transitions, and mitigate the negative influence of principal turnover on school system.

References

- Abelson, M.A., Baysinger, B.D., 1984. Optimal and dysfunctional turnover: toward an organizational level model. *Acad. Manag. Rev.* 9 (2), 331.

- Adamowski, S., Theriault, S.B., Cavanna, A.P., 2007. The Autonomy Gap: Barriers to Effective School Leadership. Thomas B. Fordham Institute and American Institutes for Research, Washington, D.C.
- Akiba, M., Reichardt, R., 2004. What predicts the mobility of elementary school leaders? An analysis of longitudinal data in Colorado. *Educ. Pol. Anal. Arch.* 12 (18).
- Baker, B.D., Punswick, E., Belt, C., 2010. School leadership stability, principal moves, and departures: evidence from Missouri. *Educ. Adm. Q.* 46 (4), 523–557.
- Behling, O., Labovitz, G., Gainer, M., 1968. College recruiting: a theoretical base. *Pers. J.* 47, 13–19.
- Bételle, T., Kalogrides, D., Loeb, S., 2012. Stepping stones: principal career paths and school outcomes. *Soc. Sci. Res.* 41 (4), 904–919.
- Chapman, J.D., International Institute for Educational Planning, 2005. Recruitment, Retention, and Development of School Principals. International Institute for Educational Planning, Paris, France.
- Copland, M.A., 2003. Leadership of inquiry: building and sustaining capacity for school improvement. *Educ. Eval. Pol. Anal.* 25 (4), 375–395.
- DeAngelis, K.J., White, B.R., 2011. Principal Turnover in Illinois Public Schools, 2001–2008. Illinois Education Research Council, Edwardsville, IL.
- Farley-Ripple, E.N., Raffel, J.A., Welch, J., 2012a. Administrator career paths and decision processes. *J. Educ. Adm.* 50 (6), 788–816.
- Farley-Ripple, E.N., Solano, P.L., McDuffie, M.J., 2012b. Conceptual and methodological issues in research on school administrator career behavior. *Educ. Res.* 41 (6), 220–229.
- Fink, D., Brayman, C., 2006. School leadership succession and the challenges of change. *Educ. Adm. Q.* 42 (1), 62–89.
- Fraser, J., Brock, B.L., 2006. Catholic school principal job satisfaction: keys to retention and recruitment. *J. Cathol. Educ.* 9 (4), 13.
- Fullan, M., 1993. *Change Forces: Probing the Depth of Educational Reform*. Falmer Press, New York, NY.
- Fuller, E.J., Young, M.D., 2009. Tenure and Retention of Newly Hired Principals in Texas. University Council for Educational Administration, Department of Educational Administration, & University of Texas at Austin, Austin, TX.
- Fuller, E., Young, M.D., Orr, M.T., 2007. Career Pathways of Principals in Texas. Paper session presented at the meeting of the American Educational Research Association, Chicago, IL.
- Fuller, E.J., Hollingworth, L., Young, M.D., 2015. Working conditions and retention of principals in small and mid-sized urban districts. In: *Leading Small and Mid-Sized Urban School Districts. Advances in Educational Administration*, vol. 22. Emerald Group Publishing, Bingley, United Kingdom.
- Gates, S.M., Ringel, J.S., Santibanez, L., 2003. Who is Leading Our Schools? An Overview of School Administrators and Their Careers. No. 1679. Rand Corporation, Santa Monica, CA.
- Gates, S.M., Ringel, J.S., Santibañez, L., Guarino, C., Ghosh-Dastidar, B., Brown, A., 2006. Mobility and turnover among school principals. *Econ. Educ. Rev.* 25 (3), 289–302.
- Grissom, J.A., Harrington, J.R., 2010. Investing in administrator efficacy: an examination of professional development as a tool for enhancing principal effectiveness. *Am. J. Educ.* 116 (4), 583–612.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9 (2), 157–191.
- Hargreaves, A., Goodson, I., 2006. Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educ. Adm. Q.* 42 (1), 3–41.
- International Labor Organization, 2015. Definition of Working Conditions. Retrieved from: <http://www.ilo.org/global/topics/working-conditions/lang-en/index.htm>.
- Leithwood, K., Seashore Louis, K., Anderson, S., Wahlstrom, K., 2004. Review of Research: How Leadership Influences Student Learning. The Wallace Foundation, New York, NY.
- Loeb, S., Kalogrides, D., Horg, E.L., 2010. Principal preferences and the uneven distribution of principals across schools. *Educ. Eval. Pol. Anal.* 32 (2), 205–229.
- Macmillan, R., 2000. Leadership succession, cultures of teaching and educational change. In: *The Sharp Edge of Educational Change: Teaching, Leading and the Realities of Reform*, pp. 52–71.
- Marzano, R.J., Waters, T., McNulty, B.A., 2005. *School Leadership That Works: From Research to Results*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Mascall, B., Leithwood, K., 2010. Investing in leadership: the district's role in managing principal turnover. *Leader. Pol. Sch.* 9 (4), 367–383.
- McKibben, S., 2013. Do local-level principal preparation programs prevent principal turnover? *Pub. Purpose* 11, 69–87.
- Miller, A., 2013. Principal turnover and student achievement. *Econ. Educ. Rev.* 36, 60–72.
- Ni, Y., Sun, M., Rorrer, A., 2015. Principal turnover: upheaval and uncertainty in charter schools? *Educ. Adm. Q.* 51 (3), 409–437.
- Ornstein, A., Levine, D.U., Gutek, G., Vocke, D., 2016. *Foundations of Education*. Nelson Education, Toronto, Ontario.
- Papa, F.C., Lankford, H., Wyckoff, J., 2002. The Attributes and Career Paths of Principals: Implications for Improving Policy. Teacher Policy Research Center, Stanford, C.A.
- Papa, F., Baxter, I., 2008. Hiring teachers in New York's public schools: can the principal make a difference? *Leadersh. Policy Sch.* 7 (1), 87–117.
- Papa Jr., F., 2007. Why do principals change schools? A multivariate analysis of principal retention. *Leader. Pol. Sch.* 6 (3), 267–290.
- Partlow, M., 2007. Contextual factors related to elementary principal turnover. *Plan. Changing* 38 (1–32), 60–76.
- Pijanowski, J.C., Brady, K.P., 2009. The influence of salary in attracting and retaining school leaders. *Educ. Urban Soc.* 42 (1), 25–41.
- Pijanowski, J.C., Hewitt, P.M., Brady, K.P., 2009. Superintendents' perceptions of the principal shortage. *Natl. Assoc. Sec. Sch. Princ. Bull.* 93 (2), 85–95.
- Pounder, D.G., Merrill, R.J., 2001. Job desirability of the high school principalship: a Job Choice Theory perspective. *Educ. Adm. Q.* 37 (1), 27–57.
- Pounder, D.G., Young, I.P., 1996. Recruitment and selection of educational administrators: Priorities for today's schools. In: *International handbook of educational leadership and administration*. Springer Netherlands, New York, NY, pp. 279–308.
- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Sable, J., Plotts, C., Mitchell, L., 2010. Characteristics of the 100 Largest Public Elementary and Secondary School Districts in the United States: 2008–09. Publication No. NCES 2011-301. National Center for Education Statistics, Washington, DC.
- School Leaders Network, 2014. CHURN: The High Cost of Principal Turnover. Retrieved from: http://connectleadsucceed.org/sites/default/files/principal_turnover_cost.pdf.
- Sun, M., Ni, Y., 2015. Work environments and labor markets: explaining principal turnover gap between charter schools and traditional public schools. *Educ. Adm. Q.* 52 (1), 144–183.
- Tekleselassie, A.A., Villarreal III, P., 2011. Career mobility and departure intentions among school principals in the United States: incentives and disincentives. *Leader. Pol. Sch.* 10 (3), 251–293.
- Vroom, V.H., 1964. *Work and Motivation*. Wiley, New York, N.Y.
- Whitaker, K., Vogel, L., 2005. Joining the ranks: opportunities and obstacles in obtaining principal positions. *J. Person. Eval. Educ.* 18 (1), 3–19.
- Yan, R., 2020. The influence of working conditions on principal turnover in K-12 public schools. *Educ. Adm. Q.* 56 (1), 89–122.
- Young, I.P., Rinehart, J.A., Place, A.W., 1989. Theories for teacher selection: objective, subjective, and critical contact. *Teach. Teach. Educ.* 5 (4), 329–336.

Review of evaluation approaches for school principals

Artemio Arturo Cortez Ochoa^a and Sally M. Thomas^b, ^a University of Cambridge, United Kingdom; and ^b University of Bristol, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	453
Methodology	454
Literature review	454
Standards for school principals	455
Methods for evaluating school principals	456
School inspections	456
Self-review	456
Written assessments	456
Surveys with the school community	457
Student learning indicators	457
School principals' professional development	458
Conclusions	458
Limitations	459
Appendix 1	459
References	465

Introduction

For the past three decades, school leadership researchers have been dedicated to studying school principals' role in developing schools into effective, efficient, and equitable learning spaces (OECD-CEPPE, 2013). Much academic attention has focused on approaches to school principals' leadership for learning, including, for example, distributed, collaborative, democratic, facilitative, participative, transformational, and ethical leadership, as well as functional, structural and cultural initiatives (Leithwood, 2001; Leithwood and Jantzi, 2005; Cheng et al., 2016; Gordon and Patterson, 2006; Catano and Stronge, 2007). Also, researchers have sought to address the effects of leadership approaches on students' outcomes (Barber and Mourshed, 2007; Leithwood et al., 2020). Yet, there seems to be a shortage of academic literature on existing approaches to assessing school principals for accountability and developmental purposes (Catano and Stronge, 2007; OECD-CEPPE, 2013). This chapter focuses on the education policy perspective and provides an overview of existing school principals' evaluation systems in several contexts worldwide.

This review is necessary because, as demonstrated in an exploration of the ERIC database in January 2021, school principals' evaluation is a less common area of research. While the term "teachers' evaluation" yielded more than 36 thousand publications since 2002, a few more than three thousand papers on school principals' evaluation were found in the database during the same period. From an education policy perspective, primary and secondary teachers have mainly been evaluated to decide on incentives, continuing professional development, and sanctions (OECD, 2005; Isoré, 2009; Villegas-Reimers, 2003; Guerriero, 2013; Martinez Rizo et al., 2016; Akiba and Liang, 2016; Polikoff and Porter, 2014; Kane et al., 2014; Kane and Douglas, 2012) and individual teacher evaluation and appraisal is seen as a key lever of school improvement (Thomas, 2020; Cortez Ochoa et al., 2018). In contrast, school principals are typically addressed differently; for instance, they are expected to achieve certifications for leading and managing via Diploma courses (Bush and Glover, 2016). In addition, they may undergo school inspections and assessments (OECD-CEPPE, 2013; DOF, 2013); however, evaluations are generally targeted at the school level rather than just the individual, as it usually happens with teachers. Given school principals' unique position as gatekeepers of education reform and promoters of productive environments for quality and equitable teaching and learning, this chapter contributes new light on current approaches to school principals' evaluations and the role this plays in educational accountability and improvement.

Regarding the accountability purpose of assessments, Leithwood (2001) identified four approaches that directly impact school leaders and schools. One of them is the creation of quasi-markets, that is, the education service emulating the assumed dynamics of the market, which allows parents to choose schools that, from their perspective, provide better quality and value. The second relates to decentralization: the devolution of responsibility of academic attainment to the schools, which holds them directly accountable. A third approach is the introduction of professional standards, which are, typically, government-created policy criteria specifying the expected performance of teachers and principals (Catano and Stronge, 2007; OECD-CEPPE, 2013), often explicitly linked to raising academic students' performance. Finally, management and strategic planning is seen as a method that school principals and staff are expected to use to maintain their institutions' working order to efficiently use resources and be effective in their outputs, i.e., educational indicators.

Although we acknowledge that school principals' evaluation, attitudes to school leadership, and impact on educational outcomes are rather inseparable, we seek to address the former area of inquiry due to its less common reference in the literature. Moreover, we recognize that while the term evaluation relates to the broader realm of performance monitoring, assessment refers to the practical and situated process of measuring or determining the extent of adherence to expected competencies and results (Goldring et al., 2009). Nevertheless, given the variety of uses given to school principals' assessments internationally, we use the terms interchangeably in this paper. The chapter is organized as follows: first, the methodology for the review is explained. Following this, a general scope of the literature is presented to highlight critical areas of school leadership assessment internationally and a brief review of 21 selected country contexts illustrating existing standards for school principals. The paper focuses on approaches to evaluation, including the most commonly used evaluation instruments. The review continues by synthesizing the link between school principals' evaluation, accountability, and professional development, and a conclusion is provided.

Methodology

This review addresses existing evaluation approaches to school principals drawing on academic research available on the internet, specialized journals, policy reports and official websites, consultations with experts in the field of school leadership, and the "gray" literature, for example, published masters and Ph.D. dissertations from 2000 to 2021. The key terms used for this review were as follows: "evaluation AND school principals" and its associated terminologies, for example, assessment, appraisal, headteacher, and school leader. Databases consulted included ERIC, ProQuest, Ethos, and Google Scholar in English. Redalyc, Scielo, and Dialnet in Spanish. Leading journals about school leadership were scanned, including the *Journal of Educational Leadership*, *International Journal of Leadership in Education*, *Educational Administration Quarterly*, *Education Policy Analysis Archives*, and *International Journal of Educational Management*.

The evidence presented focuses on 21 selected country contexts exemplifying existing school principals' evaluation approaches and the instruments most commonly employed for this purpose. Several studies precede this review, however, focusing on different aspects of school principals' accountability (OECD-CEPPE, 2013; OECD, 2013; OEI, 2017). The contribution of this chapter to existing knowledge lies in the identification of evaluation systems where there is a clear link with professional development schemes. Furthermore, the paper summarizes current practices in low- and middle-income contexts challenging to locate in the international literature.

Literature review

Educational Effectiveness Researchers agree on teacher quality and school leadership's critical role for quality education improvement (Reynolds et al., 2014; Mortimore et al., 1994; Kyriakides et al., 2018; Muijs et al., 2014; Scheerens, 2015; Day et al., 2006). Furthermore, it is believed that good school leaders can positively impact staff performance, community and parental engagement, and indirectly, pupils' academic outcomes (Goodall, 2018; Leithwood et al., 2020). Therefore, various approaches to accountability and motivation of teachers and school principals have been introduced worldwide, and Appendix 1 summarizes the 21 countries included in this review. Previous reviews in the USA indicate that although school principals are evaluated across the country, existing approaches present mixed evidence on their suitability to improve educational indicators. In many cases, they lack research-based support (Goldring et al., 2009; Davis et al., 2011). School principals' appraisals can be defined as "procedures involving external evaluators that seek to identify school leaders' strengths and weaknesses using objective criteria in order to make a judgement about their competencies for performance management, employment-related decisions and/or rewards purposes" (OECD, 2013, 492).

Like teacher evaluation systems, school principals can be subjected to assessments at two distinct points and for various reasons: before the appointment, for example, to verify suitability for the job, as well as once in the post. The latter is typically motivated by an interest in monitoring compliance to standards and regulations to determine permanency in the job, pay-related decisions, and advising on further professional development. Still, the specificities of school principals' evaluation might differ from those of teachers, depending on the country. While educational authorities centrally steer most school principal evaluation systems, there are other ways they are examined at the school community level. For instance, Gordon and Patterson (2006, 206) "view school leadership as a relational, context-specific process that considers the culture, norms, values, and beliefs of the individuals involved with the school and community". This implies that school principals may be required to meet national standards and are subject to expectations on their functions and ingrained in power negotiations at the community level (OEI, 2017).

In their literature review, Goldring et al. (2009) propose that school principals' evaluation revolves around two main aspects: what to assess and how to assess it. Regarding the former element, they suggest that four areas dominate: (a) responsibilities, (b) knowledge and skills, (c) processes, and (d) organizational outcomes. How to assess principals varies widely, and commonly a mix of instruments integrate a given evaluation system. Following Goldring et al. (2009) twofold distinction, we explore existing standards for school principals followed by current approaches to school principals' evaluation. Next, we report research on the most common evaluation instruments for principal's assessment and dedicate a section to understand the link between school principal's assessment and professional development.

Standards for school principals

Standards are typically depicted as descriptors of minimum performance in a given job. In education, standards are meant to orient teachers' and principals' day-to-day duties and might be used for hiring, evaluating achievement, make employment decisions, among other purposes (Cortez Ochoa et al., 2018; Wildy et al., 2011; Zheng, 2020; Catano and Stronge, 2007). A literature review on the topic suggests that Principal standards may focus on one or both general aspects: (i) professional functions and (ii) behavioral skills (OECD-CEPPE, 2013). Professional functions embrace setting the school's mission and liaizing with organizational and developmental aspects of teachers, for example, maintaining harmony in the school and orienting professional development. Behavioral skills include flexibility to cope with change, good communication, values, and the linkage between theory and practice (OECD-CEPPE, 2013). In this regard, Catano and Strong conclude that "defining the role of the school principal is a difficult task due to a complex set of job responsibilities, skills necessary to perform the job and values" (2007, 383).

The conceptual development of standards for school principals varies from place to place. While in some countries these are decided and enforced at the national policy level (e.g., Chile, Mexico, and Rwanda), in others, this happens at the regional policy level, regardless of the existence or not of national criteria (e.g., in Australia, Canada, Brazil, among others). Furthermore, the standards can be introduced top-down, as well as bottom-up. In some contexts, political authorities decide on the standards' content and means of observance (for example, Colombia, Ecuador, England). In contrast, in others, professional bodies or teachers' organizations can agree upon these criteria to frame their profession (e.g., Australia, Chile, New Zealand) (OECD-CEPPE, 2013). In some contexts, standards are mandated across different levels of governance (for instance, the USA). In this regard, Catano and Stronge (2007) note that the co-existence of standards of principals at national, state and local levels in the USA makes appraising the work of school principals harder.

Unlike standards expressly developed for teachers, the literature shows that school principals' standards are usually less precise in defining different performance progressions (OECD-CEPPE, 2013; Ni et al., 2019). Moreover, Romero and Krichesky (2019) found that professional standards are equal for teachers and school principals in Argentina. This lack of precision in standards to specify the career development of school principals might be because, contrarily to teachers, who might be encouraged to display incremental performance throughout a career, school principals seem to be expected to perform equally well from appointment through to retirement. Still, there are exceptions to the latter, such as Australia's Standard for Principals (AITSL, 2014). The Standard is organized into three leadership requirements and five professional practices, and seeks to inform professional profiles, particularly regarding two areas: Leadership Requirements and Professional Practices (AITSL, 2014). These criteria are adapted to the principal's school context and career stage. See Fig. 1.

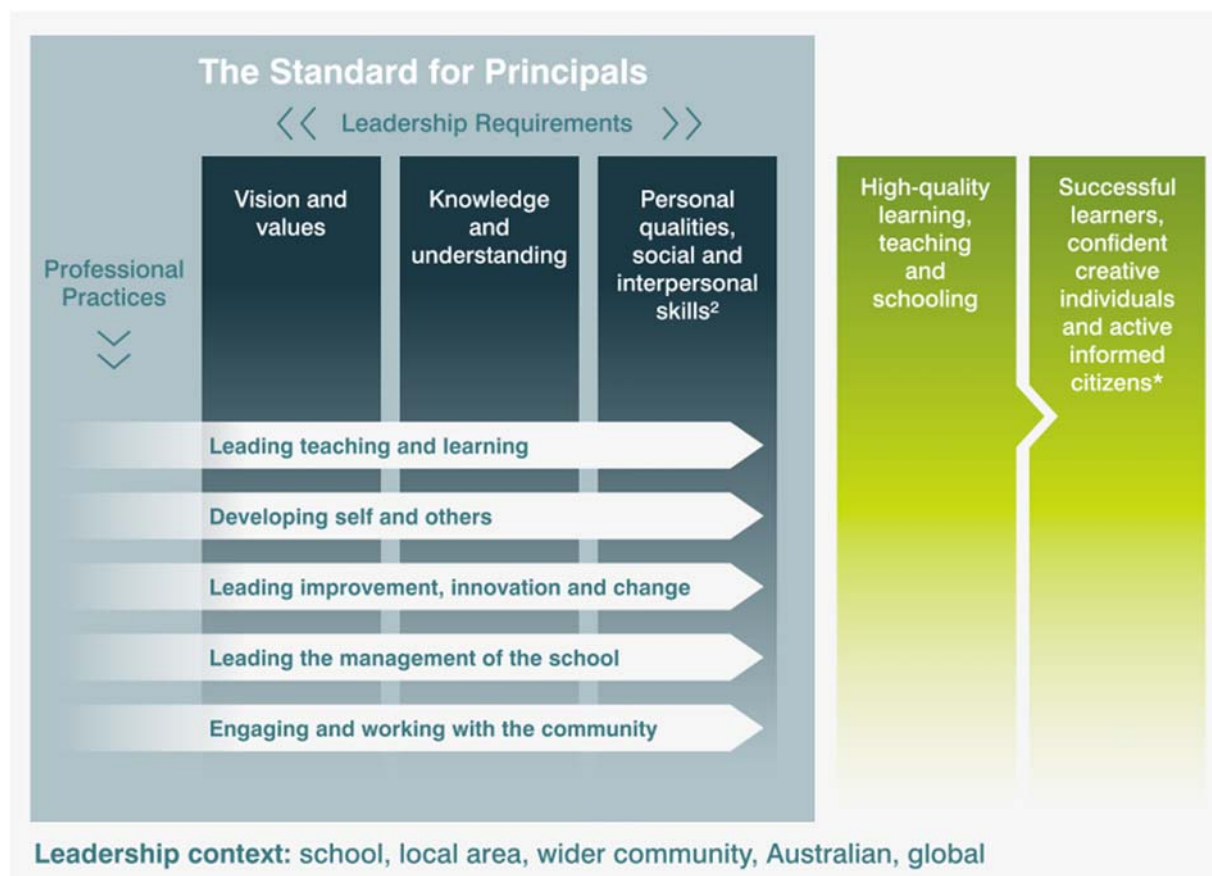


Fig. 1 The Standard for Principals—Australia. Source: AITSL (2014, 10).

Also, in Shanghai, China, the National Professional Standards for Compulsory Education Principals (NCEE, 2021) orient the progression from novice to Master Principal, and this development is explicitly structured by seniority, recognizing iterative mentoring from more experienced educators to support advancement, among other aspects that form part of a culture in this context. Shanghai is noteworthy as school principals typically transit a teacher career ladder before applying for a leadership position. This means that the extent of understanding the education system that school principals might achieve is outstanding. With the introduction of the new professional standards in 2013, principals are expected to deliver in instruction, organizational and school culture management, good relationships with the community, and professional development (Liu et al., 2017).

As can be seen in [Appendix 1](#), although the sole existence of Principal standards—or lack of them, as in Nicaragua and Costa Rica (Borden, 2002; OEL, 2017)—does not directly relate to the enactment of Principal evaluation systems (OECD-CEPPE, 2013), in some contexts, assessments have been introduced, as examined next.

Methods for evaluating school principals

Explicit professional standards are essential to designing and introducing evaluation schemes, as we noted in our previous review of teacher evaluation systems (Cortez Ochoa et al., 2018). Once criteria of what school principals should know and do are developed, these guidelines can help devise suitable methods to assess performance. In line with this view, Catano and Stronge indicate that “(1) principals’ performance evaluation should be fair and equitable; (2) [...] be based upon what they are expected to do; and (3) [...] evaluation instruments should match the expectations framed within state and professional standards” (2007, 385). Different approaches to evaluating school principals can be found, among which school inspections are one of the oldest and most commonly applied worldwide.

School inspections

School inspections have been a tradition in various European countries starting around the mid-19th century (Altrichter and Kemethofer, 2015). Through this assessment method, evaluators monitor and closely observe the school’s compliance with performance standards. Often, inspections are carried out by external personnel, such as civil servants or equivalent, who review documentation and student assessment data from the school and may also gather information from direct observations and internal school self-evaluation. Inspections typically result in an overall evaluation and potential recommendations for improvement. England has a long practice in this respect, and several examples can be found in the USA. In Latin America, this is a commonly found approach and is currently used in Brazil, Chile, Costa Rica, Cuba, Nicaragua, and the Dominican Republic. Following a recent reform in the Emirate of Abu Dhabi, school inspections were introduced. In a handful of cases, school inspections are meant to review school improvement plans; we identified this in Argentina, Chile, Costa Rica, Ecuador, Rwanda, and Spain.

Inspections can focus to differing extents on accountability or improvement purposes and contribute to School Principal professional development routes, as we discuss later. Nevertheless, the method has been found to contribute to stress and discontent. For instance, an online survey with 2300 school leaders in seven European countries revealed that among the case studies, The Netherlands and the UK reported high levels of perceived pressure coming from school inspections (Altrichter and Kemethofer, 2015). This investigation found that high-stakes consequences attached to checks contribute to principals’ perceptions of stress. Also, pressure effects on various developmental realms, such as promoting/improving self-evaluation, are found among countries with stricter high-stakes consequences from inspection outcomes. The study’s findings indicate that the perceived burden of inspections also leads to unintended consequences, such as a diminution of innovative teaching and curriculum shrinkage. There is much emphasis placed on the capacity of inspectors to exert unbiased judgments on a school’s performance and the principal’s competency to lead the institution. Thus, the reliability of this approach should be carefully considered when using inspection as a method of principal assessment (Martínez Rizo, 2016).

Self-review

Increasingly common in the toolkit for school principals’ assessments, self-evaluation is a method external evaluators use to ponder the extent of standards’ accomplishment, based on records provided by the school principal. We identified that this approach is utilized in Australia, Chile, Dominican Republic, New Zealand, Scotland, and Abu Dhabi. The specific characteristic of this technique is context-specific, including leadership portfolios, administrative records, etc. This type of assessment might be used alongside other instruments, for instance, written examinations, as in Australia and the Dominican Republic. Scotland is unique because self-evaluation is the primary method to appraise the school principals’ performance and is directly linked to professional development (The General Teaching Council For Scotland, 2012).

As much a progressive approach self-reviews appear, mainly as these are found around discourses of teacher agency, criticism persists on this method as paradoxically centering the limelight on accountability rather than professionalism (Priestley et al., 2015). Making up evidence of good performance is a potential pitfall of self-evaluation tools and approaches, especially when there are consequences for the job attached to them (Webb, 2006; Ball, 2003). Several examples from the USA and the UK provide grounds to recommend caution when self-review is meant to decide principals’ incentives and sanctions.

Written assessments

Principal exams, especially those employing multiple-choice items, were found in more than half the 21 review contexts. Although tests might be used as part of school principals’ evaluation once in the post (e.g., Brazil, Costa Rica, Cuba, Dominican Republic,

Ecuador, Mexico, among others), often, written assessments are used for hiring, as in Argentina, Australia, Colombia, Peru, and the USA. Tests provide the possibility of having a comparable measure between different candidates' competencies. Thus, it serves a practical purpose and might be cost-effective when deciding who to grant a principal's role and determine the extent of mastery of the professional duties related to the position.

We also found that, increasingly, exams integrate hypothetical scenarios or vignettes meant to assess how the respondent would react to a given situation in the school (Wildy et al., 2011). However, the intention of these fictitious circumstances might not be ideal as a proxy of actual competency. There is much debate in the literature regarding the validity and reliability of written exams to assess pedagogical knowledge, subject knowledge, and other types of skills and competencies unique to school principals (Darling-Hammond, 2001; Hill et al., 2004; Mayer and Mills, 2020; Molina et al., 2018), some of which we discussed in the previous section. Therefore, its use in combination with other indicators of suitability for the role seems appropriate (Steinberg and Sartain, 2015).

Surveys with the school community

Among the cases reviewed in this chapter, we identified seven contexts where school principals' evaluations incorporate the views of the staff and the broader community. These instances are Australia, Brazil, Chile, Cuba, England, New Zealand, and Paraguay. The procedures concerning data collection vary, from centralized electronic forms sent to the schools, to the co-participation of an inspector who conducts interviews or disseminates questionnaires. Two cases stand out due to the involvement of school boards: England and New Zealand. In these cases, parents, teachers' representatives and people from the community may provide input regarding the perceived performance of the school to hold the principal to account (National Governance Association, 2019; DfE, 2017; NZSTA, 2021).

Another notable case is Brazil, where school principals' hiring may be organized around a democratic election where parents vote for a candidate to become the new head of school (Aravena, 2020). These examples resonate with Leithwood's (2001) research regarding the role of quasi-markets, school autonomy, performance-related monitoring, and approaches to school leadership and management in the accountability of school principals. Notably, where decisions on the school principal's performance are in the hands of a broader group of individuals, appropriate training, communication, and alignment with standards should ensue.

Student learning indicators

Invariably, school principals' evaluations seek to impact students' outcomes positively. We found that in some countries, school principal evaluation utilizes students' learning indicators alongside other measures, which happens in Australia, England, Shanghai-China, and the USA. Still, how these measures are factored into a school principal's assessment varies. For example, in the USA, Chiang et al. (2016) studied the suitability of school value-added to evaluate principals' effectiveness. Using Pennsylvania's longitudinal data, they concluded that such metrics are poor indicators of principal's effectiveness. Loeb and Grissom (2013) arrived at similar conclusions. The use of school value-added, which accounts for growth in students' results in standardized tests across time, has been critiqued due to its limitations in isolating the effects of principals from anything else beyond their control. It has proved challenging to quantify the principals' effects while accounting for teachers' effects. However, a recent systematic synthesis of six longitudinal studies involving more than 22 thousand principals in the USA found that

a 1 standard deviation increase in principal effectiveness increases the typical student's achievement by 0.13 standard deviations in math and 0.09 standard deviations in reading. [In other words], replacing a below-average elementary school principal (i.e., one at the 25th percentile of effectiveness) with an above-average principal (i.e., at the 75th percentile) would result in an additional 2.9 months of math learning and 2.7 months of reading learning each year for students in that school.

Grissom et al. (2021, xiii).

Still, the authors acknowledged that more research is needed to understand how principals' effects on learning are mediated by their interventions in recruiting and maintaining effective teachers in the school. Mainly because most principal impacts on students' achievement are indirect, with some evidence of direct effects, such as behavioral management. In other words, principals' effectiveness can be challenging to measure and plausibly context-situated, fluid rather than static. It is critical to reflect on research that has used measures of school value-added to determine school principals' effectiveness, given this approach makes several assumptions about the role of school principals in matters of students' outcomes growth. For example, when it is suggested that replacing less effective with more effective principals, this idea somehow assumes that principals arrive at schools to exert changes without any resistance from existing personnel, students, and parents and that their predicted effectiveness pertains to them, and not to the various factors around them, which might also shape their approach to educational growth. Given the complexity underpinning these latter contentions, an in-depth review in a separate space would be worthy.

Our review identified several instances where the intention was to collect data on students' outcomes, however, this happens indirectly. In Abu Dhabi, Brazil, Costa Rica, Ecuador, Paraguay, Rwanda, Scotland and Spain, principals must provide administrative records of school exam performance. Research contests the limitation of this appraisal form, mainly because of the risks of cheating, especially when high-stakes consequences are attached to the evaluations (Carpenter et al., 2014; Webb, 2006; Ball, 2003).

The evaluation of school principals might lead, as mentioned earlier, to a series of decisions on their job and professional career. In addition, it has been consistently agreed that centralized evaluation systems, for example, steered at a national or state level, usually stress the accountability and high-stakes consequences side of the assessments (OEI, 2017). We identified that school

principals' assessments in less than half of the reviewed countries (8 out of 21) focus on the individual, while the rest combine school and individual evaluations. England and New Zealand stand out given their emphasis on school monitoring for accountability purposes. The following section examines the link between the school principal's evaluation and opportunities for professional development.

School principals' professional development

One of the primary reasons for school principals' evaluations is to support their professional development and, eventually, positively impact their school communities' outcomes. A review of the literature in the USA identified that school principals are more receptive to assessments that align with their expected professional duties and contribute to their professional development in areas that are relevant to their schools and contexts (Davis et al., 2011). In this research, we identified that in all reviewed countries, there exists a training component after the appointment. Still, their features could not be fully untangled via the available literature or official documentation. We noted that the main topics on offer are school management, leadership, evaluation, legislation, and less commonly, curriculum and ICT skills. In two cases: Colombia and Scotland, the evaluation system is meant for the school principal to decide and customize what professional development they need. In Costa Rica, Shanghai-China, Nicaragua and Paraguay, school principals are expected to achieve a number of professional development hours. Mentoring programs were less commonly found, only reported for England, Rwanda, Scotland, and Shanghai-China. In the case of Rwanda, development partners have introduced Diploma and Certificate CPD for current principals with The University of Rwanda-College of Education (URCE) and the Rwanda Education Board (REB) (VVOB, 2019).

In the broader literature, the professional development of school principals is typically found in any (or a mix) of the following types: traditional workshops and seminars; peer development, such as one-to-one mentorship; and programs aimed at developing the principal's investigative skills (Berkowitz et al., 2019). Mentoring schemes supported by experienced peer school leaders and professional development initiatives in everyday settings are particularly appealing to school principals as a suitable method for improvement (Hamilton et al., 2012; Christie et al., 2019). Mentoring and coaching programmes can be found in several Australasian countries (Macpherson, 2009), although the specificities and combination with other training vary from place to place. In general, opportunities for professional development must be congruent with the identified strengths and pitfalls of the school principal's performance in an evaluation (Hamilton et al., 2012). This is critical for the credibility of assessments and the eventual improvement of leadership and schools.

Conclusions

This chapter presented an overview of current school principals' evaluation approaches, drawing on examples from 21 countries. The literature from the last twenty years shows that professional standards for school principals and assessments are widely used in many contexts. The instruments for assessment vary substantially in every context analyzed; still, the most commonly identified approaches in this review include school inspections, self-reviews, written examinations, community surveys, and student learning indicators. In the cases examined, mechanisms targeted at CPD provision exist, although these varied substantially. We chose to focus on the developmental side of school principals' assessments, given their potential for school improvement.

The study of the effects of school principals on students' outcomes is an area still in exploration with no definite conclusion in some cases (Golding et al., 2009; Davis et al., 2011), although very recent findings have been more positive (Grissom et al., 2021), indicating that effective school principals matter for students' academic growth. Most importantly, given the indirect impact of principals via teachers, school management, and other realms, it is sensible to understand the positive impacts of effective school principals in their intricate nature as leaders, colleagues, administrators, and potentially many more phases. Research recommends caution in focusing on educational outcomes solely in the context of principal evaluation, potentially overlooking the rest of the school principal's tasks within the school community that also influence results (Golding et al., 2009), including those related to behavioral management (Grissom et al., 2021).

In revisiting their highly regarded paper "Seven strong claims about successful school leadership", Leithwood et al. (2020) provide a more nuanced approach to the influence of school principals in students' academic outcomes. In the light of recent evidence, they maintain that principals are important within-school factors, but reflect on current knowledge about the influence of other contextual and out of school factors that more directly explain school effects on students' attainment, in line with the Educational Effectiveness literature (Creemers and Kyriakides, 2006; Reynolds et al., 2012; Scheerens, 2015). There is enough evidence to suggest that school principals can influence students' achievement, and therefore, Leithwood and colleagues' seminal work remains relevant (OECD-CEPPE, 2013; OECD, 2013; Day et al., 2009; Grissom et al., 2021). More importantly, continuing studying school principals' evaluation can be a powerful tool to help decision-makers and researchers understand better ways to support principals' professional development, schools and learning.

Future research might need to ponder the feasibility of assessing the effects of school principals on educational outcomes, especially when leadership is increasingly a shared responsibility at the school and system levels (Daniëls et al., 2020; Leithwood et al., 2020). This might require sophisticated methodologies capable of factoring in teachers' and educational authorities' forms of leadership. Most importantly, such approaches should also offer information on how principal evaluation can provide a means to enhance teaching and learning. Some additional responsibilities being transferred to schools include the emotional development

of students and teachers, their contribution to environmental sustainability, fighting racism, among others (see also, [Facer, 2021](#); [Saeed, 2020](#); [Sriprakash et al., 2020](#)). Thus, looking more closely at the validity, reliability, and unintended consequences of school principal evaluation methods and the potential to improve the employed measures is also a critically important focus for further research.

Limitations

Given the space limits, a selection of school principals' evaluation approaches had to be made, potentially omitting other systems that could provide additional elements to our understanding of the topic. [Avalos \(2016\)](#) suggests that colonial legacies play a critical role in the standards and approaches to teacher education and regulation. In that sense, it is expected that, for instance, methods used in Canada and Australia share characteristics with those used in the UK; similarly, Latin American contexts might rely on frameworks close to those from Spain or Portugal. Research in China further shows that standards are increasingly based on Western frameworks, such as the USA's ([Liu et al., 2017](#)), suggesting that policy borrowing may also be a phenomenon for future study. Still, the school principal's evaluation is a dynamic arena that evolves to respond to context-specific needs, as demonstrated by some of the cases explored. Therefore, the review is even more relevant now.

Furthermore, there are several equally essential aspects related to school principals' evaluations that deserve tailored review. These include the governance of school principals' evaluations, their recommended frequency, and the utility of assessments for other school leaders different from the principal. Some of the references in this paper articulate these areas, *albeit* to varying extents, and thus, we encourage its exploration. We trust this chapter served its purpose of providing a much-needed update on the matter of the title and that as the field unfolds, new knowledge will be needed.

Appendix 1

Country	Standards or criteria for school principals	Evaluation methods and instruments	Relationship with professional development	Accountability at the individual level, the school level or both	Selected references
Argentina	Do not exist. Romero's and Krichesky's (2019) study in five provinces found that school principals have equal regard than teachers in current normative and that requirements depend on provincial legislation.	Entry requirements: • A teaching degree • Be a teacher before applying for a promotion to school leader • Teaching experience • Good notes on performance assessment • A written exam • An oral defense of a school improvement plan To retain the post • Once granted, a school principal remains in the position until retirement	Vary substantially from province to province and might relate to: • Management of school staff development • School leadership and the context • Curriculum management • ICT	Individual (only for hiring)	(OEI, 2017 ; Romeo and Krichesky, 2019)
Australia	National professional standards for principals and the leadership profiles	• Self-review of school improvement plan (typically reviewed annually). • Additional datasets on student achievement, teacher and community surveys. • Written assessments with hypothetical scenarios^a.	• Professional development decisions included in the Performance and Development Culture accreditation program^a.	Both	(Gurr, 2007 ; Wildy et al., 2011)

(Continued)

Country	Standards or criteria for school principals	Evaluation methods and instruments	Relationship with professional development	Accountability at the individual level, the school level or both	Selected references
Brazil	Each region monitors school principals using specific criteria	Entry requisites (Rio de Janeiro) • Five years of teaching experience and a teaching degree (direct democratic vote from school staff and the community) Other assessment instruments in different regions: • Paper-based questionnaires • Interview • Observation protocols • Administrative registers	• Traditional courses and seminars. These courses vary in different regions.	Both	(OEI, 2017; OECD, 2019; Aravena, 2020)
Chile	Marco para la Buena Dirección y el Liderazgo Escolar [Framework for Good Management and School Leadership] Practical dimensions • Constructing and implementing a strategic and shared vision • Developing professional capacities • Leading teaching and learning • Managing convivence and the school community's participation • Developing and managing the school Personal resources • Principles • Abilities • Professional knowledge	In agreement with the local authority or founder: • Self-evaluations • School development plan reports • Superintendent report • Focus groups with teachers and parents	• A mentorship scheme including technical support is currently in development. • Potential providers are universities, specialized development centers, and existing institutions concerned with teacher CPD.	Both	(MINEDUC, 2015, 2020; Benavides Moreno et al., 2019; Berkowitz et al., 2019)
Colombia	Estatuto Docente, Decreto Ley 2277 (1979) [Teaching Decree, Law 2277 (1979)].	Entry requirements: • A written exam on general and pedagogical knowledge • Critical reading • Administrative management • Academic management • Psycho-technical evaluation • Interview Continuous formative assessment • The teacher selects observable and measurable evidence of compliance to the Law. • A Personal and Professional Development Plan addressing four aspects: management, pedagogy, administrative duties, and relationship with the community.	• The Personal and Professional Development Plan is seen as contributing to the teachers' reflexivity and enhancement.	Individual	(OEI, 2017; Aravena, 2020)

(Continued)

Country	Standards or criteria for school principals	Evaluation methods and instruments	Relationship with professional development	Accountability at the individual level, the school level or both	Selected references
Costa Rica	There are no standards. Assessments are aligned with current legislation: Decreto Ejecutivo N° 38170- MEP February 13, 2014 [Executive Decree Ejecutivo N° 38170- MEP February 13, 2014].	• School supervision (observation protocols) • Administrative records • Paper-based questionnaires (aligned with learning content from Costa Rica's university programmes in education) 70 items addressing the following: • School project planning • Research • Education legislation • Administration • Professional ethics • Educational planning • Statistics • Paper-based questionnaires • Observation protocols • School management records • Administrative records • School community surveys	• Hybrid modules of 40 h seminars in total. Topics include: • Leadership, conflict management • Evaluation and monitoring of educational programmes • Educational legislation • Human resource management	Both	(Borden, 2002; OEI, 2017; Oviedo Bogantes, 2017)
Cuba	The certificate of evaluation sets the personal and professional attributes of school principals. • Personal qualities • Communicative skills • Indicators on ethics, discipline and efficiency • Organizational and managerial attributes • Results	• Paper-based questionnaires • Observation protocols • School management records • Administrative records • School community surveys	• Feedback via methodology exchange with inspectors.	Both	(Maturell Aguilera and Valiente Sandó, 2016; OEI, 2017)
Dominican Republic	Teachers and school principals performance standards. • Promote a vision and school development plan with the community • Manage, organize, and supervise teaching and learning • Procure an environment for effective learning • Promote the staff's professional development • Encourage the families and the community to participate with the school • Self-professional development	• Evaluation via a rubric by the immediate authority above the school principal • Self-evaluation • A written exam	The School of Principals for Quality Education (EDCE in Spanish) provides tailored professional development courses.	Individual	(MINERD, 2014a,b; IDEICE, 2017; OEI, 2017; Vaillant, 2019)
Ecuador	Quality standards for school principals' performance • Institutional management • Pedagogical management • Convivence, school participation and co-operation • School safety	Entry requirements • Teaching degree and a postgraduate degree in education • To pass all assessments • Five years of teaching experience • To master an indigenous language where applicable Continuous evaluation • A written exam with 60 items with four response options each. The test includes questions on the following: • Educational administration • Basic accounting • Curricular planning • Strategic planning and project design • Educational assessment • Psychology in education	• Mentorship and induction programmes • Regular workshops and Diploma courses	Individual	(Ministerio de Educación, 2010, 2017; OEI, 2017)

(Continued)

<i>Country</i>	<i>Standards or criteria for school principals</i>	<i>Evaluation methods and instruments</i>	<i>Relationship with professional development</i>	<i>Accountability at the individual level, the school level or both</i>	<i>Selected references</i>
England	The Headteachers' standards 2020 are organized in two sections: (1) Ethics and professional conduct; and (2) Headteachers' standards, which include: (1) School culture, (2) Teaching, (3) Curriculum and assessment, (4) Behavior, (5) Additional and special education needs and disabilities, (6) Professional development, (7) Organizational management, (8) Continuous school improvement, (9) Working in partnership, (10) Governance and accountability.	Governing bodies undertake tailored performance management, including: • 360-degree feedback, • Leadership portfolio, • School inspections (Ofsted), • Staff, and student interviews.	Professional development programmes, such as mentoring and specific training at accredited institutions, and via school-based development.	Both	(Altrichter and Kemethofer, 2015; DfE, 2020; Eurydice, 2021)
Mexico	Professional profile with a series of dominions, criteria, and indicators regarding: • Professional practice and professional development in line with philosophical, ethical, and legal principles of Mexico's education. • Recognition of the function to promote a school culture grounded on equity, inclusion, interculturality and excellence. • School organization to promote the integral development of girls, boys, and adolescents. • Promotion of shared responsibility between the school, families and educational authorities.	The recently introduced System for the teachers' career mandates sitting computer-based tests for promotion and salary improvement. The evaluation considers: • Professional trajectory • Content knowledge and administrative knowledge	• Individualized written report of performance • Professional development courses in centralized or regional catalogs • Dedicated YouTube channel and a website	Individual	(DOF, 2019; USICAMM, 2020)
New Zealand	Professional standards for school principals (primary, area, and secondary): • Culture • Pedagogy • Systems • Partnerships and networks	There are two types of evaluations: • An annual school self-review (supported by a Board of Trustees) The Education Review Office conducts an assessment of the school every 3 years, or earlier if development is needed. Their inspection-like procedure is based on six dimensions: • Teaching • Leading and managing • Governing • School culture • Engaging families • Paper-based questionnaires • School visits (including mentoring)	Results form part of the school improvement plans. • Several programmes are on offer	Both	(Nusche et al., 2012; ERO, 2021; Ministry of Education, 2021)
Nicaragua	There are no standards		Two sessions of two weeks each (180 h in total). Topics include: • Education in a national and global context • Pedagogy and leadership • Leadership and school improvement	Both	(Borden, 2002; OEI, 2017)

(Continued)

Country	Standards or criteria for school principals	Evaluation methods and instruments	Relationship with professional development	Accountability at the individual level, the school level or both	Selected references
Paraguay	As part of the new manual for the monitoring and management of educators and educational institutions, school principals are expected to perform in four dimensions: • Pedagogy: efficient and critical management of the national curriculum • Personal and professional development • Administration: efficacy and efficiency in resource management • Social engagement: promotes the communities' participation.	• Paper-based questionnaires • Administrative records • School community surveys	10 modules totaling 1 thousand hours of training. Topics include: • Administration for quality • Educational change and the school principal's role • Technical and pedagogical issues • School organization • Efficient school management • Strategic planning • Human resource management • Educational evaluation • Diploma in school management and leadership	Both	(Borden, 2002; OEI, 2017)
Peru	Framework for good school leader performance. It is organized into two domains, seven competencies and 21 observable indicators. The two domains are as follows. • Institutional management for learning improvement • Pedagogical management for learning improvement.	Entry requirements Centralized • To pass the unique national exam (PUN in Spanish), which includes: • Reading test • Management sample cases • Pedagogical knowledge Decentralized • Professional trajectory • Merits • Professional experience		Individual	(OEI, 2017; Aravena, 2020)
Rwanda	National School Leadership Standards 1. Creating strategic direction for the school 2. Leading learning 3. Leading teaching 4. Managing the school as an organization 5. Working with parents and the wider community	• Written evidence of school improvement plans and achievements. • Records of staff meetings • Student outputs • Records of classroom visits • Evidence of communication with the community and the staff.	CPD Diploma Program in Effective School Leadership; and CPD Certificate Program in Educational Mentorship and Coaching. As part of the CPD Diploma, headteachers complete eight practice-based written assignments and a Portfolio of Evidence. Sit a written assessment and undergo a field visit by trainers to feedback on the assessment tool. The Certificate Program includes two visits, one formative and one summative.	Both	(VVOB, 2019; REB, 2020)

(Continued)

<i>Country</i>	<i>Standards or criteria for school principals</i>	<i>Evaluation methods and instruments</i>	<i>Relationship with professional development</i>	<i>Accountability at the individual level, the school level or both</i>	<i>Selected references</i>
Scotland	Standards for leadership and management: (i) Establish, sustain and enhance the culture of self-evaluation for school improvement; (ii) Develop staff capability, capacity and leadership to support the culture and practice of learning; (iii) Ensure consistent high-quality teaching and learning for all learners; (iv) Build and sustain partnerships with learners, families and relevant partners to meet the identified needs of all learners; (v) Allocate resources effectively in line with identified strategic and operational priorities.	• Self-evaluation	• The National Model of Professional Learning • Coaching and Mentoring	Individual	(The General Teaching Council For Scotland, 2012)
Shanghai-China	National Professional Standards for Compulsory Education Principals (2013). Organized in five specific principles: 1. Morality first 2. Talent's cultivation orientation 3. Leading the development 4. Emphasis on abilities 5. Life-long learning	• Candidate's character and professionalism as well as school outcomes.	National Training Program for Primary and Secondary School Principals, based on the 2013 standards. • 300 h of training, including classwork and practicum, within their first six months on the job. Training is conducted at two national centers. • Once in the post, principals are required to take 360 h of training every 5 years. • Promotion in the position is possible via the career ladder, which has five levels, culminating in the Master Principal role.	Individual	(Liu et al., 2017; NCEE, 2021)
Spain	Each autonomous community determines them	Entry requirements: • At least 5 years of teaching experience • A certificate in school leadership • A school improvement project • To retain the post: • Interview • School visit (observation protocol) • School management records • Administrative records	Vary in every community and may include training on: • Continuous development of school leadership skills • Educational planning • Teamwork and pedagogical innovation	Both	(OEI, 2017; Díaz-Delgado and García-Martínez, 2019)

(Continued)

Country	Standards or criteria for school principals	Evaluation methods and instruments	Relationship with professional development	Accountability at the individual level, the school level or both	Selected references
United Arab Emirates (Abu Dhabi)	The professional standards for principals revolve around five areas concerned with leading: - Strategically - Teaching and learning - The organization - People - The community	- Using the so-called irtqaa tool, an electronic form that school leaders must fill out, principals prepare for school inspections. - School inspections are based on self-evaluations derived from irtqaa.	School feedback Professional development in areas such as: “creating a compelling vision, leading through technology, understanding leadership styles, promoting teamwork, supporting student-centered learning, implementing the NSM curriculum, developing high-quality teachers and building a student-centered environment” (Blaik Hourani and Litz, 2019, p. 225)	Both	(ADEC, 2011; Blaik Hourani and Litz, 2019)
United States of America ^b	–	–	–	–	(Goldring et al., 2009; Davis et al., 2011; Hamilton et al., 2012; Laing et al., 2016)

^aThis applies for school principal hiring in Western Australia; development programmes relate to schools in Victoria.

^bIn the United States of America, given their decentralized education system, there are many approaches to school principals' assessments, some of which operate at the State, District or school level. Therefore, it is advised to review the cases of interest individually.

References

- ADEC, 2011. Professional Standards for Principals. Abu Dhabi Education Council. https://www.unrwa.org/sites/default/files/adece_professional_standarda_for_principals.pdf.
- AITSL, 2014. Australian Professional Standard for Principals and the Leadership Profiles. The Australian Institute for Teaching and School Leadership (AITSL), Melbourne. https://www.aitsl.edu.au/docs/default-source/default-document-library/australian-professional-standard-for-principals-and-the-leadership-profiles652c8891b1e86477b58ff00006709da.pdf?sfvrsn=11c4ec3c_0.
- Akiba, M., Liang, G., 2016. Effects of teacher professional learning activities on student achievement growth. *J. Educ. Res.* 109 (1), 99–110. <https://doi.org/10.1080/00220671.2014.924470>.
- Altrichter, H., Kemethofer, D., 2015. Does accountability pressure through school inspections promote school improvement? *Sch. Effect. Sch. Improv.* 26 (1), 32–56. <https://doi.org/10.1080/09243453.2014.927369>.
- Aravena, F., 2020. Procesos de Selección de Directores Escolares en latinoamérica: Brasil, Chile, Colombia y Perú. *Educ. Pol. Anal. Arch.* 28 (171). <https://doi.org/10.14507/epaa.28.4666>.
- Avalos, B., 2016. Learning from research on beginning teachers. In: Loughran, J., Mary, L. (Eds.), *International Handbook of Teacher Education: Volume 1*, vol. 1. Springer, Hamilton, pp. 487–522. https://doi.org/10.1007/978-981-10-0366-0_14.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228. <https://doi.org/10.1080/0268093022000043065>.
- Barber, M., Mourshed, M., 2007. How the World's Best-Performing School Systems Come Out on Top. McKinsey & Company. <https://www.mckinsey.com/industries/education/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top>.
- Benavides Moreno, N., Donoso-Díaz, S., Reyes Araya, D., 2019. Directors at Chilean public schools seen from a performance framework. *Cad. Pesqui* 173, 130–153. <https://doi.org/10.1590/198053146317>.
- Berkowitz, D., Fernández, M., Zoro, B., 2019. Una Revisión Al Desarrollo Profesional Directivo y de Nivel Intermedio En El Contexto Del Desafío de Creación Del Sistema de Educación Pública [A Review of School Principal Professional Development]. Technical Report No 3.
- Blaik Hourani, R., Litz, D.R., 2019. Aligning professional development, school self-evaluation and principals' performance standards: a UAE case study. *Sch. Leadersh. Manag.* 39 (2), 222–249. <https://doi.org/10.1080/13632434.2018.1479848>.
- Borden, A., 2002. Directores de Escuela En América Latina y El Caribe: ¿Líderes Del Cambio ó Sujetos a Cambio? [School Principals in Latin America and the Caribbean: Change Leaders or Subjects to Be Changed?]. Inter-American Development Bank. <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=646201>.
- Bush, T., Glover, D., 2016. School leadership and management in South Africa: findings from a systematic literature review. *Int. J. Educ. Manag.* 30 (2), 211–231. <https://doi.org/10.1108/JEM-07-2014-0101>.
- Carpenter, B.W., Diem, S., Young, M.D., 2014. The influence of values and policy vocabularies on understandings of leadership effectiveness. *Int. J. Qual. Stud. Educ.* 27 (9), 1110–1133. <https://doi.org/10.1080/09518398.2014.916008>.
- Catano, N., Stronge, J.H., 2007. What do we expect of school principals? Congruence between principal evaluation and performance standards. *Int. J. Leadersh. Educ.* 10 (4), 379–399. <https://doi.org/10.1080/13603120701381782>.

- Cheng, Y.C., Ko, J., Lee, T.T.H., 2016. School autonomy, leadership and learning: a reconceptualisation. *Int. J. Educ. Manag.* 30 (2), 177–196. <https://doi.org/10.1108/IJEM-08-2015-0108>.
- Chiang, H., Lipscomb, S., Gill, B., 2016. Is school value added indicative of principal quality? *Educ. Finance Policy* 11 (3), 283–309. https://doi.org/10.1162/EDFP_a_00184.
- Christie, M., Simon, S., Graham, W., Call, K., Farragher, Y., 2019. Bungee jumping and rocket launching: transformative learning for today's transformational school leaders. *Int. J. Educ. Manag.* 33 (7), 1610–1624. <https://doi.org/10.1108/IJEM-09-2018-0288>.
- Cortez Ochoa, A.A., Thomas, S., Tikly, L., Doyle, H., 2018. Scan of International Approaches to Teacher Assessment, vol. 06. Bristol Working Papers. Bristol, UK. http://www.bristol.ac.uk/media-library/sites/education/documents/Scan_of_International_Approaches_to_Teacher_Assessment.pdf.
- Creemers, B.P.M., Kyriakides, L., 2006. Critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model. *Sch. Eff. Sch. Improv.* 17 (3), 347–366. <https://doi.org/10.1080/09243450600697242>.
- Daniëls, E., Hondégheem, A., Heystek, J., 2020. School leaders' and teachers' leadership perceptions: differences and similarities. *J. Educ. Adm.* 58 (6), 645–660. <https://doi.org/10.1108/JEA-11-2019-0199>.
- Darling-Hammond, L., 2001. Teacher testing and the improvement of practice. *Teach. Educ.* 12 (1), 11–34. <https://doi.org/10.1080/10476210123029>.
- Davis, S., Kearney, K., Sanders, N., Thomas, C., Leon, R., 2011. The Policies and Practices of Principal Evaluation: A Review of the Literature. WestEd, San Francisco, CA.
- Day, C., Stobart, G., Sammons, P., Kington, A., Gu, Q., Smees, R., Mujtaba, T., 2006. Variations in Teachers' Work, Lives and Effectiveness: Final Report for the VITAE Project. Research Report 743.
- Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K.A., Gu, Q., Penlington, C., Mehta, P., Kington, A., 2009. The impact of school leadership on pupil outcomes final report. *Nat. Coll. Sch. Leadership* 1–132. www.dcsf.gov.uk/research.
- DfE, 2017. The Constitution of Governing Bodies of Maintained Schools Statutory Guidance for Governing Bodies of Maintained Schools and Local Authorities in England. Department for Education, United Kingdom. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/640562/The_constitution_of_governing_bodies_of_maintained_schools_2017.pdf.
- DfE, 2020. Headteachers' Standards 2020. Department for Education, United Kingdom. <https://www.gov.uk/government/publications/national-standards-of-excellence-for-headteachers/headteachers-standards-2020>.
- Díaz-Delgado, M.A., García-Martínez, I., 2019. Estándares para la dirección escolar en España y México: un estudio comparativo. *Educ. Policy Anal. Arch.* 27, 113. <https://doi.org/10.14507/epaa.27.4565>.
- DOF, 2013. DECRETO Por El Que Se Expide La Ley General Del Servicio Profesional Docente. http://dof.gob.mx/nota_detalle.php?codigo=5313843&fecha=11/09/2013.
- DOF, 2019. Ley General Del Sistema Para La Carrera de Las Maestras y Los Maestros. Diario Oficial de La Federación. Mexico. http://creson.edu.mx/docs/Perfiles_profesionales_Criterios_e_Indicadores_para_Docentes_Ciclo_Escolar_2021-2022.pdf.
- ERO, 2021. Education Review Office. Education Review Office, New Zealand. www.ero.govt.nz.
- Eurydice, 2021. Continuing Professional Development for Teachers Working in Early Childhood and School Education. European Union. https://eacea.ec.europa.eu/national-policies/eurydice/content/continuing-professional-development-teachers-working-early-childhood-and-school-education-87_en.
- Facer, K., 2021. Futures in Education: Towards an Ethical Practice. Paper Commissioned for the UNESCO Futures of Education Report. <https://unesdoc.unesco.org/ark:/48223/pf0000375792.locale=en>.
- Goldring, E., Cravens, X.C., Murphy, J., Porter, A.C., Elliott, S.N., Carson, B., 2009. The evaluation of principals: what and how do states and urban districts assess leadership? *Elem. Sch. J.* 110 (1), 19–39. <https://doi.org/10.1086/598841>.
- Goodall, J., 2018. A toolkit for parental engagement: from project to process. *Sch. Leadersh. Manag.* 38 (2), 222–238. <https://doi.org/10.1080/13632434.2018.1430689>.
- Gordon, J., Patterson, J.A., 2006. School leadership in context: narratives of practice and possibility. *Int. J. Leadersh. Educ.* 9 (3), 205–228. <https://doi.org/10.1080/13603120600797831>.
- Grissom, J.A., Egalite, A.J., Lindsay, C.A., 2021. How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research. <http://www.wallacefoundation.org/principalsynthesis>.
- Guerriero, S., 2013. Teachers' Pedagogical Knowledge and the Teaching Profession: Background Report and Project Objectives. OECD Publishing. https://www.oecd.org/education/ceri/Background_document_to_Symposium_ITEL-FINAL.pdf.
- Gurr, D., 2007. Diversity and progress in school accountability systems in Australia. *Educ. Res. Policy Pract.* 6 (3), 165–186. <https://doi.org/10.1007/s10671-007-9021-2>.
- Hamilton, L.S., Engberg, J., Steiner, E.D., Awsumb, N.C., Yuan, K., 2012. Introduction. In: *Improving School Leadership Through Support, Evaluation, and Incentives*. RAND Corporation, pp. 1–9. https://www.jstor.org/stable/10.7249/mg1223pps.9?seq=1#metadata_info_tab_contents.
- Hill, H.C., Schilling, S.G., Ball, D.L., 2004. Developing measures of teachers' mathematics knowledge for teaching. *Elem. Sch. J.* 105 (1), 11–30. <https://doi.org/10.1086/428763>.
- IDEICE, 2017. Evaluación Del Desempeño 2017: Informe de Resultados de Otros Cargos Docentes [Performance Assessment 2017: A Report of Results of Other Teaching Posts]. Instituto Dominicano de Evaluación e Investigación de la Calidad Educativa. <https://ideice.gob.do/programas/evaluacion-del-desempeno-docente>.
- Isoré, M., 2009. Teacher Evaluation: Current Practices in OECD Countries and a Literature Review, vol. 23. Education Working Papers. <https://doi.org/10.1787/223283631428>.
- Kane, T., Douglas, O., 2012. Measures of Effective Teaching (MET). Gathering Feedback for Teaching: Combining High-Quality Observations With Student Surveys and Achievement Gains. Seattle, WA.
- Kane, T., Kerr, K., Pianta, R., 2014. In: Kane, T., Kerr, K., Pianta, R. (Eds.), *Designing Teacher Evaluation Systems: New Guidance From the Measures of Effective Teaching Project*. Bill & Melinda Gates foundation | Jossey-Bass A Wiley Brand, San Francisco, CA.
- Kyriakides, L., Bert, P., Creemers, M., Panayiotou, A., 2018. Using educational effectiveness research to promote quality of teaching: the contribution of the dynamic model. *ZDM* 50, 381–393. <https://doi.org/10.1007/s11858-018-0919-3>.
- Laing, D., Rivkin, S.G., Schiman, J.C., Ward, J., 2016. Decentralized Governance and the Quality of School Leadership. National Bureau of Economic Research. NBER Working Paper No. 22061.
- Leithwood, K., Jantzi, D., 2005. A review of transformational school leadership research 1996–2005. *Leadersh. Policy Sch.* 4 (3), 177–199. <https://doi.org/10.1080/15700760500244769>.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leadersh. Manag.* 40 (1), 5–22. <https://doi.org/10.1080/13632434.2019.1596077>.
- Leithwood, K., 2001. School leadership in the context of accountability policies. *Int. J. Leadersh. Educ.* 4 (3), 217–235. <https://doi.org/10.1080/13603120110057082>.
- Liu, S., Xu, X., Grant, L., James, S., Zheng, F., 2017. Professional standards and performance evaluation for principals in China: a policy analysis of the development of principal standards. *Educ. Manag. Adm. Leadersh.* 45 (2), 238–259. <https://doi.org/10.1177/1741143215587304>.
- Loeb, S., Grissom, J.A., 2013. What Do We Know About the Use of Value-Added Measures for Principal Evaluation? papers3. <http://www.publication/uuid/ADC51163-2B86-474C-A6BA-343C6E0E8AB5>.
- Macpherson, R., 2009. The professionalisation of educational leadership: implications of recent international policy research in leadership development for Australasian education systems. *J. Educ. Leader. Pol. Pract.* 24 (1), 53–117. http://www.reynoldsmacpherson.ac.nz/publications/JELPP_VOL24_1_Macpherson.pdf.
- Martínez Rizo, F., 2016. La Evaluación de Docentes de Educación Básica. Una Revisión de La Experiencia Internacional [Teacher Evaluation in Basic Education. A Review of the International Experience]. INEE, Mexico.
- Martínez Rizo, F., Taut, S., Schaaf, K., 2016. Classroom observation for evaluating and improving teaching: an international perspective. *Stud. Educ. Eval.* 49, 15–29. <https://doi.org/10.1016/j.stueduc.2016.03.002>.
- Maturell Aguilera, A., Valiente Sandó, P., 2016. Improvement of Principals' Performance in Comprehensive Schools: Its Historical Evolution in Cuba from 1899 to 2015, 25. *Avances En Supervisión Educativa* <https://doi.org/ISSN:1885-0286>.

- Mayer, D., Mills, M., 2020. Professionalism and teacher education in Australia and England. *Eur. J. Teach. Educ.* 1–17. <https://doi.org/10.1080/02619768.2020.1832987>.
- MINEDUC, 2015. Marco Para La Buena Dirección y El Liderazgo Escolar [Framework for Good Management and School Leadership]. Ministry of Education, Chile. <https://doi.org/10.1109/URSIGASS.2011.6050703>.
- MINEDUC, 2020. Actualización Estándares Indicativos de Desempeño Para Los Establecimientos Educacionales y Sus Sostenedores [Update of Performance Standards for Institutions and Funders]. Ministry of Education, Chile. <http://library1.nida.ac.th/termpaper6/sd/2554/19755.pdf>.
- MINERD, 2014a. Estándares de Desempeño de Docentes, Directores/as y Maestros/as de Centros Educativos [Performance Standards of Teachers, Principals, and Educational Centres]. Ministry of Education, Dominican Republic, Santo Domingo. <https://ideice.gob.do/documentacion/publicaciones-msg-set-id-9-art-14-estandares-de-desempeno-de-docentes-directores-as-y-maestros-as-de-centros-educativos>.
- MINERD, 2014b. Estándares Profesionales y Del Desempeño Para La Certificación y Desarrollo de La Carrera Docente [Professional Standards and Performance for the Certification and Teaching Career]. Ministry of Education, Dominican Republic, Santo Domingo. <https://www.ministeriodeeducacion.gob.do/docs/viceministerio-de-acreditacion-y-certificacion-docente/g051-estandares-profesionales-y-del-desempeno-para-la-certificacion-y-desarrollo-de-la-carrera-docentepdf.pdf>.
- Ministerio de Educación, 2010. Instructivo de La Prueba de Conocimientos Específicos Para Directores [Manual for the Specific Knowledge Exam for Principals]. Ministry of Education, Republic of Ecuador. http://web.educacion.gob.ec/_upload/Directores.pdf.
- Ministerio de Educación, 2017. Manual Para La Implementación Y Evaluación De Los Estándares De Calidad Educativa [Manual for the Implementation and Evaluation of Quality Education Standards]. Ministry of Education, Republic of Ecuador. www.educacion.gob.ec.
- Ministry of Education, 2021. Professional Standards for Primary School Principals. Ministry of Education, New Zealand. <https://www.educationalleaders.govt.nz/Leadership-development/Professional-information/Professional-standards-primary-school-principals>.
- Molina, E., Pushparatnam, A., Rimm-Kaufman, S., Wong, K.K.-Y., 2018. Evidence-Based Teaching: Effective Teaching Practices in Primary School Classrooms. <https://doi.org/10.1596/1813-9450-8656>.
- Mortimore, P., Sammons, P., Thomas, S., Mortimore, P., Sammons, P.A.M., Thomas, S., 1994. School effectiveness and value added measures school effectiveness and value added measures. *Assess. Educ.: Princ. Policy Pract.* 1 (3), 315–332. <https://doi.org/10.1080/0969594940010307>.
- Muijs, D., Kyriakides, L., Werf, G. van der, Creemers, B., Timperley, H., Earl, L., 2014. State of the Art—teacher effectiveness and professional learning. *Sch. Effect. Sch. Improv.* 25 (2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>.
- National Governance Association, 2019. What Governing Boards and School Leaders Should Expect From Each Other. <https://www.nga.org.uk/Membership/Publications.aspx%0A>. (Accessed 14 February 2021).
- NCEE, 2021. Top Performing Countries Shanghai-China. <https://ncee.org/country/shanghai-china/>.
- Ni, Y., Rorrer, A.K., Pounder, D., Young, M., Korach, S., 2019. Leadership matters: preparation program quality and learning outcomes. *J. Educ. Adm.* 57 (2), 185–206. <https://doi.org/10.1108/JEA-05-2018-0093>.
- Nusche, D., Laveault, D., Macbeath, J., Santiago, P., 2012. OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011. OECD publishing. <https://doi.org/10.1787/9789264116917-en>.
- NZSTA, 2021. A Parents' Guide to the Role of the Board of Trustees. Wellington, New Zealand. <https://www.nzsta.org.nz/assets/Information-for-parents/A-parents-guide-to-the-role-of-the-board-of-trustees.compressed.pdf>.
- OECD, 2005. Teachers Matter: Attracting, Developing and Retaining Effective Teachers. OECD Publishing. <https://www.oecd.org/education/school/34990905.pdf>.
- OECD, 2013. The appraisal of school leaders: fostering pedagogical leadership in schools. In: *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. OECD Publishing, Paris, pp. 485–582. <https://doi.org/10.1787/9789264190658-11-en>.
- OECD, 2019. TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. OECD publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>.
- OECD-CEPPE, 2013. Learning Standards, Teaching Standards and Standards for School Principals: A Comparative Study, vol. 99. EDU/WKP.
- OEI, 2017. Miradas Sobre La Educación En Iberoamérica. Desarrollo Profesional y Liderazgo de Directores Escolares En Iberoamérica [Overview of Education in Ibero-America. Development of Professional Leadership of Principals]. Organization of Ibero-American States. <http://www.oei.es/uploads/files/microsites/1/48/informe-miradas-2017.pdf>.
- Oviedo Bogantes, C.A., 2017. Informe de Evaluación Del Gestor de Centros Educativos Públicos de Costa Rica [Report on the Evaluation of Manager of Public Education Centres of Costa Rica]. Ministry of Public Education, Costa Rica. https://dgec.mep.go.cr/sites/all/files/dgec_mep_go_cr/documentos/informe_final_pedce-cr-2018.pdf.
- Polikoff, M.S., Porter, A.C., 2014. Instructional alignment as a measure of teaching quality. *Educ. Eval. Pol. Anal.* 36 (4), 399–416. <https://doi.org/10.3102/0162373714531851>.
- Priestley, M., Biesta, G., Philippou, S., Robinson, S., 2015. The teacher and the curriculum: exploring teacher agency. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. SAGE Publications Ltd., pp. 187–201. <https://doi.org/10.4135/9781473921405.n12>.
- REB, 2020. Professional Standards for Effective School Leadership in Rwanda. Rwandan Education Board, Kigali. https://rwanda.vwb.org/sites/rwanda/files/professional_standards_for_effective_school_leadership_in_rwanda_v0.0_20201201.pdf.
- Reynolds, D., Chapman, C., Kelly, A., Muijs, D., Sammons, P., 2012. Educational effectiveness: the development of the discipline, the critiques, the defence, and the present debate. *Effect Educ.* 3 (2), 1–19. <https://doi.org/10.1080/19415532.2011.686168>.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., Stringfield, S., 2014. Educational effectiveness research (EER): a state-of-the-art review. *Sch. Effect. Sch. Improv.* 25 (2), 197–230. <https://doi.org/10.1080/09243453.2014.885450>.
- Romero, C., Krichesky, G., 2019. El director Escolar En Argentina: Un actor Clave Pero (Aún) Invisible. Un Estudio Sobre Las Normas, Condiciones Laborales y Formación de Los Directores Escolares de Escuelas Públicas. *Educ. Pol. Anal. Arch.* 27 (12). <https://doi.org/10.14507/epaa.27.3576>.
- Saeed, T., 2020. Reimagining Education: Student Movements and the Possibility of a Critical Pedagogy and Feminist Praxis. Paper Commissioned for the UNESCO Futures of Education Report. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000374157.locale=fr>.
- Scheerens, J., 2015. Theories on educational effectiveness and ineffectiveness. *Sch. Effect. Sch. Improv.* 26 (1), 10–31. <https://doi.org/10.1080/09243453.2013.858754>.
- Sriprakash, A., Nally, D., Myers, K., Ramos-pinto, P., 2020. Learning With the Past: Racism, Education and Reparative Futures. Paper Commissioned for the UNESCO Futures of Education Report. UNESCO, pp. 1–13. <https://unesdoc.unesco.org/ark:/48223/pf0000374045>.
- Steinberg, M.P., Sartain, L., 2015. Does teacher evaluation improve school performance? Experimental evidence from Chicago's excellence in teaching project. *Educ. Finance Pol.* 10 (4), 535–572. https://doi.org/10.1162/EDFP_a_00173.
- The General Teaching Council for Scotland, 2012. The Standards for Leadership and Management: Supporting Leadership and Management Development. <http://www.gtcs.org.uk/web/files/the-standards/standards-for-leadership-and-management-1212.pdf>.
- Thomas, S.M., 2020. School and teacher value added performance and the relationship with teacher professional development in mainland China. In: Hall, S., Lindorff, A. (Eds.), *International Perspectives in Educational Effectiveness Research*. Springer.
- USICAMM, 2020. Marco Para La Excelencia En La Enseñanza y La Gestión Escolar En La Educación Básica: Perfiles Profesionales, Criterios e Indicadores Para El Personal Docente, Técnico Docente, de Asesoría Técnica Pedagógica, Directivo y de Supervisión Escolar. Secretariat of Public Education, Mexico. http://creson.edu.mx/docs/Perfiles_profesionales_Criterios_e_Indicadores_para_Docentes_Ciclo_Escolar_2021-2022.pdf.
- Vaillant, D., 2019. Experiencias Innovadoras En El Desarrollo Profesional de Directivos. Latin-America Development Bank. scioteca.caf.com.
- Villegas-Reimers, E., 2003. Teacher Professional Development: An International Review of the Literature. UNESCO: International Institute for Educational Planning, pp. 1–197. <http://unesdoc.unesco.org/images/0013/001330/133010e.pdf>.
- VVOB, 2019. School Leadership: Concept Note. The Flemish Association for Development Cooperation and Technical Assistance (VVOB), Kigali. https://rwanda.vwb.org/sites/rwanda/files/updated_promoting_effective_sl_v4.0en_20190814.pdf.
- Webb, P.T., 2006. The choreography of accountability. *J. Educ. Pol.* 21 (2), 201–214. <https://doi.org/10.1080/02680930500500450>.

- Wildy, H., Pepper, C., Guanzhong, L., 2011. Applying standards for leaders to the selection of secondary school principals. *J. Educ. Adm.* 49 (3), 276–291. <https://doi.org/10.1108/09578231111129064>.
- Zheng, H., 2020. Stakeholder perceptions on the role of school inspection standards in demonstrating education quality in China. *Qual. Assur. Educ.* 28 (2), 105–121. <https://doi.org/10.1108/QAE-09-2019-0093>. This page intentionally left blank

School leadership in Latin America: research and policy trends

Joseph Flessa^a, Daniela Bramwell^b, and José Weinstein^c, ^a Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada; ^b Toronto, Canada; and ^c Diego Portales University, Santiago, Chile

© 2023 Elsevier Ltd. All rights reserved.

Overview of the field of study and its relevance to the region	469
Research trends	470
Policy trends	471
Are principals professionals? Are principals educators?	471
How are principals prepared for the job?	472
Leadership frameworks: international benchmarking at the expense of local context?	472
Policy drivers and areas for future research	473
Policy drivers	473
Areas for future research	474
References	475

The purpose of this chapter is to provide an overview of school leadership research and policy trends in Latin America. The storyline of our chapter is as follows: globally there is a shared understanding that leadership matters for school improvement, and school leadership development is an important system responsibility. However, outside the Global North the school leadership research base is nascent and uneven. This means that although much school leadership research concludes that context matters, in many regions of the Global South international/borrowed policies and leadership frameworks dominate. In this paper we highlight the following tension in Latin America specifically: there are two trends that are true simultaneously, namely that the knowledge base is thin and uneven across countries *and/but* there is growing interest in school leadership among researchers and policymakers in who leaders are, what they do and with what impact, and how systems recruit, support, and evaluate them.

By reviewing trends in research and policy on school leadership in Latin America this chapter provides readers with a snapshot of the current knowledge base as well as a set of key questions to help explain key conceptual and empirical gaps. This chapter is based on our work originally published in a first-of-its-kind Special Issue on Latin America of *Educational Management Administration and Leadership* (Flessa et al., 2018). That review covered 2000–2016. As we prepared this chapter, we returned to the scholarly databases to see if our original conclusions would hold in the face of new research published between 2016 and 2021. We believe that they do.

Overview of the field of study and its relevance to the region

The study of school administration and management is not new. For example, administrative progressives in the United States, responding to expanded public schooling in the early 20th century, sought to apply the lessons of industrial efficiency to the management of schools and districts. The field of educational administration, focusing on the specialized managerial work that principals as distinct from teachers do, was born. From the start an anchoring conviction was that principals matter both for teachers' work and student achievement, and that the principalship is a profession with special knowledge and skills. Ellwood Cubberley, for example, put it this way almost a hundred years ago:

We are not likely to overestimate the importance of the office of school principal ... The knowledge, insight, tact, skill, and qualities of helpful professional leadership of the principal of the school practically determine the ideals and standards of achievement of both teachers and pupils within the school.

Cubberley (1923, pp. 349–50).

In other words, the current conviction among educators and policymakers that principals matter has a long lineage, and attention to school level leadership, especially in the United States and UK, has been undergoing a renaissance since the late 20th century as both governments (e.g., the National College for School Leadership in the UK) and foundations (e.g., the Wallace Foundation in the US) invested hugely in leadership research and development. For current readers, Cubberley's century-old assertion likely sounds much like the conclusion of Leithwood and colleagues from earlier this century: "School leadership is second only to classroom teaching as an influence on pupil learning" (Leithwood et al., 2008, p. 27). (Or, the "revisited" version: "School leadership has a significant effect on features of the school organization which positively influences the quality of teaching and learning. While moderate in size, this leadership effect is vital to the success of most school improvement efforts" (Leithwood et al., 2020, p. 6)).

Grissom et al. (2021) also "revisit, question, and extend these earlier conclusions to provide new direction for practice, policy, and research in school leadership" (p. xi) and find that "the overall impact of an effective principal can be linked to observable

behaviors. That is, *how* principals approach school leadership directly affects schools' outcomes" (p. xv). Grissom et al. (2021) have provided the most recent rigorous synthesis of 4800 documents on school leadership (in English). Their conclusions are worth quoting in full because they both reaffirm the importance of school leadership as well as redefine the knowledge base on school leadership, mostly from the Global North (p. xviii):

1. Effective principals are at least as important for student achievement as previous reports have concluded—and in fact, their importance may not have been stated strongly enough.
2. Principals have substantively important effects that extend beyond student achievement.
3. Effective principals orient their practice toward instructionally focused interactions with teachers, building a productive school climate, facilitating collaboration and professional learning communities, and strategic personnel and resource management processes.
4. Principals must develop an equity lens, particularly as they are called on to meet the needs of growing numbers of marginalized students.
5. Effective principals are not equitably distributed across schools.
6. Principals are becoming more racially and ethnically diverse, but representation gaps with students are growing, which is concerning, given the payoffs to principal diversity.
7. Research on school principals is highly variable, and the field requires new investment in a rigorous, cohesive body of research.

In short, although there is always more to know and there is always an imperative to strengthen the knowledge base, there is little debate in the English language research literature about the importance of school leadership.

But what of Latin America? As Weinstein (2021) has argued elsewhere, "very little is still known about the people that take on the leadership roles, and even less about how they develop their leadership practices, what their educational and professional motivations are, or how their efforts translate into school improvement" (p. 138). Hallinger (2020) has noted that only 5% of articles in key English language educational administration journals since 1965 have addressed Latin America; indeed, a research scan dating to 1965 showed no publications on school leadership until after 1990.

Our contribution to the knowledge base on school leadership in Latin American was to review almost 400 policy documents and research articles to understand the state of the field. The bilingual English-Spanish research team reviewed materials in both English and Spanish and, for Brazil, in Portuguese. Our conclusion, which we believe still stands, is as follows:

First, we noted that Latin America has come to school leadership as an area of research and policy focus on recently; second, a small set of Latin American countries, led by Chile, accounts for the majority of the school leadership publications in this century, meaning the published record of leadership for most countries is very thin; and, third, in the few places where there are policy efforts to emphasize leadership, their points of reference appear to be Anglo-American leadership concepts and researchers (p. 193).

In the sections that follow we first summarize the trends in educational leadership research that we uncovered in our 2018 review. After that we discuss educational leadership policy trends in the region by foregrounding the questions that the policies seek to answer. We conclude with speculation on the drivers behind the growth, however uneven, of interest in school leadership in Latin America and point to promising areas for development of future research.

Research trends

Our comprehensive review of the school leadership literature observed several trends in the research. In the sections that follow we state those trends and speculate on their meaning.

1. We found that school management research is slightly more common than school leadership research.

As Gunter (2004) and others have noted, there has been a movement away from locating the work of principals under the labels of administration/management, toward labeling the work of school principals as leadership. (As a small illustration—the home program for the lead author is now called "Educational Leadership and Policy" not "Educational Administration.") But it is certainly possible that, given the overall newness of school level leadership research in Latin America, that this movement away from administration and toward leadership is only just underway. It is also possible that the emphasis on management is an accurate reflection of prevailing ideas about school level decision making in Latin America.

2. References to school leaders and management arise primarily within the context of school change and reform.

There are two interesting aspects to this finding. First, given that the work of school level leaders, particularly principals, is important for achieving a wide range of policy priorities because principals influence the conditions of under which teachers implement (or not) policy change and reform, it stands to reason that leadership research would center reform/s. Second, the emphasis on leaders-in-a-context-of-reform suggest, perhaps worrisomely, that leaders and leadership remain an under-examined black box in their own right.

3. Studies that focus directly on leadership are often case studies of leadership characteristics, practices and problems encountered by principals and focus on leadership frameworks such as “transformational” or “instructional” leadership. Leadership teams and forms of distributed leadership are minimally examined.

Research on distributed leadership has undoubtedly been one of the most influential bodies of research in the English language school leadership literature for decades, prompting special issues of key journals, edited volumes, and texts for both academic and professional audiences, and the interest continues (see e.g., [Supovitz et al., 2019](#)). The concept is not similarly evident in the literature we examined from Latin America. We found this absence interesting for several reasons. First, in many school systems, school level leadership falls to “equipos directivos”—broadly, “leadership teams.” These teams can include principals, other administrators responsible for student conduct/discipline or administrators responsible for what in English would probably be considered “school climate” but more accurately translates to “getting-alongness.” In Chile schools have the interesting figure of the “jefe pedagógico” or pedagogical chief, an administrator responsible for working with teachers on curriculum, teaching, and learning; there is a similar figure in Brazil called the “pedagogical coordinator.” These teams are formally defined in some countries (e.g., Chile, Ecuador, Argentina) and undefined in others (e.g., Peru and Colombia). In short, although it is next to impossible to generalize about the region, there are numerous on-the-ground examples of shared or distributed leadership in many schools in Latin America. The fact that analysis examining the range of how and with what impact school level leadership is shared is in short supply perhaps derives from the simultaneous reality that policy and legislation are still fundamentally organized around a vision of the solitary school leader sitting atop the school hierarchy.

4. Latin America is diverse in terms of research production.

There are prolific countries like Chile, Brazil, Mexico (with all three participating in OECD’s TALIS study). There is less published research from Argentina, Costa Rica, Cuba, Colombia or Venezuela. Although we initially found very little from Ecuador, Bolivia, Panama, when we revisited this question to prepare this chapter we encountered several theses on “*gestión educativa*” (which roughly translated to “educational administration”). Although this work has not yet made its way into the peer reviewed or gray literatures it is possible that more work on school leadership—for example from Ecuador—will emerge from these theses.

Importantly, most of the research we encountered were small-scale studies. With the exception of Chile, with its well-established mechanisms for collection of educational statistics, we encountered very little work showing baseline demographic data (age, gender, educational qualifications) for school principals, little work linking school-level leadership to system-level priorities, little systematic research on leadership pathways. In short, even where research on school level leadership exists, we are left with gaps in our understanding of how leadership work at the school level reflects (or doesn’t) the priorities of educational systems in this populous and diverse region.

Another complementary way of viewing this issue is with respect to international tests that, although their focus is not on leadership, also collect some information on principals, as is the case of TALIS (the “Teaching and Learning International Survey” from UNESCO) or LLECE (“El Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación,” part of the “Oficina Regional de Educación para América Latina y el Caribe”, UNESCO). In some of these tests very few countries in the region participate (such as the case of TALIS) and in others (such as the case of LLECE) the information on leadership is barely analyzed from a research perspective. Meaning: regional cross-national data are both limited and, where they exist at all, under analyzed.

Policy trends

Are principals professionals? Are principals educators?

[Murillo and Román \(2013\)](#) found that about 40% of principals in their 17-country study had a second job in a school. This striking finding is potentially indicative of an important open question in Latin America about whether the principalship is a *profession*, with specialized and explicit preparation, credentialing, recruitment, and selection, and unique responsibilities, or whether the principalship is only an occupation, such as in many rural schools where teachers maybe required to take on additional mainly administrative tasks outside their classroom teaching. In our 2018 paper we concluded that “We can safely conclude that the degree to which principalship has been professionalized with specific managerial and instructional responsibilities and expertise, and standards for selection into the profession, varies greatly in Latin America” (p. 192).

Another open question is the degree to which education systems in Latin America, which range from the highly centralized (for example Cuba) to the highly localized (for example, Argentina) to the highly privatized (for example Chile), share any common perspective on what the role of the school level leader should be. Is the principal’s role best understood as administering the policy priorities of a central government ministry or as entrepreneurially solving problems according to local needs or competition for enrollment? All of the countries we examined for our 2018 paper have school leadership roles (the principalship) explicitly defined in law and regulation; those regulations typically stipulate managerial rather than pedagogical responsibilities.

Countries across the region have answered a core question about school leaders’ identity—is a school principal an educator?—differently. Most but not all countries require classroom teaching experience, and the corresponding required educational background, to become a principal. But this requirement is by no means universal. As with other parts of the Global South school principals are sometimes chosen by patronage (e.g., by a political head of local government with responsibility for schooling) where relationship to a politician or political party is more important than other qualifications (see for example [Adelman, Lemos and Teodorovicz, 2019](#); [Donoso-Díaz et al., 2019](#)). On the other hand, sometimes school principals with no background in education

are chosen via transparent civil-service-style competition, perhaps linked to ideas that schools are businesses and “strong leadership” is therefore more likely to be found outside the education sector than within.

Similarly, the processes for hiring school leaders vary greatly across Latin America. For example,

Argentina, Colombia, Ecuador, Mexico and Peru use a combination of merit-based competitions and/or examinations. Costa Rica and Chile are attempting to increase the standardization of processes by using merit-based selection assessed by central government agencies. Given the limited autonomy that principals in Latin America seem to have, we were surprised to find great diversity in the degree to which their hiring was managed centrally or are decentralized. (We expected to see mostly centralized hiring.) In Argentina local provinces regulate the eligibility process and there is no role for central government (an illustration of how many decisions in Argentina's federal system are made at the provincial level); in Chile and Peru there is a shared governance approach where the local level has the final decision; in Ecuador a school committee, made up of parents, students and teachers, participates in assessing the school management plan presented by the candidate; and in Costa Rica, Colombia and Mexico there is a fully centralized process whereby candidates are scored by judges to decide about the final appointment.

Flessa et al. (2018, pp. 191–192).

The contract for successful candidates for the school directorship varies in the region. For example, in Argentina and Colombia the number of years for a principal contract are undefined/indefinite; in Chile the contract is for five years (renewable); in Ecuador and Peru the contract is for four (renewable).

How are principals prepared for the job?

Whereas policy and research in the US, which dominates the English-language literature base, focuses a great deal on the *preparation* of school leaders, often via specialized, university-based credential programs, the policy infrastructure in Latin America is sufficiently different that to talk about “preparation” makes less sense than to talk about *formación*. To provide concrete examples, Argentina and Ecuador require some form of preservice preparation but Chile, Colombia, and Peru do not. Meaning: there is neither a policy or research base focused exclusively on preservice preparation, and, across the region the better way to understand the training that principals receive is to think about in-service professional development. There is certainly no lack of professional learning for school leaders across the region. Weinstein et al. (2018) have discouragingly found, however, that such learning opportunities for principals have little *impact* on how teachers view principals’ contribution to the professional development and collaborative work of teachers.

One way to explain this perhaps counterintuitive finding is to point out that governments have typically not invested in systematic understanding of the leadership pathway or where preparation or professional development would fit in that. And although we saw some evidence of public professional development, most in-service learning is self-funded by the leadership aspirant for programs in the private sector where oversight, alignment to state policy priorities, and quality control are minimal.

Growing interest in credentials raises the question of where knowledge resides to do this professional preparation and development well. Is it in systems, in ministries of education, in university faculties of education, in professional associations or teachers’ unions? If policy in the region were to move in the direction of more explicit requirements for leadership preparation before taking on the role of principalship, just who should provide this preparation, and how would policymakers know that this was a wise investment?

Another fundamental question that underlies any discussion of *pathways for leadership development* must be: who are school leaders? Frequently, “school leader” is used interchangeably with “principal.” But a comprehensive vision of leadership development, including the articulation of a leadership pathway, would need to consider that much of the work of school leadership is also done by teachers (“middle leaders” in roles like department chairs, see Gurr (2019)) as well as management *teams* in schools. If policymakers want investments in school level leadership to show results, such investments would need to recognize who in fact is doing the work of school leading—and that would include more than principal alone.

Finally, the articulation of a coherent leadership pipeline, including preparation and professional development expectations, must also identify—and then rectify—the obstacles, whether intentional or not, that inequitably, often on the basis of gender, ethnicity, or sexuality, exclude leadership aspirants from the role. Research that provides baseline data of school leaders’ demographic characteristics as a way of identifying systematic patterns of exclusion is a necessary but frequently missing in the region.

Leadership frameworks: international benchmarking at the expense of local context?

There is also a further question about preparation that policymakers must contend with: credentialing and preparation initiatives are preparing school leaders *to do what*? At the same time that expectations of the job are changing, so too must the expectations of how preparation programs should be organized and oriented. One possible anchor for preparation programs would be the development of leadership standards or frameworks. Chile has adopted a leadership framework (“El Marco para la Buena Dirección y el Liderazgo Escolar”) with many parallels to the Ontario (Canada) Leadership Framework (OLF) (Leithwood, 2012). The Chilean framework focuses on building a shared strategic vision, developing professional capacities, leading the teaching and learning process, building school climate and community, and effective school management to meet institutional goals. The Ontario framework uses similar categories of leadership work: setting goals, aligning resources with priorities, promoting collaborative learning

cultures, using data, and engaging in courageous conversations. The OLF and the Chilean framework have influenced the development of similar frameworks in other countries in Latin America, such as Peru and Ecuador. That said, leadership standards and frameworks are not universal; Argentina, for example, has none. In general, frameworks emphasize leadership for school improvement and pedagogical leadership. Frameworks have positive potential for creating common language in the education sector about what leadership is and what leaders do; standards express a system's priorities to which leadership preparation, recruitment, selection, and evaluation can be benchmarked. Frameworks and standards represent attempts to organize public policies, and this organization can be especially valuable to systems seeking coherence in the early phases of investment in school leadership.

That said, the body of research to illustrate that frameworks change leaders' behavior in Latin America is thin, and the question remains open whether frameworks sufficiently consider local context. In our research and policy scan of school leadership in Latin America we concluded, much like [Moorosi and Bush \(2019\)](#) observed in their work on school leadership development in Africa, that although there is near universal agreement that context matters for school leadership, imported frameworks for understanding, researching, developing, and assessing leadership dominate the field. [Hallinger \(2018\)](#) has written persuasively about this problem, arguing that "research syntheses which report what successful leaders do have tended to strip away the 'context specificity' needed to help practitioners understand how to apply findings in different schools. Metaphorically, by placing the spotlight on leadership, researchers have unwittingly relegated the context of leadership to the shadows" (p. 6). Interestingly, Hallinger discusses school contexts—plural—as influencing the work of school leaders: institutional, community, socio-cultural, political, economic, and school improvement. All of these contexts vary across this diverse region and so too should research about leadership. However, the prevailing trends in the literature are, first, that there is limited research and second, that the research that does exist relies on imported frameworks, suggesting that context is neglected in research analyses and recommendations.

This is not a generalized critique of frameworks but instead a doubt about their application in Latin America. The question of context is persistent and also inevitable. One reason we have found [Hallinger's \(2018\)](#) definition of *various* contexts to be useful is because it suggests that any authors (like us!) making arguments about "context-informed" leadership should use the word context more precisely. There's not one context, there are many—and deciding how frameworks should reflect them requires trade-offs between scope and specificity.

One concrete illustration of a realm of leadership overlooked when importing frameworks to Latin America is the leadership obligation in many competition/market-based education systems to recruit students. Student enrollments determine school budgets. This skill set is absolutely necessary for survival. [Carmen Montecinos et al. \(2015\)](#) and [Galdames et al. \(2018\)](#) have used "triple accountability" analyses to study this kind of competitive/market phenomenon but we're unaware of any frameworks or standards in Latin America that include this realm of work. For school leaders in some Latin American school systems this is the most pressing, indeed most organizationally existential, leadership obligation. It falls outside of any leadership framework we're aware of but clearly matters for defining the scope and practice of school leadership in Latin America.

We return to the question of context in our recommendations for future research.

Policy drivers and areas for future research

Given the still nascent nature of school leadership research and policy in the region, plus the vast diversity of the region's education systems and schools, uncovering a single underlying explanation for research and policy trends is not possible. Instead, we "tested" a range of alternative explanations in an attempt to better understand the policy drivers and where they come from.

Policy drivers

1. Are differences in the region explained by variations in centralization and decentralization?

School systems in the region vary in their degree of centralization; policy priorities in different countries take different approaches to how tightly coupled school and educator actions must be to rules and directions set within ministries of education. In examining our research and policy materials we wondered if greater centralization meant that school level leaders were less important, perhaps because a centralization logic would frame school leaders primarily as bureaucrats responsible for administering centralized policies. Alternately, we wondered if decentralized systems, where policy goals may be set centrally but the pathway to reaching them is left to the local level, might see greater investment in school level leaders who could guide schools' actions in response to local contexts. We found neither link in our data. Some more decentralized countries did *not* seem to focus on leadership, and others more centralized did. So, although we are confident that the degree of system centralization does affect principals' working conditions and responsibilities, it does not seem to be a primary driver of research or policy.

2. How powerful is policy borrowing in shaping the direction of school leadership development in the region?

We saw plenty of evidence of policy adoption in the region and plenty of work drawing from the same English language UK/US resources. We are unable from our review of the literature to determine what *motivated* this policy borrowing—institutional isomorphism, harmonization, response to incentives from global funding agencies, benchmarking. However, while policy adoption is clearly a phenomenon worth greater investigation, we do not conclude that it is the sole explanation for leadership development trends in the region. Other factors clearly play a role: local school and system needs; resources available for investment in this area; system coherence and strength; capacity to make change.

3. Does the amount of leadership investment in a particular country in the region somehow reflect the country's infrastructure or "developmental" readiness for leadership? Does leadership come after other priorities (e.g., universal primary education, investment in hard infrastructure like buildings or textbooks)?

The evidence from our review does not point in this direction. As we concluded in 2018, "we believe that a full explanation of the links between reform policies, school leadership, and school improvement in Latin America requires more research" (Flessa et al., 2018, p. 196).

4. Is it possible the investments in leadership are explainable as a form of technocratic problem-solving?

If one views national policy making as rational, one might conclude that policy makers see leadership as one of the best ways to solve problems such as improving education quality that have not yet been solved. There is research showing that school leadership can improve education quality, so perhaps this is the reason—or at least an important part of the reason.

5. Does the embrace of school leadership in the region mostly reflect a prevailing neoliberal orientation to policymaking, where accountability and resource scarcity are both reinforced by a focus on principals rather than, say, teachers?

We saw evidence of the link between school leadership and school accountability in materials from Chile and Brazil; principals are seen as one important lever for school level improvement aligned with state policy goals. However, there is more to neoliberalism than accountability policy; for example, one would expect to see budgetary austerity as well. And while Chile has one of the most extensive testing and school accountability regimes in the world as well as an overall privatization approach to schooling, it has also *increased* not decreased public investment in public schooling in recent years. Investments in school leadership may indeed reflect the impact of neoliberal logics among policymakers, but the literature thus far does not provide sufficient evidence for that conclusion.

A related point: important actors in the region are private foundations linked to large companies or economic groups. These actors often have great confidence in the ability of an organization's senior managers (such as CEOs) to motivate and lead changes and affect results. There is a simile here between the company and the school. In fact, business managers themselves feel that they have a lot to teach school managers in terms of leadership and management. A clear example of this impact on managers of this business aspect are the training programs that have been developed by business foundations, as has happened in Colombia (Rectores Líderes Transformadores promoted by Fundación Empresarios por la Educación).

Areas for future research

Our recommendations for future research reflect two emphases, namely to deepen the knowledge base in some areas and to initiate research in areas where we found little or none.

As we have noted, there is a nascent knowledge base concerned with what broadly could be considered leadership pipelines: the machinery of recruitment, selection, preparation and professional development, and evaluation of school leaders. This work can certainly be deepened by work that compares the various approaches—and their impacts—in the region. Such work will also require baseline data of who school leaders are and how they are prepared for the job.

There are several areas where there is minimal research, and these are true gaps in the knowledge base on school leadership in Latin America.

First, we found surprisingly little school leadership research focused on questions of equity of learning opportunity, and no work focused on leadership in or for Indigenous communities. This seems a striking oversight, given the realities of stark inequalities across the region, inequalities that have been highlighted by the impact of COVID-19 and the historic out-migration from Venezuela in recent years.

Second, research examining how school leaders work with teachers' unions and how school leaders navigate social and labor protests in the public sector are needed. Such research would not only be timely it would also provide a greater understanding of where school leadership fits within different political contexts. When the role of the state in education is contested, how do school leaders understand and enact their roles?

Third, we believe that there is space to develop a research program that examines how concepts prevalent in the English-language research literature—for example, "instructional leadership"—emerge (or don't) in Latin American schools, and why. In what ways has the political history of the region informed the way that concepts like "strong leadership" are understood? Such research would not only provide insights about the applicability of the concepts to schooling in Latin America, they would also provide insights into the concepts themselves and their limitations for English language and other audiences.

Fourth, there is space to develop research on leadership teams and leadership positions beyond the principalship, given the evidence of various school leadership roles and teams utilized in Latin American schools. This would build upon but not be the same as research into distributed leadership prevalent in the English-speaking EDLM literature.

Fifth, the topic of leadership and management in private schools and their differences with public ones is also important. The privatization of education is no longer just a Chilean issue but is advancing strongly in other countries or regions (such as Peru, Buenos Aires, etc.) and promises to be a major trend. To date however, educational policies and research tend to focus on public establishments. Therefore, considering the school system as a whole and not just one part is a requirement for future research.

Finally, we believe future research on the topic of school leaders' "sensemaking" would make valuable contributions to the knowledge base. It would contribute to both policy and practice by highlighting how school leaders, especially "*directores*" (principals) navigate the potentially contradictory incentives of the job. There is a baseline comparative education question here: which is

more significant, that principals in different systems understand their jobs differently, or that they understand them in similar ways? Is a principal a principal regardless of system, or does the system within which they work make them very different educational figures? Research on leaders' sensemaking in Latin American schools may in fact be the key body of work necessary to synthesize what is known about the particularities of local context and to apply that knowledge to broader policies like leadership frameworks. Taken together these different lines of research would build a rich knowledge base on educational leadership in Latin America and, within the right policy window, inform the development of the next generation of school leaders in the region.

In conclusion, in the published research and policy literature there is evidence of a growing interest in school leadership in Latin America. This nascent knowledge base provides valuable insights into the potential and the pitfalls of leadership development in the region and also indicates for a global scholarly audience which concepts have the most salience and which are mismatched to local context.

References

- Adelman, M., Lemos, R., Teodorovic, T., 2019. Appointing versus Locally Electing School Directors: Evidence from Brazil. Mimeo.
- Cubberley, E.P., 1923. The principal and the principalship. *Elem. Sch. J.* 23 (5), 342–352.
- Donoso-Díaz, S., Fernández-Negrete, J.P., Araya, D.R., 2019. Directores de escuelas públicas chilenas seleccionados por alta dirección removidos de su cargo antes de finalizar su período de gestión: Lecciones para América Latina. *Education Policy Analysis Archives*, p. 27.
- Flessa, J., Bramwell, D., Fernandez, M., Weinstein, J., 2018. School leadership in Latin America 2000–2016. *Educ. Manag. Adm. Leader* 46 (2), 182–206.
- Galdames, S., Montecinos, C., Campos, F., Ahumada, L., Leiva, M.V., 2018. Novice principals in Chile mobilizing change for the first time: challenges and opportunities associated with a school's readiness for change. *Educ. Manag. Adm. Leader* 46 (2), 318–338.
- Grissom, J.A., Egalite, A.J., Lindsay, C.A., 2021. How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research. The Wallace Foundation, New York, NY. Available at: <http://www.wallacefoundation.org/principalsynthesis>.
- Gunter, H., 2004. Labels and labelling in the field of educational leadership. *Discourse* 25 (1), 21–41.
- Gurr, D., 2019. School middle leaders in Australia, Chile and Singapore. *Sch. Leader. Manag.* 39 (3–4), 278–296.
- Hallinger, P., 2018. Bringing context out of the shadows of leadership. *Educ. Manag. Adm. Leader* 46 (1), 5–24.
- Hallinger, P., 2020. Science mapping the knowledge base on educational leadership and management from the emerging regions of Asia, Africa and Latin America, 1965–2018. *Educ. Manag. Adm. Leader* 48 (2), 209–230.
- Leithwood, K., 2012. The Ontario Leadership Frameworks with a Discussion of its' Research Foundations. Institute for Educational Leadership & Ontario Ministry of Education, Toronto, ON.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28 (1), 27–42.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22.
- Montecinos, C., Ahumada, L., Galdames, S., Campos, F., Leiva, M.V., 2015. Targets, threats and (dis) trust: the managerial troika for public school principals in Chile. *Educ. Pol. Anal. Arch.* 23 (87), 87.
- Moorosi, P., Bush, T. (Eds.), 2019. *Preparation and Development of School Leaders in Africa*. Bloomsbury Publishing.
- Murillo, F.J., Roman, M., 2013. La distribución del tiempo de los directores de escuelas de Educación Primaria en América Latina y su incidencia en el desempeño de los estudiantes. *Rev. Educ.* 361, 141–170.
- Supovitz, J.A., D'Auria, J., Spillane, J.P., 2019. Meaningful & Sustainable School Improvement with Distributed Leadership. CPRE Research Reports. Retrieved from: https://repository.upenn.edu/cpre_researchreports/112.
- Weinstein, J., 2021. Commentary – system thinking, school leadership and complexities in Latin America. *J. Educ. Adm.* 59 (1), 138–143. <https://doi.org/10.1108/JEA-11-2020-0247>.
- Weinstein, J., Azar, A., Flessa, J., 2018. An ineffective preparation? The scarce effect in primary school principals' practices of school leadership preparation and training in seven countries in Latin America. *Educ. Manag. Adm. Leader* 46 (2), 226–257.

Leadership in non-western countries: China

Haiyan Qian and Allan Walker, Education University of Hong Kong, Tai Po, Hong Kong

© 2023 Elsevier Ltd. All rights reserved.

The uneven rise of China	476
The complexity of educational leadership in China	478
What shapes educational leadership in China?	479
Learning from China?	480
Acknowledgment	481
References	481

The past 10 years have witnessed the growth of increasingly diverse and sophisticated literature on educational leadership in non-Western societies, particularly in Asia (e.g., Hallinger and Bryant, 2013; Liu et al., 2016; Qian and Walker, 2013, 2020; Walker and Qian, 2012, 2018a). China has contributed the largest corpus of published research on educational leadership in Asia (Dimmock, 2020; Hallinger and Bryant, 2013). Despite the quantitative expansion of research, there remains a noticeable imbalance between research generated in diverse politico-socio-cultural settings and that produced in more established societies. Dimmock (2020) holds that this curtails authentic knowledge production and transfer. A cross-culturally valid knowledge base with global relevance is yet to reach maturity.

While preparing this article we, like the rest of the world, were confronted by an unprecedented crisis which carried with it not only widespread hardship and sadness but also new levels of mistrust between nation-states. The impact of COVID-19 expanded the already festering rift between China and the West. In the longer term, the tension between Chinese and Western values, as reflected in politics, business and even research present a major challenge to the social sciences as they strive to develop a coherent knowledge base that reflects the global diversity of policies and practices. Be it in a relatively minor way, this challenge unavoidably touches understanding of and research into education including, as is our interest, educational leadership. It is therefore timely to review the development of educational leadership in China and examine this within an international frame. It is through such work that divisions of understanding can be pragmatically addressed. By exploring what educational leadership in China looks like and, as importantly, why this is so, we attempt to address how locally interpreted narratives give force to universal categories and how universal categories give force to local narratives (Stambeck, 2003, p. 157).

The purpose of the chapter is to briefly review what recent research tells us about educational leadership in China and to analyze the major contextual factors that shape its enactment. By doing this, we comment on why there appears a discrepancy between evidence-based understanding of Chinese education leadership and the stereotypical images widely presented in Western discourse. Drawing on this discussion we provide several insights into educational leadership in China. We draw evidence from our research conducted over the past decade (e.g., Qian and Walker, 2013; Qian et al., 2016; Walker and Qian, 2015, 2018a), several international comparative reports (e.g., Jensen et al., 2016; Liang et al., 2016; Tucker, 2013) and the increasing corpus of empirical studies being produced by other Chinese researchers (e.g., Liu and Hallinger, 2020; Yin et al., 2013; Zheng et al., 2019). It should be noted from the outset that China is a huge and diverse country. As in all countries, there are significant differences between regions in terms of resource availability, socio-economic-status, ethnic composition, education development, curriculum structures, policy implementation and political dynamics, to name just some. These and other contextual features impact the shape and practice of school leadership at all levels. This makes it impractical to capture a comprehensive picture of either the shape or impact of school leadership across the country. We therefore aim, within space limitations, at a general level of understanding using the major literature produced over the last decade and that most pertinent to the purpose of the article.

The uneven rise of China

Over the past decade, China has been firmly positioned as a key “reference society” (Schriewer and Martinez, 2004) due to the international attention flowing in the wake of the outstanding performance of Shanghai students on the Program for International Student Assessment (PISA) 2009 (OECD, 2010). Given its debutant status, Shanghai’s PISA results were considered outstanding. It outperformed other countries in mathematics, science and reading by a clear margin (OECD, 2010). When Shanghai participated in PISA for the second time, in 2012, it again topped the league table in all three subjects (OECD, 2013). When more provinces joined PISA in 2015 and 2018, overall student performance remained around the top on the league tables. The success went beyond academic performance; students also tended to “perform well regardless of their own background or the schools they attend[ed]” (OECD, 2010, p. 9). For example, the 2009 PISA results revealed that 76% of disadvantaged students in the Shanghai sample were considered “resilient”; that is, despite being classified as “disadvantaged” – defined by placement in the bottom quartile in terms of the distribution of socioeconomic status (SES) – they scored in the top quartile of students from all countries from similar

backgrounds. In the foreword to the 2009 volumes analyzing the PISA results, Shanghai was mentioned by name as the education system “that [has] been able to secure strong and equitable learning outcomes” (OECD, 2010, p. 3).

The afterglow of the PISA success cast Shanghai as an international model of academic success. The newly-founded visibility triggered a mixture of international admiration and intense debate. Western societies such as the US, UK and Australia began “looking East” in an attempt to identify the “formulas” driving China’s academic success (Seller and Lingard, 2013). For example, various books and international comparative reports released over the past decade either focused singularly on Shanghai or studied China in tandem with other high-performing societies. These include to name a few, a US publication that told stories of successful education performers internationally, entitled *Surpassing Shanghai* (Tucker, 2013); the book *Learning from Shanghai* written by a Singaporean researcher (Tan, 2013); and a World Bank report called *How Shanghai Does It* (Liang et al., 2016). Moreover, in addition to synthesizing and spreading “Shanghai lessons” in publications, different initiatives were set in place to mirror their practices in schools and classrooms in Western societies. For example, the *England-Shanghai Mathematics Teacher Exchange Program* instituted in 2014 aimed to promote the Shanghai method of Maths teaching in UK schools (China Daily, 2019). With the support of Peabody College of Vanderbilt University and East China Normal University in Shanghai, the State of Tennessee introduced the Shanghai model of teacher peer observation and feedback into their schools, naming it the TPEG (Teacher Peer Excellence Group) (Cravens et al., 2017).

However, the quite remarkable PISA success must be viewed within the broader context of Chinese education, such success does not hold for the entirety of China - major cities such as Shanghai and Beijing are not indicative of the country as a whole. China is a labyrinth of contextual diversity which interacts in various ways to influence school success. Perhaps the starkest differences in educational opportunity and school success are between developed urban (e.g., Shanghai) and underdeveloped rural (e.g., Guizhou) settings. They are also apparent in the provision of education for ethnic minorities and for the children of migrant workers in large urban centers (Qian and Walker, 2015). Differences between urban and rural schools endure for a number of reasons. The most telling of these is access to adequate resources. As we pointed out in an earlier study, “Principals working in cities such as Beijing and Shanghai can harness many more resources than those from rural regions in West China” (Walker and Qian, 2015, p. 484). The lack of resources plays out in rural schools in many ways. These include constant difficulty in recruiting and retaining quality teachers, time required to make the right connections to secure resources, and a lack of adequate professional development.

Government recognition of the inequities between urban and rural schools has spurred a raft of policies intended to address the situation. For example, in an effort to increase teacher quality and retention in rural schools the government instituted the Special-Position Teacher Program (*Tegang Jiaoshi*) in 2016 and the Rural Teachers’ Support Program in 2015 (Liu and Hallinger, 2017). Both sought to alleviate poverty and “improve the capacity of rural school leaders and teachers (Liu and Hallinger, 2017, p. 2). Although such policies are beginning to see glimpses of success, issues such as high drop-out rates, teacher morale, poor attendance, broken families and teachers’ preference to work in cities where the pay is better continue to typify the rural landscape (e.g., Ye and Christiansen, 2009) and consume huge amounts of principals’ time; time which urban principals use to devote to instructional matters. In summary, the much-heralded success of China on international competitive assessments, while impressive, relies on a number of large urban centers and there remains substantial variation in performance across different contexts, most notably between rural and urban centers, and this impacts the exercise of school leadership. Such variation is widely recognized in China and forms part of an ongoing dialog on the future of education in the country.

The emergence of China as an educational exemplar aroused debate and controversy both within and outside of China. For many years, China had seen herself as a learner and follower of Western theories and practices. For example, in the postscript of the Chinese version of *Surpassing Shanghai*, the translator, Chinese scholar Ke Zheng, expressed puzzlement about China becoming a new world role model for education.

How have we become the “idol” of our “idols”?... We are used to thinking that we need to learn advanced educational theories and practices from the Western societies. What has happened to this world? Our idols (Western societies) start to talk about how to learn from us? (Ke, 2013, p. 215).

Chinese American scholar Prof. Zhao Yong expressed more than puzzlement when he warned that if the West were to import Chinese educational practices, it risked “drinking poison to quench one’s thirst” because of the potentially disastrous consequences of unthinking emulation which had long been recognized in China itself (Zhao, 2017). For example, the Chinese education system was perceived as being so rigid that it could do little to promote creativity or applied problem-solving.

Putting the debates aside, the ongoing desire of Western societies to learn from China has had the effect of prompting renewed reflection on the purpose, nature and workings of the education system and of individual schools within China itself. Nationally, the central government has called on all sectors of society to “tell China Stories” (*jiang zhongguo gushi*), that is, to communicate and disseminate what is happening in China to the rest of the world. This call is echoed in the field of education where participants are encouraged to commit more effort to promote Chinese voices broadly around the world (Cui, 2017).

The policy appears to be having the desired effect in different areas. For example, a recent review showed a marked quantitative increase in the number of articles about the principalship in China in international journals in the field since 2008. Between 1998 and 2007 there were only 17 articles in the area published in major English language educational leadership journals; between 2008 and 2016, 48 papers were published in the area in the same journals (Walker and Qian, 2018a). An increasing number of

indigenous scholars have taken a more contextual view in their work and become more openly questioning about what happens in Chinese schools – commenting on both the positive and negative aspects. Their insights contribute much to the international knowledge base of what educational leaders in China do and why they do what they do.

The complexity of educational leadership in China

To outsiders, many aspects of education in China appear complex and difficult to understand. Government and education systems are centralized and generally purposefully elitist and have been so for hundreds of years (Zhao and Qiu, 2010). To some extent, China maintains a stratified school system (Cheng, 2011). Resources tend to be allocated based on the rank of the school (Walker and Qian, 2018b) and differ quite markedly between rural and urban schools. Despite widespread criticism, Gaokao, or the University Entrance Examination, still dominates the scene (Harris et al., 2009; Sargent, 2011). That is, schools must perform well in examinations if they are to be ranked as successful, regardless of whether the ranking is official or unofficial (Walker and Qian, 2012). Principals are selected, appointed and appraised by the local government agency and held tightly accountable to their superiors (Qian and Walker, 2014; Zheng et al., 2013). Teachers are also placed into different categories through a hierarchical teacher ranking system (Jensen et al., 2016; Walker and Qian, 2018b).

The factors listed above are often framed as hindering the exercise of successful school leadership, leading to the question of how school leaders in such a context can exert their roles in ways which contribute to student success. Multiple empirical studies have sought to provide contextualized understandings of the practices of educational leaders in China (e.g., Liu and Hallinger, 2017; Qian and Walker, 2013, 2020; Qian et al., 2017; Walker and Qian, 2011, 2018a, 2020). These studies have helped, at least partially, to unravel the complexities and nuances that characterize the context and practices of educational leadership in China. Several propositions capture some of the major features of educational leadership in China:

- *Chinese school leaders engage closely with teaching and learning and have detailed knowledge of classroom practice.* In contrast with much western literature which suggests that principals spend only a fraction of their time on instructional issues (Murphy et al., 2016), Chinese school leaders dedicate much more substantial time to teaching and learning and maintain high levels of visibility in classrooms. Leaders' knowledge of teaching and learning informs the establishment and operation of active structures, such as teacher collaboration, that are designed to support effective instruction (Goddard et al., 2015).
- *School leaders must meet the dual expectations of improved student exam results and holistic student development; this presents them an enduring workplace tension.* The most common way for leaders to cope with the tensions flowing from the reform environment is to speak the educational languages of different systems and actors (Qian et al., 2017). In the long run, it will continue to be a bedeviling task to strike a balance between the traditional metrics of successful student learning and the imported education policies that promote student-centredness and student holistic development.
- *A set of centrally-defined organizational routines in Chinese schools facilitates instructional leadership, but school leaders still play a key role to operationalize and enrich these routines.* Routines are an important element of a school's educational infrastructure (Cohen et al., 2018) that support and coordinate instruction, maintain instructional quality and enable instructional improvement. Along with centrally mandated routines, leaders' adoption of newer practices, such as organizing school-based research and granting teachers honorable titles, has also fostered new organizational routines in Chinese schools.
- *While principals oversee school-wide improvement activities, leadership is distributed across the school leadership team and expert teachers.* A wide range of individuals, with or without formal titles, share leadership responsibilities, such as monitoring instructional activities, modeling teaching and mentoring peers. The distribution of leadership is normally designated top-down by the principal and plugged into formal and/or informal hierarchies. Regardless of the mode of distribution, as Tucker (2014) has observed, the system allows teachers (mainly those recognized as *backbone* or famous teachers) to lead the process of improving curriculum and teaching methods and to work together as a team.
- *School leaders have expanding autonomy to make school plans, design school curriculum, and develop teachers while also being aware of the limitations and constraints of their authority.* For example, they were promised autonomy to recruit teachers, but their actual ability to do so was circumscribed by the national policy of "*bianzhi*" (teacher's posts), that is, a cap on the number of teaching posts assigned to each school by the government. They had the autonomy to design a school-based curriculum, but this only accounted for a small portion of the whole curriculum, this is particularly so in rural provinces.
- *A big part of school leaders' culture-building work is modeling their focus on learning and showing high support for teacher learning.* Only when schools and leaders demonstrate they value learning through investing time and effort will teachers be convinced that their input is welcome (Bryk et al., 2010; Louis et al., 2010). Chinese school leaders generally demonstrate high levels of support for and participation in teacher professional learning.
- *School leaders devote considerable organizational resources to building harmonious relationships and team spirit.* They display a high level of paternalistic leadership, that is, their role is akin to that of a parent in a Chinese family. They care for the teachers' livelihood and social-emotional well-being (Chen and Lee, 2008). They are highly responsive to the dependency needs of teachers. Principals understand that authority alone cannot guarantee teachers' support and trust (Walker and Qian, 2018a). They need to gain legitimacy by demonstrating expert knowledge in teaching and learning and combining sincerity with benevolence.
- *Among various instructional improvement strategies, Chinese school leaders place a high priority on developing teacher capacity and creating a supportive teacher learning environment.* Principals see quality teachers as the most valuable resource and provide impressive

support for teachers' on-the-job learning. This finding supports Robinson and her colleagues' (2008) synthesis of international research which showed that leaders promoting and participating in teacher professional development makes more of a difference than any other aspect of leadership in terms of the link with student outcomes.

- *Chinese school leaders are increasingly expected to broker and shape partnerships or networks across local communities; principals from weaker network schools are expected to be effective learners and change agents in their schools.* This systemwide policy trend relies on leveraging the resources of better-performing schools to support lower-performing schools. As such, it enables schools to run joint teacher development activities and share human and teaching resources. Consequently, this emerging "system leadership" trend (Higham et al., 2009) also raises expectations of principals to develop partnerships, to cooperate with their principal peers and to design more school-specific instructional improvement strategies.

In sum, a complex mix of leadership practices defines school principalship in China, and these can come together in different ways depending on the context within which leadership is exercised. On the one hand, their work environment is highly political - they must be conscious of their role as state employees and their accountability to and dependence on various government agencies. On the other hand, there are strong professional expectations of school principals from the system, teachers and school communities (Qian, 2018; Walker and Qian, 2018a,b). Research then seems to depict a different image of educational leadership in China from the stereotypical picture painted in the traditional Western discourse, which tends to view Chinese leaders as autocratic, hierarchical, and formal (House et al., 2004).

It is not our intention to romanticize education in China or to discount the myriad of serious issues and problems currently active in the country, we have already commented on the discrepancies in educational quality and equity between rural and urban provinces. We therefore suggest that what is judged as "good" or "bad" is relative to changing circumstances or the conditions surrounding them. This is well illustrated in a Chinese proverb *Fortune and misfortune are two buckets in a well* (*fu xi huo suo fu, huo xi fu suo yi*), which holds that every event, every condition, every bit is part of a larger whole. Different buckets dipped into the same well can reel up quite different loads, looking into one bucket risks presenting an incomplete representation (Walker and Qian, 2018b). For example, when governments are strong and directive in terms of curriculum frameworks and pedagogy, another bucket in the same well is that strong governments can easily mobilize huge amounts of monetary and personnel resources to organize cross-school and even cross-provincial teacher development programmes to promote teacher learning (Walker and Qian, 2018b, 2020). This intense and large-scale teacher learning has often been cited as critical to success (Jensen et al., 2016).

What shapes educational leadership in China?

Many different influences shape educational leadership in China. If we focus too closely on key facets of the system only, we risk ignoring that the success (or otherwise) of these particular factors is made possible only by other system features (Tucker, 2014). Whereas we don't claim these as the only contextual factors, the following three features appear key when explaining the enactment of Chinese school leadership.

- Some cultural scripts underpinning the "right" leadership as perceived by leaders and followers
- Interaction and multifarious linkages between imported Western educational policies and traditional Chinese culture
- A simultaneous push-pull effect exerted by some contextual factors

These contextual features are further elaborated below.

Feature 1: Some deep structures continue to shape school leadership and preferred leadership styles in China. For 2000 years, the ancient education system was built around the Civil Service Examination system (Gu, 1981; Mao, 1984). Confucianism underpins Chinese perceptions of the purpose of education; learning has been unambiguously connected with passing exams to bring glory to "families and ancestors" (Gu, 2006, p. 173). The learning strategies were based on repetition and memorization as in Confucian philosophy, repetition is perceived as a necessary method for deep understanding and the production of meaning (Bowden et al., 2015). Therefore, and regardless of their location, school leaders today remain caught in the tensions and conflicts arising from the lingering impact of the traditional emphasis on examinations.

Traditional expectations also set a very high bar for Chinese leaders, one replete with tension. For example, the culture reveres hierarchy but expects leaders to be responsive to the needs of followers. There is also a consensus that hierarchy and distinction should be based on some kind of merit, whether it is related to morality, competence or actual performance (Walker and Qian, 2018a). Furthermore, this merit should be achieved by individuals rather than ascribed to them through inheritance, such as birth, class or other social categories (Parsons and Shils, 1951). In short, authority is expected to combine both elitism and sympathy (Pye, 1991). Mirroring the prototype of the Chinese family, the Chinese leader is often portrayed as the father (Farh et al., 2008). As the head of the family, the father's role is to rule benevolently and justly. The other family members are expected to be obedient and loyal to the head of the family (Fu and Tsui, 2003). Similarly, the leader of an organization is expected to provide guidance, protection, nurture and care for subordinates; the subordinate, in return, is normally required to be loyal and deferential to the superior (Aycan, 2006).

Feature 2: Imported Western education policies and leadership philosophies have interacted and intertwined with traditional Chinese culture to shape leadership and organizational behaviors in schools. China is making enormous strides in education and appears committed to a path of constant improvement. This did not happen overnight - the last 30 years has witnessed gradual improvement in this area.

In the 1970s and 1980s, the push was to expand access to schools for all children. In the 1990s, the focus shifted to improving quality. A curriculum reform was piloted in Shanghai and later spread with various degrees of success around the country. There was also a major emphasis on upgrading the teaching force and trying to reduce examination pressure (Tucker, 2014). Over the past decade, improving quality and enhancing equity has become the dual goals of educational development (Ministry of Education, 2010).

Under the banner of Quality Education (*suzhi jiaoyu*), China has strived to move away from examination-oriented and teacher-centered traditions through various waves of educational reform (Cheng, 2011; Yin, 2013). While purposefully addressing a range of local issues, the reforms have drawn heavily on educational and political trends in Western societies. The central aim of these initiatives was to transform the traditional examination-oriented education system into one valuing quality-oriented, student-centered education to cultivate talented and innovative citizens for the betterment of Chinese society (Huang and Wiseman, 2011; Qian and Walker, 2013). It has not been easy for schools and systems to absorb imported Western education policies into their time-embedded educative practices. Such adaptation is often more difficult in remote and poorer provinces where traditions tend to run even deeper than in the population centers. The clash between the old and the newly imported has resulted in “the co-existence of Western-style isomorphic structures at the surface level of ministries, constitutions or management practices and, behind this front, a complex interplay of cultural deep structures, entrenched attitudes, specific meaning-processing schemata, and vested interests” (Schriewer, 2012, p. 417).

Feature 3. Some of the contextual features themselves present a double-edged sword in that they can exert enabling and/or constraining effects on school leadership. Some contextual features exert a push-pull effect on leadership. In other words, they simultaneously facilitate and constrain the principals’ leadership power to improve student learning (Qian, 2018; Walker and Qian, 2015). If we do not recognize the seemingly contradictory role of these contextual features, we may also fail to unveil the complexities facing school leaders. Two examples of such contextual features are paternalistic values and the presence of the strong state.

As stated earlier, Chinese principals tend to assume a parent-like role at their school (Chen and Lee, 2008). In other words, they expect loyalty and subordination from teachers. It is sometimes difficult for Chinese principals to introduce and implement radical changes at their schools; their concern is that these may cause potential conflict and threaten organizational harmony. At the same time, a central theme of paternalistic leadership philosophies is that leaders need to care for their subordinates’ livelihoods and social and psychological well-being (Chen and Lee, 2008). In fact, Sargent and Hannum (2005) suggested that out of concern for teacher turnover, rural principals intentionally focused on “keeping teachers happy” (also see Qian et al., 2017). Principals understand that authority alone cannot guarantee teachers’ support and trust (Qian, 2018). School leaders tend to extend their care to teachers and try to address teachers’ concerns within and beyond the schools.

With the presence of the strong state, principals’ autonomy is circumscribed, and they are held accountable across a broad range of areas. Chinese principals depend on the government for job security and promotion and are expected not to question the authority of local governments (Walker and Qian, 2018a). In short, they are pressured to prioritize requirements associated with upward accountability. However, a strong state is also a source of support for schools. Support comes primarily in the form of relatively stable government funding, including the financial and intellectual resources provided to schools. The strong state provides top-down support as it imposes firm external control (Qian and Walker, 2019).

In sum, these three contextual features underpin the way school leaders enact their practices that may seem to be complex and puzzling in the eyes of outsiders. School leaders in China are influenced by both the “east wind” (home-grown values and philosophies) and “west wind” (ideas and theories imported from the West) (Tan and Chua, 2015). A challenge facing educational leaders in China is to translate imported policies and concepts such as student-centered learning and all-round student development into school realities (Huang and Wiseman, 2011). School leaders need to widen their leadership repertoire to include new capacities such as developing and designing curriculum, sharing power and delegating responsibility and providing individualized support for teachers. At the same time, school leaders work in a hierarchical system; thus, there is a general expectation that they will comply with established social norms. The struggle between the “west wind” and the “east wind” is constant and school leaders’ work lives are often fraught with tensions and dilemmas caused by the crosswinds (Qian et al., 2017).

An implication of the complexity of Chinese principals’ worklives is that understanding of educational leadership must be firmly located within the political and educative system; the sum is more than its parts (Tucker, 2011, 2014). Many constituent parts of the system, including the emphasis on exams, high parental expectations, paternalistic culture and strong state, present a double-edged sword. If we cannot grasp holistically the way the whole comes together, we risk providing skewed descriptions of what the Chinese school leadership looks like.

Learning from China?

The outstanding performance of Asian students on international exams has prompted a heretofore unseen enthusiasm for “reverse” cross-cultural learning about education – the West learning from the East. This trend has brought both new opportunities and new challenges to everyone interested in improving educational quality both within and outside of China through comparative studies, from policymakers to classroom teachers (McNess et al., 2015). The recently-emerged evidence-based understanding of Chinese educational leadership seems to challenge the stereotypical images widely presented in Western discourse.

A possible reason for the discrepancy is that outsiders tend to “fixate on China’s authoritarian political system” and this fixed view can discount the possible value and relevance of its decisions (Johnson, 2020). Although Chinese educational leaders, to a large

extent, need to run their schools in a highly centralized system today, it would be one-sided to label them as just autocratic or hierarchical – it's much more complex than this. Chinese school leaders adopt various strategies to build and strengthen internal accountability while also responding to external demands. On the one hand, the school leaders feel their responsibility is overwhelming and their autonomy has many constraints. On the other hand, they share a common belief that they must hold themselves accountable for the state of student learning across the school (Qian and Walker, 2019). This belief helps them to internalize many external expectations for performance and use different means to make the expectations shared and internalized by school teachers (Qian and Walker, 2019; Walker and Qian, 2018a).

We understand that education is a product of social systems that are culturally, historically and politically determined by “forgotten struggles and difficulties” and “battles long ago” (McNess et al., 2015, p. 296). Thus, different practices or policies have to be understood in terms of their cultural-historical contexts. There is no one best way of teaching and learning, and no one best way of leading. Caution is warranted to avoid drawing simplistic causal relationships between an identified policy/practice and outstanding student performance (Walker and Qian, 2019). Under this caveat, it is worth revisiting what can be learned at a general level about educational leadership in China. After visiting Shanghai, Thomas Friedman (2013) wrote in *The New York Times* about the secret of Shanghai success. He believed that if there was a secret, it was a relentless focus on the basics that make for high-performing schools but are difficult to pull off consistently across an entire school system. Similarly, if there is anything that Chinese educational leaders do that makes a real difference in student learning, it is their unremitting focus and deep commitment to the basics. These include:

- Strong emphasis on classroom teaching
- Priority on teacher capacity building
- Investment of time and commitment to teacher learning
- Awareness of building their personal professional authority
- Protection and care for teachers' professional and personal lives
- Focus on relationship building within a family-like school

Like everywhere, Chinese school leaders struggle through multiple challenges to maximize student success and wellbeing and weave a professional culture among teachers. Their purpose and passion match that of their principal colleagues internationally, even as the paths they both follow and blaze differ in form and density. The increasing volume of empirical research in the area coming out of China will further inform how and why they carve their paths to success and thus contribute to the understanding of school leadership across different countries, contexts and cultures.

Acknowledgment

The authors wish to acknowledge the support of the funding support of the Research Grant Council (RGC) of Hong Kong through the General Research Fund (No. 18611117).

References

- Aycan, Z., 2006. Paternalism: towards conceptual refinement and operationalization. In: Yang, K.S., Hwang, K.K., Kim, U. (Eds.), *Indigenous and Cultural Psychology: Understanding People in Context*. Springer, New York, pp. 445–466.
- Bowden, M.P., Abhayawansa, S., Manzin, G., 2015. A multiple cross-cultural comparison of approaches to learning. *Compare* 45 (2), 272–294.
- Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., Easton, J., 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press, Chicago.
- Chen, C.C., Lee, Y.T., 2008. *Leadership and Management in China: Philosophies, Theories, and Practices*. Cambridge University Press, Cambridge.
- Cheng, K.M., 2011. Shanghai and Hong Kong: two distinct examples of education reform in China. In: OECD (Ed.), *Strong Performers and Successful Performers in Education: Lessons from PISA for the United States*, pp. 83–111. Retrieved from: https://www.oecd-ilibrary.org/education/lessons-from-pisa-for-the-united-states_9789264096660-en.
- China Daily, 2019. UK Teachers Exchanging Math Insights in Shanghai, 13 November [Online]. Available at: <https://global.chinadaily.com.cn/a/201911/13/WS5dcb09a0a310cf3e35576eb8.html>. (Accessed 9 October 2020).
- Cohen, D., Spillane, J., Peurach, D., 2018. The dilemmas of educational reform. *Educ. Res.* 47 (3), 204–212.
- Cravens, X., Drake, T.A., Goldring, E., Schuermann, P., 2017. Teacher peer excellence groups (TPEGs): building communities of practice for instructional improvement. *J. Educ. Adm.* 55 (5), 526–551.
- Cui, Y.H., 2017. Towards international sharing of Chinese experiences in education [Zouxiang jiaoyu de “zhongguo jingyan shijie fenxiang”] *J. Chin. People's Polit. Consul. Conf.* 26 July [Online]. Available at: <https://news.ecnu.edu.cn/79/56/c1835a96598/page.htm>. (Accessed 9 October 2020).
- Dimmock, C., 2020. Connecting research and knowledge on educational leadership in the West and Asia: adopting a cross-cultural comparative perspective. *Comp. Educ.* 56 (2), 257–277.
- Farh, J.L., Liang, J., Chou, L.F., Cheng, B.S., 2008. Paternalistic leadership in Chinese organizations: research progress and future research directions. In: Chen, C.C., Lee, Y.T. (Eds.), *Business Leadership in China: Philosophies, Theories, and Practices*. Cambridge University Press, Cambridge, UK, pp. 171–205.
- Friedman, T., 2013. The Shanghai secret. *N. Y. Times*, 22 October [Online]. Available at: <https://www.nytimes.com/2013/10/23/opinion/friedman-the-shanghai-secret.html>. (Accessed 9 October 2020).
- Fu, P.P., Tsui, A.S., 2003. Utilizing printed media to understand desired leadership attributes in the People's Republic of China. *Asia Pac. J. Manag.* 20, 423–446.
- Goddard, R., Goddard, Y., Kim, E.S., Miller, R., 2015. A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *Am. J. Educ.* 121 (4), 501–530.
- Gu, S., 1981. *Zhongguo lidajiaoyu zhidu* [A Historical Account of Educational Systems in China]. Jiangsu People's Publishing House, Nanjing, China.
- Gu, M.Y., 2006. An analysis of the impact of traditional Chinese culture on Chinese education. *Front. Educ. China* 2, 169–190.

- Hallinger, P., Bryant, D.A., 2013. Mapping the terrain of research on educational leadership and management in East Asia. *J. Educ. Adm.* 51 (5), 618–637.
- Harris, J., Zhao, Y., Caldwell, B.J., 2009. Global characteristics of school transformation in China. *Asia Pac. J. Educ.* 29 (4), 413–426.
- Higham, R., Hopkins, D., Matthews, P., 2009. *System Leadership in Practice*. Open University Press, Berkshire.
- House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., Gupta, V., 2004. Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies. Sage, Thousand Oaks, CA.
- Huang, T.D., Wiseman, A.W., 2011. The landscape of principal leadership development in Mainland China: an analysis of Chinese and English research. In: Huang, T.D., Wiseman, A.W. (Eds.), *The Impact and Transformation of Education Policy in China*. Emerald, Bingley, pp. 125–152.
- Jensen, B., Sonnemann, J., Roberts-Hull, K., Hunter, A., 2016. *Beyond PD: Teacher Professional Learning in High-Performing Systems*. National Center on Education and the Economy, Washington, DC.
- Johnson, I., 2020. China bought the West time. The West squandered it. *N. Y. Times*, 13 Mar [Online]. Available at: <https://www.nytimes.com/2020/03/13/opinion/china-response-china.html>. (Accessed 9 October 2020).
- Ke, Z., 2013. Postscript. In: Tucker, M. (Ed.), *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems*. East China Normal University Press, Shanghai.
- Liang, X.Y., Kidwai, H., Zhang, M.X., 2016. How Shanghai Does it: Insights and Lessons from the Highest-Ranking Education System in the World. World Bank, Washington, DC.
- Liu, S.N., Hallinger, P., 2017. Leading teacher learning in China: a mixed methods study of successful school leadership. In: Leithood, K., Sun, J.P., Pollock, K. (Eds.), *How School Leaders Contribute to Student Success: The Four Paths Framework*. Springer, Dordrecht, Netherlands, pp. 279–303.
- Liu, S.N., Hallinger, P., 2020. Unpacking the effects of culture on school leadership and teacher learning in China. *Educ. Manag. Adm. Leader* 1–20. <https://doi.org/10.1177/1741143219896042>. (Accessed 9 August 2020).
- Liu, S.N., Hallinger, P., Feng, D.M., 2016. Supporting the professional learning of teachers in China: does principal leadership make a difference? *Teach. Teach. Educ.* 59, 79–91.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., Anderson, S.E., 2010. *Investigating the Links to Improved Student Learning: Final Report of Research Findings to the Wallace Foundation*. University of Minnesota, Minneapolis.
- Mao, L., 1984. *Zhongguo jiaoyushi jianbian* [History of Chinese Education]. Education Science Publisher, Beijing, China.
- McNess, E., Arthur, L., Crossley, M., 2015. "Ethnographic dazzle" and the construction of the "other": revisiting dimensions of insider and outsider research for international and comparative education. *Compare* 45 (2), 295–316.
- Ministry of Education, 2010. *Guojia zhongchangqi jiaoyugaige He Fazhanguihua Gangyao* [Outline of the National Medium- and Long-Term Programme for Education Reform and Development (2010–2020)]. Available at: http://www.moe.edu.cn/srcsite/A01/s7048/201007/t20100729_171904.html. (Accessed 9 October 2020).
- Murphy, J., Neumerski, C.M., Goldring, E., Grissom, J., Porter, A., 2016. Bottling fog? The quest for instructional management. *Camb. J. Educ.* 46 (4), 455–471.
- OECD, 2010. PISA 2009 Results: Executive Summary. Available at: www.oecd.org/pisa/pisaproducts/46619703.pdf. (Accessed 9 October 2020).
- OECD, 2013. PISA 2012 Results in Focus: What 15-Year-Olds Know and what They Can Do with what They Know. OECD Publishing.
- Parsons, T., Shils, E.A., 1951. *Toward a General Theory of Action*. Harvard University Press, Cambridge, MA.
- Pye, L., 1991. The new Asian capitalism: a political portrait. In: Berger, P.L., Hsiao, H.H. (Eds.), *In Search of an East Asian Development Model*. Transaction Books, New Brunswick, NJ, pp. 81–98.
- Qian, H.Y., Walker, A., 2013. How principals promote and understand teacher development under curriculum reform in China. *Asia Pac. J. Teach. Educ.* 41 (3), 304–315.
- Qian, H.Y., Walker, A., 2014. Principalship in China: emerging propositions. *Lead. Manag.* 20 (2), 60–74.
- Qian, H., Walker, A., 2015. The education of migrant children in Shanghai: the battle for equity. *Int. J. Educ. Dev.* 44, 74–81.
- Qian, H.Y., Walker, A., 2019. Reconciling top-down policy intent with internal accountability: the role of Chinese school principals. *Educ. Assess. Eval. Account.* 31 (4), 495–517. <https://doi.org/10.1007/s11092-019-09309-4>. (Accessed 9 October 2020).
- Qian, H.Y., Walker, A., 2020. Creating conditions for professional learning communities (PLCs) in schools in China: the role of school principals. *Prof. Dev. Educ.* 1–13. <https://doi.org/10.1080/19415257.2020.1770839>. (Accessed 9 October 2020).
- Qian, H.Y., Walker, A., Yang, X., 2016. Building and leading a learning culture among teachers: a case study of a Shanghai primary school. *Educ. Manag. Adm. Leader* 45 (1), 101–122.
- Qian, H.Y., Walker, A., Li, X., 2017. The west wind vs. the east wind: instructional leadership model in China. *J. Educ. Adm.* 55 (2), 186–206.
- Qian, H., 2018. School principalship in China: a collective portrait. In: Kennedy, K.J., Lee, J.C.-K. (Eds.), *Routledge International Handbook of Schools and Schooling in Asia*. Routledge, New York, pp. 1015–1026.
- Robinson, V.M.J., Lloyd, C., Rowe, K., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–647.
- Sargent, T., Hannum, E., 2005. Keeping teachers happy: job satisfaction among primary school teachers in rural northwest China. *Comp. Educ. Rev.* 49 (2), 173–204.
- Sargent, T.C., 2011. New curriculum reform implementation and the transformation of educational beliefs, practices, and structures in Gansu province. *Chin. Educ. Soc.* 44 (6), 47–72. <https://doi.org/10.2753/CED1061-1932440604>. (Accessed 9 October 2020).
- Schriewer, J., Martinez, C., 2004. Constructions of internationality in education. In: Steiner-Khamsi, G. (Ed.), *The Global Politics of Educational Borrowing and Leading*. Teachers College Press, New York, pp. 29–53.
- Schriewer, J., 2012. Editorial: meaning constellations in the world society. *Comp. Educ.* 48 (4), 411–422.
- Sellar, S., Lingard, B., 2013. Looking East: Shanghai, PISA 2009 and the reconstitution of reference societies in the global education policy field. *Comp. Educ.* 49 (4), 464–485.
- Stambeck, A., 2003. World-cultural and anthropological interpretations of 'choice programming' in education. In: Anderson-Levitt, K.M. (Ed.), *Local Meanings of Global Schooling: Anthropology and World Culture Theory*. Palgrave, New York, pp. 141–160.
- Tan, C., Chua, C.S.K., 2015. Education policy borrowing in China: has the West wind overpowered the East wind? *Compare* 45 (5), 686–704.
- Tan, C., 2013. *Learning from Shanghai: Lessons on Achieving Educational Success*. Springer Science & Business Media, Dordrecht, The Netherlands.
- Tucker, M.S., 2011. *Surpassing Shanghai: An Agenda for America Education Built on the World's Leading Systems*. Harvard Education, Cambridge, MA.
- Tucker, M.S., 2013. *Tucker's Lens: Curriculum in Scotland and England: A Study in Contrasts*. Available at: <http://ncee.org/2013/07/tuckers-lens-curriculum-in-scotland-and-england-a-study-in-contrast/>. (Accessed 9 October 2020).
- Tucker, M.S., 2014. Chinese Lessons: Shanghai's Rise to the Top of the PISA League Tables. National Center on Education and the Economy, Washington, DC.
- Walker, A., Qian, H.Y., 2011. Successful school leadership in China. In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, New York, pp. 446–457.
- Walker, A., Qian, H.Y., 2012. Reform disconnection in China. *Peabody J. Educ.* 87 (2), 162–177.
- Walker, A., Qian, H.Y., 2015. Review of research on school principal leadership in mainland China, 1998–2013: continuity and change. *J. Educ. Adm.* 53 (4), 467–491.
- Walker, A., Qian, H.Y., 2018a. Deciphering Chinese School Leadership: Conceptualizations, Context and Complexities. Routledge, London.
- Walker, A., Qian, H.Y., 2018b. Exploring the mysteries of school success in Shanghai. *ECNU Rev. Educ.* 1 (1), 119–134.
- Walker, A., Qian, H.Y., 2019. Leadership and culture. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *Principles of Educational Leadership and Management*, third ed. SAGE Publications Ltd, London, pp. 311–330.
- Walker, A., Qian, H.Y., 2020. Developing a model of instructional leadership in China. *Compare*. <https://doi.org/10.1080/03057925.2020.1747396>. (Accessed 9 October 2020).
- Ye, X., Christiansen, F., 2009. China's urban-rural integration policies. *J. Curr. Chines Aff.* 38 (4), 117–143.
- Yin, H.B., Lee, C.K., Wang, W.L., 2013. Dilemmas of leading national curriculum reform in a global era: a Chinese perspective. *Educ. Manag. Adm. Leader* 42 (2), 293–311.
- Yin, H.Q., 2013 [Solving problems when educational reform comes to a difficult stage]. *Zai jiaoyu neihan fazhan de "shenshuiqu" gongjiankenan*, vol. 12B. Shanghai Education, pp. 14–16.

- Zhao, Y., Qiu, W., 2010. China as a case study of systemic educational reform. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*. Springer, Dordrecht, Netherlands, pp. 349–361.
- Zhao, Y., 2017. Fatal Attraction: Why Coping Each Other Won't Improve Education? Available at: <https://nepc.colorado.edu/blog/fatal-attraction>. (Accessed 9 October 2020).
- Zheng, Y., Walker, A., Chen, S., 2013. Change and continuity: a critical analysis of principal development policy in Mainland China (1989–2011). *Educ. Manag. Adm. Leader* 41 (4), 484–503.
- Zheng, X., Yin, H., Li, Z., 2019. Exploring the relationship among instructional leadership, professional learning communities and teacher self-efficacy in China. *Educ. Manag. Adm. Leader* 47 (6), 843–859.

Women school leadership in Malaysia: perspectives and enactments

Donnie Adams and Ravadhi Periasamy, Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	484
Research context	485
Women leadership in Malaysia	485
Methodology	485
Findings	486
Encourage teamwork	486
Accessible and caring	486
Inclusiveness	486
Instructional and curriculum knowledge	487
Firm and perfectionist	487
Good communicators	487
Diligent and hardworking	487
Reflections	487
Limitations and future studies	488
Conclusion	488
References	488

Introduction

There is extensive empirical evidence indicating that school principals are one of the most salient determinants in the improvement of school performance (Adams, 2018; Day et al., 2016; Hallinger and Heck, 2010; Harris et al., 2019; Leithwood et al., 2020; Robinson et al. 2008). Hotmire (2018) states that school principals are primarily responsible for student achievement and they dictate the implementation of any school-level reforms. Thus, in order to create schools on par with the increasingly complex demand of the 21st century, principals play a key role in improving school outcomes by influencing the motivations and capacities of teachers, as well as the school climate and environment (Sun and Leithwood, 2015).

School principals must solve problems and make decisions concerning school functions such as curriculum design, financial resourcing, and assessment of student achievement on a daily basis (Drago-Severson et al., 2018). Consequently, leadership and leadership development continue to be a priority for policy makers in many education systems as a means to improve its educational system (Harris and Jones, 2015; Walker, 2015). Likewise, increasing attention has been paid across different countries to the type or form of leadership associated with school improvement and transformation (Hallinger and Lee 2014; Robinson et al. 2008).

As education systems continue seeking emerging leadership talents, Leimon et al. (2011) claimed that women are an under-utilized resource, therefore, women should be coached to lead. Stead and Elliot (2009) reveal two crucial observations in their study of women's leadership. First, women leaders draw upon both past and present experiences in their practice of leadership. Second, relationship to work (developing professional networks), relationship to others and relationship to place (where she lives) are the important spheres and influences in women leadership. Whereas, married women leaders face the greatest challenge in trying to balance work-family responsibilities.

It is a common knowledge that women and men have certain qualities that distinguish one from the other (Msila, 2013). Valerio (2009) points out that women tend to adopt a more democratic (participative) style of leadership where else men to adopt a more autocratic (directive) style. Additionally, research has shown women principals are more of transformational leaders than male principals (Msila, 2013). However, despite a gradual change today, women leaders are still far from achieving equality in the workplace (Stead and Elliott 2009). TALIS data show that in 2018, women are a minority among school principals in around half of the participating countries and economies with only 47% of principals are women (OECD, 2020).

Given that a robust knowledge base on women leadership is yet to be established, therefore, the aim of this article is to offer new insights into the essential characteristics of women leadership in Malaysia from the perspective of headmistresses in a small sample of Tamil vernacular primary schools and to suggest implications for the development of women as school principals in Malaysia. The article commences with an overview of the research context, the literature on women leadership in Malaysia and subsequently details of the methodology are outlined, followed by a discussion of the findings. Key themes and reflections are presented with some concluding comments and observations.

Research context

Currently there are 527 Tamil vernacular primary schools in Malaysia. These schools are divided into two categories: government-aided and semi-government aided. The Tamil language is the main medium of instruction. Tamil vernacular primary schools are mostly attended by Malaysian Indian students (aged 7 years–12 years) with the majority coming from Tamil descent and very few from other races. Tamil education is only available at the primary school level. Next, students must pursue their secondary and post-secondary education in national schools, where Malay will be the medium of instruction. As for its curriculum, these schools adopt the national curriculum as other government schools where Malay and English languages are the compulsory subjects (Ibrahim, 2018).

Women leadership in Malaysia

The question of whether women and men use different approaches to leadership has been the subject of numerous studies (Kolehmaineh et al., 2014; Lyman et al., 2005; Offermann and Coats, 2018) and whether gender and leadership are interrelated or independent has always been a constant source of debate (Husu et al., 2010). Since the 20th century, women leadership has become the focus and interest among researchers especially in the education sector (Abu-Tineh, 2013; Atieno, 2013; Bissessar, 2013; Kaiser and Wallace, 2016).

Tracing the historical research of women in educational administrations, Shakeshaft (1989) explained that women do not have the opportunities, career related experiences, or the benefits as given to men. But, the 21st century has seen more women appointed in leadership positions, taking on more businesses, and procuring more degrees than ever before. With their insightfulness, interpersonal abilities, and aptitude for lifelong learning, women are well positioned to achieve even greater career heights in the coming years (Wilén-Daugenti et al., 2013). However, even though women leaders have more access to education and are considered economically and politically equal to men, they are still overlooked when it comes to leadership positions (Mendez-Morse et al., 2015).

Women leaders constantly face barriers getting into higher positions as society largely feels males are better leaders (Brinia, 2012; Cubillo and Brown, 2003; Badura et al., 2018). This continues to become a challenge to women leaders in carrying out their roles and functions as they do not fit the norms of the society (Adams and Muthiah, 2020). Crosby-Hillier (2012) adds among the other barriers faced by young women leaders are multiple roles such as being a mother, a wife and at the same time an aspiring school leader. This indicates gender plays a role in shaping the types of leadership provided by women and men (Crosby-Hillier, 2012).

Research indicates that women have a different leadership style than men. Elmuti et al. (2009) indicates that women are generally more sensitive and encouraging leaders. Female school leaders perform better in coordinating curriculum, communicating school goals, promoting professional development, instruction supervising and evaluating, being visible in school and incentive provider for learning (Kaputa, 2009; Kwadzo, 2010; Mollel and Tshabangu, 2014). Prime et al. (2009) also postulated that women school leaders always practice self-development, they are good in building lasting relationships, inspiring and motivating others, and taking initiative to drive results.

Additionally, women school leaders were also found to be more collaborative in nature and gives importance to the growth of the school community as they empower others in the organization (Makura, 2012; Perkins et al., 2013). Women school leaders tend to focus on curriculum development, instructional programs management, developing teaching and learning in a collaborative and participatory manner (Nogay and Beebe, 2008), mentoring teachers to develop their instruction in the classroom and forming professional learning communities (Arar and Oplatka, 2016).

Methodology

The study reported in this article employed a case-study approach (Strauss and Corbin, 1998; Yin, 2013). The limitations of case studies in providing any causal explanations are fully accepted and acknowledged, and the study aimed to offer new insights into the essential characteristics of women leadership in Malaysia from the perspective of headmistresses in a small sample of Tamil vernacular primary schools. The findings therefore are indicative rather than definitive.

There are 15 government-funded Tamil vernacular primary schools in Kuala Lumpur. Most school heads are women especially in Tamil vernacular primary schools. Thus, 13 headmistresses were selected as samples in the study. These headmistresses have led their respective school for more than three years. Most of them have been teaching for more than 20 years and were a senior assistant before being promoted to become a headmistress. Information about the headmistresses is presented in Table 1 below. At the time of this study, the other 2 school that were excluded from this study had no head teachers and were awaiting appointment of a head teacher for their schools from the Ministry of Education.

The interviews were conducted using a semi-structured approach along with 24 open-ended questions as a standard protocol, and each interview lasted approximately 45 min. In ensuring that this research follows ethical standards, all headmistresses completed a consent form and were offered the opportunity to check their transcripts and to correct any errors. All interviews were digitally recorded and subsequently fully translated and transcribed into English. The transcripts were then transferred into

Table 1 Demographics of the headmistresses and schools.

<i>Code</i>	<i>Experience as a headmistresses/Year of appointment</i>	<i>School size</i>	<i>Students' socio-economic status</i>
School A	11 years/2010	193 students	Mostly low income
School B	10 years/2011	450 students	Mostly middle income
School C	9 years/2012	120 students	Mostly low income
School D	16 years/2005	350 students	Mostly middle income
School E	20 years/2001	135 students	Mostly middle income
School F	9 years/2012	330 students	Mostly middle income
School G	9 years/2012	290 students	Mostly middle income
School H	16 years/2005	320 students	Mostly middle income
School I	11 years/2010	290 students	Mostly low income
School J	8 years/2013	300 students	Mostly middle income
School K	14 years/2007	320 students	Mostly middle income
School L	9 years/2012	80 students	Mostly low income
School M	10 years/2011	150 students	Mostly middle income

ATLAS.ti 8, a qualitative data analysis software. The analysis of the qualitative data followed the coding and constant comparison techniques advocated by [Strauss \(1987\)](#).

The data were analyzed using a four-step thematic analysis approach. First, the transcripts of the recorded interviews were read to ensure familiarity with the data sets. Second, a coding framework was developed based on the international literature. Third, a constant comparative analysis was used to generate key themes and propositions for further investigation and further testing against the data ([Glaser and Strauss, 2017](#)). Finally, data were synthesized for emerging patterns and to generate emergent themes. The data were reviewed once again to detect any new emerging codes.

The next sections of this article present some of the findings from this study, highlighting some of the main themes emerging from the data analysis, particularly those related to women leadership.

Findings

Encourage teamwork

The data show that headmistresses award importance to teamwork because they trust that a well-supported leadership team in schools can instil a feeling of ownership. Through teamwork, these headmistresses motivate their teachers to work hand in hand in a meaningful way toward success. The headmistresses said it's important for teachers to work together to help the school attain its goals. They further elaborated as the headmistress; they should have the ability to make teachers work together as a team to achieve the school's goals and vision.

Accessible and caring

Most of the headmistresses in this study feel that they are accessible, caring and supportive to all the stakeholders in the school. They listen to their teachers' and students' problems and are readily approachable. Teachers and students can come and see them anytime. They also make themselves accessible to parents. There is a culture of give and take among them and their teachers.

Teachers are allowed leave school early if they have any personal matters to attend to. Most headmistresses felt that greeting students and talking to parents is a way of showing that they care and are accessible for all their students and parents.

Inclusiveness

Headmistresses in this study encourage inclusiveness rather than exclusiveness in all aspects of their school management. These headmistresses involve their staff and students to maintain a closely knit organization. Teachers are involved in all the school programs. These headmistresses believe this gives their teachers higher job satisfaction and that teachers give their fullest attention to their work. On certain occasions, they take their teachers out for lunch or dinner to foster closeness and greater relationship.

Additionally, parents and Non-Governmental Organisations are encouraged to be involved in school activities such as gotong-royong, canteen day and night classes to help the headmistresses achieve what they have planned. These headmistresses shared that these efforts have encouraged parents to participate and offer their full cooperation for any of the school activities. Among the contribution from parents and the Non-Governmental Organisations are setting up reading huts in the schools by donating books and money.

Instructional and curriculum knowledge

The headmistresses in this study are experienced educators which help them to better understand their job and help them to perform their duties more effectively. Most of them had been teaching for more than 20 years and were senior assistants before being promoted to headmistress. They have wide expertise in instructional and curriculum knowledge which provides them a platform for leading instructional efforts in their schools.

The headmistresses attempt to visit classrooms daily to observe instruction and student learning and identify what are the needs and strengths of their teachers and students. They are committed to making teaching and learning a positive experience in their schools. Their vast teaching experience equips them with the skills and knowledge on teaching, instructional leadership, organizational management, and this helps bring change to their schools.

Firm and perfectionist

Findings also indicate headmistresses in this study are firm, fair, and a perfectionist. They are firm in achieving their goals and are strict when it comes to work. These headmistresses claim they need to be firm in order to push their teachers to achieve the school goals. In addition, they are a perfectionist. These headmistresses ensure everything must be completed without any flaws. They also ensure all work is done perfectly and on time. There are some instances these headmistresses are labeled as workaholic and autocratic due to their firmness and perfections.

This is to raise awareness among their teachers so that they value the schools' goals above their personal goals. Admittedly, this has resulted in few teachers being stressed or burned out, however, these headmistresses felt it is for the betterment of the school and to ensure improvement in student outcomes.

Good communicators

Headmistresses in this study feel they are good communicators with the schools' stakeholders. They shared the need for good communication skills so that all schools' stakeholders can relate to their goals and vision and work together toward achieving these goals. These headmistresses believe communication skills are important if a strong relationship is to be built. It is impossible to achieve the school goals if stakeholders in the school do not communicate effectively. Furthermore, these headmistresses often communicate with their teachers in order to build a close relationship with them and earn their trust. This also enables the headmistresses to constantly motivate their teachers and immediately solve problems arising in their schools.

Diligent and hardworking

This study also found headmistresses work hard together with their teachers in order to achieve the school goals. Their work ethics are considered exemplary by their teachers. The headmistresses shared that the expectations from the schools' stakeholders spurred them to work diligently. Most of the headmistresses prepare their own presentations for the school meetings with data in hand so that all teachers will understand their expectations and instruction clearly.

Additionally, these headmistresses work hard to plan many types of programmes for their schools as part of their school leadership and management responsibilities. They emphasize teaching and learning and are constantly searching for new ideas to bring changes to their schools. In return, teachers are expected to work equally hard with them in order to increase students' achievement. Extra emphasis is given for teachers to attend professional development courses especially during school holidays.

Reflections

The study reported in this article aimed to illuminate the essential characteristics of women leadership in Malaysia from the perspective of headmistresses in a small sample of Tamil vernacular primary schools and to suggest implications for the development of women as school leaders in Malaysia. There are four main reflections from this study. First, most of these headmistresses understood women must become visible and vocal in our educational institutions. They must display their intellectual competence, values, knowledge, and vision for the future. If women are to emerge as members of the community, respected for their contributions and achievements, they must also be given an opportunity to lead in schools, where they could influence student learning outcomes (Esa and Leong, 2013). Recent studies have shown that women leaders tend to lead their organizations by showing women leadership characteristics which affects their schools' performance and transformation positively (Kythreotis et al., 2010).

Second, these headmistresses were able to describe their leadership practices as one that conforms to the standards of "women leadership". We could suggest, based on the data, that the thirteen headmistresses in this study are strong women leaders in certain respects. For example, these headmistresses are committed to teamwork (Mooney, 2011). They often act as the coordinator and facilitator in the team and motivates their school community to practice collaborative teamwork (Makura, 2012; Perkins et al., 2013). Additionally, these headmistresses were identified as caring and creative school leaders (Coleman, 1996). In this study, the headmistresses were also found to build a good relationship with all the schools' stakeholders such as its students, teachers, parents and NGO (Non Governmental Organization) for the betterment of the school.

Third, the data highlights that headmistresses in this study have vast teaching experience and great knowledge in instructional matters which make them well equipped to lead instructional and curriculum changes in their schools (Arar and Oplatka, 2016; Grogan and Shakeshaft, 2010). In addition, these headmistresses are good communicators. Effective communication has proven to increase staff involvement and commitment in the organization (Manafa, 2018). Particularly, the headmistresses have the capacity to juggle administrative and management duties. They demonstrated dedication and commitment to work extra hours, which is an emblem of a leader's devotion to their school (Southworth, 2002). Therefore, we could argue that the practices of these headmistresses are consistent with women leadership traits in leading the capacity building of their teachers and improving school outcomes.

Limitations and future studies

This study has several limitations. First, data were collected only from thirteen headmistresses from government-funded Tamil vernacular primary schools in Kuala Lumpur. At the time of this study, all government-funded Tamil vernacular primary schools in Kuala Lumpur are led by headmistresses. As such, future studies could include representative data from headmasters in government-funded Tamil vernacular primary schools in other states of Malaysia to enable a comparative study on the nature of their leadership.

Second, the samples in this study were only headmistresses from government-funded Tamil vernacular primary schools. Considerable evidence exists in the literature that suggests school heads are inclined to overestimate their own leadership practices (see Bellibas and Gumus, 2021; Thien et al., 2021). As stated earlier, the findings of this study are indicative rather than definitive. Future studies could gather additional evidence from teachers to provide an objective view of headmistresses' leadership practices in Tamil vernacular primary schools.

Conclusion

In summary, the evidence presented in this article indicates that headmistresses in selected Tamil vernacular primary schools in Malaysia possess some of the characteristics of women leadership. However, some of their leadership traits are dependent upon the teaching and learning process in their schools. The school's success depends upon the leadership enactment of the headmistresses. When the school leader is persistent in innovating better teaching approaches to improve students' outcomes, learners are likely to achieve as the school succeeds (Adams et al., 2020; Adams, 2018; Msila, 2013).

The findings of this study have significant implications for school leadership, school leadership preparation, and practice. The improvement of school performance has less to do with the school leader's gender and more to do with commitment and vision (Msila, 2013). Women leaders are leading their schools in purposeful ways (Aaron, 2020). They adopt a democratic (participative) style of leadership, emphasize capacity building of their teachers and improving school outcomes based on sheer dedication, vision and commitment to work. The Malaysian Ministry of Education would benefit from this knowledge and use this to develop resources, and better support future and new women school leaders in Malaysia.

References

- Abu-Tineh, A., 2013. Leadership effectiveness in Jordanian educational institutions: a comparison of Jordanian female and male leaders. *Educ. Manag. Adm. Leader* 41 (1), 79–94.
- Aaron, T.S., 2020. Black women: perceptions and enactments of leadership. *J. Sch. Leader.* 30 (2), 146–165.
- Adams, D., Cheah, K.S.L., Harris, A., Jones, M., Sumintono, B., Md Yusoff, N.N., 2020. Attributes of successful school leaders. In: Tweeddale, C., Staufenberg, J. (Eds.), *Developing Effective Learning in Nepal: Insights into School Leadership, Teaching Methods and Curriculum*. British Council.
- Adams, D., 2018. *Mastering Theories of Educational Leadership and Management*. University of Malaya Press, Kuala Lumpur.
- Adams, D., Muthiah, V., 2020. School principals and 21st century leadership challenges: a systematic review. *J. Nusantara Stud.* 5 (1), 189–210.
- Arar, K., Oplatka, I., 2016. Current research on Arab female educational leaders' career and leadership: a review of extant literature and future directions for research. In: Bowers, A.J., Shoho, A.R., Barnett, B.G. (Eds.), *Challenges and Opportunities of Educational Leadership Research and Practice*, pp. 87–115.
- Atieno, O.E., 2013. Role performance of female headteachers in primary schools: a case of Bondo District, Kenya. *Adv. Women Leadersh.* 33, 96–105.
- Badura, K.L., Grijalva, E., Newman, D.A., Yan, T.T., Jeon, G., 2018. Gender and leadership emergence: a meta-analysis and explanatory model. *Person. Psychol.* 71 (3), 335–367.
- Bellibas, M.S., Gumus, S., 2021. The effect of learning-centered leadership and teacher trust on teacher professional learning: evidence from a centralised education system. *Prof. Dev. Educ.* 1–13.
- Bissessar, C., 2013. Correlation of Female Trinidadians, Bermudians, Sri-Lankans' feminist identity development, and self-esteem: reflection of social role theory. *Adv. Women Leadersh.* 33, 15–28.
- Brinia, V., 2012. Men vs women; educational leadership in primary schools in Greece: an empirical study. *Int. J. Educ. Manag.* 26 (2), 175–191.
- Coleman, M., 1996. Management style of female headteachers. *Educ. Manag. Adm.* 24 (2), 163–174.
- Crosby-Hillier, K., 2012. Women and educational leadership: exploring the experiences of current and aspiring female educational administrators. *WomenWise* 1, 1–2012.
- Cubillo, L., Brown, M., 2003. Women into educational leadership and management: international differences. *J. Educ. Adm.* 41 (3), 278–291.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Drago-Severson, E., Masling-Ostrowski, P., Blum-Destefano, J., 2018. Looking behind the curtain: principals' internal experiences of managing pressing challenges. *J. Sch. Leader.* 28, 315–343.
- Elmuti, D., Jia, H., Davis, H.H., 2009. Challenges women face in leadership positions and organizational effectiveness: an investigation. *J. Leadersh Educ.* 8 (2), 167–187.

- Esa, A., Leong, C.P., 2013. Women as a leader in school. *Int. J. Sci. Res.* 2 (11), 274–277.
- Glaser, B.G., Strauss, A.L., 2017. *Discovery of Grounded Theory: Strategies for Qualitative Research*. Routledge.
- Grogan, M., Shakeshaft, C., 2010. *Women and Educational Leadership*. John Wiley & Sons.
- Hallinger, P., Lee, M., 2014. Mapping instructional leadership in Thailand: has education reform impacted principal practice? *Educ. Manag. Adm. Leader* 42 (1), 6–29.
- Hallinger, P., Heck, R.H., 2010. Leadership for learning: does collaborative leadership make a difference in school improvement? *Educ. Manag. Adm. Leader* 38 (6), 654–678.
- Harris, A., Jones, M., Adams, D., Cheah, K., 2019. Instructional leadership in Malaysia: a review of the contemporary literature. *Sch. Leader. Manag.* 39 (1), 76–95.
- Harris, A., Jones, M., 2015. International comparisons: critique, culture and context. In: Harris, A., Jones, M. (Eds.), *Leading Futures: Global Perspectives on Educational Leadership*. SAGE, London, pp. 1–10.
- Hotmire, J., 2018. Principals' Role Perception and Implementation of Educational Reform. Retrieved from. https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=bgusu1522678378556588&disposition=inline.
- Husu, L., Hearn, J., Lämsä, A.M., Vanhala, S., 2010. Leadership through the gender lens: women and men in Organisations. In: *Immigrants and Civic Integration in Western Europe*. Retrieved from. <https://helda.helsinki.fi/bitstream/handle/10227/753/71-978-952-232-101-5.pdf>.
- Ibrahim, N., 2018. A case on Tamil education in Malaysia: is Tamil education being sidelined?. In: *SHS Web of Conferences*, vol. 53. EDP Sciences, p. 03006.
- Kaiser, R.B., Wallace, W.T., 2016. Gender bias and substantive differences in ratings of leadership behavior: toward a new narrative. *Consult. Psychol. J. Pract. Res.* 68 (1), 72.
- Kaputa, C., 2009. *The Female Brand: Using the Female Mindset to Succeed in Business*. Nicholas Brealey Publishing.
- Kolehmainen, C., Brennan, M., Filut, A., Isaac, C., Carnes, M., 2014. Afraid of being "witchy with a 'B'": a qualitative study of how gender influences residents' experiences leading cardiopulmonary resuscitation. *Acad. Med.* 89 (9), 1276–1281.
- Kwado Agezo, C., 2010. Female leadership and school effectiveness in junior high schools in Ghana. *J. Educ. Adm.* 48 (6), 689–703.
- Kythreotis, A., Pashiardis, P., Kyriakides, L., 2010. The influence of school leadership styles and culture on students' achievement in Cyprus primary schools. *J. Educ. Adm.* 48, 218–240.
- Leimon, A., Moscovici, F., Goodier, H., 2011. *Coaching Women to Lead*. Routledge, East Sussex.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22.
- Lyman, L.L., Ashby, D.E., Tripses, J.S., 2005. Leaders who dare: pushing the boundaries. *R&L Education*.
- Makura, A.H., 2012. Leadership styles of female educational leaders: in search of a gender inclusive leadership theory. *J. Soc. Sci.* 31 (3), 279–287.
- Manafa, I.F., 2018. Communication skills needed by principals for effective management of secondary schools in Anambra State. *J. Arts Manag. Soc. Sci.* 3 (2).
- Méndez-Morse, S., Murakami, E.T., Byrne-Jiménez, M., Hernandez, F., 2015. Mujeresin the principal's office: latina school leaders. *J. Latinos Educ.* 14 (3), 171–187.
- Mollel, N., Tshabangu, I., 2014. Women in educational leadership: conceptualizing gendered perceptions in Tanzanian schools. *Educ. Res. Int.* 3 (4), 46–54.
- Mooney, J.A., 2011. *Elementary School Teachers' Perceptions of Effective Leadership Practices of Female Principals* (Doctoral dissertation). Retrieved from. <http://dc.etsu.edu/>.
- Msila, V., 2013. Obstacles and opportunities in women school leadership: a literature study. *Int. J. Educ. Sci.* 5 (4), 463–470.
- Nogay, K., Beebe, R.J., 2008. Gender and perceptions: females as secondary principals. *J. Sch. Leader.* 18 (6), 583–602.
- OECD, 2020. *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*. Author, Paris.
- Offermann, L.R., Coats, M.R., 2018. Implicit theories of leadership: stability and change over two decades. *Leader. Q.* 29 (4), 513–522.
- Perkins, S.E., Phillips, K.W., Pearce, N.A., 2013. Ethnic diversity, gender, and national leaders. *J. Int. Aff.* 85–104.
- Prime, J.L., Carter, N.M., Welbourne, T.M., 2009. Women "take care," men "take charge": managers' stereotypic perceptions of women and men leaders. *Psychol. Manag. J.* 12 (1), 25–49.
- Robinson, V.M., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44 (5), 635–674.
- Shakeshaft, C., 1989. *Women in Educational Administration*. Sage Publications.
- Southworth, G., 2002. Instructional leadership in school: reflection and empirical evidence. *Sch. Leader. Manag.* 22 (1), 73–91.
- Stead, V., Elliott, C., 2009. *Women's Leadership*. Palgrave Macmillan, London.
- Strauss, A., Corbin, J., 1998. *Basics of Qualitative Research Techniques*. Sage publications, Thousand Oaks, CA.
- Strauss, A.L., 1987. *Qualitative Analysis for Social Scientists*. Cambridge university press.
- Sun, J., Leithwood, K., 2015. Direction-setting school leadership practices: a meta-analytical review of evidence about their influence. *Sch. Effect. Sch. Improv.* 26 (4), 499–523.
- Thien, L.M., Liu, S., Yee, L.Q., Adams, D., 2021. Investigating a multiple mediated-effects model of instructional leadership and teacher professional learning in the Malaysian School Context: a partial least squares analysis. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/17411432211009892>.
- Valerio, A.M., 2009. *Developing Women Leaders: A Guide for Men and Women in Organizations*. Wiley Blackwell, Oxford.
- Walker, A., 2015. Clones, drones and dragons: ongoing uncertainties around school leader development. *Sch. Leader. Manag.* 35 (3), 300–320.
- Wilen-Daugenti, T., Vien, C.L., Molina-Ray, C., 2013. *Women Lead: Career Perspectives from Workplace Leaders*. Peter Lang Publishing, Inc, New York.
- Yin, R.K., 2013. Validity and generalization in future case study evaluations. *Evaluation* 19 (3), 321–332.

Successful school principalship in Thailand: a literature review

Dhirapat Kulophas^a and Philip Hallinger^b, ^a Department of Educational Policy, Management, and Leadership, Chulalongkorn University, Bangkok, Thailand; and ^b College of Management, Mahidol University, Department of Educational Leadership and Management, University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	490
Successful school principalship: global perspectives	491
Thailand context	491
Education context in Thailand	491
Cultural context	492
Review method	493
Identification of sources	493
Findings	493
What is successful school principalship in the Thai literature?	493
Successful principal practices in Thailand	496
Leadership vision	496
Decision-making and problem-solving	496
Leading teaching and learning	496
Communication and conflict management	496
Developing the school's culture	497
Discussion	497
Conclusion	498
Acknowledgment	499
References	499

Introduction

Over the past 20 years, an extensive body of literature on educational leadership has found that the quality of school leadership has significant, albeit largely indirect effects on student learning outcomes (Hallinger, 2011; Leithwood et al., 2004; Louis et al., 2010). Although the school effectiveness research has emphasized student learning progress and attainment as key criteria for determining leader effectiveness, Day and Sammons (2016) asserted that student achievement alone should not be used as the measure of "school success". They contend that "successful schools" also embed positive social values among learners, enhance skills and attitudes that support lifelong learning, provide a foundation for successful citizenship, and support other social outcomes that meet the requirements of education in 21st society. This revised assumption has been the foundation of a global knowledge base developed under the auspices of the *International Successful School Principalship Project* (ISSPP) (e.g., Day et al., 2016; Leithwood, 2005; Mulford and Silins, 2011). At present, this worldwide study has contributed more than 100 multi-site case studies that elicit principal perspectives and practices associated with successful schools from more than 20 nations (Gurr, 2017).

The ISSPP studies find that successful school principals share many similar values, characteristics, and accomplishments across these diverse settings. In addition, the findings also reinforce assertions that the strategies which principals use to support school success are responsive to the external contexts in which they work (Day and Sammons, 2016). While this research represents the most extensive global comparative project on school leadership, it is largely comprised of research conducted in Western and English-speaking societies. Inclusion of research from emergent regions of the world may lead to a more diversified and refined view of successful school principalship (Walker and Hallinger, 2015).

Thailand, the second most populated nation in Southeast Asia, has a distinct geographical, political, and cultural context that has shaped the development of its education system (Fry and Bi, 2013). This creates a unique set of institutional and cultural forces that shape both the role definition and enactment of the principalship in Thailand (Hallinger, 2004; Kulophas and Hallinger, 2021). Thus, any discussion of "successful principalship" in the Thai context must begin with a consideration of how the principalship has evolved and the expectations that different stakeholders have for the educators holding this role. Only then, can we begin to assess the nature of practices that enable principals to meet those expectations. Thus, this review chapter addresses two research questions:

1. How is successful school principalship defined, explicitly and implicitly, in the Thai language literature on school leadership and management?
2. What constitute successful school principalship practices in Thailand?

Our approach to addressing the first research question seeks to elicit perspectives on “successful school principalship” in the Thai language literature on educational leadership and management. Following this, we review studies that have studied the practices of principals who have been characterized as “successful” based on locally-adopted criteria. Drawing upon this literature, we synthesize practices that Thai scholars suggest are associated with “successful principalship” in Thailand. In both sections of the paper, the discussion is placed within the political, societal, and cultural context of Thai education.

Successful school principalship: global perspectives

Launched in 2001, the International Successful School Principalship Project (ISSPP) has aimed to build a global body of knowledge on how principal practices and characteristics contribute to school success. The founding members came from eight nations: Australia, Canada, China, Denmark, England, Norway, Sweden, United States. At the outset, they decided to conduct a series of case studies that would investigate successful principal practices in their respective countries. They agreed to use the term “successful” instead of “effective” to incorporate not only student learning outcomes, but also the development of positive values, life-long learning, and productive citizenship behaviors (Day and Gurr, 2018; Day and Sammons, 2016). The selected “successful” schools and principals in these studies mostly demonstrated one or more of the following criteria: higher-than-expected student achievement on national tests where available, principals’ excellent recognition within the community or school system as acknowledged by peers or government officials, and other context-specific notions of success such as awards for school management, principalship or excellence projects (Gurr, 2015).

The global research associated with the ISSPP project has collectively contributed to a deeper understanding of how school leadership contributes to school success. Evidence from the international case studies echoed four common features of successful school principalship: establishing moral purpose and creating a positive school culture, developing teacher professional competence, restructuring the school organization and sharing leadership, and redesigning curriculum and enhancing teaching and learning (Leithwood and Day, 2007). Moreover, successful school principals strategically adapted their leadership practices to suit the environment in which they work (Day and Sammons, 2013).

While the current body of knowledge that has accumulated through the ISSPP project is derived from a pool of over 20 nations, most of the studies have come from Anglo-American and European societies. Thus, despite being one of the largest and most culturally diverse regions of the world, so far only three studies have been conducted in Asian societies: China (Wong, 2005), Indonesia (Raihani, 2008), and Singapore (Gurr et al., 2016). Thus, this review seeks to add further to the diversification of international knowledge on successful principalship.

Thailand context

As noted above, a fundamental assumption of the ISSPP project is that successful school principalship is contextualized (Day and Sammons, 2016). This recognizes that the practices enacted by principals is shaped by expectations embedded in the cultural, political, and institutional systems where their schools are located (Bajunid, 1996; Belchetz and Leithwood, 2007; Hallinger and Leithwood, 1996; Hallinger and Thang, 2014; Walker and Hallinger, 2015).

Education context in Thailand

Driven by the 1997 Asian economic crisis, Thailand undertook its most significant attempt at educational reform by enacting the 1999 National Education Act (NEA). This ambitious platform of reforms aimed to transform teaching and learning and overhaul the educational management system to ensure sustainable economic and social development (Fry and Bi, 2013). Despite its grand ambitions, scholars have characterized the actual reform implementation as slow, stagnant, and unsustainable (Fry and Bi, 2013; Fry and Sangnapaboworn, 2018; Hallinger and Lee, 2011). While the MOE made some progress in restructuring its administrative organization, widespread, significant improvements in the quality of teaching and learning in schools have not been realized (Fry and Bi, 2013; Hallinger and Lee, 2011; Sangnapaboworn, 2018).

For example, in 2017 an OECD report found that despite spending more time in classroom learning, Thai students still performed poorly when compared on multinational academic assessments (OECD, 2017). According to Fry and Bi (2013), classroom teaching in Thailand continues to stress teacher-centered approaches that rely on memorization and rote-learning. Active learning which engages students, enhances their critical thinking, and develops creativity have yet to gain a firm foothold among a critical mass of teachers (Hallinger and Lee, 2011).

Several system-level factors have contributed to the stagnation of Thailand’s educational reform (Fry, 2018). These include over-centralization, political interference, lack of commitment to a reform strategy, misalignment among main educational agencies, poor policy implementation, and insufficient investment in human development (Atagi, 2011; Fry and Bi, 2013; Fry and Sangnapaboworn, 2018). For example, since 1999, the MOE has been led by 21 different education ministers, averaging nine months per minister in the position. This has resulted in frequently shifting policies and eroding confidence among educators and the public. Although the NEA began to address the need for decentralization, in practice the education system remains highly centralized, particularly in the domains of finance and human resource management (Atagi, 2011; Fry and Bi, 2013).

While Thailand has invested approximately 20% of its annual fiscal budget on education (The World Bank, 2017), access to quality education remains unequal and limited among underprivileged groups of society (Fry and Sangnapaboworn, 2018; OECD/UNESCO, 2016). Moreover, due to a declining birth rate, the Thai government has sought to curb spending by reducing the recruitment of new civil servants, including teachers. This downsizing policy has caused serious teacher shortages, especially in critical subjects among rural and disadvantaged schools (Atagi, 2011).

The fast changing, competitive world economy has increased the urgency for the Thai government to transform its education system. In 2017, the current government adopted a new long-term policy called “Thailand 4.0”, which aspires to transform the economy and enable the nation to escape the “middle income trap” (Royal Thai Embassy, 2017). This policy reaffirmed the vital role of education to prepare new generations of citizens capable of meeting 21st century challenges. Ken Kay, President, Partnership for 21st Century Skills (AksornAct, 2017), asserted that Thailand 4.0 would not be possible without first developing “Educator 4.0”.

Thailand 4.0 has undoubtedly intensified the already complex role of Thai principals. Scholars have asserted that Thai school administrators lack essential competences to lead changes in teaching and learning (Hallinger, 2018c; Hallinger and Lee, 2013a). Thai principals were conventionally expected to perform administrative tasks such as resource planning, scheduling, budget management, facility management and other routine management procedures. While the concept of instructional leadership first entered the lexicon of school leadership in Thailand in the early 2000’s, only a small number of Thai principals have actively adopted this perspective on their role (Hallinger and Lee, 2013b). This has made the sustained implementation of educational reforms even more challenging (Fullan, 2002).

These continuing limitations in the capacity of Thai principals to lead and support educational reforms has been linked to systemic weaknesses (Fry and Sangnapaboworn, 2018; Hallinger, 2018b). Thai principals are enlisted by self-selection and vetted through formalistic, rules-oriented examinations (Chiangkul, 2016). In addition, the systems for preparing and developing school principals remain geared toward management (e.g., scheduling, facilities, budget, personnel) and administration (e.g., rules, processes, and policies), rather than “leadership” for teaching and learning. Thus, scholars and policymakers have concluded that Thai principals have neither been selected nor trained with the perspectives and skills needed to bridge the gap between reform policies and classroom practices (Hallinger, 2018b; OECD/UNESCO, 2016; Varavarn, 2011).

Cultural context

How Thai school principals approach their role is not only influenced by national laws and policies, but also by the cultural values and norms of Thai society (Bush, 2012; Hallinger, 2018a). It is common knowledge among Thai citizens that the three pillars of Thai society are, as represented by the three colors of Thai flag, nation (red), religion (white), and the king (blue). More than 96% of Thai citizens are Buddhist. Therefore, to understand Thai culture, it is vital to recognize how Buddhism has influenced Thai ways of living (Fry, 2018).

One of the most important teachings of Buddhism is *ariyasat sii* (four noble truths). The *ariyasat sii* state that the root cause of human suffering is craving. To be liberated from a cycle of suffering, one should practice the noble eightfold path or “the middle way”, which leads to nirvana or freedom from attachment (Tsering, 2005). Buddhist influence is evident in the redefinition of national educational goals in 1999. The NEA stated explicitly three goals that graduates of the Thai education system should possess: *geng* (capability), *dii* (morality), and *kwarm sook* (satisfaction or happiness). Thus, even when faced with the need to respond to the economic crisis of 1997, national policymakers were unwilling to sacrifice the broader mission of education to the sole purpose of academic learning.

Thailand’s late King Rama the IX developed the Sufficiency Economy Philosophy (SEP), a set of principles for living sustainably as individuals and a society. This, represents a good example of how the value of “the middle way” is widely recognized and applied in Thai society (Piboolsravut, 2004). Consistent with “the middle way” concept, the SEP asserts that, even in a fast changing and competitive world, one should live a balanced life, working and consuming moderately, responsibly, and self-consciously.

For the past several decades, schools across the country have sought to integrate the SEP in curriculum and school management. Over the past decade, the Virtuous Schools Project, supported by the King Rama IX Foundation, implemented a model of SEP schooling in more than 20,000 schools across Thailand (Dharmapiya and Saratun, 2020; Hallinger and Suriyankietkaew, 2018). The philosophy underlying this ambitious project stated that these schools should graduate students who have developed the “enduring virtues” of moderation, prudence, generosity, and cooperation. Virtuous Schools reorient their operations to meet goals related to economic development (e.g., academic and personal life skills), environmental awareness, social justice, and cultural preservation (Dharmapiya and Saratun, 2020).

Another distinguished Buddhist doctrine is called *phrom wihan sii* (four sublime attitudes). These comprise *metta* (kindness), *karuna* (compassion), *mudita* (vicarious joy), and *ubekkhā* (equanimity). *Phrom wihan sii* underlines the importance of giving, merit, tolerance, and calmness instilled in Thai culture. It is also considered an essential principle for managing schools, where principals are expected to exhibit generosity and mercy toward teachers (Rukbanglaem et al., 2015; Srichandari and Thanaphaet, 2020).

Hofstede’s cultural dimensions offer an additional lens through which to analyze the “comparative” features of Thai culture. According to Hofstede et al. (2005), the main features of Thai culture include high power distance, collectivism, and femininity, and uncertainty avoidance. As a high power distance culture, Thai people accept that power and status are distributed unevenly in society. Thai subordinates are inclined to obey their supervisors’ orders in return for security and support. A Thai custom called *kreng jai* (fearful heart) can be observed in schools where Thai teachers dutifully follow the principal’s instructions despite suppressed disagreement. This culturally ingrained norm of compliance with authority engenders a culture of passive resistance, low

commitment, and corruption throughout all levels of the Thai education and governmental system (Fry, 2018; Vichit-Vadakan, 2017).

Thailand is also a highly collectivist and feminist culture. Honoring tradition, maintaining harmony, and avoiding conflicts are essential in a collective society (Hofstede et al., 2005). In staff meetings, Thai teachers seldom ask questions or make comments for fear of creating tensions with the principal or among colleagues. In a highly feminist culture, holistic well-being, caring for others, kindness, and personal relationships are treasured. These values are encompassed in an important Thai norm of *tum bhun* or making merit by doing good deeds. This explains why, despite their low pay, Thai teachers are often motivated to work longer hours for disadvantaged students in the absence of any financial gains. Similarly, a strong need to cultivate personal relationships offers insight into why Thai principals normally spend extended hours for off-school responsibilities such as attending communal rituals and ceremonies (e.g., births, deaths, holiday ceremonies, temple ritual) which have nothing directly to do with their “in-school roles”.

Lastly, based on Hofstede et al. (2005), Thai people are generally risk averse. Moral purpose must be communicated and understood by the group before any change is initiated. To reduce uncertainty and accelerate reform, oftentimes Thai educational authorities prioritize rules, regulations, and policies. For example, professional learning communities (PLCs) were not extensively practiced in schools until MOE formally commissioned as the main professional development strategy in 2017 (Prabjandee, 2019).

Review method

Identification of sources

In order to explore the meaning of successful principalship in Thailand, we turned to relevant Thai language literature published in Thailand. Our rationale for focusing on Thai language publications was two-fold. First, as noted earlier, relatively few articles on educational leadership and management in Thailand published in the international literature have extensively referenced Thai language literature. Second, we believed that the subset of the Thai literature published in national Thai language journals might capture a somewhat different complementary perspective on the meaning of successful school principalship.

We set the start year for our search at 1999 due to its historic significance as the year the NEA was enacted. The search was conducted in the Thai-Journal Citation Index (TCI) which is the primary digital archive of academic articles published in Thai journals. Due to its formal role in determining journal impact and ranking, TCI is considered the only official journal indexing authority of Thailand.

We started by using the key words in Thai language for “success AND administrator AND school” (ความสำเร็จ ผู้บริหาร สถานศึกษา) to search for related articles. In Thai, the term “successful” could also refer to “famous” (มีชื่อเสียง) or “recognized” (ได้รับการยอมรับ). Therefore, two additional searches were conducted using the key word “famous AND administrator AND school” and “recognized AND administrator AND school”. After removing duplications, a total of 366 articles were identified. “Administrator” was used due its common usage in the Thai language when referring to school leaders, managers, or principals.

We screened titles and abstracts to eliminate ineligible articles based on the following exclusion criteria: (1) lack of focus on education; (2) focus on vocational or higher education administration; (3) insufficient focus on school administrators. According to Dixon-Woods et al. (2006), an interpretive review should prioritize “relevance” over “type of publication”. This informed our decision to include both empirical and non-empirical articles, since both could offer insight into “Thai perspectives on successful principalship”. After this screening process, a total of 32 journal articles remained (see Fig. 1).

We extracted document information such as author, year of publication, education level, participant, and methodology, and categorized these as document variables. Next, we created an analytic framework based on the existing ISSPP literature by specifying common characteristics and practices of successful school principalship. While this analytic framework was used as an initial guide for our coding, we added new codes and combined existing codes with new ones that emerged during analysis. We used MAXQDA software version 11.0 to assist in coding text and recategorizing codes during our data analysis.

The 32 eligible publications were comprised of 23 empirical and nine non-empirical articles (see Table 1). Of 23 research articles, 15 were quantitative studies including 12 that featured hypothesis testing. The other eight were either single-case or multi-site qualitative case studies. We categorized the nine non-empirical articles as “perspective” since they largely offered opinions on the characteristics school principals “should” possess and/or what practices they “should” demonstrate to be successful.

Findings

This section of the chapter seeks first to elicit a Thai perspective on the meaning of “successful school principalship” from the Thai literature in educational leadership and management. Then we examine practices that flow from this perspective.

What is successful school principalship in the Thai literature?

In the qualitative studies, most authors defined successful school leaders as someone who had been recognized as exemplary in school management, received a national or regional award, or obtained high professional rank. Principals whose schools received excellent ratings or certification by local school districts or the Office for National Education Standards and Quality Assessment

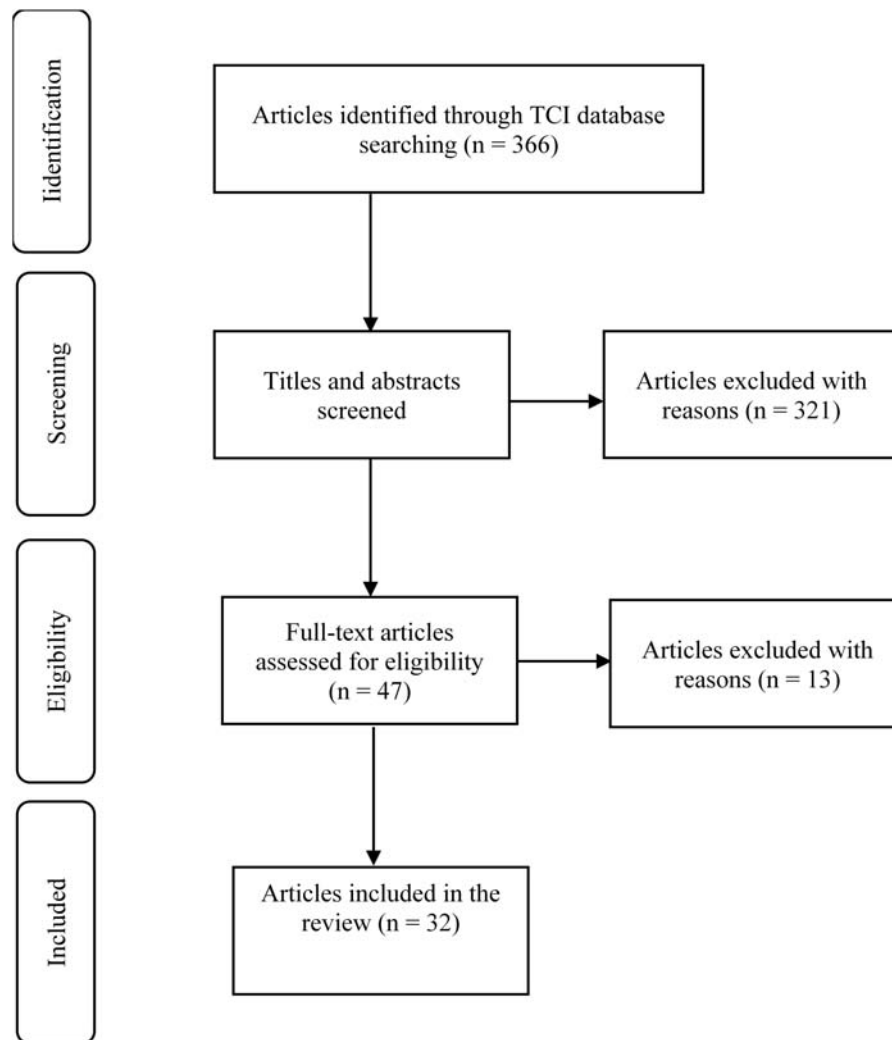


Fig. 1 Process for searching and inclusion of the articles.

(ONESQA) were identified in six qualitative articles as “successful principals” (e.g., Piungam et al., 2015; Pruetikul and Anannawee, 2014; Sankaew and Chusorn, 2016). All Thai schools must undergo educational quality audits by the ONESQA (2011) every five years. Schools that achieve ratings of “good” or “excellent” quality typically display this certification in front of the school.

Some authors described exemplary schools such as those selected by the MOE or school districts to lead high-profile academic initiatives, most notably under the auspices of the “world-class standard school” (WCSS) project (Issara and Thee-asana, 2017; Somprach, 2017). This initiative was launched in 2010 to raise the quality level of Thai education to international standards and develop students into capable global citizens. The program inspected and certified a selection of secondary schools nationwide that demonstrated outstanding improvement and innovation in curriculum, teaching and learning, and school management.

It is, however, worth noting that WCSS schools tend to be urban, advantaged, and highly selective. Thus, “success” in this program may have more to do with student composition, school culture and the environment than the principal’s practices. Moreover, due to frequent principal turnover in the Thai education system, a principal could be recognized despite having little to do with the school’s recent success.

Principals’ excellence in school administrative tasks was frequently mentioned in perspective articles as an indicator of successful school principalship. This reflects the “rules-based orientation” of school administrators within all levels of the Thai education system (Fry, 2018; Yuenyong, 2019). Excellence was most frequently determined by recommendations from school district officials without any systematic approach to assessing how principals’ performance had impacted their schools.

Receipt of national or regional awards are another prominent indicator of successful principalship. For example, to honor principals for their exemplary professional attainments, the Teacher’s Council of Thailand (TCT) and the Office of the Basic Education Commission (OBEC) present national awards on an annual basis. Accolades for school management featured in five entries as an important qualification for case selection of “successful principals” (Issara and Thee-asana, 2017; Kukkong and Eungpuang, 2014; Pruetikul and Anannawee, 2014; Sankaew and Chusorn, 2016; Thetarat and Sripokangkul, 2018).

Table 1 Characteristics of articles included in the review.

	Document	Empirical			Non-emp	Scope of study	
		Quant	Hyp test	Qual		Region	Level
1	Akarachan et al. (2018)	✓	✓			NE	Pri
2	Apiwatanasingha et al. (2008)	✓	✓			Nor	Gen
3	Banlue and Tiacharoen (2018)	✓	✓			–	Sec
4	Bulan (2017)				✓	–	Gen
5	Chaita and Anannawee (2009)	✓	✓			–	Pri
6	Chaiwong and Sanghiran (2017)	✓	✓			–	Gen
7	Chaiyakit et al. (2020)				✓	–	Gen
8	Chantanontree and Thongpan (2017)				✓	–	Gen
9	Charoenpornsakul et al. (2018)	✓	✓			–	Gen
10	Duangtongkul et al. (2016)	✓	✓			Sou	Pri
11	Issara and Thee-asana (2017)			✓		NE	Gen
12	Khemkai (2016)	✓	✓			NE	Pri
13	Kukkong and Eungpuang (2014)			✓		NE	Sec
14	Muktarakosa et al. (2018)				✓	–	Gen
15	Pechsangkoon and Thanaphaet (2020)				✓	–	Gen
16	Phochsalee et al. (2016)	✓	✓			NE	Gen
17	Phutthasen et al. (2016)	✓	✓			–	Sec
18	Piungam et al. (2015)			✓		NE	Pri
19	Prasongsuk et al. (2015)	✓				NE	Gen
20	Promkan et al. (2013)	✓	✓			NE	Sec
21	Pruetikul and Anannawee (2014)			✓		Cen	Sec
22	Raksanakorn (2020)				✓	–	Gen
23	Rukbanglaem et al. (2015)				✓	NE	Gen
24	Sankaew and Chusorn (2016)			✓		NE	Sec
25	Somboon and Kornpuang (2018)			✓		–	Pri
26	Somprach (2017)			✓		–	Gen
27	Sreebuaand and Thanaphaet (2020)				✓	–	Gen
28	Srichandari and Thanaphaet (2020)				✓	–	Gen
29	Suk-ied et al. (2015)	✓	✓			Sou	Sec
30	Sukthet et al. (2020)	✓				NE	Pri
31	Thetarat and Sripokangkul (2018)			✓		NE	Gen
32	Youngtawee and Charoenkul (2015)	✓				Cen	Sec
Total		15	12	8	9		

Nor = North; Sou = South; Cen = Central; NE = Northeast.

Pri = Primary; Sec = Secondary; Gen = General.

Five quantitative studies purposively set the scope of their research on schools that had achieved awards. Most studies selected schools that were granted “Royal School Award”, a highly prestigious award inaugurated and bestowed by the King Rama IX since 1963 (OBEC, 2017). The Royal School Awards, widely considered the epitome of success for any Thai school principal, are given to winning schools from each region and category according to academic level and school size.

Several qualitative studies used “professional rank” as a criterion for identifying successful principals. Academic status (*wittayathana*) is a professional ranking system for teachers and school administrators governed by the Office of the Teacher Civil Service and Educational Personnel Commission (OTEPC). The current professional ranking system for school principals in Thailand starts from level 1 (ordinary) to level 5 (expert).

Two qualitative studies (Kukkong and Eungpuang, 2014; Somprach, 2017) used *wittayathana* level 4 (expertize) as a selection criteria for successful principalship. It should be emphasized that progression within this ranking system results from “ticking the boxes” on measurable bureaucratic criteria such as years of experience, educational level attained, attendance at a series of Ministry-sponsored workshops, receipt of school awards, and completion of a school-based research project. Moreover, the link between acquiring awards and ranking has created a scheming of “scratching each other’s back” within a school district. It is common knowledge that awards can be given to schools or principals based on skeptical criteria to facilitate qualification for rank promotion.

To further explore the meaning of successful school principalship, quantitative studies were further examined with a focus on the operational definition of the dependent variable used to define “success”. We found that many authors used the terms “school success” and “school effectiveness” interchangeably. While “school effectiveness” was used frequently in the studies, the authors seldom applied the type of student learning criteria typically sought in the international literature on “school effectiveness” (Schrems, 2000). For example, as noted above, awards and ranking criteria seldom prioritized equity in student learning opportunities or outcomes. Instead, “school effectiveness” was treated as a catch-all term which encompassed the ability to solve school

administrative problems, use of “best practices” in school administration (e.g., technology), or school knowledge management (Duangtongkul et al., 2016; Phutthasen et al., 2016; Promkan et al., 2013; Suk-ied et al., 2015).

Successful principal practices in Thailand

Leadership vision

The authors of these articles largely concurred that successful principals had clear, long-term visions, which were used to set future direction and guide strategic planning for the school (e.g., Issara and Thee-asana, 2017; Kukkong and Eungpuang, 2014; Piungam et al., 2015; Somprach, 2017). Although Thailand is a highly hierarchical society and school principals are expected to follow the system directives, some authors observed that “successful school principals” had their own personal visions for change (Bulan, 2017; Chaityakit et al., 2020; Raksanakorn, 2020). Bulan (2017), for example, asserted that Thai principals cannot afford to be complacent, relying on “Thai-style” school management by simply reacting to orders from higher-ups and following daily management routines. They should proactively create and implement strategic plans for school improvement, following up and reporting to superiors on what they have achieved and what resources are needed (Bulan, 2017). This quotation highlights the fact that this “leadership” approach is not, however, the norm among Thai principals more broadly.

Decision-making and problem-solving

Several authors found that successful school leaders studied rules and regulations carefully, looked for rules that could leverage their own strategies, gathered information from stakeholders, and engaged them in a shared decision-making process (e.g., Kukkong and Eungpuang, 2014; Pruetikul and Anannawee, 2014; Sankaew and Chusorn, 2016). Two features stand out here. First, going beyond “knowing the rules” to adapting and leveraging them for local empowerment. Second, successful principals frequently sought input from teachers and school board members before enforcing any actions. These practices not only ensure legal compliance, but also responsiveness and acceptance from the parties responsible for support and implementation (Kukkong and Eungpuang, 2014; Thetarat and Sripokangkul, 2018).

Successful principals displayed self-efficacy and flexibility in dealing with challenges (Sankaew and Chusorn, 2016), learned from others and tried to experiment in their own schools (Somprach, 2017). Integration of digital technology in school management was cited by several authors as an example of how successful principals demonstrated continuous learning and adaptability for the benefit of their schools. For instance, Somprach (2017) found that technology integration supported the use of data-driven decision making, use of learning technologies by teachers, and enhancing within-school and school-home communications.

Leading teaching and learning

Managing curriculum and pedagogy was noted as a priority among successful school principals. One multi-site case study depicted three exemplary principals who engaged proactively in pedagogical activities at their schools (Piungam et al., 2015). They prioritized students’ learning by designing the school curriculum around the community context, planning for teacher supervision and evaluation, promoting life-long learning, and encouraging teachers to try new teaching approaches. In another study, a Buddhist school principal led teachers to change their routine practices by collaborating on a learning innovation program called *jittasuksa* or spiritual study (Somboon and Kornpuang, 2018). In this project, the principal acted as a spiritual coach encouraging teachers to reflect on and refine personal professional development and practices.

Another case focused on a principal’s efforts to lead a school that was participating in the Virtuous Schools Project with the aims of instilling self-reliance and a sustainability mindset in teachers, young children, and community members (Thetarat and Sripokangkul, 2018). The school principal worked with teachers and the community to develop a more relevant school curriculum and support the use of more active teaching and learning activities. He also carefully selected and trained teachers and made memoranda of agreement with provincial teacher colleges to lend instructors for scientific subjects with high teacher shortage.

Communication and conflict management

Another key practice for successful school principalship in Thailand concerns communicating creatively and positively. According to Pruetikul and Anannawee (2014), successful principals displayed high assertiveness while remaining approachable for teachers and parents. They chose words wisely and carefully when interacting with others, especially when giving teachers feedback. This encouraged teachers to have the courage to come forward and ask for professional or personal advice. This practice was supported by several opinion pieces (e.g., Chantanontree and Thongpan, 2017; Rukbanglaem et al., 2015; Srichandari and Thanaphaet, 2020) which asserted that school leaders should use kind words (*piyavaca*) to foster a climate of friendship (*kanlayanamit*) among school staff.

Moreover, demonstrating *piyavaca* and enabling *kanlayanamit* were highlighted as essential in managing conflicts effectively in schools. For example, Somboon and Kornpuang (2018) reported that in one case school it was common for tensions and disagreements to arise among teachers in professional learning community (PLC) meetings. The principal sought to manage conflicts productively by asking constructive questions, offering positive feedback, prompting reflection, and facilitating teachers’ engagement. Thetarat and Sripokangkul (2018) used a Thai proverb to convey the essence of a successful principal’s approach in managing conflicts: “neither let a lotus get bruised nor let water get turbid” (*bua mai hai cham naam mai hai kun*). In a collectivist culture where conflicts are often swept under the rug, this suggests a capacity for negotiating a compromise during disputes. This allows the principal to fulfill the goals of maintaining harmony while moving forward on the school’s agenda.

Developing the school's culture

The studies further highlighted practices used by successful school principals to support teacher professional growth and wellbeing, and a collaborative school culture. Muktarakosa et al. (2018) opined that, to succeed in the Thailand 4.0 era, school principals must develop teacher professional capacity to the fullest. Pechsangkoon and Thanaphaet (2020) also suggested that school principals should both encourage and advise teachers to climb up in the professional ranking system.

In the past, most Thai principals traditionally relied on MOE to arrange off-site teacher training sessions. This train-the-trainer model of teacher professional development required teachers to leave classroom for training that was often irrelevant to their needs (OECD/UNESCO, 2016). The nationwide PLC initiative, introduced by the MOE in 2015, gained tremendous support from many of the successful principals described in the qualitative studies (e.g., Piungam et al., 2015; Somboon and Kornpuang, 2018; Somprach, 2017). For example, the principal of a large provincial school received national recognition for his use of “lesson study” as a tool for teacher professional learning. Another principal formed a professional learning network with other schools, which led to an annual fair that showcased teaching and learning innovations by affiliated teachers (Somboon and Kornpuang, 2018). These examples highlight how the selected principals went “beyond compliance” to make the new PLC policy work for their teachers and schools.

One of the major tasks of successful school principals was building relationships in the school's community. Indeed, the quantitative studies emphasized community engagement as a significant factor underlying the success of principals (e.g., Banlue and Tiacharoen, 2018; Duangtongkul et al., 2016; Prasongsuk et al., 2015). Successful school principals strategically used their attendance at “culturally mandated” events in the community (e.g., funerals, weddings, holiday ceremonies) to strengthen relationships with stakeholders (e.g., parents, business leaders, alumni, monks). Investing time in relationship building was essential to obtaining the auxiliary financial, human, and materials resources needed to advance the school mission (Kukkong and Eungpuang, 2014; Piungam et al., 2015; Sankaew and Chusorn, 2016; Somprach, 2017).

One study described a path to success of a female principal, who credited community engagement as a major factor (Kukkong and Eungpuang, 2014). Throughout her career as a principal, she had applied to take part in a series of demanding school initiatives which were used as “catalysts” to energize her teachers and support school development. Participation in these initiatives imposed a higher level of public accountability in academic progress and teaching and learning standards. Realizing that participation came with both threats and opportunities, she invited a network of diverse stakeholders (i.e., teachers, parents, alumni, temples, private businesses, politicians, local authorities) to actively engage in school improvement efforts. This strong community engagement allowed the school to gain financial and public support, access cutting-edge know-how, and expedite administrative procedures.

Community engagement is rooted in Buddhist belief that right behavior is grounded in the local context (*parisannutà*) (Chanantree and Thongpan, 2017). Indeed, *parisannutà* is one of seven *sappurisa dhamma* or Buddhist principles for public administration (Sutamchai et al., 2020). Thus, proactive engagement with the community would help principals learn how to conduct themselves according to local cultural norms and traditions, thereby gradually gaining acceptance and support from the community.

Lastly, following the concept of mindfulness (Sutamchai et al., 2020), several successful principals consciously modeled values and norms through their own behavior (e.g., Prasongsuk et al., 2015; Promkan et al., 2013). For example, several principals sought to inspire to change by demonstrating ethical behavior and decision making, dressing professionally, and speaking politely (*piya-vaca*). This enabled them to gain respect and trust from peers and subordinates (Charoenpornsakul et al., 2018; Prasongsuk et al., 2015). They also acted as role models by devoting their time to worthy public causes (Duangtongkul et al., 2016).

In several qualitative studies, successful principals modeled themselves as learning leaders, sharing their own learning and growth with teachers. This gave principals more credibility when interacting with teachers, especially after the following launch of the PLC initiative. It also enabled them to develop new perspectives and skills and adapt to meet the changing demands of their profession and schools (Somprach, 2017).

One primary school principal shared his story of growing up as a *dek wat* (temple boy) in his youth to obtaining a doctoral degree in educational administration from a reputable university. In Thailand, it is not uncommon for a poor family to send their young son to become a temple boy to receive better education and opportunities (Fry and Sangnapaboworn, 2018). This principal credited his desire to learn by regularly assessing himself and finding opportunities to study from seasoned mentors.

Discussion

This review set out first to develop a deeper understanding of what it means to be a “successful school principal” in Thailand. In this regard, we feel that we have achieved partial success. Our findings suggest that “successful school principals” in Thailand are largely identified by their perceived excellence in wide-ranging school management tasks, awards gained by their schools, and the attainment of high bureaucratic rank. These indicators reflect Thailand's hierarchical culture in which stakeholders generally demonstrate respect and loyalty for people who have achieved recognition (e.g., awards) or status (e.g., rank), regardless of the basis. In a society where principals are government officials, attaining bureaucratic rank is essential to gaining power and social acceptance (Hallinger, 2004; Jirapornkul and Yolles, 2010). Thus, a school principal's merit is proven by the awards and rank attained.

While these criteria are not inconsistent with findings reported in the broader ISSPP literature, in the Thai context neither awards nor rank tend to be based upon data that describe the principal's contributions to the school's performance. Instead, the school's performance is frequently used as an indicator of the principal's performance. Similarly, we suggest that the attainment of

bureaucratic rank is more of an exercise in “ticking the boxes” than an indication of positive impact or substantive achievements by principals in their schools. This approach to using school awards and bureaucratic criteria to define successful principalship in Thailand reflects dominant values and norms of the nation’s culture. For example, Redmond (1998) observed that “face” or outward appearance central to understanding social behavior in Thailand.

Whereas Western power is held by “heads” ... Thai leaders are *hua na*—“heads in front,” [*hua* meaning head and] front being rendered by the word *na*, which means “face” ... the facial aspect of being in front indicates that leaders are meant to conform to the central values of Thai culture, namely, facial values. Leaders are neither born nor made in Thailand; like faces, they are made up. (Redmond, 1998, no page number)

This emphasis on face explains why school awards and rank play such a large role in determining who is perceived as “successful” in the Thai education context. It also explains why principals work so hard to rise in bureaucratic rank. Principals with higher rank are entitled to wear more elaborate uniforms and become eligible for placement at “better schools”. The fact that higher ranked principals tend to gain access to “better schools” reinforces the corollary trend whereby principals from award-winning schools are perceived as successful.

Our findings further suggest that the practices of successful school principals in Thailand are broadly consistent with themes identified from the ISSPP project. Indeed, many of the practices identified in this review are informed by well-known Western leadership principles. These include vision and direction setting, stakeholder engagement, capacity building, conflict management, decision-making, fostering a collaborative culture, and to a lesser degree instructional leadership.

Yet, this review has also highlighted significant ways in which these globally disseminated school leadership and management practices enacted by successful principals are “responsive to the institutional and cultural context” of Thai schools. Successful principals leverage the requirements of a highly centralized, rules-based education system. They find opportunities in the “mortar that lies between the bricks” in a tall, rigid, bureaucratic structure. Well-honed communication and conflict management skills are essential in a collectivist and feminist culture such as Thailand. Successful principals reportedly demonstrate the ability to facilitate open discussion, give constructive feedback, and manage conflicts among teachers. Notably, they used these skills to gain the most from reform policies such as professional learning communities which seek to bring principals into closer contact with teachers, and teachers with one another.

Successful Thai principals strategically utilized their time out-of-school to form strong networks that bring in fiscal support, knowledge, and cooperation from outside the school. They realize that gaining support from external stakeholders is key to achieving their visions for change. This emphasis on engaging the community concurrently serves both the cultural expectation for the *kru yai* (head teacher) to be a “visible face” in the community and the practical purpose of gaining additional resources in an underfunded education system. Lastly, these principals sought to act as ethical and professional models for values they sought to embed in their schools.

Also of interest is what we did not find in the studies. An increasingly influential leadership approach absent from mention in any of these studies was distributed leadership (Spillane, 2005). In the Thai context, the terms “leader” and “administrator” are still used synonymously, and leadership is wholly associated with position or role. Thus, it is rather hard for a Thai principal, teacher, or parent to conceive of someone being a “leader” if s/he is not an administrator. Similarly, instructional leadership, while somewhat more evident than distributed leadership, did not emerge as a significant role orientation in our findings.

Conclusion

At the outset of this paper, we noted the gap between the Thai government’s vision of education reform and actual changes in educational practices in schools and classrooms (Fry and Bi, 2013). Our analysis of the “successful principalship” in Thailand offers potential insight into why Thai schools been unable to adapt to a new mission. Thai principals remain oriented toward a role definition that is largely decoupled from success criteria and standards whose achievement would support education reform goals. This orientation is further reinforced by principal recruitment, selection, preparation, and training practices that are similarly decoupled from the espoused goals of the education system (Hallinger, 2018b). While this conclusion may seem harsh, it aligns quite closely to Meyer and Rowan’s (1977) conceptualization of school’s as “institutionalized organizations”. As institutionalized organizations, schools are perceived as “successful” when they mirror stakeholder expectations for what schools *should look like*, regardless of actual performance. In institutionalized organizations, performance is defined by process criteria rather than evidence of outcomes. This emphasis on the face that can be seen fits well with the broader social norms of Thai culture.

Three policy implications arise from the analysis. First, efforts should be made to define and disseminate a new set of professional standards for the principalship that is consistent with the nation’s education reform goals. Second, as suggested above, there needs to be a better alignment of system processes for recruitment, selection, and training of principals to national goals. Third, efforts should be made to develop an award and recognition system that draws on relevant, reliable data linked to authentic growth toward the goals of student capability, morality, and satisfaction. In a nation where the urban-rural achievement gap has become a threat to social stability, awards should target principals who have made significant contributions to their schools in rural and disadvantaged contexts.

In terms of research, we feel that the review's initial inquiry into what is meant by "successful schools principalship" in the Thai context remains incomplete. Our review concluded that the fundamental processes which guide the Thai education system in practice, and which shape the expectations for principals, are largely decoupled from the espoused educational goals. Research should be conducted that elicits stakeholder expectations of Thai schools and principals in a more explicit way.

This research also prompts us to inquire more deeply into the "indigenous" leadership practices (Bajunid, 1996) of Thai principals. Our analysis suggested that the enactment of successful leadership and management practices is clearly influenced by Buddhist principles. Successful Thai principals meet cultural expectations by preserving harmony, kindness, and face, while seeking to gain and sustain the cooperation of teachers. Qualitative research is needed that inquires into how Thai school leaders enact leadership practices that meet the cultural expectations of stakeholders while also adapting to international education norms and standards that emphasize competition and performance outcomes. These are important questions that lie at the nexus of research, policy, and practice that we hope scholars will take up in the coming decade.

Acknowledgment

This research project is supported by Ratchadapiseksompotch Fund Chulalongkorn University.

References

- Akarachan, K., Kenaphum, S., Kosolkittiampon, S., 2018. The factors influence the school administrators' competency to the educational excellence at primary schools in lower northeast of Thailand. *J. Grad. Sch. Pitchayatat, Ubon Ratchathani Rajabhat University* 13 (1), 1–8.
- AksornAct, 2017. Thailand's Education Leader Symposium 2016 (Part 6/9).
- Apiwatanasingha, D., Jampol, O.-A., Sawongpol, A., 2008. The relationship between the emotional quotient and work success of the school administrator the office of the education service area of Kamphaeng Phet 1. *The Golden Teak: Humanit. Soc. Sci. J.* 14 (2), 124–142.
- Atagi, R., 2011. Secondary Teacher Policy Research in Asia: Secondary Teachers in Thailand.
- Bajunid, I.A., 1996. Preliminary explorations of indigenous perspectives of educational management. *J. Educ. Adm.* 34 (5), 50–73.
- Banlue, A., Tiacharoen, S., 2018. Sappurisa-dhamma of administrators and organizational intelligence of school under the secondary educational service area office 8. *J. Educ. Adm.* 9 (2), 155–175.
- Belchetz, D., Leithwood, K., 2007. Successful Leadership: Does Context Matter and if So, How? Successful Principal Leadership in Times of Change. Springer, pp. 117–138.
- Bulan, K., 2017. Successful educational administrators. *Dusit Thani Coll. J.* 11 (1), 380–388.
- Bush, T., 2012. International perspectives on leadership development: making a difference. *Prof. Dev. Educ.* 38 (4), 663–678.
- Chaita, P., Anannawee, P., 2009. Causal factors affecting the success of small schools management under the office of the basic education commission. *J. Educ. Adm. Burapha* 4 (1), 51–64.
- Chaiwong, K., Sanghiran, M., 2017. Structural equation model of factors affecting managerial success of schools under the provincial administration organizations. *Soc. Sci. Res. Acad. J. [Special issue]*. 11, 43–58.
- Chaiyakit, K., Jongolnee, L.-E., Thanaphaet, C., 2020. Academic leadership and buddhism method for educational management. *NEU Acad. Res. J.* 10 (2), 165–175.
- Chantanontree, S.-a., Thongpan, S., 2017. Buddhist doctrines for education administrators. *Dhammathas Acad. J.* 17 (1), 221–233.
- Charoenpornsakul, N., Katangchol, S., Bangbon, S., 2018. Quality of management in educational institutions. *J. Pacific Ins. Manag. Sci.* 4 (2), 337–348.
- Chiangkul, W., 2016. สภาวะการศกษาไทย ปี 2557/2558 จะปฏิรูปการศกษาไทยให้ทันโลกในศตวรรษที่ 21 ได้อย่างไร [the State of Thailand Education 2014/2015 "How to Reform Thailand Education Towards 21st Century?"]. Office of the Education Council, Bangkok, Thailand.
- Day, C., Gurr, D., 2018. International network as sites for research on successful school leadership. In: Lochmiller, C.R. (Ed.), *Complementary Research Methods for Educational Leadership and Policy Studies*. Springer International Publishing, Cham, pp. 341–357.
- Day, C., Sammons, P., 2013. Successful Leadership: A Review of the International Literature. ERIC.
- Day, C., Sammons, P., 2016. Successful School Leadership. ERIC.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Dharmapiya, P., Saratun, M., 2020. Cultivating a Sufficiency Mindset in Thai Schools. Sufficiency Thinking. Routledge, pp. 129–147.
- Dixon-Woods, M., Cavers, D., Agarwal, S., et al., 2006. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med. Res. Methodol.* 6 (1), 35.
- Duangtongkul, S., Jaroensuk, B., Bunjong, W., 2016. Leadership behavior of administrators affecting success in educational management of school under surat thani primary educational service area office 3. *Suratthani Rajabhat J.* 3 (2).
- Fry, G.W., Bi, H., 2013. The evolution of educational reform in Thailand: the Thai educational paradox. *J. Educ. Adm.* 51 (3), 290–319.
- Fry, G.W., Sangnapaboworn, W., 2018. The evolution of educational policy and planning in Thailand: from nec to onee and oec. In: *Education in Thailand*. Springer, pp. 503–515.
- Fry, G.W., 2018. *Education in Thailand: An Old Elephant in Search of a New Mahout*. Springer, Singapore.
- Fullan, M., 2002. Principals as leaders in a culture of change. *Educ. Leader* 59 (8), 16–21.
- Gurr, D., Wang, L.H., Drysdale, L., 2016. Successful school leadership: case studies of four Singapore primary schools. *J. Educ. Adm.* 54 (3).
- Gurr, D., 2015. A model of successful school leadership from the international successful school principalship project. *Societies* 5 (1), 136–150.
- Gurr, D., 2017. A model of successful school leadership from the international successful school principalship project. *How School Leaders Contribute to Student Success*. Springer, pp. 15–29.
- Hallinger, P., Lee, M., 2011. A decade of education reform in Thailand: broken promise or impossible dream? *Camb. J. Educ.* 41 (2), 139–158.
- Hallinger, P., Lee, M., 2013a. Exploring principal capacity to lead reform of teaching and learning quality in Thailand. *Int. J. Educ. Dev.* 33 (4), 305–315.
- Hallinger, P., Lee, M., 2013b. Mapping instructional leadership in Thailand has education reform impacted principal practice? *Educ. Manag. Adm. Leader* 42 (1), 6–29.
- Hallinger, P., Leithwood, K., 1996. Culture and educational administration. *J. Educ. Adm.* 34 (5), 98–116.
- Hallinger, P., Suriyankietkaew, S., 2018. Science mapping of the knowledge base on sustainable leadership, 1990–2018. *Sustainability* 10 (12), 4846.
- Hallinger, P., Thang, T.D., 2014. Exploring the contours of context and leadership effectiveness in vietnam. *Lead. Manag.* 20 (2), 43–59.
- Hallinger, P., 2004. Meeting the challenges of cultural leadership: the changing role of principals in Thailand. *Discourse* 25 (1), 61–73.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2), 125–142.

- Hallinger, P., 2018a. Bringing context out of the shadows of leadership. *Educ. Manag. Adm. Leader* 46 (1), 5–24.
- Hallinger, P., 2018b. Thailand's challenge of systemic education reform: where are the leaders with the "right stuff"? In: Fry, G.W. (Ed.), *Education in Thailand: An Old Elephant in Search of a New Mahout*. Springer Singapore, Singapore, pp. 555–577.
- Hallinger, P., 2018c. Thailand's challenge of systemic education reform: where are the leaders with the "right stuff"? In: *Education in Thailand*. Springer, pp. 555–577.
- Hofstede, G., Hofstede, G.J., Minkov, M., 2005. *Cultures and Organizations: Software of the Mind*. McGraw-hill, New York.
- Issara, P., Thee-asana, K., 2017. The model to the buddhist administration of basic education administrators in khonkean elementary education service area office 3 by application philosophy of the sufficiency economy. *J. Educ. Adm.* 13 (2), 59–68.
- Jirapornkul, S., Yolles, M., 2010. Assessing values and value change in Thai organizations. *J. Organ. Transform. Soc. Change* 7 (3), 321–347.
- Khemkai, A., 2016. A causal relationship model of female directors' administrative factors affecting the effectiveness of schools under primary education service area offices in the north-eastern region of Thailand. *Ratchaphruek J* 14 (3), 65–71.
- Kukkong, K., Eungpuang, A., 2014. A paths of success in administration of feminine administrator under the office of secondary educational service area 25. *J. Educ. Grad. Stud. Res.* 8 (3), 216–225.
- Kulophas, D., Hallinger, P., 2021. Leading when the mouth and heart are in unison: a case study of authentic school leadership in Thailand. *Int. J. Leader. Educ.* 24 (2), 145–156.
- Leithwood, K., Day, C., 2007. *Starting with what We Know. Successful Principal Leadership in Times of Change*. Springer, pp. 1–15.
- Leithwood, K., Louis, K.S., Anderson, S., et al., 2004. How Leadership Influences Student Learning. Review of research. Wallace Foundation, The, p. 90.
- Leithwood, K., 2005. Understanding successful principal leadership: progress on a broken front. *J. Educ. Adm.* 43 (6), 619–629.
- Louis, K.S., Leithwood, K., Wahlstrom, K.L., et al., 2010. *Learning from Leadership: Investigating the Links to Improved Student Learning*. New York, NY.
- Meyer, J.W., Rowan, B., 1977. Institutionalized organizations: formal structure as myth and ceremony. *Am. J. Sociol.* 83 (2), 340–363.
- Muktarakosa, S., Pangthai, S., Sombatsakulkit, E.-A., 2018. The professionalism of the administrators of the Thailand education school 4.0. *Humanit. Soc. Sci.* 4 (2), 453–461.
- Mulford, B., Silins, H., 2011. Revised models and conceptualisation of successful school principalship for improved student outcomes. *Int. J. Educ. Manag.* 25 (1), 61–82.
- OBEC, 2017. คู่มือการประเมินสถานศึกษาเพื่อรับรางวัลพระราชทานระดับประถมศึกษาและมัธยมศึกษา [Handbook for Royal School Awards Assessment]. OBEC, Bangkok, Thailand.
- OECD, 2017. *Education Gps - Thailand - Student Performance (Pisa 2015)*. Available from: <http://gpseducation.oecd.org/CountryProfile?primaryCountry=THA&treshold=10&topic=PI>. (Accessed 22 October 2017).
- OECD/UNESCO, 2016. *Education in Thailand: An Oecd-Unesco Perspective*. OECD Publishing, Paris.
- ONESQA, 2011. *Manual for Assessors the Third-Round of External Quality Assessment for Basic Education 2011–2015*. ONESQA.
- Pechsangkoon, J., Thanaphaet, C., 2020. Managerial skills of school administrators and the excellence of private schools. *J. MCU Ubon Rev.* 5 (3), 577–590.
- Phochsallee, P., Wangpanich, P., Nachwichet, K., et al., 2016. Factors and indicators of women education leaders characteristic which supporting the success of school management in the northeast. *J. Educ. Adm. Khon Kaen University.* 12 (2), 147–156.
- Phutthasarn, G., Kessung, P., Surakittawon, S., 2016. A development of the causal model toward successful management of secondary schools under the secondary educational service areas in northeastern. *J. Educ. Adm.* 12 (1), 62–71.
- Piboolsravut, P., 2004. Sufficiency economy. *ASEAN Econ. Bull.* 21 (1), 127–134.
- Piungam, S., Suwannatnai, W., Pothiwat, S., et al., 2015. The effective leadership of outstanding school administrators: multi – case studies. *J. Educ. Adm.* 38 (4), 168–176.
- Prabjandee, D., 2019. Unwelcome truths of the professional learning community policy in Thailand: teacher's sensemaking. *Issues Educ. Res.* 29 (1), 223–242.
- Prasongsuk, K., Suwannatnai, W., Wonganusit, W., et al., 2015. The development model of effectiveness leadership for small school administrators under the office of the basic education commission in the inspection region no. 11. *J. Educ. Adm.* 11 (2), 83–91.
- Promkan, S., Ketsang, P., Sroinam, S., 2013. The causal relationship model of the administrators' leadership factors affecting the secondary schools effectiveness under the office of educational service areas in the northeast. *Res. Dev.* 8 (23), 49–61.
- Pruetikul, P., Anannawee, P., 2014. Communication models of success full extra large-sized secondary schools administrators under the office of the basic education. *J. Educ. Adm.* 8 (2), 138–150.
- Raihani, 2008. An Indonesian model of successful school leadership. *J. Educ. Adm.* 46 (4), 481–496.
- Raksanakorn, K., 2020. School administrators' attributes for 21st century. *J. Mod. Learn. Dev.* 5 (3).
- Redmond, M., 1998. *Wandering into Thai Culture*. Redmondian Insight Enterprises.
- Royal Thai Embassy, 2017. Thailand 4.0. Royal Thai Embassy, Washinton D.C.. Available from: <http://thaiembdc.org/thailand-4-0-2/>. (Accessed 21 October 2017)
- Rukbanglaem, C., Sungtong, E., Kerdthip, C., et al., 2015. Ethical leadership: the power driven force to leadership in educational institute. *Princess Naradhiwas Univ. J.* 8 (1), 168–181.
- Sangnapaboworn, W., 2018. The evolution of education reform in Thailand. In: *Education in Thailand*. Singapore: Springer, pp. 517–554.
- Sankaew, A., Chusorn, P., 2016. Adaptive leadership of school administrator under the office of secondary educational service area 19. *Grad. Stud. Res.* 10 (5), 180–186.
- Scheerens, J., 2000. *Improving School Effectiveness*. UNESCO.
- Somboon, P., Kompuang, A., 2018. The leadership development of school administrator by using benchmarking. *Phikun J.* 16 (2), 197–218.
- Somprach, K., 2017. Learning leadership of school principals: a grounded theory study. *Soc. Sci. Res. Acad. J.* 12 (34), 51–66.
- Spillane, J.P., 2005. Distributed leadership. *Educ. Forum* 143–150.
- Sreebuaand, S., Thanaphaet, C., 2020. Buddhist transformational leadership and school administration. *J. MCU Ubon Rev.* 5 (3), 591–604.
- Srichandari, J., Thanaphaet, C., 2020. Ethical leadership according to buddha dharma principles of school administrators. *J. Mod. Learn. Dev.* 5 (5), 321–332.
- Suk-ied, V., Aekpet, C., Narkkhan, S., 2015. Factors affecting school's development to become world class standard schools under the office secondary educational service area 14. *Narkbhut Paritat J.* 7 (2), 77–89.
- Sukthet, P., Thuanchang, W., Samniang, P., 2020. Effective model of transformation for leadership development based on formal buddhadhamma principles of school administrators basic education. *J. Educ. Rev. Faculty of Educational in MCU* 7 (3), 43–58.
- Sutamchai, K., Rowlands, K.E., Rees, C.J., 2020. The use of mindfulness to promote ethical decision making and behavior: empirical evidence from the public sector in Thailand. *Publ. Adm. Dev.* 40 (3), 156–167.
- The World Bank, 2017. *Expenditure on Education as % of Total Government Expenditure (%)* | Data. Available from: https://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS?contextual=region&end=2013&locations=TH-1W-US&start=2013&view=bar&year_high_desc=true. (Accessed 22 October 2017).
- Thetarat, W., Sripokangkul, S., 2018. Director's leadership of municipality school 5 (wat pa chitta samakkhi), nakhon ratchasima municipality, nakhonratchasima province. *Governance J.* 7 (2), 449–473.
- Tsering, T., 2005. *The Four Noble Truths: The Foundation of Buddhist Thought*, vol. 1. Simon and Schuster.
- Varavarn, K., 2011. *Educational Reform and Change in Thailand: A Retrospective View*. Keynote paper presented at the Asia Leadership Roundtable. Nakon Pathom, Thailand.
- Vichit-Vadakan, J., 2017. *Public Ethics and Corruption in Thailand*. Public Administration in Southeast Asia. Routledge, pp. 79–94.
- Walker, A., Hallinger, P., 2015. A synthesis of reviews of research on principal leadership in east asia. *J. Educ. Adm.* 53 (4), 554–570.
- Wong, K.c., 2005. Conditions and practices of successful principalship in shanghai. *J. Educ. Adm.* 43 (6), 552–562.
- Youngtawee, P., Charoenkul, N., 2015. A study of the academic leadership of successful school administrators under the office of secondary educational service area 3. *An Online J. Edu.* 10 (3), 174–188.
- Yuenyong, C., 2019. Lesson learned of building up community of practice for stern education in Thailand. In: *AIP Conference Proceedings*, p. 020002.

School leadership practice in Oman: an investigation of authentic leadership, teacher engagement and turnover intention

Yasser F. Hendawy Al-Mahdy and Mahmoud Mohamed Emam, College of Education, Sultan Qaboos University, Muscat, Oman

© 2023 Elsevier Ltd. All rights reserved.

Introduction	501
School leadership in Oman: a synopsis	502
Authentic leadership	503
Teacher engagement	503
Teachers turnover intention (TTI)	504
The relationships among authentic leadership, teacher engagement and teacher turnover intention	505
Method	505
Data collection	505
Instruments	506
Principal authentic leadership (PAL)	506
Teacher turnover intention (TTI)	506
Teacher engagement (TE)	506
Data analysis	506
Results	506
Preliminary analyses	506
Assessment of the structural model and mediating effect	507
Discussion	507
Conclusion	509
Acknowledgment	509
References	509

Introduction

It is not surprising that educational leadership and management (EDLM) including school leadership practices have become, and continue to be, buzz words in Arab-speaking countries in general and the Gulf Cooperation Council (GCC) countries in particular (Bailey et al., 2021; Hallinger and Hammad, 2019). In Oman, school leadership was given tremendous responsibility for resolving the long-standing problems of the inequitable student learning outcomes in the schooling system (Al-Mahdy and Emam, 2018). As a result of the precipitated changes that took over school leaders' practiced role in Oman in the last two decades, new policy mandates, new standards, and new demands for accountability have given a thrust to connecting school leadership practices to finding answers to big educational questions that have remained chronic problems in the educational system. These questions center around significant issues including reducing the achievement gap in general and gender achievement gap in particular (Al Jabri et al., 2018; MoE, 2018), evidence based strategies to achieve adequate yearly progress (Al-Shehhi et al., 2019), enabling diverse learners to access the curriculum including those with special educational needs (Al-Mahdy and Emam, 2018; Emam and Al-Mahdy, 2020), keeping a balance between efficiency and effectiveness (Al-Farsi, 2017), sharing leadership responsibility to develop a community of practice in which all teacher are engaged (Al-Harthi and Al-Mahdy, 2017), and finally, retaining teachers in jobs to prevent teacher turnover and retaining students in the classroom (Al Balushi et al., 2014; Al-Farsi, 2017) to prevent drop-out (World Bank Report, 2012; Education Council, 2020).

A recently conducted systematic review on school leadership models in educational research indicated that certain models have received increasing attention by researchers in different contexts. These models include distributed leadership, instructional leadership, and transformational leadership (Gumus et al., 2018) and ultimately argue for the best leadership style or the true leader who will establish the ideal leader-follower relations in school (Furman, 2012). Similar to any other part in the world, Arab countries including GCC countries and Oman focused on research that could help constructing a meaningful understanding of true school leadership. In their review of 48 article on the EDLM in the Arab region since 1990 Oplatka and Arar (2017, p. 267) stated that almost all research studies centered on identifying leadership styles and "the barriers faced by educational leaders in applying some of the leadership models originated in Western scholarship or on exploring and examining leaders' managerial perceptions and practices both in times of routine and reforms".

Islam, as a religion, is based on trusting the authenticity of Islamic leaders. Trustworthiness, from the Islamic leadership perspective, implies that a leader is expected to exert his/her efforts to guide and treat his/her followers justly. Additionally, the authenticity of an Islamic leaders is displayed to followers when the actions and practices of leadership serve the context and the support the followers to develop. Given that Oman's culture is informed by Islamic religious practices at home, we expect that teachers are likely

to search for authenticity in their school leaders and endorse it as their preferred leadership style, assuming that authentic leadership is likely to support work engagement. It could be argued, therefore, that the contextual Islamic cultural and social practices in Oman may accentuate authentic leadership practices in school and affect teacher engagement and teacher turnover intention.

In essence, the aim of the current article is twofold: (1) to contextualize the cultural, national, sociological aspects of school leadership in Oman, (2) report the results of an empirical study which explored how authentic leadership impact on teachers' engagement and turnover intentions in Oman.

School leadership in Oman: a synopsis

The Sultanate of Oman celebrated 50 years of the establishment of its modern integrated educational system which began in 1970 as Sultan Qaboos took over the throne. Oman is one of the GCC Arab-speaking countries, bordered by the United Arab Emirates, Saudi Arabia, and Yemen. In 2019, the estimated population of Oman was 4.84 million, of which expatriates represent proportionate number (1706 million), (Oman Population, 2020). The Human Development Report of 2019 indicated that Oman is a case example of the fastest progress countries in the Human Development Index. In the last three decades, Oman has risen from a poor to a rich country. This was indicated by high gross enrollment literacy rates, high income per capita, and an increase in life expectancy. In the realm of education, Oman achieved 100% access to basic education, and between 80 and 90% of them proceed to post-basic education. The progress that the educational system in Oman has made to date has been recognized in several international reports (World Bank Report, 2012; Human Development Report, 2019).

Historically, the educational system in Oman did not include any reference to school leadership until 1970s, which witnessed the beginning of the Renaissance age as His Majesty Sultan Qaboos took over the throne. Before 1970s, the sultanate of Oman had a primitive form of schooling informed by the religious culture in which learning to read and recite Quran was the ultimate goal of each family. During the early years of schooling, school leadership was embodied in the religious Imam. The paradigm-shift from schools of Quran to a typically modern school system began rapidly in the 1970s. Initially, there were a few schools, with a total of 196 schools across the country. With the increasing demand for new schools to enhance and support the expansion of the basic education system in 1997/1998, the number of school principals increased as the number of schools have increased. In 2011, the number of school principals and deputy principals was 1498. Gradually, this number increased 5% in 2019 to reach 1578 (MoE, 2019).

Oman has always adopted a centralized educational system. Therefore, the role of school principals has been to ensure the implementation of the rules and regulations envisioned a dictated by school policy makers (Al-Mahdy et al., 2018; Al-Mahdy and Emam, 2018). Culturally, school principals are perceived with appreciation and respect by the Omani society. As a result of having a central management system of schools, school principals' capabilities are mostly limited to move forward to solve complaints and challenges (Emam and Al-Mahdy, 2020). In response to his Majesty's speech in which he evaluated and reviewed Oman's educational system in 2010, and in response to World Bank Report in 2012, tremendous efforts were made to improve school leadership practices. In 2013, For example, Oman's Ministry of Education and New Zealand Education Consortium carried out an evaluative research project to review the educational system. The project focused on evaluating several aspects such as governance and management, financial affairs, student learning, teachers, school culture, curriculum, alignment between labor sector and educational system response. Findings revealed that the MoE is has mandated that in order to develop the quality of the educational system, priority should be given to enhancing and developing school leadership that could ensure school accountability (New Zealand Education Consortium, 2017).

In Oman, school leadership has been associated with increasing school effectiveness as it influences how members of an organization (e.g., schools) function (Al-Harthi and Al-Mahdy, 2017). The MoE, therefore, gave considerable attention to school leadership recognizing its importance in developing a functional system which would ensure school effectiveness and improve student learning outcomes. The MoE also introduced the Certificate of School Management in collaboration with Sultan Qaboos University. The certificate was a two-year endorsement program which targeted school principals of elementary schools. The program continued for 11 years from 2003 to 2004 to 2014 to 2015. In 2014, the Ministry of Education in Oman established the Specialized Institute for Professional Training of teachers (SIPTT) in an attempt of create an atmosphere and a platform where educational leaders will exhibit their efforts and celebrate their achievements (New Zealand Education Consortium, 2017). The establishment of the SIPTT was also seen as a major step taken by the MoE to address the needs of the educational system following the failure resulting from Oman students' low performance on international assessments such TIMSS and PIRLS (Al Shabibi and Silvennoinen, 2018; Al-Shehhi et al., 2019; MoE, 2018).

Despite the fact that there have been national initiatives that focused on enhancing school leadership (Al-Harthi and Al-Mahdy, 2017; Al-Mahdy et al., 2016) there has been little research that investigated the effectiveness of school leadership practices as perceived and envisioned by teachers and/or school leaders. We searched scientific databases and electronic resources (e.g., Scopus, PsycInfo, EBSCO, Sage Publications, Taylor and Francis, Science direct) to locate research studies which were conducted on school leadership in Oman. We located only 4 studies which examined school leadership practices in Oman (Al-Mahdy et al., 2016; Al-Harthi and Al-Mahdy, 2017; Al-Mahdy and Emam, 2018; Al-Mahdy et al., 2018). Al-Farsi (2007) investigated headteachers' perceptions of leadership role and found that the majority of them emphasized the importance of some aspects of leadership including having a shared school vision, giving teachers the opportunity to engage in teamwork, preparing an adequate school climate for work, commitment to the school, sharing leadership with teachers and delegating power. In another study, principal instructional

leadership was found to contribute to the commitment of their teachers directly through active instructional leadership practices. In addition, it also contributed to teacher commitment indirectly by influencing the collective efficacy of school teachers. (Al-Mahdy et al., 2018). The qualitative inquiry based on the data collected in the evaluation study by the New Zealand Education Consortium in 2013 in collaboration with the MOE concluded that school principals and teachers expressed lack of enthusiasm and having low morale because they felt dissatisfied with their job for various reasons (New Zealand Education Consortium, 2017). In a recent review of the leadership studies which were conducted in GCC countries and published in Arabic journals, Hammad et al. (2020) searched Arabic databases and found a total of 272 studies on EDLM written in Arabic language, 12% of which were conducted in Oman. The majority of the studies were published in the Kingdom of Saudi Arabia.

Authentic leadership

Authentic leadership implies the consistent display of explicit leadership behaviors that encourage positive psychological and ethical atmosphere with the aim of promoting self-awareness, internal moral perspective, balanced information processing and transparent communication with employees (Nasab and Afshari, 2019; Gatling et al., 2016; Azanza et al., 2015; Walumbwa et al., 2008), which makes it suitable for educational institutions (Duignan, 2014). Based on reviewing the literature, definitions of authentic leadership refer to main core components. The components include self-awareness (SA), relational transparency (RT), balanced information processing (BIP) and internal moral perspective (IMP) (Neider and Schriesheim, 2011; Walumbwa et al., 2008). The first component of SA refers to leaders' self-understanding of their strengths and weaknesses as well as of their temperament and character. The second component of RT indicates the leaders' ability to show evident communication of themselves and their genuine self to their followers. The third aspect of BIP is crystalized in providing unbiased information and less self-centered knowledge to the followers. The fourth component IMP refers to putative and genuine moral behaviors, values as well as adequate self-regulation (Neider and Schriesheim, 2011; Oh et al., 2018). The four components, arguably, may have certain degree of importance. In the field of education, SA, for example, was ranked as the highest feature used by leaders as perceived by the teachers, followed by BIP, RT and IMP respectively (Kulophas et al., 2018).

In general, studies on authentic leadership focused on investigating its relation to employees in terms of job satisfaction (Hitchcock, 2019), trust (Ahamed et al., 2013), engagement (Alazmi and Al-Mahdy, 2020; Oh et al., 2018; Ahamed et al., 2013), performance (Nasab and Afshari, 2019; Kim, 2014), turnover intention (Oh and Oh, 2017; Qureshi and Aleemi, 2018; Azanza et al., 2015; Gatling et al., 2016), academic optimism (Kulophas et al., 2018), commitment to organizations (Gatling et al., 2016) and self-efficacy (Kulophas et al., 2018; Kim, 2014). Scholars who examined authentic leadership in different contexts identified an array of characteristics and skills of authentic leaders. These include the ability to create a positive relationship with employees (Oh and Oh, 2017), the desire to provide learning opportunities and psychological support to their subordinates and retain human capital in the organization (Qureshi and Aleemi, 2018), being emulated as role models, display of morals, confidence, knowledge, optimism and resilience, serving their subordinates through increasing confidence, reducing burnout, enhancing followers' work engagement, and placing value on each person in the organization, and promoting hope, trust and positive emotions (Qureshi and Aleemi, 2018; Azanza et al., 2015). Accordingly, it has been argued that such characteristics of authentic leadership are necessarily needed in the educational settings as and that training school principals and deputy principals on authentic leadership would help them disseminate positive energy for more engaging and productive schools, and high-quality outcomes for future students (Kulophas et al., 2018).

Given that there have been proliferated efforts in Oman to highlight the role of school leadership the teacher's professional attitudes and career decisions, the leadership style of the school leaders becomes essential to examine. Although there is an array of leadership styles that have been examined in Oman, we found no records of examining authentic leadership. Our choice of investigating the impact of authentic leadership on teacher engagement and turnover intentions is based on its congruence with the cultural and religious practices in the Omani society, and therefore are inseparable from the educational and professional practices within the school context. In schools, an authentic leadership is the one whose behaviors explicitly support teachers to self-evaluate their performance against the intended goals set up by the MoE and communicated by the school leaders. Teachers trust the morals by which the school leader abides in his behaviors and decisions. Teachers also connect between the school principal's authentic leadership style and his commitment to Islamic values and believe that both are congruent with no discrepancy that could lead to a distorted message.

Teacher engagement

One of the essential factors affecting employers and employees in any workplace is work engagement. Taking school as an example, school leaders often rely on hard working teachers, who display commitment through taking responsibility and engage with the demands of the work environment. Within the collective culture in Oman and in congruence with the instruction of Islam, a leader needs truthful and honest subordinates similarly as the prophet Mohamed had companions that supported him and defended his call for Islam. Teachers showing work engagement are important resources for school leaders at times of crisis management. Work engagement focuses on the work itself and it is about the extent to which teachers are immersed and involved in their duties to achieve the school vision, mission, and intended outcomes. Additionally, work engagement in school focuses on job resources

such as autonomy, safety, and school support from school leaders and teacher interchangeably (Kim, 2014). Notably, work engagement was first investigated in relation to the effectiveness of business organizations (Kim, 2014; Oh et al., 2018), delivery of health care services (Qureshi and Aleemi, 2018) and education settings (Al-Farsi, 2007; Kulophas et al., 2018). From this perspective, work engagement has been examined as a direct or indirect outcome of different leadership styles and practices such as charismatic, transformational, and AL (Azanza et al., 2015). Leaders—who are seen as key drivers of subordinates' work engagement—perceive that work engagement is an instrument for retaining employees (Azanza et al., 2015). Alternatively, some scholars argue that leaders who display authenticity in their leadership and establish strong bonds with employees are most likely to affect their work engagement, i.e., work engagement is an outcome of positive leadership practices (Oh and Oh, 2017; Oh et al., 2018).

In school terms, work engagement implies that teachers have work positive state of mind and a motivational texture that is characterized by three main components: vigor, dedication and absorption (Oh et al., 2018; Azanza et al., 2015; Schaufeli et al., 2019). The first element, *vigor*, relates to having eminently motivated teachers with a willingness to invest their efforts to the teaching and other school related tasks. The second element, *dedication*, refers to being committed, enthusiastic, and caring to one's work and expected outcomes. The final element of work engagement is *absorption* and it refers to the state of being totally immersed in the work, showing happiness with overall work and task contexts (Kim, 2014; Kulophas et al., 2018; Oh et al., 2018).

There is evidence of significant interconnected relations between work engagement and other variables such as employees' occupational self-efficacy and role-based performance, employees' interpersonal trust (Kim, 2014), team work, job satisfaction, turnover intention, work group identity, authentic leadership (Azanza et al., 2015; Alazmi and Al-Mahdy, 2020), work core values (Qureshi and Aleemi, 2018; Oh et al., 2018) and teachers' academic optimism (Kulophas et al., 2018). The aforementioned relationships are predominantly positive, and in some cases negative as in the case of employees' turnover intention (Azanza et al., 2015). Work engagement is perceived as vital mediation mechanism between an array of variables including interpersonal trust and role-based performance, occupational self-efficacy and role-based performance, leadership and organizational citizenship behavior (Kim, 2014). Additionally, it partially mediates the relationship between authentic leadership and turnover intention (Azanza et al., 2015). In a study by Oh et al. (2018), variables served as mediators in the relations between work engagement and other variables such as practicing core values. In Thailand TE was found to result in meeting daily challenges by boosting teachers' immersion in school tasks, thereby enhancing their self-efficacy (Kulophas et al., 2018). Teacher's engagement in Oman implies involvement in school related administrative and academic tasks, negotiating how they can make right decisions on a daily basis to achieve the short and long-term goals in the areas of academic achievement, social and emotional learning, and cultural values acquisition. Al-Farsi (2007), for example, has indirectly investigated teacher's engagement by exploring how teachers are involved in teamwork in school. Based on interviewing headteachers in Oman, Al-Farsi (2007) indicated that head teachers stressed the importance of engaging teachers in some management roles with the aim of gaining teachers' support and decreasing the managerial demands. It was noted, however, that some teachers may refuse to play an administrative role due to the teaching workload, which could also lead to teacher turnover (Al Lawati, 2014).

Teachers turnover intention (TTI)

TTI refers to a voluntary proposed mental plan of an individual to search for a new job, leave or quit the current job as a result of various personal, psychological, contextual reasons (Mobley et al., 1979; Azanza et al., 2015; Qureshi and Aleemi, 2018) such as age, gender, education, attitude, intrinsic motivation, salary, promotion, the work itself, amount of work and the relationship with superiors (Mobley et al., 1979). Thus, the willingness to actual turnover process is predicted through the individual's intention to leave the job (Mobley et al., 1979; Oh and Oh, 2017; Qureshi and Aleemi, 2018).

TTI has attracted the interest of scholar in management, psychology and sociology (Azanza et al., 2015; Oh and Oh, 2017; Qureshi and Aleemi, 2018). More specifically, jobs with high rate of psychological burnout were the focus of Turnover Intention investigations, including hospitality (Gatling et al., 2016), nursing (Qureshi and Aleemi, 2018), education (Al-Mahdy and Alazmi, 2021; Hitchcock, 2019), and private and general business sectors (Azanza et al., 2015; Oh and Oh, 2017) in different countries (USA, Pakistan, Spain, Korea, Kuwait). Generally, TTI has been related to leaders' leadership style, organizational commitment, organizational performance and job satisfaction, thus throwing the ball in the manager's court (Al-Farsi, 2007; Kim, 2014; Azanza et al., 2015; Gatling et al., 2016). This association perhaps meant that positive healthy relationships, increased job satisfaction, higher level of engagement and greater retention rate or decreased burnout level reflected quality management and leadership skills (Kim, 2014; Qureshi and Aleemi, 2018; Hitchcock, 2019). In education, this argument implies that when new teachers take their teaching jobs, their satisfaction could predict their current of future intention to stay in their job (Hitchcock, 2019; Al-Mahdy and Alazmi, 2021), i.e., teachers' job satisfaction "has the advantage of reducing the cost associated with teacher turnover" (Al-Mahdy et al., 2016, p. 547). A number of studies pinpointed the variables that could curb employees' turnover intentions. Authentic leadership, for example, has been found (Gatling et al., 2016; Qureshi and Aleemi, 2018) to have an indirect negative effect on employees' turnover intention, but not the opposite. In addition, it has been found that organizational commitment is negatively connected to turnover intention and that lack of commitment leads to increased job search behavior and intention to leave (Gatling et al., 2016). Another finding indicated that employees' levels of affective commitment had a negative effect on turnover intentions and that authentic leadership had a similar effect when organization size was small to moderate (Oh and Oh, 2017). Furthermore, Azanza et al. (2015) found that the higher the employees' work engagement the lower their intention to leave their profession. Additionally, the results showed that age was negatively related to turnover intention.

In a country like the United States, shortage of qualified teachers and high teacher turnover rate emerged as a result of social context-related aspects including the increasing number of schools in poor and distant areas. Consequently, it was difficult for remote schools to develop model teaching practices and build good collaboration with parents and the local community (World Bank Report, 2012). This was typically the case in Oman where novice teacher who took their jobs in remote rural schools were likely to leave the job very soon if their wish for transfer to main city schools was not granted. This led to having less experienced teachers and expats teaching in those schools (World Bank Report, 2012).

In the Omani context, there have been several construed reasons behind teachers' intention to leave the teaching (AL-Neimi and Amzat, 2012; Al Balushi et al., 2014; Al Lawati, 2014). First, TTI may be due to environmental factors such as school size, location and school gender type and double shift-time (morning or afternoon school time) (Al Balushi et al., 2014). Second, teachers may not stand the workloads of teaching compared to the insufficient salary they are paid (AL-Neimi and Amzat, 2012). Third, it was argued that teacher–principal relationship may actually affect the teacher's decision to leave the job. Al Lawati (2014) conducted a semi-structured interview with teachers and found that teachers have the intention to quit their jobs as a result of working with uncooperative school principals. Al Lawati (2014) noted that the majority of teachers, however, may consider turnover, but most of them request to transfer to other schools before they take a final decision to actually leave the job (Al Balushi et al., 2014). Although the annual Statistics Handbook that is released by the MoE provides information on different aspects of the educational system including student dropout rate, there is no specific section on teacher turnover intention or a count of those who actually left the teaching job. To date, there is no particular study that focused on TTI in relation to school leadership in Oman, thus, providing value to the current investigation.

The relationships among authentic leadership, teacher engagement and teacher turnover intention

The relationship between authentic leadership, teacher engagement and TTI has not been studied altogether in one study, although paired relationship of the three were investigated in different recent studies (Kulophas et al., 2018; Gatling et al., 2016; Oh and Oh, 2017; Oh et al., 2018). Two studies by Qureshi and Aleemi (2018) and a study by Azanza et al. (2015), however, investigated the effect of authentic leadership on turnover intention as mediated by work engagement in different sectors such as health, education, and administration. Surprisingly, most research studies on the relationship between the three were mainly investigating the effect of authentic leadership on both turnover intention and work engagement. These studies concluded that authentic leadership has a significant effect on TTI and work engagement.

In short, there is a dearth of research examining the relationships among authentic leadership, teacher engagement, and TTI in Arab speaking countries in general and in Oman in particular. Although there are several studies that were conducted on these variables in Western countries but studies in Arab countries are at an early stage. The present study, therefore, is an attempt to explore the effect of principal authentic leadership on TTI, with a particular focus on the mediating role of teacher engagement in the context of Oman. Specifically, this study aims to answer the following three questions:

- Does principal authentic leadership significantly affect teacher turnover intention?
- What is the impact of teacher engagement on teacher turnover intention?
- Is there a mediation effect of teacher engagement on the relationship between principal authentic leadership and teacher turnover intention?

Method

The data reported in this article was drawn from a larger dataset of a research project that examined school leadership in Oman. The data was collected using a cross-sectional research design, structural equation modeling (SEM) to investigate the relationships which were indicated in the aforementioned research questions. In this section, we describe our variables and measures, procedures for collecting, analyzing, and reporting the data.

Data collection

Data were collected using a convenient sampling technique. We approached a large number of teachers in basic education middle schools (grades 5–10) in the biggest city of the Sultanate of Oman either by email or by visiting schools. Over two months during the spring semester of 2020. We sent an invitation letter, consent form and scales package electronically to teachers, requesting their participation in the survey. 398 teachers agreed to take part, this number exceeded the desired accuracy for a confidence level of 95% (Gill et al., 2010) and was sufficient to meet the minimum sample size recommended ($n = 200$) for running SEM (Kline, 2011). Of them, 224 were female (56.3%) and 174 were male (43.7%). The amount of teaching experience in the sample ranged from 1 to 27 years, with an average of 12.08 years ($SD = 5.26$). Approximately 11.3% of the participants had 1–5 years of teaching experience, 25.4% had 6–10 years of experience, and 63.3% had more than 10 years of experience.

Instruments

Principal authentic leadership (PAL)

Walumbwa et al. (2008) developed the authentic leadership questionnaire (ALQ) to measure managers' authentic leadership. Indeed, the ALQ is one of the most widely used measures of authentic leadership (Gardner et al., 2011). Although Neider and Schriesheim (2011), indicated some concerns with the ALQ but they used the same four dimensions of the ALQ, namely: self-awareness (SA), relational transparency (RT), balanced information processing (BIP) and internal moral perspective (IMP) to suggest an alternative inventory, Authentic Leadership Inventory (ALI). To limit the length of the questionnaire, the current study adopted the item selection approach for building a short-version scale used by some researchers (Alazmi and Al-Mahdy, 2020; Gillespie and Mann, 2004). One item with the highest factor loading on each dimension of the original ALQ (Walumbwa et al., 2008) and that included in the ALI (Neider and Schriesheim, 2011) was chosen to build a 4-items short form. The four items are My principal: "demonstrates beliefs consistent with their actions" (IMP); "seeks feedback to improve interactions with others" (SA); "says exactly what he or she means" (RT); "solicits views that challenge his or her deeply held positions" (PIB). Cronbach's alpha for the Arabic version of the 4-item short form was $\alpha = 0.94$.

Teacher turnover intention (TTI)

TTI was measured using 3-items theorized by Mobley et al. (1978) and found (e.g., Miller et al., 1979) to be important precursors to turnover: *thinking of quitting*; *intention to search*; and *intention to quit*. This measure has been widely used, with evidence of reliability and validity ($\alpha > 0.80$; e.g., Carmeli and Gefen, 2005; Cohen and Golan, 2007). The 3 items which comprise the TTI are: (1) "I think a lot about leaving my job", (2) "I am actively searching for an alternative to my job", and (3) "as soon as possible, I would leave my job". The α coefficient for this scale in current study was 0.89.

Teacher engagement (TE)

TE was assessed using a short-version of the Utrecht Work Engagement Scale (UWES) developed by Schaufeli et al. (2019). This scale features 3 items (UWES-3) which are: (1) "At my work, I feel bursting with energy" (vigor); (2) "I am enthusiastic about my job" (dedication); (3) "I am immersed in my work" (absorption). The Cronbach's alpha for UWES-3 was $\alpha = 0.93$ in our current investigation.

All items were rated using 1–5 Likert scales, with 1 representing a low rating, and 5 being high. Although each of these instruments has undergone extensive testing of psychometric properties in previous research, the current study tested the measurement properties using an Arabic translation of the three instruments to validate their use in the Omani context. The authors followed Brislin (1970) backwards translation method to translate the surveys from English to Arabic. We had to rephrase some items to ensure their suitability for Omani educators, but without altering their original meaning.

Data analysis

In this study, a two-step structural equation model approach was employed (Byrne, 2010; Hair et al., 2010). The Analysis of Moment Structures (IBM AMOS) version 22 was used to assess the validity of the measurement model, followed by examination of the structural model. Confirmatory factor analysis (CFA) was used to assess the validity of the measurement model by testing the construct, discriminant, and convergent validity of the scales. We used cut-off criterion of 0.60 for factor loading (Chin et al., 1997; Hair et al., 2010). Interpretation of the magnitude of standardized effect sizes (β) was informed by Cohen (1992) (i.e., 0.10–0.30 small effect, 0.30–0.60 moderate effect, >0.60 large effect).

A number of indices were used to indicate the robustness of fit, including Chi-square/degrees of freedom ratio (χ^2/df), Root Mean Square Error of Approximation (RMSEA), the Tucker–Lewis Index (TLI), Normative Fit Index (NFI) and the Comparative Fit Index (CFI). According to Schreiber et al. (2006), data fit is excellent when TLI, NFI, and CFI are greater than 0.95, and is acceptable when TLI, NFI, and CFI are no less than 0.90; for RMSEA, an excellent data fit requires it to be less than 0.06, and an acceptable fit requires it to be under 0.08.

Results

Preliminary analyses

Table 1 shows each variable's descriptive statistics, internal consistency estimates, and correlational coefficients. Descriptive statistics suggested that teachers in Oman perceived their principals to be highly authentic leaders ($M = 4.00$; $SD = 1.04$); teachers also perceived the existence of higher levels of teacher engagement ($M = 3.90$; $SD = 1.16$) and moderate levels of teacher turnover intention ($M = 2.83$; $SD = 1.45$). Regarding the normality of the variables, skewness ranges between 0.171 to -1.20 and kurtosis ranges between 0.075 and 1.383, indicating that the current data meet the assumption of a multivariate normal distribution (the standard skewness >3 and the standard kurtosis >10) (Kline, 2011). Regarding scale reliability, all measures indicated acceptable level of internal consistency estimates (Cronbach's α values ranging from 0.89 to 0.94), and the correlation coefficient estimates among the research variables (ranging from 0.230 to 0.586) were at a moderate statistically significant level. Also, all of the relationships between the variables were small enough to avoid the risk of multicollinearity >0.90 (Hair et al., 2010). While self-report surveys

Table 1 Descriptive statistics, convergent & discriminant validity.

Construct	$\bar{x}$	SD	α	AVE	1	2	3
Principal authentic leadership	4.00	1.04	0.94	0.788	<i>0.888</i>		
Teacher turnover intention	2.83	1.45	0.89	0.728	−0.230**	<i>0.853</i>	
Teacher engagement	3.90	1.16	0.93	0.809	0.251**	−0.586**	<i>0.899</i>

Notes: five-point Likert scale; teachers, $n = 398$, ** $p < 0.01$, diagonal elements in italics are the square root of AVE.

could cause a common method variance problem in the current study, the Harman's single factor test result indicated no dominant factor emerging from the factor analysis. PAL accounted for 33.64% of the variance, whereas teacher engagement and TTI accounted for 28.89% and 25.52% of the variance respectively (Podsakoff et al., 2003).

We, then conducted Confirmatory Factor Analysis (CFA) using AMOS v.22.0 to examine the one-factor structure of the PAL, teacher engagement, and TTI. The CFA results indicated an excellent data fit for PAL model, ($\chi^2/df = 1.083$, CFI = 1.000, NFI = 0.998, TLI = 1.000, RMSEA = 0.014) with factor loadings for observed PAL items ranging from 0.86 to 0.92. For the TTI scale, the CFA results indicated an excellent data fit for the TTI structure ($\chi^2/df = 1.818$, CFI = 0.999, NFI = 0.997, TLI = 0.998, RMSEA = 0.045) with factor loadings for observed items ranging from (0.92–0.95). For teacher engagement, the CFA results reflected an acceptable fit ($\chi^2/df = 2.964$, CFI = 0.998, NFI = 0.997, TLI = 0.994, RMSEA = 0.070), with factor loadings of observed items ranging from 0.88 to 0.91.

We also assessed the convergent and discriminant validity of the instruments. We determined convergent validity by examining the Average Variance Extracted (AVE) statistics. As shown in Table 1. The AVE values for the three instruments were 0.788 (PAL), 0.728 (TTI), and 0.809 teacher engagement. All values exceeded the criterion of 0.50, thereby demonstrating adequate convergent validity (Hair et al., 2010). We assessed discriminant validity by comparing the square root of the AVE statistics with correlations among the latent variables. Table 1 shows that the AVE square root values for each construct were greater than the correlations between latent constructs. These results support the overall validity of the proposed conceptual model (Hair et al., 2010).

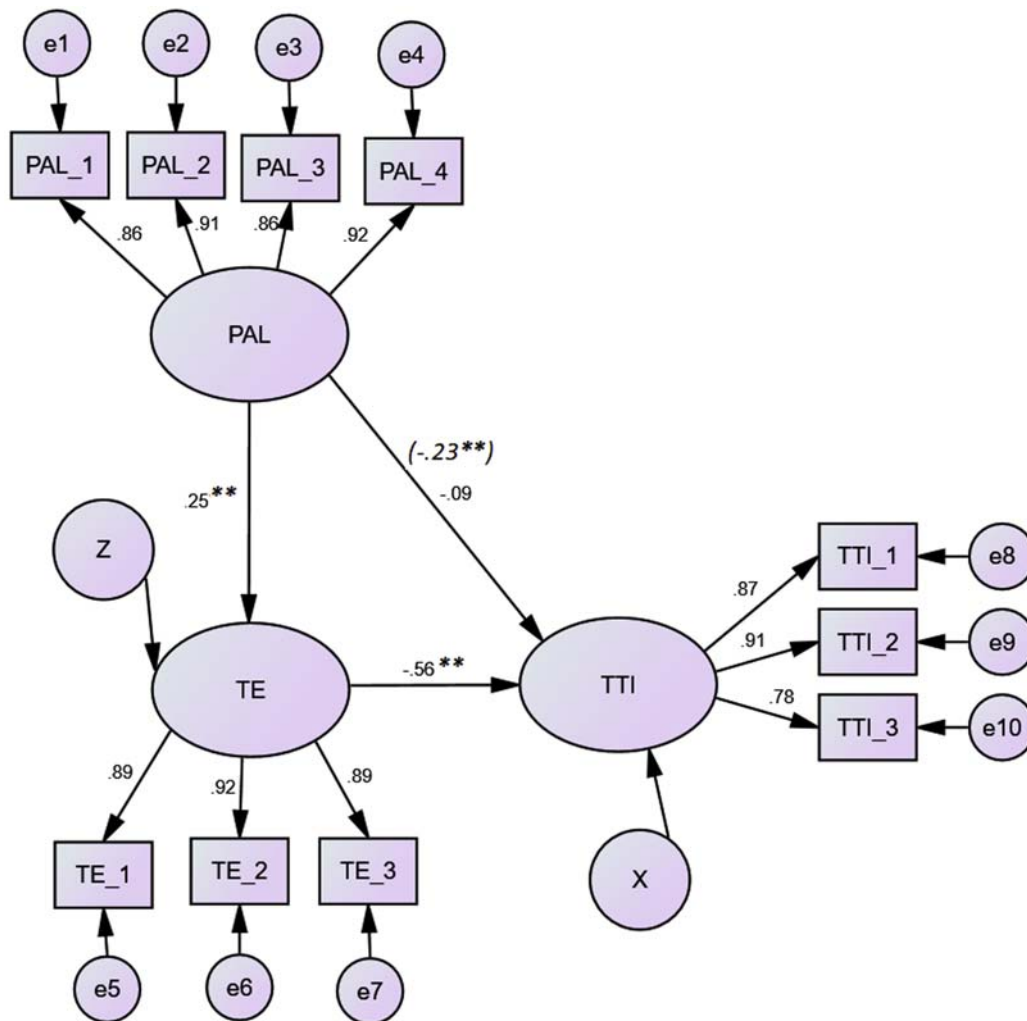
Assessment of the structural model and mediating effect

As mentioned, the focus our investigation was to explore the relationships between principal authentic leadership, teacher engagement and teacher turnover intention. Due to the strength of analyzing the complex relations among multiple latent variables simultaneously and controlling for measurement error, SEM using the Amos 22 was used to examine the influence of principal authentic leadership (PAL) on teacher turnover intention (TTI) and the mediation effect of teacher engagement. With regard to the mediation effect of teacher engagement we followed the procedure suggested by Baron and Kenny (1986) and other scholars (e.g., MacKinnon and Fairchild, 2009). Accordingly, two conditions were required for the test of mediation. First, the mediator should be significantly related to both the independent and dependent variables. Second, the mediator should significantly impact upon the dependent variables. Two SEM models, therefore, were tested to examine the relationship between PAL, TE and TTI. The first one (model A) was the direct effect model in which PAL was hypothesized to predict TTI. SEM results for this model showed good data fit ($\chi^2/df = 1.376$, CFI = 0.998, NFI = 0.992, TLI = 0.996, RMSEA = 0.031). The value of model A is shown in parentheses in Fig. 1.

The results of model B are shown in Fig. 1 and its values are shown on the figure line. According to model B, PAL had a significant positive effect on teacher engagement ($\beta = 0.25$, $p < 0.001$). While teacher engagement showed a significant negative effect on TTI ($\beta = -0.56$, $p < 0.001$), its inclusion as a mediator between PAL and TTI resulted in decreasing the regression coefficient of PAL ($\beta = -0.09$, $p < 0.064$), making it insignificant. Thus, TE fully mediated the relationship between PAL and TTI. The model B fit was excellent as indicated by various fit indices ($\chi^2/df = 1.905$, CFI = 0.991, NFI = 0.981, TLI = 0.987, RMSEA = 0.048).

Discussion

In the current article we highlighted school leadership research and practice in Oman, a context which is underrepresented in school leadership literature. We provided empirical data which explored the association between principal authentic leadership and teacher turnover intention and how teacher engagement may affect such relationship in Omani schools. The findings of our investigation pinpoint certain ideas that warrant discussion. Primarily, our empirical results showed that teachers who responded to our survey in Oman perceived their school principals as authentic leaders ($M = 4.00$; $SD = 1.04$). They also perceived themselves as highly engaged with school life activities ($M = 3.90$; $SD = 1.16$). These findings align with similar findings by Kulophas et al. (2018), who reported high engagement by school teachers in Thailand. The main finding about our empirical research questions indicated that principals' authentic leadership has is associated with a decrease in teachers' retention intentions. This result is consistent with other findings from other contexts (Gatling et al., 2016; Oh and Oh, 2017; Azanza et al., 2015; Qureshi and Aleemi, 2018; Kulophas et al., 2018) which emphasized that authentic leadership creates a buffer zone for employees who may consider quitting their jobs. Similarly, Oh et al. (2018) found a positive relationship between PAL and work engagement.



$$\chi^2=60.945; df=32; p=.002; CFI=.991; NFI=.981; RMSEA=.048; TLI=.987$$

Fig. 1 Teacher engagement as a mediator between PAL and TTI (model B). Note: Model (A) value is shown in parentheses.

The second research question investigated the impact of teacher engagement on their intention to leave the teaching profession. The teacher engagement was conversely associated with TTI ($p < 0.001$). This indicates that teacher engagement reduces turnover intentions. Similarly, Azanza et al. (2015) found a reverse impact of engagement on TTI, i.e., the higher the engagement at work, the lower the turnover intentions. In answering the third question, we examined the role of teacher engagement as a mediator between PAL and TTI. When we tested the mediational effect of teacher engagement on the association between PAL and TTI a full mediation was found. The inclusion of teacher engagement in the model made the association between PAL and TTI statistically not significant. This suggests that principal leadership influences teacher engagement, and that through teacher engagement, the principal authentic leadership is likely to reduce teacher's turnover intentions. This mediational effect of teacher engagement is in line with previous research. Azanza et al. (2015) found a mediational effect of teacher engagement on the relationship between authentic leadership and turnover intentions even though it was partial.

Overall, it could be argued that PAL contributes to teacher engagement in Oman. Engaging teachers in different roles and providing them with opportunities for continuous professional developments are essential for maintaining and sustaining school success (Hamad and Al Ani, 2016). This highlights the essential role of principals who are willing to succeed. Acting as authentic leaders is likely to help them convince teacher to engage to achieve school intended learning outcomes.

Our findings may be interpreted from the sociocultural perspective. Oman's culture which is informed by Islamic values has an impact on both individuals' professional roles in various contexts. Authentic leadership is often seen as a reflection of the trust in somebody who shows honesty, truthfulness, and genuineness to his followers in the same way the Prophet Mohamed was a role model for his followers (Aldaihani, 2020). Taken as such, school principals who advocate such characteristics, embrace transparency, and show intelligence and wisdom when leading their schools gain teachers' trust and contribute to teacher

engagement and affect teachers' turnover intentions. Perhaps this is true if we consider that in reality, the administrative system in Oman is characterized by high power distance due to centralism. School leaders tend to be the mouthpiece of decision makers and use their authority to ensure that mandates are implemented even though teacher may express discomfort. Therefore, leaders in Oman either become less integrative or embrace authentic leadership practices that paralleled with Islamic values.

Our investigation is not without limitations. First, our data was based on a randomly selected sample from the biggest city and capital of Oman. There are some cities that have different and special characteristics particularly in mountain areas. Samples from these areas need to be included in future studies. Also, this area is not preferred by the majority of teachers. Most of the teachers who are appointed in such areas leave pursue transfer to big cities and municipalities (Al Balushi et al., 2014). This was confirmed by the World Bank reports (World Bank Report, 2012). Second, we our investigation focused on testing the impact of on mediator, namely teacher engagement. Future investigations should include other possible mediators. Such as exposure to professional training on specific leadership practice in the Leadership Program by the SIPTT. This program has become compulsory for school principals and deputy principals to boost creating professional learning communities in Omani schools (Al Jabri et al., 2018). The program may have an impact of making school leaders authentic leaders and authentic practitioners. Third, our model needs to be tested again in other Arabic-speaking contexts to examine the bi-directional effects between the three variables of PAL, TTI, and teacher engagement. Finally, we relied on survey design to collection data, which has a common source variance. We relied on teachers' ratings of their school principals; future research could rely on multi-level modeling to assess separate effect at the individual (i.e., principals) and unit (i.e., school levels).

Conclusion

In the current article we highlighted school leadership in an Arab context, contextualizing the cultural, national, sociological aspects of school leadership practice in Oman, and provided empirical data on how teachers in Oman perceive authenticity in the school principals and how they perceive the effect of principal authentic leadership on teacher turnover intention. Additionally, we examined the role of teacher engagement in the relationship between principal authentic leadership and teacher turnover intention. The article argued that authentic leadership could be an effective leadership style that fits with the concept of Islamic leadership in Arab countries like Oman.

In Oman, authentic school principals have authentic values that are not compromised. Teachers trust authentic school principals because they display consistent behaviors, open communication, and empowering attitudes that enable teachers to achieve better outcomes at the personal, student, and school levels. They are those who understand that leadership is not about getting loyal teachers but about having empowered teachers and school staff. The study findings have important implications for Omani schools and for educational policy makers in Oman. First, authentic leadership is considered the blueprint of teacher engagement. Teachers feel empowered when they see authentic school principals. Training school principals on authentic leadership practices could be an important target for in-service training direct at school principals in Oman. Second, authentic leadership is close to the concept of Islamic leadership. Both center on having a leader that can build trust and that can transform his followers to authentic followers who can later lead others. School in Oman need this form of leadership particularly within the centralized system where teachers sometimes feel that they are just recipients of top-down decision and therefore, lack enthusiasm for taking actions or showing teacher agency for problem solving. Third, teacher turnover intention is a problem in Oman as it is in other places of the world. Perhaps one important finding of our investigation is the effect principal authentic leadership has on teacher turnover intention. When teacher see authentic models of leadership their engagement will increase, their intention to leave the job diminishes and they become determinant to have an effect on students and schools in which they serve. In summary, the investigation on authentic leadership in Oman could open a window for practitioners to discuss what it is that inspires teachers and make them possible authentic school leaders, thereby maximizing their influence in improving school effectiveness and student learning outcomes.

Acknowledgment

The data reported in this article is part of the research project that was funded by Ministry of Higher Education, Scientific Research & Innovation (MHECRI) (RC/RG-EDU/DEFA/19/01). The authors, therefore, would like to thank MHECRI for financial and technical support.

References

- Ahamed, F., Hassan, A., Hashim, J., 2013. Authentic leadership, trust, and employees' work engagement: a comparative study of Islamic and conventional banks in Malaysia. *J. Global Bus. Adv.* 6 (2), 152–166.
- Al Balushi, R., Idris, K., Omar, Z., Basri, R., Kumar, N., 2014. Public schools' characteristics and teacher turnover in Oman. *Int. J. Res. Stud. Educ.* 3 (1), 3–20.
- Aldaihani, S.G., 2020. Distributed leadership applications in high schools in the state of Kuwait from teachers' viewpoints. *Int. J. Leader. Educ.* 23 (3), 355–370.
- Al Jabri, M.J.H., Silvennoinen, H., Griffiths, D., 2018. Teachers' professional development in Oman: challenges, efforts and solutions. *Int. J. Learn. Teach. Educ. Res.* 17 (5), 82–103.
- Al Lawati, M.T., 2014. Leaving teaching: factors influencing Omani teachers. *Perspectives* 22 (1), 18–23.
- Al Shabibi, A.S., Silvennoinen, H., 2018. Challenges in education system affecting teacher professional development in Oman. *Athens J. Educ.* 5 (3), 261–282.

- Alazmi, A., Al-Mahdy, Y.F.H., 2020. Principal authentic leadership and teacher engagement in Kuwait's educational reform context. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/1741143220957339>.
- Al-Farsi, S.N., 2007. Omani School Head Teachers Views of Effectiveness of School Leadership of Secondary Schools in Oman. Doctoral dissertation. University of Glasgow.
- Al-Harathi, A.S.A., Al-Mahdy, Y.F.H., 2017. Distributed leadership and school effectiveness in Egypt and Oman: an exploratory study. *Int. J. Educ. Manag.* 31 (6), 801–813.
- Al-Mahdy, Y.F.H., Alazmi, A., 2021. Principal support and teacher turnover intention in Kuwait: implications for policymakers. *Leader. Pol. Sch.* <https://doi.org/10.1080/15700763.2021.1879168>.
- Al-Mahdy, Y.F.H., Emam, M.M., 2018. "Much ado about something" how school leaders affect attitudes towards inclusive education: the case of Oman. *Int. J. Incl. Educ.* 22 (11), 1154–1172.
- Al-Mahdy, Y.F.H., Al-Harathi, A.S., Salah El-Din, N.S., 2016. Perceptions of school principals' servant leadership and their teachers' job satisfaction in Oman. *Leader. Pol. Sch.* 15 (4), 543–566.
- Al-Mahdy, Y.F.H., Emam, M.M., Hallinger, P., 2018. Assessing the contribution of principal instructional leadership and collective teacher efficacy to teacher commitment in Oman. *Teach. Teach. Educ.* 69 (1), 191–201.
- Al-Neimi, R., Amzat, I.H., 2012. Teachers' perceptions of their job satisfaction and turnover: a case study of some schools in sultanate of Oman. *OIDA Int. J. Sustain. Dev.* 3 (9), 93–104.
- Al-Shehhi, S.A., Emam, M.M., Al-Otaiba, S., Ibrahim, M.M., Al-Mehrzi, R., 2019. Development of curriculum-based measurements in mathematical computations for Arab-speaking fourth grade students. *Sch. Psychol. Int.* 40 (2), 145–167.
- Azanza, G., Moriano, J.A., Molero, F., Mangin, J.P.L., 2015. The effects of authentic leadership on turnover intention. *Leader. Organ. Dev. J.* 36 (15), 955–971.
- Bailey, L., Purinton, T., Al-Mahdi, O., Al Khalifa, H., 2021. Conceptualizing school leadership in the Gulf Cooperation Council (GCC) cultures: demarcating challenges for research. *Educ. Manag. Adm. Leader* 49 (1), 93–111.
- Baron, R.M., Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Person. Soc. Psychol.* 51 (6), 1173–1182.
- Brislin, R.W., 1970. Back-translation for cross-cultural research. *J. Cross Cult. Psychol.* 1 (3), 185–216.
- Byrne, B.M., 2010. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*, second ed. Routledge, New York, NY.
- Carmeli, A., Gefen, D., 2005. The relationship between work commitment models and employee withdrawal intentions. *J. Manag. Psychol.* 20 (2), 63–86.
- Chin, W.W., Gopal, A., Salisbury, W.D., 1997. Advancing the theory of adaptive structuration: the development of a scale to measure faithfulness of appropriation. *Inf. Syst. Res.* 8 (4), 342–367.
- Cohen, A., Golan, R., 2007. Predicting absenteeism and turnover intentions by past absenteeism and work attitudes. *Career Dev. Int.* 12 (5), 416–432.
- Cohen, J., 1992. Statistical power analysis. *Curr. Dir. Psychol. Sci.* 1 (3), 98–101.
- Duignan, P.A., 2014. Authenticity in educational leadership: history, ideal, reality. *J. Educ. Adm.* 52 (2), 152–172.
- Education Council, 2020. *The Annual Report of Education in the Sultanate of Oman*. Education Council Press, Muscat.
- Emam, M., Al-Mahdy, Y.F.H., 2020. Building school capacity for inclusive education in the sultanate of Oman: a construct validation of the inclusive school climate scale. *Leader. Pol. Sch.* <https://doi.org/10.1080/15700763.2020.1770803>.
- Furman, G., 2012. Social justice leadership as praxis: developing capacities through preparation programs. *Educ. Adm. Q.* 48 (2), 191–229.
- Gardner, W.L., Coglisier, C.C., Davis, K.M., Dickens, M.P., 2011. Authentic leadership: a review of the literature and research agenda. *Leader. Q.* 22 (6), 1120–1145.
- Gatling, A., Kang, H.J.A., Kim, J.S., 2016. The effects of authentic leadership and organizational commitment on turnover intention. *Leader. Organ. Dev. J.* 37 (2), 181–199.
- Gill, J., Johnson, P., Clark, M., 2010. *Research Methods for Managers*. Sage Publications.
- Gillespie, N.A., Mann, L., 2004. Transformational leadership and shared values: the building blocks of trust. *J. Manag. Psychol.* 19, 588–607.
- Gumus, S., Bellibas, M.S., Esen, M., Gumus, E., 2018. A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educ. Manag. Adm. Leader* 46 (1), 25–48.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2010. *Multivariate Data Analysis*, seventh ed. Pearson Prentice Hall, NJ.
- Hallinger, P., Hammad, W., 2019. Knowledge production on educational leadership and management in Arab societies: a systematic review of research. *Educ. Manag. Adm. Leader* 47 (1), 20–36.
- Hamad, O.H., Al-Ani, W., 2016. School improvement in Oman: perceptions and practices. *J. Educ. Res. Rev.* 4 (1), 1–5.
- Hammad, W., Samier, E.A., Mohammed, A., 2020. Mapping the field of educational leadership and management in the Arabian Gulf region: a systematic review of Arabic research literature. *Educ. Manag. Adm. Leader.* <https://doi.org/10.1177/1741143220937308>.
- Hitchcock, T., 2019. A Study of the Relationship Between Authentic Leadership and Job Satisfaction in Public Education. Doctoral dissertation. Indiana State University.
- Human Development Report, 2019. <http://hdr.undp.org/sites/default/files/hdr2019.pdf>.
- Kim, H.K., 2014. Effect of Authentic Leadership on Employees' Attitudes, Behaviors, and Performances in a Korean Context. Doctoral dissertation. Oklahoma State University.
- Kline, R.B., 2011. *Principles and Practice of Structural Equation Modeling*, third ed. Guilford, New York, NY.
- Kulophas, D., Hallinger, P., Ruengtrakul, A., Wongwanich, S., 2018. Exploring the effects of authentic leadership on academic optimism and teacher engagement in Thailand. *Int. J. Educ. Manag.* 32 (1), 27–45.
- MacKinnon, D.P., Fairchild, A.J., 2009. Current directions in mediation analysis. *Curr. Dir. Psychol. Sci.* 18 (1), 16–20.
- Miller, H.E., Katerberg, R., Hulin, C.L., 1979. Evaluation of the Mobley, Horner, and Hollingsworth model of employee turnover. *J. Appl. Psychol.* 64 (5), 509–517.
- Mobley, W.H., Horner, S.O., Hollingsworth, A.T., 1978. An evaluation of precursors of hospital employee turnover. *J. Appl. Psychol.* 63 (4), 408.
- Mobley, W.H., Griffeth, R.W., Hand, H.H., Meglino, B.M., 1979. Review and conceptual analysis of the employee turnover process. *Psychol. Bull.* 86 (3), 493–522.
- MoE, 2018. *The National Report of the Results of Omani Students at TIMSS*. Publications of the Ministry of Education, Sultanate of Oman.
- MoE, 2019. *The Annual Educational Statistics Book (Issue 49): Educational Issues 464/2019*.
- Nasab, A.H., Afshari, L., 2019. Authentic leadership and employee performance: mediating role of organizational commitment. *Leader. Organ. Dev. J.* 40 (5), 548–560.
- Neider, L.L., Schriesheim, C.A., 2011. The authentic leadership inventory (ALI): development and empirical tests. *Leader. Q.* 22 (6), 1146–1164.
- New Zealand Education Consortium, 2017. *Evaluation of School System in Oman*. Ministry of Education, Muscat.
- Oh, J., Oh, S., 2017. Authentic leadership and turnover intention: does organizational size matter? *Leader. Organ. Dev. J.* 38 (7), 912–926.
- Oh, J., Cho, D., Lim, D.H., 2018. Authentic leadership and work engagement: the mediating effect of practicing core values. *Leader. Organ. Dev. J.* 39 (2), 276–290.
- Oman Population, 2020. *Oman Population Live*. Available at: <https://countrymeters.info/en/Oman>.
- Oplatka, I., Arar, K., 2017. The research on educational leadership and management in the Arab world since the 1990s: a systematic review. *Rev. Educ.* 5 (3), 267–307.
- Podsakoff, P.M., MacKenzie, S., Lee, J., Podsakoff, N.P., 2003. Common method biases in behavioral research. *MIS Q.* 30 (1), 115–141.
- Qureshi, M.A., Aleemi, A.R., 2018. Authentic leadership and turnover intention: mediating role of work engagement and job satisfaction in the healthcare sector of Pakistan. *Abasyn J. Soc. Sci.* 11 (2), 400–418.
- Schaufeli, W.B., Shimazu, A., Hakanen, J., Salanova, M., De Witte, H., 2019. An ultra-short measure for work engagement: the UWES-3 validation across five countries. *Eur. J. Psychol. Assess.* 35 (4), 577–591.
- Schreiber, J.B., Stage, F.K., King, J., Nora, A., Barlow, E.A., 2006. Reporting structural equation modelling and confirmatory factor analysis results: a review. *J. Educ. Res.* 99 (6), 323–337.
- Walumbwa, F.O., Avolio, B.J., Gardner, W.L., Wernsing, T.S., Peterson, S.J., 2008. Authentic leadership: development and validation of a theory-based measure. *J. Manag.* 34 (1), 89–126.
- World Bank Report, 2012. *Main Report (English). Education in Oman: The Drive for Quality*, vol. 2. World Bank Group, Washington, D.C.

The contributions of positive organizational studies to educational leadership and school improvement

Karen Seashore Louis^a and Joseph F. Murphy^b, ^aUniversity of Minnesota-Twin Cities, Minneapolis, MN, United States; and
^bVanderbilt University, Nashville, TN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	511
What do we know? A brief overview of educational research	511
What we know: some specifics about positive school leadership	512
Down in the weeds: more about positive behaviors	513
Exchange and values	514
Perceptions affect relationships	514
Transparency	514
Fairness	515
The consequences of relationship-building for school improvement	515
Developing agency	515
Identification with the school community	516
Organizational citizenship behavior	516
Job crafting and expert instruction	516
Quality matters in leader–member exchanges	517
Time and timing matter	517
Exchanges are always in context	517
Paying attention to what supports relationships	518
The importance of peers	518
Organizational outcomes of leader–member exchanges	518
Remembering reciprocity	518
Can any school leader do this?	519
References	519

Introduction

Over the past 50 years, it has become clear that relationships are at the foundation of student learning and development (Hartup, 1989; Pianta, 1999; Jennings and Greenberg, 2009; Chong et al., 2018; Epstein and Mcpartland, 1976; Walberg and Anderson, 1968; Entwistle and Ramsden, 1982). Positive organizational studies can make a significant contribution to educational leadership practice because they illuminate the ways in which leaders' work strengthens relationships, leading to increased motivation, commitment, engagement, and effort, all of which are associated with achieving the goal of student development and learning (Murphy and Louis, 2018).

This increasingly robust line of inquiry draws on positive psychology, which emerged in the mid-1990s (Seligman and Csikszentmihalyi, 2000; Gillham and Seligman, 1999). The applications of positive psychology quickly spread to organizational and leadership research (Cameron et al., 2003; Cameron and Spreitzer, 2012), and its development also overlapped with the attention to emotions in work settings, including schools (George, 2000; Barling et al., 2000; Fredrickson, 2001; Hargreaves, 1998, 2002). It was also aligned with other emerging research on school leadership practices that emphasized the importance of positive relationships within organizations, including transformational leadership (Leithwood and Jantzi, 1990; Leithwood and Steinbach, 1991), democratic and participatory leadership (Furman and Starrett, 2002; Weiss and Cambone, 1994), culturally-relevant leadership (Maxcy, 1998; Khalifa, 2018), along with others who advocated for a refocusing of schools on moral purpose (Sergiovanni, 1992). In sum, a positive psychology perspective on school leadership is not a different way of thinking about the work of running a school, but an integrative framework that may best serve as an “umbrella” under which a variety of scholars whose work is complementary may find a common space.

What do we know? A brief overview of educational research

Each of the trends briefly described above points to a shift in the role of school leaders, away from a narrow interpretation of instructional leadership that focuses almost exclusively on tested academic performance of pupils and toward a broader perspective, leadership for learning, which implies an integration of academic focus and a relationship focus (Murphy, 2013; Leithwood and Riehl,

2005). Leadership for learning also implies a responsibility for developing a school-wide culture (or climate) that fosters both academic and social development (Bryk et al., 1998; Newmann et al., 1989; Leithwood and Riehl, 2005).

Student authentic learning and socio-emotional development requires more than an improvement in technically competent instruction (Newmann and Associates, 1996) and is affected by a variety of within-school social factors such as classroom climate, peer relationships, and academic emphasis. In addition, supporting better learning environments for teachers and students requires creating facilitative school conditions (Wu et al., 2012; Lee and Louis, 2019). Principal leadership, furthermore, has a significant relationship with the school conditions that are clearly associated with student learning and development, such as teachers' sense of collective efficacy, the faculty's capacity for collective organizational learning, and commitment (Lee and Shute, 2010; Leithwood and Louis, 2011; Day et al., 2016). Nowhere is this more critical than in schools in challenging circumstances, where more is needed and expected from a school because many students come from families whose lives are more uncertain and chaotic, often for reasons outside their control (Bryk and Schneider, 2002; Goddard et al., 2001; Khalifa et al., 2016).

For example, research has focused on building high-trust school settings as a precursor to schools that work for all teachers and children (Hoy and Tschannen-Moran, 1999; Tschannen-Moran, 2009), and we have also learned that trust only develops within positive relationships (Price and Moolenaar, 2015; Tschannen-Moran, 2004). Some empirical research focuses on general perceptions of teacher–principal relationships as supportive or stimulating, as well as trusting (Fox et al., 2015; Brezicha et al., 2015; Musah et al., 2018). Studies of what principals actually do to develop the kind of relationships that are linked to school climate and student development and learning are also emerging (Hollingworth et al., 2017; Walker et al., 2021; Handford and Leithwood, 2013; Bodunrin et al., forthcoming), and have been shown to affect students but and school innovation (Bryk and Schneider, 2003; Louis, 2007; Moolenaar and Slegers, 2010).

A parallel line of work has emerged around principal caring (Smylie et al., 2016, 2019a; Louis et al., 2016; Louis and Murphy, 2017; Schechter and Shaked, 2017; Lomotey and Lowery, 2015; Eldor and Shoshani, 2016; Murphy and Torres, 2014; Beck, 1994). Again, there is evidence across countries and subcultures within countries that school leaders who express caring for both adults and students are more effective in shaping the processes and outcomes. In particular, accumulating research suggests that caring principals are more likely to contribute to school climate and instruction that is culturally responsive and supportive for students whose race/ethnicity or socioeconomic status have marginalized them (Lomotey and Lowery, 2015; Khalifa, 2018; Tichnor-Wagner and Allen, 2016; Wilson, 2016). Like trust, principal caring means wanting the best for students which often requires improvements in students' school experiences, which reflects an improvement-oriented school culture (Louis and Murphy, 2017; Kurland et al., 2010; Blase and Blase, 2000).

What we know: some specifics about positive school leadership¹

Some who make the argument for the crucial link between a principal's behaviors and leadership for learning have focused on the role of what might be called "executive leadership functions" such as motivating others, developing strategic visions and supportive structures, and hiring the right people, most of which are derived from the literature on large corporations (Kouzes and Posner, 2006; Collins, 2009). In contrast, the current weight given to the principal's role in supporting teachers assumes that he or she will spend much time discussing instruction and learning with individual teachers—hands-on, interpersonal interactions that focus on improving work routines including, but not limited to, instructional leadership (Marks and Printy, 2003). Thus, an alternative perspective suggests the explicit importance of the daily work of management and interactions with others that help to establish a school's culture (Schatz, 2003). This paper elaborates one foundational element of positive organizational research with a deep empirical base: leader–member exchanges in daily work that reinforce patterns of networked relationships within teams and the school more broadly. It will point to both the building blocks of routine relational exchange and articulate their consequences for school improvement. There is an emerging agreement about some of the key building blocks: namely that positive school leadership is:

1. *Focused on relationships as the foundation for the work of schools.* This is, of course, not a new observation (Miskel and Ogawa, 1988), but it nowhere is this more visible than in recent collections of cases of individual schools and leaders (Walker et al., 2021; Smylie et al., 2020). When aggregated to the group, a relational culture provides the foundation for asset-based experiences for all members (Burrello et al., 2016). Extensive empirical research has focused on relational trust and building high-trust school settings as a precursor to schools that work for all teachers and children (Hoy and Tschannen-Moran, 1999; Tschannen-Moran, 2009), and we have also learned that professional trust develops within positive personal relationships (Price and Moolenaar, 2015; Louis and Murphy, 2017). Some empirical research focuses on general perceptions of teacher–principal relationships as supportive or stimulating, as well as trusting (Fox et al., 2015; Brezicha et al., 2015; Musah et al., 2018), and studies that focus on how specific principals' actions help to develop the kind of relationships that are linked to school climate and student development and learning are beginning to emerge (Hollingworth et al., 2017; Walker et al., 2021; Rhodes et al., 2009)²

¹Much of this section focuses on educational research. The more extensive non-school scholarship that underlies this article is reviewed in Murphy, J., and Louis, K.S., 2018. *Positive School Leadership: Building Capacity and Strengthening Relationships*, New York, Teachers College Press.

²We acknowledge the many excellent studies that use the Multifactor Leadership Questionnaire to examine the relationship between principal behavior and school improvement. However, most of these use factors that focus on transformational leadership constructs rather than the more specific elements of principal behavior that we emphasize here.

2. *Grounded in an ethical perspective* on the purpose of schooling (and of leadership). This is, of course, fundamental to those who study ethical and value-based leadership in schools (Rosenblatt and Peled, 2002; Stefkovich and O'Brien, 2004; Norberg and Johansson, 2007), but many other leadership scholars identify ethical and value positions and values as equally fundamental to school improvement (Leithwood, 1999; Cherkowski et al., 2015) and developing a relationship-based culture. As noted earlier, in many professions an ethic of care is a central feature of ethical leadership principles (Smylie et al., 2016). Even though the virtuous leadership literature is littered with inconsistencies, this perspective informs us directly that "*school administration is a moral art*" (Greenfield, 2004) (p. 183), that positive school leaders "are more likely than others to apply moral principles and virtues to the situation to make sense of what action to take" (May et al., 2003) (p. 254). Integrity is a central aspect of positive school leadership as is "making the needs and interests of others primary" (Cooper et al., 2005) (p. 487). Humanity, altruism, trust, and service are powerful items in the virtue toolbox of positive school leaders, while many authors who write about virtues in school leadership would include empowerment of others (Leithwood and Riehl, 2005). In reviewing the positive leadership literature, Murphy and Louis (2018, pp. 29–44) synthesize nine principles, which are: asset-based, value-centered, virtue-based, transcendent, relationally grounded, focused on means as well as ends, growth-based, authentic, and service-grounded.
3. *Supportive of identification and belonging*. This theme, which emerges in all positive organization scholarship, is expressed strongly in schools—in part because schools are obligated to deal with children's emotional and behavioral development as well as cognitive growth. In particular, identification and belonging are associated with commitment to the educational purpose of schools for all members (Chan et al., 2008; Allen et al., 2018). The scholarly literature defines organizational identification in a variety of ways, but they have much in common: a feeling of openness or belonging (Beck, 1994; Mayeroff, 1971; Riley, 2013c) and inclusion (Murphy, 2016; Gonzales, 2019; Sluss and Ashforth, 2008). It is also commonly seen as a product of reciprocal leader–member identification (Avolio et al., 2004), and as a catalyst for teacher commitment (Geijsel et al., 2003) and student achievement (Reynolds et al., 2017). The longer-term development of strong interpersonal bonds and group identity increase collaboration and the willingness to take on additional roles (Angelle and Schmid, 2007) which has consequences for grade level/department teams and professional learning communities.
4. *Empowering, consistently building both self and collective efficacy over time*. Principals who empower their teachers and students take advantage of the human capital that already exists and foster the further development of that capital. One of the most consistent findings in the positive leadership research is that productive leader–member relationships increase sense of efficacy (Murphy and Louis, 2018), and have both a direct and indirect (through sense of efficacy) relationship to teacher commitment and effort (Ross and Gray, 2006). There is also ample evidence that teachers who feel empowered in their relationships with their principal and peers are more likely to experience both individual and collective efficacy while teachers' capacity to work closely with their administrators is associated with commitment to teaching (Ware and Kitsantas, 2007; Kurt et al., 2011; Goddard and Salloum, 2012; Goddard et al., 2015).
5. *Developmental, promoting continuous authentic learning for all*. For students, this means engaging classes where they feel a sense of responsibility for their own learning. The ability of formal school leaders to have an effect on classroom experiences is largely indirect, through teachers who are committed to this type of instruction (Wahlstrom and Louis, 2008; Herrington et al., 2014), which means replacing the often dreaded "professional development days" with opportunities that allow choice and increase willingness to experiment and take risks, both individual and collectively (Spreitzer, 2008; Schuttinger, 2021; Rule, 2006). There is ample evidence that school leaders have a significant effect on individual experimentation, through "job-crafting" and action research (Blase and Blase, 2000, 2002). In addition, multiple studies establish the role of the principal in promoting conditions that stimulate organizational or collective learning (Leithwood et al., 1998; Marks and Louis, 1999; Brown, 1995; Geijsel et al., 2009; Youngs and King, 2002; Supovitz, 2009; Louis and Lee, 2016). For both new and experienced teachers, it means individualized support as well as professional learning communities (Louis et al., 2016).
6. *Affected by specific features of the context*. While there are clearly uniformities in how schools are organized in modern society, community and national/local cultures influence how positive leaders behave, what teachers need in order to thrive, and how students respond to adult behavior (Hallinger et al., 1990; Hallinger and Murphy, 1986). Because the human context matters, positive school leaders must be attentive to community, whether it is urban or rural, stable or rapidly changing. Positive school leadership requires attention to changes in the demographic composition of their students' families and renewed recognition of socio-economic diversity that was always there, looking outward as well as inward (Smylie et al., 2020; Walker et al., 2021; Khalifa et al., 2016; Easton-Brooks et al., 2018; Riley, 2013a,b).

Down in the weeds: more about positive behaviors³

While a focus on assets is central, positive leadership has little to do with being a cheerleader. Rather, it is based on actions that cumulatively reinforce conditions under which all members of the school grow and flourish. Leadership is considered positive because it focuses on building assets rather than ameliorating problems (although all leaders need to do that as well ...).

³This section draws on Louis, K.S., and Murphy, J.F., 2019. El enfoque del liderazgo positivo y su relevancia para las relaciones. In: Weinstein, J., and Muñoz, G., (Eds.), *Liderazgo En Escuelas De Alta Complejidad Sociocultural*. Santiago, Chile: Universidad Diego Portales.

Much of the positive school leadership framework is based on the argument that schools, like other social institutions, function because there is agreement about appropriate relationships, typically formulated as social exchange theory (Emerson, 1976). There is obvious applicability to schools, since principals are expected to know and have some relationship with all the people who work in the school as well as many who are not daily members. The idea of relationships as a form of social exchange thus provided a foundation for understanding the ground on which trust emerges, by creating a “diffuse future obligations” reciprocity based on mutual goals (Sluss et al., 2008) (p. 458). In other words, from an exchange perspective, a leader does not “build trust” directly but engages with others on a routine basis in ways that develop *high quality, trusting* relationships over time.

But the “ah-ha” for leadership practice is the assumption that leaders form *unique relationships with each member of a group* and that *these relationships are reciprocal*. In other words, rather than focusing only on teachers as a group, principals should work with individuals in order to gain insights into how members work with each other and learn to appreciate the knowledge and insights that they gain from the people who work in their building. From an equity perspective, this may also require creating individualized relationships with parents and community members who may have experience exclusion or marginalization in educational settings (Banwo et al., forthcoming; Khalifa, 2018; Theoharis and O’toole, 2011).

By closely examining individual relationships, exchange theory points to the importance of each transaction that occurs in a school. This focus on details and individuals may seem impossible for many school leaders whose days are full with the endless array of job-related tasks (Pollock et al., 2015). However, increasing positive relationships does not always require more time, but the reflective use of the hundreds of interactions that they already undertake every day to solidify relationships and build trust.

Exchange and values

At first glance, a principal operating in a leader–member exchange framework appears instrumental, and possibly even manipulative. However, there are significant value-laden dimensions that are important in all work settings but particularly in schools, where everyone, from bus drivers to superintendents, tends to assume that their work is important because it supports children and youth. This presumption of value-based work allows every interaction to be an opportunity for verbal or non-verbal nurturing of shared values.

This observation is not new, and underpins many texts on school leadership (Sergiovanni, 1992, 1994; Deal and Peterson, 1994; Blumberg and Greenfield, 1986), as well as business (Chen et al., 2015). However, when opportunities are regularly missed, others will notice. In other words, the simple act of reinforcing the value basis of actions between self and other creates bonds that become part of the permanent fabric of a school’s life.

Since many of a principal’s dyadic encounters are visible, often when engaged in routine “management” work in the hall, lunchroom, or outside when greeting or saying goodbye to students and staff, they can be critical to setting expectations for how people work together (Lord et al., 1999). In contrast to a simple leader-centric perspective, the culmination of dyadic relationships helps to shape the way in which all members work with each other to carry out a jointly agreed upon task. In sum, the collective experience of dyadic relationships between formal school leaders and other members forms the basis for a leader’s impact on group climate and cohesiveness (Cogliser and Schriesheim, 2000).

Perceptions affect relationships

Managing other people’s perceptions is tricky business, but there is ample evidence that leaders often misjudge how others see them, and teacher deference to principals means that these are often not articulated (Kolb, 1995; Kolb and Bartunek, 1992; Bird et al., 2009). Yet we also know that leader–member relationships are critical to positive assessment of principal behavior and impacts (Price, 2011), and positive leaders are able to appreciate and understand the emotions of others as well as managing their own responses (Daus and Ashkanasy, 2005; Kotsou et al., 2019). In schools, everyone who is not a formal leader responds to (and influences) the leader–member exchanges discussed above—in fact, they are constantly making those assessments as they interact with their leader, particularly with respect to values and ethics (Mayer et al., 2009). Within a given school, a principal’s actions may be perceived as caring and visionary by some members, but flawed and deficient by others, and these differences in perception cause people to behave differently (Ehrhart, 2004). We illustrate the shifting ground of leader actions versus other’s perceptions in two areas that are considered essential to positive leadership: transparency and fairness. Both illustrate a key point: positive leadership is not a formula, but requires expert decision making in on-the-spot and ambiguous situations (Leithwood and Steinbach, 1995; Shapira and Stefkovich, 2016). Both are also particularly important during periods when there is a significant change or improvement being discussed or implemented.

Transparency

Transparency is a critical aspect of positive leadership that is seen by followers as ethical and moral (Hannah et al., 2005), and trust-inducing (Norman et al., 2010). The media consistently reminds us of the importance of transparency in the public sector, but in school settings, leaders who are more transparent are also more likely to be viewed as authentic (Srivastava and Dhar, 2019; Fox et al., 2015; Shapira-Lishchinsky and Tsemach, 2014; Bird et al., 2013). The idea of transparency can be politicized when conflated with debates about accountability (Koyama and Kania, 2014; Meijer, 2013; Hargreaves, 2015). Moreover, within a school, principals usually believe that they are transparent—but staff differences in their assessments of a school leader’s transparency are inevitable (Bird et al., 2009). Formal efforts to share information may have limited impacts unless it is also reflected in a general

willingness to be vulnerable and open within the group (Vogelgesang et al., 2013). At the same time, however, principal transparency can reduce or ameliorate within-school conflict (Klein, 2012) and can create a “culture of candor” (Bennis et al., 2008), while setting a critical groundwork for trust (Jiang and Luo, 2018).

Although important, principals can only strive to be as transparent as possible and keep testing the desire for more information. Transparency is particularly important during organizational crises or times of significant change, where it has significant impact on member’s assessment of a leader’s effectiveness (Norman et al., 2010), engagement and performance (Vogelgesang et al., 2013). Since a belief in transparency, like trust, takes time to develop, it is therefore an on-going positive leadership task.

Fairness

Fairness, like transparency, is a core element of positive psychology (Duff et al., 2016; Seligman et al., 2005) and key to teachers’ sense that they work in an ethical organization (Bogler and Somech, 2004). Both individual and groups who sense that procedures are fair are likely to put in effort beyond what is required and “they set a positive ethical standard to be followed across the organization” (Hannah et al., 2005) (p. 71). This has practical implications for carrying out key leadership obligations that are often fraught with conflict, such as allocation of scarce resources or personnel evaluations. For example, teachers who view the procedures and criteria used to assess them as fair tend to support teacher evaluation and use it for professional learning; those who see procedures as arbitrary or unfair do not (Vekeman et al., 2015; Tuytens and Devos, 2014). In addition, there are strong associations between teachers’ perceptions of their school as a just environment and their affective commitment to their school (Buluc and Gunes, 2014), as well as their willingness to work outside of their specific responsibilities to benefit the school—an essential aspect of improvement work (Ehrhart, 2004).

Fairness is, however, an area which is fraught for positive school leaders, particularly in the context of school improvement. In particular, there is the hoary issue of recurring incompatibility in different forms of fairness: procedural fairness (everyone treated alike), distributive fairness (those who need more get more) and interactional fairness (people feel equally well treated in specific situations) (Bacha and Walker, 2013). Procedural fairness is expected when it comes to teacher evaluation but, at the same time, it is recognized that there needs to be some differentiation between highly experienced and beginning teachers (Tuytens and Devos, 2019). Strict procedural fairness may be critical under conditions where trust is in question (such as an imposed new teacher evaluation), but less important in a high trust environment (Bacha and Walker, 2013). Procedural fairness needs to be balanced against specific individual needs (Cornelis et al., 2013) as well as school improvement goals, which may require allocating more resources to some teachers or groups than others. Even interactional fairness, which is an antecedent to group commitment and identity (de Lara, 2008), may be abrogated in cases of clear deviant behavior. In other words, perceived leader fairness is an essential component of school improvement—without it, teachers would be unwilling to engage in the additional effort and required collaboration—but is a mercurial aspect of positive leadership.

In addition, leaders almost always believe that they are consistently fair, while others in the organization have a different perception. Monitoring perceptions of fairness and transparency are, thus, particularly important in times of change.

The consequences of relationship-building for school improvement

Every work group is composed of many different types of relationships, but the principal–teacher relationship is particularly important for developing a wider group or organizational ethos. As noted above, as relationships deepen, they become more trusting and transformational, while instrumental exchange relationships have a narrow potential for impact on either party (Walumbwa et al., 2008). Reciprocity between principal and teacher supports the emerging agency and efficacy of both (Tourish, 2014). All evidence suggests that the effects of principals on student experiences and outcome occur primarily as a consequence of developing more engaged teachers (Hallinger and Heck, 1998; Leithwood and Louis, 2011). In other words, principals (and other school leaders) do not control the outcomes of their efforts to establish relationships with others. However, value-based social exchanges between principals and other school members can have multiplier effects when their dyadic relationships multiply, spread and change the organizational climate/culture. We will explore a few of the established outcomes of reciprocity on the potential for school improvement below.

Developing agency

Multiple publications have confirmed the importance of teacher agency, both individual and collective, to the success of school improvement efforts (Riveros et al., 2012; Datnow, 2012; Crandall et al., n.d.; Imants and Van Der Wal, 2020; Hökkä et al., 2017; Priestley et al., 2015). School-based research often identifies empowerment or shared leadership as the means of increasing teacher agency and teacher leadership, as well as improving classroom practices (Bogler and Somech, 2004; Marks and Louis, 1997; Harris, 2013). However, investigations also suggest that principal–teacher relationships provide a critical foundation for more formal initiatives to increase teachers’ contribution outside of their own classrooms. Successful teacher empowerment and shared leadership initiatives are premised on trusting relationships (Moye et al., 2005; Bush and Glover, 2012; Louis et al., 2009). Reciprocity suggests that as teachers’ agency increases, principals’ influence within the school and even in the community also grows. Empowered teachers lead to more influential leaders when relationships are caring and supportive (Printy et al., 2009).

Identification with the school community

To make a commitment to the extra effort of school improvement requires a sense of belonging in and identification with the organization. Isolated teachers may perform well in their classrooms but may make limited contributions to the overall work of the school. The relationship that a member of a group has with their supervisor is particularly important because it has “a significant influence on attitudes toward the organization, among them organizational identification” (Sluss and Ashforth, 2008) (p. 810). A study in China, for example, showed that employees who were engaged in helping relationships with their peers were also more likely to interact with a leader with whom they identified (Carnevale et al., 2019). This investigation is consistent with research in the US which indicates that teachers who see their principals as caring for them individually are more likely to work collaboratively with their peers and to support a socially just distribution of resources in the school (Louis et al., 2016). These findings suggest that if principals have reciprocal relationships with teachers, those teachers are likely to carry the values and ideals of the principal into their work with others.

In other words, collective identity has consequences. Sluss and Ashforth (2008) report that organizational identity—and the relationships that support it—are important because they are associated with how well the group manages tricky issues around diversity and social justice (p. 809) that are particularly critical in today’s schools. The longer-term development of strong interpersonal bonds and group identity increase collaboration and the willingness to take on additional roles (Angelle and Schmid, 2007) which has consequences for grade level/department teams and professional learning communities. It also suggests that principals who are focused on improvement for all students must consider how to create relationships that promote collective identification and authentic relationships with parents and other key actors (Banwo et al., forthcoming; Khalifa, 2018; Theoharis and O’toole, 2011; Riley, 2017).

At the same time, collective identity may constrain improvement initiatives when identity becomes challenged by environmental changes (Kavanagh et al., 2021). In schools, this often arises when there is rapid demographic change in a school’s student population, which may result in teachers collectively clinging to practices that “worked” for them in the past (Hargreaves and Goodson, 2006). Where teachers’ collective identity is not rooted in high quality relationships with their principal, may become oppositional, and undermine school effectiveness (Smylie et al., 2007).

Organizational citizenship behavior

School improvement requires additional initiative from members—typically work that lies outside of the formal expectations associated with the job. Organizational citizenship behavior refers to the “extent to which employees engage in behaviors beyond their prescribed roles” (Ilies et al., 2007) (p. 270). In other words, people do not engage in OCB to enhance personal benefits but because they have a positive sense of identity with the larger organization and its members (Mcneely and Meglino, 1994). The degree to which this kind of pro-social behavior occurs is deeply influenced by leader–member exchanges and perceptions of the ethical climate expectations developed by leaders, a conclusion that is affirmed in studies in the US and other nations (Hackett and Lapiere, 2004; Newman et al., 2017).

There is a wealth of scholarship on organizational citizenship behavior, and its application in educational leadership research is increasing (Dipaola and Hoy, 2013; Nutov and Somech, 2017; Somech and Ron, 2007). Educational leaders in all countries should be concerned with organizational citizenship behavior because it is associated with school climate and student learning (Dipaola and Hoy, 2005; Dipaola and Tschannen-Moran, 2014)—but also with the extra work that comes during innovation and organizational change (Bogler and Somech, 2004).

Job crafting and expert instruction

Educational research pays considerable attention to the idea of organizational learning, or the school-wide capacity to search for and incorporate new ideas (Schechter and Qadach, 2012; Leithwood et al., 1998; Louis, 2008). Even studies of instructional leadership tend to pay limited attention to the details of how principal–teacher relationships affect instructional practice.

But evidence from studies in other work settings suggest that research on job-crafting may provide clues about how principal–teacher relationships affect classrooms. Job crafting refers to the ways in which people “redesign” their work by adding or subtracting activities, or changing other aspects of timing and performance (Berg et al., 2008). In other words, it is grounded in the well-known routines of do-observe-reflect. Job-crafting is the way in which teachers and other professionals become expert in their work (Goller and Billett, 2014). Based on a large study, Marks and Printy (2003) concluded that “[w]hen teachers perceive principals’ instructional leadership behaviors to be appropriate, they grow in commitment, professional involvement, and *willingness to innovate*” (Marks and Printy, 2003) (p. 393. Italics added).

An aspect of individual job crafting that is critical for schools is that it often spills over to teams and professional learning communities (Tims et al., 2013; Schuttinger, 2021), which is where new ideas about teaching and developing students are most likely to be discussed. In other words, schools where teachers are consistently redesigning their classroom work are also likely to be active redesigning and sharing joint improvement efforts, a finding affirmed in multiple countries (Sargent, 2014; McLaughlin and Talbert, 2006; Stoll et al., 2006).

Quality matters in leader–member exchanges

Much of leadership theory (and the “wisdom” of consultants) depends on relatively simply but uniform rules. In real life, however, good principals are not rule-bound but adaptive (Heifetz, 1994). Authentic leaders take into account both individual preferences and the particular contexts of leader–member exchanges (Sternberg and Vroom, 2002; Louis et al., 2016). Exchange theory’s focus on dyads thus presumes that relationships are situational and personal. *Quid pro quo* relationships are usually low-quality and indicate a tenuous mutual commitment, while social exchange relationships tend to be “more open-ended, long-term, and less in need of an immediate ‘payoff’” (Walumbwa et al., 2011) (p. 734). There is no consensus on the full range of underlying socially valuable consequences of “high quality” leader–member exchange relationships. Often noted outcomes include: affect, loyalty, contribution, and professional respect (Ilies et al., 2006) along with mutual affection, contribution to mutual goals, and professional respect (Liden and Maslyn, 1998). But there should be no mistake: Achieving adaptive, authentic, high quality relationships on a daily basis in any school is a challenging task.

In education, the consequences of lack of individual sensitivity are serious: teachers who leave a school report that dissatisfaction with their principal is a major reason (Grissom, 2011; Boyd et al., 2011). However, not all teachers in a school will have the same reaction to a principal: any individual may have devoted followers as well as those who are more ambivalent or detached. This variability makes the work of establishing positive relationships particularly complex. Consider the implications of the statement that “leaders form different types of exchange with respective followers” (Liden et al., 2008) (p. 163).

Time and timing matter

The development of solid relationships varies not only by individual, but also over time. A new principal coming into a school where a beloved former principal has retired has one job; the new principal entering a school that was led by an incompetent or abusive leader has a different task. Both, however, must start at the beginning. Even established leaders confront inevitable turnover. One may have excellent social exchange relationships with an existing group, but a principal must bring each fresh recruit (who enters with their own set of experiences with good or bad leadership) into a unique leader–member exchange. Of course, it helps if the staff tells the new teacher that “our principal is great”, but socialization, by definition, develops through experience rather than new employee orientation or school handbooks. Wang et al. (2005) affirm a developmental approach, with three developmental stages (stranger, acquaintance, partner) that become successively more transformational (p. 423). In other words, shifting the basis of a relationship from professional obligation to collaboration is a process whose timing will vary depending on the people involved.

Four practical conclusions are of particular importance to school leaders.

- (1) *Principals should recognize differences in experience and skills.* For example, a first-year teacher requires a different kind of role-making signals than an experienced teacher–leader.
- (2) *Earlier stages in establishing a relationship, where the leader attempts to discover the relevant talents, motivations, and limits of the member, are critical.* Later stages, where there are clear mutual understandings within an established relationship, require less time (Nahrgang et al., 2009). Of course, this will have a big impact on how a principal spends his or her time depending on the frequency of staff and principal turnover in a school.
- (3) *Member views of their relationship with a leader tend to become more aligned over time:* a principal who can establish good relationships with a core group of teachers is likely (unless they represent a coalition in a conflict-filled school) may extend this to teachers who may be more reluctant to interact frequently with their formal supervisors.
- (4) Unlike many leadership narratives, individualized relationships between leaders and others must also attend to reciprocity, including how member characteristics and reactions affect leader perceptions of the members (Nahrgang et al., 2009). One specific example is that principals who see their teachers (collectively) as professionally competent are likely to have stronger individual social bonds with their teachers (Louis and Murphy, 2017).

Exchanges are always in context

We must remember that a focus on the development of relationships is not abstract theory, but the result of how real people react to real exchange settings. *At the micro level, every exchange involves an action and a reaction.* A leader acts (the principal calls a meeting); followers read and interpret those actions within the immediate context (“her meetings are usually short and useful” vs. “another wasted 45 min”). These actions and reactions have a further impact on followers in related contexts and the cycle is re-engaged as leaders respond to followers engaged or disengaged responses. In other words, the potentially transformational relationship is strengthened or eroded based on smaller transactional events. In the end, there is an organizational response based on a myriad of individual experiences in school relationships.

So how can principals help to make others see their actions as “high quality”, responsive, and adaptive? Research suggests that they must keep one eye on their daily interactions, keeping in mind the obvious generalities—reaching out whenever possible and offering care along with respect, opportunity, and trust (Walumbwa et al., 2009; Smylie et al., 2016, 2019b).

Paying attention to what supports relationships

The focus on individual relationship building will be of little use if school leaders do not attend to what supports the work of others in the school. They can also ask about what is getting in the way of teachers', counselors' or food staff's best work and adjust norms, teams, and work schedules to make sure that unnecessary barriers are removed—part of ongoing task of redesigning the organization (Walumbwa et al., 2011; Geijsel et al., 2003; Leithwood and Jantzi, 2005). In other words, a principal who wonders why the school's two-year old professional learning communities initiative has yielded very little might ask whether teachers were reluctant to tell her that 45 minutes of common meeting time was insufficient to move beyond procedural work, as well as revealing previously unacknowledged conflicts among staff (Voulalas and Sharpe, 2005). Finally, research suggests that leaders also need to pay attention to indirect supports for other members of the school, working on structures that can be almost invisible until they fail (Sluss et al., 2008). In other words, relationship building provides the base for and becomes intertwined with information gathering/asset building those principals need to due to create a high-functioning, high performing school.

The importance of peers

Schools are comprised of many individual members who also have relationships with each other. It has long been acknowledged that peer groups dominate the day-to-day life of any work group (Blau, 1955) which can have both positive and negative effects on leader–member relationships (Seashore, 1954; Podsakoff et al., 1997). As noted above, cohesive teacher groups who are strongly identified with the school can be supportive of collective goals—or may play an oppositional role. However, in general, it is well established that strong ties among teachers will have a positive effect on their satisfaction and improvement efforts (Daly et al., 2010; Louis et al., 1996; Banerjee et al., 2017). In addition, groups that have stronger ties among the members are more likely to exhibit positive behaviors that support both the organization and individuals (Cogliser and Schriesheim, 2000).

Informal teacher leaders emerge (Anderson, 2004) and effective principals often acknowledge their importance by supporting them in ways that enhance overall group effectiveness, cohesiveness, energy and effectiveness. Leithwood et al. (2007) refer to this as “taking the ego out of the system” (Leithwood et al., 2007) (p. 37). This does not reduce the importance of leader–member exchanges but acknowledges that effective principals (and formally appointed teacher leaders) need to understand not only the individual members who work in their setting, but also the informal organization. This is particularly important during any school-wide improvement initiative, where informal leaders often play a significant role (Sun et al., 2013), as well as their impact on job-crafting and individual improvement (Gordon et al., 2020).

Organizational outcomes of leader–member exchanges

There is a clear developmental cycle in a principal's relationship with others in the school that leads from individual interactions to climate/culture to commitment to the school. Research suggests that when leaders demonstrate both their *expertise* and their grounding in the *positive leadership values*, the following important organizational outcomes are likely (Murphy and Louis, 2018):

- Leader–member exchanges lead to valued *reciprocal outcomes*
- Which leads to *member satisfaction with the leader*, and *Commitment to their shared vision and rewards*,
- Which reinforces and deepens an *emotional connection* with and *trust* in the leader, leading to
- *Task performance, student learning, and more individual and collective effort extended to innovation and improvement.*

Remembering reciprocity

Critically important is the development of wide-spread reciprocity—a fundamental antecedent of the informal “psychological contract” that encourages both teacher and leader engagement in creating an organization based in positive functioning (Coyle-Shapiro, 2002; Bakker and Schaufeli, 2008). The research community often ignores the chicken-and-egg quality of the leader–member relationship, which tends to see leader behavior as a cause of other member's behavior, but solid teacher–principal relationships are based on a teacher's willingness to provide positive input to the relationship. Responsive teachers who work well with their principals encourage the principal's repeated investment in relationships. This, of course, will reinforce future value-centered and asset-based behavior on the part of the leader.

The positive impact of teachers on principal's effectiveness is under researched, but scholarship in other work settings provides evidence that “the interpersonal relation of supervisors to subordinates affects subordinates' identification and self-concepts, which in turn, are critical determinants of social and organizational processes” (Lord et al., 1999) (p. 170). In schools, teachers' extra efforts (organizational citizenship behavior) may induce the principal to increase his or her perception of the teacher's professionalism and provide additional “rewards” that give that teacher additional learning opportunities (Liden et al., 2008). Translated across many teachers, this outcome of individual leader–member exchanges may lead to increased capacity for organizational learning (Louis and Murphy, 2017), “job crafting” and investment in improvements. It is also connected to the tendency of employees to grant their leaders more latitude in acting, or a bank of “idiosyncratic credits” that allow leaders to make an occasional mistake without sacrificing all of the goodwill that has been built (Hannah et al., 2005). All of these latter points attend to the need for more research on teacher reciprocation.

Can any school leader do this?

Busy principals may shudder at the idea of adding enhanced leader-member exchange to their already full plate, but we argue that it requires only more attentiveness to the small acts that fill their day and can often be implicitly or explicitly tied to improvement objectives. Colleagues, students and others in the school are already watching, and each exchange will be filtered through the individual's lenses which, according to many of the authors we have cited, will attend to their *authenticity, meaningfulness, challenge, and relevance*. Only when those are in play will teachers reciprocate with honest feedback, which is essential to uncovering assets and locating barriers that need to be removed (Drazin and Joyce, 1979; Schein, 1999).

To be effective, the specific behaviors, words, and actions that elicit a positive response in others will vary between contexts and cultures. Much of the research that underpins this article has been carried out in the United States, but there is no evidence that we can find from studies conducted in other countries that the principles and concepts underlying positive school leadership cannot be translated to other cultures. However, we know that people's perceptions of leaders and how they prefer to relate to them vary greatly between countries (Hofstede, 2001). This is, of course, why "off the shelf" recommendations for how to reform schools often require a great deal of adjustment when they move from one country, culture or context to another. But principals, whether introverts or extraverts, must recognize that they need to get out of their offices, initiate exchanges, and do so in manner that reinforces assets. More important than personal characteristics are a strategic behavioral commitment to make every exchange, no matter how brief, count.

We conclude by returning to the distracting dichotomy between "management" and "leadership" that has permeated the leadership literature for decades, leading to the erroneous conclusion that leadership consists primarily of planning and executive action. Based on our experience working with school leaders and the positive organizational research, we argue that:

- (1) The capacity of school leaders to be influential rests on the way in which others in the organization perceive them.
- (2) The perceptions of others will be fundamentally based on accumulated observations and personal relationships that induce trust, respect, identification and ownership.
- (3) School leaders will find that the easiest way to induce these positive responses is through attention to interpersonal exchanges during routine management activities that bring them into contact with others.
- (4) Positive exchanges do not have to reinforce the status quo but can provide incentives for continuous school improvement.

This does not imply that leadership work stops with the development of good exchanges, but without a solid ground, the best improvement initiatives will be unsuccessful.

References

- Allen, K., Kern, M.L., Vella-Brodick, D., Hattie, J., Waters, L., 2018. What schools need to know about fostering school belonging: a meta-analysis. *Educ. Psychol. Rev.* 30, 1–34.
- Anderson, K.D., 2004. The nature of teacher leadership in schools as reciprocal influences between teacher leaders and principals. *Sch. Effect. Sch. Improv.* 15, 97–113.
- Angelle, P.S., Schmid, J.B., 2007. School structure and the identity of teacher leaders: perspectives of principals and teachers. *J. Sch. Leader.* 17, 771–799.
- Avolio, B.J., Weichun, Z., William, K., Puja, B., 2004. Transformational leadership and organizational commitment: mediating role of psychological empowerment and moderating role of structural distance. *J. Organ. Behav.* 25, 951–968.
- Bacha, E., Walker, S., 2013. The relationship between transformational leadership and followers' perceptions of fairness. *J. Bus. Ethics* 116, 667–680.
- Bakker, A.B., Schaufeli, W.B., 2008. Positive organizational behavior: engaged employees in flourishing organizations. *J. Organ. Behav.* 29, 147–154.
- Banerjee, N., Stearns, E., Moller, S., Mickelson, R.A., 2017. Teacher job satisfaction and student achievement: the roles of teacher professional community and teacher collaboration in schools. *Am. J. Educ.* 123, 203–241.
- Banwo, B., Khalifa, M., Louis, K.S., (forthcoming). Exploring trust: culturally responsive and positive school leadership. *J. Educ. Adm.*
- Barling, J., Slater, F., Kevin Kelloway, E., 2000. Transformational leadership and emotional intelligence: an exploratory study. *Leader. Organ. Dev. J.* 21, 157–161.
- Beck, L.G., 1994. *Reclaiming Educational Administration as a Caring Profession*. Teachers College Press, New York.
- Bennis, W., Goleman, D., O'toole, J., 2008. *Transparency: How Leaders Create a Culture of Candor*. John Wiley & Sons.
- Berg, J.M., Dutton, J.E., Wrzesniewski, A., 2008. What is Job Crafting and Why Does it Matter. Retrieved from the website of Positive Organizational Scholarship on April, 15, 2011.
- Bird, J.J., Wang, C., Watson, J.R., Murray, L., 2009. Relationships among principal authentic leadership and teacher trust and engagement levels. *J. Sch. Leader.* 19, 153–171.
- Bird, J.J., Dunaway, D.M., Hancock, D.R., Wang, C., 2013. The superintendent's leadership role in school improvement: relationships between authenticity and best practices. *Leader. Pol. Sch.* 12, 77–99.
- Blase, J., Blase, J., 2000. Effective instructional leadership: teachers' perspectives on how principals promote teaching and learning in schools. *J. Educ. Adm.* 38, 130–141.
- Blase, J., Blase, J., 2002. Teachers' perceptions of principals' instructional leadership and implications. *Leader. Pol. Sch.* 1, 256–264.
- Blau, P.M., 1955. *The Dynamics of Bureaucracy: A Study of Interpersonal Relations in Two Government Agencies*. University of Chicago Press, Chicago.
- Blumberg, A., Greenfield, W., 1986. *The Effective Principal: Perspectives on School Leadership*. Allyn & Bacon, Newton, MA.
- Bodunrin, B., Khalifa, M., Louis, K., (forthcoming). Reconciling trust and power in culturally responsive and positive school leadership. *J. Educ. Adm.*
- Bogler, R., Somech, A., 2004. Influence of teacher empowerment on teachers' organizational commitment, professional commitment and organizational citizenship behavior in schools. *Teach. Teach. Educ.* 20, 277–289.
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., Wyckoff, J., 2011. The influence of school administrators on teacher retention decisions. *Am. Educ. Res. J.* 48, 303–333.
- Brezicha, K., Bergmark, U., Mitra, D.L., 2015. One size does not fit all: differentiating leadership to support teachers in school reform. *Educ. Adm. Q.* 51, 96–132.
- Brown, J.M., 1995. Organizational transformation concepts and collaborative multi-school district efforts to improve educational programs for students with special learning needs. *J. Vocat. Spec. Needs Educ.* 18, 28–34.
- Bryk, A.S., Schneider, B., 2002. *Trust in Schools: A Core Resource for Improvement*. Russell Sage, New York.
- Bryk, A.S., Schneider, B., 2003. Trust in schools: a core resource for school reform. *Educ. Leader* 60, 40–45.
- Bryk, A.S., Sebring, P., Kerbow, D., Rollow, S.G., Easton, J., 1998. *Charting Chicago School Reform: Democratic Localism as a Lever for Change*. Westview Press, Boulder, Colo.
- Buluc, B., Gunes, A.M., 2014. Relationship between organizational justice and organizational commitment in primary schools. *Anthropologist* 18, 145–152.
- Burrello, L.C., Beitz, L.M., Mann, J.L., 2016. *A Positive Manifesto: How Appreciative Schools Can Transform Public Education*. Elephant Rock Books, Ashford, CT.

- Bush, T., Glover, D., 2012. Distributed leadership in action: leading high-performing leadership teams in English schools. *Sch. Leader. Manag.* 32, 21–36.
- Cameron, K., Spreitzer, G. (Eds.), 2012. *The Oxford Handbook of Positive Organizational Scholarship*. Oxford University Press, New York.
- Cameron, K., Dutton, J.E., Quinn, R. (Eds.), 2003. *Positive Organizational Scholarship*. Berrett-Kohler, San Francisco.
- Carnevale, J.B., Huang, L., Paterson, T., 2019. LMX-differentiation strengthens the prosocial consequences of leader humility: an identification and social exchange perspective. *J. Bus. Res.* 96, 287–296.
- Chan, W.Y., Lau, S., Nie, Y., Lim, S., Hogan, D., 2008. Organizational and personal predictors of teacher commitment: the mediating role of teacher efficacy and identification with school. *Am. Educ. Res. J.* 45, 597–630.
- Chen, Z., Zhu, J., Zhou, M., 2015. How does a servant leader fuel the service fire? A multilevel model of servant leadership, individual self identity, group competition climate, and customer service performance. *J. Appl. Psychol.* 100, 511.
- Cherkowski, S., Walker, K.D., Kutsyuruba, B., 2015. Principals' moral agency and ethical decision-making: toward a transformational ethics. *Int. J. Educ. Policy Leadersh.* 10, n5.
- Chong, W.H., Liem, G.A.D., Huan, V.S., Kit, P.L., Ang, R.P., 2018. Student perceptions of self-efficacy and teacher support for learning in fostering youth competencies: roles of affective and cognitive engagement. *J. Adolesc.* 68, 1–11.
- Cogliser, C.C., Schriesheim, C.A., 2000. Exploring work unit context and leader-member exchange: a multi-level perspective. *J. Organ. Behav.* 21, 487–511.
- Collins, J., 2009. *Good to Great: Why Some Companies Make the Leap and Others Don't*. SAGE, Thousand Oaks.
- Cooper, C.D., Scandura, T.A., Schriesheim, C.A., 2005. Looking forward but learning from our past: potential challenges to developing authentic leadership theory and authentic leaders. *Leader. Q.* 16, 475–493.
- Cornelis, I., Van Hiel, A., de Cremer, D., Mayer, D.M., 2013. When leaders choose to be fair: follower belongingness needs and leader empathy influences leaders' adherence to procedural fairness rules. *J. Exp. Soc. Psychol.* 49, 605–613.
- Coyle-Shapiro, J.A.M., 2002. A psychological contract perspective on organizational citizenship behavior. *J. Organ. Behav.* 23, 927–946.
- Crandall, D., Datnow, A., Lieberman, A., Supovitz, P., (n.d.). *Organizing Success: Dimensions of Creative Operationalism in Networked Learning Communities*. Nottingham: National College for School Leadership.
- Daly, A.J., Moolenaar, N.M., Bolivar, J.M., Burke, P., 2010. Relationships in reform: the role of teachers' social networks. *J. Educ. Adm.* 48, 359–391.
- Datnow, A., 2012. Teacher agency in educational reform: lessons from social networks research. *Am. J. Educ.* 119, 193–201.
- Daus, C.S., Ashkanasy, N.M., 2005. The case for the ability-based model of emotional intelligence in organizational behavior. *J. Organ. Behav.* 26, 453–466.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52, 221–258.
- de Lara, P.Z.M., 2008. Fairness, teachers' non-task behavior and alumni satisfaction: the influence of group commitment. *J. Educ. Adm.* 46, 514–538.
- Deal, T.E., Peterson, K.D., 1994. *The Leadership Paradox: Balancing Logic and Artistry in Schools*. Jossey-Bass Education Series. Jossey Bass, San Francisco.
- Dipaola, M., Hoy, W.K., 2005. Organizational citizenship of faculty and achievement of high school students. *High Sch. J.* 88, 35–44.
- Dipaola, M., Hoy, W.K., 2013. *Principals Improving Instruction: Supervision, Evaluation, and Professional Development*. Pearson/Allyn and Bacon, Boston.
- Dipaola, M., Tschannen-Moran, M., 2014. Organizational citizenship behavior in schools and its relationship to school climate. *J. Sch. Leader.* 11, 424–447.
- Drazin, R., Joyce, W.F., 1979. Towards a typology of resistance to change behaviors. *Acad. Manag. Proc.* 304–308.
- Duff, J., Rubenstein, C., Prilleltensky, I., 2016. Wellness and fairness: two core values for humanistic psychology. *Humanist. Psychol.* 44, 127.
- Easton-Brooks, D., Robinson, D., Williams, S.M., 2018. Schools in transition: creating a diverse school community. *Teach. Coll. Rec.* 20, 1–26.
- Ehrhart, M.G., 2004. Leadership and procedural justice climate as antecedents of unit-level organizational citizenship behavior. *Person. Psychol.* 57, 61–94.
- Eldor, L., Shoshani, A., 2016. Caring relationships in school staff: exploring the link between compassion and teacher work engagement. *Teach. Teach. Educ.* 59, 126–136.
- Emerson, R.M., 1976. Social exchange theory. *Annu. Rev. Sociol.* 2, 335–362.
- Entwistle, N., Ramsden, P., 1982. *Understanding Student Learning* (Routledge Revivals). Social Science Research Council, London.
- Epstein, J.L., Mcpartland, J.M., 1976. The concept and measurement of the quality of school life. *Am. Educ. Res. J.* 13, 15–30.
- Fox, J., Gong, T., Attoh, P., 2015. The impact of principal as authentic leader on teacher trust in the K-12 educational context. *J. Leader. Stud.* 8, 6–18.
- Fredrickson, B.L., 2001. The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am. Psychol.* 56, 218–226.
- Furman, G., Starrett, R., 2002. Leadership for democratic community. In: Murphy, J. (Ed.), *The Educational Leadership Challenge: Redefining Leadership for the 21st Century*. University of Chicago Press, Chicago.
- Geijsel, F., Sleegers, P., Leithwood, K., Jantzi, D., 2003. Transformational leadership effects on teachers' commitment and effort toward school reform. *J. Educ. Adm.* 41 (3), 228–256.
- Geijsel, F., Sleegers, P.J.C., Stoel, R.D., Krüger, M.L., 2009. The effect of teacher psychological and school organizational and leadership factors on teachers' professional learning in Dutch schools. *Elem. Sch. J.* 109, 406–427.
- George, J.M., 2000. Emotions and leadership: the role of emotional intelligence. *Hum. Relat.* 53, 1027–1055.
- Gillham, J.E., Seligman, M.E., 1999. Footsteps on the road to a positive psychology. *Behav. Res. Ther.* 37, S163.
- Goddard, R., Salloum, S., 2012. Collective efficacy beliefs, organizational excellence, and leadership. In: Cameron, K.S., Spreitzer, G.M. (Eds.), *Oxford Handbook of Positive Organizational Scholarship*. Oxford, New York.
- Goddard, R., Tschannen-Moran, M., Hoy, W., 2001. A multi-level examination of the distribution and effects of teacher trust in students and parents in urban elementary schools. *Elem. Sch. J.* 101, 3–19.
- Goddard, R., Goddard, Y., Sook Kim, E., Miller, R., 2015. A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *Am. J. Educ.* 121, 501–530.
- Goller, M., Billett, S., 2014. Agentic behavior at work: crafting learning experiences. In: Harteis, C., Rausch, A., Seifried, J. (Eds.), *Discourses on Professional Learning*. Springer, Dordrecht.
- Gonzales, S.M., 2019. Cultivating familismo: belonging and inclusion in one Latina/o learning community. *Int. J. Incl. Educ.* 23, 937–949.
- Gordon, S.P., Jacobs, J., Croteau, S.M., Solis, R., 2020. Informal teacher leaders: who they are, what they do, and how they impact teaching and learning. *J. Sch. Leader.* <https://doi.org/10.1177/1052684620924468>.
- Greenfield, W.D., 2004. Moral leadership in schools. *J. Educ. Adm.* 42, 174–196.
- Grissom, J., 2011. Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teach. Coll. Rec.* 113, 2552–2585.
- Hackett, R.D., Lapierre, L.M., 2004. A Meta-analytical Explanation of the Relationship Between LMX and OCB. *Academy of Management Proceedings*. Academy of Management, Briarcliff Manor, NY, pp. T1–T6.
- Hallinger, P., Heck, R., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Effect. Sch. Improv.* 9, 157–191.
- Hallinger, P., Murphy, J.F., 1986. The social context of effective schools. *Am. J. Educ.* 94, 328–355.
- Hallinger, P., Bickman, L., Davis, K., 1990. In: Publication Sales, National Center for Educational Leadership, Harvard Graduate School of Education (Eds.), *What Makes a Difference? School Context, Principal Leadership, and Student Achievement*. National Center for Educational Leadership, Nashville, TN.
- Handford, V., Leithwood, K., 2013. Why teachers trust school leaders. *J. Educ. Adm.* 51, 194–212.
- Hannah, S.T., Lester, P.B., Vogelgesang, G.R., 2005. Moral leadership: explicating the moral component of authentic leadership. *Leadersh. Manag.* 3, 43–81.
- Hargreaves, A., Goodson, I., 2006. Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educ. Adm. Q.* 42, 3–41.

- Hargreaves, A., 1998. The emotional politics of teaching and teacher development: with implications for educational leadership. *Int. J. Leader. Educ.* 1, 315–336.
- Hargreaves, A., 2002. The emotional geographies of teaching. *Teach. Coll. Rec.* 103, 1056–1080.
- Hargreaves, A., 2015. Autonomy and transparency: two good ideas gone bad. In: Evers, J., Kneiber, R. (Eds.), *Flip the System: Changing Education From the Ground Up*. Routledge, New York.
- Harris, A., 2013. Teacher leadership and school improvement. In: Harris, A., Day, C., Hopkins, D., Hadfield, M., Hargreaves, A., Chapman, C. (Eds.), *Effective Leadership for School Improvement*. Routledge, Abingdon, UK.
- Hartup, W.W., 1989. Social relationships and their developmental significance. *Am. Psychol.* 44, 120–126.
- Heifetz, R., 1994. *Leadership Without Easy Answers*. Belknap, Cambridge, MN.
- Herrington, J., Reeves, T.C., Oliver, R., 2014. Authentic learning environments. In: *Handbook of Research on Educational Communications and Technology*, pp. 401–412.
- Hofstede, G., 2001. *Cultures Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Sage, Thousand Oaks, CA.
- Hökkä, P., Vähäsantanen, K., Mählakaarto, S., 2017. Teacher educators' collective professional agency and identity—transforming marginality to strength. *Teach. Teach. Educ.* 63, 36–46.
- Hollingworth, L., Olsen, D., Asikin-Garmager, A., Winn, K.M., 2017. Initiating conversations and opening doors: how principals establish a positive building culture to sustain school improvement efforts. *Educ. Manag. Adm. Leader* 46, 1014–1034.
- Hoy, W., Tschannen-Moran, M., 1999. Five faces of trust: an empirical confirmation in urban elementary schools. *J. Sch. Leader.* 9, 184–208.
- Ilies, R., Judge, T., Wagner, D., 2006. Making sense of motivational leadership: the trail from transformational leaders to motivated followers. *J. Leader. Organ. Stud.* 13, 1–22.
- Ilies, R., Nahrgang, J.D., Morgeson, F.P., 2007. Leader–member exchange and citizenship behaviors: a meta-analysis. *J. Appl. Psychol.* 92, 269.
- Imants, J., Van Der Wal, M.M., 2020. A model of teacher agency in professional development and school reform. *J. Curric. Stud.* 52, 1–14.
- Jennings, P.A., Greenberg, M.T., 2009. The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev. Educ. Res.* 79, 491–525.
- Jiang, H., Luo, Y., 2018. Crafting employee trust: from authenticity, transparency to engagement. *J. Commun. Manag.* 22, 138–160.
- Kavanagh, B., Perkmann, M., Phillips, N., 2021. Collective identity and the limits of innovation: a review and research agenda. *Innovation* 23, 1–16.
- Khalifa, M., Gooden, M.A., Davis, J.E., 2016. Culturally responsive school leadership: a synthesis of the literature. *Rev. Educ. Res.* 86, 1272–1311.
- Khalifa, M., 2018. *Leadership in Troubled Times: Culturally Responsive Leadership for Minoritized Youth*. Harvard Education Press, Cambridge, MA.
- Klein, J., 2012. The open-door policy: transparency minimizes conflicts between school principals and staff. *Int. J. Educ. Manag.* 26, 550–564.
- Kolb, D., Bartunek, J., 1992. *Hidden Conflict in Organizations: Uncovering Behind-The-Scenes Disputes*. Sage, Thousand Oaks, CA.
- Kolb, J.A., 1995. Leader behaviors affecting team performance: similarities and differences between leader/member assessments. *J. Bus. Commun.* 32, 233–248.
- Kotsou, I., Mikolajczak, M., Heeren, A., Grégoire, J., Leys, C., 2019. Improving emotional intelligence: a systematic review of existing work and future challenges. *Emot. Rev.* 11, 151–165.
- Kouzes, J.M., Posner, B.Z., 2006. *The Leadership Challenge: How to Get Extraordinary Things Done in Organizations*. Jossey Bass, San Francisco.
- Koyama, J., Kania, B., 2014. When transparency obscures: the political spectacle of accountability. *J. Crit. Educ. Pol. Stud.* 12.
- Kurland, H., Peretz, H., Hertz-Lazarowitz, R., 2010. Leadership style and organizational learning: the mediate effect of school vision. *J. Educ. Adm.* 48, 7–30.
- Kurt, T., Duyar, I., Çalik, T., 2011. Are we legitimate yet? A closer look at the casual relationship mechanisms among principal leadership, teacher self-efficacy and collective efficacy. *J. Manag. Dev.* 31, 71–86.
- Lee, M., Louis, K.S., 2019. Mapping a strong school culture and linking it to sustainable school improvement. *Teach. Teach. Educ.* 81, 84–96.
- Lee, J., Shute, V.J., 2010. Personal and social-contextual factors in K–12 academic performance: an integrative perspective on student learning. *Educ. Psychol.* 45, 185–202.
- Leithwood, K., Jantzi, D., 1990. Transformational leadership: how principals can help reform school cultures. *Sch. Effect. Sch. Improv.* 1, 249–280.
- Leithwood, K., Jantzi, D., 2005. A review of transformational school leader research: 1996–2005. *Leader. Pol. Sch.* 4, 177–199.
- Leithwood, K., Louis, K.S., 2011. *Linking Leadership to Student Learning*. Jossey-Bass, San Francisco.
- Leithwood, K., Riehl, C., 2005. What do we already know about educational leadership. In: Firestone, W., Riehl, C. (Eds.), *A New Agenda for Research in Educational Leadership*. Teachers College Press, New York.
- Leithwood, K., Steinbach, R., 1991. Indicators of transformational leadership in the everyday problem solving of school administrators. *J. Person. Eval. Educ.* 4, 221–244.
- Leithwood, K., Steinbach, R., 1995. *Expert Problem Solving: Evidence from School and District Leaders*. SUNY Press.
- Leithwood, K., Leonard, L., Sharratt, L., 1998. Conditions fostering organizational learning in schools. *Educ. Adm. Q.* 34, 243–276.
- Leithwood, K., Mascall, B., Strauss, T., Sacks, R., Memon, N., Yashkina, A., 2007. Distributing leadership to make schools smarter: taking the ego out of the system. *Leader. Pol. Sch.* 6, 37–67.
- Leithwood, K., 1999. An organizational perspective on values for leaders of future schools. In: Begley, P. (Ed.), *Values and Educational Leadership*. SUNY Press, Albany.
- Liden, R.C., Maslyn, J.M., 1998. Multidimensionality of leader–member exchange: an empirical assessment through scale development. *J. Manag.* 24, 43–72.
- Liden, R.C., Wayne, S.J., Zhao, H., Henderson, D., 2008. Servant leadership: development of a multidimensional measure and multi-level assessment. *Leader. Q.* 19, 161–177.
- Lomotey, K., Lowery, K., 2015. Urban schools, black principals, and black students culturally responsive education and the ethno-humanist role identity. In: *Handbook of Urban Educational Leadership*, pp. 118–134.
- Lord, R.G., Brown, D.J., Freiberg, S.J., 1999. Understanding the dynamics of leadership: the role of follower self-concepts in the leader/follower relationship. *Organ. Behav. Hum. Decis. Process.* 78, 167–203.
- Louis, K.S., Lee, M., 2016. Teachers' capacity for organizational learning: the effects of school culture and context. *Sch. Effect. Sch. Improv.* 27, 534–556.
- Louis, K.S., Murphy, J.F., 2017. Trust, caring and organizational learning: the leader's role. *J. Educ. Adm.* 55, 103–126.
- Louis, K.S., Marks, H.M., Kruse, S.D., 1996. Teachers' professional community in restructuring schools. *Am. Educ. Res. J.* 33, 757–798.
- Louis, K.S., Mayrowetz, D., Smiley, M., Murphy, J., 2009. The role of sensemaking and trust in developing distributed leadership. In: Harris, A. (Ed.), *Distributed Leadership: Different Perspectives*. Springer, New York.
- Louis, K.S., Murphy, J., Smylie, M., 2016. Caring leadership in schools Findings from exploratory analyses. *Educ. Adm. Q.* 52, 310–348.
- Louis, K.S., 2007. Trust and improvement in schools. *J. Sch. Change* 8, 1–24.
- Louis, K.S., 2008. Changing the culture of schools: professional community, organizational learning and trust. *J. Sch. Leader.* 16, 477–489.
- Marks, H.M., Louis, K.S., 1997. Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educ. Eval. Pol. Anal.* 19, 245–275.
- Marks, H., Louis, K.S., 1999. Teacher empowerment and the capacity for organizational learning. *Educ. Adm. Q.* 35, 751–781.
- Marks, H.M., Printy, S., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39, 370–397.
- Maxcy, S.J., 1998. Preparing school principals for ethno-democratic leadership. *Int. J. Leader. Educ.* 1, 217–235.
- May, D.R., Chan, A.Y.L., Hodges, T.D., Avolio, B.J., 2003. Developing the moral component of authentic leadership. *Organ. Dynam.* 32, 247–260.
- Mayer, D.M., Kuenzi, M., Greenbaum, R., Bardes, M., Salvador, R.B., 2009. How low does ethical leadership flow? Test of a trickle-down model. *Organ. Behav. Hum. Decis. Process.* 108, 1–13.
- Mayeroff, M., 1971. *On Caring*. HarperPerennial, New York.
- McLaughlin, M.W., Talbert, J.E., 2006. *Building School-Based Teacher Learning Communities: Professional Strategies to Improve Student Achievement*. Teachers College Press.
- Mcneely, B.L., Meglino, B.M., 1994. The role of dispositional and situational antecedents in prosocial organizational behavior: an examination of the intended beneficiaries of prosocial behavior. *J. Appl. Psychol.* 79, 836.

- Meijer, A., 2013. Understanding the complex dynamics of transparency. *Publ. Adm. Rev.* 73, 429–439.
- Miskel, C., Ogawa, R., 1988. Work motivation, job satisfaction, and climate. In: Boyan, N. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York.
- Moolenaar, N.M., Sleegers, P.J., 2010. Social networks, trust, and innovation. How social relationships support trust and innovative climates in Dutch schools. *Soc. Netw. Theory Educ. Change* 97–114.
- Moye, M.J., Henkin, A.B., Egly, R.J., 2005. Teacher–principal relationships: exploring linkages between empowerment and interpersonal trust. *J. Educ. Adm.* 43, 260–277.
- Murphy, J., Louis, K.S., 2018. *Positive School Leadership: Building Capacity and Strengthening Relationships*. Teachers College Press, New York.
- Murphy, J.F., Torres, D., 2014. *Creating Productive Cultures in Schools for Students, Teachers, and Parents*. Corwin, Thousand Oaks, CA.
- Murphy, J., 2013. The architecture of school improvement. *J. Educ. Adm.* 51, 252–263.
- Murphy, J., 2016. *Understanding Schooling Through the Eyes of Students*. Corwin Press.
- Musah, M.B., Abdul Rahman, R.E., Tahir, L.M., AL-Hudawi, S.H.V., Daud, K., 2018. Headteacher–teacher relationship and its effect on teacher's trust in Malaysian high performing schools. *Int. J. Educ. Manag.* 32, 479–497.
- Nahrgang, J.D., Morgeson, F.P., Ilies, R., 2009. The development of leader–member exchanges: exploring how personality and performance influence leader and member relationships over time. *Organ. Behav. Hum. Decis. Process.* 108, 256–266.
- Newman, A., Schwarz, G., Cooper, B., Sendjaya, S., 2017. How servant leadership influences organizational citizenship behavior: the roles of LMX, empowerment, and proactive personality. *J. Bus. Ethics* 145, 49–62.
- Newmann, F., Associates, 1996. *Authentic Achievement: Restructuring Schools for Intellectual Quality*. Jossey Bass, San Francisco.
- Newmann, F.M., Rutter, R.A., Smith, M.S., 1989. Organizational factors that affect school sense of efficacy, community, and expectations. *Sociol. Educ.* 62, 221–238.
- Norberg, K., Johansson, O., 2007. Ethical dilemmas of Swedish school leaders: contrasts and common themes. *Educ. Manag. Adm. Leader* 35, 277–294.
- Norman, S.M., Avolio, B.J., Luthans, F., 2010. The impact of positivity and transparency on trust in leaders and their perceived effectiveness. *Leader. Q.* 21, 350–364.
- Nutov, L., Somech, A., 2017. Principals going above and beyond: understanding organizational citizenship behavior among school principals. *J. Sch. Leader.* 27, 184–212.
- Pianta, R.C., 1999. *Enhancing Relationships Between Children and Teachers*. American Psychological Association, Washington, DC.
- Podsakoff, P.M., Mackenzie, S.B., Ahearne, M., 1997. Moderating effects of goal acceptance on the relationship between group cohesiveness and productivity. *J. Appl. Psychol.* 82, 974–983.
- Pollock, K., Wang, F., Hauseman, D., 2015. Complexity and volume: an inquiry into factors that drive principals' work. *Societies* 5, 537–565.
- Price, H.E., Moolenaar, N.M., 2015. Principal–teacher relationships: foregrounding the international importance of principals' social relationships for school learning climates. *J. Educ. Adm.* 53 null.
- Price, H.E., 2011. Principal–teacher interactions. *Educ. Adm. Q.* 48, 39–85.
- Priestley, M., Biesta, G., Robinson, S., 2015. Teacher agency: what is it and why does it matter? In: Evers, J., Kneyber, R. (Eds.), *Flip the System: Changing Education from the Ground Up*.
- Printy, S.M., Marks, H.M., Bowers, A.J., 2009. Integrated leadership: how principals and teachers share transformational and instructional influence. *J. Sch. Leader.* 19, 504–532.
- Reynolds, K.J., Lee, E., Turner, I., Bromhead, D., Subasic, E., 2017. How does school climate impact academic achievement? An examination of social identity processes. *Sch. Psychol. Int.* 38, 78–97.
- Rhodes, J.E., Carnic, P.M., Milburn, M., Lowe, S.R., 2009. Improving middle school climate through teacher-centered change. *J. Community Psychol.* 37, 711–724.
- Riley, K., 2013a. *Leadership of Place: Stories from Schools in the US, UK and South Africa*. A&C Black.
- Riley, K.A., 2013b. *Leadership of Place: Stories from Schools in the US, UK and South Africa*. Bloomsbury Academic, London.
- Riley, K.A., 2013c. Walking the leadership tightrope: building community cohesiveness and social capital in schools in highly disadvantaged urban communities. *Br. Educ. Res. J.* 39, 266–286.
- Riley, K., 2017. *Place, Belonging and School Leadership: Researching to Make the Difference*. Bloomsbury Publishing.
- Riveros, A., Newton, P., Burgess, D., 2012. A situated account of teacher agency and learning: critical reflections on professional learning communities. *Can. J. Educ.* 35, 202–216.
- Rosenblatt, Z., Peled, D., 2002. School ethical climate and parental involvement. *J. Educ. Adm.* 40, 349–367.
- Ross, J.A., Gray, P., 2006. Transformational leadership and teacher commitment to organizational values: the mediating effects of collective teacher efficacy. *Sch. Effect. Sch. Improv.* 17, 179–199.
- Rule, A.C., 2006. The components of authentic learning. *J. Authentic Learn.* 3, 1–10.
- Sargent, T.C., 2014. Professional learning communities and the diffusion of pedagogical innovation in the Chinese education system. *Comp. Educ. Rev.* 59, 102–132.
- Schechter, C., Qadach, M., 2012. Toward an organizational model of change in elementary schools: the contribution of organizational learning mechanisms. *Educ. Adm. Q.* 48, 116–153.
- Schechter, C., Shaked, H., 2017. Leaving fingerprints: principals' considerations while implementing education reforms. *J. Educ. Adm.* 55, 242–260.
- Schein, E.H., 1999. Kurt Lewin's change theory in the field and in the classroom: notes toward a model of managed learning. *Reflections SoL J.* 1, 59–74.
- Schratz, M., 2003. From administering to leading a school. *Camb. J. Educ.* 33, 395–416.
- Schuttinger, K., 2021. Engaging teachers in innovation: positive leadership from the bottom-up. In: Walker, K., Kutsyruba, B., Cherkowski, S. (Eds.), *Positive Leadership for Flourishing Schools*. Information Age Publishing, Charlotte, NC.
- Seashore, S.E., 1954. *Group Cohesiveness in the Industrial Work Group*. University of Michigan, Institute for Social Research, Ann Arbor, MI.
- Seligman, M.E., Csikszentmihalyi, M., 2000. Positive psychology: an introduction. *Am. Psychol.* 55, 5–14.
- Seligman, M.E., Steen, T.A., Park, N., Peterson, C., 2005. Positive psychology progress: empirical validation of interventions. *Am. Psychol.* 60, 410.
- Sergiovanni, T., 1992. *Moral Leadership: Getting to the Heart of School Improvement*. Jossey Bass, San Francisco.
- Sergiovanni, T., 1994. *Building Community in Schools*. Jossey-Bass, San Francisco, CA.
- Shapira-Lishchinsky, O., Tsemach, S., 2014. Psychological empowerment as a mediator between teachers' perceptions of authentic leadership and their withdrawal and citizenship behaviors. *Educ. Adm. Q.* 50, 675–712.
- Shapiro, J.P., Stefkovich, J.A., 2016. *Ethical Leadership and Decision Making in Education: Applying Theoretical Perspectives to Complex Dilemmas*. Routledge, New York.
- Sluss, D.M., Ashforth, B., 2008. How relational and organizational identification converge: processes and conditions. *Organ. Sci.* 19, 807–823.
- Sluss, D.M., Klimchak, M., Holmes, J.J., 2008. Perceived organizational support as a mediator between relational exchange and organizational identification. *J. Vocat. Behav.* 73, 457–464.
- Smylie, M., Mayrowetz, D., Murphy, J., Louis, K.S., 2007. Trust and the development of distributed leadership. *J. Sch. Leader.* 17, 469–503.
- Smylie, M., Murphy, J., Louis, K.S., 2016. Caring school leadership: a multidisciplinary, cross-occupational model. *Am. J. Educ.* 123, 1–35.
- Smylie, M., Murphy, J., Louis, K.S., 2019a. *The Practice of Caring School Leadership*. Corwin, Thousand Oaks.
- Smylie, M., Murphy, J., Louis, K.S., 2019b. *The Practice of Caring School Leadership*. Corwin, Thousand Oaks.
- Smylie, M., Murphy, J., Louis, K.S., 2020. *Stories of Caring School Leadership*. Corwin, Thousand Oaks, CA.
- Somech, A., Ron, I., 2007. Promoting organizational citizenship behavior in schools: the impact of individual and organizational characteristics. *Educ. Adm. Q.* 43, 38–66.
- Spreitzer, G.M., 2008. Taking stock: a review of more than twenty years of research on empowerment at work. In: Cooper, C., Barling, J. (Eds.), *Handbook of Organizational Behavior*. Sage, Thousand Oaks, CA.
- Srivastava, A.P., Dhar, R.L., 2019. Authentic leadership and extra role behavior: a school based integrated model. *Curr. Psychol.* 38, 684–697.
- Stefkovich, J.A., O'Brien, G.M., 2004. Best interests of the student: an ethical model. *J. Educ. Adm.* 42, 197–214.

- Sternberg, R.J., Vroom, V., 2002. The person versus the situation in leadership. *Leader. Q.* 13, 301–323.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., Thomas, S., 2006. Professional learning communities: a review of the literature. *J. Educ. Change* 7, 221–258.
- Sun, M., Frank, K.A., Penue, W.R., Kim, C.M., 2013. How external institutions penetrate schools through formal and informal leaders. *Educ. Adm. Q.* 49, 610–644.
- Supovitz, J., 2009. Knowledge-based organizational learning for instructional improvement. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*. Springer Netherlands.
- Theoharis, G., O’toole, J., 2011. Leading inclusive ELL: social justice leadership for English language learners. *Educ. Adm. Q.* 47, 646–688.
- Tichnor-Wagner, A., Allen, D., 2016. Accountable for care: cultivating caring school communities in urban high schools. *Leader. Pol. Sch.* 15, 406–447.
- Tims, M., Bakker, A.B., Derks, D., Van Rhenen, W., 2013. Job crafting at the team and individual level: implications for work engagement and performance. *Group Organ. Manag.* 38, 427–454.
- Tourish, D., 2014. Leadership, more or less? A processual, communication perspective on the role of agency in leadership theory. *Leadership* 10, 79–98.
- Tschannen-Moran, M., 2004. *Trust Matters: Leadership for Successful Schools*. Jossey Bass, San Francisco.
- Tschannen-Moran, M., 2009. Fostering teacher professionalism in schools: the role of leadership orientation and trust. *Educ. Adm. Q.* 45, 217–247.
- Tuytens, M., Devos, G., 2014. How to activate teachers through teacher evaluation? *Sch. Effect. Sch. Improv.* 25, 509–530.
- Tuytens, M., Devos, G., 2019. Balancing differentiation and fairness in teacher evaluation: the story of Flemish secondary schools. In: *Differentiated Teacher Evaluation and Professional Learning*. Springer.
- Vekeman, E., Devos, G., Tuytens, M., 2015. The influence of teachers’ expectations on principals’ implementation of a new teacher evaluation policy in Flemish secondary education. *Educ. Assess. Eval. Account.* 27, 129–151.
- Vogelgesang, G.R., Leroy, H., Avolio, B.J., 2013. The mediating effects of leader integrity with transparency in communication and work engagement/performance. *Leader. Q.* 24, 405–413.
- Voulalas, Z.D., Sharpe, F.G., 2005. Creating schools as learning communities: obstacles and processes. *J. Educ. Adm.* 43, 187–208.
- Wahlstrom, K.L., Louis, K.S., 2008. How teachers experience principal leadership: the roles of professional community, trust, efficacy, and shared responsibility. *Educ. Adm. Q.* 44, 458–495.
- Walberg, H.J., Anderson, G.J., 1968. Classroom climate and individual learning. *J. Educ. Psychol.* 59, 414.
- Walker, K., Kutsyruba, B., Cherkowski, S. (Eds.), 2021. *Positive Leadership for Flourishing Schools*. Information Age Publishing, Charlotte, NC.
- Walumbwa, F.O., Avolio, B.J., Zhu, W., 2008. How transformational leadership weaves its influence on individual job performance: the role of identification and efficacy beliefs. *Person. Psychol.* 61, 793–825.
- Walumbwa, F.O., Cropanzano, R., Hartnell, C.A., 2009. Organizational justice, voluntary learning behavior, and job performance: a test of the mediating effects of identification and leader-member exchange. *J. Organ. Behav.* 30, 1103–1126.
- Walumbwa, F.O., Cropanzano, R., Goldman, B.M., 2011. How leader-member exchange influences effective work behaviors: social exchange and internal-external efficacy perspectives. *Person. Psychol.* 64, 739–770.
- Wang, H., Law, K.S., Hackett, R.D., Wang, D., Chen, Z.X., 2005. Leader-member exchange as a mediator of the relationship between transformational leadership and followers’ performance and organizational citizenship behavior. *Acad Manage J* 48, 420–432.
- Ware, H., Kitsantas, A., 2007. Teacher and collective efficacy beliefs as predictors of professional commitment. *J. Educ. Res.* 100, 303–310.
- Weiss, C.H., Cambone, J., 1994. Principals, shared decision making, and school reform. *Educ. Eval. Pol. Anal.* 16, 287–301.
- Wilson, C.M., 2016. Enacting critical care and transformative leadership in schools highly impacted by poverty: an African-American principal’s counter narrative. *Int. J. Leader. Educ.* 19, 557–577.
- Wu, J., Hoy, W.K., Tartar, J., 2012. Enabling school structure, collective responsibility, and a culture of academic optimism: toward a robust model of school performance in Taiwan. *J. Educ. Adm.* 51, 176–193.
- Youngs, P., King, M.B., 2002. Principal leadership for professional development to build school capacity. *Educ. Adm. Q.* 38, 643–670.

Networked systems leadership: the potential of social network theory and analysis

Yi-Hwa Liou^a and Alan J. Daly^b, ^aNational Taipei University of Education, Department of Educational Management, Taipei City, Taiwan; and ^bUniversity of California, San Diego, Department of Education Studies, La Jolla, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	524
Moving toward networked systems leadership	525
Social processes in the shaping of leadership	526
Social network theory and analysis	526
Networks	526
Social networks: theory and analysis	526
Social networks: relationship to leadership and change	527
Key insights from social network research related to leadership and change	527
Dyad level: homophily and relational ties	527
Typology of ties	528
Strength of ties	528
Node level: centrality	529
Degree centrality	529
Peripheral or isolate	530
Betweenness	530
Group and whole network level	530
Framing networked systems leadership for change	531
Promising developments of networked systems leadership and change	532
Longitudinal dynamic networks	532
Two- or multi-mode networks and multilevel networks	532
Real time, multi-data sources in profiling networked leadership phenomena	532
Conclusion	533
References	533

Introduction

Leadership is one of the most studied topics in the educational literature. Scholars in educational administration have long deconstructed and reconstructed leadership from different perspectives. Some have studied the roles, functions, traits, or styles of educational leaders (Day et al., 2016; Gardner, 2000), while others endeavored to identify the qualities of successful leaders (Deal & Peterson, 2000; Leithwood and Riehl, 2003; Murphy et al., 2017) and still others examined transactional and transformational types of leadership (Bass, 1997; Burns, 1978; Leithwood and Jantzi, 2000) just to name a few. As we look across all of the work one of the foundational threads that cuts across the varied leadership conceptualizations is that leadership implicates the role of self and other in a complex set of social interactions. However, while those sets of relationships are implicated they are often not front and center in the study of leadership (Daly, 2010; Donaldson, 2006; Leithwood, 2019). A more relational perspective on leadership foregrounds the role of social interactions and is interested in the social space between individuals and the outcomes of interactions (Liou et al., 2015). This perspective on leadership is not a new nor revolutionary one, but given new advances in both theory and methods, it is worth carefully revisiting and that is the goal of this paper.

Dating back to the 70s, scholars begun to acknowledge that the execution of leadership is not confined to those in formal administrative positions, but rather, is something that flows throughout an organization with individuals within organizations being able to exert leadership (Barnard, 1968; Thompson, 1967). As we fast forward a number of decades, the more relational perspective took root with the concepts such as distributed leadership (Spillane, 2006), networked leadership (Liou and Daly, 2019b; Díaz-Gibson et al., 2017), professional learning (Penuel et al., 2012), and networked improvement communities (Russell et al., 2017). These scholars and others view leadership as a phenomenon that can be located within the relations between and among others within organizations or systems. Leadership in this space is more of a variable that can be measured through relations at different levels of organizational systems, e.g., individual, group, organization, or systems, and related to a number of outcomes.

An increasingly growing body of research in education has begun to explore the social aspect of systems using social network theory and analysis specifically to the work of leadership in educational systems (e.g., Liou and Daly, 2014, 2019a, 2020b; Coburn et al., 2013; Daly and Finnigan, 2011, 2016; Moolenaar and Slegers, 2015; Penuel et al., 2012; Spillane and Kim, 2012), and more recently leaders' online interactional behaviors (Daly et al., 2019a,b). These lines of research view leadership as a part of a larger

social system and how it is manifest in “networks” as a key feature to systems and ultimately systemic change. Viewed through a network perspective, leadership is focused on the pattern of relationships that is likely to influence or be influenced by the beliefs and behaviors of others with whom leaders have a social tie.

Building on the social paradigm of leadership, we argue that it is important to understand leadership as being explicitly grounded in the role of interaction and interdependence and as such can be conceptualized and assessed by new advances in social network theory and methods. In making this argument, we offer a network perspective on educational leadership through introducing social network theory and analysis and its relationship to leadership and change. We discuss key insights from network literature and organize a framework that describes a social network approach to educational leadership. Finally, we conclude with several promising directions for future research in this space.

Moving toward networked systems leadership

The social paradigm of leadership shifts our attention from the individual to the sets of relationships that surround leaders. This view recognizes the social role of leadership and is interested in those in formal positions with leadership titles (formal leaders) as well as those that play important informal leadership roles as determined by their set of social ties (informal leaders) (Balkundi and Kilduff, 2006; Gardner, 2000; Leithwood, 2017; Spillane, 2017). Both formal and informal leaders are important to supporting cohesion. Scholars examining educational systems and reform have suggested that better linkages among individuals and segments of a system may result in improved communication, shared understanding, and ultimately greater coherence throughout the system (Agullard and Goughnour, 2006; Honig and Ikemoto, 2008). As these relationships are interdependent in nature and consequential to systemic and coherent change, we adopt a systems perspective to the social aspect of leadership (Liou et al., 2015; Díaz-Gibson et al., 2017; Leithwood, 2017; Leithwood et al., 2020; Pitts and Spillane, 2009; Sun et al., 2013). The balance of the systems work suggests that leadership exists at various levels in both formal hierarchical relationships and informal influence pattern within collaborative systems of organizations (Kramer et al., 2019). Scholars have noted that shifting from individual to the relational, and attending to formal and informal positions, may hold potential in better understanding the social infrastructure of systems change (Daly et al., 2017; Goldspink, 2007). Following these developments in leadership research, we view leaders in educational systems as those individuals, participants or actors that emerge in networks of social connectivity exerting (or attempting to, or having a potential to exert) influences on system-wide practice toward a purpose.

This more systems approach to leadership is driven by systems thinking (Senge, 2006)—a system of thinking about systems (Arnold and Wade, 2015)—that suggests that any change in one component part of the system may result in intended or unintended behaviors/consequences for other parts of the system (Senge, 2006). Systems thinking also focuses on the ways in which a system’s constituent parts interrelate within and respond to the context of its larger systems and as such leaders that emerge from within the systems must be taking a more systems perspective on how discreet parts take shapes a unified whole (Shaked and Schechter, 2019). Recent scholarship advocates for the use of systems thinking in educational leadership, as described by Shaked and Schechter (2019), “Systems thinking was claimed to help educational leaders to see public relations as a continual, systematic process that is essential for engaging the school community’s support to improve students’ learning” (p. 577). This view of leadership underscores the role of “relationship” in shaping systems practice toward a common purpose, and thus resonates with Donaldson’s (2006) view on educational leadership: “leadership is a relationship that mobilizes people to fulfill the purposes of education” (p. 47). Similarly, this relational emphasis of leadership corresponds to the “distributed” perspective that posits that “leadership is stretched over the work of multiple leaders” (Spillane and Diamond, 2007, p. 8) through the interactions of leaders and followers (Pitts and Spillane, 2009).

The work around “participative” systems perspective as a foundation provides some analytic purchase to further consider the relational role of leadership. Participative systems mainly refer to the cooperation or collaboration among systems of organizations as an appropriate solution to address issues related to organizational changes (Connelly, 2007; Kramer et al., 2019). Leadership in this literature focuses on the capacity for creating conditions that foster collaboration or cooperation among and between cross-sector organizations (Connelly, 2007; McCaffrey et al., 1995). Absence of relational linkages between parts of a system may limit the possibility of leading a success change (Daly and Finnigan, 2016).

Therefore, the systems approach to leadership and improvement requires a shift in the way that change strategies are formalized and enacted within a system. This shift corresponds with the social paradigm of leadership and suggests a move from a singular focus on parts of a system to engaging the entire segment of the system formed by networks of connections. Scholars taking the systems approach to organizational change indicate a wide range of strategies for building stronger intra- or inter-organizational ties (Honig, 2008; Popp et al., 2013; Tsai, 2000). These strategies include building structures to foster collaborations between levels in organizations (Daly and Finnigan, 2016), enhancing communication and knowledge sharing by bridging ties within and across organizational levels (Liou and Daly, 2020a; Honig, 2006, 2008), distributing leadership (Spillane, 2006), and creating “network initiation team” to foster commitment and motivate others, rather than using formal power to impose roles or mandate actions (Russell et al., 2017). All of these strategies suggest the importance role of social interdependencies in supporting change efforts.

These perspectives on leadership set leadership parameters at various levels of systems, e.g., individuals, relational dyads, subgroups of a system, and the whole systems. They situate leadership in a systems context and suggest that the nature of leadership is relational, distributed, participative, and networked, all of which considers networks as a feature to systems performance. As such,

leadership is embedded within networks of systems and shaped by social processes. These perspectives reinforce the important tenets of existing leadership literature and further introduce a network perspective on leadership that deserves greater attention. Taken together, we argue that leadership is a systems quality that is embedded within networks of social systems and shaped by social processes.

Social processes in the shaping of leadership

Social processes involve the complex interplay between interdependencies of individuals and their relational ecosystem toward common goals as well as their beliefs which governs their decision making and behaviors (Liou and Daly, 2020b; Siciliano et al., 2017). Discussion of social processes derives from social learning theory that posits that human behavior is constructed and shaped through direct experience and interactions with others and their surroundings (Bandura, 1977, 2002). According to Bandura (1971), "Behavior partly creates the environment and the resultant environment, in turn, influences the behavior. In this two-way causal process the environment is just as influenceable as the behavior it controls" (p. 40). In this view, social competencies can be developed, shaped through interactions that are embedded within social context and are key quality to enable learning (Bandura, 2006; Baron and Markman, 2003).

This view of human social processes helps explain the social nature of leadership. Leadership, based on the above-mentioned perspectives, is conceptualized as a process of influence toward an outcome. Specifically, such leadership largely involves social influence that occurs in sets of relationships where individuals are both influencing and influenced by one another, which affects the patterns of interactions embedded within the systems (Donaldson, 2006; Ogawa and Bossert, 2000). Central to this view of leadership in its social processes is interactions and conditions that foster such interactions (Liou and Daly, 2020b; Lin, 2009). In education, recent developments in leadership research emphasize the value of social processes and relational linkages involved in leading and change (Liou et al., 2015; Daly, 2010; Pitts and Spillane, 2009). This line of research treats leadership as based on the deployment of resources that reside in social relationships, and examine social relationships through networks. It is argued that networks of relationships may provide leaders with the necessary resources and social infrastructure to exert social influence in achieving individual and organizational goals such as implementing reforms or improving student outcomes (Liou et al., 2015; Spillane and Kim, 2012). Leadership from a network perspective emphasizes individual agency and the interdependence of social behavior that are reflected by a network of relational linkages that may moderate, influence, or inform the flow, speed, type, and level of engagement of social activities and mobilization of resources such as knowledge, information, skills and practice (Daly, 2012).

While a growing number of studies explore the social aspect of educational leaders using social network analysis (Coldren & Spillane, 2007; Daly & Finnigan, 2016; Liou, 2016; Liou & Daly, 2014, 2018, 2019a, 2020a; Pitts and Spillane, 2009; Spillane and Kim, 2012), this space still has room to grow around the study of educational leadership. In the following section, we draw on social network theory and analysis and research literature to understand the social aspect of leadership and its potential influence on change.

Social network theory and analysis

Networks

The term "networks" has been widely used in leadership and management research as a promising organizational structure for addressing problems that require the access, use, and mobilization of resources across organizational systems to engage actors in collective action toward a common purpose (Liou and Daly, 2020a; Hoberecht et al., 2011; Russell et al., 2017). A network is formed by a constellation of three or more individuals (nodes, actors, or organizations) that are connected based on types of relationships for a particular purpose (Liou and Daly, 2020a). In other words, networks consist of structures of relationships between actors that exist for common purpose(s) (Popp et al., 2013). From a systems perspective, networks can be used as a way of thinking about and configuring organizational systems (Borgatti et al., 2018).

Social networks: theory and analysis

Social network analysis is theorized based on two analytically distinct domains: network theory and theory of networks (Borgatti and Halgin, 2011). Network theory, according to Borgatti and Halgin (2011), refers to "the mechanisms and processes that interact with network structures to yield certain outcomes for individuals and groups" (p. 1). In this regard, the theory is concerned with the resulting consequences of network structures, such as being highly connected or centrally positioned (Borgatti and Ofem, 2010). Examining relational structures and their consequences becomes central to most of the network studies in this paradigm (Borgatti and Halgin, 2011). For instance, much of the network research has focused on characterizing network structures (e.g., cohesion) and actor positions (e.g., centrality), and associated these network features with actor- or group-level attributes or outcomes (e.g., Daly et al., 2020; Spillane and Kim, 2012). A theory of networks is used to explain the processes that determine a given network structure (Borgatti and Ofem, 2010). Studies of this theoretical paradigm use analysis that models formation of ties and structural effects to best characterize the network structure (e.g., Liou and Daly, 2019a).

Despite the different orientations in theorizing social networks, what remain similar are the core assumptions underpinning the study of networks. First, network research focuses on the study of sets of actors and relations/ties that connect them (Kilduff and Brass, 2010). Second, actors are embedded within a network and as such are assumed to be interdependent and interconnected rather than independent (Degenne and Forsé, 1999; Wasserman and Faust, 1994). Third, relationships between actors are regarded as “ties” (linkages or connections) that provide for the exchange of resources (Burt, 1982, 1997; Kilduff and Brass, 2010; Lin, 2009). Fourth, network structure formed by relational ties has influence on the flow of resources to and from an individual as well as across a system (Borgatti and Foster, 2003; Lin, 2009). Fifth, network structures can act as both opportunities and constraints for individual and collective action (Kilduff and Brass, 2010). These core ideas are important in grounding our discussion about leadership and its research.

Social networks: relationship to leadership and change

As discussed earlier, one of the key purposes of educational leadership is to shape organizational systems for individual and/or collective outcomes and the key medium of such leadership practice is social interaction. The practice of shaping systems involves leading and sustaining changes at scale and ensuring learning occurs within the systems (Kezar, 2018). Changing at scale (or scale-up) involves the adoption/use of practices, strategies, or reform initiatives by a broader reach of individuals over time (Coburn, 2003). In achieving scale-up, networks are found to be one of the most effective strategies, as networks help to identify opportunities and constraints for the spread of practices necessary for change (Daly and Finnigan, 2016; Kezar, 2011; Russell et al., 2017).

Studies in education and leadership have identified several ways that link networks with change (Daly, 2010; Daly and Finnigan, 2016; Russell et al., 2017): opportunities to learn, development of social capital, innovation, and collaborative capacity (Liou and Daly, 2020a). First, network research suggests that social networks provide existing and potential opportunities that enable relationships for learning—through accessibility to resources (Lin, 2009; Liou and Daly, 2018; Liou and Canrinus, 2020), diffusion, exchange, and transfer of knowledge and information (Hubers et al., 2018; Spillane and Kim, 2012), peer influence (Liou and Daly, 2019b; Supovitz et al. 2010), and shaping of norms, beliefs and attitudes (Liou and Daly, 2020b). Second, networks facilitate the shaping of social capital by providing accessibility to resources that are embedded within relationships which can be mobilized to increase the likelihood of success in purposive action of change (Daly and Finnigan, 2016; Lin, 2009). Third, as change often involves execution of innovation that entails risk-taking, network members may create a climate that is trusting, open and risk tolerant so that they can learn from others’ innovative solutions and are willing to sustain their participation in collective action (Hoberecht et al., 2011). Finally, in achieving a long-term and systemic change, networks need to develop their collaborative capacity for change (Provan and Lemaire, 2012; Weber and Khademian, 2008), which resonates with the concept of networked improvement communities (Russell et al., 2017).

Taken together, network approach to leadership and change involves several social undertakings. To begin, network members need to develop a sense of commitment to the common purpose of the network. These network members mutually engage, share useful practices, use leadership to influence connectivity, strategize collaboration, and create a mechanism that monitors performance and accountability by assessing networks’ capacity for reaching a range of network members whose participation in the network is essential for achieving the common purpose (Russell et al., 2017; Weber and Khademian, 2008). Network members, be they individuals or organizations, can exert their networked leadership influence across the systems in creating a web of resources and support, which can benefit the networks to meet specific goals and scale practices.

Key insights from social network research related to leadership and change

As relationships between actors are central and pivotal to leadership and change efforts, network studies take into account the web of relationships in which actors are embedded that both support and constrain opportunities for individual influence and overall outcomes of the network systems (Borgatti and Ofem, 2010). Those relational ties are connected between sets of actors to form paths (trials or walks required from one actor to another) that can be regarded as pipes through which resources (information, ideas, knowledge) can flow. These paths form an overall network structure in which each actor possesses a particular position and sets of actors shape up a particular subgroup. Network positions determine in part individual actors’ (dis)advantages points and network patterns of the paths influence resource flow and thus may impact network performance. As network patterning and its consequences can be understood at different levels of network systems, we follow Borgatti and colleagues’ (2010, 2018) typology of network research to organize key insights from network research in education by three levels of analysis: the dyad, the node, and the group and whole network. We include the most often used network concepts related to leadership and change in the educational literature.

Dyad level: homophily and relational ties

The dyad level mainly is concerned with pairs of actors. Studies in this space focus on (1) why two actors are connected through a particular relation or (2) the extent to which such dyadic connections influence similarity of beliefs, attitudes, or perceptions between actors. The first type is often referred to as (de)selection model and associated with the network phenomenon—homophily, which is the tendency for people to form ties with those similar to them (Lazarsfeld and Merton, 1954). Network studies have

examined homophily effect of individuals by demographic (age, gender, race/ethnicity), social status, beliefs, or geographical locations (McPherson et al., 2001). In education, studies have found that teachers tend to engage in extensive professional dialogue with colleagues who teach at the same grade and are a close friend (Frank and Zhao, 2005; Penuel et al., 2010). Teachers who attend professional development events in a similar subject area are more likely to form help-seeking ties with others in the same area (Spillane et al., 2012). Research in educational leadership that examines tie formation among a group of leadership team suggests that leaders tended to form ties with others who are physically close at work (within district or within school site) (Daly et al., 2016; Liou, 2016) and shared similar efficacy beliefs about their leadership skills (Daly et al., 2016). Other research that examines both the formation and dissolution of ties among a middle school's teachers and administrators suggests that educators who serve as a teacher leader, teach the same subject, and share different years of experiences are likely to establish and maintain new ties (Liou and Bjorklund Jr., 2021), and changing grade levels or losing leadership position predicts the dissolution of ties among elementary school staff (Spillane and Shirrell, 2017).

The homophily effect as an explanatory variable for the formation of ties contributes to the understanding of the work of leadership in terms of how leaders may consider shaping similarities in beliefs, work level, and other attributes to facilitate connections between peers. Tie formation as a result of actor homophily can speed up the flow of information necessary for change but might also result in sharing redundant information thus impeding change. In contrast, establishing ties that are heterophilous (Granovetter, 1973)—diverse in types of ties and span of sources for information and knowledge—can facilitate change by accessing new and diverse information/ideas but might slow down change processes because it takes more time for resources to travel across multiple, less tightly connected actors. Educational reform and change efforts are better advanced by developing social infrastructure that facilitates the formation of heterophilous ties among leaders and stakeholders for obtaining diverse views, resources and input (Daly & Finnigan, 2010).

The second type of network studies—*influence*—focuses on the role of dyads in explaining change in actor behaviors, beliefs, or attitudes. The idea of influence captures the processes of social influence by which an actor's beliefs/views or behaviors assimilate toward those held by members of their network (Siciliano and Thompson, 2018). Studies in educational leadership that seek to explore the influence effect are limited with only a few studies examining teacher practice. For instance, scholars have found that a teacher's own beliefs about reform initiatives tend to assimilate the level of beliefs held by those peers that this teacher is directly connected (Siciliano et al., 2017). Others have suggested that teachers' instructional practices are influenced by the behaviors of those peers with whom they are frequently interacted (Frank et al., 2013).

Typology of ties

As most of the research in education at the dyad level has to do with formation of relational ties, it is important to briefly discuss the type of ties analyzed in these studies and its function. Network researchers distinguish between two types of relational ties by their functions: instrumental ties and expressive ties (e.g., Ibarra, 1995; Lincoln and Miller, 1979). Instrumental ties tend to be work-related involving the gathering of information and resources for a particular professional purpose (Labianca and Brass, 2006). Expressive ties typically involve expression of interpersonal affect and are more normative and affect based (Umpfress et al., 2003). Expressive ties are consequential for the formation of stronger ties that can enhance team performance (Henttonen et al., 2013), whereas instrumental ties are typically information-driven and helpful for disseminating information and ideas (Labianca and Brass, 2006).

Both types of ties can be leveraged for change. For instance, educational scholars found that expressive ties such as close relationships tend to have a greater influence on teacher attitudes toward reform than instrumental ties (Cole and Weinbaum, 2010), and that teachers being identified as close friends to their peers is positively associated with the degree of accessibility to expertise inside and outside their school they perceived (Penuel et al., 2009). Shifts in educators' beliefs about mathematics instruction are associated with changes in their instrumental ties around seeking advice and information about the instruction (Spillane et al., 2018). In addition, as both types of ties often co-exist and interact in work settings, another line of research in educational leadership explores this understudied area based on positive organizational psychology literature that suggests that social support and emotional health of leaders are critical to their leadership performance (Balkundi and Kilduff, 2006). Research in this space suggests that leaders who are considered as energizers—individuals who make others feel positively inspired and motivated—are more satisfied with their job (Daly et al., 2016) and more confident about their leadership and management abilities (Liou and Daly, 2019a).

Strength of ties

Another way to understand complex relational ties is to assess the degree of tie strength. Strong ties are typically characterized by frequent interaction, prior/existing connections, intimacy (close knit), multi-layer relations (multiplexity), or mutual confiding (reciprocity) (Granovetter, 1973) and sometimes correlate with the concept of bonding social capital (Putnam, 2000). Weak links, where bridging social capital can be found, are characterized by infrequent, distance, loose relationships that are oftentimes non-reciprocated and less close (Granovetter, 1973). Strong ties are found to facilitate complex change as they are more likely to promote two-way exchanges of tacit, non-routine, and detailed knowledge and information, whereas weak ties are best suited to communicating direct, non-routine information that requires little time and efforts and are more helpful to quickly spread ideas and information, thus fostering diffusion of innovation (Tsai, 2000). While strong ties tend to be more stable and coherent due to their mutual dependency used to maintain a strong linkage (Monge and Contractor, 2003), these linkages may lead to network closure thus constraining the creation and exploration of new information necessary for innovation (Burt, 2007). Weak links typically do

not offer substantive social support or intimate connections, which also implies that individuals are not bounded by social norms, thus more open to new ideas.

In education, Coburn and Russell (2008) study a mathematic curriculum reform of elementary school teachers and measure their tie strength by combining frequent interaction and close relationship. They found that teachers' strong ties help to reshape their pedagogical approaches (Coburn and Russell, 2008). Another study by Daly and Finnigan (2011) examines tie formation in understanding change efforts among central office and site leaders in an underperforming school district. They investigated over-time changes in four types of ties (existing, reciprocal, asymmetric, and absent) and found increases in weak linkages over time but slightly drops or no change in strong ties. A more recent study by Liou and Daly (2019a) examines a strong tie of educational leaders that has occurred over a span of three years and involves work-related influence and energy relations. They found that the sustaining of these strong ties is associated with shared beliefs about leadership and geographic work location. Understanding tie strength may help leaders to assess the type and degree of social capital, which can be used as levers for innovation and change (Hoppe and Reinelt, 2010).

Node level: centrality

The node level network analysis primarily considers individual actors' network position and neighborhood. Most studies exploring this area treat an actor's network position as outcomes or as predictors. Network position, from a social network perspective, is different from the formal, contractual position discussed in most of the leadership and management literature. It refers to the social, informal position of individuals emerged in a social system which has greater influence than those with positional authority on how resources/information flows to that individual and ultimately how those resources get used in a system (Balkundi et al., 2011; Daly, 2010; Penuel et al., 2009, 2012; Spillane et al., 2012). The network position of an actor is determined by the constellation of ties that actor has, including both incoming and outgoing ties (Scott, 2017). Although there are a number of structural measures to determine an actor's network position, we focus our discussion on key centrality measures that are mostly found in educational studies: degree centrality and betweenness. We then discuss the direction of causality these measures are used in some of the educational studies.

Degree centrality

Degree centrality often refers to the number/proportion of ties a node has in a given network. In an undirected network (ties are symmetric, e.g., have a meeting with), degree centrality is the number/proportion of ties a node has regardless of directionality of ties. In a directed network (e.g., advice seeking) where ties are directional, the number/proportion of ties a node sends or receives can be calculated separately, which produces node-level outdegree or indegree centralities. The outdegree of a node indicates the node's activity, whereas a node's indegree refers to their popularity (Wasserman and Faust, 1994). In network research, degree centrality is often equated with informal leadership: individuals with higher degree centralities (more centrally located like "hubs") tend to have greater social power and influence over the network (Brass, 1992). In some studies, it is used to measure the volume of social capital a node has (Coburn and Russell, 2008; Daly et al., 2014; Liou and Daly, 2018; Penuel et al., 2009).

Another similar concept to degree centrality is the measure of closeness centrality that takes into account not only number/proportion of ties but also distance from a node to all others (Brandes et al., 2016). A node with a high closeness score means there is a short distance between them and most others, and as such they can quickly reach others to access and share relational resources in an efficient manner (Borgatti et al., 1998). Given they are highly and well connected as being points of multiple intersecting relations, these central actors tend to occupy a hub position and therefore have a greater opportunity (more social capital), and a better vantage point, to access, be exposed to, and leverage, the flow of relational resources (Kilduff and Krackhardt, 2008; Scott, 2017). However, they are also more likely to be constrained by the norms and beliefs held by their tightly connected networks, which may lead to social conformity (Krackhardt, 1999) or stagnation (Krackhardt and Kilduff, 1990).

The role of central actors as outcomes or predictors can be found in educational literature. For instance, Moolenaar et al.'s (2010) study of school principals' network position as measured by indegree centrality indicates that the indegree mediates the relationship between reported levels of transformational leadership and innovative school climate. Another study by Daly and colleagues (2014) examines network positions of both central office and site leaders as functions of seeking or providing advice around reform efforts in three districts using indegree and outdegree centralities. They found that leaders who occupy more central network positions (i.e., higher out- or in-degree) reported as being more confident in their ability to exercise leadership. Their findings also indicate that leaders' formal position (e.g., central office or school site) does not significantly contribute to explaining their informal social position, which suggests the importance of informal social positions in a districtwide reform network. As for closeness centrality, research on networks of school principal and teachers indicates a positive relationship between closeness and relational trust for teachers and principal (Liou and Daly, 2014). This work suggests closeness serves as an important proxy measure for interpersonal trust. A more recent study by Liou and Daly (2020b) uses indegree centrality to investigate the role of educational leaders' network position in shaping leader beliefs about reform. Their work indicates that indegree (i.e., being sought for advice and providing work recognition) mediates the relationship between leaders' engagement in reform efforts and their beliefs about the reform. These studies suggest the importance of "informal" social position in a network, particularly being central, as consequential to the social side of leadership.

Peripheral or isolate

Contrary to actors with high degree centrality, actors with low or zero degree centrality—often referred to as peripheral or isolates—also receive much attention in the network literature in education. Peripheral or isolated individuals represent a leadership challenge, such as disconnections between subgroups that are structured for a certain work division, or too many structural holes (Burt, 2005) in a team or organizational system that prohibit interaction, communication, or collaboration, and ultimately collective performance (Daly, 2012; Daly and Finnigan, 2010; Liou and Daly, 2018; Oh et al., 2004). These challenges also reflect opportunities for leaders to find ways to make these individuals more central if necessary in order to facilitate systemic change (Kilduff and Krackhardt, 2008). For instance, Daly and Finnigan's (2010) study on knowledge and communication networks among a district-wide leadership team indicates that the majority of school site principals tend to occupy peripheral or isolated positions compared with their central office administrative colleagues, which poses challenges to achieve changes at scale.

Betweenness

Another type of centrality that has received much attention to the study of leadership and change is betweenness centrality that measures the degree to which individuals occupy an intermediary position (brokering) between dyads of actors that are otherwise disconnected (Kilduff and Tsai, 2003). Betweenness helps identify those actors that bridge or broker and who help connect different clusters, groups or actors, and thereby play a key role in coordinating activities and information flow across a system (Balkundi and Kilduff, 2006; Hoppe and Reinelt, 2010). Few studies in educational leadership explore the broker role of educational leaders. For instance, Daly and colleagues (2014) examine brokerage patterns among superintendents in a large urban school district. They found a lack of brokering activities from area superintendents in bridging the central office and school sites. This lack of brokerage by leaders in the formal hierarchical system (area superintendents) led to a host of "informal" brokers moving resources and weakening overall district coherence around reform. Other studies that investigate betweenness of educational leaders, although reported a non-significant role of leader betweenness as a predictor (Moolenaar and Slegers, 2015) or an outcome (Wang et al., 2016), suggest the importance of betweenness as a key centrality position that cannot be overlooked in understanding the social influence of leadership.

Group and whole network level

The group and whole network level considers the overall network properties as well as subgroup structures that are embedded within the larger network. Network studies in this area focus on (1) explaining characteristics of whole network or group network structures or (2) using these network structures to predict other outcome variable(s). We discuss several network measures that are commonly used in network research in education to describe network structures: density, centralization, fragmentation, E-I index, and core-periphery. The density, centralization, and fragmentation are often used to assess cohesion of a network, while the E-I index and core-periphery analysis are mostly used to examine degree of closure for subgroup structures. The density of a network is the ratio of the observed ties to all possible ties. The centralization of a network refers to the degree to which the network is concentrated in one or a few focal nodes (Wasserman and Faust, 1994). A highly centralized network is characterized by a large variation in actor degree centralities in which a few nodes have large numbers of ties while most nodes have very few. In contrast, a less centralized network is one in which all nodes have similar degrees of connections. Focal nodes in a highly centralized network have disproportionate influence over resource flow, whereas nodes in a decentralized network share similar opportunities to access and exchange information/ideas and thus can help move resources faster from and to the periphery (Sasidharan et al., 2012).

Network fragmentation indicates the percentage of the number of disconnected dyads divided by all possible number of dyads in a network. The E-I index (external-internal index) measures the degree of closure at the whole network, subgroup, and individual levels. It takes the number of ties that are external to the groups minus the number of ties internal to the groups, and divided by the total number of ties (Krackhardt and Stern, 1988; Wasserman and Faust, 1994). The E-I index ranges from 1 (all ties are external) to -1 (all ties are internal), with 0 indicating a balance between external and internal interactions (Krackhardt and Stern, 1988). A core-periphery analysis seeks to identify two groups of actors in a network based on the degree of cohesion within each set of actors. A core-periphery network structure is defined as one that has a dense cohesive core of actors with the rest of actors loosely connected to one another as a periphery group (Everett and Borgatti, 1999).

Understanding these network/subgroup structures can help identify strengths and weaknesses in communication, information flow, and the shaping of group beliefs/norms necessary for change (Gibson and Vermeulen, 2003). In particular, subgroups (affiliated by interest, affection, or task) are often identified as important structures to the development of new ideas and problem-solving that move change forward, as they often share higher levels of trust (Finnigan and Daly, 2010). They can be capitalized to facilitate systemic changes by leveraging or connecting them (Daly, 2010; Daly et al., 2014; Liou and Daly, 2018).

A few studies in educational leadership have examined the overall network structures of educational leaders. For instance, Daly and Finnigan (2010) study the communication and knowledge network structures of central office and site leaders and suggest that the overall network tended to be sparsely connected (density) and highly centralized around a few central office leaders (centralization), with site principals tending to be on the periphery of the network. The E-I index result in their study suggested that leaders tended to interact more often with those sharing similar work level (district to district and site to site). They also demonstrated how such work level divide resulted in a core-periphery structure that prevented cross-level communication, information sharing, and ultimately systemic change. In a later work, Daly and Finnigan (2012) found a high degree of network fragmentation

among a districtwide leadership team from an underperforming school district, suggesting a lack of exchanges of leadership practices across the district.

Recent studies on educational leadership also explore these network-/group-level structures among a district's high school principals (Liou and Daly, 2018) or a cluster of schools within a community of learning (CoL) network (Sinnema et al., 2020). In Liou and Daly's study (2018), they found a lack of reform-related collaborations over time for the high school principals to connect with one another or with other levels' administrative colleagues, and an absence of these high school principals being in the core structure of the reform network. Sinnema et al.'s (2020) study focuses on the role of formal leaders (CoL leaders) and informal leaders (teachers) in a CoL network of collaborative relationships. Their findings indicate a sparse collaboration network and that only a quarter of the CoL leaders are in the core structure of this collaboration network with the majority of the core members being teachers, suggesting the emerging role of teachers in taking on leadership responsibilities that facilitate cross-school reform. Taken together, understanding the overall network structures illustrated by different metrics can help identify areas that may be better leveraged to support systems or systemic changes and improvement (Daly, 2012; Daly and Finnigan, 2016).

Framing networked systems leadership for change

Building on the abovementioned insights from existing network research in education, we provide a framework in Fig. 1 that depicts the social processes of leadership from a networked systems perspective. Leadership, in our view, is embedded not in particular individuals or roles, but in the relationships that exist among actors working toward outcomes in a networked system. This view shifts our attention to the social processes of leadership and suggests that individuals can draw on relational resources to influence or be influenced by others in order to enact their roles toward desired outcomes. Leadership thus lies in the systems of relations within which individuals are embedded. Individuals can exert leadership at different levels of the systems by their social influence affecting how their systems routines are structured and normalized, which in turn can shape the role, beliefs, and network behaviors of individuals.

Given the continuous social processes of leadership at various levels of social systems, systemic changes can be thought of as being shaped in two dimensional spaces: time and level of network. Over time, changes take shape as individuals engage in ongoing cyclic processes of social influence and (re)structuring of various levels of social systems in which they are nested. In other words, changes at the individual level will likely affect tie formation at higher-order levels (e.g., dyadic and subgroups), and in turn the overall structure of the whole network, and any changes at the network/subgroup level (e.g., internal/external policy or organizational restructuring) may also shape the attitudes, beliefs, norms, positional or social roles, and network behaviors of individuals, which continues to affect the process of network structuring at the higher levels of the systems. During the structuring process, individuals' characteristics may influence how their ties and beliefs are formed, and their beliefs, attitudes, and social positions can be further shaped by reconfiguring their social ties through networked practices (e.g., interventions to promote, constrain, or strategize connectivity). In other words, leadership is organizing dynamic social processes that is contingent upon relational resources one can draw on to enact their roles at different levels of social systems.

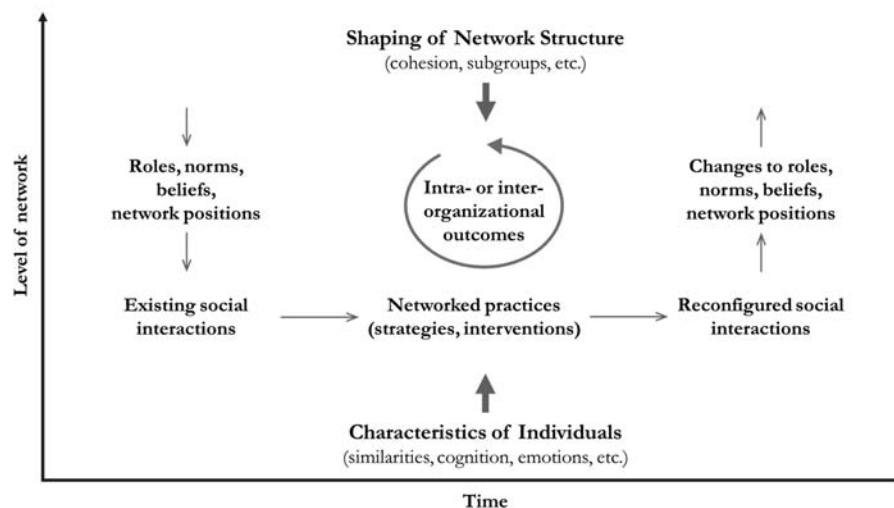


Fig. 1 Networked systems leadership for change.

Promising developments of networked systems leadership and change

As leadership is understood through complex social processes from a network perspective, social network analysis is a promising approach for the field. Although the use of social network analysis in leadership research has gained traction over the past decades, there remain areas of developments for future work, which we outline below.

Longitudinal dynamic networks

As leadership is shaped by relationships which can also shape individuals' social behaviors, a consideration of full dynamic models that account for both the selection model (individuals' social behaviors as function of their attributes or beliefs) and influence model (individuals' attributes or beliefs as function of their social behaviors) is desired. Network research often find correlations between peer influence and individuals' behaviors/views in which an actor's behaviors/views are shaped by the views/behaviors of their connected peers (social influence effect). It is also possible that it is the attributes of individuals (gender, beliefs, etc.) that affect the decision of their social selection. Such dynamic models can better represent complex social processes that are both outcomes and predictors influencing each other. A full dynamic model also requires the analysis of longitudinal data at two or more time points in order to yield stable results. Modeling these network dynamics can better understand how and why networks evolve over time, which enables educational researchers to further examine the underlying mechanisms involved in the shaping of social systems as related to leadership. Much of the existing research in educational leadership foregrounds either the selection or influence model with only a few studies employing these models (e.g., [Daly and Finnigan, 2011](#); [Siciliano et al., 2017](#)), future work may consider the use of full dynamic network models to investigate the social dynamics of leadership ([Frank et al., 2010](#)).

Two- or multi-mode networks and multilevel networks

The vast majority of network studies in education deals with one-mode (one-dimension) network in which the rows and columns in an adjacency matrix of network index one set of entities (e.g., actor by actor or event by event matrix). As social processes of leadership involve the engagement of relationships not only with people but a lot of times in meetings and/or professional development activities, information gathered based on one-mode network activities may only provide a partial view of leadership. Two-mode networks allow the field to examine relationships between sets of entities such as actors, subgroups, committees, affiliations, etc. A two-mode adjacency matrix includes the rows and columns each representing one set of entities (e.g., person by event, person by association) ([Everett and Borgatti, 2013](#)). One can imagine a two-mode network that illustrates relationships between educators and their activities, such as attending multiple professional activities or serving on more than one committee. These activities or events serve as conduits through which individuals/groups may influence one other ([Everett and Borgatti, 2013](#)). For instance, educators who attend multiple meetings have greater opportunity to share information or resources with their peers who only attend one meeting. Better understanding how information/resource flow from one meeting to another, strategies to leverage meetings with greater participation, and key informal leaders or brokers between meetings are crucial to enhance overall coherence in regard to change effort. Extending the two-mode network, one can even imagine analyzing multi-mode networks of educators, such as reform initiative, issues or concerns related to reform efforts, areas of expertise, and so forth.

While the two- or multi-mode networks address the width of activities involved in leadership, multi-level networks are concerned with the embedded nature of relationships for leadership. The studies discussed in prior section span multiple levels of the system, including: individual and dyads of teacher/administrator networks within schools or a larger district system, clusters of schools (principals or community leaders) within a district or community system, and the social or organizational system under study as a whole. These studies demonstrate that leadership and change exist at different levels of systems that are sometimes interacting with one another. As such, it is important to take into consideration data collected from multiple levels through a network perspective. As network data collects information about relationships and activities involved in leadership and is often nested (relationships in individuals in groups in systems), results from multilevel models provide a more nuanced and integrated understanding of phenomena that unfold across levels in systems.

Real time, multi-data sources in profiling networked leadership phenomena

Leadership occurs through interactions that are complex, nested, and multi-faceted in terms of data sources. We discuss a new trend in social network analysis in education that better captures the dimension (depth and width) and dynamics of social processes. The dimension of social processes involves where and how relationships take place and the nature of the relationships. It includes the collection of data from multiple sources in explaining the shaping of relationships, the content of such interactions, as well as the space where the relationships take place, e.g., online social media or physical in-person space (offline). The dynamics aspect refers to not only the social selection and influence effects as discussed earlier but also the movement among actors across different periods of time (e.g., network churn). Taken together it is the space, time, and data sources that provide a more holistic picture of leadership through a network lens.

The majority of existing research that examines the dynamic models primarily focuses on the face-to-face (offline) social activities of educators, as opposed to those that occur virtually (online) through social media (e.g., Twitter, etc.). A growing body of

scholarship calls for a full examination of educators' social activities connecting their offline and online social behaviors to best represent the social continuum (Daly et al., 2019a,b) or social dynamics of such phenomenon (Karimi et al., 2019; Pescosolido, 1992). This line of research examines educators' real time relational network data through their social media footprints to capture phenomena such as social diffusion of opinion leaders who are disseminating information and potentially influencing others' perceptions about topics (e.g., education policy) (Daly et al., 2019b), virtual communities of learning or professional groups (Liu et al., 2020). Other studies have come to integrate data sources from both spaces to examine classroom practice (Hu et al., 2020), administrators' sources of expertise and support (Daly et al., 2019a,b). Apart from the online and offline data sources, motivated by several mixed methods studies in social network analysis (e.g., Coburn and Russell, 2008; Daly and Finnigan, 2011; Froehlich et al., 2020; Judy Kamalodeen and Jameson-Charles, 2016; Liou and Daly, 2014), attending to the nature of interactions may yield more fruitful results than merely examining the presence or absence of a relational tie. Instruments that include questions concerning the depth of interactions (the when, what, why, and how of the interaction) may provide better insights into the social work of educators. This new approach to social processes is promising in disentangling complex behaviors and practices as related to networks and leadership in both online and offline spaces.

Conclusion

As the work of systemic change is ultimately enacted in systems by networks of individuals, theories and research attending to systems and social networks offer a promising lens for understanding social phenomena of change and leadership. In this article, we offer several contributions that connect social networks with leadership and change: (1) conceptualizing leader and leadership, (2) positioning leadership as a relational quality of systems, (3) linking social network concepts to leadership and change, (4) summarizing a framework supported by scholarship in education that takes a social network approach to educational leadership at different levels of systems, and (5) providing areas of developments in social network analysis in education that may offer different but complementary sources for understanding the social processes of leadership and change. Emerging themes in social network analysis supporting leadership development and systems change are important, yet underdeveloped, and as such represent high leverage theoretic and analytic points for future research in educational leadership.

References

- Aguillard, K., Goughnour, D.S., 2006. Central Office Inquiry: Assessing Organization, Roles, and Actions to Support School Improvement. WestEd, San Francisco, CA.
- Arnold, R.D., Wade, J.P., 2015. A definition of systems thinking: a systems approach. *Procedia Comput. Sci.* 44, 669–678.
- Balkundi, P., Kilduff, M., 2006. The ties that lead: a social network approach to leadership. *Leader. Q.* 17 (4), 419–439.
- Balkundi, P., Kilduff, M., Harrison, D.A., 2011. Centrality and charisma: comparing how leader networks and attributions affect team performance. *J. Appl. Psychol.* 96 (6), 1209.
- Bandura, A., 1971. Social learning theory. General Learning Press, New York.
- Bandura, A., 1977. Social Learning Theory. Prentice-Hall, Inc, Englewood Cliffs, NJ.
- Bandura, A., 2002. Social cognitive theory in cultural context. *Appl. Psychol.: Int. Rev.* 51, 269–290.
- Bandura, A., 2006. Toward a psychology of human agency. *Perspect. Psychol. Sci.* 1 (2), 164–180.
- Barnard, C.I., 1968. Functions of the Executive. Harvard University Press, Cambridge, MA.
- Baron, R.A., Markman, G.D., 2003. Beyond social capital: the role of entrepreneurs' social competence in their financial success. *J. Bus. Ventur.* 18 (1), 41–60.
- Bass, B., 1997. Does the transactional-transformational leadership paradigm transcend organizational and national boundaries? *Am. Psychol.* 52 (2), 130–139.
- Borgatti, S.P., Foster, P.C., 2003. The network paradigm in organizational research: a review and typology. *J. Manag.* 29 (6), 991–1013.
- Borgatti, S.P., Halgin, D.S., 2011. On network theory. *Organ. Sci.* 22 (5), 1168–1181.
- Borgatti, S.P., Ofem, B., 2010. Overview: social network theory and analysis. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard University Press, Cambridge, MA, pp. 17–29.
- Borgatti, S.P., Everett, M.G., Johnson, J.C., 2018. *Analyzing Social Networks*, second ed. SAGE Publications, Los Angeles, CA.
- Borgatti, S.P., Jones, C., Everett, M.G., 1998. Network measures of social capital. *Connections* 21 (2), 27–36.
- Brandes, U., Borgatti, S.P., Freeman, L.C., 2016. Maintaining the duality of closeness and betweenness centrality. *Soc. Network.* 44, 153–159.
- Brass, D.J., 1992. Power in organizations: a social network perspective. *Res. Politics Soc.* 4 (1), 295–323.
- Burns, J.M., 1978. *Leadership*. Harper & Row, NY.
- Burt, R.S., 1982. *Toward a Structural Theory of Action*. Academic Press, New York, NY.
- Burt, R.S., 1997. A note on social capital and network content. *Soc. Network.* 19, 355–373.
- Burt, R.S., 2005. *Brokerage and closure: An introduction to social capital*. Oxford University Press.
- Burt, R.S., 2007. Closure and stability: persistent reputation and enduring relations among bankers and analysts. In: Rauch, J.E. (Ed.), *The Missing Links: Formation and Decay of Economic Networks*. Russell Sage Foundation, New York, NY, pp. 100–143.
- Coburn, C., 2003. Rethinking scale: moving beyond numbers to deep and lasting change. *Educ. Res.* 32 (6), 3–12.
- Coldren, A., Spillane, J.P., 2007. Making connections to teaching practice: the role of boundary practices in instructional leadership. *Educ. Pol.* 21 (2), 369–396.
- Coburn, C.E., Russell, J.L., 2008. District policy and teachers' social networks. *Educ. Eval. Pol. Anal.* 3 (3), 203–235.
- Coburn, C.E., Mata, W.S., Choi, L., 2013. The embeddedness of teachers' social networks: evidence from a study of mathematics reform. *Sociol. Educ.* 86 (4), 311–342.
- Cole, R.P., Weinbaum, E.H., 2010. Changes in attitude: peer influence in high school reform. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard University Press, Cambridge, MA, pp. 77–96.
- Connelly, D.R., 2007. Leadership in the collaborative interorganizational domain. *Int. J. Publ. Adm.* 30 (11), 1231–1262.
- Daly, A.J., 2012. Data, dyads, and dynamics: exploring data use and social networks in educational improvement. *Teach. Coll. Rec.* 114 (11), 1–38.
- Daly, A.J. (Ed.), 2010. *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge, MA.
- Daly, A.J., Finnigan, K., 2010. A bridge between worlds: understanding network structure to understand change strategy. *J. Educ. Change* 11 (2), 111–138.
- Daly, A.J., Finnigan, K., 2011. The ebb and flow of social network ties between district leaders under high stakes accountability. *Am. Educ. Res. J.* 48, 39–79.

- Daly, A.J., Finnigan, K. (Eds.), 2016. *Thinking and Acting Systemically: Improving School Districts under Pressure*. AERA Publishing, Washington DC.
- Daly, A.J., Liou, Y., Tran, N.A., Cornelissen, F., Park, V., 2014. The rise of neurotics: social networks, leadership, and efficacy in district reform. *Educ. Adm. Q.* 50 (2), 233–278.
- Daly, A.J., Liou, Y.-H., Brown, C., 2016. Social red bull: exploring energy relationships in a school district leadership team. *Harv. Educ. Rev.* 86 (3), 412–448.
- Daly, A.J., Liou, Y.-H., Del Fresno, M., Rehm, M., Bjorklund Jr., P., 2019a. Educational leadership in the Twitter verse: social media, social networks, and the new social continuum. *Teach. Coll. Rec.* 121 (14), 1–20.
- Daly, A.J., Moolenaar, N.M., Carrier, N., Del Fresno, M., 2017. The school district as a complex adaptive system. Exploring a complexity theory perspective on policy. *J. Res. Organ. Edu.* 1, 31–59.
- Daly, A.J., Supovitz, J.A., Del Fresno, M., 2019b. The social side of educational policy: how social media is changing the politics of education. *Teach. Coll. Rec.* 121 (14), 1–26.
- Daly, A.J., Liou, Y.-H., Der-Martirosian, C., 2020. A capital idea: exploring the relationship between human and social capital and student achievement in schools. *JPCC* 6 (1), 7–28.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Deal, T.E., Peterson, K.D., 2000. Eight roles of symbolic leaders. In: Grogan, M. (Ed.), *The Jossey-Bass Reader on Educational Leadership*. Jossey-Bass, pp. 274–286.
- Degenne, A., Forsé, M., 1999. *Introducing Social Networks*. Sage, London, UK.
- Diaz-Gibson, J., Zaragoza, M.C., Daly, A.J., Mayayo, J.L., Romani, J.R., 2017. Networked leadership in educational collaborative networks. *Educ. Manag. Adm. Leader* 45 (6), 1040–1059.
- Donaldson Jr., G.A., 2006. *Cultivating Leadership in Schools: Connecting People, Purpose, and Practice*. Teachers College, New York, NY.
- Everett, M.G., Borgatti, S.P., 1999. The centrality of groups and classes. *J. Math. Sociol.* 23 (3), 181–201.
- Everett, M.G., Borgatti, S.P., 2013. The dual-projection approach for two-mode networks. *Soc. Network.* 34 (2), 204–210.
- Finnigan, K., Daly, A., 2010. Learning at a system level: ties between principals of low-performing schools and central office leaders. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge, MA, pp. 179–196.
- Frank, K.A., Zhao, Y., 2005. Subgroups as meso-level entities in the social organization of schools. In: Hedges, L.V., Schneider, B. (Eds.), *Social Organization of Schooling*. Russell Sage Foundation, New York, NY, pp. 200–224.
- Frank, K.A., Kim, C., Belman, D., 2010. Utility theory, social networks, and teacher decision making. In: Daly, A.J. (Ed.), *Social Network Theory and Educational Change*. Harvard Education Press, Cambridge, MA, pp. 223–242.
- Frank, K.A., Penuel, W.R., Sun, M., Kim, C., Singleton, C., 2013. The organization as a filter of institutional diffusion. *Teach. Coll. Rec.* 115 (1), 306–339.
- Fröhlich, D.E., Van Waes, S., Schäfer, H., 2020. Linking quantitative and qualitative network approaches: a review of mixed methods social network analysis in education research. *Rev. Res. Educ.* 44 (1), 244–268.
- Gardner, J.W., 2000. The nature of leadership. In: *The Jossey-Bass Reader on Educational Leadership*. Jossey-Bass, San Francisco, CA, pp. 3–12.
- Gibson, C., Vermeulen, F., 2003. A healthy divide: subgroups as a stimulus for team learning behavior. *Adm. Sci. Q.* 48 (2), 202–239.
- Goldspink, C., 2007. Rethinking educational reform: a loosely coupled and complex systems perspective. *Educ. Manag. Adm. Leader* 35 (1), 27–50.
- Granovetter, M.S., 1973. The strength of weak ties. *Am. J. Sociol.* 78 (6), 1360–1380.
- Henttonen, K., Janhonen, M., Johanson, J.E., 2013. Internal social networks in work teams: structure, knowledge sharing and performance. *Int. J. Manpow.* 34 (6), 616–634.
- Hoberecht, S., Joseph, B., Spencer, J., Southern, N., 2011. Inter-organizational networks: an emerging paradigm of whole systems change. *OD Practitioner* 43 (4), 23–27.
- Honig, M., 2006. Street level bureaucracy revisited: frontline district central-office administrators as boundary spanners in education policy implementation. *Educ. Eval. Pol. Anal.* 28 (4), 357–383.
- Honig, M., 2008. District central offices as learning organizations: how socio-cultural and organizational learning theories elaborate district central-office administrators' participation in teaching and learning improvement efforts. *Am. J. Educ.* 114, 627–664.
- Honig, M.I., Ikemoto, G.S., 2008. Adaptive assistance for learning improvement efforts: the case of the Institute for learning. *Peabody J. Educ.* 83 (3), 328–363.
- Hoppe, B., Reinelt, C., 2010. Social network analysis and the evaluation of leadership networks. *Leader. Q.* 21 (4), 600–619.
- Hu, S., Torphy, K., Evert, K., Lane, J., 2020. From cloud to classroom: mathematics teachers' planning and enactment of resources accessed within virtual spaces. *Teach. Coll. Rec.* 122 (6), 1–33.
- Hubers, M.D., Moolenaar, N.M., Schildkamp, K., Daly, A.J., Handelzalts, A., Pieters, J.M., 2018. Share and succeed: the development of knowledge sharing and brokerage in data teams' network structures. *Res. Pap. Educ.* 33 (2), 216–238.
- Ibarra, H., 1995. Race, opportunity, and diversity of social circles in managerial networks. *Acad. Manag. J.* 38 (3), 673–703.
- Judy Kamaldeen, V., Jameson-Charles, M., 2016. A mixed methods research approach to exploring teacher participation in an online social networking website. *Int. J. Qual. Methods* 15 (1). <https://doi.org/10.1177/1609406915624578>.
- Karimi, H., Derr, T., Torphy, K., Frank, K., Tang, J., 2019. A roadmap for incorporating online social media in educational research. *Teach. Coll. Rec.* 121 (14), 1–24.
- Kezar, A., 2011. What is the best way to achieve broader reach of improved practices in higher education? *Innovat. High. Educ.* 36 (4), 235–247.
- Kezar, A., 2018. *Scaling Improvement in STEM Learning Environments: The Strategic Role of a National Organization*. Association of American Universities, Washington, D.C.
- Kilduff, M., Brass, D.J., 2010. Organizational social network research: core ideas and key debates. *Acad. Manag. Ann.* 4 (1), 317–357.
- Kilduff, M., Krackhardt, D., 2008. *Interpersonal Networks in Organizations: Cognition, Personality, Dynamics, and Culture: Structural Analysis in the Social Sciences*. Cambridge University Press, Cambridge, UK.
- Kilduff, M., Tsai, W., 2003. *Social Networks and Organizations*. Sage, London.
- Krackhardt, D., 1999. The ties that torture: simmelian tie analysis in organizations. In: Bacharach, S.B., Andrews, S.B., Knoke, D. (Eds.), *Research in the Sociology of Organizations*. JAI Press, Stamford, CT, pp. 183–210.
- Krackhardt, D., Kilduff, M., 1990. Friendship patterns and culture: the control of organizational diversity. *Am. Anthropol.* 92 (1), 142–154.
- Krackhardt, D., Stern, R., 1988. Informal networks and organizational crises: an experimental simulation. *Soc. Psychol. Q.* 51 (2), 123–140.
- Kramer, M.W., Day, E.A., Nguyen, C., Hoelscher, C.S., Cooper, O.D., 2019. Leadership in an interorganizational collaboration: a qualitative study of a statewide interagency taskforce. *Hum. Relat.* 72 (2), 397–419.
- Labianca, G., Brass, D.J., 2006. Exploring the social ledger: negative relationships and negative asymmetry in social networks in organizations. *Acad. Manag. Rev.* 31 (3), 596–614.
- Lazarsfeld, P., Merton, R., 1954. Friendship as social process: a substantive and methodological analysis. In: Berger, M., Abel, T., Page, C. (Eds.), *Freedom and Control in Modern Society*. Van Nostrand, New York, NY, pp. 18–66.
- Leithwood, K., 2017. The Ontario leadership framework: successful school leadership practices and personal leadership resources. In: Leithwood, K., Sun, J., Pollock, K. (Eds.), *How School Leaders Contribute to Student Success: The Four Paths Framework*. Springer, Cham, Switzerland, pp. 31–48.
- Leithwood, K., 2019. Characteristics of effective leadership networks: a replication and extension. *Sch. Leader. Manag.* 39 (2), 175–197.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22.
- Leithwood, K., Jantzi, D., 2000. The effects of transformational leadership on organizational conditions and student engagement with school. *J. Educ. Adm.* 38 (2), 112–129.
- Leithwood, K.A., Riehl, C., 2003. *What We Know about Successful School Leadership*. Laboratory for Student Success, Temple University, Philadelphia, PA.
- Lin, N., 2009. *Social Capital: A Theory of Social Structure and Action*, eighth ed. Cambridge University Press, New York, NY.
- Lincoln, J.R., Miller, J., 1979. Work and friendship ties in organizations: a comparative analysis of relation networks. *Adm. Sci. Q.* 24, 181–199.
- Liou, Y.-H., Canrinus, E.T., 2020. A capital framework for professional learning and practice. *Int. J. Educ. Res.* 100, 101527. <https://doi.org/10.1016/j.ijer.2019.101527>.
- Liou, Y.-H., Bjorklund Jr., P., 2021. The rise and fall of influential expertise: Disentangling tie formation and dissolution in a professional learning school network. Paper accepted to the American Educational Research Association. Virtual Conference.

- Liou, Y.-H., Daly, A.J., 2014. Closer to learning: social networks, trust, and professional communities. *J. Sch. Leader.* 24 (4), 753–795.
- Liou, Y.-H., Daly, A.J., 2018. Broken bridges: a social network perspective on urban high school leadership. *J. Educ. Adm.* 56 (5), 562–584.
- Liou, Y.-H., Daly, A.J., 2020a. Obstacles and opportunities for networked practice: a social network analysis of an inter-organizational STEM ecosystem. *J. Educ. Adm.* 59 (1), 94–115.
- Liou, Y.-H., Daly, A.J., 2020b. Investigating leader self-efficacy through policy engagement and social network position. *Educ. Pol.* 34 (3), 411–448.
- Liou, Y.-H., Daly, A.J., Brown, C., del Fresno, M., 2015. Foregrounding the role of relationships in reform: a social network perspective on leadership and change. *Int. J. Educ. Manag.* 29 (7), 819–837.
- Liou, Y.-H., 2016. Tied to the common core: exploring the characteristics of reform advice relationships of educational leaders. *Educ. Adm. Q.* 52 (5), 793–840.
- Liou, Y.-H., Daly, A.J., 2019a. The lead igniter: a longitudinal examination of influence and energy through networks, efficacy, and climate. *Educ. Adm. Q.* 55 (3), 363–403.
- Liou, Y.-H., Daly, A.J., 2019b. The networked leader: understanding peer influence in a system wide leadership team. *Sch. Leadersh. Manag.* 40 (2-3), 163–182.
- Liu, Y., Torphy, K.T., Hu, S., Tang, J., Chen, Z., 2020. Examining the virtual diffusion of educational resources across teachers' social networks over time. *Teach. Coll. Rec.* 122 (6), 1–34.
- McCaffrey, D.P., Faerman, S.R., Hart, D.W., 1995. The appeal and difficulties of participative systems. *Organ. Sci.* 6 (6), 603–627.
- McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 27, 415–444.
- Monge, P.R., Contractor, N., 2003. *Theories of Communication Networks*. Oxford University Press, Oxford.
- Moolenaar, N.M., Slegers, P.J.C., 2015. The networked principal Examining principals' social relationships and transformational leadership in school and district networks. *J. Educ. Adm.* 53 (1), 8–39.
- Moolenaar, N.M., Daly, A.J., Slegers, P.J.C., 2010. Occupying the principal position: examining relationships between transformational leadership, social network position, and schools' innovative climate. *Educ. Adm. Q.* 46 (5), 623–670.
- Murphy, J., Louis, K.S., Smylie, M., 2017. Positive school leadership: how the professional standards for educational leaders can be brought to life. *Phi Delta Kappan* 99 (1), 21–24.
- Ogawa, R.T., Bossert, S.T., 2000. Leadership as an organizational quality. In: *The Jossey-Bass Reader on Educational Leadership*. Jossey-Bass, San Francisco, CA, pp. 38–58.
- Oh, H., Chung, M.H., Labianca, G., 2004. Group social capital and group effectiveness: the role of informal socializing ties. *Acad. Manag. J.* 47 (6), 860–875.
- Penuel, W.R., Riel, M.R., Krause, A., Frank, K.A., 2009. Analyzing teachers' professional interactions in a school as social capital: a social network approach. *Teach. Coll. Rec.* 111 (1), 124–163.
- Penuel, W.R., Riel, M., Joshi, A., Pearlman, L., Kim, C.M., Frank, K.A., 2010. The alignment of the informal and formal organizational supports for reform: implications for improving teaching in schools. *Educ. Adm. Q.* 46 (1), 57–95.
- Penuel, W.R., Sun, M., Frank, K.A., Gallagher, H.A., 2012. Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. *Am. J. Educ.* 119 (1), 103–136.
- Pescosolido, B.A., 1992. Beyond rational choice: the social dynamics of how people seek help. *Am. J. Sociol.* 97 (4), 1096–1138.
- Pitts, V., Spillane, J.P., 2009. Using social network methods to study school leadership. *Int. J. Res. Method Educ.* 32 (2), 185–207.
- Popp, J., MacKean, G., Casebeer, A., Milward, H.B., Lindstrom, R., 2013. *Inter-organizational Networks. A Critical Review of the Literature to Inform Practice*. IBM Center for The Business Development, Washington, DC.
- Provan, K.G., Lemaire, R.H., 2012. Core concepts and key ideas for understanding public sector organizational networks: using research to inform scholarship and practice. *Publ. Adm. Rev.* 72 (5), 638–648.
- Putnam, R.D., 2000. *Bowling alone, The collapse and revival of American community*. Simon and Schuster, New York.
- Russell, J.L., Bryk, A.S., Dolle, J., Gomez, L.M., LeMahieu, P., Grunow, A., 2017. A framework for the initiation of networked improvement communities. *Teach. Coll. Rec.* 119 (7), 1–36.
- Sasidharan, S., Santhanam, R., Brass, D.J., Sambamurthy, V., 2012. The effects of social network structure on enterprise systems success: a longitudinal multilevel analysis. *Inf. Syst. Res.* 23 (3-part-1), 658–678.
- Scott, J., 2017. *Social Network Analysis*, fourth ed. Sage Publications, London, UK.
- Senge, P.M., 2006. *The Fifth Discipline: The Art & Practice of the Learning Organization*, second ed. Doubleday/Currency, New York, NY.
- Shaked, H., Schechter, C., 2019. Exploring systems thinking in school principals' decision-making. *Int. J. Leader. Educ.* 22 (5), 573–596.
- Siciliano, M.D., Thompson, J.R., 2018. If you are committed, then so am I: the role of social networks and social influence on organizational commitment. *Adm. Soc.* 50 (7), 916–946.
- Siciliano, M.D., Moolenaar, N.M., Daly, A.J., Liou, Y.-H., 2017. A cognitive perspective on policy implementation: reform beliefs, sensemaking, and the common core state standards. *Publ. Adm. Rev.* 77 (6), 889–901.
- Sinnema, C., Daly, A.J., Liou, Y.-H., Rodway, J., 2020. Exploring the communities of learning policy in New Zealand using social network analysis: a case study of leadership, expertise, and networks. *Int. J. Educ. Res.* 99.
- Spillane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco, CA.
- Spillane, J.P., 2017. Leadership and learning: conceptualizing relations between school administrative practice and instructional practice. In: Leithwood, K., Sun, J., Pollock, K. (Eds.), *How School Leaders Contribute to Student Success: The Four Paths Framework*. Springer, Cham, Switzerland, pp. 49–68.
- Spillane, J.P., Diamond, J.B. (Eds.), 2007. *Distributed Leadership in Practice*. Teachers College Press, New York, NY.
- Spillane, J.P., Shirrell, M., 2017. Breaking up isn't hard to do: exploring the dissolution of teachers' and school leaders' work-related ties. *Educ. Adm. Q.* 53 (4), 616–648.
- Spillane, J.P., Hopkins, M., Sweet, T.M., 2018. School district educational infrastructure and change at scale: teacher peer interactions and their beliefs about mathematics instruction. *Am. Educ. Res. J.* 55 (3), 532–571.
- Spillane, J.P., Kim, C.M., 2012. An exploratory analysis of formal school leaders' positioning in instructional advice and information networks in elementary schools. *Am. J. Educ.* 119 (1), 73–102.
- Spillane, J.P., Kim, C., Frank, K.A., 2012. Instructional advice and information providing and receiving behavior in elementary schools: exploring tie formation as a building block in social capital development. *Am. Educ. Res. J.* 49 (6), 1112–1145.
- Sun, M., Frank, K.A., Penuel, W., Kim, C.M., 2013. How external institutions penetrate schools through formal and informal leaders. *Educ. Adm. Q.* 49 (4), 610–644.
- Supovitz, J., Sirinides, P., May, H., 2010. How principals and peers influence teaching and learning. *Educ. Adm. Q.* 46 (1), 31–56.
- Thompson, J.D., 1967. *Organizations in Action*. McGraw-Hill, New York, NY.
- Tsai, W., 2000. Social capital strategic relatedness and the formation of intraorganizational linkages. *Strat. Manag. J.* 21, 925–939.
- Umphress, E.E., Labianca, G., Brass, D.J., Kass, E., Scholten, L., 2003. The role of instrumental and expressive social ties in employees' perceptions of organizational justice. *Organ. Sci.* 14 (6), 738–753.
- Wang, Y., Sauers, N.J., Richardson, J.W., 2016. A social network approach to examine K-12 educational leaders' influence on information diffusion on Twitter. *J. Sch. Leader.* 26 (3), 495–522.
- Wasserman, S., Faust, K., 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, NY.
- Weber, E.P., Khademian, A.M., 2008. Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Publ. Adm. Rev.* 68 (2), 334–349.

Longitudinal designs for school leadership research

Ronald H. Heck, University of Hawaii at Manoa, Honolulu, HI, United States

© 2023 Elsevier Ltd. All rights reserved.

Some background	537
Features of longitudinal data	539
Longitudinal designs in school leadership research	540
Two example longitudinal studies	540
Day et al. (2009) study	540
Bryk et al. (2010) study	541
Some other useful designs	541
Quasi-experimental designs	542
Latent variable mixture models	542
Concluding thoughts	543
References	543

Over the past several decades, research on school leadership has developed a theoretically informed empirical knowledge demonstrating that school leadership contributes to student learning through strategic actions that focus on changing a constellation of sociocultural, structural, and academic processes (e.g., Bryk et al., 2010; Day et al., 2009; Hallinger and Heck, 2010; Leithwood et al., 2004, 2020; Robinson et al., 2008). Despite a well-documented substantive knowledge base describing how school leadership facilitates school improvement, scholars have highlighted two related shortcomings. First, relatively more attention has been directed at the substance of the findings than the appropriateness and quality of the methods utilized to support the knowledge claims (Goff and Finch, 2016). Stronger methods underpinning substantive results can reduce rival explanations. Previous scholars documented that such studies are in relatively short supply in the field of school leadership and improvement (Feldhoff et al., 2014; Hallinger and Heck, 2011; Thoonen et al., 2012).

Second, other than identifying potential associations between leadership and mediating variables, there is still a need to link studies of leadership and its relative distribution within the school more closely to school changes in learning practice, which can bring us closer to addressing the question of how different contextual and school conditions and patterns of school leadership interactions and strategies contribute to growth in student learning outcomes (Harris, 2008). Over the past decade, recognition that leadership can result from broad leadership sources within the school affects how school leadership is conceptualized in longitudinal studies focusing on improvement—that is, as an organizational quality of schools versus a collection of personal characteristics of the individual in the role (Ogawa and Bossert, 1995). Although scholars defined meaningful distinctions between terms such as distributed (Gronn, 2002; Spillane, 2006), shared (Marks and Printy, 2003; Pounder et al., 1995), and collaborative leadership (Bolden, 2011; Hallinger, 2013), all three terms reflect a similar concern for broadening the sources of school leadership. As Leithwood et al. (2020) recently concluded, the function of leadership at all levels, or distributed leadership, is to build the organizational conditions that foster high quality teaching and generate improvements in learner outcomes. School improvement, by definition, entails a change in the state of the organization over some period of time (Luyten et al., 2005). Yet, Leithwood et al.'s review showed that there is less understanding of how school leaders enact changes within the school in ways that facilitate improvement in educational processes that result in improved student learning. This implies the need to study school leadership and its potential to bring about changes in such organizational conditions over longer periods of time, as it takes considerable time for improvement in outcomes to unfold (Heck and Chang, 2017; Louis and Miles, 1991).

Research methods provide an important, but often unacknowledged, link to the development of a valid knowledge base underlying the practice of school leadership (Feldhoff et al., 2014; Goff and Finch, 2016; Heck and Hallinger, 1999). In one recent review of methods underlying school leadership and student learning research, Goff and Finch (2016) noted the need for stronger research designs underlying knowledge claims in the field (i.e., as most previous research was cross-sectional or case study) and greater attention directed to the temporal ordering of variables that can bring about improvement through the use of longitudinal designs and data collection. Because school improvement involves changes in the state of the school's processes, it requires designs and data that can take into account changing relationships among relevant variables (Feldhoff et al., 2014). Unfortunately, it can be challenging to obtain sufficient data that attend to both the hierarchical (and cross-classified) nature of schools as organizations through having available subject identifiers and the necessary temporal ordering of educational process variables and student achievement data over sufficient periods of time.

Given the relevance of school improvement as a policy goal in an era of global educational accountability and, in particular, school leadership, in facilitating academic improvement over time, the purpose of this chapter is to develop the rationale for the increased use longitudinal designs and data collection in future school leadership research. To begin, I provide a context for the use of longitudinal data in examining school improvement. I then comment on two longitudinal studies using varied research

designs and longitudinal data that have extended our understanding of a problem previously investigated primarily with cross-sectional data. Finally, I outline a couple of promising types of longitudinal designs that are likely to contribute to knowledge regarding the intersection between school leadership and school improvement over the next decade.

Some background

Since the mid-1960s, numerous scholars documented the substantive progress in developing the knowledge base regarding school leadership and its effects on school processes and student outcomes (e.g., Bossert et al., 1982; Boyan, 1988; Bridges, 1982; Edmonds, 1979; Erickson, 1967; Hallinger and Heck, 1996a, 1998; Harris, 2008; Hendriks and Scheerens, 2013; Leithwood, 1994; Leithwood and Duke, 1999; Leithwood and Montgomery, 1982; Leithwood et al., 2008, 2020; Marks and Printy, 2003; Ribbins and Gunter, 2002; Robinson et al., 2008; Willower and Forsyth, 1999; Witziers et al., 2003). Such reviews are useful in identifying trends in knowledge development, understanding emerging issues in the realm of practice, and critiquing research methods used in generating knowledge.

In addition to identifying studies that document pathways through which school leadership may influence community, school, and classroom processes that can enhance student learning (e.g., Bossert et al., 1982; Bridges, 1982; Hallinger and Heck, 1996a; Hendriks and Scheerens, 2013; Kruger et al., 2007; Leithwood et al., 2004, 2010, 2017, 2020; Robinson et al., 2008), a second tangible result of reviews examining scholarly progress has been a broader view of the sources of school leadership than simply the school principal (e.g., Gronn, 2002; Harris, 2008; Leithwood et al., 2020; Ogawa and Bossert, 1995; Spillane, 2006). Historically, research on school leadership focused predominantly on the principal's role, given the inclination to assume leadership was an individual quality exercised by those in formal leadership roles (Heck and Hallinger, 1999).

A third result is the recognition that school improvement, and leadership's role in facilitating it, is more of a "journey" than a specific event, and contrasting leadership configurations and strategies which may be emphasized, are in part, determined by the challenges schools may face in relation to where they are located in their particular improvement journeys (e.g., Bryk et al., 2010; Day et al., 2009; Hallinger and Heck, 2011; Heck and Reid, 2020; Jackson, 2000). At any point in time, the school's efforts to improve may be shaped by multiple factors, including its intake of students, the quality and engagement of its faculty and administrators in improvement, as well as its academic and social organization (e.g., Bossert et al., 1982; Hallinger and Heck, 2011; Heck and Reid, 2020; Leithwood et al., 2020). This implies the need to study school leadership and its potential to bring about changes in normative structures, educational environments, academic processes, and student learning outcomes over longer periods of time than commonly thought—perhaps four to seven years, or longer—rather than two to three years (Bryk et al., 2010; Heck and Chang, 2017; Louis and Miles, 1991).

Although the field of school leadership and school improvement has generated much scholarly interest over the past 50 years, most reviews of scholarly progress concluded that such reviews focused primarily on the status of substantive findings rather than methods of inquiry (e.g., Goff and Finch, 2016). There have, however, been a number of methodological reviews over the previous 50 years that attempted to integrate the substance of research findings with theoretical models and methods underpinning the individual studies in an attempt to evaluate or enhance methodological progress in developing a school leadership knowledge base (e.g., Bridges, 1982; Erickson, 1967; Everhart, 1988; Goff and Finch, 2016; Hallinger and Heck, 1996b; Halpin, 1957; Heck and Hallinger, 1999, 2005; Murphy, 1988; Pitner, 1988; Tatsuoka and Silver, 1988). Although identifying some promising approaches, the reviews suggested scholars mostly struggled to identify the field's key problems, conceptual frameworks, methods of inquiry, and a knowledge base that can directly inform improvement practices in schools.

Early reviewers concluded that the tradition of research on the impact of principals did not generally do justice to the complexity of the topic in terms of theoretical and methodological sophistication or application to school settings (Bossert et al., 1982; Bridges, 1982; Erickson, 1967). For example, in reviewing methodological progress from the mid-1960s to 1980, Bridges concluded progress was dismal, noting "Despite the rather loose definition of theory that was used in classifying the sample of research ..., most of it proved to be atheoretical. Likewise, the research seemed to have little or no practical utility" (pp. 24–25). In reviewing progress 15 years later, Phil Hallinger and I concluded that although school leadership research had contributed to substantive progress in understanding how leadership affects school outcomes through several mediating processes, several methodological shortcomings remained (Hallinger and Heck, 1996a,b). Our primary concerns were the utilization of simplified theoretical models to examine complex sets of dynamic relationships, overall reliance on a small set of analytic tools, and limitations of cross-sectional data in testing proposed theoretical relationships, including the effects of environmental and organizational conditions that moderate the school's capacity for improving student learning. Later, in revisiting methodological progress, we noted continued quantitative progress in studying selected domains of school leadership, as well as the acceptance of varied qualitative methods in examining a broader range of issues in researching school leadership (Heck and Hallinger, 1999, 2005). We concluded these latter approaches of situating school leadership research within multiple contexts was advantageous in addressing blind spots in our knowledge and disciplinary practice, for example, focusing on the moral and ethical ends of leadership in addition to its effects (Anderson, 2004; Dillard, 1995; Gronn, 2002; Gronn and Ribbins, 1996; Spillane et al., 2001).

More recently, Feldhoff et al. (2014) and Goff and Finch's (2016) reviews of methodological progress in studying school leadership and learning emphasized the need for greater attention to the conceptual frameworks, research designs, and analytic techniques that underlie the robustness of the findings of previous studies. Feldhoff et al. argued that school improvement studies need to take into account both the multilevel structure of schools and the need for longitudinal measurement of school improvement effects.

Because of the multilevel nature of educational organizations, the recognition of its hierarchical levels (e.g., students, classrooms, schools, districts) is key to advancing and properly testing propositions regarding macro (contextual) and micro (individual behavior) linkages in examining school, classroom, and student achievement processes. Such macro-level factors moderate relationships at lower levels—leadership to teacher professional practices and teacher practices to students (e.g., Heck and Hallinger, 2014; Slegers et al., 2014). Goff and Finch also concluded that previous analytic methods used to examine school leadership were generally inadequate to deal with the realities associated with leadership and school improvement.

Scholars reviewing the various contributions of previous research on school leadership, consistently highlighted three specific problems. First, many commonly accepted findings were generalized from cross-sectional studies, rather than larger-scale longitudinal studies that could better account for the temporal ordering of theoretical relationships. One primary criticism was that the research seldom gave adequate attention to the careful delineation of research methods including the measurement of key constructs and the temporal ordering of those constructs in order to link leadership to student learning and to build a sound knowledge base in school improvement research (Creemers and Kyriakides, 2008; Hallinger and Heck, 1996a; Luyten et al., 2005; Reynolds et al., 2000; Sammons and Luyten, 2009; Witziers et al., 2003). Because cross-sectional data are unable to resolve the ambiguity in correlations or other measures of association, it is more difficult to rule out alternative explanations regarding proposed relationships.

The over-reliance on cross-sectional designs hampered understanding how school leaders enact strategies and programs that affect school improvement over time (Hallinger and Heck, 2011; Harris, 2008; Scheerens, 2012). More specifically, the *snap shot* nature of cross-sectional research is useful in identifying factors that may differentiate higher producing versus lower producing schools, but it cannot inform regarding the nature of the process that leads to one set of outcomes over the other (Goff and Finch, 2016; Luyten et al., 2005). Because organizational change is a dynamic process with interdependent factors that change through time, improvement implies a change in organizational conditions (e.g., resources, capacity, attitudes, behaviors, outcomes) and, therefore, is most appropriately studied from a longitudinal perspective (Feldhoff et al., 2014).

Causal claims depend on a temporal sequence with identified causal mechanisms that are hypothesized to account for expected outcomes (Cook and Campbell, 1979; Goff and Finch, 2016). Conceptual frameworks help contribute to the accumulation of knowledge by providing a basis for comparing the results of a current study with previous studies within a logical and consistent framework. Although there are multiple research designs to get at causal inference (e.g., clinical trials, time series data with introduced treatment, regression discontinuity) one popular approach uses longitudinal panel data made up of multiple units measured at multiple occasions (Zyphur et al., 2019). It may (or may not) include an intervention at some occasion during the temporal sequence. Such data are useful for causal inference by helping to control for potential confounds. As Feldhoff et al. (2014) argued, the assignment of such attributions is facilitated by reliable theoretical justification and the measurement of independent and dependent variables at different points in time (Gustafsson, 2010; Hallinger and Heck, 2011; Marsh and Craven, 2006; Zyphur et al., 2019). Hallinger and Heck concluded that causal ordering represents an important theoretical issue with practical implications in studies of leadership related to student learning.

In the absence of experimental designs, which do not readily lend themselves to studying school improvement, such studies are largely natural experiments to improve student learning, conducted in widely different external conditions (e.g., policy environments) and with associated strengths or weaknesses in essential mediating supports (e.g., Bryk et al., 2010; Day et al., 2009; Hallinger and Heck, 2010). They require laying out expected results and building a case regarding *what* changed and *how* based on multiple types of time-related data. Importantly, innovations in both data collection and analytic approaches can lead to further insights often missed in more simplified previous studies of “associations” between variables. For example, as Bryk et al. noted in their longitudinal study of school improvement, utilizing repeated survey data on school conditions, rather than data collected at only one occasion, provided more time-sensitive measures of changes in the essential supports in explaining subsequent changes in school achievement and, therefore, greater confidence about the conclusions.

Second, early studies framed principals as the causal agents for organizational change (e.g., Edmonds, 1979). More recent studies provided increasing evidence that school leadership is an organizational quality, or unit of analysis, rather than an individual quality (Ogawa and Bossert, 1995; Spillane et al., 2001) and, as such, may be relatively more (or less) distributed throughout the school (Gold et al., 2002; Gronn, 2002; Harris, 2008; Spillane, 2006). This newer conceptualization impacted how leadership was conceptualized in longitudinal studies; that is, as a quality of schools versus a quality of the individual in the role. Initial studies identified the positive impact of leadership that was more distributed—decentralized in terms of responsibility, participation, and engagement and collaborative and interdependent in terms of involvement and monitoring of implementation—which impacted student achievement indirectly through school-wide strategic efforts aimed at improving school organizational conditions that support teaching and student learning (e.g., Camburn et al., 2003; Heck and Hallinger, 2009; Leithwood et al., 2020; Spillane et al., 2011). For example, Spillane and colleagues found leadership was distributed most successfully when it was seen to represent a form of collective responsibility, intelligence and sense making, with leadership for school improvement emerging as an interactive process of influence designed to achieve organizational ends. Leithwood et al. noted the function of leadership at all levels, or distributed leadership, is to build the organizational conditions that foster high quality teaching and generate improvements in learner outcomes, the success of which depends on opportunities for discretionary decision making by those enacting leadership.

Third, often previous research was criticized as falling short in terms of producing knowledge that can be utilized directly to improve student outcomes, given the nature of school improvement processes operating under different conditions. A simple example was studies that still reported that principals directly affect student outcomes. There is no practical meaning or application that can be attached to that finding, because there is no presumed process connecting the observed association between the effect

(i.e., be it principal experience, education, or quality) and the outcome. This challenge is to produce *usable* knowledge, which is tied to the need to understand more thoroughly how academic improvement actually unfolds over time at multiple levels within school districts (i.e., in schools, in classrooms, and among student peer and instructional groups). Of course, all research efforts to guide real-world action must acknowledge gaps between evidence, action, and its consequences (Deaton and Cartwright, 2016).

Features of longitudinal data

Edmonds (1979), in his early work on instructionally effective schools, was among the first to recommend longitudinal data in documenting school effectiveness. He noted future "... studies of school effectiveness should be multivariate in character and employ longitudinal records of pupil achievement in a variety of areas of school learning" (p. 22). Edmonds also recommended identifying school determinants of student learning in smaller-scale studies that described the nature of their school environments and associated academic processes, in contrast to the large-scale studies of the 1960s (e.g., Coleman et al., 1966) that used cross-sectional data and statistical controls to document the low academic outcomes of urban poor students ascribed to their family backgrounds. Edmonds documented several initial studies which identified strong principal leadership in setting the tone of the school, helping to decide on instructional strategies, and organizing schools resources, and the attitudes and behavior of teachers as differentiating improving versus declining schools (e.g., Brookover and Lezotte, 1977; Weber, 1971). Although the calls appeared in the literature, the challenges were great, both in obtaining longitudinal data on students and teaching processes, as well as having the technical means to deal with problems such as missing data that were prevalent in longitudinal data. Of course, during the next decade, the documented evidence of school effects (e.g., targeted professional development, teacher quality and engagement, district and principal support) on student learning began to accumulate such that the public, policymakers, educators, and researchers concluded that schools do make a difference (e.g., Berman and McLaughlin, 1977; Firestone and Corbett, 1988; Rutter et al., 1979; Scheerens and Creemers, 1989).

Despite the recognition that cross-sectional data were not well suited to investigations of processes that are assumed to be dynamic, the development of proper methods to examine change challenged researchers for many years. Time is a key factor in understanding how organizational processes unfold, as well as how their impacts may be observed. Temporal issues and the observation of changes are central to research on school leadership and improvement in school and classroom processes that lead to increased student learning. Strategies for school improvement develop, are implemented, and produce impacts which may not become apparent until a considerable amount of time has passed. A key consideration, therefore, in the design of longitudinal studies is how time will be incorporated and measured. Although time figures into the design of the study, change is often the object of the researcher's interest—change in student learning, school values and beliefs, academic processes, teacher and administrator behavior, and so forth. Newer approaches for examining longitudinal data offer greater flexibility in testing hypotheses concerning more complex relationships in the data where independent, mediating, and dependent variables may be changing over time, where changes in individuals and groups are examined simultaneously, and where separate growth trends can be compared before and after an intervention is introduced.

Although studies on school improvement routinely include longitudinal data on students, less attention has been given to monitoring changes in organizational processes at the school and classroom levels. As Griffin (1997) recognized, examining leadership effects also requires the need for longitudinal data that can describe changes in organizational processes over time. This data, however, can be difficult to obtain in ways that permit the examination of leadership effects across comparable organizational units (Hallinger and Heck, 2011). As Day et al. (2016) concluded, sustaining successful improvement effects results from schools' work in utilizing context-sensitive strategies that progressively shape and layer the school's work, culture, and achievement. We therefore must conceptualize school improvement as embedded within an organizational system that is changing over time and connect changes in the leadership and school organizational conditions to subsequent student academic growth in reading and math. Given that leadership is viewed as an adaptive process (e.g., Hallinger and Heck, 2011; Ogawa and Bossert, 1995), it is likely that these types of processes are mutually reinforcing in successful school improvement. Mutual influence implies that two (or more) variables, for example, leadership and school organization, may be both a cause and an effect of each other over time; that is, earlier temporal states of the component variables will mutually reinforce each other over a subsequent interval (time lag) or intervals (Hallinger and Heck, 2011; Heck and Reid, 2020).

Until recently, researchers did not have access to analytic tools capable of modeling more complex relationships involving successive measurements that may be mutually-reinforcing, changing at different organizational levels simultaneously, and that may affect the multiple environments (e.g., schools, classrooms) in which learning takes place. As Hallinger and Heck (2011) noted, aside from the need to study change through the measurement of component processes over time, the investigation of causal processes is also a function of selected research designs, where stronger designs (experiments and quasi-experiments) that manipulate treatments are better able to eliminate rival explanations of results and yield more defensible conclusions. School settings, however, are not readily amenable to experimental designs. Therefore, quasi-experimental research designs capable of examining temporal relationships over a number of occasions (i.e., at least two) may be useful where a particular type of intervention or strategic improvement action may be monitored before and after implementation, and possible changes in key school, classroom processes, and student outcomes observed (Goff and Finch, 2016; Heck and Chang, 2017). Degree of association between proposed variables, relative isolation from other extraneous variables, existence of a temporal relationship (which establishes the direction of influence), and replication of results enhance claims of causal relationships (Bollen, 1989; Campbell and Stanley, 1966).

Longitudinal designs in school leadership research

Building on calls for a more substantial knowledge base, during the past 15 years collective international research began to conceptualize school improvement as a process that unfolds over a number of years. These studies made use of longitudinal designs examining the impact of school leadership on strategic actions to improve academic and social processes within schools with an aim at improving student learning (e.g., Day et al., 2009; Hallinger and Heck, 2011; Heck, 2009; Heck and Chang, 2017; Heck and Hallinger, 2014; Leithwood and Louis, 2012; Mulford and Silins, 2003, 2011; Opdenakker and Van Damme, 2007; Thoonen et al., 2012). This work was facilitated by a trend whereby governmental bodies in various parts of the world began to develop substantial longitudinal datasets on school processes and outcomes (Hallinger and Heck, 2011). It also provided empirical support for the facilitative role that leadership plays in creating and sustaining a school-wide focus on learning for students, teachers and staff (Fullan, 2001; Leithwood et al., 2010; Robinson et al., 2008). Moreover, the empirical work also highlighted a capacity-building perspective especially supported by studies of transformational school leadership (Leithwood and Jantzi, 2000; Mulford and Silins, 2003).

Recent research found successful school improvement efforts targeted coordinating instructional practices with student needs and whole-school reform strategies where teachers felt the changes would benefit children, where they had raised expectations for student work and their own professional performance, and where there was increased instructional standardization, but teachers could make personal adaptations (Datnow and Castellano, 2000; Louis and Miles, 1991; Mintrop and Trijillo, 2007; Rowan and Miller, 2007). School leadership facilitated the selection, implementation, and evaluation of strategic actions that led to building the school's instructional capacity (e.g., teacher instructional skills, teaching strategies, quality of instructional experiences) and its relationship to increased student learning (Coelli and Green, 2012; Day et al., 2009; Feldhoff et al., 2014; Leithwood et al., 2004; Mulford and Silins, 2003, 2011; Murphy and Meyers, 2008; Robinson et al., 2008; Thoonen et al., 2012). This suggests viewing leadership actions as more collaborative, flexible, and contingent on schools' specific improvement needs. As scholars concluded, therefore, leadership not only impacted school actions to improve school conditions supporting academic and social organization, but it was also shaped by the context in which it was exercised (Bossert et al., 1982; Day et al., 2009; Jackson, 2000; Luyten et al., 2005; Mulford and Silins, 2011; Southworth, 2002). Moreover, external conditions such as increased accountability could result in re-defining values and routines, leadership responsibilities, interactions among individuals, and corresponding strategic actions (Stacey, 2001).

Longitudinal designs also open up the possibility of examining changing school processes (e.g., leadership, school conditions) as mutually reinforcing, which provides an alternative means of unpacking the issue of causal ordering in leadership and school improvement research, given that leadership is viewed an adaptive process (e.g., Hallinger and Heck, 2011; Ogawa and Bossert, 1995). Such processes imply that leadership and school organization, may be both a cause and an effect of each other over time (Hallinger and Heck, 2011; Heck and Reid, 2020; Marsh and Craven, 2006); that is, earlier temporal states of the component variables can mutually reinforce each other over a subsequent interval (or time "lag") or intervals. This approach has the advantage of facilitating a test of the hypothesis that initial school leadership influences subsequent school organization over time, as well as simultaneously responds to initial conditions in those organizational processes.

Two example longitudinal studies

In recent decades, longitudinal research designs have become the preferred approach for investigating individual and organizational development over time, and corresponding analytic approaches such as latent curve modeling and multilevel growth modeling have quickly become methods of choice for such research designs. In this section, I briefly discuss features and findings of two longitudinal studies among a set that examined leadership and school improvement over time (e.g., Bryk et al., 2010; Day et al., 2009; Hallinger and Heck, 2011; Heck and Hallinger, 2009; Mulford and Silins, 2003; Mulford and Silins, 2011; Opdenakker and Van Damme, 2007). Exemplars include the longitudinal examination of leadership, school practices, and school and community conditions leading to school improvement in reading and math over a seven-year period (Bryk et al., 2010) and the utilization of mixed methods in describing leadership's influence on improving school conditions in ways that lead to growth in academic outcomes over three years (Day et al., 2009).

Day et al. (2009) study

In one of the largest longitudinal studies examining improvement across some 360 elementary and secondary schools (Day et al., 2009), the research design focused on studying a subset of *improving* schools identified from an analysis of national data on primary and secondary school performance. The researchers utilized absolute student improvement and value-added contextualized improvement data to identify schools that significantly raised pupil attainment levels over a three-year period. Approximately a third of primary and secondary schools met the criteria for selection. From this set of improving schools, researchers identified three groups which made sustained improvements but from different starting points (i.e., low, moderate, and high) and contextual characteristics. Primary data collection included initial questionnaires sent to school heads and key staff (and second questionnaire a year later) combined with twenty detailed case studies, consisting of three visits per year over two years with detailed interviews.

The goal of the analyses was the identification of direct and indirect relationships between the work of effective school heads, changes in school and classroom processes and conditions, and improvements in student outcomes.

One of the primary findings was that school improvement evolved in three phases (early, middle, late), given differences in where schools initially started. Corresponding with these improvement phases, Day et al. (2009) identified four different patterns of successful school leadership, which they labeled (1) coming out of special measures (or sanctions) (2) taking ownership, (3) developing creativity, and (4) everyone a leader. Day and colleagues found certain selected strategies were prioritized across all phases, while others were given special emphasis in particular phases. For example, consistency of teaching practices and performance management of teachers was a primary feature of the early phase, and curriculum enrichment was a common emphasis in later phases. Another key finding was during the middle and later phases of school development, leaders tended to distribute leadership roles and responsibilities more widely, which led to the development of a culture of leadership trust. School heads reported most commonly focusing on improving teaching practices, followed by actions that promoted their school's academic press.

Among the strengths of the study for advancing the field's understanding of leadership and school improvement was the wealth of information collected describing the progressive distribution of leadership across phases and the use of different strategic actions in the development of school supportive conditions within each phase. Future considerations for such longitudinal studies include utilizing a comparison group of schools, perhaps by sampling across the whole spectrum of schools (i.e., including average and low growth schools) to isolate differences in school conditions related to stagnant or declining growth. Researchers might also lengthen the time in which improvement is studied beyond three years, which may not be a long enough time to illuminate fully changes in key theoretical relationships (Goff and Finch, 2016).

Bryk et al. (2010) study

In two longitudinal studies, researchers (Bryk et al., 2010; Sebring et al., 2006) examined nearly 400 neighborhood elementary schools in Chicago for gains in math, reading, and attendance. The studies utilized longitudinal data on students over a seven-year period coupled with parent, teacher, and student surveys that described changes in school processes. Their most detailed analyses focused on some 100 top-quartile public elementary schools in Chicago that improved substantially in reading and math over time contrasted with another set of 100 low-quartile schools with similar contextual factors that remained academically stagnant (Bryk et al., 2010). The researchers developed extensive information on school conditions from over 200 teacher survey questions, 70 student questions, and eight principal questions. The primary findings identified a comprehensive set of school practices and school and community conditions that differentiated high-improving and stagnant schools. In this *best* versus *worst* sample, schools demonstrating weakness on most of the core support indicators were several times more likely to stagnate than schools with strong overall organizational capacity scores. The researchers concluded substantial school improvement required building the social organization within schools and developing capacity across several domains (i.e., school leadership, professional capacity, student-centered learning climate, instructional guidance, and parent-community ties).

Among the strengths of this study for advancing the field's understanding of leadership was the repeated use of surveys to explain changes in school conditions. In their earlier study (Sebring et al., 2006), the researchers used surveys administered on one occasion, approximately in the midpoint of their study, to contextualize "improvement" in student scores. These earlier results were therefore more associational in nature; that is, measuring levels of supports at the midpoint of the study and associating different levels of the essential supports with improved longitudinal outcomes. Importantly, however, their later study provided more detailed information about teachers' work at their schools and utilized repeated surveys at two-year intervals to provide more time-sensitive measures of the school supports (Bryk et al., 2010). This allowed the researchers to confirm their primary findings over multiple replications, which provided a stronger test of time-dependent relationships. Importantly, this also facilitated examining *initial* conditions as related to subsequent changes in school academic improvement and, additionally, how *changes* in school conditions further impacted changes in school academic achievement (Bryk et al., 2010). This formulation was more consistent with theory in advancing a plausible explanation for observed changes in school supports that were related to subsequent changes in achievement. Moreover, replication of results renders alternative explanations of results less plausible. Future considerations for such longitudinal studies include sampling across the whole spectrum of schools, as comparisons between the best and worst maximizes variability that may not be present in other substantial subgroups of schools that could also be compared (Heck and Reid, 2020).

Some other useful designs

As the two previous examples illustrate, longitudinal data have been helpful in separating mediated-relationship studies from direct-effect studies. Specifically, claims of indirect effects of principal actions on outcomes can only be made in the presence of the temporal ordering of variables in a proposed model; for example, with mediators located between exogenous cause and the relevant outcome effects. This is the strongest feature of longitudinal data—the temporal ordering of variables in a proposed model—at least as related to longitudinal non-experimental designs (i.e., similar to the studies in the last section). Longitudinal data collection provides a means of examining how variables may change over time; however, it may not be able to resolve causal relationships entirely (Goff and Finch, 2016; Hallinger and Heck, 2011). In such designs, which are often referred to as *correlational* or *non-experimental* designs since they do not involve treatments, a researcher may observe an upward trend in student math scores over

time and attempt to discover school or classroom settings in which such scores were greater or less than the average trend. Although we may observe such settings with higher or lower math scores over time, we largely lack information regarding the underlying causal mechanisms that produced the observed differences due to a possible intervention, such as a change in day-to-day course delivery in some classrooms or schools that was monitored with respect to its changes in student assessments.

Dynamic theories of organizational processes have several common elements (Blalock, 1989; Kelly and McGrath, 1988). They should outline a broad plausible time parameter associated with the change being observed; they should describe the expected behavior of key variables over time and the interrelationships among variables as a series of equations; and the variables should have continuity (e.g., magnitude, rate of change, form a trend). Recently, Zyphur et al. (2019) introduced a general longitudinal cross-lagged panel model (GCLM), which can expand the conceptual and modeling options for examining relationships between variables changing over time in the absence of specific treatments introduced. GCLM provides expanded ways of thinking about relationships between variables such as leadership and school conditions by avoiding a single direction of causality and, instead, introducing concepts such as mutually-reinforcing (and reciprocal) relationships, as well as by including time lags that are consistent with causality (Hallinger and Heck, 2011; Zyphur et al., 2019). Importantly, the possibility of omitted variables that may change results observed in a particular model are a real possibility.

Quasi-experimental designs

Establishing causal relationships is, in large part, a function of the type of research design one may employ (e.g., experimental designs, quasi-experimental designs), as these latter studies incorporate some type of events (the introduction of a policy, a treatment to improve math proficiency for struggling students) assumed to induce an observable change in a series of longitudinal measurements. The ability to manipulate a type of intervention and to illuminate rival explanations are hallmarks of a strong research design. Time-series (i.e., with several measures before an event and several after) and regression discontinuity (i.e., where we can model the *process* used to assign individuals to groups before a treatment) designs provide viable options where we may use longitudinal data to examine the timing of events and subsequent changes. They can be useful where data are routinely collected over time (e.g., student test scores, school process data).

Quasi-experimental designs have the potential to uncover causal relationships between leadership, strategic interventions, and student improvement without randomized trials (Corcoran et al., 2012; Goff and Finch, 2016). They include the presence of an intervention but without random assignment to groups. The results are indicated by a discontinuity in the repeated measurements recorded at some point after the introduction of the intervention. Importantly, these designs provide greater separation of intended effects from rival explanations not typically afforded by simple longitudinal data analysis, provided certain threats to internal validity can be argued away (Cook and Campbell, 1979). The argument concerns the consideration of plausible rival explanations for any shift in the time series other than the intervention.

For example, Goff et al. (2014) used an interrupted time series to investigate the impact of a feedback and coaching intervention on principals' leadership behaviors (see also Miller's (2013) study on principal turnover). Heck and Moriyama (2010), utilized a regression discontinuity design to examine the temporal ordering of leadership, instructional practices, and added-year effects, net of the context and social composition. Heck and Chang (2017) utilized an interrupted time-series design to examine the timing of changes in key educational process indicators during NCLB in three groups of schools: a comparison group of schools that consistently met achievement targets; a group that entered mandatory restructuring due to prolonged academic failure but failed to exit; and a group that entered restructuring but managed to improve sufficiently to return to good academic standing. The design confirmed different trends in teachers' perceptions of educational processes corresponding to schools' restructuring status—a baseline trend as schools attempted to meet rising yearly performance targets, a trend of punctuated changes in educational processes after entering mandatory restructuring, and a return to relative stability in school support conditions in restructured schools that met yearly targets again and therefore exited restructuring status. The presence of a comparison group that was not restructured enhanced the design by isolating the evidence of change among *treated* groups, and the multiple years of entry and exit reduced the likelihood of incidental events affecting the observed time-series results (Cook and Campbell, 1979). Additionally, Heck and Chang noted the largest perceived change among exiting schools was in the effectiveness of the school's shared leadership and its focus on sustained strategic action for improvement.

Latent variable mixture models

Previous longitudinal studies on school leadership and school improvement suggest the influence of context in implementing changes to improve student learning. Identifying similarities in student achievement among underlying subsets of schools over a substantial period of time is an important step in understanding how different patterns of leadership and capacity building may be required for schools that are in different phases for academic improvement (e.g., Bryk et al., 2010; Day et al., 2009). Latent variable mixture modeling (LVMM) is a relatively under-utilized approach that relaxes the necessity of making "known-group" comparisons of schools' beginning achievement levels and corresponding growth rates. It facilitates identifying subgroups where membership is not known ahead of time but, rather, *emerges* from the data (Muthén, 2001). LVMM facilitates the identification of underlying latent classes, which are characterized by children or schools with similar beginning achievement levels and growth.

Where the latent classes are defined at the school level, it allows researchers to explore population heterogeneity that is caused by school-level variables. The purpose of this type of analysis is to estimate the number and size of the latent classes, assign membership to the classes, and then to identify key predictors of class membership as well as their effects on the differential class growth rates (Asparouhov and Muthén, 2014).

The primary advantage of this type of *emergent* approach for studying school improvement is that, rather than assuming one set of initial status achievement and growth estimates across a sample of schools, LVMM can identify underlying subsets of schools that share similar initial student achievement levels and growth rates within a broad distribution of schools. The approach therefore facilitates the examination of variables explaining population heterogeneity in achievement growth in more diverse samples of school growth compared to more narrow researcher-selected samples (e.g., low-vs. high-growth schools). Recent LVMM studies include Urick (2016) and Bowers et al. (2017), whose latent class analysis illustrates the type of rich contextual analysis of the relationships between teachers and leaders on a multisource survey of school leadership focused on improvement of student learning in schools.

Concluding thoughts

As longitudinal data become more readily available in upcoming years, we should observe their increased use in examining school improvement processes that result in advancing student learning. A key component, however, is the availability of more detailed information from school personnel, parents, and students on conditions in and around school communities that can help contextualize the nature of school-based changes. This provides necessary information regarding changes in processes that can be linked to subsequent changes in student outcomes. Armed with further longitudinal data and stronger research designs (e.g., quasi-experiments, cross-lagged analyses of mutual changes in variables, LVMM) researchers can contribute to improved validity of knowledge claims by reducing the plausibility of rival explanations of results.

References

- Anderson, G.L., 2004. William Foster's legacy: learning from the past and reconstructing the future. *Educ. Adm. Q.* 40 (2), 240–258.
- Asparouhov, T., Muthén, B., 2014. Auxiliary Variables in Mixture Modeling: 3-Step Approaches Using Mplus (Version 8). Muthén & Muthén, Los Angeles, CA. Retrieved from: <http://www.statmodel.com/examples/webnotes/webnote15.pdf>.
- Berman, P., McLaughlin, M.W., 1977. Federal Programs Supporting Educational Change; Volume 7. Factors Affecting Implementation and Continuation. Rand, Santa Monica, CA.
- Blalock Jr., H.M., 1989. The real and unrealized contributions of quantitative sociology. *Am. Socio. Rev.* 54, 447–460.
- Bolden, R., 2011. Distributed leadership in organizations: a review of theory and research. *Int. J. Manag. Rev.* 13, 251–269.
- Bollen, K.A., 1989. *Structural Equations with Latent Variables*. Wiley, New York, NY.
- Bossert, S., Dwyer, D., Rowan, B., Lee, G., 1982. The instructional management role of the principal. *Educ. Adm. Q.* 18, 34–64.
- Bowers, A.J., Blitz, M., Modeste, M., Salisbury, J., Halverson, R., 2017. Is there a typology of teacher and leader responders to CALL, and do they cluster in different types of schools? A two-level latent class analysis of CALL survey data. *Teach. Coll. Rec.* 119 (4), 1–66.
- Boyan, N., 1988. Describing and explaining administrative behavior. In: Boyan, N. (Ed.), *Handbook of Research in Educational Administration*. Longman, New York, pp. 77–98.
- Bridges, E., 1982. Research on the school administrator: the state-of-the-art, 1967–1980. *Educ. Adm. Q.* 18 (3), 12–33.
- Brookover, W.B., Lezotte, L.W., 1977. *Changes in School Characteristics Coincident with Changes in Student Achievement*. Michigan State University, College of Urban Development, Lansing.
- Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., Easton, J.Q., 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press, Chicago: IL.
- Camburn, E., Rowan, B., Taylor, J., 2003. Distributed leadership in schools: the case of elementary schools adopting comprehensive school reform models. *Educ. Eval. Pol. Anal.* 25 (4), 347–373.
- Campbell, D.T., Stanley, J.C., 1966. *Experimental and Quasi-experimental Designs for Research*. Houghton Mifflin, Boston, MA.
- Coelli, M., Green, D.A., 2012. Leadership effects: school principals and student outcomes. *Econ. Educ. Rev.* 31 (1), 92–109.
- Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F.D., York, R.L., 1966. *Equality of Educational Opportunity*. U.S. Office of Education, National Center for Educational Statistics, Washington, D. C.
- Cook, T., Campbell, D.T., 1979. *Quasi-experimentation. Design and Analysis Issues for Field Settings*. Houghton Mifflin, Boston, MA.
- Corcoran, S.P., Schwartz, A.E., Weinstein, M., 2012. Training your own: the impact of New York City's aspiring principals program on student achievement. *Educ. Eval. Pol. Anal.* 34 (2), 232–253.
- Creemers, B.P.M., Kyriakides, L., 2008. *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. Routledge, London.
- Datnow, A., Castellano, M., 2000. Teachers' responses to success for all: how beliefs, experiences, and adaptations shape implementation. *Am. Educ. Res. J.* 37, 775–799.
- Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., Brown, E., Ahtaridou, E., Kington, A., 2009. *The Impact of School Leadership on Pupil Outcomes. Final Report (DCSF Research Report RR108)*. Department for Children, Schools and Families (DCSF), Nottingham, London.
- Day, C., Sammons, P., Gu, Q., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Deaton, A., Cartwright, N., 2016. *Understanding and Misunderstanding Randomized Controlled Trials (Working Paper)*. Centre for Humanities Engaging Science and Society (CHESS), Durham, UK.
- Dillard, C., 1995. Leading with her life: an African American feminist (re)interpretation of leadership for an urban high school principal. *Educ. Adm. Q.* 31 (4), 539–563.
- Edmonds, R., 1979. Effective schools for the urban poor. *Educ. Leader* 37 (1), 15–18–20–24.
- Erickson, D.A., 1967. The school administrator. *Rev. Educ. Res.* 37 (4), 417–432.
- Everhart, R., 1988. Fieldwork methodology in educational administration. In: Boyan, N. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York, pp. 703–726.
- Feldhoff, T., Radisch, F., Klieme, E., 2014. Methods in longitudinal school improvement research: state of the art. *J. Educ. Adm.* 52 (5), 565–736.
- Firestone, W.A., Corbett, H.D., 1988. Planned organizational change. In: Boyan, N. (Ed.), *Handbook of Research on Educational Administration*. Longman, White Plains, NY, pp. 321–341.

- Fullan, M., 2001. *Leading in a Culture of Change*. Jossey-Bass, San Francisco.
- Goff, P., Finch, M., 2016. Challenges and opportunities for education leadership scholarship: a methodological critique. In: Bowers, A.J., Shoho, A.R., Barnett, B.G. (Eds.), *Challenges and Opportunities of Educational Leadership Research and Practice*. Information Age Publishing, Charlotte, NC, pp. 119–145.
- Goff, P.T., Guthrie, J.E., Goldring, E., Bickman, L., 2014. Changing principals' leadership through feedback and coaching. *J. Educ. Adm.* 52 (5), 682–704.
- Gold, A., Evans, J., Early, P., Halpin, D., Collabone, P., 2002. Principled Principals? Evidence From Ten Case Studies of Outstanding School Teachers'. Paper Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Griffin, M.A., 1997. Interaction between individuals and situations: using HLM procedures to estimate reciprocal relationships. *J. Manag.* 23 (6), 759–773.
- Gronn, P., Ribbins, P., 1996. Leaders in context: postpositivist approaches to understanding educational leadership. *Educ. Adm. Q.* 32 (3), 452–473.
- Gronn, P., 2002. Distributed leadership as a unit of analysis. *Leader. Q.* 13 (4), 423–451.
- Gustafsson, J.E., 2010. Longitudinal designs. In: Creemers, B.P.M., Kyriakides, L., Sammons, P. (Eds.), *Methodological Advances in Educational Effectiveness Research*. Routledge, Milton Park, pp. 77–101.
- Hallinger, P., Heck, R.H., 1996a. Reassessing the principal's role in school effectiveness: a review of empirical research, 1980–1995. *Educ. Adm. Q.* 32, 5–44.
- Hallinger, P., 2013. Reviewing reviews of research in educational leadership: an empirical assessment. *Educ. Adm. Q.* 50 (4), 539–576.
- Hallinger, P., Heck, R.H., 1996b. The principal's role in school effectiveness: an assessment of methodological progress, 1980–1995. In: Leithwood, K., Chapman, J., Corson, D., Hallinger, P., Hart, A. (Eds.), *International Handbook of Educational Leadership and Administration*. Kluwer, Dordrecht, pp. 723–783.
- Hallinger, P., Heck, R.H., 1998. Exploring the principal's contribution to school effectiveness: 1980–1995. *Sch. Eff. Sch. Improv.* 9, 157–191.
- Hallinger, P., Heck, R.H., 2010. Collaborative leadership and school improvement: understanding the impact on school capacity and student learning. *Sch. Leader. Manag.* 30 (2), 95–110.
- Hallinger, P., Heck, R.H., 2011. Conceptual and methodological issues in studying school leadership effects as a reciprocal process. *Sch. Effect. Sch. Improv.* 22 (2), 149–173.
- Halpin, A.W., 1957. A paradigm for research on educational administrator behavior. In: Campbell, R.F., Gregg, R. (Eds.), *Administrative Behavior in Education*. Harper & Row, New York, pp. 155–199.
- Harris, A., 2008. Distributed leadership: what do we know? *J. Educ. Adm.* 46, 172–188.
- Heck, R.H., 2009. Teacher effectiveness and student achievement: investigating a multilevel cross-classified model. *J. Educ. Adm.* 47 (2), 227–249.
- Heck, R.H., Chang, J., 2017. Examining the timing of educational changes among elementary schools after the implementation of NCLB. *Educ. Adm. Q.* 53 (4), 649–694.
- Heck, R.H., Hallinger, P., 1999. Next generation methods for the study of leadership and school improvement. In: Murphy, J., Louis, K. (Eds.), *Handbook of Research on Educational Administration*, second ed. Jossey Bass, San Francisco, pp. 141–162.
- Heck, R.H., Hallinger, P., 2005. The Study of educational leadership and management: where does the field stand today? *Educ. Manag. Adm. Leader* 33 (2), 229–244.
- Heck, R.H., Hallinger, P., 2009. Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *Am. Educ. Res. J.* 46, 626–658.
- Heck, R.H., Hallinger, P., 2014. Modeling the longitudinal effects of school leadership on teaching and learning. *J. Educ. Adm.* 52 (5), 653–681.
- Heck, R.H., Moriyama, K., 2010. Examining relationships among elementary schools' contexts leadership, instructional practices, and added-year outcomes: a regression discontinuity approach. *Sch. Effect. Sch. Improv.* 21 (4), 377–408.
- Heck, R.H., Reid, T., 2020. School leadership and school organization: investigating their effects on school improvement in reading and math. *Z. Erziehungswiss* 23 (5), 925–954.
- Hendriks, M.A.J., Scheerens, J., 2013. School leadership effects revisited: a review of empirical studies guided by indirect-effect models. *Sch. Leader. Manag.* 33 (4), 373–394.
- Jackson, D., 2000. The school improvement journey: perspectives on leadership. *Sch. Leader. Manag.* 20, 61–78.
- Kelly, J.R., McGrath, J.E., 1988. *On Time and Method*. Sage, Newbury Park, CA.
- Kruger, M., Witziers, B., Sleegers, P., 2007. The impact of school leader variables on school level factors: validation of a causal model. *Sch. Effect. Sch. Improv.* 18 (1), 1–20.
- Leithwood, K., 1994. Leadership for school restructuring. *Educ. Adm. Q.* 30 (4), 498–518.
- Leithwood, K., Duke, D., 1999. A century's quest to understand school leadership. In: Murphy, J., Seashore Louis, K. (Eds.), *Handbook of Research on Educational Administration*, second ed. Jossey-Bass, San Francisco, pp. 45–72.
- Leithwood, K., Jantzi, D., 2000. The effects of different sources of leadership on student engagement in school. In: Riley, K., Louis, K. (Eds.), *Leadership for Change and School Reform*. Routledge, London, pp. 50–66.
- Leithwood, K., Louis, K.S., 2012. *Linking Leadership to Learning*. Jossey-Bass, San Francisco.
- Leithwood, K., Montgomery, D., 1982. The role of the elementary principal in program improvement. *Rev. Educ. Res.* 52 (3), 309–339.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K., 2004. *How Leadership Influences Student Learning*. The Wallace Foundation, New York.
- Leithwood, K., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 28 (1), 27–42.
- Leithwood, K., Anderson, S., Maschall, B., Strauss, T., 2010. School leaders' influences on student learning: the four paths. In: Bush, T., Bell, L., Middlewood, D. (Eds.), *The Principles of Educational Leadership and Management*. Sage, London, pp. 13–30.
- Leithwood, K., Sun, J., Pollock, K., 2017. Introduction. In: Leithwood, K., Sun, J., Pollock, K. (Eds.), *How School Leaders Contribute to Student Success: The Four Paths Framework*. Springer International Publishing, Cham.
- Leithwood, K., Harris, A., Hopkins, D., 2020. Seven strong claims about successful school leadership revisited. *Sch. Leader. Manag.* 40 (1), 5–22.
- Louis, K.S., Miles, M.B., 1991. Managing reform: lessons for urban high schools. *Sch. Effect. Sch. Improv.* 2 (2), 75–96.
- Luyten, H., Visscher, A., Witziers, B., 2005. School effectiveness research: from a review of the criticism to recommendations for further development. *Sch. Effect. Sch. Improv.* 16 (3), 249–279.
- Marks, H.M., Printy, S.M., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39, 370–397.
- Marsh, H.W., Craven, R.G., 2006. Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1 (2), 133–163.
- Miller, A., 2013. Principal turnover and student achievement. *Econ. Educ. Rev.* 36, 60–72.
- Mintrop, H., Trijillo, T., 2007. The practical relevance of accountability systems for school improvement: a descriptive analysis of California schools. *Educ. Eval. Pol. Anal.* 29, 319–352.
- Mulford, B., Silins, H., 2003. Leadership for organizational learning and improved student outcomes—what do we know? *Camb. J. Educ.* 22, 175–195.
- Mulford, B., Silins, H., 2011. Revised models and conceptualization of successful school principalship for improved student outcomes. *Int. J. Educ. Manag.* 25 (1), 61–82.
- Murphy, J., 1988. Methodological, measurement, and conceptual problems in the study of instructional leadership. *Educ. Eval. Pol. Anal.* 10 (2), 117–139.
- Murphy, J., Meyers, C.V., 2008. *Turning Around Failing Schools: Leadership Lessons from the Organizational Sciences*. Corwin, Thousand Oaks, CA.
- Muthén, B., 2001. Latent variable mixture modeling. In: Marcoulides, G.A., Schumacker, R.E. (Eds.), *New Developments and Techniques in Structural Equation Modeling*. Lawrence Erlbaum, Mahwah, NJ, pp. 1–33.
- Ogawa, R.T., Bossert, S.T., 1995. Leadership as an organizational quality. *Educ. Adm. Q.* 31, 224–243.
- Opdenakker, M.C., Van Damme, J., 2007. Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *Br. Educ. Res. J.* 33, 179–206.
- Pitner, N., 1988. The study of administrator effects and effectiveness. In: Boyan, N. (Ed.), *Handbook of Research in Educational Administration*. Longman, New York, pp. 99–122.
- Pounder, D.G., Ogawa, R.T., Adams, E.A., 1995. Leadership as an organization-wide phenomena: its impact on school performance. *Educ. Adm. Q.* 31 (4), 564–588.
- Reynolds, D., Teddlie, C., Hopkins, D., Stringfield, S., 2000. Linking school effectiveness and school improvement. In: Teddlie, C., Reynolds, D. (Eds.), *The International Handbook of School Effectiveness Research*. Falmer Press, London, pp. 206–231.
- Ribbins, P., Gunter, H., 2002. Mapping leadership studies in education: towards a typology of knowledge domains. *Educ. Manag. Adm.* 30 (4), 359–386.

- Robinson, V.M.J., Lloyd, C.A., Rowe, K.J., 2008. The impact of leadership on student outcomes: an analysis of the differential effects of leadership types. *Educ. Adm. Q.* 44, 635–674.
- Rowan, B., Miller, R.J., 2007. Organizational strategies for promoting instructional change: implementation dynamics in schools working with comprehensive school reform providers. *Am. Educ. Res. J.* 44, 252–297.
- Rutter, M., Mortimer, P., Maugham, B., 1979. *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Harvard University Press, Cambridge, MA.
- Sammons, P., Luyten, H., 2009. Editorial article for special issue on alternative methods for assessing school effects and schooling effects. *Sch. Effect. Sch. Improv.* 20 (2), 133–144.
- Scheerens, J., Creemers, B.P.M., 1989. Conceptualizing school effectiveness. *Int. J. Educ. Res.* 13 (7), 691–706.
- Scheerens, J., 2012. *School Leadership Effects Revisited*. Springer, Dordrecht.
- Sebring, P.B., Allensworth, E., Bryk, A.S., Easton, J.Q., Luppescu, S., 2006. *The Essential Supports for School Improvement*. Research Report. Consortium on Chicago School Research. University of Chicago, Chicago, IL.
- Slegers, P.J.C., Thoonen, E.E.J., Oort, F.J., Peetsma, T.T.D., 2014. Changing classroom practices: the role of school-wide capacity for sustainable improvement. *J. Educ. Adm.* 52 (5), 617–652.
- Southworth, G., 2002. Instructional leadership in schools: reflections and empirical evidence. *Sch. Leader. Manag.* 22 (1), 73–91.
- Spillane, J.P., Halverson, R., Drummond, J.B., 2001. Investigating school leadership practice: a distributed perspective. *Educ. Res.* 30 (3), 23–28.
- Spillane, J.P., Healey, K., Parise, L.M., Kenney, A., 2011. A distributed perspective on learning leadership. In: Robertson, J., Timperley, H. (Eds.), *Leadership and Learning*. Sage, London, pp. 159–171.
- Spillane, J.P., 2006. *Distributed Leadership*. Jossey-Bass, San Francisco, CA.
- Stacey, R.D., 2001. Complex Responsive Process in Organizations: Learning and Knowledge Creation. Routledge, London, England.
- Tatsuoka, M., Silver, P., 1988. Quantitative research methods in educational administration. In: Boyan, N. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York, pp. 677–701.
- Thoonen, E., Slegers, P., Oorta, F., Peetsmaa, T., 2012. Building school-wide capacity for improvement: the role of leadership, school organizational conditions, and teacher factors. *Sch. Effect. Sch. Improv.* 23 (4), 441–460.
- Urick, A., 2016. The influence of typologies of school leaders on teacher retention: a multilevel latent class analysis. *J. Educ. Adm.* 54, 434–468.
- Weber, G., 1971. *Inner-City Children Can be Taught to Read: Four Successful Schools*. Council for Basic Education, Washington D.C.
- Willower, D.J., Forsyth, P.B., 1999. A brief history of scholarship in educational administration. In: Murphy, J., Seashore Louis, K. (Eds.), *Handbook of Research on Educational Administration*, second ed. Jossey-Bass, San Francisco, pp. 1–23.
- Witziers, B., Bosker, R., Kruger, M., 2003. Educational leadership and student achievement: the elusive search for an association. *Educ. Adm. Q.* 39, 398–425.
- Zyphur, M.J., Allison, P.D., Tay, L., Voelkle, M.C., Preacher, K.P., Zhang, Z., Hamaker, E.L., Shamsollahi, A., Pierides, D.C., Koval, P., Diener, E., 2019. From data to causes I: building a general cross-lagged panel model (GCLM). *Organ. Res. Methods* 1–37. <https://doi.org/10.1177/1094428119847278>.

Applying bibliometric review methods in education: rationale, definitions, analytical techniques, and illustrations

Philip Hallinger^{a,b} and Jasna Kovačević^c, ^a College of Management, Mahidol University, Bangkok, Thailand; ^b Department of Educational Leadership and Management, University of Johannesburg, Johannesburg, South Africa; and ^c Department of Management, School of Economics and Business, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

© 2023 Elsevier Ltd. All rights reserved.

Introduction	546
Rationale for bibliometric reviews	546
Identifying a document database for bibliometric review	547
Bibliometric analyses: explanations and illustrations	549
Descriptive bibliometric analyses	549
Science mapping	550
Co-citation analysis	551
Keyword co-occurrence analysis	553
Discussion	553
References	555

Introduction

Reviews of research have a long history in the field of education. Indeed, the education journal with the highest impact factor, *Review of Educational Research* (RER), is devoted exclusively to reviews of research. That said, up until the 1970s reviews of research in education, including those published in RER, tended to be of the *ad hoc* variety. This refers to reviews in which neither the procedures for selecting sources covered in the review nor the methods used to evaluate and synthesize them are made explicit and described in detail (Gough, 2007). However, since the 1970s and 1980s, authors of reviews of research in education have increasingly adopted review methods that fall within the rubric of systematic reviews of research (Foster and Hammersley, 1998; Hallinger, 2013, 2014; Mulrow et al., 1997).

Systematic reviews of research are conducted for a variety of purposes. Scoping and narrative reviews, often conducted when a research field or issue is relatively new, seek to highlight emerging trends and propose fruitful directions for future research (Foster and Hammersley, 1998; Peters et al., 2015; Sucharew and Macaluso, 2019). Critical reviews are undertaken once a body of knowledge has accumulated, with the aim of generating new theory (Paré et al., 2015). Meta-analytic reviews quantitatively integrate findings from a body of studies focused on a specific set of variable relationships with the goal of assessing evidence on impact and effectiveness (Cooper et al., 2019; Lipsey and Wilson, 2001). Bibliometric reviews analyze bibliographic data associated with a large set of documents in order to synthesize patterns of knowledge production related to a discipline or research issue (Zupic and Čater, 2015).

The emergence of these diverse approaches to research review reflects recognition of the critical role that reviews play in summarizing, critiquing, and synthesizing trends in the accumulation of knowledge (Hallinger, 2014; Mulrow et al., 1997). Indeed, Mulrow et al. (1997) referred to research reviews as, “critical links in the great chain of evidence” (p. 389). In recent years, reviews of research have taken on increased importance with the advent of “evidence-based practice” in education (Hallinger, 2014; Parkes et al., 2016), as well as other professional fields (Higgins et al., 2019). Indeed, it is for these reasons that research reviews often feature among the top-cited articles in education journals.

Among the commonly applied methods of systematic review, the bibliometric review is perhaps the least known among educational researchers. Although bibliometric reviews were conducted in education as early as the 1980s (e.g., Budd, 1988), it is only in the past several years that this mode of review has begun to be applied to educational issues with greater frequency (Aparicio et al., 2019; Hallinger and Kovačević, 2019a; Hallinger and Wang, 2020; Marín-Marín et al., 2019; Wang et al., 2017). Thus, we believe that education scholars may benefit from an explicit introduction that clarifies how to assess, interpret and execute bibliometric reviews of research.

In this article, the authors describe the rationale and method behind bibliometric reviews of research. Furthermore, we offer illustrations of bibliometric techniques from several recent bibliometric reviews of educational research. Our goal is to demonstrate the potential of this mode of research review for synthesizing patterns of knowledge production and accumulation in education. Given the increasing frequency of publication of bibliometric reviews in educational journals, readers and reviewers will benefit from a concise analysis of the rationale and methods used in this approach to systematic research review.

Rationale for bibliometric reviews

Scholars have been interested in understanding factors that shape knowledge production in education (Eidel and Kitchel, 1968; Hallinger, 2011; Metcalfe and Fenwick, 2009; Mertkan et al., 2017) as well as other disciplines since the 1960s (Gilbert, 1977; Ikpaahindi,

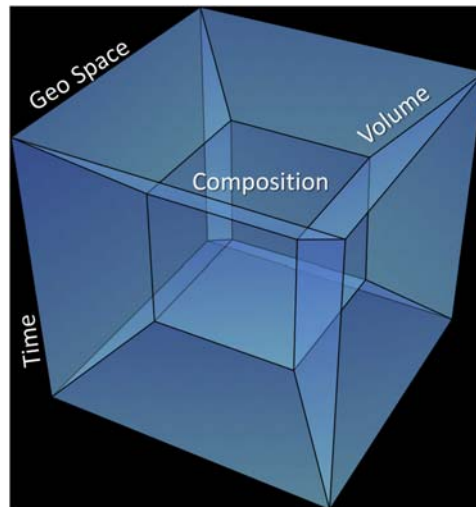


Fig. 1 Conceptual dimensions of a knowledge base.

1985; Kuhn, 2012; Merton, 1973; Small, 1973, 1999; De Solla Price, 1965). Reviews of research play a central role in this effort as they capture and make sense of knowledge accumulation within specific domains of knowledge. We assert that among existing research review methods, bibliometric reviews offer a unique approach to expanding our perspectives on the knowledge base.

The “knowledge base” in a field of inquiry can be conceptualized in terms of four dimensions (see Fig. 1). “Size” of the knowledge base refers to the volume of published research in the field or on a given topic. This takes note of the fact that the advancement of knowledge requires the accumulation of a critical mass of theory and empirical research. While the size of a knowledge base does not directly equate to knowledge advancement, it can be conceived as a necessary but insufficient condition for the development of a discipline or a line of inquiry with it.

“Time” is measured in terms of the publication trajectories that accrue on particular topics, or for the field as a whole, as the knowledge base evolves over time. Longitudinal analysis of the knowledge base enables scholars to connect current trends to past findings and codify important changes in our understanding of phenomena. This keeps scholars humble and highlights the relative breadth and depth of the knowledge base as it accumulates over time.

The “spatial” distribution of a literature can be analyzed in terms of the geographical origins of scholarship within a field. This addresses the question, “to what extent is scholarship in this field widely distributed across the world, or concentrated within a smaller number of societies?” Additional insights can be gained by analyzing the geographical distribution of a knowledge base in concert with its size and publication trajectories. This acknowledges the fact that knowledge is socially and culturally constructed, and highlights the critical goal of developing a knowledge base in education that is globally relevant and valid.

“Composition” refers to the “intellectual structure” of the knowledge base. Zupic and Čater (2015) defined intellectual structure as, “the examined scientific domain’s research traditions, their disciplinary composition, influential research topics, and the pattern of their interrelationships” (p. 435). Composition can be analyzed by reference to key documents, authors, and journals that come to be located within a knowledge base as it emerges and evolves.

Bibliometric reviews of research have the capability to gain insight into these dimensions of a knowledge base. As noted above, bibliometric reviews aim to synthesize patterns of knowledge accumulation and production through systematic means of document identification, analysis and synthesis (Hood and Wilson, 2001). As such, bibliometric reviews leverage advancements in software development (Pan et al., 2018; Van Eck and Waltman, 2009, 2017) such that reviewers are able to examine much larger document samples than are typically analyzed through other review methods.

Indeed, this capability to work with “big data” separates bibliometric reviews from other review methods. For example, in a recent bibliometric review of research on educational leadership and management, the authors used bibliometric tools to analyze 22,361 Scopus-indexed documents published over the course of 60 years (Hallinger and Kovačević, 2019b). Indeed, Waltman (March 12, 2020, Personal communication) suggested 200 documents as the “minimum document sample size” needed to reveal viable patterns for some analyses in a bibliometric review. Of course, reviews that encompass such large document samples must sacrifice depth of document examination for breadth (Hood and Wilson, 2001). This is the case with bibliometric reviews which are designed to reveal broad patterns of knowledge production and accumulation within a conceptual domain rather than evaluate specific findings.

Identifying a document database for bibliometric review

In a bibliometric review, the author begins by conceptualizing the topical focus. This is followed by a systematic search for relevant documents in a document repository such as the Web of Science, Scopus, Dimensions, or Clarivate. Google Scholar is less often used

for bibliometric reviews due to its limited capacity for exporting bibliographic data in a form that can be analyzed by bibliometric software packages.

We assert that selection of the “best” document repository is a field-dependent decision. For education and the social sciences, comparative empirical analyses have found that Scopus offers far more comprehensive coverage of “relevant journals” than the Web of Science (Falagas et al., 2008; Mongeon and Paul Hus, 2016). We learned this lesson from experience when conducting our first bibliometric review in educational leadership and management (EDLM). Whereas the Web of Science included only three of the “nine core EDLM research journals”, Scopus included all of them. Therefore, a review of EDLM research that relied on the Web of Science for the identification of the review database would fail to yield a valid picture of this field of study (see Hallinger, 2020b; Hallinger and Kovačević, 2019a).

As is with any systematic review, the reviewer must next decide which “document types” to include in the review. Relevant options offered by the major document repositories include journal articles, books, book chapters, conference papers and conference proceedings. There is no one right option in terms of which of these to include in a review. Indeed, highly-cited reviews can be identified that were limited to journals, as well as ones that included multiple document types.

Although many reviews limit the document sample to journals due to more consistent quality checks, we suggest that this is also a field dependent decision. In fields such as technology and engineering, conference papers undergo rigorous review and form an integral portion of the knowledge base. In other fields, including education, conference papers are often less central to the knowledge base. It should be noted that neither the Web of Science nor Scopus include broad coverage of books. The key point is that the reviewer should consider and explicitly justify the reasons for selecting the particular types of documents included in a bibliometric review.

The next step is to generate the “document list” in the selected document repository. Bibliometric reviews either define an explicit set of “search terms” to generate a document list, or select a delimited set of journals from which to source documents. At first glance, the comprehensiveness of the first approach would seem to make it the preferred choice. For example, one of the authors recently conducted a review of research on problem-based learning (PBL) for which the keyword-based search option was quite straightforward and yielded a comprehensive document list with relatively few ambiguities (Hallinger, 2020a).

However, for many topics, it is difficult to arrive at a set of search terms that yields a sufficiently coherent and accurate document list. In such instances, bibliometric reviewers may choose to identify a limited set of key journals for which it can be assumed that “all published documents” are relevant to the topic of the review. The authors used this approach in our review of research on educational leadership and management (EDLM), a topical domain that turned out to be exceedingly difficult to capture reliably via a keyword search (Hallinger, 2020b; Hallinger and Kovačević, 2019a).

In this case, the authors identified 22 Scopus-indexed journals that publish exclusively within the field of EDLM. Instead of using a keyword-based search, we simply exported from Scopus data associated with “all of the articles published by the 22 journals”. Although this approach resulted in the “omission” of many relevant EDLM documents published in general education and management journals, the authors were able to make a convincing argument that the 22,361 articles sourced from the 22 EDLM journals offered a sufficiently comprehensive and valid picture of the EDLM field. Note again that bibliometric reviews exploit “big data” in order to portray broad trends. Therefore, the omission of other EDLM papers that fell outside our database was determined to be non-significant.

If, however, a keyword-based search approach is adopted, the reviewer will use the conceptual definition of the review’s topical focus to guide the selection of a set of keywords. These will be input into the search engine using a combination of AND/OR/NOT operators in order to generate a document list. Other “boundaries” may also be set for the search (Hallinger, 2013). For example, if the review is conducted for a particular time frame, the researcher will enter the start and end dates into the search engine. Alternatively the review may be set up to be open-ended. If the review is delimited to a particular region or country, geographical search terms may also be defined.

We wish to reiterate that keyword-based searches vary widely in their efficiency. Authors cannot simply assume that the list generated by Scopus or the Web of Science is valid. It is likely in virtually all cases that the search engine has included irrelevant documents and omitted other relevant ones. Indeed, the authors have seen Scopus searches that generated 7000 plus documents reduced to 1200 after document inspection, filtering and elimination based on duplication and relevance.

Due to the critical nature of the document database, bibliometric reviews typically follow PRISMA guidelines for reporting the procedures used to identify the review database in systematic reviews (see Moher et al., 2009). This means that reviewers must keep a careful record of the search sequence and the results as documents are filtered and eliminated. This pertains directly to the “repliability” of the systematic review.

We reemphasize that due to the nature of a bibliometric review, special care must be taken to ensure that the document list generated by the search engine offers a “valid picture” of the knowledge base. A valid review database should meet several criteria. First, as indicated above, selection of the source document repository data must be justified on the criterion of “coverage of relevant sources”.

Second, the reviewer must ensure that the document list contains few to no “irrelevant documents”. For example, in the aforementioned bibliometric review of research on PBL (Hallinger, 2020a), the document list generated by the Scopus search included studies of a genetic compound abbreviated as “PBL”. All of these documents had to be deleted individually from the Scopus list after a careful inspection of titles and abstracts. Depending upon the “ambiguity” of the search terms, the number of documents to be eliminated will vary from relatively few to many.

Finally, concerns for validity should also prompt the reviewer to ask, “have I captured all of relevant documents contained in the document repository?” The frequency with which search engines omit potentially relevant documents is actually quite surprising. The researcher can search the data repository on the names of key authors in the field and compare the results with the initial list in order to determine if relevant papers are missing. If they are, a supplementary search may be undertaken using alternative search terms in order to address the shortcomings of the main search.

These observations suggest that keyword-based searches run the gamut from clear to ambiguous, depending upon the topic. Recently a colleague asked us to consult on a bibliometric review of bilingual education. Rather quickly it became apparent that this would be a difficult topic to capture through a keyword search. We advised that it would likely take several weeks to search and filter the results before arriving at a valid list in Scopus, if indeed that were possible.

Once the researcher has developed a solid document list, the “bibliographic data” describing the documents are exported from the repository into an intermediate file format (e.g., MS Excel or text file). As noted, the “data” analyzed in a bibliometric review do not consist of research findings. Rather they consist of “bibliographic data” that describe features of the selected research documents. These “data” typically include author(s) names, author(s) locations, author(s) affiliations, document title, document source, date of publication, citation data, co-citation dates, keywords, abstract, and funding sources.

Some descriptive analyses can be performed in Excel, but data will need to be uploaded into a bibliometric software package for the main bibliometric analyses. Currently popular software packages include VOSviewer, Citespace, BibExcel, HistCite, Pajek, Netdraw, and Publish or Perish. For the purposes of this article, the authors will use the VOSviewer (Van Eck and Waltman, 2009) program and the Scopus data repository to illustrate techniques of bibliometric analysis.

It should also be noted that the accurate execution of bibliometric analyses requires one additional step prior to data analysis, the “disambiguation of names and terms” (Strotmann and Zhao, 2012). Although space limitations preclude a detailed explication, suffice it to say that the author must take a series of steps to ensure that the bibliographic data do not contain “overlapping keywords” such as “student” and “students”, or author names such as Heck, R. and “Heck, R.H”. Otherwise the results of bibliometric analyses will be unreliable. These steps are described in detail elsewhere (Strotmann and Zhao, 2012; Van Eck and Waltman, 2013).

Bibliometric analyses: explanations and illustrations

Bibliometric analyses can be broadly classified into descriptive and science mapping analyses. Descriptive analyses focus on identifying basic trends in knowledge production. Science mapping applies network analysis techniques to illustrate “relational patterns” in the knowledge base.

Descriptive bibliometric analyses

When applied to large data sets, even simple descriptive analyses can offer useful insights with respect to the size, growth trajectory, and spatial distribution of a literature. For example, Hallinger (In press) used bibliometrics to review 14,130 Scopus-indexed documents related to problem-based learning. Simply finding a knowledge base of that size suggested that PBL has evolved from being an innovation in medical education initiated during the 1970s and 1980s to a mainstream form of learning. In order to gain further perspective on the magnitude of this knowledge base, the author ran simulated Scopus searches on “case teaching”, “cooperative learning”, “project-based learning”, “flipped classroom”, and “simulations and serious games”. Searches on these alternative forms of active learning yielded between 2000 and 4000 documents, thereby affirming the substantial scope of research and theory that has accumulated on PBL.

Analyzing the growth trajectories of published knowledge associated with a discipline, a field of study or a research topic represents another useful means of documenting the evolution of knowledge. For example, Hallinger and Kovačević (2019a,b) documented the growth of the EDLM knowledge base from its emergence as a field of study in the 1960s to the present. The authors found that the decade by decade growth trajectory of articles published in key Scopus-indexed journals rose continuously from 1544 articles in the 1960s to over 5000 articles between 2010 and 2019. Analysis of growth trajectories also provides a basis for predicting the future growth or decline of a field or a topic.

If our goal is to build a knowledge base in education that is both globally relevant and valid, then theories, policies and practices must be studied in diverse contexts (Hallinger, 2011, 2020b). Thus, geographical analysis of the distribution of articles in a knowledge base is highly relevant. The analysis of size, growth trajectories and geography can also be combined in order to document changes in the distribution of knowledge production over time. Hallinger (2020b) disaggregated the EDLM global knowledge base by regions of the world. Then he graphically depicted the longitudinal pattern of change in the proportion of the knowledge base generated in Anglo-American-European societies versus the “emerging regions” of Asia, Africa and Latin America (see Fig. 2). This analysis highlighted the unbalanced production of knowledge from the 1960s to the 2000s, with the global EDLM knowledge base dominated by Anglo-American-European studies (see Fig. 2). However, the graph shows a clear paradigm shift in the field starting in the late 2000s with more and more literature being produced in the emerging regions. This graph suggests that the center of gravity in the knowledge base has begun to shift with increasing contributions from scholars in the emerging regions of Asia, Africa and Latin America. Findings such as these have implications both for the selection of research topics and for journals which vary in their receptivity to “non-traditional” sources of knowledge (Hallinger, 2011, 2020b).

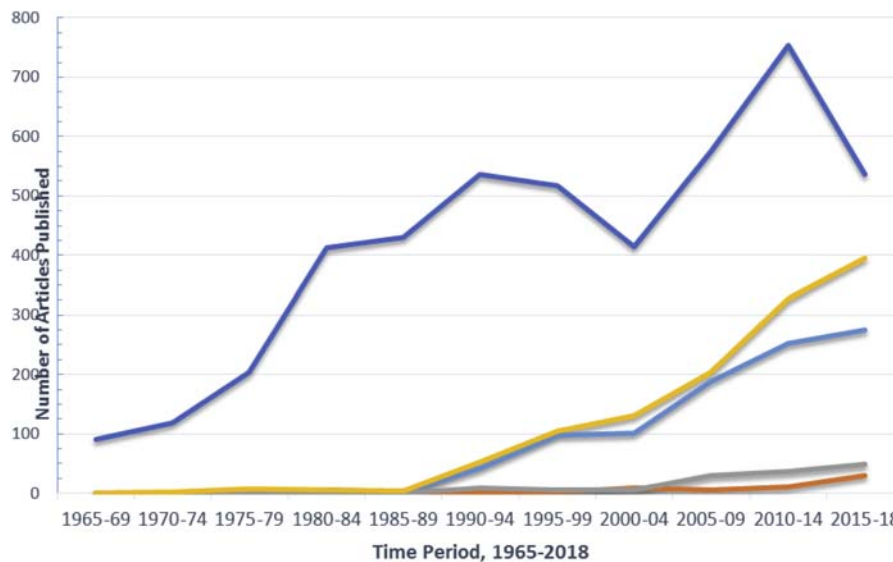


Fig. 2 Comparison of change in EDLM publication volume over time within and across different regions of the world, 1965–2018 ($n = 1171$ studies). Reproduced with permission from Hallinger (2020b), p. 2018.

We also include author and journal productivity as forms of descriptive data analysis. These productivity analyses measure the frequency of contributions of specific authors/journals to a literature. This offers insight into key contributors to a particular literature. Moreover, this type of metric becomes even more useful when combined with the results of author/journal citation and co-citation analyses.

Citation data related to authors, documents and journals are included in the Excel file exported from Scopus and Web of Science and can be analyzed either in Excel or with bibliometric software packages (e.g., VOSviewer). These citation data describe the number of times that an author/document/journal has been cited in documents included in the main document repository (i.e., Scopus, Web of Science) and are, therefore, referred to as “Scopus citations” or “Web of Science citations”. It should be noted that Scopus citations are usually larger than those generated by the Web of Science, but smaller than those yielded by Google Scholar due to differences in the size of the respective indexes.

Scholars have traditionally used citation analysis as a means of documenting the “impact” of prominent authors, documents, and journals within a domain of knowledge (Gilbert, 1977; Merton, 1973; Small, 1973). Merton asserted that citation metrics are used to, “prove the historical lineage of knowledge and to guide readers of new work to sources they may want to check or draw upon themselves” (1979, p. VI). Bibliometric scholars widely accept that frequently cited authors/documents/journals have made significant contributions to the advancement of knowledge (Hood and Wilson, 2001). This proposition rests on the assumption that authors cite other scholars or documents which contain ideas that contribute to their own work (Gilbert, 1977; Hood and Wilson, 2001; Small, 1973).

Citation analysis has been critiqued on various grounds (Weingart, 2005). Nonetheless, it remains a key metric used to gain insight into the “impact” of authors/documents/journals on knowledge accumulation (Zupic and Cater, 2015). However, we suggest that findings from citation impact analysis are best understood when assessed in concert with the results of other metrics such as author productivity, author and document co-citations (e.g., Hallinger and Kovačević, 2019a).

A variety of other descriptive analyses can be conducted focusing on patterns of funding sources, the frequency and distribution of universities that have contributed to a field of study, disaggregation of the knowledge base by subject area/discipline, and analysis of the distribution of types of documents (e.g., journal article, book, book article, conference paper). Space limitations preclude in-depth description of these analyses.

Science mapping

Advances in text mining and citation analysis have enabled bibliometrics to provide deeper and more comprehensive analyses than was possible in the past (Chen and Chen, 2003; Cobo et al., 2011; Van Eck and Waltman, 2017; Wang et al., 2017). Broadly referred to as “science mapping”, these approaches to data synthesis are capable of illuminating “relational features” of the knowledge base in a discipline (Cobo et al., 2011; De Solla Price, 1965; White, 2004; Zupic and Cater, 2015). These include co-citation analysis of authors/documents/journals, keyword co-occurrence analysis, bibliometric coupling, and co-author analysis (Ikpaahindi, 1985; Wallin, 2005; Van Eck and Waltman, 2014, 2017; Zupic and Cater, 2015). In this section of the article we define and illustrate several of the most common science mapping techniques (for information on other science mapping analyses see Zupic and Cater, 2015).

Co-citation analysis

In contrast with traditional citation analysis which scans the mother database" (e.g., Scopus or Web of Science), co-citation analysis assesses the frequency with which authors/documents/journals have been cited in the "reference lists of the review documents" themselves. This difference in the "source" of citations is fundamental to understanding one of the key difference between citation and co-citation analysis. Whereas citation analysis reveals "impact" on the broader literature, co-citation analysis reveals sources of "influence" on scholars working within a given field. Although there is always some degree of overlap between the results of citation and co-citation analysis, co-citation analysis typically reveals scholars/documents/journals from outside the field under review that have shaped the thinking of scholars working in the field under review.

For example, in Hallinger and Wang's (2020) bibliometric review of research on simulations and serious games, author co-citation analysis identified a group of highly cited cognitive and social psychologists (e.g., Bandura, Bransford, Brown, Wenger), none of whom had ever published an article on simulations and games. Thus, none of them appeared in the review database. Even so, their frequent citation in the reference lists of the review documents documented the influence of their learning theories on research into how and why simulations and games operated as modes of learning.

Similarly, in our bibliometric analyses of EDLM research (Hallinger and Kovačević, 2019b), we identified highly influential management and psychological theorists, none of whom had published articles in EDLM books or journals (e.g., Avolio, Bandura, Bass, Blau, Etzioni, Herzberg, Merton, Parsons, Simon). Yet, their theories of leadership and social organization provided the theoretical basis for understanding issues of leadership and management in schools.

While these examples have focused on "authors" as the unit of analysis, similar techniques and principles also apply to document and journal co-citation analysis. For example, the Hallinger and Kovačević (2019b) review identified Karl Weick's publication on schools as loosely coupled organizations published in the *Administrative Science Quarterly* as the "most highly co-cited article" in the field. This was despite the fact that the article had not been published in one of the "core EDLM journals" included in the review. Thus, we again emphasize that co-citation analysis offers insight into the "influence" of authors/documents/journals on scholarship in the knowledge domain under review, as opposed to the "impact" of scholars/documents/journals in the field on the broader literature.

Co-citation analysis also offers the capability to "visualize relationships" among authors/documents/journals cited within the knowledge domain under review. Small (1973) defined co-citation as the frequency with which two "units" (e.g., authors, documents, journals) are cited together in the reference lists of the review documents. In co-citation analysis the bibliometric software also calculates the frequency with which "pairs of authors/documents/journals" are cited together in the reference lists of the review documents (White and McCain, 1998; Zupic and Čater, 2015). For example, in each instance that the software finds a co-citation of two authors by a third author, each of the cited authors gains one "co-citation" and the pair accrues a "link".

For example, in Hallinger's (2020a) bibliometric review of PBL research, author co-citation analysis identified Howard Barrows and Henk Schmidt as "frequently co-cited authors" in this literature. This meant that documents included in the PBL review database which cited Barrows also frequently cited an article authored by Schmidt as well. Bibliometric scholars assume that frequently co-cited scholars/documents/journals bear an intellectual relationship or similarity (Small, 1973; McCain, 1986; Zupic and Čater, 2015). Such was the case with Barrows and Schmidt, each of whom was distinguished by his pioneering efforts at defining and designing PBL as a mode of learning.

Bibliometric software leverages this relational tendency by drawing on co-citation data to provide a visualization of similarities (VoS) among authors/documents/journals in a social network map. More specifically, co-citations and links are tracked by the bibliometric software, added into a matrix, and then transformed into a network map (Van Eck and Waltman, 2009, 2014). These visual maps of the literature allow us to gain insights into the relational structure of a knowledge base.

In order to illustrate these features of co-citation analysis, the authors ran author co-citation analysis for the Scopus-indexed EDLM literature published between 1960 and 2018 (see Fig. 3). The map in Fig. 3 is interpreted using several guidelines. First, the "size" of the author nodes reflects the relative number of co-citations accumulated by the author. On this map largest nodes belong to the three most frequently co-cited authors: Ken Leithwood (4091 co-citations), Philip Hallinger (2834), Alma Harris (2038). Based on the threshold used to create this map, the smallest nodes on the map indicate ~300 co-citations (e.g., Goddard, Printy, Silins).

The next interpretive guideline concerns the "proximity" of nodes to one another. Nodes located in close proximity bear a close co-citation relationship, and therefore, suggest a strong intellectual similarity between the relevant authors. For example, observe the proximity of Ken Leithwood and Philip Hallinger, Michael Fullan and Andrew Hargeaves, Charles Teddlie and Sam Stringfield, and Wayne Hoy and Megan Tschannen-Moran. Each of these pairs of scholars is noted for working along a similar theoretical line of inquiry. Their positions on the map reflect these close co-citation relationships empirically derived from the reference lists of the review documents.

The map also shows lines or "links" connecting pairs of authors, which as explained above, is based on their co-citation by other scholars. "Dense" links suggest a higher number of co-citations of the paired scholars. Examination of the pattern of links also offers insight into the location and breadth of influence of a scholar.

Finally, the map organizes the authors into colored "clusters" based on co-citation patterns. The author of a science mapping review must draw upon "tacit" knowledge of the field in order to identify the nature of the commonality in theoretical perspective among authors within a given cluster. Then the author will name the clusters as "schools of thought" or key theoretical streams that

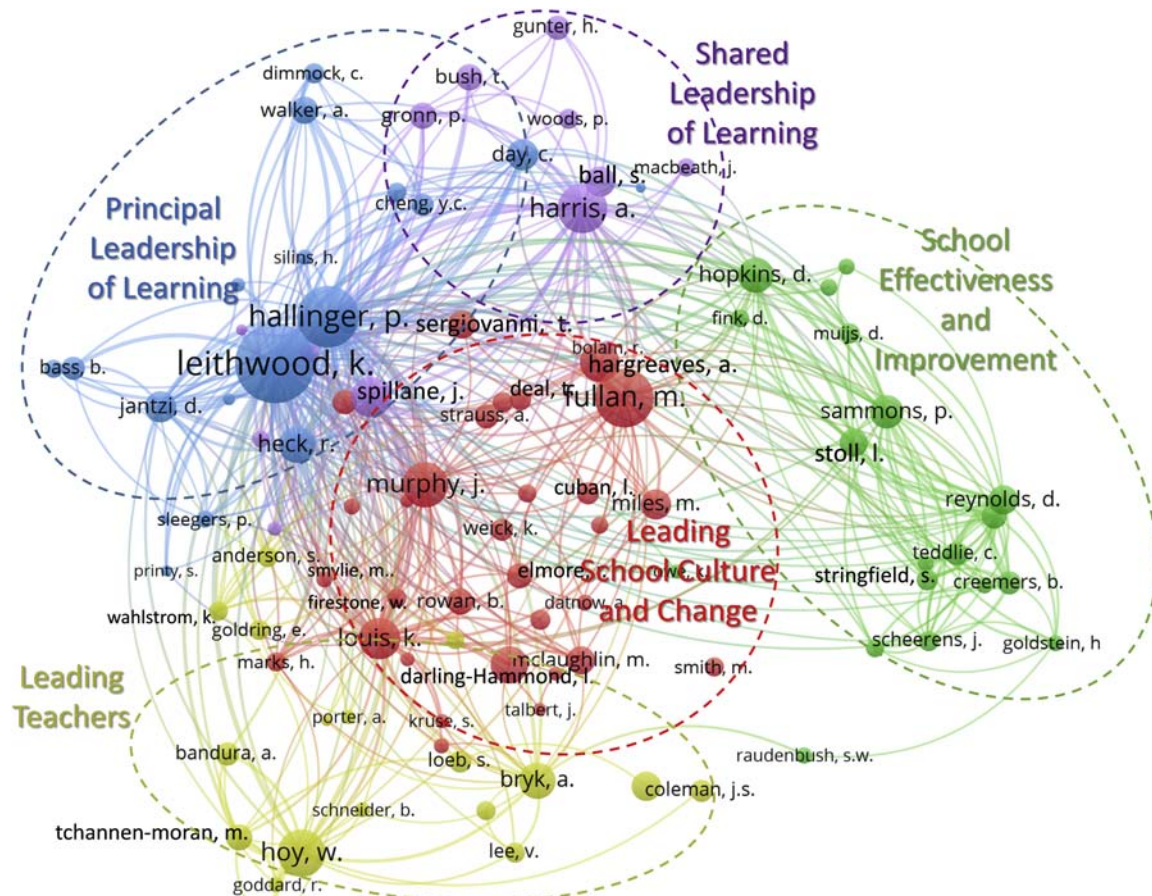


Fig. 3 Author co-citation map of educational leadership and management literature, 1960–2018 ($n = 22,361$ articles in 22 Scopus-indexed journals).

have emerged over time within the literature. This process is analogous to the process by which scholars interpret the output of exploratory factor analyses.

The map in [Fig. 3](#) highlights five schools of thought in the overall EDLM literature: Principal Leadership for Learning, Leading Teachers, Leading School Culture and Change, School Effectiveness and Improvement, Shared Leadership for Learning. These five schools capture the key theoretical directions that have evolved within this field of study during the period under review. Co-citation maps vary both in the number of clusters that emerge on a map, as well as in their degree of “coherence”.

Although the map in [Fig. 3](#) is reasonably coherent, there is some overlap between selected clusters (e.g., purple and blue, red and yellow). For example, notice that although James Spillane’s node is physically located in the Principal Leadership for Learning school, the software has assigned him, by color, to the Shared Leadership for Learning school. This reflects the fact that Spillane has published a critical mass of co-cited scholarship in both domains. Co-citation maps formed in this way have been found to provide valid conceptual depictions of disciplines ([Chen and Chen, 2003](#); [McCain, 1986](#); [Van Eck and Waltman, 2014](#)).

Synthesis of these varied features of a co-citation map can yield a range of different observations and conclusions about the nature of the knowledge base under review. For example, note that a scholar’s influence can be located primarily within his/her own cluster or school of thought (e.g., Creemers, Hanushek, Gunter) or distributed broadly across multiple schools (e.g., Leithwood, Fullan, Spillane, Murphy). Scholars with numerous links distributed across multiple clusters have been referred to as “boundary spanning scholars” whose theoretical work has shaped the field.

Because knowledge accumulation is a dynamic process, co-citation maps have also proven useful in charting evolution in the conceptual foci studied in a field over time. For example, the authors created a longitudinal series of author co-citation maps depicting the evolution of the EDLM literature over a period of six decades from the 1960s to the 2010s ([Hallinger and Kovačević, 2019b](#)). During the 1960s and 1970s, the author co-citation map revealed a lack of coherence, low author population, relatively low rates of co-citation, and few substantive themes related to the practice of educational administration. Instead, the named schools of thought tended to suggest a focus on administrative and organizational theory in education. At the same time, however, two small disconnected schools of thought had begun to emerge focusing on Organizational Change in Schools and Socio-economic Effects of Education. Subsequent maps revealed the growth of these two schools into more significant schools of thought associated with Leading School Culture and Change, and School Effectiveness and Improvement. Moreover, comparison with maps of subsequent

decades revealed a clear paradigm shift in the field occurring in the early 1980s with the advent of a new focus on leadership for teaching and learning (Hallinger and Kovačević, 2019b).

Hallinger (2020a) generated a series of longitudinal maps of the literature on problem-based learning for a similar purpose. These revealed significant stability in the foci of PBL scholarship over time with continuity of interest in PBL design and implementation as well as PBL effects on learners. At the same time, the maps also documented the evolution of an early school of thought associated with Cognitive Learning Theory into a broader domain concerned with Active Learning in Education (Hallinger, 2020a). This documented the increasing integration of other modes of active learning into PBL (e.g., flipped classroom, cooperative learning), as well as the increasing influence of PBL on other methods of learning.

The results of author co-citation analysis can also be interpreted in concert with other analyses. For example, tables that show the most productive scholars (i.e., frequency of contributions to the literature), citation impact, and co-citation influence offer multiple, complementary perspectives on author impact (see Tables 2 and 3 in Hallinger and Kovačević, 2019a, pp. 349–350). Taking these varied data into account, bibliometric scholars have identified “canonical scholars” and “canonical documents” which represent key authors/publications that have sustained high levels of impact and influence on a field across several decades (Hallinger and Kovačević, 2019a; White and McCain, 1998). Aside from recognizing key contributors to a literature, the identification of canonical scholars and documents offers a means of fast-tracking new entrants to key sources in a literature.

Keyword co-occurrence analysis

Keyword analysis has also emerged as another favored form of science mapping (Ding et al., 2001; Giannakos et al., 2020; Klavans & Boyack, 2017; Kovačević and Hallinger, 2020; Wang et al., 2017). This is a variant of text mining which uses keywords as a means of identifying topical trends within a literature. Notably, and in contrast to other bibliometric analyses, keyword analysis draws upon actual text located in the documents, which makes it a lower inference method of analysis. There are several approaches to keyword analysis that can be applied in a bibliometric review.

First, keyword analysis conducted in VOSviewer tracks the simple frequency with which keywords appear in the titles, abstracts and keyword lists of documents contained in the reviewer’s document list. Analysis of keyword frequency offers insight into the most popular topics studied within a literature.

Second, keyword co-occurrence analysis or “co-word analysis” visualizes relationships among keywords in a manner that is similar in principle to co-citation analysis. Van Eck and Waltman (2014) stated that, “The number of co-occurrences of two keywords is the number of publications in which both keywords occur together in the title, abstract, or keyword list” (p. 289). The analysis asks, “what topics are frequently studied together in this literature?” For example, in our EDLM review, the authors found that prior to 1980s “leadership” and “learning” seldom appeared together in the titles, abstracts, or keywords of the review documents. Yet during the post-1980s years, these two keywords “co-occurred” frequently.

Co-word analysis has been used to visualize the “conceptual space” of different literatures (Ding et al., 2001; Giannakos et al., 2020; Hallinger and Chatpinyakoo, 2019; Hallinger and Wang, 2020; Zupic and Čater, 2015). By way of example, the co-word map in Fig. 4 visualizes the literature on higher education for sustainable development (Hallinger and Chatpinyakoo, 2019). This map reveals three main conceptual themes within this literature: Curriculum and Instruction, Triple Bottom Line of Sustainability, Managing Universities Sustainability. The authors named these conceptual themes through a synthesis of tacit knowledge and keyword clustering.

A third approach to keyword analysis employs a ‘temporal overlay’ on the basic co-word map shown above in Fig. 4 (Van Eck and Waltman, 2014). Temporal co-word analysis links each keyword’s occurrence to the publication year of the associated documents (Van Eck and Waltman, 2014). VOSviewer software, for example, creates a distribution of publication dates for each keyword and then calculates the mean year for the distribution (Ludo Waltman personal communication, February 10, 2020). Distributions that center on a more recent year highlight topics of more current interest. A temporal co-word map differentiates the “relative recency” of keywords through color coding (see Fig. 5). As suggested above, this provides another lens through which to reveal longitudinal patterns in the topical evolution of a knowledge base.

Temporal co-word maps enable a more systematic documentation of changes in the topical foci of scholarship in a field over time (Hallinger and Wang, 2020; Van Eck and Waltman, 2014; Zupic and Čater, 2015). The temporal co-word map in Fig. 5 extends the previous figure by highlighting the time frame in which a key word was popularized. This particular review of research ran from 1998 to 2018. Therefore the dark shaded keywords reflect topics that emerged during the earlier years of the review and lighter shaded keywords over the past decade. Temporal co-word analysis has also been used to identify the “research front” or cutting edge topics of interest within a field of study or line of inquiry (Hallinger, 2020b; Hallinger and Chatpinyakoo, 2019; Hallinger and Wang, 2020; De Solla Price, 1965; Zupic and Čater, 2015).

Discussion

Bibliometric reviews of research have emerged as an increasingly popular complement to other forms of systematic review. Among extant review methods, we believe that bibliometric reviews are best suited to gaining an empirically-based, replicable, multi-faceted perspective on knowledge production within a field of study or a particular line of inquiry. Bibliometric reviews have the potential

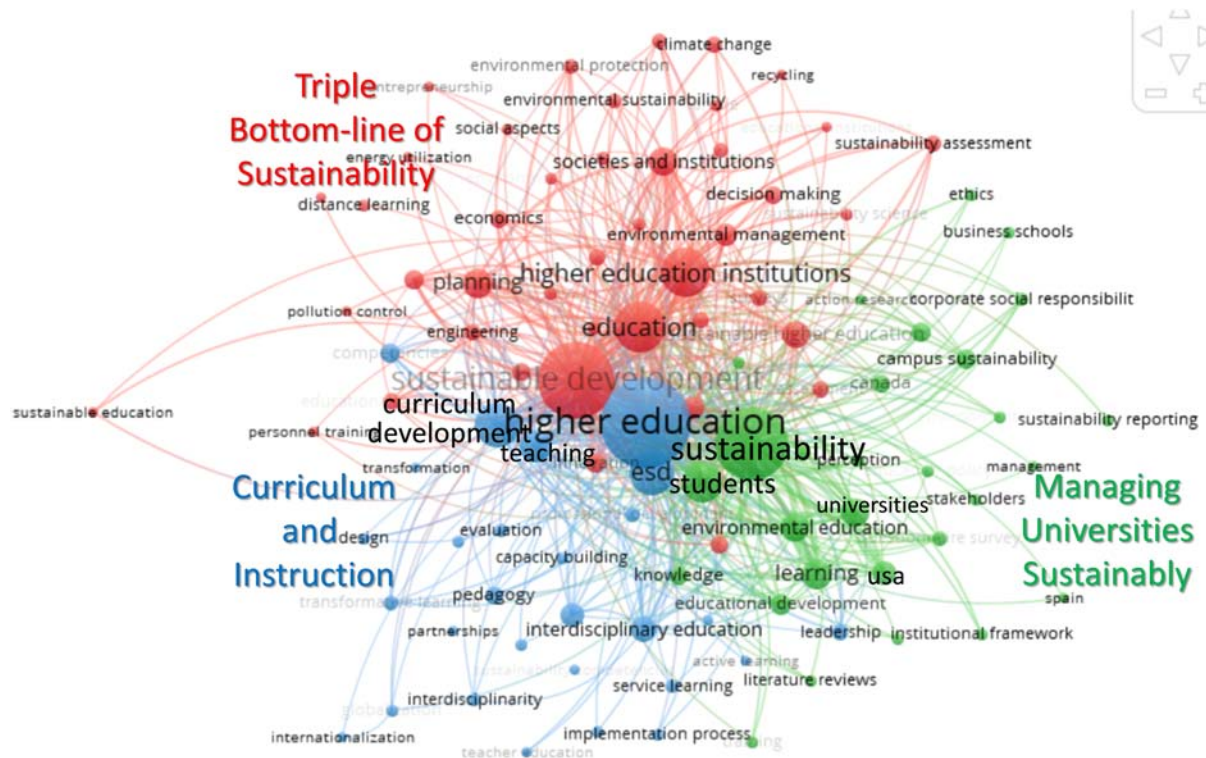


Fig. 4 Co-word map of literature on higher education for sustainable development, 1998–2018 ($n = 1459$ documents; threshold 9 co-occurrences).

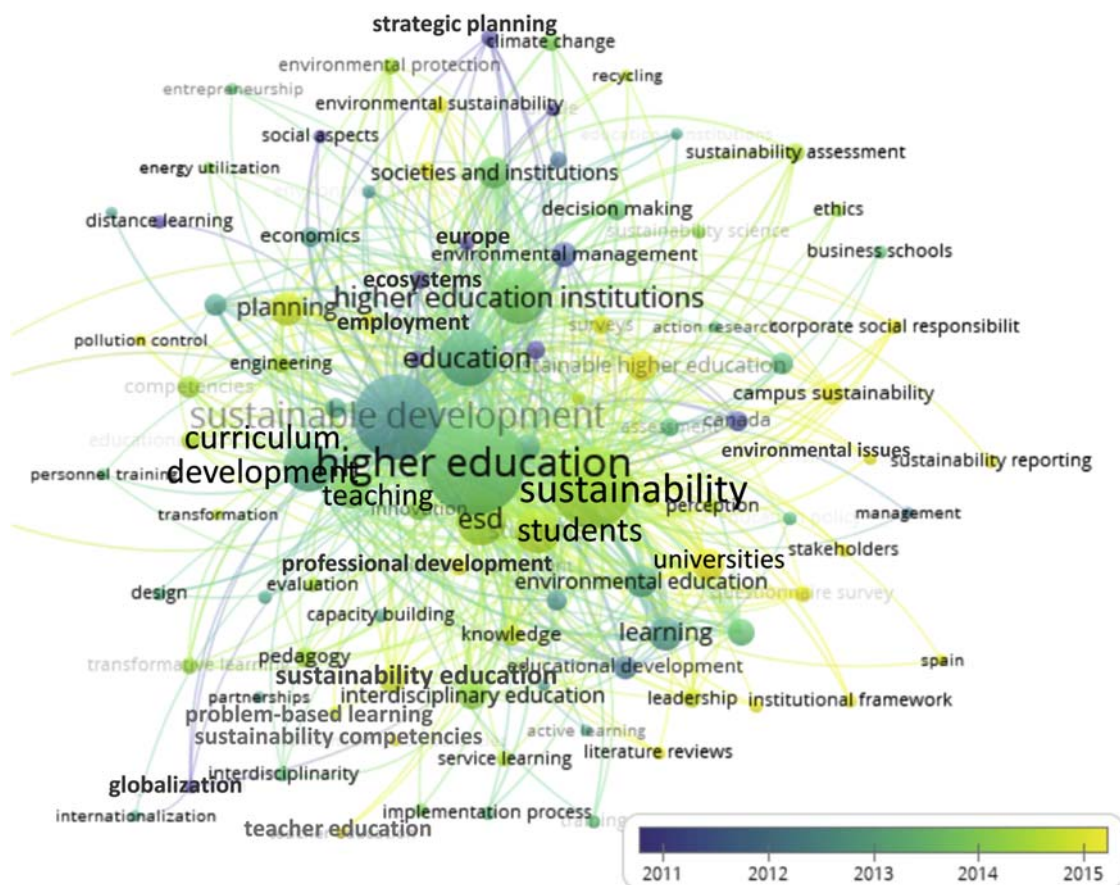


Fig. 5 Temporal overlay on a co-word map of literature on higher education for sustainable development, 1998–2018 ($n = 1459$ documents; threshold 9 co-occurrences).

to depict the theoretical evolution of a literature, identify paradigm shifts in a field of study, highlight key topics of current interest, reveal gaps in the production of knowledge, and point scholars toward the most important literature, scholars, and journals in a field of study.

It should be emphasized that the findings reported for the different bibliometric analyses often complement one another. For example, when the results of author productivity, citation and co-citation analysis are synthesized, they produce a multifaceted perspective on author influence and impact within a given literature. Similarly, combinations of analyses can be synthesized to gain insights into how scholarly capacity and contributions have been distributed geographically around the world and how they have changed over time (Hallinger, 2020b).

These capabilities of bibliometric reviews both complement and extend findings from knowledge base-oriented reviews that rely on narrative analysis, content analysis, and critical mapping. For example, the series of bibliometric reviews of the EDLM literature undertaken by the authors of this article (Hallinger, 2020b; Hallinger and Kovačević, 2019a,b) yielded perspectives and findings that both complemented and extended findings reported in prior studies of the EDLM knowledge base (e.g., Donmoyer et al., 1995; Gunter and Ribbins, 2003; Oplatka, 2010; Ogawa et al., 2000).

At the same time, as highlighted earlier, bibliometric reviews yield a “high inference” view of the field under review. They neither evaluate the quality of research through document inspection, nor are they concerned with the synthesis of specific findings from individual studies (Wallin, 2005). Thus, bibliometric reviews serve an explicit niche in the array of systematic review methods (Hood and Wilson, 2001).

We also wish to highlight three criticisms that have been lodged at bibliometric reviews. First, some authors of bibliometric reviews fail to take sufficient measures aimed at ensuring the validity of the database that forms the foundation of the review. The quantitative nature of this review approach can lead to a false sense that all is well, even when the underlying database fails to meet the desired standard. This is consistent with the “garbage in, garbage out” critique leveled at some meta-analytic reviews. The difference is that in this case the “garbage” does not lie in the quality of the studies included in the review, but rather in the validity of the database constructed for the review. This criticism reinforces the requirement that authors of bibliometric reviews offer a full, detailed explication and justification of the methods used to source the review database (Hallinger, 2013; Wallin, 2005).

Second, the authors have reviewed numerous bibliometric reviews of research that lacked sufficient depth in their conceptualization of the topic and/or interpretation of the analyses. Some authors simply pick a topic, draw a document sample, and execute the full array of analyses that can be run with the software package. The result is a series of tables and graphs that lack the benefit of context and meaningful interpretation. While this is not a limitation of the review methodology *per se*, it highlights a trap that scholars conducting bibliometric reviews can fall into all too easily. Thus, we reiterate that the analyses selected for use in a bibliometric review should be based upon a set of justified research questions, be limited in scope, and accompanied by copious thoughtful interpretation.

Finally, despite the seeming precision offered by this form of quantitative review, an author with limited knowledge of the literature will find it difficult to “make sense” of co-citation and co-word maps. At the same time, the authors have also found that conducting bibliometric reviews of a topical domain is an excellent learning experience, both for PhD students and scholars expanding their interests into a new domain. Here the “challenge” of interpreting a document or author co-citation map will lead the scholar to identify and engage specific literature that has been “validated” as valuable by other scholars. Similarly, journal citation and co-citation analyses offer specific guidance as to most receptive journals for the topic, as well as their ranking and rigor. Thus, we assert that the output of bibliometric reviews holds real value for the development of scholarship and gaining insights into the nature and evolution of the knowledge base.

References

- Aparicio, G., Iturralde, T., Maseda, A., 2019. Conceptual structure and perspectives on entrepreneurship education research: a bibliometric review. *Eur. Res. Manag. Bus. Econ.* 25 (3), 105–113.
- Budd, J.M., 1988. A bibliometric analysis of higher education literature. *Res. High. Educ.* 28 (2), 180–190.
- Chen, C., Chen, C., 2003. *Mapping Scientific Frontiers*. Springer-Verlag, London.
- Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., Herrera, F., 2011. Science mapping software tools: review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci. Technol.* 62 (7), 1382–1402.
- Cooper, H., Hedges, L.V., Valentine, J.C. (Eds.), 2019. *The Handbook of Research Synthesis and Meta-analysis*. Russell Sage Foundation, New York.
- De Solla Price, D.J., 1965. Networks of scientific papers. *Science* 150, 510–515.
- Ding, Y., Chowdhury, G.G., Foo, S., 2001. Bibliometric cartography of information retrieval research by using co-word analysis. *Inf. Process. Manag.* 37 (6), 817–842.
- Donmoyer, R., Imber, M., Scheurich, J.J. (Eds.), 1995. *The Knowledge Base in Educational Administration: Multiple Perspectives*. SUNY Press, Albany, NY.
- Eidel, T., Kitchel, J., 1968. *Knowledge Production and Utilization in Educational Administration*. ERIC, Eugene, OR.
- Falagas, M.E., Pitsouni, E.I., Malietzis, G.A., Pappas, G., 2008. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb. J.* 22 (2), 338–342.
- Foster, P., Hammersley, M., 1998. A review of reviews: structure and function in reviews of educational research. *Br. Educ. Res. J.* 24 (5), 609–628.
- Giannakos, M., Papamitsiou, Z., Markopoulos, P., Read, J., Hourcade, J.P., 2020. Mapping child–computer interaction research through co-word analysis. *Int. J. Child Comp. Interact.* 23–24, 100165.
- Gilbert, G.N., 1977. Referencing as persuasion. *Soc. Stud. Sci.* 7 (1), 113–122.
- Gough, D., 2007. Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Res. Pap. Educ.* 22 (2), 213–228.
- Gunter, H., Ribbins, P., 2003. The field of educational leadership: studying maps and mapping studies. *Br. J. Educ. Stud.* 51, 254–281.
- Hallinger, P., Chatpinyakoo, C., 2019. A bibliometric review of research on higher education for sustainable development, 1998–2018. *Sustainability* 11 (8), 2401.

- Hallinger, P., Kovačević, J., 2019a. A bibliometric review of research on educational administration: science mapping the literature, 1960–2018. *Rev. Educ. Res.* 89 (3), 335–369.
- Hallinger, P., Kovačević, J., 2019b. Science mapping the knowledge base in educational leadership and management: a longitudinal bibliometric analysis, 1960–2018. *Educ. Manag. Adm. Leader* 49 (1), 1741143219859002.
- Hallinger, P., Wang, R., 2020. The evolution of simulation-based learning across the disciplines, 1965–2018: a science map of the literature. *Simulat. Gaming* 51 (1), 9–32.
- Hallinger, P., 2011. Developing a knowledge base for educational leadership and management in East Asia. *Sch. Leader. Manag.* 31 (4), 305–320.
- Hallinger, P., 2013. A conceptual framework for systematic reviews of research in educational leadership and management. *J. Educ. Adm.* 51 (2), 126–149.
- Hallinger, P., 2014. Reviewing reviews of research in educational leadership: an empirical assessment. *Educ. Adm. Q.* 50 (4), 539–576.
- Hallinger, P., 2020a. Mapping continuity and change in the intellectual structure of the knowledge base on problem-based learning, 1974–2019: a systematic review. *Br. Educ. Res. J.* 46 (1).
- Hallinger, P., 2020b. Science mapping the knowledge base on educational leadership and management from the emerging regions of Asia, Africa and Latin America, 1965–2018. *Educ. Manag. Adm. Leader* 48 (2), 209–230.
- Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), 2019. *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons, New York.
- Hood, W., Wilson, C., 2001. The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics* 52 (2), 291–314.
- Ikpazhindi, L., 1985. An overview of bibliometrics: its measurements, laws and their applications. *Libri* 35, 163.
- Klavans, R., Boyack, K.W., 2017. Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *J. Assoc. Inf. Sci. Technol.* 68 (4), 984–998.
- Kovačević, J., Hallinger, P., 2020. Finding Europe's niche: science mapping the knowledge base on educational leadership and management in Europe, 1960–2018. *Sch. Effect. Sch. Improv.* 31 (3), 405–425.
- Kuhn, T.S., 2012. *The Structure of Scientific Revolutions*. University of Chicago press, Chicago.
- Lipsey, M.W., Wilson, D.B., 2001. *Practical Meta-analysis*. SAGE, Thousand Oaks, CA.
- Marín-Marín, J.A., López-Belmonte, J., Fernández-Campoy, J.M., Romero-Rodríguez, J.M., 2019. Big Data in education. A bibliometric review. *Soc. Sci.* 8 (8), 223.
- McCain, K.W., 1986. Cocited author mapping as a valid representation of intellectual structure. *J. Am. Soc. Inf. Sci.* 37 (3), 111–122.
- Mertkan, S., Arsan, N., Inal Cavlani, G., Onurkan Aliusta, G., 2017. Diversity and equality in academic publishing: the case of educational leadership. *Compare* 47 (1), 46–61.
- Merton, R.K., 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, Chicago.
- Metcalfe, A.S., Fenwick, T., 2009. Knowledge for whose society? Knowledge production, higher education, and federal policy in Canada. *High Educ.* 57 (2), 209–225.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151 (4), 264–269.
- Mongeon, P., Paul-Hus, A., 2016. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* 106 (1), 213–228.
- Mulrow, C.D., Cook, D.J., Davidoff, F., 1997. Systematic reviews: critical links in the great chain of evidence. *Ann. Intern. Med.* 126 (5), 389–391.
- Ogawa, R.T., Goldring, E.B., Conley, S., 2000. Organizing the field to improve research on educational administration. *Educ. Adm. Q.* 36, 340–357.
- Oplatka, I., 2010. *The Legacy of Educational Administration: A Historical Analysis of an Academic Field*. Peter Lang, London.
- Pan, X., Yan, E., Cui, M., Hua, W., 2018. Examining the usage, citation, and diffusion patterns of bibliometric mapping software: a comparative study of three tools. *J. Informetr.* 12 (2), 481–493.
- Paré, G., Trudel, M.C., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge: a typology of literature reviews. *Inf. Manag.* 52 (2), 183–199.
- Parkes, J., Heslop, J., Ross, F.J., Westerveld, R., Unterhalter, E., 2016. *A Rigorous Review of Global Research Evidence on Policy and Practice on School-Related Gender-Based Violence*. UNICEF, Paris.
- Peters, M.D., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D., Soares, C.B., 2015. Guidance for conducting systematic scoping reviews. *Int. J. Evid. Base. Healthc.* 13 (3), 141–146.
- Small, H., 1973. Co-citation in the scientific literature: a new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* 24 (4), 265–269.
- Small, H., 1999. Visualizing science by citation mapping. *J. Am. Soc. Inf. Sci.* 50 (9), 799–813.
- Strotmann, A., Zhao, D., 2012. Author name disambiguation: what difference does it make in author-based citation analysis? *J. Am. Soc. Inf. Sci. Technol.* 63 (9), 1820–1833.
- Sucharew, H., Macaluso, M., 2019. Methods for research evidence synthesis: the scoping review approach. *J. Hosp. Med.* 14 (7), 416–418.
- Van Eck, N.J., Waltman, L., 2009. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84 (2), 523–538.
- Van Eck, N.J., Waltman, L., 2013. *VOSviewer Manual*. Universteit Leiden, Leiden.
- Van Eck, N.J., Waltman, L., 2014. Visualizing bibliometric networks. In: Ding, Y., Rousseau, R., Wolfram, D. (Eds.), *Measuring Scholarly Impact: Methods and Practice*. Springer, Dordrecht, pp. 285–320.
- Van Eck, N.J., Waltman, L., 2017. Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics* 111 (2), 1053–1070.
- Wallin, J.A., 2005. Bibliometric methods: pitfalls and possibilities. *Basic Clin. Pharmacol. Toxicol.* 97 (5), 261–275.
- Wang, Y., Bowers, A.J., Fikis, D.J., 2017. Automated text data mining analysis of five decades of educational leadership research literature: probabilistic topic modeling of EAQ articles from 1965 to 2014. *Educ. Adm. Q.* 53 (2), 289–323.
- Weingart, P., 2005. Impact of bibliometrics upon the science system: inadvertent consequences? *Scientometrics* 62 (1), 117–131.
- White, H.D., McCain, K.W., 1998. Visualizing a discipline: an author co-citation analysis of information science, 1972–1995. *J. Am. Soc. Inf. Sci.* 49 (4), 327–355.
- White, H.D., 2004. Citation analysis and discourse analysis revisited. *Appl. Linguist.* 25 (1), 89–116.
- Zupic, I., Čater, T., 2015. Bibliometric methods in management and organization. *Organ. Res. Methods* 18 (3), 429–472.

FOURTH EDITION

International Encyclopedia of Education

Teachers' Lives, Work
and Professional Education

VOLUME
5

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editor

Diane Mayer



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 5

Teachers' Lives, Work, and Professional Education

VOLUME 5 EDITOR

Diane Mayer



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

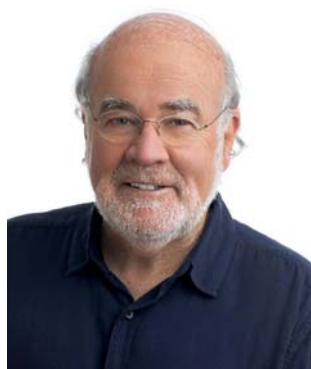
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



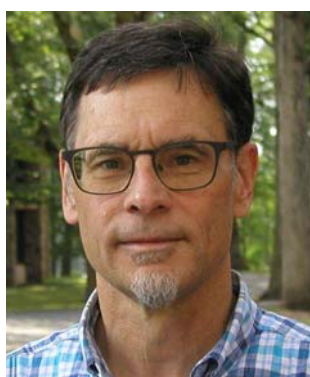
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 5: Teachers' Lives, Work, and Professional Education

Diane Mayer

© 2023 Elsevier Ltd. All rights reserved.

Introduction

Teachers' lives, work, and professional education are being impacted by ever-changing policies and reform agendas across many countries. Teacher quality, usually a perception that it needs to be improved, has become the major policy focus driving these policies and reforms. A spotlight on teacher quality is fueled by concerns about declining student test scores on international assessments and perceptions that other countries (where students have better test results) must be doing schooling better. In this context, teaching and teacher education are positioned as problems, ones that can only be fixed with new policies and reform agendas, with an accompanying sense of urgency. Various government commissioned reviews and reports from transnational organizations, usually highlighting what higher achieving countries are doing to attain their higher scores, form the basis of arguments for change. Even though claims are regularly made that these new policies are evidence-based and informed by rigorous research, at their core many are ideological in nature and political in intent (Helgetun and Menter, 2020) and based on an assumption that more complex and tighter systems of accountability will improve teacher education and teaching quality.

The chapters in this Volume interrogate teachers' lives, work, and professional education within this policy context. Together the chapters span global issues, trends, and possible futures. The authors often use their own research and local country contexts as exemplars and then make links to the international knowledge base in the field, foregrounding possible implications across countries. What emerges is a picture of connections and disconnections between and among research, policy, and practice. To frame this Volume, a brief introduction provides an overview of global policy trends in teacher education and teaching, discusses the ways in which these (re)construct professionalisms, and then highlights connections and disconnections between policy and research, many of which are explored further in the chapters that follow.

Policy Contexts of Teachers' Lives, Work, and Professional Education

Globalization and the related flows of people, knowledge, and practices have resulted in policies impacting teachers' lives, work, and professional education which look remarkably similar from country to country. An almost universal catalyst for new and emerging policies, particularly in high-income countries, is an economic imperative mobilized by a view that one's own country is, or will be, less competitive because of falling test scores on international student assessments like the Program for International Student Assessment (PISA). Countries look to others who appear to be doing better and borrow policies, or components of them. Such decontextualized policy borrowing involves "a piecemeal, 'pick n mix' approach that ignores the fact that educational policies and practices exist in ecological relationships with one another and in whole ecosystems of interrelated practices" (Chung, 2016, p. 207). However, "policies are not merely being transferred across time and space... [but] their form and their effect are transformed by these journeys" (Peck and Theodore, 2015, p. 29). In addition, global corporate actors like McKinsey and Company and transnational entities like the Organization for Economic Co-operation and Development have been regarded increasingly as authoritative voices and exert significant influence in policy development and implementation (Rautalin et al., 2019; Sellar and Lingard, 2013). Thus, the policy environments impacting teachers' lives, work, and professional education have become rather complicated, with many levers operating simultaneously.

To make these policy contexts even more complicated, individual governments aim to ground their policies and reform agendas locally, seeking country-specific justifications for change, while simultaneously looking to other countries and operating within various global influences. Thus, government reviews are announced with increasing regularity, usually involving hand-picked review panels and terms of reference aligned with political agendas. Most often, these review panels are tasked to: secure the input of various stakeholders; review policies and practices in other countries, usually those deemed to be high performing; review relevant research, often dominated by the reports from transnational entities and global actors; and make a set of workable recommendations which can be directly translated into policy and reform agendas post haste, within the existing political cycle. The review final reports regularly argue there is a problem needing to be fixed and use language imbued with a sense of urgency. A few key policy levers frequently form the core of the new policy, those that governments feel they can manipulate, and be seen to be doing so, in the name of improving teacher quality. This usually involves some version of increased regulation and tightened accountability, with the mechanisms for demonstrating accountability narrowly prescribed.

The resulting policies and reforms sometimes exhibit inherent dilemmas and contradictions, creating an entangled policy space full of competing discourses and binary debates. For example, increasingly complex systems of teacher certification and regulation for initial teacher education are being imposed in many countries, impacting on the work of existing teacher preparation programs. However, at the same time, parallel policies are promoting deregulation and opening up the teacher education market to alternative providers and pathways into teaching designed to reduce the role of universities and in large measure bypass the accountability requirements imposed on university-based teacher education programs. Another competing discourse is evident in relation to the desire for a highly qualified teaching profession. On the one hand, arguments for teachers to hold master's degree qualifications are being considered, usually citing the experiences of countries reported as high performing in international assessments. However, the realities of local teacher shortages mean that governments are implementing recruitment and employment policies that result in

increasing numbers of un-, under-, and differently qualified teachers entering the profession. The consequence is policy discourses around teacher education that have become increasingly dominated by binary and sometimes simplistic debates about whether the professional education of teachers should be in schools or in universities.

(Re)constructing Professionalisms

In these environments, professionalisms are constructed and reconstructed. For example, teacher and teacher educator professionalisms have become associated with increased levels of accountability and greater use of professional standards and measures of performance that link their work with national goals and economic agendas. Such organizational professionalism (Evetts, 2013) acts as a discourse of control by incorporating rational-legal forms of authority and hierarchical structures of responsibility and decision-making and is underpinned by a low level of trust in teachers and teacher educators. Governments set out to define the work of teachers and teacher educators (in standards) and establish accountability mechanisms by which evidence of (increased) performance is to be provided. "New professionalisms" are evident; "core" professionalisms are framed by professional standards; and "branded" professionalisms are emerging linked to new and "innovative" approaches to schooling and professional education (Whitty, 2014). The consequence can be a risk averse profession exhibiting compliant professionalism (Sachs, 2016). On the other hand, occupational professionalism is framed according to collegial authority, trust, autonomy, professional judgment, and guided by codes of professional ethics (Evetts, 2013). Similarly, Sachs' (2003) notion of an activist teaching profession incorporates democratic professionalism focusing on collegial relations and collaborative work practices. These ways of thinking of professionalism are usually favored by teachers rather than governments.

Thus, a policy context with its focus on increased accountability, standards, and performance exhibits organizational or managerial professionalism—both for teachers and teacher educators. What is the role of research?

The Role of Research: Connections and Disconnections

For some time, research on teaching has been judged through a "what works" lens and its value for teachers judged according to the use or not of prescribed methodologies (Burns and Schuller, 2007). These so-called evidence-based instructional practices are often referred to in policy documents and incorporated into the accountability mechanisms emanating from these policies. The aim is to ensure that teachers are indeed using the evidence-based practices and that teacher educators are preparing beginning teachers who know and are able to demonstrate these practices, usually as part of teaching performance assessments and observations of their emerging teaching practice. More recently, arguments have emerged for rethinking professional identity around practices that are informed and improved by and through teacher and classroom research, requiring teachers to be research literate to judge the value of publicly available research for their teaching, and also to be researchers themselves in order to investigate and improve their classroom practices. Both of these approaches frame a professionalism that involves informed professional judgment and teaching decisions designed to enhance student learning. For example, in the United Kingdom, the BERA-RSA report envisages a repositioning of teacher professionalism where "a new environment of self-improving education systems teachers will need to become research literate and have opportunities for research and inquiry. This requires that schools and colleges become research-rich environments in which to work" (British Educational Research Association, 2014, p. 5). It remains to be seen whether there will be connections between this type of research and policy.

Meanwhile, teacher education research, as distinct from research on teaching, is a relatively new field of research. While it has evolved and grown over the past decades, evaluations of this body work have concluded that it is underdeveloped, small-scale, undertheorized, fragmentary, and somewhat parochial (e.g., Menter et al., 2010; Sleeter, 2014). While the small-scale studies that dominate this work do inform local teacher education practice in helpful ways, they do not produce the data sets and findings that policymakers generally appear to be seeking. The prevailing view is that this body of work has not systematically built a knowledge base for teacher education policy (Sleeter, 2014). However, the fact that teacher education programs are constantly changing and adapting means that researching such a dynamic system is difficult when using what might be considered more traditional methods of research and analysis (Cochran-Smith et al., 2014; Gray and Colucci-Gray, 2010). What is more, there are suggestions that the current body of teacher education research has been distorted and misused (Zeichner and Conklin, 2016) in policy contexts to manufacture a narrative of failure and a rationale for tighter accountability and significant reform in teacher education (e.g., Cochran-Smith et al., 2018).

Clearly, the connections and disconnections between and among research, policy, and practice are messy. The chapters in this Volume aim to illuminate the field further.

The Contribution of this Volume

The chapters explore global trends in teaching and teacher education policy exposing the connections and disconnections between policy and research, the ways in which professionalisms are (re)constructed, and the impact on teacher education and school teaching practice. Seventy-seven chapters are included, written by authors based in more than 25 countries. The authors have often foregrounded their own and others' research in specific country contexts, then positioned this within the recent international research literature, and finally raised globally relevant implications for future policy, practice, and research in/on/about/for teachers' lives, work, and professional education. The chapters are organized in the following seven sections:

- Section 1: The Policy and Political Contexts of Teachers' Work and Professional Education
- Section 2: Teachers' Lives and Work
- Section 3: Knowledge for Teaching
- Section 4: Initial Teacher Education
- Section 5: Those Who Educate Teachers: Teacher Educators
- Section 6: Ongoing Teacher Professional Learning
- Section 7: Future and Emerging Issues: Research, Policy, and Practice

In section 1, the chapters examine the policy and political contexts of teachers' lives, work, and professional education in the neo-liberalized context of the 21st century, the impact of the Global Educational Reform Movement, the high stakes accountability environments and pervasive performative policy conditions, the politics behind definitions of quality, and also the new public management and accountability policies impacting teachers' work and teacher education. Further, the impact of teacher education policy on the professionalism of both teachers and teacher educators, including those in the early childhood sector, is examined, as are the consequences for partnerships in teacher education which have been shown to be central to teacher learning. The governance of teacher education is examined highlighting the hierarchies, networks, and markets at work and also the roles and influences of non-state actors. Challenges are highlighted in supporting teachers in low-income countries to provide quality education for all learners, especially students from disadvantaged and marginalized groups.

The chapters in section 2 focus on teachers' lives and work. Beginning with early career teachers and their teaching, issues around retention, their emerging and ongoing professional identities, and their resilience are canvassed. Further, the chapters explore support and ongoing learning and development for teaching, including coaching and teacher inquiry. Issues around job satisfaction and affirming trans and queer teachers' lives are also examined.

In section 3, the chapters provide a lens into knowledge for teaching, highlighting some of the tenants of effective teaching and inclusive practices for working with diverse groups of students. A number of aspects of knowledge for teaching are examined and discussed, including epistemic practices of science, technology integration and using mobile technologies, the role of the Arts, teacher noticing, and reconciliation pedagogies.

The chapters included in section 4 contribute to our understanding of initial teacher education. First, issues of recruitment and selection are examined, and then initial teacher education approaches are probed including partnerships, community engagement, school-based experience, case methods, anti-racist teacher education, building agency and identity, and working with diverse learners. Ways by which the effectiveness of teacher education is being judged are then discussed, including teaching performance assessments and determination of its long-term effects. A group of chapters focus on the lives of teacher educators, their use of self-study for improving practice, and the ways in which they might be prepared for their teacher education role and their ongoing professional learning. Additionally, issues related to their role in ensuring equity and social justice, as well as their growth as researchers, are analyzed.

The chapters in section 5 review various approaches to ongoing professional learning for teachers including early childhood teachers. These chapters consider situated and workplace learning, the role of middle leading, lesson study, and the impact of formative assessment. The importance of coaching and of content focused professional development is argued.

The final section of this Volume includes chapters which focus on issues related to teachers' lives, work, and professional education worthy of (re)consideration into the future—the policies that impact, the research to build our collective knowledge, and the practices likely to be relevant and effective. Global and local flows and influences in education policy are analyzed, as are considerations about the impact of the pandemic and new digital learning spaces for students as well as teachers. The importance of building teachers' research capacity, their agency, their capacity to work with increasingly mobile and diverse populations, and mentoring in new spaces are all examined. Chapters highlight the promise of teacher education for advancing social justice and of comparative teacher education research in raising possible future scenarios and highlighting the power of human agency.

Together the chapters in this Volume provide a comprehensive overview of the field and highlight issues and global trends in research on, for, and about teaching and teacher education, exposing the connections and disconnections with policy, and the impact of policy on teacher and teacher educator professionalisms and practices. Collectively, they raise internationally relevant implications for future policy, practice, and research in/on/about/for teachers' lives, work, and professional education.

Order of CHAPTERS

The Policy and Political Contexts of Teachers' Work and Professional Education

4001. The Role and Influence of Non-State Actors on Teacher's Lives and Work
Moirra Hulme and Emilee Rauschenberger

4092. Teacher Education Policy and Professionalism: A Personal Review of Teacher Education Policy Research
Eline Vanassche

4012. The 'Problem' of Teachers' Work and Learning under Performative Policy Conditions
Ian Hardy

4015. Teachers' Work Amid Global Education Reform: 'The Greatest Challenge'?

Nicole Mockler and Meghan Stacey

4007. Defiant Imaginaries: From a Politics of Identity to an Ethics of Friendship in Teaching

Anne M. Phelan, Matthew Clarke, and Lily Vincent

4004. Global Politics of Quality in Education

Jaakko Kauko

4112. Revisiting Teacher Evaluation a Decade after Reforms

Joshua Bleiberg

4008. Teacher Policy Frames, Issues, and Enactment Forms

Beatrice Ávalos and Beatriz Fernandez

4003. Hierarchies, Networks, and Markets in Initial Teacher Education

Monica E. Mincu and Min Liu

4054. Can Initial Teacher Education Partnerships Maintain their Epistemological and Pedagogical Underpinnings in Light of Market-Driven Reforms? The Case of England.

Trevor Mutton, Katherine Burn, and Hazel Hagger

4098. Equity in early childhood education: Policy framing, Translation, and Enactment

Andrea Nolan and Tebeje Molla

4065. Teachers and Teaching in Low-Income Countries

Daniel Wagner and Amy Guillotte

Teachers' Lives and Work

4019. Narrative and Biography in Teacher Work Lives and Development

Geert Kelchtermans

4075. Learning to Count What Counts: Narrative Hermeneutics for Teachers and Teaching

Dennis Sumara, Sharon Friesen, Janice Doucette, Catherine Takasugi, and Jennifer Turner

4018. Researching Teacher Learning Through Self-Video

Jenni Ingram and Alf Coles

4025. Retention of Teachers from Minority Ethnic Groups in Disadvantaged Schools

Antonina Tereshchenko and Martin Mills

4027. Early Career Teachers' Work

Anna Sullivan, Bruce Johnson, and Barry Down

4011. Coaching for Professional Learning of Early Career Teachers

Pasi Sahlberg and Andrea Stringer

4087. Early Career Teaching and Resilience

Anna Lise Gordon

4085. Beginning teachers' professional identity work and its support

Gonny Schellings and Douwe Beijaard

4057. Promoting the Development of Teacher Expertise Throughout the Career

Jan van Tartwijk, Esther E. van Dijk, Johan Geertsema, Manon Kluijtmans, and Marieke van der Schaaf

4014. Teachers' Willingness and Ability to Teach to Their Best and Well: The Person in the Professional

Christopher Day

4028. Thriving the Challenges: Understanding Teacher Resilience in Chinese Hard-to-Staff Rural Schools

Qiong Li, Xiaohong Sun, and Jingtian Zhou

4033. Affirming Trans and Queer Teachers' Lives: Rethinking Teacher Professionalization and Pedagogy

Jennifer Ingrey

4082. The Role of Job Demands and Resources for Teacher Job Satisfaction: Insights from Five European countries
Anna Toropova, Stefan Johansson, Jeffrey Casely-Hayford, Lydia Kwak and Khaleda Gani-Dutt

4026. Teacher Inquiry: By Any Other Name
Anthony Clarke

Knowledge for Teaching

4035. Enacting Inclusive Pedagogy in Teacher Education: Creating a Learning Space for Teachers to Develop their Professional Identity
Hafðís Guðjónsdóttir and Edda Óskarsdóttir

4036. Learning to Teach Epistemic Practices of Science in Secondary Schools
Sibel Erduran and Liam Guilfoyle

4037. Teacher Skills and Knowledge for Technology Integration
Leming Liang and Nancy Law

4020. The Relationship Between Knowledge Exchange and Teachers' Use of Mobile Technologies in the Mathematics Classroom
Kinga Petrovai, Gabriel Stylianides, and Niall Winters

4010. Teaching is Influenced by the Teacher
Janet Clinton, Ruth Aston, and Georgia Kate Dawson

4040. Teacher Education and the Arts: Challenges, Tensions, and Opportunities
Robyn Ann Ewing

4041. Emotion and Reconciliation Pedagogies
Michalinos Zembylas

4093. Teacher education and the development of teacher noticing
Johannes König

Initial Teacher Education

4083. Teacher Recruitment and Selection
Robert Klassen, Jade Rushby, and Tracy Durksen

4090. Professional agency of student teachers in initial teacher education
Äli Leijen

4038. Teacher Identity: Crossing the Technical-Rationalist and Affective Divide
Sue Ellis and Navan Govender

4109. Teacher Education for Diverse Learners
Katrina Liu and Richard Miller

4049. Educating the Teaching Profession
Solvi Lillejord

4045. International Perspectives on Community-Engaged Teacher Education
Rocío García-Carrión and Eugenia Allotey

4002. University-School Partnerships in Initial Teacher Education
Maureen Robinson and Jennifer Feldman

4116. Inside looking out: A third space perspective on preservice teachers' practice experiences
Viviana Daza

4051. Case Methods in Teacher Education
Misty Sato

4056. Contemporary Challenges and Approaches in Anti-racist Teacher Education
Verna St. Denis, Jenn K. Bergen, and Sharissa Unger Hanthe

4047. Teaching Performance Assessments

Larissa McLean-Davies

4021. Determining Long-Term Effects of Teacher Education

REPRINT

Those Who Educate Teachers: Teacher Educators

4016. A Lifetime of Experience: An International Study on the Lives of Teacher Educators

Clare Kosnik, Pooja Dharamshi, and Lydia Menna

4053. Self-Study of Teacher Education Practices: The Complex Challenges of Learning from Experience

Tom Russell and Rodrigo Fuentealba Jara

4050. Professional Development or Professional Learning: Developing Teacher Educators' Professional Expertise

Gerry Czerniawski

4097. Equity- and Justice-minded Teacher Educators

Jamy Stillman

4058. Improving Teacher Education through the Self-Study of Practice

Julian Kitchen

4111. Preparation of novice teacher educators for critical, justice-oriented teacher education

Hilary Conklin

4089. Teacher Educators' Professional Growth as Researchers

Leah Shagrir

Ongoing Teacher Professional Learning

4061. Questioning the Consensus on Effective Professional Development

Jenny Gore, Sally Patfield, and Leanne Fray

4062. Situating Teacher Learning in the Mathematics Classroom and Everyday Practice

Janette Bobis and Kristen Tripet

4066. Middle Leading as a Practice for Enabling Professional Learning

Karin Rönnerman and Veronica Sülau

4031. Developing Teachers: A Necessary Condition for Quality Retention

Qing Gu

4068. Contemporary Approaches to Teacher Professional Development

Karen Koellner, Hilda Borko, Jennifer Jacobs, and Nanette Seago

4069. Teacher Learning as Workplace Learning

Jeroen Imants and Klaas van Veen

4118. Lesson Study: A Practice for Teachers' Training and Professional Development at any educational level

Gabriel Hervas and José Luis Medina

4063. Examining Conceptual Frameworks for Understanding Teacher Learning in the Domain of Formative Assessment

Brent Duckor and Carrie Holmberg

4064. Conceptualizing Professional Development with and for a Diverse and Globally Mobile Early Childhood Education Workforce

Joce Nuttall, Susan Grieshaber, Sirene Lim, and Eunju Yun

Future and Emerging Issues: Research, Policy, and Practice

4094. The Global and the Local: Idea Flows, Contexts, and the Influencing of Education Policy in the 21st Century

Jo Helgetun

4091. Teachers' Research Capacity and Initial Teacher Education Policy

Clare Brooks

4070. Studying Teachers' Work and Lives: Insights from research before and during the COVID-19 pandemic
Maria Assunção Flores
4101. Teachers' Work in the Context of COVID-19
Lora Bartlett, Judith Warren Little, Alisun Thompson, Lina Darwich, and Riley Collins
4034. Reclaiming the Artistry of Teaching
Gert Biesta
4084. Tackling teacher shortages: doing the right thing at the right time and with the right evidence
Beng Huat See, Stephen Gorard, and Rebecca Morris
4048. Educating a World on the Move: Rethinking Teacher Preparation in the Context of Mass Global Migration and Increasingly Diverse Classrooms
A. Lin Goodwin, Seung Eun McDevitt, Crystal Chen Lee, and Sibel Akin-Sabuncu
4044. Research for Socially Progressive Teacher Education Programs
Aileen Kennedy
4043. The contribution of initial teacher education to social justice
Fiona Ell
4086. Online Teaching and Learning Practices in Teacher Education: Past, Present, and Future
Carmen Carrillo and Maria Assunção Flores
4009. Rethinking Teacher Education in Pandemic Times: Lessons from and for Singapore
Ee Ling Low
4017. Reinventing Teacher Agency: Shifting the Narrative of Teacher Education
Poonam Batra
4074. Mentoring 'in the Time of Cholera': Perspectives, Promises, and Challenges in the Preparation of Mentors for Diversity in the New Digital Teacher Learning Space
Lily Orland-Barak
4071. Comparative Teacher Education Research in Turbulent Times
Ian Menter and Maria Teresa Tatto

References

- British Educational Research Association, 2014. Research and the Teaching Profession: Building the Capacity for a Self-Improving Education System. Final Report of the BERA-RSA Inquiry Into the Role of Research in Teacher Education. BERA.
- Burns, T., Schuller, T. (Eds.), 2007. Evidence in Education. Linking Research and Policy. OECD Publishing.
- Chung, J., 2016. The (mis)use of the Finnish teacher education model: "policy-based evidence-making"? *Educ. Res.* 58 (2), 207–219. <https://doi.org/10.1080/00131881.2016.1167485>.
- Cochran-Smith, M., Carney, M.C., Keefe, E.S., Burton, S., Chang, W.-C., Fernandez, M.B., Miller, A.F., Sanchez, J.G., Baker, M., 2018. Reclaiming Accountability in Teacher Education. Teachers College Press.
- Cochran-Smith, M., Ell, F., Ludlow, L., Grudnoff, L., Aitken, G., 2014. The challenge and promise of complexity theory for teacher education research. *Teachers College Record* 116 (5), 1–38.
- Evetts, J., 2013. Professionalism: value and ideology. *Curr. Sociol.* 61 (5–6), 778–796. <https://doi.org/10.1177/0011392113479316>.
- Gray, D., Colucci-Gray, L., 2010. Challenges to ITE research in conditions of complexity. *J. Educ. Teach.* 36 (4), 425–439. <https://doi.org/10.1080/02607476.2010.513856>.
- Helgetun, J.B., Menter, I., 2020. From an age of measurement to an evidence era? Policy-making in teacher education in England. *J. Educ. Policy* 1–18. <https://doi.org/10.1080/02680939.2020.1748722>.
- Menter, I., Hulme, M., Elliot, D., Lewin, J., 2010. Literature Review on Teacher Education in the Twenty-First Century. Education Analytical Services, Schools Research, Scottish Government.
- Peck, J., Theodore, N., 2015. Fast Policy: Experimental Statecraft at the Thresholds of Neoliberalism. University of Minnesota Press. <http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=2050509>.
- Rautalin, M., Alasuutari, P., Vento, E., 2019. Globalisation of education policies: does PISA have an effect? *J. Educ. Policy* 34 (4), 500–522. <https://doi.org/10.1080/02680939.2018.1462890>.
- Sachs, J., 2003. The Activist Teaching Profession. Open University Press.
- Sachs, J., 2016. Teacher professionalism: why are we still talking about it? *Teach. Teach.* 22 (4), 413–425. <https://doi.org/10.1080/13540602.2015.1082732>.
- Sellar, S., Lingard, B., 2013. The OECD and global governance in education. *J. Educ. Policy* 28 (5), 710–725. <https://doi.org/10.1080/02680939.2013.779791>.
- Sleeter, C., 2014. Towards teacher education research that informs policy. *Educ. Res.* 43 (3), 146–153.
- Whitty, G., 2014. Recent developments in teacher training and their consequences for the 'University Project' in education. *Oxford Review of Education* 40 (4), 466–481. <https://doi.org/10.1080/03054985.2014.933007>.
- Zeichner, K., Conklin, H., 2016. Beyond knowledge ventriloquism and echo chambers: raising the quality of the debate in teacher education. *Teachers College Record* 118 (4), 1–38.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 5

Sibel Akin-Sabuncu

*TED University, Ankara, Turkey; and Teachers College,
Columbia University, New York, NY, United States*

Eugenia Allotey

University of Deusto, Bilbao, Spain

Ruth Aston

University of Melbourne, Melbourne, VIC, Australia

Beatrice Ávalos

*Center for Advanced Research in Education, Institute of
Education, University of Chile, Santiago, Chile*

Arnetha F Ball

*Graduate School of Education, Stanford University,
Stanford, CA, United States*

Lora Bartlett

*University of California, Santa Cruz, Santa Cruz, CA,
United States*

Poonam Batra

*India Lead, TESF Network Plus; Formerly with Central
Institute of Education, University of Delhi, Delhi, India*

Douwe Beijaard

*Eindhoven University of Technology, Eindhoven, the
Netherlands*

Jenn K Bergen

*College of Education, University of Saskatchewan,
Saskatoon, SK, Canada*

Gert Biesta

*The Moray House School of Education and Sport,
University of Edinburgh, Edinburgh, United Kingdom;
and Center for Public Education and Pedagogy,
National University of Ireland at Maynooth, Co.
Kildare, Ireland*

Joshua Bleiberg

*University of Pittsburgh, Pittsburgh, PA,
United States*

Janette Bobis

University of Sydney, Sydney, NSW, Australia

Hilda Borko

*Graduate School of Education, Stanford University,
Stanford, CA, United States*

Clare Brooks

UCL Institute of Education, London, United Kingdom

CN Brouwer

*Radboud University Nijmegen, Nijmegen, the
Netherlands*

Katharine Burn

*University of Oxford, Department of Education, Oxford,
United Kingdom*

Carmen Carrillo

*Department of Music, Plastic Arts and Physical
Education, Faculty of Educational Sciences, Autonomous
University of Barcelona, Bellaterra, Spain*

Jeffrey Casely-Hayford

*Unit of Intervention & Implementation Research for
Worker Health, Institute of Environmental Medicine,
Karolinska Institute, Stockholm, Sweden*

Anthony Clarke

*The University of British Columbia, Department of
Curriculum and Pedagogy, Vancouver, British
Columbia, Canada*

Matthew Clarke

University of Aberdeen, Scotland

Janet Clinton

University of Melbourne, Melbourne, VIC, Australia

Alf Coles

*School of Education, Helen Wodehouse Building, Bristol,
United Kingdom*

Riley Collins

*University of California, Santa Cruz, Santa Cruz, CA,
United States*

Hilary G Conklin

*DePaul University, College of Education, Chicago, IL,
United States*

Gerry Czerniawski
University of East London, United Kingdom

Lina Darwich
Lewis & Clark College, Portland, OR, United States

Georgia Dawson
University of Melbourne, Melbourne, VIC, Australia

Christopher Day
School of Education, University of Nottingham, Nottingham, United Kingdom

Viviana Daza
Department of Teacher Education and School Research, Faculty of Educational Sciences, University of Oslo, Oslo, Norway

Pooja Dharamshi
Simon Fraser University, Burnaby, Canada

Janice Doucette
Werklund School of Education, University of Calgary, Calgary, AB, Canada

Barry Down
University of South Australia, Adelaide, SA, Australia

Brent Duckor
San José State University, Department of Teacher Education, Educational Leadership, San Jose, CA, United States

Tracy L Durksen
School of Education, University of New South Wales, Sydney, NSW, Australia

Khaleda Gani Dutt
Department of Special Education, Stockholm University, Stockholm, Sweden

Fiona R Ell
Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

Sue Ellis
University of Strathclyde, School of Education, Glasgow, United Kingdom

Sibel Erduran
University of Oxford, Oxford, United Kingdom

Robyn Ann Ewing
Sydney School of Education & Social Work, NSW, Australia; and Faculty of Arts and Social Sciences, Australia University of Sydney, Sydney, NSW, Australia

Jennifer Feldman
Department of Education Policy Studies, Faculty of Education, Stellenbosch University, Stellenbosch, South Africa

Maria Beatriz Fernandez
Center for Advanced Research in Education, Institute of Education, University of Chile, Santiago, Chile

Maria Assunção Flores
Research Centre for Child Studies, Institute of Education, University of Minho, Campus de Gualtar, Braga, Portugal

Leanne Fray
Teachers and Teaching Research Centre, School of Education, University of Newcastle, Callaghan, NSW, Australia

Sharon Friesen
Werklund School of Education, University of Calgary, Calgary, AB, Canada

Rodrigo Fuentealba Jara
Facultad de Educación, Universidad Autónoma, Santiago, Chile

Rocío García-Carrión
University of Deusto, Ikerbasque. Basque Foundation for Science, Bilbao, Spain

Johan Geertsema
University Scholars Programme and Ridge View Residential College, National University of Singapore, Singapore

A Lin Goodwin
University of Hong Kong, Hong Kong, HKSAR, China

Stephen Gorard
Durham Evidence Centre for Education, Durham University, Durham, United Kingdom

Anna Lise Gordon
St Mary's University, Twickenham, United Kingdom

Jennifer Gore
Teachers and Teaching Research Centre, School of Education, University of Newcastle, Callaghan, NSW, Australia

Navan N Govender
University of Strathclyde, School of Education, Glasgow, United Kingdom

Susan Grieshaber
La Trobe University, Melbourne, VIC, Australia

Qing Gu
UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Hafþís Guðjónsdóttir
School of Education, University of Iceland, Reykjavík, Iceland

Liam Guilfoyle
University of Oxford, Oxford, United Kingdom

Amy E Guillotte
University of Pennsylvania, Graduate School of Education, Philadelphia, PA, United States

Hazel Hagger
Kellogg College, Oxford, United Kingdom

Sharissa Unger Hantke
College of Education, University of Saskatchewan, Saskatoon, SK, Canada

Ian Hardy
School of Education, The University of Queensland, Brisbane, QLD, Australia

JB Helgetun
Institute for the Analysis of Change in Historical and Comparative Societies, University of Louvain, Louvain-la-Neuve, Belgium

Michèle Hinton Herrington
Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC, Australia

Gabriel Hervas
Center for Research and Development of Higher Education, The University of Tokyo, Tokyo, Japan

Carrie Holmberg
San José State University, Department of Teacher Education, San Jose, CA, United States

Moirá Hulme
Manchester Metropolitan University, Manchester, United Kingdom

Jeroen Imants
Radboud Teacher Academy, Radboud University Nijmegen, Zeist, the Netherlands

Jenni Ingram
Department of Education, University of Oxford, Oxford, United Kingdom

Jennifer C Ingrey
Faculty of Education, Western University, London, ON, Canada

Jennifer Jacobs
Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, United States

Stefan Johansson
Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden

Bruce Johnson
University of South Australia, Adelaide, SA, Australia

Jaakko Kauko
Faculty of Education and Culture, Tampere University, Tampere, Finland

Geert Kelchtermans
Center for Innovation and the Development of Teacher and School (CIDTS) – KU, Leuven, Belgium

Aileen Kennedy
University of Strathclyde, School of Education, Glasgow, United Kingdom

Julian Kitchen
Brock University, Hamilton, ON, Canada

Robert M Klassen
Department of Education, University of York, Heslington, United Kingdom

Manon Kluijtmans
Centre for Academic Teaching, Utrecht University, Utrecht, Netherlands; and Center for Education, University Medical Center Utrecht, Utrecht, Netherlands

Jake Knaus
Minneapolis Public Schools, Minneapolis, MN, United States

Karen Koellner
Arizona State University, Tempe, AZ, United States

Johannes König
University of Cologne, Faculty of Human Sciences, Department of Education and Social Sciences, Empirical School Research, Quantitative Methods, Cologne, Germany

Clare Kosnik
University of Toronto, Toronto, Canada

Lydia Kwak
Unit of Intervention & Implementation Research for Worker Health, Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden

Nancy Law
University of Hong Kong, Pok Fu Lam, Hong Kong

Crystal Chen Lee
North Carolina State University, Raleigh, NC, United States

Äli Leijen

Institute of Education, University of Tartu, Tartu, Estonia

Qiong Li

Center for Teacher Education Research, Beijing Normal University, Key Research Institute, Ministry of Education, Beijing, China

Leming Liang

Guandong Polytechnic Normal University, Guangdong, China

Sølvi Lillejord

SLATE, Centre for the Science of Learning & Technology, University of Bergen, Bergen, Norway; and Department of Education, Oxford University, Oxford, United Kingdom

Sirene Lim

Singapore University of Social Sciences, Singapore, Republic of Singapore

Judith Warren Little

University of California, Berkeley, Berkeley, CA, United States

Katrina Liu

School of Teacher Education, Department of Teaching and Learning, University of Nevada, Las Vegas, Las Vegas, NV, United States

Min Liu

Beijing Normal University, Beijing, China

Ee-Ling Low

National Institute of Education, Nanyang Technological University, Singapore

Seung Eun McDewitt

St. John's University, Queens, NY, United States

Larissa McLean Davies

Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC, Australia

José Luis Medina

Department of Didactics and Educational Organization, University of Barcelona, Barcelona, Spain

Lydia Menna

University of Alberta, Edmonton, Canada

Ian Menter

University of Oxford, Oxford, United Kingdom; and Kazan Federal University, Kazan, Russia

Richard Miller

School of Music University of Nevada, Las Vegas, NV, United States

Martin Mills

Centre for Teachers and Teaching Research, UCL Institute of Education, London, United Kingdom

Monica Mincu

University of Torino, Turin, Italy; and University College London, London, UK

Nicole Mockler

Sydney School of Education and Social Work, University of Sydney, NSW, Australia

Tebeje Molla

School of Education, Deakin University, Geelong, VIC, Australia

Rebecca Morris

Department for Education Studies, Warwick University, Coventry, United Kingdom

Trevor Mutton

University of Oxford, Department of Education, Oxford, United Kingdom

Andrea Nolan

School of Education, Deakin University, Geelong, VIC, Australia

Joce Nuttall

Australian Catholic University, Melbourne, VIC, Australia

Lily Orland-Barak

Faculty of Education, University of Haifa, Haifa, Israel

Edda Óskarsdóttir

School of Education, University of Iceland, Reykjavík, Iceland

Sally Patfield

Teachers and Teaching Research Centre, School of Education, University of Newcastle, Callaghan, NSW, Australia

Kinga Petrovai

University of Oxford, Oxford, United Kingdom

Anne M Phelan

University of British Columbia, Vancouver, British Columbia, Canada

Emilee Rauschenberger

Manchester Metropolitan University, Manchester, United Kingdom

Maureen Robinson

Department of Curriculum Studies, Faculty of Education, Stellenbosch University, Stellenbosch, South Africa

Karin Rönnerman

*Department of Education and Special Education,
University of Gothenburg, Sweden*

Jade V Rushby

*School of Education, University of New South Wales,
Sydney, NSW, Australia*

Tom Russell

*Faculty of Education, Queen's University, Kingston,
ON, Canada*

Pasi Sahlberg

Southern Cross University, Lismore, NSW, Australia

Mistilina Sato

University of Canterbury, Christchurch, New Zealand

Gonny Schellings

*Eindhoven University of Technology, Eindhoven, the
Netherlands*

Nanette Seago

WestEd, Riverside, CA, United States

Beng Huat See

*Durham Evidence Centre for Education, Durham
University, Durham, United Kingdom*

Leah Shagrir

Levinsky College of Education, Tel Aviv-Yafo, Israel

Verna St. Denis

*College of Education, University of Saskatchewan,
Saskatoon, SK, Canada*

Meghan Stacey

*School of Education, University of New South Wales,
Sydney, Australia*

Jamy Stillman

*University of Colorado Boulder, School of Education,
Boulder, CO, United States*

Andrea Stringer

*University of New South Wales, Sydney, NSW,
Australia*

Gabriel J Stylianides

University of Oxford, Oxford, United Kingdom

Veronica Sülau

Frida Education, Vänersborg, Sweden

Anna M Sullivan

University of South Australia, Adelaide, SA, Australia

Dennis Sumara

*Werklund School of Education, University of Calgary,
Calgary, AB, Canada*

Xiaohong Sun

*Center for Teacher Education Research, Beijing Normal
University, Key Research Institute, Ministry of
Education, Beijing, China*

Catherine Takasugi

*Werklund School of Education, University of Calgary,
Calgary, AB, Canada*

Maria Teresa Tatto

*Kazan Federal University, Kazan, Russia; and Mary Lou
Fulton Teachers College, Arizona State University,
Tempe, AZ, United States*

Antonina Tereshchenko

*Centre for Teachers and Teaching Research, UCL
Institute of Education, London, United Kingdom*

Alisun Thompson

University of Puget Sound, Tacoma, WA, United States

Anna Toropova

*Unit of Intervention & Implementation Research for
Worker Health, Institute of Environmental Medicine,
Karolinska Institute, Stockholm, Sweden*

Kristen Tripet

*Australian Academy of Science, Canberra, ACT,
Australia*

Jennifer Turner

*Werklund School of Education, University of Calgary,
Calgary, AB, Canada*

Eline Vanassche

*Faculty of Psychology and Educational Sciences, KU
Leuven, Kortrijk, Belgium*

Marieke van der Schaaf

*Center for Education, University Medical Center
Utrecht, Utrecht, Netherlands; and Utrecht Center for
Research and Development in Health Professions
Education, University Medical Center Utrecht, Utrecht,
Netherlands*

Esther E van Dijk

*Department of Education, Faculty of Social and
Behavioral Sciences, Utrecht University, Utrecht,
Netherlands; Centre for Academic Teaching, Utrecht
University, Utrecht, Netherlands; and Center for
Education, University Medical Center Utrecht, Utrecht,
Netherlands*

Jan van Tartwijk

*Department of Education, Faculty of Social and
Behavioral Sciences, Utrecht University, Utrecht,
Netherlands*

Klaas van Veen

*Institute of Teacher Education, University of Groningen,
Groningen, the Netherlands*

Lily Vincent

Teach First, London, England

Daniel A Wagner

*UNESCO Chair in Learning and Literacy, University of
Pennsylvania, Graduate School of Education,
Philadelphia, PA, United States*

Niall Winters

University of Oxford, Oxford, United Kingdom

Eunju Yun

*Sookmyung Women's University, Seoul,
South Korea*

Michalinos Zembylas

*Open University of Cyprus, Nelson Mandela University,
Gqeberha, South Africa*

Jingtian Zhou

*Department of Educational Administration and Policy,
The Chinese University of Hong Kong, Hong Kong,
China*

CONTENTS OF VOLUME 5

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 5: Teachers' Lives, Work, and Professional Education</i>	<i>xix</i>
<i>Contributors to Volume 5</i>	<i>xxvii</i>

THE POLICY CONTEXT OF TEACHERS' WORK AND PROFESSIONAL EDUCATION

The role and influence of non-state actors on teachers' lives and work <i>Moirra Hulme and Emilee Rauschenberger</i>	1
Teacher education policy and professionalism: a personal review of teacher education policy research <i>Eline Vanassche</i>	10
Standardized learning: the "problem" of teachers' work and learning under performative policy conditions <i>Ian Hardy</i>	20

RESEARCHING TEACHERS, THEIR LIVES AND THEIR LEARNING

Teachers' work amid global education reform: "the greatest challenge"? <i>Nicole Mockler and Meghan Stacey</i>	29
---	----

THE POLICY CONTEXT OF TEACHERS' WORK AND PROFESSIONAL EDUCATION

Defiant imaginaries: from a politics of identity to an ethics of friendship in teaching <i>Anne M Phelan, Matthew Clarke, and Lily Vincent</i>	39
Global politics of quality in education <i>Jaakko Kauko</i>	49
Revisiting teacher evaluation a decade after reforms <i>Joshua Bleiberg</i>	54

Teacher policy frames, enactment forms and issues <i>Beatrice Ávalos and Maria Beatriz Fernandez</i>	62
The context in teacher education and professionalism: hierarchies, networks and markets in four countries <i>Monica Mincu and Min Liu</i>	72
Can initial teacher education partnerships maintain their epistemological and pedagogical underpinnings in light of market-driven reforms? The case of England <i>Trevor Mutton, Katharine Burn, and Hazel Hagger</i>	83
Equity in early childhood education: policy framing, translation and enactment <i>Andrea Nolan and Tebeje Molla</i>	93
Teachers and teaching in low-income countries <i>Daniel A Wagner and Amy E Guillotte</i>	99
TEACHERS' LIVES AND WORK	
Narrative and biography in teacher work lives and development <i>Geert Kelchtermans</i>	106
Learning to count what counts: narrative hermeneutics for teachers and teaching <i>Dennis Sumara, Sharon Friesen, Catherine Takasugi, Janice Doucette, and Jennifer Turner</i>	113
Researching teacher learning through self-video <i>Jenni Ingram and Alf Coles</i>	120
Retention of teachers from minority ethnic groups in disadvantaged schools <i>Antonina Tereshchenko and Martin Mills</i>	127
Early career teachers' work <i>Anna M Sullivan, Barry Down, and Bruce Johnson</i>	134
Coaching for professional learning of early career teachers <i>Pasi Sahlberg and Andrea Stringer</i>	144
Early career teaching and resilience <i>Anna Lise Gordon</i>	153
Beginning teachers' professional identity work and its support <i>Gonny Schellings and Douwe Beijaard</i>	161
Teacher expertise and how it develops during teachers' professional lives <i>Jan van Tartwijk, Esther E van Dijk, Johan Geertsema, Manon Kluijtmans, and Marieke van der Schaaf</i>	170
Teachers willingness and ability to teach to their best and well: the person in the professional <i>Christopher Day</i>	180
Thriving on challenges: understanding teacher resilience in hard-to-staff Chinese rural schools <i>Qiong Li, Xiaohong Sun, and Jingtian Zhou</i>	191
Affirming trans and queer teachers' lives: rethinking teacher professionalization and pedagogy <i>Jennifer C Ingrey</i>	204
The role of job demands and resources for teacher job satisfaction: insights from five European countries <i>Jeffrey Casely-Hayford, Stefan Johansson, Khaleda Gani Dutt, Lydia Kwak, and Anna Toropova</i>	213
Teacher inquiry: by any other name <i>Anthony Clarke</i>	232

KNOWLEDGE FOR TEACHING

Enacting inclusive pedagogy in teacher education: creating a learning space for teachers to develop their professional identity	243
<i>Hafðís Guðjónsdóttir and Edda Óskarsdóttir</i>	
Learning to teach epistemic practices of science in secondary schools	254
<i>Sibel Erduran and Liam Guilfoyle</i>	
Teacher skills and knowledge for technology integration	263
<i>Leming Liang and Nancy Law</i>	
The relationship between knowledge exchange and teachers' use of mobile technologies in the mathematics classroom	272
<i>Kinga Petrovai, Gabriel J Stylianides, and Niall Winters</i>	
Teaching is influenced by the teacher	287
<i>Janet Clinton, Ruth Aston, and Georgia Dawson</i>	
Teacher education and the arts: challenges, tensions, opportunities	304
<i>Robyn Ann Ewing</i>	
Emotion and reconciliation pedagogies	315
<i>Michalinos Zembylas</i>	
Teacher education and the development of teacher noticing	323
<i>Johannes König</i>	

INITIAL TEACHER EDUCATION

Teacher recruitment: a competency-based approach	329
<i>Robert M Klassen, Jade V Rushby, and Tracy L Durksen</i>	
Supporting and assessing professional agency of student teachers in initial teacher education	339
<i>Äli Leijen</i>	
Teacher identity: crossing the technical-rationalist and affective divide	346
<i>Navan N Govender and Sue Ellis</i>	
Teacher education for diverse learners	356
<i>Katrina Liu, Richard Miller, and Arnetha F Ball</i>	
Educating the teaching profession	368
<i>Sølvi Lillejord</i>	
International perspectives on community-engaged teacher education	375
<i>Rocío García-Carrión and Eugenia Allotey</i>	
University-school partnerships in initial teacher education	381
<i>Maureen Robinson and Jennifer Feldman</i>	
Inside looking out: a third space perspective on pre-service teachers' practice experiences	393
<i>Viviana Daza</i>	
Case methods in teacher education	405
<i>Mistilina Sato and Jake Knaus</i>	
Contemporary challenges and approaches in anti-racist teacher education	414
<i>Jenn K Bergen, Sharissa Unger Hantke, and Verna St. Denis</i>	

Teaching Performance Assessments (TPAs)	427
<i>Larissa McLean Davies and Michèle Hinton Herrington</i>	
Determining long term effects of teacher education	437
<i>CN Brouwer</i>	

THOSE WHO EDUCATE TEACHERS: TEACHER EDUCATORS

A lifetime of experience: an international study on the lives of teacher educators	445
<i>Clare Kosnik, Lydia Menna, and Pooja Dharamshi</i>	
Self-study of teacher education practices: the complex challenges of learning from experience	458
<i>Tom Russell and Rodrigo Fuentealba Jara</i>	
Professional development or professional learning: developing teacher educators' professional expertise	469
<i>Gerry Czerniawski</i>	
Equity- and justice-minded teacher educators	475
<i>Jamy Stillman</i>	
Improving teacher education through the self-study of practice	488
<i>Julian Kitchen</i>	
Preparing novice teacher educators for critical, justice-oriented teacher education practice	497
<i>Hilary G Conklin</i>	
Teacher educators' professional growth as researchers	504
<i>Leah Shagrir</i>	

ONGOING TEACHER PROFESSIONAL LEARNING

Questioning the consensus on effective professional development	511
<i>Jennifer Gore, Sally Patfield, and Leanne Fray</i>	
Situating teacher learning in the mathematics classroom and everyday practice	518
<i>Janette Bobis and Kristen Tripet</i>	
Middle leading as a practice for enabling professional learning	528
<i>Karin Rönnerman and Veronica Siilau</i>	
The challenge to quality retention: why schools matter	539
<i>Qing Gu</i>	
Current trends, tensions and unresolved issues in research on teacher professional learning	550
<i>Karen Koellner, Jennifer Jacobs, Hilda Borko, and Nanette Seago</i>	
Teacher learning as workplace learning	562
<i>Jeroen Imants and Klaas van Veen</i>	
Lesson study: a practice for teachers' training and professional development at any educational level	570
<i>Gabriel Hervas and José Luis Medina</i>	

Examining conceptual frameworks for understanding teacher learning in the domain of formative assessment	581
<i>Brent Duckor and Carrie Holmberg</i>	
Conceptualizing professional development with and for a diverse and globally mobile early childhood education workforce	596
<i>Joce Nuttall, Susan Grieshaber, Sirene Lim, and Eunju Yun</i>	
FUTURE AND EMERGING ISSUES: RESEARCH, POLICY AND PRACTICE	
The global and the local: idea flows, contexts, and the influencing of education policy in the 21st century	604
<i>JB Helgetun</i>	
Teachers' research capacity and initial teacher education policy	611
<i>Clare Brooks</i>	
Studying teachers' work and lives: insights from research before and during the COVID-19 pandemic	622
<i>Maria Assunção Flores</i>	
Teachers' work in the context of COVID-19	636
<i>Riley Collins, Lora Bartlett, Alisun Thompson, Judith Warren Little, and Lina Darwich</i>	
Reclaiming the artistry of teaching	648
<i>Gert Biesta</i>	
Tackling teacher shortages: doing the right thing at the right time and with the right evidence	655
<i>Beng Huat See, Stephen Gorard, and Rebecca Morris</i>	
Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms	671
<i>A Lin Goodwin, Seung Eun McDevitt, Crystal Chen Lee, and Sibel Akin-Sabuncu</i>	
Research for socially progressive teacher education programs	681
<i>Aileen Kennedy</i>	
The contribution of initial teacher education to social justice	690
<i>Fiona R Ell</i>	
Online teaching and learning practices in teacher education: past, present and future	698
<i>Carmen Carrillo and Maria Assunção Flores</i>	
Rethinking teacher education in pandemic times: lessons from and for Singapore	710
<i>Ee-Ling Low</i>	
Reinventing teacher agency: shifting the narrative of teacher education	721
<i>Poonam Batra</i>	
Mentoring "in the time of cholera": perspectives, promises and challenges in the preparation of mentors for diversity in the new digital teacher learning space	729
<i>Lily Orland-Barak</i>	
Comparative teacher education research in turbulent times	740
<i>Ian Menter and Maria Teresa Tatto</i>	

This page intentionally left blank

The role and influence of non-state actors on teachers' lives and work

Moira Hulme and Emilee Rauschenberger, Manchester Metropolitan University, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Supranational organizations	1
The global education industry	4
Private school chains and teachers' work	5
Philanthropic foundations and teacher education	6
Conclusion	7
References	7

Introduction

A growing number of non-state actors work in education, shaping the lives of teachers and the policies related to their work. Non-state actors are individual and collective actors that are unelected and independent of government and range from the well-known supranational bodies of the World Bank, the Organization for Economic Co-operation and Development (OECD) and the United Nations Educational, Scientific and Cultural Organization (UNESCO), to national and international non-profit organizations and foundations, to private philanthropists and for-profit edu-businesses. The involvement of such actors at varying levels of the educational landscape has rapidly expanded in the past two decades, both in developed and developing countries alike. This development has fueled widespread debates about the purpose of education, data use in policy deliberation, the responsibilities of national and sub-national government, and the role of the private sector.

The article is organized in two main sections. The first section outlines the key data initiatives and policy tools used to leverage system reform by influential supranational agencies. Principal among these are the OECD's Program for International Student Assessment (PISA), the Teaching and Learning International Survey (TALIS), and the World Bank's SABER initiative (Systems Approach for Better Education Results), including SABER-Teachers. The second section discusses the rise of the Global Education Industry (GEI) with its mix of non-profit and for-profit non-state actors working in dense evolving networks with the support of a variety of corporations, foundations, and private philanthropists. Particular attention is directed at how non-state actors frame the problem of teacher quality and propose market-based solutions using the examples of low cost private schools and fast track teacher preparation programs. The article concludes with consideration of the impact of supranational bodies and the Global Education Industry on teachers' work and their workplaces.

Supranational organizations

After two world wars and the formal end of European colonization, countries old and new focused on education as part of nation building while looking to international cooperation to provide support and stability. In this context, a number of supra-national organizations were established and became central to the development of a global education space in the latter half of the twentieth century. The UNESCO, OECD and the World Bank are among the most influential. These organizations provide financial, regulatory, and agenda-setting functions and play a key role in framing and communicating policy ideas influencing the selection, development, evaluation and compensation of teachers. Increasingly, they serve as knowledge producers and mobilizers—roles that are entrenching and expanding their legitimacy and power (Zapp, 2020). Their work is reshaping the work and lives of teachers across the world as school systems are accorded value in terms of measurable performance outcomes and their future contribution to national economic growth. Through this process, non-governmental actors have gained greater control of how the good teacher is classified and framed.

UNESCO was founded in 1945 by 37 countries as a human rights organization to promote international cooperation on issues related to education, science and cultural causes. Although initially primarily Western-funded and European-focused, the work of UNESCO broadened geographically and shifted to global problems of poverty, illiteracy, and underdevelopment, including education, as more developing nations joined from the 1950s onwards. UNESCO brought a global focus to education through its collaborative work setting up the World Conferences on Education for All (EFA) in Jomtien, Thailand in 1990 and Dakar, Senegal in 2000. Through this work, in collaboration with the World Bank and four other United Nations agencies, UNESCO initiated the EFA movement bringing together government and non-state actors around six goals, which include ensuring free universal primary education for both all (Verger and Novelli, 2012). With global pressure on national governments to expand their education system to meet ambitious goals, the funding and participation of private, non-state actors with a variety of interests became increasingly welcomed and normalized. From 2001, using the EFA framework and Sustainable Development Goals (SDGs), the strategic activities of the OECD and World Bank increasingly moved from ensuring access to ensuring quality (OECD, 2005, 2015).

Through its data gathering and knowledge mobilization activities, the OECD is now a central actor directing global education discourse. Sellar and Lingard (2014, p. 932) argue that the OECD exerts “infrastructural governance”, referring to the international networks and systems it has built to enable its vast data collection activities, and “epistemological governance”, referring to the OECD’s ability to influence leading education actors across contexts and levels. Established in 1961 to stimulate economic growth and world trade, the OECD originally had 19 members—16 European countries and the US, Canada, and Turkey. By 2020, the OECD had 37 member governments and offered its analytical expertise to over 100 developing and emerging market economies. In contrast to UNESCO’s emphasis on “global citizenship” and cultural interdependency, the OECD conceptualizes “global competence” narrowly as an economic necessity (Vaccari and Gardinier, 2019, p. 68). From 2002, the OECD has taken a more active role in education with a designated Education Directorate.

Perhaps the most utilized global indicator of educational quality is the OECD’s Program for International Student Assessment (PISA). Launched in 2000, PISA soon became a proxy measure for education quality aligned with human or knowledge capital. Conducted every three years, PISA tests assess the knowledge and skills of 15-year-olds in reading, mathematics and science. PISA is not simply curriculum-based like the International Association for the Evaluation of Educational Achievement’s (IEA) Trends in Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS). In contrast, PISA claims to measure how well students are equipped to meet future challenges. Optional parent and teacher questionnaires were added in 2006 and 2015 respectively, with low take up rates. Seventy-nine countries participated in PISA 2018. In addition to the main tests, the OECD has extended its reach by offering national, subnational, and school-level educational performance data through *PISA for Schools* and professional learning through the *PISA4U* online platform (Lewis et al., 2016; Lewis, 2020).

The agendas of the OECD, UNESCO and the World Bank converge in the promotion of PISA in international development. There is a shared commitment to building a global testing regime, although it should be acknowledged that the evidence used to link international test scores and economic growth over time is inconclusive (Kamens, 2015). In 2013, the PISA for Development (PISA-D) initiative—an adaptation of the existing PISA test specifically designed for low and middle-income nations—was piloted in nine countries. The OECD plans to have 170 countries participating in PISA by 2030, with the involvement of low and middle-income nations in the regular PISA cycle from 2021 onwards. While the expansion of PISA to low-income nations is espoused as demand-led, evidence from interviews with in-country actors suggests this is not always the case. For example, Auld et al. (2019) argue that the introduction of PISA-D in Cambodia in 2016 was not locally owned, but by-passed both earlier aid coordination mechanisms and the concerns of the Cambodian government that National Assessment plans would be undermined. Efforts to align national and regional assessments with PISA run the risk of disenfranchizing educators if national systems for student assessment and school inspections are considered subordinate to the interests of international policy harmonization. For example, in Ecuador and Paraguay, Addey (2017) argues that greater attention was given to aligning national assessment procedures with PISA and PISA-D than to the development of a testing system that was country-specific and policy relevant. Moreover, it is debatable whether building technical capacity for testing can be equated with building educator capacity to enhance learning.

Politicians, civil servants and journalists pay close attention to national positions within PISA performance rankings. Changes in national standing are amplified by the media in a triennial international spectacle that creates winners and losers. The education systems of high achieving countries such as Finland, South Korea and Shanghai-China attract attention for the purposes of lesson drawing (OECD, 2018). The feasibility of transference to diverse geo-political contexts is less widely discussed, despite Finnish exposition of the historical and cultural conditions that contribute over time to higher ranking (Sahlberg, 2011). The “PISA shock” resulting from lowering of a country’s comparative position evokes debates in Parliament and among teachers’ and school leaders’ unions, as was the case in Germany, Denmark, Poland, and Ireland (Michel, 2017).

Participation in high profile international assessments has implications for teachers’ practice. The high stakes attached to PISA outcomes arguably diminishes local autonomy and school control over the curriculum and textbooks, and increases pressure on educators to teach to the test to the detriment of wider curricular goals. Some governments, for example Wales, have developed a suite of PISA professional development materials for use by primary and secondary schools, local authorities/school districts and regional consortia to align local practice more closely with PISA assessments (Welsh Government, 2019). Others have produced videos to inspire the national sample of students, for example *PISA 2012—Representing Your Country* produced by the Scottish Department of Education and anchored by the Scottish National Party Minister for Education (Lingard and Sellar, 2014).

PISA outcomes have consequences for teacher policies. PISA results are a political tool deployed selectively by national policy makers as a warrant to support a desired direction of change. For example, poor PISA outcomes were used as a resource to support the case for the Gonski teaching reforms in Australia, and comprehensive reviews of national systems for teacher education in Ireland and Wales (Gonski et al., 2018; Sahlberg et al., 2014; Furlong, 2015). PISA outcomes are also deployed as a weapon by Ministers to challenge and exert pressure on political opponents within electoral cycles. For example, in the UK, PISA 2012 outcomes in Labor-majority Wales were used by the then Conservative Education Secretary in England, Michael Gove, to attack Wales as “a country going backward” that was paying the price for abandoning performance tables and external accountability (Adams, 2013).

Media amplification of PISA outcomes seldom features scientific debate. The use of PISA outcomes within policy debates has been accompanied by calls for greater transparency and acknowledgment of limitations and biases. For example, recommendations based on analysis of PISA data routinely draw on correlational rather than causal relationships (Hopman et al., 2009; Araujo et al., 2017). The analysis uses cross-sectional rather than longitudinal assessment data. The representativeness of national samples is affected if countries exclude young people with special needs and newly arrived students. Returns are influenced by school and student threshold response rates, which must reach 80%. Interpretation of results requires consideration of the long-term development trends of participating countries. A number of studies have challenged the validity of standardized testing and rank ordering of

countries, drawing attention to cultural differences between nations, a failure to consider how teacher quality is defined, or the range of student outcomes that might be valued (personal, social and academic) including the wider goals of education.

Nevertheless, PISA outcomes are increasingly used as a proxy measure for teacher quality and the quality of teacher education, despite the limitations of equating student test scores with the efficacy of teacher preparation and continuing development (Auld and Morris, 2016; Baird et al., 2016). In the past two decades, universities in developed countries have grappled with reforms, both internally developed and externally mandated reforms, to improve their teacher preparation programmes (Cochran-Smith et al., 2018; Sleeter, 2014; Mayer et al., 2015; Rauschenberger et al., 2017). Concern with Initial Teacher Education (ITE) and teacher quality along with global push for measurable indicators of performance has led to a search for ways to measure ITE outcomes quantitatively. In the US, six state governments have used new teachers' value-added measures (VAM), based on their pupils' achievement on standardized tests, to assess the effectiveness of institutions' teacher graduates in the classroom (Coggsall et al., 2012). Interestingly, a number of these assessments have shown overall greater variation among graduates of the same program and relatively little variation in graduates' effectiveness across different ITE programmes (Koedel et al., 2015). Nevertheless, the push to measure quantitatively the quality of ITE programs persists and provides the opportunity for non-state actors to demonstrate how their methods might produce teachers who garner higher pupil test scores.

In 2008, the OECD extended its education data gathering by launching the Teaching and Learning International Survey (TALIS) in 24 countries (OECD, 2009). A second wave in 2013 included 34 countries and established a TALIS-PISA link. By 2018, the TALIS survey covered about 260,000 teachers in 15,000 schools across 48 countries. The survey asks lower-secondary teachers and school leaders about working conditions and learning environments, and offers comparative analysis to support national policymaking (OECD, 2020). Analysis from TALIS informs a series of thematic reports and working papers, and a range of advisory insight and lesson reports intended to guide in-country deliberation. A *Teacher's Guide* was introduced following TALIS 2013. TALIS findings are deployed as a powerful warrant framing deliberation on teacher policy. The European Commission's *Education and Training Monitor*, an annual evaluation of education and training system across Europe, draws heavily on TALIS findings. In charting the Europeanization of education policy, Grek (2014) has noted how the European Commission found in the OECD, "not only a great resource of data to govern (which it did not have before) but also a player who would be pushing the Commission's own policy agenda forward" (p. 267). Like PISA, TALIS findings are used to establish and repeat an evidence-informed and outwardly pragmatic agenda for change.

However, critical commentators identify a normative bias toward flexibilization of employment relations in TALIS (Sorenson and Roberston, 2018; Berkovich and Benoliel, 2018). Flexibilization entails changing work practices wherein employers may no longer use internal labor markets or guarantee long-term job security, and may incentivize, re-assign and redeploy employees more easily. By focusing on teacher performance, underpinned by a utilitarian, cognitive-economic model of education, critics suggest the TALIS survey carries an individualist and anti-collectivist agenda as it migrates across contexts. TALIS findings have been used to support performance management in teaching, preparing the way for the introduction of performance-pay by reform-minded Ministers and private school chains (Froese-Germain, 2010).

Labor market flexibilization also underpins the World Bank's promotion of incentive policies that link pay or tenure to performance (World Bank, 2013). The World Bank is the foremost proponent of a neoliberal education reform agenda. While the OECD was initially oriented toward advanced economies, the World Bank was first set up in 1944 to provide loans to developing countries that could not secure such finances otherwise. Based in Washington, D.C. and often working closely with the International Monetary Fund, efficiency gains are sought using neoliberal principles of deregulation, competitiveness, and privatization. Urgent reform is justified in terms of a race to maintain economic competitiveness in the developed economies of America, Australia and Europe, and to foster economic growth in low and middle-income countries. The World Bank provides loans for educational projects, as well as technical assistance and educational research at multiple levels.

By the 1980s, the Bank had come under intense criticism for its structural adjustment policies that required developing countries to enact privatization and free-market policies in its public sectors, including education, as a condition of their loan. In addition to undermining national sovereignty, these policies caused many governments to cut funding to education, leading to greater educational disparities and paving the way for private non-state actors to fill that vacuum in Africa, Asia, and Latin America. The World Bank, the Inter-American Development Bank, the Millennium Challenge Corporation and the Asia Development Bank all use PISA results in project appraisal. International assessment results are used to provide empirical evidence of the relative effectiveness and efficiency of targeted investment. For example, PISA results were used to evaluate projects supported by the Millennium Challenge Corporation in Georgia and the World Bank in Bulgaria, the Republic of Moldova and Kyrgyzstan (Lockheed et al., 2015).

The World Bank's SABER initiative (Systems Approach for Better Education Results), established in 2011, reflects its ambition to operate as a Knowledge Bank. The SABER-Teachers initiative collects, analyses, synthesizes and disseminates information on teacher policies around the world to support evidence-informed policymaking. SABER reports are used extensively by UNESCO's Global Education Monitoring Report and EFA Teacher Task Force, the Education Commission, PISA for Development (PISA-D), the World Food Program, US AID, and Results for Development (Klees et al., 2020). Teachers and workforce development are among the thirteen education topics benchmarked by SABER. The World Bank Group has consistently argued that "most teacher training is ineffective" and in-service training too often leads to "changes in teacher knowledge but not practice" (WB, 2018: 131–132). The SABER-Teachers initiative forms part of an advisory bloc that also includes the World Bank's Systems Approach for Better Education Results Engaging the Private Sector (SABER-EPS). Policy advocacy has centered on attracting the best (highly qualified) entrants, using compensation strategies to incentivize teacher performance e.g., performance related pay and flexible employment relations within autonomous self-managing schools (WB, 2015, 2018). In collaboration with UNESCO, the World Bank also provides

Edstats, an online portal that contains country-level data on a range of education topics including access, completion rates, and expenditures. Through these knowledge-related initiatives, along with its lending practices, the World Bank is able to shape countries' education-related policies toward its economic priorities.

While there are differences in historic orientation and mission between the supranational agencies described above, the use of opinion formation and coordinating activities has accelerated convergence in policy processes. System improvement is promoted through a policy mix of high stakes external accountability, public-private partnerships, and teacher incentives (Bruns et al., 2011). The following section outlines the growth of private providers in the delivery of educational services and the implications of this for teachers' lives and work.

The global education industry

From the 1980s onward, the growing liberalization of trade worldwide resulted in the partial conversion of education to a tradeable service. Public/state education sectors are now increasingly open to non-state actors providing educational goods and services, leading to the development of what has been termed the Global Education Industry (GEI) (Tooley, 2001). Verger et al. (2016) define the GEI as, "an increasingly globalized economic sector in which a broad range of educational services and goods are produced, exchanged and consumed, often on a for-profit basis" (p. 4). At the center of the field are "edu-businesses" (Ball, 2012). Verger et al. (2016, p. 12) categorized five types of edu-business, outlined below, to which corporate-funded private organizations and individual education entrepreneurs can be added.

- Chains of private schools (e.g., GEMS, ARK, Bridge International Academies, and the Omega schools). Several of these are low-fee private schools, such as Bridge International Academies, that operate in economically depressed areas and aim to generate a profit for investors through charging students' families "affordable" fees while reducing the running costs of schools. To cut costs, low-fee private schools rely on the use of technology, centralization of operations, employment of low-wage, uncertified teachers, and/or minimizing infrastructure/building costs. (See section "**Private school chains and teachers' work**" below)
- Large education corporations and conglomerates (e.g., Pearson) that offer a broad range of educational products and services. Included in this category are software companies (e.g., Microsoft, Intel, Cisco, Hewlett Packard and Blackboard) that provide technological infrastructure for digital learning and data tracking.
- Consultancy firms (e.g., PriceWaterhouseCoopers, McKinsey, and Boston Consulting Group) and individual consultants that apply business principles to education. Whether working with national governments, NGOs, or alone, educational management consultants and large elite consultancy firms generate data and analysis of educational problems and solutions rooted in economic perspectives. In reports co-authored by Michael Barber, McKinsey has been particularly influential in shaping debates on teachers and teacher education. McKinsey's reports have reinforced the centrality of the teacher in improving pupil outcomes, consistently emphasizing the importance of recruiting and training highly talented individuals into teaching and fostering leadership within and across the school system (Barber and Mourshed, 2007; Barber et al., 2010). Their reports have called on education systems to develop globally benchmarked standards, data transparency, and effective central management but operational autonomy for schools while maintaining a focus on addressing equity and community (Barber and Mourshed, 2009). While their reports shape education discourses and agendas, consultants also help to build public-private partnerships that create new links between educational stakeholders and private sector actors, contributing to the evolution of translational networks advocating neoliberal global education reform.
- Philanthropic foundations (e.g., Bill and Melinda Gates Foundation, the William and Flora Hewlett Foundation, Eli and Edythe Broad Foundation, the Walton Family Foundation, the George Soros's Open Society Foundation) which are separate from and autonomous of any corporate entity but are more or less aligned with the business vision and strategy of their founders and donors. These organizations fund a variety of initiatives including NGOs and charter/academy schools as well as leadership programs that recruit and train teachers (e.g., Teach For All) and school system leaders (e.g., Broad Superintendents Academy). Philanthropic foundations also fund research into proposed education reforms such as teacher accountability measures and performance-related pay. (See section "**Philanthropic foundations and teacher education**" below).
- Advocacy networks, including policy networks, that coalesce and evolve among a wide range of non-state actors as their interests converge around support for particular educational reforms such as school choice, low-fee private schools, and fast-track teacher training (e.g., Teach For All). While advocacy networks working at the transnational level include civil society coalitions such as the influential Global Campaign for Education and the Inter-Agency Network (Verger and Novelli, 2012; Macpherson, 2016), networks driven by the interests of private, non-state actors tend to be more exclusive in membership and less transparent in their activities. The influence of these networks is increasing in policy-making and service delivery.

Closely related to these categories are corporate-funded private organizations such as business-linked philanthropies, think tanks, lobbying firms, and private policy institutes. Through activities such as research, knowledge mobilization, lobbying, grassroots organizing, and networking, these private organizations are influential in leveraging corporate influence at various policy levels (Fontdevila and Verger, 2019; Fontdevila et al., 2019). The New Schools Venture Fund is an example of a highly influence private organization in the US that funnels private and corporate money into education start-ups (Zeichner and Pena-Sandoval, 2015).

In addition, the GEI is populated with individual policy entrepreneurs who advocate for particular educational ideas and act as boundary spanners who connect and work across organizations and sectors to promote education innovations (e.g., James Tooley,

Michael Barber, and Wendy Kopp). All these types of organizations and individuals work through transnational networks and interact with development agencies, NGOs, and governments at various levels (Ball and Olmedo, 2011; Robertson and Verger, 2012; Macpherson, 2016). It is common for actors within this field to have held various roles in different edu-businesses as well as in think tanks, academia, and/or government.

The diverse types of non-state actors described above are linked in the GEI through dense and complex inter-locking networks (Ball, 2012; Olmedo, 2014; Verger et al., 2016). These transnational networks are formed and sustained more or less formally through various national, regional, and global events and conferences such as the Qatar's Foundation World Innovation Summit Education (WISE), the OECD's Global Education Industry Summit, International Summit on the Teaching Profession co-hosted by the OECD and Education International, and the World Economic Forum's Davos Conference (Ball and Junemann, 2012). These events provide the opportunity for business leaders, investors, policy-makers and policy entrepreneurs among others to meet up, discuss, and socialize, furthering the knowledge diffusion and commitments that so often fuels policy transfer (Stone et al., 2020). Consequently, edu-preneurs and policy advocates are able to work at local, region, and national levels through the GEI to pitch their business and policy ideas to a broader audience of corporate, government, and philanthropic elites (Ball and Olmedo, 2011; Olmedo, 2013; Rönnberg, 2017; Giridharadas, 2019). Their involvement in education promotes quasi-markets in the sector, the standardization of teachers' work and pupils' achievements, and a loosening of the regulations that typically govern public schooling.

Development agencies, philanthropists, and blue-chip companies are closely tied to the growth of the GEI through their funding of new educational ventures. This "new" philanthropy or aid—also referred to as "venture philanthropy" (Olmedo, 2018) and "philanthropy-capitalism" (Bishop and Green, 2008)—advocates market-based solutions and incentives to providing public services, often for a profit, and is an approach that "intentionally blurs the line between business, enterprise, development and the public good" (Ball and Olmedo, 2011, p. 84). This mode of financing mirrors venture capitalism as investors opt to fund mainly small, early-stage, emerging organizations that are deemed to have high impact and/or profit potential that is scalable. Unlike earlier types of educational philanthropy, venture philanthropy bestows upon donors, or "investors", an influential role in educational policy and practice as they can closely control, monitor, and evaluate the use and impact of their investment to advance their strategic objectives. As a result, Saltman (2011) highlights that venture philanthropy has the "strategic aim of 'leveraging' private money to influence public schooling" and "is contributing to both the privatization of public schooling as well as the transformation of public schooling on the model of corporate culture" (p. 3). New philanthropy financing flows toward market-based educational reforms including vouchers, privately managed public schools, low-fee private school chains, and new types of teacher training.

In the next section, the effects of these developments on teachers are explored through two illustrative examples: private school chains and fast track teacher training.

Private school chains and teachers' work

There has been an explosive rise of private schools for mostly poor but also some middle class families (e.g., Weller, 2017) in developing countries—a development that has paralleled the expansion of charter, academy, and free schools in the US and UK. Low-fee private schools, which are independently owned, managed, and financed and charge what they deem affordable fees for the economically depressed communities in which they are located, have drawn criticism for their reliance on learning technologies, large class sizes, and centralized, standardized operations (including instruction and curriculum). Low-fee private schools are prevalent in sub-Saharan Africa (e.g., Malawi, Nigeria, Kenya and Ghana), South Asia (e.g., India and Pakistan) and Latin America (e.g., Peru and Central American countries). As many low-fee private school are not regulated, and hence not included in administrative records, the size of this sector is difficult to measure. Their impact on teachers has been twofold. First, teachers have seen their work change as chains of low-fee private schools seek to reduce cost by setting curriculum, learning tasks, and assessments centrally, while also utilizing online instruction to "individualize" learning without the need for a teacher. Second, many low-fee private schools (as well as charter/academy chains) hire unqualified and/or inexperienced teachers, opting instead to provide in-house training to orient teachers to their particular learning priorities and teaching practices.

One of the fastest growing chains of low-fee private schools is Bridge International Academies (BIA), a multinational chain founded in 2007 by three American entrepreneurs. Their Academy-in-a-Box model was designed to be replicable and scalable, and is based on a highly controlled learning environment with a large target class size (Kirchgasler, 2016). As Riep (2019) explains, "pre-programmed curriculum is developed by BIA at corporate headquarters abroad and then sent electronically to each school site using web-enabled smartphones that transfer curriculum to tablet e-readers, which is then read out verbatim, word-for-word, to students by unqualified staff referred to as 'Learning Facilitators'" (p. 7). The first BIA school opened in the informal settlement of Mukuru, an area of extreme poverty in Nairobi, Kenya.

BIA's initial funding came from investors in 2008 and 2009, which included Deutsche Bank America's Foundation, Gray Ghost Ventures, the Kellogg Foundation, and the Omidyar Network (established by billionaire eBay founder Pierre Omidyar) (Reip, 2017). In the following years, BIA has attracted private equity investments from a number of entrepreneurs in the educational technology sector including Bill Gates, Facebook founder Mark Zuckerberg, and Pearson PLC. The national governments of the US and UK have invested in the chain, raising concerns about the investment of public funds in for-profit education abroad. As the international list of funders grew, BIA expanded to other sub-Saharan countries as well as India. By 2016, Riep and Machacek (2016) estimate that BIA operated in "more than 520 for-profit schools throughout Kenya, Uganda, Liberia, Nigeria, and India, serving approximately 120,000 pupils" (p. 5). Despite its growth, the chain has yet to make a profit as its model suggests that it needs

to reach between 250,000 and 500,000 students to do so (Kwauk and Robinson, 2016). Private school chains, like BIA, are shaping teachers' work in the countries where they operate by hiring under- or unqualified teachers, paying them low wages, closely controlling and monitoring their work, and offering few, if any, employee protections.

Philanthropic foundations and teacher education

In addition to the employment and working conditions of teachers, the Global Education Industry increasingly contributes to the recruitment, selection and preparation of new teachers. While NGO involvement in teacher training is common in developing countries, the emergence of fast-track teacher training routes is on the rise. Among the most controversial and well established among these is Teach For America (TFA), founded in the US in 1990 by Wendy Kopp. The fast-track program aggressively recruits high-achieving graduates of elite universities to train for a summer and then serve as teachers in under-resourced schools for two years. The high-profile model has been replicated by non-state actors in more than 53 countries worldwide since 2007 with the help of the global Teach For All organization (Crawford-Garrett and Thomas, 2018; Thomas et al., 2020). Worldwide, the programs are funded by a mix of philanthropic foundations, corporations, and governments, claims to be helping raise pupil achievement in the short term while building a pipeline of leaders to education systems worldwide. Various national TFA models collaborate with university-based ITE providers in a relationship that has often been contentious due to the model's elite positioning and its overtly neoliberal view of teaching and learning (Meyers et al., 2014).

Backed by the World Bank and OECD, Teach For All's national affiliate programs also affect wider perceptions of teaching by portraying it as a temporary service commitment through which graduates can help "other" children in communities unlike their own while positioning themselves for success outside the classroom (Ahmann, 2015; Ellis et al., 2016). The network of Teach For All programs has also supported the expansion in privately run schools, whether for-profit chains in India or publicly-funded charter and academy school in the US and UK, as the program frequently places its recruits in such schools (Kretchmar et al., 2014). Furthermore, Teach For All programs promote the deregulation of ITE in the contexts in which they operate (Kretchmar et al., 2018) while creating a cadre of neo-liberal educational leaders who go on to lead privately-supported schools, establish new edu-business, and work in philanthropy or policy (Trujillo et al., 2017).

In the US, TFA and its alumni have recently helped spur the creation of a new generation of alternative teacher preparation organizations (Cochran-Smith et al., 2020). Whether tied to specific charter school chains, including so-called graduate schools of education (i.e., Match Teacher Residency, High Tech High Teacher Center) or stand-alone institutes (i.e., Relay GSE, Sposato GSE), these new private teacher training organizations are state-authorized to provide teacher certification and/or master's degrees and are funded by a mix of private and public funds (Zeichner and Pena-Sandoval, 2015; Mungal, 2016, 2019). Their approach to training is exclusively practice-centered and school-based with graduation often tied to the ability to raise pupil test scores. Relay GSE, the most high profile of the group of new providers, was created by three charter school chains: Knowledge Is Power Program, Achievement First, and Uncommon Schools. As Stützlein and West (2014) report, these charter school chains are "closely tied to major business interests, including foundations and investors such as J.P. Morgan, the Bill & Melinda Gates Foundation, and Credit Suisse, each of whom have financially supported Relay" (p. 3). In addition, Relay GSE has close links with Teach For America, who now annually funnels hundreds of its teachers to Relay, and the New Schools Venture Fund, a venture philanthropy organization that aggressively lobbies for, funds, and trains individuals to set up new entrepreneurial teacher preparation programs (Zeichner and Pena-Sandoval, 2015).

The development of standalone GSE closely tied to charter schools in the US inspired a similar effort in the UK, with the establishment of the Institute for Teaching. Officially launched in 2017 with funding from government, hedge fund managers and philanthropic foundations, the Institute for Teaching was set up to be a "teacher-led" (rather than academic-led) and "practice-orientated" professional preparation program for new and in-service teachers (Ellis et al., 2018). Although the Institute merged in 2018 with Ambition School Leadership to form the Ambition Institute, in its new form it still represents an alternative to university-based teacher education. Both trends—tied to the Teach For All network—illustrate the interconnectedness of non-actors and their international networks. These networks advocate for deregulation, market-based solutions, and new conceptions of teacher education that primarily rely on a narrow view of pedagogy, an "impoverished" conception of experience (Ellis, 2010) and deficit thinking toward children in economically-depressed communities (Pitzer, 2014) while equating higher pupil standardized test scores with providing a high-quality education.

In addition to funding alternatives to university-based teacher education, private and corporate philanthropy has changed the educational landscape and teachers' work in a myriad of other ways. Foundation funding has promoted school choice and led to the establishment of charter schools, academies, and low-fee private schools in high-poverty communities across the US, UK and developing world respectively. Many of these schools have promoted a variant of the "no-excuses" ethos, narrow curriculum, and teaching by unqualified and/or inexperienced teachers. Philanthropists and private foundations have also provided start-up funding to edu-businesses, which shape the technology, assessments, and training teachers increasingly rely on.

Finally, foundations have a clear impact on teachers through shaping the research that informs the profession and policy debates pertaining to it. For example, starting in 2009, the Gates Foundation turned its attention to the issue of teacher quality and funded two large projects—the Measures of Effective Teaching (MET) study and Intensive Partnerships for Effective Teaching initiative. The MET project (2009–2012) involved multiple independent research teams and nearly 3000 teachers from seven school districts, and sought to find a reliable and replicable measure of teachers' effectiveness. Ultimately, researchers on the study declared teacher effectiveness measures should combine three types of measures: classroom observations, student surveys, and student achievement gains

(Kane et al., 2013). Designed to draw on the findings from the MET study, the \$212 million-dollar Intensive Partnerships for Effective Teaching initiative (2009–2016) involved three public school districts and four charter management organizations experimenting with strategies aimed to produce dramatic gains in the achievement among pupils in low-income schools by improving the teacher workforce. The strategies—which related to measures of effective teaching, staffing policies, evaluation-linked professional development, and compensation and career ladders—were ultimately unsuccessful as the evaluators of the study found no overall increase in student achievement levels or graduation rates (Stecher et al., 2018). Nevertheless, the Gates Foundation remains an influential actor in educational policy debates, particularly in respect to teacher quality (Fejerskov, 2015; McGoe, 2016).

Conclusion

In sum, the types of non-state actors making up the GEI are many—ranging from edu-businesses, policy entrepreneurs, and multinational companies to philanthropists, foundations, and think tanks—and their impact on schooling and teachers' work is becoming increasingly apparent. The effectiveness of university-led teachers' professional preparation and development is being questioned, scrutinized, and in some cases de-regulated to create a quasi-market in which new, private providers can offer their services. Such private teacher training organizations may offer preparation that focuses on teaching practices and classroom experience based on what has worked well in raising pupils' test scores. Other non-state actors of the GEI such as low-fee private schools are opting to hire less experienced or unqualified teachers to deliver scripted (teacher proof) lessons in an effort to keep costs low and fidelity with the prescribed curriculum high. In such for-profit models of schooling, students and their families become customers and teachers' performance is increasingly linked to students' satisfaction and progress on national standardized tests. This trend is growing in school systems in developed countries. For example, a number of policy-makers across the US implement value-added measures through which teachers' performance is linked to their pupils' test scores on standardized tests. The rationale for value-added measures is linked to teacher effectiveness research and the need to hold teachers and, in some instances their preparation program, accountable for measurable learning gains. The growth in standardization and performance measures, linked to the global rhetoric in education fueled by the OECD, UNESCO, and World Bank among others, provides governments with the levers through which to steer and reshape the teaching profession and open the sector further to private non-state actors that claim to be able to meet performance targets more efficiently. Paradoxically, these claims often entail the expanded use of technology, online learning, and universally applicable pedagogies, while curtailing the pedagogical authority, autonomy, pay, and professional status of teachers whom they simultaneously claim are the key driver of learning.

References

- Adams, R., 2013. Gove stands by school reforms despite Britain's failure to close gap with Asia: change needed to improve UK rankings, says minister. *Guardian*. Available at: <https://www.theguardian.com/politics/2013/dec/03/gove-defends-education-reforms>. (Accessed 22 August 2020).
- Addy, C., 2017. Golden relics & historical standards: how the OECD is expanding global education governance through PISA for development. *Crit. Stud. Educ.* 58 (3), 311–325.
- Ahmann, C., 2015. Teach for all: storytelling "shared solutions" and scaling global reform. *Educ. Policy Anal. Arch.* 23 (45), 1–27.
- Araujo, L., Saltelli, A., Schnepf, S., 2017. Do PISA data justify PISA-based education policy? *Int. J. Comp. Educ. Dev.* 19 (1), 20–34.
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education "best practice". *Comp. Educ.* 52 (2), 202–239.
- Auld, E., Rappleye, J., Morris, P., 2019. PISA for development: how the OECD and World Bank shaped education governance post-2015. *Comp. Educ.* 55 (2), 197–219.
- Baird, J.A., Johnson, S., Hopfenbeck, T.N., Isaacs, T., Sprague, T., Stobart, G., Yu, G., 2016. On the supranational spell of PISA in policy. *Educ. Res.* 58 (2), 121–138.
- Ball, S.J., Junemann, C., 2012. *Networks, New Governance and Education*. Policy Press, Bristol.
- Ball, S.J., Olmedo, A., 2011. Global social capitalism: using enterprise to solve the problems of the world. *Citizenship Soc. Econ. Educ.* 10 (2–3), 83–90.
- Ball, S.J., 2012. *Global Education Inc.: New Policy Networks and the Neoliberal Imaginary*. Routledge, Abingdon, Oxon.
- Barber, M., Mourshed, M., 2007. *How the World's Best-Performing School Systems Come Out on Top*. McKinsey & Company, London.
- Barber, M., Mourshed, M., 2009. *Shaping the Future: How Good Education Systems Can Become Great in the Decade Ahead*. McKinsey & Company, London.
- Barber, M., Whelan, F., Clark, M., 2010. *Capturing the Leadership Premium: How the World's Top School Systems are Building Leadership Capacity for the Future*. McKinsey & Company, London.
- Berkovich, I., Benoliel, P., 2018. Marketing teacher quality: critical discourse analysis of OECD documents on effective teaching and TALIS. *Crit. Stud. Educ.* <https://doi.org/10.1080/17508487.2018.1521338>.
- Bishop, M., Green, M., 2008. *Philanthrocapitalism: How the Rich Can Save the World*. Bloomsbury Press, New York.
- Bruns, B., Filmer, D., Patrinos, H.A., 2011. *Making Schools Work: New Evidence on Accountability Reforms*. World Bank, Washington DC.
- Cochran-Smith, M., Carney, M., Keefe, E., Burton, S., Chang, W., Fernández, M., Miller, A., Sánchez, J., Baker, M., 2018. *Reclaiming Accountability in Teacher Education*. Teachers College Press, New York.
- Cochran-Smith, M., Keefe, E.S., Carney, M.C., Olivo, M., Smith, R.J., 2020. Teacher preparation at new graduate schools of education: studying a controversial innovation. *Teach. Educ. Q.* 47 (2), 8–37.
- Coggshall, J.G., Bivona, L., Reschly, D.J., 2012. Evaluating the Effectiveness of Teacher Preparation Programs for Support and Accountability. Research & Policy Brief. National Comprehensive Center for Teacher Quality, Washington, D.C.
- Crawford-Garrett, K., Thomas, M.A., 2018. Teacher education and the global impact of teach for all. In: *Oxford Research Encyclopaedia of Education*. Available at: <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-417>. (Accessed 20 August 2020).
- Ellis, V., Maguire, M., Trippstad, T., Liu, Y., Yang, X., Zeichner, K., 2016. Teaching other people's children, elsewhere, for a while: the rhetoric of a travelling educational reform. *J. Educ. Pol.* 31 (1), 60–80.
- Ellis, V., Steadman, S., Trippstad, T.A., 2018. Teacher education and the GERM: policy entrepreneurship, disruptive innovation and the rhetorics of reform. *Educ. Rev.* 71 (1), 1–21.
- Ellis, V., 2010. Impoverishing experience: the problem of teacher education in England. *J. Educ. Teach.* 36 (1), 105–120.

- Fejerskov, A.M., 2015. From unconventional to ordinary? The Bill and Melinda Gates foundation and the homogenizing effects of international development cooperation. *J. Int. Dev.* 27 (7), 1098–1112.
- Fontdevila, C., Verger, A., 2019. The political turn of corporate influence in education: a synthesis of main policy reform strategies. In: *Researching the Global Education Industry*. Springer, pp. 47–68.
- Fontdevila, C., Verger, A., Avelar, M., 2019. The business of policy: a review of the corporate sector's emerging strategies in the promotion of education reform. *Crit. Stud. Educ.* 1–16.
- Froese-Germain, B., 2010. The OECD, PISA and the Impacts on Educational Policy. Canadian Teachers' Federation, Ottawa, Canada.
- Furlong, J., 2015. *Teaching Tomorrow's Teachers. Options for the Future of Initial Teacher Education in Wales*. University of Oxford Department of Education, Oxford.
- Giridharadas, A., 2019. *Winners Take All: The Elite Charade of Changing the World*. Allen Lane, UK.
- Gonski, D., Arcus, T., Boston, K., Gould, V., Johnson, W., O'Brien, L., Perry, L., Roberts, M., 2018. *Through growth to achievement: the report of the review to achieve educational excellence in Australian schools*. Commonwealth of Australia, Canberra.
- Grek, S., 2014. OECD as a site of coproduction: European education governance and the new politics of "policy mobilization". *Crit. Pol. Stud.* 8 (3), 266–281.
- Hopmann, S., Brinek, G., Retzl, M., 2009. *PISA According to PISA—Does PISA Keep What It Promises?* University of Vienna Press, Vienna.
- Kamens, D.H., 2015. A maturing global testing regime meets the world economy: test scores and economic growth, 1960–2012. *Comp. Educ. Rev.* 59 (3), 420–446.
- Kane, T.J., McCaffrey, D., Miller, T., Staiger, D., 2013. Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment, Research Paper. MET Project. Bill & Melinda Gates Foundation.
- Kirchgasler, C., 2016. The limits of "knowledge for all": historicizing transnational school reforms in Kenya. *Knowl. Cult.* 4 (2), 73–94.
- Klees, S.J., Ginsburg, M., Anwar, H., Baker Robbins, M., Bloom, H., Busacca, C., Corwith, A., DeCoster, B., Fiore, A., Gasior, S., Le, H., Primo, L., Reedy, T., 2020. The World Bank's SABER: a critical analysis. *Comp. Educ. Rev.* 64 (1), 46–65.
- Koedel, C., Parsons, E., Podgursky, M., Ehrlert, M., 2015. Teacher preparation programs and teacher quality: are there real differences across programs? *Educ. Finance Policy* 10 (4), 508–534.
- Kretchmar, K., Sondel, B., Ferrare, J.J., 2014. Mapping the terrain: teach for America, charter school reform, and corporate sponsorship. *J. Educ. Pol.* 29 (6), 742–759.
- Kretchmar, K., Sondel, B., Ferrare, J.J., 2018. The power of the network: teach for America's impact on the deregulation of teacher education. *Educ. Pol.* 32 (3), 423–453.
- Kwaak, C., Robinson, J.P., 2016. *Bridge International Academies: Delivering Quality Education at a Low Cost in Kenya, Nigeria, and Uganda*. The Brookings Institution, Washington, D.C.
- Lewis, S., Sellar, S., Lingard, B., 2016. PISA for schools: topological rationality and new spaces of the OECD's global educational governance. *Comp. Educ. Rev.* 60 (1), 27–57.
- Lewis, S., 2020. Providing a platform for "what works": platform-based governance and the reshaping of teacher learning through the OECD's PISA4U. *Comp. Educ.* <https://doi.org/10.1080/03050068.2020.1769926>.
- Lingard, B., Sellar, S., 2014. Representing for country: Scotland, PISA and new spatialities of educational governance. *Scot. Educ. Rev.* 46 (1), 5–18.
- Lockheed, M., Prokic-Bruer, T., Shadrova, A., 2015. How have PISA results informed education policy discussions and affected policy in middle-income countries?. In: *The Experience of Middle-Income Countries Participating in PISA 2000–2015*. The World Bank, Washington, D.C./OECD Publishing, Paris.
- Macpherson, I., 2016. An analysis of power in transnational advocacy networks in education. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. John Wiley & Sons, Chichester, pp. 401–418.
- Mayer, D., Allard, A., Bates, R., Dixon, M., Doecke, B., Kline, J., Kostogriz, A., Moss, J., Rowan, L., Walker-Gibbs, B., White, S., Hodder, P., 2015. *Studying the Effectiveness of Teacher Education—Final Report*. Deakin University. Available at: http://www.setearc.com.au/wp-content/uploads/2013/08/SETE_report_FINAL_30.11.152.pdf. (Accessed 18 August 2020).
- McGoey, L., 2016. *No Such Thing as a Free Gift: The Gates Foundation and the Price of Philanthropy*, reprint ed. Verso Books, London.
- Meyers, B., Fisher, T., Bloxson, K., 2014. Unfinishedness: striving for a viable partnership between TFA and its university partner. *Teach. Coll. Rec.* 116 (10), 1–32.
- Michel, A., 2017. The contribution of PISA to the convergence of education policies in Europe. *Eur. J. Educ.* 52, 206–216.
- Mungal, A.S., 2016. Teach for America, Relay Graduate School, and charter school networks: the making of a parallel education structure. *Educ. Pol. Anal. Arch.* 24 (17), 1–30.
- Mungal, A.S., 2019. The emergence of relay graduate school. *Issues Teach. Educ.* 28 (1), 52–79.
- OECD, 2005. *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. OECD, Paris.
- OECD, 2009. *Creating Effective Teaching and Learning Environments. First Results from TALIS*. OECD, Paris.
- OECD, 2015. *Education Policy Outlook 2015: Making Reforms Happen*. OECD, Paris.
- OECD, 2018. *Effective Teacher Policies: Insights from PISA*. OECD, Paris.
- OECD, 2020. *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*. TALIS. OECD Publishing, Paris. <https://doi.org/10.1787/19cf08df-en>.
- Olmedo, A., 2013. Policy-makers, market advocates and edu-businesses: new and renewed players in the Spanish education policy arena. *J. Educ. Pol.* 28 (1), 55–76.
- Olmedo, A., 2014. From England with love... ARK, heterarchies and global "philanthropic governance". *J. Educ. Pol.* 29 (5), 575–597.
- Olmedo, A., 2018. Venture philanthropy and education policy-making: charity, profit, and the so-called "democratic state". In: *The Wiley Handbook of Global Educational Reform*. Wiley Online Library, pp. 47–69.
- Pitzer, H., 2014. The authority of experience, deficit discourse and Teach for America: the risks for urban education. In: Ellis, V., Orchard, J. (Eds.), *Learning Teaching from Experience: Multiple Perspectives and International Contexts*. Bloomsbury, pp. 127–142.
- Rauschenberger, E.A., Adams, P., Kennedy, A., 2017. *Measuring Quality in Initial Teacher Education (MQUTE) Literature Review*. Scottish Council of Deans of Education, Edinburgh.
- Riep, C., Machacek, M., 2016. Schooling the poor profitably: the innovations and deprivations of Bridge International Academies in Uganda. *Educ. Int.* Available at: <https://www.right-to-education.org/resource/schooling-poor-profitably-innovations-and-deprivations-bridge-international-academies>. (Accessed 20 August 2020).
- Riep, C.B., 2017. Making markets for low-cost schooling: the devices and investments behind Bridge International Academies. *Glob. Soc. Educ.* 15 (3), 352–366.
- Riep, C., 2019. What Do We Really Know About Bridge International Academies? Education International, Brussels, Belgium, pp. 1–24.
- Robertson, S., Verger, A., 2012. Governing education through public private partnerships. In: Robertson, S., Mundy, K., Verger, A., Menashy, F. (Eds.), *Public Private Partnerships in Education: New Actors and Modes of Governance in a Globalizing World*. Edward Elgar Publishing, Cheltenham, UK, pp. 21–42.
- Rönnerberg, L., 2017. From national policy-making to global edu-business: Swedish edu-preneurs on the move. *J. Educ. Pol.* 32 (2), 234–249.
- Sahlberg, P., Broadfoot, P., Coolahan, J., Furlong, J., Kirk, G., 2014. *Aspiring to Excellence. Final Report of the International Review Panel on the Structure of Initial Teacher Education in Northern Ireland*. DEL, Belfast.
- Sahlberg, P., 2011. *Finnish Lessons. What Can the World Learn from Educational Change in Finland*. Teachers College Press, New York.
- Saltman, K.J., 2011. From Carnegie to Gates: the Bill and Melinda Gates Foundation and the venture philanthropy agenda for public education. In: Kovacs, P.E. (Ed.), *The Gates Foundation and the Future of U.S. "Public" Schools*. Routledge, Abingdon, Oxon, UK, pp. 1–20.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. *Br. Educ. Res. J.* 40 (6), 917–936.
- Sleeter, C., 2014. Toward teacher education research that informs policy. *Educ. Res.* 43 (3), 146–153.
- Sorenson, T.B., Robertson, S., 2018. The OECD program TALIS and framing, measuring, and selling quality teacher. In: Akiba, M., Le Tendre, G.K. (Eds.), *International Handbook of Teacher Quality and Policy*. Routledge, London, pp. 117–131.
- Stecher, B.M., Holtzman, D., Garet, M., Hamilton, L., Engberg, J., Steiner, E., Robyn, A., Baird, M., Gutierrez, I., Peet, E., Brodzia de los Reyes, I., Fronberg, K., Weinberger, G., Hunter, G., Chambers, J., 2018. *Improving Teaching Effectiveness: Final Report: The Intensive Partnerships for Effective Teaching Through 2015–2016*. RAND Corporation, Santa Monica, CA.
- Stitzlein, S.M., West, C.K., 2014. New forms of teacher education: connections to charter schools and their approaches. *Democr. Educ.* 22 (2), 1–10.

- Stone, D., Oliveira, O. P. de, Pal, L.A., 2020. Transnational policy transfer: the circulation of ideas, power and development models. *Policy Soc.* 39 (1), 1–18.
- Thomas, M.A.M., Rauschenberger, E., Crawford-Garrett, K. (Eds.), 2020. *Examining Teach for All: International Perspectives on a Growing Global Network*. Routledge, London.
- Tooley, J., 2001. *The Global Education Industry Lessons from Private Education in Developing Countries*, second ed. Institute of Economic Affairs, London.
- Trujillo, T., Scott, J., Rivera, M., 2017. Follow the yellow brick road: teach for America and the making of educational leaders. *Am. J. Educ.* 123 (3), 353–391.
- Vaccari, V., Gardinier, M.P., 2019. Toward one world or many? A comparative analysis of OECD and UNESCO global education policy documents. *Int. J. Dev. Educ. Global Learn.* 11 (1), 68–86.
- Verger, A., Novelli, M., 2012. *Campaigning for "Education for All": Histories, Strategies and Outcomes of Transnational Advocacy Coalitions in Education*. Springer Science, Rotterdam.
- Verger, A., Lubienski, C., Steiner-Khamsi, G., 2016. The emergence and structuring of the global education industry: towards an analytical framework. In: Verger, A., Lubienski, C., Steiner-Khamsi, G. (Eds.), *World Yearbook of Education 2016: The Global Education Industry*. Routledge, New York, pp. 3–24.
- Weller, C., 2017. A Peruvian billionaire contracted a world-famous design firm to remake his country's private school system, and the results are stunning. *Business Insider*. Available at: <https://www.businessinsider.com/innova-schools-peru-education-2017-7>. (Accessed 17 August 2020).
- Welsh Government, 2019. Programme for International Student Assessment (PISA): Training Pack. Available at: <https://www.hwb.gov.wales/repository/resource/5d56a272-3c97-40d8-8c0b-8f964819a0e2/en>. (Accessed 17 August 2020).
- World Bank, 2013. What Matters Most in Teacher Policies? A Framework Paper. World Bank, Washington, DC. Available at: http://wbfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc/Background/TCH/Framework_SABER-Teachers.pdf. (Accessed 17 August 2020).
- World Bank, 2015. The Rise of Results-Based Financing in Education. World Bank, Washington, DC. Available at: https://www.worldbank.org/content/dam/Worldbank/Brief/Education/RBF_ResultsBasedFinancing_v9_web.pdf. (Accessed 16 August 2020).
- World Bank, 2018. World Development Report 2018: Learning to Realize Education's Promise. World Bank, Washington, DC. <https://doi.org/10.1596/978-1-4648-1096-1>. Available at: file://staffhome/staff_home0/55122068/Documents/Writing/World%20Bank%202018.pdf. (Accessed 20 August 2020).
- Zapp, M., 2020. The authority of science and the legitimacy of international organisations: OECD, UNESCO and World Bank in global education governance. *Compare* 1–20. <https://doi.org/10.1080/03057925.2019.1702503>.
- Zeichner, K., Pena-Sandoval, C., 2015. Venture philanthropy and teacher education policy in the US: the role of the New Schools Venture Fund. *Teach. Coll. Rec.* 117 (6), 1–24.

Teacher education policy and professionalism: a personal review of teacher education policy research

Elaine Vanassche, Faculty of Psychology and Educational Sciences, KU Leuven, Kortrijk, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	10
Background and basic elements of critical approaches to teacher education policy analysis	11
The current teacher education	12
Teacher education is in crisis and in need of improvement	12
Teacher education should be held accountable and wants to hold itself accountable	13
Teacher education is evidence-based	13
Teacher education is practical	14
The current teacher	15
Teachers know and perform “what works”	15
Teachers are responsive and flexible	16
What’s next?	17
References	18

Introduction

This is an interesting timeframe to write about teacher education for the *International Encyclopedia of Education*. Governments globally show an increasing aspiration to intervene in order to have greater influence, if not control, over the structure, content, and delivery of pre-service or initial teacher education (Furlong et al., 2009). The aspiration to intervene is supported by the consensus, at least in policy, that the quality of an education system largely depends on the quality of teachers and their preparation and credentials (Cochran-Smith et al., 2018). Historical analyses internationally show that teacher education has always been of key strategic significance in attempts to bring about change to teachers and teaching. Since the inception of formal systems of teacher preparation, initial teacher education has led the charge to reform and transform education systems, as the work of Rowe and Skoudombis (2019) on Australia, Grimmer (2009) on Canada, Assunção Flores (2011) on Portugal, Lewin et al. (2003) on South Africa, and Zeichner (2018) on the USA shows, to offer just a few examples.

What is new about the changes in teacher education in recent decades though is what Murray et al. (2019) called “a systematic politicization of teacher education” (p. 2). A main driver for this process is the progressively complex challenges encountered in the current, globalized world which increase governments’ aspiration to intervene and regulate teacher education in many countries across the developed and developing worlds. These challenges range from the global knowledge-based economy to the technologic revolution now entering all aspects of daily life, political instability in many regions, changing migration flows, and the COVID-19 pandemic which has radically challenged existing conceptions of schooling and teaching. Government intervention in specific localized contexts is also increasingly shaped by global and international influencers such as large-scale student assessments (e.g., the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS)) which are often cited as research evidence of the quality of an education system (Mayer, 2021). Such international comparative metrics spur domestic policy reform, especially but not exclusively by those at the bottom of the rankings who borrow models and policies from top performing jurisdictions (Grek, 2009). In short, the logic underpinning current teacher education policy is that government regulation and constant and rapid reform of teacher education is necessary in these wider uncertain times for an education system and a nation as a whole to stay internationally competitive, on top of the game, and in the market.

As major reforms in teacher education became more internationally visible in the 1990s and 2000s, the field of teacher education policy analysis grew steadily. The first generation of policy researchers in the 1970s and 1980s focused almost exclusively on schools and higher education (Steiner-Khamsi, 1999). The realization that these were not the only sites where changes were felt, invited research looking specifically at teacher education policy developments in the past two decades. Some of this work adopts a functional or historical lens to explain the development and nature of teacher education systems by their function (e.g., Ginsburg, 1998). This work is not the focus of this article. This article will rather focus on teacher education policy research and scholarship which takes a more critical approach and shares a view that “teacher education has been justified in different ways at different times” (Maguire, 2000, p. 228). These policy researchers share concern that reform not only shapes and changes the structures of teacher education, but also produces and reproduces particular normative discourses around the roles and purposes of teacher education and what it means to be a teacher which are not adequately addressed in historical-chronological accounts of teacher education (Maguire, 2000). The purpose of this article is to synthesize and examine what can be learned from this critical teacher education policy research during the past 15–20 years. The review of the literature is structured around two questions: (1) What is “the current teacher education” being constructed and reconstructed in teacher education policy legislation globally?; (2) What is “the current teacher” being

constructed and reconstructed in teacher education policy legislation globally? The current teacher education is a play on Cochran-Smith's (2005) description of "the new teacher education" in her AERA presidential address in 2005 with a particular focus on the USA. This article provides a much-needed update and expansion of her work 17 years after date.

Before embarking on this work, two important remarks need to be made. First, although the focus is on (re)constructing the current teacher education and the current teacher enacted in teacher education policy and reform, the analysis contains references to both "the old" and "the new" teacher education and teacher. The latter is taken up explicitly in the closing section of the article which builds an agenda for future research. Second, despite the goal of the article to identify and synthesize common patterns and trends across countries and institutions, I acknowledge that teacher education policies and practices are locally, historically, and culturally embedded. It is not my intention to gloss over these local variations. Rather, this article adopts what could be called a "global approach" to the review in that it attempts to identify common threads and global patterns, but with references to local research showing how these global trends have played out differently across various country and institutional settings where relevant and necessary.

The article starts with discussing in a broad way critical teacher education policy research, describing some of the assumptions and goals shared by those who engage in it. It is included to provide the general contours of this scholarship as I have come to understand it, as a way of orienting the reader to what follows.

Background and basic elements of critical approaches to teacher education policy analysis

What I labeled as "critical teacher education policy research" in the opening section of this article is not a field with marked boundaries. It is indeed a label that researchers operating within this field do not use themselves to identify their own scholarship. It is my personal and perhaps idiosyncratic view of an emerging and expanding subfield of education policy research, displaying a variety of theoretical and methodological stances and considerations. What unifies this work is first and foremost the understanding that schools and higher education are not the only places where changes such as standardization, decentralization, globalization, and marketization are felt (Steiner-Khamsi, 1999). These changes cut across all policy sectors (health, education, social and care services) and levels of education, including teacher education. Second, this work starts from the idea that policymaking is not "a deliberate process, undertaken by a bounded set of actors, who use research and reason to ensure the best possible policy outcomes" (Young and Diem, 2017, p. 2). Rather, it considers policymaking as a rhetoric process concerned with negotiation, contestation, power, and ideology. Policy problems are not self-evident facts awaiting government action, but social constructions or problematizations (Bacchi, 2010), and thus contestable and open to change. These researchers engage in research "of" policy rather than research "for" policy, borrowing Moore's (1996) binary. "The former positions the standpoint of the field, the latter is positioned by it" (Moore, 1996, p. 159).

In their review of critical policy work in education, Young et al. (2019) have found that researchers tend to focus their work around five critical concerns. It could be argued that these concerns also typify critical approaches to teacher education policy analysis:

- Concern regarding the difference between policy rhetoric and practiced reality
- Concern regarding the policy, its roots, and its development (e.g., how it emerged, what problems it was intended to solve, how it changed and developed over time, and its role in reinforcing the dominant culture)
- Concern with the distribution of power, resources, and knowledge as well as the creation of policy "winners" and "losers"
- Concern regarding social stratification and the broader effect a given policy has on relationships of inequality and privilege
- Concern regarding the nature of resistance to or engagement in policy by members of nondominant groups (Young et al., 2019, p. 6).

The history of critical approaches to teacher education policy analysis is, compared to policy research focusing on schools and universities, considerably short. To the best of my knowledge, there are no reviews of this work, nor systematic attempts to define "what counts as critical teacher education policy research." This scholarship has adopted a variety of perspectives and approaches. Examples include Phelan and Morris' (2021) Foucauldian analysis of policy texts in Australia; McGarr and Emstad's (2020) use of critical discourse analysis to offer a comparative perspective to teacher education reform mandates in Ireland and Norway; or Grimmett's (2009) macro-political perspective to examine the macro-political pressures at work in Canadian teacher education. Other examples include Snoek's (2021) use of a narrative approach to examine the conception of teacher quality in the Dutch policy context; Lewis and Young's (2013) utilization of Kingdon's Multiple-Streams Framework to examine how teacher education accountability has gained prominence on the governmental decision agenda in the USA; Furlong and O'Brien's (2019) employment of a phenomenological approach to decode the teacher educator identities constructed within an Irish context of reform; and Vanasche et al.'s (2021) positioning theory analysis of the normative discourses on teacher professionalism enacted in the latest teacher education reform in Flanders (Belgium). This non-exhaustive overview shows that the emerging field of critical teacher education policy research displays the same theoretical-philosophical and methodological diversity as reviews of critical education policy research more broadly have demonstrated (e.g., Ball, 1994, 1997; Young and Diem, 2017; Young et al., 2019).

What this scholarship does very well is also one of its pitfalls. That is, it pays significant attention to the local contexts and systems in which policy is made, favoring small-scale designs, and working from a national (or regional) framework of analysis, but often fails to reach beyond those local contexts. There have been attempts to adopt an international comparative perspective

with an increasingly broad geographic range (e.g., Adams and Tulasiewicz, 1995; Brock, 1996; Furlong et al., 2009; Darling-Hammond and Lieberman, 2012; Murray et al., 2019; Symeonidis, 2021; Mayer et al., 2021). However, most of this comparative work relies on what Steiner-Khamsi (1999) called implicit comparison methodology by which she means that teacher education reform in one national context is used as an analytical framework, and other reforms are contrasted with these developments; or analyses of reforms in different national contexts are presented together in special issues or edited books, as if the diversity and contrast displayed across individual contributions speak for itself. While there is a necessity for and great value in such efforts, I share Ball's (1997) concern that such policy research is "lacking a sense of 'place'" (p. 267) by which he referred to the lack of efforts to locate educational policies in frameworks that extend beyond the local level. Ball suggests a more direct analysis of the flow of policies between nations and related to that, the mediation of the global by the local pressures, concerns, and preferences, and *vice versa*. As he argues, "this is not simply a point about empirical description, although more of that would be very welcome, it is also about theorization" (p. 267).

This article is a first attempt to start identifying global trends in the normative and performative policy discourses on teacher education and what it means to be a teacher today, drawing on critical teacher education policy research in recent years across countries and institutional contexts. It offers my read of this expanding field of research. I will contend that the current teacher education is constructed as in crisis and in need of improvement, preoccupied with evidence and research, practical, and (willing to hold itself) accountable. This, in turn, invites related notions of teacher professionalism as essentially and frontally about knowing and performing "what works," being flexible and responsive to changing demands.

This configuration of the current teacher education and the current teacher is not exhaustive, nor do I claim that it characterizes any one teacher education policy setting or program globally. Rather consider this as a living configuration which focuses on common threads and patterns in the field that take different shapes and forms when enacted in specific localized contexts.

The current teacher education

Teacher education is in crisis and in need of improvement

The current teacher education is being constructed as "at risk," "in crisis," and "in need of improvement." This is not a new discourse. Historical analyses show that teacher education has always been assumed to be in some sort of crisis. Overlooking reform dynamics in ten countries, Tatto and Plank (2007), for example, describe teacher education institutions as operating in "an environment of permanent educational crisis" (p. 269). Beauchamp et al. (2015) similarly describe teacher education as always in turbulence, marked by contest and challenge, and receiving an insistent critique. The perception of crisis has become such a defining characteristic of teacher education because of three seemingly irrefutable assumptions. First, teachers are assumed to be the single most important factor in student achievement and learning (Delandshere and Petrosky, 2004). Second, student achievement is seen as a predictor of "individual and collective productivity and wealth in the economic argument of the Knowledge Economy" (Välimaa et al., 2016, p. 19). Third, the quality of teachers is assumed to be shaped by the preparation and credentials they received. In this permanent contest for comparative advantage, "it is no longer possible for school systems to be 'good enough'" (Tatto and Plank, 2007, p. 267). In a similar way, it is no longer possible for the current teacher education to be "good enough" and deliver "good enough" teachers.

The induced sense of crisis legitimizes governments' rights to intervene in the system in crisis (Trippstad et al., 2017). The only way to deal with (the threat of) crisis is to take powerful action and speed up the reform of teacher education, as is seen in many countries, "reflecting a belief that the creation of policy in and of itself suggests order, authority and expertise" (Clarke and Phelan, 2015, p. 258). Recent policy analyses of, among others, Simpson et al. (2021) in Australia, Vanassche et al. (2019) in England, Meijer (2021) in the Netherlands, and Cochran-Smith (2020) in the USA seem to suggest that the efforts for change in teacher education are increasingly being directed at both the structure and the substance of teacher education. This includes, for example, program accreditation criteria and auditing procedures, allowable routes into the profession, but also subject knowledge, pedagogical/didactical and fieldwork requirements, and prespecified outcomes (e.g., teacher standards). The current teacher education not only needs structural change and improvement, but also more substantial improvements to the focus, content, and delivery of its curriculum. The imperative for such substantive changes is built around the accelerated pace of reform in the current, globalized world assumed to pose qualitatively new demands on all levels of education, teacher education included (e.g., Madalińska-Michalak et al., 2018).

While the literature does not deny that teacher education stands in need of improvement, it also suggests that the image of crisis is in part a rhetorical or manufactured one (Clarke and Phelan, 2015), involving persistent questioning from policymakers and public media, often driven by international metrics which are taken as solid evidence of how well systems are performing. Nordin (2014), for example, demonstrated how this crisis discourse has been used by both national and transnational actors in the EU and in Sweden seeking public legitimacy for extensive reforms. Drawing on Beck's (1998) theory of risk society, he directs attention to the social production of risks in educational policymaking. In such a context, policymaking runs the risk of becoming nothing more than crisis management and "a narrow-minded enterprise of problem-solving instead of historically informed planning for a future society always different but still recognizable" (Nordin, 2014, p. 123).

In this context of urgent reform, and the lack of hard, technical evidence about the kinds of changes that will work, policy borrowing is rife (a.o. Chung, 2016; Kuhlee, 2017). This results in a convergence or sometimes copying of teacher education practices and policies across countries. At the same time, and somewhat ironically, policies in other countries travel in completely

opposite directions as they seek to raise standards in schools (Tatto and Plank, 2007). Successive neoliberal reforms over the past two decades in England, for example, have focused on making teacher education more of a school-focused and instrumental enterprise, influenced by the idea that more practical experience in and of itself produces (better) teachers (a.o. Furlong, 2013; Murray and Mutton, 2016). On the other hand, policies in contexts like Flanders (e.g., Vanassche et al., 2021), Finland (e.g., Aspfors and Eklund, 2017) and Norway (e.g., Smith, 2020), have strengthened the idea of higher education making a distinct and necessary contribution to pre-service teacher education, resulting in corresponding expectations for teacher educators to be or become research active.

Teacher education should be held accountable and wants to hold itself accountable

The current teacher education operates in an increasingly marketized environment in which governments lend teacher education institutions more autonomy and responsibility so that the teacher education market can regulate itself. At the same time, however, the autonomy of the current teacher education is limited by increased control and accountability toward government and public, especially in the face of challenges, looming crises, and policy's insistent spotlight on teacher quality. The current teacher education should be willing to continuously account for its processes and outcomes toward government and public in return for autonomy.

How the current teacher education is being held accountable, to what outcomes, and with what consequences varies considerably though across countries. Helpful is Barbana et al.'s (2020) distinction between two distinct global narratives in accountability policies: high stakes versus low stakes accountability. Although not specifically targeted at teacher education, these narratives translate well to recent teacher education policy developments. The first, high stakes narrative is mostly associated with the English-speaking cultural world such as the USA, England, and Australia where accountability policies have been installed in teacher education from the 1980s onwards (see e.g., Bales' (2006) historic analysis of the accountability shift in the USA). In these contexts, accountability policies are based on an implicit or explicit contract with specific targets and benchmarks for teacher education institutions and educators, and the risk of sanctions if these are not met. This form of accountability tends to be measurement-driven with a particular focus on outcome measures and a reliance on "external scrutiny to achieve its ends" (Sachs, 2016, p. 253). Teacher candidates' learning and, increasingly, pupil learning outcomes and achievement scores are examples of this (Cochran-Smith et al., 2010). Research indicates that in some of these contexts, accountability has led to increasingly sophisticated systems of surveillance and external regulation where teacher educators themselves experience little opportunities for self-regulation (a.o. Murray and Mutton, 2016; Rowe and Skourdoumbis, 2019). The second narrative, low stakes accountability, is often associated with continental Europe. Here, teacher education accountability contracts are more implicit and low scores obtained by programs or institutions do not necessarily and in each case not directly result in formal sanctions "apart from an incentive to think about one's practices in the light of the negative feedback" (Barbana et al., 2020, p. 89). Such forms of responsive or reflexive accountability are less concerned with the outcomes than the process, and focus more on the ability of the current teacher education to regulate itself (Sachs, 2016).

Recent analyses seem to suggest that traditionally low-stakes accountability countries like Belgium, Norway, Sweden, Poland, Portugal, and the Netherlands, slowly move toward high(er) stakes accountability policies. Smith (2018), for example, characterizes Norwegian teacher education as operating in an increased accountability culture, inducing "a feeling of being made answerable for changes which are not yet institutionalized and resolved before new requirements are made accountable to redesign programs" (p. 22). She attributes increased accountability pressures to the uncritical adaption of international and European trends, including the framework for quality assurance systems developed by the European Commission which was obediently and faithfully implemented in the Norwegian context. The result is a tension between trust and mistrust of institutions, programs, and educators in Norwegian teacher education. Snoek and Dengerink (2016) noted similar patterns in the Netherlands. The establishment of the *Dutch Accreditation System for Higher Education* in 2002 is an early sign of the introduction of governance principles in Dutch teacher education with a tension building between, on the one hand, the government's intention to ensure quality in the face of critique and on the other hand, the importance of flexibility and autonomy for teacher education institutions for innovation and renewal. Snoek and Dengerink frame the profession's initiative to develop a *Knowledge Base for Teacher Educators* in 2012 as strategic action to avert increased government control and intervention, at times when concerns around the quality of teachers and teacher education were growing.

Teacher education is evidence-based

The current teacher education emerging from contemporary policy discourses is based on evidence. Its curricular decisions are not (only) based on custom or professional judgment, but on evidence about what strategies "work" and make the biggest impact on candidates' learning and the learning of pupils in their future classrooms. The current teacher education is "in an 'evidence era', where actions are justified through a language shrouded in talk of research, data, and best practice" (Helgetun and Menter, 2020, p. 2). The assumption is that with more and better evidence "practitioners and policymakers at all levels will make better decisions, and teacher quality will improve" (Cochran-Smith, 2005, p. 8).

What is being constructed as credible evidence for the current teacher education is two-fold. On the one hand, the literature exposes an insistent belief in research to deliver evidence. There is a call for increasingly sophisticated research approaches to inform the current teacher education. Consider, for example, the OECD report *Evidence in Education* which calls upon researchers to address the "need to make causal claims" (Cook and Gorard, 2007, p. 44) and the corresponding "ethical need for researchers to use

something like a randomized controlled trial to make these claims responsibly" (p. 44). For the current teacher education based on evidence, causal questions have a special importance and "identifying 'what works' deserves a special status among the concerns of those accountable for the quality of educational performance, as does learning about 'what works' in the most secure ways" (Cook and Gorard, 2007, p. 34). A second trusted source of evidence for the current teacher education is all other data available from internal, external, and increasingly transnational accountability activities. This includes, for example, internal data from institutions' own quality assurance procedures, external data from audits and inspections, but also more distant evidence such as pupil achievement scores on national and international metrics (e.g., TIMSS and PISA). Here, the evidence discourse combines with the accountability discourse. In the process of making teacher education quality transparent for all stakeholders involved, measurable outcomes are created, and those numbers in turn deliver evidence for the necessary changes in teacher education curricula, programs, and institutions, especially from a comparative perspective.

The current teacher education not only effectively uses evidence and research developed by others. It also creates evidence itself when asked for by external evaluators or to check whether a program meets its own standards for good teaching (Cochran-Smith, 2005). Increasingly, teacher educators engage in self-study or related forms of practitioner research to work on the improvement of their own practices and programs, and the profession as a whole (Kitchen et al., 2020). Some countries, especially those with dual systems of higher education, have also seen increasing pressures on teacher educators to become research active (a.o. Czerniawski et al., 2017).

The logic is that with more and better evidence (i.e., evidence that demonstrates the effectiveness of practices and graduates), the current teacher education will make better decisions about the content and delivery of its curriculum, but also be able to respond to the insistent critiques from the public, including parents and policymakers. Grossman (2008), for example, claims "a profession is most vulnerable to attack when serious questions can be leveled at the perceived effectiveness of its practice" (p. 13). While acknowledging that responses to critiques coming from the public "will always be driven by normative assumptions, professional judgment, and ethical commitments" (p. 13), Grossman stands that the profession's responses to critiques "would be substantially enhanced as well as more professionally responsible and publicly credible, if we had a sounder research base with which to refute their claims or support their own" (p. 13). Again, the logic is clear: the current teacher education should be based on evidence about what works, and the profession itself carries responsibility for putting in the effort and developing the skills to determine how its practices and programs make a difference. This evidence preferably takes the form of the outcomes that the public and governments care about most, that is, classroom performance of teacher candidates and pupil achievement, and preferably the relationship between both (Cochran-Smith et al., 2010).

Teacher education is practical

The final defining piece of the current teacher education is that it should be (more) practical. As Kosnik et al. (2016) articulate, "there is considerable comment around the world about the excessively theoretical and abstract nature of traditional teacher education courses" (p. 273). Their observation mirrors the conclusion of an OECD report published in 2010 which, looking across OECD countries, urged for greater "complementarity between field experience and academic studies" (Musset, 2010, p. 46). The report encourages countries to "establish shared responsibility between teacher education institutes and schools in the training of teachers, in order to fill the theory-practice gap" (p. 46). This quotation exposes the logic underpinning policy trends globally to make teacher education more practical: increasing the amount of time teacher candidates spend in schools will produce better teachers (Kidd and Murray, 2020).

This "practice turn" (Reid, 2011) has been a hallmark of teacher education policy reform globally for some time now. It "is in some respects a 'turn around'" (Reid, 2011, p. 211). The idea that newcomers in the profession can learn from the observation and interrogation of demonstrated practice of experienced teachers is a familiar discourse in teacher education policy. The principle of building the curriculum of teacher education on the activities and practices in which teachers engage is also one that, as Zeichner (2012) argues, "has been used in some form for many years" (p. 376). The pendulum swing between theory and practice not only figures in policy discourses. It also shows in the revival of scholarship on practice-based approaches to teacher education (a.o. Grossman, 2021) which first appeared in the literature in the 1970s and 1980s, particularly in the USA.

Like the other pieces of the current teacher education, the (re)emphasizing of practice is a global policy pattern that has played out differently across countries and institutional settings. In some countries, like England, Australia, and parts of the USA, this has resulted in "the far-reaching development toward a school-based teacher education system" (Gilroy, 2014, p. 624) with "limited (or no) emphasis on research or 'theoretical' course work" (Goodwin and Kosnik, 2013, p. 335). School-led and apprentice-style teacher education programs are providing alternative routes into the profession. The growing global network of *Teach for All* is a point in case. In countries evidencing such a strong pull to practice, "the principle of Higher Education making distinctive and necessary contributions to ITE has been steadily eroded in the multiplicity of providers and routes now involved in the 'marketplace' of school-led ITE" (Vanassche et al., 2019, p. 480). The relocating of teacher education into schools has been associated with shifting conceptions of teaching as a craft or applied science with "little theoretical learning and little learning of any kind, beyond mastering the routines of the cooperating teacher's classroom" (Zeichner, 1992, p. 297; see also White and Forgasz, 2016). This not only has consequences for the enacted discourse of teacher professionalism, as I will detail below; It also shapes and changes the workplace landscape and position of teacher educators. The current teacher education is no longer the unique province of higher education-based teacher educators. Large and various groups of school-based teacher educators join the ranks, often constructing their professional identities as teacher educators around their first-order practices in schools. Higher education-based teacher educators face

derogation of their expertise, feeling “less able to compete with school-based teacher educators in meeting the demands of immediate practice” (Brown et al., 2015, p. 7). In other countries, particularly in continental Europe, the practice turn has rather taken shape as a “practicum turn” (Mattsson et al., 2011), resulting in an increase in field experience and partnership models. In the Netherlands, for example, so-called “training schools” were established in which schools and teacher education institutions work alongside each other to educate teachers. These government-funded training schools transform the traditional practicum into guided work-based education (van Velzen, 2013).

The current teacher

Teacher education policies carry related assumptions about the kinds of teachers necessary if quality teaching is to be realized. This section aims to identify what form of teacher professionalism and what kinds of teachers are called for in teacher education policy legislation globally. In line with previous work, I use the term professionalism “as a descriptor of what makes up the essence of being a teacher and the (characteristics of the) knowledge, skills and attitudes invested in the work” (Vanassche et al., 2021, p. 28). The trust of the argument is that policy discourses of teacher professionalism have shifted the past decades from the teacher as a professional who relies upon subject matter and pedagogical knowledge and values to enact practice, toward the teacher as a technical expert who knows and can perform what works, as derived from research on effective teaching. The current teacher is also flexible and responsive to the always changing needs of the society, schools, and pupils.

Although the section below, mirroring the available scholarship, mainly discusses the effects of contemporary policy developments for teacher professionalism, it is easy to see how these discursive shifts also pertain to teacher educators. I return to this observation in the conclusion of this article.

Teachers know and perform “what works”

Current teacher education policy, and its combined focus on evidence and practice, leads into a conception of “professionalism as effective intervention” (Biesta, 2007, p. 7). It starts from “the idea that professionals do something (...) in order to bring about certain effects” (p. 7). The frequent use of terms like “effective teachers” and “effective teaching,” and the related emphasis on “evidence-based practice” and “evidence-based teaching” in recent teacher education policy documentation globally reflects this orientation. Teacher professionalism is increasingly framed “in the technical scheme of ‘what works/does not work’” (Vanassche et al., 2021, p. 35). The current teacher’s classroom strategies and behaviors are among the most efficient and effective available. In such a context, the primary task for the current teacher education is to train candidates who know what works and are able to perform and replicate that in their classrooms. This suggests a policy turn that shifts teacher education back toward a teacher training paradigm (Cochran-Smith, 2005; Beach and Bagley, 2013).

The demarcation of “what works/does not work” draws on the kind of empirical evidence also advocated for in the strengthening and development of the current teacher education. That is, “scientific research programs that seek the determinants (or variables) of effective teaching and teacher education, and build a knowledge base for use by teachers” (Vanassche and Berry, 2020, p. 11). The value of such research for teachers is increasingly “judged according to the use or not of prescribed methodologies” (Mayer, 2021, p. 6; see also Cook and Gorard, 2007). For questions of a causal nature, like effectiveness issues (i.e., does a classroom strategy produce the intended effect?), preference is given to robust experimental and quantitative designs. As Vanassche et al. (2021) have argued, this not only shapes and defines what counts as valid research, but also evokes a technical and rather linear relationship between practice and research. In such a logic, they argue, “teacher professionalism concerns the application of a research-informed body of knowledge to solve problems” (p. 35). Much like the current teacher education, the current teacher implements insights and evidence from research with fidelity. The term “fidelity” refers to the notion of treatment fidelity in clinical settings. To assess if a drug works, it needs to be administered under specified conditions in the prescribed dosage at prescribed time intervals. The current policy environment urges teachers to implement proven strategies as planned, or with fidelity. Consider, for example, Sanetti et al.’s (2021) advice to teachers, whom they address as “research consumers,” to “implement an evidence-based intervention in their classes (...) with fidelity (...) as students may not respond positively to an intervention they did not receive as intended” (p. 5). If a problem does arise under such conditions, the most likely explanation is faulty implementation rather than a faulty strategy. This shows how current policy discourses tend to minimize teacher discretion, and in some cases suppress teacher reflection (Bourke et al., 2016).

The conception of teacher professionalism as knowing and performing what works carries the related assumptions that professionalism is fairly generic and context-free, and can be mapped and checked (Vanassche et al., 2021). In many policy contexts, this is materialized in the development of teacher standards or competency frameworks. Standards offer a presumably exhaustive list of knowledge, skills, and attitudes that present best evidence and can be reasonably expected from teachers. Most countries have standards in place for experienced teachers but also slimmed down starting requirements for beginning teachers outlining what they “should know at the conclusion of a professional teacher education degree program” (Bullen et al., 2010, p. 47) and, increasingly, entry requirements for programs of pre-service teacher education serving as predictors of future teacher performance (Parker, 2018). Standards are considered key policy instruments in efforts to raise educational standards. Especially in countries with high stakes accountability narratives, the standards take a regulatory approach (Sachs, 2016) in that they are owned and imposed by the government and used to regulate the profession with consequences for those who do not fulfill the standards (i.e., access to tenure, higher

pay scales, etc.). In other, low(er) stakes contexts the standards “emerge through consultation among various stakeholders and are, in the main, profession led, rather than bureaucratically imposed” (Sachs, 2016, p. 256). In contrast to regulatory standards, such developmental standards are concerned with development and improvement of teacher and student performance. But regardless of the purposes and use of standards, the policy logic is clear: what makes up the essence of (good) teachers can be defined, and we can get better at that as scientific research makes significant progress in mapping what works and the conditions under which it works. These insights can then be used as a template for the selection and training of newcomers to the profession.

Research has documented how such a technical-instrumental view of teacher professionalism oversimplifies the teaching and learning process, because it fails to address context and makes obsolete discussions around the purposes of education in policy and practice. The instrumental view of teacher professionalism currently installed in teacher education policy globally leads to a view of the professional teacher as “one who ‘does things right’ rather than one who ‘does the right things’” (Darling-Hammond, 1990, p. 31). The question “what works” takes precedence over the question “what it should work for” (Biesta, 2007). In such discursive contexts, disagreement “becomes limited to varying views as to the best means by which instrumental goals can be achieved; dissent regarding the purposes of education becomes off limits” (Clarke and Phelan, 2015, p. 259).

Teachers are responsive and flexible

Overlooking the teacher education policy literature, researchers appear to be highlighting a contrast and in most countries a shift from two contrasting models of teacher professionalism and the virtues associated with it. Researchers in different contexts have used various nomenclatures, but broadly speaking they signal a shift from profession-oriented discourses in teacher education policy environments to market-oriented discourses, terms borrowed from Simons and Kelchtermans (2008). The conception of the teacher as professional is associated with professional virtues of expertise, responsibility, and autonomy. The professional teacher is someone who claims and owns the responsibility to teach based on his/her subject matter knowledge and pedagogical knowledge to teach a particular subject. “Policy is ‘profession-oriented’ when it acknowledges that teaching and teacher education (programs and institutes) rely upon specific (educational or pedagogic) values or upon a particular scientific, theoretical framework and particular study subjects” (Simons and Kelchtermans, 2008, p. 288). That is, policymakers acknowledge and give room for the specific values and professional knowledge which “govern” teaching and teacher education. Such a focus on professional and institutional autonomy (Olssen et al., 2004) does not prevent government intervention, but government intervention is used with caution and is positioned as a complement to the profession’s self-governance. By contrast, the kind of teacher currently called for in policy is not someone who owns the freedom to teach based on specific values and professional knowledge, but someone “who makes sure that what s/he offers provides an appropriate answer to the needs of the society, the school and the pupils” (Simons and Kelchtermans, 2008, p. 292). Current policy discourses replace professional responsibility by responsiveness and attentiveness to the always changing needs of the school system and society at large. The current teacher is on top of her/his game and able and willing to flexibly adapt to changing needs. The policy demand for such teachers is constructed around the increasingly complex challenges presented by the modern world, such as rapid developments in ICT, a changing labor market, and shifting demographics in society and schools. In sum, teacher education policy seems to promote a managerial (Sachs, 2016) and organizational (Evetts, 2009) view of professionalism, replacing profession-oriented values (i.e., expertise, responsibility, autonomy) with market-oriented or corporate values (i.e., effectiveness, responsiveness, and flexibility) (Simons and Kelchtermans, 2008).

As Gewirtz et al. (2009) argue, it is too simple to argue that one form of professionalism has completely overhauled the other “because in reality the two kinds of professionalism are more likely to coexist” (p. 7). Nevertheless, looking at the teacher education policy literature across national contexts, the balance between both forms of professionalism seems to be shifting. Consider, for example, Tessema’s (2007) examination of policy directives in Ethiopia. Reform mottos such as “active learning,” “competence” and “participatory” increasingly evidence a managerial approach to teacher education, a term he uses to denote policymakers’ focus on strategic planning and management of teacher education and teachers, reflecting a corporate style structuration of the profession. “It is a process that ignores pedagogical and educative arrangements such as teacher research or inquiry, continuous pedagogical dialogs or deliberations, pedagogical innovations and sharing, pedagogical voluntarism and activism, and collaborative engagements for increased professionalization and ownership” (p. 41). In Tessema’s view, such a conception of managerial professionalism “sidesteps pedagogues and pedagogy in favor of fulfilling instrumental, central and market-oriented agendas” (p. 29) and peripheralizes teachers in the Ethiopian context. Beach and Bagly’s (2013) two-way comparison between the discourses of professionalism enacted in Swedish and English teacher education policy mirrors these findings. “Sweden has, to a degree, resisted global neo-liberal educational policy and refracted its ambitions differently to other countries, while the UK has seemed to embrace these policies and was one of the first countries to adopt them” (p. 380). Despite these different logics, Beach and Bagly’s analysis of the effects of changing teacher education policy on the sense of professionalism in both countries revealed marked similarities in terms of an increasing need for teacher education institutions, educators, and students to flexibly adopt and “respond to (and integrate into teaching and practice) a constantly shifting prescribed list of behaviors, competencies and standards, against which their professional work will be assessed” (p. 389). These examples show evidence of the effective, responsive, and flexible teacher hinted at by Simons and Kelchtermans (2008).

The rise of the managerial prerogative across jurisdictions has been attributed to the increased focus on accountability, standards, and performance in education policy (Mayer, 2021), but also to global forces, such as globalization and neo-liberal reform agendas. Although some of these agendas have taken longer to reach some countries, particularly in continental Europe, “what we seem to be

witnessing today are the effects of a common global discourse of neo-liberal teacher education reform that varies primarily only in terms of being spoken in national, local and/or regional dialectal forms: ‘same, same, but different’, as Sjöberg (2011) so eloquently put it” (Beach and Bagly, 2013, p. 389). In sum, the teacher emerging from education policy across the world not only knows what works but is also flexible and responsive to changing demands made by pupils, society, and policy priorities.

What’s next?

The current teacher education emerging from the literature is one that draws on evidence and research to make decisions about the preparation of teachers and work on necessary improvements in the face of crisis. It legitimizes its existence based on its relevance for practice on the one hand and by accounting for the outcomes and effects it produces with teacher candidates to government and public on the other hand. Teacher education policy also carries implicit visions about desired forms of teacher professionalism which have changed considerably over the past 15–20 years. The teacher currently called for in policy texts is one that knows and can perform what works, as derived from teacher effectiveness research. S/he is also flexible and responsive to the always changing needs of the society, schools, pupils, and policymakers.

At the risk of over-simplifying, there is a basic tension at the heart of the relationship between the policy level and teacher education institutions in many countries. A tension between trust and mistrust, terms borrowed from Smith (2018). A huge amount of trust—or even faith—is placed upon teacher education and teaching as vectors of educational reforms of all kinds, of individual and collective prosperity and competitiveness. The policy spotlight on quality teaching suggests that this trust has perhaps never been greater. At the same time, however, the abilities of the current teacher education and teacher to meet those high expectations are mistrusted. The main issue causing tension between trust and mistrust in many countries are accountability systems holding teacher education and teaching accountable to increasingly narrow outcome measures, such as teacher candidates’ classroom performance and pupil learning outcomes.

This living configuration of the current teacher education and teacher across policy contexts shows the convergences in policy orientations across the world. Countries evidence a degree of cross-national attraction in the global policy arena, inviting policy borrowing, in some countries around an increasingly “limited menu of change” (Darling-Hammond, 2012, p. ix). Global and pan-European organizations such as the OECD and the European Commission have intensified these processes through sharing of best practices and evidence, often from a comparative perspective. Moreover, this research synthesis suggests that some policies travel the world in rather specific directions. There appear to be “leading countries,” that is, countries that have shown glimpses of what I described as the current teacher education and the current teacher years before it emerged as a more global configuration. A point in case are the accountability pressures detected in the USA and England in the 1980s now slowly conquering continental Europe. Critical policy research often uses these early adopters as the proverbial canaries in the coalmine for late adopters, drawing on comparative references to “caution against educational reform” (Ochs, 2006, p. 602) or, in some cases, “glorify current education at home in comparison to other nations” (p. 602). More research is needed to understand these policy dynamics.

An important question to address at this stage is what’s next? What are future directions for this growing field of research? First, I reiterate the need for more explicit and direct comparisons of teacher education policy discourses and practices across contexts. Such research would allow to better understand the traveling of policies between countries and the mediation of the global and the local in teacher education policy processes. This work necessitates a careful balance between the increasing globalization of ideas and the national, regional, and local enactments of those ideas. This balance needs to be made at every stage of the comparative work, from choosing an appropriate theoretical-analytical framework to deciding on the level of analysis and sources of data. Second, there is a need to include insiders’ perspectives (i.e., teacher educators, teachers, teacher candidates, program leaders, etc.) in the analysis of (shifting) policy orientations. I am reminded of Ball’s (1997) call to “people policy.” The phrase pushes back on the view “that policy is something that is ‘done’ to people” (p. 270) in that they either implement it or are advantaged or disadvantaged by it. The teacher education profession is shaped by, but also helps to shape the discourses and practices described above. Ball’s call to people policy research is a call to actively include the “sort of people and ‘voices’ [which] inhabit the texts of policy analysis” (p. 271) in research. For example, Webb et al. (2004) deliver a comparative analysis of primary teacher professionalism in England and Finland. They confronted policy text analysis with interviews of teachers. Interestingly, the interviews showed teachers welcomed increased government intervention in the Finnish context which had been traditionally built on notions of professional autonomy. Implementing government-imposed directions in literacy and numeracy “that deliberately targeted pedagogy—arguably the last preserve of teacher professionalism—led teachers to claim enhanced professionalism” (Webb et al., 2004, p. 102) as the reforms reduced uncertainty about their role and “[redressed] the increasing inequity and inconsistencies within the system that were of direct concern to them” (p. 103). This shows the value of including insiders’ voices in macro-debates about teacher (de-)professionalization. Third, teacher education policy legislation not only produces normative discourses around what it means to be a teacher today; It also shifts discourses around what it means to be a teacher educator. Consider, for example, Vanassche et al. (2019) analysis of how policy changes in England have resulted in narrower, more instrumental, and school-centered forms of teacher education. In such a context, legitimacy as a teacher educator increasingly centers around practical knowledge of teaching with university-based teacher educators being positioned as out of sync with the demands of the profession. This is but one example of the impact of changing policy discourses on teacher educators’ work, lives, and professionalism which need to be explored further.

References

- Adams, A., Tulasiewicz, W., 1995. *The Crisis in Teacher Education: A European Concern?* The Falmer Press, London.
- Aspfors, J., Eklund, G., 2017. Explicit and implicit perspectives on research-based teacher education: newly qualified teachers' experiences in Finland. *J. Educ. Teach.* 43, 400–413.
- Assunção Flores, M., 2011. Curriculum of initial teacher education in Portugal: new contexts, old problems. *J. Educ. Teach.* 3, 461–470.
- Bacchi, C., 2010. Why study problematizations? Making politics visible. *Open J. Polit. Sci.* 2, 1–8.
- Bales, B.L., 2006. Teacher education policies in the United States: the accountability shift since 1980. *Teach. Teach. Educ.* 22, 395–407.
- Ball, S.J., 1994. At the cross-roads: education policy studies. *Br. J. Educ. Stud.* 42, 1–5.
- Ball, S.J., 1997. Policy sociology and critical social research: a personal review of recent education policy and policy research. *Br. Educ. Res. J.* 23, 257–274.
- Barbana, S., Dumay, X., Dupriez, V., 2020. Accountability policy forms in European education systems: an introduction. *Eur. Educ. Res. J.* 19, 87–93.
- Beach, D., Bagley, C., 2013. Changing professional discourses in teacher education policy back towards a training paradigm: a comparative study. *Eur. J. Teach. Educ.* 36, 379–392.
- Beauchamp, G., Clarke, L., Hulme, M., Jephcote, M., Kennedy, A., Magennis, G., Menter, I., Murray, J., Mutton, T., O'Doherty, T., Peiser, G., 2015. *Teacher Education in Times of Change*. Policy Press, Bristol.
- Beck, U., 1998. Politics of risk society. In: Franklin, J. (Ed.), *The Politics of Risk Society*. Polity Press, London, pp. 9–22.
- Biesta, G., 2007. Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educ. Theor.* 57, 1–22.
- Bourke, T., Ryan, M., Lloyd, M., 2016. The discursive positioning of graduating teachers in accreditation of teacher education programs. *Teach. Teach. Educ.* 53, 10–19.
- Brock, C., 1996. *Global Perspectives on Teacher Education*. Triangle Books, Wallingford.
- Brown, T., Rowley, H., Smith, K., 2015. *The Beginnings of School Led Teacher Training: New Challenges for University Teacher Education*. Manchester Metropolitan University, Manchester.
- Bullen, P., Farruggia, S.P., Gómez, C.R., Hebaishi, G.H.K., Mahmood, M., 2010. Meeting the graduating teacher standards: the added benefits for undergraduate university students who mentor youth. *Educ. Horiz.* 89, 47–61.
- Chung, J., 2016. The (mis)use of the Finnish teacher education model: "Policy-based evidence-making"? *Educ. Res.* 58, 207–219.
- Clarke, M., Phelan, A., 2015. The power of negative thinking in and for teacher education. *Power Educ.* 7, 257–271.
- Cochran-Smith, M., Gleeson, A.M., Mitchell, K., 2010. Teacher education for social justice: what's pupil learning got to do with it? *Berkeley Rev. Educ.* 1, 35–61.
- Cochran-Smith, M., Stringer Keefe, E., Cummings Carney, M., 2018. Teacher educators as reformers: competing agendas. *Eur. J. Teach. Educ.* 41, 572–590.
- Cochran-Smith, M., 2005. The new teacher education: for better or for worse? *Educ. Res.* 34, 3–17.
- Cochran-Smith, M., 2020. Accountability and initial teacher education reform: a perspective from abroad. *Wales J. Educ.* 22, 59–81.
- Cook, T., Gorard, S., 2007. What counts and what should count as evidence. In: *Centre for Educational Research and Innovation. Evidence in Education: Linking Research and Policy*. OECD, Paris, pp. 33–49.
- Czerniawski, G., Guberman, A., MacPhail, A., 2017. The professional developmental needs of higher education-based teacher educators: an international comparative needs analysis. *Eur. J. Teach. Educ.* 40, 127–140.
- Darling-Hammond, L., Lieberman, A., 2012. *Teacher Education Around the World Changing Policies and Practices*. Routledge, London.
- Darling-Hammond, L., 1990. Teacher professionalism: why and how. In: Lieberman, A. (Ed.), *Schools as Collaborative Cultures: Creating the Future Now*. Falmer Press, London, pp. 25–50.
- Darling-Hammond, L., 2012. Foreword. In: Tucker, M.S. (Ed.), *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems*. Harvard Education Press, Cambridge, pp. ix–xii.
- Delandshere, G., Petrosky, A., 2004. Political rationales and ideological stances of the standards-based reform of teacher education in the US. *Teach. Teach. Educ.* 20, 1–15.
- Evetts, J., 2009. New professionalism and new public management: changes, continuities and consequences. *Comp. Sociol.* 8, 247–266.
- Furlong, C., O'Brien, M., 2019. An exploration of teacher educator identities within and Irish context of reform. In: Murray, J., Swennen, A., Kosnik, C. (Eds.), *International Research, Policy and Practice in Teacher Education: Insider Perspectives*. Springer, Cham, pp. 47–62.
- Furlong, J., Cochran-Smith, M., Brennan, M., 2009. *Policy and Politics in Teacher Education: International Perspectives*. Routledge, London.
- Furlong, J., 2013. *Education: An Anatomy of the Discipline*. Routledge, London.
- Gewirtz, S., Mahon, P., Hextall, I., Cribb, A. (Eds.), 2009. *Changing Teacher Professionalism: International Trends, Challenges and Ways Forward*. Routledge, London/New York.
- Gilroy, P., 2014. Policy interventions in teacher education: sharing the English experience. *J. Educ. Teach.* 40, 622–632.
- Ginsburg, M.B., 1998. *Contradictions in Teacher Education and Society: A Critical Analysis*. Falmer Press, London.
- Goodwin, A.L., Kosnik, C., 2013. Quality teacher educators = quality teachers? Conceptualizing essential domains of knowledge for those who teach teachers. *Teach. Dev.* 17, 334–346.
- Grek, S., 2009. Governing by numbers: the pisa "effect" in Europe. *J. Educ. Pol.* 24, 23–37.
- Grimmett, P.G., 2009. The governance of Canadian teacher education: a macro-political perspective. *Counterpoints* 334, 22–32.
- Grossman, P., 2008. Responding to our critics: from crisis to opportunity in research on teacher education. *J. Teach. Educ.* 59, 10–23.
- Grossman, P., 2021. *Teaching Core Practices in Teacher Education*. Harvard Educational Publishing Group, Harvard.
- Helgetun, J.B., Menter, I., 2020. From an age of measurement to an evidence era? Policy-making in teacher education in England. *J. Educ. Pol.* 1–18.
- Kidd, W., Murray, J., 2020. The Covid-19 pandemic and its effects on teacher education in England: how teacher educators moved practicum learning online. *Eur. J. Teach. Educ.* 43, 542–558.
- Kitchen, J., Berry, A., Bullock, S.M., Crowe, A.R., Taylor, M., Guðjónsdóttir, H., Thomas, L., 2020. *2nd International Handbook of Self-Study of Teaching and Teacher Education*. Springer, Singapore.
- Kosnik, C., Beck, C., Goodwin, L., 2016. Reform efforts in teacher education. In: Loughran, J., Hamilton, M.L. (Eds.), *Handbook on Teacher Education*. Springer, Dordrecht, pp. 207–224.
- Kuhlee, D., 2017. The impact of the Bologna reform on teacher education in Germany: an empirical case study on policy borrowing in education. *Res. Comp. Int. Educ.* 12, 299–317.
- Lewin, K., Samuel, M., Sayed, Y. (Eds.), 2003. *Changing Patterns of Teacher Education in South Africa: Policy, Practice and Prospects*. Heinemann, Sandown.
- Lewis, W.D., Young, T.V., 2013. The politics of accountability: teacher education policy. *Educ. Pol.* 27, 190–216.
- Madaliriska-Michalak, J., O'Doherty, T., Assunção Flores, M., 2018. Teachers and teacher education in uncertain times. *Eur. J. Teach. Educ.* 41, 567–571.
- Maguire, M., 2000. The state regulation of United Kingdom teacher education in the nineteenth century: the interplay of "value" and "sense". *Int. Stud. Sociol. Educ.* 10, 227–242.
- Mattsson, M., Eilertson, T., Rorisson, D. (Eds.), 2011. *A Practicum Turn in Teacher Education*. Sense, Rotterdam.
- Mayer, D., Goodwin, A.L., Mockler, N., 2021. Teacher education policy: future research, teaching in contexts of super-diversity and early career teaching. In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore, pp. 209–223.
- Mayer, D., 2021. *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore.
- McGarr, O., Emstad, A.B., 2020. Comparing discourses on reflective practice in teacher education policy in Ireland and Norway: critical reflection or performance management? *Eur. J. Teach. Educ.* 1–19.

- Meijer, P.C., 2021. Quality under pressure in Dutch teacher education. In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore, pp. 101–112.
- Moore, R., 1996. Back to the future: the problems of change and possibilities of advance in the sociology of education. *Br. J. Sociol. Educ.* 17, 145–162.
- Murray, J., Mutton, T., 2016. Teacher education in England: change in abundance, continuities in question. In: *The Teacher Education Group. Teacher Education in Times of Change*. Policy Press, Bristol, pp. 57–74.
- Murray, J., Swennen, A., Kosnik, C., 2019. *International Research, Policy and Practice in Teacher Education: Insider Perspectives*. Springer, Cham.
- Musset, P., 2010. Initial Teacher Education and Continuing Training Policies in a Comparative Perspective: Current Practices in OECD Countries and a Literature Review on Potential Effects. *OECD Education Working Papers*, No. 48. OECD, Paris.
- Nordin, A., 2014. Crisis as a discursive legitimization strategy in educational reforms: a critical policy analysis. *Educ. Inq.* 5, 109–126.
- Ochs, K., 2006. Cross-national policy borrowing and educational innovation: improving achievement in the London Borough of Barking and Dagenham. *Oxf. Rev. Educ.* 32, 599–618.
- Olssen, M., Codd, J., O'Neill, A.-M., 2004. *Education Policy: Globalization, Citizenship and Democracy*. SAGE, London.
- Parker, S., 2018. *Literature Review on Teacher Education Entry Requirements*. University of Glasgow, Glasgow.
- Phelan, A.M., Morris, J.D., 2021. Teaching and teacher education for a post-pandemic Canada: context, crisis, critique and complication. In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore, pp. 43–56.
- Reid, J., 2011. A practice turn for teacher education? *Asia Pac. J. Teach. Educ.* 39, 293–310.
- Rowe, E.E., Skourdoubis, A., 2019. Calling for “urgent national action to improve the quality of initial teacher education”: the reification of evidence and accountability in reform agendas. *J. Educ. Pol.* 34, 44–60.
- Sachs, J., 2016. Accountability, standards and activism: a challenge or opportunity for teacher education. In: Chi-Kin Lee, J., Day, C. (Eds.), *Quality and Change in Teacher Education: Western and Chinese Perspectives*. Springer, New York/London, pp. 251–262.
- Sanetti, L.M.H., Cook, B.G., Cook, L., 2021. Treatment fidelity: what it is and why it matters. *Learn. Disabil. Res. Pract.* 36, 5–11.
- Simons, M., Kelchtermans, G., 2008. Teacher professionalism in Flemish policy on teacher education: a critical analysis of the Decree on teacher education (2006) in Flanders, Belgium. *Teach. Teach.* 14, 283–294.
- Simpson, A., Cotton, W., Gore, J., 2021. Teacher education/ors in Australia: still shaping the profession despite policy intervention. In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore, pp. 11–26.
- Smith, K., 2018. Accountability in teacher education in Norway: a case of mistrust and trust. In: Wyatt-Smith, C., Adie, L. (Eds.), *Innovation and Accountability in Teacher Education: Setting Directions for New Cultures in Teacher Education*. Springer, Singapore, pp. 19–35.
- Smith, K., 2020. Expansive learning for teacher educators: the story of the Norwegian National Research School in Teacher Education (NAFOL). *Front. Educ.* 5, 1–12.
- Snoek, M., Dengerink, J., 2016. Lerarenopleidingen in het politieke speelveld. In: Snoek, M., Kool, Q., Walreven, A. (Eds.), *Kennisbasis lerarenopleiders*. Kenniscentrum Onderwijs en Vorming, Amsterdam, pp. 51–68.
- Snoek, M., 2021. Educating quality teachers: how teacher quality is understood in the Netherlands and its implications for teacher education. *Eur. J. Teach. Educ.* 44, 309–327.
- Steiner-Khamisi, G., 1999. Teacher education reform. *Comp. Educ. Rev.* 43, 353–361.
- Symeonidis, V., 2021. *Europeanisation in Teacher Education: A Comparative Case Study of Teacher Education Policies and Practices*. Routledge, New York/London.
- Tatto, M.T., Plank, D.N., 2007. The dynamics of global teaching reform. In: Tatto, M.T. (Ed.), *Reforming Teaching Globally*. Information Age Publishing, Charlotte, pp. 267–276.
- Tessema, K.A., 2007. The teacher education reform process in Ethiopia: some consequences on educators and its implications. *Teach. Educ.* 18, 29–48.
- Trippestad, T.A., Swennen, A., Werler, T., 2017. Reform as a state of exception. In: Trippestad, T.A., Swennen, A., Werler, T. (Eds.), *The Struggle for Teacher Education: International Perspectives on Governance and Reforms*. Bloomsbury Academic, London, pp. 185–191.
- Välimaa, J., Papatsiba, V., Hoffman, D.M., 2016. Higher education in networked knowledge societies. In: Hoffman, D.M., Välimaa, J. (Eds.), *Re-becoming Universities? Higher Education Institutions in Networked Knowledge Societies*. Springer, Dordrecht, pp. 13–40.
- van Velzen, C., 2013. *Guiding Learning Teaching: Towards a Pedagogy of Work-Based Teacher Education (PhD)*. Ipskamp Drukkers, Enschede.
- Vanassche, E., Berry, A., 2020. Teacher educator knowledge, practice, and S-STEEP research. In: Kitchen, J., Berry, A., Bullock, S.M., Crowe, A.R., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *2nd International Handbook of Self-Study of Teaching and Teacher Education*. Springer, Singapore, pp. 177–213.
- Vanassche, E., Kidd, W., Murray, J., 2019. Articulating, reclaiming and celebrating the professionalism of teacher educators in England. *Eur. J. Teach. Educ.* 42, 478–491.
- Vanassche, E., Bruneel, S., Christiaens, L., 2021. A critical examination of the conception of teacher professionalism enacted in current teacher education policy in Flanders (Belgium). In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, Singapore, pp. 27–42.
- Webb, R., Vulliamy, G., Härmäläinen, S., Sarja, A., Kimonen, E., Nevalainen, R., 2004. A comparative analysis of primary teacher professionalism in England and Finland. *Comp. Educ.* 40, 83–107.
- White, S., Forgasz, R., 2016. The practicum: the place of experience? In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education*. Springer, Dordrecht, pp. 231–266.
- Young, M.D., Diem, S., 2017. *Critical Approaches to Education Policy Analysis: Moving Beyond Tradition*. Springer, Cham.
- Young, M.D., Diem, S., Sampson, C., 2019. Where critical policy meets the politics of education: an introduction. *Educ. Pol.* 33, 3–15.
- Zeichner, K.M., 1992. Rethinking the practicum in the professional development school partnership. *J. Teach. Educ.* 43, 296–307.
- Zeichner, K.M., 2012. The turn once again toward practice-based teacher education. *J. Teach. Educ.* 63, 376–382.
- Zeichner, K.M., 2018. *The Struggle for the Soul of Teacher Education*. Routledge, New York.

Standardized learning: the “problem” of teachers’ work and learning under performative policy conditions

Ian Hardy, School of Education, The University of Queensland, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	20
What’s the problem represented to be?	20
The policy problem: Insufficiently detailed curriculum, lack of teaching standards, and poor student performance	21
Broader, global policy conditions	21
Standardized curriculum	22
Standardized teaching	23
Standardized assessment	24
So, what’s the problem <i>really</i> represented to be?	24
Broad principles to inform an alternative approach to educational policymaking	25
Conclusion	26
References	27

Introduction

Teachers’ work and learning are heavily influenced by the specific policy conditions within which they are practiced. Current policy conditions are characterized by increased pressures for improvement on particular measures of student attainment. Increasingly, these measures reflect broader concerns and discourses about the nature of curriculum provision, the quality of teaching, and student outcomes—particularly outcomes on national and international standardized tests of literacy and numeracy, and increasingly, science. Understanding the nature of current curriculum, teaching and assessment provision in this policy context is a first step to fostering conditions productive of more substantive alternatives.

This article draws upon Carol Bacchi (2009) influential work on policy analysis, and especially how particular issues come to be framed as “problems” of a particular kind. These understandings are applied to current conceptions of teachers’ classroom work and learning. Bacchi’s (2009) notion of “What’s the problem represented to be?” is helpful for making sense of current perceptions of teachers’ work and learning—particularly the push for increased standardization—and how such practices have come to be manifest in policy and practice.

What’s the problem represented to be?

Bacchi (2009) work implies that policies always constitute issues in particular ways, presenting a particular set of circumstances that will frame how such issues might be addressed. Bacchi (2009) research enables us to better understand “how governing takes place”, “how order is maintained”, and “how we live within and abide by rules” (p. ix). Government policy does not need to construct these issues as overt problems per se, but instead implies that particular kinds of reform need to take place in order to foster what is construed as necessary change:

Most government policies do not officially declare that there is a problem that the policy will address and remedy. Rather this is implicit in the whole notion of policy—by their nature policies make changes, implying that something *needs* to change.

Bacchi (2009, pp. ix–x).

Such an approach challenges more rationalistic notions of government as responding and responsive to some sort of readily comprehensible or generally shared understanding of an issue or concern, and enacting particular “solutions” understood as generally or widely accepted. As Bacchi (2012) argues:

It starts from the premise that what one proposes to do about something reveals what one thinks is problematic (needs to change). Following this thinking, policies and policy proposals contain implicit representations of what is considered to be the “problem” (“problem representations”) (p. 21).

Consequently, it is important to make *explicit* that which is construed as *implicit* about the particular “problem” at hand. To elicit these more explicit understandings, Bacchi (2009) proposes six questions:

- (1) What is the “problem” ... represented to be in a specific policy?
- (2) What presuppositions or assumptions underlie this representation of the “problem”?
- (3) How has this representation of the “problem” come about?
- (4) What is left unproblematic in this problem representation? Where are the silences? Can the “problem” be thought about differently?
- (5) What effects are produced by this representation of the “problem”?
- (6) How/where has this representation of the “problem” been produced, disseminated and defended? How could it be questioned, disrupted and replaced? (p. xii)

Problems are therefore constructed in particular ways, emphasizing some issues and matters over others, silencing alternatives and establishing dominant modes of understanding and interpretation. Such selectivity also reflects representations as deeply political and indicative of how specific politics at play always constitute “a web of decisions and actions that allocates values” (Easton, 1953, p. 130)—constituting some values as much more authoritative than others.

The notion of policy here is also understood as eclectic. Ball (1994) argues policy constitutes not only recognizable administrative documents oriented to the enactment of specific political prerogatives but also associated texts, including legislation as well as the broader contexts and discourses associated with such texts. Policy also entails the outcomes or effects of such texts. For Ball (1994), “[p]olicy is both text and action, words and deeds; it is what is enacted as well as what is intended. Policies are [also] always incomplete insofar as they relate to or map on to the ‘wild profusion’ of local practice” (p. 10). Such an approach makes it possible to draw upon a multitude of perspectives from a variety of actors engaged in both developmental and enactment processes and analysis of the circumstances within which such work occurs. Collectively, Bacchi (2009) and Ball (1994) ideas are particularly useful for making sense of how teachers’ work and learning are currently constituted within official policy contexts, texts and effects, and how they might be constituted differently.

I apply Bacchi (2009) theorizing to understanding the broader political and policy context in which teachers’ work and learning are currently constituted and enacted. Bacchi (2009) argues it is important to “make the ‘problems’ implicit in public policies *explicit* and to scrutinize them closely” (p. x; emphasis original). Specifically, I draw upon broader research literature about politics and policies associated with teachers’ work and learning to try to shed light upon how such work and learning are currently understood in policy, including in order to shed light upon the implications of such contextual circumstances for both teachers and students’ learning.

The policy problem: Insufficiently detailed curriculum, lack of teaching standards, and poor student performance

In relation to education provision in schools, the primary “problem” is often characterized as one of inadequate or insufficient “educational excellence” or “attainment” or “outcomes”. At national and global scales, this is manifest in the push for enhanced results on national and international large-scale assessments, and pressure to ensure a comparative advantage against other countries participating in such tests (Verger et al., 2019). These results are also linked with the nature of the curriculum developed in different national contexts and the way such curricula are enacted in teaching/pedagogical practice. Each of these “message systems” of curriculum, pedagogy and assessment (Bernstein, 1971) reflect particular concerns about educational provision. However, these practices of curriculum development, pedagogical practice and assessment do not exist in isolation but are instead reflective of the broader policy conditions within which they are expressed. These policy conditions are not only specific to particular locales, but also often “global” in orientation, even as they are expressed very differently in specific policy contexts (de Sousa Santos, 2006).

Broader, global policy conditions

More localized curriculum, teaching and assessment practices are framed within a broader policy and political context characterized by increased pressures for improved outcomes and the attainment of such outcomes as efficiently as possible. Contrasting processes of neoliberalism and managerialism operate alongside one another, with current broader policy conditions manifesting such emphases in multiple and contradictory ways. Reflecting more managerial influences, different modes of ranking and measuring have been utilized as “indicators” for the purposes of presenting and re-presenting the activities to both those within these organizations as well as to various external stakeholders and publics (Shore and Wright, 2015). Such responses are the manifestation of what have been referred to as various kinds of “audit cultures” (Power, 1997; Strathern, 2000) that seek to manage and monitor educators’ practices. The proliferation of such cultures is reflective of substantive change and reform, and not simply a passing fad or proclivity. As the name suggests, the adoption of such audit cultures also implies the infiltration of auditing practices more typically associated with accountancy and finance. Such practices are often technicist in orientation and reduce to what can be captured through specific, often metric, modes and markers of achievement. Such processes seek to exclude contestation, oriented as they are as alternatives to what is construed as problematic judgment on the part of teachers. Such judgments are

construed as needing to be not simply augmented but superseded by a much more authoritative “measure” of student learning and performance. Under such circumstances, processes of standardization and quantification become dominant, even as they fail to capture the nature of actual student learning and educational practices more broadly (Hardy, 2015), and even as educators seek to “appropriate” more performative practices for educative purposes (Hardy, 2014). The development of such numbers also present something of a quandary, given that may have the potential to be educative, even as they are simultaneously limited by the parameters of their development (Hardy, 2020).

Alongside more managerial practices and processes, more neoliberal influences have also exerted influence. Even as notions of neoliberalism are heavily contested in educational research (Rowe et al., 2019), it is possible to identify various processes of “neoliberalization” (Peck, 2010), whereby market reforms have been variously implemented, involving processes of regulatory restructuring on the part of governments as they seek to implement a multitude of reforms across the public sector—including public education. Competition between schools is encouraged as a means of ensuring greater efficiency of educational provision, and improved outcomes.

In relation to education, these broader policy conditions have been increasingly influenced by what Sahlberg (2016) describes as the “Global Education Reform Movement” (“GERM”). GERM constitutes the common education policies governments have sought to promote to improve the quality (and more recently, equity) of their schooling systems. These policies emphasize: (1) increased competition between schools and systems; (2) setting standards for schools, teachers and students, resulting in increased standardization of practices across these contexts; (3) narrowing the curriculum to focus on core subjects in schools—particularly literacy and numeracy, and science; (4) using corporate management models as a vehicle to drive improvement, and; (5) high stakes testing for accountability purposes as a vehicle to determine the quality of teaching and educational provision more broadly (Sahlberg, 2016). While such processes are sometimes seen as a vehicle to enhance autonomy of schools by reducing more bureaucratic structures and processes, they simultaneously increase external control upon what occurs in schools through stronger emphases upon accountability, increased attention to standards and more frequent testing. These processes are part of what Robertson (2016) describes as the “thickening” of global governance of teachers’ work. Such “thickenin” is evident in both developing and developed contexts through the influence of external bodies such as the World Bank and the OECD. Such bodies have variously advocated new tools for the governance of teachers’ conditions, work and pedagogical practices, which have become increasingly linked to students’ outcomes as a means of maintaining accountability. Tools such as the OECD’s *Teaching and Learning International Survey* (TALIS), the OECD’s Education GPS and the World Bank’s SABER-Teacher (Systems Approach for Better Education Results) all contribute to these global governance processes.

Standardized curriculum

Arguably, these more neoliberal and managerial approaches foster increased standardization of school curriculum provision. Concerns about curriculum have been long-standing issues within educational research with earlier concerns about the “hidden curriculum” (Apple, 1977) continuing to exert influence but now existing alongside concerns about how curriculum have become increasingly homogenized, particularly in light of the increased emphasis and influence of high stakes testing. Construed by Au (2011) as the “new Taylorism”, such approaches emphasize accounts of curriculum as dominated by testing and scripted curricula. These curricula are often of a pre-packaged, corporatized variety, and oriented to providing resources to prepare students to be successful on various high stakes state and national tests that have exerted increasing influence. This is perhaps most obvious in the United States, where curricula autonomy has been reduced since the passing of *No Child Left Behind* in 2001 in particular, and where local school districts across the United States have felt restricted in the nature of the literacy curriculum they feel they are able to provide to their students (Vaughn et al., 2019). Recent research across nine schools in the US reveal the substantive effects of such prescriptive approaches on curriculum reform at state and district levels, and how teachers are marginalized from decision-making in relation to who decides on the nature of the curriculum, and how these decisions are made (Vaughn et al., 2019). This is due in large measure to the way in which teachers have not been adequately represented in decisions about curriculum reform more broadly, as broader neoliberal logics have sought to foster a form of “closure” around what it is possible to teach (Hardy and Uljens, 2018). Furthermore, prescriptive literacy programmes have become increasingly prevalent in the US context with educators pressured to ensure they are teaching according to district and state reading guides (Powell et al., 2017).

However, these logics are always manifest within local settings, and influenced by the particular circumstances in which such ideas “touch down”. In reflecting upon curriculum reform scholarship in Europe, Yates (2016) identifies how researchers seek to make sense of broader “global” influences, but also how curricula are always national or sub-national in character, and how more localized approaches and foci need to be taken into account. Similarly, in their comparative analysis of Australian and United States national curriculum reforms, Savage and O’Connor (2015) argue a broader global policy-scape (Carney, 2012) has exercised influence over time, influencing the nature and flow of transnational policy ideas, including in relation to concerns about economic competitiveness as well as equity. However, these ideas are also always mediated at national and sub-national levels and via the varying conditions of possibility that exist in each context.

Standardized curricula arrangements can also be manifest in the form of more conservative political rationalities. Reflecting a conservative response to an increasingly fractious polity in England, Winter and Mills (2020) refer to how curricula construed as promoting “British values” have become increasingly pervasive in a broader context of concerns about terrorist attacks and potential radicalization of students in English schools. Reflecting the high-stakes nature of such concerns, the notion of “British values”

forms part of the British Government's Counter-Terrorism and Security Act, 2015. In this way, curriculum becomes narrowed to particular government foci and agenda, and often without adequate regard for subsequent effects.

These standardized curricula may also be the product of key non-governmental bodies in specific national settings. [Tarlau and Moeller \(2019\)](#) refer to how philanthropic foundations in the Brazilian context influence policy making for curriculum reform. This is part of a broader trend of increased influence of philanthropic organizations in educational provision more broadly ([Bhanji, 2016](#)). This reflects the ways in which corporate organizations are gaining increased influence in education to varying degrees in different countries around the world ([Ball, 2012](#)).

Of course, how curricula play out are still influenced by the educators involved in their enactment. From a Chinese context, [Fu and Clarke \(2019\)](#) refer to how teachers exercise moral agency even in the context of neoliberal influences in the way they seek to enact curriculum reform. While China's examination-dominated education system continues to exert strong influence, more competitive logics, including more recent neoliberal logics that contribute another dimension to such competitiveness, have been able to be challenged. Such responses build on earlier efforts to try to reduce memorization and promote increased inquiry skills among students, fostering collaboration and more independent thinking on the part of students ([Zhong, 2006](#)). The development of morals among students, together with a broad conception of knowledge and well-being are prioritized over the dominant exam-oriented curriculum even as the latter continue to exert influence, including in its more recent neoliberal instantiations.

Standardized teaching

The proliferation of teaching standards also reflects efforts to control teaching, and the pervasive influence of standardization of pedagogical practices. Such standards are designed to identify key attributes or qualities of "effective" teaching practice. However, such approaches also have the effect of compartmentalizing and atomizing teachers' work, and fail to adequately capture the necessarily creative and unpredictable nature of actual teaching practice. Increasingly, such standards have become associated with teachers' capacity to make sense of various forms of evidence/data about students' learning, particularly in relation to assessment.

The tightening focus upon assessment alongside issues of pedagogy contributes to what [Lewis and Holloway \(2019\)](#) refer to as the datafication of the teaching profession (2019). This is a teaching profession in which an emphasis upon data dominates teaching practice, and in which teachers are construed as those who "profess" data. Such an approach contributes to the reconstitution of teaching in ways that affect both the practices of teachers and how they come to understand their work. In this context, [Lewis and Holloway \(2019\)](#) argue more numeric data become "effective" at changing what is considered most important within the profession/what "counts", and "affective" in the way data constitute new expectations that teachers will constantly speak to—"profess"—the importance of data. Under such circumstances, teaching becomes increasingly "datafied". This process occurs from the earliest stages of schooling; [Bradbury and Roberts-Holmes \(2018\)](#) refer to the datafication of primary and early years' education in what they describe as the "data obsessed" school. Data in these instances become productive of data-driven subjectivities on the part of both students and teachers, as each is exposed to and seeks to be responsive to the broader external push for enhanced outcomes on specific, reified forms of data. Such processes essentially promote quantitative measures of student achievement, simplifying the complexity of actual learning and treating such numbers as proxies for forms of learning that cannot be easily captured or made "visible" to external audiences.

At the same time, teachers' work and learning are also dominated by discourses of "quality" and how to promote quality teaching and quality education. This is manifest in multiple contexts, including contexts as diverse as Norway, Netherlands, Australia ([Hardy et al., 2020](#)) and Chile ([Mizala and Schneider, 2019](#)). [Robertson \(2016\)](#) describes this focus upon "quality" of teachers in education systems around the world as a global phenomenon. While [Robertson \(2016\)](#) argues this global governance of teachers is not new, the way in which it is expressed at present is different, and worthy of further scrutiny. The earlier OECD (2005) report *Teachers Matter* was a key vehicle to focus attention upon teacher quality, with overriding concerns about economic competitiveness driving pressure for reform. More recently, the TALIS, Education GPS and SABER-Teacher have sought to provide much more detailed information about the nature of teachers and their work across national contexts, and some of this information, such as TALIS material, has been brought into conversation with PISA results as part of the engagement with such data ([Robertson, 2016](#)). These data also enable comparisons between countries in ways similar to how PISA results are deployed. However, and reflecting the inherent tensions that characterize efforts to standardize teaching practices, the kinds of teaching practices advocated in countries that have traditionally been construed as more successful in PISA (such as Finland) do not exhibit the sorts of teaching practices distilled within the OECD's programs as characterizing supposedly successful schooling systems ([Sahlberg, 2016](#)).

Further complicating this situation, these more homogenizing processes are always and everywhere influenced by a broad array of sub-national, national and supra-national actors and agencies. By way of example, and drawing upon the Australian context, [Savage and Lewis \(2018\)](#) reveal how the Australian Professional Standards for Teachers (APST), while purportedly "national" in character, reflect the influence of sub-national (New South Wales state government agencies), national (Australian federal government) and supranational bodies (OECD). The resulting "national" standards reflect the influence of each of these bodies, including the impact of various transnational entities. Such standards may be considered a form of global techno-scientific enterprise even as they are also always manifest in place-specific ways ([Lewis et al., 2019](#)).

Standardized assessment

Perhaps the most overt instance of the standardization of schooling is in the area of assessment. Educational governance is increasingly influenced by testing practices and associated enumerative accounts of achievement. Such accounts enable demands for increased performance, particularly as measured against various national and international large-scale standardized tests. In many ways, standardized testing has served as the leading motif of the Global Education Reform Movement (GERM) with its focus upon standardized test results as a key vehicle to ascertain the effectiveness or otherwise of national and sub-national educational systems, and of various curriculum and pedagogical reforms (Sahlberg, 2016). Verger et al. (2019) argue that national large-scale assessments serve as the policy vehicles that knit together the principal GERM policy principles of accountability, standards and school autonomy. Interestingly, Verger et al. (2019) argue these principles have begun to “infect” those countries that had previously been more resistant/“immune” to the push for increased pressures of accountability associated with standardized measures of attainment; Hardy et al. (under review) provide particularly salient insights in this regard in relation to the earlier Finnish “miracle”. Increasingly, national large-scale assessments have been used for accountability purposes and as a means of keeping track of educational outcomes and the extent to which more standardized curricula and learning standards have been implemented. Smith (2016) refers to a “global testing culture” which is characterized by census-style standardized testing tied to high stakes outcomes. Such tests have become increasingly high stakes not just for students but for teachers. In the United States, for example, these stakes are clearly evident in the way pressure to ensure improved student performance of students influences teachers’ understandings of their work. In a comparative analysis of an earlier Bush-administration push for enhanced standards through *No Child Left Behind* (NCLB) legislation, and the later Obama-administration *Race to the Top* (RTTT) emphasis upon enhanced “value-added” measures of teachers’ performance, Holloway and Brass (2018) identify a fundamental shift in teacher subjectivity. This entails a move from being responsive to what was considered an external intrusion upon teachers’ autonomy and sense of professional identity and practice through NCLB, to a sense that the measures put in place through RTTT became significant markers by which teacher came to know themselves and their work, and the quality of this work. Within such discourses, teachers also come to constitute themselves as constantly needing to work on themselves and their deficiencies—as “perpetually imperfect” (Holloway, 2019)—as indicated through these external markers of attainment. Effectively, the accountability logics that inform this process have come to dominate such that teachers’ very sense of self is constituted in terms of the practices and processes of standardized evaluation processes.

Reid (2019) provides interesting insights into the nature and effects of PISA as the archetypal mode of standardized testing. The way in which the test has taken on such status is reflected in how national governments have focused on the test as a vehicle to evaluate educational quality internationally. The Australian government, for example, passed the Australian Education Bill, 2013, which legislated the aspirational goal of Australia being positioned in the top five countries for PISA by 2025 (Reid, 2019). It is the comparative aspect associated with the ranking of such testing that gives PISA its authority. While PISA reports span multiple volumes and thousands of pages, including multiple supplementary files outlining how various numbers have been generated, it is “the production of a few numbers that capture most of the attention” (Sellar et al., 2017, p. 27). While it is the overall rankings of countries that are emphasized, this hides the fallacious nature of such comparisons, even just at a technical level, given the degrees of error that attend such tests; at best, error coefficients make it erroneous to rank countries beyond grouping countries as “high”, “average” and “low” performing (Sellar et al., 2017).

Nevertheless, such international numeric indicators garner much attention. The fixation upon specific numbers constitutes what Muller (2019) describes as the “tyranny of metrics”, with its capacity to capture and constitute the broader social practices in ways that better enable surveillance and regulation of human (in this case, educational) practices. This is the case in relation to all nation-states, but felt particularly keenly in those countries not considered highly “ranked”; the national eye becomes rearticulated through the global eye as comparative logics dominate (Nóvoa and Yariv-Mashal, 2003; Auld and Morris, 2014).

So, what’s the problem *really* represented to be?

Given these broader conditions, after Bacchi (2009), one can rightly argue, “what’s the problem [*really*] represented to be” (Q. 1) in how teachers’ work and learning are understood? This overview of broad insights into the nature of curriculum, pedagogical and assessment practices within and across varied contexts reveals the “problem” as one of a lack of educational quality as understood against key, comparative numeric indicators, and the need for increasingly standardized curriculum and teaching practices to foster enhanced outcomes—particularly as measured against national and international standardized assessment instruments. Concerns about teachers’ abilities to develop substantive curriculum are reflected in calls for more standardized curricula. At the same time, the emphasis upon standards development and more standardized and increasingly scripted teaching practices (e.g., “direct instruction” approaches) reveals a lack of trust in teachers’ pedagogies, and the need to isolate and codify what are construed as “effective” practices. On Bacchi (2009) analysis, the “problem” is represented to be insufficient rigorous and detailed curriculum that can be readily “transposed” across multiple and diverse contexts, to inform insufficiently knowledgeable teachers. The “problem” is one of “quality” and “standards” in relation to pedagogical/teaching practice. Finally, in relation to assessment, the “problem” also appears to be principally performative, and one of inadequate evidence of sufficient and competitive attainment on standardized national and international literacy and numeracy (and increasingly, science) tests.

In relation to [Bacchi \(2009\)](#) second question about the assumptions that underlie such an approach (Q. 2), such representations assume that shortcomings in curricula, teaching and assessment provision can be rectified through this higher level of curricula specificity, increased attention to the nature of students accepted into teaching and their subsequent pedagogical practices, and even more focus upon testing of “the basics” of literacy and numeracy. What is less recognized is the complexity of teaching practice *in situ*, and how teaching is always a grounded practice requiring substantive judgment and trust in teachers as professionals and “intelligent” systems of accountability to evaluate these practices ([O’Neill, 2013](#)); what is also lacking is a more developed sense of trustworthiness in general, particularly in relation to issues of honesty, competence and reliability ([O’Neill, 2018](#)).

How we have come to this point is an interesting question (Q. 3), and one that rewards close scrutiny of the litany of reviews and reforms that characterize educational provision throughout the world. The principles of the Global Education Reform Movement (GERM) have clearly had significant influence ([Sahlberg, 2016](#)), including in supposedly more successful jurisdictions, such as Finland ([Hardy et al., under review](#)). Constant processes of comparison between nation-states’ educational performance against standardized and limited measures of student achievement have led to advocacy for similarly standardized and limited approaches to curriculum and teaching—all in an effort to better control, so as to supposedly improve, teachers’ work and learning.

On the question of what is left unsaid through such problematization (Q. 4), this approach fails to recognize the limitations of more restricted curriculum, teaching and assessment practices. It assumes that increased standardization of curriculum will somehow automatically result in enhanced learning on the part of students, without adequate regard for the nature of the teaching transpiring, or actual evidence of its effects. The contrary effects created by the “data obsessed school” ([Bradbury and Roberts-Holmes, 2018](#)), and the unrelenting focus upon “perpetually imperfect teachers” ([Holloway, 2019](#)) are given less attention. Such an approach also fails to adequately consider the importance of context, particularly for those students who are most disadvantaged and for whom teaching practices based on constant efforts to increase standardized test results leads to an enervating experience of schooling that can hardly be described as “educational”. Such processes are exacerbated by what [Biesta \(2015\)](#) refers to as the social psychology of international large scale assessments such as PISA, with their emphasis upon the importance and value of numeric and comparative measures of achievement, the persistence of technical renderings of education, and the way in which concerns about social justice are deployed as a key justification for trying to understand and capture how schooling systems are performing.

Such a response results in unintended and often problematic effects (Q. 5). These include policy settings that seem impervious to evidence to the contrary. In the Australian context, for example, a relative decline in students’ literacy, numeracy and science results in the latest (2018) PISA results has led to increased calls for even more attention to “back-to-basics” curricula, and teaching approaches that are “evidence-based”, where, in a process of abject circularity, such evidence is limited to results on yet more standardized literacy and numeracy tests. This is the case even as there are always instances of challenges to these problematic practices (e.g., [Zhong, 2006](#)).

The issue of how/where this representation of the “problem” has been produced, disseminated and defended (Q. 6) focuses attention upon the broader politics that surround schooling provision around the world at present. Again, drawing upon the Australian case as an example, the federal government’s concern and anxiety to ensure improved outcomes in both national and international test results serves as a powerful vehicle for the continued focus upon such results, and as a challenge to alternative, more critical discourses. Even as more standardized approaches are revealed as having deeply problematic effects ([Reid, 2019](#)), they continue to be supported by governments seeking to find ways to “capture” and “improve” educational provision in schools. That this is a global problem is reflected in the way international organizations have played a key role in such codification processes. This includes through the OECD’s Teaching and Learning International Survey (TALIS), Education GPS and the World Bank’s SABER-Teacher (Systems Approach for Better Education Results)—all of which contribute to these broader global governance processes ([Robertson, 2016](#)).

Broad principles to inform an alternative approach to educational policymaking

However, even as there is evidence of such homogenizing processes, more critical approaches and alternative responses to standardization agenda exist. The final part of [Bacchi \(2009\)](#) final question (Q. 6)—How could this “problem” be questioned, disrupted and replaced? (p. xii)—demands consideration of the possibilities for alternatives. The focus upon more reductive approaches to engaging with data has not gone unchallenged. Adopting a more materialist approach, Madsen (2019) argues for a “configurative agency of metrics” as a mechanism to rearticulate excessive emphases upon the more performative processes of “governing by numbers” and various processes of governing education through data ([Ozga, 2009](#)). Such an approach is construed as a way for actors involved in education to engage with data more substantively.

Such approaches do not simply denigrate more quantified approaches to understanding student and teacher work and learning; datafication processes are not homogeneous but also reflective of the particular social, political, cultural and material conditions within which they are exercised and that contribute to their very development. As [Takayama and Lingard \(2018\)](#) argue, while existing studies:

offer valuable insights into the processes and effects of educational datafication, the lack of national and regional diversity could obscure the fact that datafication of schooling elsewhere might take different forms and assume different ontological meanings. More problematically, it could constrain our ability to explore different ways in which datafication can be conceptualized and its effects understood and to ask a different set of questions that are perhaps more relevant to other national and regional contexts (p. 3).

However, this is never a simple process. In recent work (Hardy, 2020), I have referred to the “quandary of quantification” to capture the way in which processes of the quantification of education and more performative logics intersect, constituting a problematic set of circumstances within which more enumerative renderings of practice may be understood in simply reductive terms. At the same time more quantified approaches can focus attention upon some areas of teachers’ work and learning that can benefit from further inquiry and scrutiny. But this must be undertaken in context and in the clear understanding that standardized measures of educational attainment cannot capture important and necessary practices that are genuinely educational. Any social practice, including practices of teachers’ work and learning are always undertaken in specific settings; practice is never divorced from context. Philosopher Theodore Schatzki (2002) captures this situatedness of practice through his conception of practice as situated—the “site of the social”—and Kemmis et al. (2014) have taken up this situatedness of practice in their advocacy for changed conditions for education provision *in situ*.

Such an approach is cognizant that some forms of standardized data may be able to be deployed for educative purposes. It is not simply that data derived from standardized tests are inherently “bad” but rather that such data are often/frequently used for problematic purposes, including for purposes for which they were never originally intended. This is most obvious in relation to various standardized national literacy and numeracy tests that are described as providing some policy-relevant insights into the effects of educational provision at a national level, but which may then be used to evaluate individual students, teachers and schools. Such unintended consequences have become so common in some settings as to be normalized, resulting in those in schools—students, teachers, school administrators—accepting that they should be held to account for such results, even as they simultaneously recognize the limited and reductive nature of such measures of attainment; Holloway (2019) analysis teacher evaluations as an ontological framework that dominates teachers’ practices is particularly salient in this regard.

Of course, the challenge for educators as processes of standardization and accountability exert ever greater influence is how to balance the tension that such approaches produce with the possibilities that may arise to use such data for more substantive educational purposes. As Reid (2019) argues, processes of accountability remain important, but not in the way they are currently constituted, which are antithetical to the forms of educational provision that need to characterize educators’ practices as they strive to ensure their students are as ready as possible for the uncertainties of the worlds in which they will live and work:

It is important to make the point at the outset that a critique of the neoliberal accountability regime is not an argument against accountability. Gathering evidence—both quantitative and qualitative—is crucial in order to inform all those who have a stake in the quality of our education systems about how things are going. It assures taxpayers that government funds are being spent wisely and appropriately; advises parents about the quality of their child’s education; tells educators about areas that need work; and assists policy-makers in making decisions about such matters as resource allocation. But it is the form that accountability takes that is at stake here. It is argued that when approaches to accountability are narrow and pinched—as are those associated with the standardizing agenda—they can only serve to severely constrain what is needed for a futures-focused education. That is, it is pointless devising an educational program for the future, only to allow it to be sabotaged by accountability strategies that are incompatible with it (p. 79–80).

Importantly, even as broader more performative policy logics have significant impact, counter-narratives that challenge the standardization phenomenon are evident, and exert influence. In his case for creating curriculum for an uncertain future, Tienken (2017) calls for American educators to “defy standardization” as they seek to foster curricula opportunities that take into account specific district, school, classroom and student needs. With their concern for how to foster more socially just schooling practices, Francis et al. (2017) also point to the need to engage with issues of social justice, even as such issues are fraught with contradictions and dilemmas. Gert Biesta (2019) call for “obstinate education”, which seeks to ask difficult questions about society and sustainability in the quest for more democratic and just practices, fosters the sorts of counter-narratives that have the potential to be more genuinely educational for all—not just those already privileged or able to garner the most resources, often at the expense of social, political, economic and environmental sustainability.

Conclusion

Consequently, an analysis of teachers’ work and learning under current policy conditions is powerfully informed by focused attention upon the nature of curriculum, pedagogical and assessment practices in schooling policy provision, and how such provision is enacted in practice. Bacchi (2009) provides a particularly useful analytical toolbox for better understanding the nature of education provision in and through policy, and as a stimulus for considering how to foster more substantive policy settings for alternative practices. Her insights reveal increased pressure for control over curriculum and teaching/pedagogy, and how these have been achieved through more standardized assessment practices for accountability purposes. These accountability purposes are often infused by performative rather than educative logics, and are less “intelligent” in their orientation than they need to be (O’Neill, 2013, 2018). Consequently, the “problem” of the need for increased “standardization” actually reveals a need for the very opposite—for space, and opportunity for teachers to develop context-responsive curricula and to be able to teach such curricula in engaging ways, particularly for those students most marginalized by dominant schooling

practices. Such an approach helps set the groundwork for more “authentic accountabilities” (Hardy et al., 2019; Hardy, 2021), which will rearticulate the “problem” of educational provision in schooling settings toward social, political, economic and ecological sustainability, and better focus our attention upon practices more likely to lead to enhanced learning toward such sustainability for both teachers and students.

References

- Apple, M., 1977. *Ideology and Curriculum*. Routledge & Kegan Paul, London.
- Auld, E., Morris, P., 2014. Comparative education, the “New Paradigm” and policy borrowing: constructing knowledge for educational reform. *Comp. Educ.* 50 (2), 129–155.
- Bacchi, C., 2009. *Analysing Policy: What’s the Problem Represented to Be?* Pearson, Frenchs Forest.
- Bacchi, C., 2012. Introducing the “What’s the problem represented to be?” approach. In: Betsans, A., Beasley, C. (Eds.), *Engaging with Carol Bacchi: Strategic Interventions and Exchanges*. University of Adelaide Press, Adelaide, pp. 21–24.
- Ball, S., 1994. *Educational Reform: A Critical and Post-structural Approach*. Open University Press, Buckingham.
- Ball, S., 2012. *Global Education Inc.: New Policy Networks and the Neo-liberal Imaginary*. Routledge, London.
- Bernstein, B., 1971. On the classification and framing of educational knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control: New Directions for the Sociology of Education*. Collier-Macmillan, London, pp. 47–69.
- Bhanji, Z., 2016. The business case for transnational corporate participation, profits and policy-making in education. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley, pp. 419–432.
- Biesta, G., 2015. Resisting the seduction of the global education measurement industry: notes on the social psychology of PISA. *Ethics Educ.* 10 (3), 348–360.
- Biesta, G., 2019. *Obstinate Education: Reconnecting School and Society*. Brill, Leiden.
- Bradbury, A., Roberts-Holmes, G., 2018. *The Datafication of Primary and Early-Years Education: Playing with Numbers*. Routledge, Abingdon.
- Carney, S., 2012. Imagining globalization: educational policy-scapes. In: Steiner-Khamsi, G., Waldo, F. (Eds.), *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. Routledge, New York, pp. 339–353.
- de Sousa Santos, B., 2006. Globalizations. *Theor. Cult. Soc.* 23, 393–399.
- Easton, D., 1981/1953. *The Political System: An Inquiry Into the State of Political Science*. University of Chicago Press, Chicago.
- Francis, B., Mills, M., Lupton, R., 2017. Towards social justice in education: contradictions and dilemmas. *J. Educ. Pol.* 32 (4), 414–431.
- Fu, G., Clarke, A., 2019. Teachers’ moral agency under neoliberal influences: what is educationally desirable in China’s curriculum reform? *Educ. Rev.* 71 (1), 51–66.
- Hardy, I., Uijens, M., 2018. Critiquing curriculum policy reform in Finland and Australia: a non-affirmative and praxis-oriented approach. *Tci* 15 (2), 50–82.
- Hardy, I., Reyes, V., Obaid Hamid, M., 2019. Performative practices and “authentic accountabilities”: targeting students, targeting learning? *Int. Educ. J.* 18 (1), 20–33.
- Hardy, I., Jakhell, R., Smit, B., 2020. The policies and politics of teachers’ initial learning: the complexity of national initial teacher education policies. *Teach. Educ.* <https://doi.org/10.1080/10476210.2020.1729115>.
- Hardy, I., Heikkinen, H., Pennanen, M., Salo, P., Kilakoski, K., 2020. The “spirit of the times”: fast policy for educational reform in Finland. *Pol. Futures. Educ. Internet*.
- Hardy, I., 2014. A logic of appropriation: enacting national testing (NAPLAN) in Australia. *J. Educ. Pol.* 29 (1), 1–18.
- Hardy, I., 2015. A logic of enumeration: the nature and effects of national literacy and numeracy testing in Australia. *J. Educ. Pol.* 30 (3), 335–362.
- Hardy, I., 2020. The quandary of quantification. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2019.1672211>.
- Hardy, I., 2021. *School Reform in an Era of Standardization: Authentic Accountabilities*. Routledge, London & New York.
- Holloway, J., Brass, J., 2018. Making accountable teachers: the terrors and pleasures of performativity. *J. Educ. Pol.* 33 (3), 361–382.
- Holloway, J., 2019. Teacher evaluation as an onto-epistemic framework. *Br. J. Sociol. Educ.* 40 (2), 174–189.
- Kemmis, S., Wilkinson, J., Edwards-Groves, C., Hardy, I., Grootenboer, P., Brisol, L., 2014. *Changing Practices, Changing Education*. Springer, Singapore.
- Lewis, S., Holloway, J., 2019. Datafying the teaching “profession”: remaking the professional teacher in the image of data. *Camb. J. Educ.* 49 (1), 35–51.
- Lewis, S., Savage, G., Holloway, J., 2019. Standards without standardization? Assembling standards-based reforms in Australian and US schooling. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2019.1636140>.
- Mizala, A., Schneider, B., 2019. Promotion quality education in Chile: the politics of reforming teacher careers. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2019.1585577>.
- Muller, J., 2019. *The tyranny of metrics*. Princeton University Press, Princeton, NJ.
- Nóvoa, A., Yaviv-Mashal, T., 2003. Comparative research in education: a mode of governance or a historical journey? *Comp. Educ.* 39 (4), 423–438.
- O’Neill, O., 2013. Intelligent accountability in education. *Oxf. Rev. Educ.* 39, 4–16.
- O’Neill, O., 2018. Linking trust to trustworthiness. *Int. J. Phil. Stud.* 26 (2), 293–300.
- Ozga, J., 2009. Governing education through data in England: from regulation to self-evaluation. *J. Educ. Pol.* 24 (2), 149–162.
- Peck, J., 2010. *Constructions of Neoliberal Reason*. Oxford University Press, Oxford.
- Powell, R., Cantrell, S., Correll, P., 2017. Power and agency in a high poverty elementary school: how teachers experienced a scripted reading program. *J. Lang. Literacy Educ.* 13 (1), 93–124.
- Power, M., 1997. *The Audit Society: Rituals of Verification*. Oxford University Press, Oxford.
- Reid, A., 2019. *Changing Australian Education: How Policy is Taking Us Backwards and What Can Be Done About it*. Allen & Unwin, Crows Nest.
- Robertson, S., 2016. The global governance of teachers’ work. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley-Blackwell, West-Sussex, pp. 275–290.
- Rowe, E., Lubienski, C., Skourdombis, A., Gerrard, J., Hursh, D., 2019. Exploring alternatives to the “neoliberalism” critique: new language for contemporary global reform. *Discourse* 40 (2), 147–149.
- Sahlberg, P., 2016. The global educational reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley-Blackwell, West Sussex, pp. 128–144.
- Savage, G., O’Connor, K., 2015. National agendas in global times: curriculum reforms in Australia and the USA since the 1980s. *J. Educ. Pol.* 30 (5), 609–630.
- Savage, G., Lewis, S., 2018. The phantom national? Assembling national teaching standards in Australia’s federal system. *J. Educ. Policy* 33 (1), 118–142.
- Schatzki, T., 2002. *The Site of the Social: A Philosophical Account of Constitution of Social Life and Change*. The Pennsylvania State University Press, University Park, PA.
- Sellar, S., Thompson, G., Rutkowski, D., 2017. *The Global Education Race: Taking the Measure of PISA and International Testing*. Brush Education Inc.
- Shore, C., Wright, S., 2015. Audit culture revisited. *Curr. Anthropol.* 56 (3), 421–444.
- Smith, W., 2016. An introduction to the global testing culture. In: Smith, W. (Ed.), *The Global Testing Culture: Shaping Educational Policy, Perceptions and Practices*. Symposium, Oxford.
- Strathern, M., 2000. Introduction. In: Strathern, M. (Ed.), *Audit Cultures. Anthropological Studies in Accountability, Ethics and the Academy*. Routledge, London, pp. 1–18.
- Takayma, K., Lingard, B., 2018. Datafication of schooling in Japan: an epistemic critique through the “problem of Japanese education”. *J. Educ. Pol.* 34 (4), 449–469.
- Tarlau, R., Moeller, K., 2019. ‘Philanthropizing’ consent: how a private foundation pushed through national learning standards in Brazil. *J. Educ. Policy* 35 (3), 337–366.
- Tienken, C., 2017. *Defying Standardization: Creating Curriculum for an Uncertain Future*. Rowman & Littlefield, Lanham, MD.

- Vaughn, M., Scales, R., Stevens, E., Kline, S., Barrett-Tatum, J., Van Wig, A., Yoder, K., Wellman, D., 2019. Understanding literacy adoption policies across contexts: a multi-state examination of literacy curriculum decision-making. *J. Curric. Stud.* <https://doi.org/10.1080/00220272.2019.1683233>.
- Verger, A., Parcerisa, L., Fontdevila, C., 2019. The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms. *Educ. Rev.* 71 (1), 5–30.
- Winter, C., Mills, C., 2020. The psy-security-curriculum ensemble: British values curriculum policy in English schools. *J. Educ. Pol.* 35 (1), 46–67. <https://doi.org/10.1080/02680939.2018.1493621>.
- Yates, L., 2016. Europe, transnational curriculum movements and comparative curriculum theorizing. *Eur. Educ. Res. J.* 15 (3), 366–373.
- Zhong, Q.-Q., 2006. Curriculum reform in China: challenges and reflections. *Front. Educ. China* 1 (3), 370–382.

Teachers' work amid global education reform: "the greatest challenge"?

Nicole Mockler^a and Meghan Stacey^b, ^a Sydney School of Education and Social Work, University of Sydney, NSW, Australia; and ^b School of Education, University of New South Wales, Sydney, Australia

© 2023 Elsevier Ltd. All rights reserved.

Background	30
The Australian context	30
The global education reform movement	31
The GERM and education reform in Australia	31
Competition and choice	31
Standardization of teaching and learning	32
Increased focus on reading, mathematics, and science	33
Corporate models of change	33
Test-based accountability	34
The GERM and policy effects: reshaping the practice architectures of teachers' work	34
Cultural-discursive arrangements	35
Material-economic arrangements	35
Social-political arrangements	35
Conclusion	36
References	37

The past three decades have brought educational reform on an unprecedented international scale. The advent of the "information age" and "network society" (Castells, 2010), wherein "networks constitute the new social morphology of our societies, and the diffusion of networking logic substantially modifies the operation and outcomes in processes of production, experience, power, and culture" (Castells, 2010, p. 500) has seen "educational borrowing and lending" (Steiner-Khamsi, 2004), enacted at an ever increasing pace, with policy flows ever more global in scale. At the same time, the increasing tendency of Western democracies toward neoliberal ideas in education and other areas of social policy has seen a growing fixation on the measuring and auditing of all things educational, as responsibility—but not necessarily control—over educational outcomes is shifted to the level of the local and the individual. Teachers' work has become an important point of focus in this continuing reform of education systems internationally, with the re-emergence of "teacher centrality" (Larsen, 2010, p. 208), "the assumption that the key to educational success lies with the teacher", as a driver of reform.

The contemporary manifestation of teacher centrality has been leveraged variously by national Governments (e.g., Australian Government, 2016; Department for Education UK, 2016), prominent shapers of teachers' professional knowledge and development internationally (Hattie, 2008), and other key players such that it has become consolidated as a key "regime of truth" (Foucault, 1980) within the past 30 years. In a recent OECD report, for example, Andreas Schleicher stated "it is a given that the quality of an education system can never exceed the quality of its teachers. So, attracting, developing and retaining the best teachers is the greatest challenge education systems have to face" (OECD, 2019, p. 4). Explicit efforts aimed at improving teaching practice and "quality", including upon entry to the profession, therefore constitute one aspect of the policy context for teachers. Also important, however, are reforms which while not ostensibly targeted at teachers, nonetheless impact on and shape their work, including, for example, moves toward the standardization of curriculum, pedagogy and assessment, which have been a key part of the global reform agenda.

In this article we examine the implications of the key tenets of the global reform agenda for teachers' work, taking the Australian context, where ongoing changes to education policy reflect a range of core concerns, as a case study. We suggest that local articulations of globalized education policy reshape the "practice architectures" of teachers' work, impacting the cultural-discursive, material-economic and social-political arrangements that "constitute enabling and constraining preconditions for the conduct of practices" (Kemmis et al., 2014, p. 31). These arrangements effectively operate to govern such work, opening up and closing down different possibilities of practice. While we do not assume any kind of prior "golden age" of teacher autonomy or empowerment—teachers "have always been under the yoke of surveillance and normalizing state control" (Smaller, 2015, p. 150–151)—we do argue that the current articulation of teacher centrality has particular implications for their work within the present historical moment. Through reference to the Australian experience, we posit that the sum total of these reforms looks less like a pursuit of "teacher quality" than a reshaping of their role, labor and discretion along lines of performativity and competition, unlikely to assist in the goal of "attracting, developing and retaining" (OECD, 2019, p. 4) outstanding teachers.

The article proceeds in three parts. In the first, we locate our discussion of Australian reform of teachers' work with respect to both the Australian national and educational context and the Global Education Reform Movement (GERM) (Sahlberg, 2016). In the second, we consider the aspects of contemporary reform in Australia as they connect to the tenets of the GERM, and their implications for the work of teachers. In the final section, we take a step back to consider some broader implications of these reforms, using

the theory of practice architectures as a touchstone, in terms of the reshaping of the arrangements that govern teachers' work and the implications of these for the teaching profession.

Background

The Australian context

Historically, school education in Australia is a responsibility of the states, a residual power retained after federation in 1901. Slowly but surely, this settlement has been eroded, particularly since the 1970s with increasing federal intrusion into state-run schooling (Campbell and Proctor, 2014). Perhaps the most formal avenue for such intrusion is the Education Council. While a series of inter-governmental forums for discussion of education have been in place since early in Australia's history as a nation, the establishment of the Council of Australian Governments, or COAG, was a key development. Under this umbrella, the intergovernmental Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA) was established in 1993, the body now referred to more simply as the Education Council (educationcouncil.edu.au). The Education Council is an important forum, producing National Declarations on schooling successively since 1989, with the most recent being the Mparntwe or Alice Springs Declaration (2019), successor to the Melbourne Declaration (2008). These documents remain important guiding statements on national school education.

Under the auspices of these policy documents have come a range of other federal moves in the Australian schooling space. One of the most notable has been the influence of the Rudd/Gillard Labor Government from 2007 to 2013, which came to power on an "education revolution" electoral platform and pursued vigorous reform in relation to schooling during its tenure. Under this government pre-existing, nascent plans for a national curriculum and nation-wide standardized testing were amplified, passing the Australian Curriculum, Assessment and Reporting Authority (ACARA) Act in 2008 enabling the establishment of ACARA as a significant national body. Via this "new national policy organization" (Savage, 2016, p. 837) the Australian Curriculum was progressed, and the national standardized testing program (National Assessment Program—Literacy and Numeracy, or NAPLAN) was established, its results published on a new national, comparative choice-based website, My School (www.myschool.edu.au). The Rudd/Gillard government also boosted the power of the Australian Institute for Teaching and School Leadership (AITSL), a body responsible for the management of teacher accreditation via nationally agreed teaching standards which had previously been state-based. This government also established a significant review of school funding popularly known as the Gonski Review (Gonski et al., 2011), which led to the passing of the Australian Education Act 2013—the first act of federal parliament to govern school education—and a significant revision to funding mechanisms for Australian schools, both government and non-government. Subsequent to this initial Gonski review, however, was what has become popularly known as "Gonski 2.0" (Gonski et al., 2018), primarily a review of how government funding for schools should be spent. Gonski 2.0 was initiated by the subsequent (and, at time of writing, current) Liberal-National Coalition government, elected in 2013, which was also responsible for the establishment of the Teacher Education Ministerial Advisory Group (TEMAG), reflecting a shift of focus from schools to how their teachers are educated within the tertiary sector.

It is this group of government bodies: the Education Council, ACARA and AITSL—supported by adhoc groups and committees such as TEMAG and the reviews of school funding committees including Gonski and Gonski 2.0—that constitute the main organizations responsible for shaping federal education policy in Australia today. Fig. 1 highlights the relationships between federal and state and territory governments and these bodies, and the key policy reforms each has been responsible for since 2007.

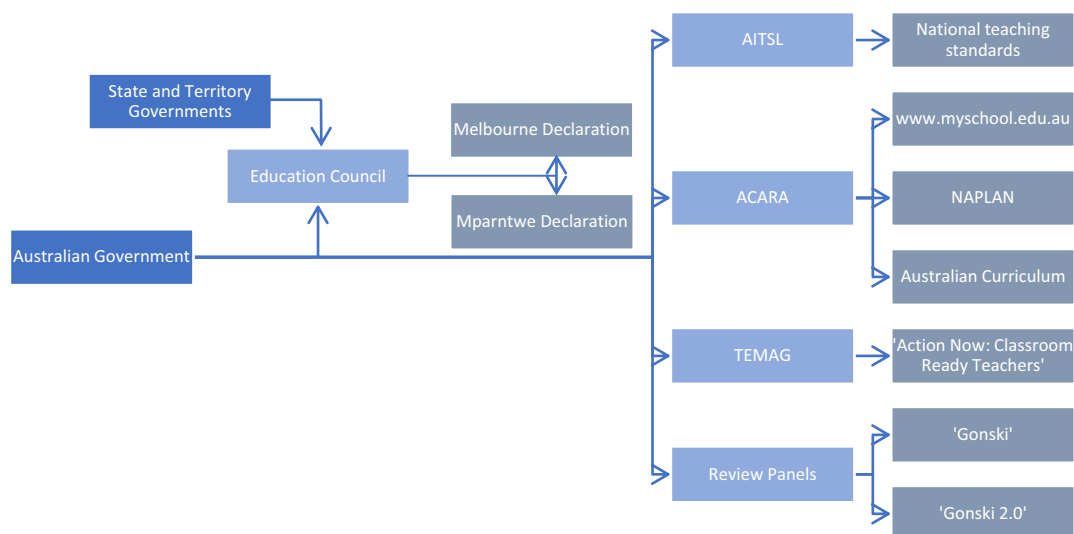


Fig. 1 Key Australian education policy reforms since 2007.

The global education reform movement

In the discussion that follows, we use Sahlberg's (2016) notion of the Global Education Reform Movement, or GERM, as an organizing framework for understanding education reform in Australia and articulating this on an internationally comparative scale. The GERM is said by Sahlberg to consist of five policy drivers: competition and choice; standardization of teaching and learning; focus on reading, mathematics, and science; corporate models of change; and test-based accountability. Sahlberg's acronym and model has spurred broad academic interest. A recent special issue of *Educational Review* similarly uses the GERM as an organizing framework (Fuller and Stevenson, 2019). In this section, we provide a brief and necessarily limited summary of some of the ways in which the GERM has been addressed within the international research literature, in relation to teachers' work.

When it comes to teachers and their work as a subject of the GERM, the international literature has tended to focus mostly on the issues of standardization, corporate models of change and test-based accountability. In relation to the standardization of teaching and learning, it is evident that the use of teaching standards is now ubiquitous (Darling-Hammond, 2012; Storey, 2006). As Sahlberg (2016, p. 142) writes, there is an assumption within GERM-related reform "that externally set teaching standards and aligned standardized testing are preconditions for success". Another arena of standardization has been within initial teacher education, with the accreditation not only of teachers but of teacher education programmes and ongoing professional development beyond initial teacher education (see e.g., Conway, 2013; Mockler, 2022). In other contexts, the "failure" of teacher education has seen the opening up of particular forms of "entrepreneurship" around alternative pathways into teaching, such as in England (Ellis et al., 2019). Elsewhere in higher education policy we have seen the introduction of Teacher Performance Assessments, argued to have arisen from "a construction of the problem of teacher education [as] lack of accountability and standardization of expectations" (Cochran-Smith et al., 2013, p. 16), and to have been based on a logic of professionalization that can, for instance, be in conflict with conceptual ideals of critical pedagogy (Sato, 2014).

Standardization of student learning has also been evident in the introduction of testing frameworks which can narrow the focus on in-school teaching and learning, with research in the United States suggesting a shift to "teaching to the test" and engaging in more teacher-centered pedagogical methods (Au, 2007; Sahlberg, 2016). This also relates to the function of corporate models of change, where research has suggested that teachers now work within forms of managerial rather than regulatory accountability in national contexts as diverse as the United States, Mexico, Chile, Brazil, China, and Germany (Holloway et al., 2017), increasingly through the use of student test scores via national large-scale assessments, as another interwoven feature of the GERM (Verger et al., 2019). Perhaps the most explicit and drastic example of test-based accountability for teachers has been seen primarily within the United States, with the controversial use of student test results as measures of teachers' work through the use of "Value-Added Modeling" tied directly to individual teachers, but with considerable problems of validity and reliability (Berliner, 2013). Also relevant in relation to corporate models of change has been devolutionary reform, as has been seen for instance in Sweden, and how this can lead to differentiated conditions for teachers (Parding et al., 2020). Globally, the international literature has been less interested in the impacts of devolution on teachers than on students or broader questions of social justice—with some exceptions (see e.g., recent research out of China critiquing the effects of decentralization and external accountability on teacher agency (Fu and Clarke, 2019)).

As we noted in our introduction, however, it is not only these aspects of reform most closely related to teaching and learning that have impacts on teachers' work. Other policy technologies, including those related to competition and choice between schools and school systems, and a focus on reading, mathematics and science, have also have effects for teachers—for instance as Sahlberg (2016, p. 138) describes, increased competition between schools can affect "the cultures of teaching". Such modes of impact on teachers have seen less attention in the international literature. In this article, we seek to take a broad view of policy informing and impacting teachers' work, and look across the major reforms in Australia to explore these effects. In particular, we focus on the major reforms to school education identified at federal level and highlighted in Fig. 1, which constitute the central and most significant national reforms over the past decade. In doing so, we highlight both consonance and dissonance across the Australian teacher policy settlement, and how this may be understood on a global scale.

The GERM and education reform in Australia

In this section, we canvas particular manifestations of the GERM in Australian school education and review the research on these policy shifts. In doing so, we also aim to highlight where and how teachers have featured within these shifts and debates. While we have structured the discussion below around the five central tenets of the GERM, we recognize that in practice, these tenets are not discrete. Test-based accountability, for example, might be seen as a technology of competition and choice, linked also to the standardization of teaching and learning. For this reason, while we have used the cultural-discursive, material-economic and social-political arrangements suggested by these tenets as a touchstone in the discussion, in the final section of the paper we provide a more integrated discussion of the practice architectures of teachers' work as suggested by the GERM in Australia.

Competition and choice

Competition and choice is a central aspect of the school education policy settlement in Australia, which boasts an "extreme" form of marketized schooling via a large, fee-charging yet publicly-funded private sector complemented by considerable differentiation

across public schools (Windle, 2015). This settlement has primarily attracted attention in terms of its impacts on schools, students and parents. Competition and choice are theoretically meant to have a social-political effect of "lifting all boats" (Chubb and Moe, 1990). The Australian research literature, however, has primarily documented the residualization of schools that are not "chosen" (Campbell and Sherington, 2013), exacerbating compositional effects for students increasingly segregated by level and nature of advantage (Bonnor and Shepherd, 2016), as parents (typically those of the middle class) engage in considerable, often gendered forms of labor (Proctor and Aitchison, 2015) to make the "right" school choice for their children (Campbell et al., 2009). Notably, this settlement remains despite the recent, large review of school funding known as the "Gonski Review" (Gonski et al., 2011), which while critiquing the ways in which funding was distributed across public and private schools did not ultimately question the provision of such funding and maintenance of a publicly-funded choice system in itself (Morsy et al., 2014).

Other research relating to choice and competition has documented impacts on teachers. With student populations segregated across school settings, "monolithic notions of the teaching experience" (Stacey, 2020, p. 164) have been shown to be inadequate, as vastly different school contexts constitute contrasting places of employment, both in a material-economic and social-political sense. Research suggests that in this divided system, teachers can experience both structural and pedagogical "lock-in" effects that limit their movement across different types of schools (Parding et al., 2017), as they come to know and feel comfortable within particular kinds of settings (Forsey, 2010; Stacey, 2020). There have also been shifts to the kinds of work undertaken in schools in relation to market positioning (Stacey, 2020).

An important part of how school choice and competition are fostered has been the advent of the My School website which presents, for each school, the results of national, standardized testing via NAPLAN (discussed further below). This move has arguably taken the technology of NAPLAN, which was designed to provide a national snapshot of student achievement, to a mode of accountability, with the public presentation of results described explicitly at the time as intended to encourage parents to "vote with their feet" (AustralianPolitics.com, 2008), driving competition between schools and their teachers. A wealth of research since has outlined these cultural-discursive, material-economic and social-political shifts and the effects on teachers of this focus on the core subjects which constitute such standardized measures, as discussed further below.

The emphasis on competition between schools that is epitomized within the Australian school choice regime is arguably consonant with a broader theme of competition between nations, also evident in the national (and international) focus on PISA rankings, where education is positioned as a tool for improved economic success in global terms. This cultural-discursive arrangement, or way of talking and thinking about the role of education, is expressed in the most recent national declarations on Australian schooling. In the 2008 "Melbourne Declaration" (Ministerial Council on Education Employment Training and Youth Affairs, 2008), some commentators have seen a primarily economic driver in the discursive constructions of the document (Buchanan and Chapman, 2011). While other analyses have seen this as balanced with a strengthening of "the individual's tendency toward compassion, the development of contemplative capacities, perseverance and open-mindedness" (Carter, 2019, p. 133), Savage (2011) has argued that the long-articulated dual goals of "equity" and "excellence" in education sit uncomfortably next to, and possibly work in tension with, each other. This cultural-discursive settlement, with attendant impacts on social-political arrangements via the instrumentalization of teachers' work, appears to be persistent. Early analysis of the subsequent Alice Springs (Mparntwe) Education Declaration (Education Council, 2019), suggests that it essentially continues the economic and human capital focus of its predecessor (Hogarth, 2020).

Standardization of teaching and learning

Similarly, moves toward standardization of learning and teaching, embodied predominantly in the Australian Curriculum, and TEMAG and Gonski 2.0 reports, suggest particular cultural-discursive, material-economic and social-political arrangements that frame and govern teachers' work, forming part of the "practice landscape" (Kemmis et al., 2014) on which their work is enacted. Indeed, to a large extent, the suite of policy reforms is underpinned by an implicit assumption that increased standardization is of benefit to students and student learning, and will shore up "teacher quality" in Australian schools.

Since its inception, the Australian Curriculum has been positioned as a "learning entitlement" for all Australian students (National Curriculum Board, 2009, p. 11), and this language of "entitlement" has increasingly been amplified in successive *Shape of the Australian Curriculum* papers (ACARA, 2010, 2011, 2012) such that in the current iteration of the document it is expressed as follows:

The Australian Curriculum describes a learning entitlement for each Australian student. It sets out what young people should be taught ... and an expectation of the quality of their learning. (2012, p. 15)

Further, the curriculum "makes clear to teachers what is to be taught" (ACARA, 2012, p. 13). This, together with the language of entitlement, points to a shrinking "discretionary space" (Comber and Kamler, 2009; Weaven and Clark, 2015) for teacher judgment and decision-making and, as Brass and Holloway (2021, p. 5) have argued in relation to the United States, a consequent displacement of teachers' "professional discretion in matters of curriculum and pedagogy" at the hands of the national curriculum.

Such standardization has of course been experienced differently across the eight Australian states and territories, whose governments hold responsibility for the provision of schooling and the oversight of curriculum. In New South Wales, traditionally thought

of as having the highest level of curriculum standardization and regulation prior to the Australian Curriculum (Reid, 2018), the shift was far more minor than other states, such as Queensland, where the adoption of the Australian Curriculum gave rise to state-developed resources known as the *Curriculum into the Classroom*, with scripted lesson plans mandated by many school principals (Barton et al., 2014; Wall, 2017).

In both the TEMAG *Action Now: Classroom Ready Teachers* (Craven et al., 2015) and "Gonski 2.0" *Through Growth to Achievement* (Gonski et al., 2018) reports, standardization of teaching and learning is linked discursively to the drive toward "teacher quality", which is underpinned by a commitment to "best practice" and "what works". An emphasis of the TEMAG report was on the role of initial teacher education to address issues of teacher quality and achieve "best practice" in both initial teacher education and schooling, and its recommendations have seen associated shifts toward control of initial teacher education entrants, discussed further below. Meanwhile the Gonski 2.0 report emphasizes the need to establish what "works best" in teaching as a foundation for proposing a "formative assessment tool", currently under development, designed to guide "evidence-based decisions" and provide "suggestions for teaching strategies, professional learning and student learning activities based on evidence and best practice" (Gonski et al., 2018, p. 65).

Increased focus on reading, mathematics, and science

One of the "local vernacular" (Rizvi and Lingard, 2010) iterations of the Global Education Reform Movement in the Australian context was the inclusion of History (along with the more commonly regarded "core" subjects of English, Mathematics and Science) in the first iteration of the Australian Curriculum. While the development of an Australian national curriculum had been on the agenda for the Keating Labor Government in the early 1990s, the project was abandoned due to the inability of the federal Government to procure agreement from the eight states and territories. The idea of a national curriculum for Australia was next advanced in 2006 when conservative Prime Minister John Howard, fueled by his enduring interest in stoking the "history wars", reignited the discussion of a national curriculum for History (Taylor, 2010). Upon Howard's defeat in the 2007 federal election, the Rudd Labor Government quickly established the beginnings of what was to become the Australian Curriculum, as part of their *Education Revolution* platform (Commonwealth of Australia, 2008) starting with the four "core" subjects of English, Mathematics, Science and History for Kindergarten to Year 10.

Despite the relative breadth of the national curriculum, however (which has now been expanded to include eight learning areas including the arts, languages, technologies, and health and physical education), NAPLAN, introduced in 2009, has focused exclusively on literacy and numeracy. While NAPLAN, and the associated MySchool website, will be further discussed below in relation to test-based accountability, together with the Australian Curriculum they have worked to shape the practice landscape of teaching such that increased focus is placed on literacy and numeracy in many schools, to the detriment of other learning areas (Polesel et al., 2014).

Furthermore, this increased focus has seen a continuing discussion and at times toxic debate around the teaching of reading in Australian schools and particularly the essential or non-essential role of synthetic phonics. While the "phonics debate" in Australia dates to at least the early 2000s (Snyder, 2008), in recent years, due partly to the attention paid in reviews such as the TEMAG review, these debates have once again taken center stage, with phonics explicitly named as prescribed content in initial teacher education programs. As a consequence of the post-TEMAG reforms, the Australian Institute for Teaching and School Leadership (AITSL) is required to report annually to the Education Council on (among five other things) "implementation progress on accreditation requirements related to reading instruction, including phonics" (AITSL, 2015, p. 30).

The final piece of the policy ensemble demonstrating this increased focus on literacy and numeracy is also a consequence of the post-TEMAG reforms, and relates to the personal literacy and numeracy knowledge of pre-service and early career teachers. The Literacy and Numeracy Test for Initial Teacher Education (LANTITE) has been administered since 2016 in an attempt to improve "pre-service teacher selection" (Craven et al., 2015), a consequence of the TEMAG recommendation that such a test should be implemented to ensure that graduating teachers are "within the top 30% of the population in personal literacy and numeracy" (p. 17). While the LANTITE test does not relate directly to the focus on literacy and numeracy in the school curriculum, it has played an important part in denoting "what matters" for pre-service teachers and their education (Barnes and Cross, 2021) and, we argue, has contributed to the shaping of the practice landscape on which early career teachers enter the profession.

Corporate models of change

Associated with the GERM are models of new public management which devolve responsibility for things such as schooling outcomes while simultaneously requiring heightened control at the center of the system. In Australia, as elsewhere in the world, this has manifested via reforms described as aiming to increase school "autonomy". Such reforms—occurring in various forms across the country, most notably the Independent Public School scheme—have been argued to both support and detract from social justice outcomes, enabling greater flexibility at the local level but, in doing so, also entailing the potential to privilege some voices over others through cultural-discursive performative logics (Keddie, 2017). Such reforms have had considerable impacts on the nature and scale of principals' work and workload (McGrath-Champ et al., 2019). In addition, research suggests that there have been particular effects on teachers of such reforms, with reports of increasing amounts of paperwork since devolutionary measures have been put into place (Fitzgerald et al., 2019; Stacey et al., 2020). This has included paperwork in relation to the monitoring of student achievement, particularly in numeric terms; what is sometimes described as the "datafication" of schooling (Hardy and Lewis, 2017),

reflecting shifting social-political arrangements with material-economic effects for teachers. Such collection and documentation is also part of monitoring oneself as a teacher, as cultural-discursive evidence of "good teaching" (Mockler and Stacey, 2021).

Part of this perceived increase in "paperwork" demands comes from the institution of teaching standards (Australian Institute for Teaching and School Leadership, 2011). The standards sit as a social-political and cultural-discursive umbrella technology under which other policy processes operate, including accreditation and the Performance and Development Framework, a process by which teachers set goals and are assessed as to their achievement of them, aligned with the Standards (Australian Institute for Teaching and School Leadership, 2012). Thus these shifts also have clear material-economic effects for teachers because, as with the Performance and Development Framework, the Standards are now linked to teachers' remuneration; in the state of NSW, for instance, teacher salaries now progress not only in relation to time in the profession but also level of accreditation. For some teachers, the accreditation process has largely been seen as "hoop jumping" (McGrath-Champ et al., 2020); it is not so much the Standards themselves that have been at issue but rather the requirement to document the ways in which one's work meets such standards, constituting what has been described as "evidence for no one" (Talbot, 2016). In this way, teaching standards and the associated accreditation procedure can be understood as a manifestation of the shifting social-political relation of teachers' work with the state, that also arguably has cultural-discursive effects of enhancing teachers' performative subjectivities.

Test-based accountability

Finally, test-based accountability realized through standardized testing in Australia as elsewhere, has had a symbiotic relationship with both the standardization of teaching and learning and an increased focus on "core" subjects. Test-based accountability, as a form of "performative accountability" (Lyotard, 1984) resonates with the "audit explosion" (Power, 2003) in schools and school systems and, through the reshaping of the practice landscape of schooling, has given rise to a range of "perverse effects" for teachers, schools and school systems (Lingard and Sellar, 2013).

As noted above, Australia's national standardized testing program for school students was introduced in 2009, a key tenet in the Rudd-Gillard push toward the increased influence of the Commonwealth Government on schooling in Australia (Savage, 2016). Australian schools were at that point no stranger to standardized testing: state and territory-based standardized tests had been undertaken by students since the mid-1990s after the Commonwealth Education Minister managed to get agreement through the intergovernmental education council, using the lure of extra funding as a "sweetener" (Lingard, 1998). However, the new national testing regime, linked in 2010 to the Commonwealth Government's new national reporting website, myschool.edu.au, quickly became high-stakes (Lingard, 2010), through the publication of school results, along with the comparison of "statistically similar" schools enabled, and indeed encouraged by the website.

Research over the past decade has indicated that the impacts of the NAPLAN/MySchool ensemble are far-reaching. A narrowing of the curriculum (Thompson, 2013), increased "teaching to the test" (Dulfer et al., 2012), raised student and teacher anxiety (Howell, 2017) and an orientation to "gaming the system" (Thompson and Cook, 2014) have all been observed as examples of "perverse effects". Other research (e.g., Gannon, 2012; Mockler, 2013, 2016) has pointed to the systemic effects of NAPLAN in terms of reshaping perceptions of teachers and schooling in the public space.

Such has been the embrace of test-based accountability in Australia that the first Commonwealth Education Act, passed into legislation in 2013, included a commitment "for Australia to be placed, by 2025, in the top 5 highest performing countries based on the performance of school students in reading, mathematics and science" ("Australian Education Act" 2013, p. 3), as measured via PISA, TIMSS and PIRLS results. This clause was removed from the Preamble of the Act when it was modified in 2018, presumably on the basis of the increasingly gloomy prospect of its aim being met.

Despite its removal, however, Australia's commitment to test-based accountability shows no signs of abating: Australia's ever-more disappointing PISA performance is an entrenched part of the ongoing rationale for education reforms invariably involving increased accountability (see, e.g., Gonski et al., 2018, p. 2). Indeed, both the "formative assessment tool" embedded in the Gonski 2.0 reforms and the LANTITE test, one of the TEMAG reforms promoted as "encouraging" in *Through Growth to Achievement* (Gonski et al., 2018, p. 72) can be viewed as new iterations of test-based accountability impacting on the practice landscapes and practice traditions at work in Australian schooling.

The GERM and policy effects: reshaping the practice architectures of teachers' work

Research over the past ten years has consistently shown that the "policy effects" (Ball, 1993) of the GERM-related policy ensemble in Australia have been far-reaching, impacting upon the work of teachers, students and school and system leaders. It is important to note in this discussion, however, that the "field" or "practice landscape" (Kemmis et al., 2014) of school education in Australia is a very complex one, comprising hugely diverse schools across eight states and territories. Accordingly, we do not suggest that the policy effects of the GERM are experienced or impact practice consistently across the field. Indeed, even in schools that "on paper" might appear to be similar, a confluence of circumstances, including the particular articulation of "practice traditions" (Kemmis et al., 2014) within each context might mean that the effects play out in very different ways. Here, we explore policy effects in the form of shifts in the cultural-discursive, material-economic and social-political arrangements that frame teaching and schooling practices, and how research has suggested these have impacted some teachers and schools, without generalizing to the entire field.

To do this, we draw on, and draw together, key examples from the previous sections exploring the articulation of the GERM in Australian education policy and discuss these as particular cultural-discursive, material-economic and social-political arrangements, highlighting the significance of each set of arrangements for teachers' work as we go.

Cultural-discursive arrangements

The "language of improvement and reform" (Mockler and Groundwater-Smith, 2018) in education embedded strongly in policies related to the GERM, has infiltrated and come to dominate many of the cultural-discursive arrangements that frame teaching practice in contemporary Australia. The "consumerization" and marketization of education resonant with the growth in competition and "school choice" and the adoption of corporate models of change has seen a cultural-discursive shift from notions of collaboration and collectivism to competition—between students, teachers, schools, school systems—and a growing fixation on rankings, both national and international. At scale, there has been a linking of education processes with the economic advance of the nation, discursively shifting the purpose and goal of teachers' labor.

The NAPLAN/MySchool policy ensemble has seen notions of competition become embedded and intertwined with public understandings of what it is to be a successful school, demonstrated by outperforming other "statistically similar" schools. At the same time, the emergence of PISA as key arbiter of educational effectiveness on an international scale (Gorur and Wu, 2015), and the consequent ranking of national and regional schooling systems has also contributed to public understandings of education as a "global race" (Baroutsis and Lingard, 2016). Finally, the language of audit and new public management has over the past two decades permeated educational spaces, fueled by reforms related to corporate models of change, test-based accountability and the discourse of competition, reminiscent of Wendy Brown's observation that neoliberalism, "configures all aspects of existence in economic terms" (2015, p. 17), therein undermining and "undoing" social and democratic institutions.

The consequences of these cultural-discursive shifts for teachers are far from semantic, and neither have they been fully revealed as yet. While a substantial amount of research has centered on the direct impact of NAPLAN, for example, on curricular and pedagogical practice (Hardy, 2014; Polesel et al., 2014; Thompson, 2013), or on the impact of professional teaching standards on teachers' work and practice (Ryan and Bourke, 2013, 2018; Talbot, 2016), cultural-discursive shifts and their impacts are more complex and likely to be more observable over time than these more direct impacts. The language of new public management has contributed to the commodification of education (school and otherwise) in Australia. The dominant metaphor embedded in many policy documents and media reports on education of the "race", comprising "winners" and "losers", and the language of educational effectiveness which imbues schooling policy at every level are products of the neoliberal turn in education that contribute to these cultural-discursive arrangements.

Material-economic arrangements

The neoliberal turn has also served to re-shape material-economic arrangements in education. Logics in support of the individual right to "choose" have and continue to impact the allocation of funding for schools, across both public and private sectors. Although this is purportedly designed around "needs-based" principles, the current model reportedly continues to fail to meet this stated goal (see e.g., Ting et al., 2019), indicating the persistent tension between the funding of fee-charging schools and the achievement of system-wide equity. A core implication of these material arrangements has been the disproportionate sorting of student bodies into particular schools along socio-economic lines, arguably a result of the ongoing encouragement of school choice practices (Campbell et al., 2009).

Material shifts have also been evident in relation to what is taught and how, as well as who. There are new curriculum documents in play, and new, and very real, regular standardized testing experiences taking place. An emphasis on the standardization of teaching and learning has changed the arrangements for "who" can enter initial teacher education, altering the material future of the teacher education space, with the increased focus on "core" subjects also impacting what initial teacher education students do once there, for instance via the phonics debate. Teachers who graduate will also encounter a new pay scale, at least within the public sector, tied to accreditation processes. It is thus clear that the GERM has not only resulted in shifting arrangements for how teachers think and speak, but also what they "do"—and how they are remunerated.

According to the research literature, shifting arrangements governing what teachers "do" are reported to have a range of fairly direct impacts (notably including the pay they receive), as well as less direct implications. Working with qualitatively different student cohorts involves different kinds of work for teachers (Stacey, 2020). In addition, school choice has been argued to impact where teachers can and will work (Parding et al., 2017). Enhanced accountability practices have also been seen to lead to an increase in "paperwork" and overall workload for teachers (Stacey et al., 2020). Thus the effects of the material-economic arrangements for teaching work under the GERM have in many ways been immediate. Like the impact of cultural-discursive arrangements, however, if left unchecked, these material-economic shifts are likely to have further impacts in future.

Social-political arrangements

Social-political arrangements concern matters of power—the relationships between people and within systems. Drives toward competition and choice change the power dynamic in and around schools, positioning teachers as service providers, with parents

as "customers" and students as "products"; a "transaction". Schools are placed in competition with each other, as the emphasis is placed on working in the interest of oneself—or one's school—rather than in the interest of the collective. While schools and teachers may wish, and indeed even act to resist this relation, they are unavoidably implicated within new market systems by virtue of the arrangements governing funding and student enrollment.

In addition, the current settlement positions schools and teachers as instruments of the state working toward economic advancement, rather than focusing on working with and within local communities. In being so positioned, these local actors are also responsibilized. This does not however mean they necessarily have greater control or autonomy in their work, but that the expectations of their responsibility for its outcomes are increased via corporate models of change. Responsibility is shifted to local levels, and balanced by enhanced accountability at the center, seen for instance through the technology of standardized testing. Corporate models of change also involve a re-arrangement of power relations within schools, due to the enhanced role of principals in matters such as staffing.

There is then, in the era of the GERM, an evident reconfiguration of the role of the teacher, with a shifting relationship between teachers and other school staff, and between teachers and their work. Research has noted a range of impacts of these changes in social-political arrangements. One impact of positioning students and parents as products and customers has been new forms of labor related to marketing and market positioning, evident in some schools more than others but a relevant concern across the board (Stacey, 2020). Meanwhile, a lack of incentive to work collaboratively across schools has been argued to be fragmenting the "social" in the system (Holloway and Keddle, 2019), while teacher and executive relationships are put under strain in the context of devolutionary governance. In such an environment, teachers are also potentially tempted to work toward more simplified, countable goals rather than more holistic ideals of equity and student wellbeing (Wilkins et al., 2020). In addition, having to engage with standardized testing and "what works" approaches, and the workload impacts of having to "prove" oneself and one's work (Stacey et al., 2020), reflect the ongoing positioning of teachers as subservient to the state through these revised power relations. It is thus clear that GERM has impacted not only how teachers think and speak, and what they "do", but also how they are positioned within hierarchies of power, from the context of the school to that of the nation.

Conclusion

In this article, we have presented the case of Australia as a national example of the global educational reform movement, or GERM, currently in operation in various forms around the world. We have highlighted the significance of teachers to this project, as responsibility for schooling outcomes is shifted to local levels including that of the teacher, entailing substantial reform that directly targets what they do, as well as how this work is positioned, encoded and recognized. However, we have also emphasized that many of the impacts of the GERM upon teachers' work come from policies which do not directly target teachers. This includes, for instance, drives toward school choice, and the reshaping of the role of principal under devolved models of governance. Nevertheless, as we have argued, these changes have impacts for teachers, affecting such things as the composition of their classrooms and their relationships with school executive, respectively. Education reform, we argue, must therefore always be considered as part of a complex system with far-reaching and indirect, and unintended as well as intended, effects. Our focus on teachers in this article serves to highlight this core tenet in approaching the analysis of policy settlements, in education and beyond.

The analysis presented in this article may strike the reader as somewhat bleak. Teachers are seen as subject to cultural-discursive, material-economic and social-political arrangements which are largely constraining: heightening responsibility for narrowly defined educational outcomes, yet simultaneously curtailing autonomy; increasing the segregation of students, yet without providing appropriate resourcing, material or otherwise; and throughout, introducing new competitive and performative tensions at classroom, school and state levels. This is at least in part the result, we argue, of the GERM, which sees the individual and the economy driving educational imperatives rather than, perhaps, more communitarian ideals.

There is some evidence that broader conversations about education may be shifting at the global level, however. Zancajo (2019) identifies an "equity turn" in global education policy, and evident in current Chilean reforms to school choice policy via the "Inclusion Law" and requirements for (some) schools to enroll a greater "social mix". Others, however, are less optimistic. Research based on North American data, for instance, has suggested that within the current environment, teachers may be becoming wholly "marketized, managed, and performative" (Holloway and Brass, 2018, p. 12), part of an ongoing shift in the nature of teachers' work in recent years. Similarly, a recent Australian publication documents the rise of the "neo-performative" teacher, whose work is aimed at narrowly-defined equity purposes of "standards" and "outcomes" of academic attainment, rather than "a broader"—and perhaps more subversive—"social justice agenda" (Wilkins et al., 2020, p. 4). A growing global focus on equity, however defined, may nevertheless constitute a move in the right direction, and may offer some hope, especially if supported by key national and global institutional bodies. In addition, our own recent research suggests that Australian teachers are somewhat wise to the landscape they have seen shifting underneath their feet, expressing some distrust and resistance toward the datafication of their work and its associated performative imperatives (Mockler and Stacey, 2021; Stacey et al., 2020). This would suggest that the trajectory of education reform over the past forty to fifty years may yet change course. An "equity turn", combined with ongoing teacher discernment and criticality, may be key ingredients in enabling a shift in the Australian education policy settlement, toward an appropriately supported, empowered and respected teaching profession.

References

- ACARA, 2010. The Shape of the Australian Curriculum V2.0. ACARA, Sydney.
- ACARA, 2011. The Shape of the Australian Curriculum V3.0. ACARA, Sydney.
- ACARA, 2012. The Shape of the Australian Curriculum V4.0. ACARA, Sydney.
- Au, W., 2007. High-stakes testing and curricular control: a qualitative metasynthesis. *Educ. Res.* 36 (5), 258–267. <https://doi.org/10.3102/0013189X07306523>.
- Australian Government, 2016. Quality Schools, Quality Outcomes. Department of Education, Canberra.
- Australian Institute for Teaching and School Leadership, 2011. Australian Professional Standards for Teachers. Ministerial Council for Education, Early Childhood Development and Youth Affairs, Victoria.
- Australian Institute for Teaching and School Leadership, 2012. Australian Teacher Performance and Development Framework. https://www.aitsl.edu.au/docs/default-source/default-document-library/australian_teacher_performance_and_development_framework_august_2012.pdf?sfvrsn=e7c2ec3c_0.
- Australian Institute for Teaching and School Leadership, 2015. Accreditation of Initial Teacher Education Programs in Australia: Standards and Procedures. Education Services Australia, Canberra.
- AustralianPolitics.com, 2008. An Education Revolution in Our Schools: Rudd and Gillard Announce School Reform Agenda. australianpolitics.com.
- Ball, S., 1993. What is policy? Texts, trajectories and toolboxes. *Discourse* 13 (2), 10–17.
- Barnes, M., Cross, R., 2021. "Quality" at a cost: the politics of teacher education policy in Australia. *Crit. Stud. Educ.* 62 (4), 455–470. <https://doi.org/10.1080/17508487.2018.1558410>.
- Baroutsis, A., Lingard, B., 2016. Counting and comparing school performance: an analysis of media coverage of PISA in Australia, 2000–2014. *J. Educ. Pol.* 32 (4), 432–449.
- Barton, G., Garvis, S., Ryan, M., 2014. Curriculum to the classroom: investigating the spatial practices of curriculum implementation in Queensland schools and its implications for teacher education. *Aust. J. Teach. Educ.* 39 (3), 166–177.
- Berliner, D., 2013. Problems with value-added evaluations of teachers? Let me count the ways. *Teach. Educ.* 48 (4), 235–243. <https://doi.org/10.1080/08878730.2013.827496>.
- Bonnor, C., Shepherd, B., 2016. Uneven Playing Field: The State of Australia's Schools accessed from the Centre for Policy Development website. <http://cpd.org.au/wp-content/uploads/2016/05/The-State-of-Australias-Schools.pdf>.
- Brass, J., Holloway, J., 2021. Re-professionalizing teaching: the new professionalism in the United States. *Crit. Stud. Educ.* 1–18.
- Brown, W., 2015. Undoing the Demos: Neoliberalism's Stealth Revolution. Zone Books, New York.
- Buchanan, R., Chapman, A., 2011. Utopia or Dystopia? A Critical Examination of the Melbourne Declaration. Philosophy of Education Society of Australasia, New Zealand.
- Campbell, C., Proctor, H., 2014. A History of Australian Schooling. Allen & Unwin, Crows Nest, Sydney.
- Campbell, C., Sherington, G., 2013. The Comprehensive Public High School: Historical Perspectives. Palgrave Macmillan, New York.
- Campbell, C., Proctor, H., Sherington, G., 2009. School Choice: How Parents Negotiate the New School Market in Australia. Allen & Unwin, Crows Nest, Sydney.
- Carter, D., 2019. Restoring purpose: applying Biesta's three functions to the Melbourne Declaration. *Curric. Perspect.* 39, 125–134. <https://doi.org/10.1007/s41297-019-00086-0>.
- Castells, M., 2010. The Rise of the Network Society, second ed. Blackwell, Oxford.
- Chubb, J.E., Moe, T.M., 1990. Politics, Markets and America's Schools. The Brookings Institute, Washington, DC.
- Cochran-Smith, M., Piazza, P., Power, C., 2013. The politics of accountability: assessing teacher education in the United States. *Educ. Forum* 77 (1), 6–27. <https://doi.org/10.1080/00131725.2013.739015>.
- Comber, B., Kamler, B., 2009. Sustaining the next generation of teacher-researchers to work for social justice. In: Noffke, S., Somekh, B. (Eds.), *The SAGE Handbook of Educational Action Research*. SAGE, Thousand Oaks, pp. 226–237.
- Commonwealth of Australia, 2008. Quality Education: The Case for an Education Revolution in Our Schools. Commonwealth of Australia, Canberra.
- Conway, P.F., 2013. Cultural flashpoint: the politics of teacher education reform in Ireland. *Educ. Forum* 77 (1), 51–72. <https://doi.org/10.1080/00131725.2013.739021>.
- Craven, G., Beswick, K., Fleming, J., Fletcher, T., Green, M., Jensen, B., et al., 2015. Action Now: Classroom Ready Teachers. https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_accessible.pdf.
- Darling-Hammond, L., 2012. The right start: creating a strong foundation for the teaching career. *Phi Delta Kappan* 94 (3), 8–13. <https://doi.org/10.1177/003172171209400303>.
- Department for Education UK, 2016. DfE Strategy 2015–2020: World-Class Education and Care. DfE, London.
- Dulfer, N., Polesel, J., Rice, S., 2012. The Experience of Education: The Impacts of High Stakes Testing on School Students and Their Families—An Educator's Perspective. <https://www.whitlam.org/publications/2017/10/17/an-educators-perspective-the-impacts-of-high-stakes-testing-on-school-students-and-their-families>.
- Education Council, 2019. Alice Springs (Mparntwe) Education Declaration. <https://docs.education.gov.au/documents/alice-springs-mparntwe-education-declaration>.
- Ellis, V., Steadman, S., Trippstad, T.A., 2019. Teacher education and the GERM: policy entrepreneurship, disruptive innovation and the rhetorics of reform. *Educ. Rev.* 71 (1), 101–121. <https://doi.org/10.1080/00131911.2019.1522040>.
- Fitzgerald, S., McGrath-Champ, S., Stacey, M., Wilson, R., Gavin, M., 2019. Intensification of teachers' work under devolution: a "Tsunami" of paperwork. *J. Ind. Relat.* 61 (5), 613–636. <https://doi.org/10.1177/0022185618801396>.
- Forsey, M., 2010. Teachers and the re-production of middle-class culture in Australian schools. *Int. Stud. Sociol. Educ.* 20 (1), 67–81. <https://doi.org/10.1080/09620211003655671>.
- Foucault, M., 1980. Power/Knowledge. Pantheon, New York.
- Fu, G., Clarke, A., 2019. Teachers' moral agency under neo-liberal influences: what is educationally desirable in China's curriculum reform? *Educ. Rev.* 71 (1), 51–66. <https://doi.org/10.1080/00131911.2019.1524205>.
- Fuller, K., Stevenson, H., 2019. Global education reform: understanding the movement. *Educ. Rev.* 71 (1), 1–4. <https://doi.org/10.1080/00131911.2019.1532718>.
- Gannon, S., 2012. My school redux: re-storying schooling with the My School website. *Discourse* 34 (1), 17–30. <https://doi.org/10.1080/01596306.2012.698861>.
- Gonski, D., Boston, K., Greiner, K., Lawrence, C., Scales, B., Tannock, P., 2011. Review of Funding for Schooling: Final Report. <https://docs.education.gov.au/system/files/doc/other/review-of-funding-for-schooling-final-report-dec-2011.pdf>.
- Gonski, D., Arcus, T., Boston, K., Gould, V., Johnson, W., O'Brien, L., et al., 2018. Through Growth to Achievement: Report of the Review to Achieve Educational Excellence in Australian Schools. https://docs.education.gov.au/system/files/doc/other/662684_tgta_accessible_final_0.pdf.
- Gorur, R., Wu, M., 2015. Learning too far? PISA, policy and Australia's "top five" ambitions. *Discourse* 35 (5), 647–664. <https://doi.org/10.1080/01596306.2014.930020>.
- Hardy, I., Lewis, S., 2017. The "doublethink" of data: educational performativity and the field of schooling practices. *Br. J. Sociol. Educ.* 38 (5), 671–685. <https://doi.org/10.1080/01425692.2016.1150155>.
- Hardy, I., 2014. A logic of appropriation: enacting national testing (NAPLAN) in Australia. *J. Educ. Pol.* 29 (1), 1–18.
- Hattie, J., 2008. Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement. Routledge, Oxon.
- Hogarth, M., 2020. Is This What Dan Tehan Means by "Back to Basics"? The Mparntwe Declaration. <https://www.aare.edu.au/blog/?tag=mellita-hogarth>.
- Holloway, J., Brass, J., 2018. Making accountable teachers: the terrors and pleasures of performativity. *J. Educ. Pol.* 33 (3), 361–382. <https://doi.org/10.1080/02680939.2017.1372636>.
- Holloway, J., Keddie, A., 2019. Competing locals in an autonomous schooling system: the fracturing of the "social" in social justice. *Educ. Manag. Adm. Leader* 1–16. <https://doi.org/10.1177/1741143219836681>.

- Holloway, J., Sørensen, T.B., Verger, A., 2017. Global perspectives on high-stakes teacher accountability policies: an introduction. *Educ. Policy Anal. Arch.* 25 (85) <https://doi.org/10.14507/epaa.25.3325>.
- Howell, A., 2017. "Because then you could never ever get a job!": children's constructions of NAPLAN as high-stakes. *J. Educ. Pol.* 32 (5), 564–587. <https://doi.org/10.1080/02680939.2017.1305451>.
- Keddie, A., 2017. School autonomy reform and public education in Australia: implications for social justice. *Aust. Educ. Res.* 44, 373–390. <https://doi.org/10.1007/s13384-017-0243-x>.
- Kemmis, S., Wilkinson, J., Edwards-Groves, C., Hardy, I., Grootenboer, P., Bristol, L., 2014. *Changing Practices, Changing Education*. Springer, Dordrecht.
- Larsen, M.A., 2010. Troubling the discourse of teacher centrality: a comparative perspective. *J. Educ. Pol.* 25 (2), 207–231. <https://doi.org/10.1080/02680930903428622>.
- Lingard, B., Sellar, S., 2013. "Catalyst data": perverse systemic effects of audit and accountability in Australian schooling. *J. Educ. Pol.* 28 (5), 634–656. <https://doi.org/10.1080/02680939.2012.758815>.
- Lingard, B., 1998. The disadvantaged schools programme: caught between literacy and local management of schools. *Int. J. Incl. Educ.* 2 (1), 1–14.
- Lingard, B., 2010. Policy borrowing, policy learning: testing times in Australian schooling. *Crit. Stud. Educ.* 51 (2), 129–147. <https://doi.org/10.1080/17508481003731026>.
- Lyotard, J.F., 1984. *The Postmodern Condition: A Report on Knowledge*. University of Minnesota Press, Minneapolis.
- McGrath-Champ, S., Stacey, M., Wilson, R., Fitzgerald, S., Rainnie, A., Parding, K., 2019. Principals' support for teachers' working conditions under devolution: insights from two Australian states. *Educ. Manag. Adm. Leader* 47 (4), 590–605.
- McGrath-Champ, S., Gavin, M., Stacey, M., 2020. Strategy and policy: the case of an Australian teachers' union. In: Lansbury, R., Johnson, A., Van den Broek, D. (Eds.), *Contemporary Issues in Work and Organizations: An Integrated Approach*. Routledge, Abingdon, pp. 110–126.
- Ministerial Council on Education Employment Training and Youth Affairs, 2008. Melbourne Declaration on Educational Goals for Young Australians. http://www.curriculum.edu.au/verve/_resources/national_declaration_on_the_educational_goals_for_young_australians.pdf.
- Mockler, N., Groundwater-Smith, S., 2018. Questioning the Language of Improvement and Reform in Education: Reclaiming Meaning. Routledge, London.
- Mockler, N., Stacey, M., 2021. Evidence of teaching practice in an age of accountability: when what can be counted isn't all that counts. *Oxf. Rev. Educ.* 47 (2), 170–188. [10.1080/03054985.2020.1822794](https://doi.org/10.1080/03054985.2020.1822794).
- Mockler, N., 2013. Reporting the "education revolution": MySchool.edu.au in the print media. *Discourse* 34 (1), 1–16.
- Mockler, N., 2016. NAPLAN and the problem frame: exploring representations of NAPLAN in the print media, 2010 and 2013. In: Lingard, B., Sellar, S., Thompson, G. (Eds.), *National Testing in Schools: An Australian Assessment*. Routledge, London, pp. 181–198.
- Mockler, N., 2022. Teacher professional learning under audit: reconfiguring practice in an age of standards. *Prof. Dev. Educ.* 48 (1), 166–180. <https://doi.org/10.1080/19415257.2020.1720779>.
- Morsy, L., Gulson, K., Clarke, M., 2014. Democracy, "sector-blindness" and the delegitimization of dissent in neoliberal education policy: a response to Discourse 34(2), May 2013. *Discourse* 35 (3), 444–461. <https://doi.org/10.1080/01596306.2014.890267>.
- National Curriculum Board, 2009. *The Shape of the Australian Curriculum V1.0*. ACARA, Sydney.
- OECD, 2019. *TALIS 2018 Results (Volume 1): Teachers and School Leaders as Lifelong Learners*. OECD Publishing, Paris, France.
- Parding, K., McGrath-Champ, S., Stacey, M., 2017. Teachers, school choice and competition: lock-in effects within and between sectors. *Pol. Futures Educ.* 15 (1), 113–128. <https://doi.org/10.1177/1478210316688355>.
- Parding, K., McGrath-Champ, S., Stacey, M., 2020. Governance reform in context: welfare sector professionals' working and employment conditions. *Curr. Sociol.* <https://doi.org/10.1177/0011392120909859>.
- Polesel, J., Rice, S., Duffer, N., 2014. The impact of high-stakes testing on curriculum and pedagogy: a teacher perspective from Australia. *J. Educ. Pol.* 29 (5), 640–657.
- Power, M., 2003. Evaluating the audit explosion. *Law Pol.* 25 (3), 185–202.
- Proctor, H., Aitchison, C., 2015. Markets in education: "School choice" and family capital. In: Meagher, G., Goodwin, S. (Eds.), *Markets, Rights and Power in Australian Social Policy*. Sydney University Press, Sydney, pp. 323–341.
- Reid, A., 2018. The journey towards the first Australian curriculum. In: Reid, A., Price, D. (Eds.), *The Australian Curriculum: Promises, Problems and Possibilities*. ACSA, Canberra, pp. 1–16.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, Abingdon.
- Ryan, M., Bourke, T., 2013. The teacher as reflexive professional: making visible the excluded discourse in teacher standards. *Discourse* 34 (3), 411–423.
- Ryan, M., Bourke, T., 2018. Spatialized metaphors of practice: how teacher educators engage with professional standards for teachers. *Crit. Stud. Educ.* 59 (2), 167–186.
- Sahlberg, P., 2016. The global education reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. John Wiley & Sons, New Jersey, pp. 128–144.
- Sato, M., 2014. What is the underlying conception of teaching of the edTPA? *J. Teach. Educ.* 65 (5), 421–434.
- Savage, G.C., 2011. When worlds collide: excellent and equitable learning communities? Australia's "social capitalist" paradox? *J. Educ. Pol.* 26 (1), 33–59. <https://doi.org/10.1080/02680939.2010.493229>.
- Savage, G.C., 2016. Who's steering the ship? National curriculum reform and the re-shaping of Australian federalism. *J. Educ. Pol.* 31 (6), 833–850.
- Smaller, H., 2015. The teacher disempowerment debate: historical reflections on "slender autonomy". *Paedagog. Hist.* 51 (2), 136–151. <https://doi.org/10.1080/00309230.2014.997752>.
- Snyder, I., 2008. *The Literacy Wars*. Allen & Unwin, Crows Nest, Sydney.
- Stacey, M., Wilson, R., McGrath-Champ, S., 2020. Triage in teaching: the nature and impact of workload in schools. *Asia Pac. J. Educ.* <https://doi.org/10.1080/02188791.2020.1777938>.
- Stacey, M., 2020. *The Business of Teaching: Becoming a Teacher in a Market of Schools*. Palgrave Macmillan, Cham, Switzerland.
- Steiner-Khamisi, G., 2004. *The Global Politics of Educational Borrowing and Lending*. Teachers College Press, New York.
- Storey, A., 2006. The search for teacher standards: a nationwide experiment in The Netherlands. *J. Educ. Pol.* 21 (2), 215–234.
- Talbot, D., 2016. Evidence for no-one: standards, accreditation, and transformed teaching work. *Teach. Teach. Educ.* 58, 80–89. <https://doi.org/10.1016/j.tate.2016.05.006>.
- Taylor, T., 2010. Constructing the Australian school history curriculum: ideology, high politics and the history wars in the howard years. In: *Globalization, Ideology and Education Policy Reforms*. Springer, Netherlands, pp. 19–37.
- Thompson, G., Cook, I., 2014. Manipulating the data: teaching and NAPLAN in the control society. *Discourse* 35 (1), 129–142. <https://doi.org/10.1080/01596306.2012.739472>.
- Thompson, G., 2013. NAPLAN, MySchool and Accountability: teacher perceptions of the effects of testing. *Int. Educ. J.* 12 (2), 62–84.
- Ting, I., Palmer, A., Scott, N., 2019. Rich School, Poor School: Australia's Great Education Divide. <https://www.abc.net.au/news/2019-08-13/rich-school-poor-school-australias-great-education-divide/11383384>.
- Verger, A., Parcerisa, L., Fontdevila, C., 2019. The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms. *Educ. Rev.* 71 (1), 5–30. <https://doi.org/10.1080/00131911.2019.152204>.
- Wall, L., 2017. Institutional logics and curriculum decision making: enacting the Australian Curriculum English and NAPLAN literacy. *Aust. Educ. Res.* 44 (4–5), 391–407.
- Weaven, M., Clark, T., 2015. Discretionary space: English teachers discuss curriculum agency. *Aust. J. Lang. Literacy* 38 (3), 162–170.
- Wilkins, C., Gobby, B., Keddie, A., 2020. The neo-performative teacher: school reform, entrepreneurialism and the pursuit of educational equity. *Br. J. Educ. Stud.* <https://doi.org/10.1080/00071005.2020.1739621>.
- Windle, J., 2015. *Making Sense of School Choice: Politics, Policies, and Practice under Conditions of Cultural Diversity*. Palgrave Macmillan, New York.
- Zancajo, A., 2019. Drivers and Hurdles to the Regulation of Education Markets: The Political Economy of Chilean Reform. <http://ncspe.tc.columbia.edu/working-papers/WP239.pdf>.

Defiant imaginaries: from a politics of identity to an ethics of friendship in teaching

Anne M. Phelan^a, Matthew Clarke^b, and Lily Vincent^c, ^a University of British Columbia, Vancouver, British Columbia, Canada; ^b University of Aberdeen, Scotland; and ^c Teach First, London, England

© 2023 Elsevier Ltd. All rights reserved.

Introduction	39
Engaging a politics of identity: toxic teaching in England's academy schools	40
Teachers' relationships to curriculum and pedagogy	41
Teachers' relationships with leadership and colleagues	42
Teachers' understanding of and relationship to themselves	43
Foucault's challenge: friendship as a form of life	43
Reconstituting a relation to curriculum, pedagogy and assessment via betrayal	44
Reconstituting non-hierarchical relationships with colleagues via distancing	45
Reconstituting a relation with self via frankness	46
Conclusion: friendship as a form of negative ethics in teaching	47
References	48

Not only do we have to defend ourselves, not only affirm ourselves, as an identity, but as a creative force.

(Foucault, 1997a, p. 164).

Introduction

In this article we consider the current state of teacher professionalism in the neoliberalized context of 21st century education, characterized by intensified cultures of competitive performativity. As part of our discussion, we draw on interview data from early career teachers working in academies in the north of England, setting these alongside Foucault's thinking on friendship as a way of life, in order to explore how teachers' relationships to their professional work, to their colleagues and to themselves might be conceived and enacted otherwise.

Schools today, operating within the mantra of global economic competitiveness in many national contexts, function within a neoliberal policy environment that has been significantly re-shaped by a range of factors. These factors include the deregulation and privatization of school governance, increasing curricular prescription, the power of high-stakes assessment, growing popularity of authoritarian approaches to discipline, the intensification of competitive pressures within and between schools, and the rise of alternative routes into teaching. A key instrument of these reforms has been the "new professionalism" (Phillips and Furlong, 2002), characterized by governments' attempts to address alleged problems of teacher quality, attrition and retention; to differentiate the profession in terms of tasks and rewards through chartered teacher schemes; to assert a coherent, evidence-based knowledge for teachers; and to remedy teachers' perceived lack of accountability (Furlong, 2008). The proclaimed goal of the "new professionalism" has been to offset the otherwise "mystificatory and self-serving ideologies" (Gilbert, 2014, p. 44)—an autonomous teaching subject with responsibility for the public good and whose competence and conduct could be harnessed to political ends (Fournier, 1999)—reminiscent of the earlier liberal form.

Whatever the ideology governing bodies such as the state can only manage identifiable subjects. Educational policy relies on and operates through a politics of teacher identity (Clarke, 2009; Day et al., 2005; Mockler, 2011). Within the regime of "new professionalism," "policy discourse constitutes the practice of teaching as the ground of student achievement, an objective ground that justifies the normative arranging and ordering of teachers" (Carusi, 2017, p. 638). As such, the teacher becomes indistinguishable from the law (i.e., teaching standards; evidence-based practice; effectiveness) and its desired results—the extraction of the educational life of the young as human capital in the interests of global capitalism—the teacher's life and the law become one; the teacher is her effects (Lewis, 2015). Policy—formation, implementation, enactment, and dispersion—foment the idea that effective teachers are "what leads to student achievement and what student achievement reveals" (Carusi, 2017, pp. 638–639). Education is reduced to a project of calculation—a strategy for realizing predetermined ends (i.e., high test scores)—with little intrinsic value (Biesta, 2010; Pinar, 2006; Taubman, 2009). Students and parents are seduced with cruelly optimistic promises of success, "a cure of sorts to a politics of fear and exclusion" (Roach, 2012, p. 166). Similarly deceived by the promise of and desire for certainty, and often fearing for their own job prospects, teachers side-line their moral judgment about what matters educationally. The ethico-political dimension of teaching is disavowed (Clarke and Phelan, 2017).

It is not unreasonable that teachers might adopt a "stable and positive [professional] identity obtained through identification with an existing socio-political order" (Biesta, 2011, p. 145; see also Day et al., 2006). However, recognition is always ambivalent, bearing both the promise of one's independence while highlighting one's reliance on the other for acknowledging

that so-called independent state (Butler, 1997). Asserting that everything—from high achievement on national and international tests to economic happiness—depends on the teacher paves the way for the standardization of teaching and teacher education and development (Larsen, 2010). The intellectual—the teacher and not the student is the site of education—and political—the teacher is scapegoated for the failures of the system—traps for teachers are lethal (Pinar, 2006). In the USA, teachers deemed ineffective are quickly taken out of the classroom, livelihoods threatened, teachers' credibility obliterated, and professional judgment erased (Lewis, 2015). In the UK "experienced teachers who have vocationally, ethically and professionally based identities" (Courtney and Gunter, 2015, p. 414) and who resist the reduction of their work to "effects", are "purged" from the workforce in favor of younger practitioners who see teaching as "just a job" and are more inclined to accept the "dominant discourses of audit, performativity and standards" (p. 414). We are witnessing a model of power that constructs the teaching subject through rejection, disqualification and exclusion by imposing a monoculture of thought and practice (Clarke and Phelan, 2017). The upshot is an educational policy context that privileges teacher compliance and conformity, while restricting their creativity and criticality. As such "teaching as a form of life-practice" (Smith, 2000, p. 16) wherein teachers are charged with the pursuit and sharing of truth with colleagues and students—what Britzman (2009) refers to as "the very thought of education"—is endangered.

Deeply concerned about the current politics of teacher identity, the central organizing questions for this chapter are: *What impact does a politics of identity—effective teacher—have on the professional lives of teachers? What constitutes an ethico-political response to the normative order of effective teaching, and with what consequences?*

We begin by illustrating the exclusionary, inequitable and "vicious hierarchies of identitarian difference" (Roach, 2012, p. 131) undergirding neo-liberal policy and playing havoc with teachers' professional lives. The site of exploration is the English academy schools where, one could argue, we find the most extreme form of neo-liberal schooling and a normalizing politics of identity—effective teacher—at full throttle. We then consider Foucault's challenge to disengage from a politics of identity—to dis-identify—by adopting an ethics of friendship. As a form of relation that transcends the confines of codified (professional) obligations and duties, friendship can play a significant role in re-figuring institutional norms and generating new practices of living together in schools (McLaren, 2007). We explore how three curious aspects of Foucauldian friendship—betrayal, distancing, and frankness—might generate non-normalizing relations that surpass institutional utility and pre-determined teacher identity. We conclude by positing a negative ethics of friendship for teaching in our times.

Engaging a politics of identity: toxic teaching in England's academy schools

We should fight against the impoverishment of the relational fabric.

(Foucault, *Essential*, VI: Ethics 158).

England has been in the vanguard of education developments based on a politics of (teacher) identity, borrowing and adapting policies and practices from other global contributors to the ever-evolving neoliberal playbook. Its national curriculum is highly prescriptive, while the combination of high stakes assessments, league tables and a punitive inspection regime triangulate to ensure a maximum degree of compliance on the part of schools, educators, parents and pupils. Since the 2010 Academies act in the UK, 77% of Secondary Schools have become academies (Department for Education, 2020), schools that are state funded but outside local government control, instead having direct responsibility to the Secretary of State for Education. Drawing inspiration from international models including Swedish Free Schools and the US Charter School movement, and inflected with a powerful dose of neo-colonial nostalgia (Clarke and Mills, 2021; Shahjahan, 2011), this program of academization has served to install a highly "unpublic", quasi-privatized form of public education in England (Thomson, 2019).

Our discussion below draws on interviews with a number of early career teachers with experience of working in academies in the North of England. Telephone interviews of 1 hour duration sought, through narrative inquiry, to establish teachers' perspectives on matters of policy, professionalism, and professional relationships. The participants taught at a number of different secondary schools serving students between eleven to sixteen years of age, from socioeconomically disadvantaged communities; the majority of these schools had been forced into academy status following poor inspection outcomes. Reflecting a pattern found in England but also in other international contexts, such as the USA, many schools serving such communities frequently employ an educational model that is grounded in demands for strict compliance with detailed curricular and pedagogical requirements and with highly specified and prescriptive disciplinary codes.

Such attention to the micro-details of institutional life has significant implications, not only for the students attending such schools but, for the teachers who work in them and who are required to promote, monitor and enforce the codes as part of their professional practice. Within such cultures, what counts as legitimate knowledge and understanding, what is required in relation to the micro-details of appearance, behavior and comportment, as well as what is expected to comprise the content of each person's thoughts and beliefs, are all specified, shared and sacrosanct. In other words, such cultures have implications at multiple levels including knowledge, interpersonal relations and individual subjectivity.

In thinking about these implications for teachers' professional lives, Foucault's notion of the possibility of conducting a historical ontology of the self is particularly valuable, since it allows us to bring together the pedagogical, political and ethical aspects of teaching as a socially situated and culturally constructed practice. Specifically, we base the following discussion on Foucault's (1983, p. 237) argument that we can conduct such an ontological inquiry on three domains:

First, an historical ontology of ourselves in relation to truth through which we constitute ourselves as subjects of knowledge; second an historical ontology of ourselves in relation to a field of power through which we constitute ourselves as subjects acting on others; third an historical ontology in relation to ethics through which we constitute ourselves as moral agents.

(Foucault, 1983, p. 237).

These three domains of historical ontology, i.e., knowledge and truth; power relations; and ethics as a practice of self-formation, provide the structure within which we present the analysis of our data from interviews with early career teachers. Specifically, we consider the domain of knowledge and truth in terms of teachers' relationships to curriculum, pedagogy and assessment; we examine the domain of power relations in terms of teachers' relationships with school leadership and colleagues; and we consider the domain of ethics in terms of teachers' understanding of and relationship to themselves. For analytical purposes, we begin with the first domain, while mindful that in practice the three domains are often inextricably entwined. For each domain, we illustrate the assimilationist, exclusionary, and hierarchical qualities of identitarian politics and policies in play—teachers' experience of subjection, and their struggles for subjectivation, as they encounter the politically administered identity of "effective teacher."

Teachers' relationships to curriculum and pedagogy

Common to the experience of many of the teachers we interviewed in our study were strict requirements to teach particular content in very specific ways. As Emma put it, "there were a lot of non-negotiables that you had to make sure were implemented in your lessons ... it was very much like, you had to do these things. Some of them were behavioral, some of them teaching". As Emma notes here, concerns about curriculum and pedagogy in the schools our interviewees worked in typically went hand in hand with highly specified behavioral codes. Such requirements reflect a belief that consistency is the key to both effective teaching and school improvement.

As Miranda explained:

It's about consistency. And that comes out of a real lack of consistency in the school's history. They were a failing school ... an inadequate school, for many years and nobody wanted to go there. And now, they've had this huge success, become a good school, and with that it's like, we've got to be consistent, we've got to be consistent. So, I think that's part of the reason it's so prescribed.

Just as staff were expected to ensure consistency in their supervision of children's learning and behavior, so too the teachers themselves were continuously monitored and surveilled to make sure they were teaching in the prescribed fashion. Yet the purposes of particular observations were often not shared with those being observed: They did a lot of unannounced drop ins constantly but they never told you what they were looking for or what you were being graded against or marked against so you always came away with some "not met" somewhere. Sarah's experiences were similar to those described by Miranda who gave a chilling account of never knowing where the monitorial gaze would fall and the consequences this could have if her pedagogical practice was found to fall short of expectations: "just picking up a book ... and if they're out of cycle – which is about a three-week cycle – that's not good enough". Miranda's reference to a three-week cycle highlights the quantifiable, and hence controllable, dimension of marking which helps explain why it was such a focus for pedagogical guidelines and requirements; yet, as the teachers' comments also imply, it is an overly simplistic and reductive proxy for good teaching and learning.

For a number of teachers in the study, such practices served to undermine their professional aspirations: "So, they have someone really controlling the curriculum in that sense, and then I just make a PowerPoint." Jackie described a similar "tick box" response to the pedagogical and professional culture in her school: "she'll want to see very much in the way of direct instruction. If I know someone's coming in, or if they pop in randomly, I will just try and figure out what they want, and make sure I tick that box." Perhaps unsurprisingly, this highly prescriptive, "teacher proof" approach to teaching, by underestimating the ability of the teachers, often undermined its own aims as Henry admits: "I know I've got to do it, so I just do it ... it does make me detest the task ... which then makes me cut corners, which makes it less effective, because I might not be marking it as precisely, or as accurately, as I would do otherwise." The effect of the school policy environment in terms of undermining teacher professionalism was evident in other comments, for example, in Miranda's comments in response to an invitation to share her thoughts on potential alternative feedback and assessment practices: "You know, I haven't even thought. I think I'd like to just trial some other ways of giving feedback ... I haven't really looked into much research around marking feedback". This de-professionalization was characterized by Sarah as a loss of professional voice: "What I would call draconian measures

about marking and lesson planning. And then we had no voice about that. So then, I think our relationship with senior leaders became a bit more strange because we didn't agree with what they were doing but it was like this is what we are doing so you have to live with it or you leave".

Teachers' relation to the work of curriculum, pedagogy and assessment is that of compliance, of blind loyalty to externally imposed mandates. When school improvement and student success are conflated with consistency of curriculum content and processes across teachers at the same grade level, to think and act otherwise constitutes a betrayal of one's students, colleagues and school leadership.

Teachers' relationships with leadership and colleagues

As others have noted (e.g., Courtney and Gunter, 2015), many English schools are pursuing a reform agenda that clearly conveys a zero-sum, take it or leave it message to staff (and students) in relation to various aspects of school policy. This was captured in Emma's reflections on the leadership in her school: "We had a caretaker head teacher or two, and then there was all this consultation, which was really just consultation by name. The trust were coming in and telling people how it was going to be. There were quite a few of those meetings, and it was interesting, because I think the phrase was "you either get on the bus, or you get off". Such slogans serve to close down discussion, providing no space for teachers to contribute to policy or to engage in professional debate and deliberation. Instead, they position teachers as mere minions implementing and adhering to the letter of policies that are decided "higher up". Indeed, Sarah used the word "minion" to describe the position of teachers in typical English school management systems as she explained the frosty reception she received when trying to contribute to policy discussions: "I offer opinions and a different opposing view and see things and I don't think they like that. I think they are not expecting that because I am just a teacher, I am not a leader in terms of management. I suppose as a teacher you are, you're just a minion really to senior leaders sometimes."

The phrase, "just a teacher", is telling, capturing the diminishment of teachers' professional agency and identities. It also suggests an expendability. This is a point noted by Courtney and Gunter (2015) and also highlighted by many of the teachers we interviewed. As Miranda observed, "If he doesn't like you, then you won't have a job. And it'll be done well enough above board and everything, but people come and go very quickly if their face doesn't fit." Olivia reported remarks made by members of her school leadership team that "no one is irreplaceable". Similarly, Sarah explained how a culture of fear and anxiety ensured compliance: "There was the fear that if you didn't meet those deadlines and targets, what would happen? Would you get sacked? Would you be put on competency?" Ensuring that non-compliant teachers could be removed and replaced quickly, easily and efficiently required that professional unions had no role in schools. As Evan put it, "You know, union's a dirty word. If you mention the union, it's just not a thing ... it's just not there".

The undermining of solidarity represented by the proscribing of unions had effects on the teachers' professional relationships with each other. Sarah described how her peers reacted to a sense that she was not one of the "in crowd" as far as management were concerned: "I think they saw me as kind of like a weak element and therefore they didn't want to be associated with a weak element so they kept away from me. They saw like I wasn't getting on with the line managers ... so they distanced themselves and kind of allied themselves more with the popular teachers". Such competitive and threatening school environments also undermined collaboration, as Jackie noted: "It's a culture of competition, and alienation. It's sink or swim, people will not give you their stuff, people will not share". Alienating teachers may not promote collaboration - or stimulate creativity and criticality - but it is a powerful tool of control.

The control of practice and the diminishment of teachers' professional agency went all the way down to relatively trivial matters like the details of dress. Exemplifying Foucault's insights regarding how power is exercised on the body, and through the micro dimensions of practice, Miranda noted the attention to detail:

Yeah, they're just very detailed again, much like the kids' uniform policy. Staff uniform is you've got to either wear a, for females anyway, a skirt obviously, on the knee. No elbows can be on show at any time. So, all your tops have to be longer than elbows, or you wear your jacket all day, basically ... They say jackets can be a pinstripe, or any professional check, but I couldn't wear a floral jacket, for example, or I couldn't wear a jazzy jacket.

Unsurprisingly, the effect of these controlling practices went beyond diminishing professional agency and contributed to a sense of alienation, where, as Miranda observes, teachers no longer recognized, and felt estranged from, themselves: "Really you know that something's not quite right, but you can't. It's so strange. Nobody says no, even though you think, in real life, you would say no, and people should say no. But nobody says no". The sense of alienation was reflected in the comparisons some of the teachers grasped for to capture their experience, as in the following remarks from Emma: "It felt like Big Brother. There are a lot of in-jokes among some of us, about it being like the Soviet Union". Miranda described the effect of the constant monitoring in a similar way: "There's that fear of big brother's watching you, for want of a better kind of phrase".

The normalizing gaze of colleagues and school leaders fostered a community of restrictive, conditional belonging (Roach, 2012) wherein teachers' sense of self was rendered fragile, fearful and manipulable. Autonomous judgment, which requires both close engagement with and distance from the ideas of others, was always in danger of being mis-read as non-compliance.

Teachers' understanding of and relationship to themselves

The policies and practices operating in the schools in which the teachers worked had a powerful effect on their sense of themselves. This sometimes reflected a deliberate intention to remould teachers' subjectivities. Sarah goes on to explain that, in her first year, she was told "we'll break you down, and we'll rebuild you". Part of this metaphorical process of "breaking you down" and "rebuilding you" involved creating subjects who internalize the external monitoring and surveillance and continuously monitor their own performance "because that is always there in the background, which I think fuels your self-motivation as well. 'What are the consequences of not doing this?'" Another violent metaphor was used by Sarah later on in the interview to describe the devastating emotional impact her experiences had on her confidence and, indeed, her whole sense of herself.

It actually destroyed me emotionally for quite some time. I actually thought, I can't teach, "I can't teach, I can't do this", and I was just distraught that five years of my life, and everything I'd worked so hard for was just destroyed just like that.

Like Sarah, Evan felt that the heavy scrutiny he experienced as a new member of staff had a significant impact on his emotional state and his sense of self: "I have, for the first time, experienced stress and anxiety. I think there are just things, language they use, and the way they kind of do things, it just keeps you on edge. I don't feel comfortable". Anxiety was something foregrounded by other teachers as well. For instance, Miranda linked anxiety to the long-term impact of the culture of monitoring and surveillance in her school: "I just feel very anxious around ... if someone's going to come in ... what's going to happen as a result of that. That was terrible ... and I think it leads you to a lot of second guessing, reflecting". For Jackie, a sense of anxiety and unease inhibited her professional creativity and led, as was no doubt intended, to compliance and conformity: "Wary ... I don't want to step a toe out of line. I'm very much always thinking about how I can make sure that everything is just spot on". For all these teachers a pervasive anxiety curtailed their professional growth, while having a corrosive effect on their mental wellbeing.

For others, such as Emma, it was the experience of being reprimanded for exercising agency and initiative that impacted most on their subjectivity: "I'm literally just coming in and doing what I'm told. I'm not engaging with this. I'm not allowing myself to emotionally invest in this, or mentally give my cognitive energy to this anymore. I'm just literally an automaton, and that, I think, sums it up". Ironically, after adopting this automaton state, Emma was "lauded" as an exemplary middle leader in the school by her manager. Like the workers in Fritz Lang's (1927) dystopian film, *Metropolis*, these eager and enthusiastic young teachers have been reduced to robotic cogs in an uncaring and instrumental performative machine.

When asked "what needs to change in teaching?", the teachers interviewed were unanimous despite their varied experiences. Common to all was the sense that teachers need to be trusted and to be granted greater agency in their work, as Emma explains:

There has got to be a sense of like, if I'm going to go back, I want to be clear I'm going to have a voice, because I've always been proud of the voice that I have [F]undamentally, there needs to be trust in the profession, and that runs from ... the Government need to trust the people in the profession, senior leaders need to trust their teachers, and so on and so forth.

Thus, it is precisely the lack of trust in teachers' voices, within multi-academy "trusts", which has undermined their relationships with pedagogy, compromised their relationship with leadership and, ultimately, eroded their professional self-concept. The gradual dismantling of teachers' confidence in their ability to think, make good judgments, and act in the best interests of students weakens teachers' ability to resist norms and to question their own or their colleagues' practices.

Our analysis of teachers' relations to their professional work, their colleagues and themselves highlights the destructiveness of a politics of identity in teaching in academy schools. Ironically, the sought after "effective teacher" in the academies is also expendable, interchangeable, and increasingly disengaged. We now turn to Foucault whose work invites us to consider friendship as a literal and metaphorical mode of being together in schools. We imagine non-normalizing relations to curriculum, colleagues and one-self—curiously wrought with betrayal, distance, and frankness—that Foucauldian friendship might offer teachers. Our intent is not to prescribe an alternative set of social relations in schools—a counter to what currently exists—but to acknowledge the possibility of a multiplicity of ethical relations and positions. Our project is about challenging the relational fabric of schools, without forming a program or agenda, by identifying the conditions for an experimental approach to relationships in "marginal spaces" (Kingston, 2009, p. 16) that are only minimally linked to the larger system.

Foucault's challenge: friendship as a form of life

The only acceptable motivation for following others into their worlds is friendship.

(Ellsworth, 1989, p. 317, fn 66).

Central to Foucault's work is the interest in overcoming "biopolitically administered identity at any cost, often without conscious knowledge of exactly why or of what lies beyond" (Roach, 2012, p. 130). In other words, he is interested in the politics of identity in

relation to various normative realms such as the law, health, gender and sexuality. In relation to the latter, rather than emphasize sexual desire as the basis of “an essential mode of being,” homosexuality is, for Foucault, a call to inventiveness and creativity; an opportunity to articulate a way of life beyond a politics of identity and “in the process multiply the ways one relates to others” (Albrecht-Crane, 2005, p. 429). Foucault offered the practice of friendship as a release from “the shackles of identity and interiority” (Roach, 2012, p. 5) and as such as a way of life.

A way of life can be shared among individuals of different age, status, and social activity. It can yield intense relations not resembling those that are institutionalized. It seems to me that a way of life can yield a culture and an ethics.

(Foucault, 1997a, p. 138).

Friendship, in this sense, is a form of localized resistance to institutional normalization, an affective relationship involving “the collaborative creation of new subjectivities and relationships as participants struggle to come to terms with one another” (Kingston, 2009, p. 9). Rather than falling back on social norms, Foucault’s concept of friendship privileges experimentation with many different kinds of relationships over traditional, institutional bonds. To suggest alternative relationships is not only to threaten the stability of normalizing and normalized relationships but also to threaten the social (and educational) order they support.

Foucault (1997a,b) was clear, therefore, that institutional relations such as family, profession and any kind of “obligatory camaraderie” (p. 136) has nothing to do with his concept of friendship. Institutional relations, in his view, are overly burdened by formalities that aim to establish and sustain prescribed roles, identities and outcomes and which rely on distancing authority. We would argue, however, that there is a need to mitigate “the dangers presented by the excessive and harmful normalization of relationships” wrought by institutional life (Kingston, 2009, p. 16). As the case of the English academy school has illustrated vividly, institutional life produces and reproduces hardened identities (e.g., “(in)effective teacher”) and “excessively rigid relationships” (e.g., teacher-principal hierarchy) and “perpetuate[s] practices that ought to be discarded” (e.g., high stakes testing and performance reviews) (p. 8). Friendship as a non-normalizing relationship might assuage those effects. This is not to say that friendships are devoid of power relations but that they may be “disconnected from the most totalizing and normalizing systems of power relations, and characterized instead by very dynamic and unstable systems of power relations” (Kingston, 2009, p. 8). The norms that apply may be fewer or of the kind that enable a temporary suspension of rules. As such, while institutional aims, norms and actions do not change drastically, friendship takes place apart from those aims and norms, “where context and chance constitute the ethical specifically because they are not *for* anything in particular” (Carusi, 2017, p. 644). This is not to say that friendship offers nothing: there is “a dynamic and dialectical interconnection between practices of the self and politics because one’s ability to resist normalization relies upon non-normalizing social relations” (McLaren, 2007, p. 196).

Yet Foucauldian friendship is anything but utopian. It is “founded in estrangement” and suspended between betrayal and loyalty, distance and closeness, brutal honesty (frankness) and sincere concern, the impersonal and the intimate (Roach, 2012, p. 7). Complicated affects—discomfort, frustration, love—are produced and impersonal ethics—nonrecognition, openness to alterity, respect for absolute difference—are encouraged. It is these curious, anti-social qualities of “friendship as shared estrangement” (p. 7)—and their a-subjective force—that create a surplus of affect uncontainable by institutional forms. In this excess lies the potential to imagine, to make, do, say, and think differently, calling into question understandings of a manipulable teacher identity characterized by blind loyalty, restrictive belonging, and compliant speech in the academy. Instead, Foucauldian friendship between and among teachers and school leaders invites betrayal, distancing, and frankness as tactics that work on, with and against the dominant school order. In the process teachers glimpse disidentification from the order of things.

Reconstituting a relation to curriculum, pedagogy and assessment via betrayal

Relations to knowledge inevitably shape teachers’ relation to the normalizing authority of the school as an institution (Castioradis, 1968). We have illustrated how steadfast adherence to a mandated curriculum, pedagogy and assessment defines teachers’ relationship to their professional work in the English academies. The fatalistic sense that “unless we do X, we will fall behind” (Smith, 2000, p. 18) and unless we do X within this time slot, destroys any particular, local relationship of the teacher (and ultimately the student) to knowledge and experience—as living practices of self- and social formation. The formulaic relation that does exist is non-negotiable, leading to either silent deference, on the one hand, or secret-keeping and fear of exposure, on the other.

Consistency, and the blind loyalty it requires, is at odds with thought. Thought requires betrayal—a form of indifference to norms and normalizing forces, if you will—one condition of Foucault’s concept of friendship. Unlike other relations such as marriage that constitute and are enacted in the private sphere, friendship is properly communal and signifies the common (Roach, 2012). Not to be confused with the “sameness” imposed on teachers in the academies by curriculum prescription and pedagogical predictability, the common is constituted in and as a space of movement, fluidity, and incoherence that “passes under the radar and off the grids of biopower” (Haver, Reading Foucault’s Genet Lectures, p. 7 in Roach, p. 146). At a tactical level, betrayal can be as simple as intermittently posing difficult questions or sharing an alternative point of view during a faculty meeting thereby diverting the school agenda. It might involve a teacher inviting school leaders into one’s classroom on occasion, rather than waiting for them to show up unannounced; and rather than accepting feedback on performance, manipulating those moments as opportunities for dialog about practice. Surprising a teaching colleague with an offer of collaboration can betray the competitive ethos of the school

temporarily. Small and seemingly insignificant acts of betrayal gradually dismantle the “architecture of fortification” (Foucault, 1988, p. 139) that the dominant school order seeks to maintain because there are no bannisters to support nor expectations to guide; each occasion for thought relies on openness to alterity and respect for difference (Roach, 2012).

Moreover, efforts to switch the register from compliance to conversation invite others to give up their grounding in any particular ideology—to disidentify from the school order temporarily—and to become a person who enters into “a movement of thought that is groundless” (Kaufman, 2001, p. 131).

These thinkers obey impersonal rules of hospitality—hospitality to the other as a form of thought and not as a person ... a new form of intellectual hospitality, a mode of being in common that is not a form of correcting or out-mastering the other, but rather a way of joining with the other in language or in thought so that what is created is a community of thought that knows no bounds, a hospitality that liquidates identity, a communism of the souls.

(Kaufman, 2001, p. 141).

Teachers engaging in “impersonal friendship that traverses thought” (Sellar, 2012, p. 71) experience the boundlessness of thought; they become “intoxicated,” “driven by curiosities that are closer to pleasure, to play and to wandering” (Arsenjuk and Koerner, 2009, p. 8), and infectious as they convey such curiosity to their students. Such engagement offers a temporary suspension of economic, social, cultural, religious or political forces pervading schools that otherwise wish to direct teachers.

For teachers engaged in dialog, “the world becomes something of a concern, something to think about [together], something that provokes ...” (Masschelein, 2011, p. 535); and, we would add, evokes a surplus of affect that cannot be contained by institutional forms and therein lies the potential to act, make, say and think differently. Teachers are reminded of their own desire to think and to care about their educational decisions. In the back and forth of dialog, for example, each interlocutor moves in and out of control of discourse. The identities and positions established and operating within and through institutional norms—experienced leader versus inexperienced neophyte, or, progressive versus reactionary—become disjoined and participants “produce new modes of being together” (Albrecht-Crane, 2005, p. 492). The preservation of one another’s perspectives—one another’s strangeness (otherwise annihilated through identification with the school’s singular ideology)—becomes paramount while one’s positional identity becomes ambiguous. This is because during dialog identity is not at stake, the world is.

The teacher as thinker has a relation to thought that is not reducible to the accumulation of information (the cumulative effect of power) or to the logic of (de)professionalization that governs the work of the English academy school. Teachers as thinkers betray their initial thinking; they have second thoughts about the tradition or ideology they have heretofore accepted. With non-reliance on identity (or ideology) there may be “a loss of self-assurance” (Sellar, 2012, p. 72) but the form of power at play is neither that which induces an ontological insecurity (Ball, 2003) nor evokes guilt, for being inconsistent with school policy. Rather the teacher is released from predetermined identification and association and invited to participate in creative and critical thinking with others.

Reconstituting non-hierarchical relationships with colleagues via distancing

Disciplinary power links the forces of hierarchical observation, normalizing judgment, and examination together to generate and sustain “the normal”—what is considered “effective.” In the English academy, disciplinary power is a crucial means of ensuring teachers’ compliance to the school dress code, pedagogical creed, and curriculum practices (Foucault, 1975). In the process not only is the “normal” teacher produced but the “abnormal” teacher—the one who does not “fit”—is easily identifiable, differentiated and excluded. Confessional in quality, hierarchical relations offer only conditional belonging, reflecting the power of the school leadership to withdraw their favor when the teacher fails to comply. Teacher autonomy is traded for a particular kind of belonging—a calculable love (Phelan and Vintimilla, 2020). To think for oneself, to offer an opposing view, or to practice in ways inconsistent with the school mandate is to experience the power of the leader who has the authority to establish blame and banish at will (Foucault, 1997a,b).

Friendship as shared estrangement can offer some respite from hierarchical power because it is not based on identity-based hierarchical divisions but suggests a way of life shared among individuals of different roles and statuses and so forth; involving “polymorphic, varied and individually modulated relationships” (Foucault, 1997a, p. 139). There is no getting away from power differentials in institutions but this is not to say that mutual trust is incompatible with unequal power. Perhaps this is so because friendship need not be “an ‘all or nothing’ relationship” (Sellar, 2012, p. 70). There may be some aspects that teachers and school leaders share—their interest in the well-being of students—but the enactment of that interest may vary. To share some sense of educational purpose can be a source of belonging for educators but this need not diminish nor erase teachers’ capacity to deliberate, to offer opposing views, to think and judge otherwise on occasion. The educational challenge is not to discover the truth in one’s teaching or leading, but rather to use both to arrive at “a multiplicity of truth-making relationships” (Albrecht-Crane, 2005, p. 493). If teachers are not to be simply objects of policy but advocates of education beyond specific institutional concerns, they have to engage in debate not only with one another (at grade, school, and district levels) but also with a range of professional others. Existing educationally as well as politically in this way demands that teachers are able to stand alone, together (Phelan, 2015).

As an ambivalent social relation, Foucauldian friendship offers the possibility of not belonging too much. As such, it “deliberatively avoids relationships that lead to attachment or possessiveness” (Roach, 2012, p. 146); friends reduce their reliance on recognition by others for a sense of their own competence and worth—a subtraction of power from one’s “destructive ego” (p. 146).

Preserving distance while working closely with colleagues in this way requires entrustment (Roach, 2012), meaning that each is secure in their distinct role in the collaborative effort of generating truth-making relationships. Entrustment is an experience in difference where each is engaged in a mutual striving toward new selves for all involved.

The possibility of generating such a multiplicity of truth-making relations is already evident in the English academy school in the form of “playful irreverence” to norms and identities (Albrecht-Crane, 2005, p. 494). Power shifts slightly when teachers in the academy take recourse to humor with their “in-jokes” about “Big Brother” and the “Soviet Union.” That such excess is possible, though privately, attests to the flexibility and instability of institutional structures and the possibility of friendship among colleagues as shared estrangement, of community that does not cohere in identity. The opening is a form of relation that exceeds institutionalized utility; a relation of “unbelonging” “generated in common” and which can “regenerate the common” (Roach, 2012, p. 15), referred to in the previous section. What if a few teachers arrive at school with elbows showing and refuse to wear their jackets all day as required? What if someone wears “a jazzy jacket”? We are reminded of tee-shirts worn by teachers, in one school district in British Columbia, Canada, reciting the ironic words of Dr. Seuss’ (1958) feisty little turtle who dared to question the demands of the despotic King Yertle: “I know up on the top you are seeing great sights, but down here on the bottom, we too should have rights” (Phelan, 2015).

In the context of dialog among teachers and school leaders in the staffroom, for example, or at professional development workshops, such playfulness might have a broader application. A playful response at a faculty meeting—spoken in jest but also in earnest—gives colleagues pause, leaving them surprised and even speechless. The assumed consensus is interrupted. Yet the teacher’s joke is not simply negative; it is generative and affirmative because it ushers in other ways of responding to the status quo. Play introduces a difference, an unexpected, unplanned, and unanticipated encounter between bodies and in the moment of mutual discomfort there is the possibility of singular and collective experimentation, other pleasures, other concerns, and other selves.

Reconstituting a relation with self via frankness

In the English academy school teachers feel silenced yet wonder why no one says “no;” teachers knowingly comply but express anxiety in doing so; teachers internalize the external gaze yet openly critique it during conversations with us. The teachers’ utterances suggest a preoccupation with being compliant and docile subjects but they also suggest a struggle against their subjection as such. Teacher’s reflections indicate the possibility of cultivating “a relation of reflexivity” (Halperin, 1995, p. 76–77), “using one’s relation to oneself as a potential resource with which to construct new modalities of subjective agency and new styles of personal life that may enable one to resist or even to escape one’s social and psychological determinations” (p. 76–77). The kind of reflection at play here is one of a “detached and disciplined concern” much like a sculptor stepping back to view her work (Roach, 2012, p. 49).

Friendship as shared estrangement can provide conditions for reflexivity because it relies on frankness or truth-telling (*parrhesia*) in Foucauldian terms. *Parrhesia* is similar to but not the same as conscience; it is a virtue, a quality, and a technique that helps us overcome self-flattery. In a moment of truth-telling, a teacher challenges a colleague, be it a teacher or school leader, to go beyond their limits, their former reliance on the comfort of past practices, for example. As friends, colleagues are confronted with the possibility and the challenge of self-transformation. Unlike confessional discourse that engenders dependence on another (who has power to absolve “sins” or not) and requires the objectification of the self to speak its truth, in truth-telling “[t]he self is not objectified but subjectivated, the self becomes the subject of true discourse and is transformed in the truth’s enunciation” (p. 28). *Parrhesia* must be dictated by generosity and is neither manipulation via persuasion nor imposition of one’s will (p. 26). The intent is to encourage someone toward self-sovereignty and not to provoke “a guilty interdependence” (p. 28). Of course, flattery can also create a dependence whereby teachers rely too much on one another to think well of themselves and to think for themselves.

In friendship as shared estrangement, Foucault marks a shift from a solitary to a collaborative esthetics of existence via *parrhesia*. As an anti-confessional discourse, friendship entails a mutual striving toward new selves (Roach, 2012, p. 37–38); hence there is a different modality of knowledge and a different relation to truth involved. Truth-telling or *parrhesia* involve the obligation to be forthright with the other and to say whatever is on one’s mind.

If the self is only a relation between subjection and subjectivation, then the true friend, the *parrhesiastes*, pushes the subject beyond historical determinants in encouraging “processes of subjectivation over procedures of subjection.”

(Roach, 2012, p. 41).

If teachers are to become friends, in the Foucauldian sense, then each must push the other beyond historical determinants in encouraging “processes of subjectivation over procedures of subjection. The friend is thus the *sine qua non* of self-transformation and immanent salvation” (Roach, 2012, p. 41). Teachers and school leaders compel one another to understand the self and the world in a relational manner.

Friendship as shared estrangement can be thought of as creating the conditions for a “concrete form of freedom”—a place of “extensive work by the self on the self” in the company of others. The goal is to shape “an ethos that is good, beautiful, honorable, estimable, memorable and exemplary” (Foucault, 1997a, p. 286).

Fearless speech among colleagues always carries a risk of course—hurting others by our remarks, damaging a fragile relationship, or losing the friend altogether. Brutal honesty may be necessary because as Forti (2015) argues the root of systemic evil is passivity and obedience, the “normativity of non-judgment,” which arises from the inability to “create friction with the present” (p. 253). Fearless speech requires that teachers be honest with one another and with their leaders, often against their own better judgment about what might be good for them in the future—garnering rewards and recognition; keeping their jobs as teachers; and achieving promotion. The impetus to “stay calm and carry on” is culturally embedded and very reasonable but such desires must become the target of critique if teachers are to defy ethically and educationally unendurable conditions for themselves and their students in the present (Woodford, 2020).

Our point, however, is not to mount resistance to schools as instruments of capital, social engineering and ideological stultification. We are not in pursuit of the teacher hero with a moral obligation to speak truth to power and to save themselves and their students in the process. We are asking, however: “what is it about identities that elicit such strong, visceral, overwhelming responses?” (Albrecht-Crane, 2005, p. 496). Does the identity offered by the English academy school—achievement-oriented, effective, one of us—offer a sense of certainty, safety and stability? Or is it that “the process of identity formation, predicated on a dualistic logic of inclusion and exclusion, makes its way into the body, into the ways individuals govern themselves” (p. 497). Feelings of disenchantment or despair are extremely difficult to live with; the only recourse may be to go along. The same feelings, however, may have a provocative and critical role in removing teachers from the false promise of certainty. In fact, Carusi (2017) argues that the repudiation of their profession—either by leaving it or not recommending it others—is an ethical response to the normative order of effective teaching. Teachers who choose defection or exodus from the profession are critiquing the present—speaking truth to power and about power—and interrupting the smooth functioning of the normative order. In doing so these teachers, many of whom leave with no certainty about what lies ahead, are also “provok[ing] movement without yet projecting a future” (p. 633). Here we witness the politics of friendship as a form of negative ethics.

Conclusion: friendship as a form of negative ethics in teaching

A politics of identity can have a devastating impact on the professional lives of teachers. Friendship as shared estrangement may be one ethico-political response to the normative order of “effective teacher.” Embracing friendship as a politics of the profession constitutes a form of negative ethics.

Friendship cannot be programmed nor prescribed. As a way of life, it is non-coincident with the normative system within which it is embedded. It is an attempt to live suspended between the norm (and actualization on its terms) and a blatant resistance to the norm—a refusal in the form of “I’d prefer not to” (Agamben, 1999). As a negation of the positivity produced by the norm, refusal is an ethical response to and within the normative; however, it “breaks from the normative in such a way that [the] normative order is unable to recuperate that ethical break” (Carusi, 2017, p. 636). Having no content per se, the ethical arrives as a negation to the normative order.

The ethical undoes the context of the normative and actively interrupts the smooth operation of a normative order. The ethical displaces norms and normative orders in ways that “go beyond that positivity” and by “detect[ing] the points it [positivity] fails to constitute itself” (Laclau, 2014, p. 128) we can align the ethical with negativity. The ethical acts against the normative not as a competing positive expression, but as an active “not that”. The ethical is precisely what is unimaginable by and does not follow from the normative.

(Carusi, 2017, p. 636).

Foucauldian friendship—as a form of shared estrangement—can displace but *not* replace norms and normative orders. It offers a radical, but temporary and retrospective, break. There is no logical transition between the political identity of *effective teacher* to experimental relation(s) of friendship as shared estrangement. There is, however, antagonism and provocation. Once we begin viewing friendship as shared estrangement as a new normative—an alternative positivity—it “sheds its negativity and status as ethical” (p. 637). Our intention, therefore, is not to prescribe a given content to teacher identity but to invite consideration of what Agamben terms a “whatever-being”—a whatever teacher, if you will. The term illustrates that it is a being that is outside binary-logics (logics that are installed, for example, by law and norms) and it is a being that is defined in its negativity (“as not”), a being, without a particular identity, personhood, belonging and capacity (Zevnik, 2009, p. 90). Such a being does not make any sense within the existing social and symbolic order. On the contrary, a “whatever-teacher” cannot be grasped within a given educational order because one is not recognizable and cannot be placed either “outside” or “inside” this order. No clear borders between inside and outside can be installed.

In friendship as shared estrangement, Foucault offers a way of life as a practice of freedom that forges a line of flight from the regime of identification as “effective teacher” by creating a critical distance, deflecting commonality and refusing transcendence. His model of friendship becomes “a micro-model for alternative communal and political forms” (Roach, 2012, p. 109) in our schools.

References

- Agamben, G., 1999. *Potentialities: Collected Essays in Philosophy* (D. Heller-Roazen, Trans.). Stanford University Press, Stanford, CA.
- Albrecht-Crane, C., 2005. Pedagogy as friendship: identity and affect in the conservative classroom. *Cult. Stud.* 19 (4), 491–514.
- Arsenjuk, L., Koerner, M., 2009. 'Introduction.' *Polygraph* 21, 1–13.
- Ball, S.J., 2003. *Education policy and social class: The selected works of Stephen J. Ball*. Routledge, London.
- Biesta, G., 2010. *Good Education in Age of Measurement: Ethics, Politics, Democracy*. Paradigm Publishers, Boulder, CO.
- Biesta, G., 2011. The ignorant citizen: Mouffe, Ranciere and the subject of democratic education. *Stud. Philos. Educ.* 30 (2), 141–153.
- Butler, J., 1997. *The Psychic Life of Power*. Stanford University Press, Stanford, CA.
- Britzman, D., 2009. The adolescent teacher: a psychoanalytic contribution to developing education. In: Paper Presented at the Mathematics Education and Contemporary Theory Conference Manchester Metropolitan University.
- Carusi, F.T., 2017. Why bother teaching? Despairing the ethical through teaching that does not follow. *Stud. Philos. Educ.* 36, 633–645.
- Castoriadis, C., 1968. The anticipated revolution. In: Curtis, D.A. (Ed.), *Political and Social Writings*, 3: 1961–79: Recommending the Revolution: From Socialism to the Autonomous Society. University of Minnesota Press, Minneapolis, MN.
- Clarke, M., Mills, M., 2021. 'We have never been public': continuity and change in the policy production of 'the public' in education in England. *Eur. Educ. Res. J.* 1–16. <https://doi.org/10.1177/1474904121990477>.
- Clarke, M., Phelan, A., 2017. *Teacher Education and the Political: The Power of Negative Thinking*. Routledge, London, UK.
- Clarke, M., 2009. The ethico-politics of teacher identity. *Educ. Philos. Theor.* 41 (2), 185–200.
- Courtney, S.J., Gunter, H.M., 2015. Get off my bus! School leaders, vision work and the elimination of teachers. *Int. J. Leader. Educ.* 18 (4), 395–417.
- Day, C., Elliot, B., Kington, A., 2005. Reform, standards and teacher identity: challenges of sustaining commitment. *Teach. Teach. Educ.* 21 (5), 563–577.
- Day, C., Kington, A., Stobart, G., Sammons, P., 2006. The personal and professional selves of teachers: stable and unstable identities. *Br. Educ. Res. J.* 32 (4), 601–616. (Accessed 21 June 2020).
- Department for Education, 2020. Open Academies, Free Schools, Studio Schools and UTCs [internet]. <https://www.gov.uk/government/publications/open-academies-and-academy-projects-in-development>. (Accessed 29 May 2020).
- Ellsworth, E., 1989. Why doesn't this feel empowering? The repressive myths of critical pedagogy. *Harv. Educ. Rev.* 59 (3), 297–302.
- Forti, S., 2015. *New Demons: Rethinking Power and Evil Today*. Stanford University Press, Stanford, CA.
- Foucault, M., Sheridan, Alan, 1975. *Discipline and Punish: The Birth of the Prison*. trans. Vintage Books, New York.
- Foucault, M., 1983. On the genealogy of ethics: an overview of work in progress. In: Dreyfus, H., Rabinow, P. (Eds.), *Michel Foucault: Beyond Structuralism and Hermeneutics*, second ed. University of Chicago Press, Chicago.
- Foucault, M., 1988. trans Richard Hurley. *The Care of the Self: The History of Sexuality*, vol. 3. Vintage Books, New York.
- Foucault, M., 1997a. Friendship as a way of life. In: Rabinow, P. (Ed.), *Michel Foucault: Ethics, Subjectivity and Truth*. The New Press, New York.
- Foucault, M., 1997b. The hermeneutics of the subject. In: Rabinow, P. (Ed.), *Michel Foucault: Ethics, Subjectivity and Truth*. The New Press, New York.
- Fournier, V., 1999. The appeal to 'professionalism' as a disciplinary mechanism. *Socio. Rev.* 47 (2), 280–307.
- Furlong, J., 2008. Making teaching a 21st century profession: Tony Blair's big prize. *Oxf. Rev. Educ.* 34 (6), 727–739.
- Gilbert, J., 2014. *Common Ground: Democracy and Collectivism in an Age of Individualism*. Pluto Press, London, UK.
- Halperin, D., 1995. *Saint Foucault: Towards a Gay Hagiography*. Oxford University Press.
- Kaufman, E., 2001. *The Delirium of Praise: Bataille, Blanchot, Deleuze, Foucault, Klossowski*. The Johns Hopkins University Press, Baltimore and London.
- Kington, M., 2009. Subversive friendships: Foucault on homosexuality and social experimentation. *Foucault Stud.* (7), 7–17.
- Laclau, E., 2014. *The Rhetorical Foundations of Society*. Verso, London.
- Lang, F., 1927. *Metropolis*. Kino International, In. Germany.
- Larsen, M., 2010. Troubling the discourse of teacher centrality: a comparative perspective. *J. Educ. Pol.* 25 (2), 207–231.
- Lewis, T.E., 2015. Suspending the ontology of effectiveness in education: reclaiming the theatrical gestures of the ineffective teacher. In: Lewis, T.E., Lavery, M.J. (Eds.), *Art's Teachings, Teachings' Art: Contemporary Philosophies and Theories in Education*. Springer, Dordrecht, pp. 165–178.
- Masschelein, J., 2011. *Experimentum scholae: the world once more ... but not (yet) finished*. *Stud. Philos. Educ.* 30 (5), 529–535.
- McLaren, M.A., 2007. *Feminism, Foucault, and Embodied Subjectivity*. SUNY Press, Albany, NY.
- Mockler, N., 2011. Beyond 'what works': understanding teacher identity as a practical and political tool. *Teach. Teach.* 17 (5), 517–528.
- Phelan, A., 2015. *Curriculum Theorizing and Teacher Education: Complicating Conjunctions*. Routledge, London.
- Phelan, A.M., Vintimilla, C.D., 2020. Autonomy as responsibility in professional life: deconstructing teaching standards. *Curr. Res. J.* 1 (2), 26–37.
- Phillips, R., Furlong, J., 2002. *Education, Reform and the State: Twenty-Five Years of Politics, Policy and Practice*. RoutledgeFalmer, London, UK.
- Pinar, W.F., 2006. *The Synoptic Text Today and Other Essays: Curriculum Development after the Reconceptualization*. Peter Lang, New York.
- Roach, T., 2012. *Friendship as a Way of Life: Foucault, AIDS, and the Politics of Shared Estrangement*. SUNY Press, New York.
- Sellar, S., 2012. 'It's all about relationships' Hesitation, friendship and pedagogical assemblage. *Discourse* 33 (1), 61–74.
- Seuss, Dr., 1958. *Yertle the Turtle and Other Stories*. Random House Books, New York, NY.
- Shahjahan, R.A., 2011. Decolonizing the evidence-based education and policy movement: revealing the colonial vestiges in educational policy, research, and neoliberal reform. *J. Educ. Pol.* 26 (2), 181–206.
- Smith, D., 2000. The specific challenges of globalization to teaching and vice versa. *Alberta J. Educ. Res.* 46 (1), 7–26.
- Taubman, P., 2009. *Teaching by Numbers: Deconstructing the Discourse of Standards and Accountability in Education*. Routledge, New York.
- Thomson, P., 2019. Oh to be in England? The production of an un-public state system. In: Wilkinson, J., Niesche, R., Eacott, S. (Eds.), *Challenges for Public Education: Reconceptualising Educational Leadership, Policy and Social Justice as Resources for Hope*. Routledge, Abingdon, pp. 101–115.
- Woodford, P., 2020. Review of: Forti, simona. In: *New Demons: Rethinking Power and Evil Today*. Stanford University Press, Stanford, CA.
- Zevnik, A., 2009. Sovereign-less subject and the possibility of resistance. *J. Int. Stud.* 36 (1), 83–106.

Global politics of quality in education

Jaakko Kauko, Faculty of Education and Culture, Tampere University, Tampere, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	49
Peeling the evaluation onion	50
Conclusion	52
References	52

Introduction

Quality of education is political. This becomes apparent when we analyze the history of each word in the previous sentence. During the 14th and 15th centuries, “quality” began to change from something unattainable (as Plato described) to something marking social rank, only to become something defined and measured during the modern period (Dahler-Larsen, 2019, 31–36). In the same period, “education,” alongside the purpose of learning, has become a vehicle for the state to foster economic growth, control key professions, and indoctrinate citizens (e.g., Foucault, 1977; Apple, 1980; Meyer et al., 1992). The massification of education has led Meyer (1986) to argue that education is the “religious basis” of the modern state. What constitutes the “political” is also under constant debate and contestation. In the long view of history, its definition has expanded from distant state affairs to the personal. Thus, quality of education is not only political, but is reforming in the flow of history.

In seeking to understand what quality of education is, it is worth distinguishing between the questions “what is quality education?” and “how is quality of education measured?” In studying quality, an abstract and fluid concept that usually has a positive connotation, one soon discovers that it has no single definition. It sometimes refers to something you simply recognize when you see it (an argument coined by some school inspectors, for example), but this both underlines the subjectivity of quality and its conceptual roots as something unattainable. When quality is operationalized as something measurable, its nature changes. Measurement inevitably changes the power relations between the subjects and objects of measurement, or as Dahler-Larsen (2012) conceptualizes it, creates “constitutive effects.” The politics of quality has observable power effects when quality is operationalized. In other words, the act of measurement makes visible the politics behind the subjective definition of quality.

The concept of Quality Assurance and Evaluation (QAE) has been used to cover the varying array of quality measurement activities, from self-evaluation to national large-scale assessments, and with different purposes, from development to control. When an umbrella concept like this is used, interest shifts to the effects of an array of such different types of evaluation, such as school inspection, auditing, different benchmarks, and indicators. In this sense, QAE is an attempt to combine understandings of the dynamics of what happens when quality is evaluated (Centeno et al., 2018; see Ozga et al., 2011). It carries the risk of essentialization, but it also provides a partner large enough for the concept of quality. We should keep this in mind when examining the different manifestations of QAE.

Quality of education therefore has two important and very diverse aspects: what is considered quality; and how it is manifested in QAE. In combining these two aspects, the understanding of politics from Giorgio Agamben’s *Means without End* (2000, 115–116) is helpful:

What is in question in political experience is not a higher end but being-into-language itself as pure mediality, being-into-a-mean as an irreducible condition of human beings. *Politics is the exhibition of a mediality: it is the act of making a means visible as such.* Politics is the sphere neither of an end in itself nor of means subordinated to an end; rather, it is the sphere of a pure mediality without end intended as the field of human action and of human thought. [Original emphasis]

For Agamben, politics is the revealing of this mediality. Analogically, the politics of quality is making visible that QAE mediates what quality is. When the mediation of what quality is becomes visible, the understanding of quality is no longer something unattainable or abstract, but becomes political and concrete. When we observe what QAE is in concrete terms, and how it functions, we leave behind the idealistic assumptions that quality is good for everyone and its unpoliticized claim. In very simple terms, the politics of quality is understanding what QAE is, and the kind of repercussions it has for action (see Kauko et al., 2018a).

This understanding of the politics of quality is not aligned with the communication of what constitutes quality. Indeed, the usual take on the politics of quality hides its mediality. The current agenda of international organizations working in the field of education is that education needs more quality, and that progress in this can be measured. The Incheon Declaration (2016, 66) by UNESCO, other UN organizations, and the World Bank Group claims that “... research and assessment culture is necessary at the national and

international levels.” Ozga et al. (2011, 2) note that “the integration of quality assurance into organizational ... thinking has interesting implications. Quality now clearly becomes something that cannot be left to chance. ... It is no longer extraordinary. It becomes simply what can be expected. ... Quality is conformity with standards.” For a policy to work smoothly, the politicizing around it should be minimized. In quality of education, the more the technical and bureaucratic dimensions are emphasized, the less room there is to ask “why?”

The two complementary yet opposite understandings of politics of quality, revealing the mediality of QAE and the agendas that cover this mediality, communicate that politics is the organization and reorganization of contingency (Kauko, 2014; Palonen, 1993, 2003, 2006). It is an attempt to channel education policy contingencies with the help of quality. Quality’s positive and malleable attributes make it a fine tool for this purpose.

Education-related testing was already in place in medieval times, for example, in the Imperial Examination in China (Zhou, 2019, 28), but quality as an indicator began to play a larger role in education policy in the Global North during the 1970s and 1980s. The main examples are the foundation of the IEA (International Association for the Evaluation of Educational Advancement) and its evaluations since 1958, the OECD’s (Organisation for Economic Co-operation and Development) social indicator programs during the 1970s and its International Indicators of Education Systems (INES) program from 1988, and the 1983 United States’ Nation at Risk Report. Different countries from every part of the world have steadily followed suit, not necessarily joining the international large-scale assessments, but testing evaluation nationally (Kamens and McNeely, 2010). To name only a couple of examples: Portugal has witnessed a growth in the use of national testing since the 1990s (Carvalho et al., 2020); while following a wave of decentralization in the 1990s, all the Nordic countries rushed toward more QAE with varying intensity (Dovemark et al., 2018). The understanding that education quality needs to be tested and measured is expanding, and it has also reached places previously untouched by the idea of accountability and standardization (Verger et al., 2018). Since the 1990s, southeastern sub-Saharan African countries have started to develop assessment (Kamens and McNeely, 2010). Sahlberg (2016) describes a “Global Educational Reform Movement” that supports competition, choice, prescribed curricula, standardized testing, and privatization. It is fair to say that the global mainstream in education favors more QAE, rather than less. In this model, there are usually elements of accountability, standards, and decentralization.

Drawing on previous projects in which my colleagues and I have been involved, this article takes the examples of Finland, Brazil, China, and Russia to analyze the politics of quality in education. It uses these examples in seeking to understand the extent to which there is a global education quality policy. Except perhaps for Finland, the case countries are not common references in the western literature, but they represent a significant proportion of the world’s population. The main focus is on general school education at primary and secondary levels.

Peeling the evaluation onion

The onion, a simple but versatile ingredient in many traditional cuisines, is a delicious metaphor for QAE. QAE in education has many layers, and their properties and use vary, depending on the context. As with quality, most people would claim to recognize an onion when they see one, despite the richness of varieties. When starting to peel the layers, it becomes clear that the layers are more interesting than its core, quality. Depending on the onion, the useable part can be the leaves (chive), inner layers (common onion), cloves (garlic), or the whole plant (scallion). Onions, like QAE systems, share some properties like layers, but their flavor and use vary according to context. A basic structure of QAE systems is participation in international large-scale assessments, different national QAEs, and regional or local QAE.

Finnish education is steered at the national level, and in the autonomous municipalities at the local level. The post-Second World War expansion, centralization, and nationalization of education faced a new era after the 1980s and 1990s. During this period, the role of the municipalities in education governance became more important as their autonomy was strengthened: It entailed the gradual abolition of school and textbook inspection, and a move toward lump-sum funding for education providers. This was soon countered by an increasing emphasis on local-level QAE, but without directive national guidelines. Currently, the main providers of basic education are the more than 300 municipalities. Kauko et al. (2020) describe Finnish QAE in education as triple layered. The first layer is the international evaluations, in which Finland has been very active. Despite the fame of Finland as a high achiever in PISA, this outer layer is politically detached from the inner national layers. In this approach, national data is considered more important and relevant. In other words, international evaluations are less important for steering than national ones (see also Seppänen et al., 2019). External evaluation’s power is limited in national QAE as well. There is systematic gathering of data as well as frequent and strategy-based evaluations. However, there are no high-stakes test-based institutional accountability measures and no national standardized tests for age cohorts (the standardized general upper secondary exam, matriculation, is an important tool for monitoring outcomes, but not for controlling resources, for example). The remaining tests are sample-based, and their use has followed principles of no public data, no ranking, no standardized tests, and no inspection (Kauko et al., 2020; Varjo et al., 2016). The third layer of self-evaluation by education providers plays an important role in evaluation. They are obliged to evaluate their own education and its effectiveness, as well as to take part in external sample-based and thematic evaluations. In sum, the Finnish layers are rather separate: The international layer has less effect on local-level schooling, national QAE serves as an information provider for national decisions, and local-level QAE is intended to help the local authorities in their work.

The organization of Brazilian education QAE is too complex to describe in full. It would require an account of the organization of education at the federal level, in 26 states and a federal district, and in around 5500 municipalities. These different entities can have their own QAE measures (Rambla, 2012). Cooperation between these entities varies and can be very limited (Azevedo and Santos, 2012). Brazil has had an influential education community, which developed QAE measures during the 20th century. This was facilitated by the federal government, with varying connections with the USAID and the World Bank (Kauko et al., 2016; Centeno, 2010). One turning point was decentralization, a common feature in post-dictatorial South American countries that was recommended by the World Bank (Meade and Gershberg, 2008), which then created a typical environment requiring a soft steering tool such as federal QAE. During the 1990s, Brazil developed federal standardized tests. At the turn of the millennium, it participated in UNESCO and OECD international large-scale assessments. After this, the systematization, strengthening, and broadening continued nationally (Kauko et al., 2016; Suominen et al., 2018). At the same time as Brazil has developed its federal QAE capacity, the role of the international organizations has diminished and focused on state and municipal projects. In addition, the role of civil society movements utilizing QAE data has become significant. (Kauko et al., 2018b). However, many governments in the past have aspired to improve the country's PISA ranking (Kauko et al., 2016). In general, the federal governments of Brazil have had an interest in being part of the PISA survey and have also been active in other ILSAs. Internally, the federal standardized tests, and different state and municipal level QAE measures, create a web of interconnected testing and a strong regime for their use in distributing resources and local decision making.

Chinese education is a huge entity. The current system is organized under the national Ministry of Education, with provincial, city, and county or district administrative layers (Zhou, 2019). In the twentieth century, after the establishment of the People's Republic of China in 1949, the imperial tradition—a decentralized, private, and elitist system—was transformed into a nationalized, centralized, and reformed education system, following the Soviet model. The country's turmoil during the Cultural Revolution seriously disrupted education and the inspection office (Suominen et al., 2018). A new development slowly started after the opening-up policy of the 1980s and 1990s. The opening-up policy and the intensification of the education and quality reforms since the 1990s create an important backdrop for Chinese QAE politics. On one hand, the reforms have bolstered the official position of the party state, but on the other, there are more implementation responsibilities at the local level (Suominen et al., 2017). Historically, Chinese QAE has involved inspection and testing. The current QAE structure consists of an inspection and school quality management system, as well as teacher training evaluation. They target students, teachers, headteachers, schools, and education governments and departments (Zhou, 2019, 72). In addition to selection, testing includes different types of evaluation of learning and school outcomes, as well as high-stakes tests. Public rankings have been officially forbidden since 2010, but they do exist (Zhou, 2019, 43). There is a certain degree of decentralization to the county level in basic education (Kauko et al., 2018b), but the state role is strong compared with Brazil (Rinne et al., 2018). The data-driven large-scale assessment layer is fairly new in China. The introduction of NAEQ (National Assessment of Education Quality) was a marked shift. It was first piloted in this sense in 2007. Its purpose is data gathering, with standardized tests for official and political use (Suominen et al., 2018, 55). (Zhou, 2019, 50) summarizes the work of other scholars by saying that the large-scale assessment, NAEQ, is a result of international emulation by China. NAEQ was also seen as a solution for systematic, nationwide, and first-hand data (Piattoeva et al., 2018). China as a whole has not participated in international large-scale assessments, but its regions have been involved in PISA. The diverse Chinese education landscape is subject to many different QAE measures. However, international ones are not strongly present, although there is some evidence that they may have affected the interests in strengthening the national evaluation scheme.

The Russian education system is also a large entity, with two federal ministries and an agency for control and supervision, as well as state, municipal, and private education providers. During the Soviet era, Russian QAE was based on inspection, school reports, and a school census. Much of QAE also worked *ex ante* as trust in the centralized curriculum and teacher training, or relied on the delegation of supervision to the local authorities (Suominen et al., 2018, 56–57). Under-resourced education faced inequality and quality problems. Reforms were started during *perestroika* and implemented after the dissolution of the Soviet Union, and they resulted in a strong decentralization of education, accompanied by partial privatization and commercialization. The reform era of the 1990s was accompanied by the involvement of the OECD and the World Bank, which advocated a standardized testing regime (Suominen et al., 2018, 58). Russian QAE has evolved from the efficiency idea of the 1990s and the broadening and systematizing QAE between 2000 and 2010 into an era of more nuanced data use (Gurova et al., 2015). The Russian QAE reforms in the 2000s were partly a response to the criticism of decentralization. The reform package of the 2000s introduced the influential high-stakes Unified State Exam (USE) and the State Final Attestation (SFA) (Gurova et al., 2018, 2015). Test scores are used as part of quality evaluation for students, school staff, and the education authorities. The local use of their data varies (Gurova et al., 2018). Russia has been actively involved in PISA and TIMSS, and their data is valued. However, national-level experts also emphasize the importance of national traditions (Kallo et al., 2018, 173). In sum, the Russian quality of education landscape is multifaceted in its plurality of decentralized education providers and local influences, while strongly aspiring at the state level to gain control with the aid of QAE.

In peeling the QAE onions of Finland, Brazil, China, and Russia, we can recognize the different layers of QAE. At the same time, the complexity of the structures is striking. The accounts above have been far from exhaustive, but it is hoped they have demonstrated that each country has traditions in education QAE that have not faded in the globalized environment. While education research sometimes tends to focus on the discussion of international large-scale assessments, their role in the national landscape, and especially in relation to everyday schooling, is very limited. It is therefore worth reconsidering what we can say about the politics of quality in education.

Conclusion

The understanding in this article is that the global politics of quality in education is linked to the idea of controlling contingency and politics as mediativity. The attempts to control contingency are manifold. Many QAE measures were introduced after increased contingency in decentralization reforms, and they were attempts to regain control for central administrations. However, the plurality of QAE and the unpredictability of its results have not fulfilled this aim. Instead, the attempts to control contingency with QAE increase the likelihood of “shared and self-reinforcing goal setting,” “authorizing, but diverted governance,” and “destabilizing and reorganized role setting” (Kauko et al., 2018c). The shared QAE agenda is embedded in the networks that involve actors, from the local to the global. This does not imply a causality of agenda and means, but more that an understanding that QAE can solve the question education faces (Kauko et al., 2018c, 183). The authorization QAE affords can help actors in different places specifically as a tool for “inserting general political aspirations into the education system,” but the complexity of implementation usually diverts the effects (Kauko et al., 2018c, 185). When implemented, QAE changes power relations and creates instabilities, as well as supports different actors, as Dahler-Larsen (2012) presents, to the extent that some of its properties as a governance tool can be questioned. In strong top down QAE systems, destabilization usually happens for the subjects of the measurement (Kauko et al., 2018c, 187).

QAE policies have become unavoidable for countries that aim to be taken seriously internationally. Their use extends to the realm of education, including other goals, and it shapes the power of the actors involved. In this regard, the global politics of quality lie not in the policy itself, but more in the politicization it creates: Global understanding has made QAE unavoidable. Its implementation reformulates education itself and the relations of central education actors. The attempt to understand what the politicization of quality entails requires an understanding of what QAE mediates. International large-scale assessments create different kinds of essentialization. There is an understanding that countries and regions compete as equal entities on the global scale—a feature criticized as methodological nationalism in comparative education. Another mediated understanding is that QAE can increase the quality of education. At its very core, this is a circular deduction if we agree that QAE defines the quality of education. This naked understanding of what quality is tears down the facade that QAE builds for quality: There is no other quality than what we define it to be. Quality of education is fictional and only conforms with our political definition. In this regard, the most political, and too seldom asked, question is: How does this definition happen?

References

- Apple, M.W., 1980. *Ideology and Curriculum*. Routledge, London.
- Azevedo, J.M.L., Santos, A.L.F., 2012. Influências do poder central no planejamento da educação dos municípios da região metropolitana do Recife. *Educ. Soc.* 33 (119), 551–573.
- Carvalho, L.M., Costa, E., Sant’Ovaia, C., 2020. Depicting the faces of results-oriented regulatory processes in Portugal: national testing in policy texts. *Eur. Educ. Res. J.* 19 (2), 125–141. <https://doi.org/10.1177/1474904119858799>.
- Centeno, V.G., Kauko, J., Candido, H.H.D., 2018. Quality assurance and evaluation through Brazilian lenses: an exploration into the validity of umbrella concepts. *Comp. Educ.* 54 (2), 132–158. <https://doi.org/10.1080/03050068.2017.1348084>.
- Centeno, V.G., 2010. Agency in establishing a transnational space: Brazilian interest in US-American Educational Assistance Programs (1956–1974). In: Engel, U., Middell, M. (Eds.), *World Orders Revisited*. Leipziger Universitätsverlag, Leipzig, pp. 155–176.
- Dahler-Larsen, P., 2012. Constitutive effects as a social accomplishment: a qualitative study of the political in testing. *Educ. Inq.* 3 (2), 171–186.
- Dahler-Larsen, P., 2019. *Quality: from plato to performance*. Springer Nature/Palgrave Macmillan, Cham.
- Dovemark, M., Kosunen, S., Kauko, J., Magnúsdóttir, B., Hansen, P., Rasmussen, P., 2018. Deregulation, privatization and marketization of Nordic comprehensive education: social changes reflected in schooling. *Educ. Inq.* 9 (1), 122–141.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison*. Allen Lane, London.
- Gurova, G., Piattoeva, N., Takala, T., 2015. Quality of education and its evaluation: an analysis of the Russian academic discussion. *Eur. Educ.* 47 (4), 346–364.
- Gurova, G., Candido, H., Zhou, X., 2018. Effects of quality assurance and evaluation on schools’ room of action. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 137–160.
- Kallo, J., Takala, T., Centeno, V.G., Suominen, O., 2018. Established and emerging actors in the national political arenas. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 161–179.
- Kamens, D., McNeely, C., 2010. Globalization and the growth of international educational testing and national assessment. *Comp. Educ. Rev.* 54 (1), 5–25. <https://doi.org/10.1086/648471>.
- Kauko, J., Centeno, V.G., Candido, H., Shiroma, E., Klutas, A., 2016. The emergence of quality assessment in Brazilian basic education. *Eur. Educ. Res. J.* 15 (5), 558–579.
- Kauko, J., Takala, T., Rinne, R., 2018a. Comparing politics of quality in education. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 1–17.
- Kauko, J., Suominen, O., Centeno, V.G., Piattoeva, N., Takala, T., 2018b. Established and emerging actors in the national political arenas. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 71–90.
- Kauko, J., Rinne, R., Takala, T., 2018c. Conclusion. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 180–190.
- Kauko, J., Varjo, J., Pitkänen, H., 2020. Quality and evaluation in Finnish schools. In: Pink, W. (Ed.), *Oxford Encyclopedia of School Reform*. Oxford University Press, New York. <https://doi.org/10.1093/acrefore/9780190264093.013.1451>.
- Kauko, J., 2014. Complexity in higher education politics: bifurcations, choices and irreversibility. *Stud. High Educ.* 39 (9), 1683–1699.
- Meade, B., Gershberg, A.I., 2008. Making education reform work for the poor: accountability and decentralization in Latin America. *J. Educ. Pol.* 23 (3), 299–322.
- Meyer, J., Ramirez, F., Soysal, Y., 1992. World expansion of mass education, 1870–1980. *Sociol. Educ.* 65 (2), 128–149. <https://doi.org/10.2307/2112679>.
- Meyer, J.W., 1986. Types of explanation in the sociology of education. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, Westport, CT, pp. 341–359.
- Ozga, J., Dahler-Larsen, P., Segerholm, C., Simola, H., 2011. *Fabricating Quality in Education: Data and Governance in Europe*. Routledge, London.
- Palonen, K., 1993. Introduction: from policy and polity to politicking and politicization. In: Palonen, K., Parvikko, T. (Eds.), *Reading the Political: Exploring the Margins of Politics*. Finnish Political Science Association, Helsinki, pp. 6–16.

- Palonen, K., 2003. Four times of politics: policy, polity, politicking, and politicization. *Alternatives* 28, 171–186.
- Palonen, K., 2006. *The Struggle with Time: A Conceptual History of "Politics" as an Activity*. Verlag Münster, Hamburg.
- Piattoeva, Nelli, Centeno Vera, G., 2018. Suominen, Olli and Rinne, Risto Governance by data circulation? The production, availability, and use of national large-scale assessments data. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study on Brazil, China, and Russia*. Routledge, pp. 115–136.
- Rambla, X., 2012. "Soft power", educational governance and political consensus in Brazil. *Int. Stud. Sociol. Educ.* 22 (3), 191–212.
- Rinne, R., Zhou, X., Kauko, J., Normand, R., Medvedeva, A., Santos, I., 2018. Changing expertise and the state. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 91–114.
- Sahlberg, P., 2016. The global educational reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *Handbook of Global Education Policy*. John Wiley and Sons, Incorporated, Somerset, pp. 180–196.
- Seppänen, P., Rinne, R., Kauko, J., Kosunen, S., 2019. The use of PISA results in education policy-making in Finland. In: Waldow, F., Steiner-Khamsi, G. (Eds.), *Understanding PISA's Attractiveness: Analysis in Comparative Policy Studies*. Bloomsbury, London, pp. 137–160.
- Suominen, O., Kallio, J., Rinne, R., Fan, Y., 2017. Subtle convergences: locating similarities between Chinese educational reforms and global quality assurance and evaluation trends. *Qual. Assur. Educ.* 25 (2), 146–160.
- Suominen, O., Centeno, V.G., Gurova, G., Kallio, J., Zhou, X., 2018. Historical paths to shared interest in quality assurance and evaluation. In: Kauko, J., Rinne, R., Takala, T. (Eds.), *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia*. Routledge, London, pp. 44–70.
- Varjo, J., Simola, H., Rinne, R., 2016. Arvioida ja hallita—perään katsomisesta informaatio-ohjaukseen suomalaisessa koulupolitiikassa [To evaluate and govern—from "looking after" to management by data in Finnish education politics]. *Educ. Res.* 70.
- Verger, A., Parcerisa, L., Fontdevila, C., 2018. The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms. *Educ. Rev.* 71 (1), 5–30.
- Zhou, X., 2019. *Changed and unchanged: the transformation of educational policies on assessment and evaluation in China*. University of Turku, Turku.

Revisiting teacher evaluation a decade after reforms

Joshua Bleiberg, University of Pittsburgh, Pittsburgh, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	54
Research question	54
Methods	55
Mechanisms	55
History of state adoption	55
Political dynamics causing state teacher evaluation reform	56
Downward vertical diffusion	56
Upward vertical diffusion	56
Intermediary organizations	57
Internal determinants	57
Implementation of state reforms	57
Evaluation system bias	58
Teacher labor market outcomes	58
Achievement effects outcomes	58
Discussion	59
References	59

Introduction

In 2009, Washington D.C. updated their teacher evaluation system, which preceded a decade long period of reform by states through the country. For the first time schools had collected years of student achievement data allowing for the calculation of Value-Added Measures (VAM). States also invested considerable resources in developing new observation protocols. Teacher evaluation reform also included policies intended to hold teachers accountable and offered incentives tied to their evaluation rating. Incentivized by \$4.35 billion in federal funding from the Race to the Top (RTTT) initiative and federal waivers from NCLB accountability sanctions, 46 states implemented major reforms to their teacher evaluation systems between 2009 and 2016 (Steinberg and Donaldson, 2016). Nationwide schools paid billions of dollars to increase the rigor and scope of teacher performance reviews and build data systems to estimate teacher VAM (Chambers et al., 2013). The reforms led to conflicts between education reformers and teacher unions resulting in prolonged litigation (Paige, 2020).

Researchers have studied teacher evaluation systems since the passage of DC Impact in 2009. Studies have focused on the political dynamics that caused the rapid and ubiquitous spread of state evaluation reforms (Bleiberg and Harbatkin, 2020; Howell and Magazinnik, 2017). Additionally, researchers have examined the degree to which districts implemented state level reforms (Government Accountability Office, 2015; Kraft and Gilmour, 2017; NCTQ, 2019). Studies have also developed theories about how teacher evaluation systems could influence students and teachers (Goldhaber and Hansen, 2010; Liebowitz, 2021; Loeb et al., 2015; Rothstein, 2015). Scholars have also examined the degree of bias in teacher evaluation systems and the degree to which evaluation reforms promoted more equitable education systems (Campbell, 2020; Chi, 2021). Finally quasi-experimental and experimental studies have examined the effects of teacher evaluation on student outcomes and teacher labor markets (Adnot et al., 2016; Cullen et al., 2021; Dee and Wyckoff, 2015; Kraft et al., 2020; Sartain and Steinberg, 2016).

In this literature review, I look back over the past decade of research to better understand why teacher evaluation was pursued and its influence on education systems throughout the whole country. A considerable portion of the research on teacher evaluation reform was conducted in specific districts or states. I synthesize results across from a variety of local contexts and disciplines. Summarizing patterns across diverse literatures will help to produce nationally generalizable and policy relevant conclusions about teacher evaluation.

Research question

I ask several questions about teacher evaluation systems:

1. What was the theory of action underlying teacher evaluation?
2. What is known about the causes of teacher evaluation adoption and the fidelity of implementation?
3. What was the effect of teacher evaluation on students and teachers?

In the subsequent sections I describe the procedure used to review research on teacher evaluation. I then discuss the theory of change motivating teacher evaluation reform. Subsequently, I summarize the political forces that led states to pursue teacher evaluation reforms and the evidence for whether the reforms took root. I then summarize the empirical literature on how teacher evaluation influenced student outcomes and teachers. I conclude by discussing the implications for future reforms to teacher evaluation systems.

Methods

I began by conducting a systemic review of research studies that were published in peer reviewed journals (Petticrew and Roberts, 2005). I collected articles using the Education Resources Information Center (2022) descriptor “teacher evaluation”. I then conducted a supplementary search for “teacher evaluation” using Google Scholar to ensure that all relevant journal and “gray literature” research was identified. The first step of the search yielded 2388 peer reviewed articles. After screening the abstracts for relevance, 192 articles remained. After a second more detailed examination of the reading and methods the remaining 86 studies were synthesized. I then grouped the findings into categories based on the research questions and summarized the results.

Mechanisms

Researchers have identified two main channels through which teacher evaluation reforms may influence the effectiveness of teachers and subsequently student outcomes. First, teacher evaluation may alter the composition of the education workforce by using incentives to attract candidates and exiting teachers with low ratings. Evaluation ratings were used for a variety of purposes in high stake systems. School leaders could use evaluation ratings to determine who to hire, layoff in case of a reduction of force, grant tenure, promote, and dismiss (Dee and Wyckoff, 2015; Loeb et al., 2015; Neumerski et al., 2018). Evaluation systems also offered incentives to effective teachers. The combination of financial incentives and high-stake accountability was theorized to keep high quality teachers in the classroom and induce the exit of low-quality teachers (Goldhaber and Hansen, 2010; Staiger and Rockoff, 2010). A study in New York City Public Schools found providing principals with value-added scores can improve teacher quality overall by leading to the exit of ineffective teachers and retention of effective teachers (Rockoff et al., 2012). The evidence from the VAMs changed how some principals appraised the quality of teacher’s instruction.

Teacher evaluation will improve teacher quality overall if it causes the exit of ineffective teachers and leads to the recruitment and retention of effective teachers. Simulation studies have found that the conditions under which high stakes teacher evaluation will improve teacher quality are infeasible in many contexts (Liebowitz, 2021; Rothstein, 2015; Sartain and Steinberg, 2021). High-stakes systems that tied bonuses and granting tenure to evaluation ratings only improve teacher quality if paired with very high salaries and excess teacher supply (Rothstein, 2015). Simulation studies have also found that high stakes evaluation systems with instructional supports for all teachers are less effective than other policy designs (Liebowitz, 2021). The optimal evaluation reform would focus accountability policies on only low performing teachers while offering instructional supports to all non-low performing teachers. Teacher labor supply would exceed the quality of ineffective teachers in only a portion of larger urban metropolitan areas (Sartain and Steinberg, 2021; Stecher et al., 2018). Overall simulation studies make important assumptions about the composition of the local labor market that are unlikely to hold in many contexts (e.g., outside of large cities).

Second, teacher evaluation may improve the performance of current teachers by improving instructional practice. More specifically evaluation systems can provide teachers with the resources they need to improve their practice through data, coaching, or other professional development (Donaldson and Firestone, 2021; Galey-Horn and Woulfin, 2021; Mintrop and Trujillo, 2007; Woulfin and Rigby, 2017). States invested considerable resources in training evaluators to provide high quality feedback (Kraft and Christian, 2022). Providing detailed and teacher specific feedback was theorized to be a powerful engine for instructional improvement (Hunter and Springer, 2022).

The high-stakes accountability provisions of teacher evaluation systems (e.g., dismissal, granting tenure, merit pay) could also improve teacher quality by improving extrinsic motivation (Firestone, 2014). Teacher evaluation systems may include a variety of high-quality data including student outcomes, observation protocols, and student surveys. Systems that use all three measures of teacher quality produce the most reliable evaluation ratings (Doan et al., 2019; Kane et al., 2013). That implies that teacher evaluation systems do contain information that is useful for improving instructional practice.

History of state adoption

Beginning in 2010, RTTT offered states that were struggling during the Great Recession large grants in return for adopting education reforms including teacher evaluation. (Bleiberg and Harbatkin, 2020; Howell and Magazinnik, 2017; McGuinn, 2012). Concurrent with RTTT, the Bill and Melinda Gates Foundation invested billions to encourage states to pursue teacher evaluation (Reckhow et al., 2021). The implementation of teacher evaluations systems started to increase in 2010 and spiked in 2012 and 2013. Between 2011 and 2016, the vast majority of states adopted teacher evaluation reforms (Steinberg and Donaldson, 2016). Beginning in 2012, states seeking a waiver from No Child Left Behind (NCLB) requirements to reach 100% proficiency by 2014 were required to adopt teacher evaluation reform. By 2016, 44 states had passed major teacher evaluation reform legislation

(NCTQ, 2016)—shifting away from evaluation systems that relied primarily on observations and were not tied to consequences for teachers (Steinberg and Donaldson, 2016; Weisberg et al., 2009). For states that adopted, implementation continued largely through the 2017 school year, after which states submitted their Every Student Succeeds Act (ESSA) plans. Under ESSA, states recovered more autonomy in setting education policy and many amended their evaluation systems (Ross and Walsh, 2019). School districts also pursued teacher evaluation reforms independently of state efforts. Many school districts made changes to their systems in advanced of state reforms.

Political dynamics causing state teacher evaluation reform

A complex set of intergovernmental dynamics led states to adopt teacher evaluation reforms. The political forces that drove adoption of teacher evaluation systems include: downward vertical diffusion, upward vertical diffusion, intermediary organizations, and internal determinants. Research has examined the role that each dimension played (Bleiberg and Harbatkin, 2020; Howell and Magazinnik, 2017, 2020; McGuinn, 2016; Wong, 2015).

Downward vertical diffusion

Downward vertical diffusion describes a “top-down” approach where the federal government compels the adoption of a policy. The Constitution limits the ability of the federal government to directly proscribe state education policy. To circumvent that constraint, the federal government requires states to pursue specific education reforms in return for accepting funding. Vertical diffusion is a more potent force during recessions, when states want to raise funds without raising taxes. The housing market crash during the Great Recession caused a corresponding decrease in state education funding that relied on property taxes (McGuinn, 2016). This unique set of political circumstances temporarily decreased congressional gridlock and led to the passage of the American Recovery and Reinvestment Act. The legislation prioritized quick action by the Obama administration and included billions in education via the RTTT competitive grant program. RTTT’s competitive grant making process is an example of “executive federalism”, which describes when the President strategically employs the powers of the office to encourage state reforms (Gais and Fosssett, 2005; Howell and Magazinnik, 2017). Executive federalism employs two complementary approaches to encourage state policy adoption. Voluntary compliance is essential because the executive lacks the power to compel action. Simultaneously, the federal government can install inflexible goals for the types of policies that states may put into place. In their pursuit of teacher evaluation reform, the Obama administration followed a strategy that was “tight” on goals but “loose” in how states could demonstrate compliance.

RTTT was a highly effective example of executive federalism. RTTT caused states to adopt teacher evaluation through exploiting fiscal pressure (Howell and Magazinnik, 2017). RTTT also indirectly raised the salience of teacher evaluation for the media and state legislators. The effect of the RTTT on adoption of teacher evaluation policies is clearly observed from the times when federal and state policy occurred. RTTT was announced in 2009 and the first round of competitive grants was awarded in March 2010, which was followed by round 2 in August 2010. 9 out of 10 RTTT grant winners from rounds 1 and 2 adopted teacher evaluation reform. After the first two rounds states (i.e., Alaska and North Dakota) that were not competing for RTTT grants began pursuing teacher evaluation (Meredith, 2013). Two states adopted despite not applying for RTTT in either round. California and Maine repealed laws that prohibited using student data for teacher evaluations (GAO, 2011). The Obama administration announced the NCLB waivers in 2012. To receive a waiver, states were required to reform their teacher evaluation systems. The waivers were a relatively less effective tool for compelling states to adopt teacher evaluation (Wong, 2015). In part because the waivers did not come with the financial incentives offered by RTTT. Additionally, most states that were willing to adopt teacher evaluation had already pursued that reform by 2012.

Upward vertical diffusion

Upward policy diffusion occurs if states adopted teacher evaluation in reaction to district implementation. State leaders monitor local politics closely to help assess support for potential future action (Kingdon, 2002). In the case of teacher evaluation, the implementation of district systems provides valuable data to assess the viability and support for state wide teacher evaluation reforms. Additionally, after district implementation, policy entrepreneurs will engage in “venue shopping” or engaging in efforts scale up a reform (Mintrom, 1997; Shipan and Volden, 2006). Teacher evaluation advocates build their internal policy networks at the state level and then engaged local political groups to strengthen their coalition (Mintrom and Vergari, 1998). Legislators engage in a similar process by borrowing political support for their own policy objective (Manna, 2006).

The effect of district implementation of teacher evaluation on state reforms appears to be relatively weaker than vertical diffusion down (Bleiberg and Harbatkin, 2020). In seven states where large districts implemented teacher evaluation there was no subsequent policy by the state. However, large districts implemented prior to state adopt of reforms in five states (i.e., Connecticut, Georgia, North Carolina, Pennsylvania, Texas) and districts adopted the same year as the state in Colorado and Florida. Bleiberg and Harbatkin (2020) find that the odds of teacher evaluation adoption were 6 times higher in states in which a large district implemented. In many cases early state adoption preceded district implementation, making it impossible to know whether vertical diffusion could have played a role.

Intermediary organizations

Foundations and the non-profits they fund played a key role in leading states to adopt teacher evaluation. Donations to education focused foundations have increased substantially over the last quarter century (Reckhow and Snyder, 2014). Foundations built strategic partnerships with districts and states that had interest in teacher evaluation prior to federal reform efforts (Reckhow, 2012). Education foundations leverage the fundraising networks that exist between foundations and other non-profits. Foundations frequently give to the same organizations, which multiplies the effect of education reform efforts (Reckhow and Snyder, 2014).

The contributions from the Gates Foundation were particularly impactful because they came before federal reform efforts. In 2009 the Gates Foundation provided funding to 5 states and 6 districts that would each subsequently reform their evaluation systems (Bleiberg and Harbatkin, 2020). The Gates Foundation appears to have caused states to adopt teacher evaluation both directly through their grants and providing technical support (Meredith, 2013). The Gates Foundation also provided grants to districts that implemented teacher evaluation systems and ultimately resulted in upward diffusion.

Internal determinants

Prior research on the innovation and diffusion of education reforms has shown that internal determinants can drive states to pursue state education policies. For example, a weak labor supply could have caused states to pursue teacher evaluation (Berry, 1994; Berry and Berry, 1990; Doyle, 2006). State politics can also play an important role. More specifically a gubernatorial candidate may choose to focus on education issues (Henig, 2013). Alternatively, teacher evaluation reforms may fail due to political organizing by the teacher's union (Moe, 2011). About one in three states reported challenges related to maintaining teachers' union support for teacher evaluation reforms (Government Accountability Office, 2015). States were more likely to adopt teacher evaluation if the Great Recession caused a relatively modest decline in state education spending and were more responsive to RTTT if they had previously invested in education reform (Howell and Magazinnik, 2020). However, Bleiberg and Harbatkin (2020) found no economic or political measures that were correlated with state teacher evaluation adoption. In the context of teacher evaluation, it appears that the intergovernmental political dynamics were the more salient factor.

Implementation of state reforms

Teacher evaluation system design and implementation varied considerably across and within states. The vast majority of states provided local districts with some degree of autonomy in implementing teacher evaluation, either by allowing local discretion within a state-designed system or allowing districts to develop their own systems given a set of guidelines (Steinberg and Donaldson, 2016). As a result, the implementation of teacher evaluation was uneven both across and within states.

The federal government promoted a specific model of teacher evaluation through RTTT and the waivers. However, states chose to adopt some evaluation components more frequently than others. Most states that won RTTT grants chose to (1) require student achievement growth, (2) use multicategory rating systems, (3) require annual evaluations, (4) mandate that evaluations were used for professional development, and (5) to use evaluations as a component of dismissal decision (Dragoset et al., 2016). Far fewer states chose to use evaluations to determine bonuses and career advancement. Many also did not include student surveys that were found to improve the reliability of evaluation systems (Kane et al., 2013).

Most states provided school districts with discretion over the implementation of teacher evaluation systems (Steinberg and Donaldson, 2016). States frequently required districts to design their own systems that met state guidelines. For example, some states required the use of student data but were agnostic to whether districts used very different measures (i.e., Value-Added Measures vs. Student-Learning Objectives). School district documents from the National Council for Teacher Quality (NCTQ) teacher contract database provide the clearest view of the evaluation systems that were implemented by school districts (NCTQ, 2022). The database includes Collective Bargaining Agreements, teacher evaluation handbooks, and other district materials from about 150 school districts including the 100 largest districts in the country, the largest district in each state, and member districts of the Council of Great City Schools. The vast majority of school districts have policies in place requiring teacher evaluation. About nine in ten required annual evaluations for non-tenured teachers and about one in two states required annual evaluations for tenured teachers. Similarly, the overwhelming majority (about nine in ten) have procedural rules in place that require districts to share evaluation data either through written feedback or a conference. Far fewer states require teacher evaluations systems to be used for accountability purposes, which is reflected in district policies. About three out of four school districts required student achievement to be included as a component in teacher evaluation systems. About six in ten school districts required evaluations to be used as a criterion for dismissal and four in ten offered bonuses for strong evaluations.

The varied state approaches to system design and voluntary implementation by districts created a national teacher evaluation policy patchwork quilt. Only one in five states installed a unitary teacher evaluation system that districts were required to use. Many states made small changes to their evaluation systems and others allowed districts tremendous flexibility with system design (Bleiberg et al., 2021). As a result, teacher evaluation systems that were intended to be revolutionary in design were quite similar to status quo systems. Teacher evaluation reforms were supposed to identify greater numbers of ineffective teachers and incentivize their exits (Weisberg et al., 2009). However, less than one percent of teachers received unsatisfactory rating in states that reformed their evaluation systems and educators were rarely dismissed for receiving low ratings (Kraft and Gilmour, 2017). Large financial

incentives were theorized to be a powerful force to attract and retain teachers. Most states did not offer bonuses and the financial incentives in states that did were far too small to change behavior (NCTQ, 2019). In Michigan many administrators were unwilling to implement teacher evaluation despite state regulations (Pogodzinski et al., 2015). Michigan required school districts to use student data when making high-stakes personnel decisions. However, many school districts implemented systems that deviated considerably from the mandated policy.

A more practical challenge for teacher evaluation systems are the weaknesses of VAMs used by many evaluation systems. VAM's can be an important component of a system that identifies struggling teachers on average for a whole state (Doan et al., 2019; Kane et al., 2013). But it may still be inaccurate or imprecise for determining whether a specific teacher's performance is unsatisfactory (Ballou and Springer, 2015). Student assignment to teachers is non-random, which likely biases VAM estimates. Teachers can also leverage social networks to choose smaller classes with fewer students that can bias VAM estimates of whether a teacher is effective (Guarino et al., 2014).

Relatedly designing systems that provided teachers with high quality and actionable feedback also proved to be challenging. Training programs and observation protocols that are designed to facilitate instructional improvement for teachers were found to improve student outcomes (Kraft and Christian, 2022; Mihaly et al., 2018). Evaluation ratings do not appear to change their professional improvement activities (Koedel et al., 2019). In Tennessee teacher a substantively large proportion of teachers received feedback that was unaligned with research-based best practices (Hunter and Springer, 2022).

Evaluation system bias

Evidence from diverse contexts and methodological approaches shows that teacher evaluation systems are biased against Black teachers. The effectiveness of Black teachers is exceptionally strong as proven by their capacity to improve student achievement, social-emotional learning, and attendance for students overall and especially for Black students (Beauboeuf-Lafontant, 2002; Campbell, 2021; Cherng and Halpin, 2016; Dee, 2004; Dunbar, 2008; Egalite and Kisida, 2018; Gershenson et al., 2021; Lindsay and Hart, 2017). Administrators gave systematically lower observation ratings to Black and White teachers that had similar value-added measures (Campbell, 2020; Chi, 2021). Teacher evaluation systems like DC IMPACT that swiftly exited teachers with low performance ratings were "harsh" and inequitably affected non-White teachers (Campbell et al., 2021). "Race neutral" evaluation systems may encourage principals suppress efforts to incorporate culturally responsive pedagogy (Donaldson, 2020).

Unfortunately, retention of early career Black teachers is lower than their peers accounting in part for there being too few Black teachers (Gershenson et al., 2021; Sun, 2018). The shortage of Black teachers is likely caused in part by evaluation systems that are biased against Black teachers (Campbell, 2020, 2021; Chi, 2021). Teacher evaluation increased the proportion of schools with unfilled teacher positions and the effects were larger in schools that served Black and Hispanic/Latino students (Kraft et al., 2020). Black teachers are systematically more likely to work in schools with relatively larger proportions of Black and Hispanic/Latino students (Boyd et al., 2011), which implies that teacher evaluation increased attrition for Black students.

Teacher labor market outcomes

Teacher evaluation was hypothesized to improve teacher quality by changing the composition of who teaches. Research from Chicago, Washington D.C., and Houston suggests that teacher evaluation increased the likelihood that the lowest performing teachers voluntarily exited the district (Adnot et al., 2016; Cullen et al., 2021; Dee and Wyckoff, 2015; Sartain and Steinberg, 2016). Evaluation systems also increased the quality of teacher candidates (Kraft et al., 2020). However, research from Pittsburgh, Memphis, and Tampa finds that teacher evaluation had no effect on teacher quality (Stecher et al., 2018).

The inclusion of accountability provisions in teacher evaluation systems likely explains why the effect of evaluation varied across district contexts. In New York City, the teacher evaluation policy required teachers to demonstrate effectiveness through their evaluation to earn tenure. Teachers who did not earn tenure due to low evaluation ratings were more likely to leave their positions (Loeb et al., 2015). Washington D.C.'s evaluation system known as DC IMPACT dismissed teachers who received multiple ineffective ratings. DC IMPACT paired high-stakes personnel decisions with a policy that offered very large financial incentives to the most effective teachers. Taken together the research suggests that receiving a low teacher evaluation rating increases the likelihood that a teacher will exit when combined with other accountability provisions (Dee and Wyckoff, 2015).

Achievement effects outcomes

Rigorous quasi-experimental studies have shown that teacher evaluation improved student outcomes in some but not all districts. On average throughout the whole country, it appears that teacher evaluation did not have an effect (Bleiberg et al., 2021). Evidence from Cincinnati, Chicago, and Washington D.C., suggests that teacher evaluation can improve student outcomes. Taylor and Tyler (2012) use Cincinnati Public Schools data from 2003 to 2004 through 2009 to 2010 showed that a teacher evaluation reform including classroom observations and formative feedback increased math but not reading test scores.

Steinberg and Sartain (2015) used an experiment to show that a pilot evaluation system in Chicago caused large increases in reading achievement, but substantively insignificant results for math. The evaluation system produced significant improvements in reading achievement and positive but imprecisely estimated effects in math in the first year. No positive effects were found in the second year, likely due to challenges relating to scaling the evaluation system or an inadequate supply of high-quality teachers. Teacher evaluation reforms in Chicago emphasized providing formative feedback to improve instruction. However, a second study of the Chicago teacher evaluation pilot found that a likely explanation for the positive effects was teachers choosing to leave after receiving low ratings (Sartain and Steinberg, 2016). Chicago's pilot evaluation system did not include a provision to dismiss ineffective teachers. However, teachers that received low evaluation ratings were 10% more likely to voluntarily exit Chicago Public Schools.

DC IMPACT included intensive professional development for teachers with large bonuses and dismissal for repeatedly ineffective teachers (Adnot et al., 2016; Dee and Wyckoff, 2015; Dotter et al., 2021; James and Wyckoff, 2020). Dee and Wyckoff (2015) found that DC IMPACT increased student outcomes through changing the composition of the teacher workforce. Student achievement increased when minimally effective teachers voluntarily exited (Adnot et al., 2016). Data from observations shows that teachers spend time preparing for evaluations, which improved their subsequent ratings (Phipps and Wiseman, 2021). The positive effects of DC IMPACT appear to be sustainable after many years of reform and iterative changes to evaluation policy (Dee et al., 2021).

Several studies have also found that teacher evaluation had no effect on student outcomes or in some cases small negative effects. One nationwide study examined the effect of evaluation systems in Pittsburgh, Memphis, Tampa, and four Charter Management Organizations (CMOs) supported by the Bill and Melinda Gates Foundation. Researchers found evidence that teacher evaluation systems were fully implemented including incentives and high-stakes accountability, but standardized test scores remained about the same compared to the period prior to the reform (Stecher et al., 2018). The evaluation systems in these districts attempted to tailor professional learning based on the needs of individual teachers by providing feedback through several observations. Teachers viewed these changes as unsubstantial and noted that most teachers had the same evaluator multiple times. Cullen et al. (2021) found that teacher evaluation reforms in Houston increased turnover for the least effective teachers. However, the resulting improvement in teacher quality was insufficiently large to substantially increase student achievement. A study of teacher policy reform in Michigan that included teacher evaluation found no effect of the reform on student achievement (Anderson et al., 2021). A recent national study of state teacher evaluation reform found no effect on student achievement or attainment (Bleiberg et al., 2021). They estimate null effects with enough precision to rule out positive effects larger than 1.5% of a standard deviation for achievement and 1% point for high school graduation and college enrollment.

Discussion

In the few locations that were able to implement reforms with fidelity there is evidence that teacher evaluation benefitted students. Students can benefit from a DC IMPACT style evaluation system, but only if policymakers wholistically embrace DC's approach to reform. DC IMPACT created historically large financial incentives for teachers. Teachers who received a highly effective rating for two consecutive years were eligible for a base pay increase of between \$10,000 to \$450,000 (Dee and Wyckoff, 2015). Funding for the bonuses was provided by philanthropic foundations and considerable human resource investments were made in the central office to administrate DC IMPACT. DC also committed to using highly qualified peer evaluators to conduct observations. The level of organizational effort to implement and organize DC IMPACT is unique and unlikely to be replicated.

The improbability that other states or districts would pursue DC style evaluation reform is motivated by the history of past efforts. After the Great Recession, cash strapped states faced unprecedented fiscal and political pressure to reform evaluation systems in exchange for RITT dollars. Most state chose systems that differed substantively from DC and did not include changes to compensation or high-stakes accountability. State policymakers calculated that the political and financial costs exceeded the potential benefit. Today in the absence of the political dynamics that led dozens of states to pass evaluation reforms, the likelihood of a renewed investment in improving evaluation systems is diminished.

Reformers hoped that changes to teacher evaluation would improve teacher quality and close achievement gaps while also raising student outcomes overall (Weisberg et al., 2009). Overall, throughout the country it appears teacher evaluation did not improve outcomes or benefit minoritized groups. There are many legitimate uses for the evaluation systems that were created. Rigorous observation protocols and coaching have the potential to help teachers improve. Data on teacher quality can be a valuable for researchers to help improve education systems. Given the lack of substantially large positive effects and the evidence that evaluation systems are biased states should consider removing high-stakes accountability for teachers.

References

- Adnot, M., Dee, T.S., Katz, V., Wyckoff, J., 2016. Teacher turnover, teacher quality, and student achievement in DCPS. *Educ. Eval. Pol. Anal.* 39, 54–76.
- Anderson, K.P., Cowen, J.M., Strunk, K.O., 2021. The impact of teacher labor market reforms on student achievement: evidence from Michigan. *Educ. Finance Policy* 1, 1–43.
- Ballou, D., Springer, M.G., 2015. Using student test scores to measure teacher performance: some problems in the design and implementation of evaluation systems. *Educ. Res.* 44, 77–86.

- Beauboeuf-Lafontant, T., 2002. A womanist experience of caring: understanding the pedagogy of exemplary Black women teachers. *Urban Rev.* 34, 71–86.
- Berry, F.S., 1994. Sizing up state policy innovation research. *Pol. Stud. J.* 22, 442–456.
- Berry, F.S., Berry, W.D., 1990. State lottery adoptions as policy innovations: an event history analysis. *Am. Polit. Sci. Rev.* 84, 395–415.
- Bleiberg, J., Brunner, E., Harbatkin, E., Kraft, M.A., Springer, M., 2021. The Effect of Teacher Evaluation on Achievement and Attainment: Evidence From Statewide Reforms. Annenberg Institute at Brown University. [EdWorkingPapers.com](https://www.edworkingpapers.com).
- Bleiberg, J., Harbatkin, E., 2020. Teacher evaluation reform: a convergence of federal and local forces. *Educ. Pol.* 34, 918–952.
- Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., Wyckoff, J., 2011. The role of teacher quality in retention and hiring: using applications to transfer to uncover preferences of teachers and schools. *J. Pol. Anal. Manag.* 30, 88–110.
- Campbell, S.L., 2020. Ratings in black and white: a quantcrit examination of race and gender in teacher evaluation reform. *Race Ethn. Educ.* 1–19.
- Campbell, S.L., 2021. Shifting teacher evaluation systems to community answerability systems: (re)imagining how we assess black women teachers. In: *The Educational Forum*. Taylor & Francis, pp. 80–92.
- Campbell, C.M., Parker, C., Shand, R., Kelly-Massoud, A., Fashola, T., Blanc, J., 2021. Perspectives on DCPS IMPACT Teacher Evaluation System: Findings From Teachers and School Leaders. American University.
- Chambers, J., Brodzia de los Reyes, I., O'Neil, C., 2013. How Much Are Districts Spending to Implement Teacher Evaluation Systems. RAND, Washington, DC.
- Cherng, H.-Y.S., Halpin, P.F., 2016. The importance of minority teachers: student perceptions of minority versus White teachers. *Educ. Res.* 45, 407–420.
- Chi, O.L., 2021. A classroom observer like me: the effects of race-congruence and gender-congruence between teachers and raters on observation scores. *Educ. Finance Policy* 1–67.
- Cullen, J.B., Koedel, C., Parsons, E., 2021. The compositional effect of rigorous teacher evaluation on workforce quality. *Educ. Finance Policy* 16, 7–41.
- Dee, T.S., 2004. Teachers, race, and student achievement in a randomized experiment. *Rev. Econ. Stat.* 86, 195–210.
- Dee, T., Wyckoff, J., 2015. Incentives, selection, and teacher performance: evidence from IMPACT. *J. Pol. Anal. Manag.* 34, 267–297. <https://doi.org/10.1002/pam.21818>.
- Dee, T.S., James, J., Wyckoff, J., 2021. Is effective teacher evaluation sustainable? Evidence from District of Columbia Public Schools. *Educ. Finance Policy* 16, 313–346.
- Doan, S., Schweig, J.D., Mihaly, K., 2019. The consistency of composite ratings of teacher effectiveness: evidence from New Mexico. *Am. Educ. Res. J.* 56, 2116–2146.
- Donaldson, M., 2020. *Multidisciplinary Perspectives on Teacher Evaluation: Understanding the Research and Theory*. Routledge, New York.
- Donaldson, M., Firestone, W., 2021. Rethinking teacher evaluation using human, social, and material capital. *J. Educ. Change* 22, 1–34.
- Dotter, D., Chaplin, D., Bartlett, M., 2021. Impacts of school reforms in Washington, DC on student achievement. *Mathematica*.
- Doyle, W.R., 2006. Adoption of merit-based student grant programs: an event history analysis. *Educ. Eval. Pol. Anal.* 28, 259–285.
- Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., Boyle, A., Tanenbaum, C., Giffin, J., Upton, R., 2016. *Race to the Top: Implementation and Relationship to Student Outcomes*. National Center for Education Evaluation and Regional Assistance.
- Dunbar, R.B., 2008. *Finding Their Way: A Critical Ethnography of Five of African American Women Educators' Early Experiences to Develop Into Culturally Relevant Pedagogues*. Georgia State University.
- Education Resources Information Center, 2022. <https://eric.ed.gov/>.
- Egalite, A.J., Kisida, B., 2018. The effects of teacher match on students' academic perceptions and attitudes. *Educ. Eval. Pol. Anal.* 40, 59–81.
- Firestone, W.A., 2014. Teacher evaluation policy and conflicting theories of motivation. *Educ. Res.* 43, 100–107.
- Gais, T., Fossett, J., 2005. Federalism and the executive branch. In: *The Executive Branch*, pp. 486–524.
- Galey-Horn, S., Woufflin, S. I., 2021. Muddy waters: the micropolitics of instructional coaches' work in evaluation. *Am. J. Educ.* 127, 441–470.
- GAO, 2011. *Race to the Top: Reform Efforts Are Under Way and Information Sharing Could be Improved*.
- Gershenson, S., Hansen, M., Lindsay, C.A., 2021. *Teacher Diversity and Student Success: Why Racial Representation Matters in the Classroom*. Harvard Education Press, Cambridge, Massachusetts.
- Goldhaber, D., Hansen, M., 2010. Using performance on the job to inform teacher tenure decisions. *Am. Econ. Rev.* 100, 250–255.
- Government Accountability Office, 2015. *Race to the Top: Survey of School Districts' Capacity to Implement Reform* (GAO-15-317SP, April 2015), an E-supplement to GAO-15-295.
- Guarino, C.M., Reckase, M.D., Wooldridge, J.M., 2014. Can value-added measures of teacher performance be trusted? *Educ. Finance Policy* 10, 117–156.
- Henig, J.R., 2013. *The End of Exceptionalism in American Education: The Changing Politics of School Reform*. Harvard Education Press, Cambridge, Massachusetts.
- Howell, W.G., Magazinnik, A., 2017. Presidential prescriptions for state policy: Obama's race to the top initiative. *J. Pol. Anal. Manag.* 36, 502–531.
- Howell, W.G., Magazinnik, A., 2020. Financial incentives in vertical diffusion: the variable effects of Obama's race to the top initiative on state policy making. *State Polit. Pol. Q.* 20, 185–212.
- Hunter, S.B., Springer, M.G., 2022. Critical feedback characteristics, teacher human capital, and early-career teacher performance: a mixed-methods analysis. *Educ. Eval. Pol. Anal.* 44 (4). <https://doi.org/10.3102/01623737211062913>.
- James, J., Wyckoff, J., 2020. Teacher evaluation and teacher turnover in equilibrium: evidence from DC public schools. *AERA Open* 6 (2).
- Kane, T.J., McCaffrey, D.F., Miller, T., Staiger, D.O., 2013. *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment*. Research Paper. MET Project, Bill & Melinda Gates Foundation. Bill & Melinda Gates Foundation.
- Kingdon, J.W., 2002. *Agendas, Alternatives, and Public Policies*, second ed. Pearson, New York.
- Koedel, C., Li, J., Springer, M.G., Tan, L., 2019. Teacher performance ratings and professional improvement. *J. Res. Educ. Eff.* 12, 90–115.
- Kraft, M.A., Brunner, E.J., Dougherty, S.M., Schwegman, D.J., 2020. Teacher accountability reforms and the supply and quality of new teachers. *J. Publ. Econ.* 188, 104212.
- Kraft, M.A., Christian, A., 2022. Can teacher evaluation systems produce high-quality feedback? An administrator training field experiment. *Am. Educ. Res. J.* 59, 500–537.
- Kraft, M.A., Gilmour, A., 2017. Revisiting the widget effect: teacher evaluation reforms and the distribution of teacher effectiveness. *Educ. Res.* 46, 234–249.
- Liebowitz, D.D., 2021. Teacher evaluation for accountability and growth: should policy treat them as complements or substitutes? *Lab. Econ.* 102024.
- Lindsay, C.A., Hart, C.M., 2017. Exposure to same-race teachers and student disciplinary outcomes for Black students in North Carolina. *Educ. Eval. Pol. Anal.* 39, 485–510.
- Loeb, S., Miller, L.C., Wyckoff, J., 2015. Performance screens for school improvement: the case of teacher tenure reform in New York City. *Educ. Res.* 44, 199–212.
- Manna, P., 2006. *School's In: Federalism and the National Education Agenda*. Georgetown University Press.
- McGuinn, P., 2016. From No Child Left Behind to the Every Student Succeeds Act: federalism and the education legacy of the Obama administration. *Publius* 46, 392–415.
- McGuinn, P., 2012. Stimulating reform: race to the top, competitive grants and the Obama education agenda. *Educ. Pol.* 26, 136–159.
- Meredith, J.K., 2013. *State Policy Diffusion and Race to the Top: The Impact of Federal Competitive Grants on State Policymaking*. University of Pittsburgh.
- Mihaly, K., Schwartz, H.L., Oppen, I.M., Grimm, G., Rodriguez, L., Mariano, L.T., 2018. Impact of a Checklist on Principal-Teacher Feedback Conferences Following Classroom Observations (No. 285). Regional Educational Laboratory Southwest, Austin, TX.
- Mintrom, M., 1997. Policy entrepreneurs and the diffusion of innovation. *Am. J. Polit. Sci.* 41, 738–770.
- Mintrom, M., Vergari, S., 1998. Policy networks and innovation diffusion: the case of state education reforms. *J. Polit.* 60, 126–148.
- Mintrop, H., Trujillo, T., 2007. The practical relevance of accountability systems for school improvement: a descriptive analysis of California schools. *Educ. Eval. Pol. Anal.* 29, 319–352.
- Moe, T.M., 2011. *Special Interest: Teachers Unions and America's Public Schools*, reprint ed. Brookings Institution Press, Washington, DC.
- NCTQ, 2016. *State-by-State Evaluation Timeline Briefs*.
- NCTQ, 2019. *Teacher & Principal Evaluation Policy*. State of the States 2019. National Council on Teacher Quality.
- NCTQ, 2022. *Teacher Contract Database [WWW Document]*.

- Neumerski, C.M., Grissom, J.A., Goldring, E., Rubin, M., Cannata, M., Schuermann, P., Drake, T.A., 2018. Restructuring instructional leadership: how multiple-measure teacher evaluation systems are redefining the role of the school principal. *Elem. Sch. J.* 119, 270–297.
- Paige, M., 2020. Moving forward while looking back: lessons learned from teacher evaluation litigation concerning Value-Added Models (VAMs). *Educ. Pol. Anal. Arch.* 28, 64.
- Petticrew, M., Roberts, H., 2005. *Systematic Reviews in the Social Sciences: A Practical Guide*, first ed. Wiley-Blackwell, Malden, MA, Oxford.
- Phipps, A.R., Wiseman, E.A., 2021. Enacting the rubric: teacher improvements in windows of high-stakes observation. *Educ. Finance Policy* 16, 283–312.
- Pogodzinski, B., Umpstead, R., Witt, J., 2015. Teacher evaluation reform implementation and labor relations. *J. Educ. Pol.* 30, 540–561. <https://doi.org/10.1080/02680939.2014.999827>.
- Reckhow, S., 2012. *Follow the Money: How Foundation Dollars Change Public School Politics*, first ed. Oxford University Press, Oxford, England; New York, NY.
- Reckhow, S., Snyder, J.W., 2014. The expanding role of philanthropy in education politics. *Educ. Res.* 43, 186–195.
- Reckhow, S., Tompkins-Stange, M., Galey-Horn, S., 2021. How the political economy of knowledge production shapes education policy: the case of teacher evaluation in federal policy discourse. *Educ. Eval. Pol. Anal.* 43, 472–494.
- Rockoff, J.E., Staiger, D.O., Kane, T.J., Taylor, E.S., 2012. Information and employee evaluation: evidence from a randomized intervention in public schools. *Am. Econ. Rev.* 102, 3184–3213.
- Ross, E., Walsh, K., 2019. *State of the States 2019: Teacher and Principal Evaluation Policy*. National Council on Teacher Quality, Washington, DC.
- Rothstein, J., 2015. Teacher quality policy when supply matters. *Am. Econ. Rev.* 105, 100–130.
- Sartain, L., Steinberg, M.P., 2016. Teachers' labor market responses to performance evaluation reform: experimental evidence from Chicago public schools. *J. Hum. Resour.* 51, 615–655.
- Sartain, L., Steinberg, M.P., 2021. Can Personnel Policy Improve Teacher Quality? The Role of Evaluation and the Impact of Exiting Low-Performing Teachers. Annenberg Institute at Brown University. [EdWorkingPapers.com](https://edworkingpapers.com).
- Shipan, C.R., Volden, C., 2006. Bottom-up Federalism: the diffusion of antismoking policies from US cities to states. *Am. J. Polit. Sci.* 50, 825–843.
- Staiger, D.O., Rockoff, J.E., 2010. Searching for effective teachers with imperfect information. *J. Econ. Perspect.* 24, 97–118.
- Stecher, B.M., Holtzman, D.J., Garet, M.S., Hamilton, L.S., Engberg, J., Steiner, E.D., Robyn, A., Baird, M.D., Gutierrez, I.A., Peet, E.D., Brodziak de los Reyes, I., Fronberg, K., Weinberger, G., Hunter, G.P., Chambers, J., 2018. *Improving Teaching Effectiveness: Final Report: The Intensive Partnerships for Effective Teaching Through 2015–2016*. RAND, Washington, DC.
- Steinberg, M.P., Donaldson, M., 2016. The new educational accountability: understanding the landscape of teacher evaluation in the post-NCLB era. *Educ. Finance Policy* 11, 340–359.
- Steinberg, M.P., Sartain, L., 2015. Does teacher evaluation improve school performance? Experimental evidence from Chicago's excellence in teaching project. *Educ. Finance Policy* 10, 535–572.
- Sun, M., 2018. Black teachers' retention and transfer patterns in North Carolina: how do patterns vary by teacher effectiveness, subject, and school conditions? *AERA Open* 4. <https://doi.org/10.1177/2332858418784914>.
- Taylor, E.S., Tyler, J.H., 2012. The effect of evaluation on teacher performance. *Am. Econ. Rev.* 102, 3628–3651.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., 2009. *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*. Executive Summary. New Teacher Project.
- Wong, K.K., 2015. Federal ESEA waivers as reform leverage: politics and variation in state implementation. *Publius* 45, 405–426.
- Woulfin, S.L., Rigby, J.G., 2017. Coaching for coherence: how instructional coaches lead change in the evaluation era. *Educ. Res.* 46, 323–328.

Teacher policy frames, enactment forms and issues

Beatrice Ávalos and Maria Beatriz Fernandez, Center for Advanced Research in Education, Institute of Education, University of Chile, Santiago, Chile

© 2023 Elsevier Ltd. All rights reserved.

Introduction	62
Teacher-related policy: contemporary issues and discussions	62
The irruption of COVID-19 and challenges to existing policy and practices	63
Teacher policy frames in the 2000s	63
Teacher career policies	66
Teacher transnational policy and international organizations	67
Arguable matters: global conversations about teachers and teacher education	67
Social justice, inclusion and teachers	68
Conclusions	68
Acknowledgment	69
References	69

Introduction

This article, centered on teacher policy frames, its issues and enactment forms, is inevitably limited in coverage. Its focus is mainly on frames considered important for their social and international relevance. Broadly, we understand frames as social constructs that underlay teacher-related policy rather than as the specifically pedagogic ones, although these are embedded in social constructs. More specifically, we use certain areas of teaching and teacher education to illustrate how they respond to policy and how policy affects them. Policy formulation is marked by issues reflecting broader theoretical positions about the nature of teaching (professional or technical), its quality control (professional or accountable/measurable), and its regulatory perspectives (market/private, justice/equity). Policy frames surround the immediateness of educational and societal issues such as international/national tests and teaching quality or education for all and teacher provision. They also serve to deal with the unexpected as exemplified in the COVID-19 pandemic and with future projections, building on evidence of the skills and capacities needed.

We have not been able to cover in this article all of these diverse frames in the extension and depth that would be required. However, we have selected overarching frames such as New Public Management (NPM) and Accountability to examine their contemporary impact on teachers and teacher education across different countries. We have examined enactment of such frames in teacher career systems around the world, and also referred to large scale testing (national and international) and to arguments regarding its effect on teacher and teacher education assessment. We have also moved the focus of our analysis to the role of international organizations in broadening policy frames along the lines of teacher preparation for inclusion and the enactment of inclusive classrooms, as well as in mapping the future and teacher preparation in this respect. We have finally highlighted that social justice and equity principles and practices, must find their place in teacher education and professional development in order to render inoperative in classrooms, the “silent exclusion of student identities.”

Teacher-related policy: contemporary issues and discussions

Teacher policy, whether international or local, reinstates traditional themes and attends to new or more specific policy issues. Thus, its formulations cannot but address questions about who should become a teacher, how should teachers be prepared, what criteria is needed to assess teaching quality and who should assess it, as well as how and under what conditions do teachers perform competently. But adequate policy responses to these traditional issues also need to be relevant to changing educational conditions and contexts as well as to the foreseeable future. More specifically, while teacher policies must rest on the longstanding acknowledgment of teaching as a profession ([International Labour Organization, 2016](#); [Hargreaves, 2000](#); [Darling-Hammond and Sykes, 1999](#); [Hoyle, 1974](#)) they should also take on discussions and research evidence on the changing forms and important nuances of teaching. This means that teacher policy needs to be broad enough to allow for divergence in its enactment, relevant to teaching in poor societies and economies with low education levels ([UNESCO, 2014](#)), as well as adventurous in promoting the development of “21st century skills” in all school populations ([IDB, 2019](#); [Reimers and Chung, 2016](#)).

We begin this section, by taking on a theme which was absent at the moment of planning this article: the irruption of COVID-19 and its challenge to existing policy and practices. We then move on to discuss the nature of teacher policy discussions, frames and teacher research in the 2000s. And finally, we provide an up-dated view of teacher policies from the perspective of their global nature and impact.

The irruption of COVID-19 and challenges to existing policy and practices

In 2020 closure of schools around the world thrust teachers, students and their families, into a re-designing of their teaching and learning modalities, and the developing of acute teacher capacity to gauge students' response, both cognitive and emotional, to the changed forms of schooling. Research and studies had alerted schools and teachers that communications were being revolutionized by technology and that quite different modes of accessing and managing communication were expanding globally (Rivas, 2019). Yet, in his analysis of teacher preparation for demands of COVID-19, Schleicher (2020) concluded that most education systems covered by the OECD's round of PISA 2018 were not ready for digital learning in school situations. Only younger teachers using technology in teaching could be considered prepared. Faced with the pandemic, global institutions such as the United Nations, UNESCO and OECD rapidly produced broad orientations to guide the thinking and acting of policy makers as well the enactment of relevant school practices. Also, in the light of the pandemic's effect over inequality, discrimination and student exclusion in different contexts and situations, a dynamic state of policy development reaffirmed the *raison-d'être* of education as a human right (United Nations, 2020), and its broadening to include the right to "connectivity" and access to knowledge and information via free and open source technologies (UNESCO, 2020a). There were suggestions on how to manage teaching where internet connection lacked or failed and global institutions encouraged the prompt opening of schools with important and needed safeguards in place (OECD, 2020). Also affirmed was the importance of the teaching profession and of teacher collaboration in conditions of autonomy and flexibility (Reimers and Schleicher, 2020; UNESCO, 2020a) as well as trust in teacher capacity to lead schools "forward in the post-pandemic world" (Sahlberg, 2020).

Reports and studies have increased with some examining the opportunity provided for teachers to establish closer connections with parents and develop different and creative ways of teaching cooperatively (Hundred Org, 2020; CIAE et al., 2020). Others have endeavored to learn how teachers manage the new context and its complexities, reaching similar findings about the importance of personal and pedagogic experiences. Different countrywide surveys provided knowledge on how teachers face COVID both personally and professionally. In India, Hassan et al. (2020) reported on teacher problems with bandwidth, internet and support in use of online platforms, also occurring in all less affluent countries. In Sweden, Bergdahl and Nouri (2020) reported on teacher dissatisfaction with their pedagogical preparation in facing the new demands. And Kraft et al. (2020) studied conditions in the United States (USA) enabling teachers to cope, such as good communication systems in their schools, collaboration and principal leadership. The rest of the story is still to be written.

Teacher policy frames in the 2000s

Moving back from COVID-19 we deal here with teacher policy accountability frames and persistence of new public management approaches. These policies are contrasted with those based on human rights' orientations, sustainable development, social justice and equity. As extremes of a continuum these policy approaches rest on different theoretical assumptions and empirical evidence regarding the purposes of education and its processes. In their pure forms they do not mark the major orientation of education systems, but in some they tend to be more pronounced than in others. We examine below the nature of discussions around these orientations.

New public management policies

In the second decade of the 2000s educational policy analysis and discussions around teachers continued to be centered on the impact of neo-liberalism on education. Critical assessments focused on its re-shaping and influence over potential loss of teacher autonomy in Hong Kong (Lee et al., 2019), the marketization of schooling in Australia (Hogan and Thompson, 2019), and on "fast policy" effects in Sweden over teachers as individuals and on democratic and professional values (Hardy et al., 2019b; Bergh et al., 2019). More broadly, in the context of neo-liberal policies, recent policy analysis and teacher research in different education systems have taken up persisting issues related to new public management and teacher performance accountability.

New public management (NPM) policies apply management theories and practices to public service administration, including education. Their effect has been to alter the longstanding view of education as a public good, and to foster policy beliefs that education services are more efficiently and effectively regulated by market forces (Grace, 1994 cited in Tolofari, 2005). In an educational market system, schools are classified according to test results that drive parental school choice, and public funding follows student numbers (Tolofari, 2005). Broadened to include the notion of "performativity", NPM conceives educational services as managed by a:

combination of enforced standards and accountability (e.g., national tests, school inspections), devolution and diversity (different types of schools running core curriculum programs and specializing in some areas of their choice), flexibility of employment (for teaching professionals at different levels), greater choice (wider variety of schools and the consumer having real power to choose) and managerialism.

Peters (2004) in Tolofari (2005, p. 86).

The above description of NPM policies applies to policy frames in several systems around the world, including England after the 1988 Education Reform Act (Gewirtz et al., 1995), the United States and Australia (Dunleavy et al., 2006) as well as in Asia, Africa and Latin America (Kapucu, 2006). Enactment of neo-liberal and NPM policies in Chile from the early 1980s, included the introduction of a voucher system of school funding and the implementation of large-scale school testing (Ávalos and Bellei, 2019). Brazil modified its public administration and education system along managerial concepts with the enactment in 1995 of the Public

Administration Reform Plan, and the opening of greater private participation in the conduction of public policy (da Rosa, 2019). Brazilian reforms also introduced large-scale assessments to monitor school quality and student learning, limiting the taught curriculum to testing demands, performance bonuses and external auditing (da Rosa, 2019).

Recent critical analysis of NPM policies has remarked on its narrow interpretation of the word “evidence”, a concept increasingly used to declare the worthwhileness of educational practices (Helgetun and Menter, 2020). According to this interpretation, evidence is a quantifiable expression of policy application influencing what Mockler and Stacey (2020) call its “datafication”. Such misuse of the word “evidence” overlooks much of what is understood as such in educative acts: substantive knowledge, social interaction, power, and the values and norms that guide or should guide education activities (Biesta, 2010).

In its relation to teachers as professionals, NPM policies have impacted on the occupational value of teacher professionalism (Evetts, 2014). They have done so via the introduction of controls over teaching such as standards and external forms of regulation while maintaining the traditional discourse that extolls teacher identity, authority, work culture and collegial relations, among others.

Teacher accountability policy mutations

To be subject to the supervision of another and not do as one sees fit, is altogether convenient.

Aristotle, *Politics* VI, 4.

Government authorities and education policy-makers have always sought evidence for the impact of education and the good use of public investment in schools, teaching and learning. For most of the 20th century this task was entrusted to school authorities and to external inspectors who observed classrooms and took note of teaching quality and its effects. In the United States (USA) as documented by Jewell (2017) teachers considered to lack competence could be subject to dismissals from the school, transfers from one school to another, changes in teaching subject or in class responsibilities. In England an inspectorate system monitored school and teacher performance (Her Majesty’s Inspectorate) until 1992 when it was replaced by the Office for Standards in Education, Children’s Services and Skills (Ofsted). Inspectorates also existed in European countries (Pons, 2014) charged with monitoring school administrative and educational performance and continue to be in place in many Latin American education systems. However, as we shall see below accountability progressively took on new forms of examining teacher performance and evaluating its quality. We examine below accountability policies as developed and implemented in three English speaking countries, which are selected because their systems have influenced such policies in other geographical contexts.

Teacher accountability in the US, England and Australia

With the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983) policy attention in the USA was directed to furthering teacher professionalism based on the analysis of what is involved in good teaching and the use of relevant generic teaching standards to assess teacher performance (Hunter, 1982; Danielson, 2007). Teacher performance assessment could involve the use of observation tools and appropriate interviews as a responsibility of school authorities or other qualified professionals. However, with the passing of the *No Child Left Behind Act* (2001) there were changes to the education system that affected teacher evaluation. These included the use of “standardized tests aligned with challenging academic achievement standards” (US Department of Education, 2002 cited in Hursh, 2005) and the requirement of more precise measures of teachers’ performance. Teachers were gradually incorporated into a test-based teacher and school accountability system that included Value-Added measures (VAM) to determine teachers’ impact on student learning (Holloway et al., 2017). The later *Federal Race to the Top* initiative (RTTT, 2009) moved the accountability system further by providing financial rewards to states that linked student data to teacher evaluation scores (Jewell, 2017). The next policy instrument, known as *Every Student Succeeds Act* (ESSA, 2015) introduced some moderation to teacher evaluation by moving it away from uniform student national testing toward state assessments, without necessarily requiring the association of teacher performance with student test results. As a whole, these policies resulted in largely different systems of teacher performance assessment and accountability across the USA states (Jewell, 2017).

Teacher quality and its effect on economic earnings add another perspective to teacher accountability discussions. On the basis of research examining student scores on standardized tests and their longtime effects on economic gains (individual and national growth rates) together with teacher value-added analysis, Hanushek (2019) concludes that very effective teachers have a bigger impact over their student future earnings compared to impact of the least effective ones. The same would be true of the aggregate impact of effective teachers over all USA students and their higher incomes and therefore over national economic growth. Solutions based on hiring good graduates from teacher education and/or improving teacher quality through professional development would not be effective given inconclusive evidence cited about the quality of these programs. Therefore, Hanushek (2019) concludes that his evidence establishes a strong case for judging the efficacy of policies such as initial teacher education or professional development. It should be noted though, that Komatsu and Rappleye (2017) contested on statistical grounds part of the evidence cited by Hanushek linking student scores on standardized tests to economic growth. They further contend that linking school results to economic growth may be empirically verifiable, but that it is altogether different from propagating this as a view of education. Setting economic gain as the goal for education may be detrimental to the wider aims of student learning.

The British school system was substantially altered along market principles and structures with the 1988 Education Reform Act passed during Margaret Thatcher's government (Gewirtz et al., 1995). Its key change was to modify the school funding system on the basis of weighted per capita (Acquah, 2013), promoting a context of school league tables and competition (1992). Schools became accountable to the Office for Standards in Education (Ofsted) on the basis of national tests and examination results (Acquah, 2013). By 1995 the English system had introduced ranking by means of Value-Added Measures (VAM), although between 2006 and 2010, the VAM was reconfigured as "contextual value-added", taking into account student prior achievement, eligibility for free school meals and lack of spoken English at home (Acquah, 2013). Teachers, as individuals, in the English system were not subjected to accountability forms similar to those in the USA. They continued to be evaluated by school authorities on the basis of performance standards established by the Department of Education (Department of Education, June 2011). However, strong pressure on teachers to improve school results and position on league tables, raised their efforts to ensure curriculum delivery rather than attend to student understanding, thus encouraging teaching-to-the-test practices (Pring et al. in Acquah, 2013). As reported in Keating's (2015) case study of the English accountability system, teachers' lost autonomy and possibility to innovate in their teaching became aligned with their emphasis on teaching only core subjects examined.

In Australia, the National Assessment Program-Literacy and Numeracy school testing system (NAPLAN) in place since 2008 affected curriculum approaches and forms of teaching, especially in states where school results did not correspond to expectations (Hardy et al., 2019a). Concerned with increased "datafication" in the Australian system Mockler and Stacey (2020), interviewed 21 teachers recruited via social media to learn about their disposition to use quantified data and/or research as sources for valid and reliable evidence of practice, and also surveyed a sample of 524 teachers on these issues. They found that the teachers studied were less enthusiastic about working around data indicators, although disposed to work with research evidence as they were to do so with data.

Teacher performance assessment and accountability demands in other countries have followed to some degree the USA path, though stopping short of using standardized test results as part of teacher performance assessment. Seeking to unveil what kind of systems and national contexts use performance-based pay in their teacher evaluation systems, Aoki and Rawat (2020) examined 2012 PISA teacher data, distinguishing between more and less liberal types of government, and other country-related variables. The researchers found that performance-based teacher pay tended to be used more in less liberal countries as compared to more liberal ones, probably as a way of controlling their public servants (Aoki and Rawat, 2020). There also is evidence that performance-based pay is used in countries with higher commitment to market ideology, and in countries and cultures with varying commitment to reduced state role (Rivas and Sánchez, 2020; Tsz-lok Lee et al., 2019; Appel, 2020).

Accountability in teacher education: holistic or practice oriented

Similar to the narrowing of what is considered as competent teacher capacity, teacher education has been under pressure to account for a narrower range of future teacher competencies than those that were part of standards formulations in the 1990s and early 2000s years. This is the case in different USA states as also in England, with policies that give greater importance to developing specific teaching competencies (lesson preparation, instruction and assessment) and lesser to more broad professional ones. Cochran-Smith et al. (2018) attribute this policy emphasis in the USA to a powerful narrative, that presents teacher education programs as uninterested in selecting good candidates for teaching and that prepare them inadequately for teaching and classroom management (Zeichner and Peña Sandoval, 2015). To deal with these "shortcomings" of teacher education, policies promote the use of strong accountability instruments able to hand-out hard data-driven evidence about the quality of teacher education. This emphasis is noticeable in two USA institutions: The Council for the Accreditation of Educator Preparation (CAEP) and The National Council on Teacher Quality (NCTQ). The CAEP conducts teacher education accreditation in states that require it, while the NCTQ is a private advocacy organization centered on holding public teacher education accountable to the public. Both institutions, despite their differences, share the notion that quality improvement is driven by data (Payne, 2019). Critical analysis by Cochran-Smith et al. (2018) concludes that CAEP is a weak institution in that teacher education accreditation is not required in all USA states, observing also that its accreditation standards and procedures lack "complete support of the profession" (p. 76). As for the NCTQ, Cochran-Smith et al. (2018) point to its market-oriented logic evident in its public displaying of program evaluations based on inputs (syllabi and textbooks), thus expecting to influence program quality on the basis of consumer choice and shaming. To overcome these forms of accountability approaches, Cochran-Smith et al. (2018) argue in favor of what they label as democratic accountability for teacher education. This would require fostering internal accountability of teacher education programs oriented to achieve democratic goals, enhance professionalism, and promote broader notions of teaching and learning, as well as democratic participation and trust.

In England and Wales, university-based teacher education eroded from the late 1980s onwards, due to policy indictment of its quality and the view that schools might do better in training teachers. Thus, currently, acquiring a Qualified Teacher Status in England requires the completion of an undergraduate degree and post-graduate teacher preparation in one of three types of institutions: universities, school consortia or a school-based salaried route into teaching (Meyer and Mills, 2020; Prospects, 2020). In addition, the current guidelines for teacher preparation (DFE, 2019) have been criticized for their narrow emphasis on "training", low expectations about teaching as a profession and insufficient research base (Turvey et al., 2019 in Mayer and Mills, 2020). In contrast to these English teacher education policies, Scotland and Northern Ireland maintain their university conducted Post-Graduate Degree in Education (PGCE) as a requisite for teaching.

In Australia, all teacher education is university-based, although it can be pursued through a number of pathways. However, as discussed by [Mayer and Mills \(2020\)](#), teacher education accountability was tightened in the light of recommendations from the TEMAG 2014 Report (Teacher Education Ministerial Advisory Group) and other documents dealing with accreditation and standards for initial teacher education. These documents have constructed the “right” sort of teacher to be prepared as somebody “who can demonstrate appropriate teaching practice against standards within a system that determines performance indicators and mechanisms for classroom readiness” ([Mayer and Mills, 2020](#), p. 8).

These narrower conceptions of teacher preparation centered more on competence in what is required of classroom teaching or “classroom readiness” but less so on the broader professional capacities of teachers are recognizable in a new instrument originated in the USA as the Capstone Teacher Performance Assessment (EdTPA) and recently adapted for use in Australia (TPA). The material of the USA assessment which is administered by Pearson, a private corporation, consists of a summative gathering by future teachers, of portfolio teaching evidence (assessments, teaching plans, unedited video recordings of teaching, etc.) which are linked to a planned sequence of lessons ([Charteris, 2019](#)). The use of EdTPA in the USA and Australia has been the subject of strong controversy, ([Cochrane-Smith et al., 2018](#); [Charteris, 2019](#)). While acknowledging that teachers are valuable and capable professionals who can be evaluated formatively, the EdTPA also embraces elements of compliance-based accountability and neo-liberal ideology ([Cochrane-Smith et al. \(2018\)](#)). More directly [Charteris \(2019\)](#) views its Australian version as an instrument designed to serve accountability purposes and competing with many other more appropriate measures used in Australia to assess new teacher preparedness (2019). A comprehensive study of peer reviewed articles over ten years ([Stacey et al., 2019](#)) about EdTPA in the USA and its predecessor the Performance Assessment for California Teachers (PACT), examined its relevance for three areas of teacher education: (1) planning and preparation, (2) evidence of teaching practice and (3) student work samples. On the basis of the studies reviewed, the authors conclude that the EdTPA instrument is not able to capture teacher influence over their students’ learning growth over a period of time nor their capacity to teach students with diverse learning needs in diverse contexts. The authors conclude that the TPAs disregard the “context-specific” nature of teaching, are not sufficiently linked to initial teacher education course work, are not able to capture the extent of actual work in the classroom, and assume that teacher candidates are already autonomous agents, which they are not entirely so ([Stacey et al., 2019](#)).

Teacher career policies

Progressively, policies centered on teacher performance quality, work responsibility and performance improvement have taken the form of career paths in a number of countries.

[Tournier et al. \(2019\)](#) carried out a comprehensive review of ten second-generation country career systems, all designed under the influence of neo-liberal principles and use of accountability procedures such as performance evaluation, ladders and merit pay. These included four career systems in Latin America, South Africa, Lithuania, Ethiopia, Thailand and two non-country systems in the state of New York and Scotland. Analysis of the ten systems as well as in-depth cases studies of three of them highlighted a diversity of issues regarding teacher careers, and allowed the authors to regard the Scottish system as a competent system and well supported by teachers.

Many of the countries with teacher careers are in Latin America, and in general these systems have tended to link pay rewards to externally managed performance assessments. Rewards include social benefits such as prizes and cultural goods, social security benefits, salary bonuses and scholarships ([Cuenca, 2018](#)). Mexico was one of the first countries to establish teacher career progression ladders in 2004 and 2005 with pay rewards for teachers externally assessed and who were judged to perform competently. However, from 2013 onwards a new career system took on a much stronger performative emphasis, adding on “consequences” for poorly evaluated teachers such as change of classes, of school or dismissals as such ([Echavari and Peraza, 2017](#)). The new career system has generated widespread teacher protests against what is regarded as “standardization” of their work and weaknesses in the management of the system ([Tournier et al., 2019](#)).

Other systems in Latin America include Colombia, which also has external teacher performance assessment linked to salary increases. The system, however, does not “live up to its promises”. Given that the career’s legislation made salary bonuses subject to availability of public funds, rewards in Colombia are not always handed out arguing lack of funding ([Decreto 1278, 2002](#); [Palabra Maestra, 2017](#)). Chile, legally established a formal teacher career system in 2016 made up of three compulsory and two voluntary stages covering up to twelve years of service. Assessment for progression includes evidence of content knowledge and teaching capacity exhibited in portfolio evidence with standards-based indicators used to classify teachers according to degrees of competence. Teachers failing to pass the third compulsory stage, may not continue to teach in the publicly-subsidized school sector.

Based on their review, [Tournier et al. \(2019\)](#) conclude that while career systems provide more comprehensive modes of evaluating teachers and of rewarding performance quality, their implementation does not always follow the proposed or desirable procedures. Key to good career systems are carefully selected evaluation criteria reflecting desirable performance stages (standards definition), consultation of these with teachers, appropriate evaluation procedures, trained evaluators and financial resources, both to implement the system as well as to comply with any promised salary improvements linked to career progression ([Tournier et al., 2019](#)).

Teacher transnational policy and international organizations

Teacher transnational policy was formally expressed in the 1966 *Recommendations Concerning the Status of Teachers* and in the later concept of social dialog (International Labour Organization, 2016). To monitor compliance with the Recommendations and other matters related, both the International Labor Organization (ILO) and UNESCO coordinate periodic meetings of expert teacher representatives of each of these two organizations. International organizations such as the World Bank and increasingly the OECD take on leadership positions in teacher policy formulations and are subject to differing assessments about their importance and relevance. We consider below ways in which teachers are currently the subject of different transnational institution policies. Teachers encompass a specific policy area in the *Education 2030 Framework for Action* (United Nations, 2015), designed to further progress in achieving “Education for All”, a commitment of international and national government representatives in 1990, in Jomtien, Thailand. The teacher policy focus of the *Framework for Action* is to increase the number of prepared teachers in order to reach the more than 300 million out-of-school children around the world (see: <http://uis.unesco.org/en/topic/out-school-children-and-youth>), a situation that is being substantially worsened due to the impact of COVID-19. Following, on these concerns, the more recently established *International Teacher Taskforce* is charged with furthering the policy of having a “qualified, motivated and empowered teacher in every classroom”. The Task Force consists of representatives of national governments, international organizations, non-governmental organizations (NGOs), civil society and global teacher organizations as well as foundations (155 members). Its secretariat is located in UNESCO headquarters (<https://teachertaskforce.org/who-we-are/our-mission>).

The OECD, largely a European institution in its origins, has nevertheless exerted since the 1990s a transnational influence over education. It has done so through its capacity to generate and disseminate educational information and knowledge and to act as an evaluator (on request) of country education systems (Rinne et al., 2004). Its two main international evaluations—the Program for International Student Assessment (PISA) and its Teaching and Learning International Survey (TALIS)—provide compelling information on student learning as well as on teachers and teaching.

A third important transnational institution as far as teachers are concerned is Education International (EI), the global federation of over 400 teacher unions in 171 countries and territories. Its key role is to safeguard the professional conditions of teachers around the world as established in 15 aims including those in the 1966 Declaration on Teachers, (https://www.ei-ie.org/en/detail_page/4360/principal-aims). Besides this teacher advocacy role, the EI furthers relevant knowledge through research and dissemination and seeks to influence more targeted policy-making through reports for organizations such as the G20 ministers’ conference (Education International, 2018). EI has been a permanent observer on the board of OECD’s Teaching and Learning Survey (TALIS) since its early applications. It has justified this role as crucial, in that it allows the union to exert leverage over data shaping debates and policy around teachers, whether in the form of new possibilities or constraints for its member unions (Sorensen, 2020). More recently, it may have contributed to TALIS’ shift from an emphasis on inquiring about teacher effectiveness and student learning to teacher professional development, self-efficacy, job satisfaction, teaching for inclusiveness and diversity and teacher leadership (Sorensen, 2020).

A fourth type of influential organizations are the World Bank as well as regional banks such as the Interamerican Development Bank (IDB). These bodies lay out education policy including teachers and fund education improvement through loans and grants. The World Bank recently issued what it describes as its first report entirely covering education issues: “Learning to realize education’s promise” (World Bank, 2018). The IDB, in turn, has produced a report on skills development for the 21st century. Both these documents deal with the role of teachers (IDB, 2019).

Arguable matters: global conversations about teachers and teacher education

The influence of international organizations through teacher-related studies, reports and measurements has been the subject of critical policy discussions occurring in different world locations such as Chile and Ireland, (Conway, 2013; Fernández, 2018; Stern, 2013). A contentious issue, for example, refers to the policy listing of “best practices” to be enacted as models for teaching and teacher education. Paine and Zeichner (2012) contend that practice is always “situated” or shaped by the local contexts in which it takes place. Thus, to recommend best forms of practice to be implemented in different geographical contexts is prone to lack relevance. In her comprehensive analysis of global forces and frames that influence teacher policy at national levels, Robertson (2012) equals rankings and benchmarks to global drones that collect information of the terrain, but cannot see the details that make the difference. On the other hand, international country studies not necessarily conclude with a list of worldwide solutions to what may be valid educational problems. They can be genuinely comparative, as was the case of the IEA TEDS-M study of 17 country teacher education systems (Tatto et al., 2012). The study produced broad descriptions and country evidence of teacher education programs and their forms, allowing for analysis of differences and similarities, adequately contextualized.

Of concern also, in relation to international organizations and other global players, is the manner in which reputable educational goals are justified as instruments to foster countries’ global competitiveness (Cheng, 2017; Rönnerström, 2015). Although not necessarily intended to be so, this concern applies to the world-wide emphasis on development of 21st century skills such as creativity and critical thinking, which are central to the OECD Future of Education and Skills Project (OECD, 2018a), or on “global competence” as conceptualized and evaluated in PISA 2018 (Ledger et al., 2019; OECD, 2018b). While agreeing on the key importance of developing 21st century skills, as well as providing compelling evidence for this, Rivas (2019) nevertheless asks an important question regarding its development: who is to control this? Accepting the inescapable move from pencils and blackboards to public education platforms and hybrid modes of teaching, -even before COVID schools were being thrust in this direction- Rivas (2019) answers by pointing at educators. They are those principally charged with setting the scene for 21st century

skills' development, through their: (a) openness to reflect and consider alternative roads, (b) knowledge base, pedagogy and scientific inquisitiveness, (c) empathy and concern for social justice, and (d) mediating the new knowledge world accessed via the information technologies (Rivas, 2019). In a similar vein, Paine (2019) argues in pro of teachers not only taking part but leading discussions on the policy moves which are key for the future. This requires not just "being there", but mainly being heard (Paine, 2019; Horsford et al., 2019).

Social justice, inclusion and teachers

Issues related to inclusion or treatment of differences at school level not only shape national and local discourses, but also travel across countries and influence international policy (Hajisoteriou and Angelides, 2020; Seddon et al., 2017).

Inclusion as a social justice concern is not directly apparent in the concept of education as a basic universal human right, though certainly embedded in this notion (United Nations, 2020). This is because its interpretation is contextual. In its wider meaning, inclusion refers to acceptance and embracing of diversity, as well as correction of excesses against rights as set forth in the United Nations' Declaration of Human Rights (United Nations, 1948). But in its specificity, inclusive education involves recognition in its contents and teaching of a plurality of diversity forms: physical, affective, intellectual, gender-related, ethnic, racial, cultural, among others. Inclusive education is attuned to social justice when it builds on diversity and furthers capability development (Sen, 1999) as also when it contributes to unmask injustice in its various forms (Zemblyas and Keet, 2019, p. 152). Inclusive education embraces the ways in which different student populations bring lived socio-cultural experiences into their learning processes. It does so by respecting the social norms and cultural beliefs of parents and groups to which students belong and that influence their learning (Robeyns, 2006). Teachers and schools must observe and correct in their practices what Srivastava and Sutherland (2017) refer to as the "silent exclusion" of students' identities, be this in relation to gender, class, home language, religion, caste or tribe. More recent challenges for teachers and schools are being brought about by increased world-wide migration.

Equity arguments are not always fair in addressing inclusion. Under a social justice guise, equity arguments were used to justify the development of private schools in England (Verger et al., 2016) presenting race and class inequality as the problem and school segregation as the solution (Horsford et al., 2019; Paine, 2019). Such arguments have also supported the second-class modes of school privatization in Asia and Africa, as well as allocation of "reserved seats" in private fee-paying schools in India for students who belong to "socially and economically disadvantaged groups" (Srivastava and Sutherland, 2017). Clearly, such policies do not serve what has been described above as inclusive nor what is understood as democratic education (Ball, 2007).

The worldwide importance of inclusive education policies in the light of social justice principles underlies the focus of the two most recent Global Monitoring Reports (GMR) (UNESCO, 2018, 2020b) addressing migration and inclusive education needs. Referring to the Sustainable Development Goals (United Nations, 2015), the 2020 GMR clearly defines inclusion not as equality but as equity. From this perspective, the Report deals with the challenges involved in achieving inclusive quality education, and firmly denounces the lack of attention to teacher capability development needs in this respect. More to the point, the GMR version for Latin America provides evidence of insufficient or non-existent teacher preparation for inclusive education in seven countries of the region (UNESCO, 2020c).

Teachers are clearly key players in implementing inclusive policies but, in many countries, they teach in accountability policy contexts that underscore the social and economic factors that impact on student learning gaps (Cochran-Smith and Fries, 2011). In this respect, Hajisoteiou and Angelides (2020) ask an interesting question: "are teachers trapped in a trade-off between efficiency and social justice?". Their response is that teachers can be "empowered" to bargain efficiency and equity if their capabilities are supported by appropriate initial and continuing education experiences, and by their moral and social justice values (Hajisoteiou and Angelides, 2020). Policies can also be appropriated, resisted or modified at the micropolitical school level through joint collaboration of key social actors such as school leaders, teachers, parents and students.

But, are teachers being sufficiently prepared to teach inclusively? A worldwide review of the literature on teacher initial and continuing education by Kubacka and D'Addio (2020) concludes that teacher preparation is mainly restricted to teaching learners with specific disabilities. The authors partly attribute this neglect to the breadth of exclusion forms that need to be considered such as language of instruction, gender differences in STEM (Science, Technology, Engineering and Mathematics)', curriculum subjects learning, inclusion and safety for LGBTI (lesbian, gay, bisexual, transgender and intersex) students, cultures within cultures among others. But also, the suggestion is that faced with the complexity of preparing future teachers to take in the diverse forms of student exclusion, there is need for more reflective modes of learning as in workshops and seminars in order to develop reflective practitioner capacities. Kubacka and D'Addio (2020)'s review did note that teacher education in a number of countries had made some progress in developing attitudes and values related to diversity as well as teacher collaborative practices and on how to create supportive work environments.

Conclusions

The overview of global teacher policy in the first two decades of the 21st century points to a continued impact of the broad frames affecting teacher preparation and teacher work that were subject to critical discussion in the 1990s. Although the recent unexpected global impact of the COVID-19 pandemic over teaching and student learning is and will continue to produce a re-examination of such frames.

Neo-liberal and new public management policies, entrenched in many education systems around the world, have facilitated an increase in school competitiveness and school privatization, especially in Chile and Brazil. National standardized testing has enhanced comparisons among schools putting pressure on teachers to use teach-to-the test practices in order to retain schools' positioning in league tables, and avoid loss of students and school closures. Though, a positive move as far as teacher formal evaluation is concerned, has been the use of contextual value-added measures to assess teacher contribution to learning results, adjusting for prior student academic achievement, achievement of student peers and family background.

Teacher performance evaluation has extended its geographical scope and become part of teacher careers in a number of countries with effects over teacher conditions and contracts. These effects may include salary bonuses or other monetary prizes for good performance, or on the contrary diminished teacher status and position as well as school dismissals if performance is deemed unsatisfactory. Teacher education has also been subject to performance evaluation and accountability through tightened systems of accreditation and standards, and external verification of student teacher content and practical knowledge, in Chile through knowledge testing (Ávalos and Bellei, 2019) and in the USA and Australia through teaching enactment of portfolio evidence. As recorded by researchers in diverse world locations, from South East Asia, Asia, Europe and Latin America, there are many discontented voices regarding teacher accountability forms. Using an international lens, we have referred to different institutions setting global agendas that impact on teachers, teacher preparation and teaching. UNESCO, ILO and Education International continue to express policy concerns about erosions on teacher professionalism, as well as on their working conditions. Through their contribution to evidence on global education needs and issues, the Global Monitoring Reports have examined and proposed relevant teacher-related policy changes. The 2030 Sustainable Development Goals (SDGs) are in danger of not being met given the millions of out-of-school children in poor world locations, unless trained teacher provisions are greatly increased. Teacher policy analysis and proposals in the light of the SDGs make up the substance of discussions by an international task force that includes UNESCO, civil society organizations, NGOs, and foundations.

OECD offers a different approach to teachers and teaching centered on producing relevant information from member and non-member countries on a variety of issues related to teachers' work, education, professionalism and self-efficacy among others. Beyond this role, the OECD, as also the World Bank, produce studies on teacher policy themes of which the most recent ones look into the future and discuss the development of 21st century teaching and learning capacities.

A lingering question related to policy refers to the extent of its implementation. Bruns et al. (2019) present a rather doubtful view of how successful this is. Teacher policy implementation, according to these authors' analysis, is affected by interests and policy views of stakeholders, the opaque character of classroom events as sites of reform as well as the long-term nature of policy impact. Arguably, however, Bruns et al. (2019) also attribute problems with policy enactment to teachers' unions opposition, depending on how professional or militant these are. This and other evidence suggest that policy enactment is never a match application of policy intentions. School communities and teachers reconfigure and recreate policy in the light of conditions and experiences, and it is through discussion with them that policy-makers can reassess or examine implementation effects.

Social justice and inclusion give body to the final policy frame we discuss in this article. Building on a clear distinction between equity and inclusion, we argue for inclusion as furthering social justice by means of valuing all forms of diversity. This goal, however, poses a huge challenge, as in many countries, teachers are not sufficiently prepared to educate inclusively, and/or are under pressure to respond to test results and school classification concerns more than to their students' particular identities.

A final thought as we finish this review, is that there is much to learn from innovative teacher practices and the effect of good policies in supporting these, which we have not addressed. Most recently, faced with the enormous change in schooling produced by the COVID-19 pandemic, and the need to design new forms of teaching using technology if possible, or devising other means, teachers around the world have risen to the challenges involved with greater or lesser success.

Acknowledgment

The authors gratefully acknowledge support for the preparation of this article from: PIA-CONICYT Basal Funds for Centers of Excellence Project FB0003.

References

- Acqua, S., 2013. School Accountability in England: Past, Present and Future. Technical Report. Centre for Education Research and Policy, Manchester. https://www.researchgate.net/publication/323611355_School_Accountability_in_England_Past_Present_and_Future.
- Aoki, N., Rawat, S., 2020. Performance-based pay: investigating its international prevalence in light of national contexts. *Am. Rev. Publ. Adm.* 50 (8), 865–879. <https://doi.org/10.1177/027507402091999>.
- Appel, M., 2020. Performativity and the demise of the teaching profession: the need for rebalancing in Australia. *Asia Pac. J. Teach. Educ.* 48 (3), 301–315. <https://doi.org/10.1080/1359866X.2019.1644611>.
- Ávalos, B., Bellei, C., 2019. Recent education reforms in Chile: how much of a departure from market and new public management systems? In: Ornelas, C. (Ed.), *Politics of Education in Latin America: Reforms, Resistance and Persistence*. Brill-Sense, Leiden, pp. 43–71.
- Ball, S.J., 2007. *Education Plc. Understanding Private Sector Participation in Public Sector Education*. Routledge, New York.
- Bergdahl, N., Nouri, J., 2020. Covid-19 and crisis-prompted distance education in Sweden. *Technol. Knowl. Learn.* <https://doi.org/10.1007/s10758-020-09470-6>. Published online.
- Bergh, A., Löfdahl Hultman, A., Englund, T., 2019. Local enactment of the Swedish' advanced teacher reform. *J. Curric. Stud.* 51 (3), 326–341.

- Biesta, G.J.J., 2010. Why what works still won't work: from evidence-based education to value-based education. *Stud. Philos. Educ.* 29, 491–502.
- Bruns, B., Macdonald, I.H., Schneider, B.R., 2019. The politics of quality reforms and the challenges for SDGs in education. *World Dev.* 118, 27–38. <https://doi.org/10.1016/j.worlddev.2019.02.008>.
- Charteris, J., 2019. Teaching performance assessments in the USA and Australia. Implications of the “bar exam for the profession”. *Int. J. Comp. Educ. Dev.* 21 (4), 237–250. <https://doi.org/10.1108/IJCED-10-2018-0039>.
- Cheng, K.M., 2017. Advancing 21st century competencies in East Asian educational systems. *Asia Soc. Center Glob. Educ.* <https://asiasociety.org/files/21st-century-competencies-east-asian-education-systems.pdf>.
- CIAE, Centro Educación Inclusiva PUCV, EduGlobal, OREALC, 2020. Nuevos contextos, nuevas demandas y experiencia docente en Chile (Santiago, Chile).
- Cochran-Smith, M., Fries, K., 2011. Teacher education policy and social justice. In: Earley, P.M., Imig, D., Michelli, N. (Eds.), *Teacher Education Policy in the United States: Issues and Tensions in an Era of Evolving Expectations*. Routledge, New York, pp. 182–207.
- Cochran-Smith, M., Carney, M.C., Keefe, E.S., Burton, S., Chang, W.C., Fernández, M.B., Sánchez, J.G., Baker, M., 2018. *Reclaiming Accountability in Teacher Education*. Teachers College Press, New York, NY.
- Conway, P.F., 2013. Cultural flashpoint: the politics of teacher education reform in Ireland. *Educ. Forum* 77 (1), 51–72. <https://doi.org/10.1080/00131725.2013.739021>.
- Cuenca, R., 2018. La cuestión docente en América Latina y el Caribe: Tendencias y desafíos. Informe de apoyo técnico para la Reunión Ministerial E2030, Sucre, Bolivia 2018. <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/10-R-Cuenca-cuestion-docente-ALC-ESP.pdf>.
- Da Rosa, S.S., 2019. Políticas regulatórias, subjetividade e os entraves a democracia a escola pública brasileira: Contribuições à pesquisa curricular. *Currículo sem Fronteiras* 19 (3), 844–867. Retrieved 29 November 2020: <http://curriculosemfronteiras.org/vol19iss3articles/rosa.pdf>.
- Danielson, C., 2007. *Enhancing Professional Practice: A Framework for Teaching*, second ed. ASCD, Alexandria, VA.
- Darling-Hammond, L., Sykes, G. (Eds.), 1999. *Teaching as the Learning Profession. Handbook of Policy and Practice*. Jossey-Bass, Inc, San Francisco.
- Decreto 1278, 2002. Estatuto de Profesionalización docente. Retrieved 2 January 2021. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=5353>.
- Department of Education, 2011. *Teachers' Standards. Guidance for School Leaders, School Staff and Governing Bodies*. Retrieved 7 January 2021. <https://www.gov.uk/government/publications/teachers-standards>.
- DFE, 2019. *ITT Core Content Framework*. Retrieved 13 January 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/919166/ITT_core_content_framework.pdf.
- Dunleavy, P., Margetts, H., Bastow, S., Tinkler, J., 2006. New public management is dead—long live digital era governance. *J. Publ. Adm. Res. Theor.* 16 (3), 467–494. <https://doi.org/10.1093/jopart/mui057>.
- Echávari, J., Peraza, C., 2017. *Modernizing Schools in Mexico: The Rise of Teacher Assessment and School-Based Management Policies*, vol. 25. Retrieved 11 January 2021. <http://epaa.asu.edu/ojs/>.
- Education International, 2018. How to Ensure a Cadre of Competent, Effective and Well-Supported Teachers: A Policy Paper for the G20. Statement to G20 Education and Employment Ministers. https://www.ei-ie.org/media_gallery/f59e3.pdf.
- ESSA, 2015. US Department of Education. Retrieved 7 January 2021. <https://www.ed.gov/essa?src=rn>.
- Evetts, J., 2014. The concept of professionalism: professional work, professional practice and learning. In: Billet, S., et al. (Eds.), *International Handbook of Research in Professional and Practice-Based Learning*. Springer, Dordrecht, pp. 29–56.
- Fernández, M.B., 2018. Framing teacher education: conceptions of teaching, teacher education, and justice in Chilean national policies. *Educ. Pol. Anal. Arch.* 26 (34). <https://doi.org/10.14507/epaa.26.2806>.
- Gewirtz, S., Ball, S.J., Bowe, R., 1995. *Markets, Choice and Equity in Education*. Open University Press, Buckingham.
- Hajisoteriou, C., Angelides, P., 2020. Efficiency versus social justice? Teachers' roles in the epoch of globalization. *Educ. Citizen Soc. Justice* 15 (3), 274–289. <https://doi.org/10.1177/1746197919852564>.
- Hanushek, E.A., 2019. Testing accountability and the American economy. *Annals, AAPSS* 683. <https://doi.org/10.1177/0002716219841299>.
- Hardy, I.J., Reyes, V., Hamid, M.O., 2019a. Performative practices and “authentic accountabilities”: targeting students, targeting learning? *Int. Educ. J.* 18 (1), 20–33. <https://openjournals.library.sydney.edu.au/index.php/IJ/article/view/13354>.
- Hardy, I., Rönnerman, K., Beach, D., 2019b. Teachers' work in complex times: the “fast policy” of Swedish school reform. *Oxf. Rev. Educ.* 45 (3), 350–366.
- Hargreaves, A., 2000. Four ages of professionalism and professional learning. *Teach. Teach.* 6 (2), 151–182.
- Hassan, M.M., Mirza, T., Hussain, M.W., 2020. A critical review by teachers on the online teaching-learning during COVID-19. *Int. J. Educ. Manag.* 5, 17–27. <https://doi.org/10.58125/ijerme.2020.05.02>.
- Helgetun, J.B., Menter, I., 2020. From an age of measurement to an evidence era? Policy-making in teacher education in England. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2020.1748722>.
- Hogan, A., Thompson, G., 2019. The quasi-marketization of Australian public schooling: affordances and contradictions of the new work order. *Asia Pac. J. Educ.* 39 (3), 391–403.
- Holloway, J., Sorensen, T.B., Verger, A., 2017. Global perspectives on high-stakes teacher accountability policies: an introduction. *Educ. Pol. Anal. Arch.* 25 (85). <https://doi.org/10.14507/epaa.25.3325>.
- Horsford, S.D., Scott, J.T., Anderson, G.L., 2019. *The Politics of Education Policy in an Era of Inequality: Possibilities for Democratic Schooling*. Routledge, New York.
- Hoyle, E., 1974. Professionalism, professionalism and control in teaching. *Lond. Educ. Rev.* 3 (2), 15–70.
- Hundred Org., 2020. *Spotlight: Quality Education for All During Covid-19 Crisis*. Retrieved 23 November 2020. <https://hundred.org/en/collections/quality-education-for-all-during-coronavirus>.
- Hunter, M., 1982. *Mastery Teaching*. Corwin Press, Thousand Oaks, CA.
- Hursh, D., 2005. Neoliberalism, markets and accountability: transforming education and undermining democracy in the United States and England. *Pol. Futures Educ. Internet* 3 (1), 3–15. <https://journals.sagepub.com/doi/pdf/10.2304/pfie.2005.3.1.6>.
- IDB, 2019. *The Future is Now. Transversal Skills in Latin America and the Caribbean in the 21st Century*. Interamerican Development Bank. <https://www.iadb.org/>.
- International Labour Organization, 2016. *The ILO/UNESCO Recommendation Concerning the Status of Teachers (1966) and the UNESCO Recommendation Concerning the Status of Higher-Education Teaching Personnel (1997)*, revised ed. 2016. ILO, Geneva.
- Jewell, J.W., 2017. From inspection, supervision and observation to value-added evaluation: a brief history of US teacher performance evaluation. *Drake Law Rev.* 65, 363–419. Retrieved 4 January 2021. <https://lawreviewdrake.files.wordpress.com/2015/01/jewell-final.pdf>.
- Kapucu, N., 2006. New public management: theory, ideology and practice. In: Farzmand, A., Pinkowski, J. (Eds.), *Handbook of Globalization, Governance, and Public Administration*. Routledge, New York, pp. 889–901.
- Keating, I., 2015. English Case Study. Professional Autonomy, Accountability and Efficient Leadership and the Role of Employers Organisations, Trade Unions and School Leaders. Retrieved 7 January 2021. https://www.csee-etu.org/images/attachments/RP_Professional_Autonomy_Accountability.pdf.
- Komatsu, H., Rappleye, J., 2017. A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth. *Comp. Educ.* 53 (2), 166–191. <https://doi.org/10.1080/03050068.2017.1300008>.
- Kraft, M., Simon, N.S., Lyon, M.A., 2020. Sustaining a Sense of Success: The Importance of Teacher Working Conditions During the COVID-19 Pandemic (Ed Working Paper: 20–279). <https://doi.org/10.26300/35nj-v890>. Retrieved from Annenberg Institute at Brown University.
- Kubacka, K., D'Addio, A.C., 2020. Targeting teacher education and professional development for inclusion. *J. Int. Coop. Educ.* 22 (2/3), 89–106.
- Ledger, S., Thier, M., Bailey, L., Pitts, C., 2019. OECD's approach to measuring global competency: powerful voices shaping education. *Teach. Coll. Rec.* 121, 080302.

- Lee, T.T.L., Kwan, P., Li, B.Y.M., 2019. Neoliberal challenges in context: a case of Hong Kong. *Int. J. Educ. Manag.* 34 (4), 641–652. <https://doi.org/10.1080/03057925.2020.1756745B.Y.M>.
- Mayer, D., Mills, M., 2020. Professionalism and teacher education in Australia and England. *Eur. J. Teach. Educ.* <https://doi.org/10.1080/02619768.2020.1832987>.
- Mockler, N., Stacey, M., 2020. Evidence of teaching practice in an age of accountability: when what can be counted isn't all that counts. *Oxf. Rev. Educ.* <https://doi.org/10.1080/03054985.2020.1822794>. Published online 6.10.2020.
- National Commission on Excellence in Education, 1983. *A Nation at Risk: the imperative for educational reform*. Elem. Sch. J. 84 (2), 113–130.
- No Child Left Behind, 2001. Elementary and Secondary Education Act. US Department of Education. <https://www2.ed.gov/nclb/landing.jhtml>.
- OECD, 2018a. The Future of Education and Skills: Education 2030. Directorate for Education and Skills. <https://www.oecd.org/education/2030-project/>.
- OECD, 2018b. Preparing Our Youth for an Inclusive and Sustainable World. The OECD Global Competence Framework. OECD, Paris. <https://www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf>.
- OECD, 2020. TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals. OECD, Paris.
- Paine, L.W., Zeichner, K., 2012. The local and the global in reforming teachers and teacher education. *Comp. Educ. Rev.* 56 (4), 569–583.
- Paine, L.W., 2019. Missing voices and possible dialogues: problems and possibilities for teacher education. *Teach. Teach.* 25 (6), 684–702. <https://doi.org/10.1080/13540602.2019.1643712>.
- Palabra Maestra, 2017. ¿Cuál es el estado actual del proceso de evaluación de los maestros del 1278? ¿Cuál es la situación de los maestros del 2277? ¿Cuál es la política detrás del congelamiento salarial. Compartir Palabra Maestra. Retrieved 2 January 2021. <https://www.compartirpalabramaestra.org/actualidad/columnas/que-pasa-con-la-carrera-docente-en-colombia-parte-i>.
- Pons, X., 2014. Les trajectoires des inspections scolaires en Europe: analyses comparatives. *Rev. Fr. Pédagog.* 186, 5–10. <https://doi.org/10.4000/rfp.4387>.
- Prospects, 2020. Routes Into Teaching. <https://www.prospects.ac.uk/jobs-and-work-experience/job-sectors/teacher-training-and-education/routes-into-teaching>.
- Reimers, F.M., Chung, C.K., 2016. Mapping the landscape of teaching and learning for the twenty-first century in Massachusetts in the context of US educational reform. In: Reimers, F.M., Chung, C.K. (Eds.), *Teaching and Learning for the Twenty-First Century. Educational Goals, Policies, and Curricula from Six Nations*. Harvard University Press, Boston.
- Reimers, F., Schleicher, A., 2020. A Framework to Guide an Education Response to the COVID-19 Pandemic of 2020. OECD, Paris. https://globaled.gse.harvard.edu/files/geii/files/framework_guide_v2.pdf.
- Rinne, R., Kallo, J., Hokka, S., 2004. Too eager to comply? OECD education policies and the Finnish response. *Eur. Educ. Res. J.* 3 (2), 454–485.
- Rivas, A., Sánchez, B., 2020. Race to the classroom: the governance turn in Latin American education. The emerging era of accountability, control and prescribed curriculum. Compare. <https://doi.org/10.1080/03057925.2020.1756745>.
- Rivas, A., 2019. ¿Quién Controla el Futuro de la Educación? Siglo Veintiuno, Buenos Aires.
- Robertson, S., 2012. Placing teachers in global governance agendas. *Comp. Educ. Rev.* 56 (4), 584–607. <https://www.journals.uchicago.edu/doi/pdf/10.1086/667414>.
- Robeyns, I., 2006. Three models of education: rights, capabilities and human capital. *Theor. Res. Educ.* 4 (1), 69–84. <https://doi.org/10.1177/1477878506060683>.
- Rönström, N., 2015. Educating competitive teachers for a competitive nation? *Pol. Futures Educ. Internet* 13 (6), 732–750. <https://doi.org/10.1177/1478210315595171>.
- RTTT, 2009. US Department of Education. Retrieved 7 January 2021. <https://www2.ed.gov/programs/racetothetop/index.html>.
- Sahlberg, P., 2020. Does the pandemic help us make education more equitable? *Educ. Res. Pol. Pract.* 20, 11–18. <https://doi.org/10.1007/s10671-020-09284-4>.
- Schleicher, A., 2020. The Impact of COVID-19 on Education. Insights from Education at a Glance 2020. OECD Publishing, Paris.
- Seddon, T., Ozga, J., Levin, J.S., 2017. Global transitions and teacher professionalism. In: Seddon, T., Levin, J. (Eds.), *World Yearbook of Education 2013. Educators, Professionalism and Politics: Global Transitions, National Spaces and Professional Projects*. Routledge, New York, pp. 3–24.
- Sen, A., 1999. *Development as Freedom*. Oxford University Press, Oxford.
- Sorensen, T.B., 2020. The space for challenge in transnational education governance: the case of Education International and the OECD TALIS program. *Discourse*. <https://doi.org/10.1080/01596306.2020.1718611>.
- Srivastava, P., Sutherland, L., 2017. Beyond Access: Silent Exclusion and Meaningful Access in Government and Private Schools in India. Paper Presented at the 2017 UKFIET Conference, Oxford, England, 5–7 September, 2017.
- Stacey, M., Mayer, D., Talbot, D., 2019. The development of an Australian teacher performance assessment: lessons from the international literature. *Asia Pac. J. Teach. Educ.* 23, 9–19. <https://www.tandfonline.com/doi/10.1080/1359866X.2019.1669137>.
- Stern, M., 2013. Bad teacher: what race to the top learned from the “race to the bottom”. *J. Crit. Educ. Pol. Stud.* 11 (3), 194–229. <http://www.jceps.com/wp-content/uploads/PDFs/11-3-09.pdf>.
- Tatto, M.T., Schwillie, J., Senk, S., Ingvarson, L., Rowley, G., Peck, R., 2012. Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M). International Association for the Evaluation of Educational Achievement, Amsterdam.
- Tolofari, S., 2005. New public management and education. *Pol. Futures Educ. Internet* 3 (1), 75–89.
- Tournier, B., Chimier, C., Childress, D., Raudonyte, I., 2019. Teaching Career Reforms: Learning from Experience. International Institute for Educational Planning, Paris.
- Tsz-lok Lee, T., Kwan, P., Yuet Man Li, B., 2019. Neo-liberal challenges in context: a case of Hong Kong. *Int. J. Educ. Manag.* 34 (4), 641–652. <https://doi.org/10.1108/IJEM-06-2019-0220>.
- UNESCO, 2014. *Global Education Monitoring Report: Teaching and Learning: Achieving Quality for All*. UNESCO, Paris.
- UNESCO, 2018. *Global Education Monitoring Report 2019: Migration, Displacement and Education—Building Bridges, Not Walls*. UNESCO, Paris.
- UNESCO, 2020a. Education in a Post-COVID World: Nine Ideas for Public Action. International Commission on the Futures of Education. UNESCO, Paris. <https://en.unesco.org/news/education-post-covid-world-nine-ideas-public-action>.
- UNESCO, 2020b. *Global Education Monitoring Report: Inclusion and Education: All Means All*. UNESCO, Paris.
- UNESCO, 2020c. *Global Education Monitoring Report 2020—Latin America and the Caribbean—Inclusion and Education: All Means All*. UNESCO, Paris.
- United Nations, 1948. Universal Declaration of Human Rights: Article 26. <https://www.right-to-education.org/page/united-nations-instruments>.
- United Nations, 2015. Transforming Our World: The 2030 Agenda for Sustainable Development. <https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981>.
- United Nations, 2020. Policy Brief: Education During COVID 19 and Beyond. Retrieved: 23 November 2020. <https://unsdg.un.org/resources/policy-brief-education-during-covid-19-and-beyond>.
- Verger, A., Fontdevila, C., Zancajo, A., 2016. *The Privatization of Education: A Political Economy of Global Education Reform*. Teacher College Press, New York.
- World Bank, 2018. *Learning to Realize Education's Promise*. International Bank for Reconstruction and Development/The World Bank, Washington, D.C.
- Zeichner, K., Peña-Sandoval, C., 2015. Venture and teacher education policy in the U.S: the role of the new schools venture fund. *Teach. Coll. Rec.* 117 (6).
- Zembylas, M., Keet, A., 2019. *Critical Human Rights Education: Advancing Social-Justice-Oriented Educational Practice*. Springer, Cham, Switzerland.

The context in teacher education and professionalism: hierarchies, networks and markets in four countries

Monica Mincu^{a,c} and Min Liu^b, ^a University of Torino, Turin, Italy; ^b Beijing Normal University, Beijing, China; and ^c University College London, London, UK

© 2023 Elsevier Ltd. All rights reserved.

Introduction: context and comparison in teacher education	72
Theoretical framework: new forms of governance and professionalism	73
Methodological note	75
School networks for ITE in England	75
Local hierarchies and the market of employers in the French ITE	76
Specialized networks of professionals in ITE for primary school teachers in Italy	77
Multilevel hierarchies and institutional networks in Chinese CPD	78
Discussion and conclusion	79
References	81

Introduction: context and comparison in teacher education

There is a growing international awareness that effective teacher policies significantly improve the quality of education (European Commission, 2012). However, the governance of TE varies between and within countries and much can be learned from comparison. The institutional arrangements and the control of such arrangements through the governance and the policy mechanisms, can be conceptualized as the “context of TE”. More traditionally, the context of education is discussed in terms of policy regulations or as the influence of larger social, economic and cultural factors (e.g., unions and political parties) and conceptually elaborated at different scales: micro, meso and macro. The context is more generally identified as country context and variously conceptualized and taken into account in quantitative and qualitative studies. In more qualitative studies the context is theoretically elaborated as cultural, social and policy context at national or international level or at the intersection of global and local factors, either as case studies (Paine and Fang, 2006; Steiner-Khamsi, 2005) or as comparative scholarship (Mayer and Mills, 2021; Osborn and McNess, 2005; Tatto and Mincu, 2009), and as context of teaching (Zhao, 2018). One review synthesizes several contextual factors from a comparative education perspective (see Tatto, 2008) such as the socio-economic status of parents, children, as well of schools themselves (Robinson, 1994; Fuller et al., 1999), of schooling (Mak, 1999), the political dimension (Ginsburg and Lindsay, 1995; Judge et al., 1994) and culture and language. A more traditional field specific way to tackle with the contextual dimension is the immediate institutional and pedagogical environment of TE programs (context-based programs Munoz-Izquierdo, 1994 in Tatto, 2008; type of providers context Davies et al., 2016; type of programs contexts Goldhaber, 2018). For Cochran-Smith and Fries (2009), the policy and politics can be identified as “shifting population patterns, major swings in the political pendulum, the growth of a complex global economy, and a number of educational and political movements [such as] the standards movement, the accountability movement, initiatives to privatize education and other public services, the press for market-based education reform, and the emergence of access to quality education as a civil rights issue” (p. 39). Another example is to consider the context in terms of new generations and new types of working demands (Mayer, 2006). A recent OECD study identifies implicitly the context with the factors affecting the governance of the teaching profession (Guerriero and Deligiannidi, 2017). In this case, some of these determinants have a larger societal character (e.g., teacher unions, the government and the administration) while others are rather field specific (inspection and standards, for instance).

An alternative way to conceptualize the context is the identification of “hierarchies, networks and markets”. The advantages of this approach is the possibility of combining policy and governance arrangements as well as political, economic and institutional cultures. Rhodes (1997) proposes three modes of governance: hierarchy, markets and networks, while arguing that network is becoming the dominant form within TE. Ball (2009) goes further and proposes a new way of approaching networks as “heterarchies”, referring to a mix of hierarchy, market and network to allow differentiated elements in the policy process that co-exist and co-operate. This concept links the economy, ideology and institutions to the actors and their agencies, both as individuals or policy elites.

An “interactive type of governance” (Torfing et al., 2012) in which networks play an increasingly important role is created by the contribution of many players and stakeholders in a social interactive arena. In this sense, the role of non-governmental actors, such as international organizations and think tanks, replacing bureaucrats and university researchers and scholars as educational experts is paramount (Han and Ye, 2017). The context of education and TE in particular is a renewed space formed by a plurality of social actors and their web of interactive networks.

The context is particularly relevant when adopting a comparative approach of TE in various countries or as single case studies that are oriented toward understanding the dynamics, configuration and outcomes of TE in various locations and is essentially

a sociological endeavor. The identification of the “context” of TE in one country, be it ITE or CPD, requires the selection of factors that can explain specific educational phenomena or institutional arrangements that are the object of the inquiry. In other words, “the context” is not a generic category that can include any macro-factor without distinction, however relevant for education in that country. The researcher contributes carefully to sketching the most pertinent context given the subject and the object of investigation, that is the focus and the theoretical issue at stake (Thomas, 2011). For instance, TE in the English case requires the elaboration of a specific context for which the neoliberal politics or the politics of the coalition or right-wing governments cannot adequately provide. In fact, a more pertinent selection of the relevant societal factors, beyond the type of approach, subject and methodology, may include the presence of hierarchies (inside school or as external accountability mechanisms, i.e., Ofsted), the presence of significant educational markets (in general and TE related), as well as of relevant institutional networks (teaching school alliances and other partnerships). Moreover, the investigation of a specific institutional aspect or pedagogical culture linked to TE entails the analysis of various contextual levels: the more immediate organizational level, the school system level, local or national government and even the global or transnational level.

The contextual dimensions and levels can vary according to number and type depending upon the perspective and approach undertaken. For instance, from the world-system or world culture perspective, the context is seen as an outcome of convergent forces and the school system exhibiting a similar *modus operandi* (Meyer and Ramirez, 2000). The theories that consider the political authority as influenced by a major ideology (e.g., nationalism, socialism), the school systems arrangements reflect the functioning of the State and its over-arching political philosophy. Other theories may consider other influential contextual forces that are not necessarily linked to the State authority.

This article is a review of previous and ongoing works while it also contributes to the advancement of theory on the role of the context and its conceptualization in TE. From a case study perspective (Thomas, 2011) the “subject” is the context of a form of TE, either ITE or CPD, while the “object” consists of the new theoretical elaboration of the context, in terms of the interactive governance of hierarchies, market, networks and the way their interplay shapes professionalism. Given the available space the cases are instrumental snapshots that serve to sketch TE arrangements in contexts with the purpose to illustrate theory building.

Theoretical framework: new forms of governance and professionalism

A major aspect to be considered when dealing with the “context” of education, and teacher education in particular, is the governance of the public systems that is actually reshaping and remodeling the traditional horizontal and vertical mechanisms of control, institutional partnering and competition. The State “steers at a distance” (Clarke et al., 2000) and recently it does so through new forms of governance, most often partnerships and networks, thus overcoming a traditional hierarchical model in the direction of a “polycentric governance in Western countries” (Ball and Exley, 2010, 151). In the same vein, Ehren et al. (2017) suggest that such configurations are to be seen as rather polycentric and therefore different from the more classical and monocentric forms, since decision making is now shared between the State and other institutional actors. The *partners* are, at the same time, deciding the rules and also ensuring they are respected via a complex system of interdependent relationships. Such governance tendencies have a deep impact on the stakeholders and can produce new institutional configurations and interactions in the field. In continuity with the previous New Public Management (NPM) phase (Theisens, 2016). New Public Governance is therefore characterized by horizontally organized social systems that can operate as networks. The governments—traditionally acting as hierarchies—can either enter to become part of such networks, monitoring the networks’ initiatives and decide the framework of their actions and collaboration from a distance, as well as facilitating the sharing of knowledge (Thompson, 2003).

As a key feature of current education contexts, networks are managerial forms that bring together various organizations such as public and private agencies and departments, in order to facilitate public action. In education this means that the responsibility pertains to the network and not to single institutions. As a case in point from the TE field, this trend is clearly visible in the creation of teaching school alliances in England, delivering both school-led alternative ITE and CPD. A second example comes from the dual partnering of universities and local regional authorities/academies in both ITE (INSPE) and CPD in France. In addition, in France the traditional hierarchical function of the school inspection has been transformed into a more supportive role for the teaching profession. Thereby, the network of professionals involved with CPD is composed of academics, local state representatives, inspectors and teachers as trainers.

Even in centralized countries where hierarchies are traditionally prominent, networks are becoming the key mechanisms of governance in the delivery of ITE and thus significantly transform the context of teacher preparation. Such recent policy transformation is backed by the premise that networks can be more significantly motivated by a desire for public service, compared to hierarchies (*bureaucracies*) or markets. Teachers and schools are placed in the “front line” of the education service, and this position would imply the possibility of securing access to valid and effective knowledge that can be more easily applied to specific situations, compared to the knowledge offered by those who are less well placed, such as within universities or training centers. Research conducted in England (Mincu and Davies, 2019) did not confirm the idea that networks of actors placed in the front line are morally superior when it comes to a public service motivation. The pressure exerted by markets and hierarchies on the training alliances involved with ITE might have influenced motivation. More autonomy does not immediately translate into a better service toward the public, even if the actors themselves may be genuinely convinced of this.

Recent reforms have strongly engaged with the context of TE and specifically with the forms of governance most appropriate to promoting a certain idea of teacher professionalism. In fact, different forms of governance of TE (both CPD and ITE) have

implications for teachers' capabilities, motivation and practice. Building on how providers shape new teachers' knowledge and professionalism, this article will add to the evidence base by presenting case studies of governance regimes in Europe and China. Given the cost of educating new teachers, it is unsurprising that there have been changes in the governance of ITE. At the same time, in some countries, CPD is traditionally a matter of partnerships between university, public centers or inspection services, private providers in free markets, schools and even teachers themselves.

NPM is a contextual factor that explains the specific institutional arrangements and types of professionalism. Even if NPM is a rather generic factor contextually, its impact is unequal and evolves in different directions in various countries. English governments have shown the most enthusiasm for NPM, experimenting with different forms of governance for ITE (.). Change in public sector governance in France has been described as a "decline in Napoleon's bureaucracy" and in Italy as deregulation. Despite being different cases historically, both France and Italy have introduced accountability measures into education, moving to a more organizational or "business like" teacher professionalism. A comparative review of the different trajectories in the UK, France and Italy should therefore yield valuable insights for research and policy. A previous analysis of NPM reforms has largely relied on idealizations of either a new configuration of governance through hierarchy and markets or an idealization of governance through professional networks. Chinese education is one sector of the public policy that has transformed from top-down state-directed into network governance (Han and Ye, 2017) clearly being influenced by the NPM model (Zhang and Bray, 2017). A pluralization of the public space in China implies that more ministries are cooperating or competing on various matters, that more matters are decided by the State council and not directly by the Party itself, more governance levels are involved (province and district, beyond the national), there is more institutional partnering in education and in the preparation of teachers.

The transformations of the context of TE, policy and governance first of all, but also the wider political and economic environment over the last decades has had huge implications for the model of teacher professionalism. A call to increased collaboration to "involve all key actors" (European Commission, 2012) in ITE recognizes that increasingly complex education systems are characterized by various levels of control and multiple stakeholders. A similar focus on strong partnership is at work in the recent policies of CPD in China (Liu, 2020).

At a theoretical level, teacher professionalism is a key outcome of the form of governance experienced and has been typically defined in relation to such macro-contextual factor (Evetts, 2009). I have summarized three of these definitions in Table 1. Hoyle (1974) vision of extended professionalism assumes that governance of ITE will be exercised through a professional network embedding public service values so that self-directed teachers will naturally serve the public good. His distinction between "extended" and "restricted" professionalism asserts that collaboration beyond the school influences new teachers' beliefs about relationships between knowledge and practice. Evetts (2009) account portrays managerial professionalism in the context of hierarchy and market governance through NPM. The depiction of collaborative professionalism (Whitty, 2008) adjusts the NPM model to include a requirement for teachers to be ready and able to collaborate with other professions when education is seen as part of social welfare. The fourth column summarizes definitions such as "occupational professionalism", "democratic professionalism" or "activist professionalism" (Sachs, 2001), which are considered as rather overlapping (see Mayer and Mills, 2021). Each definition is associated with a perspective on (1) the knowledge that teachers need (2) the values that will direct teacher activity in the absence of external constraints and incentives (3) the degree to which they should exercise self-direction in their work and (4) the role that collaboration should play in their work. Gaps in the table have been left where reading of previous literature has left the stance on this point unclear.

Each column in Table 1 presents a view of teacher professionalism with specific implications for the status of teaching as a career, the role that teachers are expected to play in forming societies and the independence that the teaching profession will enjoy. The first row invites investigation of the relationships between governance and new teachers' epistemological beliefs (Sing Chai et al., 2009). In particular, where they look for sources of "expertise". The second row invites an investigation of (new) teacher values orientation. Are new teacher values stable (as implied by research on values' orientation, indicating this issue can only be addressed at recruitment)? What qualitative differences can be identified in new teachers' self-direction? (Oosterheert et al., 2002) and to what extent does new teachers' capacity for self-direction change during ITE? What kinds of collaboration do new teachers believe are

Table 1 Forms that professionalism may take: networks of professionals/governments/agencies.

	<i>Extended professionalism</i>	<i>Managerial professionalism (NPM responding to the public choice critique)</i>	<i>Collaborative professionalism (variant of NPM)</i>	<i>Occupational/activist/democratic professionalism</i>
Knowledge	Professional judgment/application of disciplinary (HE-led) knowledge	Government defined standards for practice	Education as a strand within social welfare	Professional judgment/research-based oriented
Values assumed	Public service motivation required	Self-interest	Public service motivation	Public service motivation
Self-direction	Complex work requires capacity for self-direction	Adherence to best procedures	Adherence to best procedures	Autonomy
Collaboration	Between teachers and between teachers and academics to improve practice	Contrived collegiality	Multi-agency work for public service	Collegial authority

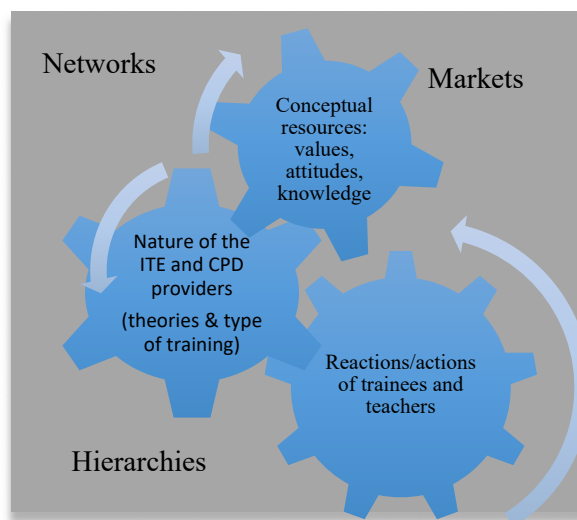


Fig. 1 The context, the nature of ITE/CPD providers and conceptual resources.

worthwhile and how does ITE influence their practice in, and beliefs about, collaboration (Musset, 2010)? Do the forms of collaboration in providing ITE/CPD play any role in forming (new) teachers' ideas about collaboration in teaching (Grimaldi, 2011)?

The nature of the providers and their institutional dynamics mediates the types of knowledge integrated into the teachers' theories of practice (Timperley et al., 2007) which are particularly relevant to inform effective teaching. The way professionalism is represented and produced, how (new) teachers' learning is approached and professional knowledge is developed are all intertwined. As showed in Fig. 1, conceptual resources are mediated by the nature of the ITE and CPD providers. The larger context of hierarchies, markets and networks is key to understanding professionalism and the way it is shaped by specific institutions. Knowledge can emerge, mainly but not exclusively, as inside/outside practitioner research through inquiry (Cochran-Smith and Lytle, 2009) or as academic research that is usually mobilized from outside (when HEIs or other expert institutions are ITE or CPD providers) or through the school system. Knowledge can also be mediated and shared by a network of private providers and trainers in a free market.

Methodological note

This article draws on teacher education scholarship, comparative education studies and education policy field. The cases exemplify the context of teacher education in countries that are different from this point of view. From a welfare state perspective (Esping-Andersen, 1990; Ferrera, 1996), all countries belong to different types—Southern European (Italy), Conservative-corporatist (France), liberal (England), a combination of social democracy, liberal and conservative-corporatist (China). From a governance perspective (Windzio et al., 2005), the types are as follows: East-Asian (China), Anglo-Saxon and Eastern European (England), State based (France) and South European (Italy). A diversity of contextual arrangements allows for theory illustration. A renewed way to interpret the context in teacher education is also functional to a new way to interpret and typify professionalism as main outcome of contextual variation. At the same time, the authors draw on data collected in these countries and further elaborate a conceptual approach to context initially developed by Mincu and Davies (2019).

School networks for ITE in England

Since 2010, education policy in England shifted ITE toward a "school-led" system with an emphasis on an apprenticeship model. The term "school-led" draws attention to the locus of decision-making in schools rather than higher education (HE). A HE-led ITE system may involve substantial "school-based" experience (or "practicum") for future teachers. What has changed in this school-led system is the way that the ITE is governed: who has control of how it is provided and how this control is exercised. The post-2010 system has favored governance through networks of schools. While the apparatus of hierarchical control has remained, the focus of the narrative in England since 2010 has shifted to giving control to local networks of schools (Mutton et al., 2017). This policy change was justified by a claim that schools rather than HE held the powerful knowledge necessary for future teachers (Gove, 2010).

The narrative trusts that governments can use contracting to reconcile a desire to exert hierarchical control with a devolution of governance to local school networks. In a contracting model the government delegates responsibility for provision to "for-profit" or "not-for-profit" organizations that are required to deliver contractual outcomes in return for payments from government. The

English government concluded that it was not able to make the ITE system work through contracts with HEIs, but does believe it can work with schools. This is consistent with a simple transfer of market thinking to teacher education: **powerful knowledge about teaching is generated by competing providers.**

Advocacy of school networks in England (e.g., Hill, 2011; Matthews and Berwick, 2013) has claimed that these networks will be motivated by a “moral imperative” in the public interest. This belief directly contradicts the assumption of self-interest that has driven NPM and the contracting model. It also overlooks the way in which the contracting model has opened the provision of schooling and teacher education to individuals, organizations and networks whose experience, beliefs and practices are rooted in fields beyond education (Ball and Junemann, 2012; Zeichner and Pena-Sandoval, 2015; Ellis et al., 2019).

Opportunities for groups of schools to form a partnership for teacher training in England were created in the early 1990s. These partnerships provided “School Centered Initial Teacher Training” (SCITTs) a school-based, employment route to qualified teacher status, the Graduate Teachers’ Program (GTP) introduced in 1998. Individuals recruited on this route received a salary from the school (heavily subsidized by the government) during their training, prior to certification. This scheme grew more rapidly than the SCITT program and by 2005, employment-based routes into teaching accounted for roughly a fifth of all ITT places in England and Wales (Smith and McLay, 2007). A pivotal moment in England arrived with the election of a coalition government in 2010 and a Secretary of State for Education who was committed to school networks and an apprenticeship model for ITT (Gove, 2010). The idea of “Teaching schools” as beacons of good practice located the apprenticeship model in institutions rather than individual mentors. Teaching schools would act as leaders of school networks, which became known as “Teaching School Alliances” (TSAs).

In 2012/13 the government introduced a new route into teaching referred to as “School Direct”. This program intended to shift the leadership of ITT from universities to schools to a much greater extent than had been achieved through the SCITT schemes. The establishment and growth of TSAs provided a realistic platform for achieving this objective and within 2-years, School Direct accounted for 30% of all ITT recruitment (Davies et al., 2016). This case study has portrayed a micro-network of schools that has achieved sustained success in the context of these struggles (as indicated by stable membership of the network, outstanding rating in government inspections of teacher education, and securing new contracts for ITE). And yet it does not seem to epitomize **the vision of an empowered local school network that has been depicted by policy (Gove, 2010).** The network was **driven by the need to adhere to government standards and to secure praise from government inspectors.**

In a study by Mincu and Davies (2019) the role of the contexts has been observed, particularly the impact of hierarchical and market governance. In fact, school networks are the favored providers of ITE in England in a “school-led” system. The evidence from participant observations and interviews conducted over a 5-month period has been analyzed as strategic intent and an enactment of a School Direct program. The policy of “school-led” ITE has been directed by statements suggesting that governance will operate through local networks. Some consequences of such arrangements could be observed in terms of the theoretical and evidence bases for the teaching practices of future teachers, the relevance of hierarchies and centralized knowledge for the teaching profession in spite of expectations for more localized forms of knowledge, their enculturation in certain organizational norms and the possibility of serving society.

Local hierarchies and the market of employers in the French ITE

For education in France, control lies strongly with the government. Nevertheless, the regional level *académie* has increasingly become a key player in an emerging “territorial governance” despite being “unofficial and not recognized” (Bouvier, 2012). This is particularly evident in ITE governance, where employers and networks of school professionals jointly direct university programs that exhibit many aspects of the English-style alternative/employment route. In fact, the context for public services in France has changed with the nature of the state which now plays a strategic and regulative role, distinguishing between commissions, operations and evaluation and with a significant portion of these functions being externalized.

A new territorial model based on networks is also very present in the public services domain in France, even if that is less visible and more incipient in education (Bouvier, 2012). These new realities run parallel to a hierarchical mode of governance based on “circulaires”/ministry decrees to the schools or the lower territorial levels. Decentralization went through a process of “territorialization” (Act 2004, see Bouvier, 2012, p. 67) that means an adaptation of the politics to respond to the specific needs of local communities. The territorial dimension of education is very complex with each territorial unit having a specific educational competence. The two main levels are the regions and the “departments”; while the first may promote more decentralization, the latter, given their longer history, operate more on a centralization model at the lower level. The three main decision-making levels in education are: central government, the academies and schools. “The *académie* has become the main place for regulation and coherent implementation of the education program” (Solau in Bouvier, 2012, p. 234).

The lower territorial level (*académies*) entertains a contractual relationship with the central ministerial office, based on an educational project and 4 strategic fields, following a logic of local strategic management (*project stratégique*) with the aim of transiting toward a results-based type of public management. The application of a contractual logic to the relationship between the State and the academies has been gradually initiated since 1998. A more hierarchical and bureaucratic State has been transformed into one more polycentric (Bouvier, 2012), also on the input of internationalization trends of what means being a professional (Wenzel, 2015; Malet, 2015). As suggested by Malet and Mincu (2018), at the territorial level the State acts not just as employer (*l’État-employeur*), but mostly as a State charged with the training (*l’État-formateur*) and it does so through local institutional networks and partnerships between local education authorities (*académies*), universities and schools.

The ITE system was radically transformed on the creation of the *Instituts Universitaires de Formation des Maîtres* IUFM (1990); institutions outside university system training that offer same level preparation to both primary and secondary teachers. Since 2005 there has been a trend of incorporation at master's level and integration into the university, a subsequent creation (2013) of the *École Supérieure du Professorat et de l'Éducation* (ESPE) aiming to combine academic preparation focused on practice with increased professionalization. The most recent reform (2019) was implemented with the *Institut National Supérieure du Professorat et de l'Éducation* (INSPE). Based on a 2-year training model inherited from the IUFM, the ESPE (now INSPE) training occurs in 2 phases: a first year of preparation toward the national recruitment competition, followed by pedagogical preparation. Following this, a second year leading to the master's degree. During the second year, students are paid as trainee teachers (and they have to teach part-time in a school). The global scheme is quite similar to that of the IUFM: after 3 years at university, students enter an ESPE and undergo 2 years of training. A reference table of teacher competences and a national curriculum framework for the ITE in ESPE serve as accountability systems that reinforce the professionalization side. Such training schools emphasize the role of research as well as the internationalization component, while also being tied to a university study path linked to the world of professions. It configures as a professional master's, different from other research master's options and striving to strike the right balance (Esterle, 2017).

The interesting aspect in terms of relevance in the context of ITE is the creation of an institutional arena where a variety of professionals and local representatives can jointly decide on the type of training. The INSPE council is made up of university representatives and external personalities representing the territory (at least 30%), while the president is named by the rector of the local *académie*, which is the employer in this case. An administrative council is chaired by the school director who is now directly nominated jointly by the higher education and the national education ministries. Half of the scientific council of the school includes external personalities. Consequently, the contribution of “territorial governance” and the local network of school inspectors, trainers and experts is fully recognized in the administration of this school that is incorporated into the university. The fragmentation of its internal cultures has been a limitation and has been opposed since its inception for its clear efforts to link the various contributions to the preparation of new teachers, such as “sandwich training, linking theory and practice very strongly” (Cornu, 2015, p. 304). A tradition of practice analysis seminars, as well as other initiatives were in place to link the more general pedagogical preparation (*trunc commun*) to teaching practice in schools. A significant percentage of the staff includes lecturers with many years of school experience, who are temporarily detached and contribute to the delivery of the training modules between theory and practice. Nevertheless, university based staff remains rather disconnected to the actual practice of trainees in schools (Cornu, 2015).

Specialized networks of professionals in ITE for primary school teachers in Italy

The incorporation of higher education for primary teacher preparation has been consolidated over the last 18 years. Fully integrated university courses—called “educational sciences for primary teaching” (*scienze della formazione primaria*)—have been established in more than 30 universities throughout the peninsula. Admissions numbers are established centrally by the ministry itself, based on a forecast of the need for new teachers in the 20 Italian regions. As a consequence of the Bologna process of the structural harmonization of higher education paths in the European Union, since 2010 the program has followed a unique model which includes 3 years at undergraduate level, plus 2 years of postgraduate studies.

Italian governance is built on a hybrid of devolution and centralized bureaucracy and the centralization model has been considered similar to a “harlequin type”, with significant margins of diversification (Semeraro, 1998). In fact, financial devolution implies a strong state, a weaker market, and a significant contribution from the community, especially from parents. At a justificatory level, Italian trends toward devolution and education markets, as key components of deregulation are in line with a global “disillusionment with the national modernization programs of the post-war period” and a “loss of legitimacy of the welfare state in general” (Whitty et al., 1998, 132).

Although deregulation appears to be weaker in coordinated market economies, such as Italy, in comparison with liberal market economies (Hall and Soskicze, 2001), it nevertheless triggers greater inequality in education. From the 1990s onwards, a deregulation phase brought major school autonomy reform. Influential Catholic groups endorsed liberal and neo-liberal arguments and allowed more involvement of local communities and families (Semeraro, 1998). Unquestionably, a clear hallmark of a renewed institutional context is the sweeping federalization reform, under way since 2001, with a change to the Fundamental Law (*The Fifth Title*). Linked to this, the school autonomy reform of 1997, various neoliberal initiatives, and recent fiscal federalization in education are concrete steps toward partial deregulation. The federalization of education (Bordignon and Fontana, 2010) as a process of financial and institutional decentralization at the regional level has been silently ongoing—through inaction—but as an intentional process, after years of indecision and lack of public debate (Mincu, 2015).

NPM in education has been visible since the newly established accountability system in 2011 (Law no. 10) of a National Evaluation System which aims to help schools monitor the efficiency and efficacy indicators of their educational offerings and undertake improvement in an autonomous manner; to provide school administration with useful information (Freddano and Pastore, 2020; Fiore et al., 2017) to promote cultural, social and economic growth and to introduce the evaluation of head teachers. In addition, and as a consequence of a developing accountability culture, the Good School Law introduced internal teacher assessment and the concept of performance-based pay (Mincu, 2019). The head teacher evaluates teachers and rewards them according to criteria set by each school for the first 3 years of the performance-based pay policy and thereafter on centrally agreed indicators. All schools are required to develop a self-evaluation report and a school improvement plan, as well as to

measure and assess overall performance of the administration itself, its offices and employees, with reward systems for the best performers (Law 150/2008).

In primary and pre-primary preparation at university level, an emerging accountability system of teacher education is rather formal and quantitative in nature: it specifies the number of credits, the types of areas to be covered and levels of staffing resources. In the absence of teaching standards, it does not regulate the content of teacher education but merely provides orientation toward competences that are locally identified by each primary school-teacher program. The 5-year bachelor's and master's level courses are closely affiliated to educational sciences courses and reflect an academic tradition in teaching and learning.

Teacher preparation is university-led with 300 h of practical experience during the 5-year program in schools. Trainees have the opportunity to undertake "indirect" *practicum* (lesson observation) or "direct" *practicum* (lesson delivery), but they are never fully responsible for a class during their training. Their practical preparation over the 5 years is strong on the academic side (educational sciences applied to the teaching profession) but weak on the number of subjects to be covered on all the 5 school years they are required to teach. Three types of training figures are responsible for their preparation: the university professors (who are on a full academic track and do not possess direct knowledge of the school world), contract university professors (temporary academic professionals, some with school experience) in charge of workshops and university tutors (teachers working temporarily for the university) in charge of their practical experience and responsible for a group of students, for or periodical meetings at the university and school visits. A fourth figure is the assigned teacher inside the school of practice, who agrees to be a local supervisor.

Collaboration between the university-based professionals is rather weak. The contract professors charged with the workshops that are selected by the university professors themselves and the curricular offer can be jointly agreed. The profile of the knowledge offered is theory-oriented, and research-based with little input from the practice in schools. The curricular autonomy is wide either for courses or for workshops, since the offer is weakly oriented toward a list of teacher competences. Agreed standards for the teaching profession are not as yet in place. The links with the university tutors (former school teachers) are rather casual. **The contribution of the four networks of professionals is delivered in separation rather than collaboration. The networks work inside a shared culture that is derived from their professional profile and duties, but there are no formal opportunities to share their practices and understandings. The university environment is largely based on academic autonomy, which is the main criteria to decide and deliver the curriculum and then to assess the students' progress.** However, the team of university tutors is an exception. They are usually coordinated by a lead tutor, and the tutoring office decides on common criteria of action. Nevertheless, the tutors are trained within a similarly very strong culture of teacher autonomy inside schools and so their practices may be weakly aligned when acting in their tutorship role.

To sum up, the preparation of the new teachers at this level is not influenced by market forces, since the admission numbers and procedures are centrally established. The school and the university hierarchies are rather weak, since accountability mechanisms are very new devices. One major type of contextual force is clearly visible: the specialized networks of professionals operating with wide margins of individual autonomy and weak coordination inside and between such networks. It has to be mentioned that the preparation of secondary school teachers is highly unstable and therefore the overall picture of the Italian TE leads to fragmented professionalism (Mincu, 2018).

Multilevel hierarchies and institutional networks in Chinese CPD

As showed by Han and Ye (2017), Chinese education policy-making has changed significantly over recent decades: from Party-dominant to primarily driven by central government, an enhanced role for higher education institutions (HEIs), scholars as "professional interest group" and the inclusion of non-governmental actors in the policy-making process. The transition from a highly hierarchical model of governance to one that is more inclusive and diversified with the presence of networks has evolved in 4 phases: (1) pre-1978: top-down state-directed (2) 1978–1987: transition from state-directed to policy community (3) 1988 to 1997: the evolution of a policy community (4) 1998–2010: the emergence of issue networks (Han and Ye, 2017). At the same time, in-service teacher education in China has undergone 4 phases of substantial transformation: (1) "recovery and reconstruction", (2) "reform and expansion", (3) "comprehensive transformation" (4) "coordinated development" (Liu, 2020).

Over the last 20 years the context for CPD has changed through a series of policy measures; mainly the expansion of teacher education by institutional diversification and structural changes through upgrading, merging and closing institutions and schools. At the same time, considerable efforts have been put into the innovations in the modes of teacher preparation, curricula and teaching practicum. The "transformation" of in-service teacher education has implied 3 types of transition: (1) academic compensation education to continuing education (2) normal teacher education in a diversity of institutions (3) a separate TE to integrate pre-service and in-service teaching.

The methods toward improving the quality of training has changed over time. After the introduction of the new curriculum in 2001, the Ministry of Education provided top-down mandated training for all teachers on the new curriculum. Hence, a cascade model was adopted (Paine and Fang, 2006), in which the central and provincial governments ran national and provincial programs for teacher trainers, in order to train other teachers across the country (Zhou, 2014). Training activities included listening to the introduction of government policy and regulations and lectures provided by outside experts, in which the content was theoretical

or even abstract without any practical application (Feng, 2003). As Zhou (2014) says “training was regarded as ‘telling’ so teacher learning was considered as ‘being told’ and the teachers were not empowered” (p. 511).

With the launch of the “National Training Program” CPD has entered a new phase of “cooperative development” right from the top, down. *The Teacher Education Revitalization Action Plan* (2018–2022) was issued by the Ministry of Education, the National Development and Reform Commission, the Ministry of Finance, the Ministry of Human Resources and Social Security, and the Central Organization. This policy document asked to innovate the teacher training model and establish a new mechanism for the joint training of teachers between higher education institutions, local governments, primary and secondary schools (kindergartens, vocational schools). The *Teacher Education Revitalization Action Plan* (2018–2022) further emphasizes the importance of building a collaborative teacher education system, based on normal universities, high-level comprehensive universities, teacher development institutions, and high-quality primary and secondary schools (Liu, 2020).

In recent years, various projects have all embodied the concept of **collaborative development** of teachers, such as “sending to the countryside”, “backbone teacher studio”, “teach and research in joint school district”, “online and offline mixed training” and “master of education”. Such new forms of teacher education have sprouted mostly in urban centers. For example, teachers can dispose of “learning communities” that are organized in school alliances, while designated “master-teachers” organize their own “master work studios” to train future teachers (Zheng et al., 2019). This new collaborative model was preceded in the early 80s by a “three components communication” model of learning promoted by the Open University, the Evening University and centered on adult self-study. This has developed into network training and mixed learning modalities. Other forms of collaboration include **urban—rural partnerships** to improve the quality of education. From the mid-1980s, education lecturers supported education and training in remote mountain villages. A learning community of urban and rural teachers has been built around the idea of partnership: senior students are sent to schools in rural areas for one semester to replace selected school teachers who then go to the colleges for training. Other forms of partnerships that changed the CPD landscapes are built around **University–Government–School partnership**, such as in one initiated by the Northeast Normal University (Jilin Province) in 2007.

These changes aim to “flexibly open up” the system of TE through diversification of its course providers. It also aims to foster competition in a system that was unable to satisfy the educational demands of an emergent knowledge economy. At the same time, the TE scenario has been changed using **accountability measures**, such as the introduction of standards in TE and the inspection of providers. New notions of professionalism are visible in the Teacher Education Curriculum Standards, that promoted learner-centered, practice-oriented and lifelong teacher learning.

Looking into the relationship between hierarchies, markets and networks, the Chinese situation is particularly complex. First, the arrangement of the central power administration is dual-faceted, given the coexistence of the Party and the government, with the latter holding the reins of power in education policy making. (Han and Ye, 2017). At the same time, even if the dominant decision-making power rests ultimately in the hands of the State, the wider context has greatly changed over time. The government cooperates with professional academic groups when undergoing reform (Han and Ye, 2017) and the wider public can be, at times, even be greatly involved, as was the case with *The outline of China’s national plan for medium and long-term education reform and development (2010–2020)*.

In China, there is a trend of power fragmentation and diversification that coexists with an **elite hierarchical power structure**. The unprecedented degree of public participation in policy matters has made education less of a political issue, and more a social concern that requires the support of scientific research.

Discussion and conclusion

The focus of this comparison is the role of the context in various types of training and its relationship to various professionalism aspects. As illustrative examples this article has considered, alternative routes in England, a HE-led model of ITE in France based on partnership with the territorial system/employer, a primary school ITE path that is HE-led in Italy, as well as the complex collaboration inside and outside hierarchies in China to deliver CPD.

One major driving force in all contexts is the NPM in education even if that assumes various manifestations and combinations of the three contextual mechanisms. While the contracting out mechanisms of the State are more visible in England and France, the “State controlling at a distance” model through accountability mechanisms is the preferred way in both China and Italy; the Italian ITE being rather autonomous and HE led, with more formal accountability mechanisms in place (see Mincu, 2019).

An inter-disciplinary “interactive governance” perspective recognizes the interactions of governance through hierarchy, markets and networks, while being sensitive to the problems and solutions offered by each combination. Advocacy of network governance for ITE in England (e.g., Hill, 2011; Matthews and Berwick, 2013) has neglected the implications of pressures from hierarchical governance (from central government) and market governance (through the supply of teachers and competition with other schools). Successive policy shifts in England have encouraged network governance within the context of market governance while also retaining hierarchical governance by government. Interactions between these varying forms of governance are inevitable. But when policy ignores these interactions it is bound to create problems. However valiant and well-intentioned schools may be, they are being asked to deal with a set of contradictory pressures. In more polycentric systems there can be various “hierarchical

arrangements” in mixed network form and this seems to be the case with the new territorial governance in France, more generally speaking, and in particular in the area of teacher preparation. The relevance of local representatives on the various committees is indicative of a model of strong networks inside hierarchies. The Italian case of primary ITE is rather singular in its simplicity. In the Italian case, networks of professionals deliver highly academic training. The Ministry guidelines about the contents are rather loose and not informed by professional teaching standards. These refer mainly to the time to be assigned to courses, workshops and practicum and roughly indicates the academic research areas that can contribute in full local autonomy. A HE hierarchy leads this model. The Chinese context of CPD is characterized by strong hierarchies that are adopting a network logic with the academic world and widely privileges partnerships and diversification.

Which profile takes the interactive governance model of the contextual forces in the four countries? Describing the context in terms of types of governance and relationship to professionalism does not have any implications on the quality of the TE. In addition, the characterization of various public systems can be very different and the situation of education can be in line with what happens in other areas or a singularity, a case per se. This can be said about China’s more networked governance, in spite of a more general hierarchical pattern. In France the transformation of the state is rather diffused but somehow unrecognized, because the current discourse continues to emphasize traditional centralization under strong State control.

England is the clearest case in point for very strong networks as major contextual factor. These manifest in solid relationships with hierarchies at school level and system level (notably with the Office for Standards in Education—Ofsted), in a market oriented ITE model, in a country with a wide diversification of training routes and clear preferences for employment-based routes. France draws upon a local partnership between HE and academies and a more territorial hierarchy that is active in its role as employer of teachers. Networks of professionals manifest themselves inside the training institutions. The tensions in France are a sign of a more successful opening up of the academic world to the school networks of professionals compared to Italy. French universities propose a short route model and privilege a hands-on and a “teaching in responsibility” model, much more similar to the employment based-routes in England. The Italian ITE shows that only HE hierarchies can have a say and the network of academics inside it are key players. HE has led teacher training on a very autonomous model path, in a context in which hierarchies and accountability mechanisms are weak and the market logic irrelevant, since numbers are centrally determined, as are the providers. China’s CPD displays collaboration between multi-level hierarchies, schools included. The central authorities are key players but the model overall has been pluralized and diversified from top to bottom with an extension of the partnerships between urban and rural areas of the country.

But what is the relevance of the context and how does it contribute to teacher professionalism in the four cases? A first consideration is that assessing this link does not imply any judgment about the quality of the training. The proposed model of professionalism as promoted by the context identified 4 models and 4 dimension (see [Table 2](#)). The fourth model represents rather an ideal type. Each country may exemplify some of its dimensions but no one can typify it. The English case of ITE appears as a perfect case of managerial professionalism across all the 4 dimensions with a transition toward more collaboration (as inter-agency). More generally, China moves from a universal and state welfare model to one which is market-based ([Guo et al., 2013](#)). At the same time, its definitively collaborative type of professionalism is rather mandated ([Ke et al., 2019](#)). Thus, this case represents a partial collaborative professionalism model across the three dimensions with a mix of managerial when it comes to collaboration, which sometimes appears as rather contrived. The French case is a mix of HE-led model as in extended professionalism but low teacher collaboration nevertheless, and relevant accountability as standards oriented TE. The importance of employment and the relevance of the territorial level renders it as a rather imperfect model of managerialism. The Italian case of primary teacher ITE is a partially extended professionalism model, because of the lack of collaboration inside schools and its strongly HE leadership.

This analysis helps to associate the various types of professionalisms with governance combinations, drawing on the empirical evidence provided derived from four country-cases. The ideal types of professionalism can be characterized in terms of governance as indicated in [Table 2](#). Then, it is possible to note how the ITE and CPD contexts, in relation to the type of professionalism, can be further interpreted at the interplay of knowledge, values, self-direction and collaboration.

Table 2 Forms of professionalism and contextual arrangements in 4 countries.

	<i>Extended professionalism HE hierarchies Networks within schools</i>	<i>Managerial professionalism (NPM responding to the public choice critique) markets, school/system hierarchies</i>	<i>Collaborative professionalism (variant of NPM) Networks between institutions & professionals</i>	<i>Occupational/activist/democratic professionalism</i>
Knowledge	Italy, France	France, England	China	England, China
Values assumed	Italy	England	China, France	Italy
Self-direction	Italy	France	China	Italy, France
Collaboration		China, England	China, England	England

References

- Ball, S.J., Exley, S., 2010. Making policy with "good ideas": policy networks and the "intellectuals" of new labor. *J. Educ. Pol.* 25 (2), 151–169.
- Ball, S.J., Junemann, C., 2012. *Networks, New Governance and Education*. Policy Press, Bristol.
- Ball, S.J., 2009. Beyond networks? A brief response to "which networks matter in education governance?". *Polit. Stud.* 57, 688–691.
- Bordignon, M., Fontana, A., 2010. *Federalismo e istruzione. La scuola italiana nell'ambito del processo di decentramento istituzionale* [Federalism and education: the Italian school and institutional decentralisation]. Working Paper, Fondazione Giovanni Agnelli, Torino.
- Bouvier, A., 2012. *La gouvernance des systèmes éducatifs*. PUF, Paris.
- Clarke, J., Gewirtz, S., McLaughlin, E., 2000. *New Managerialism, New Welfare?* Open University & SAGE, London.
- Cochran-Smith, M., Fries, K., 2009. The AERA panel on research and teacher education: context and goals. In: Cochran-Smith, M., Zeichner, K. (Eds.), *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education*. Lawrence Erlbaum, Mahwah.
- Cochran-Smith, M., Lytle, S., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York.
- Cornu, B., 2015. Teacher education in France: universitization and professionalization—from IUFMs to ESPes. *Educ. Inq.* 6 (3), 289–307.
- Davies, P., Connolly, M., Nelson, J., Hulme, M., Kirkman, J., Greenway, C., 2016. "Letting the right one in": provider contexts for recruitment to initial teacher education in the United Kingdom. *Teach. Teach. Educ.* 60, 291–302.
- Ehren, M.C.M., Janssens, F.J.G., Brown, M., McNamara, G., O'Hara, J., Shevlin, P., 2017. Evaluation and decentralized governance: examples of inspections in polycentric education systems. *J. Educ. Change* 18 (3), 365–383.
- Ellis, V., Steadman, S., Trippstad, T.A., 2019. Teacher education and the GERM: policy entrepreneurship, disruptive innovation and the rhetorics of reform. *Educ. Rev.* 71 (1), 101–121.
- Esping Andersen, G., 1990. *The Three Worlds of Welfare Capitalism*. Princeton University Press, Princeton.
- Esterle, M., 2017. Où va la formation des enseignants? *Chronique d'un passage tourmenté*. Pétra, Paris.
- European Commission, 2012. *Supporting the Teaching Professions for Better Learning Outcomes*. Author, Brussels.
- Evetts, J., 2009. New professionalism and new public management: changes, continuities and consequences. *Comp. Sociol.* 8 (2), 247–266.
- Feng, C., 2003. Where does the teacher training for new curriculum go? *Basic Educ. Res.* 16 (7–8), 47–48.
- Ferrera, M., 1996. The "southern model" of welfare in social Europe. *J. Eur. Soc. Pol.* 6 (1), 17–37.
- Fiore, B., Poliandri, D., Romiti, S., Giampietro, L., 2017. Il feedback di INVALSI nei Piani di Miglioramento V&M, Scuola democratica. *Learn. Democr.* 8 (1), 141–164.
- Freddano, M., Pastore, S., 2020. Il sistema nazionale di valutazione e l'uso dei risultati delle Rilevazioni nazionali INVALSI: quali cambiamenti? In: Falzetti, P. (Ed.), *Il dato e il miglioramento delle scuole. Il seminario "I dati INVALSI: uno strumento per la ricerca"*. Franco Angeli, Milano, pp. 9–32.
- Fuller, B., Dellagnelo, L., Strath, A., Santana Barretto Bastos, E., Holanda Maia, M., Lopes de Matos, K.S., et al., 1999. How to raise children's early literacy? The influence of family, teacher, and classroom in northeast Brazil. *Comp. Educ. Rev.* 43 (1), 1–35.
- Ginsburg, M., Lindsay, B. (Eds.), 1995. *The Political Dimension in Teacher Education: Comparative Perspectives on Policy Formation, Socialization and Society*. Falmer Press, Bristol, PA.
- Goldhaber, D., 2018. Evidence-based teacher preparation: policy context and what we know. *J. Teach. Educ.* <https://doi.org/10.1177/0022487118800712>.
- Gove, M., 2010. Speech at the National College Annual Conference, Birmingham. <https://www.gov.uk/government/speeches/michael-gove-to-the-national-college-annual-conference-birmingham>. (Accessed 25 November 2010).
- Grimaldi, E., 2011. Governance and heterarchy in education: enacting networks for school innovation. *Ital. J. Sociol. Educ.* 2 (1), 114–149.
- Guerriero, S., Deligianni, K., 2017. The teaching profession and its knowledge base. In: Guerriero, S. (Ed.), *Pedagogical Knowledge and the Changing Nature of the Teaching Profession*. OECD, Paris.
- Guo, S., Guo, Y., Beckett, G., Li, Q., Guo, L., 2013. Changes in Chinese education under globalization and market economy: emerging issues and debates. *Compare* 43 (2), 244–264.
- Hall, P.A., Soskicze, D., 2001. An introduction to varieties of capitalism. In: Hall, P.A., Soskicze, D. (Eds.), *Varieties of Capitalism. The Institutional Foundations of Comparative Advantage*. Oxford University Press, Oxford, pp. 1–68.
- Han, S., Ye, F., 2017. China's education policy making: a policy network perspective. *J. Educ. Pol.* 32 (4), 389–413.
- Hill, R., 2011. *The Importance of Teaching and the Role of System Leadership*. National College for School Leadership, Nottingham.
- Hoyle, E., 1974. *Professionalism, Professionalism and Control in Teaching*. Falmer Press, London.
- Judge, H., Lemosse, M., Paine, L., Sedlak, M., 1994. *The University and the Teachers* (Oxford Studies in Comparative Education). Symposium Books, Oxford, UK.
- Ke, Z., Yin, H., Huang, S., 2019. Teacher participation in school-based professional development in China: does it matter for teacher efficacy and teaching strategies? *Teach. Teach.* 25 (7), 821–836.
- Liu, M., 2020. *Retrospect and Prospect of In-Service Teacher Education in China*. Beijing Normal University, Beijing.
- Mak, G.C.L., 1999. Classroom management in China. In: Shimahara, N. (Ed.), *Politics of Classroom Life*. Garland, New York, pp. 239–260.
- Malet, R., Mincu, M., 2018. Quelle professionnalisation des enseignants à l'heure de la nouvelle gouvernance publique? Perspectives comparatistes européennes et franco-italiennes. In: Des Jardins, J., Dejemeppe, X., Guibert, P., Maulini, O. (Eds.), *Questionner et valoriser le métier d'enseignant. De Boeck Supérieur, Bruxelles*.
- Malet, R., 2015. La diffusion du modèle de professionnalisation et ses variations. In: Wenzel, B., Lussi Borer, V., Malet, R. (Eds.), *Professionnalisation de l'enseignement. Fondements et retraductions*. Nancy: PUN.
- Matthews, P., Berwick, G., 2013. *Teaching Schools. First Among Equals?* National College of Teaching Leadership, London.
- Mayer, D., Mills, M., 2021. Professionalism and teacher education in Australia and England. *Eur. J. Teach. Educ.* 44 (1), 45–61.
- Mayer, D., 2006. The changing face of the Australian teaching profession: new generations and new ways of working and learning. *Asia Pac. J. Teach. Educ.* 34 (1), 57–71.
- Meyer, J.W., Ramirez, F.O., 2000. The world institutionalization of education. In: Schriewer, J. (Ed.), *Discourse Formation in Comparative Education*. Peter Lang Publishers, Frankfurt am Main.
- Mincu, M., Davies, P., 2019. The governance of a school network and implications for initial teacher education. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2019.1645360>.
- Mincu, M., 2015. The Italian middle school in a deregulation era: modernity through path-dependency and global models. *Comp. Educ.* 51 (3), 446–462.
- Mincu, M., 2018. Overcoming fragmented professionalism? Accountability for improvement in teacher preparation in Italy. In: M Carvalho, L., Levasseur, L., Liu, M., Normand, R., Oliveira, D.A. (Eds.), *Education Policies and the Restructuring of the Educational Profession*. Springer.
- Mincu, 2019. Learning to teach in Italy: reviewing policy and research trends. In: Tatto, T., Menter, I. (Eds.), *Knowledge, Policy and Practice in Teacher Education*. Bloomsbury.
- Munoz-Izquierdo, C., 1994. Determinantes de las desigualdades educacionales con énfasis en los grupos de menor desarrollo socioeconómico [Determinants of Educational Inequality With Emphasis on Economically Underdeveloped Groups]. Cuadernillos de la Educación, 3. Instituto Estatal de la Educación Pública de Oaxaca, Oaxaca, Mexico.
- Musset, P., 2010. Initial Teacher Education and Continuing Training Policies in a Comparative Perspective. OECD Education Working Papers, 48. OECD Publishing, Paris.
- Mutton, T., Burn, K., Menter, I., 2017. Deconstructing the carter review: competing conceptions of quality in England's "school-led" system of initial teacher education. *J. Educ. Pol.* 32 (1), 14–33.
- Oosterheert, I.E., Vermunt, J.D., Denessen, E., 2002. Assessing orientations to learning to teach. *Br. J. Educ. Psychol.* 72, 41–64.
- Osborn, M.J., McNess, E.M., 2005. The cultural context of teachers' work: policy, practice and performance. In: Bascia, N., Cumming, A., Datnow, A., Leithwood, K., Livingstone, D. (Eds.), *International Handbook of Educational Policy*, vol. 2. Springer, pp. 507–525.

- Paine, L.W., Fang, Y., 2006. Reform as hybrid model of teaching and teacher development in China. *Int. J. Educ. Res.* 45 (4–5), 279–289.
- Rhodes, R.A.W., 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability*. Open University Press, London.
- Robinson, J., 1994. Social status and academic success in South Korea. *Comp. Educ. Rev.* 38 (4), 506–530.
- Sachs, J., 2001. Teacher professional identity: competing discourses, competing outcomes. *J. Educ. Pol.* 16 (2), 149–161.
- Semeraro, A., 1998. *Il Sistema Scolastico Italiano. Profilo Storico* [The Italian System of Education. A Historical Perspective]. Carocci, Roma.
- Sing Chai, C., Teo, T., Beng Lee, C., 2009. The change in epistemological beliefs and beliefs about teaching and learning: a study among pre-service teachers. *Asia Pac. J. Teach. Educ.* 37 (4), 351–362.
- Smith, K., McLay, M., 2007. Curates' eggs? Secondary trainee teachers' experience of the graduate teacher program and the postgraduate certificate in education. *J. Educ. Teach.* 33 (1), 35–54.
- Steiner-Khamisi, G., 2005. Vouchers for teacher education (non)reform in Mongolia: transitional, postsocialist or antisocialist explanations? *Comp. Educ. Rev.* 49 (2), 148–172.
- Tatto, M.T., Mincu, M. (Eds.), 2009. *Reforming Teaching and Learning: Comparative Perspectives in a Global Era*. Sense Publishers & WCCES, Rotterdam.
- Tatto, T., 2008. Teacher policy. A framework for comparative analysis. *Prospects* 38, 487–508.
- Theisens, H., 2016. Hierarchies, networks and improvisations in educational governance. In: Burns, T., Köster, F. (Eds.), *Governing Education in a Complex World*. OECD Publishing, Paris, pp. 55–69.
- Thomas, G., 2011. A typology for the case study in social science following a review of definition, discourse, and structure. *Qual. Inq.* 17 (6), 511–521.
- Thompson, G.F., 2003. *Between Hierarchies and Markets: The Logic and Limits of Network Forms of Organization*. Oxford University Press, Oxford.
- Timperley, H., et al., 2007. *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration*. Ministry of Education, Wellington, New Zealand.
- Torfling, J., Peters, G.J., Pierre, J., Sørensen, E., 2012. *Interactive Governance: Advancing the Paradigm*. Oxford University Press, Oxford/New York.
- Wenzel, B., 2015. Internationalization de la professionnalisation des enseignants: éléments d'analyses et de synthèse. In: Wenzel, B., Lussi Borer, V., Malet, R. (Eds.), *Professionnalisation de l'enseignement. Fondaments et retraductions*. Nancy: PUN.
- Whitty, G., 2008. Changing modes of teacher professionalism: traditional, managerial, collaborative and democratic. In: Cunningham, B. (Ed.), *Exploring Professionalism*. Bedford Way Paper. Institute of Education, London.
- Whitty, G., Power, S., Halpin, D., 1998. *Devolution and choice in education: the school, the state and the market*. Open University Press, Buckingham.
- Windzio, M., Sackmann, R., Martens, K., 2005. Types of Governance in Education: A Quantitative Analysis. *TranState Working Papers*, No. 25, Collaborative Research Center 597—Transformations of the State, Bremen. <https://www.econstor.eu/bitstream/10419/28275/1/501321926.PDF>.
- Zeichner, K., Pena-Sandoval, C., 2015. Venture philanthropy and teacher education policy in the US: the role of the new schools venture fund. *Teach. Coll. Rec.* 117 (6), 1–44.
- Zhang, W., Bray, M., 2017. Micro-neoliberalism in China: public-private interactions at the confluence of mainstream and shadow education. *J. Educ. Pol.* 32 (1), 63–81.
- Zhao, Y., 2018. The changing context of teaching and implications for teacher education. *Peabody J. Educ.* 93 (3), 295–308.
- Zheng, X., Zhang, J., Wang, W., 2019. Teacher learning as boundary crossing: a case study of Master Teacher Studios in China. *Teach. Teach.* 25 (7), 837–854.
- Zhou, J., 2014. Teacher education changes in China: 1974–2014. *J. Educ. Teach.* 40 (5), 507–523.

Can initial teacher education partnerships maintain their epistemological and pedagogical underpinnings in light of market-driven reforms? The case of England

Trevor Mutton^a, Katharine Burn^a, and Hazel Hagger^b, ^a University of Oxford, Department of Education, Oxford, United Kingdom; and ^b Kellogg College, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	83
Policy background	83
Understanding the nature of ITE partnerships	84
Learning to teach within ITE partnerships and the role of clinical practice	84
Challenges inherent within partnership working	85
Partners may not necessarily share the same conceptualization of what constitutes effective teacher preparation	85
There may be tensions between the differing goals of preparing a teacher to be “school-ready” or “profession-ready”	86
Partners have differing priorities and differing levels of accountability	86
Schools often engage in complex partnership arrangements with different providers	87
The implications for partnership in ITE as an “epistemological and pedagogical concept”	87
Principles to inform ITE partnership working	88
Conclusions	90
References	91

Introduction

This chapter arises from an in-depth case study of the way in which initial teacher education (ITE) partnership working has evolved in light of significant reforms to teacher education policy over the past decade in England. The case study focuses on a specific school (Waterside Academy – a pseudonym) that works within a well-established university-school partnership, yet which has, at the same time, become a government-accredited SCITT provider in its own right. SCITT (School-Centred Initial Teacher Training) providers offer postgraduate teacher training programs that lead to Qualified Teacher Status (QTS) with some SCITT providers also linking with a university to offer a Postgraduate Certificate in Education (PGCE). The research from which the case study is drawn was used originally as the source material for a guide produced to inform teacher educators’ understanding of partnership working (Mutton et al., 2018). Writing from our own perspective as teacher educators within one of the earliest ITE partnership models in England – the Oxford Internship Scheme (Benton, 1990) – we here examine how policy change in the past decade, and enactment of that policy, have led to changes in the way that partnership has been conceptualized.

Working from the premise that the term “partnership”, in the context of ITE, is used “to carry certain theories about the nature of learning to become a teacher” (Brisard et al., 2005, p. 5), the chapter illustrates the way in which such theories encompass an understanding not just of what beginning teachers need to know and be able to do, but also of the processes by which their professional learning takes place (Burn et al., 2015). While the importance of partnership working has long been recognized in official government policy in England, many well-established partnerships between higher education institutions and schools could be characterized as essentially university-led (Furlong et al., 2000), with insufficient integration between the university-based and school-based parts of the program. Focusing on the processes by which beginning teachers develop the knowledge and skills that they require in order to become accomplished teachers, we highlight the way in which the potential lack of a shared understanding as to how these processes might best be enacted will inevitably lead to tensions – in relation both to the organizational aspects of partnership and to pedagogical concerns. We conclude by offering some key principles which, we suggest, might usefully inform the way in which ITE partnerships operate.

Policy background

The “school-led” policy drive in England, launched by the government White Paper “The Importance of Teaching” (DFE, 2010) and developed in subsequent policy documents (DFE, 2011a, 2011b) attempted to re-align the balance of power in teacher education by setting up structures to give schools more responsibility for the recruitment and training of new teachers. While the language used was still that of partnership, the assumptions that underpinned this drive were, however, clearly encapsulated in the Secretary of State’s conception of teaching as a “craft” (Gove, 2010). Such a conception underpinned the subsequent development of “School Direct” and the expansion of School-Centred Initial Teacher Training (SCITT) provision, essentially prioritizing an apprenticeship model of professional learning and has been the predominant influence on teacher education policy over the past ten years (for a

more detailed account of policy over this period see [Menter et al., 2019](#)). This policy drive has arguably made it extremely difficult for partnerships that rest on alternative epistemological and pedagogical underpinnings to maintain an approach that both respects and seeks to integrate the distinctive contributions of school- and university-based teacher educators. It could equally be argued, however, that an increasing emphasis on the role of school-based teacher education has, in fact, led to the development of more collaborative approaches to partnership working in England over the past decade, as schools engaged in the new School Direct models negotiated their relationship with university partners. Partnership working was seen to be at the heart of these reforms and the government gave a clear signal that it wanted to “encourage more universities to follow the example of the integrated working of the best university-school partnerships” (DFE, 2011a,b,c, p. 11). Integration is seen as being key to the new model in order to deal with the perceived problem of schools within traditional university-led partnerships ([Furlong et al., 2000](#)) being somewhat marginalized in terms of their contribution to the conceptualization, planning, delivery and evaluation of the individual ITE program. Partnership working was to emulate what were considered to be more collaborative models and, in order to ensure that partnerships responded in an appropriate way, the government indicated that it would ... also change the inspection of ITT so that the quality of partnership working is treated as more important. (DFE, 2011b: para. 15).

It should be noted that ITT (initial teacher training) and ITE (initial teacher education) are both used to describe the preservice period of teacher preparation. ITT is the term used by the government in England. We use the term ITE (preferred by many university providers) unless referring to government publications.

Understanding the nature of ITE partnerships

The question as to what constitutes quality in ITE partnerships is partly addressed through consideration of different models of partnership. The Modes of Teacher Education (MOTE) project ([Furlong et al., 1996, 2000](#)), carried out in two phases in order to study the impact of earlier government policies on models of teacher education in England at the time, identified three predominant models: collaborative partnership; HEI (Higher Education Institution) led partnership; and separatist (or complementary) partnership.

Collaborative partnerships embody a “commitment to develop a training program where students are exposed to different forms of educational knowledge, some of which come from school, some of which come from HE or elsewhere” ([Furlong et al., 1996](#), p. 44). At the time, ITE providers were predominantly within the Higher Education (HE) sector and the authors therefore identify how school-based and university-based colleagues might work together to plan and deliver the ITE curriculum in such a way as to enable student teachers to critique what they learn from both settings and are thereby able to “build up their own body of professional knowledge” (1996, p. 44).

The HEI-led model sees the school’s role as acting as a resource to provide learning opportunities for its student teachers within a university-designed ITE curriculum. Within the model there is a strong emphasis on quality assurance to ensure that student teachers receive comparable training opportunities.

The separatist, or complementary model is one “where school and HE are seen as having separate and complementary responsibilities but where there is no systematic attempt to bring these two dimensions into dialogue” (1996, p. 47). Program integration is not a necessary element within such a model and it is up to the individual student to achieve any such integration.

The MOTE researchers identified that any single ITE program might embody aspects of each model, leading to a “continuum of possible relationships” ([Furlong et al., 2000](#), p. 77), although they also found that the HEI-led model was predominant, in spite of frequent rhetoric referring to collaborative partnerships. All three models carry with them, however, some sense of partnership operating as an “epistemological and pedagogical concept” ([Furlong et al., 2008](#), p. 309); that is to say an acceptance that preservice teachers require access to different forms of knowledge which neither the university nor the school can exclusively provide ([Hagger and McIntyre, 2006](#)). It is, however, only within a truly collaborative model that full integration is likely to take place. While there is almost universal agreement that effective partnership working is the key to success of any ITE program, it also needs to be acknowledged that the “rhetorical power of partnership” ([Conroy et al., 2013](#), p. 558) can hide more general concerns about the way in which teacher education pedagogy plays out in practice across different teacher education settings.

Learning to teach within ITE partnerships and the role of clinical practice

Such concerns are not unique to the English ITE context, where there has been an emphasis on promoting not just quality but also quantity (that is to say, securing the engagement of enough schools as partners in ITE programs in order to maintain the supply of new entrants to the profession). Elsewhere, there has been an increasing trend toward more school experience within programs of preparation for preservice teachers, a trend often referred to as the “practicum turn in teacher education” ([Mattsson et al., 2011](#), p. 17). Underpinning such initiatives has been a need to address issues of integration, or coherence within programs, which [Hammerness \(2006\)](#) defines as “the alignment of key ideas and goals across coursework and clinical work” (2006, p. 1244). Internationally, partnership development has been a key feature of initiatives such as Professional Development Schools in the United States,

which identified as central to their success the need for coherence between the taught program at the university and the school (i.e., clinical) experience (Darling-Hammond, 2006). Similar clinical practice models, which emphasize integration through effective university-school partnerships, have been developed elsewhere internationally, for example in Australia (see McLean Davies et al., 2013 for details), in the Netherlands through the introduction of *academische opleidingsscholen* (Hammerness et al., 2012), and in Finland, through designated training schools (Sahlberg, 2012).

It needs to be understood, however, that the term “clinical practice” designates not only the school-based elements of any ITE program but also underpins a particular approach to teacher education partnership working within a clinical practice model. Alter and Cogshall (2009) summarize the characteristics of all clinical practice professions, including teaching, as being: the “centrality of the clients” (in this case, school pupils); a complex professional knowledge base; the use of evidence and judgment (rather than pure technical skill); a community of practice operating with shared standards; and “rigorous academic and practical training” (2009, p. 9). Beginning teachers need, therefore, to have access to appropriate learning opportunities that allow them to develop as full professionals within such a framework, in order to enable them both to hone the practical skills needed as a teacher but also to enable them “to engage in the creative processes of interpretation, intervention and evaluation, drawing on diverse sources of knowledge that include research evidence as well as student data” (Burn and Mutton, 2015, p. 219).

Returning to the policy landscape in England, there is some limited evidence to suggest that the increased involvement of schools in ITE programs as a result of the implementation of the 2010 government reforms has led to greater collaboration around program design and delivery (Universities UK, 2014). A study of the impact of “School Direct” in England (Brown et al., 2015) concludes, however, that the balance of power within partnership arrangements has altered and that “(t)he composition of trainee pedagogical experience is being reconfigured” (2015, p. 23) because of differing beliefs about the nature of the learning experiences that beginning teachers may need. The authors of the report argue that the reforms have not led to greater levels of integration within ITE partnerships but have, instead, resulted in a situation where:

trainees were having to assimilate quite different understandings of pedagogy between school-based and university-based training sessions whilst efforts to mesh what was being taught within each setting was often limited. It was not uncommon for trainees to feel that content had been repeated by the school and/or university. (2015, p. 23).

There are clearly challenges to ITE partnership working within the current policy context in England but many of these challenges are not new. Almost 40 years ago, immediately prior to a previous significant wave of teacher education reform in England, Alexander (1984) was warning that the “comfortable language of partnership” that policy-makers and others so easily use, could conceal “more intractable issues” (1984, p. 142).

Challenges inherent within partnership working

From our experience as teacher educators and drawing on what research tells us, we have tried to identify what these “intractable issues” might be – both in relation to partnership working generally and, more specifically, in light of the current ITE policy context in England. We identify four such challenges:

Partners may not necessarily share the same conceptualization of what constitutes effective teacher preparation

It cannot be assumed that there will necessarily be shared understandings within an ITE partnership in terms of either the knowledge base that beginning teachers require, or the learning processes by which they acquire professional knowledge and skills to enable them to teach effectively, or both. Within what might be seen as more traditional programs there has always been the likelihood of a potential “disconnect between the campus and school-based components of programs” (Zeichner, 2010, p. 89). The introduction of “School Direct” has required schools in England to take a more dominant role in partnership working, but may not have reduced such a “disconnect” since it does not necessarily guarantee agreement between partners as to the most effective way of preparing new teachers, nor led to greater innovation within partnership working. What it has done, however, is place more emphasis on the “practical components” of ITE programs and less emphasis on “theory and analysis” (Brown et al., 2015, p. 5).

There is also an argument that effective ITE needs an agreement between partners first as to what constitutes good teaching and then, in turn, how to prepare beginning teachers to teach in this way. Where there is no “common, clear vision of good teaching that permeates all coursework and clinical experiences, creating a coherent set of learning experiences” (Darling-Hammond, 2014, p. 548) then program goals risk being compromised (Bartholomew and Sandholtz, 2009) and teacher learning fragmented (Taylor, 2008).

The evidence from our case-study of Waterside Academy would suggest that there is a high level of consistency of views within the SCITT program as to what constitutes effective ITE. The ethos underpinning the program is pragmatic and reflects the program goals of producing teachers who will work in the specific local contexts in which they train. There is agreement among those working on the program that it needs to develop independence, resilience, initiative and a sense of realism in the teachers it is training. For

those working at the same school within the university PGCE partnership there is, however, more evidence of a disconnect between the program goals and the views of the school-based teacher educators. Those goals are focused not only on developing the beginning teachers' classroom competence but also on developing their capacity to evaluate their own teaching critically. As part of this evaluation they need to be able to recognize the provisional nature of the suggestions for practice that they're offered and also appreciate that these may be open to critique and challenge in relation to different criteria. The PGCE program aims to equip them not only with the skills they need to teach but also with the capacity to go on learning as professionals by taking increased responsibility for identifying and meeting their own developing professional needs. While acknowledging the need for beginning teachers to draw on a range of sources (both theoretical and practical) to inform their developing thinking and practice, there is nevertheless a sympathy among school-based colleagues who work within the university PGCE partnership with the approach taken by the SCITT, which they see as being more focused on immediate classroom-based experience. The evidence suggests that new "school-led" models of partnership may have greater levels of internal coherence, but that such coherence is premised predominantly on a craft model of teaching, reflecting Brown et al's (2015) conclusions that practical-based training is privileged over theory and analysis.

Finally, we would argue that ITE programs need to offer more than just a model of reflective practice. Beginning teachers need to be offered "suggestions for practice", including those based on theory and research as well as those put forward by experienced school-based colleagues, but all such suggestions need to be subject to critical reflection (McIntyre, 1995). Although many teachers might justifiably endorse the value of reflective practice, McIntyre argued that this was not, in itself, sufficient as a model for beginning teachers' learning since the latter need to evaluate *critically* all suggestions for practice that they receive, and need to do so against criteria that are informed by both practice and research. While it is expected that all ITE programs will encourage beginning teachers to be reflective, the focus of this reflection needs, at some point, to go beyond a "technical" level (Carr and Kemmis, 1986) and, we would argue, needs to develop their capacity to identify and resolve issues in their teaching; to evaluate effectively the outcomes of innovations to their practice using appropriate criteria for such evaluation; and to consider the impact of their teaching in relation to wider societal issues and policy contexts.

There may be tensions between the differing goals of preparing a teacher to be "school-ready" or "profession-ready"

One effect of the ITE reforms in England has been that schools are able to work either with school-centred providers (i.e., SCITTs) or with university providers, either within a School Direct model or a traditional partnership model. Unsurprisingly, the opportunity to continue to work with ITE providers while at the same time having greater autonomy in determining the nature of the training program has been popular, particularly when the School Direct model is premised on beginning teachers securing employment in the partnership schools in which they train. This can, however, lead to tensions.

If the purpose of the ITE program is seen as being to prepare teachers to work in a specific context, then the focus is likely to be on the need to induct teachers quickly into the practices of that particular setting. Of a lesser priority, perhaps, will be the development of teachers' capacities to be able to adapt when faced with a different context or a new set of curriculum demands. While all beginning teachers need to learn how to work efficiently and effectively, they also need to develop the capacity to adapt and to "move beyond existing routines ... to rethink key ideas (and) practices" (Hammerness et al., 2005, pp. 358–359). Furthermore, they also need to develop the "capacity for critical engagement with suggested innovations in classroom practice" (Hagger and McIntyre, 2006, p. 37). Partnerships need, therefore, to do more than just ensure that beginning teachers are "classroom ready"; they need to provide the learning opportunities which support these teachers in developing not only their capacity for innovation but also the appropriate criteria (both academic and practical) by which innovations in their teaching can be evaluated. Otherwise the risk is that the beginning teacher is only ever expected to deliver the school curriculum in a specified way; what Edwards and Protheroe (2004) refer to as "teaching by proxy" (2004, p. 183). McIntyre (2009) also highlights this risk, arguing against partnership models which are "aimed only at preparing beginning teachers for the status quo, and very deliberately being planned to avoid them being encouraged to think critically of that status quo" (2009, p. 603).

At Waterside Academy the SCITT program is, as noted above, focused on preparing teachers to work within a specific context and success is measured by their capacity, as one mentor put it, to "fit in", having been "nurtured and fostered" by the school. Teachers working within the program explicitly stated that it is unhelpful for trainee teachers to want to introduce innovative methods into their teaching. Therefore, within the SCITT program, there are few apparent tensions around the implementation of the program goals and the way in which training leads to these goals being realized, and little acknowledgment of the need for teachers (and new entrants to the profession in particular) to develop what might be called "adaptive expertise" (Berliner, 2001). While those involved in the program have confidence that the teachers will be well prepared to work in partnership schools, there needs also to be an awareness of the limitations of an approach which carries the potential risks identified by both McIntyre (2009) and Edwards and Protheroe (2004).

Partners have differing priorities and differing levels of accountability

Whatever a school's level of involvement in ITE, the fact remains that educating teachers can never be the primary focus of the school's activity, which always remains the education of its own pupils. Notwithstanding the perceived advantages of such involvement, a school's ITE involvement can be demanding and can potentially be viewed as a distraction from a core focus on pupil learning. Schools remain responsible for the performance of their pupils, while accredited ITE providers are

accountable elsewhere – namely for the quality of the ITE program and the effective working of the partnership. Partnerships must demonstrate not only compliance with a range of statutory requirements, but also that they prepare their beginning teachers sufficiently well to enable them to meet the relevant professional standards. New models of partnership (particularly those associated with the development of School Direct in England) can also be “*extremely resource-intensive in terms of administration for both universities and schools*” (McNamara and Murray, 2013, p. 16), meaning an additional demand on the limited resources available.

Just as a school’s priority will inevitably remain the education of its own pupils, there will also be related priorities within universities that offer teacher education programs. It has to be acknowledged that those working in university ITE partnerships will also have a wider responsibility within the institution in which they operate to engage in research and scholarship. Indeed, it is precisely because of their familiarity with, and continued production of, educational research, that universities have a role to play in ITE. Much of that research will be both driven by and inform the ITE programs to which they contribute. For many, this is not an optional extra, but a core part of their role as university-based teacher educators.

Where schools are simultaneously engaged with different models of partnership, which is by no means unusual, comparisons between individual partnership models are inevitable. At Waterside, our case study school, where the SCITT partnership is essentially managed by the school itself, there are fewer obvious conflicts in terms of priorities than if the program were being managed elsewhere. Channels of communication and line management structures are already established internally, and the immediacy of the partnership and the close involvement in it of staff within the school appear to give a sense of collegiate working with common goals. Furthermore, although resourcing is highlighted as an issue and, in particular, the time required to mentor trainee teachers effectively, there is a recognition that program directors are sensitive to such issues and are responsive, wherever possible. This responsiveness depends largely on established relationships within the school so that accountability for the ITE program becomes part of the general accountability within the school as a whole. Where schools are working in partnership with external accredited providers (such as a university), accountability across the partnership as a whole for the quality of the ITE program may be less easy to achieve.

Schools often engage in complex partnership arrangements with different providers

In England, schools (and in particular secondary schools) have become used to working in partnership with a range of different providers often leading to quite complex partnership arrangements (Mutton and Butcher, 2008). Differences may be evident in program structures and models of delivery, program documentation, the nature and frequency of the school-based experience, approaches to assessment, mentoring approaches etc. The practicality of negotiating a range of partnership relationships with a number of different providers may lead to attempts to rationalize school-based provision wherever possible, but there is a potential danger here in that “(t)he complexity and contestability of professional knowledge is therefore no longer seen to be at the heart of what partnership is about; professional knowledge becomes simplified, flattened – it is essentially about contemporary practice in schools” (Furlong et al., 2006, p. 41).

The policy changes at national level described above have enabled schools such as Waterside to seize the opportunity to become more actively involved in the training of teachers through its own SCITT program while at the same time continuing to be part of an existing and well-established partnership relationship. In many ways the tensions that could arise are not evident in this specific school context because it operates both ITE programs almost entirely separately. Subject departments within the school may host trainees from either program, or even from both at the same time, but in the case of the latter, school-based teacher educators will work only within one program (for example, separate mentors will work with the beginning teacher from each program respectively). In this case, rather than looking to align school-based provision across both programs, keeping provision within each program distinct has enabled the school to manage any potential complexity.

The implications for partnership in ITE as an “epistemological and pedagogical concept”

While the case of Waterside Academy illustrates the way in which one school has negotiated the complexity of partnership working there is a more general question as to the extent to which beginning teachers, who are training within programs which have developed as a result of recent government reforms in England, are able to draw effectively on sufficient sources of knowledge and, additionally, are provided with opportunities to interrogate critically the knowledge gained from these sources. The Universities’ Council for the Education of Teachers (UCET) has recently argued that, within collaborative partnership models, the focus is on preparing teachers who are,

- *competent and confident professionals who recognize and understand that educating is a professional, thoughtful and intellectual endeavor. They learn from research, direct experience, their peers and other sources of knowledge.*
- *epistemic agents, who act as independent thinkers, recognizing that knowledge, policy and practice are contestable, provisional and contingent. As such, teachers search for theories and research that can underpin, challenge or illuminate their practice. They are able to analyze and interrogate evidence and arguments, drawing critically and self-critically from a wide range of evidence to make informed decisions in the course of their practice.*

- **able to engage in enquiry-rich practice** and have a predisposition to be continually intellectually curious about their work with the capacity to be innovative, creative and receptive to new ideas emerging from their individual or collaborative practitioner enquiries.
- **responsible professionals** who embody high standards of professional ethics. They act with integrity and recognize the social responsibilities of education, working toward a socially just and sustainable world. (UCET, 2020, p. 1, original emphasis)

If these professional qualities are seen as desirable, then the question remains whether ITE programs can provide sufficient access to rich learning opportunities that will enable beginning teachers to develop such qualities. Evidence suggests that policy reforms in England, and in particular the introduction of School Direct, have led to a context in which:

(i)ncreasingly, teaching is conceived in craft-based, technicist terms strengthened by increasing prescription and performativity measures, which require teachers to present and shape knowledge in particular ways. Within this context, conceptions of the relationship between theory and practice have been progressively replaced by conceptions of practice that integrate situated conceptions of theory responsive to the needs of practice.

(Brown et al., 2015, p. 5).

Jackson and Burch (2016), however, argue that ITE policy reforms, and in particular the introduction of School Direct, provide an opportunity for a new conceptualization of the teacher educator working within ITE partnerships, which they refer to as “*hybridization*” (p. 517). They suggest that within such a model, previous unhelpful distinctions between theory and practice might be reduced, through a re-conceptualization of the role of the teacher educator (whether school-based or university-based), to enable both to work together to “*provide a principled pedagogical path through a changing landscape of education policy*” (p. 511). While such a suggestion might indeed address some of the inherent challenges of partnership working, it does appear somewhat idealistic, given that one of the perennial, and perhaps intractable, issues faced by ITE programs is variability in the quality of school-based mentoring (Hobson et al., 2009; Jones and Straker, 2006). (There may also, of course, be significant variation in the quality of the work of university-based teacher educators, although the fact there are fewer of them may make such difficulties less apparent). Nevertheless, the binaries and dichotomies that Jackson and Burch identify lie at the heart of the epistemological and pedagogical challenges within ITE partnerships. We would suggest, therefore, that a useful starting point to address these challenges might be to identify some key principles that could underpin ITE partnership working and apply equally to all contexts and models.

Principles to inform ITE partnership working

Taking account of what we already know about ITE partnerships in terms of policy, pedagogy and practice, we have identified the following principles which we see as being essentially pragmatic:

Principle 1: Partnerships need to have clear agreement in terms of the programme aims (and specifically a shared view of what the desired trainee outcomes should be).

We cannot assume that all ITE partnerships adhere to the same common goals. While all ITE programs in England have to comply with specific regulatory requirements (DFE, 2021) and are also expected to deliver the ITT Core Content Framework (DFE, 2019) as they prepare beginning teachers to meet the statutory Teachers’ Standards (DfE, 2011c), there is no requirement to adopt one particular approach to the preparation of new teachers. Informed by differing pedagogical principles, ITE partners may have different views as to what beginning teachers should know (beyond the government-defined ITT Core Content Framework) and be able to do. If program goals are, however, explicitly articulated and agreed, then a common understanding can be developed and any potential differences mediated by the agreed program goals. The strength of the SCITT program at Waterside Academy could, to some extent, be attributed to commonly understood and enacted program goals, albeit ones that focus primarily on enabling trainee teachers to make the transition easily from the training year to the first year of teaching in a partnership school.

The goals of the university PGCE program (with which the school is also engaged) imply, however, a different approach which requires beginning teachers on the program to draw on a range of *different* sources, both theoretical and practical and, through a process of “practical theorizing” (McIntyre, 1995; Hagger and McIntyre, 2006) to engage in a “*process of hypothesis-testing, requiring interpretation and judgment in action, rather than the routinized application of learned repertoires*” (Burn and Mutton, 2015, p. 221). While these program aims are agreed across the partnership, evidence from our case studies suggests that for some school-based teacher educators this approach may be viewed as an optional extra, rather than a core approach of the ITE program. We must, therefore, be aware that even when program aims are explicitly stated there still needs to be regular review of the way in which these aims are being enacted across the partnership.

Principle 2: Partnerships need a clearly defined learning programme in order for their trainee teachers to be able to capitalize fully on their school-based experiences.

While the temptation might be for school-led partnerships, in particular, to focus predominantly on the classroom performance of beginning teachers, there is a risk of the ITE curriculum becoming more focused on *doing* than on *learning*. To counter this risk, the curriculum needs to be constructed in such a way as to allow for a range of progressive learning opportunities for the beginning teachers, while also ensuring that they have sufficient time both to learn how to plan and to evaluate their teaching experiences effectively.

The challenge for any ITE program lies, however, in balancing two competing tensions: on the one hand not to overwhelm beginning teachers by the scope of what they need to learn and, on the other hand, acknowledging that the nature and complexity of what needs to be learnt is challenging and has to be planned effectively. Beginning teachers might naturally tend to attribute more value to those aspects of the program that they see as having more immediate application. Any ITE program has, therefore, to be both comprehensive in scope, yet able to respond to the teacher's current needs and priorities. Practice-based learning is then "*structured as much as is necessary to maximize learners' cognitive access to the full normal realities of doing the work, but must not distort these realities*" (Hagger and McIntyre, 2006, pp. 49–50). Finally, evidence suggests that, if the capacity for beginning teachers to learn from experience is to be realized fully, they need to be able to draw on as wide a range of sources as possible, including both the expertise of those within the ITE partnership, as well as wider research and scholarship (Hagger et al., 2008).

Principle 3: Partnerships need to consider when and how their trainees will develop the adaptive expertise necessary for sustaining a career in teaching.

As we have already noted, teachers need to be able to move beyond established routines and develop the capacity to innovate (Hammerness et al., 2005), as well as being able to engage critically with suggestions for innovation (Hagger and McIntyre, 2006). Likewise, Berliner (2001), in describing the nature of adaptive expertise, argues that "*expert, domain specific contextualized knowledge can often be a limited kind of knowledge*" (2001, p. 473) and argues that teaching expertise requires the ability to relate innovative practices to an existing knowledge base. Teachers need, of course, to develop efficiency, but also the wherewithal to engage critically with new ideas and suggestions for practice. The latter, Hammerness et al. (2005) argue, is dependent on a complex interplay between four different factors, namely understanding (i.e., a deep knowledge of content, pedagogy, students and social contexts); practices (developing, practicing and enacting a beginning repertoire); tools (the conceptual/theoretical and practical resources to use); and dispositions (habits of thinking and action). More recently, and in a similar vein, Orchard and Winch (2015) have argued that good teaching needs to be underpinned by the principled thinking of practitioners who are able to: understand key educational concepts and the major debates about aims, curriculum and pedagogy that have shaped practice in schools; engage with empirical educational research and be capable of assessing its quality and its relevance to their practice; and think through the ethics of teaching and the challenges of ethical decision making (adapted from Orchard and Winch, 2015, p. 5). This inevitably raises questions as to when, and how, beginning teachers acquire this adaptive expertise.

It could be argued that ITE partnerships should steer clear of the complexity inherent in the process of developing teachers' adaptive expertise and simplify the process of learning to teach as much as possible during the period of initial preparation by focusing predominantly on the acquisition of basic classroom competence. Such an approach does not necessarily imply a rejection of research evidence and theory per se, but instead reflects more a sense that practice – extensive experience and regular feedback – should take priority; research-based understandings or more theoretical considerations can, it could be argued, come later. The question remains, however, as to how teachers are prepared for other contexts where things may be done differently, or where the needs of specific learners may require alternative approaches. Our view is that there needs to be clear agreement within partnerships as to when and how such learning takes place.

Principle 4: Roles and responsibilities within the partnership need to be clearly defined and careful consideration given to the way in which responsibility for different aspects of the programme is managed.

Given that distinctive elements of an ITE partnership will provide a range of different contributions to the learning of beginning teachers, it is important that different types of expertise are recognized and valued and collectively provide an appropriate range of perspectives on which to draw.

Notwithstanding the many strengths of locating teacher education programs more fully within the school context, it is important that school-based training is not wholly focused on preparing new teachers to work in the same, or a similar context to that in which the ITE program is taking place. Partnership directors have an important role to play in ensuring that the model of the ITE program goes beyond that of learning "by imitation" (i.e., where the main approach is to adopt an apprenticeship model).

Furthermore, we would highlight the centrality of the role of the ITE mentor in the overall success of any partnership, although, as we have noted above, the quality of mentoring may be variable across programs. While there are mentoring standards in operation in England (DFE, 2016) that define the characteristics of mentors and the general level of support that they provide, they do not refer to any wider involvement in the planning or delivery of the whole ITE program. Strong partnerships will set out such roles clearly, going beyond the day-to-day support that the mentor provides and including their wider pedagogical role.

Principle 5: Sufficiently high levels of quality assurance need to be in place to ensure consistency, equity and compliance.

Strong partnerships depend on effective quality assurance procedures, with partners taking shared responsibility for regulatory compliance, curriculum content, program delivery and program outcomes. This necessitates a range of partnership-wide features. First, a shared understanding of any statutory frameworks within which training takes place. Second, regular review and evaluation of the program itself to ensure that quality is maintained, taking account of the broader ITE curriculum and its aims, as well as specific elements within it. Third, regular monitoring to ensure that beginning teachers in the program are treated equitably and have access to the full range of learning opportunities available to them. Finally, program innovations need to be monitored carefully and evaluated against previously designated anticipated outcome measures.

Conclusions

ITE partnership working is a complex activity with high levels of risk, given that much depends on its success – not least the need to equip prospective teachers with the knowledge, skills and understanding needed to sustain their future careers. These new teachers need to be able to make sound professional judgements, a process which [Kriewaldt and Turnidge \(2013\)](#), call “*clinical reasoning*” and which they define as the “*analytical and intuitive cognitive processes that professionals use to arrive at a best judged ethical response in a specific practice-based context*” (2013, p. 106). This clearly requires ITE programs to be structured in such a way as to give beginning teachers the opportunity to develop this level of professional competence, which, we would argue, can only be achieved through strongly collaborative partnerships with a high degree of program integration. We have outlined some of the challenges to effective partnership working, particularly in light of recent ITE policy reform in England where a “craft model” has been promoted by the government as the desirable norm. Different models of partnership will, as [Brisard et al. \(2005\)](#) suggest, each inherently contain particular theories about teacher learning and what beginning teachers specifically should know and be able to do, but we have to take care lest, as Alexander warned so long ago, the “*comfortable language of partnership*”, masks potentially damaging “*intractable issues*” (1984, p. 142).

We need, however, to look forward and acknowledge that, in spite of these issues and challenges, there is the opportunity for ITE partnerships to evolve and strengthen. Given the diverse nature of ITE provision in England, we have to embrace the differences in ITE program structures and their related pedagogical approaches but, nevertheless, seek to work in a principled way, whatever the nature of the individual partnership. We would argue that the principles we have set out above are pragmatic, and will enable all ITE partnerships, wherever they are situated and however they are administered, to flourish. We would also argue that ITE partnerships have much to offer in terms of school development within a self-improving school system ([Hargreaves, 2010](#); [Gilbert, 2012](#)), a policy which has been at the forefront of government policy since 2010. Hargreaves identifies four “*building blocks*” of such a system:

– capitalising on the benefits of clusters of schools

– adopting a local solutions approach

– stimulating co-construction between schools

– expanding the concept of system leadership. (2010, p. 5)

It is not difficult to see how ITE partnerships could play a significant part within such a conceptualization of school improvement in a genuinely school-led system, particularly in light of the report of the BERA-RSA Inquiry (2014) which identified “*research literate*” teachers as being key to developing capacity within a “*self-improving education system*” (2014, p. 5). However, in their analysis of the self-improving school-led system in England, [Greany and Higham \(2018\)](#) identify the problematic influences of “*hierarchy*”, “*markets*” and “*networks*” (2018, pp. 10–11) on its effective operation. It is not unreasonable to see similar systemic challenges affecting ITE partnership working, but in setting out our principles we are trying to look beyond hierarchical models of ITE. Tensions will inevitably arise if either ITE providers or school partners see themselves as the “*dominant*” partner, or if one approach to teacher learning is privileged over another. Instead we look positively to recent developments in Wales where collaborative partnership arrangements have been enshrined in the accreditation criteria for ITE providers ([Education Workforce Council, 2017](#)). Here, the

vision is of teachers whose ITE programs have specifically prepared them to meet the demands of a new school curriculum and to operate as agile and innovative intellectuals. ITE is seen in this context as an integral part of a national education reform movement, where both HEIs and schools are *jointly* accountable for the preparation of new teachers through a process of co-construction. Early indications are that the newly accredited ITE partnerships in Wales have established programs which are focused on developing teachers who are both research literate and research active (Harris et al., 2021), through collaborative programs in which university and schools work together as equal partners. Such a vision is also possible for ITE partnerships in England but, returning to the question we posed in the chapter's title, it may be much harder to maintain epistemological and pedagogical underpinnings when the nature of partnerships is determined to such a large extent by market-driven reforms rather than principle.

References

- Alexander, R., 1984. Innovation and continuity in the initial teacher education curriculum. In: Alexander, R., Craft, M., Lynch, J. (Eds.), *Change in Teacher Education. Context and Provision since Robins*. Holt, Rinehart and Winston, London, pp. 103–160.
- Alter, J., Cogshall, J.G., 2009. Teaching as a Clinical Practice Profession: Implications for Teacher Preparation and State Policy. Issue Brief. National Comprehensive Center for Teacher Quality.
- Bartholomew, S.S., Sandholtz, J.H., 2009. Competing views of teaching in a school-university partnership. *Teach. Teach. Educ.* 25 (1), 155–165.
- Benton, P. (Ed.), 1990. *The Oxford Internship Scheme: Integration and Partnership in Initial Teacher Education*. Calouste Gulbenkian Foundation, London.
- Berliner, D.C., 2001. Learning about and learning from expert teachers. *Int. J. Educ. Res.* 35 (5), 463–482.
- Brisard, E., Menter, I., Smith, I., 2005. Models of Partnership in Programmes of Initial Teacher Education. Full report of a systematic literature review commissioned by the General Teaching Council for Scotland, GTCS Research, Research Publication No. 2. General Teaching Council for Scotland, Edinburgh.
- Brown, T., Rowley, H., Smith, K., 2015. The Beginnings of School Led Teacher Training: New Challenges for University Teacher Education. School Direct Research Project Final Report. Manchester Metropolitan University. Available at: <https://espace.mmu.ac.uk/602385/2/School%20Direct%20Research%20Report.pdf>. (Accessed 8 April 2021).
- Burn, K., Hagger, H., Mutton, T., 2015. Beginning Teachers' Learning: Making Experience Count. Critical Publishing, Northwich.
- Burn, K., Mutton, T., 2015. A review of 'research-informed clinical practice' in initial teacher education. *Oxf. Rev. Educ.* 41 (2), 217–233.
- Carr, W., Kemmis, S., 1986. *Becoming Critical: Education Knowledge and Action Research*. The Falmer Press, London.
- Conroy, J., Hulme, M., Menter, I., 2013. Developing a 'clinical' model for teacher education. *J. Educ. Teach.* 39 (5), 557–573.
- Darling-Hammond, L., 2006. *Powerful Teacher Education: Lessons from Exemplary Programs*. Jossey-Bass, San Francisco, CA.
- Darling-Hammond, L., 2014. Strengthening clinical preparation: the holy grail of teacher education. *Peabody J. Educ.* 89 (4), 547–561.
- Department for Education, 2010. *The Importance of Teaching: The Schools White Paper 2010*. Her Majesty's Stationery Office, London.
- Department for Education (DfE), 2011a. Training our next generation of outstanding teachers - an improvement strategy for discussion. Available at: <https://www.gov.uk/government/publications/training-our-next-generation-of-outstanding-teachers-an-improvement-strategy-for-discussion>. Retrieved on 8th April 2021.
- Department for Education (DfE), 2011b. Training our next generation of outstanding teachers: implementation plan. Available at: <https://www.gov.uk/government/publications/training-our-next-generation-of-outstanding-teachers-implementation-plan>. Retrieved on 8th April 2021.
- Department for Education (DfE), 2011c. Teachers' standards. Available at: <https://www.gov.uk/government/publications/teachers-standards>. Retrieved on 8th April 2021.
- Department for Education (DfE), 2016. National Standards for School-Based Initial Teacher Training (ITT) Mentors. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/536891/Mentor_standards_report_Final.pdf. Retrieved on 8th April 2021.
- Department for Education (DfE), 2019. Initial teacher training (ITT): core content framework. Available at: <https://www.gov.uk/government/publications/initial-teacher-training-itt-core-content-framework>. Retrieved on 8th April 2021.
- Department for Education (DfE), 2021. Initial teacher training (ITT): criteria and supporting advice. Available at: <https://www.gov.uk/government/publications/initial-teacher-training-criteria/initial-teacher-training-itt-criteria-and-supporting-advice>. Retrieved on 8th April 2021.
- Education Workforce Council (EWC), 2017. Submitting a programme of initial teacher education to EWC for accreditation. Available at: <https://gov.wales/sites/default/files/publications/2018-09/criteria-for-the-accreditation-of-initial-teacher-education-programmes-in-wales.pdf>. (Accessed 18 April 2021).
- Edwards, A., Protheroe, L., 2004. Teaching by proxy: understanding how mentors are positioned in partnerships. *Oxf. Rev. Educ.* 30 (2), 183–197.
- Furlong, J., Barton, L., Miles, S., Whiting, C., Whitty, G., 2000. *Teacher Education in Transition: Reforming Professionalism*. Open University Press, Buckingham.
- Furlong, J., Campbell, A., Howson, J., Lewis, S., McNamara, O., 2006. Partnership in English initial teacher education: changing times, changing definitions. Evidence from the teacher training agency's national partnership project. *Scot. Educ. Rev.* 37, 32–45.
- Furlong, J., McNamara, O., Campbell, A., Howson, J., Lewis, S., 2008. Partnership, policy and politics: initial teacher education in England. *Teach. Teach. Theor. Pract.* 14 (4), 307–318.
- Furlong, J., Whitty, G., Whiting, C., Miles, S., Barton, L., Barrett, E., 1996. Re- defining partnership: revolution or reform in initial teacher education? *J. Educ. Teach.* 22 (1), 39–56.
- Gilbert, C., 2012. Towards a Self-Improving System: The Role of Accountability in Schools. Nottingham National College for School Leadership. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/329743/towards-a-self-improving-system-school-accountability-thinkpiece.pdf. Retrieved on 18th April 2021.
- Gove, M., 2010. Speech to the national college annual conference 16th June 2010. Available at: <http://www.education.gov.uk/inthenews/speeches/a0061371/michael-gove-to-the-national-college-annual-conference-birmingham>. (Accessed 8 April 2021).
- Greany, T., Higham, R., 2018. Hierarchy, Markets and Networks Analysing the 'Self-Improving School-Led System' Agenda in England and the Implications for Schools. UCL, IoE Press, London. Available from: <https://mk0nuffieldfounpg9ee.kinstacdn.com/wp-content/uploads/2019/11/Hierarchy-Markets-and-Networks.pdf>. Retrieved on 18th April 2021.
- Hagger, H., Burn, K., Mutton, T., Brindley, S., 2008. Practice makes perfect? Learning to learn as a teacher. *Oxf. Rev. Educ.* 34 (2), 159–178.
- Hagger, H., McIntyre, D., 2006. *Learning Teaching from Teachers: Realizing the Potential of School- Based Teacher Education*. Open University Press, Maidenhead.
- Hammerness, K., 2006. From coherence in theory to coherence in practice. *Teach. Coll. Rec.* 108 (7), 1241–1265.
- Hammerness, K., van Tartwijk, J., Snoek, M., 2012. Teacher preparation in The Netherlands. In: Darling-Hammond, L., Lieberman, A. (Eds.), *Teacher Education Around the World: Changing Policies and Practices*. Routledge, Abingdon.
- Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., Zeichner, K., 2005. How teachers learn and develop. In: Darling-Hammond, L., Bransford, J., LePage, P., Hammerness, K., Duffy, H. (Eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, San Francisco, CA.
- Hargreaves, D., 2010. Creating a Self-Improving System. Nottingham National College for School Leadership. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/325873/creating-a-self-improving-school-system.pdf. Retrieved on 18th April 2021.
- Harris, A., Jones, M., Furlong, J., Griffiths, J., Hannigan-Davies, C., 2021. A research informed approach to initial teacher education in Wales: intentions, examples and reflections. In: Mayer, D. (Ed.), *Teacher Education Policy and Research: Global Perspectives*. Springer, pp. 195–208.
- Hobson, A.J., Ashby, P., Malderez, A., Tomlinson, P.D., 2009. Mentoring beginning teachers: what we know and what we don't. *Teach. Teach. Educ.* 25 (1), 207–216.

- Jackson, A., Burch, J., 2016. School Direct, a policy for initial teacher training in England: plotting a principled pedagogical path through a changing landscape. *Prof. Dev. Educ.* 42 (4), 511–526.
- Jones, M., Straker, K., 2006. What informs mentors' practice when working with trainees and newly qualified teachers? An investigation into mentors' professional knowledge base. *J. Educ. Teach.* 32 (2), 165–184.
- Kriewaldt, J., Turnidge, D., 2013. Conceptualising an approach to clinical reasoning in the education profession. *Aust. J. Teach. Edu.* 38 (6), 103–115.
- McIntyre, D., 1995. Initial teacher education as practical theorising: a response to Paul Hirst. *Br. J. Educ. Stud.* 43 (4), 365–383.
- McIntyre, D., 2009. The difficulties of inclusive pedagogy for initial teacher education and some thoughts on the way forward. *Teach. Teach. Edu.* 25 (4), 60–608.
- McLean Davies, L., Anderson, M., Deans, J., Dinham, S., Griffin, P., Kameniar, B., Page, J., Reid, C., Rickards, F., Tayler, C., Tyler, D., 2013. Masterly preparation: embedding clinical practice in a graduate pre- service teacher education programme. *J. Educ. Teach.* 39 (1), 93–106.
- McNamara, O., Murray, J., 2013. The School Direct Programme and its Implications for Research- Informed Teacher Education and Teacher Educators. Higher Education Academy, York.
- Mattsson, M., Ellertson, T., Rorison, D. (Eds.), 2011. *A Practicum Turn in Teacher Education*. Sense, Rotterdam.
- Menter, I., Mutton, T., Burn, K., 2019. Learning to teach in England - reviewing policy and research trends. In: Tatto, T., Menter, I. (Eds.), *Knowledge, Policy and Practice in Learning to Teach: A Cross-National Study*. Bloomsbury, London, pp. 60–80.
- Mutton, T., Burn, K., Hagger, H., Thirlwall, K., 2018. *Teacher education partnerships: Policy and practice*. Critical Publishing.
- Mutton, T., Butcher, J., 2008. 'we will take them from anywhere': schools working within multiple initial teacher training partnerships. *J. Educ. Teach.* 34 (1), 45–62.
- Orchard, J., Winch, C., 2015. What training do teachers need?: why theory is necessary to good teaching. *Impact (Number 22)*, 1–43.
- Sahlberg, P., 2012. The most wanted: teachers and teacher education in Finland. In: Darling-Hammond, L., Lieberman, A. (Eds.), *Teacher Education Around the World: Changing Policies and Practices*. Routledge, Abingdon, pp. 1–21.
- Taylor, A., 2008. Developing understanding about learning to teach in a university-schools partnership. *Br. Educ. Res. J.* 31 (1), 63–90.
- Universities' Council for the Education of Teachers (UCET), 2020. UCET – the intellectual base of teacher education. Available at: <https://www.ucet.ac.uk/11675/intellectual-base-of-teacher-education-report-updated-february-2020>. Retrieved on 8th April 2021.
- Universities UK, 2014. The impact of initial teacher training reforms on English higher education institutions. Available at: <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2014/impact-of-itt-reforms-on-english-heis.pdf>. Retrieved on 8th April 2021.
- Zeichner, K., 2010. Rethinking the connections between campus courses and field experiences in college and university-based teacher education. *J. Teach. Educ.* 61 (1), 89–99.

Equity in early childhood education: policy framing, translation and enactment

Andrea Nolan and Tebeje Molla, School of Education, Deakin University, Geelong, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	93
Equity work	94
Policy framing	94
Contextual translation	95
Homogenized ECEC workforce and policy enactment	95
Moving the equity agenda forward	96
Conclusion	97
References	97

Introduction

The importance of the early years of a child's life to their development, learning and wellbeing is unquestionable. We know that the first 1000 days of a child's life are critical for their neurological and biological development (Moore et al., 2017). We also know that the effect of trauma during these early years can have a significant impact on the child (Cummings and Swindell, 2019). Drawing on a longitudinal empirical analysis, in *Giving Kids a Fair Chance*, Nobel Prize-winning economist James Heckman writes:

Kids born into disadvantaged environments are at much greater risk of being unskilled, having low lifetime earnings, and facing a range of personal and social troubles, including poor health, teen pregnancy, and crime. While we celebrate equality of opportunity, we live in a society in which birth is becoming fate.

Heckman (2013:1).

Research has shown that high-quality early childhood programs can support all aspects of children's learning and development (Sylva et al., 2010), and that the quality of adult-child interactions are of paramount importance (Tyler, 2016). This points to the need for a skilled, knowledgeable and professional workforce who are guided by policies that are geared toward equity provision for all children.

Like many other OECD countries, early childhood education and care (ECEC) in Australia is a key policy priority area that has passed through a series of reforms in the last decade. Recent government initiatives focus attention on the provision of high-quality services for young children and their families. They call for the raising of standards to ensure the delivery of consistent, accessible, high quality services across the country, and establish processes for continuous improvement and consistency in these services. One policy example is the "universal access to early childhood education" (CoAG, 2008) policy aimed to ensure that all children can participate in a quality preschool program. The National Partnership Agreement on Universal Access to Early Childhood Education focuses on supporting children's early learning and development through maintaining:

Universal access to quality early childhood education programs for all children enrolled in the year before full-time school for 600 hrs per year, delivered by a qualified early childhood teacher who meets National Quality Framework (NQF) requirements, and with a focus on participation by Indigenous children, vulnerable children and children experiencing disadvantage, regardless of the setting in which programs are delivered.

CoAG (2020:3).

Although the "universal access" policy has cost the government over 4.2 billion dollars in the last 12 years (2009–2020), inequality still persists in the ECEC sector (AEDC, 2016; Nolan and Lamb, 2019; Nous, 2020), with poor developmental and learning outcomes for children experiencing vulnerability and disadvantage. Inequality in the ECEC sector continues for children with low socio-economic status, Indigenous status, and non-English language backgrounds (Productivity Commission, 2014; AEDC, 2016). Geographic location also remains a factor of inequality (Raban and Kilderry, 2017). In other words, there are "cohorts of children who are either not accessing a preschool education, or who are not accessing the hours of quality early education needed to make a difference to their development" (O'Connell et al., 2016: vii).

The question remains unanswered: why children experiencing disadvantage still miss out in early learning notwithstanding the successive reforms, the universal access agenda, and the amount of public investment in the sector? This chapter begins by exploring how equity is framed in policies that guide educator and teacher practice. The term "Educator" refers to early childhood practitioners who work directly with children in early childhood settings and hold a Certificate or Diploma (non-teaching) qualification. The term "Teacher" refers to practitioners who hold a teaching Degree qualification. The discussion then moves to identify and explore

the issues that this policy framing misses in terms of contextuality. It concludes by proposing ways for moving the equity agenda forward.

Equity work

The identification of “issues for policy attention” in the OECD report for Australia (OECD, 2001, 2006), along with strong evidence of the major challenges faced by early childhood education and care services, acted as a catalyst for action. These challenges included a lack of access to quality services for children from disadvantaged backgrounds (Press and Hayes, 2001), the fragmentation of the sector’s policy and governance (OECD Review Team, 2001), and Australia’s low public spending on early childhood education among OECD member countries (OECD, 2001, 2006). This triggered the Commonwealth, State and Territory Governments of Australia to embark on a reform agenda. The Council of Australian Governments (COAG from June 2020, replaced by the National Federation Reform Council, NFRC) endorsed the National Partnership Agreement on Early Childhood Education, confirming their commitment to achieving “universal access” to quality early childhood education (CoAG, 2008). Under subsequent National Partnership Agreements on Universal Access in Early Childhood Education (CoAG, 2013, 2015, 2016, 2020), Commonwealth, State and Territory governments committed to ensuring that accessible quality early childhood education is available for all children regardless of socio-economic status or geographic location.

Elsewhere we have written in detail about the rationales and instruments that underpin the “universal access” policy initiative and have problematized related policy documents (see Molla and Nolan, 2019a,b). In this section of the chapter, we will briefly summarize our analysis to make clear the logics of the policies that support equity provision and argue that the policy framings are underpinned by economic, educational and social rationales that impact practice. We begin by outlining how equity issues are framed in policies. Then we move to illustrate the ramifications of such equity framings for the translation of policy to the practice context, and the missed opportunity that is generated by the homogenized representation of the workforce.

Policy framing

The policy frame analysis literature (e.g., Bacchi, 2009; Rein and Schön, 1996; Wagenaar, 2015) suggests that equity instruments made available by governments reflect how inequality is conceived and framed as a problem. That is, the way a public issue is framed has a direct bearing on what strategies are designed in response. A critical frame analysis aims at revealing whose values are authoritatively allocated, and whose interests are excluded in the plan of action. Guided by this assumption, we sought to make sense of the representation of the problem of inequality in Australian ECEC, and to unmask assumptions and rationales underpinning equity provisions in the sector. What our analysis has shown is that disadvantage in the sector has been framed as a lack of access, limited navigational capacity and cultural exclusion, with these policy framings underpinned by economic, educational and social rationales.

The National Partnership Agreement on Universal Access in Early Childhood Education aims at ensuring that Indigenous children and children experiencing vulnerability and disadvantage “have access to and participate in an affordable, quality early childhood education program” (CoAG, 2013: 4). Framing devices such as “universal access,” “qualified early childhood teacher,” “all children,” “cost,” and “barrier” suggest that the governments recognize that there is a lack of access to quality early childhood education. The reforms also recognize inequality in navigational capacity among parents of different socio-economic backgrounds. Scattered ECEC services mean that parents of disadvantaged background (e.g., Indigenous communities in remote areas and recently resettled migrants and refugees) found it difficult to navigate and compare programs and providers. The reform movement in Australia’s ECEC sector was, therefore, initiated in part to “fix” fragmented services and navigational difficulty (CoAG, 2008). Finally, frames such as “cultural safety,” “culturally inclusive,” and “culturally appropriate” presuppose a problem associated with cultural exclusion or domination. Historically, Indigenous people’s knowledge and culture remained “invisible,” or were misrepresented in Australian curricula and educational experiences (Wilson, 2016). Against this backdrop, the Government’s universal access and Closing the Gap policies emphasize the importance of creating culturally inclusive learning environments for Indigenous children.

As noted above, notwithstanding the equity provisions put in place, inequality in early childhood education has persisted along the lines of low SES, Indigenous status, and non-English language backgrounds (AEDC, 2016; Productivity Commission, 2014). One possible explanation for the persisting inequality can be “frame conflict” (Schön and Rein, 1994), which is expressed in contradictions between broad political economy goals and equity needs, and/or between those who construct the policy agenda and those who enact the agenda. For example, framing devices such as accountability and efficiency reflect the prevalence of a neoliberal political economy in the national policy space, which runs counter to the basic premise of distributive justice that prioritizes the standpoint of the least advantaged (Connell, 2009). As Forest and Lombardo (2012) argue, “broader hegemonic discourses [...] define the horizon against which individual frames take place” (p. 17). Our research shows that the social justice goals in the ECEC sector are strategically framed to resonate with the hegemonic economic discourse that has pervaded the education system in general. For example, there is widespread consensus that equitable access to quality education at an early stage in life could contribute to reducing the negative effects of disadvantage in adult life thus making people more productive citizens. This aligns with the government’s economic agenda of enhancing national productivity through investment in education (Molla and Nolan, 2019a). This has been referred to by Cook et al. (2016) as the “neoliberal future-worker-investment discourses” (p. 58) that extend rationality to the ECEC sector, and links inequality in access to quality education as a loss of human capital.

The national and international research literature points to a link between the provision of quality early childhood programs and workforce qualifications. Generally speaking, the quality discourse in the ECEC sector focuses on workforce development, quality assurance and child-centered curriculum. In other words, key elements of quality ECEC services include the availability of adequately qualified early childhood professionals, the provision of stimulating learning experiences for children, lower child-to-staff ratios and active parental engagement (CoAG, 2009: 8–9). In this respect, tackling inequality or widening access to quality ECEC services means ensuring that every child has access to qualified educators, stimulating pedagogical practice, and culturally responsive and inclusive learning experiences.

Contextual translation

Policy statements represent intentions; the transformation of intentions into results, depends on contextual policy translation. For instance, in enacting equity policies, teachers and educators determine “what counts as social justice in their contexts of practice” as they act to interpret policy through the institutional ethos of their workplace, the resources of their setting, and their own social justice dispositions (Molla and Gale, 2019: 858). This active policy engagement positions teachers and educators as key agents in translating nationally prescribed equity-related policies. Here we name the discord between policy and practice to show the impact of framings on the equity work of ECEC educators.

In our research we have identified discords between policy expectations and educator experience and practice that work against an equity agenda. Specifically, we have been able to identify misalignments between policy and practice occurring in relation to making professional judgements, recognizing teachable moments, working closely with families, and recognizing and responding to barriers to children achieving success (Molla and Nolan, 2019a,b). This becomes apparent as educators navigate official expectations of their roles while relying heavily on what they value and choose to follow in their pedagogic practices (Buckler, 2016). Taking a capability approach to human development as the analytical framework for one of our studies, we considered what constituted valued and valuable professional “beings and doings” (Sen, 2009) in the ECEC sector from the view of policy expectations and educator experiences. We did this to closely examine the implication of the neoliberal logic of professionalization for professional capabilities valued in the policy space as well as in the field of practice. The findings highlight that as early childhood educators navigate official expectations of their role, they rely heavily on what they value and choose to pursue in their pedagogic work. This implies that effective implementation of the professionalization initiative is mediated by subjective meaning systems of the educators and objective structures in the context of practice. In other words, policy gets lost in the translation to the context. In our research, we saw tensions between policy and practice in aspects such as what it meant to hold expertise, recognition of the work, and different interpretations of responsiveness to children’s learning. These differences were attributed to a lack of professional agency, policy literacy, and the social justice dispositions of the educators.

Here we focus on social justice. For many early childhood teachers and educators in our studies, there is a propensity to be responsive to certain equity groups at the expense of others. In the professionalization discourse, the Australian Government frames responsiveness of educators in terms of broad inclusion goals along the lines of socio-economic status, cultural background and special learning needs. Practice guidelines propose that teachers and educators hold high expectations for the learning of all the children with whom they work, and “recognize and respond to barriers to children achieving educational success” (DEWR, 2009:13). However, the educators’ accounts of responsiveness in the studies that we have been involved with (Nolan, 2021; Nolan and Lamb, 2019) are narrowly focused. For example, one study while providing many real strengths that the educators were saying, observing, reflecting on and doing in relation to inclusive practice, none of the educators mentioned children from socio-economically disadvantaged backgrounds as targets of their social justice work yet the majority of these educators worked in low socio-economic communities (Nolan and Lamb, 2019). What this points to is a contradiction between those who construct the policy and those who enact it. Without attention to the context within which practice occurs, there is the risk of policy being mis-interrupted, misrepresented, or viewed by practitioners as unrelated to their work context.

As discussed above—not recognizing that practice is mediated by subjective meaning systems and objective structures—highlights that the intent of policy can become lost in translation. Without considering the context such as the life experiences of teachers and educators and what they draw on to inform their practice at the point of policy framing the translation of equity policies can be problematic. In the following section we illustrate how the diverse experiences and life contexts of the ECEC workforce are currently compartmentalized leading to unitary understandings of quality practice. This conceptualization misses the opportunity to build from the strengths and uniqueness of each educator to support policy enactment.

Homogenized ECEC workforce and policy enactment

The professionalization of ECEC has been a focus of attention and debate internationally for a number of years (Dalli and Urban, 2008). In Australia (SCSEEC, 2012), the professionalization of the ECEC workforce through upskilling, is seen as a way of ensuring the effective implementation of the policy reforms which include catering for the diversity of children and families who engage with early childhood services (Torrii et al., 2017; Productivity Commission, 2011). As noted elsewhere (Nolan and Molla, 2018), the professionalization of the ECEC workforce focuses on building pedagogical competence and cultivating professional dispositions. This includes social justice intentionality and cultural sensitivity. As we contend elsewhere (Nolan and Lamb, 2019), the reform agenda aspires for all early childhood educators to be skilled, reflective professionals who can actively support the learning of all children. Educators provide learning experiences for children to achieve equity goals as outlined in practice guidelines which

promotes social justice. Through critical reflection, educators are expected to identify children who could benefit from additional support in their learning, provide that support, and assist families to access specialist services if required (DEEWR, 2009). As the Government put it, “The intent is to ensure that all children’s experiences are recognized and valued” (DEEWR, 2009: 24).

However, the professionalization process in general and the expected equity work tend to homogenize the workforce. The reform agenda links building human capital to ECEC workforce reform based on “scope for the sector to shift from a path of skill atrophy toward a path of skill growth ... along with the scope for professionalization” (Bretherton, 2010: 8). The ECEC workforce is diverse in qualifications, experience, backgrounds and roles undertaken within organizations, yet all educators and teachers are expected to engage with policy documents that guide practice. The ECEC workforce in Australia comprises of staff holding certificate to Masters level qualifications with no differentiation of expectations by Governments on their ability to follow government guidelines and implement government initiatives. For example, the national framework to guide practice (DEEWR, 2009) labels all workers, regardless of qualifications or position, as “educators.” While this might be an attempt to raise the status of the vocationally qualified workers, an unintended consequence may be the lowering of the status of early childhood teachers. This framework fails to recognize the complexity and diversity of the workers; the fragmented nature of the sector; or the alignment of professional identities based on the emotional labor of care in a highly feminized workforce (Dalli, 2002; Blackmore, 2008; Brennan, 1998). Such homogenizing framing does not consider or value the uniqueness of the individual teacher/educator and their engagement with policy discourse. In doing this, an opportunity is missed to harness the rich diversity and the “professional capital” (Hargreaves and Fullan, 2012) of the workforce.

While a link has been established between quality early childhood programs and educator qualifications, educators at all levels of qualification have value to add to current notions of inclusive practice, especially when diversity is the lens (Cherrington and Shuker, 2012). Early childhood educators determine the “discourses that they make available to children and those that they silence, through their daily practices, pedagogies and curricula” (Robinson, 2002:416). As such they should be self-aware of their “unique likes and dislikes” (Vuckovic, 2008:12) as this is considered as “indispensable if the learning environment is intended to be culturally sensitive” (Vuckovic, 2008:12). Through recognition and support of the diversity of the ECEC workforce, it is hoped that other discourses more familiar to the diversity of child and family backgrounds and experiences will be enabled and valued. This means considering early childhood professionalism as “a situated concept” (Oberhuemer, 2008:13) that is influenced by historical, socio-cultural, economic and political contexts (Dalli and Urban, 2008; Oberhuemer, 2008). As Bankovic (2014) reflects, this enables educators immediate and societal contexts to be brought to the understanding and definition of professionalism. A focus on the diversity of the workforce can therefore disrupt unitary understandings of quality practice and hold ramifications for rethinking policy and professional learning.

Moving the equity agenda forward

The United Nations Convention on the Rights of the Child acknowledges “the inherent dignity and the equal and inalienable rights of all members of the human family” (Preamble). This includes young children having the right to an education that helps them fully develop their personalities, talents and abilities (Article 29), with all rights applying to all children whatever their race, religion, abilities, whatever they think or say, whatever type of family they come from (Article 2). Australian research shows that children living with disability, in poverty and in jobless families face multiple, complex and deeper levels of deprivation than their peers (Sollis, 2019). It is estimated that one in five Australian children could be considered as being multi-dimensionally deprived (Sollis, 2019). However, research shows that quality ECEC services can have a profound impact on all children’s outcomes, especially for those children from disadvantaged backgrounds (Elliott, 2006; Harrison et al., 2012; PricewaterhouseCoopers [PwC], 2014). It seems more important than ever to deliver quality early childhood services underpinned by responsive policies that recognize the contexts of children’s lives and the capabilities of the ECEC workforce.

With the policy agenda tied to the delivery of quality ECEC services and aligning quality with the professionalization of the ECEC workforce, there are implications for how equity is framed in policy. In Australia, professionalism within the ECEC sector is currently narrowly framed where measurable outcomes are valued over other more non-measurable attributes such as professional dispositions. In our research these professional dispositions were identified as being caring, having and displaying empathy, demonstrating resilience, patience, having confidence, passion, commitment, showing kindness, and a positive attitude. This seems a missed opportunity to draw on educators’ professional capital to work with children and families and to support them in their professional development as effective early childhood educators. Once workforce diversity is acknowledged, and the alignment of professional attributes valued in the policy space and in the field of practice to understand educator agency are identified, professional learning can be tailored to be meaningful. This learning can better support the professionalization of the ECEC workforce in understanding, enhancing and implementing the equity policy agenda. This is considered an important step “to better understand how practitioners can work effectively with children, families, and colleagues from diverse backgrounds” (Cherrington and Shuker, 2012, p. 89).

For professional learning to be effective there needs to be an association between educators’ aspirations and systematic requirements, a sense of collegiality between participants, critical deliberation in “safe” learning environments where experiences and knowledge can be shared without fear of retribution, and acknowledgment that learning takes place in the domains of professional dispositions, pedagogical knowledge and social capital (Nolan and Molla, 2018). However, unfortunately, our research has highlighted that there is currently no equal access to professional learning opportunities for educators (Nolan and Molla, 2019). This

means that some educators/teachers have very little agency to strengthen their contribution to the early childhood program, advance their professional growth and engage with policy in a more informed way.

Conclusion

In this chapter we have specifically argued that the framing of equity shapes practice in ways where contextually emerging issues such as child poverty are in danger of being overlooked or downplayed. This can lead to educators/teachers having a propensity to be responsive to certain equity groups at the expense of others. We have emphasized that factors such as a lack of expansive policies, the interplay and translation between policy and practice, the unitary understandings and expectations of quality practice for a diverse workforce, and the ability of educators and teachers to access effective professional learning opportunities that build on their professional capital, raise challenges for quality service provision for all children in Australia.

References

- AEDC [Australian Early Development Census], 2016. Australian Early Development Census National Report 2015: A Snapshot of Early Childhood Development in Australia. Commonwealth of Australia, Canberra.
- Bacchi, C., 2009. Analysing Policy: What's the Problem Represented to Be? Pearson, Melbourne.
- Banković, I., 2014. Early childhood professionalism in Serbia: current issues and developments. *Int. J. Early Years Educ.* 22 (3), 251–262.
- Blackmore, J., 2008. Measures of hope and despair: emotionality, politics, and education. In: Samier, E., Schmidt, M. (Eds.), *Emotional Dimension of Educational Administration and Leadership*. Routledge, London, pp. 109–124.
- Brennan, D., 1998. *The Politics of Australian Child Care: Philanthropy to Feminism and Beyond*. CUP, Cambridge.
- Bretherton, T., 2010. Developing the Child Care Workforce: Understanding “fight” or “flight” Amongst Workers. A National Vocational Education and Training Research and Evaluation Program Report. Commonwealth of Australia.
- Buckler, A., 2016. *Quality Teaching and the Capability Approach: Evaluating the Work and Governance of Women Teachers in Rural Sub-Saharan Africa*. Routledge, London.
- Cherrington, S., Shuker, M.J., 2012. Diversity amongst New Zealand early childhood educators. *N. Z. J. Teach. Work* 9 (2), 76–94.
- CoAG [Council of Australian Governments], 2008. National Partnership Agreement on Early Childhood Education. Australian Government, Canberra.
- CoAG [Council of Australian Governments], 2009. Investing in the Early Years: A National Early Childhood Development Strategy. Australian Government, Canberra.
- CoAG [Council of Australian Governments], 2013. National Partnership Agreement on Universal Access to Early Childhood Education (Phase I). Australian Government, Canberra.
- CoAG [Council of Australian Governments], 2015. National Partnership Agreement on Universal Access to Early Childhood Education (Phase II). Australian Government, Canberra.
- CoAG [Council of Australian Governments], 2016. National Partnership Agreement on Universal Access to Early Childhood Education (Phase III). Australian Government, Canberra.
- CoAG [Council of Australian Governments], 2020. National Partnership on Universal Access to Early Childhood Education—2018–2021. Australian Government, Canberra.
- Connell, R., 2009. Good teachers on dangerous ground: towards a new view of teacher quality and professionalism. *Crit. Stud. Educ.* 50 (3), 213–229.
- Cook, K., Corr, L., Breitzkreuz, R., 2016. The framing of Australian childcare policy problems and their solutions. *Crit. Soc. Pol.* 36 (4), 1–22.
- Cummings, K.P., Swindell, J., 2019. Using a trauma-sensitive lens to support children with diverse experiences. *Young Except. Child.* 22 (3), 139–149.
- Dalli, C., Urban, M., 2008. Towards new understandings of the early years profession: the need for a critical ecology. *Eur. Early Child. Educ. Res. J.* 16, 150–155.
- Dalli, C., 2002. Constructing identities: being a “mother” and being a “teacher” during the experience of starting childcare. *Eur. Early Child. Educ. Res. J.* 10 (2), 85–101.
- DEEWR [Department of Education Employment and Workplace Relations], 2009. *Belonging, Being and Becoming: The Early Years Learning Framework for Australia*. DEEWR, Canberra.
- Elliott, A., 20. Early childhood education: pathways to quality and equity for all children. *Aust. Educ. Rev.* 50, 1–64.
- Forest, M., Lombardo, E., 2012. The Europeanization of gender equality policies: a discursive-sociological approach. In: Lombardo, E., Forest, M. (Eds.), *The Europeanization of Gender Equality Policies: A Discursive-Sociological Approach*. Palgrave Macmillan, London, pp. 1–27.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press and Ontario Principals' Council, New York.
- Harrison, L.J., Goldfield, S., Metcalfe, E., Moore, T., 2012. Early learning programs that promote children's developmental and educational outcomes. In: Resource Sheet No. 15. Produced for the Closing the Gap Clearinghouse. Australian Institute of Health and Welfare and Melbourne: Australian Institute of Family Studies, Canberra.
- Heckman, J.J., 2013. *Giving Kids a Fair Chance*. The MIT Press.
- Molla, T., Gale, T., 2019. Positional matters: school leaders engaging with national equity agendas. *J. Educ. Pol.* 34 (6), 858–876.
- Molla, T., Nolan, A., 2019a. The problem of inequality in Australia's Early Childhood Education and Care sector: a policy frame analysis. *Aust. J. Educ.* 63 (3), 322–339.
- Molla, T., Nolan, A., 2019b. The “universal access to early childhood education” agenda in Australia: rationales and instruments. *Educ. Res. Pol. Pract.* 18 (1), 1–16.
- Moore, T.G., Arefadib, N., Deery, A., West, S., 2017. *The First Thousand Days: An Evidence Paper*. Centre for Community Child Health, Murdoch Children's Research Institute, Parkville, Victoria.
- Nolan, A., Lamb, S., 2019. Exploring the social justice work of early childhood educators. *Pol. Future Educ.* 17 (5), 618–633.
- Nolan, A., Molla, T., 2018. Teacher professional learning as a social practice: an Australian case. *Int. Stud. Sociol. Educ.* 27 (4), 352–374.
- Nolan, A., Molla, T., 2019. Supporting teacher professionalism through tailored professional learning. *Lond. Rev. Educ.* 17 (2), 126–140.
- Nolan, A., 2021. Addressing inequality: educators responding to the contexts of young children's lives? *Child. Soc.* 35 (4), 519–533.
- Nous, 2020. Universal Access National Partnership Review: Final Review Report. COAG Education Council.
- Oberhuemer, P., 2008. Who is an early years professional? Reflections on policy diversity in Europe. In: Miller, L., Cable, C. (Eds.), *Professionalism in the Early Years*. Hodder Education, London, pp. 131–141.
- O'Connell, M., Fox, S., Hinz, B., Cole, H., 2016. *Quality Early Education for All. Fostering Creative, Entrepreneurial, Resilient and Capable Learners*. Mitchell Report 01/2016. The Mitchell Institute, Victoria University, Melbourne.
- OECD Review Team, 2001. OECD Country Note: Early Childhood Education and Care Policy in Australia. Retrieved on 29 May 2018, from: <http://www.oecd.org/australia/2673543.pdf>.
- OECD, 2001. *Starting Strong: Early Childhood Education and Care*. OECD Publishing, Paris.
- OECD, 2006. *Starting Strong II: Early Childhood Education and Care*. OECD Publishing, Paris.
- Press, F., Hayes, A., 2001. OECD Thematic Review of Early Childhood Education and Care Policy: Australian Background Report. Department of Education, Training and Youth Affairs, Canberra.
- PricewaterhouseCoopers (PwC), 2014. Putting a Value on Early Childhood Education and Care in Australia. September 2014. Viewed 1 February 2020. <https://www.pwc.com.au/pdf/putting-value-on-eccec.pdf>.

- Productivity Commission, 2011. Early Childhood Development Workforce. Commonwealth of Australia, Melbourne.
- Productivity Commission, 2014. Childcare and Early Childhood Learning, vol. 1. Commonwealth of Australia, Canberra, 73.
- Raban, B., Kilderry, A., 2017. Early childhood education policies in Australia. In: Li, H., Park, E., Chen, J.J. (Eds.), Early Childhood Education Policies in Asia Pacific: Advances in Theory and Practice. Springer, Berlin, pp. 1–30.
- Rein, M., Schön, D., 1996. Frame-critical policy analysis and frame-reflective policy practice. *Knowl. Pol.* 9 (1), 85–103.
- Robinson, K.H., 2002. Making the invisible visible: gay and lesbian issues in early childhood education. *Contemp. Issues Early Child.* 3 (30), 415–434.
- Schön, D.A., Rein, M., 1994. Frame Reflection: Toward the Resolution of Intractable Policy Controversies. Basic Books, New York.
- SCSEEC [Standing Council on School Education and Early Childhood], 2012. Early Years Workforce Strategy, 2012–2016. SESEEC, Melbourne.
- Sen, A., 2009. The Idea of Justice. Belknap Press, Cambridge, MA.
- Sollis, K., 2019. Measuring Child Deprivation and Opportunity in Australia: Applying the Nest Framework to Develop a Measure of Deprivation and Opportunity for Children Using the Longitudinal Study of Australian Children. ARACY, Canberra, Australia.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., 2010. Early Childhood Matters: Evidence From the Effective Pre-school and Primary Education Project. Routledge, Abingdon.
- Taylor, C., 2016. Reforming Australian early childhood education and care provision (2009–2015). *Austral. J. Early Childh.* 41 (2), 27–31.
- Torii, K., Fox, S., Cloney, D., 2017. Quality is Key in Early Childhood Education in Australia. Mitchell Institute Policy Paper No. 01/2017. Mitchell Institute, Melbourne.
- Vuckovic, A., 2008. Making the multicultural learning environment flourish. The importance of the child-teacher relationships in educating young children about diversity. *Aust. J. Early Child.* 33 (10), 9–16.
- Wagenaar, H., 2015. Meaning in Action: Interpretation and Dialogue in Policy Analysis. Routledge, London.
- Wilson, C.S., 2016. Cultural Learning for Aboriginal and Torres Strait Islander Children and Young People: Indigenous Knowledges and Perspectives in New South Wales Schools. Unpublished PhD thesis. Southern Cross University, Lismore, Australia.

Teachers and teaching in low-income countries

Daniel A. Wagner^a and **Amy E. Guilloitte^b**, ^a UNESCO Chair in Learning and Literacy, University of Pennsylvania, Graduate School of Education, Philadelphia, PA, United States; and ^b University of Pennsylvania, Graduate School of Education, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	99
International policy framework for teachers	100
The challenge of inclusive education	100
Current context	100
The role of teachers in inclusive access	101
Diversity of teachers	101
Language and teachers	101
Educational quality	102
Qualifications and pre-service training	102
In-service teacher professional development	102
Teacher deployment	102
Contemporary change	102
Educational technology	102
COVID-19 and other mega-trends	103
Conclusions	104
References	104

Introduction

International and local organizations, governments, communities, families, and especially teachers have worked since the founding of the United Nations in 1946 to support education that will transform lives, contribute to building peace, reduce poverty, and provide sustainable development across the globe (UNESCO, 2020). Today, efforts to realize this educational future continue on many fronts, especially through the 2030 UN Sustainable Development Goals (SDGs). The SDGs, adopted in 2015 by all UN member states, are a set of 17 goals “to end poverty, protect the planet and improve the lives and prospects of everyone, everywhere” (United Nations, 2015). Among these, SDG 4 calls for inclusive and quality education for all.

At the same time, the global COVID-19 pandemic has had a wide ranging impact on students, teachers, and education systems globally. Before the pandemic, the United Nations had projected that by 2030 more than 200 million children would be out of school and only 60% of young people globally would complete upper secondary education due to teacher shortages at all levels (United Nations, 2020). This has now been exacerbated by the pandemic, where school closures have affected more than 90% of children worldwide with no clear end in sight as of early 2022. While distance and remote learning has been offered in many countries and communities, it is estimated that more than 500 million children, especially those in low-income countries (LICs), have not had access to digital learning (Vegas, 2020).

The pandemic has also exacerbated existing educational inequities for marginalized students. For example, in low-income countries, 79% of children from the wealthiest quintile of the population complete primary school, while the primary completion rate is only 34% for students from the poorest quintile of the population (United Nations, 2020). In order to address these gaps, governments, global organizations, and development and community organizations have begun to focus on learning obstacles and opportunities for the most marginalized students, or those “learning at the bottom of the pyramid” (Wagner et al., 2018, 2022). In focusing on improving educational access and quality for the most marginalized learners within a country, policy makers can address equity issues and raise levels of learning overall. This joint call for both access and quality for the most marginalized students is an ambitious goal, especially in Africa (Bashir et al., 2018).

When considering these long- and short-term challenges, teachers and teaching will play critical roles. Teachers are key actors in supporting student learning and inclusion, and many broader factors and events shape their work with students in classrooms, especially their work with marginalized students. The current global policy landscape is now oriented beyond access to education, and working toward quality and inclusive learning as, and for, development (Wagner, 2018a; World Bank, 2018).

This review focuses on teachers who work with the most disadvantaged and marginalized learners, across and within low-income countries, and begins with an overview of how national and global policies influence the work of teachers.

International policy framework for teachers

Broadly defined, teachers are people who are, formally or informally, expected to be change agents who support children's learning. They work across varied contexts, such as in preschools, non-formal education programs, religious schools, and primary and secondary schools. When considering teachers as agents of change and of learning, it is important to examine current and emerging policy trends that shape the context and content of teaching and instruction.

Teachers represent, as a group, one of the largest professions in developing countries. UNESCO estimates that as of 2020, there are 71 million primary and secondary teachers worldwide. This is more than a 40% increase from 2000, a growth that matches the increase in numbers of children in school (ITFTE and UIS, 2021). Still, more teachers will need to be recruited to support increasing school-age populations and address teacher attrition, particularly in sub-Saharan Africa. Sub-Saharan African countries will need to recruit 15 million additional teachers by 2030; some countries will need more than a 5% annual increase in the number of teachers in their workforce if they are to meet teacher enrollment targets by 2030 (See Fig. 1; ITFTE and UIS, 2021). Unfortunately, such needs are likely to be accompanied by similarly high rates of teacher attrition.

While there are contemporary challenges that influence teachers and teaching in common ways across the world, there are particular educational factors related to context, systems, and resources that impact low-income countries (LICs). In shifting the conversation to attend to global variation between and within countries, international organizations and national policy makers need to refocus attention on the most marginalized learners. In addressing teaching and teachers in low-income countries, we seek to highlight some of the current and emerging issues that shape teachers' work in LICs.

Basic education and literacy have been central to a global policy agenda since the ratification of the 1948 UN Universal Declaration of Human Rights, and re-emphasized at UN-organized conferences in Thailand (1990) and in Dakar (2000). While many governments, international organizations, and development agencies have worked toward realizing this human right to universally accessible, free, quality basic education, the goal will not likely be met for many learners in LICs, at least in part due to the need for improvement in teacher training and teaching.

Global educational development work is shaped in large part by international support for SDG 4, with specific targets for improved learning outcomes for students, with an emphasis on teacher training and improved teaching.

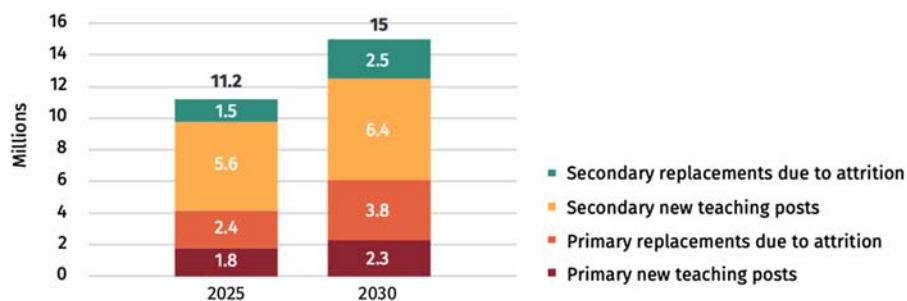
The challenge of inclusive education

Current context

UNESCO's 2020 Global Education Monitoring Report focuses on issues related to inclusion, highlighting the enduring challenges of equity in access to quality education within and across countries globally (UNESCO, 2020). While the SDGs draw attention to equity, inclusion, diversity, and non-discrimination, many learners remain excluded from educational systems. Though progress has been made in improving access to education, even before the 2020 COVID-19 pandemic 258 million children were out of school and remain excluded from a quality education, about 17% of children, adolescents, and youth globally.

Further, when educational development goals focus on national average improvements across education indicators, they can and do miss important variations in learning within countries (Wagner et al., 2018). For example, even though progress has been made in national average school completion rates in LICs, students within the wealthiest quintile in LICs are three times as likely as those from the poorest quintile to complete lower secondary schooling (UNESCO, 2020). Conversely, children living in poverty are the most likely to be disadvantaged when it comes to school access. And in sub-Saharan Africa, nearly 49% of children are affected by extreme poverty.

Migration and displacement impact inclusive education in multiple ways. Nearly half of all child refugees worldwide (especially those who reside in conflict zones) are out of school (UNHCR, 2021). Refugee learners often experience substantial gaps in their



Note: Projections based on UNESCO Institute for Statistics data and methodology in UIS. 2010. [Technical Paper No.3](#).

Fig. 1 Projected teacher recruitment needs in sub-Saharan Africa to meet universal primary and secondary education, 2025 and 2030 Reproduced from ITFTE and UIS (2021).

education due to displacement, and inadequate funding and policies in host country policies (United Nations Stats, 2018). Learners with disabilities are often excluded from formal schooling; 40% of learners with disabilities are out of school in LICs countries at the primary level, and this rises to more than 55% at the secondary level (Global Partnership for Education, 2020). Disadvantaged students – including learners who experience marginalization based on wealth, gender identity and sexual orientation, disability, ethnicity, remoteness, language, displacement, and religion – continue to experience exclusion from formal schooling, even in contexts where overall national progress in access to education has been made; see, for example, the case of Mexico (Schmelkes del Valle et al., 2022).

The role of teachers in inclusive access

It is important to consider how increased school enrollments (driven in part by UN goals), combined with budget constraints in LICs, impact teachers and their work. Countries that expand primary and secondary education access – without concomitantly increasing their teacher workforces – often experience higher pupil-teacher ratios, resulting in lower quality of education (EFA GMR, 2010). In response, some countries reduce entry requirements for teacher training programs or hire contract teachers to meet demand (UNESCO-IIEP, 2021). While the global pupil-teacher ratio in primary education is currently 27:1, in sub-Saharan Africa it is 56:1, and 38:1 in South Asia (ITFTE and UIS, 2021). Larger class sizes negatively impact teaching quality due to limitations on teachers' time with students and families, as well as challenges in classroom management, student participation and engagement, and individualized instruction.

Even as greater numbers of children and youth have access to formal schooling across the globe, inclusion also implies that every learner feels respected and valued, and has a clear sense of belonging. (UNESCO, 2020). Teachers need to be able to ensure that students feel welcomed and included in classroom and school communities, and are able to access content in languages they speak and that builds on what they already know. Further, pedagogical and curricular choices need to be relevant to students and provide opportunities for them to engage in the work of the classroom. To achieve this inclusive form of education, teachers themselves need to represent, as much as possible, the diversity of the students themselves.

Diversity of teachers

Who are the teachers in LICs, and how do their identities and experiences influence inclusive education? Women comprise the majority of the global teaching force; in LICs 88% of pre-primary teachers are female, decreasing to only 41% of primary teachers, and 23% of secondary teachers (ITFTE, UIS, and GEMRT, 2019). This pattern has potential long-term implications for inclusive and positive educational experiences, especially for girls. In addition, refugee educators may be better able to create classroom environments that are perceived by refugee students as more supportive than classrooms of host country national teachers (Richardson et al., 2018). Further, there are more than 240 million children with disabilities globally; and teachers with disabilities themselves have been shown to be able to provide a more welcoming context (UNICEF, 2021). Broadly speaking, research suggests that teacher diversity in terms of social identities – for example, by gender, ethno-linguistic categories, and immigration status) can support students' learning and sense of belonging in the classroom (Evans and Le Nestour, 2019; Gershenson et al., 2017).

Clearly, professional training is designed to overcome common differences between teachers and students. But research shows that about 70% of teachers at the primary school level in LICs have not received formal training, and these teachers primarily work in schools serving disadvantaged learners (UNESCO, 2014). Others have observed primary school classrooms in LICs, and found that students received just under 3 hours per day of teaching, which was about half of scheduled instructional time, and that few teachers evaluated students' progress, checked for student understanding, or gave feedback to students (Bold et al., 2018). Effective teachers can be one of the most important in-school factors to improve student learning (Wagner, 2018a). Diversifying the teacher workforce and providing teacher professional development remain significant challenges for education systems globally, and especially in LICs.

Language and teachers

Language is a critical factor that influences inclusive and effective teaching and learning. The changing demography of student populations, as described above, intersects with the languages used at home and in classrooms (Dryden-Peterson, 2011). There is now growing consensus that children should begin school by learning in their own mother-tongue (World Bank, 2021). This has led to an increased focus on language-in-education policies and practices to meet the changing needs of multilingual and linguistically diverse student populations. Teachers need to be prepared to work in multilingual settings and meet the needs of multilingual learners in their contexts.

This is no easy task for teachers. For example, international policymakers (such as at the UNHCR) recommend including refugees as teachers in host country national education systems to foster integration. But managing the integration of refugees and displaced learners in host country schools, while at the same time promoting the use of mother-tongue instruction, is a major barrier for many teachers. Teachers have to consider their own language practices, students' language practices, and the language policies, resources, and goals of the school, curricula, and community when they teach multilingual students.

Educational quality

Qualifications and pre-service training

To achieve SDG 4, teachers will need training and support. Yet, within the present context of increased demand for quality teachers, along with budget constraints, and growing enrollments, trained teachers have become a scarce commodity. Teachers in LICs, in particular, hold low levels of qualifications; only 70% of primary school teachers and 64% of secondary school teachers in LICs hold minimum qualifications, considerably lower than teachers in middle- or high-income countries (ITFTE 2030 and UIS, 2021). In LICs, most teacher training occurs in vocational upper secondary education, post-secondary, non-tertiary settings, or in short-cycle tertiary settings (typically 2–3 years). However, the possession of a minimum educational qualification is not a highly reliable predictor of teaching quality since a teacher's pre-service program alone does not ensure teaching quality (Akyeampong, 2022).

In-service teacher professional development

As has often been said, teachers need to be learners themselves (Feiman-Nemser, 2012). And, they need ongoing support to do this challenging work, especially if they are going to be prepared to work with diverse learners and provide quality and inclusive education. Teachers in LICs often report limited opportunities for in-service teacher professional development (TPD) programs, especially as it relates to working with marginalized learners (Akyeampong, 2022). During the COVID-19 pandemic in 2020–21, only 29% of countries in sub-Saharan Africa provided professional development to teachers during school closures and the return to schools process (ITFTE and UIS, 2021). In examining TPD programs across low and middle-income countries, teacher TPD programs that are associated with significant growth in student learning have shared various characteristics, including: linking career opportunities (improved teacher status, salary, or promotion) to TPD participation; targeting TPD toward early career teachers; providing complementary curricula materials to TPD participants to use, focusing on a subject-area (instead of general teaching); providing consecutive days of in-person training and follow-up support; and creating opportunities for teachers to practice lesson enactment during the TPD (Popova et al., 2018). There is significant overlap with research in the US on important features for effective professional learning globally (Darling-Hammond et al., 2017). In-service TPD has the potential to support teachers in improving the quality and inclusivity of their teaching practices.

Teacher deployment

The inequitable deployment and allocation of trained teachers also impacts educational quality and inclusion of marginalized learners in the classroom. In order for LICs to provide quality learning to all, schools serving marginalized and disadvantaged students need equitable allocations of trained teachers. For example, Bashir et al. (2018) describe how the Burundian government increased the positioning of trained teachers to rural and under-resourced schools. Subsequently, students from these schools made gains in numeracy scores, relative to their peers in schools with relatively greater levels of resources. While challenges persist, LIC policy makers can provide improved quality education through more carefully designated deployment of teachers.

Contemporary change

Educational technology

Today over half of the total global population uses the Internet; however in LICs this drops to 20% (ITU, 2020). Additionally, significant variation in school-based access to electricity, computers, and Internet in LICs persists, especially in rural schools and at primary grade levels (UIS, 2015). Information and communications technologies (ICTs) can have significant benefits for learning, in and out of schools (Wagner, 2018b), but ICTs also raise many questions about the possibilities and limitations of technology in teaching and learning in LICs.

While interest and scrutiny of ICTs in education has dramatically increased during the global COVID-19 pandemic, policymakers, researchers, and educators continue to ask questions such as: How can ICTs help to improve teaching, learning, and educational quality, and which ICTs should schools and teachers adopt (Hennessy et al., 2021)? What are the consequences for equity and the “digital divide” when ICTs are taken up by some teachers and learners, but not others? How are ICTs relevant for teachers in assessing student learning, communicating with families, working within multilingual contexts, and for teachers' own learning?

These are questions that researchers, policymakers, and educators continue to grapple with as globalization and technological advances continue. Particularly salient examples of the potential for ICT-assisted learning in LICs today include ICTs that can support teachers in schools with high linguistic diversity where teachers are not fluent in the multiple languages spoken by students in classrooms (Wu et al., 2021).

In some settings, ICT-assisted interventions can support teachers working with multilingual learners. For example, Castillo and Wagner (2019) carried out a program in low-performing, rural South African schools that showed promise for supporting students' in early reading development in both English and indigenous languages. Students used a software program in school computer labs to move through self-paced lessons, resources, and practice activities in English and local indigenous languages. This type of ICT use can support teachers in improving the quality of the education that multilingual, language minoritized students receive in their

classrooms (especially when teachers themselves do not share the same linguistic repertoires as all of their students). More generally, such efforts can improve educational attainment for historically marginalized students, such as indigenous and non-dominant language speakers.

ICTs can also be used to support teacher learning and development in settings where teachers have limited access to training and support, such as in refugee-serving and conflict-affected contexts. Mendenhall (2018) reviewed, for example, a teacher development program called for Teachers for Teachers in Kakuma, a refugee camp in Kenya where almost 200,000 refugees from Somalia, South Sudan, and 20 other countries live (UNHCR, 2020). This program used a combination of face-to-face training and peer coaching, in combination with group mentoring via WhatsApp and mobile devices (Mendenhall, 2018). Researchers found that teachers reported improved knowledge and skills, increased motivation and self-confidence, and improved teacher-student relationships. Clearly, ICTs have the potential to support inclusive and quality educational opportunities for refugee teachers and learners through supporting teacher mentorship and learning through digital tools (UNESCO, 2018).

Such efforts using ICTs are particularly salient because of their use with historically marginalized learners in under-resourced settings, rural, indigenous language-speaking students and refugee educators in camp settings. Many LICs have increasingly multilingual populations and are home to significant numbers of displaced and conflict-affected students and teachers. There is little doubt that ICTs will be increasingly deployed in the coming years.

COVID-19 and other mega-trends

The COVID-19 pandemic continues to affect, teachers, teaching, and students globally. During the start of the pandemic in 2020, about 90% of all students (1.57 billion children and youth) were out of school worldwide. Only 25% of LICs offered remote learning to students, primarily through broadcast television and radio programming, compared to almost 90% of high-income countries, which mostly utilized a variety of digital learning approaches (Vegas, 2020).

With schools closures during the pandemic, teachers in many countries pivoted to lead or support remote learning. However, in LICs, where education systems relied primarily on distance radio and television broadcasting to deliver remote learning, teachers received less guidance and support on how to engage students. For example, in sub-Saharan Africa, during the early days of the COVID-19 pandemic in spring 2020, no sub-Saharan African countries provided teacher training on how to best support learners during the closure of school buildings (Vegas, 2020). When teachers are not included in distance learning plans, students may not receive tailored instructional support, instead relying on broadcast TV and radio programming (if they can access these at all).

Even so, some programs provided teachers with remote learning support during the pandemic. The Speed School, an accelerated primary school program for out-of-school learners age 8–14 in Uganda and Ethiopia, has demonstrated success by providing support for teachers to conduct weekly one-to-one visits to students homes or lead “micro-classes” of up to five students in the community, in order to check in with students and parents, provide worksheets and written instructions, and support independent learning (Muskin, 2020). Teacher inclusion in distance learning plans in low-income countries is critical for students to receive individualized support in times of uncertainty. In turn, teachers themselves need support to change how they teach during times of crisis and change.

The prioritization of teachers as central figures who can support education recovery post-pandemic as students return to classrooms has varied widely. Fewer than 25% of sub-Saharan African countries reported recruiting additional teachers to support school re-opening efforts (ITFTE and UIS, 2021). Investments in sustaining and growing the teacher workforce in LICs will be critical to address unfolding demographic shifts, teacher attrition, and increased student need due to the global pandemic, even as two-thirds of low and lower middle-income countries have reduced education budgets since the start of the COVID-19 crisis (ITFTE and UIS, 2021). Impacts of the global pandemic on learners, teachers, and teaching in LICs are yet to be fully understood, but going forward, the international policy community and national governments should consider the potential roles of teachers in remote learning and student support.

In addition to the COVID-19 pandemic, other mega-trends such as globalization, migration, climate change, and sustainability will continue to impact the local, classroom-level work of teachers and teaching in LICs. Globalization has led to advances in science, technology, information flows, transportation, and commerce, which shape education in many overlapping ways. Globalization also shapes teaching and teachers’ work more directly, through the dissemination and promotion of curricula and education initiatives through UN organizations, the OECD, and international organizations, as well as by shaping families’ and societal expectations for learning for 21st century employment (Burn and Menter, 2021; Sorensen and Dumay, 2021; Wagner et al., 2020). For the most marginalized learners in LICs, globalization has also created and revealed inequities in education systems, as learners at the bottom of the pyramid fall further behind shifting international norms (Wagner et al., 2022).

UNHCR estimates that almost 80 million people are forcibly displaced worldwide, and displacement mainly affects developing countries and LICs. Over 20% of the global refugee population reside in LICs, which are home to only 10% of the world’s population, and more than half of the forcibly displaced are under the age of 18 (UNESCO, 2019) as discussed above. This has significant implications for multilingual, inclusive, and quality teaching and teachers’ work.

Additionally, environmental sustainability (or instability) and climate change are major policy priorities across the globe today, and much remains to be done to refocus educational efforts to center environmental adaptation. Climate change is a particular threat in LICs, through damage from weather-related disasters and ecological changes, alongside limited resources to support adaptations and mitigation. Teaching content and curricula about climate change and environmental sustainability is limited across the

globe (Opertti and Brylinski, 2016), and needs to better support students in climate action (Kwauk, 2020). Each of these megatrends needs to be considered as we look to the future of teaching and teachers' work in LICs, in our interconnected world.

Conclusions

As the 1948 Declaration of Human Rights envisioned, education can be transformative for individual learners across their lifespan, and also foster inclusion, equity, peace, development, and understanding within and across communities and countries. These ambitious goals are challenged by persistent inequities based on ethnolinguistic difference, gender identities, socioeconomic status, disability, immigration status, and religion that are deeply entrenched in our global, national, and local education systems. Teachers can and will play roles that impact on societies' ability to sustain and improve education systems.

To address the complex challenges of inclusive and quality education for all, it is critical to understand: (a) the important roles of teachers and teaching within classrooms in LIC contexts; (b) the policies and supports that teachers need to do this their work; and (c) the ways that global megatrends shape every teacher's contexts for teaching. Teachers have the potential, and obligation, to be change agents in finding paths to achieve the UN global goals.

References

- Akyeampong, K., 2022. Teaching at the bottom of the pyramid: teacher education in poor and marginalized communities. In: Wagner, D.A., et al. (Eds.), *Learning, Marginalization, and Improving the Quality of Education in Low-Income Countries*. Open Book Publishers, London.
- Bashir, S., Lockheed, M., Nihan, E., Jee-Peng, T., 2018. Facing Forward: Schooling for Learning in Africa. World Bank.
- Bold, T., Filmer, D., Martin, G., Molina, E., Rockmore, C., Stacy, B., Svensson, J., Wane, W., 2018. What Do Teachers Know and Do? Does it Matter? Evidence from Primary Schools in Africa. Background Paper to the 2018 World Development Report. World Bank Group, Washington, DC.
- Burn, K., Menter, I., 2021. Making sense of teacher education in a globalizing world: the distinctive contribution of a sociocultural approach. *Comp. Educ. Rev.* 65 (4), 770–789. <https://doi.org/10.1086/716228>.
- Castillo, N.M., Wagner, D.A., 2019. Early-grade reading support in rural South Africa: a language-centered technology approach. *Int. Rev. Educ.* 65, 389–408.
- Darling-Hammond, L., Hyster, M.E., Gardner, M., 2017. Effective teacher professional development. Learning Policy Institute.
- Dryden-Peterson, S., 2011. Refugee education: a global review. UNHCR. <https://www.unhcr.org/research/evalreports/4fe317589/refugee-education-global-review-sarah-dryden-peterson-november-2011.html>.
- Education For All Global Monitoring Report (EFA GMR), 2010. Education for All Global Monitoring Report 2010: Reaching the Marginalized. UNESCO. Available at: <http://www.unesco.org/en/efareport/reports/2010-marginalization/>.
- Evans, D., Le Nestour, A., 2019. Are female teachers better for girls' education? Center For Global Development. <https://www.cgdev.org/blog/are-female-teachers-better-girls-education>.
- Feiman-Nemser, S., 2012. Teachers as Learners. Harvard Education Press, Cambridge, MA.
- Gershenson, S., Hart, C.M.D., Lindsay, C.A., Papageorge, N.W., 2017. The Long-Run Impacts of Same-Race Teachers. IZA Institute of Labor Economics, Deutsche Post Foundation. <https://ftp.iza.org/dp10630.pdf>.
- Global Partnership for Education, 2020. Inclusive education for children with disabilities. <https://www.globalpartnership.org/sites/default/files/document/file/2020-10-GPE-factsheet-inclusive-education.pdf>.
- Hennessy, S., Jordan, K., Wagner, D.A., 2021. Problem Analysis and Focus of EdTech Hub's Work Technology in Education in Low- and Middle-Income Countries. Working Paper. UKaid/World Bank, London, pp. 1–73. https://www.repository.cam.ac.uk/bitstream/handle/1810/319313/Problem%20Analysis_Technology%20in%20Education_Revised%20Final_1.pdf?sequence=1.
- International Task Force on Teachers for Education 2030 and UNESCO Institute for Statistics (ITFTE and UIS), 2021. World teachers' day 2021 fact sheet. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000379187.locale=en>.
- International Task Force on Teachers for Education 2030, UNESCO Institute for Statistics, and Global Education Monitoring Report Team (ITFTE, UIS, and GEMRT), 2019. World teachers' day 2019 fact sheet. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000370921>.
- International Telecommunications Union (ITU), 2020. Press Release: household Internet access in urban areas twice as high as in rural areas. ITU. <https://www.itu.int:443/en/mediacentre/Pages/pr27-2020-facts-figures-urban-areas-higher-internet-access-than-rural.aspx>.
- Kwauk, C., 2020. Roadblocks to Quality Education in a Time of Climate Change. Brookings Institution, Washington, DC.
- Mendenhall, M., 2018. Teachers for teachers: advocating for stronger programs and policies for and with refugee teachers in Kakuma refugee camp, Kenya. *Studies in Social Justice* 12 (2), 8.
- Muskin, J., 2020. COVID-19 and education: emphasizing the journey, not the destination. Geneva Global. <https://www.genevaglobal.com/blog/covid-19-and-education-emphasizing-the-journey-not-the-destination>.
- Opertti, R., Brylinski, E., 2016. Global Monitoring of Target 4.7: Themes in National Curriculum Frameworks. GEM Background Paper. UNESCO-IBE, Paris.
- Popova, A., Evans, D.K., Breeding, M.E., Arancibia, V., 2018. Teacher Professional Development Around the World: The Gap between Evidence and Practice [Working Paper]. World Bank. <https://doi.org/10.1596/1813-9450-8572>.
- Richardson, E., MacEwen, L., Naylor, R., 2018. Teachers of refugees: a review of the literature. Education Development Trust.
- Sorensen, T.B., Dumay, X., 2021. The teaching professions and globalization: a scoping review of the anglophone research literature. *Comp. Educ. Rev.* 65 (4), 725–749. <https://doi.org/10.1086/716418>.
- Schmelkes del Valle, S., Robles Vásquez, H., Santos, A., 2022. Mexico: education and learning at the bottom of the pyramid. In: Wagner, D.A., Castillo, N.M., Grant Lewis, S. (Eds.), *Learning, Marginalization, and Improving the Quality of Education in Low-Income Countries*. Open Book Publishers, London.
- UNESCO, 2014. Global education monitoring report, 2013–2014: teaching and learning: achieving quality for all. <https://unesdoc.unesco.org/ark:/48223/pf0000225660>.
- UNESCO, 2018. A lifeline to learning: leveraging mobile technology to support education for refugees. <https://unesdoc.unesco.org/ark:/48223/pf0000261278>.
- UNESCO, 2019. Global education monitoring report, 2019: migration, displacement and education: building bridges, not walls—UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000265866>.
- UNESCO, 2020. Global Education Monitoring Report, 2020: Inclusion and Education: All Means All. <https://unesdoc.unesco.org/ark:/48223/pf0000373718>.
- UNESCO Institute for Statistics (UIS), 2015. Information and Communication Technology (ICT) in Education in Sub-Saharan Africa: A Comparative Analysis of Basic E-Readiness in Schools. <https://doi.org/10.15220/978-92-9189-178-8-en>. *Information Paper No. 25*.

- UNESCO-IIEP, 2021. Teacher recruitment and retention. <https://learningportal.iiep.unesco.org/en/issue-briefs/improve-learning/teacher-recruitment-and-deployment>.
- UNHCR: The UN Refugee Agency, 2020. Kakuma Refugee Camp and Kalobeyei Integrated Settlement. UNHCR Kenya. <https://www.unhcr.org/ke/kakuma-refugee-camp>.
- UNHCR: The UN Refugee Agency, 2021. UNHCR Education Report 2021: "Staying the Course" - the Challenges Facing Refugee Education. UNHCR. <https://www.unhcr.org/publications/education/612f85d64/unhcr-education-report-2021-staying-course-challenges-facing-refugee-education.html>.
- United Nations, 2015. The Sustainable Development Agenda. <https://www.un.org/sustainabledevelopment/development-agenda/>.
- United Nations, 2020. The Sustainable Development Goals Report: 2020. <https://unstats.un.org/sdgs/report/2020/>.
- United Nations Stats, 2018. Expert Group on Refugee and Internally Displaced Persons Statistics: International Recommendations on Refugee Statistics. Retrieved on November 7, 2021 from. <https://unstats.un.org/unsd/statcom/49th-session/documents/BG-Item3m-RefugeeStat-E.pdf>.
- UNICEF, 2021. Seen, Counted, Included: using data to shed light on the well-being of children with disabilities. <https://data.unicef.org/resources/children-with-disabilities-report-2021/>.
- Vegas, E., 2020. School closures, government responses, and learning inequality around the world during COVID-19. Center for Global Development. <https://www.cgdev.org/blog/are-female-teachers-better-girls-education>.
- Wagner, D.A., 2018a. Learning as Development: Rethinking International Education in a Changing World. Routledge.
- Wagner, D.A., 2018b. Technology for education in low-income countries: supporting the UN sustainable development goals. In: Lubin, I. (Ed.), ICT-supported Innovations in Small Countries and Developing Regions: Perspectives and Recommendations for International Education. Springer, New York, NY.
- Wagner, D.A., Wolf, S., Boruch, R.F. (Eds.), 2018. Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries. UNESCO International Institute for Educational Planning. <http://unesdoc.unesco.org/images/0026/002655/265581E.pdf>.
- Wagner, D.A., Castillo, N., Zahra, F.T., 2020. Global learning equity and education: looking ahead. In: Paper Commissioned for the UNESCO Futures of Education Report. UNESCO, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000375000>.
- Wagner, D.A., Castillo, N.M., Grant Lewis, S., 2022. Learning, Marginalization, and Improving the Quality of Education in Low-Income Countries. Open Book Publishers, Cambridge, UK. <https://books.openbookpublishers.com/10.11647/obp.0256.pdf>.
- World Bank, 2018. Learning to Realize Education's Promise. World Development Report 2018. The World Bank, Washington DC.
- World Bank, 2021. Loud and Clear: Effective Language of Instruction Policies for Learning. World Bank, Washington, DC.
- Wu, A., Gonsalves, S.-L.R., Wagner, D.A., 2021. The impact of technology on a multilingual world: problems and Opportunities. In: Lubin, I. (Ed.), ICT and International Learning Ecologies: Representation and Sustainability across Contexts. Routledge, New York.

Narrative and biography in teacher work lives and development[☆]

Geert Kelchtermans, Center for Innovation and the Development of Teacher and School (CIDTS) – KU, Leuven, Belgium

© 2023 Elsevier Ltd. All rights reserved.

This is an update of G. Kelchtermans, Narratives and Biography in Teacher Education, in International Encyclopedia of Education (Third Edition), Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier Ltd., 2010, <https://doi.org/10.1016/B978-0-08-044894-7.00666-7>.

The narrative – biographical perspective	106
The promise and potential of narratives	107
Teacher development in action and thinking: the personal interpretative framework	108
Narrative and biography in pedagogical practices for teacher development	109
The critical incident vignette: a storied example	109
The rationale in narrative pedagogies	110
References	111
Further reading	112

Becoming a teacher doesn't end with the graduation from teacher education. Teachers continue to learn and develop based on their daily experiences in the job. Apart from this informal learning, most teachers will over time also participate in more formal training and learning opportunities (in-service training, school development programs, etc.). The ongoing character of this learning throughout teachers' work lives is highlighted in the concept "continuing professional development" (see e.g. Day and Sachs, 2004).

Yet, teacher development doesn't start with initial teacher training either. Student teachers do not enter teacher education as blank slates, but rather bring with them about 15 years of personal experiences in schools. They have lived and learned in many classrooms, experienced dozens of different teachers, and thus inevitably have built a personal set of ideas and beliefs about what it means to be a teacher, to be teaching, as well as to be a pupil or a student.

Hence teacher development can be defined as teachers' lifelong learning and development that results from their meaningful interactions with their professional context. Understanding this process is of crucial importance for research and theory building on teaching, teacher education and school improvement, but also for practices of the curriculum design and the pedagogical enactment of professional learning opportunities in both initial teacher education and in-service training. Unpacking teacher development implies answering central questions like: What does it mean to live a teacher's life? What does it mean to take on a job as a teacher, to stay in teaching, or to leave the profession? Why are teachers doing what they are doing the way they are doing it? How does teachers' expertise and motivation develop over time and how does this development affect their daily professional practices? In my scholarly work I have addressed these questions through the lens of what I have called the *narrative-biographical perspective* (Kelchtermans, 1993, 2009).

The narrative – biographical perspective

When I started my study of teacher development in the mid-80s, the so-called "teacher thinking"-research had gained momentum (see e.g. Clark and Peterson, 1986; Craig et al., 2013). In line with the developments in cognitive psychology, educational researchers had become convinced that in studying teaching as well as the training and development of teachers a central place had to be given to the cognitions and beliefs that guided their judgments and actions. This line of work stressed the "importance and significance of the pre-active, interactive and post-active dimensions of teachers' personal, practical and professional thinking about the events that defined life in and around classrooms, schools and universities." (Kompf and O'Connell Rust, 2013, p. 23).

In their search for relevant conceptual and methodological approaches of this "thinking" and its manifestation in teachers' discourse and practices, several scholars became interested in *narratives*. Since "narrative is the discourse structure in which human action receives its form and through which it is meaningful" (Polkinghorne, 1988, p. 135), narratives were considered as a powerful way to unravel and understand the complex processes of sense-making that constitute teaching. Teachers' talk about their professional lives and practices is very often spontaneously framed in narrative form. They use anecdotes, metaphors, images, and other types of storytelling to recall, share, exchange, or account for their experiences in classrooms and schools. Storytelling is the natural

[☆]*Change History:* April-June 2021. The text has been restructured and rewritten in a more academic and personal style. The latter implies that I explicitly present an argument as author, based on my own work, while situating it in the wider international research interest and literature on narrative/biography in teacher development research and practice. The introduction and the first paragraph are almost entirely new. §2 is an extended version of the paragraph entitled "Narrative language, context, and voice". §3 is added and a new paragraph. §4 is a slightly rewritten and updated version of the paragraphs "The critical incident vignette", §4.2. retakes most of the content of "Narrative practices with student teachers", as well as large chunks from "Narratives, teacher thinking and professional identity".

way through which people make sense of the events, situations, and encounters they find themselves in. “Humans are storytelling organisms who, individually and socially, lead storied lives. The study of narrative, therefore, is the study of the ways humans experience the world” (Connelly and Clandinin, 1990, p. 2). This evolution came to be known as the *narrative turn* and it constituted an important theoretical, epistemological and methodological development in the way teachers and their work were conceived of (see for overviews: e.g. Carter, 1993; Carter and Doyle, 1996; Casey, 1995; Clandinin, 2007).

Both parallel to and in response to the narrative turn, researchers with an interest in teacher development explicitly linked this approach to teachers’ *biographies*. Especially the revival of the life history research (Ball and Goodson, 1985; Goodson, 1984, 1992) as well as studies of teachers’ careers and work lives (Huberman et al., 1993; Nias, 1989; Sikes et al., 1985) created rich opportunities for conceptual cross-fertilization.

As a theoretical approach the narrative-biographical perspective is characterized by a narrative, constructivist, interactionist, contextualized and dynamic stance (Kelchtermans, 1993). The *narrative* aspect refers to the central role of stories and storytelling in the way teachers deal with the professional experiences throughout their career. Narrative accounts of these experienced interactions reveal the particular meaning they have or had for the people involved. In fact, the perspective focuses on teachers’ work lives as they are lived in time and is less interested in conceiving of it in terms of a formal career (the chronological list of positions a teacher takes up over the years). The concept of work lives explicitly acknowledges the relevance of the sense-making by the person involved and as such rather looks at what could be called the “subjective career”, i.e. teachers’ personal experiences in their professional lives over time (Kelchtermans, 2009). In other words, teachers’ “subjective careers” are reflected in and represented through their narrative accounts of those experiences. *Biographical* refers to the fact that teachers –like all human beings–live their lives between birth and death (see e.g. Goodson, 1984; 1992; Kruger and Marotzki, 1996). Human existence is fundamentally characterized by historicity: people are born at some point in time, live during a particular number of years, and then they pass away. Biographical narratives reflect that in human lives and development context has a temporal as well as a spatial dimension. In these narrative accounts there is always a when and a where. The “when” refers to a particular moment or period in time. The “where” includes the organizational, institutional, political, social, cultural and material environments and conditions teachers work in (e.g. the school with its particular staff, school population, infrastructure, embedded in particular ways in the local communities and in a particular prevailing policy environment – see e.g. Kelchtermans, 2017).

Because human beings are gifted with the capacity to remember and make sense of past experiences, their interpretations, thoughts, and actions in the present are influenced by their experiences from the past and expectations for the future. This influence, however, is not to be thought of in a deterministic or causal sense, but rather rests on complex processes of interpretation and sense-making.

Interpretation and sense-making as crucial for understanding human actions runs as a red thread throughout the narrative-biographical perspective. Through collecting “career stories” or “professional biographies” (Kelchtermans, 2009) it seeks to capture the meaning teachers attribute to their experiences in the job over time. In fact, the story-telling is itself seen as an active and constructive process of sense-making: teachers actively (re)construct their experiences into a narrative that makes sense to them. Career stories are constructed in the act of telling: they are told and can be retold. Their importance and relevance lies not so much in their historical truth, but rather in their power to reveal the particular meaning events have had for the teacher.

This further reflects an anti-deterministic stance: the person is the story of his/her life as s/he tells it. The meaning of (career) experiences for a person thus may change over time. This *dynamic* aspect acknowledges that we don’t need to look for stories as the one (historically) true representation of one’s experiences, but that stories are acts of sense-making and thus may be re-storied over time. In other words, the narrative-biographical perspective on teachers’ work lives “is centered on the practical understandings that teachers develop as they enter into and begin to teach and on the ways in which beginning and/or experienced teachers come to frame their understandings within their life stories or life experiences” (Carter and Doyle, 1996, p. 129).

It is obvious that this meaning making is rooted in the *interactions* with others and the context. Teachers are not only actors, but inevitably also inter-actors, who cannot but engage with others in particular professional contexts.

The promise and potential of narratives

Although teachers’ professional actions are, to an important degree, determined by the structural conditions of policy regulations, institutional environments, and organizational settings, the narrative turn explicitly reminds us that teaching and learning to teach cannot be properly understood without acknowledging the obvious role of *agency*. Teachers are capable of perceiving and interpreting a professional situation giving meaning to it and identifying what issues are at stake. This further implies the capacity to judge the situation, to consider and deliberate between several alternative actions and to eventually decide on a response and enact it in practice. Educating children thus involves planning, knowledgeable deliberation, purposeful intervention, skillful action, as well as the capacity and willingness to take responsibility for it. In other words, teaching as a professional action requires and is constituted by *expertise*.

However, because the ‘objects’ of this action are unique human beings and its eventual goal their individual education and formation, teachers’ work is never entirely instrumental. It is inevitably also a moral endeavor of personal involvement and commitment with the educational needs of the ones that are entrusted to the teachers’ care (see e.g. Fenstermacher, 1990; Noddings, 1984). As a consequence, apart from expertise, professional teaching also requires ethical *engagement or commitment*. Because of this close

connection of expertise and engagement a technical, objective language is too limited to capture the full experience and responsibility of teaching. Because of a series of particular characteristics, narrative language provides ways to get beyond these limits.

First, a core process in narratives is the emplotment: by situating experiences in a narrative, the narrator organizes them in a particular linguistic frame and order (beginning–middle–end) through which the particular meaning of his/her experiences emerges (plot) (see e.g. Brooks, 2003). As such the plot also operates rhetorically and ideologically in that it presents an account of the events that through its coherence and meaningfulness seeks to convince and engage the audience or readership.

Narrative language further performs at the same time a referential and an evaluative function. The referential function describes events and experiences from the past in a temporal order. The evaluative function links the events with the moment of narrating by revealing what the experiences have meant for the people involved in the present (Labov and Waletzky, 1973).

Third, narratives thus always link past experiences to the actual social context. They even constitute that context, because they reflect a need to communicate, to share a meaningful experience with others (Hoepfel, 1983). In the storytelling a particular social setting emerges with different actors and roles. On the one hand, there is the narrator or storyteller who provides the narrative. His/her narrative account allows for the personal, subjective voice to be spoken and heard. Yet, on the other hand, in the act of telling the story there is always an audience already implied. Narratives are told or written by someone, who always by that time has an explicit or implicit idea of possible audiences. As such, the narratives are fundamentally intersubjective or interpersonal in nature. This applies to the direct act of storytelling as well as to different forms of written narrative accounts.

Fourth, the contextualizing characteristic of stories, however, not only includes its spatial dimension, but also the temporal dimension. Stories situate experiences both in space and time. This is most obvious in (auto)biographical stories where experiences from one's personal life are being recalled (life stories). The (auto)biographical stories hence reveal how the narrator's understanding of the present is influenced by experiences from the past as well as his/her expectations about the future.

Fifth, the use of stories and narratives is also valued for its emancipatory power. Narratives allow for the practitioner's voice to be heard and taken seriously in educational research and theory development. Yet, although acknowledging the emancipatory potential, several authors have warned for the trap of subjectivism or a romantic misconception of that emancipation. Creating a space for the voices that are often silenced in educational research is important, but it is equally crucial to recognize that these voices are always interpreted and understood as situated in and thus influenced by political, ideological, institutional, and structural contexts. For that reason authors like Goodson (1992, 2013) have argued for the need to turn life stories into life histories: linking the story to its socio-historical context.

Teacher development in action and thinking: the personal interpretative framework

Teacher development becomes visible in their professional practices as they unfold over the time of their work lives, as well as in teachers' accounts of their actions and experiences. Or -to phrase it in line with the "teacher thinking" legacy-teacher development results in qualitative changes both in actions and in thinking over time (Kelchtermans, 2009).

Based on my own narrative-biographical work I have labeled this thinking with the concept *personal interpretative framework*. This framework encompasses "a set of cognitions, of mental representations that operates as a lens through which teachers look at their job, give meaning to it and act in it. This framework thus guides their interpretations and actions in particular situations (context), but is at the same time also modified by and resulting from these meaningful interactions (sense-making) with that context. As such it is at the same time both a *condition for and a result of* the interaction, and represents the –always preliminary – 'mental sediment' of teachers' learning and developing over time" (Kelchtermans, 2009, pp. 260–261).

The systematic analysis of the data from biographical interviews (Kelchtermans, 1994) led me to conclude that the personal interpretative framework encompasses two closely interwoven domains: the professional self-understanding and the subjective educational theory (Kelchtermans, 2009). The *professional self-understanding* refers to teachers' sense of identity as teachers, the way they conceive of and represent themselves in their professional practice. Within this self-understanding five components could be distinguished: self-image, self-esteem, task perception, job motivation and future perspective. The *self-image* is the descriptive component, the way teachers characterize themselves as teachers. *Self-esteem* refers to the evaluative component of the self-understanding, reflecting one's appreciation of one's actual job performance (how well am I doing in my job as a teacher?). Thirdly, the normative component –the *task perception*–constitutes a teacher's personal professional agenda: what must I do to be a proper teacher?; what are the duties I have to take on if I want to rightly feel that I am doing well as a teacher?; what do I consider as legitimate tasks to perform and what do I refuse to see as part of 'my job'? This normative component clearly reflects the idea that teaching and being a teacher is not a neutral, technical endeavor, but implies value-laden choices, moral considerations, and ethical stances. As a fourth component the *job motivation* (or conative component) refers to the motives that make people choose to become a teacher, to stay in teaching or to give it up for another career. The job motivation is strongly determined by the agenda in the task perception as well as by the degree to which the specific working conditions in one's school permit to operate according to that personal normative program. Finally the *future perspective* encompasses teachers' expectations about their future in the job ('how do I see myself as a teacher in the years to come and how do I feel about it?'). In this last component the dynamic character of the self-understanding is explicitly emphasized: teachers' understanding of themselves as teachers is the result of ongoing processes of interaction and sense-making over time.

Although the five components can be analytically distinguished, they need to be understood as intertwined and interdependent, since -as 'components-together they make up the self-understanding. Furthermore, by using the term "self-understanding" I stress

that it refers to both the understanding one has of one's 'self' at a certain moment in time (*product*), as well as to the fact that this product results from an ongoing *process* of making sense of one's experiences and their impact on the 'self'. This is even more strongly expressed in the original word in Dutch (my mother tongue): "zelfverstaan" is at the same time a noun and a verb and as such already captures the idea of both process and product in its morphology. The narrative nature of the concept self-understanding helps to avoid the possible essentialist pitfall in conceptualizing 'identity'. In this view, there should be no looking for a 'deep', 'essential' or 'true' personal core that makes up the 'real' self (Kelchtermans, 2009, 2018).

By the *subjective educational theory* -the second domain in teachers' personal interpretive framework- I mean "the personal system of knowledge and beliefs about education that teachers use when performing their job." (Kelchtermans, 2009, p. 263). It constitutes teachers' professional know-how, on the basis of which they decide how to act or react in a particular situation. The content is a mixture of on the one hand explicit and formalized insights and understandings as derived from training or personal study (knowledge), and on the other hand beliefs referring to more experience-based, idiosyncratic convictions, built up through different career experiences over time. Although this juxtaposition of knowledge and belief suggests that it concerns two strictly separated categories, in reality teachers' expertise is much more a mix and intertwining of both. The actual border between knowledge (grounded in and based on research) and more personal beliefs (based on individual experience, single cases, 'hear say') is not always that clear. The subjective educational theory reflects the teacher's personal answer to the questions: 'how should I deal with this particular situation?' (= what to do?) and 'why should I do it that way?' (= why do I think that action is appropriate now?). The *content* of the subjective educational theory is largely idiosyncratic and based on personal experiences. Formal knowledge (for example from the curriculum of teacher education programs or in-service training) takes root in the subjective educational theory if (student) teachers have experienced that 'it works for them' or 'is found to be true in their practice'. The same applies to the beliefs (e.g. suggestions or rules of thumb inherited from more experienced colleagues). The epistemological status of the subjective educational theory is that its content 'holds true' for the teacher involved (Kelchtermans, 2009, p. 264).

Narrative and biography in pedagogical practices for teacher development

So far I have presented the narrative-biographical perspective in its potential for research and teacher development. However, its rationale can easily be translated into concrete forms of pedagogy (see e.g. Craig et al., 2014, p. 155 ff). Before elaborating on this rationale in practice, I'll start with presenting a short story of one particular narrative-biographical technique in action.

The critical incident vignette: a storied example

The seminar room fills up with chatter as about 20 student teachers enter, move around, and find a seat in the circle of tables and chairs. Throughout the chatter, I pick up excited comments on their experiences during their recently ended internships. After a brief welcome, I announce that the exercise I want them to engage in will encompass three phases. An individual phase, a group phase, and finally a plenary phase that will take place during the next meeting in the seminar.

As a start I ask all of them to take paper and pen for an individual autobiographical reflection. They are asked to recall the best or worst teacher (the choice is theirs) they have ever experienced during their lives as a pupil or a student. The task is to focus on one typical event or experience with that former teacher and write a short narrative account of the event. A possible alternative assignment is to have them reflect back on their best or worst experience in their school lives as a pupil or a student. The assignment is most often reacted on by the student teachers with surprise, some hesitating smiles or raised eyebrows, but they all soon concentrate on the paper. After only a few minutes everyone is busy writing. Deep sighs, grim looks, and also amused smiles can be observed.

As most of them are about to end their writing, I ask them to find a catchy title for the piece. The title should both trigger the reader's curiosity and summarize the key idea that is supposed to become evident from the vignette. Finally, I ask all the participants to turn the sheet and use the back to elaborate in a more systematic and explicit way what the experience or event has meant to them, what constituted its particular significance, and what possible relevance the story does have for their actual work lives as (student) teachers.

At the end of the meeting, instructions are given for the second phase of the exercise. The participants are being asked to meet in groups of three to perform a triangle discussion. They should themselves pick a time and place to meet and should allow for about one and a half hour for the discussion. During that discussion the participants are required to perform and strictly stick to one of the three prescribed roles: storyteller, sounding board, or listener. The first role is that of the storyteller: s/he starts with sharing the title and the story, followed by her/his reflective analysis of its actual significance. Once the storyteller has finished, the sounding board starts to play her/his role. The person in this role can ask clarifying questions about context, particular details, etc., which the storyteller then answers or explains. Comments or interpretations are to be strictly avoided. They are the privilege of the listener, the third role. S/he can suggest different interpretations or open up alternative ways of viewing the experiences from the story. However, all reflections or interpretations have to remain grounded in the story. No noncommittal speculations are allowed. Finally, the storyteller gets a chance to react with final comments.

Once this round is over, everyone changes roles and the cycle is repeated until every participant in the triangle has performed each of the three roles. The strict distinction between the role of sounding board and listener may have an artificial feel, but it is deliberately included to force the participants to suppress their spontaneous tendency to immediately start interpreting and commenting on the story against the background of their own experiences. Postponing this interpretation adds to the necessary active listening as well as to a growing awareness about one's temptation to engage in overhasty commenting.

The third and last phase in the exercise is a plenary discussion during the next seminar meeting and concentrates on the insights that were revealed through the structured storytelling and analysis. I facilitate the discussion in such a way that the focus is on the participants' actual thinking about teaching and about themselves as teachers as it became apparent from the storytelling. Furthermore, the public reflection should also include relevant elements of the professional context and the relationship between that context and the narrator. A climate of trust and safety is established for the discussion and all stories and reflections are, in principle, being accepted and valued. At the same time, however, critical questions are being asked, alternative interpretations of the incident provided, different types of arguments juxtaposed and commented on. Throughout the discussion, links are made to relevant theoretical perspectives and concepts to (re)frame issues, or to deepen the reflective analysis.

This brief account of the critical incident vignette, as I developed and used in both pre-service and in-service training, is meant as an example of how biographical narratives can be purposefully used to stimulate teacher development. It allows the participants to experience how a personal anecdote reveals not only its meaning and significance at the time when it happened, but also -and more importantly- its relation with the present. The stories from the past reveal a lot about the way the storyteller thinks about her/himself in the present, especially when used with experienced teachers or with student teachers in the final phase of their training (after internship experiences). The context (for e.g., other people in the scene, the organizational conditions, the social and cultural environment in and around the school, etc.) is also included and its particular meaning and impact on the storyteller become evident in the emerging picture. Taking up the different roles contributes to an increased awareness of the social, interactive character of the shared sense-making that takes place. Furthermore, the interactions entail the possibility of breaking up taken-for-granted interpretations, providing alternative understandings of that past experience and – as a consequence – of one's actual situation.

The rationale in narrative pedagogies

The actual practices of using narratives and biographical accounts for teacher development (both in pre-service and in-service training) are quite diverse. Both research and experience from practice show that life stories most often contain events that appear to have had a very significant meaning for the narrator. Often these events are all but extraordinary or spectacular. Sometimes they may even look trivial to others (Kelchtermans, 2009). But for the storyteller these events, encounters, experiences have been turning points, which made them reflect and reconsider their taken-for-granted beliefs and ideas. These so-called *critical incidents* are "key events in an individual's life, and around which pivotal decisions revolve. They provoke the individual into selecting particular kinds of actions, which lead in particular directions. (...) Critical incidents are a useful area to study, because they reveal, like a flashbulb, the major choice and change times in people's lives" (Sikes et al., 1985, p. 57).

The vignette is one exercise that takes up the idea of critical incidents. The assignment to reflectively recall, narratively share and analyze a particularly meaningful event can take different forms. Yet it always involves three steps: (1) description of the event, (2) identification of the issue it raises, and (3) a statement of the sense the narrator made of it with respect to his/her process of learning to teach (Carter, 1993). It thus always aims at triggering a reflective move from the storied event to an analysis of its particular meaning and relevance to the narrator's present thinking and actions.

In a similar vein, another group of narrative techniques or exercises draws on *metaphorical thinking*. Participants are invited to come up with a metaphor to typify themselves as teachers or to define what, in their opinion, constitutes the core of good teaching. These personal teaching metaphors (Bullough and Gitlin, 2001) thus represent, in a condensed, narrative way, their key ideas about themselves as (student) teachers. They may also reveal how these ideas can change over time. As such it makes sense to engage (student) teachers in exercises with personal teaching metaphors or accounts of critical events at different moments during the process of learning to teach (e.g., as an element in a reflective portfolio which they compose during their teacher education or as part of a professional diary). Comparing the differences between the metaphor chosen in the present and the ones developed at earlier stages helps to explore and understand one's personal learning process as well as the experiences that triggered the changes over time.

As such, the use of narratives in teacher education *per se* is also in line with and contributes to the emphasis on *teachers' reflectivity*. Reflection, both as an attitude and skill refers to people's capacity to look back on and think about themselves, their actions, and the situation they find themselves in. Because of the dynamic, highly unpredictable, and complex character of teaching and learning processes, reflectivity is a crucial tool for teachers to properly understand professional situations, choose appropriate actions, and learn from these experiences. Through reflection, teachers thoughtfully examine new knowledge and practical experiences, and integrate them with deeply held beliefs and opinions (Kelchtermans, 2009; see also Lyons, 2010; Schön, 1983).

Narratives and different forms of storytelling have proven to be powerful tools to engage (student) teachers in reflection. Narrative methods allow their reflections to go beyond the technical or instrumental issues of properly using knowledge and skills to achieve effective teaching performances. The latter often dominate (student) teachers' reflections as well as the feedback from teacher educators or trainers on how to improve the effectiveness of one's future teaching performance. Although the importance of this technical dimension in teacher professionalism and development goes undisputed, there is more to good teaching than skillful performance and efficiently dealing with the "how to" questions. Teaching and being a teacher inevitably also involves deciding on moral dilemmas, living through emotional experiences, and dealing with political issues of power and influence. Coping with these moral, emotional, and political dimensions of teaching constitutes an equally important agenda for teacher education and development as do the issues of teaching skills and (content) knowledge (Hargreaves, 1995). Through narrative accounts of teaching experiences, (student) teachers can reflectively address their moral puzzlements (e.g., conflicting norms about good teaching), emotional concerns (e.g., the experience of powerlessness or self-doubt when pupils do not show the learning outcomes one had aimed for), or political actions (e.g., conflicts with colleagues or principals about the most appropriate content or

methodology in teaching). Sharing and discussing these dimensions help student teachers to accept these dimensions as inherent to the teaching profession and thus essential to their self-understanding as teachers. In other words, narrative assignments help to ensure that the content of the reflections is sufficiently *broad* (not only technical, but also moral, emotional and political) as well as *deep* (going beyond the level of actions, to also include one's self-understanding and subjective educational theory) (Kelchtermans, 2009).

In order to achieve this, the different narrative techniques in teacher education and development show a common underlying pattern, composed of two complementary elements. First, the techniques constitute a vehicle for helping the narrators express through different narrative formats their deeply held ideas about themselves as a teacher, their values, and beliefs about (good) teaching, in short their personal interpretative framework. Because of its biographical, experiential, and largely idiosyncratic nature, this framework and the sense-making it engages in may be one-sided, limited, or simply wrong. As such it may negatively bias (student) teachers' learning and sense-making of the formal training curriculum and especially of the experiences during practical training (internships). Several authors have argued that, the often observed 'washing out'-effect of teacher education programs once teachers have started their career, is due to the fact that the implicit personal, interpretative lenses of the student teachers were not deliberately made explicit or discussed. In order to control possible biases and ensure a sustainable, in-depth integration of the training curriculum by the (student) teachers, the pedagogies used have to aim at making this implicit knowledge explicit and bringing it up for critical scrutiny and public discussion. Only this way (student) teachers' personal professional knowledge, as the basis for their professional decisions and actions, can be validated and justified.

Second, this individual, personal storytelling, always needs to be complemented by particular interventions and reactions from the audience, that is, peers and/or teacher educators/trainers. Their comments and questions on the publicly shared narrative accounts aim at questioning, probing, challenging, and deepening the reflective analysis. The storytelling in itself is not enough. Many authors stress the need to link story-telling with forms of narrative inquiry, in which the stories and the beliefs they reveal are critically analyzed, questioned, re-storied. Interpersonal dialog, narrative accounts, and the critical use of theoretical knowledge thus have to go hand in hand (see e.g. Bullough, 2008; Clandinin, 2013; Lyons, 2010).

The striving for authenticity as well as the critical probing of the accounts constitute two essential and necessary principles in narrative pedagogy if it wants to avoid the pitfalls of simply confirming existing ideas and beliefs. Critical and public examination may result either in a more informed conviction and awareness of one's ideas, or in a thoughtful reconsideration and rethinking of what was until then taken for granted.

Although an environment that is supportive, trusting, and positive is a necessary condition to allow the sharing of and learning from narratives to take place, it remains a continuous task for the users of narratives to avoid the trap of uncritical self-affirmation. The way to achieve this is to create and maintain a productive level of discomfort (Kelchtermans, 2009) that helps the critical reflection to move beyond the story as such.

A final reason for the importance of narrative and biography in the pedagogy of teacher education and development is the recognition that the person of the teacher is a constitutive part of the educational situation. Therefore, teachers' ideas of themselves as teachers, their sense of identity is of much greater importance to them as practitioners than in occupations where the person can easily be separated from the craft. When speaking about their job, teachers demonstrate "persistent self-referentialism" (Nias, 1989): their self as a person is always implicated. Developing their professional self-understanding constitutes a red thread for teachers' learning and development throughout their teacher education as well as their work life. (Auto)biographical reflection and narrative accounts allow for this professional self-understanding to become conscious, explicit as well as critically studied and reflected on.

References

- Ball, S., Goodson, I. (Eds.), 1985. *Teachers' Lives and Careers*. Falmer Press, London, UK.
- Brooks, P., 2003. *Reading for the Plot: Design and Intention in Narrative*. Harvard University press, Cambridge, Mass.
- Bullough, R., Gitlin, A., 2001. *Becoming a Student of Teaching. Linking Knowledge Production and Practice*, second ed. Routledge, New York, NY.
- Bullough, R., 2008. The writing of teachers' lives: where personal troubles and social issues meet. *Teach. Educ. Q.* 35, 7–26.
- Carter, K., Doyle, W., 1996. Personal narrative and life history in learning to teach. In: Sikula, J., Buttery, T.J., Guyton, E. (Eds.), *Handbook of Research on Teacher Education*, second ed. Macmillan, New York, pp. 120–142.
- Carter, K., 1993. The place of story in the study of teaching and teacher education. *Educ. Res.* 22 (1), 5–18.
- Casey, K., 1995–1996. The new narrative research in education. *Rev. Res. Educ.* 21, 11–253.
- Clandinin, J., 2007. *Handbook of Narrative Inquiry. Mapping a Methodology*. Sage, Thousand Oaks.
- Clandinin, J., 2013. *Engaging in Narrative Inquiry*. Left Coast Press, Walnut Creek, CA.
- Clark, C., Peterson, P., 1986. Teachers' thought processes. In: Wittrock, M. (Ed.), *Handbook of Research on Teaching*, third ed. Macmillan, New York, pp. 255–296.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14.
- Craig, J.C., Meijer, P.C., Broeckmans, J. (Eds.), 2013. *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community*. Emerald Group, Bingley.
- Craig, C.J., Orland-Barak, L., 2014. *International Teacher Education: Promising Pedagogies (Part A)*. Emerald Group, Bingley.
- Day, C., Sachs, J., 2004. *International Handbook on the Continuing Professional Development of Teachers*. Open University Press, Maidenhead.
- Fenstermacher, G., 1990. Some moral considerations on teaching as a profession. In: Goodlad, J., Soder, R., Sirotnik, K. (Eds.), *The Moral Dimensions of Teaching*. Jossey-Bass, San Francisco, CA, pp. 130–151.
- Goodson, I., 1984. The use of life histories in the study of teaching. In: Hammersley, M. (Ed.), *The Ethnography of Schooling*. Nafferton Books, Driffield, UK, pp. 129–154.
- Goodson, I. (Ed.), 1992. *Studying Teachers' Lives*. Routledge, London.
- Goodson, I., 2013. *Developing Narrative Theory. Life Histories and Personal Representation*. Routledge, London-New York.

- Hargreaves, A., 1995. Development and desire: a postmodern perspective. In: Guskey, T.R., Huberman, M. (Eds.), *Professional Development in Education: New Paradigms and Perspectives*. Teachers College Press, New York, NY, pp. 9–34.
- Hoepfel, R., 1983. Perspektiven der erziehungswissenschaftlichen Erschließung autobiographischer Formen der Selbstreflexion (Perspectives for the Use of Autobiographical Self-reflection in Educational Research). *Z. für Pädagogik* 18 (Beiheft), 307–312.
- Huberman, M., Grounauer, M., Marti, J., 1993. *The Lives of Teachers*. Cassell, London, UK.
- Kelchtermans, G., 1993. Getting the story, understanding the lives. From career stories to teachers' professional development. *Teach. Teach. Educ.* 9, 443–456.
- Kelchtermans, G., 1994. Biographical methods in the study of teachers' professional development. In: Carlgren, I., Handal, G., Vaage, S. (Eds.), *Teachers' Minds and Actions. Research on Teachers' Thinking and Practice*. Falmer, London, pp. 93–108.
- Kelchtermans, G., 2009. Who I am in how I teach is the message. Self-understanding, vulnerability and reflection. *Teach. Teach. Theor. Pract.* 15, 257–272.
- Kelchtermans, G., 2017. Studying teachers' lives as an educational issue. Autobiographical reflections from a scholarly journey. *Teach. Educ. Q.* 44, 7–26.
- Kelchtermans, G., 2018. Professional self-understanding in practice: narrating, navigating and negotiating. In: Schutz, P.A., Hong, J., Cross Francis, D. (Eds.), *Research on Teacher Identity. Mapping Challenges and Innovations*. Springer, Cham (Switzerland), pp. 229–240.
- Kompf, M., O'Connell Rust, F., 2013. The international study association on teachers and teaching (ISATT): seeing tracks and making more. In: Craig, C.J., Meijer, P.C., Broeckmans, J. (Eds.), *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community*. Emerald Group, Bingley, pp. 3–38.
- Krüger, H.-H., Marotzki, W. (Eds.), 1996. *Erziehungswissenschaftliche Biographieforschung*. Leske-Budrich, Opladen.
- Labov, W., Waletzky, J., 1973. Erzählanalyse. Mündliche versionen persönlicher Erfahrung. In: Ihwe, J. (Ed.), *Literaturwissenschaft und Linguistik*, vol. 2. Athenäum, Frankfurt am Main, pp. 78–126.
- Lyons, N. (Ed.), 2010. *Handbook of Reflection and Reflective Inquiry. Mapping a Way of Knowing for Professional Reflective Inquiry*. Springer, Dordrecht.
- Nias, J., 1989. *Primary Teachers Talking. A Study of Teaching as Work*. Routledge, London.
- Noddings, N., 1984. *Caring. A Feminine Approach to Ethics and Moral Education*. California University Press, Berkeley, CA.
- Polkinghorne, D., 1988. *Narrative Knowing and the Human Sciences*. State University of New York Press, Albany, NY.
- Schön, D., 1983. *The Reflective Practitioner. How Professionals Think in Action*. Temple Smith, London.
- Sikes, P., Measor, L., Woods, P., 1985. *Teacher Careers. Crises and Continuities*. Falmer, London.

Further reading

- Bruner, J., 1986. *Actual Minds, Possible Worlds*. Plenum, New York.
- Bullough, R.V., 1991. Exploring personal teaching metaphors in preservice teacher education. *J. Teach. Educ.* 42, 43–51.
- Day, C., Gu, Q., 2010. *The New Lives of Teachers*. Routledge, London-New York.
- Doyle, W., Carter, K., 2003. Narrative and learning to teach: implications for teacher-education curriculum. *J. Curric. Stud.* 35, 129–137.
- Freeman, M., 2017. *Modes of Thinking for Qualitative Data Analysis*. Routledge, New York-London.
- Gudmundsdottir, S., 1991. Story-maker, story-teller. Narrative structures in curriculum. *J. Curric. Stud.* 23, 207–218.
- Hanne, M., Kaal, A. (Eds.), 2019. *Narrative and Metaphor in Education. Look Both Ways*. Routledge, London-New York.
- Kelchtermans, G., 2009. Career stories as gateway to understanding teacher development. In: Bayer, M., Brinkjaer, U., Plauborg, H., Rolls, S. (Eds.), *Teachers' Career Trajectories and Work Lives*. Springer, Dordrecht, pp. 29–47.
- Lakoff, G., Johnson, M., 1980. *Metaphors We Live by*. University of Chicago Press, Chicago-London.
- Loughran, J., Hamilton, M.L., LaBoskey, K.V., Russell, T. (Eds.), 2004. *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer Academic Publishers, Dordrecht.
- Personal Narratives Group, 1989. *Interpreting Women's Lives: Feminist Theory and Personal Narratives*. Indiana University Press, Bloomington (IN).
- Sparkes, A., 1994. Life histories and the issue of voice: reflections on an emerging relationship. *Int. J. Qual. Stud. Educ.* 7, 165–183.
- Webster, L., Mertova, P., 2007. *Using Narrative Inquiry as a Research Method. An Introduction to Using Critical Event Narrative Analysis in Research on Learning and Teaching*. Routledge, London-New York.

Learning to count what counts: narrative hermeneutics for teachers and teaching

Dennis Sumara, Sharon Friesen, Catherine Takasugi, Janice Doucette, and Jennifer Turner, Werklund School of Education, University of Calgary, Calgary, AB, Canada

© 2023 Elsevier Ltd. All rights reserved.

Narrative hermeneutics	113
Other peoples' children	114
Reading an archive of feelings	114
Enabling constraints	115
Learning to count what counts	115
Possible under the circumstances	116
The new precariate	117
Narrative hermeneutics for teachers and teaching	119
References	119

Knowledge [*savoir*] cannot generally be reduced to science, nor even to learning [*connaissance*].

Lyotard (1984, p. 18).

As we considered what we might say about the future of research on teachers' work and lives we realized the topic is rife with ongoing unresolved status and power differentials between university professors and school teachers. In addition, there is an underlying assumption that work and lives can be considered apart from one another. In our research and teaching we make no distinction between the two, understanding that although there are differences between our public and personal lives, they are always intertwined and exist as co-emergent phenomena.

In this article we describe and use a narrative hermeneutic approach to emphasize the importance of prioritizing the local, situated knowledges of teachers when researching their work and lives. Through first person narratives of how teachers and leaders in the city of Fort McMurray in Canada and Saitama City in Japan are using forms of ethical know-how to guide their work and lives in the midst of Covid-19, we show how the co-emergence of nation, city, neighborhood, land, language and culture influence the actions of educators within the context of a global disruption. This context suggests productive approaches to working with teachers to better understand the intertwined situation between their professional and personal histories as these are nested within past and current narratives that shape the social, cultural and political contexts of their work and lives.

Narrative hermeneutics

Emerging from the philosophical writings of Heidegger (1966) and Gadamer (1990) and hermeneutics in education (Carson and Sumara, 1997; Gallagher, 1992; and Smith, 1991) hermeneutic inquiry posits that all experience emerges from and requires ongoing interpretations and that these interpretations are primarily structured through narrative forms and conditions. As Meretoja (2018) explains:

Conceptualizing narratives as practices of interpretation, enmeshed with other socially and historically situated dialogical practices, allows us to see how it is an inherent feature of every narrative that it can be told in different ways, and how our condition as storytelling animals is one of always being in the middle of a dialogue and struggle of interpretations (p. 85).

Narrative hermeneutic processes not only are focused on guiding processes of how stories are told and which stories are told, they also require attention to the ways these stories are influenced by history and culture, including the lived situations of those conducting the interpretations. It is not enough to interpret what was said; it also is necessary to interpret the historical influences leading to that story, the position of the narrator, the power dynamics at play, the ways in which prejudgments influence the narrated story and the storyteller, and the cultural and social contexts of the story and its telling (Meretoja, 2018).

As Smith (1991) has argued, there is a requirement on the part of the researcher using a hermeneutic approach to offer an account of how the process of engaging in research influences the researcher. There can be no separation of researcher from research. As new insights and knowledge emerge, the stories guiding the identities of those involved in the research are changed. Hermeneutic interpretations are not merely reports on knowledge; they emerge from processes that begin with the phenomenological question: "What is this experience?" Along with the hermeneutic question "Why is it like that?" the hermeneutic researcher also must consider the historical, social, cultural, political and ideological influences at play that conditions the phenomenological qualities of experience.

For the purposes of this article, we begin with first-person narratives written by three front-line educators and leaders—Janice, Jennifer and Catherine—each of whom have been thinking and writing about how Covid-19 has influenced their work and lives. We then follow with hermeneutic interpretations of what their narratives reveal when considered within the nested historical, social, cultural and political contexts of their narrated experiences.

Other peoples' children

I (Janice) remember when the Horse River Fire crossed the Athabasca River and made its way into our subdivisions on May 3rd, 2016. We teaching and support staff were at our junior high school that afternoon and, beginning shortly after the lunch period, we became responsible for keeping students calm while trying to reunite them with their families. We also had to worry about connecting with our own children who were scattered about the city at other schools and daycare homes. Staff members were legally obligated to stay in the building and take care of other people's children before our own. It was an intensely emotional time with numerous unknowns.

Our School Principal (Scott), led the reunification efforts that afternoon. At official dismissal time, many staff left the building, anxious to tend to their own children and escape before the flames blocked the only highway out of town. Those who did not have immediate obligations remained onsite to help with the dozen or so kids who had not yet been united with their families. I stayed until about 4:30 that day, at which time there were about half a dozen students remaining in the school.

I learned weeks later that Scott stayed at the school that night with five students who had not yet been claimed because their parents were out of town on business. He was willing to stay with these students that night at the school, but the RCMP came to the school at 10:30 that evening and ordered them to leave. Scott left with those students in his car and made the nearly 5 h drive to the city of Edmonton to reunite them with their parents. I learned later that principals of other schools left students unattended to fend for themselves.

Four years later, we faced a new threat: the Covid-19 pandemic. Many of the same feelings that accompany navigating the unknown surfaced again; disbelief, the burden of responsibility for other people's children, and concern for our own safety became paramount once more. One major difference is that our enrollment is 30% higher than it was four years ago; we now have close to 1000 students. How could we keep Covid-19 out of our building? Amazingly, we are over three months into the school year with no major changes in program delivery and we have not had one known case to date.

From the onset, Scott emphasized that he did not have all the answers, and he invited our input to help plan our re-entry plan for September 2020. I believe that our collaborative approach is part of our current success. I also know that the culture of caring and reciprocity that our students have experienced has primed them to respect the re-entry rules. Ours is a culture of mutual respect and Scott is a leader who practices service-oriented leadership. Most certainly Covid-19 cases will emerge, but we will be able to handle the outbreak. We've stood on the threshold of trauma and uncertainty before, and our staff and students made it through them. We are stronger and more unified for having done so.

Reading an archive of feelings

I (Jennifer) arrived in Fort McMurray to begin my new role as Superintendent nine days before the forced cancellation of all K-12 classes due to the Covid-19. Intellectually I understood that my new role situated me in a community that experienced the greatest natural disaster in recent Canadian history; the 2016 Horse River fire that destroyed hundreds of homes and businesses. Still, I had not yet understood what lasting impact would remain for the people and community or what may be needed from me as their new leader.

With intention, I set out to nurture personal connection and engage in reflective dialog, to learn about the people and community that I would join, what their needs may be, and how I might positively contribute to this system. I wanted to hear their stories. I needed to hear them. Principals, teachers, senior staff members, and community members generously shared their account of that day, describing the lasting impact on their daily lives and professional practice. The initial understanding of the local people and context gained through trust-building and relationship-building has proven to be invaluable as our school system responded to the increasing demands to isolate due to Covid-19. Forces of nature and economics would layer additional pressures and demands on community members as we responded to a 100-year flood during my second month, all the while in a community impacted by a long and deep recession as a result of losses in mining, oil, and gas industries.

In very amplified ways I was living out what I had read about how challenging circumstances require leaders to interpret and make sense of those challenges and be able to communicate those interpretations within the context of daily decisions. My own experience had taught me that small practices between individuals and small groups that demonstrate a strong ethic of care and a narrative of high expectations are necessary within a challenging and complex environment. During my conversations, initially in-person and then online, with principals, coordinators, members of senior leadership and community partners, I listened to not only the words but also took notice of the emotional responses as individuals shared their narrations. Not everything that they needed to communicate could be communicated verbally and I needed to be aware that those silences were forms of knowledge, all of which contributed to an archive that I needed to understand. Those many small actions have made a big difference for all of us as we continue to find ways to ensure health safety along with our other more typical commitments.

Enabling constraints

In Saitama City, Japan after 3 months of voluntary self-restraint all levels of public school resumed. My (Catherine) and my children's return to teaching and school process was gradual and cautious with rules layered upon an already rule heavy system. In a country where a dense population lives, moves, and disposes of garbage on a land mass that fits into the area of California, rules are deemed necessary and children, neighbors, and elders inherently learn to gently pressure deviators into conformity. As a natural freedom-seeking diversity-loving creative soul, this obedience and linearity has often left me exasperated.

At our local elementary school there is colored tape running down the center of the hallways delineating on which side a student should walk, running is of course forbidden. Schools decree size, color, shape of all manner of appropriate school items from lunch aprons, indoor shoes, and gym uniforms, characterless pencils and erasers; even pencil case style is pre-determined. Rules are what keeps the flow from day to day. It also makes school life predictable and safe. Questioning the why, how, or for whom is not built into the curriculum nor the psyche (nor is it typically encouraged), and to be honest, serves a great purpose when a nation is fighting a pandemic (or facing earthquakes, typhoons, or tsunami). When, the call was made across the nation to close schools, there was surprise and apprehension, but very little resistance.

Going back to school came with a new set of rules that I confidently knew would be strictly followed: open windows in the classroom, masks on at all times, water bottles brought from home rather than water fountain drinking, paired desks now separated by a measured margin, no touching or socializing with friends, sitting forward in desks, playground equipment off limits, and no use of instruments, no singing or loud voices. Every morning, students also had to report their temperature in written form to the classroom teacher. Now, in November 2020, with slightly more known about Covid-19 transmission, most of these rules are still in place but not quite as steadfastly as before. Masks continue to be ubiquitous, physical contact is minimized, windows are still open.

Japan's mask culture emerged decades before the Covid-19. Hygiene, flu season, bad breath, anonymity, acne; these are all perfectly acceptable reasons to wear a mask in Japan. Hay fever season alone forces a large percentage of the population to don a mask for several weeks or months every year. As early as January 2020, many families had boxes of paper masks prepared at home and drug store shelves were stocked with these necessary seasonal items well before they were suggested as a possible pandemic counter measure. When the order came, the population at large had no issue with full-time mask wearing in public. We moved directly to the second phase of the discussion: breathability, viral penetrability, washability, ideal fit and comfort, environmental concerns, and of course style.

Japan has a culture that rewards discipline, stamina, and self-sacrifice for a greater good. Awards are given to children who have continuous attendance at school for their entire 6 years of elementary school. That means no sick days are officially taken, or more accurately, plenty of sick days are experienced, but children attend school anyway. This attitude and behavior is rewarded as these children are publicly honored for their robust nature. Having not grown up in Japan and having witnessed my own sickly husband travel an hour on public transport, work a typical 12 h day, come home in worse shape, all in the name of loyalty and self-sacrifice, I have always questioned this germ dispersing practice. Covid-19 is re-shaping this attitude. Now, if a student or worker is unwell in any capacity, they are expected to remain at home.

Learning to count what counts

In the late 1970s the government of the province of Quebec in Canada commissioned the philosopher Jean-Francois Lyotard to write a report on knowledge, which eventually was published as *The Postmodern Condition* (Lyotard, 1984). In that report, Lyotard discusses the status of science, technology and the arts within the context of the emergent post-industrial age and the evolution of what counts as knowing, knowledge, information and legitimation. Pointing to what he describes as an "incredulity to metanarratives" he predicts that:

Knowledge in the form of an informational commodity indispensable to productive power is already, and will continue to be, a major—perhaps the major—stake in the worldwide competition for power (p. 5).

Lyotard argues that these metanarratives must be overtaken by what in the French language are called *les petit recits*—the small stories of everyday life that revolve around the ways we negotiate friendships, intimacies, desires, longings, worries, hopes. He uses the French words *savoir* (know-how) and *connaissance* (personal knowledge) to depict how meanings attributed to similar words can be very different across different languages. Know-how (*savoir*) and personal knowledge (*connaissance*) are not so much associated with measurable facts or normalized conceptions of truth in the form of meta discourses; they are best represented through *les petit recits*—the small stories that orient the micro identities we experience in our moment to moment micro worlds. They are often quiet, subtle stories, rich with detail. As Grumet (1988) has argued, these are the stories about getting the groceries, tending the gardens, sharing a secret, feeling hurt when one's birthday is not remembered. They are personal stories, relational ways of being, long associated with women and children and, as a consequence, not considered to be of much value.

A half-century later, we notice the dramatic amplification of how this now operates within and across so-called formal and informal communication networks. Indeed, advances in digital technologies have not only changed social interactivity they have

changed human consciousness (Edelman, 2004; Hayles, 2005). The radical freedoms of the digital age have become what Foucault (1988) described as the panoptic gaze which he argued was amplified in prisons, hospitals, and schools, but which now has become ubiquitous in all aspects of our digitized cultures and identities. While it is the case that there is more sharing of personal narratives in many ways than ever before, these have not fundamentally displaced the metanarratives and ensuing power plays that continue to orient governments or the policies they create (Harrari, 2018).

In her public annual public Canadian Massey Lectures later published as *The Real World of Technology* (1990) Ursula Franklin argues that technology as a practice permeates and integrates with all aspects of our lives, mostly in ways that become normalized and, as a consequence, not noticed. As new technologies emerge, the qualities of our personal and public experiences change. How do we experience our lives, our work, our identities in this not well-established and rapidly evolving post-modern, post-industrial, post-human world? Franklin argues that the concepts, principles, practices and values associated with the Cartesian oriented Industrial age resulted in the forms of instrumental, prescriptive practices that are amplified in most forms of formal public education. But of course none of this is new or news. However compelling they might be, these philosophical discussions, do not necessarily disclose much about the experience of what Berlant (2011) describes as amplifying precarity in everyday life:

Precarity is an economic and political condition suffered by a population or by the subjects of capitalism generally; or a way of life; or an affective atmosphere; or an existential truth about the contingencies of living, namely, that there are no guarantees that the life one intends can or will be built (p. 192)

Or, we might add, a global pandemic. The Covid-19 pandemic has shown that precarity plays no favorites; there are no walls or categories or systems that can shield individuals, families, communities or nations from its reach. Indeed, Covid-19 has highlighted that there exists an inter-class we could name the precariate—one characterized by heightened insecurity, worry, anxiety and fear. Ironically, what most has united us has emerged at the same time as the imperative to be socially distant from one another. In both inspiring and dispiriting ways, Covid-19 has revealed the fault lines of what we have prioritized and valued and what we have not.

In Canada we witnessed our failure to provide everyday necessities to those in elder care homes, to First Nations communities, to homeless populations, to early years day care facilities, to the poor. At the same time, we have witnessed ordinary citizens—notably front-line health care workers and teachers—respond to the immense challenges of ensuring health safety in communities and schools. For the first time in any of the authors' memories the nightly news features nurses, doctors, teachers. We hear not only their advice, we also hear their stories—ones that give us hope that their local knowledge of patients, students, schools, hospitals—their local contexts—might indeed be what must be prioritized. However, as we have witnessed, the power and impact of the situated knowledges of teachers can only be considered within the histories and metanarratives that guide what counts as normal, correct, valued and needed.

Possible under the circumstances

In a 1960s national radio contest, Peter Gzowski of the Canadian Broadcasting Corporation challenged the nation: "Complete the adage, As Canadian as" As it turned out listeners heard the contest as a quest for something quintessentially Canadian—a symbol to fit our nation. Most submissions were predictable (e.g., a fresh snowfall, the Royal Canadian Mounted Police). In the end, the winner was not an image or a metaphor but, instead an ironic (some would say satiric) commentary on the Canadian character: "As Canadian as possible ... under the circumstances."

Owing to Canada's historical emergence as a nation that embraced cultural and social accommodation over assimilation, a cultural mindset of ongoing adaptation became the norm (Sumara et al., 2001). However, this may not be a uniquely Canadian quality. One could argue that all nations continue to inhabit and adapt national identity in ways that are *possible under the circumstances*. What we notice, however, is that there are profound differences in the ways in which narratives that guide national identity influence responses to emerging circumstances.

With the industrialization of education came an assumption that replicable across contexts best practices could be applied and lead to desired learning outcomes (Clifford and Friesen, 1993; Davis et al., 2000). For most of the latter part of the 20th century and the first part of the 21st century we have seen movement to continue to standardize not only curriculum and learning outcomes, but also approaches to teaching—and more than we might care to admit—teachers (Friesen and Jacobsen, 2020). However, what might work in one national or regional context may not work or be appropriate in another. Catherine's narrative, for example, demonstrates clearly how and why Japan has been successful ensuring high degrees of health and psychological safety in the midst of Covid-19 threats in schools and elsewhere. Most notable is that there was never a required or enforced national shut down or quarantine but, instead, a request that all citizens comply with whatever was necessary to ensure safety for all citizens. This is paired with a national education system, a national curriculum and national education policies.

Canada has not been as successful. Owing to our decentralized confederation of provincial jurisdictions approach to democracy, each of the ten provinces and three territories hold authority for health care and education. Needless to say there have been dramatic differences to the approaches taken—with some provinces requiring many more restrictions in business, community and school settings. In the province of Alberta where four of us reside, the approach has been more aligned with an American-inspired rugged individualism. To date, there has been no provincial mandate to wear masks and only very limited restrictions on businesses and

community engagements. However, even across Alberta there are differences in the approaches used to ensure the health safety of citizens and students in schools. In the city of Fort McMurray, there are differences to be noted—ones which we argue are intricately linked to the city of Fort McMurray's history and, as well, the specific historical teaching and leadership approaches and practices used in different schools.

The new precariate

In 2020 the Covid pandemic showed clearly the precarity and fragility of human life. Indeed, it amplified what [Berlant \(2011\)](#) has argued is the new interclass of the precariate—since Covid knows no class or geographic or political distinctions. We are connected to a planetary ecology in ways that cannot be avoided and, indeed, are necessary ([Bateson, 1972](#)). We also learned a new way of thinking about safety in schools—not only safety from threats of exclusion and bullying but also immediate threats to health. Ironically, it has been a biological certainty about the outcomes of non-compliance to safety that has shone a light on the importance of front-line teachers, teacher assistants and school/district leaders. These complex new challenges and realities in schools all over the world have highlighted three intertwined factors that have implications for how we are thinking about the future of research on the work and lives of teachers:

First, the historical intertwined co-emergence of a nation/region of that nation and curriculum/education policy and practice has a profound influence on how structural and/or procedural changes occur—including those changes that become suddenly urgent and necessary ([Davis et al., 2012](#)). The prevailing discourses on learning and teaching are strongly influenced by those historical developments, including the ways in which teachers are regarded by society and the ways in which their personal and professional ways of knowing are either counted or not counted at the level of local and national decisions on approaches to education. For example, there are marked differences between the status of teachers in Japan and in Canada. Teachers have very high status in Japan and are considered to be reliable authorities by parents and members of the public. In part, this is linked to Japan's view of the school as a moral community, with teachers as leaders not only of academic development and achievement, but also of ensuring adherence to Japanese cultural and social values. While this is also true to some extent in Canada, the authority of the teacher is much more limited. Indeed, the ongoing contestations about where the line is drawn between and among parental, religious and cultural values and those included in programs of study become areas of tension and contestation for teachers ([Cereluk, 2008](#); [Sumara, 1996](#)).

Catherine's narrative about her experience as European descended Canadian born woman who for more than 20 years has taught and parented in Japan shows how the specific histories of what it is like to be considered a good citizen in Japan interweave with what it is like to be a good teacher. Although she admits to having felt somewhat constrained by the many rules in Japanese society and schools, she also explains how they became important enablers as Japan learned to adapt to daily life that included Covid-19. In Japan, the duty to nation both in and out of school conditioned a much safer and less volatile experience of teaching and learning. At this writing, Japan continues to maintain low Covid-19 transmission rates, with all schools fully open and operational.

With over 20 years experience as a classroom teacher, Janice describes how her high school of over 1000 students in the city of Fort McMurray in the northern part of the Province of Alberta has remained virtually free of any cases of Covid-19 since resuming classes in September. This is an incredible outcome given that the Government of Alberta has not required mandatory mask wearing. Despite dramatically escalating cases of Covid-19 in the last few weeks of 2020 the province has not required any kind of mandatory lock down, even in major cities where the cases are escalating at alarming rates. Ironically, they *have* closed secondary schools requiring online learning for the remainder of this term, but have *not* closed major retail ("big box") outlets. The small stories leading to deep culture of ethical know-how and trust that created conditions for Janice's high school to remain Covid-free will *not* be counted as an extraordinary success story in the Canadian province of Alberta where personal freedoms to choose have taken priority over health and psychological safety.

Over the past several decades the city of Fort McMurray has been in the news many times, owing to the fact that it is the primary urban center for the surrounding oil and natural gas industries. Originally inhabited by the Cree First Nations peoples who used the surface oil deposits to waterproof their canoes, the area that became the city of Fort McMurray was central to the fur trade in the 1800s. Commencing in the 1920s with dramatic expansion from the 1970s to the 1990s Fort McMurray became the service city for the lucrative oil and gas energy industries. However, over the past 20 years it has become severely challenged by massive wild-fires, floods and the collapse of oil and gas prices and demand. Despite this, the citizens of Fort McMurray, many of whom came there from other parts of Canada and all over the world, have demonstrated what Canadians would identify as an optimism handed down from histories of learning how to thrive in new challenging situations—as *Canadian as possible under the circumstances*.

In her narrative, Janice describes the kind of trust and collective commitment that has emerged over the years owing to the leadership style of their school Principal. She attributes the success they have had with maintaining not only health but also psychological safety in their high school as connected to a service-leadership orientation by their Principal that is not limited to creating conditions for academic success. That their high school has remained Covid-19 free while others in the city and in the Province have had major outbreaks can only be explained as the outcome of what [Varela \(1999\)](#) describes as ethical know-how in action:

Truly expert people act from extended inclinations, not from precepts and thus transcend the limitations inherent in a repertoire of purely habitual responses (pp. 30–31).

For Varela, ethics is not something known in advance and then applied uniformly when needed; instead, ethics is knowing the right thing to do in a particular situation. Furthermore, ethical know-how must always consider the micro identities and micro worlds that add up to our individual and collective experiences. The practice of ethical know-how from Scott the school Principal had a profound impact on the lives of teachers and students. It conditioned and sustained their ability to remain safe during a time of disruption in a city that has experienced unnecessarily heightened precarity owing to weak provincial government leadership.

Jennifer's narrative tells us what it was like for her as the new Superintendent of Schools for Fort McMurray to enter a school and urban system that was about to experience another highly precarious situation after a decade of working through the devastations of catastrophic floods, fires and economic collapse. Not revealed in her narrative is her many years experience as a certified school psychologist. With in-depth experience working with students and families presenting unique and challenging social, emotional and intellectual needs, Jennifer brings to her system leadership a well-developed capacity to listen closely for signs of the effects of trauma—whether these be the result of one major event or what Cvetkovich (2003) describes as a more insidious type emerging from daily experiences of emotional and psychological neglect and/or abuse. These forms of trauma emerging from systemic exclusions and bullying are difficult to remedy owing to the ways in which individuals and groups are silenced within an organizations or institution (Probyn, 2005; Sumara, 2020).

The prevailing discourses on education in any nation will be closely linked to the primary discourses about what counts as success and the way these influence perspectives on the future. In Japan the meta-narratives about what counts as successful can be attributed to deep cultural beliefs about the importance of the nation as a whole and the duty of all citizens to work toward ensuring continued stability and prosperity. In part, Japan's success with Covid-19 containment emerges from the strong way Japanese citizens without force by the government of Japan have complied with all of the health safety recommendation. In a nation that to outsiders might seem to be overly constrained with many rules, the success with maintaining full business and education operations throughout most of the pandemic has demonstrated how compliance to a collective ethic of care can be enabling. What to some Canadians might feel like lack of individual free choice has created conditions for a less precarious existence in Japan.

Second, the growing resistance to discourses and practices of evidence-based practice in schools where what counts as data are those things normally considered measurable in the absence of more embodied, intuitive forms of knowing that guide daily decision making becomes amplified during times of increased precarity. In the province of Alberta, for example, standardized achievement exams were canceled in June, with teachers being asked to determine final grades based on classroom work, not unlike the approaches to education experienced in Alberta in the 1930s and 40s when Deweyan inspired progressive education trusted teachers to teach in ways that were more student and local community centered. Over the past few months there has been virtually no discussion in the media of trying to measure or count learning or achievement. Instead, the more relational, context specific local matters related to learning and teaching have been in the foreground. Owing in part to the blurring of boundaries between home and school during periods of quarantine and school closure more holistic and integrated forms of education were seen as not only preferred but necessary. The work and lives of teachers became much more obviously entwined with the work and lives of parents, caregivers and students. Entire communities stretched to understand what counted as education when that education could not be in a school classroom.

Perspectives on and practices of education in schools have not been subject to the same scrutiny in Japan owing in part to the link between Japanese cultural mindsets about the purposes of education, the high status and authority of teachers, and consistent ways citizens of Japan avoided Covid-19 volatility through compliance to health safety requirements both in school, at home, and in the community. As Catherine's narrative shows, the clarity of the story about what counts as a good citizen is linked to a national curriculum and associated pedagogies that are not usually questioned or contested. Although the programs of studies for all school subjects are the same for the entire country, there also is a requirement that teachers adapt the curricular requirements to the specifically local context and, importantly, that these local accommodations be communicated to parents and the local community. This historical culture of valuing local knowledge and context through explicit links between parent and teacher, school and home, and school and community remained relatively stable and unchanged with the introduction of Covid-19 health risks. Being as *Japanese as possible under the circumstances* has led to relatively smooth adaptive transitions during the last year of Covid-19 disruption.

Third, the emerging skepticism about the liberal democratic story and globalization is losing its luster. As Harrari (2018) has argued, the liberal story has been until recently read as a better alternative to the fascist or the communist story. However, as another metanarrative mostly developed by and for those who already have power and wealth, the liberal story has failed those on the other side of that power. Covid-19 has shown the limitations of all three of these grand narratives—and, in particular, the inabilities of the liberal democratic story at a basic level to respond to the kind of precarity associated with a global health crisis. Indeed, what has become clear is that while the crisis is global, the response and solutions must originate at the level of the local as situated within the regional and national.

Jennifer, Janice and Catherine's narratives about what it has been like for them to teach and lead in the midst of a global health disruption demonstrate the complex ways personal and professional know-how of teachers and leaders are situated in and conditioned by global, national, regional and local histories. They also show how what appears to be the same outcomes (Covid-free schools in Saitama City and a Covid-free high school in Fort McMurray) emerge from different circumstances and for different reasons. We are not suggesting that one or the other situation is better or worse than the other. We are suggesting that when studied hermeneutically through the situated local knowledges depicted in narratives written by front line teachers and administrators our understanding of *why* becomes as important as our understanding of *what* or *how*.

Narrative hermeneutics for teachers and teaching

In this article we have tried to show how the study of the small stories of teachers and teaching that are mostly undisclosed in the context of pandemic can lead to insights that normally are not noticed and/or counted because they are not easily noticed and/or counted. It is important to note that not even *once* in the news have we heard about the success of one high school in Fort McMurray that has maintained health safety in the midst of unreliable and mostly politicized government messaging on Covid-19. Completely hidden—not written anywhere, not counted—is Principal Scott's long standing service-oriented, ethical know-how approach to school leadership. Until we were well into writing this article we did not fully appreciate and understand why Saitama City (and indeed all of Japan) has avoided high infection and death rates related to Covid-19. What from the outside might seem like an overly rule-bound constrained system upon closer scrutiny reveals itself to be more enabling than disabling. Then, of course, there is the question of whether or not we care about or count the importance of intentional, highly relational leadership practices. What does it mean for us to notice and *count* how as Superintendent of Schools, Janice's scholarly and practical knowledge as a school psychologist enabled her not only to be aware at a visceral level of the effects of long-term insidious trauma in her system, but also to know how to create novel pedagogies in the midst of ongoing precarity? How does her ability to notice and read signs of distress influence how she makes daily decisions in the midst of a major systemic disruption? Or at anytime?

Caputo (1987) argues that hermeneutics is for the hardy; it is radical thinking that aims to perceive and understand the tangled interplay among histories, cultures, contexts, identities and lived experiences. Perhaps most radical about hermeneutic inquiry is its reliance on the small stories; the details of everyday life of teachers and education leaders as they move between and among their microworlds of teaching and not teaching, home and school, public and private, know-what and know-how and know-why.

Hermeneutics is not a research methodology per se, nor does it have much to say theoretically about how to teach or how to learn. It is a philosophical orientation that assumes the inseparability of knowing, doing and being that can be applied to different forms of research. When applied to the study of narrative depictions of experience hermeneutic inquiry can create interpretive conditions for the interplay among the historical, the cultural, the political, the geographical, the social (to name a few) to be noticed as these occur in a momentary micro world. It can help us to understand why, for example, the successes with Covid-19 health safety in one high school in Fort McMurray and all high schools in Saitama City cannot really be compared to one another. What *can* be compared are the radical differences that have led to the same outcome—which, as argued by Derrida (1978) means not merely comparing one to another in order to record their dissimilarities. A radical approach to hermeneutic inquiry also requires attention to what has been deferred during processes of regulating right from wrong, normal from abnormal, important from not important, counted from not counted. We argue that the future of researching teachers work and lives must take seriously the idea that the simply told story by a teacher of what happened must count, must be counted, must matter.

References

- Bateson, G., 1972. *Steps to an Ecology of Mind*. Ballantine Books, New York.
- Berlant, L., 2011. *Cruel Optimism*. Duke University Press, Durham, London.
- Caputo, J., 1987. *Radical Hermeneutics*. Indiana University Press, Bloomington, IN.
- Carson, T., Sumara, D., 1997. *Action Research as a Living Practice*. Peter Lang, New York.
- Clifford, P., Friesen, S., 1993. Managing on the twelfth: a curious plan. *Harv. Educ. Rev.* 63 (3), 339–358.
- Cvetkovich, A., 2003. *An Archive of Feeling: Trauma, Sexuality, and Lesbian Public Cultures*. Duke University Press, Durham, London.
- Davis, B., Sumara, D., Luce-Kapler, R., 2000. *Engaging Minds: Learning and Teaching in a Complex World*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Davis, B., Sumara, D., D'Amour, L., 2012. Understanding school districts as learning systems: some lessons from three cases of complex transformation. *J. Educ. Change* 13, 373–399.
- Derrida, J., 1978. *Writing and Différance*. Routledge, London.
- Edelman, G., 2004. *Wider than the Sky: The Phenomenal Gift of Consciousness*. Yale University Press, New Haven, London.
- Foucault, M., 1988. Technologies of the self. In: Luther, M., Gutman, H., Hutton, P. (Eds.), *Technologies of the Self: A Seminar with Michel Foucault*. University of Massachusetts Press, Amherst, MA, pp. 16–49.
- Franklin, U., 1990. *The Real World of Technology*. Anansi Press, Concord, ON.
- Friesen, S., Jacobsen, M., 2020. The education system of Canada: foundations of the Canadian education system. In: Jornitz, S., Parreira do Amaral, M. (Eds.), *Global Education Systems: The Education Systems of the Americas*. Springer.
- Gadamer, H.G., 1990. In: Weinsheimer, J., Marshall, D. (Eds.), *Truth and Method*, second revised ed. Crossroad, New York.
- Gallagher, S., 1992. *Hermeneutics and Education*. SUNY Press, Albany.
- Gereluk, D., 2008. *Symbolic Clothing in Schools: What Should Be Worn and Why*. Continuum, London.
- Grumet, M., 1988. *Bitter Milk: Women and Teaching*. University of Massachusetts Press, Amherst.
- Harrari, Y., 2018. *21 Lessons for the 21st Century*. Random House, Toronto.
- Hayles, N.K., 2005. *My Mother was a Computer: Digital Subjects and Literary Texts*. University of Chicago Press, Chicago.
- Heidegger, M., 1966. *Being and Time*. Harper and Row, New York.
- Liotard, J.-F., 1984. *The Postmodern Condition: A Report on Knowledge*. Minnesota Press, Minneapolis.
- Meretoja, H., 2018. *The Ethics of Storytelling: Narrative Hermeneutics, History and the Possible*. Oxford University Press, New York.
- Probyn, E., 2005. *Blush: Faces of Shame*. University of Minnesota Press, Minneapolis.
- Smith, D., 1991. The hermeneutic imagination and the pedagogic text. In: Short, E. (Ed.), *Forms of Curriculum Inquiry*. SUNY Press, New York.
- Sumara, D., Davis, B., Laidlaw, L., 2001. Canadian identity and curriculum: a postmodern perspective. *Can. J. Educ.* 26, 21–34.
- Sumara, D., 1996. *Private Readings in Public: Schooling the Literary Imagination*. Peter Lang, New York.
- Sumara, D., 2020. On the power of not passing: a queer hermeneutics of higher education leadership. *J. Educ. Hist. Adm.* URL <http://mc.manuscriptcentral.com/cjeh>.
- Varela, F., 1999. *Ethical Know-How: Action, Wisdom, and Cognition*. Stanford University Press.

Researching teacher learning through self-video

Jenni Ingram^a and Alf Coles^b, ^a Department of Education, University of Oxford, Oxford, United Kingdom; and ^b School of Education, Helen Wodehouse Building, Bristol, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	120
Theoretical frameworks	121
Designs	122
Tensions, challenges and future directions	123
Conclusion	124
References	125

Introduction

This article summarizes research that considers teacher learning through professional development involving teacher self-video, that is, where teachers reflect upon and work with videos of their own practice. We examine the theoretical frameworks and models of teacher learning that underpin the design of this research and the associated professional development activities, and the influences these designs have on the roles that researchers, facilitators, experts and teachers have in the learning process. Drawing upon recent literature reviews and meta-studies of video use in teacher professional development, examples of aspects of different designs are used to illustrate the distinctions and commonalities that influence teacher learning from videos of their own teaching. To end we raise issues and challenges both researchers and teachers confront when working with self-video and the tensions that arise.

The use of video in teachers' learning and growth is increasing and has generated much interest since the last edition of this encyclopedia. The availability and capability of video technology has also evolved, coinciding with increasing concerns around privacy and ethics in video use in particular. There is accumulating evidence of the potential of video-based teacher learning (Borko et al., 2008; Sherin and van Es, 2009). However, the opportunities and challenges associated with using video in teacher professional development are continuously evolving which can pose a challenge to both the teachers involved, and the researchers examining the complex roles that videos may play. In this article the focus is on teacher learning through self-video, Major and Watson's (2018) review shows this to be the most widely used way of working with videos in teacher professional development.

In this article, which is focused on reviewing a range of recent research, we do not define what we take learning, change or growth to mean. We note they can be used synonymously and furthermore we concur with the conclusion of a recent review: "Given that the aims of video use center around change for teachers, leading to change in the classroom and change for learners, some theory of change or learning must be present, either implicitly or explicitly" (Coles et al., 2019, p. 127). In the section on theoretical frameworks, below, we aim to draw out some of these theories.

Videos are amenable to being used in a variety of ways by both teachers and researchers. Yet, in all professional development research there is a tension between the roles played by teachers and researchers. If the primary purpose of using video is teachers' learning, then as researching this process is secondary there will be restrictions and limitations on the data available and the questions that can be asked. The richness of the ways videos capture classroom interactions, including interactions with visual or embodied representations, allows for both researchers and teachers to revisit the data with different perspectives, focuses and purposes.

The complexity of the data collected through videos, as does the act of videoing itself, also involves extensive data reduction as some aspects will be foregrounded over others. A considerable amount of research into the use of videos, including self-videos, in teacher education has consequently focused on noticing and awareness (e.g. van Es and Sherin, 2002; Mason 2002; Ingram, 2015). This research is largely driven by two aims. Firstly, supporting teachers to notice a particular, possibly desirable, aspect of practice such as students' thinking, pedagogic practices that are the target of the professional development program, or cultural practices that contrast with their own. Secondly, starting from where the teacher is and what they notice and working with this to explore ways of acting differently (Mason 2002; Ingram and Andrews, 2020). This second approach needs researchers or facilitators to adapt and personalize (Muir et al., 2010) to each individual teacher as they reflect upon videos of their own practice, which does not bring the stability and controls that researchers often seek, and limits the potential of generalizing beyond the individual teacher.

Video enables teachers and teacher educators to watch, and re-watch episodes, practices or phenomena of interest which opens up opportunities to explore, prove and reflect on classroom practices, either a teacher's own practices, or in collaboration with others (Sherin, 2004). As technology develops teachers have access to richer representations of classroom activity but this is accompanied by increasingly complex ethical considerations, the development of new ways of working with teachers, and further challenges which are explored in the final section.

Theoretical frameworks

A variety of theoretical frameworks are drawn upon by researchers interested in video in teacher learning. Many studies bring different theories together, with one theory influencing the design in relation to teacher learning, and another relating to the focus of what is being watched or reflected upon. The use of video is not separable from a theoretical orientation (Coles et al., 2018), and the focus on video in researching teacher learning highlights the diversity of theoretical approaches where terms, such as noticing, may be used in different ways or different words may be used to describe or label similar phenomena (Bikner-Ahsbabs, 2009). Research with teachers (rather than research on teachers) tends also not to make the theoretical basis of the research explicit as there is a need for the research to be of interest and use by teachers themselves in diverse contexts. For example, much of the research discussed below draws from a phenomenological stance where teachers engage with tasks they might use with their own students that position the teachers themselves as learners. This perspective addresses one of the key difficulties researchers face when working with video, that of ensuring that the teachers feel that the videos and tasks associated with video watching, are relevant to their actual classroom practice (Hamel and Viau-Guay, 2019).

Research on teacher learning through self-video has led to the development of frameworks specifically focusing on the professional development process. Examples include the “learning to notice” framework developed by van Es and Sherin (2002) or the “Lesson Analysis Framework” developed by Santagata and Angelici (2010). These frameworks bring together the need to have a purpose to the work with video, the need to choose suitable videos, clips or materials, and tasks that are consistent with this purpose and consideration of the role of the facilitator or expert in this process, which is explored further in the design section below. Many of these frameworks have developed out of work with video cases of “expert” teaching, but are also structures that underpin much of the work with self-video or other types of professional development that focus on the teacher’s own practice. The shift from videos of other teachers to self-video often also shifts the focus to the problems of classroom practice, rather than on “best practice” (e.g., Horn, 2019; Zhange et al., 2011). The purpose becomes that of solving problems, developing teacher decision making (Schoenfeld, 2017) or pedagogical judgment (Horn, 2019).

Case methods have a long history in professional development, including teacher professional development (Sato and Rogers, 2010). With the rise of video use in teacher education, cases have shifted from being predominantly narrative written cases to video-based materials, though written and visual vignettes remain a core research tool in educational research (Skilling and Stylianides, 2019). In teacher education, video work based on the case method approach often focused on the use of videos of “expert” or desirable practice (e.g. Chen et al., 2020; Gröschner et al., 2015). More recently, there has been a shift toward cases where the teachers have selected what the clip is a “case of”, usually an aspect of practice that is the focus of the professional development they are engaging in, or a problematic moment from the teachers’ own lessons (e.g., Hollingsworth and Clarke 2017; Ingram and Andrews, 2020). These approaches to working with cases of practice have led to the development of frameworks that focus on broader aspects of practice to reflect the need to work with cases that teachers find relevant to their practice (e.g., Hollingsworth and Clarke, 2017; Prediger and Neugebauer, 2020; Schoenfeld, 2018). This way of working with case methods has also led to the development of models and theories of learning that we turn to next.

Teacher learning is a complex area of research that continues to evolve over time. There has been a shift in focus from teaching as a set of skills and to a focus on connecting theory and practice, collaborative discussions, and learning continuing over time (e.g., Mutton et al., 2017). While there is considerable research around teachers’ knowledge, attitudes or beliefs, in isolation of practice, research focusing on the use of self-video by teachers often brings beliefs, goals and practice together. Two models of teacher learning widely used in research in teacher learning using video are the Interconnected Model of Teacher Professional Growth developed by Clarke and Hollingsworth (2002) and Schoenfeld’s theory of goal-oriented decision making (2010). In both these models there is a focus on the relationship between knowledge, beliefs and goals (what Clarke and Hollingsworth refer to as the *personal domain*) with teachers’ practice (the *domain of practice*), and the role of reflection or monitoring and enactment or decision making in changing the personal domain and/or the domain of practice. These models allow research to consider teacher learning and growth over time while retaining the complexity of the relationships between knowledge, beliefs, goals and practice. Where they differ perhaps is in the role that the researcher or facilitator plays in teacher learning, and we return to this aspect below.

Noticing underpins a great deal of research into teacher learning using video. Yet researchers have conceptualized teacher noticing in a variety of ways, focusing on the sense teachers make of what they notice in the classroom (van Es and Sherin, 2002), to the classroom objects which teachers attend to (e.g., Star and Strickland, 2008; Luna et al., 2018). Mason (1984, 2002) argues that in order for teachers to develop their professional practice, they need to first develop their own awareness and noticing of experiences in the classroom. This sense of working on or with teachers’ awareness and sensitivities is drawn upon in the “learning to notice” framework described above. Yet for Mason the starting point is what teachers attend to or notice in their own practice and then working systematically, in a disciplined and intentional manner, to notice opportunities for acting differently in the future. While there are similarities here to Sherin (2004)’s ideas of considering antecedents, implications and action possibilities, there is a fundamental difference in what is being noticed. For Sherin (2004) and others drawing on her work, the research into teachers’ learning focuses on their noticing of student thinking often related to a targeted topic of the researcher’s choosing (e.g., Lam and Chan, 2020).

The theoretical frameworks and models used when researching teacher learning through self-video underpin both the design of research studies, and the design of the professional development within which the teachers are learning. It is these designs that we turn to next.

Designs

In this section, we aim to offer examples of how choices of theoretical framework influence the design of ways of working with video. However, before getting to this detail, we offer some general findings, based on meta-studies of relevant literature, about the design of work with teachers and video. Reviews of video-based work with teachers (Major and Watson, 2018; Hamel and Viau-Guay, 2019) report that a key effect of using video is a broadening of teachers' modes of reflections on teaching, or, changes in perception of the classroom, for instance, increased awareness of student reasoning (Major and Watson, 2018). Relatively few studies to date have traced the impact of work with video into the classrooms of participating teachers. Hamel and Viau-Guay (2019) found that reflections based on video were judged to be better grounded than those based on recalling incidents, or on written case studies. Major and Watson (2018) similarly found evidence of enhanced reflection. In other words, we could surmise that the promise of the use of video, which was noted as far back as the 1960s, seems to be well-founded. In particular, video-based work offers powerful opportunities for developing skills of reflection, which are vital for on-going teacher learning. However, it is also clear that there is nothing inevitable about teacher learning as a result of engaging in work with video (Major and Watson, 2018).

Tripp and Rich (2012)'s review, which focused on what we are calling self-video, found that working on video as part of a group, where there were interactions about the video, was preferred by teachers, compared to engaging as an individual; a finding confirmed in Hamel and Viau-Guay (2019). And, where there was group interaction, a consistent finding is that high quality support (e.g., in the form of a skilled facilitator) is necessary to enhance the quality of teacher reflections and subsequent learning (Gaudin and Chaliès, 2015; Major and Watson, 2018; Hamel and Viau-Guay, 2019). It is also the case that few studies document the facilitator's contribution to the mechanism of video-based reflection. The role of the facilitator of teacher discussion of video appears as a critical and yet under-researched issue. We know a facilitator can have a significant impact on teacher learning from video, but we do not yet know how or why – this is an issue we return to in the final "Challenges" section.

Several of the reviews cited above distil dimensions of video use. For instance, Tripp and Rich (2012) offer the following questions, to guide decisions about mechanisms or designs of video-based work with teachers, which offer a neat summary of common issues:

What type of reflection tasks will I ask teachers to engage in during their video reflections? Will I provide teachers with a framework to guide their reflections? Will I ask teachers to reflect individually, collaboratively, or both? What length of video will teachers use for reflection? How many times will teachers reflect on their videos? What methods will I use to determine if video was beneficial for teacher reflection?

Source: From p. 681.

Different approaches and answers to these questions will be appropriate in different contexts. In the remainder of this section, we show how answers to these questions are also driven by theoretical frameworks. We consider three frameworks: two approaches to noticing and one based on the Theory of Didactical Situations (Brousseau, 1997). We cannot be exhaustive in considering frameworks, hence this choice of focus was made to exemplify how frameworks influence design.

As noted above, the word "noticing" is used commonly in research on video with teachers, particularly within mathematics education and increasingly within science education and other sub-fields of education. There are at least two distinct groupings of meaning and each one brings its own implications for the design of work on video and for the roles of participants and facilitators. The two groupings, which we discuss further below, are *teacher noticing*, the phrase adopted by Sherin et al. (2012), and the *discipline of noticing*, in the work of Mason (2012).

The notion of teacher noticing is an aspect of teachers' expertise that includes teachers' decision making during the process of teaching, i.e., what teachers attend to, in-the-moment of teaching, how they reason about what they notice and how they make choices about what to do next. This image of the process of moving from perception to action mirrors a view of cognition as the receipt of information from the environment, the processing of that information internally and a subsequent response. Given this perspective on cognition, it makes sense that work with video can be directed toward teachers becoming more skilled at noticing particular features of the classroom environment. van Es and Sherin (2002), for instance, worked with teachers to develop their noticing of students' mathematical reasoning. When video work has a specific focus, the facilitator needs to be expert at noticing the particular focus and then needs skills to guide and direct teachers toward what might be available to be learnt from the video. Teachers will be in the role of expanding their noticing through interaction with a more expert other. The choice of video clip is significant, in allowing opportunities for the noticing of what is in focus.

For Mason (2012), noticing is about intentional acts: "This is the essence of the discipline of noticing: arranging to alert yourself in the future so as to act freshly rather than automatically out of habit" (p. 37). In other words, the noticing within the discipline of noticing is not so much a feature of expertise (as it is in *teacher noticing*) but a feature of learning to become expert. Mason's form of noticing is about learning from experience and what we can do as individuals in order to learn from experience. The discipline of noticing, as a perspective on human learning, sparked a particular manner of working with video (Jaworski, 1990; Coles, 2013) with a method of attending to the detail of events, as a mechanism to subvert participants' habitual responses to what they observe. Work on video, using this perspective, begins with a reconstruction of what took place (and hence only short video clips of up to 4 min can be used) before moving into a more interpretive phase of discussion about the clip. "Instead of getting people to analyze practices observed on the tapes, we found it more effective to use the tapes as stimuli to get people to recall

and then analyze related incidents from their own teaching.” (Mason, 2012, p. 36). The facilitator role, within the discipline of noticing, requires expertise in managing the two phases of discussion, but does not require expertise in the particular focus of discussion. The focus for the facilitator needs to be on working with what the teachers notice and supporting teachers in alerting themselves to future possibilities. Mason’s approach to noticing has distinct methodological implications. Firstly, the emphasis on the teachers’ own noticing requires that the teachers themselves select the clips from videos of their practice based on where they might want to act differently, not the researchers. It is also the teachers’ reactions to these videoed events that is of interest, not the nature of the event itself.

The Theory of Didactical Situations (TDS) (Brousseau, 1997) is a perspective on teaching and learning that is used to inform work with video, particularly in France, sometimes alongside influences from activity theory (Horoks and Robert, 2007). A key principle for video use, deriving from the TDS, is the need for viewers of a classroom video to perform (prior to watching) an *a priori analysis* of the task which students will be working on during the video clip. This analysis includes the identification of what knowledge students are likely to have, which will need to be adapted, in order to be successful on the task they are given. It is apparent therefore that significant detail about the context is required for this analysis. Some further elements of the TDS may also be introduced to teachers prior to their watching of video, to inform and enrich the subsequent viewing. Quite specific demands are placed on suitable videos for watching using the TDS. A task such as “How many numbers are there between zero and one?” is too broad to allow *a priori* analysis (Coles et al., 2018), i.e., it would not be possible to analyze the adaptations students might need to make to their existing knowledge to answer this, since their responses are potentially so diverse. The challenge (and hence video) would need to be on a question such as “Give me a fraction between a third and a half?” (Coles et al., 2018), where likely responses could be more easily predicted. The facilitator role, using the TDS with video, requires expertise in the TDS itself and a subsequent careful choice of video clip and preparation of pre-watching tasks. There is potentially a delicate skill in terms of how much of the TDS to introduce in any given context. Participant teachers will be in the role of both learning from the video itself and also learning about the TDS.

The brief illustrations, of how three different theoretical frameworks influence quite different designs for working with video, show that decisions in relation to questions such as those posed by Tripp and Rich (2012), quoted above, are not made in isolation. We suggest that working effectively with video flows from principled design, in which ways of working are consistent with a perspective on learning that informs the actions of the facilitator.

Tensions, challenges and future directions

In this final section we turn to discussions around the tensions and challenges we have personally experienced working as researchers, teacher educators, and as teachers ourselves when working with self-video in a variety of ways. These include considering the role of the facilitator when working collaboratively with videos, the nature of expertise within research on teacher learning through self-video, issues around mimicry or practice, and the complex ethical processes involved.

One challenge we have already alluded to in more than one place concerns the role of the facilitator of teachers working with video. We know this role is vital if teachers are to learn from video (Major and Watson, 2018, p. 65) and yet we know little about it. This is an area of research being given increasing attention, particularly within mathematics education. A significant contribution to the literature is a framework developed by van Es et al. (2014), which articulates four central practices and 14 facilitation “moves”. The four practices are: orienting the group to the video analysis task; sustaining an inquiry stance; maintaining a focus on the video and the (subject); supporting group collaboration. While the word “video” appears twice in these practices, they appear to capture quite general aspects of facilitating discussion, not just within mathematics education. A significant question arising from this framework is how a facilitator comes to gain the kind of expertise represented by the practices and moves.

Research into the ways facilitators acquire the knowledge they need points to the significance of reflection on their practice (Beswick and Goos, 2018), either alone (e.g., Krainer, 2008) or supported by other, experienced colleagues (e.g., Chen et al., 2018). The significance of reflection mirrors the benefits, for teachers, of working with self-video. Indeed, research points to opportunities for facilitators to learn about their own practice from researching their own work with teachers (Chapman, 2008; Chen et al., 2018; Jaworski, 2003). There are a small number of programmes designed to enhance facilitators’ skills (e.g., Childs et al., 2014). Another example is the VIDEO-LM project (Viewing, Investigating and Discussing Environments of Learning Mathematics), developed at the Weizmann Institute of Science in Israel. VIDEO-LM makes use of a bespoke analytic framework of six “lenses”, shared with participating teachers, to guide discussion (Karsenty and Arcavi, 2017; Karsenty, 2018). The project also developed six lenses to inform facilitators. The scale of the work (reaching over 1000 teachers to date) necessitated training teachers to facilitate other teachers. The lenses offered to teachers were adapted for facilitators, to support their facilitation (Karsenty, 2020).

It is important to note that not all contexts need facilitation, for instance, when all participants in a video-viewing group are skilled at providing appropriate critique (Barton et al., 2015). In such cases facilitation can become a task shared and led by participants, often the teacher represented in the video (who might select the clip to be watched and the focus for observation).

We are conscious that the research we have reviewed has largely been conducted in the global North, and predominantly in English. There are significant research questions about the role and influence of different cultural contexts and traditions of teacher leadership or development. For instance, there is a huge wealth of expertise within China and Japan around the facilitation of lesson study group discussions (following a live viewing of a lesson rather than a video recording) and we suggest that insights from this body of work could be highly relevant to increasing our understanding of facilitation of work with teachers on self-video.

One key difference between teacher professional development that makes use of videos of “expert” practice and teacher professional development that uses self-video is in the conceptualization of expertise. This difference is also evident when distinguishing between research with teachers and research on teachers. In the intersection between teacher professional development and research, the purposes or intentions of each participant during the project may differ and perhaps contrast. This refers to what [Ingram and Andrews \(2018\)](#) call the “tension of intentions” where the teacher educators, researchers and teachers have different intentions and different perceptions of others’ intentions.

One of the places where these tensions are most evident is in the recommendations made by many researchers around the use of video. For example, [Sherin et al. \(2009\)](#) identify three features of video clips that they argue support productive discourse around students’ thinking: (a) the extent to which the video clip provides windows into student thinking; (b) the depth of thinking shown; (c) the clarity of student thinking shown in video. Yet it is the researcher who makes the decisions around which clips have these properties which pre-supposes the expertise of the researcher and involves value judgements and interpretations of what students say and do, while the teacher who has worked with the class for some time may identify different clips drawing upon their own expertise of the classroom norms and the students within the class. Researchers can also avoid clips where there is ambiguity in what students may be thinking, which can narrow the possibilities for making conjectures, exploring different interpretations or perspectives.

One potential danger of the increasing number and detail of frameworks, both for teachers using video and to guide the actions of facilitators, is highlighted by [Gaudin and Chalies \(2015\)](#) as the “mimicking” effect (p. 56). They summarize this effect as follows: “the high degree of scaffolding used in many studies may (...) be prompting (teachers) to simply express more of what they perceive are the “right” reform-minded ideas” (p. 56). In that meta-study, the authors propose giving space to allow teachers to explore and express what they observe, within the language and practical or theoretical frameworks they use to make everyday decisions in the classroom. We return to the question of the link between working with video and classroom practice in the Conclusion.

The rise of video and other technology has been accompanied by a growing concern around privacy and the ethical obligations to both teachers and students. In the use of self-video in research, there is not a clear dividing line between research, professional development, and normal practice. This makes the articulation of what students and teachers are consenting to participate in challenging and poses problems for ensuring that consent (or not) arises from an informed decision that is freely given ([Fancourt, 2017](#); [Locke et al., 2013](#)). Teacher learning and growth is also not predictable or controllable, which makes anticipating changes to practice during the research an additional challenge. Technology, and in particular videos, also complicates the need to protect the confidentiality and anonymity of the teachers and students whose lessons are being recorded.

Researching teacher learning occurs in a context where the values of the researcher, teacher(s) and communities intersect. These values influence what is noticed, what is “normal” and what is emphasized in work with self-video. There is a risk that both the researcher and the teacher will normalize what fits with their own cultural and educational values, while pathologizing other ways of knowing or working. School disciplines, for example science ([Otrell-Cass, 2020](#)) or history ([Mycock, 2017](#)), also have particular societal and political consequences that reach beyond the classroom that are often visible through the enactment of values in research and teaching.

The ethics of analysis and interpretation come to the fore in professional development itself and the research of teachers’ learning through PD. Each participant is involved in the interpretation of the videos recorded. Teachers engage in self-analysis, which may or may not be influenced by their learning during the professional development. Researchers interpret and analyze the actions of the teachers and students within the video, and the reactions and re-actions of the teachers when viewing the videos. Yet the ethical dimensions of these interpretations are undertheorized. Ethical processes are often restricted to the stage before data collection begins where consent is sought, but the dynamic nature of teacher learning means that the research teachers are participating and engaging with changes throughout the process – so what is it that teachers are consenting to?

Values are also fundamental to the way in which the research is reported. There are layered obligations in all research involving teachers. Appreciative models of enquiry work with teachers, to showcase what they do well. Yet the job of research is to do more than this. Research provides alternative views and potentially a more critical lens which may frame some practices in a negative light, what [Locke et al. \(2013\)](#) refer to as “bad news”. Discussions around practice often need to involve questioning and challenging, contrasting views and potential actions ([Grossman et al., 2001](#)). Video-based learning also involves teachers making their practice public which involves risk ([Vedder-Weiss et al., 2019](#)) and some teachers may feel threatened ([Borko et al., 2008](#)).

Conclusion

In this article we have summarized research that considers teacher learning through professional development involving teacher self-video. We have used examples from research to illustrate the theoretical frameworks and models of teacher learning that underpin recent designs of this research and the associated professional development activities, and the influences these designs have on the roles that researchers, facilitators, experts and teachers have in the learning process. These illustrations highlight the distinctions and commonalities that influence teacher learning from videos of their own teaching. As the use of technology and video evolves and develops the complexities of research designs, including considerations of the roles of researchers and teachers in this process, also need to evolve and develop. The use of self-video has been shown to be an effective influence on teacher learning, particularly when combined with facilitation consistent with what the teachers identify as relevant to or the problems of their own practice. Yet this effectiveness requires researchers to be flexible and responsive in their design and facilitation, raising

new and complex challenges and tensions in this research process. Further research is needed both to theorize use of self-video in teacher learning and into designs of facilitation that support this learning.

References

- Barton, B., Oates, G., Paterson, P., Thomas, M.O.J., 2015. A marriage of continuance: professional development for mathematics lecturers. *Math. Educ. Res. J.* 27 (2), 147–164.
- Beswick, K., Goos, M., 2018. Mathematics teacher educator knowledge: what do we know and where to from here? *J. Math. Teach. Educ.* 21 (5), 417–427.
- Bikner-Ahsbals, A., 2009. Networking of theories—why and how? Special plenary session. In: Durand-Guerrier, V., Soury-Lavergne, S., Lecluse, S. (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education*. ERME, Lyon, France, pp. 6–15.
- Borko, H., Jacobs, J., Eiteljorg, E., Pittman, M.E., 2008. Video as a tool for fostering productive discussions in mathematics professional development. *Teach. Teach. Educ.* 24 (2), 417–436.
- Brousseau, G., 1997. *Theory of didactical situations in mathematics: didactique des mathématiques (1970–1990)*. Springer, Rotterdam.
- Chapman, O., 2008. Mathematics teacher educators' learning from research on their instructional practices. In: Jaworski, B., Wood, T. (Eds.), *International Handbook of Mathematics Teacher Education: Vol. 4: The Mathematics Teacher Educator as a Developing Professional*. Sense Publishers, Rotterdam, pp. 110–129.
- Chen, J.-C., Lin, F.-L., Yang, K.-L., 2018. A novice mathematics teacher educator—researcher's evolution of tools designed for in-service mathematics teachers' professional development. *J. Math. Teach. Educ.* 21 (5), 517–539.
- Chen, G., Chan, C., Chan, K., Clarke, S., Resnick, L., 2020. Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *J. Learn. Sci.* 29 (4–5), 642–680. <https://doi.org/10.1080/10508406.2020.1783269>.
- Childs, A., Hillier, J., Thornton, S., Watson, A., 2014. Building capacity: developing a course for mathematics and science teacher educators. In: Maaß, K., Barzel, B., Törner, G., Wernisch, D., Schäfer, E., Reitz-Koncebowski, K. (Eds.), *Educating the Educators: International Approaches to Scaling-Up Professional Development in Mathematics and Science Education*. WTM, Munster, pp. 47–55.
- Clarke, D., Hollingsworth, H., 2002. Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18 (8), 947–967.
- Coles, A., 2013. Using video for professional development: the role of the discussion facilitator. *J. Math. Teach. Educ.* 16 (3), 165–184. <https://doi.org/10.1007/s10857-012-9225-0>.
- Coles, A., Horoks, J., Chesnais, A., 2018. Theory and the role of the mathematics teacher educator: a comparison of teacher education sessions in France and UK using video. *Annales de didactique et de Sci. Cogn. Special Issue English-French*, 119–144. Retrieved from <https://mathinfo.unistra.fr/irem/publications/adsc/volumes/#c14903>.
- Coles, A., Karsenty, R., Beswick, K., Oates, G., Abdulhamid, L., 2019. The use of video for the learning of teachers of mathematics. In: Graven, M., Venkat, H., Essien, A.A., Vale, P. (Eds.), *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education*, Pretoria, South Africa, 7–12 July 2019, vol. 1. PME, Pretoria, South Africa, pp. 107–136.
- Fancourt, N., 2017. Research ethics in closing the gap. In: Childs, A., Menter, I. (Eds.), *Mobilising Teacher Researchers: Challenging Educational Inequality*. Routledge, London.
- Gaudin, C., Chaliès, S., 2015. Video viewing in teacher education and professional development: a literature review. *Educ. Res. Rev.* 16, 41–67.
- Gröschner, A., Seidel, T., Kiemer, K., Pehmer, A.K., 2015. Through the lens of teacher professional development components: the 'dialogic video cycle' as an innovative program to foster classroom dialogue. *Prof. Dev. Educ.* 41 (4), 729–756. <https://doi.org/10.1080/19415257.2014.939692>.
- Grossman, P., Wineburg, S., Woolworth, S., 2001. Toward a theory of teacher community. *Teach. Coll. Rec.* 103, 942–1012.
- Hamel, C., Viau-Guay, A., 2019. Using video to support teachers' reflective practice: a literature review. *Cogent Educ.* 6 (1), 1673689.
- Hollingsworth, H., Clarke, D., 2017. Video as a tool for focusing teacher self-reflection: supporting and provoking teacher learning. *J. Math. Teach. Educ.* 20, 457–475. <https://doi.org/10.1007/s10857-017-9380-4>.
- Horn, I.S., 2019. Supporting the development of pedagogical judgment. In: Lloyd, G.M., Chapman, O. (Eds.), *International Handbook of Mathematics Teacher Education*, second ed., vol. 3, pp. 321–342.
- Horoks, J., Robert, A., 2007. Tasks designed to highlight task-activity relationships. *J. Math. Teach. Educ.* 10 (4–6), 279–287.
- Ingram, J., 2015. Supporting student teachers in developing and applying professional knowledge with videoed events. *Eur. J. Teach. Educ.* 37 (1), 51–62.
- Ingram, J., Andrews, N., 2018. Tensions and opportunities when working in a collaborative video group. In: Curtis, F. (Ed.), *Proceedings of the British Society for Research into Learning Mathematics*, vol. 38 (2).
- Ingram, J., Andrews, N., 2020. Trying to improve communication skills: the challenge of joint sense making in classroom interactions. In: Ingram, J., Erath, K., Rønning, F., Schüller-Meyer, A.K. (Eds.), *Proceedings of the Seventh ERME Topic Conference on Language in the Mathematics Classroom*. University of Montpellier and ERME, Montpellier, France, pp. 147–154.
- Jaworski, B., 1990. Video as a tool for teachers' professional development. *Br. J. Serv. Educ.* 16 (1), 60–65.
- Jaworski, B., 2003. Research practice into/influencing mathematics teaching and learning development: towards a theoretical framework based on co-learning partnerships. *Educ. Stud. Math.* 54, 249–282.
- Karsenty, R., Arcavi, A., 2017. Mathematics, lenses and videotapes: a framework and a language for developing reflective practices of teaching. *J. Math. Teach. Educ.* 20 (5), 433–455.
- Karsenty, R., 2018. Professional development of mathematics teachers: through the lens of the camera. In: Kaiser, G., Forgasz, H., Graven, M., Kuzniak, A., Simmt, E., Xu, B. (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education*. Springer, Hamburg, pp. 269–288.
- Karsenty, R., 2020. The role of frameworks in researching knowledge and practices of mathematics teachers and teacher educators. In: Potari, D., Ribeiro, M., Zehetmeier, S. (Eds.), *Mathematics Teaching and Mathematics Teacher Professional Development*. Routledge, London.
- Krainer, K., 2008. Reflecting the development of a mathematics teacher educator and his discipline. In: Jaworski, B., Wood, T. (Eds.), *International Handbook of Mathematics Teacher Education*, vol. 4. Sense Publishers, Rotterdam, pp. 177–199.
- Lam, D.S.H., Chan, K.K.H., 2020. Characterising pre-service secondary science teachers' noticing of different forms of evidence of student thinking. *Int. J. Sci. Educ.* 42 (4), 576–597. <https://doi.org/10.1080/09500693.2020.1717672>.
- Locke, T., Alcorn, N., O'Neill, J., 2013. Ethical issues in collaborative action research. *Educ. Action Res.* 21 (1), 107–123. <https://doi.org/10.1080/09650792.2013.763448>.
- Luna, M.J., Selmer, S.J., Rye, J.A., 2018. Teachers' noticing of students' thinking in science through classroom artifacts: in what ways are science and engineering practices evident? *J. Sci. Teach. Educ.* 29 (2), 148–172.
- Major, L., Watson, S., 2018. Using video to support in-service teacher professional development: the state of the field, limitations and possibilities. *Technol. Pedagog. Educ.* 27 (1), 49–68.
- Mason, J., 1984. Towards one possible discipline for mathematics education. In: Steiner, H. (Ed.), *Theory of Mathematics Education*. Institute for the Didactics of Mathematics, Bielefeld University, Bielefeld, Germany, pp. 42–55.
- Mason, J., 2002. *Researching Your Own Practice: The Discipline of Noticing*. Routledge-Falmer, New York.
- Mason, J., 2012. Noticing: roots and branches. In: Sherin, M.G., Jacobs, V.R., Philipp, R.A. (Eds.), *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*. Erlbaum, Mahwah, New Jersey, pp. 35–50.

- Mutton, T., Burn, K., Menter, I., 2017. Deconstructing the Carter Review: competing conceptions of quality in England's 'school-led' system of initial teacher education. *J. Educ. Pol.* 32 (1), 14–33. <https://doi.org/10.1080/02680939.2016.1214751>.
- Muir, T., Beswick, K., Williamson, J., 2010. Up, close and personal: teachers' responses to an individualised professional learning opportunity. *Asia Pac. J. Teach. Educ.* 38 (2), 129–146. <https://doi.org/10.1080/13598661003677598>.
- Mycock, A., 2017. After Empire: the politics of history education in a post-colonial world. In: Carretero, M., Berger, S., Grever, M. (Eds.), *Palgrave Handbook of Research in Historical Culture and Education*. Palgrave Macmillan, London. https://doi.org/10.1057/978-1-137-52908-4_21.
- Otrell-Cass, K., 2020. The performativity of ethics in visual science education research: using a material ethics approach. In: Otrell-Cass, K., Andrée, M., Ryu, M. (Eds.), *Examining Ethics in Contemporary Science Education Research*. Cultural Studies of Science Education, vol. 20. Springer, Cham. https://doi.org/10.1007/978-3-030-50921-7_12.
- Prediger, S., Neugebauer, P., 2020. Capturing teaching practices in language-responsive mathematics classrooms Extending the TRU framework "teaching for robust understanding" to L-TRU. *ZDM Math. Edu.* 53, 289–304. <https://doi.org/10.1007/s11858-020-01187-1>.
- Santagata, R., Angelici, G., 2010. Studying the impact of the lesson analysis framework on preservice teachers' abilities to reflect on videos of classroom teaching. *J. Teach. Educ.* 61 (4), 339–349. <https://doi.org/10.1177/0022487110369555>.
- Sato, M., Rogers, C., 2010. Case methods in teacher education. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed., pp. 592–597.
- Schoenfeld, A.H., 2017. Uses of video in understanding and improving mathematical thinking and teaching. *J. Math. Teach. Educ.* 20, 415–432. <https://doi.org/10.1007/s10857-017-9381-3>.
- Schoenfeld, A.H., 2018. Video analyses for research and professional development: the teaching for robust understanding (TRU) framework. *ZDM Math. Edu.* 50, 491–506.
- Sherin, M.G., van Es, E.A., 2009. Effects of video club participation on teachers' professional vision'. *J. Teach. Educ.* 60 (1), 20–37. <https://doi.org/10.1177/0022487108328155>.
- Sherin, M.G., 2004. New perspectives on the role of video in teacher education. In: Brophy, J. (Ed.), *Using Video in Teacher Education*. Emerald Group Publishing Limited, pp. 1–28.
- Sherin, M., Jacobs, V., Philipp, R., 2012. Situating the study of teacher noticing. In: Sherin, M.G., Jacobs, V.R., Philipp, R.A. (Eds.), *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*. Erlbaum, Mahwah, New Jersey, pp. 3–14.
- Sherin, M.G., Linsenmeier, K.A., van Es, E.A., 2009. Selecting video clips to promote mathematics teachers' discussion of student thinking. *J. Teach. Educ.* 60 (3), 213–230. <https://doi.org/10.1177/0022487109336967>.
- Skilling, K., Stylianides, G.J., 2019. Using vignettes in educational research: a framework for vignette construction. *Int. J. Res. Method Educ.* 43 (5), 541–556. <https://doi.org/10.1080/1743727X.2019.1704243>.
- Star, J., Strickland, S., 2008. Learning to observe: using video to improve preservice mathematics teachers' ability to notice. *J. Math. Teach. Educ.* 11, 107–125.
- Tripp, T., Rich, P., 2012. Using video to analyze one's own teaching. *Br. J. Educ. Technol.* 43, 678–704.
- van Es, E.A., Sherin, M.G., 2002. 'Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *J. Technol. Teach. Educ.* 10 (4), 571–596.
- van Es, E., Tunney, J., Goldsmith, L.T., Seago, N., 2014. A framework for the facilitation of teachers' analysis of video. *J. Teach. Educ.* 65 (4), 340–356.
- Vedder-Weiss, D., Segal, A., Lefstein, A., 2019. Teacher face-work in discussions of video-recorded classroom practice: constraining or catalyzing opportunities to learn? *J. Teach. Educ.* 70 (5), 538–551. <https://doi.org/10.1177/0022487119841895>.
- Zhang, M., Lundeborg, M., Eberhardt, J., 2011. Strategic facilitation of problem-based discussion for teacher professional development. *J. Learn. Sci.* 20 (3), 342–394. <https://doi.org/10.1080/0508406.2011.553258>.

Retention of teachers from minority ethnic groups in disadvantaged schools

Antonina Tereshchenko and Martin Mills, Centre for Teachers and Teaching Research, UCL Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	127
Potential explanations of why minority teachers leave	128
School demographic characteristics	128
Teacher characteristics	128
School working conditions	129
Racial climate in schools	129
A case study of teacher diversity and retention in English schools	129
Where minority ethnic teachers work	130
What factors matter for staying	131
Conclusion	132
Acknowledgment	132
References	132
Relevant websites	133

Glossary

BAME This is a UK term for non-White British population—Black, Asian and Minority Ethnic

Teacher retention This term refers in different contexts to both teachers staying in the profession and leaving the school sector

Teacher turnover This term refers to teachers moving between schools

Teacher attrition This term often refers to teachers leaving the profession all together

Introduction

This article provides a review of research on the topic of minority ethnic teacher retention as well as an overview of the current research on the state of teacher diversity and minority ethnic teacher retention in England.

There are many good reasons for having a diverse teaching workforce. Available research has demonstrated that teachers from minority ethnic groups are in a position to providing culturally sensitive teaching, understand, communicate, or identify with students from nondominant cultural and linguistic backgrounds the way nonminority educators are often not able to do (see the case for teacher diversity in [Albert Shanker Institute, 2015](#)). International evidence suggests that exposure to teachers from a similar background has potential to help disadvantaged students from minority groups in danger of school exclusion and drop-out attain greater educational success. [Gershenson et al. \(2017\)](#) evidence from multiple settings in the US has demonstrated that a disadvantaged black male's exposure to at least one black teacher in elementary school reduces his probability of dropping out of high school by 40%. There is also research that shows that minority teachers use their cultural knowledge and own experiences as "other" to develop deeper understandings of minority students and their needs ([Gershenson et al., 2016](#); [Santoro, 2015](#)). Students of all, including nonminority, backgrounds benefit from diverse educational experiences.

Teachers from minority ethnic groups continue to represent a small portion of the teaching workforce in the increasingly diverse Anglophone societies such as USA ([Ingersoll and May, 2011](#)), UK ([Haque, 2017](#); [Menter et al., 2006](#)), Canada ([Dandala, 2020](#)), Australia ([Santoro, 2015](#)) and New Zealand ([Howard, 2009](#)). And despite an increase in the proportion of minority teachers over the last decade, a gap persists between the percentage of minority ethnic students and the percentage of minority ethnic teachers, especially in some ethnic sub-groups. The governments have sought to address this student-teacher imbalance through the agenda to boost numbers of minority teachers entering the profession. However, key authors in the field argue that the recruitment of new teachers alone will not solve shortages ([Ingersoll and May, 2011](#)). Scholars and commentators often employ "a revolving door" or "a leaky pipeline" metaphor to emphasize the problem of high turnover and attrition of new teachers once they join the profession. An approach focused on understanding the causes of minority teacher staffing problems in relation to the institutional characteristics and the culture of the workplaces where teachers are employed has been useful to study the phenomenon ([Ingersoll and May, 2011](#)). This article adopts this lens.

International research reviewed in the following section suggests that minority teacher retention and turnover have important implications for social justice in education. These teachers tend to work in urban schools with high-minority and more disadvantaged intakes, which also experience higher teacher turnover rates. Retaining minority teachers is thus crucial for managing teacher

supply in these schools to reduce the negative impact of turnover on the outcomes of disadvantaged children. For example, in England, London experiences acute teacher retention challenges (Worth et al., 2018), while minority ethnic teachers are more open and motivated than White British teachers to remain in teaching in London. There is also some evidence from the US that a diverse teaching staff in high-poverty, high-minority schools promotes retention of teachers regardless of their race (Djonko-Moore, 2016). This suggests that school staff diversity is also an important workplace issue.

Potential explanations of why minority teachers leave

There have been various explanations identified in the literature for the retention and turnover among teachers from minority ethnic groups in the Anglophone countries. The key findings are discussed in turn, although some of them are overlapping and confounding. It is notable that most research in this field focuses on the United States context.

School demographic characteristics

A substantial body of research reveals that a high proportion of minority ethnic teachers work in disadvantaged schools, which predominantly fall into the category of “hard-to-staff” schools. Allen et al. (2018) study of teacher turnover in England found that deprived schools hired younger and more ethnically diverse teachers than affluent schools did. Ingersoll et al. (2019) found that between 1987 and 2012 two-thirds of minority teachers employed in the US public sector worked in high-minority, high-poverty, and urban schools. Moreover, minority teachers were two to three times more likely than White teachers to work in these schools. A US study focused on Black teachers found that while a quarter of all teachers in 2011–2012 taught in schools that were both high-poverty and high-minority schools, about 60% of Black teachers and 62% of Black female teachers (Carver-Thomas and Darling-Hammond, 2017). This study and other research conducted in Canada (Ryan et al., 2009) have also pointed out that teachers of color were more likely than other teachers to work in schools located in large cities.

These patterns of work in schools that face additional challenges could be an explanation for higher rates of minority teacher turnover. Yet, a review of the 70 studies conducted in the US found that “teachers of color are less likely than White colleagues to leave schools serving high proportions of students from low-income and racially and culturally nondominant communities” (Achinstein et al., 2010, p. 90). Looking at the relationship between school characteristics and turnover, Ingersoll et al. (2019) found that minority teachers were slightly less likely than nonminority teachers to move from schools with higher poverty-level enrollments, as well as from larger schools. When they did move from one school to another, six times as many minority teachers did so from high-poverty schools to other high-poverty schools than from high-poverty schools to low-poverty schools, and almost three times as many moved from urban schools to other urban schools than from urban to suburban schools (Ingersoll and May, 2011).

On the contrary, in England minority ethnic and younger teachers had a higher turnover rate in schools situated in deprived neighborhoods than other teachers (Allen et al., 2018). However, while early career retention rates were found to be significantly lower for minority ethnic teachers across all training routes in England (Allen et al., 2016), newly qualified teachers from minority ethnic groups had a better retention rate in urban and diverse Inner and Outer London than White teachers (DfE, 2018). Personal background of minority teachers thus appear to be involved in turnover and retention.

Teacher characteristics

Some research identifies notable trends in minority ethnic teacher retention and turnover by the demographic characteristics of teachers. In Ingersoll et al. (2019) analysis of large datasets gender appears to be a salient predictor in the US schools, with the odds of male minority teachers departing over 50% higher than for female minority teachers. In the same study the age of teachers was a predictor of the likelihood of turnover, but only for nonminority teachers. They found that younger (less than 30) and older (greater than 50) minority ethnic teachers did not depart at higher rates than middle-aged (31–50) minority teachers. However, overall, the data both in the US and UK show that, over the past decades, minority teachers were more likely to depart from their schools than nonminority teachers, with younger teachers more affected (Achinstein et al., 2010). There is also evidence from the US that younger teachers from minority ethnic groups were more likely to be certified via alternative pathways than other teachers (Carver-Thomas and Darling-Hammond, 2017; Redding and Smith, 2016).

A new area of research, relational demography, examines the impact of the race/ethnicity congruence between teachers and school principals on minority ethnic teacher retention and turnover. Nguyen et al. (2020) recent meta-analysis found that the odds of teacher turnover may be smaller when there is congruence relative to incongruence. Grissom and Keiser (2011) have demonstrated, using US nationally representative data, that teachers who are race-congruent with their principal are less likely to leave than non-congruent colleagues in the same school. These findings held true for Black teachers in a subsequent longitudinal analysis of personnel records from both the States of Missouri and Tennessee (Bartanen and Grissom, 2019). It was found that Black principals decreased Black teacher mobility, significantly reducing the probability that a Black teacher changes schools. They also found that teachers gave higher ratings to school leadership, school climate, and their own job satisfaction in years in which they had a race-congruent principal. The next two sections delve further into what might drive these findings.

School working conditions

While the demographic characteristics of schools appear to be important to minority ethnic teachers' initial employment decisions, their later decisions about whether to depart are tied to working conditions in their schools (Achinstein et al., 2010). There is no evidence to suggest that these teachers leave because of dissatisfaction with disadvantaged students. Ingersoll et al. (2019) analysis of the US national teacher survey shows that particular school working and organizational conditions are strongly related to minority teacher departures. The data indicate that minority teachers departed at higher rates because the schools in which they were employed tended to have less positive organizational conditions, including salary, student discipline problems, school leadership support, school resources, faculty influence, and teacher autonomy. For minority ethnic teachers, predicted annual turnover rates were only 12% in the schools with the best organizational conditions versus nearly 21% in schools with the worst organizational conditions (compared to 12% in the best schools to 15% in the worst schools for nonminority teachers). The strongest organizational factors for minority ethnic teachers were the levels of collective faculty decision-making influence in their school and the degree of individual classroom autonomy. For example, a one-unit increase in reported faculty influence between schools (on a four-unit scale) was associated with a 37% decrease in the odds of a minority teacher departing, and a one-unit difference in reported teacher classroom autonomy was associated with a 40% difference in the odds of a minority teacher departing. The authors make a powerful observation based on their simulation analysis that had the schools afforded minority teachers the classroom autonomy held by teachers employed in schools that were in the top 10th percentile of teacher classroom autonomy, it is conceivable that the US would have had nearly a quarter million more minority teachers within a decade.

Racial climate in schools

Many studies that draw on the accounts of teachers conclude that the racism teachers from minority groups are exposed to in schools take a toll on their well-being, progression, and retention. The body of research on minority ethnic pre-service teachers in the UK showed that the participants were negatively affected by placement in predominantly white schools (Basit et al., 2006; Wilkins and Lall, 2011; Bhopal, 2015; Maylor, 2018). Although statistical studies such as by Ingersoll et al. (2019) did not find that minority ethnic teachers departed schools with different proportions of minority faculty at different rates, Bristol (2018) summary of the survey findings across an urban district in the US suggested that racially isolated school spaces promoted Black teachers to leave. Those who were the only Black male teachers on the faculty were more likely than those who were among many more Black teachers to cite tense relationships with colleagues and administrators as their two primary reasons for wanting to leave. They were also more likely to believe that being Black caused people to fear them in their schools.

It is important to note that diverse schools in urban areas have also been perceived as racially hostile, stressful and disrespectful workplaces for teachers from minority ethnic backgrounds. Even in schools serving multicultural communities teachers from minority ethnic groups work within a predominantly White profession leaving many feeling like outsiders. Kohli (2018) research with 218 urban American teachers of color identified two typical forms of racism that affected their professional experiences and retention: (i) color blindness and (ii) racial microaggressions. Schools that adhere to the color blind racist standpoint claim race neutrality, leaving teachers of color isolated and devalued in their efforts to raise equity concerns about the attainment gap between White and other students and to advocate for culturally relevant solutions to various educational issues. The instances of microaggressions refer to the covert forms of systemic racism in schools, including calling teachers by the wrong surname, exposing assumptions based on cultural stereotypes and lack of knowledge about different communities, questioning the relevance of culturally sensitive pedagogy in the climate of accountability, and the like. Microaggressions directed at minority ethnic students are also deeply offensive to these teachers. The findings with respect to invisibility, exclusion and professional "othering" within the "mainstream" teaching community are echoed in research with indigenous and minority ethnic teachers in Australia (Burgess, 2017; San-toro, 2015).

Recent studies have demonstrated that teachers end up leaving schools serving urban students from nondominant cultural communities because of the cumulative impact of "racial battle fatigue", defined by Smith (2009) as "the psychological, emotional, physiological, energy, and time-related cost of fighting against racism" (cited in Pizarro and Kohli, 2020, p. 969). Among these studies is that of Hancock et al. (2020) which explore how the pressure of having to "contend with the mental and emotional stress of navigating whiteness, white fragility as it relates to curriculum content, and constant disrespect for nonwhite cultural experiences and ways of knowing" impacted on one Black teacher's decision to take a prolonged career break (p. 5). Bristol and Goings (2019) interview research demonstrates that having to strategically navigate racial boundaries in the workplace expends considerable energy on the part of Black male teachers. Over time, the battle for recognition and promotion within the racialized power structures of the schools leaves teachers "burned" (Wallace, 2020). These qualitative findings add important explanatory power to statistical analyses of retention factors about what specifically in the school climate encourages minority ethnic teachers to explore opportunities in other schools or leave the profession.

A case study of teacher diversity and retention in English schools

The issues related to minority teacher retention are nuanced by national and local contexts. In this section we explore the situation in England, drawing on our own analyses of systemic data and interviews with minority ethnic teachers. This exploration

demonstrates both the commonalities of employment patterns across contexts, and the specifics of the English context, including regional variations between London and the rest of the country. We also note some of the key factors affecting the retention of both early career and experienced teachers. These include the diversity of the school, relationships with the school leadership, and promotional opportunities.

As is the case for other countries, reviews of the composition of the teacher workforce in England have shown a persistent shortage of Black, Asian and other minority ethnic (BAME) teachers in relation to the minority ethnic student population (Steel, 2015; Haque, 2017; Haque and Elliott, 2017; Joseph-Salisbury, 2020). For example, in 2018, the proportion of minority students in primary education was 33.1% and in secondary 30.3%, whereas the proportion of teachers from ethnic minority origins was estimated to be 11.3% and 17.4% respectively (Gov.UK, 2020a, b). Table 1 shows that this overall student–teacher gap in state funded schools also exists for minority subgroups, as well as for the proportion of the corresponding working age population (with the exception of the White Irish group). Minority ethnic pupils therefore do not see themselves represented in the teaching workforce. Some disadvantaged students such as from Traveler, Gypsy and Roma backgrounds do not meet any teachers from their ethnic group, despite making up 7% of the pupils (Gov.UK, 2020b).

Beyond policy pronouncements about training more minority ethnic teachers to address shortages reflected in the headline figures (Whittaker, 2019), there has been little empirical investigation of how the racial/ethnic character of the English teaching force has changed over time, where BAME teachers tend to be employed, and what happens to these teachers once they are employed. These are important questions from the organizational point of view of teacher retention (Ingersoll and May, 2011) given the existence of evidence that nationally retention is lower for minority ethnic teachers (Allen et al., 2016; DfE, 2018) and that teachers from different minority groups experience unique problems in their careers (Haque and Elliott, 2017).

Guided by these concerns, our research (Tereshchenko et al., 2020) used the recent 2018 School Workforce Census and related administrative school census datasets to model the distribution of teachers from minority ethnic groups across English schools. To our knowledge, this is the first analysis of this kind, including a total of 21,746 schools. As part of this research we also conducted 24 narrative interviews with teachers from different ethnic and professional backgrounds to explore factors shaping their retention. At risk of overgeneralizing, the findings summarized below relate to non-White British teachers as a whole.

Where minority ethnic teachers work

The employment of minority ethnic teachers is not uniform across England. Our analysis showed that around 46% of schools in England had no teachers from minority ethnic groups. Only 16% of schools employed more than one in five teachers from minority ethnic backgrounds. The most influential characteristic of schools with a large proportion of minority ethnic teachers was the location of the school. On average around 35% of teachers in Outer London schools and over 40% of teachers in Inner London schools were from non-White British background, a far greater concentration than any other English region. The data indicate that 96% of London schools employed at least one non-White British teacher, twice the rate of schools outside of London.

Our linear regression model identified the factors associated with the likelihood of a school employing a large proportion of BAME teachers in schools across England, and, to account for a strong London effect, separately in London and non-London schools. The featured factors are listed in Table 2 in order of magnitude of association. The key finding is that the presence of minority ethnic teachers occurs in secondary schools with strong minority representation among other staff, including teaching assistants, non-classroom support and auxiliary staff. The ethnic background of staff also appeared to mirror that of students, as minority ethnic teachers were more likely to be found in schools with a high proportion of students from non-dominant racial and cultural backgrounds using English as an additional language (EAL). Notably, this pattern persists nationally, in London and outside of the capital city.

In addition to the strong associations related to diversity, we found that, nationally and outside of London, minority ethnic teachers were more likely to be working in schools run by the local authorities (LA) compared to academies. Within London, we also found some evidence that schools with a high proportion of minority ethnic teachers were more likely to be located in the areas of higher income deprivation affecting children (IDACI). These schools were rated lower for effectiveness of leadership and management by Ofsted (the chief inspection and regulation department responsible for inspecting educational institutions in England). We found some additional evidence which suggests that, in England, schools rated lower for overall effectiveness by Ofsted were associated with a larger proportion of BAME teachers. This is of concern given the importance of school context in teacher retention.

Table 1 Percentage of teachers, students and the working age population, by ethnicity.

	White British total	Minority ethnic groups						
		White Irish	White other	Black	Asian	Chinese	Mixed	Other
Students	65.4	0.3	6.7	5.7	11.4	0.5	6.1	2.0
Teachers	85.9	1.6	3.9	2.2	4.4	0.2	1.3	0.6
Working age population	78.5	1.0	5.7	3.6	7.5	0.9	1.8	1.1

Sources: Gov.UK (2020a,b).

Table 2 Descriptors of schools closely associated with a large proportion of BAME teachers.

England overall	London schools	Non London schools
1. High proportions of BAME teaching assistants	1. High proportions of BAME teaching assistants	1. High proportions of BAME teaching assistants
2. London location	2. High proportions of EAL pupils	2. High proportions of BAME non-classroom based school support staff
3. High proportions of BAME non-classroom based school support staff	3. High proportions of teaching staff who work part-time	3. Secondary schools
4. Secondary schools	4. High proportions of BAME non-classroom based school support staff	4. High proportions of BAME auxiliary staff
5. High proportions of EAL pupils	5. Secondary schools	5. High proportions of EAL pupils
6. High proportions of BAME auxiliary staff	6. Effectiveness of leadership and management is lower	6. Higher number of teachers in the leadership group (FTE)
7. LA maintained schools	7. Higher IDACI quintile	7. Free schools
8. Overall school effectiveness is lower	8. Higher proportions of full-time temporarily filled posts	8. LA maintained schools

What factors matter for staying

Focusing on the prominent diversity predictors cited above, teachers from minority ethnic groups tend to reside in high diversity urban areas in England. This may partly explain why these teachers gravitate toward diverse schools in their communities. However, there is evidence from a relatively dated teacher survey that minority ethnic staff in the UK express a *preference* for culturally diverse schools (Davidson et al., 2005). Our interviews also clearly indicated that minority ethnic teachers purposefully sought out diverse schools and enjoyed working there. A recently qualified Black male teacher told us: “I’ve basically said to my fiancé, I’m probably going to move out of London but I want to work in London, because of the demographics”. This comment refers to the student demographics and supports the literature around humanistic inclinations of minority teachers to make a difference in the lives of young people from low-income and culturally and linguistically non-dominant communities (Achinstein et al., 2010). The importance of staff diversity, alongside that of students, was also extremely important for many of the teachers. They linked more diverse school contexts to reduced “racial stress” in their professional lives with implications for retention (see Hancock et al., 2020). Of course, given that only 16% of schools in England employ a fifth or more of their teachers from minority ethnic groups, in the same urban schools where BAME teachers are disproportionately employed, the teaching staff overall is nevertheless predominantly White British. Yet, in the words of one newly qualified primary teacher, having even a small support network of colleagues “that look like me, having the same experiences as me” was key for her decision to move to a London school. The following statements from the interviews are representative of the other reflections on the importance of diversity:

I was at school today actually for my first day and you could just see everyone and they’re so friendly and it’s like majority Asian. There’s quite a few black people and it’s like “this is nice”.

Charlotte, Black African heritage.

They were actively saying, “We want our teachers to be reflective of our students.” The deputy head was, I think, Indian. The head teacher, I think she was from Eastern Europe, so it was very mixed. Actually I did enjoy my time at that school.

Olivia, Black Caribbean heritage.

So for me, when choosing a school, yes, it does matter what their culture is and how they view ethnic minorities and if you walk around on interview and don’t see that reflected in students or staff then I probably will be less likely to go and work there.

Farid, Asian heritage.

In alignment with a body of work examining the impact of organizational characteristics of schools on teacher retention, we found that although school demographic characteristics mattered to minority teachers, this was not enough to keep them in schools long-term. Relationships with the leadership team emerged in our study as being the main retention factor for these teachers, with all of our respondents referring to this element, regardless of the number of years worked and type of school worked in. Teachers wanted senior leaders to appreciate, respect, value and support them. This worked particularly well when school leaders had developed their racial literacy to enable a dialogue about ‘race’/ethnicity and equity in school, to support culturally relevant teaching, and to end stereotypical approaches to minority ethnic staff development (such as direction into pastoral roles). These interpersonal and professional qualities were critical for teachers perceiving a comfortable social environment to work in, having possibilities for establishing dialog with school leadership, resolving conflicts with colleagues, having the “freedom” to teach, their sense of self-esteem and even their salary.

Our interviews suggested that as their tenure progressed, teachers developed a particularly keen eye for discerning the place of white privilege in school power hierarchies. In this respect, we observed clear differences between respondents who were early career teachers and those who were experienced teachers. Those who had been in profession for over five years highlighted their concerns about barriers to getting rewards and promotions, making a lack of advancement the key retention factor for experienced teachers from minority ethnic groups. One Black interviewee linked disciplinary roles assigned to minority (and male) staff to their poor promotion record (see also [Hargreaves, 2011](#); [Haque, 2017](#); [Wallace, 2020](#)): “They’re constantly being given those same classes that nobody else wants to deal with but they’re not being given any sort of opportunity for progression to move up, because dealing with those kind of classes, it’s going to be harder for you to be a literacy coordinator or to be a numeracy coordinator or for you to be an arts coordinator, science coordinator, phase leaders, etc.” Notably in our research, teachers who had been in the school system for a number of years were successful in obtaining middle leadership roles. For some it took less than five years to secure mid-level leadership appointments such as heads of the year group and pastoral units, subject leaders, curriculum coordinators, but they then often felt stuck in those posts for years.

Many of our interviewees were firm in the belief that their racialized status prevented them from further career progression into what they perceived as all white, and predominantly male, senior leadership circles. This finding reinforces the claim by [Wallace \(2020\)](#) about “a diversity trap” existing for Black teachers in the UK whereby their inclusion “in entry-level leadership tracks may change perceptions of the power of Whiteness in schools and to wider publics, but does not change the institutional arrangements that preserve White privilege” (p. 359). Feeling unfulfilled and/or undermined in their effort to advance within schools, some experienced teachers we interviewed indicated, for example, that they had considered or were considering retraining as school inspectors, becoming a supply teacher or an educational consultant, teaching abroad, or doing a doctorate. The concerns expressed by our interviewees are reflected in the literature. An existing body of research has identified the underrepresentation of educators from minority ethnic groups in school leadership positions in England (see [Elonga Mboyo, 2019](#)). The emerging evidence from the US suggests that the presence of Black leaders in school decrease Black teacher turnover ([Bartanen and Grissom, 2019](#)). It would appear that the challenges of promotion and progression need to be placed on the policy agenda to mitigate the turnover and attrition of experienced and indeed other teachers from minority ethnic groups.

Conclusion

The overwhelming conclusion of research on the supply of teachers is that problems of minority ethnic teacher shortages are attributable to poor retention. Results of reviewed studies show the existence of several salient commonalities among factors affecting minority teacher retention in different countries. These commonalities include the tendency to work in urban, diverse and disadvantaged schools with additional pressures and unfavourable organizational conditions while having to navigate racism within the predominantly White teaching profession. Our case study conducted in England provided additional empirical evidence of the importance of diverse and inclusive workplaces for the retention of teachers from minority groups with the particular emphasis on staff diversity and equal representation on all levels of the school hierarchy.

Looking ahead, most literature on the retention of minority ethnic teachers focuses on the United States context. More research is required on the retention of minority ethnic teachers in England and other increasingly diverse countries in Europe. The lack of empirical studies using nationally representative data of what happens to teachers from nondominant ethnic group once they are employed and what drives their turnover makes it difficult to formulate policies around staffing urban disadvantaged schools with high quality teachers and retaining them. Such analysis could also help explore and potentially raise equity concerns with respect to the relegation of these teachers to schools with lower levels of various capitals. The trends must be examined within specific ethnic subgroups to account for different histories of disadvantage. Additionally, while common retention patterns across schools and across minority ethnic subgroups can be informative, there remains a gap in qualitative research on successful retention practices of minority ethnic teachers in specific schools. Much could be learned from documenting and disseminating successful and potentially transferable practices.

Acknowledgment

The research toward this publication was supported by the Department for Business, Energy and Industrial Strategy through a British Academy/Leverhulme Small Research Grant Program.

References

- Achinstein, B., Ogawa, R., Sexton, D., Freitas, C., 2010. Retaining teachers of color: a pressing problem and a potential strategy for “hard-to-staff” schools. *Rev. Educ. Res.* 80 (1), 71–107.
- Albert Shanker Institute, 2015. *The State of Teacher Diversity in American Education*. Albert Shanker Institute, Washington, DC.
- Allen, R., Bibby, D., Parameshwaran, M., Nye, P., 2016. *Linking ITT and Workforce Data: Initial Teacher Training Performance Profiles and School Workforce Census*. NCTL.
- Allen, R., Burgess, S., Mayo, J., 2018. The teacher labor market, teacher turnover and disadvantaged schools: new evidence for England. *Educ. Econ.* 26 (1), 4–23.

- Bartanen, B., Grissom, J., 2019. School Principal Race and the Hiring and Retention of Racially Diverse Teachers (EdWorkingPaper No.19-59). Annenberg Institute at Brown University. Retrieved from: <http://edworkingpapers.com/ai19-59>.
- Basit, T.N., Roberts, L., McNamara, O., Carrington, B., Maguire, M., Woodrow, D., 2006. Did they jump or were they pushed? Reasons why minority ethnic trainees withdraw from initial teacher training courses. *Br. Educ. Res. J.* 32 (3), 387–410.
- Bhopal, K., 2015. Race, identity and support in initial teacher training. *Br. J. Educ. Stud.* 63 (2), 197–211.
- Bristol, T., Goings, R., 2019. Exploring the boundary-heightening experiences of black male teachers: lessons for teacher education programs. *J. Teach. Educ.* 70 (1), 51–64.
- Bristol, T.J., 2018. To be alone or in a group: an exploration into how the school-based experiences differ for black male teachers across one urban school district. *Urban Educ.* 53 (3), 334–354.
- Burgess, C., 2017. Having to say everyday ... I'm not black enough ... I'm not white enough. Discourses of Aboriginality in the Australian Education Context *Race Ethn. Educ.* 20 (6), 737–751.
- Carver-Thomas, D., Darling-Hammond, L., 2017. Why black women teachers leave and what can be done about it. In: Farinde-Wu, A., Ayana Allen-Handy, A., Lewis, C.W. (Eds.), *Black Female Teachers, Advances in Race and Ethnicity in Education*, vol. 6. Emerald Publishing Limited, pp. 159–184.
- Dandala, S., 2020. Public accountability and workforce diversity in Canadian public education sector. *Educ. Rev.* 72 (5), 650–662.
- Davidson, J., Powney, J., Wilson, V., Hall, S., Mirza, H.S., 2005. Race and sex: teachers' views on who gets ahead in schools? *Eur. J. Teach. Educ.* 28 (3), 311–326.
- DfE, 2018. Analysis of Teacher Supply, Retention and Mobility. Teachers Analysis Compendium 4. Available at: <https://www.gov.uk/government/statistics/teachers-analysis-compendium-4>.
- Djonko-Moore, C.M., 2016. An exploration of teacher attrition and mobility in high poverty racially segregated schools. *Race Ethn. Educ.* 19 (5), 1063–1087.
- Elonga Mboyo, J.P., 2019. School leadership and Black and minority ethnic career prospects in England: the choice between being a group prototype or deviant headteacher. *Educ. Manag. Adm. Leader* 47 (1), 110–128.
- Gershenson, S., Holt, S.B., Papageorge, N.W., 2016. Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Econ. Educ. Rev.* 52, 209–224.
- Gershenson, S., Hart, C.M.D., Lindsay, C.A., Papageorge, N.W., 2017. The Long-Run Impacts of Same-Race Teachers. IZA – Institute of Labor Economics, Bonn.
- Gov.UK, 2020a. Ethnicity Facts and Figures: School Teacher Workforce. Available at: <https://www.ethnicity-facts-figures.service.gov.uk/workforce-and-business/workforce-diversity/school-teacher-workforce/latest>.
- Gov.UK, 2020b. Academic Year 2019/20: Schools, Pupils and Their Characteristics. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics>.
- Grissom, J.A., Keiser, L.R., 2011. A supervisor like me: race, representation, and the satisfaction and turnover decisions of public sector employees. *J. Pol. Anal. Manag.* 30 (3), 557–580.
- Hancock, S.D., Showunmi, V., Lewis, C., 2020. Teaching while black and female: navigating mental, curricular, and professional aggression. *Theory Pract.* <https://doi.org/10.1080/00405841.2020.1773184>.
- Haque, Z., Elliott, S., 2017. Visible and Invisible Barriers: The Impact of Racism on BME Teachers. The Runnymede Trust and NUT. Available at: <https://neu.org.uk/media/2936/view>. (Accessed 7 August 2020).
- Haque, Z., 2017. Visible Minorities, Invisible Teachers: BME Teachers in the Education System in England. Runnymede Trust and NASUWT. Available at: <https://www.runnymedetrust.org/uploads/Runnymede%20ReportNEW.pdf>. (Accessed 7 August 2020).
- Hargreaves, L.M., 2011. The status of minority ethnic teachers in England: institutional racism in the staffroom. *DEDICA. J. Educ. Hum.* 1, 37–52.
- Howard, J., 2009. Pride and prejudice...and other barriers to teaching for Asian New Zealanders, Asia Pacific. *J. Educ.* 29 (2), 143–157.
- Ingersoll, R.M., May, H., 2011. Recruitment, Retention and the Minority Teacher Shortage. Consortium for Policy Research in Education. CPRE Research Report #RR-69.
- Ingersoll, R., May, H., Collins, G., 2019. Recruitment, employment, retention and the minority teacher shortage. *Educ. Pol. Anal. Arch.* 27 (37). <https://doi.org/10.14507/epaa.27.3714>.
- Joseph-Salisbury, R., 2020. Race and Racism in English Secondary Schools. Runnymede Trust.
- Kohli, R., 2018. Behind school doors: the impact of hostile racial climates on urban teachers of color. *Urban Educ.* 53 (3), 307–333.
- Maylor, U., 2018. Black male student teachers: tomorrow's teachers? *J. Multicult. Educ.* 12 (2), 111–125.
- Menter, I., Hartshorn, B., Hextall, I., Howell, I., Smyth, G., 2006. Widening access to the teaching profession: perspectives from Scotland. *Eur. J. Teach. Educ.* 29 (3), 271–285.
- Nguyen, T.D., Pham, L.D., Crouch, M., Springer, M.G., 2020. The correlates of teacher turnover: an updated and expanded meta-analysis of the literature. *Educ. Res. Rev.* 31. <https://doi.org/10.1016/j.edurev.2020.100355>.
- Pizarro, M., Kohli, R., 2020. "I stopped sleeping": teachers of color and the impact of racial battle fatigue. *Urban Educ.* 55 (7), 967–991.
- Redding, C., Smith, T., 2016. Easy in, easy out: are alternatively certified teachers turning over at increased rates? *Am. Educ. Res. J.* 53 (4), 1086–1125.
- Ryan, J., Pollock, K., Antonelli, F., 2009. Teacher diversity in Canada: leaky pipelines, bottlenecks, and glass ceilings. *Can. J. Educ.* 32 (3), 591–617.
- Santoro, N., 2015. The drive to diversify the teaching profession: narrow assumptions, hidden complexities. *Race Ethn. Educ.* 18 (6), 858–876.
- Smith, W.A., 2009. Campuswide climate: implications for African American students. In: Tillman, L.C. (Ed.), *The SAGE Handbook of African American Education*. Sage, Los Angeles, CA, pp. 297–309.
- Steel, S., 2015. Race to the Top 2: Diversity in Education. Elevation Networks.
- Tereschenko, A., Mills, M., Bradbury, A., 2020. Making progress? Employment and Retention of Minority Ethnic Teachers in England. UCL Institute of Education, London.
- Wallace, D., 2020. The diversity trap? Critical explorations of black male teachers' negotiations of leadership and learning in London state schools. *Race Ethn. Educ.* 23 (3), 345–366.
- Whittaker, F., 2019. Teacher Shortage Won't Be Addressed Without More BAME Teachers, Warns Hinds. School Weeks. Available at: <https://schoolsweek.co.uk/teacher-shortage-wont-be-addressed-without-more-bame-teachers-warns-hinds/>.
- Wilkins, C., Lall, R., 2011. You've got to be tough and I'm trying: black and minority ethnic student teachers' experiences of initial teacher education. *Race Ethn. Educ.* 14 (3), 365–386.
- Worth, J., Lynch, S., Hillary, J., Rennie, C., Andrade, J., 2018. Teacher Workforce Dynamics in England. NFER, Slough.

Relevant websites

- Online Resource by the Education Trust on African American Teachers "Through Our Eyes: Perspectives and Reflections From Black Teachers": <https://edtrust.org/resource/eyes-perspectives-reflections-black-teachers/>.
- Online Resource by the Education Trust on Latino Teachers "Our Stories, Our Struggles, Our Strengths": <https://edtrust.org/resource/our-stories-our-struggles-our-strengths/>.

Early career teachers' work

Anna M. Sullivan, Barry Down, and Bruce Johnson, University of South Australia, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	134
What is the early career teacher induction phase and why does it matter?	134
Who are early career teachers and what is their employment status?	135
What is happening to early career teachers and their work?	135
What research has been conducted on early career teachers and the induction phase?	136
How are early career teachers currently supported, or not?	137
Aims of induction programs	137
Common strategies and components used in induction programs	137
The role of school leaders in induction programs	138
The influence of school culture on induction programs	138
Mentoring	138
How might early career teachers' work be reimaged?	139
Acknowledgments	140
References	140

Introduction

Internationally, early career teachers have increasingly been framed as a “problem” that limits the provision of high quality education in schools (Sullivan et al., 2019b). Concerns have centered on how they can transition to the profession to assume full teaching roles just like more experienced teachers (Down and Sullivan, 2019). Extensive research has examined various induction practices, including mentoring, coaching, reduced workloads, and professional development, aimed at supporting early career teachers to successfully make this transition (see the review by Long et al., 2012). However, it is clearly understood that policy initiatives aimed at supporting early career teachers to transition successfully to the profession have largely been based on deficit views of them (Kelchtermans, 2019).

This chapter examines the international literature on teachers' work, particularly what we know about their work during the early career induction phase. Understanding the context and work of early career teachers is important especially because too few people are entering the profession and, worryingly, far too many teachers are leaving early (Sullivan et al., 2019a) especially during the current COVID-19 pandemic (Chrisafis, 2022; Rivers, 2022; Sainato, 2021; Shearing, 2021).

In this context, the chapter considers what can be done to reframe the work of early career teachers based on an alternative democratic imaginary of teachers' work beyond the limitations of neoliberal policy regimes and the constraints of performative cultures (Angus, 2012). We argue that early career teachers' work needs to be radically reconceptualized in ways that reflect the complexity, fragility, uncertainty, and relational nature of teaching (Connell, 1985; Lortie, 2002 [1975]). This involves an activist professionalism whereby early career teachers are viewed more as active agents or knowledge producers than mere deliverers and technicians (Ball, 2016; Sachs, 2003). This is a more participatory, collegial, and reflective response to the needs of early career teachers based on the values of trust, respect, and care. We deploy the notion of a critical learning community as a means of fostering this kind of vision whereby early career teachers can engage with the broader philosophical, democratic, and progressive purposes of education (Ayers, 2004; Dewey, 1944 [1916]; Thomas and Schubert, 2001). In this way early career teachers can engage with colleagues around contemporary issues of immediate relevance and worth to them. In these spaces early career teachers can begin to produce new knowledge about curriculum, pedagogy, and assessment practices. We conclude that there are unexpected and paradoxical consequences of trying to reframe early career teachers' work in the context of increased performativity and work intensification.

This chapter is organized around six key questions:

- What is the early career induction phase and why does it matter?
- Who are early career teachers and what is their employment status?
- What is happening to early career teachers and their work?
- What research has been conducted on early career teachers and the induction phase?
- How are early career teachers being supported, or not?
- How might early career teachers' work be reimaged?

What is the early career teacher induction phase and why does it matter?

The first few years of a teacher's career as they transition from pre-service teacher education into the profession is known as the “induction phase” (Kelchtermans, 2019). This phase is characterized by “hopefulness, energy and enthusiasm, a high degree of

motivation and commitment, as well as some new approaches to pedagogy and classroom management" (Johnson et al., 2016, p. 38). Research has identified the conditions under which early career teachers can thrive during this phase. For example, we know that consideration needs to be given to the broader policy environment, the school culture, the relationships that are formed, the work that they are expected to undertake, and the support given to teachers (Johnson et al., 2015, 2016). However, it is also known as a time when many new teachers experience stress and praxis shock, leading to teacher burnout (Clandinin et al., 2015; Lambert and Gray, 2021; Veenman, 1984). In most countries, early career teachers face unrealistic expectations that they will perform like more experienced teachers from the first day of employment (e.g., Johnson et al., 2016); their work is being increasingly de-professionalized (Down and Sullivan, 2019); and their employment conditions are declining (Preston, 2019; Sutherland et al., 2016). Of particular concern is the escalating levels of casualization and contractualization, which serve to undermine the notion of relational trust and a sense of belonging (Smyth, 2012). Research has examined the experiences of new teachers and identified that they are subjected to increasing pressures due to work intensification and performativity (e.g., Flores, 2019; Gewirtz et al., 2009; Johnson et al., 2016). More recently, research has identified that teachers experience high levels of scrutiny and surveillance, which leads to uncertainty about professional performance and a lack of agency (Perryman et al., 2017; Sullivan et al., 2021).

The threats to teachers' wellbeing and professional judgment have been well researched and have led to a range of attempts to support teachers during the induction phase. Kelchtermans argues that teacher induction should provide "purposeful support and guidance" to teachers as they learn the particular requirements of a school and facilitate student learning (2017) and, more broadly, that teacher induction is "the socialization of new members of the teaching profession in their job" (2019, p. 83). Kearney (2021) argues that "Induction in recent years has come to be an umbrella term for orientation, mentoring and support of teachers, both new to the profession and/or new schools (Wong, 2004) and has been linked to professional accreditation/registration" (p. 143).

Early career teachers are critical to the future of education and the teaching profession; therefore, there is an obligation to provide the conditions necessary to support them in times of uncertainty. If we wish to prepare children extraordinarily well for success in family, careers, and life, then the induction of early career teachers is essential. The importance of early career teachers is heightened by serious concerns about an international current and impending teacher supply crisis compounded by the COVID-19 pandemic (McGowan and Rose, 2022; Sullivan et al., 2019a; Wilson and Carabetta, 2022). Thus, understanding the ways in which new teachers can be supported during their induction phase is critical to their success.

Who are early career teachers and what is their employment status?

In order to understand what is happening to teachers and their work it is important to consider who early career teachers are and what their employment status is. The teaching profession continues to be composed mainly of women and it is widely recognized as a feminized profession (Moreau, 2019). Research shows that internationally men are less likely than women to consider teaching as a career pathway (Han et al., 2020). Moreau (2019) explains that the "overall proportion of women in teaching fluctuates considerably across segments of the teaching labor market with, however, women consistently concentrating in the less prestigious and (financially) less rewarding segments" (p. 9). This is illustrated in Australia, for example, where 96% of early childhood pre-service teachers are women, whereas 57% of secondary pre-service teachers are women, reflecting differences in status and pay opportunities (Australian Institute for Teaching and School Leadership, 2020). Furthermore, early career teachers are most likely to be women in their 20s (Australian Institute for Teaching and School Leadership, 2020).

The nature of the employment status of new teachers is changing. Firstly, teaching is becoming a profession in which people are increasingly employed on a part-time basis. For example, in Australia 27% of primary teachers and 20% of secondary teachers are employed part-time (Weldon, 2015, p. 13). Secondly, and of more concern, job security is becoming more precarious. The casualization of teachers' work is particularly affecting early career teachers. For example, the latest data in Australia shows that 60% of newly graduated teachers are employed on contracts of less than 1 year (30%) or as casual teachers (30%) and it can take several years to secure long-term employment (Preston, 2019). We know that early career teachers find their work incredibly demanding, but the demands are exacerbated for precariously employed teachers. As they transition into the profession, they are expected to do the same work as more experienced teachers with very little support. They embark on an extremely steep learning curve and find the work relentless. Early career teachers are doing this intense work under close scrutiny as they try to demonstrate that they are performing to a long list of teacher registration standards and, at the same time, vying for secure and ongoing employment (Sullivan et al., 2021).

What is happening to early career teachers and their work?

In a number of countries around the world, the reputation of teachers is not only being questioned but it has become highly politicized. For example, recently in the US, Donald Trump Jr. gave a speech to a crowd in which he stated:

I love seeing some young conservatives, because I know it's not easy. Keep up that fight. Bring it to your schools. You don't have to be indoctrinated by these loser teachers that are trying to sell you on socialism from birth.

Strauss (2019).

Reports indicate such attacks on teachers are occurring in other countries as well. For example, politicians in the Netherlands called on students not to put up with teachers who were brainwashing them in relation to climate change, in Hungary the government censored textbooks, and in Turkey teachers were sacked for allegedly having links to a political group (Strauss, 2019).

Such political statements and actions frame teachers as a “problem” and are aimed at creating derision and uncertainty in the broader public. Smyth and Shacklock (1998) state that such “incomplete views of social reality are constructed to serve particular political interests” (p. 70) and they draw on the work of Wallace (1993, as cited in Smyth and Shacklock, 1998) and Ball (2006) to label this as a “discourse of derision.” Unfortunately, public statements by powerful people continue to politicize teachers and their work. This politicization influences who is attracted into the teaching profession and how they do their work, particularly those teachers at the beginning of their careers.

Neoliberal and neoconservative political views of education are not new; rather they have been influencing teachers and their work for decades (Apple, 2001; Berliner and Biddle, 1995). Such views focus on individual and deficit perspectives of early career teachers in the context of concerns about the “quality” of teacher candidates, university initial teacher education programs, and falling student performance on standardized test scores. This perspective is disseminated in a range of “nation at risk” discourses encapsulated in national and international reports describing the deteriorating quality of teachers and teacher education, for example, the Carter Review of Initial Teacher Education in England (Carter, 2015), the Teacher Prep Review in the United States (Greenberg et al., 2014) and the *Action now: classroom ready teachers* report in Australia (Teacher Education Ministerial Advisory Group, 2014).

Ball (2006) argues that such views are intentional and aimed at undermining perceptions of the success of education systems to bring about more traditional approaches to schooling. Regular attacks on student performance, teacher quality, teacher education, academic standards, teaching methods, and school discipline occur in many countries around the world (Down and Sullivan, 2019). This provides fertile ground for a range of political remedies to “fix” the problem of incompetent and ill-prepared teachers by reasserting control over teachers’ work and focusing on traditional teaching methods such as scripted curriculum, testing, rewards and sanctions, behavior management, and explicit instruction (e.g., Donnelly, 2004). In the context of the wider “culture wars” (Shor, 1992) and “conservative restoration” (Apple, 1996), as well as ongoing debates about the so-called “literacy crisis” and teaching of Indigenous histories, these kinds of instructional strategies are neither innocent nor neutral. In this environment, therefore, it is very easy to slide into disparaging and demeaning public discourses about the declining quality of early career teachers and the profession more generally (Graham, 2015; Mockler, 2021).

Johnson et al. (2019b) describe this as a “negative macro policy environment which implicates early career teachers in the perceived decline in educational standards in developed countries” (p. 213) leading to a de-professionalization of early career teachers’ work and a lack of trust in the work early career teachers do (see also Down and Sullivan, 2019). This negative macro policy environment has led to increased levels of scrutiny and performativity, and spiraling work intensification for teachers, which has a significant impact on new teachers (Johnson et al., 2016).

In this policy environment, early career teachers face numerous personal and professional dilemmas, tensions, and conflicts (Bates and Townsend, 2007; Menter et al., 1997) as they “battle” to survive “in the trenches” (Bezzina, 2006) and navigate their way into the profession in circumstances not always of their own choosing (Johnson et al., 2015). They encounter a school environment in which centralized, top-down, command and control approaches lead to the deskilling and de-professionalization of their work (Bates, 2007; Clarke et al., 2000; Gewirtz et al., 2009; Gleeson and Husbands, 2001; Smyth, 2001). In short, teachers’ work is “over determined and over regulated” through unprecedented levels of accountability, regulation, surveillance, standardization, and high-stakes testing (Ball, 1993, p. 106).

What research has been conducted on early career teachers and the induction phase?

The issues related to early career teachers and their transition to the profession have been well researched. Research has documented and analyzed the challenges that early career teachers face as they transition to the profession and the conditions that support them during this phase. For example, forty years ago Veenman (1984) summarized the literature on the litany of problems neophyte teachers face as they confront the realities of school and classroom life. These include managing student behavior, motivating and engaging students, acknowledging and accommodating students with different abilities, planning appropriate curricula, assessing and reporting learning, interacting with parents, and satisfying external accountability expectations. More recently, Stokking et al. (2003) maintain that the problems that arise when early career teachers are expected to do this work as soon as they are appointed to a teaching position demonstrate that teaching is “simply too demanding” (p. 330). The transition from the relative comfort of university life to “the complexities as well as the full responsibilities of an entire classroom is for many early career teachers a confronting, challenging sometimes deeply disturbing experience” (Kelchtermans, 2019, p. 84).

Kelchtermans (2019, pp. 84–85) argues that there are “four different thematic lines” of research conducted on early career teachers. The first theme of research is individual teachers’ sometimes dramatic and traumatic experiences of “reality shock,” “transition shock,” or “practice shock” (Veenman, 1984, pp. 143–144) when the stark realities of teaching lead to the collapse of idealism and progressive beliefs formed during pre-service teacher education. Problems are manifested as “complaints about workload, stress, and psychological and physical” troubles (p. 144). This theme of research is very problem focused, individualistic, and classroom oriented.

The second theme is the challenges associated with developing “professional relationships with diverse other individuals and groups (colleagues, principals, parents) in the school, with different interests and expectations” (Kelchtermans, 2019, p. 85). This invariably involves negotiating the micropolitics of school life (Johnson, 2004; Johnson et al., 2019a) by developing “micropolitical literacy” (Kelchtermans and Ballet, 2002a) that enables early career teachers to learn to deal with “issues of power, influence and control” (Kelchtermans and Ballet, 2002b, p. 755) as they strive for recognition as competent and valued professionals.

The third theme is teacher attrition and retention and is the most frequently discussed and researched. It is well documented in many countries that the loss of early career teachers to the profession is a profoundly wasteful and threatening trend that will exacerbate looming teacher shortages (Sullivan et al., 2019a). The focus of much research on early career teachers, then, is on the ways to better support them to stay in the profession (e.g. Clandinin et al., 2015; Lavigne, 2014; Long et al., 2012).

The final research category draws on all three former themes but centers on specific strategies within formal teacher induction programs designed to support newly qualified teachers, principally through various forms of “mentoring.” In these arrangements, experienced teachers support novice teachers by providing orientation guidance, information about the school and its culture, explicit coaching, and emotional assistance. A significant body of research documents the most important aspects of early career teacher induction programs and strategies, to which we now turn.

How are early career teachers currently supported, or not?

A considerable amount of research has been undertaken into the induction of early career teachers. Making sense of a large body of research has been made easier by a number of timely reviews of the field over the past 20 years (Frederiksen, 2020; Hobson et al., 2009; Howe, 2006; Ingersoll and Strong, 2011; Long et al., 2012; Schaefer et al., 2012; Strong, 2005; Wang et al., 2008). These reviews have synthesized numerous studies of teacher induction in diverse contexts with the aim of producing “evidence-based” strategies that best support early career teachers’ transition to the profession. For example, in her “review of reviews,” Frederiksen (2020) provides a thematic analysis of seven previous reviews that examined the findings of over 330 studies of the induction process in primary and secondary schools “with their various content and contexts” (p. 49). We present insights from these reviews under the following headings: Aims of Induction Programs; Common Strategies and Components Used in Induction Programs; The Role of School Leaders in Induction Programs; The Influence of School Culture on Induction Programs; and Mentoring.

Aims of induction programs

As we discussed earlier in the chapter, the most common justification for initiating early career teacher induction programs is based on deficit views of graduate teachers; the belief that they lack the knowledge, skills and experience needed to be “classroom-ready teachers” (Down and Sullivan, 2019) and are therefore in need of remedial interventions from their more experienced and competent colleagues. Induction programs based on assumptions of ineptitude feature structured supervisory observation of teaching to identify and rectify teaching skill deficiencies, co-planning activities to ensure that curriculum and lesson preparation are “adequate,” and regular performance appraisal meetings to confirm compliance with school and system requirements (Johnson et al., 2019a, p. 117).

Deficit-based induction programs have been stridently critiqued by researchers (e.g., Kelchtermans, 2019) and school leaders who endeavor to “construct an alternative narrative about early career teachers that speaks against traditional deficit-oriented portrayals” (Johnson et al., 2019b, p. 214). These researchers and leaders recognize the strengths of early career teachers – their enthusiasm, energy, and willingness to innovate to “address the changing needs of twenty-first-century learners” (Johnson et al., 2019a, pp. 108–109), and the fact that they “represent rich networks and bring valuable ideas and practices to the school” (Kelchtermans, 2019, p. 94). Induction programs based on a strengths approach seek to invest in the ongoing professional development of early career teachers. The early career teachers in Johnson et al. (2019a) study received systematic, coordinated and strategic professional development opportunities as part of their induction to ensure that they contributed in positive ways to the achievement of their schools’ aims.

Common strategies and components used in induction programs

Research into early career teacher induction has identified a wide range of activities that are usually presented as lists of “positive things to do” to support early career teachers (Davenport, 2009; Frederiksen, 2020). It is worth noting, however, that most research into early career teacher induction focuses on teachers who are employed on long or ongoing contracts when, in many countries, many beginning teachers are employed on insecure, short-term, and tenuous contracts (Preston, 2019). Given this caveat, the components of “good” induction programs identified across many reviews (Frederiksen, 2020; Hobson et al., 2009; Howe, 2006; Ingersoll and Strong, 2011; Long et al., 2012; Schaefer et al., 2012; Strong, 2005; Wang et al., 2008) include:

- conducting a comprehensive initial orientation to the school
- planning a long-term induction program spanning 2–3 years

- providing instructional support
 - sharing curriculum and teaching materials
 - implementing mutual lesson observations
- establishing collaborative working arrangements
 - team teaching
 - co-planning curricula
 - forming network groups with peers (early career teachers and more experienced teachers or “mentors”)
- modifying work arrangements
 - reducing workload
 - providing extra release time (e.g., to write student reports)
 - de-privatizing early career teachers’ work by creating physical spaces that promote and enable teachers to collaborate more
- promoting teacher agency
 - seeking teachers’ views about their induction needs
 - offering choices about induction activities
- monitoring performance
 - observing and providing feedback on classroom teaching
 - structuring regular “performance review” meetings
 - encouraging reflective practices.

The role of school leaders in induction programs

School leaders have a pivotal role in developing early career teachers’ professional competencies and commitment. Focusing explicitly on the micropolitical work of school leaders (Blase and Bjork, 2010), Johnson et al. (2019a) identified the strategies they use to develop and retain the teachers they want:

They modify early career teachers’ work arrangements to reduce their isolation, promote their involvement in collaborative teaching arrangements and reduce their workload by providing extra release time and shared curriculum resources. ... [T]hey invest heavily in developing early career teachers ... by encouraging their involvement in formal professional learning activities, publicly recognising and rewarding their contributions and providing opportunities for future leadership. (pp. 118–119)

Understanding the way in which school leaders can structure and influence the work of early career teachers is complex, but important. School leaders are responsible for establishing the conditions that enhance the retention of quality teachers in their school (Schleicher, 2012), but systems are responsible for establishing such conditions across schools.

The influence of school culture on induction programs

Closely linked to school leaders’ micropolitical work is research into the impact of school culture and school context on the professional development of early career teachers. Frederiksen (2020) summarizes four of the reviews she examined that address the role of school culture in teacher induction:

These four reviews indicate that new teachers have very different opportunities to develop professionally, depending on school culture. The effect of the teacher induction program depends on the social, cultural and organisational context in which it is situated. ... [A] supportive, interested and encouraging management team is identified as a key factor in developing an open, collegial and collaborative school culture. (p. 60)

However, early career teacher induction can be “undermined if there is poor leadership and a non-conducive culture at the school, with a lack of ... professional communities” (Frederiksen, 2020, p. 61).

Mentoring

As Orland-Barak (2016) points out, “mentoring” is the most common label given to a suite of induction support strategies for early career teachers. Kelchtermans (2019) argues:

Most support initiatives for early career teachers encompass some sort of mentoring, in which a senior, more experienced teacher (the mentor), engages with novice colleagues (the mentees) by providing different forms of assistance, including information, coaching, technical training and moral and emotional support. (p. 85)

While there is considerable debate in the literature about the efficacy of mentoring (see Kardos and Johnson, 2010), Hobson et al. (2009) and Shockley et al. (2013) have identified the following features of positive mentoring arrangements (summarized by Frederiksen, 2020, pp. 57, 62):

- The mentor needs to be trained for the role.
- The school culture needs to value and support collegiality and cooperation between staff.
- Mentors are given extra time to prepare and perform their mentoring role.
- Mentors and mentees meet during school time.
- Mentors and mentees collaboratively plan induction activities and scheduling.
- Both mentors and mentees have access to other supports in the school beyond the mentor relationship.

How might early career teachers' work be reimagined?

While many of the above initiatives are welcome and even useful, they are insufficient to address the wider structural and institutional arrangements in which early career teachers find themselves. In this final section, we wish to advance some alternative ways of thinking about early career teachers' work and induction processes. In this task, we endeavor to progress a *democratic* teacher-driven approach which attempts to address some of the shortcomings of existing explanations of early career teacher quality and what might be done about it. We argue that it is timely to shift away from the limiting effects of management pedagogies, and instead think anew about what it means to teach. This involves creating a more participatory, critically engaged, and activist approach to early career teacher induction. In other words, we want to imagine early career teachers' work in ways that begin the process of reclaiming and reimagining the democratic purposes of education (Apple and Beane, 2007; Dewey, 1944 [1916]). In pursuing this task, we find the idea of a "critical learning community" useful because the "indigenous knowledge" of teachers is valued, trusted, and respected (Nieto, 2003; Sachs, 2003; Sergiovanni, 1994; Smyth, 1998; Wenger, 1998). Here, early career teachers learn to work together and sustain the conditions necessary for learning (Day, 1999; Nias et al., 1992). They have ownership and agency over their work including curriculum, pedagogy, and assessment, and are seen as valuable contributors to ongoing school improvement (Day et al., 2006; Day et al., 2006; Hargreaves and Fullan, 2012). Behind this empowering approach to teacher professionalism is a commitment to building both individual and collective capabilities, which Sen (1992) defines as those "valuable functionings that make up our lives, and more generally, our freedom to promote objectives we have reasons to value" (p. 3). Central to this broader project is the "quality (the 'well-ness', as it were) of the person's being" (p. 39).

Thomas and Schubert (2001) expand on this bolder vision when they describe three main anchor points of teaching beyond the current fetish with standards and management pedagogies. First, teachers should be "engaged in philosophic inquiry," that is, "investigating the value assumptions of their students, their colleagues, and their own metaphysical, epistemological, and axiological convictions" (p. 234). Second, teachers need to develop as "democratic connoisseur[s]" or "critical interpreters of existent curriculums and creators of new curriculums, novel forms of instruction, and appropriate methods of assessment" (p. 235). Third, teachers should see themselves as "progressive activist[s]" committed to "democratic practice understood as ... public advocacy for social policies that attempt to redress injustice and public criticisms of state actions that oppress or institutionalize inequality" (p. 235; see also Ayers, 2004; Freire, 1998).

There are numerous ideas and strategies available in the literature to support this kind of transformational work, for example, "teachers as intellectuals" (Giroux, 1988), "teachers as knowledge workers" (Kincheloe, 1993, 2001), "activist teacher professionals" (Sachs, 2003), "teachers as policy actors" (Ball et al., 2011; Gale and Densmore, 2003), "teachers as cultural workers" (Freire, 1998), and "becoming a critically reflective teacher" (Carr and Kemmis, 1986; Kress, 2011; Smyth, 2021). Giroux (1988) explains in a little more detail how

teachers as intellectuals will need to reconsider and, possibly, transform the fundamental nature of the conditions under which they work. That is, teachers must be able to shape the ways in which time, space, activity, and knowledge organize everyday life in schools. More specifically, in order to function as intellectuals, teachers must create the ideology and structural conditions necessary for them to write, research, and work with each other in producing curricula and sharing power. ... As intellectuals, they will combine reflection and action in the interest of empowering students with the skills and knowledge needed to address injustices and to be critical actors committed to developing a world free of oppression and exploitation. (p. xxxiv)

While these emancipatory ideas have been around for some time, the challenge of questioning the dominant ways in which teachers' work is organized is fraught with difficulties and resistance (Diem and Helfenbein, 2008). Said (1994) explains how the process of conservatism in the profession is maintained as teachers "become tame and accepting of whatever the so-called leaders in the field will allow" (p. 57). As a result, the "moral agency" of teachers and their identities is threatened because "praxis demands creative thinking, care, compassion, and critical consciousness – thinking outside or beyond the box" (Kemmis and Smith, 2008, p. 5).

If we are serious about changing the conditions under which early career teachers work, then it will require profound changes in the ways in which schools and school systems are organized, led, and imagined (Gale and Densmore, 2000; Goodlad et al., 2008). In the Australian context, for example, Johnson et al. (2016) identify six important practices necessary to building a more democratic school culture:

- engaging teachers in participatory democratic decision making;
- promoting a sense of connectedness and belonging to place;
- encouraging a spirit of collective ownership, responsibility and solidarity;
- developing awareness of agency and power to make a difference;
- building a culture of collaboration, innovation and risk taking; and
- producing professional knowledge based on the concerns of classroom teachers (p. 140).

At heart, therefore, is an awareness that “another [kind of] school is possible” (Wrigley, 2006). This raises questions about whether another kind of education system is possible, one based on the values of critical inquiry, social justice, democracy, compassion, care, and respect for all. In these schools, in the words of Freire (2007 [1974]), “people [teachers] develop their power to perceive critically *the way they exist* in the world *with which* and *in which* they find themselves” (p. 83). Early career teachers flourish in schools (and systems) where they are encouraged to conduct local investigations into the conditions and circumstances in which they (and their students) find themselves (Kelchtermans, 2021; Loughran, 1999). In the words of Kress (2011), this involves

breathing new life into research methods for teachers [in a way that] requires scholar-practitioners to develop critical consciousness about who they are in relation to their students and the larger society [whilst] conducting research that is fair, ethical, and empowering for all stakeholders. (p. 10)

However, as Connell (1985) reminds us, schools are workplaces in which the “political order” reflects “the pattern of authority and consent, alliance and co-operation, resistance and opposition that characterize the institution as a whole” (p. 131). Unsurprisingly, early career teachers flourish best when they are invited to participate in reciprocal and horizontal learning relationships in which their knowledge, expertise, and energy are valued (Johnson et al., 2014, 2016). In this sense, schools are deeply relational places in which early career teachers begin to make sense of themselves, their context, and what it means to teach (Day and Gu, 2014). How early career teachers can do this across multiple schools remains uncertain.

Of course, this is a lot to expect of early career teachers and involves a degree of personal and professional discomfort in terms of the emotional “dark side” (Brookfield, 1995) of undertaking critical self-reflection including questioning some deeply entrenched assumptions, values, and beliefs about teaching. Nonetheless, Brookfield (1995) believes there are five major benefits to be gained:

- It helps us take informed actions.
- It helps us develop a rationale for practice.
- It helps us avoid self-laceration.
- It grounds us emotionally.
- It enlivens our classrooms.
- It increases democratic trust (pp. 22–26).

Yet, there is a widespread political and popular view that teaching is not complex work, and that it can be reduced to a list of standards which are identifiable, calculable, and transparent, and thus more rigidly controlled and scripted (Fitzgerald and Savage, 2013). In contrast, a democratic approach to early career teacher induction draws attention to the complexity and fragility of teachers' work (Lortie, 2002 [1975]), as well as the emotional demands of the job (Connell, 1993, p. 63) and the centrality of critical learning communities in supporting early career teachers.

Acknowledgments

This paper is an outcome of the Early Career Teacher Resilience and Retention Study (LP0883672) and the Retaining Quality Teachers Study (LP130100830), both funded by the Australian Government through the Australian Research Council Linkage Scheme. The Early Career Teacher Resilience and Retention Study also received funds and/or in-kind support from eight education sector organizations in South Australia and Western Australia. The Retaining Quality Teachers Study received funds and in-kind support from the Principals Australia Institute. Note: the views expressed in this article do not necessarily reflect any of these educational organizations' policies.

References

- Angus, L., 2012. Preparing teachers as informed professionals: working with a critical ethnographic disposition and a socially just democratic imaginary. In: Down, B., Smyth, J. (Eds.), *Critical Voices in Teacher Education: Teaching for Social Justice in Conservative Times*. Springer, Dordrecht, pp. 45–61.
- Apple, M.W., 1996. *Cultural Politics and Education*. Open University Press, Buckingham, UK.

- Apple, M.W., 2001. *Educating The "Right" Way: Markets, Standards, God and Inequality*. Routledge Falmer, New York, NY.
- Apple, M.W., Beane, J. (Eds.), 2007. *Democratic Schools: Lessons in Powerful Education*. Heinemann, Portsmouth, NH.
- Australian Institute for Teaching and School Leadership, 2020. *National Initial Teacher Education Pipeline: Australian Teacher Workforce Data Report 1*. Education Services Australia, Melbourne.
- Ayers, W., 2004. *Teaching toward Freedom: Moral Commitment and Ethical Action in the Classroom*. Beacon Press, Boston, MA.
- Ball, S., 2006. *Education Policy and Social Class: The Selected Works of Stephen Ball*. Routledge, London.
- Ball, S., Maguire, M., Braun, A., Hoskins, K., 2011. Policy subjects and policy actors in schools: some necessary but insufficient analyses. *Discourse: Stud. Cultur. Polit. Educ.* 32 (4), 611–624.
- Ball, S.J., 1993. Education policy, power relations and teachers' work. *Br. J. Educ. Stud.* 41 (2), 106–121.
- Ball, S.J., 2016. Neoliberal education? Confronting the slouching beast. *Pol. Futures Educ. Internet* 14 (8), 1046–1059.
- Bates, R., 2007. Regulation and autonomy in teacher education: system or democracy? In: Townsend, T., Bates, R. (Eds.), *Handbook of Teacher Education: Globalisation, Standards and Professionalism in Times of Change*. Springer, Dordrecht, Netherlands, pp. 127–140.
- Bates, R., Townsend, T., 2007. The future of teacher education: challenges and opportunities. In: Bates, R., Townsend, T. (Eds.), *Handbook of Teacher Education: Globalisation, Standards and Professionalism in Times of Change*. Springer, Dordrecht, Netherlands, pp. 727–737.
- Berliner, D., Biddle, B., 1995. *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. Basic Books, New York, NY.
- Bezzina, C., 2006. Views from the trenches: beginning teachers' perceptions about their professional development. *J. Serv. Educ.* 32 (4), 411–430.
- Blase, J., Bjork, L., 2010. The micropolitics of educational change and reform: cracking open the black box. In: Hargreaves, A. (Ed.), *Second International Handbook of Educational Change*. Springer, Dordrecht, Netherlands, pp. 237–258.
- Brookfield, S., 1995. *Becoming a Critically Reflective Teacher*. Jossey-Bass, San Francisco, CA.
- Carr, W., Kemmis, S., 1986. *Becoming Critical: Education Knowledge and Action Research*. Routledge, London.
- Carter, A., 2015. *Carter Review of Initial Teacher Training (ITT)*. Department for Education, London.
- Chrisafis, A., 2022. French Teachers Walk Out of Classrooms in Strike over Covid Strategy. *The Guardian*. January 14. <https://www.theguardian.com/world/2022/jan/13/half-of-french-primary-schools-expected-to-close-teachers-strike-protest-covid-education>.
- Clandinin, D.J., Long, J., Schaefer, L., Downey, C.A., Steeves, P., Pinnegar, E., McKenzie Robblee, S., Wnuk, S., 2015. Early career teacher attrition: intentions of teachers beginning. *Teach. Educ.* 26 (1), 1–16.
- Clarke, J., Gewirtz, S., McLaughlin, E., 2000. *New Managerialism, New Welfare?* Sage, London.
- Connell, R., 1985. *Teachers' Work*. Allen & Unwin, North Sydney, NSW.
- Connell, R., 1993. *Schools and Social Justice. Our Schools/Our Selves*. Education Foundation, Toronto.
- Davenport, S., 2009. *An Investigation of the Strategies Used by One School to Support Early Career Teachers in their First Year of Teaching*. Honours thesis. University of South Australia.
- Day, C., 1999. *Developing Teachers: the Challenges of Lifelong Learning*. Falmer Press, London.
- Day, C., Gu, Q., 2014. *Resilient Teachers, Resilient Schools*. Routledge, London.
- Day, C., Kington, A., Stobart, G., Sammons, P., 2006a. The personal and professional selves of teachers: stable and unstable identities. *Br. Educ. Res. J.* 32 (4), 601–616.
- Day, C., Stobart, G., Sammons, P., Kington, A., Gu, Q., Smees, R., Mujtaba, T., 2006b. *Variations in Teachers' Work, Lives and Effectiveness*. Department for Education and Skills, London.
- Dewey, J., 1944 [1916]. *Democracy and Education: an Introduction to the Philosophy of Education*. Free Press, New York.
- Diem, J., Helfenbein, R.J. (Eds.), 2008. *Unsettling Beliefs: Teaching Theory to Teachers*. Information Age Publishing, Charlotte, NC.
- Donnelly, K., 2004. *Why Our Schools are Failing*. Duffy & Snellgrove, Sydney, NSW.
- Down, B., Sullivan, A., 2019. "Classroom-ready teachers": gaps, silences and contradictions in the Australian report into teacher education. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 39–61.
- Fitzgerald, T., Savage, J., 2013. Scripting, ritualising and performing leadership: interrogating recent policy developments in Australia. *J. Educ. Adm. Hist.* 45 (2), 126–143.
- Flores, M.A., 2019. Unpacking teacher quality: key issues for early career teachers. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 15–38.
- Frederiksen, L., 2020. Support for newly qualified teachers through teacher induction – a review of reviews. In: Olsen, K., Bjerkholt, E., Heikkinen, H. (Eds.), *New Teachers in Nordic Countries – Ecologies of Mentoring and Induction*. Cappelen Damm Akademisk, Norway, pp. 49–70.
- Freire, P., 1998. *Teachers as Cultural Workers*. Westview Press, Boulder, CO.
- Freire, P., 2007 [1974]. *Education for Critical Consciousness*. Continuum, New York, NY.
- Gale, T., Densmore, K., 2000. *Just Schooling: Explorations in the Cultural Politics of Teaching*. Open University Press, Buckingham, UK.
- Gale, T., Densmore, K., 2003. Democratic educational leadership in contemporary times. *Int. J. Leader. Educ.* 6 (2), 119–136.
- Gewirtz, S., Mahony, P., Hextall, I., Cribb, A., 2009. *Changing Teacher Professionalism: International Trends, Challenges and Ways Forward*. Routledge, London.
- Giroux, H., 1988. *Teachers as Intellectuals: Toward a Critical Pedagogy of Learning*. Bergin & Harvey, New York, NY.
- Gleeson, D., Husbands, C., 2001. *The Performing School: Managing, Teaching, and Learning in a Performance Culture*. Routledge, London.
- Goodlad, J.I., McDaniel, B., Soder, R. (Eds.), 2008. *Education and the Making of a Democratic People*. Paradigm Publishers, Boulder, CO.
- Graham, L.J., 2015. Lay off Blaming New Teachers for "Falling" Standards. *EduResearch Matters*, March 29. <http://www.aare.edu.au/blog/?p=964>.
- Greenberg, J., Walsh, K., McKee, A., 2014. *Teacher Prep Review: a Review of the Nation's Teacher Preparation Programs*. National Council on Teacher Quality, Washington, DC.
- Han, S.W., Borgonovi, F., Guerriero, S., 2020. Why don't more boys want to become teachers? The effect of a gendered profession on students' career expectations. *Int. J. Educ. Res.* 103, 1–17.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press, New York, NY.
- Hobson, A.J., Ashby, P., Malderez, A., Tomlinson, P.D., 2009. Mentoring beginning teachers: what we know and what we don't. *Teach. Teach. Educ.* 25 (1), 207–216.
- Howe, E.R., 2006. Exemplary teacher induction: an international review. *Educ. Philos. Theor.* 38 (3), 287–297.
- Ingersoll, R.M., Strong, M., 2011. The impact of induction and mentoring programs for beginning teachers: a critical review of the research. *Rev. Educ. Res.* 81 (2), 201–233.
- Johnson, B., 2004. Local school micropolitical agency: an antidote to new managerialism. *Sch. Leader. Manag.* 24 (3), 267–286.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A., Pearce, J., Hunter, J., 2016. *Promoting Early Career Teacher Resilience: A Socio-Cultural and Critical Guide to Action*. Routledge, London.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A.M., Pearce, J., Hunter, J., 2014. *Promoting early career teacher resilience: a framework for understanding and acting*. *Teach. Teach. Theor. Pract.* 20 (5), 530–546.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A.M., Pearce, J., Hunter, J., 2015. *Early Career Teachers: Stories of Resilience*. Springer, Singapore.

- Johnson, B., Sullivan, A., Simons, M., 2019a. How school leaders attract, recruit, develop and retain the early career teachers they want: positives and paradoxes. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 101–122.
- Johnson, B., Sullivan, A., Simons, M., 2019b. Teacher retention: some concluding thoughts. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 211–220.
- Kardos, S.M., Johnson, S.M., 2010. New teachers' experiences of mentoring: the good, the bad, and the inequity. *J. Educ. Change* 11 (1), 23–44.
- Kearney, S., 2021. The challenges of beginning teacher induction: a collective case study. *Teach. Educ.* 32 (2), 142–158.
- Kelchtermans, G., 2017. "Should I stay or should I go?" Unpacking teacher attrition/retention as an educational issue. *Teacher. Teach.* 23 (8), 961–977.
- Kelchtermans, G., 2019. Early career teachers and their need for support: thinking again. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 83–98.
- Kelchtermans, G., 2021. Keeping educational research close to practice. *Br. Educ. Res. J.* 47 (6), 1504–1511.
- Kelchtermans, G., Ballet, K., 2002a. Micropolitical literacy: reconstructing a neglected dimension in teacher development. *Int. J. Educ. Res.* 37 (8), 755–767.
- Kelchtermans, G., Ballet, K., 2002b. The micropolitics of teacher induction. a narrative-biographical study on teacher socialisation. *Teach. Teach. Educ.* 18 (1), 105–120.
- Kemmis, S., Smith, T., 2008. Praxis and praxis development. In: Kemmis, S., Smith, T. (Eds.), *Enabling Praxis: Challenges for Education*. Sense, Rotterdam, pp. 3–13.
- Kincheloe, J., 1993. *Toward a Critical Politics of Teacher Thinking: Mapping the Postmodern*. Bergin & Garvey, Westport, CT.
- Kincheloe, J., 2001. *Getting beyond the Facts: Teaching Social Studies/social Sciences in the Twenty-First Century*. Peter Lang, New York, NY.
- Kress, T., 2011. *Critical Praxis Research: Breathing New Life into Research Methods for Teachers*. Springer, Dordrecht, Netherlands.
- Lambert, K., Gray, C., 2021. Hyper-performativity and Early Career Teachers: Interrogating Teacher Subjectivities in Neoliberal Educational Assemblages. <https://doi.org/10.1080/01596306.2021.1933912>. *Discourse: Studies in the Cultural Politics of Education*. Advance online publication.
- Lavigne, A.L., 2014. Beginning teachers who stay: beliefs about students. *Teach. Teach. Educ.* 39, 31–43.
- Long, J.S., McKenzie-Robbles, S., Schaefer, L., Steeves, P., Wnuk, S., Pinnegar, E., Clandinin, D.J., 2012. Literature review on induction and mentoring related to early career teacher attrition and retention. *Mentor. Tutoring: Partner. Learn.* 20 (1), 7–26.
- Lortie, D., 2002 [1975]J. *School-teacher: a Sociological Study*. University of Chicago Press, Chicago, IL.
- Loughran, J.J. (Ed.), 1999. *Researching Teaching: Methodologies and Practices for Understanding Pedagogy*. Falmer Press, London.
- McGowan, M., Rose, T., 2022. "Clash of Two Crises": Fears for NSW Schools as Covid Pandemic and Ongoing Teacher Shortages Collide, Australian edition. *The Guardian*. January 19.
- Menter, I., Muschamp, Y., Nicholls, P., Ozga, J., Pollard, A., 1997. *Work and Identity in the Primary School: a Post-Fordist Analysis*. Open University Press, Buckingham, UK.
- Mockler, N., 2021. *Constructing Teacher Identities: How the Print Media Defines and Represents Teachers and Their Work*. Bloomsbury, London.
- Moreau, M.-P., 2019. *Teachers, Gender and the Feminisation Debate*. Routledge, Abingdon, Oxon.
- Nias, J., Southworth, G., Campbell, P., 1992. *Whole School Curriculum Development in the Primary School*. Routledge, London.
- Nieto, S., 2003. *What Keeps Teachers Going?* Teachers College Press, New York, NY.
- Orland-Barak, L., 2016. Mentoring. In: Loughran, J.J., Hamilton, M. (Eds.), *International Handbook on Teacher Education*. Springer, Singapore, pp. 105–141.
- Perryman, J., Ball, S.J., Braun, A., Maguire, M., 2017. Translating policy: governmentality and the reflective teacher. *J. Educ. Pol.* 32 (6), 745–756.
- Preston, B., 2019. Reforming replacement teaching: a game changer for the development of early career teaching? In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 161–191.
- Rivers, H., 2022. Omicron-fuelled Teacher Shortage Predicted to Drive More School Closings. *The London Free Press*. January 14. <https://lfpres.com/news/local-news/teacher-shortage-predicted-to-drive-more-school-closings-as-omicron-spread-widens>.
- Sachs, J., 2003. *The Activist Teaching Profession*. Open University Press, Buckingham, PA.
- Said, E., 1994. *Representations of the Intellectual*. Vintage Books, New York, NY.
- Sainato, M., 2021. "Exhausted and Underpaid": Teachers across the US are Leaving Their Jobs in Numbers. *The Guardian*. October 4. <https://www.theguardian.com/world/2021/oct/04/teachers-quitting-jobs-covid-record-numbers>.
- Schaefer, L., Long, J., Clandinin, D.J., 2012. Questioning the research on early career teacher attrition and retention. *Alberta J. Educ. Res.* 58 (1), 106–121.
- Schleicher, A. (Ed.), 2012. *Preparing Teachers and Developing School Leaders for the 21st Century: Lessons from Around the World*. OECD Publishing, Paris.
- Sen, A., 1992. *Inequality Re-examined*. Harvard University Press, Cambridge, MA.
- Sergiovanni, T.J., 1994. Organizations or communities? Changing the metaphor changes the theory. *Educ. Adm. Q.* 30 (2), 214–226.
- Shearing, H., 2021. Covid: Doubts that Ex-Teachers Will Return by January. *BBC News*. December 23. <https://www.bbc.com/news/education-59761859>.
- Shockley, R., Watlington, E., Felsher, R., 2013. Out on a limb: the efficacy of teacher induction in secondary schools. *NASSP Bull.* 97 (4), 350–377.
- Shor, I., 1992. *Culture Wars: School and Society in the Conservative Restoration*. University of Chicago Press, Chicago, IL.
- Smyth, J., 1998. The school as a critical learning community. In: Down, B., Hogan, C., Swan, P. (Eds.), *Reclaiming Professional Knowledge: New Ways of Thinking about Teachers' Learning*. Murdoch University, Perth, WA, pp. 7–14.
- Smyth, J., 2001. *Critical Politics of Teachers' Work: an Australian Perspective*. Peter Lang, New York, NY.
- Smyth, J., 2012. Problematising teachers' work in dangerous times. In: Down, B., Smyth, J. (Eds.), *Critical Voices in Teacher Education: Teaching for Social Justice in Conservative Times*. Springer, Dordrecht, Netherlands, pp. 13–25.
- Smyth, J., 2021. *Critical Research and Theorizing: a Collection of Essays from the Critical Pedagogy Networker, 1988–2002*. DIO Press, New York, NY.
- Smyth, J., Shacklock, G., 1998. *Remaking Teaching: Ideology, Policy, and Practice*. Routledge, London.
- Stokking, K., Leenders, F., De Jong, J., Van Tartwijk, J., 2003. From student to teacher: reducing practice shock and early dropout in the teaching profession. *Eur. J. Teach. Educ.* 26 (3), 329–350.
- Strauss, V., 2019. Why Donald Trump Jr.'s "Loser Teachers" Comment was "a Chilling Moment" for Educators Around the World. *The Washington Post*. February 16. <https://www.washingtonpost.com/education/2019/02/16/why-donald-trump-jr-loser-teachers-comment-was-a-chilling-moment-educators-around-world/>.
- Strong, M., 2005. Teacher induction, mentoring, and retention: a summary of the research. *N. Educat.* 1 (3), 181–198.
- Sullivan, A., Johnson, B., Simons, M. (Eds.), 2019a. *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore.
- Sullivan, A., Johnson, B., Simons, M., 2019b. Introduction. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers: Issues and Opportunities*. Springer Nature, Singapore, pp. 1–11.
- Sullivan, A., Johnson, B., Simons, M., Tippet, N., 2021. When performativity meets agency: how early career teachers struggle to reconcile competing agendas to become "quality" teachers. *Teacher. Teach.* 27 (5), 388–403.
- Sutcher, L., Darling-Hammond, L., Carver-Thomas, D., 2016. *A Coming Crisis in Teaching? Teacher Supply, Demand, and Shortages in the U.S.* Learning Policy Institute, Palo Alto, CA.
- Teacher Education Ministerial Advisory Group, 2014. *Action Now: Classroom Ready Teachers*. Department of Education, Canberra.
- Thomas, T.P., Schubert, H., 2001. Reinterpreting teacher certification standards: locating limitations and possibilities. In: Kincheloe, J., Weil, D. (Eds.), *Standards and Schooling in the United States*. ABC-CLIO, Santa Barbara, CA, pp. 229–243.
- Veenman, S., 1984. Perceived problems of beginning teachers. *Rev. Educ. Res.* 54 (2), 143–178.

- Wang, J., Odell, S.J., Schwill, S.A., 2008. Effects of teacher induction on beginning teachers' teaching: a critical review of the literature. *J. Teach. Educ.* 59 (2), 132–152.
- Weldon, P., 2015. The Teacher Workforce in Australia: Supply, Demand and Data Issues (Policy Insights, Issue 2). Australian Council for Educational Research, Melbourne, Vic.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge.
- Wilson, R., Carabatta, G., 2022. COVID and schools: Australia is about to feel the full brunt of its teacher shortage. In: *The Conversation*. January 19. <https://theconversation.com/covid-and-schools-australia-is-about-to-feel-the-full-brunt-of-its-teacher-shortage-174885?>
- Wrigley, T., 2006. *Another School Is Possible*. Trentham Books, Stoke-on-Trent, UK.

Coaching for professional learning of early career teachers

Pasi Sahlberg^b and Andrea Stringer^a, ^a University of New South Wales, Sydney, NSW, Australia; and ^b Southern Cross University, Lismore, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction to coaching for teacher development	144
Induction, mentoring, and coaching in teachers' professional development	145
Australian case study: coaching early career teachers	148
Purpose of coaching is clear but is not the same among school	148
Coaching is an investment to professional capital in school	149
Coaching in action	149
Lessons for teacher development through coaching	150
References	151

Introduction to coaching for teacher development

A frequently heard adage is that teaching is not rocket science, it is harder than that. Keeping 25–30 students with different interest and abilities engaged in productive learning lesson after lesson in schools is not an easy task.

Then there are parents, colleagues, administrators, and politicians who expect teachers to go above and beyond what is often reasonable to expect. Consequently, in many countries today teachers feel powerless in front of these external expectations. Declining job satisfaction and even burn-out have made many teachers to decide to leave teaching (OECD, 2019). Initial teacher education providers in most countries have difficulties recruiting smart, motivated people into their teacher education programs. In the United States and Australia, for example, there have been significant declines in these enrolments during the past decade. Even in countries where teaching has previously been an attractive career choice for young people, more experienced teachers are leaving the teaching profession for growing dissatisfaction and less young people consider teaching in school as a competitive career option.

There is no silver bullet that would fix the challenges the teaching profession is facing around the world. One important part of the solution is to develop initial teacher education so that it is akin to ways other professionals, medical doctors, lawyers, and architects, for example, are educated. In this chapter we will focus the other part of the solution that would help newly educated teachers in the early part of their teaching careers to get engaged in and remain committed to teaching in their first schools. This includes different practices of induction, mentoring, and coaching, which can be critically important in retaining qualified, passionate teachers in schools.

Based on our recent research and evidence for the current knowledge base we argue, that investing in early career teachers systematically through induction and professional learning that includes school embedded coaching can have a significant positive impact on teacher's professional growth, motivation levels and contribute to retaining beginning teachers in schools.

Coaching as part of professional learning and induction of teachers has been researched and practiced since the 1980s (Cox et al., 2014; Showers and Joyce, 1996; Whitmore, 2009). Yet, there is no universal definition or practice of coaching, that makes also the research of coaching complicated. What is considered coaching by one person might mean mentoring to another. In one school, coaching may mean improving teacher practice, student learning, and wellbeing. In another coaching may mean developing school culture and its organisational functions.

It is important to decide the purpose of coaching when selecting or designing a coaching program for a school. While there are many practices and definitions of coaching, teacher development professionals agree that coaching often sits on a continuum of a more directive and less directive stance (Downey, 2003; Parsloe and Leedham, 2009; Passmore and Sinclair, 2020). Facilitative coaching, instructional coaching, and mentoring are often identified on this continuum. The less directive stance of a facilitative coach, according to Whitmore (2009) does not teach but rather helps people learn and unlock hidden potentials and opportunities to professional growth.

The following examples of coaching practice are common in contemporary professional and research literature on teacher development and school improvement.

1. *Instructional or Teacher Coaching* focuses on improving quality of teaching and students' learning through regular one-to-one coach-teacher sessions in school. Instructional coaching have found to have positive impact of the quality of teaching. For example, in Kraft et al. (2018) meta-analysis of 60 studies of teacher coaching programs found larger positive effects on teacher instruction and smaller positive effects on student achievement. The University of Kansas Center for Research on Learning (2022) and Jim Knight (2007, 2009, 2011, 2018, 2019) have studied instructional coaching, communication, and other forms of professional development for over twenty years reporting similar findings.
2. *Mentoring and coaching* are situated on a continuum and designed to meet educators at their point of need. Rachel Lofthouse at the Carnegie School of Education at Leeds Beckett University directs CollectivEd: The Mentoring and Coaching Hub (2022)

where she connects professionals, academic and students with shared interests. [Hollweck and Lofthouse \(2021\)](#) and [Lofthouse \(2019a,b, 2014, 2013\)](#) has studied mentoring and coaching in practice and developed ways to support the implementation of coaching within schools ([Lofthouse et al., 2010](#)).

3. *Cognitive Coaching* is a professional learning process where teachers examine the thinking behind their practice. It emphasizes the development of teachers as self-directed learners with dispositions for lifelong professional learning ([Costa and Garmston, 2015](#)). Through self-management, self-monitoring, and self-modification, the cognitive coaching approach aims to increase an individual's potential for self-directed learning.
4. *Positive Psychology Coaching* is "a managed conversational process that supports people to achieve meaningful goals in a way that enhances their wellbeing" ([van Nieuwerburgh and Biswas-Diener, 2021, 315](#)). As an academic, researcher, executive coach and consultant van Nieuwerburgh is interested in the integration of coaching and positive psychology in education and health settings ([van Nieuwerburgh, 2014](#)). Another example is Coaching Psychology developed by Anthony Grant ([2018, 2011, 2014, 2016, 2021](#)) who has also led evidence-based approaches to coaching in the workplace, with executives and organisations.

In the next section we provide a more detailed conceptual description of the key terminology related to coaching early career teachers.

Induction, mentoring, and coaching in teachers' professional development

When new graduates in other professions enter their first workplace, they often get support in their early careers through various mentoring programs or internships. This is not always that case in education and teaching has become an increasingly complex profession. OECD's Teaching and Learning International Surveys ([2014, 2019](#)), for example, clearly show the lack of consistent and well-developed professional support for early career teachers. Early career teachers' experiences and professional growth depend on many variables, but a sharp increase in the continuation of learning or professional growth is an expectation. In addition, once accredited, educators are often required to provide evidence of professional learning for the maintenance of teacher accreditation. Common elements, according to [Darling-Hammond et al. \(2017\)](#), of effective professional learning include coaching, expert support, and time for feedback and reflection. But still today many teachers start their teaching careers without adequate guidance. Coaching would give them a stronger start and probably impact their adaptation into schools.

There has been positive progress in many countries during the past three decades to provide early career teachers with opportunities of support from trained and experienced coaches and proficient teachers, while also affording opportunities for reflection and feedback. "Sink or swim" is the metaphor often used to depict a teacher's first few years of practice ([Glazzard and Coverdale, 2018; Howe, 2006; Varah et al., 1986](#)). In many countries that have less-consistent professional scaffolds for new teachers, up to half of these early career teachers leave the profession before they finish five years of teaching ([Darling-Hammond et al., 2017](#)).

Supporting teachers with specific induction programs can have a positive impact on beginning teachers ([AITSL, 2016; Hay-Group, 2014; Ingersoll and Strong, 2011; Ramsey, 2000](#)). However, the induction process design is inconsistent and unregulated nationally and internationally ([Craven et al., 2014; OECD, 2018, 2019; Paniagua and Sánchez-Martí, 2018](#)). Offering a fresh perspective, early career teachers need support, guidance, and a degree of agency, as they bring current pedagogical knowledge, technological awareness, and enthusiasm to the school context.

Schools today increasingly develop induction policies and programs to support beginning and experienced teachers who are new to the school. It is vital for school leaders to create a collaborative culture in a positive working environment to build beginning teachers' resilience, improve efficacy, and develop a sense of belonging ([Gu and Day, 2013](#)). While induction programs promote teacher quality and job satisfaction ([Ingersoll and Strong, 2011; Paniagua and Sánchez-Martí, 2018](#)), the inconsistency is evident, as professional and emotional support is insufficient for some early career teachers ([Gallant and Riley, 2014](#)), and isolation and frustration is experienced ([Buchanan et al., 2013](#)). Programs and processes designed to encourage teacher professional growth must eliminate these levels of frustration and generate more satisfaction in the teachers' nascent years.

Mentoring is an important part of induction programs, and the terms induction and mentoring are frequently used synonymously ([Ingersoll and Smith, 2004; Long et al., 2012](#)). Mentoring and coaching are commonly viewed as one supportive structure situated on a continuum ([Downey, 2003; Parsloe and Leedham, 2009; Passmore and Sinclair, 2020](#)) and often identified collectively as coaching. Quality mentoring or coaching programs are essential to support early career teachers ([Mercieca, 2018; Riddle, 2015](#)), as these programs build teacher and teaching capacity ([Barkley and Bianco, 2010; Knight, 2009](#)). Frequently fixated on the deficiencies of beginning teachers ([Martinez, 1994; März and Kelchtermans, 2020](#)), educators, schools, and authorities need to identify and build on their strengths.

Coaching and mentoring as strength-based approaches both seek to support the professional learning of others through a one-to-one conversation and are anchored in a collaborative and reflective relationship where the focus is on personal and professional growth. Various professions and organisations view both differently and definitions are often blurred. Mentoring programs are designed to support early career teachers ([Ingersoll and Strong, 2011](#)), however, mentoring is employed differently for varied purposes ([Kemmis et al., 2014](#)). For example, in Australia and Finland, varying models of mentoring exist within induction. [Kemmis et al. \(2014\)](#) found that schools implement mentoring to either supervise, support, or assist teachers in self-development by collaborating with others. In mentoring, a more experienced person guides professional learning through conversations ([van](#)

Nieuwerburgh and Barr, 2017). Both mentoring and coaching aim to support growth through conversations but mentors distinct from coaches, require and draw on their expertise. Effective educational programs differ depending on the context and purposeful intent.

To capture the interconnection of coaching and mentoring, Hollweck (2017) designed a mobius strip (Fig. 1) illustrating an educators first few years of teaching.

As we stressed earlier, it is vital in the initial stages to determine how coaching (or mentoring) is defined in the specific context, and the purpose of program. Teachers are more likely to embrace a program if its purpose is well-defined and the process transparent.

The purpose of mentoring varies widely from developing management skills and supporting career transition to the socialization of new staff (Joo, 2005), while the primary purpose of coaching in education focuses on improving student learning and wellbeing through teacher professional development and self-awareness (Grant, 2006; Van Nieuwerburgh et al., 2021). When the focus is on improved student learning and wellbeing outcomes, selecting either teacher learning, student learning or both is imperative when initiating and evaluating a coaching program. Similarly, for measured effectiveness, improved wellbeing needs to target either teachers, students, or both. Coaching also has other benefits to teachers and schools. It can improve teachers' communication skills, develop deeper engagement in school improvement, and advance collaboration among staff, or support a more inquiry approach into their practice.

According to Joo (2005), another difference between coaching and mentoring is coaches frequently use structured processes, strategies, tools and techniques (Cox et al., 2014; van Nieuwerburgh, 2014). Schools or systems of schools (e.g., school districts, networks or clusters) often design processes and templates for their specific context and purpose and use various frameworks that initiate cycles of improvement or scaffold conversations. In contrast, mentoring is often an individual request or an acquisition of knowledge where the results are less tangible and more personal.

In teacher development, as Hollweck and Lofthouse (2021) note that in education there is no one right way to coach. They suggest that customizing to the learning environment and contributing further to the context, is critical in designing effective coaching programs. Aligning a purposeful evaluation process as part of coaching programs provides data to inform, improve and refine the practice. Embedding a coaching program requires school funding to train coaches and schedule coaching sessions.

Subsequently, funded initiatives often demand evidence to illustrate the benefits gained from the investment. Specific tools such as teacher surveys or student achievement offer consistent and current feedback. As an embedded coaching program becomes an iterative process, ongoing evaluation, check-ins, and feedback are necessary. Consistent change occurs in schools with mandated



Fig. 1 Mentoring and coaching during teacher induction. The interconnection of coaching and mentoring by (Hollweck, 2017). Reprinted from Hollweck, T. (2017) Threading the needle: examining the mentoring and coaching fellowship of the western Quebec school board. In: Kutsyruba, B., Walker, K.D (eds.), *The Bliss and Blisters of Early Career Teaching: A Pan-Canadian perspective*. pp. 205–226, Word & Deed Publishing. Burlington Ontario.

requirements, curriculum modifications, staff turnover, and the growth and learning progression of all learners, including staff and students. Therefore, the coaching program should adapt and evolve to the contextual needs.

While mentors draw on their prior professional expertise, coaches need to be cautious of providing expert advice, as this often confuses the relationship and shifts the power dynamics and stance. In a more directive stance, a mentor supports through knowledge transfer, as the more experienced person. In a less directive stance, [van Nieuwerburgh \(2012\)](#) argues that a coach supports the person to develop self-awareness and responsibility. There is a difference between taking responsibility and taking advice. [Whitmore \(2009\)](#), p. 37 argues that “If advice is given by the coach and it fails, blame may be placed on the coach.” The coach’s advice is exchanged for the coachee’s responsibility, and this is not an optimal outcome. Furthermore, if the coachee blames themselves for the failure, self-efficacy may be affected. However, both stances and approaches are appropriate, as the early career teacher may be unconsciously incompetent. At times, beginning teachers don’t know, what they don’t know.

As a powerful conversational intervention, coaching supports teachers to accomplish better results for the organisations and for themselves ([van Nieuwerburgh, 2014](#)). [Cox \(2016\)](#) notes that a coach aspires to build on strengths and broaden perspectives. Additional professional learning is an anticipated precedent for all beginning teachers, but schools also need to tap into the resources beginning teachers possess, as they bring additional strengths to the school context.

Depending on the purpose, coaching can have different forms, for example cognitive, executive, psychological, solution-focused, facilitative, and instructional. In contemporary theory and praxis, coaching is often classified into three conversational approaches, facilitative, instructive, and directive. Although there are no commonly agreed definitions of coaching ([Knight, 2018](#); [Ryan and Deci, 2019](#); [van Nieuwerburgh, 2014](#)), some common threads emerge when observing various definitions used in literature. For the purpose of this chapter, the following definition of coaching can be derived from the current literature and practice:

Coaching is a confidential one-to-one collaborative process that focuses on developing learning and wellbeing through increasing knowledge, skills, and self-awareness, while providing a sense of choice, competence, and connection.

Available knowledge base of coaching in school context reveals nuances that are relevant to the context of teacher development. For example [Whitmore \(2009\)](#) emphasizes the concept of coaching as unlocking people’s potential, while [van Nieuwerburgh \(2012\)](#) highlights the coaching skills of questioning and actively listening, as well as the importance of teachers’ self-directed learning. If this learning is self-directed, the coach is the decision maker and has agency. [Tschannnen-Moran and Tschannnen-Moran \(2011\)](#) describes coaching as a way of being that evokes motivation and actions in people through conversation. Teacher autonomy and professionalism are paramount in effective coaching for teacher development.

Based on early literature on coaching (e.g., [Rogers, 1980](#)) coaches require attributes such as humility, respect, integrity, care, and confidence in one’s ability to coach, and confidence in the potential of the coachee ([van Nieuwerburgh, 2014](#)). The coaching way of being ([van Nieuwerburgh, 2014](#)) aligns well with the philosophy of the Partnership Principals ([Knight, 2011](#)) that include attributes that promote effective interaction, such as equality, choice, voice, dialogue, reflection, praxis and reciprocity. [Stringer \(2020\)](#) notes that through coaching, reciprocity of learning occurred for both the coach and coachee. Collectively, these dispositions of a coach are essential for an effective coaching relationship.

In the coaching way of being, the coach needs to believe in the teacher’s ability to achieve the goal ([van Nieuwerburgh et al., 2019](#)). By extending this to include the coachee’s self-efficacy or sense of competence, reciprocity of belief in achieving the goal is then accomplished ([Ryan and Deci, 2019](#)). Coaching encompasses a focused conversation drawing on various coaching skills such as active listening, skillful questioning, paraphrasing, and summarizing and providing feedback ([van Nieuwerburgh et al., 2019](#)). Training to develop these skills, knowledge and characteristics is favourable. Educators with a shared understanding of coaching comprehend the expectations, the process, and the personal and professional benefits for them and their school community.

Finally, coaching programs can be formal or informal and may incorporate contracts or agreements, coaching frameworks, allocated time, coaching qualifications, and other supportive documentation. The coaching frameworks that schools often adopt identify the type of coaching embedded within the school and provide information about the coaching purpose and practice. Again, the importance of purpose, implementation and practice come into essential play, as the stance within a coaching cycle may transfer from more directive to less directive, depending on the knowledge, skills, and experience the beginning teacher brings to the conversation. When coaching early career teachers, there are times when it is required to share expertise and knowledge. Instead of telling the early career teachers what works best, suggestions made may be accepted or rejected, implemented or not. Intrinsic motivation, that is a cornerstone of school cultures in Finland ([Sahlberg and Walker, 2021](#)), for example, can be sustained when suggestions or options originate from the early career teacher, while still meeting their professional or personal needs.

Often the professional needs of an early career teacher may suit the well-organized framework of teacher or instructional coaching. This draws on [Bandura’s \(1989\)](#) theory of observational learning to improve educators’ self-perceptions of competence by watching instructional delivery. Identifying the purpose, positioning coaching in the school’s strategic plan and allocating provisions is pivotal when embedding a coaching program, as it provides clarity. The alignment of the coaching purpose, implementation and practice should be transparent and understood by all stakeholders. The program’s effectiveness is evaluated based on the purpose and practice of coaching, and these data provides a baseline for this reiterative process. Coaching programs should not be identical but contextual; as all contexts differ, learning environments change, and all learners’ needs consistently vary.

Australian case study: coaching early career teachers

The power of coaching lies in its potential to provide professional and personal support to teachers in their early careers. By understanding specific coaching practices applied within a school, leaders and coaches can adopt, adapt, or even disregard techniques, frameworks, and tools according to their context and professional needs. While advancing conceptual and practical coaching knowledge in the school context, it is important to acknowledge the influential factors that motivate and nurture early career teachers as self-determined learners. Coaching supports a sense of choice, competence and connectedness, and self-determined learners grow and flourish when their psychological needs are met (Ryan and Deci, 2019). As processes and conditions can undermine or facilitate self-motivation, the learning environment or context is pivotal.

We recently conducted a study in New South Wales in Australia to find out how to support early career teachers for the development of their professional knowledge and skills in teaching. In addition to professional growth, the support incorporates the best approach to sustain and nurture teachers' motivation levels, while providing a sense of choice, competence, and belonging. More specifically, our study examined how the purpose, implementation, and practice of coaching that we discussed above influenced early career teachers' professional growth that included both their professional learning and wellbeing. Using qualitative methodologies, we investigated how different people perceive coaching in schools, how coaching is organized and embedded within each school, and what or who motivates early career teachers to continue to learn and grow professionally.

Just like in many other countries, not much is known about the role of coaching in Australian schools. While Mansfield and Gu (2019) conducted research in Western Australia to investigate the professional learning activities of early career teachers, and the perceive impact of these on their professional development, our study explored coaching early career teachers for professional growth. The aim of our study was to advance the conceptual and practical knowledge of coaching by interpreting the perception, implementation, and practice of coaching early career teachers for professional growth.

The shift from mentoring to coaching early career teachers for professional learning transfers the stance from more directive to less directive. Many view coaching as a dialogical partnership and sharing collegial experience for professional growth and impact. The coach does not give advice but offers options and choices, nourishing autonomy.

We included three different purposefully selected schools in this case study to portray differences and similarities, existing in Australian education sectors, namely Independent, Catholic, and government school sectors. The diversity across these three schools provides an opportunity to learn about the complexities of implementing coaching in various schools. The school location, sector in which it belongs, and level of schooling are the identified variables in our study.

The range of variation regarding the coaching program and how they were implemented in different schools was unknown from the outset. The composition of the three schools consisted of one secondary (Years 7–12) and two K-12 schools located in rural, suburban, and urban settings in New South Wales, Australia. These schools provide an in-depth description of how the purpose, implementation and practice of coaching influence the professional growth of beginning teachers. Here is an overview of the initial findings in this study.

Purpose of coaching is clear but is not the same among school

The schools in this case study explained that coaching focused on either learning or wellbeing, whether it be teacher or student, and could be a combination of both. Coaching outcomes contributed to the positive learning environment that incorporated a belief in educators, research, and coaching, while being committed to a collaborative culture, where learning and adapting to challenges developed. Argenta School (school names are changed) introduced teacher coaching with the goal of supporting teacher learning and wellbeing. It has since expanded the focus to include teachers coaching students. The capacity for teachers to coach students has seen an amplification in this school. Therefore, an increase in the allocation and training of coaches for students resulted. However, the allocation of one coach for teachers remains unchanged.

At Argenta, many teachers are trained as coaches and take a "coaching approach" to their conversations around professional development plans and yearly goals. This is not a formalized coaching session per se, but a coaching approach that utilizes transferable elements of coaching (Campbell and van Nieuwerburgh, 2018). The principal of the Argenta calls coaching a "bit of an octopus" with many legs, and it "sort of evolves." She maintained that the coaching program is responsive, and you must be strategic about what is working and what is not working. Purely optional, the teacher-coach provides formalized coaching, while other trained coaches take an informal approach to determine yearly goals for professional learning plans. Other trained coaches work formally with students on their learning goals.

The second school that we call Banksia initially focused on teacher learning, yet extended its coaching program to include teacher wellbeing and made coaching available to all non-teaching staff employed by the school. The coach at the third school, Columbus, focused solely on teacher learning with the goal of improving student outcomes. While the program has evolved, the interviews revealed that in the future, the purpose might include teachers coaching peers, and coaching to support teacher wellbeing. All three schools heightened the implementation of coaching based on feedback provided from school-initiated surveys or focus groups or both. With the increased investment, the purpose of coaching extended to either include additional focus areas or funded more release time for coaching and training.

Principals position coaching in the strategic plan so that it typically lies with the leadership and coaching team. All three principals had utilized and advised that research is needed. To understand the purpose of implementing coaching, principals asked what

problem they were trying to solve and if coaching could be the answer. Consequently, the coaches and teachers needed to comprehend and exemplify the purpose in similar ways. The Columbus principal claimed, “If teachers do not feel they need coaching, it will take a long time get off the ground”. Argenta School principal stated, that “You can’t have flourishing students without flourishing teachers”. The Columbus principal proclaimed, that “Coaching empowers teachers to accelerate and amplify their impact on student learning”. Common views among these school leaders were that coaching makes good teachers into great teachers and harnesses the experience of others to leverage their development in particular areas of perceived need.

Comparatively, the purpose of coaching varies in these three schools, yet each school had a program with a clear purpose. We found that each coaching program was contextually designed around what was needed, and the needs evolved. Coaching serves multiple purposes, including teacher learning, teacher wellbeing, student learning, student wellbeing, while also influencing the school culture. How a school implements coaching may change and adapt due to the meaningful feedback from teachers, coaches and even students.

Coaching is an investment to professional capital in school

It was evident that principals had considerable influence over the implementation of a coaching program. For coaching to succeed, allocating funds for school embedded professional learning is pivotal. Expenses associated with implementing a coaching program include training teachers to become coaches and allocated release time for coaches and teachers. An overwhelming barrier for many schools is manipulating timetables to afford the time to coach. There is a cost to replace teachers from face-to-face teaching for either training or coaching sessions. The availability of casual teachers is also an issue. Making provisions for training coaches requires an allocation of time and funding, whether delivered in-house or by an outside commercial provider. The assumption that an expert teacher equates to an expert coach is a concern. All schools shared the importance of training coaches and using criteria for selecting coaches, as an effective coach requires interpersonal and professional skills, and associated dispositions. Coaching incorporated into the strategic plan elevates coaching into a high priority position.

Two of the three schools emphasized the importance of including the coaching program into the strategic plan, and when designated as a priority, coaching becomes strategically resourced. The principal from Columbus School shared that although their priorities are strategically resourced, a strategic plan is not required. He viewed strategic plans as too slow for the current era, and stated, “With a commitment to entrepreneurial thinking, our approach is to ensure the vision is well known and that decisions about learning and care are future thinking focused, data informed, and action research based”.

Typically, when integrated into the strategic plan, the principal allocates funds and establishes an evaluation process to determine the program’s effectiveness. In all schools, the evaluation process evolved to reflect a coaching program that continually developed to reflect the contextual needs and recent research. Some may view the evaluation as a form of accountability, driven by the system. Nevertheless, the principal and leadership team believed in the efficacy of the coaching program for the investment was significant, and the objectives long-term. The importance of how coaching builds on or supports already established programs was evident. The three schools did not align with one commercial program or coaching philosophy, with each program being agile, bespoke, and contextual to meet the individual school’s needs.

While perceived as a worthwhile investment, coaching demands school time and funding. Principals, coaches, and teachers all commented on the scarcity of time. This challenge is a constant tension not only with coaching but with everything that goes on in schools. The coaching process requires allocated time for coaches and teachers within busy school days. Columbus School principal stated, “Time is the greatest investment we can make in all our teachers”, and he schedules time to participate in coaching conversations. He argued that if coaching is going to be a priority, principals need to be a part of that priority, and it is one way of “connecting with what is happening at the grassroots”. Banksia School principal stated, “We’re sitting within the body of research that suggests equipping teachers is vital, and coaching teachers is a valuable instrument in that purpose”. A critical component in this coaching investment is the time the coaching program takes to mature. All three schools had embedded coaching for at least two years, while most schools for many more.

All the schools declared their coaches received training. Two schools sought training from an outside provider and explained the importance of consistency. However, both lead coaches in those schools continue to personally upskill their coaches to extend their knowledge, practice, and skills. Columbus welcomed two staff members who were previously trained over numerous years as instructional leaders. Their knowledge and experience contributed to the expansion of the coaching program. Along with the principal, these staff members delivered professional learning by facilitating an in-house coaching program that drew on their experiences and research. This course reflected the importance of consistency and clarification around the purpose and practice of coaching. These provisions reflect the purpose of coaching, with training being fundamental to the implementation of coaching that allows for coaching to be put into action.

Coaching in action

Argenta School had a less directive stance in coaching, also known as a facilitative approach to coaching. This facilitative approach places the goal purely at the coachee’s discretion. The sole teacher-coach relies on coaching expertise more than teaching expertise.

When examining the collected data on coaching goals, the teacher-coach recently developed three themes: teachers' effectiveness in their role, a future career focus, and work-life balance.

During the Australian bushfires and floods, and the disruption of teaching and learning due to COVID-19, the overwhelming theme in teacher development was work-life balance. According to the teacher-coach, a teacher shared that she "feels as if coaching is a gift, she is giving herself". The coach explained that she denied a couple of teachers who had left the school for various reasons as, they "actually asked if they could still be coached". She expressed that she believes that coaching, long term, would have an impact on retaining teachers at the school. While the purpose of coaching at Argenta was about the "whole person", the teacher-coach is interested in moving from a purely facilitative approach to a more instructional or teacher approach, as she views it as a blend of mentoring and coaching. The teacher-coach articulated her awareness of the expenses associated when adding another component to coaching and questioned the possibility.

The Banksia School did not classify the coaching program as instructional and preferred their coaches to be recognized as GROWTH coaches, who take a less directive stance. A facilitative coach does not need expertise in the area when coaching. Nevertheless, the school identifies their coaches as expert teachers. They have coaching that focuses on teaching and learning while providing their staff autonomy throughout the coaching cycle. Coaching is mandatory for all early career teachers for one year, with time allocation provided by the school. Some may identify tension in the idea of the directive stance of expert teachers versus the less directive stance of an expert coach.

Banksia School exemplified the coaching continuum. When needed, the expert teacher shared their expertise by providing options, yet holds judgment. This school has strategically embedded coaching for more than five years, with the program evolving throughout the years. It also highlighted the importance of staff wellbeing and the learning environment with the lead coach sharing how they support and deeply care for their staff and invest in coaching because they know what works. The Banksia principal discussed losing beginning teachers and how the prevailing perception is due to pay and working conditions. He said, "Often, in the past, it's been sink or swim" and a "lack of in-school support for emerging teachers". According to the lead coach in that school,

We know it [coaching] doesn't just have impact on student outcomes, but it has an impact on your wellbeing, and has impact on your professional expertise, and your professional growth. And it also allows you to then have the confidence to contribute back to our community in a professional manner.

A teacher leaving Banksia School expressed to her coach that "it's a shame I can take you with me. I would really love to be able to keep doing this [coaching]". The coach voiced how teaching is a "long haul", and that we are witnessing more teachers drop out early in their careers. She shared her wonder of whether a fortnightly coaching conversation could help retain more employees while also benefiting students.

The Columbus School also had a coaching program that evolved, but the purpose has not changed. According to one coach, the principal is genuinely committed to the coaching model and resources and has a sharp vision and understanding of learning for all. The power in this coaching model lies within the responsive approach by the leaders and the inclusion of all voices in decision making. At Columbus, coaching is mandatory, and the leaders declared that coaching is a huge part of the school. It is purely focused on teaching and learning and concentrates on data and evidence to demonstrate its effectiveness.

At Columbus, there appeared to be some tension between the need to focus purely on teacher instruction. One coach shared that while we are all working toward the same goal of improved student outcomes and teacher practice, a teacher is more than their instructional routines. It was articulated that coaching could include goals such as developing organization and conversational skills. Another coach maintained that the difference between a coach and an instructional leader is that coaching establishes relational trust and identifies teachers' strengths first. She expanded on her comments maintaining that coaching involves the heart, with practices such as empathy and trust. Motivated by wanting teachers to stay in this profession, this educator found her "niche" in coaching. While the program has evolved, the interviews revealed that the purpose might someday include teachers coaching peers and coaching to support teacher wellbeing.

In each school, whether the goal is learning or wellbeing, the coaching process identified components within frameworks or models that provided structure and guidance while ensuring reflection. The coach's dispositions or "way of being" were identified characteristics demonstrated by the coach to build a trusting relationship. Faith in others, confidence and experience, ability to care for and develop others, approachability, empathy, gratitude, humility, and integrity were highlighted as characteristics of the coaches. The learning environment appears to strongly influence the experience and expectations of beginning teachers. The coaching programs reflected the concept of contextual coaching (see [Valentine, 2019](#)).

Lessons for teacher development through coaching

Learning new teaching strategies is an essential part of professional growth of early career teachers. As the evidence from the three schools above shows, coaching is more about stepping back and looking at the bigger picture, as beginning teachers often get caught up in the day-to-day of teaching. Early career teachers find coaching sessions helpful, they understand the importance of peer support and coaching, and how reflective processes strengthen collective autonomy at school. When teachers engage in coaching

voluntarily, their motive is autonomous. If we want beginning teachers to grow and learn professionally, then there is a need to view and support them as self-determined learners.

For the early career teachers, their students were a strong motivational influence, along with their colleagues, especially within their own department. Many of these teachers think they themselves are the primary source of motivation, especially those for whom teaching was not their first career. Most of the beginning teachers in our case study shared how the first few years are a combination of proving and improving yourself as a teacher. If early career teachers fail to be intrinsically motivated, it increases possibility of teaching becoming an unsatisfying profession. Working as a temporary teacher influenced the need to prove themselves, with one beginning teacher sharing that she was all about proving herself, her worth, and her value. Reassurance and recognition from coaches and colleagues, and strengths identified built early career teachers' confidence and competence.

All stakeholders in the professional learning process must hear the voices of the beginning teachers when determining the role of coaching for professional growth. Early career teachers mostly viewed the purpose of coaching as a process that improved their teaching and learning. Their motivation was strongly influenced by a positive learning environment that provided coaching.

Professional growth was achieved through a coaching partnership that was based on trust and required constant reflection. Within this dialogical partnership, the coach provided emotional support to set a clear path to achieve the self-determined goal that can be critically important regarding teacher retention. Coaching empowers individuals to develop positive and optimistic situations.

With the alignment of purpose, implementation, and practice, it is necessary to determine, develop, or borrow tools, models, and frameworks. The schools included in our case study produced several contextual resources by incorporating or adapting various available frameworks and altered some tools, templates, and processes to suit the ever evolving or expanding needs of their school. The coaching approach reflects the purpose and the implementation and practice follow. Coaching requires teachers to have additional skills and dispositions that is reflected in selection criteria. Confidentiality and trust were considered to be vital in coaching.

So, does coaching help teachers and schools to do better? While the schools struggle to show evidence of the impact of coaching on student learning, all consistently endeavor to achieve this. There appears to be tension between the type of data collected, as some strive for the elusive quantitative data. One participant stated that the impact of coaching on student learning is not always evident and not as noticeable as the impact on teacher practice. The theory of action reasons that improving teacher practice will improve student learning. For each school, external validation through surveys, school visits, reports and university partnerships became a vital component to confirm coaching effectiveness within their context. While each school has processes in place to determine the effectiveness of coaching, all aligned with their purpose.

Overall, the purpose of implementing coaching is often about student and teacher learning or wellbeing, or both. Well-designed and resourced coaching supports teachers' professional development and contributes to the positive learning environment. Notably, all the schools mentioned the importance of developing, retaining, and supporting beginning teachers and the positive contribution effective coaching makes to education.

References

- AITSL, 2016. Graduate to Proficient: Australian Guidelines for Teacher Induction into the Profession. <https://www.aitsl.edu.au/tools-resources/resource/graduate-to-proficient-australian-guidelines-for-teacher-induction-into-the-profession>.
- Bandura, A., 1989. Human agency in social cognitive theory. *Am. Psychol.* 44 (9), 1175–1185.
- Barkley, S., Bianco, T., 2010. *Quality Teaching in a Culture of Coaching*, second ed. Rowman & Littlefield Education.
- Buchanan, J., Prescott, A., Schuck, S., Aubusson, P., Burke, P., Louviere, J., 2013. Teacher retention and attrition: views of early career teachers. *Aust. J. Teach. Educ.* 38 (3).
- Campbell, J., van Nieuwerburgh, C., 2018. *The Leader's Guide to Coaching in Schools: Creating the Conditions for Effective Learning*. Corwin.
- Carnegie School of Education, 2022. CollectivED – The Centre for Mentoring, Coaching and Professional Learning. <https://www.leedsbeckett.ac.uk/research/collectived/>.
- Costa, A.L., Garmston, R.J., 2015. *Cognitive Coaching: Developing Self-Directed Leaders and Learners* (Hawker Brownlow).
- Cox, E., 2016. Working with emotions in coaching. In: Bachkirova, T., Spence, G., Drake, D. (Eds.), *The SAGE Handbook of Coaching*. SAGE Publications.
- Cox, E., Bachkirova, T., Clutterbuck, D., 2014. The complete handbook of coaching: introduction. In: Cox, E., Bachkirova, T., Clutterbuck, D. (Eds.), *The Complete Handbook of Coaching*, second ed. Sage Publications, pp. 1–18.
- Craven, G., Beswick, K., Fleming, J., Fletcher, T., Green, M., Jensen, B., Leinonen, E., Rickards, F., 2014. Action Now: Classroom Ready Teachers. [https://www.aitsl.edu.au/docs/default-source/default-document-library/action_now_classroom_ready_teachers_accessible-\(1\)da178891b1e86477b58fff00006709da.pdf?sfvrsn=9bffc3c_0](https://www.aitsl.edu.au/docs/default-source/default-document-library/action_now_classroom_ready_teachers_accessible-(1)da178891b1e86477b58fff00006709da.pdf?sfvrsn=9bffc3c_0).
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A.L., Hammerness, K., Low, E.-L., McIntyre, A., Sato, M., Zeichner, K., 2017. *Empowered Educators: How High-Performing Systems Shape Teaching Quality Around the World*. Jossey Bass, San Francisco.
- Downey, M., 2003. *Effective Coaching: Lessons from the Coaches' Coach*, second ed. Thompson.
- Gallant, A., Riley, P., 2014. Early career teacher attrition: new thoughts on an intractable problem. *Teach. Dev.* 18 (4), 562–580.
- Glazzard, J., Coverdale, L., 2018. "It feels like its sink or swim": newly qualified teachers' experiences of their induction year. *Int. J. Learn. Teach. Educ. Res.* <https://doi.org/10.26803/ijlter.17.11.7>.
- Grant, A., O'Connor, S., 2018. Broadening and building solution-focused coaching: feeling good is not enough. *Coaching* 11 (2), 165–185. <https://doi.org/10.1080/17521882.2018.1489868>.
- Grant, A.M., 2006. A personal perspective on professional coaching and the development of coaching psychology. *Int. Coach. Psychol. Rev.* 1 (1), 12–22.
- Grant, A.M., 2011. Is it time to REGROW the GROW model? Issues related to teaching coaching session structures. *Coach. Psychol.* 7 (2), 118–126.
- Grant, A.M., 2014. Autonomy support, relationship satisfaction and goal focus in the coach–coachee relationship: which best predicts coaching success? [Article]. *Coaching* 7 (1), 18–38. <https://doi.org/10.1080/17521882.2013.850106>.
- Grant, A.M., 20. What constitutes evidence-based coaching? A two-by-two framework for distinguishing strong from weak evidence for coaching. *Int. J. Evid. Based Coach. Mentor.* 14, 74–85.
- Grant, A.M., Atad, O.I., 2021. Coaching psychology interventions vs. positive psychology interventions: the measurable benefits of a coaching relationship. *J. Posit. Psychol.* 1–13. <https://doi.org/10.1080/17439760.2021.1871944>.

- Gu, Q., Day, C., 2013. Challenges to teacher resilience: conditions count. *Br. Educ. Res. J.* 39 (1), 22–44. <http://www.jstor.org/stable/24464800>.
- Hay-Group, 2014. Building the Right Foundation - Improving Teacher Induction in Australian Schools. AITSL, Melbourne.
- Hollweck, T., 2017. Threading the needle: examining the mentoring and coaching fellowship of the western Quebec school board. In: Kutsyuruba, B., Walker, K.D (Eds.), *The Bliss and Blisters of Early Career Teaching: A Pan-Canadian perspective*. Word & Deed Publishing. Burlington Ontario, pp. 205–226.
- Hollweck, T., Lofthouse, R.M., 2021. Contextual coaching: leveraging and leading school improvement through collaborative professionalism. *Int. J. Mentor. Coach. Educ.* 10, 399–417. <https://doi.org/10.1108/ijmce-01-2021-0019>.
- Howe, E.R., 2006. Exemplary teacher induction: an international review. *Educ. Philos. Theor.* 38 (3), 287–297. <https://doi.org/10.1111/j.1469-5812.2006.00195.x>.
- Ingersoll, R., Smith, T., 2004. Do teacher induction and mentoring matter? *NASSP Bull.* 88 (638), 28–39.
- Ingersoll, R.M., Strong, M., 2011. The impact of induction and mentoring programs for beginning teachers. *Rev. Educ. Res.* 81 (2), 201–233. <https://doi.org/10.3102/0034654311403323>.
- Joo, B.-K., 2005. Executive coaching: a conceptual framework from an integrative review of practice and research. *Hum. Resour. Dev. Rev.* 4 (4), 462–488. <https://doi.org/10.1177/1534484305280866>.
- Kemmis, S., Heikkinen, H.L.T., Fransson, G., Aspors, J., Edwards-Groves, C., 2014. Mentoring of new teachers as a contested practice: supervision, support and collaborative self-development. *Teach. Teach. Educ.* 43, 154–164. <https://doi.org/10.1016/j.tate.2014.07.001>.
- Knight, J., 2007. *Instructional Coaching*. Corwin Press.
- Knight, J., 2009. In: Knight, J. (Ed.), *Coaching: Approaches and Perspectives*. Corwin Press.
- Knight, J., 2011. *Unmistakable Impact: A Partnership Approach to Dramatically Improving Instruction*. Corwin.
- Knight, J., 2018. *The Impact Cycle: What Instructional Coaches Should Do to Foster Powerful Improvements in Teaching*. Corwin Press, Thousand Oaks.
- Knight, J., 2019. Instructional coaching for implementing visible learning: a model for translating research into practice. *Educ. Sci.* 9 (2), 101. <https://doi.org/10.3390/educsci9020101>.
- Kraft, M.A., Blazar, D., Hogan, D., 2018. The effect of teacher coaching on instruction and achievement: a meta-analysis of the causal evidence. *Rev. Educ. Res.* 88 (4), 547–588. <https://doi.org/10.3102/0034654318759268>.
- Lofthouse, R., 2019a. Coaching in education: a professional development process in formation. *Prof. Dev. Educ.* 45 (1), 33–45. <https://doi.org/10.1080/19415257.2018.1529611>.
- Lofthouse, R., 2019b. Using Mentoring and Coaching to Focus on the Curriculum in Action. <https://www.bera.ac.uk/blog/using-mentoring-and-coaching-to-focus-on-the-curriculum-in-action>.
- Lofthouse, R., Hall, E., 2014. Developing practices in teachers' professional dialogue in England: using Coaching Dimensions as an epistemic tool. *Prof. Dev. Educ.* 40 (5), 758–778. <https://doi.org/10.1080/19415257.2014.886283>.
- Lofthouse, R., Leat, D., 2013. An activity theory perspective on peer coaching. *Int. J. Mentor. Coach. Educ.* 2 (1), 8–20.
- Lofthouse, R., Leat, D., Towler, C., 2010. *Coaching for Teaching and Learning: A Practical Guide for Schools*.
- Long, J.S., McKenzie-Robblee, S., Schaefer, L., Steeves, P., Wnuk, S., Pinnegar, E., Clandinin, D.J., 2012. Literature review on induction and mentoring related to early career teacher attrition and retention. *Mentor. Tutoring: Partnersh. Learn.* 20 (1), 7–26. <https://doi.org/10.1080/13611267.2012.645598>.
- Mansfield, C., Gu, Q., 2019. "I'm finally getting that help that I needed": early career teacher induction and professional learning. *Aust. Educ. Res.* <https://doi.org/10.1007/s13384-019-00338-y>.
- Martinez, K., 1994. Teacher induction revisited. *Aust. J. Educ.* 38 (2), 174–188. <https://doi.org/10.1177/000494419403800207>.
- März, V., Kelchtermans, G., 2020. The networking teacher in action: a qualitative analysis of early career teachers' induction process. *Teach. Teach. Educ.* 87, 102933. <https://doi.org/10.1016/j.tate.2019.102933>.
- Mercieca, B., 2018. Want to keep early career teachers in the job? Do this. *AARE EduRes. Matters.* <https://www.aare.edu.au/blog/?tag=mentoring-early-career-teachers>.
- OECD, 2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*. OECD, Paris.
- OECD, 2018. Chapter 4: Attracting and Effectively Preparing Candidates. *Teaching and Learning International Survey (TALIS) Report, Issue.* https://www.oecd-ilibrary.org/sites/1d0bc92a-en/1/2/5/index.html?itemid=/content/publication/1d0bc92a-en&_csp_=1418ec5a16ddb9919c5bc207486a271c&itemID=0ecd&itemContentType=book.
- OECD, 2019. *Supporting and Guiding Novice Teachers: Evidence from TALIS 2018*. OECD Publishing. <https://doi.org/10.1787/fe6c9c0c-en>. *Teaching in Focus*, No. 29.
- Paniagua, A., Sánchez-Martí, A., 2018. Early Career Teachers: Pioneers Triggering Innovation or Compliant Professionals?. In: *OECD Education Working Papers*, vol. 190, pp. 0_1–52. <https://doi.org/10.1787/4a7043f9-en> (190).
- Parsloe, E., Leedham, M., 2009. *Coaching and Mentoring: Practical Conversations to Improve Learning*. Kogan Page Publishers.
- Passmore, J., Sinclair, T., 2020. *What Is Coaching?* Springer International Publishing, pp. 7–13. https://doi.org/10.1007/978-3-030-53161-4_2.
- Ramsey, G., 2000. *Quality Matters: Revisiting Teaching: Critical Times, Critical Choices*. Report of the review of teacher education in New South Wales, Sydney. <https://www.det.nsw.edu.au/teachrev/reports/reports.pdf>.
- Riddle, S., 2015. Fears of Looming Teacher Shortage as Student Population Soars. <http://www.abc.net.au/news/2016-01-18/fears-of-looming-teacher-shortage-as-student-population-soars/7096102>.
- Rogers, C., 1980. *A Way of Being*. Houghton-Mifflin.
- Ryan, R.M., Deci, E.L., 2019. Supporting autonomy, competence, and relatedness: the coaching process from a self-determination theory perspective. In: English, S., Sabatine, J., Brownell, P. (Eds.), *Professional Coaching: Principles and Practice*. Springer Publishing Company, pp. 231–246.
- Sahlberg, P., Walker, T., 2021. In: *Teachers We Trust: The Finnish Way to World-Class Schools*. W.W. Norton and Company.
- Sharratt, L., Fullan, M., 2012. *Putting FACES on the Data: What Great Leaders Do!* Corwin Press, Thousand Oaks.
- Showers, B., Joyce, B., 1996. The evolution of peer coaching. *Educ. Leader* 53, 12–16.
- Stringer, A., 2020. *Coaching: A Reciprocal Learning Partnership. A Practice Insight Working Paper*. In: *CollectiveEd Working Papers*, pp. 47–54 (11).
- Tschannnen-Moran, B., Tschannnen-Moran, M., 2011. *Evocative Coaching: Transforming Schools One Conversation at a Time*. Jossey-Bass.
- University of Kansas Center for Research on Learning, 2022. *Instructional Coaching Group*. <https://www.instructionalcoaching.com/research/>.
- Valentine, M., 2019. Contextual coaching: leveraging context for alignment in the system. *PhilosoCoach: Int. J.* 4 (1), 93–106. <https://doi.org/10.22316/poc/04.1.07>.
- van Nieuwerburgh, C., 2012. *Coaching in Education: Getting Better Results for Students, Educators and Parents*. Karnac.
- van Nieuwerburgh, C., 2014. *An Introduction to Coaching Skills: A Practical Guide*. Sage.
- van Nieuwerburgh, C., Barr, M., 2017. *Coaching in education*. In: *The SAGE Handbook of Coaching*. Sage Publishers, pp. 505–520.
- van Nieuwerburgh, C., Barr, M., Fouracres, A.J.S., Moin, T., Brown, C., Holden, C., Lucey, C., Thomas, P., 2021. Experience of positive psychology coaching while working from home during the COVID-19 pandemic: an Interpretative Phenomenological Analysis. *Coaching* 1–18. <https://doi.org/10.1080/17521882.2021.1897637>.
- van Nieuwerburgh, C., Biswas-Diener, R., 2021. Positive psychology approach. In: Passmor, J. (Ed.), *The Coaches' handbook: The Complete Practitioner Guide for Professional Coaches*. Routledge, pp. 314–321.
- van Nieuwerburgh, C., Knight, J., Campbell, J., 2019. *Coaching in education*. In: English, S., Sabatine, J., Brownell, P. (Eds.), *Professional Coaching: Principles and Practice*. Springer Publishing Company.
- Varah, L.J., Theune, W.S., Parker, L., 1986. Beginning teachers: sink or swim? *J. Teach. Educ.* 37 (1), 30–34. <https://doi.org/10.1177/002248718603700107>.
- Whitmore, J., 2009. *Coaching for Performance: Growing Human Potential and Purpose*. The Principles and Practice of Coaching and Leadership, fourth ed. Nicholas Brealey Publishing.

Early career teaching and resilience

Anna Lise Gordon, St Mary's University, Twickenham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	153
Resilience	153
Purposefulness	154
Confidence	155
Social support	155
Agency	156
Culture	157
Concluding reflections	158
References	159

Introduction

The complexity of teaching makes it both a challenging and rewarding profession (Kincheloe, 2008) so everyone connected with early career teachers has a part to play in inspiring “a sense of confidence, resilience, and self-efficacy” (Waddell, 2007: 125). My own research (Gordon, 2020a) focused on how initial teacher education provision might foster resilience among early career teachers. Clearly, this is only ever a starting point on the professional journey and should contribute to career development pathways for teachers in the longer-term, as noted by Gu and Day’s (2013: 40) comment that “teacher educators need to become engaged in future thinking with pre-service students on internal and external factors which influence resilience in teachers, so that they are better prepared for the realities of teaching.”

The connection between early career teaching and resilience is of interest to everyone in education as we respond to the global challenges of teacher recruitment and retention (UNESCO, 2017). In their early years in the profession, teachers face high levels of work-related stress (Vesely et al., 2014) and must respond to a range of complexities in the school environment as well as an ever-changing educational context (Howes and Goodman-Delahunty, 2015).

Against this backdrop, resilience is one matter among many pertinent to professional satisfaction, motivation and longevity in the profession. The chapter opens with an overview of literature related to early career teaching and resilience, beginning with definitions of resilience. Five key themes are then explored—purposefulness, confidence, social support, agency and culture—as central to the resilience of early career teachers. Specific examples from England and Australia are given to illustrate key points that may be transferable to other global contexts around the world. The voices of early career teachers appear in italics throughout.

Resilience

There has been increasing literature about teacher resilience, closely linked to wellbeing, ranging from research-informed articles (Gu and Day, 2013; Farrell, 2016) to practical guidance for teachers (Cowley, 2019; Holmes, 2019). Challenges of definition abound with resilience often described simply as “bouncing back” in the face of adversity. Other definitions see resilience as a dual quality, being both “resistant” and “elastic”. Whatever definition is used, there is an acknowledgment that resilience can be learned and developed over time, as it is both “improvable and teachable” (Zolli and Healy, 2012: 4).

Although it is widely accepted that resilience can be developed, there remains considerable uncertainty about the concept. Neenan (2009: 3) refers to resilience as “intriguing yet elusive” and Chandler (2014: 46) describes resilience as “an increasingly ubiquitous concept”, and Thom (2020:14) specifically addresses the school context and the “elusive and underexplored quality of resilience that is the secret ingredient all teachers need in order to survive in our profession.” As highlighted by Grotberg (1995), the transformative potential of resilience as we grow more as a result of life’s experiences is a characteristic of the concept.

Drawing on my own research with early career teachers over many years, I am at ease with the notion of “secret ingredient” but have come to a more nuanced understanding of resilience as a continuum, with toughness and determination at one end, and with vulnerability and openness at the other end. This is in part because different aspects of resilience are required at various stages of the professional journey which are more complex than simply “bouncing back” which could be considered a reductionist definition of resilience. Moreover, for the transformative power of resilience to thrive, it is important to note that a “multi-faceted construct like resilience requires a multi-faceted approach” (Cooper et al., 2013: 95) and an intentional, “deliberate cultivation of resilience” (Lucas et al., 2013: 12). For this reason, all those involved in early career teacher development should be aware of key factors which have the potential to hinder or enhance resilience.

Hargreaves and Fullan (2012: vx) define “professional capital” in teaching as the combination of human, social and decisional capital, which includes resilience as a key dynamic. In this regard, as different forces mediate the professional lives of early career teachers, it is important to recognize that resilience is not a final goal, but something to develop throughout professional lives and it may be more helpful, as Chandler (2014: 47) suggests, to “approach resilience as a way of thinking about how we think about the being of being.”

Early career teachers face “transition trauma” (Johnson et al., 2014: 531) and many professional demands, dealing with complexity and unpredictability on a daily basis. On occasions, this can be overwhelming as the experiences may contrast with their preconceived notions of the role (Kelchtermans, 2017; Sabar, 2004). A growing resilience may be needed to thrive in the profession in the longer term, including through pivotal transition phases in the early stages of the profession (Gordon, 2020b).

Nurturing early career teachers as they enter the profession is essential (Allen and Sims, 2018) as robust education systems around the world require experienced teachers. Attrition among early career teachers is detrimental to the educational enterprise, leading to fewer experienced teachers, including those who may aspire to leadership roles, and reduced stability in the learning environment for children and young people. While Smith and Ulvik’s study (2017) in Norway illustrates that we can recognize proactive choice and agency to leave the profession as a sign of resilience, a lack of experienced teachers (those with more than five years of teaching) will have a negative impact on student outcomes (Day and Gu, 2007). Hence, this chapter explores some of the key factors that nurture resilience to encourage early career teachers to stay and thrive in profession.

There is a need to be cautious, even vigilant, about a possible commodification of resilience as a construct, noting Kincheloe and Berry’s (2004: 98) encouragement to be alert to the “silent and invisible ways that power operates via structures, narratives, and meta-narratives to constrain our understandings of the world.” This chapter does not intend to promote resilience above other teacher qualities but rather to explore the potential impact of resilience on the wellbeing and effectiveness of early career teachers.

Mindful, then, that the literature identifies a number of dimensions affecting teacher resilience, including personal, relational and organizational dimensions (Gu and Day, 2007), this chapter explores purposefulness, confidence, social support, agency and school culture, insofar as they influence the ability of an early career teacher to be resilient. The impact of positive mentoring is noted as a decisive factor and indeed emerges as something of a leitmotif throughout the chapter.

Purposefulness

Central to resilience is a clear sense of purpose which, for many teachers, can provide motivation and passion throughout their professional lives. This notion is echoed in the writing of Zolli and Healy (2012: 7) who describe resilience as “the capacity of a system enterprise, or a person, to maintain its core purpose and integrity in the face of dramatically changed circumstances.”

Almost all participants in my research with early career teachers highlighted a strong sense of purpose, enthusiasm and energy at the start of their teacher education program: “*My life has purpose again!*” was a typical response from an early career teacher. This positive attitude at the outset is perhaps unsurprising among those who have made a deliberate career choice to enter the teaching profession. The key challenge is that of maintaining such positivity over time when some of the felt burdens of teaching can become overwhelming.

As part of my research inquiry, early career teachers engaged in creative writing activities at various stages of their teacher training program (Gordon, 2020a), providing deep insights into the emotional aspects of their experience. This haiku poem, written by one participant at the mid-point in the final school placement, captures many of the sentiments of other early career teachers:

A train on a track
Can’t go back, can I go on?
I’ll do it but how?

The sense of purpose for this early career teacher remains, but is being tested by the relentless movement of the train (and, by inference, the early career teacher) on the journey toward its destination. In some ways, the haiku signals the potential loss of personal voice and agency, as the individual is drawn along by an external, pre-determined force. The feeling of pressure is tangible, as the early career teacher struggles to cope. The development of teaching skills takes time and is never a linear process, posing a challenge to resilience at particular moments. Implicit in this haiku is a tacit acknowledgment, or even an acceptance, of the constant demands and challenges of teaching. There is also a hint of self-sacrifice (Thom, 2020) in the haiku and this may challenge a teacher’s resilience; the early career teacher needs to know when to seek support and guidance. If the environment and support around an early career teacher are positive, progress should be smooth, but there is no sense that these factors are present, and so the nurturing of this particular early career teacher’s resilience is jeopardized. The haiku poem echoes Hargreaves and Fullan’s (2012) suggestion that teachers’ commitment and effectiveness are particularly

challenged at the start of their careers which may lead to missed opportunities with regard to teacher retention and longer-term impact on learners' progress.

A strong sense of purpose is vital, not just at the start of the teacher journey, but throughout a teacher's career. Day et al. (2006: 8) note that teachers' effectiveness is "mediated by: ... the extent to which teachers sustain their commitment (i.e., are more resilient)." In a similar vein, Hong (2012) notes that students' needs are most effectively met when taught by a positive teacher with a clear sense of professional purpose and confidence.

Confidence

An early career teacher's reflections offer an entry point for considering the importance of confidence, reflecting that:

resilience, confidence, work-life balance and so on are all big issues for early career teachers. I have often thought that getting off to a good start early on in a career, and feeling successful early on, is really important and can make or break a career. Given the stresses and extremely high learning curve in teaching, I think this is particularly relevant.

Early career teachers have access to numerous books about surviving in teaching (Allies, 2021; Cowley, 2019; Holmes, 2019) and about teacher stress (Galton and MacBeath, 2008), as well as articles and blogs on professional wellbeing. It is perhaps unsurprising that early career teachers in my research regarded teacher resilience as more aligned with "coping" rather than a developmental process leading to personal and professional growth. Particular capacities emerge as related to confidence—dealing with uncertainty, autonomy and agency, reflective practice and mentoring support—but confidence is identified as a key aspect of early career "coping" and preparedness for, and early success in, the profession.

The identification of resilience with coping may be in part because the early career teacher occupies a liminal space, no longer protected by the structure and support of the initial teacher education program and not yet able to draw on a significant level of teacher experience. This lack of a professional hinterland poses difficulty in times of uncertainty or in responding to "contested issues" (Carter, 2015: 8) and the early career teacher can become reactive in the face of the inevitable disruptions that are part of a teacher's daily life (Britzman, 2006). Since effective education systems need experienced teachers to enable children and young people fully to progress (Ingersoll, 2002), it is essential that early careers teachers are nurtured and valued to be able to deal with "both push and pull factors" (Newton, 2021: 43) of professional experience. This is particularly important in contexts where negativity around the teaching profession is prevalent and where the "drip-drip effect of professional intent is eroded" (Newton, 2021: 37). Increasing prescription in terms of curriculum content in some countries, combined with a performativity culture (Ball, 2003), may lead to feelings of powerlessness and loss of professional autonomy (Sullivan et al., 2019) for the early career teacher who needs to be able to navigate this complex landscape with an ever increasing level of confidence.

The importance of reflective practice is closely linked to the confidence of early career teachers as they navigate what Schön describes as the "wampy lowlands" (1983: 42) of professional practice. The links between reflective practice and agency are well-rehearsed (Hargreaves and Fullan, 2012; Korthagen et al., 2001; Ward and McCotter, 2004) as early career teachers explore issues and broaden their understanding of their work. Developing as a teacher is more than becoming a compliant technician and includes the need to expertly employ pedagogy, develop subject knowledge, set realistic targets for development, draw on research evidence, and prioritize high levels of self-awareness as an early career teacher (La Velle and Kendall, 2021).

Appropriate training and support are essential in building confidence and, by extension, resilience of early career teachers around the world (Van Nuland et al., 2021). My own research highlighted the impact of negative feedback on the confidence of early career teachers, for example: "(I)t affected my confidence and I began to question my capabilities" and "No matter what I do, it won't ever fully satisfy the expectations of my mentor." In contrast, the overwhelming value of positive and constructive mentor relationships emerged as crucial in terms of fostering and nurturing the resilience of an early career teacher in the initial stages of a school placement. This metaphor from a participant in my research illustrates the power of a mentor to redress a lack of confidence in the early stages for an early career teacher:

... like your way up to the (bungee) jump where you feel like you can do it, but the higher you go the more the doubts creep in about your ability to jump. Then you jump and it's exciting and dangerous. Then near the end of the jump you feel the tug of the safety cord, you know you are going to survive and, even though there are still more ups and downs, confidence is higher and you can really start to just enjoy the jump.

Social support

Furlong and Maynard's comment (1995: 195) that "the quality of the next generation of teachers will, in large part, depend on the quality of the mentoring they are given" continues to have resonance. Indeed, in England, there is a renewed emphasis on high-quality, sustained mentoring in the Core Content Framework (DfE, 2019a) for initial teacher training and the Early Career

Framework (DfE, 2019b). Effective mentoring is a necessary aspect of support for the early career teacher, minimizing feelings of vulnerability and stress. Mentoring has the potential to reduce the risk of attrition in the early stages of a teaching career as new teachers feel supported to remain in teaching for the medium and longer term (Ingersoll and Strong, 2011). Well-structured mentoring is of particular value where early career teachers are dependent on short-term contracts and leaders are “reluctant to invest time, energy or resources in induction programs” (Farrell, 2016: 13). Such a lack of support for early career teachers with temporary contracts was noted in Australia (Gordon, 2020b), adversely affecting new teachers’ sense of belonging and professional progress. It should also be noted that mentors of early career teachers deserve appropriate training, recognition and time for their work (Mackie, 2020), not least as one early career teacher noted: “While I have ... a supportive department, I feel that everyone is so busy that there isn’t enough time to clarify my various (endless) queries!”. Any lack of mentoring time for the early career teacher has the potential to undermine resilience and lead to professional dissatisfaction and attrition in teaching.

In this context, Kelchtermans’ description (2009: 265) of vulnerability as a “structural characteristic of the profession” is significant. Although effective mentoring is necessary for early career teacher development, a key feature of being resilient is having the assurance to ask for help when needed, recognizing that learning and growing as a new teacher requires a patient and pro-active approach (Day et al., 2011). An effective mentoring relationship is two-way which means that early career teachers should be agentic and take ownership of their professional development. For teachers to develop resilience, “they need to be empowered and trusted” (Thom, 2020: 183), knowing that support and challenge is on hand.

One aspect of social support that is perhaps underdeveloped in the literature about early career teacher development is the importance of communities of practice (Wenger, 1998) where teachers have the opportunity for collaborative development. Hargreaves (1995) lists numerous benefits of professional collaboration, including social and moral support, increased efficiency and effectiveness, sharing of professional wisdom, continuous learning and opportunities for deeper reflection. Cordingley and Crisp (2021: 140) explore the impact of professional learning in global contexts on retention in the profession, noting that the relationship between the two is “complex and non-linear,” requiring encouragement, time, dialog, patience, and space to effect positive change.

In many countries, early career teachers are encouraged to join professional subject-specific associations, providing opportunities to learn, explore and share fresh ideas with other teachers, often working in a different context, thereby reducing the risk of professional isolation or stagnation which can happen even in a large and vibrant school. However, this may be overwhelming for primary early career teachers in particular who are faced with many subjects, and so more generic networks and communities of practice may be more beneficial in the early stages.

Organizations such as The Chartered College of Teaching in England prioritize support for early career teachers. This support includes information deriving from contemporary research, strategies for managing workload, insights about teaching approaches and an encouragement to be connected in order to shape the future of the profession, thereby building social and professional capital. Such professional communities of practice are necessary, as they foster “teachers building resilience together” (Younie and Preston, 2021: 79). As noted by Fletcher-Wood and Zuccollo (2020), well-structured continuous professional development has potential in influencing retention in the profession. Supportive networks enhance the sense of professional development, well-being and satisfaction for early career teachers.

Agency

A further important aspect of resilience is being able to think and act flexibly, to adapt in the face of new or challenging situations, and to develop professional agency and autonomy through the early years in the profession. This aligns closely with the idea of self-efficacy (Bandura, 1982; Biesta, 2020). Garvey et al. (2014: 112) highlight the need for teachers to have “strong and stable personalities and to be able to tolerate complexity” and Neenan (2009) reminds us that people are likely to develop resilience by stepping out of their taken-for-granted frames of understanding. Given this context, it is concerning that a sense of teacher agency and autonomy are under-developed among early career teachers (Worth and Van den Brande, 2020) and may lead to loss of professional satisfaction for an individual or a de-professionalized profession more broadly. The reasons are perhaps unsurprising, noting Philion and Connelly’s (2004: 468) observation that early career teachers “want answers and methods. They want to be certain, to know; but certainty does not generate the flexibly inquiring attitude required by learning.”

Accepting one’s mistakes with a positive mindset, using them to learn, is a vital part of being resilient, where early career teachers make uneven progress in their development as a teacher. Learning to adapt and take control of circumstances is essential for early career teachers as they respond to the flow and inevitable uncertainties of being a teacher. These skills lie at the heart of Gu and Day’s (2013: 26) argument that teacher resilience is: “not primarily concerned with the capacity to ‘bounce back’ or recover from highly traumatic experiences and events but, rather, the capacity to maintain equilibrium and a sense of commitment and agency in the everyday worlds in which teachers teach.”

As noted earlier, the importance of reflective practice for early career teachers is well-documented (Schön, 1983; Pollard, 2018; McGregor and Cartwright, 2011) and must be nurtured and encouraged as an integral part of an early career teacher’s practice. This is more nuanced than the requirement to “reflect systematically” as noted in the Teachers’ Standards in England (DfE, 2012: 6) as such mechanistic language loses the richness engendered by deep thought, critical engagement with literature about teaching and learning, and collaborative work with others over time. At its best, reflective practice prompts all teachers, including early career teachers, to think differently, to adapt, to re-frame, to take ownership and, above all, to develop agency in their work.

Johnson et al. (2015: 103) capture the power of reflective practice at the heart of early career teacher resilience, noting that it allows for “the emergence of a robust teacher identity.” A reflective approach minimizes the risk of authority-dependence, not in a confrontational sense, but rather in the spirit of transformational agency.

One reflection by an early career teacher, comparing her experience as a teacher to an orchestra conductor, serves as a powerful illustration of the importance of agency. Indeed, a sense of the joy of teaching may often be overlooked as central to a teacher’s resilience:

While I had been learning much about rhythm, one essential thing was missing. That is the beauty of the music. This beauty might have been missing because so much focus was put on rhythm, counting the beats, respecting the pause, which are all important elements in an orchestra of this size. By now, I knew that what I would be looking for when putting together my own orchestra was beauty above all.

Constraints on agency and possible solutions must be considered. Feedback from my own research participants cited the performativity culture of initial teacher education as a particular challenge—“every move is watched,” “so many people to please” and “death by targets” were common themes. One research participant reflected on the overall system, “I wonder if a better long-term approach would be to look at why teaching requires such extreme amounts of resilience and think of ways to improve the situation.” This situation may be less acute in other countries than England, where prescriptive Teachers’ Standards (DfE, 2012) need to be evidenced to gain qualified teacher status. The climate of prescription and compliance in England is unlikely to improve with the introduction of the ITT Core Content Framework (DfE, 2019a) with its requirements for early career teachers to “Learn that ...” and “Learn how to ...” across key areas of professional learning. Against this backdrop, the importance of school culture and support from mentors and school leaders cannot be underestimated, as they will add the much-needed personalized support to foster early career teacher resilience.

Culture

There is ample evidence to suggest that “conditions count” (Gu and Day, 2013: 22) in relation to teacher resilience and the ability to manage complex interactions in the shifting worlds of early career teachers. A sense of belonging among early career teachers is nurtured by positive working conditions (Skaalvik and Skaalvik, 2011), both physical and social. However, it is the broader school culture and leadership, including mentors, which have an even more profound effect on professional wellbeing, particularly regarding relationships which have the power to sustain commitment and promote resilience among early career teachers (Peters and Pearce, 2012). This echoes Luthar and Brown’s (2007: 947) premise that “relationships lie at the ‘roots’ of resilience ... the presence of support, love, and security fosters resilience in part, by reinforcing people’s innate strengths (such as self-efficacy, positive emotions and emotion regulation).”

A positive school culture is created over time, intentionally, by the senior leadership team in collaboration with all colleagues. It is vital for everyone’s voice to be heard and valued, for colleagues to feel trusted and energized in their work, leading to professional satisfaction (Cowley, 2019). Such an approach must include early career teachers who must be valued as colleagues and require a positive environment in which to grow and flourish.

The renewed focus on the importance of professional mentoring in many countries is to be welcomed. Indeed, a well-supported mentor relationship which encourages critical reflection, including uncertainty, may paradoxically strengthen the early career teacher when faced with varied demands and expectations. As noted earlier, there is need to move beyond a formulaic, “tick box” approach to reflection, and this applies equally to the mentoring relationship. The real significance of the mentoring relationship emerges from the bespoke approach to individual early career teachers, exploring their experiences and emotions to gain valuable insights into how to support them more effectively than any “one size fits all” prescriptive approach is likely to achieve. It is well-considered collaborative approaches by an expert mentor that build professional knowledge and skills, as well as nurture the resilience and self-esteem of early career teachers (Lofthouse, 2018).

Creative approaches to reflection as part of the mentoring relationship may prove beneficial in this regard. As Clandinin and Connelly (2000: xxvi) note, “(s)toories lived and told educate the self and others.” Herein lies perhaps one of the biggest challenges for the profession—how to create sufficient time for mentors to engage with the early career teacher to complete prescribed mentoring tasks and nurture and empower the new teacher in their care. For example, as part of my research, an early career teacher wrote a short reflection on the initial weeks of a teacher education course, full of highly-charged emotion and perceptive insights about how her values were challenged and re-focused as she adapted rapidly during this early stage in a new career. When starting her first school placement, this experience may be overlooked by a busy mentor who is keen to provide induction materials, teaching resources and so on. A mentor who understands the prior experiences of an early career teacher, taking the time to explore the starting point for their relationship in depth, is more likely to make a difference to their professional development.

The voices of two early career teachers, one negative and one positive, may serve as a summary, underlining the power of the mentor to “make or break” an early career teacher. The first new teacher found it challenging to manage the (perhaps perceived) relentless demands of the mentor, commenting “expectations are so high” not only from the mentor but also self-imposed, adding

to the overwhelming sense of anxiety at work. Writing a short story as part of her reflection on the initial teacher education program, the reader gains a real understanding of the challenges faced by this early career teacher—“*long wavy grass*,” “*fast deep river*,” “*deep dark forest*” and, perhaps most poignantly, “*a mentor-filled snow storm*.” Teaching may have been the right profession for this individual with her strong subject knowledge and commitment to making a difference to the lives of children and young people, but it was ultimately the detrimental cost of a performativity culture to the individual’s wellbeing which led to leaving the profession. In contrast, there are many stories told by early career teachers of how an inspirational mentor has rescued them in times of need and doubt, challenged them to strive to be even better as a teacher, and empowered them to take ownership of their own professional development. Most early career teachers will experience moments of self-doubt and anxiety about their chosen profession and the mentor has a positive role to play in boosting resilience at such times, as noted in this haiku written by an early career teacher:

*Tired, tears and doubt.
You can do it, mentor said.
Self-belief, vital.*

Concluding reflections

Nurturing the resilience of early career teachers is both a moral and strategic imperative with clear potential to enhance professional satisfaction and retention. Although there is some dissonance between policy makers and experts on teacher education around the world (Ovenden-Hope and Passy, 2021), a key focus on the resilience of early career teachers who are valued and supported remains important in all contexts.

The growing global concern for early career teachers (Paniagua and Sánchez-Martí, 2018) offers some hope that the issue of retention of the newest members of the profession is being accorded its deserved significance. International studies are informative (Hong, 2012; Johnson et al., 2015; Smith and Ulvik, 2017) in exploring ways of supporting early career teachers in a wide range of contexts. A clear sense of purpose among new teachers needs to be extended to policy directives, underpinned by appropriate time and resource for mentoring of early career teachers. For example, Johnson et al. (2014) proposed a framework of conditions which support early career teacher resilience in Australia. Such research has informed current policy in Australia which requires employing schools “to strengthen their role in providing access for early career teachers to high quality induction and mentoring” (Education Council, 2018: iii).

In England, the Early Career Framework (DfE, 2019b: 4) affirms that teachers “deserve high quality support throughout their careers, particularly in those first years of teaching when the learning curve is the steepest.” Such an approach acknowledges the need for ongoing input and support through the early years of the profession on issues around workload and challenges, such as managing behavior and workload. That said, tensions abound at every turn. For example, the ITT Core Content Framework in England has laudable intent, requiring early careers teachers to learn how to “protect time for rest and recovery” (DfE, 2019a), presumably from a standpoint that an overly-stressed teacher is unlikely to be resilient and thrive in the profession. And yet, as one early career teacher commented about the importance of developing resilience, “... it’s a bit of a case of managing the symptom rather than dealing with the cause of the problem.”

Above all, with a plethora of frameworks intended to support early career teachers, vigilance is needed to ensure that any interest in resilience does not become a manipulative tool (Cooper et al., 2013) for merely ensuring that early career teachers are better equipped to cope with ever-increasing challenges in their work. There is a danger that we delimit education and, specifically, the development of early careers teachers, if we assume that resilience “is more a code for social compliance than a political ambition” (Evans and Reid, 2014: 102) as we seek to transform the professional journey of early career teachers in a positive way.

Systemic change is needed, but there are also more immediate and localized actions that may foster resilience of early career teachers. As discussed, these actions include exploration of varied and creative approaches to reflective practice, a personalized stance toward the emotional and affective aspects of early career teaching, an integrated and sustained approach between initial teacher education providers and employing schools, bespoke mentoring support, as well as a carefully sequenced transition for early career teachers in the first few years of teaching. The overall aim must be to avoid the “recruit—burnout—replace” model (Allen and Sims, 2008: 20) with one that focuses more firmly on continuous development, mentoring and nurture for early career teachers (Gordon, 2020b). Early career teachers deserve professional consideration and personalized mentoring, combined with empathy and appropriate challenge, so as to develop resilience and thrive in the profession.

To this end, there are two main reasons why resilience should be a priority in relation to early career teachers. First, there is evidence that resilience building has the potential to “enhance productivity” and “foster workplace well-being and engagement” (Robertson et al., 2015: 559). As a result, early career teachers should remain a prime concern for all those involved in their development. Secondly, as noted by Bajorek et al. (2014), teacher wellbeing, including resilience, has a positive impact on teacher effectiveness and student outcomes. Quite simply, a resilient teacher is more likely to thrive in the profession and make a difference to the lives of children and young people.

I conclude this chapter, not with the words of an early career teacher, but with the words of a fictional character in Foden's (2010: 315) novel, entitled *Turbulence*. The words capture the challenge for humankind in an ever-changing world, including the world of education for early career teachers, to maintain and develop resilience. It is often the incompleteness of the melody, or perhaps losing sight of the beauty of the music, that hinders resilience most.

How in a world of disintegration and endless renewal—a continuum, a world of flow—one must find one's own rhythm exactly by recognizing the incompleteness of the melody.

References

- Allen, R., Sims, S., 2018. *The Teacher Gap*. Routledge, Abingdon, Oxon.
- Allies, S., 2021. *Supporting Teacher Wellbeing*. Routledge, Abingdon, Oxon.
- Bajorek, Z., Gulliford, J., Taskila, T., 2014. *Healthy Teachers, Higher Marks? Establishing a Link Between Teacher Health and Wellbeing and Student Outcomes*. The Work Foundation, London.
- Ball, S., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Bandura, A., 1982. Self-efficacy mechanism in human agency. *Am. Psychol.* 37, 122–147.
- Biesta, G., 2020. Risking ourselves in education: qualification, socialization, and subjectification revisited. *Educ. Theor.* 70, 89–104.
- Britzman, D., 2006. Teacher education as uneven development: toward a psychology of uncertainty. *Int. J. Leader. Educ.* 10 (1), 1–12.
- Carter, A., 2015. *Carter Review of Initial Teacher Training (ITT)*. Department for Education, London.
- Chandler, D., 2014. *Resilience—The Governance of Complexity*. Routledge, London.
- Clandinin, D., Connelly, F., 2000. *Narrative Inquiry*. Jossey-Bass, San Francisco.
- Cooper, C., Flint-Taylor, J., Pearn, M., 2013. *Building Resilience for Success*. Palgrave Macmillan, New York.
- Cordingley, P., Crisp, B., 2021. Professional learning and recruitment and retention: what global regions can tell us. In: Ovenden-Hope, T., Passy, R. (Eds.), *Exploring Teacher Recruitment and Retention*, pp. 131–147.
- Cowley, A., 2019. *The Wellbeing Toolkit*. Bloomsbury, London.
- Day, C., Gu, Q., 2007. Variations in the conditions for teachers' professional learning and development: sustaining commitment and effectiveness over a career. *Oxf. Rev. Educ.* 33 (4), 423–443.
- Day, C., Stobart, G., Sammons, P., Kington, A., Gu, Q., Smees, R., Mujtaba, T., 2006. *Variations in Teachers' Work, Lives and Effectiveness*. Final Report for the VITAE Project. DfES, London.
- Day, C., Edwards, A., Griffiths, A., Gu, Q., 2011. *Beyond Survival—Teachers and Resilience*, Key Messages From an ERSC-Funded Seminar Series. University of Nottingham, Nottingham.
- Department for Education, 2012. *Teachers' Standards*. DfE, London.
- Department for Education, 2019a. *ITT Core Content Framework*. DfE, London.
- Department for Education, 2019b. *Early Career Framework*. DfE, London.
- Education Council, 2018. *One Teaching Profession: Teacher Registration in Australia*. Education Services Australia, Victoria.
- Evans, B., Reid, J., 2014. *Resilient Life—The Art of Loving Dangerously*. Polity Press, Cambridge.
- Farrell, T., 2016. Surviving the transition shock in the first year of teaching through reflective practice. *System* 61, 12–19.
- Fletcher-Wood, H., Zuccollo, J., 2020. *The Effects of High-Quality Professional Development on Teachers and Students*. Educational Policy Institute, London.
- Foden, G., 2010. *Turbulence*. Faber and Faber, London.
- Furlong, J., Maynard, T., 1995. *Mentoring Student Teachers*. Routledge, London.
- Galton, M., MacBeath, J., 2008. *Teachers Under Pressure*. SAGE, London.
- Garvey, B., Stokes, P., Megginson, D., 2014. *Coaching and Mentoring: Theory and Practice*. SAGE, London.
- Gordon, A.L., 2020a. *Creative Approaches to Nurturing Resilience of Trainee Teachers*. In: *Impact*. Chartered College of Teaching, London.
- Gordon, A.L., 2020b. Educate-mentor-nurture: improving the transition from initial teacher education to qualified teacher status and beyond. *J. Educ. Teach.* 46 (5), 664–675.
- Grotberg, E., 1995. *A Guide to Promoting Resilience in Children: Strengthening the Human Spirit*. Bernard Van Leer Foundation, The Hague.
- Gu, Q., Day, C., 2007. Teacher resilience: a necessary condition for effectiveness. *Teach. Teach. Educ.* 23 (8), 1302–1316.
- Gu, Q., Day, C., 2013. Challenges to teacher resilience: conditions count. *Br. Educ. Res. J.* 39 (1), 22–44.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Columbia University. Teachers College Press, New York.
- Hargreaves, A., 1995. Beyond collaboration: critical teacher development in the postmodern age. In: Smyth, J. (Ed.), *Critical Discourses on Teacher Development*. Cassell, pp. 149–179.
- Holmes, E., 2019. *A Practical Guide to Teacher Wellbeing*. SAGE, London.
- Hong, J., 2012. Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teach. Teach.* 18 (4), 417–440.
- Howes, L., Goodman-Delahunty, J., 2015. Teachers' career decisions: perspectives on choosing teaching careers, and on staying or leaving. *Issues Educ. Res.* 25 (1), 18–35.
- Ingersoll, R., Strong, M., 2011. The impact of induction and mentoring programs for beginning teachers: a critical review of the research. *Rev. Educ. Res.* 81 (2), 201–233.
- Ingersoll, R., 2002. The teacher shortage: a case of wrong diagnosis and wrong prescription. *NASSP Bull.* 86 (631), 16–31.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A., Pearce, J., Hunter, J., 2014. Promoting early career teacher resilience: a framework for understanding and acting. *Teach. Teach. Theor. Pract.* 20 (5), 530–546.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A., Pearce, J., Hunter, J., 2015. *Early Career Teachers Stories of Resilience*. Springer Briefs in Education, Perth, Australia.
- Kelchtermans, G., 2009. Who I am in how I teach is the message: self-understanding, vulnerability and reflection. *Teach. Teach.* 15 (2), 257–272.
- Kelchtermans, G., 2017. "Should I stay or should I go?" Unpacking teacher attrition/retention as an educational issue. *Teach. Teach.* 23, 961–977.
- Kincheloe, J., Berry, K., 2004. *Rigour and Complexity in Educational Research*. Open University Press, Maidenhead.
- Kincheloe, J., 2008. *Critical Pedagogy*, second ed. Peter Lang Publishing, New York.
- Korthagen, F., Kessels, J., Koster, B., Lagerwerf, B., Wubbels, T., 2001. *Linking Practice and Theory: The Pedagogy of Realistic Teacher Education*. Routledge, London.
- La Velle, L., Kendall, A., 2021. A high-status, research-informed profession. In: Ovenden-Hope, T., Passy, R. (Eds.), *Exploring Teacher Recruitment and Retention*. Routledge, Abingdon.
- Lofthouse, R., 2018. A Virtuous Circle: Making the Most of Mentoring. *The Profession*, pp. 109–111.
- Lucas, B., Claxton, G., Spencer, E., 2013. *Expansive Education*. Open University Press, Maidenhead.
- Luthar, S., Brown, P., 2007. Maximizing resilience through diverse levels of inquiry: prevailing paradigms, possibilities, and priorities for the future. *Dev. Psychopathol.* 19, 931–955.

- Mackie, L., 2020. Understandings of mentoring in school placement settings within the context of initial teacher education in Scotland: dimensions of collaboration and power. *J. Educ. Teach.* 46 (3), 263–280.
- McGregor, D., Cartwright, L., 2011. *Developing Reflective Practice: A Guide for Beginning Teachers*. Open University Press, Maidenhead.
- Neenan, M., 2009. *Developing Resilience—A Cognitive Behavioral Approach*. Routledge, Hove.
- Newton, G., 2021. Why are teachers leaving teaching in England? In: Ovenden-Hope, T., Passy, R. (Eds.), *Exploring Teacher Recruitment and Retention*. Routledge, Abingdon.
- Ovenden-Hope, T., Passy, R., 2021. *Exploring Teacher Recruitment and Retention*. Routledge, Abingdon.
- Paniagua, A., Sánchez-Martí, A., 2018. Early Career Teachers: Pioneers Triggering Innovation or Compliant Professionals? OECD Working Paper No.190 OECD, Paris.
- Peters, J., Pearce, J., 2012. Relationships and early career teacher resilience: a role for school principals. *Teach. Teach. Theor. Pract.* 18 (2), 249–262.
- Phillion, J., Connelly, F., 2004. Narrative, diversity, and teacher education. *Teach. Teach. Educ.* 20, 457–471.
- Pollard, A. (Ed.), 2018. *Reflective Teaching in Schools*, fifth ed. Bloomsbury Academic, London.
- Robertson, I., Cooper, C., Sarkar, M., Curran, T., 2015. Resilience training in the workplace from 2003–2014: a systematic review. *J. Occup. Organ. Psychol.* 88, 533–562.
- Sabar, N., 2004. From heaven to reality through crisis: novice teachers as migrants. *Teach. Teach. Educ.* 20 (2), 145–161.
- Schön, D., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- Skaalvik, E., Skaalvik, S., 2011. Teacher job satisfaction and motivation to leave the teaching profession: relations with school context, feeling of belonging, and emotional exhaustion. *Teach. Teach. Educ.* 27 (6), 1029–1038.
- Smith, K., Ulvik, M., 2017. Leaving teaching: lack of resilience or sign of agency? *Teach. Teach. Theor. Pract.* 23 (8), 928–945.
- Sullivan, A., Johnson, B., Simon, M., 2019. *Attracting and Keeping the Best Teachers*. Springer, Singapore.
- Thom, J., 2020. *Teacher Resilience—Managing Stress and Anxiety to Thrive in the Classroom*. Woodbridge, Suffolk. John Catt Educational Ltd.
- UNESCO, 2017. *Moving Forward the 2030 Agenda for Sustainable Development*. UNESCO, Paris.
- Van Nuland, S., Whalen, C., Majoche, E., 2021. Teacher recruitment and retention in Canada. In: Ovenden-Hope, T., Passy, R. (Eds.), *Exploring Teacher Recruitment and Retention*. Routledge, Abingdon.
- Vesely, A., Saklofske, D., Nordstokke, D., 2014. EI training and pre-service teacher wellbeing. *Pers. Indiv. Differ.* 65, 81–85.
- Waddell, J., 2007. The time is now. In: Davies, D. (Ed.), *Resiliency Reconsidered: Deconstructing the Policy Implications of the Resiliency Movement*. Information Age Publishing, Charlotte, North Carolina.
- Ward, J., McCotter, S., 2004. Reflection as a visible outcome for preservice teachers. *Teach. Teach. Educ.* 20 (3), 243–257.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, New York.
- Worth, J., Van den Brande, J., 2020. *Teacher Autonomy: How Does It Relate to Job Satisfaction and Retention?* National Foundation for Educational Research, Slough.
- Younie, S., Preston, C., 2021. Understanding the contribution of professional communities of practice in education technology in influencing teacher recruitment and retention. In: Ovenden-Hope, T., Passy, R. (Eds.), *Exploring Teacher Recruitment and Retention*, pp. 73–83.
- Zolli, A., Healy, A., 2012. *Resilience*. Headline Publishing Group, London.

Beginning teachers' professional identity work and its support

Jonny Schellings and Douwe Beijaard, Eindhoven University of Technology, Eindhoven, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	161
Conceptualizing beginning teachers' professional identity	161
Constraining and cultivating professional identity experiences	162
Identity issues beginning teachers face	163
Experiences that confirm or reinforce beginning teachers' identity	163
Moments of contact with students	163
Positioning toward students, colleagues and school leaders	164
Stories to live by	164
Supporting beginning teacher's professional identity development	164
Aim of identity work	164
Characteristics of identity work	165
The role of reflection	165
Tailor-made support aimed at strength	165
Recognizing growth	165
Safe environment	165
Program aimed at identity work	166
Implications	167
References	168

Introduction

Learning to be a teacher is a complex and multi-faceted process associated not only with the acquisition of professional knowledge and skills; it is also a process of building a realistic professional teacher identity (Arvaja, 2016; Beijaard et al., 2004). Many other, particularly personal and situational or contextual aspects that, together and in interaction with each other, make that teachers are who they are, can be or wish to become. Developing a professional identity is a key element of the transition into the teaching profession, yet teacher education program seem to focus predominantly on acquiring knowledge and skills imposed by the professional community, while neglecting or undervaluing the process of professional identity formation (Cobb, 2020; Rodgers and Scott, 2008).

This chapter highlights the professional identity of beginning teachers and how to support them in their work on their professional identity. In the literature, beginning teachers are named differently, for example: early career teachers, new teachers, or newly qualified teachers. Generally, most resources (implicit or explicit) see this group as a group of teachers, who, after graduation from teacher education, are in their first 3 to 5 years of working as teacher with full responsibility of their classes (Helms-Lorenz et al., 2016). Beginning teachers are often perceived as one group entering the profession with similar challenges and concerns, yet the group is composed of different individuals with different personal and educational backgrounds, needs and interests and sometimes different ages as well, but certainly with different reasons to become a teacher. Actually, beginning teachers together are a very heterogenous group, a feature reinforced by the different schools they enter after their graduation from teacher education.

This chapter starts with an explanation of how the authors of this chapter understand professional identity and its role in becoming a teacher. Next attention is paid to experiences that fuel the identity of beginning teachers while growing into the profession. Beginning teachers may face identity issues that stem from negative experiences between what is contextually possible and personally found relevant or inversely. Research has mapped out a number of these issues. While being important, up till now far too little attention has been paid to the positive experiences of beginning teachers regarding the ways in which person and context align. As far as possible, beginning teachers' positive experiences that influence their professional identity development as well are also discussed. This chapter continues with the aim and characteristics of identity work together with examples of how to support identity development and ends with some implications for identity work and related research.

Conceptualizing beginning teachers' professional identity

Beginning teachers, on average, show a steep learning curve during their 1st years of working as teachers, also because during this period relatively many of them drop out (Helms-Lorenz et al., 2016). It is particularly their 1st years in which they work on the teacher they want to be, in their work with students, colleagues and significant other people such as their school managers and/or leaders and parents (e.g., Cobb, 2020; van der Wal et al., 2019).

Making a professional identity by beginning teachers includes more than focusing on subject matter, pedagogical content knowledge, theories of teaching and learning, and skills to turn that all into practical action. Beginning teachers' professional identity development results from the complex interaction between their personal beliefs, ideals, backgrounds, norms and values they hold, including their experiences as student teachers, and requirements from the context of their schools and the practical situations they face, including widely accepted standards for the teaching profession in general (Beijaard et al., 2004; Garner and Kaplan, 2019). Professional identity understood in this way leads to a number of characteristics that can be attached to it.

Firstly, beginning teachers' professional identity is both a product and a process. As a product, professional identity might be seen as a complex configuration of personal and professional aspects that together make who a teacher is in a given situation on a given moment in time; as a process, beginning teachers' professional identity is constantly subject to change based on the interpretation and reinterpretation of experiences (Olsen, 2010; Schellings et al., 2021).

Secondly and in line with the above, beginning teachers' professional identity and its development are to a large extent situated or contextualized, both as a teacher in the classroom and as a member of the school team. Through reflection on their experiences and setting goals for new experiences based on their own motives, needs or interests, beginning teachers also have an influence on who they want to be as a teacher. Exercising a degree of control over their own work as a teacher and through that staying true to themselves, in the literature often referred to as being agentic (e.g., Etelepälto et al., 2015), is essential to developing a realistic professional identity that fits the individual. In professional identity making, agency ("what comes from within") and context ("what comes from outside") are dialectically related. Some schools offer (beginning) teachers more room for being an agent of their own work as teacher than other schools.

Thirdly, teachers' professional identity is not static but dynamic, depending on the situation in which they find themselves. Different types of students that are taught, collaboration with colleagues, conversations with parents about students, etc. make that teachers' professional identity is multi-faceted. It might therefore be argued that teachers' professional identity consists of sub-identities that are more or less in harmony with each other (Beijaard et al., 2004). Different situations in which teachers find themselves appeal to different aspects of their professional identity. In this context, the literature also refers to various I-positions that teachers take in different situations of their work (Akkerman and Meijer, 2011). Developing a harmonious professional identity can be quite a challenge for beginning teachers during the early years of their career.

Fourthly, due to the many personal and contextual aspects that configure a beginning teacher's professional identity and the way it develops, professional identity is not only complex in nature but (partly) unique as well. As indicated in the introduction, beginning teachers differ from each other in many respects, including differences in backgrounds and opinions about their work as teachers, for example as a result of their own experiences as a student and their upbringing, i.e., their learning histories and biographies in general (cf. Kissling, 2014).

Fifthly, beginning teachers' professional identity manifests itself in their image of themselves as teachers, i.e., how they see themselves as teacher or the teacher they want to be (Akkerman and Meijer, 2011; Beijaard et al., 2004). Beginning teachers' perceptions of experiences in practice that do not match their professional identity may lead to internal or mental conflicts that, in turn, may negatively influence their professional self-image. This can lead to all kinds of reactions, such as denying the relevance of an experience and putting it aside or seeing it as a challenge to work on. Conversely, perceptions of experiences that match well with their professional identity usually confirm or strengthen their professional self-image without any difficulties or concerns. Essentially, beginning teachers' professional identity functions as filter for professional learning and development. Experiences that impact one's professional identity may lead to (dis)continuities in learning and development (e.g., Leeferink et al., 2019; van Rijswijk et al., 2018).

Sixthly and finally, beginning teachers' professional identity development is not merely a cognitive process; it is also an emotional process (cf. Zwart et al., 2015). It goes without saying that certain experiences that lead to significant discontinuities in beginning teachers' professional identity can be accompanied by heavy emotions such as frustration, dissatisfaction, uncertainty, anger or even disgust (Hargreaves, 2001; Zembylas, 2003). Particularly in situations in which beginning teachers experience strong negative frictions between the personal and contextual. In the next section, these experiences are discussed in terms of identity issues that many beginning teachers face. On the other hand, when there is a harmonious connection between personal and contextual aspects, teachers' ideals and beliefs act as a driving force for engaging in the environment, an environment in which the teachers flourish and grow as professionals (Zwart et al., 2015). Unfortunately, without receiving the appropriate support, positive experiences that create feelings of flow may eventually result in negative emotions as well (Hoekstra and Korthagen, 2011).

Constraining and cultivating professional identity experiences

Beginning teachers' professional identity making is seen in this chapter as a complex, to a large extent unique and continuous process of reconciling the personal and contextual. Teachers' contexts either provide them with opportunities to cultivate or to constrain their professional identity development in terms of the teacher they want to be. As written above, feelings of constraints are often accompanied by negative emotions. On the other hand, emotions like joy and satisfaction occur when things are going well, such as experiencing good relationships with students and colleagues, which fuel feelings of being a successful teacher (Vähäsantanen et al., 2020). Below attention will be paid to experiences that both constrain and cultivate or strengthen professional identity development; the first ones in terms of identity issues, the second ones in terms of experiences that confirm or reinforce identity.

Identity issues beginning teachers face

Research on beginning teachers' professional identity has revealed insight into what issues beginning teachers face when reconciling the personal and contextual in becoming the teacher they want to be. These insights particularly pertain to:

1. The transition phase from a student who is still in teacher education to a beginning teacher in school with all associated responsibilities, which can be experienced as overwhelming and, through that, lead to experiencing an imbalance between private life and work and regularly characterized by emotional exhaustion and symptoms of burn-out (e.g., Pillen et al., 2013; Schaefer and Clandinin, 2019).
2. The ways beginning teachers position themselves in their work toward relevant others. Teachers' position taking depends on the situation they find themselves in and the people in that situation (i.e., students, colleagues, school leaders and others, including parents in particular). Ideally the positions they take coincide with "who they are" and that they are also recognized as such by others. In their work with students, they want to be recognized as a competent teacher. Students are particularly for beginning teachers very "critical reality definers" (Beijaard, 1995); their relationship with students touches upon the core values of their work, including accompanying feelings, and may therefore enhance a deeper awareness and image of themselves as teachers (Mommers et al., 2021). By colleagues and school leaders they want to be heard, accepted, and have the opportunity to share their concerns and ideals with them. It is important that teachers are able and given space to make their own voice heard and to position themselves in ways they want to (Akkerman and Meijer, 2011). Conflicts between how beginning teachers want to position themselves or what is possible in reality, may threaten the development of their confidence, self-esteem and well-being often accompanied by negative emotions (e.g., Hargreaves, 2001).
3. The confrontation with the micropolitical reality in schools may hinder or promote beginning teachers' socialization in schools. Among other things, this is reflected in the atmosphere in the school, in how teachers collaborate, how new teachers are seen and introduced in a school, and in leadership in the school. For beginning teachers, working conditions like these are essential for an optimal growth as professionals in schools. Negative experiences with these conditions are often accompanied by negative feelings as well and by feelings of vulnerability (e.g., Kelchtermans and Ballet, 2002). Working conditions also influence the ways in which beginning teachers can and may position themselves in the school as described above.
4. The commonly shared beliefs, views, and values about teaching and education; they are collective stories shaped by traditions of schools and the place of schooling in society, i.e., the professional landscape that metaphorically can be understood as a "storied landscape" in which teachers live and work (Connelly and Clandinin, 1997, 1999). "Stories to live by" can provide beginning teachers with useful guidelines for growing into their work as teachers, but they can also be experienced by them as persuasive, demanding and restrictive for the teachers they are or want to be (see also Schaefer and Clandinin, 2019).

The above-mentioned types of issues are described in more detail and illustrated with examples elsewhere (Beijaard et al., 2022). How and the extent to which they occur in beginning teachers often remains hidden, despite their sometimes significant impact on the teacher one is and wants to become. It turns out that dealing with such issues is often quite an internal struggle for a large number of them and should therefore not be left to themselves. Making experiences regarding identity issues explicit and sharing them, for example with peers and a mentor in school, can help beginning teachers to develop a professional identity that fits the person and their context, including alignment with relevant professional knowledge and skills that are important for the teaching profession as will be discussed later in this chapter.

Experiences that confirm or reinforce beginning teachers' identity

Experiences that have a positive influence on the professional identity of beginning teachers are to a certain extent the flip side of the identity issues described above. To date, however, research on beginning teachers' processes of reconciling the personal and contextual has mainly focused on identity issues. In other words, experiences that foster continuity in professional identity development, while identity issues predominantly cause discontinuity. Hardly no research is available that reports on experiences that show how the personal and contextual influence each other in a positive way, thereby confirming or reinforcing beginning teachers' professional identity and thus fostering its further development. For this reason the authors draw back on one of their own studies (Mommers et al., 2021).

This study started from research literature on contextual or situational influences on the professional identity of beginning teachers and expressed in their professional self-image. Much of this research has already been referred to in the overview of professional identity issues in the preceding section. Against this background, research was conducted into: (1) the influence of significant others, focusing on beginning teachers' contact with students and their positioning toward students, colleagues and school leaders, and (2) stories beginning teachers "live by." One cohort of 46 beginning teachers from 11 secondary schools voluntarily participated in the study. As part of a 3-year lasting induction program in the Netherlands these teachers attended a number of so-called "identity workshops." Before these workshops the teachers carried out three writing assignments; for each assignment they were requested to write a narrative of approximately 400–600 words.

Moments of contact with students

Moments of contact, either individually or in groups, with students are essential for the development of teachers' professional identity. All beginning teachers not only experienced one or more tensions in their moments of contact with students, most of them

caused by discipline problems, they also and especially reported on successful experiences in their moments of contact and wrote about these moments as being conducive for their satisfaction, self-confidence, and well-being in general. These positive feelings particularly arose when their contact with students was mutually found important and characterized by exchange of personal information. Teachers also reported on feelings of success in case of improvement of difficult relationships with students and growth of bonds, for instance, when students start co-operating in class or share personal information. They felt they could really make a difference.

Positioning toward students, colleagues and school leaders

The positions teachers take in their work ideally coincide with “who they are,” regardless of the situations they find themselves in. Most beginning teachers’ positions toward students were congruent with their desired positions such as enacting a coaching role, fostering students’ own responsibility for their learning, and being the “driving force” behind students’ learning routes. Only a few of them felt discomfort in their positioning toward students, particularly due to problems with classroom management and working conditions like having full classes and demands to teach in a certain way. Almost all beginning teachers positioned themselves also positively toward colleagues and reported on a good cooperate with them, for example, by developing new curricula together. Having a good and productive relationship with colleagues appeared to evoke feelings of success and the ability to perform on a higher level than working alone. A few teachers, however, explicitly positioned themselves as alone and indicated having a low status in the group. Some of them did not feel free to express their own opinion either. A large number of beginning teachers also positioned themselves positively toward their school leaders, and explained that they collaborated with them or that their school leaders took their innovative ideas very seriously. These beginning teachers mostly felt to be part of the “driving force” in the school and being able to express the teacher they are. Only a small group of teachers explained their positioning toward school leaders not so much as negative but more as compliant or cautious, for example, due to uncertainty about continuation of their job.

Stories to live by

Teachers draw on stories to live by; they use them as frames of reference for their interpretation of personal and social experiences and, through that, make sense of themselves as teachers. The beginning teachers reported both “in-classroom” and “out-of-classroom” stories to live by. In-classroom stories pertained to being part of current education “in transition” with a focus on the implementation pedagogies that activate student learning and give students more individual support. All beginning teachers experienced these stories to live by as positive challenges to their work in the classroom. Nevertheless a large number of them also reported on the pressure of high standards they had to meet, based on feelings of high expectations regarding their work held by others as well as by themselves. Out-of-classroom stories pertained to teachers being at the center of change in today’s education (more focus on students’ interdisciplinary competencies and personalized education). This change appeared to match many beginning teachers’ desires to innovate, meanwhile being realistic about the practical problems and struggles that must be overcome for this. Another out-of-classroom story pertained to the need and relevance of cooperating with colleagues, thus being a member of a school team. As written before, through collaboration with colleagues many beginning teachers experienced feelings of success and were able to perform on a higher level than working alone. The other side of this story, though much less reported on, regarded teachers who used to work individually without mutual consultation or contact, often accompanied by feelings of disappointment.

The results of the study reported on are far from representative. They indicate that beginning teachers do experience identity issues when reconciling the personal and contextual, but that this is not so bad and seems to be less than much research literature suggests, both in general and more individually. The beginning teachers who took part in the above-mentioned study did not report serious identity issues, with few exceptions. On the contrary, many of them experienced positive challenges showing much continuity in their professional identity and its development. It therefore seems important in supporting beginning teachers to pay attention not only to identity issues but also to challenges that foster their professional identity development.

Supporting beginning teacher’s professional identity development

Identity development is an internal and above all personal process that might remain hidden. Making challenges or issues that are typical for one’s professional identity and its development explicit and sharing them with peers or a mentor is not common practice and may even be experienced as awkward by beginning teachers. As process, working on one’s identity takes time and usually does not result in quick answers or solutions to one’s immediate issues or challenges, such as how to engage *all* students in learning when introducing a new topic. Identity work or being actively engaged in professional identity formation, requires deep reflection upon oneself in different situations. Some beginning teachers work on their professional identity on their own while others need to be encouraged to do so. Nevertheless, it is argued that they all need support with their identity work. The above mentioned characteristics that stem from our perspective on professional identity help to define the aim and settings of “formally” organized program that support identity work.

Aim of identity work

Teachers’ identity work requires from teachers to process their understanding of themselves as professionals and, as such, building realistic “images-of-themselves-as-teacher,” i.e., the teachers they are and want to be. The core of this process is unraveling the

experiences stemming from the interplay between the personal and contextual in being and becoming a teacher. Understanding oneself as a teacher means acknowledging that one's functioning and wellbeing strongly depend on the teacher as person and his/her personal interpretations of experiences (Schaefer and Clandinin, 2019; Schellings et al., 2021).

Formally organized identity work may particularly support beginning teachers in their negotiating part of constructing a professional identity (Beijaard and Meijer, 2017; Vähäsantanen et al., 2020). Such identity work supports beginning teachers to stay true to themselves and being valued as such. By far identity work does not mean that beginning teachers learn to copy or imitate their colleagues' thoughts, beliefs, and teaching practices. Identity work means that beginning teachers are searching and finding their personal way of teaching, respect and value collegial differences, reinforce each other's strengths, or (possibly) even search for other contexts that fit their professional identity better.

Particularly within the transition phase from student to full responsible teacher, identity work is important. However, professional identity work is an important ongoing process in further career stages as well with a focus on sustaining and maintaining a positive professional identity (Schaefer and Clandinin, 2019; van der Want et al., 2018). Professional identity continues to develop because of changes in personal life or because of new experiences, either self-initiated on the basis of personal experiences and changes in knowledge and beliefs or under the influence of the context such as having to teach students with other ages or level than usual (cf. Lee, 2013). When beginning teachers (experience to) value the merits of identity work in the beginning of their career, they might be more willing to engage in identity work during their midcareer or late career.

Characteristics of identity work

To reach the above aims of identity work, beginning teachers need to engage in deep reflection on who they are, want to be, and/or can be. Formally organized program supporting identity work, must find ways to foster this kind of reflection, be tailor-made focusing on strengths, reveal growth, and take place in a safe environment.

The role of reflection

A realistic professional self-image results from an internal process of framing and reframing experiences often fueled by reflective discussions with others (cf. Korthagen and Vasalos, 2005). In our studies (Mommers et al., 2021; Schellings et al., 2021), peer discussions helped beginning teachers to reflect on differences between aspects they personally find important for their work and how these aspects are shaped or filled in by the context of their school. Reflection tools (for example, writing in a diary about experiences and how these are influenced by personal and contextual aspects or filling out specific questionnaires regarding identity issues) certainly helped them to think about their professional identity, but the (peer) discussions about the completed tools within these studies "forced" the beginning teachers to explain or elaborate on their (written) thoughts to their peers guided by the mentor who moderated the discussion.

Tailor-made support aimed at strength

Identity development is a personal and unique process. Its support needs to be tailor-made, even when plenary "identity meetings" are organized for a group of beginning teachers. Identity work requires the focus on beginning teachers' individual strengths and qualities, needs to emphasize positive challenges that the person has experienced and to embrace the innovative potential the teacher may have and bring to the school (Kelchtermans, 2019; Zeichner and Gore, 1990; Zwart et al., 2015). While the role of positive experiences in identity work has to become more prevalent, the issues that beginning teachers face can of course not be ignored. Beginning teachers learn from their issues with the appropriate support (Pillen et al., 2013). Through identity work beginning teachers learn to conceive their issues as intriguing challenges or as learning goals that stimulate them to grow as professional, especially when the issues are shared and discussed with their peers.

Recognizing growth

Beginning teachers need to be encouraged to reflect upon their professional growth enhanced by experiences in different situations over time: what was improved in this particular situation compared to the previous one? Recognizing growth can be hampered when growth consists of small continuous steps or when growth takes a nonlinear route (Garner and Kaplan, 2019; Leeferink et al., 2019). To grow as professional, beginning teachers themselves need to see and appreciate the value of what they know and do as being in line with their professional self-image, because it is particularly this image that fuels their professional learning and development (Pinnegar, 2005). It is important that they feel responsible for their own identity making and therefore take the lead in their growth as teacher by making own choices in their professional development (Beijaard and Meijer, 2017; Vähäsantanen et al., 2020).

Safe environment

Dealing with emotions, either pleasant or unpleasant ones, is inherent in identity work and have a bearing on one's expression of identity in front of others and the shaping of it (Beauchamp and Thomas, 2009; Nichols et al., 2017). Identity work must provide a safe surrounding in which participants can freely and openly discuss their issues or challenges and related experiences. Establishing a safe surrounding is one of the tasks of a mentor, who will have to take an unbiased and independent perspective in helping beginning teachers to interpret the specific school context and its unique interplay with the particular teacher. Beginning teachers need to experience unconditional support of their mentor and not feeling judged on possible deficits.

Program aimed at identity work

Different programs for identity work are published, all aiming at becoming aware of and reconciling the influence of personal and contextual aspects on the teacher one is, wants to be or can be through deep reflection on experiences. For example, [Stenberg \(2010\)](#) organized about 20 meetings for beginning teachers discussing their autobiographical stories (reflections on life-experiences) and video diaries (reflections on classroom practices) in order to unravel their core values and decisions made within their school contexts. The results showed that “transformative” reflection, a phase in which inner beliefs, values and understandings are critically considered by the person was hard to reach (two of four teachers did). The teachers mainly explained the video diaries by means of their existing thoughts and practices; they did not consider to change their perspective in making sense of their experiences.

In the identity work of [Leijen and Kullasepp \(2013\)](#) organized in the form of peer dialogs, 11 beginning teachers focused on finding solutions for experienced tensions (for example, feeling a gap between the theoretical teacher education program and teaching practices in school). The results showed that teachers' solutions were twofold, some teachers found an acceptable equilibrium between their tensions, other teachers adopted a new personal pedagogy in teaching. For example, one teacher experienced a tension between her personal conviction and the ideas of the teacher education program about the value of collaborative assignments. She originally considered collaborative assignments as time consuming and resulting in limited student learning. During the identity work, she came up with some additional conditions in her conviction to reduce her tension. She now believed that collaborative work is useful in some classes where students are more capable in working together and only when there is sufficient time. A few other teachers adopted a new personal pedagogy in teaching to solve their tensions. For example, a teacher, whose tension was about classroom management, established some personal rules to maintain discipline in her class. She realized that her personal rules differed from her colleagues' rules and, probably because of these differences in rules, she had to repeat her rules in her class again and again.

In another program ([Vähäsantanen et al., 2017](#)), four different professional groups (academic teachers, administrative personnel, nurses and physicians) portrayed their professionalism in an arts-based assignment, named “the Professional Body.” The participants received a drawing of a human body on paper, which they personalized by bringing to it different symbols (e.g., in writing, painting, pictures, clippings, and drawings) that were related to their professional identity. These self-chosen symbols connected to different identity questions, such as who I am, what is my history, what do I know, what are my dreams, etc. Next to “the Professional Body,” an experienced coach offered other less-explained assignments (such as drawing autobiographical timelines) that stimulated the participants sharing experiences and collectively searching for how they could understand their own professional identity. In interviews following the program, the participants told they had learned about themselves as professional, about their professional relationships, about personal competencies, and about themselves as a person.

To illustrate more in depth how to engage in identity work, we present two of our own studies into formally organized identity work for beginning teachers in which different assignments were used to start peer discussions about the participants' professional identity. Both studies were part of a more comprehensive induction program for beginning teachers. The first study consisted of a 1-year lasting series of school transcending identity workshops ([Schellings et al., 2021](#)) in which beginning teachers were stimulated to thoroughly reflect upon both personal and contextual aspects influencing their professional identity. During each workshop, the teachers were asked to (individually) reflect on these aspects with help of diverse assignments. In one workshop, the assignment concerned filling out a frequently used and validated Identity Tension Questionnaire developed by [Pillen et al. \(2013\)](#). A second assignment, in another workshop, concerned detecting personal key values in a challenging situation (for example, the situation of motivating all students to learn independently). The next assignment in a next workshop, asked from the beginning teachers to reduce their tensions by searching for a feasible balance between conflicting perceptions on three elements: the teacher's role, the role of all persons, and the perceived task in a challenging situation ([Schellings et al., 2021](#)). In the last workshop, the teachers were asked to draw and explain their storyline in which they visualized their growth as beginning teacher in the past year (cf. [Beijaard et al., 1999](#)). All assignments, although different in form, had in common that they triggered the teachers to explicate their perception of what aspects are important for the teachers they are and/or wish to be. Reflection outcomes were thoroughly discussed during the workshops, alternately in small peer groups and plenary. After discussing, the teachers wrote down their newly acquired insights about their professional identity in their personal notebook (for example, after a discussion in the workshop a teacher wrote: *“I have trouble in balancing work and private life. I sacrifice too much spare time to prepare my lessons according to my standard”*). At the end of each workshop, they were asked to formulate a personal learning goal to explicate what to focus on in the in-between periods of the identity workshops (for example, the teacher quoted above wrote: *“Good is good enough, I have to learn to lower my standards”*).

In our second study regarding aspects that influence beginning teachers' professional identity ([Mommers et al., 2021](#)), three digital assignments were administered and afterward discussed in a peer group. Several empirical outcomes are already presented in [Experiences that confirm or reinforce beginning teachers' identity](#) section to demonstrate the impact of focusing on beginning teachers' positive experiences. Each assignment started with a short description to relate the particular task to professional identity: the influence of significant others (in moments of contact and in positioning toward others), and the impact of the teachers' professional landscape (stories to live by). In the assignment “contact with significant others,” teachers described a moment of contact with their students triggered by open ended questions: *“What do you intend with this moment of contact?”*; *“What is the situation at hand?”*; *“What do you find important within this moment?”*; *“How do you feel when the moment turned out to be different as you expected?”* The teachers were also deliberately asked to describe their feelings of success: *“How do you feel when this moment of contacts proceeds well?”*.

In the assignment “positioning toward significant others,” beginning teachers choose three images to symbolize their positioning toward students, colleagues, and school leaders. Then, they explained why they had chosen the particular image and reflected on the relationship between the three chosen images. For example, for positioning toward students: a red fish swimming in a school of green fishes that symbolized the feeling of being equal to students while being their leader at the same time. For colleagues, for example: a circle of colored persons who stand hand in hand to demonstrate both collaboration and the important role of individuals’ contribution. For school leaders, for example: an image of chess play to illustrate the political climate within the school organization. In our study, the beginning teachers took strong and equal positions toward each group and their positioning appeared to align with whom they are and want to be both as a person and professional.

In the introduction of the third assignment, “stories to live by,” it was explained that the professional landscape of the teacher can be metaphorically understood as a “storied landscape” (see [Experiences that confirm or reinforce beginning teachers’ identity](#) section). By presenting two stories, one about “teaching in class” and one about “out-of-classroom” work, the teachers described stories of their landscapes that they experienced as important for their work as teacher. In composing the stories, the beginning teachers were asked to use a metaphor. The metaphors mentioned by the teachers were, for example, oceans with raging waves, high mountains with cliffs, and blooming gardens. After each assignment, the teachers explained their written thoughts and, with help of a mentor, explained to each other what the completed assignment meant to them professionally. For example, one fragment from a story about teaching in class: *“Education is like a funfair. It is up to me to guide the students past popular attractions, I have to let them experience several rides, train them and let them attain skills to throw the rings or drive bumper cars. I hope to teach them which attractions are helpful and which distractions they better avoid”*. By creating a personal bond with her students, this beginning teacher helped students to stay focused and active in their learning. The impact of this story on the teacher was positive, because she strongly cared for the wellbeing of her students.

Above program require beginning teachers to think in depth (and to discuss) about issues, challenges or experiences that touch upon their professional identity. They all seem to share a focus on supporting how to become aware of their identity and personal and contextual aspects that influence this identity, which is the first aim of identity work. The second aim, the negotiating part, is hardly supported. Therefore, new program should focus more on the goals the beginning teachers set themselves for becoming the teacher they want to be and which experiences they need in order to pursue own motives, needs or interests in becoming that teacher. Identity work helps beginning teachers to work on personal factors, such as lowering one’s personal standards, but identity work also support them to work on “troubling” contextual factors, such as reducing organizational tasks (for example, organizing school excursions or leading a school musical) when time for teaching seems fading out. Being aware of one’s identity helps the person to stand for oneself by being “agentic” or “resilient.” Literature refers to “being agentic” as taking some control of contextual factors and pursuing personal goals, and to “being resilient” as positively dealing with identity issues ([Etelepälto et al., 2015](#); [Johnson et al., 2014](#)). Teachers, who are socially and emotionally responsive in their professional relationships and who have a personal commitment to teaching, are more likely to transform identity issues into positive challenges that take account of the person ([Johnson et al., 2014](#)). Identity work must support all beginning teachers to figure out whether and how they take their role in finding their own solutions in coping with issues.

Implications

There is a gap between theoretical perspectives on the importance of professional identity development and actual mentoring program for beginning teachers. It is therefore essential to find settings in which beginning teachers are supported to reflect on their professional identities, such as induction program specifically meant for them. Program that support professional identity development need initiatives among which tools that can be used for identity work and direct (peer) discussions about teachers’ experiences that constrain, cultivate or strengthen their professional identity development. In other words, these (peer) discussions are about identity issues and positive experiences. To be more specific, identity issues such as being a full responsible teacher but still having strong feelings of being a student in teacher education and balancing private life and work; positive experiences like having established a good bond with students and being seriously heard by colleagues and school leaders.

Narrative data sources are very helpful for these discussions, such as writing autobiographical stories and demonstrating video diaries ([Stenberg, 2010](#)), drawing and explaining storylines ([Meijer et al., 2011](#)), constructing autobiographical timelines ([Vähäsantanen et al., 2017](#)), keeping digital logs and biographical interviews ([Leeferink et al., 2019](#)). Also specific assignments that lead to narrative data can be applied to stimulate these discussions, for example, choosing and explaining symbols or metaphors like the ones described earlier in this chapter ([Mommers et al., 2021](#)) or “the Professional Body” ([Vähäsantanen et al., 2017](#)). Narrative methods have in common that they direct the user (here the beginning teacher) to take an overall view upon personal experiences and to explicate or (re)define those experiences, and, with or without the help of a mentor, make sense of them. Making sense is triggered by the relational, temporal and situational dimensions found within a narrative, particularly in stories ([Clandinin and Connelly, 2004](#)). The relational dimension pertains to personal and contextual factors that are described, the temporal dimension refers to the continuity of experiences in past, present or future, and the situational dimension refers to the situation in which the experience took place. Narrative methods thus offer a rich starting point for understanding and interpreting events as experienced and told by the teacher ([Clandinin and Connelly, 2004](#)).

The mindset about professional identity, allocated time and even some facilities, such as a place at school, are important to engage in identity work but, unfortunately, not often recognized in school’s practices. Teachers, coaches and school leaders need

to acknowledge that working on a professional identity does not provide hands-on answers to urgent matters during a school day. Working on professional identity supports the teacher to grow as a professional with a well-considered and realistic self-image, who is able to act accordingly and who is recognized by students, their parents, and colleagues as such. Teacher educators and/or coaches in schools need a more explicit and well-developed mindset about the relevance of identity work and learn how to provide beginning teachers with support in this field. Formally organized identity work goes beyond guiding professional knowledge and skills, it also and particularly is about coaching the person to become the teacher one wishes to be and can be. One important goal is that mentors succeed beginning teachers in appreciating to engage in identity work. Next, mentors need to learn how to guide (peer) discussions about revealing one's professional identity by analyzing and interpreting experiences and how to deal with teachers accompanying emotions that these processes may trigger. Important is how to organize tailor made identity program; for example, uncertain teachers need to be triggered to become seriously heard within the organization, innovative teachers must be stimulated to set up new collaborations with colleagues to contribute to school development, and perfectionists can be helped to set more realistic aims in their teaching (cf. Schellings et al., 2021).

In program for beginning teachers, some attention for professional identity development is gradually appearing, yet systematic knowledge about designing *and* evaluating formally organized identity work is still lacking. Important research questions are how and to which extent coaches in schools are currently organizing and doing identity work? What is the impact of formal identity program on beginning teachers' functioning and wellbeing, at the start and further during their careers? In this chapter, some examples of identity program are given, but what are the similarities and differences within these and other identity program and which elements of these program are most powerful in constructing a positive, realistic professional identity? What makes that beginning teachers and their coaches become to understand the value of identity work in addition to developing competent teachers? What takes away the "awkward" feelings or "feelings of resistance" to engage in identity work? What knowledge and skills do mentors need to support identity work with beginning teachers? We emphasized the role of positive experiences during identity work, however the impact of supporting individual qualities on professional identity is less examined.

Answers to such questions help to design induction program in which identity work takes a (more) central part. We suggest that an induction program starts with one or more identity workshops about beginning teachers' experiences and interpretations of these experiences regarding ways in which they reconcile the personal and contextual and what issues or challenges this raises. Ideally, beginning teachers have to be enabled to meet peers who work in different school contexts, which makes the impact of their specific configuration of personal and contextual factors on themselves even more clear. Since professional identity is an ongoing process, identity work is not restricted to the induction phase, it is also valuable at other phases during one's career. Formally organized identity work must be inserted in school practices and policies. Yet formally organized identity work is just a (powerful) tool, in the end it is the (beginning) teacher who has to engage in identity work and experience this as a relevant aspect not only of becoming but also of sustaining and remaining a professional teacher.

References

- Akkerman, S.F., Meijer, P.C., 2011. A dialogical approach to conceptualizing teacher identity. *Teach. Teach. Educ.* 27 (2), 308–319.
- Arava, M., 2016. Building teacher identity through the process of positioning. *Teach. Teach. Educ.* 59 (1), 392–402.
- Beauchamp, C., Thomas, L., 2009. Understanding teacher identity: an overview of issues in the literature and implications for teacher education. *Camb. J. Educ.* 39 (2), 175–189.
- Beijaard, D., Meijer, P.C., 2017. Developing the personal and professional in making a teacher identity. In: Clandinin, D.J., Husu, J. (Eds.), *The Sage Handbook of Research on Teacher Education*. Sage, London, pp. 177–192.
- Beijaard, D., van Driel, J., Verloop, N., 1999. Evaluation of story-line methodology in research on teachers' practical knowledge. *Stud. Educ. Eval.* 25 (1), 46–62.
- Beijaard, D., Meijer, P., Verloop, N., 2004. Reconsidering research on teachers' professional identity. *Teach. Teach. Educ.* 20 (2), 107–128.
- Beijaard, D., Koopman, M., Schellings, G., 2022. Reframing teacher professional identity and learning. In: Menter, I. (Ed.), *The Palgrave Handbook of Teacher Education Research*. Springer, Cham in press and online first.
- Beijaard, D., 1995. Teachers' prior experiences and actual perceptions of professional identity. *Teach. Teach.* 1 (2), 281–294.
- Clandinin, D.J., Connelly, F.M., 2004. *Narrative Inquiry: Experience and Story in Qualitative Research*. John Wiley & Sons.
- Cobb, D.J., 2020. Initial teacher education and the development of teacher identity. In: Peters, M.A. (Ed.), *Encyclopedia of Teacher Education*. Springer, Singapore, pp. 1–5.
- Connelly, F.M., Clandinin, D.J., 1997. Teachers' personal practical knowledge on the professional knowledge landscape. *Teach. Teach. Educ.* 13 (7), 665–674.
- Connelly, F.M., Clandinin, D.J. (Eds.), 1999. *Shaping a Professional Identity: Stories of Educational Practice*. Teachers College Press, New York.
- Etelepälto, A., Vähäsantanen, K., Hökka, P., 2015. How do novice teachers in Finland perceive their professional agency? *Teach. Teach.* 21 (6), 660–680.
- Garner, J.K., Kaplan, A., 2019. A complex dynamic systems perspective on teacher learning and identity formation: an instrumental case. *Teach. Teach.* 25 (1), 7–33.
- Hargreaves, A., 2001. The emotional geographies of teachers' relations with colleagues. *Int. J. Educ. Res.* 35 (5), 503–527.
- Helms-Lorenz, M., van de Grift, W., Maulana, R., 2016. Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *Sch. Effect. Sch. Improv.* 27 (2), 178–204.
- Hoekstra, A., Korthagen, F., 2011. Teacher learning in a context of educational change: informal learning vs. systematically supported learning. *J. Teach. Educ.* 62 (1), 76–92.
- Johnson, B., Down, B., Le Cornu, R., Peters, J., Sullivan, A., Pearce, J., Hunter, J., 2014. Promoting early career teacher resilience: a framework for understanding and acting. *Teach. Teach. Educ.* 20 (5), 530–546.
- Kelchtermans, G., Ballet, K., 2002. The micropolitics of teacher induction. A narrative-biographical study on teacher socialisation. *Teach. Teach. Educ.* 18 (1), 105–120.
- Kelchtermans, G., 2019. Early career teachers and their need for support: thinking again. In: Sullivan, A., Johnson, B., Simons, M. (Eds.), *Attracting and Keeping the Best Teachers, Professional Learning and Development in Schools and Higher Education*, vol. 16. Springer, Singapore, pp. 83–98.
- Kissling, M.T., 2014. Now and then, in and out of the classroom: teachers learning to teach through the experiences of their living curricula. *Teach. Teach. Educ.* 44, 81–91.
- Korthagen, F., Vasalos, A., 2005. Levels in reflection: core reflection as a means to enhance professional growth. *Teach. Teach.* 11 (1), 47–71.
- Lee, I., 2013. Becoming a writing teacher: using "identity" as an analytic lens to understand EFL writing teachers' development. *J. Sec Lang. Writ.* 22 (3), 330–345.
- Leeferink, H., Koopman, M., Beijaard, D., Schellings, G.L.M., 2019. Overarching professional identity themes in student teacher workplace learning. *Teach. Teach. Educ.* 25 (1), 69–89.

- Leijen, Ä., Kullasepp, K., 2013. Unlocking the potential of conflicts: a pilot study of professional identity development facilitation during initial teacher education. *IJDS* 7 (1), 67–86.
- Meijer, P.C., de Graaf, G., Meirink, J.A., 2011. Key experiences in student teachers' development. *Teach. Teach.* 17 (1), 115–129.
- Mommers, J., Schellings, G., Beijard, D., 2021. Professional identity assignments to support beginning teachers' growth into the profession. *J. Teach. Educ. Educ.* 10 (2), 151–173.
- Nichols, S.L., Schutz, P.A., Rogers, K., Bilica, K., 2017. Early career teachers' emotion and emerging teacher identities. *Teach. Teach.* 23 (4), 406–421.
- Olsen, B., 2010. *Teaching for Success: Developing Your Teacher Identity in Today's Classroom*. Paradigm Publishers, Boulder/London.
- Pillen, M., Beijard, D., den Brok, P., 2013. Tensions in beginning teachers professional identity development, accompanying feelings and coping strategies. *Eur. J. Teach. Educ.* 36 (3), 240–260.
- Pinnegar, S., 2005. Identity development, moral authority and the teacher educator. In: Hoban, G.F. (Ed.), *The Missing Links in Teacher Education: Developing a Multi-linked Conceptual Framework*. Springer, Dordrecht, pp. 259–279.
- Rodgers, C.R., Scott, K.H., 2008. The development of the personal self and professional identity in learning to teach. In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D.J. (Eds.), *Handbook of Research on Teacher Education*. Macmillan, New York, pp. 732–755.
- Schaefer, L., Clandinin, D.J., 2019. Sustaining teachers' stories to live by: implications for teacher education. *Teach. Teach.* 25 (1), 54–68.
- Schellings, G., Koopman, M., Beijard, D., Mommers, J., 2021. Constructing configurations to capture the complexity and uniqueness of beginning teachers' professional identity. *Eur. J. Teach. Educ.* <https://doi.org/10.1080/02619768.2021.1905793> online version.
- Stenberg, K., 2010. Identity work as a tool for promoting the professional development of student teachers. *Reflective Pract.* 11 (3), 331–346.
- Vähäsantanen, K., Hökkä, P., Paloniemi, S., Herranen, S., Eteläpelto, A., 2017. Professional learning and agency in an identity coaching programme. *Prof. Dev. Educ.* 43 (4), 514–536.
- Vähäsantanen, K., Hökkä, P.K., Paloniemi, S., 2020. University teachers' professional identity work and emotions in the context of an arts-based identity coaching program. In: McKay, L., Barton, G., Garvis, S., Sappa, V. (Eds.), *Arts-Based Research, Resilience and Well-Being Across the Lifespan*. Palgrave Macmillan, Cham, pp. 233–256. Cham.
- van der Wal, M.M., Oolbekkink-Marchand, H.W., Schaap, H., Meijer, P.C., 2019. Impact of early career teachers' professional identity tensions. *Teach. Teach. Educ.* 80 (1), 59–70.
- van der Want, A.C., Schellings, G.L.M., Mommers, J., 2018. Experienced teachers dealing with issues in education: a career perspective. *Teach. Teach.* 24 (7), 802–824.
- van Rijswijk, M.M., Bronkhorst, L.H., Akkerman, S.F., van Tartwijk, J., 2018. Changes in sensed dis/continuity in the development of student teachers throughout teacher education. *Eur. J. Teach. Educ.* 41 (3), 282–300.
- Zeichner, K.M., Gore, J.M., 1990. Teacher socialization. In: Houston, W.R. (Ed.), *Handbook of Research on Teacher Education*. MacMillan, New York, pp. 329–348.
- Zembylas, M., 2003. Emotions and teacher identity: a poststructural perspective. *Teach. Teach.* 9 (3), 213–238.
- Zwart, R.C., Korthagen, F.A., Attema-Noordewier, S., 2015. A strength-based approach to teacher professional development. *Prof. Dev. Educ.* 41 (3), 579–596.

Teacher expertise and how it develops during teachers' professional lives

Jan van Tartwijk^a, Esther E. van Dijk^{a,b,c}, Johan Geertsema^d, Manon Kluijtmans^{b,c}, and Marieke van der Schaaf^{c,e},

^a Department of Education, Faculty of Social and Behavioral Sciences, Utrecht University, Utrecht, Netherlands; ^b Centre for Academic Teaching, Utrecht University, Utrecht, Netherlands; ^c Center for Education, University Medical Center Utrecht, Utrecht, Netherlands;

^d University Scholars Programme and Ridge View Residential College, National University of Singapore, Singapore; and ^e Utrecht Center for Research and Development in Health Professions Education, University Medical Center Utrecht, Utrecht, Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	170
The nature of expertise from a cognitive-psychological perspective	170
Experts' use of memory resources	171
Expert teachers use of memory when dealing with classroom situations	171
The development of expertise	172
Experience and deliberate practice	172
Experience	172
Deliberate practice	172
The development of teacher expertise	173
Stages in the development of teachers as professionals	173
The development of teacher expertise for specific tasks	174
Deliberate practice and reflection in teacher education and teacher professional development	174
Adaptive expertise	175
Implications for teacher education and teacher professional development	176
Developing routine and adaptive expertise	176
Practicing for adaptive expertise	176
An alternative perspective	177
References	177

Introduction

The knowledge, skills, and characteristics of individuals that determine the quality of how they perform tasks, is referred to with the concept of "expertise" by Ericsson and his colleagues (Ericsson, 2018b). Ericsson defines as "experts" in a domain, those who are consistently able to exhibit superior performance for tasks that are representative for that domain. Ericsson et al. use a cognitive psychological perspective when studying expertise and its development. A vast amount of research has been carried out on expertise and its development from this perspective. This research is summarized in a series of handbooks that were published over the years (Chi et al., 1988; Ericsson et al., 2006, 2018; Ericsson, 2009; Ward et al., 2020). Much of this research focuses on what sets apart experts from others who are not that proficient (yet) in different domains such as chess, music, sport, computer programming, and radiology. Although the attention for expertise and expert performance of teachers is limited to one article in one of these handbooks (Stigler and Miller, 2018), this literature is important for a better understanding of the nature of teacher expertise and how it develops. In this contribution, we will highlight some findings of this research and apply it to teacher expertise and its development.

Despite its contributions, research on expertise from a cognitive psychological perspective has also been criticized for several reasons. A first reason is that this research tends to focus on expert performance in tasks which are relatively stable in time and across contexts, with relatively easy to measure and unambiguous outcomes. Examples of such tasks are performance in chess and playing a musical instrument. However, in most domains, like teaching, experts will be frequently confronted with new insights or changes in the context in which they carry out their tasks. Those people who can respond well to these changes are referred to as "adaptive experts" and are distinguished from "routine experts". They have a set of core competencies that they apply throughout their lifetime with greater and greater efficiency (Bransford et al., 2005; Hatano and Inagaki, 1986). In this contribution, we will elaborate on the differences between adaptive and routine experts. We will also discuss the implications for teacher education and teacher professional development.

The nature of expertise from a cognitive-psychological perspective

Just after the second world war, the Dutch chess master and psychologist Adriaan De Groot finished his thesis (De Groot, 1948/1965). In this thesis, he described experiments in which grand masters of chess and proficient amateurs in the game were asked to recall realistic and random board positions. The grand masters turned out to be much better than the amateurs in remembering

the realistic board positions, but differences were small when the chess pieces were randomly placed on the board. From his experiments, De Groot concluded that, contrary to popular beliefs, may be it was not thinking ahead that made of most of the difference between grand master and master chess players and others, but perception and knowledge of board positions.

Gobet and Charness (2018) and many others regard De Groot's experiments as the starting point for modern-age research on expertise from a cognitive-psychological perspective. In more than 70 years of research on expertise from this perspective, much has been learned about the general mechanisms underlying expert performance and what distinguishes experts from others. Also, a large body of literature is available on how expertise develops and what can be done to stimulate this development.

Experts' use of memory resources

When experts are defined as those who perform the tasks in their domain more effectively and efficiently than others, as a consequence the research focus will be on what distinguishes them from others. When taking a cognitive psychological perspective, research will concentrate on the way in which experts structure information in their memory and use, and by doing so circumvent the limitations of human memory (Feltovich et al., 2018). To understand those limitations, it is first important to know more about how memory works.

Human memory consists of different memory systems. Baddeley et al. (2015) distinguish between two main systems: Long-term memory and working memory. Within long-term memory, two subsystems can be distinguished. The first subsystem is the explicit or declarative memory, from which people can retrieve facts or (conscious) memories of events. The second subsystem is the implicit or non-declarative memory that forms the basis for automated skills such as reading, walking or driving. The second memory system is the working memory. People use this part of the memory to perform cognitively complex tasks, such as reasoning, learning and understanding. A first subsystem of the working memory is the visuo-spatial sketchpad, where people temporarily store visual, spatial information. A second subsystem within working memory is short-term memory, in which people temporarily keep small amounts of information required for a cognitive task.

A first and fundamental limitation of human memory is the capacity of short-term memory. Humans can only keep about seven units active in it (Miller, 1956). That means that a series of numbers such as 1 4 9 2 0 1 9 6 9 (nine units) is almost impossible to remember for most of us without specific training. A work-around to overcome this limitation is clustering units into a smaller number of meaningful units, to which Baddeley et al. refer as "chunks". In our example, meaningful chunks could be years: 1492 0 1969. These years are easier to remember if meaning is attached to them: 1492 is the year in which Columbus was the first European (on record) to set foot on American soil; 0 is the beginning of our era; and 1969 is the year the first man set foot on the moon.

Unlike short-time memory, human long-term memory has an almost unlimited capacity in practice, but here the limitation is that it is sometimes difficult to retrieve information quickly. The way in which information is organized in long-term memory can remedy this. Chess grandmasters, for example, perceive positions of chess pieces on the board as one configuration that they recognize in a fraction of a second from the tens of thousands of configurations they have stored in their long-term memory as chunks (Gobet and Charness, 2018). In addition, they have developed the ability to immediately see when a chess piece has an exceptional position that presents an opportunity or threat, which requires action. Experts have highly automated the recognition of such configurations (Feltovich et al., 2018). How this knowledge is structured in memory and how these knowledge structures develop, has been extensively investigated in the domain of medicine (Boshuizen et al., 2020; Schmidt and Rikers, 2007). In this research, it was established that the relationship between case recall and level of expertise is not linear but curvilinear: case recall and the use of biomedical and clinical details in case descriptions increase during the first years of training (up to 6 years), but after that point it decrease with experience. This referred to as the "Intermediate effect" (Schmidt and Boshuizen, 1993). Boshuizen et al. (2020) describe how this decrease in recall of details is accompanied by an increasing use of macro-concepts with experience, for which the term "knowledge encapsulation" is used: "a prolonged process of both abbreviation of reasoning lines and integration of new knowledge parts under the same umbrella term." (p. 3). They suggest that chunking in chess memory but also step-skipping in problem solving are manifestations of knowledge encapsulation under the umbrella of such macro-concepts.

Expert teachers use of memory when dealing with classroom situations

For teachers, just like other professionals, it is essential to avoid working memory overload. This is not easy in the everyday classroom. Doyle (2006, 1986) for instance describes the classroom as a place where twenty to thirty students get together with one teacher. The teacher has the task to help all students achieve learning goals by planning and monitoring classroom activities. Doyle identified a number of features of the classroom that are relevant for how teachers carry out this task. A first feature is that a lot of things happen at the same time in the class. He describes how a teacher, while helping an individual student during seatwork, must monitor the rest of the class, acknowledge other requests for assistance, and keep track of time. Also, during a discussion, a teacher must listen to student answers, watch other students for signs of comprehension or confusion, formulate the next question, and scan the class for possible misbehavior. A second feature is the rapid pace of classroom events. The average teacher has over 500 interactions with individual students in a single day, Doyle reports. Because so many things that demand attention happen at the same time and with high speed, the limited working memory capacity of a teacher is easily exceeded. For that reason, novice teachers often feel overwhelmed. Their more experienced colleagues, however, seem to be able to deal with the demanding classroom situations relatively easy.

Early experimental research, comparing how expert teachers, advanced beginners and novices made sense of hectic classroom situation, was carried out by Sabers et al. (1991). The expert teachers in these experiments were very carefully selected to ascertain that they were indeed experts. The researchers compared the skill of these three groups of teachers in recognizing important information in realistic classroom situations. The classroom situations were videotaped from different angles and were shown simultaneously on three monitors with the sound coming from the middle monitor. In comparison with novices and beginners, expert teachers scanned the three monitors more frequently whereas the novices and beginners focused on the middle monitor. The expert teachers turned out to be better able to monitor, understand, and interpret simultaneous, multiple events within a classroom. However, when comparing the three groups of teachers, the differences in how many specific details they remembered were minor. The analogy with the chess players involved in the research by De Groot, that we described earlier, is striking. Both chess masters and expert teachers have developed more effective and efficient routines in recognizing and interpreting important situations than novices and beginners, but the difference is much smaller for insignificant details.

Similar studies comparing experts and novices were carried out by Van den Bogert, Wolff and colleagues using eye-tracking techniques. In line with the earlier experiments by Sabers et al. (1991), these studies show that expert teachers process visual information faster, and consistently check up on pupils more regularly (Van den Bogert et al., 2014). The expert teachers monitored more areas than novices and focused their attention on areas where relevant information was available. Also, their perception of events appeared to be more knowledge-driven whereas novices' perception appeared more image-driven (Wolff et al., 2016). Both Van Tartwijk et al. (2009) and Wolff et al. (2015) focused on expert teachers' knowledge that guides their classroom management strategies using teachers' verbalization of classroom events visible on video recordings of classroom events. Wolff et al. found that experts focused more on learning in the classroom and the teacher's ability to influence learning, and novices were more concerned with maintaining discipline and behavioral norms when watching a series of clips. Van Tartwijk et al. (2009) asked teachers who were identified as experts in classroom management in multicultural classes to comment on the classroom management strategies they used in their own videotaped lessons. The teachers were aware of the importance of providing clear rules and correcting student behavior whenever necessary, but they also wanted to reduce potential negative influences of corrections on the classroom atmosphere. They aimed at developing positive teacher-student relationships and adjusted their teaching methods anticipating students' responses. The classroom management strategies seemed to be guided by script-like if-then expectations about student responses which are likely to be based on previous experiences (Baldwin, 1992).

The development of expertise

Experience and deliberate practice

Experience

People build a knowledge base in their long-term memory through experience, but that is not enough to reach expert level. For example, after taking driving lessons for a specific period and obtaining their driving license, most people no longer pay much attention to further improving their driving skills. Of course, they build up routine by making the necessary mileage, but their development as a driver will stabilize over time. In a similar vein, teachers may start working after completing teacher education at an acceptable level of proficiency, and even though they are likely to develop routine and become more efficient, they are unlikely to improve their performance after a certain level by experience alone. Day (2008) found that most teachers maintain their effectiveness throughout their professional lives, but do not necessarily become more effective over time. He even found that teachers in later years of their professional lives are at greater risk of becoming less effective. This has also been found in the medical domain, where Choudhry et al. (2005) found that the performance of many medical doctors declines with experience. The next section explores how this stagnation or decline in performance can be countered according to researchers using a cognitive psychological perspective.

Deliberate practice

Experts reach their high level in task performance by extending their active period of expertise development by years or even decades (Ericsson et al., 2006, p. 691). Professional musicians and top athletes do this by engaging in series of highly structured activities with increasing complexity designed by a coach, with the aim of improving performance. In between meetings with the coach, they have to monitor their performance and make adjustments to meet the practice goals. This requires concentration. During sessions with the coach, feedback is provided, and new activities are introduced. By going through such training cycles, they postpone full automatization. The term "deliberate practice" was coined by Ericsson and his colleagues for this cyclical process of improvement (Ericsson et al., 1993; Ericsson, 2018a). As a result, experts in domains such as music and sports do not only develop routines that enable them to perform tasks efficiently, but the level of their effectiveness is much higher than that of others. Ericsson (Ericsson, 2018a) illustrates how the course of improvement may differ between experts and others for everyday activities (see Fig. 1).

Ericsson writes that the goal for most individuals is to reach a satisfactory level that is stable for everyday skills. After a period of practicing an activity, its performance will become virtually automatic (the light gray end of bottom arrows of the graph in Fig. 1). This means that a minimal amount of cognitive effort is invested in performance. Experts, however, postpone automatization and remain in phases of deliberate practice. These cognitive efforts result in the development of increasingly complex mental

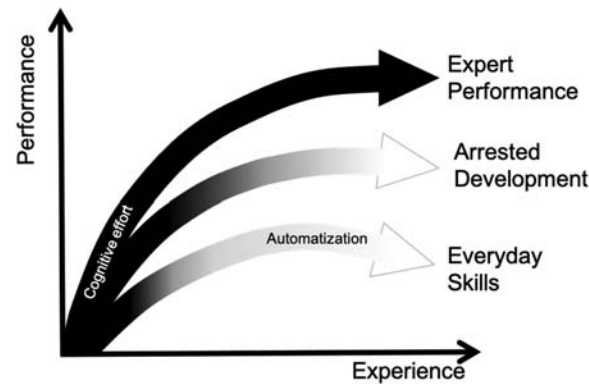


Fig. 1 An illustration of the qualitative difference between the course of improvement of expert performance in everyday activities. Adapted from Ericsson (1998).

representations of tasks and high levels of control of their performance. The arrow in the middle depicts the development of the group that “gave up” practicing, which results in premature automatization. Again, we use learning to drive a car as an example of learning to master an everyday skill. Most of us who have taken driving lessons will remember the first lessons as incredibly complicated, because our working memory capacity had to be divided between paying attention to traffic around us (including the traffic behind us, visible in the mirrors), steering, operating the direction indicator light, controlling the gas and brakes pedals, and -for the Europeans among us-handling the manual transmission at the right moment. All of that is accompanied by a driving instructor giving distracting commands. After a while, however, most of these tasks become automatized and require far less working memory capacity. In this autonomous phase, it is even possible to have a relaxed conversation with a passenger. Because the quality of this routinized task performance is not “maintained”, it may even decrease. Professional formula one drivers need to postpone automatization to reach the high levels of expert performance that are necessary to participate at the top level of their sport. This requires years of deliberate practice.

The concept of deliberate practice has sparked a lot of interest in the popular literature for expert performance (cf. Gladwell, 2008), but cannot be applied to most professional tasks. One reason for this was put forward by Macnamara et al. (2014). Based on a systematic review of the empirical literature, they argue that the association between deliberate practice and performance has been primarily researched in domains where the predictability of the task environment is high (i.e., the task environment does not change much while the performer is planning and executing an action). Examples of such domains are music and sports, but this association is much lower in domains such as education and other professions where the predictability of the task environment is much lower. Another reason is that professional musicians and top athletes can practice and will get feedback before they perform the tasks in a concert or competition. However, professionals in other areas, such as teachers, have far fewer opportunities to systematically practice and receive feedback before carrying out their tasks. Fadde and Klein (2010) proposed the term “deliberate performance” as an alternative for deliberate practice when professionals systematically work on improving performance in the workplace. They define deliberate performance as “the effort to increase domain expertise while engaged in routine work activity” (p. 6) and suggest that it includes developing exercises that use every-day job tasks instead of training exercises (without adding excessive time) and do not require expert judgment for feedback. Professionals themselves analyze their performance.

The development of teacher expertise

Stages in the development of teachers as professionals

In the previous section, we gave the example of how difficult it is to learn how to drive a car. It requires lots of working memory capacity to be able to execute all the tasks at the same time. After a while, the execution of tasks becomes automatized, and driving can even be relaxing. This course of improvement which people go through when learning to drive a car, resembles the stage that those of us who are (or have been) teachers in primary or secondary teachers may have experienced in the first stage of teaching when we had to learn to deal with the hectic classroom where so many things happen at the same time, which Doyle (2006) so eloquently described.

There is a long tradition of research which aims to describe the development of teachers' expertise throughout their professional lives. Much of this research was inspired by the work of Fuller and her colleagues (Fuller, 1969; Fuller and Bown, 1975), who identified several stages in the concerns teachers have about their development and functioning throughout their professional lives. As a first stage, they distinguish the period before prospective teachers start teaching. In this stage, they are primarily concerned with their own progress as students and do not yet think from a teacher's perspective, but from a student's perspective. In this stage, they can be very critical of teachers. However, once they have started teaching, the second stage, most of these beginning teachers' concerns are related to their own survival in the hectic classroom. In this second stage, beginning teachers for instance focus on class control and being liked by students and on the quality of the way they teach the class. Subsequently, a mastery stage can be

distinguished. In this phase the concerns of novice teachers focus on being able to perform well as a teacher. Concerns are less emotionally colored than in the survival phase and there is room to think about the best pedagogical approach. In the final stage, teachers' concerns focus more on student learning and development. Authors such as Fessler and Christensen (1992), Day and Gu (2010) and Van der Lans et al. (2017) have built on the work of Fuller and her colleagues.

The development of teacher expertise for specific tasks

All studies using the conceptual framework of stages in the development of teacher expertise, have in common that classroom management, including developing a positive working relationship with students, is teachers' most important concern in the first stages of their professional lives. Being able to manage a classroom at an acceptable level is a necessary condition for being able to carry out other teaching activities such as instructing a class. In other words, it is logical that stages can be distinguished in the development of teacher expertise. However, a pitfall when conceptualizing the development of teachers' expertise in terms of stages, is that it might easily lead to the misconception that once a teacher has mastered a routine that requires most attention in specific stage, there is no need to be concerned about this routine anymore and it is time to move on to the next stage. When a teacher would no longer pay attention to the further development of a routine, the risk is that they will not develop it anymore and their level of expertise won't increase anymore, or they may even become less effective later in their professional life (Day, 2008).

Research on the development of teacher expertise related to specific routines throughout teachers' professional life, is relatively scarce. An exception is research by Mieke Brekelmans and her colleagues, who gathered data on the development of teacher expertise in being able to build a positive interpersonal relationship with students (Brekelmans et al., 2005). Being able to build such a positive working relationship is an important component of classroom management (Wubbels et al., 2006). Brekelmans et al. used cross-sectional data among the students of 2388 student-teachers and 3813 Dutch secondary school teachers with 1–36 years of experience. They also gathered longitudinal data among a limited number of teachers (86 teachers for at least 4 years, 37 teachers for 6 years or more). These data were gathered with the Questionnaire on Teacher Interaction (QTI, for an overview see Wubbels et al., 2006) which were completed by students of these teachers by indicating if statements such as "This teacher is a good leader", "This teacher is uncertain", or "This teacher is friendly" applied to their teacher. Underlying this questionnaire is interpersonal theory (cf. Horowitz and Strack, 2011). At the core of this theory is the premise that two independent dimensions are both necessary and sufficient to describe the interpersonal meaning of behavior and the position of individuals in interpersonal relationships (e.g. Leary, 1957; Sadler et al., 2011). Depending on the context (e.g., psychiatry, family, education) the dimensions are designated differently with agency and communion as meta-level labels (Gurtman, 2009). Agency refers to influence and control, whereas communion refers to friendliness and affiliation. A classroom where the social climate is characterized by low levels of teacher agency combined with low levels of communion between teacher and students, is usually a classroom where the teacher has serious problems with classroom management. Brekelmans et al. (2005) established that, during the first 3 years of teaching, mean students' perceptions of teacher influence significantly increased every successive year. After these 3 years, mean differences between years got smaller but still increased until about 6 years of experience, and then stabilized for most teachers. Their analyses showed however, that there was significant individual variation around the mean. Compared with students' perceptions of agency, the students' perceptions of their teachers' interpersonal styles in terms of communion showed far less variation across the teaching career. On average perceptions of communion slightly increased during the first 10 years of teaching and became slightly lower afterward. Differences were large between teachers, but the variations across years for specific teachers were small. These findings suggest that the development pattern of teacher agency during the professional life of teachers does match the stage models described earlier, with a strong development in the first stages of the career, but also resembles the pattern in the development of expertise in other domains with a strong increase of expertise in the first period of professional practice, a slower increase later, and then a stabilization.

Note that the teachers involved in the Brekelmans-study were not selected as experts. In this respect they differ from the expert-teachers involved in the studies to which we referred earlier (Sabers et al., 1991; Van den Bogert et al., 2014; Van Tartwijk et al., 2009; Wolff et al., 2015; Wolff et al., 2016; Bronkhorst et al., 2014) who turned out to be very skilled in monitoring, understanding, and interpreting multiple events within a classroom.

Deliberate practice and reflection in teacher education and teacher professional development

In the domain of teacher education and teacher professional development, the term "deliberate practice" has not been used much to describe the systematic efforts of (prospective) teachers to improve their teaching. One exception is the work of Bronkhorst and her colleagues (Bronkhorst et al., 2011, 2014) who defined deliberate practice as "deliberated activities intended for self-improvement, but also intended to foster pupil learning and development." (p. 20). Such activities are realistic tasks in the workplace instead of training tasks, and therefore this conceptualization of deliberate practice resembles the concept of deliberate performance that Fadde and Klein propose as an alternative for deliberate practice in the workplace (Fadde and Klein, 2010). Bronkhorst et al. describe four characteristics that define deliberate practice: (1) practice should be designed for self-improvement, (2) practice should be repeated to enable successive refinement, (3) practice should be followed by immediate, informative feedback, and (4) practice requires significant effort and concentration. According to these authors, designing such learning activities can be regarded as enacting a (personal) theory of practice. They refer to this as "anticipatory reflection" which they distinguish from "retrospective reflection".

The concept "reflection" is far more common in the domain of research on teachers and their development. Although the concept of reflection has a rich history and many definitions, a common essence is that reflection processes contribute to a learner's development (Dewey, 1933; Van Manen, 1977; Schön, 1983; Poldner et al., 2014). After describing the relevant literature and identifying the various ways in which the concept of reflection is used, Korthagen et al. (2001) conclude that there seems to be a consensus that teachers' professional behaviors are rooted in mental structures which are not static but can be created or altered through experiences and influenced by reflection. They then define reflection as "the mental process of trying to structure or restructure an experience, a problem, or existing knowledge or insights." (p. 58). The process is of a cyclical nature and based on learning activities that are carried out by the actor or subject (e.g., a teacher) during the reflection process. Examples of learning activities are selecting (e.g., certain information), analyzing, synthesizing, evaluating, concluding, and intending (e.g., setting new goals) (Oosterbaan et al., 2010). The outcome of this process, e.g., new insights, beliefs or plans for action, is an indication for learning. Reflection is always related to an "object", e.g., reflection on one's development, beliefs or performance. The object of reflection can also involve critical thinking about the premises upon which a teacher has based his or her analyses or view. By means of reflection, these objects move or transform in several ways to support development, e.g., certain experiences are restructured resulting in transformed cognitive schemes, new perspectives on personal and professional effectiveness, or (tacit) intentions for acting.

Hatton and Smith (1995) distinguish five levels in the object of reflection that they related to the development of concerns during teachers' professional lives that Fuller distinguishes, which we introduced earlier. The first level which Hatton and Smith distinguish is "technical rationality", where the focus is on one's own technical use of essential skills. For the next levels, they build on the work of Donald Schön (1983, 1987), who distinguished between professionals' "reflection-on-action", which refers to looking back on task performance sometime after it has taken place, and "reflection-in-action" which refers to conscious thinking and modification while carrying out a task. Hatton and Smith refer to reflection-on-action as the second level of reflection and write that this type of reflection focusses on the impact of action. Within the level of reflection-on-action, they distinguish three sublevels. At the lowest sublevel, to which Hatton and Smith refer as the "descriptive" sublevel, the focus is on the best possible practice and one's own performance in the professional role. At the second "dialogical" sublevel, teachers reflect on alternative solutions for professional problems. During reflection at the third and highest sublevel, to which they refer as "critical reflection", teachers reflect on the effects of their actions on others taking while taking social, political and cultural "forces" into account. Hatton and Smith consider reflection-in-action as the highest type of reflection because they consider the application of technical reflection and reflection-on-action as a situation is unfolding as the most advanced form of reflection.

Adaptive expertise

Whereas dialogical and critical reflection seem to pertain to thinking activities that facilitate adaptation or change of professional routines, the concepts deliberate practice and deliberate performance refer to conscious practice with the goal of highly efficient and effective routines for task performance. For beginning teachers, developing efficient and effective routines is their first concern, because in the highly demanding task environment of the classroom, it is a prerequisite for professional survival.

Routines, however, have major disadvantages. A first disadvantage is that routines may become so intuitive and tacit that experts don't even realize that they are part of their repertoire (Bransford et al., 2005). It is difficult for many expert teachers to articulate in detail what the critical clues are which they perceive in class situations why they act in the way they do. This is a problem for two reasons. The first reason is that this makes it much harder for them to teach beginning teachers how to teach if they can't articulate why they do what they do. The second reason is that it makes it much harder to change teaching routines if a teacher is hardly aware of them. Researchers very often need special techniques to get teachers to make their teaching strategies and the underlying knowledge explicit, such as stimulated recall (Yinger, 1986), asking teachers to comment on their own classroom management strategies visible on video (Van Tartwijk et al., 2009) or concept mapping (Vink et al., 2016).

Here, it is important to distinguish types of experts that have been mentioned in the literature: routine and adaptive experts (Bransford et al., 2005; Hatano and Inagaki, 1986; National Research Council, 2000). *Routine experts* have a set of core competencies that they develop throughout their lives with growing efficiency, whereas *adaptive experts* are much more likely to change their core competencies and expand and restructure their expertise. According to Bransford et al. this restructuring may reduce efficiency in the short run, but makes adaptive experts more flexible. They emphasize that these processes of restructuring "often have emotional consequences that accompany realization that cherished beliefs and practices need to be changed" (Bransford et al., 2005, p. 45).

A review by Bohle Carbonell et al. (2014) identified a number of differences between routine and adaptive experts. The first of these differences is that, although adaptive and routine experts have a similar extent of knowledge, the knowledge of adaptive experts seems to be more theoretical. Adaptive experts also seem to have more cognitive flexibility and more problem-solving skills. They rely more on analogical reasoning in which they use their knowledge base. According to Chi (2011), this more abstract knowledge base helps them understand the principles and mechanisms that determine why and how routines work. "With such deeper understanding, adaptive experts obviously can "generalize" their skills to other non-routine problems." (Chi, 2011).

The National Research Council (2000) of the USA discusses adaptive expertise as the golden standard for learning. A framework for promoting the development of adaptive expertise in teacher education and professional development is crucial, according to the council.

Implications for teacher education and teacher professional development

Developing routine and adaptive expertise

Referring to them as experts, implies that both routine and adaptive experts are consistently able to exhibit superior performance for tasks that are representative for that domain compared with others, and that they remain in the cognitive/associative phase longer than others and by doing so avoid premature automatization of routines (Ericsson, 2018a). However, there are important differences between how routine experts and adaptive experts develop expertise.

The way in which routine experts develop their expertise, prepares them for performing tasks in domains that do not change. To develop routine expertise deliberate practice is appropriate. Earlier in this article, we have described the process of developing expertise through deliberate practice. Examples of techniques to which the definition of deliberate practice would apply (a cyclical process of consciously and specifically training to improve performance, often with the help of a coach) are video interaction guidance (e.g. Fukkink and Tavecchio, 2010), in ear-coaching (e.g. Coninx et al., 2013), and providing feedback using augmented reality techniques (e.g. Theelen et al., 2020).

Bransford et al. (2005) emphasize that a good balance between the development and renewal of routines is important for the development of adaptive expertise at all levels. Too much emphasis on innovation at a time when a teacher is not ready, due to the lack of routine, can lead to frustration. Too much emphasis on mere routine can result in the teacher no longer being able to adjust his or her task performance properly if necessary. The zone in which the optimal balance has been found between gaining routine and renewal is referred to as the optimal innovation corridor in the literature on adaptive expertise. In Fig. 2, which is adapted from Bransford et al. this is depicted. In this figure, we added the stairs-like shape in the optimal innovation corridor. By inserting this shape, we aimed to visualize the stages in the development of (prospective) teachers: they first need to master a specific routine at a satisfactory level before they can spend their "limited working memory capacity" on the development of other routines. The challenge is, of course, not to prematurely automatize the first routine when the focus shifts to another. This requires reflection in the way it was defined by Korthagen et al. (2001).

Bransford warns that reflection on routines once they are mastered can have emotional consequences. Instead of simply training to improve routines (prospective) teachers are asked to be critical about the routines and the underlying assumptions that they and their colleagues use in practice. This emphasizes that it is crucial to stay within the optimal innovation corridor depicted in Fig. 2 to avoid frustration beyond a level that is unproductive. For this reason, Korthagen proposes to organize the curriculum for teacher education in such a way, that it can be tailored to student teacher concerns (Korthagen et al., 2001; Brouwer and Korthagen, 2005). Of course, this is also important advice when designing professional development for (beginning) teachers.

Practicing for adaptive expertise

In a recent review of the literature, Kua et al. (2020) provide practical advice for courses, programs and professional development that aim to stimulate the development of adaptive expertise. They provide four design strategies that emerged from their analyses of the literature on such interventions: (1) present abstract materials early in the training, (2) facilitate the integration of conceptual ideas with existing knowledge, (3) guide discovery with hypothesis testing and problem solving, and (4) provide instruction on metacognitive processes such as reflection.

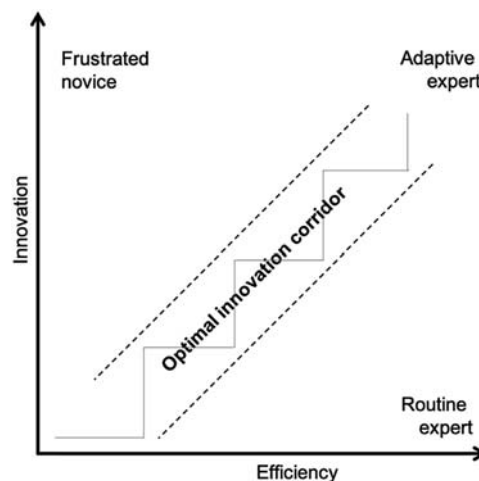


Fig. 2 The dimensions of adaptive expertise and trajectories for development. Adapted from Bransford et al. (2005).

Based on a review of the literature as well, Ward et al. (2018) also provide a list of principles for training adaptive skills: (1) provide flexibility-focused feedback, (2) relate concepts to cases in such a way that they illustrate the structural patterns in situations, (3) present tough cases with limited time available for solving the problem, (4) provide cases that are tailored to current needs and proficiency levels, (5) preserve the complexity of the situation in which learning takes place, and (6) promote reflective practice prior to, during and following task performance.

Both the work of Korthagen et al. and the lists provided by Kua and Ward and their colleagues seem to focus on providing theory early but in connection with the context where tasks need to be carried out. This should be done in such a way that learners recognize the structure of the situation which also makes theory more meaningful. Guided discovery and hypothesis testing, solving problems presented in cases, and reflection on cases, all seem to be strategies to train learners in analyzing situations in an inquiry-based manner. Finally, Bransford, Korthagen and Ward all emphasize that providing opportunities to develop (adaptive) expertise should be challenging for learners, but that unproductive frustration because tasks are too hard should be avoided.

An alternative perspective

The literature and the research finding presented in this article until now, all focused on the development of expertise in individuals. In much of this research, task performance of novices and experts is compared. The implicit assumption on which this literature is based, is that expertise is the realm of the individual who strives for the highest possible level of performance in sport, music or in professions in domains such as health care or education.

This assumption is contested by scholars such as Engeström (2018). Focusing on the cognitive foundation of expertise alone ignores that thinking and knowledge are embedded in social situations, practices, and cultures, he asserts. Referring to the work of scholars such as Lave and Wenger (1991), he suggests that expertise should not be conceived as a quality of an individual but that “the proper unit of analysis of skilled human activity is a community of practice rather than an isolated individual.” (Engeström, 2018). Also, Engeström refers to the work of Bereiter and Scardamalia (1993), when he proposes that expertise should be reconceptualized as a process of progressive problem-solving. This resonates with a plea by Ward et al. (2018), who write that being able to adapt routines should be a condition “sine qua non” for designating someone as an expert. Engeström proposes activity theory as an alternative framework for conceptualizing expertise. In this framework, expertise is conceived to reside in an activity system which a heterogeneous group of subjects form together, which is oriented at an object aiming to generate a certain outcome, and which is embedded in a larger community, with a division of labor and specific rules.

It is beyond the scope of this contribution to explore this perspective further, but we introduce it here to point at the importance of other perspectives on expertise than the highly individualized cognitive psychological perspective which also tends to ignore context (Stigler and Miller, 2018). The common practice now is that teachers tend to perform their primary tasks in a classroom setting where they are alone with over twenty students. Their expertise is related to their task performance in that particular context. However, that context is not self-evident: it may for instance differ between urban, suburban and rural schools or between countries. Also, new developments, for instance in technology, may fundamentally change how education is organized. It is the shared responsibility of teachers and others involved in the action system aimed at providing education for all students, to use new opportunities for improving practice in various contexts.

References

- Baddeley, A., Eysenck, M.W., Anderson, M.C., 2015. Memory. Psychology Press, New York.
- Baldwin, M.W., 1992. Relational schemas and the processing of social information. *Psychol. Bull.* 112, 461–484.
- Bereiter, C., Scardamalia, M., 1993. *Surpassing Ourselves: An Inquiry Into the Nature and Implications of Expertise* Chicago. Open Court.
- Bohle Carbonell, K., Stalmeijer, R.E., Könings, K.D., Segers, M., Van Merriënboer, J.J.G., 2014. How experts deal with novel situations: a review of adaptive expertise. *Educ. Res. Rev.* 12, 14–29.
- Boshuizen, H.P.A., Gruber, H., Strasser, J., 2020. Knowledge restructuring through case processing: the key to generalise expertise development theory across domains? *Educ. Res. Rev.* 29, 100310.
- Bransford, J., Derry, S., Berliner, D.C., Hammerness, K., Beckett, K.L., 2005. Theories of learning and their roles in teaching. In: Darling-Hammond, L., Bransford, J., Lepage, P., Hammerness, K., Duffy, H. (Eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, San Francisco.
- Breklemans, M., Wubbels, T., Van Tartwijk, J., 2005. Teacher-student relationships across the teaching career. *Int. J. Educ. Res.* 32, 55–71.
- Bronkhorst, L.H., Meijer, P.C., Koster, B., Vermunt, J.D., 2011. Fostering meaning-oriented learning and deliberate practice in teacher education. *Teach. Teach. Educ.* 27, 1120–1130.
- Bronkhorst, L.H., Meijer, P.C., Koster, B., Vermunt, J.D., 2014. Deliberate practice in teacher education. *Eur. J. Teach. Educ.* 37, 18–34.
- Brouwer, N., Korthagen, F.A.J., 2005. Can teacher education make a difference? *Am. Educ. Res. J.* 42, 153–224.
- Chi, M.T.H., Glaser, R., Farr, M.J. (Eds.), 1988. *The Nature of Expertise*. Lawrence Erlbaum, Hillsdale, NJ.
- Chi, M.T.H., 2011. Theoretical perspectives, methodological approaches, and trends in the study of expertise. In: Li, Y., Kaiser, G. (Eds.), *Expertise in Mathematics Instruction*. Springer, New York.
- Choudhry, N.K., Fletcher, R.H., Soumerai, S.B., 2005. Systematic review: the relationship between clinical experience and quality of health care. *Ann. Intern. Med.* 142, 260–273.

- Coninx, N., Kreijns, K., Jochems, W., 2013. The use of keywords for delivering immediate performance feedback on teacher competence development. *Eur. J. Teach. Educ.* 36, 164–182.
- Day, C., Gu, Q., 2010. *The New Lives of Teachers*. Routledge, New York.
- Day, C., 2008. Committed for life? Variations in teachers' work, lives and effectiveness. *J. Educ. Change* 9, 243–260.
- De Groot, A.G., 1948/1965. *Thought and Choice in Chess*. Mouton Publishers, The Hague.
- Dewey, J., 1933. *How We Think: A Restatement of the Relation between Reflective Thinking to the Educative Process*. Heath, Boston.
- Doyle, W., 1986. Classroom organization and management. In: Wittrock, M.C. (Ed.), *Handbook of Research on Teaching*, third ed. McMillan, New York.
- Doyle, W., 2006. Ecological approaches to classroom management. In: Evertson, C.M., Weinstein, C.S. (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Engeström, Y., 2018. *Expertise in Transition: Expansive Learning in Medical Work*. Cambridge University Press, Cambridge.
- Ericsson, K.A., Krampe, R.T., Tesch-Römer, C., 1993. The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363.
- Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (Eds.), 2006. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, Cambridge.
- Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), 2018. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, Cambridge.
- Ericsson, K.A., 1998. The Scientific Study of Expert Levels of Performance: general implications for optimal learning and creativity. *High. Abil. Stud.* 9, 75–100.
- Ericsson, K.A. (Ed.), 2009. *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*. Cambridge University Press, Cambridge.
- Ericsson, K.A., 2018a. The differential influence of experience, practice and deliberate practise on the development of superior individual performance of experts. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Ericsson, K.A., 2018b. An introduction to the second edition of the Cambridge Handbook of Expertise and Expert Performance: its development, organization, and content. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Fadde, P.J., Klein, G.A., 2010. Deliberate performance: accelerating expertise in natural settings. *Perform. Improv.* 49, 5–14.
- Feltovich, P.J., Prietula, M.J., Ericsson, K.A., 2018. Studies of expertise from psychological perspectives: historical foundations and recurrent themes. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Fessler, R., Christensen, J., 1992. *The Teacher Career Cycle: Understanding and Guiding the Professional Development of Teachers*. Allyn & Bacon.
- Frenzel, A.C., Taxer, J.L., Schwab, C., Kuhbandner, C., 2019. Independent and joint effects of teacher enthusiasm and motivation on student motivation and experiences: a field experiment. *Motiv. Emot.* 43, 255–265.
- Fukink, R.G., Tavecchio, L.W.C., 2010. Effects of video interaction guidance on early childhood teachers. *Teach. Teach. Educ.* 26, 1652–1659.
- Fuller, F.F., Bown, O.H., 1975. *Becoming a teacher*. In: Ryan, K. (Ed.), *Teacher Education*. University of Chicago Press, Chicago.
- Fuller, F.F., 1969. Concerns of teachers: a developmental conceptualization. *Am. Educ. Res. J.* 6, 207–226.
- Gladwell, M., 2008. *Outliers: The Secret of Success*. Penguin, London.
- Gobet, F., Charness, N., 2018. Expertise in chess. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Gurtman, M.B., 2009. Exploring personality with the interpersonal circumplex. *Soc. Psychol. Compass* 3, 1–19.
- Hatano, G., Inagaki, K., 1986. Two courses of expertise. In: Stevenson, H., Azuma, H. (Eds.), *Child Development in Japan*. W.H. Freeman, New York.
- Hatton, N., Smith, D., 1995. Reflection in teacher education: towards definition and implementation. *Teach. Teach. Educ.* 11, 33–49.
- Horowitz, L.M., Strack, S. (Eds.), 2011. *Handbook of Interpersonal Theory: Theory, Research, Assessment, and Therapeutic Interventions*. John Wiley & Sons, Hoboken, NJ.
- Korthagen, F.A.J., Kessels, J., Koster, B., Lagerwerf, B., Wubbels, T., 2001. *Linking Practice and Theory: The Pedagogy of Realistic Teacher Education*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Kua, J., Lim, W.-S., Teo, W., Edwards, R.A., 2020. A scoping review of adaptive expertise in education. *Med. Teach.* 1–30.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Leary, T., 1957. *An Interpersonal Diagnosis of Personality*. The Ronald Press Company, New York.
- Macnamara, B.N., Hambrick, D.Z., Oswald, F.L., 2014. Deliberate practice and performance in music, games, sports, education, and professions: a meta-analysis. *Psychol. Sci.* 25, 1608–1618.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81.
- National Research Council, 2000. *How People Learn: Brain, Mind, Experience, and School*. National Academy Press, Washington D.C.
- Oosterbaan, A.E., Van Der Schaaf, M.F., Baartman, L.K.J., Stokking, K.M., 2010. Reflection during portfolio-based conversations. *Int. J. Educ. Res.* 49, 151–160.
- Poldner, E., Van Der Schaaf, M., Simons, P.R.J., Van Tartwijk, J., Wijngaards, G., 2014. Assessing student teachers' reflective writing through quantitative content analysis. *Eur. J. Teach. Educ.* 37, 348–373.
- Sabers, D.S., Cushing, K.S., Berliner, D.C., 1991. Differences among teachers in a task characterized by simultaneity, multidimensionality and immediacy. *Am. Educ. Res. J.* 28, 63–88.
- Sadler, P., Ethier, N., Woody, E., 2011. Interpersonal complementarity. In: Horowitz, L.M., Strack, S. (Eds.), *Handbook of Interpersonal Psychology: Theory, Research, Assessment, and Therapeutic Interventions*. John Wiley & Sons, Hoboken, NJ.
- Schmidt, H.G., Boshuizen, H.P.A., 1993. On the origin of intermediate effects in clinical case recall. *Mem. Cognit.* 21, 338–351.
- Schmidt, H.G., Rikers, R.M.P.J., 2007. How expertise develops in medicine: knowledge encapsulation and illness script information. *Med. Educ.* 41, 1133–1139.
- Schön, D., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- Schön, D., 1987. *Educating the Reflective Practitioner*. Jossey Bass, San Francisco.
- Stigler, J.W., Miller, K.F., 2018. Expertise and expert performance in teaching. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge.
- Theelen, H., Van Den Beemt, A., Den Brok, P., 2020. Developing preservice teachers' interpersonal knowledge with 360-degree videos in teacher education. *Teach. Teach. Educ.* 89, 102992.
- Van Den Bogert, N., Van Bruggen, J., Kostons, D., Jochems, W., 2014. First steps into understanding teachers' visual perception of classroom events. *Teach. Teach. Educ.* 37, 208–216.
- Van Der Lans, R.M., Van De Grift, W.J.C.M., Van Veen, K., 2017. Individual differences in teacher development: an exploration of the applicability of a stage model to assess individual teachers. *Learn. Individ. Differ.* 58, 46–55.
- Van Manen, M., 1977. Linking ways of knowing with ways of being practical. *Curric. Inq.* 6.
- Van Tartwijk, J., Den Brok, P., Veldman, I., Wubbels, T., 2009. Teachers' practical knowledge about classroom management in multicultural classrooms. *Teach. Teach. Educ.* 25, 453–460.
- Vink, S., Van Tartwijk, J., Verloop, N., Gosselink, M., Driessen, E., Bolk, J., 2016. The articulation of integration of clinical and basic sciences in concept maps: differences between experienced and resident groups. *Adv. Health Sci. Educ.* 21, 643–657.

- Ward, P., Gore, J., Hutton, R., Conway, G.E., Hoffman, R.R., 2018. Adaptive skill as the conditio sine qua non of expertise. *J. Appl. Res. Mem. Cognit.* 7, 35–50.
- Ward, P., Schraagen, J.M., Gore, J., Roth, E.M. (Eds.), 2020. *The Oxford Handbook of Expertise*. Oxford University Press, Oxford.
- Wolff, C.E., Van Den Bogert, N., Jarodzka, H., Boshuizen, H.P.A., 2015. Keeping an eye on learning: differences between expert and novice teachers' representations of classroom management events. *J. Teach. Educ.* 66, 68–85.
- Wolff, C.E., Jarodzka, H., Van Den Bogert, N., Boshuizen, H.P.A., 2016. Teacher vision: expert and novice teachers' perception of problematic classroom management scenes. *Instr. Sci.* 44, 243–265.
- Wubbels, T., Brekelmans, M., Den Brok, P., Van Tartwijk, J., 2006. An interpersonal perspective on classroom management in secondary classrooms in The Netherlands. In: Evertson, C., Weinstein, C. (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Yinger, R.J., 1986. Examining thought in action: a theoretical and methodological critique of research on interactive teaching. *Teach. Teach. Educ.* 2, 263–282.

Teachers willingness and ability to teach to their best and well: the person in the professional

Christopher Day, School of Education, University of Nottingham, Nottingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	180
Framing the debate	180
Teachers who teach to their best and well	181
Teachers' lives: different journeys	183
Commitment	184
Professional identities: sites of struggle	185
Resilience	186
The school leadership influence	186
Conclusions	188
References	188

Introduction

Every parent wants their child to be taught by a “good” teacher. Yet understanding what this means is not straight forward. From the perspective of some parents, teachers who help students achieve academic outcomes which enable them to enter the workplace or undertake advanced study may be defined as “good”. For others, values-based education, development of independent thinking, or an appreciation of culture and society may be equally important in defining “good” teachers. In short, they may want a more rounded education for their children. Governments, on the other hand, may want all students to have the same opportunities for education and achievement, regardless of birth, gender and religion, “stock” the needs of a changing workforce, and also be prepared to participate in and contribute to the kind of society which they define as “good”.

From these perspectives, “good” teachers are those who are committed to their work, knowledgeable about the subject(s) which they teach, pedagogically adept, have a strong sense of positive professional identity, are clear about their educational (moral, civic, academic) purposes, those of the organization, parents, and their governments, (though some or all of these may be conflicting), willing and able to manage these, regardless of personal and professional circumstance, (in the face of changes in demands from governments and expectations of parents), and in classrooms in which students are likely to exhibit the results of a variety of motivations, abilities and behaviors.

Over the years, there has been a growing acknowledgment of the complex factors and contexts which teachers must navigate over a career if they are to learn, develop, and sustain “good” teaching. Researchers have focused variously upon the influence of biography, identity, policy, culture, organization, leadership, and moral purpose on teachers’ beliefs, purposes and practices. Such research has usually been carried out as discrete strands of inquiry, and so findings have been valuable, but limited. The same would apply to the impact of much educational small-scale, “cottage industry” research of these kinds on policy and practice. What is still largely lacking is a more integrated, holistic research-informed view which explores how these influences affect teachers’ work and lives. This article will seek to bring these strands together in order to discuss their effects, and how the extent to which they are managed by teachers and their school leaders influences their willingness and ability to teach to their best and well.

Framing the debate

Any discussion of teachers’ willingness and ability to teach to their best and well (good teachers and teaching) needs to be framed within changes in the larger national and international social and policy contexts, many of which influence government policy, and thus schools, directly or indirectly. For example, in this century, many school systems in the West and beyond have been subject to increasing devolution of decision making by governments with regard to governance systems, curriculum, and classroom pedagogies, alongside the development of increasingly sophisticated measures of external surveillance and accountability.

The leaders of most education systems believe now that they are able to judge accurately the successes attained by students, particularly in core areas of literacy, numeracy and science, through explicit criteria of school and teacher standards and student performance in schools both within and between countries, as a result of the development of international tests by the OECD (PISA, TIMMS and PIRLS). What some researchers term “regimes of numbers” have, it has been asserted, become “a resource through which surveillance can be exercised” and “data [have] become the resource for comparison” (Ozga, 2008: 264, 267), Governments’ responses to results in schools in their own countries, and comparisons of these with those in other countries are passed to schools in the form of new demands, and these are then passed on to teachers for action.

In effect, now and in the foreseeable future the work of teachers and the value they bring to students' learning and performance in examinations has become more transparent, "visible," and so more subject to both external and internal interrogation and judgment (Hall and Noyes, 2009). Governments and the public at large have come to assume that there is a direct cause and effect relationship between the quality of teaching, learning and student outcomes, rather than adopt research-informed findings that students' motivation, engagement, learning progress and achievements are likely to be mediated by both the personal and professional qualities of teachers, external social and peer influences.

Although curriculum and system changes have varied in countries and jurisdictions with different political, cultural and educational histories, practices and aspirations, all may be said to have complicated and enlarged teachers' work. Such complications have been further compounded by changes in attitudes and expectations over time in societies, and the persistence of technological advances which enable more information to be accessed, often more quickly by students than teachers, and thus are challenging traditional teaching and learning pedagogies. Taken together, these changes have added to teachers' intellectual, emotional and practical workloads. Because good teaching demands the investment of teachers' personal as well as "professional" selves, they have resulted in increased pressures upon teachers' in managing workloads, and the levels of energy they are able to bring to these. There is now widespread evidence internationally of increases in levels of stress and an erosion of morale in many teachers. In England, a recent survey found that 98% of teachers who responded reported that they were under increasing pressure, 82% reported that their workload had become unmanageable, and 75% reported that their mental health was adversely affected. Only 12% reported that they had a good work-life balance, and 43% were planning to leave within the next five years (Lightfoot, 2016). In England, also, among a growing number of countries, there is now a crisis of teacher recruitment and retention. Seventy-nine percent of schools reported that they were struggling to recruit to retain teachers (Lightfoot, 2016). A similar situation exists in the USA, one of many other countries which report low teacher morale and problems of recruitment and retention. It has been argued that many teachers who work in such 21st century environments are struggling to seek, "a sense of coherence, worth and belonging" (Lumby and English, 2009: 95). Yet, there is widespread evidence that not all teachers have, "set aside personal beliefs and commitments to live an existence of calculation" (Ball, 2003: 215).

I have written elsewhere of five important consequences for teachers' work and mental health, as a result of both the increased dominance of policy voices and changes in the needs of students:

- (i) Increases in the demands on teachers in terms of workload and bureaucracy associated with classroom practices which focus on close monitoring of "visible", measurable student progress on the mistaken assumption that all learning is "linear";
- (ii) The management of teachers' professional learning and development opportunities, so that they are closely aligned to the "functionalist" needs of the school, often allied closely with student performance expectations
- (iii) Participation in collaborative activities has become less of a choice that individual teachers make and more of a formal requirement to be met;
- (iv) The traditional isolation (and as a consequence the "individual" autonomy) of teachers has diminished;
- (v) There are increased pressures on teachers to enlarge the range of their classroom teaching and learning approaches; Such consequences can be found among reports about teacher morale in many countries

Yet while the flow out of teaching, especially among early career teachers in England, USA and Australia, has become alarming, most teachers in most countries stay. Those teachers experience the same shifts in policy, the same increases in workloads, the same dilemmas experienced by all teachers that have been claimed to have led to "de-professionalization", "technicisation", "demoralisation", "stress" and "attrition".

Teachers who teach to their best and well

An important question for those concerned with the quality of teaching and learning, and who wish to understand factors which influence this in changing times and changing circumstances is, therefore, "What influences their willingness and abilities to teach to their best and well?" It is important to use these two terms, because teaching to one's best does not necessarily mean that teachers will be "performing" well, however "performing" is defined. Teachers may be highly motivated, but their ability to help students meet the required output standards may be stronger or weaker depending upon their personal (mental and physical) health, reserves of energy (capacities for resilience), professional competences (knowledge and skills, relational connectedness with students and colleagues), and sense of positive professional identity (efficacy, agency). The extent to which these increase, sustain, or impede performance in its broadest sense may be mediated by external factors (policy, society) or/and internal factors (school leadership, available resources, and professional support and learning cultures). Research internationally into the values, characteristics and behaviors of teachers who teach to their best and well has found that they are those who first and foremost have an enduring internally driven belief that they can make a positive difference to the learning lives and achievements of students, through adding value to their self-belief, aspirations, learning and achievements; and that they do this through who they are (their beliefs, attitudes, personal and professional values, efficacy, sense of agency), what they know (knowledge, strategies, skills) and how they teach—their attitudes, relationships and behaviors in practice settings.

But what does "teaching to one's best and well" really mean? First, teachers can reasonably be expected to focus on increasing students' academic learning and achievement. In these terms, it is difficult to argue with Hattie's general premise that, "All students deserve at least a year's progress for a year's input, no matter where they start" (Hattie, 2015, p. 3). "Value added" techniques for

measuring the rate of students' progress have been developed in recent years, and are used in an increasing number of countries. These enable judgments of teachers' ability to "add value" to the academic level at which their students entered their classrooms to be made. Using these techniques, students' progress and performance may be measured against their own previous performance and home histories. This is known as "context value added" (CVA). Although CVA measures have been criticized for, "the extent to which they successfully control for time-invariant factors" (Gorard et al., 2013, in Morris et al., 2018: 726), they are widely used within schools in England, for example, as a reliable means of measuring student progress that is, "independent of the selection of pupils, background characteristics and innate ability" (Morris et al., 2018: 726; Chetty et al., 2014). This is regarded as a fairer, more equitable way of assessing their progress, and their teachers' contribution to this, than raw attainment measures. Individual results can also be aggregated at school level and compared with results from schools that serve similar socio-economic communities. In this sense, schools and their teachers can be identified as performing well or under-performing.

Yet measuring significant progress, even in terms of a "value-added" comparison between students' academic standing at the beginning and the end of each year, provides an insufficiently nuanced result. It is likely that there will be variations in the pace of student learning, their motivations, the varied effects of being taught by different teachers, and unavoidable unpredictabilities of health and home life. Reports by the OECD on student relative progress and performance suggest, also, that the variance in standards of achievement in reading is greater within than between schools, e.g., 36% versus 64% across OECD countries (OECD, 2009a,b). This may suggest a variance in the quality of teaching and/or other learning resources.

Policy voices across the world emphasize the need for students improved academic achievements. However, while these are fundamentally important, research suggests that we should avoid the temptation of making judgments about quality that rely solely on statistical measurements which apply only to one part of students' progress and achievement and not others. Fenstermacher and Richardson point to the dangers of assuming a direct causal relationship between teaching and learning:

There is currently a considerable policy focus on quality teaching, much of it rooted in the presumption that the improvement in teaching is a key element in student learning. We believe that this policy focus rests on a naïve conception of the relationship between teaching and learning. This concept treats the relationship as a straightforward causal connection such that if it could be perfected, it could then be sustained under almost any conditions, including poverty, vast linguistic, racial, or cultural differences, and massive differences in the opportunity factors of time, facilities and resources. Our analysis suggests that this presumption of simple causality is more than naïve; it is wrong ... good teaching is but one of four ingredient in the mix. The others are that the learner desires to learn and expends the necessary effort to do so; that the social surround of family, community, and peer culture support all assist in learning; and that there are sufficient facilities, time and resources (opportunities) to accomplish the learning that is sought

Fenstermacher and Richardson (2005: 3).

In addition to the issue of "causality", we should also avoid the temptation of talking or writing about "teachers" as if they were all the same. They are not. Not all are likely to hold the same educational values and beliefs, or be in possession of the same qualities, knowledge and skills necessary to teach to their best and well in changing conditions in the workplace, and all teachers will be challenged to maintain their commitment to their work over the whole course of their careers. Teachers who teach to their best and well can be differentiated from other teachers in their schools, not because they are extraordinary, or as some would claim, "great" but because of their abiding strength of purpose, their passion and courage (Day, 2008; Palmer, 2007) commitment of care (Noddings, 1992), career-long, "growth" mindsets (Dweck, 2008; Faulkner and Latham, 2016), capacities for resilience (Day and Gu, 2014), and their insistence in exercising individual and collective agency (Hargreaves and Fullan, 2012).

Research suggests that teachers who teach to their best and well are attentive to students' academic progress and attainment, but they are not those for whom, *measurement has become an end in itself rather than a means to achieve a good education in the fullest and broadest sense of the term*' (Biesta, 2015: 83). These teachers are driven by strong "moral" and "ethical" purposes (Day, 2017), regardless of personal, and professional circumstance. These purposes are at the center of their understanding and application of their responsibilities to act in the broader rather than narrower interests of the students. They inform their motivations, purposes, actions, and professional identities (Day et al., 2007; Nias, 1981); and they "enact" rather than "implement" externally initiated reforms, interpreting and contextualizing, rather than thoughtlessly or coercively complying (Ball et al., 2012). "Good" teaching has been described also as being more than assisting students to achieve high test scores. Rather, it consists of a combination of three elements: *logical acts* (explaining, interpreting, correcting, demonstrating); *psychological acts* (motivating, encouraging, rewarding, punishing); and *moral acts* (honesty, courage, tolerance, compassion, respect, and fairness) (Palmer, 2007; Fenstermacher and Richardson, 2005).

The synchrony between teachers and students indicates how much rapport they feel; studies in classrooms show that the closer the movement co-ordination between teacher and student, the more they feel friendly, happy, enthused, interested, and easygoing while interacting ... [the level of] ... Synchrony reflects the depths of engagement between the parties ...

Goleman (1995: 116–117).

A key factor that underpins the work of teachers who teach consistently to their best and well, is their inner strength of educational purpose, their "call to teach" (Hansen, 1995). Self-efficacy and agency are characteristics of these teachers. They are expressed in their strong sense of care about every student, an emotional commitment that goes beyond any contractual obligation to care for.

To these teachers, relationships are as important as rationality. They exercise professional empathy (apprehending another's reality), where their understandings result in context-sensitive action in the classroom.

Research consistently reports that teachers who develop and sustain a willingness and ability to teach to their best and well, demonstrate a strong sense of efficacy and agency as, within clear moral/ethical sets of values and educational beliefs, they navigate changes in their personal, professional, organizational and policy environments, while moving through different life phases and working contexts. What differentiates teachers who are able to teach to their best and well from others is that:

When faced with obstacles, setbacks, and failures, those who doubt their capabilities slacken their efforts, give up, or settle for mediocre solutions. Those who have a strong belief in their capabilities redouble their effort to master the challenges

Bandura (2000: 120).

Teachers' individual and collective efficacy (merging diverse individual goals in support of common organizational goals) is closely associated with their individual and collective sense of agency (Etelapelto et al., 2013), defined as, "being exercised in particular when professional subjects and/or communities exert influence, make choices, and take stances on their work and/or professional identities" (Etelapelto et al., 2013: 57–58). Many empirical studies have found that high levels of teacher efficacy and agency are associated with a strong sense of professional identity. However, any discussions of the nature of teacher identity must include a consideration of the extent to which individual teachers—and whole schools—may exercise efficacy and agency within changing economic, cultural, and social structures (Hitlin and Elder, 2007).

Changes in external environments and conditions of work which appear to promote narrow, "functionalist", "managerialist dominated" and "technicist" cultures of teaching and learning do not always result in feelings of helplessness among all teachers, or *determine* how all individuals feel or act:

It is not that oppressive structures and discourses do not exist, and that these do not affect teacher professionalism and identity. Rather, it is not necessarily the case that their existence, by definition, must always result in negative effects on all teachers. There is a wide range of international research on improving, effective, and successful schools which demonstrates how teachers mediate, resist, and find productive room to manoeuvre, asserting their individual and collective sense of agency within a sense of positive identity

Day (2018: 67).

To date, there is relatively little large scale empirical research which investigates whether, to what extent, and how teachers build and sustain their willingness and emotional and intellectual capacities to teach to their best and well, and, in doing so, challenge the claims of some academics that in recent times, their abilities to make professional judgments (decisional capital) and exercise self-efficacy and agency within a strong, positive sense of professional identity have been "subverted" and that, the "technical" aspects of teaching are now privileged ... to the neglect of the relational (Sachs and Mockler, 2012: 33).

Teachers' lives: different journeys

Teachers vary in the students they serve, the extracurricular duties they take on, the amount of time they have for planning ... Certainly some teachers are better able than others to accommodate the variety of stresses and strains they may face in their work. And to the extent that this is so, these differences may be due to differences in personal qualities ... We measure and track their value-added test scores but we do not measure their teaching loads, planning time, student absences, proportion of difficult to teach or resistant students, frequency of outside interruptions, access to textbooks or equipment of good quality.

Kennedy (2010: 596–597).

Over the last 20 year period especially, researchers have identified differences between teachers' work experiences, career trajectories, and teacher quality, and connected these to the positive and/or negative effects of national policy changes, conditions of service, school cultures, personal biographies and life changes. The most influential early research which focused on enhancing understandings of Swiss secondary school teachers' lives (Huberman, 1993) identified five different career phases from early survival and discovery (year 1–3), through stabilization (year 4–6), experimentation (year 7–8), serenity and conservatism (year 19–31) and disengagement (year 32+). Huberman found differences, however, between and within these groupings. Later he built upon this research, noting a "dynamic ebb and flow" (p. 42) among teachers as they responded to changed personal and organizational contexts. Subsequent research identified associations between teaching experience and levels of commitment (Day et al., 2007; Ben-Peretz and McCullough, 2009; Worth et al., 2018), and teachers' "best loved selves" (Craig, 2013).

Building on Huberman's classic work, a large scale four -year nationally representative mixed methods study of 300 teachers in 100 primary and secondary schools in England of teachers' work and lives and effectiveness was undertaken. It identified six "professional experience phases": commitment, support and challenge (year 1–3), Identity and efficacy (year 4–7); growing tensions and transitions (year 8–15); challenges to motivation and commitment (year 16–23); challenges to sustaining motivation (years 24–30); and declining motivation (year 31+).

Eight of the most important findings from that research were:

- (i) Within each professional experience phase there were teachers whose commitment to their work was rising, stable and declining
- (ii) Years of teaching experience did not necessarily lead to increased expertise
- (iii) There were perceived and significant statistical relationships between teachers' levels of commitment, capacities for resilience, and effectiveness in terms of student measurable attainments over a three year consecutive period
- (iv) Commitment and capacities for resilience varied over time within and between their professional experience phases
- (v) Teachers' sense of professional identity was an important part of their sense of professionalism
- (vi) This was influenced positively or negatively by policy, organizational, and personal/professional contexts
- (vii) School contexts (i.e., Supportive/unsupportive leaders, and or colleagues, student classroom behavior), and personal life events were perceived to have the biggest influence on teacher commitment, capacity for resilience, sense of positive or negative professional identity, and effectiveness.
- (viii) Despite what might be considered as considerable pressures to abandon or at least minimize their commitment to broader purposes of education, most of these teachers continued to uphold them (Day et al., 2007).

In subsequent years, policy interest in the impact of teacher commitment, professional identity, resilience, and relationships between these with their willingness and ability to teach well and to their best (effectively) has grown, particularly in countries which face continuing difficulties in teacher recruitment and retention. The next part identifies what research continues to show are the most important features that they exhibit, which underpin and inform their work, and what are the positive and negative influences that challenge or support these: (1) Commitment; (2) Professional identity; (3) Resilience; (4) School Leadership

Commitment

Commitment has been defined as a predictor of teachers' performance, as well as having an important influence on students' cognitive, social, behavioral, and affective outcomes (Louis and Kruse, 1998; Day et al., 2007). It is a term often used by teachers to describe themselves and each other, and is a part of their sense of professionalism.

In this article, I take commitment to mean a strong and enduring confidence by teachers that they can make a positive difference to students' motivation to learn, engagement in learning, academic progress and personal and social well-being; and that this will, in the longer term, also contribute to the "good" of society (moral/ethical/spiritual purpose). Teacher commitment has been identified repeatedly over the years by scholars as one of the most critical factors in the positive difference that teachers' can bring to their work. Crosswell (2006) proposed six "indicators" of teacher commitment:

- passion for a certain aspect of teaching and the teachers see it as a motivating or a sustaining factor
- investment of extra time—time donated beyond the usual contact hours with the students
- focus on individual needs, both socio-emotional and academic, of the students.
- maintaining professional knowledge and quest for further professional learning
- taking responsibility for transmitting knowledge and attitudes, values and beliefs to prepare students to take responsibility for future endeavors.
- willingness to engage through interaction with colleagues in the school and its community members, and to stay working at a particular school as an organization.

More recently, Thein et al. (2014) identified four inter-related dimensions of teacher commitment. It is interesting to note their inclusion of commitment to school as one of these

1. *Commitment to teaching*: related to the psychological link between teachers and their teaching work. This can be checked by seeing the willingness of the teachers in exerting significant efforts in being engaged in teaching and in devoting extra time for the students for improved results.
2. *Commitment to students*: related to the involvement and responsibility assumed by teachers in student development and achievement by making learning meaningful.
3. *Commitment to school*: related to the teachers associating their own beliefs and values with the goals and values of school and remaining with the school for the achievement of these goals.
4. *Commitment to the profession*: related to the extent of motivation the teacher has to developing their knowledge of teaching and learning.

We can see from the above what researchers have found to be the values, characteristics and behaviors of committed teachers. However, what they do not include are the mediating influences of their environments, how these might change over time, and how teachers themselves may change. It is self-evident that the depth of teacher commitment will be associated with the strengths of their beliefs and purposes, their self-efficacy, agency and professional identity, and that building and sustaining these may be enhanced by their capacities for resilience, or diminished by school contexts (student behavior, collegial and administrative support, parental demands). Nevertheless, it seems that committed teachers are able to teach to their best and well despite, as well as because, of factors in their working environments. They possess what has been defined as "professional capital". Hargreaves and Fullan (2012) proposed that professional capital combines "human" capital (qualifications, experience, knowledge and skills),

“social” capital (frequent, productive interactions with colleagues inside and outside school), and “decisional” capital (the ability to make wise judgments about teaching and learning in the classroom). Investment in the development of professional capital, they suggest, is necessary if students are to be taught well.

When the vast majority of teachers come to exemplify the power of professional capital, they become smart and talented, committed and collegial, thoughtful and wise. Their moral purpose is expressed in their relentless, expert driven pursuit of serving their students and their communities, and in learning, always learning, how to do that better

Hargreaves and Fullan (2012: 5).

As evidence to support associations between human and social capital and student achievement, they drew upon empirical research which found that teachers with high social capital increased their students’ mathematical scores by 5.7% more than teachers with lower social capital, and that students of teachers with both high human capital and high social capital, achieved the highest gains. Indeed, they found that so-called “low ability” (low human capital) teachers were able to achieve as well as teachers of so-called “average ability”, when they had strong social capital in their school (Leana, 2011: 34). Although a direct cause and effect relationship between professional capital and thoughtfulness and wisdom is unlikely, it is reasonable to suggest that there will be close associations.

Professional identities: sites of struggle

Developing a professional identity involves finding a balance between the personal and professional side of becoming a teacher

Pillen et al. (2012: 8).

Professional identities have been defined by some as “sites of struggle” (MacLure, 1993: 314). Teachers’ identity is now recognized as not necessarily being fixed over a career, but negotiated in changing personal, social, organizational and policy contexts. One or more of these may threaten to disrupt the stability of existing identities (Beijaard et al., 2004). It is the extent to which teachers are willing and able to manage the uncertainties of short, or longer term disruptions to their existing working lives that will determine whether their professional identities become or remain positive and stable, or become or remain negative-stable, or negative-unstable. Their ability to retain a strong sense of identity may be mediated positively or negatively by colleagues, school leaders, their performance in the classroom, their students’ responses, or the strength of their vocation. If, for example, alongside their confidence in their content and pedagogical knowledge, they have strong professional capital, then their ability to navigate the winds of change in their schools and personal lives is likely to be enhanced (Holstein and Gubrium, 2000).

Understanding teachers’ professional identities is important because they are sources of meaning for them. As with commitment, research recognizes the relationship between the rational and the emotional. Sustaining a strong and stable professional identity “requires the management of both by the individual” (Zembylas, 2005: 24). Their sense of professional identity may motivate them if they are positive, or de-motivate them if they are negative. They may contribute to their high self-efficacy or to low self-esteem. Professional identity should not be confused with “role”, which is the means by which people’s functions within organizations are organized (Castells, 2004). The ways in which teachers form their professional identities are not only influenced by how they feel about themselves and how they feel about their students (James-Wilson, 2001), but also about how they feel about themselves as members of a wider communities of teachers, the success of their work in the classroom, their relationships with colleagues, and the value that is placed upon them by their school leaders, parents, governments, and the community at large.

Teachers’ identities are a core element of their professionalism. Hoyle’s early description of the key differences is worth recollecting:

By restricted professionalism, I mean a professionalism which is intuitive, classroom-focused, and based on experience rather than theory [The teacher] is unencumbered with theory, is not given to comparing his work with that of others, tends not to perceive his classroom activities in a broader context, and values his classroom autonomy The extended professional, on the other hand, is concerned with locating his classroom teaching in a broader educational context, comparing his work with that of other teachers, evaluating his own work systematically, and collaborating with other teachers

Hoyle (1980: 49).

There is now considerable body of research which extends and nuances what Hoyle called “extended professionalism” to teachers with a “strong professional identity.” Not all teachers always demonstrate the same degree of professionalism. Some, for example, may show by their attitudes and behaviors, “restricted” professionalism, while others may show “extended” professionalism; and this may vary depending on circumstance at different points in their professional experience, and for different reasons.

If they have the support of colleagues, leaders and good relationships with students, then this will strengthen them further and reinforce their commitment and capacities for resilience.

Resilience

Resilience is not a quality that is reserved for the heroic few. Rather, it can be shared by many ... teachers who remain ... committed to serving the learning and achievement of the children on an everyday basis and also over the course of their professional lives

Day and Gu (2014: 3).

Early studies of resilience originated in the disciplines of psychiatry and developmental psychology, and either focused upon examining the personalities or traits of children or adults who had suffered significant physical or mental traumas and recovered (Howard et al., 1999). Researchers sought to understand how and why some were able to adapt and thrive, while others were not. More recently, the focus changed to one which sought to understand the positive characteristics of resilient people, in order to inform what kinds of interventions would be most likely to assist recovery. At this point, resilience began to be understood not only as a (fixed) personality trait, but as a characteristic or quality which could be enhanced so that it could function as a reserve that could be drawn upon, "to improve the odds of successful coping and survival ... through experiences of positive emotions" (Fredrickson, 2004: 1367–9).

A capacity for resilience has been found to be associated with optimism, altruism, self efficacy and emotional wellbeing—all qualities of teachers who teach well and to their best. For teachers, resilience:

... is not primarily associated with the capacity to "bounce back" or recover from highly traumatic experiences and events but, rather, the capacity to maintain equilibrium and a sense of commitment and agency in the everyday worlds in which teachers teach

Day and Gu (2014: 26).

Thus, resilience has become recognized as a social construction, influenced by multi-dimensional factors that are likely to be unique to particular contexts (Unger, 2004). When examined in teachers' working contexts, researchers have found that:

- (i) Resilience is not a quality that is innate
- (ii) "Coping" with the challenges of classroom and school environments is necessary but insufficient if teachers are to teach to their best and well. They need to manage these successfully
- (iii) The extent to which they are willing and able to do so is influenced directly by their sense of professional identity (commitment, self-efficacy and agency), and indirectly by their colleagues' support, school leaders, and their personal support they receive outside the school
- (iv) Teachers' capacities for resilience are likely to fluctuate over the course of a career
- (v) Because schools are socially dynamic settings (in the classroom, in the staffroom, in the playground), and teaching to one's best and well demands relational and intellectual connectedness, teachers need to be able to draw upon reserves of resilience on an everyday basis

There is no doubt that teaching can be a stressful work. Classroom and staffroom relationships are not always predictable (Patterson and Kelleher, 2005), and may not always be productive, personal life circumstances may be problematic, school leaders may not always be understanding or supportive, policy driven changes, or challenges of career role or phase may challenge teachers' beliefs and practices (Day and Gu, 2010; Mansfield et al., 2012). The strength and depth of teachers' commitment may be challenged by one or more of these. Capacities for resilience will vary within and between individuals.

The school leadership influence

So far, this article has cited research which has revealed the complexities of teaching, the positive and negative influences that changing personal, situational and policy contexts exercise upon teachers' commitment, efficacy, professional identities, agency and capacities for resilience. It has posited these are essential values and qualities displayed by those teachers who, often against the odds, remain persistently willing and able to teach to their best and well. What they do, how they feel about what they do, how they do it, and how successful they are, however, will not be independent of significant others. Among the most significant will be school leaders. Their success is likely to be associated, directly and indirectly, with the quality of leadership that they experience.

There is voluminous and varied research on school leadership, and much of this focuses upon principal leadership, since researchers seem to agree that it is they who are primarily responsible for establishing the vision, values, direction, structures, relationships, and learning cultures of the schools that they lead. Culture is referred to as "the way we do things around here", the patterns and expectations that, "permeate everything: the way people act, how they dress, what they talk about and consider taboo, whether they seek out colleagues or isolate themselves, whether they work together" (Deal and Peterson, 2016: 7). Research suggests that the nature of the culture can influence teachers' physical and emotional retention through, for example, strengthening or weakening ties of loyalty and contributing positively or negatively to the strength of teachers' commitment; that culture is dynamic, and can change; and that school principals are primarily responsible for creating and working with it (Waldron and McLeskey, 2010).

Research on what is known as “effective” and “successful” principal leadership has grown exponentially over the last thirty years, and a large corpus of knowledge is now accessible. Much of this consists of competing leadership “models”. There are so many, that one researcher critiqued their proliferation:

Within the literature on school leadership there also remains an overwhelming endorsement of one-style-fits-all leadership, often accompanied with varying degrees of evangelical zeal. Promoted are a range of adjectival leaderships: “democratic” leadership, “strategic” leadership, “breakthrough” leadership, and so on. While we know that leadership does not automatically take on a new meaning simply when a new adjective precedes the term, there remains a predominant view that the “right” leadership style, if found, practiced and implemented, in a strong, unequivocal manner will make all the difference

Mulford (2008: 38).

Since then, the most influential adjectival “models” have been associated with “transformational” (Leithwood and Riehl, 2005) and/or “instructional” leadership (Robinson et al., 2009). Within the “transformational” leadership model, Leithwood and Riehl identified core practices of successful principals:

- (i) Setting directions: Leaders set a clear direction by developing and communicating shared goals that encourage a sense of common purpose and high performance expectations
- (ii) Developing people: offering teachers intellectual stimulation and individual support
- (iii) (Re) designing the organization: designing their schools as professional learning communities, by modifying school structures and processes to increase professional collaboration and dialog among teachers and to improve home school relationships
- (iv) Managing the teaching and learning program

Both transformational and instructional models have been found to characterize leadership which has achieved and sustained success (Marks and Printy, 2003; Day et al., 2016). Both highlight the importance of the principals’ values, beliefs and practices in shaping the preferred workplace culture and meeting successfully external requirements (Day et al., 2011; Gu et al., 2014). That body of research reports that principals influence students’ learning progress and achievement indirectly, through their direct influence on teachers (Leithwood et al., 2006). A “best evidence” synthesis of the available quantitative research internationally identified five key dimensions of “effective” leadership, calculated statistically by “effect size”. Of these, “encouraging and participating in professional development” had the highest effect size of all (0.84). Leithwood et al. (2010), extended the dimensions of successful leadership by conceptualizing its indirect influence as flowing along four paths toward student learning: the rational, emotional, organizational, and family paths. Leaders, they have suggested, increase student learning by improving the condition or status of the variables on all these paths. This added to research which identified the importance of the socio-cultural dimension of successful leadership (Mulford and Silins, 2011; Bennett et al., 2014). Researchers in the USA, and England have also found reliable associations between qualitative and quantitative elements of successful leadership and teaching. For example, a key feature of successful, sustained improvement in students’ achievements in reading and mathematics in elementary schools in the USA is “relational trust” between teachers, principals and teachers, and students and teachers (Bryk and Schneider, 2002; Tschannen-Moran, 2014); and, noted earlier in this article, statistically significant associations between teacher commitment and student performance in examinations in primary and secondary schools in England have been found (Day et al., 2007).

There is no space in this article to detail all the merits and demerits of leadership models, though like all leadership models, they are general, do not, and cannot represent the nuances of successfully leading what have become characterized as “complex adaptive systems” (CAS), defined as, “systems that adapt and evolve in the process of interacting with dynamic environments” (Merali and Allen, 2011, cited in van der Steen et al., 2013: 556). In these authors’ views, thinking of schools as complex adaptive systems re-defines linear cause and effect relations as “reciprocal and interactive.” Rather than seeking direct cause and effect relationships between different factors, what van der Steen et al. [p. 55] term as “linear conceptions of causality”, an alternative body of leadership research has sought to understand how a range of factors interact and influence each other in predictable and unpredictable ways. These leadership studies are based upon an acknowledgment of schools, and therefore, relationships within schools, as complex adaptive systems. They have found that:

the process of organizing involves developing relationships from a shared sense of purpose, exchanging and creating information, learning constantly, paying attention to the results of our efforts, co-adapting, co-evolving, developing wisdom as we learn, staying clear about our purpose, being alert to changes from all directions

Wheatley (2005: 27).

Researchers who are members of a twenty five country “International Successful School Principalship Project” (ISSPP) network have identified successful leaders as those who focus upon managing a range of interactive variables and having much broader purposes than raising the academic profiles of their schools. Successful schools have been found, for example, to strive to educate their pupils by promoting positive values (integrity, curiosity, compassion and fairness), love of lifelong learning, as well as fostering citizenship and personal, economic and social capabilities (Mulford and Silins, 2011; Ishimaru, 2013; Day and

Leithwood, 2007). These social outcomes are deemed by successful leaders to be as important as fostering students' academic outcomes (Robinson et al., 2009; Bryk et al., 2010; Sammons et al., 2011; Bruggencate et al., 2012). The findings echo those of Biesta (2015), who claims that the purposes of schools include but are much broader than the "functional" Seeking to judge school success in those terms is likely to result in a narrow, possibly oversimplified, and even distorted understanding of principals' relationally oriented work with teachers. As Finnegan and Daly have observed:

... our findings suggest that the structure and quality of social relationships within schools ... play a crucial role in schools' capacities for organizational learning and improvement

Finnegan and Daly (2012: 65).

Teachers who are willing and able to teach to their best do not do so in isolation of the prevailing school culture. It is important to note, also, that *personal efficacy* beliefs are not fixed or necessarily stable, but subject to social influence, as is the capacity for resilience, agency and commitment and professional identity. Successful leaders are those, research suggests, who attend to these, develop a synergistic interplay between vision and values, strategies, actions and interpersonal relationships (Day et al., 2016), who create and sustain a climate of trust, foster individual and collective ambitions to improve, and build teachers' intellectual and emotional capacities through ensuring opportunities for inter-connections between staff and between staff and students and staff and parents (Mitchell and Sackney, 2000)—in short, provide every opportunity to teachers to teach to their best and well, in the knowledge that it is the teachers who have a direct effect upon student learning, progress and achievement through the quality of their relationships and ability to stimulate new knowledge generation.

Conclusions

The research reported about teachers who teach to their best and well in this article is located in the psychological, sociological, organizational, national policy, and school leadership contexts in which they work. In doing so, it reveals the personal, organizational and policy challenges which they face and overcome through their constant commitment to their educational values and beliefs, their resilience, professional identities, self-efficacy and agency. Their ability to do so is, research reveals, enhanced by the quality of school leadership. Research reveals the importance of leadership support, in particular by the principal in creating and managing cultures of trust and providing informal and formal opportunities for renewal, in conceptualising good teaching as encompassing the person in the professional. Teachers who teach to their best and well in today's contexts possess more than the subject knowledge and technical competences associated with the notion of teaching as a "craft," and fulfilling only "functional" purposes. Debates about the relative importance of "craft" over "artistry", "competences" over "qualities", or "compliance" over "commitment," are unhelpful in defining characteristics of "good" teachers and good teaching, since they oversimplify the complex nature, purposes and practices of "good" teachers and "good" teaching. While narrowly framed competences are rightly emphasized as being vital for "good" professionals, by themselves they do not meet the requirements of good teachers and good teaching. To fully understand these requires an examination of teachers' motivations, commitment, professional identities and capacities for resilience over a career. Good teaching by good teachers will include meeting sets of external and professionally identified capabilities or standards related to their expected contribution to students' learning and academic progress and attainment, but it is also more than this. Schools are socially interactive places in which emotions are embedded, and so classrooms demand teachers' cognitive and emotionally skilled engagement (Schutz and Zembylas, 2009; O'Connor, 2008). Good teachers educate the "hearts" of students as well as their "heads".

To teach to their best and well, teachers will need to possess and apply context sensitive combinations of both task and student-centric moral/spiritual and ethically embedded values, personal and professional qualities and skills, deep knowledge of subject and student, enduring commitment, a strong sense of professional identity, and a capacity for resilience which will enable them to authentically care for the student as person, in the knowledge that there is a symbiotic relationship between "cognition" and "affect", the emotional as well as cognitive selves. In doing so, they need the support of their colleagues and their leaders. Yet, even now, policy makers and even school principals whose support is necessary in different phases of their professional lives, seem not yet to fully understand that educating teachers to teach to their best and well means focusing more closely upon enhancing their levels of commitment, positive sense of professional identity, efficacy and agency, and building their capacities for resilience—in short, educating the person in the professional.

References

- Ball, S., Maguire, M., Braun, M., 2012. *How Schools Do Policy: Policy Enactment in Secondary Schools*. Routledge, London.
- Ball, S., 2003. The teachers' soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Bandura, A., 2000. Cultivate self-efficacy for personal and organizational effectiveness. In: Locke, E.A. (Ed.), *Handbook of Principles of Organization Behavior*. Blackwell, Oxford, UK, pp. 120–136.
- Beijaard, D., Meijer, P.C., Verloop, N., 2004. Reconsidering research on teachers' professional identity. *Teach. Teach. Educ.* 20 (4), 107–128.

- Ben Peretz, M., McCullough, G., 2009. International perspectives on veteran teachers. *Teach. Teach. Theor. Pract.* 15 (4), 403–408.
- Bennett, J.V., Yilmak, R., Dugan, Y.M., Brunderman, L.A., 2014. Developing the potential for sustainable improvement in underperforming schools: capacity building in the socio-cultural dimension. *J. Educ. Change* 15, 377–409.
- Biesta, G., 2015. What is education for? On good education, teacher judgement, and educational professionalism. *Eur. J. Educ.* 50, 75–87. <https://doi.org/10.1111/ejed.12109>.
- Bruggencate, G., Luyten, H., Scheerens, J., Slegers, P., 2012. Modeling the influence of school leaders on student achievement: how can school leaders make a difference? *Educ. Adm. Q.* 48, 699–732.
- Bryk, A., Schneider, B., 2002. *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation.
- Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., Easton, J., 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press, Chicago, IL.
- Castells, M., 2004. *The Information Age: The Power of Identity*. Blackwell Publishing, Oxford.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014. Measuring the impacts of teachers I: evaluating bias in teacher value-added estimates. *Am. Econ. Rev.* 104, 2593–2632. <https://doi.org/10.1257/aer.104.9.2633>.
- Craig, C.J., 2013. Teacher education and the best-loved self. *Asia Pac. J. Educ.* 33, 261–272. <https://doi.org/10.1080/02188791.2013.788476>.
- Crosswell, L., 2006. *Understanding Teacher Commitment in Times of Change*. Doctor of Education, Queensland University Technology Thesis, Australia.
- Day, C., Gu, Q., 2010. *The New Lives of Teachers*. Routledge, London.
- Day, C., Gu, Q., 2014. *Resilient Teachers, Resilient Schools: Building and Sustaining Quality in Testing Times*. Routledge, New York, NY.
- Day, C., Leithwood, K. (Eds.), 2007. *Successful School Principals in Times of Change: An International Perspective*. Springer, Dordrecht, Netherlands.
- Day, C., Sammons, P., Stobart, G., Kington, A., Gu, Q., 2007. *Teachers Matter: Variations in Work, Lives and Effectiveness*. Open University Press, Maidenhead.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *Successful School Leadership: Linking with Learning and Achievement*. McGraw Hill Open University Press, Maidenhead, England.
- Day, C., Gu, Q., Sammons, P., 2016. The impact of leadership on student outcomes: how successful school leaders use transformational and instructional strategies to make a difference. *Educ. Adm. Q.* 52 (2), 221–258.
- Day, C., 2008. Committed for life? Variations in teachers' work, lives and effectiveness. *J. Educ. Change* 9 (3), 243–260.
- Day, C., 2017. *Teachers' Worlds and Work: Understanding Complexity, Building Quality*. Routledge, London, UK.
- Day, C., 2018. Professional identity matters: agency, emotions, and resilience. In: Shutz, P., Hong, J., Cross Francis, D. (Eds.), *Research on Teacher Identity*. Springer, Dordrecht, pp. 61–70.
- Deal, T.E., Peterson, K.D., 2016. *Shaping School Culture*, 3rd Edition. Jossey Bass, New York.
- Dweck, C., 2008. *Mindset: The Difference Between Good Teachers and Great Teachers*. Ballantine Books, New York.
- Etelapelto, A., Vähäsantanen, K., Hökkä, P., Paloniemi, S., 2013. What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65.
- Faulkner, J., Latham, G., 2016. Adventurous lives: teacher qualities for 21st century learning. *Aust. J. Teach. Educ.* 41 (4).
- Fenstermacher, G.D., Richardson, V., 2005. On making determinations of quality in teaching. *Teach. Coll. Rec.* 107 (1), 186–215.
- Finnegan, K.S., Daly, A.J., 2012. Mind the gap: organizational learning and improvement in an underperforming system. *Am. J. Educ.* 119, 41–70.
- Fredrickson, B.L., 2004. The broaden-and-build theory of positive emotions. *Roy. Soc.* 259, 1367–1377.
- Goleman, D., 1995. *Emotional Intelligence*. Bantam, New York.
- Gorard, S., Hordosy, R., Siddiqui, N., 2013. How unstable are “school effects” assessed by a value-added technique? *J. Stud. Int. Educ.* 6, 1–9. <https://doi.org/10.5539/ies.v6n1pl>.
- Gu, Q., Day, C., Walker, A., Armstrong, P., Bryant, D., Ko, J., et al., 2014. How Successful Secondary Schools in England and Hong Kong Reshape External Educational Reforms: Key Messages From the Research. Symposium Conducted at the Symposium at the BERA 2014 Annual Conference, Institute of Education, London.
- Hall, C., Noyes, A., 2009. New regimes of truth: the impact of performative school self evaluation systems on teachers' professional identities. *Teach. Teach. Educ.* 25, 850–856.
- Hansen, D.T., 1995. *The Call to Teach*. Teachers College Press, New York.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press, New York, NY.
- Hattie, J., 2015. *What Doesn't Work in Education: The Politics of Distraction*. Pearson, London.
- Hittlin, S., Elder, G.H., 2007. Time, self, and the curiously abstract concept of agency. *Socio. Theor.* 25 (2), 170–191.
- Holstein, J.A., Gubrium, J.F., 2000. *The Self We Live by: Narrative Identity in a Postmodern World*. Oxford University Press, New York.
- Howard, S., Dryden, J., Johnson, B., 1999. Childhood resilience: review and critique of literature. *Oxf. Rev. Educ.* 25 (3), 307–323.
- Hoyle, E., 1980. Professionalization and de-professionalization in education. In: Hoyle, E., Megarry, J. (Eds.), *World Yearbook of Education 1980: The Professional Development of Teachers*. Kogan Page, London.
- Huberman, M., 1993. *The Lives of Teachers*. Cassell, London.
- Ishimaru, A., 2013. From heroes to organizers: principals and education organizing in urban school reform. *Educ. Adm. Q.* 49 (1), 3–51.
- James-Wilson, S., 2001. The Influence of Ethnocultural Identity on Emotions and Teaching. Paper Presented at the Annual Meeting of the American Educational Research Association, New Orleans.
- Kennedy, M.K., 2010. The uncertain relationship between teacher assessment and teacher quality. In: Kennedy, M. (Ed.), *Teacher Assessment and the Quest for Teacher Quality*. A Handbook. Jossey Bass, San Francisco, CA, pp. 1–6.
- Leana, C.R., 2011. The missing link in school reform. *Stanford Soc. Innovat. Rev.* 9 (4), 30–35.
- Leithwood, K., Riehl, C., 2005. What we know about successful school leadership. In: Firestone, W., Riehl, C. (Eds.), *A New Agenda: Directions for Research on Educational Leadership*. Teachers College Press, New York, pp. 22–47.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. Seven Strong Claims About Successful School Leadership. National College for School Leadership, Nottingham.
- Leithwood, K., Patten, S., Jantzi, D., 2010. Testing a conception of how school leadership influences student learning. *Educ. Adm. Q.* 46 (5), 671–706.
- Lightfoot, L., 2016. Nearly Half of England's Teachers Plan to Leave in the Next Five Years. *The Guardian*.
- Louis, K.S., Kruse, S.D., 1998. Creating community in reform: images of organizational learning in inner-city schools. In: Leithwood, K., Louis, K.S. (Eds.), *Organizational Learning in Schools*. Swets & Zeitlinger, Lisse, pp. 17–45.
- Lumby, J., English, F.W., 2009. From simplicity to complexity in leadership identity and preparation: exploring the lineage and dark secrets. *Int. J. Leader. Educ.* 12 (2), 95–114.
- MacLure, M., 1993. Arguing for your self: identity as an organizing principle in teachers' jobs and lives. *Br. Educ. Res. J.* 19 (4), 311–322.
- Mansfield, C.F., Beltman, S., Price, A., McConney, A., 2012. “Don't sweat the small stuff”: understanding teacher resilience at the chalk face. *Teach. Teach. Educ.* 3, 357–367.
- Marks, H., Printy, S., 2003. Principal leadership and school performance: an integration of transformational and instructional leadership. *Educ. Adm. Q.* 39 (3), 370–397.
- Merali, Y., Allen, P., 2011. Complexity and systems thinking. In: Allen, P., Maguire, S., McKelvey, B. (Eds.), *Sage Handbook of Complexity and Management*, pp. 31–52.
- Mitchell, C., Sackney, L., 2000. *Profound Improvement: Building Capacity for a Learning Community*. Swets & Zeitlinger, Lisse, Netherlands.
- Morris, T.T., Davies, N.M., Dorling, D., Richmond, R.C., Davey Smith, G., 2018. Testing the validity of value-added measures of educational progress with genetic data. *Br. Educ. Res. J.* 44 (5), 725–747.
- Mulford, B., Silins, H., 2011. Revised models and conceptualization of successful school principalship for improved student outcomes. *Int. J. Educ. Manag.* 25 (1), 61–82.
- Mulford, B., 2008. *The Leadership Challenge: Improving Learning in Schools* (Australian Education Review No. 53). Camberwell. Australian Council for Educational Research, Victoria, Australia. Retrieved from: www.acer.edu.au/research_reports/AER.html.
- Nias, J., 1981. Commitment and motivation in primary school teachers. *Educ. Rev.* 33 (3), 181–190.
- Noddings, N., 1992. *The Challenge to Care in Schools: An Alternative Approach to Education*. Teachers College Press, New York, NY.
- O'Connor, K.E., 2008. You choose to care: teachers, emotions and professional identity. *Teach. Teach. Educ.* 24 (1), 117–126.

- OECD, 2009a. PISA 2009 Results: Learning Trends: Changes in Student Performance since 2000, vol. v.
- OECD, 2009b. Teaching and Learning International Survey (TALIS): Creating Effective Teaching and Learning Environments. OECD, Brussels.
- Ozga, J., 2008. In: Allen, J., Smyth, G. (Eds.), *Social Capital, Professionalism and Diversity: Studies in Inclusive Education*. Sense Publishers, Rotterdam.
- Palmer, P., 2007. *The Courage to Teach*, tenth ed. Jossey-Bass, San Francisco.
- Patterson, J., Kelleher, P., 2005. *Resilient School Leaders: Strategies for Turning Adversity Into Achievement*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Pillen, M., Beijard, D., den Brok, P., 2012. Tensions in learning to teach: accommodation and the development of teacher identity. *Eur. J. Teach. Educ.* 55 (1), 8–24.
- Robinson, V., Hohepa, M., Lloyd, C., 2009. *School Leadership and Student Outcomes: Identifying What Works and Why*. Best Evidence Syntheses Iteration (BES), Ministry of Education, New Zealand.
- Sachs, J., Mockler, N., 2012. Performance cultures of teaching: threat or opportunity? In: Day, C. (Ed.), *International Handbook of Teacher and School Development*. Routledge, London, pp. 33–43.
- Sammons, P., Gu, Q., Day, C., Ko, J., 2011. Exploring the impact of school leadership on student outcomes: results from a study of academically improved and effective schools in England. *Int. J. Educ. Manag.* 25 (1), 83–101.
- Schutz, P.A., Zembylas, M., 2009. *Advances in Teacher Emotion Research: The Impact on Teachers' Lives*. Springer, Dordrecht, The Netherlands.
- Thein, L.M., Razak, N.A., Ramayah, T., 2014. Validating teacher commitment scale using a Malaysian sample. *SAGE Open* 17 (2), 1–9.
- Tschannen-Moran, M., 2014. Trust Matters: Leadership for Successful Schools. John Wiley & Sons.
- Unger, M., 2004. A constructionist discourse on resilience multiple contexts, multiple realities among at-risk children and youth. *Youth Soc.* 35 (3), 341–365. <https://doi.org/10.1177/0044118X03257030>.
- van Der Steen, M., van Twist, M., Fenger, M., Le Cointre, S., 2013. Complex causality in improving underperforming schools: a complex adaptive systems approach. *Pol. Polit.* 41 (4), 551–567.
- Waldron, N.L., McLeskey, J., 2010. Establishing a collaborative school culture through comprehensive school reform. *J. Educ. Psychol. Consult.* 20 (1), 58–74. <https://doi.org/10.1080/10474410903535364>.
- Wheatley, M., 2005. *Finding Our Way: Leadership for an Uncertain Time*. Berrett-Koehl, San Francisco.
- Worth, J., Lynch, S., Hillary, J., Rennie, C., Andrade, J., 2018. *Teacher Workplace Dynamics in England*. NFER, Slough.
- Zembylas, M., 2005. Discursive practices, genealogies, and emotional rules: a poststructuralist view on emotion and identity in teaching. *Teach. Teach. Educ.* 21 (8), 935–948.

Thriving on challenges: understanding teacher resilience in hard-to-staff Chinese rural schools

Qiong Li^a, Xiaohong Sun^a, and Jingtian Zhou^b, ^a Center for Teacher Education Research, Beijing Normal University, Key Research Institute, Ministry of Education, Beijing, China; and ^b Department of Educational Administration and Policy, The Chinese University of Hong Kong, Hong Kong, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	191
Policy context for hard-to-staff rural schools in China: an alternative hiring plan of <i>Special Teaching Position</i> (特岗计划)	192
Conceptual framework	192
Conceptualizing teacher resilience	192
Factors influencing the development of teacher resilience	193
Methodology	194
Participants	194
Data collection	194
Data analysis	196
Results	196
Rural teacher resilience across professional life phases	196
Group one: low-to-high	196
Group two: U-curve	196
Group three: moderate level with occasional waves	197
Group four: sustained moderate level	197
Group five: sustained high level	197
Factors influencing rural teacher resilience	197
Individual factors affecting rural teacher resilience	197
Environmental factors affecting rural teachers' resilience development	199
Discussion	200
Teacher resilience: a multidimensional and developmental construct	200
Building teacher resilience in hard-to-staff rural schools	201
References	202

Introduction

Teaching has been increasingly considered a physically and emotionally challenging profession (Hargreaves, 1994; Kyriacou, 2000; Troman and Woods, 2001). Therefore, teacher stress, job burnout and attrition have been important topics in international research on teacher professional development for more than two decades (Imazeki, 2005; Kyriacou, 1987, 2001). Studies have explained why some teachers gradually lose their teaching enthusiasm and stagnate or leave. Although teachers are confronted with challenges, many maintain their passion and commitment and try their best to teach well (Gu and Li, 2013; OECD, 2005). An international teacher survey showed that 78% of lower secondary teachers would still choose teaching if they were making a new career decision (OECD, 2014). However, there has been very little research concerning why these teachers maintain consistent enthusiasm and strong resilience despite the challenges in schools. Teacher resilience is thus a relatively new area that explores why teachers show tenacity and adaptability when they face pressure, challenges or adversity (Beltman et al., 2011). The focus on resilience provides researchers with a positive psychological perspective to develop methods of helping teachers maintain their enthusiasm and passion in adverse environments and ways to support teachers' professional development (Beltman et al., 2011). While researchers have attempted to identify personal and contextual risks and protective factors that influence teacher resilience, studies of how teacher resilience develops across their professional life phases and the factors influencing the development of teacher resilience remain insufficiently explored.

The rationale for our focus on teacher resilience in the Chinese context requires further explanation. First, resilient qualities have been emphasized for thousands of years in Chinese traditional culture. As part of character development, resilience has always been used to describe people who are indomitable and unyielding in adversity. For instance, "I Ching" (易经, one of the oldest Chinese classics) notes, "As heaven's movement is ever vigorous, so must gentlemen ceaselessly strive along" (天行健, 君子以自强不息) (Li et al., 2019). Chinese teachers have been immersed in such resilience education from a young age, which leaves potential impact on the way they respond to challenges in their work. In a context of rapid economic development, policy and social changes, Chinese teachers are confronted with many education reforms and challenges such as heavy workload, moving from an examination-oriented environment to student-centered learning assessments, and managing high parental expectations (Gao,

2008; Liu and Onwuegbuzie, 2012). Considering such work conditions, aforementioned resilience culture may have an influence on Chinese teachers' attitudes towards these challenging external requirements. Given that Chinese teachers' resilience is an under-researched area, it is necessary to amplify related studies in Chinese context.

Second, we were motivated by the need to explore how teacher resilience is related to the retention of quality teachers in the Chinese education system. Until the end of the 1990s, Chinese teachers were permanent employees by local educational authorities (Gao, 2008). Although new contract employment practices have been adopted in schools over the last decade, teaching is still seen as a stable profession. While the rates of teacher attrition in China are much lower than other countries in general (Gu and Li, 2013; Liu and Onwuegbuzie, 2012), the loss of rural teachers poses a serious problem. However, the loss of rural teachers in China poses a serious problem. Due to the long-standing urban-rural divide, rural schools are often staffed with significantly fewer quality teachers than their urban counterparts (Han and Li, 2013). As Chinese teachers in rural areas face higher pressure and more challenges, how do they maintain enthusiasm and commitment to teaching in times of change? This crucial question makes research into the resilience of Chinese rural teachers vital, especially as the answers may provide productive insights for school leaders and policymakers to enhance teaching and learning quality.

Questions remain about how and why teachers maintained resilience in hard-to-staff schools. Drawing on analysis of interview results from 12 rural teachers from a national project called *Special Teaching Position Plan* in a county in remote southwest China, this paper examines teacher resilience in hard-to-staff Chinese rural schools. We aim to explore the factors that affect resilience across the rural teachers' professional life phases so as to consider implications of these insights for the professional development of teachers working in hard-to-staff schools. The specific research questions are: What are the characteristics of rural teachers' resilience development across their professional life phases? How do internal and external factors influence rural teacher resilience?

Policy context for hard-to-staff rural schools in China: an alternative hiring plan of *Special Teaching Position* (特岗计划)

An alternative hiring policy is a teacher supplement strategy used to solve the serious problems of teacher shortage and limited budgets (Chudgar et al., 2014). These alternative routes may have different names, but their central goal is to staff schools with sufficient and qualified teachers (Liao, 2019).

Aiming to attract teachers to economically underdeveloped rural areas, the policy on *The Implementation of the Special Teaching Position (STP) for School Teachers in Rural Compulsory Education* was issued by the Chinese government in 2006. The implementation principles are stipulated for a fixed-term and offered all participating teachers the opportunity to be tenured after three years. Central and local governments are involved in the recruitment of STP teachers, and central government provides the funds to these teachers for their first three years. Provincial governments take responsibility for matters like analyzing the number of STP teachers needed, setting up the recruitment criteria and organizing recruitment examinations and interviews (Liao, 2019). Local governments are responsible for the management and professional development of all STP teachers and pay for the tenured STP teachers' salaries from the fourth year (Ministry of Education of China, 2006).

Even though alternative teacher hiring policies are not uncommon worldwide, the STP is different from those used in other countries, and the Chinese government invested considerable support in the plan. This support can be reflected in the investment of special funds and the number of STP teachers. From 2006 to 2015, the central government arranged nearly 30 billion CNY for the STP plan and recruited more than 570,000 STP teachers. Since the implementation of the STP in 2006, nearly 300,000 STP teachers have been recruited to more than 30,000 rural schools in 22 provinces in Chinese central and western regions, and the retention rate had reached 87% for three consecutive years. STP teachers face many challenges, including their work-life balance. Liao (2019) described the emerging phenomenon of "weekday rural teachers, weekend urban spouses and parents" faced by STP teachers who stayed in rural schools. Inconveniences associated with daily living are a continual challenge as remote rural schools lack sufficient resources, are far from the city and do not satisfy teachers' basic needs (Wang et al., 2017).

Conceptual framework

Conceptualizing teacher resilience

The concept of resilience originates from psychiatry and developmental psychology which study personality characteristics of children who adapt well despite being in a disadvantaged environment (Howard et al., 1999). These studies have found that some children living in high-risk environments have good adaptability. Instead of being overwhelmed by difficulties or challenges, they adapted to overcome difficulties and develop positively (Masten, 2001). Early resilience research paid attention to both static internal protective factors (e.g., optimism, humor, and persistence) and external protective factors (e.g., seeking help and social support) influencing individuals' strengths and potential (Garmezy, 1984). At the same time, some researchers believed that resilience reflected knowledge, abilities, attitudes, traits and skills that could protect individuals from harm irrespective of circumstances (Garmezy, 1985). Scholars found that simply focusing on protective factors is insufficient to fully understand individual resilience, or devise effective strategies to support its development (Mansfield et al., 2016). Thus, researchers

started to pay attention to the interaction between protective and risk factors, aiming to reveal the relationship between internal and external factors that affect the development of resilience.

Four main definitions of teacher resilience can be summarized from the literature. The first is the process view, that is, teacher resilience is a dynamic process observed when teachers face a complex and rapidly changing external environment (Luthar et al., 2010). For example, some researchers pointed out that resilience developed and changed dynamically and was a process of interaction between individuals and the environment (Richardson, 2002). The second is the ability view, which sees teacher resilience as a kind of ability to manage an uncertain teaching situation and maintain a sense of commitment and responsibility (Gu and Day, 2011). From this perspective, resilience is regarded as to withstand high levels of destructive change while exhibiting few negative behaviors (Werner, 1995), recover from a negative experience and adapt flexibly to the changing environment (Lazarus, 1993). Some researchers described resilience as the positive trait that helps individuals overcome difficulties, face adversity, and continue to live a meaningful life with a healthy attitude (Turner, 2001). These researchers tended to regard resilience as an individual ability and cognitive or emotional psychological traits, including personality traits and self-concepts (Rutter, 2000). The third perspective is the result view, which means, teacher resilience is the result of adaptation, emphasizing teachers' positive condition after experiencing adversity (Tait, 2008; Masten and Coatsworth, 1998). In other word, resilience is not a fixed ability but the result of the interaction between the individual and the environment (Johnson and Wiechelt, 2004). The fourth is the multidimensional view. Teacher resilience is constructed by the combined action of individual and environmental factors could be understood as a comprehensive body of ability, process, and results (Luthar et al., 2010).

In summary, the process view emphasizes the dynamic and developmental nature of resilience; the ability view focuses on the importance of individual factors, while the result view stresses individual development. The multidimensional view is a synthesis of the other three views regarding teacher resilience as process, results and ability.

Factors influencing the development of teacher resilience

Researchers began to explore the relationship between protective and risk factors from the 1990s and proposed many models to describe how protective factors reduce the disadvantages caused by risks. In 1990, Rutter suggested that research should explore the development and situational mechanisms of the protection process (Kaplan, 1999). In subsequent studies, the interaction between the environment and the individual was emphasized. Researchers proposed that resilience was not a fixed or absolute ability that enabled individuals to resist stress or overcome difficulties whatever the situation is. In fact, resilience and its effects could change as individuals interact with the environment. (Rolf et al., 1990). This view is supported by Rutter (Rutter, 1985) who believed that the operation of the protective mechanism was the result of the interaction between the environment and the individual. Working together, both aspects could buffer the negative impact of risk factors on the individuals.

Studies have explored the influence of risk and protective factors on teacher resilience development. Risk factors refer to physical, psychological, cognitive or environmental factors that hinder teacher professional development, even leading to burnout and other undesirable results (Beltman et al., 2011). Heavy teaching workload, classroom management difficulties, lack of support and limited resources are all potential threats to resilience development. Beltman et al. (2011) summarized previous research, dividing the risk factors into personal and environmental risk factors. Personal risk factors include negative self-belief, difficulties encountered when seeking help, and the contradiction between personal beliefs and reality. Environmental risk factors include pre-service teacher education environment (e.g., unstable curriculum structure, inappropriate training curriculum system, or excessive academic pressure), family environment (e.g., negative pressure, lack of family support, or the imbalance between family and work), school or classroom environment (e.g., classroom management difficulties, destructive student behavior, or time-consuming non-teaching work) and professional work environment (e.g., professional development challenges, lack of practical teaching experience, work pressure, lack of support from school leaders, or lack of practical training) (Beltman et al., 2011). Risk factors can also be considered from a broad social perspective, with external factors such as changing external policies, teacher accountability, decentralization and education reforms influencing teacher resilience (Kelchtermans, 2005). Challenges or pressures affecting the development of teacher resilience teachers are mainly from three aspects: the personal level (e.g., lack of confidence, physical health challenges such as insomnia or anxiety), the school level (e.g., lack of sufficient resources and limited supportive leadership relationships), and the social level (e.g., threats from neighbors) (Maring and Koblinsky, 2013). The pressure teachers face will also continue to change with teacher development. Therefore, it is very important to clarify the source of pressure at each stage of teacher development to identify the risk factors of teacher resilience.

Protective factors can interact with risk factors to reduce or eliminate their negative influence. Protective factors may include personal altruistic attributes, strong internal motivations, and support from school leaders, peers or family. Some researchers pointed out that protective factors were more important for teacher resilience than risk factors (Knight, 2007). Therefore, many studies explored the resources that could enhance teacher resilience by asking what helped teachers maintain their teaching enthusiasm and quality. These studies can be roughly divided into two groups: those that focus on personal factors and those centering on environmental aspects. Personal factors include motivation (Watt and Richardson, 2008) and self-efficacy (Hoy and Spero, 2005), while environmental factors contain professional development (Anderson and Olsen, 2006), instructors (Shank, 2005), and trusted interpersonal relationships with students (Hirschhorn, 2009). Thus, protective factors of teacher resilience consist of internal protective factors and external protective factors. Internal protective factors refer to psychological characteristics and abilities to adapt quickly in adverse environments (Beltman et al., 2011). For teachers, the core internal protective factors are their altruistic motives and strong internal motives for teaching (Beltman et al., 2011). Teachers' self-efficacy is another

important internal protective factor. Day (2008) believes that teachers with strong resilience can be confident and competitive. There are six types of teachers' internal protective factors: (1) personal traits, e.g., gender, sense of humor, strong internal motivation, optimism, emotional intelligence, flexibility, and persistence (Beltman et al., 2011); (2) self-efficacy and self-confidence (Day, 2008); (3) coping skills, e.g., using coping strategies, active problem-solving skills, help-seeking, high-level interpersonal skills, and social support networks (Sammons et al., 2007); (4) teaching skills, e.g., understanding students, being creative, proficiency in teaching, and paying attention to student learning (Yost, 2006); (5) professional reflection and growth, e.g., paying attention to professional development, self-insight, self-evaluation, reflection, professional freedom, and lasting professional learning (Beltman et al., 2011); and, (6) self-care, e.g., taking responsibility for one's physical and mental health and actively creating favorable conditions for oneself (Beltman et al., 2011). The more protective factors he/she has, the better they can recover from adversity and difficulties. It is also important to note that the role of protective factors varies from person to person, depending on the teacher's career, the specific challenges they face and how they respond.

The external protective factors of resilience refer to the resources in the social environment that reduce the impact of individual risk factors. Protective environmental factors that can help teachers recover quickly occur at three levels: school, family, and society. First, strong, caring leadership is the main support resource for teacher resilience at the school level. Strong and caring school leaders are democratic, open, and fair, so that they create a pleasant and comfortable working environment for teachers. They also provide consistent support and timely, positive feedback for teachers, helping them overcome difficulties encountered in teaching and promote their professional development. (Gu and Li, 2013). Especially for new teachers, school leaders could support them by not assigning additional teaching tasks, making sure they are not given difficult-to-manage classes, and recognizing their contributions to the school. Ultimately, strong leaders can improve teachers' teaching effectiveness and professional commitment. The second factor is the proactive support of instructors from the same field, which is especially important for pre-service and new teachers. Mentoring relationships can help new teachers quickly adapt to the school environment and improve their self-reflection and problem-solving skills. Their self-esteem, self-confidence and teaching enthusiasm will be enhanced too. Support from peers and colleagues also plays an important role and can motivate teachers, help them overcome work difficulties, and maintain their enthusiasm, especially when facing difficulties. Finally, students are both a challenge and a source of support for teachers. A positive teacher-student relationship can help teachers overcome difficulties, and contribute to their well-being.

At the family level, the support from family and friends can reduce teachers' worries and enable them to balance work and life. They may then be better able to complete teaching tasks and maintain their teaching enthusiasm. The care of family members can give teachers a platform for emotional adjustment, so as to maintain positive emotion and confidence. At the societal level, it is necessary for society to create an atmosphere of respect and safety for teachers to improve their status. More relevant policies and regulations should be promulgated to promote teacher retention and development as well.

Methodology

Participants

This study adopted criteria sampling (Patton, 2002) and selected 12 STP teachers from rural schools as research participants (Table 1). As the research focused on outstanding or "retainable and teach-well" STP teachers, the participants were required to meet specific criteria. They were (a) teachers who are "retainable", that is, teachers who have been employed for more than 10 years in the same rural school, and (b) teachers who are teaching leaders now. We chose "retainable" teachers because they had shown strong resilience. Recruiting the teaching leaders was not only because they were outstanding STP teachers, but it would also allow us to analyze how teacher resilience changes across different career stages. Given the qualitative emphasis on contextual features, we chose to focus on just one Chinese province to attend to its local features in-depth and investigate how those features impacted the teachers' resilience development. Twelve STP teachers in County B of Province G met the sampling criteria. We approached the 12 teachers in early 2020, and they all agreed to participate in this study voluntarily.

Data collection

Both individual and focus group interviews were held to generate data about STP teachers' resilience development and the reasons for development. The authors visited G province, B county, from October to December 2020, where they conducted face-to-face individual and group interviews with participants. First, researchers provided information about the study, collected basic information about the participants and obtained verbal consent for their participation. Then in-depth semi-structured individual interviews took place with 12 STP rural teachers. Each interview lasted one to two hours to ensure a full understanding of the participants' willingness to stay or leave, as well as factors and key events that influenced their willingness, such as influencing factors and key events, including changes of emotions, attitudes, and agency since they chose to be an STP teacher. A focus group interview with five participants was also hosted to collectively share their experiences and feelings about the development of resilience.

A month later, telephonic or online interviews were conducted based on the preliminary findings. Here, all participants were invited to provide more details about their experiences and feelings. This second round of interviews offered an opportunity to refine the emerging themes and delve deeper into the factors and critical incidents affecting STP teachers. The main interview questions focused on the reasons for choosing a teaching job, as an STP teacher, the difficulties they faced in different career stages in hard-to-staff schools and how they overcome these difficulties.

Table 1 Background of the participants.

	<i>ZHANG</i>	<i>PANG</i>	<i>DAI</i>	<i>LI</i>	<i>LIU</i>	<i>TAN</i>	<i>HE</i>	<i>CHEN</i>	<i>JIN</i>	<i>AN</i>	<i>MA</i>	<i>ZHONG</i>
Entry time	2008	2008	2008	2008	2008	2009	2008	2007	2010	2010	2007	2008
Teaching age	12	12	12	12	12	11	12	13	10	10	13	12
Educational background	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate	Undergraduate
Subject	Chinese	Chinese	Math	Math	Math	Math	Math	Chinese	Math	Physical	Geography	Math

Data analysis

This study reports primarily on the qualitative thematic analysis of the transcriptions of interviews with 12 rural teachers. The interviews were recorded and transcribed verbatim, and then listened to a second time to add non-verbal information such as changes in the interviewee's tones, moods, pauses, and silences during the interviews. The details supplemented the analysis. The researchers sorted and numbered the original data. In the data collection and preliminary analysis stage, this study mainly used the situational analysis method to describe the rural teachers' resilience development. Initial coding of 12 teacher interview transcripts and a constant comparative approach were used for further interpretation and understanding through sorting, reconstruction and re-coding an agreed but still evolving index. In the stage of in-depth analysis and concluding, the qualitative analysis was performed using MAXQDA software, through which the participants' responses were coded and analyzed with the generic analysis method, involving searching for themes that capture important elements within the data and focusing on the relationships between themes (Nowell et al., 2017).

The validity of the study was ensured through meticulous examination of the data by two researchers. An independent examination of the data was carried out in the beginning of the analysis, and a brainstorming meeting was subsequently convened, focusing on the research questions in light of the interview analysis. Interviews were jointly examined at this stage, in an iterative process, until agreement was reached.

Results

Rural teacher resilience across professional life phases

Taking the progression of the tenacity of rural teachers from STP as the ordinate, and the teachers' working years as the time abscissa, Fig. 1 shows the resilience development of the 12 rural teachers. Five patterns of STP teachers' resilience development emerged: *low-to-high*, *U-curve*, *occasional waves*, *stable development*, and *high sustainable enthusiasm* (Fig. 1 Shown).

Group one: low-to-high

ZHANG and PANG were both reluctant to be STP teachers at first. They passively chose this career under the advice of family members. ZHANG's parents believed that *"teaching is a stable profession"*, and PANG's parents believed that *"being an STP teacher would help me be closer to home"*. As a result, these two teachers were less resilient at the start of their careers. As ZHANG said:

At that time, I didn't want to be an STP teacher at all. My heart was full of resistance. I wanted to leave every day, and I didn't want to stay there for a moment.

PANG also described how eager he was to leave at first. However, both these teachers gradually felt a sense of achievement about their work. ZHANG found that *"children are so cute"* and felt *"a sense of accomplishment"*. These participants noted that teachers' salaries and social status were higher than before, and local people believed that teaching is *"the profession that everyone admires"*. The support from school leaders and colleagues also improved the two teachers' ability to cope with difficulties. In other words, these teachers showed enhanced resilience. Therefore, the resilience development of these two teachers can be characterized as moving from *"low-to-high"*.

Group two: U-curve

Both LIU and CHEN experienced high enthusiasm, disillusionment/loss of ideals and renewed fighting spirit in the professional development process. Their resilience development showed typical characteristics of a *"U-curve"*. The two teachers *"love education"* and were *"also very fond of children"* before starting work. Thus, they were *"full of confidence when they first entered the job, and curious about everything"*. However, the harsh environment of the schools quickly extinguished their enthusiasm. LIU observed:

The environment was so bad that you can't imagine it. At the beginning, the dormitories we lived in were transformed from classrooms. The teaching building was dilapidated, and the floor of the classroom had potholes. There were even all kinds of poultry in the classroom! I was in a very low mood and really wanted to leave.

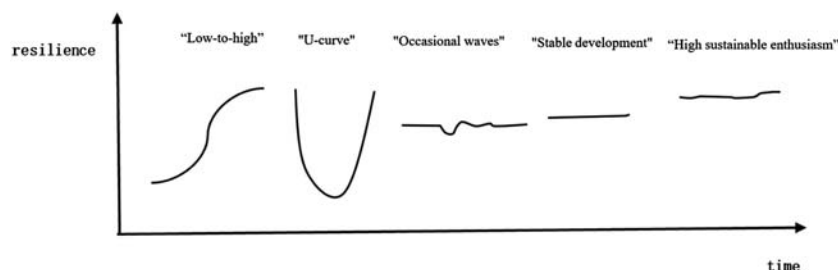


Fig. 1 Five groups of teachers' resilience development.

Similarly, CHEN experienced serious burnout not long after entering the job – attributed to difficulties in class management and lack of subject knowledge and teaching skills. Gradually, however, the teachers adapted to the work and lives in rural areas. They rekindled their fighting spirit and devoted themselves to teaching because of the harmonious working atmosphere, the gradually improved school environment and most importantly, their love for education and students.

Group three: moderate level with occasional waves

Some rural teachers showed relatively strong resilience throughout their careers with occasional fluctuation. We termed this the “occasional waves” of resilience development in this research. TAN and AN had been passionate about teaching for more than ten years. In recent years, TAN’s increased administrative load was taking up much of her time.

Most of my time is spent dealing with affairs unrelated to teaching, so I am a little tired physically and mentally.

In contrast, AN is “*not a local person*”, and she occasionally wants “*to go back to my hometown*”. Although the two teachers have tend to leave from time to time, they still feel infinitely motivated when they stand on the podium to teach.

Group four: sustained moderate level

The resilience of this group remained a moderate level at each stage – thus, characterized as “sustained moderate level”. JIN and ZHONG can be classified into this type. JIN has been an STP rural teacher for ten years, and has been doing his job diligently. He always overcomes difficulties in teaching or life, but noted:

In fact, I don’t like education much, but I still hope to do my job well. I think it is my responsibility – the matter handed over to me – and I should try my best to do it well.

ZHONG has always responded to problems in teaching – and life – calmly. He takes responsibility for doing his job, saying: “*Fight fire with water. It is a natural feeling, and I don’t think about it so much*”.

Group five: sustained high level

Four teachers v DAI, LI, HE, and MA, have shown sustained high level of resilience throughout their careers, coping actively with difficulties in teaching and life. During the interviews, we found these teachers had the following characteristics in common:

First, all started their jobs with a clear goal: “*desiring to contribute to their hometown*”; “*rural children’s desire for knowledge*”; “*hoping to go back to the village elementary school in my hometown*”; “*born in the countryside and knowing clearly what children in the countryside need*”. Second, they persevered to face difficulties: “*It is normal to meet some difficulties at the start of one’s profession*”; “*From the beginning of my choice in education, I persevered unswervingly*.”; “*When you meet difficulties, you must take the initiative and find ways to overcome them*”; “*If you need to deal with some difficult problems, please consult colleagues or respected teachers in the same discipline*”. Third, all are reflective practitioners: “*Teachers should always reflect on themselves and their teaching, otherwise they will never be able to improve their teaching level*”; “*Frequently conduct teaching research*”; “*Listen to the lessons of experienced teachers, learn class management experience from predecessors, and reflect on yourself from the teacher training provided by the school*”; They showed strong resilience from the outset and had worked hard consistently in their teaching positions for many years. Although these four teachers have experienced many difficulties in their careers, they have always maintained a positive attitude to learning and teaching. They showed strong resilience from the outset and had worked consistently hard in their teaching positions for many years.

Factors influencing rural teacher resilience

Although outstanding STP teachers have built resilience by the later stages of their careers, the levels of their resilience varied at different stages of their journeys. The teachers’ resilience development also showed different characteristics. Why does resilience differ? STP teacher resilience mainly comes from two sources: individual factors and environmental factors. These two factors interacted closely, leading to unique outcomes for each individual.

Individual factors affecting rural teacher resilience

There are four individual factors that affect STP teacher resilience: individual characteristics, internal emotions, professional abilities and life needs. Every aspect shows the interaction between protective and risk factors.

Individual characteristics: suitability for teaching versus non rural life trajectory

The rural teachers’ individual characteristics (e.g., personality, working style, self-efficacy, and background) directly impact upon their resilience. Protective factors are reflected in teachers being “*suitable for teaching*”. Some teachers “*have been confident enough*”,

or “*peaceful about everything*”, or “*have a cheerful personality*”. These positive traits generate enthusiasm to help teachers manage difficulties. Similarly, the teachers’ own style of coping also reflects their “*suitability for teaching*”. For example, one participant said:

I think I should try my best to do all the things I have been told, either not to do them or to do the best. This is my principle.

Similarly, another participant suggested:

I want to make my own mark, so I would not leave if I can’t do it.

Teachers with a high sense of self-efficacy also have a strong ability to rebound when encountering difficulties and show strong resilience. As MA said, “*I am very good at learning. I don’t think there is any difficulty*”. However, “teachable” teachers also face individual risks, such as regional issues, which means they were not born or raised in the countryside. Whether they can adapt and settle in the rural areas is an important factor. ZHANG grew up in the county town, and “*never having lived in such an environment before, there was a big sense of disappointment when I first came here*”. ZHONG had the same feeling. When he first arrived at the rural school, he could not even understand the local dialect and noted: “*Communication is always a problem*”.

For rural STP teachers, the “teachable” characteristic is a protective factor, while the regional problem is the risk factor. Teachable characteristics can stimulate teachers’ inner teaching enthusiasm and improve resilience. However, risk factors like regional problems may affect their decision to stay, reducing their resilience.

Inner emotion: occupational enthusiasm versus career burnout

Some teachers took the position initially with the idea of “*making a contribution to my hometown*” or “*returning to my hometown and devoting myself to local education*”. After starting the job, their inner emotions were reflected in their commitment. As participants said:

I like being with children. That makes me feel energetic.

I am full of expectations for the arrival of every new student.

I always feel a sense of fulfilment when teaching.

After several years, the teachers felt that they “*have strong emotions for the school*” and “*want to change the mindset of more children*”. Inevitably, some teachers experienced burnout during the process. Some teachers said that “*Students are difficult to manage and sometimes I feel very tired*”. Positive emotions give STP teachers a high degree of identification with their work and gradually lead to a sense of belonging and mission. This kind of emotion enables STP teachers to calmly face the pressures of work and life, but burnout consumes their enthusiasm.

Professional competence: reflection versus teaching problems

In terms of professional ability, risk factors are problems encountered by STP teachers in teaching, and protective factors refer to learning and reflection methods for solving these problems. Almost all STP teachers had teaching difficulties when they first entered their jobs. This occurred as they struggled with their new identity as a teacher, or because the subject they were teaching differed from their major. As ZHANG said:

My major was history, so I was not professional enough to teach Chinese.

Faced with such predicaments, excellent STP teachers will find ways to solve the problems. For example, one participant said:

By observing some famous teachers’ classes, analysing how the same lesson is taught, and then reflecting on what is wrong with my class, I improved my teaching skills.

STP teachers may also receive “supports from school leaders and colleagues”, “participate in lectures” or “participate in teaching competitions” to improve their teaching skills. Administrative load was frequently cited as a major hindrance. Some teachers complained that “There are many administrative affairs now. It seems that administration is my job and teaching is the supplement”. Teaching is the teachers’ main task, and whether the problems in the teaching process can be solved smoothly is a key factor for STP teachers to sustain their professional development motivation. Although teaching difficulties cannot be avoided, most outstanding STP teachers could seek solutions and achieve professional growth through active learning and reflection.

Life needs: pursuit of ideals versus shock of reality

Whether the life needs of STP teachers can be met is also an important factor influencing their resilience. Life needs include personal needs (e.g., shopping or travel) and relationships (e.g., taking care of elderly parents, living apart from their loved ones). Almost all the teachers mentioned life needs in their interviews. For example, ZHANG wanted to leave when he first entered the profession because “the accommodation conditions were pretty bad and it was not convenient to go out. Only the principal’s car could take teachers out”. PANG’s reason for wanting to leave was more private: “My parents, children and lover are all in another place”. HE and his spouse do not work at the same place, leading to strong desire to move elsewhere. These teachers faced dilemmas, but none of them left. Their pursuit of educational ideals supported them and encouraged them to deal with the difficulties. As ZHANG said:

“I really don’t want to leave my students. I hope to lead them to see the bigger world”.

Life needs are the basis for individual survival and development, and the tension between reality and their pursuit of ideals is an important factor in their resilience development.

Environmental factors affecting rural teachers’ resilience development

The environmental factors that affect STP teachers’ resilience development mainly include environmental atmosphere, external support and others. The interaction between protective and risk factors is reflected in each aspect.

The effect of significant others: modeling and leading versus negative suggestions

Support from significant others is indispensable in the process of rural teachers’ professional development. Significant others include family members, school leaders, and colleagues. The effect of family members is embodied in their attitudes toward teachers.

My family thinks that teaching is a stable job. In fact, the social status of the teaching profession in the local area is very high, and my family will be proud of me as a teacher.

My father is a teacher. He has been my role model since I was very young, and I hoped to be like my father. Therefore, I chose to go back to my hometown to be a teacher because of my father’s influence. My family also hoped that I could return to work in the countryside.

These show that family members’ suggestions and roles are important in influencing STP teacher retention. The leadership style of respected teachers also has an impact on teacher professional development. Effective leadership and appropriate attention can provide an atmosphere conducive to the development of STP teachers. For example, LIU noted:

The school principal takes good care of STP teachers, especially their work and life and has taken the lead in developing a harmonious atmosphere in the school.

Similar to school leaders, the relationship among colleagues helps maintain a strong momentum for rural teachers’ resilience development. *We often talk about students, life, and the future together. Everyone gets along very well, and it feels like home in the school.* Of course, some rural teachers have received negative suggestions, such as “my family hopes that I can find a chance to transfer back to the county town”. This kind of suggestion may have a negative influence on resilience development.

Environmental atmosphere: Professional atmosphere: positive versus negative atmosphere

The harmonious professional atmosphere is usually created by the harmonious relationship between the STP teacher and others. The professional atmosphere in this study mainly refers to class, school and local atmosphere. For example, some STP teachers said:

It is difficult for me to manage the noisy class. Sometimes I don’t know how to manage the class and how to coordinate the relationship between students. The class is always in a mess, which puts me in a bad mood.

The school atmosphere is great. The headmaster and colleagues are so nice that we are like family and we often cook and chat together.

The local villagers are very nice, and they often give us vegetables and sweet potatoes they grow. Local teacher status is pretty high, and teachers are always respected a lot.

The surrounding environment significantly affects the resilience development of STP teachers. A positive atmosphere can provide STP teachers with the basis for professional growth, and help teachers maintain their optimistic and upward development. On the contrary, a negative environment is an important reason for teacher burnout which is not conducive to supporting teachers to resist pressure and maintain strong resilience.

External support: sufficient support versus lack of support

External support, including professional support and life support, is indispensable for the resilience development of STP teachers. Professional support include the various professional training and learning activities that the teachers participated in after they started their jobs. For example, new teachers would be “guided by experienced teachers” to learn about the experience and later “dispatched to the school,” and take part in “school-wide lesson observation activities” to promote their teaching competencies. Life support refers to the salary, housing, living allowance and other support that STP teachers receive. One teacher suggested that “the salary has increased a lot in recent years, and the standard of living has also improved”. Adequate support can help teachers concentrate on their teaching. In contrast, when this support is lacking due to increased administrative work and lack of development assistance, their resilience is reduced.

Discussion

Teacher resilience has become an important global research field in recent years, but the development of resilience and its influencing factors in hard-to-staff schools remains poorly understood. This study presented several characteristic patterns of resilience development in STP rural teachers, which we described as *low-to-high*, *U-curve*, *moderate level with occasional waves*, *sustained moderate level* and *sustained high level*. These patterns are a result of the interaction result between teachers’ protective and between internal and external factors.

Teacher resilience: a multidimensional and developmental construct

Findings show that instead of being a stable characteristic, the resilience of STP teachers can be developed as they encounter challenges and difficulties throughout their professional development trajectory. It is affected by various sources of teacher resilience at different stages of teacher professional development. Day (2008) summarized the challenges faced by teachers at each stage of professional development based on teachers’ experience: 0–3 years – teacher commitment; 4–7 years – identity and self-efficacy; 8–15 years – adaptation to changes in roles and status; 16–23 years – work-life tension; 24–30 years – maintenance motivation, and approaching retirement/after 31 years – reduction of motivation (Day, 2008). The different challenges faced by teachers at different stages of professional development constitute diverse sources of resilience, that contribute to teacher resilience development. In addition, the resilience development of STP teachers who perform well in teaching has different characteristics. This conclusion is similar to that of other research. For example, researchers pointed out that each teacher’s “resilience story” is unique because resilience is a process in which multiple factors interact in a complex manner (Galea, 2018). Some researchers have also emphasized that teacher resilience is a social construction affected by unique multidimensional factors in different situations (Ungar, 2016). Of the three characteristics of teacher resilience discovered by Day (2017), context can explain the different characteristics of teacher resilience development. Contextual specificity emphasizes that teacher resilience should be understood according to specific family, class, school and work situations (Beltman et al., 2011). Even if STP teachers are in similar professional work situations, their school, family, and class situations are different, and these situations can affect teacher resilience. Studies show that harmonious cultural atmosphere in schools can help enhance the sense of belonging of teachers, strengthen the emotional connection between teachers and other groups, and encourage them to teach for a long term (Wang et al. 2018). Similarly, in this study, non-local STP teachers’ desire to resign is stronger, as they may find it hard to fit into the local culture and establish their identities in the rural schools, which might lead to professional burnout if handled unsuccessfully (Fu and Yao, 2018). Moreover, difficulties in classroom management and teaching encountered by STP teachers due to insufficient preservice preparation are also important factors for a decrease in teacher resilience (Wang et al., 2017). These factors prompt us to consider the living environment, identity development and professional learning of STP teachers when making policies. Therefore it is vital to strengthen STP teachers’ resilience through creating a positive school atmosphere, developing a strong sense of belonging to a school and a region, and providing sufficient professional development opportunities.

Building teacher resilience in hard-to-staff rural schools

STP teachers' resilience is largely the result of the interaction between protective and risk factors and between internal and external factors. Studies indicate that the resilience process involves an interactive relationship between protective and risk factors faced by teachers over time (Beltman et al., 2011) – which coincides with the conclusions of the present research. The interaction between individual and environmental factors as well as that between protective and risk factors also occur simultaneously. Environmental support and significant others can strengthen the positive influence of individual protective factors, thereby enhancing STP teacher resilience (Windle, 2011). In return, individual protective factors can also reinforce the protective effects of the external environment so that teachers perceive changes in the external environment in a more positive way. The transactional socioecological framework has been effectively used to analyze teacher resilience, indicating that teacher resilience is derived from personal characteristics and related to previous experience and specific social and cultural background. In conclusion, teacher resilience results from the dynamic interaction between the internal and external environment (Masten, 2001; Windle, 2011) (as shown in Fig. 3). The implications for policy making are that internal factors of STP teachers should be considered in the selection and recruitment of STP teachers (e.g., personality, motivation, social and emotional ability, teaching enthusiasm, and ability to reflect on learning). Policymakers should also consider external factors to support the development of STP teachers, such as increasing salaries appropriately, providing professional development opportunities, and promoting the living environment of STP teachers.

In short, the resilience development of STP teachers presents complex dynamic changes with continuous interaction between protective and risk factors, and internal and external factors over each phase of the teaching journey. The patterns of resilience development of STP teachers are diverse, and there is no set standard resilience path that can be applied to all. Resilience development involves spiral interweaving and dynamic interaction among individual protective factors, individual risk factors, environmental protective factors, and environmental risk factors which they co-constructed a relationship of tension and symbiosis (shown in Fig. 2). The dynamic interaction ultimately constructs the teachers' resilience. For example, ZHANG loves children and has always had a sense of accomplishment about her work. However, she was born and raised in the city, and the Chinese subject she taught was not her undergraduate major, so she lacked "professional knowledge". However, with favorable national policies and the improvement of rural school conditions, she gradually adapted to rural life and grew in her teaching through her strong learning and reflection skills and the training opportunities provided by the school. ZHANG's "loving children" and "teaching accomplishment" are individual protective factors, while "not enough professional knowledge" and "non-village people" constitute individual risk factors. The external support she received exemplifies environmental protection factors. What is reflected here is the dynamic interaction between the constituent factors in the resilience development, affecting ZHANG's resilience. Just as social learning theory uses personal factors, environmental conditions, learning experience, cognitive emotions and other factors to explain teacher performance, these factors often form unique combinations to determine professional performance. The combinations of factors and interactive methods here help us better understand the dynamic mechanism of the factors building resilience of teachers in hard-to-staff rural schools.

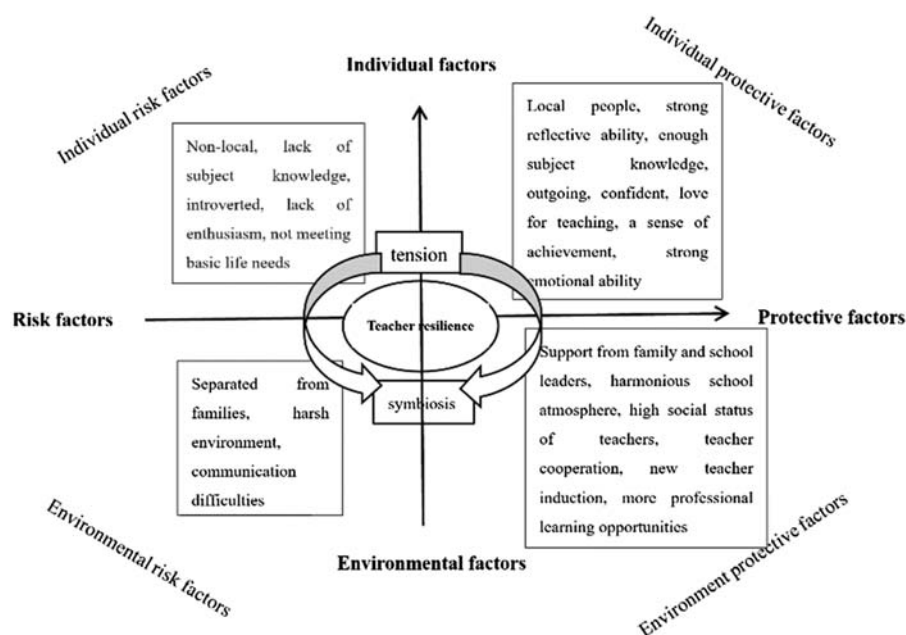


Fig. 2 Schematic diagram of the influence of factors on STP teachers' resilience development.

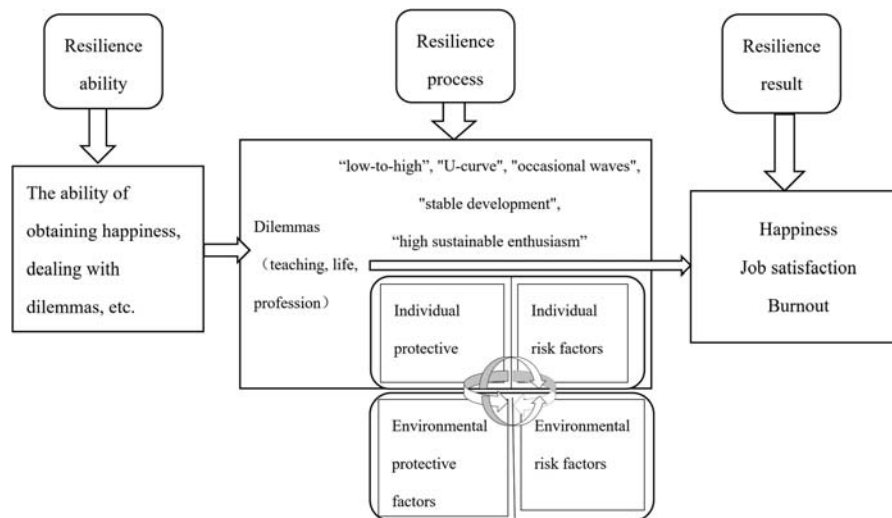


Fig. 3 A diagram of STP teachers' resilience development mechanism.

References

- Anderson, L., Olsen, B., 2006. Investigating early career urban teachers' perspectives on and experiences in professional development. *J. Teach. Educ.* (4), 359–377.
- Beltman, S., Mansfield, C., Price, A., 2011. Thriving not just surviving: a review of research on teacher resilience. *Educ. Res. Rev.* 6 (3), 185–207.
- Chudgar, A., Chandra, M., Razzaque, A., 2014. Alternative forms of teacher hiring in developing countries and its implications: a review of literature. *Teach. Teach. Educ.* 37, 150–161.
- Day, C., 2008. Committed for life? Variations in teachers' work, lives and effectiveness. *J. Educ. Change* 9 (3), 243–260.
- Day, C., 2017. School leadership as an influence on teacher quality. In: Xudong Zhu, A.L.G.H. (Ed.), *Quality of Teacher Education and Learning: Theory and Practice*. Springer, pp. 70–74.
- Fu, W.Q., Yao, Y., 2018. A study of the relationship between regional integration and job burnout of special post teachers: based on empirical studies in 13 provinces. *J. Educ. Stud.* 14 (2), 89–96 (in Chinese).
- Galea, K., 2018. *Teachers' Narratives of Resilience: Responding Effectively to Challenging Behaviour*. Springer International Publishing.
- Gao, X., 2008. Teachers' professional vulnerability and cultural tradition: a Chinese paradox. *Teach. Teach. Educ.* 24 (1), 154–165.
- Garnezy, N., 1984. Stress-resistant children: the search for protective factors. In: Stevenson, J.E. (Ed.), *Recent Research in Developmental Psychopathology* (Book Supplement to the *Journal of Child Psychology and Psychiatry*). Pergamon Press, Oxford, pp. 213–233.
- Garnezy, N., 1985. Stress-resistant children: the search for protective factors. *Rec. Res. Dev. Psychopathol.* (4), 213–233.
- Gu, Q., Day, C., 2011. Challenges to teacher resilience: conditions count. *Br. Educ. Res. J.* 1–23.
- Gu, Q., Li, Q., 2013. Sustaining resilience in times of change: stories from Chinese teachers. *Asia Pac. J. Teach. Educ.* 41 (3), 288–303.
- Han, Li, 2013. Is centralised teacher deployment more equitable? evidence from rural China. *China Econ. Rev.* 24, 65–76.
- Hargreaves, A., 1994. *Changing Teachers, Changing Times*. Cassell.
- Hirschhorn, M., 2009. Student-teacher relationship and teacher induction: Ben's story. *Teach. Dev.* (3), 73–82.
- Howard, S., Dryden, J., Johnson, B., 1999. Childhood resilience: review and critique of literature. *Oxf. Rev. Educ.* (3), 307–323.
- Hoy, A.W., Spero, R.B., 2005. Changes in teacher efficacy during the early years of teaching: a comparison of four measures. *Teach. Teach. Educ.* (4), 343–356.
- Imazeki, J., 2005. Teacher salaries and teacher attrition. *Econ. Educ. Rev.* 24 (4), 431–449.
- Johnson, J.L., Wiechelt, S.A., 2004. Introduction to the special issue on resilience. *Subst. Use Misuse* 39 (5), 657–670.
- Kaplan, H.B., 1999. Toward an understanding of resilience: a critical review of definitions and models. In: Kluwer, G.J. (Ed.), *Resilience and Development: Positive Life Adaptations*. Academic/Plenum Publishers, pp. 55–56.
- Kelchtermans, G., 2005. Teachers' emotions in educational reforms, self-understanding, vulnerable commitment and micropolitical literacy. *Teach. Teach. Educ.* (8), 995–1006.
- Knight, C., 2007. A resilience framework: perspectives for educators. *Health Educ.* (6), 543–555.
- Kyriacou, C., 1987. Teacher stress and burnout: an international review. *Educ. Res.* 29 (2), 146–152.
- Kyriacou, C., 2000. *Stress Busting for Teachers*. Stanley Thomas Ltd.
- Kyriacou, C., 2001. Teacher stress: directions for future research. *Educ. Rev.* 53 (1), 27–35.
- Lazarus, R.S., 1993. From psychological stress to the emotions: a history of changing outlooks. *Annu. Rev. Psychol.* (1), 1–21.
- Li, Q., Gu, Q., He, W.J., 2019. Resilience of Chinese teachers: why perceived work conditions and relational trust matter. *Measurement* 17 (3), 143–159.
- Liao, W., 2019. "Weekday rural teachers, weekend urban spouses and parents": a Chinese case of how alternative hiring policy influences teachers' career decisions. *Int. J. Educ. Dev.* 67, 53–63.
- Liu, S., Onwuegbuzie, A.J., 2012. Chinese teachers' work stress and their turnover intention. *Int. J. Educ. Res.* 53, 160–170.
- Luthar, S.S., Cicchetti, D., Becker, B., 2010. The construct of resilience: a critical evaluation and guidelines for future work. *Child Dev.* 71 (3), 543–562.
- Mansfield, C.F., Beltman, S., Broadley, T., Weatherby-Fell, N., 2016. Building resilience in teacher education: an evidenced informed framework. *Teach. Teach. Educ.* 54, 77–87.
- Maring, E.F., Koblinsky, S.A., 2013. Teachers' challenges, strategies, and support needs in schools affected by community violence: a qualitative study. *J. Sch. Health* (6), 379–388.
- Masten, A.S., Coatsworth, J.D., 1998. The development of competence in favorable and unfavorable environment: lessons from research on successful children. *Am. Psychol.* (2), 205–220.
- Masten, A.S., 2001. Ordinary magic: resilience processes in development. *Am. Psychol.* 56 (3), 227–238.
- Ministry of Education of China, 2006. *The Plan of Implementing the Special Teaching Position Policy in Compulsory Education Rural Schools*. Ministry of Education, Beijing, China. May 15 Retrieved from (in Chinese). http://www.moe.gov.cn/srcsite/A10/s7058/200605/t20060515_81624.html.
- Nowell, L.S., Norris, J.M., White, D.E., Moules, N.J., 2017. Thematic analysis: striving to meet the trustworthiness criteria. *Int. J. Qual. Methods* 16 (1), 1–13.

- OECD, 2005. *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. Author, Parris.
- OECD, 2014. *New Insights from TALIS 2013: Teaching and Learning in Primary and Upper Secondary Education*. OECD, Paris.
- Patton, M.Q., 2002. *Qualitative Research and Evaluation Methods*. Sage Publications, Thousand Oaks, CA.
- Richardson, G.E., 2002. The metatheory of resilience and resiliency. *J. Clin. Psychol.* (3), 307–321.
- Rolf, J., Masten, A.S., Cicchetti, D., Nuchterlein, K.H., Weintraub, S., 1990. Risk and Protective Factors in the Development of Psychopathology. In: *Competence under Stress: Risk and Protective Factors*, vol. 11, pp. 236–256.
- Rutter, M., 1985. Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. *Br. J. Psychiatry*, 147 (6), 598–611.
- Rutter, M., 2000. Resilience Reconsidered: Conceptual Considerations, Empirical Findings, and Policy Implications. In: *Handbook of Early Childhood Intervention*, vol. 2, pp. 651–682.
- Sammons, P., Day, C., Kington, A., Gu, Q., Stobart, G., Smees, R., 2007. Exploring variations in teachers' work, lives and their effects on pupils: key findings and implications from a longitudinal mixed-method study. *Br. Educ. Res. J.* 33 (5), 681–701.
- Shank, M.J., 2005. Mentoring among high school teachers: a dynamic and reciprocal group process. *Mentor. Tutoring* (1), 73–82.
- Tait, B.M., 2008. Resilience as a contributor to novice teacher success, commitment, and retention. *Teach. Educ. Q.* 57–75.
- Troman, G., Woods, P., 2001. *Primary Teachers' Stress*. Routledge.
- Turner, S.G., 2001. Resilience and social work practice: three cases studies. *Fam. Soc.* (5), 441–448.
- Ungar, M., 2016. A constructionist discourse on resilience. *Youth Soc.* 35 (3), 341–365.
- Wang, Y.L., Su, P., Gou, S.M., 2017. An empirical study on the special-post teachers' will of the mobility and attrition and its influencing factors: based on the survey in yunnan province. *Teach. Educ. Res.* 29 (5), 7–13 (in Chinese).
- Wang, H., Yan, Y.F., Yao, Y., 2018. Factors that influence STP teachers' intention to retain in rural areas—a logistic regression analysis based on A sample of special-post teachers of China. *Teach. Educ. Res.* 30 (1), 41–48.
- Watt, H., Richardson, P.W., 2008. Motivations, perceptions, and aspirations concerning teaching as a career for different types of beginning teachers. *Learn. Instruct.* 18 (5), 408–428.
- Werner, E.E., 1995. Resilience in development. *Curr. Dir. Psychol. Sci.* (3), 81–85.
- Windle, G., 2011. What is resilience? A review and concept analysis. *Rev. Clin. Gerontol.* 21 (2), 152–169.
- Yost, D.S., 2006. Reflection and self-efficacy, enhancing the retention of qualified teachers from a teacher education perspective. *Teach. Educ. Q.* (4), 59–76.

Affirming trans and queer teachers' lives: rethinking teacher professionalization and pedagogy

Jennifer C. Ingrey, Faculty of Education, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	204
Professionalism, pedagogy and the personal	204
Sexuality-diverse teachers as persons, professionals and pedagogues	206
Gender-diverse teachers as persons, professionals and pedagogues	208
Conclusion	210
Acknowledgment	211
References	211

Introduction

Teachers are thought to be caring, conscientious educators, but they are also trained to be and practice as professionals in education systems that reproduce cis-heteronormativity, alongside white supremacy, colonialism, classism and ableism. For teachers who are also gender and/or sexuality-diverse (GSD), the pursuit to becoming a teacher, doing the work of a teacher, in concert with (or not) navigating their personal lives may present extra and undue burden given that the notion of teacher professionalism historically excludes their full humanity (Khayatt, 1992; Wells, 2018; Keenan, 2021). While a teacher is legally a professional once hired in a public system (in countries/states that recognize teaching as a profession), their development as a professional (i.e., professionalization) continues throughout their career beginning in preservice programs in Faculties of Education and continuing through in-service professional development opportunities as well as elective continuing education in graduate programs. Professionalism is also enforced in quotidian school practice, classroom conduct, and extracurricular expectations both through power relations and panoptic surveillance tactics (Foucault, 1977, 1980). The professional teacher is tasked with upholding dominant ideologies and practices; they are state agents through the schooling system (Khayatt, 1992). It is this process of professionalization that arises from and contributes to public and policy understanding of the teacher as professional that requires interrogation for the teachers whose lives run counter to dominant gender and sexuality categories. For these teachers who are trans, queer, non-binary, and so forth, professionalization not only strategically overlooks, denies, or excludes but also dehumanizes. Even progressive human rights legislation—specifically protection against discrimination on the grounds of sexual or gender identity—do little to guarantee inclusive, fair, or authentic workplaces for trans and queer teachers (Ullman, 2020; Platero and Drager, 2015; Spade, 2015). As Khayatt (1992) claims, “a law never gets to people’s hearts” (p. 207). Globally, the OECD frames professionalism as the path to guarantee the aims of education systems to reach standards of student achievement, promote social cohesion and advance inclusive educational environments. In a paper responding to the challenges of the teaching profession post-pandemic, the OECD (2021) announces institutionalized professionalism, standardization of professional growth through collaboration and technology and the codification of learning as educational goals. To address the massive shortfalls in achievement during the pandemic, these goals are in pursuit of developing and garnering resilience in students to enhance their capacity for learning. Ironically, the resilience of the teacher, specifically who are GSD, is ignored. The teacher as professional is a disembodied component in the educational system.

This paper reviews seminal research on the lives of teachers in K-12 education who are lesbian, gay, queer, bisexual, transgender, non-binary or otherwise gender or sexuality-diverse identifying that provide specific and rich counter-narratives of professionalism through an interrogation of harmful and dehumanizing cis-heteronormative discourses, policy, and practice in schools. Highlighting the experiences of teachers who are sexuality-diverse precedes those who are gender diverse given the chronology of the publication dates. Additionally, a focus on transgender/trans*-formative pedagogies (Platero and Drager, 2015; Keenan, 2021) as a practice of possibilities for undoing the institutional harm of cis-heteronormativity concludes the paper and sets sights on a future of critical, conscious, and constant rethinking of gender and sexuality subjectivity. The paper surveys scholarship on the experiences and understandings of GSD professional teachers across Canada, US, and Australia. First, I provide an extension on the discussion connecting teachers’ personal lives with their professional and pedagogical practice.

Professionalism, pedagogy and the personal

Wells (2018) offers a triad of the personal, pedagogical and political as an analytic tool to explore the experiences of three gender diverse teachers who transitioned across three decades in Canada to provide informed strategies for future gender non-conforming or genderqueer teachers who might also need to navigate similar obstacles. I adapt Wells’ triadic tool to make connections among the teacher as professional, person and pedagogue. The teacher as professional requires a carving out to accommodate the person

who is the teacher, inclusive of their gender and sexuality; the pedagogue holds potential to integrate the personal, private lives and being through their work as a professional. Much of the scholarship on trans pedagogies (Platero and Drager, 2015; Keenan, 2021) also makes this claim that the presence of the trans teacher in the classroom informs their pedagogical work as trans-infused in a way that might be avoided for queer teachers, especially if they are not out at work. Jackson's (2007) work on the lives of gay and lesbian teachers in the US also provides a framework for thinking about the becoming of a teacher as both separate from and intertwined with the coming out of the closet as a queer person. Indeed, Jackson's framework leads to the integration of the professional teacher as gay, or, the becoming of a gay teacher.

To unpack the triad of the teacher as professional, person, and pedagogue is to start with how professionalism is understood and practiced. Provincial or state education legislation does little to define professionalism except to mandate it through teacher conduct, disposition, and practice. The teacher in K-12 schools is a professional, even a caring parent (as per *in loco parentis*), but one whose personal beliefs and values must not conflict with the public's construction of and investment in the professional teacher. In the Ontario, Canada context, legal precedents such as that deriving from the case of Paul Fromm warn teachers of off-duty conduct that can risk their professional status leading not only to job loss but also professional sanctions (Dickinson, 2003). As a classroom teacher in the Peel region, Fromm refused to halt his engagement in racist and white supremacist extra-curricular activities, which lead to the termination of his employment and teaching certificate. The case set a legal precedent given the curricular and pedagogical work of Fromm was not in question, but, rather, his duty to inculcate by precept and example was violated resulting in a poisoned classroom environment (Dickinson, 2003, p. 7). Indeed, professionalization, in many ways, equates to the becoming of a professional subject where personal ontologies are bracketed in favor of the dominant neoliberal teacher subject, one who is either asexual (Ferfolja, 2014) or heterosexual, marriageable, preferably feminine, and motherly (Khayatt, 1992).

Given the emphasis in the teaching profession on professionalization itself, it may seem counterintuitive to investigate something so anti-professional/personal in teachers' lives such as their gender or sexual identity. But, when a teacher's professional identity is formed through a strategic bracketing of one's sexual and gendered identity, the inequity lies in the disparity between gender and sexual diverse teachers versus cis-heterosexual identifying teachers. For the latter, no bracketing of sexual or gender identity is necessary in a system that privileges cis-heteronormativity. As Harris and Jones (2014) put it, "in school contexts, heteronormative personal details are not intimate, but queer ones are" (p. 23). For GSD teachers, identity management techniques may be a matter of survival (Keenan, 2021; Ullman, 2020; Wells, 2018), if not employability and, ultimately, self-determination.

What counts as teacher professionalism requires a rethinking given the ontological and material harm (Gray and Harris, 2014) that GSD teachers undergo in schools. Professionalism does not equate to a complete disavowal of the personal, but a reproduction of certain normative gender and sexuality persons, or subjectivities, and, thus, for GSD teachers, results in an erasure of self in an attempt to appear and be professional enough (Jackson, 2007; Wells, 2018). The limits to professional subjectivity must be expanded to account for gender and sexuality subjectivities. Currently the discursive construction of the professional teacher accounts for some subjectivity through the professional obligation to enact one's own professional judgment, or discretion. Subjectivity informs discretionary skills, combined, of course, with critically informed education and professional experience. Teaching itself is very much a discretionary, dialogic process, specifically if one subscribes to anti-oppressive pedagogies (Kumashiro, 2000). But the degree to which the person behind those discretionary skills is allowed to be presented as a person, whose humanity in its fullness is considered appropriate is very much governed by regimes of gender, race, class, and so forth. The normalized body who is classed, raced, sexed and gendered and so on in normative and dominant ways has fewer obstacles to being seen as a person because their embodiment does not conflict with the ways that the schooling system is also classed, raced, sexed and gendered along normative lines.

Certainly, what makes a professional teacher has long occupied scholars' attention. In his review of the debate between two educational theorists in the 1980s, Pitman (2012) highlights the tensions between Shulman's concern to standardize and emphasize the instrumentalization of teaching, versus Sockett's endorsement of the moral disposition and discretion of the teacher via phronesis. Updating the debate, the OECD advocates for the standardization of the profession, and the professionalization of the educator, not only to bolster public confidence in the teaching profession but also "as important public policy" (p. 11). On the other hand, Biesta (2015) is wary how neoliberal systems of management in education end up curtailing teacher discretion highlighting an urgency for protecting the subjectivity of the teacher in delivering "good" education. Wells (2018) builds on this urgency to insist that pedagogy is not fixed or technical but an embodied process that is informed by the critical disruption of cis-heteronormative gender and sexuality regimes.

Even beyond discretion or the practice of professional judgment, teachers' lives inform their practice as a professional in that they equip them with experiences, perspectives, and dispositions that make their practice alive and engaging for their students. Of course, I am referring here, in some ways, to the ideals of the profession that are articulated in the ethical and professional standards that mark teachers as pedagogues who are competent, active, and sensitive to students' well-being (see OCT). While there does exist recourse for professional misconduct, or varying degrees of disciplinary measures to regulate poor performance of teachers, some teachers may not teach in anti-oppressive ways but also not face professional sanctions. These teachers may also be technically competent but fail to challenge directly the oppressive systems that continue to harm and silence marginalized groups via race, class, sexuality, gender and so on, as well as the intersections therein. It is beyond the scope of this paper to discuss those teachers who fail to make their teaching progressive, inclusive, or dialogic (even though I do have concerns about this more broadly as well as for the implications for their students who are subsequently not engaged or themselves silenced by dominant narrative reproductions). Rather, this paper accounts for teachers whose lives would be harmed by this lack of

inclusion or acknowledgment of anti-oppressive, gender expansive, trans-informed or queering approaches to teaching given they are themselves located as gender or sexuality diverse. In actuality, the delimiting of GSD teacher agency and subjectivity leads to an equal delimiting of anti-oppressive, queer and trans-informed pedagogy.

What is most productive is to ask why or when a teacher's life is relevant to the work of schools and how it intersects with professional practice through critical engagement of curriculum and pedagogy through anti-oppressive approaches that are gender expansive, trans-informed and queering. For a moment, I turn to a remarkable example from not education, but another helping profession, psychiatry, that bridged the personal and the professional. In 1972, Dr. Harry Anonymous (now known to be Dr. Harry Tiebout) presented his account of being a "homosexual" psychiatrist and the injustices he and his colleagues had to face by literally masking himself, wearing a disguise of an unruly wig, oversized tuxedo, and a rubber mask of Nixon to make literal the masquerade of a working gay professional practitioner (Eugenios, 2022). His speech lasted a mere 10 minutes but was attributed to the eradication the following year of the removal of gay or queer as a pathological state from the DSM. In some ways, presenting himself as a ridiculous, disheveled yet articulate figure, even a figure of pathologization, he de-pathologized gay psychiatrists, decoupled their sexuality and sexual lives from the diagnosis of deviance. The life of the practitioner is most at risk when it is institutionally silenced and negated; an examination into the lives of teachers who identify as trans, queer, or gender and/or sexual diverse is to make their professional lives, and their capacity as professionals, better.

To review scholarship in teachers' work is most often an investigation into pedagogy. More personal explorations abound around attitudes, perceptions or beliefs that impact teachers' ability to support students who identify as gender diverse or queer. Much of the literature on trans or queer issues in the K-12 classroom excludes the gender or sexual identity of the teacher and focuses instead on students (Harris and Jones, 2014). But how do teachers who are trans, gender diverse or queer navigate their teaching in a climate that is historically and currently cis-heteronormative? And why is it important to think about this navigation? It is about being recognized for one's full humanity, as per Dr. Anonymous' plea: any less recognition leads to a dehumanization. Where cis-heteronormativity is written into institutional discourse, gender and sexuality diverse identities are consequently suppressed and silenced.

Key contributors have offered their own stories to describe the nuances, tensions and complexities they have experienced over the decades. This paper will review those key contributors to pay attention not only to their gains, but to also develop an appreciation for the struggles these educators have endured, all the while remaining silent or being silenced.

Sexuality-diverse teachers as persons, professionals and pedagogues

In many cases, research on the lives of teachers who are GSD are accounts offered by researchers, or, witnessed by researchers who identify also as GSD. These researchers engage in critical reflexivity by naming their situatedness to conduct ethical and anti-oppressive research. For example, Laing (2021) explains that her research on two-spirit youth who are seeking spaces for community building and political mobilization is done well only because she is able to account for herself as a whole person. Her lived experiences and identity as a queer Indigenous researcher provide her a view that enables understanding where she is aligned with her participants but also presents opportunity when she needs to learn more from them. These scholars are clear that their insights as insiders do not overwrite the stories from their participants who tell multiple, diverse, and complex tales of their own authorship. Indeed, trans and queer subjectivities and experiences are not monolithic. While they may share a common concern to disrupt cis-heteronormative colonialist structures and systems, teachers and researchers, and students who identify as GSD are also uniquely experiencing their own sexuality and gender lives.

Khayatt's (1992) book not only broke ground for articulating the presence and action of lesbian teachers in schools, whether they were out or not, she also offers personal reflection through standpoint methodology and Foucaultian analysis of regimes of power to articulate her own navigations of sexuality and professionalism. In attempt to fill the gap in existing literature, she constructed an ethnographic study of 1980s Ontario teachers who identified as lesbian (or agreed to participate in a study of lesbian teachers) in publicly funded elementary and secondary schools (p. 6). In one instance, for example, she recalls her own early career challenges of being outed unwillingly by a student:

I was caught between, on the one hand, wanting to be out, needing to claim my lesbian identity, compelled to articulate that hidden part of my experiences, and even thinking—in my innocence—that I could patch up the fragments that were my life then, and, on the other hand, having to face the terror that the label *lesbian* held for me, a teacher (p. 2).

But in these acts of hiding and protection, there are "rules" of complicity that involve not only the person who is hiding, but also the one from whom one is hiding, from whom things are hidden who have to comply with this façade and refuse to perforate its veneer (p. 3). Khayatt's methodology adapts Dorothy Smith's notion of splits or ruptures in experiences as a lesbian between ideology and felt experience "to make visible the ways in which our lives have been denied" (p. 95).

The purpose of Khayatt's (1992) book in one instance is to extract the experiences of lesbian teachers from those of gay men. It may also be as she has expressed in another introductory anecdote about her own failure or inability to confront the ideas she found perverse, specifically about how women and lesbians were being represented by the media. She explains not being able to voice her

concerns because she did not yet have the language to express herself nor the frameworks to understand on what basis were her feelings of disgust or dissent. Even when she was able to understand, as she progressed in her education as a graduate student, she found she still had very little sway or influence to enact change, specifically as a teacher (p. 3). Her book is about the persistence of lesbian teachers feeling they must be silent, that they have been silenced, even if and as they choose to be self-protective (p. 5). She provides a historical overview to explain from where this self-subscribed silencing derives both individually and socially. The nexus of gender and sexuality in lesbian teachers marks the problem not that these teachers are themselves deviant, but that the social structures upon which sexuality and gender is regulated as well as the "rejection of patriarchal prescriptions (which would keep them and other women under male control)" (p. 7) is what is deviant, in Khayatt's estimation. That is, Khayatt situates lesbian oppression, denial, silencing "within the [broader] framework of women's oppression" (p. 7) rather than homophobia. Indeed, lesbian teachers' sexuality is doubly regulated by way of patriarchal heteronormativity as well as patriarchal notions of professionalism. That Khayatt even has to refute the deviance claim highlights the era in which this study was constructed as well as the remarkable gains the field of feminism and lesbian studies has contributed to academia. Not that deviance is no longer of concern, and not that feminism itself can account for gender expansiveness or trans-informed ontologies and education, but that the lesbian teacher is not explicitly disrupting a claim of deviance. Rather, homophobia and heteronormativity operate to conceal their effects that render the lesbian teacher excluded, if not deviant.

Helpfully, Khayatt (1992) reviews the history of teachers in Ontario to situate how professionalism became entwined with the cis-heteronormative model of the male teacher or administrator with authority paired with a female teacher denoting a nurturing, marriageable and motherly caregiver, especially in elementary grades—a trope that exists currently to inform "institutional heterosexuality" (p. 61). To chart the emergence of lesbian teachers is to start with the history of women teachers. Teachers were predominantly women because of their affordability as workers versus men who were often relegated to higher-paying positions in administration given the expectation they would head families as breadwinners. Ironically, women's unions were the first to form and professionalize teaching and yet, these same women were relegated to the "worst working conditions" (p. 50) of anyone in the profession; their very inferiority propped up the professional esteem of their male counterparts. The construct of women as natural nurturers became a societal justification for their rightful place as educators, albeit underpaid and enduring employment precarity. Their capacity as heterosexual became central to socio-economic organization dependent on a certain division of labor along gendered lines. The presence of the lesbian teacher disrupts said construction as one who refused participation in the hegemonic heterosexual matrix of society (p. 5). Furthermore, Khayatt's study focuses on K-12 teachers because their legal mandate to care for children via *in loco parentis* combined with their claimed lesbian status disrupts the affirmed teacher-as-carer position. Ultimately again, the historical basis for condemnation of lesbianism derives from a heterosexual framing of women as the center of family and whose existence is tied to their capacity as procreators. In another article about strategic gender essentialism, Petersen (2000) describes this historical demarcation of gender along dualisms that are hierarchized: women are valued according to the degree they are marriageable and fertile, men are valued as persons unto themselves. Transgressions to the gender (and sexuality) order were seen to be acts against nature rendering these "rules" as universal (p. 23).

Further, in this denigration of lesbianism, women's sexual pleasure is either denied entirely or pathologized if perceived to be a threat to male sexuality (p. 17). To frame the then contemporary context of lesbian teachers in Ontario in the 1980 and early 1990s, Khayatt (1992) provides a comprehensive overview of a long history of the abjection of female sexuality and female homosexuality. Beginning with the classical interpretation of sex acts among men, Khayatt addresses the biblical regulation of sex for the purposes of procreation, highlights historical records moving to the sexologists whose work focused primarily on male homosexuality and leading to women's rights movements in the early and into the mid- 20th century where the lesbian presents as a threat only when she is able to exist on her own without a male mate.

Thus, the lesbian teachers in Khayatt's (1992) study navigated these spaces of deviance and perceived failure at heterosexual womanhood. It is key that in this history, Khayatt charts not only the construction of lesbians, but also the emergence of lesbian as an identity, something her participants grappled with or accepted partially given their own upbringing in homophobic contexts that conceptualized lesbian or same-sex female relations as moral ineptitude or abnormal mental condition. She explains both this reticence to fully identify as lesbian as well as their complicity in institutional heterosexuality through Gramscian analytics of ideology and hegemony that frame heterosexual privilege through institutional bodies of knowledge that reproduce patriarchy as neutral (p. 67). The teachers in Khayatt's study whose identities she protected with vigilance understood their existence as defying the state order of cis-heteronormativity and whose very survival depended on silence and an "invisibility [that] is historically rooted" (p. 74). They internalized shame, stigma, and engaged in actions to pass, to pursue legitimate behaviors fitting a professional teacher as a form of deliberate closeting (p. 79).

Detailing rich, diverse negotiations of understanding their sexual experiences and navigating professional teaching as a lesbian-identified or same-sex preference teacher, Khayatt's (1992) 18 participants ranged from small town to larger urban centers across Ontario teaching in multiple subject areas at varying points of their careers. Her challenges in recruiting participants were overcome only by stealth, word of mouth, trust and insider insights. Remarkably, Khayatt reworked the research process to ensure privacy of her participants from removing names from transcripts lest they be court subpoenaed to altogether avoiding pseudonym usage, thus strategically concealing their identities as they had learned to do for themselves. Khayatt presents their experiences not as representative but as an entry point to present how lesbian teachers begin to locate oneself through coping practices at school. Khayatt's participants conveyed their personal tales of coming out, remaining closeted, and every state in between through their negotiations of school spaces and interactions with school staff and students; what they did not develop was a pedagogy informed by their lesbian status.

Similar to [Khayatt \(1992\)](#), [Jackson's \(2007\)](#) work is also informed in part by standpoint theory providing some insights as a lesbian former teacher who experienced institutionalized homophobia resulting in self-silencing and denial of agency. Interviewing nine lesbian and gay teachers in an American context and collecting various curricular and pedagogical material, Jackson's work, *Unmasking Identities*, references both strategic closeting for queer teachers as well as growing rates of disclosure of sexuality despite their fears (p. 1) and the risks at coming out for the purposes of creating "a safe space for students to explore their own identities and assumptions" (p. 5). Deliberately making the connection between teachers' lives and their professional work, Jackson disrupts the de-sexed, heterosexual context of teaching to insist all teaching should sexuality affirming. [Jackson's \(2007\)](#) participants tell counter stories as professional teachers who are gay or lesbian or bisexual and who contend with barriers to becoming a complete self. These testimonios illustrate the complexity of coming out as gay, arriving at understandings of teacher identity, and consolidating queer (or gay) and teacher identity.

[Jackson \(2007\)](#) presents three broad stages of gay teacher identity formation: pre-teaching (which involves coming into gayness); closeted teaching (which involves teachers on the verge of coming out as well as those who refrain from disclosure); and post-coming out (which involves the gay poster child phase as well as the authentic phase). In the pre-teaching phase, participants' testimonios relayed a "pattern of realization, disavowal and acceptance" (p. 50) where gayness and a pursuit of teaching as an occupation align or follow similar evolutions in discoveries of self. The closeted teaching phase marked each participant's entry into teaching in that they often were confronted with the notion of the professional teacher as compulsorily heterosexual and shared a desire to secure and retain employment despite their gayness. This resulted in a bracketing of their full humanity and a fragmentation of self. This phase also re-incited trauma for those teachers who recalled their experiences in schools as gender or sexuality non-conforming students (see also [Keenan, 2017](#)).

In the latter two phases, closeted and post-coming out teaching, [Jackson's \(2007\)](#) participants remarked on fluid cycling between as well as potential for regression. The participants engaged in compensatory practices as a super teacher while still closeted, which included over-commitment to extra-curriculars and taking a protective stance to appear "beyond reproach" (p. 62). This phase expired only when the "constant vigilance about monitoring appearance" (p. 62) and self-editing to prevent "gayness from seeping into teaching" (p. 62) led to exhaustion and burnout. The post-coming out phase allowed for an integration of gayness with professional identity to merge the personal with the professional as a way to inform the pedagogical. Only at this stage did participants claim how their teaching, classroom interactions, and curriculum not only became more deliberately queer, but also could be considered good teaching due to the participation of their full humanity.

Jackson's framework is helpful if not formulaic. While allowing for regression and nuance, Jackson's framework attempts to fix, if not standardize, experiences of teachers who navigate their gay and lesbian identities alongside or despite their professional teacher identities. Responding to accusations of representing gay identity formation projects as too simplistic or reductionist (p. 79), Jackson intends instead to capture nuance and make clear the connection that the gay self is part of the teacher self through integrated authentic teaching. Her intent is to provide a starting point for interrogating institutionalized homophobia and heterosexism through the recognition of diverse sexuality teachers as professionals.

While [Jackson's \(2007\)](#) data is almost two decades more recent than [Khayatt's \(1992\)](#), both articulated the distinction of the label of gay or lesbian versus the experiences of being in same-sex relations or intimacies. Not all of Khayatt's participants were comfortable claiming the lesbian label; indeed, multiple and contested understandings of the label exist and for some it did not feel suitable. Arriving at a self-claimed identity label of both lesbian/gay and teacher is the process of becoming that [Jackson \(2007\)](#) is interested in charting, and what [Wells \(2018\)](#) calls a political act, for whom trans teachers have little room to thwart given their embodied difference when in the midst of transitioning in schools and as professionals. For queer teachers who are ready to claim gay or lesbian identity, queering may not be secured but at threat, "fleeting, contingent, and just as rapidly revoked" by state powers that legalize homophobia ([Gray and Harris, 2014](#), p. 9) or fail to ensure inclusive, comfortable climates for queer professionals ([Jackson, 2007](#)).

Gender-diverse teachers as persons, professionals and pedagogues

The connection among the teacher as person, professional and pedagogue rearticulates with an emphasis on trans-infused pedagogies for trans and gender-diverse teachers. The requisite embodiment of gender diversity, especially for trans teachers who fail to or do not pass as gender conforming distinguishes their potential from queer teachers, something [Keenan \(2021\)](#) calls a gift to students who need to see gender diversity in their teachers whom they trust and respect as good teachers. Describing his own teaching practice as both trans-informed and decolonizing, Muñoz positions his teaching as the political act of embodying space in the classroom as teacher who is both a person of color and genderqueer ([Muñoz and Garrison, 2008](#)). In his account, teaching is inseparable from the body and life of the teacher. Indeed, the roundtable discussion that is the basis of his article concludes by insisting the questions around trans and queer bodies as teachers in classrooms require an interrogation into mis-readings as well as the possibilities that might emerge for their continual engaged presence.

Situated as a suitable middle ground between sexuality and gender diverse teachers, [McCarthy's \(2003\)](#) article features one transgender teacher's experiences that emerged through coming out as lesbian and moving toward feeling more transgender despite remaining closeted at school for many years. What is most significant is the way that Kelly navigated schooling, initially grappling with coming out as a lesbian when they identified as female, deliberately wearing more feminine clothing as compensatory practice to avoid speculation over their sexuality and feeling isolated socially from their colleagues. In these early years, Kelly also refrained

from disclosing their sexuality to their students, but not without trepidation and difficulty: "I would tremble when a kid would call me a dyke and it would be because they were trying to insult me, not because they had any idea that I was really gay. That used to really scare me" (p. 173).

As Kelly's professional status became more secure, so too did their exploration of their gender expression; Kelly began wearing more male clothing and eventually came out at school to many colleagues as lesbian, but kept their emerging trans identity hidden. Kelly explains that much heterosexual discomfort around gay and lesbian people come from fears that gender lines are crossed; if cisgender status can be confirmed, the threat is removed. For Kelly, the discomfort from colleagues could not be assuaged because Kelly did not have a female identity. The questions around gender identity were valid: "it's more threatening because their judgments, their discomfort, their ignorance, and the ideas they have within that ignorance are somewhat true, in my case ... I am wearing my identity. I literally wear clothing to project a certain look that matches my gender identity" (p. 174).

While Kelly expresses an emerging trans identity, they still resort to calling their gender expression "gay," and calling themselves a woman, which coincides with McCarthy's (2003) explanation that Kelly is limited by available language. What is more concerning perhaps is that McCarthy frames transgender itself in limiting terms where it stands only to challenge the gender binary, and must be consequently limited therein (p. 180). This article is relatively new in the field of trans studies; however, its construction of trans as defying the gender binary ignores those trans folks who embrace the gender binary, gender norms, and identify not as their assigned sex but their own claimed gender identity.

Setting these concerns aside, if possible, this chapter is about illustrating varying ways GSD teachers navigate their personal lives where relevant in the classroom. For Kelly (McCarthy, 2003), meeting new students and staff who confronted them about their gender diverse appearance meant a new struggle, one that exhausted Kelly even though they responded that they hoped to develop a relationship with their students where questions like that might be better addressed: "And once I get to know the kids, ... they respect me and they know that I care about them. And it doesn't matter what I'm wearing; they don't care anymore. ... And having a teacher who creates an environment that is comfortable and enjoyable and relaxed and stimulating, in terms of learning, they so quickly forget how I'm dressed and how I look" (McCarthy, 2003, p. 174). However, this navigating takes its toll: "But it's lonely, it's lonely to feel like I do, and know that there's definitely not any place for that in the work environment. And I don't even get sad about it; it's just the reality" (p. 174). Kelly also described a vivid scene where a group of students violated her car with urine, saliva and food in retaliation for being suspended for an earlier misdemeanor. To add to the distress, the vandalism was ignored: both the administration at the school and the police advised Kelly not to press charges to avoid further and more severe retaliation.

McCarthy's (2003) article is notable for the uninterrupted passages of Kelly's experiences navigating interactions with students who are both challenging (and even violent) and also accepting. Ultimately, however, the school environment did little to support Kelly in their own trans becoming, similar to the experiences of Ullman's (2020) trans participants. Kelly's support system at home, their built-up confidence in delivering professional development workshops, and their overall love of the job provided protective barriers to the systemic and normalized homophobia, transphobia and cisgenderism in their workplace. In looking to the future, Kelly expressed required education for all teachers to learn not just about gender discrimination, but also disrupting harmful gender norms.

Not all trans or gender diverse lives may be represented fairly or fully in scholarship. Mere inclusion of trans stories does not guarantee the interpretive material will be critically informed or trans affirmative. Gender diverse and trans voices must be included with care, under-interpreted, and left to speak for themselves, or, better still, interpreted by themselves as is the case for trans researchers who also relay their testimonios as trans teachers (see Muñoz and Garrison, 2008; Keenan, 2017, 2021). Where McCarthy's (2003) paper is situated in an equity lens, it fails to account for the gendered nuance of the participant even despite the extent to which it presents the words unfiltered from the trans participant. It seems that through an attempt to analyze Kelly's ideas, McCarthy re-establishes cisnormativity and biological determinism by calling Kelly by female pronouns despite their repeated expression of their identity as trans and masculine.

In long, uninterrupted excerpts presented as counter stories, Wells' (2018) article retains the authenticity of the participants' ideas but also frames them through their pedagogical and political potential to inform gender affirming education. Refusing to rewrite their tales, Wells' concern is instead to highlight their contributions as trans teachers who have historically been denied their full humanity, struggled to be known and constructed as monstrous, if not erased entirely. Their testimonios are also an active resistance to "(re)make and (re)define their personal and professional identities" (p. 1545), to rewrite their own intelligibilities in light of the ongoing regulation of cisnormative gender regimes in schools. Scaffolding the experiences of three trans MTF (male to female) teachers who transitioned in the 1980s, 1990s, and 2000s in Canada, Wells provides the triadic model as introduced above through the personal, pedagogical and political. Conversely to Khayatt's (1992) discreet methods to protect the identities of her participants, Wells' (2018) participants refused erasure insisting on the publication of their real names despite the ongoing risk, threat, and potential trauma that anonymity might protect against. In this act of self-writing, they enact their roles as "resilient survivors and pioneers in education who have transgressed a culturally ingrained pedagogy of impossibility in order to transform bodies, identities, and minds" (p. 1546).

Like Jackson's (2007) deliberate unmasking, these trans teachers refuse "to carry the stigma of silence, invisibility, shame and denial" (Wells, 2018, p. 1551) and instead wish to pursue both "a desire for recognition, ...[and] a quest for personhood" (p. 1551). Gayle, Angela and Carol relayed stories of rupture where students recognized their gender identity in glimpses, where fears around job loss and professional sanctions especially post-transition remained, and where the pain of living not as their authentic gendered selves caused too much pain to bear. They transitioned in different contexts: one where legal protections saved Carol's job as a teacher, if only technically; one where medicalization justified Gayle's transition to colleagues and employers; and one where

school health supports eased Angela's transition at school, even though she still experienced intolerance and cisnormative resistance. All experienced "uncertainty, hyperawareness and fear" (p. 1555) and had to confront the "threat of misrecognition and the specter of the monstrous" (p. 1557) when colleagues, administrators or parents expressed mistrust post-transition. Most productively, these testimonios share also trans-infused curriculum and pedagogy where pedagogy is understood to be "a continual state of being ... [and] an emergent process" and for whom trans "bodies become the site for an engagement with a pedagogy of (im)possibility in which gender is cast from its binary moorings and the conditions for existence are transformed" (p. 1562). These trans teachers infused trans embodied pedagogy through their classroom management, interactions with students and curricular content.

Keenan (2017) writes his own perspective on trans-infused curriculum through his lived experiences as a trans teacher in elementary education. Keenan embraces the complex and "limitless possibilities of bodily expression" (p. 539) in trans folks denying universality except in the common ways that trans, gender diverse, or queer people "challenge dominant paradigms of identity categorization" (p. 539). He is also careful to insist that he speaks only for himself and not for the community of queer and trans people. It is this balance between complexity of gender identities and common ground that makes Keenan's self-expression meaningful to the purpose of developing critical trans pedagogies. Again, these pedagogies derive from the lived experiences of the trans educators who practice them. They cannot be separated out from the lives of trans folks.

Keenan (2017) also writes his own narrative, similar to Kelly's from McCarthy's (2003) article except that Kelly was responding to interview questions and Keenan (2017) is both participant and researcher. He presents early memories of dressing and styling his hair as masculine; he also encountered challenges to this expression when he entered school: "my way of expressing myself challenged a system that just didn't make a whole lot of sense to me" (p. 541), similar to Kelly whose gender expression as a transitioning teacher confronted cis-hetero gender norms followed by many of their colleagues. Keenan's (2017) stories are presented as learning moments of "embodied knowledge" (p. 542) informing his teaching. Indeed, much of Keenan's life as a student in K-12 schooling he explains taught him "to hate my genderqueerness ... to hate myself" (p. 542). Thus, his teaching as an adult of K-12 students is an active pursuit in breaking those scripts of cis-heteronormativity to break the effects of self-hatred, transphobia, and cisgenderism for young people. To be a trans educator in schools whose pedagogy is deliberately informed by one's own life as a trans person is also, for Keenan, to return to "the site of much of [their] own trauma" (p. 544).

How Keenan (2017) understands, expresses and is read as gendered, raced, and classed also implicates his professional practice. For one, he acknowledges his white and male privilege delivers status and readability, but may conceal his trans-ness, and present a potential barrier for students who are looking to teachers for support and understanding in their own gender explorations. Unscripting gender acknowledges that gender is complex and "always limited by the social and geographic location of our bodies in the world and the history of how we came to be where we are" (p. 547). It is not easily scripted, defined, fixed, neither is it biologically determined. Instead of ignoring the gendered realities of trans and queer teachers, Keenan insists "we need pedagogies that concentrate more of our efforts on inviting people to be with each other in our full humanity" (p. 553).

Answering the call to infuse trans pedagogies through education is the meta presentation of trans pedagogy in action, a dialog between Platero and Drager (2015) acting as both model and content of trans pedagogical practice. These trans educators account for where teachers' lives intersect to inform pedagogy and to provide a queer future that relies on coalition, consensus, critical and constant practice of trans pedagogies. Their accounts interweave the personal with the professional and pedagogical work of teachers, especially as they "embody an interruption to the norm" (p. 450). They live and work in Spain and contest with political, colonial, economic and linguistic limitations to their teaching as trans and via trans pedagogies. They also explain that the pedagogy of discomfort (Boler and Zembylas, 2003) means an active embrace of the uncomfortable, those sorts of questions students and children ask as curious learners, in order to disarm the unknown and to continue to foster critical understanding of gender diversity and possibilities. As trans educators, their embodied presence in the classroom may be enough to start trans-infused pedagogies, but they also must actively work to claim their intelligibility, otherwise, "gender becomes something to be ignored, rather than discussed" (Platero and Drager, 2015, p. 453). Their hypervisibility as trans teachers risks a strategic invisibility lest they actively disrupt cis-heteronormativity.

In an occasional paper series, Keenan (2021) takes his critical trans pedagogies to the next level to write in no uncertain terms a survival guide for trans educators. Entitled, "Keep Yourself Alive: Welcoming the Next Generation of Queer and Trans Educators," Keenan relays a history of trans rights interwoven with his own history of gendered and trans subjectivity. Similar to Khayatt's (1992) history of lesbian teachers, albeit shorter, both provide a context and legacy for the work they contribute to their fields. Keenan (2021) acknowledges his own teaching capacity as a former teacher and current professor who is also trans is built on the backs of forebearers in trans studies and activism; to carry the possibilities of trans-infused pedagogies further is to own one's responsibility in this pursuit.

Conclusion

Teachers who are gender or sexuality-diverse are people whose lives are curtailed, delimited, misrecognized and erased, or rendered monstrous if seen at all, but they are also professionals who not only teach students but must prepare them to live in a world that will continue to harm unless active disruption to cis-heteronormative, racist, colonialist practices are part of their pedagogy. Teachers' lives always inform their pedagogy by virtue of their ability to live as full and authentic human beings; for GSD teachers, this actualization continues to be at risk.

Teachers who are trans, queer, non-binary, genderqueer, and so on may serve as role models, but better still, represent a diversity of gender and sexuality to students that is possible to claim. GSD teachers navigate a fraught landscape of teaching to celebrate the gains, successes and disruptions of cis-heteronormalization. This paper has reviewed queer and trans teachers through their lives as presented in academic research.

While GSD teachers' lives are already oriented to doing anti-oppressive work, this remains not the burden of GSD teachers alone. Rather, being a professional begins with a legal and professional obligation to uphold the ethical and professional standards of the teaching profession, regardless of their own personal ideologies, belief systems, or lived experiences. Certainly, to have one's personal belief system align with professional standards is an easier journey, but to do critical anti-oppressive gender expansive and queering work, attention must be paid to the disposition of all teachers as well as the work that teacher does to shift their own understanding toward critical trans and queer informed pedagogy and curriculum. To articulate a framework for engaging with teacher beliefs is that of the pedagogy of discomfort from Boler and Zembylas (2003) wherein a teacher must "identify unconscious privileges as well as invisible ways in which one complies with dominant ideology" (p. 111). Only through this practice of critical discomfort is one able to apply critical frameworks informed by trans affirmation and queering to their teaching. The role of the pedagogy of discomfort is less about the discomfort of teachers whose lives have been systemically oppressed due to their sexual and gender diversity and more about how other teachers who have been in the service of the dominant ideologies around gender and sexuality must work to disrupt that service and attachment thereby experiencing discomfort. The lives of teachers are affected not only if they are minoritized; indeed, for gender expansive education and policy in schools to be successful, all teachers must take close study of their own lives to see how they are positioned as oppressors in order to undo that harmful practice.

The implications for students also remain because if teachers whose gender identity or sexuality is unsanctioned in schools come to work as incomplete, fearful, and constantly guarded, the effect for their students who might also feel the same is only reproduced or exaggerated. For others, the climate of intolerance, hate and cis-heteronormativity remains uninterrogated and stable. Students who identify, or are coming to do so, as GSD are equally fearful, guarded or feeling incomplete in schools where teachers and pedagogy are cis-heteronormative. Claiming that GSD teachers are not able to contribute their full professional selves in cis-heteronormative schools is not to be misinterpreted as an association between performance competence and teacher identity because that can lead to dangerous ground. This slippage would alight on conclusions that would only be homophobic, racist, transphobic, ableist and so on. Thus, it is important to insist that it is not teacher performance that is at risk if teachers are otherwise engaged in self-protective strategies due to their gender or sexual diversity. Rather, the health, authenticity and self-actualization of GSD teachers, students, and the entire school climate is at risk when GSD teachers feel their lives are inappropriate, pathological, deviant or monstrous. In the words of Platero and Drager (2015), critical trans (and queering) pedagogies require no "end point to this process of imagining and striving ... and pushing for a future and/or different world" (p. 459) that fails to reproduce shame, harm, and moves toward gender and sexuality affirmation, integration, and authenticity.

Acknowledgment

The author wishes to acknowledge the kind and ongoing academic and personal support of Dr. Wayne Martino in the conceptualization of this paper as a focus on trans and queer teachers' lives.

References

- Biesta, G., 2015. What is education for? On good education, teacher judgement, and educational professionalism. *Eur. J. Educ.* 50 (1), 75–87.
- Boler, M., Zembylas, M., 2003. Discomforting truths: the emotional terrain of understanding difference. In: Trifonas, P.P. (Ed.), *Pedagogies of Difference: Rethinking Education for Social Change*. Routledge Falmer.
- Dickinson, G.M., 2003. Thoughts, words and deeds: limiting teachers' free expression—the case of Paul Fromm. *Educ. Law J.* 13 (1), 131–155.
- Eugenios, J., 2022. "I Am a Homosexual. I Am a Psychiatrist": How Dr. Anonymous Changed History. NBC News. <https://www.nbcnews.com/nbc-out/out-news/-homosexual-psychiatrist-dr-anonymous-changed-history-rcna26836>.
- Ferfolja, T., 2014. Reframing queer teacher subjects: neither in nor out but present. In: Harris, A., Gray, E.M. (Eds.), *Queer Teachers, Identity and Performativity*. Palgrave Pivot, London, pp. 29–44.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison*. A. Sheridan, Trans., second ed. Vintage Books, Random House, Inc, New York, NY.
- Foucault, M., 1980. In: Gordon, C., Marshall, L., Mephram, J., Soper, K., Gordon, C. (Eds.), *Power/knowledge: Selected Interviews and Other Writings, 1972–1977*. Pantheon Books, New York, NY.
- Gray, E.M., Harris, A., 2014. Introduction: marked presence/unremarkable absence: queer teachers, "identity" and performativity. In: *Queer Teachers, Identity and Performativity*. Palgrave Pivot, London, pp. 1–10.
- Harris, A., Jones, T., 2014. Trans teacher experiences and the failure of visibility. In: *Queer Teachers, Identity and Performativity*. Palgrave Pivot, London, pp. 11–28.
- Jackson, J.M., 2007. *Unmasking Identities: An Exploration of the Lives of Gay and Lesbian Teachers*. Lexington Books.
- Keenan, H., 2017. Unscripting curriculum: toward a critical trans pedagogy. *Harv. Educ. Rev.* 87 (4), 538–556. <https://doi.org/10.17763/1943-5045-87.4.538>.
- Keenan, H., 2021. Keep Yourself alive: welcoming the next generation of queer and trans educators. *Occas. Pap. Ser.* 2021 (45), 12.
- Khayatt, M.D., 1992. *Lesbian Teachers: An Invisible Presence*. State University of New York Press.
- Kumashiro, K.K., 2000. Toward a theory of anti-oppressive education. *Rev. Educ. Res.* 70 (1), 25–53.
- Laing, M., 2021. *Urban Indigenous Youth Reframing Two-Spirit*. Routledge.
- McCarthy, L., 2003. Wearing my identity: a transgender teacher in the classroom. *Equity Excell. Educ.* 36 (2), 170–183.
- Muñoz, V., Garrison, E.K., 2008. Transpedagogies: a roundtable dialogue. *WSQ* 36 (3), 288–308.
- OECD, 2021. *Building the Future of Education*. <https://www.oecd.org/education/future-of-education-brochure.pdf>.

- Petersen, A., 2000. From Essentialism to scepticism. In: *Unmasking the Masculine. "Men" and "Identify" in a Sceptical Age*. Sage, Thousand Oaks, CA, pp. 19–40.
- Pitman, A., 2012. Professionalism and professionalisation: hostile ground for growing phronesis?. In: *Phronesis as Professional Knowledge*. Brill Sense, pp. 131–146.
- Platero, R., Drager, E.H., 2015. Two trans* teachers in Madrid: interrogating trans* formative pedagogies. *Transgender Stud. Q.* 2 (3), 447–463.
- Spade, D., 2015. *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*. Duke University Press.
- Ullman, J., 2020. Present, yet not welcomed: gender diverse teachers' experiences of discrimination. *Teach. Educ.* 31 (1), 67–83. <https://doi.org/10.1080/10476210.2019.1708315>.
- Wells, K., 2018. Transgender teachers: the personal, pedagogical, and political. *J. Homosex.* 65 (12), 1543–1581. <https://doi.org/10.1080/00918369.2017.1380989>.

The role of job demands and resources for teacher job satisfaction: insights from five European countries

Jeffrey Casely-Hayford^a, Stefan Johansson^b, Khaleda Gani Dutt^c, Lydia Kwak^a, and Anna Toropova^a, ^a Unit of Intervention & Implementation Research for Worker Health, Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden; ^b Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden; and ^c Department of Special Education, Stockholm University, Stockholm, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	213
Background	214
Teacher job satisfaction	214
The job-demands resource model	214
Relationship between job demands, resources and teacher job satisfaction	215
Country settings	215
Overview of Malta's educational system	216
Overview of Estonia's educational system	216
Overview of Slovenia's educational system	217
Overview of Austria's educational system	217
Overview of Sweden's educational system	217
Method	218
Data sources and sample	218
Measures	218
Teacher job satisfaction	219
Needs for professional development	219
Disciplinary climate	219
Workload stress	219
Teacher professional collaboration	220
Teacher autonomy	220
Teacher self-efficacy	220
Data analysis	221
Results	222
Relationships between job demands, resources and teacher job satisfaction	222
Discussion	223
Job demands and teacher job satisfaction	223
Job resources and teacher job satisfaction	224
Limitations and further research	224
Conclusions	225
Appendices	225
References	228

Introduction

The [European Commission \(2021\)](#) has highlighted the presence of motivated teachers as a crucial prerequisite of any successful education system wherein students, regardless of their socioeconomic backgrounds, are able to be successful in their educational endeavors. However, the teaching profession has been enduring a longstanding vocational crisis, illustrated by high levels of teacher turnover and attrition ([European Commission, 2021](#)). A recent study by [Federičová \(2021\)](#) examining the occupational trajectories of teachers in nineteen European countries estimated the rate of teacher turnover to be around 28% and found that the rates of turnover varied across the continent, ranging from 15% to 37%. Previous studies have shown how high rates of turnover and attrition has a disruptive impact on schools as it diminishes the social cohesion and quality of tuition provided at schools ([Carver-Thomas and Darling-Hammond, 2019](#); [Kelchtermans, 2017](#); [Ronfeldt et al., 2013](#)). A report by the [OECD, 2016a](#) corroborated the disruptive impact of teacher turnover and estimated that one of ten students in Europe attended a school with teacher shortages. In addition to this, several studies have also suggested that high rates of teacher turnover reduce the level of educational equity as lower performing schools and schools serving students from socio-economically disadvantaged backgrounds are disproportionately affected by turnover ([Johnson et al., 2011](#); [Ronfeldt et al., 2013](#)). As such, the vocational crisis in the teaching profession poses a significant challenge for education systems in Europe.

The research literature has pointed to teachers' dissatisfaction with their working environment and the diminishing status of the teaching profession as being salient factors influencing teachers' retention outcomes (Nguyen et al., 2019; Nguyen, 2021; Viac and Fraser, 2020). This has in turn directed the attention toward policy measures aimed at improving the attractiveness of the teaching profession in order to facilitate teacher recruitment and retention (European Commission, 2013; See et al., 2020). Consequently, teachers' perceptions about the teaching profession and their well-being at work has risen to the top of European and national policy agendas and teachers working conditions has been identified as an important element to improve in order to enhance the attractiveness and status of the profession (European Commission, 2021; Viac and Fraser, 2020). Job satisfaction has been highlighted as a "*decisive element*" involving positive or negative appraisals that influences teachers' attitudes and performance at work (Caprara et al., 2003, p. 823). Teacher job satisfaction is a conceptualization of teachers' affective responses to their teaching role and can be defined as "a state of mind determined by the extent to which the individual perceives her/his job-related needs to be met" (Evans, 1997, p. 328). Teacher job satisfaction also involves teachers' perceptions of job comfort and job fulfillment; two key components which capture how satisfied an individual is with their job conditions and circumstances, and the extent of their satisfaction derived from personal accomplishments within meaning areas of their teaching role (Evans, 1997). The importance of teacher job satisfaction has been emphasized in the research literature and it has been found to be associated with favorable working environments (Toropova et al., 2021; Wang et al., 2020), better occupational health outcomes (De Simone et al., 2016; Skaalvik and Skaalvik, 2009, 2011a, 2017b, 2020), teacher wellbeing (Shoshani and Eldor, 2016), higher levels of self-efficacy (Burić and Kim, 2021; Gil-Flores, 2017; Skaalvik and Skaalvik, 2014), better student outcomes (Collie et al., 2012), and crucially a willingness to remain in the profession (Björk et al., 2019; Madigan and Kim, 2021; McCarthy et al., 2014; Skaalvik and Skaalvik, 2017a,b).

Empirical evidence derived from large scale studies such as the Teaching and Learning International Survey (TALIS) conducted by the Organization for Economic Co-operation and Development (hereafter OECD) have amplified the voices of school leaders and teachers by placing their views at the forefront of educational policy analysis and development (OECD, 2019c). Results from TALIS surveys have influenced educational policy across the world and has driven the interest in comparative research comparing key indicators of effective teachers and teaching, such as teachers' job satisfaction, cross-nationally (OECD, 2020a). A recent study by Zakariya et al. (2020) ranked 38 countries in relation to the levels of teacher job satisfaction on the basis of TALIS 2018 data. In light of these results, the current study aims to investigate the relative importance of school working conditions for teachers' job satisfaction by comparing data from countries representing the highest (Austria) and lowest mean (Malta), as well as countries situated in the lower, middle, and upper parts of the distribution: Sweden (28), Estonia (19) and Slovenia (9). By doing so, this study aims to contribute to the literature by providing insights into which factors in the school work environment stand out in the countries with higher versus lower levels of teacher job satisfaction.

Background

This section provides an overview of the conceptualization of teacher job satisfaction and discusses how teachers' appraisals of their teaching conditions influences their job satisfaction. Following this, an overview of the education systems in the countries of interest to this study is provided.

Teacher job satisfaction

Teachers' satisfaction with their working environment and occupational role is influenced by numerous factors and can be assessed in different ways. The literature has conceptualized and studied teacher job satisfaction as both a facet-specific measure of teachers' satisfaction with aspects of their job and as an overall measure of teachers' satisfaction with their teaching role and conditions (Skaalvik and Skaalvik, 2009, 2011b). Most studies have however favored measuring teacher job satisfaction as an overall sense of satisfaction as facet-specific measures of job satisfaction are susceptible to individual differences and are dependent on the value placed on specific aspects of the job by the individual teacher (Skaalvik and Skaalvik, 2009, 2011b). Indeed, despite teachers' job satisfaction being the result of the combination of factors across levels, empirical evidence has consistently shown an association between school-based factors and teachers job satisfaction (Gil-Flores, 2017; Lopes and Oliveira, 2020; Sun and Xia, 2018).

The job-demands resource model

The job-demands-resource model (JDR) is a theoretical framework that has commonly been used to examine how teachers' perception about their teaching role and their well-being at work influences a range of organizational outcomes (Bakker and Demerouti, 2017; Dicke et al., 2018; Granziera et al., 2020). The model, which enables the investigation of how job characteristics influences occupational well-being and job performance, suggests that working conditions can be classified as either job demands which refer to "physical, social, or organizational aspects of the job that require sustained physical or mental effort and are therefore associated with certain physiological and psychological costs" or job resources which refer to "physical, psychological, social, or organizational aspects of the job that may [...] be functional in achieving work goals, reduce job demands and its related costs, or stimulate personal growth and development" (Demerouti et al., 2001, p. 501). Applied to a teaching context, job demands would encompass

factors such as quantitative workload, role conflict, time pressure and challenging student behavior (Hakanen et al., 2006), whereas job resources involves factors such as autonomy, social support from colleagues and school leadership and other factors that can facilitate occupational goals and professional development (Granziera et al., 2020).

The model suggests potential interactions between job demands and job resources in which the presence of a job resource can mitigate job demands and facilitate occupational well-being, motivation, and better health outcomes (Bakker et al., 2007; Bakker and Demerouti, 2017). For instance, Bakker et al. (2007) examined the interactions between job demands (student misbehavior) and job resources (job control, supervisory support, school climate, innovativeness, and appreciation) in a sample of 805 elementary, secondary, and vocational school teachers. Their findings indicated the importance of the job resources included in the study for teachers as they all provided a buffer against the negative influence of student misbehavior on teachers work engagement.

In addition, advances in the conceptualization in the JDR model has emphasized the importance of personal resources in shaping workplace experiences. Personal resources, which refer to self-evaluation of one's ability to influence one's environment, has been found to influence the effect of job demands and job resources on organizational outcomes (Collie et al., 2011; Xanthopoulou et al., 2007). In their study, Dicke et al. (2018) examined how teachers' personal resources (self-efficacy in classroom management) moderated the relationship between job demands (classroom disturbances) and emotional exhaustion and found that self-efficacy reduced the negative association observed between classroom disturbances and emotional exhaustion while simultaneously enhancing work engagement when classroom disturbances were particularly high. Taken together, these results suggest that job resources and personal resources can safeguard teachers from the negative effects of job demands and at the same time enhance teachers' sense of satisfaction at work.

Relationship between job demands, resources and teacher job satisfaction

Studies adopting the conceptual framework provided by the JDR model have consistently provided empirical evidence illustrating the differential effects of job resources and job demands on teacher job satisfaction (Madigan and Kim, 2021). Despite teaching being a rewarding occupation, it is widely acknowledged that teaching involves a high degree of job stress (Ramberg et al., 2021). Job demands such as workload stress, characterized by a high quantitative workload and an increased focus on administrative tasks, impedes teachers' ability to rest and recover (Arvidsson et al., 2016), facilitates emotional exhaustion and burnout resulting in job dissatisfaction (Skaalvik and Skaalvik, 2020). Student discipline and classroom management problems are another source of occupational stress for teachers which inhibits their goal achievement, influences their self-efficacy beliefs, and accelerates job dissatisfaction (Collie et al., 2012; Klassen and Chiu, 2010; Skaalvik and Skaalvik, 2011a; Wang et al., 2018). Moreover, job dissatisfaction results in decreasing motivation and enthusiasm for teaching which can in turn lead to higher attrition rates for the teacher profession (Buric and Moe, 2020; Madigan and Kim, 2021).

Most teachers are successful in their coping of job demands and employ a wide array of strategies such as seeking social and emotional support from colleagues (Hong, 2012; Thomas et al., 2021) however, the inability of coping successfully with these job demands invariably results in feelings of burnout (McCarthy et al., 2009). Indeed, given the inherent social connectedness found in the educational milieu, several studies have pointed to support from colleagues and school leadership as being a particularly important job resource that buffers against workload stress and facilitates job satisfaction and retention (Pomaki et al., 2010; Thomas et al., 2019). A teaching environment characterized by support from colleagues and the school leadership is indicative of a positive school climate which also positively influences job satisfaction (Aldridge and Fraser, 2015; Lopes and Oliveira, 2020). A positive school environment is also important for boosting teachers' sense of self-efficacy, which further results in job satisfaction among teachers (Malinen and Savolainen, 2016; Skaalvik and Skaalvik, 2014). Another important job resource that has been identified by the literature is teacher autonomy and studies have shown that teachers experience higher levels of job satisfaction when they are able to exercise their autonomy and self-determination (Collie et al., 2018; Skaalvik and Skaalvik, 2014).

In accordance with JDR model and on the basis of previous research, this study investigates how various job demands and job resources in the school work environment, including a personal resource of teacher self-efficacy beliefs, are related to teacher job satisfaction. The hypothesized relationships are explored for each of the countries of interest according to the conceptual model presented in Fig. 1.

Country settings

The importance of context is critical in comparative research and is essential to deepen one's understanding of the ways by which global and national policies influence the quality of the education system in different countries. International comparative studies illustrate the uniqueness of the regional locations while simultaneously pointing out their similarities and differences (Dutt, 2017). The locational comparisons draw out the diversity of the regions, socio-economic context, cultural peculiarities and highlight the points of intersection to compare and contrast educational systems. This regional analysis not only deepens the understanding of different systems but also identifies a new basis for cross-cultural discussions (Bray and Thomas, 1995). Teaching profession is not an exception when studied cross-culturally, and the context of the education system becomes an integral part of such investigation (Steiner-Khamsi, 2014). This section provides a brief account of the education system and teacher situation in Austria, Estonia, Malta, Slovenia and Sweden.

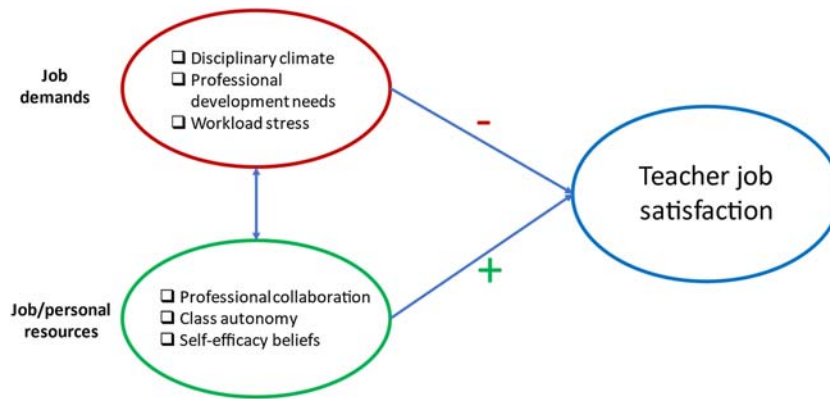


Fig. 1 Conceptual model of job demands, job and personal resources and teacher job satisfaction.

Overview of Malta's educational system

Malta has always had a strongly centralized governance of education in which the Maltese Ministry of Education and Employment has been responsible for the administration, organization, and financial management of schools. However, recently the Maltese educational system has been undergoing an educational reform process in which the governance of education is becoming increasingly decentralized (Cutajar et al., 2013; Mifsud, 2016). Schools in Malta are divided into three categories: church schools, independent schools, and state schools. State schools are free of charge and are networked into ten regional colleges each led by a College Principal (Sultana et al., 2019).

The Maltese education system consists of compulsory schooling between the age of 5 and 16 and the first cycle of formal compulsory schooling starts from year 1 to year 6. The secondary cycle involves a 5-year program in lower secondary schools to the age of 16, followed by national examinations (Secondary Education Certificate) in curricular subjects. Malta is a bilingual country, and the instruction is held in both English and Maltese. Moreover, the Maltese educational system is mostly based on the British curriculum.

Participation in international studies and student assessments have highlighted several challenges facing the Maltese education system. In 2014 Malta's Ministry for Education and Employment released the "Framework for the Education Strategy for Malta 2014–2024" which outlined Malta's strategy for overcoming the achievement gap between students attending state schools and those in non-state schools. In addition to this, the reform also outlined Malta's strategy to deal with: the high rate of early school leavers, of which Malta ranks the highest across the European Union; inclusive education; and the introduction of the Learning Outcomes Framework (Sultana et al., 2019). The Maltese government also implemented a new reform in secondary schools in 2019 titled "My Journey: Achieving through different paths," introducing three streams offering students additional vocational and applied pathways in order to encourage student retention. Some scholars have argued that these major educational reforms have been introduced without providing adequate preparations for schools and teachers to address these changes, resulting in feelings of ambiguity among teachers regarding their role and responsibility (Cutajar et al., 2013; Mifsud, 2016).

Overview of Estonia's educational system

In contrast to Malta, the Estonian education system is highly decentralized and the division of responsibility between the state, municipalities and schools is clearly defined. The overarching educational policy and national strategy, the Estonian Education Strategy 2021–35, is set by the Estonian government and the Ministry of Education and Research who are also responsible for ensuring the provision of quality education and setting the laws and regulations for education funding and quality assessment. Municipalities own the schools in their district and are responsible for their administration, regulation, and financial management. The Estonian education system consists of compulsory basic education schooling from year 1 to year 9. Secondary education is a 3-year program divided into general secondary education, provided by upper-secondary schools, and vocational upper-secondary education, provided by vocational schools (Santiago et al., 2016).

According to OECD (2020b) reports, Estonian schools rank the highest in the OECD for autonomy with 58.3% of key decisions made by Estonian schools compared to the OECD average of around 34%. School principals in Estonian schools are also highly autonomous and are responsible for the financial management and quality assurance of the school. In 2014 the national curriculum for basic and upper secondary schools underwent a reform consolidating the central role students have in their own learning (OECD, 2020b). In addition to this, the curricular reforms also enhanced the autonomy for schools and teachers by giving them free rein over how to organize the learning process according to the school context and the student's educational needs (Erss, 2018).

Despite cutting back on education expenditure and having a below-average education expenditure in comparison to other OECD countries, Estonia has performed well in international studies and student assessments (OECD, 2020b). In addition to these reforms, Estonia has made great strides in improving inclusive education and teachers' wages, resulting in an increased proportion of teachers feeling that the teaching occupation was valued by society in TALIS 2018.

Overview of Slovenia's educational system

The governance and funding of education in Slovenia is shared by the state and the municipalities. The state sets the educational policy and is also responsible for establishing public upper-secondary schools, higher education colleges, special education institutions and student accommodation. The Slovenian Ministry of Education, Science and Sport is responsible for the evaluation and implementation of regulations, compulsory basic schooling, and upper secondary school (OECD, 2016b, 2019b). Other bodies such as the National Education Institute, the Inspectorate for Education and Sport, and the National examinations center also play a part in shaping and implementing educational policy. The most recent national curriculum (2011/12) introduced a range of key competencies to be developed in general education. The Opening up Slovenia reform was introduced in 2014 and aimed to complement the national curriculum by enabling innovative and dynamic learning approaches (OECD, 2017).

Municipalities are responsible for establishing and providing kindergarten, basic schools, and adult education institutions. School leaders are appointed by school councils and are employed for a 5-year term. Compulsory education is referred to as basic school and starts from age 6 up until 15. The majority of schools in Slovenia are public and parents are able to choose which schools to enroll their children in. Schools in Slovenia have an average level of organizational autonomy and are responsible for most schooling decisions such as the organization of instruction and resource management (OECD, 2016b).

However, in contrast to Estonia, Slovenian schools have less influence over the curriculum and student assessment. One of the main challenges facing the Slovenian education system is the need to improve teaching conditions and the quality of teaching. According to OECD (2016b) reports, the teaching conditions in Slovenia include fewer teaching hours and smaller class-sizes in comparison to other OECD countries.

Overview of Austria's educational system

In Austria, the state holds the overall responsibility for the Austrian education system. The Federal Ministry of Education are responsible for the national curriculum in compulsory education, school quality development, school legislation, and staffing matters. In addition to this the Federal Institute for Research on Education, Innovation and Development of the Austrian School System collects educational data about students, teachers and schools which is used for school evaluations and policy development. Provincial authorities are responsible for implementing educational legislation and staff provision for compulsory schools in their district. At the municipal level, local authorities are responsible for the maintenance of school buildings. Austria's school governance is fragmented and complex as responsibilities are shared between the state, provinces and municipalities and further differ according to school type and education level (OECD, 2016b).

Compulsory education in Austria comprises of 9 years in primary school and lower-secondary. Most schools are publicly funded however Austria also has several fee-paying private schools and allow home-schooling for children. School choice is limited in primary education as students are assigned to schools in their geographical catchment areas. However, school choice is allowed during secondary schooling. Austria is one of two countries within the OECD that allows academic streaming at an early age and students transition to secondary education either via a compulsory secondary school or an academic secondary school (OECD, 2016b, 2017).

One of the main challenges facing the Austrian education system is reducing the achievement and gender gaps in student achievement by promoting effective and equal opportunities for all children regardless of their cultural or socioeconomic background. Student performance has been declining across PISA cycles and reports have also shown how the influence of students' socioeconomic background on their student achievement is higher than the OECD average (OECD, 2017). This has prompted the Austrian government to invest heavily in initiatives and programmes, such as The New Secondary School and language support courses, that aim to reduce the influence of early academic streaming and enhance the provision of equitable learning outcomes for all students regardless of their background.

Teaching is highly regulated in Austria and teachers' workload, worktime, and teaching time among others are regulated in teacher service codes, which in turn differ for federal teachers and teachers at provincial schools (OECD, 2016b). According to reports by the OECD (2017), Austria has a large proportion of teachers nearing retirement-age and attracting teachers to the profession is one of the main challenges facing the Austrian education system.

Overview of Sweden's educational system

Sweden's education system is one of the most decentralized education systems in the world. Since the 1990s Sweden has undergone a long process of educational reforms and restructuring and is today one of the most decentralized education systems in the world. The state sets national objectives and student learning outcomes, consolidated in legislation and steering documents, and determines educational policy which is carried out and implemented by the municipalities (OECD, 2015, 2019a). The municipalities in Sweden are responsible for the financial management and administration of schools in their district.

The Swedish education system is a market-based education system characterized by a high degree of free school-choice. School funding is attached to students rather than to individual schools which has led to a high degree of competition between schools. Compulsory schooling in Sweden comprises of 10-year schooling and most of the schools are public schools however, there is a large proportion of independent schools which are also publicly funded. Municipalities are required to ensure that schools adhere to and follow national objectives however schools and municipalities are awarded the freedom to decide how to achieve national objectives and targets (OECD, 2019a). The quality of education provided by Swedish schools and their compliance with regulations are further assessed and evaluated by the Swedish Schools Inspectorate and the Swedish National Agency for Education.

According to the [OECD \(2019a\)](#), some of the main challenges facing the Swedish education system is the need to improve students learning environment; enhance educational equity by addressing school segregation; implement effective resource allocation processes and improve teaching conditions. In order to address these challenges, Sweden has implemented several initiatives and programmes, such as the subject-specific knowledge boost programmes, which aim to strengthen learning in school environments and improve learning outcomes. Improving teaching conditions is an important challenge facing the Swedish education system as the amount of certified teachers working in the profession has reduced over time ([Adermon and Laun, 2018](#)), and the availability and distribution of qualified teachers is one of the factors driving school segregation ([Skolverket, 2020](#)). Teachers job dissatisfaction has received a lot of attention in Sweden and several initiatives and measures, such as the First Teacher reform and the Boost for Teachers program, have been implemented to improve teaching conditions.

Even though the countries of interest to this study are all situated in Europe, they represent quite diverse geographical regions as well as different educational and cultural contexts which are important for our understanding of the factors in the school work environment that relate to teacher job satisfaction. This study aims to investigate the relationship between job demands and resources (including personal resource of self-efficacy beliefs) and teacher job satisfaction in Austria, Estonia, Malta, Slovenia and Sweden.

Method

Data sources and sample

The study is based on a secondary analysis of data from The Teaching and Learning International Survey (TALIS)—a large-scale study of school work environment and learning environments as experienced by teachers and school principals ([OECD, 2020a](#)). These groups fill in a comprehensive questionnaire on their demographic information, professional characteristics as well as their perceptions on a wide array of school work conditions, job satisfaction and perceived value of the teaching profession, among others. TALIS has been conducted since 2008 on a 5-year cycle, with the next survey scheduled for 2024. Around 260,000 teachers in 15,000 schools in 48 nations were covered by the latest 2018 assessment ([OECD, 2019c](#)). Thus, TALIS offers an opportunity to compare school work conditions in a number of education systems worldwide.

TALIS selected teacher populations in each respective country using a stratified two-stage probability design, meaning that schools were randomly selected at the first stage, followed by a random selection of teachers from the above schools ([OECD, 2019d](#)). The target population included schools offering ISCED 2 educational level (representing lower secondary education level, which corresponds to grades 7–9 in most countries), their teachers and principals. For the purpose of the current study we use teacher data at ISCED 2 schools in Austria, Estonia, Malta, Slovenia and Sweden. An overview of the number of participating schools and teachers in the investigated countries is shown in [Table 1](#) below.

Measures

Many items in TALIS teacher and principal questionnaires serve as indicators of latent constructs, which are not directly observable. One example of such construct is teacher job satisfaction. By using a set of items from the TALIS teacher questionnaire, it is possible to formulate latent constructs. TALIS uses the process referred to as scaling to combine several items from a question reflecting a particular latent trait into a single scale score. In brief, scale scores were computed by conducting confirmatory factor analysis with latent variables. A detailed procedure for scale validation and scale score construction is presented in TALIS 2018 technical report ([OECD, 2019d](#)).

To enable the analysis, in which results from different countries can be compared and contrasted, TALIS conducted measurement invariance testing. The statistical procedure, conducted with the help of confirmatory factor analysis ensures that latent scales in different countries are comparable. There are several levels of comparability, referred to as measurement invariance levels, specifically, configural, metric and scalar levels ([Cheung and Rensvold, 2002](#)).

Each level of measurement invariance dictates possibilities for comparisons across groups (e.g., cross-country data analysis and reporting). Configural invariance presupposes comparisons only on a conceptual level; the configuration of items in each group is the same but statistical comparisons are unjustified. Hence, results (e.g., regression coefficients) can only be discussed through reference to each specific country. Metric invariance, in turn, allows comparisons of associations between scales in different countries. Finally, scales with scalar measurement invariance allows for comparisons of the mean scale scores between the participating countries ([OECD, 2019d](#)). For the purpose of the analysis in this study, scales with metric measurement invariance were used, which allow for comparisons of correlations and linear regression analyses.

Table 1 Participating schools and teachers in TALIS 2018 in selected countries.

	<i>Austria</i>	<i>Estonia</i>	<i>Malta</i>	<i>Slovenia</i>	<i>Sweden</i>
Schools	277	195	54	119	171
Teachers	4255	3004	1656	2094	2782

Source: [OECD, 2019d](#).

Table 2 TALIS composite scale teacher job satisfaction.*T3JOBSA: Job satisfaction, composite***T3JSENV: Job satisfaction with work environment (subscale)**

TT3G53: We would like to know how you generally feel about your job. How strongly do you agree or disagree with the following statements?

Response options: "Strongly disagree" (1), "Disagree" (2), "Agree" (3), "Strongly agree" (4).

TT3G53C*	I would like to change to another school if that were possible
TT3G53E	I enjoy working at this school
TT3G53G	I would recommend this school as a good place to work
TT3G53J	All in all, I am satisfied with my job

T3JSPRO: Job satisfaction with profession (subscale)

TT3G53: We would like to know how you generally feel about your job. How strongly do you agree or disagree with the following statements?

Response options: "Strongly disagree" (1), "Disagree" (2), "Agree" (3), "Strongly agree" (4).

TT3G53A	The advantages of being a teacher clearly outweigh the disadvantages
TT3G53B	If I could decide again, I would still choose to work as a teacher.
TT3G53D*	I regret that I decided to become a teacher
TT3G53F*	I wonder whether it would have been better to choose another profession

* Reverse coded item.

Source: OECD (2019d), p. 302.

This study employs TALIS-constructed scales (OECD, 2019e) from the teacher questionnaire representing constructs such as teacher job satisfaction, job demands (disciplinary climate, needs for professional development, workload stress), job resources (teacher professional collaboration, classroom autonomy) and a personal resource of self-efficacy beliefs.

Teacher job satisfaction

Teacher job satisfaction was the outcome variable in the study and is a composite score of the two TALIS subscales representing teachers' satisfaction with their school work environment and teachers' satisfaction with the profession. More information about the composite scale and constituting subscales is shown in Table 2 below.

Needs for professional development

The need for professional development represented one of the job demands in the school work environment and included items about teachers' perceived need in the development of content and pedagogical content knowledge. A more detailed information about the scale is provided in Table 3 below.

Disciplinary climate

Teachers' perceived disciplinary climate represented another job demands subdomain and was measured by the corresponding TALIS scale. The item regarding students striving for a pleasant learning classroom climate was reverse coded, so that the direction of the responses pointed toward an unfavorable classroom climate. The scale is described in Table 4 below.

Workload stress

Teachers' perceived workload stress represented one of the job demands and was measured by the corresponding TALIS scale. Information about the scale is provided in Table 5 below.

Table 3 TALIS scale need for professional development in subject matter in pedagogy.*T3PDPED: Needs for professional development in subject matter and pedagogy*

TT3G27: For each of the areas listed below, please indicate the extent to which you currently need professional development.

Response options: "No need at present" (1), "Low level of need" (2), "Moderate level of need" (3), "High level of need" (4).

TT3G27A	Knowledge and understanding of my subject field(s)
TT3G27B	Pedagogical competencies in teaching my subject field(s)
TT3G27C	Knowledge of the curriculum
TT3G27D	Student assessment practices
TT3G27F	Student behavior and classroom management

Source: OECD (2019d), p. 265.

Table 4 TALIS scale teachers' perceived disciplinary climate.*T3DISC: Teachers' perceived disciplinary climate*

TT3G41: How strongly do you agree or disagree with the following statements about this <target class>?

Response options: "Strongly disagree" (1), "Disagree" (2), "Agree" (3), "Strongly agree" (4).

TT3G41A	When the lesson begins, I have to wait quite a long time for students to quieten down
TT3G41B*	Students in this class take care to create a pleasant learning atmosphere
TT3G41C	I lose quite a lot of time because of students interrupting the lesson
TT3G41D	There is much disruptive noise in this classroom

* Reverse coded item.

Source: OECD (2019d), p. 334.

Table 5 TALIS scale teacher workload stress.*T3WLOAD: Workload stress*

TT3G52: Thinking about your job at this school, to what extent are the following sources of stress in your work?

Response options: "Not at all" (1), "To some extent" (2), "Quite a bit" (3), "A lot" (4).

TT3G52A	Having too much lesson preparation
TT3G52B	Having too many lessons to teach
TT3G52C	Having too much marking
TT3G52D	Having too much administrative work to do (e.g., filling out forms)
TT3G52E	Having extra duties due to absent teachers

Source: OECD (2019d), p. 319.

Table 6 TALIS scale professional collaboration in lessons among teachers.*T3COLES: Professional collaboration in lessons among teachers (subscale)*

TT3G33: On average, how often do you do the following in this school?

Response options: "Never" (1), "Once a year or less" (2), "2–4 times a year" (3), "5–10 times a year" (4), "1–3 times a month" (5), "Once a week or more" (6)

TT3G33A	Teach jointly as a team in the same class
TT3G33B	Provide feedback to other teachers about their practice
TT3G33C	Engage in joint activities across different classes and age groups (e.g., projects)
TT3G33H	Participate in collaborative professional learning

Source: OECD (2019d), p. 252.

Teacher professional collaboration

Teacher professional collaboration represented one of the job resources in the school work environment and was measured by a TALIS scale composed of items concerning lesson collaboration, joint teaching in the same class and providing feedback to each other. The scale is described in Table 6 below.

Teacher autonomy

Teacher autonomy represented one of the job resources and was measured by a TALIS scale composed of items about teachers' perceived control over, for example, course context, teaching, assessment and classroom management methods. A more detailed information about the scale is provided in Table 7 below.

Teacher self-efficacy

Teacher self-efficacy represented a personal resource and was measured by a composite TALIS scale comprised of three subscales reflecting the corresponding theoretical sub-domains: instructional, classroom management and student engagement self-efficacy (Tschannen-Moran and Hoy, 2001). More information about the composite scale and the constituting subscales is provided in Table 8 below.

Table 7 TALIS scale satisfaction with target class autonomy.*T3SATAT: Satisfaction with target class autonomy*

TT3G40: How strongly do you agree or disagree that you have control over the following areas of your planning and teaching in this <target class>?
 Response options: “Strongly disagree” (1), “Disagree” (2), “Agree” (3), “Strongly agree” (4).

TT3G40A	Determining course content
TT3G40B	Selecting teaching methods
TT3G40C	Assessing students’ learning
TT3G40D	Disciplining students
TT3G40E	Determining the amount of homework to be assigned

Source: [OECD \(2019d\)](#), p. 302.

Table 8 TALIS composite scale teacher self-efficacy.*T3SELF: Teacher self-efficacy, composite***T3SECLS: Self-efficacy in classroom management (subscale)**

TT3G34: In your teaching, to what extent can you do the following?

Response options: “Not at all” (1), “To some extent” (2), “Quite a bit” (3), “A lot” (4).

TT2G34D	Control disruptive behavior in the classroom
TT2G34F	Make my expectations about student behavior clear
TT2G34H	Get students to follow classroom rules
TT2G34I	Calm a student who is disruptive or noisy

T3SEINS: Self-efficacy in instruction (subscale)

TT3G34: In your teaching, to what extent can you do the following?

Response options: “Not at all” (1), “To some extent” (2), “Quite a bit” (3), “A lot” (4).

TT2G34C	Craft good questions for students
TT2G34J	Use a variety of assessment strategies
TT2G34K	Provide an alternative explanation, for example when students are confused
TT2G34L	Vary instructional strategies in my classroom

T3SEENG: Self-efficacy in student engagement (subscale)

TT3G34: In your teaching, to what extent can you do the following?

Response options: “Not at all” (1), “To some extent” (2), “Quite a bit” (3), “A lot” (4).

TT2G34A	Get students to believe they can do well in school work
TT2G34B	Help students value learning
TT2G34E	Motivate students who show low interest in school work
TT2G34G	Help students think critically

Source: [OECD \(2019d\)](#), p. 285.

Descriptive statistics for the above scales as well as zero-order correlations between the scales are presented in [Appendices A and B](#) respectively. Results from the correlational analysis (see [Appendix B](#)) reveal that the associations between job demands and job resources in the teacher work environment follow the expected direction, i.e., there was a negative relationship between these two domains. Some interesting deviations from this pattern could be observed, however. In Slovenia and Estonia, teachers’ needs for professional development were positively linked with higher levels of professional collaboration between teachers. In all of the countries under analysis, the relationship between job demands, job resources, a personal resource of teacher self-efficacy beliefs and teacher job satisfaction was in the expected direction, i.e., job demands were negatively linked to teacher job satisfaction, while teacher job and personal resources had a positive association with teacher job satisfaction.

Data analysis

Hierarchical linear regression was the main method used in this study. In the first step, job demands in the school work environment were related to teacher job satisfaction. In the next step job resources were entered in the model, including teacher personal resource of self-efficacy beliefs. The analyses were performed with the help of the IEA International Database Analyzer (IDB analyzer) and SPSS 27 software (IBM SPSS Statistics®). The IDB Analyzer is a software that facilitates analyses from international large-scale studies, such as TALIS and was designed by the IEA Data Processing and Research Center ([OECD, 2019e](#)). IDB Analyzer employs stratification weights and accounts for the complex sampling design (teachers nested in schools) in a given analysis. The α -level for all statistical significance tests was set at 0.05.

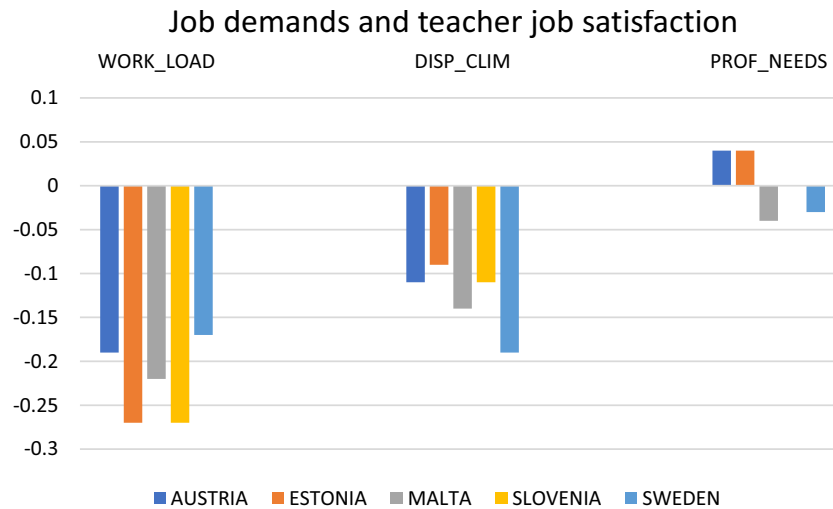


Fig. 2 Summary of the associations between job demands and teacher job satisfaction.

Results

The relationships between teacher job satisfaction and job demands and resources in the school work environment of the investigated countries are summarized in Figs. 2 and 3 below.

As can be noted from Fig. 2, workload had a negative relation with job satisfaction in all countries. The negative relationship was in fact stronger than that of disciplinary climate which was also negative for all countries. Furthermore, professional development needs were not significantly related to job satisfaction in most countries. In general, teachers who reported having adequate job resources also reported higher levels of job satisfaction. For example, Fig. 3 shows that the more efficacious teachers reported higher job satisfaction although there is some variation across countries. In the following section, results of the regression analyses (provided in Appendix C) for each of the investigated countries are described in more detail.

Relationships between job demands, resources and teacher job satisfaction

In Austria, job demands explained about 8% of the variance in teacher job satisfaction, with workload stress having the strongest negative relationship (-0.20) with job satisfaction, followed by the unfavorable disciplinary climate (-0.17). Teacher's needs for professional development in subject matter and pedagogy were not related to teacher job satisfaction.

Adding job resources to the model explained an additional 5% of the variance in teacher job satisfaction. Teacher self-efficacy beliefs and teachers' perceptions of class autonomy had a positive association, of equal strength, with teacher job satisfaction (0.14), followed by teacher professional collaboration (0.08).

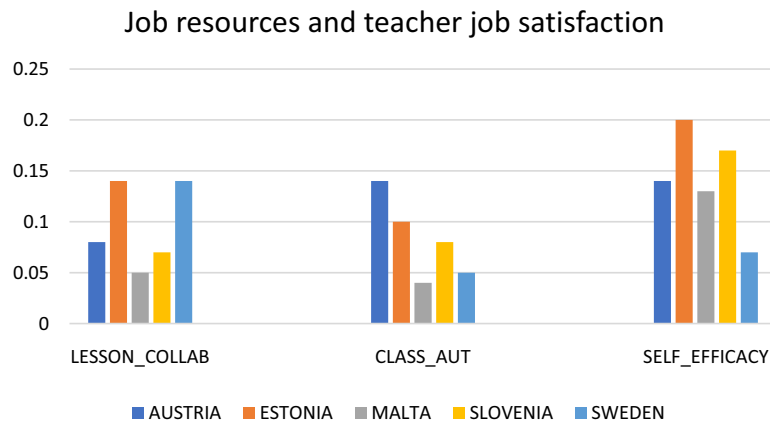


Fig. 3 Summary of the associations between job resources and teacher job satisfaction.

In **Estonia**, job demands explained 9 % of the variance in teacher job satisfaction. Workload stress had the strongest negative association with teacher job satisfaction (-0.27), followed by the disciplinary climate (-0.09). Professional development needs were not related to teacher job satisfaction.

After entering job resources in the model, the explained variance in teacher job satisfaction amounted to 18%. Teacher self-efficacy beliefs had the strongest positive association with teacher job satisfaction (0.20), followed by teacher professional collaboration (0.14), and perceptions of class autonomy (0.10).

In **Malta**, job demands explained about 9% of the variance in teacher job satisfaction, with workload stress having the strongest negative relationship with job satisfaction (-0.21), followed by the unfavorable disciplinary climate (-0.17). Teachers' needs for professional development in subject matter and pedagogy were negatively related to teacher job satisfaction (-0.06).

Adding job resources to the model explained an additional 2% of the variance in the model. Teacher self-efficacy beliefs had the strongest positive association with teacher job satisfaction (0.13). Teacher professional collaboration and teachers' perceptions of class autonomy did not demonstrate any association with teacher job satisfaction.

In **Slovenia**, job demands explained about 11 % of the variance in teacher job satisfaction, with workload stress having the strongest negative relationship with job satisfaction (-0.25), followed by the unfavorable disciplinary climate (-0.14). Professional development needs were not related to teacher job satisfaction.

Adding job resources to the model explained an additional 4% of the variance in teacher job satisfaction. Teacher self-efficacy beliefs had the strongest positive association with teacher job satisfaction (0.17), followed by teachers' perceptions of class autonomy (0.08), and teacher professional collaboration (0.07).

In **Sweden**, job demands explained about 9% of the variance in teacher job satisfaction. The unfavorable disciplinary climate had the strongest negative association with job satisfaction (-0.22), followed by workload stress (-0.18). There was no association between professional development needs and teacher job satisfaction.

Job resources contributed with an additional 3% of the explained variance in teacher job satisfaction. Teacher professional collaboration had the strongest positive association with teacher job satisfaction (0.14), followed by teacher self-efficacy beliefs (0.07). Teacher perceptions of class autonomy were unrelated with teacher job satisfaction.

Discussion

This study aimed to investigate the relationship between job demands and resources (including a personal resource of self-efficacy beliefs) and teacher job satisfaction in five European countries: Austria, Estonia, Malta, Slovenia and Sweden. Results reveal both similar and differing national patterns for the above relationships, which will be discussed in a greater detail in the following sections.

Job demands and teacher job satisfaction

Among the examined job demands in the school work environment, teacher-perceived workload stress was the strongest negative predictor of teacher job satisfaction in all countries under investigation, except for in Sweden. However alarming, this result was not unexpected, as previous research has consistently named teacher workload, and the resulting stress from it, as one of the most common reasons for teacher mental ill-health, burnout, job dissatisfaction and ultimately, attrition (Klassen and Chiu, 2010; Skaalvik and Skaalvik, 2020; Nguyen et al., 2019). Among the analyzed countries, the strongest negative relationship between workload stress and teacher job satisfaction was observed among teachers in Slovenia and Estonia. Even though there is limited research in the field for the above countries, existing studies do point out that the many reforms in the education sector have led to an increase in work demands in general and perceptions of workload in particular, in Slovenia (Rijavec et al., 2016). Similarly, in a study based on Estonian teachers, workload was the main predictor of teacher intentions to leave the profession (Saks et al., 2021). The problem of increased teacher workload due to accountability pressures and a culture of performativity (Perryman and Calvert, 2019) has been discussed in relation to Estonian (Tinn and Ümarik, 2021), Swedish (Allodi and Fischbein, 2012) and Maltese teachers (Cassar and Debono, 2000).

The next job demand in the school work environment -unfavorable student discipline—was the strongest negative predictor of teacher job satisfaction in Sweden, followed by Malta. For the rest of the countries under analysis, student discipline was the second strongest negative predictor of teacher job satisfaction. Student misbehavior has consistently appeared in the literature about Swedish teachers' stress, burnout, job dissatisfaction and attrition (Jacobsson et al., 2001; Toropova et al., 2021). For instance, a study by Jacobsson et al. (2001) using data from 928 comprehensive school teachers in Sweden found that student misbehavior was one of the strongest correlates of teacher stress, together with work demands, and also predicted work demands. In a similar vein, a study on Maltese teachers' job satisfaction and career choices highlighted how student misbehavior and a high teacher workload increased occupational stress and job dissatisfaction among teachers (Borg et al., 1991).

With regards to teacher needs for professional development, no statistically significant associations with teacher job satisfaction were demonstrated in all of the countries but Malta, where a modest link was revealed. Results for Malta find support in previous research, which emphasized both the needs for and provision of teacher professional development opportunities (Bezzina, 2002).

For other countries under investigation, these results were not anticipated, as professional development is known to be an important precursor of teacher professional commitment, effectiveness, and job satisfaction, whereas the lack of opportunities for professional growth are known to lead to work disengagement and job dissatisfaction. The key role of teacher professional development for the success of teachers, students and schools overall is reflected in research on Swedish (Karlberg and Bezzina, 2020), Slovenian (Rijavec et al., 2016), Estonian (Saks et al., 2021) and Austrian teachers (Raggl, 2019).

One of the explanations for the absence of the associations may be the measure “needs for professional development in subject matter and pedagogy,” which taps into the current perceived gaps in professional development in this area. Such needs may be viewed as an integral part of teacher professional lives (Karlberg and Bezzina, 2020), and are thus unrelated to satisfaction with the teaching job. The question which could be more pertinent for their job satisfaction is whether the opportunities to address this and other professional needs are available for teachers, which is better captured by the scale “Barriers to professional development.” Barriers to professional development demonstrated links with teacher job satisfaction in previous research (e.g., Admiraal, 2022). However, this scale could not be employed in the current study due to lack of measurement invariance, which would enable cross-country comparison.

Job resources and teacher job satisfaction

The pattern of the relationships between job resources in the school work environment and teacher job satisfaction was less consistent compared to that of job demands. A personal resource of teacher self-efficacy beliefs emerged as the strongest predictor of teacher job satisfaction in all of the countries, but Sweden. The strongest relationships between teacher self-efficacy and teacher job satisfaction was found in Estonia and Slovenia. The findings for these countries are well in line with previous research (Rijavec et al., 2016; Saks et al., 2021). A meta-analysis of studies involving secondary analysis of TALIS data examining the links between teacher self-efficacy and teacher job satisfaction found a positive association in all of the studies included in the analysis (Kasalak and Dagyar, 2020). Results showing a lack of association in Sweden were unexpected, since positive relationships between self-efficacy beliefs and teacher job satisfaction were found in a number of national and cross-national studies (Björk et al., 2019; Kasalak and Dagyar, 2020; Toropova et al., 2021). One of the explanations could be that the relationship is more complex, with teacher self-efficacy being a mechanism through which job resources influence job satisfaction (Katsantonis, 2020).

Teacher professional collaboration was the strongest predictor of teacher job satisfaction in Sweden, and the second strongest in Estonia. While these results find support in research on Swedish teachers (Jacobsson et al., 2001; Paulsrud and Wermke, 2020), results from studies on Estonian teachers are less consistent. On one hand, the lack of cooperative culture among Estonian teachers, exacerbated by the performativity demands placed on each individual teacher, has been reported (Saks et al., 2021). On the other hand, Estonian teachers may be dissatisfied with the high levels of competition, (Tinn and Ümarik, 2021) and would welcome opportunities for closer professional collaboration, as suggested by the results of this study. No links between teacher professional collaboration and teacher job satisfaction were found in Malta. These results can be explained by the lack of attention to teacher collaborative learning practices in Maltese schools and the related “teacher isolation” phenomena (Attard Tonna and Shanks, 2017).

Teachers’ perceptions of class autonomy was the strongest predictor of teacher job satisfaction in Austria, at par with teacher self-efficacy beliefs, and second strongest in Slovenia. In Malta and Sweden, teachers’ perceptions of class autonomy did not reveal any association with teacher job satisfaction. These results are rather unexpected since international research has pointed out the importance of teacher autonomy for teacher job satisfaction. For instance, a study on Maltese teachers found that teachers’ decision latitude was substantially positively associated with teacher job satisfaction (Cassar and Debono, 2000). In a similar vein, a recent study by Attard Tonna and Calleja (2021) found that a lack of autonomy, occupational stress and a perceived lack of prestige and respect for the teaching occupation contributed to teachers job dissatisfaction and desire to leave the profession in Malta.

A study by Saks et al. (2021) examining teacher retention in Estonia suggested that although teachers may be free in making classroom-based decisions they are limited by the external demands of accountability and performativity, and are thus only able to exercise “restricted autonomy.” These tensions between decision-making and control have been previously explored in the study of autonomy perceptions by the Swedish teachers (Wermke et al., 2019). In a similar vein, the difference in the relationships between teacher autonomy and job satisfaction in Austria and Sweden, with relatively high levels of teacher autonomy on the international scale, may be due to the different governance regimes resulting in complex interactions between decision-making and control in the teaching profession.

Limitations and further research

This study has a number of limitations. First, the cross-sectional design of this study does not allow for causal inferences to be made, and the associations between job demands, job resources and teacher job satisfaction may be bi-directional. Thus, teacher perceptions of job satisfaction may shape their perceptions of job demands and resources. Previous research, however, confirms the direction from the perceptions of work environment factors to job satisfaction for teachers as well as a plethora of other occupations.

Next, the scope of this study only allowed for the investigation of the most salient factors highlighted in previous research as related to teacher job satisfaction. However, teachers experience a myriad of demands and resources, and their complex inter-relationships in their everyday work life. Future research should investigate additional job demands as well as job and personal resources, including the more nuanced relationships between them. For example, examining job resources' potential as a buffer against job demands in the school work environment is of particular importance for teacher job satisfaction.

Further, teacher job satisfaction cannot solely be seen as a function of factors on one level, but rather is the result of a combination of factors on multiple levels. Reflecting an overall complexity of the teaching profession, teacher job satisfaction should be examined considering national, regional, school, classroom and individual teacher factors, and the multiple and complex interrelations between them. Factors like national variations in the organization of the educational system, teacher professional status, school leadership, as well as teacher gender, educational level and teaching experience, are only among the few important factors in explaining differences in teacher job satisfaction. This calls for more nuanced national and cross-cultural research in the area.

Conclusions

This study aimed to investigate the relationship between job demands and resources (including a personal resource of self-efficacy beliefs) and teacher job satisfaction in five European countries: Austria, Estonia, Malta, Slovenia and Sweden. Results reveal both similar and differing national patterns for the above relationships. Overall, job demands demonstrated a more consistent pattern of relationships with teacher job satisfaction across countries, while the associations between job resources and teacher job satisfaction were more varied. Despite different rankings in the teacher job satisfaction levels, teachers in the investigated countries perceived workload stress and classroom discipline as the most critical job demands, while teacher self-efficacy beliefs stood out as an important personal resource.

Results of the study suggest that one way to enhance teacher job satisfaction is to reduce demands faced by the teachers in their school work environment. Another way is to augment teacher work and personal resources, which may potentially act as buffers against high job demands, thus safeguarding teacher job satisfaction. As teacher job satisfaction is closely related to teacher retention, an issue high on the global educational agenda, results of this study could inform policy measures in the investigated countries toward enhancing teacher job satisfaction. It is also hoped that this study encourages more cross-cultural research on teacher job satisfaction and its relation to school work environment factors, in particular in those parts of the world where knowledge in this area is lacking.

Appendices

See [Appendix A–C](#).

Appendix A Zero-order correlations between the study variables.

	1.	2.	3.	4.	5.	6.
Austria						
1. Professional collaboration	1					
2. Class autonomy	0.023	1				
3. Teacher self-efficacy	0.193*	0.246*	1			
4. Need for professional development	−0.012	−0.150*	−0.229*	1		
5. Disciplinary climate	−0.044*	−0.167*	−0.300*	0.192*	1	
6. Workload stress	0.026	−0.109*	−0.066*	0.120*	0.127*	1
7. Teacher job satisfaction	0.094*	0.193*	0.218*	−0.045*	−0.192*	−0.227*
Malta						
1. Professional collaboration	1					
2. Class autonomy	0.024	1				
3. Teacher self-efficacy	0.133*	0.193*	1			
4. Need for professional development	0.003	−0.099*	−0.159*	1		
5. Disciplinary climate	0.017	−0.205*	−0.210*	0.115*	1	
6. Workload stress	−0.014	−0.062*	0.095*	0.090*	0.132*	1
7. Teacher job satisfaction	0.052*	0.122*	0.177*	−0.106*	−0.194*	−0.252*

(Continued)

Appendix A Zero-order correlations between the study variables.—cont'd

	1.	2.	3.	4.	5.	6.
Estonia						
1. Professional collaboration	1					
2. Class autonomy	−0.050*	1				
3. Teacher self-efficacy	0.226*	0.256*	1			
4. Need for professional development	0.088*	−0.109*	−0.013	1		
5. Disciplinary climate	−0.054*	−0.214*	−0.113*	0.121*	1	
6. Workload stress	0.060*	−0.017	0.063*	0.087*	0.105*	1
7. Teacher job satisfaction	0.157*	0.158*	0.245*	−0.018	−0.173*	−0.257*
Slovenia						
1. Professional collaboration	1					
2. Class autonomy	−0.032	1				
3. Teacher self-efficacy	0.175*	0.240*	1			
4. Need for professional development	0.048*	−0.106*	−0.098*	1		
5. Disciplinary climate	−0.021	−0.200*	−0.306*	0.202*	1	
6. Workload stress	0.042	−0.009	0.010	0.092*	0.134*	1
7. Teacher job satisfaction	0.078*	0.141*	0.228*	−0.072*	−0.208*	−0.298*
Sweden						
1. Professional collaboration	1					
2. Class autonomy	0.022	1				
3. Teacher self-efficacy	0.181*	0.235*	1			
4. Need for professional development	−0.010	−0.132*	−0.233*	1		
5. Disciplinary climate	−0.040	−0.192*	−0.294*	0.168*	1	
6. Workload stress	−0.086*	−0.006	−0.008	0.128*	0.065*	1
7. Teacher job satisfaction	0.194*	0.126*	0.165*	−0.105*	−0.227*	−0.230*

* Significant at 0.05 level.

Appendix B Descriptive statistics for the study variables.

Variable	N	Min	Max	Mean	SD
Austria					
Teacher job satisfaction	4112	3.64	14.11	12.08	1.93
Need for professional development	4153	6.90	14.00	9.49	1.53
Disciplinary climate	3595	6.00	14.63	8.81	2.19
Workload stress	4122	6.85	14.04	9.19	1.47
Teacher professional collaboration	4160	5.35	15.14	9.02	1.91
Teacher autonomy	3595	3.52	14.43	12.74	1.92
Teacher self-efficacy	4148	4.88	16.18	12.67	1.70
Estonia					
Teacher job satisfaction	2923	4.86	15.75	12.10	1.79
Need for professional development	2941	5.67	13.76	9.50	1.89
Disciplinary climate	2572	6.16	15.06	8.78	1.99
Workload stress	2940	6.29	15.01	9.10	2.03
Teacher professional collaboration	2938	5.67	15.89	9.12	1.80
Teacher autonomy	2578	3.16	14.88	12.70	1.84
Teacher self-efficacy	2935	3.93	16.52	12.65	1.82
Malta					
Teacher job satisfaction	1562	4.60	15.95	12.13	2.08
Need for professional development	1587	6.56	14.21	9.46	1.77
Disciplinary climate	1142	5.19	14.06	8.74	2.20
Workload stress	1570	4.56	13.23	9.17	2.22
Teacher professional collaboration	1588	7.16	17.97	9.15	2.00
Teacher autonomy	1150	4.40	15.68	12.76	2.07
Teacher self-efficacy	1584	2.35	15.88	12.70	2.24

Appendix B Descriptive statistics for the study variables.—cont'd

<i>Variable</i>	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>SD</i>
Slovenia					
Teacher job satisfaction	2024	4.28	15.59	12.08	1.97
Need for professional development	2062	6.87	15.00	9.48	1.90
Disciplinary climate	1919	5.45	14.22	8.79	1.97
Workload stress	2033	6.33	14.24	9.18	1.57
Teacher professional collaboration	2050	6.86	15.10	9.14	1.29
Teacher autonomy	1927	4.18	14.75	12.75	1.67
Teacher self-efficacy	2047	5.52	16.41	12.67	1.89
Sweden					
Teacher job satisfaction	2524	4.16	15.81	12.14	2.16
Need for professional development	2579	6.51	14.38	9.43	1.87
Disciplinary climate	1987	5.58	14.49	8.77	2.24
Workload stress	2527	5.48	13.39	9.20	1.85
Teacher professional collaboration	2579	5.41	14.14	9.06	1.79
Teacher autonomy	1991	2.77	14.05	12.71	1.70
Teacher self-efficacy	2558	6.42	16.43	12.72	1.94

Appendix C Summary of main effects hierarchical linear regression.

<i>Step and variables</i>	<i>Stand. coefficient beta</i>	<i>SE</i>	<i>t-value</i>	<i>R²</i>	<i>R² change</i>
Austria					
1. Job demands				0.08	
Disciplinary climate	−0.17	0.02	−8.23		
Prof. Development needs	0.001	0.02	0.13		
Workload stress	−0.20	0.03	−9.07		
2. Job demands				0.13	0.05
Disciplinary climate	−0.11	0.02	−5.08		
Prof. Development needs	0.04	0.02	2.38		
Workload stress	−0.19	0.02	−9.03		
Job resources					
Teacher collaboration	0.08	0.02	4.48		
Classroom autonomy	0.14	0.02	5.76		
Teacher self-efficacy beliefs	0.14	0.02	6.54		
Estonia					
1. Job demands				0.09	
Disciplinary climate	−0.14	0.02	−7.21		
Prof. Development needs	0.04	0.02	1.77		
Workload stress	−0.25	0.02	−13.79		
2. Job demands				0.18	0.09
Disciplinary climate	−0.09	0.02	−4.34		
Prof. Development needs	0.04	0.02	1.94		
Workload stress	−0.27	0.02	−16.88		
Job resources					
Teacher collaboration	0.14	0.02	6.46		
Classroom autonomy	0.10	0.02	4.67		
Teacher self-efficacy beliefs	0.20	0.02	9.88		
Malta					
1. Job demands				0.09	
Disciplinary climate	−0.17	0.04	−4.49		
Prof. Development needs	−0.06	0.02	−2.98		
Workload stress	−0.21	0.03	−7.33		

(Continued)

Appendix C Summary of main effects hierarchical linear regression.—cont'd

<i>Step and variables</i>	<i>Stand. coefficient beta</i>	<i>SE</i>	<i>t-value</i>	<i>R²</i>	<i>R² change</i>
2. Job demands				0.11	0.02
Disciplinary climate	−0.14	0.04	−3.51		
Prof. Development needs	−0.04	0.02	−1.72		
Workload stress	−0.22	0.03	−7.36		
Job resources					
Teacher collaboration	0.05	0.03	1.82		
Classroom autonomy	0.04	0.04	1.10		
Teacher self-efficacy beliefs	0.13	0.03	5.05		
Slovenia					
1. Job demands				0.11	
Disciplinary climate	−0.14	0.02	−6.13		
Prof. Development needs	0.04	0.02	−0.47		
Workload stress	−0.25	0.02	−10.30		
2. Job demands				0.15	0.04
Disciplinary climate	−0.11	0.03	−4.05		
Prof. Development needs	0.00	0.02	0.09		
Workload stress	−0.27	0.02	−11.83		
Job resources					
Teacher collaboration	0.07	0.02	3.38		
Classroom autonomy	0.08	0.02	3.88		
Teacher self-efficacy beliefs	0.17	0.03	6.34		
Sweden					
1. Job demands				0.09	
Disciplinary climate	−0.22	0.03	−6.68		
Prof. Development needs	−0.04	0.03	−1.49		
Workload stress	−0.18	0.04	−4.49		
2. Job demands				0.12	0.03
Disciplinary climate	−0.19	0.04	−5.02		
Prof. Development needs	−0.03	0.03	−0.92		
Workload stress	−0.17	0.04	−4.22		
Job resources					
Teacher collaboration	0.14	0.03	5.12		
Classroom autonomy	0.05	0.03	1.64		
Teacher self-efficacy beliefs	0.07	0.03	2.33		

References

- Adermon, A., Laun, L., 2018. Bristyrken i offentlig verksamhet: Var arbetar de utbildade? Retrieved from Uppsala, Sweden. <https://www.ifau.se/globalassets/pdf/se/2018/r-2018-19-bristyrken-i-offentlig-verksamhet.pdf>.
- Admiraal, W., 2022. Teachers' job demands, resources and job satisfaction: secondary analyses of TALIS 2018 data from Flanders and the Netherlands. *Eur. J. Educ. Stud.* 9 (2).
- Aldridge, J.M., Fraser, B.J., 2015. Teachers' views of their school climate and its relationship with teacher self-efficacy and job satisfaction. *Learn. Environ. Res.* 19 (2), 291–307. <https://doi.org/10.1007/s10984-015-9198-x>.
- Allodi, M., Fischbein, S., 2012. Teachers' perceptions of their work environment in Swedish junior high schools. *Res. Comp. Int. Educ.* 7 (3), 376–393.
- Arvidsson, I., Håkansson, C., Karlson, B., Björk, J., Persson, R., 2016. Burnout among Swedish school teachers—a cross-sectional analysis. *BMC Publ. Health* 16 (1), 823. <https://doi.org/10.1186/s12889-016-3498-7>.
- Attard Tonna, M., Shanks, R., 2017. The importance of environment for teacher professional learning in Malta and Scotland. *Eur. J. Teach. Educ.* 40 (1), 91–109.
- Attard Tonna, M., Calleja, J., 2021. An investigation of the professional behaviour, status, career and identities of teachers in Malta. *Eur. J. Teach. Educ.* 1–19. <https://doi.org/10.1080/02619768.2021.1889508>.
- Bakker, A.B., Demerouti, E., 2017. Job demands-resources theory: taking stock and looking forward. *J. Occup. Health Psychol.* 22 (3), 273–285. <https://doi.org/10.1037/ocp0000056>.
- Bakker, A.B., Hakanen, J.J., Demerouti, E., Xanthopoulou, D., 2007. Job resources boost work engagement, particularly when job demands are high. *J. Educ. Psychol.* 99 (2), 274–284. <https://doi.org/10.1037/0022-0663.99.2.274>.
- Bezzina, C., 2002. Rethinking teachers' professional development in Malta: agenda for the twenty-first century. *J. In-Service Educ.* 28 (1), 57–78.
- Björk, L., Stengård, J., Söderberg, M., Andersson, E., Wastensson, G., 2019. Beginning teachers' work satisfaction, self-efficacy and willingness to stay in the profession: a question of job demands-resources balance? *Teach. Teach.* 25 (8), 955–971. <https://doi.org/10.1080/13540602.2019.1688288>.
- Borg, M.G., Riding, R.J., Falzon, J.M., 1991. Stress in teaching: a study of occupational stress and its determinants, job satisfaction and career commitment among primary schoolteachers. *Educ. Psychol.* 11 (1), 59–75. <https://doi.org/10.1080/0144341910110104>.

- Bray, M., Thomas, R.M., 1995. Levels of comparison in educational studies: different insights from different literatures and the value of multilevel analysis. *Harv. Educ. Rev.* 65 (3), 472–490.
- Burić, I., Kim, L., 2021. Job satisfaction predicts teacher self-efficacy and the association is invariant: examinations using TALIS 2018 data and longitudinal Croatian data. *Teach. Educ.* 105, 103406. <https://doi.org/10.1016/j.tate.2021.103406>.
- Buric, I., Moe, A., 2020. What makes teachers enthusiastic: the interplay of positive affect, self-efficacy and job satisfaction. *Teach. Teach. Educ.* 89.
- Caprara, G., Barbaranelli, C., Borgogni, L., Steca, P., 2003. Efficacy beliefs as determinants of teachers' job satisfaction. *J. Educ. Psychol.* 95, 821–832. <https://doi.org/10.1037/0022-0663.95.4.821>.
- Carver-Thomas, D., Darling-Hammond, L., 2019. The trouble with teacher turnover: how teacher attrition affects students and schools. *Educ. Pol. Anal. Arch.* 27. <https://doi.org/10.14507/epaa.27.3699>.
- Cassar, V., Debono, M., 2000. Influences of empowerment dimensions in an educational context: an empirical investigation amongst Maltese teachers. *Mediterr. J. Educ. Stud.* 5 (1), 177–189.
- Cheung, G., Rensvold, R., 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model.* 9 (2), 233–255.
- Collie, R.J., Shapka, J.D., Perry, N.E., 2011. Predicting teacher commitment: the impact of school climate and social-emotional learning. *Psychol. Sch.* 48 (10), 1034–1048. <https://doi.org/10.1002/pts.20611>.
- Collie, R.J., Shapka, J.D., Perry, N.E., 2012. School climate and social-emotional learning: predicting teacher stress, job satisfaction, and teaching efficacy. *J. Educ. Psychol.* 104 (4), 1189–1204. <https://doi.org/10.1037/a0029356>.
- Collie, R.J., Granziera, H., Martin, A.J., 2018. Teachers' perceived autonomy support and adaptability: an investigation employing the job demands-resources model as relevant to workplace exhaustion, disengagement, and commitment. *Teach. Teach. Educ.* 74, 125–136. <https://doi.org/10.1016/j.tate.2018.04.015>.
- Cutajar, M., Bezzina, C., James, C., 2013. Educational reforms in Malta: a missed opportunity to establish distributed governance. *Manag. Educ.* 27 (3), 118–124. <https://doi.org/10.1177/0892020613490872>.
- De Simone, S., Cicotto, G., Lampis, J., 2016. Occupational stress, job satisfaction and physical health in teachers. *Eur. Rev. Appl. Psychol.* 66 (2), 65–77. <https://doi.org/10.1016/j.erap.2016.03.002>.
- Demerouti, E., Bakker, A.B., Nachreiner, F., Schaufeli, W.B., 2001. The job demands-resources model of burnout. *J. Appl. Psychol.* 86 (3), 499–512.
- Dicke, T., Stebner, F., Linniger, C., Kunter, M., Leutner, D., 2018. A longitudinal study of teachers' occupational well-being: applying the job demands-resources model. *J. Occup. Health Psychol.* 23 (2), 262–277. <https://doi.org/10.1037/ocp0000070>.
- Dutt, K.D., 2017. The Role of Adult Literacy in Transforming the Lives of Women in Rural India: Overcoming Gender Inequalities. Comparative Case Studies in Bhilwara District Rajasthan & Howrah District West Bengal India. Doctoral Thesis in International and Comparative Education at Stockholm University, Sweden 2017.
- Erss, M., 2018. "Complete freedom to choose within limits"—teachers' views of curricular autonomy, agency and control in Estonia, Finland and Germany. *Curric. J.* 29 (2), 238–256. <https://doi.org/10.1080/09585176.2018.1445514>.
- European Commission, 2013. Study on Policy Measures to Improve the Attractiveness of the Teaching Profession in Europe. Retrieved from Luxembourg.
- European Commission, 2021. Teachers in Europe Careers, Development and Well-Being. Retrieved from Luxembourg. https://eacea.ec.europa.eu/national-policies/eurydice/sites/default/files/teachers_in_europe_2020_foreword.pdf.
- Evans, L., 1997. Addressing problems of conceptualization and construct validity in researching teachers' job satisfaction. *Educ. Res.* 39 (3), 319–331. <https://doi.org/10.1080/0013188970390307>.
- Federičová, M., 2021. Teacher turnover: what can we learn from Europe? *Eur. J. Edu.* 56 (1), 102–116. <https://doi.org/10.1111/ejed.12429>.
- Gil-Flores, J., 2017. The role of personal characteristics and school characteristics in explaining teacher job satisfaction. *Revista de Psicodidáctica* 22 (1), 16–22. <https://doi.org/10.1387/RevPsicodidact.15501>.
- Granziera, H., Collie, R., Martin, A., 2020. Understanding teacher wellbeing through job demands-resources theory. In: Mansfield, C.F. (Ed.), *Cultivating Teacher Resilience: International Approaches, Applications and Impact*. Springer Singapore, Singapore, pp. 229–244.
- Hakanen, J.J., Bakker, A.B., Schaufeli, W.B., 2006. Burnout and work engagement among teachers. *J. Sch. Psychol.* 43 (6), 495–513. <https://doi.org/10.1016/j.jsp.2005.11.001>.
- Hong, J.Y., 2012. Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teach. Teach.* 18 (4), 417–440. <https://doi.org/10.1080/13540602.2012.696044>.
- Jacobsson, C., Pousette, A., Thylefors, I., 2001. Managing stress and feelings of mastery among Swedish comprehensive school teachers. *Scand. J. Educ. Res.* 45 (1), 37–53.
- Johnson, S.M., Kraft, M.A., Papay, J.P., 2011. How context matters in high-need schools: the effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teach. Coll. Rec.* 114 (10), 1–39.
- Karlberg, M., Bezzina, C., 2020. The professional development needs of beginning and experienced teachers in four municipalities in Sweden. *Prof. Dev. Educ.* 1–18.
- Kasalak, G., Dagyar, M., 2020. The Relationship between teacher self-efficacy and teacher job satisfaction: a meta-analysis of the teaching and learning international survey (TALIS). *Educ. Sci.: Theory Pract.* 20 (3), 16–33.
- Katsantonis, I.G., 2020. Investigation of the impact of school climate and teachers' self-efficacy on job satisfaction: a cross-cultural approach. *Eur. J. Investig. Health Psychol. Educ.* 10 (1), 119–133. <https://doi.org/10.3390/ejihpe10010011>.
- Kelchtermans, G., 2017. "Should I stay or should I go?": unpacking teacher attrition/retention as an educational issue. *Teach. Teach.* 23 (8), 961–977. <https://doi.org/10.1080/13540602.2017.1379793>.
- Klassen, R.M., Chiu, M.M., 2010. Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102 (3), 741–756. <https://doi.org/10.1037/a0019237>.
- Lopes, J., Oliveira, C., 2020. Teacher and school determinants of teacher job satisfaction: a multilevel analysis. *Sch. Effect. Sch. Improv.* 31 (4), 641–659. <https://doi.org/10.1080/09243453.2020.1764593>.
- Madigan, D., Kim, L., 2021. Towards an understanding of teacher attrition: a meta-analysis of burnout, job satisfaction, and teachers' intentions to quit. *Teach. Teach. Educ.* 105, 103425. <https://doi.org/10.1016/j.tate.2021.103425>.
- Malinen, O.P., Savolainen, H., 2016. The effect of perceived school climate and teacher efficacy in behavior management on job satisfaction and burnout: a longitudinal study. *Teach. Teach. Educ.* 60, 144–152. <https://doi.org/10.1016/j.tate.2016.08.012>.
- McCarthy, C.J., Lambert, R.G., O'Donnell, M., Melendres, L.T., 2009. The relation of elementary teachers' experience, stress, and coping resources to burnout symptoms. *Elem. Sch. J.* 109 (3), 282–300. <https://doi.org/10.1086/592308>.
- McCarthy, C.J., Lambert, R.G., Reiser, J., 2014. Vocational concerns of elementary teachers: stress, job satisfaction, and occupational commitment. *J. Employ. Counsel.* 51 (2), 59–74. <https://doi.org/10.1002/j.2161-1920.2014.00042.x>.
- Mifsud, D., 2016. "Decentralised" neoliberalism and/or "masked" re-centralisation? The policy to practice trajectory of Maltese school reform through the lens of neoliberalism and Foucault. *J. Educ. Pol.* 31 (4), 443–465. <https://doi.org/10.1080/02680939.2015.1121409>.
- Nguyen, T.D., Pham, L., Springer, M., et al., 2019. The factors of teacher attrition and retention: an updated and expanded meta-analysis of the literature. *EdWorkingPaper* 19–49. <https://doi.org/10.26300/cdf3-4555>.

- Nguyen, T.D., 2021. Linking school organizational characteristics and teacher retention: evidence from repeated cross-sectional national data. *Teach. Teach. Educ.* 97, 103220. <https://doi.org/10.1016/j.tate.2020.103220>.
- OECD, 2015. Improving Schools in Sweden: An OECD Perspective. Retrieved from Organisation for Economic Cooperation and Development (OECD). <http://www.oecd.org/education/school/improving-schools-in-sweden-an-oecd-perspective.htm>.
- OECD, 2016a. PISA 2015 Results (Volume II): Policies and Practices for Successful Schools. OECD Publishing, Paris.
- OECD, 2016b. Education Policy Outlook in Slovenia. Retrieved from Paris. <https://www.oecd-ilibrary.org/content/paper/9d472195-en>.
- OECD, 2016c. OECD Reviews of School Resources: Austria. OECD Publishing, Paris.
- OECD, 2017. Education policy outlook: Austria. In: Education Policy Outlook 2019: Working Together to Help Students Achieve Their Potential. OECD Publishing, Paris, p. 28.
- OECD, 2019a. Education Policy Outlook: Sweden. Retrieved from Paris. <https://www.oecd-ilibrary.org/content/component/84003fa5-en>.
- OECD, 2019b. Slovenia. In: Education Policy Outlook 2019: Working Together to Help Students Achieve Their Potential. OECD Publishing, Paris, pp. 490–497.
- OECD, 2019c. TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. Retrieved from Paris. <https://www.oecd-ilibrary.org/content/publication/1d0bc92a-en>.
- OECD, 2019d. TALIS 2018 Technical Report. OECD, Paris. Retrieved from. http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf.
- OECD, 2019e. TALIS 2018 and TALIS Starting Strong 2018 User Guide. Retrieved from. https://www.oecd.org/education/talis/TALIS_2018-TALIS_Starting_Strong_2018_User_Guide.pdf. OECD, Paris.
- OECD, 2020a. TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals. Retrieved from Paris. <https://www.oecd-ilibrary.org/content/publication/19cf08df-en>.
- OECD, 2020b. Education Policy Outlook in Estonia (13). Retrieved from Paris. <https://www.oecd-ilibrary.org/content/paper/9d472195-en>.
- Paulsrud, D., Wermke, W., 2020. Decision-making in context: Swedish and Finnish teachers' perceptions of autonomy. *Scand. J. Educ. Res.* 64 (5), 706–727.
- Perryman, J., Calvert, G., 2019. What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *Br. J. Educ. Stud.* 68 (1), 3–23. <https://doi.org/10.1080/00071005.2019.1589417>.
- Pomaki, G., DeLongis, A., Frey, D., Short, K., Woehrlé, T., 2010. When the going gets tough: direct, buffering and indirect effects of social support on turnover intention. *Teach. Teach. Educ.* 26 (6), 1340–1346. <https://doi.org/10.1016/j.tate.2010.03.007>.
- Raggl, A., 2019. Small rural schools in Austria: potentials and challenges. In: *Geographies of Schooling*. Springer, Cham, pp. 251–263.
- Ramberg, J., Läftman, S.B., Nilbrink, J., Olsson, G., Toivanen, S., 2021. Job strain and sense of coherence: associations with stress-related outcomes among teachers. *Scand. J. Publ. Health*. <https://doi.org/10.1177/14034948211011812>, 14034948211011812-14034948211011812.
- Rijavec, M., Pečjak, S., Jurčec, L., Gradišek, P., 2016. Money and career or calling? Intrinsic vs. extrinsic work orientations and job satisfaction of Croatian and Slovenian teachers. *Croat. J. Educ.* 18 (1), 201–223.
- Ronfeldt, M., Loeb, S., Wyckoff, J., 2013. How teacher turnover harms student achievement. *Am. Educ. Res. J.* 50 (1), 4–36. <https://doi.org/10.3102/0002831212463813>.
- Saks, K., Hunt, P., Leijen, A., Lepp, L., 2021. To stay or not to stay: an empirical model for predicting teacher persistence. *Br. J. Educ. Stud.* 1–25.
- Santiago, P., Levitas, A., Radó, P., Shewbridge, C., 2016. OECD Reviews of School Resources: Estonia 2016. OECD Publishing, Paris.
- See, B.H., Morris, R., Gorard, S., Kokotsaki, D., Abdi, S., 2020. Teacher recruitment and retention: a critical review of international evidence of most promising. *Intervention* 10 (10), 262.
- Shoshani, A., Eldor, L., 2016. The informal learning of teachers: learning climate, job satisfaction and teachers' and students' motivation and well-being. *Int. J. Educ. Res.* 79, 52–63. <https://doi.org/10.1016/j.ijer.2016.06.007>.
- Skaalvik, E.M., Skaalvik, S., 2009. Does school context matter? Relations with teacher burnout and job satisfaction. *Teach. Teach. Educ.* 25 (3), 518–524. <https://doi.org/10.1016/j.tate.2008.12.006>.
- Skaalvik, E.M., Skaalvik, S., 2011a. Teacher job satisfaction and motivation to leave the teaching profession: relations with school context, feeling of belonging, and emotional exhaustion. *Teach. Teach. Educ.* 27 (6), 1029–1038. <https://doi.org/10.1016/j.tate.2011.04.001>.
- Skaalvik, E.M., Skaalvik, S., 2011b. Teachers' feeling of belonging, exhaustion, and job satisfaction: the role of school goal structure and value consonance. *Anxiety Stress Coping* 24 (4), 369–385. <https://doi.org/10.1080/10615806.2010.544300>.
- Skaalvik, E.M., Skaalvik, S., 2014. Teacher self-efficacy and perceived autonomy: relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychol. Rep.* 114 (1), 68–77. <https://doi.org/10.2466/14.02.PR0.114k14w0>.
- Skaalvik, E.M., Skaalvik, S., 2017a. Motivated for teaching? Associations with school goal structure, teacher self-efficacy, job satisfaction and emotional exhaustion. *Teach. Teach. Educ.* 67, 152–160. <https://doi.org/10.1016/j.tate.2017.06.006>.
- Skaalvik, E.M., Skaalvik, S., 2017b. Still motivated to teach? A study of school context variables, stress and job satisfaction among teachers in senior high school. *Soc. Psychol. Educ.* 20, 15–37. <https://doi.org/10.1007/s11218-016-9363-9>.
- Skaalvik, E.M., Skaalvik, S., 2020. Teacher burnout: relations between dimensions of burnout, perceived school context, job satisfaction and motivation for teaching. A longitudinal study. *Teach. Teach.* 26 (7–8), 602–616. <https://doi.org/10.1080/13540602.2021.1913404>.
- Skolverket, 2020. Skolverkets lägesbedömning 2020. Retrieved from Stockholm. <https://www.skolverket.se/publikationsserier/rapporter/2020/skolverkets-lagesbedomning-2020>.
- Steiner-Khamis, G., 2014. Comparison and context: the interdisciplinary approach to the comparative study of education. *Curr. Issues Comp. Educ.* 16 (2), 34–42.
- Sultana, R.G., Gellel, A.M., Caruana, S., 2019. Teacher education in Malta. In: Karras, K.G., Wolhuter, C.C. (Eds.), *International Handbook of Teacher Education*. H.M. Studies & Publishing, Nicosia, Cyprus, pp. 397–414.
- Sun, A.N., Xia, J.G., 2018. Teacher-perceived distributed leadership, teacher self-efficacy and job satisfaction: a multilevel SEM approach using the 2013 TALIS data. *Int. J. Educ. Res.* 92, 86–97. <https://doi.org/10.1016/j.ijer.2018.09.006>.
- Thomas, L., Tuytens, M., Moolenaar, N., Devos, G., Kelchtermans, G., Vanderlinde, R., 2019. Teachers' first year in the profession: the power of high-quality support. *Teach. Teach.* 25 (2), 160–188. <https://doi.org/10.1080/13540602.2018.1562440>.
- Thomas, L., Rienties, B., Tuytens, M., Devos, G., Kelchtermans, G., Vanderlinde, R., 2021. Unpacking the dynamics of collegial networks in relation to beginning teachers' job attitudes. *Res. Pap. Educ.* 36 (5), 611–636. <https://doi.org/10.1080/02671522.2020.1736614>.
- Tinn, M., Ümarik, M., 2021. Looking through teachers' eyes - investigating teacher agency. *Br. J. Educ. Stud.* 1–17.
- Toropova, A., Myrberg, E., Johansson, S., 2021. Teacher job satisfaction: the importance of school working conditions and teacher characteristics. *Educ. Rev.* 1–27. <https://doi.org/10.1080/00131911.2019.1705247>.
- Tschannen-Moran, M., Hoy, A.W., 2001. Teacher efficacy: capturing an elusive construct. *Teach. Teach. Educ.* 17 (7), 783–805.
- Viac, C., Fraser, P., 2020. Teachers' Well-Being: A Framework for Data Collection and Analysis. OECD Education Working Papers No. 213. Retrieved from. https://www.oecd-ilibrary.org/education/teachers-well-being_c36fc9d3-en.
- Wang, K., Chen, Z., Luo, W., Li, Y., Waxman, H., 2018. Examining the differences between the job satisfaction of STEM and non-STEM Novice teachers with leaving intentions. *Eurasia J. Math. Sci. Technol. Educ.* 14. <https://doi.org/10.29333/ejms/89516>.
- Wang, K., Li, Y., Luo, W., Zhang, S., 2020. Selected factors contributing to teacher job satisfaction: a quantitative investigation using 2013 TALIS data. *Leader. Pol. Sch.* 19 (3), 512–532. <https://doi.org/10.1080/15700763.2019.1586963>.

- Wermke, W., Olason Rick, S., Salokangas, M., 2019. Decision-making and control: perceived autonomy of teachers in Germany and Sweden. *J. Curric. Stud.* 51 (3), 306–325. <https://doi.org/10.1080/00220272.2018.1482960>.
- Xanthopoulou, D., Bakker, A.B., Demerouti, E., Schaufeli, W.B., 2007. The role of personal resources in the job demands-resources model. *Int. J. Stress Manag.* 14 (2), 121–141. <https://doi.org/10.1037/1072-5245.14.2.121>.
- Zakariya, Y.F., Bjørkestøl, K., Nilsen, H.K., 2020. Teacher job satisfaction across 38 countries and economies: an alignment optimization approach to a cross-cultural mean comparison. *Int. J. Educ. Res.* 101, 101573. <https://doi.org/10.1016/j.ijer.2020.101573>.

Teacher inquiry: by any other name

Anthony Clarke, The University of British Columbia, Department of Curriculum and Pedagogy, Vancouver, British Columbia, Canada

© 2023 Elsevier Ltd. All rights reserved.

Teacher inquiry	232
Make explicit	232
Make sense	233
Make public	234
A color fan of possibilities	235
Action research (AR) ... is distinguished by its emphasis on “teacher action”	236
Quality teaching rounds (QTR) ... is distinguished by its emphasis on “rounds of teaching”	236
Lesson study ... is distinguished by its emphasis on “lesson design”	236
Learning study ... is distinguished by its emphasis on the “object of learning”	236
Scholarship of teaching and learning (SoTL) ... is distinguished by its shift from scholarly approaches to teaching to the “scholarship” of teaching	237
Appreciative inquiry (AI) ... is distinguished by its initial emphasis on “what’s working”	237
Self-study ... is distinguished by its emphasis on “the self in teaching”	237
Insider research ... is distinguished by its emphasis on “positionality”	238
Narrative inquiry ... is distinguished by its emphasis on “story”	238
Pedagogical narration ... is distinguished by its use of “multimedia” to tell a story	238
Autobiography ... is distinguished by its emphasis on “pattern”	239
Autoethnography ... is distinguished by its emphasis on “the self in context”	239
Life history ... is distinguished by its emphasis on “temporality”	239
Memoir ... is distinguished by its emphasis on “memory”	240
Portraiture ... is distinguished by its emphasis on “esthetics”	240
Professional practice: what’s inquiry got to do with it?	241
References	241

Teacher inquiry

Teacher inquiry is a form of professional development that enables, enriches, and extends a teacher’s understanding of their professional practice. As such, teacher inquiry is knowledge generation about one’s practice. There is general consensus that there are three elements to teacher inquiry. The first is the ability to make explicit the particulars of one’s practice relative to the inquiry being undertaken. Second is the ability to make sense of those elements in light of past experience and current knowledge toward new understandings about one’s practice. Finally, it is important to make public those understandings so that they can be part of a larger conversation about professional practice that prompts further refinement and reflection on practice. These elements are presented as distinct from one another in this article but in reality they overlap and unfold in unexpected ways with each interacting and informing the other. Although difficult to visualize, one depiction of teacher inquiry that comes close is that of a Mobius strip (Fig. 1).

It is challenging for any abstraction to fully capture the complexity of practice; for example, not all teacher inquiries unfold in a linear fashion. Also, while there might be a nominal start and end point to a particular inquiry, in truth the idea that sparks one inquiry likely has traces that reach back in time and typically continues to evolve into other questions or inquiries well after the initial inquiry concludes.

Make explicit

The starting point for most teacher inquiry studies is an attempt to make explicit one or more aspects of practice that surprise, intrigue or challenge us in our daily work as teachers. These aspects are often embedded in stories or narratives of practice. As teachers, we have many stories that accumulate over time. Some stories stick in our mind or pique our curiosity more so than others. Often we can’t fully make sense of them in the immediate retelling of the story after the event. However, there is often a nugget within the story that prompts further thinking. In this sense, making explicit our stories of practice is an important first step in attempting to better understand our practice. It is only when we articulate and make explicit aspects of our practice that our stories are available for closer scrutiny. One of the stories from my practice is about Lara, a Year 10 student at East Doncaster High School in Australia.

One Monday during the morning recess break Lara ran up to me in the courtyard and asked excitedly: “Mr. Clarke, Mr. Clarke, did you hear about the contest? Can we enter it?” I had taught Lara a few years previously but hadn’t had a lot of contact with her



Fig. 1 Three elements of teacher inquiry: make sense, make explicit, and make public. (Thanks to Kieran Forde for suggesting this image).

since then. So her sudden approach and exclamation took me by surprise. The competition she referred to was sponsored by a local radio station and the school with the greatest number of empty 'Big M' milk containers would win a live concert at their school by Australia's most popular rock band at that time, the Angels (www.theangels.com.au). After several discussions with the school administration and my teaching colleagues my response to Lara was "Yes."

The above paragraph is an excerpt from my attempt to use teacher inquiry to better understand and make sense of this story in light of my teaching practice (Clarke, 2014). Prior to framing this as a teacher inquiry, I had shared this story of the Angels performing at our school many times with colleagues and friends but I always told it with an emphasis on the Angels. However, after repeated retelling I became increasingly curious about Lara's role in the story. As a result, when I deliberately began to make explicit the key details about this story, its essence shifted and I found myself asking: What if I had said "No" to Lara? The challenge at this point was to make sense of this question in terms of my practice as a teacher at East Doncaster High School.

Make sense

Making sense of one's practice means analyzing or interrogating more fully that which has been made explicit. This move—from identification to interpretation—represents a second important element of teacher inquiry. To appreciate the significance of this shift, it is helpful to understand the emergence of teacher inquiry within the broader context of educational research over time and, in particular, in relation to four distinct trends: (1) research on pupils, (2) research on teachers, (3) research with teachers, and (4) research by teachers (Fig. 2).

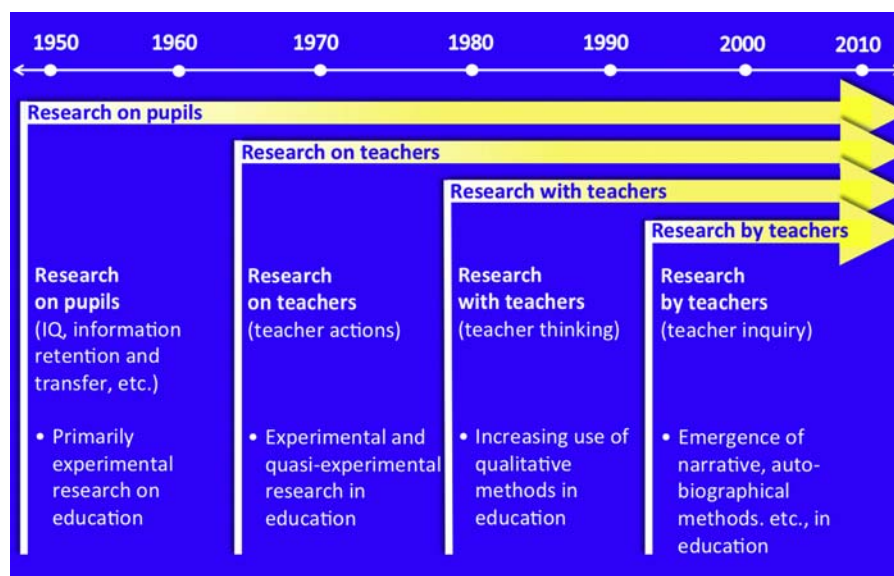


Fig. 2 Four trends in educational research over time.

The first trend is associated with a natural sciences approach to educational research that was prevalent in the 1950s. Teacher education research at this time was imbued with a strong cognitive psychology approach to student learning (e.g., a focus on I.Q. tests, knowledge retention, knowledge transfer). This research, referred to as experimental research, was characterized by attempts to isolate elements of learning into discrete units so that they could be the focus of intensive statistical analysis. While it had an impact on education (e.g., policy development), it is generally regarded as having had a relatively minor influence on teaching and learning in the classroom. A second trend emerged in the late 1960s and coincided with a dramatic shift in public interest in education. Events such as Sputnik suddenly brought a renewed focus on schools and schooling, particularly in North America. As a result, educational researchers were forced to rethink their approach to research from focussing almost exclusively on student learning to include teacher actions (Erickson, 1986). Still, many of these experimental and quasi-experimental efforts were based upon linear causal models that implied that professional practice could be regarded as the field of theoretical application (Connelly and Clandinin, 1987), and further, that the knowledge, skills, and competencies required by teachers could be precisely specified in advance (Zeichner, 1987). Much of this process-product, teacher effectiveness, and teacher competency research was based upon a technical rational approach to teaching and learning (Shulman, 1981; Boydell, 1986).

The next trend appeared in the 1980s. As dissatisfaction with a technical rational approach to education became more widespread, alternative research methods emerged (Houston et al., 1989). It was during these years that Zeichner and others sought approval by the American Educational Research Association (the largest educational research association in North America) for a Research Division specifically devoted to teacher education (Zeichner, 1998). The AERA governing body rejected their proposal on the basis that there was insufficient scholarship in the field of teacher education to warrant such a division. However, the governing body felt that if the proposal was to include the “study of teaching” as well as the “study of teacher education” then there would be sufficient grounds for establishing a new division. As a result, a division called “Teaching and Teacher Education” (Division K) was created within AERA in 1984. Division K, with its broader mandate, prompted a shift from research on teachers (which, among other things, was largely objective analyses of teaching actions) to research with teachers (which acknowledged the legitimacy of more subjective analyses of teaching practice) with a corresponding increase in the use of more qualitative research methods.

The early 1990s saw a further evolution of educational research to include inquiries “by teachers.” Evidence of this fourth trend can be found in special theme issues in professional journals at that time, for example, volume 22, issue 3, of *Teacher Education Quarterly*, and the development of special interest groups (SIG) such as the Self-Study of Teacher Education Practices (S-STEP) at AERA. More significantly, an even broader range of research methods was now deemed legitimate in education (e.g., for the first time the use of case study research for a Ph.D. degree in education was allowed in many North American universities). The significance of the fourth trend was that it opened up a color fan of possibilities for teacher inquiry in the new millennium (Tidwell et al., 2009) (Fig. 3).

It is worth noting that although this trend was only formally recognized within academia in the 1990s, it confirmed the earlier writings of Stenhouse (1975), one of the original advocates of teacher inquiry. This long overdue recognition of Stenhouse’s pioneering work confirms, even to this day, the glacial pace of innovation within the academy (Diamond, 2006).

The color fan of possibilities depicted in Fig. 3 lists fifteen forms of teacher inquiry. This is not an exhaustive or definitive list. However, each possibility provides a springboard for making sense of practice. Further, each form provides guidance and direction for advancing various types of teacher inquiry. Some forms require the collection of additional information or data. Others draw on existing sources (e.g., journals, artifacts, memories). The choice of form most appropriate for a particular inquiry is not always obvious at the outset and it may take some time before finding the most suitable form for the questions one asks and the claims one wishes to make. In my case, it took some time for Lara’s story to land on narrative inquiry as the most appropriate form for my teacher inquiry.

Narrative inquiry demands that I set the stage and paint the backdrop to portray as vividly as possible the essence of the key elements of the story. Narrative inquiry uses writing in the form of storytelling as a way to refine and reflect upon practice. When writing Lara’s story, I had to go beyond my own local knowledge and experience and engage with the literature on teaching and learning as the story unfolded. As a result I was better able to situate my inquiry and draw upon a frame for making sense of that inquiry. The outcome of my deliberations was an acknowledgment that I was teaching students, not subjects (Moje, 1996; Palmer, 1997; Sparks and Pole, 2019); a maxim that is old as teaching itself but it took my analysis of Lara’s story for this to have personal relevance for me. This realization was facilitated by my reading of Fenstermacher’s (1992) work which makes the distinction between “a system of schooling” and “an educative agenda.” This distinction highlights the difference between being a teacher and being an educator. Further the evolution of my inquiry was very much a result of sharing my developing understandings with others which brings us to the third element of teacher inquiry.

Make public

For our investigations to be considered teacher inquiry, the outcomes need to be shared with knowledgeable others; that is, we deliberately move the emerging results of our inquiries from the private to the public domain. As noted at the outset, teacher inquiry is about knowledge generation. LaBosekey (2004) argues that an important part of any knowledge generation is that “we formalize our work and make it available to our professional community” (p. 860). Further, as Davis et al. (2003) remind us, the intelligence of the group is always greater than the intelligence of the individual.



Fig. 3 A color fan of possibilities for teacher inquiry.

The use of the phrase “knowledgeable others” is deliberate here. Knowledgeable others might include our colleague in the next classroom. It might include members of a study group within our school. It might include teachers attending a district workshop. It could also include readers of a teacher magazine or professional journal where our work is published. All these are ways of making public our teacher inquiries. As such, they provide opportunities for those with a knowledge of teaching to react and respond to our work which, in turn, presents opportunities to extend and further our own and also the profession’s understanding of practice.

By sharing Lara’s story with knowledgeable others, I was moving my understandings of teaching and learning from the private to the public domain. In doing so, I was opening myself (and my practice) up to scrutiny. Making my story public by sharing it with knowledgeable others and then listening to their feedback was absolutely critical to the new understandings that I gained about my practice. It helped me to think more deeply about *why I do what I do* as a member of the teaching profession.

I first shared fragments of Lara’s story with a small study group who were exploring the “inner work” of teachers and its role in successful teaching practice (Cohen and Bai, 2012). The group asked a number of questions about the story and encouraged me to capture it more fully in written form. After much to-ing and fro-ing, I decided on narrative inquiry as the most appropriate form for doing this. When my writing was sufficiently coherent, I shared Lara’s story with other teachers and colleagues at a local workshop and later at a provincial conference. The feedback I received prompted further analysis and refinement of the story (Clarke, 2014).

As noted above, narrative inquiry is one of a range possibilities for teacher inquiry. In the section that follows, a summary of fifteen other forms of teacher inquiry is provided as a stepping stone for thinking further about teacher inquiry as a form of professional development in practice settings.

A color fan of possibilities

The color fan of possibilities depicted in Fig. 3 lists various forms of teacher inquiry that have been defined and used in educational research. As such, they provide substantive examples of how teacher inquiry has been taken up by practitioners. There are undoubtedly other forms of teacher inquiry that are being explored at the present time and will at some point become part of the educational research lexicon. A quick perusal of the fifteen forms suggests some overlap between forms and, in some cases, only very slight or subtle shades of difference. For example, narrative inquiry and pedagogical narration share storytelling as a central component of the inquiry process. Sometimes teacher inquirers use variations of 15 forms or combine them with other social science practices to create a distinct approach suitable for their context. For example, Participatory Action Research (MacDonald, 2012) is a combination of action research and a “Communities of Practice” (CoP) model for professional development (Lave and Wenger, 1991). Some forms are difficult to neatly categorize because of their broad ambit (e.g., Scholarship of Teaching and Learning or SoTL).

Nonetheless, each form provides teachers with a way of advancing their understandings of practice. The descriptions provided here are meant as starting points for exploring possible forms of teacher inquiry that can be taken up in educational settings.

Action research (AR) ... is distinguished by its emphasis on “teacher action”

Action research, as the name implies, places a strong emphasis on “action.” In particular, the anticipated outcome of action research is a change in one’s practice. Action research is a cyclical process with four distinct phases per cycle: plan, act, observe, and reflect. The outcomes from one cycle are used as the basis for continued exploration of practice in the following cycle. Thus, each cycle potentially leads to new understandings of practice. One of the greatest appeals of action research is tangible improvement in practice with direct benefits for both the teacher and the students. One of the earliest attempts to articulate action research as a form of inquiry was by [Lewin \(1946\)](#). Other useful references for AR include:

Carr, W., Kemmis, S., 1986. *Becoming Critical: Education, Knowledge and Action Research*. London, Falmer. doi: 10.4324/9780203496626.

McNiff, J., 2013. *Action Research: Principles and Practice*, third ed. Abingdon, Routledge. doi: 10.4324/9780203112755.

Quality teaching rounds (QTR) ... is distinguished by its emphasis on “rounds of teaching”

Quality Teaching Rounds (QTR) is a recent Australian initiative and occurs within the context of Professional Learning Communities (PLCs). QTR is based on series of instructional rounds focused on a particular aspect of pedagogy as decided by the PLC. QTR has echoes of hospital teaching rounds used in medicine where the immediacy of the action setting is the crucible for learning. An instructional round is held over the course of a single day and involves reading a text or related article, an observation of a lesson taught by one of the teachers from the group, followed by “a rigorous diagnostic professional conversation” ([Gore et al., 2017](#), p. 100) by the group. Each member of the group teaches at least one lesson as part of the QTR process and the members stay together for an entire set of rounds. A key characteristic of QTR is the intense juxtapositioning of literature, observation, and conversation. Useful references for QTR include:

Bowe, J., Gore, J., 2016. Reassembling teacher professional development: the case for quality teaching rounds. *Teach. Teach.* 23(3), 352–366. doi: 10.1080/13540602.2016.1206522.

Bowen, L., 2020. Time to talk about teaching: the power of quality teaching rounds. *J. Prof. Learn. Semester 1*, 1–7. Retrieved from: <https://cpl.asn.au/journal/semester-1-2020/time-to-talk-about-teaching-the-power-of-quality-teaching-rounds>.

Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., Lubans, D., 2017. Effects of professional development on the quality of teaching: results from a randomized controlled trial of quality teaching rounds. *Teach. Teach. Educ.* 68, 99–113. doi: 10.1016/j.tate.2017.08.007.

Lesson study ... is distinguished by its emphasis on “lesson design”

Lesson Study (“jugyuu kenkyuu”) originated in Japan in the 1990s. It is teacher-led inquiry where a small group of teachers collectively identify an aspect of a lesson that they are seeking to better understand or improve through the development of a more appropriate lesson design. They plan, teach, observe, and reflect upon the same “research lesson” ([Fernandez, 2002](#), 394) as taught by various (but not necessarily all) members of the group. Thus, lesson design and re-teaching of the same lesson are central to lesson study. Useful references for lesson study include:

Elliott, J., 2019. What is lesson study? *Eur. J. Educ.* 54(2), 175–188. doi: 10.1111/ejed.12339.

Fernandez, C., 2002. Learning from Japanese approaches to professional development: the case of lesson study. *J. Teach. Educ.* 53(5), 393–405. doi: 10.1177/002248702237394.

Lewis, C., Lee, C., 2017. The global spread of lesson study: contextualization and adaptations. In: Akiba, M., G. K. Letendre, G. K. (Eds.), *International Handbook of Teacher Quality and Policy*, New York, Routledge, pp. 185–203. doi: 10.4324/9781315710068-13.

Learning study ... is distinguished by its emphasis on the “object of learning”

Learning study originated in Hong Kong and is similar to lesson study in many ways. However, a distinguishing feature is that it always takes the “object of learning” as its starting point (i.e., what is to be learned). Further, learning study is always grounded in a theory of learning. To date, proponents of learning study have typically drawn upon variation theory ([Martin, 2015](#)), which has its basis in phenomenography ([Martin 1986](#)). Proponents of variation theory contend that learning results from the experience of difference, and that difference is what allows learners to discern the critical features of a particular phenomenon. [Runesson \(2015\)](#) argues that “variation theory synthesized with other theories might contribute to a further development of learning study” (e.g., the use of cultural historical activity theory or CHAT).

Useful references for learning study include:

- Lo, M.L., 2009. The development of the learning study approach in classroom research in Hong Kong. *Educ. Res. J.* 24(1), 165–184.
- Wood, K., 2017. Is there really any difference between lesson and learning study? Both focus on neriage. *Int. J. Lesson Learn. Stud.* 6(2), 118–123. doi: 10.1108/IJLLS-02-2017-0008.
- Xu, L., 2019. Relating teaching and learning of science through the lens of variation theory. *Scand. J. Educ. Res.* 63(1), 145–162. doi: 10.1080/00313831.2017.1357143.

Scholarship of teaching and learning (SoTL) ... is distinguished by its shift from scholarly approaches to teaching to the “scholarship” of teaching

SoTL is a form of inquiry into practice undertaken by college and university professors in Higher Education. SoTL emphasizes the difference between scholarly approaches to teaching and learning and the scholarship of teaching and learning (Richlin, 2001). SoTL researchers view classrooms and related university learning spaces (such as labs, tutorials, field experiences) as sites for inquiry into practice. SoTL also encompasses other forms of inquiry into practice such as the Scholarship of Curriculum Practice (SoCP) and the Scholarship of Educational Leadership (SoEL). SoTL practitioners often begin with investigation methods consistent with the signature pedagogies in their academic disciplines and then expand beyond those methods over time and as appropriate. Not all SoTL inquiries fall strictly within the definition of teacher inquiry, however the majority do. Useful SoTL references include:

- Boyer, E.L., 1991. The scholarship of teaching from: scholarship reconsidered: priorities of the professoriate. *Coll. Teach.* 39, 11–13. doi: 10.1080/87567555.1991.10532213.
- Huber, M.T., Hutchings, P., 2005. *The Advancement of Learning: Building the Teaching Commons*. San Francisco, CA, Jossey-Bass.
- Shulman, L., 2011. Feature essays: the scholarship of teaching and learning: a personal account and reflection. *Int. J. Sch. Teach. Learn.* 5(1), Article 30. doi: 10.20429/ijstl.2011.050130.

Appreciative inquiry (AI) ... is distinguished by its initial emphasis on “what’s working”

In stark contrast to many approaches to teacher inquiry that start by identifying an issue, problem or challenge, AI begins by articulating “what’s working” (Kung et al., 2013, p. 3) in one’s practice. Practitioners then use that as the starting point and platform for their inquiries into practice. Because of this very distinct approach to teacher inquiry, AI is referred to as a strengths-based, positive approach to change. AI is often associated with system-level change within organizations but the principles of AI are equally applicable to individual classroom contexts. AI typically has four stages that practitioners move through during their inquiries: discovery, dream, design, and destiny. The beginnings of appreciative inquiry can be traced back to the early work of Cooperrider and Srivastva (1987). Useful references for AI include:

- Giles, D., Kung, S., 2010. Using appreciative inquiry to explore the professional practice of a lecturer in higher education: moving toward life-centric practice. *Aust. J. Adult Learn.* 50(2), 308–322.
- Kaplan, C.S., 2014. An appreciative inquiry approach to reflecting on teaching. *Lang. Educ.* 9(3), 44–47.

Self-study ... is distinguished by its emphasis on “the self in teaching”

Self-study as a form of educational research came to prominence in the 1990s when a group of teacher educators in colleges and universities began to deliberately explore, critique, and publish inquiries about their teaching practices (Loughran, 2004). In 1994, the American Educational Research Association (AERA)—arguably the largest association of educational researchers in North America—permitted these professors to establish a Special Interest Group (SIG) within AERA. The SIG, called as Self-Study of Teacher Education Practices (S-STEP), has since developed forums, journals, and international handbooks to support this form of teacher inquiry. Today, the ripple effect of S-STEP is felt across all levels of education as self-study has moved from the shadows of academia into the full light of legitimate educational research.

In the beginning “the self” for S-STEP practitioners was understood to be the individual. However, debates quickly arose within S-STEP about what constituted the unit of analysis for “the self.” For example, could a science department using the principles of S-STEP conduct a self-study where “the self” was the department? Subsequent S-STEP studies suggest this is possible. Finally, proponents of self-study have steadfastly held to the term self-study to describe their inquiries while drawing upon a multitude of approaches and strategies for those inquiries. Useful references for self-study include:

- Loughran, J.J., 2004. A history and context of self-study of teaching and teacher education practices. In: Loughran, J.J., Russell, T.L., (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, Springer, Dordrecht, pp. 7–39. doi: 10.1007/978-1-4020-6545-3_1.
- Austin, T., Senese, J., 2004. Self-study in school teaching: teachers’ perspectives. In: Loughran, J.J., Russell, T.L., (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, Springer, Dordrecht, pp. 1231–1258. doi: 10.1007/978-1-4020-6545-3_32.
- Samaras, A., Roberts, L., 2011. Flying solo: teachers take charge of their learning through self-study research. *Learn. Prof.* 32(5), 42–45. Retrieved from: <https://eric.ed.gov/?id=EJ965106>.

Insider research ... is distinguished by its emphasis on “positionality”

Insider research is the analysis of practice from the unique perspective of someone from “inside” the practice. As an insider, the practitioner possesses intimate knowledge of the practice under study. Thus positionality is an important aspect of insider research:

Positionality refers to the stance or positioning of the researcher in relation to the social and political context of the study—the community, the organization or the participant group. The position adopted by a researcher affects every phase of the research process, from the way the question or problem is initially constructed, designed and conducted to how others are invited to participate, the ways in which knowledge is constructed and acted on and, finally, the ways in which outcomes are disseminated and published.

Rowe (2014, p. 628).

Further, insider knowledge may differ from one insider to another insider within the same practice. Thus, the imperative for those engaged in insider research is that “you have to position yourself somewhere in order to say anything at all” (Hall, 1990, p. 18). Cochran-Smith and Lytle (1993) were among the first to write about insider research as a form of teacher inquiry. Other useful references include:

- Finefter-Rosenbluh, I., 2017. Incorporating perspective taking in reflexivity: a method to enhance insider qualitative research processes. *Int. J. Qual. Methods*, 16(1), 1–11. doi: 10.1177/1609406917703539.
- Greene, M.J., 2014. On the inside looking in: methodological insights and challenges in conducting qualitative insider research. *Qual. Rep.* 19(29), 1–13. Retrieved from: <http://www.nova.edu/ssss/QR/QR19/greene15.pdf>.
- Schaenen, I., Kohnen, A., Flinn, P., Saul, W., Zeni, J., 2012. “I” is for “Insider”: practitioner research in schools. *Int. J. Act. Res.* 8(1), 68–101. doi: 10.1688/1861-9916_IJAR_2012_01_Schaenen.

Narrative inquiry ... is distinguished by its emphasis on “story”

Narrative inquiry is the study of experience interpreted by and through stories of practice. For Connelly and Clandinin (1990), “Humans are storytelling organisms who, individually and collectively, lead storied lives” (p. 2). Stories embody knowledge accumulated and experience gained over time. Narrative inquiry is one way of making sense of that knowledge and experience. In the process of telling a story, the teacher inquirer sets the stage and paints the landscape as vividly as possible to portray the essence of the event(s) that are the focus of the inquiry. The inquirer emphasizes some elements and relegates others to background color. Each of these decisions requires one to pause, think, and then act—that is, to make judgments about what is important and relevant to the story as it unfolds. Useful references for narrative inquiry include:

- Bell, J.S., 2002. Narrative inquiry: more than just telling stories. *TESOL Quart.* 36(2), 207–213. doi: 10.2307/3588331.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19(5), 2–14. doi: 10.3102/0013189X019005002.
- Huber, J., Caine, V., Huber, M., Steeves, P., 2013. Narrative inquiry as pedagogy in education: the extraordinary potential of living, telling, retelling, and reliving stories of experience. *Rev. Res. Educ.* 37(1), 212–242. doi: 10.3102/0091732X12458885.

Pedagogical narration ... is distinguished by its use of “multimedia” to tell a story

Pedagogical Narration is the systematic documentation of teaching practice through multimedia, including text, images, and artifacts. Using multimedia to make learning visible emerged in the late 1970s within the context of early childhood settings, in particular, Reggio Emilia (Giudici et al., 2001). Pedagogical narration allows for the assemblage of one’s practice in ways that are not always possible in more traditional forms of documentation. This assemblage provides alternative ways of seeing and noticing that practice, leading to new insights and understandings. Pedagogical narration is not a bulletin board of disconnected artifacts. Rather it is a deliberate and intentional construction in terms of inviting curiosity, wonder, and engagement. As Browne et al. (2010) argue, “when we document children’s learning, we consider and reflect on what the children may be thinking, doing, and/or learning and then apply these insights to the unique manner of teaching them” (p. 49). Thus, pedagogical narration has the potential to make both children’s and teachers’ learning visible. Useful references for pedagogical narration include:

- Edwards, C., Gandini, L., Forman, G. (Eds.), 2012. *The Hundred Languages of Children: The Reggio Emilia Experience in Transformation*. third ed. Oxford, Praeger, Cop. Retrieved from <https://eric.ed.gov/?id=ED531297>.
- Turner, T., Wilson, D., 2010. Reflections on documentation: a discussion with thought leaders from Reggio Emilia. *Theor. Pract.* 49(1), 5–13. Retrieved from: <https://eric.ed.gov/?id=EJ879851>.
- Wien, C.A., 2013. *Making Learning Visible Through Pedagogical Documentation*, 1–4. Toronto, Ontario, Queen’s Printer for Ontario. Retrieved from: <https://cste-educ.sites.olt.ubc.ca/files/2019/07/Berger-2015-Pedagogical-Narration.pdf>.

Autobiography ... is distinguished by its emphasis on “pattern”

Autobiography provides the opportunity to understand “the self” over time by exploring the patterns that define one’s life. It is argued that an understanding of oneself is critical in any attempt to understand others and the world about us (Böckler et al., 2017). Autobiography is a highly personal endeavor: “it is ... simultaneously historical record and literary artifact, psychological case history and spiritual confession, didactic essay and ideological testament,” (Stone, 1981, p. 80). Thus, what one includes and why it is included are all “in play” in an autobiographical account. Sometimes the veracity of autobiographical accounts cannot be confirmed externally because they are stories written by oneself about oneself:

Autobiography is not the story of a life; it is the recreation or the discovery of one. In writing of experience, we discover what it was, and in the writing create the pattern we seem to have lived. Simply put, autobiography is a reckoning.

Heilbrun (1995, p. xvii).

Autobiographies are not always linear in their construction and it is the patterns that emerge that say more about one’s teaching life than the fullness of the account itself. Useful references for autobiography include:

- Davis, N.T., 1996. Looking in the mirror: teachers’ use of autobiography and action research to improve practice. *Res. Sci. Educ.* 26(1), 23–32. doi: 10.1007/bf02356961.
- Graham, R.J., 1989. Autobiography and education. *J. Educ. Thought*, 23(2), 92–105. Retrieved from: <https://www.jstor.org/stable/23768385>.
- Ramsey, S.J., 2004. Autobiography of a teacher: a journey toward critical multiculturalism. *Sch. Pract. Quart.* 2(3), 45–52. Retrieved from: <https://eric.ed.gov/?id=EJ796119>.

Autoethnography ... is distinguished by its emphasis on “the self in context”

Mishler’s (1979) rhetorical question: “Meaning in context: is there any other kind?” speaks to the importance of context in making sense of who we are and what we do as educators. Thus, for autoethnographers, the self is part of a larger picture. Denshire (2014) argues that autoethnographic accounts “can function as something of a corrective” (p. 833) to highly personalized and therefore potentially decontextualized accounts of practice. In this sense, autoethnographic writing is a relational endeavor seeking to understand one’s practice in light of the political, social, and cultural contexts in which it is embedded. Useful references for autoethnography include:

- Ellis, C., Adams, T., Bochner, A., 2011. Autoethnography: an overview. *Forum*, 12(1), Article 10, pp. 1–13. Retrieved: <http://nbn-resolving.de/urn:nbn:de:0114-fqs1101108>.
- Starr, L.J., 2010. The use of autoethnography in educational research: locating who we are in what we do. *Can. J. New Sch. Educ.* 3(1), 1–9. Retrieved from: <https://journalhosting.ualgary.ca/index.php/cjnse/article/view/30477>.
- Woodley, H., 2016. The use of autoethnography in classroom based practitioner research. *Ann. Rev. Educ. Commun. Lang. Sci.* 13, 44–56. Retrieved from: <https://research.ncl.ac.uk/media/sites/researchwebsites/arecls/Helen%20Woodley.pdf>.

Life history ... is distinguished by its emphasis on “temporality”

Those who take up life history as a research method have a particular interest in understanding the connection between people, places and events over time. Thus, temporality is central to life history. Bullough and Gitlin (1995) argue that “To know the past is to know oneself as an individual *and* as a representative of a socio-historical moment in time; like others each person is a victim, vehicle, and ultimately a resolution of a culture’s dilemmas.” (p. 25). As with any historical account there is the expectation that life history is based and draws upon on reliable and verifiable sources. For some researchers, such as Watson and Watson-Franke (1985), life history is “any retrospective account by the individual of his life in whole or in part, in written or oral form, that has been elicited or prompted by another person” (p. 2). Thus, traditionally, a life history is written by somebody other than the person whose history is being written. However, it is on this point that Fitzpatrick’s (2011) question becomes relevant here: “Can You Write a History of Yourself?” Although Fitzpatrick found this difficult to do (she made the shift to autobiography), Samaras et al. (2004) argue that:

Personal history self-study is increasingly becoming an essential methodology towards teacher educators’ personal and professional growth and especially to improving their teaching practice and impacting their students’ learning. Through a personal history self-study approach, professors and their students are able to reconstruct significant life events to inform them of their professional identity formation and to help them make meaning of their pedagogy and the connections of their practice to theory. (p. 905)

Useful references for life history include:

- Chambers, C., Hasebe-Ludt, E., Leggo, C., Sinner, A., 2012. *A Heart of Wisdom: Life Writing as Empathetic Inquiry*, New York, Peter Lang.
- Goodson, I., 1980. Life histories and the study of schooling. *Interchange*, 11(4), 62–76. Retrieved from <https://eric.ed.gov/?id=EJ244869>.
- Gouthro, P., 2014. Stories of learning across the lifespan: life history and biographical research in adult education. *J. Adult Contin. Educ.* 20(1), 87–103. doi: 10.7227/jace.20.1.6.
- Knowles J.G., 1993. Life-history accounts as mirrors: a practical avenue for the conceptualization of reflection in teacher education. In: Calderhead, J., Gates, P. (Eds.), *Conceptualizing Reflection in Teacher Development*, London, The Falmer Press, pp. 70–92. doi: 10.4324/9780203209851_chapter_5.

Memoir ... is distinguished by its emphasis on “memory”

Memoir draws on a variety of elements in making sense of one’s practice. Perhaps its most defining characteristic is the use of memory. Keightley (2010) notes: “The turn to memory in social research enables the role of the past in the present to be explored. ... Not simply concerned with the past, memory studies are concerned with the ways memories respond to the demands of current experience and future desires ...” (p. 67). One reason for this response is that, as humans, we try to overcome cognitive dissonance and therefore attempt to align our emotions, behaviors, and practices in terms of what we remember (Ghalipour, 2017). Therefore, memoirs are more reflective than reportorial; they reveal as much about feelings as they do about facts (Radden and Varga, 2013). Useful references for memoir include:

- Boylorn, R., 2008. Memoirs. In: Given, L. (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods*, Thousand Oaks, SAGE Publications, Inc., pp. 505–506.
- Keightley, E., 2010. Remembering research: memory and methodology in the social sciences. *Int. J. Soc. Res. Methodol.* 13(1), 55e70. Retrieved from: <https://doi.org/10.1080/13645570802605440>.
- Maniam, V., 2014. The memoir method in educational research from an Australian perspective. *Qual. Sociol. Rev.* 10(3), 49–60. Retrieved from: http://www.qualitativesociologyreview.org/ENG/archive_eng.php.

Portraiture ... is distinguished by its emphasis on “esthetics”

Lawrence-Lightfoot and Davis (1997) argue that portraiture “blurs the boundaries of esthetics and empiricism in an effort to capture the complexity, dynamics, and subtlety of human experience” (p. xv). The portraiture combines scientific aspects of inquiry through systematic use of observation, interviews, and other forms of data collection to describe phenomena, while simultaneously capturing the beauty and esthetic properties of phenomena (Quigley et al., 2015, p. 21). This form of teacher inquiry is perhaps the most expressive, and possibly artistic, of all the forms discussed thus far. Carl Leggo (2014), a renowned educator and poet, exemplifies this sort of writing in a recent publication that is replete with photographs and poetry in combination with regular text:

I grew up in a working-class home where few photographs were taken. It was too expensive to pay for film and developing. So, the entire collection of photos from my childhood is approximately fifty images. This is the whole of a working class family archive. Therefore, the school photos of elementary school are especially significant because they are a part of a small collection. I don’t have photos of high school classes. Most of the photos are black and white, and most are blurry or faded. Nevertheless, when I look at the photos, I am struck by the still seriousness in so many of the images. Like Barthes (2010), I see in these images something irreducible in my childhood: “everything which is still in me, by fits and starts” is “in the child” (p. 22). Above all, I see fear in the face of the child in so many of the photos. (p. 63)

Useful references for portraiture include:

- Dixon, A.D., Chapman, T.K., Hill, D.A., 2005. Research as an esthetic process: extending the portraiture methodology. *Qual. Inq.* 11(1), 16–26. doi: 10.1177/107780040427083.
- Hancheruk, H., McBride, J., Witczak, K., 2016. Re-locating the self: portraiture for teacher professional growth. *Learn. Landsc.* 9(2), 285–300. doi: 10.36510/learnland.v9i2.776.
- Lawrence-Lightfoot, S., Davis, J.H., 1997. *The Art and Science of Portraiture*. San Francisco, Jossey-Bass. Retrieved from: <https://www.amazon.com/Art-Science-Portraiture-Sara-Lawrence-Lightfoot/dp/0787910643>.

All of the above forms of teacher inquiry require practitioners to make explicit, make sense, and then make public their explorations of practice. These explorations can include personal and professional dimensions in various measures. The outcomes contribute to one’s understandings of practice and encourage all of us to think more deeply about *why we do what we do*. Further, the outcomes advance understandings of the profession where both teachers and students are the beneficiaries.

Professional practice: what's inquiry got to do with it?

For someone to be a member of a profession it is incumbent upon them to meet a number of expectations. For example, a professional needs to meet the qualifications stipulated for entry into a particular profession, abide by the ethical principles of that profession, and be accountable to the regulatory body that oversees the profession. Furthermore, and perhaps most importantly, professionals must always be knowledgeable about the latest developments in their field and, in turn, ensure that that knowledge becomes a part of and is embedded in their daily work. In short, they must be curious, inquisitive, and generative in sustaining and maintaining their professional practice. Failure to do so means that their practice becomes perfunctory, routinized, or even duplicative.

Practice becomes perfunctory when we lose interest in *why we do what we do*. Our practice becomes routinized when we simply do tomorrow exactly what we did today regardless of the contexts or circumstances in which we find ourselves; we become automatons. Our practice becomes duplicative when we surrender our agency as professionals and mindlessly copy or mimic what others do. When our practice is perfunctory, routinized, or duplicative it becomes labor or technical work. Laborers follow a predetermined plan of action provided to them by others in order to achieve a particular goal. Technicians have the opportunity to choose a plan of action from among a set of predetermined plans which are provided to them in order to achieve a particular goal. Professionals differ from laborers or technicians in that they have the freedom to plan their own course of action to achieve a particular goal. For example, the Ministry of Education might set out particular curriculum goals but teachers have freedom to develop their own plans of action for implementing those goals in their classrooms. As noted above, exercising that right requires teachers to be curious, inquisitive, and generative in the daily enactment of their professional practice. These dispositions are embedded in professional standards for teachers in most jurisdictions throughout the world. For example, in the United States, the National Board for Professional Teaching Standards (NBPTS) notes that teachers as professionals:

... serve as paradigms of lifelong learning and achievement. ... Moreover, they epitomize the intellectual capacities they foster: the ability to reason carefully, consider multiple perspectives, question received wisdom, adopt an inquiry-based approach, solve problems, and persevere. In all aspects of their action and demeanor, accomplished teachers convey the significance of reflection and learning ...

NBPTS, p. 33.

Similarly, the Australian Institute for Teaching and School Leadership (AITSL), which is responsible for establishing the standards for teaching in Australia, stipulates that it is the responsibility of all teachers to “identify their own learning needs and analyze, evaluate and expand their professional learning both collegially and individually” (AITSL, 2018, p. 5). In sum, teacher inquiry—the desire to make explicit, make sense, and make public—is what provokes, nourishes, and advances deeper understandings and broader conceptualizations of teaching and learning. In sum, inquiry is a defining feature of professional practice.

References

- Australian Institute for Teaching and School Leadership, 2018. Australian Professional Standards for Teachers. AITSL, Melbourne, Australia. Retrieved from: https://www.aitsl.edu.au/docs/default-source/national-policy-framework/australian-professional-standards-for-teachers.pdf?sfvrsn=5800f33c_64.
- Barthes, R., 2010. Roland Barthes by Roland Barthes. Macmillan.
- Böckler, A., Herrmann, L., Trautwein, F.M., Holmes, T., Singer, T., 2017. Know thy selves: learning to understand oneself increases the ability to understand others. *J. Cogn. Enhanc.* 1 (2), 197–209. <https://doi.org/10.1007/s41465-017-0023-6>.
- Bowne, M., Cutler, K., DeBates, D., Gilkerson, D., Stremmel, A., 2010. Pedagogical documentation and collaborative dialogue as tools of inquiry for pre-service teachers in early childhood education: an exploratory narrative. *J. Scholarsh. Teach. Learn.* 10 (2), 48–59. Retrieved from: <https://files.eric.ed.gov/fulltext/EJ890715.pdf>.
- Boydell, D., 1986. Issues in teaching practice supervision research: a review of the literature. *Teach. Teach. Educ.* 2 (2), 115–125. [https://doi.org/10.1016/0742-051X\(86\)90010-7](https://doi.org/10.1016/0742-051X(86)90010-7).
- Bullough, R.V.J., Gitlin, A., 1995. *Becoming a Student of Teaching: Methodologies*. For Exploring Self and School Context. Garland Publishing, Inc, New York. <https://doi.org/10.4324/9781315023243>.
- Clarke, A., 2014. What if I had said “No”. In: Cohen, A., Bai, H., Leggo, C., Porath, M., Meyer, K., Clarke, A. (Eds.), *Speaking of Learning: Recollections, Revelations, and Realizations*. Sense Publishers, Rotterdam, pp. 103–116.
- Cochran-Smith, M., Lytle, S.L. (Eds.), 1993. *Inside/Outside: Teacher Research and Knowledge*. Teachers College Press, New York.
- Cohen, A., Bai, H., 2012. Minding what really matters: relationship as teacher. In: Chambers, C., Hasebe-Ludt, E., Sinner, A., Leggo, C. (Eds.), *A Heart of Wisdom: Life Writing as Empathic Inquiry*. Peter Lang, New York, pp. 257–266.
- Connelly, F.M., Clandinin, D.J., 1987. On narrative method, biography and narrative unities in the study of teaching. *J. Educ. Thought* 21 (3), 130–139. Retrieved from: <https://www.jstor.org/stable/23768646>.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14. <https://doi.org/10.3102/0013189X019005002>.
- Cooperrider, D.L., Srivastva, S., 1987. Appreciative inquiry in organizational life. In: Woodman, R.W., Pasmore, W.A. (Eds.), vol. 1. JAI Press, Stamford, CT, pp. 129–169. Retrieved from: https://www.oio.nl/wp-content/uploads/APPRECIATIVE_INQUIRY_IN_Orgnizational_life.pdf.
- Davis, B., Sumara, D., Simmt, E., 2003. Complexity and collectivity: on the emergence of a few ideas. In: Paper Presented at Complexity Science and Educational Research Conference, Edmonton, Canada. Retrieved from: http://www.complexityandeducation.ualberta.ca/conferences/2003/Documents/CSER_Davis.pdf.
- Denshire, S., 2014. On autoethnography. *Curr. Sociol. Rev.* 62 (6), 831–850.
- Diamond, R., 2006. Why colleges are so hard to change? *Inside High. Ed* (1–8). Retrieved from: <http://www.insidehighered.com/views/2006/09/08/diamond>.
- Erickson, P., 1986. Qualitative methods in research on teaching. In: Wittrock, M. (Ed.), *Handbook of Research on Teaching*, third ed. Macmillan, New York, pp. 119–161.
- Fenstermacher, G., 1992. Policy Development and Teacher Education: An Educative Agenda vs. A System of Schooling. Invited Presentation. University of British Columbia, Vancouver, BC.

- Fernandez, C., 2002. Learning from Japanese approaches to professional development: The case of lesson study. *J. Teach. Educ.* 53 (5), 393–405.
- Fitzpatrick, S., 2011. Can you write a history of yourself? *Griffith Rev.* 33, 1–7. Retrieved from: <https://www.griffithreview.com/articles/can-you-write-a-history-of-yourself>.
- Ghalipour, B., 2017. Writing a memoir is a strange psychological trip through your past. *N. Y. Times Mag.* Retrieved from: <https://www.thecut.com/2017/04/what-happens-to-your-memories-when-you-write-a-memoir.html>.
- Giudici, C., Rinaldi, C., Krechevsky, M., Barchi, P., Gardner, H., Filippini, T., 2001. Making Learning Visible: Children as Individual and Group Learners. Project Zero, Harvard Graduate School of Education, Cambridge, MA.
- Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., Lubans, D., 2017. Effects of professional development on the quality of teaching: results from a randomised controlled trial of quality teaching rounds. *Teach. Teach. Educ.* 8, 99–113.
- Hall, S., 1990. Cultural identity and diaspora. In: Rutherford, J. (Ed.), *Identity: Community, Culture, Difference*. Lawrence & Wishart, London, England, pp. 2–27.
- Heilbrun, C., 1995. *The Education of a Woman: The Life of Gloria Steinem*. Dial Press, New York.
- Houston, W.R., Haberman, M., Sikula, J., 1989. Teacher education as a field of scholarly inquiry. *Action Teach. Educ.* 11 (2), 19–24. <https://doi.org/10.1080/01626620.1989.10462720>.
- Keightley, E., 2010. Remembering research: memory and methodology in the social sciences. *Int. J. Soc. Res. Methodol.* 13 (1), 55–70.
- Kung, S., Giles, D., Hagan, B., 2013. Applying an appreciative inquiry process to a course evaluation in higher education. *Int. J. Teach. Learn. High. Educ.* 25 (1), 29–37. Retrieved from: <https://files.eric.ed.gov/fulltext/EJ1016417.pdf>.
- LaBoskey, V.K., 2004. The methodology of self-study and its theoretical underpinnings. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.L., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Springer, Dordrecht, pp. 817–869. https://doi.org/10.1007/978-1-4020-6545-3_21.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511815355>.
- Lawrence-Lightfoot, S., Davis, J.H., 1997. *The Art and Science of Portraiture*. Jossey-Bass, San Francisco. Retrieved from: <https://www.amazon.com/Art-Science-Portraiture-Sara-Lawrence-Lightfoot/dp/0787910643>.
- Leggo, C., 2014. Learning poetically: remembering school. In: Cohen, A., Bai, H., Leggo, C., Porath, M., Meyer, K., Clarke, A. (Eds.), *Speaking of Learning: Recollections, Revelations, and Realizations*. Sense Publishers, Rotterdam, pp. 43–69.
- Lewin, K., 1946. Action research and minority problems. *J. Soc. Issues* 2 (4), 34–46. <https://doi.org/10.1111/j.1540-4560.1946.tb02295.x>.
- Loughran, J.J., 2004. A history and context of self-study of teaching and teacher education practices. In: Loughran, J.J., Russell, T.L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Springer, Dordrecht, pp. 7–39. https://doi.org/10.1007/978-1-4020-6545-3_21.
- MacDonald, C., 2012. Understanding participatory action research: a qualitative research methodology option. *Can. J. Action Res.* 13 (2), 34–50.
- Marton, F., 1986. Phenomenography—a research approach to investigating different understandings of reality. *J. Thought* 21 (3), 28–49.
- Marton, F., 2015. *Necessary Conditions of Learning*. Routledge, New York. <https://doi.org/10.4324/9781315816876>.
- Mishler, E., 1979. Meaning in context: is there any other kind? *Harv. Educ. Rev.* 49 (1), 1–19. <https://doi.org/10.17763/haer.49.1.b748n4133677245p>.
- Moje, E.B., 1996. “I teach students, not subjects”: teacher-student relationships as contexts for secondary literacy. *Read. Res. Q.* 31 (2), 172–195. <https://doi.org/10.1598/RRQ.31.2.4>.
- National Board for Professional Teaching Standards, 2016. What Teachers Should Know and Be Able to Do. NBPTS, Arlington, Virginia. Retrieved from: www.accomplishedteacher.org/wp-content/uploads/2016/12/NBPTS-What-Teachers-Should-Know-andBe-Able-to-Do-.pdf.
- Palmer, P.J., 1997. The heart of a teacher identity and integrity in teaching. *Change* 29 (6), 14–21. <https://doi.org/10.1080/00091389709602343>.
- Quigley, C., Trauth-Nare, A., Beeman-Cadwallader, N., 2015. The viability of portraiture for science education research: learning from portraits of two science classrooms. *Int. J. Qual. Stud. Educ.* 28 (1), 21–49. <https://doi.org/10.1080/09518398.2013.847507>.
- Radden, J., Varga, S., 2013. The epistemological value of depression memoirs: a meta-analysis. In: Fulford, K., Davies, M., Gipps, R., Graham, G., Sadler, J., Stanghellini, G., Thornton, T. (Eds.), *The Oxford Handbook of Philosophy and Psychiatry*. Oxford University Press, Oxford, pp. 99–115.
- Richlin, L., 2001. Scholarly teaching and the scholarship of teaching. In: Kreber, C. (Ed.), *Revisiting Scholarship: Identifying and Implementing the Scholarship of Teaching*, New Directions for Teaching and Learning. Jossey-Bass, San Francisco, pp. 57–68. <https://doi.org/10.1002/tl.16>.
- Rowe, W.E., 2014. Positionality. In: Coghlan, D., Brydon-Miller, M. (Eds.), *The SAGE Encyclopedia of Action Research*. SAGE Publications Ltd, London, pp. 628–631. <https://doi.org/10.4135/9781446294406>.
- Runesson, U., 2015. Pedagogical and learning theories and the improvement and development of lesson and learning studies. *Int. J. Lesson Learn. Stud.* 4 (3), 186–193. <https://doi.org/10.1108/JLLS-04-2015-001>.
- Samaras, A.P., Hicks, M.A., Berger, J.G., 2004. Self-study through personal history. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Springer Press, Dordrecht, Netherlands, pp. 905–942. https://doi.org/10.1007/978-1-4020-6545-3_23.
- Shulman, L., 1981. Disciplines of inquiry in education: an overview. *Educ. Res.* 10 (6), 5–23. <https://doi.org/10.3102/0013189X010006005>.
- Sparks, D.M., Pole, K., 2019. “Do we teach subjects or students?” Analyzing science and mathematics teacher conversations about issues of equity in the classroom. *Sch. Sci. Math.* 119 (7), 405–416. <https://doi.org/10.1111/ssm.12361>.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development*. Heinemann Educational Books, London.
- Stone, A. (Ed.), 1981. *The American Autobiography: A Collection of Critical Essays*. Prentice-Hall, Englewood Cliffs, NJ.
- Tidwell, D., Heston, M., Fitzgerald, L. (Eds.), 2009. *Research Methods for the Self-Study of Practice*, vol. 9. Springer: Science & Business Media.
- Watson, L., Watson-Frank, M.-B., 1985. *Interpreting Life Histories: An Anthropological Inquiry*. Rutgers University Press, New Brunswick, NJ.
- Zeichner, K., 1987. Alternative paradigms of teacher education. *J. Teach. Educ.* 34 (3), 3–9. <https://doi.org/10.1177/002248718303400302>.
- Zeichner, K., 1998. The New Scholarship in Teacher Education. A Paper Presented at the Annual Meeting of the American Educational Research Association. American Educational Research Association, San Diego, California.

Enacting inclusive pedagogy in teacher education: creating a learning space for teachers to develop their professional identity

Hafþís Guðjónsdóttir and Edda Óskarsdóttir, School of Education, University of Iceland, Reykjavík, Iceland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	243
Theoretical framework	244
Inclusion	244
Inclusive pedagogy	244
Teachers' professional identity	245
Teacher education pedagogy	246
Methodology and methods	246
Participants	247
Data sources	247
Data analysis	247
Ethical issues	247
Presentation of findings	247
The course: working in inclusive practices	247
About the course	248
Understanding inclusion and inclusive education	249
Professionalism in teaching diverse groups of students	250
Discussions	252
References	252

Introduction

An essential element of inclusive education involves ensuring that all teachers are prepared to teach all students, as worldwide they are working with increasingly diverse groups of students in inclusive practices. This calls for teachers to have opportunities to engage with inclusive education theories, policies, and examples of practice through their professional education, both initial education and professional development.

Inclusion can be realized when teachers see themselves as agents of change, with the values, knowledge and attitudes that allow every student to succeed (UNESCO, 2020). Preparing teachers reflects the standards and qualifications in each country context. But a common theme is that education systems are moving away from focusing on the learner as a problem and toward “identifying barriers to learning and participation and providing anticipatory responses, planning for all learners up front” (European Agency for Special Needs and Inclusive Education, 2015, pp. 14–15).

A recent audit of the Icelandic system of inclusive education, conducted by the European Agency for Special Needs and Inclusive Education, concluded that there is a strong foundation in legislation and policy in line with the relevant international conventions (European Agency for Special Needs and Inclusive Education, 2017). The legislation builds on a broad conceptualization of inclusion as an education system built on justice and equity, enabling all learners to participate, make progress and enjoy learning. Findings from the audit show, however, that the implementation of the legislation into practice has not been straightforward. Furthermore, many teachers and administrators question the degree to which their initial education and/or ongoing professional development opportunities have prepared them for the realities of inclusive practices and to what extent they are enabled to implement inclusive education as a rights-based approach for all learners (European Agency for Special Needs and Inclusive Education, 2017).

According to Óskarsdóttir (2014), teachers report that they are not prepared to teach diverse groups of students. They believe they need more knowledge on different disabilities, learning difficulties, behavior problems and how to work with learners with diverse language backgrounds. However, the problem is not that inclusive education is not addressed in the teacher education program; the problem is that there is no consensus on the emphasis in the program because teacher educators have different views as to how to prepare teacher students to work in inclusive settings. Some courses at the University of Iceland respond to inclusive education by focusing on developing teachers' knowledge of various learning disabilities, while others focus on building teacher competences in designing their practice to be effective for all learners (Guðjónsdóttir and Óskarsdóttir, 2020). This is similar to what can be seen out in the field, different understandings result in different responses (Cochran-Smith and Dudley-Marling, 2012; Forlin and Chambers, 2011; Florian and Pantić, 2017).

Courses offered at UI will continue to emphasize different aspects and build on different theoretical approaches (Forlin and Chambers, 2011). However, each course should be seen as a piece in the puzzle of a complete teacher education that prepares student teachers for their role in the inclusive education system. This can be achieved by enabling them to build their personal theoretical framework, choose pedagogical approaches and to take a disciplinary stance.

In this article we explore a graduate course taught at the School of Education, University of Iceland, grounded in a pedagogy that integrates professional knowledge about teaching, learning and child development, and involving an ethical and social commitment to all students. We examine two main themes: inclusive pedagogy, and the professionalism of teachers. Through these two themes, the aim is to clarify important issues in the preparation and ongoing professional development of teachers working in inclusive practices in Iceland. Our research question is: What kind of learning environment in teacher education would push our students to take a stance and responsibility toward inclusive practices? Analyzing and interpreting data from the course, we will discuss how professionals are tutored to realize their theoretical and pedagogical stance and to be critical as they influence the teaching, learning and reconstruction of inclusive schools.

Theoretical framework

The theoretical framework for this paper centers around inclusion, inclusive pedagogy, teachers' professional identity and teachers' education pedagogy. Each of these concepts is briefly defined in this section.

Inclusion

Inclusive practice is fundamentally grounded in the ideologies of social justice, democracy, human rights, and full participation of all (UNESCO, 2020). These ideologies are connected and interdependent in various ways. A socially just education system is premised on the idea that quality education is the democratic right of all, rather than a prize to be competitively fought for (Reay, 2012). This idea is integral to inclusive education systems as those advocating for social justice seek a world that is fair and equitable toward everyone, not a world where everyone gets the same opportunities in order to reach the same goals.

The development of how inclusion is defined focuses on the notion of diversity, rather than disability, and how schools respond to and value a diverse group of pupils (Florian, 2021). Diversity is a natural attribute of a school community, mirroring the wider community, and can be explained as the range of elements that result in a perception of difference among people. Inclusion is aimed at steering attention toward inequalities presented in marginalization, exclusion and discrimination against diversities such as social and ethnic circumstances, religion, gender, ability of students and their families (UNESCO, 2020). Thus, inclusive education can be seen as a movement against marginalization and exclusion and an adverse reaction to social inequality (UNESCO, 2015). As such, it is an ongoing process focusing on continuous development, working against inequality, and enhancing people's sense of belonging in school and society.

Furthermore, inclusion implies a shift from emphasizing the source of learning differences or difficulties in school as coming from within the pupil or stemming from his/her social circumstances to viewing the problem as the influence of the system of education or the environment (Schuelka et al., 2019). School is then a place where difference is understood as fundamental to human development and the aim is to create rich learning environments for all students through differentiation, focusing on what is to be taught rather than who is to learn it (Óskarsdóttir et al., 2018). Developing inclusive practice demands that teachers and other school personnel maintain a welcoming disposition toward diversity, an understanding of learning as a creation of meaning and an assumption that all pupils are competent (Biklen, 2020; Guðjónsdóttir and Óskarsdóttir, 2016; Hart et al., 2007). So, rather than expecting pupils to assimilate to existing school and subject structures, the curricula, teaching practices and learning situations must mirror the inherently diverse human nature.

Inclusive pedagogy

The term pedagogy appears in the educational literature to explain the various and composite tasks involved in the act of teaching. Three consistent uses of the term "pedagogy" can be found in the literature; (a) to cover teaching methods, instructional programs and curricula; (b) as an all-embracing term for education in poststructuralist thought; and (c) to express and address moral education and discourse about teaching and learning (Bruner, 1996; Freire, 2005; Van Manen, 1991, 1999). In this article the term pedagogy is understood in accordance with the definition by Alexander (2013) as composed of the act of teaching and the ideas, values and beliefs informing, sustaining, and justifying that act.

The distinction of pedagogy employed in an inclusive school is that it is based on the ideas of equity and a fundamental change in how teachers approach and think about pedagogy:

- From a focus on traditional, individualized approach to learner diversity that begins by preparing for most pupils and then offering something additional (special) to some learners identified as having special needs.
- Toward an inclusive pedagogical approach that begins with the learning of everybody, focusing on how to create rich learning opportunities and make them available to all learners so that everybody can participate in the classroom community (Black-Hawkins, 2017).

This means that teachers (and others) need to move away from a focus on diagnostic labels attached to pupils. Categories of diagnosis and their associated labels can negatively affect how teachers think about a pupil's ability and are problematic for a number of reasons. They support a deficit model of learning where the pupil is viewed as "faulty" and in need of repair, rather than focusing on an exploration and development of the learning context. Diagnostic categories can overlap, and there are huge variations between people identified as being members of the same group. Focusing on those categories and labels provides an oversimplified view of a person and lacks the richness and complexity inherent in everyone's life (Black-Hawkins, 2017). The very act of focusing on difference intensifies the isolation and marginalization of children and adds to the social construction of disability (Grenier, 2010).

The alternative is to focus on what is to be taught and how this can be presented to accommodate diverse ways of learning (Florian and Spratt, 2013). The practice of teaching diverse groups of pupils is grounded in a pedagogy that includes more than a skill in using prescribed instructional practices. Rather, this practice integrates professional knowledge about teaching, learning and child development, and involves an ethical and social commitment to children. Pedagogical qualities of the responsive professional teacher are witnessed in teachers who understand child development and individual differences, are committed to the education of all pupils, and who have a knowledge base enabling them to differentiate between pupils as they develop a curriculum for all pupils (Guðjónsdóttir, 2000).

Teaching for equity, where teaching practices that presume diversity from the beginning are employed and built on a recognition of pupil resources, can be considered simply as good teaching for all. All pupils bring valuable resources and experiences to the classroom. These resources are their talents, strengths, and skills, built upon their personal experience, knowledge, and beliefs. Teachers who understand their pupils' resources can better attune their teaching to the resources pupils bring to the classroom. Gonzalez et al. (2005) see culture as "funds of knowledge" or, in other words, a resource to draw upon in the process of learning and empowering of pupils.

These connotations of inclusion challenge all teachers to investigate their values and beliefs, to reflect on their understanding of teaching, learning and curriculum, and to reinvent their roles as participants in school change. Through the recognition of their identities, their unique qualities and their roles, teachers have more opportunities to empower their pupils. Realizing that they are the makers of their culture and by showing their knowledge and understanding of the complexity and diversity of the world in which their pupils live they can become cultural workers (Freire, 2005). In working with diverse pupils, it is critical that teachers understand how their own identities influence the ways they work and how they respect the value of diverse identities. People tend to look at those who are different from themselves as inferior (Freire, 2005). Teachers that become aware of this central problem, aware of their own prejudices, take a special stance and make sure their beliefs do not hinder them from transferring their cultural work.

Teachers' professional identity

Understanding teachers' professional identity defined as the beliefs, values, and commitments an individual holds toward being a teacher can help them to understand who they are as teachers (Sachs, 2001). For teachers, professional identity is mediated by their own experience in schools and outside of schools, as well as their own beliefs and values about what it means to be a teacher and the type of teacher they aspire to be. Identity is not something one has, but something that develops during one's whole life and is not a fixed entity but an ongoing process of interpretation and re-interpretation of experiences, beliefs and knowledge (Sachs, 2005). Teacher values are important resources to enrich the quality of teaching and learning experiences for teachers and their students. Knowledge of both themselves and their personal values can help teachers handle the challenges of their practice and well-being (Barni et al., 2019).

Agency is an important element of professional identity, meaning that teachers have to be active in the process of professional development (Coldron and Smith, 1999). Teacher agency is not something that teachers can have, rather it is something they do or achieve (Biesta et al., 2015). It is a developing phenomenon with a focus on past, future and present, but only acted in the present as it is about the interplay between what teachers bring to their practice and what the practice brings to them. If teacher agency today is influenced by experiences from the past, then it can be assumed that today's contexts will impact the future agency of teachers (Priestley et al., 2015). Risk taking and changing practice calls for a professional agency where individuals act independently and make their choices. Teacher professional agency represents how teachers define themselves both to themselves and to others and how they respond to theoretical ideas or government acts (Biesta et al., 2015). Teachers practice professional agency as they influence, make choices, and take a stance in a way that affects their work and their professional identities (Eteläpelto et al., 2013). It is formed through their lived experiences, dialog and internal conversation and its development is shaped by historical, sociological, psychological, and cultural influences (Dalmau and Guðjónsdóttir, 2017). By creating a space in teacher education for student teachers and their tutors to reflect on their practical experience, educational theories, or knowledge, and ethical or moral principles they are given an opportunity to recognize their professional identity. Using multifaceted ways together with multi-dimensional, multi-layered, and dynamic processes student teachers and their tutors can identify their professional agency (Dalmau and Guðjónsdóttir, 2017).

Teacher education pedagogy

The philosopher Skúlason (1989) had the following to say about the art of teaching:

To teach is even harder than to learn. People know this, but seldom think about it [...]. Teaching is harder than learning because teaching means getting people to learn [...]. A teacher is not ahead of students except to the extent that he has to learn much more than they do because he must get them to learn (p. 10).

Teaching people to become teachers can, therefore, be a complex profession because it is more than simply delivering information about the content of a subject or about teaching methods (Loughran, 2013). The pedagogy of teacher education focuses on the complex interplay between human, material, and non-tangible elements (Hordvik et al., 2020). It is about understanding how this complex interplay influences the relationship between teacher and student, student and student, and learner and subject matter, encompassing both context and content (Loughran, 2013). In their teaching, teacher educators have to take into account the different ways people learn when making decisions about how they teach people to become teachers. There is value in giving student teachers space to understand what kind of teachers they aspire to be, and in that space teacher educators give precedence with their teaching, showing how learning influences teaching and the way we teach. On the other hand, teacher educators need to acknowledge the fact that it is beyond their agency to address all the uncertainty, ambiguity, and unpredictability that characterizes the practice of teaching (Hordvik et al., 2020). In the face of this, teacher educators need to support teacher students in developing a mindset of hope, of looking at challenges as opportunities for growth and new learning.

Preparing teachers for working in inclusive school settings is a key factor in teacher education; a factor that requires teachers and teacher educators to take the initiative to move out of their comfort zone, risking uncertainty while exploring new or diverse ways of working, planning, and organizing (Guðjónsdóttir and Óskarsdóttir, 2020). Seeking collaboration with colleagues to problem solve, debate, and celebrate success, where each contributes through individual strengths and interests, can support teachers in this endeavor. The effort to develop inclusive practice will be rewarded with more effective classrooms that can reach and engage a broad spectrum of pupils in an equitable way.

Responsiveness, coupled with pedagogical skills, enables teachers to use pupil differences, contextual issues, cultural and community events, subject matter, and problems and challenges as opportunities for effective teaching and learning. However, responsive teachers go beyond acknowledging and respecting differences as they create the curricula. They focus on pupils and what they bring into the classroom in the shape of their resources (ability, attitude, background, experience, interest, knowledge, and skills), and respond by designing a universal learning environment that supports all pupils in expanding their learning (Guðjónsdóttir, 2000). Responsive teachers are skilled in creating a curriculum of learning activities and environments in which all pupils can succeed. This means that the responsive teachers' preparations are solidly grounded in major theoretical perspectives on child development and their practical applications, in an understanding of the social and moral issues inherent in individual differences, as well as in a strong professional commitment to the education of pupils (European Agency for Development in Special Needs Education, 2011; Guðjónsdóttir, 2003; Guðjónsdóttir et al., 2007).

Methodology and methods

The research builds on the methodology of self-study of teaching and teacher education practices (S-STTEP), where we as the researchers and teacher educators focus on inclusive education in our practice through the critical and collaborative lens of self-study (Bodone et al., 2004). The critical focus was employed to reveal the position of dominant practice to support us in becoming more conscious of both the transformation that takes place and the process that leads to that transformation (Kincheloe, 2003; Pin-negar and Hamilton, 2009). The S-STTEP methodology was a frame for us to examine the course; it was a way to think about and discuss our practice in collaboration (Tidwell and Staples, 2017). Thus, we shared reflections about our responses and the development of practice. The purpose of this research was to gain an insight into how teacher education can respond to the call for preparing teachers to be responsive, see themselves as agents of change, with the values, knowledge and beliefs that allow every learner to succeed. The aim was to cast a light on important issues in the preparation and ongoing professional development of teachers and student teachers. Our research question was: What kind of learning environment in teacher education would push our students to take a stance and responsibility toward inclusive practices?

The focus was on a graduate course, *Working in Inclusive Practices* (WIP), taught at the School of Education, University of Iceland, at graduate level. The course was established in 1999 and through the methodology of self-study of teacher educational practice it has been under constant reconstruction. The course is grounded in a pedagogy that integrates professional knowledge about teaching, learning and child development, and involves an ethical and social commitment to all students. Through analyzing data from the course, we will discuss how professionals are enlightened and encouraged to realize their theoretical and pedagogical stance and be critical as they influence teaching, learning and reconstruction of inclusive practice.

Participants

There are six teachers teaching this course and two of them participate in this research. Hafdís, who created the course in the beginning and has led it since, and Edda, who has been an assistant teacher for a decade and will be leading the course next term. About 60 students attend the course each year and the group is always diverse. Some are in their initial teacher education, others are coming back for their graduate studies, some are pre-school teachers and others are compulsory or secondary level teachers. Some are distance students, while others take the course on campus. All students sign a consent form at the beginning of the course that we can use their tasks, *Tickets Out of Class* (TOCs) and comments. All names that appear in the findings are pseudonyms.

Data sources

The data in focus was collected in WIP for four academic terms (2017–2021). The course is 10ECTS and taught every fall term for 15 weeks. Data collection consisted of:

- Research journal. We write a journal about our teaching, collaboration, meetings, and thoughts. They are our personal accounts that we analyze individually and share with each other. We also communicate through e-mails and messages on Microsoft Teams. We collected those data sources for analysis and interpretation.
- Students' tasks. These were collected through the four years. Our students submit their tasks online, making these and their discussions available. Through the course students submit four main projects. The last task for students is to assess their learning with relevant arguments.
- TOC (ticket out of class). At the end of each class, students turn in a ticket where they reflect on what they learned that day and what they would like to focus on in the next course session. We keep all these tickets and analyze them after each class before we plan the next one.
- The course schedule, lesson plans and presentations were also part of the data source.

Data analysis

To ensure trustworthiness we analyzed the data individually and in partnership aiming for transparent and systematic data analysis (Wolcott, 2001). The methods for transforming the data into findings, and the linkages between data, findings, and collective interpretations provided explanatory and meaningful insights into our practice (LaBoskey, 2004). At the first stage of our systematic analysis, we read through all our data to locate the learning activities focused on inclusive education and/or professionalism.

We employed document analysis as we read through our data and made a list of these learning activities, looking for evidence of how we teach about inclusion and how we support our participants in order to enable them to teach and work with diverse groups of students. At the next stage we analyzed the work of participants, such as their projects, online discussions and TOCs, through the lens of inclusive education and professionalism. The third stage involved analysis of our personal journals. The final stage of the process focused on interpreting our collective analysis to gain an in-depth understanding and transparent view of how we support participants in creating inclusive practices.

Ethical issues

In the research project, we followed ethical standards for research, the Act on Personal Data Protection and the Processing of Personal Data (No. 90/2008) together with the University of Iceland's Code of Ethics (2014). From the outset of teaching the course, we have collected data on our teaching practice. To ensure that our research was ethically sound, we discussed the research plan with all participants at the beginning of each term and asked the willing participants to sign an informed consent form regarding our data collection. The anonymity of participants was ensured by giving them pseudonyms and de-identifying the data. All data have been kept in a secure place and only the authors have access to them.

Presentation of findings

The following article is organized into three sections based on the main themes. The first focuses on overall information about the course, the next focuses on how we work with the understanding of inclusion and inclusive practice and the third on professionalism and the professional working theory. The article ends with a summary.

The course: working in inclusive practices

Diverse groups of students in inclusive schools call for professional teachers who have a strong grasp of pedagogical competence and content, an understanding of child development and individual differences and who are devoted to the education of all their students. In this article we (the researchers) will introduce the WIP course as an example, the structure, and the tasks. We use quotes from the course material and participants' assignments to describe and illustrate our interpretation of our findings.

About the course

As we began to develop WIP, the aim was to prepare participants to draw on their resources in their work with diverse groups of students. Resources refer to individual talents and skills, knowledge and beliefs, experience, and culture. Participants bring valuable resources to the course and our intention is to draw upon these as we establish a learning environment that supports them in making meaning and acting.

The course is taught simultaneously on-line and on campus and the content is divided into four main themes as stated in the course syllabus: inclusion, professionalism and leadership, inclusive pedagogy, and collaboration. The themes relate theory and practice in inclusive educational settings, aiming at learning outcomes that give participants the competence to critically engage with issues and ideas inherent in learning and teaching in inclusive schools. This includes being able to design an educational environment, practice and pedagogy that considers students' diverse resources. Through this work and reflecting on their professional identity and beliefs the participants build their professional working theory (PWT).

The course methodology is centered on ideas from inclusive pedagogy and innovation education. Both build on social constructivist approaches, which are exemplified by reflection, collaborative discussions, and opportunities for students to develop new understandings and shape their learning. Innovation education focuses on creativity and knowledge in meeting needs or solving problems identified as important. The aim is to help people develop critical and creative thinking through dealing with every-day situations and gain capacity for action, using creative and critical problem solving. The core pedagogy builds on the ideas of learners as creative explorers and teachers as facilitators. Inclusive pedagogy emphasizes holistic and creative approaches to teaching practice, assessment, and collaboration within and around school. It begins with the learning of everybody and how the teacher can create rich learning opportunities available for all learners in the classroom to ensure their participation. It involves developing a system of offering equitable learning opportunities built on students' resources and the premises of universal access, active participation and belonging in school. Through the pedagogy of innovative and inclusive education, the goal is to facilitate the development of creative problem solving to enable participants to create learning spaces for diverse groups of students in inclusive schools.

The learning space created in the course is designed to provide participants with opportunities to reflect on their learning after each session through TOCs and by listing the teaching methods used. The TOCs are designed to give everyone a voice and a conscious participation in the course development. For each TOC, participants reply to the questions: What do you take with you from today's lesson? What do you want to know more about? Their replies are anonymous and here are examples of responses:

I learnt about teacher ethics in professional theory. Interesting.

I want to learn more about the ethical part of the PWT.

I learnt that my PWT is work in progress—one must remember that.

I want to learn more about collaboration between school and parents.

Our response is to adjust the next session, based on what was called for in the TOCs, either with increased emphasis on what was already discussed or with a change in focus.

At the end of each session, participants list the different teaching methods we employ. The list is compiled on the whiteboard and posted on Canvas. After one such session the list read:

- Group work
- Problem solving
- Think-Pair-Share
- Walk and Talk
- ...

The aim of the list is to give an overview of, as well as model, the pedagogy we use. Furthermore, our aim is for participants to practice identifying and discussing teaching methods.

One essential ingredient of the learning space created is the freedom participants are allowed in their assignments. The freedom involves stipulating in the assignment descriptions that creativity is encouraged, such as using prose, video, artwork, or some other innovative ways in their assignments. Many participants enjoyed the freedom and flexibility of the course. Stella stated she “really

enjoyed the creativity. I have not had the same opportunity to play like this and create in other courses.” For others the freedom meant that they were outside their comfort zone, as Sigrún phrased it: “Creativity is not something I would consider my strong side at all,” but agreed that it helped their learning.

While participants are encouraged to be creative, they must provide an argument for their work and link it to literature as needed. As an example, participants evaluate their participation and learning through a reflective self-assessment at the end of the course. This learning activity is a critical reflection on and evaluation of their work during the course and on what they take away with them from their participation in the various tasks and discussions. Rebekka wrote in her self-assessment:

For students to evaluate their own progress in the course and give themselves a grade is a good example of inclusive practice. I have never done this before and feel this is a great idea that challenges the student to look inward, at their moral values, independence, self-confidence and much more.

Participants are required to support their reflection with descriptive examples from their activities, such as excerpts from their study journal, examples from online discussions or critical moments in the course work. While we have found that the self-assessment is a practical tool for reflection that supports participants’ personal growth and professional development, many participants find it difficult to reflect on their own learning but are ready to evaluate our work instead.

Understanding inclusion and inclusive education

People have different understandings of inclusive education, what it is and what it calls for in practice. At the outset of the course, the participants are required to read selected texts and take part in an online discussion where they explore their understanding of inclusive education. This is done so we can analyze the understanding in the group—where they are at—so that we can begin working with them from that point. In that exercise, many express the opinion that inclusive education is a beautiful idea, but that it does not work in practice. Many participants agree that “the ideology behind inclusive education is beautiful and kind but the schools lack financing and resources to be able to work in that spirit” as Hulda, a student teacher, wrote in online class discussions.

Through a brainstorming exercise, the participants are asked to list the challenges of inclusive practice in schools. The list usually has issues such as lack of financing, professionals in schools, resources, and equipment, learning material and knowledge of disabilities. Participants also list the challenge of dividing attention and quality education between students with various abilities. In the words of Linda, “neither I nor other teachers can focus on one individual while we have 23 others that also need our assistance.” The aim is for participants to see how they can influence the challenges they listed, and through creative problem solving they look at their own contribution in creating inclusive practices. We organize a walk and talk, hang thirteen posters around the school building and arrange the participants into an equal number of groups. Each poster poses a challenge from the brainstorming exercise and participants use creative problem solving to come up with solutions to the challenge. The groups rotate through the thirteen questions, walk and talk, and at the end bring completed posters back to the whole group for sharing. While writing their solutions, they are asked to focus on issues within their own power to change, not on issues such as more money, resources or adding people into the schools.

Throughout the course participants write columns. In these columns they focus on their ideas, experiences or examples of practice supported by arguments and by referring to literature. There are four columns in all, each linked to the course themes: (1) theories of inclusive education and innovation education, (2) professionalism in teaching diverse groups of students, (3) inclusive teaching practices and pedagogy, and (4) collaboration in the inclusive school. For each column, the participants focus on the theme, reflect on their understanding and beliefs, experience and knowledge, and take a stance. The columns are posted online for others in the group to comment on and discuss. This process takes a week—from when the column is posted to when the discussions have ended. The aim of the task is to give everyone a chance to communicate their ideas and take part in an informed discussion with other participants. Everyone is given a space to have an opinion in a constructive conversation.

Participants write about various aspects of each theme and as Sigga wrote: “I admit that I was scared of writing the columns, but when I read the instructions and how I could use the writing to connect to my work I was unstoppable, and I wanted to write so much more.” Another participant, Jón, who is a school principal, realized that:

Reading other participants’ columns gave me an insight into the mindset of the teacher students who are in their first years of teaching, which is a good reminder for me that as a principal I am also “dealing with” diversity in my staff.

Here he is referring to the multiple viewpoints, beliefs and discussions that appeared in the columns as he acknowledged the difference in participant experience and backgrounds, which he can also expect in his staff.

In the assignment: Learning and Teaching Project, everyone chooses an issue related to the teaching of diverse groups of students, presenting and discussing it in a critical way. The project can be an individual project or a small piece of group work. Participants can focus on a specific teaching method, teaching methods, assignments, study materials, different pedagogical emphases, or other

topics/issues in connection with the teaching of diverse groups of students in inclusive schools. Examples of the assignments turned in are: Teaching deaf children how to read; Cooperative learning; Using multiple intelligence in the preschool; Icelandic saga: Njála for all.

Toward the end of the course, a transformation can be witnessed in participants' understanding of inclusive practice and pedagogy. Many state that their views and knowledge of the practical aspects of organizing inclusive schools have changed as they progressed through the course. Andri admitted that:

In my work I have been about to give up and sometimes wonder whether a student belongs in a traditional school system. After this course, my attitude is different and I believe it is possible to work with all individuals [...] but you need to be solution-oriented and open-minded.

Some also made connections to the importance of being aware of and working against prejudices visible in schools toward students who are different, have a foreign background or find it hard to fit into the system. Helga "wondered about the prejudice I have witnessed in teachers and other school staff and attitudes toward inclusive education. It is important to be vigilant toward these." Learning to meet the diverse needs of students is one of the greatest challenges of being a teacher. Participants have found that they have gained the "knowledge to design learning environments that build on and attend to students' resources" which is one of the central competence criteria for the WIP course.

Professionalism in teaching diverse groups of students

The course is based on the fundamental belief that awareness of professional identity is the core competence of teachers in inclusive practices. To create this awareness, we design tasks where participants take part in ongoing and critically reflective professional dialogs as they build their professional working theory (PWT). Guðný stated:

Since the beginning of the semester, I have had this project at the back of my mind. I have been wondering what kind of teacher I am, what kind of teacher I want to be, what I am doing well and what I could do better.

PWT is a living phenomenon rooted in practice, theory and ethics and is continuously evolving in the interplay between these aspects. The interplay means that the PWT is constantly being shaped as participants develop their understanding and are encouraged to reflect on how their life experiences affect their identities and professionalism and vice versa in the way professional knowledge, educational theories and practice influence beliefs and identity. Hermann wrote in his assignment that "In this project I have the opportunity to form an opinion of what kind of teacher I want to be and what expectations I have for the job." He expressed the thinking that the PWT has been in development throughout the course. PWT is explicitly developed through systematic and comprehensive critical reflection and dialog with colleagues as it contributes to the construction of professional identity, the creation of professional knowledge, and the transfer of collaborative approaches into practice. Participants are encouraged to use metaphors to explain their PWT and Hrefna compared her work with children to knitting:

... as I enjoy knitting just as much as working with children, I build on previous experience with new methods and solutions. Sometimes the methods I have been using do not work and then I have to go back (unstash) and try a new method that works better. The tasks can be demanding and then it is necessary to have a good dose of patience and resilience in order to be able to complete them. The joy of creation is often experienced with children, and it is always exciting to start a new project. The loops on the knitting needle are all the links that are related to the work and form the whole piece.

The PWT enables participants to systematically explore socio-cultural and historical influences on the practice of teaching. Like a braid, the PWT has three key strands: professional practice, educational theories or knowledge and ethics or moral principles. In the course, participants are asked to reflect on each strand to identify their professional identity:

- By looking at their practice, participants are provided with a space to explore their experience of their professional work and roles, including teaching, assessment, evaluation, collaboration with colleagues, and relationships with students and parents.
- By looking at theories the participants are provided with a space to reflect on their knowledge and the way they understand and relate practice to theory. This reflects teachers' theoretical framework and their explanations of what happens in the classroom as well as representing their method of relating self-understanding and reflective practice to theory.
- By looking at ethics, participants are provided with space to explain the reasons behind their practice and relate to their beliefs and values. This reflects what they are becoming, and what they want to be as teachers.

The course aim is to support participants to make their PWT explicit so they can continue to develop it for their future professional lives. The work takes time and calls for various methods and strategies. Therefore, allocated time to develop the PWT through discussions and tasks is organized throughout the course. Among the tasks are the Three Step Interview, column writing, poetic analysis, and core qualities, and we will introduce these shortly here.

The Three-Step Interview (a Participatory Interview) was originally designed to support participative, equitable and reflective approaches to life history interviews in a small group. Here this task is employed to provide participants with the space to explore the first strand: their professional practice. Participants, arranged in groups of three, are given three rotating roles of being an interviewer, interviewee and reflective note-taker and take turns in each role. For the Three Step Interview to be effective calls for good listening and interviewing skills, using open-ended and follow-up questions. In his TOC Fannar wrote: "The interview is a perfect way to begin analyzing one's own PWT, not only when I was interviewed but also in listening to the others being interviewed." Participants are encouraged to record the interviews on their phones so that they can revisit them for deeper reflection. The activity can be useful in working with students in teacher education because it provides opportunities to organize personal learning based on information from each student. It provides the opportunity for all to get to know each other on a deeper level and to get to know themselves, their strengths, and what they need to work on. Gunna stated that: "In my interview—I pleasantly surprised myself." For many, the interview is the first time they get a venue to openly reflect on and discuss their practice, how they think about it and what they value.

Writing the columns supports the participants in focusing their theoretical understanding and relating to their practice. The columns are the material for connecting and finding harmony between practice, theory and participants. They gain an insight into how theory can explain practice and be a rationale for practice. Ása wrote: "I used the columns to reflect on my practice and since I have very limited teaching experience, I enjoyed reading the other participants' responses, viewpoints and stances on my writing." Meeting the purpose of writing the columns, the participants took different learning paths, as Reynir put it:

I used them to take a closer look at certain aspects, innovation, inclusive education, professionalism and what professionalism in work really is and how it can be increased and maintained in terms of important cooperation and collaborative school and school communities.

In connection with the column writing, we also used the method "Poetic analysis." Poetic analysis can be defined as a method that creatively constructs text into expressive and artful representations. The purpose of using this method is to support participants in looking creatively at issues from different viewpoints and to identify the core meaning of their writing. Anna stated: "This method has completely crept under my skin, so to speak." She decided to use this method to express her PWT as she summarized her thoughts and ideas. So, the poetic analysis can be seen as both being used to analyze meaning and to bring together participants' ideas and theories. Her poem from her PWT follows (the first verse in our translation):

Positive leader, determined but fair

We learn and play inside.

For all it is best to learn and participate,

The young learn best through play and games.

Summarizing the columns is the foundation for the theoretical framework that they build their practice on. The summary is then braided into the PWT as one of the key factors of the PWT braid.

The final strand of the PWT braid are the ethics and to begin with it looks like this is more challenging for participants than reflecting on their practice. Through teaching the course, we have developed ways to work with this challenging factor. As a starting point participants explore their own core qualities. Core qualities are individual strengths, what they are good at or what others praise them for. Using Think-pair-share to support the participants to reflect on their core qualities, they are given a list of 39 descriptive words. During the "think" part of the activity they select five words from the list, then in small groups they introduce their core qualities and give feedback to each other (pair). The activity ends by everyone reporting to the larger group (share). Participants usually enjoy this task as it allows them to focus on their own qualities and how they affect people and to keep them in mind when it comes to working with others. They often state that this task opens their eyes to issues that one does not usually think about. As Heiða stated: "Finally it is all right to acknowledge your qualities and be who you are in your teaching."

The final PWT projects are submitted in various forms. Helgi stated that "I could express my thoughts and ideas not only in writing but also in artistic ways—my strength is creativity." By means of the PWT project participants are given the chance to reflect on their values, ethics and professionalism, and Mary admitted "I had not realized before that I build on certain values or ethics in my teaching." By creating a capacity in teacher education for student teachers and teachers to reflect on their professional identity and define themselves as teachers, they can develop an understanding of their teacher professional agency and how it influences their work and allows them to take a stance.

In the findings, we have introduced the course and examined two main themes: inclusive pedagogy, and the professionalism of teachers. We have compiled, analyzed, and interpreted data to show how we have designed a course that aims to enlighten professionals as they form their theoretical and pedagogical stance and create inclusive practice for all students.

Discussions

The purpose of this self-study of teacher education practice was to untangle how we as teacher educators respond to the call for preparing teachers to work in inclusive practices. Our intention was to explore one course, WIP, and analyze critical issues in the preparation of student teachers. Thus, we will discuss the learning environment we created in the teacher education at graduate level to push our students to take a stance and become responsible toward inclusive practices.

In our theoretical framework, we explored the challenge of teaching student teachers to teach and the complex interplay between students and teachers, and students and subjects (Hordvik et al., 2020; Loughran, 2013). Teaching people to become teachers is not straightforward (Loughran, 2013) and to become a responsive teacher in inclusive settings is even more complicated because people do not necessarily refer to their own experience of diverse classrooms. Our student group is diverse according to teaching experience and age, teaching level and subjects, cultural and language background and we need to respond to these differences. Our teaching team is also diverse, and every team member contributes a valuable resource of knowledge and experience (Loughran, 2013). Behind our thinking as we develop and teach the WIP course is to “walk the talk,” be inclusive, creative problem solvers and build our teaching on students’ resources. However, it can be demanding to create a learning environment of trust, collaboration and problem solving in one term. Reconstructing thinking, beliefs and learning takes time (Dalmau and Guðjónsdóttir, 2017) and we suggest that all courses in the teacher education program should consider inclusive pedagogy as a foundation for practice. In feedback from the majority of participants through the years, we see are that they are comfortable with the learning spaces created. Nevertheless, it often takes them time to get used to the freedom, find the courage to step out of their “box” and approach tasks and projects in creative ways. It can be challenging, both for participants and teacher educators, to move out of the comfort zone, risk uncertainty and explore new ways of planning, teaching, and learning (Guðjónsdóttir and Óskarsdóttir, 2020). We believe, however, that this is the basis for creating a responsive and inclusive learning environment.

Through the years, many participants have attended the course in the assumption that inclusive education is a beautiful idea that does not work in practice. Participants build their beliefs on the dominant discourse in society, in schools and among other professionals and parents (Óskarsdóttir, 2014). Many teachers and administrators also report that their initial education or ongoing professional development has not prepared them for inclusive practices (European Agency for Special Needs and Inclusive Education, 2017), and others call for more knowledge on different disabilities, learning difficulties, behavior problems or how to work with students with diverse cultural backgrounds or languages (Óskarsdóttir, 2014). The participants in the course also point out that there is lack of finance, support personnel and paraprofessionals and that teachers do not know how to teach students with diverse needs. We have grasped the need to bring this discourse on inclusion to a new level—or out of the deficit discourse that leads to pessimism. For that purpose, we have worked with participants on different tasks that gradually move them to view inclusive education from a different angle and understand their responsibility in implementing it. Introducing the methods of innovation education to the course has been instrumental in moving this thinking forward. Identifying barriers to participation and employing creative problem solving has shown participants how they can plan inclusive learning environments for all learners (European Agency for Special Needs and Inclusive Education, 2015). According to our data, it is clear that through the course participants transform their thinking about inclusive pedagogy and professionalism. They approach inclusive education as a task for them to solve and not something that others have to solve for them, thereby conveying that they are willing to take a stance and let it guide their decisions and practice.

Teachers play a key role in responding to the call for schools based on inclusion and need to see themselves as agents of change (UNESCO, 2020). This is achieved as teachers recognize their professional and personal resources, become aware of their professional identity, rationalize their practice with educational theories and acknowledge the moral principles that affect their practice (Barni et al., 2019; Eteläpelto et al., 2013; Sachs, 2001). We introduced the PWT, a living phenomenon rooted in practice, theory and ethics, as a framework for participants to identify their professional identity. Thus, by beginning to critically reflect on their practice and then relate it to educational theories and their ethics, the participants develop their professional agency (Biesta et al., 2015).

Our interpretation of the findings is that the pathway from the beginning to end of the course is not linear. In our learning environment multiple paths can be taken based on the needs, interests, and resources participants bring to the course or choose to reach their destination. As student teachers travel along these diverse pathways, they are supported to become critical and creative problem solvers and be aware of their professional identity. Their chosen pathways support them in exploring how they can influence teaching and learning and become involved in the reconstruction in Iceland of inclusive practice which offers rich learning opportunities available to all learners (Black-Hawkins, 2017).

References

- Alexander, R., 2013. *Essays on Pedagogy*. Routledge.
- Barni, D., Danioni, F., Benevene, P., 2019. Teachers’ self-efficacy: the role of personal values and motivations for teaching. *Front. Psychol.* 10 (1645). <https://doi.org/10.3389/fpsyg.2019.01645>.

- Biesta, G., Priestley, M., Robinson, S., 2015. The role of beliefs in teacher agency. *Teach. Teach.* 21 (6), 624–640. <https://doi.org/10.1080/13540602.2015.1044325>.
- Biklen, D., 2020. Presuming competence, belonging, and the promise of inclusion: the US experience. *Prospects* 49, 233–247. <https://doi.org/10.1007/s11125-020-09510-0>.
- Black-Hawkins, K., 2017. Understanding inclusive pedagogy. In: Plows, V., Whitburn, B. (Eds.), *Inclusive Education*. Sense publication, pp. 13–28.
- Bodone, F., Guðjónsdóttir, H., Dalmau, M.C., 2004. Revisioning and recreating practice: collaboration in self-study. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, vol. 12. Kluwer Academic Publishers, pp. 743–784.
- Bruner, J.S., 1996. *The Culture of Education*. Harvard University Press.
- Cochran-Smith, M., Dudley-Marling, C., 2012. Diversity in teacher education and special education: the issues that divide. *J. Teach. Educ.* 63 (4), 237–244. <https://doi.org/10.1177/0022487112446512>.
- Coldron, J., Smith, R., 1999. Active relocation in teachers' construction of their professional identities. *J. Curric. Stud.* 31 (6), 711–726. <https://doi.org/10.1080/002202799182954>.
- Dalmau, M.C., Guðjónsdóttir, H., 2017. From the beginning to the future: professional working theory emerging. In: Dalmau, M.C., Guðjónsdóttir, H., Tidwell, D. (Eds.), *Taking a Fresh Look at Education: Framing Professional Learning in Education Through Self-Study*. Sense Publishers, pp. 129–148.
- Eteläpelto, A., Vähäsantanen, K., Päivi Hökkä, P., Paloniemi, S., 2013. What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65. <https://doi.org/10.1016/j.edurev.2013.05.001>.
- European Agency for Development in Special Needs Education, 2011. *Teacher Education for Inclusion Across Europe—Challenges and Opportunities*. European Agency for Development in Special Needs Education.
- European Agency for Special Needs and Inclusive Education, 2015. *Empowering Teachers to Promote Inclusive Education*. Literature Review. European Agency for Special Needs and Inclusive Education.
- European Agency for Special Needs and Inclusive Education, 2017. *Education for All in Iceland: External Audit of the Icelandic System for Inclusive Education*. European Agency for Special Needs and Inclusive Education.
- Florian, L., Pantić, N., 2017. Teacher education for the changing demographics of schooling: policy, practice and research. In: Florian, L., Pantić, N. (Eds.), *Teacher Education for the Changing Demographics of Schooling: Issues for Research and Practice*. Springer International Publishing, pp. 1–5.
- Florian, L., Spratt, J., 2013. Enacting inclusion: a framework for interrogating inclusive practice. *Eur. J. Spec. Needs Educ.* 28 (2), 119–135. <https://doi.org/10.1080/08856257.2013.778111>.
- Florian, L., 2021. The universal value of teacher education for inclusive education. In: Köpfer, A., Powell, J.W., Zahnd, R. (Eds.), *Handbuch Inklusion International/International Handbook of Inclusive Education*. Verlag Barbara Budrich, pp. 89–106.
- Forlin, C., Chambers, D., 2011. Teacher preparation for inclusive education: increasing knowledge but raising concerns. *Asia Pac. J. Teach. Educ.* 39 (1), 17–32. <https://doi.org/10.1080/1359866X.2010.540850>.
- Freire, P., 2005. *Education for Critical Consciousness*. Continuum.
- Gonzalez, N., Moll, L.C., Amanti, C., 2005. *Funds of Knowledge*. Routledge.
- Grenier, M., 2010. Moving to inclusion: a socio-cultural analysis of practice. *Int. J. Incl. Educ.* 14 (4), 387–400. <https://doi.org/10.1080/13603110802504598>.
- Guðjónsdóttir, H., Óskarsdóttir, E., 2016. Inclusive education, pedagogy and practice. In: Markic, S., Abels, S. (Eds.), *Science Education Towards Inclusion*. Nova Science Publishers, Inc, pp. 7–22.
- Guðjónsdóttir, H., Óskarsdóttir, E., 2020. "Dealing with diversity": debating the focus of teacher education for inclusion. *Eur. J. Teach. Educ.* 43 (1), 95–109. <https://doi.org/10.1080/02619768.2019.1695774>.
- Guðjónsdóttir, H., Kristinsdóttir, J.V., Óskarsdóttir, E., 2007. Mathematics for All: Preparing Teachers to Teach in Inclusive Classrooms. *Proceedings of the 3rd Nordic Research Conference on Special Needs Education in Mathematics*, pp. 123–137.
- Guðjónsdóttir, H., 2000. *Responsive Professional Practice: Teachers Analyze the Theoretical and Ethical Dimensions of Their Work in Diverse Classrooms*. Unpublished Dissertation. University of Oregon.
- Guðjónsdóttir, H., 2003. Viðbragðssnjallir kennarar [Responsive teachers]. *Glæður* 13 (1 & 2), 59–66.
- Hart, S., Drummond, M.J., McIntyre, D., 2007. Learning without limits: constructing a pedagogy free from determinist beliefs about ability. In: Florian, L. (Ed.), *The Sage Handbook of Special Education*. Sage, pp. 499–514.
- Hordvik, M., MacPhail, A., Ronglan, L.T., 2020. Developing a pedagogy of teacher education using self-study: a rhizomatic examination of negotiating learning and practice. *Teach. Educ.* 88, 102969. <https://doi.org/10.1016/j.tate.2019.102969>.
- Kincheloe, J.L., 2003. *Teachers as Researchers: Qualitative Inquiry as a Path to Empowerment*. Routledge Falmer.
- LaBoskey, V., 2004. The methodology of self-study and its theoretical underpinnings. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, vol. 12. Kluwer Academic Publishers, pp. 817–869.
- Lög um grunnskóla nr. 91/2008, 2008. Act on Personal Data Protection and the Processing of Personal Data.
- Loughran, J., 2013. Pedagogy: making sense of the complex relationship between teaching and learning. *Curric. Inq.* 43 (1), 118–141. <https://doi.org/10.1111/curi.12003>.
- Óskarsdóttir, E., Guðjónsdóttir, H., Tidwell, D., 2018. Breaking free from the needs paradigm: a collaborative reflection. *Stud. Teach. Educ.* 15 (1), 44–55. <https://doi.org/10.1080/17425964.2018.1541290>.
- Óskarsdóttir, G., 2014. Starfshættir í grunnskólum við upphaf 21. aldar [Teaching and Learning Procedures in Icelandic Compulsory Schools in the Beginning of the 21st Century]. Háskólaútgáfan.
- Pinnegar, S., Hamilton, M.L., 2009. *Self-Study of Practice as a Genre of Qualitative Research: Theory, Methodology, and Practice*. Springer.
- Priestley, M., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Publishing.
- Reay, D., 2012. What would a socially just education system look like? Saving the minnows from the pike. *J. Educ. Pol.* 27, 587–599. <https://doi.org/10.1080/02680939.2012.710015>.
- Sachs, S., 2001. Teacher professional identity: competing discourses, competing outcomes. *J. Educ. Pol.* 16 (2), 149–161. <https://doi.org/10.1080/02680930116819>.
- Sachs, J., 2005. Teacher education and development of professional identity: learning to be a teacher. In: Denicolo, P.M., Kompf, M. (Eds.), *Connecting Policy and Practice: Challenges for Teaching and Learning in Schools and Universities*. Routledge, pp. 5–21.
- Schuelka, M., et al., 2019. Introduction: scholarship for diversity and inclusion in education in the 21st century. In: Schuelka, M.J., Johnstone, C.J., Thomas, G., Artiles, A.J. (Eds.), *The Sage Handbook of Inclusion and Diversity in Education*. Sage, pp. xxi–xlili.
- Skúlason, P., 1989. Pælingar II [Reflections]: Safn Erinda Og Greinastúfa [Collection of Lectures and Short Articles] (Reykjavík).
- Tidwell, D., Staples, A., 2017. The collaborative process in educators' self-study of practice. In: Dalmau, M.C., Guðjónsdóttir, H., Tidwell, D. (Eds.), *Taking a Fresh Look at Education*. Brill Sense, pp. 89–111.
- UNESCO, 2015. *Education 2030: Incheon Declaration and Framework for Action*. <http://www.uis.unesco.org/Education/Documents/incheon-framework-for-action-en.pdf>.
- UNESCO, 2020. *Global Education Monitoring Report 2020: Inclusion and Education: All Means all*. <https://unesdoc.unesco.org/ark:/48223/pf0000373718>.
- University of Iceland's Code of Ethics, 2014. *Visindasiðareglur Háskóla Íslands*. https://www.hi.is/sites/default/files/atli/pdf/log_og_reglur/vsih_sidareglur_16_1_2014.pdf.
- Van Manen, M., 1991. *The Tact of Teaching: The Meaning of Pedagogical Thoughtfulness*. State University of New York Press.
- Van Manen, M., 1999. The language of pedagogy and the primacy of student experience. In: Loughran, J.J. (Ed.), *Researching Teaching: Methodologies and Practices for Understanding Pedagogy*. Falmer Press, London, pp. 13–27.
- Wolcott, H., 2001. *Writing Up Qualitative Research*. Sage Publications.

Learning to teach epistemic practices of science in secondary schools

Sibel Erduran and Liam Guilfoyle, University of Oxford, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	254
Epistemic practices in science education	255
Teachers' epistemologies and epistemic practices of science	256
Argumentation as an epistemic practice: teaching in different domains of knowledge	257
Nature of argumentation in science versus other domains	257
OARS project: argumentation in different school subjects	258
Science versus RE teachers' views of argumentation	258
Conclusions and implications	260
Acknowledgment	261
References	261

Introduction

Historically, science curricula around the world have emphasized the learning of science content knowledge. Science teaching often positions students as receivers of “final forms of scientific knowledge” (Duschl, 1990) without engaging learners in how such knowledge has been generated and why this knowledge is credible. For example, how do we know that matter is made of tiny, indivisible particles? Why do we believe that the planet earth rotates around the sun? Delivering unquestionable “facts” to students without engaging them in evidence-based reasoning is essentially akin to dogma, and as such, is far from the scientific worldview that values reason and evidence. The aspects of science that emphasize knowledge production processes are related to what is often considered as “epistemic practices” of science.

Epistemic practices thus defined relate to those activities that help mediate the development of scientific knowledge which is characterized by theories, models and laws (Erduran and Dagher, 2014). They can include processes such as modeling and argumentation. Argumentation can be defined as the justification knowledge claims with evidence (Driver et al., 2000; Erduran and Jiménez-Aleixandre, 2007; Jimenez-Aleixandre et al., 2006) and includes variety of discursive practices such as “*assessing alternatives, weighing evidence, interpreting texts, and evaluating the potential viability of scientific claims*” (Driver et al., 2000, p. 288). Models, on the other hand, are forms of knowledge that can be defined as an abstract representation of reality and can be “*an idea, object, event, process, or system*” (Justi and Gilbert, 2000, p. 994). Modeling involves a range of practices including data aggregation and representation. All forms of scientific knowledge essentially provide explanations to account for how and why particular natural and physical phenomena occur (Chinn and Brown, 2000; Nagel, 1961).

Although historically, epistemic practices of science have been neglected in science curricula, more recent curriculum standards document have included references to the significance of learning epistemic aspects of science. Consider, for instance, the Framework for K-12 Science Education developed by the National Research Council (NRC, 2012) of National Academy of Sciences designed in order to create standards of science education. The Framework is based on evidence-based findings of current research and introduced three major dimensions of science education. Scientific practices were recommended as one of them and defined as “(a) the major practices that scientists employ as they investigate and build models and theories about the world” (NRC, 2012, p. 30). National Research Council (NRC, 2012) specified eight practices that propose the necessary components of science curriculum. These are:

- Asking questions (for science) and defining problems (for engineering);
- Developing and using models;
- Planning and carrying out investigations;
- Analyzing and interpreting data;
- Using mathematics and computational thinking;
- Constructing explanations (for science) and designing solutions (for engineering);
- Engaging in argument from evidence;
- Obtaining, evaluating, and communicating information (p. 49).

Although advancement in curriculum reform have thus advanced epistemic practices as learning outcomes, it is not entirely clear how teachers may teach epistemic practices in their lessons. Curriculum revision by itself may not be enough for ensuring that students are taught for understanding and experiencing of epistemic practices. Teachers as the major actors of learning environment should also have an understanding of epistemic practices. Considering most teachers' own learning experiences would have preceded such recent reform initiatives, the question remains as to how teachers can be supported in their learning to teach epistemic practices of science.

The purpose of this article is thus to explore how teachers' learning of epistemic practices of science can be supported. The particular focus will be on argumentation in science as an epistemic practice. Following a review of the literature on epistemic practices including examples of argumentation, modeling and explanations, we will thus focus on teachers' learning of argumentation. An important element of teachers' understanding of argumentation *in science* as an epistemic practice involves how they might potentially distinguish it from argumentation in other subject domains. In other words, given a broad set of subjects can be considered from an argumentation perspective (e.g., religious education, history, English), what differentiates argumentation as an epistemic practice of science? We will address this question by contrasting how argumentation is framed broadly in epistemic practices of two contrasting school subjects: science and religious education. Following this theoretical account, we will then draw on some empirical data from a funded project based in England to illustrate how in-service science teachers view argumentation and engage in its teaching.

Epistemic practices in science education

Epistemic practices in science could involve a range of knowledge production and evaluation processes such as coordination of theory and evidence; and making sense of patterns in data (Kelly and Licona, 2018). Epistemic practices in science education refer to those cognitive and discursive activities that engage learners in the knowledge construction processes of science. Teachers and students may have their own personal epistemologies related to a range of subjects as well (Hofer, 2001). Conventionally in school science, students' engagement in epistemic practices of science and in argumentation in particular has been scarce, and science teachers' own background in learning to teach epistemic practices tends to be fairly limited (Kelly and Licona, 2018). Despite years of reform across the world to improve science teaching and learning, the epistemic aspects of science teaching and learning continue to be of concern to many science education researchers (e.g., Chinn and Maltotra, 2002; Jimenez-Aleixandre et al., 2006; Sandoval and Reiser, 2004).

Apart from an explicit focus on epistemic practice, several other lines of research in science education are concerned with epistemic practices of science. These include argumentation (Erduran et al., 2019), modeling (Justi and Gilbert, 2000) and explanations (Chinn and Brown, 2000). Argumentation can be defined as the justification knowledge claims with evidence (Jimenez-Aleixandre et al., 2006) and includes variety of discursive practices such as "*assessing alternatives, weighing evidence, interpreting texts, and evaluating the potential viability of scientific claims*" (Driver et al., 2000, p. 288).

Models and modeling literature in science education led to substantial amount of insight into how learners construct knowledge representations. Models are forms of knowledge that can be defined as an abstract representation of reality and can be "*an idea, object, event, process, or system*" (Justi and Gilbert, 2000, p. 994). Modeling enables an unknown reality (target) to be explained with a more familiar reality (source) (Duit and Glynn, 1996). By focusing on specific aspects of a phenomenon, models simplify the phenomenon (Duschl and Erduran, 1996; Ingham and Gilbert, 1991). Different forms of knowledge such as models offer explanations which describe how and why something happens (Chinn and Brown, 2000). Typically, scientists explain phenomena by determining how and why they occur along with the conditions surrounding the observed events (Nagel, 1961). Explanations can be articulated relative to disciplinary conceptual frameworks (Dagher and Erduran, 2014) highlighting variations in how explanations are defined in different fields of science such as biology and chemistry.

Apart from their epistemic dimensions, lines of research on argumentation, modeling and explanations may involve other aspects as well. For instance, argumentation is not only an epistemic practice but also it can be framed from a linguistic and social perspective, which underpin the work of many science educators working on argumentation (Erduran et al., 2015). In our previous work, we have thus developed a heuristic based on a review of literature on philosophy of science (Erduran and Dagher, 2014) that consolidates the cognitive, epistemic and social aspects of practices in science. The heuristic called the Benzene Ring Heuristics (BRH) includes epistemic components such as scientific activities, data, model, explanation and prediction (see Fig. 1). Cognitive and social components respectively refer to representations and reasoning and social dissemination and certification of scientific claims.

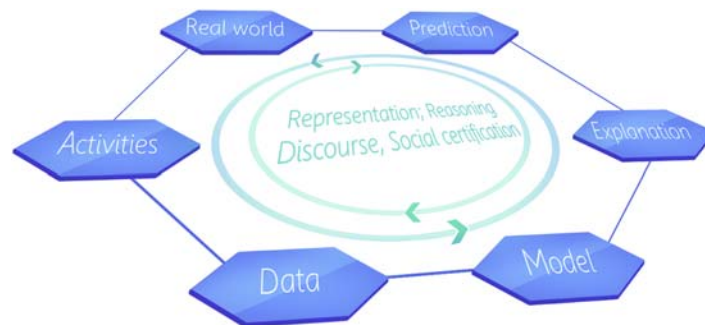


Fig. 1 "Benzene Ring" heuristic of scientific practices. From Erduran and Dagher (2014, p. 82).

BRH suggests that all features of scientific practices are related to each other and these relationships do not have a linear order. In other words, holistic understanding of scientific practices was emphasized with the representation referring to an analogy that is Benzene Ring. In this heuristic, components of scientific practices such as modeling are related to the social components (representation, reasoning, discourse, social certification) in a dynamic and interactive manner. The heuristic emphasizes modeling of data generated through a set of activities such as experimentation, observation and classification. Modeling serves to explain and predict the phenomena under investigation.

The models help explain different scientific concepts. For example, the atomic model can help explain concepts such as electrons, protons and neutrons. Furthermore, it illustrates the connections between these concepts. In the example of the atomic model, the Bohr model of the atom can help predict chemical bonding between different atoms and explain chemical properties. Argumentation is a mediating process that interlinks the epistemic and cognitive aspects of science such as prediction and explanation. For example, atomic structures yielding different molecular combinations can be argued about. Various stages of modeling, predicting and explaining can involve argumentation or the justification with evidence of the claims being made.

Teachers' epistemologies and epistemic practices of science

Incorporating frameworks such as the Benzene Ring Heuristic (see Fig. 1) in science teaching and learning requires that teachers are familiar not only with the underlying concepts such as models and explanations but also the epistemic practices through which these concepts are generated, revised and evaluated. In other words, there are demands on teachers to interpret and teach epistemic practices of science. However, teachers themselves bring with them to the science lessons their own personal epistemologies. Arguments for the importance of teachers' understanding of the epistemic dimensions of their subject are not new. Almost half a century ago, Schwab (1962) had argued that expertise in teaching requires both knowledge of content of a domain and knowledge about the epistemology of that domain. Teachers develop the necessary capability of transforming subject into teachable content only when they know how the disciplinary knowledge is structured (Erduran and Kaya, 2019). For example, a chemistry teacher who understands why and how models contribute to chemical knowledge will be better able to facilitate students' understanding of the role of models in chemistry.

Such understandings are often referred to in educational psychology and science education as *personal epistemologies* or *epistemic beliefs* (Hofer, 2016). There are many other terms, but in most instances the core idea refers to a set of beliefs about the nature of knowledge and nature of knowing. For example, Hofer (2000) refers to these beliefs as consisting of a set of beliefs about the certainty of knowledge (whether knowledge is fixed and stable or tentative and evolving over time), simplicity of knowledge (whether knowledge is an accumulation of isolated facts or an interconnected web), justification for knowing (beliefs about how we come to know and justificatory standards for saying we know something) and source of knowledge (whether we source knowledge from the testimony of others such as experts or from within through perception, reasoning, memory, etc.).

Significant effort has gone into researching the current landscape of teachers' epistemic beliefs, their development by natural means and through intervention, and their implications for teacher education. There are at least 3 ways to consider the influence of epistemic beliefs and personal epistemologies on teaching and learning. These are the influences on (1) students' learning (2) teachers' learning and (3) teachers' teaching. There is currently sufficient evidence in the research literature to consider epistemic beliefs to be particularly worthy of attention in teacher education. In a synthesis of the literature on personal epistemology in learning mathematics, Muis (2004) found that all 33 studies in the review demonstrated significant relationships between personal epistemology and other relevant factors to teaching and learning, such as motivation and academic achievement. This resonates with findings that epistemic beliefs can be influential for one's own strategies toward learning (Windschitl and Andre, 1998; Ravindran et al., 2005; Bondy et al., 2007; Bromme et al., 2008), critical thinking (Bråten and Strømsø, 2006), motivation (Beghetto and Baxter, 2012), metacognition (Yenice, 2015), views on the Nature of Science (Yenice, 2015), and sense of self-competence (Di Martino and Zan, 2011; Roesken et al., 2011).

Teachers' epistemic beliefs have been associated with their understanding of effective pedagogy (Yadav and Koehler, 2007), their own self-efficacy in teaching (Chrysostomou and Philippou, 2010), their own learning strategies (Peng and Fitzgerald, 2006), and their support of student autonomy (Weinstock and Roth, 2011). These beliefs can also mediate how those who are learning to teach judge their own success in their teaching. Yilmaz-Tuzun and Topcu (2008) found, for example, that pre-service science teachers felt successful in their science teaching when students memorized facts, if they held an epistemic belief that scientific knowledge is an accumulation of unchanging facts. Furthermore, epistemic beliefs have been found to be influential for a teacher's conception of teaching and learning (Wong et al., 2009; Yilmaz and Şahin, 2011), their preference in teaching practices (Windschitl et al., 2008; Chai and Khine, 2008; Çetin-Dindar et al., 2014) and their actual teaching practices (Kang, 2008).

Despite the rather extensive quantities of research that ultimately support Schwab's (1962) argument that expertise in teaching requires both knowledge of content of a domain and knowledge about the epistemology of that domain, it is often found that teachers may have underdeveloped epistemic beliefs in their subject domain (e.g., Kang and Wallace, 2005; Kang, 2008; Liu and Tsai, 2008; Markic and Eilks, 2012; Topcu, 2013). In such a context, the question for teacher education is how to support the development of epistemic beliefs in pre-service and in-service teachers. The field has generated research on this topic too—focusing both on understanding the natural development of trainee teachers during their initial teacher education and the influence of interventions for epistemological development. Walker et al. (2012), reporting a longitudinal study of epistemological development in

teacher education, noted a number of factors that appeared to “play a role” in development. For example, teaching experience on school placement, increasing subject knowledge, engaging in reflective practice, and encountering contradictions in theory or opinion in their initial teacher education (ITE) program. Guilfoyle (2018), in another longitudinal study, pointed to the research project component of the ITE program as having potential for epistemological development within the discipline being researched. Group discussions and explicit written reflection on one’s own epistemic beliefs may also be helpful in developing epistemic beliefs (Brownlee et al., 2001; Baxter Magolda, 2002), particularly those including comparison of multiple perspectives (Marra and Palmer, 2011). Such pedagogical interventions are most effective, it is suggested, where PSTs can “defend their opinions in response to counter-opinions ... [and] consider counter-claims” (Topcu, 2011, p. 117). Related to this, Iordanou et al. (2016) suggest that epistemic beliefs develop “in sync” with argumentative reasoning and so argumentation activities may aid in the development of epistemic understanding (Iordanou, 2010). Modeling intervention activities, too, have been associated with the development of epistemological development in science teacher education (Windschitl et al., 2008). We focus our attention more closely on these two elements, argumentation and modeling, to draw out more specific implications for teacher education and the relation with learning to teach epistemic practices in secondary science classrooms.

Argumentation as an epistemic practice: teaching in different domains of knowledge

Nature of argumentation in science versus other domains

It has long been recognized that argumentation can have features that are similar across a range of fields, domains or disciplines (field-invariant) and features that are specific to the arguments of particular fields (field-dependent) (Toulmin, 1958). For instance, a field invariant feature may be a structure involving the coordination of information (facts, concepts, principles, etc) to create claims that are supported by evidence, where the relevance of the evidence to the claim is established by means of a warrant. But field-dependence may emerge in differences in the nature of the claims, evidence, and warrants that are deemed appropriate for a particular disciplinary context and these represent important epistemological principles for that discipline. We introduced argumentation, above, with respect to science specifically. Now, we will explore how the nature of argumentation in science compares with the nature of argumentation in other domains in the secondary school context.

Goldman et al. (2016, 2018), as part of a wider framework for disciplinary literacy, compared the epistemology, argumentation, and inquiry practices across three disciplines (i.e., science, history and literature). Their theoretical work serves as a useful tool for our current purpose. They tease apart the similarities and differences between these domains by looking for distinctions in common features. For example, while each domain is concerned with advancing claims of some sort, the nature of these claims often distinguishes them. In the claims advanced in study of literature are about personal relevance, structure, rhetoric, theme, and connections to historical, political, ideological, or theoretical positionings.

The evidence drawn upon will be segments of text as representations or illustrations of patterns, messages, structures, etc. The warrants may be rooted in personal beliefs, life experiences, theory, or philosophical tenets. In the study of history, the claims that are advanced are interpretations of the past based on available evidence of the historical record. They draw on primary sources such as documents like diaries, letters, census data, government documents, maps, photographs, video or other physical artifacts. Evidence may in some instances come from secondary or tertiary sources too. There is an acknowledge of the perspective of these sources or the that the historical record may in many cases be incomplete. Historians can use a range of interpretative frameworks in warranting their arguments. For example, they can generate interpretations in terms of societal systems, government systems, relations between phenomena (e.g., cause-and-effect or chance), or particular themes (e.g., industrialization, migration, etc).

In all instances the interpretation must account for all the remnants of history. That is to say, one cannot ignore a piece documentary evidence in advancing an argument because it does not fit. In science, claims are generally about the nature and operation of phenomena in the natural world. These claims may be in the form of explanations, theories, laws, models, or predictions about these phenomena or the relationships between phenomena. Evidence is generally that which can be gathered or generated through reliable processes that can be corroborated. Evidence is represented in a variety of forms, including but not limited to tables, graphs, equations, diagrams, models, simulations and written narratives. Warrants are provided through principles that are enduring in the field.

Religious Education (RE), as a school subject, provides some interesting scope for developing an understanding of epistemic practices across a range of disciplines as they are view in together or in comparison. RE in England, as in other jurisdictions globally, is a multi-disciplinary study of religions, beliefs, and worldviews (Fancourt, 2015). It is generally conceived as being pluralistic with the ambition to foster tolerance, understanding, respect for varying religious beliefs and none (Jackson, 2004). Studying religions and worldviews like this will naturally draw on philosophy, history, sociology, psychology, and other cognate disciplines (Freathy and John, 2019). Natural science also features in RE both in terms being considered a “worldview” (the scientific worldview) but also how science has something to say about some of the claims made by religions like, for example, creation. This “melting pot” of disciplinary perspectives provides a rather rare opportunity to emphasize disciplinary differences in epistemic practices. Though, if learning to teach one subject area of substantive and syntactic knowledge (Schwab, 1962) weren’t enough, a practitioner would need to understand the syntactic knowledge of these each of these cognate disciplines in order to facilitate students’ development of these epistemic understandings. This is where the strength of a collaborative approach can come to the fore. Next, we will discuss a funded project in the University of Oxford, which brings together both science and RE teachers from the same schools to focus on interdisciplinary and cross-curricular collaboration for the teaching and learning of argumentation.

OARS project: argumentation in different school subjects

The Oxford Argumentation in Religion and Science (OARS) project is a three-year research and development project funded by the Templeton World Charity Foundation's Big Questions in Classrooms initiative (Chan and Erduran, 2022; Chan et al., 2020; Erduran et al., 2020; Erduran et al., 2019; Godfrey and Erduran, 2022; Guilfoyle and Erduran, 2021; Guilfoyle et al., 2021). The OARS project brings together teachers and researchers from science and RE to focus on the teaching and learning of argumentation within and across these subjects. A Continuous Professional Development (CPD) program, consisting of six workshops, was organized by researchers and teacher educators at the Department of Education, University of Oxford. Pairs of science and RE teachers were invited from schools within a reasonable distance of the University to attend this program of evening sessions to think, talk, and learn about argumentation in their own subjects and in the "other" subject.

The CPD focused on supporting students in learning how to make complex judgments in science, religious education and, importantly, in the issues where these subjects may overlap (Erduran, 2020). For example, in discussing the origins of the universe and life, or whether or not genes should be cloned. Such skills are important for achieving curriculum goals in both subjects (Chan et al., 2020; Erduran et al., 2019) but also for nurturing children as future citizens with the capacity to have productive conversations across divides of opinion and discipline. In sessions, we recognize the distributed expertise (Childs et al., 2014) where each individual in attendance is bringing worthwhile expertise to the collective learning. Of the 30 teachers who were participating, many were highly experienced practitioners, and some had a Ph.D. in their content area. As such, participating teachers were often in a position to significantly contribute to their peers' learning within and outside their subject area, and within and outside their school contexts. This was a key advantage of the collaborative process of the CPD project. A natural in-school peer community was created by virtue of the fact that teachers were coming to the project as pairs, or sometimes small groups, from their school, and coming together during CPD sessions to form a wider community with other schools (Vangrieken et al., 2017). In the next section, we will draw on qualitative data from interviews with OARS project teachers to highlight the teachers' views of argumentation and some of their expectations of engaging in collaborative work between teachers of science and RE. These serve to illustrate some of the potential of engaging in this work and raise some questions about the challenges involved in teachers' learning to teach epistemic practices.

Science versus RE teachers' views of argumentation

The collaborative element of the OARS project opens conversations between RE teachers and science teachers. Not each expression by the teachers' is necessarily true or accurate, particularly as they characterize the "other" subject, but these characterizations could often lead to challenge by other teachers. This process, ultimately a discursive argumentation exercise with the common goal of ascertaining the nature of argumentation for their subjects, drew out a deeper thinking about epistemic practices and criteria. Take, for example, the following quote.

I think listening to the RE teachers at the sessions, we were, you know, trying to **pin each other down on what we mean by evidence**. And, of course, coming from a scientific perspective, having to sort of broaden that out into things that could include doctrines and life experience and stuff like that.

School B—Science.

Here we see a science teacher reflecting on their scientific perspective on evidence because of an interaction with RE teachers about what counts as evidence for arguments in their subject. There is not yet a conclusion here, but it raises questions in their mind about what other disciplines consider acceptable evidence. Such thoughts about the epistemic criteria in science might not otherwise be elucidated but for the conversations with RE teachers. It alerts the science teacher, at least, to the fact that students are exposed to arguments with different justificatory standards in other areas of the curriculum. The notion of "evidence" is not unique to science. Instead, this teacher feels the need to "broaden that out", and instead think about the standards of acceptability of evidence.

When talking about how argumentation features within their classes, some RE teachers highlight how they need to engage with epistemic criteria about what counts for evidence for knowing (addressing Hofer's justification and source dimensions of epistemic beliefs).

I was talking with one group last week about life after death. And one of them said, "but there's no evidence!". And then we looked at sort of personal experiences, and they went, "well, that's not evidence". And it's **that what will you and won't you accept as evidence** and personal experience ... I said, "Well, if I came in tomorrow, and told you something, what would you say?" And they sort of laughed at me and said, "We probably think you had too much wine".

School B—RE.

It is worth noting that even though RE teachers discuss and recognize the wider forms of evidence invoked in arguments of the cognate disciplines or of religious individuals, they often held aspirations that their students' arguments within their subject would be strongly substantiated.

Our pupils choose evidence that's quite woolly. They will often prefer to describe beliefs, which they think they can describe more freely than facts. I think that if our pupils were more factual within the describing language that would be an improvement for argument.

School F—RE.

But RE teachers emphasized that while there is differing evidence acceptable within their subject, it is not solely evidence that is the marker of a quality argument, but also the development and structure of the argument.

There's differences in the type of evidence you use. So for instance, if you're having an ethical argument, you are **not relying on necessarily on proven fact**, and researched information, you're **relying on logical development** in arguments. Whereas there perhaps isn't an opportunity within science Similarly, in RE, there's opportunities for using I would say "proper" evidence. But I think it's weighted much more heavily on one than the other. I guess.

School C—RE.

This sort of focus on the structure and development of arguments in the form of warranting (i.e., establishing the link between evidence and claim) was something that some science teachers felt was lacking in their subject when they compared it to RE.

So it's just the argumentation is actually very different. It's sort of "here's the evidence". Well, there you go, then. It's sort of ... there's not [in RE teacher]'s terms, there's not much in terms of the warrant. In other words, how well does the evidence relate to the claim? That's not well taught in science. It is pretty much: "Here's the claim and here's the evidence that shows it QED".

School R—Science.

Some science teachers wondered if the sorts of arguments they engaged in within science really "counted" as arguments at all, as they felt there was a lack of a controversial nature to the claims involved and that there exists a correct, verifiable answer.

So I was saying earlier, so we think substance A is ionic, because it's got a high melting point, because it conducts when it's in solution. So the idea of giving reasons, from experimental evidence for saying the thing you're saying. So we have quite a lot of opportunity for that. But that doesn't feel to me, like really, what we're talking about here in terms of framing an argument.

School C—Science.

This idea that arguments in science, or at least school science, are "unwarranted" seems to relate to the ideas about the accessibility or complexity of scientific knowledge for this age group. There is a sense that science needs to be simplified to be understood, and there for it needs to be accepted ("source" beliefs, Hofer, 2000) with little provision of the link between claim and evidence (justification beliefs).

So, there might be, you know, a page in the textbook on different views of the atom and how that's changed over time. But yeah, right now, it's pretty much as much as you get. Whereas some quite important things are based on some quite complex arguments, currently. And it is simplified to be understandable, but you lose that sense of the link between evidence and why you believe things.

School R—Science.

This view of school science can lead to an epistemic message that science is effectively an accumulation of knowledge (simplicity beliefs) that is unquestionable (certainty beliefs) and is received from authority (source beliefs).

Science deals with facts, and it's really the regurgitation of those facts in a kind of simple formula. And that's what science is until you are at University, probably beyond your undergraduate. You're basically learning the kind of bedrock of stuff.

School M—RE.

However, this idea seems to run counter to beliefs of teachers about what school, and the teaching and learning of argumentation, ought to achieve. The following pair of teachers from School C were in agreement that emphasizing the skills of argumentation in class was about getting students to understand what was necessary to warrant a claim and to be able to assess the presence or absence of an appropriate warrant for the claims they were faced with in life. While this is clearly about the justification of knowledge, it also underscores beliefs about the source of knowledge in not blindly receiving knowledge from an authority source but building a capacity to self-assess arguments.

I would have to say “Is this a logically ...” ... Does that evidence actually backup that claim? Because we get a lot of claims ... And that’s a really valuable thing, I think, for students to take away from their education, because then they’re presented with a vast number of claims, and mostly through social media, there’s a huge number of completely invalid claims. Supposedly, with evidence and yet it doesn’t back it up.

School C—Science.

And being able to see that the evidence that is being quoted is irrelevant to the claim that is being made ... that skill of spotting it, I think, is a skill. I think for a lot of people they are just kind of nodding when they hear things.

School C—RE.

This is echoed by other science teachers who say that “*it’s one of the main reasons I want students to learn science*” to “*hopefully provide them with skills to not take everything for granted*” (School G—Science). There is a real sense of wanting to develop some general epistemic beliefs about the source of knowledge being more nuanced than just from authority (“*scrutinize the conclusion that being presented to you ... in the era of fake news and spurious arguments, made by politicians*”, School L—Science). But given the tension of simplification, science may struggle to do this (“*I don’t think we’re very successful to do it*”, School L—Science).

So [there] does come a point where we just have to trust arguments or information to some extent, if you’re not an expert in that particular topic area, and how does one make a judgment about that, I suppose.

School N—Science.

Science teachers at times seemed surprised to note from RE teachers just how much justification and argumentation processes occurred in RE classrooms (“*There’s a lot more of that in your subject [to RE teacher]. Than I had realized*”, School A—Science). RE teachers often echoed many of the same goals for their students in terms of learning critical thinking skills to dissect the arguments that they encounter in their lives, beyond the content area of the subject. There is a real sense from the teachers that they are working to the same ends:

Being able to read something, or to listen to something and then to be able to express their ... and uncover the hidden assumptions in the line of argument, whether that was a bit of fallacious reasoning or something like that. That’s ultimately taking the skills base that we work on. And not restricting it just to what we think about [in the subject] but to other things as well. Take it to the real world.

School L—RE.

These quotes provide an illustration of the early stage conversations and thoughts that occurred in the collaboration process. There are numerous rather promising indications in these early remarks. First, is that the collaboration between teachers seemed to open up a dialog (through and about argumentation) that promoted thinking and reflection on their own epistemic beliefs of their subject area. This is done in the context of teaching and learning of students, so it becomes directly applicable to their teaching of the syntactic knowledge of their subject or discipline. Secondly, these teachers began to gain an understanding of each other’s subject areas. Strengthening this understanding means that teachers are more aware of what their students experience in others subjects and what arguments or forms of evidence they are exposed to. This understanding could be capitalized upon to make clearer the similarities and differences between subjects. Finally, teachers uncovered common underlying goals for students to be critical thinking citizens who can manage in a “post-truth” world surrounded by “fake news”. This can highlight the benefit of working together to achieve these goals. In fact, in doing so, students may have a more nuanced capacity to deal with problems in the real world if they have the ability to discern the arguments of different disciplines apart in multi- or inter-disciplinary contexts. Cross-curricular work can then have the potential to strengthen multi- and inter-disciplinary thinking.

Conclusions and implications

The article provided an overview of some key epistemic practices of science including modeling and argumentation. Following a review of literature on teachers’ own personal epistemologies, we presented data from a funded research project that aimed to incorporate argumentation in science and religious education teachers’ lessons. The data illustrated that teachers of both subjects value argumentation as a component of their subject. Furthermore, they were keen to exchange professional experiences and knowledge with teachers of the other subject. Engagement in collaborative dialog can expose teachers to how argumentation may be situated in not only their own subject but also in a different school subject, potentially resulting, through a comparative analysis, in a deeper understanding of the epistemic practices of their own subject including its particularities. Teachers’ engagement in other epistemic practices such as modeling across school subjects may also provide fruitful in fostering understanding of nuances of disciplinary characteristics of knowledge.

Although some epistemic practices such as argumentation may provide points of conjecture between different school subjects, there may also be fundamentally different components to the epistemology of the subject. For instance, how data and modeling are used in science may not necessarily apply in a subject such as RE. Further studies are needed to elaborate on how subject epistemologies might compare across a whole range of aspects. Theoretical accounts on comparative analyses of subject epistemologies can help clarify what epistemic practices can be prioritized in teachers' learning. Even within a particular school subject such as science, there may be domain-specific ways of thinking and reasoning. Consider, for example a contrast of epistemological features of biology and chemistry. When the types of explanation in chemistry and biology are contrasted, the result is a diversity of types that are distinctive to the science in question (Dagher and Erduran, 2014). While biological explanations include viability and developmental explanations that draw closely from the nature of biological content, chemical explanations focus on the structural and representational explanations that are based on either quantum mechanical or simple chemical models. The context of debates around the nature of biological and chemical laws and explanations are also rather particular. Whereas reference to principles is common in biologists' discussions of laws, the chemists are preoccupied with questions regarding axiomatization and approximation (Dagher and Erduran, 2014).

Finally, there are implications for not only continuous professional development programmes of science teachers and pre-service teacher education but also curriculum reform. Our discussion suggests that teacher education programmes to benefit from a systematic approach to representing the epistemic practices of school subjects. A systemic endorsement of the value of subject epistemology across the whole school curriculum is likely to improve the likelihood of uptake of this approach in actual lessons. The epistemological aspects of school subjects can be used to empower both teachers and students in understanding how and why knowledge develops in different lines of enquiry (Erduran et al., 2007).

Acknowledgment

The research reported in this paper has been supported by funding from the Templeton World Charity Foundation (Grant Number TWC0238).

References

- Baxter Magolda, M.D., 2002. Epistemological reflection: the evolution of epistemological assumptions from age 18 to 30. In: Hofer, B., Pintrich, P.R. (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*. Lawrence Erlbaum, Mahwah, NJ.
- Beghetto, R.A., Baxter, J.A., 2012. Exploring student beliefs and understanding in elementary science and mathematics. *J. Res. Sci. Teach.* 49 (7), 942–960.
- Bondy, E., Ross, D., Adams, A., Nowak, R., Brownell, M., Hoppey, D., Kuhel, K., McCallum, C., Stafford, L., 2007. Personal epistemologies and learning to teach. *Teach. Educ. Spec. Educ.* 30 (2), 67–82.
- Bråten, I., Strømso, H.I., 2006. Effects of personal epistemology on the understanding of multiple texts. *Read. Psychol* 27 (5), 457–484.
- Bromme, R., Kienhues, D., Stahl, E., 2008. Knowledge and epistemological beliefs: an intimate but complicate relationship. In: Erduran, S., Jimenez-Aleixandre, M.P. (Eds.), *Argumentation in Science Education*. Springer, The Netherlands, p. 423441.
- Brownlee, J., Purdie, N., Boulton-Lewis, 2001. Changing epistemological beliefs in pre-service teacher education students. *Teach. High. Educ.* 6 (2), 247–268.
- Çetin-Dindar, A., Kurbulut, Z.D., Boz, Y., 2014. Modeling between epistemological beliefs and constructivist learning environment. *Eur. J. Teach. Educ.* 37 (4), 479–496.
- Chai, C.S., Khine, M.S., 2008. Assessing the epistemological and pedagogical beliefs among pre-service teachers in Singapore. In: Khine, M.S. (Ed.), *Knowing, Knowledge and Beliefs: Epistemological Studies Across Diverse Cultures*. Springer, The Netherlands, pp. 287–299.
- Chan, J., Fancourt, N., Guilfoyle, L., 2020. Argumentation in religious education in England: an analysis of locally agreed syllabuses. *Br. J. Relig. Educ.* 1–14. <https://doi.org/10.1080/01416200.2020.1734916>.
- Chan, C., Erduran, S., 2022. The impact of collaboration between science and religious education teachers on their understanding and views of argumentation. *Res. Sci. Educ.* <https://doi.org/10.1007/s11165-022-10041-1>.
- Childs, A., Edwards, A., McNicholl, J., 2014. Developing a multi-layered system of distributed expertise: what does cultural historical theory bring to understandings of workplace learning in school-university partnerships? In: McNamara, O., Murray, J., Jones, M. (Eds.), *Workplace Learning in Teacher Education: International Practice and Policy*. Springer Netherlands, Dordrecht, pp. 29–45.
- Chinn, C., Brown, D.E., 2000. Learning in science: a comparison of deep and surface approaches. *J. Res. Sci. Teach.* 37 (2), 109–138.
- Chinn, C.A., Malhotra, B.A., 2002. Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Sci. Educ.* 86, 175–218.
- Chrysostomou, M., Philippou, G.N., 2010. Teachers' epistemological beliefs and efficacy beliefs about mathematics. *Proc. Soc. Behav. Sci.* 9, 1509–1515.
- Dagher, Z., Erduran, S., 2014. The role of disciplinary knowledge in science education: the case of laws and explanations in biology and chemistry. In: Matthews, M. (Ed.), *Handbook of Research on History, Philosophy and Sociology of Science*. Springer, Dordrecht, pp. 1203–1234.
- Di Martino, P., Zan, R., 2011. Attitude towards mathematics: a bridge between beliefs and emotions. *ZDM* 43 (4), 471–482.
- Driver, R., Newton, P., Osborne, J., 2000. Establishing the norms of scientific argumentation in classrooms. *Sci. Educ.* 84 (3), 287–312.
- Duit, R., Glynn, S., 1996. Mental modeling. In: Welford, G., Osborne, J., Scott, P. (Eds.), *Research in Science Education in Europe: Current Issues and Themes*. Falmer Press, Bristol, PA, pp. 166–176.
- Duschl, R., Erduran, S., 1996. Modeling the growth of scientific knowledge. In: Welford, G., Osborne, J., Scott, P. (Eds.), *Research in Science Education in Europe: Current Issues and Themes*. Falmer Press, London, pp. 153–165.
- Duschl, R., 1990. *Restructuring science education: the importance of theories and their development*. Teachers College Press, New York.
- Erduran, S., Dagher, Z., 2014. *Reconceptualizing the Nature of Science for Science Education: Scientific Knowledge, Practices and Other Family Categories*. Springer, Dordrecht, The Netherlands.
- Erduran, S., Jiménez-Aleixandre, M.P. (Eds.), 2007. *Argumentation in Science Education: Perspectives from Classroom-Based Research*. Springer, Netherlands.
- Erduran, S., Kaya, E., 2019. *Transforming Teacher Education Through the Epistemic Core of Chemistry: Empirical Evidence and Practical Strategies*. Springer, Dordrecht.
- Erduran, S., Aduriz-Bravo, A., Mamluk-Naaman, R., 2007. Developing epistemologically empowered teachers: examining the role of philosophy of chemistry in teacher education. *Sci. Educ.* 16 (9–10), 975–989.
- Erduran, S., Ozdem, Y., Park, J.Y., 2015. Research trends on argumentation in science education: a journal content analysis from 1998 to 2014. *Int. J. STEM Educ.* 2 (5).

- Erduran, S., Guilfoyle, L., Park, W., Chan, J., Fancourt, N., 2019. Argumentation and interdisciplinarity: reflections from the Oxford argumentation in religion and science project. *Discip. Interdiscip. Sci. Educ. Res.* 1 (1).
- Erduran, S., Guilfoyle, L., Park, W., 2020. Science and religious education teachers' views of argumentation and its teaching. *Res. Sci. Educ.* <https://doi.org/10.1007/s11165-020-09966-2>.
- Erduran, S., 2020. Argumentation in science and religion: match and/or mismatch when applied in teaching and learning? *J. Educ. Teach.* 46 (1), 129–131.
- Fancourt, N., 2015. Re-defining 'Learning about Religion' and 'Learning from Religion': a study of policy change. *Br. J. Relig. Educ.* 37 (2), 122–137.
- Feathery, R., John, H.C., 2019. Religious education, big ideas and the study of religion(s) and worldview(s). *Br. J. Relig. Educ.* 41 (1), 27–40.
- Godfrey, H., Erduran, S., 2022. Argumentation and intellectual humility: a theoretical synthesis and empirical study about students' warrants. *Res. Sci. Educ.* <https://doi.org/10.1080/02635143.2021.2006622>.
- Goldman, S.R., Britt, M.A., Brown, W., Cribb, G., George, M., Greenleaf, C., et al., 2016. Disciplinary literacies and learning to read for understanding: a conceptual framework for disciplinary literacy. *Educ. Psychol.* 51 (2), 219–246. <https://doi.org/10.1080/00461520.2016.1168741>.
- Goldman, S.R., Ko, M.L.M., Greenleaf, C., Brown, W., 2018. Domain-specificity in the practices of explanation, modeling, and argument in the sciences. In: Fischer, F., Chinn, C.A., Engelmann, K., Osborne, J. (Eds.), *Scientific Reasoning and Argumentation: The Roles of Domain-Specific and Domain-General Knowledge*. Routledge, London.
- Guilfoyle, L., Erduran, S., 2021. Recalibrating the evolution versus creationism debate for student learning: towards students' evaluation of evidence in an argumentation task. *Int. J. Sci. Educ.* 43 (18), 2974–2995. <https://doi.org/10.1080/09500693.2021.2004330>.
- Guilfoyle, L., Erduran, S., Park, W., 2021. An investigation into secondary teachers' views of argumentation in science versus religious education. *J. Beliefs Values.* 42 (2), 190–204. <https://doi.org/10.1080/13617672.2020.1805925>.
- Guilfoyle, L., 2018. A longitudinal study exploring the influence of epistemic beliefs on pre-service science teachers' perceptions of education studies. Unpublished PhD thesis. University of Limerick, Limerick, Republic of Ireland.
- Hofer, B.K., 2000. Dimensionality and disciplinary differences in personal epistemology. *Contemp. Educ. Psychol.* 25 (4), 378–405. <https://doi.org/10.1006/ceps.1999.102>.
- Hofer, B.K., 2001. Personal epistemology research: implications for learning and instruction. *Educ. Psychol. Rev.* 13, 353–382.
- Hofer, B.K., 2016. Epistemic cognition as a psychological construct: advancements and challenges. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 19–38.
- Ingham, A.I., Gilbert, J.K., 1991. The use of analogue models by students of chemistry at higher education level. *Int. J. Sci. Educ.* 13 (2), 203–215.
- Iordanou, K., Kendeou, P., Beker, K., 2016. Argumentative reasoning. In: Greene, J., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 39–53.
- Iordanou, K., 2010. Developing argument skills across scientific and social domains. *J. Cogn. Dev.* 11 (3), 294–327.
- Jackson, R., 2004. Rethinking religious education and plurality. Routledge Falmer, London.
- Jimenez-Alexandre, M.P., Buggalo-Rodriguez, A., Duschl, R., 2006. Contextualizing practices across epistemic levels in the chemistry laboratory. *Sci. Educ.* 90, 707–733.
- Justi, R., Gilbert, J.K., 2000. History and philosophy of science through models: some challenges in the case of "the atom". *Int. J. Sci. Educ.* 22 (9), 993–1009.
- Kang, N.H., Wallace, C.S., 2005. Secondary science teachers' use of laboratory activities: linking epistemological beliefs, goals, and practices. *Sci. Educ.* 89 (1), 140–165.
- Kang, N.H., 2008. Learning to teach science: personal epistemologies, teaching goals, and practices of teaching. *Teach. Teach. Educ.* 24 (2), 478–498.
- Kelly, G.J., Licona, P., 2018. Epistemic practices and science education. In: Matthews, M. (Ed.), *History, Philosophy and Science Teaching. Science: Philosophy, History and Education*. Springer, Dordrecht.
- Liu, S.Y., Tsai, C.C., 2008. Differences in the scientific epistemological views of undergraduate students. *Int. J. Sci. Educ.* 30 (8), 10551073.
- Markic, S., Eilks, I., 2012. A comparison of student teachers' beliefs from four different science teaching domains using a mixed methods design. *Int. J. Sci. Educ.* 34 (4), 589–608.
- Marra, R.M., Palmer, B., 2011. Personal epistemologies and pedagogy in higher education: did we really mean to say that to our students? In: Brownlee, J., Schraw, G., Berthelsen, D. (Eds.), *Personal Epistemology and Teacher Education*. Routledge, New York, pp. 129–145.
- Muis, K.R., 2004. Personal epistemology and mathematics: a critical review and synthesis of research. *Rev. Educ. Res.* 74 (3), 317–377.
- Nagel, E., 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace and World, New York, NY.
- National Research Council, 2012. A framework for K-12 science education: practices, crosscutting concepts and core ideas. National Academy Press, Washington, DC.
- Peng, H., Fitzgerald, G., 2006. Relationships between teacher education students' epistemological beliefs and their learning outcomes in a case-based hypermedia learning environment. *J. Technol. Teach. Educ.* 14 (2), 255285.
- Ravindran, B., Greene, B.A., DeBacker, T.K., 2005. Predicting preservice teachers' cognitive engagement with goals and epistemological beliefs. *J. Educ. Res.* 98 (4), 222–232.
- Roesken, B., Hannula, M.S., Pehkonen, E., 2011. Dimensions of students' views of themselves as learners of mathematics. *ZDM* 43 (4), 497–506.
- Sandoval, W.A., Reiser, B.J., 2004. Explanation-driven inquiry: integrating conceptual and epistemic scaffolds for scientific inquiry. *Sci. Educ.* 88, 345–372.
- Schwab, J., 1962. The teaching of science as enquiry. In: Schwab, J.J., Brandwein, P.F. (Eds.), *The Teaching of Science*. Simon and Schuster, New York, pp. 1–103.
- Topcu, M., 2013. Preservice teachers' epistemological beliefs in physics, chemistry, and biology: a mixed study. *Int. J. Sci. Math. Educ.* 11 (2), 433–458.
- Topcu, M.S., 2011. Turkish elementary student teachers' epistemological beliefs and moral reasoning. *Eur. J. Teach. Educ.* 34 (1), 99–125.
- Toulmin, 1958. *Uses of Argument*. Cambridge University Press, Cambridge.
- Vangrieken, K., Meredith, C., Packer, T., Kyndt, E., 2017. Teacher communities as a context for professional development: a systematic review. *Teach. Teach. Educ.* 61, 47–59.
- Walker, S., Brownlee, J., Whiteford, C., Exely, B., Woods, A., 2012. A longitudinal study of change in preservice teachers' personal epistemologies. *Aust. J. Teach. Educ.* 37 (5), 24–35.
- Weinstock, M., Roth, G., 2011. Teachers' personal epistemologies as predictors of support for their students' autonomy. In: Brownlee, J., Schraw, G., Berthelsen, D. (Eds.), *Personal Epistemology and Teacher Education*. Routledge, New York, pp. 165–179.
- Windschitl, M., Andre, T., 1998. Using computer simulations to enhance conceptual change: the roles of constructivist instruction and student epistemological beliefs. *J. Res. Sci. Teach.* 35 (2), 145–160.
- Windschitl, M., Thompson, J., Braaten, M., 2008. How novice science teachers appropriate epistemic discourses around model-based inquiry for use in classrooms. *Cogn. Instruct.* 26 (3), 310–378. <https://doi.org/10.1080/07370000802177193>.
- Wong, A.K.Y., Chan, K.W., Lai, P.Y., 2009. Revisiting the relationships of epistemological beliefs and conceptions about teaching and learning of preservice teachers in Hong Kong. *Asia Pac. Educ. Res.* 18 (1), 1–19.
- Yadav, A., Koehler, M., 2007. The role of epistemological beliefs in preservice teachers' interpretation of video cases of early-grade literacy instruction. *J. Technol. Teach. Educ.* 15 (3), 335–361.
- Yenice, N., 2015. An analysis of science student teachers' epistemological beliefs and metacognitive perceptions about the nature of science. *Educ. Sci. Theory Pract.* 15 (6), 1623–1636.
- Yilmaz, H., Şahin, S., 2011. Pre-service teachers' epistemological beliefs and conceptions of teaching. *Aust. J. Teach. Educ.* 36 (1), 73–88.
- Yilmaz-Tuzun, O., Topcu, M.S., 2008. Relationships among preservice science teachers' epistemological beliefs, epistemological world views, and self-efficacy beliefs. *Int. J. Sci. Educ.* 30 (1), 65–85.

Teacher skills and knowledge for technology integration

Leming Liang^a and Nancy Law^b, ^aGuandong Polytechnic Normal University, Guangdong, China; and ^bUniversity of Hong Kong, Pok Fu Lam, Hong Kong

© 2023 Elsevier Ltd. All rights reserved.

This is an update of N. Law, Teacher Skills and Knowledge for Technology Integration, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 211-216, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00746-6>.

Introduction	263
Trends in learning technology development	264
Digital literacy and computational thinking as core 21st century skills relevant for all teachers	265
Educational digital competence	266
Lifelong learning capacity for a future-ready teaching profession	268
References	269

Introduction

The core skills and knowledge that teachers require for integrating digital technology into their teaching vary depending on the perceived role of technology in education. One common categorization distinguishes three different roles: learning *about*, *with*, and *through* technology (Pelgrum and Law, 2003; Erstad and Voogt, 2018). Learning *about* technology refers to technology as part of the school curricula for students to learn, which used to include technology-specific knowledge domains such as programming, computer operations and information literacy. Learning *with* technology refers to the use of technology as a productivity tool for learners and teachers to enhance existing methods of teaching and learning, such as the use of office applications, multimedia, internet, and interactive whiteboard. Learning *through* technology refers to technology as an essential tool integrated into curricula and the process of teaching and learning, without which the curriculum and/or the intended learning experience cannot be realized. For instance, simulations and modeling tools provide opportunities for learners to engage in a wide range of experiences that would otherwise not be possible, such as controlling a nuclear power plant, and exploring the consequences of different severities of social distancing measures on the spread of a pandemic. Access to large authentic databases and analytic tools also provides learners with different ways to explore “what if” questions in the same way that natural and social scientists do in their authentic investigations rather than simply receiving canned answers from teachers.

The different roles that technology is called upon to play in educational contexts also represent different pedagogical paradigms and visions about what learning could be like when technology is used. The key role of teachers’ pedagogical knowledge is highlighted. The success of technology integration in education depends on whether teachers have the competence to intentionally use technology for pedagogical purposes. Technological Pedagogical Content Knowledge (TPACK, Mishra and Koehler, (2006)) is a popular competence framework that describes the knowledge teachers need to acquire for effective learning *with* and *through* technology. This framework highlights the need for teachers to develop knowledge and skills that situate in the interaction space among the three knowledge domains of technology, subject matter content, and pedagogy. In a comprehensive review of international comparative studies of digital technology use and student learning, Law and Liang (2019) reported a shift in the focus of these studies from assessing students’ computing skills (following the learning *about* technology paradigm) to a focus on the pedagogical innovations that may accompany technology use for learning since the turn of the millennium, reflecting a shift toward learning *with* and *through* technology. In terms of student achievement, there is a diversification in foci to include computer-based assessment that tests students’ ability (OECD, 2010) to use digital tools to arrive at an answer (OECD, 2017), as well as students’ collaborative problem-solving ability. These reflect a priority at the policy level for digital technology use to support students in developing 21st century skills (Dede, 2014; Albion and Tondeur, 2018).

However, with the rapid progress and pervasive adoption of technology in all aspects of life, the paradigm of differentiating learning *about*, *with* and *through* technology and the TPACK framework are no longer adequate in providing a comprehensive list of knowledge and skills required by teachers for competent professional practice. The differentiation of technology use into learning *about*, *with* and *through* technology was proposed in the early 2000s (see Pelgrum and Law, 2003) when technology adoption in classrooms was either to serve additional curriculum goals (about technology) or as a tool to augment and enhance teaching and learning. However, two decades on, digital technology has been used pervasively and become an integral part in mediating communication, producing new knowledge, and augmenting everyday life and working experiences. Technological changes take place all the time, and this constant and fast changing technological landscape calls for a core competence to make use of digital technology for creative problem solving in the contexts relevant to the individual, including workplace, social and personal contexts. The technological environment for teaching and learning has also become richer. Effective integration of technology in education requires improving the digital competence of stakeholders at multiple levels, including individual teachers, school leaders, and system level leadership. Like “the fish do not know they are in water”, both teachers and students learn, work, and live in an

ecosystem permeated with digital technology. Thus, the contexts for learning *about*, *with* and *through* becomes intertwined and their differentiation becomes not that important. Technological knowledge and its deployment can no longer be isolated from other domains of knowledge, and digital competence is essential for every citizen's well-being. Teaching and learning without any use of technology greatly limits the educational possibilities available to the learner and the teacher. While teachers' pedagogical knowledge and skills remain important for the use of digital technology in education, their understanding of technology integration should become much broader than as additional augmentations in classroom teaching.

Both the European Commission (Carretero et al., 2017) and UNESCO (Law et al., 2018) have put forward digital competence/digital literacy frameworks that are applicable for everyone in the contemporary society. For teachers and associated education practitioners, this includes the ability to communicate with learners, parents and third parties; to engage in collaboration and knowledge sharing with other educators; and to use digital resources for professional development and classroom teaching (Punie and Redecker, 2017). In recent years, computational thinking (CT) has also been included as a core element of digital literacy that students should learn about. Emerging technologies, such as mobile and wearable devices, virtual reality (VR), augmented and mix reality (AR, MR), robotics, Internet of things with sensors, data science, and artificial intelligence (AI), have been gradually integrated into daily life. The implication of these emerging technologies for education is that more technology contents need to be introduced into the learning *about* technology curriculum (e.g., AI techniques, machine learning, and ethics in AI). On the other hand, a wide range of AI-driven and autonomous applications and services have been (and are being) developed to support student learning *with* and *through* technology. For example, the use of AI-powered assistive tools can help to automate learning tasks. Data-driven learning analytics tools can suggest for students a personalized and adaptive learning pathway. VR technology support synchronous and asynchronous modes of learning. Teaching and learning mediated entirely through digital technologies have made it possible for educational services to be maintained during crisis periods such as the COVID-19 pandemic.

We argue that conceptualizing teachers' competence for technology integration should go beyond TPACK and take a broader perspective of ensuring students' well-being in preparation for an uncertain future in a digital society. In this article, students' well-being is defined based on the concept of "the four pillars of education" put forward by UNESCO (Delors, 1996). According to the concept of "the four pillars of education", teachers should teach students or provide learning experience for them to "learn to know", "learn to do", "learn to live together" and "learn to be". Learning to know refers to competence in learning subject matter knowledge and skills. Learning to do is a generic competence in working with others and handling situations arising from teamwork. Learning to live together is the most important for the well-being of the society, on both local and global dimensions. It requires not only the competence to deal with conflicts, but also the appreciation of our interdependence as a global citizen and a respect for different cultures and other diversities. Learning to be is the ability to reflect and understand oneself, to act with autonomy and personal responsibility. This article provides an overview of teachers' competence needed for technology integration that is future oriented in view of the changing technological, societal and educational contexts.

Trends in learning technology development

Emerging technologies refer to the applications and services that are driven by artificial intelligence (AI), virtual, augmented, and mix reality (VR, AR, and MR), wearable devices with displays and sensors, Internet of things with 5G mobile standard, and social robotics (Vuorikari et al., 2019). These emerging technologies are part of an evolving, digitally infused, ecosystem that influence the social, mental, physical, and emotional well-being of students, and indeed teachers and other inhabitants of the "global village". In educational contexts, emerging technologies can provide immersive learning experiences, support data-driven decision-making, and AI-powered automated learning support. Head-mounted devices such as Google Glass provides a first-person view of particular learning scenarios, which can be used for the education of surgeons (Alaker et al., 2016). In a systematic review and meta-analysis of virtual reality (VR) for training in laparoscopic surgery, Alaker et al. (2016) found that wearable technologies can provide not only simulated surgical experiences, but also enable students to review the simulated surgery from different perspectives based on the self-recording of the training sessions. Bower and Sturman (2015) described applications of wearable devices in the context of augmented reality (AR) for enhanced learning experiences such as in field trips. Learners can access multimodal information (videos, photos, textual information, etc.) using diverse platforms and devices. Another important area that emerging technologies can contribute to is supporting assessment and feedback through the application of data science and learning analytics. Tracing students' performance changes over time and how such changes connect with the learner's actions and interactions based on real time multimodal data provides new possibilities for formative assessment practices, both in terms of ongoing direct feedback to learners as well as enabling teachers to redesign learning experiences.

In recent years, the application of AI to education (aka, AI in education, AIEd) has received great attention from two perspectives. The first is how AI technologies can help us understand better how student learn and provide a more adaptive and personalized learning experiences to students (Luckin et al., 2016). AI-driven educational applications are usually built on voice and facial recognition, educational data mining, computer modeling and machine learning technologies. Voice recognition allows students to ask questions and seek answers from AI-powered database. Facial recognition can be used to authenticate students' identity, Analysis of students' behavioral and performance data support autonomous decision-making and providing intelligent tutoring support (Vuorikari et al., 2019).

In addition to supporting learning and teaching, AIEd has also been introduced as a subject matter component in basic education in the K-12 school curriculum, when the goal of education is not on preparing students to become AI experts. Touretzky and Gardner-McCune (2022) identify five types of knowledge of AI that K-12 students should know, which are described as Five Big

Ideas: perception, representation and reasoning, learning, natural interaction, and societal impact. Apparently, four of the Big Ideas (except for societal impact) are about understanding how AI as computer systems can interact with the world like a human. Touratzky and Gardner-McCune argued that studying AI in K-12 should go beyond learning technology to also help students appreciate the complexity of how people think, learn, make decisions, respond, and interact with another human. The last Big Idea, societal impact, concerns the ethics of AI used, and the potential benefits and threats of AI, data privacy, etc.

For teachers and educators, learning about ethics of AI *in general* as a subject matter is not enough to teach AI. Teachers should also know about ethics of AIED *in particular* (Holmes et al., 2021) when applying AI-powered technologies in the process of teaching and learning. Ethics of AIED describes how teachers make technological and pedagogical decisions that are ethical. Holmes et al. (2021) put forward a preliminary framework for ethics of AIED. The framework identifies three foci of ethical issues of AIED: big data, algorithms and computation, and education, based on the assumption that ethics of AIED is not simply about data and computational approaches but needs to consider the ethics of education (equity, assumptions of students' characteristics, the choice of pedagogy, etc.). The Institute for Ethical AI in Education (2021) launched a comprehensive ethical framework for AIED comprising nine principles to ensure that all students can benefit optimally from AIED while being protected against adverse effects such as students losing their agency and autonomy in learning, being subjected to biases due to the data or algorithms used in the AI applications, and loss of personal privacy. The nine principles are (1) achieving educational goals, (2) forms of assessment, (3) administration and workload, (4) equity, (5) autonomy, (6) privacy, (7) transparency and accountability, (8) informed participation, and (9) ethical design. Teachers would need to have an appropriate level of understanding of AI and its implications in order to be able to make judgment and informed decisions on the selection and adoption of AIED systems and applications. Unscrupulous avoidance or adoption would harm learners' well-being in at least three of the four pillars of education: learning to know, learning to do, and learning to be.

In fact, not only AIED, but all uses of digital technology in education requires teachers/educators to consider the associated ethical consequences as well as the need to protect children's rights in the digital world. The United Nations Committee on the Rights of the Child (2021) issued the General comment No. 25 on children's digital rights that need to be respected and protected, including access to information, freedom of expression, right to privacy, digital birth registration and right to identity. The 5Rights Foundation, an organization established to promote the rights of children online, developed technical standards and protocols to ensure digital services and products cater for children and young people (5Rights foundation, 2022). It advocates regulations and international agreements in four priority areas: (1) Design of Service, (2) Child Online Protection, (3) Children and Young People's Rights, and (4) Data Literacy. Recognizing and respecting children's digital rights, and being knowledgeable about the areas of action necessary to ensure that such rights are protected are important aspects of teachers' professional competence in the 21st century, current developments regarding children's digital rights have implications for teachers' professional practice, including (1) the use of digital technology should cater for the vulnerabilities, needs, and rights of children, and help children flourish; (2) measures should be in place to protect students from exposure to age-inappropriate or harmful materials, unsafe contact from adults, cyberbullying and exhortations to gamble; (3) measures to protect children from digital harm should not jeopardize children's rights and autonomy in the digital environment; and (4) the need to equip students with adequate digital literacy so that they can make use of digital tools "knowledgeably, creatively, and fearlessly" (5Rights foundation, 2022, para. 5).

Digital literacy and computational thinking as core 21st century skills relevant for all teachers

New literacies and competencies are needed for students to thrive in the digital society, and education systems have to respond to such new requirements and renew school curricula by re-defining what needs to be learned and taught (Erstad and Voogt, 2018). Substantial changes have occurred in terms of the scope of what needs to be learned for students' well-being and for them to be able to exercise their citizenship rights and responsibilities (Law et al. (2018); Law and Lee (2023)). These changes have made it necessary for some level of digital literacy and computational thinking to become a core competence for all teachers rather than only for teachers of computing subjects.

The term "digital citizenship" emerged in the educational vocabulary since the turn of the millennium, and the increasing prominence of this strand of research is accompanied by a strong realization that the rights and responsibilities of individuals cannot be exercised without an adequate capacity to handle the complexities of living in a world where the physical and the digital are inextricably intertwined (Chen et al., 2021). Digital citizenship as a human capacity goes far beyond the technical skills needed for digital operations. While definitions of digital citizenship differ in terms of their scope, digital literacy (DL)/digital competence (DC) lies at the core of all the different definitions. The former (DL) is the term used in North America and the latter (DC) is the term commonly adopted in Europe. In this article, we use these two terms interchangeably, depending on the source we reference. The DigComp framework developed by the European Commission (EC) Joint Research Center (Carretero et al., 2017) comprises five key competence areas: information and data literacy, communication and collaboration, digital content creation, safety, and problem solving. Recognizing that the DL needed may be dependent on the infrastructural, socioeconomic and cultural contexts involved, a study to develop a Digital Literacy Global Framework (DLGF) was commissioned by UNESCO (Law et al., 2018) to take account of the development needs and realities, particularly for developing countries and economies. The main difference between DLGF and DigComp is that the former comprises two additional competence areas: hardware and software operations and career-related competences. More importantly, DLGF provides evidence that the same functional operations (e.g., a digital financial operation) may involve different digital operations and hence different kinds of DL performance depending on the digital

devices (e.g., mobile phones vs. smart phones) and the network infrastructure available (e.g., whether the available mobile network can connect to the Internet or not). It also highlights the need for DL curricula and training to be sensitive to the predominant as well as changing technological, socioeconomic, and cultural contexts.

With the rapid development of emerging technologies, such as AI, data science, robotics and IoT, there is a need to update citizens' DL so that they can engage with these technologies in different technology use contexts for learning and work. EC has been engaged in the progressive refinement of the DigComp framework over the past decade. DigComp 2.2, due to be launched in February 2022, provides an elaboration of the framework by providing examples of knowledge, skills and attitudes (referred to as dimension 4 in the framework) applicable to each of the five competence areas (i.e., information and data literacy, communication and collaboration, digital content creation, safety, and problem solving) in three specific areas of technology use for learning, for work, and for participation in society: AI systems, media and information literacy, and protecting health and well-being. It is clear from the dimension 4 elaborations that the digital competences targeted focus on ensuring individual and social well-being, and that these are relevant not only for children and youth, but for all citizens. It is also evident that the knowledge, skills and attitudes associated with digital competence would need to be integrated with those relevant to the workplace context of the individuals concerned for them to be of value to learning and work. This implies that for teacher professional development, there would be a core level of competence expected of all teachers, as well as specialized knowledge, skills and attitudes relevant to the roles of the teachers concerned.

Computational thinking (CT) refers to the ability to address real-world problems by re-formulating them into computational problems and designing algorithmic solutions that can be carried out by computers (Fraillon et al., 2019). While CT originated in the computer science field (Wing, 2008; Barr and Stephenson, 2011), the rationale for its importance in basic education goes beyond algorithmic thinking, and includes cognitive skills such as problem decomposition, problem abstraction, pattern recognition, and generalization (Wing, 2008; Grover and Pea, 2013). CT is considered as a sub-area under problem-solving in the DigComp framework (DC4JOBS, 2020). It is often considered an important 21st century competence as it is assumed that everyone will need to communicate, learn, and work with computers in the future. It has been integrated into computer science/STEM curricula in K-12 settings (Yadav et al., 2018). CT will be included as an integral part of DL in the ICILS 2023 assessment (Fraillon et al., 2019). Thus, similar to other domain aspects of digital competence, all teachers need to have a basic level of literacy in CT, while teachers with specialist roles in teaching computational subjects would need further levels of competence and possibly in a wider range of knowledge and skills.

Educational digital competence

Teachers' attainment of an adequate level of digital literacy is a precondition for their ability to adequately carry out their professional responsibilities. Several competence frameworks have been proposed to define and describe the types of teacher digital literacy that align teacher knowledge and skills to students' 21st century learning (Caena and Redecker, 2019). The European Framework for the Digital Competence of Educators (aka. DigComEdu, Redecker, 2017) was developed for the purpose of guiding teacher professional development. It comprises 22 elementary competences organized in six areas that involve the use of digital technology: (1) professional engagement, (2) digital resource, (3) teaching and learning, (4) assessment, (5) empowering learners, and (6) facilitating learners' digital competence. Areas (2), (3) and (4) are related to teachers' use of digital technologies to achieve their "traditional" roles in designing learning materials, preparing lessons and evaluating students' learning outcomes, Area 6 is about teachers' ability to foster students' DL (as described in the DigComp framework) as an extension of the curriculum responsibilities that a teacher is expected to achieve. Area 5 concerns teachers' ability to provide adaptive and personalized learning experiences to ensure the accessibility and inclusiveness of learning resources and activities offered to all learners such that students' transversal skills, deep thinking and creative expression can be fostered as an important aspect of student well-being. Area 1 stipulates that teachers need to be able to engage in professional learning and interactions with relevant communities of practice via digital technology. This area of competence ensures that teachers can have the conditions for sustainable, lifelong learning in learning to know, to do, to be and to live together as a member of the teaching profession.

The framework also puts forward a six-level proficiency model indicative of a pathway for teachers' DC development, progressing from two beginner levels—Newcomer and Explorer, to two intermediate levels—Integrator and Expert—when teachers can use technology in meaningful ways in increasingly broader contexts, to two advanced levels—Leader and Pioneer—when teachers are able to reflect on and improve their own practices, share their expertise with others, develop and lead more complex and advanced ways of technology use.

The US national educational technology standards on computational thinking for educators (ISTE, 2017) also puts forward a similar list of professional roles that teachers may play in education: learner, leader, collaborator, designer and facilitator. These standards involve a more comprehensive set of 21st-century oriented skills that go beyond subject-content and pedagogical competence and highlights the capacity of teachers as life-long learners and problem solvers. To be able to foster the development of computational thinking, teachers need to develop resilience and perseverance in addressing open-ended and ambiguous problems, be able to cope with failure and treasure such experiences as valuable opportunities to learn. In terms of professional practice, teachers should learn (1) approaches to establish an inclusive culture that respects diversity, (2) collaboration skills to work with colleagues and students to create meaningful learning opportunities, and (3) a creative design mindset to inspire students in the process of problem solving.

The key drivers of economic growth have changed from manufacturing and services to information and knowledge, and collaboration, information sharing, and self-management have become increasingly important abilities in the workplace (Kozma, 2009). The UN Agenda 2030 for Sustainable Development delineates 17 sustainable development goals (SDGs) which are perceived as important for education systems worldwide. Teachers need to be well-prepared to guide students to achieve these goals with the support of technology. UNESCO launched the third version of the ICT Competency Framework for Teachers (ICT-CFT) in 2018 to take account of these goals. There are six aspects of teacher knowledge and skills under each stage: understanding ICT in education policy; curriculum and assessment; pedagogy; application of digital skills; organization and administration; and teacher professional learning.

According to the 2018 ICT-CFT framework, pedagogy is important for technology use to support student-centered, collaboration and inquiry-based problem-solving to foster students' 21st century skills in alignment with national policy documents. With regard to curriculum and assessment, teachers should be able to identify the types of technology conducive to the achievement of specific curriculum goals and how technology use can support assessment. This requires that teachers take into consideration emerging technologies relevant to different aspects of the competence framework, with a focus on those that foster students' 21st century skills through learning experiences that engage learners in problem-solving, communication, collaboration, experimentation, critical thinking, and creative expression. For organization and administration, teachers should be able to engage in social networking for knowledge management and professional development, and to use AI-supported software to streamline administrative management processes. At the leadership level, teachers need to learn how to transform schools into learning organizations capable of actualizing the implementation of e-learning. Similar to the other teacher competence frameworks, the teacher professional learning aspect also focuses on the use of technology that empowers teachers to be life-long learners for sustainable professional development.

In the teacher learning literature, Technological Pedagogical Content Knowledge (TPACK) is a competence framework that describes the knowledge teachers need to teach competently using digital technologies (Mishra and Koehler, 2006). This framework highlights the need for teachers to develop knowledge and skills that situate in the intersection space among the three knowledge domains of technology, subject matter content, and pedagogy. There are four areas of interaction in this framework: technological content knowledge (TCK), technological pedagogical knowledge (TPK), pedagogical content knowledge (PCK), and all three taken together as technological pedagogical content knowledge (TPCK).

Yeh et al. (2014) emphasized the importance of practical skills for competent implementation. They put forward a TPACK-Practical framework comprising eight practice dimensions, namely, using ICT to understand content, using ICT to understand learners, planning ICT-infused curriculum, using ICT representations, using ICT-integrated teaching strategies, applying ICT to instructional management, infusing ICT into teaching contexts, and using ICT to assess students. Ethical issues are also an important aspect that teachers need to address when using emerging technology. Benton-Borghi (2013) proposed an infusion of universal design for learning (UDL) with the TPACK framework in order to cater to the needs of all learners, including marginalized groups (e.g., those from low SES backgrounds, racial/ethnic/linguistic/cultural minorities, those with disabilities). This extension guides teacher's pedagogical decision-making regarding delivery models, suggests multiple options for the representation of and ways to access the content.

While the TPACK framework specifies the knowledge domains necessary for teachers to integrate technology into teaching and learning, teachers' design competence is also crucial for their effective mobilization in different course and learner contexts. When knowledge creation and authentic problem solving are key competences to be fostered in students, the teacher's role needs to change from an instructor who delivers well-defined facts and skills to a designer who carefully crafts the technology-enhanced learning environment within which students can learn proactively and productively through self-directed inquiry, problem-solving and creative (co-) constructions. Learning design is a field that engages in "devising effective learning experiences aimed at achieving defined educational objectives in a given context" (Mor et al., 2015, p. 221). Learning design is complex because it involves different levels of decision-making: course/curriculum unit, curriculum component, and learning tasks (Law and Liang, 2021). Design decisions at hierarchically nested levels need to be guided by pedagogical philosophy, approach, strategy, and tactics respectively (Goodyear, 2005). Moreover, learning design competence includes the ability to make use of learning analytics to inquire into student learning (Hansen and Wasson, 2016) for continuous improvement of design principles and practice. Teachers' participation in a co-design process to choose and make sense of learning analytics tools is proven to be crucial for effective adoption and integration of LA technologies (Rosé et al., 2019).

An important aspect of teacher learning regarding emerging technologies is whether the user recognizes the pedagogical affordance of the technology. Here, the word "affordance" is interpreted as defined by Norman (1988), which is the perceived affordances (aka possibilities) rather than the designed affordances that are most important in influencing their usage/application. Angeli and Valanides (2018) hypothesize a five-stage process for the pedagogical affordances of technologies to be actualized. The first stage requires the necessary technical affordances to be made available as built-in functions made available in the form of educational applications/tools. The second stage involves the teachers' recognition and familiarization of the technical affordances of these applications. For example, concerns about the privacy, ethical and social issues brought about by the capacity of wearable technologies to instantaneously broadcast information about the wearers (Gill, 2008) can only be raised by those who recognize the possibilities that such technologies bring. To advance to the third stage requires that the teacher perceives the pedagogical affordance of the tools, i.e., the pedagogical actions that can be enabled through the technical affordance. For example, a teacher may be able to perceive and make use of natural language processing AI tools such as Amazon Alexa as a handy dictionary in her own work. For teachers to recognize that this same technology can be used to help students to improve their pronunciations if they can be

encouraged to talk to these AI agents requires a third stage of teacher learning advancement—perception of the pedagogical affordance that can be leveraged through the technology. While the first three stages involve conceptual recognition, the fourth stage concerns the actualization of the pedagogical potential of the technologies concerned. This requires the fulfillment of additional conditions, including the teacher possessing the necessary instructional design and implementation skills, the availability of the necessary technological infrastructure and resources, as well as the compatibility of the affordance with the school's curriculum and developmental priority. Actualization would be accompanied by visible observations of the impact on student learning, which in turn deepens the understanding about the affordance, leading to a fifth stage when those involved in the technology-supported pedagogical innovation can reflect on the experience and revise their perception of the possible affordances and pedagogical designs in the deployment of the technology.

Angeli and Valanides' (2018) five stage framework indicates that TPACK only covers the first three stages, and is inadequate for the effective actualization of the educational potentials of digital technology. Han and Patterson's (2020) empirical study of the challenges encountered by an elementary teacher in his attempts to use VR in his teaching found that the obstacles were multifaceted. These included not only teachers' TPACK in integrating VR into the school curriculum, but also some generic skills. For example, the teacher had to communicate with parents and students to make explicit the expected process and outcomes of the VR lessons. He also needed to make an alternative plan to cater for students with motion sickness. Managing the budget, space, and time limit for the actual implementation of VR were additional concerns for the teacher. These findings are consistent with EC's DigComEdu framework, which stipulates that teachers' digital competence need to go beyond technology and pedagogy, such as interacting and sharing with colleagues, collaboration within schools, and communicating with parents. Porras-Hernández and Salinas-Amescua (2013) enriched the TPACK framework by including two dimensions of contextual knowledge: scope (encompassing three nested levels: macro or system level, meso or school level, and micro) and actor (students' and teachers' inner and external context). These two dimensions are intertwined, with teachers' and students' characteristics, needs, preferences, prior knowledge, self-efficacy, pedagogical beliefs, subject discipline, school culture, ethnicity, community culture, and socioeconomic background being prominent contextual aspects to be taken into consideration in the actualization process. In addition to knowledge.

In summary, a teacher's educational digital competence involves knowledge and skills in the intersection of technology, pedagogy, and subject matter content, as well as learning design competence to construct and orchestrate appropriate learning environments and learner experiences for students to achieve the desired outcomes, including achieving appropriate levels of digital competence. A core challenge for teachers as the learning designers is to provide learning experiences that foster learner agency, protect learners from cyber risks while ensuring their rights to participate and benefit from engagement and interactions in cyberspace. In addition, for teachers to actualize the affordances of digital technologies for learning, they need communication and collaboration skills to work with parents, colleagues, and other stakeholders to create the necessary conditions to implement technology-enhanced learning innovations. These are tall orders that can only be achieved through appropriate leadership at the institution and system levels that builds a collaborative culture and supportive communities of practice for teachers to experiment and share their visions and experiences. Teacher's effective use of technology also depends on the technical infrastructure available to them and the social infrastructure to provide appropriate administrative and professional support.

Lifelong learning capacity for a future-ready teaching profession

It is clear from the preceding discussion that the knowledge and skills required of the teacher are evolving with the rapid development of technology, and there is an increasing recognition that the use of technology in education should prepare students for the future. At the turn of the millennium, learning about technology was considered as a low level of technology integration in education (cf. Law, 2010). With the rapid advances in digital technology and their pervasive deployment in all aspects of our lives, at work, for learning, socialization, entertainment, as well as to exercise our rights and responsibilities as a citizen, the ability to master the use of emerging technologies in daily life and work is crucial for the well-being of individuals and societies. Education systems need to build in mechanisms for agile adoption and adaptation of new technologies to support learning and teaching that supports professional learning and cares for teachers' well-being in the process.

The COVID-19 pandemic serves as a case to demonstrate how digital technology can help schools and universities to provide instructional services to students even under strict social distancing measure. At the same time, it also shows that the effectiveness of the online learning arrangements depends on how well-prepared different systems and institution are regarding the use of digital technology in learning and teaching before the pandemic started. The preparedness includes the availability of appropriate technology infrastructure, access to devices, digital competence of teachers and students, as well as their prior experience in e-learning (Law et al., 2021a). The pandemic was the Black Swan that forced teachers worldwide to seek solutions from technology, which was not necessary before the pandemic. Teachers have to move existing face-to-face classroom practices to synchronous (e.g., Zoom, Google Meet) and asynchronous online tools (e.g., Moodle, Google Classroom) as an emergency and tactical response to crisis. In contrast to the well-planned blended or fully online learning designs, the shift to online instructional delivery due to the pandemic is often referred to as emergency remote teaching (ERT) in the literature (Hodges et al., 2020; Reynolds and Chu, 2020), instituted to maintain as far as possible the existing educational service during the crisis situation. On the other hand, there are others that took the forced use of digital technology for teaching and learning as an opportunity for long-term innovation of the educational ecosystem (Hodges et al., 2020), and coined the vision of a "New Normal" after the pandemic (Law et al., 2021a;

Berwick, 2020). While ERT may provide some stopgap relief, teachers' digital literacy and TPACK are crucial for technology deployment to effectively achieve the intended outcomes. During the pandemic, digitally competent teachers were able to use social media (e.g., twitter, Trust et al., 2020) and online learning communities (Borup et al., 2020) to strengthen interaction, share and collaboratively design with colleagues. 360 video recordings (Zolfaghari et al., 2020) and VR simulations (Sasaki et al., 2020) can be used to provide immersive learning environments for teachers to implement and reflect on their teaching practices. König et al. (2020) conducted a survey on 165 early career teachers in Germany who adopted online teaching during COVID-19 school closure. They found that teachers' TPACK was significant in predicting their ability to maintain social contact and to provide task differentiation (through online adaptive teaching).

Since the role of TPACK is still prominent in influencing teachers' responses to the crisis, digital divide among teachers was found to be a barrier for the successful implementation of ERT. Iivari et al. (2020) conducted case studies on teachers' response to the crisis in Finland and India, and found that teachers who had a high level of digital literacy showed greater resilience, creativity and perseverance in responding to the challenging situations due to COVID-19. Gudmundsdottir and Hathaway (2020) investigated US and Norwegian teachers' readiness and agency during the pandemic using a four-dimensional model of competence: pedagogical, ethical, attitudinal, and technical. They found that most due to teachers in their study were well-prepared in the dimension of technical and attitudinal dimensions, but not in the pedagogical and ethical dimensions.

The COVID-19 pandemic sends a clear message that we need to be prepared for the unexpected. Code et al. (2020) argued that teachers who were better prepared in responding to emergencies and continue to cater for their students' well-being do not only need digital competence, as the abrupt transition to online teaching is a process of transformative learning for the teachers concerned (Mezirow, 2006). During such transitions, teachers experience "a deep structural shift in thought and action" that force them to rethink how they teach and how students learn. Frameworks from the field of learning design may provide a useful lens to support teachers in making flexible use of online technology tools and resources in an agile manner to transform their teaching into blended/fully online modes. Code et al. (2020) proposed a learning design framework that highlighted six aspects that teachers need to reconsider in their design of blended learning: authentic backbone (the core task to be performed); agency (the roles of students and teachers); spatial (the locations for learning to take place); instrumental (the kinds of tools to be used); temporal (the timeframe and sequence of learning tasks); and quick wins (solutions and resources that can be quickly deployed).

For ERT to be a successful process of transition into a New Normal in which the online learning provided to students are transformative in ensuring student agency and catering for diversities requires not only preparedness at the teacher level, but also holistic changes at the school leadership level to support teachers' technology integration practices. In a study to investigate how students, parents, teachers, and school leaders in K-12 schools in Hong Kong were responding to the crisis situation during the first half of 2020, and the extent to which well-being was maintained for different stakeholders in the process, Law et al. (2021b) made several important observations: (1) teachers' online teaching preparedness at the school level has a strong influence on students' online learning participation, academic self-efficacy and the learning of DL skills; (2) schools' e-learning strategic plans and their implementation before the pandemic, including the e-learning coordination team composition (e.g., membership includes senior leaders in the school) and functions (e.g., the ability to allocate human and physical resources, to make decisions on changes in school routines), infrastructure priorities, and teacher professional development provisions were significant predictors of teachers' online teaching self-efficacy and students' self-reported improvements in DL during the pandemic; and (3) a school-wide culture of openness to innovation and a climate of collaboration, mutual trust, and collaboration at the school level were key to predicting teacher preparedness for online teaching.

Results of the Teaching and Learning International Survey 2018 (TALIS 2018) show that just over half of the surveyed teachers in OECD (53%) reported frequent use of technology for students' class work (OECD, 2019). How should education in the post pandemic period prepare educators and students for an unpredictable future? What kinds of teacher and student capacity can deal with rapid, unpredictable change in the workplace and in society? Richards and Dede (2020) proposed a "60-Year Curriculum" (60YC) as a new model of lifelong learning to cater for unpredictable and constant transitions. Lifelong learning takes place in a global network where learning content (i.e., the knowledge and skills required by the work) and processing tools (technologies and resources that support the work) are available for those prepared for "just-in-time" learning. To prepare students for an unpredictable future as a citizen living in a society in which the physical and the digital are merged and no longer distinguishable (the so-called metaverse), it is even more important for teachers to help students build and exercise transversal skills (often referred to as 21st century skills) in the cognitive, intrapersonal, and interpersonal domains. In this context, teachers' continuous professional learning needs to be supported by various kinds of network and collaboration partnership that provides opportunities to interact and learn from other education professions, including pedagogical coaches, experienced teachers, teacher educators, and educational researchers.

References

- 5 Rights Foundation, 2022. About 5Rights Foundation. Retrieved from. <https://5rightsfoundation.com/about-us/>.
- Alaker, M., Wynn, G.R., Arulampalam, T., 2016. Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. *Int. J. Surg.* 29, 85–94.
- Albion, P., Tondeur, J., 2018. Information and communication technology and education: meaningful change through teacher agency. In: Voogt, J., Knezek, G., Christensen, R., Lai, K.W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer, pp. 381–396.

- Angeli, C., Valanides, N., 2018. Knowledge base for information and communication technology in education. In: Voogt, J., Knezek, G., Christensen, R., Lai, K.W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer, pp. 398–414.
- Barr, V., Stephenson, C., 2011. Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *ACM Inroads* 2 (1), 48–54.
- Benton-Borghi, B.-H., 2013. A Universally Designed for Learning (UDL) infused Technological Pedagogical Content Knowledge (TPACK) practitioners' model essential for teacher preparation in the 21st century. *J. Educ. Comput. Res.* 48 (2), 245–265.
- Berwick, D.M., 2020. Choices for the "new normal". *JAMA* 323 (21), 2125–2126.
- Borup, J., Jensen, M., Archambault, L., Short, C.R., Graham, C.R., 2020. Supporting students during COVID-19: developing and leveraging academic communities of engagement in a time of crisis. *J. Technol. Teach. Educ.* 28 (2), 161–169.
- Bower, M., Sturman, D., 2015. What are the educational affordances of wearable technologies? *Comput. Educ.* 88, 343–353.
- Caena, F., Redecker, C., 2019. Aligning teacher competence frameworks to 21st century challenges: the case for the European Digital Competence Framework for Educators (DigCompEdu). *Eur. J. Educ.* 54 (3), 356–369.
- Carretero, S., Vuorikari, R., Punie, Y., 2017. *DigComp 2.1: The Digital Competence Framework for Citizens*. Publications Office of the European Union, Luxembourg.
- Chen, L.L., Mirpuri, S., Rao, N., Law, N., 2021. Conceptualizations and measurement of digital citizenship across disciplines. *Educ. Res. Rev.* <https://doi.org/10.1016/j.edurev.2021.100379>.
- Code, J., Ralph, R., Forde, K., 2020. Pandemic designs for the future: perspectives of technology education teachers during COVID-19. *Inf. Learning Sci.* 121 (5/6), 419–431.
- DC4JOBS (Digital Competence for Jobs), 2020. Digital Competence 5: Problem Solving. Retrieved from: <http://dc4jobs.eu/wp-content/uploads/2020/01/5.pdf>.
- Dede, C., 2014. The Role of Digital Technologies in Deeper Learning. Students at the Center: Deeper Learning Research Series. Jobs for the Future.
- Delors, J., 1996. *Learning: The Treasure Within*; Report to UNESCO of the International Commission on Education for the Twenty-First Century. UNESCO Publishing, p. 46.
- Erstad, O., Voogt, J., 2018. The twenty-first century curriculum: issues and challenges. In: Voogt, J., Knezek, G., Christensen, R., Lai, K.W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer, pp. 19–36.
- Frailon, J., Ainley, J., Schulz, W., Duckworth, D., Friedman, T., 2019. Computational thinking framework. IEA International Computer and Information Literacy Study 2018 Assessment Framework. Springer International Publishing, Amsterdam NL, pp. 25–31.
- Gill, S.P., 2008. Socio-ethics of interaction with intelligent interactive technologies. *AI Soc.* 22 (3), 283–300.
- Goodyear, P., 2005. Educational design and networked learning: patterns, pattern languages and design practice. *Australas. J. Educ. Technol.* 21 (1), 82–101.
- Grover, S., Pea, R., 2013. Computational thinking in K-12: a review of the state of the field. *Educ. Res.* 42 (1), 38–43.
- Gudmundsdottir, G.B., Hathaway, D.M., 2020. "We always make it work": teachers' agency in the time of crisis. *J. Technol. Teach. Educ.* 28 (2), 239–250.
- Han, I., Patterson, T., 2020. Teacher learning through technology-enhanced curriculum design using virtual reality. *Teach. Coll. Rec.* 122 (11), 1–34.
- Hansen, C.J., Wasson, B., 2016. Teacher Inquiry into student learning: the TISL heart model and method for use in teachers' professional development. *Nordic J. Dig. Literacy* 11 (1), 24–49.
- Hodges, C., Moore, S., Lockee, B., Trust, T., Bond, A., 2020. The Difference Between Emergency Remote Teaching and Online Learning. EDUCAUSE, US. Review Report, Retrieved June 1, 2020 from: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S.B., Koedinger, K.R., 2021. Ethics of AI in education: towards a community-wide framework. *Int. J. Artif. Intell. Educ.* 1–23.
- Iivari, N., Sharma, S., Ventä-Olkkonen, L., 2020. Digital transformation of everyday life—how COVID-19 pandemic transformed the basic education of the young generation and why information management research should care? *Int. J. Inf. Manag.* 55, 102183.
- International Society for Technology in Education (ISTE), 2017. ISTE Computational Thinking Competencies. Retrieved from: <https://www.iste.org/standards/iste-standards-for-computational-thinking>.
- König, J., Jäger-Biela, D.J., Glutsch, N., 2020. Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *Eur. J. Teach. Educ.* 43 (4), 608–622.
- Kozma, R.B., 2009. The knowledge ladder: using ICT and education reform to advance economic and social development goals. In: *Symposium on Education Technology in Schools: Converging Innovation and Creativity*.
- Law, N., Lee, W.O., 2023. Curriculum and digital citizenship. In: Tierney, R., Rizvi, F., Ercikan, K. (eds.), *International Encyclopedia of Education* (4th ed.), Elsevier. (in press).
- Law, N., Liang, L., 2019. Digital technologies and educational transformation. In: Suter, L., Smith, E., Denman, B. (Eds.), *SAGE Handbook of Comparative Studies in Education*. Sage, pp. 508–531.
- Law, N., Liang, L., 2021. A multilevel framework and method for learning analytics integrated learning design. *J. Learning Anal.* 7 (3), 98–117.
- Law, N., Woo, D., Wong, G., 2018. A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2. UNESCO, p. 146. Information Paper No. 51. <http://uis.unesco.org/sites/default/files/documents/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf>.
- Law, N., Tan, C.Y., Pan, Q.Q., Lan, M., 2021a. From Evidence-Based Community-Wide Collaborative Action to Recommendations for a Better New Normal. Retrieved from Hong Kong. https://360-cms.ecitizen.hk/uploads/bulletin06_en_e329864845.pdf.
- Law, N., Tan, C.Y., Lan, M., Pan, Q., 2021b. eCitizen Education 360 Project—Bulletin 4: Multi-level School Leadership for Online Learning Preparedness. Faculty of Education, The University of Hong Kong, Hong Kong. https://360-cms.ecitizen.hk/uploads/bulletin04_en_ac768f3e5b.pdf.
- Law, N., 2010. Teacher skills and knowledge for technology integration. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, vol. 8. Elsevier, Oxford, pp. 211–216.
- Luckin, R., Holmes, W., Griffiths, M., Forcier, L.B., 2016. *Intelligence Unleashed: An Argument for AI in Education*. Pearson Education, London.
- Mezirow, J., 2006. An overview of transformative learning. In: Sutherland, P., Crowther, J. (Eds.), *Lifelong Learning: Concepts and Contexts*. Routledge, New York, NY, pp. 24–38.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a new framework for teacher knowledge. *Teach. Coll. Rec.* 108 (6), 1017–1054.
- Mor, Y., Ferguson, R., Wasson, B., 2015. Editorial: learning design, teacher inquiry into student learning and learning analytics: a call for action. *Br. J. Educ. Technol.* 46 (2), 221–229.
- Norman, D., 1988. *The Psychology of Everyday Things*. Basic Books, New York.
- OECD, 2010. *PISA Computer-Based Assessment of Student Skills in Science*. OECD Publishing, Paris.
- OECD, 2017. *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, revised ed. OECD Publishing, Paris. <https://doi.org/10.1787/9789264281820-en>.
- OECD, 2019. *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*. TALIS, OECD Publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>.
- Pelgrum, W.J., Law, N.W.Y., 2003. ICT in Education Around the World: Trends, Problems and Prospects. UNESCO: International Institute for Educational Planning.
- Porras-Hernández, L.H., Salinas-Amescua, B., 2013. Strengthening TPACK: a broader notion of context and the use of teacher's narratives to reveal knowledge construction. *J. Educ. Comput. Res.* 48 (2), 223–244.
- Punie, Y., Redecker, C. (Eds.), 2017. *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union, Luxembourg.
- Redecker, C., 2017. *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union, Luxembourg, ISBN 978-92-79-73494-6. <https://doi.org/10.2760/159770>. JRC107466.
- Reynolds, R., Chu, S.K.W., 2020. Guest editorial: introduction to the special issue on emergency remote teaching (ERT) under COVID-19. *Inf. Learning Sci.* 121 (5/6), 233–239. <https://doi.org/10.1108/ILS-05-2020-144>.

- Richards, J., Dede, C., 2020. The 60-year curriculum: a strategic response to a crisis. *Educause Rev.* Retrieved from <https://er.educause.edu/articles/2020/10/the-60-year-curriculum-a-strategic-response-to-a-crisis>.
- Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R., 2019. Explanatory learner models: why machine learning (alone) is not the answer. *Br. J. Educ. Technol.* 50 (6), 2943–2958. <https://doi.org/10.1111/bjet.12858>.
- Sasaki, R., Goff, W., Dowsett, A., Paroissien, D., Matthies, J., Di Iorio, C., Montey, S., Rowe, S., Puddy, G., 2020. The practicum experience during Covid-19—supporting pre-service teachers practicum experience through a simulated classroom. *J. Technol. Teach Educ.* 28 (2), 329–339. Retrieved March 22, 2021 from: <https://www.learntechlib.org/primary/p/216244/>.
- The Institute for Ethical AI in Education, 2021. The Ethical Framework for AI in Education. Retrieved from: <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>.
- Touretzky, D.S., Gardner-McCune, C., 2022. Artificial intelligence thinking in K-12. In: Kong, S.-C., Abelson, H. (Eds.), *Computational Thinking Education in K-12: Artificial Intelligence Literacy and Physical Computing*. The MIT Press, Cambridge, MA (Chapter 8).
- Trust, T., Carpenter, J.P., Krutka, D.G., Kimmons, R., 2020. #RemoteTeaching & #RemoteLearning: educator tweeting during the COVID-19 pandemic. *J. Technol. Teach Educ.* 28 (2), 151–159. Retrieved March 22, 2021 from: <https://www.learntechlib.org/primary/p/216094/>.
- United Nations Committee on the Rights of the Child, 2021. General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment. Retrieved from: <https://digitallibrary.un.org/record/3906061?ln=en>.
- UNESCO, 2018. The ICT Competency Framework for Teachers (ICT CFT) Version 3. UNESCO.
- Vuorikari, R., Punie, Y., Cabrera, M., 2019. *Emerging Technologies and the Teaching Profession: Ethical and Pedagogical Considerations Based on Near-Future Scenarios*. Publications Office of the European Union, Luxembourg.
- Wing, J.M., 2008. Computational thinking and thinking about computing. *Phil. Trans. Math. Phys. Eng. Sci.* 366 (1881), 3717–3725.
- Yadav, A., Krist, C., Good, J., Caeli, E.N., 2018. Computational thinking in elementary classrooms: measuring teacher understanding of computational ideas for teaching science. *Comput. Sci. Educ.* 28 (4), 371–400.
- Yeh, Y.-F., Hsu, Y.-S., Wu, H.-K., Hwang, F.-K., Lin, T.-C., 2014. Developing and validating technological pedagogical content knowledge—practical (TPACK-practical) through the Delphi survey technique. *Br. J. Educ. Technol.* 45 (4), 707–722.
- Zolfaghari, M., Austin, C.K., Kosko, K.W., Ferdig, R.E., 2020. Creating asynchronous virtual field experiences with 360 video. *J. Technol. Teach Educ.* 28 (2), 315–320. Retrieved March 22, 2021 from: <https://www.learntechlib.org/primary/p/216115/>.

The relationship between knowledge exchange and teachers' use of mobile technologies in the mathematics classroom

Kinga Petrovai, Gabriel J. Stylianides, and Niall Winters, University of Oxford, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	272
Factors that impact teachers' use of technology	273
Developing teachers' use of technology	274
Teachers' learning needs	274
Role of professional development	274
Professional learning networks for teachers' development	275
Methodology	275
Group meetings	275
Purpose	275
Conducting the group meetings	276
Developing the group meeting coding scheme	276
Results	277
Factors	277
Knowledge Exchange addressing factors	281
Group meetings	282
Addressing factors	282
Discussion	284
Creating learning networks in schools	284
Curating tablet resources for teachers	285
Conclusion	285
References	285

Schools are increasingly incorporating mobile digital tools into classroom teaching that have diverse uses. On the one hand, there has been some criticism that these digital tools are used to replace physical resources (e.g., books) without making teaching more effective. On the other hand, some research on the use of digital tools in mathematics classrooms has shown that they can have a positive impact on teaching and learning. A deeper understanding of the factors impacting teachers' use of digital tools is needed to help better realize the potential of digital tools to influence classroom practice. This paper contributes to this domain of research by investigating the factors impacting the way four secondary mathematics teachers in England have used tablets in their teaching for a period of 18 months, and the role knowledge exchange played in their developing use. The study introduced a structured method of knowledge exchange, which used guiding questions to facilitate teachers' discussions and encourage knowledge exchange, in order to develop teachers' use of tablets in teaching mathematics. Our analysis is based on individual interviews, classroom observations, post observation interviews, and conversations among the teachers in group meetings which we organized to enable teachers to share their knowledge and experiences in using tablets. In this paper we present the aspects of *factors* and *knowledge exchange* of the Teacher Adoption of Digital Technologies framework, which we developed to articulate and examine how digital tools can be valuable to teachers even when they may appear similar to physical tools. Our findings show that there are seven categories of factors that impact teachers' use of digital tools and that a structured knowledge exchange among teachers can help teachers address underlying issues related to these factors. Implications of teacher professional development and future research are discussed in light of these findings.

Introduction

The Organization for Economic Co-operation and Development (OECD) has declared that "despite considerable investments in computers, Internet connections and software for education use, there is little solid evidence that greater computer use among students leads to better scores in mathematics and reading" (OECD, 2015, p. 145). This outcome raises questions regarding the reasons why powerful technologies, that have the ability to perform complex computations and demonstrate concepts in new ways, have not had an impact on scores in mathematics. This is particularly pertinent as the history of mathematics—which is the discipline this study is focused on—has examples showing that the availability of certain tools has influenced the course of the conceptual development of mathematics as a scientific discipline (Laborde and Sträße, 2010).

It is increasingly being recognized that more focus needs to be placed on the teacher's role in how to enhance teachers' professional knowledge in relation to the use of new technologies (Noss and Hoyles, 1996, p. 184). However, traditional professional development, the most widely used way of helping teachers integrate new technologies, is often ineffective (Opfer and Pedder, 2011) and is also hard to scale up (e.g., because of financial constraints) to help large numbers of teachers. The approach of creating opportunities for teachers to meet and exchange their knowledge and experiences around the use of technology in their teaching is easier to scale up and has been found to hold some promise (Clark-Wilson et al., 2015; Ertmer and Ottenbreit-Leftwich, 2010). Indeed, teaching networks have been found to be extremely useful, both in person and online, by many teachers (Trust et al., 2016) in that they provide a peer community in which knowledge, experiences, and resources can be shared. Unfortunately, though, having a learning community either online or in-person, is not the norm for most teachers.

One of the reasons, as Joubert (2012) suggests, for why technology is not incorporated effectively in education, include research not sufficiently accounting for the interests, issues, concerns, and developments of the researchers and stakeholders in mathematics education. This study aims to contribute toward addressing these gaps by focusing on key stakeholders, being the teachers, and their experiences of using technology in their daily classroom teaching. There also needs to be further research in naturalistic school settings on the link between teachers' technology use, particularly technology that offers multiple non-mathematical programs like tablet, and their classroom instructional practices (Rakes et al., 2006) as well as how teachers become absorbed into these existing and evolving practices (Merchant, 2012). This was the approach used in this study, where all interactions by the researcher were designed to have minimal interference with the natural setting of the school.

The broader study from which this paper is drawn (Petrovai, 2018) was designed to address the concerns and opportunities outlined above. In partnership with a comprehensive British school that has implemented a one-tablet-per-student program, the study investigated the way four Year 8 (12–13 year-olds) mathematics teachers' use of, and views on, tablets in their teaching developed over time, and the factors impacting them. The teachers in the study were given pseudonyms to mask their identity and gender. The study was situated in a natural school setting, with the added dimension that teachers met four times as a group to discuss their use of, and views on, tablets while exchanging knowledge about using tablets in their teaching. These meetings were facilitated by the first author, whose role was to ask the teachers questions that guided them to reflect on their practice and discuss ways of using tablets in a group setting, but not to provide any information. The data collected included audio recordings of the group meetings, individual teacher interviews at the start of the study, classroom observations, and short interviews with teachers directly following each observed lesson.

The result of the study this paper is drawn from was the development of a framework called the *Teacher Adoption of Digital Technology* (TADT) framework (Petrovai, 2018). This framework was designed to capture the teachers' experience of using a digital technology (in this particular case a tablet technology) and examine how this use developed over time. The TADT framework is intended to be a living framework, that can expand and change, used as a guide to facilitate conversations among teachers to help them learn from each other. There are four distinct parts of the TADT framework:

1. **Use** is broken down into Laborde (2001) framework of instrumental evolution. As this framework is progressive, with the tasks using technology in progressively more complex ways, the categories are represented in step format. This step format is to clearly indicate that the categories of tasks are progressive.
2. **Views** captures the ways teachers see tablets and their use of them.
3. **Factors** are a cumulative list of factors that teachers indicate as important in defining the way they use and view the tablets in teaching mathematics. Although the factors in the initial framework were drawn from the literature, additional factors emerged from the particular teachers in the study.
4. **Knowledge Exchange** captures the instances when teachers share exchange knowledge during the group meetings on how they use and manage tablets in their classroom.

This paper focuses on the Factors and the Knowledge Exchange portions of the TADT framework, that is, it aims to understand the factors that impact teachers' use and views of digital technologies as well as how structured conversations among teachers can help in knowledge exchange and addressing the factors impacting them. The literature in which this study is situated is outlined in the following sections.

Factors that impact teachers' use of technology

Tools have greatly influenced how mathematics is understood and taught in the classroom, an example of this is the ruler and compass for the development of geometry (Laborde and Sträße, 2010). However, there are questions on the level of impact one of the most recent tools—mobile technologies—may have on the way mathematics is taught and learned. Mobile technologies are increasingly being introduced in schools with little evidence to guide how they are implemented (Kiger et al., 2012).

There are numerous factors that may impact the way teachers use technology. In a study investigating the integration of online learning in 26 Australian schools, lack of time was found to be a key factor (Neyland, 2011). In this same study, teachers were found to be very frustrated by technical issues they encountered, both with the infrastructure and the network. Moersch (1995) discusses five distinct problem areas of computer usage in education: (1) staff development was usually insufficient and misdirected; (2) computers tended to be used for isolated activities unrelated to central instructional themes or concepts; (3) computer usage was one step removed from the classroom teacher; (4) technology usage sustained the existing curricula rather than being a catalyst for change; and (5) technology plans failed to establish a link between the need for technology and identifiable instructional priorities. In his 2007 study, Franklin indicated additional barriers to success. Even teachers who were well educated in the use of educational

technologies and graduated from a university that won numerous awards for preparing teachers to teach with technology, those teachers did not always use their skills once they started teaching. Several authors have also written about the lack of effective teacher training on the use of classroom technologies (Guha, 2003; Mentz and Mentz, 2003; Tsitouridou and Vryzas, 2004; Pelgrum, 2001).

There are many ways of looking at factors that impact teachers, which can vary depending on a number of reasons such as the training teachers received, the type of technology they are using, the school environment, or the teacher's personal experiences and beliefs. The research reviewed in the previous paragraph often points to factors that are not necessarily ones given by teachers as impacting their choices. Although all factors need to be understood, it is important that the factors teachers feel impact their decisions of if, and how, they use technology in the classroom, are also understood. In this study the factors that are captured are those that the participating teachers themselves raised as impacting their use of tablets. Gathering this information gives insight into their experiences but can also be used to address their further development.

Developing teachers' use of technology

There can be many reasons and factors contributing to why teachers may, or may not, develop their use and views on tablets over time. Even if they do develop in some way, research indicates that any change without some type of assistance takes a long time (Laborde, 2001). There are many ways that professionals can learn in the workplace, but they are not without challenges. Schools can be particularly challenging places as there are many different demands placed on a teacher's time. This section first outlines the literature on teachers' learning needs with regards to integrating technology into the classroom, followed by the role of professional development in schools, and thirdly the concept of Professional Learning Networks and how they can be helpful to teachers.

Teachers' learning needs

Both the success of integrating technology into the classroom, and the degree to which students can benefit from a technology-enriched environment, depends largely on what the teacher does with these tools. This is well understood, and the development of teacher competencies for technology use has been the focus of current educational reforms worldwide (Chronaki and Matos, 2014). Tools are not neutral: they change the actions of the users as well as significantly affect their conceptualization of reality (Vérillon and Rabardel, 1995). Central to the idea of integration of technology in teaching is the understanding and developing of teacher professional development (English, 2009). Several authors have written about the lack of effective teacher training on the use of classroom technologies (Guha, 2003; Mentz and Mentz, 2003; Tsitouridou and Vryzas, 2004).

The teachers in our study will discuss ideas together, but will design the lessons outside of the group meetings. Some may assume that this type of learning is only needed for teachers who are unfamiliar with technology. However, Ma et al. (2005) found that, although young teachers have grown up with computers, they may not automatically use technology effectively to improve the way students learn. It is unlikely that students, who have grown up with digital technologies, will learn to use computers in ways that benefit their learning as adults, unless their teachers have modeled effective technology use in schools (Fuller, 2000). This is important for our study, since some of the participating teachers were young enough to have had digital technologies in their school as students, however that does not mean they learned to teach with them. Like with any skill, guidance and training is needed for all teachers to know how to use technology in a way that will elevate their teaching and enable greater learning. Professional development plays an important role in this type of growth.

Role of professional development

As mentioned above, the main audience for teacher training on how to use technology in the classroom is not limited to those who are unfamiliar with digital technologies. Even among teachers who use technology in their existing instructional practices, some have found that they rarely take full advantage of its potential or use the technology to teach in new ways (Dawson, 2012). One factor may be the shortcomings of professional development and lack of support for teachers, which is already well documented (Ertmer and Ottenbreit-Leftwich, 2010; Opfer and Pedder, 2011). In mathematics, as in other STEM (Science, Technology, Engineering, and Mathematics) subjects in the K-12 classroom, there have been found to be significant challenges for teachers. Steep learning curves and the challenge of gaining confidence and skill at guiding students through complex activities within the constraints of the K-12 classrooms all add to the difficulty of using mobile devices (Parker et al., 2015). It is not that teachers are luddites, it is often likely that educational practitioners are already using 21st century methods and strategies in varying degrees and simply need additional support and professional development to ensure full application (Gunn and Hollingsworth, 2013).

Professional development needs to move away from the traditional model of sending teachers on an occasional course to develop their skills. Gunn and Hollingsworth (2013) have found that, "while sporadic professional development has been known to incite temporary adoption of new methods or innovations, widespread teacher change is only possible through thoughtful and systematic implementation of critical knowledge, skills, and strategies" (p. 204). There are many ways that such professional development may be achieved.

Heitink et al. (2016) stated that "[t]eachers' effective use of technology might be improved if they become better able to articulate the reasoning behind the use of technology in their teaching, share this reasoning with colleagues and confront their reasoning with findings from research." (p. 82) This suggests that one way in which professional development may be effective is by having teachers share knowledge and experience among themselves, through learning networks. Having such networks in the workplace may also be beneficial as Short (2014) found through a review of literature on learning in the workplace, that trust is extremely important and face-to-face connection with colleagues helps build this trust.

Professional learning networks for teachers' development

The notion of *Professional Learning Networks* (PLN), also referred to as Personal Learning Networks, was first coined by Tobin (1998). Tobin defined PLN as a "group of people who can guide your learning, point you to learning opportunities, answer your questions, and give you the benefit of their own knowledge and experience" (Tobin, 1998, para. 1). Such a network can take many different forms. These networks may include online chat rooms, conference gatherings, or formalized networks within an organization. What they all include is an ongoing communication and exchanging of knowledge among the different members. Communication has been found to be an integral part of learning, and in a study by Swallow (2015) it was found that communication between and among teachers and students was a consistent critical component in the integrating mobile devices into the classroom. Ruthven (2014) review of thirteen mathematics teacher networks found that most described some form of positive change in classroom practice.

Both in England and around the world, teachers who have not found the support they needed have taken the initiative to take their own professional learning and development in their hands by developing online PLNs (Trust, 2012). The flexibility of digital technologies to have anytime, anywhere availability of expansive PLNs, and their capacity to respond to educators' diverse interests and needs may have the ability to fulfill some of the professional growth requirements a teacher might have. Jones (2011) went as far as to say that "while access to computing power may shortly no longer be a source of an unbridgeable 'digital divide', differential access to networks of people that provide support for, and nurturing of, educational innovation via ICT may be an emerging new form of 'digital divide'" (p. 43).

Although there are many benefits, joining online PLNs does require a certain level of being digitally savvy and having the time to initiate such involvement. Incorporating the essence of a PLN into a school setting would remove the barriers of entry that some teachers may face. In addition, fostering a learning network among the teachers at a school has the added benefit that they share the challenges and opportunities of the context they are in. In addition, everyone would benefit from the knowledge members acquire outside of the school. As part of this study, we were influenced by the idea of learning networks, and we introduced and fostered a PLN among the participating teachers. This network was fostered through regular face-to-face group meetings among the participants facilitated by the first author. The essence of the group meetings is for the teachers to learn from each other. To have a quantitative representation of this learning, we will capture the instances they exchange knowledge with each other, which we will refer to as *knowledge exchange*. Exchanging knowledge among peers is important aspect of this study as it is very popular with teachers and it is an effective way of developing existing practice (Trust et al., 2016; Clark-Wilson et al., 2015; Ertmer and Ottenbreit-Leftwich, 2010).

Methodology

The broader study from which this paper is drawn had two phases. Phase 1, referred to as *Setting the Foundation* phase, aimed to identify the context in which the study was run and define the baseline for how the teachers used tablets, and their views on, tablets at the early stage of the study. The three methods used to achieve this were, school visits, classroom observations and teacher interviews. The results of Phase 1 informed the final design of the methods used in Phase 2, as well as the creation of the revised TADT framework. This enabled the methods of Phase 2 to be further embedded in the context of the regular daily setting of the school within which my study was situated. This was an important aspect of how the study was designed. Phase 2, referred to as the *Evolution of Use* phase, aimed to identify how over the course of six months the teachers developed, or stayed the same, in their use of, and views on, tablets in the teaching of mathematics. The three methods used to achieve this included classroom observations, post observation interviews, and group meetings. The methods that included data for Factors and Knowledge Exchange were the following: teacher interviews, post observation interviews, and group meetings.

- *Teacher interviews* were used to include the teachers' perspective, at an early stage of the study. Interviewing each teacher individually set a starting point from which any possible development in their use of and views about tablets might be seen over the course of the study.
- *Post observation interviews* were used to dig deeper into what the first author observed in the classroom and understand the teachers' views on how tablets were used, the reasons they chose to do what they did, and the factors that contributed to their choices.
- *Group meetings* aimed to give the teachers an opportunity to reflect on, and develop, the way they used tablets in their mathematics teaching and learn from their peers. The first author facilitated the meetings by asking specific questions to guide the conversation.

Group meetings

Purpose

As outlined in the literature review, teachers' professional development is a critical part of integrating technology into their teaching. Basing it on the literature on Professional Learning Networks, the group meetings are aimed to create a similar environment in which teachers can learn from each other and share their experiences of using tablets to teach mathematics. The three main reasons for having group meetings were as follows:

1. The meetings served as group interviews, which have the benefit of allowing participants to elaborate on each other's ideas and experiences, giving depth to the individual interviews.

2. The meetings served as a dedicated time in which teachers could develop their use of tablets in their teaching. Other studies have shown (Dwyer et al., 1990; Laborde, 2001; Ertmer and Ottenbreit-Leftwich, 2010) that without any training or guidance, it takes years for teachers to change and develop their use of new technology in their teaching. A major aspect of this study was to situate it in a school setting in such a way that would be sustainable for any school to adopt the framework that emerges. For this reason, rather than having specific external training for teachers, the group meetings served as peer training. There are three main reasons for this approach. First, the cost-effective nature of this approach would enable other schools to use this model. Second, it allows schools to tap into the diverse knowledge and experience that teachers already have, and share it among each other. Third, it ensures that the training takes into account the specific context of the school and the specific circumstances and challenges that the teachers face, which is likely to impact the way they use tablets in their teaching.
3. The meetings provide added value to the participating teachers. In the initial informal conversations teachers expressed a desire to learn and develop their use of tablets in their teaching, which is why having time to train was something that they found valuable.

Conducting the group meetings

The first author took on the role of the facilitator of the meetings. Her role was to set goals, ask questions, allow the conversation to develop, and pick at important points that were raised to understand the teachers as much as possible. The facilitator ensured that all teachers had a voice in the conversation and could express their thoughts. The goals were stated at the start of each meeting, and possibly gently nudged the conversation while it was taking place, which was the extent of my intended interference. It is possible that the information they provided was influenced by the facilitators presence, but great care was taken to minimize involvement. If the conversations were about the use of tablets in their teaching allowed conversations to develop. It was assumed that the teachers expressed what was most important to them. To capture the most honest data possible, it was important to allow the group meetings to be organic and allow teachers to express their thoughts and feelings.

In the following paragraphs outlines the themes of the questions asked during each meeting. The approach was to be flexible and alter the questions for the meetings following the second meeting based on the outcomes of the previous meeting.

Group meeting 1

The goal of the meeting was to establish a benchmark of how the teachers were using and thinking about the use of tablets in their teaching. To start the conversation, the guiding questions begin by asking the teachers' opinion of the tablet program and how they are currently using the tablets. The questions go further to ask how they do/would like to use the tablets and the reasons why. The meeting ends by asking the teachers what their expectations are of the group meetings.

Group meeting 2

The goal of this meeting was to think about concepts that will be taught in the upcoming lessons, which could benefit from the use of tablets. The questions went from reflecting and exchanging ideas on how tablets were being used, to discussing upcoming mathematical topics being taught that may benefit from using tablets. The final part of the meeting challenged the teachers to think about differentiation and how tablets can be used to address not only difficult mathematical concepts but also how they may address different type of students learning needs.

Group meeting 3

This meeting was aimed at continuing to push everyone's thinking about ways in which the tablets could be used in the classroom by digging deeper into more details. Teachers were asked to reflect on their recent practice, successes and challenges of using tablet, as well as new ideas they can use in the future, with prompting questions on the benefits, drawbacks and details of how it would help students.

Group meeting 4

The goal of the final meeting was to reflect on the journey taken by each teacher during the study and to discuss ideas for future use. Prior to the final meeting, the school unexpectedly decided to change the tablet program and remove tablets during the transition. The questions in this meeting focused on reflecting on how the teachers changed their use of tablets over time, the impact they saw on teaching and learning, and the value they saw in using tablets to enhance mathematical understanding.

Developing the group meeting coding scheme

The criteria for coding something as a knowledge exchange was when one teacher makes a statement about the way they use, or wish they could use, tablets in their teaching and another teacher added to the conversation either by providing ideas, sharing experiences, or giving instruction on how to use the tables in similar situations. This type of exchange identifies how the factor of working with colleagues can influence the teachers' considerations in how they use the tables.

Table 1 Definition of factors.

<i>Factors</i>	
Time	Time available to prepare or learn to use tablets
Resources	Resources, software or hardware, available to use tablets
Technical issues	Issues concerning technical challenges with the hardware, software, or infrastructure
Mathematics misunderstood	Issues related to actions or policy that does not fully consider the needs of mathematics teaching and learning
Support	Support provided to the teachers from the leadership or technical team
Management	Decisions made by school leadership that impact teachers' use of tablets
Classroom management	Classroom management concerns caused by the use of tablets

Results

There are many factors, and ways of looking at factors, that impact teachers, as shown by the literature, but addressing those factors at a particular school may not be relevant. There are also factors that are important but other research have not highlighted. This study found seven factors that the teachers indicated as having an impact on their work, as defined in [Table 1](#). As outlined in the previous section, the results came from individual teacher interviews, post observation interviews, and the knowledge exchange group meetings. The abundance of knowledge exchange conversations in the group meetings are categorized into three overarching themes. In these three themes of conversations, the seven factors that teachers expressed as impacting their use of tablets in teaching mathematics are addressed. In this section each of the factors and themes of knowledge exchange are discussed in more detail ([Table 2A–C](#)).

[Table 2A–C](#) presents the results of the study for Factors and Knowledge Exchange. As can be seen, the most frequently mentioned factors are slightly different for each method. [Table 2A](#) shows that classroom management and time were the two most frequently mentioned factors in the individual teacher interviews, each being mentioned by all the teachers. In the post observation interviews, [Table 2B](#) shows that classroom management and resources were mentioned significantly more than others, with 17 and 12 mentions respectively. In the group interviews, [Table 2C](#) shows that the most frequent factors discussed were resources and support, with each being mentioned by at least two teachers in each of the four group meetings. Although the frequencies are important, the greater importance is on the factors that emerged from teachers and the abundance of the knowledge exchange conversations in each group meeting. The following two sections explain in more details the results for Factors and Knowledge Exchange.

Factors

Factors emerged from all of the interactions with the teachers. During the group meetings most of the conversations among the teachers focused on the factors impacting the use of tablets in their teaching. This is not surprising as [Mumtaz \(2000\)](#) also outlined a multitude of factors that was synthesized from literature. One of the greatest factors impacting teachers' use of tablets was time because the "pressures of work inside and outside the classroom, timetabling does not allow time for such learning [about ICT]" (p. 334). This was certainly a factor that came up often among the teachers.

Time available for teachers to prepare lessons or learn and use tables was a factor mentioned by all teachers. In the initial interviews, Jordan said that she saw great value in tablets providing individualized learning, something she valued. Although she was incorporating some individualized learning with digital worksheets, but it was not to the extent she could imagine. However, it was the lack to time that she said really prevented her from implementing that, as she described.

Table 2A Results for individual teacher interviews.

		<i>Jordan</i>	<i>Sam</i>	<i>Alex</i>	<i>Charlie</i>	<i>Total</i>
Factors	Time	1	1	2	3	7
	Resources	1	2	1	2	6
	Technical issues		1	1		2
	Mathematics misunderstood		1	1		2
	Support					
	Management	2	1	1	1	5
	Classroom management	4	3	3	1	11
Knowledge exchange	Conversations in which help was offered or knowledge was exchanged					

Table 2B Results for post observation interviews.

[illegible]

Table 2C Results for group meetings.

		Jordan					Sam					Alex					Charlie					
		Group meetings																				
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Total
Factors	Time			1		1	1				2	2	1			3	1		1		2	8
	Resources			2	2	4	1	1	1	1	4	2	2	5	1	10	1	1	1	2	5	23
	Technical issues	2		3		5	2	2		4	1		2	2	5	1		1		2	16	
	Mathematics misunderstood	1		1		2			1	1					0			1		1	4	
	Support	1	2		1	4	1		1	1	3	1	3		3	7			1	3	4	18
	Management	2		1		3					0	1		1		2					0	5
	Classroom management					0		3			3	2	1	1		3	1	1			1	7
Knowledge exchange	Conversations in which help was offered or knowledge was exchanged	4		13	2	19		13	14		27	5	15	12	1	33	1	12	14	2	29	33

I can happily create traditional resources that are read on an iPad, such as a PDF, so that is still a step up from just giving a worksheet, but it's still not interactive. So, in terms of creating resources, that is something that would be useful, but I don't have the capacity at the moment to create an interactive resource that is built for an iPad.

Charlie also said that lack of time was the greatest factor that impacted her ability to research tools and resources to use on tablets. However, she said that having the first group meeting was extremely beneficial and it did start the ball rolling toward more integrated use of tablets in her mathematics class.

The biggest positive for me is that I have seen it work. I'm enthused enough to know that I would like to do some more. If I could cut down the amount of time, or find a shorter way to get really good resources, I would use it more often. It's the sort of think you feel that actually, you can have someone go through and give a list of resources.

Although Charlie was enthusiastic about the possibility of using tablets in her teaching, the barrier of entry was too high for her at the start of the project. The lack of time to explore the resources, lack of knowledge and experience in using tablets, and the lack of opportunity to receive guidance prevented Charlie from adopting the tablet in her teaching.

Resources, in terms of software or hardware available to use tablets, was the most frequently discussed factor in the group meetings. Specifically, the resource the teachers were discussing was the funding to pay for apps. The teachers often discussed that they knew of apps that would be good for their class, but the fact that there was a fee, regardless how nominal, made it an inhibiting factor. The teachers discussed the dilemmas of paying for an app that would rarely be used, or the cost of having every student purchase a costlier app. There were also discussions about whether it would be useful to buy only one version of an app for the teacher and project it on the board. However, it was agreed that this would make any app much less useful if students could not use them individually. This issue of fees on apps raises an important issue for tablets in the classrooms. There needs to be a school policy or process for how much funds can be devoted to apps, and how to select which apps to pay for. It seemed to be an unspoken rule, that only free apps would be used unless special permission was granted by the school. This presents a logistical hurdle and possibly prevents the best software from being used on these powerful devices.

Lack of quality apps available, or not finding the type of app the teachers imagined being useful for a lesson, also became a central discussion. Without knowing of a specific app or resource, all the teachers imagined ways in which the tablets would help their class. Jordan captured the essence of this when she said that: "Sometimes we have an image but no app for it. Like in algebra with the balancing scales, but we don't have scales." But the frustration of not being able to find exactly what they were looking for became a major debilitating factor for the teachers. Jordan stated that "there's an equal measure of over and under abundance. There is an overabundance of some stuff." Not only is there a sea of resources that teachers need to search through, but what is available is disproportionately distributed to lower mathematical skills. The teachers talked about not being able to find resources for their higher achieving students in Year 8, when they saw many wonderful apps for younger grades, and also having some topics that they could find any apps for. Even when they did find apps they did not fit all the requirements they had, as Charlie describes her experience:

The interesting thing is that there are hundreds to choose from but when you start looking at them, they really don't do what you want them to do. They hit the surface but what we really need is to get to the conceptual bit. I didn't find anything that was the best, but I found something that would do. I went down the route of what if I paid a bit to get something better.

It became clear that having access to a curated list of useful resources would be essential for the teachers, but none of them have found a list of that type of material. There are certainly lists of the top ten best apps for this, or that, which can be found on some blogs or websites. Jordan did mention having found some apps on similar lists, but these lists are not well aligned to what the teachers are teaching, nor are they easy to find for topics they need them for. Rather than lists, what the teachers were describing was a customized curation that can be easily accessed for the topics they were teaching. Such a resource would greatly reduce the amount of time they were spending to find resources, and the barrier of entry for teachers who may not know what they are looking for.

Technical Issues concerns technical challenges with the hardware, software, or infrastructure faced by the teachers. This was one of the issues that was given by all teachers in each of the data collection methods. The issues ranged from insufficient infrastructure in terms of losing internet connection to the unpredictable nature of how apps work. Teachers expressed frustration at not being able to predict how the technology would work even when they prepared and practiced using it before the lesson. This unpredictable nature of the technology not only adds a lot of time to the preparation needed by teachers, but the extra work they spend on diligently preparing does not eliminate the disruption caused by the technology not working during the lesson. It is this unpredictability that was given by teachers as they reason that they have to prepare two lessons, one if using the technology and another traditional lesson in case the technology stopped working. Although this is a difficult factor to address, as there are many reasons for this unpredictability, but from the amount of frustration expressed by the teachers it is easy to see why this factor alone might hinder the initiative to use technology in the classroom. Possibly as technology advances, and the teachers and school become more accustomed to using the technology, these issues will subside, but this was a significant inhibiting factor expressed in this study.

Mathematics Misunderstood, refers to factors that have the characteristic of not fully understanding the needs of mathematics subject teaching and learning. One example given by Alex, was that the training provided for teachers did not address the teaching needs of mathematics subject teachers. The apps that were being taught to be used on the tablets were not relevant to mathematics teachers. Sam also mentioned that she, or the school overall, have not found really mathematics specific applications to use.

A lot of the conversation in the group meetings focused on was that mathematics teachers felt disconnected from other subjects and the leadership that was implementing the one-to-one tablet program in the school. "We do swim in a slightly different boat a lot of the time" stated Alex, to which everyone agreed. The professional development offered to the teachers seemed to miss the requirements that the subject of mathematics has. Jordan expressed frustration at the IT director, who tried to insist that all teachers try to be creative and use programs that make movies and collages as a way of engaging students. The mathematics teachers agreed that they could only do so much with these programs that do not offer any insight into the understanding of mathematics and on many occasions they found traditional lessons conveyed the learning more effectively. As an example, the teachers mentioned iMovie program as being highly recommended by school. The teachers said that they could see how this might be useful in other subjects such as history, but they could not do too much with this program. They said that although it might be captivating for the students to make a short movie demonstrating their understanding of a mathematical concept, their mathematical knowledge could not be pushed too far using such programs.

The teachers said that this disconnect was actually preventing them from using tablets at all and some of them have been using tablets significantly less than they did early on when the program was introduced. The narrow vision that the leadership was promoting was inhibiting subjects like mathematics. Rather it should have been recognized that mathematics has special needs. Sam suggested that if there were someone made responsible for finding applications specifically for mathematics, that would make a big difference in enabling the teachers to use tablets more widely and successfully.

Support provided to the teachers from the leadership or technical team was an important factor that was raised in different contexts. One important example that unfolded over the course of the study was Alex's experience. This is a significant example because Alex was very enthusiastic supporter of the tablets being introduced into the school. She was a young teacher who said that she was given a computer to use when she was in school as a learning aid because she has dyslexia. She said that the computer made a big difference in her learning and she was excited to use tablets to help her students learn better. But in the year since the tablets were introduced and this study began, Alex had almost altogether stopped using tablets for many reasons.

At the first group meeting Alex was very encouraged and enthusiastic about learning from the other teachers. Her enthusiasm sparked an idea, but the lack of support from school leadership was the factor that impacted her. In the first meeting, hearing the concerns other teachers had, Alex started getting an idea.

Alex: Just thinking about what these guys are saying. Jordan, imagine a world where you didn't have to research for digital worksheets that you could write your own. So I can write the questions that I want my class to answer

Jordan: Just a static question and they type in an answer?

Alex: Yes, they type in answer and then they know if it's write or wrong

Jordan: Maybe you can use Socrative so maybe it's worth us thinking about how we can do that and they can do it self-paced

Alex: But just listening to you I can't help thinking it would be great if we could create our own

Jordan: Yes yes!

From this one meeting, Alex picked up on a need that she and the other teachers were experiencing and already started to test her idea with the other teachers. She knew her boyfriend was a good programmer, so she asked for his help. Together, they designed the program and within a few weeks he made it. Alex was thrilled to see her idea coming to life and she knew this would be a great help for her and the other teachers. She describes the ins and outs of process when the other teachers inquired about when the program would be ready:

It's all there, it's all been written, I just can't get the IT department. It's just that classic thing. I went to finance and they said that before they pay for it, I need to go to IT and they need to sign off on it that it's for work. That was a long time ago. I even got IT manager involved and I said this is the way we want to go in math, and every week he said he will get back to me, but he still hasn't replied. This is the way IT department work, they get things done very quickly that they want to get done but not others.

It was not a matter of funding, as the license was quite inexpensive, and it was not a matter of other teachers not supporting her idea, because they were enthusiastically waiting for her program to be available. Yet, over the months of this study, Alex grew more and more frustrated at the hurdles she had to jump over. Before the tablets were taken back by the school, Alex was still trying to get the school to install this program, and may have continued to do so until she was successful, but all the tablets were recalled in the school. The lack of support Alex faced dampened an incredible turnaround in enthusiasm for the tablet program.

Management are factors that relate to decisions made by school leadership that impact teachers' use of tablets. Some of these decisions included what paid programs teachers had access to, whether students each had their own tablet or there were class sets, and the type of professional development is offered.

The teachers in this study had several management related factors. One example is that the school had decided to give one laptop to every study, that they were able to take home. The impact of this decision was that often students forgot to bring their tablets, did not have them charged, or did not load the apps that the teacher requested to be loaded for the day. The teachers often talked about how they came to prefer the idea of class sets, rather than a tablet for every child. The ownership that was it gave the students had the possibility of causing classroom disruptions. In several of the observed lessons in this study, it was noted that the teacher had to pass out her tablet, group students together for activities that were intended to be individual, and on occasion even hand out her personal smart phone because several students did not have their tablets. The trickle effect of this management decision to provide each student with a tablet caused unexpected inhibiting factors in the classroom.

Classroom management concerns, such as disruptions or students being distracted, caused by the use of tablets were given by teachers, often in connection to another factor. Some examples involve students being disruptive when a technology failure allowed for a pause in the lesson. The teachers were hesitant to use tablets in their less well behaved classes as they feared losing control over the class. To some extent this concern may be alleviated as everyone becomes more accustomed to using tablets, but a loss of control is significant as an important element of a teacher's job is to keep the class focused on the lesson.

Another example of classroom management factors relate to the example given previously about the students having ownership of their individual tablets. In one example Jordan talked about the compulsive ownership students had over their tablet, many having to physically touch the tablet even when it was not being used. Something she thinks might not be an issue if it were a class set. Teachers also complained that students had filled their tablets with games, did not remember their passwords, or were distracted by the programs they had on them. Although classroom management is a key concern for teachers even when there is no new technology introduced, any change or disruption does have the added impact of disrupting the flow of a class.

Knowledge Exchange addressing factors

Knowledge Exchange is defined in this study as conversations in which help was offered and knowledge was exchanged. It draws on the extensive literature that teaching networks, where teachers learning from each other, have been found to be extremely useful by many teachers in developing their professional practice (Trust et al., 2016; Dawson, 2012). In this study *Knowledge Exchange* is captured by counting each time help is offered and knowledge is exchanged between teachers, as defined in Table 3. By not capturing the context of the information exchanged, the focus is on the exchange between the teachers rather than judging or evaluating the help or knowledge being exchanged.

Table 3 Definition of knowledge exchange.

Knowledge exchange	Definition: Conversations in which help was offered or knowledge was exchanged. Example: Anne “after a week of searching for an app that would dynamically demonstrate the properties of a circle, I finally found one.” Tim “what is it called? I have been searching for one to use in my next class.” Anne “I’ll send it to you.” Example: Ken “I tried to use the graphing app on the tablets but I wasn’t able to graph more complex equations.” Sarah “I found a way to do that last week, let me show you.”
--------------------	---

Group meetings

All the teachers in this study expressed a great appreciation for having time to learn from each other. In her initial interview, Sam said would help her improve her practice was the opportunity to share and learn from her colleagues. Sam felt that it would be useful to learn how her colleagues dealt with difficult classroom management situations caused by the tablets as well as to get tips on the tools they found useful in their lessons. Unfortunately, she said that the nature of the school schedule does not give teachers the opportunity to share their learning, which she hopes the group meetings will provide.

The four group meetings quickly developed into dynamic discussion in which the teachers opened about their experiences using tablets, their worries, and shared their knowledge. In this learning network, the teachers exchanged many ideas on how to use the tables in the classroom, resources, and hands-on demonstrations to teach each other how to use the tables. At the start of the first group meeting the first author asked the teachers what they wanted to get from these group meetings. Unanimously they said that they wanted to get ideas and inspiration on how to use tablets. Jordan said that she also wanted to have shared experiences, where she got to talk to others who were further on in their practice of using tablets in certain ways. As she explained, everyone agreed, they were looking to learn from each other. Charlie added that she just wants to grow her confidence. When the first author asked her how this group will help with that, she said the following:

It’s quite nice that not everyone here is a real wizard. Because when you go to any of the training I just sit and feel completely and utterly overwhelmed. I watch it all going on and it’s like I’m not even in the room, I’m on the ceiling thinking oh my, I will never be able to do that. So it’s nice to hear that other people also feel that.

This speaks to the importance of the shared experience and understanding that was central to the design of the group meetings, which intend to create a professional learning network for the teachers. Everyone felt that they had a lot to learn about the use of tablets in teaching and although they had different levels of experiences in this, they also had something to share.

The open-ended structure of the meetings allowed for these conversations to grow and flow as the teachers felt it was necessary, allowing them to discuss the aspects of teaching with tablets that was most important and valuable to them. With this structure, general themes emerged for each meeting. The first meeting resulted in discussing the factors that impacted the teachers’ use of tables, the second meeting was focused on exchanging ideas on how to use the tablets, the third meeting was the most vibrant and rich in ideas on how tablets could be used and exchanging ideas and resources on using tablets. The fourth meeting focused on factors that impacted the use of tablets and reflecting on how tablets can be used, this was a more reflective conversation because the tablets had already been taken out of the school. It was an unplanned change in the school policy to turn the one-tablet-per-child program into classroom sets. The tablets were revoked from the school for several months to refurbish. Had there been an opportunity to have more meetings, it is possible that the teachers would have progressed to discussing ways in which the tablets could be used in increasingly more innovative ways. Unfortunately, the tablets were suddenly removed from the school before all the meetings could be completed.

All the teachers were very engaged in each of the meetings, and the conversations flowed with ease and ideas regularly grew bouncing from one person to the next. In terms of building knowledge, the third meeting was the most productive and vibrant. This hints at how rich, productive, and innovative these meetings could have become had the tablets not been removed from the school after the third meeting. However, the learning network that was created with these teachers during these meetings still had significant impact on their use and views of the tablets.

Addressing factors

The knowledge exchange conversations during the group meetings can be categorized in three distinct ways in which they address the factors that teachers were mentioning throughout the study, by sharing resources, providing support, and speaking the same language. Some factors were addressed by one or more of these categories, but these are the overarching themes of the conversations that addressed the factors. Each is described in more detail.

Sharing resources

Conversations where teachers shared useful resources in the apps or websites they were using or the strategies they incorporated in how to manage tablets in their lessons. One example of a strategy was when Alex shared that she always instructed students to have tablets face down on the desk to avoid distraction.

Providing support

Support was provided in a variety of ways, including sharing stories of similar challenges as well as giving hands on help on how to use a certain app or program. The group meetings often served as making the connections to enable further support to be provided after the group meetings. Often one teacher would offer to set up a time outside of the meeting to help another teacher learn to use a certain program.

Speaking the same language

Having the common ground of all participants in the group meeting being teachers in the same school, same subject, and same year group, provided a lot of opportunity to share valuable resources and provide support. The importance of having a common understanding of what they were each experiencing was often mentioned and also given as something missing in other professional development settings.

Although the four teachers were already colleagues in the same school, the time spent on discussing experiences and ideas created common ground. The first author rarely had to ask questions or direct the conversation because the four teachers easily built on each other's answers. There was a sense of understanding and camaraderie as they all shared the same goal of growing their practice in using tablets to teach Year 8 mathematics. By the end of the conversation there were ideas and resources being shared among the teachers, and some of them were making plans to meet at a later date so they could spend more time demonstrating to each other how to use certain programs on the tablet.

As the meeting came to an end, the first author wanted to understand what the teachers felt about having group meetings and participating in this project. She encouraged the teachers to be critical and give suggestions as the structure may help other teachers develop their own learning networks in their schools. The first author wanted to find out if, and how, these meetings impacted them. The teachers unanimously agreed that the meetings were extremely valuable to their development of using the tablets, and as Alex described in the first group meeting, she had stopped using the tablets altogether before joining my study. Some of the reasons why the teachers found this format useful are best described in their own words.

Jordan: It was a stimulus to think and try. It was like well it's coming up, if nothing else it was the fact that the meeting was coming up and required some thinking and planning. I like the idea of using an iPad, sometimes you get bogged down because it doesn't work.

Sam: And then it takes someone else to get inspired again.

Alex: I just enjoyed the dedicated contact, dedicated time to improving this, and being on a separate site and the lack of conversations I have on anything related to pedagogy, means that this was like ok for the next hour I get to think about this. Sometimes in my own teaching practice you just get a nugget. It might be a single phrase of something someone said that gave you the confidence to go and try it and then use it in your own way and push it forward.

Charlie: What I would say is that this was absolutely fundamental to me doing anything. If I hadn't come I wouldn't have tried. The main thing was that everyone was massively supportive and didn't make me feel like I was horrendous. Also, I think it's nice having a small group and everyone is not techy, well Jordan is. But in trainings here, everyone is going a million miles an hour showing things on the iPad and I can't keep up with it. Actually, I don't take much in.

As Charlie stated, the format of traditional professional development courses can be overwhelming and out of context for some people. Research has also shown that these forms of professional development are less effective than anticipated. This does not mean they should be eliminated. Having external professionals, who understand the technology and are avid users, can share a wealth of information that would be useful for teachers. But maybe that benefit is missed if those sessions are not supported by regular opportunities for teachers to discuss and share their understanding among themselves in these type of learning networks. Sam articulated this when she described what she thought would be helpful for the teachers moving forward in using tablets in their teaching.

What would really help and what would really be useful would be if we continued doing that and continued to meet. This is the only time we can focus on this. The shared document, ya it's great, but we need to look at it at the same time. We all need to be thinking at the same time and responding at the same time.

Having dedicated time is not the ultimate solution, but we would suggest it is the final piece in the puzzle that makes all the other great resources stick together, enabling teachers to use technology in the best way possible. This format, of having dedicated time for guided discussions, provides the glue and the spark, as Charlie summed up "Having these conversations sparked off the thinking. The individual discussions overlapped and sparked off together."

Discussion

Throughout this study, seven key factors emerged that contributed greatly to the teachers' considerations of how, and why, they used and viewed tablets in their teaching of mathematics. It became clear to us that the following two initiatives—creating learning networks in schools and curating tablet resources for teachers—may address the factors that teachers indicated as having an impact on the way they use of tablets in mathematics. With the teachers indicating a positive experience of having structured knowledge exchange meetings, implementing learning networks would continue that experience, as well as addressing some of the factors that emerged. The use of the TADT framework, that was developed in the study this paper is drawn, can be used as a tool to facilitate the learning networks among the teachers. Curating tablet resources for teachers would also address several of the factors that emerged, particularly the factors involving time, support, and resources. Together these two initiatives may address many of the challenges that teachers faced in using tablets to teach mathematics.

Creating learning networks in schools

Developing a learning network within a school, such as the group meetings used in this study, can be the glue that holds together all the other valuable resources and investments that schools make in the implementation of technology in teaching. Having a technology manager in the school, external professional development, and purchasing resources for the tablets, would all be more valuable if teachers had the opportunity to discuss their use, and experience of using technology in their teaching, and develop their skills. Having this type of forum allows teachers to process what they learned in professional development classes, help them learn to use the electronic resources that the school purchased, and discuss the benefits and opportunities of the initiatives introduced by the technology manager.

Through this study, it became clear that giving the opportunity for teachers to work together, to develop their use of tablets, greatly impacted how much they used the tablets, their level of innovation, and it also improved their self-efficacy. Alex and Charlie are good examples of this. Alex stopped using the tablets before this study started and Charlie never got the courage to start using the tablets. During the first group meeting Charlie and Alex spoke at length on why having a learning network would be valuable to them. This is one example of many in which teachers highlighted the value they saw from having dedicated time to discuss and learn from each other.

Charlie: When you go to any of the trainings I feel completely and utterly overwhelmed. I watch it all going on and I'm not in the room, I'm on the ceiling, and I'm watching it all going on and I'm thinking "oh my god I would never be able to do that". So it's just nice to be able to hear that not everything is absolute instant hit.

Alex: I am hoping through doing this I can spend some time literally in a room with people who are also trying to teach math with iPads because of my timetable this year I don't spend any time on this site and I teach in a vacuum. For me this is literally the time I get to interact with people and prioritize these thoughts.

Although there are many different professional learning networks online for teachers, fostering this type of network in a school provides many advantages. To begin with, it incorporates teachers who otherwise would not find such a network online. In order to participate in such an online community, someone would need a certain level of engagement with technology. Having this network within a school, includes everyone who wishes to participate, not only those who are more comfortable using technology.

Time is another factor that would greatly inhibit many teachers. As was found in the study, teachers are greatly restricted by the amount of time they have to prepare their lessons. Without easily accessible and pre-arranged group meetings, finding time to engage in online communities is unlikely to happen for many teachers.

Relevant context is another important reason to have learning networks for teachers in schools. As previously mentioned, different school contexts have different obstacles and opportunities in the way technology is used in the classroom. Having teachers from the same school, who know the same students, meet regularly and learn from each other, means that they have the same cultural reference and context in which they apply the technology. For all these reasons, having a learning network for teachers within a school can provide a glue for all the other resources that schools provide to support the technology used in teaching. Teachers can leverage their different skills and knowledge to each develop further their use of technology in teaching mathematics. Schools spend significant resources on technology, the supporting structures, and the professional development to train teachers to use the technology. Strategically leveraging the knowledge and imagination of the teachers in the school can make the difference between a successful integration of technology or wasting valuable investment on technology that will quickly fade into the back of the classroom. Providing the type of group meeting structure that was in this study, could be the glue that holds together all the other investments they have made in technology.

One important direction for future research would be to investigate how the group meetings could develop over time and how the knowledge exchanged in the meetings might impact the way teachers use technology. It would be valuable for future research to

incorporate more group meetings over a longer period, and for group meetings to also increasingly focus on specific mathematical topics. This may enable teachers to design innovative uses for tablets that can help students to interact with mathematics in ways that would not be possible without a tablet.

Curating tablet resources for teachers

The greatest factor inhibiting teachers' use of tablets in this study was resources. Everyone agreed that there was an abundance of resources available, but the concern was the quality, usability, price, and disparate sources of countless resources. This factor impacted the amount of time teachers had to spend to find what they needed and also the self-efficacy was hindered as they got lost in a sea of apps, webpages, and programs. When, during a group meeting, we tried to find an app to suit a particular need teachers were discussing, Charlie summed up the frustration teachers face in the when she said that apps are "incredibly time consuming to find. We probably just spend 15 min demonstrating that we are lost in a world of a million apps, or no apps."

Having a curated list of useful apps, websites, and programs that teachers can use was a desire expressed in different ways throughout the course of the study. The teachers have never found a good forum to curate material online, but they also expressed the need for this to be a role that the technology manager could fill. They particularly made it clear that it was critical for anyone who did this, to be very sensitive and knowledgeable about the specific subject needs. The challenge is to know what mathematics teachers need and have specific understanding of what features programs need to have.

Although the teachers specifically mentioned that they would like the technology manager to help them curate the apps available to them, this role could also be taken on by the software industry. It would be useful to have a reliable and well-informed resource where mathematics teachers could find programs that best suited the topic they were teaching. In addition to curation, it might also be beneficial to build into programs clear teaching guides that include the rationale for the design of the program that helps students learn mathematics. This type of guides may help teachers select programs more effectively and also learn ways of implementing them in their class. This type of material, that aims to promote teacher learning alongside student learning, is often referred to as *educative*. An example of such an educative material was used to help teachers teach proof in mathematics (Stylianides and Stylianides, 2014) by providing annotated PowerPoint slides with the annotations detailing what to do as well as why/how to do it. These types of instructions, included in programs and apps, could greatly aid teachers in how they use technology. This is only one part of the greater ecosystem required for integrating technology into teaching, which includes professional development, resources, and infrastructure, among others, but it could be a significant contribution to helping teachers find and use the resources they would like.

Conclusion

This study, by being situated in the natural school setting and focusing on the teacher, provides insights to help integrate technology into mathematics education in the school setting. By focusing on teachers with different levels of experience and comfort in using technology to teach mathematics, this study highlights the factors that impact their decisions when using tablets in their teaching of mathematics. Through regular group meetings that fostered a learning network, the teachers focused on reflection and exchanging knowledge about their use of tablets, and in this process their use tablets developed. These insights are hoped to inform future research and implementation of mobile devices in schools, as well as the implementation of ongoing learning networks in schools that may help teachers further develop their use of technology in mathematics education.

References

- Chronaki, A., Matos, A., 2014. Technology use and mathematics teaching: teacher change as discursive identity work. *Learn. Media Technol.* 39, 107–125.
- Clark-Wilson, A., Hoyle, C., Noss, R., Vahey, P., Roschelle, J., 2015. Scaling a technology-based innovation: windows on the evolution of mathematics teachers' practices. *ZDM Math. Educ.* 47, 79–92.
- Dawson, K., 2012. Using action research projects to examine teacher technology integration practices. *J. Digit. Learn. Teach. Educ.* 28 (3), 117–124.
- Dwyer, D.C., Ringstaff, C., Haymore, J., 1990. The Evolution of Teachers' Instructional Beliefs and Practices in High-Access-to-Technology Classrooms (ACOT Report #8). Apple Computer, Inc, Cupertino, CA.
- English, L., 2009. Setting an agenda for international research in mathematics education. In: English, L., Bartolini Bussi, M., Jones, G., Lesh, R., Srirnam, B. (Eds.), *Handbook of International Research in Mathematics Education*, second ed. Routledge, New York, pp. 3–19.
- Ertmer, P.A., Ottenbreit-Leftwich, A.T., 2010. Teacher technology change: how knowledge, confidence, beliefs, and culture intersect. *J. Res. Technol. Educ.* 42, 255–284.
- Fuller, H.L., 2000. First teach their teachers: technology support and computer use in academic subjects. *J. Res. Comput. Educ.* 32 (4), 511–535.
- Guha, S., 2003. Are we technologically prepared?—teachers' perspective on the causes of comfort or discomfort in using computers at elementary grade teaching. *Inf. Technol. Child. Educ. Annu.* (1), 317–349.
- Gunn, T.M., Hollingsworth, M., 2013. The implementation and assessment of a shared 21st century learning vision: a district-based approach. *J. Res. Technol. Educ.* 45 (3), 201–228.
- Heitink, M., Voogt, J., Verplanken, L., van Braak, J., Fisser, P., 2016. Teachers' professional reasoning about their pedagogical use of technology. *Comput. Educ.* 101, 70–83.
- Jones, K., 2011. The value of learning geometry with ICT: lessons from innovative educational research. In: Oldknow, A., Knights, C. (Eds.), *Mathematics Education With Digital Technology*. Continuum, London, UK, pp. 39–45.
- Joubert, M., 2012. Using digital technologies in mathematics teaching: developing an understanding of the landscape using three "grand challenge" themes. *Educ. Stud. Math.* 82, 341–359.
- Kiger, D., Herro, D., Prunty, D., 2012. Examining the influence of a mobile learning intervention on third grade math achievement. *J. Res. Technol. Educ.* 45 (1), 61–82.
- Laborde, C., Sträbe, R., 2010. Place and use of new technology in the teaching of mathematics: ICMI activities in the past 25 years. *ZDM Math. Educ.* 42, 121–133.

- Laborde, C., 2001. Integration of technology in the design of geometry tasks with cabri-geometry. *Int. J. Comput. Math. Learn.* 6, 283–317.
- Ma, W.W., Andersson, R., Streith, K., 2005. Examining user acceptance of computer technology: an empirical study of student teachers. *J. Comput. Assist. Learn.* 21, 387–395.
- Mentz, M., Mentz, K., 2003. Managing technology integration into schools: a South African perspective. *J. Educ. Adm.* 41 (2), 186–200.
- Merchant, G., 2012. Mobile practices in everyday life: popular digital technologies and schooling revisited. *Br. J. Educ. Technol.* 43 (5), 770–782.
- Moersch, C., 1995. Levels of technology implementation (LoTi): a framework for measuring classroom technology use. *Learn. Lead. Technol.* 23 (3), 40–42.
- Mumtaz, S., 2000. Factors affecting teachers' use of information and communications technology: a review of the literature. *J. Inf. Technol. Teach. Educ.* 9 (3), 319–342.
- Neyland, E., 2011. Integrating online learning in NSW secondary schools: three schools perspectives on ICT adoption. *Aust. J. Educ. Technol.* 27 (1), 152–173.
- Noss, R., Hoyle, C., 1996. *Windows on Mathematical Meanings: Learning Cultures and Computers*. Kluwer Academic, Dordrecht.
- OECD, 2015. *Students, Computers and Learning: Making the Connection*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264239555-en>.
- Opfer, V.D., Pedder, D., 2011. Conceptualizing teacher professional learning. *Rev. Educ. Res.* 81 (3), 376–407.
- Parker, C.E., Stylianides, C.D., Bonney, C.R., Schillaci, R., McAuliffe, C., 2015. Examining the quality of technology implementation in STEM classrooms: demonstration of an evaluative framework. *J. Res. Technol. Educ.* 47 (2), 105–121.
- Pelgrum, W.J., 2001. Obstacles to the integration of ICT in education: results from a worldwide educational assessment. *Comput. Educ.* 37 (2), 163–178.
- Petrovai, K., 2018. *A Study Investigating Teachers' Use of, and Views on, Tablets in the Teaching of Mathematics* [PhD thesis]. University of Oxford.
- Rakes, G.C., Fields, V.S., Cox, K.E., 2006. The influence of teachers' technology use on instructional practices. *J. Res. Technol. Educ.* 38 (4), 409–424.
- Ruthven, K., 2014. Characteristics and impact of the Further Mathematics Knowledge Networks: analysis of an English professional development initiative on the teaching of advanced mathematics. *Teach. Math. Appl.* 33, 137–149.
- Short, H., 2014. A critical evaluation of the contribution of trust to effective Technology Enhanced Learning in the workplace: a literature review. *Br. J. Educ. Technol.* 45 (6), 1014–1022.
- Stylianides, G.J., Stylianides, A.J., 2014. The role of instructional engineering in reducing the uncertainties of ambitious teaching. *Cogn. Instr.* 32 (4), 374–415.
- Swallow, M., 2015. The year-two decline: exploring the incremental experiences of a 1:1 technology initiative. *J. Res. Technol. Educ.* 47 (2), 122–137.
- Tobin, D.R., 1998. *Building Your Personal Learning Network*. Retrieved from: <http://www.tobincls.com/learningnetwork.htm>.
- Trust, T., Krutka, D.G., Carpenter, J.P., 2016. "Together we are better": professional learning networks for teachers. *Comput. Educ.* 102, 15–34.
- Trust, T., 2012. Professional learning networks designed for teacher learning. *J. Digit. Learn. Teach. Educ.* 28 (4), 133–138.
- Tsitouridou, M., Vryzas, K., 2004. The prospect of integrating ICT into the education of young children: the views of Greek early childhood teachers. *Eur. J. Teach. Educ.* 27 (1), 29–47.
- Vérillon, P., Rabardel, P., 1995. Cognition and artifacts: a contribution to the study of thought in relation to instrumented activity. *Eur. J. Psychol. Educ.* 9 (3), 77–101.

Teaching is influenced by the teacher

Janet Clinton, Ruth Aston, and Georgia Dawson, University of Melbourne, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	287
Methodology	288
Studies reviewed	288
Quality assessment of reviewed studies	288
Characteristics of effective teachers	288
Cognition	288
Social and emotional competence	289
Self-efficacy	291
Communication	292
Motivation	292
Cultural competence	292
Personality	293
Self-reflection and reflection	293
Collegiality	294
Attitudes, beliefs & expectations	295
Values, morals and ethics	296
Summary	296
A model of teacher effectiveness	297
Trait dimensions	297
State dimensions	297
The psychosocial characteristics related to the self	298
Social and emotional competency	298
Cognitive characteristics	299
Capability dimensions	299
Cultural competence	299
Commitment and motivation	299
Critical reflection	299
Attitudes, beliefs, expectations	299
Behavioral dimensions	299
Communication	299
Collegiality	299
Competence	299
Understanding the juxtaposition between effective teachers and teaching	299
Understanding effective teachers and teaching	300
A final word	300
Acknowledgment	300
References	301
Sources identified for supplementary evidence	302

Introduction

Teachers have been shown to elicit the greatest impact of in-school variables on student outcomes (Hattie, 2009). Darling-Hammond (2000) noted that the effects of quality teaching on student outcomes are greater than associated with students' backgrounds. This impact may be positive or negative, and the effect of poor-quality teaching has shown to be debilitating and cumulative on the student outcomes (Darling-Hammond, 2000). Slater et al. (2012) noted that students had almost a 12-month difference in academic progression when being taught by a low-performing teacher compared with a high-performing teacher. As such, teacher and teaching quality and effectiveness have become critical components of school improvement.

There is an important distinction between teaching quality as defined by practice and instruction, and teacher quality as defined by personal attributes, competency, skill, and dispositions (Darling-Hammond, et al., 2012). There are extensive reviews and research on teaching quality, and many existing standards define quality teaching practices, however, there are comparatively few reviews of research on teacher quality.

While all teachers have their unique personalities and identities, until now it has not been clear what types of communication style, beliefs, attitudes, and personality attributes are common among high-quality teachers, and therefore contribute to more

effective teaching. It is also unclear how and how much the teacher as a person relates to teaching practice, and thence to student outcomes. This review aims to explore these key characteristics of effective teachers and consider the juxtaposition of teacher and teaching quality.

Methodology

Given the complexity and inconsistency in the use of terminology related to teacher effectiveness, a systematic review appears an appropriate method to explore the influence of teacher attributes. The research strategies employed in this review utilised a search strategy that included a combination of database-specific subject headings and natural language keywords. Strategies used reflect Cochrane systematic review recommendations to identify as many relevant records as possible, with searches comprising a combination of subject terms selected from the controlled vocabulary or thesaurus ("exploded" where appropriate) with a wide range of free-text terms.

Inclusion criteria employed during the source screening phase of the review included the country in which the study was conducted, type of study, teacher variables, and association with outcomes of interest (such as student learning outcomes). To ensure that the findings of the review were based on relevant quality evidence, the review only included studies on teacher evaluation tools that were published in the last 10 years, between 2008 and 2018. Considering the breadth of studies that were included in this review, a quality assessment approach was created for different studies. Two quality appraisal checklists were used: The Grading of Recommendations Assessment, Development and Evaluation [GRADE] and the Critical Appraisal Skills Programme of Qualitative Studies [CASP]. Screened sources meeting the inclusion criteria were fully reviewed in detail, and data were extracted and coded to identify teacher characteristics and the impact and strength of the relationship, or association between these characteristics with the outcomes of interest. Where available statistical data including measures of outcomes in teacher quality and student learning were extracted to investigate the impact of teacher characteristics on specific outcomes of interest. Qualitative data were also extracted following specific coding categories and procedures.

Studies reviewed

Of the 12,317 articles across twelve databases, identified through the application of the search strategies, only 52 met the inclusion criteria. As illustrated by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) diagram (Fig. 1), most sources ($n = 10,982$) were excluded at the initial abstract screening stage as they did not meet the inclusion criteria of discussing teacher characteristics or were not primary research studies. Most of the included sources utilised qualitative research designs ($n = 23$), with the majority being classified as narrative reviews ($n = 12$). Only four of the included articles were systematic reviews or meta-analyses.

Quality assessment of reviewed studies

GRADE and an adaptation of the CASP checklist were used to assess the quality of the 52 included studies. Based on the CASP quality assessment, all qualitative studies were deemed of sufficient quality to be included in the review. For the quantitative reviewed studies, the quality assessment using GRADE indicated that across all studies there were no serious limitations based on the study design and results presented. However, the quantitative studies varied considerably in quality which was largely explained by differences in the depth of information reported, and the measures conducted (or not) to address sources of bias. In light of these results, the specific effect size estimates should be interpreted with some caution, however the GRADE quality assessment findings were not concerning enough to warrant exclusion of the four quantitative studies included in the meta-analysis.

Characteristics of effective teachers

A total of 11 teacher characteristics were identified through the analytic process. These characteristics were cognition, social and emotional competence, self-efficacy, communication, motivation, cultural competence, personality, self-reflection, collegiality, attitudes, beliefs, expectations, values, morals, and ethics. Evidence for each of these is summarized in the accompanying tables and judgments about whether there is a direct or indirect relationship between the characteristic and effectiveness outcomes (teaching quality and student learning) are made based on the systematically reviewed evidence.

Cognition

Cognition was defined as the teachers' creativity, critical thinking, and competence. Of the 52 included studies, five investigated teacher characteristics associated with cognition and how they relate to teaching quality outcomes, including managing student behavior and developing inclusive education environments. Overall, the studies synthesized utilized a range of review designs and were of moderate to high quality based on the CASP and GRADE criteria. Teacher cognition was minimally addressed in the reviewed primary sources, however its related aspects of creativity and teacher credibility were explored in greater detail. Based

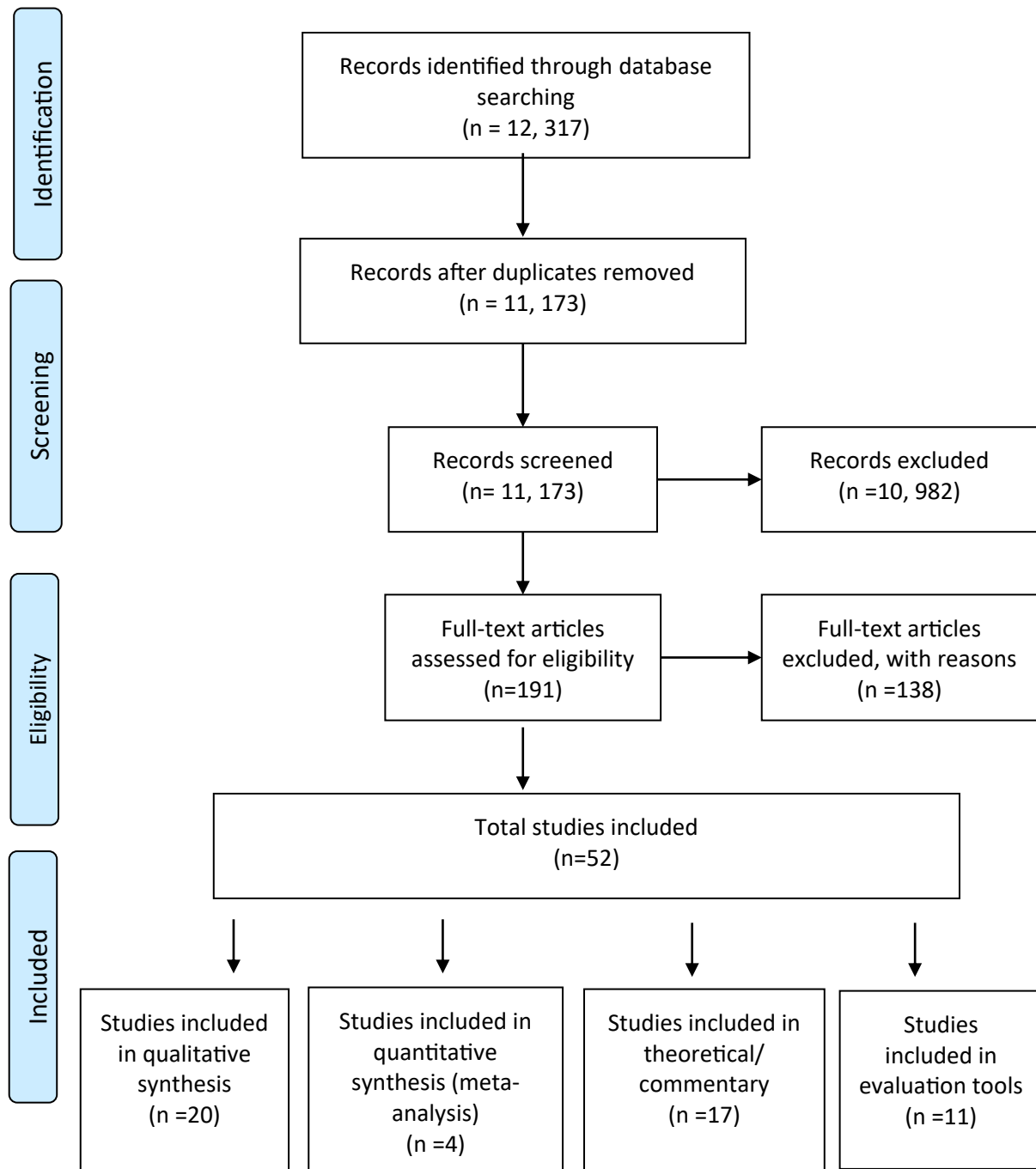


Fig. 1 PRISMA diagram.

on the review of primary sources, creativity is considered an attribute that is broadly useful and significant for teacher effectiveness (Table 1).

Social and emotional competence

This was defined by five core SEL competencies (also referred to throughout as SEC) that comprise this characteristic and included self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (Table 2). Of the 52 included studies, six investigated the importance and relationship between *teacher social social-emotional learning and competency* and student outcomes including student learning, school climate, and teacher outcomes. The studies synthesized utilized varied study designs and were of varying quality based on the CASP and GRADE criteria (2018; 2013). Based on the synthesis of six studies

Table 1 Cognition.

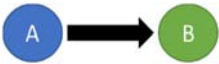
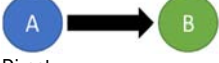
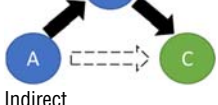
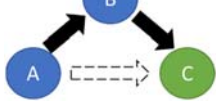
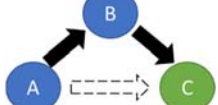
	Direct/indirect	Effect size (Cohen's d) & qualitative findings/direction of the relationship
Teacher ability to teach children with behavioral problems	Direct 	Teacher creativity, imagination, insight, and openness were related to their ability to teach children with behavioral problems and manage behavior within a mainstream classroom. $d = 0.35$
Classroom management	Direct 	
Community building/collaboration	Direct 	Creative, open, and imaginative teachers are more likely to engage in efforts to build communities of practice and collaborate with more and less experienced educators to learn. Open and creative teachers tend to promote inclusivity in their classroom readily.
Inclusive education	Direct 	
Student-centered teaching	Direct 	$d = 0.36$
Teacher judgment, decision making & student achievement	Direct 	$d = 0.63$ Teachers who adopt more evidence-informed judgments improved student achievement.

Table 2 Social and emotional competence.

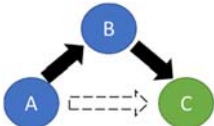
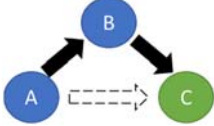
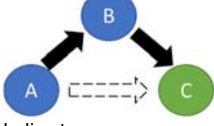
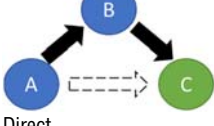
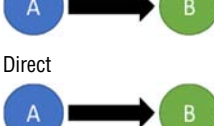
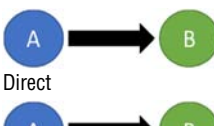
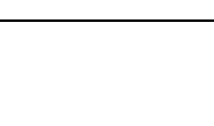
	Direct/indirect	Effect size (Cohen's d) & qualitative findings/direction of the relationship
Student relationships	Direct 	$d = 0.52$
Relationships with other teachers/colleagues	Direct 	SEC improves teachers' relationship building with colleagues.
Relationships with the broader school community	Direct 	SEC improves teachers' relationship building with parents, caregivers, and the broader community.
Teacher personal wellbeing	Indirect 	High SEC teachers tend to report greater personal wellbeing.
Student social and emotional learning	Indirect 	High SEC teachers have a greater capacity to implement SEL programs in the classroom.
Student achievement	Indirect 	High SEC teachers who implement SEL in their classroom tend to promote more positive development which can improve academic performance.
Student behavior	Indirect 	$d = 0.56$ High SEC teachers promoting more positive development in the classroom can help manage student behavioral problems and promote social capital.

in this area, SEC of the teacher is a characteristic that is related to teacher quality, which is correlated to known outcomes that are causally related to student achievement, and this is particularly notable for the contribution of teacher's SEL to the development of beneficial student-teacher relationships (Hattie, 2018).

Self-efficacy

This is a teacher's belief in his or her innate ability or capability to organize and execute the courses of action required to manage prospective situations and have a high impact on student learning (Table 3). Eight of the 52 included sources that investigated the relationship between *teacher self-efficacy* and outcomes. Overall, the studies synthesized displayed a combination of systematic and narrative designs as detailed below. The quality of studies varied based on the CASP criteria with a combination of high and moderate-quality reviews. Two discussion papers were judged to be of poor quality, largely due to insufficient reporting of study design, analyses, and the process for review of primary research (2018; 2013). Based on the synthesis, teacher self-efficacy can be considered a teacher characteristic that is related to teacher quality, including student achievement, and has a large effect on student self-efficacy (Visible Learning, 2017). In comparison to the studies synthesized that discuss other identified characteristics, the evidence for teacher self-efficacy is of high quality, and there is sufficient evidence to suggest that teacher self-efficacy is an important characteristic for teachers, teacher practice, and student learning.

Table 3 Self efficacy.

	Direct/indirect	Effect size (Cohen's d) & qualitative findings/direction of the relationship
Student expectations	Indirect 	Linked with teacher expectations. Teachers with high TSE tend to have high expectations of students and in turn influence students' expectations of their ability.
Student motivation	Indirect 	$d = 0.42$
Student self-efficacy	Indirect 	$d = 0.92$
Student achievement [teacher self-efficacy & student achievement]	Indirect 	$d = 0.21$ (0.22) Teachers with high TSE have been positively associated with student achievement in STEM areas.
Classroom management	Direct 	$d = 0.35$ High TSE is associated with classroom practices, particularly in managing problem behavior, and promoting a cohesive environment.
Teacher wellbeing	Direct 	High TSE is associated with teacher wellbeing, particularly in terms of confidence and resilience to coping with stress.
Teacher expectations of students	Direct 	$d = 0.43$ (0.08)
Inquiry-based teaching	Direct 	$d = 0.40$ Teachers with high TSE tend to promote student participation and discussion in the classroom, and therefore inquiry-based learning.

Communication

Communication is the act of transmitting information and includes verbal and non-verbal types of communication such as communication style, body language, appearance, and instructional communication (Table 4). Across the reviewed sources, there appears to be strong evidence for the importance of communication in influencing teacher pedagogy and practice, and ultimately student learning outcomes. Eight papers described the importance of communication as a characteristic of effective teachers. Most of these papers were qualitative study designs, although communication was also explored in two of the meta-analyses. Collectively, the evidence suggested communication is a critical characteristic of effective teaching, “decades of research leave no doubt that communication is central to the teaching-learning process” (Webster, 2010, p. 33). Based on the synthesis, communication can be considered a teacher characteristic that is related to outcomes of interest, which is correlated to known outcomes that are causally related to student achievement, namely student-teacher relationships (Hattie, 2018).

Motivation

Generally, motivation is defined as the drive that influences behavior; there are various dimensions of motivation theory including intrinsic motivation and self-determination theory (Table 5). When considering teacher motivation with other identified characteristics, there was variable evidence found for motivation. This conclusion, however, is based on only one qualitative source that focused on teacher motivation as a key teacher characteristic. Han and Yin (2016) conducted a literature review focused on the “abundance of approaches to explain motivation” (p. 12), aiming to conceptualize a framework for further research. This review was graded as moderate-high quality. Overall, teacher motivation is linked with student outcomes, Teacher motivation can be understood through self-determination theory, which also correlates with identified demotivating factors, such as lack of autonomy.

Cultural competence

This is awareness of one’s cultural competence and views on diversity, and one’s ability to successfully teach students who identify with a culture or cultures other than those of the teacher (Table 6). One article among the 52 included studies discussed cultural competence. Cultural competence was found to both *directly* and *indirectly*, impact student learning and other outcomes. Porter-Samuels (2013) described the importance of cultural responsiveness specifically concerning Māori and Pasifika students in a New Zealand context. Based on the review of secondary and supplementary research literature, cultural competence is an important characteristic for creating an empathetic, caring classroom environment, for all students and especially for minority students or students whose background does not necessarily match the teaching methods and culture(s) of their teachers.

Table 4 Communication.

	Direct/indirect	Effect size (Cohen's d) & qualitative findings/direction of relationship
Student achievement [clarity]	Indirect	$d = 0.75$
Student achievement [verbal ability]	Indirect	$d = 0.22$
Student motivation	Direct	Relational and rhetorical communication techniques can allow teachers to increase their communication effectiveness, and subsequently increase student motivation in class.
Positive classroom/school climate	Direct	Communication plays an important role in promoting positive home-school relationships and a positive classroom climate

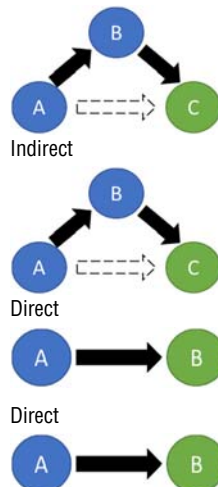


Table 5 Motivation.

	Direct/indirect	Effect size/qualitative findings
Student achievement	Indirect <pre> graph LR A((A)) --> B((B)) B((B)) --> C((C)) A((A)) -.-> C((C)) </pre>	Teacher motivation is linked with student learning through enhancing classroom effectiveness (positive classroom climate for learning)

Table 6 Cultural competence.

	Direct/indirect	Effect size (Cohen's d) & qualitative findings/direction of the relationship
Student relationships	Direct <pre> graph LR A((A)) --> B((B)) </pre>	Culturally responsive teachers can build a stronger rapport with students from culturally diverse backgrounds.
Student engagement	Indirect <pre> graph LR A((A)) --> B((B)) B((B)) --> C((C)) A((A)) -.-> C((C)) </pre>	Cultural responsiveness can increase student engagement, through the inclusion of culturally relevant materials.
Classroom climate	Direct <pre> graph LR A((A)) --> B((B)) </pre>	Culturally responsive teachers can create “culture of care,” an inclusive, caring, and supportive environment.

Personality

This is the combination of those characteristics or qualities that make up an individual's character (Table 7). There are several models of personality, most notably the five-factor model of personality which includes: extraversion, agreeableness, openness to experience, conscientiousness, and neuroticism. A total of eight articles included in the review examined personality. There was variance in the strength of the evidence among the eight articles. While not directly observable in a teaching and learning context, the evidence base suggests personality and certain dispositions such as extraversion have been linked to positive classroom behaviors, and can influence student learning outcomes. Empirical studies illustrate that personality factors are related to teaching behaviors and practice that in turn impact students, in both affective outcomes and cognitive outcomes. While there is limited effect size information, Hattie (2018) suggests the magnitude of the influence of personality on student outcomes is quite small. This is likely due to the distal nature of the relationship, where personality is expressed through teacher behavior and practice, which in turn influence student outcomes (Hattie, 2018). There are few dimensions of teacher personality systematically related to achievement effects, and at best there are indirect effects on other attributes (as noted in Fig. 3). From the breakdown in Table 7, positive effects on student achievement relate to teacher's personal confidence in themselves, demonstrating enthusiasm for the job, and those most able to identify and make judgments about situations. The other personality factors have low to zero correlations with student achievement.

Self-reflection and reflection

This relates to a teacher's reflective processes in identifying their strengths and weaknesses in their teaching practice, and their ability to adapt or correct them accordingly (Table 8). Three articles within the review examined the impact of self-reflection and reflection on teaching practice with one. Narrative review rated with moderate to high evidence strength. Teacher self-reflection is an important tool that is closely linked to teacher self-efficacy. The belief in one's capabilities to execute a course of action, to manage prospective situations requires the teacher to assess their own past experiences to inform their future decisions and practice. Self-reflection, therefore, is correlated with quality teaching practice. Teacher self-reflection and its relationship with student outcomes are not particularly comprehensively explored in the reviewed studies. However, there was consensus across reviewed literature that teacher self-reflection is influential on teaching practice and therefore is an important teacher characteristic.

Table 7 Personality.

	Direct/indirect	Effect size (Cohen's <i>d</i>) & qualitative findings/direction of relationship
Classroom management	Direct 	Personality traits, such as extraversion have been linked to positive classroom behaviors, and reduced reports of aggression and negative behaviors in children.
Student affect	Direct 	Teacher's personality traits were found to impact student affect and emotions.
Student motivation	Direct 	Teacher's personality traits were found to impact student motivation and engagement.
Communication	Direct 	Teacher's personality traits were found to impact teacher's communication styles and practices.
Teacher attitudes and beliefs	Direct 	Personality is related to the attitudes and beliefs that teachers have about students and their learning.
Student achievement	Indirect 	Teacher personality attributes are indirectly related to student achievements, $d = 0.25$. This relationship is mediated by other characteristics, such as confidence and knowledge.
Student achievement	Indirect 	Certain personality types may be more willing to engage in self-reflection and adaptive teaching strategies, ultimately improving student learning outcomes.
Student engagement	Direct 	All five personality dimensions (except for neuroticism) had a positive relationship with a teacher's ability to engage students with behavioral problems.

Collegiality

Collegiality defines the cooperative relationship between teachers and their colleagues, and in this context the collaborative nature, encouragement, and optimization of peers, especially with teacher leaders (Table 9). Of the 52 reviewed studies, three investigated teacher characteristics associated with collegiality, leadership, and community and their respective relationship with student outcomes. The studies reviewed included a narrative review and two theoretical commentaries, which were of poor to moderate quality based on the CASP criteria (2018). Collegiality and the constructs within this category have several direct and indirect relationships with student outcomes. Collegiality is also linked to other characteristics, such as teacher self-efficacy and self-reflection. Based on the synthesis of reviewed research, collegiality can be considered a characteristic that is related to outcomes of interest, however, more high-quality empirical research is required to definitively conclude that collegiality is important for student learning.

Table 8 Self-reflection and reflection.

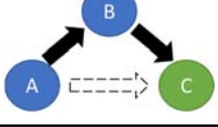
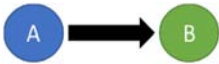
	Direct/indirect	Effect size (Cohen's <i>d</i>) & qualitative findings/direction of the relationship
Student achievement	Indirect 	Reflective practices are linked to student learning outcomes, as teachers adapt their classroom practices to better meet the needs of their students.

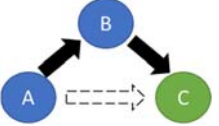
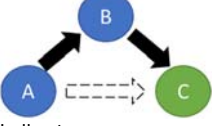
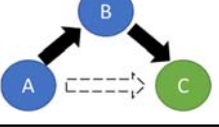
Table 9 Collegiality.

	Direct/indirect	Effect size (Cohen's <i>d</i>) & qualitative findings/direction of a relationship
Social capital	Direct 	Teacher social capital and information sharing, as well as trust, are directly related through collegiality and the illustration of these characteristics through the development of professional learning communities.
Co- or team-teaching	Direct 	$d = 0.19$ Collegiality encourages the implementation of co or team-teaching, more collaborative teaching, and educators are encouraged to observe colleagues teaching.
Student-centered teaching	Direct 	$d = 0.36$

Attitudes, beliefs & expectations

This is a teachers' belief in the importance of accomplished teaching and the learning and achievement potential of all students (Table 10). Of the 52 included studies, six investigated teacher characteristics associated with attitudes, beliefs, and expectations and their respective relationships with student outcomes. Teacher attitudes, beliefs, and expectations have both direct and indirect relationships with outcomes for students and the classroom learning environment. However, the strength of evidence varied considerably among the studies synthesized. Based on the synthesis of six secondary studies, teacher attitudes, beliefs, and expectations can be considered a characteristic that is related to student learning. However, it is noted that the evidence for some outcomes, for instance, student cognitive learning, is of poor quality and therefore the relationship between this characteristic and student learning should be interpreted with some caution. However, isolating the effects for goal setting and teacher expectations specifically, which are large and reflected in multiple sources synthesized, the evidence is strong (Hattie, 2012). Overall, there is sufficient evidence to

Table 10 Attitudes, beliefs and expectations.

	Direct/indirect	Effect size (Cohen's <i>d</i> , Fisher's <i>Z</i> [<i>r</i>]) & qualitative findings/direction of the relationship
Teacher expectations	Direct 	$d = 0.43$ Teacher expectations of their teaching ability and the capability of their students have an impact on student achievement.
Teacher goal-setting for students	Direct 	$d = 0.59$ Teachers who hold positive attitudes, beliefs, and high expectations tend to set regular goals for their students that align with an ambitious achievement vision. This goal-setting has an impact on student achievement.
Student cognitive learning	Indirect 	$r = 0.05$ Teacher beliefs and expectations about their students' learning have an impact on cognitive learning outcomes.
Student confidence	Indirect 	Teacher beliefs and expectations when aligned to a positive view of the learning potential of students can increase student confidence in their learning.
Student motivation and affective learning	Indirect 	$d = 0.44$ $r = 0.01$ Teacher beliefs and expectations about their students' learning have an impact on affective learning outcomes.

indicate that teacher attitudes, beliefs, and expectations are important characteristics as they support the establishment of a learning environment that encourages students to progress in their learning and work towards achieving their potential.

Values, morals and ethics

Teacher values can be defined as the beliefs and virtues that teachers may hold about issues, such as religion or environmental sustainability, and extend to a teacher's ethics and morality. Surprisingly, the characteristics of values, morals and ethics did not emerge from the literature reviewed. Given that it is considered an important dimension of a teacher's character in policy and resources relating to teacher standards, an additional search for supplementary evidence was conducted and yielded 52 results. Nine documents were analyzed of varying significance and quality regarding the value, moral, and ethical dimensions of teacher characteristics and its influence on student learning outcomes. No empirical evidence was found in reviewed studies however, the targeted review of supplementary research illustrated the ways in which teachers' values, morals and ethical position can influence their teaching practice, and relationships with their students. As with areas such as personality and cognition, teacher values, morals and ethics may have a more distal relationship to outcomes of interest but could be an important mechanism through which other characteristics are expressed, for instance, communication could be influenced by teacher values, morals, and ethics. More high-quality empirical research is needed to investigate and evaluate these mechanisms in more depth to make a definitive conclusion about the importance of teacher values, morals, and ethical position for teaching practice, student, and school outcomes (Table 11).

Summary

The evidence for the identified teacher characteristics is considered to be of moderate strength and quality. While some characteristics were evidenced by a large number of reviewed articles, such as self-efficacy and communication, others had very few articles providing evidence, such as motivation, collegiality, self-reflection, cognitive capability, and cultural competence. These characteristics were supplemented with a further search for primary or supplementary sources, which increased the breadth of literature for those characteristics (Table 11). Note that Values, morals and ethics are not included in the table due to insufficient evidence.

The 11 teacher characteristics identified together influence a wide range of outcomes categorized as student, teacher, or school outcomes. These influences can be either direct or indirect. Student outcomes are perhaps the most common and visible of the outcomes, as teacher practice is generally geared toward the end goal of student achievement. Some characteristics directly influence student outcomes such as student engagement, student affect, and student motivation; teacher characteristics that directly influence these outcomes include personality, cultural competence, communication. It is more common, however, for teacher characteristics to exert influence over student outcomes indirectly; this may be because teacher characteristics influence teaching practice, which in turn affects student experiences and achievement. All characteristics, except for collegiality, indirectly influence at least one student outcome.

Teacher outcomes include outcomes that affect the teacher as an instructor, for example, teacher wellbeing or teacher expectations of students. Classroom outcomes are an extension of teacher outcomes but are treated separately here to distinguish between the teacher as an instructor and the classroom as an instructional environment. Teacher characteristics generally have a direct influence over classroom outcomes. For example, teacher self-efficacy has a direct effect on teacher wellbeing, teacher expectations of students, classroom management, and inquiry-based teaching. All characteristics, except for motivation, directly or indirectly influence at least one classroom outcome.

School outcomes take a broader perspective and tend to apply to the school as a system. Despite the complex nature of schools, the identified school outcomes are under the direct influence of the teacher characteristics. Characteristics such as communication,

Table 11 Summary of findings of characteristics.

Characteristic	Number of primary/supplementary sources providing		Strength of evidence
	Number of secondary sources	evidence	
Cognitive capability	5	152	Moderate
Social and emotional learning/competence	6	N/A	Moderate
Self-efficacy	8	570	Moderate-high
Communication	8	233	Low-moderate
Attitudes, beliefs, and expectations	6	381	Moderate
Motivation	4	12	Moderate
Cultural competence	1	8	Moderate
Personality	8	61	Moderate
Self-reflection and reflection	3	5	Low
Collegiality	3	5	Moderate
Total	52	1427	Moderate

social and emotional learning, and cognitive capability all have a direct influence on school outcomes, including school climate, relationships with colleagues and the wider school community, and inclusive education.

The relational map of teacher characteristics illustrates a summary view of the key connection and outcomes (Fig. 2).

A model of teacher effectiveness

The findings from a review of the literature provides an opportunity to postulate a model that outlines the process in which the characteristics relate to each other and ultimately impact student, teacher, and school outcomes. The model is an attempt to make meaning of the characteristics identified in the systematic review and those lacking from the evidence base. It draws on many integral and well-established relationships; however, we do not suggest that the systematic review findings alone provide explicit, specific evidence for the construction and design of this model. More empirical evidence would enable the categorization of characteristics and order of the relationships to be tested.

It is suggested that the characteristics can be ordered to reflect a relational process. As such, there have been changes to the characteristic groupings presented in the results, to ensure that the characteristics align with the overarching dimension categories. Importantly, this process is underpinned by context and the specific teaching and learning environment, as well as contextual factors related to the individual teacher (e.g., career stage). Therefore, it must be considered as a dynamic and interactive model, which varies across teachers, time, and place. It is suggested that there are trait and state dimensions of teacher behavior that interact with the teaching environment. These are mediated by capability, attitudes and beliefs that in turn manifest behaviors (Fig. 3).

Trait dimensions

Trait aspects of a teacher's character are those personality variables that relate to extraversion, openness, conscientiousness, neuroticism, and agreeableness. These are those factors that are often hidden from sight, but which are relatively stable throughout one's life span. It is these elements that make up the character of a person and often influence thoughts, behaviors, and emotions. Hence, they are often thought of as inherent and distinguishing. While the evidence regarding the impact of these personality variables on teaching is not strong, they are often seen as characteristics that predict the communication and behavior of the teacher. Furthermore, there is strong evidence in the psychology literature that these traits and personality constructs relate to a number of the psychosocial constructs highlighted within the model, such as self-awareness and self-efficacy (Brown and Cinamon, 2016; Strecher et al., 1986; Lord et al., 1986). In the model, these trait variables are considered central to the teacher characteristics that facilitate effective teaching. However, we acknowledge that they may not always be measurable.

State dimensions

We have used the term "state dimensions" as a means of describing those characteristics of a teacher that are movable, dynamic, and dependent on certain contexts. State characteristics are often defined as temporary behaviors or feelings that depend on a person's situation, mood and motivations at a particular time and in a particular context. In this case, these psychosocial resources, such as

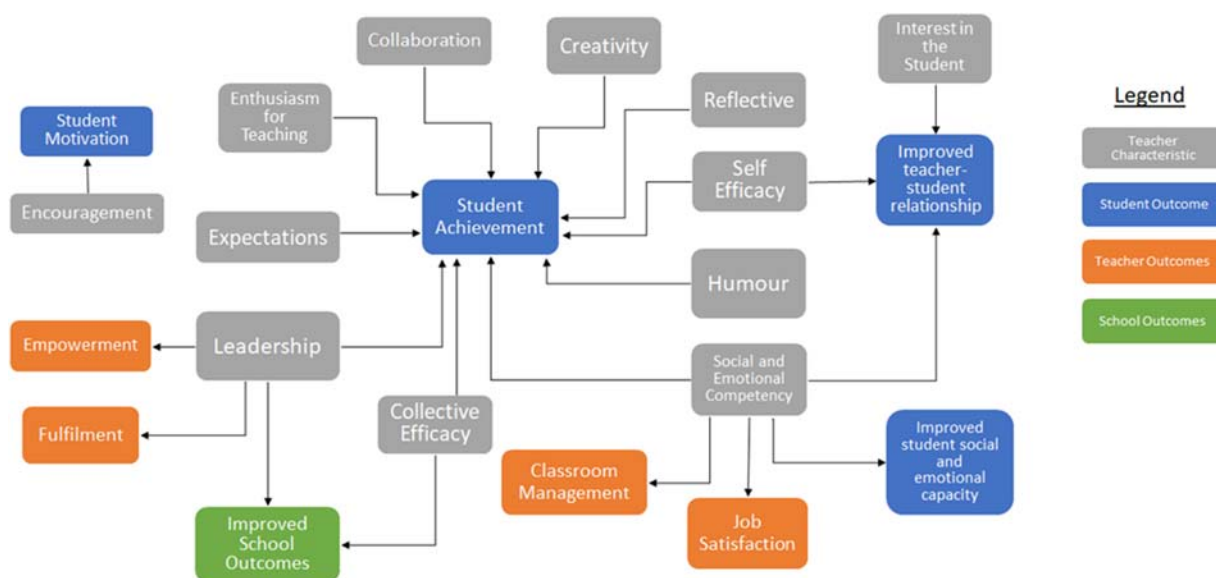


Fig. 2 A relational map of teacher characteristics.

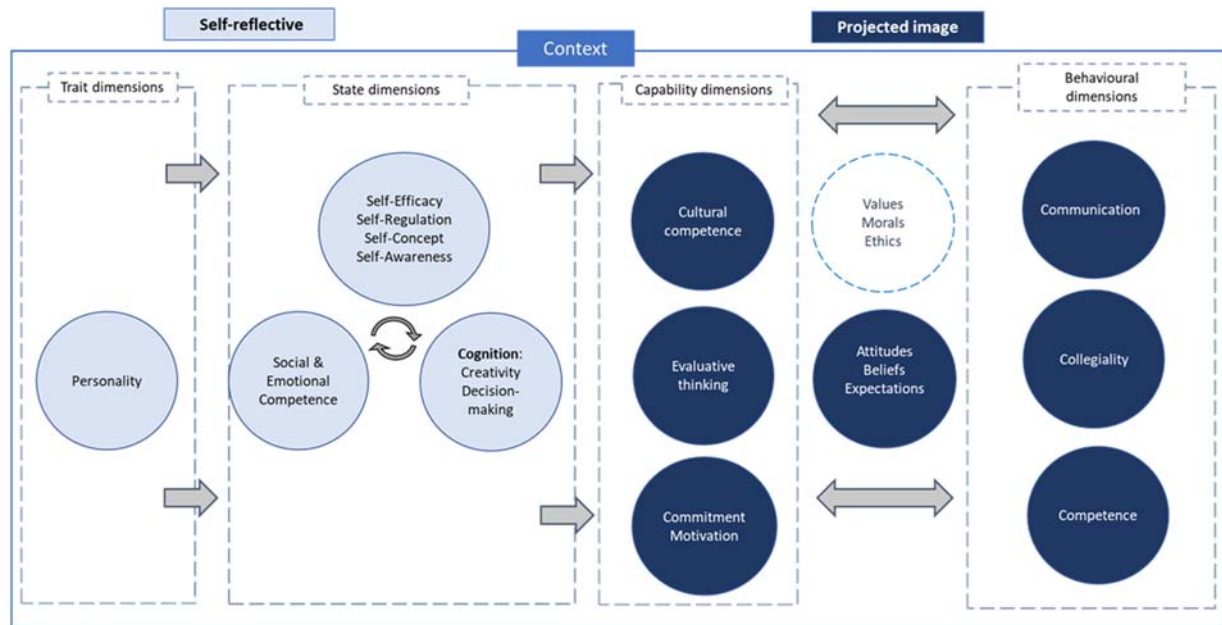


Fig. 3 Model of teacher characteristics.

self-esteem and self-efficacy, are those characteristics that are particularly influenced by the context of teaching. We acknowledge that these are intrinsically related to all personality variables but are more especially relevant to the context of teaching. The model demonstrates key dimensions concerning the self, one's social and emotional competency, and the cognitive dimensions relating to such characteristics as critical thinking and creativity.

The psychosocial characteristics related to the self

These are characteristics that are individual and often relate to social situations. Self-efficacy, regulation, and concept awareness are thought to have beneficial effects on mental and physical health outcomes. In the case of teachers, it is well documented that teacher self-efficacy is related to job satisfaction, wellbeing and teacher retention (Klassen and Chiu, 2010; Caprara et al., 2003, 2006; Skaalvik and Skaalvik, 2010).

Social and emotional competency

This refers to the skills that help individuals to interact in positive ways with others and manage their emotions. These skills are varied and include, among others, relationships skills and social sensitivity. These are extrinsically related to the self and cognitive variables.

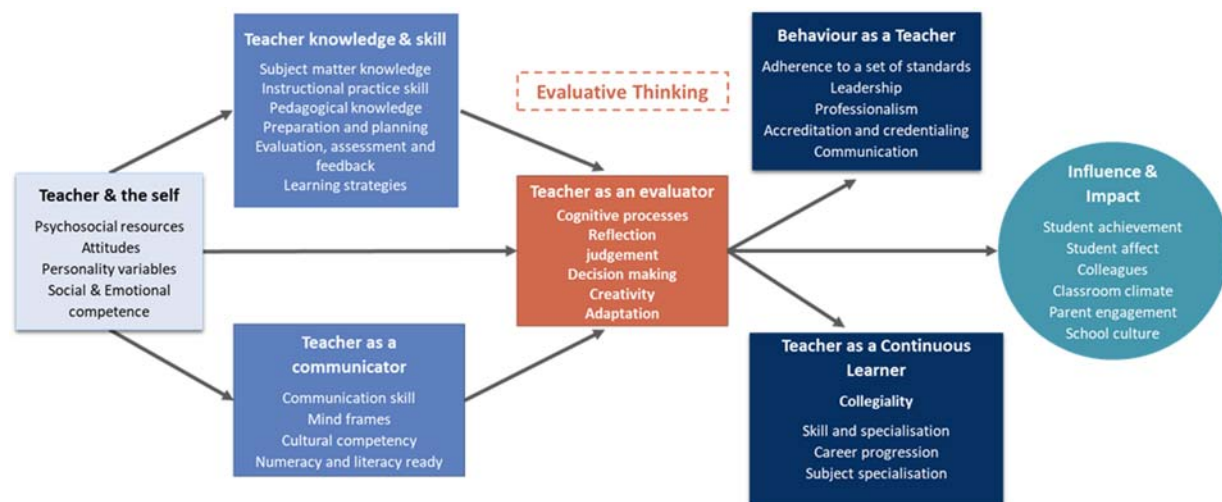


Fig. 4 The relationship between teaching and the teacher.

Cognitive characteristics

These characteristics relate to factors such as creativity, openness, decision-making, or inquisitiveness. The model suggests that there is a strong interaction between the self, social and emotional competency, and cognitive characteristics. These individual dimensions, or a combination of these dimensions, come to the fore in various situations and contexts, for example, when working collaboratively with colleagues on an innovative teaching initiative.

Capability dimensions

The next level of dimensions relates to a teacher's capacity to engage in evaluative thinking, cultural competency, and commitment to and motivation for teaching. These dimensions are defined as capabilities in that they encompass the ability to demonstrate a certain behavior or characteristic. In essence, a capability or skill relating to oneself can be learned and developed.

Cultural competence

The idea of cultural competence relates to three characteristics: the awareness, knowledge, and skill to act appropriately and competently in specific cultural situations.

Commitment and motivation

These capabilities allow for a teacher to demonstrate a level of commitment and the motivation to understand and further education in a variety of settings.

Critical reflection

Critical reflection is a collective of characteristics that allow a teacher to review and modify their practice in a particular setting through the process of evaluative thinking (Yost et al., 2000).

Attitudes, beliefs, expectations

The model then suggests that trait, state, and capability dimensions relate to teaching and consequently shape a range of individual attitudes, beliefs, and expectations as a teacher. In the reviewed literature there was little evidence to distinguish clearly between attitudes and beliefs. There is, however, a strong literature base relating teacher expectations to student outcomes. Closely associated with attitudes, beliefs, and expectations are the morals and ethics that define values often held by teachers. These characteristics were not specifically identified in the systematic review; however, it is a widely discussed area within the standards of behavior for teachers. We suspect this dimension relates to the affective nature of teachers and draws on many of the personal characteristics and the contexts that teachers often find themselves in. Further analysis and research on this dimension are necessary.

Behavioral dimensions

The final level of dimensions within the model relates to projected behaviors: how a teacher communicates; their level of collegiality, for example, their collective efficacy working within teams and school contexts; and finally, the professional image and competence they project. These key aspects of this dimension are those most easily and appropriately evaluated.

Communication

This specifically relates to the skills and behaviors that a teacher may use in a variety of situations to communicate, for example, classroom discourse.

Collegiality

Collegiality relates to the professional, collaborative relationships a teacher may have or how they may project their professionalism to their peers.

Competence

Competence is the confident and skillful image projected by a teacher, or, in other words, behaving and acting as a teacher competently and confidently.

Understanding the juxtaposition between effective teachers and teaching

The current study has focused on teacher attributes as a dimension of effective teaching, as this was an area requiring more fine-grained analysis. As a result of an increased understanding of the dimensions that comprise effective teaching, the conceptual model has been developed to illustrate the nature of the interactions between all dimensions and their impact on a school community, and student learning. Within this overall model, we have been able to determine the significance of the cognitive component of the teacher characteristics. In the model, a higher-order structure is presented where several dimensions are highly related and influence

students, teachers, school climate, parents, and community. This current synthesis has afforded the exploration of the relationships within the model. As previously suggested, this is not a linear structure given the complexity of teaching and learning, however, there is evidence to suggest that many dimensions are related to specific outcomes for students as well as community and school climate. Furthermore, many dimensions interact to mediate influence on these outcomes as indicated in the figure above.

Understanding effective teachers and teaching

Fig. 4 illustrates that there are several key characteristics that a teacher brings to the classroom including knowledge and personal attributes. These attributes include knowledge, mindsets, life experiences, personality, and dispositions. It is suggested that many of these characteristics relate to a person's behavior as a teacher. This behavior and its ensuing outcomes are mediated by a collective of cognitive constructs that can be thought of as evaluative thinking. It is suggested that teachers use internal processes that draw on their knowledge, skill, and personal characteristics to make decisions about a way of behaving. It is this way of behaving concerning teaching, instruction, communication, and interpersonal relationships that ultimately influences the education community.

The importance of a teacher's thought processes has been investigated by many. Hattie (2012) suggests that all teachers are evaluators and are constantly engaging in a process of evaluating their impact. Teachers' thinking processes or cognitive skills have been demonstrated as a significant construct in influencing decisions that teachers make in planning and determining how to adapt their lessons for different students within their charge. Risko et al. (2005) go one step further to describe these characteristics as "overlapping constructs that involve deliberate, evaluative, and constructive activity." These meta-cognitive components of teaching can be thought of as evaluative thinking.

Evaluative thinking stems from the discipline of evaluation and is considered a cognitive construct. That Tom Archibald defines as:

Evaluative Thinking is a cognitive process in the context of evaluation, motivated by an attitude of inquisitiveness and a belief in the value of the evidence, that involves skills such as identifying assumptions, posing thoughtful questions, pursuing deeper understanding through reflection and perspective-taking and making informed decisions in preparation for action.

Archibald (2013).

Further, Patton (2008) suggests it is a way of thinking that infuses everything. The idea of evaluative thinking has a strong synergy with education as it describes that an element of effective teaching is a way of knowledge, decision-making and agility in the application of knowledge and decision-making to the context of a classroom. Wyatt (2017) suggests that "Evaluative thinking is a disciplined approach to inquiry and reflective practice that helps us make sound judgments using good evidence, as a matter of habit" (2017, p. 79). Wyatt goes on to demonstrate how the disciplines of evaluative thinking are woven throughout the school planning process (NSW Department of Education, 2018; Wyatt, 2017). We suggest that it is the synthesis of understanding, knowledge, skill, mindset and the self as a teacher to enact behavior that creates an impact. Within this proposed model of effective teaching, evaluative thinking is the mediator or the conduit between the teacher and teaching.

A final word

This systematic review has concentrated on those cognitive and non-cognitive characteristics of teachers that are thought to have an impact on or influence the school community, in particular, students. It is but one element of the model of effective teachers (Clinton et al., 2017).

The synthesis was devised to progress our understanding of the complex world of teachers and ultimately attempt to predict their impact. Sources of evidence for this rapid review have primarily derived from secondary research sources such as meta-analyses. Systematically reviewing the evidence in this way has allowed us to develop our understanding of those fundamental characteristics of the teacher that are significant. The depth and breadth of the evidence that emerged from the review in itself reveals the state of the research on the characteristics of teachers. There are many gaps and quality is variable.

The overall model of the path of teacher characteristics is derived from the triangulation of evidence extracted from reviewed secondary sources. It is an attempt to demonstrate how certain characteristics relate to or influence others and importantly how we might see some of these characteristics manifest in behavior.

Acknowledgment

This work was funded by a research grant from the Australian Government Department of Education and Training.

References

- Archibald, T., 2013. Evaluative Thinking. Free Range Evaluation, 11th November 2013. Retrieved from. <https://tgarchibald.wordpress.com/2013/11/11/18/>.
- Ashton, P., 1984. Teacher efficacy: a motivational paradigm for effective teacher education. *J. Teach. Educ.* 35 (5), 28–32.
- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84 (2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>.
- Bolyard, J.J., Moyer-Packenham, P.S., 2008. A review of the literature on mathematics and science teacher quality. *Peabody J. Educ.* 83 (4), 509–535.
- Brimhall, J., 2014. Characteristics of music teachers who effectively promote social capital: a literature review. In: *Update Applications of Research in Music Education*, vol. 33, pp. 42–48, 1.
- Brown, D., Cinamon, R.G., 2016. Personality traits' effects on self-efficacy and outcome expectations for high school major choice. *Int. J. Educ. Vocat. Guid.* 16 (3), 343–361.
- Buttner, S., Pijl, S.J., Bijstra, J., van den Bosch, E., 2015. Personality traits of expert teachers of students with behavioural problems: a review and classification of the literature. *Aust. Educ. Res.* 42 (4), 461–481.
- Caprara, G.V., Barbaranelli, C., Borgogni, L., Steca, P., 2003. Efficacy beliefs as determinants of teachers' job satisfaction. *J. Educ. Psychol.* 95 (4), 821.
- Caprara, G.V., Barbaranelli, C., Steca, P., Malone, P.S., 2006. Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: a study at the school level. *J. Sch. Psychol.* 44 (6), 473–490.
- Chichekian, T., Shore, B.M., 2016. Preservice and practicing teachers' self-efficacy for inquiry-based instruction. *Cogent Educ.* 3 (1).
- Clinton, J., Anderson, M., Dawson, G., Dawson, A., Bolton, S., Mason, R., 2017. Teacher Effectiveness Systems, Frameworks and Measures: A Review. Centre for Program Evaluation, Melbourne, Australia.
- Cooke, S., 2017. The moral work of teaching: a virtue-ethics approach to teacher education. In: Clandinin, J., Husu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*. SAGE.
- Creemers, B., Kyriakides, L., 2015. Process-product research: a cornerstone in educational effectiveness research. *J. Classr. Interact.* 50 (2), 107–119.
- Danielson Group, 2011. The 2011 Framework for Teaching Evaluation Instrument. Retrieved from. <https://www.danielsongroup.org/resource-item/the-framework-for-teaching-evaluation-instrument-2/>.
- Darling-Hammond, L., 2000. Teacher quality and student achievement. *Educ. Policy Anal. Arch.* 8 (1).
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., Rothstein, J., 2012. Evaluating teacher evaluation. *Phi Delta Kappan* 93 (6), 8–15.
- Duffy, G.G., Miller, S.D., Parsons, S.A., Meloth, M., 2009. Teachers as metacognitive professionals. In: Hacker, D.J., Dunlosky, J., Graesser, A.C. (Eds.), *Handbook of Metacognition in Education*. Lawrence Erlbaum, Mahwah, NJ, pp. 240–256.
- Finn, A.N., Schrod, P., Witt, P.L., Elledge, N., Jernberg, K.A., Larson, L.M., 2009. A meta-analytical review of teacher credibility and its associations with teacher behaviors and student outcomes. *Commun. Educ.* 58 (4), 516–537.
- Graham, L., Berman, J., Bellert, A., 2015. Sustainable Learning. Cambridge University Press, Melbourne.
- Goddard, R.D., Hoy, W.K., Hoy, A.W., 2000. Collective teacher efficacy: its meaning, measure, and impact on student achievement. *Am. Educ. Res. J.* 37 (2), 479–507.
- Goe, L., Bell, C., Little, O., 2008. Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. Retrieved from. <http://search.proquest.com.ezp.lib.unimelb.edu.au/docview/881466794?accountid=12372>.
- Goe, L., Stickler, L., 2008. Teacher Quality and Student Achievement: Making the Most of Recent Research. Retrieved from.
- GRADE, 2013. In: Schunemann, H., Brozek, J., Guyatt, G., Oxman, A. (Eds.), *GRADE Handbook*. Retrieved from. <http://gdt.guidelinedevelopment.org/app/handbook/handbook.html>.
- Han, J., Yin, H., 2016. Teacher motivation: definition, research development and implications for teachers. *Cogent Educ.* 3 (1).
- Hattie, J., 2009. Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement. Routledge.
- Hattie, J., 2012. Visible Learning for Teachers: Maximizing Impact on Learning. Routledge, London.
- Hattie, J., 2018. The Major Influences of the Teacher. Unpublished: personal communication.
- Jennings, P.A., Greenberg, M.T., 2009. The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev. Educ. Res.* 79 (1), 491–525. <https://doi.org/10.3102/0034654308325693>.
- Jones, S., Bouffard, S., Weissbourd, R., 2013. Educators' social and emotional skills vital to learning. *Kappan* 94 (8).
- King, S.H., Watson, A., 2010. Teaching excellence for all our students. *Theor. Pract.* 49 (3), 175–184.
- Klassen, R.M., Chiu, M.M., 2010. Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102 (3), 741.
- The Australian Institute for Teaching and School Leadership, 2017. Australian Professional Standards for Teachers.
- Lord, R.G., De Vader, C.L., Alliger, G.M., 1986. A meta-analysis of the relation between personality traits and leadership perceptions: an application of validity generalization procedures. *J. Appl. Psychol.* 71 (3), 402.
- Lumpkin, A., Claxton, H., Wilson, A., 2014. Key characteristics of teacher leaders in schools. *Adm. Issues J. Educ. Pract. Res.* <https://doi.org/10.5929/2014.4.2.8>.
- MacSuga-Gage, A.S., Simonsen, B., Briere, D.E., 2012. Effective teaching practices that promote a positive classroom environment. *Beyond Behav.* 22 (1), 14–22.
- McDonald, C.V., 2016. STEM education: a review of the contribution of the disciplines of science, technology, engineering and mathematics. *Sci. Educ. Int.* 27 (4), 530–569.
- Miller, P., 1987. Ten characteristics of a good teacher. *Engl. Teach. Forum* 25 (1).
- Moore, J., Whitefield, V., 2008. Musing: a way to inform and inspire pedagogy through self-reflection. *Read. Teach.* 61 (7), 586–588.
- Moye, J., 2010. Making Your Classes Come Alive. Techniques, April.
- Muñoz, M.A., Chang, F.C., 2007. The elusive relationship between teacher characteristics and student academic growth: a longitudinal multilevel model for change. *J. Person. Eval.* 20 (3–4), 147–164.
- NSW Department of Education, 2018. Evaluative Thinking. Retrieved from. <https://education.nsw.gov.au/teaching-and-learning/professional-learning/evaluation-resource-hub/evaluative-thinking>.
- Nixon, A., Dam, M., Packard, A., 2010. Teacher dispositions and contract non-renewal. *Plan. Changing* 41 (3), 210–219.
- Okoli, A.C., 2017. Relating communication competence to teaching effectiveness: implication for teacher education. *J. Educ. Pract.* 8 (3), 150–154.
- Patton, M.Q., 2008. Utilization-Focused Evaluation, fourth ed. Sage, Thousand Oaks, CA.
- Pajares, M.F., 1992. Teachers' beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62 (3), 307–332.
- Popham, J., 2013. On serving two masters: formative and summative teacher evaluation. *Princ. Leadersh.* 13 (7), 18–22.
- Porter-Samuels, T., 2013. Raising Pasifika achievement: teacher cultural-responsiveness. *Kairaranga* 14 (2), 17–25.
- Reilly, R.C., Lilly, F., Bramwell, G., Kronish, N., 2011. A synthesis of research concerning creative teachers in a Canadian context. *Teach. Teach. Educ.* 27 (3), 533–542.
- Risko, V.J., Roskos, K., Vukelich, C., 2005. Reflection and the self-analytic turn of mind: toward more robust instruction in teacher education. In: Israel, S.E., Collins Block, C., Bauserman, K.L., Kinnucan-Welsch, K. (Eds.), *Metacognition in Literacy Learning: Theory, Assessment, Instruction, and Professional Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 315–333.
- Schonert-Reichl, K.A., 2017. Social and emotional learning and teachers. *Future Child.* 27 (1), 137–155.
- Seidel, T., Shavelson, R.J., 2007. Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Rev. Educ. Res.* 77, 454–499.
- Sherman, J., Rasmussen, C., Baydala, L., 2008. The impact of teacher factors on achievement and behavioural outcomes of children with Attention Deficit/Hyperactivity Disorder (ADHD): a review of the literature. *Educ. Res.* 50 (4), 347–360.
- Skaalvik, E.M., Skaalvik, S., 2010. Teacher self-efficacy and teacher burnout: a study of relations. *Teach. Teach. Educ.* 26 (4), 1059–1069.

- Slater, H., Davies, N.M., Burgess, S., 2012. Do teachers matter? measuring the variation in teacher effectiveness in England. *Oxf. Bull. Econ. Stat.* 74 (5), 629–645.
- Slatter, C.J., 2008. Talent Development for the Talent Developers: Identifying and Developing the Qualities Of Effective Teachers of Gifted Students, p. 15.
- Steele, N.A., 2010. Three characteristics of effective teachers. In: *Update: Applications of Research in Music Education*, vol. 28, pp. 71–78, 2.
- Strecher, V.J., McEvoy DeVellis, B., Becker, M.H., Rosenstock, I.M., 1986. The role of self-efficacy in achieving health behavior change. *Health Educ. Q.* 13 (1), 73–92.
- Sudkamp, A., Kaiser, J., Moller, J., 2012. Accuracy of teachers' judgments of students' academic achievement: a meta-analysis. *J. Educ. Psychol.* 104 (3), 743–762.
- Titsworth, S., Mazer, J.P., Goodboy, A.K., Bolkan, S., Myers, S.A., 2015. Two meta-analyses exploring the relationship between teacher clarity and student learning. *Commun. Educ.* 64 (4), 385–418.
- Vieira, C.R., Gaspar, M.F., 2013. PLENITUDE teacher education for effectiveness and well-being with neuro-linguistic programming. *US China Educ. Rev.* 1–17.
- Visible Learning, 2017. 250+ Influences on Student Achievement. Retrieved from: <https://www.visiblelearningplus.com/sites/default/files/250%20Influences.pdf>.
- Warren, J.M., Hale, R.W., 2016. The influence of efficacy beliefs on teacher performance and student success: implications for student support services. *J. Ration. Emot. Cogn. Behav. Ther.* 34 (3), 187–208.
- Webster, C.A., 2010. Increasing student motivation through teacher communication: six essential skills every physical educator should master. *J. Phys. Educ. Recreat. Dance* 81 (2), 29–33.
- Westwood, P., 2016. What Teachers Need to Know About Differentiated Instruction. Australian Council for Educational Research, Melbourne.
- Williams, M., Burden, R., 1997. *Psychology for Language Teachers*. Cambridge University Press, Cambridge, UK.
- Wyatt, T., 2017. Developing Evaluative Thinking and Evidence-Based Practice: A Synthetic Case Study. Retrieved from: https://research.acer.edu.au/cgi/viewcontent.cgi?article=1316&context=research_conference.
- Yost, D.S., Sentner, S.M., Forlenza-Bailey, A., 2000. An examination of the construct of critical reflection: implications for teacher education programming in the 21st century. *J. Teach. Educ.* 51 (1), 39–49.
- Zee, M., Koomen, H.M.Y., 2016. Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: a synthesis of 40 Years of research. *Rev. Educ. Res.* 86 (4), 981–1015.

Sources identified for supplementary evidence

Creativity, decision making

- De Bruin, L.R., Harris, A., 2017. Fostering creative ecologies in australasian secondary schools. *Aust. J. Teach. Educ.* 42 (9), 23–43.
- Ghanizadeh, A., Jahedizadeh, S., 2016. EFL teachers' teaching style, creativity, and burnout: a path analysis approach. *Cogent Educ.* 3 (1), 1151997.
- Khodabakhshzadeh, H., Hosseinnia, M., Moghadam, H.A., Ahmadi, F., 2018. EFL teachers' creativity and their teaching's effectiveness: a structural equation modelling approach. *Int. J. Instruct.* 11 (1), 227–238.
- Prayekti, 2018. The influence of cognitive learning style and learning independence on the students' learning outcomes. *High. Educ. Stud.* 8 (2), 37–46.
- Ramos, N.C., McCullick, B.A., 2015. Elementary students' construct of physical education teacher credibility. *J. Teach. Phys. Educ.* 34 (4), 560–575.

Motivation

- Abazaoglu, I., Aztekin, S., 2016. The role of teacher morale and motivation on students' science and math achievement: findings from Singapore, Japan, Finland and Turkey. *Univers. J. Educ. Res.* 4 (11), 2606–2617.
- Appova, A., Arbaugh, F., 2018. Teachers' motivation to learn: implications for supporting professional growth. *Prof. Dev. Educ.* 44 (1), 5–21.
- Daniels, E., 2017. Curricular factors in middle school teachers' motivation to become and remain effective. *RMLE Online* 40 (5).
- Keller, M.M., Neumann, N., Fischer, H.E., 2016. The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *J. Res. Sci. Teach.* 52 (5), 586–614.
- Shifrer, D., Turley, R., Heard, H., 2017. Do teacher financial awards improve teacher retention and student achievement in an urban disadvantaged school district? *Am. Educ. Res. J.* 54 (6), 1117–1153.

Cultural competence

- Brinkley, B., Hines, E., Jones, A., McMillan, E.G., Sturdivant, B., Walker, M., 2018. Fixing systems, not kids: changing the narrative of black males in Guilford county schools. *VUE* (28), 19–25.
- Gambrell, J.A., 2017. Is Culturally Responsive Pedagogy Enough? Toward Culturally "Real"-Evant Curriculum. *Democracy & Education*.
- Landa, M.H., Odone-Holm, J., Shi, L., 2017. Education abroad and domestic cultural immersion: a comparative study of cultural competence among teacher candidates. *Teach. Educ.* 52 (3), 250–267.
- Lewthwaite, B., Owen, T., Doiron, A., Renaud, R., McMillan, B., 2014. Culturally responsive teaching in Yukon First Nation settings: what does it look like and what is its influence? *Can. J. Educ. Adm. Pol.* 155.
- Mette, I.M., Nieuwenhuizen, L., Hvidston, D.J., 2016. Teachers' perceptions of culturally responsive pedagogy and the impact on leadership preparation lessons for future reform efforts. *NCPEA Int. J. Educ. Leadersh. Prep.* 11 (1).
- Stowe, S., 2017. Culturally responsive teaching in an Oglala Lakota classroom. *Soc. Stud.* 108 (6), 242–248.
- Warren, C.A., 2018. Empathy, teacher dispositions and preparation for culturally responsive pedagogy. *J. Teach. Educ.* 69 (2), 169–183.
- Zygmunt, E., Cipollone, K., Tancock, S., Clausen, J., Clark, P., Mucherah, W., 2018. Loving out loud: community mentors, teacher candidates, and transformational learning through a pedagogy of care and connection. *J. Teach. Educ.* 69 (2), 127–139.

Self-reflection and reflection

- Ashraf, H., Zolfaghari, S., 2018. EFL teachers' assessment literacy and their reflective teaching. *Int. J. Instruct.* 11 (1), 425–436.
- Farrell, T.S., 2016. TESOL, a profession that eats its young! The importance of reflective practice in language teacher education. *Iran. J. Lang. Teach. Res.* 4 (3), 97–107.
- Kheirzadeh, S., Sistani, N., 2018. The effect of reflective teaching on Iranian EFL students' achievement; the case of teaching experience and level of education. *Aust. J. Teach. Educ.* 43 (2), 143–156.
- Lotter, C.R., Miller, C., 2016. Improving inquiry teaching through reflection on practice. *Res. Sci. Educ.* 47, 913–942.
- Yildirim, T., 2017. An examination of geography teachers' reflective thinking tendencies. *Int. J. High. Educ.* 6 (6), 78–90.

Collegiality

- Banerjee, N., Stearns, E., Moller, S., Mickelson, R.A., 2016. Teacher job satisfaction and student achievement: the roles of teacher professional community and teacher collaboration in schools. *Am. J. Educ.* 123 (2).

- Kunnari, I., Ilomaki, L., Toom, A., 2018. Successful teacher teams in change: the role of collective efficacy and resilience. *Int. J. Technol. High. Educ.* 30 (1), 11–126.
- Noland, A., Richards, K., 2015. Servant teaching: an exploration of teacher servant leadership on student outcomes. *J. Scholarsh. Teach. Learn.* 15 (6), 16–38.
- Sebastian, J., Allensworth, E., Huang, H., 2016. The role of teacher leadership in how principals influence classroom instruction and student learning. *Am. J. Educ.* 123 (1), 69–108.
- Sun, M., Loeb, S., Grissom, J., 2015. Building Teacher Teams: Evidence of Positive Spillovers From More Effective Colleagues. CEPA Working Paper no. 15–20.

Values, morals and ethics

- Benninga, J.S., Sparks, R.K., Tracz, S.M., 2011. Enhancing teacher moral judgment in difficult political times: swimming upstream. *Int. J. Educ. Res.* 50, 177–183.
- Chang, D., Heesoon, B., 2016. Self-with-other in teacher practice: a case study through care, Aristotelian virtue, and Buddhist ethics. *Ethics Educ.* 11 (1), 17–28.
- Crossleya, M., Koya Vaka'utab, C.F., Lagib, R., McGrathc, S., Thamanb, K.H., Waqailitib, L., 2017. Quality education and the role of the teacher in Fiji: mobilising global and local values. *Compare* 47 (6), 872–890.
- Fenech, M., Lotz, M., 2018. Systems advocacy in the professional practice of early childhood teachers: from the antithetical to the ethical. *Early Years* 38 (1), 19–34. <https://doi.org/10.1080/09575146.2016.1209739>.
- Jackson, R., Everington, J., 2017. Teaching inclusive religious education impartially: an English perspective. *Br. J. Relig. Educ.* 39 (1), 7–24.
- Nucci, L., Creane, M.W., Powers, D.W., 2015. Integrating moral and social development within middle school social studies: a social cognitive domain approach. *J. Moral Educ.* 44 (4), 479–496.
- Samanci, N.K., 2015. A study on the link between moral judgment competences and critical thinking skills. *Int. J. Environ. Sci. Educ.* 10 (2), 135–143.
- Switala, E., 2015. Teachers' values related to sustainable development in polish and Latvian secondary schools. *Discourse Commun. Sustain. Educ.* 6, 127–133.

Teacher education and the arts: challenges, tensions, opportunities

Robyn Ann Ewing^{a,b}, ^a Sydney School of Education & Social Work, NSW, Australia; ^b Faculty of Arts and Social Sciences, Australia University of Sydney, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	304
Defining “the arts”	305
A brief snapshot of relevant research	305
Embedding the creative arts as a core formal curriculum component: exemplars	306
Wales	307
Finland	307
Singapore	307
Challenges and tensions for teacher education in the arts	307
Understanding teaching as artistry	309
Opportunities	309
Teacher education and arts professional learning during the pandemic	311
Concluding comments	311
References	312

Glossary

Terms important in understanding the tenuous relationships between the Arts and education and teacher education as discussed in this article include:

Arts and education is an overarching term emphasizing equal status and acknowledging a reciprocal and interactive relationship.

Arts in education and education through the Arts capture the concepts of engaging in arts experiences and using arts processes and strategies as pedagogical tools to facilitate learning and foster creative and flexible thinking connecting different kinds of knowledge

Education in the Arts and arts education describe the teaching and learning *about* different arts disciplines

quality in arts learning experiences manifest all-encompassing engagement; involvement with authentic, artistic processes and materials; an exploration of “big ideas” about both art and human experience; and direct experiences with completed or in-process works of art (Seidel et al., 2009)

Teacher education refers to both initial or pre-service education needed to gain a qualification to teach as well as ongoing professional learning for experienced teachers

Teaching Artist defined by Booth (2010/2020) as a person who is “... a practicing professional artist ... who can effectively engage a wide range of people in learning experiences in, through, and about the arts” (para 5)

Introduction

Arts education matters because people trained in the arts play a significant role in the innovation process the arts are an essential part of human heritage and of what makes us human, and it is difficult to imagine an education for better lives without arts education.

(Winner et al., 2013, p. 13–14)

The Arts are as old as humanity because they are central to who we are as human beings: they connect to our imaginations and creative and expressive capacities. Reflect for a moment on the emotional way we respond to a moving orchestral, dance, film, a theater performance or an artwork; the curiosity or wonder triggered by an imaginative arts process or experience; our expanded capacity for empathy and compassion on identifying with a compelling fictional character in a novel; and the personal satisfaction that arises when we find a way to express a challenge or dilemma provocatively through one or several artforms.

Research and scholarly writing over the last three decades have unequivocally supported the imperative of embedding the Arts throughout our lives to ensure our social and emotional wellbeing (for example, *A New Approach*, 2020; *World Health Organisation*, 2019). The Arts are important for who we are and must never be viewed solely as “instrumental” to the achievement of other learning outcomes (McCarthy et al., 2004).

Nevertheless, in the education context, it is increasingly clear from international research and practice that engaging teachers and learners in quality arts processes and experiences fosters imaginative learning and creative thinking (Martin et al., 2013; Gibson and Ewing, 2020; Ewing 2010), transforms outdated transmissive curricula (for example, Burnard and Colucci-Gray, 2020), enriches our understanding of ourselves (Eisner, 2002), enables us to develop compassion for others (Biesta, 2013) and provides new possibilities for the way we live in our global communities (Greene, 1995).

This article first explains the use of the term “The Arts” and provides a brief snapshot of relevant research before arguing that pre-service and ongoing in-service teacher education should be arts-rich – both in specific arts disciplines and as critical, quality pedagogy (Ewing, 2002, 2006, 2019) across the curriculum. All teachers must have the confidence and expertise to provide quality arts learning experiences in their classrooms at all levels of education. Some of the major challenges and tensions that frequently obstruct the realization of this proposal are then discussed. Finally, current opportunities are considered through several contemporary teacher education initiatives.

Defining “the arts”

Cultural and social contexts shape and sustain and, conversely, inhibit the flourishing of different arts disciplines (Bamford, 2006) in different communities. Conscious of the ever-changing nature of contemporary arts experiences and practices, the term “the Arts” has been used throughout this article as a collective, shorthand term to refer to major creative arts disciplines and the embodied knowing they develop. The following arts disciplines are named in many formal curriculum documents all over the world and are the key artforms for the Arts referred to in this article:

- Dance involves moving the body expressively and intentionally to make meaning, represent, question and celebrate experience.
- Drama employs dramatic elements and processes to create and explore roles, characters, relationships as well as actual and fictional worlds.
- Literature includes imaginative novels, short stories, plays, poetry and picture books that use evocative, expressive language to explore and communicate experiences, perspectives and ideas.
- Media arts incorporate traditional and new technologies to tell stories creatively and includes, for example, photography, film, video and digital animation.
- Music is aurally based and involves listening, composing, performing and critical appreciation.
- Visual arts includes two- and three-dimensional design and art-making (painting, drawing and sculpture) to communicate, challenge, express and appreciate ideas.

Each arts discipline represents different ways of knowing (Habermas, 1972), making meaning and communicating about ourselves, others and the world, and so each embodies a different literacy (NAAE, 2019). All art forms must therefore be studied, understood and appreciated for their intrinsic worth (Bryce et al., 2004), discrete knowledges, understandings and skills. At the same time, all art forms involve processes that clearly resonate with creative dispositions and 21st century skills (National Education Association, 2013; Jefferson and Anderson, 2017; Gibson and Ewing, 2020). These processes include imaginative play; experimentation or exploration; provocation; use of metaphor; expression or representation; communication; design; and the artistic or esthetic shaping of the body or other media (Ewing, 2010). Teachers need to embed these processes and acknowledge how they are inextricably linked to both their own already rich imaginative and creative abilities and the nurturing of their learners’ social and emotional wellbeing and academic success.

While the knowledge about discrete disciplines of art-making processes must be upheld, this author strongly believes that artistic and educational collaborations should also be based on authentic interdisciplinary and multidisciplinary explorations (Harris and Carter, 2021; Gibson and Ewing, 2020; Bresler, 2004). It is damaging when competitive arguments about which arts discipline should receive the most attention dominate the educational agenda. Interdisciplinary, organic, ways of thinking about the Arts have the potential to provide a much-needed shift in epistemological grounding (Gadsden, 2008).

Developing teacher confidence and expertise in different artforms is insufficient, however, if pre-service and experienced teachers do not understand and celebrate their own artistry as educators. For the Arts to play an important role in formal educational contexts it is imperative that both pre-service and well-established teachers see their teaching as artful. In the author’s experience, teachers sometimes describe themselves as “uncreative” or “not artistic”, and are unaware of the esthetic (Hobbs, 2012) that underpins their teaching.

A brief snapshot of relevant research

Decades of international research have highlighted the potential role that the Arts can and should play across early childhood, elementary, secondary and tertiary education to motivate and engage learners (Fleming et al., 2016) as well as to enhance their social and emotional wellbeing and identity development that, in turn, ensures improvement in academic achievements more generally (for example, Saunders, 2019; Catterall, 2013, 2009; Deasy, 2002). Quality arts processes and experiences are inextricably

linked to the development of dispositions/habits of mind that are critical for successfully living with post-normal (Sarder, 2010) 21st century uncertainties (for example, Gibson and Ewing, 2020; Winner et al., 2013; Ewing, 2019; 2010; Bamford, 2006; UNESCO, 2006; Miller and Saxton, 2016, 2004) and improve our achievement in other academic areas.

Findings from a range of research and practitioner sources over more than two decades demonstrate that the Arts, while in no way the solution for all educational and social problems, have the potential to help us address many of the habitual issues routinely embedded in current educational institutions. These include the motivation and engagement of young people in authentic learning (Wyn, 2009) and the inequities of the competitive academic curriculum (Connell, 1994). Many successful arts projects also depict the Arts as providing potential catalysts for social transformation in the community more generally (for example, see 25 years of projects coordinated by Big hArt, <https://www.bighart.org/what-we-do/past-projects/>).

Arts-rich experiences can give teachers and students the courage to make and represent meanings aesthetically (Greene, 1978), to take risks and fail memorably (Zusak, 2014) to genuinely collaborate (Heath, 2000), and, critically, to develop the resilience to try again (O'Connor et al., 2020). The process of creating or responding to an artwork, a theater or dance performance, a piece of music, a film or writing a story, can trigger the asking of big questions that inevitably challenge the stereotypes that prevail too often within individual and community life. Experiencing and engaging in quality arts processes and experiences are therefore necessary for all children and young people as part of their everyday learning – through these processes they gain access not only to the real world but to other possible imaginative worlds, and begin to make coherent meanings about their own identities, their worlds and the others in these worlds.

Increasingly, neuro-scientific research confirms that genuine collaborative strategies developed through creative dance, music, visual arts and drama activities can achieve an increase in brain synchrony encouraging the breaking up of old patterns of behavior (Damasio, 2004; Freeman, 1995). The ability to empathize or walk in someone else's shoes emphasized in educational drama is also important in understanding diversity and the need to be inclusive (O'Toole and Dunn, 2020; Karaolis, 2020; Dunn et al., 2019, 2020; Campbell and Hogan, 2019; Donelan, 2012; Heath, 2000).

Thousands of arts and education programs and initiatives all over the world strive to improve educational outcomes or change problematic social situations through embedding arts experiences in formal or informal learning contexts (A New Approach, 2020; Ewing, 2010). Too numerous to mention, these programs and initiatives have varying specific goals and objectives primarily concerned with improving the life chances of children or adults considered to be vulnerable or "at risk". Given the growing research evidence about the centrality of the Arts for cognitive and emotional well-being and the impact it can have on deep learning, arts education needs to be seen as a core curriculum component and be accessible to all students from all backgrounds at each level of education. Yet, despite the importance of such programs to their participants and their potential generalized social impact, many of them are initiated and implemented by not-for-profit, philanthropic organisations rather than education departments.

Despite growing evidence that the Arts can initiate curriculum reform, provision of arts education and education through the arts professional learning for pre-service and experienced teachers has been continually cut alongside funding for many arts organisations in western countries over the last three decades. While specialist arts secondary teachers usually qualify through a first degree in their chosen arts field(s), this article particularly highlights the need for *all* teacher education courses to incorporate meaningful education about, in and through the Arts.

Almost in contradiction to the rhetoric that proclaims the importance of the Arts and their relationship to creativity, problem solving and flexibility, neoliberal government policies seek linear recipes for success in literacies defined in very reductive ways (Darling-Hammond, 2010) through assessment of achievement based on high-stakes testing (Nussbaum, 2010).

Although the Arts are flourishing in some countries, including Wales, Finland and Singapore, they have been radically cut in others. In Australia for example, the pandemic saw many Creative Arts Degrees and professional learning programs cut across the country. Resource funding for quality arts education in schools has historically been unevenly applied and distributed and consequently, and even more so since COVID-19. It is often the most vulnerable school communities that miss out-as many advantaged schools can ask parents to pay for extra-curricular arts programs and experiences – when those who are disadvantaged are the ones that need these experiences the most (Vinson and Rawsthorne, 2015; Ewing, 2010).

Acknowledging the value of arts education in developing creativity emphasizes the importance of the interdisciplinary nature of key competencies, including cultural awareness (Harris and Carter, 2021). They represent an important lens through which to view the Arts in teacher education (President's Committee on the Arts and Humanities, 2011; Council of the European Union's Agenda for Culture, 2007).

The following section highlights several examples where the Arts has been foregrounded as a core curriculum component in schools alongside the provision of ongoing teacher education.

Embedding the creative arts as a core formal curriculum component: exemplars

Despite the trends mentioned above, there are a number of countries where arts in education has been prioritized to ensure it is a core curriculum component and where teacher professional learning needs have been addressed. Several examples are highlighted briefly below.

Wales

In 2015 the need for the centrality of the Arts in the formal Welsh curriculum was recognized and has been actioned by the Welsh government over the last few years (Welsh Government, 2015). *Successful Futures* (Donaldson, 2015) reviewed curriculum and assessment in Wales building on an earlier report (Smith, 2013) to strongly endorse the recommendation to place arts and creativity at the heart of the Welsh curriculum. Expressive Arts is now one of six “areas of learning and experience” in the Welsh national curriculum for ages 3–16 adopted in place of traditional, often silo-ed curriculum subjects. The report recommended that the entirety of the school curriculum should be designed to help all children and young people become:

- ambitious, capable learners, ready to learn throughout their lives
- enterprising, creative contributors, ready to play a full part in life and work
- ethical, informed citizens of Wales and the world, ready to be citizens of Wales and the world
- healthy, confident individuals, ready to lead fulfilling lives as valued members of society.

In the Expressive Arts curriculum, it is intended that learners develop their creative appreciation and talent, and their artistic and performance skills. It spans the making, performance, expression and appreciation of art, drama, music, dance, film and digital media and “many opportunities” are provided for students to visit theaters and galleries. Funding is available to bring artists and musicians into the classroom. The *All Wales Arts and Education Program* supports the partnerships between artists, creative and heritage organisations and schools by building bridges to support teachers and enrich arts teaching and experiences across the curriculum in schools. Professional development events for teachers “focus on contemporary arts practices and historic arts and culture to equip expressive arts teachers with confidence to present and respond to the professional practice of artists and creative practitioners”. “Arts Champions” visit school and cultural venues to mentor expressive arts teachers. See <https://nawr.cymru/creative-learning/>.

Finland

In Finland, regarded as a world leader educationally, the Arts also have traditionally played an important role in both school and cultural life. This is often attributed to the influence of the concept of the Scandinavian folk schools. Finland education authorities have long considered the Arts in education vital across their system. Interestingly there has not been the same privileging of STEM (science, technology, engineering, and mathematics) over the arts as has been the case in many other western countries. The Finnish Ministry (2010), acutely aware of the inequitable distribution of artforms and arts resources across different regions, implemented strategies to widen the opportunities for children to pursue artistic interests, notably by establishing a network of 11 children’s cultural centers.

Singapore

In Singapore over the last decade there has been a stronger and explicit focus on the importance of quality arts experiences for all students across the formal mainstream curriculum from preschool to the conclusion of secondary school. There is also an emphasis on providing greater opportunities for sustained partnerships between Singapore artists and arts educators in school to deepen their creative practices and increase learners’ engagement in and through the Arts. Professional learning initiatives have been developed to support both artists and teachers.

Yet there are many other examples worldwide where the arts in-and-through education are not flourishing despite the policy rhetoric. The following section considers some of the ongoing challenges and tensions that continue to work against ensuring all teachers are well prepared to embed the Arts in learning.

Challenges and tensions for teacher education in the arts

The challenges and tensions for teacher education in the Arts continually recur in any conversations about arts and education. Although these are inter-related, they are briefly discussed separately below:

1. A perception that arts disciplines represent a form of cultural elitism and are only accessible to the privileged. The persistence of the continuing hierarchy of curriculum subjects in many neoliberal western education systems – with the most notable priorities being narrow understandings of literacy and numeracy often measured by high stakes tests (Barton and Ewing, 2017). Competitive academic curricula (Connell, 1994) is often the main “passport” for entering tertiary education. It also underlines that the Arts are often still regarded as elite or as a “soft” or extra-curricular offering rather than core to a child’s emotional and social wellbeing and academic success.
2. In many western education systems, teachers are constantly exhorted to improve literacy and numeracy test results rather than to focus on a rich integration of curriculum areas. Winner et al. reported in 2013 that the share of time allocated for arts education for children aged between 9 and 11 has diminished in 10 of the 18 OECD nations. This observation is supported by evidence from individual countries:

- i) The [University of Warwick Report \(2015\)](#) argued that between 2003 and 2011 there was evidence that the Arts and creative processes were progressively being “squeezed out” of the English curriculum.
 - ii) In the United States of America the *No child left behind Act (2001)* has been blamed for diminished arts offerings in many public schools. In 2014 The US National Center for Education reported a huge decline of the arts in public schools with six percent of public elementary schools (serving 1.3 million students) failing to offer music, 17% (serving nearly 4 million students) not offering visual arts and ninety-seven percent of public elementary schools (serving 23 million students) not offering any instruction in dance or theater. The secondary public school offerings are not much better with 88% (18 million students) excluding dance and 55% theater.
 - iii) In Australia in 2020, the then federal Education minister recently flagged a review of what he described as “the crowded curriculum” calling for a return to the core subjects of mathematics and English. In a news report he bemoaned the sacrifice of “quality over quantity” and called for a concentration on “essential content” expressing wariness of so called “soft skills”, of “problem solving, critical thinking, creativity and teamwork” ([Hunter, 2011](#)). Pervasive attitudes like these also contribute to a regime of standardized testing that denies a focus on more creative approaches to teaching and learning.
 - iv) In South Africa [Vermeulen’s \(2009\)](#) doctoral research found that few teachers had expertise in more than one arts discipline and that practical arts activities were often sacrificed for more literacy and numeracy instruction.
 - v) For thousands of years in China, Confucianism provided a basis for a holistic approach to Chinese Arts education. “Six Arts” (etiquette, music, archery, charioteering, calligraphy and mathematics) were emphasized from early childhood. The *Annual Report of Chinese Arts Education (2012)* foregrounds music and visual arts as the main components of preschool arts education ([Li, 2013](#); [Zhang, 2012](#)). Some Chinese educators suggest there is a biased focus on a skills-only approach to teaching in the Arts and claim that many children are expected to follow the teacher’s instruction while more spontaneous exploration and appreciation are ignored ([Tau, 2019](#); [Zhang, 2012](#)).
3. The literary arts are often pitted against other arts disciplines ([Barton and Ewing, 2017](#)) and often excluded from the arts syllabus documents (even though they are usually an important component of English syllabi). This is another example of the lack of attention to interdisciplinary studies. There are therefore ongoing tensions between the Arts and literacy – in fact they are sometimes conceived as competing for curriculum time, resources and physical space.
- Allocated time for the Arts in pre-service teacher education and professional learning programs in many countries has continued to be cut over several decades. Inadequate or almost non-existent pre- and in-service teacher education and professional learning in the Arts for early childhood, elementary and non-arts secondary teachers has resulted in many teachers articulating a lacking of confidence and expertise to embed the Arts in their pedagogy in one or more artforms ([Ewing and Saunders, 2016](#)) suggesting there is little chance for Arts pedagogy to improve. At Murdoch University in Western Australia, for example, elementary pre-service teachers have just 6 hours in their program to learn about the Creative Arts ([Pascoe, 2015](#)). Reduced time allocation for learning about specific arts disciplines in teacher education courses has resulted in many early-career teachers reporting that they lack the knowledge, skills and understanding in one or more of the art forms that they are expected to teach. Additionally some research suggests that pre-service teachers do not always see the relevance of the Arts (see, for example, [Lemon and Garvis, 2013](#)) for their classrooms.
4. In many schools’ arts activities can be “soul destroying and meaningless”, lacking in depth and dimension ([Boyd, 2009](#)).
 5. At times intense political and often destructive arguments within and between the arts disciplines themselves ([Pascoe, 2015](#)) have created other kinds of tensions. The paucity of funding for arts and education programs has contributed to some re-emergence of fragmentation and competition between different art forms as they fight for the meagre dollars available, with the consequent assertion that some arts disciplines are more important than others (e.g. [Thomson and Sefton-Green, 2011](#)).
 6. The assessment of esthetic knowledge is complex for schools and teachers, because it is wholistic engaging the brain, body and emotions simultaneously in a range of symbolic languages and forms. Current traditional assessment systems concentrate on those cognitive aspects of knowledge that can be made explicit and learned propositionally, just in words or numbers.
 7. Increasing pressure for schools to include the study of new media (including film, photography and digital arts) in their curricula, and to enable students to use ICT as part of the creative process is also costly and very often not viable for some less advantaged schools, privileging once again those students whose parents can afford to pay for additional teachers.
 8. Misconceptions about teaching arts disciplines continue. One such misconception is the deeply held perception that being able to do or teach an art form requires special innate gifts only attainable by, or given to, a chosen few. Such feelings are sometimes influenced by unfortunate prior personal arts experiences in the educators’ own learning journeys.
 9. There is also evidence of a mismatch between what is expected of early childhood and generalist elementary teachers in teaching about, in and through the Arts and what they themselves believe they can do ([Alter et al., 2009](#); [Garvis and Riek, 2010](#)).
 10. While there is much correlational evidence about quality arts and the improvement of student learning across the curriculum, more systematic large-scale research quantifying their impact ([Ewing, 2010](#)) in ways preferred by governments is needed.

Although policy rhetoric often values the place of the Arts in curricula, there needs to be ongoing and sustained teacher education in the Arts, adequate resourcing, teacher professional learning and prioritizing if the Arts are to take their rightful place in schools and

other learning institutions including universities. The next section particularly focuses on the need for teachers to understand their own artistry as educators.

Understanding teaching as artistry

As mentioned above, often generalist early childhood, elementary and non-arts secondary teachers at all stages of their careers articulate a lack of confidence and expertise in teaching different art forms (Gibson and Ewing, 2020; Uptis et al., 2010). Many bring baggage from childhood experiences that have affected their confidence with one or more artforms.

In addition to developing knowledge and expertise about the teaching of mandated arts disciplines in early childhood and elementary contexts, teachers must see their teaching as artistry and become aware of the role the Arts and arts processes can play in enriching their creative pedagogical perspectives in their curriculum area(s). As Dinham (2019) argues, all pre-service and experienced teachers need the time and resources within their teacher education to engage in learning through the Arts and to reflect on these experiences. They need to be challenged to think creatively about their own teaching and the esthetics of teaching.

Offering innovative units of study that encourage early-career teachers to engage in Creative Arts experiences empowers them to be actively involved in designing their learning against a backdrop of traditional transmissive educational practices commonplace in many institutions of higher education. Young teachers need to see meaningful connections across the Creative Arts and experience different routes of learning that can be accessed through the Creative Arts.

If early-career teachers are to feel empowered to change transmissive and traditional educational practices that often exist in schools, it is vital that they acknowledge, understand and question their own emerging teacher identity and many taken-for-granted aspects of teaching practice. More than a decade ago Ewing and Gibson (2007/2015) documented a then five year journey to design and develop two elective final year Bachelor of Education Creative Arts units. These aimed to encourage final year pre-service elementary teachers to teach creatively through integrated arts processes and experiences. It was demonstrated that, through such processes, early career teachers reconnected with their own creative potential. Engaging in meaningful artistic creative processes helped them consider different ways of learning, understand the importance of the Arts for all learners and inspired them to offer similar learning opportunities for the children they subsequently taught.

This section has emphasized the importance of quality education in and through the Arts, and has highlighted the insecurities many teachers articulate about their ability to use arts processes in their teaching. Governments and universities must re-consider the initial preparation of all teachers to give them confidence to embed the Arts in their teaching and learning practices. Funding for ongoing professional development to enable both experienced generalist and arts teachers to continue to update their knowledge, expertise and skills in embedding the Arts in learning, and to undertake their own research and reflection on the impact of the Arts as pedagogy should also be prioritised.

The next section looks at some innovative teacher professional learning initiatives which suggest broader potential opportunities.

Opportunities

As demonstrated earlier, provision of quality teacher preparation in the Arts for early-career early childhood, elementary and non-arts secondary teachers has diminished and ongoing professional learning in the Arts has been almost non-existent. It has therefore been necessary for arts teachers, artists and teaching artists (Booth, 2010) to find innovative ways to share their skills and expertise by working collaboratively across universities, schools and the broader community. Many such initiatives involve teaching artists and artists-in-residence programs shared with community arts organisations, galleries and museums to underpin and support sustained teacher education in the Arts. This has been termed a co-participative approach (Nilson et al., 2013). Where possible the involvement of families in developing creative arts skills and interest is also developed.

Uptis et al. (2010) examined four North American professional programs aimed at upskilling generalist teachers in the Arts. While a number of teachers reported personal transformations, others reported that they experienced increased self-confidence and a greater awareness of their creativity and self-expression as well as moments of frustration. In many instances participants made a long-term commitment to the Arts in their classrooms and shared these new understandings with their students' parents. In Canada, masterclasses, teacher resources and internships for emerging artists are funded by *Arts Alive* (ArtsAlive.ca) established by the National Youth and Education Trust.

In the United Kingdom, *The Creative Partnerships Program* (2002–11) aimed to develop the arts skills and knowledge of more than a million young people and 90,000 teachers across England by funding 8000 innovative partnerships between creative professionals and schools. Many were long term and aimed to raise the aspirations of young people to give them more options for their future. Improvements in student attendance, engagement, behavior and academic achievement were documented (Thomson and Sefton-Green, 2011). Looking particularly at teacher professional learning, researchers noted that, for significant and effective professional learning to occur, teachers needed to go beyond the application of skills learned to understand the deeper underlying principles of the processes so they could generate their own planning and implementation of this work subsequently (Hatcher, 2011).

Most recently a very comprehensive longitudinal research project was conducted by the [Arts Council, England \(2019\)](#). The Royal Shakespeare Company (RSC), TATE and the University of Nottingham partnered to investigate the benefits of arts and cultural education for students and their teachers from 2015–2018. *Tracking Arts Learning and Engagement* (TALE) listened to the voices and opinions of young people aged between 14 and 18. [Cole and Thomson \(2021\)](#) analyzed students' responses as well as those of 63 teachers to provide new understandings about and insights into why sustained engagement with arts and cultural education makes such a difference. The clear overall message from students was that arts and cultural learning fosters their imaginations and creative intuitions as well as their self-efficacy in ways that other learning does not. The TALE project also investigated the different but successful models of continuing professional learning for teachers offered by Tate and the RSC.

The Kennedy Center in Washington DC has offered a variety of arts teacher education programs, *Changing Education through the Arts* (CETA) since 1976 aiming to enable all teachers to ensure the Arts are an integral part of every child's education. These intensive professional development programs are designed to teach educators about the Arts and about arts integration techniques. They include mentoring and co-teaching by expert teaching artists in the classroom. CETA partners with schools that have determined a school-wide integrated focus on the Arts and concentrates on teacher professional learning. It aims to: establish and nurture a network of schools committed to arts integration as an approach to teaching and learning; build teacher planning, teaching and assessment skills and practices in arts integration; create supportive school cultures that enable sustained teacher professional learning and collaboration; and share knowledge with other organisations with similar programs. Evaluations of CETA to date have reported that participant teachers nurture their own creativity as well as knowledge, skills and appreciation of the Arts and become more confident in working collaboratively to meaningfully connect art forms with other curriculum subjects ([Silverstein et al., 2010](#)). Similarly the [Lincoln Centre](#) in New York offers *Activate* programs that include speakers, immersive workshops and mentoring that connect arts professionals with the education community.

In Australia there are also many teacher professional learning programs involving artists-in-residence in schools that aim to develop teacher knowledge, expertise and confidence in embedding the Arts in the curriculum ([Hunter, 2005, 2011](#)).

Y-Connect was a unique three year (2016–18) government funded multi-arts project partnership between Griffith University and Yeronga State High School, a multicultural Queensland school with students from over 60 countries including a significant proportion from refugee backgrounds ([Dunn et al., 2019](#)). It aimed to support students' learning and sense of belonging by using arts-based pedagogy across the curriculum. Artists and artist-teachers partnered with teachers to deliver both classroom-based learning and extra-curricular activities. Six case studies have been developed to exemplify the outcome. One focuses on the impact of the project on both mainstream and English as an Additional Language/Dialect students. Another examines small groups of newly arrived year 10 students who are using drama based approaches to learning English. As the project lead researcher, Julie Dunn, reported:

Maths lessons involving dance artists, English lessons co-taught by actors and playwrights, students singing and composing with Opera Queensland or doing acrobatic tricks with Circa Circus team members, there is no such thing as a normal school day for students at Yeronga State High School ... In the 21st century, finding ways to engage young people in learning is always a challenge, so we are very pleased to report the energising impact the project has had on student learning and the flow-on effect this has had on their academic achievement.

(Dunn et al., 2019).

Since 2009 the Sydney Theater Company (STC) in partnership with The University of Sydney, has developed and implemented a teacher professional learning program, the *School Drama Program*. This program fosters teacher confidence and expertise in the use of drama processes with quality literature to enhance student identity, empathy and compassion as well as deepening their English and literacy skills. It is based on and extends [Ewing's \(2002, 2006\)](#) academic co-mentoring work with teachers over 17 years at Sydney's Curl Curl North Public School.

In the Classic form of *School Drama*, actors or teaching artists work alongside participant elementary teachers in their classrooms over seven weeks, modeling the use of educational drama as a powerful medium for improving students' English and literacy learning ([Ewing and Saunders, 2016, 2017](#)). At the heart of the program is ongoing teacher professional learning and a co-mentoring relationship between elementary teachers and experienced actors that re-awakens the teachers' own artistry. Using drama strategies to improve student engagement in a selected area of English and literacy, in turn improves the overall learning outcomes of students ([Ewing et al., 2015; Saunders, 2015, 2019](#)). The program now reaches all education sectors in every Australian state. Two of the pilot study schools remain involved in the program more than a decade later. All participant teachers and actors are initially involved in 3 h professional learning workshop so they experience the drama strategies and conventions first hand and reflect on how they apply to language and literacy learning. In addition principals and executive are encouraged to participate in these initial workshops.

The teaching artist and class teacher then plan and team teach a seven week program to meet the selected English and literacy outcome(s) incorporating relevant drama strategies and authentic literary texts. Each weekly session spans 60–90 min and after each session the teacher and teaching artist engage in a short debriefing session. Benchmarking of the students' capacities in the identified literacy area occurs before and after the program to assess student starting points and progress over the program. Pre- and post-program surveys are conducted with teachers and debriefing focus sessions are undertaken with teaching artists after each program. A meta-analysis ([Gibson and Smith, 2013](#)) and twelve case studies have been undertaken ([Karaolis, 2020; Beaumont, 2020;](#)

Saunders, 2015, 2019; McAtamney, 2018; Hankus 2016; Smith, 2014; Sze, 2013; Robertson, 2010) to explore various aspects of the program. At the time of writing nearly 35,000 teachers and their students had been involved in the program.

It has been critical to monitor the sustainability of such professional learning over an extended time frame. A variety of methods have been used to collect data from principals and executive staff, teachers, students and teaching artists to ensure that all participant voices and perspectives are heard. It does appear that this program is an innovative way of developing transformative teacher professional practice as well as deep literacy learning understanding for students. To date, based on information collected and analyzed, it is clear that the co-mentoring model is a highly effective professional learning approach. Both teachers and teaching artists express surprise at how valuable it is for them as well as the participating students. Teachers certainly value the distinction between building expertise and confidence through co-teaching and learning rather than, as in some artist in residence programs, a program that emphasizes *learning from* an expert. Teachers consolidate their learning between and beyond the teaching artists' visits. This has been shown to be at the heart of the creative transformation of teachers' practice. The teaching artists are also surprised by the depth of the new learnings and understandings of using their craft and expertise in this way.

Teachers consistently report improved engagement, empathy and confidence as well as literacy outcomes for their students. They note as well as a growth in their own confidence, understanding and expertise in using drama strategies in their programming across the curriculum. They frequently acknowledge a growth in or a return to their own creative pedagogical processes. The research findings also underline that the participant teachers' excitement with and enthusiasm for being associated with a working theater and professional actors alongside their new-found expertise in their use of drama.

Genuine collaboration between the teacher and teaching artist is both critical for and central to enabling and empowering teacher knowledge, understanding and confidence. Working together, learning from each other and sharing new ideas grows experience and expertise, and professional dialogue nurtures the development of authentic learning communities. With such opportunities, teachers can build courage and capacity to meaningfully embed drama in both their literacy program and across other subjects in the curriculum, and to articulate these philosophies and processes in their school communities.

It is also clear that teachers who have been involved in the *School Drama* program continue to use the drama strategies modeled for them in their subsequent English and literacy programs indicating the sustainability of the processes. Both text (Ewing and Saunders, 2016) and online resources (Ewing and Saunders, *The School Drama Companion*, 2016b) have been developed and the *School Drama Hub* professional learning program has been created to address the demand from teachers who have been involved in the Classic program to continue their professional development and involvement in the learning community. Attending twilight workshops to learn more, trying their learning out in their specific contexts and reporting back in the next workshop enables both a consolidation of skills and understandings as well as renewal. New units of work are jointly written and shared.

Further indications of the success of *School Drama* include its adaptation for use in intensive English classes in secondary schools for EALD students (Beaumont, 2020; McAtamney, 2018); in juvenile justice centers for disengaged students; in adult English classes for refugees (Hogan, 2020; Campbell and Hogan, 2019) and in early childhood contexts (Karaolis, 2020).

Teacher education and arts professional learning during the pandemic

Most recently crises including the pandemic and the increasing number of global crises like earthquake, bushfires and floods exacerbated by climate change have seen a number of very successful arts informed professional learning programs developed to help teachers work with children as they return to school following major traumatic and life threatening events. Arts processes provide powerful strategies to engender hope and restore wellbeing after disaster. *Te Rito Toi*, for example, was founded by Peter O'Connor in New Zealand drawing on his experience after the Christchurch and Mexico earthquakes. It now provides resources and support for teachers all over the world (<https://www.teritotoi.org>). Another example *The Banksia Initiative* <https://www.artshealthnetwork.com.au/about-the-banksia-initiative-recovering-rebuilding-reimagining/> was a cross disciplinary alliance of researchers, educators and artists to support Australian classroom teachers and communities as they recovered from the bushfires.

Core to the success of arts informed teacher professional learning programs are the respect for the teacher's artistry and professionalism. Bentley's (2015) key features of an effective professional learning framework have been reinterpreted by Abbey MacDonald (2018) for arts professional learning and are relevant here. The framework's features include time for consultation, planning and debriefing around shared purpose and vision; the forging of a genuine partnership and the development of collegiality and trust to learn from each other. Most important is the recognition of the intertwining of teacher, artist and student wellbeing, a growth in confidence and, consequently, a growth in identified knowledges, skills and understandings. Such programs also inform teacher (and learner) self-efficacy and the development of positive beliefs about how arts education and education through the Arts can be realized. Mentoring needs to be envisioned for all art forms and funding needs to be provided by the educator sector to augment the funding provided by arts foundations and philanthropists.

Concluding comments

If arts education is to be made accessible to all students from all backgrounds at each level of education, there are direct implications for teacher education and for school curricula. Because many official curriculum documents fail to emphasize arts processes and experiences, many teacher education programs do not provide adequate time and resourcing in arts education. If a creative culture is to be developed through arts education among teachers, principals and other leaders, governments and education systems must

reshape the competing competitive academic discourse that currently works against achieving such a culture. Cultivating imagination and creativity needs to become meaningfully integrated with other espoused priorities. Interdisciplinary experiences, including digital and other emerging art forms, need to be deemed important, not marginalized. More cross-curricular work involving arts and other (non-arts) subject areas, working together on creative and/or cultural themes, such as domestic violence, the refugee crisis, the environment and sustainable development, global citizenship and the need for peace, would initially place new demands on teachers and schools. It would require both intelligent school leadership and support at a policy level as well as adequate funding. The Arts are demonstrably well-placed to facilitate change in educational systems and structures. Communicating the individual and social impact of arts education, to raise community awareness about the intrinsic benefits of the Arts is also imperative.

A considerable and growing body of research supports the assertion that the Arts should be playing an important role in all areas of school and university education and community development. This cannot be actualized until there is adequate teacher education in and through the Arts.

All learners are entitled to access knowledge and skills that will enable their well-being in such a rapidly changing world. Nevertheless the position of the Arts in teacher education is under threat or already diminishing in many countries. With notable exceptions, it appears that in many education systems some art forms are encouraged (usually music and visual arts) while others are seen as optional, including dance and drama. In a crowded curriculum, and with worldwide cuts to education spending, minimal attention and funding is being provided for meaningful education in and through the Arts, limiting children's opportunities and experience in music, drama, creative writing and quality literature. We must research the impact of teacher education programs that do embed quality arts education and processes on both the teachers themselves and their students. If we are to realize the transformative potential of the Arts and education, we must move beyond rhetoric in policy about its importance, to action, by ensuring the provision of adequate funding for career-long teacher professional learning in and through the Arts.

References

- A New Approach, 2020. Transformative: The Impacts of Culture and Creativity. Insight Research Series 2. Australian Academy of the Humanities Council, Sydney. https://www.humanities.org.au/wp-content/uploads/2020/06/ANA-InsightReportTwo-FullReport_VO.1.pdf.
- Alter, F., Hays, T., O'Hara, R., 2009. Creative arts teaching and practice: critical reflections of primary school teachers in Australia. *Int. J. Educ. Arts* 10 (9). Retrieved [date] from <http://www.ijea.org/v10n9/>.
- Arts Council, England, 2019. Time to listen. Evidence from tracking arts learning and engagement (TALE) project. Retrieved from <https://researchtale.files.wordpress.com/2019/03/time-to-listen-background-report.pdf>.
- Bamford, A., 2006. The Wow Factor: Global Research Compendium on the Impact of the Arts in Education. Waxmann Verlag, Berlin, Germany.
- Barton, G., Ewing, R., 2017. Encouraging a dynamic. Relationship between the arts and literacy. In: Barton, G., Baguley, M. (Eds.), *The Palgrave Handbook of Global Arts Education*. Palgrave Macmillan, London, pp. 221–243.
- Beaumont, N., 2020. Unpublished M.Ed. (Research). In: *Drama and Oral Literacy in the Early Primary Years*. University of Sydney, Sydney.
- Bentley, T., 2015. Achieving More Together: Reflecting on Collaboration and Education in Tasmania's Future. Peter Underwood Centre, University of Tasmania, Hobart. Retrieved from http://www.utas.edu.au/_data/assets/pdf_file/0005/1045517/Tom-Bentley-Visiting-Scholar-Report.pdf.
- Biesta, G., 2013. *The Beautiful Risk of Education*. Paradigm Publishers, London.
- Big hART <https://www.bighart.org>.
- Booth, E., 2010/20. The history of teaching artistry. Retrieved from <https://ericbooth.net/the-history-of-teaching-artistry-revised-2020>.
- Booth, E., 2010. *The Music Teaching Artist's Bible: Becoming a Virtuoso Educator*. Oxford University Press, Inc, New York.
- Boyd, B., 2009. *On the Origin of Stories*. Harvard University Press.
- Bresler, L., 2004. *Knowing Bodies, Moving Minds: Towards Embodied Teaching and Learning*. Kluwer, Dordrecht, Netherlands.
- Bryce, J., Mendelovits, J., Beavis, A., McQueen, J., Adams, I., 2004. *Evaluation of School-Based Arts Education Programmes in Australian Schools*. Australian Council of Educational Research, Melbourne.
- Burnard, P., Colucci-Gray, L., 2020. *Why Science and Art Creativities Matter*. Brill, Netherlands.
- Campbell, V., Hogan, Z., 2019. Pandora and the tiger's whisker: stories as a pretext in two adult language learning contexts. *NJ: Drama Australia* 42 (1).
- Catterall, J., 2013. Getting real about the E in STEAM. *STEAM J.* 1 (1). <https://doi.org/10.5642/steam.201301.06>. <http://scholarship.claremont.edu/steam/vol1/iss1/6>.
- Catterall, J., 2009. *Doing Well and Doing Good by Doing Art: The Long Term Effects of Sustained Involvement in the Visual and Performing Arts during High School*. Los Angeles Imagination Group, Los Angeles, CA.
- Cole, C., Thomson, P., 2021. Making the most of school arts education partnerships. *Curric. Perspect.* 41 (1), 101–106.
- Connell, R., 1994. Poverty and education. *Harv. Educ. Rev.* 64 (2), 125–147.
- Council of European Union, 2007. *European Agenda for Culture*. Accessed at <http://eur-lex.europa.eu/legal-content/GA/ALL/?uri=celex:52007DC0242>.
- Damasio, A., 2004. *Descartes Error. Emotion, Reason and the Human Brain*. Avon Books, New York.
- Darling-Hammond, L., 2010. *The Flat World and Education: How America's Commitment to Equity Will Determine Our Future*. Teachers College Press, New York.
- Deasy, R., 2002. *Critical Links. Learning in the Arts and Student Academic and Social Development*. Arts Education Partnership, Washington DC.
- Dinham, J., 2019. *Delivering Authentic Arts Education*, fourth ed. Cengage, Melbourne.
- Donaldson, G., 2015. *Successful Futures. Independent Review of Curriculum and Assessment Arrangements in Wales*. <http://gov.wales/docs/dcells/publications/150225-successful-futures-en.pdf>.
- Donelan, K., 2012. Arts education as intercultural and social dialogue. In: Sinclair, C., Jeanneret, N., O'Toole, J. (Eds.), *Education in the Arts*, second ed. Oxford University Press, Melbourne.
- Dunn, J., Jones, A., Zink, L., 2020. Beyond Y connect – sustaining artist-teacher partnerships. University of Sydney, CREATE Centre Partnerships Webinar, 8 November. <https://yconnectproject.com/2020/11/08/y-connect-partnerships-webinar-beyond-y-connect-sustaining-artist-teacher-partnerships/>.
- Dunn, J., Jones, L., Buncy, P., Stinson, M., Lazaroo, N., 2019. *Creating Critical Connections Through the Arts: The Y Connect Key Findings and Significance Report. Examining the Impact of Arts-Based Pedagogies and Artist/teacher Partnerships in One Australian Secondary School*.
- Eisner, E., 2002. What can education learn from the arts about the practice of education? John Dewey Lecture, Stanford University *J. Curric. Superv.* 18 (1), 4–16. Retrieved from <http://artstart2011.pbworks.com/w/file/40364564/Eisner-ed%20learn%20from%20arts.pdf>.
- Ewing, R., 2019. *Arts-rich Pedagogy and Becoming Deeply Literate*. Drama Australia, Brisbane.

- Ewing, R., 2010. The Arts and Australian Education: Realising Potential. Australian Education Review 58. Australian Council for Educational Research, Melbourne, Australia.
- Ewing, R., 2006. Reading to allow spaces to play. In: Ewing, R. (Ed.), Beyond the Reading Wars: Towards a Balanced Approach to Helping Children Learn to Read. Primary English Teaching Association Australia, Camdenville, Sydney, pp. 141–152.
- Ewing, R., 2002. Framing a professional learning community: an Australian case study. Curr. Perspect. 22 (3), 23–32.
- Ewing, R., Gibson, R., Campbell, V., Saunders, J., Hristofski, H., 2015. School drama: towards state of the art in drama professional learning? In: Anderson, M., Roche, C. (Eds.), The State of the Art: Teaching Drama in the 21st Century. Sydney University Press, Sydney, pp. 25–48.
- Ewing, R., Gibson, R., 2007/2015. Creative teaching or teaching creatively? Using creative arts strategies in preservice teacher education. Waikato J. Educ. 13, 159–178.
- Ewing, R., Saunders, J.N., 2017. Towards 'grown-up ness in the world' through drama as critical, quality pedagogy. In: Biesta, G., Naughton, C., Cole, D. (Eds.), The Arts, Artists and Pedagogy (Chapter 10). Routledge, London.
- Ewing, R., Saunders, J.N., 2016a. The School Drama Book: Drama, Literature & Literacy in the Creative Classroom. Currency Press, Sydney.
- Ewing, R., Saunders, J.N., 2016b. The school drama companion. <https://www.sydneytheatre.com.au/education/teacher-learning/the-school-drama-companion>.
- Ewing, R., Simons, J., Hertzberg, M., Campbell, V., 2015. Beyond the Script. Take 3. Drama in the Classroom, third ed. Primary English Teaching Association Australia, Sydney.
- Finnish Ministry of Education and Culture, 2010. Arts Education and Cultural Education in Finland. Policy Analysis Report of the Ministry of Education in Finland. <http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2010/liitteet/okmpol022010.pdf?lang=en>.
- Fleming, J., Gibson, R., Anderson, M., 2016. How Arts Education Makes a Difference. Research Examining Successful Classroom Practice and Pedagogy. Routledge, Abingdon, Oxon.
- Freeman, W., 1995. Societies of Brains: A Study in the Neuroscience of Love and Hate. Lawrence Erlbaum, New York.
- Gadsden, V., 2008. The arts and education: knowledge generation, pedagogy and the discourse of learning. Rev. Res. Educ. 32, 29–61.
- Garvis, S., Riek, R., 2010. Improving generalist teacher education in the arts. Int. J. Art Soc. 5 (3), 159–169.
- Gibson, R., Ewing, R., 2020. Transforming the Curriculum through the Arts, second ed. Palgrave Macmillan, London.
- Gibson, R., Smith, D., 2013. School Drama: A Meta-Analysis. Unpublished report. University of Sydney, Sydney.
- Greene, M., 1995. Releasing the Imagination: Essays on Education, the Arts and Social Change. Jossey Bass, San Francisco.
- Greene, M., 1978. Landscapes of Learning. Teachers College Press, New York, NY.
- Habermas, J., 1972. Knowledge and Human Interests (J. Shapiro, trans.). Heinemann, London.
- Hankus, N., 2016. It Made Me Feel like I Lost Something: Engaging Students in Learning through Process Drama (Unpublished honours dissertation). University of Sydney, Sydney, Australia.
- Harris, A., Carter, M., 2021. Applied creativity and the arts. Curr. Perspect. 41, 107–112.
- Hatcher, R., 2011. Professional learning for creative teaching and learning. In: Sefton-Green, J., Thomson, P., Jones, K., Bresler, L. (Eds.), The Routledge International Handbook of Creative Learning. Routledge, London and New York, pp. 404–413.
- Heath, S., 2000. Seeing our way into learning. Camb. J. Educ. 30 (1), 121–131.
- Hobbs, L., 2012. Examining the aesthetic dimensions of teaching: relationships between teacher knowledge, identity and passion. Teach. Teach. Educ. 28 (5), 218–227.
- Hogan, Z., 2020. Connected: Personal, Social and Creative Experiences of English Language Learning with Adults from Refugee Backgrounds. University of Sydney (PhD Proposal).
- Hunter, M., 2011. Australia Council Community Partnerships Artist in Residence Program Evaluation Report. Australian Council for the Arts.
- Hunter, M., 2005. Education and the Arts: Research Overview. A Summary Report Prepared for the Australia Council for the Arts. Australia Council for the Arts. Australian Government and Australian Council for the Arts, Sydney. Canberra.
- Jefferson, M., Anderson, M., 2017. Transforming schools. Bloomsbury, London.
- Karalis, O., 2020. Everybody in! Drama as a Pedagogy for Inclusion (Unpublished doctoral thesis). The University of Sydney, Australia.
- Lemon, N., Garvis, S., 2013. What is the role of the arts in a primary school? An investigation of pre-service teachers in Australia. Aust. J. Teach. Edu. 38 (9), 1–9.
- Li, H.J., 2013. 2012 Annual report of Chinese arts education – preschool. Art Critic. 5, 64–66.
- Lincoln Center <http://lincolncenter.org/lincoln-center-at-home>.
- McCarthy, K., Ondaatje, E., Zakaras, L., Brooks, A., 2004. Gifts of the Muse: Reframing the Debate About the Benefits of the Arts. RAND, Santa Monica, CA.
- Macdonald, A., 2018. Collaboration: a delicate dance, light feel and a gentle touch. In: Paper Presented at ACERN. University of Sydney, Sydney.
- McAtamney, O., 2018. School Drama: A Case Study of Using Drama for Oracy in the EAL/D Classroom (Unpublished Hons dissertation). The University of Sydney, Australia, Sydney.
- Martin, A.J., Mansour, M., Anderson, M., Gibson, R., Leim, G.A.D., 2013. The role of arts participation in students' academic and nonacademic outcomes: a longitudinal study of school, home, and community factors. J. Educ. Psychol. 105 (3), 709–727.
- Miller, C., Saxton, J., 2016. Into the Story 2: More Stories, More Drama. Intellect, Bristol, UK.
- Miller, C., Saxton, J., 2004. Into the Story: Language in Action through Drama. Heinemann, Portsmouth, NH.
- National Advocates for the Arts in Education, 2019. More than Words Can Say: A View of Literacy through the Arts. National Affiliation of Arts Educators, Canberra.
- National Education Association, 2013. An educator's guide to the four C's. Preparing 21st century students for a global society. Retrieved from. www.nea.org/tools/52217.htm.
- Nilson, C., Fetherston, C., McMurray, A., Fetherston, T., 2013. Creative arts: an essential element in the teacher's toolkit when developing critical thinking in children. Aust. J. Teach. Edu. 38 (7), 1–17.
- Nussbaum, M., 2010. Not for Profit: Why Democracy Needs the Humanities. Princeton University Press, Princeton.
- O'Connor, P., Anderson, M., Freebody, K., Ginns, P., 2020. Replanting Creativity during Post-normal Times. Centre for Arts and Social Transformation, Auckland. https://www.teritoti.org/wp-content/uploads/2020/10/Replanting-Creativity-during-post-normal-times_FINAL-opt_2020.pdf.
- O'Toole, J., Dunn, J., 2020. Stand up for Literature. Currency Press, Sydney.
- Pascoe, R., 2015. The transformative change in teacher education: AIR at Murdoch. In: Paper Presented at Australian Association in Education Annual Conference, Perth, 30 November–3 December.
- President's Committee on the Arts and the Humanities, 2011. Reinvesting in Arts Education: Winning America's Future Through Creative Schools. Author, Washington, DC.
- Robertson, A., 2010. The School Drama Experience: A Case Study of Learning in and Through the Art of Drama (Unpublished Hons dissertation). The University of Sydney, Australia, Sydney.
- Sardar, Z., 2010. Welcome to postnormal times. Futures 42 (5). June 2010. Retrieved from. <http://ziaddinsardar.com/2011/03/welcome-to-postnormal-times/>.
- Saunders, J.N., 2019. Dramatic Interventions: A Multi-Site Case Study Analysis of Student Outcomes in the School Drama Program (Unpublished PhD thesis). University of Sydney, Sydney.
- Saunders, J.N., 2015. Unpublished M.Ed (Research) dissertation. In: Academic and Non-academic Outcomes for Students Engaged in School Drama: A Case Study. University of Sydney, Sydney.
- Seidel, S., Tishman, S., Winner, E., Hetland, L., Palmer, P., 2009. The Qualities of Quality. Understanding Excellence in Arts Education. Project Zero, Harvard Graduate School of Education and The Wallace Foundation, Cambridge, Mass.
- Smith, D., 2014. School Drama Program Sustainability Case Study. The University of Sydney, Sydney.
- Smith, D., 2013. An Independent Report for the Welsh Government into Arts Education in Welsh Schools. Accessed at. <http://gov.wales/docs/dcells/publications/130920-arts-in-education-en.pdf>.
- Silverstein, L., Duma, A., Layne, S., 2010. Arts Integration Schools: what, Why and How? Changing Education Through the Arts. The Kennedy Center, Washington.
- Sze, E., 2013. Sustainable Professional Development: A Case Study on Quality Arts Partnerships in the Primary Classroom (Unpublished Hons dissertation). The University of Sydney, Sydney.

- Tau, X., 2019. 'I want the whole world to be musical'. The role of the Arts in preschool education. Music curriculum in Australia v. China. Unpublished PhD, University of Sydney.
- Thomson, P., Sefton-Green, J. (Eds.), 2011. *Researching Creative Learning: Methods and Issues*. Routledge, London.
- University of Warwick, 2015. *Enriching Britain: Culture, Creativity and Growth*. http://www2.warwick.ac.uk/research/warwickcommission/futureculture/finalreport/warwick_commission_report_2015.pdf.
- UNESCO, 2006. Road map for arts education. Retrieved March 26, 2017 from. http://www.unesco.org/fileadmin/multimedia/HQ/CLT/CLT/pdf/Arts_Edu_RoadMap_en.pdf.
- United States Department of Education, 2001. No Child Left Behind. Elementary and Secondary Education Act. <https://www2.ed.gov/hclb/landing.jhtml>.
- United States National Centre for Education, 2014. The Condition of Education 2014. US Department of Education, Washington. <https://nces.ed.gov/pubs2014/2014083.pdf>.
- Uptitis, R., Atri, A., Keely, C., Lewis, A., 2010. Teachers' experiences of professional development programs in the arts: generalist teachers and arts advocates. *Univ. Melb. E-J.* 1 (5), 1–27.
- Vermeulen, D., 2009. *Integrating Music in an Integrated Arts Curriculum for South African Primary Schools*. Unpublished PhD dissertation. University of Pretoria, Pretoria.
- Vinson, T., Rawsthorne, M., 2015. Dropping off the Edge 2015. Persistent Communal Disadvantage in Australia. Jesuit Social Services/Catholic Social Services Australia, Curtin, ACT, Australia, pp. 6–148.
- Welsh Government, 2015. Creative learning through the arts – an action plan for Wales 2015–2020. Accessed at. http://www.arts.wales/c_engagement-and-participation/action-plan-creative-learning-through-the-arts.
- Winner, E., Goldstein, T., Vincent-Lancrin, S., 2013. Arts for Art's Sake? the Impact of Arts Education. OECD, Paris. Accessed at. www.oecd.org/e_du/cei/arts.htm.
- World Health Organisation, 2019. What Is the Evidence on the Role of the Arts in Improving Health and Well-Being? A Scoping Review. <https://www.euro.who.int/en/publications/abstracts/what-is-the-evidence-on-the-role-of-the-arts-in-improving-health-and-well-being-a-scoping-review-2019>.
- Wyn, J., 2009. Touching the Future: Building Skills for Life and Work. *Australian Education Review*, 55. Australian Council for Educational Research Press, Melbourne.
- Zhang, J., 2012. Contemporary preschool music education in China. *Science and Technology Information* 33, 196.
- Zusak, M., 2014. The Failurist. TEDx, Sydney, Australia. Retrieved March 26, 2017 from. https://www.youtube.com/watch?v=A-_8Qldm4hA.

Websites

- Kennedy Center Changing Education through the Arts (CETA) <https://www.kennedy-center.org/education/networks-conferences-and-research/networks-and-strategic-leadership/ceta-schools>.
- Te Rito Toi <https://www.teritotoi.org>.
- The Banksia Initiative <https://www.artshealthnetwork.com.au/about-the-banksia-initiative-recovering-rebuilding-reimagining/>.
- The Maxine Greene Institute <https://maxinegreene.org>.
- The School Drama Program <https://www.sydneytheatre.com.au/schooldrama>.

Emotion and reconciliation pedagogies

Michalinos Zembylas, Open University of Cyprus, Nelson Mandela University, Gqeberha, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	315
Emotion and reconciliation	315
Reconciliation pedagogies in posttraumatic contexts	317
Reconciliation pedagogies: three pedagogical principles	318
The significance of pedagogic discomfort	319
The pedagogical principle of mutual vulnerability	319
The value of compassion and strategic empathy	320
Conclusion	321
References	321

Introduction

Educators are increasingly faced with what [Worsham \(2006\)](#) has described as “posttraumatic” cultural moments, namely, moments that are marked by “unprecedented historical trauma” (*ibid.* p. 170)—such as wars, conflicts, and genocides around the world. Traumatic events and posttraumatic experiences raise tremendous challenges for educators because they often create situations that are conducive to adopting reductive binary positions of we-and-they ([Zembylas, 2008, 2015a](#)). However, at the same time, traumatic events and posttraumatic experiences can also offer pedagogical opportunities for a critical exploration of the emotional complexities involved in dealing with the aftermath of these events. This exploration can be the point of departure for developing more productive ways of engaging students with posttraumatic experiences to support the long and painful process of reconciliation ([Bekerman and Zembylas, 2012; Zembylas et al., 2016](#)).

Although the chicken-and egg-problem is real and has consequences—that is, a level of sociopolitical reconciliation may need to be reached before educators touch on the sensitive issues of trauma in the classroom—curriculum and pedagogy can introduce students to some aspects of the difficult process of reconciliation ([Cole, 2007](#)). This is interpreted here to mean that educators and students engage with narratives of trauma and suffering in pedagogically sensible ways that might contribute to the broader process of reconciliation, yet without superficializing or sentimentalizing either the process of reconciliation or pedagogical efforts related to this ([Zembylas, 2013a, 2016](#)). The term “reconciliation pedagogies” has been suggested in reference to the pedagogical practices that encourage the exploration of how we might live together in societies of ever-increasing complexities and conflicts, focusing attention on aspects “such as dialog, the ‘discourse of possibility’, remembering and witnessing, and the affective dimensions of difficult, contested knowledge” ([Hattam et al., 2012](#), p. 6).

Reconciliation pedagogies, then, constitute intervention activities that are planned and organized at the everyday level of life in schools and other educational settings (formal or informal), aiming to challenge the reductive binaries that result from conflict situations and establish a *reconciliatory ethos*. A reconciliatory ethos in curriculum and teaching helps students become aware of the shared meanings, emotional complexities, and ethical interdependence that can constrain as well as expand mutual understanding and interaction with Others who may be perceived as ‘enemies’. Reconciliation pedagogies encourage students to make choices that facilitate their struggles to deal with the emotional complexities of the process of reconciliation in the aftermath of posttraumatic experiences, yet without ‘imposing’ the value of reconciliation.

The purpose of this chapter is to discuss the role of emotion in reconciliation and analyze how reconciliation pedagogies may be designed and enacted, taking into consideration the emotional complexities involved in reconciliation processes. The chapter is divided into three parts. The first part discusses the entanglement between emotion and reconciliation to show how the former is an inextricable part of the latter; as it is shown, this entanglement has important consequences for pedagogical practices. The second part of the chapter takes on reconciliation pedagogies in posttraumatic contexts and explores their potential as well as their challenges and risks. The final part outlines three pedagogical principles of reconciliation pedagogies emerging from my long-time research in this area; their discussion is meant as illustrative of the pedagogical openings that reconciliation pedagogies create to address the complex emotional issues emerging in posttraumatic contexts. At the same time, it is suggested that the limitations and risks of using reconciliation pedagogies in posttraumatic contexts need to be kept in mind.

Emotion and reconciliation

In many conflict-affected societies, one of the emerging social and political movements is that of *reconciliation* ([Auerbach, 2009](#)). Reconciliation as a term and a process is a dynamic, complex and long-term process that does not have a fixed definition across different settings ([Cole, 2007](#)). In general, reconciliation is understood as a sociopolitical and emotional process in which

antagonistic parties “acknowledge the reality of the terrible acts that were perpetrated; accept with compassion those who committed injurious conduct, as well as acknowledging each other’s sufferings; believe that their injustices are being redressed and anticipate mutual security and well-being” (Kriesberg, 1998, pp. 351–352). Proponents of this movement stress the role of reconciliation as an alternative to revenge and spirals of violence, on the one hand, while accepting that for this to happen some degree of putting aside the past may be needed, on the other (Flam, 2013).

Traumatic events, such as wars, conflicts, genocides, and terrorist attacks, generate powerful emotions—most notably, fear, grief, anger, shame, resentment, and hatred (Hutchison and Bleiker, 2008). These emotions fuel people’s perceptions and actions and make conflicts, wars, and genocides imaginable as well as the prospects of reconciliation more difficult (Des Forges, 1999). Emotions, then, are important components of historical consciousness in conflict-affected societies, and they strongly shape collective narratives, identities, and memories. In fact, some scholars (e.g., Retzinger and Scheff, 2000) argue that it is often these powerful emotions, rather than material interests, that pose the most serious obstacles to peacebuilding, healing, and reconciliation efforts. A thorough understanding of the powerful but often neglected role of emotions, write Hutchison and Bleiker (2008), is essential to increase the possibilities of establishing a culture of healing and reconciliation. Rather than assuming though beforehand that reconciliation should necessarily imply the management of certain feelings (e.g., fear, anger and resentment), it is crucial to explore *how* such feelings may become a productive part or an obstacle of the reconciliation process (Zembylas, 2013b). This argument suggests that if conflict-affected societies are to come together and engage in reconciliation, there have to be social, political and pedagogical spaces that allow individuals and groups to share their traumas, emotions and memories rather than closing down such possibilities (Bekerman and Zembylas, 2012; Zembylas, 2008, 2015a).

Hutchison and Bleiker (2008) write that there are two key issues in considering the linkages between emotion and reconciliation in conflict-affected societies. The first involves the question of what kinds of emotions may facilitate the collective reckoning with trauma and the past; the second entails how mediators (e.g., educators) might cultivate these emotions. With regard to the first issue, it is suggested that emotions such as empathy and compassion may help generate the space conducive to the acknowledgment of victimhood and the rehumanization of the Other (Halpern and Weinstein, 2004; Schaap, 2006). In particular, Halpern and Weinstein (2004) suggest that there are three important qualities in empathy that makes it particularly important for reconciliation. First, it entails seeking the individual perspective of another; this can serve as an important motivation and encouragement to place one’s self in another’s circumstances and accepts that he or she possesses the same rights. Second, empathy involves a genuine effort to get to know the Other and his or her perspectives. In contrast, polarized trauma narratives close one’s mind toward the Other’s experiences. Third, empathy involves emotional as well as cognitive openness, and the toleration of ambivalence; ambivalence entails viewing differences as an enriching part of creating an ongoing workable relationship. For these reasons, teaching and learning how to empathize with the Other’s experiences and narratives is probably one of the most challenging aspirations for education (Zembylas et al., 2016).

The second issue, namely, how mediators (e.g., educators) might cultivate emotions such as empathy and compassion, prompts broader questions concerning the task of educators to create spaces in which such emotions are worked through without establishing emotional regimes that either ‘impose’ reconciliation or undermine it (Zembylas, 2013b, 2018a,b). At stake here is not the suspension of emotions such as fear, anger and resentment in the name of the (noble) cause of reconciliation, but rather, and crucially, that the *emotional regimes* established in the process do not moralize or sentimentalize the reconciliation process in ways that end up having the opposite results. To make this argument clearer, I will recount more precisely what I mean by “emotional regimes” and what this analysis shows concerning the ways in which the institutionalization of reconciliation processes molds particular emotional regimes (e.g., in schools and classrooms).

Scholars in the history of emotion have worked on different concepts to describe and theorize how emotional regimes are established and sustained. Reddy (2001) defines ‘emotional regime’ as the set of normative emotions and the official rituals, practices, and expressions that underpin any political regime—at the micro- or macro-level. *Emotionology* is another term developed in the field of history of emotions (see Stearns and Stearns, 1985) that refers to the emotional rules constructed within a particular culture—i.e., the rules or norms by which people are supposed to shape their emotional expressions and respond to the expressions of others (Plamper, 2010). Finally, Rosenwein (2002) uses the term *emotional communities* to denote how emotional expressions and feeling rules are interwoven: “what these communities (and the individuals within them) define and assess as valuable or harmful to them; the evaluations that they make about others’ emotions; the nature of the affective bonds between people that they recognize; and the modes of emotional expression that they expect, encourage, tolerate, and deplore” (p. 842). All of the above terms—namely, emotional regime, emotionology, and emotional community—highlight the fundamental idea that *feeling rules* (Hochschild, 1979) are inevitable in any emotional expressions within a community and that these rules are expressions of power relations.

Flam (2013) suggests that the moral feeling rule produced for dealing with the past is to place the burden on the victim from whom more are demanded in terms of courage, endurance and emotions management than from the perpetrators. Flam writes about the feeling rules formulated for the perpetrators and the victims:

The obligations and feeling rules formulated for the *perpetrators* make a short list. They are to tell the truth, and to admit guilt rather than forget. [...] In an effort to convince them to reveal the truth about their deeds, it is argued that even if it is painful for them to remember or they fear the consequences of admitting their guilt, they should testify for the sake of truth, justice and reconciliation, and also because telling their story will have a healing, therapeutic effect. (2013, pp. 373–374, original emphasis)

On the other hand, the list of obligations and feeling rules for the victims is much longer, argues Flam:

They are told that they should forego bitterness and hatred, and give up revenge since revenge does not restore but only leads to escalation, excesses and mutual destruction. It destroys one's own life and upsets the society. The victims should not express rage. Instead, they should recognize the legitimate limits of feeling over the wrongdoing they experienced. (2013, p. 374)

Most importantly, the victims are caught in a double bind: on the one hand, they are told they should testify for the sake of reconciliation, even if it is painful for them and they might face consequences; even if they would rather forget, they are told they should remember, tell and heal. On the other hand, the victims are told they should leave their trauma behind and are told not to degrade or deny the dignity of the wrongdoer. The burden of reconciliation, then, is placed on the victims who, although they have the right to rage, are told that they can only achieve empowerment when they forgive the perpetrators (Flam, 2013).

Overall, according to Flam (2013), there is what she calls a *therapeutic discourse* of reconciliation imposing many more feeling rules on the victims than on the perpetrators, telling the victims how to think and reframe the situation, what to remember and tell, what to heal and forget. The victims are told that they should engage in emotions management, and get rid of their destructive emotions such as rage, hatred and resentment, for the sake of reconciliation. This therapeutic discourse calls for sensitivity and for a giving a chance to the perpetrators, while telling victims and perpetrators alike they can achieve reconciliation when they remember, tell and forgive. This joint therapeutic-reconciliation perspective establishes an emotional regime which imposes certain emotions on the victims and discourages other emotions in the name of reconciliation. The imposition of this therapeutic-reconciliation perspective on the victims has caused much frustration and disappointment among those who feel that the burden of reconciliation should not be placed on the victims (Chouliaraki, 2013; Cohen, 2001).

Arguably, then, emotional regimes are expressions of power relations because they prescribe specific feeling rules which are indicative of who holds power in a given community. This idea was originally formulated by Kemper (1978) who pointed out that power holders define cultural rules, including these pertaining to emotional expression (Flam, 2013). By recounting how emotional regimes impose certain feeling rules on the victims, one can begin to engage in a deeper theorization of events such as secondary victimization in which these feeling rules might lead to further emotional suffering for the victims. The concept of emotional regimes, therefore, can become a valuable conceptual tool in efforts to engage in critical analysis of the history and politics of emotions in conflict-affected societies.

In light of the above analysis, the next part of the chapter makes an attempt to respond to the following questions: How can educators acknowledge the emotional complexities of *difficult* events such as that of reconciliation and, specifically, how can they contribute to reconciliation through their pedagogies? Furthermore, what emotional regimes might be invoked or molded in classrooms through reconciliation pedagogies, and what are the challenges and risks involved?

Reconciliation pedagogies in posttraumatic contexts

The purpose of this part of the chapter is to show: first, the pedagogical consequences of acknowledging the emotional complexities of *difficult knowledge* (Britzman, 2000, 2013; Simon, 2011a, 2011b), that is, knowledge which is difficult to teach/learn (e.g., reconciliation) because it entails serious affective, epistemological and other challenges; and second, to argue that it is precisely this acknowledgment that enriches the radical potentiality of reconciliation pedagogies to inspire transformative emotional practices in schools and classrooms (Zembylas, 2014). Before doing so though, I want to introduce another conceptual tool that is going to be helpful in further understanding the power and tenacity of difficult knowledge in posttraumatic situations. This is the notion of difficult knowledge as a “technology of affect” (Hook, 2005); that is, how difficult knowledge implies certain mechanisms that work through the body, conduct and being to establish particular ideological and affective investments that are difficult to challenge (Zembylas, 2015c).

First of all, it is important to clarify the meaning of the term *technology*. Technology is a key term in Foucault's later work and refers to any assemblage of knowledges, practices, techniques and discourses, used by human beings on others or on themselves to achieve particular ends (Foucault, 1977; Hook, 2007; Rose, 1998). *Technology of affect*, in Hook's (2007) terminology, refers to the mechanisms through which emotions come to be instrumentalized, containing certain social norms and dynamics of inclusion/exclusion with respect to one's self and the Other. Given that emotion is never simply individual or internal (Ahmed, 2004), difficult knowledge can be understood as a form of affective technology, an affective mode of self-practice situated in a circuit of social and political meaning.

For example, consider the context of post-apartheid South Africa. The difficult knowledge of the traumatized White or Black student—who is traumatized for different reasons and in different ways (Jansen, 2009)—is manifested affectively on the basis of norms that appeal to feelings of whiteness or blackness. White and Black students carry the burden of this difficult knowledge with considerable consequences for pedagogical purposes. White students may feel resentment for being vilified for their parents' history; Black students may feel resentment when they realize that very little has changed since the apartheid era. The complexity and diversity of

these feelings have to be acknowledged by everyone involved in this process, especially in terms of the forms of exclusion and inclusion they create. As Hook writes:

It is important to acknowledge here that we may ready our affective postures in how we position ourselves in given social encounters. We may as such assume certain affect-positions (fear, anger, irritation, love) which then become the proof of affect for a given ideological proposition, for a categorical relationship of entitlement, exclusion, belonging, etc. So: that I feel threatened by an influx of immigrants is proof enough of their moral dubiousness, proof enough also of why they—and others like them—should be prevented any rights of access. (2005, p. 93)

Hook's analysis allows for the theorization of difficult knowledge as an affective technology that maintains certain social norms and structures of feeling. For example, feelings of resentment for the Other (e.g., immigrants, asylum seekers, refugees) within a country which accepts a growing number of immigrants, work as an ideological and affective position of exclusion for those who are considered *different* and inclusion for those deemed to be *similar* to the self (Zembylas, 2010). An understanding of difficult knowledge and its various complexities and nuances enables a critical consideration of how difficult knowledge is emotionally produced and maintained as forms of inclusion and exclusion. For educators to disrupt resentment, abjection and other emotions about the Other, they must understand difficult knowledge as affectively, socially and politically produced (Zembylas, 2016).

This understanding implies that a conception of reconciliation pedagogies that is merely grounded in critical pedagogy ideas—that is, negative dialectics and counter-affirmation that put an emphasis on universalizing tenets about emancipatory forces and oppressive processes, ideologies and identities—has little value in posttraumatic societies (Jansen, 2009; Zembylas, 2013c, 2018b). Needless to say, this is not to deny the systemic and institutionalized character of oppression, identity and social injustice; rather, it is to highlight that classrooms are not homogeneous environments with a common understanding of oppression, but deeply divided places where contested narratives are steeped in the politics of emotions to create complex emotional and intellectual challenges for teachers. Teachers can be themselves carriers of difficult or *troubled knowledge*, as Jansen (2009) calls it, and this has serious implications for the form that reconciliation pedagogies take in posttraumatic societies.

Consequently, a premise of reconciliation pedagogies in posttraumatic contexts should not be simply to question the dominant educational arrangements (curricula, textbooks); it should also be “the people there, the bodies in the classroom, who carry knowledge within themselves that must be engaged, interrupted, and transformed” (Jansen, 2009, p. 258). These bodies and their difficult knowledge constitute the starting point for reconciliation pedagogies in this context. Thus, taking sides early on based on justice and democracy ideals, maintains Jansen, may not be such a wise move, when there is clash of embodied knowledges and memories. For instance, a teacher who takes sides early on in a posttraumatic context in which two conflicting communities have been both *perpetrators* and *victims* may end up intensifying unsafe classroom spaces. It has long been acknowledged in critical pedagogy rhetoric that classroom is not a safe space and that some teachers and groups of students may be (unintentionally) burdened, for example, through their minority status as representatives of their group—e.g., literature that highlights the ways in which students of color in the United States are burdened with the responsibility of teaching White students about diversity (see Boler, 2004).

It is my contention here that reconciliation pedagogies in posttraumatic situations need to move a step further and be attentive and cautious of the increased complexities that difficult knowledge adds to (already) unsafe spaces in the classroom. Hence, I want to emphasize the role of reconciliation pedagogies in constructing (more) *safe* classroom spaces within which the wounded of all sides can engage in critical and productive dialog (Zembylas, 2015b). By choosing to adopt a reconciliatory ethos in a posttraumatic context, the teacher is committed to interrogate how received knowledges and re-constructed emotions underlie structures of feeling in the teaching of various narratives, without adopting a self-righteous or a therapeutic perspective.

In other words, the emotional labor that is demanded for reconciliation pedagogies in posttraumatic contexts is not intended as a self-justified approach or a therapeutic intervention that seek to validate some selected emotions or dismiss and pacify some others in the name of healing past wounds (see also Amsler, 2011). The “healing of past wounds” approach, as Mohanty (2003) and Oliver (2001) aptly point out, is problematic because it hides or dismisses the relationship between trauma and power relations. Also, other critics such as Ecclestone and Hayes (2009) emphasize that therapeutic approaches, introduced under the banners of preoccupation with students' well-being, deliberately decrease the possibility that students will experience complexity, multiplicity or discomfort in their studies; preoccupation with emotional fragility offers a deeply diminished view of the human subject and undermines the ways in which these feelings contribute unwittingly to the perpetuation of power inequalities and hierarchies.

The critical interrogation of difficult knowledge in posttraumatic contexts marks a valuable intervention in the broad domain of reconciliation pedagogies precisely by focusing on identifying and challenging the affective technologies and emotion-informed ideologies that underlie possible responses toward difficult knowledge—by students and teachers alike—and seeking to make a concrete difference in the lives of those who still suffer from carrying the burden of this knowledge. The process of dissolving categorizations of *us* and *they* in posttraumatic settings is a matter of observing very carefully the consequences of the underlying ideological and affective investments carried in the classroom. These investments need to be engaged and interrupted in sensitive, critical and strategic ways.

Reconciliation pedagogies: three pedagogical principles

In the last part of this chapter I discuss some ideas and examples that have emerged in my long-standing research in this area. In particular, I will discuss three pedagogical principles that may be used to embrace the idea of acknowledging and working with the

emotional complexities of difficult knowledge in the context of reconciliation pedagogies. These principles are: (1) the significance of pedagogic discomfort; (2) the pedagogical principle of mutual vulnerability; and, (3) the value of compassion and strategic empathy. Each of these aspects—along with their strengths and limitations—are discussed below in terms of how they help navigate difficult knowledge in sensitive and productive, yet not therapeutic or sentimental, ways.

The significance of pedagogic discomfort

In recent years, there is increasing research that acknowledges how challenging students and teachers beyond their *comfort zones* and pushing them to deconstruct the ways in which they have learned to see, feel and act constitutes a valuable pedagogic approach in social justice, citizenship and *anti-racist* education (e.g., see Boler, 1999, 2004; Boler and Zembylas, 2003; Zembylas, 2015b, 2017, 2018a; Zembylas and McGlynn, 2012; Zembylas and Papamichael, 2017). In fact, it has been argued that if a major purpose of teaching is to unsettle taken for granted views and emotions, then a *pedagogy of discomfort* is not only unavoidable but may also be necessary (Berlak, 2004). “A pedagogy of discomfort begins,” explains Boler, “by inviting educators and students to engage in critical inquiry regarding values and cherished beliefs, and to examine constructed self-images in relation to how one has learned to perceive others” (1999, pp. 176–177).

For example, individuals who belong to the hegemonic culture experience discomfort when having to confront their privileges in relation to educational and social inequities (see e.g., Leibowitz et al., 2010). Leibowitz and her colleagues, who write in the context of post-apartheid South Africa, demonstrate how a pedagogy of discomfort is valuable in uncovering and questioning the deeply embedded emotional dimensions that shape some individual and group privileges (e.g., those of White students) through daily habits and routines. By closely problematizing these emotional habits and routines and their attachments to whiteness and structural injustices, it is shown that teachers and students in a posttraumatic context can begin to identify the invisible ways in which they comply with dominant ideologies.

Additionally, Zembylas and McGlynn (2012) discuss the emotional tensions, ethical dilemmas and transformative possibilities of using pedagogy of discomfort in a post-conflict society (Belfast, N. Ireland). This work emphasizes that discomforting pedagogies require considerable vulnerability and thus the ethical responsibility of the teacher becomes a complex issue (see also, Zembylas, 2015b). Similarly, Zembylas et al. (2016) study how teachers’ discomforting emotions are manifest in a teacher education setting and how teacher educators might engage with discomfort pedagogically. These observations about the ethical, political and pedagogical challenges reiterate once again the complexities of handling difficult emotional knowledge in posttraumatic societies (Zembylas, 2017, 2018a,b).

When examined through the lens of a posttraumatic context, then, it becomes clearer how and why certain features of pedagogic discomfort concern students and teachers beyond divisions of the world into rival camps (e.g., *oppressors* and *victims*). Students and teachers come into the classroom carrying their difficult knowledge about “conquest and humiliation, struggle and survival, suffering and resilience, poverty and recovery, black and white” (Jansen, 2009, p. 361). Unsettling this knowledge demands emotional effort, careful listening to each other’s traumatic experiences, and explicit discussion of the potential and the harm that difficult knowledge stimulates.

The value of pedagogic discomfort in posttraumatic contexts cannot be overstated though. This process should not be assumed to be always already transformative, and beyond question. In other words, there are no guarantees for change in the social and political status quo; a pedagogy of discomfort, especially in light of the tensions identified here, demands time and realistic decisions about what can and what cannot be achieved. Also, concerns about the ethical implications of pedagogic discomfort must be foremost on the agenda (Zembylas, 2015b). That is, is pedagogic discomfort concerning difficult knowledge always appropriate and effective? How far can reconciliation pedagogies engage students’ discomfort without violating some ethical sensibilities? Needless to say, not all students will respond in the same way or benefit from pedagogic discomfort in the same manner; some may adopt some sort of change, others may resist, and still others may experience distress (Kumashiro, 2002).

The pedagogical principle of mutual vulnerability

The second aspect which enriches the task of reconciliation pedagogies in posttraumatic contexts concerns the pedagogical principle of *mutual vulnerability* (e.g., see Keet et al., 2009; Zembylas, 2009, 2014). The notion of mutual vulnerability is grounded in the idea that there is interdependence between human beings and that the recognition of all people as *vulnerable* has important pedagogical consequences concerning the possibility of assuming critical responsibility toward one’s own life and the lives of others in a community. The argument that is developed here is grounded theoretically in the work of Butler (2004; see also Butler, 2009, 2012, 2016).

Butler (2004) presents a number of examples to show that “each of us is constituted politically in part by virtue of the social vulnerability of our bodies ... Loss and vulnerability seem to follow from our being socially constituted bodies, attached to others, at risk of losing those attachments, exposed to others, at risk of violence by virtue of that exposure” (p. 20). This is evident, for instance, in the experience of losing someone whom one is attached to; thus each one of us is mutually obliged to others because of this common vulnerability. The denial of such vulnerability unleashes violence against others whereas its acknowledgment creates openings for an ethical encounter with others. Consequently, “we might critically evaluate and oppose,” Butler emphasizes, “the conditions under which certain human lives are more vulnerable than others, and thus certain human lives are more grievable than others” (p. 30). Once we consider how hegemonic power relations determine “who will be a grievable human” and what “acts”

are “permissible” for “public grieving” (p. 37), then we may begin to realize how a prohibition of grieving others’ lives extend the aims of violence and conflict.

The notion of vulnerability has important pedagogical consequences because the mutual experience of loss and mourning reveals the possibility of an alternative moral responsibility and sense of community (Vlieghe, 2010). Butler’s theorization of common vulnerability constitutes the point of departure for a renewed pedagogical politics of recognition in posttraumatic contexts—a politics that is not founded in a rationalist or individualistic morality, but in the experience of vulnerability as such (Zembylas, 2009, 2019). The notion of mutual vulnerability, then, enriches reconciliation pedagogies because it disrupts normative frames of community on the basis of rationality and self-advancement and puts forward the notion of community on the basis of loss. This idea does not imply, however, an equalization of vulnerability, but the recognition that there are different forms of vulnerabilities in posttraumatic contexts.

Furthermore, mutual vulnerability provides “a new grammar” for reconciliation pedagogies in posttraumatic contexts because the experiences of “the traumatized and the distressed and the micro-politics of peoples’ struggles for survival become its central pre-occupation” (Keet et al., 2009, p. 116). As Keet and his colleagues point out, writing in the context of post-apartheid South Africa, the notion of mutual vulnerability enhances the emancipatory interests of education “because the frames of its interlocutor are made vulnerable and therefore hospitable to moderation” (p. 116.). In this sense, there is more exchange of difficult knowledge between different sets of meaning-making frames; thus, in pedagogic terms, the notion of mutual vulnerability can be the starting point for acknowledging the burden of difficult knowledge carried more or less by all participants in posttraumatic settings.

The value of compassion and strategic empathy

Finally, the third principle I want to suggest here concerns the value of compassion and strategic empathy for reconciliation pedagogies in posttraumatic contexts (e.g., Zembylas, 2012, 2013b,c). In light of the discussion so far, it is evident that difficult knowledge provokes strong emotional reactions in the posttraumatic classroom—reactions that could be quite discomforting for teachers and students alike. Teachers have to find ways to handle constructively these reactions as well as the discourses that overly focus on one’s own traumatic experiences and ignore the other’s sufferings. A constructive relationship with difficult knowledge, suggest Bonnell and Simon (2007), is based on the process of confronting and dismantling any taken for granted assumptions upon encountering such unfamiliar knowledge. Importantly, then, teachers need to establish trust in the classroom, develop strong relationships and enact compassionate understanding in every possible manner. Reconciliation pedagogies in the context of posttraumatic situations require the strategic use of those pedagogical resources to enable the formation of new affective alliances among members of traumatized communities.

As noted earlier, attentiveness to mutual vulnerability is an important component of reconciliation pedagogies in posttraumatic contexts. Students are enabled to establish and maintain this attentiveness, when they begin to question and challenge arguments based on binaries like us/them, perpetrators/victims, friends/enemies, and good/evil, a stereotyping of groups considered to be *more* or *less* grievable (Butler, 2004, 2009). Jansen (2009) highlights two pedagogical tactics that I find particularly useful in teaching students how to learn compassion by challenging these binaries: first, the acknowledgment of brokenness by all sides, that is, the idea that humans are prone to failure and incompleteness and as such we constantly seek a higher order of living which cannot be accomplished without being in communion with others. Second, a pedagogical reciprocity is also required, that is, everyone carrying the burden of difficult knowledge has to move toward each other. As Jansen puts it in the context of post-apartheid South Africa: “the white person has to move across the allegorical bridge toward the black person; the black person has to move in the direction of the white person” (2009, p. 268).

To promote the prospects of compassion in the posttraumatic classroom, I also argue that one of those pedagogical resources that will be needed is strategic empathy. Strategic empathy is essentially the use of empathetic emotions in both critical and strategic ways (Lindquist, 2004); that is, it refers to the willingness of the teacher to make herself strategically skeptical (working sometimes against her own emotions) in order to empathize with the difficult knowledge students carry with them, even when this knowledge is disturbing to other students or to the teacher. The use of strategic empathy can function as a valuable pedagogical tool that opens up affective spaces which might eventually disrupt the emotional roots of difficult knowledge—an admittedly long and challenging task (Zembylas, 2012, 2013b,c). Undermining the emotional roots of difficult knowledge through strategic empathy ultimately aims at helping students integrate their troubled views into compassionate and socially just perspectives.

Compassion is further cultivated if students begin to understand the conditions (structural inequalities, poverty, globalization etc.) that give rise to difficult knowledge and suffering and acknowledge some sort of human connection (e.g., mutual vulnerability) between themselves and others. But mere understanding is clearly not enough; students will become more susceptible to affective transformation, when they enact compassionate action early on in their lives starting with simple things such as learning to be more patient and tolerant with peers who do not grasp a difficult concept in language or mathematics. As they grow up, children are offered opportunities to enact more complex manifestations of compassion that include action to alleviate the suffering of people who experience difficult times, no matter which community they come from. What needs to follow the acknowledgment of mutual vulnerability is taking action that dismisses essentialized categories of victims and oppressors and highlights instead the impact of solidarity on reducing everyday inequalities.

Conclusion

This chapter has constructed an argument about the importance of reconciliation pedagogies in posttraumatic contexts when emotional complexities and their consequences are taken into serious consideration. As it has been emphasized, this argument entails making pedagogical spaces for understanding difficult knowledge in more nuanced terms, a process enabled by “a pedagogy of strategic performance, in which teachers work to tactically position themselves as conduits of students’ affective responses” (Lindquist, 2004, p. 189). What this means for the practice of reconciliation pedagogies is that emotions are inevitably implicated in the way that difficult knowledge operates in a posttraumatic context. Without a fundamental revision of our thinking about difficult knowledge and its consequences, the radical potential of reconciliation pedagogies to reconstitute the emotional connections of traumatized students and teachers may be compromised, in spite of our best intentions (Zembylas, 2013a,b,c, 2018b). This task requires a constant reconsideration of new pedagogical resources to enhance the potential of reconciliation pedagogies such as the pedagogical principles discussed here; undoubtedly, there are many more ideas that can contribute toward this effort.

At the same time, the limitations and risks of using reconciliation pedagogies in posttraumatic contexts need to be kept in mind. There are important unanswered questions in the effort to make the emotional complexities of difficult knowledge explicit and constructive: How can teachers avoid becoming some sort of therapists for their students, when there is pressure in current times to provide therapeutic education? What are the limits of discomfort, mutual vulnerability and compassion in different posttraumatic settings so that reconciliation pedagogies retain their radical but also pragmatic potential? How far can teachers push their students to problematize their emotional responses toward difficult knowledge and to critique self and otherness in light (or in spite) of open wounds within a traumatized community? These and other important questions offer possibilities to broaden the ways that incorporate the emotional implications of difficult knowledge into our understanding of reconciliation pedagogies in posttraumatic situations.

References

- Ahmed, S., 2004. *The Cultural Politics of Emotion*. University of Edinburgh Press, Edinburgh.
- Amsler, S., 2011. From ‘therapeutic’ to political education: the centrality of affective sensibility in critical pedagogy. *Crit. Stud. Educ.* 52 (1), 47–63.
- Auerbach, Y., 2009. The reconciliation pyramid—a narrative-based framework for analyzing identity conflicts. *Polit. Psychol.* 30 (2), 291–318.
- Bekerman, Z., Zembylas, M., 2012. *Teaching Contested Narratives: Identity, Memory and Reconciliation in Peace Education and beyond*. Cambridge University Press, Cambridge.
- Berlak, A., 2004. Confrontation and pedagogy: cultural secrets and emotion in antioppressive pedagogies. In: Boler, M. (Ed.), *Democratic Dialogue in Education: Troubling Speech, Disturbing Silence*. Peter Lang, New York, pp. 123–144.
- Boler, M., Zembylas, M., 2003. Discomforting truths: the emotional terrain of understanding differences. In: Tryfonas, P. (Ed.), *Pedagogies of Difference: Rethinking Education for Social Justice*. Routledge, New York, pp. 110–136.
- Boler, M., 1999. *Feeling Power: Emotions and Education*. Routledge, New York.
- Boler, M., 2004. Teaching for hope: the ethics of shattering world views. In: Liston, D., Garrison, J. (Eds.), *Teaching, Learning and Loving: Reclaiming Passion in Educational Practice*. RoutledgeFalmer, New York, pp. 117–131.
- Bonnell, J., Simon, R., 2007. “Difficult” exhibitions and intimate encounters. *Mus. Soc.* 5 (2), 65–85.
- Britzman, D.P., 2000. If the story cannot end: deferred action, ambivalence, and difficult knowledge. In: Simon, R.I., Rosenberg, S., Eppert, C. (Eds.), *Between Hope and Despair: The Pedagogical Encounter of Historical Remembrance*. Rowman & Littlefield Publishers, Lanham, MD, pp. 27–57.
- Britzman, D.P., 2013. Between psychoanalysis and pedagogy: scenes of rapprochement and alienation. *Curric. Inq.* 43 (1), 95–117.
- Butler, J., 2004. *Precarious Life: The Powers of Mourning and Violence*. Verso, London.
- Butler, J., 2009. *Frames of War: When Life is Grievable*. Verso, London.
- Butler, J., 2012. Precarious life, vulnerability, and the ethics of cohabitation. *J. Specul. Philos.* 26 (2), 134–151.
- Butler, J., 2016. Rethinking vulnerability and resistance. In: Butler, J., Gambetti, Z., Sabsay, L. (Eds.), *Vulnerability in Resistance*. Duke University Press, Durham, NC, pp. 12–27.
- Chouliaraki, L., 2013. *The Ironic Spectator: Solidarity in the Age of Post-humanitarianism*. Polity Press, Cambridge, UK.
- Cohen, S., 2001. *States of Denial: Knowing about Atrocities and Suffering*. Polity Press, Cambridge.
- Cole, E., 2007. Introduction: reconciliation and history education. In: Cole, E. (Ed.), *Teaching the Violent Past: History Education and Reconciliation*. Rowman & Littlefield, Lanham, MD, pp. 1–28.
- Des Forges, A., 1999. *Leave None to Tell the Story: Genocide in Rwanda*. Human Rights Watch, New York, NY.
- Ecclestone, K., Hayes, D., 2009. *The Dangerous Rise of Therapeutic Education*. Routledge, Abingdon, UK.
- Flam, H., 2013. The transnational movement for Truth, Justice and Reconciliation as an emotional (rule) regime? *J. Polit. Power* 6 (3), 363–383.
- Foucault, M., 1977. *Discipline and Punish (Trans.)*. In: Sheridan, A. (Ed.). Vintage Books, New York.
- Halpern, J., Weinstein, H.M., 2004. Rehumanizing the other: empathy and reconciliation. *Hum. Right Q.* 26, 561–583.
- Hattam, R., Atkinson, S., Bishop, P., 2012. Rethinking reconciliation and pedagogy in unsettling times. In: Ahluwalia, P., Atkinson, S., Bishop, P., Christie, P., Hattam, R., Matthews, J. (Eds.), *Reconciliation and Pedagogy*. Routledge, London, pp. 1–9.
- Hochschild, A., 1979. Emotion work, feeling rules, and social structure. *Am. J. Sociol.* 85 (3), 551–575.
- Hook, D., 2005. Affecting whiteness: racism as technology of affect. *Int. J. Crit. Psychol.* 16 (1), 74–99.
- Hook, D., 2007. *Foucault, Psychology and the Analytics of Power*. Palgrave, New York.
- Hutchison, E., Bleiker, R., 2008. Emotional reconciliation: reconstituting identity and community after trauma. *Eur. J. Soc. Theor* 11 (3), 385–403.
- Jansen, J., 2009. *Knowledge in the Blood: Confronting Race and the Apartheid Past*. Stanford University Press, Stanford.
- Keet, A., Zinn, D., Porteus, K., 2009. Mutual vulnerability: a key principle in a humanising pedagogy in post-conflict societies. *Perspect. Educ.* 27 (2), 109–119.
- Kemper, T., 1978. Toward a sociology of emotions: some problems and some solutions. *Am. Sociol.* 13, 30–41.
- Kriesberg, L., 1998. Coexistence and the reconciliation of communal conflicts. In: Weiner, E. (Ed.), *The Handbook of Interethnic Coexistence*. Continuum, New York, NY, pp. 182–198.
- Kumashiro, K.K., 2002. Against repetition: addressing resistance to anti-oppressive change in the practices of learning, teaching, supervising, and researching. *Harv. Educ. Rev.* 72 (1), 67–92.

- Leibowitz, B., Bozalek, V., Rohleder, P., Carolissen, R., Swartz, L., 2010. "Ah, but the whiteys love to talk about themselves": discomfort as a pedagogy for change. *Race Ethn. Educ.* 13 (1), 83–100.
- Lindquist, J., 2004. Class affects, classroom affectations: working through the paradoxes of strategic empathy. *Coll. Engl.* 67 (2), 187–209.
- Mohanty, C., 2003. *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity*. Duke University Press, Durham.
- Oliver, K., 2001. *Witnessing: Beyond Recognition*. University of Minnesota Press, Minneapolis.
- Plamber, J., 2010. The history of emotions: an interview with William Reddy, Barbara Rosenwein, and Peter Stearns. *Hist. Theor.* 49 (2), 237–365.
- Reddy, W.M., 2001. *The Navigation of Feeling: A Framework for the History of Emotions*. Cambridge University Press, Cambridge, Eng.
- Retzinger, S., Scheff, T., 2000. Emotion, alienation, and narratives: resolving intractable conflict. *Mediat. Q.* 18, 71–85.
- Rose, N., 1998. *Inventing Ourselves: Psychology, Power, and Personhood*. Cambridge University Press, Cambridge, UK.
- Rosenwein, B., 2002. Worrying about emotions in history. *Am. Hist. Rev.* 2, 821–845.
- Schaap, A., 2006. Agonism in divided societies. *Philos. Soc. Critic.* 32 (2), 255–277.
- Simon, R.I., 2011a. A shock to thought: curatorial judgment and the public exhibition of 'difficult knowledge'. *Mem. Stud.* 4 (4), 432–449.
- Simon, R.I., 2011b. Afterword: the turn to pedagogy: a needed conversation on the practice of curating difficult knowledge. In: Lehrer, E., Milton, C.E., Patterson, M.E. (Eds.), *Curating Difficult Knowledge: Violent Pasts in Public Places*. Palgrave Macmillan, New York, pp. 193–209.
- Stearns, P.N., Stearns, C.Z., 1985. Emotionology: clarifying the history of emotions and emotional standards. *Am. Hist. Rev.* 90, 813–836.
- Vlieghe, J., 2010. Judith Butler and the public dimension of the body: education, critique and corporeal vulnerability. *J. Philos. Educ.* 44 (1), 153–170.
- Worsham, L., 2006. Composing (identity) in a posttraumatic age. In: Williams, B.T. (Ed.), *Identity Papers: Literacy and Power in Higher Education*. Utah, State University Press, Logan, pp. 170–181.
- Zembylas, M., McGlynn, C., 2012. Discomforting pedagogies: emotional tensions, ethical dilemmas and transformative possibilities. *Br. Educ. Res. J.* 38 (1), 41–60.
- Zembylas, M., Papamichael, E., 2017. Teaching for anti-racism by troubling teachers' comfort zones: discomfort and empathy as pedagogies in teacher professional development. *Intercult. Educ.* 28 (1), 1–19.
- Zembylas, M., Charalambous, C., Charalambous, P., 2016. *Peace Education in a Conflict-Troubled Society*. Cambridge University Press, Cambridge.
- Zembylas, M., 2008. *The Politics of Trauma in Education*. Palgrave, MacMillan, New York.
- Zembylas, M., 2009. Counter-narratives of mourning the missing persons in Cyprus: pedagogical limits and openings for reconciliation education in conflict-ridden societies. *Perspect. Educ.* 29 (2), 120–132.
- Zembylas, M., 2010. Agamben's theory of biopower and immigrants/refugees/asylum seekers: discourses of citizenship and the implications for curriculum theorizing. *J. Curriculum Theor.* 26 (2), 31–45.
- Zembylas, M., 2012. Pedagogies of strategic empathy: navigating through the emotional complexities of antiracism in higher education. *Teach. High. Educ.* 17 (2), 113–125.
- Zembylas, M., 2013a. Critical emotional praxis: rethinking teaching and learning about trauma and reconciliation in schools. In: Trifonas, P., Wright, B. (Eds.), *Critical Peace Education: Difficult Dialogues*. Springer, Dordrecht: The Netherlands, pp. 101–114.
- Zembylas, M., 2013b. The 'crisis of pity' and the radicalization of solidarity: towards critical pedagogies of compassion. *Educ. Stud.* 49, 504–521.
- Zembylas, M., 2013c. The emotional complexities of "our" and "their" loss: the vicissitudes of teaching about/for empathy in a conflicting society. *Anthropol. Educ. Q.* 44 (1), 19–37.
- Zembylas, M., 2014. Theorizing 'difficult knowledge' in the aftermath of the 'affective turn': implications for curriculum and pedagogy in handling traumatic representations. *Curric. Inq.* 44 (3), 390–412.
- Zembylas, M., 2015a. *Emotion and Traumatic Conflict: Re-claiming Healing in Education*. Oxford University Press, Oxford.
- Zembylas, M., 2015b. 'Pedagogy of discomfort' and its ethical implications: the tensions of ethical violence in social justice education. *Ethics Educ.* 10 (2), 1–12.
- Zembylas, M., 2015c. Rethinking race and racism as technologies of affect: theorizing the implications for antiracist politics and practice in education. *Race Ethn. Educ.* 18 (2), 145–162.
- Zembylas, M., 2016. Emotion, trauma, and critical pedagogy: implications for critical peace education. In: Bajaj, M., Hantzopoulos, M. (Eds.), *Peace Education: International Perspectives*. Bloomsbury, New York, pp. 19–34.
- Zembylas, M., 2017. Practicing an ethic of discomfort as an ethic of care in higher education teaching. *Crit. Stud. Teach. Learn.* 5 (1), 1–17.
- Zembylas, M., 2018a. Affect, race, and white discomfort in schooling: decolonial strategies for 'pedagogies of discomfort'. *Ethics Educ.* 13 (1), 86–104.
- Zembylas, M., 2018b. Understanding the emotional regimes of reconciliation in engagements with 'difficult' heritage. In: Smith, L., Wetherell, M., Campbell, G. (Eds.), *Emotion, Affective Practices and the Past in the Present*. Routledge, New York, pp. 197–212.
- Zembylas, M., 2019. The ethics and politics of precarity: risks and productive possibilities of a critical pedagogy for precarity. *Stud. Philos. Educ.* 38 (2), 95–111.

Teacher education and the development of teacher noticing

Johannes König, University of Cologne, Faculty of Human Sciences, Department of Education and Social Sciences, Empirical School Research, Quantitative Methods, Cologne, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	323
Conceptualizing teacher noticing	323
Different perspectives on teacher noticing research	323
Cognitive-psychological perspective	324
Socio-cultural perspective	324
Expertise-related perspective	324
Prevalence of perspectives	324
Facets of teacher noticing	324
Noticing as part of teacher professional competence	325
Modeling and measuring teacher noticing	325
Teacher noticing in mathematics	325
Teacher noticing for effective classroom management	326
The development of teacher noticing during teacher education	326
Intervention studies	326
Expert-novice comparisons	326
Conclusion	327
References	327

Introduction

In their daily work, teachers are confronted with complex and unpredictable situations. To master demands that are typical for their profession, teachers are required to pay attention to relevant classroom processes and events of social interaction, select important information, draw appropriate conclusions, and decide quickly on how to proceed with their instruction. The process-related description of underlying abilities and cognitive skills are frequently referred to by the term ‘teacher noticing’. Since they are assumed to be specific for successful teaching, teacher noticing is considered being essential for the professional teacher and contributing to his or her professional competence.

Over the last two decades, research on teacher noticing has remarkably increased (König et al., 2022). Several literature reviews have been published (Amador et al., 2021; König et al., 2022; Santagata et al., 2021; Stahnke et al., 2016), summarized in encyclopedia entries (Choy and Dindyal, 2020; Jacobs et al., 2018). Moreover, teacher noticing has been discussed as a core constituent of teacher professional competence (Kaiser and König, 2019), significantly impacting the research discourse on teacher competence and teacher education effectiveness (König and Blömeke, 2020).

Various approaches to conceptualize teacher noticing exist, making it difficult to provide a single clear-cut definition of teacher noticing as a scientific term. One reason is that teacher noticing is not one-dimensional, but comprises several facets such as attention, perception, interpretation, and decision-making. Another reason is the diversity in empirical research approaches and research foci deployed by teacher noticing researchers. Over years, different perspectives have shaped research on teacher noticing (König et al., 2022; Santagata et al., 2021), leading to a heterogeneous state of research.

Moreover, teacher noticing is not limited to in-service teachers. In fact, numerous empirical studies focus on pre-service teachers learning to notice (König et al., 2022). Intervention studies show the impact specific learning opportunities have on such pre-service teacher learning (see, for mathematics, the review by Santagata et al., 2021). As will be shown in the following, teacher noticing has its place in teacher education.

The present article first provides insight into how teacher noticing has been conceptualized. It then uses two examples from teacher noticing research illustrating how empirical research has proliferated our understanding of teacher noticing, also with regards to teacher professional competence. An overview is given on how teacher education may support pre-service teacher learning to notice. We conclude with a discussion on perspectives for future research and needs for teacher education.

Conceptualizing teacher noticing

Different perspectives on teacher noticing research

In their systematic literature review on teacher noticing, which was based on about 180 publications on teacher noticing, König et al. (2022) state that a common understanding of the construct of noticing in the literature is missing. Based on existing literature

reviews and survey papers, König et al. (2022) therefore suggest to differentiate between relevant perspectives on teacher noticing research such as a “cognitive psychological perspective” referring to the work by van Es and Sherin (2002), a “socio-cultural perspective” originating from the work by Goodwin (1994) on professional vision and an “expertise-related perspective” resulting from the research paradigm of differences between novice and expert teachers as discussed by Berliner (1988). (For a fourth perspective, too detailed to present here as well, see König et al., 2022).

Cognitive-psychological perspective

According to König et al. (2022), a cognitive-psychological perspective can be identified that focuses on noticing in terms of what teachers attend to and make sense of. This perspective is closely related to the well-known work by van Es and Sherin (2002: 573) who proposed the following conceptualization of teacher noticing: “(a) identifying what is important or noteworthy about a classroom situation; (b) making connections between the specifics of classroom interactions and the broader principles of teaching and learning they represent; and (c) using what one knows about the context to reason about classroom interactions”.

Over the last two decades, this conceptualization has been used as a theoretical reference for numerous empirical studies. As a frequently cited article, van Es and Sherin (2002) might be considered as a cornerstone in teacher noticing research. Meanwhile, further differentiations of this conceptualization have been suggested as well indicating that “no consensus has yet been reached on the different facets of noticing” (König et al., 2022).

Socio-cultural perspective

Emphasis on the social and situated nature of teacher noticing has been given with the well-known conceptualization of “professional vision” as it was introduced by Goodwin (1994). He proposed that “the ability to see a meaningful event” is “a socially situated activity” which evolves through discursive practices of professionals (Goodwin, 1994: 606). These “discursive practices are used by members of a profession to shape events in the domains subject to their professional scrutiny. The shaping process creates the objects of knowledge that become the insignia of a profession’s craft: the theories, artifacts, and bodies of expertise that distinguish it from other professions.” (Goodwin, 1994: 606). As Goodwin further explains, on the basis of these discursive practices, “participants build and contest *professional vision*, which consists of socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular group” (Goodwin, 1994: 606). Thus the conceptualization of “professional vision” relates to the group-specific perspectivation and privileges, for example in terms of expertise, thus serving the demarcation from persons not belonging to the relevant profession (e.g., clients or laypersons) (Lefstein and Snell, 2011).

It is important to note that the conceptualization of “professional vision” by Goodwin (1994) and its development was not specified for the teaching profession. As König et al. (2022) outline, later on “professional vision” was incorporated into the discussion of teacher noticing as it emerged through the previously described cognitive-psychological perspective. As a consequence, numerous authors relate their work on teacher noticing to the conceptualization of professional vision (e.g., Sherin and van Es, 2009; Jacobs et al., 2010).

Expertise-related perspective

In applying and specifying the expertise research paradigm to teachers, Berliner and colleagues (Berliner, 1988, 2001, 2004; Carter et al., 1988; Sabers et al., 1991) created important approaches to teacher noticing as well. Although teacher expertise research as a paradigm may not directly focus on teacher noticing (Bromme, 2001; Stigler and Miller, 2018), teacher expertise and its development from novice to expert forms an important reference supporting the conceptualizations from cognitive-psychological and socio-cultural perspective. Therefore, Lachner et al. (2016: 198) consider expertise research “as precursors” to research on teacher noticing.

Prevalence of perspectives

In their systematic literature review of about 180 publications on teacher noticing, König et al. (2022) analyzed how the different perspectives have shaped empirical research studies on teacher noticing over the last two decades. It turned out that the cognitive-psychological perspective has strongly come to the foreground, whereas combinations with the other perspectives can be observed and interpreted as enrichment of conceptualizing teacher noticing. In contrast, the socio-cultural perspective related to professional vision has been applied by empirical studies to a far less extent. As König et al. (2022) state, only few researchers actually followed the original approach initiated by Goodwin (1994) such as McDonald (2016), Lefstein and Snell (2011) or Louie (2018). At the same time, the notion of professional vision has been taken up in quantitative studies, mainly originating from Germany, with a close link to concepts of instructional quality (e.g., Seidel and Stürmer, 2014), leading, however, into a direction that clearly differs from the Goodwin’s (1994) original approach (see, for more details, König et al., 2022).

Facets of teacher noticing

To understand conceptualizations of teacher noticing more closely, it is important to look at differentiations into facets as suggested by teacher noticing researchers. As König et al. (2022) conclude from their systematic review of about 180 publications on teacher noticing, empirical approaches principally differ with respect to whether researchers have investigated noticing as a “holistic concept” or developed a differentiation into facets, typically “attending/perceiving”, “interpreting/reasoning”, and

“decision-making”. For example, [Kaiser et al. \(2015\)](#) identified perception, interpretation, and decision making as facets of noticing. The cognitive processes behind these faces are also well-known from teacher reflection approaches, which underlines the relevance of differentiating teacher noticing in facets. For example, in her “reflective cycle”, [Rodgers \(2002\)](#) suggested to distinguish among three phases of teacher’s action: (1) describing significant events of teaching, (2) assigning meaning to these events, (3) making decisions for how to proceed. While a large number of empirical studies deploy qualitative methodology and investigate teacher noticing holistically, [König et al. \(2022\)](#) identified an increase of quantitative studies that are committed to develop standardized measures with the explicit objective to differentiate various facets of teacher noticing.

Noticing as part of teacher professional competence

In recent studies on teacher professional competence, teacher noticing has become part of modeling and measuring competence. In particular, [Blömeke et al. \(2015a\)](#) presented a framework of teacher competence that accounts for the interaction of personal, situational, and social characteristics. Starting from the argument that former conceptualizations of competence neglect either the stable dispositional or the more variable situational competence facets and suggesting approaches should complement each other, [Blömeke et al. \(2015a\)](#) describe competence along a continuum from personal dispositions, comprising teacher professional knowledge and beliefs, to situation-specific cognitive skills, comprising perception, interpretation, and decision-making, that finally predict the observed teacher performance in the classroom (as a similar competence as continuum model see the “cascade model” by [Krauss et al., 2020](#)). As these situation-specific skills not only cover the perception and interpretation of situations as well as subsequent decision-making, but are also based on teachers’ professional knowledge, the framework developed by [Blömeke et al. \(2015a\)](#) overlaps conceptually with teacher noticing ability to pay selective attention to critical locations in a complex classroom environment as well as “the ways in which a teacher reasons about what is noticed based on his or her knowledge and understanding” ([Sherin and van Es, 2009](#): 22). Thus, current advances in modeling teacher professional competence not only integrate conceptualizations of teacher noticing, but they also link teacher noticing to teacher professional knowledge as described and classified in teacher knowledge research ([Shulman, 1987](#)) as well as referenced by teacher expertise research ([Bromme, 2001](#); [Stigler and Miller, 2018](#)). Moreover, the function of teacher noticing for successful teacher performance comes to the foreground as well. That means, teacher noticing is not only a construct to be investigated as such, but teacher noticing researchers are also encouraged to provide evidence for its predictive validity in classroom teaching and student learning ([König et al., 2021](#)).

Modeling and measuring teacher noticing

When collecting empirical research data on teacher noticing, many researchers typically deploy video material of teaching as a means of stimulating noticing among pre-service or in-service teachers ([Gaudin and Chaliès, 2015](#)). As quantitative approaches to model and measure teacher noticing have increased recently ([König et al., 2022](#)), two examples are provided in the following. The first approach was developed in the context of the Teacher Education and Development Study in Mathematics (TEDS-M) research program ([Kaiser and König, 2019](#)), with specific focus on mathematics education. The second approach relates to effective classroom management as one of the core demands teachers have to master during their daily work ([König, 2015](#)).

Teacher noticing in mathematics

Between 2011 and 2014, a follow-up of the comparative large scale assessment TEDS-M 2008 was conducted in Germany. Besides broad assessments of teacher knowledge, a novel approach was developed to assess noticing of early career teachers. Following the strategy of TEDS-M to recruit samples that would represent a country’s teacher candidate population for teaching mathematics, the follow-up survey, conducted about four years after TEDS-M, addressed former participants of the study TEDS-M from all over Germany who had entered the teaching profession by 2012.

Teacher noticing was conceptualized for mathematics and pedagogy. Three staged video-clips with a length of about 3–4 min were created as stimuli, followed by closed and open response tests items to measure situation-specific skills of perception, interpretation, and decision-making. Some of these test items were related to mathematics pedagogical content knowledge (MPCK), others to general pedagogical knowledge (GPK) of teachers. The video-clips were construed in a way that they relate to lengthy intervals of teaching a lesson, although providing thin slices of classroom events and procedures.

Test items had two different formats. Some of them were multiple-choice items with a Likert-rating scale (four categories) ranging from “completely agree” to “completely disagree”. They were used to measure the precision of teacher perception (e.g., “The teacher presents the lesson’s task visually AND acoustically”) or the holistic nature of teacher perception (e.g., “Most students take an active part in the lesson”). An expert review was conducted to assure the correct solution of each item’s rating scale (see for details, [Hoth et al., 2016](#)).

Open-response items were applied to measure the situation-specific skills of interpretation and decision-making. These items were designed to enable teachers in providing cognitively more complex statements. For example, one of the video-clips focuses on students computing the volume of a box during pair-work. The video-clip shows three different pairs and visualizes different kinds of cooperation. One of the open-response items then asks the testtakers for an in-depth analysis of cooperation from a pedagogical perspective. Based on test items and scaling analyses, reliable total scores of teacher noticing were created,

one for mathematics pedagogy, another one for general pedagogy (Blömeke et al., 2015b; König et al., 2014). More specific scaling analyses on teacher noticing can be found in König et al. (2014). First, structural analysis showed perception and interpretation are two situation-specific processes that can empirically be separated thus confirming important theoretical underpinnings of the noticing construct. Second, declarative teacher knowledge seems to be more closely related to processes of interpretation than to processes of perception. Further findings from this kind of research on the nature of teacher noticing as a construct has been currently proliferated in a more recent analysis of the TEDS-M research program by Bastian et al. (2022), who investigated how teachers' noticing can be empirically differentiated and subscales of perception, interpretation, and decision-making be used for a description of growth toward expertise in secondary mathematics teaching.

Teacher noticing for effective classroom management

In a different project (König, 2015), an approach was developed to specifically measure and model teacher noticing for effective classroom management. Using the expertise-related perspective (see Section **Expertise-related perspective**), the focus was on measuring expert-novice differences in the perception and interpretation of teacher-student interaction situations typical for effective classroom management. As the test instrument primarily targeted on in-service teachers with sufficient teaching experience (as a precondition to expertise, but not being the same, see Caspari-Sadeghi and König, 2018), the test instrument was denoted as Classroom Management Expertise (CME).

Similar to the video-based instruments in the TEDS-M follow-up study (see previous section), four video-clips served as stimulus, followed by test items relating to each video-clip specifically. From a content-related perspective, the situations depicted in the video-clips refer to typical classroom management situations. Following classifications of typical classroom management situations found in the literature (Hawk and Schmidt, 1989; Doyle, 2006), the video clips had to represent classroom instruction in which a teacher is challenged (1) to manage transitions, (2) to manage instructional time, (3) to manage student behavior, and (4) to manage instructional feedback to students. Test development comprised several steps: First, an expert review before the procedure of selecting appropriate video clips was started. Second, test items were developed to cover situation-specific skills of perception and interpretation. Third, coding rubrics were developed for the open-response items. Fourth, in a pilot study rubrics were examined and finalized for a coding manual that could be used in training raters. Based on a sample of about 120 in-service teachers, findings showed that teacher noticing for effective classroom management can be measured in a reliable way using the CME test (König, 2015).

Further studies using the CME test show that teacher noticing for effective classroom management is specifically related to declarative pedagogical knowledge in the area of effective classroom management rather than pedagogical knowledge of structuring lessons, dealing with heterogeneous learning groups, or assessing students (König and Kramer, 2016). The higher a teacher scores in the CME test, the better his or her instructional quality is rated by external observers, which in turn contributes to learning progress of students, as the finding from multilevel-analyses by König et al. (2021) can show.

The development of teacher noticing during teacher education

Intervention studies

Reviews have highlighted the overall support of using videos in teacher education, also to stimulate pre-service teachers' learning to notice (Gaudin and Chaliès, 2015; Santagata et al., 2021). More specifically, Santagata et al. (2021) analyzed a selection of 35 empirical studies that deployed video as a supporting tool in mathematics teacher education. It turned out that there were several benefits in using video as a tool, for example for "anchoring discussions of teaching and learning around specific evidence" (Santagata et al., 2021: 128), and the underlying frameworks for watching videos provided a helpful orientation for learning to notice.

Using about 70 articles for their literature review of intervention studies in teacher noticing during teacher education, König et al. (2022) analyzed what the authors of each study themselves reported and interpreted about the improvement in noticing skills. Findings showed that about at least half of the studies comprised descriptions on the improvement of (pre-service) teachers' skills in the teacher noticing area of attending and/or reasoning. Only few studies reported on more complex skills such as proposing alternatives which was found in less than 10% of the studies. So further empirical investigation into how teachers make decisions based on their noticing seems to be a research desideratum. Another problematic issue is that authors used different interventions designs and noticing measures, making it difficult to reliably compare findings across different studies.

Expert-novice comparisons

Following the expertise-related perspective, some studies on teacher noticing have investigated possible learning progress in contrasting pre-service and the more experienced in-service teachers. König et al. (2022) systematically identified a group of studies (Gold and Holodynski, 2017; Jacobs et al., 2010; Meschede et al., 2017; Star and Strickland, 2008; Wolff et al., 2016; Dreher and Kuntze, 2015; Reuter, 2017a, b) that all approached expert-novice contrasts by using standardized measures for teacher noticing. These studies share the common finding that experts outperformed novices, although different constructs and facets of

noticing were deployed by the studies. This allows to conclude that differences between novices and experts seem to be quite stable across different subjects and research designs (see, for more details, König et al., 2022).

Another example for differences between novice teachers and experienced teachers is provided by König and Kramer (2016) using the CME measure (see Section **Teacher noticing for effective classroom management**) for investigating teacher noticing for effective classroom management. Experienced teachers also outperformed pre-service teachers, showing that teacher noticing for effective classroom management “is much more dependent on the expertise level acquired during professional development” (König and Kramer, 2016: 149).

Conclusion

Teacher noticing is an important construct to describe teacher professionalization and teacher professional competence. Research on teacher noticing has remarkably increased over the last two decades, proliferating empirical evidence on its relevance for teaching and teacher education. In particular, quantitative studies with standardized measurement of teacher noticing that target on pre-service teachers learning to notice provide substantial evidence through interventions or expert-novice difference analyses.

Although this increase and proliferation of findings on teacher noticing and its development during initial teacher education is promising, the state of research clearly has its limitations. For example, while van Es and Sherin (2002: 573) highlighted “identifying” and “making connections” as key aspects of noticing, the systematic review conducted by König et al. (2022) showed that only about 10% of all studies provided an analysis of how noticing is connected with professional teacher knowledge. This indicates a research desideratum indicated as early as by Schoenfeld (2011). During the last two decades, standardized assessments of teacher knowledge have increased (Kaiser and König, 2019), therefore it might be just the next step to figure out the relationship between knowledge and noticing in near future research (Yang et al., 2020).

In a similar way, only few studies examined how teacher noticing is actually related to teaching practice. Competence models such as conceptualized by Blömeke et al. (2015a) emphasize the significance of noticing skills like perception, interpretation, and decision-making for observed teacher behavior which in turn should contribute to the quality of learning opportunities delivered to students as well as the learning progress of students. However, only few studies have actually provided in-depth analysis of the underlying chain that comprises measures of noticing, measures of teaching practice, and student learning outcomes. For example, Kersting et al. (2016) provided empirical evidence for mathematics teacher noticing can predict instructional practice. Similar, in a recent analysis, König et al. (2021) demonstrated that teacher noticing for effective classroom management (as measured via the CME instrument, see Section **Teacher noticing for effective classroom management**) was predictive for instructional quality which in turn contributed to student learning in mathematics. However, correlations were only of small to moderate effect size, and not all hypotheses developed by authors could be confirmed, making it necessary to intensify this specific research field. Moreover, competence models themselves should further be examined, for example by more studies using the expert-novice comparison principle. Finally, although the number of studies trying to distinguish various facets of teacher noticing have increased and may contribute to our understanding of the cognitive processes underlying noticing, our theoretical and empirical knowledge seems to be limited. Researchers should be encouraged to develop and apply innovative empirical approaches in the near future, such as using new digital technologies to provide authentic stimuli (e.g., using virtual reality glasses or eye tracking, see Smith, 2012; Wolff et al., 2016). But the further elaboration of theoretical frameworks on teacher noticing seems also a necessity to show future directions of teacher noticing research.

References

- Amador, J., Bragelman, J., Castro Superfine, A., 2021. Prospective teachers noticing: a literature review of methodological approaches to support and analyze noticing. *Teach. Educ.* 99, 103256.
- Bastian, A., Meyer, D., Schwarz, B., Kaiser, G., König, J., 2022. Teachers' noticing and its growth toward expertise: an expert-novice comparison with (pre-service) secondary mathematics teachers. *Educ. Stud. Math.* 110, 205–232. <https://doi.org/10.1007/s10649-021-10128-y>.
- Berliner, D.C., 1988. The development of expertise in pedagogy. *Am. Assoc. Coll. Teach.* AACTE Publications, Washington DC.
- Blömeke, S., Gustafsson, J.-E., Shavelson, R.J., 2015a. Beyond dichotomies: competence viewed as a continuum. *Z. für Psychol.* 223 (1), 3–13.
- Blömeke, S., Hoth, J., Dörmann, M., Busse, A., Kaiser, G., König, J., 2015b. Teacher change during induction: development of beginning primary teachers' knowledge, beliefs and performance. *Int. J. Sci. Math. Educ.* 13 (2), 287–308.
- Bromme, R., 2001. Teacher expertise. In: Smelser, N.J., Baltes, P.B. (Eds.), *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier, Amsterdam, Netherlands, pp. 1545–1546.
- Carter, K., Cushing, D., Stein, P., Berliner, D., 1988. Expert-novice differences in perceiving and processing visual classroom information. *J. Teach. Educ.* 39, 25–31.
- Caspari-Sadeghi, S., König, J., 2018. On the adequacy of expert teachers: from practical convenience to psychological reality. *Int. J. High. Educ.* 7 (5), 1–19.
- Choy, B.H., Dindyal, J., 2020. Teacher noticing, mathematics. In: Peters, M. (Ed.), *Encyclopedia of Teacher Education*. Springer.
- Doyle, W., 2006. Ecological approaches to classroom management. In: Evertson, C.M., Weinstein, C.S. (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues*. Erlbaum, Mahwah, NJ, pp. 97–125.
- Dreher, A., Kuntze, S., 2015. Teachers facing the dilemma of multiple representations being aid and obstacle for learning: evaluations of tasks and theme-specific noticing. *Journal für Mathematik-Didaktik* 36 (1), 23–44.
- Gaudin, C., Chaliès, S., 2015. Video viewing in teacher education and professional development: a literature review. *Educ. Res. Rev.* 16, 41–67.
- Gold, B., Holodynski, M., 2017. Using digital video to measure the professional vision of elementary classroom management: test validation and methodological challenges. *Comput. Educ.* 107, 13–30.

- Goodwin, C., 1994. Professional vision. *Am. Anthropol.* 96 (3), 606–633.
- Hawk, P.P., Schmidt, M.W., 1989. Teacher preparation a comparison of traditional and alternative programs. *J. Teach. Educ.* 40 (5), 53–58.
- Hoth, J., Schwarz, B., Kaiser, G., Busse, A., König, J., Blömeke, S., 2016. Uncovering the predictors of disagreement: ensuring the quality of expert ratings. *ZDM – Math. Educ.* 48 (1), 83–95.
- Jacobs, V.R., Lamb, L.L., Philipp, R.A., 2010. Professional noticing of children's mathematical thinking. *J. Res. Math. Educ.* 41 (2), 169–202.
- Jacobs, V.R., Philipp, R.A., Sherin, M.G., 2018. Noticing of mathematics teachers. In: Lerman, S. (Ed.), *Encyclopedia of Mathematics Education*. Springer.
- Kaiser, G., Busse, A., Hoth, J., König, J., Blömeke, S., 2015. About the complexities of video-based assessments: theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *Int. J. Sci. Math. Educ.* 13 (2), 369–387.
- Kaiser, G., König, J., 2019. Competence measurement in (mathematics) teacher education and beyond: implications for policy. *High Educ. Pol.* 32 (4), 597–615.
- Kersting, N.B., Sutton, T., Kalinec-Craig, C., Jablon Stoeck, K., Heshmati, S., Lozano, G., Stigler, J.W., 2016. Further exploration of the classroom video analysis (CVA) instrument as a measure of useable knowledge for teaching mathematics: taking a knowledge system perspective. *ZDM – Math. Educ.* 48, 97–109.
- König, J., 2015. Measuring classroom management expertise (CME) of teachers: a video-based assessment approach and statistical results. *Cog. Educ.* 2 (1), 991178.
- König, J., Blömeke, S., 2020. Wirksamkeits-Ansatz in der Lehrerinnen- und Lehrerbildung. [Teacher education effectiveness approach.]. In: Cramer, C., König, J., Rothland, M., Blömeke, S. (Eds.), *Handbuch Lehrerinnen- und Lehrerbildung*. [Handbook teacher education.]. Klinkhardt/UTB, Bad Heilbrunn, pp. 172–178.
- König, J., Blömeke, S., Jentsch, A., Schlesinger, L., Felske, C., Musekamp, F., Kaiser, G., 2021. The links between pedagogical competence, instructional quality, and mathematics achievement in the lower secondary classroom. *Educ. Stud. Math.* 107, 189–212.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., Kaiser, G., 2014. Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teach. Teach. Educ.* 38, 76–88.
- König, J., Kramer, C., 2016. Teacher professional knowledge and classroom management: on the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM – Math. Educ.* 48 (1), 139–151.
- König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., Kaiser, G., 2022. Teacher noticing: a systematic literature review on conceptualizations, research designs, and findings on learning to notice. *Educ. Res. Rev.* <https://doi.org/10.1016/j.edurev.2022.100453>.
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N., Blum, W., 2020. Competence as a continuum in the COACTIV study: the "cascade model". *ZDM – Math. Educ.* 52 (2), 311–327.
- Lachner, A., Jarodzka, H., Nückles, M., 2016. What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instr. Sci.* 44, 197–203.
- Lefstein, A., Snell, J., 2011. Professional vision and the politics of teacher learning. *Teach. Teach. Educ.* 27 (3), 505–514.
- Louie, N.L., 2018. Culture and ideology in mathematics teacher noticing. *Educ. Stud. Math.* 97 (1), 55–69.
- McDonald, S.P., 2016. The transparent and the invisible in professional pedagogical vision for science teaching. *Sch. Sci. Math.* 116 (2), 95–103.
- Meschede, N., Fiebranz, A., Möller, K., Steffensky, M., 2017. Teachers' professional vision, pedagogical content knowledge and beliefs: on its relation and differences between pre-service and in-service teachers. *Teach. Teach. Educ.* 66, 158–170.
- Reuter, S., 2017a. The noticing of physical education teachers: a comparison of groups with different expertise. *Phys. Educ. Sport Pedagog.* 22 (2), 150–170.
- Reuter, S., 2017b. The knowledge-based reasoning of physical education teachers: a comparison between groups with different expertise. *Eur. Phys. Educ. Rev.* 23 (1), 3–24.
- Rodgers, C., 2002. Seeing student learning: teacher change and the role of reflection. *Harv. Educ. Rev.* 72 (2), 230–253.
- Sabers, D.S., Cushing, K.S., Berliner, D.C., 1991. Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *Am. Educ. Res. J.* 28 (1), 63–88.
- Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A.-K., Yang, X., Kaiser, G., 2021. Mathematics teacher learning to notice: a systematic review of studies of video-based programs. *ZDM – Math. Educ.* 53 (1), 119–134.
- Schoenfeld, A.H., 2011. Noticing matters. A lot. Now what? In: Sherin, M.G., Jacobs, V.R., Philipp, R.A. (Eds.), *Mathematics Teacher Noticing. Seeing through Teachers' Eyes*. Routledge, pp. 223–238.
- Seidel, T., Stürmer, K., 2014. Modeling and measuring the structure of professional vision in preservice teachers. *Am. Educ. Res. J.* 51 (4), 739–771.
- Sherin, M.G., van Es, E.A., 2009. Effects of video club participation on teachers' professional vision. *J. Teach. Educ.* 60 (1), 20–37.
- Shulman, L.S., 1987. Knowledge and teaching: foundations of the new reform. *Harv. Educ. Res.* 57 (1), 1–22.
- Smith, B., 2012. Eye tracking as a measure of noticing: a study of explicit recasts in SCMC. *Lang. Learn. Technol.* 16 (3), 53–81.
- Stahnke, R., Schueler, S., Roesken-Winter, B., 2016. Teachers' perception, interpretation, and decision-making: a systematic review of empirical mathematics education research. *ZDM – Math. Educ.* 48 (1–2), 1–27.
- Star, J.R., Strickland, S.K., 2008. Learning to observe: using video to improve preservice mathematics teachers' ability to notice. *J. Math. Teach. Educ.* 11 (2), 107–125.
- Stigler, J.W., Miller, K.F., 2018. Expertise and expert performance in teaching. In: Ericsson, A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, pp. 431–452.
- van Es, E.A., Sherin, M.G., 2002. Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *J. Technol. Teach. Educ.* 10 (4), 571–596.
- Wolff, C.E., Jarodzka, H., van den Bogert, N., Boshuizen, H.P., 2016. Teacher vision: expert and novice teachers' perception of problematic classroom management scenes. *Instr. Sci.* 44 (3), 243–265.
- Yang, X., Kaiser, G., König, J., Blömeke, S., 2020. Relationship between pre-service mathematics teachers' knowledge, beliefs and instructional practices in China. *ZDM – Math. Educ.* 52 (2), 29–42.

Teacher recruitment: a competency-based approach

Robert M. Klassen^a, Jade V. Rushby^b, and Tracy L. Durksen^b, ^a Department of Education, University of York, Heslington, United Kingdom; and ^b School of Education, University of New South Wales, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	329
Two sides of teacher recruitment	329
Four opportunities to build the teaching workforce	330
Teacher attraction	330
Teacher selection	331
Person-vocation fit and competency-based recruitment	331
Examples of competency-based recruitment outside of education	331
A competency-based approach for teacher recruitment	332
Testing a person-environment fit model with Canadian teachers	332
Attracting prospective teachers using a competency-based approach	333
Selecting prospective teachers using a competency-based approach	334
Situational judgment tests	334
Situational judgment tests for teacher selection	334
Multiple mini-interviews for teacher selection	335
Implementing a competency-based approach for selecting teachers	336
Conclusion and next steps	336
Final thoughts	337
References	337

Introduction

Teachers play a key role in fostering positive student outcomes, but does it matter *who* is recruited into the teaching workforce? And if the answer is *yes, it does matter*, are there desirable personal characteristics that can be identified and evaluated in the recruitment process? Evaluating prospective teachers' *competencies*—the academic and non-academic knowledge, skills, abilities, and values—presents an ongoing challenge and is not without controversy. First, there is a multitude of personal characteristics and contextual factors to consider, and deciding which competencies are (a) most important, and (b) reliably evaluated is difficult. Second, recruitment decisions involving people who are at the beginning of their careers will always prove challenging with a large degree of “fuzziness.” Third, there are multiple definitions of “success” or “effectiveness” for teachers, and making predictions during recruitment about success in initial teacher education (ITE) may involve different factors than making predictions about success at Year 1 or Year 5 of professional practice. Teacher recruitment is a complex and multivariate problem, but building on a foundation of theory- and research-supported competencies increases the likelihood of building a strong and effective teacher workforce. This article describes a “competency-based approach” for two facets of teacher recruitment: *attracting* applicants to consider entering the profession and *selecting* the best possible candidates from a pool of applicants.

Two sides of teacher recruitment

There are two sides to teacher recruitment—attraction and selection—and the relative weighting between these two recruitment activities varies according to national and regional contexts. In some settings (e.g., Singapore and Finland), there is not a very strong need to attract more applicants, and so the recruitment emphasis falls heavily on the side of selection. In these “selection-focused” settings, entrance into ITE involves a rigorous selection process designed to identify the most promising candidates from a deep pool of qualified candidates, and less explicit attention is paid to attracting applicants to consider the profession. In “attraction-focused” settings (e.g., some Asian and European countries, some parts of the United States, Australia, the United Kingdom), the applicant pool for ITE and jobs may not be very deep, and the recruitment emphasis is weighted toward attracting more high-quality applicants. In these settings, many of the policy and practice discussions focus on ways to attract high-quality candidates into considering the profession, with recruitment strategies centering on short-term solutions involving financial incentives that may not recruit candidates who are well-suited to the profession. Whether the recruitment focus falls on attraction or selection, a competency-based teacher recruitment program is built on the belief that candidates come with a wide range of background experiences and personal competencies, and that these individual differences play a role in predicting future teaching behaviors. Although some of these individual differences are fleeting (e.g., mood, interest), others are more enduring (personality and aptitude), and are at the heart of recruitment activities.

Four opportunities to build the teaching workforce

There are four key opportunities to build a strong teacher workforce (see Fig. 1): during recruitment (attraction), selection, ITE, and professional practice. Most efforts to improve the teacher workforce have focused on improving ITE (Greenberg et al., 2015) and in-service professional development (e.g., Gore, 2021), with less systematic attention paid to improving the way that we attract and select individuals into the profession. Delivering high-quality development programs during ITE and professional practice is important to build a strong workforce, but can be expensive and sometimes ineffective, with teachers' professional development especially suffering from poor returns on investment (Greenberg et al., 2015). Strengthening the teacher workforce during ITE and in professional practice is critical and can bring lasting improvements. However, building better attraction and selection methods represents a "quick win" in building the teacher workforce, with system-wide benefits attained at a relatively modest cost. Each of these workforce-building opportunities can be built on a foundation of key competencies, but attraction and selection systems especially benefit from being built with an explicit link to the academic and non-academic competencies that form the bedrock of successful teaching (see Fig. 1).

Teacher attraction

Decisions about entering the teaching profession are influenced by candidates' perceptions of the *utility values* associated with embarking on a teaching career; namely *personal utility values* and *social utility values* (see Fig. 2). Personal utility refers to beliefs about the personal benefits of the profession (e.g., stability, progression opportunities, working conditions), or in some settings, consideration of financial incentives such as grants or bursaries to enter training. Teacher attraction strategies are frequently based on highlighting the personal utility associated with the profession. See et al. (2020) reviewed teacher attraction methods across 20 studies and found that most teacher attraction strategies depended on offering financial incentives, but warned that bursary schemes and other financial incentives tended not to have lasting effects. They found no strong evidence that other personal utility approaches, for example, offering better mentoring, induction, or development programs, made a lasting impact on attracting and retaining new teachers.

Another influence on career decision-making is perceptions of the social utility of teaching; that is, perceptions of how teaching makes a social contribution through improving the lives of children and young people. Teacher attraction advertising campaigns are frequently built on heightening perceptions of the social utility associated with the profession. Klassen et al. (2021a,b) examined

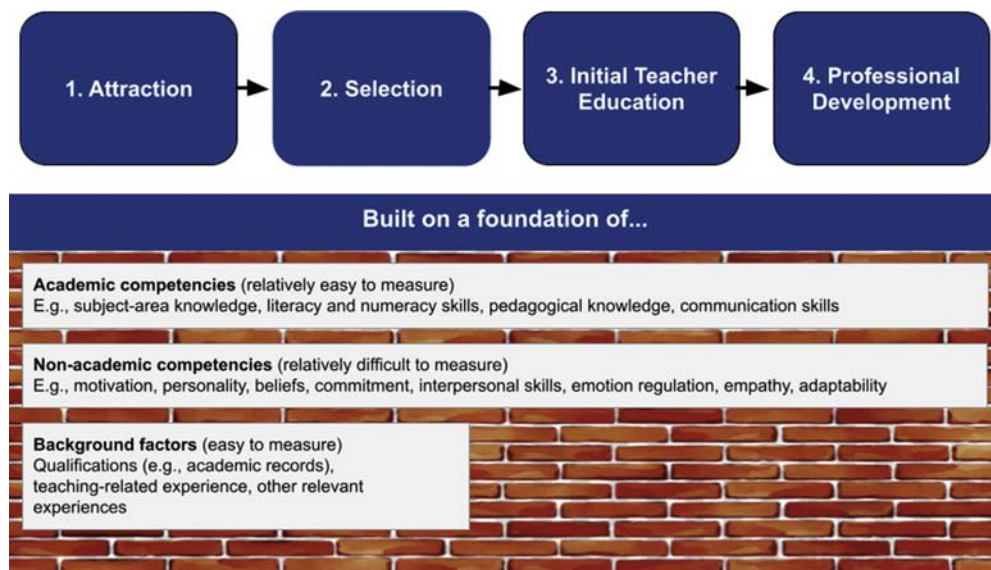


Fig. 1 Four opportunities to strengthen the teacher workforce.

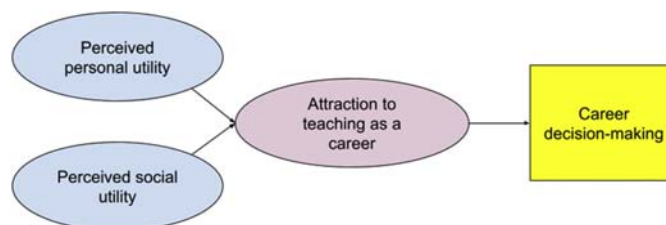


Fig. 2 Career decision-making model.

recruitment messages in public-facing media campaigns for two major education players in the UK (the Department for Education [DfE] and Teach First) and contrasted these campaigns with those from the National Health Service (NHS). They found that DfE advertising campaigns (e.g., *Every lesson shapes a life*) emphasized the social utility of a teaching career, whereas the Teach First advertisements emphasized both personal utility (e.g., *in terms of career progression ... it's been a superb choice*), and social utility (e.g., *Let's give opportunities to kids who wouldn't normally have these opportunities*). In contrast, the NHS recruitment campaigns were built on a “values-based approach” in which candidates were urged to consider the fit between their own values and those of the organization. Basing teacher recruitment campaigns on personal and social utility can help attract people into considering the profession, and may encourage some candidates into moving toward the selection phase of recruitment.

Teacher selection

Current teacher selection methods have been described as “ad hoc” (Goldhaber et al., 2014), with weak empirical and theoretical foundations underpinning their use (Klassen and Kim, 2019). This lack of research on analyzing the effectiveness of selection methods is surprising considering the importance of the interest in building high-quality education systems worldwide, and in light of the considerable research on selection in other fields. When selecting candidates for ITE or jobs, most education systems evaluate one or more academic competencies (e.g., subject knowledge and/or basic skills), non-academic competencies (such as motivation and resilience), and background factors such as qualifications and teaching-related experiences. Assessing academic competencies and background factors is relatively straightforward: records of academic competencies such as academic transcripts are readily available, and if further data is desired, standardized subject knowledge and literacy/numeracy tests are readily available.

Programs face a greater challenge when evaluating the non-academic competencies that they deem to be important for success as novice teachers. Conducting reliable and valid assessments of empathy, interpersonal relationships, and motivation is more difficult to carry out during screening or selection than assessing academic skills. Teacher selection methods may pay lip service to teaching standards as behavioral definitions reflective of multiple competencies and values (e.g., Casey and Childs, 2017), but the methods chosen for selection may not reliably assess these competencies. In any field, selection methods require regular evaluation of their reliability (consistency over time), validity (evidence of predictive utility), and fairness for all applicants, regardless of age, gender, ethnicity, sexuality, and socioeconomic status. An evidence-supported competency-based approach to teacher recruitment first identifies a framework of desired competencies based on theory and research, and then develops reliable, valid, and unbiased methods to assess these competencies.

Person-vocation fit and competency-based recruitment

A competency-based approach to teacher recruitment is built on person-environment fit (PE fit) theory which holds that the fit between people and their environments is a key determinant of human flourishing (Darrow and Behrend, 2017). Sub-categories of PE fit include person-vocation fit (PV fit), defined as the fit between an individual's personal characteristics (e.g., interests, abilities) and the demands of a career path, and person-organization fit (PO fit), defined as the fit between an individual and a particular organization (e.g., school or education system). Research suggests a strong relationship between PV fit and applicants' beliefs about a job before considering an occupation, and between PV fit and on-the-job behaviors related to job performance and at-work attitudes (Vogel and Feldman, 2009). A meta-analysis conducted in 2012 showed that perceived fit was a strong predictor of job attraction through multiple stages of the recruitment process (Uggerslev et al., 2012). Another meta-analysis (Chapman et al., 2005) showed that although job-related and organizational characteristics were important factors in applicants' career decisions, perceptions of fit were the strongest predictors of career decisions. The fit between individual and vocation serves as the foundation for competency-based recruitment strategies in a range of professional disciplines.

Examples of competency-based recruitment outside of education

Recruitment research and practices have seen a shift toward competency-based (and values-based) approaches in health and business fields in the last decade. The National Health Service (NHS) in the UK is a prime example of a health organization that has shaped its recruitment, selection, training, and professional development activities around a core of key competencies, or “values” that underpin its service. A *values-based* approach to recruitment involves recruiting and selecting individuals for roles based on the fit between organizational values—variously described as “enduring beliefs” or “motivational goals”—and an individual's professional values (what we term competencies in this article), such as integrity, compassion, and commitment. The NHS builds its recruitment activities, including attraction, screening, and selection, on six core values (*Respect and dignity, Commitment to quality of care, Compassion, Improving lives, Working together for patients, Everyone counts*). Following the recruitment process, the six values underpin post-selection activities such as education, training, and establishing organizational culture (NHS, 2021). Other organizations, including the College of Policing in the UK (2018), have developed a combined competency and values-based framework that underpins their organization, with six competencies (e.g., innovative, emotionally aware, collaborative) encompassing four values (e.g., transparency and integrity). These values and competencies are further defined by their operationalization as behaviors, that is, the behavioral manifestations in policing practice. Similar competency-based (or values-based) approaches are embraced in specific health-related fields (e.g., midwifery; Power and Clews, 2015), and also in large corporations (e.g., KPMG, 2021). The identified competencies across occupational groups are often widely similar (e.g., integrity, empathy), and can be used as a point of

reflection when building competency-based attraction and selection programs in education. Competency-based approaches from other disciplines can serve as models for teacher recruitment strategies.

A competency-based approach for teacher recruitment

Although most research on competency-based recruitment has been conducted in settings outside of education, a handful of studies with teachers or prospective teachers have examined how perceived fit with desired competencies influences recruitment outcomes. De Cooman et al. (2009) found that teachers who perceived a fit between their personal values and the perceived values of their schools were more likely to stay in their jobs. Player et al. (2017) found that the person-vocation fit was predictive of retention: teachers with perceived fit one standard deviation above average were 43% less likely to leave the profession. From a theoretical perspective, these perceptions of fit mediate the relationship between attraction to the profession and the decision-making process. Fig. 3 shows graphically how perceived personal and social utility influence attraction, and how the relationship between attraction to the profession and career decision-making is mediated by perceptions of fit.

When a candidate considers the fit between their own competencies and those salient to teaching, they seek out information from multiple sources: their past experience in schools, advice from their social network, and from recruitment messages delivered by education organizations (e.g., ITE programs and national education bodies). Some ITE programs explicitly target a set of non-academic competencies that they want to see reflected in their preservice and practicing teachers. For example, one program in the UK lists several competencies that applicants should demonstrate, or show the potential to demonstrate: *humility, respect and empathy; interaction; understanding and motivation; leadership; planning and organizing; problem-solving; resilience; and self-evaluation*. In this program, the targeted competencies are evaluated explicitly in selection activities that include, three “motivation statements,” individual interviews, a group activity, and a self-evaluation exercise, each of which targets a particular non-academic competency. Another ITE program in the UK is focused on developing new teachers who are *creative, proactive, and reflective*, with selection activities that are focused on evaluating these competencies. In Australia, several ITE programs have adopted a set of 10 academic and non-academic competencies, including academic competencies such as *ethical behavior, behavioral stability, reflective skills, and relational skills*. Identifying a set of key competencies is the first step in developing a competency-based selection process, with selection activities focused on evaluating these competencies.

It is one thing for education organizations to list a set of desired competencies in their recruitment process, but explicit consideration of these competencies and the methods used to evaluate these competencies is relatively rare. More typical is the example of a large and high-ranking ITE program on the west coast of the United States, where non-academic competencies are assessed through a “goal statement” that is designed to assess “*a passion for teaching ... and an interest in issues of social justice*” with little attention paid to the reliability and validity of the methods. This is ineffective since personal statements are notoriously unreliable predictors of future performance (Patterson et al., 2016). The explicitness of the target competencies is not always made apparent to applicants, and the methods and scoring approach used to evaluate the competencies are not always very specific. Recruiting and selecting new teachers is a “predictive exercise” where education organizations are attempting to attract and identify candidates who will experience success as teachers. Building a framework of the non-academic competencies that are key for this success can underpin the recruitment, selection, and development phases of a teaching career.

Testing a person-environment fit model with Canadian teachers

An education-relevant person-environment fit model was recently tested in a sample of 1086 practicing Canadian teachers (Wang and Klassen, 2021). The researchers analyzed data from a two-wave longitudinal study that tested how teachers’ utility values were related to their quitting intentions. Personal utility values were operationalized as *prestige, autonomy, and job security*, while social utility values were operationalized as *altruism, student relationships, and co-worker relationships*. Results from structural equation modeling (see Fig. 4) showed that teachers who valued autonomy, prestige, and job security (personal utility values) were more likely to express intentions to leave the profession. Teachers who put more value on their relationships with students and peers, and who valued the altruistic values of the job (social utility) were more likely to express the intention to stay in their current position and in the profession overall. Teachers’ perceptions of their fit with their organization (i.e., school) mediated the relations between social utility values and attrition intentions, but did not have the same mediating effect for the relations between personal utility and quitting intentions; i.e., teachers’ valuing of prestige, autonomy, and job security had a direct effect on their intentions to

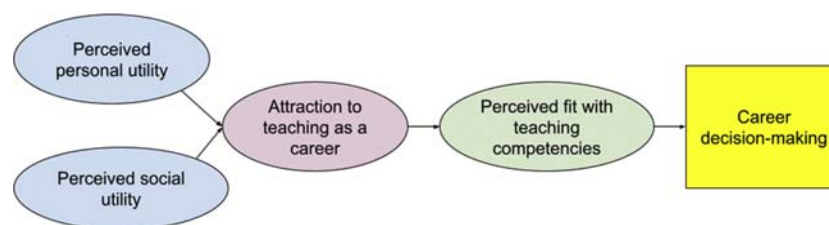


Fig. 3 A mediated model of career decision-making.

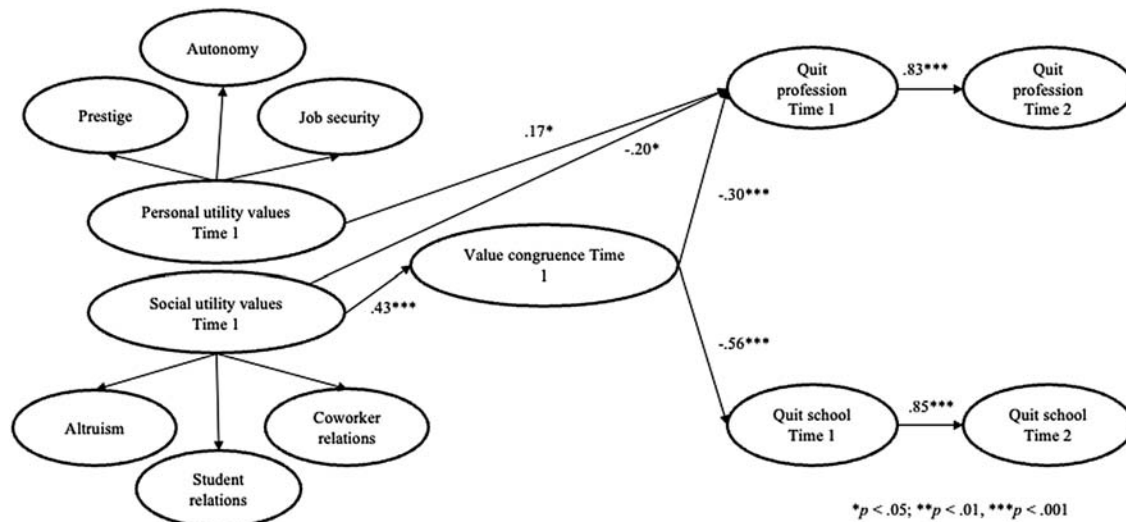


Fig. 4 Utility values, value congruence, and quitting intentions.

leave the profession. Of the three personal utility values, a perceived lack of autonomy support had the strongest effects on intention to leave the profession, suggesting that organizations may wish to consider how impositions on teachers' professional autonomy have serious consequences for teacher retention.

Attracting prospective teachers using a competency-based approach

Although teacher attraction strategies based on personal and social utility may be effective at least in the short term, they can also attract applicants who have an unrealistic view of the profession, resulting in high attrition rates due to eventual feelings of mis-fit (Baur et al., 2014). For example, a decision to enter the profession based on bursaries for shortage subjects (i.e., based on personal utility) may result in high attrition rates once bursary payments are no longer in effect (See et al., 2020). A decision to enter the profession based on altruistic reasons (*I'll really be making a difference in children's lives!*) may spur some candidates to apply for training, but without consideration of personal fit with the competencies needed for success, these social utility values may lose their motivational appeal when the "reality shock" of teaching is felt in the challenges typically experienced in the 1st year. Teacher recruitment (attraction) strategies that capitalize on personal and social utility models may be effective in the short term to inspire candidates to enter the profession, but building a recruitment strategy that also includes consideration of person-vocation fit and the competencies needed for teaching promises a more robust, theoretically-sound, and sustainable recruitment program.

A recent intervention study (Klassen et al., 2021a,b) was conducted by attracting university students in STEM (Science, Technology, Engineering, and Mathematics) programs and inviting them to consider teaching as a profession with *realistic job previews*, or RJPs built on a person-vocation fit foundation. RJPs are a recruitment method where potential applicants engage with realistic workplace scenarios, with the over-arching question, *How well do you fit with this job?* A fit message provides participants with an indication of how well their responses fit with expected responses from those already in the field. Research using an RJP approach has been conducted for more than five decades, with results showing increases in better integration once in the job, along with lower attrition rates and better workplace outcomes (Baur et al., 2014). RJPs can be delivered through a variety of methods, including video and text, and have been used for recruitment purposes in a range of fields, but not previously in education.

Klassen and colleagues recruited UK undergraduates in STEM-related fields to participate in the study. The brief RJP intervention was delivered online (to participants' phone, tablets, or computers) using four animated classroom scenarios. Participants (a) viewed the animated video, (b) rated the appropriateness of three possible responses, (c) wrote a reflective rationale for their response, (d) received real-time feedback on the alignment between their ratings and those of experienced teachers, and (e) were provided with a rationale from experienced teachers for the appropriate responses to the scenario. The study results showed that the intervention had a significant positive effect on participants' interest in pursuing a teaching career ($r = 0.26$); the results were robust after controlling for prior career intentions. Analysis of subsequent individual interview data examined how the intervention was perceived by participants: *The activity showed that I had a similar ideology as a teacher so made me think that maybe I would be suitable; it really helped me think about how teachers think.* The study showed how a brief, cost-effective and scalable intervention based on person-vocation fit could supplement current teacher attraction strategies based on personal and social utility values.

Selecting prospective teachers using a competency-based approach

Teacher selection involves making predictions about how candidates will fare in the short-term (i.e., in training), and longer-term (i.e., in professional practice). A robust and defensible process for selection is needed when the number of applicants is greater than the number of places (for training or for jobs), when there is a need to screen out unsuitable applicants, and when there is a need to identify applicant strengths and weaknesses for future development. Although competency-based selection methods in other fields have been the subject of in-depth research attention in recent years, these methods have infrequently been applied to teacher selection. Most current teacher selection methods—academic records, personal statements, group interviews, reference letters, lightly-structured interviews—have been shown to lack predictive power and to be subject to biases against certain groups of candidates (Patterson et al., 2016). Two competency-based methods—situational judgment tests (SJTs) and multiple mini-interviews (MMIs)—have been successfully implemented in fields outside of education for the last two decades and have recently been developed and tested in education settings.

Situational judgment tests

Situational judgment tests (SJTs) have become increasingly popular in the last two decades because, in comparison to other selection tools, they show high levels of predictive validity and can be built on a set of desired competencies. SJTs consist of a stem scenario that depicts a challenging workplace situation, and a series of response options outlining possible courses of action, and can be delivered in text-only or video + text formats. Early versions of SJTs were used to evaluate the situational judgment of officer candidates in World War II, and subsequently have been used for selection into medical training, the civil service, and large corporations. The theories underpinning SJTs include *implicit trait policy* (Whetzel and McDaniel, 2009), which refers to an individual's implicit beliefs about the advisability of expressing specific personality traits and motivations in particular situations. SJTs are typically designed to capture information about how particular competencies (e.g., empathy, adaptability, commitment) are expressed in particular contexts. Other tools used for selection may be “context-free” (e.g., tests of personality or emotional intelligence) but it is the interaction between context (e.g., a classroom situation in a teacher selection SJT) and the targeted competency that separates SJTs from other recruitment tools.

Recent research and meta-analyses show that SJTs administered as selection tools at the beginning of training programs can be robust predictors of job performance (e.g., Patterson et al., 2013) and are less prone to group biases that plague other selection methods. The development of SJTs for selecting prospective teachers promises the advantages of clear links to key competencies, efficiency (i.e., modest time investment for implementation after development, and especially helpful for screening large numbers of applicants), and effectiveness, with SJTs shown to be one of the most predictive tools available for selection (Patterson et al., 2016). However, SJTs are relatively new in teacher education circles, possibly due to a historical reluctance to move away from more “intuitive” (but largely ineffective) selection approaches (e.g., Davies et al., 2016).

Situational judgment tests for teacher selection

SJTs have not often been implemented for selection into teacher education or for entry into the profession, but the last few years have seen an increase in their use (see Fig. 5 for a still shot from a video SJT from the Teacher Selection Project). Klassen and Kim



You are teaching a 9th-grade science lesson when you notice one of your students, Jake, using his mobile phone. Your school has a ‘no phone in class’ policy, and you have already warned Jake about using it in your classroom. Jake seems to have forgotten your warning, and is now openly using the phone while you are teaching.

	Inappropriate	Somewhat inappropriate	Somewhat appropriate	Appropriate
Go to get assistance from a senior member of staff	☐	☐	☐	☐
Ignore that Mark is using his phone and continue with the lesson	☐	☐	☐	☐
Ask Mark to put the phone away again, but more sternly	☐	☐	☐	☐
Give school sanctions for having a mobile phone in class and for refusing the teacher's instruction	☐	☐	☐	☐

How appropriate are the following responses?

Fig. 5 Example of a video-based SJT from teacher selection project.

(2019, 2021) found that it was rare for teacher education programs to be built on methods with a published evidence base (i.e., using SJTs, MMIs, or other evidence-based methods) but research on using SJTs for selecting teachers is gradually emerging, with most studies in the last decade coming from the Teacher Selection Project group in the UK (<https://www.teacherselect.org/>), and some from the Teacher Capability Assessment Tool in Australia (e.g., Bowles et al., 2014). Results from research in the UK (Klassen et al., 2014) showed that SJTs administered to primary and secondary teaching applicants were generally well-received, with applicants for both primary and secondary programs expressing favorable opinions of the test. Two additional studies showed that SJTs administered at the point of selection successfully predicted performance in the second stage of selection (Klassen et al., 2020) and up to 6 months into teacher training (e.g., Klassen and Rushby, 2019).

Linking desired competencies to the content of teacher selection SJTs begins at the very beginning of the development process. Although SJTs have been criticized for their lack of clearly defined factor structure, largely due to the verisimilitude and complexity of the scenarios that make up the heart of the method (i.e., multiple constructs are represented in the “real-life” scenarios), key non-academic competencies form the foundation of the method. Multiple steps are involved in developing SJTs, including selecting and defining the key non-academic competencies on which to build the test. Klassen et al. (2017) described four steps in the development of an integrated (inductive and deductive) preservice teacher SJT:

1. **Teacher observation.** A job analysis is conducted through shadowing an early career and mid-career teacher in order to ground the research team in a shared understanding of the activities and tasks involved in a teacher’s day.
2. **Focus groups and individual interviews with stakeholders.** Step 2 involves an inductive approach to identify the key competencies needed for successful preservice teaching. Focus groups with experienced teachers and teacher educators are used to generate and discuss key competencies (e.g., *What are the individual characteristics you see in successful preservice teachers?*).
3. **Research review and creation of target competencies.** In this phase relevant research outlining key competencies (i.e., exploration of research examining non-academic predictors of teacher effectiveness) are added to the mix of competencies identified in the previous step to form a comprehensive list of target competencies.
4. **Review of target competencies.** In this phase, the inductively-derived (from focus groups) and deductively-derived (i.e., from research review) competencies are integrated to form an initial set of competencies used as a foundation for building SJT content.

Following the steps of SJT development, Klassen et al. (2017) identified three clusters of competencies: *empathy and communication*, *organization and planning*, and *resilience and adaptability*. Each of the SJT items, formed through a critical incident process, was built on one of these three competency clusters. After development, the teacher selection SJT was pilot-tested on incumbents and after revision, used for “live selection” as a screening tool (Klassen et al., 2020). In general, SJTs are most often used for large-scale screening rather than for more intensive selection purposes, for which approaches such as *multiple mini-interviews* (MMIs) are more appropriate.

Multiple mini-interviews for teacher selection

Traditional interviews (i.e., low-to medium-structured format with a high degree of interviewer flexibility) are one of the most used formats for selecting candidates into teacher training in some settings (Davies et al., 2016), but are generally labor-intensive and poor predictors of subsequent performance (Dana et al., 2013). One of the more recent developments in recruitment research is the development of multiple mini-interviews (MMIs), a method that has been widely used for selection into medical education (e.g., Eva et al., 2019) and more recently for selection into teacher education (Metsäpelto et al., 2020). MMIs used for selection into teacher education can be built on core competencies and are shown to be more reliable, valid, and fair than interview methods used in the past.

MMIs were developed by Eva et al. (2004) in an attempt to improve the fairness of the interview experience for applicants to medical school. The distinguishing features of MMIs (compared to traditional interviews) are that they use independent, highly-structured stations with scoring protocols that are also highly structured. MMIs, by definition, include multiple stations (from 3 to 10), each targeting a particular non-academic competency. Each station is independent and interviewers do not have knowledge of the interviewees’ profiles or performance in other stations, and the station content and scoring protocols are highly structured with standardized scoring schemes. MMIs are typically co-developed by experts with a diverse group of program staff who first identify the key target competencies, and then develop and test stations designed to assess these competencies.

Similar to the SJT development process, there are multiple steps involved in the development of a competency-based MMI for selecting teacher candidates:

1. **Identify key competencies.** This first step involves a series of focus groups and individual interviews with reference to relevant research (i.e., on competencies related to success in teacher education programs), policy documents, and program goals all under consideration. Ideally, a set of 4–8 key competencies are identified for consideration (e.g., empathy, professionalism, communication, diversity and social justice, mindset, and teacher beliefs).
2. **Assessment of interview logistics.** In this step, the logistics of the MMIs are considered, with questions about the number of applicants and available resources needing consideration, as well as the number of stations that can be accommodated (typically 4–8 stations in total).
3. **Station-writing workshops.** Each station is created to target one of the key non-academic competencies. Activities may include card-sorting, discussions with picture stimuli, simulated classroom situations, and structured interviews.

4. **Scoring guide development.** A systematic and detailed scoring rubric with positive and negative indicators is developed for each station task along with a summary of each station task, a list of specific questions to be asked and additional prompt questions, and a scoring rubric
5. **Pilot testing and review.** In this step, the MMI is pilot-tested with a small group of incumbent students with changes based on feedback from students and examiners regarding content and logistics.

MMIs provide an in-depth, competency-based approach to in-depth interviews, and as such, are an ideal complement to SJTs for larger-scale screening. The work conducted in Finland (Metsäpelto et al., 2020) and in the UK by the Teacher Selection Project provides the first examples of a competency-based and validated approach to in-depth teacher selection approaches. Using evidence-supported methods such as SJTs and MMIs represents a positive change for education organizations, but the implementation of competency-based recruitment methods takes careful planning and testing of the methods to be implemented.

Implementing a competency-based approach for selecting teachers

There are two key questions to answer when developing a competency-based recruitment and selection program (see Fig. 6). The first question posed by an education organization is, *What are the key competencies that we should target in our recruitment program?* The question is not trivial, because the competencies considered at this stage will potentially influence the attraction, selection, and development activities in the program. There is no single set of competencies that fits all organizations: an ITE program attracting and developing teachers for rural and remote Australian schools may choose to focus on competencies that reflect the unique challenges of the target context (Durksen and Klassen, 2018). These key competencies can be chosen through a review of research and theory on teacher effectiveness (i.e., top-down or deductively), and through a consideration of the particular foci and themes of the organization (i.e., bottom-up or inductively). The choice of key competencies can be developed using an integrated deductive and inductive approach, where research- and theory-supported competencies are integrated with the competencies developed through an inductive data-gathering phase (e.g., see Klassen et al., 2020).

The second question to consider during the implementation phase is how to develop reliable, valid, and unbiased methods to evaluate these competencies. The second question posed by the organization is *What selection methods should be considered?* and, *What is the research evidence for their effectiveness?* Constraints related to resources (i.e., time, space, personnel, cost) play a part in determining the viability of selection methods, but *status quo* approaches (e.g., depending solely on academic records, personal statements, and conventional interviews) have been shown to be ineffective in teacher selection (Klassen and Kim, 2019, 2021). The key factors to consider by an organization in implementation are first, working together to establish a set of agreed-upon competencies, and second, considering the best (reliable, valid, and viable) ways to evaluate these competencies. Recent research in medical education shows that selection tools such as SJTs and MMIs are preferable to more traditional methods (Patterson et al., 2016). Including the assessment of a range of competencies using multiple methods at each point in the recruitment process increases the likelihood of predicting future success in any professional field (e.g., Lievens and Sackett, 2017). Implementing tested and validated recruitment methods provides an effective and efficient way to improve the teacher workforce; building research-supported recruitment programs can provide lasting benefits to organizations and education systems.

Conclusion and next steps

In this article, we have considered how a competency-based approach to teacher recruitment—attraction and selection—can lead to improvements in the teacher workforce. The recruitment challenges facing educational organizations are real: recruiting the best

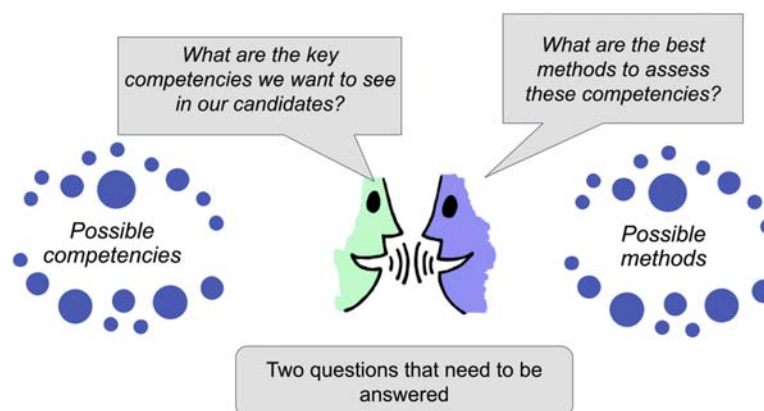


Fig. 6 Choosing key competencies and evidence-supported methods.

possible teachers represents a complex multivariate problem, but recent advances in research and theory can contribute to solving recruitment problems. For educational organizations, the adoption of a competency-based approach coupled with the use of recent methodological innovation can provide a “quick win” to strengthening the teacher workforce, with the additional advantage of providing a reliable and valid approach to recruitment and selection.

We suggest that the next steps for research on competency-based approaches to teacher recruitment will focus on:

1. **Building a better understanding of key competencies.** Cross-cultural work exploring the universality and cultural specificity of key competencies associated with effective teaching practices. Limited work has been conducted in this area, and although some competencies are universal—empathy and resilience, for example—the salience of some competencies varies according to context (e.g., Klassen et al., 2018). From a research standpoint, understanding how individual characteristics might vary across contexts is of real interest; from an applied standpoint, understanding which competencies should be targeted during recruitment, and how these competencies might change over time and place is important for the development of recruitment activities.
2. **Development and testing of innovative recruitment methods.** Although many educational organizations moved to online recruitment during the COVID-19 pandemic, potential technological advances in recruitment go far beyond online delivery. One potential area for growth is *gamified assessments* (serious games) that incorporate gaming elements in non-game contexts (e.g., Woods et al., 2020). For example, SJTs and RJP's can be “gamified” by placing prospective teachers in a virtual school classroom using virtual reality technologies, and work on “branched” scenarios can provide a more interactive and tailored selection process (e.g., Reddock et al., 2020). Current work on branched scenarios for teacher recruitment and development is taking place in Australia by researcher Jade Rushby and colleagues at UNSW. One immersive virtual teaching environment is SimLab, developed in the US (<https://kognito.com/>) and recently piloted for teacher education at Murdoch University in Australia (Ledger and Fischetti, 2020). In the last few years, advances in virtual reality and simulated teaching environments have spurred on new ways to select prospective teachers (e.g., Ke and Xu, 2020), and we expect teacher recruitment methods to be influenced by these innovations.
3. **Improving the recruitment (attraction and selection) of school leaders.** An important, but largely neglected, area of growth is the recruitment of school leaders. We see a worldwide shortage of high-quality school leaders and current recruitment processes can be arbitrary and opaque (Yang et al., 2021). Lessons learned from teacher recruitment (e.g., building from a foundation of key competencies, using evidence-informed methods, measuring the validity of methods) can improve the approaches used to attract and select prospective school leaders.

Final thoughts

Recruiting the next generation of teachers presents an opportunity to positively influence the educational outcomes of children and adolescents around the world. A carefully constructed, tested, and refined teacher recruitment program identifies key teacher competencies and uses the best possible methods to recruit and retain outstanding teachers. In this article, we have argued that working to build the best possible recruitment programs is a worthwhile endeavor that has often been neglected in education. Efforts to build a rigorous and robust recruitment program will pay off with a stronger teacher workforce, a higher quality education system, and a more productive and prosperous society.

References

- Baur, J.E., Buckley, M.R., Bagdasarov, Z., Dharmasiri, A.S., 2014. A Historical Approach to Realistic Job Previews: An Exploration Into Their Origins, Evolution.
- Bowles, T., Hattie, J., Dinham, S., Scull, J., Clinton, J., 2014. Proposing a comprehensive model for identifying teaching candidates. *Aust. Educ. Res.* 41, 365–380.
- Casey, C., Childs, R.A., 2017. Teacher education program admission criteria and what beginning teachers need to know to be successful teachers. *Can. J. Educ. Adm. Pol.* 67, 1–24.
- Chapman, D.S., Uggerslev, K.L., Carroll, S.A., Piasentin, K.A., Jones, D.A., 2005. Applicant attraction to organizations and job choice: a meta-analytic review of the correlates of recruiting outcomes. *J. Appl. Psychol.* 90, 928–944.
- College of Policing, 2018. Values-Based Recruitment and Selection. Retrieved from: https://paas-s3-broker-prod-lon-6453d964-1d1a-432a-9260-5e0ba7d2fc51.s3.eu-west-2.amazonaws.com/s3fs-public/2020-11/Values-Based_Recruitment_Guidance-1.pdf.
- Dana, J., Dawes, R., Peterson, N., 2013. Belief in the unstructured interview: the persistence of an illusion. *Judgm. Decis. Mak.* 8, 512–520.
- Darrow, J.B., Behrend, T.S., 2017. Person-environment fit is a formative construct. *J. Vocat. Behav.* 103, 117–131.
- Davies, P., Connolly, M., Nelson, J., Hulme, M., Kirkman, J., Greenway, C., 2016. “Letting the right one in”: provider contexts for recruitment to initial teacher education in the United Kingdom. *Teach. Teach. Educ.* 60, 291–302.
- De Cooman, R., De Gieter, S., Pepermans, R., Hermans, S., Du Bois, C., Caers, R., Jegers, M., 2009. Person–organization fit: testing socialization and attraction–selection–attrition hypotheses. *J. Vocat. Behav.* 74 (1), 102–107.
- Durksen, T.L., Klassen, R.M., 2018. The development of a Situational Judgment Test (SJT) of personal attributes for quality teaching in rural and remote Australia. *Aust. Educ. Res.* 45, 255–276.
- Eva, K.W., Rosenfeld, J., Reiter, H.I., Norman, G.R., 2004. An admissions OSCE: the multiple mini-interview. *Med. Educ.* 38, 314–326.
- Eva, K.W., Macala, C., Fleming, B., 2019. Twelve tips for constructing a multiple mini-interview. *Med. Teach.* 41, 510–516.
- Goldhaber, D., Grout, C., Huntington-Klein, N., 2014. It's Selective But Is It Effective? Exploring the Predictive Validity of Teacher Selection Tools. CEDR Policy Brief 2014-9. University of Washington, Seattle, WA.
- Gore, J.M., 2021. The quest for better teaching. *Oxf. Rev. Educ.* 47, 45–60.
- Greenberg, J., Walsh, K., McKee, A., 2015. 2014 Teacher Prep Review: A Review of the Nation's Teacher Preparation Programs. National Council on Teacher Quality. Retrieved from: <https://www.nctq.org/publications/Teacher-Prep-Review-2014-Report>.

- Ke, F., Xu, X., 2020. Virtual reality simulation-based learning of teaching with alternative perspectives taking. *Br. J. Educ. Technol.* 51, 2544–2557.
- Klassen, R.M., Kim, L.E., 2019. Selecting teachers and prospective teachers: a meta-analysis. *Educ. Res. Rev.* 26, 32–51.
- Klassen, R.M., Kim, L.E., 2021. *Teacher Selection: Evidence-Based Practices*. Springer. ISBN 978-3-030-76186-8.
- Klassen, R.M., Rushby, J.V., 2019. Can We Predict Preservice Teachers' Performance in Teaching Placements? The Validity of ITE Selection Methods. TSP Working Paper. <https://bit.ly/2SeQkVv>.
- Klassen, R.M., Durksen, T.L., Rowett, E., Patterson, F., 2014. Applicant reactions to a situational judgment test used for selection into initial teacher training. *Int. J. Educ. Psychol.* 3, 104–125. <https://doi.org/10.4471/ijep.2014.07>.
- Klassen, R.M., Durksen, T.L., Kim, L.E., Patterson, F., Rowett, E., Warwick, J., Warwick, P., Wolpert, M.A., 2017. Developing a proof-of-concept selection test for entry into primary teacher education programs. *Int. J. Assess. Tools Educ.* 4, 96–114.
- Klassen, R.M., Durksen, T.L., Al Hashmi, W., Kim, L.E., Longden, K., Metsäpelto, R.-L., Poikkeus, A.M., Györi, J., 2018. National context and teacher characteristics: exploring the non-cognitive attributes of prospective teachers in four countries. *Teach. Teach. Educ.* 72, 64–74.
- Klassen, R.M., Kim, L.E., Rushby, J., Bardach, L., 2020. Can we improve how we screen applicants for initial teacher education? *Teach. Teach. Educ.* 87, 102949.
- Klassen, R., Bardach, L., Rushby, J., Durksen, T.L., 2021a. Examining teacher recruitment strategies in England. *J. Educ. Teach.* <https://doi.org/10.1080/02607476.2021.1876501>.
- Klassen, R.M., Granger, H., Bardach, L., 2021b. Attracting prospective STEM teachers using realistic job previews: a mixed methods study. *Eur. J. Teach. Educ.* <https://doi.org/10.1080/02619768.2021.1931110>.
- KPMG, 2021. Our Values. Retrieved from. <https://www.kpmgcareers.co.uk/experienced-professional/applying-to-kpmg/our-values/>.
- Ledger, S., Fischetti, J., 2020. Micro-teaching 2.0: technology as the classroom. *Australas. J. Educ. Technol.* 36, 37–54.
- Lievens, F., Sackett, P.R., 2017. The effects of predictor method factors on selection outcomes: a modular approach to personnel selection procedures. *J. Appl. Psychol.* 102, 43–66.
- Metsäpelto, R., Utriainen, J., Poikkeus, A., Muotka, J., Tolvanen, A., Warinowski, A., 2020. Multiple Mini Interview as a Selection Tool for Initial Teacher Education Admissions. Retrieved from. psyarxiv.com/ph5t8.
- NHS, 2021. Values-Based Recruitment. Retrieved from. <https://www.hee.nhs.uk/our-work/values-based-recruitment>.
- Patterson, F., Lievens, F., Kerrin, M., Munro, N., Irish, B., 2013. The predictive validity of selection for entry into postgraduate training in general practice: evidence from three longitudinal studies. *Br. J. Gen. Pract.* e734–e741. <https://doi.org/10.3399/bjgp13X674413>. November.
- Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousins, F., Cleland, J., 2016. How effective are selection methods in medical education? A systematic review. *Med. Educ.* 50, 36–60.
- Player, D., Youngs, P., Perrone, F., Grogan, E., 2017. How principal leadership and person-job fit are associated with teacher mobility and attrition. *Teach. Teach. Educ.* 67, 330–339.
- Power, A., Clews, C., 2015. Values-based recruitment and the NHS constitution: making sure student midwives meet the brief. *Br. J. Midwifery* 23, 818–820.
- Reddock, C.M., Auer, E.M., Landers, R.N., 2020. A theory of branched situational judgment tests and their applicant reactions. *J. Manag. Psychol.* 35 (4), 255–270.
- See, B.H., Morris, R., Gorard, S., El Soufi, N., 2020. What works in attracting and retaining teachers in challenging schools and areas? *Oxf. Rev. Educ.* 46, 678–697.
- Uggerslev, K.L., Fassina, N.E., Kraichy, D., 2012. Recruiting through the stages: a meta-analytic test of predictors of applicant attraction at different stages of the recruiting process. *Person. Psychol.* 65, 597–660.
- Vogel, R.M., Feldman, D.C., 2009. Integrating the levels of person-environment fit: the roles of vocational fit and group fit. *J. Vocat. Behav.* 75, 68–81.
- Wang, H., Klassen, R.M., 2021. Longitudinal Influences of Teachers' Utility Values on Their Quitting Intentions: A Person-Organization Fit Perspective. Manuscript submitted for publication.
- Whetzel, D.L., McDaniel, M.A., 2009. Situational judgment tests: an overview of current research. *Hum. Resour. Manag. Rev.* 19, 188–202. <https://doi.org/10.1016/j.hrmr.2009.03.007>.
- Woods, S.A., Ahmed, S., Nikolaou, I., Costa, A.C., Anderson, N.R., 2020. Personnel selection in the digital age: a review of validity and applicant reactions and future research challenges. *Eur. J. Work. Organ. Psychol.* 29, 64–77.
- Yang, M., Lee, S.W., Goff, P.T., 2021. Labor dynamics of school principals in rural contexts. *AERA Open* 7, 1–18.

Supporting and assessing professional agency of student teachers in initial teacher education

Äli Leijen, Institute of Education, University of Tartu, Tartu, Estonia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	339
Teacher agency	339
Ecological model of teacher agency	340
Supporting teacher agency in initial teacher education	340
Competency base	340
Professional purpose	341
Environmental conditions	341
Noticing and assessing student teachers' agency in initial teach education	343
Conclusion	343
Acknowledgment	344
References	344

Introduction

Today, professionals from different areas are expected to keep up to date with the most recent developments in the field and continuously develop their knowledge and skills. This reality has also changed the perception of who a professional is and in fact has started to define it. While in the past, in many areas, basic knowledge and key skills could be acquired during initial training, and this knowledge may have been enough to pursue a lengthy career in a field, nowadays this is no longer possible. Today we assume that every professional, including a teacher, is a “learning professional” (Simons and Ruijters, 2014). More specifically, Simons and Ruijters point out that being a professional demands two moral requirements: commitment and integrity. These two moral requirements are necessary to develop four qualities of a professional: body of knowledge, theory of action, field of expertise, and engagement in professional community. All these relate to different ways of learning and results in two benefits for a professional: autonomy and authority. In the case of autonomy, teachers can make decisions about their work and professional development, and with authority they can contribute to the development of the profession, train novice teachers and experienced colleagues, and advice the general public.

The model of the “learning professional” requires a teacher to take an active role both in their daily work and in shaping the teacher profession more broadly. This active role can manifest itself in commitment, purposiveness, intentionality, and influence, as well as resisting or transforming existing practices. This manifestation has been conceptualized as human agency (Archer, 2000; Emirbayer and Mische, 1998; Eteläpelto et al., 2013; Moate and Ruohotie-Lyhty, 2014). As I will elaborate in the next sections, agency can derive from individual capacities but can also be supported or hindered by structural, cultural, or even material conditions in the environment (Priestley et al., 2015). As such, it is very important to understand how teacher's work is guided and structured by educational policies, legislative frameworks, and everyday practices as well as how teachers enter into the profession in the context of initial teacher education. Drawing on our previous publications (Leijen et al., 2020, 2021), this article will propose how to understand student teacher agency and how to support developing agency in the initial teacher education context.

Teacher agency

The concept of agency is closely related to identity, and from a broad understanding identity refers to who the person is. Agency, subsequently, indicates what a person does—taking an active role and influencing realities. Different research traditions conceptualize agency differently. For example, psychological traditions tend to view agency as a characteristic or capacity of a person. Related to this approach, Bandura (2009) defined human agency as “the human capability to exert influence over one's functioning and the course of events by one's actions” (p. 8). This indicates that one's actions depend primary on their personality and individual capacities, e.g., how determined one is, which concrete skills or knowledge one has, or which beliefs one carries. On the other hand, sociological, socio-cultural, and pragmatist traditions emphasize the relationship between the person and environment and argue that what a person does cannot depend only on the persons themselves but also on the environment around them. Followers of the socio-critical tradition seek for possibilities to change certain social structures and deliberate individuals and groups from their influence. Followers of the socio-cultural and pragmatist approaches do not address the influence of structures and systems so radically, rather, they see human agency as something that is made possible by the environment. Socio-cultural approaches tend to stress the mediated nature of agency, focusing on how contexts and cultural tools shape human beliefs, values, and activities (Lasky,

2005; Wertsch, 1993). However, the socio-cultural approaches are not uniform, and some authors neglect individual agency in theories of object-oriented activity while others focus on active agency while accepting interdependence with the social context (see Eteläpelto et al., 2013 for an overview). Drawing on the latter approach, Juutilainen et al. (2018) and Lipponen and Kumpulainen (2011) emphasize the social and contextual factors in developing student teachers' agency based on their qualitative research studies. Similarly to the subject-oriented socio-cultural views, the approaches derived from pragmatism focus primarily on the person's temporary constructed engagement with the environment and define agency as a potential to act by means of the environment (Biesta et al., 2015). The attainment of teacher agency is influenced by individual efforts, resources, and cultural and structural factors as shown in several empirical studies guided by this approach (Leijen et al., 2021; Oolbekkink-Marchand et al., 2017; Oosterhoff et al., 2020; Priestley et al., 2015).

Since it is important to understand how agency is achieved by student teachers from a broad perspective and what can facilitate or hinder this process in the teacher education context, I will proceed from the pragmatist approach and elaborate on an ecological model related to this tradition more thoroughly.

Ecological model of teacher agency

Drawing on the earlier work of Emirbayer and Mische (1998), Priestley et al. (2015) proposed an ecological model that describes the achievement of agency as a temporal-relational phenomenon through three dimensions: iterational, projective, and practical-evaluative. Firstly, the iterational dimension indicates that agency is rooted in past experiences. Thinking about teachers, we can see that their activities are influenced by their knowledge base, which contains personal life histories and also knowledge, attitudes, and skills acquired during teacher education and in their work context. In a way, the iterational dimension of agency guards for stability in teacher's actions since teachers tend to repeat activities that have proven successful in the past. Secondly, agency is oriented toward the future. The projective dimension contains short-term and long-term purposes which allow to shape actions in line with a teacher's plan for the future; thus, guarding for potential change in teacher's actions. Thirdly, the practical-evaluative dimension entails teacher's capacity in a present situation to make decisions for action. This decision-making process entails that a teacher considers alternatives for action, which are on the one hand guided by the iterational and projective dimensions and on the other hand by the cultural, structural, and material conditions, which could be enablers, constraints, or resources for action. For example, horizontal relations and close ties among teachers, as well as student teachers, can function as enablers of agency (see Juutilainen et al., 2018), while performance-oriented culture and hierarchical relations among people can act as constraints for teacher agency as noted earlier (see Priestley et al., 2015). Therefore, a student teacher achieves agency when they can consider alternatives and is able to judge which of them would be the most appropriate in a given practical situation in light of their greater professional purpose and competency base. "Agency is not present if there are no options for action or if the teacher simply follows routinized patterns of habitual behavior with no consideration of alternatives" (Priestley et al., 2015, p. 141). In the next section, I will further analyze this model and discuss how to support teacher agency in the context of initial teacher education.

Supporting teacher agency in initial teacher education

Below I will highlight a few ways how student teacher agency could be strengthened across these three dimensions of the ecological model: iterational, projective, and practical-evaluative.

Competency base

One of the oldest tasks of initial teacher education is developing student teachers' competency base. This is often addressed through foundation courses on subject disciplines as well as courses on educational psychology, sociology of education, and more practical methods courses which focus more concretely on mastering teaching activities. Therefore, on the one hand, initial teacher education emphasizes the value of formal knowledge which is important for providing perspectives and concepts for teachers for understanding their work (Biesta et al., 2015). On the other hand, it is important to support developing student teachers' practical knowledge (Darling-Hammond et al., 2005; Jensen et al., 2018; Meijer et al., 2002) to make learning experiences more personal and meaningful. What is crucial regarding both formal and practical knowledge in relation to agency is making it explicit and therefore more useable while constructing different alternatives for action. Developing formal and practical knowledge regarding teachers' work has, for a long time, been the main focus of teacher education. More recently, initial teacher education has started to emphasize student teachers' self-knowledge (Fairbanks et al., 2010), their strengths, and core qualities (Korthagen and Vasalos, 2005) as resources for professional work. This means that teacher education needs to provide space for exploring one's personal identity and for exploring the relationship and fit between their professional and personal competences (Anspal et al., 2019; Leijen et al., 2018).

In our teacher education context, we have encouraged student teachers to voice their strengths, beliefs, and values regarding education in individual assignments and asked them to explore where these come from. We have also asked them to describe their role models and professional role expectations following teacher standards. Moreover, we have asked student teachers to explore how their personal competencies fit with the role expectations and which aspects they would like to reach a better fit. These assignments are carried out alongside assignments focusing on professional purposes.

Professional purpose

Research has shown that long-term purposes are not often visible in teacher's work (Priestley et al., 2015) and not consistently considered by student teachers (Leijen et al., 2021). This indicates that a professional vision has remained implicit in many teacher education programs and perhaps stayed in the shadows of professional competences. There are, however, several scholars (Feinam-Nemser, 1990; Hammerness, 2006; Klette and Hammerness, 2016) who have argued for the importance of articulating a vision for a teacher education program and carrying it throughout different activities. These authors stress that a program vision should be visible from program documents such as course outlines, shared among the teaching staff, including the mentor teachers from practice schools, and embedded in different teaching and learning activities. Making the vision of teacher education program explicit also helps to model the importance of wider purposes in education for student teachers. Biesta (2009) has proposed three broad purposes of education: qualification, socialization, and subjectification (becoming a subject) which could be used as a wider frame for specifying more concrete long-term and short-term purposes of teacher education and consequently also for education in general. We have argued (Leijen et al., 2020) that besides these normative purposes of education articulated from the learner's perspective, student teachers need space for formulating more personal long-term purposes as well (Damon et al., 2003; Tirri et al., 2019). Finally, similarly as exploring the alignment between personal competences and professional role expectations, student teachers can be guided to synthesize their unique professional purpose.

In our context, we have encouraged student teachers to think about their own vision of education by paying attention to the three dimensions: focus, range, and distance as proposed by Hammerness (2003) in individual assignments. Following, we have asked student teachers to consider the three broad aims of education proposed by Biesta (2009), analyze how purposes are formulated in legislative frameworks and teacher standards, and finally to explore how well their personal vision is aligned with the professional purposes of education and whether they wish to increase the coherence between some aspects. Following these individual assignments, we have asked student teachers to reflect on their competence base (see previous section) and purposes and create a digital story about who they are and what their pedagogical vision is and to share this within a homegroup (voluntary formed group of 6–8 student teachers). Student teachers do these assignments at the end of the base module of teacher education. We see, however, that it would be useful to return to these assignments after some time e.g., once in a semester.

Environmental conditions

Although individual characteristics such as competencies and purposes are very important for teacher agency, these are not sufficient in case the environmental conditions, such as cultural, structural, or material, are limiting or restricting teacher agency. Starting from cultural conditions, it is evident that collaboration-oriented cultures are supporting teacher and student teacher agency. For these reasons, collaborative activities and student teachers' learning communities are recognized as an important context for sustaining and strengthening student teacher agency (Juutilainen et al., 2018; Lipponen and Kumpulainen, 2011; Toom et al., 2017; Vähäsantanen, 2015). More specifically, Juutilainen et al. (2018) showed that student teacher agency is supported by well-functioning peer groups where members feel emotionally safe and respected. A lack of emotional safety, unequal treatment, and comparisons between peer groups restricted student teacher agency. The latter indicates that a competition-oriented culture lowers trust and can raise anxiety among student teachers. For these reasons, it is not recommended to promote comparisons and competitions in a teacher education context and in an educational context more generally (see also the works of Pekrun (2016) regarding academic emotions). A collaboration-oriented culture in an initial teacher education setting also necessitates supportive relationships between teacher educators and student teachers. Core reflection, proposed by Korthagen and Vasalos (2005), and video-based guided reflection, developed in a context of an international project called ACITEA (Allas et al., 2017, 2020; Husu et al., 2008; Knezic et al., 2019; Toom et al., 2019), are good examples of reflection procedures that can be used by mentor teachers to empower student teachers. Although these procedures are different in terms of what is seen as the major resource for understanding and solving pedagogical situations (personal strengths and core qualities in the case of the former and theoretical knowledge and concepts learned during teacher education in the case of the latter), they both aim to empower student teachers and foster developing their views and voices in dialog with a mentor teacher or a peer student. This leads to structural conditions which support student teacher agency. Priestley et al. (2015) found that horizontal, reciprocal, and symmetric relations between staff are supporting agency where hierarchical and asymmetric relations between staff members are limiting teacher agency. Similarly, Juutilainen et al. (2018) showed that in order to support student teacher agency in initial teacher education context, it is important to break traditional teacher-students authority relationships. Other examples of structural conditions are, for example, the systems implemented to monitor and control educational practices. Regarding this, formative assessment with feedback from peers and teacher educators and reflection assignments are supporting student teacher agency while summative assessments leave less room for this.

The above-described possibilities to support teacher agency focused on the level of teacher education programs. We should also consider the macro level cultural and structural conditions and their influence on meso and micro levels (see for example a study conducted by Tinn and Ümarik (2021)). For example, some countries have a stronger competition- and performance-oriented culture than others. An example can be seen in countries where schools and universities are ranked in league tables. Or countries where regional level competitions for different areas are organized, targeting individual teachers, schools, and also individual pupils and students. There are also considerable differences between countries and regions regarding macro level structural condition in terms of monitoring and controlling teachers' work and related to material conditions of teachers work. The 2020 TALIS report (see OECD, 2020) showed that several countries who are known for a strong performance-oriented culture in education, such as

Singapore, Estonia, England, and the United States, reported high percentages (24% and above, OECD average was 16%) of young teachers wishing to leave the profession. This despite students of these countries often doing very well in international comparison tests such as PISA. Among other aspects, these tests also indicate that teachers are competent in their job. Contrary to these countries, only 10% of young teachers in Finland indicated they would like to leave the profession in the coming 5 years. Although this could be related to a high quality teacher education, it most probably also reflects the contextual factors of teachers' work. For example, although country level results about pupils' performance are collected in Finland, these results are only used for feedback purposes for individual schools and are not made publicly available. Similarly, teachers in the Finnish context are trusted and are not monitored by external evaluators.

Besides paying attention to how we can support student teacher agency in initial teacher education, we should also encourage student teachers to learn how to notice different aspects of their environment. Noticing different aspects of one's environment, alongside professional language and concepts (Biesta et al., 2015), is a prerequisite for being able to start negotiating work conditions with colleagues and policy makers. For these reasons, we proposed (Leijen et al., 2020) to distinguish between noticing (perceiving) and interpreting while evaluating environmental conditions prior to deciding for action in teacher education context (see Fig. 1 which captures the ecological model of teacher agency). Video-based reflection procedures (e.g., ACTEA mentioned earlier) have shown positive results to expand student teacher's scope of noticing in analyzing pedagogical situations (see e.g., Allas et al., 2020) and guidance procedures with peer students and mentor teachers have also shown positive results regarding developing practical knowledge and linking it with theoretical concepts (see e.g., Allas et al., 2017; Knezic et al., 2019).

In brief, while discussing how to strengthen student teacher agency across its three dimensions, we have shown how reflection and collaboration with peers and mentors can foster agency. In terms of types of reflection (see Leijen et al., 2020 for a more extensive overview) these examples have mostly covered supporting reflection-in action (also referred to as technical reflection) which is utilized in the decision-making in practical situation and reflection-on action (also referred to as practical reflection) which is conducted retrospectively. Both technical reflection and practical reflection foster learning about teacher's work within cultural and structural conditions. The third type, critical reflection, is needed in case we wish to guide more deep questioning of existing cultures and structures, reveal assumptions and values, which underlie people's actions and decisions, and propose alternatives. Mezirow's (1991) guide for transformative learning that was developed for the adult learning context could be useful for that purpose. This time-consuming and resource-intensive procedure is especially valuable to initiate deeper changes in groups of student teachers.

We have explored ways how to support teacher agency in today's initial teacher education. Some ways how student teacher agency dynamics could be noticed and assessed in the context of their initial education will be considered in the final section of this article.

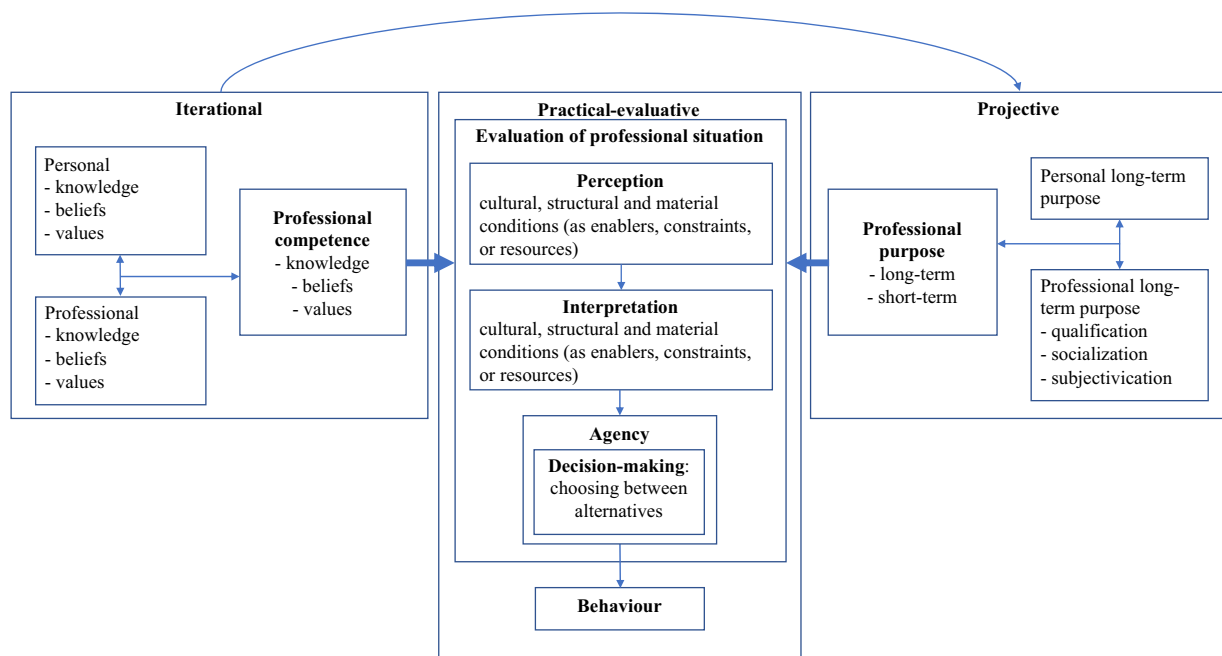


Fig. 1 Model of the formation of teacher agency. Based on Leijen et al. (2020).

Noticing and assessing student teachers' agency in initial teacher education

Most studies on teacher agency are qualitative. Several studies have investigated how to notice signs of student teacher agency. For example, [Juutilainen et al. \(2018\)](#) found that student teacher agency is manifested in evolving understandings, contesting prevalent views, and experiencing change in thinking. Similarly, [Lipponen and Kumpulainen \(2011\)](#) identified different forms of agency that emerged in the collective discussions of the classroom community evidencing pre-service teachers' transformative agency (making considerable changes in thinking and acting), relational agency (being able to utilize the support given by others as well as being a resource for others), and epistemic agency (taking responsibility over one's own learning). Somewhat differently, [Ukkonen-Mikkola and Varpanen \(2020\)](#) identified professional competence, reflection, and working with others as different dimensions of professional agency.

Some studies have also utilized questionnaires to investigate student teacher agency. For example, [Soini et al. \(2015\)](#) and [Toom et al. \(2017\)](#) investigated student teacher agency using an instrument previously developed for in-service teachers: the Teacher's Sense of Professional Agency survey (TPA) ([Pietarinen et al., 2013](#)) that focuses on motivation to learn, efficacy beliefs about learning, and intentional acts for facilitating and managing learning in the classroom. [Toom et al. \(2017\)](#) proposed a five-factor model of student teacher agency and found that first year student teachers in a Finnish university context had relatively high scores of agency. More specifically, they expressed a strong sense of community, found feedback useful for developing their instruction, practiced active help-seeking to develop their instruction, appreciated school development work, and somewhat less also appreciated collective efficacy. [Jääskelä et al. \(2021, 2017\)](#) proposed a more detailed scale for studying university student' agency, which could also be used in initial teacher education. Their instrument distinguishes between individual resources (interest and motivation, self-efficacy, competency beliefs, participation activity), relational resources (equal treatment, teacher support, peer support, trust), and contextual resources (opportunities to influence, opportunities to make choices) of student agency and suggests monitoring these in different university programs.

Both above-described instruments seem to focus mostly on student teacher's competency base, including knowledge and affective-motivational competences and beliefs, environmental enablers of agency, and relational aspects of agency. For our study of student teacher agency ([Leijen et al., 2021](#)), we developed a questionnaire based on the ecological model of agency to capture student teachers' competency and self-efficacy beliefs, professional purposes, and school level environmental enablers and resources of agency in three domains: planning of teaching and learning activities, teaching diverse ability students in the same class, and using ICT in teaching. These distinct areas were selected because agency is domain specific, and these domains represent different important aspects of today's teachers work. Planning of teaching and learning activities could be understood as a traditionally important activity teachers need to carry out. Teaching diverse ability students and utilizing ICT in teaching could be understood as a recently increasing demand describing teachers' work in many countries. Regarding each domain, the questions were similar. The 3 times 10-item questionnaire proved reliable and valid since we could confirm the previously proposed three dimensions of agency regarding all three domains. Based on an empirical study conducted in the Estonian context, we saw that students reported the highest agency related to planning of teaching and learning activities. We found statistically significant differences between student teachers with different teaching experience, whereas less experienced teachers had lower agency than more experienced teachers. More specifically, we found that gaining teaching experiences is positively related to two dimensions of planning related agency which target accumulated competences and professional purposes. The lowest agency among student teachers was reported regarding teaching diverse ability students in the same class. Moreover, gaining of teaching experience was not related to an increased agency in this domain, which could be explained by relatively late implementation of inclusive education in Estonia. Agency related to utilizing ICT in teaching was also relatively high, moreover younger students had higher agency in this area. However, student teachers evaluated their competences and environmental conditions higher than their purposes related to using ICT for teaching and learning, which brings back the importance of pedagogical vision and long-term purposes of education. Although these results were very informative, the formulation of items might be somewhat too general. For example, in the current version we have a general question about long-term purposes of a teacher, but we do not specify different long-term goals as elaborated above ([Biesta, 2009](#)). Therefore, we are currently developing a new instrument with more specifically formulated items. In future studies we would like to compare these instruments and find out whether they are both suitable for monitoring teacher agency dynamics in initial teacher education and during the professional career as a teacher.

Conclusion

This article explored how to conceptualize teacher agency and how to support and assess it in initial teacher education context. I showed that thinking about one's own and others' intentions and activities and creating a broader meaning about them is an important condition for achieving agency. As we have seen from this article, decision-making and acting on them is not solely dependent on an individual student teacher but reflecting will certainly help to better understand the situation, which is a prerequisite to start negotiating conditions and implementing changes in one's environment. Major structural and cultural changes are difficult to achieve as an individual, which is why it is very important to collaborate with others. Collaboration is also more likely to lead to collective agency, showing teachers actively guiding and shaping their daily work and teaching profession more broadly. This idea brings us back to the model of the "learning professional" which was introduced at the beginning of the article.

Acknowledgment

I would like to thank student teacher and teachers who have shared their conceptions and experiences in different studies which this article is largely based on. I would also like to thank colleagues and students for their insights and contributions in studying teacher agency.

References

- Allas, R., Leijen, Ä., Toom, A., 2017. Supporting the construction of teacher's practical knowledge through different interactive formats of oral reflection and written reflection. *Scand. J. Educ. Res.* 61, 600–615.
- Allas, R., Leijen, Ä., Toom, A., 2020. Guided reflection procedure as a method to facilitate student teachers' perception of their teaching to support the construction of practical knowledge. *Teach. Teach.* 26, 166–192.
- Anspal, T., Leijen, Ä., Löfström, E., 2019. Tensions and the teacher's role in student teacher identity development in primary and subject teacher curricula. *Scand. J. Educ. Res.* 63, 679–695.
- Archer, M.S., 2000. *Being Human: The Problem of Agency*. Cambridge University Press.
- Bandura, A., 2009. Agency. In: Carr, D. (Ed.), *Encyclopedia of the Life Course and Human Development*. Macmillan, Michigan, pp. 8–11.
- Biesta, G., Priestley, M., Robinson, S., 2015. The role of beliefs in teacher agency. *Teach. Teach.* 21, 624–640.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21, 33–46.
- Damon, W., Menon, J., Cotton Bronk, K., 2003. The development of purpose during adolescence. *Appl. Dev. Sci.* 7, 119–128.
- Darling-Hammond, L., Hammerness, K., Grossman, P., Rust, F., Shulman, L., 2005. The design of teacher education programs. In: *Preparing Teachers for a Changing World: What Teachers Should Learn Be Able Do*, pp. 390–441.
- Emirbayer, M., Mische, A., 1998. What is agency? *Am. J. Sociol.* 103, 962–1023.
- Eteläpelto, A., Vähäsantanen, K., Hökkä, P., Paloniemi, S., 2013. What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65.
- Fairbanks, C.M., Duffy, G.G., Faircloth, B.S., He, Y., Levin, B., Rohr, J., Stein, C., 2010. Beyond knowledge: exploring why some teachers are more thoughtfully adaptive than others. *J. Teach. Educ.* 61, 161–171.
- Feinam-Nemser, S., 1990. Teacher preparation: structural and conceptual analysis. *Handb. Res. Teach. Educ.* 212–233.
- Hammerness, K., 2003. Learning to hope, or hoping to learn? The role of vision in the early professional lives of teachers. *J. Teach. Educ.* 54, 43–56.
- Hammerness, K., 2006. From coherence in theory to coherence in practice. *Teach. Coll. Rec.* 108, 1241–1265.
- Husu, J., Toom, A., Patrikainen, S., 2008. Guided reflection as a means to demonstrate and develop student teachers' reflective competencies. *Reflective Pract.* 9, 37–51.
- Jääskelä, P., Poikkeus, A.-M., Vasalampi, K., Valleala, U.M., Rasku-Puttonen, H., 2017. Assessing agency of university students: validation of the AUS Scale. *Stud. High. Educ.* 42, 2061–2079.
- Jääskelä, P., Heilala, V., Kärkkäinen, T., Häkkinen, P., 2021. Student agency analytics: learning analytics as a tool for analyzing student agency in higher education. *Behav. Inf. Technol.* 40, 790–808.
- Jenset, I.S., Klette, K., Hammerness, K., 2018. Grounding teacher education in practice around the world: an examination of teacher education coursework in teacher education programs in Finland, Norway, and the United States. *J. Teach. Educ.* 69, 184–197.
- Juutilainen, M., Metsäpelto, R.-L., Poikkeus, A.-M., 2018. Becoming agentic teachers: experiences of the home group approach as a resource for supporting teacher students' agency. *Teach. Teach. Educ.* 76, 116–125.
- Klette, K., Hammerness, K., 2016. Conceptual framework for analyzing qualities in teacher education: looking at features of teacher education from an international perspective. *Acta Didact. Nor.* 10, 26–52.
- Knezic, D., Meijer, P., Toom, A., Leijen, Ä., Mena, J., Husu, J., 2019. Student teachers' self-dialogues, peer dialogues, and supervisory dialogues in placement learning. *Eur. J. Teach. Educ.* 42, 539–556.
- Korthagen, F., Vasalos, A., 2005. Levels in reflection: core reflection as a means to enhance professional growth. *Teach. Teach.* 11, 47–71.
- Lasky, S., 2005. A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teach. Teach. Educ.* 21, 899–916.
- Leijen, A., Kullasepp, K., Toompalu, A., 2018. Dialogue for bridging student teachers' personal and professional identity. In: *The Dialogical Self Theory in Education*. Springer, pp. 97–110.
- Leijen, Ä., Pedaste, M., Lepp, L., 2020. Teacher agency following the ecological model: how it is achieved and how it could be strengthened by different types of reflection. *Br. J. Educ. Stud.* 68, 295–310.
- Leijen, Ä., Pedaste, M., Baucal, A., 2021. Assessing student teachers' agency and using it for predicting commitment to teaching. *Eur. J. Teach. Educ.* 1–17.
- Lipponen, L., Kumpulainen, K., 2011. Acting as accountable authors: creating interactional spaces for agency work in teacher education. *Teach. Teach. Educ.* 27, 812–819.
- Meijer, P.C., Zanting, A., Verloop, N., 2002. How can student teachers elicit experienced teachers' practical knowledge? Tools, suggestions, and significance. *J. Teach. Educ.* 53, 406–419.
- Mezirow, J., 1991. *Transformative Dimensions of Adult Learning*. ERIC.
- Moate, J., Ruohotie-Lyhty, M., 2014. Identity, agency and community: reconsidering the pedagogic responsibilities of teacher education. *Br. J. Educ. Stud.* 62, 249–264.
- OECD, 2020. *TALIS 2018 Results (Volume II)*. <https://doi.org/10.1787/19cf08df-en>.
- Oolbekkink-Marchand, H.W., Hadar, L.L., Smith, K., Helleve, I., Ulvik, M., 2017. Teachers' perceived professional space and their agency. *Teach. Teach. Educ.* 62, 37–46.
- Oosterhoff, A.M., Oenema-Mostert, I.C., Minnaert, A.E., 2020. Aiming for agency. The effects of teacher education on the development of the expertise of early childhood teachers. *Teach. Teach. Educ.* 96, 103176.
- Pekrun, R., 2016. Academic emotions. *Handb. Motiv. Sch.* 2, 120–144.
- Pietarinen, J., Pyhältö, K., Soini, T., Salmela-Aro, K., 2013. Reducing teacher burnout: a socio-contextual approach. *Teach. Teach. Educ.* 35, 62–72.
- Priestley, M., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Publishing.
- Simons, P.R.J., Ruijters, M.C., 2014. The real professional is a learning professional. In: *International Handbook of Research in Professional and Practice-Based Learning*. Springer, pp. 955–985.
- Soini, T., Pietarinen, J., Toom, A., Pyhältö, K., 2015. What contributes to first-year student teachers' sense of professional agency in the classroom? *Teach. Teach.* 21, 641–659.
- Tinn, M., Ümarik, M., 2021. Looking through teachers' eyes—investigating teacher agency. *Br. J. Educ. Stud.* 1–17.
- Tirri, K., Moran, S., Mariano, J.M., 2019. *Education for Purposeful Teaching Around the World*. Routledge.
- Toom, A., Pietarinen, J., Soini, T., Pyhältö, K., 2017. How does the learning environment in teacher education cultivate first year student teachers' sense of professional agency in the professional community? *Teach. Teach. Educ.* 63, 126–136.
- Toom, A., Tiilikainen, M., Heikonen, L., Leijen, Ä., Mena, J., Husu, J., 2019. Teacher candidate learning of action-oriented knowledge from triggering incidents in teaching practice. *Teach. Teach.* 25, 536–552.

- Ukkonen-Mikkola, T., Varpanen, J., 2020. Integrated initial and continuing training as a way of developing the professional agency of teachers and student teachers. *Teach. Teach. Educ.* 96, 103189.
- Vähäsantanen, K., 2015. Professional agency in the stream of change: understanding educational change and teachers' professional identities. *Teach. Teach. Educ.* 47, 1–12.
- Wertsch, J.V., 1993. *Voices of the Mind: Sociocultural Approach to Mediated Action*. Harvard University Press.
- This page intentionally left blank

Teacher identity: crossing the technical-rationalist and affective divide

Navan N. Govender and Sue Ellis, University of Strathclyde, School of Education, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction and historical overview	346
Professional identity and professional learning	347
From identity research to a pedagogy of identity: the rise of identity work in teacher learning	348
Theoretical framework: rationalist and affective models for teacher identity construction	348
Rationalist models of teacher identity	349
Affective models of teacher identity	350
An integrated approach	351
Conclusion	352
References	353

Introduction and historical overview

The concept of identity has wide application across all landscapes in education research: we talk of learner identities to help us understand the learning pathways of children and young people; we talk of institutional and community identities to understand the distinctive practices and beliefs in schools, colleges, districts or universities, and we talk of professional identities to understand the particular characteristics and responses of both the teaching profession as a whole and of individual teachers or subject specialist groups. As a complicated construct, identity research is rooted in disciplines as diverse as philosophy, anthropology, sociology, psychology, education and political science. Each brings a distinct disciplinary framework, history, and discourse, which makes research into the professional identity of educators a complex and multi-faceted area of study. Multiple disciplinary approaches are evidenced, even when research is focused on teachers at a particular career-stage e.g. [Leeferink et al. \(2019\)](#); [Joseph \(2019\)](#); [Ritacco and Bolivar \(2019\)](#); on teachers working in particular subject areas (e.g. the [Widodo et al. \(2020\)](#) study on language teachers or [Karaolis and Philippou \(2019\)](#) and [Boaler \(2003\)](#) studies on mathematics teachers), and on the wider population of teachers (e.g. [Berger and Lê Van, 2019](#)). In its broadest sense, research on teachers' professional identity seeks to explain how teachers are positioned by, and come to position themselves in relation to, the cultural and social worlds they experience. In exploring the interface between teachers, their teaching environments and wider lives, professional identity offers the potential to contribute to our understanding of teacher professional development, teacher efficacy, agency, resilience and retention, and through these, inform our understanding of how to frame effective education policy, school leadership and curriculum guidance.

In this article we provide a brief historical overview of research on teacher professional identity before highlighting some current empirical studies and research themes. We explore professional identity as a dichotomy between rationalist and affective models of professionalism, arguing that an attempt to consolidate the rationalist/affective nexus allows for a more holistic and education-focused approach to teachers' professional identity. We then illustrate the possibilities of crossing the reason-emotion divide in professional learning and knowledge development.

The roots of much empirical identity research lie in social psychology and anthropology. Both disciplines seek to understand the intersection between the personal world of individuals and the collective meanings, social relationships and histories of the communities in which those individuals live and work. As a social psychological concept, identity research helps explain some socio-cognitive processes of teachers' professional development and some of the psychodynamic processes that arise from being a member of a professional group (see for example, [Clandinin et al., 1999](#)). As an anthropological concept, identity studies describe how individuals draw on cultural resources to make: "sense of" (about) themselves and represent the dilemmas of their social situations "for themselves" ([Holland et al., 2001](#), p. 8). Current social psychology and anthropology research perspectives suggest that identity is fluid, a fundamental act of being and a dynamic process that requires internal agency. It is malleable, shaped by social structures, interactions and experiences, but also exerts its own social and personal influence: "People tell others who they are but more importantly, they tell themselves, and they then try to act as though they are who they say they are" ([Holland et al., 2001](#), p. 3).

Over time, research on teacher professional identity has evidenced a gradual shift from static perspectives in which identity was used as a concept to capture the teaching profession's beliefs, values, and practices toward poststructuralist understandings of identity as a dynamic and fluid construct. [Beijaard et al. \(2004\)](#) undertook a systematic review of early research on teacher professional identity and noted that many early studies adopted a cognitive focus that positioned professional identity as the personal practical knowledge of teachers rather than a socially-negotiated construct. Psychological studies of teachers interpreted professional identity as an "image of self" that determines how teachers teach and respond to educational change whereas studies of teacher education linked professional identity to the teachers' professional roles or to processes such as reflection and self-evaluation. Both kinds of study assumed that professional identity was relatively stable and a stabilizing influence. In contrast, more recent research views professional identity as fluid, social, multifaceted, linked to learning, to affect and to feelings of belonging, and that it is "neither intrinsically stable nor intrinsically fragmented ... [but] more, or less, stable and more or less fragmented at different times and in

different ways according to a number of life, career and situational factors.” (Day et al., 2006, p. 601). Although a systematic review of recent research on student teachers’ professional identity (Rodrigues and Mogarro, 2019) uncovered a range of definitions and investigative purposes for studying identity, there was wide agreement that identity is malleable, “formed through socio-cultural situated practical experiences and depends on the complex negotiation between professional experiences and external factors.” (Rodrigues and Mogarro, 2019, p. 6). This shift is evidenced in social psychology frameworks that seek to understand how identity contributes to teachers’ sense of: belonging (see for example, Brown and Everson, 2019); their emotional engagement in teaching (see for example, Wang et al., 2019; Nichols et al., 2017), and their professional knowledge and craft (see for example, Leeferink et al., 2019). It is also evidenced in anthropological frameworks that describe how interactions between culture and agency shape how teachers think, feel, act and learn (see for example, Sachs, 2001), and in sociological frameworks that chart the impact of power dynamics and social positioning on teacher identity (see for example, Moghadam et al., 2020; Day et al., 2006). Despite these common understandings, Rodrigues and Mogarro (2019) argue that identity research in teacher education is “still hostage of the other social sciences,” and that “a unique and distinct discourse within teacher education has not yet been developed.” (ibid 2019, p. 10).

Professional identity and professional learning

A unique discourse for education must draw on affective and cognitive views to explain the power of both individual agency and institutions to shape professional identity and professional learning. The concept of “communities of practice” (Lave and Wenger, 1991) offers an explanation of how professional identities are forged from the social interactions between community members who share a common interest and competence. It suggests that shared knowledge and informal, situated learning are negotiated through joint activities so that the community grows, legitimates and regulates the knowledge, values, skills, resources, and know-how of individuals. Apprenticeship into the community begins with legitimate peripheral participation, following non-linear pathways to full participation and acceptance into the community. Wenger (1999, 2010a) argues that this social participation and informal, work-based learning allow identity to be negotiated through words and actions, as individuals work out “ways of being a person in that context” (Wenger, 2010a, p. 133). Boaler (2003) refers to this as a “dance of agency” that requires active engagement in community practices, histories, learning and stories, alongside some compliance to community norms and practices and to the hierarchies of power. Wenger-Trayner and Wenger-Trayner (2015) suggest that multiple communities of practice, including policy, regulatory and research communities, form a “landscape of practice”. Education professionals navigate this landscape, working across community boundaries to align different knowledge-domains within specific contexts to forge a meaningful moment of practice. Different alignments create the possibility of different actions and outcomes (which neatly explains professional innovation), but all alignment requires imagination to extrapolate experience across contexts and practical engagement. The “landscapes of practice” model suggests teachers do not simply “put theory into practice” and it positions teacher agency and identity as intrinsic to active knowledge-production. The concept of professional knowledge generated in a landscape of practice chimes with understandings of trans-disciplinary distributed “Mode 2” knowledge produced in its context of use (Gibbons, 2000). In this, it provides a model of teacher identity-work that does two things: Firstly it interweaves professional identity with narratives of *how* teachers come to know, and secondly, it harnesses professional identity to concepts of professional agency, reflection and knowledge through “engagement, imagination and alignment” (Wenger, 2010b, p. 184) in which teachers attune their practices to the needs of their students and the communities the school serves.

However, when professional knowledge, learning and identity are socially negotiated, the pathways to adaptive professional competence can seem opaque and uncertain. Practitioners may choose to resist, reshape or side-line new and potentially useful knowledge-communities (Farnsworth et al., 2016) or adopt intuitively appealing but ineffective practices. Education research to problematize and understand how professional identity operates to expand or inhibit professional learning, knowledge-formation, efficacy, and values in complex contextualized situations is vital. In some situations the desire to adopt a particular professional identity may inhibit learning, agency, enquiry, reflection and curiosity: Hall et al. (2012) report that student teachers in practicum schools were marginalized rather than engaged as the legitimate peripheral participants that Lave and Wenger (1991) suggest. Feeling too vulnerable to display uncertainty or to position themselves as learners, the student teachers forged narrow teaching identities that reduced professional competence to “class control” and children’s learning to “skill development”. Research into identity and professional values is equally complex and presents difficulties of identification and theorization. The theory of implicit bias, suggests that practitioners can harbor unconscious “discriminatory biases based on implicit attitudes or implicit stereotypes [which] ... can produce behavior that diverges from a person’s avowed or endorsed beliefs or principles.” (Greenwald and Krieger, 2006, p. 951). Murphy and Wyness (2020) report from England that teachers’ grade predictions for university admission are overwhelmingly optimistic for all socioeconomic groups except high-attainers from disadvantaged backgrounds, the only group significantly more likely to have pessimistic grade predictions.

Given such uncertainties, policy and regulatory communities in some countries have imposed competency-based or content-focused frameworks that centrally direct teachers’ professional actions and learning and do not engage professional identity. Such technical-rationalist innovations introduce standardized curricula, scripted teaching, high-stakes assessment or target-setting accompanied by standards-based managerialism and funding regimes (Hall and McGinity, 2015). The incongruity is clear: increasing attainment, social justice, equity, inclusion and wellbeing in different contexts requires responsive teachers who are noticing, creative, critical and inquiring, who bring agency, knowledge and expert judgment to adapt teaching practices, and

who challenge bias and learn in context. Contextual affordances and constraints may require professionals to adopt different pathways to a common outcome, top-down technical-rationalist interventions that impose a single pathway diminish those very aspects of professional identity - agency, reflection and criticality - that are necessary.

From identity research to a pedagogy of identity: the rise of identity work in teacher learning

Bensimon (2007) argues, albeit for higher education, that policy analysts' focus of attention on tuition policies, incentive funding, accountability and regulatory systems overlooks practitioner sense-making frames as institutional agents of change. Despite terms such as "self evaluation", "active reflection", "collaboration" and "peer regulation" entering the policy discourse (see for example OECD, 2019, 2020) we need better understanding of the kind of healthy professional identities that can do this, how they can be grown at scale across multiple contexts, and the kinds of learning, actions and values required to re-shape schools to deliver positive educational outcomes and social justice.

Whilst Rodrigues and Mogarro (2019) are right to argue that education research on professional identity has yet to hit its own stride, studies of identity formation do chart some of the tensions and contradictions at career transition points that impact identity and prompt professional learning (Gaudelli and Ousley, 2009). These highlight affective influences on the kinds of teachers that practitioners imagine themselves becoming (e.g. Dinham et al., 2017) and the kinds of discourse that affect teacher identity (e.g. Haniford, 2010). In policy studies on teacher stress and burnout professional identity both measures impact and can operate as a protective factor in how teachers internalize challenging work conditions, systems or environments (e.g. Kangas-Dick and O'Shaughnessy, 2020; Hong, 2010; Dassa and Derosé, 2017; Skinner et al., 2018; Nichols et al., 2017; Dinham et al., 2017). We know that moments of tension can prompt reflection and learning, but an overload creates anxiety and stress, impacting negatively on self-efficacy, professional learning and, ultimately, pupil attainment (Glazzard and Rose, 2020).

Research has identified some effective and ineffective features that promote professional identity and learning through mentoring (Orland-Barak and Wang, 2021; Anspal et al., 2019); course design (Alderton, 2020); system design (Hökkä et al., 2019); and self care systems (McKay, 2019). Specific pedagogical tools have been designed to prompt "identity work" in the form of deliberate, conscious reflection and engagement, to promote self-aware, noticing and adaptive practitioners. The studies tend to be small-scale but include identity-work tools such as: protocols for reflection (Pomerantz and Ippolito, 2015); auto-ethnography (Lavina and Lawson, 2019); collaborative reflection (Maclean and White, 2007); lesson study (Lieberman, 2009); action research (Estola, 2003; Webb, 2005; Cattley, 2007; Maclean and White, 2007; Vavrus, 2009) as well as case-study and focused experiential learning; professional (face-to-face and virtual) networks (Lundin et al., 2020). The concept of "ethical practitioner enquiry is underpinned by an epistemology which challenges both technical and practical views of education" (Lofthouse, 2014, p. 15) and necessarily engages with issues of power, identity, values, and socio-cultural context in order to avoid unwittingly reproducing problematic ideologies and practices.

Seemingly advocating for a critical approach to practice-based research and professional development, Lofthouse (2014, p. 17) states that "practitioner enquiry rests on the proposition that those in practice are able to take informed intentional actions, explore their effects and form judgments of their value" as part of the standard social practices in education. Here, exploring effects, if viewed in relation to questions of ethics, must engage with the social effects of teaching and learning practices on the lived experiences of learners and teachers alike. Both these social effects and the judgments that enquirers make, are thus evaluated in relation to the ideologies (dominant or not) that they represent or challenge. This article reveals that a critical approach that bridges rationalist-affective domains of being and doing are necessary for constructing teacher identities that are professionally *and* ethically driven.

As both a research methodology and a classroom pedagogy (for higher education student teachers and school-going learners), critical reflection thus becomes a useful means for identifying, interpreting, and deconstructing teacher identities in context (Avraamidou, 2019; Cochran-Smith et al., 2020). Critically reflective journal writing, framed by action research, critical practitioner enquiry, or autoethnography, enables this:

By looking at critical incidents and experiences (as texts) in relation to issues of power, teachers and students might interrogate the ways in which social and political power influence the everyday construction of individual identities, how individuals take up or resist hegemonic orders, and the ways in which individuals occupy positions of privilege and subordination at different times and in different spaces. (Student) teachers can record their own identity shifts over time, in relation to lived experience, using critical analytical tools such as critical discourse analysis (Fairclough, 2001) to reveal how their own experiences, uses of language, and ways of thinking are situated in systems of power. That power can be defined in relation to gender and sexual diversity, or other identity categories, and contributes toward (student) teachers gaining the confidence and social consciousness to construct more socially just identities for themselves

(Govender, 2019, p. 353).

Theoretical framework: rationalist and affective models for teacher identity construction

In this section, we briefly outline rationalist (derived from cognitivist understandings of professionalism and identity) and affective (associated with socio-cultural and poststructuralist understandings of professionalism and identity) models for teacher identity construction. This is followed by our attempt to consolidate these two models, and the divide they represent, into an integrated

model for teacher identity construction characterized by fluidity, critical reflexivity, and the understanding that identity and practice are intrinsically connected to issues of power. Where a rationalist model prioritizes teachers' identities as disciplinary experts and their practices as measurable against predetermined national standards (Edwards and Blake, 2007) by taking on cognitivist definitions of identity, an affective model foregrounds multiplicity, discontinuity, and sociality (Arvaja, 2016) by taking on a poststructuralist perspective of identity. In trying to consolidate these two seemingly disparate models, we offer an approach that sees teachers' identity construction as socio-culturally situated, pluridimensional, and as deeply bound to issues of power in and out of the classroom.

Rationalist models of teacher identity

From a rationalist, or cognitivist, understanding of teacher identity and professionalism, it becomes easy to identify and critique the characteristics of a "good" teacher. Such characteristics are described in varying detail across international educational policies through professional standards for teacher registration or regulation and the assessment criteria of teacher education programmes (Mockler, 2010; Buchanan, 2014). In this way, by articulating "good teachers" and "good teaching" in cognitivist, measurable ways (Benesch, 2018), emerging (student) teachers and teacher educators are at risk of reproducing established, naturalized, and sometimes problematic notions of teacher identity and practice (Edwards and Blake, 2007; Arvaja, 2016; Gonçalves et al., 2013). Across the literature, three main contributing factors can be identified in the process of "professionalizing" teacher identity: (1) teachers are constructed as disciplinary experts, (2) teacher competency and professionalism are conceptualized and driven by pre-determined professional standards issued by policy-makers, and (3) professional practice, as interconnected with professional identity, is legitimated through particular ways of doing evidence-based practitioner enquiry.

Firstly, if teachers, particularly in secondary schooling, are constructed as masters of specialist knowledge rather than practitioners of or agents acting *through* particular disciplines, they may also inadvertently become constructed as the sole intellectual authorities in the classroom. In this sense, teachers may see themselves as holders of knowledge obtained and validated through university degrees. In Freirean terms, this risks positioning both knowledge and teaching/learning in problematic ways: firstly by foregrounding knowledge over people, and secondly by positioning those people (teachers and learners respectively) as depositors and receptacles of knowledge rather than co-constructors of meaning (Moje, 2008).

Moje (2007, p. 1) raises this issue by questioning how teachers "deal with conflicts between the intellectual work of teaching content concepts and the moral work of teaching those content concepts to human beings ... with varying perspectives on the value of the content, varying skill sets and ways of knowing that they brought to their learning, and rich and full lives that might or might not intersect with the content under study." In this sense, the rationality of intellectual work without the affective consideration for ethics in education risks decontextualizing classroom content, pedagogy, and assessment. On the other hand, teachers who see themselves as experts of their discipline may also gain confidence and agency in their classrooms. Where literacy and language teachers, for example, situate their knowledge and practice in established theoretical perspective *and* as part of communities of practice, they can begin to make decisions about effective literacy and language pedagogy that resonates with their learners in meaningful ways (Ellis and Smith, 2017; Schneider Kavanagh and Danielson, 2020; Moje, 2008).

The second contributing factor to developing teacher identities involves how teacher competency and professionalism are conceptualized and driven by professional standards, often found in national policy and teacher education assessment. Because such standards are shaped by existing dominant ideologies, they can have a strong influence on how teachers regulate themselves and each other. Furthermore, teacher educators and experienced in-service teachers use these as assessment tools to provide new and student teachers with feedback on their performance and progression, socializing emerging teachers into potentially narrow versions of professionalism.

If these standards are conceptualized from a rationalist perspective, they can serve to narrow what counts as appropriate, professional teacher identities (Hall and McGinity, 2015). For instance, to what extent are innovation and creativity measured and awarded? What kinds of social consciousness in teaching and learning are prioritized or excluded? What version of professionalism is legitimated through policy documents and regulatory boards? And, what kinds of (inter)actions are considered "appropriate" within the frame of the standards that govern? These issues highlight the tension between teachers and broader educational institutions:

Teachers' professional identities are being reworked from inside and outside of the teaching profession. Teachers as professionals are increasingly subject to external standards and codes of practice while, on the other hand, they are struggling to construct their own professional identities by exercising their own agency, prioritising the needs and interests of their students and classroom interactions and promoting "appropriate" professional practice

(Gonçalves et al., 2013, p. 57).

Teachers must therefore navigate the contentious space between the personal and the professional. This is particularly true in contexts where national professional standards do not recognize how teachers' own value systems, cultures, and ways of knowing and how these impact on the identity-practice nexus. For example, if teachers' identities are homogenized through policy and regulatory processes, both the regulatory boards and teacher education institutions risk losing critical access to the productive power of

difference (of perspective, of approach, of interpretation, ideology, or of subjectivity) and overvaluing sameness (O'Connor, 2006). Edwards and Blake (2007) as well as Johnson and Golombek (2020) argue that it is the critical engagement with difference and diversity in teachers' identities, knowledges, and practices that drive education forward.

The third component concerns how teachers evidence and measure the efficacy of their own practice, through evidence-based practitioner enquiry, against policy requirements and dominant research. While the intention to engage teachers with this kind of reflective work "contends that the education of teachers can be realized neither through the delivery of a pedagogical blueprint of techniques, nor through the disinterested delivery of disciplinary knowledge" (Edwards and Blake, 2007, p. 38) is a critical one, there are still risks depending on how that enquiry is ideologically framed. From a rationalist perspective, teachers' enquiries into their own classrooms may be limited to the reasoned analysis of practice, where practice is measured against policy and research without enough engagement with the socio-political context of either. Without a critical approach to enquiry, that sees ideology and power as necessary dimensions to research and pedagogy, practitioner enquiry risks becoming a mechanism for teachers to legitimate their practices by narrowly aligning them to the established policies. "Good" findings may end up being used to exemplify the successes of policy, vindicating teachers and governments of inadequacy. The question this raises is whether or not teachers might feel that they are able to interpret the policy standards, or what counts as evidence for effective teaching and learning, in creative and critical ways.

According to Edwards and Blake (2007, p. 34), one mechanism for the regulation of teachers' identities "is the persistence in education of a belief in an objective, transcendent rationality, the assumption that there is a neutral vantage point from which student teachers, by submitting to the technology of reason and logic, may discern order, laws and clarity across teaching dilemmas within particular contexts." Where teachers are conceived of (and where they conceive themselves as) masters of disciplinary knowledge, performers of professional standards, and practitioners of reasoned enquiry, purely rationalist models of teacher identity formation become insufficient. This is not to say that teachers should not feel like powerful actors within their disciplines, or that their interrogation of their professional practices are not valuable, or that it does not lead to valuable learning. Instead, this is to say that a rationalist model provides only a partial view of what it means to be a teacher and do teaching (Aronson, 2020). In the next section, we outline the possibilities for emotional competency as part of a more holistic approach to understanding teacher identity formation.

Affective models of teacher identity

Who teachers are, their relationship to the content knowledge they teach, the contexts and institutions of learning within which they work, and the relationships they establish with staff and students matter (Aronson, 2020). Zembylas (2007) provides a useful conception of how an "emotional ecology" contributes to educational work in the classroom as well as teachers' identity formation. Defined as "a teacher's (or learner's) emotional knowledge in a particular social and political context, including rich connections to emotional experiences, and relationships with others (e.g. students, colleagues, parents) over time" (Zembylas, 2007, p. 357), the emotional ecology of a place of learning (from classrooms to whole school experiences) aid in enabling or disabling effective teaching and learning. Such an ecosystem is therefore also characterized by, at the very least, (1) issues of investment (Norton, 1995; Darvin and Norton, 2015), where teachers' relationships to content knowledge and pedagogical approaches influence the extent to which they are able to work productively and in culturally sustainable ways (Ladson-Billings, 2014), and (2) how teachers position themselves, taking up or resisting identities.

In the first instance, issues of investment in content knowledge and pedagogical approaches suggest that having knowledge of what to teach and how to teach is not enough. The emotional attachment that teachers have to their subject area, their profession, and how they understand the impact of their teaching on learners is paramount to whether or not teachers themselves might see their identities and practices as worthwhile. Here, "investment regards the learner [and the teacher] as a social being with a complex identity that changes across time and space and is reproduced in social interaction" (Darvin and Norton, 2015, p. 37). By recognizing whether or not, and to what extent, teacher's emotional investment in their work affects their practice, teachers themselves can begin to interrogate more deeply the socio-cultural and political agendas that underlie their practice. For instance, in a climate where assessment-driven teaching dominates, teachers who resist the culture of teaching-to-the-test may experience certain emotional tensions that "compel them to realign their pedagogical practices and beliefs" (Loh and Liew, 2016, p. 269) to the underlying neo-liberal ideology at work. In negotiating these tensions, teachers are required to evaluate their own beliefs and reasons for teaching while taking seriously the effects that their own practices may have on their learners. Teachers who take up critical approaches to teaching, then, will have their commitment to social justice education tested by institutions bent on achieving high attainment levels that are narrowly defined, colleagues who regard successful teaching as only that which results in learners' high attainment scores, or even learners who cannot see the connection between classroom content and their current or imagined lives.

The second component to an emotional ecology (Zembylas, 2007) is that of agency and the positions that teachers take up or resist in relation to professional identity. Related to investment, teachers' capacity to recognize, acknowledge, and confront their own subject positions (values, beliefs, performances) as they emerge in specific socio-cultural contexts enables teachers to understand how their own identities, practices, and emotional investments reproduce or challenge certain social norms. In this sense, teachers are *positioned* (Janks, 2010) by policy constructs of professional identity, institutional cultures of teaching and learning, and regulatory boards. They are also *positioning* (Janks, 2010) through the impact of their pedagogical choices in the classroom that enable or disable particular ideologies, their selection of content, and the relationships they help to establish with learners in the classroom that may validate or stigmatize certain learner-identities. For example, consider how student teachers in one South African teacher education program critically reflected on their own gendered identities in order to consider how their beliefs and

socio-cultural context positioned them in particular ways. Their engagement with identity exposed possibilities for how critically reflective work might be used to rethink their classroom practice in ways that might enable their imagined learners to think more critically about language and gender (Govender, 2019). In this study, the emotionally-laden reflective writing by student teachers revealed both the conceptual and emotional work involved in facing social issues such as gender and sexual diversity, and how this might be used by those student teachers to (re)position themselves (and each other) in more socially just ways.

Therefore, the “rich connections” (Zembylas, 2007) that teachers make with their learners and other teachers, across educational contexts and cultures, can be understood as intrinsically linked to their ways of identifying (being) and acting (doing) where “emotions are regulated socioculturally whereby members of discourse communities are taught which emotions are considered appropriate in particular social situations and which are viewed as improper” (Jaggar, 1988, cited in Benesch, 2018, p. 61). Understanding how emotional ecologies are constructed in the interests of some and used to police (student) teachers into particular ways of teaching, thinking, doing, and believing is therefore necessary in professional identity construction and development that takes on a social justice agenda (Aroson et al., 2020; Vasquez et al., 2019). Validating and negotiating the emotional ties and responses that teachers have to social issues, content knowledge, pedagogical approaches, and so on, in socio-cultural context may enable those teachers to take hold of their agency and construct identities and practices that are more culturally responsive, sustainable, and socially just.

An integrated approach

An integrated approach to understanding teacher identity and its development over space and time recognizes how reason and emotion, rationality and affectation, inform each other across socio-cultural context, individual and collective identities, and disciplines or communities of practice. It also advocates for the critical engagement of both self and socio-cultural context by pointing at the insufficiency of technical-rationalist approaches alone (Mockler, 2010). Table 1, below, attempts to map the rationalist, affective, and rationalist-affective crossovers evident in the literature on teacher identity.

From this outline, it becomes easy to see how the underpinning theoretical perspectives on teacher identity, and its development, influence what constitutes a “good” teacher, and by effect “good” teaching. This starting point therefore has a ripple effect on the ways in which disciplinary knowledge, professional standards, change, reflection and practitioner enquiry, and social context and power are perceived as either having or not having an effect on how teachers see themselves, their work, and their impact in and out of the classroom. For instance, cognitive approaches to disciplinary knowledge foreground knowledge over people while affective approaches position teachers as emotional beings who reflect on their emerging proficiency in disciplinary knowledge. At the crossover, a rationalist-affective model places teachers in the role of knowledge co-constructors. Here, alongside their learners, teachers can begin to negotiate what knowledge counts, how it is used, and the effect such knowledge has on their own identities as teachers as well as its effect on learners’ understanding of the world (Moje, 2008). Over and above reflecting on their emotions, teachers must also interrogate their investment (Darvin and Norton, 2015) in the field and its role in learners’ lives (that is, learners’ investments).

Rationalist-affective approaches, then, emerge from a tradition of poststructuralism in education. Identity, from this perspective, is fluid and continuously changes in relation to:

1. *Changing policies and professional standards.* At this level, policy constructs of professionalism are interpreted and critically evaluated in relation to teachers’ contexts of work as well as their own ideological or belief systems. What counts as “professional identity”, then, is constructed by the teacher in relation to external forces such as national professional standards, initial teacher education assessment criteria (Cochran-Smith, 2003; Cochran-Smith et al., 2020; Buchanan, 2014) and ethics (Lofthouse, 2014), and not the other way around.
2. *Developments in disciplinary knowledge.* As co-constructors of disciplinary knowledge and practitioners within disciplinary communities of practice, teachers should not only measure their technical proficiency of the discipline but also their investment in the discipline and the social effect of their approach to disciplinary knowledge on learners (Ellis and Smith, 2017). That is, are teachers merely socializing learners into existing communities of practice, or are teachers engaging learners in processes of knowledge construction in socially sustainable ways?
3. *Changing learner populations and contexts of teaching and learning.* Teachers do not work in a vacuum, and so how they perceive and engage with difference, diversity, and power is affected by their own subject positions. That is, teachers’ personal beliefs, dispositions, and emotions affect their pedagogies and educational approaches (Hinojosa Pareja and López López, 2018). Critically reflective teachers (Govender, 2019; Kubanyiova, 2020; Aronson, 2020), then, actively interrogate how their own subject positions affect their pedagogical decisions and change over time with changing educational contexts.
4. *Reflective practice and practitioner enquiry that foregrounds teachers’ agency.* Reflection and practitioner enquiry practices are processes for engaging with teacher identity, and necessarily crosses the boundaries of what teachers do in schools and the effects of these actions (Lassonde and Israel, 2008). However, it is also important to make clear how these actions are imbued with ideologies, and therefore how an investigation into those actions constitutes a “critical inquiry” (Cochran-Smith et al., 2020).
5. *How teachers view themselves in relation to issues of power in and out of the education system.* Because teaching involves both ways of doing and being (Vasquez et al., 2019), teachers’ identities are intrinsically bound to issues of power. Race/ethnicity, gender, sexuality, nationality, age, (dis)ability, and so on, all contribute to how professionalism and educational practice happen (Lawrence and Nagashima, 2020). Rationalist-affective models of teacher identity, then, actively engage with these issues in context and in relation to their social impact on learners and learning environments by taking specifically socio-cultural and poststructuralist perspectives.

Table 1 Mapping the rationalist, affective and rationalist-affective crossovers in teacher identity.

	<i>Rationalist models</i>	<i>Affective models</i>	<i>Rationalist-affective crossovers</i>
Theoretical perspective	Teacher identity development is conceptualized from cognitivist perspectives	Teacher identity development is conceptualized as emotional and/or pastoral	Teacher identity development is conceptualized from socio-cultural and poststructuralist perspectives
Disciplinary and pedagogical knowledge	Teachers as experts of disciplinary and pedagogical knowledge define their identity in relation to the mastery of knowledge and practice	Teachers' investment (Darvin and Norton, 2015; Norton Pierce, 1995) in disciplinary and pedagogical knowledge and practice is foregrounded	Teachers are constructed as co-constructors of disciplinary and pedagogical knowledge and practice, which is influenced by both their (perceived) mastery and investment in the field
Professional standards	Professional standards provided by national and international policy construct teacher identity development as the learning and performance of pre-defined, measurable, competencies and skills	Navigating the emotional ecology (Zembylas, 2003, 2007) of one's personal and professional contexts is necessary in unpacking and constructing changing teacher identity	Professional standards are interrogated in relation to the socio-cultural contexts they are meant to represent, while the performance and evidencing of professional standards are interrogated in relation to identity, the context of performance, ethics, and power
Change	Teacher identity is perceived as developmental and measurable, as described in professional standards and initial teacher education assessment criteria	Teacher identity is perceived as developmental and measurable, as described in professional standards and initial teacher education assessment criteria, but it also involves taking seriously the emotional impact of teaching on teachers and learners	Teacher identity is envisaged as a fluid construct that changes with contexts of teaching and learning, foregrounding teachers' ongoing negotiation of knowledge, practice, investment, institutional culture, and so on.
Reflection and practitioner enquiry	A professional standards-based approach focuses on teachers' identity as primarily connected to their performance of established criteria	Teacher-reflection is advocated as a means to identify and unpack how teachers, their feelings, and the impact of their teaching, influence teacher identity development	Critical reflection and critical approaches to practitioner enquiry recognize that teachers, as emerging and agentive participants in communities of practice (Lave and Wenger, 1991), must negotiate not only professional standards and initial teacher education assessment criteria but also their own ideological positions (including ways of doing, thinking, being, reading, writing, speaking, believing)
Social context, ideology and power	Teacher identity is situated in the mind and is measurable in relation to skills competency	Teacher identity is constructed in relation to their contexts of teaching (such as learners' needs and the labels used to describe their contexts in relation to socio-economic status, deprivation or privilege, resources, and so on)	Teachers' identities inform and are informed by their contexts of work and life (including under systemic relations of power related to socio-economic equity, race/ethnicity, gender, sexuality, age, region, language, and so on)

Teacher identity, then, and its development involve ongoing, critical negotiations of who teachers are, who they want to be, the effects they have on their learners and social context, and the ways in which they navigate the plethora of ideologies already embedded in everyday educational practice. This is all done in view of the external and internal forces that teachers negotiate on a day-to-day basis.

Conclusion

This article has explored the politics of teacher identity, and its development, in the research available. By paying particular attention to the role of rationalist-technical and affective models, it becomes clear how important it is that emerging teachers and teacher education are able to critically engage both rationalist and affective approaches to professional identity. This is doubly important when one considers the extent to which identity and socio-cultural context can shape the lived experiences and practices of teachers and their learners.

From the literature reviewed in this article, six domains that shape, and are shaped by, teacher identity and identity development have emerged: (1) theoretical perspective, (2) disciplinary and pedagogical knowledge, (3) professional standards, (4) change, (5)

reflection and practitioner enquiry, and (6) social context, ideology, and power. Each of these domains, as discussed, are interdependent and bound to particular contexts, communities, and cultures. Interrogating them, then, becomes vital in understanding how teachers understand themselves and themselves within and without the teaching profession and in relation to society, culture, ideology, policy, and investment. This has implications for transforming classroom practice and teacher education into becoming more sustainable, inclusive, and socially just.

Recent research has suggested that a move toward integrated, rationalist-affective border-crossing perspectives on teacher identity and teacher identity development have been taken up in teacher education programmes, especially where this relates to socio-cultural models of literacy and literacy education (Burgess and Rowsell, 2020; Johnson and Golombek, 2020; Gupta, 2020; Aronson et al., 2020; Lawrence and Nagashima, 2020; Hong and Cross Francis, 2020; Kocabaş-Gedik and Ortaçtepe Hart, 2020; as well as others). This, it would seem, speaks back to dominant neoliberal moves (Winter, 2000; Edwards and Blake, 2007) in educational policy across global contexts and advocates for the recognition of (student) teachers' diverse contexts and cultural identities as resources for educational practice – especially where they have a direct impact on the lives and education of children and young people. Furthermore, such recognition may enable increased teacher agency and renewed understandings of professionalism: from ongoing self-reflection and practice-based inquiry that is socially and culturally situated, to the construction of socially and culturally sustainable curriculums that engage both intellectual and emotional faculties. This may also have implications for research and research practices that engage more fully with the needs of particular contexts and peoples in relation to education (O'Keeffe and Skerrett, 2021; Medina, 2020).

While the current research paints a promising picture for teachers and teacher education, work needs to continue and systemic changes need to be made. By actively engaging in rationalist-affective crossovers, teachers, teacher educators, and the students who are meant to benefit most from education might be able to spot the issues in the system, pose critical questions, research and problem solve, and effect change on their circumstances and environments in sustainable ways.

References

- Alderton, J., 2020. Kelly's story: transformative identity work in primary mathematics teacher education. *Gend. Educ.* 32 (2), 145–160.
- Anspal, T., Leijen, Å., Löfström, E., 2019. Tensions and the teacher's role in student teacher identity development in primary and subject teacher curricula. *Scand. J. Educ. Res.* 63 (5), 679–695.
- Aronson, B., Meyers, L., Winn, V., 2020. "Lies my teacher [educator] still tells": using critical race counternarratives to disrupt whiteness in teacher education. *Teach. Educ.* 55 (3), 300–322. <https://doi.org/10.1080/08878730.2020.1759743>.
- Aronson, B.A., 2020. From teacher education to practicing teacher: what does culturally relevant praxis look like? *Urban Educ.* 55 (8–9), 1115–1141. <https://doi.org/10.1177/0042085916672288>.
- Arvaja, M., 2016. Building teacher identity through the process of positioning. *Teach. Teach. Educ.* 59, 392–402.
- Avraamidou, L., 2019. Stories we live, identities we build: how are elementary teachers' science identities shaped by their lives experiences? *Cult. Stud. Sci. Educ.* 14, 33–59.
- Beijaard, D., Meijer, P.C., Verloop, N., 2004. Reconsidering research on teachers' professional identity. *Teach. Teach. Educ.* 20 (2), 107–128.
- Benesch, S., 2018. Emotions as agency: feelings, rules, emotional labor, and english language teachers' decision-making. *System* 79, 60–69.
- Bensimon, E.M., 2007. The underestimated significance of practitioner knowledge in the scholarship on student success. *Rev. High. Educ.* 30 (4), 441–469.
- Berger, J.L., Lê Van, K., 2019. Teacher professional identity as multidimensional: mapping its components and examining their associations with general pedagogical beliefs. *Educ. Stud.* 45 (2), 163–181.
- Boaler, J., 2003. Studying and capturing the complexity of practice—the case of the. *Int. Group Psychol. Mathemat. Educ.* 1, 3–16.
- Brown, S., Everson, J., 2019. Belonging, becoming, owning: the interplay between identity and self-esteem in trainee teachers in post-compulsory education and training. *Res. Post-Compulsory Educ.* 24 (2–3), 231–250.
- Buchanan, R., 2014. Teacher identity & agency in an era of accountability. *Teachers Teaching* 21 (6), 700–719.
- Burgess, J., Rowsell, J., 2020. Transcultural-affective flows & multimodal engagements: reimagining pedagogy and assessment with adult language learners. *Lang. Educ.* 34 (2), 173–191. <https://doi.org/10.1080/09500782.2020.1720226>.
- Cattley, G., 2007. Emergence of professional identity for the pre-service teacher. *Int. Educ. J.* 8 (2), 337–347.
- Clandinin, D.J., Connelly, F.M., Bradley, J.G., 1999. Shaping a professional identity: stories of educational practice. *McGill J. Educ.* 34 (2), 189.
- Cochran-Smith, M., Grudnoff, L., Orland-Barak, L., Smith, K., 2020. Educating teacher educators: international perspectives. *N. Educ.* 16 (1), 5–24.
- Cochran-Smith, M., 2003. Learning & unlearning: the education of teacher educators. *Teaching Teacher Educ. Int. J. Scholarsh. Stud.* 19 (1), 5–28.
- Darvin, R., Norton, B., 2015. Identity and a model of investment in applied linguistics. *Ann. Rev. Linguist.* 35, 36–56.
- Dassa, L., Derosé, D.S., 2017. Get in the teacher zone: a perception study of preservice teachers and their teacher identity. *Issues Teach. Educ.* 26 (1), 101–113.
- Day, C., Kington, A., Stobart, G., Sammons, P., 2006. The personal and professional selves of teachers: stable and unstable identities. *Br. Educ. Res. J.* 32 (4), 601–616.
- Dinham, J., Chalk, B., Beltman, S., Glass, C., Nguyen, B., 2017. Pathways to resilience: how drawings reveal pre-service teachers' core narratives underpinning their future teacher-selves. *Asia Pac. J. Teach. Educ.* 45 (2), 126–144.
- Edwards, G., Blake, A., 2007. Disciplining the practice of creative inquiry: the suppression of difference in teacher education. *Int. J. Res. Method Educ.* 30 (1), 33–55.
- Ellis, S., Smith, V., 2017. Assessment, teacher education & the emergence of professional expertise. *Literacy* 51, 84–93.
- Estola, E., 2003. Hope as work—Student teachers constructing their narrative identities. *Scand. J. Educ. Res.* 47 (2), 181–203.
- Fairclough, N., 2001. *Language and Power*. Pearson Education.
- Farnsworth, V., Kleanthous, I., Wenger-Trayner, E., 2016. Communities of practice as a social theory of learning: a conversation with Etienne Wenger. *Br. J. Educ. Stud.* 64 (2), 139–160.
- Gaudelli, W., Ousley, D., 2009. From clothing to skin: identity work of student teachers in culminatingfield experiences. *Teach. Teach. Educ.* 25 (6), 931–939.
- Gibbons, M., 2000. Mode 2 society and the emergence of context-sensitive science. *Sci. Publ. Pol.* 27 (3), 159–163.
- Glazzard, J., Rose, A., 2020. The impact of teacher well-being and mental health on pupil progress in primary schools. *J. Publ. Ment. Health* 19 (4), 349–357. <https://doi.org/10.1108/JPMH-02-2019-0023>.
- Gonçalves, T.N.R., Azevedo, N.R., Alves, M.G., 2013. Teachers' beliefs about teaching & learning: an exploratory study. *Educ. Res. J.* 2 (1), 54–70.

- Govender, N., 2019. Critical literacy and critically reflective writing: navigating gender and sexual diversity. *Engl. Teach. Pract. Critiq.* 18 (3), 351–364.
- Greenwald, A.G., Krieger, L.H., 2006. Implicit bias: scientific foundations. *Calif. Law Rev.* 94 (4), 945–967.
- Gupta, A., 2020. Preparing teachers in a pedagogy of third space: a postcolonial approach to contextual & sustainable early childhood teacher education. *J. Res. Child. Educ.* 34 (1), 43–58. <https://doi.org/10.1080/02568543.2019.1692108>.
- Hall, D., McGinity, R., 2015. Conceptualizing teacher professional identity in neoliberal times: resistance, compliance and reform. *Educ. Pol. Anal. Arch.* 23, 88.
- Hall, K., Conway, P.F., Murphy, R., Long, F., Kitching, K., O'sullivan, D., 2012. Authoring oneself and being authored as a competent teacher. *Ir. Educ. Stud.* 3 (2), 103–117.
- Haniford, L.C., 2010. Tracing one teacher candidate's discursive identity work. *Teach. Teach. Educ.* 26 (4), 987–996.
- Hinojosa Pareja, E.F., López López, M.C., 2019. Interculturality & teacher education: a study from pre-service teachers' perspectives. *Austr. J. Teach. Educ.* 43 (3), 74–92.
- Hökkä, P., Rautiainen, M., Silander, T., Eteläpelto, A., 2019. Collective agency-promoting leadership in Finnish teacher education. In: *International Research, Policy and Practice in Teacher Education*. Springer, Cham, pp. 15–29.
- Holland, D.C., Lachicotte Jr., W., Skinner, D., Cain, C., 2001. *Identity and Agency in Cultural Worlds*. Harvard University Press.
- Hong, J., Cross Francis, D., 2020. Unpacking complex phenomena through qualitative inquiry: the case of teacher identity. *Educ. Psychol.* 55 (4), 208–219. <https://doi.org/10.1080/00461520.2020.1783265>.
- Hong, J.Y., 2010. Pre-service and beginning teachers' professional identity and its relation to dropping out of the profession. *Teach. Teach. Educ.* 26, 1530–1543.
- Jaggar, A.M., 1988. Love and knowledge: emotion in feminist epistemology. *Inq.* 32, 151–176.
- Janks, H., 2010. *Literacy and Power*. Routledge, London & New York.
- Johnson, K.E., Golombek, P.R., 2020. Informing and transforming language teacher education pedagogy. *Lang. Teach. Res.* 24 (1), 116–127. <https://doi.org/10.1177/1362168818777539>.
- Joseph, D., 2019. Trying to keep afloat: early career teaching in an Australian setting. *Issues Educ. Res.* 29 (2), 444–457.
- Kangas-Dick, K., O'Shaughnessy, E., 2020. Interventions that promote resilience among teachers: a systematic review of the literature. *Int. J. Sch. Educ. Psychol.* 8 (2), 131–146.
- Karaolis, A., Philippou, G.N., 2019. Teachers' professional identity. In: *Affect and Mathematics Education*. Springer, Cham, pp. 397–417.
- Kocabaş-Gedik, P., Ortaçtepe Hart, D., 2020. "It's not like that at all": a poststructuralist case study on language teacher identity and emotional labor. *J. Lang. Ident. Educ.* <https://doi.org/10.1080/15348458.2020.1726756>.
- Kubanyiova, M., 2020. Language teacher education in the age of ambiguity: educating responsive meaning makers in the world. *Lang. Teach. Res.* 24 (1), 49–59. <https://doi.org/10.1177/1362168818777533>.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lassonde, C.A., Israel, S.E., 2008. *Teachers Taking Action*. International Reading Association, Newark, DE.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge university press.
- Lavina, L., Lawson, F., 2019. Weaving forgotten pieces of place and the personal: using collaborative auto-ethnography and aesthetic modes of reflection to explore teacher identity development. *Int. J. Educ. Arts* 20 (6).
- Lawrence, L., Nagashima, Y., 2020. The intersectionality of gender, sexuality, race, and native-speakerness: investigating ELT teacher identity through duoethnography. *J. Lang. Ident. Educ.* 19 (1), 42–55. <https://doi.org/10.1080/15348458.2019.1672173>.
- Leeferink, H., Koopman, M., Beijard, D., Schellings, G.L., 2019. Overarching professional identity themes in student teacher workplace learning. *Teachers Teaching* 25 (1), 69–89.
- Lieberman, J., 2009. Reinventing teacher professional norms and identities: the role of lesson study and learning communities. *Prof. Dev. Educ.* 35 (1), 83–99.
- Lofthouse, R., 2014. Engaging in educational research and development through teacher practitioner enquiry: a pragmatic or naïve approach? *Educ. Today* 64 (4), 13–19.
- Loh, C.E., Liew, W.M., 2016. Voices from the ground: the emotional labour of english teachers' work. *Teach. Teach. Educ.* 55, 267–278.
- Lundin, M., Lantz-Andersson, A., Hillman, T., 2020. Teachers' identity work in a professional Facebook group. *J. Inf. Technol. Educ.* 19, 205–222.
- Maclean, R., White, S., 2007. Video reflection and the formation of teacher identity in a team of pre-service and experienced teachers. *Reflect. Practice* 8 (1), 47–60.
- McKay, L., 2019. Supporting intentional reflection through collage to explore self-care in identity work during initial teacher education. *Teach. Teach. Educ.* 86, 102920.
- Medina, R.A., 2020. Designing, facilitating, and supporting for the critical engagement of self-reflection, critical dialogue, and justice-oriented teaching. *Teach. Educ. Q.* 47 (1), 117–122.
- Mockler, N., 2010. Beyond 'what works': understanding teacher identity as a practical and political tool. *Teachers Teaching* 17 (5), 517–528. <https://doi.org/10.1080/13540602.2011.602059>.
- Moghadam, H.H., Adel, S.M.R., Ghaniabadi, S., Amirian, S.M.R., 2020. A Bourdieusian analysis of the educational field and professional identity of EFL teachers: a hermeneutic phenomenological analysis. *Qual. Res. J.* 19 (2), 156–170. <https://doi.org/10.1108/QRJ-12-2018-0002>.
- Moje, E.B., 2007. Developing socially just subject-matter instruction: a review of the literature on disciplinary literacy teaching. *Rev. Res. Educ.* 31, 1–44.
- Moje, E.B., 2008. Foregrounding the disciplines in secondary literacy teaching and learning: a call for change. *J. Adolesc. Adult Literacy* 52 (2), 96–107.
- Murphy, R., Wyness, G., 2020. Minority report: the impact of predicted grades on university admissions of disadvantaged groups. *Educ. Econ.* 1–18.
- Nichols, S.L., Schutz, P.A., Rodgers, K., Bilica, K., 2017. Early career teachers' emotion and emerging teacher identities. *Teachers Teaching* 23 (4), 406–421.
- Norton Pierce, B., 1995. Social identity, investment, and learning language. *Tesol Q.* 29 (1), 9–31.
- O'Connor, K.E., 2006. "You choose to care": teachers, emotions and professional identity. *Teach. Teach. Educ.* 24, 117–126.
- OECD, 2019. *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS. OECD Publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>.
- OECD, 2020. *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS. OECD Publishing, Paris. <https://doi.org/10.1787/19cf08df-en>.
- O'Keeffe, S., Skeritt, C., 2021. Exploring teacher identity using poststructural tools. *Int. J. Res. Method Educ.* 44 (2), 179–192.
- Orland-Barak, L., Wang, J., 2021. Teacher mentoring in service of preservice teachers' learning to teach: conceptual bases, characteristics, and challenges for teacher education reform. *J. Teach. Educ.* 72 (1), 86–99.
- Pomerantz, F., Ippolito, J., 2015. Power tools for talking: custom protocols enrich coaching conversations. *J. Staff Dev.* 36 (1), 40–43.
- Ritaco, M.J., Bolivar, A., 2019. A dual and discontinuous professional identity: school principals in Spain. *Int. J. Educ. Manag.* 33 (5), 806–827.
- Rodrigues, F., Mogarro, M.J., 2019. Student teachers' professional identity: a review of research contributions. *Educ. Res. Rev.* 28, 100286.
- Sachs, J., 2001. Teacher professional identity: competing discourses, competing outcomes. *J. Educ. Pol.* 16 (2), 149–161.
- Schneider Kavanagh, S., Daniels, K.A., 2020. Practicing justice, justifying practice: toward critical practice teacher education. *Am. Educ. Res. J.* 57 (1), 69–105. <https://doi.org/10.3102/0002831219848691>.
- Skinner, B., Leavey, G., Roth, D., 2018. Managerialism and teacher professional identity: impact on well-being among teachers in the UK. *Educ. Rev.* 1–16.
- Vasquez, V.M., Janks, H., Comber, B., 2019. Critical literacy as a way of being and doing. *Lang. Arts* 96 (5), 300–311.
- Vavrus, M., 2009. Sexuality, schooling, and teacher identity formation: a critical pedagogy for teacher education. *Teach. Teach. Educ.* 25 (3), 383–390.
- Wang, X.Q., Zhu, J., He, X., Hu, Y.Y., Li, F., Liu, M.F., Ye, B., 2019. Affective and cognitive empathy in pre-teachers with strong or weak professional identity: an ERP study. *Front. Hum. Neurosci.* 13, 175.
- Webb, M., 2005. Becoming a secondary-school teacher: the challenges of making teacher identity formation a conscious, informed process. *Issues Educ. Res.* 5 (2), 206–224.
- Wenger, E., 1999. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge university press.

- Wenger, E., 2010a. Conceptual tools for CoPs as social learning systems: boundaries, identity, trajectories and participation. In: *Social Learning Systems and Communities of Practice*. Springer, London, pp. 125–143.
- Wenger, E., 2010b. Communities of practice and social learning systems: the career of a concept. In: *Social Learning Systems and Communities of Practice*. Springer, London, pp. 179–198.
- Wenger-Trayner, E., Wenger-Trayner, B., 2015. Learning in landscapes of practice. In: *Learning in Landscapes of Practice. Boundaries, Identity, and Knowledgeability in Practice-Based Learning*, pp. 13–30.
- Widodo, H.P., Fang, F., Elyas, T., 2020. The construction of language teacher professional identity in the Global Englishes territory: 'we are legitimate language teachers'. *Asian Englishes* 1–8.
- Winter, C., 2000. The state steers by remote control: standardising teacher education. *Int. Stud. Sociol. Educ.* 10 (2), 153–175. <https://doi.org/10.1080/09620210000200059>.
- Zembylas, M., 2003. Emotions & teacher identity: a poststructuralist perspective. *Teachers Teaching* 9 (3), 213–238.
- Zembylas, M., 2007. Emotional ecology: the intersection of emotional knowledge & pedagogical content knowledge in teaching. *Teach. Teach. Educ.* 23, 355–367.

Teacher education for diverse learners

Katrina Liu^a, Richard Miller^b, and Arnetha F. Ball^c, ^a School of Teacher Education, Department of Teaching and Learning, University of Nevada, Las Vegas, Las Vegas, NV, United States; ^b School of Music University of Nevada, Las Vegas, NV, United States; and ^c Graduate School of Education, Stanford University, Stanford, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	356
Student diversity	356
Preparing future teachers to teach diverse students	357
Preparing white Teachers for diverse students	357
Multicultural education	357
Culturally relevant pedagogy	358
Culturally sustaining pedagogy	358
Recruiting, preparing, and retaining teachers of color to teach diverse learners	359
Addressing the roots of inequity through transformative action	360
Transformative Teacher Education for Educational Equity	361
The roadmap for transformation: Critical Race Theory	361
The vehicle for transformation: culturally relevant and sustaining community-based teacher education	362
The engine for transformation: critical reflection and generative change for transformative action	362
Conclusion	365
References	365

Introduction

Over the past fifty years, teacher educators in the United States have come to acknowledge the diversity among K-12 student populations, as well as the need to take that diversity into account in both K-12 teaching and teacher education in order to achieve educational equity and broader social justice for all students and their communities. Unfortunately, efforts to address these realities through teacher education, when taken as a whole, have ranged from minimally effective to actively harmful. In this article we briefly examine efforts by teacher educators in the US to address student diversity, educational equity, and social justice with the goal of drawing lessons that may be of value globally. We begin by considering what teacher educators mean by “student diversity,” then review and critique the primary methods teacher educators have developed to attempt to prepare teachers who can work with and for a diverse student population to achieve educational equity and social justice. We end by considering the common features of the challenges and failures of these approaches to teacher education, and suggest a transformative teacher education framework as a key to surmounting these challenges by synthesizing Critical Race Theory as the roadmap, culturally relevant and sustaining community-based teacher education as the vehicle, and critical reflection for transformative learning as the engine.

Student diversity

A review of literature on education reform in the US shows that the 1969 Othaniel Smith Report, *Teachers for the Real World*, is one of the earliest recognitions of student diversity, although the word “diversity” does not appear in that work. Rather, Smith wrote of “disadvantaged youth,” a term that moved from a sociological association with poverty to encompass other kinds of disadvantage in the 1960s. According to Smith, disadvantaged youth are: (1) denied choice in careers; (2) denied any semblance of political power and even denied the opportunity to learn democratic decision making; (3) stigmatized as culturally disadvantaged, excluded from the broader cultural activities, and have their own culture denigrated; (4) denied inter- and intra-personal competence, segregated from the larger society, and labeled as socially or emotionally disturbed if they protest against the inequity of the situation (pp. 3–4). Smith’s definition did not specify groups that could be labeled “disadvantaged,” instead focusing on the social costs they face by being so labeled. More recent notions of diversity tend to focus on a list of the ways in which groups diverge from a norm that itself remains undefined. For example, Zeichner (1996) lists social class, ethnicity, race, and language (p. 133). Banks and Banks (2010) list gender, social class, ethnic, racial, or cultural characteristics (p. 3). The National Council for Accreditation of Teacher Education (NCATE, 2006) defines “diversity” as “differences among groups of people and individuals based on ethnicity, race, socioeconomic status, gender, exceptionalities, language, religion, sexual orientation, and geographical area” (NCATE, 2006).

For the purposes of this article, we define diversity in terms of the spectrum of educational disadvantage articulated by Smith (1969), understanding that the basis for this spectrum, while reflecting global factors that produce some commonality from place to place (colonialism, neo-liberal political economy, White supremacy, and so forth) is primarily determined by regional and local socio-political realities. The primary examples in this article are drawn from the US which, primarily constructs the spectrum of disadvantage in terms of race, gender, and language.

According to statistics kept by the National Center for Education Statistics, the US K-12 public school student population continues to become more diverse. White students represented 88% of the school population in 1972, but as of 2020 have become a minority, representing 46% of the total public student population. This number is expected to drop to 43% by 2030. Significantly, Hispanic students represented 28% of the total public student population, followed by Black student (15%) and Asian students (5%) (NCES, 2022a). At the same time, the nation's teaching force remains primarily White. In 2017–2018, 79% of America's teachers were White (NCES, 2022b). Hispanic teachers represent only 9% of the teacher workforce, followed by Black teachers (7%) and Asian teachers (2%) (NCES, 2022b). Similarly, the number of students labeled as English learners (ELs) or speakers of non-standard variants of English is significant and growing. In fall 2019, the percentage of US public school students who were English learners (ELs) was 10.4% (5.1 million students), higher than in fall 2010 (9.2%, or 4.5 million students). In some states such as Texas and California, students who were ELs were approaching 20%.

In spite of data clearly showing that the growth in diversity has emphasized race, ethnicity, and linguistic diversity, research indicates that teachers tend to avoid using race and ethnicity in their conception of diversity, preferring socially safer terms such as “culture” (Ladson-Billings, 2009). Moreover, teachers' ideas of culture tend to be “reductive,” essentializing minoritized groups through stereotyping (Howard, 2013, p. 2001) in a way that reinforces a “deficit model” (Valencia, 2010) that lays the blame for educational inequities on the students, their families, and their communities instead of on schooling and the larger society. It is likely that these tendencies are also found in other countries that use other criteria to minoritize portions of their population—avoiding key criteria used to disadvantage students in favor of a focus on broad generalities, and constructing deficit views of those minoritized students, families, and communities. Teacher educators in the US, seeing these problems, have worked to combat these tendencies, albeit with limited success, by preparing new teachers with the tools to recognize the habits of mind and patterns of acting that systematically disadvantage minoritized students, families, and communities (Liu and Ball, 2019). On the belief that understanding the challenges, successes, and failures of US teacher educators' efforts might be instructive for teacher educators in other countries, we now examine attempts to prepare future teachers to teach a diversity of students.

Preparing future teachers to teach diverse students

In the US, there have been two common approaches to teacher education for equity and social justice across diverse communities: (1) educating prospective teachers from dominant groups to work with and for minoritized communities, and (2) recruiting, preparing, and retaining prospective teachers from minoritized groups to work with and for their communities (Ball et al., 2021, Liu, 2017; Liu and Ball, 2019; Gist, 2017; Gist et al., 2014; Ingersoll and May 2011, 2016; Kohli, 2019; Philip, 2013; Sleeter, 2008). Because the key elements of diversity in the US are race, ethnicity, and linguistic diversity, these two approaches are typically described as “preparing White teachers for diverse students” and “recruiting, preparing, and retaining teachers of color.”

Preparing white Teachers for diverse students

In the US, there are three major approaches to preparing White teachers for diverse students, the trajectory of which may be instructive in other parts of the world in which the dominant and minoritized populations are structured along lines other than race and ethnicity: Multicultural Education, Culturally Responsive Pedagogy, and Culturally Sustaining Pedagogy.

Multicultural education

Multicultural education arose more or less simultaneously in the 1970s in the US, the UK, and Australia (Banks, 1974; Grant, 1975; Gay, 1977; Bullivant, 1972) in response to a general concern that K-12 education was too narrowly focused on White middle-class suburban norms and experiences, and therefore students in minoritized groups were not being equitably educated. The core idea is to bring a diverse range of experiences into the classroom in order to better include minoritized students while addressing fundamental ignorance about minoritized groups endemic among students (and teachers) in the dominant group. Several broad literature reviews by Sleeter and Grant (1987) and Bennett (2001) demonstrate the enormous scope of research and pedagogy generated over several decades under the rubric of multicultural education—they do not, however address teacher preparation. They do demonstrate that, in spite of the use of the word “culture,” in the beginning multicultural education was more closely linked with notions of race, ethnicity, and linguistic diversity, and had a somewhat contentious relationship with the identity politics of gender, sexual orientation, and disability. As early as 1974, Banks warned against the expansion of multicultural education beyond race and ethnicity, arguing that race and ethnicity are so fundamental to American inequity that they must be addressed first, and that bringing in other ways people identify in groups would harm the causes of both sides (Banks, 1974). A year later, Grant warned that instrumentalizing multicultural education as a set of skills transmissible through professional development would not address the elimination of race-based inequities built into the American school system (Grant, 1975).

Notwithstanding early warnings by Banks and Grant, and repetitions over the years (see Milner, 2017), the focus of multicultural education in both practice and research has shifted away from race and ethnicity as it has become a standard pedagogical subject in US teacher education programs. As a result, the majority of teacher education programs—and the majority of research on multicultural teacher education—focuses on the efficacy of teaching White teachers to teach “students of diversity” by changing their mindset through some combination of self-reflection and exposure to students of color. How to accomplish this has remained unclear; Banks has long advocated teaching “multiple dimensions” of multicultural education in his influential textbook (Banks, 1995),

whereas Sleeter argued two decades later for targeting specific resistant practices of White preservice teachers, focusing their learning about multicultural education on understanding interest convergence, color blindness, and experiential knowledge—concepts drawn from Critical Race Theory (Sleeter, 2017).

Regardless of the specific conception of multicultural education, recent examples of coursework in the literature still emphasize self-reflection and/or exposure to students of color. For example, Davis (2009) used “critical consultation” between White preservice teachers and students of color in their placements to encourage preservice teachers to discard their habitual deficit thinking. Souto-Manning (2010) built a course using theatrical role-playing methods to raise the awareness of racism among White preservice teachers, and Matias and Mackay (2016) constructed a course focused on self-reflection to try to accomplish the same ends. Several scholars advocate autobiography or other narrative inquiry methods to encourage self-reflection around multicultural issues (Fits Fulmer, 2012; Rieger, 2015). In many teacher education programs, as Grant and Koskela (1986), Zeichner (1993), and others have observed, multicultural teacher education tends to be handled with a single course (or inservice workshop) rather than being integrated into the entire teacher education curriculum. Given the centrality of self-reflection and exposure to students of color to these multicultural teacher education programs, it is reasonable to predict that this “one-shot approach” (Zeichner, 1993) would have limited effectiveness—and, indeed, several scholars (Amos, 2010; Evans-Winters and Hoff, 2011; Philip and Benin, 2014) suggest that, far from infusing White teachers with a sense of social justice, isolating multicultural content from the rest of the curriculum reinforces unexamined assumptions about race and ethnicity.

Parallel to the multicultural education courses in teacher education programs, research since the 1990s on the efficacy of such courses has focused on assessing changes in White teachers’ attitudes toward students of colors (see, for example, Gomez, 1994; Goodwin, 1994; Deering and Stanutz, 1995). This research relies primarily on the self-reports of White teachers, with researchers interpreting statements made by teachers in surveys, interviews, and professional portfolios. For example, Horton and Scott (2004) employed Freire’s generative themes to examine the responses of four White preservice teachers to a multicultural education program, locating each participant in one of the stages of racial identity development suggested by Helms (1990). They found that, far from encouraging notions of pluralism, short-term multicultural education courses seemed to stimulate even more rigid notions of White identity. Amos (2010), working with a framework of Critical Whiteness theory, observed similar behavior, and went so far as to title his article “They Don’t Want To Get it!” using a quote from one of the preservice teachers of color about their White classmates to summarize the problem: White preservice teachers dominate discussion in multicultural education classes, in part because the structures of teacher education encourage such dominance (Rieger, 2015).

In sum, it appears that consciousness raising through dedicated coursework and exposure to students of color has not produced the desired transformation of White teachers that early multicultural education theorists once envisioned. Moreover, even when White teachers claim transformation, their actual classroom behavior has not been examined systematically.

Culturally relevant pedagogy

Students of color and their families have been historically viewed through deficit approaches, their language, literacies, and cultures considered deficiencies requiring replacement with schooling “standards” embodying the language, literacy, and cultural practices of the dominant majority (Leonardo, 2004; Paris and Ball, 2009). The result in the US is a “grammar to normalize the standards of white supremacy as the standard for all sorts of everyday transactions” (Bonilla-Silva, 2012, p. 1) that determine what to teach, how to teach it, and how to assess student learning. The learning outcomes of students of color based on this curriculum and pedagogy are then used as evidence to further support a deficit framework and justify their exclusion from learning opportunities in areas such as STEM (Riegle-Crumb et al., 2019). Problems such as these stimulated concerns that multicultural education placed too much emphasis on fitting students of color into the standards of Whiteness, which hindered the ability of students of color to achieve to their academic potential. In response to this concern, Gloria Ladson-Billings penned a seminal article “Toward a Theory of Culturally Relevant Pedagogy” (1995), drawing from the practices of successful teachers of African American students to develop a coherent theoretical and pedagogical foundation for academic achievement. This *culturally relevant pedagogy* attempted to make teaching and learning relevant and responsive to the languages, literacies, and cultural practices of students of color to “produce students who can achieve academically ... who demonstrate cultural competence, and ... who can both understand and critique the existing social order” (p. 474). Culturally relevant pedagogy also calls for students to develop critical perspectives that challenge societal inequalities. Since Ladson-Billings called for culturally relevant pedagogy the approach has been elaborated via a range of frameworks (Brown-Jeffy and Cooper, 2011) and textbooks (Ladson-Billings and Paris, 2021) into an important element of many teacher education programs.

Culturally sustaining pedagogy

Culturally relevant pedagogy was a step forward in reducing the deficit view of students of color, and bringing their life experiences into multicultural education as an active element in teaching and learning. However, as happened earlier to multicultural education, over time culturally responsive pedagogy as implemented in teacher education programs failed to live up to its promise. Ladson-Billings observed this firsthand in teacher education, noting that many practitioners, and those who claim to translate research to practice, seem stuck in very limited and superficial notions of culture Even when people have demonstrated a more expansive knowledge of culture, few have taken up the sociopolitical dimensions of the work, instead dulling its critical edge or omitting it altogether.” (Ladson-Billings, 2014, p. 77).

The result was “teachers who had good intentions toward the students and wanted to embrace culturally relevant pedagogy” but who “rarely pushed students to consider critical perspectives on policies and practices that may have direct impact on their lives and

communities” (Ladson-Billings, 2014, pp. 77–78). In other words, culturally relevant pedagogy had become another simplistically applied classroom technique, accepted in theory by preservice and inservice teachers, but not implemented in practice.

At about the same time, and seeing many of the same issues with both multicultural education and culturally relevant pedagogy as practiced in teacher education as well as K-12 education, Django Paris (2012) argued that educators should not just work to help minoritized students adapt to the culture of the dominant majority, but should also “ensure the valuing and maintenance of our increasingly multiethnic and multilingual society” (p. 94). He therefore suggested the idea of *culturally sustaining pedagogy*, which focuses on supporting the learning of students of color in their home languages and cultures as the base for academic achievement in both the dominant and minoritized cultures. In terms of teacher education, culturally sustaining pedagogy attempts to

offer pre-service and in-service teachers a way of both naming and conceptualizing the need to meaningfully value and maintain the practices of their students in the process of extending their students’ repertoires of practice to include dominant language, literacies, and other cultural practices [it] seeks to perpetuate and foster—to sustain—linguistic, literate, and cultural pluralism as part of the democratic project of schooling. (p. 94).

Paris has continued to elaborate culturally sustaining pedagogy both individually (Paris, 2021) and in collaboration with H. Samy Alim (Paris and Alim, 2017) the approach has been applied by other educators and has informed an effort to revitalize culturally responsive teaching (Ladson-Billings, 2014).

Recruiting, preparing, and retaining teachers of color to teach diverse learners

The approaches to preparing teachers for a diverse student population described above were all developed under the reality that the majority of K-12 teachers in the US are White; indeed, the current teaching force, although slightly more diverse than in past years, still remains about 79% White (NCES, 2022b). However, efforts have been made to recruit, prepare, and retain teachers of color—an approach that seeks to change the makeup of the teacher corps, and not accept the overwhelming Whiteness of education as natural. Villegas and Irvine (2010) identified three research-based rationales for increasing the supply of teachers of color: (1) teachers of color serve as role models for all students; (2) teachers of color can improve the academic outcomes and school experiences of students of color; and (3) more teachers of color are needed to reduce the acute shortage of educators for high-needs urban schools (pp. 176–186). Research shows that teachers of color in the US can benefit all students, providing caring, decriminalizing and humanizing education to students of color historically subjected to intergenerational legacies of violence and punishment often leading to incarceration (Basile, 2021; Bryan, 2020, 2021; Siddle Walker, 2001). Moreover, recruiting, preparing, and retaining teachers of color can lessen the multigenerational socialization of White oppression over students of color passed down from White teachers to White students (Bryan, 2021). Further, Villegas and Irvine (2010) found that teachers of color have more favorable views of students of color, as well as greater recognition of their academic potential.

For these reasons, efforts have been made over several decades to recruit, prepare, and retain teachers of color. For example, Villegas et al. (2012) summarized four specific recruitment strategies promoted by the Education Commission of the States (1990) to recruit teachers of color which are still employed today: (1) early recruitment and college entrance support programs targeting promising pre-college students; (2) partnerships between two- and four-year colleges to facilitate the transition of community colleges students into four-year colleges; (3) career ladder programs for paraprofessionals; and (4) financial incentives such as scholarships, forgivable loans, and signing bonuses. Literature indicates that teacher education programs have adopted all of these strategies in the past two decades to recruit prospective teachers of color (see Gist, 2014; Schultz et al., 2011; Toshalis, 2012; Valenzuela, 2017; Villegas and Davis, 2008). Once recruited into teacher education programs, preservice teachers of color are faced with curricula, instructional strategies, and fieldwork constructed in the context of the White, middle-class preservice teachers (Cochran-Smith et al., 2015) who form the bulk of their fellow students and their teacher educators. As a result, teacher education programs tend to make few concessions to the life experiences, cultures, and languages of prospective teachers of color (Gist, 2014, 2017; Kohli, 2019), nor do they address the specific needs of prospective teachers of color in themselves learning to teach a diverse student population (Brown, 2014; Liu, 2020). Instead, the programs have primarily focused retention efforts on providing academic support through tutoring, learning communities, and exam preparation workshops, and in some institutions they social support by identifying cohorts of students of color on which to base cultural events, counseling, and advisement (Gist, 2014) including peer and near-peer mentorship programs (Berry et al., 2008; Carver-Thomas, 2018).

A review of the literature on the recruitment, preparation, and retention of prospective teachers of color primarily reveals that most research focuses on two areas: (1) experiences of prospective teachers of color in teacher education programs; and (2) how teacher educators support prospective teachers to develop awareness of their identity, cultural and linguistic assets, and/or their bias and limitations. The first area focuses on the cultural and social isolation (Ladson-Billings, 2001; Schultz et al., 2008), institutional racism (Gay, 2005; Irizarry, 2007; Villegas and Davis, 2008; Woodson and Pabon, 2016), and ideological and curricular challenges (Milner, 2008; Philip, 2010) that prospective teachers of color experience in teacher education programs. Ladson-Billings (2001), for example, documented the experience of a prospective teacher of color who considered dropping out due to the lack of diversity in the program, remaining only after two other prospective teachers of color joined the program and they were able to support each other in their study. Irizarry (2007) and Kohli (2014) also documented how prospective teachers of color endured institutional racism such as overtly racist comments or actions and racial hierarchies. Similarly, Kohli (2019) studied 11 women

teachers of color about their teacher preparation experiences, finding that they believed their teacher preparation programs not only marginalized them during their preparation, but continued to negatively affect their careers as they entered the field with limited tools to resist racialized institutional structures and practices (p. 10). Gist (2017) exploring the program and pedagogical experiences of teacher candidates of color, found that prospective teachers experience a lack of culturally and linguistically responsive learning in their teacher education programs. In parallel findings, Philip (2013) documented the experience of a prospective teacher of color who entered the teacher education program with a strong commitment to equity and social justice, but found that the strengths gained through college student activism were overlooked and the teacher education program failed to help her understand the nuances, strengths, and limitations of her activist stances within the context of classroom teaching and institutional change within schools.

The second area of research on prospective teachers of color focuses on how teacher educators facilitate prospective teachers of color in developing a deeper understanding of who they are, where they are situated in society, what assets they have, and what biases they bring to education. This has been addressed through critical reflection and collaborative dialogue (Clark and Flores, 2001; DePalma, 2008; Kohli, 2012, 2014; Tellez, 1999). Villegas and Davis (2008) argued that teacher educators need to create a safe space conducive to critical dialogue in the preparation of prospective teachers of color. Similarly, Clark and Flores (2001) focused on how providing pedagogical approaches can support the ethnic identity development of Latinx teacher candidates to increase confidence in their ability to enact culturally responsive pedagogy. Kohli (2014), as well, found that prospective teachers of color felt that critical dialogue on internalized racism within teacher preparation was essential to developing pedagogy that challenged racial inequality. These approaches provide valuable pedagogical guidance for teacher educators regarding how to recognize and cultivate the cultural and linguistic strengths of prospective teachers of color.

Addressing the roots of inequity through transformative action

Considering together the two approaches to preparing teachers for a diverse student population—working to improve the practice and disposition of White teachers and preparing teachers of color—there appear to be clear limits to what either approach can accomplish. Multiple studies (Evans-Winters and Hoff, 2011; Horton and Scott, 2004; Matias and Mackey, 2016; DePalma, 2008) indicate that the habits of privilege on the part of White preservice teachers and the general structures and assumptions of Whiteness in teacher education programs combine to amplify the voices and biases of White preservice teachers and teacher educators and silence those of preservice teachers and faculty of color, reducing or eliminating elements of critique in favor of symbolic incorporation stereotyped notions of “culture.” Moreover, programmatic efforts to recruit and retain prospective teachers of color give little attention to preparing them to teach students of color, assuming that their non-White identity is sufficient to enable classroom success. Thus, although the practice of recruiting more students of color to teacher education programs has, to certain extent, diversified the K-12 teaching force, without further changes to teacher education curriculum, pedagogy, and instruction, is unlikely to produce significant transformation of schooling in the US.

A review of the literature on preparing teachers for a diverse student body indicates three fundamental problems, all stemming from a loss of emphasis on the deep roots of inequity in schooling, and a concomitant lack of focus on transformative action by teachers. First is the fundamental problem of Whiteness in teacher education. Prospective teachers of color need support to navigate the racialized teacher education programs and critique the dominant social order. If teacher educators wish to teach prospective teachers to implement culturally relevant and sustaining pedagogy in their future teaching, they should implement culturally relevant and sustaining pedagogy throughout the programs, building curriculum and field experience on prospective teachers of color’s cultural capital (Yosso, 2005) or funds of knowledge (Moll et al., 1992). In addition, once students of color become teacher education candidates, their programs are often predominantly White (Ladson-Billings, 1995) with curriculum, pedagogy, methods, and field experiences constructed to prepare White teachers, and thus fail to provide the types of social, cultural, emotional, and academic support that students of color need to successfully complete the program (Bryan, 2021; Gist, 2014; Sleeter and Thao, 2007). Prospective teachers of color then encounter marginalization and discrimination in teacher education programs similar to their experiences in the K-12 system (Liu, 2020; Burant, 1999; Parker and Hood, 1995; Sleeter, 2001).

Second is the problem of interest convergence. Although research shows that exposure to diverse students and communities dramatically improves the ability of White prospective teachers to empathize with students of color and implement culturally responsive teaching (Milner, 2006), scholars have also repeatedly observed that race rarely has a central place in American teacher education (Milner, 2003, 2017; Milner and Laughter, 2015).

The third problem is the lack of attention to how preservice teachers implement actions to support diverse learners in their classroom practice (Anderson and Stillman, 2013; Cochran-Smith et al., 2015). By *practice*, we mean “teacher candidates learning how to do the actual tasks of teaching” (Cochran-Smith et al., 2015, p. 117). However, much research that claims to study how teacher preparation influences prospective teachers’ professional practice does not investigate their actions in classroom contexts because practice is considered “teacher candidates being engaged as reflective and inquiring professionals” (Cochran-Smith et al., 2015, p. 117). This type of research has primarily engaged prospective teachers in talking about, analyzing, or reflecting on their teaching, and thus research data on practice is primarily based on prospective teachers’ reflections, self-reports, surveys, and interviews. Based on a thorough review of research on the preparation of teachers for urban and high-needs contexts, Anderson and Stillman (2013) find a disproportionate emphasis on belief and attitude change among prospective teachers with relatively little evidence for the development or change in actual teaching practice. They also find a tendency toward reductive views of culture and context.

Cochran-Smith et al. (2015) argued that “we need more research that goes beyond assuming that changing teacher candidates’ beliefs necessarily leads to different behaviors and actions in their classrooms” (p. 117). Thus teacher educators need to move beyond research on prospective teachers’ attitudes and perspectives about teaching diverse students to research on how they teach, what they teach, and how their teaching impacts the learning of diverse learners based on their day to day practices. This type of classroom practice-based research is time consuming and needs to deliberately take time into consideration in research design in order to see prospective teachers’ growth and transformation. It is, however, vital for the life and learning of marginalized learners, whose school experiences should be taken seriously and be used as first-hand data for teacher educators to transform their curriculum, pedagogy, and instruction.

Transformative Teacher Education for Educational Equity

In order to address the limitations of current approaches to teacher education for diverse student populations, we suggest a comprehensive reimagination of teacher education starting from the theoretical orientation through the institutional structures and practices down to the preparation of the individual preservice teacher. We envision this reimagination as a set of nested practices, each of which performs a different function while supporting the functioning of the other practices, much like physical road travel requires a roadmap, a vehicle, and an engine (Fig. 1). Here, however, the destination is not a physical one but a social and political one: transformation of the educational system to effect educational equity for diverse students.

The roadmap for transformation: Critical Race Theory

Critical Race Theory of education should be integrated into teacher education programs as an overarching framework to help preservice teachers understand race as the central construct to analyze school inequity, deciphering “the social-structural and cultural significance of race in education” (Ladson-Billings and Tate, 1995, p. 50). Ladson-Billings and Tate (1995) argued that although gender and class intersect with race, explanations based on class and gender differences are not powerful enough to explain all the difference in school experience and performance, including but not limited to higher dropouts, higher rates of out-of-school discipline, and failure of Black and Brown students. Critical Race Theory reveals that schools are designed to protect White economic interests through disparate educational opportunities (Ladson-Billings and Tate, 1995).

Critical Race Theory, in contrast to Multicultural Education, prioritizes the experience of minoritized communities of color over that of White teachers, even in the branches of Critical Race Theory that are specifically concerned with the concepts and mechanisms of Whiteness (such as Critical Whiteness Theory). From the standpoint of Critical Race Theory, educational institutions and practices function to perpetuate White supremacy, so the goal of teacher education should be to enable prospective teachers to recognize, acknowledge, and change those institutions and practices. This can be achieved by using history and lived experiences as counternarratives (Miller et al., 2020) to support prospective teachers in analyzing historical practices of slavery and the appropriation of the bodies and labor of enslaved people, the annexing of the indigenous land and genocide of the original inhabitants, the dehumanizing of immigrants, the segregation and resegregation of students of color, and also current unjust practices such as defunding schools in communities of color and overdisciplining students of color to push them into the carceral system.

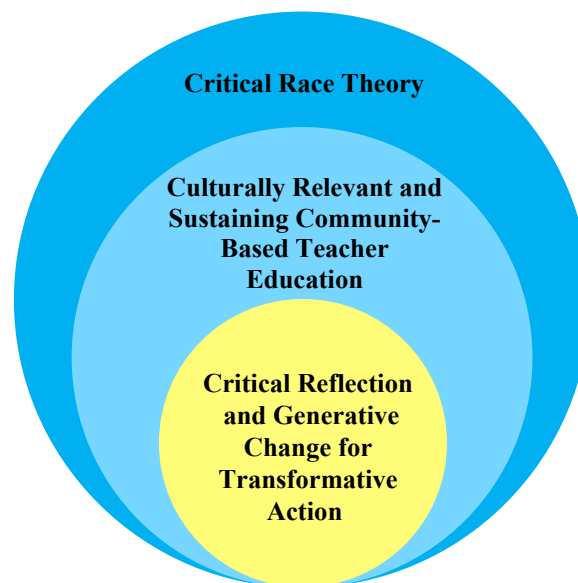


Fig. 1 Transformative teacher education for educational equity.

The vehicle for transformation: culturally relevant and sustaining community-based teacher education

Critical Race Theory, with its emphasis on the ubiquity (and hence centrality) of the lived experience of racism, and its reliance on counter-narrative as both a source of knowledge and a mechanism for critique, should be placed at the center of *community-based teacher education* that deploys culturally relevant and sustaining pedagogy (Ladson-Billings, 2014; Ladson-Billing and Paris, 2021; Paris, 2012) in deep collaboration with local communities of color. All of the approaches to preparing White teachers for diverse student populations discussed so far assume a “traditional,” university-based teacher education model. However, there are approaches to preparing teachers for diverse students that do not assume a university-centric teacher education program, drawing from ideas of Community-Based Learning in an effort to address the needs of communities of color (Ball, 1995; Graue, 2005; Graue and Brown, 2003). As Graue (2005) observes, “there are significant discrepancies in perceptions between school people and parents about school efforts and family involvement in education” (p. 158), further exacerbated by racial and cultural differences between poor parents of color and the largely White middle class teacher corps (Bryk and Schneider, 2002; Meier, 2002; Noguera, 2001; Tschannen-Moran, 2004). Baquedano-López et al. (2013) revealed in their historical analysis that current education reform and practice on parent involvement has been dominated by a discourse incorporating White middle-class values, seeing parents from minoritized communities as problems in need of remedy. This deficit view fails to address the experiences of many parents and students from nondominant backgrounds through intersectional lenses of race, class, and immigration, evaluating nondominant families to abide by White, middle-class values and expectations. When nondominant families do not meet these expectations, educators consider them uninterested in their children’s education and thus blamed them for the school’s failure in educating their children (Ishimaru et al., 2019).

Zeichner et al. (2016) suggested that addressing these problems requires that teachers know about the communities in which their students grow, develop respectful and trusting relations with students’ families and other adults in their communities, and make use of this knowledge and these relationships to support their students’ learning (p. 277). Murrell (2000) calls this type of teacher a “community teacher” who

develops the contextualized knowledge of culture, community, and identity of children and their families as the core of their teaching practice. Community teachers possess the “multicultural competence” they need for accomplished practice in the communities they teach. Community teachers are individuals who typically live and work in the same under resourced urban neighborhoods and communities where students from diverse backgrounds live and go to school (p. 340).

In order to implement community-based teacher education, Baquedano-Lopez et al. (2013) argued that teachers need a paradigm shift from seeing parents as a deficit to parents having agency to advocate for their children and resist inequity, a type of teacher-family-community collaboration that Zeichner et al. (2016) termed as “teacher-family-community solidarity” (p. 278). Solidarity-oriented community-based teacher education has been implemented through community panels involving parents, community leaders as panelists (Zeichner et al., 2016), engaging community members as mentors for prospective teachers (Guillen and Zeichner, 2018; Zeichner et al., 2016; Zygmunt et al., 2018), and facilitating community and teacher educators in co-constructing teacher education programs (Ball et al., 2021; Fickel et al., 2018; Zeichner et al., 2016). These approaches represent a more democratic and transformative framework of teacher education, decentering the power of teacher educators and deconstructing the traditional White curriculum of teacher education programs by integrating the cultures, experiences, and practices of the communities of color.

Restructuring teacher education from the standpoint of Community-Based Learning is predicated on the belief “in the educability of all students, thus providing students from different social backgrounds, with diverse levels of ability and behavioral dispositions, opportunities to learn together to live together” (Lopez et al., 2012, p. 23). Aiming toward social justice and equity for marginalized students and communities, community-based teacher education on the one hand has proved effective in preparing White preservice teachers to teach in diverse communities (Guillen and Zeichner, 2018; Zeichner et al., 2016; Zygmunt et al., 2016; Zygmunt et al., 2018) and on the other hand, has the potential to encourage teacher education programs to actively recruit, prepare and retain teachers of color from and for marginalized communities. The combination, then, of community-based teacher education informed by Critical Race Theory and employing culturally relevant and sustaining pedagogy holds promise to dismantle the institutional and societal attitudes and practices of deficit thinking and a savior mentality toward marginalized students and communities.

The engine for transformation: critical reflection and generative change for transformative action

As critiqued earlier, there is a lack of attention to prospective teachers’ classroom actions; rather, teacher education practice and research tend to focus on changing prospective teachers’ attitudes and perceptions based on their self-reported evidence. In order to support prospective teachers to bridge reflection with action and develop from changed attitudes and perceptions to transformed action, we argue here that critical reflection for transformative learning and generative change are the ideal engine for community-based teacher education. Critical reflection is a hermeneutic approach to individual learning going back at least to Donald Schön (1983), if not John Dewey (1933), and is typically considered to involve repeated reexamination of one’s assumptions about knowledge and understanding, particularly those that are socially, politically, or culturally based. However, as Ball (2012) insisted

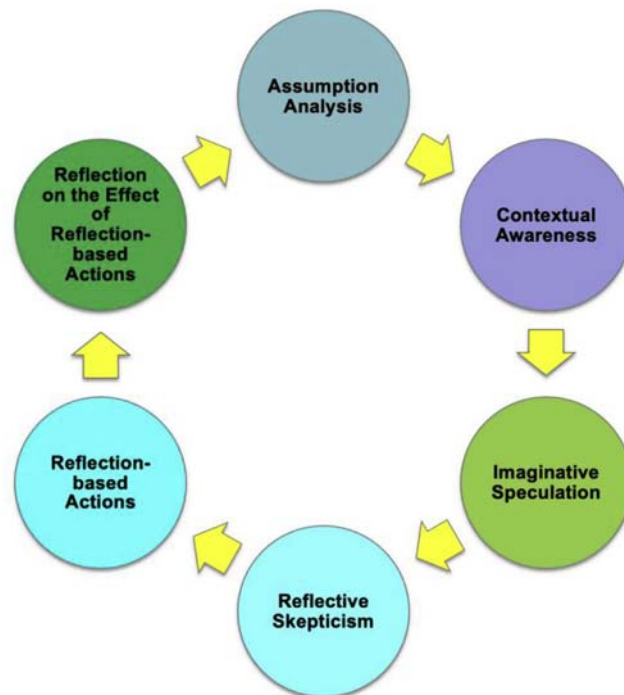


Fig. 2 The hermeneutic cycle of critical reflection for transformative learning. Adapted from Liu (2015, p. 148).

in her AERA Presidential Address, knowing is not enough: There must be a movement from thinking to action, from critique to intervention. For example, Brookfield argued (1995) that “reflection in and of itself is not enough; it must always be linked to how the world can be changed” (p. 217), a position that resonates with Mezirow’s (1990) idea that critical reflection is necessary for changing both how teachers teach and learners learn—what he termed *transformative learning*: “Reflective discourse and its resulting insights alone do not make for transformative learning. Acting upon these emancipatory insights, a praxis is also necessary” (Mezirow, 1990, p. 354). Furthermore, that praxis requires goals which, in the context of preparing teachers for diverse student populations must rest in educational equity and social justice—otherwise the “emancipatory insights” of critical reflection, no matter how important to the individual, cannot translate to collaborative action for community benefit. Accordingly, Liu (2015) links the practice of critical reflection with the goal of transformative action for social justice, defining critical reflection as

a process of constantly analyzing, questioning, and critiquing established assumptions of oneself, schools, and the society about teaching and learning, and the social and political implications of schooling, and implementing changes to previous actions that have been supported by those established assumptions for the purpose of supporting student learning and a better schooling and more [just] society for all children. (pp. 10–11).

Liu (2015) articulated a full processual model of critical reflection for transformative learning based on Brookfield’s work, including a cycle of six steps of assumption analysis, contextual awareness, imaginative speculation, reflective skepticism, reflection-based action, and reflection on reflection-based action (See Liu, 2015, p. 148 for detailed description of the six steps). Fig. 2 provides a visual representation of the hermeneutic cycle of critical reflection for transformative learning.

Critical reflection can effect changes at the classroom level, but what about the larger institutional changes that are also necessary? Here we turn to generativity for an answer. Generativity is the generation of new or novel behavior in problem solving (Epstein, 1996), a complex psychosocial construct that describes how an individual responds to societal demands, inner desires, conscious concerns, beliefs, and commitments. The concept includes productivity and creativity to make the world a better place; it also manifests as a concern for one’s legacy that leads to concrete goals and transformative action.

Research on generativity over the past seven decades suggests generativity is a formal, predictive theory of creative behavior or activity on the part of individuals. Ball (2009), noting that teachers are faced with societal demands to which they must respond creatively, proposed that generativity provides a framework explaining the process by which teachers engage in transformative change of the classroom lives of diverse student populations. Given that nearly 40% of teachers entering the classrooms report feeling unprepared for the challenges that await them, generativity could also guide teacher educators in their transformation of program practices. Ball (2009) thus combined generative theory and teacher efficacy in a model designed to prepare teachers who believe in their potential ability to affect positive change in the lives of their students and who also think in generative ways that result in transformative action in their classrooms. Ball’s model of generative change includes four stages facilitating teachers’ development of a change-oriented mindset, prompting metacognitive awareness, a sense of agency, a sense of advocacy, and the development of a personal voice concerning the education of diverse students. The figure below illustrates Ball’s model as

Model of Generative Change: The Processes Through Which Teachers Develop Voice, Generativity, and Efficacy in Their Development Toward Transformative Change

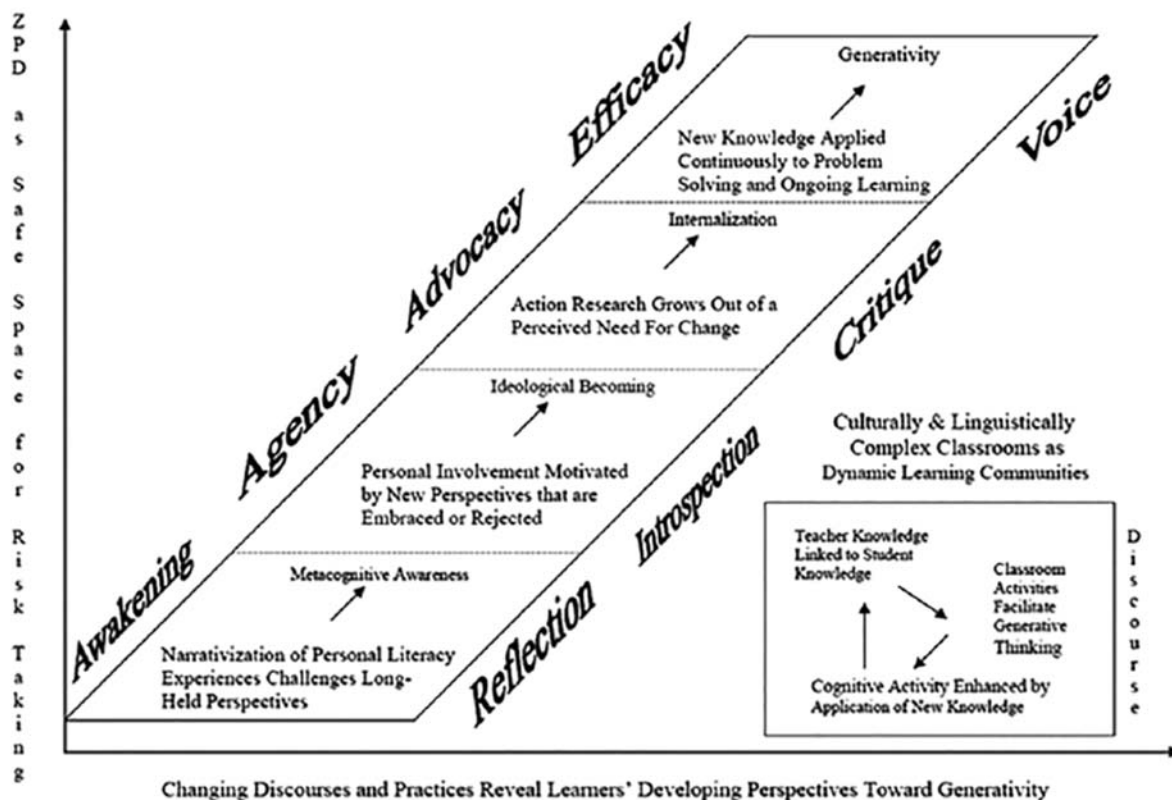


Fig. 3 Model of generative change (Ball, 2009).

a series of steps by which individuals move from metacognitive awakening to their own sense of efficacy to becoming transformative intellectuals (Giroux, 1988) able to reshape curriculum, pedagogy, institutional structures, and classroom practices. The result is the preparation of teachers who are agents of change rather than objects of change, forging new relationships between teachers and students, schools and communities (Fig. 3).

In the context of preparing teachers for a diverse student population through community-based teacher education, it is important first that the steps be infused with insight from Critical Race theory, and second that there is systematic movement from rethinking basic assumptions and obtaining critical awareness to generating new knowledge and taking action to transform learning undertaken in collaboration by both education professionals and the students and families in the community. Combining this collaborative process with Critical Race theory provides two vital affordances to community-based teacher education. First, it provides clear guidance for all parties in identifying the problems afflicting the educational system and actively searching for alternatives to the ways in which students and families are failed by our current educational systems. Second, the frameworks of critical reflection for transformative learning, and generative change when adopted collaboratively by educators and communities, can help all parties generate new knowledge and develop into transformative intellectuals (Ball, 2009; Giroux, 1988). According to Ball (2009), the knowledge becomes generative “when the teacher continues that learning by making connections with his or her students’ knowledge and needs and begins planning the teaching based on what he or she is learning” (p. 48).

Finally, critical reflection for transformative learning and generative change provide valuable frameworks for teacher educators to conduct research on the content of and procedures for prospective teachers’ reflection on teaching diverse students (see, for example, Ball, 2009; 2021; Liu, 2015; 2017; 2020). It further provides a critical lens to guide teacher educators to conduct research on the real actions of prospective teachers in their teaching—whether or not their reflection brings about transformative and generative actions in teaching diverse students. Research based on classroom actions as well as written reflections can help teacher educators gain a better understanding of how to improve their own support for prospective teachers to achieve transformative learning. Therefore, the frameworks, on the one hand, point out the importance of focusing on prospective teachers’ teaching practices and, on the other hand, guide teacher educators to analyze prospective teachers’ actions to determine whether or not they are transformative in terms of teaching diverse learners.

Conclusion

Given the experiences in the US with preparing teachers for diverse student populations, both the positive efforts as well as the limited results, what lessons can be drawn that would be valuable in other countries? First, in the US, race and racism are the central division between the dominant majority and the minoritized populations. Because US racism developed inseparably with more the broader practices of Euro-American colonialism—fundamentally linked through systemic slavery at home and abroad—US-style race and racism are common worldwide, as evidenced by phenomena such as the global practice of Blackface in popular culture, which perpetuates the pervasive stereotypes of Africans and African Diasporans as backward, primitive, highly sexualized, and dangerous (Thelwell, 2020). However, there are parts of the world in which this US.-centric notion of race is not central to the denial of social justice and educational equity. Nevertheless, if we abstract from the US case to think in terms of dominant majorities versus disadvantaged minoritized groups, the insights drawn from the example of the US can still apply. The educational system in the PRC, for example, is primarily constructed with reference to the urban population; rural schools are often treated as an afterthought in terms of funding and specialized curricula (Zhang et al., 2015), and students from rural areas who migrate to urban areas are typically approached from a deficit view (Meng, 2008), much as minoritized students are treated in the US system.

From this standpoint, the limited success of multicultural education stemming from its instrumentalization by teacher education programs and the subsequent weakening (or eliminating) of its critical edge hold a warning for other parts of the world: attempting to address minoritized students by “sensitizing” a system of schooling that remains primarily run by and for the dominant group is unlikely to produce transformative results. Similarly, the role of Critical Race Theory as a vital roadmap to understanding, critiquing, and intervening in unjust schooling in the US cannot be applied unmodified in other parts of the world. However, the basic tenets of Critical Race Theory—that one should focus on the actual lived experience of minoritized groups under schooling designed by and for the dominant groups, and that the fact of that domination conditions the systems and practices throughout schooling—can form the basis for tailoring the other aspects of Critical Race Theory to specific contexts.

Finally, the models of critical reflection and generative change, coupled with the goals of social justice and educational equity, are not tied to the concepts and circumstances found in US education, but are broadly applicable worldwide. Indeed, although the models were built with specific reference to the Euro-American tradition of reflection, critique, and praxis, similar traditions are found around the world whether based in religious disputation (Buddhism and Islam, for example) or more secular educational traditions (such as Confucianism and related philosophical traditions in China and across East Asia). In fact, because the key relationship to be invigorated is that between schools, teacher educators, and the community in which they are based, localization of both inspiration and working method is vital to the successful preparation of teachers for diverse student populations.

References

- Amos, Y.T., 2010. They don't want to get it! Interaction between minority and White preservice teachers in a multicultural education class. *Multicult. Educ.* 17 (4), 31–37.
- Anderson, L., Stillman, J., 2013. Student teaching's contribution to preservice teacher development: a review of research focused on the preparation of teachers for urban and high-needs contexts. *Rev. Educ. Res.* 83 (1), 3–69.
- Ball, A.F., 1995. Community-based learning in urban settings as a model for educational reform. *Appl. Behav. Sci. Rev.* 3 (2), 127–146.
- Ball, A.F., 2009. Toward a theory of generative change in culturally and linguistically complex classrooms. *Am. Educ. Res. J.* 46 (1), 45–72.
- Ball, A.F., 2012. Presidential address: to know is not enough: knowledge, power, and the zone of generativity. *Educ. Res.* 41, 283–293.
- Ball, A.F., Greene, D.M., Friedman, J.S.L., Dray, B.J., 2021. The Trifecta framework: preparing agents of change in urban education. *Urban Educ.* <https://doi.org/10.1177/00420859211023113>.
- Banks, J.A., 1974. *Multicultural Education: In Search of Definitions and Goals*. National Academy of Education, Syracuse.
- Banks, J.A., 1995. Multicultural education and curriculum transformation. *J. Negro Educ.* 64 (4), 390–400.
- Banks, J.A., Banks, C.A.M., 2010. *Multicultural education: Issues and perspectives*, 8th ed. Wiley.
- Basile, V., 2021. Decriminalizing practices: disrupting punitive-based racial oppression of boys of color in elementary school classrooms. *Int. J. Qual. Stud. Educ.* 34 (3), 228–242.
- Baquedano-Lopez, P., Alexander, R.A., Hernandez, S.J., 2013. Equity issues in parental and community involvement in schools: what teacher educators need to know. *Rev. Res. Educ.* 37, 149–182.
- Bennet, C., 2001. Genres of research in multicultural education. *Rev. Educ. Res.* 71 (2), 171–217.
- Berry, B., Montgomery, D., Curtis, R., Hernandez, M., Wurtzel, J., Snyder, J., 2008. *Creating and Sustaining Urban Teacher Residencies: A New Way to Recruit, Prepare, and Retain Effective Teachers in High-Needs Districts*. Aspen Institute and the Center for Teaching Quality, Washington, DC. Available at: <https://files.eric.ed.gov/fulltext/ED512419.pdf>.
- Bonilla-Silva, E., 2012. The invisible weight of whiteness: the racial grammar of everyday life in America. *Mich. Sociol. Rev.* 26 (1), 1–15.
- Brown, K.D., 2014. Teaching in color: a critical race theory in education analysis of the literature on preservice teachers of color and teacher education in the US. *Race Ethn. Educ.* 17 (3), 326–345.
- Brown-Jeffy, S., Cooper, J.E., 2011. Toward a conceptual framework of culturally relevant pedagogy: an overview of the conceptual and theoretical literature. *Teach. Educ. Q.* 65–84. Winter.
- Bryan, N., 2020. Shaking the *bad boys*: troubling the criminalization of black boys' childhood play, hegemonic white masculinity and femininity, and the school playground-to-prison pipeline. *Race Ethn. Educ.* 23 (5), 673–692.
- Bryan, N., 2021. *Toward a BlackBoyCrit Pedagogy: Black Boys, Male Teachers, and Early Childhood Classroom Practices*. Routledge.
- Bryk, A., Schneider, B., 2002. A grounded theory of relational trust. In: Bryk, A., Schneider, B. (Eds.), *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation, pp. 124–144.
- Bullivant, B.M., 1972. The cultural reality of curriculum development. *Educ. News* 13 (9), 14–17.
- Burant, T.J., 1999. Finding, using, and loving voice: a pre-service teacher's experiences in an urban educative practicum. *J. Teach. Educ.* 50 (3), 209–219.
- Carver-Thomas, D., 2018. *Diversifying the Teaching Profession: How to Recruit and Retain Teachers of Color*. Learning Policy Institute, Palo Alto, CA. Available at: <https://learningpolicyinstitute.org/product/diversifying-teaching-profession-report>.
- Clark, E.J., Flores, B.B., 2001. Who am I? The social construction of ethnic identity and self-perceptions in Latino preservice teachers. *Urban Rev.* 33 (2), 69–86.

- Cochran-Smith, M., Villegas, A.M., Abrams, L., Chavez-Moreno, L., Mills, T., Stern, R., 2015. Critiquing teacher preparation research: an overview of the field, Part II. *J. Teach. Educ.* 66 (2), 109–121.
- Davis, D.E., 2009. Preparing white student teachers through a critical consultative model. *Int. J. Prog. Educ.* 5 (2), 1–23.
- Deering, T.E., Stanutz, A., 1995. Preservice field experience as a multicultural component of a teacher education program. *J. Teach. Educ.* 46 (5), 390–394.
- DePalma, L., 2008. The voice of every black person: brining authentic minority voices into multicultural dialog. *Teach. Teach. Educ.* 24 (3), 767–778.
- Dewey, J., 1933. *How We Think: A Statement of the Relation of Reflective Thinking to the Educative Process*. D.C. Heath & Co, Boston.
- Epstein, R., 1996. *Cognition, creativity, and behavior*. Praeger.
- Evans-Winters, V.E., Hoff, P.T., 2011. Aesthetics of white racism in pre-service teacher education: a critical race theory perspective. *Race Ethn. Educ.* 14 (4), 461–479.
- Fickel, L., Abbiss, J., Brown, L., Astall, C., 2018. The importance of community knowl- edge in learning to teach: foregrounding Maori cultural knowledge to support preservice teachers' development of culturally responsive practice. *Peabody J. Educ.* 93 (3), 285–294.
- Fits Fulmer, D.E., 2012. *Autobiographical Meaning Making, Practitioner Inquiry, and White Teachers in Multicultural Education*. Unpublished EdD dissertation. University of Pennsylvania, Philadelphia.
- Gay, G., 1977. Changing conceptions of multicultural education. *Educ. Perspect.* 16 (4), 4–9.
- Gay, G., 2005. Politics of multicultural teacher education. *J. Teach. Educ.* 3 (56), 221–228.
- Giroux, H., 1988. *Teachers as Intellectuals: Toward a Critical Pedagogy of Learning*. Bergin & Garvey.
- Gist, C., 2014. *Preparing Teachers of Color to Teach: Culturally Responsive Teacher Education in Theory and Practice*. Palgrave Macmillan.
- Gist, C., 2017. Voices of aspiring teachers of color: unraveling the double bind in teacher education. *Urban Educ.* 52 (8), 927–956.
- Gist, C.D., Flores, B.B., Claes, L., 2014. Competing theories of change: critical teacher development. In: Sleeter, C., Neal, L., Kumashiro, K. (Eds.), *Addressing the Demographic Imperative: Recruiting, Preparing, and Retaining a Diverse and Highly Effective Teaching Force*. Routledge, pp. 19–31.
- Gomez, M.L., 1994. Teacher education reform and prospective teachers' perspectives on teaching "Other people's" children. *Teach. Teach. Educ.* 10 (3), 319–344.
- Goodwin, A.L., 1994. Making the transition from self to other: what do preservice teachers really think about multicultural education? *J. Teach. Educ.* 45 (2), 119–131.
- Grant, C.A., 1975. *Sifting and Winnowing: An Exploration of the Relationship between Multi- Cultural Education and CBTE*. University of Wisconsin Teacher Corps Associates Program, Madison.
- Grant, C.A., Koskela, R.A., 1986. Education that is multicultural and the relationship between preservice campus learning and field experiences. *J. Educ. Res.* 79 (4), 197–204.
- Graue, E., 2005. Theorizing and describing preservice teachers' images of families and school- ing. *Teach. Coll. Rec.* 107 (1), 157–185.
- Graue, E., Brown, C.P., 2003. Preservice teachers' notions of families and schooling. *Teach. Teach. Educ.* 19 (7), 719–735.
- Guillen, L., Zeichner, K., 2018. A university-community partnership in teacher education from the perspective of community-based teacher educators. *J. Teach. Educ.* 69 (2), 140–153.
- Helms, J., 1990. *Black and White Racial Identity: Theory, Research and Practice*. Greenwood Press.
- Horton, J., Scott, D., 2004. White students' voices in multicultural teacher education preparation. *Multicult. Educ.* 11 (4), 12–16.
- Howard, T.C., 2013. *Black Male(d): Peril and Promise in the Education of African American Males*. Teachers College Press, New York, NY.
- Ingersoll, R., May, H., 2011. Recruitment, Retention and the Minority Teacher Shortage. Consortium for Policy Research in Education, Philadelphia, PA. Available at: http://www.cpre.org/sites/default/files/researchreport/1221_minorityteachershortagereporttr69septfinal.pdf.
- Ingersoll, R., May, H., 2016. Minority Teacher Recruitment, Employment, and Retention: 1987- 2013 (Research Brief). Available at: https://learningpolicyinstitute.org/sites/default/files/product-files/Minority_Teacher_Recruitment_Employment_Retention%20_BRIEF.pdf.
- Irizarry, J., 2007. Home-growing teachers of color: lessons learned from a town-gown partnership. *Teach. Educ. Q.*, fall 87–102.
- Ishimaru, A., Bang, M., Valladares, M.R., Nolan, C.M., Tavares, H., Rajendran, A., Chang, K., July 2019. Recasting Families and Communities as Co-designers of Education in Tumultuous Times. National Education Policy Center, Boulder, CO.
- Kohli, R., 2012. Racial pedagogy of the oppressed: critical interracial dialogue for teachers of color. *Equity & Excell. Educ.* 45 (1), 1–16.
- Kohli, R., 2014. Unpacking internalized racism: teachers of color striving for racially just classrooms. *Race Ethn. Educ.* 17 (3), 367–387.
- Kohli, R., 2019. Lessons for teacher education: the role of critical professional devel- opment in teacher of color retention. *J. Teach. Educ.* 70 (1), 39–50.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491.
- Ladson-Billings, G., 2001. *Crossing over to Canaan*. Jossey-Bass, San Francisco, CA.
- Ladson-Billings, G., 2009. Race still matters: critical race theory in education. In: Apple, M.W., Au, W., Gandin, L.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, pp. 110–122.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: A.k.a the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Ladson-Billings, G., Paris, D., 2021. *Culturally Relevant Pedagogy: Asking a Different*.
- Leonardo, Z., 2004. The color of supremacy: beyond the discourse of 'white privilege. *Educ. Philos. Theor.* 36 (2), 137–152.
- Liu, K., 2015. Critical reflection as a framework for transformative learning in teacher education. *Educational Review* 67 (2), 135–157.
- Liu, K., 2017. Creating a dialogic space for prospective teacher critical reflection and transformative learning. *Reflective Practice* 18 (6), 805–820.
- Liu, K., 2020. Critical reflection for transformative learning: Understanding ePortfolios in teacher education. Springer. <https://link.springer.com/book/10.1007/978-3-319-01955-0>.
- Liu, K., Ball, A., 2019. Critical reflection and generativity: Toward a framework of transformative teacher education for diverse learners. *Review of Research in Education* 43 (1), 68–105.
- Lopez, V., Montecinos, C., Rodriguez, J.I., Calderon, A., Contreras, J.F., 2012. Enacting solidarity to address peer-to-peer aggression in schools: case studies from Chile. In: Sleeter, C., Soriano, E. (Eds.), *Creating Solidarity across Diverse Communities*. Teachers College Press, New York, NY, pp. 23–43.
- Matias, C.E., Mackey, J., 2016. Breakin' down whiteness in antiracist teaching: introducing critical whiteness pedagogy. *Urban Rev.* 48 (1), 32–50.
- Meier, D., 2002. In: *Schools We Trust: Creating Communities of Learning in an Era of Testing and Standardization*. Beacon, Boston, MA.
- Meng, D., 2008. The attitudes of primary school teachers toward inclusive education in rural and urban China. *Front. Educ. China* 3 (4), 473–492.
- Mezirow, J. (Ed.), 1990. *Fostering Critical Reflection in Adulthood: A Guide to Transformative and Emancipatory Learning*. Jossey-Bass, San Francisco, CA.
- Miller, R., Liu, K., Ball, A., 2020. Critical counter-narrative as transformative methodology for educational equity. *Review of Research in Education* 44 (1), 269–300.
- Milner IV, H.R., 2003. Reflection, racial competence, and critical pedagogy: how do we prepare pre-service teachers to pose tough questions? *Race Ethn. Educ.* 6 (2), 193–208.
- Milner IV, H.R., 2006. Preservice teachers' learning about cultural and racial diversity: implications for urban education. *Urban Educ.* 41 (4), 343–375.
- Milner IV, H.R., 2008. Critical race theory and interest convergence as analytical tools in teacher education policies and practices. *J. Teach. Educ.* 59 (4), 332–346.
- Milner IV, H.R., 2017. Where's the race in culturally relevant pedagogy? *Teach. Coll. Rec.* 119 (1), 1–32.
- Milner IV, H.R., Laughter, J., 2015. But good intentions are not enough: preparing teachers to center race and poverty. *Urban Rev.* 47 (2), 341–363.
- Moll, L.C., Amanti, C., Neff, D., González, N., 1992. Funds of knowledge for teach- ing: using a qualitative approach to connect homes and classrooms. *Theory Into Pract.* 31 (2), 132–141.
- Murrell Jr., P.C., 2000. Community teachers: a conceptual framework for preparing exem- plary urban teachers. *J. Negro Educ.* 69 (4), 338–348.
- National Center for Education Statistics, 2022a. *Characteristics of Public School Teachers*. Condition of Education. U.S. Department of Education, Institute of Education Sciences. July 2nd, 2022. Available at: <https://nces.ed.gov/programs/coe/indicator/clr>.

- National Center for Education Statistics, 2022b. Racial/Ethnic Enrollment in Public Schools. Condition of Education. U.S. Department of Education, Institute of Education Sciences. July 2nd, 2022. Available at: <https://nces.ed.gov/programs/coe/indicator/cge>.
- National Center for Education Statistics, 2022c. English Learners in Public Schools. Condition of Education. U.S. Department of Education, Institute of Education Sciences. July 2nd, 2022. Available at: <https://nces.ed.gov/programs/coe/indicator/cgf>.
- Noguera, P., 2001. Transforming urban schools through investments in the social capital of parents. In: Saegart, S., Thompson, J.P., Warren, M. (Eds.), *Social Capital and Poor Communities*. Russell Sage Foundation, pp. 189–212.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Paris, D., 2021. Culturally sustaining pedagogies and our futures. *Educ. Forum* 85 (4), 364–376.
- Paris, D., Alim, H.S., 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press.
- Paris, D., Ball, A., 2009. Teacher knowledge in culturally and linguistically complex classrooms: lessons from the golden age and beyond. In: Morrow, L.M., Rueda, R., Lapp, D. (Eds.), *Handbook of Research on Literacy Instruction: Issues of Diversity, Policy, and Equity*. Guilford, pp. 379–395.
- Parker, L., Hood, S., 1995. Minority students vs. majority faculty and administrators in teacher education: perspectives on the clash of cultures. *Urban Rev.* 27 (2), 159–174.
- Philip, T.M., 2010. Moving beyond our progressive lenses: recognizing and building on the strengths of teachers of color. *J. Teach. Educ.* 62 (4), 356–366.
- Philip, T.M., 2013. Experience as college student activists: a strength and liability for prospective teachers of color in urban schools. *Urban Educ.* 48 (1), 44–68.
- Philip, T.M., Benin, S.Y., 2014. Programs of teacher education as mediators of White teacher identity. *Teach. Educ.* 25 (1), 1–23.
- Rieger, A., 2015. Making sense of White identity development: the implications for teacher education. *Multicultural Learning and Society* 10 (2), 211–230.
- Riegle-Crumb, C., Morton, K., Nguyen, U., Dasgupta, N., 2019. Inquiry-based instruction in science and mathematics in middle school classrooms: examining its association with students' attitudes by gender and race/ethnicity. *AERA Open* 5 (3). Available at: [10.1177/2332858419867653](https://doi.org/10.1177/2332858419867653).
- Schön, D., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.
- Schultz, B., Gillette, M., Hill, D., 2008. A theoretical framework for understanding Grow Your Own Teachers. *The Sophist's Bane* 4 (1), 69–80.
- Schultz, B.D., Gillette, M.D., Hill, D.A., 2011. Teaching as political: theoretical perspectives for understanding the grow your own movement. In: *Grow Your Own Teachers: Grassroots Change for Teacher Education*. Teachers College Press, pp. 5–21.
- Siddle Walker, V., 2001. African American teaching in the south, 1940–1960. *Am. Educ. Res. J.* 38 (4), 751–779.
- Sleeter, C.E., 2001. Preparing teachers for culturally diverse schools: research and the overwhelming presence of Whiteness. *J. Teach. Educ.* 52 (2), 94–106.
- Sleeter, C.E., 2008. Preparing White teachers for diverse students. *J. Teach. Educ.* 59 (3), 212–219.
- Sleeter, C.E., 2017. Critical race theory and the whiteness of teacher education. *Urban Educ.* 52 (2), 155–169.
- Sleeter, C.E., Grant, C., 1987. An analysis of multicultural education in the United States. *Harv. Educ. Rev.* 57 (4), 421–445.
- Sleeter, C.E., Neal, L.I., Kumashiro, K.K., 2014. *Diversifying the Teacher Workforce: Preparing and Retaining Highly Effective Teachers*. Routledge, New York, NY.
- Sleeter, C.E., Thao, Y., 2007. Guest editors' introduction: diversifying the teaching force. *Teach. Educ. Q.* 34 (4), 3–8.
- Smith, B.O., 1969. *Teachers for the Real World*. American Association of Colleges for Teacher Education.
- Souto-Manning, M., 2010. Playing with power and privilege: theatre games in teacher education. *Teach. Teach. Educ.* 27 (6), 997–1007.
- Téllez, K., 1999. Mexican-American preservice teachers and the intransigency of the elementary school curriculum. *Teach. Teach. Educ.* 15, 555–570.
- Thelwell, C., 2020. *Exporting Jim Crow: Blackface Minstrelsy in South Africa and beyond*. University of Massachusetts Press.
- Toshalis, E., 2012. The rhetoric of care: preservice teacher discourses that depoliticize, deflect, and deceive. *Urban Rev.* 44 (1), 1–35.
- Tschannen-Moran, M., 2004. *Trust Matters: Leadership for Successful Schools*. Jossey-Bass.
- U.S. Department of Education, National Center for Education Statistics, 2003. Prepared by. In: Hurst, D., Tan, A., Meek, A., Sellers, J. (Eds.), *Overview and Inventory of State Education Reforms: 1990 to 2000*.
- Valencia, R., 2010. *Dismantling Contemporary Deficit Thinking: Educational Thought and Practice*. Routledge.
- Valenzuela, A., 2017. *Grow Your Own Educators Programs: A Review of the Literature with an Emphasis on Equity-Based Approach* (Intercultural Development Research Association Report). Available at: <http://www.idraeacsouth.org/wp-content/uploads/2017/12/Grow-Your-Own-Educator-Programs-Lit-Review-IDRA-EAC-South-2017.pdf>.
- Villegas, A.M., Davis, D., 2008. Preparing teachers of color to confront racial/ethnic disparities in educational outcomes. In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, J. (Eds.), *Handbook of Research in Teacher Education: Enduring Issues in Changing Contexts*. Earlbaum, Mahwah, NJ, pp. 583–605.
- Villegas, A.M., Irvine, J.J., 2010. Diversifying the teaching force: an examination of major arguments. *Urban Rev.* 42 (3), 175–192.
- Villegas, A.M., Strom, K., Lucas, T., 2012. Closing the racial/ethnic gap between students of color and their teachers: an elusive goal. *Equity & Excell. Educ.* 45 (2), 283–301.
- Woodson, A., Pabon, A., 2016. "I'm none of the above": Exploring themes of heteropatriarchy in the life histories of Black male educators. *Equity & Excell. Educ.* 47 (1), 57–71.
- Yosso, T., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Zeichner, K.M., 1993. *Educating Teachers for Cultural Diversity: NCRTL Special Report*. National Center for Research on Teacher Learning.
- Zeichner, K., 1996. Educating teachers for cultural diversity. In: Zeichner, K., Melnick, S., Gomez, M.L. (Eds.), *Currents of reform in preservice teacher education*. Teachers College Press, pp. 133–175.
- Zeichner, K., Bowman, M., Guillen, L., Napolitan, K., 2016. Engaging and working in solidarity with local communities in preparing the teachers of their children. *J. Teach. Educ.* 67 (4), 277–290.
- Zhang, D., Li, X., Xue, J., 2015. Educational inequality between rural and urban areas of the People's Republic of China, migrant children education, and some implications. *Asian Dev. Rev. Stud. Asian Pac. Econ. Issues* 32 (1), 196–224.
- Zygmunt, E., Clark, P., Clausen, J., Mucherah, W., Tancock, S., 2016. *Transforming teacher education for social justice*. Teachers College Press.
- Zygmunt, E., Cipollone, K., Tancock, S., 2018. Loving out loud: community members, teacher candidates, and transformational learning through a pedagogy of care and connection. *Journal of Teacher Education* 62 (2), 127–139.

Educating the teaching profession

Sølvi Lillejord^{a,b}, ^a SLATE, Centre for the Science of Learning & Technology, University of Bergen, Bergen, Norway; and
^b Department of Education, Oxford University, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Aim of the article	368
Teacher education—a scattered picture	368
Characteristics of professions and semi-professions	369
Professional education—historic background	369
Perspectives on the organization of schools	370
Practice-shock—the theory-practice problem	371
Way forward: more symmetric partnerships in teacher education	372
References	373

Aim of the article

The article aims at comparing traits of the teaching profession and more established professions to identify how teacher education programs can support the development of a *knowledge-producing* teaching profession. Historic relations between church and school are outlined, the school organization and relationship between teachers and their leaders discussed and traits of the teaching profession compared to those of more mature professions. In conclusion, prerequisites for a university-based study program that can strengthen the teaching profession are presented by indicating how school-university partnerships should be organized, structured, and led to cumulatively build a knowledge base, upon which teachers can enact their professional discretion.

Teacher education—a scattered picture

The last decades, many countries have reformed their teacher education, in the hope that program change can improve schools, but with few attempts to identify lessons learned from specific reform efforts and from developments in other professions (Zeichner and Liston, 1990). It is a problem that many of the reforms have not addressed fundamental structural problems or contributed to increased coherence in teacher education programs (Grossman et al., 2008). Mayer and Oancea (2021) noted that quite often, reviews of teacher education research find the field underdeveloped, small scale, undertheorized, fragmentary, and somewhat parochial. They therefore concluded that much teacher education research has minimal influence on policy other than being used as a weapon to advance competing views or manufacture narratives of failure. For instance, Cochran-Smith et al. (2017, 572) described how the US approach early 2000 was to “hold teacher education accountable,” by assuming that enhanced teacher education quality depended on evaluation and monitoring of outcomes related to institutions, programs, and teacher candidates. In Australia, Mayer et al. (2017) found teacher education positioned as a policy problem and newly educated teachers were expected to be instantly classroom ready, and effectively able to enhance students’ learning outcomes.

In the mid-1980s, Darling-Hammond (1985) suggested to professionalize teaching and teacher education, and early 2000, Cochran-Smith and Fries (2001) noted international moves in this direction. In many countries today, those who want to be teachers can choose integrated programs at universities or colleges. They can, however, also choose among a variety of short courses under umbrella terms such as Teach for America, Teach for All, School Direct and Teach First. These programs are typically based on training in school by more experienced teachers and summer school courses. On this background, Darling-Hammond (2010) noted that teaching is currently where medicine was in 1910, when aspiring doctors could take 3-week training programs, memorizing lists of symptoms and cures or they could attend Johns Hopkins University graduate school, that prepared doctors in clinical practice and the sciences of medicine.

In terms of speed and ambition, the development of teacher education programs differs across nations. In Finland, politicians decided to move teacher education from seminaries to universities in the 1970s (Simola et al., 1997). Iceland transferred teacher education to tertiary level in 1971 (Bjarnadóttir, 2008), and a 1977 reform in Sweden, gave the responsibility of educating teachers to universities and university colleges (Lundahl et al., 2010). Around the turn of the century, Erixon et al. (2001) noted that while all teacher education programs (preschool, primary school, upper secondary school) in Finland and Sweden were integrated and university-based, Denmark had no university-based teacher education program. During the 1970s, Norwegian teachers were prepared in two- or three-year *teacher education schools*, that replicated the seminar tradition. The first Norwegian White Paper with references to the *teaching profession* was published in 2009 (Ministry of Education and Research, 2008–2009), and in 2017, some forty years after neighboring Nordic countries, Norwegian politicians decided that primary teacher education should be at Masters’ level.

In England, a 2011 White Paper referred to teaching as a *craft*, with skills that are better trained in schools (Browne and Reid, 2012). The idea that central government is not only deciding *what* should be taught, but also *how*, is, according to Furlong et al. (2008) a result of Thatcherism and New Labor's Third Way-policy, and a contribution to *de*-professionalizing teachers (Beck, 2008). Having compared reform initiatives in the UK and Sweden, Weiner (2002) concluded that while Britain chose to respond to problems in the 1990s by creating a crisis of confidence in teachers and teacher education, impose regulations and limit teachers' autonomy, Swedish reforms aimed at greater coherence between different routes and raising the status of teachers by substantially increased research funding connected to schools and teachers' work.

According to Darling-Hammond (2012), exemplary teacher education programs are characterized by regarding teachers as professionals, a focus on learner- and learning-centered pedagogy, and holistic performance assessments. Strong university-based programs are mostly staffed by researchers who study teacher education, demonstrate a thoughtful balance between theory and practice, coherent curricula, and carefully structured clinical experiences with partner schools. They draw on complex and varied knowledge and see students' development as integral to the social contexts. Having compared three teacher education programs (Utrecht University in the Netherlands, Queen's University in Canada, and Monash University in Australia), Korthagen et al. (2006) argued that there would be no real changes in teacher education, unless teacher educators started modeling the practices they talked about. Having analyzed 20 articles published between 1983 and 2002, they found that good teacher education programs helped students develop their professional discretion. These programs centered around the student teachers, who worked inquiry based and collaborated in professional learning communities where they learned how research informed their practice. A core characteristic of programs was strong partnerships between universities and schools, and a joint understanding of the relationships between teaching and learning.

To sum up: There are, globally, huge differences in how teachers are educated and followed up. Problems related to shifting education policies, the low status of teacher education at universities and teachers' weak professional status, contribute to shaping the teaching profession. It is therefore interesting to compare the teaching profession to other professions.

Characteristics of professions and semi-professions

Researchers (e.g. Brante, 2013; Darling-Hammond, 2013; Morris et al., 2006; Simons and Ruijters, 2014; Lillejord and Børte, 2020b) agree on these characteristics of professions: They are (1) exposed to long, specialized, systematic, academic training followed by exams and certification that distinguish them from the non-certified. Professions develop their leaders through specialization and qualification and their associations ensure the quality of the professions' work and sanction unprofessional conduct. (2) Professions use their discretion to decide upon proper interventions on a joint knowledge base from research and experience. The professional collective shares the responsibility to renew and maintain the knowledge base, agrees on characteristics of reliable and valid knowledge, and develops standards for good practice. (3) Professions adhere to a societal mandate; they help people and provide services of high quality and societal relevance. (4) Professions are collegially organized with a high degree of autonomy. Their work is complex, tasks are heterogeneous and best solved collectively, in accordance with ethical guidelines.

In contrast to this, Etzioni (1969) labeled occupations rooted in care and calling such as teachers, nurses, and social workers, as *semi-professions*, based on the following arguments: semi-professions have shorter education than the established professions, they follow standards and work procedures developed by external actors and communicate knowledge rather than create it. A fragmented teaching corps, internal disagreement on fundamental issues, and the absence of professional organization positioned teachers not as respected experts of the field, but as "street-level bureaucrats" (Lipsky, 1980), responsible for implementing decisions from above.

When elaborating on the idea of an occupation as a "calling," Gustafson (1982) noted that *calling* refers to moral and religious motives and to a vision of the larger ends and purposes that work serves. In contrast, professions are characterized by mastery of technical knowledge, concepts and theories that guide choices, institutionalization that exercises social controls, and a service orientation. Therefore, he argued, "calling" without professionalization is inept, and a profession *without* a calling lacks moral and humane roots, loses human sensitivity, and restricts the vision of the purposes of human good that are served. Core characteristics of professions are therefore certification (Freidson, 1988), long, specialized education, a societal mandate, a value foundation, ethics and collective, continuous knowledge development. Taking a historical perspective on knowledge and learning, this article focuses primarily on the latter.

Professional education—historic background

Universities have a long tradition of educating professionals, with theology, medicine, and law as the medieval and early modern *learned* professions. Research-based education, global agreement on standards and professional organization, has proven successful for physicians and lawyers, who still rank high on lists showing the relative prestige of occupations (Ingersoll and Collins, 2018). Compared to the professions that emerged during the 20th Century, most teachers were, historically, educated in seminaries. Except for teachers who taught students in the medieval Cathedral schools, and teachers in contemporary upper secondary, who hold degrees in their teaching subjects, teachers' education has, in general, been shorter, less

research-informed, and often focused on instructional delivery. Teachers' main task has been to *communicate* knowledge to students, and to test how well students remember what they have been taught. This understanding of what it means to know and learn can have emerged from the historic ties between the Church and the schools that for centuries have influenced educationalists' perceptions of teaching.

Mamadailiev et al. (2019) described how the elementary school system in Prussia, Germany evolved during the 1700. Instruction centered on catechism and studies of the Bible. Teachers were candidates of theology, and the first German seminary for school-teachers was established in 1778. Gradually, central government took control over educational institutions. Under the chairmanship of the Minister of State, and directly answerable to the King, several secular members constituted local educational authorities. Cremin (1953) described US education as characterized by vast differences in quality, availability, and organization. There were three kinds of schools: elementary schools (reading, writing and religion); secondary or grammar schools (preparing young men for college) and colleges which taught liberal arts programs (languages, mathematics, and the classics). Many US teachers were clerics who prepared for teaching through seminaries and shorter courses (Mehta, 2013). As the need for education increased, secular subjects were introduced. Men unfit for manual work and more females entered teaching, and between 1880 and 1920, US education was institutionalized in a process that, according to Mehta (2013, p. 11) transformed a "nation of one-room school-houses into a network of district school systems." Mostly male superintendents were empowered as CEOs to lead mostly female teachers who were expected to follow rules and programs prescribed by their superintendent. From this emerged the idea that the complex practice in need of attention was not teaching, but *administration*. A new class of male administrative professionals elevated themselves by distancing themselves from the teachers.

In his analysis of educational organizations, Weick (1976) noticed that in schools, two loosely coupled systems appear to work separately from each other, with little transfer of knowledge across systems. One system consists of teachers, parents, students and curriculum, another of the principal, vice-principal, and superintendent. The two systems are somehow attached, but each part preserves its own identity, and there is little transfer of knowledge. While teachers are responsible for the schools' "technical core," the principal holds the authority of office. Teachers' autonomy and reluctance to change is often cited as the prime reason for reform implementation failure in schools. Instead, Mehta (2013) argued that a root problem in the educational sector is that it is organized as a bureaucracy rather than as a profession and suffers under the absence of an "educational infrastructure" to support practice.

Perspectives on the organization of schools

All *waste* in education, argued Dewey (1915), is the result of organization, or the lack of it, as organization is often motivated by economy and efficiency. The educational system has, historically, grown from the top-down, with different historical purposes and each part isolated from the others. While the university has its roots in the Middle Ages, kindergartens were established in the 1900s. As Latin, and Greek, during the Renaissance, connected people with the culture of the past, this was reflected in the grammar school, which primarily aimed at *cultivation*. The rationale behind the primary school was *utility*, as it became a business necessity to know how to read, write and figure during the 1500s. Late 1700 and early 1800, the training schools for teachers, should train persons *how* to teach, while the college taught them *what* to teach—with almost a contempt for methods of teaching: "In this division between what to teach and how to teach, each side suffers from the separation" (Dewey, 1915, p. 44).

The historic ties between the Church and school brought clerical practices to schools. The links between preachers and teachers, theology, and pedagogy, positioned the Church as the school's parent. Pedagogy was not among the learned sciences, and teaching was modeled on preaching and lecturing. As secular subjects were introduced, textbooks substituted the catechism, and subject knowledge substituted Biblical knowledge. In both cases, knowledge was regarded as external, transported into schools by teachers, who interpreted, adapted, and transmitted it to students. Discipline and upbringing were important; students listened to the teachers' lectures, were expected to remember what was taught, and their knowledge tested. For centuries, this has influenced our ideas about how we learn. Using examples from college teaching, Smith and Waller (1997) referred to this understanding of teaching as *the old paradigm*: Teachers *give* knowledge, students *get* it. Students sit quietly at their seats, eyes front, feet on the floor, as passive recipients of knowledge. Assessment is "objective" and students expected to memorize and recall relevant information in tests. Anyone with expertise in their field can teach without training.

In his analysis of how teachers taught between 1890 and 1980, Cuban (1984) concluded that teacher-centered practices dominated. Characteristics of these practices were whole-class instruction, that teacher talk exceeded student talk during instruction, teachers determined the use of class time, and the classroom was usually arranged into rows of desks facing a blackboard. Later, Cuban (2013) argued that the dominant policy strategy, to change teachers, had failed because it did not take the situation in which teachers find themselves, their *context*, into consideration. He referred to this as a fundamental attribution error, undervaluing the power of the age-graded school, particularly the way it isolates teachers from one another and discourages collaboration. Schools have a long history of isolating people from one another and separating teachers from their leaders.

Mehta (2013) argued that, from a sectoral perspective, American schooling is *organized backwards*, when compared to both medicine and higher achieving educational systems. By this, he meant that the educational authorities aim to accomplish on the back end—by holding teachers accountable—what is not done on the front end—building a system that is reasonably expected to consistently create good results. He asked if the downwards spiral of distrust, external regulations, and increased accountability, which

makes teaching less attractive to talented people, could explain the under-professionalization of teachers. [Ingersoll and Collins \(2018\)](#) argued that mature professions exert substantial control over the curriculum, admissions, and accreditation of professional training schools; set and enforce behavioral and ethical standards for practitioners; and exert substantial control over who their future colleagues are to be. As teachers, in contrast, have limited authority over their work, they cannot be held accountable for activities they cannot control.

Practice-shock—the theory-practice problem

As both teachers ([Colognesi et al., 2020](#)), nurses ([Duchscher and Windey, 2018](#)) and health professionals ([Opoku et al., 2020](#)) are reported to experience practice-shocks when they start working, some argue that their education should be more practice-relevant. The practice- or transition-shock is related to the theory—practice-problem. The idea that theory can be transferred to students through lectures, and that student teachers can “use” the knowledge from the lectures in their teaching practice, was labeled *technical rationality* by [Schön \(1983\)](#), and by [Doyle \(1990\)](#) a fragmented and mechanistic understanding of teaching. As indicated, it can also be a cultural reminiscence (relict) of clerical practices and a result of how schools are organized and led—on principles of economy and efficiency.

The practice-shock is part of a problem complex encompassing high rates of dropouts and too many newly qualified teachers who quit. [Kelchtermans \(2017\)](#) observed that while teacher attrition and retention were presented as complex, multilayered issues, the concepts were not clearly defined. He therefore suggested to see teacher attrition/retention as two sides of the same coin, where *attrition* describes teachers who leave the profession before having reached the age of retirement, and *retention* refers to keeping teachers in teaching. As we can imagine situations where it is favorable for both teachers and students if teachers leave the job, both attrition and retention is meaningless without a simultaneous reference to what constitutes good teaching. [Kelchtermans \(2017\)](#) therefore suggested seeing questions of attrition and retention as the need to prevent good teachers from leaving the job for the wrong reasons. Also, [Ingersoll \(2001\)](#) argued for a broader perspective, for instance to study teacher attrition/retention at the school organizational level.

[Voss and Kunter \(2020\)](#) found that studies of beginning teachers often followed two research strands, one focusing on teachers’ beliefs, the other on beginning teachers’ emotional exhaustion and stress. Induction periods and mentoring are frequently suggested to remedy their situation ([Paula and Grinfelde, 2018](#)). However, [Kelchtermans \(2019\)](#), argued that many support initiatives are caught up in deficit and remedial thinking with negative and even counter-productive side effects. Instead of interventions to support early career teachers in the induction phase, he suggested to see them as *sense-making agents*, *networkers*, and *assets* to the school. Instead of assuming that new teachers *lack* competence, schools should emphasize their agency: the capacity to interpret and judge the situations they find themselves in and to act upon it. As individual actors, teachers are situated in complex social relationships. Early career teachers often question existing practice, and should be regarded as assets to the school. In a study of how primary schools in the northern part of Norway welcomed the first students with master theses from the university, [Jakhelln et al. \(2016\)](#) concluded that neither teachers nor school leaders showed interest in the topics of their thesis. They also ascribed little value to the research skills that the students had developed during their education.

Many countries have tried to solve the theory-practice problem by transferring the responsibility for teacher education from universities to schools. This, argued [Ellis \(2010\)](#), will not solve the more fundamental problem, which is how teachers’ professional knowledge can be strengthened through a synthesis of research (theory) and practice (experience). While students learn a lot from practical experiences in schools, they do not learn how they can criticize practice to improve it, a prerequisite in the development of professional knowledge. Similar arguments were put forward by [Lillejord and Børte \(2020a\)](#), who conducted a systematic review and found that middle leaders in school rely solely on experience in their practice. [Ellis \(2013\)](#) argued for rethinking the division of labor between universities and schools as university staff spend too much time on students during their school placements, and [Ellis and McNicholl \(2015\)](#) suggested that the time researchers at universities spend on “relational work” with frustrated students, they should rather spend on research. This indicates a need to strengthen schools’ awareness of their important contribution to teacher education and schoolteachers’ competence as teacher educators.

[Grossman et al. \(2008\)](#) argued that the move from a teacher education curriculum organized by knowledge domains to a curriculum organized around professional practices requires that teacher educators *first*, must develop programs that undo the historical separation between foundation and methods courses; *second*, help novice teachers develop and refine a set of core practices for teaching. This requires changes both at the pedagogical and organizational level of teacher education, and at the collaborating institutions.

This section has shown that historically, knowledge has been regarded as external to the school. It has been transported into schools in books, and teachers have interpreted and adapted the subject knowledge students are expected to learn. Students come to school to get knowledge, and early career teachers to get teaching experience. Both groups are treated as if they *lack* knowledge and skills. The schools’ response to this deficit is supervision and mentoring by experienced teachers. There are complaints about the low practice-relevance of teacher education, and these complaints often come from schools, with the responsibility for the practicum part of teacher education. This paradox can indicate that schools need to learn about their own knowledge processes and develop a meta-perspective on their knowledge work.

Way forward: more symmetric partnerships in teacher education

Having outlined historical, structural, and organizational problems in teacher education, this section discusses prerequisites for how university-based study programs can contribute to strengthening the teaching profession. How school-university partnerships should be structured, organized, and led to cumulatively build teachers' knowledge base will be emphasized.

Maandag et al. (2007) compared partnership models in England, France, Germany, The Netherlands, and Sweden, and identified these partnership models: Work placement, training schools and training in and within schools. The dominant model was work placement. Mehta (2013) suggested to change the current model where teacher training institutions are responsible for teacher preparation during student teaching or the first year of teaching, and after that, districts are responsible for induction. Instead, he suggested, teacher education could learn from medicine and develop an apprenticeship program akin to the medical residency system. It would make more sense, he argued, to unite these functions under single institutions that would oversee teachers' development during their first three years of teaching.

In a research mapping of studies on partnerships in teacher education, Lillejord and Børte (2016) found that researchers described partnerships as complex and resource-intensive cross-institutional infrastructures for knowledge sharing, with the ambition to bridge theory and practice, enhance the practice-relevance of teacher education, and support mentoring and professional learning. An early model of integrated teacher education was the Oxford Internship Scheme, developed in the mid-1980s to strengthen connections between the theory students met at university and what they experienced in schools (BERA, 2014). Around the same time, in the US, the Holmes Group developed teacher education partnerships based on clinical practice (Kochan and Kunkel, 1998). How well school-university partnerships function, depend on how they are structured and led, responsibilities defined, and work divided. Studies included in the research mapping revealed asymmetric power-relations and tensions at all levels and emphasized the need for competent academic leadership in the establishment, running and renewal of partnerships. A major challenge is how to establish and maintain productive learning relations between the partners.

Zeichner (2010) argued that weak integration of the schools' contribution to teacher education is a big problem. The triad student-mentor-supervisor frequently causes misunderstandings and disagreement and real collaboration between schools and teacher education institutions is difficult to achieve. Allen et al. (2010) found that teachers and academics lacked understanding of the other party's knowledge culture. Partners appear to live in separate worlds, with separate communities. Strategic leadership is needed to improve the dialog between the parties, and the optimal solution appears to be concrete collaborative projects in which the partners jointly create a shared product. Studies reported reluctance in challenging the tradition where students are first presented the theory they later are expected to "practice." An alternative to this model is to design a study program where students simultaneously investigate instructional practice through first-hand experience and read relevant research. Here, research and practice are perceived as two different, but equally important sources of knowledge. This approach gives students an opportunity to synthesize experience-based knowledge and knowledge from research and prepares them for challenges they will encounter as teachers.

The mapping revealed that the concept "third space" is frequently used in research on partnerships, based on the assumption that collaboration need neutral ground for deliberation. In *third space* university supervisor, school mentor and student step out of their own field of knowledge, to jointly create new knowledge. While the ideal is symmetric relations between participants, the triadic collaboration caused tensions and disagreements in the relations between mentor-student, mentor-supervisor and supervisor-student. As the student is dependent on both supervisor and mentor, the three parties are not partners on equal terms. Students were often entangled in conflicts of loyalty and had to choose side. Problems were related to asymmetric power relations, imbalance between the professional and personal, in-fighting over who was the expert and different ideas and expectations on how to teach and develop professionally.

As some partnership models can be dysfunctional, there is an obvious need for innovative thinking in teacher education partnerships. Lillejord and Børte (2016) suggested to place the students' learning at the center of activities in a model where both mentors in school and supervisors at university would have to relate to both research and practice. The tension between theory and practice can be eased when none of the parties may claim exclusive "ownership" of one knowledge area. Such a model also supports knowledge development because all actors must cross boundaries and relate to various types of knowledge.

This article has asked if a core problem in teacher education is that teachers in school primarily perceive knowledge as external to the school. While they interpret and communicate this knowledge, they do not, themselves, systematically *produce* knowledge. This is one of the traits used to distinguish semi-professions from more mature professions.

Reflective practice has been promoted as an approach to increase teachers' understanding of their work. Russell (2013), however, asked if the strong belief in reflective practice has done more harm than good because teacher educators rarely explain what they mean by reflective practice or model it, and generally fail to link reflection clearly and directly to professional learning. Therefore, teacher candidates tend to complete the program with a muddled and negative view of what reflection is and how it might contribute to their professional learning.

Professions' knowledge base consists of various knowledge sources, the two most important are experience and research. When it comes to research use, teachers are reported to lack information literacy (Williams and Coles, 2007). Wieser (2016) noted that teachers used practical knowledge when planning their teaching, and personal knowledge when reflecting on it. They do not have the competence to interpret research so that it becomes relevant for their own practice, which they see as essentially unique, extremely complex, and unsolvable by research (Cain, 2016). One solution to this problem is to establish partnerships where teachers and student teachers collaboratively work to investigate and document their practice. Professional practice is a knowledge

form that should be collected, discussed in the professional learning communities, criticized, compared to findings from research and improved. Increased understanding of the schools' own knowledge can support teachers' use of research and data literacy.

To sum up: There is, globally, huge differences in how teachers are educated and followed up. Problems related to shifting education policies, the antagonistic relationship between school leaders and teachers, coupled with the low status of teacher education at universities, contribute to shaping the teaching profession. In addition, a strong belief in external knowledgeable experts such as consultants or market agents, has hampered the development of a strong teaching profession (Lillejord and Børte, 2020b). As established professions have systematically built their knowledge from within, this is also the way forward for the teaching profession.

References

- Allen, J.M., Butler-Mader, C., Smith, R.A., 2010. A fundamental partnership: the experiences of practicing teachers as lecturers in a pre-service teacher education program. *Teach. Theor. Pract.* 16 (5), 615–632.
- Beck, J., 2008. Governmental professionalism: re-professionalizing or de-professionalizing teachers in England. *Br. J. Educ. Stud.* 56 (2), 119–143.
- BERA, 2014. The Role of Research in Teacher Education: Reviewing the Evidence. Interim Report of the BERA-RSA-Inquiry. <http://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-RSA-Interim-Report.pdf>.
- Bjarnadóttir, K., 2008. Societal demands on the profession of the mathematics teacher in Iceland in a historical context. *Int. J. Hist. Math. Educ.* 3 (2), 73–82.
- Brante, T., 2013. The professional landscape: the historical development of professions in Sweden. *Prof. Prof.* 3 (2). <https://doi.org/10.7577/pp.558>.
- Browne, L., Reid, J., 2012. Changing localities for teacher training: the potential impact on professional formation and the university sector response. *J. Educ. Teach.* 38 (4), 497–508.
- Cain, T., 2016. Research utilization and the struggle for the teacher's soul: a narrative review. *Eur. J. Teach. Educ.* 39 (5), 616–629.
- Cochran-Smith, M., Fries, M.K., 2001. Sticks, stones, and ideology: the discourse of reform in teacher education. *Educ. Res.* 30 (8), 3–15.
- Cochran-Smith, M., Baker, M., Burton, S., Chang, W.C., Cummings Carney, M., Fernández, M.B., et al., 2017. The accountability era in US teacher education: looking back, looking forward. *Eur. J. Teach. Educ.* 40 (5), 572–588.
- Colognesi, S., Van Nieuwenhoven, C., Beausaert, S., 2020. Supporting newly qualified teachers' professional development and perseverance in secondary education: on the role of informal learning. *Eur. J. Teach. Educ.* 43 (2), 258–276.
- Cremin, L.A., 1953. The heritage of American Teacher Education. *J. Teach. Educ.* 4 (2), 163–170.
- Cuban, L., 1984. *How Teachers Taught: Constancy and Change in American Classrooms, 1890–1990*. Longman, New York & London.
- Cuban, L., 2013. *Inside the Black Box of Classroom Practice: Change Without Reform in American Education*. Harvard Education Press.
- Darling-Hammond, L., 1985. Valuing teachers: the making of a profession. *Teach. Coll. Rec.* 87 (2), 205–218.
- Darling-Hammond, L., 2010. Teacher education and the American future. *J. Teach. Educ.* 61 (1–2), 35–47.
- Darling-Hammond, L., 2012. *Powerful Teacher Education: Lessons From Exemplary Programs*. John Wiley & Sons, San Francisco, CA.
- Darling-Hammond, L., 2013. *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement*. Teachers College Press, New York.
- Dewey, J., 1915. *The School and Society & the Child and the Curriculum*. Unabridged Republication. Dover Publications Inc, Mineola, New York.
- Doyle, W., 1990. Themes in teacher education research. In: Houston, W.R., Haberman, M., Sirkula, J. (Eds.), *Handbook of Research on Teacher Education*. Macmillan, New York, pp. 3–23.
- Duchscher, J.B., Windey, M., 2018. Stages of transition and transition shock. *J. Nurses Prof. Dev.* 34 (4), 228–232.
- Ellis, V., McNicholl, J., 2015. *Transforming Teacher Education: Reconfiguring the Academic Work*. Bloomsbury, London.
- Ellis, V., 2010. Impoverishing experience: the problem of teacher education in England. *J. Educ. Teach.* 36 (1), 105–120.
- Ellis, V., 2013. Teacher education in the public university: the challenge of democratising knowledge production. *Pedagog. High. Educ.* 198–214.
- Erixon, P.O., Frånberg, G.M., Kallós, D., 2001. Postgraduate studies and research in teacher education within the European Union. In: *The Role of Graduate and Postgraduate Studies and Research in Teacher Education Reform Policies in the European Union*, pp. 47–60.
- Etzioni, A., 1969. *The Semi-professions and Their Organization: Teachers, Nurses, Social Workers*. Free Press, New York.
- Freidson, E., 1988. Theory and the professions. *Ind. Law J.* 64, 423.
- Furlong, J., McNamara, O., Campbell, A., Howson, J., Lewis, S., 2008. Partnership, policy and politics: initial teacher education in England under New Labor. *Teach. Theor. Pract.* 14 (4), 307–318.
- Grossman, P., Hammerness, K.M., McDonald, M., Ronfeldt, M., 2008. Constructing coherence: structural predictors of perceptions of coherence in NYC teacher education programs. *J. Teach. Educ.* 59 (4), 273–287.
- Gustafson, J.M., 1982. Professions as "callings". *Soc. Serv. Rev.* 56 (4), 501–515.
- Ingersoll, R.M., Collins, G.J., 2018. The status of teaching as a profession. In: Ballantine, J., Spade, J., Stuber, J. (Eds.), *Schools and Society: A Sociological Approach to Education*, sixth ed. Pine Forge Press/Sage Publications, CA, pp. 199–213.
- Ingersoll, R., 2001. Teacher turnover and teacher shortages: an organizational analysis. *Am. Educ. Res. J.* 38, 499–534.
- Jakhelln, R.E., Bjørndal, K.E., Stølen, G., 2016. Masteroppgaven—relevant for grunnskolelæreren? (The Master Thesis—relevant for primary school teachers?). *Acta Didact. Nor.* 10 (2), 193–211.
- Kelchtermans, G., 2017. "Should I stay or should I go?" Unpacking teacher attrition/retention as an educational issue. *Teach. Theor. Pract.* 23 (8), 961–977.
- Kelchtermans, G., 2019. Early career teachers and their need for support: thinking again. In: *Attracting and Keeping the Best Teachers*. Springer, Singapore, pp. 83–98.
- Kochan, F.K., Kunkel, R.C., 1998. The learning coalition: professional development schools in partnership. *J. Teach. Educ.* 49 (5), 325–333.
- Korthagen, F., Loughran, J., Russell, T., 2006. Developing fundamental principles for teacher education programs and practices. *Teach. Theor. Pract.* 22 (8), 1020–1041.
- Lillejord, S., Børte, K., 2016. Partnership in teacher education—a research mapping. *Eur. J. Teach. Educ.* 39 (5), 550–563.
- Lillejord, S., Børte, K., 2020a. Middle leaders and the teaching profession: building intelligent accountability from within. *J. Educ. Change* 21 (1), 83–107.
- Lillejord, S., Børte, K., 2020b. Trapped between accountability and professional learning? School leaders and teacher evaluation. *Prof. Dev. Educ.* 46 (2), 274–291.
- Lipsky, M., 1980. *Street-Level Bureaucracy: The Dilemmas of Individuals in Public Service*. Russell Sage Foundation, New York.
- Lundahl, L., Arreman, I.E., Lundström, U., Rönnerberg, L., 2010. Setting things right? Swedish upper secondary school reform in a 40-year perspective. *Eur. J. Educ.* 45 (1), 46–59.
- Maandag, D.W., Deinum, J.F., Hofman, A.W., Buitink, J., 2007. Teacher education in schools: an international comparison. *Eur. J. Teach. Educ.* 30 (2), 151–173.
- Mamadaliyev, A.M., Svechnikova, N.V., Miku, N.Y.V., Mădico, A., 2019. The German system of public education in the period between the 15th and early 20th centuries. Part 2. *Eur. J. Contemp. Educ.* 8 (3), 646–654.
- Mayer, D., Oancea, A., 2021. Teacher education research, policy and practice: finding future research directions. *Oxf. Rev. Educ.* 47 (1), 1–7. <https://doi.org/10.1080/03054985.2021.1853934>.
- Mayer, D., Cotton, W., Simpson, A., 2017. Teacher education in Australia. In: *Oxford Research Encyclopedia of Education*.
- Mehta, J., 2013. From bureaucracy to profession: remaking the educational sector for the twenty-first century. *Harv. Educ. Rev.* 83 (3), 463–488.

- Ministry of Education and Research (MER), 2008–2009. The Teacher—The Role and the Education (Whitepaper 11 2008–2009). St.meld. nr. 11 (2008–2009)—regjeringen.no.
- Morris, P.W., et al., 2006. Exploring the role of formal bodies of knowledge in defining a profession—the case of project management. *Int. J. Proj. Manag.* 24 (8), 710–721. <https://doi.org/10.1016/j.ijproman.2006.09.012>.
- Opoku, E.N., Khuabi, L.A.J.N., Van Niekerk, L., 2020. Exploring the Factors that Affect the Transition From Students to Health Professionals: An Integrative Review. Preprint Research Square 170c2a43-58e7-40f8-a991-a98dd9b1b8f5.pdf. researchsquare.com.
- Paula, L., Grinfelde, A., 2018. The role of mentoring in professional socialization of novice teachers. *Probl. Educ.* 76 (3), 364.
- Russell, T., 2013. Has reflective practice done more harm than good in teacher education? *Phronesis* 2 (1), 80–88.
- Schön, D.A., 1983. *The Reflective Practitioner. How Professionals Think in Action.* Jossey-Bass, San Francisco.
- Simola, H., Kivinen, O., Rinne, R., 1997. Didactic closure: professionalization and pedagogic knowledge in Finnish teacher education. *Teach. Teach. Educ.* 13 (8), 877–891.
- Simons, P.R.J., Ruijters, M.C., 2014. The real professional is a learning professional. In: *International Handbook of Research in Professional and Practice-Based Learning.* Springer, Dordrecht, pp. 955–985.
- Smith, K.A., Waller, A., 1997. Afterward: new paradigms for college teaching. In: Campbell, W.E., Smith, K.A. (Eds.), *New Paradigms for College Teaching.* Interaction Book Company, pp. 269–281.
- Voss, T., Kunter, M., 2020. “Reality shock” of beginning teachers? Changes in teacher candidates’ emotional exhaustion and constructivist-oriented beliefs. *J. Teach. Educ.* 71 (3), 292–306.
- Weick, K.E., 1976. Educational organizations as loosely coupled systems. *Adm. Sci. Q.* 21 (1), 1–19.
- Weiner, G., 2002. Uniquely similar or similarly unique? Education and development of teachers in Europe. *Teach. Educ.* 13 (3), 273–288.
- Wieser, C., 2016. Teaching and personal educational knowledge—conceptual considerations for research on knowledge transformation. *Eur. J. Teach. Educ.* 39 (5), 588–601.
- Williams, D., Coles, L., 2007. Teachers’ approaches to finding and using research evidence: an information literacy perspective. *Educ. Res.* 49 (2), 185–206.
- Zeichner, K.M., Liston, D.P., 1990. Traditions of reform in US teacher education. *J. Teach. Educ.* 41 (2), 3–20.
- Zeichner, K., 2010. Rethinking the connections between campus courses and field experiences in college- and university-based teacher education. *J. Teach. Educ.* 61 (1–2), 89–99.

International perspectives on community-engaged teacher education

Rocío García-Carrión^a and Eugenia Allotey^b, ^a University of Deusto, Ikerbasque. Basque Foundation for Science, Bilbao, Spain; and ^b University of Deusto, Bilbao, Spain

© 2023 Elsevier Ltd. All rights reserved.

Introduction	375
Why do we need community-engaged teacher education?	376
The power of dialogic interactions between teachers, students, and communities	377
Raising critical consciousness by discussing evidence with social impact	377
Funds of knowledge and cultural intelligence to dismantle deficit thinking	378
Conclusion	379
References	379

Introduction

Addressing educational inequalities in the 21st century need to have teachers prepared to work effectively with students, their families and their communities, especially with those from vulnerable populations. Indeed, the role of teachers is crucial to creating optimal conditions for learning and development. Granted that without authentic knowledge of the community culture prospective teachers will not have the context they require to be effective teachers (Zygmunt and Cipollone, 2019), current research in the science of learning and development emphasizes the importance of connecting teachers with family and community contexts (Darling-Hammond et al., 2020). Hence, educators need to be prepared to develop strong and respectful relationships by understanding and building on children's knowledge and experience, which are shaped by teachers' interactions with their families and their communities in a particular socio-cultural context (Vygotsky, 1978; Flecha, 2015). In the current situation, where educational inequalities have been exacerbated even more due to global crises, such as the pandemic or the war, community-engaged teacher education emerges as a transformational pathway to prepare teachers to work effectively with students, particularly with those from vulnerable communities, and to deeply engage with those who have suffered systematic marginalization. This should ultimately contribute to guaranteeing the right of every single child to an inclusive and quality education for all (United Nations, 2015).

Several social-justice scholars have developed frameworks and programs to prepare excellent, socially just and equity focused teachers to work for and alongside vulnerable communities (Lampert, 2021; Zygmunt et al., 2022; Cochran-Smith et al., 2016; Villegas, 2007). These models are conceived within innovative and transformative paradigms aiming at advancing educational equality. Thus, the Australian approach to building community engagement starts from the premise that "deeper engagement with vulnerable communities has the potential to develop community-engaged, culturally responsive [anti-deficit] teachers who are prepared to both reach and teach children" (Lampert, 2021, p. 457). It aimed at developing teacher education programs in Australia that train teachers to respect local and contextual differences of communities and incorporate what they have learned into their work. Along these lines, the Nexus program (Lampert and Browne, 2022) was designed to prepare a small, carefully mentored cohort of teachers for hard-to-staff, low socio-economic status schools in urban and rural Victoria in Australia by incorporating a new school-embedded Master of Secondary Teaching with community engagement at its core (p. 149). It examines applicants background, commitment to social justice, among others to consciously recruit a high number of culturally diverse teachers (Lampert and Browne, 2022) paving way for the debate on teacher quality and efficiency (Anagnostopoulos et al., 2016). Nonetheless, the approach enabled designers of the program to understand candidates' beliefs and to explore their interpretation of social justice in order to design a teacher education course on the principles of community engagement and social justice (Lampert and Browne, 2022).

This is consistent with the community-engaged teacher preparation approach that has envisioned a better future for minority students in American schools by privileging the knowledge and expertise of their communities and elevating people's funds of knowledge (Moll et al., 1992) and community cultural wealth (Yosso, 2005). Importantly, community-engaged approach to teacher preparation differs from what is known as community-based initiatives, which simply considers being placed in a community, and it emphasizes the need to "authentically integrating educator preparation programs into historically marginalized communities" (Zygmunt et al., 2022). Likewise, engaging pre-service teachers with adult community members as part of their preparation in a predominantly African American neighborhood in the United States led to teachers' becoming more adept in cultural understanding and acquiring an asset-based mindset as opposed to a deficit one (Keller, 2019). In South Africa, the model's positive impact contributed to the development of policies, standards and guidelines for the implementation of community engagement in higher education at the national level (Lazarus et al., 2008). Considering the history of African Americans in the United States and the Apartheid regime in South Africa, these demonstrably, reveal an approach to teaching and learning that advocates for equity among vulnerable groups.

Similarly, the dialogic approach to community-engaged teacher education has aimed at engaging teachers in transformative dialogs and interactions among them and within those communities who have systematically been marginalized to ultimately achieve

children's educational and social success (García-Carrión et al., 2017). Underlying this dialogic approach and aligned with other international perspectives on community engaged teacher preparation, there is the Freire dialogic action theory (Freire, 2003). Particularly, his notion of praxis—"the action and reflection of men and women upon their world in order to transform it" (p. 79) becomes essential for preparing teachers as agents of change in and within communities, especially with those systematically marginalized. For that goal to be achieved, community-engaged teacher education must be informed, on the one hand, by scientific evidence that has demonstrated to report educational and social improvements and, on the other, by the knowledge co-created in dialog with the community. This implies monitoring teacher education to achieve social impact (Flecha et al., 2019).

Indeed, providing the teachers with relevant knowledge is not enough: such knowledge has to be effectively used and fully transferred into practice to make the world a better place (Ball, 2012). This process requires, on the one hand, raising teachers' critical consciousness by discussing social impact research in education through egalitarian dialog, and on the other hand, the co-creation of new knowledge with families and communities for dismantling prejudices and deficit thinking (Valencia, 2010). This also implies a dialog that values their funds of knowledge (Moll et al., 1992) and builds on their cultural intelligence (Flecha, 2000). Within this approach all social actors -teachers, researchers, families, and communities-are mutually engaged in a partnership or network of knowledge-co-creation through dialogic learning as they have the shared purpose of creating optimal conditions for every single child to thrive.

This dialogic approach to community-engaged teacher education has been developed over three decades of collaboration between the Community of Research on Excellence for All (CREA) along with teachers, schools, and communities, including those systematically excluded and marginalized, living in high poverty areas. Indeed, this approach has nurtured community-engaged teacher education in Europe and South America. Transferred to many diverse contexts worldwide, the approach described here has shown to be applicable to any educational attempt in which children can benefit from teachers who are equipped with the knowledge and the commitment to transform schools into learning communities oriented to achieving educational and social success for all (García-Carrión et al., 2017).

Why do we need community-engaged teacher education?

Research have widely demonstrated that teacher education is a key factor in achieving school success, because it has the potential to influence students' learning and academic performance (Baker et al., 2008; Mutton et al., 2010; Herrera, 2018) and have a stronger effect on a students' academic outcomes than, for example, the pupils' socioeconomic status (Darling-Hammond, 2000). However, scientific evidence on teacher education might be insufficient if it does not reach vulnerable communities and address some of their most pressing challenges and needs. Indeed, tackling the challenge of preparing teachers to work in culturally and linguistically complex classrooms or in high-poverty schools have reflected the crucial role of preparing teachers within those communities (Ball, 2009; Lampert et al., 2017). Thus, teachers and teacher educators must work with under-represented communities to identify respectful models of engagement, especially with those people whose voices have been rarely included in educational discourses and practices (Lampert, 2021). Although community-engaged teacher preparation programs have already pointed out the need for establishing meaningful synergies among teachers-to-be (coming from the university) and high poverty schools, less is known about the most effective mechanisms for establishing such positive cooperation.

In this regard, a dialogic approach to teacher education opens the possibility for engaging in transformative interactions between teachers, students and community members, and it creates affordances for community-engaged teachers to act against inequity and exclusion. This approach is grounded on Flecha's (2000) dialogic learning. His conceptualization articulates seven principles that define those dialogs and interactions as (1) being egalitarian, (2) capitalizing on every participant's cultural intelligence, (3) provoking critical transformation instead of adaptation to the environment, (4) fostering the instructional dimension of dialog, (5) being solidarity-based, (6) sources of meaning making, and (7) allowing equality of differences. These principles guide the interactions and dialogs that teachers and learners engage with, not only among themselves, but also with their families and other community members (Flecha, 2015).

Hence, the community plays a major role and, if teachers want their children to thrive, they need to engage with their communities to transform the sociocultural context. Since cognitive development cannot occur outside of this context or independent of social interaction (Vygotsky, 1978), teachers need to engage with their students' cultural context and to consider those relationships in their preparation and in their professional development. Decades of research highlights the potential of this approach to enhance academic achievement and social cohesion even in schools serving pupils belonging to socio-economically disadvantaged households (Roma, migrants and refugees, among others) (Flecha et al., 2019; García-Carrión et al., 2017).

In this regard, community-engaged teachers develop specific structures and criteria that ensure the involvement of migrant and vulnerable groups in school daily life reporting benefits in student learning and social cohesion (Soler et al., 2020). Specifically, they encouraged families and community members to engage in educational activities, capitalizing on families' and communities' funds of knowledge (Moll et al., 1992) and dismantling the deficit thinking (Valencia, 2010). Accordingly, families and community members are perceived as a great asset and become teachers' allies to improve students' education but also reporting psychosocial benefits to the families and community themselves (García et al., 2019; Gírbés-Peco et al., 2019). As a historically situated social practice, community engaged teacher education is needed to ensure high-quality education for all, but specially, for students from historically marginalized groups.

The power of dialogic interactions between teachers, students, and communities

Using the dialogic approach to community engaged teacher education has enabled teachers to better understand how to face challenges they encounter when entering schools serving historically marginalized groups or located in high-poverty areas (García-Carrión et al., 2020). Furthermore, this approach has been transferred to many diverse schools worldwide and it has been widely applied and researched within the project “Schools as Learning Communities” which are grounded on the dialogic learning approach (Flecha, 2000). Currently, they form an international network of schools across countries and across educational levels from early childhood education to adult education, that allows for all children to achieve success regardless of their socioeconomic, ethnic, geographic or minority background, reversing the reproduction of social and educational inequalities (Soler-Gallart and Rodrigues de Mello, 2020). Dialogic interactions and the participation of the community are two main cornerstones of this project where schools implement Successful Educational Actions (Flecha, 2015) that are based on scientific evidence of the most fundamental theoretical contributions in education as well as on the most recent research findings. As a result of its impact, the European Commission (2011) recommended this model from preventing early school leaving that mainly affect vulnerable students, and it is defined as follows:

Schools as “learning communities” agree on a common vision, basic values and objectives of school development. It increases the commitment of pupils, teachers, parents and other stakeholders and supports school quality and development. “Learning communities” inspire both teachers and pupils to seek improvement and take ownership of their learning processes. It creates favourable conditions also for reducing school drop-out and for helping pupils at risk of dropping out.

European Commission (2011).

By engaging teachers in dialogic interactions with the communities, a dialogic space is created for them to overcome the pervasive deficit thinking about those historically marginalized and the “cruel optimism” of involving with vulnerable groups in superficial ways (Rasmussen, 2015). This occurs, for example, with in-service teachers in Schools as Learning Communities when they create a dialogic space to dream together with the entire community the school they want for their children and for themselves. Hence, they create a collective understanding of people’s needs and desires which often challenges prejudices, specially toward historically disadvantage groups such as the Roma. Indeed, the study conducted by García Yeste et al. (2018) showed a positive impact on Roma students and their families’ expectations to higher education, and how these high expectations affected also students’ outcomes. Thus, this model tears down the walls of hermetic schools where only high-educated ones can be part of the school (Souto-Manning and Swick, 2006). Since learning and development lie in the interactions children have with every person who is part of their contexts (not only schools, but playgrounds, neighborhood, families’ workplaces, among others), community-engaged teachers can open this opportunity to access high-quality and diverse interactions.

Pre-service teachers can also learn and engage in these dialogic interactions. Recent research involving 129 student teachers, 20 schoolteachers and 3 headteachers in Spain found that the gap between theory and practice in current teacher education can be reduced when student teachers are involved in schools where practices based on scientific evidence are implemented (Ríos González et al., 2019). In this regard, a volunteering program on Schools as Learning Communities was designed and implemented with student teachers who wanted to take part. Indeed, volunteering programs involving pre-service teachers have shown evidence of the improvement of children’s learning. By volunteering at the schools, pre-service teachers not only acquire meaningful and relevant professional training but also help the pupils to progress in their learning processes (Ríos González et al., 2019). This win-win equation is promoted under this dialogic approach to teacher education in Initial Teacher Training, through programs that aim to facilitate pre-service teacher involvement in Schools as Learning Communities.

Indeed, the experienced of pre-service teachers participating as volunteers in School as Learning Communities is shaped by transformative educational practices based on scientific evidence. This has been especially helpful to challenge their previous assumptions, such as misconceptions concerning the placement of migrant pupils in catch-up groups. As a result, the dialogic teacher training approach has changed the students teachers’ views of children with minority background, nurturing high expectations toward them. These first experiences are crucial since they will shape pre-service teachers’ future believes and practices (Ríos González et al., 2019).

Raising critical consciousness by discussing evidence with social impact

According to Freire (2003) critical consciousness is needed for the teachers to intervene in the world and transform that world. In his original term, *conscientização* enables a critical position in relation to power since it invites people to recognize reality and name social injustice. Dialogic teacher education fosters raising teachers’ critical consciousness as defined by Flecha et al. (2019: 2), it allows for “spaces of collective construction of knowledge through the connection of theory, practice, and research.” Importantly, those dialogs are oriented to use scientific knowledge to support teacher development toward a successful education for all, moving from assumptions in education to those actions that have provided evidence of educational and social improvements (Flecha et al., 2019). This is aligned with the international movement “scientific evidence-based teacher education,” which puts international scientific evidence at the service of educational improvement (Georgiou et al., 2020). However, accessing the knowledge is not

enough. Creating dialogic spaces in teacher education to successfully work with diverse students and communities is essential. This is precisely what Dialogic Pedagogical Gatherings (DPG hereinafter) have offered to foster community-engaged teacher education.

In this dialogic space grounded on an egalitarian dialog, scientific literature on evidence-based approaches to teaching and learning is read and discussed. Particularly, the most important contributions in the educational field are selected for discussion in each session of DPG, such as reference books on the latest contributions to learning and teaching, as well as research published in international peer-reviewed journals whose results have reported social impact (Rodríguez-Oramas et al., 2021). In such context, participants are exposed to a collective and collaborative reasoning regarding teaching, learning, and educational actions that have previously been implemented and that have improved learning outcomes (Roca-Campos et al., 2015). The readings discussed at the DPG are chosen by all participants, who often include not only teachers but also family and community members. Research conducted with in-service teachers (Roca-Campos et al., 2021; Rodríguez-Oramas et al., 2021) have reported a positive impact on teachers' ability to address current school problems and to transfer new knowledge to the classroom which turned into improving student learning. In addition, DPG also open the possibility to build partnerships in the community and create networks among pre-service and in-service teachers to discuss lesson plans and practices with colleagues, drawing on scientific evidence (Rodríguez et al., 2020).

Similar research on DPG has been conducted with Master students in their preparation to become secondary school teachers in a Spanish university (García-Carrión et al., 2021). This case involved 40 student teachers who engaged in meaningful and egalitarian dialog with their peers. Participants read and discussed Paulo Freire's "Teachers as Cultural Workers: Letters to Those Who Dare Teach" (Freire, 2009) on weekly basis. The DPG allow them to reflect and gain critical awareness in a space where all voices and beliefs, supported by arguments, are welcome. Interactions among pre-service teachers fostered their critical thinking and created opportunities for students to critically reflect on teacher skills and values, their role as teachers and their social self, which allow them to reshape their future practices (García-Carrión et al., 2021). Moreover, by engaging in egalitarian dialog through these readings, student teachers felt free to express their thoughts, feelings and beliefs regarding key issues related to education (such as expectations toward students, relationships with families, or dialogs with the school staff). In this dialogic space, future teachers' own educational experiences emerged intertwined with assumptions about ethnic minorities or other vulnerable groups. Once these thoughts were openly shared during the DPG, Freire's text itself and teacher students' desire to be coherent teachers soon allowed them to overcome certain prejudices and misconceptions.

Funds of knowledge and cultural intelligence to dismantle deficit thinking

Through the dialogic interactions they engage in and the common vision of the school they construct, teachers, families, community members and researchers develop trust-based relationships through egalitarian dialog (Khalfaoui et al., 2020). "Dialog is egalitarian when it takes different contributions into consideration according to the validity of their reasoning, instead of according to the positions of power held by those who make the contributions" (Flecha, 2000, p. 2). In egalitarian dialogs individuals support their views with arguments and hold an attitude of openness to rectification on the basis of the arguments provided by others. To achieve egalitarian dialog people should be allowed to engage in dialog using their funds of knowledge (Moll et al., 1992) creating an appropriate context for people to develop enough confidence and use their cultural intelligence (Flecha, 2000). Similarly, to the funds of knowledge, which are the "historically accumulated and culturally developed bodies of knowledge and skills essential for household or individual functioning and well-being" (Moll et al., 1992, p. 133), the cultural intelligence is, a set of academic, communicative, and practical knowledge and abilities that develop through experiences in life, and are used to solve problems in everyday life (Flecha, 2000).

By reflecting and interacting directly with people from historically marginalized groups in a context where people's funds of knowledge and cultural intelligence are valued and privileged, diversity constitute a cultural and cognitive resource with great potential for teaching and learning. Thus, teachers and children can capitalize on those resources to enhance their learning in the classroom. This occurs, for example, when teachers organize the classroom in Interactive Groups and volunteers from the community, including illiterate Roma grandmothers participate supporting small heterogeneous groups of students while they carry out curricular tasks (Valls and Kyriakides, 2013). This dialogic approach in the classroom counteracts the pervasive "deficit thinking" which "blames the victim" instead of examining how teachers and schools can be structured and transformed, such as in the Interactive Groups, to support vulnerable students in learning and succeeding at school (Valencia, 2010).

In this sense, prejudices and deficit views about minorities and vulnerable populations fade as the priorities of the community align with those of community-engaged teachers committed to provide the best education for all children. When those biases are overcome and the voices of all people are considered in an equal epistemological level, co-creation processes emerge. Such processes acknowledge that all individuals have inherent capacities for communication and social interaction, and that everyone can have their own understanding of the world, generate knowledge, and change social structures. This paradigm is reflected in the Communicative Methodology of research (Gómez et al., 2019), which has been used in the research conducted in Schools as Learning Communities being a useful tool for achieving socially useful results to achieve equity and inclusion, involving the vulnerable populations in the entire research process.

Schools that open their doors and value the cultural intelligence of all communities become a real arena where prejudices and deficit thinking are dismantled. These schools open meaningful opportunities for vulnerable groups to participate and be listened to and make available adults with different cultural backgrounds and profiles to interact with students on equal terms to promote the

success of all the children. This context is an optimal setting in which teachers can become a great resource for overcoming prejudice, precisely by engaging in real educational actions that transform difficulties into opportunities.

Conclusion

Community-engaged teacher education is a transformational pathway necessary to prepare educators to work effectively with learners from vulnerable communities and to extensively engage with students who have suffered systematic exclusion and marginalization. This is based on the principles of social justice and the premise that teacher education should involve and be engaged with the community in which the school is located to create opportunities for transforming current educational inequalities into academic and social success, thus, contributing to the social obligation of ensuring an inclusive quality education for all (SDG 4).

Despite the creation of innovative and transformative programs to advance educational equality based on community-engaged teacher education, there is an urgency for the most effective mechanisms of establishing such positive synergies between teachers and the community. The dialogic community-engaged teacher education approach projects a cross-cultural approach based on dialogic learning that addresses this challenge. Implemented in over a thousand schools worldwide, it creates an opportunity for teachers, students and community members to engage in transformative interactions establishing trust-based relationships. This creates affordances for community-engaged teachers to act against inequity and exclusion. Scientific literature emphasizes that this dialogic approach has great potential to enhance in-service and pre-service teachers' critical consciousness, transform their teaching practices to improve children's academic performance and social cohesion in culturally diverse settings.

If teachers want their children to thrive in an increasingly diverse society, they need to engage with their communities by valuing their funds of knowledge and cultural intelligence to transform the sociocultural context. Providing teachers with that knowledge is not enough, such knowledge has to be effectively used and fully transferred into practice to change schools and communities into places where the dream of an equitable and successful education for all becomes a reality.

References

- Anagnostopoulos, D., Lingard, B., Sellar, S., 2016. Argumentation in educational policy disputes: competing visions of quality and equity. *Theory Pract.* 55 (4), 342–351.
- Ball, A., 2009. Toward a theory of generative change in culturally and linguistically complex classrooms. *Am. Educ. Res. J.* 46 (1), 45–72. <https://doi.org/10.3102/0002831208323277>.
- Ball, A.F., 2012. To know is not enough: knowledge, power, and the zone of generativity. *Educ. Res.* 41 (8), 283–293.
- Baker, C., Johnson, G.S., Williams, L., Perkins, D.G., Rainey, S.A., 2008. The highlander research and education center: utilizing social change-based models for public policy. *Race Gend. Class* 15 (3–4), 308–334.
- Cochran-Smith, M., Ell, F., Grudnoff, L., Haigh, M., Hill, M., Ludlow, L., 2016. Initial teacher education: what does it take to put equity at the center? *Teach. Teach. Educ.* 57, 67–78. <https://doi.org/10.1016/j.tate.2016.03.006>.
- Darling-Hammond, L., 2000. How teacher education matters. *J. Teach. Educ.* 5 (1), 166–173. <https://doi.org/10.1177/0022487100051003002>.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D., 2020. Implications for educational practice of the science of learning and development. *Appl. Dev. Sci.* 24 (2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>.
- European Commission, 2011. Tackling Early School Leaving: A Key Contribution to the Europe 2020 Agenda. European Commission, Brussels, Belgium. COM(2011) 18 Final.
- Flecha, R., 2000. *Sharing Words*. Rowman & Littlefield, Lanham, M.D.
- Flecha, R., 2015. *Successful Educational Action for Inclusion and Social Cohesion in Europe*. Springer Publishing Company, Berlin and Heidelberg.
- Flecha, R., Roca, E., López de Aguilera, G., 2019. Scientific Evidence-Based Teacher Education and Social Impact. *Encyclopedia of Teacher Education*. Springer Science and Business Media LLC, Berlin/Heidelberg, Germany, pp. 1–6.
- Freire, P., 2003. *Pedagogy of the Oppressed*, 30th anniversary ed. Continuum, New York.
- Freire, P., 2009. *Teachers as Cultural Workers: Letters to Those Who Dare Teach*. Westview Press, New York.
- García, C., Ruiz, L., Comas, M.A., 2019. Formación de Familiares y Voluntariado. Los Beneficios de la Solidaridad en la Escuela. *Multidiscip. J. Educ. Res.* 9 (2), 144–168. <https://doi.org/10.17583/remie.2019.4188>.
- García-Carrión, R., Gomez, A., Molina, S., Ionescu, V., 2017. Teacher education in schools as learning communities: transforming high-poverty schools through dialogic learning. *Aust. J. Teach. Educ.* 42 (4), 44–56. <https://doi.org/10.14221/ajte.2017v42n4.4>.
- García-Carrión, R., Padrós Cuxart, M., Alvarez, P., Flecha, A., 2020. Teacher induction in schools as learning communities: successful pathways to teachers' professional development in a diverse school serving students living in poverty. *Sustainability* 12 (17), 7146.
- García-Carrión, R., Khalfaoui, A., Marauri, J., Santiago-Garabeta, M., 2021. Alumnado y profesorado universitario dialogando juntos: innovación docente a través de las tertulias dialógicas. In: Bezanilla, M.J., García-Olalla, A., Arruti, A., Aláez, M. (Eds.), *Respuestas innovadoras al nuevo perfil del estudiante universitario*. Síntesis, pp. 77–86.
- García Yeste, C., Morlà, T., Ionescu, V., 2018. Dreams of higher education in the Mediterranean school through family education. *Front. Educ.* 3 (79). <https://doi.org/10.3389/educ.2018.00079>.
- Georgiou, D., Mok, S.Y., Fischer, F., Vermunt, J.D., Seidel, T., 2020. Evidence-based practice in teacher education: the mediating role of self-efficacy beliefs and practical knowledge. *Front. Educ.* 5, 559192. <https://doi.org/10.3389/educ.2020.559192>.
- Girbés-Peco, S., Gairal-Casadó, R., Torrego-Egido, L., 2019. Participación de mujeres gitanas y marroquíes en la formación de familiares: beneficios educativos y psicosociales. *Cultura y Educación* 1–26. <https://doi.org/10.1080/11356405.2019.1656487>.
- Gómez, A., Padrós, M., Ríos, O., Mara, L.C., Pukepuka, T., 2019. Reaching social impact through communicative methodology. Researching with rather than on vulnerable populations: the roma case. *Front. Educ.* 4, 1–8. <https://doi.org/10.3389/educ.2019.00009>.
- Herrera, A.L., 2018. Decálogo de Competencias y sus Indicadores para Gestión de Capital Humano Universitario. *Multidiscip. J. Educ. Res.* 8 (1), 56–85. <https://doi.org/10.17583/remie.2018.2990>.
- Keller, D.B., 2019. Requisite community engagement for teacher education: a different take on service learning. *Athens J. Educ.* 6 (2), 93–110. <https://doi.org/10.30958/aje.6-2-1>.
- Khalifaoui, A., García-Carrión, R., Villardón-Gallego, L., 2020. Bridging the gap: engaging Roma and migrant families in early childhood education through trust-based relationships. *European. Early Child. Educ. Res. J.* 28 (5), 701–711. <https://doi.org/10.1080/1350293X.2020.1817241>.

- Lampert, J., 2021. A community-engaged framework for the preparation of teachers for high-poverty communities. *Aust. Educ. Res.* 48, 449–466. <https://doi.org/10.1007/s13384-020-00406-8>.
- Lampert, J., Browne, S., 2022. Examining Teacher Candidates' Backgrounds, Experiences, and Beliefs as Precursors for Developing Dispositions for Democracy. *Teachers College Record*, 01614681221086995.
- Lampert, J., Burnett, B., Comber, B., Ferguson, A., Barnes, N., 2017. "It's not about punitive": exploring how early-career teachers in high-poverty schools respond to critical incidents. *Crit. Stud. Educ.* 61 (2), 149–165. <https://doi.org/10.1080/17508487.2017.1385500>.
- Lazarus, J., Erasmus, M., Hendricks, D., Nduna, J., Slamati, J., 2008. Embedding community engagement in South African higher education. *Educ. Citizen Soc. Justice* 3 (1), 57–83.
- Moll, L.C., Amanti, C., Neff, D., González, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141. <http://www.jstor.org/stable/1476399>.
- Mutton, T., Burn, K., Hagger, H., 2010. Making sense of learning to teach: learners in contexts. *Res. Pap. Educ.* 25 (1), 73–91. <https://doi.org/10.1080/02671520802382912>.
- Rasmussen, M., 2015. "Cruel optimism" and contemporary Australian critical theory in educational research. *Educ. Philos. Theor.* 47 (2), 192–206. <https://doi.org/10.1080/00131857.2013.793929>.
- Ríos González, O., García, C., Jiménez, J., Ignatiou, Y., 2019. Student teachers volunteering in pre-service programmes in successful schools: contributing to their successful training. *Educación XX1* (1), 22. <https://doi.org/10.5944/educxx1.22620>.
- Roca-Campos, E., Gómez, A., Burgués, A., 2015. Luisa, transforming personal visions to ensure better education for all children. *Qual. Inq.* 21 (10), 843–850. <https://doi.org/10.1177/1077800415614026>.
- Roca-Campos, E., Renta-Davids, A.I., Marhuenda-Fluixá, F., Flecha, R., 2021. Educational impact evaluation of professional development of in-service teachers: the case of the dialogic Pedagogical Gatherings at Valencia "on giants' shoulders". *Sustainability* 13, 4275. <https://doi.org/10.3390/su13084275>.
- Rodríguez, J.A., Condom-Bosch, J.L., Ruiz, L., Oliver, E., 2020. On the shoulders of giants: benefits of participating in a dialogic professional development program for in-service teachers. *Front. Psychol.* 11, 5. <https://doi.org/10.3389/fpsyg.2020.00005>.
- Rodríguez-Oramas, A., Morla-Folch, T., Vieites Casado, M., Ruiz-Eugenio, L., 2021. Improving students' academic performance and reducing conflicts through family involvement in primary school learning activities: a Mexican case study. *Camb. J. Educ.* 52 (2), 235–254. <https://doi.org/10.1080/0305764X.2021.1973374>.
- Souto-Manning, M., Swick, K.J., 2006. Teachers' beliefs about parent and family involvement: rethinking our family involvement paradigm. *Early Child. Educ. J.* 34 (2), 187–193. <https://doi.org/10.1007/s10643-006-0063-5>.
- Soler-Gallart, M., Rodrigues de Mello, R., 2020. Schools as learning communities. DIO Press Incorporated.
- Soler, M., Morla-Folch, T., García-Carrión, R., Valls, R., 2020. Transforming rural education in Colombia through family participation: the case of school as a learning community. *J. Soc. Sci. Educ.* 4, 67–80. <https://doi.org/10.4119/jsse-3251>.
- United Nations General Assembly, Transforming Our World: The 2030 Agenda for Sustainable Development, 2015. A/RES/70/1, Available at. <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>. (Accessed 10 June 2022).
- Valencia, R.R., 2010. Dismantling Contemporary Deficit Thinking: Educational Thought and Practice. Routledge.
- Valls, R., Kyriakides, L., 2013. The power of interactive groups: how diversity of adults volunteering in classroom groups can promote inclusion and success for children of vulnerable minority ethnic populations. *Camb. J. Educ.* 43, 17–33. <https://doi.org/10.1080/0305764X.2012.749213>.
- Villegas, A.M., 2007. Dispositions in teacher education: a look at social justice. *J. Teach. Educ.* 58 (5), 370–380.
- Vygotsky, L.S., 1978. *Mind in Society: Development of Higher Psychological Processes*. Harvard university press.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91. <https://doi.org/10.1080/1361332052000341006>.
- Zygmunt, E.M., Cipollone, K., 2019. Community-engaged teacher education and the work of social justice. *J. Fam. Consum. Sci.* 111 (1). <https://doi.org/10.14307/jfcs111.1.15>.
- Zygmunt, E., Cipollone, K., Clark, P., Tancock, S., Community-Engaged Teacher Preparation. Oxford Research Encyclopedia of Education, Retrieved 10 Jun. 2022, from <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-476>.

University–school partnerships in initial teacher education

Maureen Robinson^a and Jennifer Feldman^b, ^a Department of Curriculum Studies, Faculty of Education, Stellenbosch University, Stellenbosch, South Africa; and ^b Department of Education Policy Studies, Faculty of Education, Stellenbosch University, Stellenbosch, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	381
Purpose of university–school partnerships	382
Benefits of university–school partnerships	382
Models of partnership	383
HEI-led	383
Collaborative	383
Separatist	383
A complementary model	383
A collaborative model	384
An equivalent model	384
A community of practice model	384
Research and innovation	384
Democratic professionalism	384
Key concepts	385
Cultural historical activity theory	385
“Third space”	385
Boundary crossing	385
Expansive learning	386
Horizontal expertise	386
Common knowledge and relational expertise	386
Sociocultural perspectives on professional learning	386
Teacher education pedagogy	387
Key considerations for universities and schools	387
Oversight and responsibility	388
Shared vision	388
Trust	388
Time	388
Sustainability	389
Equal partners	389
Teacher mentors	389
Acknowledgment and recognition of teacher mentors	389
Professional inquiry	390
Concerns and caveats	390
Concluding comments	390
References	391

Introduction

This article provides an overview of core issues in research on university–school partnerships in initial teacher education. The article shows how the notion of university–school partnerships can be conceptualized in several ways, with varying permutations and models of such partnership relationships. The article further identifies the key considerations for universities and schools, and outlines some of the challenges related to the implementation of such partnerships.

A few clarifications and delineations are relevant. The structure of initial teacher education programs varies widely across the globe, including in number of years of the qualification and length of time spent in schools. As it is impossible to do justice to all such structures, the discussion here is aimed more at principles of partnerships than at examples of specific strategies or programs. In certain instances, though, examples of interventions are given by way of illustration.

Initiatives to promote university–school partnerships can occur within the full range of a teacher education curriculum, including the subject-specific as well as the foundational disciplines. The focus here is on the school-based period of initial teacher education as this is where most of the relevant writings can be found. Different terminology for this school-based period is used across the globe, including practicum, Teaching Practice, clinical experience, clinical practice, field experiences, and work-

integrated learning; this article uses “practicum” to denote all of these terms. It is important, however, that the notion of “partnership” not be conflated with that of the practicum. While the practicum assumes a relationship between universities and schools, a partnership refers, at the very least, to a more systematic and formalized arrangement, with a recognized level of shared responsibility between universities and schools for the education of student teachers.

The field under consideration is very vast, and it is not possible to report on the full range of research available. The information presented has been gathered through cross-referencing of core articles in academic journals and edited books and a selection of country reports. Emphasis has been placed on overview writings, rather than on outlines of specific research projects or case studies, of which there are very many. Interestingly, the emphasis on academic articles as source material has meant that most of the information presented is from the perspective of the university; further referencing of perspectives from schools and teachers is encouraged.

Purpose of university-school partnerships

Forms of university-school partnerships may differ, nevertheless, all share certain common purposes. The most prevalent of these is to build bridges between the oft-cited divide between the world of “theory” (the university) and the world of “practice” (the school). A partnership is intended to create learning opportunities that enable student teachers to draw on the conceptual as well as the practical resources of the university and the school and, in so doing, to advance their understanding, insights and competences for teaching. Through gaining access to theoretical resources as well as to the professional knowledge of teachers, student teachers are seen to be in a stronger position to develop their own teacher identities and competences.

Despite this common purpose, the interest in establishing university-school partnerships may derive from different perspectives on the roles and contributions of each of the partners. In one approach, partnerships are seen as a way to address the critique that university-based teacher education prioritizes theory over practice (Cooper and Grudnoff, 2017), and the resultant disconnect between what students learn in initial teacher education and what they need to do in practice:

The notion of a “theory-practice gap” is based on the perception that university models of teacher education overly emphasize theory, values and beliefs at the expense of actual teaching practice, thus leaving new teachers on their own to implement or translate university-produced theory into classroom-ready practice

Cochran-Smith et al. (2020, 47).

A second purpose for partnerships argues for a reciprocal relationship between theory and practice, as opposed to a “theory-into-practice” or a solely “practice-based” approach (Cochran-Smith et al., 2020). Attempting to reverse a theory-practice binary, theory from this perspective sees practice as the source and center for professional reflection and development (Mitchell and Reid, 2017). Theory and practice are thus not seen as dichotomous, nor is practice seen as non-theoretical; rather theory and practice are seen as mutually and synergistically related (Cochran-Smith et al., 2020). Through building on such synergies, the intention is that both academic knowledge and pedagogical skills are strengthened.

In both approaches, the combination of theory and practice is seen as a way of preparing future teachers better for the world of work. The first approach, however, can be said to be more instrumental, as the emphasis is on the mastery of skills, and socialization into the world of the school (Orrell and Higgs, 2012). The second approach cautions against an uncritical celebration of practice, arguing rather for “new kinds of pedagogical and curriculum practices to emerge from a new configuration” (Boud, 2012: 55).

Benefits of university-school partnerships

University-school partnerships have been argued to offer significant benefits to the student teacher, the school and the university, thus holding the potential for supporting and improving teacher quality. The following are some of the benefits as highlighted in the literature:

- For student teachers, partnerships are said to offer more authentic learning experiences, more structured teaching practicum experiences and feedback from teacher mentors, and more varied opportunities to involve themselves in school activities (Darling-Hammond, 2010).
- Partnerships can provide opportunities for inquiry that brings together students’ university course-work theory with school-based practice (Zeichner, 2010).
- By integrating university theory with school-based practice students learn to “handle the problems of everyday teaching through theory-guided action” (Korthagen et al., 2006: 1021). In so doing, students learn to analyze and apply theory, reflect on practice, and have additional opportunities to retry and improve in their practice (Darling-Hammond, 2006).
- For the schools, the main benefit of partnership agreements is the opportunity that they provide for teachers to take time to reflect on their own practice as they mentor the students in lesson planning and classroom management (Green et al., 2020).

- Partnerships encourage teacher mentors to “consistently look at new and engaging pedagogies, as well as providing the stimulus to question traditional methods” thus contributing to the school-based educators’ professional development (Miller et al., 2015: 63).
- For university lecturers, the benefit of partnerships includes the development of contextual knowledge for teaching. Teacher educators are encouraged to keep up to date with the developments in schools and connect their teaching more closely with the complex realities within schools (Toon and Jensen, 2017).
- Authentic partnerships involve activities such as designing and evaluating programs, sharing data and information, observing and sharing teaching practices. When this takes place, Toom (2017) argues that partnerships play an important role in modeling teacher collaboration and collective curriculum work, as well as utilizing research on teaching and learning to develop innovative pedagogical programs through networked expert communities that involve university teacher educators as well as school-based educators.

The combination of these benefits is well summed up as follows:

Partnerships play an important role in producing and utilizing research on teaching and learning, implementing innovative pedagogical developments, modeling teacher collaboration and collective curriculum work, producing teaching and learning materials, as well as constructing and developing networked expert communities of TEs and school-based teacher educators. Furthermore, this helps teacher educators to keep up to date with school-based developments and contributes to school-based teacher educators’ professional learning and growth

Toom (2017).

Models of partnership

Partnerships may be constructed according to different models, each of which brings different priorities, expectations, roles and responsibilities for universities and schools. A number of different models as identified in the literature are outlined here; it should be noted though that these models constitute typologies. In practice, institutions are likely to work with a hybrid set of models as circumstances change over time, as different contexts and requirements emerge, or as particular relationships develop. Mixed models may even exist, as universities engage in a variety of arrangements with specific schools.

Furlong et al. (1996) have identified three “ideal-typical” models of partnership that were operating in the United Kingdom by the mid-1990s, namely higher education institution (HEI)-led, collaborative and separatist.

HEI-led

In the HEI-led partnership the school is seen as a resource that can provide practice-based learning opportunities for student teachers. The university defines the responsibilities of the school, and school-based mentors are expected to work according to these indicated tasks. Planning and monitoring are the responsibility of the university and is aimed at ensuring that students receive comparable learning opportunities at the different schools. While mentor teachers might be somewhat involved in the assessment of student teachers, this remains the primary responsibility of the university.

Collaborative

Within the collaborative model, teachers are seen as having an equally legitimate, but different body of professional knowledge from those in higher education. There is a negotiated relationship between the university and the school, and teachers and university lecturers work together to plan an integrated program of work for the students. Student teachers draw on, or may even critique, the different bodies of knowledge to develop their own professional identities and insights. Assessment is a collaborative responsibility of both the university and the school.

Separatist

In the separatist model schools and higher education institutions have different but complementary roles. Broad planning allows for the identification of agreed areas of responsibility. There is no systematic attempt to bring the worlds of school and university learning together and it is left to the student teacher to achieve integration of the two knowledge domains. Few opportunities for dialog exist between the university and the school and mentoring and assessment are primarily the responsibility of the school.

A complementary model

A second set of typologies identifies three conceptual models of university-school partnerships, namely collaboration, complementarity, and equivalence (Edwards et al., 2009). To this, these authors have added the concept of “community”, which they see as assimilating the characteristics of the other three concepts.

In a complementary partnership, the school and university are understood as distinctive entities with separate but complementary responsibilities. While no systematic attempt is made to bring the two dimensions together, a process of mutual renewal is envisaged through the symbiotic relationship that is established.

A collaborative model

In a collaborative partnership, teachers and lecturers work and plan together on a more regular and coordinated basis. Such partnerships require “personnel from schools and universities to engage mutually in negotiating the meaning of a shared agenda in which neither party’s meaning is privileged” (Edwards et al., 2009: 13). Collaborative partnerships include all members of the partnership—student teacher, teacher mentor, teacher educator—who work as a team toward a common professional goal. Collaboration between all stakeholders would therefore assist with the development of respect and understanding of one another’s roles and professional learning then becomes a non-hierarchical endeavor where each person’s knowledge and expertise is acknowledged equally (Payne and Zeichner, 2017).

An equivalent model

In an equivalent partnership, partners have equal status and share responsibilities in all areas, including planning, supervision and assessment.

A community of practice model

Drawing on critiques of the above models, Edwards et al. (2009) outline a shift in research literature toward the notion of “community” in university-school partnerships. Key concepts here are communities of practice, professional learning communities, or communities of inquiry. Learning is understood as a social process that involves “mutual engagement in a joint enterprise and participation in the negotiation of the meanings of a shared repertoire” (Edwards et al., 2009: 14).

Research and innovation

In this next conceptualization of partnership, the emphasis is on the promotion of research and/or curriculum innovation. Opportunities are created for student teachers and teacher educators to work with practicing teachers toward the advancement of knowledge on aspects of teaching and learning. The partnership is based on bringing “research-based understandings of teaching and learning into dialog with the professional understandings of experienced teachers” (Burn and Mutton, 2015: 219), thus promoting research-informed clinical practice. Examples of this model include the Oxford Internship scheme in the United Kingdom, Professional Development Schools in the USA, realistic teacher education in the Netherlands, and Teacher Training Schools in Finland (Burn and Mutton, 2015).

Darling-Hammond (2014) describes highly developed professional development school models in the USA as places where curriculum reforms are supported by the school and the district; teams from the university and the school work on curriculum development, school reform and action research; university faculty are involved in teaching and professional development at the school; and school educators often teach on the teacher education program. Most classrooms in these schools are sites for student teaching placements and cooperating teachers are trained as teacher educators and mentors. Some such schools explicitly include an equity agenda, thus helping student teachers to understand the broader institutional and social context of schooling.

A further example of a partnership aimed at promoting research and innovation can be found at the University of Johannesburg in South Africa (Henning et al., 2015). The Bachelor of Education program is purposefully designed around the centrality of child study; a key part of professional learning is seen as having a sound understanding of child development, “how children, learn, change and develop over time” (Gravett et al., 2019: 3). The aim is to develop strong professional practice knowledge through an inquiry-orientation within a school placement.

Democratic professionalism

A further typology is deliberately aimed at building inclusive spaces for the active participation of non-dominant families and communities in teacher education, including in the conceptualization of the program. The cornerstones of such partnerships are respect for community knowledge and a mutually beneficial interaction of teacher educators, families, and communities:

[this is] a form of democratic professionalism ... where current power hierarchies are reduced and where teachers, teacher unions, and teacher educators work in more genuinely collaborative ways with families and communities in schools, in the programs that educate teachers, and in communities
 Zeichner (2019: 5).

An example of such an approach are some of the urban teacher residencies (UTRs) in the United States. Teacher candidates here spend time in schools, often seeking shared responsibility for teacher education among university, school and, in certain cases, local communities. While UTRs differ in design and ideology, there are those that seek deliberately “to be more relevant and connected to problems of practice in public schools and to successfully draw on the expertise that is located in universities, schools and local communities under a new more democratic form of teacher education governance” (Zeichner, 2014: 262). The intention here is to integrate academic, practitioner, community and student knowledge, and not to privilege one over the other (Klein et al., 2016).

Another example of a partnership that works within this approach can be cited from South Africa. [Silbert \(2019\)](#) discusses a university-school partnership at the University of Cape Town that works across a strong social divide and stark socioeconomic differences between the university and the school. Reflecting on these power differentials, the project sought to move beyond normative assumptions of reciprocity and mutuality to the co-creation of spaces for the collective empowerment and solidarity of partners. In so doing, the project hoped to establish a “communicative ethic which encompasses dialogic engagement and moral humility” ([Silbert, 2019: 12](#)).

The variations in these models of partnership give rise to different expectations and structures for schools and universities. However, even within these differences, the broad purpose remains the same, namely, to provide opportunities for perspectives on teaching and learning to be shared in mutually beneficial ways across schools, universities and, in some instances, communities.

Key concepts

This section outlines some of the more prominent conceptual categories of analysis that are drawn on in the research on university-school partnerships.

Cultural historical activity theory

Much of the conceptual work on university-school partnerships draws on the framework of cultural historical activity theory (CHAT) associated with the work of Engeström (see [Ellis et al., 2010](#)). CHAT is a useful framework, as it helps to identify the spaces that need to be traversed when working across different systems, as in a university and a school, each with its respective cultures, histories, organizational arrangements, norms, etc.

Six analytic categories are associated with CHAT, namely subject, object, outcome, instruments/tools, rules, community, and division of labor ([Tsui et al., 2009: 32](#)). These categories might be differently understood and enacted within the purpose, organization and culture of universities and schools, and this may well create tensions in school-university partnerships ([Edwards et al., 2009](#)). These differences include purpose, function, structure, rules and regulations, calendars and schedules, work routines, orientation to teaching. For example, university academics and schools might share the same commitment to quality education (object), as well as similar expectations for the partnership of knowledge sharing, practical experience, and interaction with experienced teacher mentors (outcome). At the same time, however, different mediating tools might be in use (for example, lesson planning formats, time for meetings), different rules (lines of accountability), or divisions of labor (for example, responsibilities of mentors).

The extent to which these different expectations and understandings inhibit or advance dialog, cooperation and collaboration will depend on the commitment of the partners to sharing their perceptions, and to using the opportunity of working together to help explore new possibilities in their practice.

“Third space”

An oft-cited concept in the literature on university-school partnerships is that of a “third space”. The idea of a third space draws from a hybridity theory that recognizes that individuals draw on multiple discourses to make sense of the world. Third spaces reject binaries such as practitioner and academic knowledge or theory and practice and involve an integration of what are often seen as competing discourses. In this manner, the either/or perspective is transformed into a both/also point of view ([Zeichner, 2010](#)). Practitioner knowledge and academic knowledge, which are often seen as separate domains, are brought together to create new and richer learning opportunities for prospective teachers.

The notion of a “third space” is also an epistemological one; it breaks down assumptions of a hierarchy between the academic knowledge of the university and the professional knowledge of the school. Authority is seen to reside in both forms of knowledge; through an integration of these different sources of knowledge, students gain access to expanded learning opportunities.

Boundary crossing

Working across the activity systems of the school and the university is likely to require navigation across different patterns of interaction, cultures, structures, and expectations. This navigation, or “boundary-crossing”, refers to the engagement that occurs as members of the different communities (or activity systems) come together to consider matters of mutual concern.

Boundary objects are the tools or artifacts that mediate boundary crossing; an example here is lesson study, where co-planning and reflection by different participants encourages the sharing of perspectives on conceptions of teaching and learning ([Law and Tsui, 2009](#)). Processes embedded in boundary crossing include identification of the diverse practices, coordination of cooperative exchanges, reflection and expansion of perspectives, and transformation through collaboration and co-development of new practices ([Akkerman and Bakker, 2011](#)).

Expansive learning

Boundary crossing may bring to the fore competing discourse and assumptions, thereby potentially giving rise to tensions in the relationship. However, the engagement also has the potential to transform the learning experience for all parties in positive and mutually beneficial ways and to open opportunities for new learnings; crossing such boundaries “forces participants to take a fresh look at their long-standing practices and assumptions, and can be a source of deep learning” (Tsui and Law, 2007: 1290).

From this perspective, even dissonance becomes a potential source of change and innovation; “it is in the process of resolving contradictions that activity systems are transformed and expansive learning takes place” (Edwards et al., 2009: 17). Sorensen (2014) provides an example here, where he outlines how discussions between student teachers and their mentor teachers around the positioning of desks led to deeper considerations of classroom strategies, as well as of policies and practices to support teaching and learning within the school.

Horizontal expertise

Engagement and negotiation around boundary objects in a partnership requires that teachers and teacher educators bring their different forms of expertise together. As the resources of the different settings combine into new hybrid solutions, the work of the teachers and teacher educators is enriched and expanded:

Co-creating boundary objects—rather than using existing ones—facilitates the mutual engagement, negotiations, and hybridizations that characterize and generate horizontal expertise. It is this jointly created expertise—produced through negotiating language, practices, and tools—that facilitates cross-institutional communication and collaboration

Anagnostopoulos et al. (2007: 139–140).

Situations where such horizontal expertise would come into play include, for example, the co-creation of assessment rubrics between teachers and teacher educators. The negotiation of the social language of practice, as well as the articulation of competing views, creates opportunities to enrich conversation around the purpose and form of such rubrics (Anagnostopoulos et al., 2007).

Common knowledge and relational expertise

Horizontal linkages and the distribution of expertise is enhanced when different players in a partnership appreciate what is important for the other and why, and can work productively across potential differences. Edwards (2012) identifies a set of concepts to explain how interpretations of problems and alignments of practices are mediated across professional boundaries. The key concept here is that of common knowledge: “the motives that take forward each contributing practice ... [and that] mediate interactions across practice boundaries and give strategic direction to activities in and across services which are in the process of integration” (Edwards, 2012: 22).

The building of common knowledge requires relational expertise—the ability to recognize what the other partner values, to be able to take their standpoint, and to align one’s motives in order to pursue the engagement in mutually productive ways (Edwards, 2012). In this sense, relational expertise goes beyond understanding the work of another, or simply sharing one’s specialist expertise (as might happen in a discussion about pedagogical practices between teachers and teacher educators); the emphasis rather is on the ability to work in cooperative ways toward the integration of knowledge, “recognizing what others can offer a shared enterprise and why they offer it; and being able to work with what others offer while also making visible and accessible what matters for you” (Edwards, 2012: 26).

The establishment of the Oxford Deanery (a university-school partnership in the United Kingdom) provides an example of seeking common knowledge through relational expertise. Participants in initial teacher education, continuous professional development, as well as research, were invited to jointly construct the Deanery, so that it would reflect what mattered to all involved (Fancourt et al., 2015). In this way, student teachers were not seen as “boundary crossers”, but rather as participants in a shared and mutually constructed endeavor.

Sociocultural perspectives on professional learning

A further conceptual perspective on university-school partnerships focuses more deliberately on the learning process within the partnership. Drawing on a sociocultural perspective on learning, Tsui et al. (2009) argue that teacher learning is a function of both individual cognitive processes as well as broader social processes of which the student teachers and teachers are a part. They draw on the theoretical concepts of communities of practice, legitimate peripheral participation, and identity, all of which refer to the ways in which meanings are negotiated and identities shaped as student teachers participate in the life of a school:

How participants position themselves in the social space that is created as community boundaries are crossed, how they make sense of new forms of activities that emerge from the interactions, how they appropriate the sociocultural tools that are brokered in boundary crossing and how they negotiate the meanings of the joint enterprise are critical processes which constitute the learning experience and shape their identities as teachers

Tsui et al. (2009: 42).

Examples of such processes include being invited to participate in the full repertoire of teaching (Edwards and Tsui, 2009), a three way discussion after lesson observation of a student teachers' lesson by the student teacher, mentor teacher and university lecturer (Lopez-Real et al., 2009) and collaborative inquiry of classroom teaching through lesson study (Law and Tsui, 2009).

Teacher education pedagogy

The discussion thus far has focused on the practicum, however coherence and integration across course work and clinical work are central to quality teacher education (Darling-Hammond, 2006). This means that the structure and purpose of university-school partnerships in the practicum should by rights have implications for pedagogy within the entire teacher education program. Some of these implications are outlined here.

The first implication is that program design will—to a greater or lesser extent—include networks that bring “teacher educators, disciplinary faculty, school-based leaders and teachers, teachers’ union representatives, community and/or municipality representatives, and student teachers together to exchange ideas, concerns, values, and visions” in program design, evaluation and decision-making (Cochran-Smith et al., 2020: 114).

Certain approaches to teaching and learning within the partnership are likely to be prioritized by different models of teacher education. While various models exist in the literature, we use those of Cooper and Grundroff (2017) to illustrate the point. The first model, the effective teacher, emphasizes skills, content, performativity and measurement. The model of the reflective teacher also pays attention to skills and content, but includes knowledge about learners, and consideration of the underlying values and purposes of education. The inquiring teacher emphasizes systematic enquiry, deploying research and evaluation methods and techniques, while the transformative teacher adopts a critical enquiry approach, looking beyond the classroom to consider social context, moral and ethical issues (Cooper and Grundroff, 2017). Across these approaches, different pedagogical strategies will be employed in the partnership, ranging from teachers providing models of classroom practice to shared dialogs around dilemmas of practice, to critical reflection on ethics and values, to active engagement of student teachers in the community surrounding the school.

Some examples of pedagogy-in-action are described here to further illustrate learning across the partnership. Kameniar et al. (2017) describe a task where teacher candidates participate in an oral examination that assesses their capacity to use the theoretical and research discourses of the profession when speaking about practice. Assessment here focuses on the teacher candidates’ reasoning and pedagogical choices, and their capacity for informed reflection and action.

Another well-known strategy is that of the tripartite conference, where a student teacher, mentor teacher, and university teacher conduct a three-way conversation after the presentation of a lesson by the student teacher (Lopez-Real et al., 2009). Open and genuine conversations are encouraged, with the emphasis both on supporting and critically evaluating classroom practice.

The practice turn (Reid, 2011) in teacher education “suggests that views of teacher learning that emphasize cognition at the expense of the body may ultimately be inadequate—as the tacit knowledge of teachers actually resides in the body, in the situation of practice, and in the practice of teaching itself” (Mitchell and Reid, 2017: 46). Within this approach to teacher education pedagogy, the practices of teaching are explicitly studied, and observation, reflection and enactment are brought together, so that novice teachers come to understand how teachers work in embodied ways, at the same time as they make connections to teacher knowledge and reflexivity (Mitchell and Reid, 2017).

Reid describes a process whereby “core practices of teaching are ‘deconstructed’ and modeled for students, then rehearsed and approximated by students for their peers, under the gaze of an expert professional mentor giving feedback and coaching” (2011: 306). She describes the potential of this pedagogical strategy for shared dialog and critical reflection by the various partners:

In this case too, the expert teacher mentors with whom we are working are also not in their own habituated places, so that the practice being studied is similarly estranged and becomes an object of study for them as coaches. They can reflect, theoretically and phronetically, on why one way of practice might be more desirable, more effective, and more professional, in a particular situation

Reid (2011: 307).

Darling-Hammond et al. (2017) support such an inquiry-based approach in their research that studied high-performing countries in teacher education. What they found was that in countries where the focus was on supporting students to become teacher-researchers who actively and regularly engaged in teacher research, action research and other forms of practice related inquiry during the course of their teacher education programs, the quality of teaching overall improved. Such programs would necessarily need to be situated within well-structured university-school partnerships to support the teacher-research component across both fields.

Key considerations for universities and schools

University-school programs that exist in local contexts to support reform and innovation are often sustained through individual and institutional initiative, goodwill and commitment. In contrast, “country-level, system-wide examples of teacher education supported by genuine university-school partnerships with a focus on inquiry-rich school experiences are much harder to find”

(Cochran-Smith et al., 2020: 92). By inquiry-rich, the authors of the report refer to student teacher practicum experiences that provide opportunities for students' ongoing reflection, the opportunity to raise questions and make connections between theory (university learning) and practice (school-based learning), as well as having the opportunity to collect and interpret practice-based data that informs their thinking on classroom teaching, school policy and practice.

Several structural, cultural, systemic and relational considerations influence the longer-term sustainability of partnerships. These "mediating preconditions of practice [include] cultures and discourses, social and political structures and dynamics, and material and economic conditions" (Kemmis, 2009: 33). This section outlines in more detail some of the more significant conditions as outlined in the research, bearing in mind that these may vary significantly across different countries.

Oversight and responsibility

Part of the collaborative efforts among all stakeholders involves oversight and responsibility for the partnership, in other words, who is responsible for the different aspects of how the partnership is established and managed. This can be complex, given that universities and schools are usually governed by different national or local departments, each of which has its own policies, procedures and practices. In addition, different and sometimes conflicting perspectives among stakeholders about what is important for student teachers' learning and development, may exist (Cochran-Smith et al., 2020).

A further challenge relates to the fact that most often the primary responsibility for student teacher learning is located in the university with the schools serving as the venue where students complete a requisite number of school-based days as mandated by the national policy that guides teacher education in the country. How this arrangement has been established and is enacted means that in most cases the university has primary control over how the student's field experience is organized and what is required of the student during their practicum. This places the primary oversight within the university and schools may then defer to the university rather than engaging as equal collaborative partners. It is, therefore, important that all stakeholders' roles and responsibilities are clearly articulated to ensure that each role player knows who is accountable for what, to whom, and for what purpose in order for all role players to be empowered and actively involved in the knowledge building and learning of student teachers.

Shared vision

A central aspect of university-school partnerships involves the developing of a shared vision regarding the principles that guide the purpose of the partnerships (Green et al., 2020). The establishment of a shared vision between the school and university is, however, only the starting point. The implementation of a shared vision is a complex and long-term process that should inform the teacher education program design, curriculum and pedagogy and guide what and how student teachers learn in practice within the partnership agreement (Cochran-Smith et al., 2020).

In addition, a shared understanding of the developmental progression of student teachers' learning, as well as what constitutes effective teaching is important (Jones et al., 2016). This refers to what student teachers should be learning in practice during their school experience in relation to their university studies in the teacher education program. Developing a sequence of learning that connects foundational and theoretical knowledge from the university course-work with students' school-based practicum will assist to close the theory/practice gap often experienced by student teachers and holds the potential to reject the assumption that theory and practice operate in two separate worlds (Cochran-Smith et al., 2020).

Trust

A key aspect of a successful partnership is that of trust between the various role players. This includes a sense of mutuality and reciprocity (Kruger et al., 2009) where trust is constructed through shared understanding, and where there are benefits gained for each partner. Mutuality refers to each partner understanding the conjoint outcomes that can be achieved by collaboration while reciprocity is achieved when each stakeholder recognizes and acknowledges the value and expertise that each partner brings to the partnership (Kruger et al., 2009). Thus, trust moves beyond what can be mandated by policy and involves a willingness of all role players to bring commitment and expertise to the partnership to the benefit of all involved.

Time

Issues of time for all role players can be a challenging aspect in university-school partnerships. This involves time for university and school-based educators to engage in regular discussions regarding the planning and design of student teachers' learning experiences during their practicum. Green et al. (2020) note that when working with multiple parties in partnership agreements a deeper level of engagement and discussion is involved which usually takes more time and effort than the basic conversations or correspondence that might be the norm between universities and school-based teacher mentors. To build truly collaborative university-school partnerships co-operative dialoging, and an opportunity for differences in agendas and perspectives to be shared and acknowledged, needs to take place (Cooper and Grudnoff, 2017).

Sustainability

For partnerships to be sustainable, shared governance with dedicated resources and funding that foster sustainability within the partnerships is necessary (Burns et al., 2016). By shared governance, Burns et al. (2016) refer to aspects of decision-making that develop shared routines, rituals and celebrations. Sharing governance of partnerships assists to develop genuine commitment between schools and universities which is needed to establish and build long-term relationships and partnerships that are built on mutual professional respect and trust between all participants (Cooper and Gudnoff, 2017).

Partnerships will struggle to be sustainable if they are developed only through the “determined efforts by inspired individuals” (Kruger et al., 2009: 50). Ongoing and enduring institutional support is essential for the sustainability of partnerships both at university and school level. Green et al. (2020: 418), discussing the sustainability and expansion of partnerships, note that issues of systems within institutions, changing personnel, insecurity of funding, as well as the scalability of partnerships, are all aspects which need to be taken into account.

Equal partners

Within partnership agreements, schools and school personnel often do not perceive themselves to be equal partners in the teacher education process. This may lead to schools choosing not to prioritize resources to develop or extend partnerships with universities (Cochran-Smith et al., 2020). In addition, due to the sometimes perceived uneven status of the role players in the partnerships (student teachers, university staff members and the school-based teacher mentor), tensions may develop which could hamper a genuine sharing of ideas and cooperation across universities and schools (Cohen et al., 2013). Thus, truly equal partnerships between schools and universities can be difficult to achieve as teacher knowledge and expertise is at times devalued or marginalized and university ideas and theories are often given precedence by teacher educators.

To counter power relationships Payne and Zeichner (2017) argue for “collaborative egalitarianism” that invites universities and schools to work together on student teacher practicum as an authentic problem to which both university and school personnel share expertise to solve the issues that arise. Similarly, Burns et al. (2016) argue that all stakeholders should be involved in developing university-school agreements. By all stakeholders, as per Payne and Zeichner’s concept of “collaborative egalitarianism”, Burns et al. (2016) suggest that university-school partnerships should include engagement between the school and local community where the community is conceptualized more broadly than educational practitioners within the school walls. In other words, it involves connecting and building bridges with local communities, homes, parents and so forth. Simply developing partnership agreements between universities and schools should not be seen as true collaboration unless all parties involved are given equal “voice” in how the student practicum process is established (the vision and aims of the process), unfolds, and responds to challenges that arise. When teaching practicums are developed as “collaborative egalitarianism” university academic knowledge, school-based practitioner knowledge and community knowledge are given equal status and a non-hierarchical interplay between academic, practitioner and community expertise is established (Payne and Zeichner, 2017).

Teacher mentors

While university-school partnerships involve multiple role players, the role that the teacher mentor plays has a significant impact on student teachers’ school-based experiences. Being an effective teacher mentor is not the same as being a good teacher, nor does it simply mean making classroom teaching opportunities available to students. Mentoring student teachers during their practicum involves a form of professional practice that requires the teacher mentor to draw on their knowledge of the practice of teaching to create learning opportunities with focused feedback that supports and extends a student teachers’ practice (Feiman-Nemser, 2001; Feiman-Nemser and Beasley, 2007). As such, mentoring student teachers in their school-based teaching experience has its own skills, knowledge and practices and ideally teacher mentors should be given formal systematic training which includes mentor teachers engaging in their own critical reflection on the mentoring process (Aspfors and Fransson, 2015). Thus, as suggested by Cooper and Gudnoff (2017) there should be a more rigorous process of identification, selection, training and resourcing of teacher mentors.

Acknowledgment and recognition of teacher mentors

Internationally, an ongoing concern is that teacher mentors are expected to mentor student teachers in addition to their normal teaching responsibilities often without a clear description of what their role as a teacher mentor entails or an understanding of the process involved (Cohen et al., 2013). Mentoring students requires additional preparation and time out of a teacher’s already busy day and there are rarely any formal benefits for a teacher agreeing to mentor student teachers. Within formal university-school partnership agreements, teacher mentors might also be expected to meet regularly with university staff to discuss the teaching practicum process, adding extra time, preparation and responsibility to a teacher’s already full load (Cohen et al., 2013).

A similar concern exists for university-based teacher educators with respect to the acknowledgment and recognition of the work of establishing and maintaining partnerships. Few incentives exist for faculty members to engage in the work of coordinating, mentoring and monitoring field-based placements and this work is often carried out by adjunct staff who may not be in touch with the core principles of the teacher education program (Zeichner, 2010). Communication, developing relationships and the building of

trust—all key aspects of partnerships—are hardly valued in a competitive higher education context, where academics are rewarded for generating funding and publishing in high-status journals (Ellis and McNicholl, 2015). This may be exacerbated where a separation exists between so-called academic and professional workers, either within the division of labor of the university (Gunn et al., 2020) or within the self-identities of the academics themselves (Robinson and McMillan, 2006).

Professional inquiry

Policies and mandates, while important for sustainability, are in themselves no guarantee of successful implementation of partnership agreements. Even where enabling structural arrangements like time, resources and memoranda of understanding are in place, meaningful opportunities for a theory-practice dialog will depend on the active and willing engagement and agency of the participants in such encounters. For both universities and schools, a move within a partnership to integrate theoretical knowledge, situational awareness and practical know-how (Winch et al., 2015) places responsibilities on the inquiry orientation of both parties. For universities, this means embracing an epistemology that recognizes the expertise of schools and communities as co-existing on a more equal plane with academic knowledge (Zeichner, 2010). For schools, it becomes important that “inquiry as stance” (Cochran-Smith and Lytle, 2009: 44) is part of the ethos and daily functioning of the school so that teachers are encouraged to share concerns, deliberate on dilemmas, and explore innovative practices among their colleagues, as well as with the university. Where such inquiry cultures exist, the possibilities for mutually beneficial dialogs becomes so much stronger.

Several factors influence the capacity for schools to indeed establish a culture of inquiry. At a constraining level, these factors could include heavy workloads, demanding schedules, too much administrative work, little time for collaborative engagement, poor leadership, innovation overload, or a lack of trust and collegiality (Robinson, 2009). Furthermore, Hogan (2015) has pointed out that educational reform measures that homogenize and standardize teaching, and which encourage teaching-to-the-test pose a considerable risk to pedagogical inquiry on the part of teachers.

However, as suggested by Lund and Erikson (2016), if universities and schools work toward a more co-configuring and reconfiguring of partnership agreements teacher mentors may be capacitated to see their role as not merely recipients of student teachers, but as being part of the teacher education program and practitioner inquiry, recognizing their role as teacher educators as equal but different to the university. This focus moves from a systemic structure that requires teachers to mentor student teachers during their practicum to one that emphasizes the development of ongoing school-based learning and teacher professionalism, with a focus on learning for all stakeholders (see Kruger et al., 2009). This may assist teachers to view practitioner inquiry not as an additional workload but as forming part of the co-construction and integration of knowledge and research competency that creates inquiry-rich school-based experiences and a more professionally-orientated teaching force (Cochran-Smith et al., 2020).

Concerns and caveats

University-school partnerships have been presented as a powerful way to enhance teacher, student teacher and university faculty learning. However, the partnership needs to be envisaged as more than student teachers spending time in schools; in fact, various researchers have cautioned against the assumption that spending more time in schools will lead to student teachers becoming better teachers.

Ronfeldt and Reininger (2012), for example, investigated whether lengthening student teachers’ practicum improves their preparedness, or perception of preparedness, efficacy and outcomes. The study findings suggest that the duration of the practicum program has little effect on teacher outcomes, however, what does impact on student teacher outcomes is the quality of teaching experienced by the student during the practicum experience.

Zeichner (2014) points out how shifting the location of teacher education more to schools can indeed strengthen the educative value of the preparation but can also uncritically glorify learning from experience. He argues the importance of “unpacking and analyzing the underlying commitments to particular purposes of public schooling and the meaning of teaching, learning, the teacher’s role and the process of learning to teach (2014: 264).”

Working in the United Kingdom, with its well-established history of school-based teacher education, Ellis (2010) critiques what he calls an impoverished notion of learning often embedded in arguments for more school-based work in teacher education. Experience in schools, he argues, has the potential to lead to the reproduction of routines and the development of “bureaucratic virtues” like compliance. He highlights some tensions identified within the Oxford Internship Scheme, a program with a strong base of schools working with the university. These include, for example, contradictory conceptions of knowledge and learning, and issues of relative power between university academics and teacher mentors. Ellis (2010) argues for university-school partnerships to open up potential barriers to dialog so that novice teachers are encouraged to engage with the social systems in which they find themselves, so that teacher education becomes a way of “using experience as a means to develop an idealized vision of the future” (Ellis, 2010:117).

Concluding comments

This article has outlined some of the issues identified in research on partnerships. These issues are both conceptual and practical, and include how the partners envisage the purpose of the partnership, as well as how different institutions work together. Not all

partnership models developed can be generalized. In other words, some aspects of partnerships models are site-specific due to particular institutional arrangements. However, aspects of how a partnership model has been established and operates successfully can be helpful as a sort of “template” for other partnership models to be developed in their own settings (Fancourt et al., 2015: 369).

What is clear in the international literature is that “partnerships between universities and schools are central to the success of almost all education reforms, even though establishing true partnerships is very difficult work” (Cochran-Smith et al., 2020: 90). However, at the very least, it is clear that university-school partnerships need to be founded on a sense of mutual benefit, with key elements of trust, vision, conceptions of teaching and learning and enabling structural conditions being central. Partnerships take time, effort, resources and ongoing collaboration by all involved. Unless all role players agree on the focus and intention of the partnership, partnerships may struggle to be sustained. Partnerships also need to move beyond individuals within the partnerships. Partnership agreements that are in place at a more systemic level rather than through the efforts and interests of individuals have a better chance of being sustained over time. In addition, developing the capabilities and incentives of those involved in university-school partnerships may assist to support and develop deeper levels of engagement and learning by all stakeholders over time.

References

- Akkerman, S., Bakker, A., 2011. Boundary crossing and boundary objects. *Rev. Educ. Res.* 81 (2), 132–169. <https://doi.org/10.3102/0034654311404435>.
- Anagnostopoulos, D., Smith, E., Basmadjian, K., 2007. Bridging the university-school divide: horizontal expertise and the “two worlds pitfall”. *J. Teach. Educ.* 58 (2), 138–152. <https://doi.org/10.1177/0022487106297841>.
- Aspfors, J., Fransson, G., 2015. Research on mentor education for mentors of newly qualified teachers: a qualitative meta-synthesis. *Teach. Teach. Educ.* 48, 75–86. <https://doi.org/10.1016/j.tate.2015.02.004>.
- Boud, D., 2012. Problematising practice-based education. In: Higgs, J., Barnett, R., Billet, S., Hutchings, M., Trede, F. (Eds.), *Practice-Based Education: Perspective and Strategies*. Sense Publishers, Rotterdam, pp. 55–68.
- Burn, K., Mutton, T., 2015. A review of “research-informed clinical practice” in initial teacher education. *Oxf. Rev. Educ.* 41 (2), 217–233. <https://doi.org/10.1080/03054985.2015.1020104>.
- Burns, R., Jacobs, J., Baker, W., Donahue, D., 2016. Making muffins: identifying core ingredients of school-university partnerships. *Sch. Univ. Partnersh.* 9 (3), 81–95.
- Cochran-Smith, M., Lytle, S., 2009. Teacher research as stance. In: Noffke, S., Somekh, B. (Eds.), *The SAGE Handbook of Research on Teacher Education*. SAGE Publishing, London, pp. 39–49.
- Cochran-Smith, M., Alexandersson, M., Ellis, V., Grudnoff, L., Hammerness, K., Oancea, A., Toom, A., 2020. *Transforming Norwegian Teacher Education: The Final Report of the International Advisory Panel for Primary and Lower Secondary Teacher Education*. Norwegian Agency for Quality Assurance in Education, Norway.
- Cohen, E., Hoz, R., Kaplan, H., 2013. The practicum in preservice teacher education: a review of empirical studies. *Teach. Educ.* 24 (4), 345–380. <https://doi.org/10.1080/10476210.2012.711815>.
- Cooper, B., Grudnoff, L., 2017. Redesigning authentic collaborative practicum partnerships: learnings from case studies from two New Zealand universities. In: Peters, M., Cowie, C., Menter, I. (Eds.), *A Companion to Research in Teacher Education*. Springer, Singapore.
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A.L., Hammerness, K., Low, E., McIntyre, A., Rothman, R., Sato, M., Zeichner, K., 2017. *Empowered Educators: How High-Performing Systems Shape Teaching Quality Around the World*. John-Wiley & Sons, Hoboken.
- Darling-Hammond, L., 2006. Constructing 21st-century teacher education. *J. Teach. Educ.* 57 (3), 300–314. <https://doi.org/10.1177/0022487105285962>.
- Darling-Hammond, L., 2010. Teacher education and the American future. *J. Teach. Educ.* 61 (1–2), 35–47. <https://doi.org/10.1177/0022487109348024>.
- Darling-Hammond, L., 2014. Strengthening clinical preparation: the holy grail of teacher education. *Peabody J. Educ.* 89 (4), 547–561. <https://doi.org/10.1080/0161956X.2014.939009>.
- Edwards, G., Tsui, A., 2009. On becoming a member of a community of practice. In: Tsui, A., Edwards, G., Lopez-Real, F. (Eds.), *Learning in School-University Partnership: Sociocultural Perspectives*. Routledge, New York, pp. 47–68.
- Edwards, G., Tsui, A., Stimpson, P., 2009. Contexts for learning in school-university partnership. In: Tsui, A., Edwards, G., Lopez-Real, F. (Eds.), *Learning in School-University Partnership: Sociocultural Perspectives*. Routledge, New York, pp. 3–24.
- Edwards, A., 2012. The role of common knowledge in achieving collaboration across practices. *Learn. Cult. Soc. Interact.* 1 (1), 22–32. <https://doi.org/10.1016/j.lcsi.2012.03.003>.
- Ellis, V., McNicholl, J., 2015. *Transforming Teacher Education: Reconfiguring the Academic Work*. Bloomsbury, London.
- Ellis, V., Edwards, A., Smagorinsky, P. (Eds.), 2010. *Cultural-Historical Perspectives on Teacher Education and Development: Learning Teaching*. Routledge, London and New York.
- Ellis, V., 2010. Impoverishing experience: the problem of teacher education in England. *J. Educ. Teach.* 36 (1), 105–120.
- Fancourt, N., Edwards, A., Menter, I., 2015. Reimagining a school-university partnership: the development of the Oxford Education Deanery narrative. *Educ. Inq.* 6 (3), 353–373. <https://doi.org/10.3402/edui.v6.27724>.
- Feiman-Nemser, S.F., Beasley, K., 2007. Discovering and sharing knowledge: inventing a new role for cooperating teachers. In: Carroll, D.E., Featherstone, H.E., Featherstone, J.E. (Eds.), *Transforming Teacher Education: Reflections From the Field*. Harvard Education Press, Cambridge, MA, pp. 139–160.
- Feiman-Nemser, S.F., 2001. From preparation to practice: designing a continuum to strengthen and sustain teaching. *Teach. Coll. Rec.* 103 (6), 1013–1055.
- Furlong, J., Whitty, G., Whiting, C., Miles, S., Baron, L., Barrett, E., 1996. Redefining partnership: revolution or reform in initial teacher education? *J. Educ. Teach.* 22 (1), 39–55. <https://doi.org/10.1080/02607479650038418>.
- Gravett, S., Petersen, N., Ramsarop, S., 2019. A university and school working in partnership to develop professional practice knowledge for teaching. *Front. Educ.* 3 <https://doi.org/10.3389/educ.2018.00118>.
- Green, C.A., Tindall-Ford, S.K., Eady, M.J., 2020. School-university partnerships in Australia: a systematic literature review. *Asia Pac. J. Teach. Educ.* 48 (4), 403–435. <https://doi.org/10.1080/1359866X.2019.1651822>.
- Gunn, A., Hill, M., Berg, D., Haigh, M., 2020. *The Promise and Practice of University Teacher Education: Insights From Aotearoa New Zealand*. Bloomsbury, New York.
- Henning, E., Petker, G., Petersen, N., 2015. University-affiliated schools as sites for research learning in pre-service teacher education. *S. Afr. J. Educ.* 35 (1), 1–8.
- Hogan, P., 2015. Cultivating human capabilities in venturesome learning environments. In: Heilbronn, R., Foreman-Peck, L. (Eds.), *Philosophical Perspectives on Teacher Education*. Wiley Blackwell, Chichester, pp. 132–151.
- Jones, M., Hobbs, L., Kenny, J., Campbell, C., Chittleborough, G., Gilbert, A., Herbert, S., Redman, C., 2016. Successful university-school partnerships: an interpretive framework to inform partnership practice. *Teach. Teach. Educ.* 60, 108–120. <https://doi.org/10.1016/j.tate.2016.08.006>.
- Kameniar, B., Davies, L.M., Kinsman, J., Reid, C., Tyler, D., Acquaro, D., 2017. Clinical praxis exams: linking academic study with professional practice knowledge. In: Peters, M., Cowie, B., Menter, I. (Eds.), *A Companion to Research in Teacher Education*. Springer, Singapore. https://doi.org/10.1007/978-981-10-4075-7_4.

- Kemmis, S., 2009. Understanding professional practice: a synoptic framework. In: Green, B. (Ed.), *Understanding and Researching Professional Practice*. Sense Publishers, Rotterdam, pp. 19–38.
- Klein, E.J., Taylor, M., Onore, C., Strom, K., Abrams, L., 2016. Exploring inquiry in the third space: case studies of a year in an urban teacher-residency program. *N. Educat.* 12 (3), 243–268. <https://doi.org/10.1080/1547688X.2016.1187980>.
- Korthagen, F., Loughran, J., Russell, T., 2006. Developing fundamental principles for teacher education programs and practices. *Teach. Teach. Educ.* 22 (8), 1020–1041. <https://doi.org/10.1016/j.tate.2006.04.022>.
- Kruger, T., Davies, A., Eckersley, B., Newell, F., Cherednichenko, B., 2009. *Effective and Sustainable University–School Partnerships—Beyond Determined Efforts by Inspired Individuals*. Teaching Australia, Canberra.
- Law, D., Tsui, A., 2009. Mutual engagement and boundary crossing in lesson studies. In: Tsui, A.B.M., Edwards, G., Lopez-Real, F. (Eds.), *Learning in School–University Partnership: Sociocultural Perspectives*. Routledge, New York, pp. 111–131.
- Lopez-Real, F., Law, D., Tang, R., 2009. Tripartite conferences: positioning and activity systems. In: Tsui, A., Edwards, G., Lopez-Real, F. (Eds.), *Learning in School–University Partnership: Sociocultural Perspectives*. Routledge, New York, pp. 90–110.
- Lund, A., Erikson, T.M., 2016. Teacher education as transformation: some lessons learned from a centre for excellence in education. *Acta Didact. Norge* 10 (2), 53–72. <https://doi.org/10.5617/adno.2483>.
- Miller, J., Haynes, J., Pennington, J., 2015. A partnership aimed at improving health and physical education at a rural school: impacts on pupils, university students, teachers and academics. *Aust. Int. J. Rural Educ.* 25 (2), 56–72.
- Mitchell, D.M., Reid, J., 2017. (Re)turning to practice in teacher education: embodied knowledge in learning to teach. *Teach. Teach.* 23 (1), 42–58. <https://doi.org/10.1080/13540602.2016.1203775>.
- Orrell, J., Higgs, J., 2012. Social and political change: implications for professional and practice-based university education. In: Higgs, J., Barnett, R., Billet, S., Hutchings, M., Trede, F. (Eds.), *Practice-Based Education: Perspective and Strategies*. Sense Publishers, Rotterdam, pp. 41–54.
- Payne, K., Zeichner, K., 2017. Multiple voices and participants in teacher education. In: Clandinin, D.J., Hsu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*. SAGE Publications, London, pp. 1101–1116.
- Reid, J., 2011. A practice turn for teacher education? *Asia Pac. J. Teach. Educ.* 39 (4), 293–310. <https://doi.org/10.1080/1359866X.2011.614688>.
- Robinson, M., McMillan, W., 2006. Who teaches the teachers? Identity, discourse and policy in teacher education. *Teach. Teach. Educ.* 22, 327–336. <https://doi.org/10.1016/j.tate.2005.11.003>.
- Robinson, M., 2009. Practitioner inquiry in South African schools: what, how and why (not). *Educ. Action Res.* 17 (1), 121–131. <https://doi.org/10.1080/09650790802667519>.
- Ronfeldt, M., Reininger, M., 2012. More or better student teaching? *Teach. Teach. Educ.* 28, 1091–1106. <https://doi.org/10.1016/j.tate.2012.06.003>.
- Silbert, P., 2019. From reciprocity to collective empowerment: re-framing university–school partnership discourses in the South African context. *Gatew. Int. J. Community Res. Engagem.* 12 (1), 1–16. <https://doi.org/10.5130/ijcre.v12i1.6620>.
- Sorensen, P., 2014. Collaboration, dialogue and expansive learning: the use of paired and multiple placements in the school practicum. *Teach. Teach. Educ.* 44, 128–137. <https://doi.org/10.1016/j.tate.2014.08.010>.
- Toom, A., 2017. Teachers' professional and pedagogical competencies: a complex divide between teacher work, teacher knowledge and teacher education. In: Clandinin, D.J., Hsu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*. SAGE Publications, London, pp. 803–819.
- Toon, D., Jensen, B., 2017. *Teaching Our Teachers: A Better Way—Developing Partnerships to Improve Teacher Preparation*. Learning First. Available at: <http://www.learningfirst.com/wp-content/uploads/2018/03/2columnsITECoPPaper2PartnershipsFINAL17Nov17.pdf>.
- Tsui, A.B.M., Law, D.Y.K., 2007. Learning as boundary-crossing in school–university partnership. *Teach. Teach. Educ.* 23, 1289–1301. <https://doi.org/10.1016/j.tate.2006.06.003>.
- Tsui, A.B.M., Lopez-Real, F., Edwards, G., 2009. Sociocultural perspectives of learning. In: Tsui, A.B.M., Edwards, G., Lopez-Real, F. (Eds.), *Learning in School–University Partnership: Sociocultural Perspectives*. Routledge, New York, pp. 25–43.
- Winch, C., Oancea, A., Orchard, J., 2015. The contribution of educational research to teachers' professional learning: philosophical understandings. *Oxf. Rev. Educ.* 41 (2), 202–216. <https://doi.org/10.1080/03054985.2015.1017406>.
- Zeichner, K., 2010. Rethinking the connections between campus courses and field experiences in college- and university-based teacher education. *J. Teach. Educ.* 61 (1–2), 89–99. <https://doi.org/10.1177/0022487109347671>.
- Zeichner, K., 2014. *The politics of learning to teach from experience*. In: Ellis, V., Orchard, J. (Eds.), *Learning Teaching From Experience: Multiple Perspectives and International Contexts*. Bloomsbury Academic, London, New York.
- Zeichner, K., 2019. Preparing teachers as democratic professionals. *Action Teach. Educ.* 1–10. <https://doi.org/10.1080/01626620.2019.1700847>.

Inside looking out: a third space perspective on pre-service teachers' practice experiences

Viviana Daza, Department of Teacher Education and School Research, Faculty of Educational Sciences, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	393
Partnerships in teacher education	394
Partnerships in the third space	394
Communities of practice: a third space perspective	395
Inside looking out: PSTs' voices and experiences	396
Methodology	396
Sample and data collection	396
Data analysis	397
Findings	397
Descriptive overview	397
Voices and experiences	397
Epistemic negotiations	397
Relational factors	400
Identity development	401
Discussion	401
Conclusion	402
Acknowledgment	403
References	403

Introduction

Pre-service teachers' (PSTs) central position in practice provides an "inside looking out" perspective that expands the more theoretical "third space" concept, which is increasingly used to describe less hierarchical partnerships in Teacher Education (TE). The empirical case presented here highlights PSTs' views on their practice experiences to illuminate the inner dynamics of these partnerships. PSTs' voices are essential to achieving a pragmatic understanding of how the third space works and how their perspectives contribute to forming an "inside looking out" narrative of the dynamics of becoming a teacher.

In the first part, this study discusses the notion of partnerships in TE and the concept of the third space in TE research and practice. The study then examines the possible compatibility of the third space with the theory of Communities of Practice (Lave and Wenger, 1991; Wenger, 1998), showing how they complement and enrich each other when dealing with partnerships in TE. Later, we draw from an empirical study with PSTs to present an "inside looking out" perspective that helps us understand the opportunities and challenges related to experiencing practice in the third space. Thereby, this study aims to contribute to the knowledge of how the third space operates and what elements PSTs perceive to be relevant for their learning process. By using the concept of Communities of Practice as a framework, this study also aims to expand our understanding of how partnerships can further support professional development in TE.

TE combines core elements of theoretical and practical knowledge that prospective teachers need in order to develop their teaching in the field. It is therefore crucial to ensure a strong flow between these elements. Contemporary educational research highlights partnerships as an effective and sustainable way to strengthen the connection between schools and TE institutions (Cochran-Smith and Lytle, 1999; Darling-Hammond, 2006; Allen et al., 2013a; Bloomfield and Nguyen, 2015; Lillejord and Børte, 2016; Miles et al., 2019b). Partnerships in TE are institutionally mandated networks that develop relationships based on mutual support, shared responsibility, and ambitions (Bloomfield and Nguyen, 2015) on the one hand, and encourage less hierarchical relations between partners that acknowledge practice expertise and support hybrid professionals on the other (Zeichner, 2010; Zeichner and Payne, 2013; Allen et al., 2013b; Beaton and Hope, 2022). The latter aspect has received significant attention in educational research worldwide with the aim of fostering effective partnerships that move away from vertical forms to connect and balance the dominant narratives of the two separate spaces of schools and TE institutions.

One of the main objectives of partnerships is to enhance the quality of TE and support PSTs' learning and transitioning into the teaching profession. Although partnership collaboration intends to connect schools and TE institutions, these do not necessarily result in the desired outcomes (White et al., 2010; Walsh and Backe, 2013). Hence, the following section will present key aspects of partnerships in TE and explore successful collaboration stories and challenges associated with their development.

Partnerships in teacher education

As practice is a fundamental element for teachers' professional development, partnerships need to genuinely promote collaboration to improve the quality of TE and provide meaningful practice experiences for future teachers (Darling-Hammond, 2014; Zeichner, 2010, 2012; Barragan, 2017; Broadley et al., 2019; Chan, 2019). Learning to teach requires a sound integration of theoretical knowledge and coherent and authentic experiences from practice. In response to this, educational research endorses the establishment of alliances that foster this integration (Darling-Hammond, 2010, 2017; Le Cornu, 2010, 2015).

An essential characteristic of these partnerships is the implementation of a practice that facilitates the transition into the teaching profession within an integrated learning community that disrupts boundaries across institutions (Jackson and Burch, 2019; Lemon et al., 2018; Campbell-Perry, 2022). Other researchers have found that such partnerships can be successful when the partners share goals and merge knowledge sources. For instance, Vagi et al. (2019) and Allen and Wright (2014) demonstrate how PSTs benefit from and improve their teaching quality through prolonged student residencies and embedded coursework assessments, respectively. Such joint activities provide opportunities to adapt knowledge from school classrooms to campus and vice-versa through exploration, communication, and shared experiences. This research indicates that a solid collaboration in partnerships, as organizational tools (Lillejord and Børte, 2016), can provide PSTs with access to an authentic learning experience that facilitates their development, instills confidence in their knowledge, and improves their readiness to teach.

However, research has found that partnerships do not automatically increase collaboration or guarantee successful cooperation. The traditional academic hegemony of TE institutions maintains hierarchical relationships that hinder the creation of sustainable bridges between practices at schools and TE institutions (Snow-Geron, 2005; Barak et al., 2016; Zeichner et al., 2015).

Over the last decade, educational researchers have been increasingly interested in partnership models that support the third space. The third space in this context denotes a meeting point where equal participation and status are granted to schools and TE institutions. The third space for collaboration in partnerships favors PSTs' development, as it benefits the schools and TE institutions by improving teachers' practices, pupils' performance, teacher retention rates, and cooperation (Darling-Hammond, 2010; Zeichner, 2012; Allen et al., 2013b). As Barak et al. (2016, p. 38) argue, "For partnerships to flourish, the existing institutional borders have to become penetrable to enable interactions among the practice and discourse in the different institutions." Hence, the third space has become an alternative model of collaboration between schools and TE institutions that promises an epistemic change and a new way to enact practice.

Partnerships in the third space

In essence, the third space is used to examine the areas "in-between" and the hybridity of two or more discourses or binaries (Bhabha, 1994). This postcolonial concept has been adopted in various disciplines and with different purposes. American scholar Kenneth Zeichner introduced the third space concept to TE discourse in 2010. In the TE field, partnerships in the third space seek to interconnect complementary learning experiences from theory and practice by developing a common language about the ways of being and becoming a teacher, blurring boundaries and favoring merging ideas, discourses, and epistemologies between schools and TE institutions (Zeichner, 2010; Ramsaroop and Gravett, 2017; Beck, 2018). In particular, partnerships in the third space can potentially sustain less hierarchical relationships among partners. Such characteristics benefit PSTs' learning processes and professional development, and disrupt the way we think about practice (Elsden-Clifton and Jordan, 2016, 2019; Klein et al., 2013, 2016; Cuenca et al., 2011; Daza et al., 2021; Buckreis, 2012).

The third space can help PSTs see connections, grasp contradictions, and form a third understanding of the teaching activity, intersecting theoretical and practical epistemologies (Zeichner, 2010). It thus assigns the responsibility of preparing teachers to both schools and TE institutions, referred to here as School-Based Teacher Educators (SBTEs) and University-Based Teacher Educators (UBTEs). SBTEs and UBTEs are hybrid teachers who are tasked with the mediating role of merging practice and theory knowledge and enacting brokering work across them. From a third space perspective, the learning activities in TE that are developed on campus and in school classrooms (i.e., coursework, assessment, instruction, courses taught on-site in schools, practice, and so on) should seek to integrate epistemologies, cross boundaries, open up communication, navigate tensions, and aim to achieve greater coherence (Elsden-Clifton and Jordan, 2019; Daza et al., 2021).

Daza et al. (2021) found that the negotiation of identities through boundary-crossing and hybridity work are essential features of partnerships in the third space. Hence, boundary brokers, hybrid teacher educators, and hybrid PSTs are essential members of partnerships in the third space where they enact their roles as teachers, students, and equal partners. Such partnerships also allow for new pedagogical possibilities to emerge and invigorate practice by co-creating activities that enrich each other in productive ways (Steele, 2017; Beck, 2018; Lee, 2018; Chan, 2020), furthering the idea of "working together rather than working with" (Jackson and Burch, 2016, p. 522).

However, the utopian nature of the third space brings about challenges for partnerships in TE as well, such as adapting to new roles and identities, managing power relations among partners, and sustaining the quality of collaboration (Martin et al., 2011; McDonough, 2014; Daza et al., 2021). Hence, a constant negotiation of tensions and process evaluation are crucial factors in developing and maintaining effective collaboration between schools and TE institutions. Despite these challenges, however, educational research continues to explore the promising third space concept in order to facilitate unique opportunities to rethink and reconfigure TE.

One way to understand the nature of partnerships in the third space is to acknowledge them as complex networks of ways of knowing and doing in TE. From this perspective, schools and TE institutions are two separate communities of practice that interweave their experiences and knowledge to achieve the common goal of an integrated practice for PSTs. Therefore, the concepts of Situated Learning and Communities of Practice (Lave and Wenger, 1991; Wenger, 1998) are used here as frameworks to help us understand the role of practice in communities and the gradual process of acquiring mastery and becoming active community members.

Communities of practice: a third space perspective

The concept of communities of practice developed by Lave and Wenger (1991) and Wenger (1998) is at the core of situated approaches to learning, and it has been adopted across numerous disciplines, such as management, social sciences, and educational sciences, among others (Barton and Tusting, 2005; Lea, 2005). Lave and Wenger view participation in a community as a crucial aspect of human interaction, where identities form through mutual engagement and shared practice in social learning systems that aim to develop members' skills and exchange knowledge. Thereby, they are "groups of people informally bound together by shared expertise and passion for a joint enterprise" (Wenger, 2000, p. 139).

We propose that the concept of communities of practice can be understood, within the context of third space partnerships in TE, as a blended social learning system that integrates practical and theoretical expertise from schools and TE institutions. What makes third space partnerships "blended" is the inherently dynamic nature of TE and the "back and forth" exchange of expertise across the two separate yet interlinked communities of schools and academia. Consequently, PSTs become part of this blended community of practice, wherein they conduct practice; share a common goal; have a shared repertoire of languages, norms, and artifacts; actively choose to participate; and mutually engage in building new knowledge.

Lave and Wenger (1991) suggest the concept of "legitimate peripheral participation" to describe the situatedness of learning and the first step that takes the apprentice from the edge of practice toward the center of the activity in determined communities of practice. This peripheral zone facilitates the integration of newcomers into the mainstream in an environment that is both more tolerant to inside and outside influences and less hostile to learning and adapting to the practices in the community.

The authors describe several communities of practice (i.e., midwives, tailors, alcoholics) where newcomers take steps from peripheral participation to gradually acquiring mastery. In this process, newcomers collaborate with more experienced peers—old-timers—and participation goes beyond mere observation. It also involves doing and participating as a way of learning by "both absorbing and being absorbed in the culture of practice" (Lave and Wenger, 1991, p. 95). We believe that the concept of communities of practice that comprise peripheries in the process of "becoming" is helpful for understanding third space partnerships in TE. However, we suggest a broader approach encompassing the constant evolution of the teaching profession and the dialectics of the two communities of practice working together in third space partnerships.

The dynamics of partnerships in TE are complex and frequently pose challenges due to its transdisciplinary nature, making up intricate contexts within which PSTs gradually integrate their professional development to become in-service teachers. Wenger and Lave's more unambiguous communities of practice thus do not sufficiently capture the dynamics of third space partnerships. The factors that characterize communities of practice from a third space perspective call for a broader approach to discern the complexities of a hybrid, egalitarian community that works together but which are communities of practice in their own right.

Several essential aspects of communities of practice have implications for third space partnerships in TE. First, when PSTs and teacher educators work together in practice, they negotiate meaning as they further develop their identity. Wenger (2010) proposes that learning in communities of practice transcends new skills and building understanding and involves active participation, negotiation, and the transformation of identities. It also involves "becoming a certain person—a knower in a context where what it means to know is negotiated" (p. 2). For instance, when PSTs commute within and across third space partnerships, they "become" hybrid learners and teachers. They begin to identify and negotiate their roles, increasingly "becoming" more knowledgeable teachers and members of the teaching community.

As the negotiation of meaning and participation are central and continuous processes involved in communities of practice, we will further expand on them. For Wenger (1998), "participation refers to a process of taking part and also to the relations with others that reflect this process. It suggests both action and connection" (p. 55). Both ideas of negotiation of meaning and participation are at the heart of Wenger's communities of practice. They are used here as a framework to understand partnerships in the third space:

- (1) Negotiation of meaning: partnerships in the third space constantly face tensions due to clashes of power, disagreements on how to approach or enact practice, and the need to adapt to new roles. Such challenges are part of the hybrid work that partners develop in practice and are meant to be managed, navigated, and constantly revised. The sustainability of partnerships depends on the negotiation of meanings and the partners' will to maintain meaningful learning experiences to improve collaboration with and respect toward different knowledge and expertise. Such negotiations occur through the interplay of lived experiences that take form and become organized while participating in the community of practice. Therefore, participation is a crucial ongoing process involving personal memberships and community involvement (Wenger, 1998).
- (2) Participation is the crucial element that weaves the community together through collaborative practice. To think about participation includes thinking of peoples' "doings" and how these fit into the doings of groups and communities (Keating, 2005). From this point of view, participation creates a learning environment where common understandings emerge from the

inside and the outside, shaping a hybrid form of collaboration that integrates theoretical and practical knowledge in third space partnerships. Hence, expertise is not uniquely limited to the teacher educators as old-timers but is possessed by all members involved in the community of practice. Less hierarchical participation in third space partnerships presupposes a more hybrid and collaborative peripheral zone where PSTs and teacher educators participate and further legitimize their membership.

However, as noted earlier, this participation in the third space comes with various tensions and challenges. As a group of partners mutually participate to negotiate meaning together through some endeavor over time, they need to develop a common understanding and adopt processes and constructs for participation as shared practices. In this way, both participation and negotiation of meaning are fundamental processes by which communities of practice form. These two criteria for communities of practice are fundamental in third space partnerships. Therefore, the aim is to continuously work together toward achieving effective collaboration and sustainability rather than arriving at the metaphorical third space “podium.”

Inside looking out: PSTs' voices and experiences

The present study highlights the voices and experiences of a group of PSTs regarding practice. The overarching aim is to approach the inner workings of third space partnerships from PSTs' perspectives and understand how they integrate their practice experiences into their development as teachers. It also considers how aspects of participation, negotiation of meaning, and collaboration are crucial factors related to the enculturation process that PSTs find relevant in third space partnerships. Therefore, our research question is twofold: *How do pre-service teachers perceive practice, and how do they integrate their practice experiences into their professional development?*

Methodology

Since we aim to explore a complex social phenomenon from PSTs' perspectives, we selected the case study as a suitable research design (Cohen et al., 2018; Yin, 2014). The focus of the study are two cohorts of PSTs attending a TE program at the University of Oslo, Norway. The participants were enrolled in the Practical Pedagogical Education Program, a profession-oriented and experience-based TE program that qualifies teachers to work in upper primary and secondary schools, as well as in adult education. All participants hold a master's degree in their disciplines. The program's curriculum offers teaching qualifications and competence in subject didactics and pedagogy. This enables PSTs to plan, implement, and critically analyze teaching based on reflections from classroom experience. The program offers two alternatives: full-time study (1 year) and part-time study (1.5 years) and grants 60 European Credit Transfer and Accumulation System (ECTS) credits upon completion. During the program, PSTs are required to complete 12 weeks of professional practice.

Sample and data collection

All of the PSTs in the 2020–2021 cohort were invited to participate in an online survey regarding their perceptions of practice. The survey was conducted in two phases: before and after practice. This intended to capture as much information as possible from PSTs' experiences and perceptions. The data collection was conducted in fall 2020 and spring 2021. At the time, all the participants had experienced a practice placement at least once and were halfway through their TE program. Ethical approval for data collection was received in advance from the Norwegian Center for Research Data.

The online surveys served as a starting point to capture the general picture of PSTs' perceptions about and experiences with practice. The surveys included open-ended questions and four-point Likert scale questions. The multiple-choice questions served to gain insight into PSTs' perceptions on crucial aspects of practice and encouraged them to choose a standpoint (Clason and Dormody, 1994). The open-ended questions, on the other hand, were designed to explore their perceptions using their own words and ideas.

In line with an explorative approach (Kvale, 2007), we prepared a semi-structured interview guide based on the survey answers. The objective was to ask general follow-up questions, requiring participants to further reflect on their responses, address possible contradictions, speak up, and provide an “inside looking out” perspective of practice in the third space. Therefore, the semi-structured interviews are the backbone of the data corpus, and the surveys serve to provide general background information. All PSTs consented to participate in the study, and complete anonymity was guaranteed. In total, 184 surveys were answered before and after practice. As the surveys were anonymous, it is impossible to say how many answered prior and after. Thus, there are 184 unique answers, but the overlap is unknown. 28 PSTs participated in the interview conducted after practice using the cloud-based video conferencing system Zoom, and each lasted an average of 60 min. For an overview, see Table 1.

Table 1 Overview of data collection.

<i>Method-data tool</i>	<i>Participants</i>	<i>Quantity</i>	<i>Interview length</i>	<i>Total</i>
Surveys	Part-time cohort	77		184
Surveys	Full-time cohort	107		
Interview	Part-time	14	13 h 28 min	28
Interview	Full-time	14	14 h 12 min	

Data analysis

In the first phase of data analysis, the SPSS software was used to distribute and generate descriptive statistics of the Likert scale questions. The open-ended questions were categorically analyzed and used to assign abstract codes to ideas, phrases, and significant words. This initial approximation of the data analysis aimed to provide a general understanding of PSTs' lived experiences of practice and their perceptions about their learning process in the TE program. This first survey analysis helped to design the interview guide (Patton, 2015). All participants who completed the surveys were invited to participate in an individual follow-up interview, which was semi-structured to facilitate the free expression of participants' thoughts (Cohen et al., 2018). Twenty-eight PSTs agreed to participate in the interview and consented to their responses being audio-recorded. Notes taken during interviews were anonymized and included keywords, ideas, and verbatim quotes.

A research assistant transcribed the interview data. An inductive approach was then taken to code the transcripts and identify categories related to participants' practice experiences. The initial process included a vertical analysis (Miles et al., 2019a) where each interview transcript was analyzed separately and both NVivo and abstract codes were assigned to significant ideas and phrases. The second process involved comparing and contrasting the data to develop broad categories and identify relevant themes in and between transcripts (Coffey and Atkinson, 1996). Finally, the analysis involved identifying similarities between the themes generated by the two data sets (interviews and surveys). Then, the whole data set was scanned once more to ensure that all commonalities and relationships across themes were included. Pseudonyms are used throughout the reporting of results.

Findings

Descriptive overview

Participants undoubtedly find that practice is a crucial component of their TE. In fact, for most of the respondents, practice is the space where they can "apply" what they have learned on campus and take a closer look into the teaching profession. From all participants in the survey, 77 indicated having had a good experience with practice, which directly relates to the perception of 95 participants who marked that good communication and collaboration with the SBTEs had a significant impact on their positive perception of practice.

Feeling prepared to conduct practice professionally, emotionally, and practically was crucial in determining the outcomes of participants' practice experiences. Almost half of the participants responded that they felt underprepared. However, there was a significant difference between their readiness perceptions from the first practice experience and the second one. One of the reasons for this was that having conducted one practice boosted their confidence, as they now knew how it could unfold. For others, the challenges to be ready for practice were more linked to frustration and issues connected to the particular circumstances of conducting practice in the midst of the COVID-19 pandemic. This finding highlights how PSTs' psychological well-being is essential to their learning process.

Regarding the integration of knowledge sources, the data shows that 29 participants could interconnect expertise from practice and theory to a large degree and 51 to some degree. On the other hand, 18 participants noted that they could only do so to a small degree and found what they learned on campus to be of little relevance to what they ended up doing in practice. This was further explored in the interviews in relation to the challenges participants faced in practice.

Voices and experiences

Four major themes were generated from the data analysis: epistemic negotiations, relational factors, identity development, and opportunities and challenges. These themes are composed of several sub-themes that reflect how PSTs perceive practice and integrate it into their professional development (see Fig. 1). Fig. 1 shows, at the center and in constant movement, the third space dynamism where all community members relate to each other and engage in continuous participation and negotiation of meaning. The boundaries of the four themes are blurry and often overlap. Hence, though the themes are presented separately for the sake of clarity, some aspects within one will be relevant to and intertwined with the others. This is why the opportunities and challenges theme is presented as an integrated part of the other themes.

Epistemic negotiations

Epistemic negotiations represent a fundamental factor in participants' practice experiences. The findings indicate that participants highly value practice as an opportunity to integrate knowledge in a school setting. These negotiations occur both at the individual and at the community levels. At the individual level, participants attempt to draw connections between theoretical and practical expertise based on their own experience. However, the negotiations gain greater relevance when the participants observe, discuss, compare, and collaborate with other members of the partnership. Epistemic negotiations are closely associated with the themes of relational factors and identity development. However, the focus here is on the opportunities and challenges participants face when interconnecting knowledge sources.

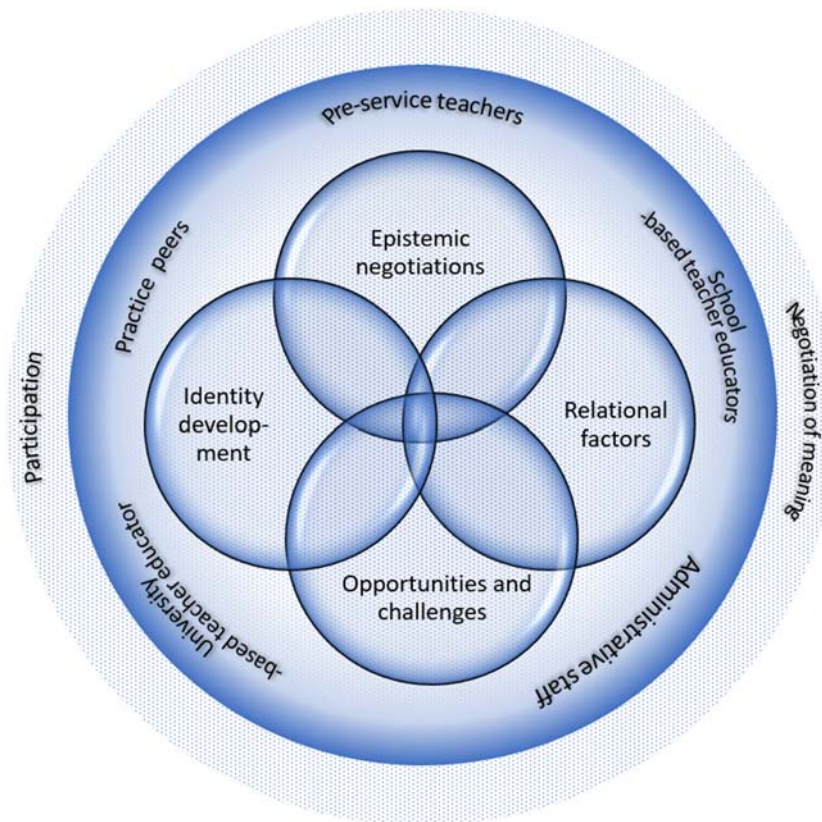


Fig. 1 Inside looking out. The third space partnership as a community of practice.

Interconnections, feedback, and other blending opportunities

The findings show that learning by observing and doing bolsters the interconnection between theoretical and practical expertise. Some participants initiated their practice by observing the SBTE teach lessons before trying on their own. The participants argue that such activities provide them with tools to weave theory and practice together, improve their confidence, and reduce the initial pressure of conducting a whole teaching lesson on their own. Noah explains this in the following excerpt:

It was also very useful to have practice. Because then you can see the connection between the theoretical knowledge you have from the university and how to relate it into practice and use it at school; everything makes more sense ... We could experience before our own eyes how it is done in practice ... We had conversations to summarize and discuss what we had seen. So, in a way, this was a feedback session where we discussed our observations in light of the theory.

Participants also report that feedback sessions and conversations that connected knowledge sources and emphasized concrete aspects, such as relationship-building, motivation, and classroom management, often occurred before and after teaching pupils. The PSTs had a shared meeting with their practice peer and the SBTE where they focused on identifying the positive outcomes of the teaching session as well as possibilities for improvement. On some occasions, participants expressed that they could discuss relevant pressing issues from pedagogy or didactics. For instance, Noah, his practice peer, and the SBTE dedicated a few of the discussion sessions to work on adapted education and planned activities that could be implemented with pupils. Noah's practice experience resonates with the idea of weaving expertise by observing and doing, which can occur at different stages of the TE process at schools and on campus.

However, these opportunities are curtailed when the partners do not engage in discussions that highlight both the theoretical and practical aspects of teaching and are not receptive of the vital role they play in their professional development. The participants who emphasize observations, feedback, and discussion report that some of the teacher educators have been educated in supervision, which could suggest that they are cognizant of their roles as teacher educators. This is not to say that teacher educators without supervision education do not provide connecting opportunities to PSTs. Instead, it highlights the relevance of observations,

feedback, and collaboration in PSTs' professional development. After a challenging first practice period, Emma draws on interesting ideas regarding the implications for her professional development:

For me, practice is learning, it is a connection opportunity. I felt a bit like a "fish on land" in the first practice. I stood there [*in the classroom*] and asked myself: "How do I do this? How to deal with discipline?" We did not have a model to observe first ... I was lucky with my supervisors the second time; they talked to us, they showed us how [*to do it*]. Theory and practice seemed much more connected. I should say that practice is a very safe environment to test oneself and share information with other knowledgeable [*people*] before trying it out as a new teacher on my own.

Blending opportunities occur when "doing" as well. Beyond observations, Olivia notes that conducting the teaching lessons herself, with the support of the SBTE and practice peers, enriched her experiences and allowed her to enact her role as teacher and become increasingly confident in her pedagogical and didactical knowledge. Olivia states:

I got to practice a lot as the teacher. I developed a slightly different relationship with the pupils. The supervisor sat in the back of the classroom, and if there were any inquiries to the supervisor, he would say: "I'm not the one who has this class." So I knew I was *the teacher* and that I had his support in case something went wrong. I think it was very rewarding to try out my ideas in a safe environment. I now have a different perspective on what it means to teach a class.

Other opportunities that foster interconnections are the microteaching sessions and the authentic cases or examples that SBTEs and UBTEs use in lectures, seminars, feedback, and assessments. In the microteaching sessions, PSTs plan and conduct a teaching segment of 6–8 min to a small group of PSTs and a session leader. The session leader can be either a UBTE or SBTE. Participants find that the microteaching sessions allow them both to try out the teacher role in a safe environment before going to practice, and also serve as their first opportunity to put some of the theoretical concepts they learned into action and then discuss these with others in the community.

On negotiations

As we have explored so far, negotiations occur at schools and on campus in dynamic relationships with teacher educators, peers, pupils, and other members of the partnership community. Recognizing the different approaches to teaching appears to be essential for participants' practice experiences. For instance, Martha illustrates an explicit epistemic negotiation with the SBTE that blends with challenges and power tensions, and showcases how participants negotiate meaning:

He [*The SBTE*] talked about a bit of his experience and why he did things the way he did. Among others, he had moved away from talking about curriculum goals and such in class. He instead had "today's question." He had a bit of his own philosophy. And then I told him: "We have been told [*on campus*] that we should anchor the teaching to the curriculum and learning objectives." And then he said that "that does not exactly motivate the students." ... So, he rather uses "today's questions," which comprise the objectives, but the pupils are not made aware of that. And I wondered: "Should I listen to them [*UBTEs*], or should I listen to the supervisor [*SBTE*]?" But then I had to be a mentee! So I chose to follow the supervisor, but I was a little unsure if that was the right thing to do then.

In this example, Martha eloquently explains how she engaged in negotiations of meaning with her SBTE. As conflicting approaches to starting a lesson surfaced, she was left wondering which approach would be more appropriate. In the excerpt, we can see how Martha decided to follow the SBTE's approach and be a "mentee," demonstrating the power relations in TE. The negotiation of meaning that Martha carried out later in her practice included trying out both ways to initiate a lesson (i.e., "today's question" and presenting the learning goals) with her practice peer and arriving at her own conclusions on what could work best based on her experience.

James's practice experiences also illustrated epistemic negotiations in the third space. James and his practice peer wanted to try out different theoretical concepts, such as consolidation and differentiation. Although the topics and lesson goals were already established for James's classes, he engaged in negotiations with the SBTE and included activities that highlighted such concepts. When asked about why this was relevant for him, James answered that these concepts are often abstract, and he needed to put them to work in practice to be able to grasp the connections. However, James further reflects on how this critical aspect of practice is not always achievable, as one is dependent on the school's plans and the SBTE's willingness to negotiate and foster meaningful learning experiences.

Anna's negotiation in practice had to do with the lesson plan. Anna argues that the detailed lesson plan promoted in lectures, seminars, and other activities on campus was not equally underlined at the school. While planning the lessons together, Anna and the SBTE agreed on the importance of learning how to narrow down competence goals, establish learning objectives for the lessons, and bear in mind the pupils' needs. However, they decided to develop a more simplified plan that functioned just as well in fast-paced teaching work. In this negotiation, Anna's SBTE had an essential role in supporting Anna's participation and interweaving the knowledge that Anna had been acquiring on campus with the SBTEs' expertise from school classrooms.

Certainly, teacher educators' role in promoting meaningful learning experiences and extending meaning is crucial for participants' perceptions of practice. Sonia validates this with the following excerpt: "If there is a big gap between what you learn in theory and practice, that's where the teacher [teacher educator] comes in and builds that bridge. They can explain why they teach this or that way, bring 'real' cases to campus, bring the campus to school. Make those connections visible."

In general, participants perceive both the practical and theoretical elements of TE as constituents of their professional development. While some of their experiences reflect the complexity of interconnecting knowledge sources in practice, others found that certain activities at schools and on campus could potentially enhance the intersection of epistemologies in TE. Such opportunities integrate authentic cases to exemplify theoretical concepts on campus and vice versa at schools by taking advantage of the hybrid expertise of third space partnerships.

Relational factors

Communicating, collaborating, and cultivating relationships with other members of the third space partnership is a requisite for participation and managing the tensions that emerge when working together. This section will explore the opportunities and challenges related to forming relations with teacher educators at schools and on campus, practice peers, pupils, and others in practice. These relations are linked to perceptions of value, integration, feeling welcome, knowledge creation and sharing, and PSTs' psychological well-being.

Working together or working with?

The findings indicate that the most relevant relationships participants establish are with their SBTE and their practice peers. The relationships with the UBTEs are also relevant but are mentioned to a lesser degree. Based on the quality of these relations, the participants perceive the practice either as positive or negative or a composite of both, which denotes the complexity of relations and the role that power, tensions, and other challenges play in shaping PSTs' perspectives and learning opportunities.

The main factors that participants highlight are communication and feelings of inclusion, respect, trust, and constant support. Besides the SBTEs, these factors relate to the partnership as a whole (other teachers at school, UBTEs, administrative staff at both school and university). As Emma puts it, positive relations in practice mean feeling "*valued and seen as practice students, colleagues, and future teachers.*" Hence, participants perceive that when relations with SBTEs are positive, the negotiations and participation flow more naturally, enhancing learning opportunities and making room for fruitful discussions.

Frida explains how she developed a very good relationship with her SBTE and practice peer from day one:

I had a great experience with practice. They [all members at the partner school] are wonderful—everyone, from the principal to the secretary, to the other teachers, to my supervisor [SBTE] and practice peer. I felt welcomed from the beginning. The school was prepared to receive us and showed us all around. The supervisor built a very good relationship with us: excellent practice supervisor, very good support, and suggestions for teaching. I could not expect more. Everything was very good.

Frida's practice experience elucidates the vital role of relations in practice in third space partnerships. James holds a similar perception when referring to the practice peer. Working together with his SBTE and practice peer promoted a sense of collaboration and confidence: "The practice peer and I worked together a lot. In meetings with the supervisor [SBTE] and him [*my practice peer*], I got many confirmations that meant a lot to me: that this is the right way to go and that I have what it takes to be a teacher."

It appears that PSTs' psychological well-being is fundamental to their professional development. Feeling welcomed goes beyond the material logistics that facilitate teaching. It involves the integration of PSTs into the mainstream of the community in an environment that is tolerant and that supports participants' progressive learning and continued adaptation.

Another relevant factor to effectively integrating PSTs in partnerships is trust, which is related to the feeling of assurance participants experience when given more and greater responsibilities and opportunities to interact with the school community in and outside of the classroom. Jacob highlights this by stating that he was able to try out new ideas in the classroom with the SBTEs' support and interact with pupils and other teachers in various activities.

However, not all participants were as satisfied with their relations in practice. When PSTs experience a lack of communication, unwillingness to engage in discussions, unpreparedness, or any other impediment to establishing a relationship with the SBTE, the overall perception of the practice and their opportunities to integrate prior knowledge is negative. In this context, participants refer to dissonance, miscommunication, chaos, unpreparedness, and lack of guidance and report feeling "alone" in practice. Sara explains her experience with miscommunication and unpreparedness:

We [the PSTs] were told that the supervisor [SBTE], or at least the school, would contact us before practice. And it never happened. We sent an e-mail to the supervisor a few days before, and we did not get an answer. So when we got there [the school], the supervisor said, "Oh yes, it was *today* you start!" Everything in practice was a bit disorganized, Very "ad hoc," I would say.

Negative practice experiences are also often related to dissonance: the difficulties PSTs have in connecting the curriculum and the guidelines with actual implementation in practice. For some of the participants, the practice expectations were vague, as they

expressed having limited knowledge of what to observe, what to ask from SBTEs, and a general sense of confusion about the transition from theoretical conceptualizations to practice. For instance, one participant, Louise, wonders, "How many hours should the SBTE guide me? Should I try to apply the theoretical knowledge from lectures to my lesson plans? Am I allowed to suggest topics? And what if they do not fit with the school plan? What exactly will I teach?" Louise further explains that she felt like a "burden" to the SBTE, like she was in the way of his/her busy schedule. She also found scanty assistance from the campus. Louise could not engage in feedback sessions with her practice peer, as the SBTE could not do so due to time constraints, only having a few general feedback sessions.

Margarita says she felt "alone" during practice, having limited contact with her SBTE and no practice peer at all. Regarding the latter, she argues that she wished she had a partner with her in practice, expressing that she missed out on the learning opportunities of having someone to discuss and plan lessons with. Similar to Louise's example above, Margarita's SBTE could not dedicate as much time to feedback sessions or observations. Instead, Margarita was left to teach several lessons without the SBTEs' presence: "They [*the school*] say that things are different in practice than what we learn at the university and that it is not like that in reality; she [*SBTE*] trusted me, I think. She left me alone to teach several times. I felt a bit like a substitute and not a student in practice." Margarita's practice experience shows how opportunities to negotiate meaning and weave knowledge are sometimes missed due to a lack of communication and guidance. In addition, she perceived her role as one of a substitute because she was left to teach alone several times. Correspondingly, performing different and hybrid roles is equally relevant in PSTs practice experiences.

Identity development

When participating and negotiating meaning in third space partnerships, all members of the community of practice commute across blurred boundaries and perform hybrid roles. Such dynamic movements and enactments foster PSTs' identity development.

Moving back and forth

As seen in the previous sections, the roles of SBTEs, UBTEs, practice peers, and other members in the partnerships are crucial to help PSTs interweave knowledge and establish connections. In this sense, participants actively perform hybrid roles in practice, as they are at the same time learners, colleagues, teachers, and partners.

Luna's comment illuminates these movements: "Sometimes I felt like a colleague, planning the lesson together with my supervisor [*SBTE*] and my practice peer, discussing possible issues, reflecting on my teaching and how to best improve it. In a way, I was both a student and a teacher." With this, Luna highlights the relevance of good communication and collaboration in practice and places herself in a hybrid role that constructed her identity as an active member of the teaching community. Other participants, such as Martin and Olivia, express similar ideas about their hybrid roles and their constant commute between different settings and practices. Martin says, "I think I was closer to being a colleague than an apprentice. I got to be the teacher who is responsible for what happens in the classroom. It's empowering; it showed that she [*the SBTE*] trusted us [*PSTs*]. We were included in meetings, and they asked for our opinions. We were not only taken into account but also heard. Like we were teachers there [*at the school*]."

Martin's excerpt further demonstrates how roles overlap and interrelate with one another. Martin seemed to have a good relationship with his SBTE and practice peer, as they created opportunities to interconnect knowledge and collaborate. These factors are embedded in participants' movements back and forth between school and academia; as a result, they enacted hybrid roles that supported their professional development.

The extent to which PSTs feel encouraged to try out new ideas, make mistakes, and later reflect on them enhances their perceptions of achievement, stimulates connections, and integrates PSTs into the community. Conversely, as we have examined before, faulty relations and a lack of participation or negotiation negatively affect PSTs' overall perception of practice and hinder the movements between theoretical and practical knowledge. For instance, Andrea explains, "It was difficult to implement what we had read in both pedagogy and subject didactics; the framework at school was narrow, and we [*my practice peer and I*] had to work on this alone. Sometimes we felt unprepared to face key issues of practice, we had not talked about that on campus. I feel I did not deliver good work; I was not a good teacher but not a good student either. It was frustrating, so the practice became demotivating." There are several implications for Andrea's professional development in the reflection above. Andrea found it difficult to draw connections between theoretical and practical knowledge, indicating challenges in epistemic negotiations and relations. Feeling alone indicates that Andrea required guidance and more opportunities to weave together knowledge sources from both the practice school and from campus. Commuting in the third space partnership and performing hybrid roles can further challenge the integration of knowledge and the identity development of PSTs who, like Andrea, face frustration and limited opportunities to participate and negotiate.

Discussion

This study contributes to and supports the growing body of research suggesting the third space as an effective partnership model to support teachers' professional development. Although the model may present challenges regarding tensions in correspondence, implementation, relations, and resources, the emphasis on horizontal relationships promises to balance the core elements of TE, granting equal value to theoretical and practical expertise. If we genuinely want to be part of the change process and influence the coming generation of teachers, we should aspire to break the vicious circle of disconnections and hierarchies that have dominated educational discourse (Barak et al., 2016).

This study reinforces four pivotal yet practical considerations for designing third space partnerships based on the “inside looking out” perspectives and experiences of PSTs.

- (1) **Fostering epistemic negotiations:** PSTs in this study implied on numerous occasions the indisputable relevance of weaving together knowledge from schools and campus in meaningful ways. Based on such negotiations, the PSTs are able to make sense and integrate the different knowledge sources into their development as teachers. Therefore, the design of practice and collaboration should include clear guidelines that stimulate the movements and integration of knowledge back and forth between these institutions.

The design of third space partnerships should consider the benefits of observation, feedback, and discussions in professional development, as highlighted by the participants. These support PSTs' learning, improve coherence, and involve all community members in participation, negotiation, and identity development. Such dialectical opportunities should be practiced on campus and at schools to promote epistemic movements within, across, and beyond the peripheries of practice. In this way, negotiating meaning in third space partnerships is an iterative process that requires active engagement and participation (Wenger, 1998).

- (2) **Cultivating relationships:** effective participation and epistemic negotiations depend to a large extent on the relations that participants establish in third space partnerships. Cultivating collaboration and professional relations, as this study demonstrated, is key to weaving knowledge, negotiating meaning, fostering PSTs' confidence, and ensuring their psychological well-being. Relations are, however, finely balanced and are often strained or limited due to time constraints, unclear expectations, and miscommunication. Relations are the backbone of third space partnerships in a community of practice that welcomes participation and work together in the process of “becoming.” Thus, cultivating them would enable participation and negotiation among partners. For PSTs, the essential relation factors in practice involved communication and feelings of inclusion, respect, trust, and support. Therefore, third space partnerships should enrich communication channels that connect participants and create more precise agreements that cultivate relations and encourage connection.
- (3) **Developing identities:** participation is a key component in third space partnerships. In weaving the community together through collaborative practice and the continuous negotiation of meaning, PSTs gradually move beyond the periphery toward the core of practice (Lave and Wenger, 1991). By crossing the blurred boundaries and performing hybrid roles in practice, the PSTs in this study suggest that they developed their identity as learners, teachers, and colleagues, negotiating meaning. When performing hybrid roles, participants (as well as other members of the teaching community) negotiate their identities by doing, discussing, feeling, belonging, and working together (Wenger, 1998), increasingly becoming more knowledgeable teachers and community members. The back and forth exchange of expertise in third space partnerships bound the community members together and promote hybrid roles that ultimately shapes their identities.
- (4) **Navigating challenges:** even if third space partnerships appear promising, their establishment and sustainability face numerous challenges. The challenges explored in this study suggest that, despite efforts from the partnerships, connections are not robust enough to foster epistemic negotiations, cultivate relations, and develop identities for all PSTs on an equal basis. While some PSTs' experiences were positive and fruitful, others found practice frustrating. These challenges are closely related to the absence or poor execution of the three themes presented above. Hence, a possible way to navigate such challenges would be to strengthen the third space partnerships by enhancing open communication and creating opportunities for negotiation and development.

Third space partnerships are blended Communities of Practice that bolster members' agency to work together in the process of “becoming” and navigating the impending challenges associated with this collaboration. It involves action and connection (Wenger, 1998). While developing their professional identity, PSTs integrate their practice experiences into their professional repertoire by participating, negotiating, and exchanging expertise with other community members. PSTs become increasingly knowledgeable on the institutional regulations and formal requirements of the teaching profession while developing a broader sense of the epistemologies of their school subjects, the pedagogy, and didactics. PSTs also develop relations that promote negotiations and involve producing, redirecting, and extending meaning. Seen as a community of practice, the third space supports an integrated way to connect minds and bodies, including emotions and social connections, while bolstering PSTs' professional development.

As increased coherence and collaboration in partnerships are central elements in research and policy (Hermansen, 2020; Farrell, 2021; OECD, 2015), more significant efforts are needed to improve these alliances to support PSTs' professional development. Consequently, this study has attempted to understand partnerships from within, listening to PSTs' voices and experiences regarding their roles, expectations, participation, and negotiations. Such explorations may shed further light on how third space partnerships collaborate and the crucial factors that determine successful collaboration.

Conclusion

In summation, this third space take on PSTs' practice reveals the continuous challenges and opportunities involved in becoming teachers in more democratic environments. In providing empirical evidence of how PSTs experience practice, we hope to have contributed to clarifying the complex dynamics of third space partnerships and identifying the key ingredients to its development and sustainability. The study supports earlier claims that the third space in TE can bolster connections between schools and TE

institutions and enhance collaboration (Jackson and Burch, 2019; Chan, 2020; Lee, 2018) while also suggesting that PSTs' practice experience requires sound opportunities for epistemic negotiation, active participation, and collaboration that foster their identities and professional development.

Acknowledgment

I would like to thank the pre-service teachers for participating in the study. I also wish to thank Professors Greta B. Gudmundsdottir and Andreas Lund and the other members of the TEPEC research group and the University of Oslo for commenting on an earlier version of this study.

References

- Allen, J.M., Wright, S.E., 2014. Integrating theory and practice in the pre-service teacher education practicum. *Teach. Teach.* 20, 136–151.
- Allen, J.M., Ambrosetti, A., Turner, D., 2013a. How school and university supervising staff perceive the pre-service teacher education practicum: a comparative study. *Aust. J. Teach. Educ.* 38 (4), 108–128.
- Allen, J.M., Howells, K., Radford, R., 2013b. A "partnership in teaching excellence": ways in which one school-university partnership has fostered teacher development. *Asia Pac. J. Teach. Educ.* 41, 99–110.
- Barak, J., Gorodetsky, M., Hadari, H., 2016. In-between school and college. In: Barak, J., Gidron, A. (Eds.), *Active Collaborative Education: A Journey Towards Teaching*. SensePublishers, Rotterdam.
- Barragan, A.G.C., 2017. Building third space in the initial teacher training. A practical experience from funds of digital knowledge and identity perspective. *Revista Latinoamericana De Tecnologia Educativa-Relatec* 16, 35–49.
- Barton, D., Tusting, K., 2005. *Beyond Communities of Practice: Language Power and Social Context*. Cambridge University Press.
- Beaton, F., Hope, J., 2022. Exploring the identities of blended professionals in higher education. In: McIntosh, E., Nutt, D. (Eds.), *The Impact of the Integrated Practitioner in Higher Education: Studies in Third Space Professionalism*, first ed. Routledge, London.
- Beck, J.S., 2018. Investigating the third space: a new agenda for teacher education research. *J. Teach. Educ.* 71 (4), 379–391.
- Bhabha, H.K., 1994. *The Location of Culture*. Routledge, London.
- Bloomfield, D., Nguyen, H.T.M., 2015. Creating and sustaining professional learning partnerships: activity theory as an analytic tool. *Aust. J. Teach. Educ.* 40 (11), 22–44.
- Broadley, T., Martin, R., Curtis, E., 2019. Rethinking professional experience through a learning community model: toward a culture change. *Front. Educ.* 4 (22), 1–15.
- Buckreis, S., 2012. *Education as a Journey: Exploring the Third Space of Hongyu Wang's the Call From the Stranger on a Journey Home*. Foundation for Curriculum Theory, Rochester.
- Campbell-Perry, S., 2022. The blended professional. Barriers and boundaries to collaborative institutional cultures. In: McIntosh, E., Hope, J. (Eds.), *The Impact of the Integrated Practitioner in Higher Education. Studies in Third Space Professionalism*, first ed. Routledge, London.
- Chan, C., 2019. Crossing institutional borders: exploring pre-service teacher education partnerships through the lens of border theory. *Teach. Teach. Educ.* 86.
- Chan, C., 2020. I know how it feels: how online mentors help pre-service teachers negotiate practicum tensions in the third space. *Mentor. Tutoring* 28, 189–210.
- Clason, D.L., Dormody, T.J., 1994. Analyzing data measured by individual Likert-type items. *J. Agric. Educ.* 35 (4), 31–35.
- Cochran-Smith, M., Lytle, S.L., 1999. Relationships of knowledge and practice: teacher learning in communities. *Rev. Res. Educ.* 24, 249–305.
- Coffey, A., Atkinson, P., 1996. *Making Sense of Qualitative Data: Complementary Research Strategies*. SAGE, Thousand Oaks.
- Cohen, L., Manion, L., Morrison, K.R.B., 2018. *Research Methods in Education*, eighth ed. Routledge, London.
- Cuenca, A., Schmeichel, M., Butler, B.M., Dinkelman, T., Nichols, J.R., 2011. Creating a "third space" in student teaching: implications for the university supervisor's status as outsider. *Teach. Teach. Educ.* 27, 1068–1077.
- Darling-Hammond, L., 2006. Constructing 21st-century teacher education. *J. Teach. Educ.* 57, 300–314.
- Darling-Hammond, L., 2010. Teacher education and the American future. *J. Teach. Educ.* 61, 35–47.
- Darling-Hammond, L., 2014. Strengthening clinical preparation: the holy grail of teacher education. *Peabody J. Educ.* 89, 547–561.
- Darling-Hammond, L., 2017. Teacher education around the world: what can we learn from international practice? *Eur. J. Teach. Educ.* 40, 291–309.
- Daza, V., Gudmundsdottir, G.B., Lund, A., 2021. Partnerships as third spaces for professional practice in initial teacher education: a scoping review. *Teach. Teach. Educ.* 102.
- Elsden-Clifton, J., Jordan, K., 2016. Reframing professional experience. In: Barkatsas, T., Bertram, A. (Eds.), *Global Learning in the 21st Century*. SensePublishers, Rotterdam.
- Elsden-Clifton, J., Jordan, K., 2019. Who is the hybrid teacher educator? Understanding professional identity in school-university partnership. In: Gutierrez, A., Fox, J., Alexander, C. (Eds.), *Professionalism and Teacher Education: Voices From Policy and Practice*. Springer Singapore, Singapore.
- Farrell, R., 2021. The school-university nexus and degrees of partnership in initial teacher education. *Ir. Educ. Stud.* 1–18.
- Hermansen, H., 2020. In pursuit of coherence: aligning program development in teacher education with institutional practices. *Scand. J. Educ. Res.* 64, 936–952.
- Jackson, A., Burch, J., 2016. School Direct, a policy for initial teacher training in England: plotting a principled pedagogical path through a changing landscape. *Prof. Dev. Educ.* 42, 511–526.
- Jackson, A., Burch, J., 2019. New directions for teacher education: investigating school/university partnership in an increasingly school-based context. *Prof. Dev. Educ.* 45, 138–150.
- Keating, M.C., 2005. The person in the doing: negotiating the experience of self. In: Barton, D., Tusting, K. (Eds.), *Beyond Communities of Practice: Language Power and Social Context*. Cambridge University Press, Cambridge.
- Klein, E.J., Taylor, M., Onore, C., Strom, K., Abrams, L., 2013. Finding a third space in teacher education: creating an urban teacher residency. *Teach. Educ.* 24, 27–57.
- Klein, E.J., Taylor, M., Onore, C., Strom, K., Abrams, L., 2016. Exploring inquiry in the third space: case studies of a year in an urban teacher-residency program. *New Educ.* 12, 243–268.
- Kvale, S., 2007. *Doing Interviews*. SAGE, Los Angeles.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Le Cornu, R., 2010. Changing roles, relationships and responsibilities in changing times. *Asia Pac. J. Teach. Educ.* 38, 195–206.
- Le Cornu, R., 2015. Professional experience: learning from the past to build the future. *Asia Pac. J. Teach. Educ.* 44, 80–101.
- Lea, M.R., 2005. "Communities of practice" in higher education: useful heuristic or educational model? In: Barton, D., Tusting, K. (Eds.), *Beyond Communities of Practice: Language Power and Social Context*. Cambridge University Press, Cambridge.
- Lee, R.E., 2018. Breaking down barriers and building bridges: transformative practices in community- and school-based urban teacher preparation. *J. Teach. Educ.* 69, 118–126.
- Lemon, N., Wilson, A., Oxworth, C., Zavros-Orr, A., Wood, B., 2018. Lines of school-university partnership: perception, sensation and meshwork reshaping of preservice teachers' experiences. *Aust. J. Teach. Educ.* 43, 81.
- Lillejord, S., Børte, K., 2016. Partnership in teacher education—a research mapping. *Eur. J. Teach. Educ.* 39, 550–563.

- Martin, S.D., Snow, J.L., Torrez, C.A., 2011. Navigating the terrain of third space: tensions with/in relationships in school-university partnerships. *J. Teach. Educ.* 62, 299–311.
- Mcdonough, S., 2014. Rewriting the script of mentoring pre-service teachers in third space: exploring tensions of loyalty, obligation and advocacy. *Stud. Teach. Educ.* 10, 210–221.
- Miles, M., Huberman, A.M., Saldaña, J., 2019a. *Qualitative Data Analysis: A Methods Sourcebook*. SAGE, Los Angeles.
- Miles, R.H., Garoni, S., Knipe, S., 2019b. Learning to be a professional: bridging the gap in teacher education practice. In: Guitierrez, A., Fox, J., Alexander, C. (Eds.), *Professionalism and Teacher Education: Voices From Policy and Practice*. Springer Singapore, Singapore.
- OECD, 2015. *Education Policy Outlook 2015. Making Reforms Happen*. OECD Publishing, Paris.
- Patton, M.Q., 2015. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. SAGE, Los Angeles.
- Ramsaroop, S., Gravett, S., 2017. The potential of teaching schools in enabling student teacher learning for the teaching profession. *J. Curric. Stud.* 49, 848–865.
- Snow-Gerono, J.L., 2005. Professional development in a culture of inquiry: PDS teachers identify the benefits of professional learning communities. *Teach. Teach. Educ.* 21, 241–256.
- Steele, A.R., 2017. An alternative collaborative supervision practice between university-based teachers and school-based teachers. *Issues Educ. Res.* 27, 582–599.
- Vagi, R., Pivovarova, M., Barnard, W., 2019. Dynamics of preservice teacher quality. *Teach. Teach. Educ.* 85, 13–23.
- Walsh, M.E., Backe, S., 2013. School–university partnerships: reflections and opportunities. *Peabody J. Educ.* 88, 594–607.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge.
- Wenger, E., 2000. *Communities of Practice and Social Learning Systems*, vol. 7. Organization, London, England, pp. 225–246.
- Wenger, E., 2010. *Communities of Practice and Social Learning Systems: The Career of a Concept*. Springer London, London.
- White, S., Bloomfield, D., Cornu, R.L., 2010. Professional experience in new times: issues and responses to a changing education landscape. *Asia Pac. J. Teach. Educ.* 38, 181–193.
- Yin, R.K., 2014. *Case Study Research: Design and Methods*. SAGE, Los Angeles.
- Zeichner, K., Payne, K., 2013. Democratizing knowledge in urban teacher education. *J. Teach. Educ.* 66 (2), 122–135.
- Zeichner, K., Payne, K.A., Brayko, K., 2015. Democratizing teacher education. *J. Teach. Educ.* 66 (2), 122–135.
- Zeichner, K., 2010. Rethinking the connections between campus courses and field experiences in college- and university-based teacher education. *J. Teach. Educ.* 61 (1–2), 89–99.
- Zeichner, K., 2012. The turn once again toward practice-based teacher education. *J. Teach. Educ.* 63, 376–382.

Case methods in teacher education

Mistilina Sato^a and Jake Knaus^b, ^a University of Canterbury, Christchurch, New Zealand; and ^b Minneapolis Public Schools, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of M. Sato, C. Rogers, Case Methods in Teacher Education, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 592-597, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00662-X>.

Case methods in teacher education	405
Recent research on the impact of case methods in teacher education	406
Recent research on the construction of case studies in teacher education	407
What have we learned from these recent research studies?	408
Proposing next steps for case methods practice and research	409
Exploring a case-based knowledge system for teaching	409
Curating a case-based knowledge system for teaching	410
Narrative spaces in teacher education	411
Conclusion	411
References	411

Case methods in teacher education refers to the use of case studies as narratives drawn from real-life instances and used as teaching tools to catalyze group discussion and individual reflection. Using case discussion, case writing, and case analysis is a pedagogical approach that aims to bring real-world events and experience vicariously into the formal teaching and learning spaces of teacher education. Case methods as pedagogy is distinct from case study as a research activity, while both are grounded in epistemologies of narrative and bounded stories.

In this article, we aim to do three things in order to review the topic of case methods in teacher education. First, we will review the prior encyclopedia article (Sato and Rogers, 2010) to provide the grounding that is already in place with regard to case methods. Second, we will provide a brief review of the literature from 2010 to 20 to bring us up to date on the current status of research in this area. Finally, we will propose a way forward for research and practice in case methods as we look to the future and what it might hold for using narrative as a form of holding and sharing knowledge for teaching.

Case methods in teacher education

In the 2010 volume of this Encyclopedia, Carrie Rogers and I summarized case methods in teacher education (Sato and Rogers, 2010). We began by pointing out that case methods have been used in teacher education since the 1980s and developed a substantial research and practice base as a pedagogy throughout the 1990s as it emerged as a promising pedagogy for teacher education (Grossman, 2005; Lundeberg et al., 1999; Merseeth, 1996). Indeed, *Teacher Education Quarterly* dedicated an entire volume to Case Methodology in the Study and Practice of Teacher Education in 1990 (Doyle, 1990).

In that 2010 article, we presented several frameworks for how case methods in teacher education have been framed through a variety of empirical studies with regard to what teachers learn through case-writing and through case analysis. For example, Merseeth (1996) suggested three purposes for the use of case methods: as exemplars or models of best practices; as opportunities for analysis of a complex situation and to practice decision-making; and to stimulate reflection. Adding to the call for research on case study methods, Lundeberg et al. (1999) suggested that the field needed to better understand *what* pre-service teachers actually learn from case methods and *how* that learning occurs. These authors suggested that teachers can learn reasoning ability, problem-framing skills, reflective mindsets, subject-matter knowledge, and pedagogical repertoires through the use of case methods, but these skills, knowledge, and mindsets had not yet been particularly well-defined by the field of teacher education. We also reported that several types of cases have been used within case pedagogy, including exemplars of practice, dilemmas of practice, historical issues, and representations of everyday practice; as well as several structures for presenting cases including written forms, anecdotes, and multimedia representations drawing on student work samples, classroom video, teacher narratives, and reflections.

In the 2010 Encyclopedia article (Sato and Rogers, 2010), we also presented a novel unpacking of the various theoretical frameworks that have implicitly guided some of the empirical work on the effectiveness, outcomes, and structuring of case pedagogy. Using case methods as a pedagogy has been located in a variety of learning theories to support their use as a learning tool. We reviewed situated cognition theory (Lave and Wenger, 1991), cognitive flexibility theory (Spiro et al., 1987), social cognitive theory (Bandura, 1986), reflective practice (Dewey, 1933; Schon, 1987), and practical reasoning (Dunne, 1993, based on Aristotle's concept of *phronesis*). We found that the empirical articles we reviewed often conflated multiple theoretical framings for why and how cases were used to support pre-service teacher learning. The articles that drew across multiple theoretical frames tended

to be less theoretically robust than those that clearly identified a specific learning theory for exploring what learning was occurring and how that learning took shape.

We highlighted a few studies that clearly grounded their research within a particular learning or reasoning theory in order to illustrate how particular frames on learning supported a more consistent and deeper analysis of how cases were mobilized pedagogically and defined the learning that resulted within specific contexts. Ultimately, we concluded that case methods have been adopted as a pedagogy because they tend to represent the complexity of practice and provide an opportunity for pre-service teachers to locate their learning in the particulars of practice while also drawing on educational theories to understand their practice. We summarized: "Thus, case methods have the potential to not only situate the learning of the teacher in a particular context but also to provide opportunities to theorize from practice while practicing theory (Sato and Rogers, 2010, p. 596). This conclusion drew on the work of Judith (Judy) Shulman who wrote about a "vision of teaching as a deeply professional and profoundly complex form of theoretically enriched practice." (2003, p. 2). We want to acknowledge Judy Shulman's foundational idea about learning to teach here and dedicate this article to her in recognition of her passing in 2021. She is a great loss to the field of teacher education research, especially with regard to case methods as a pedagogy.

Recent research on the impact of case methods in teacher education

For this article, two advanced searches of standard education databases and a Google Scholar search using "case method" and "teacher education" as keyword search criteria since 2010 were conducted. We found that no systematic research review on case methods in teacher education has been published in educational research journals since the 2010 Encyclopedia article was published. Through this review, 34 papers that represent empirical studies of case methods in teacher education published in English were found, as well as a review of case methods and its relationship specifically to preservice teacher decision-making covering the time period of 1990–2010 (Çevik, 2013). We first summarize the review by Çevik (2013) and then summarize the 34 research studies.

Çevik (2013) defined teacher decision-making as teachers' thought processes, reasoning, and instructional decision-making (p. 94). The author divided the 20 reviewed studies into four groupings: four studies of pre-service teachers using video recordings to analyze their own teaching; two studies of pre-service teachers analyzing exemplar case studies; eleven studies that used cases as sources of inquiry and problem-solving; and three studies in which pre-service teachers constructed their own cases for analysis. Çevik (2013) found that the use of case study can influence the growth and development of a preservice teachers' decision-making capabilities, such as seeing situations through multiple perspectives, using evidence to support claims and interpretations, and their ability to generate alternative solutions to problems and situations. Some limitations to the use of cases for the development of teacher decision-making included how the preservice teacher often relies on the word of perceived experts rather than their own careful case analysis, superficial analysis of complex situations due to the inexperience of the preservice teacher, and pre-existing beliefs and prior experiences guiding preservice teacher decision-making. Using the framework of novice-to-expert development, Çevik (2013) concluded that the use of case study pedagogy can enhance the development of expertise in teaching, but that the development of expertise still takes time. The particularities of the pedagogical approach of case method that were reported in the research studies was also an important factor in preservice teacher learning. Çevik reports that the integration of case discussion and analysis enhanced learning and multiple exposures to case analysis over time resulted in stronger learning outcomes.

The 34 research studies published since 2010 show that case methods are being used around the world as the studies are located in Australia, Austria, Canada, two from Germany, Hong Kong, Ireland, Israel, Japan, Malaysia, three from Taiwan, South Africa, Switzerland, five from Turkey, Great Britain, and 13 from the United States. In terms of teacher learning through case methods, these studies explore learning in curriculum areas of English, physical education, mathematics, music, reading, the sciences, and social studies. Other areas of teaching knowledge were also present in this set of studies, including professional ethics and moral reasoning, adolescent development, classroom management, culturally relevant pedagogy, and technology integration. Aspects of teaching such as understanding the complexity of teaching, teacher problem-solving, analytical thinking within situations, and the special context of teaching in rural areas were also represented. We also noted cases across the schooling sectors of early childhood, special education, secondary, and primary within this collection.

Substantively, these studies tend to ask questions about the effectiveness of the use of case studies as a pedagogy in terms of how they support a learning outcome for preservice teachers. Twenty of the studies report a positive outcome for preservice teacher learning in areas such as improving attitudes toward chemistry (Ayyildiz and Tarhan, 2012); using theoretical understandings in decision-making (Ching, 2014); learning about how students learn math (Osmanoğlu and Koç, 2013); analytic thinking ability (Frommelt et al., 2019; Zottmann et al., 2011); and fostering learning of culturally responsive pedagogies (Gunn et al., 2015).

Three studies showed that preservice teachers do not always respond to learning from case studies (Liu and Chen, 2019; Shin et al., 2019; Tal, 2010). Given that these studies are counter-narratives to the general consensus of effectiveness of case methodology, we will look at these three studies more closely. Liu and Chen (2019) set out to study how the use of case methods impacts preservice teachers' reflective thinking through their own action research in a teacher education course on "education issues" in a Taiwanese national university. Two issues were addressed using video cases followed by discussion and a group written reflection. Using a questionnaire about the preservice teachers' awareness of the education issue and qualitative analysis of the post-case discussion analysis, the authors concluded that the use of case study was effective at supporting and developing reflective thinking, but it did not alter students' fundamental viewpoints on the particular issues that the case addressed.

Tal (2010) asked students in a second-year teacher education course in Israel to construct cases from practice as a way of exploring and learning about classroom management as a complex activity. The cases were written using social episode analysis, in which they chose an episode from their experience, interpreted and analyzed it in terms of the classroom management theories learned in their course, their own effectiveness, and what could be improved in the future. The preservice teachers then selected a second episode to reflect on how they had improved their coping strategies within a similar situation. Of the 34 cases written by preservice teachers, the author chose two to showcase in this study. One case illustrated strong learning and development based on the course expectations and that classroom management practices began to align with the theoretical perspectives of moral and ecological approaches taught in the course. The second case illustrated how a preservice teacher seemed to miss learning opportunities in the case analysis, made leaps and assumptions about the students she was teaching, and stood firm in her personal beliefs about corrective punishment and obedience to authority as the classroom management approach she was trying to enact, which did not align with what she was taught in her teacher education course. The authors conclude that the case writing was beneficial to the students to develop their understanding and provided an avenue for analyzing practice. However, the case writing exercise was not sufficient for ensuring that all students fully embrace particular teaching approaches taught in the course, especially when fundamental beliefs are held onto. The authors also suggest that more time or more coaching of the preservice teacher who wrote the second case highlighted in the study could help further develop understanding and improve practice.

The third study in this set illustrated how preservice teachers' perceptions of the cases themselves can influence their learning process. Shin et al. (2019) explored the perceptions and experiences of the use of technology-enhanced cases with two cohorts of students in a social studies methods course in a United States university. The students reported both positive and negative perceptions. The preservice teachers reported that case methods provided students with a medium for sharing ideas and an authentic window into the details of teaching. Some of the preservice teachers within both cohorts also reported that they did not believe the cases were authentic to real practice they had observed in schools and that the classrooms in the videos appeared to be too controlled. The preservice teachers suggested in interviews that the lack of classroom management issues in the videos made it more difficult for them to engage in case analysis. The teacher educator facilitating the case discussions reported that even though the video cases were real teaching videos and not staged, the preservice teachers seemed to hold a view of classroom management that is uncoupled from well-designed and effective teaching that engages students. The authors suggest in using case study pedagogy, having case materials is not enough on their own to support student engagement; learning needs to be supported through providing a purpose, context, scaffolding, and facilitation.

Finally, one study stood out as atypical in its approach to examining case methodology. Burns et al. (2012) provided a philosophical argument about the assumptions embedded in a widely-used education ethics textbook that relies on case method (Strike and Soltis, 2009). This paper argues that the case method approach that Strike and Soltis (2009) advocate, which seems to be based on highly structured cases, develops dialogic competence but falls short of developing practical competence in ethical decision-making. The authors suggest other pedagogies that examine real situations from practice or more nuanced accounts from literature rather than the constructed and guided case analysis provided by the textbook under review.

Recent research on the construction of case studies in teacher education

Within this group of 34 studies, eight focus on the construction of cases and the features of case studies that contribute to teacher learning. For example, Kurz and Batarelo (2010) explored the features of case studies that primary school preservice teachers perceived as beneficial in video case studies and found three features that were considered constructive: (1) modeling of teaching techniques; (2) supplemental materials to support the video case; and (3) observing classroom management. McLinden et al. (2010) reported on the development of online, problem-based learning case scenarios for use in a distance education program for teachers of students with visual impairments. The authors tested the prototype cases with a group of 36 teachers enrolled in an advanced teaching qualification program in a UK University. Using questionnaires, the authors were able to gather feedback on how the online case design aligned with their design principles and get recommendations for further development. Saltan et al. (2016) used action research cycles to design and develop online case-based learning environments for primary preservice teachers at a university in Turkey. The online environment allows the preservice teacher to watch video cases and analyze them using scaffolds such as a structure for the case analysis, peer discussions, question prompts, and an expert response to the case. Using data from interviews and questionnaires, the authors identified six factors that contributed to the authenticity of the cases: commonness, providing different perspectives, filling emotion, holding experience, reflecting social facts, and being multimedia.

The self-construction of cases by preservice teachers has also been further explored in two of the studies located in this review. Rosaen et al. (2009) asked what teaching interns learn when they make their own video cases from their own teaching. The study zeroes in on how the perceived audience of the video cases shaped how the teaching interns presented their cases, what they chose to focus on, and what they learned from the case construction. The authors found that while all the preservice teachers demonstrated learning through the process of making a video case, the question of who was the audience for the video cases was important for the preservice teachers to understand in order to support their learning in particular areas.

Tulu (2016) explored how preservice teachers in an English language teaching program in Istanbul engaged in writing cases based on their practicum observations and reflecting on the cases collaboratively. Using post-course written reflection and focus group interviews, the author sought to identify the effect that the case-writing had on the students' learning. Preservice teachers reported that writing the cases based on their own practicum observations motivated them to better engage in the practicum

experience, to notice more during their observations, and that writing the cases created a meaningful learning experience for them. Constraints identified by the preservice teachers mainly centered on how to manage the time for sharing the cases written by the preservice teachers to ensure that all the cases get shared and discussed as part of their learning process.

Tinker Sachs and Ho (2011) describe the process of developing and field-testing cases used in a Teaching English as a Second Language teacher education program in Hong Kong. The cases were subsequently edited and published in a book (Tinker Sachs and Ho, 2007). The authors describe five stages of development of the case studies: (1) identifying the teacher participants; (2) interviewing the teacher participants for contextual background; (3) videoing lessons of the teacher participants; (4) preparing cases through episode selection and question generation; and (5) testing and refining the cases with the teachers and with preservice teachers. The article also describes the case structure and outlines the case pedagogy.

A novel use of cases in teacher education is to feature them as an evaluation or assessment tool within a teacher education program. Wagler (2009) used case narratives as an evaluation tool with primary preservice teachers taking a science methods teaching course in a US university. The author constructed two cases that depicted a positive and a negative science teaching case narrative based on using pedagogies of scientific inquiry (National Research Council, 1996). Preservice teachers were presented these cases randomly and asked to evaluate them in light of effective science teaching criteria as a means of determining if the preservice teacher could discern this type of science teaching.

The question of differential learning by preservice teachers with different backgrounds was explored by Liu (2013). This study examined the differences that preservice teacher characteristics make in terms of learning outcomes using dilemma-based cases in primary teacher education. Men achieved better results than women (contradicting previous research reports); those with no prior teaching experience achieved higher than those with prior experience; and the academic level of the teacher education program did not make a difference in terms of learning outcomes.

What have we learned from these recent research studies?

This collection of studies continues to support the efficacy of case study methods as a pedagogy that supports preservice teachers in developing their understanding of concepts and knowledge base, reasoning ability, and analytic thinking. Case methods are generally engaging for preservice teachers as they are based in authentic experiences, either derived from others or harvested from their own experiences. And case methods provide a tangible approach for building the relationship between theoretical understandings about teaching and learning and the practical decision-making in the enactment of teaching in everyday terms.

The limitations of the use of case methods are both pedagogical and practical. Pedagogically, the teacher educator must know how to position the case method in a scaffolded and facilitated way that helps the preservice teacher understand the purpose and the context of the case (Shin et al., 2019). Details about how to introduce and engage in case discussions continue to add to the field's understanding of how to enact this pedagogy (Tinker Sachs and Ho, 2011). The pedagogical practice of using case methods can be particularly rewarding and challenging when preservice teachers are asked to construct their own cases (Rosaen et al., 2009; Tuluze, 2016). This requires teacher educators to be clear on the purpose and audience for the case so that the preservice teacher can select and construct a case narrative that allows them to explore the required theoretical aspects of teaching, analyze and reflect on their own learning at the depth expected by the profession, and to share their case appropriately with others for collaborative learning from practice.

As a set of research studies, it appears that the field of teacher education is not fully embracing a research program (Sato and Sheppard, 2008; Shulman, 2004) around case methods in teacher education. Many of the papers find their roots in case methods in teacher education by citing some of the original arguments and research base for using cases as a pedagogy in teacher education (Lundeberg et al., 1999; Merseth, 1996; Shulman 1986). From there, the studies tend to diverge across a variety of theoretical frameworks such as professional socialization (Hemphill et al., 2015; Richards et al., 2012); reflective learning (Liu and Chen, 2019; Rosaen et al., 2009; Tuluze, 2016); narrative learning (Ayyildiz and Tarhan, 2012); phronesis and reasoning (Burns et al., 2012); planned behavior (Dündar, 2019); professional vision (Frommelt et al., 2019); cognitive learning (Kale and Whitehouse, 2012); transformational learning (Kindle and Schmidt, 2011); ill-defined problems (Koç, 2012); problem-based learning (McLinden et al., 2010); cognitive moral reasoning (O'Flaherty and McGarr, 2013); gamification (Heffernan et al., 2016); computer mediated communication (Osmanoğlu and Koç, 2013); social episode analysis (Tal, 2010); and cognitive flexibility (Zottmann et al., 2011). The most common theoretical space for these studies was in social constructivism (Ching, 2014; Florez, 2011; Gravett et al., 2016; Howell, 2012; Ohara and Edahiro, 2010) which is a broader epistemological stance as compared to the more specific concepts grounded in learning that the other studies used. Two studies had no discernible theoretical framework (Kurz and Batarello, 2010; Tinker Sachs and Ho, 2011) and five studies argued for case methods being used and studied within a framework of professional knowledge (Günther et al., 2019; Schritteser, 2014), exemplary practice (Mayall, 2010), and the theory-practice gap (Gravett et al., 2016; Shin et al., 2019). Two studies by the same author (Liu, 2013, 2019) went as far as to name the use of cases as a pedagogy as "case learning," maybe inadvertently making a move to name a new way of theorizing learning from cases, but the author did not go into an explicit discussion about how this might actually be introducing a learning theory in and of itself. This suggests that individual researchers are studying a pedagogical approach and its outcomes without much concerted focus on theory building within the field.

Proposing next steps for case methods practice and research

With only 34 research studies identified from a ten-year period, we suggest that case methods in teacher education seems to be so well accepted as a standard pedagogy, the allure of exploring its nuances through research activity seems to be waning. During a period of teacher education that saw a turn back toward practice (Ball and Cohen, 1999; Ball & Forzani, 2009; Grossman, 2011; Reid, 2011; Zeichner, 2012), we wonder if case pedagogy has become an everyday teaching practice and is not deemed as “research-worthy” as it was when it was novel in the 1990s. We contend that there is still much more work to do in understanding the role of, the practice of, and the impact of case methods in teacher education, as we discuss in the next section. To propose the next stages of a research agenda for case methods in teacher education, we return to the original theorizing of case methods in teacher education from Lee Shulman (1986).

Exploring a case-based knowledge system for teaching

When Shulman (1986) introduced the idea of the use of cases in teacher education, he originally spoke of cases as one way to shape the knowledge base for teaching. Shulman provided a framework for the knowledge base for teaching and proposed three forms of teacher knowledge—propositional knowledge, case knowledge, and strategic knowledge. Propositional knowledge is descriptive or declarative knowledge; the ideas and facts that we hold to be “true,” or more often described in teacher education as “research-based.” Strategic knowledge is how you know what to do in a situation; knowledge that supports the actions you choose such as contextual knowledge and knowledge from experience. Shulman described *case knowledge* as “knowledge of specific, well-documented, and richly described events... [in which] the knowledge they represent is what makes them cases” (p. 11). Shulman spoke to both the substance of the cases and the pedagogical function that a case can serve by describing three types of cases: prototypes that exemplify theoretical principles; precedent cases that capture and communicate principles of practice or maxims; and parables that convey norms or values.

What we have seen in teacher education since Shulman’s original proposal (1986) are discussions and debates about a “knowledge base” for teaching that has been organized around paradigmatic content themes such as disciplinary subject matter knowledge, pedagogical knowledge, and knowledge of learning and child development (see, for example, Council of Chief State School Officers, 2011). Shulman’s work itself led teacher education down this path with his 1987 article *Knowledge and Teaching: Foundations of the New Reform*. Identification of a knowledge-base for teaching has grown into the global phenomenon of professional teaching standards, with many jurisdictions around the world having constructed their frameworks for what teachers should know through the framing of professional standards for teachers (Darling-Hammond et al., 2017; Sato and Kemper, 2017).

There have been some attempts to disrupt the idea of a knowledge base that is thematically organized into discrete domains. Hugh Sockett (1987) wrote a rebuttal to Shulman’s knowledge-base argument and asked, “has Shulman got the strategy right?” Sockett posed his questions from a moral philosophy lens and asked, if we use a metaphor of a knowledge base, and we rely on the technical language stemming from the field of psychology, might we be putting the moral dimensions of teaching out of view? More recently, Sinnema et al. (2017) posed a new framework for professional standards for teaching based in a framework of professional inquiry that is a deeply embedded concept in teaching practice in New Zealand where the authors are located. These authors attempt to disrupt the commonly accepted approach to constructing standards as lists of knowledge domains and indicators that appear discrete in how they are described and assessed. They outline an integrated model that foregrounds the active and complex nature of teaching and teacher learning through inquiry as it is experienced in practice and often through the lens of values, ethical considerations, and dispositional considerations.

Shulman and Wilson (2004) argued that a vital source of legitimate knowledge for teachers, as professionals, comes from the “wisdom of the practice itself” (pp. 232–233). This is to say that the knowledge base for teaching could be viewed less as a collection of things to “know,” and more as the practical understandings and wisdom that result from the “process” of becoming a teacher itself. Shulman (2004) also insisted that stories from teachers that involved pedagogical action, thinking, and reasoning are not isolated tidbits of information collected over the course of an education program or a career, but rather are constitutive of and resultant from the practice of teachers. This suggests that teacher narratives are simultaneously the knowledge base for teaching as well as the process by which a teacher develops competency and fluency in teaching. In other words, a teacher develops the “wisdom of the practice” through practice, and is, at the same time, defining what that wisdom is as they work to actively construct a teacher life (Knaus, 2020).

The epistemology of story suggests that meaning is found in the stories that we tell, the stories that we live, and the stories that we experience—in the particulars of a situation rather than general principles. Bruner (1986) advocated for the storied nature of human life as how humans understand, interpret, and learn from experience by constructing stories and listening to the stories of others. From this narrative stance, knowledge and the meaning of that knowledge is communicated as much through story as through the construction of logical arguments. Sharing knowledge then becomes a process of story construction and storytelling. The power of the story for sharing knowledge is linked to its detailing of events that unfold over time in a particular place (Clanin and Conelly, 2000). Knowledge as story engages our attention, captures our interest, allows us to relate to other human experience in intimate ways, and has the potential to create a connection that helps us retain the story in our memories. However, sense-making from stories can be challenging in academic spaces. We often revert to propositional knowledge and interpret the

story through the lens of accepted theoretical perspectives, categorical coding, and technical definitions. Once we make this interpretive move—transforming a story from its whole form to a set of categories or lessons to be learned—we run the risk of decontextualizing the knowledge that is bound up in the story through the very nature of the story being the knowledge, not the source of knowledge.

We propose that the field of teacher education could more systematically explore what it might mean if we returned to Shulman's original notion of cases as representing the knowledge of teachers. The literature on case methods in teacher education has primarily (or possibly only) focused on the method as pedagogy—for learning through case construction, examination, and analysis. We are proposing that the identification of cases as a way to structure the knowledge for teaching through narrative forms should be more carefully considered within teacher education, as [Shulman \(1986\)](#) originally proposed.

The knowledge base for teaching as based in this narrative epistemology would not be a monolithic entity. It would be created and re-created in each teacher as they persist in becoming a teacher. This is not to assert, however, that the knowledge that teachers possess and use in their teacher lives ([Knaus, 2020](#)) is unique to them, and different from that of every other teacher. There is a reason why the stories that teachers tell resonate with other teachers. They have shared, common themes and understandings, however colored in ways that are specific to each and the contexts in which they teach. Teachers then find meaning, opportunities for growth and purpose in their own stories, the stories of other teachers, and the connections that they make between them. This conception of teacher knowledge is what [Fenstermacher \(1994\)](#) described as “knowledge of teachers” (p. 11), the practical knowledge that comes from practice. This stands in contrast to what he termed “knowledge for teachers” (p. 9), which is formal “theoretical and propositional knowledge” ([Verloop et al., 2001](#), p. 444) that is scientifically and empirically derived. There is room for both in the broad view of the knowledge base for teaching. But what if we focused on the knowledge base for teaching as understood through the medium and epistemology of narrative, or cases? We might study case methods with questions such as: What do the stories we select as cases illustrate about teaching practice? How do the stories we select help preservice teachers understand the complexity of teaching? Are there iconic cases that sit at the core of our teacher education programs; that convey the values and meaning of learning to teach? Can we represent the complexity of teaching narratively?

Curating a case-based knowledge system for teaching

The idea of curating teachers' stories in the service of constructing teacher knowledge has three elements to it. The first is on a personal level. Humans internally curate our own experiences, both consciously and subconsciously. Preservice and inservice teachers as well as teacher educators have a multitude of stories and pedagogical episodes that we could document and share. We select stories to tell others. Researchers have also identified that teachers have secret stories, lived out in their classrooms and that stay relatively secret within those autonomous spaces ([Craig and Olson, 2002](#)). When these stories are told to other teachers, they are told generally in “safe spaces that they or others have created or found” ([Craig and Olson, 2002](#), p. 25). The second type of curation is interpersonal. This is curation done in conjunction with other people, for example in inquiry groups, learning communities, and amongst professional colleagues. This mutual curation allows one story to connect to and elicit another story. Teacher education programs create these opportunities when they ask their preservice teachers to construct their own cases (e.g., [Rosaen et al., 2009](#); [Tuluze, 2016](#) in the set of studies reviewed in this paper). A caution to note in this interpersonal curation is that when teachers move out of the classroom into more public spaces in their professional knowledge landscapes ([Clandinin and Connelly, 1995](#)), teachers “often live and tell cover stories, stories in which they portray themselves as experts, certain characters whose teacher stories fit within the acceptable range of the story of school being lived in the school” ([Craig and Olson, 2002](#), p. 25).

The final form of curation is formal curation. This is a more structured and “official” form of curation that might occur in a teacher education program or a professional development scheme for teachers. This form of curation is what we see in case books (e.g., the work of July Shulman), collections of video cases of teaching practice (e.g., [Tinker Sachs and Ho, 2007](#) in the set of studies reviewed in this paper), and websites (e.g., [Hatch and Pointer, 2007](#)). Compiling these collections is intense work and they are necessary tools that advance the pedagogical practice of case methods. We propose that research on case methods in teacher education could examine the processes of curating cases with attention to questions of how these cases represent the knowledge base for teaching. Through the act of curating teachers' stories, teacher educators can build on [Shulman's \(1986\)](#) vision for types of case knowledge that serve as prototypes that exemplify theoretical principles; that set precedent by capturing principles of practice or maxims; and that hone norms or values through parable.

The role of a curator is a complex one. Professional curators and docents at art museums select the works and arrange tours of the museum in a way that presents the works as parts of a larger whole, as building blocks for constructing a larger understanding of art and how the viewer constructs meaning. As Lucy Worsley, a professional curator in the United Kingdom explains, “people think that curating just means choosing nice things. But this is only half of it. Our real job as museum curators is to look after artifacts from the past, yet also to be the repository of knowledge about them” (“People Think Curating Just Means Choosing Nice Things”, 2016, emphasis in original). This suggests there is a layer of understanding involved in formal curation of knowledge for teaching through cases. Curating stories is, like teaching, an intensely human and personal endeavor, and the possibility for one's biases to influence the process is high. Similarly, there is a risk of one's own views of worthy cases of becoming the dominant (or only) source of narratives. Who chooses the cases? When preservice teachers construct their own cases, what do the cases represent? A critical lens on questions of curation will be imperative.

Narrative spaces in teacher education

Creating narrative spaces within teacher education programs is necessary if case methods are to be used seriously and to their full potential. The contours of narrative spaces, like most aspects of education, is context-dependent. That is to say that what a narrative space might look like—physically, emotionally, and intellectually—may differ from program to program. More in-depth study of how these spaces function as a narrative space is another aspect of case methods that appears under-represented in the research studies we examined here. These studies might include how cases are embedded programmatically in teacher education; the ways that programs or individual courses do (or do not) promote the telling and hearing of teachers' stories; the purposes for which these stories are shared; how stories compete with other aspects of the program; and how stories are cared for in their whole form as well as the deconstruction that takes place during case analysis.

Other larger contexts and forces that influence learning in teacher education could also be examined through studies of narrative spaces. For example, the examination and understanding of the role that race, language, culture, religion, class, as well as a host of other issues, such as immigration status, first generation college graduate status play in the ways in which teachers' stories are told and heard give meaning to the social construction of knowledge and experience. It is important to note that the telling and hearing of teachers' stories is not done only by teachers, but by the public as well. This suggests that there may be an opportunity to explore cases from media sources, policy discourse, and communities that further illustrate teachers' work and practice outside of the first-hand lived experience.

Conclusion

From the collection of studies in this review, we conclude that case methods in teacher education remain a pedagogical approach that has impact on learning and is still developing in terms of understanding the nuance of the construction of cases for learning. The field of teacher education also seems to struggle with establishing broader programs of research across multiple investigations that build theory or use theory to develop a more robust analysis than is possible with one researcher working in what appears to be isolated contexts.

Schritteser (2014) (one of the studies in this review) provides a conceptual discussion about the use of case study in teacher education, summarizing that case methods should allow teacher education:

- to cultivate a receptive, context-sensitive attitude, and the comprehension skills which are essential to encouraging deep learning in students;
- to foster observational and diagnostic skills which are central to monitoring student progress and potential;
- to learn to analyze the specifics of a situation in order to find out what is the case and why it is an instance of a more general question, which becomes relevant given the dual commitment of teachers both to society and to the individual student;
- to try and bridge the gap between doing practice ones self and reflecting on practice incidents as an exercise in deliberate practice, which in turn is crucial for the growth of expertise in teaching. (pp. 192–193)

These elements of case pedagogy are reminiscent of Shulman's (1987) original arguments for case methods in teacher education and the empirical literature seems to be bearing out these early propositions. We propose, however, that there is room to advance case methods in teacher education to be more akin to Richert's (1991) argument that the descriptive nature of cases "provides an opportunity for presenting teachers' work in all of its complexity" (p. 136). If we accept story as knowledge, then how can teacher education better represent teacher knowledge through these stories? This may be our next big challenge.

References

- Ayyildiz, Y., Tarhan, L., 2012. Effect of case studies on primary school teaching students' attitudes toward chemistry lesson. *Hacettepe Univ. J. Educ.* 43, 62–70.
- Ball, D.L., Cohen, D.K., 1999. Developing practice, developing practitioners: toward a practice-based theory of professional education. In: Darling-Hammond, L., Sykes, G. (Eds.), *Teaching as a Learning Profession*. Jossey-Bass, pp. 3–32.
- Ball, D., Forzani, F., 2009. The work of teaching and the challenge for teacher education. *J. Teach. Educ.* 60, 497–510.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall.
- Bruner, J., 1986. *Actual Minds, Possible Worlds*. Harvard University Press.
- Burns, D.P., Leung, C., Parsons, L., Singh, G., Yeung, B., 2012. Limitations of the case study approach to pedagogical ethics education. *Transformat. Dialog.* 6 (1), 1–10.
- Çevik, Y.D., 2013. How case methods are used to examine and enhance preservice teacher decision-making? *Hacettepe Univ. J. Educ.* (1), 94–108.
- Ching, C.P., 2014. Linking theory to practice: a case-based approach in teacher education. *Proc. Soc. Behav. Sci.* 123, 280–288. <https://doi.org/10.1016/j.sbspro.2014.01.1425>.
- Clandinin, D.J., Connelly, F.M., 1995. *Teachers' Professional Knowledge Landscapes*. Teachers College Press.
- Clandinin, D.J., Connelly, F.M., 2000. *Narrative Inquiry: Experience and Story in Qualitative Research*. Jossey-Bass.
- Council of Chief State School Officers, 2011. *Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards: A Resource for State Dialogue*. Author.
- Craig, C.J., Olson, M.R., 2002. The development of teachers' narrative authority in knowledge communities: a narrative approach to teacher learning. In: Lyons, N., LaBoskey, V.K. (Eds.), *Narrative Inquiry in Practice: Advancing the Knowledge of Teaching*. Teachers College Press, pp. 115–132.
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A.L., Hammerness, K., Low, E.L., McIntyre, A., Sato, M., Zeichner, K., 2017. *Empowered Educators: How High-Performing Systems Shape Teaching Quality Around the World*. Jossey-Bass.
- Dewey, J., 1933. *How We Think*. Heath & Company.

- Dunne, J., 1993. *Back to the Rough Ground: Practical Judgment and the Lure of Technique*. University of Notre Dame Press.
- Doyle, W., 1990. Case methods in the education of teachers. *Teach. Educ. Q.* 17 (1), 7–15.
- Dündar, Ş., 2019. An evaluation of the intentions of preservice teachers to use the case study method in social studies classes. *Issues Educ. Res.* 29 (2), 384–401.
- Fenstermacher, G., 1994. The knower and the known: the nature of knowledge in research on teaching. *Rev. Res. Educ.* 20, 3–56.
- Florez, I.R., 2011. Case-based instruction in early childhood teacher preparation: does it work? *J. Early Child. Teach. Educ.* 32 (2), 118–134. <https://doi.org/10.1080/10901027.2011.572227>.
- Frommelt, M., Hugener, I., Krammer, K., 2019. Fostering teaching-related analytical skills through case-based learning with classroom videos in initial teacher education. *Journal for Educational Research Online* 11 (2), 37–60.
- Gravett, S., de Beer, J., Odendaal-Kroon, R., Merseeth, K.K., 2016. The affordances of case-based teaching for the professional learning of student-teachers. *J. Curric. Stud.* 49 (3), 369–390. <https://doi.org/10.1080/00220272.2016.1149224>.
- Grossman, P., 2005. Research on pedagogical approaches in teacher education. In: Cochran-Smith, M., Zeichner, K. (Eds.), *Review of Research in Teacher Education*. American Educational Research Association, Washington, DC, pp. 425–476.
- Grossman, P., 2011. A framework for teaching practice: a brief history of an idea. *Teach. Coll. Rec.* 113 (12).
- Gunn, A.A., Peterson, B.J., Welsh, J.L., 2015. Designing teaching cases that integrate course content and diversity issues. *Teach. Educ. Q.* 42 (1), 67–81.
- Günther, S.L., Fleige, J., Zu Belzen, A.U., Krüger, D., 2019. Using the case method to foster preservice biology teachers' content knowledge and pedagogical content knowledge related to models and modeling. *J. Sci. Teach. Educ.* 30 (4), 321–343. <https://doi.org/10.1080/1046560x.2018.1560208>.
- Hatch, T., Pointer, D., 2007. *Making Teaching Public: A Digital Exhibition*. Teachers College Record.
- Heffernan, A., Fogarty, R., Sharplin, E., 2016. G'aim'ing to be a rural teacher? Improving pre-service teachers' learning experiences in an online rural and remote teacher preparation course. *Australian and International Journal of Rural Education* 26 (2), 49–61.
- Hemphill, M.A., Richards, K.A., Gaudreault, K.L., Templin, T.J., 2015. Pre-service teacher perspectives of case-based learning in physical education teacher education. *Eur. Phys. Educ. Rev.* 21 (4), 432–450. <https://doi.org/10.1177/1356336x15579402>.
- Howell, P.B., 2012. Conceptualizing developmentally responsive teaching in early field experiences. *Middle Grades Res. J.* 7 (4), 43–55.
- Kale, U., Whitehouse, P., 2012. Structuring video cases to support future teachers' problem solving. *J. Res. Technol. Educ.* 44 (3), 177–204.
- Kindle, K.J., Schmidt, C.M., 2011. Outside in and inside out: using a case study assignment in a reading methods course. *Teach. Educ. Q.* 38 (3), 133–149.
- Knaus, J., 2020. *Teacher Life: A Narrative Inquiry into the Storied Knowledge of Teachers*. Doctoral Dissertation. University of Minnesota.
- Koç, K., 2012. Using a dilemma case in early childhood teacher education: does it promote theory and practice connection? *Educ. Sci. Theor. Pract.* 12 (4), 3153–3163.
- Kurz, T.L., Batarelo, I., 2010. Constructive features of video cases to be used in teacher education. *TechTrends* 54 (5), 46–63.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Liu, W.Y., 2013. Using the case method in Taiwan teacher education: the influence of pre-service teacher's characteristics on case learning. *J. Int. Soc. Teach. Educ.* 17 (1), 159–167.
- Liu, W.Y., 2019. Can we use case method for all preservice teachers? The influence of preservice teacher's characteristics on case learning. *Adv. Soc. Sci. Res. J.* 6 (8), 95–108. <https://doi.org/10.14738/assrj.68.6900>.
- Liu, W.Y., Chen, P., 2019. The use of the case method to promote reflective thinking in teacher education. *Adv. Soc. Sci. Res. J.* 6 (7), 547–557. <https://doi.org/10.14738/assrj.67.6831>.
- Lundeberg, M.A., Levin, B.B., Harrington, H.L., 1999. *Who Learns what from Cases and How? the Research Base for Teaching with Cases*. Lawrence Erlbaum Associates.
- Mayall, H.J., 2010. Integrating video case studies into a literacy methods course. *Int. J. Instr. Media* 37 (1), 33–42.
- McLinden, M., McCall, S., Hinton, D., Weston, A., 2010. Developing authentic online problem-based learning case scenarios for teachers of students with visual impairments in the United Kingdom. *J. Vis. Impair. Blind.* 104 (1), 30–42.
- Merseeth, K.K., 1996. Cases and case methods in teacher education. In: Sikula, J. (Ed.), *Handbook of Research on Teacher Education*, second ed. Macmillan; Association of Teacher Educators, pp. 722–741.
- National Research Council, 1996. *National Science Education Standards*. National Academy Press.
- O'Flaherty, J., McGarr, O., 2013. The use of case-based learning in the development of student teachers' levels of moral reasoning. *Eur. J. Teach. Educ.* 37 (3), 312–330. <https://doi.org/10.1080/02619768.2013.870992>.
- Ohara, Y., Edahiro, K., 2010. Effects of case method in elementary teacher training: approach to authentic classroom practice. *NUE J. Int. Educ. Coop.* 5, 61–66.
- Osmanoğlu, A., Koç, Y., 2013. Investigation of using online video case discussions in teacher education: sources of evidence of mathematics learning. *Educ. Sci. Theor. Pract.* 13 (2), 1295–1303.
- Reid, J., 2011. A practice turn for teacher education? *Asia Pac. J. Teach. Educ.* 39 (4), 293–310. <https://doi.org/10.1080/1359866x.2011.614688>.
- Richards, A.K., Templin, T.J., Eubank, A.M., Hemphill, M.A., 2012. Student-authored case studies as a learning tool in physical education teacher education. *J. Phys. Educ. Recreat. Dance* 83 (3), 47–52. <https://doi.org/10.1080/07303084.2012.10598747>.
- Richert, A.E., 1991. Case methods and teacher education: using cases to teach teacher reflection. In: Tabachnick, B.R., Zeichner, K.M. (Eds.), *Issues and Practices in Inquiry-Oriented Teacher Education*. The Falmer Press, pp. 130–150.
- Rosaen, C.L., Lundeberg, M., Terpstra, M., Cooper, M., Fu, J., Niu, R., 2009. Seeing through a different lens: what do interns learn when they make video cases of their own teaching? *Teach. Educat.* 45 (1), 1–22. <https://doi.org/10.1080/08878730903386849>.
- Saltan, F., Özden, M.Y., Kiraz, E., 2016. Design and development of an online video enhanced case-based learning environment for teacher education. *J. Educ. Pract.* 7 (11), 14–23.
- Sato, M., Kemper, S., 2017. Teacher assessment from pre-service through in-service teaching. In: Clandinin, J., Husu, J. (Eds.), *SAGE Handbook of Research on Teacher Education*. Sage Publications, pp. 944–962.
- Sato, M., Rogers, C., 2010. Case methods in teacher education. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed.
- Sato, M., Sheppard, M., 2008. Bridging the gap between teacher preparation and beginning teacher practice: reviewing the program of research [Paper presentation]. In: *Annual Meeting of the American Educational Research Association*. New York.
- Schritter, I., 2014. The role of case study in teacher education: an attempt to bridge the gap between action and reflection. *Glob. Educ. Rev.* 1 (4), 187–194.
- Shin, S., Brush, T.A., Saye, J.W., 2019. Using technology-enhanced cases in teacher education: an exploratory study in a social studies methods course. *Teach. Teach. Educ.* 78, 151–164. <https://doi.org/10.1016/j.tate.2018.11.018>.
- Schon, D.A., 1987. *Educating the Reflective Practitioner*. Jossey-Bass Publishers.
- Shulman, J.H., 2003. From practice to theory and back again: cases and portfolios as instruments for professional development. *WestEd*.
- Shulman, L., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14.
- Shulman, L., 1987. Knowledge and teaching: Foundations of the new reform. *Harv. Educ. Rev.* 57 (1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>.
- Shulman, L., 2004. Truth and consequences: inquiry and policy in research on teacher education. *J. Teach. Educ.* 53 (3), 248–253.
- Shulman, L., Wilson, S., 2004. *The Wisdom of Practice: Essays on Teaching, Learning, and Learning to Teach*. Jossey-Bass.
- Sinnema, C., Meyer, F., Aitkne, G., 2017. Capturing the complex, situated, and active nature of teaching through inquiry-oriented standards for teaching. *J. Teach. Educ.* 68 (1), 9–27.
- Socket, H., 1987. Further comment: has Shulman got the strategy right? *Harv. Educ. Rev.* 57 (2), 208–220. <https://doi.org/10.17763/haer.57.2.n026767181833w77>.

- Spiro, R.J., Vispoel, W.P., Schmitz, J.G., Samarapungavan, A., Boerger, A.E., 1987. Knowledge acquisition for application: cognitive flexibility and transfer in complex content domains. In: Britton, B.K., Glynn, S.M. (Eds.), *Executive Control Processes in Reading*. Lawrence Erlbaum, pp. 177–199.
- Strike, K.A., Soltis, J.F., 2009. *The Ethics of Teaching*, fifth ed. Teachers College Press.
- Tal, C., 2010. Case studies to deepen understanding and enhance classroom management skills in preschool teacher training. *Early Child. Educ. J.* 38 (2), 143–152. <https://doi.org/10.1007/s10643-010-0395-z>.
- The Guardian, 2016, January 22. "People Think Curating Just Means Choosing Nice Things"—Secrets of the Museum Curators. Available from: <https://www.theguardian.com/culture-professionals-network/2016/jan/22/museum-curator-job-secrets-culture-arts>.
- Tinker Sachs, G.M., Ho, B., 2007. ESL/EFL Cases: Contexts for Teacher Professional Discussion. City University of Hong Kong.
- Tinker Sachs, G.M., Ho, B., 2011. Using cases in EFL/ESL teacher education. *Innovat. Lang. Learn. Teach.* 5 (3), 273–289. <https://doi.org/10.1080/17501229.2010.537755>.
- Tülüçe, H.S., 2016. Using the case story method in a teacher education practicum: affordances and constraints. *Educ. Sci. Theor. Pract.* 16 (4), 1275–1295. <https://doi.org/10.12738/estp.2016.4.0155>.
- Verloop, N., Van Driel, J., Meijer, P., 2001. Teacher knowledge and the knowledge base of teaching. *Int. J. Educ. Res.* 35 (5), 441–461.
- Wagler, R., 2009. Using science teaching case narratives to evaluate the level of acceptance of scientific inquiry teaching in preservice elementary teachers. *J. Sci. Teach. Educ.* 21 (2), 215–226. <https://doi.org/10.1007/s10972-009-9160-9>.
- Zeichner, K., 2012. The turn once again toward practice-based teacher education. *J. Teach. Educ.* 63 (5), 376–382. <https://doi.org/10.1177/0022487112445789>.
- Zottmann, J.M., Goeze, A., Frank, C., Zentner, U., Fischer, F., Schrader, J., 2011. Fostering the analytical competency of pre-service teachers in a computer-supported case-based learning environment: a matter of perspective? *Interact. Learn. Environ.* 20 (6), 513–532. <https://doi.org/10.1080/10494820.2010.539885>.

Contemporary challenges and approaches in anti-racist teacher education

Jenn K. Bergen, Sharissa Unger Hantke, and Verna St. Denis, College of Education, University of Saskatchewan, Saskatoon, SK, Canada

© 2023 Elsevier Ltd. All rights reserved.

What is anti-racist education?	414
Theoretical underpinnings of anti-racist education	415
The Whiteness of teaching and teacher education	415
Experiences of Black, Indigenous, and teacher candidates of color	416
The necessity of anti-racist teacher education	416
Programmatic challenges	416
Lack of adequate anti-racist policies	416
Siloed equity courses and practica structure	417
Resistance by teacher educators	417
Racist discourses and ideologies	417
Dysconscious racism and lack of critical understandings of race	417
Color-evasive racism	417
Meritocracy and democratic racism	418
Deficit discourses	418
White saviorism, false empathy, and uncritical allies	418
Teacher candidate resistance	418
“Good people”, emotionalities, and the maintenance of innocence	418
Denial and defensiveness	419
Distancing strategies	419
Strategic silence	419
Saying the “right thing”	419
Compliance and disrespect	420
Anti-racist approaches in teacher education	420
Creating safe spaces where injustice can be critically examined from multiple perspectives	420
Critical self-reflection, examination, and autobiography	420
Discomfort and processing emotions	420
Investigate the survival strategies of those with lived experiences of racism	421
Counter-narratives	421
Perspective-taking and building empathy	421
Engaging those who experience racism and members of the dominant culture in examining how racism is resisted	421
Media literacy and analysis	421
Analyzing teaching materials	421
Conceptual frameworks and access to new language	421
Facilitating skill development in order to resist oppressive conditions	422
Analysis of classrooms and schools	422
Peer mentorship	422
Challenging racism in all its forms through informed collective action	422
Participatory action research	422
Service learning and social action curriculum projects	422
Summary and recommendations	423
Recommendations for teacher training policy and practice	423
Prioritize Black, Indigenous, and teacher candidates of color	423
Hire and train anti-racist teacher educators	423
Scaffolded equity courses and collective learning	423
References	424

What is anti-racist education?

Anti-racist education confronts and dismantles institutional Whiteness and systemic racism and oppression in education (Dei, 1995, p. 13). It focusses not only on individual discrimination and prejudices, but examines how “racist ideas and individual actions are entrenched and unconsciously supported in institutional structures” (Dei, 1995, p. 13). The work of Barbara Thomas

(1984) identifies five conditions that must be met for socially transformative anti-racist education to take place, which include: (1) Creating safe spaces where injustice can be critically examined from multiple perspectives; (2) Investigating the survival strategies among those with lived experiences of racism; (3) Facilitating opportunities for skill development in order to resist oppressive conditions; (4) Engaging members of the dominant culture and those who experience racism in examining how racism is resisted; and (5) Challenging racism in all its forms through informed collective action (Thomas, 1984 as cited in Fredericks et al., 2001).

Theoretical underpinnings of anti-racist education

Anti-racist education has roots in a number of theoretical traditions, and has evolved from context-specific fields of practice including Culturally Responsive Pedagogy, Ethnic Studies, multicultural education, liberatory education, Critical Race Theory, Critical Whiteness Studies, Intersectionality, and anti-oppressive approaches to education.

Culturally Responsive Pedagogy (CRP) and multicultural education seek to improve the academic performance and curricular inclusion of Indigenous, Black, and students of color, and critical forms of CRP hold overlapping values and strategies with anti-racist education. When done in isolation from specific anti-racist education, however, “cultural” approaches used by White teachers can reassert the very dominance that they seek to address (Gebhard, 2018; St. Denis, 2007). That is, uncritical CRP and multicultural education locate cultural difference as the problem, rather than systemic racism, and their basis for change focuses on techniques alone. These approaches are foundationally different from critical approaches, whose praxis are based in social critique and engagement in informed action. In our review, Critical Race Theory and Critical Whiteness Studies were the most commonly cited theoretical basis for anti-racist education work.

Critical Race Theory (CRT), which has roots in legal studies, has offered anti-racist education several key concepts. The principle of “racial realism” contends that although race is not a biological construct, it is a permanent and enduring feature of societal structures (Bell, 1992). CRT also stresses the necessity of “counter-narratives” and storytelling, as a way of describing the racial discrimination that Indigenous, Black, and people of color experience in their everyday lives, in an effort to make racial biases, discrimination, and unstated normative assumptions more visible (Brown and Jackson, 2013, pp. 19–20).

Ladson-Billings and Tate (1995) describe Critical Race Theory as attending directly to institutional and structural racism in education. Solórzano (1998) further argues that CRT “challenges the dominant discourses on race and racism as they relate to education by examining how educational theory, policy, and practice are used to subordinate certain racial and ethnic groups” (p. 122). By recognizing race as an ideological construct that has every-day consequences for people and groups, and by working to change policies and practices that uphold racism and other interlocking forms of oppression on individual, cultural, and systemic levels, CRT in education is “active”, which is to say that it works through particular practices and has specific outcomes (Delgado and Stefancic, 2001; Ladson-Billings, 1998; Ladson-Billings and Tate, 1995; Omi and Winant, 1994; Solórzano, 1998).

Theorizing about systems of Whiteness, such as through the early works of racial scholars like Du Bois (1903/2017) and Baldwin (1962), has offered further avenues for critical race studies. Harris’ (1993) concept of “Whiteness as property” describes the way that legal systems have been structured to protect property, and that Whiteness allows for greater social and economic security as a result (Brown and Jackson, 2013, pp. 19–20). Mills (1997) contends that White supremacy, as a political system, constitutes a “racial contract”, where dominance and privilege have been so normalized for White people that they do not see this domination as political (pp. 1–2). Goldberg’s (1993) work on historical installation implicates educational institutions, in particular, in the reproduction of racializations and racisms as “common sense” (Goldberg, 1993, p. 12). Critical Whiteness Studies (CWS) builds on these works, including through Frankenberg’s (1993, 2001) feminist theory of Whiteness, Leonardo’s (2004, 2009) theory of Whiteness as ideology, and through what are considered “second-wave” practitioners, who have shifted a focus from White privilege to White supremacy (Leonardo, 2013; Jupp and Lensmire, 2016; Mason, 2016).

Frankenberg (2001) describes Whiteness as the relationship of dominance between White people and the Other (racially minoritized groups), enacted through individual, interpersonal, cultural, and institutional levels. Whiteness is at once a structural advantage, a set of cultural practices, and the unmarked, unnamed norm upon which all others are compared and subjected to (Baldwin et al., 2011). Leonardo (2004, 2009) similarly describes Whiteness as an ideology, forged through historical domination, and kept in place through an enduring series of actions, policies, and laws.

Aveling (2006) argues that “the critical examination of Whiteness is part of a larger project of anti-racism because for too long we have focused on the shortcomings perceived in the ‘Other’ and equated race with ‘studying down’ in the power structure” (p. 263). A focus on Whiteness turns the gaze toward those teachers and educators who play a role in sustaining racism (Carr and Lund, 2007; Case and Hemmings, 2005; Evans-Winters and Hoff, 2011; Fine et al., 2004; Frankenberg, 1993). While the first wave of Critical Whiteness Studies focused on White privilege discourses, second-wave CWS focuses on opening up possibilities for White people to become racially conscious, and to learn about White dominance, supremacy, structural injustice, and the socially constructed nature of identities, rather than individual privilege (Jupp and Lensmire, 2016). In this way, CWS asks teacher candidates to understand the discursive nature of their identity construction, and that racial identities are continually performed and can be transformed. Critical Whiteness Studies also seeks to understand White teacher candidates’ resistance to learning about these discursive reproductions, institutionalized racism, and White supremacy, in an attempt to uncover ways of penetrating this resistance in education.

The Whiteness of teaching and teacher education

The homogeneity of teaching forces in settler colonial and “Western” countries, where teachers are overwhelmingly White, reinforces a structural chasm between the experiences and knowledge of Black, Indigenous, and students of color and their White

teachers (see Miller, 2016; Ryan et al., 2009; Sleeter, 2017). Teacher education programs have similar structural problems. Teacher candidates of color report a lack of peers with similar perspectives and experiences (Brown, 2014), sparse minority role models (Szecsi and Spillman, 2012), and a lack of diversity among teacher educators (Jackson, 2015).

The need for diverse instruction and anti-racist education is not only a demographic issue, as the culture of Whiteness also pervades the structure of education programs with discursive, curricular, and pedagogical implications (Brown, 2014; Cheruvu et al., 2015). Anti-racist education must also therefore attend to the structural power dynamics that demographics often represent, in order to reduce the likelihood that White teachers “will endorse a political and social system that perpetuates White nationalism” (Dei, 2001, p. 150). Levine-Rasky (2000a) has argued that “programmes oriented toward preparing educators to effectively teach to diversity are themselves identified as a crucial element in the failure of equity education” (p. 272), and that it is only through an identification of the ways that Whiteness pervades process, policies, and practices that teacher education can begin to challenge and change these structures in education (Levine-Rasky, 2000a).

Experiences of Black, Indigenous, and teacher candidates of color

Schools and teacher education programs are reported to participate in silencing Black, Indigenous, and teacher candidates of color. The Whiteness of teacher education curriculum, the student body, and faculty ensure that the experiences and knowledge of teacher candidates of color are rarely acknowledged, much less affirmed. Programs “do not extend [build] upon life experiences and cultural knowledge bases of teacher candidates of color” (Jackson, 2015, p. 225.), and neglect to prepare these teacher candidates for the realities of a predominantly White profession (Kohli, 2019, p. 48). Subsequently, teachers of color often cite school climate and lack of voice as reasons for leaving the profession (Gist, 2017, p. 951).

The reaffirmation of the dominance of Whiteness occurs in various explicit and implicit ways. For example, students of color are often positioned as objects—but ignored as participants—in multicultural teacher education (Montecinos, 2004). Through “excluding, silencing and ignoring the presence of preservice teachers of color ... teacher education is, paradoxically, securing the norm of Whiteness in teacher preparation” (Montecinos, 2004). Black and teacher candidates of color are also frequently positioned as “newcomers”. This framing means they “must primarily rely on the ‘expertise’ of predominantly White teacher educators to learn how to teach in socially just and culturally relevant ways” (Pham, 2018, p. 55). Despite having experience that would position them as “knowers”, or even experts, professors hold lower expectations of these candidates, and they are often constructed as having a deficit (in particular if they speak another language) (Brown, 2014).

This Whiteness of teacher education courses also serves to deny racism and perpetuate racial stereotypes. African American preservice teachers attending a predominantly White teacher education program, for example, witnessed White classmates who simultaneously dismissed the continued existence of racism in US society, while also expressing stereotypes. The African American teacher candidates felt unfairly burdened by the responsibility to counter these racial stereotypes (Brown, 2014, p. 335). Teacher candidates of color also battle assumptions that they will automatically become culturally responsive or anti-racist teachers because of their race, ethnicity, or culture. Even being asked to speak their voice – sincere as it might be – can tokenize and humiliate teacher candidates who do not offer a critical analysis of their marginalization. The result can be an added level of feeling alienated and inadequate, and causes them “to develop feelings of distrust and alienation toward their teacher education program” (Brown, 2014, p. 335).

Teacher candidates also struggle to embrace their own cultural and linguistic heritage while negotiating the expectations of White teacher education programs (Haddix, 2012, p. 170). Latinx teacher candidates, for example, feel pressure to accommodate White normative behaviors in order to succeed in their program, including acting in ways that are culturally unfamiliar and that feel inauthentic to the participants (Brown, 2014, p. 335). Such conflicting demands on teacher candidates of color “wears them down” (Achinstein and Ogawa, 2011 in Pham, 2018, p. 54), and they are then more likely to become “absorbed in the system through a process in which they internalized institutional values” (Gist, 2017, p. 946).

Given the obstacles faced by Black, Indigenous, and teacher candidates of color, their inclusion in teacher education programs remains tokenistic. While being celebrated and acknowledged for diversifying primarily White programs, teacher candidates themselves perceive that their presence “is only for statistics” and that their well-being is not a concurrent consideration (Jackson, 2015, p. 231).

The necessity of anti-racist teacher education

Programmatic challenges

The manifestation of a culture of Whiteness is observable in teacher education programs through a host of ways including: policy, recruitment and retention of diverse faculty and teacher candidates; through course curricula and sequence; assumed course audiences; the nature and structure of practica and community placements; how equity and methods courses are integrated (or not); and the degree to which teacher educators resist or are willing to engage in anti-racist work in their own courses. All of these routine processes and programs can be examined for how they enact a culture of Whiteness in education.

Lack of adequate anti-racist policies

Ineffective or non-existent anti-racist policy at the program level may work to counteract other equity program goals (Gist, 2017). For example, “invisible practices of faculty that deny or ignore cultural and linguistic differences among teacher candidates may offset the culturally responsive program design efforts” (Gist, 2017, p. 945). Course evaluation policies, for example, can penalize

teacher educators who teach racial equity courses, by receiving negative feedback from candidates who are not willing to learn about racism (Milner and Howard, 2013; Evans-Winters and Hoff, 2011). Programs may also primarily support research that is interested in how White teacher candidates think and develop their pedagogy (Jackson, 2015), while devaluing cultural and linguistic connections of diverse students through a lack of faculty and student diversity (Gist, 2017, p. 946).

Siloed equity courses and practica structure

Equity content is typically siloed into specific, often one-term courses with the unfortunate effect that they may be inadequately supported by managers and colleagues who are dismissive of anti-racism (Jackson, 2015). Typically, equity and diversity courses also assume a White audience, and are not being created with students of color in mind (Gist, 2017; Jackson, 2015; Kohli, 2019). In addition, programs that have not been designed in consultation with the communities that they seek to serve can reinforce unconscious racism on the part of both teacher candidates and teacher educators. Teaching placements, for example, where mentor teachers and student teaching supervisors unconsciously and dysconsciously adopt racist perspectives can perpetuate their erroneous perspectives with teacher candidates, even though the mentors may be committed to working with diverse populations (Cheruvu et al., 2015, p. 257).

Resistance by teacher educators

Teacher educators who are not specifically anti-racist educators may lack anti-racist perspectives or be unsupportive of equity efforts (Chapman, 2011; in Gist, 2017), resulting in the marginalization of equity courses, and the normalization of methods courses as core knowledge of the program. Additionally, teacher educators who lack anti-racist training focus instead on what has become known as “White privilege pedagogy”, which lacks adequate historical, social, and political context (Levine-Rasky, 2000a, p. 274). Lensmire et al. (2013) argue that, “to an amazing degree, the antiracist aspirations and responsibilities of White teacher educators have been concentrated in McIntosh’s discussion of White privilege,” (p. 411) via her infamous “Unpacking the invisible knapsack” article. Leonardo (2004) argues that teacher education should move past White privilege pedagogy, which asks White people to confess their privilege, and (at its worst) allows this confession to stand in for anti-racist action. Instead, teacher education programs need to move toward accountability for participation in unjust and White-dominant structures.

Racist discourses and ideologies

In addition to programmatic challenges, a significant challenge for anti-racist teacher educators are the racist discourses and ideologies that teacher candidates arrive to their teacher training programs with. These interrelated discourses include dysconscious racism (King, 1991), “color-evasiveness” (Bonilla-Silva, 2006), the myth of meritocracy (Schick and St. Denis, 2003), beliefs in the importance of neutrality and objectivity (Ladson-Billings and Tate, 1995), deficit discourses (Valencia, 2010), White saviorism (Cammarota, 2011), and false empathy (Warren and Hotchkins, 2015).

Dysconscious racism and lack of critical understandings of race

Teacher candidates lack racial consciousness, including a failure to recognize Whiteness as a racialized identity, and ignorance of their own racial positioning. King (1991) describes “dysconscious racism” as the process where White people tacitly or unknowingly accept dominant White norms and privileges, and do not think critically about racial inequity, aside from accepting “certain culturally sanctioned assumptions, myths, and beliefs that justify the social and economic advantages White people have as a result of subordinating diverse others” (p. 135).

Many teacher candidates are superficially aware of race but lack a critical understanding of race and racism, may lack an awareness of the pervasiveness of racism in educational systems, or are unable to understand race in historical context. This can include being unaware of the pervasiveness of racism in their own country (Solomon and Daniel, 2007), or a lack of socio-historical knowledge regarding the realities caused by race toward Indigenous Peoples in their region (Cherubini, 2011). A lack of socio-historical context for race also produces teachers that are unprepared to teach students of color due to a lack of training and knowledge of “theories of race, racism, and inequities that recognize the historical legacy of Whiteness as property” (Annamma, 2015, p. 310).

Color-evasive racism

Color-evasive discourses, formerly known as ‘color-blind’ discourses, are one of the most well-documented challenges in anti-racist teacher education (Bonilla-Silva, 2006, 2018; Ladson-Billings and Tate, 1995; Ledesma and Calderón, 2015). Color-evasive racism includes the assertion that teachers “don’t see color”, and therefore promote equal opportunity and reward individual merit and hard work (Ullucci and Battey, 2011). Color-evasive discourses ignore structural and institutional racism, and construct racism as having occurred only in the past. These discourses also position race as innate or biological rather than socially constructed, and make use of culturally-based stereotypes to justify racist beliefs about groups of people, which can lead to believing that racial inequality is natural, or “just the way things are” (Bonilla-Silva, 2006).

Teachers that hold color-evasive views focus on individual acts of racism rather than structural racism, and locate the problem of racism as belonging to the individual only. When students or parents try to point out structural racism, these claims are constructed as pulling the “race card” (McCreary, 2011). Teachers simultaneously displace the source of racism, by claiming it comes from the community or home, rather than from societal or school structures (McCreary, 2011). Some teachers are also afraid to discuss race and treat Whiteness as the assumed norm (Annamma, 2015), or claim that they will “treat all of their students the same”, in an effort to construct themselves as fair and non-discriminatory (Flintoff, 2018; Schick and St. Denis, 2003). Teacher candidates also may

believe that acknowledging the racial background of their students would make the candidates racist, and that they should focus on teaching their subject matter only, because race is irrelevant (Milner and Laughter, 2015).

Meritocracy and democratic racism

Meritocratic beliefs are closely linked to color-evasive racism. The ideological assumption behind meritocracy holds that “everyone has equal opportunity because we are all basically the same; all that is required to get ahead is hard work, talent, and effort” (Schick and St. Denis, 2003), which ignores the many ways that institutional racism counteracts individual efforts. Meritocratic and individualist discourses also assert that democratic institutions are neutral and fundamentally “fair”, and as such, will be able to mitigate all forms of oppression, including racism (Henry and Tator, 1994). A belief in the democratic values of justice and equality, while disavowing that educational systems have been designed to support White students, results in what has been called “democratic racism” (Henry and Tator, 1994).

Teacher candidates’ beliefs in meritocracy are grounded in entrenched ideas about individual merit and “liberalist notions of individualism” (Solomon et al., 2005, p. 153), asserting that racism resides in individual students and is not embedded in educational institutions (Wilson and Kumar, 2017). This belief then locates failure on the part of minoritized students as being an individual failure. In addition, even though teacher candidates locate the problem of racism in the individual, they resist individual self-implication by embodying a “removedness” from racism (Wilson and Kumar, 2017). Many teacher candidates, when asked to define social justice, also use language that suggests meritocratic beliefs, such as “treating everyone fairly” and references to equality (but not equity) (Bergen, 2020).

Deficit discourses

Deficit thinking is an additional long-held, racist, ideological discourse that locates school failure in the internal deficits or deficiencies of individual students (usually racially minoritized students and those of low-socioeconomic status), rather than looking at the systemic factors that produce school failure (Valencia, 2010, pp. 6–7). Perceived deficits such as limited intellectual abilities, linguistic ability, lack of motivation, and bad behavior are attributed to students’ genetics, culture, class, and familial socialization, rather than to the constructed superiority of the educator and school systems (Valencia, 2010, p. 7). These assumptions serve to blame the victim for their own failings, reinforce negative biases, result in low expectations of students, and create an additional form of oppression that morphs depending on the context (Valencia, 2010). For example, Dunne et al. (2018) found that various deficit discourses about students were prevalent among the majority of 250 White teacher candidates that they surveyed. Gebhard (2018), also found that teachers held low expectations specific to their First Nations, Métis, and Inuit students, and constructed these Indigenous students as a “problem”.

Teacher candidates of color are also constructed as having deficits by classmates and professors. Professors refer to additional language learner backgrounds as possible “barriers”, and teacher candidates of color felt that they needed to work twice as hard to prove their competence (Szecsi and Spillman, 2012, pp. 26–27). Teacher candidates of color also feel they need to defend against deficit, stereotypical beliefs held about them (Brown, 2014, p. 337), experienced low expectations from faculty, staff, and White peers (Brown, 2014, p. 334), and received negative reactions from parents about their accents (Szecsi and Spillman, 2012, p. 26).

White saviorism, false empathy, and uncritical allies

White saviorism and false empathy represent unique racist discursive challenges in teacher education. White saviorism, often grounded in narcissism or an elevated sense of self-importance, describes White teachers who understand themselves as caring or altruistic, while not experiencing or understanding the risks present in this perspective for their students of color (Matias, 2016). Other characteristics of White saviorism can include focusing on “saving” or “helping” students rather than transforming or addressing oppressive structures, avoiding understanding patterns of injustice (espousing a color-evasive view), not recognizing the need for collective action, and focusing on good deeds rather than underlying conditions of oppression (Cammarrata, 2011). White teacher candidates may assume a White savior position by expressing emotions such as care or love for their students, which in reality serve to mask hidden expressions of disgust for the Other (Matias and Zembylas, 2014). Similar rhetorics of “care” have also been shown to conceal deficit discourses, pathologize students, and construct teachers as victims (Toshalis, 2012).

False empathy or false generosity describes teachers who “think that they are building cross racial relationships with others, but fail to negotiate the terms of that relationship with the person they hope to be in relationship with, [and make] unsubstantiated assumptions about what that individual needs” (Warren and Hotchkins, 2015, p. 270). False empathy produces teachers who believe that they are allies, or possess more empathy than is confirmed by the supposed beneficiaries of their actions, but lead to “a failure to recognize, as a person with power and privilege, the multiple ways [they] subordinate others” (p. 269). Thompson (2008) describes this identity as the “lone White hero” self-construction, where teachers see themselves as the rare White person who “gets” it, and who contrasts themselves with less enlightened White people. Rather than examining their own complicity in racism, these teachers may engage in surveillance of others, burden Indigenous people and people of color with requests for affirmation, and demand that their own anti-racist activism be acknowledged rather than collaborating on collective and collegial responses to racism.

Teacher candidate resistance

“Good people”, emotionalities, and the maintenance of innocence

Teacher candidates do not want to be constructed as racist or “part of the problem” (Leonardo and Zembylas, 2013). Faced with anti-racist pedagogy, Schick (2000) found that teacher candidates’ “overwhelming desires for their productions of ‘good’ and

‘innocent’ identities ... is at odds with their desires not to know, with their will to ignorance” (p. 86), and that teacher candidates will employ various strategies to maintain this “goodness”, “innocence” and “respectability”. McIntyre (1997) terms these strategies “White talk”, which include “the infinite number of ways we manage to ‘talk ourselves out of’ being responsible for racism” and “serves to insulate White people from examining their/our individual and collective role(s) in the perpetuation of racism” (p. 45). Characteristics of White talk and moves to innocence include denials and defensiveness, evading or being unwilling to talk about racism, remaining silent or emotionally divesting from discussions about race and racism, distancing oneself from racist White people, and “colluding with each other in creating a ‘culture of niceness’” (p. 46).

Many of these strategies are the result of, or accompanied by, White emotionalities that manifest in different ways depending on the context. Typically these emotionalities show up when deconstructing race, racism, and bias. These can include feelings of distress, guilt, shame, discomfort, uncertainty, helplessness and self-victimization (Clark and Zygmunt, 2014; Matias, 2016). Of these, White guilt is particularly powerful (Matias et al., 2014), as it has immobilizing effects on teacher candidates. These emotions are also often given more space and attention than the emotionalities of BIPOC, and whose emotions teacher education programs cater to therefore becomes political (Matias, 2016).

The strategies that teacher candidates use to maintain the image of themselves as an innocent, good people are discussed individually next, but in practice these will often occur simultaneously. For example, Yu (2012) found that White teachers in a graduate level critical multicultural education class felt uncomfortable about engaging in race talk, responded defensively, and acted as though their credentials or teaching abilities were under attack. Their responses focused on claiming that they already knew “how” to teach, and that they had “love” for their students - strategies that serve to maintain their sense of goodness without examining their own complicity in perpetuating racism.

Denial and defensiveness

Teacher candidates’ denials can include disregarding or disbelieving evidence that indicates that White teachers show a preference for White students (Clark and Zygmunt, 2014), denials of White privilege or structural racism (Solomon et al., 2005), negating their White capital (Solomon et al., 2005), and denying “the role that Whiteness plays in shaping White educators’ lives” (Higgins et al., 2015, p. 251). White teacher candidates also display defensiveness (Solomon and Daniel, 2007; Yu, 2012), don’t feel that they should be responsible for incorporating Indigenous content into their teaching because they aren’t Indigenous (Higgins et al., 2015), don’t feel that as White people, it is “their place” to teach about race and racism (Bergen, 2020), or assume that they will be at primarily White schools, so learning about race was inapplicable to them (Nash, 2013). Some White teacher candidates may also use self-victimization as a way of turning the focus to how they have also been harmed by Whiteness (Matias, 2016).

Distancing strategies

Teacher candidates also seek to create distance between themselves and racism and the responsibility for addressing racism. For some, this might mean distancing themselves from overtly racist White people (Higgins et al., 2015; Schick, 2000), or constructing social distance from racism through time or place (Levine-Rasky, 2000b). Distancing strategies might also involve using a combination of silence and social disassociation in order to distance themselves from responsibility (Case and Hemmings, 2005), or being conflicted about their Whiteness and significant critiques of Whiteness (“White anxiety”) (Levine-Rasky, 2000b). Teacher candidates may also use “strategic ignorance” to distance themselves from the responsibility to meaningfully engage in critical education about treaties (Tupper, 2011), or perform helplessness in order to evade responsibility for figuring out anti-racist actions for themselves (Matias, 2016). In order to avoid feeling uncomfortable, teacher candidates may also try to shift conversations about race and racism to other axes of experience, such as gender or culture (Solomon et al., 2005), or to their subject area training, such as claims that teaching about racism takes away from teaching about literacy (Nash, 2013).

Strategic silence

Both teachers and teacher candidates will also use silence as an avoidance strategy. This can be exemplified by being unwilling to talk about race and racism (Gebhard, 2018), a desire to avoid being uncomfortable (Yu, 2012), withdrawing from meaningful conversations about race (Case and Hemmings, 2005; Dunne et al., 2018; Evans-Winters and Hoff, 2011) or emotionally disinvesting from these conversations (Matias et al., 2014), and not engaging in online discussions in order to resist learning information that challenges their worldviews (Lewis Grant et al., 2018). Teacher candidates also use silence or the omission of ideas by using phrases like “you know”, which enables them to make their views about race and racism known without having to feel uncomfortable by naming specifics. In this way, teacher candidates can maintain a sense of safety while expressing ideas they believe non-White groups may find problematic (Fasching-Varner, 2013).

Saying the “right thing”

Due to fears of being called racist, teacher candidates also use coded language, or want to know how to say the “right thing” when talking about race and racism. For example, teacher candidates wanted to be “equipped” with the ability to teach about race and were uncomfortable not knowing the “right” terminology or language to use (Dunne et al., 2018; Schick, 2000). Teacher candidates also use coded language in order to avoid using the wrong language, such as the use of “diversity discourses”: “while they often openly reject colorblindness, words like ‘urban’ and ‘inner-city’ stand in a coded language to signify racial categories without actually talking about race” (Mason, 2016, p. 1046). Bergen (2020) found that White teacher candidates also drew on these kinds of diversity discourse and other forms of “social justice” language in order to mask a lack of critical racial analysis. In part, this may be due to what Solomon et al. (2005) term “ideological incongruence”, where teacher candidates profess to be open to anti-racist education, but locate “the

problem of racism as a lack of knowledge about other cultures, thereby removing the need for an analysis of issues of power, dominance, historical colonization and oppression fostered primarily by Whites against other groups" (p. 158).

Compliance and disrespect

Finally, some teacher candidate resistance shows up in the form of compliance or outright disrespect for the course professors. Wilson and Kumar (2017), for example, found teacher candidates who were compliant in completing their equity course material because they knew they needed to pass the course. Similarly, teacher candidates can also perform "simulated tolerance" in class, but then use course evaluations to lash out against their course professors as a result of their discomfort with having to confront ideas about White supremacy (Evans-Winters and Hoff, 2011). Finally, some teacher candidates may also show outright and blatant disrespect toward their course professors and colleagues (Matias, 2016).

Anti-racist approaches in teacher education

Anti-racism in teacher education challenges the discourses and ideologies that reinforce Whiteness and racial discrimination in educational structures, including those that maintain other forms of discrimination, based on class, gender, sexuality, culture, and ability. Young (1995) positions anti-racist education as fundamentally different from other forms of education, as it requires an understanding that teaching is a political activity, rather than a vocational one (p. 61). Teacher candidates need anti-racist education to enable them to be agents of social justice so that all students can be school achievers. All teachers need to understand that schooling is a political activity, and that with knowledge of anti-racist education, they and their students can learn how and why they may participate in a more equitable society.

There are no universal approaches to anti-racism in teacher education. Seeking to identify and address multiple, context-specific racisms and forms of oppression will look differently depending on community location, program structure, professors, and students. In addition, some of the discourses and forms of resistance reviewed above can be more easily countered with knowledge-building and analysis than those that pose a fundamental threat to a teacher candidates' sense of self. Laughter (2011) asserts that White preservice teachers, in particular, may require a variety of approaches to engage in conversations about racism and anti-racism.

The next section outlines course-based and integrated or field-experience approaches to anti-racist education with teacher candidates (approaches to program reform are discussed in the recommendations section). Approaches are organized using an adapted form of Thomas' (1984) five conditions for anti-racist education. It is important to note that the boundaries between the categories overlap substantially, and that anti-racist education fundamentally requires a life-long commitment to un/learning and strategic action informed by Black, Indigenous, and people of color.

Creating safe spaces where injustice can be critically examined from multiple perspectives

Immersing teacher candidates in anti-racist literature does not adequately equip them to engage in the critical self-analytical reflection necessary for gaining anti-racist competencies (Cochran-Smith, 2000; Cheruvu et al., 2015; Yu, 2012). To become racially conscious and critically examine understandings of social difference and oppression, traditional lectures are insufficient, and teacher educators must call upon a variety of pedagogical practices to encourage active learning (Adams et al., 1997). White teacher candidates engaged in such reflection can challenge notions of their own White neutrality (Ullucci and Battey, 2011), ask critical questions about the process of their own racialization (Pollock, 2008), and develop an analysis of their intersectional identities and the consequent power-related implications (Schick and St. Denis, 2005).

Critical self-reflection, examination, and autobiography

Critical self-reflection can be initiated through storytelling and autobiography exercises (Bell, 2010; Yu, 2012), provided that they help teacher candidates connect these stories to structural thinking (Crowley and Smith, 2015). Schick and St. Denis (2005), for example, engaged teacher candidates in "rewriting" national and personal narratives regarding their social location based on various identities (race, gender, class), and to reflect on the ongoing implications of these identities within systems of power. Lund and Carr (2018) have found that engaging teacher candidates in identifying and explaining the context of experiences that have shaped who they are can provide inroads into talking about race and racism. Pollock (2008) describes using a similar approach that analyzes "opportunity denial", through asking critical questions about the racialized way that people have experienced the world, and describing what accounts for these differences. Pennington and Brock (2012) found autoethnography to be a useful tool to engage White students in moving beyond confessional tales, which assuage White guilt, toward developing the capacity to critically reflect on their own racial experiences in daily teaching practices.

Discomfort and processing emotions

White emotionalities outlined earlier, such as anger, shame, fear, guilt, and emotional distancing hinder teacher candidates from deeper emotional investment in anti-racism (Matias and Zembylas, 2014). The classroom can be a place where teacher candidates can reflect, process emotions, and work on overcoming deficit orientations and other racist discourses (Bialka, 2015). Teacher candidates who are actively engaged in a "pedagogy of discomfort" (Zembylas and Boler, 2002) are better positioned to understand their

personal investments in racism. A pedagogy of discomfort “requires that individuals step outside of their comfort zones and recognize what and how one has been taught to see (or not to see)” (Zembylas and Boler, 2002, p. 5). Through creating a space for discomfort while engaging in forms of critical self-examination, teacher candidates can begin to problematize “individual and collective emotions, histories, and sense of self” (p. 6). This process involves providing teacher candidates with “alternative affective investments on the basis of solidarity and social justice” (Leonardo and Zembylas 2013, p. 162), in order to cultivate race-conscious and aspirational anti-racist identities, and to help teacher candidates understand that this process is never finished, as it requires identifying and undoing the social, structural, and discursive reproductions of oneself.

Investigate the survival strategies of those with lived experiences of racism

Counter-narratives

Counter-narratives are stories that detail the experiences and perspectives of those who are historically oppressed, excluded, or silenced in educational settings. Reading, listening to, and being immersed in counter-narratives can help equity courses and teacher education programs validate the experiences and perspectives of Black, Indigenous, and people of color (Ullucci and Battey, 2011), and to (re)frame these sources of knowledge as valuable and necessary in classrooms and in the field (Cheruvu et al., 2015, p. 261). In addition, when a majority of teacher candidates are White, counter-narrative perspectives are valuable sources of knowledge, and tools that can disrupt, disturb, and challenge common hegemonic narratives (Cheruvu et al., 2015; Milner and Howard, 2013). Counter-narratives can be used in course readings and viewings, teaching materials, guest speakers, and immersive field experiences.

Perspective-taking and building empathy

Engaging teacher candidates in perspective-taking work, often through the use of counter-narratives, can build empathy and cultivate cultural humility (Warren, 2018). Empathy plays a significant role in preparing teacher candidates to understand their future students, and to build solidarity across race with their students of color (Ullucci and Battey, 2011). For example, DiAngelo and Sensoy (2014) found that using vignettes about minoritized people’s life experiences can allow teacher candidates to analyze situations, unsettle their expectations, and build understanding that can be drawn upon when they are being directly implicated in racism. The “through other eyes” approach developed by Andreotti and de Souza (2008) also includes perspective-taking, by allowing teacher candidates to imagine and encounter systems from other points of view, encouraging both empathy development and critical analysis, and endeavoring to cultivate ethical frameworks for engagement. In order to avoid charity-motivated responses in perspective-taking work, ensure that counter-narratives of Black, Indigenous, and people of color focus on stories of resistance and transformation, rather than stereotypical or “stock stories” (Bell, 2010).

Engaging those who experience racism and members of the dominant culture in examining how racism is resisted

Media literacy and analysis

In order to build critical understandings of how racism is resisted, Bennett et al. (2019) recommend doing media analysis, analysis of statistics, and other forms of research with teacher candidates to inform action opportunities (letter writing or advocacy campaigns, for example). Exercises can include engaging teacher candidates in discourse analysis and the use of anti-racist lenses with news stories, images, and commentary (Lund and Carr, 2018), which can allow them to more readily identify stereotypes, to see how racialized gender roles are reinforced, to understand how the media replicates racist narratives, and to connect these narratives to political parties (Housee, 2012).

Analyzing teaching materials

Racial literacy skills can also be developed through analyzing teaching materials for the “hidden curriculum”, which can include bias, stereotyping, and racism, among other forms of oppression (Tator and Henry, 1991). While stressing the importance of diverse representations in children’s literacy materials is clear (Martin and Spencer, 2020), diverse course materials do not necessarily lead to an anti-racist analysis, and it is important to train teacher candidates how to disrupt uncritical narratives in teaching materials (Beneke, 2017). For example, Mosley and Rogers (2011) engaged White teacher candidates with a framework that analyzed book characters for either passive or active racism or anti-racism, enabling them to identify which teaching materials were more actively anti-racist.

Conceptual frameworks and access to new language

Understandings of systemic racism can be stimulated by equipping teacher candidates with new frameworks to think through systemic problems (Bergen, 2020; Gillies, 2021). For example, teacher candidates who learn to apply a structural determinism conceptual framework are more able to understand racism as systemic. This framework includes looking at how discourses inform policies, practices, and laws, which then dictate outcomes, and subsequently feed back into discourses (Gillies, 2021). By applying this framework to different problems, teacher candidates can understand complex systems that were previously unknown or indescribable to them (Bergen, 2020).

Similarly, teacher candidates need access to new language in order to identify and name injustice. In Smith and Glenn (2019), for example, graduate education students did a two-part reflective assignment, where they wrote an initial reflection about their identity, and then re-visited this reflection at the end of the course using the lenses and frames they had since learned about,

including about power, difference, and cultural identity development. DiAngelo and Sensoy (2014) also use “silence breakers” to create constructive learning environments for White students, by equipping students with constructive phrases prior to discussions about race, so that they can articulate claims and questions in ways that open rather than shut down discussions. These can include “I have always heard that ... What is an anti-racist perspective on that?” or “I am having a ‘yeah but,’ moment, can you help me work through it?” (p. 194).

Facilitating skill development in order to resist oppressive conditions

Analysis of classrooms and schools

Bennett et al. (2019) found that much of the research with teacher candidates is idealist (focuses on changing individual beliefs) rather than realist (encouraging social action), and teacher education programs must try to bridge this gap. Investigations into one’s own classroom or school can have a sizable impact on how students understand their own capacity for identifying and challenging unjust structures (Pollock, 2008). In one course, teacher candidates were asked to observe the diversity of student identities and power dynamics (including racial and gender dynamics) in their practica classrooms, and to design lessons that would address the dynamics they had observed. During this process, teacher candidates reported being more committed to anti-oppressive approaches in their future teaching (Kelley and Brandes, 2010). Similarly, Pollock et al. (2010) found that teacher candidates who understood the tension between theory and action in anti-racist education were the most committed to serving students of color in their future classrooms. In order to have teacher candidates begin to understand racialized policies and structures in classrooms, Shah and Coles (2020) used a “racial noticing” module in a methods course, which included training students to notice racial identity and positionality, social interactions, and structural practices, artifacts, and norms.

Peer mentorship

Peer mentorship is an important component of anti-racist education for both White students and teacher candidates of color. Participation in collaborative social justice support groups, for example, results in new teachers becoming reflective in their journey, and strengthens commitments to teaching and social justice (Picower, 2011). Peer groups can also be structured as professional learning communities that include partners in teacher education, mentor teachers, and teacher candidates that are committed to disrupting Whiteness in their teaching (Little, 2002 as cited in Wetzel, 2020). Extra-curricular fellowship models that combine peer learning with teacher-community inquiry meetings also help teacher candidates develop racial literacies (Solic and Riley, 2019), and practices of “critical friendship” can provide structure and support for White teacher candidates who are working to develop anti-racist frameworks. The need for peer affinity spaces for teachers of color, in particular, is discussed in the recommendations section.

Challenging racism in all its forms through informed collective action

Participatory action research

Engaging in action research with teacher candidates provides them with experience in ways to challenge injustices in their future teaching. The action research by Anderson et al. (2015) worked with teacher candidates to interrogate the sites of learning about social and environmental justice issues in their own program. The process of this project led candidates to consider issues of power and privilege, and to brainstorm solutions together that could be implemented in the following years. This process highlighted that it is necessary for pedagogy to engage both the intellect of students and their emotions. Butler et al. (2017) describe how White teacher candidates were tasked with learning from urban high school students, through a student voice project, by inquiring about students’ learning experiences, and asking what advice they had for teacher candidates. The authors found that, over the course of the year, teacher candidates cultivated deeper relationships with students than those that would normally be constrained by a typical practicum placement. Teacher candidates learned about the positive place-based associations that students had with their school, and this in turn allowed the teacher candidates to develop deeper understandings of themselves and their students.

Service learning and social action curriculum projects

Service learning and social action curriculum projects (SACP) that give teacher candidates the responsibility of defining the “problem” through a critical pedagogical lens can engage them in interrogating their own beliefs (Ponder and Veldt, 2011; Shultz and Baricovich, 2010; Stenhouse and Jarrett, 2012). Stenhouse and Jarrett’s (2012) Problem Solutions Project and Shultz and Baricovich’s (2010) SACP’s saw teacher candidates identify problems as a group, brainstorm solutions, implement action, and write reflections on the process. Projects ranged from a lack of resources and books in elementary schools to hunger rates faced by teacher candidates’ in their own communities. Stenhouse and Jarrett’s (2012) findings revealed that teacher candidates felt a sense of empowerment as a result of being engaged in issues that affected their own realities.

Similar to SACP’s, Ponder and Veldt (2011) revealed that asking teachers to implement service-learning projects in their classrooms enabled them to effectively foster a democratic classroom environment where student-led initiatives could be facilitated, to make connections and build partnerships with communities outside of the classroom, and to reflect on their own growth throughout the process. Field experiences in communities of color can also build capacity for naming racist educational practices and developing a race-conscious repertoire (Ullucci and Battey, 2011; Waddell, 2011), provided that these cultural immersion experiences are designed in conjunction with the communities they serve.

Summary and recommendations

The sections above describe practices and strategies that have been shown to increase racial literacy, racial consciousness, and commitments to anti-oppression on the part of teacher candidates. These goals are achieved through being specific about identifying and disrupting problematic racist discourses, including dysconscious racism, color-evasive racism, meritocracy, deficit thinking, and White saviorism or uncritical allyship. An important component of this process is challenging White teacher neutrality by racializing and dissecting Whiteness, and naming racist educational practices and policies (Ullucci and Battey, 2011). In addition, preparing for resistance from teacher candidates and strong emotional reactions should be built into these processes.

The good practices identified use strategies that allow for critical self-evaluation, and immersion in counter-narratives, to equip teacher candidates with new conceptual frameworks and the ability to identify and name injustices. Without these new frameworks, teacher education programs run the risk of producing students who are confident that they can be anti-racist educators, even though they may lack an accompanying understanding and analysis or continue to hold individualistic views about racism (Bergen, 2020). Importantly, teacher educators must ensure that their courses and materials are being created to validate the experiences and perspectives of Black, Indigenous, and teacher candidates of color (Ullucci and Battey, 2011). Lastly, teacher educators must also work to intentionally equip teacher candidates with skills to challenge and change oppressive structures and not just their beliefs.

Recommendations for teacher training policy and practice

Prioritize Black, Indigenous, and teacher candidates of color

Teacher training must prioritize the recruitment and retention of Black, Indigenous, and teacher candidates of color, and must also ensure that their knowledge and experiences are valued in teacher training (Brown, 2014), by recognizing that teacher training does not occur in a race-neutral context, and by supporting and validating collectivism and collaboration. In part, this can be done by creating race affinity spaces, as “involvement in such networks can be seen as an important factor in a successful teacher education experience, both socially and professionally” (Cheruvu et al., 2015, p. 250). Kohli (2019) recommends that these affinity spaces be legitimized with course credit, as teacher candidates of color may need a reprieve from being minoritized (Kohli, 2019) and an opportunity to do collaborative group work with their own racialized cohort (Tileston and Darling, 2008 in Szecsi and Spillman, 2012, p. 27). Learning with other minority students has also served as a “comfort zone” in otherwise White-dominated spaces (Szecsi and Spillman, 2012, p. 27).

Meaningful engagement with the experiential knowledge of teacher candidates of color serves to recognize them as both learners and experts with unique insights (Pham, 2018). Teacher education can honor the life experiences of students of color while not assuming that they may always recognize or understand how racism manifests in their own experiences (Brown, 2014, p. 336). In order to do this, programs must offer engagement with critical theory and BIPOC theorists (Kohli, 2019), and explicitly equip teacher candidates of color with strategies for understanding the double standards that they will experience in education and as teachers (Gist, 2017, p. 947).

Hire and train anti-racist teacher educators

Many authors have highlighted the necessity of problematizing the racial demography of teacher educators (Milner and Howard, 2013, p. 550), and of increasing the number of BIPOC university faculty, supervisors, and mentor teachers (Cheruvu et al., 2015; Szecsi and Spillman, 2012). Without faculty diversity, the cultural and linguistic diversity of the program structure is at risk (Gist, 2017). Teacher candidates of color often cite a lack of diversity among teacher educators as evidence of their program’s lack of commitment to culturally responsive pedagogy (Jackson, 2015, p. 232). Teacher candidates also appreciate mentors who are minority teachers (Szecsi and Spillman, 2012, p. 27), and value mentor teachers who were from the same community as their students (Pham, 2018). To be effective, both graduate and undergraduate programs should ensure that existing faculty recognize, support, and practice anti-racist teaching.

Scaffolded equity courses and collective learning

Often, equity or foundations courses are expected to cover all the facets of social injustice in a single course (Milner and Howard, 2013). Program claims of infusing these topics throughout a program are often unfounded or exaggerated (Milner and Howard, 2013, p. 553), and seldom do programs include social justice or anti-racist materials in methods courses other than English language arts or social studies (Jackson, 2015, p. 230). In addition, Brown (2014) asserts that “developing teachers that both hold a pedagogic commitment and understand how to effectively enact such socially just and relevant teaching in the classroom takes deliberate, targeted training” (p. 340), and many have argued that teacher education programs need to emphasize a deeper study of race across all courses (Milner and Laughter, 2015).

Scaffolding equity courses throughout a program, similar to the scaffolded nature of many subject area methods courses, can encourage teacher candidates to build their content knowledge and skills over multiple semesters and years, while building relationships with mentors in these areas. Courses that merge equity content and methods, such as “content-agnostic methods course focused on race at the classroom level” (Shah and Coles, 2020, p. 597) can also disrupt the theory/methods dichotomy in

many programs. Finally, Solic and Riley (2019) recommend that teacher education programs should cultivate multidimensional opportunities beyond classroom walls, such as the peer mentorship models discussed earlier, in order to better develop racial literacies.

References

- Achinstein, B., Ogawa, R.T., 2011. *Change(d) agents: New teachers of color in urban schools*. Teachers College Press, New York, NY.
- Adams, M., Bell, L.A., Griffin, P., 1997. *Teaching for Diversity and Social Justice: A Sourcebook*. Routledge.
- Anderson, V., McKenzie, M., Allan, S., et al., 2015. Participatory action research as pedagogy: investigating social and ecological justice learning within a teacher education program. *Teach. Educ.* 26 (2), 179–195.
- Andreotti, V., de Souza, L.M.T.M., 2008. *Learning to Read the World Through Other Eyes*. Global Education, Derby.
- Annamma, S.A., 2015. Whiteness as property: innocence and ability in teacher education. *Urban Rev.* 47 (2), 293–316.
- Aveling, N., 2006. "Hacking at our very roots": rearticulating white racial identity within the context of teacher education. *Race Ethn. Educ.* 9 (3), 261–274.
- Baldwin, A., Cameron, L., Kobayashi, A. (Eds.), 2011. *Rethinking the Great White North: Race, Nature, and the Historical Geographies of Whiteness in Canada*. UBC Press, Vancouver.
- Baldwin, J., November 10, 1962. Letter from a region in my mind. N. York. Retrieved from: <https://www.newyorker.com/magazine/1962/11/17/letter-from-a-region-in-my-mind>.
- Bell, D., 1992. *Faces at the Bottom of the Well: The Permanence of Racism*. Basic Books, New York.
- Bell, L.A., 2010. *Storytelling for Social Justice: Connecting Narrative and the Arts in Anti-Racist Teaching*. Routledge.
- Beneke, M.R., 2017. Race and Ability Talk in Early Childhood: Critical Inquiry Into Shared Book Reading Practices with Pre-Service Teachers. Ph.D., University of Kansas.
- Bennett, J.S., Driver, M.K., Trent, S.C., 2019. Real or ideal? A narrative literature review addressing white privilege in teacher education. *Urban Educ.* 54 (7), 891–918.
- Bergen, J., 2020. *Looking Inward/Looking Outward: Experiences of White Teacher Candidates Encountering Civic Education, Social Justice, and Anti-racist Pedagogy in Two Canadian Teacher Education Programs*. Ph.D. University of Ottawa.
- Bialka, C.S., 2015. Deconstructing dispositions: toward a critical ability theory in teacher education. *Action Teach. Educ.* 37 (2), 138–155.
- Bonilla-Silva, E., 2006. The central frames of color-blind racism. In: *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in the United States*, second ed. Rowman & Littlefield, Maryland.
- Bonilla-Silva, E., 2018. *Racism Without Racists: Colour-Blind Racism and the Persistence of Racial Inequality in America*, fifth ed. Rowman & Littlefield.
- Brown, K., Jackson, D.D., 2013. The history and conceptual elements of critical race theory. In: Lynn, M., Dixon, A. (Eds.), *Handbook of Critical Race Theory in Education*. Routledge, New York, pp. 9–22.
- Brown, K.D., 2014. Teaching in color: a critical race theory in education analysis of the literature on preservice teachers of color and teacher education in the US. *Race Ethn. Educ.* 17 (3), 326–345.
- Butler, J.K., Kane, R.G., Morshead, C.E., 2017. "It's my safe space": student voice, teacher education, and the relational space of an urban high school. *Urban Educ.* 52 (7), 889–916.
- Cammarota, J., 2011. Blind sided by the avatar: White saviors and allies out of Hollywood and in education. *Rev. Educ. Pedagog. Cult. Stud.* 33 (3), 242–259.
- Carr, P.R., Lund, D.E. (Eds.), 2007. *The Great White North? Exploring Whiteness, Privilege and Identity in Education*. Sense Publishers, Rotterdam.
- Case, K.A., Hemmings, A., 2005. Distancing strategies: White women preservice teachers and antiracist curriculum. *Urban Educ.* 40 (6), 606–626.
- Chapman, T.K., 2011. A critical race theory analysis of past and present institutional processes and policies in teacher education. In: Ball, A.F., Tyson, C.A. (Eds.), *Studying Diversity in Teacher Education*. Rowman & Littlefield, New York, NY, pp. 237–256.
- Cherubini, L., 2011. Understanding the marginalized in the mainstream: teacher education and aboriginal educational policy in Ontario (Canada). *Int. J. Educ.* 3 (2), 1–21.
- Cheruvu, R., Souto-Manning, M., Lencl, T., Chin-Calubaquib, M., 2015. Race, isolation, and exclusion: what early childhood teacher educators need to know about the experiences of pre-service teachers of color. *Urban Rev.* 47 (2), 237–265.
- Clark, P., Zygmunt, E., 2014. A close encounter with personal bias: pedagogical implications for teacher education. *J. Negro Educ.* 83 (2), 147–161.
- Cochran-Smith, M., 2000. Blind vision: unlearning racism in teacher education. *Harv. Educ. Rev.* 70 (2), 157–190.
- Crowley, R.M., Smith, W., 2015. Whiteness and social studies teacher education: tensions in the pedagogical task. *Teach. Educ.* 26 (2), 160–178.
- Dei, G.J.S., 1995. Integrative anti-racism: intersection of race, class, and gender. *Race Gender Class* 2 (3), 11–30.
- Dei, G.J.S., 2001. Rescuing theory: anti-racism and inclusive education. *Race Gender Class* 8 (1), 139–161.
- Delgado, R., Stefancic, J., 2001. *Critical Race Theory: An Introduction*. New York University Press, New York.
- DiAngelo, R., Sensoy, Ö., 2014. Calling in: strategies for cultivating humility and critical thinking in antiracism education. *Understand. Dismant. Privil.* 4 (2), 13.
- Du Bois, W.E.B., 1903/2017. *The Souls of Black Folk*. Oxford University Press, Oxford.
- Dunne, L., Kay, V., Boyle, R., Obadan, F., Lander, V., 2018. "I love a curry": student-teacher discourse around "race" and ethnicity at a UK university. *J. Educ. Teach.* 44 (2), 162–174.
- Evans-Winters, V.E., Hoff, P.T., 2011. The aesthetics of White racism in pre-service teacher education: a critical race theory perspective. *Race Ethn. Educ.* 14 (4), 461–479.
- Fasching-Varner, K.J., 2013. "Uhh, you know," don't you?: white racial bonding in the narrative of white pre-service teachers. *J. Educ. Found.* 27 (3/4), 21–41.
- Fine, M., Weis, L., Pruitt, L.P., Burns, A. (Eds.), 2004. *Off White: Readings on Power, Privilege, and Resistance*, second ed. Routledge, New York.
- Flintoff, A., 2018. Diversity, inclusion and (anti) racism in health and physical education: what can a critical whiteness perspective offer? *Curri. Stud. Health Phys. Educ.* 9 (3), 207–219.
- Frankenberg, R., 1993. *White Women, Race Matters*. University of Minnesota Press, Minneapolis.
- Frankenberg, R., 2001. Mirage of an unmarked whiteness. In: Rasmussen, B., Klinenberg, E., Nexica, I., Wray, M. (Eds.), *The Making and Unmaking of Whiteness*. Duke University Press, Durham, pp. 72–96.
- Fredericks, S., Robertson-Lyon, V., Robinson, P., 2001. What are the main features and challenges of anti-racist education? *Quest. Ans. Adult Educ.* <http://fcis.oise.utoronto.ca/~dschugurensky/faqs/qa7.html>. (Accessed 6 April 2021).
- Gebhard, A., 2018. 'Let's make a little drum': limitations and contradictory effects of cultural approaches in Indigenous education. *Race Ethn. Educ.* 21 (6), 757–772.
- Gillies, C., 2021. Curriculum integration and the forgotten indigenous students: reflecting on Métis teachers' experiences. In *Education* 26 (2), 3–23.
- Gist, C.D., 2017. Voices of aspiring teachers of Color: unraveling the double bind in teacher education. *Urban Educ.* 52 (8), 927–956.
- Goldberg, D., 1993. *Racist Culture: Philosophy and the Politics of Meaning*. Blackwell.
- Haddix, M.M., 2012. Talkin' in the company of my sistas: the counterlanguages and deliberate silences of Black female students in teacher education. *Ling. Educ.* 23 (2), 169–181.
- Harris, C., 1993. Whiteness as property. *Harv. Law Rev.* 106 (8), 1707–1791.
- Henry, F., Tator, C., 1994. The ideology of racism: democratic racism. *Can. Ethnic Stud.* 26 (2), 1–14.
- Higgins, M., Madden, B., Korteweg, L., 2015. Witnessing (halted) deconstruction: White teachers' 'perfect stranger' position within urban Indigenous education. *Race Ethn. Educ.* 18 (2), 251–276.
- Housee, S., 2012. What's the point? Anti-racism and students' voices against Islamophobia. *Race Ethn. Educ.* 15 (1), 101–120.

- Jackson, T.O., 2015. Perspectives and insights from preservice teachers of color on developing culturally responsive pedagogy at predominantly White institutions. *Action Teach. Educ.* 37 (3), 223–237.
- Jupp, J.C., Lensmire, T.J., 2016. Second-wave White teacher identity studies: toward complexity and reflexivity in the racial conscientization of White teachers. *Int. J. Qual. Stud. Educ.* 29 (8), 985–988.
- Kelly, D.M., Brandes, G.M., 2010. 'Social justice needs to be everywhere': imagining the future of anti-oppression education in teacher preparation. *Alberta J. Educ. Res.* 56 (4), 388–402.
- King, J.E., 1991. Dysconscious racism: ideology, identity, and the miseducation of teachers. *J. Negro Educ.* 60 (2), 133–146.
- Kohli, R., 2019. Lessons for teacher education: the role of critical professional development in teacher of colour retention. *J. Teach. Educ.* 70 (1), 39–50.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Ladson-Billings, G., 1998. Just what is critical race theory and what's it doing in a nice field like education? *Int. J. Qual. Stud. Educ.* 11 (1), 7–24.
- Laughter, J.C., 2011. Rethinking assumptions of demographic privilege: diversity among White preservice teachers. *Teach. Educ.* 27 (1), 43–50.
- Ledesma, M.C., Calderón, D., 2015. Critical race theory in education: a review of past literature and a look to the future. *Qual. Inq.* 21 (3), 206–222.
- Lensmire, T., McManimon, S., Tierney, J.D., et al., 2013. McIntosh as synecdoche: how teacher education's focus on White privilege undermines antiracism. *Harv. Educ. Rev.* 83 (3), 410–431.
- Leonardo, Z., Zembylas, M., 2013. Whiteness as technology of affect: implications for educational praxis. *Equity Excell. Educ.* 46 (1), 150–165.
- Leonardo, Z., 2004. The color of supremacy: beyond the discourse of 'White privilege'. *Educ. Philos. Theor.* 36 (2), 137–152.
- Leonardo, Z., 2009. *Race, Whiteness, and Education*. Routledge, New York.
- Leonardo, Z., 2013. *Race Frameworks: A Multidimensional Theory of Racism and Education*. Teachers College Press, New York.
- Levine-Rasky, C., 2000a. Framing Whiteness: working through the tensions in introducing Whiteness to educators. *Race Ethn. Educ.* 3 (3), 271–292.
- Levine-Rasky, C., 2000b. The practice of Whiteness among teacher candidates. *Int. Stud. Sociol. Educ.* 10 (3), 263–284.
- Lewis Grant, K.S., Lee, V.J., Lyttle, C.F., 2018. White preservice and inservice teachers' engagement with multicultural content in online courses. *Multicult. Educ.* 26 (1), 17–23.
- Little, J.W., 2002. Professional community and the problem of high school reform. *Int. J. Educ. Res.* 37 (8), 693–714.
- Lund, D.E., Carr, P.R., 2018. Disrupting denial and White privilege in teacher education. In: Lea, V., Lund, D.E., Carr, P.R. (Eds.), *Critical Multicultural Perspectives on Whiteness: Views from the Past and Present*. Peter Lang Publishing, New York, pp. 193–206.
- Martin, A.D., Spencer, T., 2020. Children's literature, culturally responsive teaching, and teacher identity: an action research inquiry in teacher education. *Action Teach. Educ.* 42 (4), 387–404.
- Mason, A.M., 2016. Taking time, breaking codes: moments in White teacher candidates' exploration of racism and teacher identity. *Int. J. Qual. Stud. Educ.* 29 (8), 1045–1058.
- Matias, C.E., Zembylas, M., 2014. 'When saying you care is not really caring': emotions of disgust, Whiteness ideology, and teacher education. *Crit. Stud. Educ.* 55 (3), 319–337.
- Matias, C.E., Viesca, K.M., Garrison-Wade, D.F., Tandon, M., Galindo, R., 2014. "What is critical Whiteness doing in our nice field like critical race theory?" Applying CRT and CWS to understand the White imaginations of White teacher candidates. *Equity Excell. Educ.* 47 (3), 289–304.
- Matias, C.E., 2016. *Feeling White: Whiteness, Emotionality, and Education*. Sense Publishers, Rotterdam.
- McCreary, T., 2011. Colour-blind: discursive repertoires teachers used to stave racism and Aboriginality in urban prairie schools. *Brock Educ. J.* 21 (1).
- McIntyre, A., 1997. *Making Meaning of Whiteness: Exploring Racial Identity With White Teachers*. State University of New York Press, New York.
- Miller, P., 2016. "White sanction", institutional, group and individual interaction in the promotion and progression of Black and minority ethnic academics and teachers in England. *Power Educ.* 8 (3), 205–221.
- Mills, C.W., 1997. *The Racial Contract*. Cornell University Press.
- Milner, H.R., Howard, T.C., 2013. Counter-narrative as method: race, policy and research for teacher education. *Race Ethn. Educ.* 16 (4), 536–561.
- Milner, H.R., Laughter, J.C., 2015. But good intentions are not enough: preparing teachers to center race and poverty. *Urban Rev.* 47 (2), 341–363.
- Montecinos, C., 2004. Paradoxes in multicultural teacher education research: students of color positioned as objects while ignored as subjects. *Int. J. Qual. Stud. Educ.* 17 (2), 167–181. <https://doi.org/10.1080/09518390310001653853>.
- Mosley, M., Rogers, R., 2011. Inhabiting the "tragic gap": pre-service teachers practicing racial literacy. *Teach. Educ.* 22 (3), 303–324.
- Nash, K.T., 2013. Everyone sees color: toward a transformative critical race framework of early literacy teacher education. *J. Transformative Educ.* 11 (3), 151–169.
- Omi, M., Winant, H., 1994. *Racial Formations in the United States*, second ed. Routledge, New York.
- Pennington, J., Brock, C., 2012. Constructing critical autoethnographic self-studies with White educators. *Int. J. Qual. Stud. Educ.* 25, 225–250.
- Pham, J.H., 2018. New programmatic possibilities: (re)positioning preservice teachers of color as experts in their own learning. *Teach. Educ. Q.* 45 (4), 51–71.
- Picower, B., 2011. Learning to teach and teaching to learn: supporting the development of new social justice educators. *Teach. Educ. Q.* 38 (4), 7–24.
- Pollock, M., Deckman, S., Mira, M., Shalaby, C., 2010. "But what can I do?": three necessary tensions in teaching teachers about race. *J. Teach. Educ.* 61 (3), 211–224.
- Pollock, M. (Ed.), 2008. *Everyday Antiracism: Getting Real About Race in School*. The New Press, New York.
- Ponder, J., Veldt, M.V., 2011. Citizenship, curriculum, and critical thinking beyond the four walls of the classroom: linking the academic content with service-learning. *Teach. Educ. Q.* 38 (4), 45–68.
- Ryan, J., Pollock, K., Antonelli, F., 2009. Teacher diversity in Canada: leaky pipelines, bottlenecks, and glass ceilings. *Can. J. Educ.* 32 (3), 591–617.
- Schick, C., St. Denis, V., 2003. What makes anti-racist pedagogy in teacher education difficult? Three popular ideological assumptions. *Alberta J. Educ. Res.* 49.
- Schick, C., St. Denis, V., 2005. Troubling national discourses in anti-racist curricular planning. *Can. J. Educ.* 28 (3), 295.
- Schick, C., 2000. "By virtue of being White": resistance in anti-racist pedagogy. *Race Ethn. Educ.* 3 (1), 83–101.
- Shah, N., Coles, J.A., 2020. Preparing teachers to notice race in classrooms: contextualizing the competencies of preservice teachers with antiracist inclinations. *J. Teach. Educ.* 71 (5), 584–599.
- Shultz, B., Baricovich, J., 2010. Curriculum in the making: theory, practice, and social action curriculum projects. *J. Curriculum Theor.* 26 (2), 46–61.
- Sleeter, C.E., 2017. Critical race theory and the Whiteness of teacher education. *Urban Educ.* 52 (2), 155–169.
- Smith, M.D., Glenn, T.L., 2019. "Why's this man making me feel racist?" Teacher candidates' reflections on diversity course texts. *Teach. Educ. Q.* 46 (2), 30–57.
- Solic, K., Riley, K., 2019. Teacher candidates' experiences taking up issues of race and racism in an urban education fellowship program. *Action Teach. Educ.* 41 (2), 99–116.
- Solomon, R.P., Daniel, B.J., 2007. Discourses of race and 'White privilege' in the next generation of teachers. In: Carr, P.R., Lund, D.E. (Eds.), *The Great White North? Exploring Whiteness, Privilege and Identity in Education*. Sense Publishers, Rotterdam, pp. 161–172.
- Solomon, R.P., Portelli, J.P., Daniel, B.J., Campbell, A., 2005. The discourse of denial: how White teacher candidates construct race, racism and 'White privilege'. *Race Ethn. Educ.* 8 (2), 147–169.
- Solorzano, D.G., 1998. Critical race theory, race and gender microaggressions, and the experience of Chicana and Chicano scholars. *Int. J. Qual. Stud. Educ.* 11 (1), 121–136.
- St. Denis, V., 2007. Aboriginal education and anti-racist education: building alliances across racial identity. *Can. J. Educ.* 30 (4), 1068–1092.
- Stenhouse, V.L., Jarrett, O.S., 2012. In the service of learning and activism: service learning, critical pedagogy, and the Problem Solution Project. *Teach. Educ. Q.* 39, 51–76.
- Szecli, T., Spillman, C., 2012. Unheard voices of minority teacher candidates in a teacher education program. *Multicult. Educ.* 19 (2), 24–29.
- Tator, C., Henry, F., 1991. *Multicultural Education: Translating Policy into Practice*. Multiculturalism and Citizenship Canada, Ottawa.
- Tileston, D., Darling, S., 2008. *Why culture counts: Teaching children of poverty*. Solution Tree, Bloomington, IN.
- Thomas, B., 1984. Principles of anti-racist education. *Currents* 2 (3), 20–24.
- Thompson, A., 2008. Resisting the 'lone hero' stance. In: Pollock, M. (Ed.), *Everyday Anti-racism: Getting Real about Race in School*. The New Press, New York, pp. 328–333.
- Toshalis, E., 2012. The rhetoric of care: preservice teacher discourses that depoliticize, deflect, and deceive. *Urban Rev.* 44 (1), 1–35.

- Tupper, J., 2011. Disrupting ignorance and settler identities: the challenges of preparing beginning teachers for treaty education. *Education* 17 (3), 38–55.
- Ullucci, K., Battey, D., 2011. Exposing color blindness/grounding color consciousness: challenges for teacher education. *Urban Educ.* 46 (6), 1195–1225. <https://doi.org/10.1177/0042085911413150>.
- Valencia, R.R., 2010. *Dismantling Contemporary Deficit Thinking: Educational Thought and Practice*. Routledge, New York.
- Waddell, J., 2011. Crossing borders without leaving town: the impact of cultural immersion on the perceptions of teacher education candidates. *Issues Teach. Educ.* 20 (2), 23–36.
- Warren, C.A., Hotchkins, B.K., 2015. Teacher education and the enduring significance of “false empathy.” *Urban Rev.* 47 (2), 266–292.
- Warren, C.A., 2018. Empathy, teacher dispositions, and preparation for culturally responsive pedagogy. *J. Teach. Educ.* 69 (2), 169–183.
- Wetzel, M.M., 2020. Disrupting race-evasive practices in literacy teacher education: reflections on research and implications for policy. *Lang. Arts* 97 (5), 306–316.
- Wilson, M.B., Kumar, T., 2017. Long ago and far away: preservice teachers' (mis)conceptions surrounding racism. *Int. J. Multicult. Educ.* 19 (2), 182.
- Young, J., 1995. Multicultural and anti-racist teacher education: a comparison of Canadian and British experience in the 1970s and 1980s. In: Ng, R., Staton, P., Scane, J. (Eds.), *Anti-racism, Feminism, and Critical Approaches to Education*. Greenwood Publishing Group, Westport, pp. 45–63.
- Yu, T., 2012. “What is it that I don't know”: learning with White teachers in anti-racist education. *Multicult. Educ.* 19 (4), 1–7.
- Zembylas, M., Boler, M., 2002. On the spirit of patriotism: challenges of a ‘pedagogy of discomfort’. *Teach. Coll. Rec.* 104 (5), 1–27.

Teaching Performance Assessments (TPAs)

Larissa McLean Davies and Michèle Hinton Herrington, Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	427
The policy context of TPAs in Initial Teacher Education programs	427
Perspectives on Teaching Performance Assessments	428
The common design components of TPAs	429
Lesson plans and key instructional resources and materials	430
Videos	430
Work samples from three (3) students	430
Whole class assessment	430
Other general characteristics, processes and practices associated with TPAs	430
TPA differences and distinctions	431
A range of components and cycles/processes are used	431
Moderation and assessment processes are bespoke and contextual	432
The focus and nature of the TPA credential	432
Issues and challenges for TPAs and TPA research	432
Conclusion	433
References	434

Introduction

Teaching Performance Assessments (TPAs) have been part of the initial teacher educational landscape for more than 20 years (Stacey et al., 2020). TPAs are final, summative, portfolio tasks that assemble evidence of the pre-service teachers' ability to plan, enact and reflect on the complexity of teaching for diverse learners. In this way, they are intended to be authentic assessments of teaching (Darling-Hammond and Snyder, 2000; Gallant and Mayer, 2012) and take place in the context of practice. While a final assessment of teaching and a teacher's capacity and "readiness" (Buchanan and Schuck, 2016; Darling-Hammond, 2012a,b) to graduate and enter the classroom is arguably a feature of all pre-service teacher programs, TPAs are generally understood to have a capstone function; this means they are designed as the culmination of pre-service teachers' knowledge and practice gained throughout the Initial Teacher Education (ITE) program they have undertaken. While all ITE program arguably have final teaching assessments, when discussion of TPAs occurs in public and educational discourses, what is generally being referred to are assessment instruments that are common to more than one institution, and that involve moderation and benchmarking practices beyond an individual Initial Teacher Education program. TPAs are considered "high stakes" because teaching graduates' access to professional licensure (e.g., United States of America) or provisional registration (e.g., Australia) is dependent on successful completion of this portfolio task. Drawing on international research and examples, this chapter will explore: the policy context and genesis of Teaching Performance Assessments (TPAs); the key components of the assessment instruments; and issues and challenges for the implementation and research associated with TPAs. The chapter will conclude with considerations for TPA research and practice going forward.

The policy context of TPAs in Initial Teacher Education programs

It has been well documented that international governmental concern with educational outcomes for students and, relatedly, the quality of teaching and learning in schools, has led to decades of reviews into the impact and efficacy of Initial Teacher Educations (ITE) programs (Dinham, 2013; McLean Davies et al., 2017). Indeed, many articles about Initial Teacher Education begin by acknowledging, like this one, that this is an enthusiastically reviewed and increasingly regulated space (Louden, 2008; Mayer, 2014). A key outcome of these reviews has been to introduce Teacher Professional Standards in various national contexts, which are designed to ensure that teachers have the necessary knowledge, skills, and dispositions for impactful teaching as they enter the profession and then progress through their careers. As might be anticipated, there is much scholarly literature on the standards-based reform of teacher education (for example, Allard and Doecke, 2014; Cochran-Smith, 2006; Wyatt-Smith et al., 2017). Teacher education researchers have alternatively welcomed (Shaukat and Chowdhury, 2021) and raised concerns (Allard and Doecke, 2014; Tuinamuna, 2011) about the ways Professional Standards have been developed and implemented within national contexts.

In the US, high-profile teacher education researchers, such as Linda Darling-Hammond, have been integral to educational reforms and have championed the informed use of standards as a means of raising the status of the profession; supporting the

quality of learning for diverse students in different contexts and ensuring transparency and consistency of practice that is common in other professional contexts (Darling-Hammond, 2004, 2015). From an accountability perspective, the development of Standards can also be understood as a response to concerns that the field of Initial Teacher Education is unregulated, parochial, and opaque (Darling-Hammond et al., 2014). In this light, the implementation of Standards has also been critiqued as reducing the complexity of teaching, ignoring, or flattening the disciplinary and contextual nature of teaching; homogenizing teaching and learning (Mills and Goos, 2017) and potentially devaluing teacher judgment (Mayer, 2021).

TPAs are closely linked with the introduction of Standards (AITSL, 2018), and therefore can be understood in the context of a global education reform agenda (Mayer, 2021), and what Stahlberg has labeled the Global Education Reform Movement (GERM) (Sahlberg, 2011), which emphasizes the importance of teachers to student learning and achievement (Hattie, 2012; OECD, 2005, 2009; Wenglinsky, 2002). The link between the development of Teacher Professional Standards and TPAs can be observed in the United States (Darling-Hammond, 2012c); Malaysia (Gallant and Mayer, 2012); and the Australian context (Kriewaldt, 2012; Wyatt-Smith et al., 2017) (see Table 1 below). However, while Teaching Standards are produced by Government agencies and/or education departments (depending on the national context), Teaching Performance Portfolio Assessments are not generally or exclusively determined by governments, but rather are often produced by consortia of Universities or Teaching Colleges, sometimes in collaboration with government (as in the case with Australia—see Kriewaldt, 2021), or industry (as in the case of the EdTPA, Sato, 2014).

Indeed, the development of TPAs across the globe offers evidence of Teacher Education researchers, scholars and leaders responding to in the context of Standards-based reforms by taking responsibility for the ways in which teaching will be ultimately assessed, and relevant national Standards enacted in Initial Teacher Education Programs. This was initially the case with the TPAs in California, where Darling Hammond and the team at Stanford played a key role (Darling-Hammond et al., 2013); in Australia where two consortia consisting of universities developed and piloted TPAs in 2017 (see Buckworth et al., 2018; Kriewaldt et al., 2021; Wyatt Smith et al., 2017), and in Malaysia where a TPA was co-developed with between Malaysian and Australian teacher educators (see Gallant and Mayer, 2012).

Perspectives on Teaching Performance Assessments

Given that TPAs are, in policy terms, aligned with the GERM movement, these high-stakes assessments can be understood as a material effort to regulate a profession that governments contend has relied on anecdotal evidence rather than systematically gathered data to inform decision-making (Cochran-Smith et al., 2013; Gallant and Mayer, 2012). Scholars have also noted that the emphasis on the assessment of the “impact” of pre-service teachers more generally on student learning, and government intervention in this regard, makes the educational problem teachers, and directs attention away from the complexity of education systems and the range of actors who influence student learning (Ell et al., 2019); contributing to what Power has called “audit culture” (Power, 1997), and Ball the “technology of performativity” in education (Ball, 2003). For some, the US edTPA, which is now owned and administered by the multi-national company Pearson, exemplifies these issues with externally and commercially driven Standards-based reform and has thus drawn significant criticism (Choppin and Meuwissen, 2017; Greenblatt and O’Hara, 2015; Hutt et al., 2018; Tuck and Gorlewski, 2016).

Yet, although TPAs are identified as being part of neo-liberal reform, teacher education researchers who have chosen to work with or alongside government and industry to development these assessments have argued that TPAs have significant value, as

Table 1 Examples of interconnections between the development of Teacher Professional Standards and Teacher Performance Assessments.

Country/state	Introduction of standards	Introduction and development of TPA/s
California, North America	Teacher standards for California were developed in 2007 Common Core State Standards (CCSS) adopted in 45 states in 2010.	Since 2008 4x TPAs have been utilized to assess California’s teaching performance expectations Fresno Assessment of Student Teachers (FAST); Performance Assessment for Californian Teachers (PACT), the nationally available edTPA; and the Californian Teaching Performance Assessment (CalTPA).
Malaysia	Teacher standards—2009	Malaysian TPA developed in 2009 and used in three Malaysian teacher training institutes (Gallant and Mayer, 2012)
Australia	2009—work begins on teacher standards. Teacher Professional Standards (teaching and program) introduced in 2011	Program standards revised in 2015 (to include TPA requirement—AITSL, 2015, p. 12) following a Ministerial review of initial teacher education (TEMAG, 2015). Two TPAs were initially developed for national use. The Graduate Teacher Performance Assessment (GTPA) in 2015 and: the Assessment for Graduate Teaching (AfGT, 2018) in 2016 (see summary report 2018).

they authentically capture the work of teachers (Darling-Hammond, 2012a,b) keep responsibility for teaching assessment within the profession—thus arresting the take-over of teaching by external stakeholders (Buchanan et al., 2020)—and provide evidence of how Initial Teacher Education Programs are preparing pre-service teachers to support the learning of diverse students (Darling-Hammond, 2015; Gallant and Mayer, 2012). Indeed, Bourke and others contend that although TPAs are sometimes viewed as a “compliance mechanism” (p. 40) these assessments “would be better understood as an opportunity for assessment innovation, improvement, and effective change” (Bourke, 2019, p. 40; see also Guskey, 2002; Supovitz and Taylor, 2005). To this point, Darling-Hammond has pointed out that when TPAs were developed in California, they replaced mandatory, multiple choice tests which were bespoke across North America. Darling-Hammond argues, following the policy analysis of Edward Crowe, that these multiple-choice tests had no intentional consistency; did not use evidence of practice; and often assessed teacher knowledge prior to entering their Initial Teacher Education Program (Darling-Hammond, 2010). Following the introduction of TPAs in the US context, then, Darling-Hammond asserts that the expectations of pre-service teachers’ assessment are now better aligned with those of in-service teachers/accomplished teachers in the US context, who also seek US National Board Certification via portfolio assessment which included video and lesson plans (rather than a summative test).

Like others who affirm the role and place of TPAs within pre-service programs (Caughlan and Jiang, 2014; Wei and Pecheone, 2010), Darling-Hammond also draws attention to the ways in which these assessments enable pre-service teachers to meaningfully reflect on and integrate theory and practice, and argues that data from these portfolio assessments can be utilized to

both reflect and predict teachers’ success with children so that they can not only inform personnel decisions, but also leverage improvements in preparation, mentoring, and professional development.

Darling-Hammond (2010, p. 1).

Further, those involved in designing and implementing TPAs similarly contend that these assessments, which include multiple components, are better able (than other forms of teacher practice assessment, such as summative tests) to reflect the complexity of teaching, of student learning, and importantly, of teacher judgment and decision making, which is understood as key to teachers’ work and professional development (Kriewaldt, 2012; Ravitch, 2010). In this way TPAs gather more authentic and contextual evidence of teaching and learning that can’t be captured by students’ standardized test scores (Darling-Hammond and Snyder, 2000; Tang, 2008), and usefully inform decisions, planning and assessment when pre-service teachers undertake practicum experience (Gallant and Mayer, 2012).

While reviews of Teacher Education have contributed to the “policy borrowing” (Steiner-Khamsi, 2004) across the globe, of TPAs the movement and influence of significant researchers in the field of Teacher Education have also contributed to the development of TPAs across the globe, and to “policy traveling” across diverse national contexts (Alexiadou and Jones, 2001, p. 2). For example, Professor Diane Mayer (originally from Australia) was involved in the development of PACT; the Malaysian TPA, and the Australian Assessment for Graduate Teaching (AfGT), and a TPA developed for use at Deakin University, also in the Australian context. Mayer’s commitment to TPAs as authentic assessments (Gallant and Mayer, 2012), is evident across these instruments. There are also established links within the US context regarding the development of TPAs. Hébert (2017) notes that high-profile teacher educators and researchers Percheone, Wei and Darling-Hammond all have had key roles in developing the Performance Assessment for Californian Teachers (PACT) and the national edTPA.

The common design components of TPAs

Korthagen and Vasalos note that it is challenging “to make contact with the core qualities which are of importance at that particular moment” (Korthagen and Vasalos, 2005, p. 68) when making decisions about student learning in the context of teaching, and thus a range of assessment components or elements are generally considered necessary in order for TPAs to authentically capture the complex work of teachers as they support student learning. As one might expect, however, there are commonalities among TPA components, as teacher educators and scholars have worked through issues of what should be collected as “evidence” of sound teaching preparation and performance, and how pre-service teachers will be best be supported to meaningfully analyze and reflect on their practice in different national contexts. In their systematic review of international research literature pertaining to Teaching Performance Assessment, Stacey et al. identify that in order to capture the work and complexity of teaching authentically, TPAs commonly reflect the learning and teaching cycle (Stacey et al., 2020) and include: planning and preparation documents and evidence of student learning in context—which often include critical observations of classroom practice (undertaken by a third party, such as a mentor teacher), or facilitated through video recording; samples of student work; and summative assessment data (Stacey et al., 2020, pp. 512–513). Throughout these core components, two integrated aspects of reflection are undertaken: pre-service teachers reflect on student learning, and on their own practices that have supported (or perhaps limited) learning. This capacity for profession reflection is identified as core to TPA models (Gallant and Mayer, 2012; Kriewaldt et al., 2021), and thus aligns TPAs with practitioner research and inquiry approaches and commitments (Cochran-Smith and Lytle, 2009). In this section, we will briefly outline what can be discerned as the common components of TPAs, noting the rationale for the inclusion of each component, and any challenges or considerations related to its use.

Lesson plans and key instructional resources and materials

Lesson preparation and planning is core and the most common component of TPAs (Darling-Hammond, 2012a,b). TPAs generally require PSTs to submit not only a sequence of lesson plans but a range of multimodal resources and materials that support the lesson plans, for example, reading materials, graphic organizers, slides, and supporting documents. Lesson plans and supporting materials offer evidentiary insight into what PSTs want their students to understand and develop in their learning sequence. This component is designed to capture the thinking processes of PSTs and their knowledge of content-specific pedagogy through written reflections, commentary and feedback on planned classroom tasks and the selection and developmentally appropriate content and pedagogy. In addition, some TPAs, such as the CalTPA requires the submission of a written narrative: *Getting to Know Your Students*. Here contextual information about students' assets and needs supports the rationale for developing lesson plans.

Attention has been drawn to the tensions around the time commitment and manageability of "labor-intensive" (Ginsberg and Kingston, 2014, p. 40) texts to support the content materials and resources. Sato (2014) raises another concern that PSTs may be tempted "to make compromises" about what they include in this component to meet the demands of the assessment task. TPA models like the AfGT and the CalTPA have attempted to ameliorate the labor-intensive demands of this component by including response templates.

Videos

Collecting segments of video evidence that records teaching practice offers an effective way for PSTs to review and reflect on teaching performance (for example for the edTPA see Jackson et al., 2018; Noel, 2014; for PACT see Okhremtchouk et al., 2009; and for the AfGT see Kriewaldt et al., 2021). As effective feedback can play a powerful and pivotal role in learning (Hattie and Timperley, 2007; Sadler, 1989), TPAs often require documented feedback from mentor teachers on PST teaching performance. PSTs are encouraged to provide a rationale to justify the recorded teaching moments along with the evaluation criteria, and a summary of student performance (Kriewaldt et al., 2021). Sato (2014) and others has raised concerns that video recording in classrooms have the potential to be problematic in terms of privacy and made complex by the logistics of arranging for someone else to film while the PST delivers the lesson. These issues must be negotiated with the school. Aligning with the Australian Institute of Teaching and School Leadership requirement for Australian TPAs (AITSL), and mindful of concerns around video, the GPTA does not limit the range of ways to "show" teaching practice by stipulating the use of video recording. Alternative ways to demonstrate the current range of teaching performance by a PST can be articulated when completing the GTPA, through deep reflection and observation notes with embedded images (Wyatt-Smith et al., 2020).

Work samples from three (3) students

The collection and analysis of student work is an opportunity to evidence student learning. It is common for TPAs to require PSTs to conduct an "an in-depth analysis of the work of several individual students" (Bunch et al., 2009, p. 106). A comprehensive response to this task reveals to the assessors not only the PST's knowledge and skills associated with creating appropriate assessments but, crucially the thinking processes to analyze and evaluate evidence of student learning. To support PSTs, the CalTPA model provides a template to prompt reflection on how successful the PSTs thought they were in preparing relevant and rigorous content-specific lessons for their focus students (see for example in CalTPA [2018] sample template to prompt reflection on what was learned (https://www.csusm.edu/soe/currentstudents/caltpa_c1_guide_ss.pdf)).

Whole class assessment

As part of their professional preparation, it is important that PSTs develop the ability to analyze whole class or summative assessment data, and this therefore also a common feature of TPAs. The focus of this component is often on the collection, analysis, and evaluation of aggregated assessment data. Reviews of TPA models commonly point to this component as the most challenging aspect of TPAs due to the complex and exacting language and data literacy demands facing PSTs as they engage with this data and explicate their results and findings (Kim, 2019). While challenging, the experience for PSTs is often described as informative and fulfilling and TPA templates are used to prompt and guide PSTs through aspect of the task so that they might more deeply engage with the relationship between whole class data and their teaching practices (Kriewaldt et al., 2021).

Other general characteristics, processes and practices associated with TPAs

Collaboration is a key feature of both the design and implementation of TPAs. As the intention of TPAs is to provide a common assessment of teaching performance that is both relevant to individual teaching contexts, and able to be implemented across diverse contexts, collaboration is required at each stage of TPA design, implementation, and assessment phases. This is often complex and political work, requiring Teacher Educators to broker relationships across institutions, governments, principals and teachers, and education unions, and can result in large consortia of interested parties being established for the purposes of developing the TPA.

In the American context, the Performance Assessment for Californian Teachers (PACT) offers a strong example of collaborative design. PACT was created by a consortium of teacher preparation programs including "30 universities, 1 district internship program, and 1 charter school network" (SCALE) and was led by the Stanford Center for Assessment, Learning and Equity (SCALE). The national EdTPA, drew on PACT (Hébert, 2017) in its early development, and some have referred to it as the "second-generation" PACT (Falk, 2013, p. 1) and like other TPAs, was also developed through a partnership between a university, the Stanford Center for Assessment, Learning and Equity [SCALE], and the American Association of Colleges for Teacher Education (AACTE) (Hébert, 2017;

Wyatt-Smith et al., 2022). Collaboration between institutions is necessary if a TPA is to be genuinely appropriate across contexts, and able to be taken to scale.

Developed some 10 years after PACT, in the Australian context, the AfGT, piloted in 2017, was designed by an initial consortium of 10 Universities nationally. The establishment of the consortium was catalyzed by the Australian Institute for Teaching and School Leaders' invitation to consortia of ITE providers to collectively bid for modest funding to design a TPA that would suit the diverse geographical and social ITE contexts. As well as being designed for diverse geographical and social contexts, the AfGT was designed for pre-service teachers preparing for careers in primary or secondary settings, and in programs delivered face-to-face, online or a mixture of these modes. Clear structures to enable collaborative design, and the input of all teacher educators who would be implementing the TPA (whatever the size of their pre-service teacher cohort), enabled all stakeholders to feel ownership of the process and the assessment instrument being developed (Kriewaldt et al., 2021).

Writing about the introduction of a TPA in the Malaysian context, Gallant and Mayer report that collaboration was evident in the ways in which the range of teacher educators involved in the design of the assessment instrument co-designed and tested the rubrics intended to be used to assess the TPA.

Malaysian educators worked together as a group to refine the TPA rubrics to ensure that they not only judged the work in relation to the Malaysian teacher standards but also reflected the cultural contexts within which the pre-service teachers would work.

Gallant and Mayer (2012, p. 298).

Further, the Malaysian TPA served, like other assessment tasks which are enacted in practice (see for example Kameniar et al., 2017) to serve as a connecting bridge between all stakeholders—student, mentors, or supervising teachers—who were charged with the responsibility of supporting the professional development of the pre-service teachers. Gallant and Mayer note that the Malaysian TPA impacted on professional learning in school contexts, and improved relationships between lecturers, pre-service teachers, and classroom mentors/supervisors (Gallant and Mayer, 2012, pp. 302–3).

TPA differences and distinctions

A range of components and cycles/processes are used

As indicated in Table 2 above, TPAs use some common components, but these are not standardized, mandated, or expected across instruments within or across national contexts. In some TPAs, such as the edTPA, emphasis is placed on detailed reflective

Table 2 Summary of the common TPA components.

<i>Common TPA components</i>	<i>Focus</i>	<i>Examples of TPAs with common component</i>
Planning for teaching and learning Lesson plans Supporting - Reading materials - Graphic organizers - Slides - Content supporting documents Commentary notes Rationale/justifications Supervising teacher notes	Pedagogical content knowledge and skills	California Teacher Performance Assessment (CalTPA) Assessment for Graduate Teaching (AfGT) Graduate Teacher Performance Assessment (GTPA) Education Teacher Performance Assessment (EdTPA) Performance Assessment of California Teachers (PACT)
Observation Video and audio recording Still images Annotations Reflections Mentor teacher's written feedback on video segments	Teaching performance	CalTPA AfGT GTPA (video data optional) PACT EdTPA
Selected Student work samples Collect and analyze work examples from three (3) students that demonstrate a range of achievement. Cognitive commentaries on students' work	Evidence of individual student learning during and after lesson sequence	CalTPA AfGT GTPA edTPA
Analysis of whole class assessment data	Evidence of whole class learning before and after lesson sequence	Cal TPA AfGT GTPA edTPA

commentary which accompanies the pre-service teachers' gathered evidence of planning for learning, teaching and then assessing student learning (Hébert, 2017; SCALE, 2021). Other TPAs, such as the AfGT, include elements such as a situational analysis, which provides pre-service teachers with scenarios in a component called "expanding practice" and requires them to articulate their professional judgment and reasoning in contexts they may not encounter in the course of their practicum experience (Kriewaldt et al., 2021).

While many TPAs use video, as noted above, this is optional in some TPAs (such as the Australian GTA), and the duration of the cycle and structure of the assessment itself depends on the context of the TPA. Although the CalTPA includes two cycles with a focus on content-specific instruction and then assessment (Ervin-Kassab et al., 2021), the Australian GTA and AfGT instruments conceive the assessment in terms of one cycle which includes all elements of planning, teaching, assessing, evaluating, and reflecting. Finally, TPAs have varying degrees of weight within a program. In some programs, TPAs act as the total final assessment of a PST's "readiness for practice," while in other cases, PSTs must also pass their supervising teacher/mentor teacher's individual assessment of their teaching performance, (which is the case for those PSTs undertaking the AfGT <https://education.unimelb.edu.au/research/projects/assessment-for-graduate-teaching-afgt>).

Moderation and assessment processes are bespoke and contextual

While all TPAs involve some form of external moderation, the nature of this varies across national contexts. In the case of PACT, two major dimensions are assessed: the Embedded Signature Strategies which are intended to be formative and are determined by individual institutions, and the Teaching Event/s which are assessed by discipline experts and who are trained to use bespoke assessment rubrics (Hébert, 2017; Wei and Pecheone, 2010). Hébert notes that a major distinction between PACT and edTPA pertains to scoring of materials. While PACT assessors are active in the consortium, assessors for the edTPA consist of current or retired educators who both "have content-specific expertise in the content area [they] are applying to score" and "have worked with or mentored teacher candidates or beginning teachers within the last 5 years" (Pearson Technology, n.d., para. 3 cited in Hébert, 2017, p. 76.) While it has been argued that the physical and contextual differences between assessors and the pre-service teachers producing the edTPA will ensure a robustness of the assessment instrument, and enable the edTPA to be genuinely national (Pearson Education, 2011), some scholars are concerned that this will mean that assessors miss important contextual information about programs and school communities, and that the separation between assessors and those preparing pre-service teachers to undertake the edTPA will impact on the integration of the assessment in the program (Hébert, 2017). In the case of the Australian AfGT, a much smaller TPA, assessment is undertaken locally, by trained assessors connected with the relevant programs, and then these assessments are benchmarked and moderated across the consortium, a process which leads to the refinement of the instrument itself (Kriewaldt et al., 2021; Buckworth et al., 2018).

The focus and nature of the TPA credential

The focus and credentialing offered by the TPA is different, depending on its national context. While the Australian and Malaysian TPAs are designed for general application, and utilized in disciplinary contexts, the North American TPAs assess the suitability of pre-service teacher to be considered "ready for practice" in a particular disciplinary or learning contexts, such as mathematics or English Language, or special education. For example, PACT has been "customized to seventeen different credential areas" (Pecheone and Chung, 2006).

Issues and challenges for TPAs and TPA research

There are a range of challenges facing those assembling consortia and designing, implementing, assessing, and then researching the efficacy of TPAs. One of the main intentions of TPAs is to support pre-service teachers to focus on student learning, and the teaching that supports this, rather than making the *teacher* the unit of analysis. However, Gallant and Mayer note, regarding the Malaysian TPA, that there was a danger of pre-service teachers concentrating on themselves, rather than on student learning:

One key finding from this study is that when pre-service teachers perceived the TPA as explicit documentation of teacher performance, considerations about student learning became peripheral. Data indicated that when pre-service teachers speak only about themselves and their teaching, they essentially separate the teaching act from considerations of student learning. This is worrying since the explicit purpose of the TPA was to assess teacher performance specifically related to student learning and the associated professional decision making (p. 304).

This research highlights the challenge of maintaining the intention of TPAs when the outcome of the assessment will determine the pre-service teacher's entry into the profession. It also draws attention to the challenges of "authenticity" presented by these assessments. Indeed, it might be anticipated that pre-service teachers will not successfully enhance student learning (or at least for all their target students) in a given context, over a series of lessons. While it is "authentic" for this to occur, and appropriate for pre-service teachers to learn and reflect on the complex circumstances that have contributed to this limited student progress in the target area (some of which are likely outside their control), this authenticity may sit in tension with the purpose of TPAs, which is to assess whether or not pre-service teachers are meeting national or state standards (a purpose which requires the

pre-service teacher to demonstrate successful and measurable student learning). So, it is understandably difficult for pre-service teachers to not be “distracted” by the personally and professionally high-stakes nature of the task they are undertaking.

Another challenge relates to the nature of the portfolio assessment itself. It has been argued that TPAs, even those which require use of classroom video as evidence, can be manipulated post the classroom events captured to produce a desirable outcome for the pre-service teacher (Marsh et al., 2010). For this reason, Evans (2011) has suggested that it would be valuable for TPAs to include an oral component, which he recommends would give teacher candidates the opportunity to display knowledge and skills demonstrating three components of professionalism “behavioral, intellectual and attitudinal” (Evans, 2014, p. 189), and to engage in dialog about their practice. Yet, while small institutions may be able to include an oral component in their TPA, this would be difficult to take to scale, and would produce further challenge for efficient cross-institutional moderation practices.

Further and more general challenge regarding the implementation of TPAs is bought about by consistency and resourcing. Despite the intention for TPAs to be relevant across contexts, and robustly moderated and assessed, as we have noted, a range of TPAs and assessment practices exist within national and state contexts in different geographical jurisdictions. While the intention of governments and regulators appears to be for TPAs to be robust and relevant, and an appropriate assessment of standards (DeLuca and Bellara, 2013), with the assessment undertaken by pre-service teachers able to be validated and measured, and applicable across a range of contexts, this requires significant resourcing and support, and can’t always be sustained by smaller initial teacher education providers. In Australia, for example, there has been a policy return to more individual development of final teaching assessments (Sachs, 2005), a practice that inspired the introduction of TPAs in the first place. While the Australian Institute of Teaching and School Leadership has established a separate TPA Expert Advisory Group to determine the appropriateness of TPAs that are submitted by ITE providers in the context of their program accreditation, the fact remains that the intention of introducing TPAs in Australia collaboratively, in order that they might be taken to scale, has had to be revised (AITSL).

Indeed, perhaps the most obvious challenge to success is the time and commitment needed to ensure that TPAs have fidelity across contexts, are robustly moderated, kept up-to-date and relevant to the contexts in which they are being enacted. At a time when teacher education programs worldwide are under intense resourcing pressure (Darling-Hammond and Lieberman, 2013), teacher educators are obliged to find additional time and support to develop TPAs. Again, referring to the Australian example, while the Australian Institute of Teaching and School Leadership provided some funding for those consortia initially successful in tendering to develop TPAs, this resource was in no way commensurate to the time and resource required to establish consortia and collaboratively develop the elements of the assessment instrument.

One significant criticism or challenge for TPAs, relates to the fact that much of the research into efficacy of TPAs, that is cited to encourage initial teacher education providers to adopt particular instruments in their programs, is often undertaken by the same personal involved in designing these tasks. On the one hand, this follows the long tradition of teacher educators as reflective researchers of practice and reflects the complex and institutional nature of research ethics clearance. On the other hand, it risks diminishing the claims that can be made about the impact and utility of TPAs. Hébert (2017) raises concern about the research evidence used to argue for the efficacy of the EdTPA, noting that this research is undertaken by academics central to the design of this task (Hébert, 2017, p. 69). He also notes, however, that a range of other Teacher Educators implementing the edTPA have written about its limitations in terms of ensuring pre-service teachers are appropriately “ready” for practice (citing Ledwell and Oyler, 2016); regarding the ways in which the assessment privileges certain kind of pre-service teachers (citing Greenblatt and O’Hara, 2015); and the manner in which the edTPA determines the content and focus of some ITE programs, obliging teacher educators to proverbially “teach to the test” (citing Dover and Schultz, 2016). In light of this criticism, Hébert asserts that this research, as well as more positive affirmations of the utility and relevance of the assessment, should be transparently available for future users of the TPA concerned (Hébert, 2017). This point draws attention to the somewhat conflicting motivations and imperatives for those responsible for designing and implementing TPA policy, points also acknowledged by Gallant and Mayer writing about the Malaysian TPA (p. 303). While teacher educators, in all instances, are initially involved in the design and implementation process, and are invested in researching their practice, and the broader utility of TPAs (Stacey et al., 2020), many are also implicated in the marketing of the TPA product, and obliged to be so, in order for the resourcing and upkeep of the TPA to be maintained.

Conclusion

Teaching Performance Assessments, for all the challenges of implementation and research, remain a key measure through which pre-service teachers’ ability to support diverse learners, and to assess pre-service teachers’ capacities to move into independent practice. The adoption of and investment in the development of TPAs by teacher education researchers, in collaboration with each other, governments and school-based stakeholders, highlights the ways in which these assessment tasks enable pre-service teachers to demonstrate authentic teaching performance that directs attention toward student learning and growth and away from the teacher as a “performer” (Kriewaldt et al., 2021; Stacey et al., 2020).

As we have outlined in this article, TPAs are not without challenges of implementation, most substantively concerning resourcing for development, assessment moderation and future refinement. If TPAs are ultimately to remain the remit and purview of teacher education providers and their stakeholders—rather than external commercial entities—then resourcing for infrastructure, development, data storage and analysis, will need to be prioritized by governments, and time will need to be allocated for work associated with administering and assessing TPAs within institutions.

While much research into the efficacy of TPAs has been done within specific national contexts (a practice for which teacher education research is often criticized) there seems significant potential for international and large-scale research to be undertaken which would enable initial teacher education providers to gather evidence of best practice regarding TPAs internationally. Post-pandemic, it will be valuable to better understand the factors and circumstances across contexts that are required to support the rigorous implementation of TPAs in ways that will develop pre-service teachers' ability use evidence to shape and respond to the learning needs of diverse students.

At the time of writing this chapter, the Federal government in Australia (where the authors are based), has recently released a further report into Initial Teacher Education, which, building on the previous report of 2015 (TEMAG) confirms the central place of TPAs in a reform and quality improvement agenda. Like many reports, this one directs attention to Initial Teacher Education as the panacea for educational challenges at a system level—it is imperative for scholars to resist and speak-back to assumptions made about initial teacher education and its assessments touted by such reports, and to clearly articulate the scope and purpose of ITE, and TPAs within a broader and complex education system (Ell et al., 2019). Going forward, and in the spirit of seeing ITE as the first step in a teacher's professional journey, it is worth exploring the potential of TPA-like assessment for practicing, in-service teachers. This notion has been taken up in some national contexts, see for example activities associated with the implementation of the Kentucky teacher evaluation standards (Guskey et al., 2010). Previous research undertaken in the Australian context has shown the utility of in-service teachers undertaking similar assessments to pre-service teachers, both for the integration of theory and practice across the career, and to ensure continuity of teacher professional learning (Kameniar et al., 2017; McLean Davies et al., 2017). Finally, while TPAs are, as we have noted, fundamentally tied to national professional teaching standards, further TPA generated research might be directed toward the standards themselves. Teacher education researchers might ask: "In what ways does the TPA being implemented reveal issues for standards and their interpretation?", and "What standards need modification so that teachers can authentically enact practice in social just and differentiated ways?" In this way, teacher educators can leverage their investment in TPAs to influence broader conversations about quality in teacher education and student learning.

References

- AfGT, 2018. Assessment for Graduate Teaching. Development and Implementation of the Assessment for Graduate Teaching Instrument: A Nationally Applicable Assessment of Pre-service Teachers' Performance. Melbourne Graduate School of Education, Melbourne. Available. https://education.unimelb.edu.au/_data/assets/pdf_file/0007/2887783/AfGT-Summary-Report-v1.0-0-.pdf. (Accessed 12 July 2022).
- AITSL, 2015. Australian Institute for Teaching and School Leadership. Initial Teacher Education: Data Report 2015. Education Services Australia, Carlton South.
- AITSL, 2018. Australian Institute for Teaching and School Leadership. Australian Professional Standards for Teachers. Education Services Australia, Carlton South. https://www.aitsl.edu.au/docs/default-source/initial-teacher-education-resources/eag-operational-principles.pdf?sfvrsn=b90cfd3c_12. (Accessed 20 June 2022).
- Alexiadou, N., Jones, K., 2001. Travelling Policy/Local Spaces. Paper to the Congrès Marx International 111 Le Capitale et L'Humanité University of Paris X 26–29 September. University of Paris, Paris.
- Allard, A., Doecke, B., 2014. Professional knowledge and standards-based reforms: learning from the experiences of early career teachers. *Engl. Teach. Pract. Critiq.* 13 (1), 39–54. Available. <http://education.waikato.ac.nz/research/files/etpc/files/2014v13n1art3.pdf>. (Accessed 12 July 2022).
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228. <https://doi.org/10.1080/0268093022000043065>.
- Bourke, T., 2019. The changing face of accreditation for initial teacher education programmes in Australia. In: Gutierrez, A., Fox, J., Alexander, C. (Eds.), *Professionalism and Teacher Education*. Springer, Singapore, pp. 27–45.
- Buchanan, J., Harb, G., Fitzgerald, T., 2020. Implementing a teaching performance assessment: an Australian case study. *Aust. J. Teach. Educ.* 45 (5), 74–90. <https://doi.org/10.14221/ajte.2020v45n5.5>.
- Buchanan, J., Schuck, S., 2016. Preparing a "classroom ready" teacher: the challenge for teacher educators. In: Gibbs, I.S. (Ed.), *Teacher Education: Assessment, Impact, and Social Perspectives*. Nova Science Publishers, Hauppauge, pp. 1–12.
- Buckworth, J., Crane, N., Keany, K., Ziebell, N., Tan, K., 2018. Demonstrating achievement of the graduate teaching standards: building interconnected assessments across the elements of the AfGT. Australian Association for Research in Education Conference in Sydney. Charles Darwin University, Australia. Accessed 29 June. <https://www.aare.edu.au/publications/aare-conference-papers/show/11875/demonstrating-achievement-of-the-graduate-teaching-standards-building-interconnected-assessments-across-the-elements-of-the-afgt>, 2018. (Accessed 29 June 2022).
- Bunch, G.C., Aguirre, J.M., Téllez, K., 2009. Beyond the scores: using candidate responses on high stakes performance assessment to inform teacher preparation for English learners. *Issues Teach. Educ.* 18 (1), 103–128.
- CalTPA, 2018. California Teacher Performance Assessment. California Commission on Teacher Credentialing, US. Accessed 28 June. https://www.csusm.edu/soe/currentstudents/caltpa_c1_guide_ss.pdf.
- Caughlan, S., Jiang, H., 2014. Observation and teacher quality: critical analysis of observational instruments in preservice teacher performance assessment. *J. Teach. Educ.* 65 (5), 375–388. <https://doi.org/10.1177/0022487114541546>.
- Choppin, J., Meuwissen, K., 2017. Threats to validity in the edTPA video component. *Action Teach. Educ.* 39 (1), 39–53. <https://doi.org/10.1080/01626620.2016.1245638>.
- Cochran-Smith, M., 2006. Stayers, Leavers, Lovers, and Dreamers: Why People Teach and Why They Stay. In: Cochran-Smith, M. (Ed.), 2004 Barbara Biber Lecture. Bank Street College of Education Occasional Paper, New York, NY.
- Cochran-Smith, M., Lytle, S.L., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Columbia University: Teachers College Press, New York.
- Cochran-Smith, M., Piazza, P., Power, C., 2013. The politics of accountability: assessing teacher education in the United States. *Educ. Forum* 77 (1), 6–27. <https://doi.org/10.1080/00131725.2013.739015>. Taylor & Francis Group.
- Darling-Hammond, L., 2004. Standards, accountability, and school reform. *Teach. Coll. Rec.* 106 (6), 1047–1085.
- Darling-Hammond, L., 2010. Evaluating Teacher Effectiveness: How Teacher Performance Assessments Can Measure and Improve Teaching. Center for American Progress, Washington DC. <https://files.eric.ed.gov/fulltext/ED535859.pdf>. (Accessed 12 July 2022).
- Darling-Hammond, L., 2012a. Policy frameworks for new assessments. In: *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht, pp. 301–339.
- Darling-Hammond, L., 2012c. Teacher preparation and development in the United States: a changing policy landscape. In: Darling-Hammond, L., Lieberman, A. (Eds.), *Teacher Education around the World. Changing Policies and Practices*. Routledge, London, pp. 130–150.
- Darling-Hammond, L., 2012b. Real Teacher Ed Reform. Inside Higher Ed. Accessed 20 June. <https://www.insidehighered.com/views/2012/08/13/essay-arguesreal-teacher-education-reform-going-led-profession>.

- Darling-Hammond, L., 2015. *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement*. Columbia University: Teachers College Press, New York.
- Darling-Hammond, L., Lieberman, A., 2013. *Teacher Education Around the World: Changing Policies and Practices*. Routledge, London and New York.
- Darling-Hammond, L., Snyder, J., 2000. Authentic assessment of teaching in context. *Teach. Teach. Educ.* 16 (5–6), 523–545. [https://doi.org/10.1016/S0742-051X\(00\)00015-9](https://doi.org/10.1016/S0742-051X(00)00015-9).
- Darling-Hammond, L., Newton, S.P., Wei, R.C., 2013. Developing and assessing beginning teacher effectiveness: the potential of teacher performance assessments. *Educ. Assess. Eval. Account.* 25 (3), 179–204. <https://doi.org/10.1007/s11092-013-9163-0>.
- Darling-Hammond, L., Wilhoit, G., Pittenger, L., 2014. Accountability for college and career readiness: developing a new paradigm. *Educ. Pol. Anal. Arch.* 22 (86), 1. <https://doi.org/10.14507/epaa.v22n86.2014>. https://uknowledge.uky.edu/ncie_facpub/1?utm_source=uknowledge.uky.edu2Fncie_facpub2F1&utm_medium=PDF&utm_campaign=PDFCoverPages.
- DeLuca, C., Bellara, A., 2013. The current state of assessment education: aligning policy, standards, and teacher education curriculum. *J. Teach. Educ.* 64 (4), 356–372. <https://doi.org/10.1177/0022487113488144>.
- Dinham, S., 2013. Connecting clinical teaching practice with instructional leadership. *Aust. J. Educ.* 57 (3), 225–236. <https://doi.org/10.1177/0004944113495503>.
- Dover, A.G., Schultz, B.D., 2016. Troubling the edTPA: illusions of objectivity and rigor. *Educ. Forum* 80 (1), 95–106. <https://doi.org/10.1080/00131725.2015.1102368>. Routledge.
- Ell, F., Simpson, A., Mayer, D., McLean Davies, L., Clinton, J., Dawson, G., 2019. Conceptualising the impact of initial teacher education. *Aust. Educ. Res.* 46 (1), 177–200. <https://doi.org/10.1007/s13384-018-0294-7>.
- Ervin-Kassab, L., Escalante, K., Soodjinda, D., 2021. The start of a conversation with critical friends: can the CalTPA be used as a catalyst for program and professional inquiry? *J. Crit. Issues Educ. Pract.* 11 (1), 1–15.
- Evans, L., 2011. The “shape” of teacher professionalism in England: professional standards, performance management, professional development and the changes proposed in the 2010 White Paper. *Br. Educ. Res. J.* 37 (5), 851–870. <https://doi.org/10.1080/01411926.2011.607231>.
- Evans, L., 2014. Leadership for professional development and learning: enhancing our understanding of how teachers develop. *Camb. J. Educ.* 44 (2), 179–198. <https://doi.org/10.1080/0305764X.2013.860083>.
- Falk, B., 2013. Lessons from the performance assessment of California teachers (PACT). *N. Educ.* 9 (1), 1–2. <https://doi.org/10.1080/1547688X.2013.751308>.
- Gallant, A., Mayer, D., 2012. Teacher performance assessment in teacher education: an example in Malaysia. *J. Educ. Teach.* 38 (3), 295–307. <https://doi.org/10.1080/02607476.2012.668330>.
- Ginsberg, R., Kingston, N., 2014. Caught in a vise: the challenges facing teacher preparation in an era of accountability. *Teach. Coll. Rec.* 116 (1), 1–48 doi:10.1177%2F016146811411600101.
- Greenblatt, D., O'Hara, K.E., 2015. Buyer beware: lessons learned from EdTPA implementation in New York state. *Teach. Educ. Q.* 42 (2), 57–67.
- Guskey, T.R., 2002. Professional development and teacher change. *Teach. Teach. Theor. Pract.* 8, 381–391. <https://doi.org/10.1080/135406002100000512>.
- Guskey, T.R., Swan, G.M., Jung, L.A., 2010. Developing a State Wide Standards-Based Student Report Card: A Review of the Kentucky Initiative. Paper presented at the annual meeting of the American Educational Research Association. American Educational Research Association, Denver, Colorado.
- Hattie, J., 2012. *Visible Learning for Teachers: Maximizing Impact on Learning*. Routledge, London and New York.
- Hattie, J., Temperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77 (1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Hébert, C., 2017. What do we really know about the edTPA? Research, PACT, and packaging a local teacher performance assessment for national use. *Educ. Forum* 81 (1), 68–82. <https://doi.org/10.1080/00131725.2016.1242680>. Routledge.
- Hutt, E.L., Gottlieb, J., Cohen, J.J., 2018. Diffusion in a vacuum: edTPA, legitimacy, and the rhetoric of teacher professionalization. *Teach. Teach. Educ.* 69 (1), 52–61. <https://doi.org/10.1016/j.tate.2017.09.014>.
- Jackson, E.D., Kelsey, K.D., Rice, A.H., 2018. A case study of technology mediated observation in pre-service teaching experiences for edTPA implementation. *NACTA J.* 62 (1), 1–10.
- Kameniar, B., Davies, L.M., Kinsman, J., Reid, C., Tyler, D., Acquaro, D., 2017. Clinical praxis exams: linking academic study with professional practice knowledge. In: Peters, M.A. (Ed.), *A Companion to Research in Teacher Education*. Springer, Singapore, pp. 53–67.
- Kim, N., 2019. Challenges of teaching and preparing edTPA. *Art Educ.* 72 (1), 21–27. <https://doi.org/10.1080/00043125.2019.1534439>.
- Korthagen, F., Vasalos, A., 2005. Levels in reflection: core reflection as a means to enhance professional growth. *Teach. Teach.* 11 (1), 47–71. <https://doi.org/10.1080/1354060042000337093>.
- Kriewaldt, J., 2012. Reorienting teaching standards: learning from lesson study. *Asia Pac. J. Teach. Educ.* 40 (1), 31–41. <https://doi.org/10.1080/1359866X.2011.643761>.
- Kriewaldt, J., Walker, R., Morey, V., Morrison, C., 2021. Activating and reinforcing graduates' capabilities: early lessons learned from a teaching performance assessment. *Aust. Educ. Res.* 48 (4), 681–696.
- Ledwell, K., Oyler, C., 2016. Unstandardized responses to a “standardized” test: the edTPA as gatekeeper and curriculum change agent. *J. Teach. Educ.* 67 (2), 120–134. <https://doi.org/10.1177/0022487115624739>.
- Louden, W., 2008. 101 Damns: the persistence of criticism and the absence of evidence about teacher education in Australia. *Teach. Teach.* 14 (4), 357–368.
- Marsh, B., Mitchell, N., Adamczyk, P., 2010. Interactive video technology: enhancing professional learning in initial teacher education. *Comput. Educ.* 54 (3), 742–748.
- Mayer, D., 2014. Forty years of teacher education in Australia: 1974–2014. *J. Educ. Teach.* 40 (5), 461–473.
- Mayer, D., 2021. Teacher education policy and research: an introduction. In: Mayer, D. (Ed.), *Teacher Education Policy and Research*. Springer, Singapore. https://doi.org/10.1007/978-981-16-3775-9_1.
- McLean Davies, L., Angelico, T., Hadlow, B., Kriewaldt, J., Rickards, F., Thornton, J., Wright, P., 2017. Supporting the development of the profession: the impact of a clinical approach to teacher education. In: Savage, G., Bentley, T. (Eds.), *Educating Australia: Challenges for the Next Decade*. Melbourne University Press, pp. 209–228.
- Mills, M., Goos, M., 2017. The place of research in teacher education? An analysis of the Australian teacher education ministerial advisory group report. “Action now: classroom ready teachers.” In: Peters, M.A., Cowie, B., Menter, I. (Eds.), *A Companion to Research in Teacher Education*. Springer, Singapore, pp. 637–650.
- Noel, A.M., 2014. Teacher performance assessment (edTPA): an instructor's development and evaluation of an embedded signature assessment in an early childhood literacy course. *J. Early Child. Teach. Educ.* 35 (4), 357.
- OECD, 2005. *Education at a Glance 2005: OECD Indicators*. OECD, Paris.
- OECD, 2009. *Education at a Glance 2009: OECD Indicators*. OECD, Paris.
- Okhremchouk, I., Seiki, S., Gilliland, B., Ateh, C., Wallace, M., 2009. Voices of pre-service teachers: perspectives on the performance assessment for California teachers (PACT). *Issues Teach. Educ.* 18 (1), 39–62.
- Pearson Technology, (n.d.-a). edTPA Scorer Qualification. Accessed 20 June <http://scoreedtpa.pearson.com/become-an-edtpa-scorer/edtpa-scorer-qualifications.html>.
- Pearson Education, 2011. Pearson recruits 1,200 educators to score teacher performance assessment. Pearson Education. <http://www.pearsoned.com/news/pearson-recruits-1200-educators-to-score-teacher-performance-assessment/>, 2011. (Accessed 12 July 2022).
- Pecheone, R.L., Chung, R.R., 2006. Evidence in teacher education: the performance assessment for California teachers (PACT). *J. Teach. Educ.* 57 (1), 22–36.
- Power, M., 1997. *The Audit Society: Rituals of Verification*. Oxford University Press, Oxford.
- Ravitch, D., 2010. *The death and life of the great American school system*. Basic Books, New York.
- Sachs, J., 2005. Development of professional identity: learning to be. In: Denicolo, P., Kompf, M. (Eds.), *Connecting Policy and Practice: Challenges for Teaching and Learning in Schools and Universities*. Psychology Press, pp. 5–21.
- Sadler, D.R., 1989. Formative assessment and the design of instructional systems. *Instr. Sci.* 18 (2), 119–144.
- Sahlberg, P., 2011. The fourth way of Finland. *J. Educ. Change* 12 (1), 173–185. <https://doi.org/10.1007/s10833-011-9157-y>.
- Sato, M., 2014. What is the underlying conception of teaching of the edTPA? *J. Teach. Educ.* 65 (5), 421–434. <https://doi.org/10.1177/0022487114542518>.

- Stacey, M., Talbot, D., Buchanan, J., Mayer, D., 2020. The development of an Australian teacher performance assessment: lessons from the international literature. *Asia Pac. J. Teach. Educ.* 48 (5), 508–519. <https://doi.org/10.1080/1359866X.2019.1669137>.
- Steiner-Khamsi, G. (Ed.), 2004. *The Global Politics of Educational Borrowing and Lending*. Columbia University: Teachers College Press, New York.
- SCALE, 2021. Stanford Center for Assessment, Learning and Equity. About edTPA: Outcomes. Pearsons Education, MA. <http://edtpa.aacte.org/about-edtpa#Overview-0>.
- Shaukat, S., Chowdhury, R., 2021. Pre-service teachers' perceptions of professional standards and their integration into pre-service training: a comparative study of Australia and Pakistan. *Aust. J. Teach. Educ.* 46 (11), 54–68. <https://search.informit.org/doi/10.3316/informit.274718860667219>.
- Supovitz, J.A., Taylor, B.S., 2005. Systemic education evaluation: evaluating the impact of systemwide reform in education. *Am. J. Eval.* 26 (2), 204–230.
- Tang, Y.F.S., 2008. Issues in field experience assessment in teacher education in a standards-based context. *J. Educ. Teach.* 34 (1), 17–32. <https://doi.org/10.1080/02607470701773416>.
- TEMAG, 2015. Teacher Education Ministerial Advisory Group Action. Now: Classroom Ready Teachers. Australian Government Department of Education, Canberra. Accessed 20 June. <https://docs.education.gov.au/node/36783>, 2014. (Accessed 12 July 2022).
- Tuck, E., Gorlewski, J., 2016. Racist ordering, settler colonialism, and edTPA: a participatory policy analysis. *Educ. Pol.* 30 (1), 197–217 doi:10.1177%2F0895904815616483.
- Tuinamuana, K., 2011. Teacher professional standards, accountability, and ideology: alternative discourses. *Aust. J. Teach. Educ.* 36 (12). <https://doi.org/10.14221/ajte.2011v36n12.8>.
- Wenglinsky, H., 2002. The link between teacher classroom practices and student academic performance. *Educ. Pol. Anal. Arch.* 10, 12. <https://doi.org/10.14507/epaa.v10n12.2002>.
- Wei, R.C., Pecheone, R.L., 2010. Assessment for learning in preservice teacher education: performance-based assessments. In: Kennedy, M.M. (Ed.), *Teacher Assessment and the Quest for Teacher Quality*. Jossey-Bass, San Francisco, CA, pp. 69–132.
- Wyatt-Smith, C., Alexander, C., Fishburn, D., McMahon, P., 2017. Standards of practice to standards of evidence: developing assessment capable teachers. *Assess. Educ. Princ. Pol. Pract.* 24 (2), 250–270. <https://doi.org/10.1080/0969594X.2016.1228603>.
- Wyatt-Smith, C., Humphry, S., Adie, L., Colbert, P., 2020. The application of pairwise comparisons to form scaled exemplars as a basis for setting and exemplifying standards in teacher education. *Assess. Educ. Princ. Pol. Pract.* 27 (1), 65–86. <https://doi.org/10.1080/0969594X.2020.1712326>.
- Wyatt-Smith, C., Adie, L., Haynes, M., Day, C., 2022. Three essential features in a new conceptualization of complex performance assessments: authenticity, system and site validity, and intelligent accountability. In: *Professionalizing Teacher Education: Performance Assessment, Standards, Moderation, and Evidence*, 1st. Routledge, London, pp. 48–63.

Determining long term effects of teacher education

C.N. Brouwer, Radboud University Nijmegen, Nijmegen, the Netherlands

© 2010 Elsevier Ltd. All rights reserved.

This is a reprint of C.N. Brouwer, Determining Long Term Effects of Teacher Education, In International Encyclopedia of Education (Third Edition),

Edited by : Penelope Peterson, Eva Baker, Barry McGaw, Elsevier, 2010, <https://doi.org/10.1016/B978-0-08-044894-7.00644-8>.

Introduction	437
Relevance of TE effect research	437
Aim of TE effect research	437
Contested ground. What we know about TE effects and how we know them	438
The state of the art	438
Seminal studies	438
Research reviews	439
How to research long-term effects of educating teachers?	439
Modeling the object of research	439
Methodology of exemplary studies	440
Validity requirements	441
Suggestions for conducting TE effect studies	441
Design	441
Instrumentation	441
Data collection	442
Analysis methods	442
Conclusion	442
References	442
Further reading	443
Relevant websites	444

Introduction

The existence of programs for preservice teacher education (TE) is not self-evident. The life and sustainability of such programs ultimately depend on the effects they have on teachers' in-service behavior. Effects of TE are defined here as durable consequences of the preservice acquisition of knowledge and skills and development of motives, as manifested in graduates' professional behavior. A stricter definition pertains to long-term effects of TE programs. Here, such effects are defined as durable consequences of participation in preservice programs which manifest themselves in professional behavior more than 1 year after the completion of preservice TE. It is the responsibility of educational research to demonstrate, describe, explain, and help optimize the relationship between preservice TE and the professional behavior which teachers exhibit and develop during their careers. The description and explanation of long-term effects of preservice programs lend greater credence to the utility of TE.

Relevance of TE effect research

Studying TE effects and especially long-term effects empirically is a relevant endeavor in several respects. Quantitative arguments in favor of TE effect studies deal with issues of recruitment, selection, and retention. Any country that wishes to sustain its educational system must know what contribution TE programs make to the availability and employability of its teacher workforce. Like no other component of a country's education system, TE possesses a multiplier function for the quality of schooling and learning throughout the system as a whole. TE programs deliver generations of teachers, who in turn educate generations of pupils and students, who on their part carry with them and transform the knowledge and skills acquired in school, college, or university during the rest of their lives and work. Indeed, the contributions that TE can make to the productivity of a country's educational system eminently justify research into the influence of TE programs on the teaching competence of their graduates.

Aim of TE effect research

The aim of TE effect research should be to explain the influence that specific features of preservice programs exert on teaching competence, especially in the longer run. If predictors of success in teaching can be found before and during TE as well as during beginning teaching and the further career, this evidence could underpin guidelines and strategies for improving TE learning environments. Empirical evidence about long-term TE effects could not only help increase retention in the profession, but also lay a basis for developing teachers' professional practice and well-being during all stages of their careers. The relevance of long-term effect

studies for policymaking in the domain of TE is clearly illustrated by the debate about what is an adequate or even the best architecture of TE programs, be they regular 4-year programs or alternative certification programs (e.g., [Darling-Hammond and Youngs, 2002](#)). Teacher shortages worldwide lend urgency to such questions. As yet, the empirical knowledge base for designing learning environments for teachers is growing, but seriously incomplete. There is a clear need for research which explains specifically how TE learning effects come about.

TE as a research field characterizes itself by an enormous complexity. Program effects are not exclusively influenced by program features, but these interact with the personal backgrounds of candidates and the varied contexts in which their schooling takes place. To explain the complex interplay between conditions, processes, and outcomes involved in teacher learning, an adequate research base is not yet available, the American Educational Research Association (AERA) Panel on Research and Teacher Education has concluded. However, improvements in research design are underway ([Cochran-Smith and Zeichner, 2005](#), p. 704). This conclusion is even more applicable when long-term effects of TE are taken as criteria.

This article focuses on what we know about TE effects as well as how we know it. Coming to grips with the complexity of teacher learning is crucial to scientific progress in the field of TE. The empirical evidence about effects of TE programs are therefore referenced with special attention to the strengths and weaknesses of TE effect studies. On this basis, validity requirements for TE effect research are explicated and productive research strategies are presented. For this purpose, a number of pioneering and exemplary studies serve as illustrations.

Contested ground. What we know about TE effects and how we know them

Empirical evidence about effects of TE programs has begun to emerge during the last decades of the twentieth century, both in Europe and in the English-speaking world. Toward the end of the century, growing teacher shortages in the United States spurred survey research on a larger scale into the outcomes of alternative certification programs. These programs, developed initially in the US since the 1980s, were intended to attract new candidates to the teaching profession ([Dill, 1996](#); [Zeichner and Schulte, 2001](#)) and often consisted of abbreviated or condensed versions of regular TE programs. The need to evaluate the adequacy of these programs has put effect research into TE programs, both alternative and regular programs, high on the agenda. However, long before that time, in fact since the formalization and institutionalization of TE programs began in the 1920s, qualitative research has been carried out representing the experiences of teacher educators in developing programs and of candidates while learning to teach.

The state of the art

The state of the art in TE effect research becomes apparent from seminal studies and research reviews in this domain.

Seminal studies

Seminal research exploring how teachers become and develop as teachers was reported by Lortie in his well-known book *School Teacher*. In two case studies using survey questionnaires and retrospective interviews, Lortie vividly showed and persuasively argued how a web of social and psychological forces conditions the ways in which teachers are socialized by the existing school system ([Lortie, 1975/1977, 1995](#)). Lortie's sociological angle reminds the reader of the risks of unreflective reproduction of craft traditions of teaching. In his wake, Lacey explored attitude change in aspiring teachers as they enter the school system and traced the various strategies of occupational socialization open to them. On the basis of five case studies, in which observations and questionnaires were used to follow prospective teachers as they moved from preservice TE through their first year of beginning teaching, Lacey ([1977, 1995](#), p. 619) distinguishes among internalized adjustment, strategic compliance, and strategic redefinition as possible ways for beginning teachers to deal with existing traditions in schools.

Until the 1970s, large-scale empirical evidence about long-term TE effects was virtually nonexistent; in fact, even now, shortly after the turn of the century, it is scantily available. The project Teacher Attitudes, carried out at Konstanz University in the south of Germany, was the first in Europe to demonstrate by means of path analysis on national longitudinal survey data the influence of specific features of preservice programs on outcomes in graduates ([Dann et al., 1978, 1981](#)). This project showed that program features encouraging the integration of theory and practice can delay and/or mitigate the occurrence of practice shock and the associated attitudinal adjustment to existing teaching practices. Building on work on types of attitude change, [Dann et al. \(1978, pp. 96–104\)](#) and [Müller-Fohrbrodt et al. \(1978\)](#) have identified as crucial in beginning teachers' socialization what they call discrepancy experiences; they experience a rift between idealistic notions developed during TE programs, on the one hand, and pressure from schools as institutions to rely on traditional patterns of behavior, on the other. These discrepancy experiences are strongest in situations where practical action is required ([Dann et al., 1981](#)).

This phenomenon of conservative attitude shift was demonstrated repeatedly in the US by the Pupil Control Studies initiated by Hoy ([Packard, 1988](#)). It was also documented in early qualitative research by [Corcoran \(1981\)](#), who described entry into teaching in terms of transition shock.

Perhaps the most important contribution of the Konstanz research group is that program characteristics were shown to differ in how fast and in what respects they influenced the U-curve in prospective and beginning teachers' attitude development ([Dann et al., 1978, 1981](#)). This evidence supports the conclusion that integrative approaches in TE, in which student teachers' practical experiences are closely linked to theoretical input, can strengthen graduates' innovative teaching competence ([Dann et al., 1978, 1981](#)).

Similar research and findings were reported from Hamburg by [Hinsch \(1979\)](#) and from the Netherlands by [Brouwer and Korthagen \(2005\)](#).

Research reviews

Both the European research referenced above and the comprehensive review undertaken by the AERA Panel on Research and Teacher Education in the US confirm that TE programs do influence graduates' learning outcomes. The AERA review contains indications that coursework can promote graduates' subject-matter knowledge and that methods courses combined with field work can have an impact on their professional behavior ([Cochran-Smith and Zeichner, 2005](#)). The beneficial influence of integrating theory and practice during teacher preparation on teachers' in-service performance was confirmed in the US by reviews and secondary analyses of partly cross-sectional, partly longitudinal survey data carried out by [Darling-Hammond \(1999, 2000\)](#). This synthetic work has yielded two important conclusions: (1) the amount and quality of teacher preparation are positively related to teachers' retention in the profession as well as their success in promoting student achievement and (2) teacher quality is the single most influential factor determining student achievement.

Taken together, these conclusions mean that preservice TE programs have a discernible and relevant influence on teaching and learning in schools, especially if they integrate practical and theoretical components. However, precisely how these effects come about and which specific program features are responsible for which kinds of effects are questions which require further study of a sophisticated kind.

How to research long-term effects of educating teachers?

So far, much TE research lacks the qualities needed to refine and elaborate upon the available evidence. [Wilson et al. \(2002\)](#) reviewed 57 studies about the impact of TE programs published in peer-reviewed journals mostly during the 1990s. They concluded that this research was often small scale, allowed researcher bias, employed few direct measures, and often lacked in detail in reporting about methodology. Similar criticisms were voiced by [Wideen et al. \(1998\)](#) in their review of the research about learning to teach. In their judgment, the theoretical basis of such studies should be made more explicit. They also note a lack of longitudinal research and advocate a systemic and ecological approach in TE effect research. Considering this state of affairs, the AERA Panel on Research and Teacher Education calls for research taking into account the contexts of TE programs, participants' backgrounds and characteristics, and how these relate to the processes and outcomes of TE programs. Such research requires employing longitudinal designs more frequently and a theory-driven conceptualization of outcome measures, whose operationalization should go beyond indirect verbal data.

These reviews, with quite a some consensus, point in the same direction. The outcomes of TE programs probably never result from a limited number of social conditions and personal characteristics. They had better be conceived of as resulting from a complex interplay of factors which simultaneously create and constrain opportunities for learning. To understand how this complex interplay affects teacher learning, a type of relational knowledge is needed, as described by [Apple \(1979\)](#); in other words, a systems perspective which pays attention to the manifold relationships between the components of TE as a social system. If TE effect research is to answer its questions, it would benefit not only from stronger conceptualization generally, but also, especially from using ecological and developmental concepts to theorize about its research object.

Modeling the object of research

In order to explain how all the factors operating in TE programs work together to generate teacher learning, TE effect research should rest on more systematic theorizing. A suitable vehicle for achieving this is the decision-oriented evaluation model introduced by [Stufflebeam and Webster \(1988\)](#). This model distinguishes between different types of variables, designated as context, input, process, and product (CIPP). When applied to preservice TE, context factors include the institutional and cultural constraints in schools, colleges, and universities within which TE programs are offered. Input factors include personal background characteristics of teacher candidates and the resources for training and guiding them on their way to beginning teaching. Process factors are the ways in which programs are implemented, including the degree to which and the ways in which practical exercise and theoretical study are adjusted to each other. These process factors are decisive mediators determining the nature of learning effects in graduates. These effects fall under the CIPP model component product. Models for TE effect research should take into account candidate, program, and context features and preferably include factors from each model component relevant for the study at hand. This kind of modeling encourages the researcher to examine the relationships and the balance of forces between influences emanating from TE programs and from school contexts in particular.

Ecological conceptions of the objects of social research go rather against the grain of mainstream educational psychology, but if the TE research community wants to develop a methodology that is congruent with theorizing about the complexity of teacher learning, it could consider a paradigm shift from a mechanistic toward an organicist worldview (cf. [Baltes et al., 1977](#)). In developmental psychology, ecological conceptions have long been available, notably in the work of [Bronfenbrenner \(1979\)](#), who situates the individual as one of the driving forces in a nested structure of systems, that is, the microsystem, in which he or she directly participates, the mesosystem, which constitutes the institutional constraints within which human activity unfolds, and the macrosystem on national and world levels determining the material and cultural conditions of human development. Researching the individual as an entity developing in its social context, warns Bronfenbrenner, is seeking to understand his or her exosystem, that is, how the relationships between the three subsystems distinguished mediate personal development.

Methodology of exemplary studies

What we know about TE effects is determined to an important extent by how we know them. Empirical studies are inevitably shaped by the way in which researchers relate their theoretical perspectives to their choice and use of methods. It is from this angle that we now examine an illustrative selection of exemplary TE effect studies in order to try to cull from the researchers' experience suggestions for conducting ecological inquiry into the influence of TE programs on the teacher workforce and its professional competence. The studies presented below are all examples of longitudinal research.

Phenomenographic studies

Since the 1970s, especially, there has been a host of exploratory, phenomenographic case studies into prospective and beginning teachers' experiences in becoming a teacher (see the review by [Wideen et al., 1998](#)). Most of this research is purely qualitative and involves small numbers of respondents, if only because the researchers lacked the resources to process large amounts of qualitative data. The result has been to describe in interpretive ways prospective teachers' experiences as they learned to teach. An advantage of this approach is that it can illuminate the processes and mechanisms involved in teacher learning, especially if the "inner workings" of the TE programs concerned are carefully described ([Cochran-Smith and Zeichner, 2005](#): p. 700). However, studies which restrict themselves to using only qualitative methods in small samples may fail to identify and explain which program features significantly encourage later performance in teaching practice. Without larger numbers of respondents and strategic sampling, TE effect research cannot produce generalizable knowledge about productive learning environments for teachers.

Large-scale quantitative studies

At the other extreme is purely quantitative survey research capitalizing on the law of great numbers, which uncovers patterns and relationships, notably concerning issues of recruitment (cf. the pertinent entry), selection, and retention.

A persuasive example is [Ingersoll's \(2001\)](#) demonstration of a revolving-door effect in teacher turnover. By analyzing with regression techniques the national longitudinal data sets in the US spanning 7 years, Ingersoll showed that contrary to public and political opinion, there are sufficient numbers of teachers available who are willing to fulfill schools' staffing needs. However, too many of them grow dissatisfied too soon with the work conditions in education and therefore move to different schools or leave teaching profession altogether. [Chin and Young \(2007\)](#) reported how personal backgrounds of alternative-route candidates such as ethnic status, age, gender, family obligations, and career history interact with local and regional employment opportunities as well as attractive characteristics of teaching as a profession in determining recruitment patterns.

In the area of teacher selection and self-selection, various longitudinal studies were carried out tracing the backgrounds, employment, and work experience of graduates of alternative certification programs. A review of these studies makes it clear that rather than general measures, such as grade point averages based on standardized achievement tests, candidates' cognitive skills, and levels of cognitive complexity, their self-efficacy and previous experience in instructing and coaching people in group settings have some merit as predictors of retention or at least success in teaching as rated by teacher educators and employers ([Brouwer et al., 2007](#)).

Time-dependent patterns in retention have been uncovered by [Shen \(2003\)](#), using large-scale longitudinal data. Dropout of beginning teachers was shown to increase with time, often in leaps at the beginning of a new school year or semester. In the same research, a positive relation was found between the length of training and the length of time teachers remained in the profession. In addition, support for teachers after graduation can influence their commitment to schools. Various induction packages consisting of different types of support for teachers who were new to a school were shown to differ significantly in their influence on percentages of turnover after the first year of employment. Much of the turnover occurred among teachers without any type of support during induction. Turnover was reduced among teachers who had supportive communication with school leaders, and even more so when they were given opportunities to collaborate with colleagues and had access to teacher networks and extra resources ([Smith and Ingersoll, 2004](#)).

In the area of attitude development, [Watzke \(2007\)](#) analyzed the development of work-related concerns in 79 beginning teachers using six repeated measures with a concerns inventory. [Fuller \(1969\)](#) and others have posited a stage chronology, which has tended to portray beginning teachers' development as a series of subsequent steps proceeding from concerns about maintaining a desirable presentation of self through concerns about performing teaching tasks proficiently to concerns about the adequacy of one's own teaching in promoting student learning. In contrast to this self-task-impact-stage chronology, concerns about impact on learners were shown in Watzke's sample to rank consistently highest across time. Concerns about impact, whether academic or personal in orientation, were shown to recur over time instead of characterizing one specific phase of beginning teacher development.

Mixed-methods studies

A study based on ecological assumptions and relating processes and outcomes is the reconstruction of competence development in 357 teachers from the beginning of preservice training until the third year of in-service practice reported by [Brouwer and Korthagen \(2005\)](#). Building on the work of [Dann et al. \(1978\)](#) and using a conceptual model derived from the literature on teacher socialization, this study assessed the relative influences on competence development of preservice TE programs on the one hand, and school contexts before and after graduation on the other. The methods used in this study were questionnaires, document analysis, observations, and retrospective interviews with student teachers as they developed into beginning teachers, cooperating teachers, and teacher educators. Among the main findings were that a carefully planned alternation of student-teaching and college-based periods, a gradual increase in the complexity of student-teaching activities, and the use of a clinical supervision model in cooperative small-group settings encouraged graduates' professional competence, notably their ability to activate learners.

Longitudinal mixed-methods approaches capable of addressing the complexity of teacher learning are also used on a larger scale in the Pathways into Teaching Project in New York City schools. In this project, comprehensive quantitative surveys are combined with detailed qualitative descriptions of an array of different TE programs and the learning experiences of the participants in them (Boyd et al., 2006).

Validity requirements

Against this backdrop of examples and with due caution, the methodological strengths and weaknesses of studies of long-term TE effects can be assessed as follows. Explaining relationships between TE program features and outcomes in graduates is facilitated by combining quantitative and qualitative methods. Using quantitative data from sizable samples enables the researcher to establish significant relationships and patterns and begin to generalize findings. Using qualitative data, preferably from direct observation, enables the researcher to reconstruct and infer from thick description (Geertz, 1973) in which settings and in which ways teachers acquire and transform professional teaching behaviors. TE effect research with explanatory power is further characterized by longitudinal rather than cross-sectional designs, repeated rather than one-shot measures, and multivariate correlational rather than experimental or quasi-experimental methods of analysis.

Given these experiences and trends in the conduct of TE effect research, it is appropriate to explicate what should count as valid causal explanations in an ecological conception of TE effect research. Therefore, two validity requirements are now presented for the benefit of explaining the complex whole of factors determining effects of TE programs. Conclusions about cause–effect relationships in TE effect research should be considered valid only, when:

- relationships are demonstrated on the basis of a research model incorporating context, input, process, and product factors and
- the processes responsible for these relationships are reconstructed.

The latter demand implies that explanations of human learning are incomplete, if only an experimentalist, black-box approach is taken. Such an approach would underestimate the role that subjective representations, such as perceptions, experiences, opinions, and attributions, play in human learning (cf. Maxwell, 2004). These two validity requirements embody a causal-genetic conception of scientific explanation in educational research and share important assumptions about the object of research as put forward in Bronfenbrenner's ecological conception of human development.

Suggestions for conducting TE effect studies

The validity requirements proposed above lay a basis for arriving at decisions about research design, instrumentation, data collection, and data analysis (cf. the criteria for research programs advocated by Bronfenbrenner and Läscher, 1976).

Design

An ecological conception of TE programs and effects as an object of causal-genetic explanation would entail including in the design of an empirical study the main settings of the college-based curriculum, student teaching, and beginning teaching in schools. Prospective teachers should be followed longitudinally from the beginning of training until the second year of in-service teaching, preferably longer. The instruments and measures used should allow recording and understanding the backgrounds, actions, experiences, and motives of the persons studied.

Combining qualitative and quantitative methods helps solve the breadth–depth problem in investigating TE programs and effects. If one type of method is used exclusively, either many situations and persons are studied superficially, or a few are studied thoroughly (Berger, 1974). A solution to this problem is to select, from all the programs and respondents in a study, a smaller number to form a representative subset. These may be termed the whole sample and the subsample, respectively. The whole sample can then be studied using quantitative methods, so that generalizable relationships and patterns can be discovered. The subsample can be studied additionally by means of qualitative methods, so that mediating processes can be reconstructed and explained in sufficient detail.

Instrumentation

In order to arrive at a multifaceted representation of both the outcomes and processes of long-term teacher learning, it is advisable to use multiple instruments for data collection (cf. for the original idea of a multitrait–multimethod matrix in social research, Campbell and Fiske, 1959). Not only self-report methods should be used, but also direct observation, so that attributions by any respondent group can be recognized for what they are and reliable accounts of actual professional performance can be achieved.

The operationalization of observations and measures is crucial for validity. Three epistemological perspectives, that is, assumptions about the nature of TE programs and effects can serve as a guide for operationalization: the ecological, genetic, and activity perspectives. The ecological perspective assumes that the research object in the social sciences always consists of a social system, which is at the same time internally structured and embedded within a wider, often institutional context (cf. Maschewsky, 1979; Tabachnick, 1981). The genetic perspective recognizes all learning as a set of processes, whose unfolding influences learning. This perspective assumes that understanding learning outcomes requires understanding the processes which produce them (cf.

Davydov, 1977). The activity perspective entails a focus on the actions of people as these express the continual tension between their personal motives on the one hand and contextual constraints on the other (cf. Zeichner and Tabachnick, 1985).

In quantitative inquiry, low-inference measures should be employed wherever possible. In qualitative inquiry, a balance should be sought between self-report and direct observation data. On-site visits, video recording of teaching, and retrospective and stimulated-recall interviews are therefore suitable methods.

Data collection

In an ecological inquiry of TE programs and effects, minimally, the main respondent groups should be involved, that is, prospective, beginning and/or experienced teachers, cooperating teachers, teacher educators, and/or professional development facilitators (cf. Cochran-Smith and Zeichner, 2005, pp. 331–334).

An important consideration in designing TE effect research is which time spans should be covered in longitudinal data collection. Decisions on this point are obviously made depending on the research questions and available resources, but it should be borne in mind that learning effects need time to come to fruition and to manifest themselves. To determine long-term effects of TE, it is advisable to cover time spans of at least 2 years. When following respondents over time, it is of course necessary to identify them, secure their consent with continued participation, inform them of the study's progress, and take all other measures useful for guarding against attrition generally (Mednick et al., 1984).

Analysis methods

Having based research design on ecological assumptions, TE researchers face strategic choices in analyzing pluriform data sets. Formulating and accounting for these choices with the help of a predetermined analysis plan can make or break a study's success. Evaluating any specific TE program in an ecological fashion confronts the researcher with social constellations and processes of unique complexity. Tracing distal causes and mediating factors will be prominent challenges (cf. Baltes et al., 1977). An analysis plan should therefore serve careful hypothesis generation and explicate the reasons why certain analysis techniques are selected and how their use is sequenced.

Given the validity requirement that the explanation of learning outcomes should rest on a reconstruction of the learning processes involved, it is logical to make qualitative analyses preceding quantitative ones, so that hypotheses can be generated and explored in a controlled way. If multivariate analyses yield clear indications of testable relations and/or differences, then experimental or quasi-experimental analyses can follow. This sequence conforms to the plea by Chatterji (2004, 2006) for extended-term mixed-method evaluation (ETMM) designs.

Finally, in analyzing quantitative longitudinal data, caution is needed in handling missing values. Attrition of respondents can easily cause individual data pairs to drop out of specific correlations and regressions. This risk increases to the degree that data pairs are collected further apart in time. To contain this risk, techniques can be used for estimating missing values (Peugh and Enders, 2004).

Conclusion

Qualifying and retaining teachers requires building an empirical knowledge base which can underpin program design in TE. Answering questions relevant for explaining and optimizing long-term effects of preservice TE programs can be achieved by means of carefully planned longitudinal studies combining quantitative and qualitative methods. TE researchers are well advised to conceptualize their questions using explicit, ecological models, which help them to theorize about the crucial role of mediating processes in generating learning effects in teachers. Research design, instrumentation, data collection, and analysis should fulfill validity requirements which are congruent with research models aiming at causal-genetic explanation.

If TE effect research were to succeed in mastering its challenges, evidence-based decision making could become possible in policymaking for TE. At the beginning of the 21st century, the greatest challenge for research into the long-term effects of TE is to help solve the dilemma of countering teacher shortages while simultaneously maintaining, enhancing, and assuring the quality of the teacher workforce worldwide.

References

- Apple, M., 1979. *Ideology and Curriculum*. Routledge and Kegan Paul, London.
- Baltes, P.B., Reese, H.W., Nesselroade, J.R., 1977. *Life-Span Developmental Psychology: Introduction to Research Methods*. Brooks/Cole, Monterey, CA.
- Berger, H., 1974. *Untersuchungsmethode und soziale Wirklichkeit (Research Method and Social Reality)*. Suhrkamp, Frankfurt.
- Boyd, D.J., Grossman, P., Lankford, H., et al., 2006. Complex by design: investigating pathways into teaching in New York City schools. *J. Teach. Educ.* 57 (2), 155–166.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press, Cambridge, MA.
- Bronfenbrenner, U., Läscher, K. (Eds.), 1976. *'Ökologische Sozialisationsforschung' (Ecological Socialization Research)*. Ernst Klett Verlag, Stuttgart.
- Brouwer, C.N., Korthagen, F.A.J., 2005. Can teacher education make a difference? *Am. Educ. Res. J.* 42, 153–224.

- Brouwer, C.N., Ackley, B.C., Fallon, M.A., 2007. Intake assessments for alternative teacher education: moving from legitimation towards predictive validity. *Assess Eval. High Educ.* 32 (6), 657–665.
- Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychol. Bull.* 56, 81–105.
- Chatterji, M., 2004. Evidence on “What works”. An argument for extended-term, mixed-methods (ETMM) evaluation designs. *Educ. Res.* 33, 3–13.
- Chatterji, M., 2006. Grades of evidence. Paper presented at. In: the AERA Annual Meeting 2006. San Francisco, CA.
- Chin, E., Young, J.W., 2007. A person-oriented approach to characterizing beginning teachers in alternative certification programs. *Educ. Res.* 36, 74–83.
- Cochran-Smith, M., Zeichner, K. (Eds.), 2005. *Studying Teacher Education. The Report of the AERA Panel on Research and Teacher Education*. Erlbaum, Mahwah, NJ.
- Corcoran, E., 1981. Transition shock: the beginning teacher's paradox. *J. Teach. Educ.* 32, 19–23.
- Dann, H.-D., Cloetta, B., Mller-Fohrbrodt, G., Helmreich, R., 1978. Umweltbedingungen innovativer Kompetenz. Eine Lngsschnittuntersuchung zur Sozialisation von Lehrern in Ausbildung und Beruf (Contextual Conditions of Innovative Competence. A Longitudinal Study of Teacher Socialization During Preservice Training and Beginning Teaching). Klett-Cotta, Stuttgart.
- Dann, H.-D., Mller-Fohrbrodt, G., Cloetta, B., 1981. Sozialisation junger Lehrer im Beruf. Praxischock drei Jahre spter (Occupational socialization in young teachers. Practice shock three years later). *Zeitschrift fr Entwicklungspsychologie und Pdagogische Psychologie* 13, 251–262.
- Darling-Hammond, L., 1999. *Teacher Quality and Student Achievement*. University of Washington, Center for the Study of Teaching and Policy, Seattle, WA.
- Darling-Hammond, L., 2000. How teacher education matters. *J. Teach. Educ.* 51 (3), 166–173.
- Darling-Hammond, L., Youngs, P., 2002. Defining “highly qualified teachers”: what does “scientifically-based research” actually tell us? *Educ. Res.* 31, 13–25.
- Davydov, V., 1977. *Arten der Verallgemeinerung im Unterricht (Types of Generalization in Education)*. Volk und Wissen, Berlin.
- Dill, V.S., 1996. Alternative teacher certification. In: Sikula, J., Buttery, T., Guyton, E. (Eds.), *Handbook of Research on Teacher Education*. Macmillan, New York.
- Fuller, F.F., 1969. Concerns of teachers: a developmental conceptualization. *Am. Educ. Res. J.* 6, 207–226.
- Geertz, C., 1973. Thick description: toward an interpretative theory of culture. In: Geertz, C. (Ed.), *The Interpretation of Cultures*. Basic Books, New York, NY, pp. 3–30.
- Hinsch, R., 1979. Einstellungswandel und Praxischock bei jungen Lehrern, eine empirische Lngsschnittuntersuchung (Attitude Change and Practice Shock in Beginning Teachers, An Empirical Longitudinal Study). Beltz, Weinheim.
- Ingersoll, R.M., 2001. Teacher turnover and teacher shortages: an organizational analysis. *Am. Educ. Res. J.* 38, 499–534.
- Lacey, C., 1977. *The Socialization of Teachers*. Methuen, London.
- Lacey, C., 1995. Professional socialization of teachers. In: Anderson, W. (Ed.), *International Encyclopedia of Teaching and Teacher Education*. Elsevier-Science/Pergamon, Oxford, pp. 616–620.
- Lortie, D.C., 1975. *Schoolteacher, A Sociological Study*. University of Chicago Press, Chicago, IL.
- Maschewsky, W., 1979. ‘Zur Bedingungskontrolle in der psychologischen Forschung: Rationale, Mglichkeit, Notwendigkeit’ (Experimental control in psychological research: Its rationale, possibility, and necessity). In: Jger, M., Leiser, E., Maschewsky, W., Schneider, U. (Eds.), *Subjektivitt als Methodenproblem, Beitrge zur Kritik der Wissenschaftstheorie und Methodologie der brgerlichen Psychologie*. Pahl-Rugenstein, Cologne, pp. 142–174.
- Maxwell, J.A., 2004. Causal explanation, qualitative research, and scientific inquiry in education. *Educ. Res.* 33, 3–12.
- Mednick, S.A., Harway, M., Finello, K.M., 1984. *Handbook of Longitudinal Research*. Praeger, New York.
- Mller-Fohrbrodt, G., Cloetta, B., Dann, H.D., 1978. Der Praxischock bei jungen Lehrern. Formen-Ursachen-Folgerungen. Eine zusammenfassende Bewertung der theoretischen und empirischen Erkenntnisse (Practice Shock in Young Teachers: Forms, Causes, and Consequences. An Evaluative Summary of Theoretical Insights and Empirical Evidence). Klett-Cotta, Stuttgart.
- Packard, J.S., 1988. The pupil control studies. In: Boyan, N.J. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York, pp. 185–209.
- Peugh, J.L., Enders, C.K., 2004. Missing data in educational research: a review of reporting practices and suggestions for improvement. *Rev. Res. Educ.* 74, 525–557.
- Shen, J., 2003. Characteristics of those entering teaching via traditional and alternative certification: a national, longitudinal study of fresh college graduates. Paper Presented at the 2003. In: Annual Meeting of the American Educational Research Association. Western Michigan University, Kalamazoo, MI.
- Smith, T.M., Ingersoll, R.M., 2004. What are the effects of induction and mentoring on beginning teacher turnover? *Am. Educ. Res. J.* 41, 681–714.
- Stufflebeam, D., Webster, W.J., 1988. Evaluation as an administrative function. In: Boyan, N.J. (Ed.), *Handbook of Research on Educational Administration*. Longman, New York, pp. 569–603.
- Tabachnick, B.R., 1981. Teacher education as a set of dynamic social events. In: Tabachnick, B.R., Popkewitz, T., Szekely, B.B. (Eds.), *Studying Teaching and Learning: Trends in Soviet and American Research*. Praeger, New York, NY, pp. 76–86.
- Watzke, J.L., 2007. Longitudinal research on beginning teacher development: complexity as a challenge to concerns-based stage theory. *Teach. Teach. Educ.* 23, 106–122.
- Wideen, M., Mayer-Smith, J., Moon, B., 1998. A critical analysis of the research on learning to teach: making the case for an ecological perspective on inquiry. *Rev. Educ. Res.* 68, 130–178.
- Wilson, S.M., Floden, R.E., Ferrini-Mundy, J., 2002. Teacher preparation research: an insiders' view from the outside. *J. Teach. Educ.* 53, 190–204.
- Zeichner, K.M., Schulte, A.K., 2001. What we know and don't know from peer-reviewed research about alternative teacher certification programmes. *J. Teach. Educ.* 52, 266–282.
- Zeichner, K.M., Tabachnick, B.R., 1985. Social strategies and institutional control in the socialization of beginning teachers. *J. Educ. Teach.* 11, 1–25.

Further reading

- Bourdieu, P., Passeron, J.-C., 1977. *Reproduction in Education, Society and Culture*. Sage, London.
- Brophy, J.E., Good, T.L., 1986. Teacher behavior and student achievement. In: Wittrock, C.M. (Ed.), *Handbook of Research on Teaching*, third ed. MacMillan, New York, pp. 328–376.
- Cole, A.L., Knowles, J.G., 1993. Shattered images: understanding expectations and realities of field experiences. *Teach. Teach. Educ.* 9, 457–471.
- Dangel, J.R., Guyton, E.M. (Eds.), 2005. *Research on Alternative and Non-traditional Teacher Education: Teacher Education Yearbook XIII*. Scarecrow Education/Rowman and Littlefield, Lanham, MD.
- Frech, H.W., Reichwein, R., 1977. Der vergessene Teil der Lehrerbildung: Institutionelle Bedingungen und inhaltliche Tendenzen im Referendariat der Gymnasiallehrer (The Forgotten Part of Teacher Education: Institutional Conditions and Substantive Trends in the Practical Preparation of Gymnasium Teachers). Klett-Cotta, Stuttgart.
- Fullan, M.G., Stiegelbauer, S., 1991. *The New Meaning of Educational Change*. Cassell, London.
- Korthagen, F.A.J., Kessels, J., Lagerwerf, B., Wubbels, T., 2001. *Linking Practice and Theory. The Pedagogy of Realistic Teacher Education*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Sarason, S.B., 1996. *Revisiting the Culture of the School and the Problem of Change*. Teachers College Press, New York.
- Veenman, S.A.M., 1984. Perceived problems of beginning teachers. *Rev. Educ. Res.* 54, 143–178.
- Waller, W., 1967. *The Sociology of Teaching*. Wiley, New York.

Relevant websites

www.ctpweb.org – Centre for the Study of Teaching and Policy (CTP).

www.teacherpolicyresearch.org – Teacher Policy Research (TPR). This page intentionally left blank

A lifetime of experience: an international study on the lives of teacher educators¹

Theme 2: researching teacher educators, their lives and their learning

Clare Kosnik^a, Lydia Menna^b, and Pooja Dharamshi^c, ^a University of Toronto, Toronto, Canada; ^b University of Alberta, Edmonton, Canada; and ^c Simon Fraser University, Burnaby, Canada

© 2023 Elsevier Ltd. All rights reserved.

Part one: general discussion of teacher educators' lives and journeys	445
Government intervention in teacher education	446
Moving from an apprenticeship model to an academic/intellectual endeavor	446
Dismantling traditional teacher education	446
Research on teacher educators	448
Transition from classroom teacher to teacher educator	448
Knowledge bases for teacher educators	448
Teacher educators as researchers	449
Part two: teacher educators' lives and journey	449
Goals for their courses	450
Conceptions of literacy	451
Preparation to teach mandated curriculum: conflicting agendas	452
Place of and doing research	452
Identity	453
Discussion	454
Appendix one	455
Methodology	455
References	455

Discussing teacher education is complicated because each country develops its own system and programs (Darling Hammond and Lieberman, 2012; Murray et al., 2019). Teacher education in general reflects each country's specific culture and values, context and industries/resources, and long-standing practices; these practices are usually tied to the country's short-and long-term goals (Christou, 2017; Moss, 2009, 2017). This article recognizes the tremendous variance in teacher education; even within the same university or country there are a variety of models for teacher education. Variance is useful; on the one hand, it allows universities to tailor teacher education for its teacher trainees yet on the other hand, there is the danger of teacher education programs being driven by a particular government or a Head of School/Dean's particular interests.

The difficulty in discussing teacher education globally is further hampered by a lack of consensus on such basics as terminology and more broadly on goals. For example, students in teacher education programs can be referred to as teacher trainees, student teachers, or preservice candidates. The term teacher trainees will be used throughout the article. The term teacher educator will be used to refer to those who hold university appointments (as opposed to school-based teacher educators whose primary appointment is in schools.) In order to contribute to this encyclopedia choices were made about what aspects of teacher education to discuss and present. Most of the research in this article draws on perspectives from western countries which does not disqualify the validity of work done in other regions; however, it allows the article to have coherence by presenting a number of prominent themes and developments which are applicable to specific countries. This article on teacher educators does not attempt to discuss **all** teacher educators rather it has chosen to discuss teacher educators working in specific countries and contexts. The article is divided into two parts: Part One is a General Discussion of Teacher Educators and Part Two presents two recent studies on teacher educators by the authors.

Part one: general discussion of teacher educators' lives and journeys

When discussing teacher educators two lines of inquiry must be considered: the government's role in teacher education and research on teacher educators. At times these intersect while at other times they run parallel. This article includes both lines of inquiry because this provides a more fulsome discussion of teacher educators. The article is organized into four sections: the first section provides a backdrop to the increasing politicized context of teacher education such as increased government intervention (one

¹We wish to thank the Social Sciences and Humanities Research Council of Canada for their generous funding of this project.

line of inquiry); the second section describes three key findings of research on teacher educators (second line of inquiry). These three findings were selected because they are relevant to the third section which presents findings on a study of teacher educators' journeys. The final section argues that teacher educators need to be fully recognized academics who have a broad knowledge base and should have a measure of academic freedom because they are researchers and need to be at the government table where decisions about teacher education programs are made.

Government intervention in teacher education

Over time some countries (e.g., England, Australia, United States) seem to have done a 360-degree rotation in understanding what it means to be a teacher educator. They shifted from conceptualizing teacher educators as being mentors (apprenticeship model) to teacher educators as full-fledged academics who are researchers with a degree of academic freedom to design academically challenging courses to reverting back to viewing teacher educators as mentors (Furlong, 2013; Murray and Passy, 2014; Whitty, 2014).

Moving from an apprenticeship model to an academic/intellectual endeavor

Beginning in the 19th century most preparation to be a teacher was conceptualized as an apprenticeship model (Cuban, 1987; Houston and Prentice, 1988; Labaree, 2006). The goal of teacher educators was to impart to teacher trainees "a set of independent, observable actions that together make up the coherent whole of teaching practice" (Russ et al., 2016, p. 393). Fries and Cochran-Smith (2006) note that teaching was regarded largely as a technical transmission activity which occurred in a linear manner (p. 77). Early Normal Schools in Canada and the United States which were the forerunner to teacher's colleges saw teacher education as a place to improve beginning teachers' literacy and mathematics skills (Labaree, 2000, 2006) in order for them to more capably teach children and adolescents a standardized curriculum. The teacher educators' job was to prepare teacher trainees to teach a narrow curriculum focused on basic skills such as hand-writing and oral reading. Gradually the value of teachers having some level of preparation through formal teacher education was recognized which led to most attending at least basic teacher training programs completing a diploma or certificate at Normal Schools or a free-standing teachers' college (Ducharme and Ducharme, n/d). During periods of the 20th century requiring teachers to be certified diminished for a variety of reasons (e.g., teacher shortages) which led to uncredentialed teachers being given emergency credentials (Darling Hammond, 2006; Darling Hammond and Lieberman, 2012).

The next significant step toward the professionalization of teachers was moving teacher education programs into universities. The move into universities (emerged during the 1980s in many countries and is still continuing) conceptualized teacher education as both an intellectual process and an opportunity for teacher trainees to acquire a broad range of practical skills (Dewey, 1916, 1938; Goodlad, 1998). Shulman (1986) argued that teacher education was not merely an apprenticeship and suggested that teachers and teacher educators required three types of specialized knowledge: subject matter content knowledge, pedagogical content knowledge, and curricular knowledge. Shulman (1986) provided a pathway for future research by suggesting there needs to be an examination of the specialized knowledge for teaching (e.g., Ball and Hill, 2009; Kennedy, 2006) and others such as Darling Hammond tried to demonstrate how it can be learned (2002, 2005). Adding to this body of knowledge, socio-cultural theorists (e.g., Vygotsky, 1986; Ladson-Billings, 2005) placed teaching into larger social and cultural practices (Palincsar, 1998; Lewis et al., 2020). Teacher educators and teacher trainees had to examine and understand the social, cultural, and historical context of schooling. From this intellectual or inquiry perspective, learning to teach required a broad skill set, certain dispositions, and an intellectual interrogation of teaching; therefore, teacher education had to be more than an apprenticeship. (See Rosales, 2019 for a description of history of teacher learning).

The move into universities not only expanded the curriculum of teacher education, but also led to the reconsideration of teacher educators themselves. The inherent culture of universities expects all faculty to be researchers and academics, not merely practitioners. However, the move to include teacher educators in universities was not without its problems. There have been scathing analyses of teacher educators (Ducharme and Ducharme, 1996; Labaree, 2006) questioning the legitimacy of teacher educators to be full-fledged academics. The conundrum of moving teacher education from an apprenticeship model to an intellectual endeavor housed in universities brought concomitant problems which are still being realized (Goodlad, 1998). Were teacher educators outstanding teachers or were they full-fledged academics? The education community continues to wrestle with this question.

Dismantling traditional teacher education

In many countries over the last four decades, education and teacher education have been squarely placed in the cross-hairs of government agendas further confounding the situation. Broad-sweeping government intervention into teacher education in many ways upturned teacher education:

Something dramatic has happened to teacher education policy in the last 30 years. Up until the 1980s, in virtually every country around the world, teacher education was a relative backwater in terms of educational policy. In a world of intense competition among nations, education increasingly plays a key role. National prosperity, social justice, and social cohesion are all seen to rest on the shoulders of education.

Furlong (2013, p. 29).

As Furlong notes here, the stakes in education have been raised because it is seen to be one of the saviors for society and the economy. This stance has led to significant interference from governments. For example, in England the government has determined the number of days student teachers must spend in practice teaching schools (Murray and Passy, 2014, p. 497) which has diminished their connection to the university. Gilroy (2014) wryly observes that in 1992 the Secretary of State “launched a ferocious attack on university teacher education ... where he made it clear that only a tiny rump of courses would be left in the hands of higher education, with teacher education quickly to be moved to schools” (p. 623). Following suit, the US governments (federal and state) are increasingly inserting themselves into teacher education (Kosnik et al., 2016).

Tony Blair, Prime Minister of the UK from 1997 to 2007, was one of the first to ardently argue that teaching and teacher education were tied to national prosperity (Moss, 2009, 2017) but rather than elevate the role of teacher educators his government started a march to diminish and do away with formal teacher education programs and reduce teacher educators as academics (Furlong, 2013). To ensure national prosperity the “answer” was to improving the knowledge and skills of their citizens was the re-emergence of teacher education as an apprenticeship so that teachers could teach a ready-made, pre-set curriculum devised by the government (e.g., Literacy Hour). Governments seized almost full control of teacher education but quite illogically supported alternative certification programs which are the antithesis of teacher education as an intellectual endeavor (McNamara and Murray, 2013; Mitchell and Romero, 2010). Alternate certification programs are truncated programs often only a few weeks in length with many teacher trainees being taught a “script” for teaching. Therefore, the government believed that by removing as many variables as possible (teacher educator choice) they would improve the skills of all children. One of the first significant threats to university-based teacher education came from Teach for All, a global network of alternative certification programs which currently (2020) has partners with 50 countries. For example, Teach for America and other alternative certification providers have been so successful that approximately 20–30% of teachers entering teaching in the U.S. come from an alternative pathway (Zeichner and Conklin, 2016, p. 2). A host of other alternative certification providers emerged. In Britain there are now 1500 providers for teacher certification (Gilroy, 2014).

What began in the early 1980s as a way to ward off projected shortages of teachers and replace emergency certification has rapidly evolved into an accepted model for recruiting, training, and certifying those who already have at least a bachelor's degree and want to become teachers.

National Education Association (n/d).

One of the most questionable assaults on teacher educators has emerged from so called “reformers” who could be described as edupreneurials (Anderson, 2019, p. 13). One of their initiatives is what Cochran-Smith et al. (2018) call new graduate schools of education (nGSE) “which are independent organizations that offer teacher preparation programs and use university language, such as ‘graduate school of education’ ... ‘dean’ but are not affiliated with universities” (p. 575). They do not adhere to traditional university structures (e.g., practice teaching blocks), ignore important goals (e.g., developing teachers as decision-makers), and do not address a rich and varied curriculum practices (e.g., how to teach a diverse student body).

Quite perplexingly, teacher educators in alternative certification or edupreneurial programs often do not have higher education degrees, nor are they involved in research on their venture (Anderson, 2019). “Some of the leaders of nGSEs are former [Teach for America] TFA members, while many faculty are school-based educators who focus on effective practice” (Cochran-Smith et al., 2018, p. 576). Anderson (2019) presents the New Schools Venture Fund (NSVF) as an example of an edupreneurial reformer, initially focused on expanding the number of charter schools but since 2009 “are trying to profoundly disrupt the current teacher preparation market” (p. 8). Typically, nGSEs programs “disavow theory and social foundations” (Anderson, 2019, p. 7) and prefer Lemov’s “codification of discrete techniques that can be taught quickly to teachers” (2010, p. 16). Zeichner and Conklin (2016) observe that 40% of some programs are delivered through on-line instruction (p. 13) which makes the programs cheap to run thus appealing to for-profit companies.

There has long been oversight of teacher education programs; however, the level of compliance has been vastly elevated (Sahblerg, 2012a,b). The English government established the Office for Standards in Education, Children’s Services and Skills (OFSTED) which dictated not only what needs to be taught but how it should be taught (Yandell, 2012, 2016). For example, in England the government not only describes the content of the literacy programs in schools but also how literacy should be taught in both schools and teacher education programs. In many cases, teacher educators are being asked to return to a narrow delivery content and skills.

As noted earlier, questioning the rightful place of teacher educators to be university faculty has persisted. Cochran-Smith et al. (2018) argue that “teacher educators have been positioned as the objects of reform and the implementers of reform agendas developed by others, rather than as the agents of reform” (p. 573). An intended or unintended consequence of the government support for alternative certification programs jeopardized teacher educators, in particular their role—were they trainers or academics? As academics they would be required to have higher education degrees, be research-active, be able to secure grants, participate in scholarly conferences, and infuse their teaching with research. As trainers, they act only as mentors for apprentices.

Many teacher educators responded forcefully to government intervention by defending teacher educators as full-fledged academics who needed to develop programs for the 21st century not retreat to 19th century models. While some governments are attempting to dismantle teacher education programs by allowing alternative certification programs and nGSEs to be recognized, there is increasing attention by the international teacher education community to teacher educators themselves (Ellis et al., 2014;

Goodwin and Kosnik, 2013; Info Ted, 2019; Loughran, 2006). The European Commission (2009, 2013) recognizes there needs to be more commitment for supporting teacher educator learning. The teacher educators' efforts to preserve university-based teacher education has had limited success because governments began dismantling traditional teacher education in a number of ways such as reducing the number of days spent in academic programs at the university giving schools more control over the process (Florian and Pantic, 2013).

Research on teacher educators

In this maelstrom of attacks on traditional teacher education, teacher educators were caught between two diametrically opposed views: did they construct their programs as an apprenticeship or did they design it based on an intellectual model? Grappling with this distinction has been difficult and not yet "resolved." These are not mutually exclusive but each model will drive the content and format.

Nevertheless, research on teacher educators has grown substantially in the last two decades showing that they are not a homogenous group (Kosnik et al., 2013, 2014, 2017; Mayer et al., 2011; Mayer, 2014; Murray, 2016) with there being many paths to becoming a teacher educator (Murray et al., 2019), and that the political context is impacting them (Kosnik et al., 2020; Lambert and Burnett, 2015; Loudén, 2008; Peters et al., 2017). These key findings must be fore-fronted in order to understand the complexity of the situation. Although there are many important insights gained from the research on teacher educators, only three specific aspects will be discussed. They were selected because they connect to Section Three of the article and show the short-comings of the alternative certification programs described above.

Transition from classroom teacher to teacher educator

The ground-breaking article *Becoming a Teacher Educator: Evidence from the Field* by Murray and Male (2005) showed that the transition from classroom teacher to teacher educator was not a simple process:

The transition from the first-order setting of school teaching into the second-order setting of HE (higher education) based teacher education was constructed by the majority of the interviewees as a distinct and stressful career change, characterized by high levels of uncertainty and anxiety. Recurring feelings about the early years of HE work were of being "de-skilled", of "struggle", and of "masquerading" (p. 129).

The move to higher education was in many cases difficult. Murray and Male (2005) elaborated: "[t]hemes of 'survival', anxiety about 'fitting in' and striving to make sense of HE work were dominant in the accounts of life in the first year of teacher education work. Despite their school teaching experience, 26 of the interviewees reported feeling exposed, vulnerable and uncertain about their new teaching roles in HE ... there was no straightforward 'transfer' of the pedagogical knowledge and experience acquired in and through school teaching to the HE context" (p. 130). Over the years, Murray and Male's (2005) findings have been substantiated by multiple researchers (Dinkelman et al., 2006; Hokka and Etelapelto, 2014; McKeon and Harrison, 2012) adding depth and nuance to understanding the fraught shift from classroom teacher to teacher educator. As research on teacher educators mushroomed, issues regarding being a teacher educator that had previously been "invisible" emerged. For instance, White (2016, 2019); White et al. (2018), from Australia identified the massive demands on teacher educators as they developed school university partnerships.

Further, the issue of teacher educator identity emerged as a pivotal aspect for teacher educators given the dismantling of traditional university programs. Are they classroom teachers? Are they academics? Are they both? Kosnik et al. (2013) in their paper *A Foot in Many Camps: Literacy Teacher Educators Acquiring Knowledge Across Many Realms and Juggling Multiple Identities* argued that "since much of teacher educators' work was in schools (supervising student teachers, conducting research) they must be able to 'talk the talk' with teachers. They cannot shed their teacher identity completely when they secure a faculty position" (p. 535). In the current milieu, it seems that university-based teacher educators must hold multiple identities and have a broad skillset that includes teaching, research, building school-university partnerships, working with school districts, being advocates for teacher education, and actively participating in service to the university.

Knowledge bases for teacher educators

In an apprenticeship model, teacher educators would have been expert classroom teachers but from the perspective of teacher education as an intellectual endeavor Zeichner (2005) argues that preparation for teacher educators needs to be more expansive: "If teacher education is to be taken more seriously in colleges and universities, then the preparation of new school- and university-based teacher educators needs to be taken more seriously as well" (p. 124). Cochran-Smith (2003) argues that we need to pay "more attention to what teachers of teachers themselves need to know, and what institutional supports need to be in place in order to meet the complex demands of preparing teachers for the 21st century" (p. 6). To be able to handle this complex role requires many skills and extensive knowledge; however, for most new teacher educators there is little to no preparation.

As the research on teacher education expanded Korthagen (2010) and Loughran (2006) work on a pedagogy of teacher education revealed the complexity of the task; teacher educators must teach about teaching which requires different pedagogical skills and knowledge than simply being an expert teacher. In Loughran's ground-breaking text, *Developing a Pedagogy of Teacher Education* (2006), he argued teaching about teaching is a very different enterprise than teaching school children. He states that "[b]eing a teacher educator requires an understanding of teaching that goes beyond being a good teacher. There is a need to be able to theorize practice in such a way as to know and be able to articulate the what, how, and why of teaching and to do so through the very experiences of teaching and learning about teaching" (2006, p. 14). In later work, Loughran and Hamilton (2016) argue that since this "practice is so complex ... learning to teach must indeed entail considerably more than a training regime" (p. 5). Teacher education pedagogy must also go beyond modeling good practice; Williamson (2013) notes that "it is not enough for teacher educators to model instructional practices without being transparent about their pedagogical thinking. Much about teaching is invisible to the novice eye, and explicit modeling allows educators to unpack the pedagogical decisions that they make in their planning" (p. 137).

Identifying the knowledge needed to be a teacher educator where teaching is seen as an intellectual endeavor (rather than an apprenticeship) has led to various delineations. Murray and Male (2005) identified knowledge that new teacher educators need: the acquisition of the pedagogical knowledge and experience appropriate to being a teacher educator in HE; the enhancement and generalization of their existing knowledge base of schooling; developing an identity as a researcher; developing ways of working with mentors in school-based settings; acquiring pragmatic knowledge of the HE institution and how it operates (p. 158).

Goodwin (2008, 2012) approached the issue of knowledge by identifying five domains of knowledge that teacher educators require: personal knowledge/autobiography and philosophy of teaching; contextual knowledge/understanding learners; schools and society; pedagogical knowledge/content, theories, teaching methods and curriculum development; sociological knowledge/diversity, cultural relevance and social justice; and social knowledge/cooperative, democratic group process and conflict resolution.

Kosnik et al. (2015) in their study of literacy/English teacher educators in Canada, US, England, and Australia identified four spheres of knowledge for literacy/English teacher educators: knowledge of research; knowledge of pedagogy of teacher education; knowledge of literacy and literacy teaching; and knowledge of current school and government initiatives.

Common features in these knowledge bases are: knowledge must be extensive, including practical, theoretical, and contextual. Compared to traditional academics (e.g., chemistry, history) teacher educators must know about both research and practice situated within the socio-cultural context. Compared to instructors in alternative certification programs, university-based teacher educators must be much more than practitioners.

Teacher educators as researchers

As noted earlier, the move from teachers' colleges to universities raised many thorny questions. One of the most difficult has been the extent to which teacher educators should be researchers. In the three lists above (Cochran-Smith, 2003; Goodwin, 2008; Kosnik et al., 2015; Murray and Male, 2005) knowledge of research is included but that matter is far from resolved or fully understood. Are they consumers of research? Are they generators of research? Does research guide their teaching?

Becoming and being a researcher is difficult and is not typically in the purview of the work of classroom teachers; therefore, when teachers move to HE to be teacher educators there is often sudden additional demands to be both consumers and producers of research. Lunenberg et al. (2014) report that many university-based teacher educators consider being involved in research as part of their work and Loughran (2014) suggests that teacher educators should conceptualize their growth as a "research journey" (p. 2). However, striving to be a quality researcher increases pressure on teacher educators because it is another significant knowledge base for them to acquire and an important duty to fulfill. "Some teacher educators have developed expertise as researchers before or during their work as teacher educators [but] others have not" (Tack and Vanderline, 2014, p. 299). Becoming a researcher is not a straightforward process because it includes acquiring skills, expanding your academic communities, and shifting your identity (Murray and Male, 2005; Lunenberg et al., 2014). Tack and Vanderline (2014) describe a "researcherly disposition as the habit of mind to engage in research and thus to produce both local knowledge and public knowledge on teacher education. This researcherly disposition embodies three interrelated aspects: an ability to conduct research; a sensitivity to research occasions, and an inclination toward research" (p. 301).

MacPhail et al.'s study (2018) asked 61 higher education teacher educators (from England, Ireland, Israel, Norway Scotland, and The Netherlands) about their professional development needs. They found: "[t]here is an inextricable link between teaching and research. Teacher educators were unanimous in the need to keep abreast of recent developments that would inform their teaching" (p. 8). Yet many have extremely heavy teaching loads with limited time to read research and do not have a research background or the skills to conduct research. "There was a clear frustration for a number of teacher educators that the higher education institutes' research-active rhetoric did not match the university support structures" (p. 9). Czerniawski et al. (2018, p. 141) looked at professional development needs of teacher educators in England, Ireland, and Scotland and identified four top categories: current developments in teacher education; curriculum development; scholarly writing (e.g., articles, chapters, learning materials, and researching one's own practice). Teacher educators need assistance with developing research skills.

Part two: teacher educators' lives and journey²

This article now presents data from two research studies that the authors have conducted. The first study *Literacy Teacher Educators: Their Backgrounds, Visions, and Practices* included 28 literacy/English teacher educators (LTEs) from four countries: Canada, US,

Table 1 Background of participants—data as of 2019.

Country		
Canada—15 participants US—22 participants UK—12 participants		
Australia—10 participants		
Overview of experience		
	Years as CR teacher	Years in higher education
0	2	
1–5	9	17
6–10	9	22
11–20	10	19
20+	1	20
Unknown	2	1

England, and Australia who were interviewed between 2014 and 2017. The second study, *Pedagogy of Literacy Teacher Education: Meeting the Challenges of 21st Century Literacies* recruited 30 additional literacy/English teacher educators. The literacy teacher educators (LTEs) were interviewed in 2019. The overall goals of both studies are:

- (i) Primarily, to study in depth a group of literacy teacher educators, with special attention to their backgrounds, views, and practices and their experiences, induction, and support within the university.
- (ii) Contribute to emerging theory on a pedagogy of literacy teacher education.
- (iii) Investigate factors that affect LTEs' pedagogies and views.

The 58 participants were interviewed using a semi-structured format which took approximately 60–90 min. Interview questions included those on: background (e.g., timeline of turning points), topic of Ph.D. research, experience as a classroom teacher; influences on their current practice; qualities of a teacher educator; academic community; framework and goals for their literacy course(s); pedagogies used and reasons for using them; assignments and readings; and equity and diversity. Studying the work of LTEs adds to the research base on teacher education because it looks at those in a specific discipline—literacy—and looks at how they are responding to the current intrusion of the government. **Table 1** is a summary of the backgrounds of the participants.

Goals for their courses

Across both studies there was extensive examination of goals LTEs set for their courses. All 58 LTEs stated that building knowledge of literacy was their top priority. Their goals were broad moving beyond simply teaching to a script. It shows that knowledge of literacy although the main goal is supplemented with goals that are positioning literacy within a cultural and political context. Other goals mentioned were:

- Build knowledge of pedagogy
- Student teachers adopt a professional role
- Student teachers develop a critical stance
- Build knowledge of government initiatives
- Build knowledge of digital technology
- Focus on student teacher growth

One of the goals mentioned consistently across both cohorts was helping teacher trainees become what Justin³ called “public intellectuals.” Teachers need to be much more than technicians; therefore, their teacher education program needs to be more than what Jeff calls “teacher proof.” He felt that his job was not doing “teaching for dummies” but to help teacher trainees have a much deeper understanding of teaching. Padua's aims were to prepare his teacher trainees to understand that the current national curriculum in England and Australia are “prescriptive and narrow in their conceptual arrangement” but he felt that “we are duty-bound as academics to provide [teacher trainees] with a much broader conceptual frame than some of these initiatives allow for.” In designing his courses/modules Jeff wanted to find “a place where I can have rich discussions ... to me it's what are these questions that are gnawing at you that I put in your head that is going to make it easy or hard for you to implement that curriculum.” Tessa from Canada oriented her courses to help her teacher trainees become “critical consumers ... and not technicians” of the curriculum. Through their course goals they were indirectly (and at times directly) pushing back on government mandates by opening up spaces for discussion and critique.

Over time, the impact of politics and government intervention does not seem to have lessened. Overall, the LTEs were trying to preserve the concept of teachers as intellectuals. A theme common to the 58 participants was an emphasis on teacher trainee survival. Madelaine echoed the sentiments of many:

²See appendix for methodology.

³Pseudonyms used throughout.

We have to give [teacher trainees] the survival skills, which means understanding and knowing how to implement things that maybe we don't think are ideal but at the same time preparing them to bring about a better world. At the same time, we are also saying to them what is there right now and this is how you navigate it but this is what theory and research is showing is optimal learning conditions. I don't think it is fair for us to put that utopian vision in the most vulnerable people in the population, the teaching populations, and say fix the mess that we couldn't.

Both LTEs and teacher trainees were feeling pressure to implement “reforms” that conceptualize literacy narrowly (Kosnik et al., 2020). The LTEs aimed to prepare teacher trainees in many ways: build their identity as professionals, connect theory and practice, and encourage them to think about the big questions in education.

To meet the multiple goals for their work, LTEs felt they must have a broad knowledge base that spans theory and practice, have a vast repertoire of pedagogical skills, be very thoughtful in course construction, and be aware of the current political context. They needed to create learning opportunities that maximize on the limited contact time they had with teacher trainees. The next two sections describe some of the goals for their courses/modules in more detail. These were chosen because they root their goals in a discipline (literacy) and give concrete examples of the ways LTEs and teacher trainees go beyond a scripted curriculum or teaching to the exit requirement test such as a test or portfolio in order to be credentialed.

Conceptions of literacy

For most of the 58 LTEs, literacy was not seen as a practice synonymous with the discrete and autonomous skills of simply reading and writing as Lemov (2010) does. Although many current government curricula have a narrow view of literacy, most LTEs understood literacy as a complex social practice in which historical, cultural, social, and technological perspective are at the center (Kinloch et al., 2020; Menna, 2013; Milner et al., 2015; Souto-Manning and Stillman, 2020). It is fairly accurate to describe their approach to literacy as highly progressive seeing literacy as an expansive practice that includes the influence of digital technologies on communication, attention to home languages, and power dynamics at play in literacy education. For the most part their goals were not consistent with the government-developed curricula which often focuses intently on synthetic phonics, teaching writing according to a script, while minimizing digital technologies as a form of communication. Sonia from the U.S. described the dilemma both LTEs and teacher trainees face:

the classrooms are not necessarily implementing the type of literacy instruction that I'm advocating for ... I'm telling them to do A, and they're not seeing it, they're seeing B in the classroom, so that is a huge problem ... I'm teaching culturally relevant responsive and sustaining pedagogies. They have a very, very hard time wrapping their heads around that.

Through the design of their courses, readings selected, assignments set, inclusion of guest speakers, and connections with the community many of these LTEs were moving beyond teaching to a script. For example, Karly in the US, who had a long history of being involved in activism, partnered with local schools where teacher trainees were assigned to educators who were activists and were pushing back on the prescribed curriculum. Karly felt that the combination of her academic course, which made space for critical conversations, alongside placing teacher trainees with progressive teachers made an impact on teacher trainee learning and development. She described the effects through the following themes: how teacher trainees relate to the formal government curriculum; ways that teachers can move beyond a test and drill of skills; and broadening their conceptions of literacy. Like many others who provided these rich experiences, Karly felt these were aspects were transformational. This is a stark contrast to the alternative certification programs discussed earlier which tried to compress the teacher education program into a few weeks.

Rosetta from the UK recognized that the national curriculum had a “big emphasis on phonics.” Like others she felt she needed to teach to the curriculum; however, noted: “[we] do try and get the students to critique that, so we are quite torn with this. We wouldn't advocate the way that it's suggested that it's being taught ... so we're kind of constantly asking them to rethink and think about ways in which they can contextualize what grammar is.” Robert also working in the UK identified yet another issue with the current narrow curriculum. Based on his work and research he found that school children felt that the highly prescriptive curriculum was “dull and they're jumping through hoops. The teachers fall into this pedagogy where they feel that they're doing the right thing [because] they have taken this idea from the government that they must do this.” He felt the narrow, mandated curriculum which was supposed to be focused in their teacher education programs had deleterious effects on the children, teachers, and teacher trainees. Everyone was losing. Most tried to give their teacher trainees a way to understand literacy as a complex social, relevant, and engaging complex process.

The examples listed above are a far cry from an apprenticeship model where teacher trainees learn to teach a script because; rather, the 58 university-based teacher educators in these studies have adopted a very broad approach to their work. The practices and pedagogies of the LTEs described above requires a knowledge of pedagogy of teacher education, a deep understanding of literacy, an in-depth knowledge of classroom teaching, and an awareness that teaching occurs in a very complicated context. They acquired this knowledge (and continue to grow and learn) through advanced educational studies, reflective practice, collaboration with others, reading widely, being involved in multiple communities, and conducting research. It is a life-time work and commitment.

Preparation to teach mandated curriculum: conflicting agendas

Many LTEs experienced a collision between the goals for their courses and the government mandates. Navigating this conflict was not an easy or straight-forward process because there was no “play book” for this kind of teaching. Karly reconciled this conflict because she felt that teacher preparation programs “have a responsibility to make sure students are aware of the government directives and the socio-political context in which they will be walking into.” However, she tempered this position by elaborating, “I do not think it’s the responsibility of teacher preparation programs to prepare students to implement the government’s agenda, standards, curriculum directives.” This is the horn of the dilemma in which many LTEs are caught: acquiescence or non-compliance with the government mandates.

Justin noted that he had lived through five national curricula. Similarly, Elan takes the “long-view” regarding preparing teacher trainees to focus on the current curriculum:

it’s kind of a fool’s errand to prepare them to implement any particular initiative because it’s going to be gone by the time that they’re there. They need to be able to think on their feet, and they need to be able to figure out what part of the initiative was consonant with their own beliefs about literacy development and literacy education, what parts aren’t, and how do you negotiate that before you have tenure.

Many of the LTEs lamented that it is proving to be more difficult to find the time for important and critical conversations. As Karly argued: “It’s important to embed practice within questions of the larger purpose of education and what our beliefs are and what’s most important.”

Without a doubt the LTEs’ contexts were impacting their course goals and pedagogy both overtly and subtly. For example, the oversight by OFSTED was seen as a real and menacing threat. Padua described how he was formally disciplined because the OFSTED reviewers thought he did not place sufficient emphasis on teaching synthetic phonics. LTEs in the US and Australia felt the exit outcome (e.g., [edTPA](#)⁴) caused teacher trainees to focus on what they had to do due to “pass the test” rather than have a longer-term goal of gaining a deep understanding of the teaching and learning process.

Place of and doing research

All LTEs felt that doing research and being knowledgeable about research were part of their work and role. The ability to conduct research requires skills that classroom teachers may not have acquired. All 58 LTEs had completed a Ph.D. where they developed research skills, gained a respect for research, and know how to use research to guide teaching (in both schools and in higher education). Many were part of international research communities which gave them a wider world-view so they could get a sense of what was being done in other countries. Participation in research communities beyond their university or country was described as a turning point in their work as LTEs.

The LTEs research was not kept in a “silo” rather it influenced their practice and identity. Sam spent significant time reading research because “it’s that kind of research that keeps me asking questions about effective practice.” All saw doing research as an essential part of their work. Madelaine describes how being a research affects her teaching:

doing research, both in our lab schools and other schools, how much enriched my understanding of teaching and learning and I have then translated that into my courses and across the whole department because we are integrating a different approach to digital technology across our whole department based on what I learned from my research ... doing research is [integral] to my job and my level of satisfaction.

Conducting research, however, is not neutral. For example, one LTE did critical discourse analyses of literacy policies showing its shortcomings. In some ways their research could be conceptualized as a form of resistance. For example, Padua commented “what worries me is the imposition of neo-liberal ideology, politics, economics, on the educational system” and his research revealed how this is playing itself out in schools. While Tessa who had done significant research on Indigenous students’ classroom experiences found traditional literacy programs were “harmful in the way people are storied as readers and writers impacts them for the rest of their life.” This led to her inviting Indigenous Elders to be guest speakers and in turn develop responsive practices that reflected the needs of her learners.

Tanya in the UK felt that her research opened up new avenues for understanding children and literacy. She described how her work with colleagues led her to question much of her practice:

[The researcher] took tape recorders into the homes and they interviewed the kids. They observed them several times, always in challenging areas. She was just shocked. What she saw shocked her and kind of slapped her around the face. Then when she came back to share it with me and with other teachers it shocked all of us. Everybody was shocked everywhere. Because we just had no bloody idea and we all thought we were trying to do our best.

⁴edTPA is a performance-based, subject-specific assessment and support system used by teacher preparation programs throughout the United States to emphasize, measure and support the skills and knowledge that all teachers need from Day 1 in the classroom.

This experience led to Tanya leading a country-wide initiative on reading which some saw as a very political act.

Research as a form of resistance played out differently for Alicia. “I’ve become more politicized because I’ve seen problems, I’ve seen issues being pushed through and the impact of those issues on children and on my colleagues. So I guess I just became alert through the research definitely because of my reading and also through the leadership role issue.” Having learned significantly from her research she willingly accepted a senior administrative position in her university to try to stem the tide of harmful reforms.

Table 2 provides examples of current research done by the LTEs which shows that their research moved beyond the limited confines of studying the mandated curriculum. They explored relevant timely topics which addressed the complexities of literacy:

These examples show that the LTEs’ research was contributing to a counter-narrative to the governments that are pushing through a narrow view of education. Being consumers and generators of research enhanced their knowledge base; refined their research skills; and influenced their pedagogy of teacher education. The following section describes the influence that conducting research had on their identity. They did not see themselves simply as trainers but as academics.

Identity

There were a number of questions about identity in the interviews (e.g., turning points; term you would use to describe yourself; own strengths). As described earlier, teacher educators must have multiple identities (Kosnik et al., 2013): identify as classroom teachers, researchers, and higher education instructors. Nora described her complex identity as:

I think teacher educators get to be in this amazing liminal space we cannot totally just fly into academic world because these [teacher trainees] are going to be in front of kids, and I want to show them that theory matters, but I need to make sure I’m building the bridge to how theory relates to the practice you enact every day.

These LTEs were not plucked from classrooms to teach a script or to act as mentors to apprentices; rather, they drew on their rich experiences (e.g., teaching in developing nations, working in programs with marginalized youth) when creating their courses. What is significant is not only that they had a wide-range of experiences but also, that for the most part, they examined these experiences in their doctoral programs. For example, Manuela described “a major turning point was just engaging in my Ph. D. in the College of Education. I don’t know if everybody is sort of blown over by this when they start with those early social foundation courses but going like, Oh my god. Education is not a meritocracy.” The doctoral experience also provided them with a set of colleagues with whom many still collaborate. For many their former thesis supervisor and (former) fellow doctoral students were their academic communities which further reinforced their identity as researchers.

The LTEs saw themselves as full-fledged academics believing they have a place in the university. Sam wryly stated that the government “would prefer to call us teacher trainers rather than teacher educators so I would rather be called a teacher educator than a teacher trainer.” He did not see himself as merely a trainer. For many they worked hard in their Ph.D. programs and identify as researchers so they do want to revert to becoming trainers. However, there were pressure points because of the unintended consequences of the current politicization of teacher education. For example, many felt they no longer wanted to teach literacy courses in teacher education because they believed they did not see themselves as trainers nor did they want to teach a mandated curriculum. The LTEs have complex identities because of their work which requires them to straddle both schools and the university. Ciara in Canada said that she feels like a “border crosser” which is demanding. Many described how they often felt like second-class citizens in their university (Kosnik et al., 2013) because they were closely tied to schools and with government conceptualizing and calling them trainers they felt it further marginalized them. As a result, many were opting out of mandatory literacy courses in teacher education preferring to teach electives or doctoral level courses where they had more independence, less scrutiny, and more connection to their full-time colleagues in the university. For some, they had to choose where to focus their work—on teaching or research. For many they valued both but felt forced. Some described it as a “push and pull” because their universities wanted them to increase research productivity while in other cases they had to increase their teaching load. This impacted their identity—to be a researcher or not to be a researcher? Another unintended consequence of the current context was a reduction in fewer full-time LTEs because of

Table 2 Current research being done.

<i>Karly</i>	<i>Impact of student-driven service learning projects</i>
Manuela	Identify centrally important practices or core practices for teaching English language learners
Simona	Help teachers teach literary in ways that move away from formal and academic programs
Katreena	Study of teach for America programs and teacher trainees enrolled in those programs
Giovanni	Place of community-based literacy programs
Pietro	Study literacy experiences of youth who have been incarcerated

budget cuts. To fill the void more and more sessional instructors were teaching the mandatory literacy courses. The LTEs certainly respected and valued the sessionals but there is a difference in identity between sessional/contract instructors and a full-fledged academic.

Doing a Ph.D. had a long-term impact on the LTEs. Rather than being a “one-off” experience it provided them with ways to deepen their knowledge about literacy, helped them recognize the social-political context of literacy learning/teaching, taught them how to read research, provided them the skills to conduct research and ways to use their findings in their teaching, and reinforced their identity as researchers. Having committed a minimum of five years to their doctoral studies it influenced their identity with many calling themselves researchers. Diane proudly stated: “I really value my researcher role.” Their Ph.D. and their on-going research cemented their identity as academics.

Discussion

Certainly, teacher education needs reform but rather than impose draconian reforms that are conceived of by for-profit companies or ill-informed government officials, there needs to be what Kennedy (2016) calls incremental changes. Reforms need to be based on solid research (Ping et al., 2018); teacher educators need to participate in the discussion to determine overall goals for schooling (Goodlad, 1998; Dewey, 1916, 1938).

The stark contrast between the two approaches to teacher education—apprenticeship vs. a fulsome academic intellectual endeavor—show the incredible splintering of teacher education. Teacher education has long been criticized for its lack of uniformity (Shulman, 2005); however, the variance among programs has increased almost exponentially. Is “letting a thousand flowers bloom” a realistic and responsible approach to teacher education? Children are being taught by new teachers who have had vastly different preparation. In rapidly changing times, returning to an apprenticeship model of education defies logic and seems irresponsible. The scant research on alternative certification program (Anderson, 2019; Goldhaber, 2019) suggests these graduates are less prepared than those from full teacher education programs and as Zeichner and Conklin (2016) suggest they often do not remain in the teaching profession. It takes time to learn to be a teacher (regardless of the length of your teacher education program). Teacher educators need to be addressing many key topics: changing demographics of schools, systemic racism, international competition, globalization, climate change, technology-driven economy; and increasing disparities in wealth. Teacher trainees need excellent teacher educators because they will be teaching children who need the highest quality of education that prepares them for a rapidly changing world (e.g., COVID, dominance of social media platforms, politicization of news organization). Teacher educators cannot address this array of issues in a condensed program. They need to be working together to flesh out a pedagogy of teacher education that is powerful and appropriate for the current reality.

When selecting teacher educators, universities need to select teacher educators who have thought deeply about education, have engaged in graduate study, have unpacked their own personal and professional experiences, are willing to be part of an international network of teacher educators, have a wide-range of experiences, are flexible, and are willing to grow.

The conceptualization of teacher educators is essential (trainers or academics) because this has implications for their identity and practices. The presentation of the two studies of LTEs offered compelling evidence that they should be viewed as academics, researchers, and decision-makers. They need extensive pedagogical and intellectual skills to develop courses/modules that address pressing issues (e.g., advancing social justice). The work of the 58 LTEs presented demonstrate a thoughtful and deliberate approach to their teaching. In order to do this high level of teaching, they need preparation to be teacher educators (induction) and on-going support (professional development). Induction and professional development need to be based in research (Goodwin, 2012; Kosnik et al., 2014; Murray and Male, 2005) which delineates what teacher educators need to know and be able to do. Teacher educators who are plucked from classrooms to lead alternative certification programs need an opportunity to grow beyond their own experiences.

There is no evidence that government “reforms” have “worked.” How would they define success? If governments truly want to truly “improve” teacher education they must reallocate financial resources used on testing and supporting alternative certification programs to offering the kind of professional development (Boyd and Harris, 2010; Boyd et al., 2011) for teacher educators outlined by McPhail et al. (2018) and Czerniawski et al. (2018). This would be a clear sign they understand that teacher education is not an apprenticeship but is a complex intellectual/academic endeavor that requires teacher educators who devote serious attention to both theory and pedagogy. Governments of course have a place in education but it should be helpful, enlightened, and collaborative. Sending OFSTED officers to review school of education to check off boxes on a form is an insult and counterproductive to bettering practices. Is there any “proof” that OFSTED reviews have improved our workforce or citizenry? The money allocated for external reviews could be used much more wisely (e.g., support for teacher educators to be involved in international research teams). If governments engage in discussions with teacher educators it would be an opportunity for them to broaden their understandings of teacher education including setting realistic and timely goals for education. Their current myopic view is discouraging for teachers, children, LTEs, and teacher trainees. Teaching teacher trainees to deliver a script or simply use Levov (2010) strategies completely misleads them because it ignores the complex world of education. It is no wonder that new teachers often leave teaching prematurely or experience burnout. When the education being offered is not engaging what is being accomplished?

The question of whether teacher educators should be researchers has been shown to be almost absurd (Cochran-Smith, 2012). Of course, they need to do research on their teaching, on children, in the community, on their graduates, and on families as many of the LTEs did in this study. In turn, they must use their research to inform their teaching. This makes their courses/modules far more

relevant for the 21st century; teacher education courses based on research rather than on a whim or limited by the teacher educators' own classroom teaching experiences would be far more relevant. Being involved in research including international collaborations broadens the LTEs' world-view and increases their research skills. The whole issue of teacher educators themselves as researchers seems to have been overlooked by alternative certification programs because they have limited the work of teacher educators to be implementers and technicians. Before governments invest more heavily in alternative certification programs there needs to be research (done at arm's length) in a rigorous manner to determine if their talking points on their glitzy websites are valid (Bos and Gerdeman, 2017; Mitchell and Romero, 2010). Meanwhile, traditional university-based teacher education programs need to access research on alternative certification programs to determine if there is something to be learned from them. Working in silos or parallel paths undermines efforts.

Realistically, LTEs have very little political clout; however, their research and course construction which were a form of resistance must be continued. By using broad understandings of literacy and offering rich learning experiences they are preparing teacher trainees for the world which they and their students will face. The research on LTEs presented in this article provides examples for others on how to navigate multiple demands placed on them namely teaching and research. The courage shown by LTEs presented in these two studies is inspiring because they truly have worked hard to develop their courses, were focused on the teacher trainees' learning, and went to extraordinarily lengths to make teacher education engaging and relevant. They are interested in teacher education that is not driven by neoliberal reforms of for-profit companies. It is time for courageous leadership. As Roberto in the UK states:

I use three words when I am talking about teaching as a quite basic level with the students, but I believe I bring as well. Its purpose, passion, and empowerment. Those are three key words for me, which I try and show when I teach. I teach with a passion, and you have to have passion with your subjects. There must be a purpose that we're teaching children. Children need to know the purpose, and the purpose is not good enough to be, because the government said or because it's the next objective. And then, we're empowering, I empower the students, and the students have to empower children within their classrooms.

Appendix one

Methodology

For both studies, invitations to participate in the study were sent to a small group of LTEs. This led to "snowball sampling" whereby some LTEs who had accepted the invitation then suggested a colleague who might be relevant for the study. This led to 58 participants who were interviewed using a semi-structured format which took approximately 60–90 min. The same questions were asked of all participants, but probe questions were posed and additional comments were welcomed. Most of the questions were open-ended in that they sought more than a yes/no responses or simple factual answers. Interviews were done either face-to-face or on Skype and were audio-recorded and transcribed.

Interview questions included those on: background (e.g., timeline of turning points, topic of Ph.D. research, experience as a classroom teacher; influences on your current practice; qualities of a teacher educator; academic community; framework and goals for their literacy course(s); pedagogies used and reasons for using them; assignments and readings; and equity and diversity.

Much of the methodology was qualitative as defined by Merriam (2009) and Punch (2014). Qualitative inquiry is justified as it provides depth of understanding and enables exploration of questions that do not on the whole lend themselves to quantitative inquiry. It opens the way to gaining entirely unexpected ideas and information from participants in addition to finding out their opinions on simple pre-set matters. A modified grounded theory approach was used, not beginning with a fixed theory but generating theory inductively from the data using a set of techniques and procedures for collection and analysis.

For data analysis qualitative software NVivo was used. The first level of analysis, "open coding" was used to examine properties of the data by identifying salient words and phrases, relating to the research questions and any other category or theme, which were emerging. To initially generate open codes, the clearest transcripts were analyzed; they were "richly descriptive" and had few disruptions (e.g., phrases transcribed as "XXX" because transcriber could not hear what was being said). During the open coding process, transcripts were first coded by hand. Then, all transcribed interviews were imported into NVivo12 software for analysis "to allow rigorous interrogation of the data". 164 nodes were including sub-nodes were developed.

Approximately 32 nodes (themes) that related to teacher educators' journeys. As the analysis progressed, key findings were identified and refined—adding some and deleting or merging others—through "constant comparison" of the interview transcripts. Through axial coding connections were identified across the nodes. Queries were run to determine connections between goals for literacy courses with pedagogy.

The findings are divided into three sections: goals for their courses; place of and doing research activity; and identity. These were chosen because they address the topics raised in Part One.

References

- Anderson, L., 2019. Private interests in a public profession: teacher education and racial capitalism. *Teach. Coll. Rec.* 121, 38.
 Ball, D.L., Hill, H., 2009. The curious—and crucial—case of mathematical knowledge for teaching. *Phi Delta Kappan* 91 (2), 68–71.

- Bos, H., Gerdeman, D., 2017. Alternative Teacher Certification: Does It Work? <https://www.air.org/resource/alternative-teacher-certification-does-it-work>.
- Boyd, P., Harris, K., 2010. Becoming a university lecturer in teacher education: expert school teachers reconstructing their pedagogy and identity. *Prof. Dev. Educ.* 36 (1–2), 9–24.
- Boyd, P., Harris, K., Murray, J., 2011. Becoming a Teacher Educators: Guidelines for Induction, second ed. ESCalate, Bristol www.escalate.ac.uk.
- Christou, T. (Ed.), 2017. *The Curriculum History of Canadian Teacher Education*. Routledge, New York.
- Cochran-Smith, M., Keefe, E.S., Carney, M.C., 2018. Teacher educators as reformers: competing agendas. *Eur. J. Teach. Educ.* 41 (5), 572–590.
- Cochran-Smith, M., 2003. Learning and unlearning: the education of teacher educators. *Teach. Teach. Educ.* 19 (1), 5–28.
- Cochran-Smith, M., 2012. Composing a research life. *Action Teach. Educ.* 34, 99–110.
- Cuban, L., 1987. Cultures of teaching: a puzzle. *Educ. Adm. Q.* 23 (4), 25–35.
- Czarniawski, G., Gray, D., MacPhail, A., Baid, Y., Conway, P., Guberman, A., 2018. The professional learning needs and priorities of higher education-based teacher educators in England, Ireland and Scotland. *J. Educ. Teach.* 44 (2), 133–148.
- Darling-Hammond, L., Lieberman, S. (Eds.), 2012. *Teacher Education Around the World: Changing Policies and Practices*. Routledge, New York, pp. 66–80.
- Darling-Hammond, L., Chung, R., Frelow, F., 2002. Variation in teacher preparation: how well do different pathways prepare teachers to teach? *J. Teach. Educ.* 53 (4), 286–302.
- Darling-Hammond, L., Holtzman, D.J.H., Gatlin, S.J., Heilig, J.V., 2005. Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Educ. Pol. Anal. Arch.* 13 (42), 1–34.
- Darling-Hammond, L., 2006. *Powerful Teacher Education: Lessons from Exemplary Programs*. Jossey-Bass, San Francisco.
- Dewey, J., 1916. *Democracy and Education*. Macmillan, New York.
- Dewey, J., 1938. *Experience and Education*. Collier-Macmillan, New York.
- Dinkelman, T., Margolis, J., Sikkenga, K., 2006. From teacher to teacher educator: experiences, expectations, and expatriation. *Stud. Teach. Educ.* 2 (1), 5–23.
- Ducharme, E., Ducharme, M., 1996. The development of the teacher education professoriate. In: Murray, F. (Ed.), *The Teacher Educator's Handbook: Building a Knowledge Base for the Preparation of Teachers*. Jossey-Bass, San Francisco, CA, pp. 691–714.
- Ducharme, E., Ducharme, M., (n/d). *Teacher Education: Historical Overview, International Perspective on Teacher Education*. <https://education.stateuniversity.com/pages/2479/Teacher-Education.html>.
- edTPA. <https://www.edtpa.com/>.
- Ellis, V., McNicholl, J., Blake, A., McNally, J., 2014. Academic work and proletarianization: a study of higher education-based teacher educators. *Teach. Teach. Educ.* 40, 33–43.
- European Commission, 2009. *Standards and Guidelines for Quality Assurance in Teacher Education in Europe*. Retrieved from: http://www.enqa.eu/wp-content/uploads/2013/06/ESG_3edition-2.pdf.
- European Commission, 2013. *Supporting Teacher Educators for Better Learning Outcomes*. Retrieved from: http://ec.europa.eu/education/school-education/teacher-cluster_en.htm.
- Florian, L., Pantic, N. (Eds.), 2013. *Learning to Teach: Exploring the History and Role of Higher Education in Teacher Education*. The Higher Education Academy, New York, UK, pp. 10–17.
- Fries, K., Cochran-Smith, M., 2006. Teacher research and classroom management: what questions do teachers ask. In: *Handbook of Classroom Management: Research, Practice, and Contemporary Issues*. Lawrence Erlbaum, Mahwah, NJ, pp. 945–981.
- Furlong, J., 2013. Globalization, neoliberalism, and the reform of teacher education in England. *Educ. Forum* 77 (1), 28–50.
- Gilroy, P., 2014. Policy interventions in teacher education: sharing the English experience. *J. Educ. Teach.* 40 (5), 622–632.
- Goldhaber, D., 2019. Evidence-based teacher preparation: policy context and what we know. *J. Teach. Educ.* 70 (2), 90–101.
- Goodlad, J., 1998. *Educational Renewal: Better Teachers, Better Schools*. Jossey-Bass, San Francisco.
- Goodwin, A.L., Kosnik, C., 2013. Quality teacher educators = quality teachers? Conceptualizing essential domains of knowledge for those who teach teachers. *Teach. Dev.* 17 (3), 334–346.
- Goodwin, L., 2008. Defining teacher quality: is consensus possible? In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D.J., Demer, K. (Eds.), *Handbook of Research on Teacher Education: Enduring Questions in Changing Contexts*, third ed. Routledge, New York, pp. 399–403.
- Goodwin, L., 2012. Teaching as a profession: are we there yet? In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, New York, pp. 44–56.
- Hokka, P., Etelapelto, A., 2014. Seeking new perspectives on the development of teacher education: a study of the Finnish context. *J. Teach. Educ.* 65 (1), 39–52.
- Houston, S., Prentice, A., 1988. *Schooling and Scholars in Nineteenth-Century Ontario*. University of Toronto Press, Toronto.
- Info-Ted, 2019. <https://www.emeraldinsight.com/doi/abs/10.1108/S1479-368720150000022015>.
- Kennedy, M., 2006. Knowledge and vision in teaching. *J. Teach. Educ.* 57 (3), 205–211.
- Kennedy, M., 2016. Parsing the practice of teaching. *J. Teach. Educ.* 67 (1), 6–17.
- Kinloch, V., Penn, C., Burkhard, T., 2020. Black lives matter: storying, identities, and counternarratives. *J. Literacy Res.* 52 (4), 382–405, p. 387.
- Korthagen, F., 2010. Situated learning theory and the pedagogy of teacher education: towards an integrative view of teacher behavior and teacher learning. *Teach. Teach. Educ.* 26, 98–106.
- Kosnik, C., Menna, L., Dharamshi, P., Miyata, C., Beck, C., 2013. A foot in many camps: literacy teacher educators acquiring knowledge across many realms and juggling multiple identities. *J. Educ. Teach.* 39 (5), 534–540.
- Kosnik, C., Dharamshi, P., Menna, L., Miyata, C., Cleovoulou, Y., Beck, C., 2014. Beyond initial transition: an international examination of the complex work of experienced literacy/English teacher educators. *Engl. Educ.* 48 (1), 41–62.
- Kosnik, C., Dharamshi, P., Menna, L., Miyata, C., Cleovoulou, Y., Beck, C., 2015. You teach who you are: the experiences and pedagogies of literacy/English teacher educators who have a critical stance. In: Lambert, J., Burnett, B. (Eds.), *Teacher Education for High Poverty Schools*. Springer, Rotterdam, pp. 131–151.
- Kosnik, C., Beck, C., Goodwin, L., 2016. Reform efforts in teacher education. In: Loughran, J., Hamilton, M.L. (Eds.), *Handbook on Teacher Education*. Springer Academic Publishers, Dordrecht, pp. 207–224.
- Kosnik, C., Menna, L., Dharmashi, P., 2017. So how do you teach literacy in teacher education? Literacy/English teacher educators' goals and pedagogies. *Aust. J. Lang. Literacy* 40 (1), 59–72.
- Kosnik, C., Dharamshi, P., Menna, L., 2020. Displaced academics: intended and unintended consequences of the changing landscape of teacher education. *Eur. J. Teach. Educ.* <https://doi.org/10.1080/02619768.2020.1793947>.
- Labaree, D., 2000. On the nature of teaching and teacher education: difficult practices that look easy. *J. Teach. Educ.* 53 (3), 228–233.
- Labaree, D., 2006. *The Trouble With Ed Schools*. Yale University press.
- Ladson-Billings, G., 2005. Is the team all right? Diversity and teacher education. *J. Teach. Educ.* 56 (3), 229–234. <https://doi.org/10.1177/0022487105275917>.
- Lambert, J., Burnett, B. (Eds.), 2015. *Teacher Education for High Poverty Schools*. Springer, Rotterdam, pp. 131–151.
- Lemov, D., 2010. *Teach Like a Champion: 49 Techniques that Put Students on the Path to College*. Jossey Bass, San Francisco, CA.
- Lewis, C., Enciso, P., Moje, E., 2020. *Reframing Sociocultural Research on Literacy*. Routledge, New York.
- Literacy Hour. <http://www.educationengland.org.uk/documents/literacytaskforce/implementation.html>.
- Louden, L., 2008. 101 Damnations: the persistence of criticism and the absence of evidence about teacher education in Australia. *Teach. Teach. Theor. Pract.* 14 (4), 357–368.
- Loughran, J., Hamilton, M.L., 2016. Developing and understanding of teacher education. In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education*, vol. 2. Springer, Singapore, pp. 3–22.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education: Understanding Teaching and Learning About Teaching*. Routledge, London, New York.
- Loughran, J., 2014. Professionally developing as a teacher educator. *J. Teach. Educ.* 65 (4), 271–283.

- Lunenberg, M., Dengerink, J., Korthagen, F., 2014. *The Professional Teacher Educators—Role, Behavior and Professional Development of Teacher Educators*. Sense Publishers, Rotterdam.
- MacPhail, A., Ulvik, M., Guberman, A., Czerniawski, G., Oolbekkink-Marchand, H., Bain, Y., 2018. The professional development of higher education based teacher educators: needs and realities. *Prof. Dev. Educ.* <https://doi.org/10.1080/19415257.2018.1529610>.
- Mayer, D., Mitchell, J., Santoro, N., White, S., 2011. Teacher educators and “accidental” careers in academe: an Australian perspective. *J. Educ. Teach.* 37 (3), 247–260.
- Mayer, D., 2014. The appropriation of the professionalization agenda in teacher education. *Res. Teach. Educ.* 4 (2), 39–44.
- McKeon, F., Harrison, J., 2012. Developing pedagogical practice and professional identities of beginning teacher educators. *Prof. Dev. Educ.* 36 (1–2), 5–44.
- McNamara, O., Murray, J., 2013. The school direct program and its implications for research-informed teacher education and teacher educators. In: Florian, L., Pantic, N. (Eds.), *Learning to Teach: Exploring the History and Role of Higher Education in Teacher Education*. The Higher Education Academy, New York, UK, pp. 10–17.
- Menna, L., 2013. Times are changing and you’ve got to keep up: negotiating multiple literacies within the context of teacher education. In: Kosnik, C., Rowsell, J., Williamson, P., Simon, R., Beck, C. (Eds.), *Literacy Teacher Educators: Preparing Student Teachers for a Changing World*. Sense Publishers, Rotterdam, pp. 17–30.
- Milner, R., Laughter, R., Judson, C., 2015. But good intentions are not enough: preparing teachers to center race and poverty. *Urban Rev.* 47 (2), 341–363.
- Mitchell, D., Romero, L., 2010. The politics and practice of alternative teacher education. *Educ. Adm. Q.* 46 (3). <https://doi.org/10.1177/0013161X10375272>.
- Moss, G., 2009. The politics of literacy in the context of large-scale education reform. *Res. Pap. Educ.* 24 (2), 155–174.
- Moss, G., 2017. Assessment, accountability and the literacy curriculum: reimagining the future in the light of the past. *Literacy* 51 (2), 56–64.
- Murray, J., Male, T., 2005. Becoming a teacher educator: evidence from the field. *Teach. Teach. Educ.* 21 (2), 125–142.
- Murray, J., Passy, R., 2014. Primary teacher education in England: 40 years on. *J. Educ. Teach.* 40 (5), 492–506.
- Murray, J., Swennen, A., Kosnik, C., 2019. *International Research, Policy and Practice in Teacher Education: Insider Perspectives*. Springer Academic Publishers, Dordrecht.
- Murray, J., 2016. Beginning teacher educators: working in higher education and schools. In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education*, vol. 2. Springer, Singapore, pp. 35–70.
- National Education Association, (n/d). Research Spotlight on Alternate Routes to Teacher Certification. Retrieved from. <http://www.nea.org/tools/16578.htm>.
- New Schools Venture Fund. <https://www.newschools.org/>.
- Office for Standards in Education, Children’s Services and Skills (OFSTED), <https://www.gov.uk/government/organisations/ofsted>.
- Palincsar, A.S., 1998. Social constructivist perspectives on teaching and learning. *Annu. Rev. Psychol.* 49, 345–375.
- Peters, M., Cowie, B., Menter, I. (Eds.), 2017. *Springer Companion to Teacher Education Research*. Springer Academic Publishers, Dordrecht.
- Ping, C., Schellings, G., Beijaard, D., 2018. Teacher educators’ professional learning: a literature review. *Teach. Teach. Educ.* 75, 93–104.
- Rosales, E., 2019. Opportunities for Teacher Professional Learning: Two Case Studies of Experienced Teachers in Ontario, Canada. <https://tspace.library.utoronto.ca/handle/1807/65532>.
- Russ, R.S., Sherin, B.L., Gamoran Sherin, M., 2016. What constitutes teacher learning? In: Gitomer, D.H., Bell, C.A. (Eds.), *Handbook of Research on Teaching*, fifth ed. American Educational Research Association, Washington, D.C., pp. 391–438.
- Sahlberg, P., 2012a. <https://pasisahlberg.com/text-test/>.
- Sahlberg, P., 2012b. http://www.akeducate.com/media/pdf/Sahlberg_Presentation.pdf.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15 (2), 4–31.
- Shulman, L., 2005. Teacher Education Does Not Exist Stanford Educator. <http://www.ed.stanford.edu/suse/news-bureau/educator-newsletter.html>.
- Souto-Manning, M., Stillman, J., 2020. In the pursuit of transformative justice in the education of teacher educators. *N. Educat.* 16 (1), 1–4. <https://doi.org/10.1080/1547688X.2019>.
- Tack, H., Vanderlinde, R., 2014. Teacher educators’ professional development: towards a typology of teacher educators’ researcherly disposition. *Br. J. Educ. Stud.* 62 (3), 297–315.
- Vygotsky, L., 1986. *Thought and Language*. MIT Press, Cambridge, MA.
- White, S., 2018. An insider look at the implications of “partnership” policy for teacher educators’ professional learning: an Australian perspective. In: Murray, J., Swennen, A., Kosnik, C. (Eds.), *International Research, Policy and Practice in Teacher Education: Insider Perspectives*. Springer Academic Publishers, Dordrecht.
- White, S., 2016. Teacher education research and education policymakers: an Australian perspective. *J. Educ. Teach.* 42 (2), 252–264.
- White, S., 2019. Teacher educators for new times? Redefining an important occupational group. *J. Educ. Teach.* 45 (2), 200–213.
- Whitty, G., 2014. Recent developments in teacher training and their consequences for the “University Project” in education. *Oxf. Rev. Educ.* 40 (4), 466–481.
- Williamson, P., 2013. Engaging literacy practices through inquiry and enactment in teacher education. In: Kosnik, C., Rowsell, J., Williamson, P., Simon, R., Beck, C. (Eds.), *Literacy Teacher Educators: Preparing Student Teachers for a Changing World*. Sense Publishers, Rotterdam, pp. 135–148.
- Yandell, J., 2012. Different ways of reading, or just making the right noises? *Appl. Financ. Econ.* 19 (3), 283–294.
- Yandell, J., 2016. The impact of policy on teacher education and literacy education in England: some notes from a corner of a small island. In: Kosnik, C., White, S., Beck, C., Marshall, B., Goodwin, L., Murray, J. (Eds.), *Building Bridges: Rethinking Literacy Teacher Education in a Digital Era*. Sense Publishers, Rotterdam, pp. 31–42.
- Zeichner, K., Conklin, H., 2016. Beyond echo chambers and knowledge ventriloquism: raising the quality of the debate on teacher education. *Teach. Coll. Rec.* 119 (4).
- Zeichner, K., 2005. Becoming a teacher educator: a personal perspective. *Teach. Teach. Educ.* 21, 117–124.

Self-study of teacher education practices: the complex challenges of learning from experience

Tom Russell^a and Rodrigo Fuentealba Jara^b, ^a Faculty of Education, Queen's University, Kingston, ON, Canada; and ^b Facultad de Educación, Universidad Autónoma, Santiago, Chile

© 2023 Elsevier Ltd. All rights reserved.

This is an update of T. Russell, Self-Study by Teacher Educators, In International Encyclopedia of Education (Third Edition), edited by Penelope Peterson, Eva Baker, and Barry McGaw, Elsevier Ltd., 2010, <https://doi.org/10.1016/B978-0-08-044894-7.00677-1>.

Introduction: why study our own practices?	458
Why should a teacher or teacher educator study personal teaching practices?	458
How do teachers and teacher educators learn to teach?	458
Context	459
Outline of the argument	459
Self-study as a methodology for the improvement of teacher education	460
What is self-study and how does one begin?	461
Tom's personal account of self-study of practice	461
Rodrigo's personal account of self-study of practice	462
The impact of self-study on Rodrigo's perspectives	463
REDFORMA: introducing self-study to another culture	463
The role and value of critical friendship in self-study	463
Critical friendship should be mutual	464
Resources for learning more about self-study of teacher education practices	466
Summary and conclusion	466
"You can't fix what you can't see"	466
Acknowledgment	467
References	467
Further reading	467

Introduction: why study our own practices?

Why should a teacher or teacher educator study personal teaching practices?

Many answers to this question can be found in the extensive literature of action research. The focus on self-study of teacher education practices dates to the early 1990s. There is very little in the teacher education literature about the fact that when teacher educators are teaching those who are learning to teach, they are modeling teaching as they educate about how and why we should teach in particular ways. In our view, this is one of the most compelling reasons why teacher educators should study their own teaching practices. Tom recalls well the day that a teacher candidate—the term we use here for those studying in a preservice teacher education program—asserted that professors teaching in a faculty of education should be the best teachers at the university (and he felt that this was not the case in his program).

In 2020, Self-Study of Teacher Education Practices (S-STEP) is approaching 30 years as a named field of research. In addition to countless academic publications and conference papers, the field has its own journal (*Studying Teacher Education*) begun in 2005 and has published not one but two international self study handbooks of teaching and teacher education practices (Loughran et al., 2004; Kitchen et al., 2020).

How do teachers and teacher educators learn to teach?

To many students, teaching looks easy (Labaree, 2000); first-year teachers soon discover otherwise. To many teachers, teaching teachers looks easy (once you have mastered teaching in a school). But there is much to teaching teacher candidates that is not obvious to those who have never done it. Lortie (1975) offered an essential insight when he coined the now-famous phrase, "apprenticeship of observation" (p. 61). In essence, every student learns to teach by observing numerous teachers over more than 15 years of experience in the student role. In our analysis, one key point is that students have learned what teachers do but they have not learned how teachers think or why they teach as they do. The apprenticeship is partial and incomplete in critical ways. Lortie put the problem clearly when he observed that all students, including those in pre-service teacher education programs, typically lack a great deal of information about why teachers act as they do:

Students... are not privy to the teacher's private intentions and personal reflections on classroom events. Students rarely participate in selecting goals, making preparations or postmortem analysis. Thus they are not pressed to place the teacher's actions in a pedagogically oriented framework. (p. 62)

It is improbable that many students learn to see teaching in an ends-means frame or that they normally take an analytic stance toward it. Students are undoubtedly impressed by some teacher actions and not by others, but one would not expect them to view the differences in a pedagogical, explanatory way. What students learn about teaching, then, is intuitive and imitative rather than explicit and analytical; it is based on individual personalities rather than pedagogical principles. (p. 62)

Lortie continues by explaining how an apprenticeship of observation falls far short of a true apprenticeship:

The student's learning about teaching, gained from a limited vantage point and relying heavily on imagination, is not like that of an apprentice and does not represent acquisition of the occupation's technical knowledge. It is more a matter of imitation, which... becomes tradition. It is a potentially powerful influence which transcends generations, but the conditions of transfer do not favor informed criticism, attention to specifics, or explicit rules of assessment. (p. 63)

Every teacher and teacher educator has experienced thousands of hours of observation of what teachers do. Against this unanalyzed but powerful background, a teacher education program offers a much smaller number of hours in which teacher educators present the content of educational theory and research as well as practical advice, typically without addressing how much has already been learned about teaching. This is not only a challenge for those learning to teach. It is also a challenge for teacher educators who may have extensive experience in both school and university contexts but who have never analyzed how their teaching habits were shaped by experience. Against this background reality, we see self-study of teacher education practices as critical for all teacher educators in all cultures.

When Vick (2006) provided an historical analysis of the place of the practicum in teacher education in England and Australia since 1900, he reported that despite decades of reports and calls for change, the *assumptions* underlying teacher education practices have shown little change over time:

The analysis of the history of the practicum in teacher preparation indicates how firmly contemporary proposals are embedded within the framework of assumptions and approaches that have shaped teacher education for over a century. (p. 181)

Our own self-study explorations of assumptions underlying our own teaching confirm the significance of Vick's message. We must study not only our practices but also the assumptions about learning implicit in those practices as we seek to improve our teacher education practices.

Context

We write as teacher educators located in two different countries with distinctly different languages and educational cultures. Our critical friendship has evolved from an initial meeting in Santiago in 2010; Tom has since visited Chile another 16 times, while Rodrigo has visited Canada four times. As we write in 2020, we share a strong commitment to the importance of teacher educators studying their own practices in their teacher education classrooms and in practicum supervision. Tom was a founding member of the Self-Study of Teacher Education Practices (S-STEP) special interest group within the American Educational Research Association in 1993. Rodrigo has developed a strong commitment to teacher educators' learning from their own experiences and in introducing the concept of self-study to teacher educators in a range of universities in Chile. Since 2015, Rodrigo has collaborated with the Chilean office of the Organization of Spanish-American States (Organización de Estados Iberoamericanos, OEI) for that purpose.

Outline of the argument

Authors' instructions for this article included the following statement: "This edition aims to include contributions from traditions, regions and people that have in the past been marginalized, promoting debates about the global politics of knowing, learning, teaching and governing." Authors were also invited to include a summary of recent personal research. In this article, we are guided by both suggestions. In the spirit of a more global perspective, we explore the challenges of introducing a concept that has evolved in one culture into a quite different culture. In keeping with the international character of the encyclopedia and this entry, three tables provide translations into Spanish for the benefit of Spanish-speaking readers. One of our first challenges was recognizing that the literal translation of *self-study* into Spanish as *autoestudio* is inappropriate and misleading; we use *self-study* in both languages as shorthand for self-study of teacher education practices.

We begin with a section about self-study as a methodology for improving the quality of teacher education. This is followed by further discussion of self-study and how it begins. To illustrate how self-study might begin, we provide personal accounts of our own experiences of self-study. Being more recent, Rodrigo's account is more detailed and includes a bilingual account of the impact of self-study on his perspectives as a teacher educator. Not long after learning about self-study, Rodrigo found support for creating

a network that would introduce other teacher educators in Chile to self-study, and details are provided in a second bilingual presentation of his learning from experience. The argument continues with a discussion of critical friendship based on what we have learned together from our extended period of friendship as teacher educators. The argument closes with some additional resources for exploring self-study and with a summary and conclusion.

Self-study as a methodology for the improvement of teacher education

Tom retired in 2019 after 42 years as a teacher educator in Canada; his passion for teacher education continues. As of this writing in 2020, Rodrigo has 27 years of experience as a teacher educator, including several years as Dean of a Faculty of Education in Chile. In Canada and in Chile, as in many other countries, appeals for the improvement of teacher education seem ever-present, yet teacher education continues to qualify as a field in which “the more things change, the more they stay the same,” or as Donald Schön once put it, “the more things stay the same, the more you know that there are deliberate efforts being made to keep them that way” (Unpublished lecture, Queen’s University, October 1984).

Even when pressed, many teacher candidates find it difficult to describe how and why they went about selecting goals, preparing for their teaching and then analyzing the effects of their teaching. As a case in point, pre-service candidates may seem to understand what we refer to as *reflection* but they are often challenged when trying to analyze and develop new frames for what they are doing in the practicum classroom. The innovative teaching approaches that teacher educators urge prospective teachers to consider and that they believe could improve what happens in schools have emerged from and are supported by the extensive research over the last 40 years on how people learn. Research tells us that people learn best when they are active, challenged and engaged. Research reminds us that all students come to us with prior knowledge in various subject areas. The largely invisible apprenticeship of observation generates the prior knowledge and beliefs that teacher candidates bring with them, but we as teacher educators are not necessarily adept at recognizing, exposing and responding to that knowledge and those beliefs and adapting our teaching accordingly. As already noted, we have probably not done the same ourselves.

We risk suggesting the possibility that teacher educators themselves may not understand fully why they teach as they do and what the full range of effects of their teaching may be. “Learning how to think and act in ways that achieve one’s intentions is difficult, particularly if knowledge is embedded in the practice itself.” (Darling-Hammond, 2006, p. 37). Like teachers, teacher educators do learn from the experience of their teaching, but the quality of learning from experience is not easily evaluated. Hence the need for and the importance of a methodology for studying one’s own practices as a teacher educator. Self-study of teacher education practices is not a study of the self; rather, it is the study of the teacher education practices of the individual who enacts those practices.

Consider the following critique of the state of teacher professional development and also consider whether similar comments might be made about the preparation of teacher educators to work in a field that is uniquely different from teaching the content of the school curriculum:

The typical preservice program is a collection of unrelated courses and field experiences. Most induction programs have no curriculum, and mentoring is a highly individualistic process. Professional development consists of discrete and disconnected events. Nor do we have anything that resembles a coordinated system. Universities regard preservice preparation as their purview. Schools take responsibility for new teacher induction. Professional development is everybody’s and nobody’s responsibility.

Feiman-Nemser (2001, p. 1049).

Teacher education is far more complex than it is often taken to be, if we consider not only the short period of professional preparation relative to other professions, but also the challenge of teaching someone how to do the very activity that one is engaged in with them. Labaree (2000) described the complex challenges of teaching in these memorable words:

We ask teacher education programs to provide ordinary college students with the imponderable so that they can teach the irrepressible in a manner that pleases the irreconcilable, and all without knowing clearly either the purposes or the consequences of their actions. Is it any wonder that these programs are not seen as smashing successes?... In addition, they face a situation in which the profession of teaching is generally seen to be relatively easy. And this perception is not simply characteristic of the untutored public; it is also endemic among teacher candidates. (p. 231)

When self-study of teacher education practices arrived on the landscape of teacher education, it was with a strong sense that teacher educators themselves should look within. We too learned from an Apprenticeship of Observation, for “We teach as we were taught” tends to be as true for teacher educators as it is for teachers and those learning to teach. One of the first observations that teacher candidates may make in their education courses concerns the teacher educator who uses one teaching method (often transmission-based) while advocating for another (often inquiry-based). If teacher educators cannot “practice what they preach” and enact their own values and understandings, what incentive is there for teacher candidates to attempt to do so? Teacher educators and the programs in which they teach are unlikely to excel in the preparation of teachers without considering their own development and how they arrived at and might improve their current practices.

What is self-study and how does one begin?

In the context as teacher education, self-study is research in one's own setting of practice to understand and reduce the gap between our good intentions as teacher educators and the actual learning of pre-service teacher candidates. In our view, three related steps are essential:

1. Self-study requires being willing to make fundamental changes in one's perspective on personal teaching practices. This includes openness to surprising, puzzling and unexpected events that can lead to *reframing* and what we call *repracticing*—a term that parallels reframing and reminds us that new ways of viewing practice should lead to new ways of enacting practice.
2. Self-study requires seeing those learning to teach not as empty vessels to be filled with knowledge they lack but as experienced students with extensive but incomplete and potentially misleading knowledge of teaching that must be explored and reinterpreted as part of becoming a teacher.
3. Self-study requires recognition that the teaching practices a teacher educator models for teacher candidates can send messages more powerful than the content knowledge of teaching being developed. Thus it is important for the self-study researcher to create and lead in-class opportunities for discussing and explaining the rationales for specific ways of teaching

Broadly speaking, self-study begins with discomfort and dissatisfaction with the view that learning to teach is a straightforward matter of learning how to teach by being told and then moving to practicum settings to practice what one has already learned. Self-study becomes a powerful research strategy when a teacher educator is willing to explore the complexities of learning to teach in the context of exploring simultaneously the complexities of one's own teaching. Self-study begins by attending to events, experiences and interactions that had previously been ignored, not because they were irrelevant but because they were taken for granted and accepted as familiar and unproblematic: "You can't fix what you can't see." In the next two sections we present personal accounts of our own self-study experiences.

Tom's personal account of self-study of practice

When Tom began to teach pre-service science teachers in 1977, it had been 10 years since he had taught physics in a school. He had just worked for three years with experienced teachers, who had shown him that reducing the amount of talking they do in class was quite challenging. A reduced amount of talking was not what preservice teachers expected, and he quickly began to listen in various ways to their expectations. He also observed many of his students in their practicum placements, where he saw that the practicum is not the place where they put what they had recently been taught into practice. Schön's (1983) *The Reflective Practitioner* offered the perspective he needed to move him forward. Three lectures by Schön at Queen's University in 1984 focused his attention on reflection-in-action—reframing in response to surprising, unexpected or puzzling comments by students.

As the word "reflection" swept into teacher education programs everywhere, Tom had to find ways to have his students focus on "reflection-in-action" in the practicum and away from reflection about issues raised in program classes. From 1984, he and colleague Hugh Munby used Schön's (1983) perspectives on reframing and reflection-in-action in several nationally funded research projects to better understand how individuals learn to teach. It was perhaps inevitable that he should apply the same perspectives to his own teaching. Tom's entry into self-study was pushed forward by returning to the secondary school physics classroom in 1991 and again in 1992 to teach one 75 min physics lesson every day for a full semester after a 25-year absence from the secondary school.

Self-study has enabled me to understand my teaching and my students as never before. When the term "self-study" was gaining attention and interest, I had taken myself back to the secondary school physics classroom to better understand the work I was helping people learn. Now I had a better conceptual framework for viewing both my own development as a teacher and the development of those learning to teach. I also began teaching an M.Ed. course in action research, and through successes and failures, self-study became the key to better understanding.

Hamilton et al. (1998, p. 3).

I can now credit self-study with enabling me to develop the following pedagogical principles that were present in my teaching in 1997–1998 but very absent in 1992–1993.

- Build on and trust in their extensive early teaching experiences.
- Build a community in my classroom.
- Every class includes a significant experience, with implications for how we teach.
- Gather backtalk [students' comments about the quality of their learning] often and distribute it freely.
- Identify a few central themes, post visual cues to each, and link to them as often as possible.
- Replace lectures with notes on a website; use classroom time for experiences and discussions.
- Encourage and reward self-directed learning.
- At the end of the course, ask students what they think you learned from them (Hamilton et al., 1998, p. 3).

As critical friends, Tom and Rodrigo try to remind each other of their previous work. Rodrigo recently suggested that Tom include the preceeding statements (see Table 1) that have emerged from his efforts to study his own teaching of future teachers. They represent his advice to himself and to others who might be interested.

Table 1 Suggestions for teacher educators' relationship to those they teach.

Go easy on the issue of attendance. They pay for our classes. If your teaching doesn't attract them, low attendance will tell you that you need to change, unless you command attendance. A student in a chair is no guarantee of any learning happening.	Con el tema de la asistencia. Ellos pagan nuestras clases. Si su enseñanza no los atrae, la baja asistencia le dirá que necesita cambiar, a menos que usted exija la asistencia. Un estudiante desinteresado no es garantía de que algo suceda más en el futuro.
Teacher candidates are adults, not children. Interact with them as adults; for years it was common in this building to refer to them as kids. They are younger than we are, and most lack experience outside schools, but they deserve our respect.	Los candidatos a profesores son adultos, no niños. Interactúe con ellos como adultos; por años fue común en estos espacios referirse a ellos como niños. Ellos son más jóvenes que nosotros, y tienen menos experiencia fuera de las escuelas, pero ellos se merecen nuestro respeto.
Listen to TCs' experiences in your first class after each practicum placement. Find out how their practicum experiences have changed their sense of what they need to learn from you in your course.	Escuche las experiencias de los candidatos a profesores cuando cuenten sus experiencias al regresar de su práctica. Entérese como sus experiencias de práctica han cambiado el sentido de lo que ellos necesitan aprender de Ud.
I would venture that 50% of what TCs learn from you involves what they learn by watching you teach and experiencing its impact on their learning. They are here to learn to teach, and so they watch our teaching moves far more closely than they watched previous professors. They are frustrated by teaching moves that contradict our words about teaching. Our words and our tune MUST go together consistently.	Me atrevería a decir que el 50% de lo que los candidatos a profesores aprenden de usted implica lo que aprenden al verlo enseñar y experimentar su impacto en el aprendizaje de ellos. Están aquí para aprender a enseñar, y por lo tanto observan nuestros movimientos de enseñanza mucho más de lo que observaron a profesores anteriores. Se sienten frustrados por los movimientos de enseñanza que contradicen nuestras palabras acerca de la enseñanza. Nuestro discurso y nuestra acción DEBEN ir juntos de manera consistente.
Mind the experience gap. We have teaching experience and they do not. It is very difficult to talk over, around, under or through that gap. People with experience think differently about teaching (as in not needing formal lesson plans, etc.) that those who are just beginning (Russell, 2018).	Tenga en mente la brecha de experiencia. Nosotros tenemos experiencia y ellos no. Es muy difícil hablar por encima, alrededor, debajo o a través de esa brecha. Las personas con experiencia piensan de manera diferente respecto de la enseñanza (como no necesitar planes formales de lecciones, etc.) que aquellos que recién están comenzando

In the 2018–2019 academic year, Tom conducted a final self-study of his last year as a teacher educator, introducing three new practices and studying their effects. The practices included replacing *theory* and *practice* with *book knowledge* and *craft knowledge* (following Hagger and McIntyre, 2006), modeling cognitive principles identified by Willingham (2009), and using the last 15 min of every 2-h class for a metacognitive conversation about what we had just learned and how we learned it. With ethical clearance from the university, all 14 students agreed to participate and provided valuable comments on the impact of those new practices on their learning (Russell et al., 2020). Self-study generated a memorable conclusion to a teacher educator's career.

Rodrigo's personal account of self-study of practice

In 1992, a teacher educator whom Rodrigo was assisting at the time asked him, "Have you thought about what working in teacher education means?" and thus began his efforts to answer that question. As he came to know other teachers, Rodrigo understood that knowledge of the subject was not enough for the students, and he began to pay more attention to his students' comments and to questions such as "Teacher, why did you do that? How do you know what he is going to do? How did you know when to respond?" Getting in touch with the questions and concerns of the students, as well as paying attention to his own practices, allowed him to discuss what until then was the mantra: "One can't teach what one doesn't know." Meeting Tom and developing a critical friendship encouraged actively listening to students' questions, clarifying his decisions for his students, and continuously analyzing his own practices with another teacher educator. The self-study perspective has also allowed him to understand that analysis of his own practices, and those of future teachers, refers not only to the method used but also to the assumptions that underlie teaching actions. Those who do not make themselves explicit in front of others will not have the opportunity to better understand their teaching and to model for future teachers the methods and importance of that understanding.

Self-study has allowed Rodrigo to understand the following points:

- Future teachers have a lot of experience as observers of teachers, so they need spaces for exploration in order to identify the assumptions that accompany their observations of teachers.
- I need to promote spaces for real practice by future teachers, not time for more observations.
- To the extent that I make my professional decision-making explicit in my classes, I can enter into better interaction with the students, allowing them in turn to give an account of the ways in which they learn.
- To the extent that I generate instances for listening to the voice of my students, I can better understand the decisions that accompany their practices.

- To the extent that a climate of trust and professional exchange is generated, it is possible to collect more definitive evidence of their learning than through a test. For example, asking for anonymous tickets-out-of-class about their learning has allowed me to collect valuable information regarding the quality of their learning, and this allows me to analyze their learning in relation to my practices.
- Observing a critical friend's practices helps me to identify the assumptions that have guided my professional practices.

The impact of self-study on Rodrigo's perspectives

In [Table 2](#), Rodrigo describes his interests in self-study, how his practices have changed since he became interested, and why he believes self-study of teacher education practices could be valuable to educators and education in Chile.

REDFORMA: introducing self-study to another culture

In an article prepared for an international encyclopedia, it is essential to address the issue of introducing an innovation into other cultures. Self-study of Teacher Education Practices appears to have spread relatively easily across many English-speaking countries and it has attracted interest in countries where teacher educators are comfortable working in English. It is important to note that the official S-STEP membership is steady but small and self-study is a minor but significant methodology in the educational research literature. It is always necessary to pose questions about cultural similarities and differences before attempting to introduce self-study into another academic culture as an unfamiliar research methodology and literature.

At Rodrigo's initiative, the introduction of self-study of teacher education practices into the academic culture in Chile began in 2015 and is continuing to attract interest. The initial experiences were understandably hesitant but by 2020 they have evolved and publications of self-study research are beginning to appear. [Table 3](#) provides an overview of the first five years of experience; additional details are provided by [Fuentelba et al. \(2017, 2020\)](#).

The preceding accounts of our own self-studies and of the challenges of introducing self-study into another culture are intended to illustrate the many different ways in which self-study of teacher education practices can add important new dimensions to the work of teacher educators, potentially leading to improvement not only in teacher education classrooms but also in teacher education programs. We turn next to the importance of critical friendship in self-study.

The role and value of critical friendship in self-study

Even though self-study is a study of one's own actions as a teacher educator, it is not an activity best done on one's own. [Loughran and Northfield \(1998\)](#) explained the rationale for and the importance of collaboration during self-study:

The learning in self-study is intensely personal, but self-study itself requires collaboration.... In this approach, the intensely personal aspects of the study that might otherwise be simply accepted without challenge or scrutiny are able to be professionally and constructively challenged from within the study itself. (p. 14)

There is always a danger that individuals will interpret situations in ways that reinforce existing perceptions. Genuine study of classrooms is associated with a willingness to consider alternate frames of reference, and colleagues are an important source of ideas and support as the teaching and learning are reviewed. (p. 16).

While some self-studies are carried out in isolation and privacy, it is typically both important and highly productive to identify a collaborator for on-going analysis of one's questions, perspectives, and data during self-study research.

Critical friendship is at the heart of our relationship as two teacher educators working across cultural boundaries. That friendship is centered on trust and a shared desire to help each other develop not only better actions in the teacher education classroom but also better understanding of the assumptions implicit in our actions. Although we are physically separated by a 10 h flight and a potential language barrier, mobile phones and Wi-Fi networks mean that we can be as close as the house next door, thanks to Rodrigo's facility in English. Using a combination of Skype and video recordings, Rodrigo observed most of Tom's physics methods classes in the four-year period 2015 to 2019; he also visited Tom's classes in person for one week in each of those four years. Tom made 10 visits to Chile in the same period and attended several of Rodrigo's classes and interacted with some of his students.

The exchange of visits to each other's classes inspired numerous conversations about teaching practices and provided insights into each other's educational cultures, both pre- and post-secondary. While the differences seemed dramatic at first, we soon realized that at a deeper level the differences were more subtle. In Chile, the basic pattern is that teachers talk and students listen. In Canada, as in other English-speaking contexts, that is still a basic pattern, although there are many superficial variations that can suggest

Table 2 Rodrigo's perspectives on self-study.

<i>¿Por qué mi interés en self study?</i>	<i>Why am I interested in self-study of teacher education practices?</i>
El Self-study permite centrarse en las propias prácticas de enseñanza y, mediante un examen honesto y la colaboración con otros, promover el cambio de valores profesionales y creencias personales con respecto a cómo aprendemos y enseñamos. También el Self-study permite un cuestionamiento más profundo de los supuestos subyacentes a nuestras propias prácticas de enseñanza. Esto tendría un efecto significativo en aquellos que forman parte de un entorno profesional: futuros profesores y colegas. El desarrollo de un Self-study no se centra en teorizar, sino que tiene como objetivo replantear los supuestos en los que se basan las prácticas y cómo las prácticas pueden cambiar y reinterpretarse en un contexto de aprendizaje. El Self-study es una actividad colaborativa, realizada en intercambios profesionales sobre supuestos y cambios en la práctica entre los participantes. El Self-study ve a los futuros profesores como profesionales activos con voz y no como receptores vacíos del proceso de aprendizaje que aceptan acríticamente lo que les dice su profesor	Self-study permits one to focus on one's own teaching practices and, through an honest examination and collaboration with others, to advance the change of professional values and personal beliefs with respect to how we learn and teach. Self-study also permits deeper questioning of the assumptions underlying our own teaching practices. This would have a significant effect on those who are part of one professional environment—future teachers and colleagues. The development of a self-study is not focused on theorizing but aims to reframe the assumptions that one's practices are based on and how practices can change and be reinterpreted in a learning context. Self-study is a collaborative activity, conducted in professional exchanges about assumptions and changes in practice between the participants. Self-study sees future teachers as active professionals with voice and not as empty recipients of the learning process who accept uncritically what their teacher tells them.
¿Cómo han cambiado mis prácticas? Al cederles la voz a los futuros profesores, puedo crear un ambiente de confianza profesional para el intercambio de prácticas. Soy más consciente de la importancia de explicar a mis alumnos las razones que enfatizan mis prácticas de enseñanza. Esta es una manera de modelar el proceso de aprender a pensar como maestro. Compartir mis prácticas con un amigo crítico ha sido fundamental, porque he podido aprender lo siguiente: ver mis propias prácticas desde diferentes perspectivas, desafiar mis suposiciones, hacerme nuevas preguntas y desafiarme a mí mismo a cambiar mis prácticas de enseñanza. El proceso de reenmarcamiento [reframing] de mis prácticas a nivel de supuestos respecto a enseñar y aprender no sólo tiene importancia a nivel declarativo, sino que también implica asumir riesgos con nuevas prácticas (repracticing).	How have my practices changed? By giving voice to future teachers, I am able to create an environment of professional confidence for the exchange of practices. I am more aware of the importance of explaining to my students the reasons that underlie my teaching practices. This is a way of modeling the process of learning to think like a teacher. Sharing my practices with a critical friend has been central, because I have been able to learn the following: to look at my own practices from different perspectives, to challenge my assumptions, to ask myself new questions, and to challenge myself to change my teaching practices. The process of reframing my practices in terms of assumptions about teaching and learning not only has significance at a declarative level but also implies taking risks with new practices. We call this repracticing, as it follows reframing.
¿Porqué se debe introducir el Self Study en Chile? En la formación de docente, en Chile, la tradición predominante asume que la práctica docente es un lugar para aplicar lo que previamente se ha aprendido en las aulas universitarias. Debido a que la práctica ha sido vista como un espacio para replicar los modelos previamente enseñados en la universidad, debemos replantear nuestros supuestos para ver que la práctica docente es un escenario para el propio aprendizaje y no sólo un lugar para replicar los modelos de otros. Porque en Chile, la concepción de la complejidad de los procesos de enseñanza asume que un docente novato debe abordar el proceso de pasantía de manera muy paulatina a pesar de haber observado a los docentes durante muchos años. Por lo tanto, las actividades en la escuela comienzan primero con la observación, luego se expanden a pequeñas intervenciones de enseñanza y, mucho más tarde, se hacen responsables de una clase completa.	Why should self-study be introduced in Chile? In teacher education in Chile the prevalent tradition assumes that teaching practice is a place to apply what has been learned previously in university classrooms. Because practice has been seen as a space for replicating the models previously taught in the university, we must reframe our assumptions to see that teaching practice is a setting for one's own learning and not just a place for replicating the models of others. Because in Chile the conception of the complexity of the teaching processes assumes that a novice teacher must approach the internship process very gradually despite having watched teachers for many years. Thus activities in school begin first with observation, then expand to small teaching interventions, and much later taking responsibility for a complete class.

otherwise. Over the last 20 years, Tom's practices have emphasized listening to his students (Cook-Sather, 2002) and Rodrigo found that to be an intriguing practice that he has initiated with his own students. While students may at first respond hesitantly, they soon realize that he is serious about listening to them and move on to embrace the openness that develops.

Critical friendship should be mutual

As Loughran and Northfield (1998) suggested, it is often recommended that a teacher educator identify a critical friend to assist with a self-study and many articles have been published to illustrate the value of such a friend. Our experience of sharing our teaching as

Table 3 A bilingual account of introducing self-study in Chile.

En 2015 Rodrigo compartió su creciente interés por el Self Study de las prácticas en la formación de profesores con Carolina Hirmas de la Oficina de Santiago de la Organización de Estados Iberoamericanos (OEI). La idea inicial fue crear un espacio alternativo para los formadores de formadores donde compartieran sus prácticas y otras inquietudes. Este espacio se denominó REDFFORMA (Red de Formadores de Formadores; http://www.redfforma.cl). En 2016, se publicaron una serie de artículos de Self Study traducidos del Inglés al Español, para apoyar la nueva red. En Chile, los temas más comunes entre los formadores de profesores dicen relación con los cambios en el currículum de formación y en el apoyo a los futuros profesores relacionado con aspectos relacionados con la teoría y la investigación más que con la práctica. Por lo general, no hay oportunidades para discutir las prácticas formativas o el papel del formador como modelador de prácticas docentes prometedoras. Rodrigo y Carolina comenzaron organizando reuniones entre formadores de futuros profesores que trabajan como supervisores de prácticas. En estas reuniones se compartían sus prácticas explorando preguntas planteadas por Tom, amigo crítico internacional del grupo. Algunas preguntas abordadas: * ¿Porqué quiere cambiar sus prácticas de supervisión? * ¿Qué aspectos de su práctica ha cambiado o quisiera cambiar? * ¿Cómo han respondido sus estudiantes cuando ha introducido una o más prácticas nuevas? * ¿Qué evidencia está recolectando de sus estudiantes? Preguntas como estas animaron a los miembros de REDFFORMA a centrarse en sus prácticas. En un inicio, se constató que es más fácil decirle a los futuros profesores que es importante analizar sus prácticas que analizar las propias prácticas. Los esfuerzos iniciales en el análisis de las prácticas estuvieron más centrados en perspectivas teóricas. Con el apoyo de la OEI, se convocó a la realización de Self Studies colaborativos que involucraran a formadores de docentes de diferentes universidades. Tres proyectos obtuvieron el apoyo de la OEI y Rodrigo y Carolina actuaron como amigos críticos externos de los grupos. En la medida que se trabajaba para hacer de las prácticas formativas el foco central, los formadores de formadores se enfrentaron a tres importantes desafíos: 1. Compartir no sólo las prácticas sino que también los supuestos orientadores de dichas prácticas. 2. Escuchar las voces de los estudiantes en tanto miembros de una comunidad y no como meros receptores de nuestro conocimiento. 3. Ver a Rodrigo y Carolina más como amigos críticos externos que como expertos dando consejo a los grupos. En la medida que crecía el interés entre formadores de formadores chilenos, se convocó a una segunda oportunidad para presentar propuestas. Eso permitió que Rodrigo y Carolina ajustaran su enfoque en función de los tres importantes desafíos identificados. Tom estuvo presente en una gran reunión en que los grupos compartieron sus principales puntos de vista. Esto permitió en 2020 el desarrollo de un número especial de la <i>Revista Iberoamericana de Educación</i>, 82(1) https://rieoei.org/RIE/issue/view/Self_Study Para el período 2020–2021 REDFFORMA apoya a tres nuevos grupos, con nuevas universidades, incluyendo dos formadores de formadores internacionales, una de Colombia y otra de España. Esta situación agrega nuevos desafíos para el Self Study, y específicamente para los amigos críticos, ¿Cómo en el proceso ellos desarrollarán un entendimiento compartido respecto de lo que significa ser formador de futuros profesores?, ¿Cómo cada grupo logrará una sintonía fina respecto del entendimiento sobre sus prácticas si cada uno tiene culturas y significados diferentes.	In 2015 Rodrigo shared his growing interest in self-study of teacher education practices with Carolina Hirmas at the Santiago office of Organización de Estados Iberoamericanos (OEI). The initial idea was to create an alternative space for teacher educators to share their practices and other concerns. That space is named REDFFORMA (Red de Formadores de Formadores: Network of Teachers of Teachers; http://www.redfforma.cl/). In 2016, self-study reports translated from English to Spanish were published to support the new network. In Chile the typical topics among teacher educators include changes in the teacher education curriculum and in support for future teachers related to theory and research rather than to practice. Usually, there are no opportunities to discuss teacher education practices or the role of the teacher educator in modeling promising teaching practices. Rodrigo and Carolina began by arranging meetings of teacher educators who work as practicum supervisors. At these meetings, they would share their practices by exploring questions posed by Tom as the group's international critical friend. Questions included the following: • Why do you want to change your supervision practices? • What aspects of practice will you change or have you changed? • How did your students respond when you introduced one or more new practices? • What evidence are you collecting from your students? Questions such as these encouraged members of REDFFORMA to focus on their practices. At first, they realized that it is easier to tell future teachers that it is important to analyze classroom practices than it is to analyze their own practices. The initial efforts to analyze practices were focused on theoretical perspectives. With support from OEI, a call was issued to develop collaborative self-studies involving teacher educators in different universities. Three projects obtained initial support from OEI and Rodrigo and Carolina served as external critical friends to the groups. As they worked to make their teacher education practices their central focus, the teacher educators encountered three significant challenges: 1. To share not only our practices but also their underlying assumptions 2. To listen to students' voices as members of a community rather than as recipients of our knowledge To see Rodrigo and Carolina more as external critical friends than as experts giving advice to the groups As interest in self-study grew among Chilean teacher educators, a second call for proposals was issued. This provided an opportunity for Rodrigo and Carolina to adjust their approach based on what they learned as they identified the three significant challenges. Tom was present for a large gathering in which groups shared their principal insights. This led in 2020 to the development of a special issue of the OEI journal, <i>Revista Iberoamericana de Educación</i>, 82(1): https://rieoei.org/RIE/issue/view/Self_Study For the period 2020–2021 REDFFORMA is supporting three new groups, including more universities and two international teacher educators, one from Colombia and one from Spain. This situation adds interesting new challenges, particularly for the critical friends: How will each group develop shared understanding of what it means to be teacher educator? How will each group fine tune its understanding of teacher education practices if different cultural contexts bring different interpretations of practices?
---	--

we study our own practices now leads us to recommend that the ideal critical friendship in self-study involves two teacher educators studying their own practices simultaneously, preferably over an extended period of time. Together, Rodrigo and Tom have prepared four conference presentations based on their shared self-studies (Fuentelba and Russell, 2014, 2016, 2020; Fuentelba et al., 2018). Tom has participated in several more self-studies involving critical friendship (Schuck and Russell, 2005; Martin and Russell, 2016; Martin, 2017; Martin and Russell, 2018; Russell, 2017).

The advantages of mutual critical friendship may not be obvious at first; adjusting to a developing critical friendship can itself be a challenging experience. Persistence pays rich dividends. Rodrigo and Tom have found that collaborating to construct this encyclopedia entry has been yet another remarkable and positive experience. If seeing teaching as telling (with its implicit assumption that listening is learning) is imperfect, what are the alternatives? Perhaps teaching can be seen in part as *asking good questions* (with its implicit assumption that a question stimulates a more productive cognitive activity). Reviewing our collaborative writing experience, we see that we have freely asked each other good questions. Perhaps a key to critical friendship involves asking good questions—questions that we both want to explore and for which we genuinely want to hear the other's response. When both critical friends are engaging in self-study, the flow of good questions in both directions can be particularly productive. At all times, respect for the knowledge and experience of the other is crucial.

Resources for learning more about self-study of teacher education practices

Major sources of further information about and examples of self-study are the two international handbooks (Loughran et al., 2004; Kitchen et al., 2020) and issues of *Studying Teacher Education*. In addition, there have been 13 international self-study conferences held biennially at Herstmonceux Castle in the UK since 1996; the proceedings of all 13 conferences are available at <https://www.castleconference.com/conference-history.html>. In Spanish we recommend Russell et al. (2016).

Summary and conclusion

With a history of more than 25 years (since 1993), self-study of teacher education practices has achieved considerable momentum. Two editions of an international handbook, a journal with a 15-year history and 13 international conferences with published proceedings provide a great deal of background and guidance for teacher educators who wish to better understand their own teaching practices and their impact on those learning to teach. The methodology set out by LaBoskey (2004) has stood the test of time. While self-study depends on the willingness of one teacher educator to examine not only personal practices but also the underlying assumptions, it is often much more productive and informative when conducted with a colleague, often referred to as a critical friend.

It is common to urge others to learn from experience. Learning from our own experiences is a challenging process, but it offers important opportunities to rethink how we educate future teachers, particularly in a field that has shown stability over time. Lortie's (1975) work suggests that much of what we know about teaching has been learned by observation of those who have taught us. We may have learned a great deal of education theory, but theory is rarely put into practice in some straightforward way. Schön (1995) has argued that we need a different way of knowing, an epistemology of practice, and this reminds us that we have little experience to guide the evaluation of our teaching practices, not just in terms of their influence on those we teach but also in terms of understanding the origins of our practices and how they are evolving. Learning to listen to our students (Cook-Sather, 2002) is one of the most obvious lines of inquiry.

"You can't fix what you can't see"

During our extended critical friendship, Rodrigo has emphasized the importance of identifying one's underlying assumptions, the lines of thought (of which we are typically unaware) that are implicit in our actions in the teacher education classroom. Argyris and Schön (1974) wrote about espoused theory (what we say we do) and theory-in-use (what our actions imply); they also contrasted single-loop and double-loop learning. We close with a reminder that you can't fix what you can't see, and self-study of teacher education practices is a methodology for helping us see what is normally hidden from view in our practice. Self-study has considerable potential for improving both individual teacher education practices and the all-important coherence of teacher education programs. Seeing our own practices in new ways (reframing) is the necessary prelude to generating and exploring new practices (repracticing).

Examples from our own practices and from our critical friendship involving the sharing of our self-studies are intended to provide insight into possible directions arising from self-study. The account of efforts to introduce self-study of teacher education practices into another language and culture demonstrates both the complexities and the possibilities for doing so. Although self-study in each culture will be unique and evolve in different ways, teacher educators exploring the challenges of learning from experience need not be limited to working in English.

Acknowledgment

We thank our colleague Max Lizano for his valuable assistance in confirming the accuracy of translations between English and Spanish.

References

- Argyris, C., Schön, D., 1974. *Theory in Practice. Increasing Professional Effectiveness*. Jossey-Bass, San Francisco.
- Cook-Sather, A., 2002. Authorizing students' perspectives: toward trust, dialogue and change in education. *Educ. Res.* 31 (4), 3–14.
- Darling-Hammond, L. (Ed.), 2006. *Powerful Teacher Education*. Jossey-Bass, San Francisco.
- Feiman-Nemser, S., 2001. From preparation to practice: designing a continuum to strengthen and sustain teaching. *Teach. Coll. Rec.* 103, 1013–1055.
- Fuentealba, R., Russell, T., 2016. Critical friends using self-study methods to challenge practicum assumptions and practices. In: Garbett, D., Ovens, A. (Eds.), *Enacting Self-Study as Methodology for Professional Inquiry (Proceedings of the 11th International Conference on Self-Study of Teacher Education Practices)*. University of Auckland, Auckland, New Zealand, pp. 227–233.
- Fuentealba, R., Russell, T., 2020. Telling is not teaching, listening is not learning: new teacher education practices from a critical friendship. In: Edge, C., Cameron-Standerford, A., Bergh, B. (Eds.), *Textiles and Tapestries: Self-Study for Envisioning New Ways of Knowing*. EdTech Books, Provo, UT. https://edtechbooks.org/textiles_tapestries_self_study/chapter_24.
- Fuentealba, R., Russell, T., Hirmas, C., 2017. Building a professional community of teacher education through a self-study network in Chile. In: Mena, J., García-Varcárcela, J., García-Peñalvo, F., Martín del Pozo, M. (Eds.), *Search and Research: Teacher Education for Contemporary Contexts*. Universidad de Salamanca, Salamanca, pp. 903–910.
- Fuentealba, R., Hirmas, C., Russell, T., 2018. Using a pedagogy of knowledge integration to introduce self-study to groups of teacher educators. In: Garbett, D., Ovens, A. (Eds.), *Pushing Boundaries and Crossing Borders: Self-Study as a Means for Researching Pedagogy (Proceedings of the 12th International Conference on Self-Study of Teacher Education Practices)*. University of Auckland, Auckland, pp. 497–504.
- Fuentealba, R., Hirmas, C., Russell, T., 2020. Introducing self-study of teacher education practices into another culture: the experience in Chile. In: Kitchen, J., Berry, A., Bullock, S.M., et al. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Dordrecht, The Netherlands, pp. 1355–1371.
- Fuentealba Jara, R., Russell, T., 2014. A collaborative self-study of cross-cultural critical friendship to improve teacher education practices. In: Garbett, D., Ovens, A. (Eds.), *Changing Practices for Changing Times: Past, Present and Future Possibilities for Self-Study Research (Proceedings of the 10th International Conference on Self-Study of Teacher Education Practices)*. University of Auckland, Auckland, New Zealand, pp. 91–93.
- Hagger, H., McIntyre, D., 2006. *Learning Teaching from Teachers: Realizing the Potential of School-Based Teacher Education*. Open University Press, Maidenhead.
- Hamilton, M.L., LaBoskey, V.K., Loughran, J., Russell, T., 1998. Have five years of self-study changed teacher education? Artifacts of our personal development as teacher educators. In: Cole, A.L., Finley, S. (Eds.), *Conversations in Community: Proceedings of the Second International Conference of Self-Study of Teacher Education Practices*. Queen's University Faculty of Education, Kingston, ON, pp. 1–5.
- Kitchen, J., Berry, A., Bullock, S.M., et al. (Eds.), 2020. *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer Nature, Singapore.
- Labaree, D.F., 2000. On the nature of teaching and teacher education: difficult practices that look easy. *J. Teach. Educ.* 51 (3), 228–233.
- LaBoskey, V.K., 2004. The methodology of self-study and its methodological underpinnings. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook on Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, The Netherlands, pp. 817–869.
- Lortie, D., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago.
- Loughran, J., Northfield, J., 1998. A framework for the development of self-study practice. In: Hamilton, M.L. (Ed.), *Reconceptualizing Teaching Practice: Self-Study in Teacher Education*. Falmer, London, pp. 7–18.
- Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), 2004. *International Handbook on Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, The Netherlands.
- Martin, A.K., Russell, T., 2016. A self-study of two teacher educators responding to program change. In: Garbett, D., Ovens, A. (Eds.), *Enacting Self-Study as Methodology for Professional Inquiry*. University of Auckland, Auckland, New Zealand, pp. 257–262.
- Martin, A.K., Russell, T., 2018. Does the practicum supervisor know her pedagogy? Insights gained through critical friendship. In: Garbett, D., Ovens, A. (Eds.), *Pushing Boundaries and Crossing Borders: Self-Study as a Means for Researching Pedagogy (Proceedings of the 12th International Conference on Self-Study of Teacher Education Practices)*. University of Auckland, Auckland, New Zealand, pp. 433–440.
- Martin, A.K., 2017. In search of ways to improve practicum learning: self-study of the teacher educator/researcher as responsive listener. *Stud. Teach. Educ.* 13 (2), 127–144. <https://doi.org/10.1080/17425964.2017.1342347>.
- Russell, T., Fuentealba, R., Hirmas, C., 2016. Formadores de formadores, descubriendo la propia voz a través del self-study [Teacher Educators, Discovering Their Own Voice Through Self-Study]. Organización de Estados Iberoamericanos, Santiago, Chile. <https://oei.cl/Oei/publicaciones>.
- Russell, T., Martin, A., Loughran, J., 2020. Introducing new practices in a teacher education classroom: lessons learned, insights gained. In: Edge, C., Cameron-Standerford, A., Bergh, B. (Eds.), *Textiles and Tapestries: Self-Study for Envisioning New Ways of Knowing*. EdTech Books, Provo, UT. https://edtechbooks.org/textiles_tapestries_self_study/chapter_115.
- Russell, T., 2017. Improving the quality of practicum learning: self-study of a faculty member's role in practicum supervision. *Stud. Teach. Educ.* 13 (2), 193–209. <https://doi.org/10.1080/17425964.2017.1342358>.
- Russell, T., 2018. Tom Russell's Outrageous Suggestions for Teaching Teacher Candidates. Unpublished document. Queen's University.
- Schön, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- Schön, D.A., 1995. Knowing-in-action: the new scholarship requires a new epistemology. *Change* 27 (6), 26–34.
- Schuck, S., Russell, T., 2005. Self-study, critical friendship, and the complexities of teacher education. *Stud. Teach. Educ.* 1 (2), 107–121. <https://doi.org/10.1080/17425960500288291>.
- Vick, M., 2006. "It's a difficult matter": historical perspectives on the enduring problem of the practicum in teacher education. *Aust. Pac. J. Teach. Educ.* 34 (2), 181–198.
- Willingham, D.T., 2009. *Why Don't Students Like School?* Jossey-Bass, San Francisco.

Further reading

- Aubusson, P., Schuck, S., 2006. *Teacher Learning and Development: The Mirror Maze*. Springer, Dordrecht, The Netherlands.
- Berry, A., 2007. *Tensions in Teaching About Teaching: Understanding Practice as a Teacher Educator*. Springer, Dordrecht.
- Bullough Jr., R.V., Pinnegar, S., 2001. Guidelines for quality in autobiographical forms of self-study research. *Educ. Res.* 30 (3), 13–21.
- Hamilton, M.L. (Ed.), 1998. *Reconceptualizing Teaching Practice: Self-Study in Teacher Education*. Falmer, London.

- Hoban, G.F. (Ed.), 2005. *The Missing Links in Teacher Education Design: Developing a Multi-linked Conceptual Framework*. Springer, Dordrecht, The Netherlands.
- Kosnik, C., Beck, C., Freese, A.R., Samaras, A.P. (Eds.), 2006. *Making a Difference in Teacher Education Through Self-Study: Studies of Personal, Professional and Program Renewal*. Springer, Dordrecht, The Netherlands.
- Kroll, L., Cossey, R., Donahue, D.M., et al., 2004. *Teaching as Principled Practice: Managing Complexity for Social Justice*. Sage, Thousand Oaks, CA.
- Lampert, M., 2001. *Teaching Problems and the Problems of Teaching*. Yale University Press, New Haven, CT.
- Loughran, J., Russell, T., 2002. *Improving Teacher Education Practices Through Self-Study*. RoutledgeFalmer, London.
- Loughran, J., Russell, T., 2007. Beginning to understand teaching as a discipline. *Stud. Teach. Educ.* 3 (2), 201–211.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education: Understanding Teaching and Learning About Teaching*. Routledge, London.
- Russell, T., Loughran, J. (Eds.), 2007. *Enacting a Pedagogy of Teacher Education: Values, Relationships, and Practices*. Routledge, London.
- Samaras, A.P., 2002. *Self-Study for Teacher Educators: Crafting a Pedagogy for Educational Change*. Peter Lang, New York.
- Schön, D.A., 1992. The theory of inquiry: Dewey's legacy to education. *J. Curric. Stud.* 22, 119–139.
- Tidwell, D., Fitzgerald, L. (Eds.), 2006. *Self-Study and Diversity*. Sense Publishers, Rotterdam.

Professional development or professional learning: developing teacher educators' professional expertise

Gerry Czerniawski, University of East London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	469
Professional development and professional learning – an unnecessary dichotomy	469
The transitional fears of a heterogeneous profession	470
Variability in the work of teacher educators	471
Understanding professional learning in the workplace	471
Knowledge for, in and of practice	472
School-based teacher educators' professional learning	472
Conclusion	473
Acknowledgment	473
References	473

Introduction

"Professional development", "continuing professional development" (CPD) and "professional learning" are concepts often conflated in meaning despite many writers defining these terms differently. While the body of literature relating these concepts to teacher educators is still in its infancy, it is gathering momentum (European Commission's, 2013; White et al., 2015; Murray et al., 2017; Ping et al., 2018; Byman et al., 2020). Understanding the nuanced differences that exist, in policy, practice and academia, when deploying these terms is important when considering how best we can prepare a future generation of teacher educators. When, for example, professional development is narrowly and instrumentally conceived, as a formal, external activity done to recipients then opportunities for authentic teacher educators' professional learning is limited. If, however, a wider interpretation of professional development is used embracing both formal and informal activities that enable critical professional reflection, then professional development, as a term, has validity. This point is significant for two reasons. First, because of the variegated composition of the contemporary teacher education workforce (Menter, 2018) in schools, colleges and university and their variegated professional learning needs. Second, most teacher educators do not experience supportive induction programmes or any form of formal route to their profession (Mayer et al., 2011; Czerniawski, 2013; Ping et al., 2018).

This article begins by examining what we mean by professional development and professional learning, and how both terms vary in their significance for teacher educators. The article then looks at the transitional fears experienced by neophyte teacher educators and the significance this has for their professional learning needs. Finally the article explores some of the literature on professional learning in the workplace and how that literature can contribute to increasing our understanding about teacher educators' professional learning needs.

Professional development and professional learning – an unnecessary dichotomy

Professional development is often deployed as a portmanteau term for what Day and Sachs (2009) describe as a "hugely complex intellectual and emotional endeavor" (p. 4). Typically, activities cited include observations, on-the-job coaching and mentoring, team teaching, self-directed study, in-service courses, sabbaticals, membership of working groups and networks, collaborative learning (both formally and informally), professional reflection and research (Dymoke and Harrison, 2006; Bates et al., 2011; Ping et al., 2018). However, writing about teachers, Stoll and Earl (2011) draw a distinction between professional development and professional learning:

Both are intentional, ongoing, systematic processes. Over time, however, the term "professional development" has taken on connotations of delivery of some kind of information ... in order to influence their practice whereas "professional learning" implies an internal process in which individuals create professional knowledge through interaction with this information in a way that challenges previous assumptions and creates new meanings (p. 4).

Watson and Michael (2016), writing about Scottish teacher professionalism, argue that while in the past Continuing Professional Development (CPD) became a commodified quantifiable "thing" that constructs professionalism as adherence to a contract, more recently in Scottish policy documentation (e.g. Donaldson, 2010; NPG, 2012) there has been a linguistic shift from "CPD" to "professional learning":

In contrast to CPD in which the teacher is the professional engaged in development, the discursive shift to professional learning sees the teacher embark on an endless journey of *becoming*, taking on the responsibility for their own self development

(Watson and Michael, 2016: p. 272).

Despite the growing body of literature on teacher educators' professional learning (e.g., Czerniawski, 2018; McPhail et al., 2018; White et al., 2020), this body is still relatively under-researched, with much of the literature drawing on teachers' continuing professional development (CPD) in schools. Kennedy (2005), for example, has identified nine models of teachers' CPD in international literature, classified in relation to their capacity for supporting professional autonomy and transformative practice. These teacher-based models have been usefully applied to literature on teacher educators' professional learning and development (see Bates et al., 2011):

- **Training model:** a top-down, skills-based technocratic view of professional development delivered by "experts" to passive learners e.g., early models of mentor training in which school mentors attended sessions run by university-based teacher educators and often held at the university.
- **Award bearing model:** award-bearing programmes normally validated by Higher Education Institutions (HEIs) e.g., Master's level programmes;
- **Deficit model:** positions professional development needs arising from becoming a teacher educator and acted upon either by the individual or part of a managed process on the part of the employer institution;
- **Cascade model:** in which an individual attends a professional learning event and then re-presents the content and resources of the event to other colleagues;
- **Standards-based model:** The use of standards (sometimes described as "professional frameworks") for teacher educators (e.g., as used in the Netherlands and the USA). Debates exist as to what extent these standards might be imposed on or owned by teacher educators and the extent to which they are seen as rules or guidelines accordingly.
- **Coaching model:** emphasizing the ways in which novice teacher educators work alongside more established members of the professional community or "communities of practice" (Lave and Wenger, 1991);
- **Community of practice model:** acknowledging many aspects of the coaching model above this model emphasizes the importance of "legitimate peripheral participation" (Lave and Wenger, 1991). The importance of supportive relationships in engaging with the practices of the community is emphasized. By engaging in peripheral activities new entrants become socialized into the tasks, ways-of-being, terminology and organizing principals of the community;
- **Action research model:** with its focus on reflective practice this model places high value on research engagement as a process in professional learning. While not demeaning the value of externally produced research the emphasis is on teacher educators identifying and implementing their own research activity;
- **Transformative model:** aligning professional learning to educational change this model effectively enables an opportunity to combine some or all of the above models positioning CPD as an opportunity to support and promote whole scale change in the education system.

Hadar and Brody (2017) divide professional learning programmes and models like these into two paths. The first, self-guided learning, consists of teacher educators engaging in research (often self study) individually or in groups (see Vanassche and Kelchtermans', 2015 review of teacher educators' self-study literature). The second path consists of more formalized programmes, courses and events specifically designed for teacher educators. The authors sub-divide this second classification into programmes initiated by associations for teacher educators that serve the needs of teacher educators from different institutions; and specific institutions providing in-service programmes of learning for its own employees (see below). This bifurcation is helpful when considering a further distinction made by Lipowski et al. (2011) between two forms of professional learning. First: *in-service programmes* are organized formal programmes for practitioners within the institutions where they work, considered by some to be the primary way in which they receive continuing support (Loucks-Horsley et al., 1997). Second: *continuous experiential learning* accommodates the more informal learning opportunities that contribute to everyday professional practice, the importance of which cannot and should not be underestimated when trying to understand the work teacher educators do in different national locations and the professional support they need (MacPhail et al., 2014).

The transitional fears of a heterogeneous profession

Making generalizations about teacher educators' professional learning needs is problematic. This is despite existing trends, internationally, in the implementation of professional standards' frameworks, and, in some cases the increasing take-up of school-based and school-driven teacher education (White et al., 2015). Professional learning needs vary depending on the context in which the teacher educator is situated. Universities, schools and colleges differ in their teacher education programmes; and university, college and school departments may vary in their interpretations of the knowledge, skills, practices, ethics, values and attributes that different frameworks prioritize. Murray (2014) also draws attention to how occupational change for neophyte teacher educators, usually those coming from a background in school teaching to higher education (HE) can bring with it professional uncertainty. Relatively poorly conceived university induction programmes for new staff often compound this transition, particularly for those

coming from the more practitioner-based occupations such as teaching, nursing and social work. Four transitional fears have been identified in the literature by Braund (2015) as teacher educators move into university employment.

1. Uncertainty with new professional roles (Boyd and Harris, 2010);
2. The nature of adult pedagogy (Loughran, 2006; Murray and Male, 2005);
3. Fears in relation to the adequacy of professional and academic knowledge bases (McKeon and Harrison, 2010)
4. The relative absence of collaborative working environments to facilitate reflective workplace learning (Eraut, 2004; Harrison and McKeon, 2008).

Such occupational changes can be particularly challenging for those coming from often senior positions within schools. Once in university employment, many teacher educators find themselves in relatively lower hierarchical positions in a university department and working within a discipline (education) often considered to be low status, "soft", applied and marginalized (Becher and Trowler, 2002; Murray, 2014).

Variability in the work of teacher educators

The variability in teacher educators' work roles is well documented (Lunenburg et al., 2014; Swennen et al., 2010) with each role associated with potentially a different set of professional learning needs. These work roles can include: teaching, coaching, the facilitation of collaboration between diverse organizations and stakeholders, assessment, gatekeeping (into the teaching profession), curriculum development, research, critical inquiry and writing. The nature of these roles and the professional learning associated with them will vary depending on a number of factors including the nature of employment (e.g., part time/full time; length of contract; length of time in service), the phase (e.g., primary/secondary/post-16) and whether teacher educators are situated in schools, colleges or universities. However it is the scholar and researcher roles that have been cited as the most pressing concern for novice teacher educators working in universities (Boyd et al., 2011; Braund, 2015).

Griffiths et al. (2014) make a further distinction found between two groups of university-based teacher educators. Those coming from Anglophone countries often move to universities having previously taught in schools. A second group exists from countries where teacher educators are drawn mainly from academic disciplines often lacking practical teaching experience (Griffiths et al., 2014). Both groups are likely to encounter different transitional experiences and tensions in the development of their own professional identities and their subsequent learning needs.

Understanding professional learning in the workplace

In the introduction to their book on workplace learning in teacher education, McNamara et al. (2014), draw attention to the historical trajectory England has taken in terms of workplace learning. This is a trajectory that has shifted from the 19th century classroom-based apprenticeship model of teacher education to the more theory-laden university-based model broadly adopted by the second-half of the 20th century. However since the 1990s the authors point to the turning of the tide in England (in contrast to much of the rest of Europe) with a return to a more extensive workplace-based model of teacher education, fueled by "radical ideological fervor" (ibid:1). But the authors also point out that this relocation of professional learning to the classroom is one that is qualitatively different and more complex than its 19th century ancestor, and that as such, it makes England an interesting "test-bed" for understanding the complexities in workplace learning.

The body of literature associated with workplace and practice-based professional learning frameworks continues to grow, but its capacity for theorizing such learning is still relatively under-deployed within the literature on teacher educators. As we saw earlier, Lave and Wenger's (1991) concepts of communities of practice and situated learning have dominated much of this literature. Both authors argue that this type of learning involves participation in a community of practice with learning moving from legitimate peripheral participation to central participation within that community. For example, I might start my new job at a university knowing what to teach, but it might take me several weeks to know: who to talk to, who to avoid, where to sit, what informal and formal "rules", norms and values I should adhere to if I am to be accepted as a colleague within that professional community. Communities of practice are made up of individuals who mutually engage in an activity and develop communal resources over a period of time. Learning in this sense is about participating in the practices associated with that community rather than an acquisition of knowledge per se. Learning is "situated" within the practices and social cultural context of that community (e.g., a university department, school, publishing company etc). A community is viewed as "the social configuration in which our enterprises are defined as worth pursuing and our participation is recognizable as competence" (Wenger, 1998: 5). Legitimate peripheral participation describes how novice employees start by engaging with subordinate and auxiliary practices within the periphery of that community (e.g., understanding the protocols of school visits; who to go to for important information within the department etc). As novices gain more experience they move toward the central participation within that community (e.g., being able to design and facilitate a curriculum for student teachers). Understanding the significance of different communities of practice within any organization is significant not just in providing greater insight into how individuals develop professionally but how communities of practice can contribute to the development of larger organizations (e.g., universities) of which they are a part (Philpot, 2014).

Knowledge for, in and of practice

In discussing teachers' professional practice-based learning, Cochran-Smith and Lytle (1999) distinguish between knowledge-for-practice, knowledge-in-practice and knowledge-of-practice. The authors conceive knowledge *for* practice as the formal substantive content knowledge required for teachers to improve their own practice and the learning of their students. The assumption here being that this "formal" knowledge generated by university-based researchers leads to more effective teaching and learning. Knowledge *in* practice, focuses on the day-to-day practical knowledge of the teacher "embedded in practice and in teachers' reflections on practice" (Cochran-Smith and Lytle, 1999: 250). This knowledge can be enhanced by reflecting on the work of expert colleagues and the challenging situations encountered in everyday at work. By focusing on knowledge in practice educators are encouraged to be aware and challenge their own assumptions, reasoning and decision making. Knowledge *of* practice assumes that knowledge to teach well is generated when teachers treat their own classrooms and schools "as sites for intentional investigation" while treating the knowledge produced by others as "generative material for interrogation and interpretation" (Cochran-Smith and Lytle, 1999:250). Practice-based education (PBE) locates learning in all three of these domains, acknowledging their interrelationship and significance in the process of professional reflection. In this reconceptualization of the reflective practitioner (Schon, 1983), PBE embraces a "form of social practice which shapes the educational development of individuals, framed around a perspective, model or theory of education that encompasses interactive philosophical, political, moral, technical and practical dimensions" (Higgs, 2011: 9). For the teacher educator, this emphasis on the complexity, value and significance of this type of reflection adds further weight to the distinction made earlier between teacher "training" and teacher "education". It also adds weight to the argument that there is "an urgent need to re-affirm teachers' and teacher educators' professionalism in terms of a research-based profession" (Jones et al., 2011: 272).

School-based teacher educators' professional learning

In recent years increasing attention is being paid to the shift, internationally, toward more school-based teacher education models (Boyd and Tibke, 2012; Lunenberg et al., 2014; White et al., 2015). In England, these models have included school-centered initial teacher training schemes (SCITTs), employment-based routes (EBITTs), School Direct, Teach First and the emergence of Teaching Schools. Such models mean that many teacher educators working in schools find themselves increasingly responsible for organizing all aspects of Initial Teacher Training (ITT) including the recruitment of trainees; the design, implementation and evaluation of course components and their assessment (McNamara et al., 2017; Czerniawski et al., 2017a,b). White et al. (2015: 443) divide school-based teacher educator roles into three areas, namely, those associated with a traditional mentoring role, those associated with a supervisory role across a school or schools and those associated more commonly with institute-based teacher educators (IBTEs). Teacher educators working in schools are also increasingly responsible for the continuing professional development of qualified school teaching staff within their own institutions.

Evidence of professional learning opportunities for school-based teacher educators (specifically related to their role as teacher educators) is, as yet, limited. A point emphasized by O'Dwyer and Atli (2015: 6) is that:

The lack of focus of in-service educators may represent an important omission given the current importance attached to teachers as life-long learners as a result of endemic professional change

The presumption that, if you can teach young people in schools you can therefore teach adults how to teach and how to teach more effectively, is one that has permeated teacher education systems in England since the introduction of the state system toward the end of the nineteenth century. This is perhaps one explanatory factor for why very few formal professional learning opportunities have existed for school-based teacher educators in the past. And yet the knowledge, skills, creativity and emotional endeavor required to teach adults varies considerably in content, form, process and technique to that required when teaching children. The dual identity that school-based teacher educators possess (as both teachers and teacher educators) adds additional complexity in understanding what sorts of professional learning activities are of most value to the profession as a whole.

However while the literature on school-based teacher educator professional learning might be limited it is nevertheless growing. A study of university and school-based teacher educators in the Netherlands (Dengerlink et al., 2015) has indicated that the preferences of school-based teacher educators vary over time. Teacher educators with less experience were interested in coaching skills, pedagogical content knowledge, their own role within wider communities of teacher educators and opportunities to contribute to the knowledge development of the profession as a whole. They valued participation in seminars, conferences and formal courses and desired greater opportunities for more structured learning arrangements (e.g., supervision, peer-coaching etc). However school-based teacher educators with more than 7 years experience were interested in learning about the policy context of school-university partnerships, curriculum issues at the programme level and the pedagogy of teacher education. These more experienced educators tended to place greater value on professional reading, participating in research projects and the role that experimentation can play within their own professional practice. Common to both groups was the value placed on conversations with colleagues as a significant form of professional learning.

In England, one of the outcomes of governmental policy in initial teacher education since 2010, has been the expanding provision of *professional development* for some teacher educators in some schools albeit in fragmented and particularized forms. A point emphasized by Murray et al. (2017):

The emergence of "Teaching Schools" as recognized centres of excellence, with government funding available to support school-led ITE programmes and school-focused research and development projects, has greatly increased the professional learning opportunities available to school-based educators (p. 660).

However, in England, opportunities brought about by an expanding school-based system of teacher education might not necessarily result in more effective *professional learning* of its school-based teacher educators. Earley and Bubbs (2004) distinguish "hard" economic utilitarianism, where professional *development* addresses the strategic goals of an institution, from a "softer" developmental humanism in which professional *learning* caters for valued, confident and motivated staff. There is a danger therefore that in England's marketized and competitive school system, teacher educators working in schools might be subjected to the former rather than the latter.

Conclusion

When professional development is narrowly and instrumentally conceived as a formal, external activity "done to" recipients then opportunities for authentic teacher educators' professional learning is limited. If, however, a wider interpretation of professional development is used embracing both formal and informal activities that enable critical professional reflection then professional development, as a term, has validity. However I prefer the term "professional learning" to "professional development" and define teacher educators' professional learning as both the formal and informal processes that enable them to improve their professional practice throughout their careers with a commitment to transform education for the better. By understanding, more fully, the sorts of transitional dilemmas encountered by those entering teacher education for the first time, institutional leadership teams in schools, colleges and universities are in a position to construct more meaningful professional induction and more sustained, valued, authentic and nuanced professional learning opportunities. However the article affirms existing literature that professional learning for teacher educators is often deemed for more valuable, not through formal provision and organized programmes, but through informal workplace learning (Boyd et al., 2011; Murray et al., 2017). While professional learning opportunities for teacher educators working in schools and universities are on the rise much more needs to be done to foster and recognize the professional learning needs of both groups of educators. One notable difference between these groups centers around the degree of support sought for research activities with this having greater priority for higher education-based teacher educators.

Acknowledgment

Many of the ideas explored in this article draw directly on material first published in *Teacher Educators in the Twenty-First Century*, Czerniawski (2018). St Albans: Critical Publishing, where they are discussed in greater detail.

References

- Bates, T., Swennen, A., Jones, K., 2011. *The Professional Development of Teacher Educators*. Routledge, London.
- Becher, T., Trowler, P., 2002. *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. SRHE and OUP, Buckingham.
- Boyd, P., Harris, K., 2010. Becoming a university lecturer in teacher education: expert school teachers reconstructing their pedagogy and identity. *Prof. Dev. Educ.* 36 (1–2), 9–24.
- Boyd, P., Tibke, J., 2012. Being a school-based teacher educator: developing pedagogy and identity in facilitating work-based higher education in a professional field. *Pract. Res. High. Educ.* 6 (2), 41–57.
- Boyd, P., Harris, K., Murray, J., 2011. *Becoming a Teacher Educator: Guidelines for Induction*. Higher Education Academy, Subject Centre for Education (ESCalate), University of Bristol Graduate School of Education, Bristol.
- Braund, M., 2015. Teacher educators' professional journeys: pedagogical and systemic issues affecting role perceptions. *Afr. Educ. Rev.* 12 (2), 309–330.
- Byman, R., Jyrhämä, R., Stenberg, K., Sintonen, S., Maaranen, K., Kynäslähti, H., 2020. Finnish teacher educators' preferences for their professional development – quantitative exploration. *Eur. J. Teach. Educ.* <https://doi.org/10.1080/02619768.2020.1793952>.
- Cochran-Smith, M., Lytle, S.L., 1999. Relationships of knowledge and practice: teacher learning communities. *Rev. Res. Educ.* 24, 249–305.
- Czerniawski, G., Kidd, W., Murray, J., 2017a. We are all teacher educators now – understanding school-based teacher educators in times of change in England. In: Murray, J., Swennen, A., Kosnik, C. (Eds.), *International Research, Policy and Practice in Teacher Education: Insider Perspectives*. Springer Academic Publishers, Dordrecht.
- Czerniawski, G., MacPhail, A., Guberman, A., 2017b. The professional development needs of higher education-based teacher educators: an international comparative needs analysis. *Eur. J. Teach. Educ.* vol. 40 (1), 127–140.
- Czerniawski, G., 2013. Professional development for professional learners: teachers' experiences in Norway, Germany and England. *J. Educ. Teach.* 39 (4), 383–399.
- Czerniawski, G., 2018. *Teacher Educators in the Twenty-First Century – Identity, Knowledge and Research*. Critical Publishing, St Albans.
- Day, C., Sachs, J., 2009. *International Handbook on the Continuing Professional Development of Teachers*. Open University Press, Maidenhead.
- Dengerink, J., Lunenburg, M., Kools, Q., 2015. What and how teacher educators prefer to learn. *J. Educ. Teach.* 41 (1), 78–96. <https://doi.org/10.1080/02607476.2014.992635>.
- Donaldson, G., 2010. *Teaching Scotland's Future. Report of a Review of Teacher Education in Scotland*. Scottish Government, Edinburgh.

- Dymoke, S., Harrison, J.K., 2006. Professional development and the beginning teacher: issues of teacher autonomy and institutional conformity. *J. Educ. Teach.* 32 (1), 71–92.
- Earley, P., Bubbs, S., 2004. *Leading and Managing Continuing Professional Development: Developing Teachers, Developing Schools*. Sage/Paul Chapman, London.
- Eraut, M., 2004. Informal learning in the workplace. *Stud. Cont. Educ.* 26 (2), 247–273.
- European Commission, 2013. *Supporting Teacher Educators for Better Learning Outcomes*. European Commission, Brussels.
- Griffiths, V., Thompson, S., Hryniewicz, L., 2014. Landmarks in the professional and academic development of mid-career teacher educators. *Eur. J. Teach. Educ.* 37 (1), 74–90.
- Hadar, L.L., Brody, L.D., 2017. *Teacher Educators' Professional Learning*. Routledge, New York.
- Harrison, J., McKeon, F., 2008. The formal and situated learning of beginning teacher educators in England: identifying characteristics for successful induction in the transition from workplace in schools to workplace in higher education. *Eur. J. Teach. Educ.* 31 (2), 151–168.
- Higgs, J., 2011. *Practice-Based Education: Enhancing Practice and Pedagogy*. Final Report for ALTC Teaching Fellowship. Australian Learning and Teaching Council – Australia.
- Jones, M., Stanley, G., McNamara, O., Murray, J., 2011. Facilitating teacher educators' professional learning through a regional research capacity-building network. *Asia Pac. J. Teach. Educ.* 39 (3), 263–275.
- Kennedy, A., 2005. Models of continuing professional development: a framework for analysis. *J. Serv. Educ.* 31 (2), 235–250.
- Lave, J., Wenger, E., 1991. *Situated Learning*. Cambridge University Press, Cambridge.
- Lipowski, K., Jorde, D., Prenzel, M., Seidel, T., 2011. Expert views on the implementation of teacher professional development in European countries. *Prof. Dev. Educ.* 37 (5), 685–700.
- Loucks-Horsley, S., Sewson, P.W., Love, N.B., Stiles, K.E., 1997. *Designing Professional Development for Teachers of Science and Mathematics*. Corwin, Thousand Oaks, CA.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education*. Taylor and Francis, London.
- Lunenberg, M., Dengerink, J., Korthagen, F., 2014. *The Professional Teacher Educators: Roles, Behaviour and Professional Development of Teachers*. Sense Publications, Rotterdam.
- MacPhail, A., Patton, K., Parker, M., Tannehill, D., 2014. Leading by example: teacher educators' professional learning through communities of practice. *Quest* 66, 39–56.
- MacPhail, A., Ulvik, M., Guberman, A., Oolbekkink-Marchand, H.W., Czerniawski, G., Bain, Y., 2018. The professional development of higher education-based teacher educators: needs and realities. *Prof. Dev. Educ.* <https://doi.org/10.1080/19415257.2018.1529610>.
- Mayer, D., Mitchell, J., Santoro, N., White, S., 2011. Teacher educators and "accidental" careers in academe: an Australian perspective. *J. Educ. Teach.* 37 (3), 247–260.
- McKeon, F., Harrison, J., 2010. Developing Pedagogical practice and professional identities of beginning teacher educators. *Prof. Dev. Educ.* 36 (12), 25–44.
- McNamara, O., Murray, J., Jones, M., 2014. *Workplace Learning in Teacher Education*. Springer, Singapore.
- McNamara, O., Murray, J., Phillips, R., 2017. Policy and Research Evidence in the 'Reform' of Primary Initial Teacher Education in England. Cambridge Primary Review Trust, York.
- Menter, I., 2018. Foreword. In: Czerniawski, G. (Ed.), *Teacher Educators in the 21st Century – Identity, Knowledge and Research*. Critical Publishing, St Albans.
- Murray, J., Male, T., 2005. Becoming a teacher educator: evidence from the field. *Teach. Teach. Educ.* 21 (2), 125–142.
- Murray, J., Lunenberg, M., Smith, K., 2017. Educating the educators: policies and initiatives in European teacher education. In: Peters, M., Cowie, B., Menter, I. (Eds.), *A Companion to Research in Teacher Education*. Springer Publishers, pp. 651–666.
- Murray, J., 2014. Teacher educators' constructions of professionalism: a case study. *Asia Pac. J. Teach. Educ.* 42 (1), 7–21.
- NPG (National Partnership Group), 2012. *Teaching Scotland's Future - National Partnership Group. Report to Cabinet Secretary for Education and Life-Long Learning*. Scottish Government, Edinburgh. <http://www.gov.scot/Resource/0040/00407756.pdf>.
- O'Dwyer, J.B., Atli, H.H., 2015. *Eur. J. Teach. Educ.* 38 (1), 4–20.
- Philpott, C., 2014. *Theories of Professional Learning – a Critical Guide for Teacher Educators*. Critical Publishing, Northwich.
- Ping, C., Schellings, G., Beijard, J., 2018. Teacher educators professional learning: a literature review. *Teach. Teach. Educ.* 75 (2018), 93–104.
- Schon, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- Stoll, L., Earl, L., 2011. *Realising the Power of Professional Learning*. Open University Press, Maidenhead.
- Swennen, A., Jones, K., Volman, M., 2010. Teacher educators: their identities, sub-identities and implications for professional development. *Prof. Dev. Educ.* 36 (1–2), 131–148.
- Vanassche, E., Kelchtermans, G., 2015. The state of the art in self-study of teacher education practices: a systematic review. *J. Curric. Stud.* <https://doi.org/10.1080/00220272.2014.995712>.
- Watson, C., Michael, M.K., 2016. Translations of policy and shifting demands of teacher professionalism: from CPD to professional learning. *J. Educ. Pol.* 31 (3), 259–274. <https://doi.org/10.1080/02680939.2015.1092053>.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. University Press, Cambridge.
- White, E., Dickerson, C., Weston, K., 2015. Developing an appreciation of what it means to be a school-based teacher educator. *Eur. J. Teach. Educ.* 38 (4), 445–459.
- White, S., Murray, J., Goodwin, A.L., Kosnik, C., Beck, C., 2020. On the shoulders of giants: advice for beginning teacher educators. *Asia Pac. J. Teach. Educ.* <https://doi.org/10.1080/1359866X.2020.1756223>. Available at:

Equity- and justice-minded teacher educators

Jamy Stillman, University of Colorado Boulder, School of Education, Boulder, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	475
Racism and the “overwhelming presence of whiteness” in teacher education	475
The overwhelming presence of whiteness in teacher education	476
How teacher education preserves the status quo of whiteness	476
Conceptualizing teacher educator development as a learning problem	477
The value of CHAT in theorizing teacher educators’ learning and activity	477
Critical pedagogy as a complementary frame for considering teacher educators’ learning	478
A doctoral curriculum for preparing equity- and justice-oriented teacher educators	478
What does equity- and justice-oriented teaching educating involve?	479
Three dimensions of doctoral education for equity- and justice-oriented NTEs	480
Teacher educating for equity and justice	480
Equity- and justice- oriented doctoral coursework	481
Critical inquiry communities	483
Discussion	484
Conclusion	485
Acknowledgments	485
References	485

Introduction

Since the early 2000s, teacher education scholars have argued the need to focus more concerted on teacher educator development. Despite more than 20 years of such calls, however, teacher educator preparation and support remain under-practiced, under-supported and undertheorized. Formal preparation pathways for those interested in becoming teacher educators are few and far between (Cochran-Smith et al., 2020), and teacher educators report mostly “falling into” the profession and figuring out what teacher educating entails while “already on the job” (Loughran, 2006; Goodwin et al., 2014).

Supporting teacher educators’ development matters for multiple reasons. Numerous accounts, for example, illustrate that making the transition from teacher to teacher educator is not automatic or natural, (Feiman-Nemser and Buchmann, 1985) since teacher educating requires pedagogical approaches that are distinct from classroom teaching (Berry, 2009; Korthagen, 2020; Murray and Male, 2005). Others underscore how teacher education’s complexity and contingency in relation to local contexts—including the communities for whom teachers are prepared, teacher educators’ and teachers/learners’ identities, evolving policy and political landscapes and institutional conditions—demand specially-prepared, highly skilled practitioners (Stillman and Beltramo, 2019). Still others point to claims about the ineffectiveness of teacher education as a whole as a reason for attending more seriously to teacher educators’ development. Paradoxically, claims of teacher education’s ineffectiveness often decry its wrongheadedness (Cochran-Smith et al., 2008; Levine, 2006), as well as its weak socializing impact on graduates (Zeichner and Tabachnick, 1981). Stronger teacher educator preparation could, therefore, be important for stronger teacher education overall.

Less common among these critiques are those that outline how limited attention to teacher educator preparation contributes to teacher education’s struggle to prepare equity-and justice- oriented teachers (e.g., Cochran-Smith, 2003; Conklin, 2020; Stillman and Beltramo, 2019; Stillman et al., 2019a; Beltramo et al., 2020). Yet, ample evidence suggests that teacher education often perpetuates racism and white supremacy, contributing to rather than disrupting the harm that K-12 schools have historically imposed on students from non-dominant communities (Milner, 2008; Sleeter, 2017).

This paper argues that teacher educators stand to play a crucial role in transforming teacher education from a status quo-perpetuating endeavor into an equity- and justice-oriented one and suggests that specialized teacher educator preparation is essential to realizing this possibility (Souto-Manning and Stillman, 2020; Stillman et al., 2019a,b). Following an account of how teacher education upholds racism and centers whiteness through its everyday, largely taken-for-granted practices, the chapter proposes a doctoral curriculum framework for preparing teacher educators for the transformational work the field requires.

Racism and the “overwhelming presence of whiteness” in teacher education

Although it has become fashionable for university teacher education programs to claim the mantle of social justice, teacher education often fails to disrupt and arguably contributes to injustice and growing opportunity gaps for students of color and other minoritized students (e.g., Milner, 2008; Picower, 2021; Souto-Manning, 2019). This chasm between teacher education’s aims and actions

reflects the “gross undertheorization” of race and racism in teacher education (Milner et al., 2013), as well as the field’s persistent and “overwhelming presence of whiteness” (Sleeter, 2001, 2017).

The overwhelming presence of whiteness in teacher education

At the most basic level, whiteness in teacher education is evidenced by its participants. In the US, around 74% of teacher candidates in university programs are white (U.S. Department of Education, 2016); this number stands in stark contrast to the US K-12 public school student population, of which more than 50% are students of color (National Center for Education Statistics, 2019). White teacher candidates also complete university preparation programs at a higher rate than teacher candidates of color, with more than 80% of graduates from US universities identifying as white (U.S. Department of Education, 2016). Although more recent statistics have been difficult to locate, as of around 15 years ago, the numbers looked similar for university-based teacher educators, who were then around 88% white (Ladson-Billings, 2005).

Complicating the notion of whiteness, Ohito (2020) explains that “whiteness can be considered ‘a constellation of knowledge, ideologies, norms, values, identities, and behaviors that maintains a race and class hierarchy’ (Helfand, 2009, p. 77). This definition gestures toward the entwinement of Whiteness and racism.” This “constellation” is performed in teacher education’s everyday acts through embodied enactments of Whiteness (Ohito, 2020). For example, Whiteness permeates teacher education’s curriculum and learning experiences since a mostly white teacher education faculty means programs largely reflect “white sensibilities” (Sleeter, 2017). Programs commonly treat teaching and teacher education as culture-free (Nieto and Bode, 2011) and/or subscribe to an “add-on” approach in which learning about equity and diversity are relegated to one or two courses rather than addressed program-matically (Cochran-Smith, 2000; King and Butler, 2015; Rueda and Stillman, 2012).

Even when teacher education programs address equity and diversity, whiteness still tends to take center stage. Teacher education instruction is often geared toward white teacher candidates, assumed to have little knowledge of or experience with non-dominant communities (Lowenstein, 2009). In addition, during discussions about equity and justice, white teacher educators, especially, have been shown to: enable white teacher candidates to control the narrative; privilege white candidates’ emotional needs; passively accept deficit framings of students and communities of color; and selectively address or evade racism—based on their own comfort (Chang-Bacon, 2022; Matias, 2016; Ohito, 2020; Sleeter, 2017; Teachers of Color Collective and Souto-Manning, 2022). These actions can hold true even for white teacher educators who consider themselves anti-racists (Ohito, 2020).

Meanwhile, just as teacher education is often treated as culture-free, the K-12 content that teacher education programs expect teacher candidates to learn typically mirrors the Eurocentric, white dominated K-12 curriculum (Picower, 2021). Literal and figurative distance and disconnection between university teacher education and the knowledges and experiences of communities of color likewise perpetuates whiteness’s grip (Souto-Manning and Martell, 2019). Given all this, it is no surprise that white teachers who matriculate from preservice programs typically “are not equipped to offer the racially/ethnically diverse students in schools a strong and culturally responsive education” (Sleeter, 2017, p. 157).

It is to state the obvious to acknowledge that the conditions and practices of “normalized White culture” (Brown, 2014, p. 340) in teacher education don’t just fail to adequately prepare white teacher candidates to teach students from non-dominant communities; they also negatively impact teacher candidates of color (Brown, 2014; Teachers of Color Collective and Souto-Manning, 2022). In addition to harms resulting from the patterns described above, teachers and teacher candidates of color consistently report experiencing marginalization, silencing, and a lack of belonging in teacher education (Brown, 2014; Haddix, 2012; Souto-Manning and Emdin, 2020; Teachers of Color Collective and Souto-Manning, 2022). Not only does teacher education regularly devalue the histories, knowledges, and experiences of candidates of color, it also often essentializes their identities (Matias, 2016; Teachers of Color Collective and Souto-Manning, 2022). This includes assumptions that teacher candidates of color enter teacher education already committed to teaching for social justice and ready to teach and serve as role models for minoritized children (Brown, 2014). Brown (2014) explains this “entraps these teachers to one specific role” while masking “the actual ways that they might make a pedagogic difference for both K-12 students of color, as well for their White counterparts” (Brown, 2014, p. 338). Assuming teacher candidates of color enter programs “ready” also absolves teacher educators from their core responsibility to help all candidates develop crucial dimensions of knowledge and expertise needed for effective teaching. Together, these patterns suggest that calls and efforts to “diversify” teaching and teacher education may ultimately serve white interests by making teacher education appear more inclusive, without transforming endemic structures and conditions that perpetuate racism and white supremacy (Brown, 2014; Sleeter, 2017).

How teacher education preserves the status quo of whiteness

Multiple factors sustain racism and the “overwhelming presence of whiteness” in teacher education. In addition to the field’s under-theorization of race, racism, and whiteness (Milner et al., 2013), which normalizes and obscures racism’s presence and role, teacher education’s conservatism—its endurance over time with little change—serves to maintain status quo-perpetuating pedagogies and practices (Ellis and McNicholl, 2016; Goodwin et al., 2014). The larger institutional culture and neoliberal policy contexts surrounding university teacher education likewise uphold the status quo (Picower and Kohli, 2017; Zeichner, 2010). Commonly considered universities’ “cash cows,” programs often foreclose access for students who must work while preparing to teach by scheduling courses during traditional hours and requiring extensive, unpaid field experiences. Biased subject area exams and teacher performance assessments shown to discriminate against teachers of color ensure the predominance of white teacher candidates

and advance the false narrative that they are more qualified to teach than candidates of color (Ladson Billings, 1998; Sleeter, 2017). Meanwhile, high-stakes accountability policies (e.g., standardized testing in K-12 schools; state and national accreditation regimes; etc.) ensure the teacher education curriculum remains white- and Euro-centric, and English dominant (Perez Huber et al., 2006; Souto-Manning, 2021).

At the same time, faculty reward structures—which treat teacher education as low-status and privilege research productivity over teaching and other applied work—de-incentivize teacher educators from engaging in necessary (un)learning about racism and other forms of oppression (Cochran-Smith, 2003) or time-intensive justice-oriented work like building authentic relationships with and learning from school partners and community members. Relatedly, teacher educators who express anti-racist stances are more likely to receive poor course evaluations, with particularly harmful outcomes for teacher educators of color (Jayakumar et al., 2009; Matias, 2016; Ohito and Deckman, 2018). Notably, these conditions intersect with persistent, and well-documented dilemmas of teacher education, like the apprenticeship of observation (Lortie, 1975), whereby teacher candidates, having attended status-quo perpetuating K-12 schools themselves, resist approaches to schooling with which they are unfamiliar.

As all of this suggests, it will take nothing short of transformation for teacher education to become genuinely anti-oppressive and capable of preparing equity- and justice-oriented teachers. This transformation will require pushing beyond incremental changes that involve tinkering with the field's lowest-hanging fruits (e.g., adding a few new readings; revising a single course, etc.), instead addressing the field's diseased root structures that propagate some of teacher education's most troubling features.

Some scholars have argued that highly qualified, equity- and justice-oriented teacher educators may be the linchpins for inciting such foundational change (Carter Andrews, 2021; Cochran-Smith, 2003; Goodwin and Kosnick, 2013; Souto-Manning and Stillman, 2020; Stillman and Beltramo, 2019), especially given the transformational potential of their pedagogical work (Conklin, 2015, 2020; Ellis and McNicholl, 2016). Ellis's (2018) definition of transformation as "change processes that take place *within* social systems or formal organizations and that require ... the work-on-change to be driven by the momentum of participants' agency rather than ... by external political energy," likewise suggests teacher educators are among those best positioned to stimulate consequential changes in the field.

Yet, current pathways into the profession, including learning while "on the job," offer limited opportunities for emerging teacher educators to develop the "shift in underlying values and rationalities as well as concrete labor practices" (Ellis, 2018) needed to refashion teacher education's architecture. This paper, therefore, offers ideas for a doctoral curriculum that aims to equip teacher educators for transformative possibilities.

Conceptualizing teacher educator development as a learning problem

Questions about teacher educator development and the design of teacher educator preparation are fundamentally questions about learning—what it aims to achieve and how it is organized (Korthagen, 2020). Thus, before outlining the proposed design, I turn to a brief discussion of learning theory. With the goal of catalyzing transformation in teacher education through the preparation of equity- and justice-oriented teacher educators, I lean on two learning theories: cultural historical activity theory (CHAT) (Vygotsky, 1978) and critical pedagogy (Freire, 1970). Through their complementary emphases on culture and power's roles in learning, these theories offer special purchase for grappling with equity and justice issues and for exploring transformative possibilities for the field.

The value of CHAT in theorizing teacher educators' learning and activity

As a situated learning theory, CHAT treats the development of learners, in this case teacher educators, as a social process that unfolds within the context of collective activity (Engeström, 1999, 2001). In viewing learning as collective, social, and contextual, CHAT draws attention to factors that mediate teacher educator learning, including, for example: state and district policies shaping K-12 teaching and teacher education practices; the stated missions, student populations, and geographic locations of institutions where Novice Teacher Educators (NTEs) engage in doctoral work and teacher educating; and the various participants in teacher education's activity system, including peers, teacher candidates, faculty, K-12 learners, and the local schools and communities to whom teacher education is accountable. This framing encourages learning experiences that treat NTEs as dynamic learners and cultural beings, whose identities and lived experiences and understandings mediate their learning in consequential ways. It additionally underscores the potential power of learning experiences that provide NTEs with opportunities to collectively identify and explore questions and dilemmas that emerge from the institutional, political, and social contexts of their own teacher educating experiences.

CHAT also views transformation and learning as interconnected. CHAT suggests that when existing systems "appropriate" newly introduced influences, disruptions in need of resolution emerge, which, in turn, generate opportunities for (a) expansive learning among the system's participants and (b) the system's transformation (Engeström, 1987, 1999, 2001). This conceptualization of learning underscores the transformative potential of learning experiences designed to help NTEs surface dilemmas or conflicts from their work and then muddle together toward resolutions that can engender disturbances to the system—in this case, teacher education's "historically sedimented" conditions and practices (Souto-Manning and Martell, 2019). Indeed, CHAT inspires us to imagine how we might leverage disruptions as openings for "future oriented innovative solutions" to teacher education's most stubborn challenges including those related to preparing equity- and justice-oriented teachers (Jóhannsdóttir, 2010, p. 169).

Critical pedagogy as a complementary frame for considering teacher educators' learning

Critical Pedagogy likewise views learning as a collective, social process, but as an explicitly liberatory theory, it reminds us that education is power laden and tends to socialize students into the status quo (Freire, 1970). To overcome these tendencies, critical pedagogy implores educators to use education to “desocialize” (Shor, 1987) through a process whereby learners expand their consciousness of the socio-political context and their positions and roles therein through dialogic exchanges grounded in lived experience. Paired with reflective action—what Freire termed “praxis”—this critical cycle of learning can support participants to transform their worlds (Freire, 1970; Souto-Manning, 2010).

Like CHAT, critical pedagogy points to the potential benefits of learning experiences for NTEs that are grounded in their own subjectivities and lived realities and that offer opportunities for collective problem solving in the service of transformation, such as, imagining and working toward a more just future for teacher education. Given its aim to name and challenge oppression, critical pedagogy also encourages learning experiences that invite explorations of how teacher educators—novices and veterans alike—are implicated in teacher education’s perpetuation of oppression and inequality.

Notably, CHAT and critical pedagogy’s notions of learning dovetail with scholarship that suggests that to disrupt the status quo, teacher education must resist treating knowledge as fixed or unproblematic, as well as assumptions that learning to become a responsive teacher can occur through “training” in the technical methods of teaching (Bartolomé and Trueba, 2000; Ben-Peretz, 1995; Korthagen, 2020). Goodwin et al.’s (2014) application of Cochran-Smith and Lytle’s (1999) knowledge domains for teaching similarly underscores the risks associated with privileging knowledge *about* teaching (“knowledge-for-practice”) in learning experiences for teacher educators, while pointing to the transformative power of learning experiences that support teacher educators to bring into conversation knowledge-for-practice and understandings acquired through their own teacher education practice (“knowledge-in-practice”) to generate “knowledge-of practice.”

In the subsequent section, I draw on this theoretical framework to offer design recommendations for preparing equity- and justice-oriented teacher educators. While the proposal aims to advance thinking and practice across institutions, CHAT and critical pedagogy’s notions of transformation discourage “a completed recipe for ‘rolling out’ or ‘scaling up’” teacher educator development, “as the pervasive metaphors of reform might have it” (Ellis and McNicholl, 2016, p. x). Contrary to more technical conceptions of change, the designs advanced in this chapter are meant to be taken up rather than scaled up; this includes reflecting and responding to NTEs’ identities and subjectivities and to the communities and contexts where learning is situated.

A doctoral curriculum for preparing equity- and justice-oriented teacher educators

This design proposal builds on the thoughtful work of scholars who have advocated for deliberate pathways into teacher education (e.g., Cochran-Smith, 2003; Goodwin and Kosnick, 2013; Hollins et al., 2014; Korthagen, 2020; Loughran, 2006; Stillman and Anderson, 2014; Stillman and Beltramo, 2019; etc.). Unlike accounts which explore a single dimension of teacher educator preparation (e.g., a pedagogy of teacher education course), however, this proposal offers recommendations for a reconfiguration of the positioning of doctoral students within programs of teacher education and a doctoral-level curriculum that can be re-contextualized across settings. Heeding calls to center “the why” of teacher educator preparation in program design (Conklin, 2015) and research that points to concerning limited understandings about “multiculturalism” and “diversity” among practicing teacher educators (Goodwin et al., 2014), this proposal aims to offer insights for preparing equity- and justice-oriented teacher educators who can provoke radical change in the field.

It is important to note that the proposed curriculum represents a deliberate departure from traditional doctoral education which privileges research training and treats graduate student appointments in teacher education as merely labor. These more traditional programs frame teacher educating as intuitive, low-skilled work that any doctoral student could do temporarily and assume that disciplinary expertise or experience in K-12 schools are acceptable proxies for preparedness to teach teachers. These assumptions have been shown to contribute negatively to teacher educator efficacy and teacher education quality.

Instead, this design assumes that to advance the field, teacher educators must become skilled not only as teacher education researchers but also as practitioners of teacher educating who can meaningfully link scholarship and practice (Cochran-Smith and Zeichner, 2005; Goodwin and Kosnick, 2013). These abilities are especially critical to efforts to transform the field, given the need for specialized knowledge about teaching and teacher educating for equity and justice, and the need for more and better research, grounded in problems of practice, about equity- and justice-oriented teacher education (Beltramo et al., 2020; Cochran-Smith, 2003; Goodwin et al., 2014; Stillman and Beltramo, 2019).

Whereas articulating the why and how of teacher educator preparation is relatively straightforward, questions about what teacher educator preparation ought to involve are harder to resolve. As many have argued, the field’s limited agreement about what teacher educators need to know and be able to do hampers efforts to codify a knowledge base for teacher educators (Goodwin and Kosnick, 2013; Goodwin et al., 2014; Korthagen, 2020). Others, meanwhile, have raised questions about the very premise of identifying a unitary knowledge base for teacher educators, given the mediating roles of identity and context in teacher educating and teacher educator learning, and the risks to equity and justice that a predetermined knowledge base may impose (Stillman and Beltramo, 2019; Stillman et al., 2019a,b).

What does equity- and justice-oriented teaching educating involve?

Any attempt at curriculum/program design demands some delineation of the work that equity- and justice-oriented teacher educators are responsible for. Several scholars have suggested competencies that all teacher educators need (Goodwin and Kosnick, 2013; Koster et al., 2005) with more recent empirical work identifying teacher educators' central tasks (Goodwin et al., 2014). Goodwin and Kosnick (2013), for example, organize teacher educator competencies into five knowledge categories: personal, contextual, pedagogical, sociological, and social. These categories complement Koster et al.'s (2005) competency and task domains, which similarly underscore teacher educators viewing themselves as learners (personal knowledge) and having the ability to design and deliver responsive programs (pedagogical, contextual, and sociological knowledge), where they must draw on communicative and reflective competencies (social knowledge), while participating in policy development and navigation (contextual knowledge). The description of teacher educator responsibilities I offer embraces and expands on these framings, focusing specifically on expectations for equity- and justice-oriented teacher educators.

Building on accounts of how racism and whiteness manifest in teacher education, I argue that equity- and justice-oriented teacher educators' central responsibility is to disrupt racism and the "overwhelming presence of whiteness" in the field through the preparation of anti-racist teachers. Here I borrow from Ohito's (2020) definition of antiracism through which she explains, "Antiracism is the practice of resisting or opposing racism and/or intervening in ways that subvert its impact and relax its hegemonic clasp on persons, institutions, schools, and other such entities in society. Antiracist teaching engenders the pedagogical application of an antiracist stance" (p. 19).

Preparing anti-racist teachers unfolds for teacher educators at the level of courses, programs, and institutions, and in relation to local and national policy. At the very heart of this work is developing teacher education programs that prepare teachers to "challenge the field's 'pathology of deficiency' (Ladson-Billings, 2017), and instead honor, build upon, and sustain the knowledges, beliefs, and voices of intersectionally minoritized students" (Stillman and Beltramo, 2019, p. 26). An extensive body of literature animates this charge, with emphasis on teachers recognizing, centering, and leveraging minoritized students' knowledges, cultural practices, experiences, and interests as resources for learning, rather than suppressing or exchanging them for knowledge deemed higher-status or academic (Alim and Paris, 2017; Ladson-Billings, 1995; Moll et al., 1992). Much of this work centers the dynamic role of culture in learning and calls on teachers to honor students' humanity by "produc[ing] students who can achieve academically" (Ladson-Billings, 1995, p. 474), while maintaining their "cultural integrity" (p. 476).

Many agree that equity- and justice teacher educators must also support teacher candidates to develop critical consciousness—an ability to engage in a structural analysis of inequality (Freire, 1970; Shor, 1987; Souto-Manning, 2010)—as well as the "ideological and political clarity" (Bartólome and Trueba, 2000) needed to operate as agents of change. Recent scholarship, especially, illustrates the importance of teachers engaging in advocacy in solidarity with minoritized students, families, and communities (Guillén and Zeichner, 2018; Love, 2019; Souto-Manning and Martell, 2019). These understandings and sense of agency come together when equity- and justice-oriented teachers exercise what Horn and Campbell (2015) termed "pedagogical judgment," whereby teachers learn to draw on their "situated knowledge of their particular teaching context" to "work nimbly" to respond to the "perennial puzzles of teaching" (p. 155).

In addition to developing curriculum, equity- and justice-oriented teacher educators must have the understandings and skills needed for de-centering whiteness and challenging deficit framings, while centering the knowledges of teacher candidates of color in course discussions and in feedback on candidate assignments (Brown, 2014; Sleeter, 2017; Stillman and Beltramo, 2019; Stillman et al., 2019a,b). This key dimension of teacher educator mediation, at the center of which is teacher educators' communication competencies (Ohito, 2020), often goes unrecognized in discussions of teacher educators' work (Anderson and Stillman, 2013a; Stillman and Anderson, 2016; Stillman and Beltramo, 2019).

Equity- and justice-oriented teacher educating additionally involves developing awareness of and confidence to challenge racism and whiteness across teacher education spaces, beyond courses (Stillman et al., 2019b). This includes drawing on equity- and justice-oriented understandings to inform program development and work surrounding candidate field experiences (Anderson and Stillman, 2013b; Guillén and Zeichner, 2018). Critically navigating (and resisting) institutional, policy and political conditions—increasingly crafted in the service of neo-liberal, technocratic goals that take a race-neutral view of education—presents yet another layer of teacher educator responsibility (Zeichner, 2010).

Scholarship is also clear about the importance of teacher educators—especially those committed to disrupting the status quo—wearing the complementary hats of teacher educator practitioner and researcher (Cochran-Smith, 2003; Zeichner, 2005). Not only does studying teacher education, including one's own practice, serve as vehicle for improving one's own pedagogy (Korthagen, 2020), it additionally contributes to teacher educators' identities as change agents, capable of getting and doing "better" (Lytle, 2008).

These bodies of research provide direction for designing preparation experiences that equip NTEs for transformative, anti-racist practice. In the subsequent sections, I expand on these accounts and their intersections to propose a design that aims to advance the field's thinking about doctoral education programs designed to prepare and support equity- and justice-oriented teacher educators. Throughout, my recommendations reflect my position that a neutral approach to teacher educator development—while still a better option than no formal preparation—falls short of addressing the field's most urgent needs, and that transforming the field requires learning experiences organized for that purpose.

Three dimensions of doctoral education for equity- and justice-oriented NTEs

The curriculum design includes three overarching components: (1) Teacher Educating for Equity and Justice, (2) Equity- and Justice-oriented Doctoral Coursework, and (3) Critical Inquiry Communities. Fig. 1 provides a visual representation of this design.

The triangular representation, with double-sided arrows connecting the different points, illustrates components' potential interactions with one another. One such interaction, for example, might involve NTEs learning to conduct research on teacher education through coursework, and then engaging in such research—grounded in questions and dilemmas identified from one's own teacher educating—within a critical inquiry community.

The concentric circles surrounding the triangle, meanwhile, illustrate the mediating roles of the institutional and policy contexts within which teacher educator development occurs. This visual intends to communicate that learning and teacher educating unfold contextually, that practicing across different settings requires re-contextualizing rather than replicating practice, and that learning to engage critically with institutional and political forces is central to teacher educators' work and learning. I unpack each of the three major components below.

Teacher educating for equity and justice

Teacher education programs commonly hire doctoral students as adjunct instructors and fieldwork supervisors, often depending on them to keep programs afloat. Qualifications for these positions typically include disciplinary expertise and recent or current experience as a practicing teacher. Although such knowledge and expertise may be preconditions for teacher educating, they are not proxies for teaching about teaching (Feiman-Nemser and Buchmann, 1985). Nevertheless, these students tend to be hired "as is," perhaps with opportunities to engage in generic onboarding experiences (e.g., focused on learning management systems, syllabi requirements, teaching college students, etc.), but rarely with access to ongoing support or learning focused on teacher education's complexity. This practice poses significant risks to teacher education quality, in general, but is especially problematic in relation to equity- and justice-oriented teacher educating which requires teacher educators' specialized knowledge and critical treatment—rather than ready acceptance—of traditional teacher education structures and practices.

This first component of the proposed design—Teacher Educating for Equity and Justice—recognizes that even the most novice teacher educators occupy vital positions as equity- and justice-oriented educators of teachers and therefore aims to reconfigure the positioning of doctoral students within teacher education programs. Instead of hiring doctoral students as a means of responding to labor demands or as a funding fallback for students with little interest in teacher education, this design both requires and guarantees NTEs' access to multiple, varied, extended, and supported opportunities to practice equity- and justice-oriented teacher educating. (This points to benefits of doctoral programs' proximity to equity- and justice-oriented teacher preparation programs.)

Contrary to sink-or-swim induction models common in teacher education, support would accompany NTEs through these experiences via individual mentoring by faculty or more advanced doctoral students, and mediation that the other two curriculum components—coursework and critical inquiry communities—provide. For example, specialized courses (described below) would provide opportunities for NTEs to draw connections between their experiences as teacher educators and course themes and questions. Participation in a critical inquiry community would offer a critical, collegial, environment for facilitating NTEs' inquiry into

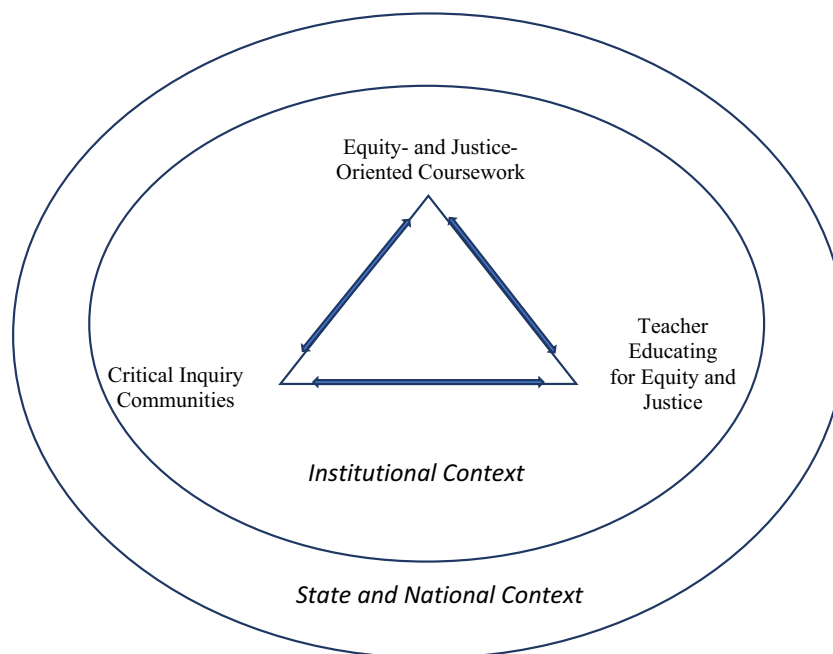


Fig. 1 A three-pronged design for specialized doctoral education

their own practice in context (also described below). These learning experiences might also be leveraged as opportunities to support (novice) school-based teacher educators (i.e., mentor teachers in K-12 settings), who are similarly almost always expected to learn to teacher educate on their own. Bringing school- and university-based teacher educators together could in turn improve equity- and justice- oriented teacher educating across contexts and strengthen university-school partnerships, with NTEs playing a genuine role in this crucial, relational work.

Relatedly, reconfiguring doctoral students' positions in teacher education can also expand the types of learning experiences equity- and justice-oriented NTEs have access to. In addition to teaching individual courses, doctoral students might, for example, be supported and compensated to contribute to curriculum development efforts aimed at deeper, equity- and justice- oriented program-level change. Meanwhile, opportunities to engage in equity- and justice-oriented practices beyond course instruction—for example, collaborating with and learning from local families and communities to rethink/reorganize programs (Zeichner et al., 2016)—would equip NTEs with experiences needed for transforming teaching education in their future positions, while also helping programs where they are working during doctoral study move toward systemic change. Notably, all of these experiences could provide rich material for collaborative studies, potentially explored through course assignments and participation in critical inquiry communities.

Unlike programs where doctoral students pass through the revolving door of adjunct instructors, this design also aims to support NTEs over time. Not only would this allow for growth and setbacks, NTEs would also be able to participate as full members of the teacher education community during their time as doctoral students. This stands to provide NTEs with authentic and supported opportunities to grapple with teacher education's many layers of complexity and to position them as leaders capable of guiding equity- and justice-oriented change.

Equity- and justice- oriented doctoral coursework

Coursework resides at the center of doctoral education, often dominating the first two or more years of study. This dimension of the proposal considers possibilities for four specialized courses for facilitating NTEs' learning about preparing equity- and justice-oriented teachers and challenging racism in the profession, more broadly. These include:

- Critical perspectives on the history and context of teacher education
- Becoming a teacher education researcher
- Understanding and designing for teacher learning
- Pedagogies of equity- and justice-oriented teacher education

Because my goal is to stimulate dialog about, rather than present a recipe for, the preparation of equity- and justice-oriented teacher education, my descriptions are merely hypotheticals that I hope provide fodder for necessary discussions in the field. One assumption I make is that these courses would run as seminars, where authentic dialog and collective problem solving could occur.

Critical perspectives on the history and context of teacher education

Scholars have argued that teacher educators' abilities to challenge the status quo of schooling hinge on opportunities for deep (un) learning about structural inequality, and about how historic and current policies and institutional conditions shape the profession (Cochran-Smith, 2003; Cochran-Smith et al., 2020; Sleeter, 2017). This course would aim to fulfill that purpose. The course, for example, might explore how the profession emerged, and how shifting politics and policies have shaped the field over time. Rather than taking a neutral stance on these issues, the course would center how racism and other forms of oppression have been perpetuated in and through teacher education and offer examples, drawn from critical scholarship, of what anti-racist teacher education could look like, including the policy and institutional changes—and teacher educators' potential roles in shaping them—needed to support it.

The course might also ask NTEs to tap into their teacher educating experiences as generative contexts for bringing these topics to life. For instance, NTEs might trace the history of a current a teacher education policy or practice of concern (e.g., the use of standardized assessments for licensing purposes; pressure to privilege the "science of reading" over responsive literacy instruction; etc.) and explore how it undermines and/or contributes to equity and justice goals. A next step might involve working with peers to imagine how to critically navigate the policy or practice, perhaps experimenting with these imaginings in and through their own practice. Such an assignment stands to deepen NTEs' understanding of the profession's current architecture, while providing opportunities to appropriate the identity of policy co-author, challenging the idea that teacher educators must accept policies and practices as permanent fixtures (Stillman and Anderson, 2015).

Becoming a teacher education researcher

The field is clear about the benefits of teacher educators becoming knowledgeable about teacher education's research base as well as active researchers of teacher education themselves (Korthagen, 2020). Such benefits include preparedness to make informed rather than anecdotal decisions about teacher education practice (Zeichner, 2005) and greater inclinations to critically analyze existing systems, structures, and practices, rather than accepting them as unproblematic (Cochran-Smith, 2003). In this respect, research facility may be crucial to teacher educators' abilities to challenge the field's conservatism, including its perpetuation of racism. Yet, while most university teacher educators complete research-focused doctoral programs, few are well-versed in teacher education scholarship, and few report applying their research knowledge to the study of teacher education (Goodwin et al., 2014).

This course aims to address this problem, building on more general methodology courses to support NTEs to design and conduct justice-oriented teacher education research. A key element of the course would involve exploring self-study methodologies, given their significance in building an empirical foundation for the field (Berry, 2007; Korthagen, 2020; Zeichner, 1999) and the opportunities they provide for practicing teacher educators to embody a critical learning stance (Cochran-Smith and Lytle, 2009). As a methodology that illustrates “that at the core of expert practice is not the strict implementation of evidence-based practices, but the need to make subtle judgments in unique situations” (Korthagen, 2020, p. 333), self-study is especially well-suited for research that advances equity and justice.

To work toward these learning goals, an assignment for this course might require NTEs to conduct practitioner inquiry projects in relation to their own teacher educating, specifically exploring questions about racism and other forms of oppression, and with particular attention to how they are implicated in perpetuating and disrupting oppression in and through practice. Engaging in self-study within the confines of a semester-long course wouldn’t allow for extensive study. However, it would provide NTEs with a foundation for later self-studies and a built-in inquiry community where collective, critical sensemaking could be facilitated. These activities would equip NTEs to participate in Critical Inquiry Communities (described below).

Despite self-study’s powerful contribution to local and public knowledge (Cochran-Smith and Lytle, 2009), scholars have also underscored its limitations, including studies’ disconnections from one another, and lacking methodological variation (Cochran-Smith and Zeichner, 2005; Korthagen, 2020). Thus, this course would offer NTEs opportunities to engage with multiple methodologies—including critical approaches to quantitative research—for the purpose of studying and transforming teacher education. The course would also privilege critical theoretical methodological and theoretical frameworks (e.g., Critical Race Theory, feminist theories, queer theory, decolonial theories; etc.) to generate understandings about teacher education, currently, and actionable insights for radical change.

Understanding and designing for teacher learning

Ironically, studies of teacher education rarely make explicit a view of teacher learning (Anderson and Stillman, 2013a; Clarke and Hollingsworth, 2002). As Korthagen (2020) explains, this tendency weakens the study and practice of teacher education by compromising efforts to delineate a pedagogy of teacher education and obscuring the difference between student and teacher learning, a key distinction in teacher educating (Feiman-Nemser and Buchmann, 1985; Murray and Male, 2005). Failing to attend explicitly to teacher learning in teacher education research additionally leaves us with vague notions of what teacher education aims to achieve and how. This is especially troubling considering the urgent need to grapple as a field with racism and whiteness, which requires establishing both clear learning goals for teacher candidates and learning experiences capable of facilitating candidates’ learning to practice responsive, anti-oppressive teaching.

This course, therefore, would aim to equip NTEs to center teacher learning in their own work. In particular, the course would introduce NTEs to, and/or build on existing knowledge from previous doctoral coursework on, a range of learning theories that could be applied to the study and practice of teacher education. A possible assignment for facilitating this learning, for example, might expect NTEs to critically analyze existing scholarship and local practices through the lens of learning theories and frameworks. Such an assignment would provide NTEs with opportunities to discern how different conceptions of learning and approaches to teacher educating may be conducive to preparing equity- and justice-oriented teachers, capable of exercising pedagogical judgment. A subsequent assignment, meanwhile, might challenge NTEs to link learning theory with program and course design. NTEs, for example, could be encouraged to draw across dilemmas of practice from their own teacher educating and various learning theories to craft possible course and program designs for facilitating teacher candidates’ equity- and justice-oriented learning.

Pedagogies of equity- and justice-oriented teacher education

No area of teacher educator development has received more attention in the literature than the pedagogy of teacher education (e.g., Goodwin and Kosnick, 2013; Korthagen, 2020; Loughran, 2006). Indeed, many have suggested that the ability to formally prepare teacher educators depends on codifying the knowledge necessary for teacher educating (Grossman et al., 2009).

Yet, there are contradictions inherent to identifying a universal knowledge base if the aim of teacher educator preparation is to transform teacher education rather than perpetuate its current structures and practices. Put another way, what might be the limitations of supporting teacher educators to dismantle the profession’s current architecture, to re-envision how and what the profession could be, using a knowledge base drawn from the profession as we know it? Relatedly, there is good reason to question the very premise of a universal knowledge base for teacher education given the transformative learning that can emerge when NTEs center their own subjectivities to co-construct teacher educator knowledge “from the ground up” (Stillman et al., 2019a,b).

The challenges and rich debates surrounding the codification of a knowledge base for teacher educators, alongside tendencies to treat disciplinary expertise as sufficient qualification for teaching teachers, are a few key reasons why a course dedicated to the critical exploration of teacher education pedagogies is essential to teacher educator preparation. Additional reasons include widely accepted understandings that teaching and teacher educating are distinct activities (Feiman-Nemser and Buchmann, 1985), since becoming a teacher educator requires learning about both “first-order,” K-12 classroom teaching, and “second-order” teaching, or teaching about that teaching (Murray and Male, 2005). Like Korthagen’s (2020) argument that challenging the status quo in teacher education hinges on critical understandings of teacher education pedagogy, I argue that efforts to transform teacher education into a more equity- and justice-oriented enterprise rests on teacher educators’ dynamic and critical pedagogical knowledge. Indeed, teacher educators must also exercise their own pedagogical judgment (Horn and Campbell, 2015) as they make contextually sensitive decisions about how to prevent and respond to racism and whiteness in courses and beyond.

Within an anti-racist teacher educator preparation program, any course on the pedagogy of teacher education would need to help NTEs deepen their knowledge of first-order teaching for equity and justice through opportunities to both unlearn allegedly “best practices” now understood to advantage white, English dominant students, and to engage in deep learning about multicultural education and culturally sustaining and critical approaches to teaching. Since these pedagogies continue to evolve, and NTEs bring with them a diverse range of identities, knowledges, experiences, and lived realities, learning experiences would need to make space for and leverage varied degrees of expertise and understanding.

This course would additionally offer learning experiences designed to foster collaborative, critical explorations of second order equity- and justice-oriented teaching. Korthagen (2020) recently outlined widely accepted teacher education pedagogical strategies, such as case methods (Richert, 1991), video analysis (van Es and Sherin, 2010), promotion of reflection (Zeichner and Liston, 2013), teacher identity work (Britzman, 2003), and practitioner research (Cochran-Smith and Lytle, 2009), among others. Although these approaches were not derived with the intention of moving teacher candidates toward anti-racist teaching, critically exploring them could provide NTEs with opportunities to consider how they stand to perpetuate race-neutral teacher education (Brown, 2014) or other forms of oppression, and/or open possibilities to stimulate anti-oppressive teaching. As one example, although Korthagen (2020) and others have lauded approximations of practice (Grossman et al., 2009) as a promising pedagogy for all, others have raised concerns about the pedagogy’s potential to impose harm on students from non-dominant communities (Beltramo et al., 2020; Daniels and Varghese, 2019; Philip et al., 2018).

While a more general approach to teacher educator preparation might stop with these widely accepted pedagogies, a program designed to facilitate equity- and justice-oriented NTEs’ development would require NTEs to explore pedagogical strategies focused on anti-racist teacher education (e.g., see Picower, 2021). Conklin (2015, 2020), for example, found that within her own doctoral course on the pedagogy of teacher education, identity circles, autobiographies, and modeling love and compassion seemed to have the greatest impact on NTEs’ equity- and justice-related learning and practice.

In addition, a burgeoning area of scholarship—much of it produced by scholars of color—offers pedagogical strategies explicitly geared toward preparing equity- and justice-oriented teachers. Some examples include principled improvisation (Philip, 2019), race-based caucuses (Varghese et al., 2019), language autobiographies (Athanasios et al., 2019), horizontal partnerships between teacher educators, and teachers and children of color (Souto-Manning and Martell, 2019), pedagogical rehearsal that cultivates epistemic disobedience (Dominguez, 2021), and humanizing pedagogy (Carter-Andrews et al., 2019). To deepen their understandings of these strategies, NTEs could identify dilemmas of practice related to equity and justice and then select a justice-oriented pedagogical strategy to address the dilemma in context. NTEs could then critically analyze the strategy’s use with their own teacher candidates.

Finally, there is some evidence to suggest that an especially challenging and prevalent aspect of equity- and justice-oriented teacher educating involves disrupting such thinking about minoritized students and their communities (Conklin, 2020; Stillman and Beltramo, 2019; Stillman et al., 2019a,b). Although the pedagogical strategies above aim to dismantle deficit framings, opportunities for teacher educators to disrupt such framings also regularly emerge in course discussions and while providing candidates with feedback. This complex mediational work isn’t intuitive, and like other dimensions of teacher educating, requires supported opportunities for NTEs’ critical exploration and practice. As described below, critical inquiry communities offer especially rich learning contexts for this work.

Critical inquiry communities

Supporting teacher educators as practitioners *and* researchers is crucial to any effort to improve the profession. Indeed, many teacher educators’ survival in academe depends on their success in publishing new research. Yet, teacher educators’ working conditions—including heavier than typical teaching loads (Zeichner, 2014)—coupled with tendencies to treat research and practice as distinct (Goodwin et al., 2014)—can constrain teacher educators’ capacities to remain research active. These are all good reasons to include in preparation efforts experiences that cultivate NTEs’ comfort with and commitment to doing teacher education research.

When teacher educators link their research and practice it can also cultivate transformational change (Goodwin et al., 2014). In particular, practitioner inquiry holds promise for facilitating teacher educators’ un-learning and critical identity development—both of which are central to advancing social justice goals (Cochran-Smith et al., 1999). However, Cochran-Smith and Lytle (2009) caution that engaging in practitioner inquiry does not necessarily challenge oppressive systems, structures, and practices, and can just as easily bolster them. For example, teacher educators often elect to engage in self-study so that they can meet publication demands despite limited funding or ample time for other types of research, rather than pursuing self-study to engender genuine learning and change (Cochran-Smith and Villegas, 2016). Cochran-Smith and Lytle (2009), thus, suggest that teacher educators participate in inquiry communities to ensure a critical lens anchors their investigations.

Recognizing the powerful learning that can happen outside of formal courses and within professional communities, this third design component aims to deepen and extend NTEs’ learning from formal coursework, by providing opportunities for NTEs to ground learning in critical investigations of their own teacher educating practice. The community would operate as a dynamic requirement throughout NTEs’ doctoral education, providing a venue to share dilemmas of practice, conduct individual and/or collaborative self-studies, and engage in collective sensemaking. To bring structure, purpose, and criticality to such as space, I propose drawing on critical pedagogical approaches.

In my own work, I have illustrated the potential of Culture Circles (Freire, 1970) and Theater of the Oppressed (*teatro*) (Boal, 1979) to ensure criticality within inquiry communities for equity- and justice-oriented NTEs. Structuring participation according to

Freire's Critical Cycle (Freire, 1970; Souto-Manning, 2010), for example, can support NTEs to: identify dilemmas of practice centrally concerned with equity, justice, and teacher learning; draw connections across questions and dilemmas of practice; co-construct knowledge for teacher educating grounded in their situated identities and experiences; and engage in critical, collective reflection and problem solving geared toward action (Stillman and Beltramo, 2019; Stillman et al., 2019a,b).

But, as Ohito (2020) points out, to enact anti-racism in practice, teacher educators also need experiences that facilitate embodied learning, since possessing an anti-racist ideology does not necessarily lead to anti-racist behavior. *Teatro*, coupled with Culture Circles can facilitate this sort of embodied learning. Specifically, *teatro* has been shown to produce learning intimately tied to NTEs' identities; to help NTEs develop and rehearse using tools for addressing equity- and justice-related challenges; and to cultivate NTEs' courage for challenging deficit ideologies and oppressive systems and structures (Stillman et al., 2019b; Beltramo et al., 2020).

Since transforming teacher education depends upon all participants' learning and improvement, this space would ideally be cross-generational and cross-contextual, including NTEs at varying stages of development, existing university teacher education faculty, and school-based teacher educators. Cross-generational learning is especially crucial in teacher educator development efforts since "practicing teacher educators apparently do not feel adequately prepared for diversity, while ... less experienced teacher educators were more likely to consider diversity and multicultural issues as more important than their more experienced peers" (Goodwin et al., 2014, p. 297). Thus, unlike traditional mentoring relationships whereby allegedly more knowledgeable veteran practitioners support novices, this arrangement would nurture bi-directional relationships and learning. Rotating facilitation each session would help to foster horizontal participation, ensuring all participants are positioned as expert and can develop a sense of ownership and leadership, and would allow for the privileging of multiple perspectives, identities, and topics.

Discussion

In this chapter, I have argued that for teacher education to become an equity- and justice-oriented field, authentically answerable to the communities of color it claims to serve, we must invest in the preparation of equity- and justice-oriented teacher educators. After describing how oppression and injustice, particularly related to racism and whiteness, manifest in US teacher education programs, I proposed a three-pronged framework for a specialized doctoral curriculum. Anchored in the situated learning theories of critical pedagogy and CHAT, the proposal offers recommendations for contextualized, future-oriented, and embodied learning experiences aimed at equipping equity- and justice-minded NTEs for the complex work needed to spark change in the profession.

In particular, the hypothetical learning experiences described are designed to support NTEs' engagement in collective problem solving and reflection, guided by course-generated understandings and grounded in dilemmas from NTEs' lived realities and teacher educating. These learning experiences, designed to foster NTEs' development of knowledge of teacher educating, reflect Cochran-Smith and Lytle's (1999) conceptions of teacher learning, as applied to teacher educator learning (Goodwin et al., 2014). My hope is that these recommendations expand discussions about how teacher education and teacher educators are implicated in the perpetuation of racism and other forms of oppression, and catalyze critical dialog about, and action focused on, teacher educator preparation as a lever for creating a different, more just future for the field.

This isn't to suggest that the existence of teacher educator preparation programs can alone resolve teacher education's historically sedimented problems. For one, teacher educator preparation cannot on its own mitigate the harms that a predominately white teacher educator workforce imposes. Without any doubt, hiring and supporting more teacher educators of color is essential to teacher education's transformation (Brown, 2014; Picower and Kohli, 2017; Teachers of Color Collective and Souto-Manning, 2022). Teacher educator preparation can begin to address this problem by recruiting into specialized doctoral programs students of color interested in anti-racist teacher educating. However, just as diversifying the teacher candidate population leaves teacher education's racist roots intact (Brown, 2014), diversifying the teacher education faculty without providing effective preparation or support, while also making other changes—including pressing and supporting white teacher educators to change—cannot eradicate the racist conditions and practices that characterize the profession.

As noted earlier, myriad institutional factors also uphold the field's predominance of whiteness and perpetuation of racism. Inadequate resources, teacher educators' heavy teaching loads and under-compensation, universities' reliance on teacher education to generate revenue, and neoliberal policies all pose challenges to building and sustaining effective, justice-oriented teacher education programs. Although teacher educators can address these problems, time and experience have proven that their efforts alone cannot overcome them.

Of course, these conditions raise questions about the degree to which equity- and justice-oriented transformation in teacher education is even possible. But acknowledging these conditions also provides insights for cultivating teacher educator preparation's success. Figuring out how to make teacher educator preparation a fixture of the profession is critical, as it is hard to imagine any fundamental change occurring without investment in teacher educators' learning—learning that would position teacher educators to generate disruption and transformation from within (Ellis, 2018; Jóhannsdóttir, 2010). This investment will need to take multiple forms.

For one, institutions must recognize and compensate faculty for the actual work required to develop and sustain programs. While the proposed curriculum includes traditional components, like courses, for which faculty can expect compensation (as part of their assigned teaching loads), like teacher education, the curriculum also includes necessary, but less traditional and, therefore, less visible roles and responsibilities. A few clear examples include multi-year faculty mentoring roles and the critical inquiry community convener, for which faculty are less likely to receive compensation. These responsibilities take enormous time and

energy and cannot simply be added teacher educators' existing workloads (Gunn et al., 2016). Thus, it would behoove faculty to preemptively—during program development—identify and make official the many roles and responsibilities that teacher educator preparation demands so that they can be accounted for when assigning courses and service responsibilities.

Institutions must also invest in their existing faculty's learning. Few universities have concentrations of faculty with the knowledge, expertise, and will needed for designing and running preparation programs for equity- and justice-oriented NTEs. In addition, most current teacher educators have fallen into the profession themselves, likewise learning on the job without structured opportunities to study the profession or their own practice. Recent research indicating that much of the predominately white teacher educator force is not knowledgeable about or interested in learning about multicultural education and diversity only compounds these issues (Goodwin et al., 2014).

These realities suggest that investing in the learning of practicing teacher educators is equally as important as developing the profession's next generation. Some such opportunities can be built into teacher educator preparation programs, for example, through the creation of cross-generational learning spaces like the critical inquiry communities described earlier. But to transform the field, practicing teacher educators additionally need access to and support for engaging in a range of professional learning experiences centered on anti-racist teacher education and critical examinations of their own practice, all with an eye toward how they would mentor others into the field.

Conclusion

If we hope to transform teacher education, teacher educator preparation is essential, and will require greater investment in the field. But these investments are not just about workloads, compensation, and learning opportunities, though all matter greatly. Rather, investments must also involve changing the nature of the profession through the humanization of teacher educators. Students from the doctoral program for NTEs at my own institution perhaps made this point best in an article they coauthored after conducting practitioner inquiry projects while engaging together in Freirean Culture Circles and Boalian theater (Freeman et al., 2020).

In the article they recount their own struggles to humanize teacher candidates at a predominately white institution, when so many of them expressed deficit ideologies, often about these NTEs' own communities. Through their individual inquiries and collective sensemaking, these NTEs began to link their own capacities for humanization to the degree to which they felt humanized in their own learning. They ultimately argued that the humanization of teacher educators is crucial to advancing teacher education toward equity and justice.

I sincerely hope that investing in the next generation of teacher educators not only sets the stage for transforming teacher education into a genuinely anti-racist enterprise, but that the next generation of teacher educators also experiences a better, more humanizing work life than the generations before them. Investing in teacher educators' learning and treating the profession as worthy of such support, is a step in that direction.

Acknowledgments

Many of the ideas in this manuscript represent collective thinking with valued colleagues, especially Dr. Melissa Braaten and Dr. Mariana Souto-Manning. I also extend my deepest gratitude to former and current doctoral students/NTEs at the University of Southern California and the University of Colorado Boulder, with whom I have had the distinct privilege to work and learn. Your brilliance, passion, and hard work move us closer to a more just and equitable world.

References

- Alim, S., Paris, D., 2017. What is culturally sustaining pedagogy and why does it matter? In: Paris, D., Alim, S. (Eds.), *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York, NY, pp. 1–21.
- Anderson, L.M., Stillman, J., 2013a. Student teaching's contribution to preservice teacher development: a review of research focused on the preparation of teachers for urban and high-needs contexts. *Rev. Educ. Res.* 83 (1), 3–69.
- Anderson, L., Stillman, J., 2013b. Making learning the object: using cultural historical activity theory to analyze and organize preservice field placements in urban, high-needs schools. *Teach. Coll. Rec.* 115 (3), 1–36.
- Athanases, S., Banes, L.C., Wong, J.W., Martinez, D.C., 2019. Exploring linguistic diversity from the inside out: implications of self-reflexive inquiry for teacher education. *J. Teach. Educ.* 70, 581–596.
- Bartolomé, L., Trueba, E.T., 2000. Beyond the politics of schools and the rhetoric of fashionable pedagogies: the significance of teacher ideology. In: Trueba, H.T., Bartolomé, L.I. (Eds.), *Immigrant Voices: In Search of Educational Equity*. Rowman & Littlefield, New York, pp. 277–292.
- Beltramo, J., Stillman, J.L., Ahmed, K.S., 2020. From approximations of practice to transformative possibilities: using theater of the oppressed as rehearsals for facilitating critical teacher education. *N. Educ.* 16 (1), 25–44.
- Ben-Peretz, M., 1995. Curriculum of teacher education programs. In: Anderson, L.W. (Ed.), *International Encyclopedia of Teaching and Teacher Education*. Elsevier Science/Pergamon, Oxford/New York/Tokyo, pp. 543–547.
- Berry, A., 2007. Reconceptualizing teacher educator knowledge as tensions: exploring the tension between valuing and reconstructing experience. *Stud. Teach. Educ.* 3, 117–134.
- Berry, A., 2009. Professional self-understanding as expertise in teaching about teaching. *Teach. Teach.* 15, 305–318.
- Boal, A., 1979. *Theater of the Oppressed*. Urizen Books, New York.
- Britzman, D.P., 2003. *Practice Makes Practice: A Critical Study of Learning to Teach*. SUNY Press, Albany.

- Brown, K., 2014. Teaching in color: a critical race theory in education analysis of the literature on preservice teachers of color and teacher education in the U.S. *Race Ethn. Educ.* 17, 326–345.
- Carter Andrews, D.J., 2021. Preparing teachers to be culturally multidimensional: designing and implementing teacher preparation programs for pedagogical relevance, responsiveness, and sustenance. *Educ. Forum* 85, 416–428.
- Carter Andrews, D.J., Brown, T., Castillo, B.M., Jackson, D., Vellanki, V., 2019. Beyond damage-centered teacher education: humanizing pedagogy for teacher educators and preservice teachers. *Teach. Coll. Rec.* 121, 1–28.
- Clarke, D., Hollingsworth, H., 2002. Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18, 947–967.
- Chang-Bacon, C.K., 2022. “We sort of dance around the race thing”: race-evasiveness in teacher education. *J. Teach. Educ.* 73, 8–22.
- Cochran-Smith, M., 2000. Blind vision: unlearning racism in teacher education. *Harv. Educ. Rev.* 70, 157–190.
- Cochran-Smith, M., 2003. Learning and unlearning: the education of teacher educators. *Teach. Teach. Educ.* 19, 5–28.
- Cochran-Smith, M., Lytle, S., 1999. Relationships of knowledge and practice: teacher learning in communities. In: Griffin, G. (Ed.), *Rev. Res. Educ.* 24, 249–305.
- Cochran-Smith, M., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York.
- Cochran-Smith, M., Villegas, A.M., 2016. Research on teacher preparation: charting the landscape of a sprawling field. In: Gitomer, D., Bell, C. (Eds.), *Handbook of Research on Teaching*, fifth ed. American Educational Research Association, Washington, DC, pp. 439–547.
- Cochran-Smith, M., Zeichner, K.M. (Eds.), 2005. *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education*. Lawrence Erlbaum, Mahwah, NJ.
- Cochran-Smith, M., Albert, L., Dimattia, P., Freedman, S., Jackson, R., Mooney, J., 1999. Seeking social justice: a teacher education faculty’s self-study. *Int. J. Leader. Educ.* 2, 229–253.
- Cochran-Smith, M., Barnatt, J., Lahann, R., Shakman, K., Terrell, D., 2008. Teacher education for social justice: critiquing the critiques. In: Ayers, W., Quinn, T.M., Stovall, D. (Eds.), *Handbook of Social Justice in Education*, first ed. Routledge, New York, pp. 625–639.
- Cochran-Smith, M., Grudnoff, L., Orland-Barak, L., Smith, K., 2020. Educating teacher educators: international perspectives. *N. Educ.* 16, 5–24.
- Conklin, H.G., 2015. Preparing novice teacher educators in the pedagogy of teacher education. *Action Teach. Educ.* 37, 317–333.
- Conklin, H.G., 2020. The preparation of novice teacher educators for critical, justice-oriented teacher education: a longitudinal exploration of formal study in the pedagogy of teacher education. *Teach. Teach.* 26, 491–507.
- Daniels, J.R., Varghese, M., 2019. Troubling practice: exploring the relationship between whiteness and practice-based teacher education in considering raciolinguicized teacher subjectivity. *Educ. Res.* 49, 56–63.
- Dominguez, M., 2021. Cultivating epistemic disobedience: exploring the possibilities of a decolonial practice-based teacher education. *J. Teach. Educ.* 72 (5), 551–563.
- Ellis, V., 2018. Transformation: What Exactly Does It Mean? Available at: <https://vivellis.org/2018/04/15/transformation-what-exactly-does-it-mean/>. (Accessed 8 February 2022).
- Ellis, V., McNicholl, J., 2016. *Transforming Teacher Education: Reconfiguring the Academic Work*. Bloomsbury, London, UK.
- Engeström, Y., 1987. *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Orienta-Konsultit, Helsinki.
- Engeström, Y., 1999. Activity theory and individual and social transformations. In: Engeström, Y., Miettinen, R., Punamäki, R.L. (Eds.), *Perspectives on Activity Theory*. Cambridge University Press, Cambridge, pp. 19–38.
- Engeström, Y., 2001. Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14, 133–156.
- Feiman-Nemser, S., Buchmann, M., 1985. Pitfalls of experience in teacher preparation. *Teach. Coll. Rec.* 87, 53–65.
- Freeman, Q., Flores, R., Garzón, D., Gumina, D., Sambolin-Morales, A.N., Silva Diaz, E., Stamatis, K.M., 2020. Collaborating towards humanizing pedagogies: culture circles in teacher educator preparation. *N. Educ.* 16, 86–100.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum Publishing, New York.
- Goodwin, A.L., Kosnik, C., 2013. Quality teacher educators = quality teachers? Conceptualizing essential domains of knowledge for those who teach teachers. *Teach. Dev.* 17, 334–346.
- Goodwin, L., Smith, L., Souto-Manning, M., Cheruvu, R., Tan, M.Y., Reed, R., Taveras, L., 2014. What should teacher educators know and be able to do? Perspectives from practicing teacher educators. *J. Teach. Educ.* 284–302.
- Grossman, P., Hammerness, K., McDonald, M., 2009. Redefining teaching, re-imagining teacher education. *Teach. Teach. Theor. Pract.* 15, 273–289.
- Guillén, L., Zeichner, K., 2018. A university-community partnership in teacher education from the perspectives of community-based teacher educators. *J. Teach. Educ.* 69, 140–153.
- Gunn, A., Berg, D., Haigh, M., Hill, M., 2016. *Work of Teacher Educators: Teaching and Learning in New Zealand University-Based Initial Teacher Education*. Teaching & Learning Research Initiative. Available at: <http://www.tlri.org.nz/sites/default/files/projects/TLRI%20Gunn%20End%20of%20project%20report%20-%20web%20ready.pdf>. (Accessed 8 February 2022).
- Haddix, M.M., 2012. Talkin’ in the company of my sistas: the counterlanguages and deliberate silences of Black female students in teacher education. *Ling. Educ.* 23, 169–181.
- Helfand, D.J., 2009. Teaching outside whiteness. In: Steinberg, S.R. (Ed.), *Diversity and Multiculturalism: A Reader*. Peter Lang, New York, pp. 77–96.
- Hollins, E., Luna, C., Lopez, S., 2014. Learning to teach teachers. *Teach. Educ.* 25, 99–124.
- Horn, I.S., Campbell, S.S., 2015. Developing pedagogical judgment in novice teachers: mediated field experience as a pedagogy for teacher education. *Pedagogies* 10, 149–176.
- Jayakumar, U., Howard, T., Allen, W., Han, J., 2009. Racial privilege in the professoriate: an exploration of campus racial climate, retention, and satisfaction. *J. High. Educ.* 80, 538–563.
- Jóhannsdóttir, T., 2010. Deviations from the conventional: contradictions as sources of change in teacher education. In: Ellis, V., Edwards, A., Smagorinsky, P. (Eds.), *Cultural-Historical Perspectives on Teacher Education and Development*. Routledge, New York, pp. 163–179.
- King, E., Butler, B.R., 2015. Who cares about diversity? A preliminary investigation of diversity exposure in teacher preparation programs. *Multicult. Perspect.* 17, 46–52.
- Korthagen, F.A.J., 2020. Pedagogy of teacher education. In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education*, vol. 1. Springer, pp. 311–344.
- Koster, B., Brekelmans, M., Korthagen, F., Wubbels, T., 2005. Quality requirements for teacher educators. *Teach. Teach. Educ.* 21, 157–176.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491.
- Ladson-Billings, G., 1998. Teaching in dangerous times: culturally relevant approaches to teacher assessment. *J. Negro Educ.* 67, 255–267.
- Ladson-Billings, G., 2005. The evolving role of critical race theory in educational scholarship. *Race Ethn. Educ.* 8, 115–119.
- Ladson-Billings, G., 2017. “Makes me wanna holler”: refuting the “culture of poverty” discourse in urban schooling. *Ann. Am. Acad. Polit. Soc. Sci.* 673, 80–90.
- Levine, A., 2006. *Educating School Teachers*. The Education Schools Project. Available at: http://edschools.org/pdf/Educating_Teachers_Report.pdf. (Accessed 8 February 2022).
- Lortie, D., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education*. Routledge, New York.
- Love, B.L., 2019. *We Want to Do More Than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon Press, Boston.
- Lowenstein, K.L., 2009. The work of multicultural teacher education: reconceptualizing white teacher candidates as learners. *Rev. Educ. Res.* 79, 163–196.
- Lytle, S.L., 2008. At last: practitioner inquiry and the practice of teaching: some thoughts on “better”. *Res. Teach. Engl.* 42, 373–379.
- Matias, C.E., 2016. *Feeling White: Whiteness, Emotionality, and Education*. Sense Publishers, Boston, MA.
- Milner IV, H.R., 2008. Critical race theory and interest convergence as analytic tools in teacher education policies and practices. *J. Teach. Educ.* 59, 332–346.
- Milner IV, H.R., Pearman III, F.A., McGee, E.O., 2013. Critical race theory, interest convergence, and teacher education. In: Lynn, M., Dixon, A.D. (Eds.), *Handbook of Critical Race Theory in Education*. Routledge, New York, NY, pp. 339–354.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31, 132–141.
- Murray, J., Male, T., 2005. Becoming a teacher educator: evidence from the field. *Teach. Teach. Educ.* 21, 125–142.

- National Center for Education Statistics, 2019. Percentage Distribution of Enrollment in Public Elementary and Secondary Schools, by Race/Ethnicity and State or Jurisdiction. Available at: https://nces.ed.gov/programs/digest/d19/tables/dt19_203.70.asp?current=yes. (Accessed 8 February 2022).
- Nieto, S., Bode, P., 2011. *Affirming Diversity: The Sociopolitical Context of Multicultural Education*. Longman, New York, NY.
- Ohito, E.O., 2020. Fleshing out enactments of Whiteness in antiracist pedagogy: snapshot of a White teacher educator's practice. *Pedagog. Cult. Soc.* 28, 17–36.
- Ohito, E.O., Deckman, S.L., 2018. Feeling Black and blue in preservice teacher education: encountering emotion and embodiment in antiracist teaching. In: Ahad-Legardy, B., Poon, O.A. (Eds.), *Difficult Subjects: Insights and Strategies for Teaching About Race, Sexuality, and Gender*. Stylus, Sterling, VA, pp. 129–144.
- Perez Huber, L., Johnson, R.N., Kohli, R., 2006. Naming racism: a conceptual look at internalized racism in US schools. *Chicana/o Latina/o Law Rev.* 26, 183–206.
- Phillip, T., 2019. Principled improvisation to support novice teacher learning. *Teach. Coll. Rec.* 121, 1–32.
- Phillip, T., Souto-Manning, M., Anderson, L., Horn, I., Carter Andrews, D., Stillman, J., Varghese, M., 2018. Making justice peripheral by constructing practice as “core”: how the increasing prominence of core practices challenges teacher education. *J. Teach. Educ.* 70, 251–264.
- Picower, B., 2021. *Reading, Writing and Racism: Disrupting Whiteness in Teacher Education and the Classroom*. Beacon Press, Boston.
- Picower, B., Kohli, R., 2017. Introduction. In: Picower, B., Kohli, R. (Eds.), *Confronting Racism in Teacher Education: Counternarratives of Critical Practice*. Routledge, New York/London, pp. 1–17.
- Richert, A.E., 1991. Case methods and teacher education: using cases to teach teacher reflection. In: Tabachnick, B.R., Zeichner, K.M. (Eds.), *Issues and Practices in Inquiry-Oriented Teacher Education*. The Falmer Press, London/New York, pp. 130–150.
- Rueda, R., Stillman, J., 2012. The 21st century teacher: a cultural perspective. *J. Teach. Educ.* 63 (4), 245–253.
- Shor, I., 1987. Educating the educators: a Freirean approach to the crisis in teacher education. In: Shor, I. (Ed.), *Freire for the Classroom: A Sourcebook for Liberatory Teaching*. Heinemann, Portsmouth, NH, pp. 7–32.
- Sleeter, C., 2001. Preparing teachers for diverse schools: research and the overwhelming presence of whiteness. *J. Teach. Educ.* 52, 94–106.
- Sleeter, C., 2017. Critical race theory and the whiteness of teacher education. *Urban Educ.* 52, 155–169.
- Souto-Manning, M., 2010. *Freire, Teaching and Learning: Culture Circles Across Contexts*. Peter Lang, New York.
- Souto-Manning, M., 2019. Transforming university-based teacher education: preparing asset-, equity-, and justice-oriented teachers within the contemporary political context. *Teach. Coll. Rec.* 121, 1–26.
- Souto-Manning, M., 2021. Righting the literacy education debt: a matter of justice. *J. Lit. Res.* 53 (4), 588–600.
- Souto-Manning, M., Emdin, C., 2020. On the Harm Inflicted by Urban Teacher Preparation Programs: Learning From the Historical Trauma Experienced by Teachers of Color. *Urban Education*. <https://doi.org/10.1177/0042085920926249>.
- Souto-Manning, M., Martell, J., 2019. Toward critically transformative possibilities: considering tensions and undoing inequities in the spatialization of teacher education. *Teach. Coll. Rec.* 121, 1–42.
- Souto-Manning, M., Stillman, J., 2020. In pursuit of transformative justice in the education of teacher educators. Introduction to the special issue: rethinking the preparation of teacher educators: centering equity and justice. *N. Educ.* 16 (1), 1–4.
- Stillman, J., Anderson, L., 2014. Commentary: preparing the next generation of teacher educators. *Teach. Coll. Rec.* (Accessed 27 June 2014).
- Stillman, J., Anderson, L., 2015. From accommodation to appropriation: teachers, identity, and authorship in a tightly coupled policy context. *Teach. Teach. Theor. Pract.* 21 (6), 720–744.
- Stillman, J., Anderson, L., 2016. Minding the mediation: examining one teacher educator's mediation of two preservice teachers' learning in context(s). *Urban Educ.* 51 (6), 659–682.
- Stillman, J., Beltramo, J.L., 2019. Exploring Freirean culture circles as a pedagogical space for preparing asset-oriented teacher educators. *Teach. Coll. Rec.* 121 (6), 1–38.
- Stillman, J., Beltramo, J., Ahmed, K.S., Castañeda-Flores, E., Garza, V.G., Pyo, M., 2019a. From the ground up: cultivating teacher educator knowledge from the situated knowledges of emerging, equity-minded teacher educators. *Asia Pac. J. Teach. Educ.* 47 (3), 265–285.
- Stillman, J., Ahmed, K.S., Castañeda-Flores, E., 2019b. On cocoons, consciousness and courage: preparing justice-oriented teacher educators through critical pedagogies. *Educ. Forum* 83 (4), 432–452.
- The Teachers of Color Collective & Souto-Manning, M., 2022. On the miseducation of teachers of color: a letter to teacher educators. *J. Teach. Educ.* 73, 66–80.
- U.S. Department of Education, 2016. *The State of Racial Diversity in the Educator Workforce*. U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, Washington, DC.
- Varghese, M., Daniels, J., Park, C., 2019. Structuring disruption within university based teacher education programs: possibilities and challenges of race-based caucuses. *Teach. Coll. Rec.* 121, 1–34.
- van Es, E., Sherin, M.G., 2010. The influence of video clubs on teachers' thinking and practice. *J. Math. Teach. Educ.* 13, 155–176.
- Vygotsky, L., 1978. *Mind in Society*. Harvard University Press, Cambridge, MA.
- Zeichner, K., 1999. The turn once again toward practice-based teacher education. *J. Teach. Educ.* 63, 376–382.
- Zeichner, K., 2005. Becoming a teacher educator: a personal perspective. *Teach. Teach. Educ.* 21, 117–124.
- Zeichner, K., 2010. Competition, economic rationalization, increased surveillance, and attacks on diversity: neo-liberalism and the transformation of teacher education in the U.S. *Teach. Teach. Educ.* 26, 1544–1552.
- Zeichner, K., 2014. The struggle for the soul of teaching and teacher education in the USA. *J. Educ. Teach.* 40, 551–568.
- Zeichner, K., Liston, D., 2013. Understanding reflective teaching. In: Zeichner, K.M., Liston, D.P. (Eds.), *Reflective Teaching: An Introduction*, second ed. Routledge, New York, pp. 1–8.
- Zeichner, K., Tabachnick, B.R., 1981. Are the effects of university teacher education “washed out” by school experience? *J. Teach. Educ.* 32, 7–11.
- Zeichner, K., Bowman, M., Guillén, L., Napolitan, K., 2016. Engaging and working in solidarity with local communities in preparing the teachers of their children. *J. Teach. Educ.* 67 (4), 277–290.

Improving teacher education through the self-study of practice

Julian Kitchen, Brock University, Hamilton, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	488
Self-study and enacting a pedagogy of teacher education	488
What is self-study?	489
Theoretical assumptions	489
Methodology	490
How does self-study improve teacher education practices? Four exemplars	490
Reconceptualizing teacher knowledge as tensions	491
Self-study, critical friendship, and the complexities of teacher education	491
From teacher to teacher educator	492
Providing students with written feedback on their assessment	493
Moving forward through self-study	494
References	494

Introduction

The self-study of practice has established itself as a robust approach to teacher education. The field was established in 1993 by teacher educators, who established the Self-Study of Teacher Education Practices (S-STEP) Special Interest Group of the American Educational Research Association (see [Loughran, 2004](#)). This practitioner research group, which initially borrowed theoretical conceptions and research methods from other fields, has over the past thirty years developed a rich literature on the practices of teacher educators working within their institutional contexts. S-STEP is the most common term for this field, as evidenced in the title of its leading journal, *Studying Teacher Education: A Journal of Self-Study of Teacher Education Practices*. At times, the term is extended to include teachers, as in the *International Handbook of Self-Study of Teaching and Teacher Education Practices* ([Loughran et al., 2004](#)) and the second edition ([Kitchen et al., 2020](#)). The term self-study of practice is employed, as in *Research Methods for the Study of Practice* ([Tidwell et al., 2009](#)), to widen the circle to include practitioners in higher education and the professions. This chapter, which builds on Tom [Russell's \(2010\)](#) overview of self-study in the third edition of this encyclopaedia, focusses primarily on how self-study aims to improve preservice teacher education practices.

Self-study and enacting a pedagogy of teacher education

In "Self-Study by Teacher Educators," [Russell \(2010\)](#) framed self-study as a response to the challenges of learning to teach. Building on fundamental problems identified by [Darling-Hammond \(2006\)](#), he wrote:

One broad conclusion that can be drawn from these researchers is that teacher education is far more complex than it has commonly been taken to be, if we consider the short period of professional preparation relative to other professions and if we consider the lack of explicit attention to the problems of apprenticeship of observation, enactment, and complexity. When self-study of teacher education practices arrived on the landscape of teacher education, it was with a strong sense that teacher educators themselves must look within.

[Russell \(2010, p. 690\).](#)

This was followed by brief profiles of four leading self-study practitioners who had attempted to develop, enact and research teacher education practices that were responsive to the complexities of teacher education and summaries of two significant documents on self-study methodology.

John [Loughran \(2006\)](#), one of the practitioners profiled by Russell, articulated the case for the development of a pedagogy of teacher education through which teacher educators understand teaching and learning about teaching. He argued that "students of teaching need to be able to see and hear the pedagogical reasoning that underpins the teaching they are experiencing" (p. 5). Teacher educators, he proposed, need to purposefully and explicitly examine the complex interplay among elements within courses and across a program in order to go beyond transmission of knowledge and modeling practice:

Unpacking teaching in ways that give students access to pedagogical reasoning, uncertainties and dilemmas of practice ... to make clear how the teaching approach purposefully encourages learning and how learning influences teaching in action. (p. 6)

Self-study, for Loughran (2005), is “a meaningful way for uncovering important facets of the knowledge of practice” as it helps teacher educators “capture, unpack and portray and the complexities of teaching and learning about teaching in ways that might lead to deeper understandings of practice” (p. 13). “The S-STEP community,” through its collaborations, conferences and publications, was regarded by Loughran (2004) as “an important audience” that supports, critiques, models and mentors (p. 29) teacher educators in enacting a pedagogy of teacher education. It also fosters “development in critical and useful ways so that the learning through self-study might be informative, accessible and useable for others” (Loughran, 2004, p. 32).

What is self-study?

The term *self-study of teacher education practices* is often abbreviated to *self-study*. This is understandable but it is wise to regularly remind oneself to complete the term. An alternative is to substitute the acronym S-STEP.

The first word, *self-study*, conveys that the *self* is engaged and directing in the *study*. Many institutions, such as universities and professional regulatory bodies, use the term “institutional self-study” to refer to comprehensive internal reviews aimed at improving programs (e.g., [Liaison Committee on Medical Education, 2017](#)). The abbreviated version of S-STEP is reasonably accurate in conveying that teacher educators engaged in this approach to practitioner inquiry are studying their practice in order to become more effective at their work. Indeed, they are both the researchers and the subjects of their studies.

The self, however, is more personally engaged in S-STEP than in institutional self-studies. Teacher educators are inquiring into their professional practices and, often, engaged in *reflection-in-action* (Schön, 1983). As [Martin and Russell \(2020\)](#) wrote, “self-study of teacher education practices is a metacognitive and reflective practice conducted by teacher educators learning from experience” (p. 1049). Undergirding much of self-study is the view that “teaching experience is acquired by investing time in the context of professional action; learning from experience demands reflection-in-action as an alternative frame of reference for personal learning” (p. 1059).

Many S-STEP researchers go beyond reflecting in and on action by also considering their personal and professional identities. They recognize, as [Connelly and Clandinin \(1988\)](#) articulated, that “‘personal knowledge’ ... determines all matters of significance relative to the planned conduct of classrooms” (p. 4). This leads many to engage, to varying degrees, in personal history, narrative and autoethnographical approaches to self-study ([Samaras, 2002](#)). As co/autoethnographers [Coia and Taylor \(2009\)](#) wrote, “Our pedagogical choices, our perceptions of the challenges we face, all involve our values, beliefs, and prior experiences” (p. 5). Coia and Taylor, in the next sentence, cautioned “clearly that is not all that is involved (p. 5). They then acknowledged their subjectivity and “the impact of the world on constructing the personal” (p. 5).

[Loughran \(2004\)](#), in his history of self-study, stressed the importance of collaboration in “the checking of data and interpretations” (p. 20) and “framing and reframing” (p. 21) understandings. As a result, collaboration and critical friendships have increasingly become means of demonstrating trustworthiness in self-study ([Schuck and Russell, 2005](#)).

Study conveys that S-STEP is a research method. Researchers collect and analyze information in order to more deeply understand themselves and their practices, students and contexts. Most practitioners are guided by an understanding of teacher knowledge as “knowledge that comes from experience ... learned in context, and ... “expressed in practice” ([Clandinin and Connelly, 2004](#), p. 579). [LaBoskey \(2004\)](#) identified “self-initiated and focused” as one of the predominant characteristics of self-study. Through self-study methods, teacher educators consider multiple dimensions as they “provoke, challenge and illuminate” on their paths to “hard won, elusive, and generally more tentative” ([Bullough and Pinnegar, 2001](#), p. 20) conclusions about their lived experiences in practice settings.

Teacher education is the context in S-STEP, although there are also self-studies of teaching, higher education and other professions ([Kitchen, 2020c](#)). Self-study, however, acknowledges that each local context is different and that the practitioner is well positioned to understand and adapt to contextual factors. More importantly, it helps teacher educators “maintain a focus on their teaching and their student’s learning” while also being “informative, accessible and useable for others” ([Loughran, 2004](#), p. 31).

As practitioner researchers, teacher educators in S-STEP are interested in identifying effective *practices* that prepare graduates to be successful teachers. As [Loughran \(2006\)](#) stated, there is a need to develop a pedagogy of teacher education. In enacting a pedagogy of teacher education, S-STEP teacher educators share with other practitioner researchers an interest in identifying specific practices that work ([Cochran-Smith and Lytle, 2004](#)). The difference is that self-study brings teacher educators to forefront alongside their practices, and calls on practitioners to be critical of their roles as practitioners and inquirers ([Feldman et al., 2004](#)).

S-STEP simultaneously considers all elements of the education included in the term. In practice, however, individual publications may foreground one element. For example, while [Kitchen’s \(2005a,b\)](#) “relational teacher education” is a conception that considers self, students, practice and context in relation to each other, individual articles and chapters often focus primarily on a single piece of the larger puzzle. His body of self-study work includes explorations of his personal and professional identity (e.g., [Kitchen, 2016](#)), methodology (e.g., [Kitchen, 2009](#)), research ethics ([Kitchen, 2019](#)), collaboration ([Kitchen et al., 2008](#)), student voices (e.g., [Kitchen, 2020a](#)), and specific practices (e.g., [Kitchen, 2008](#); [Kitchen and Bellini, 2012](#)).

Theoretical assumptions

Self-study emerged as responses to the technical-rational assumption that research knowledge can be applied to practical problems with little reference to people or context ([Schön, 1983](#)). [Russell \(2004\)](#) argued that “our growing awareness of the importance of *self*

in teacher education" was "not only necessary but also inevitable" (p. 1192) given the failure of the technical rational model to serve the needs of practitioners. This led many educational theorists and teacher educators to consider thinkers such as Dewey, Schwab and Schön who acknowledged the complexities of teaching and learning. Dewey (1938) regarded educative experiences as those that lead to growth through teachers who are responsive to "the situations in which interaction takes place" (p. 45). Emphasizing the centrality of practice, Schwab (1971) claimed that theories are "in the main, ill-fitted and inappropriate to problems of actual teaching and learning" (p. 287). Practical problems for Schwab "arise from states of affairs in relation to ourselves" (p. 189) and, thus, the "method of the practical ... is, then, not at all a linear affair proceeding step-by-step, but rather a complex, fluid, transactional discipline aimed at identification of the desirable and at either attainment of the desired or at alteration of desires. (p. 291). Schön's (1983) offered a new way of approaching the theory-practice by making the case that reflective practitioners in classrooms carry the "authority of experience" (Munby et al., 2001). Connelly and Clandinin's (1988) on the "personal practical knowledge" of teachers' and their authority curriculum-makers also resonated with the founders of S-STEP, as it reinforced the importance of knowing oneself in order to understand and effectively teach students in classrooms (Connelly and Clandinin, 1988).

These and related theories justify research that is initiated by teachers, teacher educators and other practitioners, focused on themselves and their practice (LaBoskey, 2004) in relation to their students (Bullough and Pinnegar, 2001), and grapples with the complexities of practice in real world contexts (Loughran, 2004).

Methodology

Self-study, according to Loughran (2004), began as "a 'coming together' of like-minded people with similar interests, issues and concerns" (p. 13) open to employing "a remarkable range of methods (p. 17) selected based on the "question/issue/concern under consideration" (p. 17). Self-study, according to Bullough and Pinnegar (2001), is a "mongrel" as "the study is always of practice, but at the intersection of self and other and its methods are borrowed" (p. 15). Today, despite a growing body of self-study research, "one of the defining features of S-STEP is its hybridity" (Fletcher, 2020, p. 270). For Fletcher, this hybridity is evident in the theoretical frames, analytical tools and methodologies (and methods) used. Common methods used in S-STEP include reflective practice; narrative inquiry, personal history and autoethnography; action/teacher/practitioner research; dialogic communities, portfolios; and arts-based inquiry. (Loughran, 2004; LaBoskey, 2004; Fletcher, 2020).

While hybridity and an interest in coming together are strengths, they exist in tension with calls for academic rigor. LaBoskey (2004), in "The Methodology of Self-Study and Its Theoretical Underpinnings," addresses the need for well-conceptualized and executed research design. She identifies five criteria for designing and assessing self-studies:

- *Self-initiated and focused*: teacher educator as "both who is doing the research and who is being studied (p. 842);
- *Improvement-aimed*: "evidence of progress and ... adjustments to our understanding of practice" (p. 859);
- *Interactive*: "interactions with our colleagues near and far" (p. 859), including students, critical friends and educational literature;
- *Multiple, primarily qualitative, methods*: "competent use ... of multiple methods ... to gain different, and thus more comprehensive perspectives" (pp. 859–860);
- *Exemplar-based validation*: "formalize our work and make it available to our professional community for deliberation, further testing, and judgment" (p. 860).

Similarly, in order to promote rigor, Bullough and Pinnegar (2001) identified guidelines for autobiographical writing in self-study, including that it "ring true," "promote insights and interpretation," "take an honest stand," and help "improve the learning situation not only for the self but for the other". (pp. 16–17).

The importance of methodological rigor in self-study today is nicely summarized in "Methods and Tools of Self-Study" by Tidwell and Jónsdóttir (2020):

Self-study is a methodology that offers a variety of research designs and methods that often challenge the known, and push researchers to think deeper about their own practice, their own context, and their understanding about the complex roles in and purposes of education. ... As the qualitative nature of self-study and its ingrained call for social justice and change often disrupts and deconstructs what is known, pushing further into the unknown, it opens the possibilities and need for more creative and innovative approaches to the study of practice ... It is critically important for self-study research to include the theoretically grounded research question(s), the data choices and collection process, and the data analyses used to make meaning of the data. It is this full reporting of the data process that enables us to make sense of what we know, and that enables us to confirm the efficacy of our research practice. (p. 418)

Self-study's hybridity and commitment to rigor provide both a firm foundation for future research and the flexibility to adapt to changes in the field of education.

How does self-study improve teacher education practices? Four exemplars

S-STEP as a discourse community engages in the collective development of knowledge. Given the focus of self-studies, the common purpose is to improve teacher education with a view to enhancing student learning in schools. Articles and chapters in self-study

play an important role in advancing this agenda. In this section, four articles published in *Studying Teacher Education* are considered as exemplars of how self-study improves teacher education.

Reconceptualizing teacher knowledge as tensions

Dewey, Schwab and Schön were among the theorists who inspired teacher educators to study their practices and make sense of the complexities of teaching and learning in teacher education and beyond. Amanda Berry, in an effort to improve her practice, grappled with “the ambiguities and complexities inherent in teacher educators’ work” (Berry, 2007a, p. 121). This led her to reconceptualize and reframed “tensions” as “an analytic tool for studying practice. Berry (2007b), in *Tensions in Teaching about Teaching: Understanding Practice as a Teacher Educator*, identified six tensions that together offer “a helpful frame for interpreting and organizing teacher educators’ accounts of their experience (2007a, p. 120). In “Reconceptualizing Teacher Educator Knowledge as Tensions: Exploring the Tension between Valuing and Reconstructing Experience,” one of the ten most cited in *Studying Teacher Education* (Kitchen and Berry, 2020), she focused on one of these tensions illustrate how this frame can serve to draw out the complexities of practice, The impetus for Berry’s self-study was “the dissonance created between what [she] sought to do in [her] sought to do in [her] teacher education practices and the means taken to achieve [he] goal” (Berry, 2007b, p. 118). Recognizing the limitations of her “pedagogy-of-presentation,” she sought to develop “a pedagogy of teacher education ... consistent with the view of learning [she]was presenting in teacher education, that high school students needed opportunities to construct their knowledge of biology through engaging in the process of meaning making for themselves” (p. 118). Berry’s self-study methodology for this inquiry was rigorous and hybrid. Her data sources included an autobiographical account of her experiences, videotaped lessons, journals, field notes, observations by colleagues, student assignments and correspondence with a teacher candidate.

The main significance of this study was the emergence of “tensions” as a conceptual frame and analytical tool. Berry wrote:

Analysis of the data developed alongside a review of the research published by other teacher educators also motivated to study their practice. I recognized similarities between my own experiences and what these teacher educators reported from their practice. I also saw a broader framework that connected these elements of teacher educators’ practice. It was apparent that teacher educators regularly experienced particular *tensions* as they attempted to manage complex and conflicting pedagogical and personal demands within their work as teachers of prospective teachers. The notion of tensions offered a useful way of describing teacher educators’ experience of their practice (including my own). It captured well the feelings of internal turmoil experienced by teacher educators as they found themselves pulled in different directions by competing pedagogical demands in their work and the difficulties they experienced as they learnt to recognize and manage these demands. (p. 119)

While the recognition of tensions is not unique to Berry, her framework and six illustrations make these tensions. This articulation, by both aiding recognition of tensions and offering insights into managing them, has aided many teacher educators in their research and practice (e.g., Kitchen and Bellini, 2012; Kitchen, 2020a).

The body of the article explores the tension between valuing “personal experience in learning to teach” and challenging those experiences “in ways that move [teacher candidates] beyond simply (re)confirming their existing beliefs about teaching and learning” (Berry, 2007b, p. 121) experience. Berry then drew on her data sources to explore how this tension played out in both “deliberate pedagogical structures” (p. 121) such as a personal learning review (teacher candidate responses to questions about their schooling in sciences), a drawing task, and peer microteaching. In parallel, she considered new approaches to reconstructing experience through “through an unplanned activity of sharing experiences of teaching with Lisa, a student in the class” (p. 121). The email correspondence with Lisa served as valuable for of interaction that provided Berry with insights into issues and validated thinking about her enactment of teacher education pedagogy (LaBoskey, 2004).

The article concluded with Berry considering how this research, particularly the idea of tensions, contributes to self-study and teacher education. She emphasized the importance of sharing both “phronesis” (personal perceptions) and “episteme” (findings) with a wider discourse community o as to improve teacher education pedagogy and programs. As tensions “serve both as a language for describing practice and as a frame for studying practice” (Berry, 2007b, p. 132), it is helpful in capturing and holding the complexities of practice without reducing or resolving them. This contribution makes “Reconceptualizing Teacher Educator Knowledge as Tensions” an exemplar of theory that illuminates teacher educator identity and improves self-study practice.

Self-study, critical friendship, and the complexities of teacher education

A second way in which self-study improves teacher education is by providing practitioner with methodologies that draw out their insights as teacher educators and make them accessible to others in the field. As LaBoskey (2004) noted, interactivity with others is an important research design characteristic of self-study. One means of collaborating with others is through critical friendships in which two or more teacher educators interact personally and professionally to make sense of individual or shared experiences as revealed in research data and analysis. “Self-Study, Critical Friendship, and the Complexities of Teacher Education” by Schuck and Russell (2005), the most highly cited article in *Studying Teacher Education* (Kitchen and Berry, 2020) exemplifies how a self-study article can improve research practice.

Sandy Schuck and Tom Russell “as teacher educators in Australia and Canada” joined together as critical friends out of a shared concern “with studying and improving our teaching” (p. 107). Tom “enlisted Sandy’s aid as a critical friend for one semester” (p.

117) to help him make sense of “an unexpected teaching assignment” (p. 117) in secondary school science classes. He hoped that Sandy’s insights from afar would help him make sense of the data and construct meaning. Soon after, Tom took the opportunity to work at Sandy’s university, thus affording Sandy the opportunity to engage Tom as a critical friend to observe and help make sense of her mathematics lessons.

The article, after a brief introduction, begins by identifying problematic issues and assumptions about critical friendship that emerged from their collaboration. They identify “the difficulty of assessing one’s own practice and reframing it” (p. 108) as a problematic issue. This means that it is vitally important that the critical friend be trusted, supportive, and willing to ask provocative questions. While both were willing to engage in this manner, implicit assumptions emerged that needed to be addressed. The first was that Sandy was unsure how to meet Tom’s expectations of a critical friend. The second was that Sandy had to confront her perception that she possessed lower academic status, even though these “issues, that were quite significant for Sandy, were distant or invisible for Tom” (p. 109).

The body of the article consists of data collection and analysis of the first phase involving Tom’s teaching, followed by a summary of outcomes and consideration of insights. A similar structure was then employed for the second phase on Sandy’s university teaching. These sections of the paper are rich in detail on their experiences, the data collection process, and the lessons learnt. Excerpts from their reflective journals bring the experiences and interactions to life. They “show” more than “tell” how a critical friendship is experienced. By identifying the problematic and making it front-and-centre, they help the reader grapple with the complexities of developing authentic critical relationships.

In offering insights into how their critical friendships might be improved, they wrote:

The issue of how best to act as a critical friend continues to intrigue us. Supportive and encouraging comments are more easily received than ones that challenge our thinking and beliefs. We both experienced this in the role of the practitioner, and we both wondered how best to challenge gently as the critical friend. (p. 117)

This led Sandy and Tom to construct meaning to be shared from their experiences over the two phases. They identified differences between face-to-face interactions and one’s by email from a distance. While email ensured a detailed record, the types of observations and interaction is limited by the control over the themes to be explored. They also contrasted teaching with critical friendship, noting similar commitments to relationship and growth, yet a need for the critical friend to be both sensitive and critical. Particularly significant for them was the “dichotomy between being told new insights and working out new insights for oneself” (p. 119). They ended with eleven conclusions about critical friendships within teacher education and self-study that are evident from data shared and, based on its citation record, helpful to teacher educators engaging in critical friendship.

Schuck and Russell’s account provide insights into the experience of being in a critical friendship. The conclusions offer valuable takeaways on how employ this methodology. The article is also a model of working through the complexities of teacher education through self-study, as it is clearly improvement-aimed, interactive, validated by exemplars, and varied in the qualitative methods used while remaining real and accessible.

From teacher to teacher educator

Teacher educators are drawn to S-STEP by their interest in improving their professional practice. Novice teacher educators experience a genuine need for guidance, especially as there is little formal preparation in developing and enacting a pedagogy of teacher education (Loughran, 2006). Half of the ten most cited articles in *Studying Teacher Education*, including the two-part “From Teacher to Teacher Educator” study by Todd Dinkelman et al. (2006a,b), relate to becoming a teacher educator (Kitchen and Berry, 2020).

These two articles report on “a study of two beginning teacher educators making the transition from classroom teacher to university-based teacher educator” (Dinkelman et al., 2006a, p. 5). The two novices, Jason and Karl, were supported by Todd, a university-based teacher educator and researcher. The first article, “From Teacher to Teacher Educator: Experience, Expectations, and Expatriation,” reported on their questions, framing within research literature, and methodologies, as well as results related to “their institutional context and shifting role identification” (p. 5). The second, “From Teacher to Teacher Educator: Reframing Knowledge in Practice” (Dinkelman et al., 2006b), built “on these insights by addressing the frames of understanding and knowledge employed in this transition and by describing how these frames informed decisions in the arena of teacher education practice” (2006a, p. 5). The began by identifying a dearth of research on the professional development of teacher educators and, in particular, the challenges for those transitioning from elementary and secondary school teaching. They identified S-STEP as “the most powerful impetus for the growing body of research into the development of teacher educator identity, competence and practice” (2006a, p. 7), before framing their work in relation to the teacher socialization literature. The authors identified case study as their primary methodology, with weekly reflective journals and semi-structured interviews being key data sources to be coded.

Three common themes are the focus of the study. First, as graduate students engaged in teacher preparation, they found the institutional context to be both supportive and challenging. It was supportive in that it offered time for reflection and, through Todd’s mentoring, support in making the transition. On the other hand, they felt challenged by “an environment that simultaneously valued and devalued their work” (2006a, p. 14) and lacked education and professional guidance on the day-to-day practice of teacher educators.

Second, they were challenged by their shifting role identification. While their classroom experience formed a strong foundation for their university work, the two identities did not always correspond. Karl reflected in an interview:

Becoming a teacher educator is more complicated ... I'm trying to think at the same time about whether or not I'm practicing what I'm preaching ... So the process, this becoming process is a lot more complicated for me, than the process of becoming a teacher was ... They are also complicated by the fact I'm doing many things at the same time. (2006a, p. 20)

They were discovering, as Loughran (2006) was conceptualizing at the time, that teachers modeling practice was not enough, that teacher education required a different identity and a different pedagogical stance.

Third, as outlined in the second part, Karl and Jason needed to reframe their understanding and knowledge in order to successfully transition to teacher education. This extended section is broken into nine subsections, including *street credibility and survival strategies* (e.g., drawing on prior experiences as teachers and using these to demonstrate credibility); a shift from *learning by doing* to *learning by thinking* thanks to reflection on their teacher education practices: *reframing puzzles* through their journaling and meetings as part of this study; a dearth of formal opportunities to learn how to become a teacher educator in graduate programs; and the importance of *modeling* as teacher educators not just as teachers. Specific examples from journals and interviews demonstrate Jason and Karl's growth and give texture and nuance.

The second article ends with an extended discussion section that considers several important themes. The first is an exploration of the need to better prepare novice teacher educators through graduate courses that help them develop a pedagogy of teacher education. This is followed by consideration of how to most effectively draw on the sources of knowledge derived from teaching in schools. Finally, the authors reflect on the shifting of professional identity from that of teacher to teacher educator: "What is at stake if teacher educators identify with one world or the other?" (p. 133).

These articles are particularly important as they offer insights, supported by artifacts of experience, into how to transition from teacher to teacher educator. The article concludes by reminding readers that the "initial experience of doing teacher education is a powerful force in shaping the professional practice of teacher educators over their careers" (p. 135). Thus; research is vitally important as a means of understanding novice teacher educators and identifying effective approaches to helping them develop and enact a pedagogy of teacher education.

Providing students with written feedback on their assessment

While many S-STEP articles focus on the teacher educator identity and presentation in teacher preparation programs, other studies begin with a problem of practice—for example, feedback on reflective writing (Kitchen, 2008) or teaching about queer issues (Kitchen and Bellini, 2012)—or a strategy for more effectively delivering content or skills, such as exit cards (Kitchen et al., 2020). A good example of this is "Providing Students with Written Feedback on Their Assessment: A Collaborative Self-Study Exploring the Nexus of Research into Practice" by Sharon Pittaway and Tony Dowden (2014).

The authors begin by identifying "dissatisfaction with written feedback on their assessment" as "a perennial concern in university learning contexts" (p. 197). The article that follows addresses improving the quality of written feedback through the practice of Pittaway and Dowden over five years. Whereas most such studies are modest and local, this study took place in two institutions, over five years, involved the collection of demographic data, and was supported by a teaching development grant. What makes this a self-study is that the authors, in addition to drawing on data, engaged in reflective writing on their respective pedagogical practices and documented in detail their professional conversations over those years. Through these processes, Sharon and Tony explored their "beliefs and practices and the ways these changed and developed as a result of [their] professional conversations" (p. 201).

The analysis of their growth that each provides is rich in complexity and offers insight into the value of studying one's practice. Sharon's account speaks to the value of shifting through self-study from "being the teacher" to "being the learner" (p. 203). This led her to be "more specifically aware of [student] experiences and perceptions or, put differently, I took more notice of their voices" (p. 202). This, among other learnings, inspired her to expand her feedback vocabulary, craft feedback of greater clarity and precision, and led her to structure her course so that "there is ample opportunity to provide [formative] feedback on technical aspects of writing during weekly discussion forums" (p. 204), as well as general feedback posts, and specific responses to individual students. Tony too grew substantially through reflection and professional conversations. The "revelation" for Tony was that he became "definitely more attuned to the whole notion and purpose of feedback and feed-forward" (p. 205). He wrote in his critical reflection that "There has been a sea-change in my practice as I've realized this old-school approach not only doesn't wash; it actually does more harm than good" (p. 205). Self-study led him to the revelation that "full and accurate critique" is not sufficient, that "emotion and cognition mediate each other and, therefore, it is [his] professional responsibility to provide feedback that students will find acceptable, helpful and supportive" (p. 206).

This article is notable as an exemplar of self-study in at least three respects. First, it highlights the value of deeply examining a particular practice. Second, it demonstrates the value of conducting self-study in parallel with other research approaches, as the authors published related studies in assessment journals. Third, in the spirit of self-study as a collaborative venture, through the "collaborative and occasionally confrontational nature" of their research, Pittaway and Dowden were "better able to grapple with difficult realities in [their] pedagogical practices" and enhance "the student learning experience" through sharing "with trusted colleagues" (p. 207).

Moving forward through self-study

Self-study of teacher education practices is an approach to improvement of practice that simultaneously studies practices in the teacher education classroom and the teacher educators role in them. Self-study, which is initiated by practitioners in order to better understand themselves and their practice, is improvement-aimed, interactive, uses multiple qualitative methods, and validated through exemplars. S-STEP typically adopts an ontological stance that places experience above theory as practitioners puzzle over practice in context, often with the support of critical friends and collaborators. As conveyed by the four exemplars above, self-studies improve teacher education by deepening theoretical understanding, offering hybrid methodologies for studying practice, critically and reflectively storying the experiences of teacher educators, and/or closely examining the enactment of pedagogies and practices.

Moving forward, what more can self-study do to improve teacher education and contribute to enhanced practice in other contexts?

The work of the self-study community has validated and enriched the individual and collective work of teacher educators. Zeichner (2007) credited S-STEP for adding “dignity to the important yet universally undervalued work of those faculty and staff who educate a nation’s teachers” and contributing to “many improvements in the quality of teacher education programs” (p. 37). Yet Zeichner and others have challenged S-STEP to offer more generalizable and verifiable findings that can be replicated and implemented on a large scale to improve practice overall. Zeichner and Conklin (2005) made a case for research that helps identify program structures that are particularly effective in preparing teachers, especially in the face of politically-driven challenges to conventional programs. Similarly, from within the S-STEP community, Korthagen and Lunenburg (2004) suggested that self-study needs to build from improving one’s own practice to helping improve teacher education within one’s institution, and to offering collective gains within a wider community of teacher educators. Clift (2004), drawing on exemplars of program-level self-studies, made a strong case for self-studies that go beyond course-based inquiry to improving key program elements and “continuous self-study with an entire program” (p. 1360). Kitchen (2020b), in the second *International Handbook of Self-Study of Teaching and Teacher Education Practices*, argued for more self-studies of exemplary teacher education programs that feature shared vision, faculty members committed to the shared vision, academic rigor, and integrated approaches to learning across all university and field-based elements of the program. From them teacher educators can learn how to do multiple things well simultaneously, not just attending to one element at a time as can happen in individual self-studies. Petrarca and van Nuland (2020), in the same volume, make a strong case for more self-studies on the complexities of practice teaching and finding ways to make meaningful connections across program elements. Similarly, O’Connor et al. (2020) make a case for more self-studies employing teaching portfolios as integrative assessment tools to be used across courses. Also, it is important to consider teacher education leadership through a self-study lens. Allison and Ramirez (2020) propose that self-studies of leadership can “promote positive change within teacher education, the academy, and beyond to the work of teachers and schools in society” (p. 1220). Taken together, these handbook chapters offer constructive suggestions for programs of research within the S-STEP community.

It is also important to “look beyond teacher education to possibilities for self-study in schools and higher education” as “self-study has much to offer teachers, college and university instructors, and other professions” (Kitchen, 2020c, p. 1025). While self-study has always been concerned with teaching, there has been little growth in published self-studies by elementary and secondary educators. Whitehead et al. (2020), advocate for self-study as an approach to inquiry in schools by experienced school-based educators. Similarly, Samaras and Pithouse-Morgan (2020) have worked to extend the reach of self-study into higher education and beyond. They argue that transdisciplinary self-study “offers a broader inclusiveness to practitioners from multiple fields of professional expertise and from diverse disciplinary backgrounds” (p. 1294). Hopefully, the self-study community can reach out more to practitioners in schools, institutions of higher education and professions, who could benefit from self-study’s theoretical frames, methodologies and exemplars of practitioners researching themselves and their practices. Although self-study is still largely an English-speaking world phenomenon, there is considerable room for self-studies in other languages, from different cultural perspectives and from around the world (Thomas and Guðjónsdóttir, 2020). Interest across cultures and languages is evident in a volume of seminal self-study translated into Spanish (Russell et al., 2016) and an entire section of the second *International Handbook of Self-Study of Teaching and Teacher Education Practices* devoted to self-study across languages and cultures (Kitchen et al., 2020).

In conclusion, the self-study of teacher education practices has come a long way from its inception in the early 1990s and since the publication of the first *International Handbook of Self-Study of Teaching and Teacher Education Practices* in 2004. The ongoing work of S-STEP as a community of practice and scholarship dedicated to studying and improving teacher education practices is worthy of acknowledgment. Its development as a methodology and its considerable body of published work merits celebration. At the same time, we need to take stock of the ways in which we as a self-study community can work toward improving education practice and programs at all levels—teaching, teacher education, community-based, and higher education.

References

- Allison, V., Ramirez, L., 2020. Using self-study to navigate roles and responsibilities as teacher educator administrators. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1199–1223.

- Berry, A., 2007a. Reconceptualizing teacher educator knowledge as tensions: exploring the tension between valuing and reconstructing experience. *Stud. Teach. Educ.* 3 (2), 117–134. <https://doi.org/10.1080/17425960701656510>.
- Berry, A., 2007b. Tensions in Teaching about Teaching: Understanding Practice as a Teacher Educator. Springer, Dordrecht.
- Bullough Jr., R.V., Pinnegar, S., 2001. Guidelines for quality in autobiographical forms of self-study research. *Educ. Res.* 30 (3), 13–21.
- Clandinin, D.J., Connelly, F.M., 2004. Knowledge, narrative and self-study. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practice*. Springer, Dordrecht, pp. 575–600.
- Clift, R.T., 2004. Self-study research in the context of teacher education programs. In: Loughran, J.J., Hamilton, M.L., LaBoskey, V.K., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Springer, Dordrecht, pp. 1333–1366.
- Cochran-Smith, M., Lytle, S.L., 2004. Practitioner inquiry, knowledge, and university culture. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, pp. 601–649.
- Coia, L., Taylor, M., 2009. Co/autoethnography: exploring our teaching selves collaboratively. In: Tidwell, D., Heston, M., Fitzgerald, L. (Eds.), *Research Methods for the Self-Study of Practice*. Springer, Dordrecht, The Netherlands, pp. 3–16.
- Connelly, F.M., Clandinin, D.J., 1988. *Teachers as Curriculum Planners: Narratives of Experience*. OISE Press, Toronto.
- Darling-Hammond, L., 2006. *Powerful Teacher Education: Lessons from Exemplary Programs*. Jossey-Bass, San Francisco, CA.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York.
- Dinkelman, T., Margolis, J., Sikkenga, K., 2006a. From teacher to teacher educator: experience, expectations, and expatriation. *Stud. Teach. Educ.* 2 (1), 197–209. <https://doi.org/10.1080/17425960600557447>.
- Dinkelman, T., Margolis, J., Sikkenga, K., 2006b. From teacher to teacher educator: reframing knowledge in practice. *Stud. Teach. Educ.* 2 (2), 197–209. <https://doi.org/10.1080/17425960600983155>.
- Feldman, A., Pugh, P., Mills, G., 2004. Self-study through action research. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, pp. 943–977.
- Fletcher, T., 2020. Self-study as hybrid methodology. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 269–297.
- Kitchen, J., Bellini, C., 2012. Making it better for lesbian, gay, bisexual and transgender students through teacher education: a collaborative self-study. *Stud. Teach. Educ.* 8 (3), 209–225. <https://doi.org/10.1080/17425964.2012.719129>.
- Kitchen, J., Berry, A., 2020. Maximizing the impact of your article in *Studying Teacher Education*. *Stud. Teach. Educ.* 16 (1), 1–5. <https://doi.org/10.1080/17425964.2020.1720088>.
- Kitchen, J., Ciuffetelli Parker, D., Gallagher, T., 2008. Authentic conversation as faculty development: establishing a self-study group in an education college. *Stud. Teach. Educ.* 4 (2), 157–171. <https://doi.org/10.1080/17425960802433637>.
- Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), 2020. *International Handbook of Self-Study of Teaching and Teacher Education*, second ed. Practices. Springer, Rotterdam, The Netherlands.
- Kitchen, J., 2005a. Looking backwards, moving forward: understanding my narrative as a teacher educator. *Stud. Teach. Educ.* 1 (1), 17–30. <https://doi.org/10.1080/17425960500039835>.
- Kitchen, J., 2005b. Conveying respect and empathy: becoming a relational teacher educator. *Stud. Teach. Educ.* 1 (2), 194–207. <https://doi.org/10.1080/17425960500288374>.
- Kitchen, J., 2008. Using written feedback to promote critical reflection: a teacher educator responds to reflective writing by preservice teachers. *Excelsior* 2 (2), 37–46.
- Kitchen, J., 2009. Passages: improving teacher education through narrative self-study. In: Tidwell, D., Heston, M., Fitzgerald, L. (Eds.), *Research Methods for the Self-Study of Practice*. Springer, Dordrecht, The Netherlands, pp. 35–51.
- Kitchen, J., 2016. Inside out: my identity as a queer teacher educator. In: Kitchen, J., Tidwell, D. (Eds.), *Self-Study and Diversity II. Sense*, Rotterdam, The Netherlands, pp. 11–26.
- Kitchen, J., 2019. Ethical issues in reporting on teacher candidate perspectives in a cultural diversity course: increasing trustworthiness, improving practice, and protecting participants. In: Brandenburg, R., McDonough, S. (Eds.), *Ethics, Self-Study Research, Methodology and Teacher Education*. Springer Nature, Singapore, pp. 97–115.
- Kitchen, J., 2020a. Attending to the concerns of teacher candidates in a social justice course: a self-study of a teacher educator. *Stud. Teach. Educ.* 16 (1), 6–25. <https://doi.org/10.1080/17425964.2019.1691134>.
- Kitchen, J., 2020b. Pedagogy of teacher education in exemplary programs. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1253–1289.
- Kitchen, J., 2020c. Self-study in teacher education and beyond. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1023–1044.
- Korthagen, F., Lunenberg, M., 2004. Links between self-study and teacher education reform. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, pp. 421–449.
- LaBoskey, V., 2004. Moving the study of self-study research and practice forward: challenges and opportunities. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, pp. 817–869.
- Liaison Committee on Medical Education, 2017. *Guide to the Institutional Self-Study*. Author, Washington, DC. https://www.upstate.edu/ume/pdf/cme/guide_to_institutional_self_study.pdf.
- Loughran, J.J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), 2004. *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht.
- Loughran, J., 2004. A history and context of self-study of teaching and teacher education practices. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dordrecht, pp. 7–39.
- Loughran, J., 2005. Researching teaching about teaching: self-study of teacher education practices. *Stud. Teach. Educ.* 1 (2), 5–16. <https://doi.org/10.1080/17425960500039777>.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education: Understanding Teaching and Learning About Teaching*. Routledge, London & New York.
- Martin, A., Russell, T., 2020. Advancing the epistemology of practice for research in self-study of teacher education. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1045–1073.
- Munby, H., Russell, T., Martin, A.K., 2001. Teachers' knowledge and how it works. In: Richardson, V. (Ed.), *Handbook of Research on Teaching*, fourth ed. American Educational Research Association, Washington, DC, pp. 877–904.
- O'Connor, K., Sterenberg, G., Vaughan, N., 2020. The portfolio and self-study: an epistemology of practice. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1135–1157.
- Petrarca, D., Van Nuland, S.V., 2020. Initial teacher education practicum 2.0: learning through self-study in a professional setting. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1103–1133.
- Pittaway, S., Dowden, T., 2014. Providing students with written feedback on their assessment: a collaborative self-study exploring the nexus of research into practice. *Stud. Teach. Educ.* 10 (3), 197–209. <https://doi.org/10.1080/17425964.2014.949659>.
- (Compilers) Russell, T., Fuentealba, R., Hirmas, C., 2016. *Formadores de formadores, descubriendo la propia voz a través del self-study [Teacher Educators, Discovering Their Own Voice Through Self-Study]*. Organización de Estados Iberoamericanos, Santiago, Chile. <https://oei.cl/en/Oei/News/formadores-de-formadores-descubriendo-la-propia-voz>.

- Russell, T., 2004. Tracing the development of self-study in teacher education research and practice. In: Loughran, J., Hamilton, M.L., LaBoskey, V., Russell, T. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Kluwer, Dodrecht, pp. 1190–1210.
- Russell, T., 2010. Self-study by teacher educators. In: *International Encyclopedia of Education*, third ed. Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00677-1>.
- Samaras, A.P., Pithouse-Morgan, 2020. Polyvocal self-study in transdisciplinary higher education communities. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1291–1322.
- Samaras, A.P., 2002. *Self-Study for Teacher Educators: Crafting a Pedagogy for Educational Change*. Peter Lang, New York.
- Schön, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York, NY.
- Schuck, S., Russell, T., 2005. Self-study, critical friendship, and the complexities of teacher education. *Stud. Teach. Educ.* 1 (2), 107–121. <https://doi.org/10.1080/17425960500288291>.
- Schwab, J.J., 1971, 1978. The practical: a language for curriculum. In: Westbury, I., Wilkof, N.J. (Eds.), *Science, Curriculum, and Liberal Education: Selected Essays*. University of Chicago Press, Chicago, pp. 287–321.
- Thomas, L., Guðjónsdóttir, H., 2020. Self-study across languages and cultures. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1325–1337.
- Tidwell, D.L., Jónsdóttir, S.R., 2020. Methods and tools of self-study. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 377–426.
- Tidwell, D., Heston, M., Fitzgerald, L. (Eds.), 2009. *Research Methods for the Self-Study of Practice*. Springer, Dordrecht, The Netherlands.
- Whitehead, J., DeLong, J., Huxtable, M., Campbell, L., Griffin, C., Mounter, J., 2020. Self-study in elementary and secondary teaching: a living theory approach. In: Kitchen, J., Berry, A., Bullock, S., Crowe, A., Taylor, M., Guðjónsdóttir, H., Thomas, L. (Eds.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*, second ed. Springer, Rotterdam, The Netherlands, pp. 1253–1289.
- Zeichner, K.M., Conklin, H.G., 2005. Teacher education programs. In: Cochran-Smith, M., Zeichner, K.M. (Eds.), *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education*. Erlbaum, Mahwah, NJ, pp. 645–736.
- Zeichner, K.M., 2007. Accumulating knowledge across self-studies in teacher education. *J. Teach. Educ.* 58 (1), 36–46. <https://doi.org/10.1177/0022487106296219>.

Preparing novice teacher educators for critical, justice-oriented teacher education practice

Hilary G. Conklin, DePaul University, College of Education, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Becoming a teacher educator: a personal account	497
Existing forms of learning and development for teacher educators	498
Self-study and collaborative inquiry	498
Formal study	499
Varying emphases of professional learning for teacher educators	499
A case of formal preparation for novice teacher educators focused on critical, justice-oriented aims	500
Course background and focus	500
Studying novice teacher educators' learning and practice	501
Learning conceptual and practical tools	501
Unplanned course learning	501
Implications of this case	501
Conclusion	502
References	502

In recent decades, the body of research on how to prepare teachers has continued to deepen, expand, and evolve, particularly in relation to approaches that can support teachers to enact humanizing, culturally sustaining, justice-oriented pedagogy (cf., [Carter Andrews, 2021](#); [Paris, 2012](#)). Yet while this body of literature on the preparation of teachers has proliferated, an essential companion to this research literature has remained persistently understudied: how do we prepare the people who prepare teachers? That is, how do we prepare teacher educators? And in particular, given persistent societal injustices, structural inequities, and systemic racism, how do we prepare teacher educators who will pursue transformative justice in the education of teachers ([Souto-Manning and Stillman, 2020](#))? In this article, I consider some of the research that sheds light on these questions. I begin by offering a personal illustration as a way of setting the stage for the scholarship in this domain.

Becoming a teacher educator: a personal account

When I reflect on my own journey of learning to become a teacher educator, a variety of sites, sources, and emphases stand out as contributing to my learning along the way. The first moment that I recognized I was becoming a teacher educator began more than 20 years ago during my time as a classroom teacher: a new colleague joined my team of seventh grade social studies teachers, and I began sharing my nascent social studies teaching practices with her in an effort to support her developing practice. That early professional experience of beginning to articulate my work as a teacher in order to foster my colleague's learning marked an initial stage of my journey as an educator of educators. Further, the experience inspired me to pursue doctoral studies at a university—focused on teacher education—where I rapidly expanded my opportunities to learn and grow as a teacher educator.

As part of my doctoral program, I enrolled in doctoral courses that helped me understand the history and politics of teacher education, practices of field-based teacher education supervisors, and approaches to action research in teacher education. As I pursued this coursework, I took part in a scholarly analysis of research about teacher education programs and practices under the guidance of an experienced and deeply knowledgeable mentor ([Zeichner and Conklin, 2005](#)). Alongside these experiences, I also practiced as a teacher educator, supervising teaching candidates as they learned to teach in a local middle school and leading a site-based weekly seminar for teaching candidates with a local teacher. At the same time, as a doctoral student, I conducted my own reflective action research on my emerging teacher education practice, collecting and analyzing artifacts of my practice as a way to deepen my learning and development as a field supervisor.

That intensive learning period launched me into the university-based work I do now as a teacher educator, teaching novice social studies teachers to develop their craft. Over more than 15 years in this role as a faculty member, I have continued my learning in various ways: through reading research on various aspects of teacher education, writing about dilemmas I have encountered and visions of practice I espouse ([Conklin, 2008](#)), participating in professional conferences, deliberately reflecting on my ever-evolving practice, partnering with a colleague and intentionally studying our own teacher education practices ([Conklin and Hughes, 2016](#)), and teaching a formal course to novice teacher educators ([Conklin, 2015, 2020](#)). Through all of these experiences, my work has been rooted in an overarching aim of creating more equitable, just, and humane learning spaces for young people in schools—a vision that has continually evolved, deepened, and clarified over time.

My intent in sharing the story of my own path is to highlight the many potential formal and informal contexts in which teacher educators may learn to do their work (cf., Murray, 2017), the varied forms of learning that may spur teacher educators' growth and development (cf., Hadar and Brody, 2017; Pinnegar and Hamilton, 2017), and the aims around which their learning may be centered. Yet while my personal story showcases many ways in which teacher educators might be prepared for and continually learn about their practice, it is also a very unusual story. A nagging theme in the research literature is that most teacher educators have very little preparation to help them learn how to become skilled teacher educators and very few intentional opportunities to further develop their practice (Goodwin and Kosnik, 2013; Goodwin et al., 2014; Ping et al., 2018; Stillman and Beltrano, 2019; Zeichner, 2005). And while there is limited research focused on the preparation of teacher educators in general, even less research exists that contributes to our understanding of how to prepare teacher educators who will challenge educational inequities, foster asset-oriented teaching practices, and disrupt the status quo (Stillman et al., 2019; Stillman and Beltrano, 2019).

In this chapter, I consider the anchoring question: How should we prepare novice teacher educators to foster critical, justice-oriented practice among teachers? I begin by offering an overview of research that has explored the learning and development of teacher educators. With this broader research landscape as a backdrop, I then draw from my own research to illustrate one approach to preparing novice teacher educators that highlights a subset of the dimensions of the learning and development of novice teacher educators. Specifically, I consider the affordances of a formal course of study in a university setting for novice teacher educators that focuses on pedagogical practices for critical, justice-oriented goals.

Existing forms of learning and development for teacher educators

While a wide range of people find themselves in the position of being teacher educators, very few have specific preparation for the role. Instead, research has confirmed that most teacher educators learn about the complex work of doing teacher education on the job (cf., Goodwin et al., 2014; Ping et al., 2018). Because many teacher educators have been classroom teachers at the primary or secondary level, there is often an assumption that the knowledge and skills they have acquired from being a classroom teacher will readily translate into becoming proficient teacher educators (Goodwin et al., 2014; Korthagen et al., 2005; Zeichner, 2005). Yet the existing scholarship on teacher educator learning and development offers a compelling case that the work of teaching children and teaching teachers—while similar in some ways—has important differences (cf., Bullock and Christou, 2009; Zeichner, 2005). For example, it's important for teacher educators to view teachers, themselves, as learners (Conklin, 2020; Stillman and Beltramo, 2019), and teacher educators also need to interrogate their own personal biographies and beliefs about issues like culture, race, class, and language in order to facilitate such learning among teachers (Cochran-Smith et al., 2020). Without opportunities for intentional training, most novice teacher educators are, at best, under-prepared for their roles—and at worst, they are in positions to reinforce problematic patterns of discrimination and inequity (Cochran-Smith et al., 2020).

Self-study and collaborative inquiry

While a lack of formal preparation for teacher educators is the norm, some teacher educators take it upon themselves to gain a better understanding of the important work they do. Thus, the primary way that teacher educator learning and development has occurred has been through self-initiated, individual or collaborative study of practice (Cochran-Smith et al., 2020; Hadar and Brody, 2017). Such examination of teacher educator practice has substantially increased in the last two decades and has led to a strong—though still small relative to the vast numbers of teacher educators—international community of teacher educators invested in their own and others' learning and development. The journal *Studying Teacher Education*, for example, was launched in 2005 to create a scholarly forum for sharing work focused on topics including self-study of teacher education practices, the transition from teacher to teacher educator, and the professional learning and development of teacher educators. This journal is a publication of the American Educational Research Association's Special Interest Group "Self-Study of Teacher Education Practice" (S-STEP), a community of educators dedicated to making contributions and improvements to theory and practice of teacher education. Scholarly communities like these have played a critical role in bringing attention to and advancing the learning and development of teacher educators.

Research in this area reveals that many teacher educators have fostered their own learning about teacher educating by closely examining and reflecting upon their practice as teacher educators (cf., Berry, 2009; Ritter, 2010). This robust body of work has shed light on how the practice of self-study offers an important pathway for teacher educators to surface their tacit understandings of teaching so that they can make those assumptions more explicit as they work with novice teachers (cf., Berry, 2009; Bullock, 2009; Loughran, 2007; Ritter, 2010). At the same time, self-study may also reveal the limits of one's knowledge as a teacher educator and highlight the complexity, challenge, and uncertainty of the work (e.g., Ritter, 2007). Further, such "self-education" may not result in the desired shifts in teachers' learning that teacher educators aim for, particularly with regard to issues of equity and justice (Cochran-Smith et al., 2020; Cochran-Smith and Villegas, 2016).

A closely related approach to the work of self-study is teacher educator development that occurs through collaboration among teacher educators. This research literature has demonstrated the value in collaborative inquiry into practice for teacher educators' development. Research has showcased teacher educators' learning through the formation of teacher educator inquiry groups (cf., Cochran-Smith, 2003; Hadar and Brody, 2013; Jacobs et al., 2015; Koster et al., 2008), faculty learning communities (Dover et al., 2020), collaboration with a critical friend or fellow teacher educator (cf., Conklin and Hughes, 2016; Peercy and Troyan, 2017; Vogel and Bartlett, 2013; Wolkenhauer and Hooser, 2021), or through participation in distinctive forms of inquiry like

Freirean culture circles and Boalian theater (Stillman and Beltramo, 2019; Stillman et al., 2019). Not only can these mutually vulnerable spaces foster important learning, this body of scholarship on collaborative inquiry has demonstrated the power of teacher educators' collective examination of problems of practice: when colleagues conduct this work together, they may be empowered to enact changes in the local contexts in which they work (Jacobs et al., 2015; Stillman et al., 2019).

Taken together, the learning that these different forms of inquiry into teacher educators' practice can generate is powerful. And importantly, the work in this domain represents teacher educators at many stages of professional learning: in some cases, it is novice teacher educators who take up such study (e.g., Ritter, 2007; Wolkenhauer and Hooser, 2021), while at other times, it is already experienced teacher educators who seek to continually improve their craft (e.g., Koster et al., 2008). While this body of work and the practices it showcases are vital to the field, such learning opportunities are not available to—or taken up by—all teacher educators: typically, this inquiry happens informally and is motivated by those teacher educators who take the work upon themselves or seek out like-minded colleagues for collaboration.

Formal study

Although some researchers have made compelling arguments that the primary content of teacher educators' learning and practice should focus on community-based knowledge derived from the local context (Stillman et al., 2019; Stillman and Beltramo, 2019), other scholars have argued that learning to become teacher educators should include not only inquiry into one's own practice in a particular setting but also formal preparation through coursework or other organized systems (Goodwin et al., 2014; Karagiorgi and Nicolaidou, 2013; Zeichner, 2005). And indeed, there is evidence of the need for the need for formal preparation for teacher educators. For example, Goodwin et al.'s (2014) research surfaced the finding that teacher educators had little preparation for enacting particular teacher education pedagogies, supporting the need for preparation in both the research on instructional strategies as well as intentional mentoring and support focused on pedagogies of teacher education (Goodwin et al., 2014; Lunenberg, 2002; Zeichner, 2005).

In various countries around the world, there have been some efforts to provide systematic education for teacher educators, such as in Israel, through the MOFET Institute (Cochran-Smith et al., 2020; Hadar and Brody, 2017), as well as in the Netherlands (Koster et al., 2008). In the United States, some doctoral programs focus specifically on teacher education (like the one in which I enrolled), while in Norway, teacher educators are encouraged to pursue doctorates, although the emphasis of these degrees is primarily on research (Cochran-Smith et al., 2020). Yet most researchers who study the education of teacher educators agree that none of these existing systems is sufficient; the existing programs either reach a very limited number of teacher educators or do not include the full range of professional learning that is necessary to become thoughtful, skilled, and justice-oriented teacher educators (Cochran-Smith et al., 2020).

Advocates of more formalized teacher educator learning and preparation have suggested specific types of understandings that teacher educators should hold for their work. These include knowledge about teacher education policy debates, the history of teacher education, research on teacher mentoring and teacher education pedagogies, as well as training and experience in particular pedagogies of teacher education (Cochran-Smith et al., 2020; Goodwin et al., 2014; Lunenberg, 2002; Zeichner, 2005). Further, scholars have suggested that preparing teacher educators should involve helping them gain in-depth understandings of human learning and development situated in complex socio-cultural contexts; expertise in diversity, cultural relevance, and social justice; contextual, socio-political, and sociological knowledge; and research capacity rooted in an inquiry stance (Cochran-Smith et al., 2020; Goodwin and Kosnik, 2013). Such scholarship has also included the recommendation for teacher educators to have ongoing opportunities to examine their own assumptions and biases related to issues like inequality, culture, race, sexual orientation, and language and how such biases relate to policies and practices in schools (Cochran-Smith et al., 2020). Finally, some scholars have focused on the need for teacher educators to hold a specialized form of expertise that matches the complexity of teaching; such understandings would enable teacher educators to nurture novice teachers' thoughtful decision-making within a vision of teaching that emphasizes its complexity (Goodwin and Kosnik, 2013; Loughran and Menter, 2019).

Varying emphases of professional learning for teacher educators

The calls for particular formal areas of study in the preparation and development of teacher educators point to the many different emphases of the various forms of professional learning reported in the research literature. Some scholarship, for example, has highlighted the process of becoming teacher educators in particular content areas, such as learning to become mathematics teacher educators (cf., Yow et al., 2016) or developing the pedagogical content knowledge needed to be successful science teacher educators (cf., Abell et al., 2009). In other cases, researchers have focused on learning to use particular pedagogical approaches as teacher educators, such as documenting the complexity of learning to use core practices in the context of preparing novice teachers of English Language Learners (Peercy and Troyan, 2017).

Other scholars have turned their attention to the broader aims of the enterprise of preparing teacher educators. In particular, a variety of scholarship in more recent years has taken up the question of how to prepare teacher educators who will work toward justice and transformative goals—an essential emphasis, given the vast, systemic inequities in our current schools. Stillman and Beltramo (2019), for example, documented how their group of teacher educators used Freirean culture circles and Boalian theater over 3 years to foster deeper understandings of the critical, contingent nature of asset-oriented teacher education and the necessity of focusing on relationships and community contexts in order to disrupt the inequitable status quo. Dover et al. (2020) used their

critical, social justice-oriented faculty learning community to support their efforts as teacher educators to teach for social justice. As [Carter Andrews \(2021\)](#) argued, this kind of focus on the development of teacher educators' sociopolitical consciousness is essential for enacting humanizing, justice-oriented teacher education pedagogies.

Thus, in considering the research base on the preparation of novice teacher educators for critical, justice-oriented aims, there are limited but emerging insights. I turn now to offer one case from my own research and practice that demonstrates the role a formal course of study can play in fostering learning among novice teacher educators. Given the particular needs established for teacher educators to learn to enact pedagogies of teacher education ([Goodwin et al., 2014](#)) that enable them to pursue transformative justice ([Souto-Manning and Stillman, 2020](#)), this case highlights a structured set of learning experiences designed to emphasize the pedagogy of teacher education anchored by critical, justice-oriented goals.

A case of formal preparation for novice teacher educators focused on critical, justice-oriented aims

Course background and focus

The research I describe here focuses on a course I taught in 2009 to university doctoral students who were learning to become teacher educators ([Conklin, 2015](#)), what the doctoral students learned from the course, and how this learning shaped their teacher education practice over time ([Conklin, 2020](#)). The course I designed—titled “The Pedagogy of Teacher Education”—sought to help novice teacher educators learn about *how* to teach novice teachers. The pedagogy of teacher education refers to the range of instructional strategies, tasks, assignments, and interactions that teacher educators use in teacher education contexts ([Grossman, 2005](#)), such as using cases for teachers to explore the dilemmas and complexities of practice (e.g., [Darling-Hammond and Hammerness, 2002](#)), modeling practices (e.g., [Loughran and Berry, 2005](#)), or engaging teachers in action research (e.g., [Price and Valli, 2005](#)). In this course, my particular aim was to enable novice teacher educators to gain both theoretical and practical knowledge—conceptual and practical tools ([Grossman et al., 1999](#))—of teacher education pedagogy, rooted in a vision of compassionate, critical, justice-oriented teacher education ([Conklin, 2008](#)).

The 15-week course, which I taught to ten novice teacher educators pursuing doctoral degrees in education at a US university, began by engaging students in exploring conceptual frameworks of teacher learning and scholarship focused on teacher education pedagogy. The course then went on to examine scholarship that emphasized the purposes of justice-oriented teacher education, focusing on approaches that both embrace diversity and critically examine structural inequities (e.g., [Banks et al., 2005](#); [Gay, 2002](#); [McDonald and Zeichner, 2009](#); [Sleeter, 2008](#); [Villegas and Lucas, 2002](#)). Another conceptual thread that ran through the course was the importance of relationships and compassion as elements of the pedagogy of justice-oriented teacher education ([Conklin, 2008](#); [Hanh, 1987](#); [hooks, 2003](#)). These theoretical emphases were complemented by regular class experiences that served as conceptual tools, such as collective discussion of teacher education tasks and assignments in which we used a common set of analytical questions about the goals of the assignment, its key features, and conceptions of learning embedded within it. All of these conceptual frameworks were intended to support the novice teacher educators in better grasping the aims of justice-oriented teacher education and providing them with theoretical guides for their future teacher education practice.

Throughout the course, I paired these conceptual tools with practical tools to help the novice teacher educators envision and experience pedagogies of teacher education in tangible ways. While course readings featured scholarship on specific teacher education pedagogies, we also participated in these pedagogies during class sessions whenever possible. For example, we paired reading about the use of cases in teacher education (e.g., [Darling-Hammond and Hammerness, 2002](#); [Goodwin, 2002](#); [Moje and Wade, 1997](#)) with a class analysis of a teacher education case dilemma I had encountered in my teacher education practice: how to address a white pre-service teacher who enacted racist caricatures in class. Similarly, we read scholarship on pedagogies such as the use of role playing/theater of the oppressed ([Cahnmann-Taylor et al., 2009](#)) and conducting critical readings of texts ([Jones, 2006](#))—practical pedagogical tools aimed at preparing critical, justice-oriented teachers. We also participated in other pedagogies as part of our course sessions, such as using a structured, equity-oriented discussion protocol called the “Final Word” that equalizes participants' opportunities to share ideas and counters traditional patterns of classroom dominance ([Fischer-Mueller and Thomson-Grove, n.d.](#)). All of these practical tools were aimed at offering the novice teacher educators concrete pedagogies that they could enact with novice teachers that were aligned with the course's conceptual underpinnings.

As our collective participation in these practices suggests, a key pedagogy of the course was the modeling of practice ([Loughran and Berry, 2005](#)) and making key features of practice explicit ([Loughran and Berry, 2005](#); [Lunenberg et al., 2007](#))—an approach to fostering teacher learning supported by research. Thus, I engaged novice teacher educators in specific pedagogies in our classroom as a way of demonstrating possible practices they could use as teacher educators; this modeling of practice was also intended to illustrate to the novice teacher educators that they, too, could and should model specific pedagogies as part of their work preparing future teachers. Further, following each demonstration of a teacher education pedagogy, I also modeled the analysis and evaluation of the practices we had used from a pedagogical perspective in order to make its features explicit ([Loughran and Berry, 2005](#)). In some cases, I used this opportunity to reflect aloud and offer critiques of my own practice. Again, this analytical practice of debriefing was aimed at showing novice teacher educators how they could make their own pedagogical choices transparent to novice teachers.

Studying novice teacher educators' learning and practice

Five years after teaching this course, I set out to examine what the novice teacher educators had learned from our formal study of the pedagogy of teacher education and how, if at all, the coursework shaped their own practice as teacher educators. Of the ten students who had enrolled in the course, four were in roles as practicing teacher educators and agreed to be part of the research. Each novice teacher educator completed an initial survey that asked about their current role and the courses they taught. The survey also invited them to reflect on influences on their teacher educator work, provide examples of teacher education practices they were currently using that they deemed successful and unsuccessful, and consider what role—if any—the doctoral coursework had in shaping their practice as teacher educators. Following their survey completion, I interviewed each participant, asking them to discuss influences on their teacher education practice further, talk about the goals of their work, and share sample assignments and syllabi from their practice as teacher educators. The participants provided me with sample assignments they had completed in our course as a way of documenting changes in their learning over time. While only possible in the case of one participant, I also observed a class session of the participant teaching pre-service teachers as one further source of data to trace any relationship between the formal course learning and the novice teacher educators' learning and practice.

Although this exploratory case study only included a small group of participants, the results provided a window into both the potential of formal coursework for teacher educators' learning and valuable practices for the development of justice-oriented teacher educators. While each novice teacher educator's learning was shaped by their own identities, experiences, prior learning, and the contexts in which they worked, the data illuminated both conceptual and practical tools from the coursework that influenced the novice teacher educators' learning and practice (Conklin, 2020).

Learning conceptual and practical tools

First, the research with these participants demonstrated that they learned a range of conceptual tools from the Pedagogy of Teacher Education course that offered helpful guidance in their work as teacher educators. The course's range of teacher learning theories and conceptions of the various purposes for teacher education provided novice teacher educators with theoretical frameworks that enabled them to make intentional choices as teacher educators; these conceptual tools also offered the novice teacher educators a clear language and rationale that helped them provide reasoning for the practices they implemented and justify both pedagogical and programmatic decisions. The theories of teacher learning featured in the course also helped the novice teacher educators to view teachers as learners, something they had not all considered before. Further, the value of modeling as part of teacher education practice was also salient for the novice teacher educators; they came to understand the importance of being intentional as a teacher educator and making these intentions transparent in their work with novice teachers. Finally, the coursework's conceptual underpinnings of compassionate, critical, justice-oriented teacher education showed up in various aspects of the novice teacher educators' learning and practice, shaping their interactions and pedagogies with novice teachers, the purposes of the teacher education experiences they designed, and their teacher education course planning (Conklin, 2020).

Second, this exploratory study of novice teacher educators' learning from the formal coursework revealed practical tools, too, that contributed to the teacher educators' development. In many cases, the participants reported directly using particular strategies or assignments that they learned in our coursework in their work with novice teachers. They also noted that they used the practice of modeling, debriefing practices, and making pedagogical decisions explicit, following the use of these practices in our coursework. Thus, the research highlighted the value of offering—and modeling—specific teacher education pedagogies for novice teacher educators. These specific pedagogies served not as a prescribed set of practices for teacher educators but instead gave novice teacher educators an array of concrete possibilities for how they might facilitate the learning of novice teachers (Conklin, 2020).

Unplanned course learning

This exploratory case study also revealed that the novice teacher educators gained additional learning from the course that had not been a planned, intentional focus. Specifically, these novice teacher educators reported finding value in the ways that I, as an instructor, demonstrated a willingness to take pedagogical risks, show vulnerability, and be explicit about times I thought my instruction had not succeeded. The study participants noted that, particularly as novice teacher educators who were nervous about their developing craft, it was powerful for them to see another teacher educator be open about the unpredictability of teacher education practice and the insecurities that come along with our work. Additionally, the novice teacher educators in the study noted that the course helped them understand the incredible complexity of the work of teacher education—something which I also see as a fundamental quality of educating teachers, but a quality I had not made an intentional goal for the course learning.

Implications of this case

The case of this formal coursework focused on the pedagogy of teacher education demonstrates one avenue for supporting the learning of novice teacher educators and helping those who are new to the craft develop and enact pedagogies for teacher education focused on justice-oriented goals. Specifically, the study indicated that the intentionally designed course experiences enabled novice teacher educators to learn both conceptual and practical tools for teacher education pedagogy. These tools included the recognition of teachers as learners; critical, justice-oriented theories of teaching and teacher education; and pedagogical tools that support those justice-oriented goals. Further, in the face of increasingly reductive, technical approaches to teaching that have gained traction globally (e.g., Lemov, 2010), this formal coursework provided one approach to helping novice teacher educators understand the

complexity of teacher education practice, resist simplistic models of practice, and learn tools to transform teacher education (Loughran and Menter, 2019).

Looking more broadly, this example points to a variety of implications for the preparation of justice-oriented teacher educators. Providing novice teacher educators with both conceptual and practical tools that are connected to critical, justice-oriented goals offers a crucial pairing; as others have argued, in order to reform teacher education, it is essential for the field to have not only clear conceptualizations of socially-just, transformative goals, but also corresponding teacher education practices that support these goals (Cochran-Smith et al., 2009). Indeed, many scholars have made the critical point that teacher educators need concrete, practical approaches that are clearly connected to theoretical constructs that allow for decision-making in complex and changing contexts (Cochran-Smith et al., 2020; Goodwin and Kosnik, 2013; Peercy and Troyan, 2017). The research case I describe also adds to other scholarship that indicates the potentially important role of modeling vulnerability and risk-taking as key pedagogies in the preparation of justice-oriented teacher educators (cf., Berry and Russell, 2016; Dale and Frye, 2009; hooks, 2003; Peercy and Troyan, 2017)—pedagogies that should likely go along with modeling flexibility, openness, a commitment to continual learning, and humility (Cochran-Smith et al., 2020). Additionally, this research supports the idea that formally organized coursework can play an important role in teacher educators' development, alongside more contextually-driven study into one's own practice (Goodwin et al., 2014; Zeichner, 2005).

Conclusion

With this case of research and practice in mind, then, along with the broader research on teacher educators' learning and development, how should we prepare novice teacher educators to foster critical, justice-oriented practice among teachers? In part, this body of work suggests that advancing the aims of critical, justice-oriented teacher education will require a multitude of coordinated approaches for preparing and developing teacher educators. Such approaches should equip teacher educators with clear conceptual frameworks to guide their practice along with practical approaches that are well-aligned with these theoretical underpinnings and that highlight the complex, contextual, socio-political, and relational nature of transformative teacher education (Cochran-Smith et al., 2020; Stillman and Beltrano, 2019).

The existing research also suggests that the preparation and development of teacher educators must be planned intentionally and prioritized, particularly to support critical, justice-oriented goals. Providing more just and equitable learning experiences for all K-12 students relies on well-prepared teachers with strong sociopolitical consciousnesses—which, in turn, relies on well-prepared teacher educators. While the research literature illustrates that there are important efforts underway in contexts around the globe to support novice teacher educators' learning, such efforts remain small and scattered relative to the vast need that exists. Prioritizing the professional learning of teacher educators in order to support the work of asset-oriented, culturally sustaining teaching is an urgent need.

References

- Abell, S.K., Park Rogers, M.A., Hanuscin, D.L., Lee, M.H., Gagnon, M.J., 2009. Preparing the next generation of science teacher educators: a model for developing PCK for teaching science teachers. *J. Sci. Teach. Educ.* 20, 77–93.
- Banks, J., et al., 2005. Teaching diverse learners. In: Darling-Hammond, L., Bransford, J. (Eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey Bass, San Francisco, pp. 232–274.
- Berry, A., 2009. Professional self-understanding as expertise in teaching about teaching. *Teach. Teach. Theor. Pract.* 15 (2), 305–318. <https://doi.org/10.1080/13540600902875365>.
- Berry, A., Russell, T., 2016. Self-study and vulnerability. *Stud. Teach. Educ.* 12 (2), 115–116.
- Bullock, S.M., 2009. Learning to think like a teacher educator: making the substantive and syntactic structures of teaching explicit through self-study. *Teach. Teach. Theor. Pract.* 15 (2), 291–304. <https://doi.org/10.1080/13540600902875357>.
- Bullock, S.M., Christou, T., 2009. Exploring the radical middle between theory and practice: a collaborative self-study of beginning teacher educators. *Stud. Teach. Educ.* 5 (1), 75–88. <https://doi.org/10.1080/17425960902830468>.
- Cahnmann-Taylor, M., Wooten, J., Souto-Manning, M., Dice, J.L., 2009. The art and science of educational inquiry: analysis of performance-based focus groups with novice bilingual teachers. *Teach. Coll. Rec.* 111 (11), 2535–2559.
- Carter Andrews, D., 2021. Preparing teachers to be culturally multidimensional: designing and implementing teacher preparation programs for pedagogical relevance, responsiveness, and sustenance. *Educ. Forum* 85 (4), 416–428. <https://doi.org/10.1080/00131725.2021.1957638>.
- Cochran-Smith, M., 2003. Learning and unlearning: the education of teacher educators. *Teach. Teach. Educ.* 19 (1), 5–28. [https://doi.org/10.1016/S0742-051X\(02\)00091-4](https://doi.org/10.1016/S0742-051X(02)00091-4).
- Cochran-Smith, M., Villegas, A., 2016. Research on teacher preparation: charting the landscape of a sprawling field. In: Gitomer, D., Bell, C. (Eds.), *Handbook of Research on Teaching*, fifth ed. American Educational Research Association, Washington, DC, pp. 439–547.
- Cochran-Smith, M., Barnatt, J., Lahann, R., Shakman, K., Terrell, D., 2009. Teacher education for social justice: critiquing the critiques. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, New York, pp. 625–639.
- Cochran-Smith, M., Grudnoff, L., Orland-Barak, L., Smith, K., 2020. Educating teacher educators: international perspectives. *N. Educ.* 16 (1), 5–24.
- Conklin, H.G., 2008. Modeling compassion in critical, justice-oriented teacher education. *Harv. Educ. Rev.* 78 (4), 652–674.
- Conklin, H., 2015. Preparing novice teacher educators in the pedagogy of teacher education. *Action Teach. Educ.* 37 (4), 317–333. <https://doi.org/10.1080/01626620.2015.1078753>.
- Conklin, H., 2020. The preparation of novice teacher educators for critical, justice-oriented teacher education: a longitudinal exploration of formal study in the pedagogy of teacher education. *Teach. Teach. Theor. Pract.* 26 (7–8), 491–507. <https://doi.org/10.1080/13540602.2021.1889497>.
- Conklin, H., Hughes, H., 2016. Practices of compassionate, critical, justice-oriented teacher education. *J. Teach. Educ.* 67 (1), 47–60. <https://doi.org/10.1177/0022487115607346>.
- Dale, M., Frye, E., 2009. Vulnerability and love of learning as necessities for wise teacher education. *J. Teach. Educ.* 60 (2), 123–130.

- Darling-Hammond, L., Hammerness, K., 2002. Toward a pedagogy of cases in teacher education. *Teach. Educ.* 13 (2), 125–135. <https://doi.org/10.1080/1047621022000007549>.
- Dover, A., Kressler, B., Lozano, M., 2020. "Learning our way through": critical professional development for social justice in teacher education. *N. Educ.* 16 (1), 45–69.
- Fischer-Mueller, J., Thomson-Grove, G., (n.d.). The Final Word. National School Reform Faculty, https://www.nsrharmony.org/wp-content/uploads/2017/10/final_word_0.pdf.
- Gay, G., 2002. Preparing for culturally responsive teaching. *J. Teach. Educ.* 53 (2), 106–116. <https://doi.org/10.1177/0022487102053002003>.
- Goodwin, A.L., 2002. The case of one child: making the shift from personal knowledge to professionally informed practice. *Teach. Educ.* 13 (2), 137–154. <https://doi.org/10.1080/1047621022000007558>.
- Goodwin, A., Kosnik, C., 2013. Quality teacher educators = quality teachers? Conceptualizing essential domains of knowledge for those who teach teachers. *Teach. Dev.* 17, 334–346. <https://doi.org/10.1080/13664530.2013.813766>.
- Goodwin, A.L., Smith, L., Souto-Manning, M., Charuvu, R., Tan, M.Y., Reed, R., Taveras, L., 2014. What should teacher educators know and be able to do? Perspectives from practicing teacher educators. *J. Teach. Educ.* 65 (4), 284–302. <https://doi.org/10.1177/0022487114535266>.
- Grossman, P., 2005. Research on pedagogical approaches in teacher education. In: Cochran-Smith, M., Zeichner, K.M. (Eds.), *Studying Teacher Education*. Lawrence Erlbaum, Mahwah, NJ, pp. 425–476.
- Grossman, P., Smagorinsky, P., Valencia, S., 1999. Appropriating tools for teaching English: a theoretical framework for research on learning and teaching. *Am. J. Educ.* 108, 1–29.
- Hadar, L., Brody, D., 2013. The interaction between group processes and personal professional trajectories in a professional development community for teacher educators. *J. Teach. Educ.* 64 (2), 145–161.
- Hadar, L., Brody, D., 2017. Professional learning and development of teacher educators. In: Clandinin, D.J., Hsu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*, first ed. SAGE Publications Limited, Thousand Oaks, CA, pp. 1049–1064.
- Hanh, T.N., 1987. *Being Peace*. Parallax, Berkeley, CA.
- hooks, B., 2003. *Teaching Community: A Pedagogy of Hope*. Routledge, New York.
- Jacobs, J., Yendol-Hoppey, D., Dana, N.F., 2015. Preparing the next generation of teacher educators: the role of practitioner inquiry. *Action Teach. Educ.* 37, 373–396. <https://doi.org/10.1080/01626620.2015.1078758>.
- Jones, S., 2006. Lessons from Dorothy Allison: teacher education, social class, and critical literacy. *Chang. Engl.* 13 (3), 293–305. <https://doi.org/10.1080/13586840600971802>.
- Karagiorgi, Y., Nicolaidou, M., 2013. Professional development of teacher educators: voices from the Greek-Cypriot context. *Prof. Dev. Educ.* 39 (5), 784–798.
- Korthagen, F., Loughran, J., Lunenberg, M., 2005. Teaching teachers—studies into the expertise of teacher educators: an introduction to this theme issue. *Teach. Teach. Educ.* 21, 107–115.
- Koster, B., Dengerink, J., Korthagen, F., Lunenberg, M., 2008. Teacher educators working on their own professional development: goals, activities and outcomes of a project for the professional development of teacher educators. *Teach. Teach. Theor. Pract.* 14 (5–6), 567–587. <https://doi.org/10.1080/13540600802571411>.
- Lemov, D., 2010. *Teach Like a Champion*. Jossey-Bass, San Francisco.
- Loughran, J., 2007. Researching teacher education practices: responding to the challenges, demands, and expectations of self-study. *J. Teach. Educ.* 58 (1), 12–20. <https://doi.org/10.1177/0022487106296217>.
- Loughran, J., Berry, A., 2005. Modelling by teacher educators. *Teach. Teach. Educ.* 21 (2), 193–203. <https://doi.org/10.1016/j.tate.2004.12.005>.
- Loughran, J., Menter, I., 2019. The essence of being a teacher educator and why it matters. *Asia Pac. J. Teach. Educ.* 47 (3), 216–229. <https://doi.org/10.1080/1359866X.2019.1575946>.
- Lunenberg, M., 2002. Designing a curriculum for teacher educators. *Eur. J. Teach. Educ.* 25 (2&3), 263–277.
- Lunenberg, M., Korthagen, F., Swennen, A., 2007. The teacher educator as a role model. *Teach. Teach. Educ.* 23 (5), 586–601. <https://doi.org/10.1016/j.tate.2006.11.001>.
- McDonald, M., Zeichner, K., 2009. Social justice teacher education. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, New York, pp. 595–610.
- Moje, E.B., Wade, S.E., 1997. What case discussions reveal about teacher thinking. *Teach. Teach. Educ.* 13 (7), 691–712. [https://doi.org/10.1016/S0742-051X\(97\)00015-2](https://doi.org/10.1016/S0742-051X(97)00015-2).
- Murray, J., 2017. Defining teacher educators: international perspectives and contexts. In: Clandinin, D.J., Hsu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*, first ed. SAGE Publications Limited, Thousand Oaks, CA, pp. 1017–1032.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97. <https://doi.org/10.3102/0013189X12441244>.
- Peercy, M., Troyan, F., 2017. Making transparent the challenges of developing a practice-based pedagogy of teacher education. *Teach. Teach. Educ.* 61, 26–36. <https://doi.org/10.1016/j.tate.2016.10.005>.
- Ping, C., Schellings, G., Beijard, D., 2018. Teacher educators' professional learning: a literature review. *Teach. Teach. Educ.* 75, 93–104. <https://doi.org/10.1016/j.tate.2018.06.003>.
- Pinnegar, S., Hamilton, M.L., 2017. The promise of the particular in research with teacher educators. In: Clandinin, D.J., Hsu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*, first ed. SAGE Publications Limited, Thousand Oaks, CA, pp. 1065–1080.
- Price, J., Valli, L., 2005. Preservice teachers becoming agents of change: pedagogical implications for action research. *J. Teach. Educ.* 56 (1), 57–72. <https://doi.org/10.1177/0022487104272097>.
- Ritter, J.K., 2007. Forging a pedagogy of teacher education: the challenges of moving from classroom teacher to teacher educator. *Stud. Teach. Educ.* 3 (1), 5–22. <https://doi.org/10.1080/17425960701279776>.
- Ritter, J.K., 2010. Revealing praxis: a study of professional learning and development as a beginning social studies teacher educator. *Theor. Res. Soc. Educ.* 38 (4), 545–573.
- Sleeter, C., 2008. Preparing White teachers for diverse students. In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, J. (Eds.), *Handbook of Research on Teacher Education: Enduring Issues in Changing Contexts*, third ed. Routledge, New York, pp. 559–582.
- Souto-Manning, M., Stillman, J., 2020. In the pursuit of transformative justice in the education of teacher educators. *N. Educ.* 16 (1), 1–4.
- Stillman, J., Beltramo, J.L., 2019. Exploring Freirean culture circles and Boalian theatre as pedagogies for preparing asset-oriented teacher educators. *Teach. Coll. Rec.* 121, 1–38.
- Stillman, J., Ahmed, K.S., Beltramo, J.L., Cataneda-Flores, E., Garza, V.G., Pyo, M., 2019. From the ground up: cultivating teacher educator knowledge from the situated knowledges of emerging, asset-oriented teacher educators. *Asia Pac. J. Teach. Educ.* 47 (3), 265–285. <https://doi.org/10.1080/1359866X.2019.1600187>.
- Villegas, A.M., Lucas, T., 2002. Preparing culturally responsive teachers: rethinking the curriculum. *J. Teach. Educ.* 53 (1), 20–32.
- Vogel, A., Bartlett, A., 2013. Passing on a teacher educator's knowledge: self-study of a doctoral teaching internship. *Stud. Teach. Educ.* 9 (2), 130–140. <https://doi.org/10.1080/17425964.2013.808048>.
- Wolkenhauer, R., Hooser, A., 2021. Becoming clinically grounded teacher educators: inquiry communities in clinical teacher preparation. *J. Teach. Educ.* 72 (2), 168–179.
- Yow, J.A., Eli, J.A., Beisiegel, M., McCloskey, A., Welder, R.M., 2016. Challenging transitions and crossing borders: preparing novice mathematics teacher educators to support novice K-12 mathematics teachers. *Math. Teach. Educ. Dev.* 18 (1), 52–69.
- Zeichner, K.M., 2005. Becoming a teacher educator: a personal perspective. *Teach. Teach. Educ.* 21 (2), 117–124. <https://doi.org/10.1016/j.tate.2004.12.001>.
- Zeichner, K., Conklin, H.G., 2005. Teacher education programs. In: Cochran-Smith, M., Zeichner, K. (Eds.), *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education*. Lawrence Erlbaum, Mahwah, N.J.

Teacher educators' professional growth as researchers

Leah Shagrir, Levinsky College of Education, Tel Aviv-Yafo, Israel

© 2023 Elsevier Ltd. All rights reserved.

Introduction	504
Professional growth	505
Three-stage model of scholarly growth in teacher education - ISE	505
The basic stage of the ISE Model – building professional identity as a researcher - I	506
The intermediate stage of the ISE Model – conduct as a scholar in higher education – S	507
The third stage of the ISE Model – conduct as an expert – E	508
Concluding remarks	508
Why is it important to get to know the ISE Model?	508
How should the ISE Model be used?	509
What is the contribution of the ISE Model to teacher educators?	509
References	509

Introduction

Faculty members in higher education are required to excel in teaching, undertake scholarship activities, present research excellence, and work intensively to become experts in their profession. Academic promotion is measured by achievements reached by faculty as lecturers, as researchers contributing to the body of knowledge, and as contributors to varied services' activities (Brew, 2010; Halse et al., 2007; Shagrir, 2015a).

As faculty members in higher education, teacher educators are required to act at all times as professionals researching their work, their teaching, their professional pathways, theirs' and their students' achievements, and the body of knowledge they teach. Researchers emphasize focusing on conducting research and investigating teaching as activities that should be part of the professional daily routine of those who prepare students to be professionals in the teaching profession (Griffiths et al., 2014; Hökkä et al., 2012; Murray, 2010).

Researchers have found that teacher educators for whom undertaking research is the norm and resolute and long-term occupation with activities of a research nature optimize their functioning in teaching and working with students (Russell, 2010; Tack and Vanderlinde, 2016), and they serve as good role models for students who will be, in the future, teachers like them (Berry and Scheele, 2007; Loughran, 2014; Loughran and Berry, 2005; Shagrir, 2010a).

Teacher educators are called upon to enlist in this matter. The more they continue to research in their fields and areas, and the more they add important outputs, findings, and conclusions, so will opportunities for them to expand their scholarship achievements increase to expand the profession's body of knowledge and improve the status of teacher education as a profession among all other academic professions (Cochran-Smith, 2005; Lunenberg et al., 2014; Murray, 2010; Shagrir, 2011).

Teacher education is a profession in and of itself (Bridges, 1999; Shulman, 1998) and its features require those who are occupied in it to be qualified and master specific skills (Hadar and Brody, 2017; Koster et al., 2005; Loughran, 2006). Similar to academic programs educating students for professional occupations such as medicine, para-medical professions, social work and others, teacher education programs also aspire to instill in students mastery of the profession's secrets, existing educational doctrines, and theories, as well as mastery of the practical skills of teaching and learning serving those occupied in the teaching profession. However, teacher education programs have a unique nature, and components that are not included in academic education programs for other professions; therefore, teacher educators who lead them require unique qualifications. These unique qualifications are essential by virtue of teacher educators serving in their teaching and professional conduct as role models for students and by virtue of they being examples encouraging future teachers to act as teacher-researchers as well (Berry and Scheele, 2007; Lunenberg et al., 2007; Shagrir, 2010b).

Since the teaching profession is in essence a practical profession, teacher educators are required to demonstrate expertise both in constructing and leading quality programs and courses, and in building an envelope including everything that must be experienced in teaching practice that allows students to progress toward being quality teachers (Korthagen, 2014; Korthagen et al., 2005; Swennen et al., 2008). Likewise, teacher educators must ensure that students acquire knowledge in subject areas taught in schools, acquire the qualifications, tools, and skills necessary for their work and quality teaching, learn how to apply the knowledge, theories and professional doctrines in education and pedagogy, get to know the educational field in which they will have to operate and more (Shagrir, 2013a).

Loughran (2014) added that over and above teaching skills and working with students, teacher educators must possess abilities to carry out researching teacher practices, skills and means of attaining them can be acquired through professional development

activities. Therefore, educational systems, institutions, unions and professional bodies are called upon to continue to encourage involvement, learning and scholarship activities among teacher educators. They are called upon to mark out cornerstones and develop a variety of means that lead to offering interesting and fruitful professional and scholarship activities (Griffiths et al., 2014; Shagrir, 2017c).

Professional growth

Faculty members interested in being experts in their profession, and who want to advance in their academic career, must meet promotion requirements published by each academic institution. Although promotion requirements differ between institutions, most require faculty members to invest in activities leading to professional development and do this continuously to achieve full mastery of skills, content, and unique features of the profession (Avalos, 2011; Boyd et al., 2007; Murray, 2010). Guidelines call on faculty members to participate in professional development activities, in other words, to be consumers using up to date knowledge emerging from studies conducted by others, and also to initiate conducting long and short-term professional and quality research (Loughran, 2014; Shagrir, 2017c).

Participation in professional and scholarship activities is seen as contributing to teacher educators building their professional identity as teacher education professionals (Dinkelman, 2011; Griffiths et al., 2014; Hökkä et al., 2012), and is perceived to influence ways of teaching, and thus also students' learning achievements. Long-term professional development is likely to lead to expertize in levels of teaching performance, which will result in students' meaningful learning (Jarvis et al., 2013; Shagrir, 2015b).

Enlisting in the scholarly growth journey, is viewed as contributing to improving and enhancing teacher educators' work and professionalism on several dimensions (Shagrir, 2010a). One dimension is structuring the professional self. This is affected by the professional nature and level of activities carried out, and the effect of consolidating conceptual role perceptions in relation to the educational world, the world of teacher education, the role of teacher education, and the professional theories and doctrines on which it is founded. Another dimension is expanding professional relationships established over time and increasing the intensity of these relationships. This dimension is expressed in the nature of collaboration with partners such as, students, professional colleagues and professional community members, and the nature of cooperation with educational institutions, education systems and professional bodies and unions (Berry and Scheele, 2007). Participation in scholarship activities deepens another dimension, the personal dimension expressed in the degree, scope, power, and willingness of individuals to commit and dedicate themselves to scholarly growth and their professional development (Shagrir, 2010a).

Researchers emphasize that teacher educators should constantly engage in their own professional development and in research as long as they serve as teacher educators. They should publish the professional knowledge they possess, reveal, and clarify their insights, and be aware of the impact of their work on students (Boyd et al., 2007; Hökkä et al., 2012; Koster et al., 2008; Lunenberg and Hamilton, 2008; Lunenberg, 2010; Shagrir, 2013b, 2015a).

In conclusion, teacher educators should meticulously investigate their work, and constantly develop new directions of improving it. All these will contribute to their progress toward expertize, meeting promotion requirements, academic promotion and career advancement, will promote their and their institution's national and international reputation, and in general contribute to improving the quality of teacher education and affect education systems in which they operate (Berry, 2009; Hadar and Brody, 2017; Lunenberg and Hamilton, 2008).

Three-stage model of scholarly growth in teacher education - ISE

The three stage Model of Scholarly Growth in Teacher Education - ISE, follows processes occurring during faculty members' development journey and was built after accumulating decades' experience working in teacher education. The model is based on a profound, reflexive analysis (Attia and Edge, 2017; Berger, 2015; Brand, 2015; Palaganas et al., 2017) of the author's professional development steps, and an analysis of what teacher educators from diverse institutions and countries have stated in research studies including dozens of conversations and interviews. Research findings reflected development and professionalization processes among colleagues helped to build the foundation of the ISE Model.

The published model (Shagrir, 2020) consists of three stages: the first is the stage of induction into teaching in higher education and formulating worldviews and ways of working as members of the profession. The second stage expands the range, intensity and depth of professional development and scholarship activities undertaken to acquire expertize and academic promotion. The third stage presents the role and contribution of those who have reached achievements and are considered experts in the profession in which they are occupied, experts in the area of knowledge they research and teach, and highly experienced researchers.

Details and more depth in describing the features of each of the three stages in the ISE Model establish its important meaningfulness as a tool for scholarly growth, and a professional tool helping faculty members to get to know the spectrum of opportunities available to them in an academic career. The model includes signposts, serves to direct growth, and teaches about challenges and milestones facing those who want to grow and reach desired academic promotion.

The ISE Model is presented as a linear-processive model. This mean of presentation allows expanding and adding explanations clarifying the place and importance of each stage. Each stage in the model derives from, relies on and grounds its previous stage. Every stage presents the unique features and categories added to faculty growth and its involvement in research activities. The model

describes how progress is created in research work, and the effects of professional growth on shaping professional views, shaping professional identity, and establishing research skills. The model's description provides details about tools and skills researchers require at different stages and includes options available to them to deepen and navigate their research work, and to focus on collaborative research with students, colleague, experts from similar and different areas of interest.

The description here expands on what has been presented in the past (Shagrir, 2020) mainly for the purpose of examining scholarship and research activities, which are considered a central component of growth in professional functioning for teacher educators who are also researchers.

The basic stage of the ISE Model – building professional identity as a researcher - I

Candidates for the role of teacher educators in most higher education institutions must have a higher academic degree. For most, induction into the role does not require designated previous education including knowing the role's demands and accompanying challenges (Loughran, 2006; Murray and Male, 2005; Swennen et al., 2009). In some teacher education institutions, there are no acceptance conditions requiring candidates to have previous experience in teaching, for example teaching children at schools. Hence, research literature often addresses suggestions and advice for successful induction into the profession and raises the difficulties and challenges facing those deciding to join the role of teacher educators (Boyd et al., 2007; Swennen and van der Klink, 2009; van Velzen et al., 2010).

As previously mentioned, upon induction into the role, and throughout the years of work in higher education, faculty members should aspire to academic promotion, reach professional achievements, acquire academic and professional expertise in their teaching areas and the teacher education body of knowledge. Upon induction, teacher educators are busy building their professional identity including acquiring suitable teaching tools to teach teaching, building ways of working and tools to implement pedagogical considerations as role models for students, building working relationships with students, developing reflective skills and tools to integrate theory and practice, building organized and established doctrines as part of their professional credo, developing suitable values, formulating professional conceptions and the like (Dinkelman, 2011; Hökkä et al., 2012; McKeon and Harrison, 2010; Williams and Ritter, 2010). The entirety of these tasks and challenges facing early career teacher educators are included in the basic stage of the growth model, where they acquire and build their professional identity as members of the profession. This stage is called Identity and is symbolized in the model's name with the capital letter I.

The building professional identity stage also includes establishing a foundation for professional growth as early career researchers. One can see that in this I stage teacher educators acquire basic tools to serve them as researchers learning the secrets of the academic world which they have entered, and bodies of knowledge they teach (Beaudry and Miller, 2016).

The basic tools early career researchers employ include intelligent consumerism of studies published on academic stages such as journal articles, textbooks, lectures, presentations and so on. As intelligent consumers, they will utilize up to date research findings and conclusions and integrate them into various courses and programs they teach. At the I stage, faculty will develop skills of reading, listening, and observing academic publications presented in various formats, will acquire the language, concepts and terminologies of the research world and academic publications, and will constantly be updated about developments in the professional knowledge world in their taught subjects. On the basis of the knowledge and skills acquired, they will also update their students on professional innovations both in different teaching situations and syllabi.

In most cases researchers in their Identity stage have experienced in the past conducting several in-depth and broad ranging studies, mainly in writing their theses as part of the requirements for higher academic degrees. Thesis research leads to a significant change in the way researchers perceive themselves as researchers, how they perceive the need for time and resources to conduct research, and their awareness of everything needed for research to be reliable, valid and professional to meet the requirements of academic research.

In a few cases, there are faculty members who avoid undertaking any task related to research after completing their theses in academic frameworks. The intensity, energies and resources needed to carry out studies create among some people a fear of not being able to withstand again such tasks even if they are smaller in scope (Shagrir, 2017c).

Murray (2008) pointed out that not all teacher educators are meant to undertake research, but all of them must have insights about studies and materials published, must be involved and aware of innovations emerging from research, and persist in enquiries into personal practice.

Those who are interested in progressing as researchers invest efforts and time resources in a range of activities such as initiating and planning studies, conducting disciplinary and/or pedagogical research, acquiring the correct tools to submit applications for research grants, responding to calls for proposals to conduct studies, expanding collaborations with fellow researchers, promoting research in their knowledge areas, and the like.

Some of those interested in progressing along academic scales see research conducted in their theses as an inviting event with whose conclusion, and because of it, they can be part of the research world and are able to advance it and come out with additional research work. Such work includes, for example, investing in disseminating thesis research outputs on accepted academic stages and turning their accumulated knowledge into knowledge publicly available, and especially to their colleagues in the profession (Hadar and Brody, 2017; Murray, 2010). Another direction taken by researchers in their early professional careers is conducting studies based on their theses research results, such as follow-up or comparative studies. Furthermore, they conduct new research, sometimes using research methodologies known to them, and sometimes experiencing and acquiring tools to employ new research methodologies or explore areas they have not investigated before.

Institutions view scholarship and research activities as part of what is required of early career teacher educators, and failure to participate in such activities is likely to threaten their continued employment in an academic framework. Most academic institutions make available to staff infrastructure and platforms to encourage research activity, and even offer researchers assistance, learning, and support under research centers for faculty. Among others, professional support is offered by colleagues and experienced role holders, research budgets, calls to conduct research and more. Some proposals call to conduct studies within employing institutions. Such studies focus, for example, on investigating institutional worldviews, curricula, or vision, or on studying populations operating in an institution such as lecturers or students. Some proposals call to conduct studies about populations and topics outside an institution, such as systemic, social, educational, or disciplinary topics. Some proposals call to conduct studies in the framework of interinstitutional and/or international partnership (Brooks, 2005; Geerdink et al., 2016; Rosser and Johnsrud, 2003).

Such conduct in institutions encourages early career researchers to continue to study, to undertake diverse opportunities to conduct research, and thus to exploit them as opportunities contributing to developing and establishing their professional confidence as researchers and as professionals (Geerdink et al., 2016; Murray and Male, 2005; Shagrir, 2010a; Williams and Ritter, 2010).

Achievements during development at the basic I stage lead researchers to continued professional development and promotes their growth to the next, second stage in the Model of Scholarly Growth - ISE.

The intermediate stage of the ISE Model – conduct as a scholar in higher education – S

At the next stage of the Model of Scholarly Growth - ISE, faculty members focus on a broader scope of scholarship and research activities continuing from those carried out at the Identity stage. Together with performing constant research activities based on experience, knowledge and tools acquired at the previous stage, faculty also invest resources and time in expanding their professional conduct as professional scholars. This stage of professional growth is called Scholar and is therefore symbolized in the model's name by the capital letter S.

Researchers who have reached the S stage already have experience in organizing and leading studies, sometimes have even led teams of researchers, and have mastered research skills and met deadlines set for completing research. They apply the worldview consolidated and professional conceptions they had built at the previous stage as a foundation for their perception as researchers, and basis for their view of research's place and contribution. At this stage, researchers enhance their acquired abilities and their undertakings also make a relevant contribution to the world of content and that of teaching, and present informed recommendations with a broad impact and challenging ideas for those interested in future research. Researchers carrying out collaborative studies develop abilities to establish meaningful work relationships during research and ongoing cooperation. They develop abilities to resolve challenges arising from disagreements that sometimes occur between researchers, learn to contain others' ideas and how to clearly present their ideas (Koster et al., 2008; Murray et al., 2009; Shagrir, 2017a).

Expanding professional conduct and professional identity as scholars means operating in parallel in two dimensions that fit together and nourish each other: the one is conduct as professional scholars and the other - conduct as professional researchers. The latter dimension at the S stage means employing various research tools, expanding use of diverse research methodologies, choosing diverse research populations, enriching examined areas, choosing a range of research partners, conducting studies requiring long-term work such as in-depth studies, follow-up studies, comparative studies, and the like (Swennen et al., 2017; Tack and Vanderlinde, 2014).

Conduct as professional scholars is characterized by also addressing activities that interface with the world of research and mainly derive from it. Professional scholars labor to disseminate the fruit of their studies, invest efforts to make them known to the public, to persuade others that the fruits of their studies are important and meaningful, and track how much and to what extent these fruits indeed influence the body of knowledge in the scientific world (Denemark and Espinoza, 1974; Haviland et al., 2010; Rubin and Zadik, 2015). Faculty at stage S are meant to be experienced in clear and comprehensive presentation of their worldviews and professional visions to others and diverse institutions, bodies, and systems. Activities they perform as scholars will also leverage and promote their academic reputation and that of the academic institution where they work and research (Swennen et al., 2017).

To illustrate: mass presentation of research outputs while using specific skills such as writing publications for prestigious academic professional forums. Being keynote speakers at conferences and study days. Leading core agenda in social, academic, or professional areas, and more. Such platforms require scholars to invest time and learn how to do it and acquire specific skills and abilities that will enable them to meet their rigorous demands.

Another example of conduct as professional scholars is being concerned with activities leading to securing resources and grants to conduct research. The essence of these activities is mostly developing abilities to respond to calls for proposals from research foundations or invitations to conduct research for bodies, associations, institutions etc. These activities mostly have a systemic nature, whether institutional, national, or international and sometimes require collaboration with and including research colleagues from diverse institutions, associations, and organizations.

Further and many examples teach us that this layer of work as researchers requiring partnerships of such scope deepen and expand professional worldviews and enable researchers to study a range of research populations, and meet teacher education cultures different to that which is known. In this way researchers operate as part of the scientific community and work to promote their place and name in the community and contribute to the scientific world.

Successes and achievements in diverse and extensive activity performed by scholars at the S stage are the basis for transition to the third and advanced stage of the ISE Model of professional and academic scholarly growth.

The third stage of the ISE Model – conduct as an expert – E

The third stage of the Model of Scholarly Growth - ISE is the stage at which faculty function as both experts in knowledge areas they teach and research, and experts who have mastered the enigma of the academic research world. Hence the third stage of the model is called the Expert stage and is symbolized in the model's name by the capital letter E.

This stage is attained by researchers who possess a consolidated professional identity, professional confidence, rich teaching experience, broad and profound mastery of knowledge areas, qualifications, and skills, as detailed in the characteristics of the previous two model stages (Bradley et al., 2006; Kinchin et al., 2008).

At the E stage, faculty continue to function persistently and in great depth as researchers and at the same time invest resources in additional dimensions. They invest mainly in activities where expertise is demanded and their contribution as experts influencing the professionalism of education systems and higher education, as helping to build education systems' worldviews and policies, as partners in building policies, visions and work plans for influential and leading bodies. The nature of activities in stage E changes from those focusing mainly on achievements in one's personal professionalization and academic-profession progress, to activities with a nature of service and contribution to benefit the profession, the scientific world, and society. Through such activities experts influence the areas in which they have worked and specialized, and become those leading the scientific world, profession, and research in higher education (Shagrir, 2017b).

Below are some examples of scholarship activities of a nature that contributes to and influences others and the world of knowledge in which experts have specialized: writing and publishing position papers, opinions, and reviews in professional areas. Active leadership of academic conferences, interest groups, forums and think tanks. Active membership of journal and academic books' editorial boards. Taking part as members or managers of professional bodies and associations. Heading grant award committees and leading professional working bodies as known experts. Main undertakings are for the benefit of building, leading and implementing professional agenda and systemic policy in education systems, higher education systems and/or teacher education systems.

As mentioned, most these activities, which are only a few examples of possible undertakings, are not always considered part of the required obligations of those operating in the promotion process framework and personal promotion in academic institutions' professorial levels. Nonetheless, these activities certainly advance the body of knowledge in teacher education, the world of higher education and thus society too.

Research literature has taught us that some expert researchers saw their achievements as a good opportunity to conduct reflexive research on their scholarly growth (Loughran, 2014; Russell, 2018; Young and Erickson, 2011; Zeichner, 2005). Reflexive research enables an examination of the overall activities, experiences and achievements reached over years. It allows identification of the nature of contributions made to the academic world and body of knowledge and thus reach conclusions regarding the next steps of their scholarly growth. Using reflexive research methodologies, researchers examine and analyze their career progression, stations and junctions of their progress and growth, difficulties and challenges encountered, solutions they chose, and degree of expertise reached and its benefits. In light of findings emerging from reflexive analyses, experts can plan their continued professional journey and possibilities available to them in their academic-professional journey. They can decide in which areas and topics it is worthwhile investing resources and plan the pathways on which they will focus to continue to influence with the expertise they have accumulated over years.

Those who define themselves as professionals may conduct a reflexive examination of their professional paths at an earlier stage of their development and analyze their challenges and successes by revealing generic insights assisting them to explore directions for continued professional undertakings. Reflexive research carried out at the third stage of scholarly growth helps identify historical growth steps, examine them in comparison to professional colleagues, and through profound introspection produce broad and universal insights about the uniqueness of their journey and voice, insights that are likely, among others, to help colleagues at earlier stages of their growth (Brand, 2015; Day, 2012).

Concluding remarks

The three stage Model of Scholarly Growth - ISE presents the features of faculty members' conduct and activities at different stages of their academic career development, and their professionalization as professionals and experts. The way the model is built raises the highlights the benefits of scholarly growth when it occurs as processive growth, which consolidates over time through participation in diverse scholarship and research activities.

Answering three key questions will help teacher educators identify the contribution of the ISE Model and perhaps to adopt the directions it outlines:

Why is it important to get to know the ISE Model?

How should the ISE Model be used?

What is the contribution of the ISE Model to teacher educators?

Why is it important to get to know the ISE Model?

The ISE Model outlines numerous ways to professional growth including specialization in quality teaching, consolidating established professional conceptions and insights, mastering skills and required qualifications, and developing diverse professional

abilities. The model constitutes a tool to identify the development and growth stage in which individuals find themselves, present meaningful ways in which they can strive toward scholarly growth and progress to the next stages, and it constitutes a tool to understand challenges and benefits of the stage to which they strive to advance. Because ISE Model is built as a processive-linear development model, it demonstrates the nature of scholarly growth and how it expands in different working directions, using a range of platforms and work intensity at every growth stage. To focus on growth in the research world and research activity, the model demonstrates activities enabling faculty members to grow as expert researchers at various levels and dimensions of expertise.

How should the ISE Model be used?

The distinction between the ISE model stages is clearly dichotomous, but it is recommended it be viewed as a modular distinction. Growth steps are built layer after layer, with each layer founded on the previous one and a basis for continued growth. It is important to consider the notion that transitions between stages are not always visible, clear and/or unequivocal. In the daily reality of academic work, these transitions occur over time and naturally depend on researchers' personalities, their willingness to cooperate with others, to invest efforts and time, and to work to disseminate the fruits of their expertise as researchers (Avalos, 2011; Dinkelman, 2011; Murray, 2014).

The ISE Model stages can serve as milestones for faculty interested in examining their progress and achievements and exploring how it is worthwhile, desired and possible to continue to invest resources to continue growing.

What is the contribution of the ISE Model to teacher educators?

The ISE Model highlights the important place of teacher educators functioning both as excellent and quality lecturers and as expert researchers. This importance derives from the nature of their profession and being role models for the next generation of teachers. Teacher educators should prompt their students operate as teacher researchers themselves, exploring their professionalism, their work, their achievements, and the achievements of their pupils. Conduct as researchers carries a number of benefits: it allows examination from time to time of professional and academic progress, enables acquisition of diverse, innovative and up to date teaching and learning skills, enables acquisition of means and tools to meet the changing needs of students, allows them to be lecturers updated in the world of knowledge in their specialist teaching areas and the field of teacher education, it allows them to get to know profoundly the teaching and education systems, and enables teacher educators to be among those leading the body of knowledge in the teacher education profession and education systems and to make their mark.

References

- Attia, M., Edge, J., 2017. Be(com)ing a reflexive researcher: a developmental approach to research methodology. *Open Rev. Educ. Res.* 4 (1), 33–45.
- Avalos, B., 2011. Teacher professional development in teaching and teacher education over ten years. *Teach. Teach. Educ.* 27 (1), 10–20.
- Beaudry, J.S., Miller, L., 2016. *Research Literacy: A Primer for Understanding and Using Research*. The Guilford Press, New York, London.
- Berger, R., 2015. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qual. Res.* 15 (2), 219–234.
- Berry, A., 2009. Professional self-understanding as expertise in teaching about teaching. *Teach. Teach. Theor. Pract.* 15 (2), 305–318.
- Berry, A., Scheele, S., 2007. Constructing our professional knowledge as teacher educators: a collaborative self-study. In: *The International Study Association for Teachers and Teaching Biennial Meeting*. Brock University, Ontario, Canada, July 5–9.
- Boyd, P., Harris, K., Murray, J., 2007. *Becoming a Teacher Educator: Guidelines for the Induction of Newly Appointed Lecturers in Initial Teacher Education*. University of Bristol, Bristol.
- Bradley, J.H., Paul, R., Seeman, E., 2006. Analyzing the structure of expert knowledge. *Inf. Manag.* 43 (1), 77–91.
- Brand, G., 2015. Through the *looking glass space* to new ways of knowing: a personal research narrative. *Qual. Rep.* 20 (4), 516–525.
- Brew, A., 2010. Transforming academic practice through scholarship. *Int. J. Acad. Dev.* 15 (2), 105–116.
- Bridges, D., 1999. Teacher education and training: the location of professionalism. In: *The 3rd International Conference on Teacher Education*. Beit-Berl College, Israel.
- Brooks, R.L., 2005. Measuring university quality. *Rev. High. Educ.* 29 (1), 1–21.
- Cochran-Smith, M., 2005. Teacher educators as researchers: multiple perspectives. *Teach. Teach. Educ.* 21 (2), 219–225.
- Day, S., 2012. A reflexive lens: exploring dilemmas of qualitative methodology through the concept of reflexivity. *Qual. Sociol. Rev.* 8 (1), 60–85.
- Denemark, G.W., Espinoza, A., 1974. Educating teacher educators. *Theor. Pract.* 13 (3), 187–197.
- Dinkelman, T., 2011. Forming a teacher educator identity: uncertain standards, practice and relationships. *J. Educ. Teach.* 37 (3), 309–323.
- Geerdink, G., Boei, F., Willemse, M., Kools, Q., Vlokhoven Van, H., 2016. Fostering teacher educators' professional development in research and in supervising student teachers' research. *Teach. Teach. Theor. Pract.* 22 (8), 965–982.
- Griffiths, V., Thompson, S., Hryniewicz, L., 2014. Landmarks in the professional and academic development of mid-career teacher educators. *Eur. J. Teach. Educ.* 37 (1), 74–90.
- Hadar, L., Brody, D.L., 2017. Professional learning and development of teacher educators. In: Clandinin, D.J., Hsu, J. (Eds.), *Handbook of Research on Teacher Education*. Sage, UK, pp. 1049–1064.
- Halse, C., Deane, E., Hobson, J., Jones, G., 2007. The research-teaching nexus: what do national teaching awards tell us? *Stud. High Educ.* 32 (6), 727–746.
- Haviland, D., Shin, S.H., Turley, S., 2010. Now I'm ready: the impact of a professional development initiative on faculty concerns with program. *Innovat. High. Educ.* 35 (4), 261–275.
- Hökkä, P., Etelapelto, A., Rasku-Puttonen, H., 2012. The professional agency of teacher educators amid academic discourses. *J. Educ. Teach.* 38 (1), 83–102.
- Jarvis, J., Dickerson, C., Stockwell, L., 2013. Staff-student partnership in practice in higher education: the impact on learning and teaching. *Proced. Soc. Behav. Sci.* 90, 220–225.
- Kinchin, I.M., Cabot, L.B., Hay, D.B., 2008. Visualising expertise: towards an authentic pedagogy for higher education. *Teach. High. Educ.* 13 (3), 315–326.
- Korthagen, F.A.J., 2014. Promoting core reflection in teacher education: deepening professional growth. In: Craig, C.J., Orland-Barak, L. (Eds.), *International Teacher Education: Promising Pedagogies*. Advances in Research on Teaching, Emerald Group, West Yorkshire, pp. 73–89.

- Korthagen, F.A.J., Loughran, J., Lunenberg, M., 2005. Teaching teachers - studies into the expertise of teacher educators: an introduction to this theme issue. *Teach. Teach. Educ.* 21 (2), 107–115.
- Koster, B., Brekelmans, M., Korthagen, F.A.J., Wubbels, T., 2005. Quality requirements for teacher educators. *Teach. Teach. Educ.* 21 (2), 157–176.
- Koster, B., Dengerink, J.J., Korthagen, F.A.J., Lunenberg, M., 2008. Teacher educators working on their own professional development: goals, activities and outcomes of a project for the professional development of teacher educators. *Teach. Teach. Theor. Pract.* 14 (5–6), 567–587.
- Loughran, J., 2006. *Developing a Pedagogy of Teacher Education: Understanding Teaching and Learning about Teaching*. Routledge, New York.
- Loughran, J., 2014. Professionally developing as a teacher educator. *J. Teach. Educ.* 65 (4), 271–283.
- Loughran, J., Berry, A., 2005. Modelling by teacher educators. *Teach. Teach. Educ.* 21 (2), 193–203.
- Lunenberg, M., Dengerink, J.J., Korthagen, F.A.J., 2014. *The Professional Teacher Educator Roles, Behaviour, and Professional Development of Teacher Educators*. Sense, Rotterdam.
- Lunenberg, M., Hamilton, M.L., 2008. Threading a golden chain: an attempt to find our identities as teacher educators. *Teach. Educ. Q.* 35 (1), 185–205.
- Lunenberg, M., Korthagen, F.A.J., Swennen, A., 2007. The teacher educator as a role model. *Teach. Teach. Educ.* 23 (5), 586–601.
- Lunnenberg, M., 2010. Characteristics, scholarship and research of teacher educators [liner notes]. In: *International Encyclopedia of Education*. Elsevier Ltd.
- McKeon, F., Harrison, J., 2010. Developing pedagogical practice and professional identities of beginning teacher educators. *Prof. Dev. Educ.* 36 (1–2), 25–44.
- Murray, J., 2008. Teacher educators' induction into higher education: work-based learning in the micro communities of teacher education. *Eur. J. Teach. Educ.* 31 (2), 117–133.
- Murray, J., 2010. Towards a new language of scholarship in teacher educators' professional learning? *Prof. Dev. Educ.* 36 (1–2), 197–209.
- Murray, J., 2014. Teacher educators' constructions of professionalism: a case study. *Asia Pac. J. Teach. Educ.* 42 (1), 7–21.
- Murray, J., Male, T., 2005. Becoming a teacher educator: evidence from the field. *Teach. Teach. Educ.* 21 (2), 125–142.
- Murray, J., Swennen, A., Shagrir, L., 2009. Understanding teacher educators' work and identities. In: Swennen, A., Klink, M.V.D. (Eds.), *Becoming a Teacher Educator: Theory and Practice for Novice Teacher Educators*. Springer, New York, pp. 29–43.
- Palaganas, E.C., Caricativo, R.D., Sanchez, M.C., Molintas, M.V.P., 2017. Reflexivity in qualitative research: a journey of learning. *Qual. Rep.* 22 (2), 426–438.
- Rosser, J.V., Johnsrud, K.L., 2003. Academic deans and directors: assessing their effectiveness from individual and institutional perspectives. *J. High Educ.* 74 (1), 1–25.
- Rubin, O., Zadik, R., 2015. The identity of the teacher-educator researcher: the teacher educators' perspective. *Dapim J. Stud. Res. Educ.* 60, 36–58.
- Russell, T., 2010. Improving the quality of teaching and learning: lessons learned as a teacher educator. In: *Australian Teacher Education Association 2010 Annual Conference: Teacher Education for a Sustainable Future*. Australian Teacher Education Association (ATEA), Townsville, Australia.
- Russell, T., 2018. A teacher educator's lessons learned from reflective practice. *Eur. J. Teach. Educ.* 41 (1), 4–14.
- Shagrir, L., 2010a. Professional development of novice teacher educators: professional self, interpersonal relations and teaching skills. *Prof. Dev. Educ.* 36 (1–2), 45–60.
- Shagrir, L., 2010b. Professional development of novice teacher educators: professional self, interpersonal relations and teaching skills. In: Bates, T., Swennen, A., Jones, K. (Eds.), *The Professional Development of Teacher Educators*. Taylor & Francis, London, pp. 55–70.
- Shagrir, L., 2011. Professional development of the teacher educator: orientations and motivations. *Int. J. Univ. Teach. Faculty Dev.* 18 (1), 17–32.
- Shagrir, L., 2013a. Teacher educators and the practical component in teacher education. *J. Educ. Pract.* 4 (27), 172–184.
- Shagrir, L., 2013b. Three professional development patterns among faculty in higher education. *Int. J. Univ. Teach. Facul. Dev.* 4 (2), 1–13.
- Shagrir, L., 2015a. Factors affecting the professional characteristics of teacher educators in Israel and in the USA: a comparison of two models. *Compare* 45 (2), 206–225.
- Shagrir, L., 2015b. Working with students in higher education – professional conceptions of teacher educators. *Teach. High. Educ.* 20 (8), 783–794.
- Shagrir, L., 2017a. Collaborating with colleagues for the sake of academic and professional development in higher education. *Int. J. Acad. Dev.* 22 (4), 331–342.
- Shagrir, L., 2017b. The professional characteristics and nature of professional development of experts. In: Alegre de la Rosa, O.M. (Ed.), *Research on University Teaching and Faculty Development: International Perspectives*. Nova Science Publishers, New York, pp. 139–153.
- Shagrir, L., 2017c. Teacher educators' professional development - motivators and delays. In: Boyd, P., Szplit, A. (Eds.), *Teachers and Teacher Educators Learning through Inquiry: International Perspectives*. Wydawnictwo Attyka, Kielce-Krakow, pp. 159–180.
- Shagrir, L., 2020. Three-phase model of scholarly growth in teacher education. *Eur. J. Teach. Educ.* 44 (2), 271–291.
- Shulman, L.S., 1998. *Teaching and Teacher Education Among the Professions*. American Association of Colleges for Teacher Education (AACTE), New Orleans, Louisiana.
- Swennen, A., Geerdink, G., Volman, M., 2017. Developing a researcher identity as teacher educator. In: Boyd, P., Szplit, A. (Eds.), *Teachers and Teacher Educators Learning through Inquiry: International Perspectives*. Wydawnictwo Attyka, Kielce-Krakow, pp. 143–157.
- Swennen, A., Shagrir, L., Cooper, M., 2009. Becoming a teacher educator: voices of beginning teacher educators. In: Swennen, A., Klink, M.v. d. (Eds.), *Becoming a Teacher Educator: Theory and Practice for Novice Teacher Educators*. Springer, New York, pp. 91–102.
- Swennen, A., van der Klink, M. (Eds.), 2009. *Becoming a Teacher Educator: Theory and Practice for Novice Teacher Educators*. Springer, New York.
- Swennen, A., Volman, M., van Essen, M., 2008. The development of the professional identity of two teacher educators in the context of Dutch teacher education. *Eur. J. Teach. Educ.* 31 (2), 169–184.
- Tack, H., Vanderlinde, R., 2014. Teacher educators' professional development: towards a typology of teacher educators' researcherly disposition. *Br. J. Educ. Stud.* 62 (3), 297–315.
- Tack, H., Vanderlinde, R., 2016. Measuring teacher educators' researcherly disposition: item development and scale construction. *Vocat. Learn.* 9 (1), 43–62.
- van Velzen, C.P., van der Klink, M., Swennen, A., Yaffe, E., 2010. The induction and needs of beginning teacher educators. *Prof. Dev. Educ.* 36 (1&2), 61–75.
- Williams, J., Ritter, J.K., 2010. Constructing new professional identities through self-study: from teacher to teacher educator. *Prof. Dev. Educ.* 36 (1–2), 77–92.
- Young, R.J., Erickson, B.L., 2011. Imagining, becoming, and being a teacher: how professional history mediates teacher educator identity. *Stud. Teach. Educ.* 7 (2), 121–129.
- Zeichner, K., 2005. Becoming a teacher educator: a personal perspective. *Teach. Teach. Educ.* 21 (2), 117–124.

Questioning the consensus on effective professional development

Jennifer Gore, Sally Patfield, and Leanne Fray, Teachers and Teaching Research Centre, School of Education, University of Newcastle, Callaghan, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	511
Investing in effective professional development	511
What is effective professional development?	512
Interrogating the evidence for effective professional development	512
Interrogating the primacy of curriculum content in effective PD	513
Interrogating the role of coaching in effective PD	514
Where to next?	515
References	516

Introduction

Continuing professional development (PD) lies at the heart of efforts to maintain and improve teachers' knowledge and practice once they enter the workforce, complementing and extending their initial teacher education and induction experiences. Across the OECD, participation in PD is near universal, with 94% of teachers and 99% of principals, on average, engaging in at least some form of ongoing professional development (OECD, 2019). Indeed, PD is mandatory in many jurisdictions, with teachers required to undertake a minimum number of hours in a designated period if they are to retain registration, renew their teaching license, or apply for promotion. Increasingly, PD is considered essential to, even inseparable from, school improvement and large-scale educational reform (Garet et al., 2001; King, 2014) and a key mechanism in improving student outcomes (Yoon et al., 2007; Darling-Hammond et al., 2017). As a result, teachers are explicitly positioned not only as lifelong learners, but also as fundamental enactors of education policy and change (OECD, 2019).

Traditionally, PD has been conceptualized as teachers learning new knowledge and skills to put into practice upon returning to their classrooms. The key measure of its success stemmed from self-reported teacher satisfaction (Desimone, 2011; Hill et al., 2013). More recently, the rising accountability policy climate has spurred on efforts to directly link PD to measurable student impact, using empirical evidence gathered in classrooms and schools. Arguably, this newer expectation of PD is grounded in a process-product view of causality (Opfer and Pedder, 2011) linking engagement in "effective" PD with improvements in instructional practice and, by extension, student learning and achievement gains. In this way, what is "effective" has become synonymous with measurable impact, and substantial research attention is now directed toward understanding how and why PD "works." In the current moment, high quality PD programs are increasingly seen as those that significantly improve student outcomes, in addition to—and as a direct result of—building teachers' capacity for effective teaching.

Investing in effective professional development

This shift in focus has been accompanied by a widespread view among scholars and policymakers that governments should be investing, and encouraging participation, in "effective PD" (Garet et al., 2001; Hill, 2009; Desimone, 2009). Given that PD is estimated to cost billions of dollars annually (Birman et al., 2007; Wei et al., 2009), it makes sense to direct money toward activities that generate decent return on investment. However, the true costs of PD are likely to be vastly underestimated, given difficulties in capturing consistent measures of expenditure that account for the full range of funding costs and sources, as well as the lack of rigorous, comparable data across jurisdictions (Hawley Miles et al., 2004; Hill, 2009). For example, the budget for PD in one US school district was initially reported to be around USD \$460,000 but, after closer scrutiny, later calculated to be USD \$8.9 million (Miles and Hornbeck, 2000). Such discrepancies occur because the substantial teacher and school time required for PD is not fully acknowledged—such as regularly funding teacher release and the costs associated with employing coaches, mentors, and curriculum specialists (Jacob and McGovern, 2015).

Directing investment toward impactful forms of PD is hampered by the fact that, in many countries, continuing professional development for teachers is essentially a "marketplace" (Hill, 2009). A wide array of competing providers vie for money, time, and attention, but not all deliver high quality programs. Indeed, much PD has been described as "intellectually superficial," "woefully inadequate" (Borko, 2004) and "uninspired" (Hill, 2009); criticisms that continue to be made even after the shift away from the dominance of episodic, individualistic, and "one-shot" forms of PD traditionally associated with in-service staff development (Borko et al., 2010; Avalos, 2011). Furthermore, PD is often conducted without any formal evaluation, so that there is little rigorous evidence to demonstrate success (or failure) in improving teachers' knowledge and practice, let alone student outcomes. As Hill (2009) succinctly puts it, "the challenge [now] is how to improve the quality, not the quantity, of what is available" (p. 473).

Not surprisingly then, in the current PD marketplace, school systems tend to see meagre returns on investment. This point has been starkly illustrated in a well-known report from the education non-profit organization, The New Teacher Project (TNTP), which invokes the analogy of a “mirage” in arguing that we are yet to develop a sound evidence base for directing money toward effective PD, particularly at scale (Jacob and McGovern, 2015). Based on an examination of three public school districts in the US, Jacob and McGovern (2015) found that the vast majority of teachers did not significantly improve their practice despite a sizable investment of USD \$18,000, on average, per teacher, per year; representing 6–9% of a district’s annual operating budget. Rather, approximately five out of every 10 teachers in the study maintained their current level of instructional performance over several years while teaching quality actually declined for two of every 10 teachers.

Governments and policymakers across the globe want to know how to strengthen the capacities of teachers and achieve substantial impact on student learning. Given that rigorous research into effective PD is relatively scant, researchers now have even more reason for investigating the effects of PD on teachers and students. Meanwhile, however, an informal consensus has emerged concerning key elements or design features that are thought to underpin, or even maximize, effectiveness. Our aim in this article is to interrogate that consensus. In so doing, we raise questions about rigor and methodology, highlight recent advances, and identify important future directions in continuing to build knowledge in the field.

What is effective professional development?

In an attempt to identify best practice, the phrase “effective PD” has become a mainstay in the literature on teacher PD. While “effectiveness” can have many meanings, it is typically used to refer to PD activities associated with positive changes in teacher knowledge and practice, and/or positive changes in student outcomes. More often it is the former, rather than the latter, that is the topic of investigation, or at least the subject of published results.

The primary focus of researchers and policymakers has been to identify the key design features or characteristics of PD programs that have shown success, usually derived from conducting literature reviews, synthesizing empirical data and/or reviewing policy documents. By and large, this approach is grounded in a belief that there is a core set of characteristics that work to make PD effective, particularly in terms of its structural and logistical dimensions. Typically expressed as lists, these articulations of core characteristics arguably amount to a new consensus, emerging among scholars, about specific factors that are crucial to PD’s effectiveness (Penuel et al., 2007; Desimone, 2009).

In general, the consensus revolves around the following features: a focus on curriculum content, including discipline-specific content knowledge and pedagogy; active learning; collaboration between teachers across grade levels, schools, or districts; a sustained duration over an extended period of time; modeling of best practice; the use of coaches or expert support; opportunities for feedback and reflection; and coherence with school and teacher priorities, so that PD is aligned with broader policies and/or reform efforts (Garet et al., 2001; Desimone et al., 2002; Desimone, 2009, 2011; Darling-Hammond et al., 2017). Importantly, these features reflect the shift away from earlier didactic approaches, long condemned for being intellectually superficial, top-down, rigid, disconnected from classroom practice and school context, and providing little to no room for teachers to exert any control over their own learning (Ball and Cohen, 1999; Wilson and Berne, 1999).

The salience of this consensus as a new orthodoxy is manifest in arguments made by a number of prominent scholars. It has been claimed, for example, that these are *the* features of PD worth testing (Wayne et al., 2008), and that we currently have enough empirical evidence so that future studies ought to narrow their focus to these specific elements in order to better understand their relative importance (Desimone, 2009). Further, that these features should form the basis of a shared conceptual framework for studying PD in order to move the field forward (Desimone, 2009). Such a plan would certainly provide a clear basis for what to measure and subsequently attain a degree of consistency across studies, including those conducted in different jurisdictions and contexts. But it may also limit new ways of thinking about powerful PD. And what if the evidence on effectiveness is itself questionable or less than ideally rigorous?

Interrogating the evidence for effective professional development

Despite the new orthodoxy, PD is yet to have the kind of impact sought by policymakers, researchers and teachers themselves. Indeed, the consensus on core features of effective PD has not escaped criticism, which calls into question whether it is productive to take these lists as a starting point for conceptualizing and measuring effective PD. We highlight five main concerns.

First, the evidence base that underpins the consensus is relatively weak. Soon after articulations of effective PD began to appear, doubts were raised about the criteria on which effectiveness had been judged. In particular, Guskey (2003) identified two methodological issues that compromised the rigor of contributing studies: (1) many of these lists were simply derived from self-reported survey data gathered from teachers about changes in their knowledge and instructional practice; and, (2) several of the lists had been formulated based on the views of PD researchers following scans of the literature. Rarely were the key elements of effective PD derived from empirical research that rigorously established a relationship between PD and gains in student achievement. As Guskey (2003) succinctly put it at the time, “the best that can be said, therefore, is that researchers and educators *generally favor* these characteristics and *believe* they are important, despite a lack of verifying evidence” (p. 9; emphasis added). And yet, the very same design features have continued to underpin articulations of effective PD some 15 years later (Darling-Hammond et al., 2017).

Second, rigorous experimental and quasi-experimental designs in PD studies challenge the consensus on effective PD. A review of more than 1300 studies claiming to demonstrate the effect of PD on student achievement found that only nine studies actually met rigorous evidence standards according to the US Department of Education's *What Works Clearinghouse* (Yoon et al., 2007). Given this tiny proportion of acceptable studies, no clear patterns were discernible in terms of key design elements. The PD in these successful studies varied by content (including curriculum content and pedagogy), contact hours (from as few as 5 h to 100 h), and duration (from four weeks to one year). These studies were also typically derived from a single PD activity in a single context, and delivered by the researchers or their associates—common methodological weaknesses in the broader field of teacher PD. Given that PD is often locally developed and research is often undertaken by the developer of the PD program (Borko, 2004; Wayne et al., 2008; Hill, 2009), such studies typically fail to address how the local success of a PD intervention might translate into effectiveness in other broader contexts.

Third, the consensus tends to gloss over the fact PD which includes the desired design features is often unsuccessful at improving student outcomes. This point is highlighted in the results from three large-scale randomized controlled trials (RCT) conducted by the Institute of Education Sciences in the US (Garet et al., 2016). The PD under examination included the design elements that have come to be associated with effective PD, namely: a focus on curriculum content and content-specific pedagogical knowledge; opportunities for collaboration; sustained duration over the course of a school year; the use of experts to lead the intervention workshops; and, in-school coaching throughout the year involving modeling and feedback (Garet et al., 2008, 2010, 2016). While, on average, the PD in these studies produced a significant positive impact on teachers' knowledge and instructional practice, it did not translate into improvements in student achievement (Garet et al., 2016).

Fourth, the consensus lacks clarity of focus (Guskey, 2003; Wayne et al., 2008). If, for example, a focus on curriculum content is critical to effective PD, which curriculum design features matter most? What exactly should be the focus? Studies demonstrating a positive link to student outcomes are wide-ranging and discipline-specific, including a focus on student learning and assessment in mathematics (Carpenter et al., 1989), "learning cycles" in science (Marek and Methven, 1991), and instructional routines in reading (Mcgill-Franzen et al., 1999). Similarly, if collaboration among teachers is important, what method is most effective? There are many options here, including communities of practice (Gallagher et al., 2017), shared lesson planning (Gersten et al., 2010), or interactive forums (Shaha and Ellsworth, 2013). And, if it is recommended that coaches are used, what should their engagement look like? A range of possibilities have been suggested, such as onsite or remote observations (Powell et al., 2010), coaching sessions involving reflection and questioning (Allen et al., 2011), and the analysis of videoed lesson segments (Allen et al., 2015). Such questions apply to every one of the key features that are said to be essential for the design and implementation of "effective PD".

Finally, most PD programs incorporate a host of features and design elements—what Hill (2009) refers to as a "package"—making it difficult to discern which elements actually contribute to success (or failure) and how. This challenge is exacerbated because the PD package also extends to the theory of teacher change and theory of instruction embedded within the design of the program (Wayne et al., 2008), core components of PD which are often disregarded when only the design elements are privileged in the new orthodoxy. How then might governments and school systems ensure investment in PD is directed appropriately?

The complexities outlined above signal an urgent need to keep interrogating the current orthodoxy, rather than simply continuing efforts to refine features of PD which might have gained undeserved or unwarranted status. In the remainder of this article, we contribute to such an examination by problematizing two of the most widely cited and accepted design elements which consistently appear in lists of effective PD: the centrality of curriculum content and the need for coaching. Through this analysis, we aim to demonstrate why it is important to think differently about what constitutes effective PD and raise questions about the new orthodoxy.

Interrogating the primacy of curriculum content in effective PD

A focus on curriculum content is a central plank of the new orthodoxy, a design feature which appears in almost all articulations of effective PD. Indeed, it has been proposed that discipline-specific content knowledge and/or pedagogies is the "most influential" feature of successful PD programs (Desimone, 2009) and even the "most relevant and striking" characteristic associated with success (van Veen et al., 2012). The premise here is that increasing teachers' knowledge of their subject and how to teach it is key to improving their practice, thus helping to achieve positive gains in students' subject-specific outcomes. This view is similarly evident in recent TALIS data indicating that teachers believe content-driven PD to be one of the most impactful forms of PD they have experienced (OECD, 2019). Content-focused PD is also highly attended by teachers, with approximately 75% of teachers on average across the OECD reporting that they have participated in PD focused on understanding their subject or discipline, and 72% of teachers reporting that their PD activities have covered subject-specific pedagogical competencies (OECD, 2019). These figures are not surprising given the importance of subject-matter content and pedagogical content knowledge (Shulman, 1986) to teaching, and teachers.

However, despite this perceived impact on practice, demand for content-driven PD is relatively low. In fact, the same TALIS data demonstrates that very few teachers report needing this form of PD, instead wanting more pedagogy-focused PD, such as teaching students with special needs, using ICT in the classroom, or teaching in a multicultural or multilingual setting (OECD, 2019). On the one hand, this inverse relationship may simply be a function of supply and demand; teachers might not report a desire for content-focused PD if it is already provided on a regular basis. On the other hand, the mismatch between supply and demand potentially signals teachers' need for different forms of PD not directly connected to a specific subject or discipline. Plausibly, the high degree of impact believed to be associated with content-focused PD might therefore be due, at least to some degree, to provision.

The very nature of content-focused PD also means that the PD activity narrowly targets a specific group of teachers often working in a single discipline. Most PD interventions shown to have been effective—resulting in changes in teachers’ practices and student learning outcomes—are in the disciplines of mathematics, science, or English/language arts (Darling-Hammond et al., 2017; Yoon et al., 2007). While academic achievement in these disciplines is often easier to measure than in other subjects due to the availability of standardized tests, the dominance of these content-focused interventions essentially reflects the hegemony of “core subjects” in many state and national curricula.

With content-focused PD dominating the literature on effective PD, much less has been made about the value of pedagogy-focused PD. This relative inattention is surprising given the emergence and growing popularity of many “generic” observational instruments used to measure effective teaching, such as the Framework for Teaching (FFT) and the Classroom Assessment Scoring System (CLASS). While FFT focuses on the effectiveness of teachers in a wide range of areas, including planning and preparation, the classroom environment, instruction, and professional responsibilities (Danielson, 2007), CLASS looks more specifically at daily teacher–student interactions, particularly in terms of emotional support, instructional support, and classroom organization (Pianta et al., 2008). As such, both frameworks can be—and are—used *across* subjects and disciplines.

Research on these “generic” observational instruments, which can be applied in all teaching contexts, is promising in terms of pinpointing “what works” to foster student achievement. For example, the large-scale Measures of Effective Teaching (MET) project found that both FFT and CLASS are just as correlated to increases in student outcomes in mathematics and English/language arts as subject-specific observational instruments (Bill and Melinda Gates Foundation, 2012). Although the MET project focused on observation rather than professional development, this finding suggests that teachers can use a generic pedagogical model to examine—and improve—their practice in different subjects, rather than having to use a subject-specific instrument aligned with the curriculum area being taught.

Rigorous studies of pedagogy-focused PD have also demonstrated gains in student achievement, further calling into question whether a focus on discipline-specific content is essential to effective PD. Evidence from Yoon et al.’s (2007) early review of the field identifies two generic forms of PD that show positive effects on student outcomes under RCT conditions. Both forms of PD focus on teaching behaviors and instructional strategies applicable to all subjects (Cole, 1992; Sloan, 1993). More recently, our own research on Quality Teaching Rounds (QTR)—a pedagogy-focused approach to PD that combines the benefits of teaching rounds and professional learning communities with a comprehensive pedagogical framework—has similarly demonstrated noteworthy improvements in student achievement. Using a cluster RCT design, students of teachers in the QTR group made 25% more progress in mathematics than the control group, or approximately two months additional growth over an 8-month period (Gore et al., 2021). This form of PD has also shown significant positive effects on the quality of teaching, teacher morale, and sense of recognition under RCT conditions (Gore et al., 2017), illustrating that the impact of pedagogy-focused PD can be significant for both students and teachers.

Pedagogy-focused PD has also been shown to have additional benefits. The capacity to collaborate is often constrained when participating in content-focused PD, because teachers usually work with colleagues from the same subject or discipline. By contrast, pedagogy-focused PD, such as the QTR approach, has shown the powerful benefits of working with, and observing, teachers across subject boundaries (Gore and Rosser, 2020). Participating in QTR PD not only enhanced collegiality across grades and subjects, but it helped teachers refine their practice by generating new insights about their pedagogy and their students—in large part because they were working with *diverse* rather than *similar* colleagues from across the whole school (Gore and Rosser, 2020).

Overall, we argue that the potential of pedagogy-focused PD—for both students and teachers—has largely been overlooked in literature on effective PD. Instead, most research has focused on content-focused PD, producing more published results on the effects of content-focused PD, leading to further research, thus embedding content-focused PD as core to the consensus. Our brief analysis here begins to problematize whether a focus on curriculum content is necessary for effective PD.

Interrogating the role of coaching in effective PD

A second PD design feature common to almost all articulations of effective PD is the need for coaching, a practice that has dramatically increased in the past two decades (Mangin, 2009; Biancarosa et al., 2010; Coburn and Woulfin, 2012; Jacobs et al., 2017). Coaching is underpinned by two key premises. First, that teachers who take part in coaching activities become more adept with planned, intentional practice; and second, by breaking teaching into its parts, and practicing those parts, coaching can be a powerful means of achieving professional learning (Gore, 2020).

The increasing acceptance of coaching as a component of effective PD is reflected in the many forms of coaching that currently exist (Deussen et al., 2007), including peer coaching (Joyce and Showers, 1980, 1982), cognitive coaching (Costa and Garmston, 2002), and instructional coaching (Desimone and Pak, 2017; Woulfin and Rigby, 2017). The role of coaches varies significantly based on the requirements of policymakers, departments of education and school leaders. Coaches can work with teachers one-on-one or in a group, to assist with lesson planning and assessment (West and Staub, 2003), student diagnostics (Vogt and Rogalla, 2009) and identifying goals for teacher learning (Borko et al., 2014). Coaches engage with teachers to foster active learning (Kretlow and Bartholomew, 2010) with the most common form being one-on-one communication between coach and teacher (Killion, 2012). In some contexts, coaches also provide continuing PD, lead professional learning communities, co-teach in classrooms and work with small groups of children (Lesley et al., 2020).

Despite the prevalence of coaching and widespread support for it as a key component of effective PD, four main areas of concern arise. First, critics argue that there is a lack of empirical evidence on why, when and how coaching works (Desimone and Pak, 2017), with effects on student outcomes remaining largely unsubstantiated (Mangin, 2009). For example, Neuman and Cunningham (2009) found that a combination of coaching and coursework was required to improve the quality of practice in schools, but they did not provide evidence of impact on student outcomes. In a review of effective PD, Darling-Hammond et al. (2017) found that 30 of the 35 studies included some form of coaching or expert advice. However, these results should be treated with caution. The review does not provide the total number of studies considered (beyond the 35 included in the review), provide an assessment of the quality of evidence provided in the reviewed articles, or consider if coaching was the key factor impacting student outcomes or whether it was a combination of factors (Hill et al., 2013). More recently, in a meta-analysis of 60 studies, Kraft et al. (2018) found that coaching does have positive effects on teacher instruction and student achievement, but the effectiveness of such programs diminishes significantly when attempting to take the programs to scale.

Second, coaching is expensive to implement and to scale. One of the greatest criticisms of coaching relates directly to the allocation of resources, particularly the time and monetary investment required for coaches to build trusting relationships (Knight, 2006). In particular, the financial investment in coaching can be up to 12 times the cost of other forms of professional development (Mangin, 2009). Likewise, inadequate funding can limit the effective implementation of coaching programs (Mangin, 2009).

Third, coaching relies heavily on outside “experts,” which can lead to teachers becoming reliant on those who know better, diminishing their confidence and leaving them uncertain of their skills and in need of sustained external endorsement of their teaching practices (Gore, 2020). Extreme forms of coaching, such as bug-in-ear feedback (Scheeler et al., 2006) can further undermine engagement with the multifaceted nature of teaching (Hargreaves and Dawe, 1990; Horn and Little, 2010).

Last, on a practical level, several factors can mitigate the success of coaching. There is considerable variability in coaching activities and their implementation (Atteberry and Bryk, 2011). Teachers can be unwilling to work with coaches to change their instructional methods (Deussen et al., 2007). Coaching can slip into supervisory practices aimed at supporting specific performance targets (Lofthouse and Leat, 2013). And, coaches are often classroom teachers who may not have the skills and training necessary to be effective in those roles (Gallucci et al., 2010; Poglinco and Bach, 2004).

Arguably, the focus on coaching as a key aspect of effective PD has obscured the impact of other programs which might exclude coaching but show demonstrable positive impacts on student outcomes. For example, a program of research examining the impact of a web-based, on-demand form of professional development on student achievement has consistently found positive effects, (Shaha and Ellsworth, 2013; Shaha et al., 2015; Shaha and Glassett, 2016) despite not including a coaching component in the design of the PD program. Other successful programs such as QTR are designed to foster peer support to do the “heavy lifting” of PD and can be enacted almost entirely by teachers without external input (Gore et al., 2017). While Kraft et al.’s (2018) meta-analysis found that coaching works, the positioning of coaching as the latest holy grail of effective PD may be ill-considered. The 60 studies in this metanalysis were primarily literacy coaching programs provided to teachers of pre-kindergarten and elementary school students in the United States, without wider applicability to other contexts or subject areas. In sum, we question whether the enduring consensus on coaching as a key feature of effective PD, like the centrality of content, might be obscuring the significant benefits of other approaches.

Where to next?

Our analysis of just two of the design features that constitute the new orthodoxy on effective PD demonstrates how important it is to keep raising questions and conducting research to explore what constitutes effectiveness. Despite the many benefits associated with structuring and framing research and practice in PD, we propose that the field has embraced the informal consensus prematurely and uncritically. This is not to undo or undermine the advances made. Rather, we call for scholarship and research, both conceptual and empirical in nature, that opens up the field with ongoing exploration of how to meet teachers’ professional learning needs. At the same time, we call for greater sophistication and rigor in how PD is conceived and researched.

The problem of PD not being effective—that is, not impacting positively on teachers or students—may have more to do with the nature of what is on offer than it does with particular design features (Gore and Rickards, 2020). Regardless of the specific design features utilized in a PD program, teachers may resist the PD for a host of reasons including: to regain agency and carve out their own learning trajectories (Achinstein and Ogawa, 2006; Flint et al., 2011; Musanti and Pence, 2010); because they are unwilling to take the attendant risks (such as threats to their experience, to power relations, and their perceived expertise) (Zimmerman, 2006); or, on principle, given a fundamental difference in educational beliefs (Achinstein and Ogawa, 2006). Indeed, many teachers are resistant even to newer, more collaborative forms of PD that directly align with the current consensus (Achinstein and Ogawa, 2006; Musanti and Pence, 2010). As such, we must think of PD as more than a set of design features. It needs to be positioned as part of a system, embedded within policies and school cultures and its success needs to be recognized as contingent on teachers’ willingness to actively engage.

At the same time, if we are to learn more about effective PD in ways that inform the investments of governments and systems, then we need to examine both the methodologies being used and the quality of evidence underpinning claims to effectiveness (Sims and Fletcher-Wood, 2020). This does not mean buying simplistically into accountability culture or only supporting experimental studies. Instead, as researchers and PD providers in the field, we contend that we have responsibility for the quality of our work as well as our greater responsibility to teachers and their students. Moving forward, we must therefore continue to pursue knowledge and interrogate current understandings and new orthodoxies.

In sum, the enduring focus on lists of key design features of effective PD may be obfuscating and restraining the field rather than moving it forward. More than seven years ago, Heather Hill and colleagues (Hill et al., 2013) argued that we had reached a crossroads in PD research, with an urgent need to provide “more rigorous evidence” (p. 476). While there are signs that the field is slowly moving toward ensuring higher quality evidence (see for example, Kraft et al., 2018), there is still much to be done to support our understanding of what makes PD effective. Such efforts, we argue, should include rigorous trialing and evaluation of PD, including forms that do not align with the current consensus. This requires thinking about new ways of creating, designing and implementing PD, rather than too-quickly subscribing to any particular orthodoxy. Ultimately, it means encouraging innovation, inviting the field to think beyond current practice, and inspiring the next generation of researchers to contribute fresh insights to the evolution of the field.

References

- Achinstein, B., Ogawa, R.T., 2006. (In) fidelity: what the resistance to new teachers reveals about professional principals and prescriptive education policies. *Harv. Educ. Rev.* 76 (1), 30–63.
- Allen, J.P., Pianta, R.C., Gregory, A., Mikami, A.Y., Lun, J., 2011. An interaction-based approach to enhancing secondary school instruction and student achievement. *Science* 333 (6045), 1034–1037.
- Allen, J.P., Hafen, C.A., Gregory, A.C., Mikami, A.Y., Pianta, R., 2015. Enhancing secondary school instruction and student achievement: replication and extension of the My Teaching Partner-Secondary intervention. *J. Res. Educ. Eff.* 8 (4), 475–489.
- Atteberry, A., Bryk, A.S., 2011. Analyzing teacher participation in literacy coaching activities. *Elem. Sch. J.* 112 (2), 356–382.
- Avalos, B., 2011. Teacher professional development in Teaching and Teacher Education over ten years. *Teach. Teach. Educ.* 27 (1), 10–20.
- Ball, D.L., Cohen, D.K., 1999. Developing practices, developing practitioners: toward a practice based theory of professional development. In: Sykes, G., Darling-Hammond, L. (Eds.), *Teaching as the Learning Profession: Handbook of Policy and Practice*. Jossey-Bass, San Francisco, CA, pp. 30–32.
- Biancarosa, G., Bryk, A.S., Dexter, E.R., 2010. Assessing the value-added effects of literacy collaborative professional development on student learning. *Elem. Sch. J.* 111 (1), 7–34.
- Bill & Melinda Gates Foundation, 2012. *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*. Bill & Melinda Gates Foundation, Seattle, WA.
- Birman, B.F., Le Floch, K.C., Klekotka, A., Ludwig, M., Taylor, J., Walters, K., Wayne, A., Yoon, K.S., Vernez, G., Garet, M.S., O'Day, J., 2007. *State and Local Implementation of the No Child Left Behind Act: Volume II—Teacher Quality Under NCLB: Interim Report*. U.S. Department of Education, Washington, DC.
- Borko, H., Jacobs, J.K., Koellner, K., 2010. Contemporary approaches to teacher professional development. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*. Elsevier, Oxford, pp. 548–556.
- Borko, H., Koellner, K., Jacobs, J., 2014. Examining novice teacher leaders' facilitation of mathematics professional development. *J. Math. Behav.* 33, 149–167.
- Borko, H., 2004. Professional development and teacher learning: mapping the terrain. *Educ. Res.* 33 (8), 3–15.
- Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C.P., Loef, M., 1989. Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *Am. Educ. Res. J.* 26 (4), 499–531.
- Coburn, C.E., Woulfin, S.L., 2012. Reading coaches and the relationship between policy and practice. *Read. Res. Q.* 47 (1), 5–30.
- Cole, D.C., 1992. The effects of a one-year staff development program on the achievement test scores of fourth-grade students. *Diss. Abstr. Int.* 53.
- Costa, A.L., Garmston, R.J., 2002. *Cognitive Coaching: A Foundation for Renaissance Schools*. Christopher-Gordon Publishers, Norwood, MA.
- Danielson, C., 2007. *Enhancing Professional Practice: A Framework for Teaching*, second ed. Association for Supervision and Curriculum Development, Alexandria, VA.
- Darling-Hammond, L., Hyler, M.E., Gardner, M., 2017. *Effective Teacher Professional Development (Report)*. Palo Alto, CA.
- Desimone, L.M., Pak, K., 2017. Instructional coaching as high-quality professional development. *Theor. Pract.* 56 (1), 3–12.
- Desimone, L.M., Porter, A.C., Garet, M.S., Yoon, K.S., Birman, B.F., 2002. Effects of professional development on teachers' instruction: results from a three-year longitudinal study. *Educ. Eval. Pol. Anal.* 24 (2), 81–112.
- Desimone, L.M., 2009. Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educ. Res.* 38 (3), 181–199.
- Desimone, L.M., 2011. A primer on effective professional development. *Phi Delta Kappan* 92 (6), 68–71.
- Deussen, T., Coskie, T., Robinson, L., Autio, E., 2007. “Coach” Can Mean Many Things: Five Categories of Literacy Coaches in Reading First (Issues & Answers Report, REL 2007–No. 005). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest, Washington, DC.
- Flint, A.S., Zisook, K., Fisher, T.R., 2011. Not a one-shot deal: generative professional development among experienced teachers. *Teach. Teach. Educ.* 27 (8), 1163–1169.
- Gallagher, H.A., Woodworth, K.R., Arshan, N.L., 2017. Impact of the national writing project's college-ready writers program in high-need rural districts. *J. Res. Educ. Eff.* 10, 570–595.
- Gallucci, C., Van Lare, M.D., Yoon, I.H., Boatright, B., 2010. Instructional coaching. *Am. Educ. Res. J.* 47 (4), 919–963.
- Garet, M.S., Porter, A.C., Desimone, L.M., Birman, B.F., Yoon, K.S., 2001. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38 (4), 915–945.
- Garet, M.S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H.S., Doolittle, F., Zhu, P., Szejnberg, L., 2008. *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement (NCEE 2008-4030)*. U.S. Department of Education, Institute of Education Sciences, Washington DC.
- Garet, M.S., Wayne, A.J., Stancavage, F., Taylor, J., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., Doolittle, F., Warner, E., 2010. *Middle School Mathematics Professional Development Impact Study Findings After the First Year of Implementation. Executive Summary. (NCEE 2010-4010)*. U.S. Department of Education, Institute of Education Sciences, Washington DC.
- Garet, M.S., Heppen, J.B., Walters, K., Smith, T.M., Yang, R., 2016. *NCEE Evaluation Brief. Does Content-Focused Professional Development Work? (NCEE2016-4010)*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC.
- Gersten, R., Dimino, J., Jayanthi, M., Kim, J.S., Santoro, L.E., 2010. Teacher study group: impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *Am. Educ. Res. J.* 47, 694–739.
- Gore, J.M., Rickards, B., 2020. Rejuvenating experienced teachers through quality teaching rounds professional development. *J. Educ. Change* 1–20. Advance online publication.
- Gore, J., Rosser, B., 2020. Beyond content-focused professional development: powerful professional learning through genuine learning communities across grades and subjects. *Prof. Dev. Educ.* Advance online publication.
- Gore, J.M., Lloyd, A., Smith, M., Bowe, J., Ellis, H., Lubans, D., 2017. Effects of professional development on the quality of teaching: results from a randomized controlled trial of Quality Teaching Rounds. *Teach. Teach. Educ.* 68, 99–113.
- Gore, J.M., Miller, A., Fray, L., Harris, J., Prieto, E., 2021. Improving student achievement through professional development: results from a randomized controlled trial of Quality Teaching Rounds. *Teach. Teach. Educ.*
- Gore, J.M., 2020. Why isn't this empowering: the discursive positioning of teachers in efforts to improve teaching. In: Brown, A., Wisby, E. (Eds.), *Knowledge, Policy and Practice in Education and the Struggle for Social Justice: Essays Inspired by the Work of Geoff Whitty*. UCL Press, London, pp. 199–216.

- Guskey, T.R., 2003. What makes professional development effective? *Phi Delta Kappan* 84 (10), 748–750.
- Hargreaves, A., Dawe, R., 1990. Paths of professional development: contrived collegiality, collaborative culture, and the case of peer coaching. *Teach. Teach. Educ.* 6 (3), 227–241.
- Hawley Miles, K., Odden, A., Fermanich, M., Archibald, S., Gallagher, A., 2004. Inside the black box of school district spending on professional development: lessons from comparing five urban districts'. *J. Educ. Finance* 30 (1), 1–26.
- Hill, H.C., Beisiegel, M., Jacob, R., 2013. Professional development research: consensus, crossroads, and challenges. *Educ. Res.* 42 (9), 476–487.
- Hill, H.C., 2009. Fixing teacher professional development. *Phi Delta Kappan* 470–476.
- Horn, I.S., Little, J.W., 2010. Attending to problems of practice: routines and resources for professional learning in teachers' workplace interactions. *Am. Educ. Res. J.* 47 (1), 181–217.
- Jacob, A., McGovern, K., 2015. *The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development*. TNTP, Brooklyn, NY.
- Jacobs, J.K., Boardman, A., Wang, C., 2017. Understanding teacher resistance to instructional coaching. *Prof. Dev. Educ.* 44 (5), 690–703.
- Joyce, B., Showers, B., 1980. Improving in-service training: the messages of research. *Educ. Leader* 37 (5), 379–385.
- Joyce, B., Showers, B., 1982. The coaching of teaching. *Educ. Leader* 4–10.
- Killion, J., 2012. Coaching in the K-12 context. In: Fletcher, S., Mullen, C.A. (Eds.), *SAGE Handbook of Mentoring and Coaching in Education*. Sage, London.
- King, F., 2014. Evaluating the impact of teacher professional development: an evidence-based framework. *Prof. Dev. Educ.* 40 (1), 89–111.
- Knight, J., 2006. Instructional coaching. *Am. Assoc. Sch. Admin.* 63 (4), 36–40.
- Kraft, M.A., Blazar, D., Hogan, D., 2018. The effect of teacher coaching on instruction and achievement: a meta-analysis of the causal evidence. *Rev. Educ. Res.* 88 (4), 547–588.
- Kretlow, A.G., Bartholomew, C.C., 2010. Using coaching to improve the fidelity of evidence-based practices: a review of studies. *Teach. Educ. Special Educ.* 33 (4), 279–299. <https://doi.org/10.1177/0888406410371643>.
- Lesley, M., Beach, W., Smit, J., 2020. "You can't put everything I've been through into professional development": transformative literacy coaching in an "underperforming" high school. *Lit. Res. Instr.* 1–23. Advance online publication.
- Lofthouse, R., Leat, D., 2013. An activity theory perspective on peer coaching. *Int. J. Mentor. Coach. Educ.* 2 (1), 8–20.
- Mangin, M.M., 2009. Literacy coach role implementation: how district context influences reform efforts. *Educ. Adm. Q.* 45 (5), 759–792.
- Marek, E.A., Methven, S.B., 1991. Effects of the learning cycle upon student and classroom teacher performance. *J. Res. Sci. Teach.* 28 (1), 41–53.
- Mcgill-Franzen, A., Allington, R.L., Yokoi, L., Brooks, G., 1999. Putting books in the classroom seems necessary but not sufficient. *J. Educ. Res.* 93 (2), 67–74.
- Miles, K.H., Hornbeck, M., 2000. Rethinking District Professional Development Spending to Support a District CSR Strategy (Resource Reallocation, Issue #3). New American Schools, Arlington, VA.
- Musanti, S.I., Pence, L., 2010. Collaboration and teacher development: unpacking resistance, constructing knowledge and navigating identities. *Teach. Educ. Q.* 73–89.
- Neuman, S.B., Cunningham, L., 2009. The impact of professional development and coaching on early language and literacy instructional practices. *Am. Educ. Res. J.* 46 (2), 532–566.
- OECD, 2019. *TALIS 2018 Results (Volume 1): Teachers and School Leaders as Lifelong Learners*. OECD Publishing, Paris, France.
- Opfer, D., Pedder, D., 2011. Conceptualizing teacher professional learning. *Rev. Educ. Res.* 81 (3), 376–407.
- Penueel, W.R., Fishman, B.J., Yamaguchi, R., Gallagher, L.P., 2007. What makes professional development effective? Strategies that foster curriculum implementation. *Am. Educ. Res. J.* 44 (4), 921–958.
- Pianta, R.C., La Paro, K.M., Hamre, B.K., 2008. *Classroom Assessment Scoring System: Manual K-3*. Paul H Brookes Publishing, Baltimore, MD.
- Poglinco, S.M., Bach, A.J., 2004. The heart of the matter: coaching as a vehicle for professional development. *Phi Delta Kappan* 85 (5), 398–400.
- Powell, D.R., Diamond, K.E., Burchinal, M.R., Koehler, M.J., 2010. Effects of an early literacy professional development intervention on head start teachers and children. *J. Educ. Psychol.* 102, 299–312.
- Scheeler, M.C., McAfee, J.K., Ruhl, K.L., Lee, D.L., 2006. Effects of corrective feedback delivered via wireless technology on preservice teacher performance and student behavior. *Teach. Educ. Spec. Educ.* 29 (1), 12–25.
- Shaha, S.H., Ellsworth, H., 2013. Predictors of success for professional development: linking student achievement to school and educator successes through on-demand, online professional learning. *J. Instr. Psychol.* 40, 19–26.
- Shaha, S., Glassett, K., 2016. Maximizing educator enhancement: aligned seminar and online professional development. *Contemp. Issues Educ. Res.* 9 (1), 15–22.
- Shaha, S.H., Glassett, K.F., Ellsworth, H., 2015. Long-term impact of on-demand professional development on student performance: a longitudinal multi-state study. *J. Int. Educ. Res.* 11 (1), 12–25.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14.
- Sims, S., Fletcher-Wood, H., 2020. Identifying the characteristics of effective teacher professional development: a critical review. *Sch. Eff. Sch. Improv.* 1–17. <https://doi.org/10.1080/09243453.2020.1772841>.
- Sloan, H.A., 1993. Direct instruction in fourth and fifth grade classrooms. *Diss. Abstr. Int.* 54.
- van Veen, K., Zwart, R., Meirink, J., 2012. What makes teacher professional development effective? In: Kooy, M., van Veen, K. (Eds.), *Teacher Learning that Matters: International Perspectives*. Routledge, London, pp. 3–21.
- Vogt, F., Rogalla, M., 2009. Developing adaptive teaching competency through coaching. *Teach. Teach. Educ.* 25, 1051–1060.
- Wayne, A.J., Yoon, K.S., Zhu, P., Cronen, S., Garet, M.S., 2008. Experimenting with teacher professional development: motives and methods. *Educ. Res.* 37 (8), 469–479.
- Wei, R.C., Darling-Hammond, L., Andree, A., Richardson, N., Orphanos, S., 2009. *Professional Learning in the Learning Profession: A Status Report on Teacher Development in the U.S. And Abroad*. National Staff Development Council, School Design Network, Dallas, TX.
- West, L., Staub, F.C., 2003. *Content-Focused Coaching. Transforming Mathematics Lessons*. Heinemann, Portsmouth, NH.
- Wilson, S.M., Berne, J., 1999. Teacher learning and the acquisition of professional knowledge: an examination of research on contemporary professional development. *Rev. Res. Educ.* 24, 173–209.
- Woulfin, S.L., Rigby, J.G., 2017. Coaching for coherence: how instructional coaches lead change in the evaluation era. *Educ. Res.* 46 (6), 323–328.
- Yoon, K.S., Duncan, T., Wen-Yu Lee, S., Scarloss, B., Shapley, K.L., 2007. Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement (REL2007-No.033). Institute for Education Sciences, US Department of Education, Southwest Regional Laboratory at Edvance Research Inc, Washington, DC.
- Zimmerman, J., 2006. Why some teachers resist change and what principals can do about it. *NASSP Bull.* 90 (3), 238–249.

Situating teacher learning in the mathematics classroom and everyday practice

Janette Bobis^a and Kristen Tripet^b, ^a University of Sydney, Sydney, NSW, Australia; and ^b Australian Academy of Science, Canberra, ACT, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	518
Teacher professional learning	518
Teacher learning as situated in the classroom	519
Teacher noticing	520
Exemplar from a classroom-based research project	521
Developing the learning trajectory	521
Implementing the learning sequence	521
Teaching episode 1—noticing students' use of the distributive property	522
Teaching episode 2—noticing students develop an understanding of the associative property	522
Teaching episode 3—noticing students' misconceptions with multi-digit numbers	523
Teacher learnings from noticing	524
Learning from teaching	524
Conclusion	526
References	526
Further Reading	527

Introduction

The common discourse surrounding teaching quality stresses the connections between quality teaching and on-going teacher learning (Avalos, 2011; Kennedy, 2016; Organization for Economic and Cooperation and Development [OECD], 2016) for the ultimate purpose of improving student learning. The literature on teacher learning predominantly associates teacher learning as a consequence of some type of formally structured professional development (OECD, 2014; Kwakman, 2003; Richter et al., 2019) but recognition of the need to ensure teacher learning is not only sustainable but relevant to their contexts has caused researchers to explore new approaches to teacher professional learning (Superfine, 2019). In a recent review of Australasian research on teacher professional learning and development, Bobis et al. (2020) noted increased attention on teacher noticing as a strategy for their professional learning. This growth in attention was regularly grounded in the importance of teachers learning from their everyday practices and the presumption that the support mechanisms surrounding such learning were more sustainable than that stemming from structured programs of professional development (Chan et al., 2018). Despite the growing recognition of teachers' in-situ learning and the role it can play in improving classroom practices in the research literature, the authors concluded that there was "an evidence void" surrounding "the ways in which teachers learn from their daily practice" (p. 139). Bobis et al. called for increased attention to research in this area to build our understanding of how we might improve the capacity of teachers to learn from their everyday practice.

In this paper we explore the emerging field of teacher noticing as a vehicle for on-going professional learning that is situated in everyday mathematics classrooms. Our aim is to build upon and add new insights to teachers' professional learning by proposing strategies and mechanisms to support teachers learn how to notice what matters most in the mathematics classroom, and how such learning can be used to improve their practice. First, we review relevant theory and literature on teacher professional learning spotlighting in-situ learning and merge it with recent insights from research on teacher noticing. We then draw upon an example from our own classroom-based research to elaborate and reflect upon important aspects of teacher noticing relating to our focus.

Teacher professional learning

The terms professional development and professional learning are often used interchangeably in the literature. Increasingly professional learning is recognized as the product of external and workplace-embedded professional development activities that assist teachers acquire new knowledge and help them improve their practices in ways that will benefit student learning (Bobis et al., 2020; Darling-Hammond et al., 2017). Therefore, professional learning is much broader than professional development, but may not always result from a teacher's involvement in such activities.

The benefits of teacher professional learning have been associated with enhanced student outcomes and teaching quality (Darling-Hammond et al., 2017), and the development of teacher knowledge, beliefs and practices (Beswick et al., 2016). There are also cases where studies of teacher professional development outcomes have shown little or no positive effects (Garet et al., 2010) even when the programs in question incorporated certain characteristics associated with high quality professional learning (Richter et al., 2019; Bobis et al., 2020). The reasons for these conflicting findings vary. Kennedy (2016) argues that it is not

necessarily the program characteristics that are important but the underlying theories of action that determine their success. Others, such as [Lindvall and Ryve \(2019\)](#), suggest that the various outcomes might derive from the fact that researchers in the field define and apply the characteristics associated with so-called “effective” programs of professional development differently. For example, Lindvall and Ryve conducted a systematic review of professional learning and development literature focusing on the feature of coherence—the degree to which a program of teacher learning aligned with other aspects of teachers’ work and contexts. The researchers found that the term was conceptualized in many different ways, with the result that studies claiming to investigate the same feature were in fact studying very different things and were therefore unsurprised that findings varied across studies. They reflected on the fact that teacher professional learning occurs in a wide range of real-world contexts and that research in the past often neglected to account for the complexity of teachers’ workplaces. This neglect has occurred despite an awareness that from the perspective of teachers, professional learning is considered more effective the closer it is connected to the workplace and situated in their daily practices ([Beswick et al., 2017](#); [Kwakman, 2003](#)).

Teacher learning as situated in the classroom

Inspired by both situated and cognitive perspectives of learning, we view teacher professional learning as an agentic, constructive and social process in which the opportunity for learning is embedded in everyday practices. [Clarke and Hollingsworth’s \(2002\)](#) interconnected model of teacher growth aligns with such a perspective of teacher professional learning. Characterized as four domains—personal, external, practice and consequences—it focuses on the various relationships between different aspects of professional learning and the multiple pathways that such learning can take (see [Fig. 1](#)). The model acknowledges teachers as active and reflective learners who are influenced by society and their learning environment. Teacher agency is emphasized by the importance placed on professional experimentation, with enactment and teacher reflection proposed as the essential connecting mechanisms for teacher learning. The centrality of knowledge and beliefs emphasizes the construction of teacher knowledge as both a consequence of, and stimulus for, reflective involvement in professional learning activities and practice. For example, depending on existing knowledge and prior beliefs, a teacher may value higher order reasoning and thus regularly provide her students with mathematical opportunities requiring them to justify their thinking. As a consequence, she is also more likely to notice such thinking (or lack of) in her students and seek out professional learning resources to experiment with in the classroom and, depending on the salient outcomes, modify her practices accordingly. More importantly, Clarke and Hollingsworth’s model conceptualizes the process of teacher learning as interconnected and multi-directional across domains. This implies that learning in both the personal domain and the domain of practice can conceivably emanate from the very act of teaching and not, like in other linear models of teacher learning (e.g., [Guskey, 1986](#)), as a precondition for that practice to occur.

Although the classroom might be considered in theory an ideal and powerful context for teachers to learn while on-the-job, it seems that in practice, few professional learning activities or research projects have focused on teacher learning from the act of teaching itself ([Kwakman, 2003](#); [Margolinas et al., 2005](#)). Recently, there has been growing recognition of the important role social processes involving collegial interactions, self-directed learning and reflective practice play in teacher learning, which has seen an increase in professional learning approaches that are focused on situating teacher learning in the classroom. Two such approaches are lesson study (e.g., [Choy, 2016](#); [Groves and Doig, 2016](#)) and teacher noticing ([Chan et al., 2018](#); [Hollingsworth and Clarke, 2017](#)).

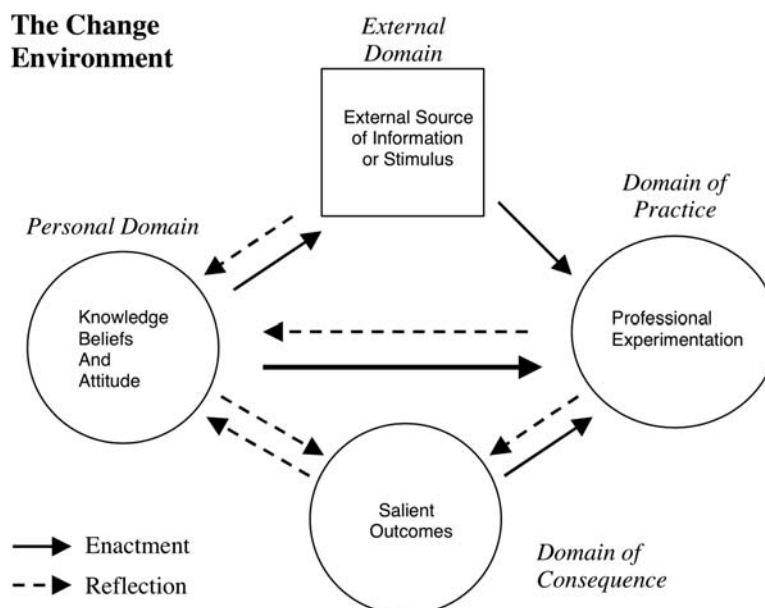


Fig. 1 Interconnected model of teacher professional growth. Reproduced with permission from Elsevier, [Clarke and Hollingsworth’s \(2002, p. 951\)](#).

Taking account of the growing body of literature surrounding teacher noticing, Chan et al. (2018) posited that when “teaching a lesson, teachers learn from the things to which they attend” (p. 92) and that a teacher’s attention is influenced by their prior knowledge and beliefs, the intentions of the lesson, and the context in which the lesson is enacted. Building on Clarke and Hollingsworth’s (2002) model of teacher professional growth, Chan et al. proposed a framework to conceptualize teacher learning as situated in their everyday practice (see Fig. 2).

Central to Chan et al.’s. (2018) conceptualization of situated learning is teacher selective attention—a major component of teacher professional noticing.

Teacher noticing

Noticing in the classroom is a common skill of teaching that has long been acknowledged, discussed and studied by mathematics educators and researchers for its potential to improve the quality of teaching and, thus, student learning (e.g., Franke et al., 1998; Mason, 2002; van Es and Sherin, 2008). It is also acknowledged that learning to notice the things that matter most in classrooms is complex, challenging and an important part of developing expertise as a teacher (Jacobs et al., 2010). It is this capacity to notice the important things in classrooms that sets apart an expert teacher from what a novice teacher or lay person might notice. This type of expert noticing is variously referred to in the literature as professional vision (van Es and Sherin, 2008), professional noticing (Jacobs et al., 2010) and intentional noticing (Mason, 2002). In this paper, we adopt Mason’s intentional noticing terminology to refer to the collection of practices teachers use to purposefully notice features of teaching and learning that are valued.

Central to the discipline of noticing is the necessity of paying attention to something (Mason, 2011). Research consistently confirms just how challenging it is for even experienced teachers to filter out what is worthy of their attention from the complex array of classroom events—particularly while in the very act of teaching (e.g., Jazby and Widjaja, 2019; van Es and Sherin, 2008). What teachers attend to in the classroom is influenced by their knowledge, beliefs and prior experiences of teaching. Therefore, noticing is a subjective process that is influenced not only by what one knows but also by what one believes to be important. For example, if a teacher believes that mathematics is a set of facts and procedures to be learned, they might pay attention to instances when students obtain correct answers to number facts and not attend to the point that some students used inefficient counting-by-ones strategies to do so. Moreover, research findings from the field of psychology (e.g., Simons, 2000) and mathematics education (e.g., Mason, 2011) tell us that people generally do not attend to events that they do not expect to see or if they “do” notice them, their response is either an inappropriate habitual response or appropriate but provided too late.

Simon (2007) argued that because teacher in-situ learning is limited by what they notice in the classroom, which itself is constrained by their conceptions and expectations, significant change in teachers is “unlikely to happen solely as a result of a teacher learning from her own teaching” (p. 141). To assist teachers focus their attention on what matters, researchers have suggested conceptualizations and frameworks to not only fine-tune their noticing practices (e.g., Jacobs et al., 2010; Mason, 2011) but to guide the selectivity of their attention-seeking gazes (e.g., Ivars et al., 2018). For instance, Ivars et al. combined Jacobs et al. conceptualization of teachers’ professional noticing of children’s mathematical thinking (attending, interpreting and deciding how to respond) with a hypothetical learning trajectory of part-whole fractions to guide preservice primary teachers’ noticing of students’ mathematical thinking. The researchers found that the learning trajectory helped focus the teachers’ attention on important milestones in students’ understandings of fractions.

Similarly, the process of intentional noticing is conceptualized by Mason (2011) as involving three interrelated practices: “preparing” to notice, “post-paring” by reflecting on recent events to select what one wants to notice, in order to “pare”, or to notice in the moment. Preparation to notice and respond are key activities in this process. During these preparation phases, Mason stresses the importance of pre-selecting the events likely to occur as a consequence of the stimuli provided by the teacher. Such stimuli could include the entire learning context but particularly the topic being studied, tasks, questions and resources used throughout the lesson. Pre-selecting which event teachers wish to focus upon means that they can pre-prepare what to notice. For instance, if a teacher pre-selects to focus upon students’ reasoning about, and strategies for mentally solving 2-digit by 1-digit multiplication problems, she might then prepare to notice using the following steps. She first reflects upon student responses to recent tasks on the same or a similar topic and draws upon her existing knowledge about student thinking strategies, possibly referring to an existing research-based learning trajectory to guide her pre-selection. She might also draw upon her curriculum and mathematical content knowledge. Based on this knowledge and experiences, she then anticipates three or four likely strategies her students will use and their explanations for using them.

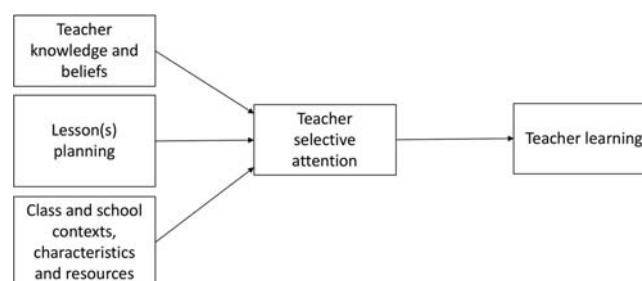


Fig. 2 Conceptualization of teacher learning from teaching. Modified from Chan et al. (2018, p. 92).

Next to each of these anticipated responses she prepares her own responses with the intention of eliciting deeper thinking in the students and providing appropriate prompts that will lead them to using more efficient mental strategies; consulting a hypothetical learning trajectory once again to provide greater directionality in her prompts for student strategy progression.

We agree with [Chan et al. \(2018\)](#) and consider teachers' capacities to select what to attend to as a critical determinant of not only their ability to notice what matters in the classroom but to learn from their everyday classroom experiences. Starting with Chan et al.'s conceptual framework ([Fig. 1](#)), we add the practice of anticipation ([Stein et al., 2008](#)) and combine it with [Mason's \(2011\)](#) approach of pre-paring a collection of appropriate teacher responses. Using this collection of practices to purposefully plan a lesson will reduce the likelihood of a teacher responding habitually to an event or with an inappropriate response such as simply "telling" students answers. To further assist teachers learning from these everyday classroom events, we recommend teachers reflect post-lesson upon what "surprised" them during each lesson—What happened that they did not anticipate and how will this new information impact what they do differently (or keep the same) in the future. Such reflective prompts might bring to the surface a new and important experience for a teacher to learn from and improve their instruction that would otherwise have gone unnoticed. This type of in-situ learning is only enacted and observable when a subsequent lesson is employed to consolidate the learning of students and the teacher.

In the next section we present an example of how one elementary teacher combined the principles and practices of intentional noticing previously outlined with a hypothetical learning trajectory for multi-digit multiplication. The study was conducted as part of Tripet's larger doctoral project with the dual aims of exploring student learning and her own learning from the everyday practices of planning and teaching mathematics.

Exemplar from a classroom-based research project

In this study, I took on multiple roles: teacher, learner and researcher. The focus of intentional noticings and the incidental observations that were made in the classroom were greatly influenced by the theoretical perspective of the research. Realistic Mathematics Education (RME) was used to guide the design of the classroom-based study and analysis of data. RME is founded on the belief that mathematics is not a closed body of knowledge to be transmitted. Rather, it is an exercise in which learners are active participants ([van den Heuvel-Panhuizen, 2003](#)) as they "reinvent" conventional mathematics for themselves, as part of a process of mathematization ([Gravemeijer, 2004](#)).

As the teacher, I consider my role in the classroom as a proactive one in which I need to guide my students' activity of reinvention. My constant challenge is to pay careful attention to students' thinking and to capitalize on inventions as means of developing their understanding of important mathematical concepts. To realize this, I regularly use existing learning trajectories to form the basis of my teaching programs. As "a vehicle for planning learning of a particular concept" ([Simon and Tzur, 2004](#), p. 93), a learning trajectory identifies the key mathematical understandings in a given domain and hypothesizes the path by which student learning might proceed. As such, trajectories shift the focus of the teacher from transmitting content to noticing and making sense of students' cognitive constructions.

In my study, a hypothetical learning trajectory was used to map the predicted path of my students' reinvention in the domain of multi-digit multiplication. I used the trajectory to help anticipate students' likely constructions, to guide their constructions and support their progressive mathematization.

Developing the learning trajectory

Hypothesizing the learning trajectory was akin to Mason's "pre-paring". The trajectory highlighted that which was mathematically significant in the domain, and therefore, that which was important for me to notice. Since an existing learning trajectory for multi-digit multiplication did not exist at the time of this study, I used a sequence of anticipatory thought experiments to delineate my own. First, I conducted a detailed review of relevant literature. The purpose of this review was to draw on existing knowledge of students' thinking in the domain and to improve my own mathematical content knowledge of multi-digit multiplication, with the intention of improving my capability to notice mathematical significant happenings in the classroom. Using this literature review, a draft trajectory was constructed for how students' learning might proceed. The draft identified the key developmental understandings of distributivity and associativity and the mathematical significant aspects of students' constructions of which I should be mindful.

The draft trajectory was then placed in the context of the classroom in which I was teaching. One-to-one interviews were conducted with all the students to assist my understanding of their current levels of reasoning in multiplication. Learning goals and instructional starting points for the sequence were clarified based on these interviews, and tasks designed to elicit the type of thinking desired. Key to this stage was my anticipation of students' solution strategies for each task, and how these strategies were indicative of the key developmental understandings noted on the trajectory. For each strategy anticipated, possible responses were also prepared. Planning in this way was intended to encourage a shift from responding habitually to students and events in the classroom to responding purposefully.

Implementing the learning sequence

The learning sequence was implemented over a two-week period and comprised four teaching episodes. Each episode was characterized by a focus on a distinct concept and spanned two or three 1 h lessons. An episode commenced with a contextual problem to

which students were asked to provide a solution and an explanation of how their solution worked. This section reports on the first three teaching episodes and describes the significant aspects of students' constructions that were noticed and the subsequent instructional decisions that I made as a consequence.

Teaching episode 1—noticing students' use of the distributive property

The first teaching episode focused on building students' understanding of the distributive property of multiplication. Using the context of a bakery, the students were presented with a narrative that prompted them to calculate the total number of cupcakes baked in a day. The cakes were presented pictorially to students as a large array.

Strategies that drew on distributivity were anticipated in the trajectory as studies indicate that the distributive property is an intuitive concept for many students (Ambrose et al., 2003; Barmby et al., 2009; Izsak, 2004). This was realized with many students using the structure of the array to partition it into place value parts to create manageable calculations. However, one method of partitioning the array surprised me. A small number of students used place value to accurately solve the problem but did not utilize the structure of the array in their partitioning. Instead they removed a section of a row, as illustrated in Fig. 3. This manner of partitioning indicated a disconnect between students' procedural competence and their understanding of the multiplicative structure. The disconnect was reinforced as I further noticed students use a finger to count by ones to determine the number of cakes to separate.

The use of this ineffective strategy determined the focus of the whole class discussion conducted at the start of the following lesson. Initially, I had planned for this discussion to focus on the fact that the array can be partitioned to support efficient computation. Based on my previous noticing, I chose to direct students' attention differently: connecting distribution to the multiplicative structure. During the discussion, I presented a number of strategies that used partitioning to the class, including the aforementioned strategy. Comparing the commonalities and differences between strategies, I was able to guide students' attention to that which was mathematically significant, that is, how the structure of multiplication enables efficient distribution, as shown through the array representation. Directing the students' attention in this way prompted them to modify and refine their own thinking and enabled them to build a greater appreciation of how the coordinated rows and columns of an array could be used to aid computation.

Teaching episode 2—noticing students develop an understanding of the associative property

The second teaching episode focused on the associative property of multiplication. For me, a significant learning from the previous task was the importance of focusing on the multiplicative structure, as represented by the array. Episode 2 continued the context of the bakery and, again, the students were required to calculate the total number of cakes in a given collection. In this instance, the cakes appeared in boxes and so were not visible—the array was partially abstracted. Upon solving the problem, the students realized that the total number of cakes in this instance, 16×12 , was the same as the total number of cakes in the first teaching episode, 24×8 . This led to an investigation of a second mathematical goal—why did $16 \times 12 = 24 \times 8$?

During the course of investigation, a number of students noted a certain relationship between the numbers involved, which is reflected in the following excerpt:



Fig. 3 Not separating a full row or column.

Ryan: 12 is half of 24.

Dylan: You could join two of the boxes together to make 24 then ... wait, that's 8 groups ... yeah ... that's 8 groups because 8 twos are 16.

Ryan: So then, 8 is also half of 16 so they are both half.

Dylan: Yeah, I think that's right.

The students had recognized an important numerical relationship, but they had not yet made a connection to the multiplicative structure. I prompted the students to consider how the numerical relationship could be represented using the array. The pair of students took a significant amount of time to make the connection, before dividing the array in half and rearranging it to transform 16×12 into 24×8 (Fig. 4).

Noticing the boys' dynamic use of the array informed my focus of the subsequent lesson. A sequence of associated multiplication problems, specifically focusing on halving and doubling, was presented to the students. In order to highlight that which was mathematically significant, I asked students to represent the different problems using grid paper and to manipulate the arrays to show how each one was connected. Using symbolic recording, and through the course of whole class discussion, students recognized the process of halving and doubling as the associative property of multiplication.

Teaching episode 3—*noticing students' misconceptions with multi-digit numbers*

The intent of the third teaching episode was to extend students' use of the distributive property to two-digit numbers. I presented students with a drawing of the trays inside different cupcake boxes, with the array presented as an area model, or open array. The students discussed why one array was skewed and the other was not (see Fig. 5). The students offered a variety of explanations as to the use of a skewed array, with some hypothesizing that this array was smaller in area and therefore would be cheaper to make. This hypothesis became the focus of the teaching episode.

The initial plan for the class activity was to have students use their own strategies to work out which tray had the larger area. The students were asked to predict if it was true that a tray with dimensions $28 \text{ cm} \times 24 \text{ cm}$ was larger than a tray with dimensions $25 \text{ cm} \times 25 \text{ cm}$. The students' responses were interesting:

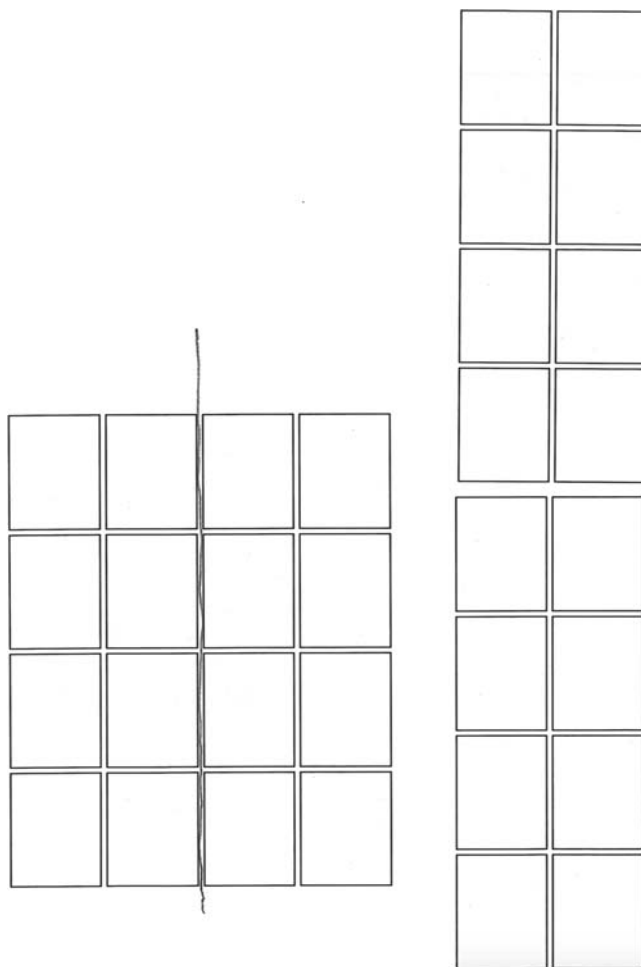


Fig. 4 Ryan and Dylan's representation of the associative property.

Response 1:

Ashley: Well, both of them are just 20 multiplied by 20. Then ... well, with that one [indicating the 25 cm $\times$ 25 cm array] you do 5 times 5 and then with that one (indicating the other array) you do 4 times 8 and I think ... um ... wait, it's, yeah, it's 32 and the other one is 25, and 32 is bigger than 25.

Response 2:

Mia: 28×24 is like 26×26 .

Teacher: Why is that? What are you thinking?

Mia: Well, if you take 2 off the 28 that makes 26, and then you just put that 2 with the 24, then it is 26 too. So, it is 26 and 26 and you just times them together and it's obvious ... 26 and 26 will make more than 25 and 25.

Up to this point in the sequence, the students' responses had demonstrated sound mathematical understandings. However, in this instance, both explanations were based on incorrect mathematical reasoning. The literature reviewed indicated that students do not always recognize all partial products in a multi-digit multiplication (Murray et al., 1994; Young-Loveridge and Mills, 2009). As such, Ashley's response had been represented in the trajectory and I had therefore anticipated it. However, Mia's response was surprising—it was not anticipated. Her misconception centered on compensation and I knew that it is incorrect to use additive compensation strategies in a multiplicative context.

Noticing the students' misconceptions led to a significant and spontaneous instructional decision: the problem shifted from having students work out the total areas of each tray, to asking students to show how their explanations worked mathematically and whether they were correct. The students chose to use the array to make sense of their explanations. Although the context of the investigation used an open array, most students selected to use grid paper which displayed the rows and columns of all individual squares in the array. This noticing was significant because it led me to important new insights about how students learn multiplication: students required the flexibility to move between different forms of the array as they engaged in sense-making. In this instance, students made sense of what was occurring mathematically using a more simplistic form of the array. Once they understood this, their learning moved forward again to the more abstract form of the open array.

Teacher learnings from noticing

On reflection of the instructional sequence, two key learnings emerged from my noticing of students' constructions with the array. First, students needed the opportunity to connect multiplicative calculations with the multiplicative structure. In Teaching Episode 1 and 2, students were able to perform calculations accurately, yet were not always able to represent how these calculations worked using the array. Students needed to be afforded time to make these connections for themselves.

The second significant noticing was students' movement back and forth between simplistic and abstract forms of the array. When faced with their own incomplete understandings, students would revert back to the realistic form of the representation: the array with all parts visible. Students would use this form of the array to make sense of what was happening mathematically.

As a teacher, these two learnings had significant implications for my practice. Students need the opportunity to use the array to explore the multiplicative structure. In their explorations, students should not be restricted to one representation of the array. Rather, flexibility is needed. Students should be afforded the opportunity to select and use different representations of the array (from simplistic with all parts visible to more stylistic and abstract forms), recognizing that different representations will serve different functions in the process of sense making.

Learning from teaching

In this paper we have sought to explore the emerging field of teacher noticing as a vehicle for on-going professional learning that is situated in everyday mathematics classrooms. Although our argument for such teacher learning builds on existing research and theory, we have been selective in our use of evidence from just one case study teacher. We acknowledge that learning to notice the things that matter most in classrooms is complex and challenging, but that it is also a critical part of developing the expertise of all teachers and is therefore both an incredibly important research agenda to pursue and practical discipline to refine. Our intention here was not to be conclusive about how teachers learn from their everyday practices. Instead, we aimed to illustrate how providing teachers with a set of interrelated

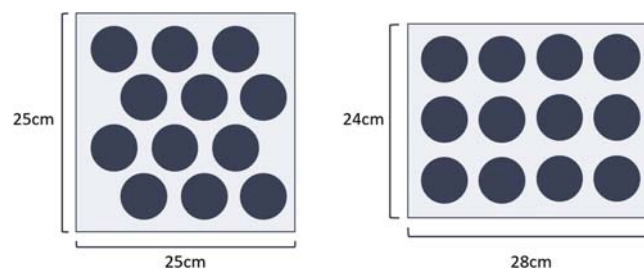


Fig. 5 Comparing the area of two trays in teaching episode 3.

practices emanating from the work of Chan et al. (2018), Mason (2011) and Stein et al. (2008) can help fine-tune their noticings in their classroom with the potential to stimulate teacher learning and improve their teaching practice. In summary, this set entails teachers:

- (1) Pre-selecting the event to focus upon;
- (2) Pre-paring what to attend to by:
 - (a) Post-paring or reflecting upon student responses to recent tasks on the same or a similar topic, drawing upon existing knowledge and new knowledge from research or curriculum documents;
 - (b) Anticipating likely responses by students and how they are indicative of substantial mathematical ideas/concepts;
 - (c) Preparing meaningful teacher responses;
- (3) Attending to in-the-moment events in the classroom;
- (4) Attending to surprising events post-lesson; and
- (5) Reflecting upon what was attended to and interpreting the implications for future instruction.

Given findings that teacher noticings are limited by their mathematical knowledge and conceptions of what is important (Simon, 2007), we suggest juxtaposing this series of practices with an existing or teacher-generated learning trajectory of student understandings to not only guide the selectivity of teachers' attention-seeking gazes but to support the direction in which their instructional sequences should progress. In some countries, such trajectories are emerging in the form of curriculum materials (e.g., Australian National Numeracy Learning Progressions www.australiancurriculum.edu.au/resources/national-literacy-and-numeracy-learning-progressions/) but also exist in research publications (van den Heuvel-Panhuizen, 2003) and evidence-based texts for practitioners (Clements and Sarama, 2009).

Recognizing the complexity of teachers' work and the intensity of demands on their time and attention in any single lesson, we have embedded the above set of noticing practices into a lesson planning protocol (see Fig. 6). This protocol is intended to be adaptable to teacher contexts and is proposed as an additional support tool for teachers as they learn to embed noticing practices into their everyday work of teaching.

Grade/Class:		Date:
Mathematical focus		Materials:
Learning Intentions/ success criteria At the end of this lesson students can: •		
Language to emphasize		
Lesson Launch / Introduction	Launch the task without 'telling' students how to do the mathematics. Clarify terms to ensure students understand the task.	
Anticipate How might my students solve this problem?		How will I respond? (Enabling & Extending prompts)
Anticipate a common error or misconception		Example questions/prompts you can use to help students who get stuck.
Anticipate an advanced strategy or response		Example questions/prompts that you can use to extend students thinking and deepen their reasoning.
Anticipate the response you expect/want most students to use		Example questions/prompts that will focus students' attentions on the important concepts and help elicit their reasoning.
Things I want to attend to while students are exploring the mathematics: Watch how students are ... Look for students who are ...		
Key mathematical idea(s) to highlight during the summarizing phase(s) of the lesson: What mathematics do you want the students to notice and understand?		
Closing the lesson/revising student thinking	Questions or prompts to get students to reflect on their learning and/or revise their thinking. Students complete statements: "I learnt how to"; "I want to better understand", "Before I thought ... but now I think ..."	
Post-lesson teacher reflection	Something that surprised me during the lesson was What happened that I had not anticipated? What does it reveal about student key understandings of mathematical concepts and processes? What will I change (or keep the same) in this or a subsequent lesson? Why?	

Fig. 6 Lesson protocol to support teacher noticing and learning from teaching.

Conclusion

In the past, teachers have been considered passive participants in teacher professional learning initiatives. Consistent with more recent views of teacher learning, we have portrayed teacher professional learning as an agentic and constructive process in which teachers can seek opportunities for learning in their everyday work. We hope that our proposals surrounding the practices of noticing, the use of learning trajectories and a protocol to support teachers integrate and embed these tools into everyday practice will assist their in-situ learning. While our example of how this might occur and what the process of learning might look like was selective, we hope that what we have proposed will entice other researchers and practitioners to continue systematically exploring the powerful potential teacher noticing can have for sustained teacher learning that is embedded in the complexity of their own classrooms.

Finally, we do not advocate that all teacher learning can or should take place in-situ. A healthy balance of professional learning opportunities centered in and out of the classroom is more likely to be effective and practical than too heavy a focus on just one approach. The important determining element for the choice of an individual teacher's professional learning context should be whether the learning focus and approach aligns with her or his needs.

References

- Ambrose, R., Baek, J.M., Carpenter, T., 2003. Children's invention of multiplication and division algorithms. In: Baroody, A.J., Dowker, A. (Eds.), *Studies in Mathematical Thinking and Learning. The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise*. Erlbaum, Mahwah: NJ, pp. 305–336.
- Avalos, B., 2011. Teacher professional development in Teaching and Teacher Education over 10 years. *Teach. Teach. Educ.* 27, 10–20.
- Barnby, P., Harries, T., Higgins, S., Suggate, J., 2009. The array representation and primary children's understanding and reasoning in multiplication. *Educ. Stud. Math.* 70, 217–241.
- Beswick, K., Anderson, J., Hurst, C., 2016. The education and development of practicing teachers. In: Makar, K., Dole, S., Visnovska, J., Goos, M., Bennison, A., Fry, K. (Eds.), *Research in Mathematics Education in Australasia 2012–2015*. Springer, Singapore, pp. 329–352.
- Beswick, K., Fraser, S., Crowley, S., 2017. The best ever: mathematics teachers' perceptions of quality professional learning. In: Kaur, B., Ho, W.K., Toh, T.L., Choy, B.-H. (Eds.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics Education. PME, Singapore*, pp. 169–176.
- Bobis, J., Kaur, B., Cartwright, K., Darragh, L., 2020. Teachers' professional learning and development in mathematics education. In: Way, J., Attard, C., Anderson, J., Bobis, J., McMaster, H., Cartwright, K. (Eds.), *Research in Mathematics Education in Australasia: 2016–2019*. Springer, Singapore, pp. 117–146.
- Chan, M.C.E., Clarke, D.J., Clarke, D.M., Roche, A., Cao, Y., Peter-Koop, A., 2018. Learning from lessons: studying the structure and construction of mathematics teacher knowledge in Australia, China and Germany. *Math. Educ. Res. J.* 30, 89–102.
- Choy, B.H., 2016. Snapshots of mathematics teacher noticing during task design. *Math. Educ. Res. J.* 28, 421–440.
- Clarke, D., Hollingsworth, H., 2002. Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18, 947–967.
- Clements, D.H., Sarama, J., 2009. *Learning and Teaching Early Math: The Learning Trajectories Approach*. Routledge, New York, NY.
- Darling-Hammond, L., Hyler, M.E., Gardner, M., 2017. *Effective Teacher Professional Development*. Learning Policy Institute, Palo Alto, CA.
- Franke, M.L., Carpenter, T., Fennema, E., Ansell, E., Behrend, J., 1998. Understanding teachers' self-sustaining, generative change in the context of professional development. *Teach. Teach. Educ.* 14, 67–80.
- Garet, M.S., Wayne, A.J., Stancavage, F., Taylor, J., Walters, K., Song, M., Hurlburt, S., 2010. *Middle School Mathematics Professional Development Impact Study: Findings After the First Year of Implementation*. U.S. Department of Education, Washington, DC.
- Gravemeijer, K., 2004. Local instruction theories as means of support for teachers in reform mathematics education. *Math. Think. Learn.* 62, 105–128.
- Groves, S., Doig, B., 2016. The role of the knowledgeable other in post-lesson discussions in lesson study. In: Csikos, C., Rausch, A., Szitanyi, J. (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education. PME, Szeged, Hungary*, pp. 315–322.
- Guskey, T.R., 1986. Staff development and the process of teacher change. *Educ. Res.* 15, 5–12.
- Hollingsworth, H., Clarke, D., 2017. Video as a tool of focusing teacher self reflection: supporting and provoking teacher learning. *J. Math. Teach. Educ.* 20, 457–475.
- Ivrs, P., Fernandez, C., Llinares, S., Choy, B.H., 2018. Enhancing noticing: using a hypothetical learning trajectory to improve pre-service primary teachers' professional discourse. *Eurasia J. Math. Sci. Technol. Educ.* 14, 1–16.
- Izsak, A., 2004. Teaching and learning two-digit multiplication: coordinating analyses of classroom practices and individual student learning. *Math. Think. Learn.* 6, 37–79.
- Jacobs, V.R., Lamb, L.L., Philipp, R.A., 2010. Professional noticing of children's mathematical thinking. *J. Res. Math. Educ.* 41, 169–202.
- Jazby, D., Widjagi, W., 2019. Teacher noticing of primary students' mathematical reasoning in a problem-solving task. In: Hine, G., Blackley, S., Cooke, A. (Eds.), *Mathematics Education Research: Impacting Practice: Proceedings of the 42nd Annual Conference of the Mathematics Education Research Group of Australasia. MERGA, Perth, Australia*, pp. 380–387.
- Kennedy, M., 2016. How does professional development improve teaching? *Rev. Educ. Res.* 86, 945–980.
- Kwakman, K., 2003. Factors affecting teachers' participation in professional learning activities. *Teach. Teach. Educ.* 19, 149–170.
- Lindvall, J., Ryve, A., 2019. Coherence and the positioning of teachers in professional development programs. A systematic review. *Educ. Res. Rev.* 27, 140–154.
- Margolinas, C., Coulange, L., Bessot, A., 2005. What can the teacher learn in the classroom? *Educ. Stud. Math.* 59, 205–234.
- Mason, J., 2002. *Researching Your Own Practice: The Discipline of Noticing*. Routledge, London.
- Mason, J., 2011. Noticing: roots and branches. In: Sherin, M., Jacobs, V., Philipp, R. (Eds.), *Mathematics Teacher Noticing: Seeing Through Teachers' Eyes. Studies in Mathematical Thinking and Learning*. Routledge, New York, pp. 35–50.
- Murray, H., Olivier, A., Human, P., 1994. Fifth graders' multi-digit multiplication and division strategies after five years' problem-centered learning. *Int. Group Psychol. Math. Educ.* 18, 399–406.
- OECD [Organisation for Economic Co-operation and Development], 2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*. OECD Publishing, Paris.
- OECD [Organisation for Economic Cooperation and Development], 2016. *High-Quality Teacher Professional Development and Classroom Teaching Practices: Evidence From Talis 2013 (OECD Education Working Papers No. 141)*. OECD Publishing, Paris.
- Richter, D., Kleinknecht, M., Gröschner, A., 2019. What motivates teachers to participate in professional development? An empirical investigation of motivational orientations and the uptake of formal learning opportunities. *Teach. Teach. Educ.* 86, 102929.
- Simon, M., Tzur, R., 2004. Explicating the role of mathematical tasks in conceptual learning: an elaboration of the hypothetical learning trajectory. *Math. Think. Learn.* 6, 91–104.
- Simon, M., 2007. Constraints on what teachers can learn from their practice: teachers' assimilatory schemes. In: Woo, J.H., Lew, H.C., Park, K.S., Seo, D.Y. (Eds.), *Proceedings of the 31st Annual Conference of the International Group for the Psychology of Mathematics Education. PME, Seoul, South Korea*, pp. 137–141.
- Simons, D.J., 2000. Attentional capture and inattention blindness. *Trends Cogn. Sci.* 4, 147–155.

- Stein, M.K., Engle, R.A., Smith, M.S., Hughes, E.K., 2008. Orchestrating productive mathematical discussions: five practices for helping teachers move beyond show and tell. *Math. Think. Learn.* 10, 313–340.
- Superfine, A.C., 2019. Reconceptualizing ways of studying teacher learning: working with teachers rather than conducting research on teachers. *J. Math. Teach. Educ.* 22, 1–4.
- van den Heuvel-Panhuizen, M., 2003. The didactical use of models in realistic mathematics education: an example from a longitudinal trajectory on percentage. *Educ. Stud. Math.* 54, 9–35.
- van Es, E.A., Sherin, M.G., 2008. Mathematics teachers' "learning to notice" in the context of a video club. *Teach. Teach. Educ.* 24, 244–276.
- Young-Loveridge, J., Mills, J., 2009. Teaching multi-digit multiplication using array-based materials. In: Hunter, R., Bicknell, B., Burgess, T. (Eds.), *Crossing Divides: Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia*. MERGA, Palmerston North, NZ, pp. 635–643.

Further Reading

- Clements, D.H., Sarama, J., 2014. Learning trajectories: foundations for effective, research-based education. In: Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age, Charlotte, NC, pp. 1–30.
- Nabb, K., Hofacker, E., Ernie, K., Ahrendt, S., 2018. Selecting and sequencing student work with cognitively demanding tasks in a group environment can teach important mathematical ideas. *Math. Teach.* 111, 366–373.
- Schack, E.O., Fisher, M.H., Wilhelm, J.A. (Eds.), 2017. *Teacher Noticing: Bridging and Broadening Perspectives, Contexts and Frameworks*. Springer, Singapore.

Middle leading as a practice for enabling professional learning

Karin Rönnerman^a and Veronica Sülau^{b, a} Department of Education and Special Education, University of Gothenburg, Sweden; and
^b Frida Education, Vänersborg, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Professional learning – a brief overview	528
Methods and methodology	530
The site	530
The theory of practice architectures	531
Data collection	532
Findings	532
Organizing for professional learning: A model where two levels of middle leading co-exist	533
Middle leading as managing professional learning	533
Middle leading as facilitating professional learning	533
Middle leading at different levels	534
The model enables and constrains the professional learning practice	534
Number of levels/positions - boundary management	535
Scaffolding - to create structures for learning	535
Shared spaces for communication - possibility to calibrate the initiative	535
Distinctness and stringency in the commission statement	535
Positive culture of change - trust and confidence	535
Concluding remarks	536
Middle leading practice at two levels	536
Trust	537
The theory of practice architectures - a way of understanding middle leading	537
References	537

This article examines the discourse on professional learning by exploring how it is enacted at a school site in Sweden. The term collaborative learning (in Swedish called “collegial learning”, *kollegialt lärande*) was introduced in Sweden by the Swedish National Agency for Education (in Swedish *Skolverket*) in 2010. During the last decade, it has been used as a central term in Swedish reforms of continuing professional development (CPD) as something to strive for and to be fulfilled (Kirsten and Carlbaum 2020). The arguments are based on research on professional learning communities (PLCs) and teacher learning communities (TLCs).

For a better understanding of the study, a contextualization of Swedish policy reforms during the last decade is needed. In 2010, a rewriting of the Swedish Education Act (SFS 2010:800) emphasized a focus on education to be built on research and proven experiences. Along with this, a new career pathway reform for teachers (in Swedish *karriärvägar för lärare*) was established in 2013 (SFS, 2013:70). The reform was legislated by the Government and enacted by the National Agency of Education. It addressed teachers working in schools, with the purpose of creating possibilities for supporting school development and professional learning at the local school (Skolverket 2020). The reform included two possible ways of career paths: (1) teachers studying at university to acquire a Ph.D. exam and to be employed as a lecturer in schools; (2) teachers applying to the responsible authorities to become a “First Teacher” (in Swedish *förstelärare*) at their local school. No academic training is required for being a First Teacher, but certain criteria must be met: being certified as a teacher, having a minimum of four years of educational work in a school, improving student academic outcomes, being interested in developing teachers’ professional learning and being considered appropriate to teach. Both positions are subject to a distinct monthly salary rise (for a lecturer €1000 and for a First Teacher €500). This reform lies well in line with similar global movements in other countries regarding teacher professional standards, e.g., in USA, Australia and Scotland (Hardy and Rönnerman 2018).

In this article we will focus only on the First Teacher and how this part of the reform was enacted in an independent school company, specifically in an upper secondary school. The purpose of this article is to elaborate on how professional learning is organized at a specific school site, and addresses two questions:

- How does a specific school site organize for professional learning through middle leading?
- How does the specific organization of the middle leading practices enable and constrain professional learning among the teachers?

Professional learning – a brief overview

Professional learning has been widely researched globally. During the last decades, there has been a shift in the meaning of this term, from in-service training via professional development to professional learning. While in-service training focused on the

individual teacher, professional learning has a stronger focus on interrelationships with teacher learning, teacher leading and educational change (Edwards-Groves and Rönnerman 2013). Similarly, Timperley (2011) suggests the change in concept from professional development to professional learning, which emphasizes that the teachers have an active role in their own learning.

In contemporary literature, professional learning is also combined in a number of ways with concepts such as collegiality, collaboration, collectivity and community. In a Swedish study, Granberg and Ohlsson (2016) differentiate the terms collegial, collaborative and collective by their different meanings. They argue that collective and collaborative learning has to do with individuals working together to develop mutual knowledge, while collegial learning is more referred to as a method to achieve specific theories or teaching methods to be used individually in the classroom. In the Swedish educational discourse, the concept “collegial learning” was introduced by the Swedish National Agency of Education to be used for national school development programs. The concept is used both as an aim and as a strategy: as an aim to achieve a different culture of learning in schools where teachers collaborate, and as a strategy used in CPD programs with a specific model to promote teacher learning. These programs concern different aspects of educational work, for instance numeracy and literacy. They are structured in a certain way encompassing four steps: (a) individual preparation (reading texts), (b) collegial conversations (discussing content with colleagues and planning a lesson/activity), (c) lesson activity (performing the planned lesson/activity in the classroom) and (d) collegial follow-up (discussing and reflecting on the performed lesson/activity) (Sülau 2019, 106). Within these programs teachers are gathered in groups facilitated by a colleague, who have participated in a short training supplied by the National Agency of Education on how to promote collegial learning (Skolverket 2020). Hardy and Rönnerman (2018) explored how reforms in Sweden have reconstituted teachers’ work. By examining the different reforms following the Education Act, they claim that current reforms reflect neoliberal policy that seek to modify teachers’ practices at a micro level, more than earlier reforms have.

Hargreaves and Fullan (2012) point to the importance of teachers’ professional learning in order to achieve successful student outcomes. They refer to the professional learning that occurs when teachers continuously and collaboratively explore and improve their own teaching. In the same way Timperley (2011) emphasizes the importance of inquiry and argues strongly on using students’ results as the main basis in the teachers’ discussions to be able to learn and develop their teaching. From these perspectives, professional learning is about collaborative action that includes collaborative activities and follow-ups from what is done in practice. Learning might otherwise be limited to the individuals, and the collaborative processes highlighted by Timperley risks getting lost.

In the literature, it is also evident that professional learning is connected to community. Some factors that come through in different studies as important for professional learning communities to be successful, are shared norms and values, collaborative focus on student learning, sharing experiences and dialogues (Fischer, 2013). Some studies even find a relation between PLC’s and student results (McLaughlin and Talbert 2010; Opfer and Pedder 2011). Opfer and Pedder (2011) argue further that teachers learn best when involved in actions together and when these actions are related to everyday activities. Such a community is characterized by collaborative structures and cultures, shared leadership, shared values and visions and collaborative decision-making (Stoll et al. 2006; Nehring and Fitzsimons, 2011). In a recent study, Sülau (2019, 43–46) summarizes six factors for successful professional learning communities, namely purpose and focus, relationships, collaboration, enquiry, leadership and building capacity and support. Most of these factors coincide with other research as presented above, but in this article two are of specific interest: leadership and capacity building in terms of trust.

Leading connected with trust has become more essential in recent times. In Sweden, the two concepts have merged into the concept of “trustbased leading” (in Swedish *tillitsbaserat ledarskap*). This can be seen as a reaction to New Public Management models. Instead of steering toward goal and result, steering by trust focuses on the practices and the need among its participants, where relations are put to the fore. Ståhlkrantz (2021) argues that in the trust-based steering model, every level of decision-making acts to secure the quality of all students’ development and learning. The model strives for collaboration between professionals on a horizontal level to secure collective quality at the same time as it collaborates with the vertical steering chain.

Likewise, Edwards-Groves et al. (2016) show how middle leaders take responsibility for these factors by showing empathy related to trust and confidence to colleagues. A middle leader is in this case a teacher, who most of the time is working in the classroom but also has a role of facilitating learning with colleagues - leading from the middle (Grootenboer et al. 2015). Middle leaders by this definition open up and keep safe places for collaboration and conversations and thereby show togetherness by sharing their practices. Furthermore, a middle leader shows professional knowledge around on-going activities and leads change that is practical, relevant, realistic and reachable. Trust becomes important for this to happen. Jederlund (2019) analyzed school development from the concepts of collegial and collective trust. While “collegial trust is a prerequisite for collaboration, collective trust is a result by collaboration” (Jederlund, 2019, translated from Swedish by the authors, italics in Swedish original). Trust can be strengthened over time, by leaders securing visions and through organizing and supporting collaboration. Jederlund (2019) argues for the importance of organizing and supporting a practice for learning to occur among teachers.

However, as Harris et al. (2019) argue, the term “middle leader” is problematic as one can say there is always a middle leader when you view an organization. They argue that a more diverse set of middle leadership has emerged in the literature (Harris et al., 2019). For example, Fullan (in Harris et al., 2019) defines a middle leader in the regional system as someone having an overall view in the relation between the system goals and the local needs. In the same way Hargreaves and Ainscow (2015) points to middle leaders who are at the district level being the collective drivers of change and development. Harris et al. (2019) also refer to research that underlines how middle leaders have an important role in enhancing students learning. It seems that the term “middle leader” is used at different levels in the school system. Harris et al. (2019) found that research focusing on leadership responsibilities, action and the roles of different middle leaders has increased in the last ten years (Harris et al., 2019, 269).

In this article, we explore the middle leading practice as it is organized at a specific school site and started our work to follow the definition of middle leading from Grootenboer et al. (2015, 524) talking about three characteristics of middle leading as being:

- (1) Positionally – middle leaders are structurally and relationally situated “between” the school senior management and the teaching staff. They are not in a peculiar space of their own, but rather they are practising members of both groups.
- (2) Philosophically – middle leaders practice their leading from the center or alongside their peers. In this sense they are not the “heroic crusader” leading from the front, but rather are alongside and in collaboration with their colleagues.
- (3) In practice – middle leading is a practice and is understood and developed as a practice. To this end, the focus is on the sayings, doings and relatings of leading rather than the characteristics and qualities of middle leadership.

At this site, the middle leading practice unfolds at two levels: in the organization and among the teachers. Our focus is on exploring how the organization of the middle leading practice enables and constrains professional learning. In the following section, we will turn to the site in which this study took place and present its organization followed by a brief introduction to the theory of practice architectures, which is used for analyzing the results. Thereafter, methods used for data collection will be presented followed by the results and conclusions.

Methods and methodology

The site

Frida Education is an independent education company, founded by a few teachers during the “free school reform” (in Swedish *Fris-kolereformen*) that was implemented by the Swedish Government in 1992. The reform aimed at increasing the variety of schools and to create competition between schools by giving students and parents the opportunity to have a free choice of schools. The independent schools were allocated the same resources per student as in the municipal schools. For the teachers who founded Frida Education, the reform made it possible to build a school with a particular focus on the writings in the new curriculum (Lpo 94) about fundamental values and tasks of the school, where concern for the well-being and development of each individual, an equivalent education and student responsibility and influence were among the leading concepts. The vision - *Educating backpackers* - was formed together with some guidelines on how to create a good environment for learning. The guidelines permeate all levels of the company and are the foundation in all activities and relations among the children, students, staff and parents. Some of the key words in the guidelines are: social openness, strong trust, collaboration, flexibility and constant dialogue (Fridaskolan, 2020).

Today, Frida Education consists of different subsidiaries: Frida schools and Didaktikcentrum¹ (explained further down). The company has grown successively since it started in 1993. At the time of the current study, there were Frida schools in five different municipalities in western Sweden, with education from preschool (children from the ages of 1–5) to upper secondary school. About 3500 children and students and 600 employees (teaching staff, service staff and administration) were at this time a part of Frida Education.

When the new career pathway reform for teachers was implemented, Frida Education decided to build a system of First Teachers within the organization. In total, the company was entitled funding for 33 First Teachers; 6 at each elementary school unit and 3 at the upper secondary school. The First Teachers were appointed through an internal application based on the criteria set out in the state directives (especially skilled teachers). The principals at each school unit conducted interviews and decided which teachers should be assigned the commission. The First Teachers at Frida Education have no reduction in working hours, but receive the state grant of €500 as an increased monthly salary premium. 31 of the First Teachers at Frida Education have a 100% employment as teachers at the six different Frida school units. Their commission as a First Teacher is to facilitate professional learning among their colleagues in the current CPD initiative. In this article they are named Facilitators (F).

Two of the First Teachers have a specific commission to work as development leaders (in the text referred to as DL) with a specific responsibility for the company’s internal CPD work. They have a 40% employment at Didaktikcentrum, while the remaining 60% they are employed as teachers at the Frida schools. The Development Leaders are planning and implementing the school board’s decisions of the upcoming two-year CPD cycle. They have a specific responsibility to organize educational meetings (training) on the current professional learning content for the 31 First Teachers acting as facilitators within the organization.

At Didaktikcentrum, a group of people with different competences and responsibilities work with organization and school development issues. There are two researchers with a Ph.D. in Education respective Educational Work, two Development Leaders (DLs) and one person in charge of international projects and relations. In order to lead relevant and important development work, one of Didaktikcentrum’s remits is to keep abreast of current national and global trends in educational research, for example through networking with researchers, politicians and the education community and through canvassing academic literature. Didaktikcentrum conducts an ongoing dialogue with the company board about possible content for upcoming CPD initiatives. The goal is to plan these initiatives about a year in advance before they are enacted in the schools.

Frida Education has a model for change and improvement work, which is known within the organization. CPD initiatives are performed over cycles of at least two years, building on the experience that structural changes and processes take time, and need to be given time in order to create sustainable change in all parts of the organization. The content for the professional learning is decided

¹In 2022, Didaktikcentrum changed names to Frida Professional in order to better describe its focus and purpose. The organisation and the staff as described in the study remains the same.

based on the needs of the local schools, but also on influences by the research and development team at Didaktikcentrum, the National Agency of Education and current educational research literature. Experiences from previous CPD initiatives are taken advantage of, based on how successful they have been in relation to what effects they have generated within the organization. As a part of the company model for change and improvement work, there are also set times for professional learning among the staff. Every Wednesday between 2 p.m. and 4 p.m. is scheduled for all teaching staff to be engaged in educational work conferences. Each principal plan and leads these conferences with a content relevant to the local school. During a school year, 11 of these conferences are predetermined by the company board, and are related to the specific CPD content for the current period. These conferences are to be planned by the principal together with the facilitators at each school, with the support from the development leaders at Didaktikcentrum.

During the school years 2017/18 and 2018/19 the CPD work within the organization focused on *collegial learning*. The purpose of the initiative was described by the school board as “developing a general competence among the staff, which is about trying to understand how to work better together” (Educational Manager) in order to bring about “development and change in the classroom with the students” (Regional Manager). Collegial learning has been part of the company’s idea of how development work should be conducted since its inception, but has not been systematized and organized the way it was during this time period.

The theory of practice architectures

This research focuses on how one national reform, the career pathways (SFS 2013:70), was enacted in an independent school company in western Sweden. We found the theory of practice architectures (TPAs) helpful in understanding how professional learning is being enabled and constrained within an organization (Kemmis et al., 2014). This theory provides a valuable lens for understanding practices (in this case the middle leading) as they happen in real places (like in collegial meetings among teachers or First Teacher meetings). TPA is a site-based theory where the happenings in the practice are in focus. A practice is understood as socially constructed by human beings who come together as individuals in particular sites (like schools or classrooms) and under certain situations and activities (like collegial meetings). A practice consists of *sayings*, *doings* and *relatings* with a specific goal, the project. In many ways, practices are organized around the everyday routines of daily life. Wenger (1998) claims we go in and out of practices all day long. For example, the everyday practices of a teacher can be teaching students in the classroom, meeting with colleagues in the staff-room and joining meetings for professional learning facilitated by a colleague. Because a practice is a social phenomenon based on happenings in the intersubjective spaces between human beings, a practice is never fixed, it is an open question about how things will unfold in real life in real time. Therefore, practices can be understood in routines and predictability, but are prepared for the extraordinary and unexpected at the same time.

A practice is always developed in its local context. In other words, the middle leading practice in the actual organization is situated within a complex of other education practices, and is influenced (both supported and challenged) by these other practices. This means middle leading is mediated, shaped or influenced by the very particular needs of the other individuals with whom middle leaders work (students, colleagues, principals, other administrative staff) and by specific circumstances in the places (or sites) where the middle leading occurs. These specific circumstances are the practice architectures, which inform and are informed by the practice.

The practice architectures consist of three different types of arrangements all enabling and constraining the practice. The *cultural-discursive arrangements* occur in the realm of an intersubjective space where language and thoughts (*sayings*) are used to create shared meanings and mutual understandings about what is being done in the practice. At this specific site, one example is the shared vision and values in the school organization and how they are understood and discussed among teachers. The *economic-material arrangements* occur in the realm of physical space-time where particular activities (*doings*) happen as colleagues meet in shared places in space and time to work with one another and being involved in collaborative activities. In our case, the teachers take part in collegial meetings led by the facilitator at certain times in a specific place. In the same way the facilitators meet each principal to discuss and plan how to continue with the professional learning program. The third type of arrangements, the *social-political arrangements*, occur in social space time where relationships (*relatings*) exist or develop as teachers and facilitators form diverse kinds of collegial relationships through their collaborative meetings. At the site of the current study, the meetings between the teachers and the facilitator can be an activity where power and solidarity is enacted. Therefore, the facilitator needs to be conscious of the roles and the relationships between the teachers so that professional learning is the focus and not the person *per se*.

In actual sites, practices are always influenced by other things that are both local and external, that both enable and constrain what can possibly be done. That is to say, practices of any kind are influenced by conditions, in the theory described as practice architectures. The theory of practice architectures (Kemmis and Grootenboer 2008; Kemmis et al., 2014) lead educators toward a theory of social practice that enables us to consider broader questions concerning the cultural-discursive, material-economic and social-political arrangements (the practice architectures) that shape the practices, for changing educational practices in schools (see Fig. 1).

Kemmis et al. (2014, 6) describe the purpose of the theory of practice architectures as threefold. First, it provides a theoretical language for researchers in sociology and educational sciences to interpret and describe the world. Second, it is a practical tool to use when working with change processes in these scientific fields. In order to change the way a practice come about, one must first identify the specific arrangements that keep it in place. This conclusion leads to the third purpose of the theory, which is to provide a critical perspective toward the relationship between the practice and the practice architectures. By identifying the practice architectures of a practice and how these enable and constrain the practice in reaching its goal (project), it is also possible to be creative in changing these arrangements to be either over won or strengthened.

The *sayings, doings* and *relatings* of a practice are made possible and held in place by *practice architectures*

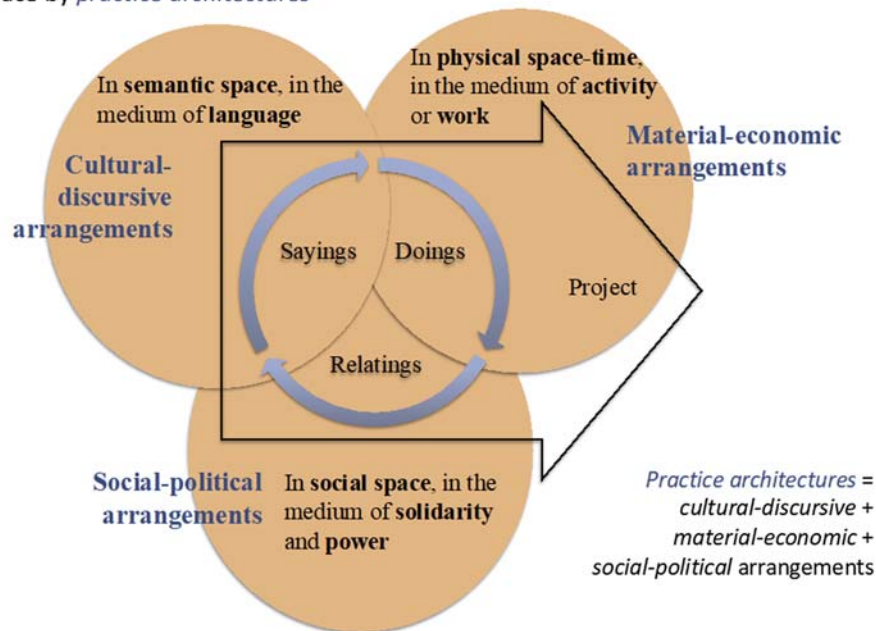


Fig. 1 How a practice is held in place by its practice architectures. Adapted from Kemmis et al., 2014, 34.

Data collection

The case study in this research is one upper secondary school within an independent school company, Frida Education. It can be considered a unique case in terms of how the school board since the start more than 25 years ago has had a conscious strategy and organization for CPD work, based on the vision *Educating backpackers* (described at www.fridaskolan.se). The vision is grounded in the idea that learning takes place in social interaction, based on communication and relations. This means, among other things, that the work within the school company is organized in teams in all parts of the organization, both within the board and school management as well as among teachers and staff in offices and services. In this study, we explore how this organization is enabling and constraining professional learning, with a specific focus on the middle leading practices. Our main data sources are interviews with persons having leading roles in different parts of the organization. The following people were interviewed:

- The educational manager (EM) at Frida Education
- The regional manager (RM) at Frida Education
- The principal (P) at the upper secondary school (Fridagymnasiet)
- The two first teachers acting as development leaders (DL) at Didaktikcentrum
- The two first teachers acting as facilitators (F) in the upper secondary school

The Educational and the Regional Manager, the principal and the development leaders were interviewed once. The two facilitators were interviewed twice, the first time together and the second time separately. The two researchers (the authors of this article) were both present at most interviews. Exceptions were the interviews with the principal and the development leaders and the second interview with one of the facilitators, when only one of the researchers attended, due to time and distance to the school. Prior to conducting the interviews, the two researchers together worked on the interview schedule. While one of them conducted the interview, the other took notes. All interviews were audio recorded and kept in a safe place.

No interviews were transcribed in full. Based on the notes taken while interviewing and repeatedly listening to the audio recordings, selected data was transcribed and used for analyses and excerpts in the findings. According to the research questions, the focus in this article is on the middle leading practices and how they were enabled and constrained in leading professional learning in the organization.

Findings

In this study, concepts of the theory of practice architectures are used as a lens for analyses (Kemmis et al., 2014). The theory of practice architectures is based on the idea that people meet and understand each other in what Kemmis et al. (2014, 4) call *inter-subjective spaces*. These spaces are always arranged in certain ways; through language, place, time, material things and through social relations. Therefore, how we understand each other and how things happen in a specific practice, is conditioned and prefigured but not determined by these arrangements.

In this study, a point of departure is to illuminate how the “First Teacher” reform in Sweden can be used to create conditions for professional learning. We are specifically focusing on the middle leading practice within an independent school company, to explore how this practice enable and constrain professional learning among teachers. The theory of practice architectures has been used to identify the arrangements that in different ways enable and constrain the project of the middle leading practice, that is to promote professional learning. This makes it possible to draw conclusions on the conditions necessary for the practice to reach its goal, and what changes that might need to come about in order for the practice to work differently.

Organizing for professional learning: A model where two levels of middle leading co-exist

In the first part of the findings, the way that the school company organizes the middle leading practice is presented. The results show that the middle leading practice within the organization is enacted at two levels with different roles and functions; middle leading as managing (development leaders) and middle leading as facilitating (facilitators) (Fig. 2). The interaction between these types is shown in the form of a model (Fig. 3).

Middle leading as managing professional learning

The development leaders within the organization describe their mission as managing the first teachers in their role as facilitators in the collegial groups among the staff. They also support the facilitators in planning school conferences together with the respective principals at the six school units within the company. On eight occasions during one school year, the development leaders meet with the facilitators. On these occasions, seminars and discussions are held about the facilitator’s role in leading professional learning.

In the interviews, the development leaders emphasized that their area of responsibility is only linked to the *structure* of the CPD initiative, that is, the collegial conversations among the teachers. As far as the *content* of the collegial conversations is concerned, it is up to each principal to decide for him/herself based on the needs of the local school. The principals and the facilitators are then expected to plan the professional learning conferences together to unite structure and content in a relevant and functional way.

Within the organization, it is the regional managers who have continuous contact with the principals; they serve as support both in the CPD work and in other parts of the principal’s assignment. The development leaders thus constitute an important support for the regional managers in understanding and communicating the CPD work toward the principals. It is through the meetings with the facilitators that the development leaders are given the opportunity to follow the CPD work, how it is conducted and organized at the various school units.

Middle leading as facilitating professional learning

The facilitators are described by the development leaders as having a significant role in the CPD initiative and for professional learning to be an integral part of the organization. They are in a position between the teachers and the principals in the specific assignment to lead collegial conversations with the teachers.

The two first teachers at the upper secondary school described their mission as facilitating professional learning among their peers. Preparations for this take place at the monthly meetings they have with the development leaders. There is an agenda for the meetings and the facilitators are often given tasks to prepare in advance. It may involve reading some literature and reflecting on the content in relation to one’s own teaching. They also have specific responsibilities for the two joint professional learning days each academic year where all teaching staff within the company participates.

The facilitators suggested that the collegial work in the schools is not a new phenomenon *per se*. Conversation and dialogue are important parts of all conferences and meetings. The difference, the facilitators suggested, is that it now takes place more

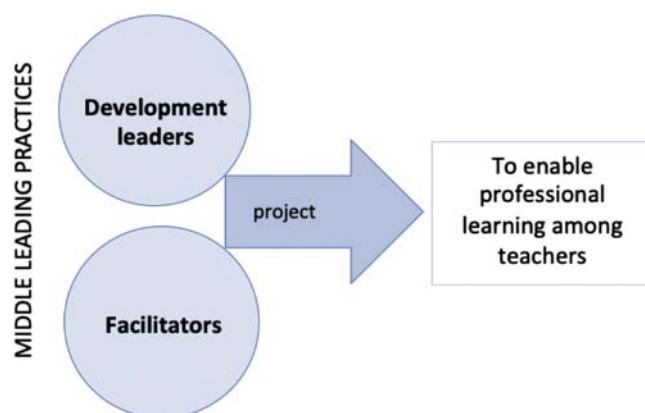


Fig. 2 The two middle leading practices within Frida Education.

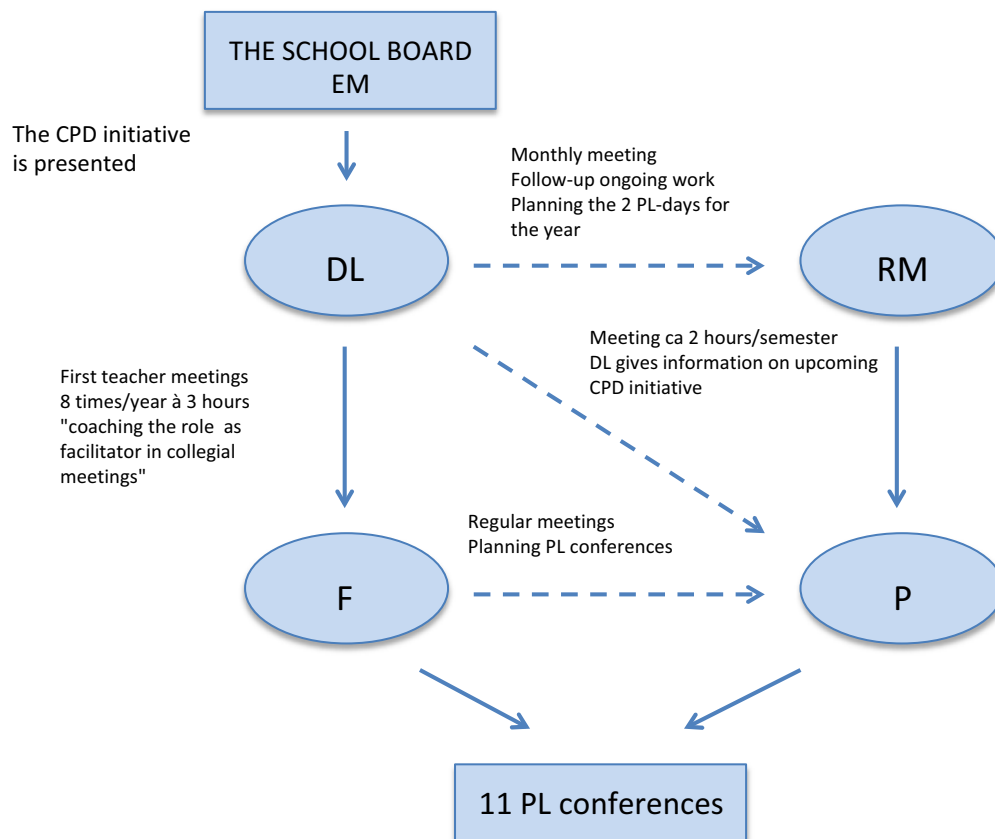


Fig. 3 The model of the organization of professional learning at Frida Education. (DL, Development leader; EM, Educational Manager; F, Facilitator; P, Principal; RM, Regional Manager).

systematically and in a more structured way. As facilitators, the first teachers are given the opportunity to create a structure in the collegial meeting based on various success factors described in the literature on professional learning.

The two first teachers at the upper secondary school described that they applied for the assignment mainly because they had worked for a number of years within the company and felt a need to take new steps in their careers. They suggested that the assignment as a facilitator gave them the opportunity to be challenged and to take on a greater responsibility. Through the regular meetings with the principals, the facilitators gained greater transparency and participation in the CPD work. Being able to attend and lead conferences and collegial conversations is perceived as a motivating factor. Through the assignment, the facilitators were able to access to new information and knowledge in the meetings with the development leaders, which they also believed developed their personal competence.

Middle leading at different levels

The results show that middle leading within the organization is enacted at two different levels with different assignments and objectives. Some of these are more directed toward managerial tasks, while others are directed toward the practice and the professional learning among teachers. The following model crystallizes:

The model enables and constrains the professional learning practice

The school group's organization of the middle leading practice conditions the way professional learning is and can be enacted in the organization. In this section, the results on how the model enables and constrains professional learning among the teachers are presented.

The middle leading practice in the current school group includes two different levels of middle leading with different assignments and roles (Fig. 2). However, both levels work toward the same goal, or project, to enable professional learning among the teachers. An analysis of the interviews using the theory of practice architectures shows how the middle leading practice is held in place by different arrangements, which enable and constrain this practices in reaching its goal. In this section, the arrangements that are specifically linked to organization are highlighted.

Number of levels/positions - boundary management

The school organization's model for professional learning contains a variety of levels, where different positions are responsible for different staff groups. It provides an opportunity to distribute the workload, but it also requires the different groups to understand the system and have confidence in each other. The term "boundary management" is used by the employees to point out differences between the various positions and assignments. Boundary management, according to the respondents, means that one position should not work with detail-controlled issues that fall within someone else's area of responsibility. At the same time, different positions need to be aware of details in order to be able to organize the work in the best way. In the interviews, it is shown that the balance between governance and information in these boundary managements sometimes constrains the middle leaders' work with professional learning. The development leaders suggested that when the principals have too little knowledge of the professional learning and the content of the meetings with the facilitators, it creates too much room for the principals to pursue the development work in their own direction that does not harmonize with the intended whole. The facilitators also described ending up between two different boundary managements - the development leaders and the principals - where they sometimes had to compromise between the input they received at the facilitator meetings and the output that the principals wanted to address at the conferences.

Scaffolding - to create structures for learning

All groups in the model (Fig. 3) expressed the importance of supporting professional learning with help of different structures. For instance, this meant creating spaces for communication and collaboration, setting agendas for meetings and creating synergies and consistency between the different leading practices. Ultimately, this responsibility lies with the Education Manager at the school board, who says that he is constantly asking himself, "Is it organized well enough to create consistent change?"

An important part of these structures is about anticipation in relation to the commission. Both development leaders and facilitators reported that they needed time to understand the commission (from the school board) in order to read relevant literature and to plan appropriate content in the facilitator meetings and PL conferences respectively. The Regional Manager also emphasized the importance of, in the beginning of the school year, being able to communicate the content for the coming academic years' 11 PL conferences to principals. By having an established structure for CPD initiatives, the principals at the individual school units do not have to spend time creating this structure themselves. However, it also means that time and space must be given to create a common understanding of what is to be achieved and in what way it should be done.

Shared spaces for communication - possibility to calibrate the initiative

A prominent part of the arrangements that stipulate the extent to which the middle leading practices can enable professional learning is about shared spaces for communication. In the model above (Fig. 3), the different spaces for communication are shown in the form of solid and dashed arrows. The solid arrows mark pronounced communication paths between different boundary managements, while the dashed arrows indicate communication paths that implicitly and on a very limited scale, intersect between the different boundary managements. All respondents expressed the importance of increasing these spaces to allow professional learning.

Meetings between the different groups fulfill two functions: firstly, to create opportunities to convey the professional learning content in a similar way, and secondly to be able to calibrate the initiative if any group deviates from the intended content. However, these spaces or meetings do not occur spontaneously, but need to be planned in advance.

Distinctness and stringency in the commission statement

Another important success factor in enabling middle leading practices to promote professional learning is how the CPD initiative is communicated, both in terms of structure and content. All respondents discussed the importance of expectations of the different roles and functions included in the model being clearly stated for the CPD initiative to work. Especially for the middle leading practice, these expectations become evident. The development leaders expressed the importance of the role of the facilitator and his/her commission in relation to the CPD initiative being clearly stated and defined. Should the facilitator act as support for the principal in leading professional learning among teachers or as a catalyst that drives and shapes professional learning at one's own school? Depending on the perspective, the result will be different. It also has consequences for the professional learning in the middle leading practice. When the expectations of the facilitator's role differ in the different school units, it becomes difficult for the development leaders to create a common content for the facilitators. Facilitators also run the risk of being trapped between development leaders' expectations and principals' expectations, which will constrain the professional learning. Thus, it is crucial that the functions and assignments of the first teachers are clearly communicated by the school board at all stages of the organization.

Positive culture of change - trust and confidence

According to the respondents, the organization with facilitators as key persons to support professional learning among the staff is based on a high degree of trust and confidence. Facilitators must be given a mandate to pursue professional learning, both through a formal mandate from the principal and through a more informal mandate from their fellow colleagues. The interviews show that this trust exists in the organization. The Educational Manager stated: "Compared with the investigations and evaluations that are written nationally and in the media about this (the First Teacher reform), there seems to be a relatively high degree of trust among other teachers within Frida Education". One of the facilitators said in the interview: I have received a good response from my

colleagues. They have been positive. It is important that you are accepted by your colleagues, especially since we (Frida Education) strive to be a flat organization where the salaries don't differ too much'. The facilitators also discuss how the trust they have been given in being appointed a First Teacher influence their work. They said: "It also feels important for us that we handle that task [facilitating] and take it seriously. You want to live up to it and feel that it was right that they chose me". Both the facilitators as well as the development leaders and the principal reported that they received positive feedback from the staff, both in terms of the facilitator's role as a leader in professional learning and in the relevance of the content of the CPD initiative. The development leaders highlighted that the culture of the organization is positive as well and that all staff are very active in change work.

The respondents described the teachers' high confidence in the school management and their decision to select a relevant CPD content together with the possibility and ability to adapt the content to the different school units, as a success factor. At the same time, they say, it is a balance between central control and local adaptation. Content that is too general risks being perceived by the teachers as irrelevant, while a local adaptation that is too loose can make it difficult to create a common foundation for professional learning. In this work, the facilitator has an important node for shaping and keeping the professional learning practice in place.

Concluding remarks

The aim of the study was to explore how a specific independent school company has implemented the national career pathway reform to organize professional learning through middle leading. The independent school company in focus for the study has a long history in organizing for professional learning. When a new reform was introduced (career path ways) by the Government, the school company had to find ways to adapt to the new criteria, which meant to find ways to introduce a new professional role - the First Teacher as a facilitator. The Educational Manager emphasized the view that it was important to find a relevant relation between the new role and the existing ideas in the company on how to promote professional learning. Many dialogues were held between the school board and the principals on how to define the role and task of the First Teachers.

Middle leading practice at two levels

As we have shown in this case study, the middle leading practice at Frida Education involves two levels. The organization of the First Teacher reform at the specific site has created two types of First Teachers who coexist in this practice: the development leaders having more of a managerial role and the facilitators leading their colleagues. Both levels form their own practices within the middle leading practice; but are both driven by the same project, to promote professional learning among teachers (Fig. 2). In the literature the concept of middle leading is not clear-cut. Hargreaves and Ainscowe (2015) use it to refer to leaders outside the actual school with more of an administrative role in organizing how to disseminate knowledge in a region. Additionally, Fullan (2015) points to this role as well and emphasizes the necessity of this role getting more concordance in the school system through strengthening the goals and local needs. Grootenboer et al. (2015) use a different definition of middle leading as part of teachers work when facilitating professional learning of others. In this case the middle leader works in a classroom but has an additional task to also lead their colleagues' professional learning. In this study we have observed a third way where middle leading can be viewed as managing and facilitating in the same organization. While the facilitating role match how Grootenboer, Edwards-Groves and Rönnerman define middle leading as closed to the teachers and classroom, the managerial role match Fullan's point on being aware of the school system through goals and local needs.

In the analyses of this study, this became an important outcome in the way the school company had appointed First Teachers at two levels in the organization, where the first teachers as development leaders can be considered middle managers and the first teachers appointed as facilitators are closer to the definition of middle leaders used by Grootenboer et al. (2015). Although acting at different levels in the organization, both practices of middle leading strive for the same goal or project, to promote professional learning among teachers (see Fig. 2). Acknowledging the situatedness of practice (Schatzki, 2002), we do not seek to generalize from our results, but instead put forward an understanding of middle leading as a practice dealing with both managing and leading. This involves for example planning and structuring collegial meetings, using tools and concepts to enhance professional learning, and creating a positive culture of trust among colleagues.

In the literature on middle leading, tension is described connected to the specific role. The middle leader acts at a level between the principal and their colleagues, and is thereby simultaneously handling different practices. Rönnerman et al. (2018) describe these practices as leading and teaching, organizing and facilitating, and cooperating and communicating (see also Grootenboer et al. 2020). In this way, the middle leading practice is conditioned by all these practices and is also enabled and constrained by the practice architectures of these practices. In a review of Bennett (in Harris et al., 2019) middle leaders describe tension connected to the task of leading colleagues. In the current case study, the middle leaders at both levels describe certain tensions that can be connected to their specific role and place in the organization. In the interviews, they talk about these tensions in terms of "expectations" from both leaders as well as colleagues, implicit expectations that the middle leader feels he or she must fulfill. Unlike the middle leaders described in Bennett's review (Harris et al., 2019), the middle leaders at Frida Education don't describe this tension in relation to the task of leading their colleagues. The results show that there is a high level of trust toward the school board, the principals and the middle leaders, which is a probable reason for this.

Trust

Throughout the data, trust with one another and loyalty to the company is obvious. Leaders at all levels in the school company express trust both upward and downward in the hierarchical chain of the organization. Trust is expressed both toward the way different people manage their commission, and to the specific content of the CPD program. [Jederlund \(2019\)](#) claims that trust is a predictor for school development. From the interviews we can see how the respondents talk about a shared vision of the organization that tie the teachers together. One of the development leaders said in the interview: “Despite the fact that we are so many and different school units, we succeed to create and shape what we think is typically ‘Frida’ in all of them”. There is trust to both principals and other leaders in the hierarchy that enable the strategies for the CPD program to be fulfilled. This contributes to the culture where teachers want to collaborate and take responsibility. Furthermore, [Jederlund \(2019\)](#) describes how teachers’ collaboration and collective learning in local CPD is affected by trust on three different kinds: interpersonal relation (collegial trust), in the teacher team (collective trust) and in organizing CPD over time (trust in the process). In another study where a middle leader facilitated teachers, [Edwards-Groves et al. \(2016, 381\)](#) found five dimensions of relational trust, namely interpersonal, interactional, intersubjective, intellectual and pragmatic trust. These dimensions work for CPD that is site based and point out that leading change in schools involves more than just the actual doings. The practice architectures in form of the cultural-discursive, material-economic and social-political arrangements enable and constrain what is happening in the professional learning practice at the school and in the conversations between the development leader and principals and the development leaders and the facilitators. In the school company, there is a chain of regular meetings that enable the leaders and first teachers to stay on track carrying through with the decided CPD content without being hierarchical. In these meetings the conversation goes both ways and develops a shared language on the actual CPD. Relational trust thus seems to be important and from our data we can see that teachers in the middle leading practices (see [Fig. 2](#)) are critical agents for nourishing such a culture in a school and a school company.

The theory of practice architectures - a way of understanding middle leading

In this study, the theory of practice architectures has been used to identify the arrangements that in different ways enable and constrain the project of the middle leading practices, to promote professional learning. The study object has been the way that a specific school company organizes for the middle leading practices. [Harris et al. \(2019, 259\)](#) differentiate between literature related to “middle leader/middle leadership” and “leading from the middle”. The organization of middle leading at Frida education (first teachers acting as development leaders at Didaktikcentrum and first teachers acting as facilitators in the schools) is a conscious strategy to increase capacity and internal cohesion. This can be understood as what both [Grootenboer et al. \(2015\)](#) and [Fullan \(2015\)](#) describes and refers to as “leading from the middle”. In this respect, this study relates to this field, but with a specific focus on the role, practice and behavior of the middle leader, that is a combination of the two perspectives described by [Harris et al. \(2019\)](#). By taking as a starting point the organization of middle leading, and not the middle leader *per se*, this study makes a contribution to the research field as expressed by [Harris et al. \(2019\)](#). It is important to nuance and widen the research on middle leading, since it tends to have a great impact on school through different kinds of reforms. Many studies draw positive conclusions on middle leading, although the empirical base to support these statements is relatively thin, according to [Harris et al. \(2019\)](#). This study tries to answer to this critique by focusing not on middle leading to be implemented in the organization, but on focusing on the organization to understand more about middle leading.

References

- Edwards-Groves, C., Rönnerman, K., 2013. Generating leading practices through professional learning. *Prof. Dev. Educ.* 39 (1), 122–140. <https://doi.org/10.1080/19415257.2012.724439>.
- Edwards-Groves, C., Grootenboer, P., Rönnerman, K., 2016. Facilitating a culture of relational trust in school-based action research: recognising the role of middle leaders. *Educ. Action Res.* 24 (3), 369–386. <https://doi.org/10.1080/09650792.2015.1131175>.
- Fischer, K.B., 2013. Fostering Teacher Learning Communities: A Case-Study of a School-Based Leadership Team’s Action Research. PhD diss. University of Maryland.
- Fullan, M., 2015. Leadership from the Middle. A System Strategy. Education Canada. https://michaelfullan.ca/wp-content/uploads/2015/12/LeadershipfromtheMiddle_EdCan_v55no4.pdf.
- Fridaskolan, 2020. https://fridaskolan.se/wp-content/uploads/2017/10/Riktlinjer_WEBB_uppslag.pdf.
- Grootenboer, P., Edwards-Groves, C., Rönnerman, K., 2015. Leading practice development: voices from the middle. *Prof. Dev. Educ.* 41 (3), 1–19. <https://doi.org/10.1080/19415257.2014.924985>.
- Grootenboer, P., Edwards-Groves, C., Rönnerman, K., 2020. Middle Leadership: Leading Professional Learning. Springer, Singapore.
- Harris, I., Rönnerman, K., 2018. A “deleterious” driver. The “first” teacher reform in Sweden. *Scand. J. Educ. Res.* 63 (5), 805–818. <https://doi.org/10.1080/00313831.2018.1452289>.
- Hargreaves, A., Fullan, M., 2012. Professional capital. In: *Transforming Teaching in Every School*. Teachers College Press, New York.
- Hargreaves, A., Ainscove, M., 2015. The top and bottom of leadership and change. *Phi Delta Kappan* 97 (3), 42–48.
- Harris, A., Jones, M., Ismail, N., Nguyen, D., 2019. Middle leaders and middle leadership in schools: exploring the knowledge base (2003–2017). *Sch. Leader. Manag.* 39 (3–4), 255–277. <https://doi.org/10.1080/13632434.2019.1578738>.
- Jederlund, U., 2019. Tillit som förutsättning för skolutveckling. En studie av skolutveckling genom kollektivt lärande i arbetslag (Trust as a prerequisite for school development. A study of school development through collective learning in teacher teams). *Pedagogisk forskning i Sverige* 24 (3), 7–34.
- Kemmis, S., Grootenboer, P., 2008. Situating praxis in practice. In: Stephen, K., Smith, T.J. (Eds.), *Enabling Praxis Challenges for Education*. Sense Publications, Rotterdam, pp. 37–62.

- Kemmis, S., Wilkinson, J., Edwards-Groves, C., Hardy, I., Grootenboer, P., Bristol, L., 2014. *Changes Practices, Changing Education*. Springer, Singapore.
- Kirsten, N., Carlbam, S., 2020. Kompetensutveckling för professionella lärare?: introduktionen av kollegialt lärande i svensk skola. *Pedagogisk forskning i Sverige* 25 (1), 7–34. <https://doi.org/10.15626/pfs25.01.01>.
- McLaughlin, M.W., Talbert, J.E., 2010. *Building School-Based Teacher Learning Communities*. Teachers College Press, New York.
- Nehring, J., Fitzsimons, G., 2011. The professional learning community as subversive activity: countering the culture of conventional schooling. *Prof. Dev. Educ.* 37 (4), 513–535.
- Opfer, V.D., Pedder, D., 2011. Benefits, status and effectiveness of continuous professional development for teachers in England. *Curric. J.* 21 (4), 413–431.
- Granberg, O., Ohlsson, J., 2016. Vad är kollektivt lärande? In: Granberg, O., Ohlsson, J. (Eds.), *Kollektivt lärande – i arbetslivet*. Studentlitteratur, Lund, pp. 9–18. *Collective Learning*.
- Rönnerman, K., Edward-Groves, C., Grootenboer, P., 2018. Att leda från mitten – lärare driver professionell utveckling (Leading from the middle-Teachers drive professional development). Lärarförlaget, Stockholm.
- Schatzki, T., 2002. *The Site of the Social: A Philosophical Account of the Constitution of Social Life and Change*. University Park. Pennsylvania State University Press.
- SFS, 2010:800. Skollag. Education Act. Utbildningsdepartementet [Ministry of Education, Stockholm].
- SFS, 2013. 70. Karriärvägar för lärare (Teacher career pathways). Utbildningsdepartementet [Ministry of Education], Stockholm.
- Skolverket, 2020. <https://www.skolverket.se/skolutveckling/kurser-och-utbildningar/matematik-kompetensutveckling-i-matematikdidaktik>.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., Thomas, S., 2006. Professional learning communities: a review of the literature. *J. Educ. Change* 7, 221–258.
- Ståhlkrantz, K., 2021. Skolutveckling ur huvudmannaperspektiv (School development from the view of the school authorities' perspective). In: Hirsh, Å., Anette, O. (Eds.), *Skolutveckling i teori och praktik*. Gleerups, Malmö.
- Sulau, V., 2019. Vad händer i lärares kollegiala samtalspraktik? PhD diss. In: *En studie av mötet mellan en nationell kompetensutvecklingsinsats och en lokal fortbildningspraktik*. University of Gothenburg (Gothenburg Studies in Educational Sciences, 437). https://gupea.ub.gu.se/bitstream/2077/60280/2/gupea_2077_60280_2.pdf.
- Timperley, H., 2011. *Realizing the Power of Professional Learning*. Open University Press, Mc Graw Hill House.
- Wenger, E., 1998. *Communities of Practices. Learning, Meaning and Identity*. Cambridge University Press.

The challenge to quality retention: why schools matter

Qing Gu, UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

The retention challenge	539
Why <i>quality</i> retention matters	540
Retention first, recruitment second	540
Teacher quality revisited: purpose, value and beliefs at the core	542
Purpose and values: the soulfulness of quality	542
Efficacious beliefs: psychological functioning that enhances quality	544
Developing teacher quality: why schools matter	544
A social ecological approach to researching teacher quality	544
Why schools matter in developing and sustaining teacher quality	545
Concluding remarks	546
References	546

The retention challenge

For nearly two decades there has been no shortage of policy and research attempts to improve the financial and non-financial attractiveness of teaching in response to stubborn teacher supply challenges across the four UK nations (e.g., [European Parliament, 2019](#); [Department for Education, 2020](#); [EPI, 2021a](#); [Worth and Faulkner-Ellis, 2022](#)). However, despite intense investments, almost 40% of newly qualified entrants to teaching in 2009 have left the profession after 10 years ([Department for Education, 2019](#)) and the overall number of teachers has in general failed to keep pace with increasing pupil numbers since 2011 ([House of Commons, 2021](#)). At a time when the UK education system is still stretched by trying to make up for the effects of Covid lockdowns on pupils' learning, the sector is unlikely to be able to recruit its way out of the teacher supply crisis in a foreseeable future.

Investing in teachers' on-going development has been seen by researchers, thinktanks and policy makers as a cost-effective approach to improving teacher retention ([DfE, 2019](#); [EPI, 2020, 2021b,c](#); [RAND, 2021](#); [PBE, 2022](#); [Perry et al., 2022](#)), and leveling up educational opportunities for every child in all parts of the UK ([IOP, 2020](#); [EEF, 2021](#); [HM Government, 2022](#)). However, to date, extensive efforts to improve schools—especially those serving disadvantaged and marginalized communities—by increasing teachers' human capital (i.e., sum of teachers' knowledge, skills, capabilities and dispositions) appear to have failed to bring about the desired result. Examples include the £355 m funds for strategic school improvement and teaching and leadership innovation in England, the four-year Wellcome CPD Challenge to increase the quantity of subject-specific professional development.

The failure of the teacher-centered human capital strategy to transforming teacher quality and reforming schools is not limited to England alone. In the USA for example, [Johnson \(2019\)](#) questioned the lack of success in reaching its goal of improving student achievement and graduation rates of the intense and costly effort of the Bill & Melinda Gates Foundation's six-year, \$575 million intervention to improve human capital in three large urban school districts and four charter management organizations. This and other UK-wide initiatives suffer from a profound conceptual limitation and have struggled to identify sustained impact on improvement: By focusing narrowly on individual teachers' learning entitlements and increasing the total sum of teachers' knowledge, skills and dispositions, these initiatives have largely ignored the integral role of the school organization in teacher retention. Schools are the finders and keepers of teachers ([Johnson, 2004](#)): teachers who excel and whose students excel with them do not just do so on their own ([Day and Gu, 2010, 2014](#); [Day, 2017](#); [Gu et al., 2021](#)). Johnson argues that

Relying exclusively on efforts to improve human capital in this way rests on the shaky assumption that a teacher's success is independent of the school context in which he works, that a teacher who is effective in one school will be equally effective in any other. It is as if the features of schools that teachers regularly report matter to them—for example, the knowledge and skills of the principal, the effectiveness of schoolwide order and discipline, how time is used, whether they have a curriculum and what it is—have no influence on teachers' practice or their ability to successfully educate their students. According to this view, schools are little more than collections of classrooms, each housing a teacher and her students, who work together unaffected by others.

Johnson (2019: 3–4).

Building and sustaining the quality of teachers is not only an agenda for the future ([Day, 2017](#)), but equally importantly, for the wellbeing, learning and achievement of the children who attend schools today. Improving education quality holds the hope of ending learning poverty for the most disadvantaged and marginalized children and young people around the world. For these children especially, school represents an *oasis of safety and hope* where they can learn, play, grow and achieve. Although schools alone

cannot address many of the centuries' old issues of educational and social inequalities that still challenge many children's fundamental right to quality education in modern times, they are spaces where many committed and caring teachers are dedicated to inspiring the learning and achievement of young minds.

The question of how to retain the hearts, minds and effectiveness of teachers who work in intellectually, emotionally and physically challenging places called "school" (Goodlad, 2004) has been contemplated for many years in many countries. In essence, this is a *quality* retention question. In confronting this question, we encounter more fundamental issues of what constitutes teacher quality, why quality retention matters (as opposed to teachers' *physical* retention), and what schools and systems can do to enable them to learn to live new lives in which they are able to sustain their capacity to teach to their best over time.

Why *quality* retention matters

The important role of a high-quality teaching profession in raising standards and transforming educational outcomes cannot be better emphasized in research papers nationally and internationally. Research on teacher effectiveness consistently reports that teachers' classroom practices have the largest effects on student learning and achievement (Rockoff, 2004; Hallinger, 2005; Rivkin et al., 2005; Leithwood et al., 2006). The positive effects of high-quality teaching are especially significant for pupils from socio-economically disadvantaged backgrounds. Evidence suggests that when taught by very effective teachers, pupils can gain an extra year's worth of learning (Hanushek, 1992, 2011; Sutton Trust, 2011). Internationally, comparative research evidence from the Organization for Economic Co-operation and Development (OECD) affirms that "teacher quality" is the single most important school variable influencing student achievement (OECD, 2005). Indeed, its Teaching and Learning International Surveys (TALIS) have repeatedly concluded that making teaching a more attractive and more effective profession must be the priority in all school systems if they are to secure and enhance effective learning (OECD, 2009, 2011, 2018a).

It remains the case, however, for diverse and complex socio-economic and political reasons, that for many countries retaining and developing committed and effective teachers is a real challenge. In many low- and medium-income countries, for example, where school enrollment is on the rise, an acute shortage of primary teachers represents one of the greatest hurdles to providing education for all school-age children (UNESCO, 2011, 2014, 2015). A lack of resources and financial incentive packages to attract qualified personnel into teaching has meant that quantity, rather than quality, continues to be a primary concern in their efforts to provide basic education. This has meant that, unfortunately, children in countries needing teachers the most, tend to be taught by the least qualified personnel (UNESCO, 2006, 2015, 2020).

In contrast, in the world of high-income countries, such as the USA, the UK and many European countries, shortage of teacher supply tends to be a particularly pressing problem for core subject areas such as maths, modern foreign languages and science (European Commission, 2012, 2014; European Parliament, 2019) and for schools serving socioeconomically deprived communities (Ingersoll, 2001; Guarino et al., 2006; Boyd et al., 2008; Allen and McInerney, 2019). There are also troubling indicators which suggest that leadership turnover is especially higher and teacher quality is especially lower in schools serving high-need communities (Loeb et al., 2005; Boyd et al., 2008; Goldhaber and Hansen, 2009; Sibiet, 2020) where most children, who are already disadvantaged in accessing or benefiting from rich capital and social capital in their early years, are then denied access to the quality education to which they are entitled when entering the formal school system.

Retention first, recruitment second

Concerns about teacher supply and quality are not new. Significant falls in the proportion of graduates applying for teacher training programs across many European countries have culminated in an urgent call to increase efforts to transform the conditions of teaching and through this, attract more suitably qualified people to the profession (Auguste et al., 2010; OECD, 2011; European Commission, 2012; European Parliament, 2019). At the same time, the aging population of the existing teaching workforce (Grissmer and Kirby, 1997; Guttman, 2001; Chevalier and Dolton, 2004; OECD, 2005; Matheson, 2007; Aaronson, 2008; European Commission, 2012, 2014) poses pressing challenges to the nature of its composition in the future. Only 7% of all primary and secondary teachers in the EU are under 30 years old, while around more than 1 in 3 (36%) are 50 or older (European Parliament, 2019). In England, for example, close to half of the full-time teachers (46%) are aged over 40, with 19% of these aged over 50 (Department for Education, 2020). This situation is the most pronounced in Italy where more than half of all teachers are older than 50 (European Commission, 2012, 2014; European Parliament, 2019).

Moreover, high rates of attrition of teachers in their first five years of teaching (Darling-Hammond, 1997; Ingersoll, 2003; Kados and Johnson, 2007; Burghes et al., 2009; Shen and Palmer, 2009; OECD, 2005, 2011; Department for Education, 2020) remain a persistent teacher retention problem. Although some studies show that on average, early career leavers tend to be less effective than stayers as measured by the test score gains of the students in their classrooms (Henry et al., 2011; Goldhaber et al., 2011; Boyd et al., 2011), others suggest that teachers with stronger qualifications and more competitive university backgrounds are more likely to exit early (Lankford et al., 2002; Boyd et al., 2005; Guarino et al., 2006; Feng and Sass, 2011).

The reasons behind teachers' decision to leave are complex. A common critique has, at least in part, attributed teacher attrition to a whole-sale redefinition of teacher professionalism driven by "a culture of accountability, performance, and measurability" (Luke, 2011, p. 370; see also Rots and Aelterman, 2008; Smith and Kovacs, 2011). Luke (2011) laments that "the normative, the ethical,

the cultural—matters of value—have quietly slipped from policy discussion (Ladwig, 2010), overridden by a focus on the measurable, the countable, and what can be said to be cost efficient and quality assured” (2011, p. 368). Within such a performativity culture, teacher professionalism has become more closely aligned with national educational policy which tends to define educational success in relatively narrow, instrumental terms (Furlong, 2008). Alongside this, the power of government regulatory bodies for the setting and adjudication of standards has been increased significantly over time. This has been complemented by more detailed and bureaucratic monitoring of what teachers do in their classrooms and how they do it—irrespective of whether what they are required to do is educationally and/or culturally meaningful (Luke, 2011). Thus, whether or not teachers agree with the centrally prescribed policy agendas and strategies, they are expected to conform to them in their day-to-day practice (Furlong, 2008).

It is perhaps, then, no surprise that this highly prescriptive culture of neoliberal accountability and performativity has been criticized by scholars as having contributed to a wide-spread lack of deep trust in teachers’ professional standing, judgment and capability. For some years now, a substantial number of studies on teacher retention have been overly concerned with a narrow problem-focussed agenda. Although this body of research has improved our understanding of the factors which may cause some teachers to struggle and/or exit, it offers rather limited explanation as to why so many other teachers who are working with similar challenges embedded in outcomes driven educational systems are willing, able and committed to continue to teach to their best. What tends to be absent from many of the investigations of teacher stress, attrition and job satisfaction of individual teachers is the integral role that teachers’ professional, role and organizational identities, wellbeing and sense of vocation play in enabling them to meet the daily challenges of teaching and learning (Day and Gu, 2014); and the part played by school leaders in mediating the sometimes negative effects of educational policies and through this, shaping and influencing many, if not all, teachers’ sense of job fulfillment, commitment and effectiveness (Gu et al., 2018).

The positive impact of strong leadership on student learning through building supportive school culture and creating favorable working conditions for teachers is well documented in the teacher development, school improvement and school effectiveness literature (Hallinger, 2005; Johnson, 2004; Leithwood et al., 2004, 2006; Gu et al., 2008; Day et al., 2011; Sammons et al., 2011). Griesom’s (2011) study found that higher teacher turnover rates in disadvantaged schools result in part from the ineffectiveness of the principal. His analysis of national school and teacher surveys showed that teacher satisfaction is likely to be lower, and that the probability for teachers to leave schools is significantly greater when the leadership of the principal is weak and ineffective.

Recent research on scaling up innovations for better teaching and learning outcomes in England also points to the significance of leadership in enabling sustained teacher professional learning, development and change (Gu et al., 2019, 2020, 2021). The research shows that leadership support for professional learning and development is a prerequisite for change in school culture which shapes the intellectual, social and organizational environments that are necessary for change in teacher practice and improvement in pupil engagement in learning. Put simply, “A ‘good’ research-informed innovation can rarely travel into the day-to-day realities of classrooms on its own merits without school leaders that can help teachers engage with it and apply and adapt it to their own classroom contexts” (Gu et al., 2021: 19). In line with these research findings is the observation of the strong and positive associations between school leaders’ administrative support and low teacher retention rates (Ladd, 2009; Boyd et al., 2011). A strong sense of staff collegiality has been found to be crucial in building intellectual, emotional and social capital in schools so that teachers, and especially those working in schools serving socioeconomically deprived communities, are able to maintain their integrity and commitment in times of change (Gu and Day, 2007; Allensworth et al., 2009; Day et al., 2011; Holme and Rangel, 2012).

In his reflection of his leadership journey in education, Sir Dunford (2016) argued that schools and the education system at large should place “retention first” and “recruitment second” in their endeavor to secure an effective workforce to meet the learning entitlements of every student in every school in every country of the world. It would indeed be more fruitful and educationally more meaningful if greater attention were paid to the factors which enable teachers who decide to stay to maintain committed to their own learning and the learning and achievement of their pupils. This is, in essence, a quality retention issue because, as Johnson et al. (2005) have argued, the physical retention of teachers, “in and of itself, is not a worthy goal”:

Students are not served well when a district retains teachers without regard to quality. Little can be achieved (and much might be lost) when a district succeeds in reducing teacher turnover if some of those teachers are incompetent, mediocre, disengaged, or burnt out. Instead, student learning is the goal, and schools must seek to retain teachers who demonstrate that they are skilled and effective in the classroom, are committed to student learning, and are ready and able to contribute to the improvement of their school.

Johnson et al. (2005, p. 2).

However, we all know that to teach at one’s best over time is not easy. In reflecting on her professional work with teachers over the course of her career, Nieto’s (2011) expressed the greatest respect for such teachers: “My belief in teachers is stronger than ever because I have seen the best of them do unbelievable work in sometimes harsh circumstances” (2011, p. 133). These are the teachers who give witness to the essential meaning of “everyday resilience” that Day and Gu (2014) have elaborated in their work on the work, lives and effectiveness of teachers.

Over the years, scholars have used different conceptual and methodological lenses to explore issues around improving the quality of provision of teaching and learning in schools. We have chosen *resilience* in our recent work because it provides a useful conceptual lens for teachers, school leaders and policy makers to understand how and why many teachers have managed to weather the often unpredictable “storm” of school and classroom life (Patterson and Kelleher, 2005) and sustain their commitment and motivation in the profession. It also enables us as researchers to probe deep into teachers’ inner and external professional worlds

to explore why many are able to remain committed and passionate about making a real difference and continue to do so—irrespective of the unpredictable nature of every school and the many physical, emotional and intellectual challenges that are associated with this. More importantly, we know from research that pupils of highly committed and resilient teachers are more likely to perform better academically (Day et al., 2007).

In the remainder of the chapter, I associate teacher quality with the purpose, value and beliefs in teachers' inner worlds which enable them to endure the persistence of hope and endeavor to do their best to reach and engage every student in learning, regardless of the challenges which this may bring. I will do so in the knowledge that teacher quality is not fixed as individuals' capacity to express the conviction of their vocational commitment is almost certain to fluctuate over the course of a career. The extent to which individual teachers are able to exercise their capacity to be committed and resilient will depend not only upon their individual histories and their personal resources, but also upon the influence of the school environment, their colleagues and the quality of their school leadership.

Teacher quality revisited: purpose, value and beliefs at the core

There is strong and consistent research evidence that teachers' classroom practices have the largest effects on student learning achievement (Rockoff, 2004; Hallinger, 2005; Rivkin et al., 2005; Leithwood et al., 2006; Barber and Mourshed, 2007; Mourshed et al., 2010; Leithwood, 2019). At the systems level, the two McKinsey's reports (Barber and Mourshed, 2007; Mourshed et al., 2010) show that getting the right people into the teaching profession and developing them to become effective teachers have played a central role in enabling the world's most improved school systems to come out on top, and more importantly, keep getting better. Put simply, teacher quality is the most critical school-level and system-level predictor of student success.

Depending on the theoretical and methodological interests and preferences of educational researchers, the definitions and measures of teacher quality vary considerably. At times, the meaning of *teacher* quality is used interchangeably with that of *teaching* quality. For example, in some studies where the research focus was on investigating increases in the quality of teaching, value-added scores and learning gains were used to measure and categorize the quality of teachers rather than what they do in the classroom (e.g., Chetty et al., 2014; Bowen and Mills, 2017).

It is necessary to make a distinction between teacher quality and teaching quality because, as Kennedy (2010) argues, "the qualities teachers bring with them to their work are not always enough to ensure better teaching practices" (2010: 591). A profound difference is that the former is essentially concerned about the person in the professional teacher, while the latter is about the practice of the professional. Put differently, teacher quality concerns *who the teacher is*, *what being a teacher means to the person*, and *what the values, capacity and, capabilities they have* that influence, positively or negatively, student learning and achievement. In contrast, teaching quality is primarily concerned with *what the teacher does in the classroom* (i.e., processes and activities of instruction) and *how well they do it* (i.e., impact of instruction on student learning). Thus, quality teaching as against quality teachers will be influenced, as Kennedy (2010) reminds us, not only by the "enduring personal qualities that they bring with them" but will also be "a function of schedules, materials, students, institutional incursions into the classroom, and the persistent clutter of reforms that teachers must accommodate" (2010: 597).

Teacher quality has remained central to policy debate about teacher supply and demand in many countries. This is because, at least in part, teacher quality attends to the ethical and vocational values, passion and resilience, as well as cognitive, social and psychological resources and dispositions of the person *in* the teacher—all of which influence, powerfully and profoundly, teachers' capacity to teach well over the course of their professional lives. Their commitment, enthusiasm and resilience—qualities that are harder to measure—are fundamental to how teachers feel about their work, how they think about themselves as professionals, and importantly, how they are (or are not) able to fulfill their professional values and core purposes by making a real difference to children's learning and achievement.

In this chapter, I focus on two inner resources of teacher quality: *altruistic values* and *efficacious beliefs*. The reasons are twofold. First, these inner resources are expressions of the personal values and strengths of the *person* in the professional teacher that define the ethical character and professional quality of the teacher. Second, and related to the first, they enthuse and enable teachers to improve their capabilities to fulfill the personal and social meaning of teaching, and therefore explain *why* some can "maintain equilibrium and a sense of commitment and agency in the everyday worlds in which teachers teach" (Gu and Day, 2013, p. 26) while others cannot.

Purpose and values: the soulfulness of quality

In his seminal work on teaching as a calling and vocation, Hansen (1995, 2021) argues persuasively that the idea of vocation underscores the fact that it is not the teaching *role* itself but "the person within the role and who shapes it who teaches students" (1995: 17). In describing teaching a calling and vocation, Hansen invites us to probe deeply into understanding the soulfulness and passion in teaching as well as the social origins of the *practice* of teaching which embodies teachers' esthetic (e.g., having a feeling and a sense for work done well), moral, and intellectual endeavors and inspirations to enact their inner urge to make a difference to the learning and achievement of their students. The language of vocation takes us *inward* to recognize that when the person is called to "*be* a teacher rather than just to *do* or *perform* a job that others might easily accomplish in their place (Sherman, 2013, 2020)"

(Hansen, 2021: 23), we see the identity of the person in the teacher and get closer to understand why for many teachers, teaching is much larger than a job.

The sense of vocation finds its expression at the crossroads of public service and personal fulfilment. It takes shape through involvement in work that has social meaning and value. ... Vocation is also expressed over time. ... a person cannot will a sense of service into existence nor wake up one day and decide to be of service. Those dispositions grow and take shape over time, through interaction with people and through the attempt to perform the work well.

Hansen (2021: 3–4).

What follows is that the person's altruistic values and principles shape the ways in which they derive the social, intellectual and moral meanings from their role as a teacher. Recognizing how these inner meanings motivate teachers to improve their ability to teach well is integral to understanding why the teaching profession is regarded by many committed and dedicated teachers as a "life-style" (Day and Gu, 2014).

Altruism as a philosophical concept has long been used in education to describe an ethical principle and moral purpose that drive many teachers to build caring relationships with their students (Nias, 1989; Noddings, 1992), and remain committed and passionate about giving their best to the students whom they serve (Ayers, 2010; Day, 2004; Gu and Li, 2013; Hansen, 1995; Nieto, 2003, 2015). Altruistic principles and values are internal to the calling that has drawn many teachers into the practice of teaching. However, to view teaching as an "altruistic vocation" (Schwarz and Alberts, 1998: 155) is not, as skeptics might argue, an unsustainable "ethical ideal" which requires teachers to sacrifice a high degree of selflessness (e.g., Higgins, 2011). Nor does it deny that teachers' material needs (such as pay and income) are important. Our research on teachers' subjective wellbeing in China shows that teachers' altruistic values have no direct associations with their income satisfaction, but positive and direct associations with their satisfaction with work (Song et al., 2020). When teachers' salaries have met the "basic threshold" of income, those with stronger altruistic values and higher efficacious beliefs tend to report smaller effects of income on their work satisfaction.

Over the last decade evidence from our own research in England and China and international surveys of teachers (OECD, 2014, 2018b) repeatedly shows that the majority of teachers in the profession are still hopeful, committed and passionate about making a difference to the learning and lives of the children—for whom they care and feel responsible. Bullough and Pinnegar (2009) argue that it is in moments when teachers were meeting their ethical obligations to children that they felt elevated and were reminded of why they were first "called to teach" (2009: 246).

Therefore, in our analysis of teachers' work and lives, we consider altruism as the essence of teachers' ethical values, principles and obligations, which according to Hansen (1995), enable them to see work as a "service to others" and gain "personal satisfaction in the rendering of that service" (1995, p. 3). This view "turns the focus of perception in such a way that the challenges and the complexity in teaching become sources of interest in the work, rather than barriers or frustrating obstacles to overcome" (Hansen, 1995: 144). Such a view of teaching leads to a positive and productive emphasis on learning, growth and development—which regards teachers' need for intellectual challenge and capacity building as an indispensable part of their moral responsibility. It reveals profoundly that the good life of teaching entails both an *intellectual* act and a *moral* enterprise (Huberman, 1993; Palmer, 2007; Day and Gu, 2010, 2014; Goodlad et al., 1990; Hansen, 2021). It is this inner meaningfulness—intellectually, emotionally and spiritually—that influences, deeply and powerfully, teachers' own sense of identity and wellbeing.

Becoming a wonderful teacher, or a great or awesome teacher, is a lifetime affair. This is because good teaching is forever pursuing better teaching; it is always dynamic and in motion, always growing, learning, developing, searching for a better way. Teaching is never finished, never still, never easily summed up. "Wonderful Teacher" might be inscribed on someone's lifetime achievement award, printed on a retirement party banner, or etched on a tombstone, but it is never right for a working teacher.

Ayers (2010: p. 16).

From the perspective of social change and evolution, Welzel and Inglehart's (2010) seminal work on values and wellbeing reminds us of the importance of considering the social markers and cultural sources of values when investigating teachers' internal worlds. This is because "humans internalize most of their values fully unconsciously in an unquestioned process of socialization"—through which they are "familiarized with what is socially accepted in their society" (Welzel and Inglehart, 2010, p. 47). For example, in the Chinese culture and society where the self is viewed as "a connected, fluid, flexible, and committed being who is bound to others" (Lu, 2010, p. 335), and where the prevalent Confucian moral discourse has long stressed the importance of suppressing selfish desires to serve the collective, the origins of Chinese teachers' sense of purpose and meaningfulness as well as intellectual and social connections in their worlds of work reside not only in their personal motives and goals, but more profoundly, the fulfillment of their moral obligations as a professional for the society. In this sense, Chinese teachers' satisfaction with the quality of their professional lives encapsulates the attainment of purpose and goals that meet both their own intrinsic professional needs and those of the larger professional community and society—to which they belong. The identification of culture-specific values in shaping Chinese teachers' perceptions of the meaning and goals of their work highlights the need for a deep understanding of *what matters to whom* and *in what contexts* in future research on teachers and teaching.

Efficacious beliefs: psychological functioning that enhances quality

Bandura's (1986) social cognitive theory defines perceived self-efficacy as individuals' beliefs about their capabilities to exercise control over events that affect their lives, to produce expected performance, and to influence others. People with strong beliefs in their capabilities make things happen by setting themselves challenging goals, thinking strategically in the face of difficulties, and remaining resilient, motivated and task-focused (Bandura, 1997). Bandura (1997) thus rejects the view that efficacy beliefs are mere inert predictors of performance accomplishment. Rather, he argues that they are "a vital personal resource" (2000: 120) which function as a key contributor to effective functioning and human flourishing.

When faced with obstacles, setbacks, and failures, those who doubt their capabilities slacken their efforts, give up, or settle for mediocre solutions. Those who have a strong belief in their capabilities redouble their effort to master the challenges.

Bandura (2000: 120).

Bandura's argument, together with findings from the latest research on teachers' self-efficacy (e.g., Klassen et al., 2011; Renshaw et al., 2015; Zee and Koomen, 2016), offers conceptual and empirical ground for the argument that teachers' efficacy beliefs are their self-referent judgments of their own capacity and capability to perform at certain levels (Kelley and Finnigan, 2003; Zee and Koomen, 2016). Over the last thirty years research has consistently reported that self-efficacious teachers tend to suffer less emotional exhaustion and burnout symptoms, and experience more confidence in mastering challenging tasks and higher levels of responsibility for teaching, commitment, personal accomplishment, and job satisfaction (e.g., Klassen and Chiu, 2010; Pillay et al., 2005; Skaalvik and Skaalvik, 2010; Schwerdtfeger et al., 2008; Zee and Koomen, 2016; Song et al., 2020). These findings lend support for the view that the strengths of teachers' judgments of their performance capability enhance, or constrain, the strengths of their aspiration, commitment, and effort to fulfill different task demands in given circumstances and through these, influence, positively or negatively, their emotional outlook and satisfaction with the quality of their work and life.

Values and efficacious beliefs are both key aspects of psychological functioning (OECD, 2013). Our recent research on teachers' wellbeing (Song et al., 2020) offers additional evidence emphasizing the importance of psychological functioning in explaining human flourishing (see also, Diener et al., 2010; OECD, 2013). Although strong altruistic values can fulfill teachers' satisfaction of "becoming the kind of teacher and person they desire to be," (Bullough and Pinnegar, 2009, p. 246), the enactment and fulfillment of these values is mediated by their efficacious beliefs. On the one hand, purpose and meaning, as "powerful regulators of human behavior" (Welzel and Inglehart, 2010, p. 47), empowers (or limits) individuals' efficacious beliefs in their capacity to perform proficiently in context. On the other hand, teachers need to feel efficacious in what they do in order for there to be a positive effect on work satisfaction from their strong values. Put another way, only when teachers' care and love for students are enacted in ways which enable them to master the challenges and thus continue to teach to their best can they experience higher levels of satisfaction with the quality of their work as a teacher.

Developing teacher quality: why schools matter

A social ecological approach to researching teacher quality

Over the years, we have worked with teachers also who, despite their hard work, commitment, moral purpose and dedication to the education of their students, struggle to improve outcomes. Those teachers who excel and whose students excel with them do not just do so on their own. While it is clearly the responsibility of each teacher to teach to their best, it is the responsibility of each individual school, school district and national government to ensure that they are able to do so through high-quality leadership and the provision of physical and other resources. Strong and consistent evidence from educational research suggests that the social and intellectual organization of the school, and by extension, schools' connections with local communities, society, and the policy structure—when characterized with supportive and trusting relationships between different stakeholders—fosters teachers' collective capacity, commitment, resilience and effectiveness (Bryk and Schneider, 2002; Day and Gu, 2010, 2014; Sammons et al., 2007; Tschannen-Moran and Barr, 2004).

Such evidence invites us to use a social-ecological approach to analyze how the environments in which teachers work and live influence, and are influenced by, their capacity to teach well over the course of their professional lives. The relevance of the social-ecological approach to researching teachers' quality are at least threefold.

First, an *environment-centered approach* to human development reinforces the importance of emphasizing the impact of multilevel contexts on the growth of teachers over the course of their professional lives, especially in terms of their capacity to maintain a sense of commitment and agency in the everyday worlds in which they teach. The theoretical underpinning of the social ecology of teachers' work, lives and professional development emerged from Bronfenbrenner's (1979) ground-breaking work on human development which is, in essence, concerned with the interconnectedness and interactions between multilevel systems and the ways they shape the course of human development throughout the life span. Bronfenbrenner defined human development as "the person's evolving conception of the ecological environment, and is relation to it, as well as the person's growing capacity to discover, sustain, or later its properties" (1979, p. 9).

The same argument applies to teacher development. The quality of the reciprocating systems in education influences the quality of the intellectual, social and organizational conditions in which teachers work—which in turn impact, positively or negatively, on their professional identities and their capacity to be committed, efficacious and effective (Belman, 2015; Day and Hong, 2016; Gu, 2014; Gu and Day, 2007, 2013; Johnson et al., 2016). Among these, the environments of schools are found to be particularly significant in shaping why and how many teachers are able to continue to commit their time, energy and passion to a profession which can make them “feel a sense of invisibility and powerlessness” (Nieto, 2015, p. 252; see also Gu and Johansson, 2013; Johnson, 2004; OECD, 2016; UNESCO, 2015). I will unpack why schools and school leadership especially, matter to teacher development and teacher quality in more detail later.

Second and related to the first, the social ecological model of person $\times$ environment interaction (Ungar et al., 2013) enables us to place teachers in their complex worlds of work and analyze the ways in which their capacity to teach to their best *influences* and is *influenced* by their professional worlds. Unlike children whose individual resources (e.g., optimism and efficacy) are believed to be “only as good as the capacity of his or her social and physical ecologies that facilitate their expression and application to developmental tasks” (Ungar, 2011: 6), teachers are “moral and political agents” (Soder, 2004: 11) whose moral values form the hub of a school’s culture which energize teachers’ collaboration, learning and development. For teachers, if their role expectations were primarily defined by prescribed standards, methods and techniques, then these expectations would have missed a simple but fundamental truth about “a good teacher”:

Good teachers possess a capacity for connectedness. They are able to weave a complex web of connections among themselves, their subjects and their students so that students can learn to weave a world for themselves. ... The connections made by good teachers are held not in their methods but in their hearts—meaning heart in its ancient sense, as the place where intellect and emotion and spirit and will converge in the human self.

Palmer (2007: 11).

The voice of the *heart* flows from the identity and integrity of the inward teacher (Palmer, 2007) and speaks out why many teachers in many countries regard teaching as “a way to live in the world” (Nieto, 2015, p. 1). As I have elaborated earlier in this chapter, understanding the role of moral values and vocation in influencing what many teachers do and why they do it in their schools and classrooms is key to understand the dynamic and complex nature of the interaction between the teacher and their context of work. Deal and Peterson (2009) argue that “like a butterfly, a school must be nurtured by its inner energy in order to thrive” (2009: 180). This reinforces the observation that teachers’ actions, commitment and beliefs also influence, individually and collectively, the social and organizational culture and contexts in which they work and live. Buchanan (2015) found in her research on teacher identity and agency that the discourse of education policy and structures do not completely redefine teachers’ self-concepts: “Teachers are more complex and multifaceted than that” because individual teachers bring with them “a unique mix of personal and professional experiences and commitments” (Buchanan, 2015: 700):

Teachers therefore confront the policies and professional discourses they encounter not as *tabulae rasae*, but rather actively use their own preexisting identities to interpret, learn from, evaluate, and appropriate the new conditions of their work in schools and classrooms. In this process, their identities are reformed and remade—and professional agency is carved out.

Buchanan (2015: 700).

Last but not least, the social ecological understanding of teachers’ work and lives emphasizes a *process-oriented, developmental-contextual* approach to viewing and supporting teacher development. Teacher quality, or put differently, teachers’ capacity to develop and sustain their qualities to teach well is nurtured, learned and acquired in their context of work. Over time such capacity unfolds progressively in a developing individual (Schoon, 2012) as they learn to navigate and negotiate their way to the intellectual, social/relational and technical resources in their immediate external environment that enable them to continue to learn, to grow, and to enact their values and capacity to teach well. For many teachers who have managed to sustain their commitment and motivation in the profession, the ability to weather the often unpredictable “storm” of school and classroom life (Patterson and Kelleher, 2005) is not an option, but a necessity. Sustained pursuit of continuing professional learning and development is what makes good teachers and great teaching. This is a career-long moral commitment and “a sustainable investment for professional capital” (Hargreaves and Fullan, 2012: 186).

Why schools matter in developing and sustaining teacher quality

The conceptual strengths of the social ecological approach to researching teachers’ work and lives highlight the significance of schools as “finders and keepers” (Johnson, 2004) of committed, resilient and quality teachers. Despite the reform and societal pressures on teachers, research consistently shows that three interrelated conditions—teachers’ vocational selves, social and professional relationships with colleagues, and leadership support and recognition—are found to be integral in enabling them to sustain their educational purposes and successfully manage the “unavoidable uncertainty” (Shulman, 1987, p. 1) inherent in the everyday life of a teacher (Gu and Day, 2013). Among the three, school leadership comes first and acts as a necessary organizational condition that enables teachers to fulfill their professional commitments in collaboration and partnership with their colleagues.

Recent research evidence on teacher retention shows that teacher leavers are not necessarily “escaping” from pupils’ poor behavior. Rather, they are escaping from poor leadership and dysfunctional school cultures. Toxic cultures devalue our need of others, impede our ability to turn to them for support, and challenge our capacity to form supportive and collegial relationships to learn and develop (Jordon, 2006, 2012). At a result, with a heavy heart, we continue to observe in our research that some teachers seeing teaching behind closed doors as a reality of the profession. Many of these teachers, more often than not, have struggled to understand why their sheer hard work has failed to bring about the levels of progress and achievement that they would like to see in their students (Matthews et al., 2014).

Our decades of research on school leadership and school improvement shows that principal leadership is instrumental in bringing about improved learning outcomes in schools. They contribute to student learning largely *indirectly* through leadership activities and influence, particularly through building school capacity and (re)designing structural, sociocultural, and relational processes that are conducive to promoting professional learning communities and raising the quality of teaching and learning (Bryk et al., 2010; Day et al., 2011; Gu and Johansson, 2013; Hallinger and Heck, 2010a,b; Leithwood et al., 2010, 2017; Sun and Leithwood, 2012, 2015). The quality and professional capability of leadership also defines the extent to which individual schools are to be engaged with externally initiated improvement effort and the ways in which the new learning is to be aligned with, as well as advance, the existing culture and capacity for change and sustained improvement.

In this process, school leaders, and school principals especially, are the *architects* of social relations and learning cultures in their schools. Their ability to drive professional development and improvement in their own schools (leader self-efficacy) and to develop sustained collaboration with other leaders and schools (leader collective efficacy) shapes the *base* capacity to be engaged with improvement initiatives within local school systems and “the *new* capacities that must be developed to sustain and extend these initial efforts over time” (Bryk et al., 2010, p. 220, italics in original; also Leithwood and Louis, 2012). It is perhaps then no surprise that in high-performing schools a key defining aspect of principal leadership is an unrelenting focus on fostering consistent values, expectations, and standards, and through these, empowering and transforming staff capacities and organizational conditions to embrace change and improvement.

Concluding remarks

A central task for all concerned with enhancing quality and standards in schools is, therefore, not only to have a better understanding of what influences teachers’ quality over the course of a career, but also the means by which the teacher quality necessary for these to be sustained may be nurtured in the contexts in which they work. Promoting and cultivating healthy individual and collective learning and achievement cultures in schools is essential to how they feel about themselves as professionals.

Establishing consistency and coherence in school structures, cultures, and improvement processes holds the key to engaging the heads, hearts and hands of the school community to achieve sustained and sustainable performance over time. Such consistency and coherence were reflected in understandings—between those who led and those who were led—of how and why their schools were able to become successful and, as importantly, to stay successful. They helped to shape the day-to-day practice of school life and enabled schools in our research to remain focused on the core business of schools—quality provision of teaching and learning—irrespective of external pressures of change (Gu et al., 2018).

References

- Aaronson, D., 2008. The Impact of Baby Boomer Retirements on Teacher Labor Markets, vol. 254. Chicago Fed Letter. September. http://www.chicagofed.org/publications/fedletter/cflseptember2008_254.pdf.
- Allen, B., McInerney, L., 2019. The Recruitment Gap: Attracting Teachers to Schools Serving Disadvantaged Communities. Sutton Trust, London.
- Allensworth, E., Ponisciak, S., Mazzeo, C., 2009. The Schools Teachers Leave: Teacher Mobility in Chicago Public Schools. Consortium on Chicago School Research, Chicago. Retrieved 1st April 2012, from: http://ccsr.uchicago.edu/publications/CCSR_Teacher_Mobility.pdf.
- Auguste, B., Kihn, P., Miller, M., 2010. Closing the Talent Gap: Attracting and Retaining Top-Third Graduates to Careers in Teaching. McKinsey & Company, New York.
- Ayers, W., 2010. To Teach: The Journey of a Teacher, third ed. Teachers College Press, New York, NY.
- Bandura, A., 1986. Social Foundation of Thought and Action: A Social Cognitive Theory. Prentice Hall, Englewood Cliffs, NJ.
- Bandura, A., 1997. Self-Efficacy: The Exercise of Control. W. H. Freeman, New York, NY.
- Bandura, A., 2000. Cultivate self-efficacy for personal and organizational effectiveness. In: Locke, E.A. (Ed.), Handbook of Principles of Organization Behavior. Blackwell, Oxford, UK, pp. 120–136.
- Barber, M., Mourshed, M., 2007. How the World’s Best Performing School Systems Come Out on Top. McKinsey and Company, New York, NY.
- Beltman, S., 2015. Teacher professional resilience: thriving not just surviving. In: Weatherby-Fell, N. (Ed.), Learning to Teach in the Secondary School. Cambridge University Press, Melbourne, Aust, pp. 20–38.
- Bowen, D., Mills, J., 2017. Changing the education workforce? The relationships among teacher quality, motivation and performance pay. Teach. Coll. Rec. 119, 1–32.
- Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J., 2005. Explaining the short careers of high-achieving teachers in schools with low-performing students. American Economic Review, 95, 166–171.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., Wyckoff, J.H., 2008. Who Leaves? Teacher Attrition and Student Achievement. NBER Working Paper No. W14022.
- Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., Wyckoff, J., 2011. The role of teacher quality in retention and hiring: using applications-to-transfer to uncover preferences of teachers and schools. J. Pol. Anal. Manag. 30 (1).
- Bronfenbrenner, U., 1979. The Ecology of Human Development: Experiments by Nature and Design. Harvard University Press, Cambridge, MA.
- Bryk, A.S., Schneider, B.L., 2002. Trust in Schools: A Core Resource for Improvement. Russell Sage Foundation, New York.
- Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., Easton, J., 2010. Organising Schools for Improvement: Lessons From Chicago. University of Chicago Press.

- Buchanan, R., 2015. Teacher identity and agency in an era of accountability. *Teach. Teach. Theor. Pract.* 21 (6), 700–719.
- Bullough, R., Pinnegar, S., 2009. The happiness of teaching (as eudaimonia): disciplinary knowledge and the threat of performativity. *Teach. Teach. Theor. Pract.* 15 (2), 241–256.
- Burghes, D., Howson, J., Marenbon, J., O'Leary, J., Woodhead, C., 2009. *Teachers Matter: Recruitment, Employment and Retention at Home and Abroad*. Politeia, London.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014. Measuring the impacts of teachers ii: teacher value-added and student outcomes in adulthood. *Am. Econ. Rev.* 104 (9), 2633–2679.
- Chevalier, A., Dolton, P., 2004. Teacher Shortage: Another Impending Crisis? *CentrePiece*, pp. 15–21. Winter.
- Darling-Hammond, L., 1997. *Doing What Matters Most: Investing in Quality Teaching*. National Commission on Teaching & America's Future, New York.
- Day, C., 2004. *A Passion for Teaching*. RoutledgeFalmer, London and New York.
- Day, C., 2017. *Teachers' Worlds and Their Work*. Routledge, London and New York.
- Day, C., Gu, Q., 2010. *The New Lives of Teachers*. Routledge, Abingdon, Oxon.
- Day, C., Gu, Q., 2014. *Resilient Teachers, Resilient Schools: Building and Sustaining Quality in Testing Times*. Routledge, London.
- Day, C., Hong, J., 2016. Influences on the capacities for emotional resilience of teachers in schools serving disadvantaged urban communities: challenges of living on the edge. *Teach. Teach. Educ.* 59, 115–125.
- Day, C., Sammons, P., Stobart, G., Kington, A., Gu, Q., 2007. *Teachers Matter: Connecting Lives, Work and Effectiveness*. Open University Press, Maidenhead.
- Day, C., Sammons, P., Leithwood, K., Hopkins, D., Gu, Q., Brown, E., Ahtaridou, E., 2011. *School Leadership and Student Outcomes: Building and Sustaining Success*. Open University Press, Maidenhead.
- Deal, T.E., Peterson, K.D., 2009. *Shaping School Culture*, second ed. Jossey-Bass, San Francisco, CA.
- Department for Education, 2019. *School Workforce in England: November 2019*. DfE, London.
- Department for Education, 2020. *School Workforce in England*. <https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england>.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., Biswas-Diener, R., 2010. New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc. Indic. Res.* 97 (2), 143–156.
- Dunford, J., 2016. *The School Leadership Journey*. John Catt Educational Ltd., London.
- Education Endowment Foundation, 2021. *Effective Professional Development: Three Recommendations for Designing and Selecting Effective Professional Development*. EEF, London.
- Education Policy Institute, 2020. *The Effects of High-Quality professional Development on Teachers and Students: A Rapid Review and Meta-analysis*. Education Policy Institute, London.
- Education Policy Institute, 2021a. *Policy Analysis: Teacher Recruitment and Retention in the Eye of the pandemic*. Education Policy Institute, London. <https://epi.org.uk/publications-and-research/recruitment-and-retention-in-the-eye-of-the-pandemic/>.
- Education Policy Institute, 2021b. *The Cost of High-Quality professional Development for Teachers in England*. Education Policy Institute, London.
- Education Policy Institute, 2021c. *The Effects of High-Quality professional Development on Teachers and Students: A Cost-Benefit Analysis*. Education Policy Institute, London.
- European Commission, 2012. *Key Data on Education in Europe 2012*. Brussels: Education. Audiovisual and Culture Executive Agency.
- European Commission, 2014. *Study on Policy Measures to Improve the Attractiveness of the Teaching Profession in Europe*. European Commission.
- European Parliament, 2019. *Teaching Careers in the EU*. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/642220/EPRS_BRI\(2019\)642220_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/642220/EPRS_BRI(2019)642220_EN.pdf).
- Feng, L., Sass, T., 2011. Teacher Quality and Teacher Mobility. National Center for Analysis of Longitudinal Data in Education. Research Working Paper 57.
- Furlong, J., 2008. Making teaching a 21st century profession: Tony Blair's big prize. *Oxf. Rev. Educ.* 34 (6), 727–739.
- Goldhaber, D., Hansen, M., 2009. National board certification and teachers' career paths: does NBPTS certification influence how long teachers remain in the profession and where they teach? *Educ. Finance Pol.* 4 (3), 229–262.
- Goldhaber, D., Gross, B., Player, D., 2011. Teacher career paths teacher quality and persistence in the classroom: are public schools keeping their best? *J. Pol. Anal. Manag.* 30 (1), 57–87.
- Goodlad, J., 2004. *A Place Called School*. McGraw-Hill, New York.
- Goodlad, J.I., Soder, R., Sirotnik, K.A., 1990. *Places Where Teachers are Taught*. Jossey-Bass, CA.
- Grissmer, D., Kirby, S.N., 1997. Teacher turnover and teacher quality. *Teach. Coll. Rec.* 99 (1), 57–61.
- Grissom, J.A., 2011. Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teach. Coll. Rec.* 113 (11), 2552–2585.
- Gu, Q., 2014. The role of resilience in teachers' career long commitment and effectiveness. *Teach. Teach. Theor. Pract.* 20 (5), 502–529.
- Gu, Q., Day, C., 2007. Teachers resilience: a necessary condition for effectiveness. *Teach. Teach. Educ.* 23, 1302–1316.
- Gu, Q., Day, C., 2013. Challenges to teacher resilience: conditions count. *Br. Educ. Res. J.* 39 (1), 22–44.
- Gu, Q., Johansson, O., 2013. Sustaining school performance: school contexts matter. *Int. J. Leader. Educ.* 16 (3), 301–326.
- Gu, Q., Li, Q., 2013. Sustaining resilience in times of change: stories from Chinese teachers. *Asia Pac. J. Teach. Educ.* 41 (3), 288–303. <https://doi.org/10.1080/1359866X.2013.809056>.
- Gu, Q., Sammons, P., Mehta, P., 2008. Leadership characteristics and practices in schools with different effectiveness and improvement profiles. *Sch. Leader. Manag.* 28 (1), 43–63.
- Gu, Q., Sammons, P., Chen, J., 2018. How principals of successful schools enact education policy: perceptions and accounts from senior and middle leaders. *Leader. Pol. Sch.* 17 (3), 373–390.
- Gu, Q., Hodgen, J., Adkins, M., Armstrong, P., 2019. *Incentivising Schools to Take Up Evidence-Based Practice to Improve Teaching and Learning: Evidence From the Evaluation of the Suffolk Challenge Fund: Final Project Report*. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Campaigns/Suffolk_Challenge_Fund.pdf.
- Gu, Q., Rea, S., Seymour, K., Smethem, L., Bryant, B., Armstrong, P., Ahn, M., Hodgen, J., Knight, R., 2020. *The Research Schools Network: Supporting Schools to Develop Evidence-Informed Practice*. Evaluation Report. Education Endowment Foundation, London. https://educationendowmentfoundation.org.uk/public/files/RS_Evaluation.pdf.
- Gu, Q., Seymour, K., Rea, S., Knight, R., Ahn, M., Sammons, P., Kameshwara, K., Hodgen, J., 2021. *The Research Schools Programme in Opportunity Areas: Investigating the Impact of Research Schools in Promoting Better Outcomes in Schools*. Evaluation Report. Education Endowment Foundation, London.
- Guarino, C.M., Santibanez, L., Daley, G.A., 2006. Teacher recruitment and retention: a review of the recent empirical literature. *Rev. Educ. Res.* 76 (2), 173–208.
- Guttman, C., 2001. *A Hard Sell for Teaching*. The Courier UNESCO. October.
- Hallinger, P., 2005. Instructional leadership and the school principal: a passing fancy that refuses to fade away. *Leader. Pol. Sch.* 4 (3), 1–20.
- Hallinger, P., Heck, R., 2010a. Collaborative leadership and school improvement: understanding the impact on school capacity and student learning. *Sch. Leader. Manag.* 30 (2), 95–110. <https://doi.org/10.1080/13632431003663214>.
- Hallinger, P., Heck, R.H., 2010b. Leadership for learning: does collaborative leadership make a difference? *Educ. Manag. Adm. Leader.* 38 (6), 654–678. <https://doi.org/10.1177/1741143210379060>.
- Hansen, D., 1995. *The Call to Teach*. Teachers College Press, New York, NY.
- Hansen, D., 2021. *Reimagining the Call to Teach*. Teachers College Press, New York, NY.
- Hanushek, E.A., 1992. The trade-off between child quantity and quality. *J. Polit. Econ.* 100 (1), 84–117. February.
- Hanushek, E.A., 2011. The economic value of higher teacher quality. *Econ. Educ. Rev.* 30 (3), 466–479.
- Hargreaves, A., Fullan, M., 2012. *Professional Capital: Transforming Teaching in Every School*. Teachers College Press, New York, NY.

- Henry, G.T., Bastian, K.C., Fortner, C.K., 2011. Stayers and leavers: early career teacher effectiveness and attrition. *Educ. Res.* 40 (6), 271–280.
- Higgins, C., 2011. *The Good Life of Teaching: An Ethics of professional practice*. Wiley-Blackwell, Chichester, West Sussex.
- HM Government, 2022. White Paper: Levelling Up the United Kingdom. Presented to Parliament by the Secretary of State for Levelling Up, Housing and Communities by Command of Her Majesty. <https://www.gov.uk/government/publications/levelling-up-the-united-kingdom>.
- Holme, J., Rangel, V., 2012. Putting school reform in its place: social geography, organizational social capital, and school performance. *Am. Educ. Res. J.* 49 (2), 257–283.
- House of Commons, 2021. *Teacher Recruitment and Retention in England*. House of Commons, London.
- Huberman, M., 1993. *The Lives of Teachers*. Continuum, Falmer.
- Ingersoll, R.M., 2001. Teacher turnover and teacher shortages: an organizational analysis. *Am. Educ. Res. J.* 38 (3), 499–534.
- Ingersoll, R., 2003. Is There Really a Teacher Shortage? Consortium for Policy Research in Education, University of Pennsylvania, Philadelphia, PA. Retrieved from: <http://www.gse.upenn.edu/pdf/rmi/Shortage-RMI-09-2003.pdf>.
- Institute of Physics, 2020. *Subjects Matter*. IOP, London.
- Johnson, S.M., 2004. *Finders and Keepers*. Jossey-Bass, San Francisco, CA.
- Johnson, S. (2019). *Where Teachers Survive: Organising Schools for Success*. Cambridge, Massachusetts: Harvard Education Press.
- Johnson, S.M., Berg, J.H., Donaldson, M.L., 2005. A Review of the Literature on Teacher Retention. Harvard Graduate School of Education.
- Johnson, B., Down, B., Cornu, R., Peters, J., Sullivan, A., Pearce, J., Hunter, J., 2016. *Promoting Early Career Teacher Resilience*. Routledge, New York and London.
- Jordan, J., 2006. Relational resilience in girls. In: Goldstein, S., Brooks, R.B. (Eds.), *Handbook of Resilience in Children*. Springer, New York, NY, pp. 79–90.
- Jordan, J., 2012. Relational resilience in girls. In: Goldstein, S., Brooks, R.B. (Eds.), *Handbook of Resilience in Children*, second ed. Springer, New York, NY, pp. 73–86.
- Kardos, S.M., Johnson, S.M., 2007. On their own and presumed expert: new teachers' experience with their colleagues. *Teach. Coll. Rec.* 109 (9), 2083–2106.
- Kelley, C.J., Finnigan, K., 2003. The effects of organisational context on teacher expectancy. *Educ. Adm. Q.* 39 (5), 603–634.
- Kennedy, M.M., 2010. Attribution error and the quest for teacher quality. *Educ. Res.* 39 (8), 591–598.
- Klassen, R.M., Chiu, M.M., 2010. Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102 (3), 741–756.
- Klassen, R.M., Tze, V.M., Betts, S.M., Gordon, K.A., 2011. Teacher efficacy research 1998–2009: signs of progress or unfulfilled promises? *Educ. Psychol. Rev.* 23, 21–43.
- Ladd, H., 2009. Teachers' perceptions of Their Working Conditions: How predictive of policy-Relevant Outcomes (National Center for Analysis of Longitudinal Data in Education Research Working Paper No. 33). CALDER, Washington, DC.
- Ladwig, J.G., 2010. Beyond academic outcomes. *Rev. Res. Educ.* 34, 113–141.
- Lankford, M., Loeb, S., Wyckoff, J., 2002. Teacher sorting and the plight of urban schools: a descriptive analysis. *Educ. Eval. Pol. Anal.* 13 (3), 256–268.
- Leithwood, K., 2019. *Leadership Development on a Large Scale: Lessons for Future Success*. Corwin, Thousand Oaks, CA.
- Leithwood, K., Louie, K.S., 2012. *Linking Leadership to Student Learning*. Jossey-Bass, San Francisco, CA.
- Leithwood, K., Seashore Louis, K., Anderson, S., Wahlstrom, K., 2004. *How Leadership Influences Student Learning*. Center for Applied Research and Educational Improvement, St. Paul, MN.
- Leithwood, K., Day, C., Sammons, P., Harris, A., Hopkins, D., 2006. *Seven Strong Claims About Successful School Leadership*. National College for School Leadership, Nottingham.
- Leithwood, K., Harris, A., Strauss, T., 2010. *Leading School Turnaround: How Successful Leaders Transform Low-Performing Schools*. Jossey-Bass, San Francisco.
- Leithwood, K., Sun, J., Pollock, K. (Eds.), 2017. *How School Leaders Contribute to Student Success: The Four Paths Framework*. Springer, Dordrecht, Netherlands.
- Loeb, S., Darling-Hammond, L., Luczak, J., 2005. How teaching conditions predict teacher turnover in California schools. *Peabody J. Educ.* 80 (3), 44–70.
- Lu, L., 2010. Chinese wellbeing. In: Bond, M.H. (Ed.), *The Oxford Handbook of Chinese Psychology*. Oxford University Press, Oxford, pp. 327–342.
- Luke, A., 2011. Generalizing across borders: policy and the limits of educational science. *Educ. Res.* 40 (8), 367–377.
- Matheson, I., 2007. Current Demographics in the School Teacher population in Scotland. Paper presented at the Scottish Educational Research Association conference.
- Matthews, P., Rea, S., Hill, R., Gu, Q., 2014. *Freedom to Lead: A Study of Outstanding Primary School Leadership in England*. National College for School Leadership, Nottingham.
- Mourshed, M., Chijioke, C., Barber, M., 2010. *How the World's Most Improved School Systems Keep Getting Better*. McKinsey and Company, New York, NY.
- Nias, J., 1989. *Primary Teachers Talking*. Routledge, London.
- Nieto, S., 2003. *What Keeps Teachers Going?* Teachers College Press, New York, NY.
- Nieto, S., 2011. Critical hope, in spite of it all. In: Elmore, R.F. (Ed.), *I Used to Think ... and Now I Think*. Harvard Education Press, Cambridge, MA, pp. 127–133.
- Nieto, S., 2015. *Why We Teach Now?* Teachers College Press, New York, NY.
- Noddings, N., 1992. *The Challenge to Care in Schools*. Teachers College Press, New York.
- OECD, 2005. *Teachers Matter*. OECD, Paris.
- OECD, 2009. *Creating Effective Teaching and Learning Environments: First Results From TALIS*. OECD, Paris.
- OECD, 2011. *Building a High-Quality Teaching Profession: Lessons From Around the World*. OECD, Paris.
- OECD, 2013. *OECD Guidelines on Measuring Subjective Well-Being*. OECD Publishing, Paris.
- OECD, 2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*. OECD, Paris.
- OECD, 2016. *Education at a Glance: OECD Indicators*. OECD Publishing, Paris.
- OECD, 2018a. *Where All Students Can Succeed*, vol. II. OECD, Paris.
- OECD, 2018b. *TALIS 2018 Results: Teachers and School Leaders as Lifelong Learners*, vol. 1. OECD, Paris.
- Palmer, P., 2007. *The Courage to Teach*, tenth ed. Jossey-Bass, New York, NY.
- Patterson, J., Kelleher, P., 2005. *Resilient School Leaders: Strategies for Turning Adversity Into Achievement*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Perry, E., Halliday, J., Higginson, J., Patel, S., 2022. *Meeting the Challenge of Providing High-Quality Continuing Professional Development for Teachers*. Wellcome, London.
- Pillay, H., Goddard, R., Wilss, L., 2005. Well-being, burnout and competence implications for teachers. *Aust. J. Teach. Educ.* 30, 22–33.
- Pro Bono Economics (PBE), 2022. *Learning to Save: Teacher CPD as a Cost-Effective Approach to Improving Retention*. Pro Bono Economics, London.
- RAND, 2021. *Understanding Teacher Retention: Using a Discrete Choice Experiment to Measure Teacher Retention in England*. Office for Manpower Economics, Cambridge.
- Renshaw, T.L., Long, A.C.J., Cook, C.R., 2015. Assessing adolescents' positive psychological functioning at school: development and validation of the student subjective wellbeing questionnaire. *Sch. Psychol. Q.* 30 (4), 534–552.
- Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools and academic achievement. *Econometrica* 73 (2), 417–458.
- Rockoff, J.E., 2004. The impact of individual teachers on student achievement: evidence from panel data. *Am. Econ. Rev. Papers Proc.* 94 (2), 247–252.
- Rots, I., Aelterman, A., 2008. Two profiles of teacher education graduates: a discriminant analysis of teaching commitment. *Eur. Educ. Res. J.* 7, 523–534.
- Sammons, P., Day, C., Kington, A., Gu, Q., Stobart, G., Smees, R., 2007. Exploring variations in teachers' work, lives and their effects on pupils: key findings and implications from a longitudinal mixed-method study. *Br. Educ. Res. J.* 33 (5), 681–701.
- Sammons, P., Gu, Q., Day, C., Ko, J., 2011. Exploring the impact of school leadership on pupil outcomes. *Int. J. Educ. Manag.* 25 (1), 83–101.
- Schoon, I., 2012. Temporal and contextual dimensions to individual positive development: a developmental-contextual systems model of resilience. In: Ungar, M. (Ed.), *The Social Ecology of Resilience: A Handbook of Theory and Practice*. Springer, Dordrecht, pp. 143–156.
- Schwarz, G., Alberts, J., 1998. *Teacher Lore and Professional Development for School Reform*. Bergin & Garvey, Westport, CT.
- Schwerdtfeger, A., Konermann, L., Schönhofen, K., 2008. Self-efficacy as a health-protective resource in teachers? A biopsychological approach. *Health Psychol.* 27, 358.
- Shen, J., Palmer, L.B., 2009. Inadequate preparation does impact teacher attrition. In: Shen, J. (Ed.), *School Leaders: Professional and Demographic Characteristics*. Peter Lang, New York, pp. 125–140.

- Sherman, S., 2013. *Teacher Preparation as an Inspirational Practice: Building Capacities for Responsiveness*. Routledge, New York, NY.
- Sherman, S., 2020. Evolving enactments of personal fulfilment and service in teaching. In: De Marzio, D.M. (Ed.), *David Hansen and the Call the Teach: Renewing the Work That Teachers Do*. Teachers College Press, New York, NY, pp. 13–26.
- Shulman, L.S., 1987. Knowledge and teaching. *Harv. Educ. Rev.* 57, 1–22.
- Sibieta, L., 2020. *Teacher Shortages in England: Analysis and Pay Options*. Education Policy Institute.
- Skaalvik, E., Skaalvik, S., 2010. Teacher self-efficacy and teacher burnout: a study of relations. *Teach. Teach. Educ.* 26, 1059–1069.
- Smith, J.M., Kovacs, P.E., 2011. The impact of standards-based reform on teachers: the case of “No Child Left Behind”. *Teach. Teach. Theor. Pract.* 7 (2), 201–225.
- Soder, R., 2004. When I get my own classroom. In: Goodlad, J., McMannon, T. (Eds.), *The Teaching Career*. Teachers College Press, New York, NY, pp. 1–18.
- Song, H., Gu, Q., Zhang, Z.H., 2020. An exploratory study of teachers' subjective wellbeing: understanding the links between teachers' income satisfaction, altruism, self-efficacy and work satisfaction. *Teach. Teach. Theor. Pract.* <https://doi.org/10.1080/13540602.2020.1719059>.
- Sun, J., Leithwood, K., 2012. Transformational school leadership effects on student achievement. *Leader. Pol. Sch.* 11 (4), 418–451.
- Sun, J., Leithwood, K., 2015. Leadership effects on student learning mediated by teacher emotions. *Societies* 5, 566–582. <https://doi.org/10.3390/soc5030566>.
- Sutton Trust, 2011. *Improving the Impact of Teachers on Pupil Achievement in the UK - Interim Findings*. London: The Sutton Trust.
- Tschannen-Morgan, M., Barr, M., 2004. Fostering student learning: the relationship of collective teacher efficacy and student achievement. *Leader. Pol. Sch.* 3 (3), 189–209.
- UNESCO, 2006. *Teachers and Educational Quality: Monitoring Global Needs for 2015*. UNESCO, Paris.
- UNESCO, 2011. *The Global Demand for Primary Teachers—2011 Update*. UNESCO, Paris.
- UNESCO, 2014. *Wanted: Trained Teachers to Ensure Every Child's Right to Primary Education*. UNESCO, Paris. Montreal, Canada/Paris, UNESCO Institute for Statistics/EFA Global Monitoring Report. (Policy Paper 15/Fact Sheet 30.).
- UNESCO, 2015. *Education for All 2000–2015: Achievements and Challenges*. UNESCO, Paris.
- UNESCO, 2020. *Global Education Monitoring Report: Inclusion and Education: All Means All*. UNESCO, Paris.
- Ungar, M., 2011. The social ecology of resilience: addressing contextual and cultural ambiguity of a nascent construct. *Am. J. Orthopsychiatr.* 81 (1), 1–17.
- Ungar, M., Ghazinour, M., Richter, J., 2013. Annual research review: what is resilience within the social ecology of human development? *JCPP* 54 (4), 348–366.
- Welzel, C., Inglehart, R., 2010. Agency, values, and well-being: a human development model. *Soc. Indic. Res.* 97 (1), 43–63.
- Worth, J., Faulkner-Ellis, H., 2022. *Teacher Labour Market in England*. National Foundation for Educational Research, Slough.
- Zee, M., Koomen, H.M.Y., 2016. Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher wellbeing: a synthesis of 40 years of research. *Rev. Educ. Res.* 86 (4), 981–1015.

Current trends, tensions and unresolved issues in research on teacher professional learning

Karen Koellner^a, Jennifer Jacobs^b, Hilda Borko^c, and Nanette Seago^d, ^a Arizona State University, Tempe, AZ, United States; ^b Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, United States; ^c Graduate School of Education, Stanford University, Stanford, CA, United States; and ^d WestEd, Riverside, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of H. Borko, J. Jacobs, K. Koellner, Contemporary Approaches to Teacher Professional Development, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 548-556, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00654-0>.

Trend: practice-based professional development	551
Artifacts of practice	551
Co-design of instructional resources	552
Tensions and unresolved issues	552
Trend: online professional development	552
Forms of online professional development	553
Effectiveness of online professional development	553
Tensions and unresolved issues	554
Trend: facilitation of professional development	554
The facilitator's role in supporting learning and community development	554
The preparation of PD facilitators	554
Frameworks for effective PD facilitation	555
Tensions and unresolved issues	555
Trend: professional development with a focus on equity	555
Content-specific professional development through an equity lens	556
Content-agnostic professional development through an equity lens	556
Tensions and unresolved issues	557
Discussion	557
References	557

The literature on formally structured professional learning opportunities for in-service teachers has continued to grow steadily over the past decade. In our article on contemporary approaches to professional development (PD) in the International Encyclopedia of Education ten years ago, we examined the content and processes commonly seen in high quality professional development (Borko et al., 2010). Since that time, a steadily growing body of research has confirmed the importance of the characteristics of effective PD that we highlighted: situated in practice, focused on student learning, models preferred instructional practice, active teacher inquiry, collaborative learning communities, located in appropriate settings, and sustained over time (c.f., Asterhan and Lefstein, 2020; Lefstein et al., 2020; Bautista et al., 2017; Darling et al., 2017; Desimore and Garet, 2015; Hubers et al., 2020).

There continues to be a widespread effort within the international research community - across content areas and topics of interest - to carefully consider how to support teachers' on-going professional development (c.f. Bautista and Ortega-Ruiz, 2015; Kennedy, 2016; Lieberman and Mace, 2010). There is agreement that teachers can benefit from engaging in opportunities to study and reflect on instructional practices, student learning, and the broader educational context. To this end, intentionally organized structures (such as practice-based learning and online learning environments) and domain expertise (such as knowledgeable facilitators and coaches) are critical elements. There is also a growing recognition of the importance of centering teacher learning around issues of equity and supporting diverse students.

In our earlier article we provided examples of PD efforts within each of three research phases, as delineated by Borko (2004), noting that most research on teacher professional learning prior to 2010 could be classified as Phase 1 or proof-of-concept studies. Since that time, other researchers have showcased the large increase in Phase 2 (single program/multiple sites) and 3 studies (multiple programs and multiple sites), particularly in the area of mathematics education (Karsenty and Arcavi, 2017; Santagata et al., 2019; Sztajn et al., 2017). Therefore, in this article, we move beyond describing and cataloging PD efforts and instead shift our focus to several of the most salient and innovative topics in PD literature since 2010. Our intention is to highlight trends and unresolved issues within and across these topics, to showcase how the field has matured, and to spotlight issues that are likely to become increasingly relevant over the next decade.

Thousands of journal articles, monographs, books, and dissertations are published about PD each year, as well as numerous reviews of this literature. This article is not intended to serve as another review, but rather as a vehicle to consider a small number of topics that have emerged, in our view, as taking on critical importance in the field over the past decade. To determine which topics

to focus on, we primarily considered existing literature reviews, syntheses, meta-analyses, and research studies pertaining to PD that were published from 2010 to 2020. Our reading of the literature was not exhaustive, but rather in pursuit of our goal to select a set of topics for inclusion in this article. Ultimately, we arrived at the following four topics: (1) practice-based professional development; (2) online professional learning; (3) facilitation of professional development, and (4) PD focused on supporting diverse learners. Within each topic this article highlights active areas of investigation and selected research findings, and identifies questions yet to be answered.

We elected to include these four topics due to their importance as trending issues in the field, meaning that there exists a substantive body of literature that has been built up over the past decade (or longer), and because we anticipate that they will be critical topics in the field going forward. Just as context influences decisions about professional learning experiences, it informed our decisions when writing this article. Thus, we would like to acknowledge the fact that this article was written at the height of the COVID-19 pandemic by an authorship team based in the United States. The pandemic has had a profound and enduring effect on teaching and professional learning (Ferdig et al., 2020; Hartshorne et al., 2020; Minkos and Gelbar, 2021), in particular by centering the importance of online tools for both students and educators. Also critical is that the article was written as the Black Lives Matter movement (Garza, 2014) gained prominence in the United States and other countries. This grass roots movement served to highlight the systemic injustices perpetrated against people of color, drawing the attention of both teachers and researchers to critical issues around race, equity, and social justice (Mayorga and Picower, 2018). These historical events contributed to our selection of topics to address in this article, and motivated our thinking about the topics that are likely to be particularly relevant in the decade ahead.

Trend: practice-based professional development

Over the past decade practice-based PD has become increasingly popular as a way to organize teachers' professional learning, highlighting classroom instruction and students' learning. This trend comes as the primary focus of teacher education has shifted from specifying the knowledge necessary for teaching to supporting teachers in becoming more reflective about the everyday practice of teaching (McDonald et al., 2013; Sztajn et al., 2020). As the National Academies of Sciences, Engineering and Medicine (2020) documented in their recent report on teachers and teacher education over the last 20 years, "advances in PD practice and research have been prominently marked by the potential for teachers' learning in and from practice" (p. 6).

Practice-based PD situates professional learning opportunities in the central activities of teaching (Ball and Cohen, 1999) and supports teachers to "develop or enhance skills by connecting them to current classroom practices" (Hirsch et al., 2020, p. 2). Although many PD models employ practice-based approaches, they are most commonly incorporated into ongoing or iterative models such as plan-do-study-act continuous improvement cycles (Bryk et al., 2015; Lewis, 2015), lesson study (Lewis and Perry, 2017) and the Problem-Solving Cycle/Dialogic Video Cycle (Borko et al., 2015; Gröschner et al., 2015; Borko et al., 2017).

The variety of approaches to practice-based PD highlights the creativity of both teachers and PD designers as they explore new avenues for teachers to attend to specific aspects of their practice. In this section we focus on two of these approaches: (1) examining artifacts of practice during PD, and (2) collaboratively designing instructional materials during PD.

Artifacts of practice

Classroom artifacts of practice are physical records that capture aspects of the work that happens during lessons such as video of classroom discussions, samples of students' written work, and transcripts of students' small-group problem solving (Goldsmith and Seago, 2013). During professional development, classroom artifacts provide a common point of reference to anchor teachers' collaboration and community development and to focus their discussions on content, learning, and teaching (Roth et al., 2011). Recent research has sought to understand how bringing artifacts of practice into PD can offer authentic and rich learning experiences for teachers, as well as how teachers can be supported to meaningfully engage with those artifacts (Jilk, 2016). For example, there is an emerging body of literature on what teachers notice when they attend to artifacts of practice during professional learning workshops (Dreher and Kuntz, 2015; Jilk, 2016; Santagata and Yeh, 2014; Sherin and van Es, 2009). The concept of teacher noticing includes both what teachers observe and how they make sense of teaching and learning as they interact with artifacts of practice (Sherin et al., 2011). Van Es (2011) developed a framework for learning to notice student thinking, which she found to be a useful tool for scaffolding teacher learning.

Video is perhaps the most commonly used and studied artifact of practice in the current literature on practice-based PD. Several recent, comprehensive literature reviews highlight the substantial promise of video-based PD (Gaudin and Chaliès, 2015; Major and Watson, 2018; Marsh and Mitchell, 2014; Tripp and Rich, 2012). These reviews conclude that video is a powerful and effective tool to promote improvements in teaching practice. Furthermore, video-based PD has the potential to enhance teacher cognition and, in some cases, impact student learning outcomes. A number of reports offer guidance on designing PD that integrates video as part of the learning environment (Blomberg et al., 2013; Roth et al., 2017; van Es et al., 2019).

A wide range of professional development models, including both adaptive and specified programs, make use of video to encourage teacher inquiry into their practice, either through individual or collective reflection (Borko et al., 2011b; Koellner and Jacobs, 2015). Adaptive models derive learning goals and resources from the local context and shared video is generally from the classroom of one of the participating teachers. Well-researched adaptive programs that incorporate video include video clubs

(Luna and Sherin, 2017; Sherin et al., 2009), instructional coaching (Desimone and Pack, 2017), and the Problem-Solving Cycle/Dialogic Video Cycle (Borko et al., 2015; Gröschner et al., 2015). Specified programs that incorporate video often provide published materials that specify teacher learning goals and include pre-selected resources, tasks and video clips. Well-researched specified programs that incorporate video include Cognitively Guided Instruction (Carpenter et al., 1989), Learning and Teaching Linear Functions (Seago et al., 2004), and Learning and Teaching Geometry (Seago et al., 2017). Science Teachers Learning through Lesson Analysis (STeLLA) is a well-researched program using video that incorporates both specified and adaptive components (Taylor et al., 2017; Roth et al., 2019).

Video can be used to encourage teacher inquiry into their own instructional practice or the practice of their colleagues or unknown teachers. In general, watching video from their own or other teachers' classrooms has proven beneficial for teacher learning. There is evidence for the positive benefits of both approaches and at present there is not a consensus as to which is best (Beisiegel et al., 2018; Gaudin and Chaliès, 2015; Kleinknecht and Schneider, 2013; Seidel et al., 2011). For example, Seidel et al. (2011) found that when teachers analyze their own videos versus other teachers' videos they had higher engagement and noticed more relevant aspects of teaching and learning, however they were less self-reflective. In contrast, Kleinknecht and Schneider (2013) reported that teachers' emotional and motivational engagement was higher when they viewed video of unknown teachers versus their own teaching. Based on the studies they reviewed, Gaudin and Chaliès (2015) cautioned that video of other teachers can seem removed from teachers' own classrooms and recommended that the educational context in which the video was made be described when video of other teachers is used.

Co-design of instructional resources

Practice-based PD often involves the study and reflection on classroom instruction that has already taken place. However, it can also include learning activities that prepare teachers for future classroom routines and practices. One example of a future-oriented, practice-based activity that can be brought into PD settings is the collaborative development of instructional resources (Voogt et al., 2015). Collaborative design or "co-design" refers to the active participation of teachers in the generation of educational materials or innovations. Similar to the arguments for using artifacts of practice as a tool for teacher learning, co-designing (or re-designing or co-adapting) curriculum and other instructional resources situates teachers' involvement in the learning process as directly meaningful to their everyday practice (Voogt et al., 2011).

Co-design can occur within a variety of formats that intentionally foster a design community in which teachers share knowledge and perspectives, with the aim of generating a resource to support student learning. For example, co-design can be part of iterative inquiry cycles such as lesson study (Wake et al., 2016), and research-practice partnerships that are organized around solving educational problems of practice (Penuel, 2019). Co-design projects fall on a spectrum that ranges from a primary focus on teacher professional learning (such as lesson study, in which teachers collaboratively develop exemplar lessons; Huang et al., 2016; Huang and Shimizu, 2016) to a primary focus on product development (such as producing innovative curricular materials; Gueudet et al., 2016). However, even when the focus is solidly on product development, the participating teachers are engaging in professional learning (Westbroek et al., 2019). By contributing their expertise to the design of resources that will be used in practice, co-design has been demonstrated to have a number of positive impacts on teachers - including an impact on their knowledge (Voogt et al., 2016), perceptions of agency (Severance et al., 2016), and their instruction (Voogt et al., 2016).

Tensions and unresolved issues

There are a range of challenges in designing, implementing, and researching practice-based PD. For example, determining which goals make sense within a particular professional learning context (van Es et al., 2019), identifying the most appropriate artifacts to use (Sherin et al., 2009), and designing effective PD around the artifacts (Beisiegel et al., 2018) are complex, interrelated decisions. These types of decisions typically have multiple options, with both affordances and constraints, that may influence teacher learning opportunities and well as research questions and analyses. Once a design is decided upon, facilitating the PD has its own set of challenges such as how to guide productive discussions that scaffold teacher noticing and support their reflections around the artifacts of practice (Jilk, 2016), and that foster both personal and social growth (Westbroek et al., 2019). More research is needed that examines and compares the affordances, challenges, efficacy and impact of different types of practice-based PD.

Trend: online professional development

Among the biggest concerns of teachers for improving their practice is the need for more and flexible time to access and process new information (Merritt, 2016). The COVID-19 pandemic accelerated the shift from face-to-face to online PD programs and increased and normalized the availability of video conferencing platforms (Ferdig et al., 2020; Lockee, 2020). Online PD has the potential to provide anytime, anywhere access that alleviates teachers' time constraints. As more professional development moves online to take advantage of its flexibility, increased access, and scalability, a tendency may be to view online and face-to-face PD as distinct from each other (Peltola et al., 2017). However, Fishman (2016) argued that they are not different and in fact "online teacher professional development is *professional development*" (p. 14). The quality or success of a PD program is not a function of whether it is online or in person or a combination of the two, but rather a function of its instructional design (Dede et al., 2016).

Forms of online professional development

Numerous online professional learning options are available for teachers including online courses, workshops, and professional communities; curated resources such as video libraries; and virtual instructional coaching (see [Schachter et al., 2019](#) for more specific examples). This plethora of options ensures there are multiple avenues for teachers to take part in self-guided, personalized learning or to participate in an online network or professional community ([Macià and García, 2016](#)). Fully online and hybrid PD courses have become increasingly prevalent over the past decade ([National Academies of Sciences, Engineering and Medicine, 2020](#)). Fully online PD experiences can be structured as synchronous, with all participants interacting in real time; or asynchronous, where participants work at their own pace; or some mix of synchronous and asynchronous. Hybrid or blended learning models offer a combination of face-to-face and online experiences.

In the last several years, a number of successful face-to-face PD models have been adapted or redesigned to become online learning experiences for teachers. Examples of materials adapted for synchronous courses include iRaise, a science literacy program ([Jaciw et al., 2016](#)), and GENA's lesson study, a mathematics lesson study program ([Schoenbach et al., 2016](#); [Francis and Jacobsen, 2013](#)). Adaptations for the development of asynchronous courses are beginning to take place as well, and corresponding research on their effectiveness and impact is underway ([Seago et al., 2020](#)).

The nature of synchronous PD, in which teachers come together in real time in a shared virtual space, often for an extended period of time, offers more opportunities for peer interactions than asynchronous PD. Thus, synchronous courses generally encourage discussion, collaboration and community building, as well as individual accountability for their participation in collective activities ([Schoenbach et al., 2016](#)). Technological tools such as audio and video conferencing, text chats, and collaborative document editing provide a variety of avenues for communication between participants as they teach and learn from each other synchronously ([Fishman, 2016](#)). In addition to increasing the demand and urgency for online professional learning opportunities, the COVID-19 pandemic spurred improvements to video conferencing platforms, enabling increased collaboration and flexibility through features such as screen sharing and breakout rooms, as well as the ability to link easily to tools like whiteboards, Padlets, and Google Jamboards.

Asynchronicity allows participants access to PD at any time, in any location with available internet, and can potentially eliminate the often-mentioned lack of scheduling flexibility and geographic distance ([Meritt, 2016](#); [Pletola et al., 2017](#)). Asynchronous forms of online learning have resulted in positive findings related to teachers attitude and self-efficacy ([An, 2018](#)) as well as high satisfaction and relatively high levels of information sharing ([Yoon et al., 2020](#)). Although asynchronous online PD is highly valued by teachers ([Parsons et al., 2019](#)), the focus on self-directed learning can prove challenging for collaboration and interactivity ([Altman and Harsh, 2017](#)). An emerging body of literature speaks to the potential of asynchronous PD, despite concerns regarding community development and feelings of isolation ([Yuan and Kim, 2014](#)).

Interest in providing scalable, accessible and cost-effective PD has resulted in the rapid evolution of massive open online courses (MOOCs) and MOOCs for Educators (MOOC-Ed) ([Kleiman and Wolf, 2016](#)). MOOCs - which may be synchronous, asynchronous, or some combination - have clear potential for engaging and educating large numbers of teachers around the world ([Misra, 2018](#)). Although there are important concerns such as low completion rates and uneven quality, researchers are beginning to outline design considerations to help ensure the courses follow best practices and offer meaningful experiences (e.g., [Hodges et al., 2016](#)).

Professional organizations such as the National Science Teachers Association (NSTA) imagine the future of teacher learning as a hybrid model that successfully integrates online learning with site-based efforts to provide convenient, immediate, and self-directed learning opportunities ([Byers and Menez, 2016](#)). NSTA has formally collaborated with school districts and universities that use its Learning Center in order to extend and enhance local face-to-face efforts in a hybrid form. For example, after taking part in in-person workshops and incorporating what they learned into their classroom practice, teachers can participate in moderated online discussions around shared student work ([Byers and Menez, 2016](#)). Blending face-to-face and online PD in this manner has led to beneficial learning outcomes documented in numerous studies ([Bowen et al., 2012](#); [Murphy et al., 2014](#); [Fishman, 2016](#)).

Effectiveness of online professional development

Although the research literature on online PD is limited, there is increasing evidence as to its positive effects for the participating teachers, even when compared to face-to-face options ([Fishman et al., 2013](#); [Ilaria, 2017](#); [O'Dwyer et al., 2010](#)). For example, Fisher and colleagues ([Fisher et al., 2010](#)) report that online courses designed based on high-quality learning materials can produce learning outcomes that are similar to or better than face-to-face options. Multiple large scale randomized control studies have demonstrated that cohort-based, facilitated online PD sessions can have significant positive impacts on teachers' knowledge and practice as well as their students' achievement ([Dash et al., 2012](#); [Masters et al., 2012](#); [O'Dwyer et al., 2010](#)). On the other hand, a meta-analysis of STEM teacher PD and curriculum improvement programs by [Lynch et al. \(2019\)](#) found that programs in which teachers completed a portion of the PD online led to poorer student outcomes compared to programs without an online component.

While more research is needed to ascertain the effectiveness and impact of various forms of online PD, it is clear they offer a valuable option for large numbers of teachers. Online PD supports equitable access to resources that may not be available to teachers locally, such as teachers residing in rural communities ([Erickson et al., 2012](#)). In addition, online learning can be a good fit for teachers who are reluctant to share ideas in face-to-face settings but may be more comfortable doing so in digitally mediated interaction ([Dede et al., 2009](#)). Moreover, online PD is considerably more scalable than comparable face-to-face PD, and in many cases is

subject to fewer monetary and logistical constraints for teachers (Killion, 2013). Finally, especially relevant for teachers around the world who recently shifted to virtual instruction due to COVID-19, online PD can serve as important preparation to effectively lead an online classroom environment (An, 2020; Philipsen et al., 2019) and provide a social community for teachers working in isolation (Charteris et al., 2021; Lieberman and Mace, 2010).

Tensions and unresolved issues

A challenge to online PD design, enactment, and research is ensuring that teachers have the knowledge and skills to capitalize on the dynamic nature of technological advancement (Schneider et al., 2016). It is still unclear whether online PD in general is as effective as face-to-face options. However, the nature and design of online PD programs are just as varied and complex as face-to-face programs, and the outcome variables of interest are as wide-ranging. Certainly not all distance learning is effective, just as not all in-person workshops are effective. There is not yet a large enough body of research on how to design and implement effective online PD, but the demands and challenges currently faced by educators necessitate increased availability of professional learning opportunities, including remote options. More research is needed to understand the affordances and limitations of different online and hybrid PD models and delivery platforms to meet a variety of learning goals, including the role they might play across the trajectory of teachers' careers.

Trend: facilitation of professional development

As the demand for PD has grown, efforts to bring PD to scale have increased in the US (Marrongelle et al., 2013) and around the world (Krainer, 2015). Related to these efforts is a recognition of the critical role played by the individuals who serve as PD facilitators and the need for well-trained and knowledgeable PD facilitators leaders, and coaches (Bates et al., 2011). Skillful PD facilitators are critical to ensuring that the PD is carried out in a meaningful, engaging, and productive manner that ensures all teachers have the opportunity to learn (Borko et al., 2011a). Although the research base on PD facilitation has grown in recent years, it is not nearly as robust as the research on practice-based or online professional learning. In this section we consider three themes related to the facilitation of PD that are beginning to be more widely addressed: the facilitator's role, the preparation of facilitators, and frameworks for effective facilitation.

The facilitator's role in supporting learning and community development

Facilitators of professional learning play a variety of roles depending on the educational context, the goals of the PD, and the particular needs and interests of the teachers with whom they are working (Jensen et al., 2016). Facilitation roles are multilayered and include "listener, expert, critical friend, coach and mentor as well as teacher and workshop leader" (Perry and Bolan, 2018, p. 256). Typically, PD leaders support communities of teachers to examine their instructional practice and student thinking, which requires positioning the participants as both experienced professionals as well as adult learners (Ince, 2017). Facilitators play a key role in organizing and managing the conditions within which reflective professional inquiry takes place. For example, facilitators help to establish norms for collaborative discourse and guiding productive communication during workshops (Lefstein et al., 2020). By building a community that fosters confidence and capacity among the participating teachers, facilitators can serve as a social agent to empower teachers to transform not only their practice but also problematic aspects of their work environment (Goncalves et al., 2020; Luguetti et al., 2020).

The preparation of PD facilitators

Although there is general agreement facilitators play an essential role in providing all teachers with the opportunity for ongoing professional learning (Hadar and Brody, 2017), research on how to educate facilitators is sparse (Even, 2014; Sztajn et al., 2017). Yet, training and support for novice facilitators is critical given how new the field is and the fact that it is increasingly common for teachers to lead PD for their peers, often as part of school-based professional learning communities. There is a growing recognition that these teacher leaders need considerable guidance as they shift from implementing lessons in classrooms with students to leading workshops with fellow teachers (Huggins et al., 2017). Allen (2016) argued that this shift not only involves new resources and skills, but also the formation of a new professional identity commensurate with their facilitation responsibilities. More generally, educators of practicing teachers require a specialized knowledge base that is beginning to be better understood and articulated by the research community (Selmer et al., 2016). Perry (2020) has recently classified the knowledge needed by PD facilitators as knowledge and skills of teaching, facilitation knowledge and skills, and knowledge of professional development.

Several research programs offer insights into effective strategies for facilitator preparation, such as modeling, rehearsals and critically analyzing video cases. Modeling of PD activities involves the novice facilitators participating as teachers while more experienced facilitation leaders model or showcase exemplary practices (Even et al., 2003; Jackson et al., 2015). In some cases, modeling is followed by the novice facilitators collectively planning for and rehearsing their facilitation strategies in preparation for leading PD sessions (Borko et al., 2014, 2017). Working with a more experienced facilitator, Jacobs et al. (2017) relied only on the rehearsal of PD activities. Several research teams have designed video-based learning opportunities for facilitators, in which they viewed videos

of themselves or others leading professional development workshops and critically analyzed the facilitation moves and practices (Lesseig et al., 2017; Perry and Boylan, 2018; Schueler and Roesken-Winter, 2018). Similar to many of the PD programs highlighted above, recently developed programs for facilitator preparation tend to be practice-based, bringing activities and resources from teacher PD into the facilitator PD experiences (Prediger et al., 2019).

Following from the literature on teacher “noticing” during professional development, teacher leader “noticing” has begun to emerge as an important construct related to the preparation of PD facilitators. Lesseig et al. (2017) argued that PD leaders can benefit from structured opportunities to notice or attend to effective facilitation skills and resources in order to become more knowledgeable, purposeful, and effective. They offered several design principles for workshops that support PD leaders, including the importance of establishing clear learning goals, targeting the development of specialized content knowledge, and using artifacts of practice. Other research highlights not only the importance of facilitators noticing teachers’ ideas and thinking, but also gaining knowledge about how to productively act on those ideas (Schueler and Roesken-Winter, 2018; Tekkumru-Kisa and Stein, 2017).

Frameworks for effective PD facilitation

Several recently developed frameworks for effective facilitation shed light on the specific practices that are likely to foster teacher participation in and learning from PD (e.g., Arya et al., 2014; Gonzales et al., 2016; Tekkumru-Kisa and Stein, 2017; Van Es et al., 2014; Zhang et al., 2011). Most of these frameworks have a particular focus on leading practice-based PD, especially prompting verbal interactions during video-based workshops. Van Es et al.’s (2014) framework includes four facilitation practices aimed at generating productive conversations around student thinking: (1) orienting the group, (2) sustaining an inquiry stance, (3) maintaining a focus on the video and the mathematics, and (4) supporting group collaboration. Gröschner et al. (2014) additionally emphasize the importance of facilitators’ “mindfulness” as they use specific moves to foster a trustful, community-building atmosphere within a collaborative learning community. Arya et al. (2014) found that modeling, scaffolding, and co-construction of ideas were facilitation practices that benefited teachers when engaging in video-based discussions.

Tensions and unresolved issues

Reflecting global efforts to scale teacher professional learning is the need for increased attention to, and additional research on, the facilitation of high-quality PD (Kennedy, 2016; National Academies of Science, Engineering and Medicine, 2015; Prediger et al., 2019). In particular, there is an urgent need for additional research building on what is currently known effective facilitation strategies and how to prepare facilitators to skillfully utilize those strategies in a variety of contexts with diverse teachers. As many PD efforts pivot to address challenges such as COVID-19 and the pursuit of racial equality within educational systems, facilitators are engaging with new formats (i.e., online learning) and expanding professional learning goals. Flexibly and appropriately responding to teachers’ complex and rapidly changing circumstances mandates new types of facilitation roles, knowledge, and frameworks. It is incumbent on researchers in the field to address these evolving needs to ensure that facilitators have the tools and skills to meaningfully carry out their work even in uncertain times.

Trend: professional development with a focus on equity

The field of education has an obligation to attend to issues of equity, access, and justice to reduce the racial disparities inherent throughout our societies and especially in our schools. Recent cultural shifts and grassroot movements, such as Black Lives Matter, have contributed to an increased demand for PD models that ensure teachers have the necessary knowledge, skills, and pedagogical resources to meet this critical need across the globe. There is a small but accumulating body of literature about PD efforts to support teachers in their efforts to work equitably with racially, culturally, and linguistically diverse groups of students (Parkhouse et al., 2019). We label this research as having a focus on equity, while recognizing that others have used a variety of labels. For example, teacher professional learning programs that highlight inclusion and equity have been labeled as multicultural education (Banks and Banks, 2019), intercultural education (Cloonan et al., 2017), and peace education (Verma, 2017), among others. These programs generally feature strategies teachers can use to bring students’ home cultures and languages into the classroom, countering implicit biases and oppressive ideologies concerning people of color, and frequently highlighting instructional practices labeled as culturally relevant pedagogy (Ladson-Billings, 1992, 1995, 2014), culturally responsive teaching (Gay, 2002), or culturally sustaining pedagogy (Ladson-Billings, 2014; Paris, 2017; Paris and Alim, 2017).

Research suggests that most teachers will require substantial professional learning supports to respond productively to rapidly changing student demographics as well as increased expectations that they effectively address the needs of all students (Battey and Franke, 2015; Cobb and Couch, 2018; Osborne et al., 2019). Considering the manner in which teachers learn to become more equitable in their classroom instruction, Carter Andrews et al. (2016) argued that “practicing teachers cannot develop and maintain a humanizing pedagogy through experiences in one ‘diversity’ or ‘multicultural’ professional development seminar” (p. 171). Correspondingly, there has been a slow but steady increase in both PD opportunities and the related literature on how best to guide teacher learning about social justice, inclusion and equity in a manner that leads to classroom change (e.g., Foote, 2010; Parkhouse et al., 2019; Strutchens et al., 2012; Wager and Foote, 2013).

PD programs that highlight the importance of moving toward greater educational equity can motivate teachers to become more attentive to the increasingly diverse student populations that are in their classrooms (National Academy of Sciences, 2020). Further, such PD may serve as inspiration for teachers to consider ways to actively disrupt the systemic inequality that currently exists in K-12 schooling (Ladson-Billings, 2014; Martin, 2012). Providing teachers with opportunities for professional collaboration, dialog, and reflection around issues of (in)equity both empowers and legitimizes their role as agents of change (Hawkins, 2011), and supports movement away from assimilationist teaching toward the “curricularization of equity” (Paris, 2017).

We consider the current literature on PD with an equity focus through two lenses: content-specific PD with an equity focus and content-agnostic PD with an equity focus. Within both categories, some PD efforts are primarily concerned with increasing teachers’ knowledge and attentiveness to equity-related issues, while others move beyond awareness and delve into strategies for planning, teaching and assessing in ways that actively promote equity.

Content-specific professional development through an equity lens

Although not yet commonplace, there are an increasing number of content-specific PD efforts that have an equity focus, with the goal of integrating subject matter content and attention to equity within a given professional learning endeavor. Here we discuss the research on some of these PD programs for mathematics, computer science and science teachers, as these areas have been the focus of most published research, although certainly there are important efforts in other content areas.

In a recent review of the mathematics education literature, Yolcu (2019) argues: “Improvement of teaching practices has become necessary in the context of social perspectives and reform-oriented curriculum not only to address but also to facilitate the creation of socially just and equitable opportunities to learn mathematics. For example, one of the trends in the professional development programs focuses on the improvement of cooperative learning opportunities and meaningful group work experiences for all students” (p. 708). Others in mathematics education have developed PD frameworks that incorporate key aspects of culturally relevant pedagogy and studied teachers’ uptake of these practices in their implementation of mathematics lessons (Rubel and Chu, 2012; Timmons-Brown and Warner, 2016). Battey and Franke (2015) describe another approach in which teachers constructed narratives around race that challenged their thinking about the abilities and participation of students of color in their mathematics classrooms.

The field of computer science is marked by a pronounced racial participation gap, prompting a variety of efforts to broaden participation, accessibility, and inclusion (K-12 Computer Science Framework, 2016). Similar to mathematics education, researchers in computer science have developed a number of PD programs that encourage computer science teachers to integrate culturally responsive pedagogical practices (Coddling et al., 2019; Goode2020; Scott et al., 2010). For example, Goode and colleagues (Goode et al., 2020; Goode et al., 2020) developed a PD program for computer science teachers that focuses on disrupting “colorblind discourse” and instead encourages critical conversations about race. As Goode et al. (2020a,b) explain: “In an educational discipline marked by power, privilege, and pervasive and systemic patterns of underrepresentation, this study demonstrates how long-term professional development can support teachers in developing an equity-based orientation toward computer science education” (p. 30). They report that this effort was successful in shifting teachers’ equity-based attitudes and beliefs, as well as moving teachers toward culturally responsive computer science pedagogy.

In science education, there have been a number of PD efforts drawing on a variety of design principles to encourage attention to equity (Bancroft and Nyirenda, 2020). For example, similar to mathematics and computer science, science educators have developed PD programs that support teachers to incorporate culturally responsive teaching practices (Brown and Crippen, 2016; Johnson, 2011). In a program entitled “Teaching for Equity in Mathematics and Science Education,” Bianchini et al. (2015) foregrounded issues of equity and diversity, using a variety of strategies that prompted teachers to engage in conversations about equity as well as having them conduct their own equity research projects. A different area of recent research in mathematics and science PD involves utilizing learning analytics technology to provide teachers with student data disaggregated in particular ways, such as by race or gender, and encouraging the noticing of patterns related to inequity or bias in their own classrooms (Raza et al., 2020; Reinholz et al., 2020). The variation across these programs suggests that there is an array of PD designs that hold promise in supporting teachers to become better equipped to intentionally disrupt inequality in their classrooms.

Content-agnostic professional development through an equity lens

A small number of scholars have focused on designing PD that centers equity, independent of a subject matter content focus. For example, transformative PD is a model that encourages teachers’ transformative beliefs and instructional practices, where “transformative” is defined as attending to equity and social justice in multicultural classrooms (Kose and Lim, 2010). “Teaching Respect for All” (UNESCO, 2014) is a transformative PD program that has been successfully used in several countries including South Africa to empower teachers as agents of social cohesion (Mogliacci et al., 2016). Peace education is another example of an educational and PD models that has gained considerable attention internationally (Bajaj and Hantzopoulos, 2016; Burns and Aspeslagh, 2014). Peace education emphasizes conflict resolution, particularly in response to violent conflicts or wars (Salomon and Cairns, 2011). PD that supports teachers to implement curriculum and practices aligned with peace education highlights the importance of becoming committed to and adept in classroom instruction that aligns with the goals of peacemaking, democratic dialog, and restorative justice (Parker and Bickmore, 2020; Verma, 2017).

Cochran-Smith (2015) describes teacher learning communities that formed with the explicit goal of promoting equity within their schools and classrooms, addressing concerns that emerged from the participating teachers' practice and considering how to best foster learning opportunities for all students, particularly those who have been historically marginalized by schools. Carter Andrews and Richmond (2019) add that it is not enough to focus on equity goals for students; to be effective, equity-focused PD must include equity goals for both student learners and teacher learners. They suggest that PD activities be designed to help teachers develop an asset-based pedagogical orientation and shared goals related to sustaining students' cultural assets and ensuring that all students have access to learning opportunities that will maximize their academic success and life opportunities. Larrivee (2000) further suggests the necessity for teachers to be willing to critically reflect on their own bias and ways in which their personal and professional identities inform their (in)ability to meet the needs of diverse student populations.

Zeichner (2011) makes a related point, arguing that teachers need to understand the communities and neighborhoods in which their children live. He proposes that teachers spend time interacting with community activists and people from neighborhoods who experience inequalities and injustices. This immersive approach allows teachers to see first-hand the conditions and environment in which their children live every day. A number of programs for both preservice and inservice teachers have provided immersive experiences in an effort to help educators humanize their students of color and cultivate praxis or critical reflection on the world in order to change it (Aujla-Bhullar, 2011; Stacy et al., 2020). In a similar vein, international field trips have been studied as a means to support teachers' intercultural awareness and competencies, particularly trips to developing countries (Alfaro and Quezada, 2010; Baecher and Chung, 2020).

Tensions and unresolved issues

PD focused on equity prompts teachers to articulate the underlying theories and reasoning behind their craft with respect to reaching all students, to confront the existing obstacles and tensions, and to deeply consider pedagogical improvements that promote social justice. Although this is a trending topic in the recent education literature, teacher participation in such PD is still relatively rare. Banilower et al. (2018) reported that only a quarter of science teachers and 20% of mathematics teachers indicated that they had participated in PD with a focus on incorporating students' cultural backgrounds into instruction. Clearly there is a need to scale up PD related to equity, along with clarity as to what such PD should look like. In particular, more research is needed to understand the design criteria for effective PD in this area, and helps teachers shift their beliefs and practices related to ensuring an equity focus in their classrooms and in their students' learning outcomes.

Discussion

The start of the new century has brought exciting new technological innovation, diffusely organized and disruptive social movements, and most recently, the greatest pandemic that modern society has faced - all of which have heavily impacted schools, teachers and students. Educators have begun generating wide-ranging strategies to address these novel challenges, including moving the field forward with expanded professional learning opportunities. In this article we have reviewed several of the salient and trending topics related to PD that have been taken up across the globe in the last decade. We anticipate that in the next decade the field will make considerable progress in each of these areas, leading to a better understanding of how to support teachers to learn and thrive in diverse contexts, even during a time of rapid globalization, change, and uncertainty.

This article spotlights four PD trends: practice-based professional development, online professional learning, facilitation of professional development, and PD focused on equity. We highlight current research in each of these areas, while also pointing out some of the existing tensions and unresolved issues that may be fodder for future lines of research. Practice-based learning by design requires thoughtful attention to the specific, local contexts in which teachers work, including using artifacts of practice to reflect on the needs of their students and collaboratively designing high quality instructional materials to meet those needs. The demand for online PD has greatly increased, prompting researchers to move quickly to delineate the affordances and limitations of this modality and speak to its potential in supporting a variety of learning outcomes of interest. Most PD programs, regardless of focus or modality, are led by facilitators who support teachers' community development and learning, thereby driving the need for research pertaining to the development of facilitation skills and expertise. Lastly, there is a heightened interest in PD that addresses issues of equity, inclusion and social justice. This type of PD helps teachers both to become better able to meet the needs of the increasingly diverse student populations in their classrooms and to play an advocacy role for underrepresented youth more generally. Although the article attends to each of these four trends separately, they are in many ways intertwined. Each trend has a unique scope of work and draws on a particular extant literature base, however future research may benefit from efforts to combine them in creative ways – such as preparing facilitators to lead online, practice-based professional development that focuses on equity.

References

- Alfaro, C., Quezada, R.L., 2010. International teacher professional development: teacher reflections of authentic teaching and learning experiences. *Teach. Educ.* 21 (1), 47–59.
- Allen, D., 2016. The resourceful facilitator: teacher leaders constructing identities as facilitators of teacher peer groups. *Teach. Teach.* 22 (1), 70–83.
- Alterman, R., Harsch, K., 2017. A more reflective form of joint problem solving. *Int. J. Comp. Support. Collabor. Learn.* 12 (1), 9–33.

- An, Y., 2018. The effects of an online professional development course on teachers' perceptions, attitudes, self-efficacy, and behavioral intentions regarding digital game-based learning. *Educ. Technol. Res. Dev.* 66, 1505–1527.
- An, Y., 2020. A response to an article entitled "Improving teacher professional development for online and blended learning: a systematic meta-aggregative review". *Educ. Technol. Res. Dev.* 1–4.
- Arya, P., Christ, T., Chiu, M.M., 2014. Facilitation and teacher behaviors: an analysis of literacy teachers' video-case discussions. *J. Teach. Educ.* 65 (2), 111–127.
- Asterhan, C., Lefstein, A., 2020. Evidence-Based Design Principles For Effective Professional Development: A Critical Appraisal of the Evidence.
- Aujla-Bhullar, S., 2011. Deconstructing diversity: professional development for elementary teachers. *Diaspora, Indig. Minority Educ.* 5 (4), 266–276.
- Baeher, L., Chung, S., 2020. Transformative professional development for in-service teachers through international service learning. *Teach. Dev.* 24 (1), 33–51.
- Bajaj, M., Hantzopoulos, M. (Eds.), 2016. *Peace Education: International Perspectives*. Bloomsbury Publishing.
- Ball, D.L., Cohen, D.K., 1999. Developing practice, developing practitioners: toward a practice-based theory of professional education. In: Darling-Hammond, L., Sykes, G. (Eds.), *Teaching as the Learning Profession: Handbook of Policy and Practice*. Jossey-Bass, San Francisco, pp. 3–32.
- Bancroft, S.F., Nyirenda, E.M., 2020. Equity-focused K-12 science teacher professional development: a review of the literature 2001–2017. *J. Sci. Teach. Educ.* 31 (2), 151–207.
- Banilower, E.R., Smith, P.S., Malzahn, K.A., Plumley, C.L., Gordon, E.M., Hayes, M.L., 2018. Report of the 2018 NSSME+. Chapel Hill, NC: Horizon Research, Inc.
- Banks, J.A., Banks, C.A.M. (Eds.), 2019. *Multicultural Education: Issues and Perspectives*. John Wiley & Sons.
- Bates, T., Swennen, A., Jones, K., 2011. Teacher educators—a professional development perspective. In: Bates, T., Swennen, A., Jones, K. (Eds.), *The Professional Development of Teacher Educators*. Routledge, London, pp. 7–19.
- Bathey, D., Franke, M., 2015. Integrating professional development on mathematics and equity: countering deficit views of students of color. *Educ. Urban Soc.* 47 (4), 433–462.
- Bautista, A., Ortega-Ruiz, R., 2015. Teacher professional development: international perspectives and approaches. *Psychol. Soc. Educ.* 7 (3), 240–251.
- Bautista, A., Yau, X., Won, J., 2017. High-quality music teacher professional development: a review of the literature. *Music Educ. Res.* 19 (4), 455–469.
- Beisiegel, M., Mitchell, R., Hill, H.C., 2018. The design of video-based professional development: an exploratory experiment intended to identify effective features. *J. Teach. Educ.* 69 (1), 69–89.
- Bianchini, J.A., Dwyer, H.A., Brenner, M.E., Wearly, A.J., 2015. Facilitating science and mathematics teachers' talk about equity: what are the strengths and limitations of four strategies for professional learning? *Sci. Educ.* 99 (3), 577–610.
- Blomberg, G., Renkl, A., Sherin, M.G., Borko, H., Seidel, T., 2013. Five research-based heuristics for using video in pre-service teacher education. *J. Educ. Res.* 5 (1), 90. Online.
- Borko, H., Jacobs, J., Koellner, K., 2010. Contemporary approaches to teacher professional development. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, vol. 7. Elsevier, Oxford, pp. 548–556.
- Borko, H., Koellner, K., Jacobs, J., 2011a. Meeting the challenges of scale: the importance of preparing professional development leaders. *Teach. Coll. Rec.* Date Published: March 04, 2011 <http://www.tcrecprd.org> ID Number: 16358.
- Borko, H., Koellner, K., Jacobs, J., Seago, N., 2011b. Using video representations of teaching in practice-based professional development programs. *Zentralblatt für Didaktik der Math.* 43 (1), 175–187.
- Borko, H., Koellner, K., Jacobs, J., 2014. Examining novice teacher leaders' facilitation of professional development. *J. Math. Behav.* 33, 149–167.
- Borko, H., Jacobs, J., Koellner, K., Swackhamer, L., 2015. *Mathematics Professional Development: Improving Teaching Using the Problem-Solving Cycle and Leadership Preparation Models*. Teachers College Press, Columbia University, NY.
- Borko, H., Carlson, J., Mangram, C., Anderson, R., Fong, A., Million, S., Villa, A.M., 2017. The role of video-based discussion in model for preparing professional development leaders. *Int. J. STEM Educ.* 4 (1), 1–15.
- Borko, H., 2004. Professional development and teacher learning: mapping the terrain. *Educ. Res.* 33 (8), 3–15.
- Bowen, W.G., Chingos, M.M., Lack, K.A., Nygren, T.J., 2012. *Interactive Learning Online at Public Universities: Evidence From Randomized Trials*, vol. 22. Ithaca S+ R, pp. 1–53.
- Brown, J.C., Crippen, K.J., 2016. Designing for culturally responsive science education through professional development. *Int. J. Sci. Educ.* 38 (3), 470–492.
- Bryk, A.S., Gomez, L.M., Grunow, A., LeMahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press.
- Burns, R.J., Aspeslagh, R., 2014. *Three Decades of Peace Education Around the World: An Anthology*. Routledge.
- Byers, A., Mendez, F., 2016. Blended professional learning for science educators. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA, pp. 167–185.
- Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C.P., Loe, M., 1989. Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *Am. Educ. Res. J.* 26 (4), 499–531.
- Carter Andrews, D.J., Richmond, G., 2019. Professional development for equity: what constitutes powerful professional learning. *J. Teach. Educ.* 70 (5), 408–409.
- Carter Andrews, D.J., Bartell, T., Richmond, G., 2016. Teaching in dehumanizing times: the professionalization imperative. *J. Teach. Educ.* 67 (3), 170–172.
- Charteris, J., Berman, J., Page, A., 2021. Virtual Team Professional Learning and Development for Practitioners in Education. *Professional Development in Education*, pp. 1–13.
- Cloonan, A., Fox, B., Ohi, S., Halse, C., 2017. An analysis of the use of autobiographical narrative for teachers' intercultural learning. *Teach. Educ.* 28 (2), 131–144.
- Cobb, D.J., Couch, D., 2018. Teacher education for an uncertain future: implications of PISA's Global Competence. In: *Teacher Education in and for Uncertain Times*. Springer, Singapore, pp. 35–47.
- Cochran-Smith, M., 2015. Teacher communities for equity. *Kappa Delta Pi Rec.* 51 (3), 109–113.
- Codding, D., Mouza, C., Pollock, L., Sheridan, S., 2019. Culturally responsive and equity-focused computer science professional development. In: *Society for Information Technology and Teacher Education International Conference. Association for the Advancement of Computing in Education (AACE)*, pp. 648–656.
- Darling-Hammond, L., Hyler, M.E., Gardner, M., 2017. *Effective Teacher Professional Development*. Learning Policy Institute.
- Dash, S., de Kramer, R.M., O'Dwyer, L.M., Masters, J., Russell, M., 2012. Impact of online professional development on teacher quality and student achievement in fifth grade mathematics. *J. Res. Technol. Educ.* 45 (1), 1–26. <https://doi.org/10.1080/15391523.2012.10782595>.
- Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., McCloskey, E.M., 2009. A research agenda for online teacher professional development. *J. Teach. Educ.* 60 (1), 8–19.
- Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), 2016. *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA.
- Desimone, L.M., Garet, M.S., 2015. Best practices in teachers' professional development in the United States. *Psychol. Soc. Educ.* 7, 252–263.
- Desimone, L.M., Pak, K., 2017. Instructional coaching as high-quality professional development. *Theory Into Pract.* 56 (1), 3–12.
- Dreher, A., Kuntze, S., 2015. Teachers' professional knowledge and noticing: the case of multiple representations in the mathematics classroom. *Educ. Stud. Math.* 88 (1), 89–114. <https://doi.org/10.1007/s10649-014-9577-8>.
- Erickson, A., Noonan, P., McCall, Z., 2012. Effectiveness of online professional development for rural special educators. *Rural Spec. Educ. Q.* 31, 22–32. <https://doi.org/10.1177/875687051203100104>.
- Even, R., Robinson, N., Carmeli, M., 2003. The work of providers of professional development for teachers of mathematics: two case studies of experienced practitioners. *Int. J. Sci. Math. Educ.* 1 (2), 227–249.
- Even, R., 2014. Challenges associated with the professional development of didacticians. *ZDM Math. Educ.* 46 (2), 329–333.
- Ferdig, R.E., Baumgartner, E., Hartshorne, R., Kaplan-Rakowski, R., Mouza, C. (Eds.), 2020. *Teaching, Technology, and Teacher Education during the COVID-19 Pandemic: Stories from the Field*. Association for the Advancement of Computing in Education (AACE). Retrieved from. <https://www.learnlib.org/p/216903/>.
- Fisher, J.B., Schumaker, J.B., Culbertson, J., Deshler, D.D., 2010. Effects of a computerized professional development program on teacher and student outcomes. *J. Teach. Educ.* 61 (4), 302–312.

- Fishman, B., Konstantopoulos, S., Kubitskey, B.W., Vath, R., Park, G., Johnson, H., Edelson, D.C., 2013. Comparing the impact of online and face-to-face professional development in the context of curriculum implementation. *J. Teach. Educ.* 64 (5), 426–438.
- Fishman, B., 2016. Possible futures for online teacher professional development. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA, pp. 13–30.
- Foot, M. (Ed.), 2010. *Mathematics Teaching and Learning in K-12: Equity and Professional Development*. Palgrave Macmillan, New York, NY.
- Francis, K., Jacobsen, M., 2013. Synchronous online collaborative professional development for elementary mathematics teachers. *Int. Rev. Res. Open Dist. Learn.* 14 (3), 319–343. <https://doi.org/10.19173/irrodl.v14i3.1460>.
- Garza, A., 2014. A herstory of the #BlackLivesMatter movement. In: *Are All the Women Still White? Rethinking Race, Expanding Feminisms*, pp. 23–28.
- Gaudin, C., Chaliès, S., 2015. Video viewing in teacher education and professional development: a literature review. *Educ. Res. Rev.* 16, 41–67.
- Gay, G., 2002. Preparing for culturally responsive teaching. *J. Teach. Educ.* 53 (2), 106–116.
- Goldsmith, L., Seago, N., 2013. *Examining Mathematics Practice through Classroom Artifacts*. Pearson, Boston, MA.
- Gonçalves, L.L., Parker, M., Liguetti, C., Carbinatto, M., 2020. The facilitator's role in supporting physical education teachers' empowerment in a professional learning community. *Sport Educ. Soc.* 1–14.
- González, G., Deal, J.T., Skultety, L., 2016. Facilitating teacher learning when using different representations of practice. *J. Teach. Educ.* 67 (5), 447–466.
- Goode, J., Ivey, A., Johnson, S.R., Ryoo, J.J., Ong, C., 2020a. Racializing to computer science for all: how teachers talk and learn about equity in professional development. *Comput. Sci. Educ.* 1–26.
- Goode, J., Johnson, S.R., Sundstrom, K., 2020b. Disrupting colorblind teacher education in computer science. *Prof. Dev. Educ.* 46 (2), 354–367.
- Gröschner, A., Seidel, T., Pehmer, A.K., Kiemer, K., 2014. Facilitating collaborative teacher learning: the role of "mindfulness" in video-based teacher professional development programs. *Gruppendynamik Organ.* 45 (3), 273–290.
- Gröschner, A., Seidel, T., Kiemer, K., Pehmer, A.K., 2015. Through the lens of teacher professional development components: the 'Dialogic Video Cycle' as an innovative program to foster classroom dialogue. *Prof. Dev. Educ.* 41 (4), 729–756.
- Gueudet, G., Pepin, B., Sabra, H., Trouche, L., 2016. Collective design of an e-textbook: teachers' collective documentation. *J. Math. Teach. Educ.* 19 (2–3), 187–203.
- Hadar, L.L., Brody, D.L., 2017. Professional learning and development of teacher educators. In: *Handbook of Research on Teacher Education*, vol. 2, pp. 1049–1064.
- Hartshorne, R., Baumgartner, E., Kaplan-Rakowski, R., Mouza, C., Ferdig, R.E., 2020. Special issue editorial: preservice and inservice professional development during the COVID-19 pandemic. *J. Technol. Teach. Educ.* 28 (2), 137–147.
- Hawkins, M.R. (Ed.), 2011. *Social Justice Language Teacher Education*. Multilingual Matters, Tonawanda, NY.
- Hirsch, S.E., Randall, K.N., Common, E.A., Lane, K.L., 2020. Results of practice-based professional development for supporting special educators in learning how to design functional assessment-based interventions. *Teach. Educ. Spec. Educ.* 43 (4), 281–295.
- Hodges, C., Lowenthal, P., Grant, M., 2016. Teacher professional development in the digital age: design considerations for MOOCs for teachers. In: *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE), pp. 2075–2081.
- Huang, R., Shimizu, Y., 2016. Improving teaching, developing teachers and teacher developers, and linking theory and practice through lesson study in mathematics: an international perspective. *ZDM Math. Educ.* 48 (4), 439–587.
- Huang, R., Gong, Z., Han, X., 2016. Implementing mathematics teaching that promotes students' understanding through theory-driven lesson study. *ZDM* (4), 425–439.
- Hubers, M.D., Enderdijk, M.D., Van Veen, K., 2020. Effective characteristics of professional development programs for science and technology education. *Prof. Dev. Educ.* <https://doi.org/10.1080/19415257.2020.1752289>.
- Huggins, K.S., Lesseig, K., Rhodes, H., 2017. Rethinking teacher leader development: a study of early career mathematics teachers. *Int. J. Teach. Leader.* 8 (2), 28–48.
- Ilaria, D., 2017. The efficacy and impact of a hybrid professional development model on handheld graphing technology use. *Contemp. Issues Technol. Teach. Educ.* 17 (2), 194–204. Waynesville, NC USA: Society for Information Technology & Teacher Education. Retrieved December 3, 2020 from: <https://www.learntechlib.org/primary/p/149873/>.
- Ince, A., 2017. Managing risk in complex adult professional learning: the facilitator's role. *Prof. Dev. Educ.* 43 (2), 194–211.
- Jaczi, A.P., Schellinger, A., Lin, L., Zacamy, J., Toby, M., 2016. Effectiveness of Internet-Based Reading Apprenticeship Improving Science Education (iRAISE): A Report of a Randomized Experiment in Michigan and Pennsylvania (Empirical Education Rep.No. Empirical_ iRAISE-7027-FR1-Y2-O.3). Empirical Education Inc, Palo Alto, CA. Retrievable at: <http://empiricaleducation.com/pdfs/iRAISEfr.pdf>.
- Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C., Applegate, M., 2015. Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM Math. Educ.* 47 (1), 93–104.
- Jacobs, J., Seago, N., Koellner, K., 2017. Preparing facilitators to use and adapt mathematics professional development materials productively. *Int. J. STEM Educ.* 4, 30. <https://doi.org/10.1186/s40594-017-0089-9>.
- Jensen, B., Sonnemann, J., Roberts-Hull, K., Hunter, A., 2016. Beyond PD: Teacher Professional Learning in High-Performing Systems. *Teacher Quality Systems in Top Performing Countries*. National Center on Education and the Economy, Washington, DC.
- Jilk, L.M., 2016. Supporting teacher noticing of students' mathematical strengths. *Math. Teach. Educ.* 4 (2), 188–199.
- Johnson, C.C., 2011. The road to culturally relevant science: exploring how teachers navigate change in pedagogy. *J. Res. Sci. Teach.* 48 (2), 170–198.
- K-12 Computer Science Framework Steering Committee, 2016. *K-12 Computer Science Framework*. Retrieved from: <https://k12cs.org/>.
- Karsenty, R., Arcavi, A., 2017. Mathematics, lenses and videotapes: a framework and a language for developing reflective practices of teaching. *J. Math. Teach. Educ.* 20, 433–455. <https://doi.org/10.1007/s10857-017-9379-x>.
- Kennedy, M.M., 2016. How does professional development improve teaching? *Rev. Educ. Res.* 86 (4), 945–980.
- Killion, J., 2013. *Meet the Promise of Content Standards: Tapping Technology to Enhance Professional Learning*. Learning Forward, Oxford, OH.
- Kleiman, G., Wolf, M.A., 2016. Going to scale: the MOOCs for educators initiative. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*, pp. 49–68.
- Kleinknecht, M., Schneider, J., 2013. What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teach. Teach. Educ.* 33, 13–23.
- Koellner, K., Jacobs, J., 2015. Distinguishing models of professional development: the case of an adaptive model's impact on teachers' knowledge, instruction, and student achievement. *J. Teach. Educ.* 66 (1), 51–67.
- Kose, B.W., Lim, E.Y., 2010. Transformative professional development: relationship to teachers' beliefs, expertise and teaching. *Int. J. Leader. Educ.* 13 (4), 393–419.
- Krainer, K., 2015. Reflection on the increasing relevance of large-scale professional development. *ZDM Math. Educ.* 47 (1), 143–151. <https://doi.org/10.1007/s11858-015-0674-7>.
- Ladson-Billings, G., 1992. Reading between the lines and beyond the pages: a culturally relevant approach to literacy teaching. *Theory Into Pract.* 31 (4), 312–320. <https://doi.org/10.1080/00405849209543558>.
- Ladson-Billings, G.J., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 35, 465–491.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: A.k.a. the remix. *Harv. Educ. Rev.* 84 (1), 74–84. <https://doi.org/10.17763/haer.84.1.p2rj131485484751>.
- Larriave, B., 2000. Transforming teacher practice: becoming the critically reflective teacher. *Reflective Pract.* 1, 293–307.
- Lefstein, A., Louie, N., Segal, A., Becher, A., 2020. Taking stock of research on teacher collaborative discourse: theory and method in a nascent field. *Teach. Teach. Educ.* 88, 102954.
- Lesseig, K., Elliott, R., Kazemi, E., Kelley-Petersen, M., Campbell, M., Mumme, J., Carroll, C., 2017. Leader noticing of facilitation in videocases of mathematics professional development. *J. Math. Teach. Educ.* 20 (6), 591–619.
- Lewis, C., Perry, R., 2017. Lesson study to scale up research-based knowledge: a randomized, controlled trial of fractions learning. *J. Res. Math. Educ.* 48 (3), 261–299.

- Lewis, C., 2015. What is improvement science? Do we need it in education? *Educ. Res.* 44 (1), 54–61.
- Lieberman, A., Pointer Mace, D., 2010. Making practice public: teacher learning in the 21st century. *J. Teach. Educ.* 61 (1–2), 77–88.
- Lockee, B.B., 2020. Shifting digital, shifting context(re) considering teacher professional development for online and blended learning in the COVID-19 era. *Educ. Technol. Res. Dev.* 1–4.
- Lugueti, C.N., Oliver, K., Parker, M.A., 2020. Facilitation as an act of love: a self-study of how a facilitator's pedagogy changed over time in the process of supporting a community of learners. *J. Teach. Phys. Educ.* 1–8.
- Luna, M.J., Sherin, M.G., 2017. Using a video club design to promote teacher attention to students' ideas in science. *Teach. Teach. Educ.* 66, 282–294.
- Lynch, K., Hill, H.C., Gonzalez, K.E., Pollard, C., 2019. Strengthening the research base that informs STEM instructional improvement efforts: a meta-analysis. *Educ. Eval. Pol. Anal.* 41 (3), 260–293.
- Macià, M., García, I., 2016. Informal online communities and networks as a source of teacher professional development: a review. *Teach. Teach. Educ.* 55, 291–307.
- Major, L., Watson, S., 2018. Using video to support in-service teacher professional development: the state of the field, limitations and possibilities. *Technol. Pedagog. Educ.* 27 (1), 49–68.
- Marrongelle, K., Sztajn, P., Smith, M., 2013. Scaling up professional development in an era of common state standards. *J. Teach. Educ.* 64 (3), 202–211.
- Marsh, B., Mitchell, N., 2014. The role of video in teacher professional development. *Teach. Dev.* 18 (3), 403–417.
- Martin, D.B., 2012. Learning mathematics while Black. *Educ. Found.* 26, 47–66.
- Masters, J., Magidin de Kramer, R., O'Dwyer, L., Dash, S., Russell, M., 2012. The effects of online teacher professional development on fourth grade students' knowledge and practices in English language arts. *J. Technol. Teach. Educ.* 20 (1), 21–46. Waynesville, NC USA: Society for Information Technology & Teacher Education. Retrieved December 4, 2020 from: <https://www.learntechlib.org/primary/p/35346/>.
- Mayorga, E., Picower, B., 2018. Active solidarity: centering the demands and vision of the Black Lives Matter movement in teacher education. *Urban Educ.* 53 (2), 212–230.
- McDonald, M., Kazemi, E., Kavanaugh, S., 2013. Core practices of teacher education: a call for a common language and collective activity. *J. Teach. Educ.* 64 (5), 378–386.
- Merritt, E.G., 2016. Time for teacher learning, planning critical for school reform: students aren't the only ones who need more time to learn; teachers also need more and better time for learning and for planning. *Phi Delta Kappan* 98 (4), 31–36.
- Minkos, M.L., Gelbar, N.W., 2021. Considerations for educators in supporting student learning in the midst of COVID-19. *Psychol. Sch.* 58 (2), 416–426.
- Misra, P., 2018. MOOCs for teacher professional development: reflections and suggested actions. *Open Prax.* 10 (1), 67–77.
- Mogliacci, R.J., Raanuis, J., Howell, C., 2016. Supporting teachers in becoming agents of social cohesion: professional development in Post-apartheid South Africa. *Educ. Change* 20 (3), 160–179.
- Murphy, R., Snow, E., Mislevy, J., Gallagher, L., Krumm, A., Wei, X., 2014. Blended Learning Report. Retrieved from Michael and Susan Dell Foundation. <https://www.msdf.org/whitepapers/blended-learning-report/>.
- National Academies of Sciences, Engineering, and Medicine, 2015. Science Teachers Learning: Enhancing Opportunities, Creating Supportive Contexts, Committee on Strengthening Science Education Through a Teacher Learning Continuum. In: Board on Science Education and Teacher Advisory Council, Division of Behavioral and Social Science and Education. The National Academies Press, Washington, DC.
- National Academies of Sciences, Engineering, and Medicine, 2020. Changing Expectations for the K-12 Teacher Workforce: Policies, Preservice Education, Professional Development, and the Workplace. The National Academies Press, Washington DC.
- O'Dwyer, L.M., Masters, J., Dash, S., De Kramer, R.M., Humez, A., Russell, M., 2010. E-learning for Educators: Effects of On-Line Professional Development on Teachers and Their Students. Technology and Assessment Study Collaborative, Chestnut Hill, MA.
- Osborn, J.F., Borko, H., Fishman, E., et al., 2019. Impacts of a practice-based professional development on elementary teachers' facilitation of and student engagement with scientific argumentation. *Am. Educ. Res. J.* 56 (4), 1067–1112.
- Paris, D., Alim, H.S. (Eds.), 2017. Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World. Teachers College Press.
- Paris, D., 2017. On Educating Culturally Sustaining Teachers. Equity by Design. Equity Assistance Center Region III, Midwest and Plains Equity Assistance Center.
- Parker, C., Bickmore, K., 2020. Classroom peace circles: teachers' professional learning and implementation of restorative dialogue. *Teach. Teach. Educ.* 95, 103129.
- Parkhouse, H., Lu, C.Y., Massaro, V.R., 2019. Multicultural education professional development: a review of the literature. *Rev. Educ. Res.* 89 (3), 416–458.
- Parsons, S.A., Hutchison, A.C., Hall, L.A., Parsons, A.W., Ives, S.T., Leggett, A.B., 2019. U.S. teachers' perceptions of online professional development. *Teach. Teach. Educ.* 82, 33–42.
- Peltola, P., Haynes, E., Clymer, L., McMillan, A., Williams, H., 2017. Opportunities for Teacher Professional Development in Oklahoma Rural and Nonrural Schools (REL 2017–273). Department of Education, Institute of Education Sciences. National Center for education evaluation and regional assistance, regional educational Laboratory southwest. Retrieved from: <http://ies.ed.gov/ncee/edlabs>.
- Penuel, W.R., 2019. Co-design as infrastructuring with attention to power: building collective capacity for equitable teaching and learning through design-based implementation research. In: Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning. Springer, Cham, pp. 387–401.
- Perry, E., Boylan, M., 2018. Developing the developers: supporting and researching the learning of professional development facilitators. *Prof. Dev. Educ.* 44 (2), 254–271.
- Perry, E., 2020. The roles of professional development leaders. *Sci. Teach. Educ.* (88), 10–16.
- Philipsen, B., Tondeur, J., Roblin, N.P., Vanslambrouck, S., Zhu, C., 2019. Improving teacher professional development for online and blended learning: a systematic meta-aggregative review. *Educ. Technol. Res. Dev.* 67 (5), 1145–1174.
- Prediger, S., Roesken-Winter, B., Leuders, T., 2019. Which research can support PD facilitators? Strategies for content-related PD research in the Three-Tetrahedron Model. *J. Math. Teach. Educ.* 22, 407–425.
- Raza, A., Penuel, W., Jacobs, J., Sumner, T., 2020. Supporting equity in using visual learning analytics to understand learners' classroom experiences. In: Schmidt, M., Tawfik, A., Jahnke, A., Earnshaw, I. (Eds.), Learner and User Experience Research: An Introduction to the Field of Learning Design and Technology. Edtechbooks.
- Reinholz, D.L., Stone-Johnstone, A., Shah, N., 2020. Walking the walk: using classroom analytics to support instructors to address implicit bias in teaching. *Int. J. Acad. Dev.* 25 (3), 259–272.
- Roth, K.J., Garnier, H.E., Chen, C., Lemmens, M., Schwiller, K., Wickler, N.I.Z., 2011. Videobased lesson analysis: effective science PD for teacher and student learning. *J. Res. Sci. Teach.* 48 (2), 117–148.
- Roth, K.J., Bintz, J., Wickler, N.I., Hvidsten, C., Taylor, J., Beardsley, P.M., et al., 2017. Design principles for effective video-based professional development. *Int. J. STEM Educ.* 4 (1), 1–24.
- Roth, K.J., Wilson, C.D., Taylor, J.A., Stuhlsatz, M.A., Hvidsten, C., 2019. Comparing the effects of analysis-of-practice and content-based professional development on teacher and student outcomes in science. *Am. Educ. Res. J.* 56 (4), 1217–1253.
- Rubel, L.H., Chu, H., 2012. Reinscribing urban: teaching high school mathematics in low income, urban communities of color. *J. Math. Teach. Educ.* 15 (1), 39–52.
- Salomon, G., Cairns, E. (Eds.), 2011. Handbook on Peace Education. Taylor & Francis.
- Santagata, R., Yeh, C., 2014. Learning to teach mathematics and to analyze teaching effectiveness: evidence from a video-and practice-based approach. *J. Math. Teach. Educ.* 17 (6), 491–514.
- Santagata, R., Lee, J., Sandoval, C., 2019. Practice partnerships in mathematics teacher education. In: International Handbook of Mathematics Teacher Education, vol. 3. Brill Sense, pp. 183–210.
- Schachter, R.E., Gerde, H.K., Hatton-Bowers, H., 2019. Guidelines for selecting professional development for early childhood teachers. *Early Child. Edu. J.* 47 (4), 395–408.

- Schneider, S.A., Lepori, K.L., Carroll, C.E., Ramirez, A.B., Knotts, A.K., Silberglitt, M.D., Gale, M.A., Salguero, K., Luttgen, K.M., Hauk, S., Ringstaff, C., 2016. Exploring models of online professional development. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA, pp. 215–234.
- Schoenbach, R., Greenleaf, C., Brown, W., Howlett, H., 2016. Supporting deep change with an Immersive online course. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA, pp. 69–86.
- Schueler, S., Roesken-Winter, B., 2018. Compiling video cases to support PD facilitators in noticing productive teacher learning. *Int. J. STEM Educ.* 5 (50). <https://doi.org/10.1186/s40594-018-0147-y>.
- Scott, K.A., Clark, K., Hayes, E., Mruczek, C., Sheridan, K., 2010. Culturally relevant computing programs: two examples to inform teacher professional development. In: Society for Information Technology & Teacher Education International Conference. Association for the Advancement of Computing in Education (AACE), pp. 1269–1277.
- Seago, N., Mumme, J., Branca, N., 2004. *Learning and Teaching Linear Functions: Video Cases for Mathematics Professional Development*. Heinemann, San Diego, pp. 6–10 (University Foundation).
- Seago, N., Jacobs, J., Driscoll, M., Callahan, P., Matassa, M., Nicola, J., 2017. *Learning and Geometry: Video Cases for Mathematics Professional Development*. WestEd, San Francisco, CA.
- Seago, N., Knotts, A., Carroll, C., 2020. Designing flexible, asynchronous online professional development to impact teacher learning and practice. In: *Proceedings of the 13th International Conference of Education, Research and Innovation*, November 9–10, 2020.
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., Schwindt, K., 2011. Teacher learning from analysis of videotaped classroom situations: does it make a difference whether teachers observe their own teaching or that of others? *Teach. Teach. Educ.* 27, 259–267.
- Selmer, S., Bernstein, M., Bolyard, J., 2016. Multilayered knowledge: understanding the structure and enactment of teacher educators' specialized knowledge base. *Teach. Dev.* 20 (4), 437–457.
- Severance, S., Penuel, W.R., Sumner, T., Leary, H., 2016. Organizing for teacher agency in curricular co-design. *J. Learn. Sci.* 25 (4), 531–564.
- Sherin, M.G., van Es, E.A., 2009. Effects of video participation on teachers' professional vision. *J. Teach. Educ.* 60 (1), 20–37.
- Sherin, M.G., Linzenmeier, K.A., van Es, E.A., 2009. Issues in the design of video clubs: selecting video clips for teacher learning. *J. Teach. Educ.* 60 (3), 213–230.
- Sherin, M.G., Jacobs, V., Philipp, R., 2011. *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*. Routledge, New York.
- Stacy, J., Fernández, Y., McGovern, E.R., 2020. El instituto: centering language, culture, and power in bilingual teacher professional development. *J. Cult. Val. Educ.* 3 (2), 120–137.
- Strutchens, M., Bay-Williams, J., Civil, M., Chval, K., Malloy, C.E., White, D.Y., et al., 2012. Foregrounding equity in mathematics teacher education. *J. Math. Teach. Educ.* 15 (1), 1–7.
- Sztajn, P., Borko, H., Smith, T.S., 2017. Research on mathematics professional development. In: Cai, J. (Ed.), *Compendium for Research in Mathematics Education*. National Council of Teachers of Mathematics, Reston, VA, pp. 213–243.
- Sztajn, P., Dick, L., Alnizami, R., Heck, D., Malzahn, K., 2020. Controlled implementations. Teaching practice to practicing mathematics teachers. In: Linares, S., Chapman, Y.O. (Eds.), *International Handbook of Teacher Education, Tools and Processes in Mathematics Teacher Education*, second ed., vol. 2. Brill, Leiden, The Netherlands, pp. 285–310.
- Taylor, J.A., Roth, K., Wilson, C.D., Stuhlsatz, M.A., Tipton, E., 2017. The effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *J. Res. Educ. Effect.* 10 (2), 241–271.
- Tekkumru-Kisa, M., Stein, M.K., 2017. A framework for planning and facilitating video-based professional development. *Int. J. STEM Educ.* 4 (1), 28.
- Timmons-Brown, S., Warner, C., 2016. Using a conference workshop setting to engage mathematics teachers in culturally relevant pedagogy. *J. Urban Math. Educ.* 9 (1), 19–47.
- Tripp, J.R., Rich, P.J., 2012. The influence of video analysis on the process of teacher change. *Teach. Teach. Educ.* 28 (5), 728–739.
- UNESCO, 2014. *Teaching Respect for All: Implementation Guide*. UNESCO, Paris. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000227983>.
- Van Es, E.A., Tunney, J., Goldsmith, L.T., Seago, N., 2014. A framework for the facilitation of teachers' analysis of video. *J. Teach. Educ.* 65 (4), 340–356.
- van Es, E.A., Tekkumru-Kisa, M., Seago, N., 2019. Leveraging the power of video for teacher learning: a design framework for mathematics teacher educators. In: *International Handbook of Mathematics Teacher Education*, vol. 2. Brill Sense, pp. 23–54.
- van Es, E.A., 2011. A framework for learning to notice student thinking. In: Sherin, M.G., Jacobs, V., Philipp, R. (Eds.), *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*. Routledge, New York, NY, pp. 134–151.
- Verma, R., 2017. *Critical Peace Education and Global Citizenship: Narratives from the Unofficial Curriculum*. Taylor & Francis.
- Voogt, J., Westbroek, H., Handelzalts, A., Walraven, A., McKenney, S., Pieters, J., De Vries, B., 2011. Teacher learning in collaborative curriculum design. *Teach. Teach. Educ.* 27 (8), 1235–1244.
- Voogt, J., Laferrière, T., Breuleux, A., Itow, R.C., Hickey, D.T., McKenney, S., 2015. Collaborative design as a form of professional development. *Instr. Sci.* 43 (2), 259–282.
- Voogt, J.M., Pieters, J.M., Handelzalts, A., 2016. Teacher collaboration in curriculum design teams: effects, mechanisms, and conditions. *Educ. Res. Eval.* 22 (3–4), 121–140.
- Wager, A.A., Foote, M.Q., 2013. Locating praxis for equity in mathematics: lessons from and for professional development. *J. Teach. Educ.* 64 (1), 22–34.
- Wake, G., Swan, M., Foster, C., 2016. Professional learning through the collaborative design of problem-solving lessons. *J. Math. Teach. Educ.* 19 (2–3), 243–260.
- Westbroek, H., de Vries, B., Walraven, A., Handelzalts, A., McKenney, S., 2019. Teachers as co-designers: scientific and colloquial evidence on teacher professional development and curriculum innovation. In: *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham, pp. 35–54.
- Yolcu, A., 2019. Research on equitable mathematics teaching practices: insights into its divergences and convergences. *Rev. Educ.* 7 (3), 701–730.
- Yoon, S.A., Miller, K., Richman, T., Wendel, D., Schoenfeld, I., Anderson, E., Shim, J., 2020. Encouraging collaboration and building community in online asynchronous professional development: designing for social capital. *Intern. J. Comput. Support. Collab. Learn.* 15, 351–371. <https://doi.org/10.1007/s11412-020-09326-2>.
- Yuan, J., Kim, C., 2014. Guidelines for facilitating the development of learning communities in online courses. *J. Comput. Assist. Learn.* 30 (3), 220–232.
- Zeichner, K., 2011. Teacher education for social justice. In: *Social Justice Language Teacher Education*, pp. 7–22.
- Zhang, M., Lundeberg, M., Eberhardt, J., 2011. Strategic facilitation of problem-based discussion for teacher professional development. *J. Learn. Sci.* 20 (3), 342–394.

Teacher learning as workplace learning[☆]

Jeroen Imants^a and Klaas van Veen^{b, a} Radboud Teacher Academy, Radboud University Nijmegen, Zeist, the Netherlands; and
^b Institute of Teacher Education, University of Groningen, Groningen, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

This is an update of J. Imants, K. van Veen, Teacher Learning as Workplace Learning, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 569-574, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00657-6>.

Introduction	562
Teachers' professional development and learning	563
Workplace learning	564
Teachers' workplace	564
Workplace conditions in schools	565
Learning potential of the task	565
Opportunities for feedback, evaluation, and reflection on the outcomes of work actions	566
Formalization of work processes	566
Teacher participation in handling problems and developing work processes	566
Shared norms	567
Learning resources	567
Conclusions	568
References	569

Introduction

A central notion in theories of workplace learning is the intrinsic and mutual relationship between working and learning. In teacher learning as workplace learning three themes come together: (1) the content and organization of teachers' work, (2) the content and process of teachers' learning in the context of their classrooms and schools, and (3) teacher learning as continuous professional development during the teacher career.

During the last four decades, workplace learning has been an important item in discussions on vocational education and professional development in industrialized societies. These discussions cover the broad field of economic and cultural developments, technological and organizational innovations, knowledge economy and knowledge creation, and professional development and vocational learning in schools and work organizations. In the field of education, workplace learning has shown to be a promising approach for the development of productive learning contexts in teacher education and continuing professional development.

In this article, teachers' workplace learning is defined as changes in teaching practices in classrooms and schools that are mediated through individual and collective teacher learning and problem-solving processes in the school (Ellström, 2001). At the core of workplace learning are teaching practices in classrooms and schools, and insights in how these practices are related to student learning.

Teacher professional development is the generic term under which programs for teacher learning and formalized learning activities are organized and discussed. These programs and activities can be organized inside or outside the school. Teachers' professional development can be defined as those processes and activities designed to enhance the professional knowledge, skills, and attitudes of educators so that they might, in turn, improve the learning of students (Guskey, 2000). In a broader approach professional development is linked to the professional career of teachers, from initial teacher training to leaving the profession: a process filled with plateaus, discontinuities, regressions, spurts, and dead ends (Huberman, 1995). This process of lifelong learning also includes a variety of informal learning experiences in school as well as in private life outside the school. According to Kyndt et al. (2016) informal learning results from engagement in daily work-related activities in which learning is not the primary goal. Informal learning is undertaken autonomously, either individually or collectively, but without an instructor. It often happens spontaneously and unconsciously. From the learner's perspective, it is unintentional, and outcomes are unpredictable.

In this article we focus on teacher learning in the context of the school. We discuss teacher learning in the context of the school as a dimension ranging from formalized training and professional development to informal learning in classrooms and schools. This

[☆] *Change history:* December, 2020. Structure, focus and authors of the article are completely maintained, as compared to the 2010 version of the article in the third edition of the Elsevier Encyclopedia (volume 7, pp. 569–574). When we prepared the manuscript for the third edition in 2008 teachers' workplace learning was a relatively new topic of study. During the last decades a growing body of research results has been published. This is accompanied by the development of new theoretical perspectives on teacher learning and professional development, and on research approaches that differ from the traditional evaluation method for professional development research. These developments are given attention in the actualized article for the fourth edition. Nevertheless, the main conclusions of the third edition article have been maintained.

also includes off site programs in which implementation in classroom and school is part of the course. In many instances, teacher workplace learning is a mix of formal and informal learning activities.

Recently, professional development programs often build on workplace learning insights. The aim of this article is to explore the field of teachers' workplace learning in classrooms and schools, as it builds on, and adds new insights to teachers' professional development and teacher learning in schools. The question is how and in what directions teacher learning can be promoted in the context of work in classrooms and schools. First, the state of the art in teachers' professional development and learning will be shortly characterized. Next, recent insights and research results in teachers' workplace learning will be discussed.

Teachers' professional development and learning

Teachers' professional development and teacher learning literature shows a magnitude of insights regarding views of learning (acquisition vs. constructivism), designs (fragmented vs. ongoing and systematic), and opportunities (formal and informal, mandatory and voluntary, serendipitous and planned). Lifelong curriculum of teachers' professional development seems fragmented and incoherent, as [Borko \(2004\)](#) observed, which seems to be caused not so much by a lack of understanding, rather a lack of a collective sense of urgency of teacher professional development and learning.

This sense of urgency and the rationales for professional development diverge considerably among schools, districts and countries. In general, two main rationales can be distinguished, which can differ depending on the quality of schools and teachers. A first rationale, often used in contexts where the quality of schools and teachers strongly (are assumed to) differ, argues that educational reforms are needed and require a great deal of learning on the part of teachers. Teachers need support and advice to understand the reform and work according to the reform. A second rationale, often found in contexts where the quality of schools and teachers is (assumed to be) high in general, starts from the nature of teachers' work. Student learning itself and teaching as the organization of student learning processes are assumed to be complex processes, in which teachers can continuously learn, develop and improve themselves. The first rationale assumes that teachers show personal inadequacies in need of repair, while the second rationale assumes that teacher learning contributes to fulfilment as a practitioner in support of students' learning.

Attention is growing for teacher learning and teacher professional development programs situated in the workplace in contrast to off-site programs. The assumption is that on-site would be more effective than off-site. However, in general there is hardly any empirical evidence for this assumption, because effectiveness of both is determined by the learning goals and both are found to be relatively effective in this respect. Off-site activities entail workshops, conference sessions, seminars, lectures, and other short-term training events on subject matter issues and topics like cooperative learning and classroom management. Design elements and conditions that can enhance the effectiveness of these activities are longer duration, connection with teachers' work context, collective participation, focus on subject matter and especially how students learn specific content, and emphasis on analysis and reflection ([Desimone, 2009](#); [Van Driel et al., 2012](#)). By identifying these design elements and conditions for off-site learning, it is a small step toward recent on-site professional development methods that are embedded in daily work in classrooms and schools. These work-embedded methods include training within the school or local context, creation of ongoing professional communities, teacher design teams, lesson study, and inquiry groups. In addition to the previous remarks this shift toward workplace learning can be summarized as:

- from a focus on teacher knowledge, skills, and teaching competencies, including new instructional methods, to a focus on student learning and specific teacher problems,
- from single sessions or a short series to long-term and ongoing,
- from a focus on change as something that is done to teachers and programs that change teachers as passive participants to change as a complex process that involves learning and teachers as active learners,
- from a focus on individual teacher learning to collective teacher learning in terms of professional learning communities.

As far as research has been executed regarding the effectiveness of learning activities that are embedded in daily work, it shows that teacher learning is promoted by a focus on subject matter knowledge for teaching, understanding of students' thinking, and instructional practices. Further, the process of meaningful learning is slow and uncertain, and some elements of teacher knowledge are easier to change than others. It takes time to develop a community. Teachers have little experience engaging in a professional discourse that is public and critical of their work and the work of colleagues ([Desimone, 2009](#); [Wilson and Berne, 1999](#)). In general, several methods for workplace learning have been developed (*coaching, supervision, and collegial consultation; mentoring and master-mate learning; development tasks and action-research; instruction in the workplace and application in the workplace; inquiry groups as data teams, learning studios and lesson study*).

Despite an increasing body of knowledge on design elements and conditions, there is little consensus about how professional development works, how the different elements relate, what teachers learn, how they learn, and how the professional development and this knowledge improve their practice. As [Kennedy \(2016\)](#) argues, all those lists of design elements and conditions are not really helpful in understanding how and why professional development works. Rather, what is needed is an understanding of the underlying theories of action. Such theory of action includes the central problems of practice that it aims to inform and the pedagogy that will help teachers enact new ideas, while translating these ideas into the context of their own practice. In addition to this theory of action, it is also useful to understand the theories of instruction and theories of context. Theory of instruction focuses on student results and refers to the relationships between the features of the professional development intervention, the intended changes in

teacher knowledge and instruction, and the expected changes in student outcomes. Theory of context refers to the school's organizational conditions that are necessary for implementation and sustainability of high-quality professional development in the school or in the classrooms.

Regarding the theory of context, most professional development research hardly takes the conditions of the daily workplace into account, while these conditions strongly determine the opportunities and limitations of the specific interventions and teacher learning (Bryck, 2015; Van Driel et al., 2012). Often the focus is only on the characteristics of the professional development initiative as an explanation for its effectiveness. When the intervention does not produce positive results, either teachers are blamed or the initiative is quickly discarded for the next new promising initiative. By contrast, examining the role of the organizational context can provide a more comprehensive understanding of the practice of how professional development unfolds (Kennedy, 2016). A useful concept to think through the organizational context for teacher learning is organizational routines, derived from organizational literature (Spillane et al., 2011). Organizational routines are generally defined as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (Feldman and Pentland, 2003, p. 95). The concept allows to move beyond the general lists of organizational conditions and enables an investigation of how work in organizations gets done: how it is created and planned, implemented, maintained and developed. Effective organizational routines are perceived as useful and more or less self evident. For example, in cases where professional development is implemented in the school and the classroom without aligning existing organizational routines with the professional development content, the position of the professional development effort will be isolated, resulting in low effectiveness and finally extinction. Another example refers to the role of leadership: organizational routines can reveal not just *that* leadership matters but also *how*. The concept has been used to show how leaderships' management of a new professional development promoted or constrained its implementation and sustainability.

In summary, the literature on teacher professional development and learning in schools has shown a shift toward aspects of workplace learning. Aspects of workplace learning have been implemented in professional development practices. At the same time a lack of understanding of the organization of professional development and learning opportunities and conditions of the workplace itself are observed. In the next sections the analysis of teachers' professional development and learning will be broadened by taking the mutual relationship between teachers' working and learning as the starting point.

Workplace learning

The identification of professional development and teacher learning in workplace contexts depends on the perspective from which learning is analyzed. Perspectives on workplace learning differ with regard to goals, outcomes, rationales, and managerial views. *Firstly*, workplace learning varies from adaptive to developmental (Ellström, 2001). In adaptive learning tasks, methods, and results are given, for example a new teacher in a school who is acquiring skills in working with this school's curriculum and instruction method. In developmental learning some or all of these aspects are not given, for example teachers in a subject department developing and evaluating innovative higher order thinking assignments for students.

Secondly, a broad range of possible learning outcomes for teachers in the workplace can be found, ranging from habitual reaction in which conservation is central to innovative reflective practices and inquiry (Bredeson, 2003). *Thirdly*, these outcomes can be realized at different interrelated levels of the school:

- (1) Individual or personal learning by teachers or school leaders within the context of the school,
- (2) Social learning in small groups or teams of teachers,
- (3) Learning that occurs across the school organization as a whole.

Regarding the level of learning outcomes in the school as a whole, a recurring problem is the lack of so called "snowballing". This is the diffusion through the whole school of new successful practices that were developed by an individual teacher or a small team. Moreover, in the context of social learning in teams, it is not uncommon that teachers report their learning results individually as personal learning results (Imants et al., 2020). *Finally*, as was discussed in the introduction, workplace learning ranges between formal and informal learning.

Teachers' workplace learning occurs in diverging interrelated contexts. In teacher education settings, workplace learning primarily is regarded as preparation for working, and qualification is the goal. In ongoing work settings in schools, teachers' workplace learning is regarded as contributing to productivity of schools, often as informal learning. This learning is assumed to be an intrinsic quality of professional work. In school reform settings, teachers' workplace learning is closely linked to the vitality of the school and to school improvement. In the individual perspective, teachers' workplace learning is regarded as learning for life and individual professional development.

Teachers' workplace

Formally the school as a workplace consists of organizational structures, like separate units that are charged with diverging tasks and responsibilities. However, the workplace of teachers is not restricted to these formal structures. Informally the school contains several spheres of influence and interaction, for example groups of teachers across formal units, or personal links between teachers and school leaders. For teachers, the work with students, inside and outside the classroom, makes up the

core of their work. Opportunities to enhance students' learning are central drives for teachers to participate in professional development. The largest part of their working day, teachers work individually with (a) group(s) of students, rather isolated from their colleagues. The most meaningful work experiences that promote teachers' workplace learning will therefore be gained in three interrelated activities:

- in subject and task specific instruction of students and in the guidance of these students,
- in the student related work with colleagues,
- in the ongoing enactments of organizational routines that are closely linked to the work with students and colleagues.

Together, this is the community of practice in which individual teachers participate (Wenger, 1998), and in which teachers' workplace learning can be expected to be primarily located. These communities include the teachers' own spheres of influence and interaction, and the formal structures in which their work with students and colleagues is organized.

School conditions and processes that are located in the administrators' sphere can be assumed to affect teacher learning processes and results in an indirect way, more specifically, as mediated by processes of reinvention and reinterpretation by teachers in their own sphere (Weick, 1995). For example, an educational policy document of a school or district will not influence teachers' learning directly. However, depending on how teachers interpret the school's educational policy, this document will affect teacher learning indirectly in foreseen and unforeseen directions (Luttenberg et al., 2013).

Schools as workplaces for teachers afford opportunities for learning. From their part, individual teachers elect how to engage in work activities and how to deal with the guidance provided by the workplace. This dual base of affordance and engagement for participation at work in the school, including the interactions between affordance and engagement, is central to understanding the kinds of teacher learning that schools are able to provide (Billett, 2001). Teacher agency plays a central role in how teachers perceive and enact these affordances. At the same time availability of resources and opportunities to enact these resources function as constraining or stimulating factors in teachers' engagement for learning (Eteläpelto et al., 2013; Imants and Van der Wal, 2020).

Workplace conditions in schools

The educational systems in industrialized countries are characterized by a long period of restructuring and standardization of work in schools, aimed at improved student learning results and the promotion of a new professionalism in the teachers' profession. Restructuring and standardization did have a significant but contradictory impact on the content, processes, and effects of teachers' workplace learning in schools. An example is the result of short-term improvements in scores on standardized tests for students combined with long-term tendencies of bureaucratization in schools and intensification in teachers' work. As a result, the long-term actual impact can be opposed to the formal aims of reform. This complex relationship between working and learning will be elaborated in this section.

In his review of general workplace learning research, Ellström (2001) distinguishes five groups of factors that can be assumed to be critical for *facilitating* and *constraining* the integration of learning and work: (1) the learning potential of the task, (2) opportunities for feedback, evaluation, and reflection on the outcomes of work actions, (3) formalization of work processes, (4) employee participation in handling problems and developing work processes, and (5) learning resources. Besides, schooling is a strongly moral and normative enterprise. In overviews of workplace conditions for teacher learning shared norms are often mentioned as a critical factor. Shared norms are added as a group of school factors to Ellström's general overview.

These general factors reasonably summarize the results of overviews of workplace learning in schools from the last 20 years. More specifically, the results of two recent review studies on informal and semi-formal forms of workplace learning (Kyndt et al., 2016), and teacher collaboration (Vangrieken et al., 2015) are frequently used in the following sections. Besides, we refer several times to the OECD (2013) TALIS study (Teaching and Learning International Survey). This large scale study presents an overview of workplace learning activities and outcomes for lower secondary teachers in 17 OECD countries.

Learning potential of the task

Generally, task complexity, variety, autonomy, and being in control are regarded as important determinants of the learning potential of the workplace. Traditionally, teacher work is viewed as complex, varied, and autonomous. In that respect, the school seems to be a favorable context for workplace learning. A recent review shows that individual teachers' autonomy serves as a foundation for informal learning activities like observation of colleagues without direct interaction, and individually consulting sources and media. However, while individual autonomy can be a positive workplace context for learning activities, teacher isolation within the school often is regarded as a restrictive condition. TALIS data indicate that most teachers in lower secondary education are still teaching largely in isolation, as over half of the teachers report very rarely or never team-teaching with colleagues, and two-thirds report the same rates for observing their colleagues teach. In general, developments in organizational routines can serve as impediments and opportunities for workplace learning. For example, with regard to variety as a determinant for workplace learning, studies on teachers' work show loss of variety (a combination of narrowing and formalizing the traditional teaching tasks), and simultaneously unwanted variety (adding new non-teaching tasks to teachers' work).

Opportunities for cross-school and cross-department exploration of teaching tasks expand workplace learning. Limited access to task accomplishment in other contexts and a narrow focus on individual task performance restrict workplace learning. Positive recognition of high level skills for special task performance expands workplace learning. On the other hand, workplace learning

is restricted when special skills for task performance are taken for granted. In this respect the egalitarian culture in schools can serve as a hindrance for effective workplace learning. Redesign of work in schools can expand effective workplace learning when the focus is on crucial problems of curriculum and instruction, and when teacher agency and engagement are promoted.

Opportunities for feedback, evaluation, and reflection on the outcomes of work actions

Activities in which teachers reflect on students' performance, work and thinking can promote teacher learning. Reflection in action frequently is identified as an activity in which teachers engage during informal learning. Reflection in action is promoted by activities like observing colleagues, encountering difficulties and unplanned situations, and developing and applying new practices combined with testing by experiencing and experimenting.

Feedback plays a crucial role in these reflective activities. Feedback on the results of work activities is considered to be necessary for learning to occur. The effects of feedback can be cognitive as well as motivational. About half of the teachers in the TALIS study report never receiving feedback on their teaching from their school leader, or other members of the school management. When these teachers receive formal or informal feedback on their practice, it is in a variety of ways, from a variety of sources. Almost 80% report receiving feedback following classroom observation, and nearly two-thirds report receiving feedback following an analysis of student test scores. Nearly nine in ten of these teachers report that student performance, teachers' pedagogical competency in their subject, and classroom management are emphasized in the feedback they receive.

The function of feedback depends on the existence of clear goals. At the level of classrooms and schools educational goals often are vague or inconsistent, and they can change during teaching. The problem of absence of clear feedback because of unclear goals can be an obstacle for teacher learning. On the other hand, explicating specific goals and aligning instruction to these goals can be assumed to promote teacher learning.

In schools, much information is available that could serve as feedback and promote reflection on the quality of teaching and learning, for example school self-evaluation reports, reports by the inspectorate, files of student assessment outcomes, aggregated student test results, and teachers' accounts of classroom experience (Little, 2012). Systematic use of these data on student performance, work, and thinking as sources for teacher and team learning hardly occurs in schools. Work with "data teams" shows that data-based decision making in schools is a promising instrument for teacher and student learning (Schildkamp et al., 2013).

Formalization of work processes

The daily educational and pedagogical practices in the school are important opportunities and constraints for teacher learning. This learning is focused on subject content, pedagogic and instructional methods, and classroom management (Scribner, 1999). Traditional curricular forms of formalization (fixed content, regulations, and procedures) and new forms of standardization (testing and output control) can be distinguished. Both formalization and standardization of daily work processes and outcomes in schools tend to dominate instruction, classroom management, and organization in schools. For example, a persistent problem in education is the limited flexibility to deal with student diversity in schools and classrooms, notwithstanding inclusive aims. Moreover, dominant images of how schools and classrooms should be structured and function still build on traditional scientific management principles, combined with the new public management model (NPM). A gap exists between the innovative expectations on workplace learning in schools, and the highly formalized instructional practices and images of schools.

However, starting from the observations that (a) work in schools is strongly regulated by informal and recurring practices and by accompanying implicit expectations, and (b) that the daily knowledge base of teachers is strongly tacit, a collaborative process of aligning instruction and guidance practices can also contribute to teacher learning. This learning by joint formalization should not be confused with simply replacing an old instructional method by a new one (Smith and Gillespie, 2007). Interactions and discussions with colleagues and others are often mentioned activities for informal learning. Mutual adjustment of instruction and guidance to shared, explicit goals for student learning can serve as the starting point for a collegial dialogue on effective instruction. Fixed instructional practices are questioned and tacit beliefs are made explicit and open for discussion. Examples of contexts for such professional discussions are Lesson Study, teacher design team or data team, and PLC. In that case, the aim is for teachers to get a better understanding of (1) what and how their students' learn and think, and (2) how student learning is promoted by teachers' instruction and guidance.

Teacher participation in handling problems and developing work processes

Learning is facilitated if teachers have access to and are able to successfully participate in problem handling and activities aimed at the development of improvements in instruction and guidance, both formally and informally. Interactions, discussions, and sharing with colleagues are frequently observed informal learning activities. This also holds for collaboration in committees and networks. Teachers share positive perceptions of participation in decision making in schools with regards to instructional, pedagogical, and curricular topics, but they perceive participation in decision making on administrative topics less favorably.

Teacher participation in decision making can create frustration, anger, and distrust when the implementation of decisions that were made is blocked or not promoted by school management. Teachers should be able to participate in all the steps of cycles in which improvements are developed: planning, development, performance, and evaluation. When the majority of teachers is only

expected to perform instructional tasks that are planned, developed, and evaluated by teacher leaders and managers, learning will decrease.

Task interdependence in teachers' work

Three subject teachers are asked by their unit leader to grade tests of thirty students who made the test for a colleague, and who is absent due to illness. Because of the study progress of these students and the school regulations regarding grading there is time pressure on this task. In such situations teachers often collaborate in the form of division of tasks that are executed individually. In such case every teacher takes ten tests for grading. However, these three teachers recognize that the absent colleague prepared the test from a specific point of view that differs from their own viewpoints. So they decided that before grading, they should explore and discuss the specific approach of their colleague, to realize a fair and uniform grading of the students' work. As a result the grading transformed from a divided quick task for individuals with low interdependence into a joint problem solving task with high interdependence. A rich learning context was created based on their own enactment of the unit leader's request. Besides interdependence, this case illustrates the role of affordances and teachers' agency and engagement.

Teachers' collegial interactions vary in task interdependence and alignment (synergy in directions of aims and actions of different actors). High alignment and interdependence, for example in the context of joint problem solving, are associated with richer and more stimulating learning environments for teachers. This collaboration can produce shared understanding and investment, thoughtful development, and a fair, rigorous test of selected ideas. Lower levels of interdependence and alignment occur when teachers are exchanging stories, tips, or materials. This can result in reinforcement of and emotional support for maintaining low-quality practices (Rosenholtz, 1989). Moreover, contrived forms of collaboration and balkanization within the school of isolated groups with groupthink show the risk of counterproductive effects of collaboration.

Shared norms

Shared norms that affect teacher learning in the workplace can be distinguished in two types: norms regarding student learning and development, and norms regarding teachers' work and relationships with colleagues. Collective focus on, and shared responsibility for student learning are identified as important characteristic of teachers' professional communities in schools. However, according to Lortie (2002) the development of these shared norms is seriously hindered by the absence of a shared conceptual and technical language or a common technical culture in schools. Nevertheless, in-depth analysis of teachers' professional conversations shows that skepticism on the opportunities in schools for developing shared norms over-simplifies and underestimates the potential of teachers' collegial practices for teacher learning (Horn and Little, 2010).

Isolation from colleagues has a negative effect on teacher learning by creating "invisible walls" between teachers and diminishing the valuable role of collaborative activities (Scribner, 1999). This individualism, in which personal experience plays a central role, is regarded as a defining feature of the teaching profession, mainly because of the still dominant way of organizing teaching in regular classrooms and subjects with only one teacher responsible for student learning. In strong traditional cultures teachers collectively defend this individualism and existing conceptions of appropriate curriculum and instruction, even in the face of student failure. On the other hand, strong innovative and professional cultures can contribute to norms of collegiality and experimentation, and to teachers' expansive, informal learning. School leaders play a significant role in establishing an inhibiting or facilitating school culture. Besides they function as relevant sources of information, direction, and feedback.

Learning resources

Four resources for workplace learning are (1) teachers' prior knowledge, (2) curricular and instructional materials, (3) time, and (4) support and rewards. Teacher learning by experience presupposes "conceptual tools" (e.g., pedagogical content knowledge) and explicit knowledge about student learning and how student learning is promoted by instruction and pedagogic methods. These conceptual tools and knowledge are used to identify and interpret relevant experiences while teaching, and they play an essential role in workplace learning. Besides willingness to learn, proactivity, openness, extraversion, and outgoing disposition are relevant individual characteristics for teachers' informal learning.

Varied sources of instructional and curricular materials promote reinvention of instructional practices and lower barriers for experimentation. These practical materials provide opportunities for teachers to experience "what works" in daily classroom practice. However, when implemented superficially, the availability of the same sources bears the risk of counterproductive learning effects, because the anticipated student learning results are not realized in the case of superficial implementation.

Time to learn is an important resource for participation in formal and informal learning activities, as well as for reflection as part of the ongoing work process. The role of time often plays out in combination with related factors, for example school schedules, pressure from management regarding replacement of absent colleagues, etc. Professional development activities should be conducted often and long enough to ensure gains in knowledge, skills, and confidence. Formal availability of time for professional development is necessary but not sufficient as a resource for teachers to learn. Meaningful and time consuming collaboration is often squeezed to margins of teachers' work lives (at Sundays afternoons or weekly evenings in teachers' houses). The organization of time for work in schools is dominantly based on teaching schedules. Opportunities for teacher learning in schools will improve substantially, when time for teacher learning is the starting point of the scheduling processes instead of being the last thing to be added after all other activities have been organized. Without these organizational protections, time

for learning will disappear under the pressure of chains of incidents and daily routines. Therefore, a reassessment of the complex trade-off between time for production and time for learning in schools is needed for teacher learning to be well-organized within schools.

Scheduling is one of many factors influencing teacher learning. Others are full-time employment, which is identified as a facilitating condition for informal learning and continuity of staff, which is important for productive teacher collaboration. Moreover, teachers' informal learning can be restricted or supported by (lack of) freedom to experiment and to decide, (un)supportive relationships and (absence of) social support, quality of communication, and availability of common room for joint work. Identified rewards for informal learning are appreciation and recognition, and career opportunities. Lack of these rewards has counterproductive effects. Formal teacher induction programs are shown to be important for supporting beginning teachers. However, many teachers do not take advantage of this opportunity. Not even half of teachers in 17 OECD countries, on average, report that they participated in induction activities during their first regular teaching job. The school leader plays a decisive role in the availability of support and rewards. As such, the starting point for beginning teachers' use of induction activities is in the priority and facilities that school leaders give to these types of activities.

Conclusions

Before presenting conclusions, it should be noted that the insights presented in this article differ in the solidity of their base of research. In the past 10–15 years a growing body of professional development research has been reported on teacher workplace learning topics. An explicit shared conceptual language is developing. However, empirical research can lack the rigour that is needed for drawing general conclusions. Moreover, the national and regional educational contexts in which professional development and workplace learning are explored can differ considerably in terms of sense of urgency, rationale for professional development, and even in basic assumptions about teacher and school quality, making it sometimes complicated to compare.

The shift in professional development research toward workplace learning topics, insights and methods sheds light on restrictions in the traditional and dominant evaluation approach of professional development research. Firstly, in this evaluation approach the aspired for student learning results often are the only criterion for the measurement of effectiveness of teacher learning. In such evaluations often the focus is on the narrow defined *product* of professional development. The *process* of the learning of the teachers and the changes in their behaviors that lead to improved student learning remain largely unaddressed. Secondly, student learning results that fall outside the scope of originally aspired for results are not included in the evaluation. Thirdly, in the evaluation approach effects often are measured at an ideal-typical scale ranging from 0 (no effect) to +1 (optimal effect). Fourthly, the relationship between intervention and effect is often analyzed based on a one-way linear machine-like model. In such models separate conditions can only have one relationship with effects, either a positive or a negative relationship. In a more dynamic, interactive model that is based on workplace learning insights, effects in student and teacher learning range from -1 (negative or unwanted effects) to +1 (positive effects, including unforeseen effects). In addition, workplace conditions can be both expansive and restrictive. The direction depends on specific designs and inventions of these conditions in schools, and on interpretations and enactments of these conditions by agents in the school.

Summarizing the evidence, two basic perspectives on teachers' workplace learning and work-embedded professional development can be distinguished. Each perspective has particular assumptions on manageability and predictability of learning processes and outcomes in workplace contexts. In one view, teacher workplace learning is a specific activity that can be managed in the direction of formalized school outcomes and desired change. In the other view, teacher learning outcomes range from improvement in school performance to counterproductive effects on teaching and student learning. In this latter view, teacher learning is emergent, and steering opportunities for management to intervene effectively in this process are weak. An unanswered question is if these perspectives are complementary or contradictory toward each other.

Several potentially effective methods for workplace learning have been identified. The effectiveness of these methods should be regarded within the specific workplace contexts in which they are applied.

Six factors have been identified for the integration of teachers' working and learning in schools. Often, these factors are regarded as objective characteristics of the schools in which they emerge. However, it can be assumed that the impact of these factors is located in how they are interpreted and reinvented by teachers, and how teachers make sense of these factors in their daily work. Simply stated, one teacher will experience half a day in the week for professional development as a luxury, while another teacher will regard this time for professional development far too limited. One teacher will experience opportunities for collaboration as enriching, while for another teacher these opportunities will feel contrived. Teachers not only interpret the working conditions in the school differently at a specific moment, in terms of a continuing sense-making process they strongly differ in how they reinvent and reinterpret the content and strategy of workplace learning initiatives and professional development programs.

The relationships between these six factors and teacher learning are complex and paradoxical. These factors play a multidimensional role in teacher workplace learning: they potentially promote opposite directions of learning (reinforcing existing practices or promoting innovations) and they can have opposite side effects on teacher workplace learning (commitment or isolation, empowerment or alienation, etc.).

Traditionally, teachers have been regarded as passive participants in professional development and workplace learning programs, lacking knowledge and skills. Recently, teachers are regarded as active learners, and as agents in co-determining their working and learning goals, contents, processes and outcomes. In this respect, the six factors for the integration of working and

learning in schools also should be regarded as co-determined by the teachers themselves. Starting from insights in sense making in organizations and schools, all aspects of working and learning in schools can be regarded as part of an ongoing process of reinvention and reinterpretation by teachers and other agents in and around the school.

Generally, high-commitment, flexible, and learning-intensive work systems are assumed to promote the integration of learning and working, but scientific management principles still tend to dominate the design of many workplaces (Ellström, 2001). The same phenomenon seems to hold for schools. Rosenholtz (1989) distinguished learning enriched and learning impoverished schools. Rosenholtz and other researchers found far more learning impoverished schools than learning enriched schools (cf. Little, 2012). In reality many schools do not reflect ideal situations like learning organization, PLC, or partner in a learning network. Judging from the current organization of learning resources for teachers in schools, schools are still primarily organized to ensure student learning. Teacher learning remains an ad hoc and often individual affair.

Workplace learning in schools has potential advantages, like facilitated transfer of learning, accommodation of different needs and learning styles, ability to tap collective knowledge, and the natural promotion of collaboration and school renewal. But there are also potential problems in workplace learning. Not all learning “at work” is good, as is shown by negative lessons learned by (student) teachers, reinforcement of existing biases, and refinement of poor practices (Bredeson, 2003). Moving all learning to the workplace might create a too heavy burden on the workplace. Effective teacher learning in schools implies the search for a productive balance and creative swing between adaptive and innovative varieties of individual and collective learning, and on-site and off-site learning.

References

- Billett, S., 2001. Workplace affordances and individual engagement at work. *J. Workplace Learn.* 13, 209–214.
- Borko, H., 2004. Professional development and teacher learning: mapping the terrain. *Educ. Res.* 33, 3–15.
- Bredeson, P., 2003. *Designs for Learning: A New Architecture for Professional Development in Schools*. Corwin Press, Thousand Oaks CA.
- Bryk, A.S., 2015. 2014 AERA distinguished lecture: accelerating how we learn to improve. *Educ. Res.* 44, 467–477.
- Desimone, L., 2009. Improving impact studies teachers' professional development: toward better conceptualizations and measures. *Educ. Res.* 38, 181–199.
- Ellström, P., 2001. Integrating learning and work: problems and prospects. *Hum. Resour. Dev. Q.* 12, 421–435.
- Eteläpelto, A., Vähäsantanen, K., Hökkä, P., Paloniemi, S., 2013. What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65.
- Feldman, M., Pentland, B., 2003. Reconceptualizing organizational routines as a source of flexibility and change. *Adm. Sci. Q.* 48, 94–118.
- Guskey, T.R., 2000. *Evaluating Professional Development*. Corwin Press, Thousand Oaks CA.
- Horn, I., Little, J., 2010. Attending to problems of practice: routines and resources for professional learning in teachers' workplace interactions. *Am. Educ. Res. J.* 47, 181–217.
- Huberman, M., 1995. Professional careers and professional development: some intersections. In: Guskey, T., Huberman, M. (Eds.), *Professional Development in Education, New Paradigms and Practices*. Teachers College Press, New York NY, pp. 193–224.
- Imants, J., Van der Wal, M., 2020. A model of teacher agency in professional development and school reform. *J. Curric. Stud.* 52, 1–14.
- Imants, J., Meijer, P., Blankesteyn, E., 2020. Expansive learning in teacher education's hybrid spaces: the challenges and possibilities in and beyond Learning Studios. *Front. Edu.* 5 (64), 1–13.
- Kennedy, M., 2016. How does professional development improve teaching? *Rev. Educ. Res.* 86, 1–36. <https://doi.org/10.3102/0034654315626800>.
- Kyndt, E., Gijbels, D., Grosemans, I., Donche, V., 2016. Teachers' everyday professional development: mapping informal learning activities, antecedents, and learning outcomes. *Rev. Edu. Res.* 86, 1111–1150.
- Little, J.W., 2012. Professional community and professional development in the learning-centered school. In: Kooy, M., Van Veen, K. (Eds.), *Teacher Learning that Matters: International Perspectives*. Routledge, New York NY, pp. 22–46.
- Lortie, D., 2002. *Schoolteacher*, second ed. Chicago University Press, Chicago IL.
- Luttenberg, J., Imants, J., Van Veen, K., 2013. Reform as an ongoing positioning process: the positioning of a teacher in the context of reform. *Teach. Teach. Theor. Pract.* 19, 293–310.
- OECD, 2013. TALIS 2013, an international perspective on teaching and learning. https://read.oecd-ilibrary.org/education/talis-2013-results_9789264196261-en#page1.
- Rosenholtz, S., 1989. *Teachers' Workplace: The Social Organization of Schools*. Longman, New York NY.
- Schildkamp, K., Lai, M., Earl, L. (Eds.), 2013. *Data-based Decision Making in Education, Challenges and Opportunities*. Springer, Dordrecht.
- Scribner, J., 1999. Professional development: untangling the influence of work context on teacher learning. *Educ. Adm. Q.* 35, 238–266.
- Smith, C., Gillespie, M., 2007. Research on professional development and teacher change: implications for adult basic education. *Rev. Adult Learn. Lit.* 7, 205–244.
- Spillane, J., Parise, L., Sherer, J., 2011. Organizational routines as coupling mechanisms: policy, school administration, and the technical core. *Am. Educ. Res. J.* 48, 586–619.
- Van Driel, J., Meirink, J., Van Veen, K., Zwart, R., 2012. Current trends and missing links in studies on teacher professional development in science education: a review of design features and quality of research. *Stud. Sci. Educ.* 48, 129–160.
- Vangrieken, K., Dochy, K., Raes, E., Kyndt, E., 2015. Teacher collaboration: a systematic review. *Educ. Res. Rev.* 15, 17–40.
- Weick, K., 1995. *Sensemaking in Organizations*. Sage, Thousand Oaks CA.
- Wenger, E., 1998. *Communities of Practice, Learning, Meaning, and Identity*. Cambridge University Press, Cambridge.
- Wilson, S., Berne, J., 1999. Teacher learning and the acquisition of professional knowledge: an examination of research on contemporary professional development. In: Iran-Nejad, A., Pearson, P. (Eds.), *Review of Research in Education*, vol. 24, pp. 173–209.

Lesson study: a practice for teachers' training and professional development at any educational level[☆]

Gabriel Hervas^a and José Luis Medina^b, ^aCenter for Research and Development of Higher Education, The University of Tokyo, Tokyo, Japan; and ^bDepartment of Didactics and Educational Organization, University of Barcelona, Barcelona, Spain

© 2023 Elsevier Ltd. All rights reserved.

Introduction	570
Background on the history and internationalization of lesson study	570
Lesson study: practice and benefits	571
Key features of LS from the epistemology of complexity	571
LS and its phases: emergence in LS	572
The RL and the ecology of action	573
Learning through cooperation and joint reciprocal reflection	573
Some results regarding the practice of LS	574
Conclusion	578
Acknowledgments	578
References	579

Introduction

This article, building on our research and practical experience, delves into lesson study (LS), a teachers' professional development practice originating in Japan. The different sections address: (a) the history and internationalization of LS; (b) its practice and benefits for the training and development of teachers at any educational level; (c) a novel theoretical approach to LS through the lens of complexity; and (d) a summary of our research and results on the practical use of LS in higher education (HE).

Background on the history and internationalization of lesson study

LS, the English translation of the Japanese concept “jugyou kenkyuu” (授業研究), is a practice whose roots stretch back to the Japanese Meiji era (1868–1912). This century of history has been addressed in different studies (i.e., [Ishii, 2016](#); [Makinae, 2019](#); [Shimizu and Chino, 2015](#)) and, less explicitly, Sato (1991) has linked it to the arrival in Japan of ideas based on the Herbartians' formal five steps. Nonetheless, differently to the previous studies, [Matoba \(2017\)](#) has briefly addressed the history of LS suggesting a more contemporary origin.

LS began to draw attention outside of Japan with Stigler and Hiebert's work (1999), which identified best practices from around the globe for improving education in the classroom. From the same year, the unpublished doctoral dissertation of [Yoshida \(1999\)](#) has also received credit as contributing to the internationalization of LS. Its popularization came after a period of growing interest in Japanese education during the 1980s and 1990s (i.e., [Beauchamp, 1991](#); [Cummings, 1980](#); [Horio, 1988](#); [Lewis, 1995](#); [Rohlen, 1983](#); [Rohlen and Björk, 1998](#); [Rohlen and LeTendre, 1996](#); [Shimahara and Sakai, 1995](#); [Stevenson, 1991](#); [Stigler et al., 1987](#)). Since that time, seeing a link between the good results of Japanese students in international tests and the training and teaching culture in the country ([Ingersoll, 2004](#)), teachers and researchers around the world have embraced and studied LS ([Lewis and Lee, 2017](#)), including in educational contexts as varied as Hong Kong ([Hargreaves and O'Connor, 2018](#)), Iran ([Moghaddam et al., 2015](#)), Uganda ([Fujii, 2014](#)), the UK ([Dudley, 2013](#)) and the US ([Akiba and Wilkinson, 2016](#)). These last two countries are where we find the most examples of LS being put into practice and of related studies.

Regarding the internationalization of LS, we wish to make a clarification for historical and scientific rigor. The well-deserved acknowledgment of [Stigler and Hiebert's \(1999\)](#) work and [Yoshida's \(1999\)](#) doctoral dissertation has sometimes come through statements that may have generated confusion. [Pang and Marton \(2003, p. 175\)](#) claimed that Stigler and Hiebert were responsible for the introduction “to the West of the Japanese approach of ‘lesson study’”, [Doig and Groves \(2011, p. 77\)](#) described Stigler and Hiebert's work as the means through which LS “was released outside Japan”, and [Crockett \(2002, p. 610\)](#) went even further and wrote that around 1997 “there were virtually no English language materials available on Japanese professional development”. These authors did not state that Stigler, Hiebert and Yoshida were the first to inform about LS outside of Japan, but their words—common in the literature about LS—could have easily produced that impression.

[☆] *Change History:* February 2021. This article includes an updated version of [Hervas and Medina \(2020a,b\)](#) in section 4 and the conclusions. The authors updated and restructured the original paper to meaningfully include it here.

We have already referred to an extensive selection of literature exploring the work and training of Japanese teachers during the 1980s and 1990s. In addition, our research (Hervas, 2022) has shown that also during the early 1990s we find studies clearly describing LS and other oblique or fragmentary references even earlier (e.g., Becker, 1989; Shimahara, 1979), many of them signed by Japanese originating authors; they all deserve credit in the history of LS overseas.

Lesson study: practice and benefits

LS is a central component of the professional development and in-service training (“kounai kenshū”) of Japanese teachers (Fernandez and Yoshida, 2004; Stigler and Hiebert, 1999), and it is used widely in Japan for teachers’ training and professional development at different institutional and administrative levels (Lewis, 2013).

LS consists of a cyclical, collaborative process of inquiry and reflection, carried out by a group of teachers who design a lesson plan named the “research lesson” (RL), teach it and observe it being taught, and analyze both the lesson itself and the process of teaching it, with the goal of improving lessons and teaching practices and, in this way, enriching students’ learning experiences (Lewis and Hurd, 2011; Fujii, 2016; Stepanek et al., 2007; Stigler and Hiebert, 1999; Takahashi et al., 2005; Takahashi and McDougal, 2016). In detail, LS involves the following phases or steps:

- Set goals for the RL according to the subject, topic and students’ knowledge gaps or “learning challenges” (Cajkler and Wood, 2016). The goals can differ depending on where LS is implemented (Lewis, 2015) and where the focus is placed (students’ learning, teachers’ training, curriculum development, etc.). In Japanese schools, for example, lesson goals often arise from transferring an institution-wide objective to the context of a particular subject.
- Design the RL (structure, methodology, activities, materials, etc.) and a proposal of inquiry (research topic, instruments to gather data, aspects to analyze, etc.) that teachers will conduct so that they can later analyze the RL. In the design of the RL, teachers also try to anticipate students’ possible approaches to the tasks in order to prepare a response.
- Implement the RL. In this phase, a teacher (usually one) teaches the lesson while the rest of the group observes and collects data according to the proposal of inquiry. On many occasions, each observer pays attention to a different aspect.
- Reflect collaboratively in a post-lesson discussion (Takahashi and McDougal, 2016) to analyze the RL based on what teachers have experienced and observed so that it can be improved and implemented in a different class—something that does not always happen and disseminated to the educational community through the public report of the lesson and what has been learned through the process.

This process brought Cheng (2019) to conceptualize LS as a knowledge management approach and is in line with other models for continuous improvement and reflective and inquiry-based educational practices such as “Plan-Do-Check-/Study-Act” (PDCA/PDSA; Langley et al., 2009), Analyze-Design-Develop-Implement-Evaluate (ADDIE; Branch, 2009) and action-research (McIntyre, 2008). Overall, LS has been considered useful for teachers’ induction and acculturation (Howe, 2006) and continues to show promising results for teaching enhancement (Hiebert and Stigler, 2017) and efficacy (Chong and Kong, 2012), for the development of teachers’ pedagogical content knowledge (PCK; Coenders and Verhoef, 2019), and for teachers’ active (Vrikki et al., 2017) and transformative learning (Wong, 2018). Also, it has been shown to offer possibilities for improvement science within education (Lewis, 2015) and potential for curriculum reform (Lewis and Takahashi, 2013). Several of these positive outcomes are related and similar to those reported in relation to the peer observation of teaching (Bell and Mladenovic, 2015; Steinert, 2010; Steinert et al., 2016).

These positive outcomes have contributed to the international growth of LS and to its transfer to different educational levels, including the training and professional development of university teachers through faculty or academic development programs. In this last scenario, we conducted a systematic review of the literature (Hervas, 2021) that evinced an increasing number of studies addressing LS as a faculty development practice, most of them from the US and concentrated in disciplines related to language and math. Although the number of studies is still limited and their results are mostly based on isolated experiences with few participants, they are also encouraging as they reveal beneficial outcomes of LS in the design of the lessons, the participants’ pedagogical knowledge, and the participants’ approach to teaching, such as those that we find among schoolteachers and preservice teachers. Nevertheless, our review also showed mixed results in relation to the participants’ reflection and collaboration, and less positive outcomes related to the organization of LS. Overall, our review made clear the need for further research to substantiate the practice of LS for faculty development in HE, a context in which the professional practice and idiosyncrasies differ from those of schools and schoolteachers.

Key features of LS from the epistemology of complexity

Despite its benefits, previous studies also reveal the need to improve our theoretical foundation to understand the purpose of LS, its features, and its use (Lewis et al., 2006; Murata, 2011; Rock and Wilson, 2005). This increasing theoretical foundation would help teachers who implement LS avoid the difficulties and misconceptions revealed by various authors (Chokshi and Fernandez, 2004; Diaz et al., 2005; Fujii, 2014; Postholm, 2019; Takahashi and McDougal, 2016; Yoshida, 2012). These situations often arise when

teachers uncritically adopt a process that, even in Japan, has not always led to reflective and collaborative teaching (Sarkar Arani et al., 2010).

In this section, we synthesize one of our articles (Hervas and Medina, 2020a), and we propose to expand the theoretical understanding of LS by observing it from the perspective of the epistemology of complexity (also known as “complex thinking” or “systemic thinking”). We adopt this viewpoint under the assumption that it is useful—though not necessarily required—for thinking about the cognitively complex teaching-learning processes (McAlpine et al., 2006) that, additionally, are executed in an educational reality that is itself unattainable and emerging.

From this unexplored angle, we make explicit tacit assumptions of LS that teachers may overlook when implementing it. When teachers do not keep in mind the underlying complexity involved in LS, they can become concerned or discouraged when their lesson plans do not develop as expected in the classroom, or when their peers' opinions sound unflattering. In contrast, recognizing the complex cognitive processes inherent to LS can help teachers carry out this process, which requires open-mindedness and the ability to tolerate unexpected outcomes. By applying the lens of complexity, teachers can capitalize on opportunities for reflection and avoid feeling overwhelmed.

Beyond its potential value to think and put into practice LS in any educational context, we propose that this perspective will be particularly useful in teaching contexts lacking any of the following habits, which are fundamental initial conditions for carrying out LS: (a) a culture of institutional support for teachers' inquiry practices (Takahashi and McDougal, 2016), (b) a routine of cooperation around creating lesson plans (Fujii, 2016), and (c) a systematic engagement in research-based professional learning (Akiba and Wilkinson, 2016).

We aim to explore LS using the framework of complexity, which would in turn allow teachers to implement it, while avoiding misunderstandings and maintaining their commitment to the process. Hence, in the following sub-sections, we interpret three key components and facets of LS from the perspective of the epistemology of complexity: its phases or steps, its product (the RL), and collaboration.

LS and its phases: emergence in LS

LS consists of a cyclical process that involves an ongoing feedback loop through which the output of the analysis of the RL is used as input for a new version of it. The phases of LS entail the creation of an original RL (and the forms of pedagogical and content-based knowledge connected to it), its deconstruction, and its reconstruction. We could imagine LS as a spiral path that, if resources permitted, would be infinite; teachers could always return to observe the RL, reflect on it, reformulate it and teach it again, cooperating with its contextualized evolution.

We argue that, to implement LS successfully, participants must tacitly acknowledge the unattainability of reality, recognizing the existence of weak emergence—phenomena and properties that, in organized systems, arise unexpectedly (Chalmers, 2006)—and of classroom dynamics beyond their control. These elements will affect their work and the design of the RL, especially because of two of the features of emergence described by Goldstein (2016): radical novelty and “ostensiveness”. Although Goldstein considers the notion of emergence inadequate and, under a different logic, talks about “self-transcending construction”, this author offers a clear definition for both features when referring to radical novelty as the unpredictability of phenomena from the initial conditions and perspective and to “ostensiveness” as the impossibility to foresee these phenomena until they happen.

Recognizing these features is key to implementing LS, because the continuity of its structural stability as a process is tied to the emergence of unpredictable events that take place and are observed in the classroom. Still, this recognition also runs the risk of making LS seem too challenging and thus diminishing teachers' commitment. We suggest that we can overcome the perception of difficulty by adopting other elements from the epistemology of complexity and situated cognition, a way of thinking in which knowledge arises as a means for coordinating an activity from within the activity itself, articulated with the social context in a specialized population niche and transformed through processes of assimilation and interpretation (Robbins and Aydede, 2009).

Classrooms are complex systems occupied by human beings that, as Morin (2007) suggests, could be understood as non-trivial machines, analytically unpredictable and indeterminable (von Foerster, 1984). These features though, do not mean that it is ineffective to try to anticipate behaviors in complex contexts; to the contrary, attuning oneself to a complex system can occur by accident, but, at a cognitive level, it can also occur through conscious reflection (Robbins and Aydede, 2009). LS practitioners should not confuse unexpectedness with nondeducibility, as phenomena can be both unexpected and, yet, deducible in principle (Chalmers, 2006). This idea reinforces the significance of two of LS's phases that involve aligning the RL with the behavior of the class as a system: (1) planning, as in Schön's (1987) “reflection for action” and van Manen's (1991) “anticipatory reflection”, and (2) revision, as in Schön's (1987) “reflection on action” and van Manen's (1991) “recollective reflection”.

However, if students are non-trivial machines, how can teachers who carry out LS “pre-decode” their actions and possible answers during the teaching of the lesson? Teaching actions can be understood as strategies (Morin, 1999/2010), adaptable rather than immutable decisions. Also, teachers—especially experienced ones—can make use of their PCK, which offers them ideas on common answers and difficulties that students face in relation to a particular content or activity. The anticipation of students' responses is considered a core practice for teaching (Grossman et al., 2009) and it is tightly connected to the practice of “press-onance”—a process that enables a contextualized anticipation of future states, based on the ability to experience situations vicariously (Robbins and Aydede, 2009)—and, in class, to “pedagogic resonance”—relational awareness of different elements that occur when interacting with students that resonates in the mind of the teachers, and which generates a pedagogic action (Trigwell and Shale, 2004). It is not that students' behaviors (individual or as a group) are entirely predictable and analytically determinable;

rather, teachers draw on professional knowledge that grants them a general idea of what to expect when presenting students with a given input (topic, activity, etc.). This awareness helps them to design the RL, foreseeing situations that might emerge, and increases their confidence in an RL that, in the end, is based on students' hypothetical performances.

During LS, teachers need not anticipate all possible outcomes that an RL could have, nor how every individual student could react to it. Despite an infinitude of situations that could arise, only some of them are highly probable and even fewer will become a reality. Even natural systems balance order and chaos (Crutchfield, 2012) and in them too we find patterns, self-organization, and specific points to which they seem to propel themselves (Boeing, 2016). Reflecting on the design of the RL helps teachers locate—at a cognitive level—critical “pressure points” from which small changes can generate a great impact in terms of students' learning. Pressure points related to content could be, for example, threshold concepts (Walker, 2013), concepts that, when understood, transform how students think about a topic.

Unpredictability and structured disorder, rather than being obstacles to LS, form the basic premise that allows LS to function. Teachers tacitly assume that in the dynamic of the class system, non-linear interactions will occur, in turn provoking unexpected behaviors that may be amplified by a self-reinforcing feedback loop. Without unpredictability and disorder—if everything in the classroom were foreseeable regularity—the processes of observation and reflection in LS would become unnecessary. Conversely, accepting that things will surely not play out exactly as we expected allows teachers to analyze what happens and re-design the RL accordingly. In short, we can only carry out LS successfully if we accept the emergent properties of classroom experience and understand that the RL is, at best, a method of experimentation.

The RL and the ecology of action

In LS, teachers witness and participate first-hand in the ecology of human actions; once these are set in motion, they are taken over by the environment, weakening our control over them and limiting our ability to understand them (Heath and Luff, 2013). From the moment the RL enters the classroom, we can imagine it as a raft floating downstream. The teacher takes actions to steer it and tries to navigate its way to port, making continuous adjustments as unforeseen currents alter the raft's path. After the RL has been taught and gets to port, the teachers reflect on and discuss the necessary modifications to make it more likely to successfully navigate the waters they have experienced. Thanks to this feedback, the RL regains a temporary state of equilibrium, ready to be launched again. Still, the waters it will navigate next time will be different and might offer new challenges. For this reason, even if the RL sometimes recovers stability, as long as it continues to be taught in new contexts with new students, it will continue being taken over by the environment, facing disequilibrium and needing to be adjusted.

As seen, the RL itself is a stable system, but it is mobile and subject to an irreversible flow, without which it cannot operate. We find in this process an example of “endo-exo-causality” (Morin, 2005a,b, p. 37), a feedback loop of causalities between the RL and the phenomena that influence it: the RL is designed for a specific context, but the RL itself brings about unexpected events, which in turn affect it and make it possible to continue the process of LS. Students, their relationship to what teachers prepare, and the circumstances all affect the RL, as teachers observe and analyze these elements in creating a modified version. This process is an example of how, through interaction with the environment, both a system overall and the product of the system experience self-organization.

When teachers remain open to unpredictability, they participate in uncertainty, incorporating it into the logic of the RL using rationality. We follow Morin (2005a,b) in understanding rationality as a quality that draws on dialog between our logical structures and reality, without pretending—unlike rationalization—that our logical structures encompass reality. In so doing, teachers consider not only the present choices, but also the decision tree they lead to, as Schön (1983) pointed out for reflection in action. Thus, in trying to anticipate contingencies that might appear when the RL is implemented, teachers construct design networks of great complexity, as if drawing what in physics (Nolte, 2010) would be a map of trajectories intersecting the phase space.

Acknowledging that not all paths and problematic situations can be identified a priori, LS participants do not envision the RL as determining a linear and immutable course of action. The RL is, on the contrary, a space of possibility; it is a product—“biodegradable” in the action and the surrounding context—in which LS practitioners are invited to embrace uncertainty as something productive (Mintz, 2016) and act as an attentive audience, ready to learn from any possible divergence from the expected.

Learning through cooperation and joint reciprocal reflection

LS is a collaborative process, but its collaborative practices and structures are sometimes left behind when LS is transferred to contexts outside Japan, revealing misconceptions in its practice and the need for adaptations (Takahashi and McDougal, 2016). When carried out in its full form, LS involves several teachers who are part of an epistemic community (sharing the same subject or a similar disciplinary field) and are subject both to a general normative framework (legal, institutional, curricular, etc.) and to their own system of (often implicit) rules, making of their collaborative process a complex system that evolves (Yuan et al., 2018).

Cooperation though, is highly shaped by context and culture. In Japan, professionals are often involved in “kaizen”—continuous improvement (Sarkar Arani et al., 2010)—a work philosophy that involves the willingness to unlearn thinking schemes and construct new ones with the help of colleagues. For this reason, we find in Japanese schools a culture of institutional support for teachers' inquiry practices (Takahashi and McDougal, 2016) and a tradition of engaging in research-based professional learning (Akiba and Wilkinson, 2016). In Japanese, we also find the concept of “hansei” to describe a practice of continual critique of our prior actions with the goal of learning from mistakes and improving our behavior in the future. In this sense, the LS group is shaped

as a setting that favors self-contemplation and openness to joint self-inspection as a path toward reconsidering what was previously known. Together, teachers co-construct knowledge through dialog and observation, under the assumption that there is no single correct way to create and conduct an RL; in this manner, LS can be described as an intimate practice (Lewis and Hurd, 2011) that requires teachers to reveal themselves professionally, contributing with their knowledge, practical experience, judgments and educational ethics and values. Overall, teachers collaboratively apply their reasoning, combining their different forms of knowledge to reach a consistent group vision of what to teach and to whom.

The RL is developed from this point of overall—though not necessarily complete—consensus; teachers must be willing to accept critique (of the work more than of the person who performs it) and to pursue self-evaluation. This process requires that participants develop a type of communication that allows them to place each other in the world cognitively, in order to coordinate their joint efforts (Robbins and Aydede, 2009). Through this joint reflection, interpersonal knowledge is engendered from the shared experience of a socially situated process (van Dijk, 2010), which has different representations in the mental models of the individual participants. This is yet another case in which the part (individual knowledge) and the whole (group knowledge) mutually influence each other, generating interactive learning paths for both the individual and the group.

It is useful to think of the group of teachers conducting LS as a system in the form of a learning community. This system is a perceived whole whose parts remain united, because, while continuously affecting each other, they operate for a common purpose. Despite resistance to change, by engaging in dialog and regulating the relations and interactions that happen in and arise from the system, the community evolves in codependence with the context. The group and its work are flexible and sensitive to the development of the RL. These traits make it possible for the group to face needs that emerge beyond what was planned and tolerate circumstances related to mistakes, “liminal spaces”—areas or states of transition and ambiguity where meanings and learners’ subjectivity are transformed (Land et al., 2014)—and “post-liminal variation”—variation in the point and state of exit out from the liminal space into a novel conceptual state (Calduch and Rattray, 2021). Despite the complexity of the definitions coming from the study of the threshold concepts, there are typical facets of any learning situation.

We should understand LS as a practice that, due to its origin, both requires and promotes a culture of cooperation among groups of teachers and at their institutions more broadly. In Japanese schools, for instance, LS groups specify broad institutional goals within their subjects through the RL; this specification involves connecting systems at different levels, contributing to develop a more organic perspective of the organization by recognizing teachers as driving forces for educational change and improvement. LS groups are then spaces guided by cooperation rather than competition, where the principles of solidarity and flexibility are applied. Practitioners strive to understand the inevitable diversity and contradictions that arise as signs of the group’s vitality. And this very diversity provides teachers with access to other viewpoints, allowing them to avoid self-deception and to use contrast to analyze their own standpoints, knowledge, and methods.

Some results regarding the practice of LS

To close this article, we present a summary of the most relevant results of our recent research on LS to complement, from an inquiry-based perspective, our readers’ understanding of LS. Recently, we analyzed the integration of an instrument called “Content Representations” (CoRes) within LS (Hervas and Medina, 2022). CoRes is a tool developed to uncover how teachers organize and conceptualize teaching and their PCK (Loughran et al., 2001, 2004). It takes the form of a table with different prompts related to the teaching of an idea and, for this reason, it is useful for lesson design (Loughran et al., 2008) and for teachers’ training, as it encourages teachers to reflect on their PCK and to reconceptualize their teaching practice (Bertram and Loughran, 2012).

Despite these positive outcomes, prior to our research, the integration of CoRes into LS had only been explored once (Juhler, 2016, 2017). Our research contributed to filling in the gap within the context of HE and revealed that the CoRes, even when approached superficially (in special, experienced teachers), enhanced the participants’ reflection and contributed to organizing ideas considering the design of the lessons in more detail and addressing newer aspects. However, our study also showed that the CoRes was perceived as time-consuming tool. For teachers’ training and professional development, we encourage the integration of CoRes into LS when we want to explore the teaching of a difficult lesson or of new contents introduced in the curriculum.

For these situations, Fig. 1 presents the instrument that we called “Content and learning outcomes representations for design and reflection” (CLOREDRE). CLOREDRE is a substantial evolution and update of the CoRes, created with the aim of facilitating deeper self-reflection and a more exhaustive process of lesson design. With that aim, CLOREDRE: (a) incorporates new prompts about teaching and learning formats, resources, learning environments, collaboration, inclusive teaching and learning, feedback, the integration of contents, and the inquiry about the lessons; (b) incorporates learning outcomes, with explicit demands for connecting them with contents and for connecting lesson and course learning outcomes; (c) departs from the use of concepts or ideas to address contents of different nature (knowledge, skills, attitudes, and values); (d) broadens the proposal of Nilsson and Loughran (2012) to include a section for self-reflection before and after the lesson; and (e) includes a new full section for making explicit the different ideas in the design of the structure of a lesson.

We have also analyzed the use of videos during LS and the potential and challenges that they pose (Hervas et al., 2020). We did this because LS might involve the recording of the teaching of the RL (to be used in the later discussion), but there is little research exploring the impact of the use of video (e.g., Lim et al., 2011). In this regard, some authors have urged caution when using videos in teachers’ training (Gaudin and Chaliès, 2015), as its effectiveness is often assumed rather than documented (Wang and Hartley, 2003).

Course title: ...	Goal(s) of the lesson (LG) State the main aim(s) of your lesson: LG1. ... LG2. ...			Connection with course goal(s) (CG) Make explicit the connection between the lesson goal(s) and the course goal(s): E.g.: LG1 relates to CG3 in ...			
Title of the lesson: ...	Intended learning outcome(s) of the lesson (LLO) Specify the learning outcome(s) that you intend your students to have achieved by the end of the lesson (they can also be expressed in terms of a competence). By the end of the lesson, students will be able to (...): LLO1. ... LLO2. ...			Connection with course learning outcome(s) (CLO) Make explicit the connection between the lesson and the course learning outcome(s): E.g.: LLO1 relates to CLO4 in ...			
(Potential) Num. of students: ...	CONTENT State what content your students should learn during the lesson to achieve the intended LLO(s). Content of any nature: knowledge, skills, attitudes, and values. Each LLO might involve different contents. Also, each content can be shared by different LLO(s). Add/eliminate columns as needed.			SELF-REFLECTION Regarding the formulated questions...			
Format (e.g.: online, hybrid, in person, hy-flex, etc.): ...				Before the lesson		After the lesson	
				A. How meaningful is this question for me when I reflect on teaching and learning (T&L)?		A. How meaningful is this question for me when I reflect on teaching and learning (T&L)?	
				B. How confident do I feel when I respond it? Rate from 1 (low) to 10 (high)		B. How confident do I feel when I respond it? Rate from 1 (low) to 10 (high)	
Where & when (e.g.: classroom number, link, dates, etc.): ...	CONTENT 1 State the content in less than one sentence ...	CONTENT 2 State the content in less than one sentence ...	CONTENT 3 State the content in less than one sentence ...				
Other relevant information: ...	LLO(s) State the LLO(s) that involve this content (e.g., LLO2, LLO3) ...	LLO(s) State the LLO(s) that involve this content (e.g., LLO2, LLO3) ...	LLO(s) State the LLO(s) that involve this content (e.g., LLO2, LLO3) ...			C. What did I learn? Describe any learning in relation to this aspect	
				A	B	A	B
1. What do you intend the students to learn about this content? Describe what your students should learn about this content.	...	...	...	...	...	...	...
2. Why is it important for your students to learn this? Justify the relevance of the content from any perspective (professional, scientific, curricular, inclusive, etc.).	...	...	...	...	...	...	...
3. What else do you know about this content that you do not intend students to learn yet? Briefly, refer other aspects of the content that would not be addressed in the lesson.	...	...	...	...	...	...	...
4. What are the challenges of T&L this content? State difficulties and limitations that you anticipate.	...	...	...	...	...	...	...
5. What do you know about your students' that might influence T&L this content? Describe what you know about your students and previous cohorts that might affect T&L this content (e.g.: previous knowledge, skills, attitudes, and values, learning in other courses, interests, schedules, past cohorts' 'relation' with the content, etc.	...	...	...	...	...	...	...
6. What T&L methods, strategies, techniques, or activities will you use to promote the learning of this content? Name them (e.g., flipped classroom, project-based learning, simulation, jigsaw, poster tour, debate, etc.) and, briefly, describe how they will be developed. Consider yours and the students' perspective.	...	...	...	...	...	...	...

Fig. 1 Content and learning outcomes representations for design and reflection (CLOREDRE).

7. How will you ascertain students' understanding of the content before, during, or after the lesson? <i>Briefly, describe any diagnostic, formative, and summative assessment, including low-stakes classroom assessment techniques (e.g., one-minute paper, questioning, quizzes, concept maps, etc.). State if peer- or self-assessment will take place and, when necessary, refer how you (or others) will deliver feedback & grades.</i>	...	...	...	...	...	...	...
						C	...
8. Which are the main materials and resources (also, technological) that you will use to promote the learning of this content? <i>Name them and, briefly, describe their use (link them, if necessary). E.g.: readings, tablets, etc.</i>	...	...	...	...	...	...	...
						C	...
9. What do you know about the learning environment that might influence T&L this content? <i>Describe what you know about the learning environment and the available resources that might affect T&L this content (e.g., classroom distribution, accessibility issues, updated software, etc.).</i>	...	...	...	...	...	...	...
						C	...
10. Are there any other people/agents that you can involve in teaching this content (hence, in designing the lesson)? <i>E.g.: peers, alumni, professionals, students, etc.</i>	...	...	...	...	...	...	...
						C	...
11. Are there any other factors that might influence T&L this content? <i>Refer any other factors</i>	...	...	...	...	...	...	...
						C	...
12. What elements will you incorporate to promote T&L this content with an inclusive perspective? <i>For each of the questions responded, if pertinent, describe any relevant element in this regard. E.g.: inclusive LLO/content, diversity of tasks, dual coding, activities that can be solved from multiple perspectives, inclusive references and examples, universal design of materials, granting accessibility to the learning environment, etc.).</i>	...	...	...	...	...	...	...
						C	...

Fig. 1 (continued).

Our study showed that the use of videos during LS has potential to improve the RL and stimulate reflection, as it: (a) grants perspective, (b) facilitates noticing, recalling, and analyzing pedagogical details, and (c) enhances communication by reducing the worries related to expressing opinions. However, we also revealed drawbacks, including the possible effect of cameras on behavior and difficulties in terms of logistics and the duration of the LS process that affected the chances of sustaining LS in the

13. When teaching these contents, how will you connect them? <i>Describe how these contents will be integrated so that the students achieve the LLO(s).</i>	...	...	...	...	...	...	...	
				C	...			
14. In case of need, what modifications would be required to adjust T&L these contents to a different format? <i>For each of the questions responded, if pertinent, describe any relevant adjustment that you would need to make. E.g.: you can refer the necessary changes to teach and learn these contents on-line, if the lesson is designed for an in-person format.</i>	...	...	...	...	...	...	...	
				C	...			
15. How are you planning to assess this lesson to improve it in the future? <i>Describe, if pertinent, how are you planning to enquire about your lesson (aims, how and what data you will collect, any dissemination plans, etc.).</i>	...	...	...	...	...	...	...	
				C	...			
Following this, make explicit the structure of the lesson fulfilling the following table <i>Add or eliminate rows as needed.</i> <i>Also, if a lesson involves more than a class, add rows as needed</i>								
When	T&L event <i>Describe, overall, what you designed to happen and how it would happen</i>	Duration <i>Minutes that the event would last from the student's perspective</i>	Teacher(s)'s actions <i>Describe what you will need to do before, during, and after the T&L event (e.g., prepare materials, organize groupwork, design assessment, etc.)</i>	Students' actions <i>Describe what your students will need to do during the T&L event</i>	Content <i>State the content addressed during the event (you can write the number used before and state previous and upcoming contents)</i>	Assessment & grading <i>State if/how the event will be assessed and graded and if/how it will incorporate feedback</i>	Resources <i>Name the resources & materials to use during the event</i>	Others
Before the class								
During the class								
After the class								

Fig. 1 (continued).

long term. For teachers' training and professional development, we encourage the use of short video segments to address specific aspects regarding the learning outcomes that we wish to promote.

Finally, we have also explored the topics that teachers discuss during LS, the learning opportunities that these conversations grant them (Hervas and Medina, 2021), and what these conversations reveal about how university teachers talk about their students

(Hervas and Medina, 2020b). Our results show that, more than any other topic, the participants discussed how to change the RL (its structure and the strategies used) in connection to promoting their students' engagement; concurrently, they also granted themselves opportunities to discuss the design process that they followed and how to improve it. In contrast, conversation about the teaching of the RL was less common and was of a descriptive nature, while conversation about their students showed room to become less teacher centered by focusing more on students' learning and thinking processes, and less on their (lack of) engagement and on general perceptions about HE students. For teachers' training and professional development, we encourage the use of conversation analysis to promote self- and peer-reflection and make explicit tacit ideas and beliefs.

Conclusion

In this article, we have attempted to offer an overall perspective of LS, building on our experience practicing it and our research on the topic. In the first section, we have described the history of LS and of its internationalization. We have recognized the research from the 1980s and the early 1990s that described the practice of LS and that, until present, has received little attention from researchers in our field.

In the second section, we have described LS, its different phases, and the benefits of it that the literature reveals. In addition, we have reported the outcomes and challenges of LS for faculty development in HE, an educational context in which LS is still making its first steps in terms of the body of evidence available.

In the third section, we have presented complex thinking as a cognitive attitude that can allow teachers to make sense of and carry out the process of LS. The goal of our theoretical analysis has been to make LS more understandable, therefore helping to confront the challenges that have arisen when teachers have tried to apply it outside of Japan.

Along the same lines, we have described how the LS cycle revolves around the RL, a product tailored to a specific context and to specific groups of students, assuming that their features are unique. This approach avoids oversimplification by paying attention to the local and particular, while simultaneously recognizing LS' embedding in a global context. LS' continuity depends on change, observation, and critical reflection and, through them, the original RL progresses. In LS, teachers forget about mechanical certainties and accept approximate knowledge and partial comprehension as starting points. From there, discoveries and learning happen, thanks to inaccurate predictions.

LS can contribute immensely to teachers' professional development, as it asks of them to be willing to shape their practices by taking advantage of emergence, boosting "pressonance" and operationalizing their own wisdom and knowledge. This process faces the complexity inherent in classroom systems by encouraging teachers to reflect on what they have experienced, embracing their own and others' reflective critique to devise strategies for improving the RL.

In this section, we have also suggested that LS is based on more an ecological than a mechanistic conception of reality. In this sense, LS is a process for constantly seeking an RL befitting a context, goal, and topic. The pursuit of an unreachable ideal—the perfect RL—is also the only way to bring teachers closer to achieving an envisioned archetypical lesson for a specific moment and situation. LS proceeds, and the RL evolves, as teachers engage in collaboration and tolerant dialog to problematize previous ways of doing (a process that benefits from the original Japanese cultural features), as they make use of strategy and situated thinking to attend to unforeseen situations and compare alternatives, and as they remain aware that they can only have ephemeral control over the actions that they carry out and the product that they create.

Deleuze said that "(...) there is no other method for finding other than a long preparation" (Deleuze and Parnet, 1977/2007, p. 7). We propose that—viewed from the perspective of the epistemology of complexity—LS is in fact a form of long preparation in which teachers move forward by developing tentative answers that lead to subtler questions.

Finally, in the last section of this article we have described the results of our previous research regarding the practice of LS, hoping to help teachers make informed decisions when practicing LS. First, we have shown the potential of integrating the CoRes within LS and the concerns that their combined use raises. Second, in a similar vein, we have described the implications and challenges of using videos during the practice of LS. Lastly, we have described what LS participants talk about and how they do so.

We hope that this review of LS, with theoretical and practical suggestions for teachers' training, will help teachers, future teachers, teacher educators and faculty developers when thinking about and implementing LS.

Acknowledgments

This chapter is derived in part (mainly section four and conclusions) from an article published in *Teachers and Teaching: Theory and Practice*, 26(1), 118–128 on March 26th, 2020, available online at <https://doi.org/10.1080/13540602.2020.1745174> and <https://www.tandfonline.com/eprint/WEFXNJ6W9BB7YUJB4MDI/full?target=10.1080/13540602.2020.1745174>. Susan Frekko provided English-language support for this chapter.

This article and our other studies here described were partly supported by the Spanish Ministry of Economy and Competitiveness under Grant [EDU2015-63712-P-BES-2016-076824].

In addition, we would like to express our utmost gratitude to our families and colleagues (friends) who, with their support and insights, contributed to enhancing our work and, more importantly, to making our journey this far a happy one. Particularly, we wish to share our gratitude with the members of the "FODIP" research and innovation lab at the University of Barcelona—very specially, with professors Francesc Imbernón

(Paco), Beatriz Jarauta, and Isaac Calduch—with Professor Mari Paz Sandín, also from the University of Barcelona, and with the members and staff of the Center for Research and Development of Higher Education at The University of Tokyo. Thank you!

References

- Akiba, M., Wilkinson, B., 2016. Adopting an international innovation for teacher professional development: state and district approaches to lesson study in Florida. *J. Teach. Educ.* 67 (1), 74–93.
- Beauchamp, E.R. (Ed.), 1991. *Windows on Japanese Education*. Greenwood Press, Westport.
- Becker, J.P. (Ed.), 1989. *Mathematics Teaching in Japanese Elementary and Secondary Schools*. Report of the ICTM Japan Mathematics Delegation. Available at: <https://files.eric.ed.gov/fulltext/ED308070.pdf>. (Accessed 1 February 2022).
- Bell, A., Mladenovic, R., 2015. Situated learning, reflective practice and conceptual expansion: effective peer observation for tutor development. *Teach. High. Educ.* 20 (1), 24–36.
- Bertram, A., Loughran, J., 2012. Science teachers' views on CoRes and PaP-eRs as a framework for articulating and developing pedagogical content knowledge. *Res. Sci. Educ.* 42 (6), 1027–1047.
- Boeing, G., 2016. Visual analysis of nonlinear dynamical systems: chaos, fractals, self-similarity and the limits of prediction. *Systems* 4 (4), 37–54.
- Branch, R.M., 2009. *Instructional Design: The ADDIE Approach*. Springer, New York.
- Cajkler, W., Wood, P., 2016. Lesson study and pedagogic literacy in initial teacher education: challenging reductive models. *Br. J. Educ. Stud.* 64 (4), 503–521.
- Calduch, I., Rattray, J., 2021. Revisiting postliminal variation in threshold concepts: issues of unexpected transformation and legitimization. *Stud. High Educ.* <https://doi.org/10.1080/03075079.2021.1910651>.
- Chalmers, D., 2006. Strong and weak emergence. In: Clayton, P., Davies, P. (Eds.), *The Re-emergence of Emergence*. Oxford University Press, Oxford, pp. 244–256.
- Cheng, E., 2019. *Successful Transposition of Lesson Study: A Knowledge Management Perspective*. Springer, Singapore.
- Chokshi, S., Fernandez, C., 2004. Challenges to importing Japanese lesson study: concerns, misconceptions and nuances. *Phi Delta Kappan* 85 (7), 520–525.
- Chong, W.H., Kong, C.A., 2012. Teacher collaborative learning and teacher self-efficacy: the case of lesson study. *J. Exp. Educ.* 80 (3), 263–283.
- Coenders, F., Verhoef, N., 2019. Lesson study: professional development (PD) for beginning and experienced teachers. *Prof. Dev. Educ.* 45 (2), 217–230.
- Crockett, M.D., 2002. Inquiry as professional development: creating dilemmas through teachers' work. *Teach. Teach. Educ.* 18, 609–624.
- Crutchfield, J.P., 2012. Between order and chaos. *Nat. Phys.* 8 (1), 17–24.
- Cummings, W., 1980. *Education and Equality in Japan*. Princeton University Press, Princeton.
- Deleuze, G., Parnet, C., 2007. *Dialogues II*. Columbia University Press, New York (Original Work Published 1977).
- Diaz, H., Fernandez, C., Gill, A., Jackson, B., Ma, L., Silva, M., 2005. The challenges of implementing lesson study. In: Wang-Iverson, P., Yoshida, M. (Eds.), *Building Our Understanding of Lesson Study*. Research for Better Schools, Philadelphia, pp. 127–137.
- Doig, B., Groves, S., 2011. Japanese lesson study: teacher professional development through communities of inquiry. *Math. Teach. Educ. Dev.* 13 (1), 77–93.
- Dudley, P., 2013. Teacher learning in lesson study: what interaction-level discourse analysis revealed about how teachers utilised imagination, tacit knowledge of teaching, and fresh evidence of pupils' learning, to develop practice knowledge and so enhance their pupils' learning. *Teach. Teach. Educ.* 34, 107–121.
- Fernandez, C., Yoshida, M., 2004. *Lesson Study: A Japanese Approach to Improving Mathematics Teaching and Learning*. Lawrence Erlbaum Associates, Mahwah.
- Fujii, T., 2014. Implementing Japanese lesson study in foreign countries: misconceptions revealed. *Math. Teach. Educ. Dev.* 16 (1), 65–83.
- Fujii, T., 2016. Designing and adapting tasks in lesson planning: a critical process of lesson study. *ZDM Mathematics Educ.* 48, 411–423.
- Gaudin, C., Chaliès, S., 2015. Video viewing in teacher education and professional development: a literature review. *Educ. Res. Rev.* 16, 41–67.
- Goldstein, J., 2016. Emergence, self-transcendence, and education. In: Koopmans, M., Stamovlasis, D. (Eds.), *Complex Dynamical Systems in Education*. Springer, Cham, pp. 39–57.
- Grossman, P., Hammerness, K., McDonald, M., 2009. Redefining teaching, re-imagining teacher education. *Teach. Teach.* 15 (2), 273–289.
- Hargreaves, A., O'Connor, M.T., 2018. Solidarity with solidity. The case for collaborative professionalism. *Phi Delta Kappan* 100 (1), 20–24.
- Heath, C., Luff, P., 2013. Embodied action and organizational activity. In: Sidnell, J., Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Wiley-Blackwell, Malden, pp. 283–307.
- Hervas, G., Medina, J.L., 2020a. Key components of lesson study from the perspective of complexity: a theoretical analysis. *Teach. Teach. Theor. Pract.* 26 (1), 118–128.
- Hervas, G., Medina, J.L., 2020b. What higher education teachers talk about when they talk about their students: nature of their conversations during lesson study. *Res. Post-Compul. Educ.* 25 (4), 461–479.
- Hervas, G., Medina, J.L., 2021. Learning and developing during lesson study through professional conversations. *Int. J. Acad. Dev.* 26 (3), 237–251.
- Hervas, G., Medina, J.L., 2022. Higher education teachers' perception and use of content representations in lesson study. *Int. J. Lesson Learn. Stud.* <https://doi.org/10.1108/IJLLS-09-2021-0077>.
- Hervas, G., Medina, J.L., Sandin, M.P., 2020. Participants' views of the use of video in lesson study in higher education in Spain: an exploratory multiple case study. *J. Res. Technol. Educ.* 52 (4), 461–473.
- Hervas, G., 2021. Lesson study as a faculty development initiative in higher education: a systematic review. *AERA Open* 7 (1), 1–19. <https://doi.org/10.1177/2332858420982564>.
- Hervas, G., 2022. The international popularisation of lesson study: early studies and their relevance in later literature. *Rev. Panam. Pedagog.* 33, 38–54.
- Hiebert, J., Stigler, J.W., 2017. Teaching versus teachers as a lever for change: comparing a Japanese and a U.S. perspective on improving instruction. *Educ. Res.* 46 (4), 169–176.
- Horio, T., 1988. *Educational Thought and Ideology in Modern Japan: State Authority and Intellectual Freedom* (Platzer, S. (Ed. and trans.)). Tokyo University Press, Tokyo.
- Howe, E.R., 2006. Exemplary teacher induction: an international review. *Educ. Philos. Theor.* 38 (3), 287–297.
- Ingersoll, R.M. (Ed.), 2004. *A Comparative Study of Teacher Preparation and Qualification in Six Nations*. Consortium for Policy Research in Education, Philadelphia.
- Ishii, T., 2016. Historical overview of lesson study. In: Tanaka, K., Nishioka, K., Ishii, T. (Eds.), *Curriculum, Instruction and Assessment in Japan*. Routledge, New York, pp. 57–72.
- Juhler, M.V., 2016. The use of lesson study combined with content representation in the planning of physics lessons during field practice to develop pedagogical content knowledge. *J. Sci. Teach. Educ.* 27 (5), 533–553.
- Juhler, M.V., 2017. Assessment of understanding: student teachers' preparation, implementation and reflection of a lesson plan for science. *Res. Sci. Educ.* 48 (3), 515–532.
- Land, R., Rattray, J., Vivian, P., 2014. Learning in the liminal space: a semiotic approach to threshold concepts. *High Educ.* 67 (2), 199–217.
- Langley, G.J., Moen, R.D., Nolan, K.M., Nolan, T.W., Norman, C.L., Provost, L.P., 2009. *The Improvement Guide*, second ed. Jossey-Bass, San Francisco.
- Lewis, C., Hurd, J., 2011. *Lesson Study Step by Step*. How Teacher Learning Communities Improve Instruction. Heinemann, Portsmouth.
- Lewis, C., Lee, C., 2017. The global spread of lesson study. In: Akiba, M., LeTendre, G.K. (Eds.), *International Handbook of Teacher Quality and Policy*. Routledge, New York, pp. 185–203.
- Lewis, C., Takahashi, A., 2013. Facilitating curriculum reforms through lesson study. *Int. J. Lesson Learn. Stud.* 2 (3), 207–217.
- Lewis, C., Perry, R., Murata, A., 2006. How should research contribute to instructional improvement? The case of lesson study. *Educ. Res.* 35 (3), 3–14.
- Lewis, C.C., 1995. *Educating Hearts and Minds: Reflections on Japanese Preschool and Elementary Education*. Cambridge University Press, Cambridge.

- Lewis, C., 2013. How Do Japanese Teachers Improve Their Instruction? Synergies of Lesson Study at the School, District and National Levels (Commissioned Paper by the National Research Council Board on Science Education). National Academy Press, Washington. Available at: https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_084385.pdf. (Accessed 1 February 2022).
- Lewis, C., 2015. What is improvement science? Do we need it in education? *Educ. Res.* 44 (1), 54–61.
- Lim, C., Lee, C., Saito, E., Syed Haron, S., 2011. Taking stock of lesson study as a platform for teacher development in Singapore. *Asia Pac. J. Teach. Educ.* 39 (4), 353–365.
- Loughran, J., Milroy, P., Berry, A., Gunstone, R., Mulhall, P., 2001. Documenting science teachers' pedagogical content knowledge through PaP-eRs. *Res. Sci. Educ.* 31 (2), 289–307.
- Loughran, J., Berry, A., Mulhall, P., 2004. In search of pedagogical content knowledge in science: developing ways of articulating and documenting professional practice. *J. Res. Sci. Teach.* 41 (4), 370–391.
- Loughran, J., Mulhall, P., Berry, A., 2008. Exploring pedagogical content knowledge in science teacher education. *Int. J. Sci. Educ.* 30 (10), 1301–1320.
- Makinae, N., 2019. The origin and development of lesson study in Japan. In: Huang, R., Takahashi, A., Ponte, J. (Eds.), *Theory and Practice of Lesson Study in Mathematics*. Advances in Mathematics Education. Springer, Cham, pp. 169–181.
- Matoba, M., 2017. Building academic-oriented lesson study. *Educ. Res. Bulletin Tokai Gakuen Univ.* 3, 120–134. Available at: http://repository.tokaigakuen-u.ac.jp/dspace/bitstream/11334/1492/1/spkiyo_003_12.pdf. (Accessed 1 February 2022).
- McAlpine, L., Weston, C., Timmermans, J., Berthiaume, D., Fairbank-Roch, G., 2006. Zones: reconceptualizing teacher thinking in relation to action. *Stud. High Educ.* 31 (5), 601–615.
- McIntyre, A., 2008. *Participatory Action Research*. Sage, Thousand Oaks.
- Mintz, J., 2016. Bion and Schön: psychoanalytic perspectives on reflection in action. *Br. J. Educ. Stud.* 64 (3), 277–293.
- Moghaddam, A., Sarkar Arani, M.R., Kuno, H., 2015. A collaborative inquiry to promote pedagogical knowledge of mathematics in practice. *Issues Educ. Res.* 25 (2), 170–186.
- Morin, E., 2005a. Introduction à la pensée complexe. Seuil, Paris.
- Morin, E., 2005b. La epistemología de la complejidad. In: Solana, J.L. (Ed.), *Con Edgar Morin, por un pensamiento complejo*. Akal, Madrid, pp. 27–52.
- Morin, E., 2007. Restricted complexity, general complexity. In: Gersheson, C., Aerts, D., Edmonds, B. (Eds.), *Worldviews, Science and Us: Philosophy and Complexity*. World Scientific, London, pp. 5–29.
- Morin, E., 2010. *La Mente Bien Ordenada*. Seix Barral, Barcelona (Original Work Published 1999).
- Murata, A., 2011. Introduction: conceptual overview of lesson study. In: Hart, L.C., Alston, A., Murata, A. (Eds.), *Lesson Study Research and Practice in Mathematics Education*. Springer, Heidelberg, pp. 1–12.
- Nilsson, P., Loughran, J., 2012. Exploring the development of pre-service science elementary teachers' pedagogical content knowledge. *J. Sci. Teach. Educ.* 23 (7), 699–721.
- Nolte, D.D., 2010. The tangled tale of phase space. *Phys. Today* 63 (4), 33–38.
- Pang, M.I., Marton, F., 2003. Beyond "lesson study": comparing two ways of facilitating the grasp of some economic concepts. *Instr. Sci.* 31, 175–194.
- Postholm, M.B., 2019. The teacher educator's role as enacted and experienced in school-based development. *Teach. Teach.* 25 (3), 320–333.
- Robbins, P., Aydede, M. (Eds.), 2009. *The Cambridge Handbook of Situated Cognition*. Cambridge University Press, Cambridge.
- Rock, T.C., Wilson, C., 2005. Improving teaching through lesson study. *Teach. Educ. Q.* 32 (1), 77–92.
- Rohlen, T.P., Björk, C. (Eds.), 1998. *Education and Training in Japan*, vols. 1–3. Routledge, New York.
- Rohlen, T.P., LeTendre, G. (Eds.), 1996. *Teaching and Learning in Japan*. Cambridge University Press, Cambridge.
- Rohlen, T.P., 1983. *Japan's High School*. University of California Press, Oakland.
- Sarkar Arani, M.R., Fukaya, K., Lassegard, J.P., 2010. "Lesson study" as professional culture in Japanese schools: an historical perspective on elementary classroom practices. *Japan Rev.* 22, 171–200.
- Schön, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- Schön, D.A., 1987. *Educating the Reflective Practitioner*. Jossey-Bass, San Francisco.
- Shimahara, N., Sakai, A., 1995. *Learning to Teach in Two Cultures: Japan and the United States*. Garland, New York.
- Shimahara, N., 1979. *Adaptation and Education in Japan*. Greenwood Press, Westport.
- Shimizu, S., Chino, K., 2015. History of lesson study to develop good practices in Japan. In: Inprasitha, M., Isoda, M., Wang-Iverson, P., Yeap, B.-H. (Eds.), *Lesson Study: Challenges in Mathematics Education*. World Scientific, Singapore, pp. 123–140.
- Steinert, Y., Mann, K., Anderson, B., Barnett, B.M., Centeno, A., et al., 2016. A systematic review of faculty development initiatives designed to enhance teaching effectiveness: a 10-year update: BEME Guide No. 40. *Med. Teach.* 38 (8), 769–786.
- Steinert, Y., 2010. Becoming a better teacher: from intuition to intent. In: Ende, T. (Ed.), *Theory and Practice of Teaching Medicine*. American College of Physicians, Philadelphia, pp. 73–93.
- Stepanek, J., Appel, G., Leong, M., Turner, M., Mitchell, M., 2007. *Leading Lesson Study. A Practical Guide for Teachers and Facilitators*. Corwin Press, Thousand Oaks.
- Stevenson, H.W., 1991. Japanese elementary school education. *Elem. Sch. J.* 92 (1), 109–120.
- Stigler, J.W., Hiebert, J., 1999. *The Teaching Gap*. The Free Press, New York.
- Stigler, J.W., Lee, S.-Y., Stevenson, H.W., 1987. Mathematics classrooms in Japan, Taiwan, and the United States. *Child Dev.* 58 (5), 1272–1285.
- Takahashi, A., McDougal, T., 2016. Collaborative lesson research: maximizing the impact of lesson study. *ZDM* 48 (4), 513–526.
- Takahashi, A., Watanabe, T., Yoshida, M., Wang-Iverson, P., 2005. Improving content and pedagogical knowledge through Kyozaikenkyu. In: Wang-Iverson, P., Yoshida, M. (Eds.), *Building Our Understanding of Lesson Study. Research for Better Schools*, Philadelphia, pp. 101–110.
- Trigwell, K., Shale, S., 2004. Student learning and the scholarship of university teaching. *Stud. High Educ.* 29 (4), 523–536.
- van Dijk, T.A., 2010. *Discurso y conocimiento. Una aproximación sociocognitiva*. Cambridge University Press, Cambridge.
- van Manen, M., 1991. *The Tact of Teaching: The Meaning of Pedagogical Thoughtfulness*. State University of New York Press, Albany.
- von Foerster, H., 1984. Principles of self-organization in a socio-managerial context. In: Ulrich, H., Probst, G.J.B. (Eds.), *Self-Organization and Management of Social Systems*. Springer-Verlag, Berlin, pp. 2–24.
- Vrikki, M., Warwik, P., Vermunt, J.D., Mercer, N., Van Halem, N.V., 2017. Teacher learning in the context of lesson study: a video-based analysis of teacher discussions. *Teach. Teach. Educ.* 61, 211–224.
- Walker, G., 2013. A cognitive approach to threshold concepts. *High Educ.* 65 (2), 247–263.
- Wang, J., Hartley, K., 2003. Video technology as a support for teacher education reform. *J. Technol. Teach. Educ.* 11 (1), 105–138.
- Wong, J.L.N., 2018. Why social capital is important for mentoring capacity building of mentors: a case study in Hong Kong. *Teach. Teach.* 24 (6), 706–718.
- Yoshida, M., 1999. *Lesson Study: A Case Study of a Japanese Approach to Improving Instruction Through School-Based Teacher Development* (Doctoral Dissertation). The University of Chicago, Chicago.
- Yoshida, M., 2012. Mathematics lesson study in the United States. *Int. J. Lesson Learn. Stud.* 1 (2), 140–152.
- Yuan, R., Zhang, J., Yu, S., 2018. Understanding teacher collaboration processes from a complexity theory perspective: a case study of a Chinese Secondary School. *Teach. Teach.* 24 (5), 520–537.

Examining conceptual frameworks for understanding teacher learning in the domain of formative assessment

Brent Duckor^a and Carrie Holmberg^b, ^a San José State University, Department of Teacher Education, Educational Leadership, San Jose, CA, United States; and ^b San José State University, Department of Teacher Education, San Jose, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Assessment reform and school reform in the US context	583
The shift from formative evaluation to formative assessment discourse	584
Characteristics, processes and facet-based perspectives on formative assessment	587
Professional growth over time and developmental trajectories perspectives on formative assessment	589
Conclusion and future directions	592
References	593

Glossary

Construct Map Construct maps, like learning progressions, represent variations in observed skills, practices, and behaviors that help us to infer who possesses more or less of a construct (Wilson, 2005). Bounded by upper and lower anchors, these maps articulate in advance what structures (levels, facets, dimensions) of teacher FA practice look like and subject our construct theories to empirical investigation, and hence grounds for falsification. Construct maps are not rubrics or coding sheets; they are hypotheses about the structure of a dimension that requires quantitative study (Duckor et al., 2009)

FA moves Duckor and Holmberg (2017) identified 7 high-leverage FA moves that make a difference in student learning when used during lessons, minute by minute. The moves are:

- Priming – Preparing the groundwork
- Posing – Using an effective range of questions
- Pausing – Giving students adequate time to think and respond
- Probing – Asking follow-up questions that elicit student thinking and deepen student understanding
- Bouncing – Sampling on a variety of responses
- Tagging – Publicly representing the variation in student thinking
- Binning – Seeing patterns and using patterns to inform “Next Steps”

While these moves are interconnected, they don’t occur in rigid sequences. Rather, they are used flexibly and creatively. They can be mobilized “on the fly” in clusters—such as posing-pausing-probing—at key hinge points in a lesson. Moves can be planned and/or improvised during a lesson

Teacher learning trajectories Applying the logic of learning progressions to the study of teacher development, teacher learning trajectories are hypothesized levels of a construct along a dimension (Duckor and Holmberg, 2019). In the domain of formative assessment, these trajectories can be specified, examined, and interpreted with qualitative and quantitative methods in terms of content and structure

Progress guides A progress guide is a formative evaluation tool. Scoring guides are less cognitively and linguistically complex than an analytic rubric; therefore, they may be employed as a more student-friendly binning strategy. Progress guides tend to contain less verbiage and dysfunctional detail that might overwhelm students, including students learning English and students with special needs

Learning progressions Learning progressions describe a trajectory of learning within a content domain that spans a much longer period of time than a lesson or single unit of instruction. The trajectory may provide a multi-year image of students’ successively more sophisticated performance levels, for example, in addressing the big idea of “Why things sink and float?” in physics. Learning progressions are mediated by targeted instruction and curriculum. Heritage (2008) notes that a progression of learning in a domain can provide the big picture of what is to be learned, support instructional planning, and act as a touchstone for formative assessment

Any review of the myriad discourses on assessment literacy must confront two problems. First, one must review the multitude of theories of assessment proficiency; second, one must ground those theories in actual trajectories of assessment practice. Our focus in this article offers a third conceptual challenge, one that is aimed at unpacking the place of teacher understanding and use of formative assessment along trajectories of practice in a dynamic, evolving learning environments.

Rather than exhaustively survey all the literature related to what has been dubbed “assessment literacy” or take up debates on what teachers need to know and be able to do with “assessment data” for example, for State or national licensure exams, or produce a definitive review of the road to “formative assessment” that grew out of the evaluation literatures of the 1960s, our goal is more circumspect, and we hope, productive. The theses we explore in this article are intended to: (a) recall the need for

assessment reform in the US prior to the naming of formative assessment as a discourse field: (b) contextualize the development of formative assessment scholarship from its roots in evaluation theory; and (c) explore the conceptual frameworks offered by scholars for unpacking teachers' growth and development in the movement toward assessment for learning orientated practice.

When examining the literature on teachers' work—whether in the contexts of pre-service, induction, and in-service settings, projects, and programs—there is a paucity of educational research aimed at unpacking teachers' learning in the domain of formative assessment, not to mention the area of classroom assessment more broadly. There are expert-novice studies on teachers and teaching that focus on lesson planning, or enactments of lessons, and processes for productive reflection (See, e.g., Borko and Livingston, 1989; Putnam and Borko, 2000; Feiman-Nemser, 2001a,b).

But a clear, unswerving look at the individual development and comparative differences between teachers' knowledge and use of assessment practices has been undertheorized. The idea of developmental trajectories of formative assessment practice lurks beneath many discussions of effective classroom strategies and high-leverage teaching practices (TeachingWorks, 2018) but too frequently the development of these strategies and practices is left unpacked by scholars. The notion that teachers have mental models p-prims, even misconceptions that they replace with other schema through experience is absent from the literature. Rather than examining teachers' developmental trajectories and learning progressions in the field of formative assessment, notions of assessment literacy have been advanced to explain differences in teachers' practices (Stiggins, 2002). This approach has further led to the assembly of "measures" of self-reported beliefs, attitudes and dispositions toward classroom assessment (Gotch and French, 2014) rather than exploration of teacher learning trajectories and investigations of teachers-as-learners who develop over time (Duckor and Holmberg, 2017).

Whatever the role of these definitions or surveys of assessment literacy (and, by presumption, assessment illiteracy) may be for experts and educational researchers' knowledge base, there remains a lack of understanding of teacher development, growth, and change as formative assessors. Observed variations in naturalistic settings taken from samples of teachers moving from novice to more expert states of formative assessment practice have been scarce. There are likely many explanations for scholarly inattention and disinterest in teachers' development as a subject of rigorous study in the formative assessment literature. Part of the scholarly inattention is rooted in the historical struggle to define the field itself—against the backdrop to testing and evaluation theory. Experts were learning to define what makes formative assessment formative.

In this article, we ask the reader to move with us beyond static pictures of characteristics, definitions and predicates of "formative assessment" (FA) practice. The aim here is to move the field toward dynamic models of "assessing formatively" that reveal unfolding levels of mastery along trajectories that respect teachers' individual zones of proximal development. To do this, we ask:

- What can educational researchers and experts tell us about expected outcomes i.e., levels of progress, for teachers who are developing their expertise with questioning techniques and strategies, making sense of data from an exit ticket, putting concept maps on the dry erase board to document variation in student thinking, using a gallery walk for a peer-based feedback protocol, or differentiating feedback through self-assessment comments added to a shared document, and so forth?
- What do we know about how, and in what ways, teachers learn and enact discreet FA practices, strategies, and tactics along various proposed dimensions? Are there facets of FA practice that resist the notion of qualitative differences and progressions? If so, which ones? Why?
- What are teachers' zones of proximal development with respect to the planning for enactment of and reflection on FA practices individually and taken as a whole? To use a mountaineering metaphor, are there "fixed lines" and "bottlenecks" and "slippage" that "naturally" accompany the individual's movement toward FA expertise?
- What can educational researchers and experts tell us about types of consistent, timely and useful feedback at a given level of FA practice for those making the journey toward learning goals in this domain that are generalizable and reproducible?

Each of these kinds of research questions, while beyond the scope of inquiry for this article, is nonetheless necessary for new, dynamic conceptualizations of teacher learning and growth in the discourse field of formative assessment.

Teachers, like students, are learners when it comes to formative assessment and each deserves to be seen through the research lens of growth as they learn to master new practices. Each and every teacher who is learning to formatively assess, for example, is developing those skills, practices, and dispositions across the arc of a career. There are pivot points and roadblocks along the path. There are breakthroughs and cul-de-sacs along the way.

We argue that teachers-as-learners are at the present moment the necessary subjects of education reform, which include but are not limited to reforms aimed at what experts have called "formative assessment" (Black and Wiliam, 1998a,b), "assessment for learning" (Stiggins, 2002), and "visible learning" (Hattie, 2009). Prior to *No Child Left Behind* (2001), it is clear that local assessment reform led by teachers was well underway. In fact, while not always visible to many outside observers, schools in the US were engaged in assessment reforms decades well before the rise of these particular FA discourse fields. Classroom teachers and school leaders were experimenting with "formative" assessment practices for children in multiple settings that delivered results without referring to the term formative assessment. Educational researchers and experts and teachers were more focused on authentic assessment and portfolio-based assessment in the US context (Darling-Hammond, 1997; Duckor and Perlstein, 2014; Meier, 1995).

Assessment reform and school reform in the US context

Teachers and teacher learning are and have been at the center of school reform efforts in the US for over a half century. Progressive educators from Dewey to Delpit have emphasized how teachers can make a positive, sustained impact on the lives of their students in the educational system—from early childhood to adulthood. The pursuit of educational excellence and equity in public education demands careful attention to the daily work of teachers in classrooms.

Reformers have focused on powerful examples of teachers whose work unites pedagogical, instructional and assessment practices, weaving them into rich learning experiences for all students. Some of these same reformers have noted that high-leverage, promising, culturally responsive practices not only elevate all students' levels of authentic engagement and achievement, but that these practices are essential to better serving historically disadvantaged minorities and minoritized students (Ladson-Billings, 1995; Hollins et al., 1994).

In the US context, the work of Darling-Hammond (1997) epitomizes the promise of teachers' work and practice for aiming at better outcomes that serve historically underserved students. In the *Right to Learn*, Darling-Hammond reminds policymakers that schools that see teachers as levers of instructional and assessment reform do produce results on multiple measures of success. In documenting examples of schools where classroom and school level assessment practices work for children of color, Darling-Hammond (1997) wrote:

The innovative curriculum and assessment efforts of several of the original Center for Collaborative Education schools served as exemplars for [New York] state's initial rethinking of its curriculum and assessment practices [citing a New York State Council report, 1994]. Portfolio and performance assessments models from CPESS, International High School and Urban Academy ... provide examples of authentic assessments to school districts across the state (p. 156).

Portfolios are rigorously evaluated and often sent back for more work; students learn what it takes to develop a piece of work that meets the standards of inquiry in a field and the standards of written and oral discourse demanded by committee members. They get feedback and revise and revise and revise; they internalize standards; they develop the capacity for sustained effort and ambitious work (p. 157).

An alternative approach to reform uses standards and assessments as means of giving feedback to educators and tools for organizing student and teacher learning, rather than as a sledgehammer to beat schools into change ... (p. 241).

It is clear that effective and equitable schools are places where teachers integrate rich learning, teaching, and assessment practices. Assessment and instruction—in any subject or discipline—are woven together in a coherent, consistent whole by these teachers. The craft of good teaching is seamlessly integrated with good assessment practices in all of these case studies.

The argument for authentic assessment in action centers education reform on teachers and students' work in classrooms that blend assessment and instruction predates much of the literature on formative assessment. Yet echos of the formative assessment ethos and references to discrete FA practices and purposes such as feedback and evaluation resound in these writings prior to landmark studies of studies.

Darling-Hammond et al. (1995) noted that:

Increasingly, local schools, schools districts, and states are experimenting with these methods and with other alternatives to standardized testing for assessing student learning and performance. Persuaded that traditional standardized tests fail to measure many of the important aspects of learning, and that they do not support many of the useful strategies for teaching, practitioners are introducing alternative approaches to assessment in classrooms—approaches that help teachers look more carefully and closely at students, their learning, and their work (p. 4).

... [M]embers of school communities must be engaged in the development and use of these assessments. This is because teacher learning about the deep structures of curriculum, the nature and nuances of student thinking, and the connections between teaching efforts and student performances derive substantially from firsthand, constructivist encounters with assessment development, and from the subsequent evaluation of student work (p. 253).

No longer left to the outside evaluation expert or standardized testing authorities, teachers' daily assessment practices came to the forefront of the struggle for excellence and equity in K-12 education. Schools like Central Park East Secondary School (CPESS) in New York City epitomize what teachers engaged in the struggle for educational equity and specifically what deep equity looked like from an assessment lens. Learning by doing became synonymous with assessment by doing in the Coalition of Essential Schools movements across the US (Duckor and Perlstein, 2014;Sizer, 1996; Wiggins, 1993).

In "Assessing Habits of Mind: Teaching to the Test" Duckor and Perlstein (2014, pp. 16–17) document that teachers' expertise in assessment for learning was indispensable to minoritized students' success and was being implemented in the late 1980s and early

1990s in the US Interviewing a pool of teachers who were deeply involved at these reform-oriented K-12 schools, Duckor and Perlstein found:

CPESS educators developed their approach to assessment to better articulate their approach to teaching and learning. They thus had a strong sense not only of what they were assessing but also of how and why they were assessing it. "CPESS," recalls English teacher Lori C., "was based on formative assessment. ... What we were doing was figuring out where the kids were at, what did they need to know, how were you going to get them and bridge the gap between where you wanted to get them with the project and where they were in their learning and understanding."

Numerous studies, books, video artifacts, and testimonies attest to deeper learning through teachers' work with students, and students' assessment of their own and others' work. Teachers exposed students to embedded authentic assessment across 6 years of schooling—from 7th to 12th grades. It was the deep immersion in formative assessment practices—supported by staff, administrators, and parents—in a community of learners committed to those assessment practices that led to achievement gains and myriad opportunities to advance individual outcomes (Bensman, 2000).

The shift from formative evaluation to formative assessment discourse

Paralleling the lived experiences and emerging literatures on authentic assessment, portfolio assessment, and feedback-driven classroom-based assessment practices from urban American primary and secondary schools in the late 1980s and early 1990s, a group of assessment experts weighed in on "what works" to advance student assessment for learning. Documenting studies from reformed classrooms and reforming school settings across the globe, these scholars began to use the term formative assessment to articulate a complimentary stance on systemic, school-based reform in the US.

Moving away from the classical era of formative evaluation (Scriven, 1966, 1973), this cadre of US, Australian, and UK experts re-centered the classical evaluation field on formative assessment; that is, they looked toward a set of teacher-driven practices that showed promise for improving student learning outcomes. A pioneer in the evaluation field, Scriven (1966) argued that "... formative evaluators should be very sharply distinguished from the summative evaluators." He further noted that "[evaluators] may certainly work in developing an acceptable summative evaluation schema, but they should of course exclude themselves from any judgmental role" (pp. 8–9).

Sending a clear warning to evaluators who fail to distinguish between assessment purposes, Scriven presciently wrote:

... testing for learning of certain rather delicate and pervasive concepts may be itself destructive, in that it makes the student too self-conscious about the role of a concept at too early a stage, thereby preventing its natural and proper development ... It would not show up on tests, indeed it might be stultified by the intrusion of tests, in any given year, but it has to be in the curriculum in order to produce the finished products that we desire. In this case, evaluation seems to be both incompetent and possibly destructive (p. 9).

Traditionally, the evaluation field had focused on the needs and designs of test makers and curriculum developers' agendas to improve the quality of results from evaluation studies. The uses of assessment data—for teachers and students—was secondary and not the subject or focus of Scriven and others' concerns in the evaluation theory field in the 1960s.

Stake (1967) is an outlier in the evaluation field. He identified the limits of classical approaches to evaluation to inform teaching and learning. Imbued with methodological formalism and the pretense of scientific method, too many evaluations emphasized the problems of stakeholders outside the messy classroom, the imperfect curriculum, and the dynamic ever-changing relationship between learners and their teachers. Stake noted: "Today, educators fail to perceive what formal evaluation could do for them. They should be imploring measurement specialists to develop a methodology that reflects the fullness, the complexity, and the importance of their programs. They are not" (p. 523).

Noting the skepticism toward evaluation by teachers who assess daily and need assessment data for learning, Stake (1967) wrote:

Formal evaluation of education is recognized by its dependence on checklists, structured visitation by peers, controlled comparisons, and standardized testing of students. Some of these techniques have long histories of successful use. Unfortunately, *when planning an evaluation, few educators consider even these four*. The more common notion is to evaluate informally: to ask the opinion of the instructor, to ponder the logic of the program, or to consider the reputation of the advocates. Seldom do we find a search for relevant research reports or for behavioral data pertinent to the ultimate curricular decisions.[Emphasis added] (p. 523).

Dissatisfaction with the formal approach is not without cause. Few highly relevant, readable, research studies can be found. The professional journals are not disposed to publish evaluation studies. Behavioral data are costly, and often do not provide the answers. Too many accreditation -type visitation teams lack special training or even experience in evaluation. Many checklists are ambiguous; some focus too much attention on the physical attributes of a school. Psychometric tests have been developed primarily to differentiate among students at the same point in training rather than to assess the effect of instruction on acquisition of skill and understanding. *Today's educator may rely little on formal evaluation because its answers have seldom been answers to questions he is asking* [Emphasis added] (p. 523).

Evaluation theory dealt with schools, classrooms, and outcomes but it was incapable of seeing evolving relationships between teachers and students engaged in classroom-based assessment. In a groundbreaking work on evaluation, Bloom et al. (1971) ventured that assessment data from evaluation projects and programs could be useful for others, including teachers and students.

In the *Handbook of formative and summative evaluation of student learning*, Bloom and his colleagues claimed a new evaluation stance that was “intended primarily for present and future classroom teachers,” and represented the “state of the art” of evaluating student learning. In this period of transition, from formal “scientific” evaluation methods to the new, classroom-level assessment theory, the challenge was: How to make ostensibly summative tests, formative? Using examples of “formative tests” from STEM subjects such as Chemistry (Fig. 1), these authors saw a unique role for teachers-as-evaluators and students-as-evaluators.

The teacher-evaluator was now expected to conduct “error analysis,” provide “corrective” feedback, and support interim unit-level interventions before the final results of learning in a given subject area. Compared to Scriven’s narrower conception of evaluation’s role in understanding schools and classrooms, Bloom and colleagues signaled a transition to classroom assessment practice as a crucial site of intervention in educational outcomes: “We regard formative evaluation as useful not only for curriculum construction but also for instruction and student learning” (p. 117).

For Bloom and colleagues, the teacher—not the expert outsider or policy maker or curriculum developer—shifted more clearly into focus. The teacher was seen as the one who monitors student progress and who supports students’ efforts at self-assessment. All of a teacher’s complex work, practice, and efforts in the classroom would soon become the rightful subject of study for formative evaluation. But was that enough for those interested in the role of assessment to shape outcomes—between teachers and students—at the classroom level?

Signaling a break with the field of the classical evaluation tradition, a new cadre of formative-oriented assessment scholars and learning theorists sought to place more effort and attention on teachers’ and students’ work in the classroom. In the 1980s and 1990s, classroom assessment and evaluation began to focus on discrete lessons and units of instruction, which were suffused with exchanges of assessment information and “turns of talk” analyses. Formative processes and cycles, as they were called, were expected to advance learning outcomes inside a learning community—not only visible with outside indicators. By advancing teaching and learning goals, the problems of testing and evaluation reporting advocates were far less important for this generation of assessment scholars. They were firmly committed to the promise of assessment for learning during instructional cycles.

Sadler’s work (1989) marks an important transition from formal evaluation study of schools to classroom-level assessment theory’s role inside them. Sadler saw teachers having an essential role to play in facilitating students’ transition from passively receiving feedback to active, self-monitoring activities during the learning segment. Teachers nor their students were merely objects for formal evaluation procedures conducted by experts. Rather, both teachers and students became the subject of situated assessment practices that influence learning outcomes in real-time. Sadler wrote:

The transition from teacher-supplied feedback to learner self-monitoring is not something that comes about automatically. For an important class of learning outcomes, the instructional system must make explicit provision for students themselves to acquire evaluative expertise. It is argued that providing direct and authentic evaluative experience is a necessary (instrumental) condition for the development of evaluative expertise and therefore for intelligent self-monitoring. It is insufficient for students to rely upon evaluative judgments made by the teacher (p. 143).

Sadler noted that the assessment practices recommended “are not radically new, and are already employed in some instructional systems.” The emerging lexicon of formatively-oriented, classroom level assessment practices were known to teachers and teacher educators and could be traced by the qualitative study of classroom artifacts, discourses, and relationships. By the 1980s and early 1990s, Sadler described these “formative assessment” practices as including: the use of descriptive statements and exemplars and “criterion sheets” to help students understand and “specify” standards; the inclusion of self-assessment routines to monitor progress; and the facilitation of peer assessment protocols to invite more feedback into the learning environment prior to any final, summative, and/or end of year evaluation.

Building on the work of Sadler and others, with the publication of “Assessment and Classroom Learning” in *Assessment in Education* in 1998a and “Inside the Black Box: Raising Standards Through Classroom Assessment” (1998b) in *Phi Delta Kappan*, the shift away from the aims and purpose of formal evaluation theory to the new assessment theory was complete. Black and Wiliam (1998a) called forth a new era:

This article is a review of the literature on *classroom formative assessment*. Several studies show firm evidence that innovations designed to strengthen the frequent *feedback* that students receive about their learning yield substantial learning gains. The perceptions of students and their role in *self-assessment* are considered alongside analysis of the strategies used by teachers and the *formative strategies* incorporated in such systemic approaches as mastery learning. There follows a more detailed and theoretical analysis of the nature of feedback, which provides a *basis for a discussion of the development of theoretical models for formative assessment* and of the prospects for the improvement of practice (p. 7) [Emphasis added].

As if to further underscore the break with the problem of making summatively designed tests “formative,” a problem which clearly occupied Bloom and colleagues in the 1970s, the cadre of newly emerging assessment experts highlighted the relationship between assessment and learning inside classrooms. In a vein similar to advocates for authentic classroom-based assessment in the US, Black and Wiliam (1998a) wrote when surveying the international scene:

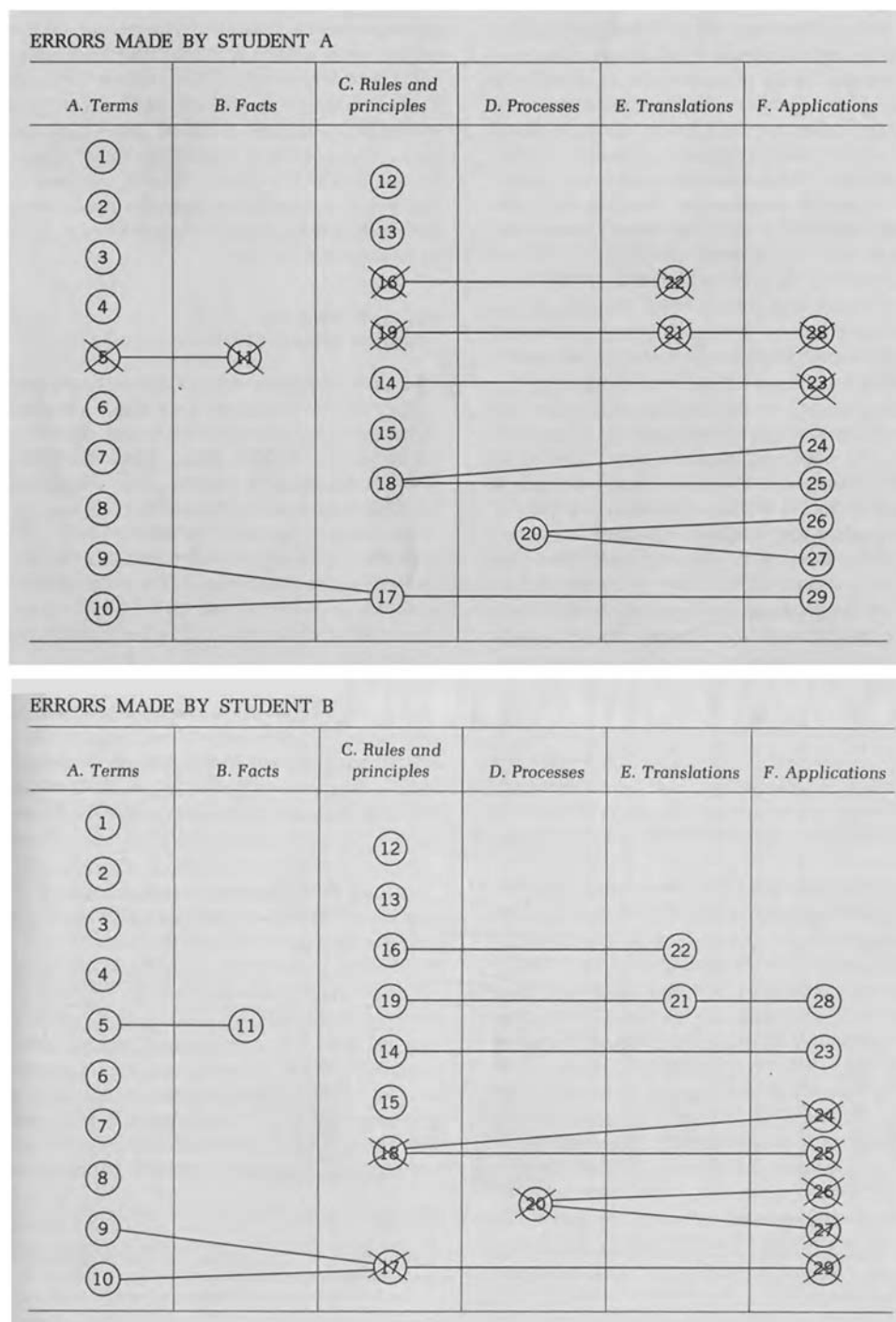


Fig. 1 Errors made by two students on the formative test in the chemistry unit.

One of the outstanding features of studies of assessment in recent years has been the shift in the focus of attention, towards greater *interest in the interactions between assessment and classroom learning and away from concentration on the properties of restricted forms of test* which are only weakly linked to the learning experiences of students. This shift has been coupled with *many expressions of hope that improvement in classroom assessment will make a strong contribution to the improvement of learning*. So one main purpose of this review is to survey the evidence which might show whether or not such hope is justified. A second purpose is to see whether *the theoretical and practical issues associated with assessment for learning* can be illuminated by a synthesis of the insights arising amongst the diverse studies that have been reported (p. 7) [Emphasis added].

By the late 1990s and early 2000s, Australian, US and UK scholars converged on the essential question: What makes formative assessment formative? Many had dropped the concern for making summative assessments formative (an agenda that continues today) and devoted themselves to the emergent field of formative assessment. Marked by a deep skepticism about the goals of evaluators and test-makers, these authors placed teachers' and students' work in assessment for learning on the same footing as the evaluation expert and curriculum designer.

Echoing school reformers in the authentic and portfolio assessment traditions of the Coalition of Essential Schools, the Career Academies, and New Tech schools in the US, these global advocates for formative assessment (FA) and assessment for learning (AFL) signaled a similar commitment to student-centered learning reform and equity-focused classroom assessment practices from the UK, Australia and Canada. Teachers, like their students, were now expected: (a) to actively participate in assessment, (b) to co-construct new identities as self-, peer-, and classroom assessors, and (c) to live and learn in the traditions of progressive education broadly defined. While it is clear that dynamic assessment (DA) provided the most explicit conceptual links to Dewey and Vygotsky, FA and AFL advocates embraced a similar if implicit commitment to constructivist teaching and teachers who were constructivists (Black and Wiliam, 2009; Leung, 2007; Perrenoud, 1998; Poehner and Lantof, 2005).

Characteristics, processes and facet-based perspectives on formative assessment

The rise of the formative assessment literature, despite its pivot away from the early classical evaluation theorists and firm rejection of standard psychometric solutions for testing, has nonetheless struggled to make its mark. Formative assessment theory and theorists have struggled—for meaning, for subjects, and for objects of study—as they built a body of professional knowledge. While tending to deemphasize the complexity of teachers' subjectivity and various trajectories of practice in complex classroom settings, scholarship in formative assessment tended toward describing sets of aspirations for best practices, on a case by case or project by project basis. The objects of study were sometimes actual groups of teachers and students, other times these studies documented sets of disembodied practices, strategies, characteristics and definitions.

Consistent with the need to define formative assessment practice as a distinct body of knowledge, and locate it within a perspective that differed from testing and evaluation theory, research has tended to focus on setting boundaries for, seeking better definitions of, and building consensus on the characteristics, processes and facets of formative assessment (Black and Wiliam, 2009; McMilian, 2010; Moss and Brookhart, 2019; Stiggins, 2002; Wiliam, 2011; Wiliam and Thompson, 2008). It has become common to catalog myriad practices, strategies and tactics that teachers and students should "know and be able to do" in the formative assessment literature.

These normative and aspirational forms of discourse are visible and traceable in discussions of key features of formative assessment today. The *characteristics-processes-facets perspective* animates much of the pioneering work in the formative assessment literary corpus. At a time when evaluation theory offered no solutions or clear directions for studying teachers' and students' work on assessment for learning, early authors worked vigorously to differentiate by careful definition and thick description what formative assessment was and was not, and how FA broadly defined looked in exemplary classrooms.

The early formative assessment researchers had to sketch outlines of what "formative" interactions might look like from general observation as they developed new theory and began documenting practice. In a 2 year investigation of secondary science teachers, Bell and Cowie (2002, pp. 61–79) identified ten characteristics that helped launch the search for authentic formative assessment. They determined that formative assessment:

- Is seen as being responsive;
- Is often a tacit process;
- Involves multiple sources of evidence related to the purposes for which the formative assessment is done;
- Relies on student disclosure;
- Uses professional knowledge and experiences;
- Is an integral part of teaching and learning;
- Is done by teachers and students;
- With attention to the purposes for engaging in formative assessment;
- Is a highly contextualized process; and
- Involves the management of dilemmas.

Hinting at the issue of teacher growth and development, the authors write "formative assessment is a highly complex and skilled activity for both the teacher and the student. Formative assessment is not something teachers are likely to learn to do in a short session in an inservice course. It is a professional skill that develops with increasing professional experience, awareness and reflection" (Bell and Cowie, 2002, p. 79).

Taking a similar approach to the *characteristics perspective*, Stiggins (2002, p. 5) lists behaviors and actions and beliefs that describe instances of authentic assessment for learning. He writes that when teachers "assess for learning, they use the classroom assessment process and the continuous flow of information about student achievement that it provides in order to advance, not merely check on, student learning." Maintaining a focus on characterizations of teachers' work, without necessarily addressing teacher's development or growth in practice, Stiggins determined that teachers "do" assessment for learning by:

- Understanding and articulating in advance of teaching the achievement targets that their students are to hit;
- Informing their students about those learning goals, in terms that students understand, from the very beginning of the teaching and learning process;
- Becoming assessment literate and thus able to transform their expectations into assessment exercises and scoring procedures that accurately reflect student achievement;
- Using classroom assessments to build students' confidence in themselves as learners and help them take responsibility for their own learning, so as to lay a foundation for lifelong learning;
- Translating classroom assessment results into frequent descriptive feedback (versus judgmental feedback) for students, providing them with specific insights as to how to improve;
- Continuously adjusting instruction based on the results of classroom assessments;
- Engaging students in regular self-assessment, with standards held constant so that students can watch themselves grow over time and thus feel in charge of their own success; and
- Actively involving students in communicating with their teacher and their families about their achievement status and improvement.

Shepard (2000, 2005, 2008), like others working to define the literature and field in the early 2000s, also sought to determine the main characteristics of formative assessment. In an effort to ensure careful delineation between assessment purposes and uses, Shepard addressed the concern with misconceptions and confusion about what formative assessment is, and what it is not.

Partly driven by a concern for teachers and students in the US who were being influenced by hyper-accountability regimes in the wake of NCLB, Shepard's scholarship established firm boundaries to preserve the gains of formative assessment advocates. Working against the marketing efforts of curriculum and testing companies that soon promised pre-packaged "formative assessments" to clients across states, Shepard (2008, p. 293) held the line for the formative assessment reform movement. Shepard argued that "effective" formative assessment is:

- More than simplistic alignment, tasks must "embody" learning goals;
- It should be curriculum-embedded (both in timing and substance). Tasks should be instructional tasks to provide insights about learning as it is occurring;
- By definition, it must enable the supportive learning processes invoked in the formative assessment literature; and
- Instructional insights should lead to coherent, theoretically sound improvements in teaching.

Heritage and associates have also held the line on the need for clear, precise and common definition of the field of formative assessment against evaluation and testing agendas (Heritage, 2008, 2010; CCSSO 2008/2018; AERA et al., 2014). The development of standard definitions of formative assessment for professional bodies and associations that emphasize the central role of the teacher and learner in assessment discourse ensured boundaries while evoking the complete transition from any dependency on evaluation or test theory.

In addition to these developments, a focus on *processes and cycles* has also accompanied the formative assessment literature (Cowie and Bell, 1999; Hattie and Timperley, 2007; Moss and Brookhart, 2019; Ruiz-Primo and Furtak, 2006, 2007; William, 2010; William and Thompson, 2008). Moving from producing control over definitions and declarative norms, researchers have also offered pictorial representations of instructional and assessment processes. These processes represent elements and facets of practice that students and teachers are expected to engage in when assessing formatively. Some processes are rooted in empirical observations of actual teachers and students' work while others are products of expected and aspirational interactions.

Cowie and Bell (1999) offered a *process* "model of formative assessment" that aimed to describe and explain the formative assessment processes as carried out by teachers (and implicitly by their students). Cowie and Bell distinguished between two kinds of FA processes that teachers and students engage in: "planned" and "interactive." The process of planned formative assessment is characterized by the teachers eliciting, interpreting and acting on assessment information. The purpose of the planned formative assessment, they maintained, is to determine how the classroom level information is collected, interpreted and acted. The four aspects of planned formative assessment, according to Cowie and Bell, are interrelated and mutually determining (Fig. 2).

Interactive FA differs from planned FA in that a specific assessment activity is not planned; rather, it arises out of a learning activity. The teachers and students within the project interacted in whole-class, small-group and one-to-one situations. The process of interactive formative assessment involved the teachers noticing, recognizing and responding to student thinking during these classroom-based interactions:

The teachers discussed how they might not notice and recognise the significance of students' comments and actions because they lacked the experience and knowledge. They mentioned that *interactive formative assessment was difficult for beginning teachers and for experienced teachers with a new class*, say at the beginning of the year (Bell and Cowie, 1997). The teachers said their awareness of student thinking (what they noticed and recognised) was often triggered if a student response was unexpected or incorrect, or if a number of students indicated that they held a similar view (Cowie and Bell, 1999, p. 109) [Emphasis added].

While these characterizations of teachers' work in the domain of formative assessment from the *process perspective* show the dynamic, contextual demands upon teachers' (and students') development, as they each practice eliciting, interpreting and acting

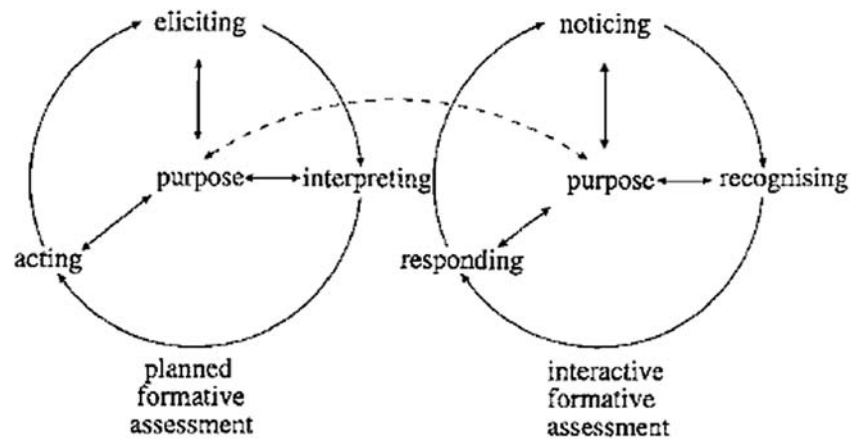


Fig. 2 A model of formative assessment.

on assessment information, it remains unclear whether change and growth are possible, not to mention probable. Too often the study of formative assessment processes and cycles have not included robust schematic diagrams of those processes from any clear, well-defined developmental perspective.

Rather than confront the difficult and challenging trajectories of “beginning,” “experienced” “novice” “expert” or any ways of describing teachers transitioning in between these liminal states, research on characteristics and processes of formative assessment offers static pictures with many predicates and pictures. The research has also tended to treat the subjects of formative assessment as “knowers and doers,” rather than as learners and emerging thinkers working to develop expertise in practice.

Absent a developmentally sensitive perspective on teachers’ growth that acknowledges transitions and bottlenecks and pitstops, too many researchers have plainly ignored the problems of practice for teachers. The fact is teachers across many settings do not always know which FA practices, cycles, and processes are most effective, when to deploy them, or why a particular combination of FA strategies actually worked for a particular student in a particular classroom (Duckor, 2014).

Professional growth over time and developmental trajectories perspectives on formative assessment

While the majority of formative assessment advocates and researchers have rarely documented teachers’ learning in the FA domain from a developmental perspective, there are exceptions (Jacobs et al., 2010; Furtak et al., 2016; Wylie et al., 2018). Shepard (2000, 2005) has described how theories of learning informed various stakeholders’ perspectives on the purposes for evaluating students. While not addressing how these learning theories may influence teachers’ work and learning in classroom assessment explicitly, the crux of the argument brings attention to the link between learning theory and assessment practice for students and presumably for their teachers.

Implicit in this schematic representation (Fig. 3) is the notion that social constructivist perspectives on children’s learning (and by association, teachers’ learning) tend to emphasize formative orientations toward classroom assessment, whereas behaviorist perspectives on learning do not. Though a full articulation of teachers’ learning trajectories in assessment for learning practices is absent, Shepard’s argument implies that classroom assessment’s alignment with learning theory is useful, if not necessary, to achieve coherence and stability within school accountability frameworks and local classroom settings (Penuel and Shepard, 2016).

The logic of Shepard’s argument—that socio-cultural approaches to growth and development matter for authentic learning—as it extends to the study of teacher learning has been largely absent from the discourses of advocates for formative assessment. Educational researchers and experts who work in the domain of formative assessment have too often flirted with behaviorist approaches to teachers’ learning by ignoring the relevance of socio-cultural accounts of teachers’ learning in the FA domain itself. Rather than offer any coherent or systematic theory as to how, why, and what teachers learn while they are becoming formative assessors, the field has been silent about trajectories of FA practice. Strikingly, while students’ ZPD is acknowledged as part of the formative assessment conceptual apparatus, teachers’ ZPD with FA practice itself is not.

We note that Wilson and Sloane (2000), similar to Shepard’s (2000) account, put teachers’ work in classroom assessment forward as key to conceptualizing educational assessment reform and deeper learning outcomes. Wilson and Sloane argue, along with other authentic assessment movements during the 1990s, that “embedded assessment” systems that connect teachers’ and students’ use of information during instruction can also improve outcomes. Wilson and Sloane (2000) wrote:

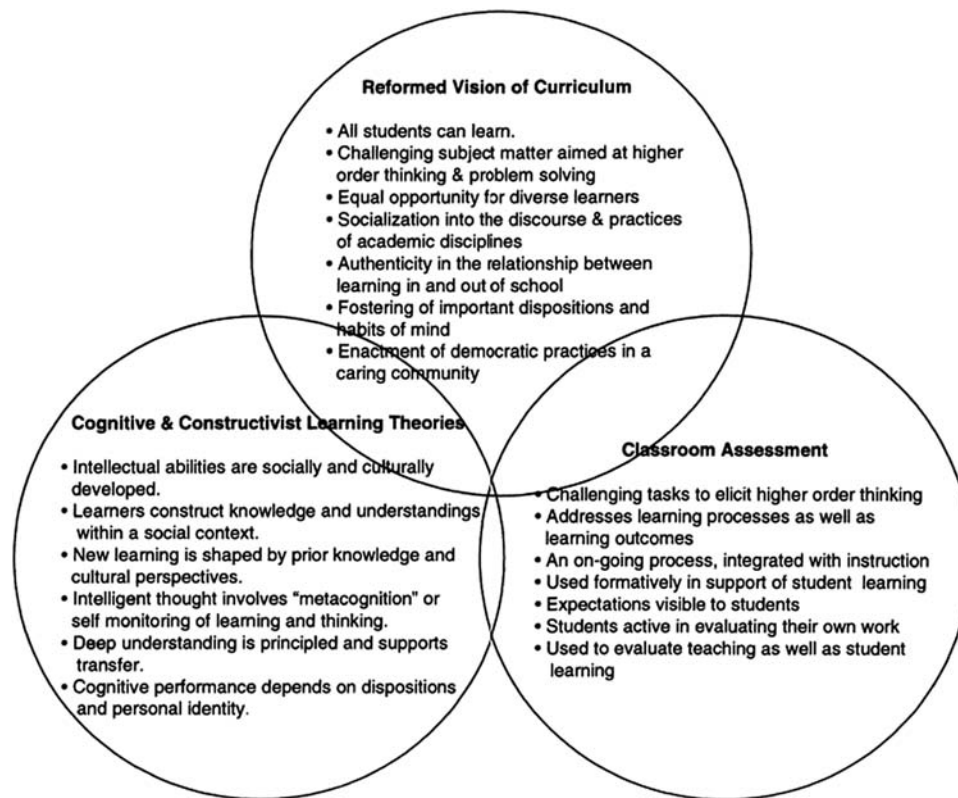


Fig. 3 Interlocking tenets of curriculum theory, psychological theories, and measurement theory characterizing the dominant 20th-century paradigm.

Explanations of the distinctions between alternative assessment procedures and more traditional forms of testing frequently emphasize the importance of teachers' roles in mediating and interpreting the alternative assessment results within the classroom context (e.g., Chittenden, 1991; Wolf et al., 1991) as well as the more immediate and meaningful uses of the alternative assessment procedures in the ongoing instructional process (Cole, 1991; Darling-Hammond and Ancess, 1996). However, it is also recognized that such integration will require new views of the teaching and learning process, new roles for (and demands on) teachers, or even a new "assessment culture" in the classroom (Brown et al., 1992; Cole, 1991; Resnick and Resnick, 1992; Torrance, 1995a,b; Zessoules and Gardner, 1991). Preparing teachers to use these types of assessments in their classroom teaching may be a difficult challenge. Teachers' understanding and acceptance of innovations are crucial to the ultimate success of change (Airasian, 1988; Stake, 1991a,b) (pp. 190–191) [Emphasis added].

While Wilson and Sloane do not propose explicit theories of teacher learning, the argument for formative uses of assessment data and protocols requires a developmental perspective on students (and teachers!) who are learning how to assist one another in such exchanges. Implicitly, these embedded assessment systems are expected to influence and shape teacher practice and therefore teacher learning and preparation. Nonetheless, while emphasizing students' growth and progress, teachers' trajectories as embedded assessors, as we might call them, remains in the background.

Recent scholarship from the *professional growth perspective* has leaned into the role of teacher development to better articulate the gaps between experts and novices in research on formative assessment practice at the classroom level. Wylie et al. (2016) use the language of "dimensions" that build upon continua of formative assessment practice.

- Learning goals
- Criteria for success
- Tasks and activities that elicit evidence of student learning
- Questioning strategies that elicit evidence of student learning
- Extending thinking during discourse
- Descriptive feedback
- Peer feedback
- Self-assessment
- Collaborative culture of learning
- Use of evidence to inform instruction

Utilizing principles of validation and consensus building across various stakeholders over pre-service, induction, and in-service years, these authors posited “ten dimensions” of FA practices and articulated four levels of accomplishment for each proposed continua.

To represent variation, that is, qualitative differences in teachers’ practice along the proposed Using the Formative Assessment Rubrics, Reflection and Observation Tools to Support Professional Reflection on Practice (Revised FARROP) model, rubrics have been developed to describe both the teacher role and the student role in a particular formative assessment “dimension.” For each rubric, there are four levels (plus a “not observed” indicator) that describe implementation of the practice labeled “beginning,” “developing,” “progressing,” and “extending.” The rubrics are intended to avoid judgment for a particular teacher’s level of expertise and indicate instead, the “level of implementation” of a particular aspect of practice.

Each level of a rubric provides two to five sentence-long descriptions to guide evaluation of levels of implementation of FA practice. For example, the “beginning level” of the “Feedback Loops During Questioning,” rubric includes two descriptors: “The teacher asks none or very few questions designed to encourage classroom discourse during the lesson,” OR “The teacher asks questions from students, but discourse focuses on a statement of correct or incorrect rather than deeper/meaningful exploration of ideas.” The “extending” level of this rubric has three descriptors:

- The teacher asks questions designed to encourage classroom discourse consistently throughout the lesson and integrates questioning and discussion seamlessly into instruction.
- The teacher and students consistently build on other students’ responses, clarify student comments, push for more elaborate answers, or engage more students in thinking about the problem.
- Extended feedback loops are used to support students’ elaboration and to have students contribute to extended conversations. Classroom discourse is characterized by the consistent use of feedback/probes that encourage deeper/more meaningful exploration of ideas.

The FARROP model was developed as a set of rubrics and observation protocols and “reflection tools” to focus attention on classroom practices in formative assessment. The commitment to supporting teachers’ professional learning, particularly through peer observation and reflection on teaching practice, explains the theory of differentiated FA practice. Rigorous study of a sample of teachers and therefore validation of the internal structure of these dimensions is not yet available. But the *professional growth over time* perspective shared by these FA researchers signals a more robust commitment to the role of teachers in developing their FA practice-based repertoires, more than previous research in the field.

Recent scholarship has leaned further into the empirical quantitative study of teacher development in classroom assessment knowledge and practices (Duckor et al., 2009, 2017). To better articulate the gaps in between experts and novices in research on formative assessment practice, Duckor and Holmberg (2019) use the language of “FA Moves” to test hypotheses about empirical trajectories and progressions of high-leverage formative assessment practices. Similar to the implicit growth modeling of the FARROP rubrics, this moves-based approach emphasizes the teacher’s development, learning, and transitions toward expertise in myriad formative assessment practices across 5 “developmentally appropriate” levels of practice.

Utilizing principles of evidence-centered design (Pellegrino et al., 2001) and the conceptual framework of learning progressions, these authors envision trajectories of high-leverage formative assessment practices. The theoretical framework of construct modeling (Wilson, 2005) places emphasis on moving from listing practices or facets of formative assessment toward the aim of articulating qualitative differences in the enactment of well-defined practices along a trajectory of moves.

Rather than seek to define an exhaustive list of “dimensions” or “facets,” Duckor and Holmberg’s research draws from consensus about the foundational role of questioning practices in FA (Duckor, 2014; Duckor and Holmberg, 2017) that are developmentally appropriate for novices and beginner teachers. Duckor and Holmberg (2019, p. 847) write:

Based on a review of the assessment for learning literature and professional teaching standards documents, questioning practices are integral to the practice of preservice and in-class FA across the P-12 curriculum (Applebee et al., 2003; Ball et al., 2009; Heritage and Heritage, 2013; Jacobs et al., 2010; Ruiz-Primo and Furtak, 2006; Smith and Stein, 2011; William, 2007). To better understand these foundational, generative questioning practices, we have decomposed them into three high-leverage moves: posing, pausing, and probing within the FA moves framework. Each construct map represents a variable in the framework.

Duckor and Holmberg developed the Posing-Pausing-Probing (PPP) construct model to examine how instructional and assessment moves can be woven together along empirically verifiable dimensions of practice. A recent quantitative study of these three distinct but related FA moves among pre-service teacher candidates indicated strong evidence for reliability and validity of score interpretations (Duckor et al., 2020a,b).

The FA moves-based approach in the PPP study articulates levels of practice along a well-defined dimension; it emphasizes the teacher’s development, learning, and transitions toward expertise in formative assessment practice without judgment from an idealized norm or standard. Empirical study of teacher artifacts (i.e., video of lesson enactments on standardized licensure teacher performance assessments) has led to valid and reliable representation of trajectories in three areas: “Posing” with a range of questions; “Pausing” for wait- and think-time for responses; and “Probing” on responses for elaboration.

In Duckor and Holmberg’s approach, there are 5 levels of differentiated practice for each construct map. These are titled “pre-structural/pre-probing,” “unistructural/emergent,” “multistructural/intentional,” “strategic/purposeful,” and “differentiated/

Probing			
More ↑ ↓ Less	<i>Respondents</i> whose probing is adaptive to evidence of individual and groups of students' present understandings in relation to the intended learning target, context, and curriculum. Can explain what they anticipate to happen as a result of probing and why. They communicate concern for responding productively to student responses they cannot anticipate and can explain strategies that support eliciting and leveraging such responses in relation to learning target. They enact probes with relevant knowledge of students and learning progressions in mind. They enact ways to support students probing each other's thinking and performances. They tend to enact lessons that reflect multifaceted purposes for probing and integration of other FA moves to support probing. They tend to enact lessons that encourage student independence and interdependence related to probing and the generation and use of formative feedback. Respondents probe equally well regardless of configuration—whole class, small group, one-on-one—and consistently probe in each of these configurations during lessons.	Differentiated "adaptive"	<i>Responses to tasks/items</i> indicate patterned and responsive probing that includes productive teacher responses to information elicited by teacher and/or student probing and incorporated into further probing. Responses indicate that probing includes focus on student self-regulation of learning and that formative feedback students receive or generate is tied to data elicited from probing and linked to learning target. Observation shows productive handling of unorthodox responses to probing. Probing enacted reflects detailed and relevant knowledge of students, context, and curriculum, and may suggest use of learning progressions. Observation demonstrates a productive balance amongst probing technique/strategy, students' affective states, and learning goals and integration of other FA moves to support probing. Observation reveals configuration (whole class, small group, one-on-one) does not influence quality of probing, which is consistently done throughout the lesson, regardless of configuration.
	<i>Respondents</i> who describe the purposes motivating probes/probing moves and attend to probing's role in generating formative feedback and in supporting student agency in efforts to progress toward intended learning target. They tend to enact lessons such that instructional decision making is contingent upon what is elicited from probing. They support student-to-student (S-2-S) probing strategically and integrate other FA moves with probing. They tend to enact probing that results in responses used by students and teacher to move students' learning forward in relation to the learning target. Respondents who use other FA moves (e.g., priming, pausing, and posing) to support probing. Respondents for whom episodes of whole class probing begin to reach the probing skill levels they demonstrate during small group and one-to-one configurations, though not consistently.	Strategic "purposeful"	<i>Responses to tasks/items</i> indicate teachers take up evidence of student performance in probing formulation or delivery. A variety of probing moves are demonstrated. Observation shows that student-to-student (S-2-S) probing occurs, supported by norms, routines, and scaffolds. Observation may show "extended episodes" of probing that are on topic and on task between teacher and students and students and students. Observation suggests attention is paid to using probing to improve formative feedback related to intended learning target that is available to students by teacher and students. Responses to task/items show other FA moves (e.g., priming, pausing, and posing) are used to support probing. Observation demonstrates episodes of whole class probing beginning to reach the probing skill levels demonstrated during small group and one-to-one configurations, though not consistently.
	<i>Respondents</i> who address probing's relationship to decision making and to probing's role in making understanding and thinking visible and why this is important. Probing's role in clarifying student understanding of intended learning target, fostering student responsibility for efforts toward the target, and informing formative feedback tends to be overshadowed by other teacher goals for probing (e.g., "uncovering misconceptions", checking understanding of academic language). They tend to enact lessons that include probing of "correct answers" and that encourage S-2-S probing. They tend to enact probing that targets either catalyzing movement toward the learning target or influencing student affect. They may be challenged to capitalize on what probing elicits. Respondents whose probing is noticeably better during one-on-one and small group configurations than it is during whole class instruction.	Multistructural "intentional"	<i>Responses to tasks/items</i> indicate probing is potentially valuable to teacher or student decision making. Probes target uncovering misconceptions. Observation shows teacher, and sometimes students, using what probes make visible in attempts to advance student performance related to learning target. Observation shows probing of "correct answers", probing to uncover misconceptions and/or checking understanding of academic language. Observation demonstrates support of S-2-S probing, although explicit connections to how S-2-S probing can support student understanding of the intended learning target and student ownership of efforts toward this target may be absent. Responses may indicate "teacher probing as formative feedback to students." Observation demonstrates teacher's quality of probing during one-on-one and small group configurations is much better than it is during whole class instruction.
	<i>Respondents</i> who contend the main purposes of probing are to spur student action and to make learners' thinking more visible. They tend to rely on generic probes (e.g., "Why?" or "Please explain."). They tend to enact lessons where probing is not explicitly tied to informing possible formative feedback and enact most of their probing in one component of the lesson. Some respondents enact lessons where the wording and/or pace of probing mirrors students' needs, but the orientation of the series of probes is still primarily from the teacher's point of view. Most and/or best quality probing occurs during one-on-one or small group configurations. Respondents who are able to probe student thinking during one-on-one conversations, but not leverage this to probing well with small groups of students or during whole class configuration. Respondents tend to complexify their probing before a student responds. Some respondents may repeat or revoice student responses as probing or proto-probing.	Unistructural "emergent"	<i>Responses to tasks/items</i> indicate probing relies on generic probing moves (e.g., "Why?" "Say more..." "What do you mean?") as probes that are beyond "probing to manage" or "probing to engage." Observation shows that some probes do make some learners' present thinking somewhat visible. Observation may show that what gets elicited via probing gets used by students or teacher, but also that probing is not explicitly tied to formative feedback. Observation often reveals probes reside in a narrow range. Some responses to task/items indicate that during teaching the wording and/or pace of probing mirrors students' needs, but still is oriented primarily to the teacher's thinking. Observation shows most and/or best quality probing happens with one student and occurs during one-on-one or small group configurations and not during whole class instruction. Observation shows teachers complexify probe(s) before students respond. Some responses to task/items indicate that the teacher repeats or voices students' words as proto-probing.
	<i>Respondents</i> for whom probing student thinking related to the learning target is not a primary concern in classroom assessment. They tend to enact lessons where student discourse is not rich and where "discussions," if they occur, exemplify "coverage and review," not processes supporting "uncovering" or "elaboration" of student responses. They do not probe student thinking regardless of configuration: whole class, small group, or one-on-one.	Prestructural "pre-probing"	<i>Responses to tasks/items</i> do not plausibly indicate that probing related to the learning target has occurred. Observations of teacher to student interactions in whole class, small group, and one-on-one configurations do not suggest probing.

Fig. 4 Construct map for identifying differentiated probing practices.

adaptive." In the example of the Probing construct, the authors carefully specify levels of progress along the proposed continuum of FA practice (Fig. 4).

The nature of the complexity of a construct map makes them (a) essential to the design of sound assessments for learning FA practices, (b) useful for instructional decision making (when instructional practices for learners are linked to vetted FA construct maps) during coursework and professional development modules, and, potentially, (c) helpful in efforts aiming to integrate site-based FA instruction with wider professional accountability in teacher licensure. As Wilson (2009, p. 3) and others note, construct maps are at once general enough to be interpretable within a particular curriculum—and possibly across curricula—and specific enough to guide the development of the curriculum task designs and scoring tools/guides for an evaluation and/or observational study design.

These new directions in FA research offer significant affordances for those interested in designing measures in preservice teacher assessment and inservice evaluation settings, where individuals may not share a common preparation curriculum or standardized professional development program. Moreover, the progress guides framework aligned with corresponding construct maps operate at a grain size intended for teacher-driven, peer and self-assessment in different contexts (See, e.g., Duckor & Holmberg, in press). In addition to aiding cooperating teachers and faculty as they observe practice, like the FARROP rubrics, these PPP guides aligned with construct maps support the process of formative feedback for, by, and among teachers working toward FA expertise.

Conclusion and future directions

Everything we know about teachers and teaching teachers tells us that they, like all professionals, are learners who are working toward mastery of a craft (Grossman, 1990; Grossman et al., 2009; Shulman, 1986, 1987). The subject who is learning about formative assessment is no different than the one learning about educational psychology, language acquisition and learning theory, culturally responsive and sustaining pedagogy or any other (dis)courses of study intended to deepen teacher preparation, support on-going in-service training and ultimately broaden professional practice.

Teachers are lifelong learners on different trajectories, much like their students; teachers, too, are apprentices, often learning with guidance and support from more knowledgeable others in the profession. Teachers' locations in particular preparation pathways, induction programs and professional development efforts are only part of a story of teacher growth. Assessment practices, routines, and procedures cut across many contexts for life-long learning. When it comes to developing expertise in formative assessment practices, we need to locate how these FA practices evolve, cohere, and ultimately become, as Vygotsky might note, internalized automatically within a learning environment.

Research on teachers and teaching has shown that formative assessment can improve student learning more than most instructional practices (Hattie, 2009, 2012). Empirical evidence spanning nearly three decades indicates that thoughtfully implemented formative assessment practices improve students' learning, yet little research work has been done on the actual development of teachers-as-formative-assessors as they learn to take up high-leverage practices in the domain of classroom-based assessment for learning across myriad subject disciplines in K-12 education.

New developments in FA research are exploring implications of teachers becoming formative assessors from pre- and in-service classroom perspectives, in part, by centering research on how high-leverage instructional and assessment moves can be modeled for study. The concept of "dimensions" or "levels" and "facets" in formative assessment requires technical evidence and must rest upon a developmental foundation. When one approaches how to model teachers' learning progressions or hypothesize trajectories of practices in the domain of formative assessment, qualitative study of "aspects" "moves" and "strategies" are at the root of deeper inquiry. Theory can only be built from the firm foundation of observational study and lived experience in the FA realm across the spectrum of teacher-learners—from those engaged in novice to expert practices, and all those emerging in between.

Stake (1967) presciently noted that, while evolving over nearly over a half century, the evaluation and by implication the assessment field has blindspots. Teachers' work, which includes their own learning and development, has always been the black box for those who fail to appreciate contingencies. Stake wrote:

Lesson planning and curriculum revision through the years has been built upon faith in certain contingencies. Day to day, the master teacher arranges his presentation and selects his input materials to fit his instructional goals. For him the contingencies, in the main, are logical, intuitive, and supported by a history of satisfactions and endorsements. Even the master teacher and certainly less-experienced teachers need to bring their intuited contingencies under the scrutiny of appropriate juries (pp. 528–529).

The question before the new assessment theorists and for those who study teachers in their full range of capacities, dispositions, orientations, and most importantly, personal trajectories of practice over a professional career is: Can we study and honor contingency while hypothesizing growth and respecting the teacher-as-learner of effective classroom assessment practices? Our teachers' need for visible scaffolds, structured feedback, and focused attention on "just-in-time" goals is ubiquitous and necessary. What is wanted for students is needed for teachers and vice-versa.

The lived, evolving, often messy practice of formative assessment—for students and teachers alike—has been largely ignored in the search for expertise and exemplars of formative assessment in the classroom over the past three decades. Without conceptual frameworks rooted in learning theory for teachers and students, the field labors to define accessible pathways, to identify proximal trajectories, and to empirically instantiate visible progress markers that honor incremental success. It may have been convenient for a season to remain agnostic about "how teachers learn" in the domain of formative assessment but that privileged perspective is untenable for those committed to novices, beginners, and most all of those teachers working to become formative assessors in today's culturally, ethnically, economically, and linguistically diverse learning environments.

References

- Airasian, P.W., 1988. Measurement-driven instruction: a closer look. *Educ. Meas.* 7, 6–11.
- American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME), 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Applebee, A.N., Langer, J.A., Nystrand, M., Gamoran, A., 2003. Discussion-based approaches to developing understanding: classroom instruction and student performance in middle and high school English. *Am. Educ. Res. J.* 40, 685–730.
- Ball, D.L., Sleep, L., Boerst, T.A., Bass, H., 2009. Combining the development of practice and the practice of development in teacher education. *Elem. Sch. J.* 109, 458–474.
- Bell, B., Cowie, B., 1997. Formative assessment and science education. In: Research Report of the Learning in Science Project (Assessment). Hamilton, New Zealand: Learning in Science Project (Assessment). University of Waikato.
- Bell, B., Cowie, B., 2002. *Formative Assessment and Science Education*. Kluwer Academic Publishers, New York.
- Bensman, D., 2000. *Central Park East and Its Graduates: Learning by Heart*. Teachers College Press, New York.
- Black, P., William, D., 1998a. Assessment and classroom learning. *Assess. Educ.* 5, 7–74.
- Black, P., William, D., 1998b. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan* 80, 139–148.
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21, 5–31.
- Bloom, B.S., Hastings, J.T., Madaus, G.F. (Eds.), 1971. *Handbook of Formative and Summative Evaluation of Student Learning*. McGraw-Hill, New York, NY.
- Borko, H., Livingston, C., 1989. Cognition and improvisation: differences in mathematics instruction by expert and novice teachers. *Am. Educ. Res. J.* 26, 473–498.
- Brown, A.L., Campione, J.C., Webber, L.S., McGilly, K., 1992. Interactive learning environments: a new look at assessment and instruction. In: Gifford, B.R., O'Connor, M.C. (Eds.), *Changing Assessments*. Kluwer Academic, Boston, pp. 121–212.
- Chittenden, E., 1991. Authentic assessment, evaluation, and documentation of student performance. In: Perrone, V. (Ed.), *Expanding Student Assessment*. Association for Supervision and Curriculum Development, Alexandria, VA, pp. 22–31.

- Cizek, G.J., 2010. An introduction to formative assessment: history, characteristics, and challenges. In: Andrade, H.L., Cizek, G.J. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, pp. 3–17.
- Cole, N., 1991. The impact of science assessment on classroom practice. In: Kulm, G., Malcom, S. (Eds.), *Science Assessment in the Service of Reform*. American Association for the Advancement of Science, Washington, DC, pp. 97–106.
- Council of Chief State School Officers [CCSSO], 2008. Attributes of effective formative assessment. In: A Work Product Coordinated and Led by Sarah McManus, North Carolina Department of Public Instruction, for the Formative Assessment for Students and Teachers (FAST) Collaborative. Council of Chief State School Officers, Washington, DC.
- Council of Chief State School Officers [CCSSO], 2018. *Revising the Definition of Formative Assessment*. Council of Chief State School Officers, Washington, DC.
- Cowie, B., Bell, B., 1999. A model of formative assessment in science education. *Assess. Educ. Princ. Pol. Pract.* 6, 101–116.
- Darling-Hammond, L., Ancess, J., 1996. Authentic assessment and school development. In: Baron, J.B., Wolf, D.P. (Eds.), *Performance-Based Student Assessment: Challenges and Possibilities*. *Ninety-Fifth Yearbook of the National Society for the Study of Education*. University of Chicago Press, Chicago, pp. 52–83.
- Darling-Hammond, L., Ancess, J., Falk, B., 1995. *Authentic Assessment in Action: Studies of Schools and Students at Work*. Teachers College Press, New York.
- Darling-Hammond, L., 1997. *The Right to Learn: A Blueprint for Creating Schools That Work*. Jossey-Bass, San Francisco.
- Duckor, B., Holmberg, C., 2017. *Mastering Formative Assessment Moves: 7 High-Leverage Practices to Advance Student Learning*. ASCD, Alexandria, VA.
- Duckor, B., Holmberg, C., 2019. Exploring how to model formative assessment trajectories of posing-pausing-probing practices: toward a teacher learning progressions framework for the study of novice teachers. *J. Educ. Meas.* 56, 836–890.
- Duckor, B., Holmberg, C., (in press). Feedback for continuous improvement in the classroom: New perspectives, possibilities, and practices. Thousand Oaks, CA, Corwin.
- Duckor, B., Perlstein, D., 2014. Assessing habits of mind: teaching to the test at central Park East secondary school. *Teach. Coll. Rec.* 116, 1–33.
- Duckor, B., Draney, K., Wilson, M., 2009. Measuring measuring: toward a theory of proficiency with the constructing measures framework. *J. Appl. Meas.* 10, 296–319.
- Duckor, B., Draney, K., Wilson, M., 2017. Assessing assessment literacy: an item response modeling approach for teacher educators. *Pensamiento Educativo* 54, 1–25.
- Duckor, B., Holmberg, C., Patthoff, A., Téllez, K., 2020a. Understanding Preservice Teachers' Formative Feedback Practices in Elementary, Middle, and High School Classrooms in High Needs Contexts. University of California California Teacher Education Research and Improvement Network (CTERIN). Retrieved from: <https://cterin.ucop.edu/resources/publications/focusvol1no3.html>.
- Duckor, B., Holmberg, C., Zhang, Y., 2020b. Novice Study of Teacher Learning Progressions in Posing, Pausing, and Probing Practices: A Multi-dimensional IRT Analysis. Invited speaker at the Berkeley Evaluation and Assessment Research Center, UC Berkeley, CA.
- Duckor, B., 2014. Formative assessment in seven good moves. *Educ. Leader.* 71, 28–32.
- Feiman-Nemser, S., 2001a. From preparation to practice: designing a continuum to strengthen and sustain practice. *Teach. Coll. Rec.* 103, 1013–1055.
- Feiman-Nemser, S., 2001b. Helping novices learn to teach: lessons from an exemplary support teacher. *J. Teach. Educ.* 52, 17–30.
- Furtak, E.M., Kiemer, K., Circi, R.K., Swanson, R., de León, V., Morrison, D., Heredia, S.C., 2016. Teachers' formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instr. Sci.* 44, 267–291.
- Gotch, C.M., French, B., 2014. A systematic review of assessment literacy measures. *Educ. Meas.* 33, 14–18.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., Williamson, P., 2009. Teaching practice: a cross-professional perspective. *Teach. Coll. Rec.* 111, 2055–2100.
- Grossman, P., 1990. *The Making of a Teacher: Teacher Knowledge and Teacher Education*. Teachers College Press, New York.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Hattie, J., 2009. *Visible Learning: A Synthesis of over 800 Meta-analyses Relating to Achievement*. Routledge, New York, NY.
- Hattie, J., 2012. *Visible Learning for Teachers: Maximizing Impact on Learning*. Routledge, New York, NY.
- Heritage, M., Heritage, J., 2013. Teacher questioning: the epicenter of instruction and assessment. *Appl. Meas. Educ.* 26, 176–190.
- Heritage, M., Wylie, E.C., 2020. *Formative Assessment in the Disciplines: Framing a Continuum of Professional Learning*. Harvard Education Press, Cambridge, MA.
- Heritage, M., 2008. *Learning Progressions: Supporting Instruction and Formative Assessment*. The Council of Chief State School Officers, Washington, DC. Retrieved from: <https://www.michiganassessmentconsortium.org/wp-content/uploads/Learning-Progressions.pdf>.
- Heritage, M., 2010. *Formative Assessment: Making It Happen in the Classroom*. Corwin, Thousand Oaks, CA.
- Hollins, E.R., King, J.E., Hayman, W.C. (Eds.), 1994. *Teaching Diverse Populations: Formulating a Knowledge Base*, vol. 1944. SUNY Press, New York.
- Jacobs, V.R., Lamb, L., Philipp, R.A., 2010. Professional noticing of children's mathematical thinking. *J. Res. Math. Educ.* 41, 169–202.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491.
- Leung, C., 2007. Dynamic assessment: assessment for and as teaching. *Lang. Assess. Q.* 4, 257–278.
- McMillan, J.H., 2010. The practical implications of educational aims and contexts for formative assessment. In: Andrade, H.L., Cizek, G.J. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, pp. 41–58.
- Meier, D., 1995. *The Power of Their Ideas: Lessons for America From a Small School in Harlem*. Beacon, Boston.
- Moss, C.M., Brookhart, S.M., 2019. *Advancing Formative Assessment in Every Classroom: A Guide for Instructional Leaders*, second ed. ASCD, Alexandria, VA.
- New York State Council on Curriculum and Assessment, 1994. *Learning-Centered Curriculum and Assessment for New York State*. New York State Education Department, Albany.
- No Child Left behind Act of 2001, P.L. 107-110, 20 U.S.C., 2002.
- Pellegrino, J.W., Chudowsky, N., Glaser, R. (Eds.), 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academies Press, Washington, DC, pp. 1–382.
- Penuel, W.R., Shepard, L.A., 2016. Assessment and teaching. In: Bell, C., Gitomer, D. (Eds.), *Handbook of Research on Teaching*. American Educational Research Association, Washington, DC, pp. 787–850.
- Perrenoud, P., 1998. From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assess. Educ. Princ. Pol. Pract.* 5, 85–102.
- Poehner, M.E., Lantolf, J., 2005. Dynamic assessment in the language classroom. *Lang. Teach. Res.* 9, 233–265.
- Putnam, R.T., Borko, H., 2000. What do new views of knowledge and thinking have to say about research on teacher learning? *Educ. Res.* 29, 4–15.
- Resnick, L.B., Resnick, D.P., 1992. Assessing the thinking curriculum: new tools for educational reform. In: Gifford, B.R., O'Connor, M.C. (Eds.), *Changing Assessments*. Kluwer Academic, Boston, pp. 37–76.
- Ruiz-Primo, A.M., Furtak, E.M., 2006. Informal formative assessment and scientific inquiry: exploring teachers' practices and student learning. *Educ. Assess.* 11, 237–263.
- Ruiz-Primo, A.M., Furtak, E.M., 2007. Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *J. Res. Sci. Teach.* 44, 57–84.
- Sadler, R., 1989. Formative assessment and the design of instructional assessments. *Instr. Sci.* 18, 119–144.
- Scriven, M., 1966. The methodology of evaluation: research report #110. Lafayette. In: Purdue University Social Science Education Consortium.
- Scriven, M., 1973. The methodology of evaluation. In: Worthen, B.R., Sanders, J.R. (Eds.), *Educational Evaluation: Theory and Practice*. Charles A. Jones, Worthington, OH.
- Shepard, L.A., 2000. The role of assessment in a learning culture. *Educ. Res.* 29, 4–14.
- Shepard, L.A., 2005. Linking formative assessment to scaffolding. *Educ. Leader.* 63, 66–70.
- Shepard, L.A., 2008. Formative assessment: caveat emptor. In: Dwyer, C.A. (Ed.), *The Future of Assessment: Shaping Teaching and Learning*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 279–303.
- Shulman, L., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14.
- Shulman, L., 1987. Knowledge and teaching: foundations of the new reform. *Harv. Educ. Rev.* 57, 1–23.
- Sizer, T., 1996. *Horace's Hope: What Works for the American High School*. Houghton Mifflin Co, Boston, MA.
- Smith, M.S., Stein, M.K., 2011. *5 Practices for Orchestrating Productive Mathematics Discussions*. National Council of Teachers of Mathematics, Reston, VA.

- Stake, R.E., 1967. The countenance of educational evaluation. *Teach. Coll. Rec.* 68, 523–530.
- Stake, R.E., 1991a. Advances in Program Evaluation: Volume 1, Part A. Using Assessment Policy to Reform Education. JAI, Greenwich, CT.
- Stake, R.E., 1991b. The teacher, standardized testing, and prospects of revolution. *Phi Delta Kappan* 71, 243–247.
- Stiggins, R.J., 2002. Assessment crisis: the absence of assessment for learning. *Phi Delta Kappan* 83, 758–765.
- TeachingWorks, 2018. High-Leverage Practices. Teaching Works, University of Michigan, Ann Arbor. Retrieved from. <http://www.teachingworks.org/work-of-teaching/high-leverage-practices>.
- Torrance, H., 1995a. The role of assessment in educational reform. In: Torrance, H. (Ed.), *Evaluating Authentic Assessment*. Open University Press, Philadelphia, pp. 144–156.
- Torrance, H., 1995b. Teacher involvement in new approaches to assessment. In: Torrance, H. (Ed.), *Evaluating Authentic Assessment*. Open University Press, Philadelphia, pp. 44–56.
- Wiggins, G., 1993. *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. Jossey-Bass, San Francisco, CA.
- William, D., Thompson, M., 2008. Integrating assessment with learning: what will it take to make it work? In: Dwyer, C.A. (Ed.), *The Future of Assessment: Shaping Teaching and Learning*. Routledge, New York, NY, pp. 53–82.
- William, D., 2007. Keeping learning on track: classroom assessment and the regulation of learning. In: Lester Jr., F.K. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, vol. 2. Information Age Publishing, Charlotte, NC, pp. 1053–1098.
- William, D., 2010. An integrative summary of the research literature and implications for a new theory of formative assessment. In: Andrade, H.L., Cizek, G.J. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, pp. 18–40.
- William, D., 2011. What is assessment for learning? *Stud. Educ. Eval.* 37, 3–14.
- Wilson, M., Sloane, K., 2000. From principles to practice: an embedded assessment system. *Appl. Meas. Educ.* 13, 181–208.
- Wilson, M., 2005. *Constructing Measures*. Lawrence Erlbaum, Mahwah, NJ.
- Wilson, M., 2009. Measuring progressions: assessment structures underlying a learning progression. *J. Res. Sci. Teach.* 46, 716–730.
- Wolf, D., Bixby, J., Glenn III, J., Gardner, H., 1991. To use their minds well: investigating new forms of student assessment. *Rev. Res. Educ.* 17, 31–74.
- Wylie, C., Lyon, C., Formative Assessment for Students and Teachers (FAST), State Collaborative on Assessment and Student Standards (SCASS), 2016. *Using the for- Mative Assessment Rubrics, Reflection and Observation Tools to Support Professional Reflection on Practice (Revised)*. Council of Chief State School Officers, Washington, DC.
- Wylie, C., Bauer, M., Bailey, A.L., Heritage, M., 2018. Formative assessment of mathematics and language: applying companion learning progressions to reveal greater insights to teachers. In: Bailey, A.L., Maher, C.A., Wilkinson, L.C. (Eds.), *Language, Literacy, and Learning in the STEM Disciplines*. Routledge, New York, pp. 143–168.
- Zessoules, R., Gardner, H., 1991. Authentic assessment: beyond the buzzword and into the classroom. In: Perrone, V. (Ed.), *Expanding Student Assessment*. Association for Supervision and Curriculum Development, Alexandria, VA, pp. 47–71.

Conceptualizing professional development with and for a diverse and globally mobile early childhood education workforce

Joce Nuttall^a, Susan Grieshaber^b, Sirene Lim^c, and Eunju Yun^d, ^a Australian Catholic University, Melbourne, VIC, Australia; ^b La Trobe University, Melbourne, VIC, Australia; ^c Singapore University of Social Sciences, Singapore, Republic of Singapore; and ^d Sookmyung Women's University, Seoul, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Introduction	596
How professional development is understood in early childhood education	596
The early childhood education professional development research literature	597
Researching the relationship between professional development and program quality	598
Diversity and global mobility in the early childhood education workforce	600
Supporting the learning of a diverse and globally mobile early childhood education workforce	600
Conclusion	601
References	601

Introduction

This chapter contributes to this volume on teachers' lives, work, and professional preparation in two ways. First, it provides a brief summary of the contemporary literature concerned with research and practice in professional development in early childhood education. Second, it offers a critique of this literature, specifically in the context of a globally mobile early childhood education workforce. Such critique is essential if educational institutions serving our youngest citizens are to recognize and value the diverse knowledges and cultural forms of *all* children and their families. In recent decades the emphasis of early childhood education policy around the world has turned to raising the quality of programs for children and their families, with attendant policy frameworks and professional development opportunities. Our overall aim is to argue for a thorough conceptualization of the nexus between professional development and quality in early childhood education in the context of a globalized and highly diverse early childhood workforce. We base this argument on a critique of the ways in which professional development has been conceptualized to date in early childhood education research and practice, and the failure of professional development policies, programs, and research to adequately address the diversity of the early childhood education workforce. We propose that addressing this gap is likely to benefit all professionals in early childhood education and the communities they serve.

We begin by explaining how the term "professional development" has typically been understood as a direct process of acquiring and developing a universal set of knowledge and skills competencies required to produce a desired set of outcomes in the early childhood education field. It is only since around 2015 that the empirical literature on early childhood educators' professional development has made a serious attempt to move away from the quest to identify effective professional development principles and program designs, and toward a more robust conceptualization of professional development as being anchored in educators' *learning*. We argue this is a necessary shift, although the number of studies since 2015 that attempt to conceptualize and theorize professional development in early childhood education still remains small. We provide evidence for this argument by examining some of the research into early childhood educators' professional development published since 2010, with a particular emphasis on professional development initiatives linked to program quality improvement. This focus and timespan are both deliberate. As we note above, many countries are now investing heavily in building the capacity of their early childhood education workforces with the aim of increasing the quality of care and education programs. Much of the literature in the decade since 2010 is, therefore, tightly situated in the relationship between professional development and concepts of quality in the provision of early childhood education and care.

We then turn to the way in which early childhood educators form part of the culturally diverse, globally mobile, and overwhelmingly female workforce that comprises the caring professions. Yet the expertise, needs, and aspirations of these early childhood educators are curiously absent from the professional development literature. In the final part of the chapter we draw together our threads of professional development, early childhood program quality, educator diversity, and workforce mobility to argue the need for new and robust theoretical underpinnings to inform the professional development opportunities necessary for all educators—migrant educators and their local colleagues alike—in the coming decades.

How professional development is understood in early childhood education

Systematic programs of professional development are a relatively new phenomenon in early childhood education internationally. In part, this is because work is still being done in many countries to establish a workforce that holds even baseline qualifications. In

schools around the world, teachers typically hold a minimum locally required credential prior to their classroom employment, but this is not the case in early childhood education and care (OECD, 2019). In Singapore, for instance, even though minimal professional qualifications are stipulated for its largely privatized childcare and kindergarten sector—10 years of formal education and an accredited professional certificate-level or diploma-level course—these expectations are significantly less compared to the undergraduate degree requirement imposed on primary school teachers in the public school system. And in Singapore, as well as in other so-called advanced economies such as Australia, concerted efforts are still needed to recruit, educate, and retain a suitably qualified workforce (Australian Children's Education and Care Quality Authority [ACECQA], 2021). We suggest this lack of a universal qualifications base is one reason why discourses of professional development in early childhood education are also less developed when compared with the schools sector.

As evidence of this relative lack of maturity, we point to a problematic feature of attempts to conceptualize professional development in early childhood education, both in research and in practice: the way professional *development* has been defined as professional development *activities* rather than as the *learning* that is necessary to develop practice. For example, in the US the National Association for the Education of Young Children [NAEYC] defines professional development as "... a continuum of learning and support *activities* designed to prepare individuals for work with and on behalf of young children and their families, as well as ongoing *experiences* to enhance this work" (NAEYC, 2022, n. p., emphasis added). Over a decade ago, Buysse et al. (2009) argued that the early childhood field needed to develop authoritative definitions of what constitutes professional development and that, without these, the field lacked "a common vision for the most effective ways of organizing and implementing professional development to improve the quality of the early childhood workforce" (pp. 235–6). However, in their attempts to define professional development, Buysse, et al. also misrecognize professional development activities as development itself:

The term *professional development* encompasses all types of facilitated learning opportunities, for example, those that result in college credit or degrees as well as those that generally are less intensive and do not yield credits or degrees, those that occur largely through formal coursework, and those that are more informal and situated in practice.

Buysse et al. (2009, p. 238, emphasis in original).

Within the same decade, Winton et al. (2016) proposed rethinking professional development in early childhood education as *learning*, in the light of challenges such as a lack of common definition about what constituted professional development and, consequently, the lack of coherent professional development systems and practices within a fragmented early childhood education landscape in the US. These authors believed that a unifying definition could promote clarity and alignment, proposing a broad definition of professional development as "facilitated teaching and learning experiences that are transactional and designed to support the acquisition of professional knowledge, skills and dispositions as well as the application of this knowledge in practice (Winton et al., 2008, p. 3)". In their 2018 meta-analysis of the relationship between professional development, quality ratings, and child outcomes, Egert, Fukkink, and Eckhardt also hint at a potential shift toward conceptualizing professional development as learning in early childhood education. They argue that "teacher outcomes" of effective professional development should comprise *awareness, orientations, knowledge, competencies, and performance* (Egert et al., 2018, p. 405). However, their analysis of professional development research suggests that evidence of the ways in which educators actually learn awareness, orientations, knowledge, competencies, and performance through professional development activities is still a "missing link" in the early childhood field. In the absence of a robust conceptualization of professional development, professional development research in early childhood education still largely relies on an assumption of a teleological (though not causal) link between professional development interventions (activities) and children's learning (outcomes), often via the category of "improved early childhood program quality." This assumption, in turn, is likely to be a consequence of policy settings (with attendant research funding) that focus on child outcomes, without first conceptualizing exactly *how* educators learn.

For the purposes of this chapter, we propose a definition of early childhood professional development that we believe has the potential to guide its conceptualization as *the learning that necessarily precedes development*. We suggest: *early childhood individual and collective professional development is a consequence of learning that fosters changes in educators' practical, cultural, and psychological grasp of their work, in ways that enhance their lives as well as the lives of children and families*. We return to this definition later in the chapter alongside calls by other scholars for the identification of promising concepts to explain early childhood educators' development and guide the activities that can foster the learning that necessarily precedes such development. Before doing so, we turn to a brief precis of the early childhood professional development research literature.

The early childhood education professional development research literature

Both the scope and scale of the professional development research literature in early childhood education are now considerable—in an initial search for their meta-synthesis of this literature, Brown and Englehardt (2016) found 974 peer-reviewed qualitative studies in the US alone—so a thorough review of the international literature is beyond the scope of this chapter. However, broadly speaking, this literature has a strong focus on evaluation of professional development program effectiveness, often with the aim

of identifying principles and practices to support effectiveness. This is unsurprising given the recency of regular professional development as an established phenomenon in the early childhood field. Since 2015 several scholars have capitalized on the size of the effectiveness literature to undertake meta-analysis and systematic review. In addition to Brown and Englehardt's (2016) meta-synthesis of qualitative studies in the US, Jensen and Iannone (2018) have undertaken a comparative review across ten EU member countries; Schachter et al. (2019) have reviewed US studies of effective professional development programs to identify their distinguishing features; and Peleman et al. (2018) have conducted a systematic review of primary research studies across 28 EU member nations, also with a view to identifying the features of successful programs.

Although the major analyses of the literature we cite above are confined to the US and EU, the geographical scope of the effectiveness literature is broad including, for example, Wolf et al. (2019) evaluation of the *Quality Preschool for Ghana* professional development program; He and Ho's (2013) study of the policy|teacher development interface in Hong Kong; Mendive et al. (2016) study of intervention fidelity in an early literacy professional development program in Chile; Nolan and Molla's (2018) study of coaching effectiveness in Australia; and Machado's (2019) study of practice development in a center in Portugal. Data generation and analysis methods are similarly wide in scope across these studies. These include case accounts (e.g. Machado, 2019; Inglis and Brown, 2013); video analysis (e.g. Harkoma et al., 2021; Mendive et al., 2016); interviews and focus groups (e.g. Nolan and Molla, 2018); standardized instruments for measuring teacher-child interaction (e.g. Torres et al., 2021; Eadie et al., 2019); and mixed methods (e.g. Linder et al., 2016). A further rich aspect of this literature is the wide range of content foci across the reported professional development programs. These foci comprise the whole range of entry points through which to understand early childhood education, including individual curriculum domains (e.g. literacy, mathematics, science); work with specific age groups (e.g. infants and toddlers; transition to school); approaches to assessment and documentation; work with families; and leadership and management.

Despite the heterogeneity of this large corpus of effectiveness studies, it is possible to identify two distinctive sub-sets within this literature. The first relates to evidence of shifts in the quality of educators' interactions with children following professional development interventions. Again, this focus is unsurprising given the intensely relational nature of teaching young children, the persuasive evidence for the relationship between classroom interaction quality and children's learning, and the interest taken by psychologists as well as educationalists in children's language development. This sub-set frequently overlaps with the second sub-set, which examines the relationship between the professional development of educators and consequential shifts in program quality. As stated at the outset, it is this relationship that we seek to critique in this chapter in the context of the diversity and global mobility of the early childhood education workforce.

Researching the relationship between professional development and program quality

Recent decades have seen a sociohistorical era in which teachers across all sectors of education and in many countries have been positioned as instruments of policy implementation who are directly accountable for children's learning outcomes, rather than as learners in their own right. With reference to the schools sector, Down (2012) calls this positioning "technical instrumentalism" (p. 65), and Angus (2012) attributes this positioning to "the current economic framing of education, in which the education policy framework has become narrow and instrumental, and principals, teachers and, most importantly, students, in keeping with a technical/managerial perspective, are treated as objects of policy" (2012, p. 50). It is not surprising, therefore, that a second distinctive sub-set of the early childhood professional development literature examines the effectiveness of professional development interventions in the context of policy mandates to improve the quality of early childhood programs.

The studies in this "professional-development-for-quality" sub-set typically invoke specific national and supra-national policies. For example, Jensen and Iannone (2018) reviewed early childhood professional development research across ten EU countries, focusing on the concept of early childhood system innovation and its relationship to quality outcomes, in response to the EU's *Europe 2020, a strategy for smart, sustainable and inclusive growth* (European Commission, 2010). Similarly, Peleman et al. (2018) made a systematic review of primary research studies across 28 EU countries, finding that long-term interventions that are well-integrated into center practice are effective, even in national early childhood policy contexts where services are poorly subsidized and have low rates of educator qualifications. Both of these EU studies point to cross-national policies in Europe as a driver of research into the relationship between professional development and early childhood program quality. In Australia, Barber et al. (2014) analyzed professional development opportunities and foci relative to Australia's *National Quality Standard* [NQS] (2012/2018), the policy statement against which mandatory assessment of center quality is made by Australian governments. Even when studies in this sub-set are anchored in state-level policies, such as Ton-yan et al. (2017) study of family childcare providers' engagement in quality improvement initiatives in California, supra-national policies and the influential research studies that have shaped these policies are often cited. These influential studies include the economic analyses of Heckman and colleagues (Heckman, 2011; Campbell et al., 2014), work on teacher-child interaction as a key variable in center quality (e.g. Sylva et al., 2004), and long-standing insights into the relationship between early childhood education participation and children's subsequent social and academic achievement (e.g. Schweinhart and Weikart, 1997).

Many national early childhood policies also seek to address social inequalities, including the overlap between inequality and cultural diversity. In Singapore, for example, the recognition that culturally diverse children have greater academic success if they retain their home language, as well as speaking the language of instruction, is embedded in education policy. This is reflected in early childhood settings through provision of educators who can teach in children's "mother tongue" (Ng, 2014). Pathways for educators to temporarily migrate from China and elsewhere to provide early childhood education in home languages are correspondingly built into Singapore's migration policies. In the EU context, Jensen and Iannone (2018) also point to why understanding children's diverse cultural contexts is important:

One important argument for considering CPD is to strengthen professionals beyond their pre-service training in their ongoing work towards high quality ECEC (Zaslow, 2009). Another addresses professionals' need to take an active part in creating and implementing culturally-specific goals for ECEC in ways that take into account different contexts of child development, including a wider context of society, reflecting the views of children as active learners (UN General Assembly, 1989), the values of families, later school requirements and the interests of society (European Commission, 2010), also in terms of social inequality and issues regarding the vulnerable (p. 24).

However, in proposing these arguments Jensen and Iannone (op. cit.) repeat the teleological assumption at the heart of much contemporary international early childhood education policy (and therefore at the heart of many accompanying professional development initiatives): that the purpose of professional development of educators is to improve their knowledge, skills, and attitudes *so that* early childhood education services improve. We suggest this is a necessary—but insufficient—rationale for the professional development of early childhood educators. Not only does this rationale fail to take early childhood educators into account as learners in their own right, it also neglects the rich resources that educators bring to the professional development process.

It also assumes that educator learning involves a technicist "input-output process," a view which has already been challenged in research outside of early childhood education in studies examining novice and experienced schoolteachers' sense-making (e.g. Clarke and Hollingsworth, 2002), teachers' emotions (e.g. Zembylas, 2011), and teacher agency (e.g. Vähäsantanen et al., 2008). Teachers' sense-making has been defined as an interaction between teachers' understandings of situational demands and their personal frames of reference, often intertwined with their feelings of being recognized or valued. Emotions are yet another fluid dimension of teaching and learning that has been found to be highly influential in teachers' professional learning, with consequences for their practices and interpretations of school reform and their understandings of power structures in schools and society (Zembylas, 2011).

Educators are notably absent from the professional-development-for-quality-improvement literature, except as sources of research data. In other words, in treating activities and outcomes as the main units of analysis for research into professional development, the professional expertise of the people who are supposed to be "developing" is rendered largely invisible. Brown and Englehardt (2016) have likewise argued that research into early childhood professional development has paid insufficient attention to educators and how educators make sense of the context-specific features for their work:

While important and informative, this empirical focus on what changes teachers' practices and/or improves children's development and learning provides limited insight into ECE PD. Missing from these analyses is an understanding of the nature of the PD sessions themselves, the teachers' experiences within these PD programs, the contexts in which the teachers work, or the impact of these PD programs on teachers' conceptions of instruction or themselves as teachers (p. 217).

We concur with this claim, but suggest there is a further weakness in the early childhood professional development literature that cannot be addressed simply by incorporating the lived experiences of educators. Background theoretical concepts related to learning and development are also absent. As we noted earlier, professional development, at least as it is defined by authoritative sources such as the NAEYC, by Buysse et al. (2009) and by Egert et al. (2018), is comprised of inputs (activities, opportunities, experiences) and outcomes (either for children or as empirical measures of program quality that stand as proxies for child outcomes). A shortcoming of this definition is that it neither employs nor develops conceptual frameworks with the potential to explain *what* professional development is in early childhood education. The identification of principles and practices of program design that has characterized research to date fails to explain how early childhood educators *learn in order to develop* in the specific contexts of early childhood education services. We agree with Brown and Englehardt (2016) when they say,

What appears to be missing from this literature is a clear understanding of how researchers frame the knowledge and skills of practicing early educators, what sets of knowledge and skills researchers think are most important for improving early educators' teaching, and what researchers see as being the best practices for delivering these sets of knowledge and skills to practicing early educators (p. 220).

Without the incorporation of strong conceptual frameworks and background theories of learning, claims about principles and practices will continue to be derived from the features of existing programs (e.g. the length of interventions; the choice of intervention methodologies, such as action research; individual vs. center-based programs; etc.). This approach presents the serious risk that claims about how effective programs *should be* designed are derived from how programs are *already* designed. In the final section of this chapter, we argue for inclusion of concepts drawn from learning and development within and across diverse cultures to address the need for a strong conceptual basis for early childhood educator professional development in the interests of program quality for children and families. Before doing so, we contextualize this argument by describing the globally mobility and diversity of the early childhood education workforce.

Diversity and global mobility in the early childhood education workforce

Early childhood educators are one of three main groups that form the international care workforce (the others being aged care workers and personal care workers). A fourth major group of similarly gendered (female) migrant workers is comprised of domestic workers in private homes. As Sheehan (2017) explains, “In light of growing international demand for child and senior care givers and the ongoing entry of women in developed nations into the professional labor market, domestic labor markers often look toward global labor pools to fill positions” (p. 314). In early childhood education, this reorientation has also been marked by a shift across time away from traditional sources of migrant educators by developed nations. In Australia before 1976, for example, 48% of Australia’s migrant care workforce was drawn from the UK; by 2016, 48% of this migrant workforce was from Southern Asia (35%) and Maritime South East Asia (13%) combined, with 28% of these workers coming from non-main English-speaking countries (Eastman et al., 2018). With specific reference to early childhood education, by 2016 approximately 35% of Australia’s childcare workforce was born outside Australia. In Singapore too, even as its early childhood sector continues to expand amid a shortage of locally trained educators, current labor laws potentially allow about a third of its early childhood workforce to be made up of migrant educators. Based on anecdotal evidence, we know that many migrant educators hold short-term (2-year) work visas and have moved from the neighboring ASEAN countries and the Peoples’ Republic of China.

Given the scale of global migration in care work, it is remarkable that so little attention has been paid to the learning needs of migrant workers in early childhood education, especially considering the wealth of literature on work with migrant children and families. Recently, Quintero et al. (2021) have argued for the convergence of leadership, teacher education, early childhood education, migration, and storytelling in the preparation of all early childhood educators, particularly those who are refugees or migrants. But how and why do these educators join the workforce in the first place? What are their professional learning needs? And how do they contribute to early childhood program quality? These too are “missing links” in early childhood education and teacher education research.

To inform the best ways to support diverse educators for contemporary early childhood education, the early childhood field needs to know more about basic phenomena such *who* these diverse educators are and *why* they have chosen to (a) migrate and/or (b) join the early childhood education and care profession. Here we deliberately situate this question as a and/or problem. This is because some migrant educators are already early childhood professionals in their country of origin and migrate via dedicated immigration pathways (as noted earlier, Singapore is one example of this approach); while other countries mainly rely on individuals who enter the country via other migration pathways then make a career change from their original profession into early childhood education. Australia is a notable example of the latter approach, having no dedicated migration pathway for early childhood professionals; only center managers and childcare group leaders (with the required qualification and at least 1 year of experience in a regulated early childhood service) are eligible for migration assessment (Australian Department of Home Affairs, 2021). This approach of relying mainly on migrants making a career switch to early childhood education post-migration results in a rich mix of experience among migrant educators; in their interviews with migrant women who had become early childhood educators after migrating to Australia, Hamilton et al. (2021) found a range of prior qualifications and work experience including journalism, pharmacy, midwifery, chemistry, and social work. However, even if migrant educators were already qualified early childhood teachers in their country of origin, systematic yet holistic support is needed to help them develop a sense of belonging, re-situate and strengthen their identities as professionals in a new cultural context, make sense of their own professional learning sufficiently to exercise teacher agency (Beijaard and Meijer, 2017), and respond to the professional norms, curricular frameworks, and cultural expectations around children, early childhood education, and parenting in their destination country.

Supporting the learning of a diverse and globally mobile early childhood education workforce

Although there is a considerable literature on the effectiveness of professional development for work with diverse children (see Bottiani, 2017) for a systematic review of this literature in school settings), as yet there is almost no literature that considers the learning needs of diverse educators. Beyond understanding the profile and motives of migrant early childhood educators, researchers need to find ways to systematically identify the expertise these educators bring from their countries of origin. This expertise includes the

range of ways in which “early childhood program quality” is conceptualized across families from diverse cultural backgrounds. In keeping with our argument in this chapter, this empirical work should be guided by background theories of learning and development that take cultural differences into account. Even without this research, we suggest there is a need to support migrant educators in their transition to working in their destination countries. However, such programs must not themselves rest on normative assumptions, i.e. that early childhood programs in destination countries are framed around “best practice” so it is the task of migrant educators to adapt and enculturate themselves to local definitions of quality in early childhood programs. We also point to the need for *local* educators in destination countries to learn to recognize and value the rich professional and cultural knowledges that can be contributed by migrant educators, and know how to mobilize these knowledges in the interests of children and families. If such programs position migrant educators as “the other,” they will not only perpetuate structural racism but perpetuate the neglect of the expertise of *all* educators we have detected in our analysis of early childhood quality policies in East Asia (Nuttall et al., 2021).

We see considerable potential in theories of learning and development developed within and across diverse cultures to address the need for robust conceptualization of professional development in early childhood education. But any call for such a conceptualization is not neutral. There is now considerable evidence for the ways in which normative (i.e. race-based) assumptions of development obscure the valuable knowledges and cultural practices of diverse children and families, with inequitable consequences for their communities (Delpit, 2002) and negative impacts on child health (Trent et al., 2019). Theorization of professional development as well as quality in early childhood education should therefore reach beyond theories of learning and development anchored in traditional developmental psychology. Instead, researchers need to develop framings that understand educator learning as embodied and relational (Cekaite and Bergnehr, 2018), gendered, classed, and raced (Wright, 2011) and context-bound (Gomez et al., 2015)—concepts it is impossible to ignore when paying attention to the work educators actually do in early childhood education. These theorizations should be complemented by use of theories that explain the systemic factors that marginalize women and migrants, such as critical race theory, sociohistorical theory, feminist economic theories, and theories developed by First Nations scholars and in the global south. It is through engagement with such background theoretical frameworks that more culturally nuanced, and therefore more representative, principles and practices of quality improvement in early childhood education can be developed. We suggest such theorization of professional development *as learning* would benefit all early childhood professionals, not just migrant educators.

Conclusion

In this chapter we have critiqued the neglect of educators’ lived experiences and expertise in early childhood professional development research and argued that this is characteristic of the studies conducted during the last 10–20 years, when governments around the world have attempted to strengthen early childhood system and center quality. We have also suggested that, by conducting professional development research within the policy|quality nexus, a closed instrumental loop—research influencing policy influencing research—educators’ expertise and needs are sidelined. Culturally diverse educators are a particular omission from this research; in our analysis of the early childhood professional development research we were unable to find any studies of professional development specifically for diverse educators. Despite attempts by policymakers to increase program quality in the interests of diverse and marginalized children and families, culturally diverse educators remain “missing in action” in the early childhood professional development research. The neglect of their learning, as well as the failure to mobilize their expertise, seems an illogical oversight in the context of increasing global mobility and diversity, as well as the need to develop rich understandings of early childhood program quality.

We have also pointed to the absence of robust theorizing within early childhood professional development research to date. Although there have been some attempts since 2015 to identify concepts that might assist researchers to describe and analyze the complex nature of adult learning and development in early childhood education, and allow educators to connect their learning to desired outcomes in their work, there has been limited progress. Two bodies of theory suggest themselves as obvious starting points: developmental theory derived from cross-cultural understandings; and sociological and other theories with explanatory power for diverse educational contexts. Only then will it be possible to understand how educators develop in early childhood education and, therefore, how truly effective professional development for quality improvement should be designed. We have argued for a definition of professional development anchored in educators’ learning as the necessary starting point for this work.

References

- Australian Children’s Education and Care Quality Authority [ACECQA], 2021. Shaping Our Future: A Ten-Year Strategy to Ensure a Sustainable, High-Quality Children’s Education and Care Workforce 2022–2031. ACECQA, Sydney, Australia.
- Angus, L., 2012. Preparing teachers as informed professionals: working with a critical ethnographic disposition and a socially democratic imaginary. In: Down, B., Smyth, J. (Eds.), *Critical Voices in Teacher Education: Teaching for Social Justice in Conservative Times*. Springer, Dordrecht, Netherlands, pp. 45–61.
- Australian Department of Home Affairs, 2021. Skilled Occupation List at Skilled occupation list. homeaffairs.gov.au. (Accessed 10 January 2022).
- Barber, H., Chrsssen, C., Church, A., 2014. Meeting the Australian National Quality Standards: a case study of the professional learning needs of early childhood educators. *Australas. J. Early Child.* 39 (4), 21–27.

- Beijaard, D., Meijer, P.C., 2017. Developing the personal and professional in making a teacher identity. In: Clandinin, D.J., Husu, J. (Eds.), *The SAGE Handbook of Research on Teacher Education*, pp. 177–192.
- Bottiani, J.H., Larson, K.E., Debnam, K.J., Bischoff, C.M., Bradshaw, C.P., 2017. Promoting educators' use of culturally responsive practices: a systematic review of inservice interventions. *J. Teach. Educ.* 69 (4), 367–385. <https://doi.org/10.1177/0022487117722553>.
- Brown, C.P., Englehardt, J., 2016. Conceptions of and early childhood educators' experiences in early childhood professional development programs. A qualitative metasynthesis. *J. Early Child. Teach. Educ.* 37 (3), 216–244. <https://doi.org/10.1080/10901027.2016.1204574>.
- Buyse, v., Winton, P.J., Rous, B., 2009. Reaching consensus on a definition of professional development for the early childhood field. *Top. Early Child. Spec. Educ.* 28 (4), 235–243. <https://doi.org/10.1177/0271121408328173>.
- Campbell, F., Conti, G., Heckman, J.J., Moon, S.H., Pinto, R., Pungello, E., Pan, Y., 2014. Early childhood investments substantially boost adult health. *Science* 343, 1478–1485. <https://doi.org/10.1126/science.1248429>.
- Cekaitea, A., Bergnehr, D., 2018. Affectionate touch and care: embodied intimacy, compassion and control in early childhood education. *Eur. Early Child. Educ. Res. J.* 26 (6), 940–955. <https://doi.org/10.1080/1350293X.2018.1533710>.
- Clarke, D., Hollingsworth, H., 2002. Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18 (8), 947–967. [https://doi.org/10.1016/S0742-051X\(02\)00053-7](https://doi.org/10.1016/S0742-051X(02)00053-7).
- Delpit, L., 2002. No kinda sense. In: Delpit, L., Dowdy, J.K. (Eds.), *The Skin That We Speak: Thoughts on Language and Culture in the Classroom*. The New Press, New York, pp. 31–48.
- Down, B., 2012. Reconceptualising teacher standards: authentic, critical and creative. In: Down, B., Smyth, J. (Eds.), *Critical Voices in Teacher Education: Teaching for Social Justice in Conservative Times*. Springer, Dordrecht, Netherlands, pp. 63–80.
- Eadie, P., Stark, H., Niklas, F., 2019. Quality of interactions by early childhood educators following a language-specific professional learning program. *Early Child. Educ. J.* 47, 251–263. <https://doi.org/10.1007/s10643-019-00929-5>.
- Eastman, C., Charlesworth, S., Hill, E., 2018. FACT SHEET 1: Migrant Workers in Frontline Care. <https://www.unsw.edu.au/arts-design-architecture/our-research/research-centres-institutes/social-policy-research-centre/our-projects/markets-migration-work-care-australia>. (Accessed 11 January 2022).
- Egert, F., Fukkink, R.G., Eckhardt, A.G., 2018. Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: a meta-analysis. *Rev. Educ. Res.* 88 (3), 401–433. <https://doi.org/10.3102/0034654317751918>.
- European Commission, 2010. Communication From the Commission: Europe 2020, a Strategy for Smart, Sustainable and Inclusive Growth. Retrieved from: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF>.
- Gomez, R.E., Kagan, S.L., Fox, E.A., 2015. Professional development of the early childhood education teaching workforce in the United States: an overview. *Prof. Dev. Educ.* 41 (2), 169–186. <https://doi.org/10.1080/19415257.2014.986820>.
- Hamilton, M., Hill, E., Adamson, E., 2021. A “career shift”? Bounded agency in migrant employment pathways in the aged care and early childhood education and care sectors in Australia. *J. Ethnic Migrat. Stud.* 47, 3059–3079. <https://doi.org/10.1080/1369183X.2019.1684246>.
- Harkoma, S.M., Sajaniemi, N.K., Suhonen, E., Saha, M., 2021. Impact of the pedagogical intervention on early childhood professionals' emotional availability to children with different temperament characteristics. *Eur. Early Child. Educ. Res. J.* 29 (2), 183–205. <https://doi.org/10.1080/1350293X.2021.1895264>.
- He, P., Ho, D., 2013. Rethinking the professional development policy for early childhood education in Hong Kong. *Australas. J. Early Child.* 38 (2), 95–102.
- Heckman, J., 2011. The Economics of Inequality. The Value of Early Childhood Education. American Educator, Spring, pp. 31–47.
- Inglis, S., Brown, A., 2013. So what happens after the event? Exploring the realisation of professional development with early childhood educators. *Australas. J. Early Child.* 31 (1), 11–15.
- Jensen, B., Iannone, R.L., 2018. Innovative approaches to continuous professional development (CPD) in early childhood education and care (ECEC) in Europe: findings from a comparative review. *Eur. J. Educ.* 53, 23–33. <https://doi.org/10.1111/ejed.12253>.
- Linder, S.M., Rembert, K., Simpson, A., Ramey, M.D., 2016. A mixed-methods investigation of early childhood professional development for providers and recipients in the United States. *Prof. Dev. Educ.* 42 (1), 123–149. <https://doi.org/10.1080/19415257.2014.978483>.
- Machado, I., 2019. Context based participatory professional development of educational teams: educators and assistants learning in companionship. *Eur. Early Child. Educ. Res. J.* 27 (3), 346–359. <https://doi.org/10.1080/1350293X.2019.1600804>.
- Mendive, S., Weiland, C., Yoshikawa, H., Snow, C., 2016. Opening the black box: intervention fidelity in a randomized trial of a preschool teacher professional development program. *J. Educ. Psychol.* 108 (1), 130–145. <https://doi.org/10.1037/edu0000047>.
- National Association for the Education of Young Children [NAEYC], 2022. Professional Development webpage. <https://www.naeyc.org/resources/pd>. (Accessed 11 January 2022).
- Ng, C.L.P., 2014. Mother tongue education in Singapore: concerns, issues and controversies. *Curr. Issues Lang. Plann.* 15 (4), 361–375. <https://doi.org/10.1080/14664208.2014.927093>.
- Nolan, A., Molla, T., 2018. Teacher professional learning in early childhood education: insights from a mentoring program. *Early Years* 38 (2), 258–270. <https://doi.org/10.1080/09575146.2016.1259212>.
- Nuttall, J., Grieshaber, S., Lim, S., Yun, E., Oh, J., Ahn, H., Sum, C.W., Yang, W., Kim, S., 2021. Workforce diversity and quality improvement policies in early childhood education in East Asia. *Early Years*. <https://doi.org/10.1080/09575146.2021.1888076>.
- OECD, 2019. Good Practice for Good Jobs in Early Childhood Education and Care. OECD Publishing, Paris. <https://doi.org/10.1787/64562be6-en>. (Accessed 7 January 2022).
- Peleman, B., Lazzari, A., Budginaite, I., Siarova, H., Hauari, H., Peeters, J., Cameron, C., 2018. Continuous professional development and ECEC quality: findings from a European systematic literature review. *Eur. J. Educ.* 53, 9–22. <https://doi.org/10.1111/ejed.12257>.
- Quintero, E.P., Callaway-Cole, L., Taha-Resnick, A., 2021. Rethinking Higher Education: Making Space for Stories Leadership in Higher Education: Learning With Migrant and Refugee Populations in Early Childhood and Teacher Education Contexts. Springer, Singapore.
- Schacter, R.E., Gerde, H.K., Hatton-Bowers, H., 2019. Guidelines for selecting professional development for early childhood teachers. *Early Child. Educ. J.* 47, 395–408. <https://doi.org/10.1007/s10643-019-00942-8>.
- Schweinhart, L.J., Weikart, D.P., 1997. The high/scope preschool curriculum comparison study through age 23. *Early Child. Res. Q.* 12, 117–143.
- Sheehan, M., 2017. Revisiting women's migration and care labor. *Migr. Stud.* 6 (2), 312–319. <https://doi.org/10.1093/migration/mnx036>.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., 2004. The Effective Provision of Pre-school Education (EPPE) Project: Findings From Pre-school to End of Key Stage 1. Department for Education and Skills, Nottingham, United Kingdom.
- Tonyan, H.A., Nuttall, J., Torres, J., Bridgewater, J., 2017. Engaging with quality improvement initiatives: a descriptive study of learning in the complex and dynamic context of everyday life for family child care providers. *Early Educ. Dev.* 28 (6), 684–704. <https://doi.org/10.1080/10409289.2017.1305152>.
- Torres, E., Marea, M., Mendive, S., 2021. Change in early childhood classroom interaction quality after a professional development programme. *J. Stud. Educ. Dev.* <https://doi.org/10.1080/02103702.2021.1972699>.
- Trent, M., Dooley, D.G., Dougé, J., 2019. The Impact of Racism on Child and Adolescent Health. Policy Statement. American Academic of Pediatrics, Itasca, IL. AAP.
- Vähäsantanen, K., Hökkä, P., Eteläpelto, A., Rasku-Puttonen, H., Littleton, K., 2008. Teachers' professional identity negotiations in two different work organisations. *Vocat. Learn.* 1 (2), 131–148.
- Winton, P.J., Snyder, P.A., Goffin, S.G., 2016. Beyond the status quo: rethinking professional development for early childhood teachers. In: Couse, L.J., Recchia, S.L. (Eds.), *Handbook of Early Childhood Teacher Education*. Routledge, New York, USA, pp. 54–68.

- Wolf, S., Aber, J.L., Behrman, J.R., Tsiniogo, E., 2019. Experimental impacts of the “Quality Preschool for Ghana” interventions on teacher professional well-being, classroom quality, and children’s school readiness. *J. Res. Educ. Eff.* 12 (1), 10–37. <https://doi.org/10.1080/19345747.2018.1517199>.
- Wright, T.S., 2011. Countering the politics of class, race, gender, and geography in early childhood education. *Educ. Pol.* 25 (1), 240–261.
- Zembylas, M., 2011. Teaching and teacher emotions: a post-structural perspective. In: Day, C., Lee, J.C.-K. (Eds.), *New Understandings of Teacher’s Work: Emotions and Educational Change*. Springer, Dordrecht, Netherlands, pp. 31–43.

The global and the local: idea flows, contexts, and the influencing of education policy in the 21st century

J.B. Helgetun, Institute for the Analysis of Change in Historical and Comparative Societies, University of Louvain, Louvain-la-Neuve, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	604
The global context: from the age of enlightenment to PISA and TALIS	605
The regional context: Bologna and the European Union: attempts to harmonize teacher education in Europe	606
The local context: England and France on diverging trajectories	607
Concluding remarks	608
Acknowledgment	609
References	609

Introduction

This essay discusses the relation between the global, regional, and local contexts of policymaking in education as observed in the two contrasting cases of teacher education in England and France. This is done by drawing on existing academic literature, and empirical examples drawn from research conducted as part of the TeachersCareers project¹ (see Helgetun, 2021 for an in-depth presentation of the research methodology that informed the discussion presented here), to construct a socio-historical narrative. The literature was identified based on a systematic literature reviews on the topic of teacher policy (Sørensen and Dumay, 2021) conducted within the project, as well as discussions with leading authors in the field in both England and France. Due to the format, some of the information presented here is brief, however, sources are provided should the reader desire to read further about for example teacher education provision and policymaking in England and France. As such, this essay can be seen as providing an overview of, and entry point for, scholars who desire further knowledge on the interaction between the global and the local.

The importance of such contrasting analysis of teacher education policy and policymaking rests in how teachers are arguably the central component of education regardless of if their role is to transfer knowledge or enable children to learn throughout their life (Brisard, 2002). Moreover, what teachers know and how they perceive their roles is rooted in the teacher's professional education and training (Beck and Young, 2005; Beck, 2009), which is shaped by policymakers (Foster, 1999; Brisard, 2002). Policymakers here refer to individuals that make de facto rules that govern the practices of others. At a state level, these individuals are usually elected officials such as politicians and/or public administrators such as civil servants. Policy on teacher education is in turn influenced by stakeholders such as teaching professionals, teacher education providers (UNESCO, 2005), or researchers (Helgetun and Menter, 2020), and reflect beliefs in a society (Menter, 2016; Tatto et al., 2016). As the discussion here shows, teacher education policymaking is largely influenced by local contexts as well as global trends, reminiscent of what Maroy et al. (2016) called a form of vernacular globalization. These developments are then discussed in relation to European integration theories, rarely used in analysis of policy and policymaking in education. What is further argued here is that globalization is nothing new, but that the form global communication takes, and as such its impact, has changed over the past centuries due to technological innovations.

To present the context and relevance of the discussion: Internationally, academic research has shown an ongoing convergence of societal beliefs on what education ought to do (Meyer et al., 1997; Schofer and Meyer, 2005; Meyer and Benavot, 2013) as well as a rise of what has been dubbed a global policy space in education (Lingard and Rawolle, 2011). Concurrently, solutions to perceived issues in teacher education have been found to be increasingly solved through policy borrowing (Steiner-Khamsi, 2010), particularly from high performers in PISA (Chung, 2016). Furthermore, education, as other aspects of public policy, is increasingly dominated by New Public Management (NPM) type policies all over the world (Hartley, 2003; Steiner-Khamsi and Stolpe, 2006; Addey et al., 2017; Dupriez and Cattonar, 2018). However, what teacher education policy is enacted and for what purpose, has been found to differ between countries and reflect the socio-cultural and historic particularities in them (Brisard, 2002). For example, France has since 1989 gone toward a university based model of teacher education provision (Sacilotto-Vasylenko and Fave-Bonnet, 2011; Prost, 2014), while England has since 1984 moved in the opposite direction from academia toward a school-led system (Furlong, 2013; Whiting et al., 2018). In England this has been addressed as a "turn to the practical" (Childs and Menter, 2013; Quirk-Marku and Hulme, 2017), while in France the trends since 1989 are seen as a form of "universitification" of teacher education (Misra, 2014; Farges, 2017). Thus, two contrasting trajectories can be observed in England and France, the former centered on an increasing role for the school, the latter on an increasing role for the university that is happening in spite of any global convergence, the Bologna process, or EU initiatives in relation to education (Helgetun, 2021). The national trajectories are here interesting in the sense that

¹See www.teacherscareers.eu.

they tell us something about the societal views on what a teacher is and what their job in the society is as expressed in different education and training models for teachers.

The next section discusses the global context from the Age of Enlightenment to PISA and TALIS, then the regional context of Bologna and the EU is presented, before the following section discusses the local contexts of England and France and how they relate to the global and regional contexts. The essay then presents a case for analyzing teacher education policy through a liberal inter-governmentalist lens (Moravcsik, 1993) before concluding remarks are provided.

The global context: from the age of enlightenment to PISA and TALIS

Different as their approaches to teacher education today are, in historic terms educational developments in both England and France as understood today originated in the Age of Enlightenment, which one may claim began with the Peace of Westphalia in 1648 and lasted until the Congress of Vienna in 1814/15. An era which laid most of the intellectual foundations for later developments in education (Tröhler, 2020). Meanwhile, the work of intellectuals such as Locke, Montesquieu, Rousseau, and Voltaire was disseminated across the Euro-centric world with an impact on how both early modern nation-states and education were conceived and constructed (Paine, 1995; Anderson, 2006; Tröhler, 2014, 2020). The Euro-centric world is here taken to mean places reached by the Nations of Europe, be it the Americas, Australia, or India. This arguably makes historical divergences, and the trajectories of socio-cultural peculiarities in the two cases as they engage with “borderless” ideas of the enlightenment, particularly interesting. For example, Jean-Jacques Rousseau’s work “*Émile, or on education*” was immediately translated from French to English upon its publication in 1762 (Rosenberg, 1981) and served as an inspiration for educational developments in both France and England, particularly through impacting on how people thought and talked about education (Morley, 1878). Meanwhile, due to the Age of Empire, the ideas of India were for example introduced to England and vice-versa (Alexander, 2001). As such, historically ideas have been hard to contain within artificial state borders, as they are disseminated by the exchange of people and literature across large distances and between diverse societies. In more recent times, this can for example be seen in how across time both England and France have engaged with remarkably similar ideas of progressive versus traditional education, such as the concept of “new mathematics” in the 1960s and 70s, or had to confront the complex opposition between practice based teacher training and academia centered teacher education that go back over 100 years in both cases (Helgetun, 2021).

The last several decades have seen the birth of different comparative approaches from the OECD’s PISA comparing student performance and well-being in schools (Ydesen, 2019), the International Association for the Evaluation of Educational Achievements (IEA) TIMSS measuring mathematics performance, as well as Eurydice presenting comparative overviews of countries education and teacher education systems (Akiba, 2017), that can be traced back (at least) to measures taken in the West following the Sputnik shock in 1957 (Tröhler, 2013). These new approaches employ mostly large-scale data sets but also qualitative measures such as the United Nation’s goal of education for all (Lingard and Rawolle, 2011; Fazekas and Burns, 2012; Robertson, 2012; Normand, 2013; Sellar and Lingard, 2013). This constitutes a major shift from the past, where ideas and ideations such as the work of Rousseau, Locke, or Montesquieu, that all affected the early global construction of the state and education, were spread by individual intellectuals rather than multi-national organizations.

What is common between today and the 18th century is a broad dissemination of ideas on education; what has shifted is the mode of dissemination (particularly speed and reach in a population) and the type of “evidence” used to support ideas on education. For example, Rousseau’s “*Émile*” has a distinctly different form of discourse than the OECD’s PISA reports. PISA, as the rest of the work of the OECD, is highly focused on measurement and statistical data, with the purpose of education framed under the mantra of a “knowledge based economy” (Godin, 2006), while *Émile* focuses on the nature of education and the nature of man to philosophize on the education of an ideal citizen. As such, they constitute two distinct “cognitive blueprints” in their own right, but also reflect fundamentally different views on the purpose of education. A cognitive blueprint is in short a structuring idea causing through mimesis (imperfect copying of a blueprint) institutionalized beliefs in something being “what works” (best practice) and therefore being what should be done (Zapp et al., 2018a,b; Helgetun and Dumay, 2021). For example, the work of Dewey—as a cognitive blueprint—has arguably fostered a belief in learning by doing, or PISA rankings and reports has fostered a belief in the Finnish model of teacher education as “best.”

In relation to teacher Education policy in the mid to late 20th century and early 21st century, global trends that have been observed are copying of high performing, mostly as measured in the OECD’s PISA and IEA’s TIMSS surveys, countries approaches to educating teachers, particularly the Finnish and Singaporean models (Steiner-Khamisi, 2010; Chung et al., 2012; Furlong, 2013; Sellar and Lingard, 2014; Chung, 2016; Adamson et al., 2017; Steiner-Khamisi et al., 2019).

In more recent years, the OECD has also begun a more targeted approach to understanding teacher’s work and their education through the Teaching and Learning International Survey (TALIS) program, the first round was in 2008, with new methodologies compared to past surveys (see Akiba, 2017), which may be argued to derive legitimacy from claims of “knowledge sharing” and listening to teachers (Sørensen, 2016). Meanwhile, the IEA launched the Teacher Education and Development Study—Mathematics (TEDS-M) in 2008 (same year as TALIS) to compare how different countries prepare their teachers to teach math’s (Akiba, 2017). It has also been argued that not all countries participated in TEDS-M under the belief that TALIS would cover the same things (so TALIS would suffice), this is reportedly not the case as TALIS and TEDS-M are distinct (Tatto, 2011). In either case, the goal of these programs is in essence to figure out “what works” in education and disseminate the results (Akiba, 2017; Tatto, 2011; Zapp, 2020). Furthermore, these new surveys of TALIS and TEDS-M may also become a political tool alongside such existing surveys as PISA and

TIMMS in new forms of education governance, centered on cultural capital and ideas more so than traditional hierarchies anchored in a legal framework (Steiner-Khamsi, 2010; Chung, 2016; Akiba, 2017).

The use of international surveys and reports to influence policymaking has been further helped along by what is reported to be an increasing network of global actors in education and teacher education (Ball and Junemann, 2012; Ball, 2016), that has emerged since the 1950s. With this increase of transnational actors, Resnik (2006) argued that international organizations have become agents in a global education arena where they act as producers of a global educational culture.

In this paradigm centered on a global or globalizing educational culture and rise in international organizations influence in education, “public” actors such as the OECD, UNESCO, and the EU, as well as “private” actors such as McKinsey, ARK, and even historically defense oriented organizations such as RAND and Defense Advanced Research Project Agency (DARPA) attempt to influence how teachers should be educated mainly through diffusing scripts of “what works,” and are reported to have the ears of people in high places (Saunders, 2007; Alexander, 2010; Coffield, 2012; Tröhler, 2013). In other words, teacher education is becoming a global policy area, where both public and private national and international actors alike contest practices and propose solutions to problems. The socio-cultural convergence, driven by IOs such as the OECD, is argued to constitute global normative, economic, and cultural pressure to adopt certain standardized approaches to education, in what has been deemed a globalizing world (Meyer and Ramirez, 2007; Zapp, 2020).

The regional context: Bologna and the European Union: attempts to harmonize teacher education in Europe

Any socio-cultural convergence in Europe today is arguably driven in part by structural harmonization attempts, primarily the Bologna process and European Union (EU) interference in education, as this section will explore.

The 1999 Bologna accord and subsequent Bologna process has sought to unify degrees across signatories, and harmonize on the 3 + 2 + 3 structure and the European Credit Transfer and Accumulation System (ECTS) where one academic year equals 60 credits. The accord was signed by both England and France, and arguably has implications for teacher education, which one may characterize today as tertiary education in both England and France. It has been observed that the Bologna process has led to convergence in higher education policy across signatory countries, and to a lesser extent also to non-signatories (Vögtle et al., 2011). Here it is important to keep in mind that general convergence in the abstract, does not automatically equate to concrete convergence in the specific. For example, teacher education *may* be loosely coupled (see (Elken and Vukasovic, 2019) for a discussion on the “loosely defined” concept of loose coupling) or de-coupled (see (Bromley and Powell, 2012) for an update of the concept, that includes not only policy-practice de-coupling but also means-end de-couplings) from higher education as a whole. Or, de-coupling may happen between different levels of policy through a form of two-level games (Putnam, 1988), thereby explaining why some countries, institutions and disciplines converge and others do not. To clarify, two-level games relates to when international considerations may dictate domestic policy and vice-versa. A de-coupling between levels would then be if one form of discourse and even policy as text is adopted at one level, but never implemented in another. In other words, the Bologna accord may become de-coupled due to a lack of national implementation in the context of practice.

Another harmonizing force is in education touching on our two cases is arguably the EU. To provide a caveat, England left the EU in 2020, however they were part of the organization for 47 years, 28 of them following the 1992 Maastricht treaty which transformed the European Communities into the European Union with the contours of a shared political framework (see Cini, 2007) for an overview of the European Community, the Maastricht treaty, and the European Union as a political entity). Therefore, we cannot ignore the history of England—EC/EU relations, but should also note that the relationship has not always been cordial, and did come to an abrupt end. As is discussed below, the European integration processes has been seen differently in England and France.

Officially, education is the prerogative of the nation state, without EU influence on the matter. However, in spite of this, the EU has been identified as a driving force behind Bologna (Simões et al., 2018). The EU has also made large inroads into coordinating and funding educational research through the Horizon (2020) program leading to some convergence in educational research (Powell, 2020), particularly as EU funds are often seen as “elite” and the pinnacle of research funding in education (Zapp et al., 2018a,b). Although the EU’s reach into teacher education is legally limited they still are very much present particularly on the agenda-setting stage of policymaking. Their power as such, seems to mostly lay in research funding and dissemination of ideas and values through publications and workshops such as the European semester whose original and official purpose is economic coordination, however, their Europe 2020 agenda identified education as one of five key areas for the semester to focus on (Wilkinson, 2019). Central to such “EU level” discourse on teacher education is arguably the perceived vital role teachers play in advancing human potential, shaping future generations and influencing society as a whole (Simões et al., 2018). A result of such emphasis has been argued to manifest as a form of “European teacher” (Simões et al., 2018). Thus, discussion of education, including teacher education and the related topics of educational research and higher education are all under the EU umbrella of coordination, standard setting, and idea dissemination.

However, it is important not to fall into a common pitfalls of neglecting cultural differences as to what “global” may mean, and how academic concepts such as “globalization,” “global/regional standardization,” “convergence” or “unification” may be at least as much constructions of the social sciences as empirical observations of “real” phenomenon (Tröhler, 2014). Instead, globalization, regionalization, and other catchwords denoting a global interdependence or a global organizational field and institutionalized environment, must be carefully analyzed and seen in relation to unique contexts that are both temporally and spatially dependent. Thus, we may talk of a global scale or level of organizational fields and institutionalized environments (both plural) pertaining to education, but as we do so we must relate it back to other levels of the scale and avoid temptations to overly

generalize without denoting local variances of importance. This is evident for instance in how similar scripts and modes of thought may influence the construction of new practices, but at the same time be tempered by local factors and political bricolage resulting in translation and shifts in meaning (Maroy et al., 2016). For these reasons, a global cognitive blueprint, such as the knowledge economy, has been found to not be uniformly adopted in specific contexts but takes on local variances (Zapp et al., 2018a,b) that arguably constitutes a separate derivate blueprint, which may again resonate differently in other contexts and again find itself translated and changed into a new and different blueprint (e.g. the knowledge economy becoming the knowledge society). In other words, there is nothing static about idea transfer and the spread of “what works.” This then leads us to the importance of not only emphasizing the global or the regional, but to always keep an eye on the local context as well to which the essay now turns.

The local context: England and France on diverging trajectories

England is characterized largely as a systemless system in education, held together by data-flows centered on certification, inspections, and prescribed standards (Lawn, 2013; Helgetun and Dumay, 2021). Within this systemless system the state has comparatively low internal capacities to directly govern teacher education and produce data for policy, instead relaying on governing at a distance mechanisms (Ball, 2016; Ozga, 2009, 2019) and external sources such as the NFER or Education Datalab to produce evidence for policymaking (Helgetun and Menter, 2020; Helgetun, 2021). As such, England has a textbook new public management system of governance in (teacher) education at present. Meanwhile, in ideological terms education, including teacher education, is increasingly framed by both the concept of the knowledge economy as well as neo-liberalism centered on a free market of provision under state prescribed rules of the game. The neo-liberal system of today, that is coupled with new public management methods of steering at a distance through inspections and data flows, arguably grew out of a largely free model of teacher education devoid of much accountability or harmonization prior to the creation of the Qualified Teacher Status certificate in 1984 (Furlong, 2013). In terms of provision, this has led to a fragmented system where teachers are educated in a myriad of ways centered either on Universities, School-Centered Initial Teacher Training providers, or true partnerships between the two (Whiting et al., 2018; Helgetun and Dumay, 2021). In other words, there is no central model for teacher education (there never has been), only the teacher standards that form the basis of the QTS to standardize and define what a teacher is in England and therefore what their education should be (Childs and Menter, 2013; Helgetun and Dumay, 2021; Helgetun, 2022). The development of the QTS was in one sense the outcome of local concerns, such as the Great Debate in Education started in 1976 by Labor Prime Minister Jim Callaghan and concerns regarding was seen as a secret garden of the curriculum and progressive excesses in education (Phillips, 2001), which can be traced back to the “Black Papers” in the 1960s (Helgetun, 2021). Meanwhile, the OECD were emphasizing the link between the economy and education and the need for state intervention in a time of economic upheaval following the 1973 oil crisis, further influencing the perceived need for the state to take on an active role in all aspects of education (Phillips, 2001), including teacher education (Furlong, 2013). Hence, there were both local and international debates shaping the need to devise some system of state intervention in the fragmented landscape of education in England, leading to the current iteration of the systemless system in teacher education (Helgetun, 2022). These changes have led to normative transformations in England, where the teacher now is arguably a form of “self-improving craftsperson” (Helgetun and Dumay, 2021), rather than the old image of a “scholar who just happened to be a teacher” that was almost taken for granted (at least by academia) prior to the Great Debate (Furlong, 2013). As such, the teacher is today defined as a professional anchored in the school more than in academia, while government policy is moving in a traditional direction.

By contrast, education in France, including teacher education, is shaped by a Jacobin logic (Normand, 2013) and meritocratic myth centered on the values of the French republic. The Jacobin approach centers on a strong educative state which aims to educate the future leaders of the country through a meritocratic approach that enables these shepherds to guide and lead the flock (Helgetun, 2021). This necessitates centralized control over an education system that is harmonious and as singular as possible (Alexander, 2001), while being rooted in the local context of *la République Française*, where technocratic mastery is displayed by the state. This view on the teacher harkens back, at least, to the French Revolution of 1789 (see e.g. Condorcet (1792)), and became cemented through the Ferry laws in 1881/2 (Prost, 2013; Helgetun, 2021). As such, teachers were, and still are, in France seen as “agents of the republic” who are educated by the French state as uniformly and harmoniously as possible to represent the values of the republic (Ozouf and Ozouf, 1992; Farges, 2017). Meanwhile, the French state takes responsibility for the large-scale production of both qualitative and quantitative data to shore up policy in education through agencies such as the national education inspectorate (IGEN) and national data analyst center in education (DEPP) (Helgetun, 2021). As such, the French model is much more static than its English counterpart at an ideological level, in spite of dramatic changes to the mode of provision instigated by the creation of the IUFM in 1989 (Prost, 2014). The changes are instead centered on improving the forming of these agents of the republic through an increased emphasis on the university (Saciolotto-Vasylenko and Fave-Bonnet, 2011; Cornu, 2015). This contrasts to England where the change was not only in terms of form of teacher education, but also in the very meaning of what a teacher was to be. To provide some context, the IUFM when created by law in 1989, was the first unified institute of teacher education in France encompassing both primary and secondary teachers. Prior to the opening of the IUFMs in 1992, primary teachers were educated in “teachers’ colleges,” while secondary teachers had no formal teacher education beyond requirements of a graduate degree in the subject they were to teach and the passing of the civil service entry examination known as “concours.” The concours is adapted to the type of civil servant one desires to be, and was as such aimed at becoming a teacher.

However, it was, and is, heavily focused on subject knowledge and not knowledge of how to teach. See for example [Helgetun \(2021, 2022\)](#) for an elaboration on the history of teacher education in France in English, or [Prost \(2014\)](#) for a more comprehensive history in French.

To further analyze the European dimension behind these shifts, as observed by John [Sayer \(2006\)](#) English policymaking has not taken European trends into account at any point in regards to teacher education. Meanwhile, according to [Helgetun and Menter \(2020\)](#) for example, references in teacher education policy documents in England are to Hong Kong and Singapore more so than European countries (with the exception of Finland) or even Wales, Northern Ireland, or Scotland. Furthermore, as noted by [Chung \(2016\)](#) the inspiration for teacher education policy following 2010 was the Finnish education model, which was much misunderstood and abused. Moreover, there are no references of note to Bologna in the national discourse, nor does teacher education follow the 3 + 2 + 3 model. Meanwhile, courses are not denoted in terms of ECTS but a national credit system. It appears then as if England is content not to adopt European models, but instead cherry picks policy ideas from across the globe. By contrast, France has engaged with Bologna, but not to actually mirror trends elsewhere as so much as implementing desired national reforms. Teacher education became systemically linked to the university in 1989 with the above-mentioned creation of the IUFM that instigated a 3 + 2 system where future teachers would first attend 3 years of undergraduate studies in a subject area, followed by a two year post-graduate education in teaching well before Bologna ([Bon, 2014](#)). However, a reinforcement of the master component of teacher education was begun in the mid-2000s, culminating in the change from the IUFM to the ESPE in 2012 and creation of a “new” Master in Education and Teaching the same year, all supported by arguments centered on the need to respect Bologna and fully incorporate the 3 + 2 + 3 system in teacher education as this was what “others” in Europe were doing ([Helgetun, 2021, 2022](#)). However, there remains limited evidence that others were in fact doing as the French were, and a closer examination of the discourse seem to indicate that Bologna was a useful argumentative tool for implementing reforms begun in 1989 ([Bon, 2014](#)).

Here, derivative theories from [Putnam’s \(1988\)](#) two-level games (e.g. liberal intergovernmentalism, see [Moravcsik, 1993, 1997](#)) may help explain these relations better than dominant models in international and comparative education rooted in supranationality or international policy-borrowing ([Steiner-Khamsi, 2010](#)), which focuses predominantly on sense-making and legitimization that makes *different modes* of engagement with Europe harder to explain. To my understanding, [Moravcsik’s \(1993\)](#) theory of liberal intergovernmentalism would see England’s ignorance of Bologna as a national play with implications on two-levels, both in terms of signaling internally that the DfE, not Brussels is in charge of education, and internationally that states do not need to harmonize in education (a wish for limited European integration). In France, by contrast, the imaginations of Europe are of a different nature, the image of Europe for French intellectuals is of France as Europe and Europe as France ([Chabert, 2011](#)). The two-level play would then harmonize around sending strong signals to Europe of the need to adopt Bologna and the 3 + 2 model, also in important national aspects such as education. The international concerns do here further nudge the national considerations, not because of what other states do nor only internal (national) concerns, but so France in the future can help push for further European integration in education.

Concluding remarks

To recapitulate and conclude, neither England nor France from their inception existed in isolation from the other, nor from the wider world, even though some accounts may create a semblance of national development existing in isolation prior to the “modern” and “post-modern” eras of globalization. Cross-national diffusion of ideas and engagements with similar cognitive blueprints is nothing new. However, what has changed are the speed of communication and the reach of dissemination. Ideas on education and the teacher is now reachable in every home with an internet connection through sites such as Wikipedia or Google where one can find for example PISA reports and rankings, Eurydice reports on the (believed) layout of an education system, or the book *Émile*.

Meanwhile, as argued in this essay, contrasts must be understood as an interplay between different levels of an environment (local, regional, global). Therefore, one may say the two cases are distinct in so far as one is turning toward the practical and the other is focused on integrating teacher education within the university, but not that they exist in isolation from one and other. This becomes apparent as we see how they engage with ideas of the enlightenment, of democracy and expertise, but in different ways and with distinct emphasizes that form different trajectories of abstract state governing logics and explicit teacher education policies. As such, the key to comprehending teacher education policy, is understanding the local socio-cultural and historic specificities which translates and transforms international trends into local outcomes (see also [Brisard, 2002](#); [Crossley and Watson, 2003](#); [Crossley, 2010](#); [Maroy et al., 2016](#)), arguably leading to what [Maroy et al. \(2016\)](#) label vernacular globalization. One way to analyze the intersection between global, regional, and local trends is through contrasting comparison with an emphasis on the historical developments and particularities of both state governing and education logics and beliefs. In other words, to analyze both ideas of (teacher) education and ideas of good governance and the purpose of the state in a given context, while keeping an eye on the wider contexts (e.g. Europe, the globe) the object of analysis is located in. This approach however, would benefit from a coupling with two-level game type theories ([Putnam, 1988](#)) such as liberal intergovernmentalism ([Moravcsik, 1993](#)), to fully analyze the complex interaction between international and national concerns that policymakers and practitioners in education constantly engage with. This provides further analytical models to comprehending why we see vernacular globalization.

Based on what has been presented in this essay, we can say that the national imaginary vis-a-vis Europe plays an important role in shaping how England and France engage with Europe differently due to both national *and* international considerations that are rooted in the distinct national socio-historic trajectories rooted on the one hand in the French revolution of 1789 which led to a strong educative state, and on the other in a liberal system of education without an overarching educative state. These national myths, as socially constructed imaginaries, can then also be said to fully constitute what Benedict Anderson (2006) called imagined communities where we imagine ourselves to share a life with our compatriots, for France parts of this community extend into “enlightened Europe” (France is Europe), while in England it ends with the Channel Islands, and for some also at Hadrian’s wall and the March of Wales. Interestingly however, England does references the far reaches of the former British Empire such as Singapore and Hong-Kong, but not Scotland or Ireland. At the same time, Bologna becomes an important cognitive blueprint (Zapp et al., 2018a,b) that is accessible for national policymakers in England and France as they try to both figure out how to do teacher education and how to justify it within an “evidence era” paradigm (Helgetun and Menter, 2020) where such blueprints are worth more than their weight in gold. Through understanding such divergent trends under common umbrellas, we begin to flesh out the context sensitivity of not only teacher education policy, but also of the teaching profession and the career of teachers that is intricately tied to the teachers’ education.

Acknowledgment

The project leading to this article has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 714641).

References

- Adamson, B., Forestier, K., Morris, P., Han, C., 2017. PISA, policymaking and political pantomime: education policy referencing between England and Hong Kong. *Comparative Education* 53 (2), 192–208.
- Addey, C., Sellar, S., Steiner-Khamisi, G., Lingard, B., Verger, A., 2017. The rise of international large-scale assessments and rationales for participation. *Compare* 47 (3), 434–452.
- Akiba, M., 2017. Editor’s Introduction: Understanding Cross-National Differences in Globalized Teacher Reforms. *Educational Researcher* 46 (4), 153–168.
- Alexander, R.J., 2001. *Culture & Pedagogy: International Comparisons in Primary Education*. Blackwell publishing, USA: Malden.
- Alexander, R.J., 2010. “World class schools”—noble aspiration or globalized hokum? *Compare* 40 (6), 801–817.
- Anderson, B., 2006. *Imagined communities*. Verso, UK: London.
- Ball, S.J., 2016. Following policy: networks, network ethnography and education policy mobilities. *J. Educ. Policy* 31 (5), 549–566.
- Ball, S.J., Junemann, C., 2012. *Networks, new governance and education*. Policy Press.
- Beck, J., Young, M.F.D., 2005. The assault on the professions and the restructuring of academic and professional identities: a Bernsteinian analysis. *Br. J. Sociol. Educ.* 26 (2), 183–197.
- Beck, J., 2009. Appropriating professionalism: restructuring the official knowledge base of England’s “modernized” teaching profession. *Br. J. Sociol. Educ.* 30 (1), 3–14.
- Bon, A., 2014. *Pavane pour une infante défunte : les IUFM de 1998 à 2009. La formation des maîtres de 1940 à 2010*. A. Prost. Presse universitaire de Rennes, Rennes.
- Brisard, E., 2002. National visions of initial teacher education? The conceptualisation of the initial preparation of secondary teachers in England, France and Scotland. *Scot. Educ. Rev.* 34 (1), 51–60.
- Bromley, P., Powell, W.W., 2012. From smoke and mirrors to walking the talk: decoupling in the contemporary world. *Acad. Manag. Ann.* 6 (1), 483–530.
- Chabert, G., 2011. *Europa, Europa: Historien om en fransk besettelse*. Tapir akademiske forlag, Trondheim.
- Childs, A., Menter, I., 2013. Teacher education in 21st century England. A case study in neoliberal public policy. *Rev. Española Educ. Comp.* 22, 93–116.
- Chung, J., Atkin, C., Moore, J., 2012. The rise and fall of the MTL: an example of European policy borrowing. *European Journal of Teacher Education* 35 (3), 259–274.
- Chung, J., 2016. The (mis)use of the Finnish teacher education model: “policy-based evidence-making”? *Educ. Res.* 58 (2), 207–219.
- Cini, M. (Ed.), 2007. *European Union Politics*. Oxford University Press, Oxford.
- Coffield, F., 2012. Why the McKinsey reports will not improve school systems. *J. Educ. Pol.* 27 (1), 131–149.
- Condorcet, 1792. *Rapport et projet de décret sur l’organisation générale de l’instruction publique : présentés à l’Assemblée nationale au nom du comité d’instruction publique*. Instruction publique. French National Assembly, Paris, p. 5.
- Comu, B., 2015. Teacher education in France: universitization and professionalization—from IUFMs to ESPEs. *Educ. Inq.* 6 (3).
- Crossley, M., Watson, K., 2003. *Comparative and International Research in Education: Globalization, Context and Difference*. RoutledgeFalmer, London and New York.
- Crossley, M., 2010. Context matters in educational research and international development: learning from the small states experience. *Prospects* 40 (4), 421–429.
- Dupriez, V., Cattonar, B., 2018. Between evidence-based education and professional judgement, what future for teachers and their knowledge? In: Normand, R., Liu, M., Carvalho, L.K., Oliveria, D.A., LeVasseur, L. (Eds.), *Education Policies and the Restructuring of the Educational Profession*. Springer.
- Elken, M., Vukasovic, M., 2019. The looseness of loose coupling: the use and misuse of “loose coupling” in higher education research. *Theor. Method High. Educ. Res.* 53–71.
- Farges, G., 2017. *Les Mondes enseignants. Identités et clivages*. PUF, Paris, France.
- Fazekas, M., Burns, T., 2012. Exploring the Complex Interaction Between Governance and Knowledge in Education. *OECD Education Working Papers* 1–43.
- Foster, R., 1999. School-based Initial Teacher Training in England and France: trainee teachers’ perspectives compared. *Mentor. Tutoring* 7 (2), 131–143.
- Furlong, J., 2013. Globalisation, neoliberalism, and the reform of teacher education in England. *Educ. Forum.* 77 (1), 28–50.
- Godin, B., 2006. The Knowledge-Based Economy: Conceptual Framework or Buzzword? *Journal of Technology Transfer* 31, 17–30.
- Hartley, D., 2003. Education as a global positioning device: some theoretical considerations. *Comp. Educ.* 39 (4), 439–450.
- Helgetun, J.B., Dumay, X., 2021. From scholar to craftsman? Constructing an accountable teacher education environment in England 1976–2019. *Eur. J. Teach. Educ.* 44 (1), 80–95.
- Helgetun, J.B., Menter, I., 2020. From an age of measurement to an evidence era? Policy-making in teacher education in England. *J. Educ. Pol.* 1–18.
- Helgetun, J.B., 2021. *Maintaining and Changing Institutionalized Environments: Teacher Education Policymaking in England and France Compared*. Doctor of Science of Education Monographic Dissertation. University of Louvain.
- Helgetun, J.B., 2022. *The Importance of Context: Teacher Education Policy in England and France Compared the Palgrave Handbook of Teacher Education Research*. I. Menter, Palgrave-Macmillan.
- Lawn, M., 2013. A systemless system: designing the disarticulation of English state education. *Eur. Educ. Res. J.* 12 (2), 231–241.

- Lingard, B., Rawolle, S., 2011. New scalar politics: implications for education policy. *Comp. Educ.* 47 (4), 489–502.
- Maroy, C., Pons, X., Dupuy, C., 2016. Vernacular globalizations: neo-statist accountability policies in France and Quebec education. *J. Educ. Pol.* 32 (1), 100–122.
- Menter, I., 2016. Introduction. In: Beauchamp, G., Clarke, L., Hulme, M., et al. (Eds.), *Teacher Education in Times of Change*. Bristol Policy Press, University of Bristol, pp. 3–18.
- Meyer, H.-D., Benavot, A., 2013. PISA and the globalization of education governance: some puzzles and problems. In: Meyer, H.-D., Benavot, A. (Eds.), *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Symposium Books, Oxford: UK.
- Meyer, J.W., Ramirez, F.O., 2007. World Institutionalization of Education. *Weltkultur und kulturelle Bedeutungswelten*. J. Schriewer 279–298.
- Meyer, J.W., Boli, J., Thomas, G.M., Ramirez, F.O., 1997. World society and the nation state. *Am. J. Sociol.* 103 (1), 144–181.
- Misra, P.K., 2014. The state of teacher education in France: a critique. *Fondation Maison Sci. l'homme* 1 (58), 1–24.
- Moravcsik, A.M., 1993. Preferences and power in the European community: a liberal intergovernmentalist approach. *J. Common. Mark. Stud.* 31 (4), 473–524.
- Moravcsik, A.M., 1997. Taking Preferences Seriously: A Liberal Theory of International Politics. *International Organization* 51 (4), 513–553.
- Morley, J., 1878. Rousseau, Kindle edition. online.
- Normand, R., 2013. The doubts and uncertainties of french educators in the face of traveling policies. In: Seddon, T., Levin, J.S. (Eds.), *Educators, Professionalism and Politics. Global Transitions, National Spaces and Professional Projects*. Routledge, New York.
- Ozga, J., 2009. Governing education through data in England: from regulation to self-evaluation. *J. Educ. Pol.* 24 (2), 149–162.
- Ozga, J., 2019. The politics of accountability. *J. Educ. Change* 21, 19–35.
- Ozouf, J., Ozouf, M., 1992. *La République des instituteurs*. Seul, Paris.
- Paine, T., 1995. *Rights of Man, Common Sense and other political writings*. Oxford University Press, USA: New York.
- Phillips, R., 2001. Education, the state and the politics of reform: the historical context, 1976–2001. In: Phillips, R., Furlong, J. (Eds.), *Education, Reform and the State: Twenty-Five Years of Politics, Policy and Practice*. Routledge, London and New York.
- Powell, J.J.W., 2020. Comparative education in an age of competition and collaboration. *Comp. Educ.* 56 (1), 57–78.
- Prost, A., 2013. *Du changement dans l'école: Les réformes de l'éducation de 1936 à nos jours*. Seuil, Paris, France.
- Prost, A. (Ed.), 2014. *La formation des maîtres de 1940 à 2010*. Presses universitaires de Rennes, Rennes, France.
- Putnam, R.D., 1988. Diplomacy and domestic politics: the logic of two-level games. *Int. Organ.* 42 (3), 427–460.
- Quirk-Marku, C., Hulme, M., 2017. Re-making teacher professionalism in England: localism and social responsibility. *RASE* 10 (3).
- Resnik, J., 2006. International organizations, the “Education-Economic Growth” black box, and the development of world education culture. *Comp. Educ. Rev.* 50 (2), 173–195.
- Robertson, S.L., 2012. Placing Teachers in Global Governance Agendas. *Comparative Education Review* 4 (56), 584–607.
- Rosenberg, A., 1981. Translating Rousseau, 50. *University of Toronto Quarterly*, pp. 339–341.
- Sacilotto-Vasylenko, M., Fave-Bonnet, M.F., 2011. Teacher education in France: persistent tensions. In: Valenčič Zuljan, M., Vogrinc, J. (Eds.), *European Dimensions of Teacher Education—Similarities and Differences*. Faculty of Education, Ljubljana. Kranj: The National School of Leadership in Education.
- Saunders, L. (Ed.), 2007. *Educational Research and Policy-Making: Exploring the Border Country Between Research and Policy*. Routledge, London & New York.
- Sayer, J., 2006. European perspectives of teacher education and training. *Comp. Educ.* 42 (1), 63–75.
- Schofer, E., Meyer, J.W., 2005. The worldwide expansion of higher education in the 20th century. *Am. Soc. Rev.* 70 (6), 898–920.
- Sellar, S., Lingard, B., 2013. The OECD and global governance in education. *Journal of Education Policy* 28 (5), 710–725.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. *British Educational Research Journal* 40 (6), 917–936.
- Simões, A.R., Lourenço, M., Costa, N. (Eds.), 2018. *Teacher Education Policy and Practice in Europe Challenges and Opportunities for the Future*. Routledge, London and New York.
- Sørensen, T.B., 2016. Teachers and the Global Educational Policy Field. In: Jules, T.D. (Ed.), *The Global Educational Policy Environment in the Fourth Industrial Revolution*, pp. 59–84.
- Sørensen, T.B., Dumay, X., 2021. The teaching professions and globalization: a scoping review of the Anglophone literature. *Comp. Educ. Rev.* 65 (4).
- Steiner-Khamsi, G., Karseth, B., Baek, C., 2019. From science to politics: commissioned reports and their political translation into White Papers. *Journal of Education Policy* 1–26.
- Steiner-Khamsi, G., Stolpe, I., 2006. *Educational Import: Local Encounters With Global Forces in Mongolia*. Palgrave Macmillan, New York.
- Steiner-Khamsi, G., 2010. The politics and economics of comparison. *Comp. Educ. Rev.* 54 (3), 323–243.
- Tatto, M.T., 2011. Reimagining the Education of Teachers: The Role of Comparative and International Research. *Comparative Education Review* 55 (4), 495–516.
- Tatto, M.T., Richmond, G., Carter-Andrews, D.J., 2016. The research we need in teacher education. *J. Teach. Educ.* 67 (4), 247–250.
- Tröhler, D., 2013. The OECD and Cold War Culture: thinking historically about PISA. In: Meyer, H.-D., Benavot, A. (Eds.), *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Symposium Books, Oxford: UK.
- Tröhler, D., 2014. The medicalization of current educational research and its effects on education policy and school reforms. *Discourse* 36 (5), 749–764.
- Tröhler, D., 2020. Learning progress and the taming of change: the educational aspirations of the age of enlightenment. In: Tröhler, D. (Ed.), *A Cultural History of Education: In the Age of Enlightenment*. Bloomsbury, London and New York, pp. 1–24.
- UNESCO, 2005. *Guidelines and Recommendations for Reorienting Teacher Education to Address Sustainability*. UNESCO, Paris.
- Vögtle, E.M., Knill, C., Dobbins, M., 2011. To what extent does transnational communication drive cross-national policy convergence? The impact of the bologna-process on domestic higher education policies. *High. Educ.* 61, 77–94.
- Whiting, C., Whitty, G., Menter, I., Black, P., Hordern, J., Parfitt, A., Reynolds, K., Sørensen, N., 2018. Diversity and complexity: becoming a teacher in England in 2015–2016. *Rev. Educ.* 6 (1), 69–96.
- Wilkinson, H., 2019. *The European Semester: A Literature Review*.
- Ydesen, C. (Ed.), 2019. *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*. Palgrave Macmillan.
- Zapp, M., Helgetun, J.B., Powell, J.J.W., 2018a. (Re)shaping educational research through “programmification”: institutional expansion, change, and translation in Norway. *Eur. J. Educ.* 53 (2), 202–217.
- Zapp, M., Marques, M., Powell, J.J.W. (Eds.), 2018b. *European Educational Research (Re)constructed: Institutional Change in Germany, the United Kingdom, Norway, and the European Union*. Symposium books, Oxford, England.
- Zapp, M., 2020. The authority of science and the legitimacy of international organizations: OECD, UNESCO and World Bank in global education governance. *Compare* 1–20.

Teachers' research capacity and initial teacher education policy

Clare Brooks, UCL Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	611
The value of research within teacher education	612
Four trends affecting research capacity	612
The diminishing role of universities in initial teacher education	613
Pre-service teachers and their expectations of teaching as a research-based profession	613
Teacher education accountability structures and the significance of research	614
Career trajectories and contractual arrangements of teacher educators	614
The ITT Market Review in England	614
The Welsh National Strategy for Education Research and Enquiry	616
Aim	617
Objectives	617
Enabling or constraining research capacity	619
Acknowledgment	620
References	620

Introduction

In 2021, a Special Issue of *Teaching Education* took its theme from Lawrence Stenhouse's statement that research should look less to the world and more to the village (1981): as a call for more teacher engagement in the world of education research. In my contribution to that special issue (Brooks, 2021b), I argued that the profession's capacity to contribute to research was being diminished due to the reduced capacity of research in initial teacher education: in short that research was being squeezed out of teacher education. My argument was based on four trends that I observed while conducting an international comparison of teacher education programs in five universities (located in Australia, Canada, England, New Zealand and the US). I outlined these four trends as follows:

The first [trend] is that the role of universities within teacher education is precarious, along with the perceived value of research in teacher education. The second argues that prospective teachers do not perceive teaching as a research-based profession and therefore expect teacher education to be practice-orientated. Teacher education accountability structures downplay the significance of research as part of a teacher's knowledge base and practical repertoire. And finally, the career trajectories and contractual arrangements of teacher educators reduces the overall research capacity of the field.

Brooks (2021b).

The combination of these four trends presented a diminishing capacity for (new) teachers to be inducted into their potential engagement with educational research, which may limit their ability to engage meaningfully in research later in their careers.

Tatto (2021) has also argued for a greater focus on research in teacher education as an important counter-balance to some of the value-added methods that are increasingly driving education policy and practice. In her paper, she outlines a growing division between the practice of teaching and educating teachers and the research that influences education and in doing so, asks important questions about the role of research within teacher education:

Who should produce the knowledge that must be used to guide policy and practice in the teaching profession? What conditions enable and constrain teachers from engaging in such research? How does generative and transformative teacher research sit alongside structures and cultures of professional compliance?

Tatto (2021).

In my paper, written during a period of COVID-19 related lockdown, I concluded on an optimistic note, recognizing a range of changes that were occurring across the education sector and the potential they provided for these trends to be reversed. Prompted by the need for immediate and swift action in the light of the pandemic, many governments were turning to "experts" for advice about how to make adaptation in many areas of life, as such research seemed to be having a renaissance. Teacher education was no different. Recognizing the need to continue to educate teachers to ensure teacher supply, and the very real challenges that lockdowns and school closures presented for doing so, in many places teacher educators were consulted and given new levels of autonomy to make adaptations to their programs and practices. Such changes, I argued, could be seen as the start of a shift toward enabling

teacher educators to redefine their practices in a less controlled environment and as such the role of research within teacher education could be foregrounded, which could enhance the research capacity of new teachers.

In this chapter I build on the arguments I have outlined above, and reflect on whether my optimism was warranted. My experience, based in England, has been dominated by the Department for Education's ITT (Initial Teacher Training) Market Review report published in July 2021 ([Department for Education, 2021](#)), and the public consultation which followed. My experience suggests that my optimism was indeed unfounded, and that the future of teacher education in England with respect to research capacity is likely to be diminished further. Alternatively, I have been interested in the developments in Wales, where the publication of the National Strategy for Education Research and Enquiry (2021) presents a very different vision for the role of research, and indeed for building the research capacity for research in teaching and in teacher education. Using the four trends highlighted in the 2021 paper as an analytical framework, I explore these two different approaches and their likely impact on new teachers' research capacity. In doing so, I argue that policy is a significant force in defining the importance of research in early teacher development and whether research is seen to make a valuable contribution to education development as a whole.

The value of research within teacher education

While there are few detractors about the value of research in teacher education, and indeed much to suggest that is a strong feature of teacher education in so-called high performing systems, ([Darling-Hammond, 2006](#); [Darling-Hammond and Lieberman, 2013](#); [Darling-Hammond et al., 2000](#)), there is some debate about the ways in which research contributes effectively, whether it is seen as "content": making up the knowledge base for teacher education through research-informed or evidence-based findings about ways to teach; or as a form of teacher education pedagogy: using research as a way of learning about teaching both in the initial stages but providing a pedagogy of professional development and improvement throughout a teacher's career.

In practical terms, how research is positioned on teacher education programs varies, as outlined by [Menter et al. \(2010\)](#) who categorized four forms of teacher professionalism showing how research can be positioned differently within them: for example:

- The effective teacher: with an emphasis on meeting standards and competences, positions research as content to be digested and understood;
- The reflective teacher: with an emphasis on individual professional development achieved through practice positions research as being a tool with which to challenge assumptions based on experience;
- The enquiring teacher: adopts an enquiry approach, which in some cases has veered toward a research orientation
- The transformative teacher: which adopts an activist stance in relation to enquiries in order to contribute to social change and sees research as emancipatory.

Research can therefore be seen as a way of challenging assumptions about teaching, as a way of investigating and inquiring into (personal) practice, but also as a way of challenging issues of equity and justice that exist within education, and providing a range of ideas possible for transformation. The BERA-RSA inquiry into research-informed clinical practice also discusses the range of different approaches to embedding research across ITE, and the impact research can have on teacher's professional learning ([Burn and Mutton, 2013](#)). How new teachers experience research can vary enormously.

Four trends affecting research capacity

In 2018–2020, I conducted a study of five large scale ITE providers in order to research the features of high-quality large-scale initial teacher education. Even though the universities I studied were situated in different contexts (Queensland University of Technology (QUT) in Australia; Ontario Institute for the Study of Education (OISE), University of Toronto, in Canada; University of Auckland, in New Zealand; UCL Institute of Education (IOE) in London, UK and Mary Lou Fulton Teachers College, Arizona State University (MLFTEC), USA), the role and function of research within their programs was a dominant theme. All the universities valued and prioritized research and saw it as an important aspect of their practice: but within that framing they all encountered very similar challenges and concerns:

- What was the best combination of research and practice within program design?
- What sort of research should new teachers be inducted too, how much research methodology did they need to understand, and to what extent was research literacy a key aspect of the assessment of their programs?
- What were the research expectations of teacher educators? How were they supported to be research active? What were the university expectations around research, and how does adopting a research perspective sit with the expectations of being a teacher educator?

A further common theme was the role of research in the accountability structures that were pertinent to teacher education, and the extent to which research was valued as part of a teachers' initial professional development. The answers to these questions reflected the different program structures and designs and the extent to which being research-aware, research literate, and even research-active was an expectation of new teachers in those contexts. Upon reflection, these contextual perspectives are perhaps one of the most influential dimensions that affect research capacity of new teachers. Throughout the research I observed teacher educators adapting

their practice in line with local knowledge and expectations, and the capacity they themselves were given to enact change (as reflected in my model of ITE practice outlined in [Brooks, 2021a](#)). However, the reason why teacher educators were needing to adapt their practice was often because of the policy context, and the climate that this created for the possibility for research. The policy context then can be seen as a key component that can enable or constrain research practices in initial teacher education.

Two recent developments have provided a good opportunity to explore this phenomenon in more detail and to exemplify how the policy context creates the conditions for research capacity. In England, under the banner of an ITT Market Review, the government's Department for Education appointed Expert Group published a series of recommendations. The subsequent public consultation prompted a range of widespread criticisms of this report by a range of stakeholders: a key theme of which was the role of evidence and research as presented in the report. Alternatively, in Wales, the National Strategy for Education Research and Enquiry foregrounds research, and promotes partnerships between universities and schools in the education of new teachers. Published in July 2021, the challenge for teacher educators in Wales is the roll out of this strategy, and to consider the implications for schools, teacher educators and new teachers to transform teaching into a research-based profession at all levels. The analysis for this paper is based on a review of relevant literature for both, which included the policy documentation, published blogs and reports and consultation documents in response, and also interviews with university-based teacher educators with responsibility for rolling out these policy initiatives. I have sought to summarize the emerging trends as I see them, but as the documents are at different stages of agreement and implementation, the trends themselves are early in formation and the impact has yet to be realized. The intention here is to see them as indicative of different policy approaches which can highlight how research capacity can be diminished or enabled and to consider what the implications of that may be. After a brief introduction to the four trends outlined in the previous paper, I will explore each of the approaches in turn and how they appear to enable or constrain research capacity.

The diminishing role of universities in initial teacher education

In my previous paper, I argued that while in many countries around the world, universities are still the main provider of initial teacher education, the role of universities has diminishing in prominence. Citing evidence on the rise of non-university-based routes into teaching (in particular the growth of the Teach First/Teach for All movement), and the increasing interest in non-university-based routes, some of which were school- or employment based, I noted that in some cases this means that routes into teaching might have little or no research content. Indeed, the growth of non-university-based research organizations represents a changing dynamic between universities and their (taken-for-granted) role in initial teacher education, and in particular could affect the role of research in the pedagogy and content of teacher education programs.

For proponents of a school-based teacher education system, the diminishing role of universities is not perceived as a problem. The role of universities has always been problematic, both from the view of universities themselves, where teacher education was considered a weak discipline with a poor track-record of research ([Labaree, 2006](#)), but also from the idea that a university-based teacher education was not sufficiently grounded in practice, and was too theoretical, abstract and irrelevant. While this is a fairly caricatured image of university-based teacher education that does not reflect the variety of knowledge traditions it encompasses ([Furlong, 2013](#); [Furlong and Whitty, 2017](#)), it is a pervasive account often used to discredit university-based providers (for a detailed account of this in the US, see [Zeichner, 2017](#)).

Research funding in England for teacher education has always been limited. This has been exacerbated by the changing dynamic of education research funding in England toward the Education Endowment Foundation (EEF), which funds a narrow range of research described as overly instrumentalized, dehumanizing and reliant on a narrow range of research methodologies such as Randomized Control Trials ([Ellis et al., 2018](#); [Godfrey, 2017](#); [Hulme et al., 2020](#); [Turvey, 2018](#); [Turvey and Pachler, 2020](#)). At the same time, there has been a growth of teacher-research orientated movements such as ResearchEd which focus on practitioner-to-practitioner research accounts, and which has received support from key politicians, while many of its proponents are highly critical of research conducted by higher education institutions ([Watson and Barnes, 2021](#)). [Tatto \(2021\)](#) also notes the influence of value-added methods and econometric based analysis on education policy, and how these tend to come from different sources than university-based teacher educators.

Universities, however, are still central to knowledge generation and its validation and verification ([Furlong and Whitty, 2017](#)) which make up the specialist knowledge necessary for the recognition of a profession, but their role in education research is diminishing. However, [Young and Muller \(2014\)](#) have argued that such a perspective is necessary to counter the anti-intellectualism of professional development, often found in fields which emphasize expertise over knowledge. Despite this academic recognition, the reality of teacher education in many international settings is an increasing erosion of the perceived importance of the university in defining the knowledge base for teacher education, or indeed as the main context for qualification to become a teacher.

Pre-service teachers and their expectations of teaching as a research-based profession

A key factor in the success and impact of research in teacher education lies in how it is perceived by new teachers. New teachers enter the profession with an image of what being a teacher will feel like and an expectation of what sort of preparation they will need to become a teacher. These expectations will vary by individual but are influenced by public images of teachers ([Moore, 2004](#); [Henry, 2020](#)) and also their own experience of having been taught ([Lortie, 1975](#)). Despite an emphasis on research and enquiry in many of the programs I researched ([Brooks, 2021a,b](#)), there was still evidence that new teachers did not perceive teaching to be a research-based profession or that undertaking research was a valuable or essential component of their programs. I argued that this was linked

to a strong public discourse about teaching as a practical activity requiring technical skills or behaviors rather than a propensity to undertake or be engaged with research. This discourse is so strong that new teachers, keen to be the best they can be, do not expect to see research as part of either their initial training or as a part of their role as a qualified teacher.

Teacher education accountability structures and the significance of research

Accountability structures shape activities, even when situated as minimal requirements. When those accountability structures downplay the role of research, and situate the (new) teacher as a passive consumer of research: this can be seen as part of the "what works" discourse that has been critiqued by [Biesta \(2007\)](#) as profoundly uneducational, in particular because it lacks the nuance required for situational judgements. This was in contrast to the teacher educators I spoke to in my research who were passionate about the role of research on their programs not just as content but as a form of professionalization, best described by the Cochran-Smith and Lytle phrase "inquiry as stance" ([Cochran-Smith and Lytle, 1999](#)).

Acknowledging the pervasiveness of how and why research is considered important in teacher education across different contexts, how research is then presented and discussed in key accountability documents can send a different message. Verb constructions in key documents that influence teacher education in England situate the relationship between the teacher and the research as a passive one; where the teacher "draws on", is "informed by" and can "demonstrate" their knowledge of research (2021). The idea of a teacher as researcher or as a critical consumer of research are absent.

Career trajectories and contractual arrangements of teacher educators

University-based teacher educators tend to come from two backgrounds: either having had previous experience as an academic researcher, or as enrolled on a doctoral program ([Cochran-Smith et al., 2019](#)), or as a former teacher who has to undergo an identity shift in order to transition into their new role ([Trent, 2013](#); [Williams, 2014](#)). For some, teacher education was not their intended career goal, and it is commonly referred to as *second-order* ([Murray and Male, 2005](#)), which reflects to an extent the lack of systematic professional development they receive ([Holme et al., 2016](#)).

In contrast to teacher educators who are based in school contexts, university-based teacher educators are part of the university faculty, and so also need to address the expectations of being an academic and the precarious nature of academic contracts. These concerns are typified by what Grudnoff ([Cochran-Smith et al., 2019](#)) has called the bifurcation of their roles: as they juggle being both academics and teacher educators. [Ellis and McNicholl \(2015\)](#) have noted the heavy workload involved in university-based teacher education (particularly when they include time-consuming school visits), and the tensions that the teacher educators face in handling research expectations as a faculty member. [Cochran-Smith et al. \(2019\)](#) highlights the challenges this presents for the teacher educators, including:

- Their ability to be recognized within their institution as researchers and academics;
- The induction and professional development they need in order to become high-quality teacher educators;
- The dissonance between the dual expectations of them as teachers of teachers and as academics and researchers.

This is further complicated by the lack of status and respect offered to education faculty and the work they undertake in the academy ([Labaree, 2006](#)). This has been compounded by changing working conditions ([Ellis and McNicholl, 2015](#); [Furlong, 2013](#)), increasingly orientated around temporary and casualized employment contracts, reducing the opportunity for professional development and career advancement.

These four trends make up a context of diminishing research capacity within teacher education, and by extension then for new teachers. This context, however, is largely driven by education policies which outline the possibilities for research to occur in each of these four areas: in the eyes of new teachers, in the careers of teacher educators, the accountability structures, and in the role given to universities in the teacher education landscape. The next two sections discuss how two contrasting policy contexts create the conditions that can enable or constrain research within teacher education.

The ITT Market Review in England

The ITT Market Review was undertaken by an Expert Group convened by the Department of Education (DfE) of the British Government during 2020/21. The ITT Market Review applies only to teacher education (or teacher training (ITT) to use the DfE's vernacular) in England, as this responsibility is devolved to respective governments in Scotland, Wales and Northern Ireland. The review is a response to the highly differentiated landscape of teacher education (taught through a combination of university led and school-based routes) that has emerged from previous policy initiatives, leading to a teacher education "market" that is described as being both diverse and complex ([Whiting et al., 2018](#)).

During the period between the establishment of the Market Review and the publication of its findings, there were a number of other related activities: Ofsted, the government's inspectorate, published a new ITT Inspection Framework, and a controversial and critical report into how ITT providers have responded to the pandemic. Teaching School Hubs, a major initiative in the teacher training infrastructure were announced, as was the national roll out of a suite of government-owned professional development programs (the Early Career Framework (ECF), and National Professional Qualifications (NPQs)), and the continuation of the

High Potential Initial Teacher Training contract awarded to Teach First now in its 19th year. The tender process for a new government-initiated flagship teacher training institution, the Institute of Teaching, had also begun. In other words, the landscape of ITT was changing and changing rapidly. With new players being introduced to the "market" and the fundamentals of that market changing substantially.

The Market Review group published their report in June 2021, which consisted of 14 recommendations, and a list of Quality Requirements which would constitute an accreditation process for all providers. In this document, the main aim of the review is described:

The central aim of the review and our recommendations is to enable the provision of consistently high-quality training, in line with the CCF, in a more efficient and effective market.

The CCF referred to above is the government's ITT Core Content Framework published in November 2019, which "defines in detail" the content of initial teacher training programs as a "minimum entitlement" which the DfE describes as drawing upon "best available evidence" although this claim is disputed. The Market Review focusses heavily on evidence (evidence is mentioned 88 times in the document, research: 40 times), which is taken from a range of government documents related to teacher education: the ITT Core Content Framework (CCF), the Early Career Framework and the National Professional Qualifications, each of which has been verified by the Education Endowment Foundation (the CEO of which was on the Expert Group that authored the ECF). Where research is mentioned in the document it is positioned as a mechanism for which "evidence" is procured and produced. For example, in the first paragraph of the document it is stated:

We must ensure that teachers receive high-quality training and professional development at every stage of their career, and fundamental to this is initial teacher training (ITT). This means ensuring that all teacher professional development, from initial training through to advanced leadership development, systematically and consistently draws on the knowledge we now have about teaching from up-to-date and high-quality research.

The emphasis in the recommendations on evidence rather than research positions the teacher passively in terms of their (direct) engagement with research. This has been a recurring theme in the English education system, promoted initially in the Carter Review of Initial Teacher Education (2015) and critiqued in the Burn and Mutton report for the BERA/RSA inquiry into research and teacher education (Burn and Mutton, 2013, 2015). The Carter Review argued that there are limitations on trainees' abilities to engage fully in research, and positioned them as "intelligent consumers of research who take an evidence-based approach to their own practice" (Carter, 2015, p. 21). The review recommends the creation of "synthesized executive summaries providing practical advice on research findings about effective teaching in different subjects and phases" (Carter, 2015, p. 8), which places the teacher as a passive consumer of research who is also unable to make selections about the sorts of research they wish to engage, and places teacher educators as the curators of such executive summaries. This perspective privileges a particular type of "what works" research which has been widely criticized for its lack of contextual detail, ethical reflections and transferability (Biesta, 2007).

However, the current policy landscape takes this one step further. The publication of the document "Delivering World Class Teacher Development" in June 2021 by the DfE includes an Annex on Delivery infrastructure which shows diagrammatically the relationship between the DfE, providers and other organizations. Placing the DfE at the head of the diagram the document, the diagram is annotated as:

The DfE created evidence informed frameworks, validated by the Education Endowment Foundation (EEF).

In other words, this infrastructure not only places the teacher as a consumer of previously curated research, but also places the DfE as the curator of that research or evidence. This is a distinct move away from the Carter Review which, while promoting the teacher as a consumer model, clearly situated the teacher educator as the curator of what was deemed appropriate evidence.

In addition, there is a distinct absence of universities in the Market Review Report (only two specific mentions in the whole document and in neither case as a provider of teacher education). However, what is perhaps of more significance is that the evidence cited in the Market Review document itself shows a preference for research generated outside of universities. The Market Review document contains citations of:

- 4 articles from peer-reviewed journals
- 4 reports from non-university organizations, and
- 18 government publications.

The dominance of research from non-university sources is indicative of a further diminishing of the perceived role of universities in their contribution to teacher education, but also in their role in generating research for that community. The preferred emphasis is on government-sponsored publications.

Curating research and presenting it as content is of limited benefit for new teachers, as it renders them unable to consider practical implications of implementing research findings or the ability to judge the merits of research on its methodological rigor or contextual or situational appropriateness. However, the ITT Market Review recommendations goes further by not just outlining the content of teacher education programs but also the ways in which that content should be presented to new teachers:

- the assessment regime of trainees has to be orientated on the content of the CCF (and only on their ability to demonstrate the Teacher Standards at the end of the program),
- the programs have to be planned in relation to a pre-defined set of sequenced intensive experiences (as opposed to be in line with student teacher needs or requirements),
- and the training of mentors and their monitoring aligned with pre-determined conditions and parameters.

In other words, the proposals outline the need for a specific, pre-determined and pre-set “sequenced curriculum” as a precondition for re-accreditation. The room for accommodation, deviation and individualization is limited, as is involvement of school partners, or student teachers in the construction of that curriculum (compounded by the very short timescales indicated). Under such a regime, it is likely that new teachers will remain passive as others curate research for them to consume and apply.

This is further compounded by the proposed accountability structures. Under the Recommendations all providers need to be reaccredited for which they would need to demonstrate that they have met a series of Quality Requirements. Not only do these requirements foreground the government's own alternative to universities, the proposed Institute of Teaching, but it also outlines that the accreditation process would involve assessment of the following:

- the development of a selection of detailed curriculum planning, including at subject or phase-specific level,
- the development of at least part of the mentor training curriculum,
- demonstration that the provider has in place the fundamentals of a workable delivery chain, including key ITT lead partners or other key parts of the delivery structure,
- a sample of the curriculum-aligned assessment framework to be used, and
- sufficient detail on quality assurance arrangements to provide assurance that the approach will meet the demands of the Quality Requirements.

The Quality Requirements detail components of the anticipated planned curriculum for new trainees for example:

Providers must identify how all components of the planned curriculum will be taught, applied to practice in a range of contexts and assimilated. The curriculum should encompass a variety of approaches, including direct explanation, deconstruction, structured and focused observation and targeted practice with systematic analysis, feedback and mentoring. At all times, the planned and sequenced curriculum must closely inform taught components, independent learning, practice and feedback.

Further detail is provided on how this is to be achieved, and where new teachers are expected to be engaged with research, that research again has been curated for them by official government organizations, such as the inspectorate Ofsted:

Subject-specific approaches must be delivered by suitably qualified experts and take full account of the evidence available for subject-specific teaching, for example Ofsted subject research reviews.

Research that has been previously curated by official government bodies has then become the content for teaching programs. The notion of inquiry, or research informed clinical practice as stance is not just absent but has been completely side-lined. The outcome of such an approach will be for new teachers to perceive themselves as passive consumers of research that has been curated for them, and presented to them for consumption and implementation. Their capacity for research will be diminished even further.

Finally, the English ITT Market Review Recommendations are silent on the issue of teacher educators, but does make extensive reference to “providers”, and “mentors”, including the provision for a monitoring role given the title of “lead mentor”, and for those that undertake these roles to follow only government-approved development programs. The review also refers to “experts”, and “subject experts” but the attribution of this role to either school-based or university-based teacher educators is not defined. There is stipulation of the time that should be allocated to certain tasks and the role of schools to ensure that such requirements are met. The impact on teacher educators themselves is not considered.

The Welsh National Strategy for Education Research and Enquiry

Developed in response to disappointing educational performance in Wales, the opportunities afforded by devolution, and a desire to empower the teaching profession, the Welsh government announced a new vision for teacher professionalism, one that is encapsulated in the National Strategy document, published in June 2021. The Welsh National Strategy document consists of

Aims, Objectives, Priorities and Domains, and outlines the necessary national infrastructure, including the roles for higher education, researchers, collaborations, Research Centers and expectations of the profession itself, and the provision of Schools as Learning Organizations. The Aims and Objectives are outlined as:

Aim

The aim of the NSERE is that educational policy and practice in Wales should be informed by the best available research evidence and disciplined enquiry undertaken by educational professionals.

Objectives

The objectives of the NSERE are to:

- further support the development of evidence-informed policy through engagement with a wide range of research by Welsh Government and other parts of the educational system in Wales
- further develop high-quality research capacity and volume (that uses a range of methodological approaches) within our HEIs that can help serve the needs of the Welsh education system
- develop an evidence-informed profession in Wales where leaders, teachers, support staff, LA and regional consortia staff and other educational professionals use high-quality research evidence and are provided with opportunities to take part in robust professional enquiry
- contribute to and learn from international research and evidence.

The National Strategy also set priorities that were aligned with other Welsh government education initiatives and recognized that different “interdependent domains” that would need to support this infrastructure. The Welsh National Strategy, as [Furlong \(2019\)](#) outlines, has to be seen as part of an integrated education strategy in Wales designed to improve the quality of education provision that also includes teacher-led school and curriculum reform. Similar to the ITT Market Review document, the National Strategy places a lot of emphasis on evidence, although as the analysis below will outline the emphasis and orientation is quite different.

The distinction between the different roles and functions of research in initial teacher education is made explicit in the Welsh National Strategy, which distinguishes between research and enquiry, and makes a clear commitment to both:

We recognise that “research” can encompass a spectrum of activities that utilise research methods. We do not subscribe to a view that research is something that can only be done by professional educational researchers and not by educational professionals. We are interested in the concept of a “continuum” of development whereby educational professionals move from being professional enquirers to becoming teacher researchers. We wish to learn from developments in this area in other professions, especially the health profession.

The National Strategy also supports the idea of teachers as both consumers and producers of research:

The developments we propose to undertake will focus on enabling practitioners to be both consumers of high-quality research and producers of professional enquiry that is informed by rigorous research methods. This is the reason why we choose to have both “research” and “enquiry” included in our strategy, so that the approach we take in Wales is inclusive and “close to practice.”

The National Strategy also acknowledges the difference between research and enquiry while emphasizing that they are “of equal value and interdependent”:

- “Academic research” is a process of investigation leading to new knowledge. It is published so that others can learn from and critique it. To be regarded of high quality it should be significant, original and rigorous. To that end, it should be peer-reviewed before it is published. This type of research is usually undertaken by academics, those studying for higher degrees and professional researchers working in the government and independent sectors.
- “Professional enquiry” is usually undertaken by practitioners within their workplace as a way of identifying problems, establishing causes, finding solutions, evaluating practice and achieving improvement. If it is to be of value it should utilize action research approaches that are systematic, cyclical and that emphasize the collection of evidence.

The National Strategy restates a commitment to universities being involved in teacher education and the importance of enhancing the research capacity of the universities to enable them to do that. The National Strategy outlines that it will do that through:

- Build[ing] upon existing strengths within the sector, including existing collaborative research structures

- Ensure[ing] that the intent expressed in the accreditation criteria for new initial teacher education (ITE) programmes in Wales (that all staff should become research active) is fully supported, allowing more of them to be included in future Research Excellence Framework (REF) exercises than has previously been the case (pages 14/15).

There is recognition that the capacity for education research in Welsh universities needs support, and also that there is a need for collaborative and supportive structures to help to develop that: the document does not presume that the universities are currently well-placed to fulfill this vision immediately, however, the intention for universities to be an important and enhanced part of the teacher education infrastructure is clear and linked to the ways in which research have been constructed within the National Strategy.

The National Strategy builds on the new Accreditation Criteria (established in 2017) upon which current ITE programs are already based, and which required all providers to be reaccredited through partnerships of both universities and schools. The Welsh Accreditation Criteria for initial teacher education was specific in this regard, laying down not only the requirements for partnerships between universities and schools but also in the role and function of research within the teacher education programs. [Furlong \(2019\)](#) aligns this approach with what [Burn and Mutton \(2015\)](#) have described as "research informed clinical practice": which they distinguish from "research-informed" practice or professional learning seen as trial and error, or reflection based on experience. This approach has implications for schools as they are given the task of providing systematic and structured learning opportunities, as well as implications for universities who have to provide access to research, to theories and to knowledge about excellent pedagogical practice ([Furlong, 2019](#)). In addition, he notes that this will also change how universities and schools will work together "to ensure that there is scope for the professional knowledge of both to be adequately represented, and how that will influence joint planning, supervision and of course assessment and accountability."

The reality of course is that such an approach takes time to embed and achieve, and will no doubt be differently expressed across the accredited providers in Wales. However, the intent to have a system based on research-informed clinical practice certainly has policy support in Wales, even though there are still questions about what is meant by the division between research and professional enquiry, as one teacher educator described:

I think the things that we're coming up against is "what is research?". It's sort of all fallen now under this umbrella of professional inquiry, and whether it needs to be redefined for the practicing teacher. Are we expecting peer reviewed papers, you know, that kind of academic research, or are we going with a more practice led, professional accounts? So, people are kind of muddling through, to be honest. And I think one of the risks is we'll lose our integrity, if we lose the peer reviewed kind of research, and the kind of research that you can submit to REF, we will start to lose the academic field if we're not careful.

The recent experience of ITE providers indicates that while the intention for a research orientation is strong in the policy, there is still some way to go in relation to changing pre-service teachers' expectations. As one colleague explained, they have to recruit teachers locally and those teachers are encultured into a previous learning culture:

In the selection criteria will look for things like openness, willingness to learn all those kinds of things. But mostly we have to recruit very locally. So, we are recruiting the product of a somewhat broken education system. They [the new recruits] are all cut from the performativity culture: "just tell me what to do. Tell me what to write" and all that kind of thing.

The new vision for education in Wales means that during the teacher education process, and through the research-orientation required in the new programs, the teacher educators are seeing a shift in the perspectives of these new teachers as they adapt to the demands of the program:

In week one of the programs, we do assumptions, beliefs, stereotypes, and we are aiming to basically change their mindsets from week one They start off with having those myths exploded and their expectations shifted. And then, as they develop their understanding and experience, by the end of it, you see quite a shift.

The basis for this shift is aligned with the focus in this particular program on four research dispositions which make up an important part of the program and are mapped across both the undergraduate and postgraduate programs:

- being skeptical;
- being ethical;
- being part of an inquiry profession, and;
- having the skills of a researcher or being a researcher,

These are broken down into dispositions and behaviors and mapped across the modules and developed through a professional learning passport and supported by other aspects of the course including professional learning conversations. It is through this

program infrastructure that the program intends to shift the expectations of the student teacher into one of a more “researcherly disposition”, which is then reflected in how they are able at the end of the program to “confidently talk about their decisions in the classroom.”

While the outcomes of the Welsh approach have yet to be fully realized, the progress so far does seem to suggest that a transformation is taking place in how the qualifying teachers at the end of their initial education are engaging with research, particularly as they work in research-engaged schools that can support their understanding of the role and function of research in relation to their practice.

Importantly, a key aspect of the Welsh National Strategy is the recognition of the professional development and investment required for teacher educators to be sufficiently research active to fulfill their obligations. The National Strategy outlined a new Master's degree in Education “that will enable large numbers of educational practitioners to develop their research and enquiry skills,” and a range of doctoral level opportunities for teacher educators and educational researchers. The aim of which is to support not only the research capacity of teacher educators but also to fund research networks, and research itself into education in Wales. It is therefore seen not only as a function support for teacher education but also as part of the educational infrastructure for Wales.

While a commitment to supporting the research capacity of teacher educators is an important initiative in order to support research capacity generally, there is still the question as to whether these proposals address the challenges teacher educators face in becoming research active.

From the experience of teacher educators in Wales, there are still challenges for the ambitious Welsh vision to be realized:

Our expectations of our staff are much more: “don't replicate the classroom here”. We get them to think about what's the university is for? What do we understand by critical thinking and critical debates? Where is that occurring in your practice? We've got better at those kinds of conversations. But I think it's very hard. ... we've just had a whole bunch of new staff. And some of them have been saying, “if only I'd had time to do this when I was a teacher. If only somebody had told me to read this or share this with me or discuss that with me, it would have changed my practice.”

But how we shape them into a teacher educator is quite different. And actually, the skills that you need to be a good teacher educator, are not necessarily the things that you need to write good research report. And those skills don't necessarily translate into how you support schools. Research can be quite an insular experience, almost selfish. Whereas if you're supporting a school, it's much more of an altruistic social collaborative experience.

The question of the capacity of teacher educators to conduct research remains an important one. However, this comparison highlights the distinctive set of skills and professional development needed to be a good teacher educator and to be a good academic researcher. Even when working collaboratively with schools, there are challenges to overcome in how teacher educators are enabled to support the expectations for their role. In particular the question remains as to how research within teacher education itself can develop while the research capacity for teacher educators remains a significant challenge.

Enabling or constraining research capacity

The Welsh National Strategy seeks to enhance and support capacity for research within and across the whole education sector, while the approaches taken by the English government could be described as more prescriptive, narrow, and conservative with a particular stance focused on curated evidence over (practitioner) research. These different examples reveal how policy can create the conditions for research capacity within initial teacher education.

The two contrasting teacher education policy perspectives outlined in detail here are indicative of how the role of research in teacher education, in two quite similar contexts can be enabled or constrained. The detail, while specific to these two geographically connected contexts, is useful in revealing how the trends identified in previous research (the role of universities, the perspective of new teachers, the preparedness of teacher educators and the accountability systems) are replicated and compounded in policy documents. Policy is central in creating the conditions in which teacher educators have to work adaptively in order to ensure that their programs are high quality and effective. Without the work of teacher educators, these policy contexts would be instrumental in determining how research is situated in the process of becoming a teacher, and ultimately how that then frames teacher knowledge and professionalism.

Tatto (2021) makes a strong case for why teacher educators and teachers need to “take back control” of research in education. She argues that without increasing research capacity in this area that education policies will “worsen rather than improve the performance of underserved populations,” that the capacity for teacher educators to be taken seriously within the academy will be reduced and that the “fragile bond” formed by school-university partnerships will be harder to maintain. Moreover:

Instead of faculty working arm in arm with future teachers and together engaging in a research-rich culture of continuous programs of research, learning and improvement, the worlds of faculty and future teachers remain separate The narrowing of the curriculum in teacher education has resulted in teachers' lack of preparation to engage in rigorous evaluation research able to challenge ill-informed non-educationist research and to explore innovative ways of teaching and learning, to, in short, reimagine education.

The analysis presented here shows that a vision of teacher education that foregrounds research is possible at a national policy level, and that while this does not necessarily mean that the documented problems of teachers' research-engagement will be resolved, it does provide an example of how research capacity can be enabled. The analysis also shows what the alternative could look like: a controlled and prescriptive, narrow curriculum which also influences the pedagogy and practice of teacher education, stripping away the role and expertise of teacher educators, ultimately placing government departments in control of what is taught to new teachers. Although contrasting in approach, the policy documents reflect the trends in professionalizing teacher education highlighted by Mayer (2017): in particular in how policies can seek to regulate provision as well as (de)regulate providers, and how Standards (and accreditation) can be used as policy technologies to affect pedagogy and practice.

While the examples of England and Wales are fairly contrasting, they reflect the discourse community of the policy contexts they have emerged from. The ITT Market Review Recommendations have yet to be implemented or even accepted by the British government (at time of writing), but in effect as a document which sets out a series of expectations about teacher education, the potential "damage" to research capacity is stark and clear. This of course has been written in response to the complex and diverse landscape of providers in the English system that has been diversified in this way to represent a faux market, driven by a technicist discourse about what it means to learn to teach. While this technicist approach may seek to solve a short-term problem about teacher supply and demand, it fails to address the much more substantive issues about teaching quality and teachers' adaptive capacity. It is precisely here that research capacity is so important.

The link between research capacity in teacher education and teacher quality are perhaps difficult to quantify. The reality of teacher development is that individual trajectories are so bound up with personal experiences that claims about one approach or another are fairly meaningless when scaled up (as concluded by Grossman and Loeb (2008)). If such a link cannot be made, then the question becomes a philosophical one: what sort of education our society privileges and how it supports teachers to contribute to that.

Acknowledgment

I am very grateful to the Welsh colleague who agreed to be interviewed for this research.

References

- Biesta, G., 2007. Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educ. Theor.* 57, 1–22.
- Brooks, C., 2021a. *Initial Teacher Education at Scale: Quality Conundrums*. Routledge, London.
- Brooks, C., 2021b. Research capacity in initial teacher education: trends in joining the "village". *Teach. Educ.* 32, 7–26.
- Burn, K., Mutton, T., 2013. Review of Research-Informed Clinical Practice in Teacher Education. Paper Submitted to the BERA-RSA Inquiry. BERA/RSA, London.
- Burn, K., Mutton, T., 2015. A review of "research-informed clinical practice" in Initial Teacher Education. *Oxf. Rev. Educ.* 41, 217–233.
- Carter, A., 2015. Carter Review of Initial Teacher Training (ITT). Department for Education, London.
- Cochran-Smith, M., Lytle, S.L., 1999. The teacher research movement: a decade later. *Educ. Res.* 28, 15–25.
- Cochran-Smith, M., Grudnoff, L., Orland-Barak, L., Smith, K., 2019. Educating teacher educators: international perspectives. *N. Educat.* 1–20.
- Darling-Hammond, L., Lieberman, A., 2013. *Teacher Education Around the World: Changing Policies and Practices*. Routledge, Abingdon.
- Darling-Hammond, L., Macdonald, M.B., Snyder, J., Whitford, B.L., Ruscoe, G., Fickel, L., 2000. *Studies of Excellence in Teacher Education: Preparation at the Graduate Level*. ATCEE, Washington D.C.
- Darling-Hammond, L., 2006. *Powerful Teacher Education: Lessons From Exemplary Programs*. John Wiley & Sons, San Francisco.
- Department For Education, 2021. *Initial Teacher Training (ITT) Market Review Report*. Department for Education, London.
- Ellis, V., McNicholl, J., 2015. *Transforming Teacher Education: Reconfiguring the Academic Work*. Bloomsbury Publishing, London.
- Ellis, V., Steadman, S., Trippstad, T.A., 2018. Teacher education and the GERM: policy entrepreneurship, disruptive innovation and the rhetorics of reform. *Educ. Rev.* 71, 101–121.
- Furlong, J., Whitty, G., 2017. Knowledge traditions in the study of education. In: Whitty, G., Furlong, J. (Eds.), *Knowledge and the Study of Education: An International Exploration*. Symposium, Oxford.
- Furlong, J., 2013. *Education—An Anatomy of the Discipline: Rescuing the University Project*. Routledge, Abingdon.
- Furlong, J., 2019. The Universities and initial teacher education; challenging the discourse of derision. *The Case of Wales. Teach. Teach.* 25, 574–588.
- Godfrey, D., 2017. What is the proposed role of research evidence in England's "self-improving" school system? *Oxf. Rev. Educ.* 43, 433–446.
- Grossman, P.L., Loeb, S., 2008. *Alternative Routes to Teaching: Mapping the New Landscape of Teacher Education*. Harvard Education Press, Cambridge, MA.
- Henry, J., 2020. The cinematic pedagogies of underprepared teachers. *Teach. Teach. Educ.* 89, 102990.
- Holme, R., Robb, A., Berry, W., 2016. Becoming a teacher educator—the motivational factors. *Eur. J. Teach. Educ.* 39, 340–354.
- Hulme, M., Wood, J., Shi, X., 2020. Measuring up? Metrics and research assessment in UK teacher education. *J. Educ. Teach.* 46, 220–239.
- Labaree, D.F., 2006. *The Trouble With Ed Schools*. Yale University Press, New Haven.
- Lortie, D.C., 1975. *Schoolteacher*. Univ of Chicago Press, Chicago.
- Mayer, D., 2017. *Professionalizing Teacher Education*. Oxford Research Encyclopedia of Education. Oxford University Press, Oxford.
- Menter, I., Hulme, M., Elliot, D., Lewin, J., Baumfield, V., Britton, A., Carroll, M., Livingston, K., McCulloch, M., McQueen, I., 2010. *Literature Review on Teacher Education in the 21st Century*. University of Glasgow.
- Moore, A., 2004. *The Good Teacher: Dominant Discourses in Teaching and Teacher Education*. RoutledgeFalmer, London.
- Murray, J., Male, T., 2005. Becoming a teacher educator: evidence from the field. *Teach. Teach. Educ.* 21, 125–142.
- Stenhouse, L., 1981. What counts as research? *Br. J. Educ. Stud.* 29, 103–114.
- Tatto, M.T., 2021. Developing teachers' research capacity: the essential role of teacher education. *Teach. Educ.* 32, 27–46.
- Trent, J., 2013. Becoming a teacher educator. *J. Teach. Educ.* 64, 262–275.
- Turvey, K., Pachler, N., 2020. Design principles for fostering pedagogical provenance through research in technology supported learning. *Comput. Educ.* 146.

- Turvey, K., 2018. Humanizing as innovation in a cold climate of [so-called-evidence-based] teacher education. *J. Educ. Teach.* 45, 15–30.
- Watson, S., Barnes, N., 2021. Online educational populism and new right 2.0 in Australia and England. *Glob. Soc. Educ.* 1–13.
- Welsh Government, 2021. In: Language, E.A.T.W. (Ed.), *The National Strategy for Educational Research and Enquiry (NSERE): Vision Document*. Welsh Government, Cardiff.
- Whiting, C., Whitty, G., Menter, I., Black, P., Hordern, J., Parfitt, A., Reynolds, K., Sorensen, N., 2018. Diversity and complexity: becoming a teacher in England in 2015–2016. *Rev. Educ.* 6, 69–96.
- Williams, J., 2014. Teacher educator professional learning in the third space. *J. Teach. Educ.* 65, 315–326.
- Young, M., Muller, J., 2014. From the sociology of professions to the sociology of professional knowledge. In: Young, M., Muller, J. (Eds.), *Knowledge, Expertise and the Professions*. Routledge, Abingdon.
- Zeichner, K.M., 2017. *The Struggle for the Soul of Teacher Education*. Routledge, Abingdon.

Studying teachers' work and lives: insights from research before and during the COVID-19 pandemic

Maria Assunção Flores, Research Centre for Child Studies, Institute of Education, University of Minho, Braga, Portugal

© 2023 Elsevier Ltd. All rights reserved.

Introduction	622
The multidimensional and changing nature of teacher professionalism	622
Teachers' work and the role of emotions	623
Being a teacher in changed circumstances	624
The studies	624
Participants	624
Methods	626
Data collection and analysis	626
Findings	626
First study	626
Appreciation of the pedagogical and collaborative aspects of teaching	626
Portraying a negative picture: heavy workload, bureaucracy and tiredness	627
Mixed feelings: positive energy versus disappointment	628
Maintaining motivation in teaching	628
Second study	629
Teaching in a changed scenario: forced digital learning during the pandemic	629
Working extra hours and frustration "at not being able to reach all the students"	629
Lack of support at policy level: reliance on collegiality and collaboration	630
Capacity for resilience and increased professional learning	631
Conclusion and discussion	631
Acknowledgments	633
References	633

Introduction

The realities of teachers' work and lives have been investigated from a variety of perspectives including the knowledge and skills required of teachers, the impact of (trans) national policies, the changing nature of their professionalism, the perceived image of the teaching profession, school and teaching cultures, etc. (Helsby, 2000; Osborn, 2006; Day et al., 2007; Kelchtermans, 2009; Flores, 2012, 2019; OECD, 2019). Quality has been the common theme, often associated with the need to raise education standards internationally, and also with issues of intensification and bureaucratization, increased forms of managerialism, greater accountability and public scrutiny (Osborn, 2006; Ávalos, 2013; Day, 2017; Flores, 2019).

Existing literature has pointed to the complexity and interplay of the variables impacting on teachers' work and lives in many parts of the world (Flores, 2012; Day, 2017), issues identified including the rise of performative cultures through increased accountability and imposition of standards (Sachs, 2016), the compliance of teachers with the demands of performance policies and the tyranny of audit (Lo, 2012), and the decline in teacher autonomy and an environment of distrust (Sachs and Mockler, 2012). More recently, the outbreak of the COVID-19 pandemic has brought about profound changes in teachers' work as a result of the forced and abrupt shift from face-to-face to remote teaching due to the compulsory closure of schools in an effort to mitigate the spread of SARS-CoV-2. In many parts of the world, the long-term effects of this changed scenario have had a marked impact on the circumstances facing teachers in their work and professional lives. This chapter reports on findings from two studies carried out in Portugal before and during the COVID-19 pandemic in order to illustrate key issues that cut across teachers' work and which need to be considered in discussions about the nature of teacher professionalism.

The multidimensional and changing nature of teacher professionalism

Teacher professionalism has been described as a contested site (Sachs, 2016) since it depends on a wide range of dynamic variables, whether political, professional, or contextual. Thus, the study of teachers' professionalism needs to go beyond a normative perspective and to take into account teachers' views and experiences of professionalism in context (Flores, 2020). Tensions have been reported in the literature between, on the one hand, managerial and top-down understandings of teacher professionalism and, on the other hand, individual, collegial and contextualized responses to the changing nature of teaching (Ávalos, 2013; Day, 2017). Looking at policies that have affected teachers' work in the Portuguese context, Estrela (2014) contends that old and new

professionalisms may be identified, representing both regression and progress in the teaching profession. She asserts that teacher professionalism remains a sensitive issue and warns that teaching might be seen as a risky profession or a profession at risk while, at the same time, it may be viewed as an opportunity for teacher re-professionalization (Estrela, 2014). Similarly, Sachs (2016) identifies the factors that shape teacher professionalism, namely performance management and performance cultures, increased accountability, and the continued imposition of teacher standards.

The two ideal types identified by Evetts (2009), namely organizational and occupational professionalism, are useful concepts for understanding teachers' work in times of intensive school reform and top-down initiatives, indicating the need to improve the quality of teaching and raise student achievement while also highlighting the importance of teacher autonomy and agency. Organizational professionalism is situated within a logic of control, greater public accountability, the use of standardized procedures within hierarchical structures of authority and decision making. This managerial perspective of looking at teacher professionalism includes external forms of regulation, target-setting and performance review and assessment. In turn, occupational professionalism draws on the discourse constructed within professional groups, in which issues of collegial authority and discretion and occupational control of teachers' work are prevalent. In addition, it entails "relations of practitioner trust from both employers and clients and it is based on autonomy and discretionary judgment and assessment by practitioners in complex cases" (Evetts, 2009, p. 23). These two discourses of professionalism identified by Evetts (2009) parallel the notions of managerial professionalism, with a focus on performance and accountability, and democratic professionalism, focusing on collegial relations and collaborative work practices, as identified by Sachs (2016).

In the Scottish context, Reeves (2009, p. 114) analyses three competing forms of teacher professionalism which are "embedded in multileveled and mixed practices in schools": (i) bureau professionalism; (ii) educational operationalism; and (iii) "new" professionalism. The discourse of bureau professionalism involves viewing teachers as experts in their classrooms working in bureaucratic structures. In this context, good teaching means displaying "certain prescribed behaviors that should be used to grade teacher performance as a prerequisite for identifying targets for improvement" (p. 107). As such, control and accountability are key features of this type of professionalism which is to be assured through a "detailed prescription and supervision of performance" within the "implementation of a central determined curriculum, thus setting clear restrictions on the exercise of practitioner judgment and agency" (p. 107). In turn, "educational operationalism" has some similarities with the positioning of teachers within bureau professionalism, namely in regard to their "lack of say over organizational goals and governance and the hierarchical nature of power relations in the education system" (pp. 107–108). Teachers are closely supervised and are seen as compliant operatives. Thus, educational operationalism accentuates the "range of mechanisms for controlling teachers' work and the intensity of their impact" representing a "material alteration of relationships within the hierarchy resulting in a loss of autonomy for teachers" (p. 108).

The notion of "new" professionalism draws on the reflective and collegial role of organizational members. Issues of collaboration and knowledge sharing are key factors as are learner-centered practice; clarity about moral and social purpose; commitment to evidence-informed practice and critical reflection; discretionary judgment exercised both collectively and individually; collegiality and collaboration with colleagues, other professionals, pupils, and parents; commitment to continuing professional development and knowledge creation (Reeves, 2009, p. 108). In Reeves' view, the tensions arising from the co-existence of these three forms of professionalism create opportunities for teachers to reinvent and revitalize new forms of professionalism through "action, argument and alliances within the field of practice" (p. 115). It is, therefore, important to examine further the nature and forms of teacher professionalism in context and the conditions for its enhancement by drawing on the views and experiences of teachers.

Teachers' work and the role of emotions

Emotions are a crucial element in teaching (Hargreaves, 1998; Nias, 1996). They need to be taken into consideration if we are to fully understand the nature of teachers' work and their professional lives. Existing research has shown the importance of emotions in the process of identity development (e.g. Flores and Day, 2006; Lasky, 2005; O'Connor, 2008; van Veen et al., 2005) and in understanding teachers' responses to school reform and educational change (e.g. Hargreaves, 2005; Darby, 2008; Schmidt and Datnow, 2005) as well as their work in the classroom setting (Bahia et al., 2013; Sutton et al., 2009).

In the UK, Jeffrey and Woods (1996) identified professional uncertainty, confusion, anomie, anxiety, mortification and doubt among teachers when they investigated primary school teachers' responses to an OFSTED inspection, associating these with "dehumanization" and "de-professionalism". Other negative emotions were: frustration; anger exacerbated by tiredness, stress and students' misbehavior; anxiety due to the complexity of the job; guilt, sadness, blame and shame at not being able to achieve ideals or targets imposed by others. The study also identified loss of pedagogic values and of harmony, as well as changed and weakened commitment. The authors argue that one possible way to avoid such negative trauma is by shifting identity and status from professional to technician.

Schmidt and Datnow (2005), drawing on sociological and educational theory, examine the emotions teachers mobilize in making sense of educational reforms. Positive and negative emotional responses were found in regard to changes in classroom practices with implications for their role and identity as teachers. Similarly, Darby (2008) looked at teachers' specific emotions in critical incidents that occurred during the period of reform and to explore the reconstruction of their professional self-understanding. She found that teachers experienced fear and intimidation when their professional self-understandings were challenged. She also identified feelings of pride and excitement, which were linked to support from a literacy coach and university faculty, helping teachers to reconstruct self-understandings and leading to improvements in student achievement and in their own instructional practices.

In a similar way, [van Veen et al. \(2005\)](#) examined one teacher's perceptions of educational reform from a social psychological perspective. The authors found both negative and positive emotions. The former were related to anxiety, anger, guilt, and shame stemming from the "lack of time, enormous number of portfolios each semester, and the lack of support from his subject colleagues, school management, and government" (p. 931) while the latter were associated with achievement of goals and purposes in teaching and opportunities to improve teaching and reinforce professional identity.

As [Sutton et al. \(2009\)](#) assert, teaching is an "emotional enterprise" as "teachers may experience happiness when an instructional objective is met or students follow directions, frustration when students cannot grasp a concept, anger with misbehavior, disappointment with lack of effort, and anxiety when competence is challenged" (p. 130). Emotions entail an individual, private and psychological dimension, as well as being socially constructed and situated in relations ([Zembylas, 2007](#)). They concern not only personal dispositions and commitments but are also a product of given ways of organizing teaching ([Hargreaves, 2001](#)). Emotions are, therefore, key to understanding the way in which teachers perceive school reform and educational change but also how they view their role and identity and how they experience teaching in a particular context.

Being a teacher in changed circumstances

The sudden and unexpected shift from face-to-face to remote teaching as a result of the closure of schools due to the COVID-19 pandemic entailed global and local responses to a scenario for which nobody was prepared. Emergency remote teaching ([Bozkurt and Sharma, 2020](#)) was rapidly put in place but a number of difficulties were identified from the outset. Technical problems, lack of adequate equipment and discrepancies in access to technological devices and to internet access have been reported ([Zhang et al., 2020](#); [Flores et al., 2021](#)) with implications for pupils' learning ([NFER, 2020](#)).

Existing literature points to the unexpected digital transformation that students, teachers, parents, school administration and society had to face regarding the development of new skills and resources ([Iivari et al., 2020](#)) and appropriate infrastructures ([Carrius, 2020](#)). Teachers were asked to continue teaching in new scenarios while having to face new challenges regarding communication with students and providing support for students' learning ([König et al., 2020](#)). Issues such as time pressure and increased workload ([van der Spoel et al., 2020](#)), knowledge and skills required in distance learning education, internet access, use of devices for distance learning education ([Alea et al., 2020](#)) and lack of interaction and spontaneity ([Niemi and Kousa, 2020](#)) have been identified. Previous experience in distance teaching explained a more positive adaptation on the part of teachers to remote teaching ([van der Spoel et al., 2020](#); [Flores et al., 2021](#)) as well as support on the part of school leadership ([Judd et al., 2020](#); [NFER, 2020](#)). In addition, teachers' digital competencies have also been explored as well as the need to develop appropriate professional development opportunities in ICT ([Huber and Helm, 2020](#); [König et al., 2020](#)). Other empirical work found that there was a reduction in the perception of teachers' well-being regarding the profession along with concerns related to their professional future ([Alves et al., 2021](#)). In face of such a challenging scenario, it is important to examine how teachers have experienced remote teaching and how it has affected their lives and sense of professionalism.

The studies

This chapter reports on findings from two wider studies carried out in Portugal. The first study was conducted in 2016/2017, focusing on professional standards and professional development, and their impact on teachers' work and lives. This research was funded under grant SFRH/BSAB/114245/2016 (from the National Foundation for Science and Technology). The second study was carried out in 2021 (May–July), after the experience of two lockdowns (2020–2021) and of face-to-face teaching under more challenging circumstances due to sanitary measures such as the wearing of masks and physical distancing. In this chapter, the following questions are addressed:

- What dimensions of their work do teachers value the most and the least, and how do they feel about being a teacher (Study 1)?
- How do teachers view the experience of being a teacher during the COVID-19 pandemic (Study 2)?
- What are the key issues characterizing teachers' work before and during the COVID-19 pandemic (Study 1 and 2)?

Participants

In the first study, 1307 teachers participated, the vast majority of them being female (70.5%) and aged over 45 (69.7%). Most of the participants have between 21 and 30 years of teaching experience (44.1%) (see [Table 1](#)).

In total, the second study had 2192 participants, drawn from all sectors of the teaching profession and most of whom are experienced teachers (see [Table 2](#)). The majority are female (82.2%) and aged over 45 (78.7%). Such figures are in accordance with official and international data. According to the recent TALIS report ([OECD, 2019](#)), the average age of teachers is 44, but in Portugal it is 49. A recent study conducted by the National Council of Education ([CNE, 2019](#)) points to the need to renew more than half of the teaching workforce in the next decade: 52.9% of teachers are aged 50 or over and only 1.1% are aged under 35 ([CNE, 2019](#)). Teaching is also characterized as follows: a highly qualified profession (a master's degree is now required for all entrants into teaching, from pre-school to secondary school), with professional stability (low teacher mobility) but also precarious working conditions (for teachers who do not have a permanent post in a school) and a lack of appreciation for the teaching profession ([CNE, 2016](#)).

Table 1 Demographic characteristics of the teachers participating in the first study.

<i>Demographic characteristics</i>	<i>f</i>	<i>%</i>
Gender		
Female	921	70.5
Male	386	29.5
Age		
18–24	0	0.0
25–34	21	1.6
35–44	375	28.7
45–54	606	46.4
55–64	300	22.9
65 or over	5	0.4
Years of experience		
1–5	15	1.2
6–20	445	34.0
21–30	577	44.1
Over 30	256	19.6
Missing	14	1.1
Teaching sector		
Pre-school	88	6.7
1st cycle (primary)	225	17.2
2nd cycle	182	13.9
3rd cycle	134	10.3
2nd and 3rd cycle	65	5.0
Secondary	214	16.4
3rd cycle and secondary	323	24.7
Other (e.g. special education, school library)	76	5.8
Total	1307	100.0

Table 2 Demographic characteristics of the teachers participating in the second study.

<i>Demographic characteristics</i>	<i>f</i>	<i>%</i>
Gender		
Female	1801	82.2
Male	391	17.8
Age		
21–25	4	0.2
26–30	17	0.8
31–35	16	0.7
36–40	114	5.2
41–45	314	14.4
46–50	461	21.0
51–55	539	24.6
56–60	476	21.7
61–65	242	11.0
Over 65	9	0.4
Years of experience		
0–5	72	3.3
6–20	432	19.6
21–30	946	43.3
Over 30	742	33.8
Teaching sector		
Pre-school	185	8.4
1st cycle (primary)	428	19.5
2nd cycle	267	12.2
3rd cycle	394	18.0
Secondary	352	16.1
More than one teaching sectors	566	25.8
Total	2192	100.0

Methods

In the first study, an online survey was designed based on a questionnaire developed and conducted by Flores and Van Nuland (2013) and Flores (2014). The questionnaire consisted of six sections and participants were asked to respond to a series of questions using a Likert scale relating to their understanding of the purposes and use of professional standards; dimensions of their work; feelings about being a teacher; opportunities for professional development and motivations to remain in or leave the teaching profession. They were asked about the nature of their work, and from a list of 15 possible areas, to select the five most important aspects of their work as a teacher. In relation to their career-long development, participants were asked to describe two of the most significant experiences of professional learning and development in their career over the last five years. Finally, where relevant, participants were also asked to comment on their experience and work as teachers. Demographic questions were also included.

The second study is based on a questionnaire drawing on previous studies (see Flores et al., 2021; Flores et al., 2021; Ávalos et al., 2021) which contained three main sections: (i) perceptions about the second period of remote teaching (2021) compared with the first one (2020); (ii) perceptions about resources and support; (iii) work experience since the eruption of the pandemic and overall impact. Teachers were asked to rate each item in each section using a Likert scale, in the case of closed-ended questions, and were also given open-ended questions offering them the chance to describe their work and first-hand experience in more detail.

Data collection and analysis

In the case of the first study, a nationwide survey was conducted through an online questionnaire (using SurveyMonkey) which was sent to the principals of all the elementary and secondary schools in mainland Portugal. The questionnaire was then distributed to the teachers in each school. Permission for administering the questionnaire in state schools was previously obtained from the Ministry of Education (ref. no. 0554600001). The voluntary nature of participation was emphasized and confidentiality of the data guaranteed.

As for the second study, the survey was carried out through Google Forms. The invitation to participate and the link to the questionnaire were sent to teachers via their professional associations, unions and teacher centers across the whole of mainland Portugal and the Islands. The project was approved by the Ethics Committee for Research on Social and Human Sciences at the University of Minho (CEICSH 055/2020).

Quantitative data were analyzed through SPSS (version 26). Qualitative data analysis was conducted in two phases: first, an analysis of data gathered from each open-ended question; and subsequently a comparative or horizontal analysis (cross-case analysis) (Miles and Huberman, 1994). In the second phase, common patterns as well as differences were examined. A semantic criterion was used to look for key themes arising from the qualitative data. Key themes and sub-themes were identified. In order to avoid subjectivity, data was validated by at least two researchers. The most frequent trends were then summarized and labeled accordingly.

Findings

In this section the key findings of the two studies are presented in order to capture the teachers' experiences and their views of being a teacher at two specific moments in time, namely before and during the pandemic. Several quotations are used to illustrate the main themes drawn from the open-ended questions.

First study

Appreciation of the pedagogical and collaborative aspects of teaching

The teachers participating in the first study (2017) were asked to select the five most important dimensions of their work from a list of possible areas (see Table 3). It is clear that teachers highlight the pedagogical and collaborative side to their work, the most important issues being motivating pupils, devising new strategies for them to learn better and reflecting on their work in the classroom. Following their work in the classroom the participants also identified their collaborative work with colleagues and professional learning as key aspects of their work. However, only 28.6% referred to participating in the decision-making process, which may be related to a logic of bureaucratic control prevailing in existing structures and a culture of curriculum execution as a result of top-down initiatives within a centralized education system. Teachers feel overwhelmed with the number of tasks they have to perform and with the endless changes in education:

There is so much bureaucracy and it seems that as a teacher you need to do everything. As a teacher, you should devote your time to teaching, planning, organizing and selecting teaching materials, having space and time to work with students ... but much of what you do is paperwork and statistics and reports ... (21 years of teaching)

Table 3 Aspects of teachers' work (Study 1).

	<i>f</i>	%
Motivating students	864	66.1
Devising new strategies for pupils to learn better	811	60.1
Reflecting on my own work in the classroom	702	53.7
Collaborating with colleagues	643	49.2
Enhancing my professional learning	602	46.1
Supporting pupils' learning	595	45.5
Dealing with pupils' behavior	469	35.9
Lesson planning	389	29.8
Participating in the decision-making process in my school	374	28.6
Creating new materials for students so that they can learn better	264	20.2
Sharing materials with my colleagues	219	16.8
Assessing pupils' learning	212	16.2
Communicating with parents	188	14.4
Using ICT in teaching	150	11.5
Doing administrative work (minutes, reports)	17	1.3

I feel unmotivated concerning the policies that are imposed on schools and the ways in which the Ministry of Education treats teachers. (21 years of teaching)

Not surprisingly, completing administrative work was viewed by teachers as the least important aspect of their work. Overall, the teachers participating in the study highlight the core issues of teaching and point out the number of non-teaching tasks that they have to perform which lead them to feelings of disappointment and tiredness as the following quotes illustrate:

Teachers are trained to teach children, to be in the classroom and to do teaching. They are not supposed to do admin work all the time and to undertake admin roles that they do not want but have to do. They have to leave behind their personal life and their family and they cannot even support their own children. (30 years of experience, Study 1)

I miss the time when I was "only" a teacher. My focus was teaching and learning and I had freedom to make choices for the betterment of my students such as projects with the community and so on. (26 years of teaching)

I feel tired, exhausted and with a sense of no hope in teaching. (26 years of teaching)

Sharing materials with colleagues, using ICT and assessing pupils' learning were also among the least appreciated aspects of their work although these are considered to be important for improving teaching and learning in schools and classrooms. This is interesting at a time when the use of ICT and the classroom of the future have been widely discussed. The overemphasis on summative (and external) assessment may explain why the participants do not attach importance to assessing pupils' work even if it plays a key role in improving teaching and adjusting pedagogical strategies in the classroom. Similarly, communicating with parents is cited by participants as one of the least appreciated aspects of their work. This picture changed after the experience of remote teaching during the COVID-19 pandemic as will be shown later in this chapter.

Portraying a negative picture: heavy workload, bureaucracy and tiredness

How teachers feel about their work and their role matters as much as the conditions in which they do their job. Teachers' interactions within their professional contexts entail a political dimension (Kelchtermans, 2005) which is related to their beliefs and emotions and the ways in which they respond to the challenges of the environments in which they work. When asked about how they felt as teachers, most of them reported negative feelings and emotions. Issues such as tiredness and fatigue, lack of energy, disappointment and disillusion, fear and insecurity, sadness, frustration and feelings of giving up, anger, loneliness, anxiety, pessimism and low morale were reiterated in most of the accounts. The following quotes are illustrative of this:

I feel tired and with no energy (the workload is the same or even greater ...); it is more and more difficult to deal with so many demands and with students who have other opportunities to learn in a far more interesting way than at the school. (29 years of teaching)

I feel tired because I have been dealing with changes in teaching for a long time and I cannot see any changes in reality ... (23 years of teaching)

As a teacher, I feel tired and sad. You cannot change what is wrong. There are no models to inspire us. (31 years of teaching)

Teachers are rather pessimistic and express negative emotions when they talk about how they feel as teachers. The reasons put forward for this vary but most of them are related to endless policy initiatives, lowering of their socioeconomic status, the increasingly bureaucratic nature of teaching and the logic of control and accountability also associated with heavy workload and both internal and external pressures to comply with regulations and top-down directives. Bureaucracy, many (unproductive) meetings, large classes, and the lack of social recognition of the profession were reiterated in their accounts. They also emphasized difficulties in maintaining a balance between work and personal life and the lack of career prospects:

Nowadays as a teacher you don't have time for anything even for your family. There is no weekend, you go to bed very late, you need to do prep for teaching, you need to do lots of reports, looking at emails ... it's too much, it's very tiring ... (19 years of teaching)

My perception is that teachers are tired professionally due to the lack of recognition, lack of career progression (in my case 12 years without any kind of progression!), ageing, instability in terms of ongoing legislative texts and top-down directives to comply with, etc. (16 years of experience)

Mixed feelings: positive energy versus disappointment

Other teachers revealed mixed feelings and spoke of both positive and negative experiences. On the one hand, they talked about disappointment and tiredness, which they relate to the policy environment, the lowering of the socioeconomic status of the profession, the lack of career prospects and the lack of appreciation for the teaching profession, especially on the part of policy makers:

I have mixed feelings and maybe even contradictory emotions; sometimes I feel motivated and energetic; sometimes I feel tired with feelings of giving up. (23 years of teaching)

I think I still have energy and willingness to teach but I feel that policy makers do not value my work. I feel tired with so much bureaucracy and I am disappointed with no career progression and having to work until the age of 66. (37 years of teaching)

On the other hand, they still find positive aspects to teaching, which they clearly attribute to their students and their work in the classroom, as well as to feelings of job satisfaction and to intrinsic rewards of teaching. This is in line with earlier research in the context of changes in educational policies marked by bureaucratization in Portugal which showed that negative emotions were related to policy reform whereas positive ones were related to classroom interactions ([Bahia et al., 2013](#)). Some of the teachers participating in the first study also spoke of their core values as teachers and their sense of professionalism, vocation and identity:

My motivation comes from my students and their willingness to learn. (21 years of teaching)

Personally, I feel motivated by teaching and my other roles at school. However, the endless reforms in education have contributed to a negative image of the teaching profession and to devalue the role of the teacher especially in relation to students and parents. (19 years of teaching)

Maintaining motivation in teaching

A minority of the teachers participating in the first study are clearly positive about their work and their role as teachers. Their accounts involve positive emotions which again they relate to a sense of vocation and to the core issues of teacher professionalism such as dedication, care, and the moral and social purposes of teaching. Making students learn, being motivated and energetic, as well as happiness and love for teaching were identified as key elements:

I feel energetic and willing to contribute to the development of my students with creative dynamics ... (21 years of experience).

I feel motivated because I love teaching. (16 years of experience)

Despite my age, I feel energy to keep going in teaching ... I feel curious about the future. (24 years of teaching)

I am definitely motivated and happy to see my students learn. (30 years of teaching)

The focus on students, on their learning, well-being and development is at the forefront of teachers' clear positive experience, which was associated with their professional values as teachers. The collaborative dimension of their work and the contribution teachers can make to the education of children are also reiterated:

I feel motivated because I still believe in teaching and in the key role of the school in children's development and education. (29 years of teaching)

I am energetic because I don't give up. I continue to fight for the betterment of students and for the teaching profession but I feel that only a few of my colleagues are with me. (42 years of teaching)

I feel good about contributing to students' learning, developing their knowledge and potential and working collaboratively with my colleagues. (27 years of experience)

Second study

Teaching in a changed scenario: forced digital learning during the pandemic

While using ICT in teaching was not considered to be a key dimension in teachers' work by the participants in 2017 (study 1), the outbreak of the pandemic in 2020 imposed a new reality. The rapid and unexpected shift from face-to-face to remote teaching as a result of the closure of schools represented a great challenge for teachers. From the very beginning it was plain to see the lack of adequate equipment and resources and the lack of digital competencies on the part of both students and teachers. The need to improvise the classroom in their homes and maintain contact with their students meant that teachers were forced to take up digital learning. This situation arose due to the "emergency remote teaching" which teachers suddenly had to provide at the beginning of the COVID-19 crisis but which also later involved the construction of online tools and resources as well as a more planned pedagogical response. The following quotes are illustrative of the "initial shock" imposed by the pandemic, but also the learning that came about from the experience of doing remote teaching along with feelings of powerlessness, tiredness and loss of boundaries between personal and professional life:

The first time was a moment of "shock" and trying to do everything that was possible to keep in touch with my students, being there for them, supporting them and their families. I was not so worried about the content or the tasks. But it was a moment lived with a lot of stress, tiredness, powerlessness, and loss of balance between my personal and professional life. (...) The second time was much more planned, organized and reflected ... problems were resolved in a quicker and more effective way. Teachers were able to focus, to better plan their work ... students had also developed competencies to deal better with distance teaching ... (22 years of experience).

During the first lockdown, the focus was on changing everything "from one day to the next". We had to adapt "overnight" in order for the students' learning to be affected as little as possible. At the same time, there was a need to manage the family situation (everybody was at home). In the second lockdown, the biggest concern was for students to participate in the synchronous sessions. (19 years of experience)

Teachers' highlight a sense of "discovering a new world" as they became familiar with digital platforms and online educational tools (GoogleMeet, Zoom, Mentimeter, Google Classroom, Teams, Socrative, ...). They highlight the development of digital competencies and the need to learn more about how to teach in the new teaching scenario.

Working extra hours and frustration "at not being able to reach all the students"

Teachers' work and lives were profoundly affected by the pandemic. The sudden shift from face-to-face to remote teaching marked the experience of being a teacher during the pandemic as did the imposition of sanitary rules which teachers had to deal with when schools reopened. The need to keep in touch with the students and get them involved in the activities so as to avoid too much

"learning loss" was stressful for teachers, whose working hours have increased due to time taken up preparing lessons, contacting pupils and their families and managing their daily professional work. Existing findings point to teachers being concerned about their students' well-being and level of learning as well as the financial situation of students and their families to the detriment of concerns about their own financial, personal and professional situation (Flores et al., 2021; Ávalos et al., 2021).

Although teachers recognize that digital learning occurred as a result of their teaching experience during the pandemic, they emphasize the heavy workload, the difficulty to find a balance between their personal and professional lives and an increase in time devoted to teaching preparation. Being a teacher in COVID-19 times, especially during the lockdowns, involved more intense work in direct contact with families in order to better support their students ("being there for them") and, therefore, communicating with parents was seen as a key aspect of their work, which stands in marked contrast to the findings of the first study. Issues reported also included lack of time, difficulties in managing the various roles that had to be performed at home, along with tiredness and fatigue:

From my experience I have learned that working until physical and psychological exhaustion is not good. I used to work for 14 or 15 hours a day during the first lockdown. And this had consequences. Preparing synchronous and asynchronous lessons with tasks for 212 pupils and also the verification of the tasks led me to take medical leave for the first time in almost 30 years of teaching. (...) For me the problem was the countless hours of preparation. (29 years of experience)

I feel overwhelmed and with a sense of professional disenchantment. As a teacher I had to respond to the queries of the various stakeholders at any time of the day or night and even during the weekend. (28 years of experience)

This perception was corroborated by other empirical studies which point to an increase in workload, stress, anxiety, and even burnout (Sokal et al., 2020; Flores et al., 2021) alongside a reduction in the perception of well-being in regard to the profession and an increase of concern in relation to one's professional future (Alves et al., 2021). The difficulty to reconcile family and professional lives and the increased workload (non-remunerated working hours) are reiterated in teachers' accounts along with the lack of support, lack of time and frustration at not being able to "reach all the students":

My experience meant more work, more tiredness and a lot of frustration at not being able to reach all of my students. (5 years of experience)

Lack of support at policy level: reliance on collegiality and collaboration

While the unexpected and sudden transition to remote teaching was marked initially by a sense of confusion and anomie, teachers showed a positive and rapid response. The mobilization of endogenous resources and the dynamics of collaboration and collegiality were reiterated in teachers' accounts. When asked about the support received and the ways in which they have handled the difficulties encountered, the majority of the teachers participating in the second study highlight factors associated with their working contexts, namely: colleagues (69.5%), tutors (teachers responsible for liaizing between students, teachers and families) (63.4%), leadership team (60.9%), department/year/cycle coordinators (59.1%), the principal (58.7%) and the support team responsible for distance teaching in the school (45.7%). Less significant were the contribution of school association centers responsible for INSET (28.6%) and the guidance provided by the Ministry of Education (15.1%). The following quotes point to the difficulties faced by teachers but also to the capacity of schools and teachers to "reinvent" themselves to meet the challenges of the new teaching scenario:

I emphasize the capacity of teachers and schools to reinvent themselves in the sense that, in record time, they were able to respond to distance teaching and all the changes it entailed. (25 years of experience)

I have learned that I need to get round these situations on my own. The Ministry of Education's response was completely lacking in terms of guidance. The documents published in this regard don't say anything. They were vague and schools had to organize themselves. (16 years of experience)

As teachers we were very professional in overcoming the difficulties and obstacles and we helped one another. (22 years of experience)

While schools and teachers were able to respond rapidly and relatively effectively to the difficulties involved in remote teaching, it is, however, possible to identify different forms and degrees of change associated with different organizational dynamics and digitalization cultures (Flores et al., 2021). Collegiality and collaboration were key in how teachers responded to the education crisis caused by the pandemic, and which arise from opportunities to expand their professional capital, particularly their social capital based on cultures of mutual support and collaboration (Hargreaves, 2021).

Capacity for resilience and increased professional learning

One of the outcomes of teachers' experience during the pandemic concerns an increase in digital competencies and greater confidence in using online tools and platforms. When asked about the impact of the pandemic on their work, teachers participating in the second study stress the following aspects: greater reflection on how to teach in new scenarios/contexts (64.9%), access to digital equipment and tools on the part of teachers (64.0%), access to digital equipment and tools on the part of students (63.4%), teachers' digital literacy (60.6%), greater use of ICT and digital tools in face-to-face teaching (54.0%), teachers' professional learning and development (51.4%) and students' digital literacy (49.0%). Their accounts point to opportunities for professional learning and development particularly regarding the online environment but also to the chance of reflecting on teaching in a changed scenario and its implications for their work as teachers and for student learning.

I now have more confidence in using ICT in teaching. I have also reflected on my students' learning needs. (24 years of experience)

My experience during the pandemic has led me to reflect on my practice and how to improve it in the new context. (23 years of experience)

From this experience I take a more profound reflection on my professional practice and how to better support my students. (21 years of experience)

However, they also emphasize, once again, the increased difficulties they had to face during remote teaching due to technical problems and lack of resources, the endless hours devoted to preparation and to establishing contact with students and their families. The experience of remote teaching has highlighted for most teachers the importance of the relational dimension in their work, and their resilience and ability to adapt to a new context:

I was impressed with my capacity for resilience and adaptation. This capacity was felt by the vast majority of my colleagues. (30 years of teaching)

I have reinforced the importance of personal contact with children and the key role of human relations. (34 years of experience)

When you need to adapt to a different reality that you are not used to, you discover creativity, resilience and courage that you didn't know you had. (33 years of experience)

I have learned to give more importance to life. I have also realized that distance teaching does not meet my students' needs. (24 years of experience)

As this last quote illustrates, teachers stress the value of face-to-face teaching which they consider to be irreplaceable even though they recognize the importance of digital learning. The lack of student feedback and the lack of support along with tiredness explain the more negative experience of remote teaching:

Distance teaching was at a given moment a real nightmare. It was very tiring and it brought me a sense of powerlessness as I knew that only a few students were listening to me and were able to follow my teaching. With this experience, which I hope not to repeat again, I have learned to appreciate even more the importance of face-to-face contact with students and colleagues. (21 years of experience)

I have learned to deal with technologies better; I have reinforced my belief about the superiority of face-to-face teaching at all levels: scientific, pedagogical and human. (38 years of experience)

Conclusion and discussion

This chapter has set out to examine the realities of teachers' work and lives drawing on data from two wider research projects. Teachers' accounts reveal that their experience of professionalism, their usual practices of teaching and ways of interacting with students and colleagues have been challenged, particularly as a result of the transition from face-to-face to remote teaching during the COVID-19 pandemic. As a dynamic and contextual process subject to different and sometimes competing

interpretations and “voices” (Helsby, 2000; Hargreaves, 2000), teacher professionalism is determined by macro-level definitions and decisions but it also depends on how teachers experience their work and see themselves as professionals.

Both studies reported in this chapter illustrate how teachers live their profession and what they understand of the experience of being a teacher before and during the COVID-19 pandemic. Findings point to differences but also to similarities. One of the features that clearly stands out from the two studies is the appreciation of collaboration and collegiality, which are seen as essential aspects of teachers' work. While the unexpected and sudden transition to remote teaching was marked initially by a sense of confusion and anomie, teachers showed a positive and rapid response which was due in great part to the mobilization of endogenous resources within schools and local communities as well as the dynamics of collaboration and collegiality. This is emphasized within the context of a perceived lack of support or clear guidance at the policy level thus testing the capacity of schools and teachers to “reinvent” themselves to meet the challenges of the new teaching environment. The translocation of teaching from a classroom setting to online platforms (Ávalos et al., 2021) was an intense experience for teachers as they felt the “initial shock” but also created an opportunity for them to collaborate in order to respond to the demands arising from the sudden closure of schools. Discussing the losses, gains and transformations affecting teachers and teaching due to the COVID-19 pandemic, Hargreaves (2021) stresses, among other features, teachers' professional skills, especially pedagogical expertise, together with their professional capital, and in particular social capital, in relation to how they share knowledge and ideas within collaborative cultures and professional networks.

While, in the first study, teachers highlight the individual and collegial aspects of their professionalism (e.g., collaborating with colleagues, motivating pupils, reflecting on their work, etc.) with a focus on what they consider to be the core issues of their profession (their work with students and peers), they also spoke of managerial and functional issues (e.g., administrative tasks), which they generally consider to be the least important aspects of their work but which influence the ways in which they exercise their professionalism. Their accounts thus reveal features of bureau professionalism and educational operationalism (Reeves, 2009) when they refer to issues of bureaucracy, the amount of paperwork and the lack of autonomy, which is situated within a logic of top-down policy initiatives driven by accountability demands. They tend to see themselves as compliant operatives and point to a loss of control over their work marked by prescriptive and narrow perspectives of teaching alongside pressure from society in general and from policy makers in particular. This trend is associated with how teachers see the profession from within and how it is viewed from the outside, particularly from the perspective of policy makers. In other words, tensions between organizational and occupational professionalism emerge (Evetts, 2009) with an emphasis on the former, as top-down policies, formal procedures, managerialism, and hierarchical and bureaucratic structures influence teachers' capacity for action. This may be seen, for instance, in the first study when teachers indicate that participating in the decision-making process is not one of the aspects of their work that they most appreciate. Such a view is to be understood within a logic of bureaucratic control prevailing in existing school structures and within a culture of curriculum execution as a result of top-down initiatives within a centralized education system. Their accounts illustrate the inner tensions of being a teacher in scenarios which are marked by contradictory trends: the focus on teaching and on students' learning needs and well-being, on the one hand, and the pressure for immediate results within a logic of greater accountability and control on the other hand (Ben-Peretz and Flores, 2018; Flores, 2020).

The complex and demanding realities of teaching are described with strong emotions. Teachers feel overwhelmed by policy initiatives and the greater demands and pressures over their work associated with the increased bureaucracy in teaching and endless reforms which impact on their subjectivities and understandings of professionalism (Gewirtz et al., 2009). In the first study, they spoke of disappointment, tiredness, lack of energy, fear, sadness, frustration, loneliness, anxiety, and feelings of giving up. These are experienced within a context of greater demands and excessive workload along with the lack of career prospects, lack of recognition of the teaching profession and the lowering of their socio-economic status (see also Flores, 2020). Issues of stress, fatigue and frustration at “not being able to reach all the students” are also reported in the second study and again corroborate the overall negative picture of the experience of being a teacher. This is in line with earlier empirical work focusing on emotions inherent in teachers' response to external pressures (Jeffrey and Woods, 1996) and with recent research which points to an increase in workload, stress, anxiety, and even burnout associated with teaching during the COVID-19 pandemic (Sokal et al., 2020; Flores et al., 2021) alongside reduced perception of well-being in regard to the profession and increased concern among teachers about their professional future (Alves et al., 2021).

However, positive aspects were also identified when teachers in both studies refer to their work with students. They talk about energy, motivation, satisfaction and willingness to carry on. The focus on students' learning and well-being are at the forefront of their accounts pointing to a sense of vocation and their identity as teachers. At the center of their professionalism are their core professional values, in particular the moral and social purposes of teaching, care, dedication and commitment to students' learning and well-being. Such findings point to the relational dimension of teaching and to the “ethics of care” (Noddings, 1984) and corroborate other research underlining how teachers' experiences of caring are influenced both by their need to sustain positive professional relationships with their students and by their individual beliefs about their role as a teacher (O' Connor, 2008). Such issues were also reinforced in the second study as students' well-being was of greater concern to teachers (“being there for them”) than their own problems (see also Flores et al., 2021; Ávalos et al., 2021). The experience of remote teaching is, therefore, accompanied by an emphasis not only on the non-cognitive aspects of teaching related to personal factors but also on the relationship with students as well as the emotional dimension of teaching (Hargreaves, 2021; Ávalos et al., 2021).

Another feature revealed by the data from the two studies concerns technology. While using ICT in teaching and communicating with parents were among the least appreciated aspects of teachers' work in the first study, the COVID-19 pandemic has completely changed this picture. Due to the shift from face-to-face to remote teaching as a result of the closure of schools, technology has assumed a pivotal role in teaching and in teachers' interaction with students, colleagues and families. The "initial shock" and the need for teachers to get round the situation overnight without the required time for preparation or reflection led them to a forced digital learning. This "discovery of a new world" meant that they had to get to know the necessary online tools and platforms in order to be able to teach and interact with students and colleagues. Clearly this experience represented an opportunity for teachers to develop digital competencies and adapt their usual ways of teaching to the new circumstances. However, as other research suggests, if the response by schools and teachers to the unexpected and sudden transition from face-to-face to remote teaching was generally effective, it also revealed different paces and processes of adaptation associated with different organizational dynamics and digitalization cultures (Flores et al., 2021). For instance, earlier empirical work concluded that prior experience in distance teaching explained a more positive adaptation to remote teaching on the part of teachers (van der Spoel et al., 2020; Flores et al., 2021b) and that support provided by school leadership (NFER, 2020) and collaboration with colleagues (Delcker and Ifenthaler, 2021) were key influential factors.

Overall, the need to develop digital competencies in teacher education, during both initial and in-service training, has been identified as one of the lessons learned during the COVID-19 pandemic. It is, however, essential to move beyond an instrumental approach to technology and to consider its ethical, political and pedagogical dimensions. This implies moving beyond the questions of "what" and "how" and to consider the "why", "who" and "for what purpose" in order to transcend the bureaucratic (Reeves, 2009) and managerial professionalism (Sachs, 2016) which is marked by a logic of standardization and replication.

If the COVID-19 pandemic had implications for the ways in which we view education and teachers' work, it is also true that we must not forget the problems and challenges that already existed prior to the pandemic. In particular, it is important to consider the conditions for teaching and learning in schools and classrooms, especially the necessary equipment and resources, and how policies affect, and even undermine, curriculum innovation and pedagogy as well as teachers' professional autonomy and collaboration, two key components of teacher professionalism. Teachers should have a say in the definition of their professionalism but also in the direction and focus of their professional growth. This is also something that teachers participating in the two studies acknowledge when they identify professional learning as one of the key aspects of their work. The experience of being a teacher during the COVID-19 pandemic has also entailed a process of learning that teachers tend to appreciate despite the challenges that they had to face when they were obliged to teach and work together in changed circumstances. Therefore, relevant support, including time and resources, and appropriate opportunities for professional development are crucial to enhance more collaborative professionalism (Hargreaves and O' Connor, 2017) and to expand new spaces for joint professional learning and collective action. As Darling-Hammond and Hyler (2020) assert, there is a need to invest in high-quality educator preparation, transforming learning opportunities for professional educators to match current needs, supporting mentoring and the development of new teacher roles, and creating time for educators to collaborate with each other and key partners. While the COVID-19 pandemic has exposed the fragilities of education systems, it has also provided an opportunity for professional learning and enabled the emergence of creative and contextualized responses that need to be continued. But for this to happen, it is necessary to go beyond contingent and temporary responses to a crisis in education such as the one that the COVID-19 pandemic has imposed on teachers. It requires the consideration of teachers' voices, their capacity for action, and investment in policies of teacher education that are consistent and congruent with sophisticated and complex views of teaching and of teachers' individual and collective agency.

Acknowledgments

Funded by the Foundation for Science and Technology (FCT) under the grant SFRH/BSAB/114245/2016 (Study 1) and within the Research Centre for Child Studies of the University of Minho - project under the reference UIDB/00317/2020, UIDP/00317/2020 (Study 2).

References

- Alea, L.A., Fabrea, M.F., Roldan, R.D.A., Farooqi, A.Z., 2020. Teachers' Covid-19 awareness, distance learning education experiences and perceptions towards institutional readiness and challenges. *Int. J. Learn. Teach. Edu. Res.* 6 (19), 127–144.
- Alves, R., Lopes, T., Precioso, J., 2021. El bienestar de los docentes en tiempos de pandemia Covid-19: factores que explican el bienestar profesional. *Int. J. Eng. Res. Innovat.* 15, 203–217.
- Ávalos, B., 2013. Héroes o vilanos? La profesión docente en Chile. *Heroes or Vilans? The teaching profession in Chile*. Editorial Universitaria, Santiago de Chile.
- Ávalos, B., Flores, M.A., Arandeda, S., 2021. Battling to keep education going: Chilean and Portuguese teacher experiences in COVID-19 times. *Teach. Teach. Theor. Pract.* <https://doi.org/10.1080/13540602.2021.2012758>.
- Bahia, S., Freire, I., Amaral, A., Estrela, M.T., 2013. The emotional dimension of teaching in a group of Portuguese teachers. *Teach. Teach. Theor. Pract.* 19 (3), 275–292.
- Ben-Peretz, M., Flores, M.A., 2018. Tensions and paradoxes in teaching: implications for teacher education. *Eur. J. Teach. Educ.* 41 (2), 202–213.
- Bozkurt, A., Sharma, R., 2020. Emergency remote teaching in a time of global crisis due to Coronavirus pandemic. *Asian J. Dist. Edu.* 15 (1), 1–7.
- Carius, A., 2020. Teaching practices in mathematics during Covid-19 pandemic: challenges for technological inclusion in a rural Brazilian school. *Am. Sci. Res. J. Eng. Technol. Sci.* 72 (1), 35–43.

- CNE, 2016. Relatório técnico. A condição docente: contributos para uma reflexão. CNE, Lisboa.
- CNE, 2019. Estudos de seleção e recrutamento do pessoal docente da educação pré-escolar e ensinos básico e secundário. CNE, Lisboa.
- Darby, A., 2008. Teachers emotions in the reconstruction of professional self-understanding. *Teach. Teach. Educ.* 24, 1160–1172.
- Darling-Hammond, L., Hyler, M.E., 2020. Preparing educators for the time of COVID ... and beyond. *Eur. J. Teach. Educ.* 43 (4), 457–465.
- Day, C., 2017. *Teachers Worlds and Work. Understanding Complexity, Building Quality*. Routledge, London.
- Day, C., Flores, M.A., Viana, I., 2007. Effects of national policies on teachers' sense of professionalism: findings from an empirical study in Portugal and in England. *Eur. J. Teach. Educ.* 30 (3), 249–266.
- Delcker, J., Ifenthaler, D., 2021. Teachers' perspective on school development at German vocational schools during the Covid-19 pandemic. *Technol. Pedagog. Educ.* 30 (1), 125–139.
- Estrela, M.T., 2014. Velhas e novas profissões, velhos e novos profissionalismos: tensões, paradoxos, progressos e retrocessos. *Investigar em Educação* 2, 5–30.
- Evetts, J., 2009. The management of professionalism: a contemporary paradox. In: Gewirtz, S., Mahony, P., Hextall, I., Cribb, A. (Eds.), *Changing Teacher Professionalism. International Trends, Challenges and Ways Forward*. Routledge, London, pp. 19–30.
- Flores, M.A., Day, C., 2006. Contexts which shape and reshape new teachers' identities: a multi-perspective study. *Teach. Teach. Educ.* 22 (2), 219–232.
- Flores, M.A., 2012. Teachers' work and lives: a European perspective. In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, London, pp. 94–107.
- Flores, M.A., 2014. Profissionalismo e Liderança dos Professores. De Facto Editores, Santo Tirso.
- Flores, M.A., 2019. O trabalho e a vida dos professores. Um olhar nacional e internacional. Lisbon International Press, Lisboa.
- Flores, M.A., 2020. Surviving, being resilient and resisting: teachers' experiences in adverse times. *Camb. J. Educ.* 50 (2), 219–240.
- Flores, M.A., Van Nuland, S., June 2013. Teachers' understanding of standards. In: 57th World Assembly of the International Council on Education for Teaching (ICET), vol. 86. Sukhotai Thammathirath Open University (STOU), Nonthaburi, Tailândia, pp. 25–28.
- Flores, M.A., Machado, E.A., Alves, P., Vieira, D.A., 2021a. Ensinar em tempos de COVID-19: um estudo com professores dos ensinos básico e secundário em Portugal. *Rev. Port. Educ.* 34 (1), 5–27.
- Flores, M.A., Barros, A., Veiga simão, A.M., Gago, M., Fernandes, E.L., Pereira, D., Ferreira, P.C., Costa, L., 2021b. Ensino remoto de emergência em tempos de pandemia: a experiência de professores portugueses. *Revista Portuguesa de Investigação Educacional* 21, 1–26.
- Gewirtz, S., Mahony, P., Hextall, I., Cribb, A., 2009. Policy, professionalism and practice: understanding and enhancing teachers' work. In: Gewirtz, S., Mahony, P., Hextall, I., Cribb, A. (Eds.), *Changing Teacher Professionalism. International Trends, Challenges and Ways Forward*. Routledge, London, pp. 3–16.
- Hargreaves, A., O'Connor, M.T., 2017. Collaborative Professionalism. WISE (World Innovation Summit for Education) report. Qatar.
- Hargreaves, A., 1998. The emotional practice of teaching. *Teach. Teach. Educ.* 14 (8), 835–854.
- Hargreaves, A., 2000. Four ages of professionalism and professional learning. *Teach. Teach. Theor. Pract.* 6 (2), 151–182.
- Hargreaves, A., 2001. The emotional geographies of teaching. *Teach. Coll. Rec.* 103 (6), 1056–1080.
- Hargreaves, A., 2005. Educational change takes ages: life, career and generational factors in teachers' emotional responses to educational change. *Teach. Teach. Educ.* 21, 967–983.
- Hargreaves, A., 2021. What the COVID-19 pandemic has taught us about teachers and teaching. In: *Children and Schools during COVID-19 and beyond: Engagement and Connection through Opportunity*. An RSC Policy Briefing, pp. 138–167.
- Helsby, G., 2000. Multiple truths and contested realities. The changing faces of teacher professionalism in England. In: Day, C., Fernandez, A., Hauge, T.E., Moller, J. (Eds.), *The Life and Work of Teachers. International Perspectives in Changing Times*. Falmer Press, London, pp. 93–108.
- Huber, S.G., Helm, C., 2020. Covid-19 and schooling: evaluation, assessment and accountability in times of crises—reacting quickly to explore key issues for policy, practice and research with the school barometer. *Educ. Assess. Eval. Account.* 32, 237–270.
- Iivari, N., Sharma, S., Ventä-Olkkonen, L., 2020. Digital transformation of everyday life – how Covid-19 pandemic transformed the basic education of the young generation and why information management research should care? *Int. J. Inf. Manag.* 55.
- Jeffrey, B., Woods, P., 1996. Feeling deprofessionalised: the social construction of emotions during an OFSTED inspection. *Camb. J. Educ.* 26 (3), 325–343.
- Judd, J., Rember, B.A., Pellegrini, T., Ludlow, B., Meisner, J., 2020. "This Is Not Teaching": The Effects of COVID-19 on Teachers. Available at: https://www.socialpublishersfoundation.org/knowledge_base/this-is-not-teaching-the-effects-of-covid-19-on-teachers/ (accessed: 20 October 2021).
- Kelchtermans, G., 2005. Teachers' emotions in educational reforms: self-understanding, vulnerable commitment and micropolitical literacy. *Teach. Teach. Educ.* 21, 995–1006.
- Kelchtermans, G., 2009. O comprometimento profissional para além do contrato: auto-comprensão, vulnerabilidade e reflexão dos professores. In: Flores, M.A., Veiga Simão, A.M. (Eds.), *Aprendizagem e desenvolvimento profissional de professores: contextos e perspectivas*. Edições Pedagogo, Mangualde, pp. 61–98.
- König, J., Jäger-Biela, D., Glutsch, N., 2020. Adapting to online teaching during Covid-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *Eur. J. Teach. Educ.* 43 (4), 608–622.
- Lasky, S., 2005. A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teach. Teach. Educ.* 21 (8), 899–916.
- Lo, L.N.K., 2012. Issues in teacher professionalism and performativity: introduction. In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, London, pp. 13–18.
- Miles, M., Huberman, A.M., 1994. *Qualitative Data Analysis*, second ed. Sage Publications, Thousand Oaks, CA.
- NFER, 2020. Schools' Responses to Covid-19. Pupil Engagement in Remote Learning. Available at: <https://www.nfer.ac.uk/schools-responses-to-covid-19-pupil-engagement-in-remote-learning/> (accessed: 20 October 2021).
- Nias, J., 1996. Thinking about feeling: the emotions in teaching. *Camb. J. Educ.* 26 (3), 293–306.
- Niemi, H.M., Kousa, P., 2020. A case study of students' and teachers' perceptions in a Finnish high school during the covid pandemic. *Int. J. Technol. Educ. Sci.* 4 (4), 352–369.
- Noddings, N., 1984. *Caring: A Feminine Approach to Ethics and Moral Education*. University of California, Berkeley, CA.
- O'Connor, K.E., 2008. "You choose to care": teachers, emotions and professional identity. *Teach. Teach. Educ.* 24, 117–126.
- OECD, 2019. TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. TALIS, OECD Publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>.
- Osborn, M., 2006. Changing the context of teachers' work and professional development: a European perspective. *Int. J. Educ. Res.* 45, 242–253.
- Reeves, J., 2009. Inventing the chartered teacher. In: Gewirtz, S., Mahony, P., Hextall, I., Cribb, A. (Eds.), *Changing Teacher Professionalism. International Trends, Challenges and Ways Forward*. Routledge, London, pp. 106–116.
- Sachs, J., 2016. Teacher professionalism: why are we still talking about it? *Teach. Teach. Theor. Pract.* 22 (4), 413–425.
- Sachs, J., Mockler, N., 2012. Performance cultures of teaching. Threat or opportunity? In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, London, pp. 33–43.
- Schmidt, M., Datnow, A., 2005. Teachers' sense-making about comprehensive school reform: the influence of emotions. *Teach. Teach. Educ.* 21 (8), 949–965.
- Sokal, L., Trudel, L.E., Babb, J., 2020. Canadian teachers' attitudes toward change, efficacy and burnout during the COVID-19 pandemic. *Int. J. Educ. Res. Open* (1), 100016. <https://doi.org/10.1016/j.ijedro.2020.100016>.
- Sutton, R., Mudrey-Camino, R., Knight, C.C., 2009. Teachers' emotion regulation and classroom management. *Theor. Pract.* 48 (2), 130–137.

- van der Spoel, I., Noroozi, O., Schuurink, E., van Ginkel, S., 2020. Teachers' online teaching expectations and experiences during the Covid19-pandemic in The Netherlands. *Eur. J. Teach. Educ.* 43 (4), 623–638.
- van Veen, K., Slegers, P., van de Ven, P.-H., 2005. One teacher's identity, emotions, and commitment to change: a case study into the cognitive–affective processes of a secondary school teacher in the context of reforms. *Teach. Teach. Educ.* 21, 917–934.
- Zembylas, M., 2007. Emotional ecology: the intersection of emotional knowledge and pedagogical content knowledge in teaching. *Teach. Teach. Educ.* 23, 355–367.
- Zhang, W., Wang, Y., Yang, L., Wang, C., 2020. Suspending classes without stopping learning: China's education emergency management policy in the covid-19 outbreak. *J. Risk Financ. Manag.* 13 (3), 55.

Teachers' work in the context of COVID-19

Riley Collins^a, Lora Bartlett^a, Alisun Thompson^b, Judith Warren Little^c, and Lina Darwich^d, ^a University of California, Santa Cruz, Santa Cruz, CA, United States; ^b University of Puget Sound, Tacoma, WA, United States; ^c University of California, Berkeley, Berkeley, CA, United States; and ^d Lewis & Clark College, Portland, OR, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction and framing	636
Shifts in teachers' work amidst COVID-19	637
Teachers' working conditions during COVID-19	638
Work intensification	638
Teacher voice and control over working condition	639
Teachers' pedagogy	640
Digital technology: a crucial resource and a learning curve	640
Adapting curriculum and instruction	641
The unique challenges of hybrid models	641
Context mattered	642
Teachers' relational work	642
Students: disrupted participation, disrupted relationships	642
Families: support, scrutiny, and changing perceptions	643
Colleagues: an important source of support	643
Looking ahead: implications for teachers and teaching	644
References	645

Introduction and framing

In March 2020, the COVID-19 pandemic brought about an immediate “quasi-systematic closure of schools” around the globe (OECD, 2021). Educators rapidly adapted their practice to the new conditions of teaching and learning. These closures affected more than 1.6 billion students and 63 million teachers worldwide (United Nations, 2021), but resulting forms of schooling varied across space and time. Although significant research began early in the pandemic to explore the effects on student learning, scholars paid less attention to the changes in teaching and the effect on teachers. Hargreaves (2021) reflected, “While a lot of the post pandemic [learning loss] debate has focused on adding learning time, increasing technology use, and retaining or reintroducing testing as ways to compensate for learning loss, far less attention has been paid to the short-term and long-term impact of the pandemic on teachers and teaching” (p. 1836).

This chapter addresses the transformation of teachers' work during the COVID-19 pandemic and the possibilities—or missed opportunities—for these shifts to translate to educational change. Birkland (1998) has examined the relationship between “focusing events” and large-scale change. Focusing events can help induce change by revealing deficiencies in existing policy and reconfiguring subsystem dynamics. This chapter draws on Birkland's framework to describe the shifts in teachers' work brought about by COVID-19 and how these short-term changes may reshape educational policy in the long run. Due to the pandemic, ruptures in teachers' work affected three principal domains: *teachers' working conditions*, the *pedagogical work* of teaching, and the *relational work* of education.

This analysis draws from a longitudinal study of teachers working in the United States during the pandemic. From spring 2020 through summer 2022, the Suddenly Distant Research Project (SDRP), designed and conducted by the authors of this chapter, chronicled and analyzed the experiences of seventy-five teachers in public schools in nine different US states. Researchers selected states for a variation on two criteria that suggested potential differences in local experiences: COVID-19 death rates as of early summer 2020 and the strength of teachers' unions (Winkler et al., 2012). Teachers were sampled to capture variations in grade-level teaching, subject matter, and community geography within each state. While not a representative sample, the research mirrors the demographics of the US teacher population (Tae and Gldring, 2020). Data collected, including teacher interviews, surveys conducted twice annually, and an ongoing policy and social context analysis, allowed detailed documentation of teachers' lived experiences over time, variation by context, and similarities that transcend context.

Findings from SDRP are the starting point for analyzing the three critical issues addressed in the sections below: working conditions, pedagogical practice, and relational experience. Each unit introduces significant empirical findings expanded from the project data to include and consider related research conducted during the pandemic. This analysis is rooted in the experience of US teachers and endeavors to situate that experience with parallel and corresponding research in other countries. By considering the extent to which SDRP findings are echoed, refuted, or complicated by related research within and across national contexts, this chapter also presents a fuller picture of teachers' work experience during the pandemic. This work highlights teacher self-reports and includes large-scale data sets that reveal patterns in work experiences, even when teachers were not the primary respondents.

Shifts in teachers' work amidst COVID-19

In the spring of 2020, the initial shutdown of brick-and-mortar schooling across the globe was centralized and swift. From March until June of 2020, all teachers in the SDRP sample taught remotely, as did nearly all teachers in the US and elsewhere (Hargreaves, 2021; Kraft and Simon, 2020; OECD, 2021). Remote teaching encompassed a range of virtual activities, both synchronous and asynchronous, but also—especially where technology resources and internet connectivity were lacking—through print materials, educational television, and radio programming (UN, 2020). According to a 2020 UNICEF study in 71 countries, less than half of the countries' population has access to the Internet. Even in countries with high internet connectivity, inequitable access to digital devices created a divide in learning opportunities between the wealthy and the poor (OECD, 2021).

Efforts to reopen schools in fall 2020 spawned a range of remote, hybrid, and in-person arrangements, each with affordances and challenges. Despite persistent surges in the pandemic, state and local systems rapidly and comprehensively restored pre-pandemic expectations concerning student attendance, standards-based curriculum coverage, grading, and testing. In the United States, politics shaped school district decision-making as political partisanship was associated with district reopening plans (Hartney and Finger, 2020; Grossmann et al., 2021). In fall of 2020, 47% of US schools were in a hybrid model (a combination of remote and in-person instruction), 33% were providing fully remote instruction and 20% were fully in-person (Diliberti and Kaufman, 2020). Hybrid models themselves varied as well. Some approaches required teachers to juggle teaching groups of in-person and remote students simultaneously, while others involved teaching in only one modality, with critical implications for teachers' work (Bartlett, 2022). While the fall of 2021 started out with more than three-quarters of U.S. districts offering only full time in person schooling, these schools still struggled with high pandemic-related student and staff absenteeism, necessitating on-going access modifications to teaching practice (Herold, 2021; Santibanez and Guarino, 2020).

The spring of 2020 initially held some promise of increased teacher flexibility and independence. School administrators entrusted teachers to design the content and structure of distance learning based on understanding of their students' needs, often with little administrative oversight or direction (Bartlett et al., 2020). However, as the pandemic progressed, many teachers reported increased administrative demands on their time, heightened pedagogical scrutiny, and lessened control in relationships with students and families (Diliberti and Kaufman, 2020; Thompson, 2021; Jones et al., 2022). These increased demands continued as teachers moved back into the classroom and were required to manage hybrid learning configurations and enforce masking and social distancing measures, often with little input into school reopening plans (Diliberti and Kaufman, 2020; Shah, 2021). We posit that teachers' work experiences during the pandemic are part of a history of teacher labor intensification and decreased teacher autonomy over the last four decades (Apple, 2009; Hargreaves, 1994; Wong, 2006; Ballet et al., 2006; Fitzgerald et al., 2019).

In the sudden shift to distance learning during the spring of 2020, teachers were charged with making critical decisions about content matter, lesson delivery format, and class structure. Teachers had little experience teaching in a fully remote capacity prior to the pandemic—for example, in a study of 239 U.S. teachers and 574 Norwegian teachers, 92% of US teachers and 67% of Norwegian teachers had not previously taught online (Gudmundsdottir and Hathaway, 2020). Despite this lack of prior experience, Gudmundsdottir and Hathaway highlighted the "agentic stance" that teachers adopted in their capacity to "go the extra mile to move teaching to online platforms" (p. 244). However, while teachers from the SDRP study (Bartlett et al., 2020) reported some transition support from technology specialists, school librarians, and other teachers, they were left mainly "to learn on the fly" during the spring of 2020. In one snapshot of this independence, 60% of teachers in the SDRP study reported a high level of flexibility in shaping their daily and weekly schedule in spring 2020. An SDRP teacher reflected, "We were just expected to make it work."

However, as the pandemic progressed, educators described increased administrative demands on their time, heightened pedagogical scrutiny, and lack of input into school reopening plans despite the earlier authority over critical decisions (Diliberti and Kaufman, 2020; Thompson, 2021; Wilson et al., 2020). Teachers' diminished sense of autonomy over working conditions, pedagogical decisions, and relations with students during COVID-19 can best be understood in the longer history of teacher labor intensification (Apple, 1986; Hargreaves, 1994; Ballet et al., 2006; Fitzgerald et al., 2019) and deskilling of teachers' work (Apple, 1986; Wong, 2006; Chakraborti, 2013). The notion of teacher work intensification, originally outlined by Larson (1980) and further developed by Apple (1986) and Ballet et al. (2006), pointed to the "bureaucratically driven escalation of pressures, expectations and controls concerning what teachers do and how much they should do within the teaching day" (Hargreaves, 1989, p. 23), which is seen "most visibly in the chronic sense of work overload that has escalated over time" (Apple, 2009, p. 188). As these demands on teachers' time have risen, teacher control over curricular and other classroom decisions has eroded. Scholars trace the disempowerment and deskilling of teachers to a process that began in the 1980s (Chakraborti, 2013, p. 185).

These changes occurred within the market-centered transformations of schooling throughout the past four decades, which brought about attempts to tailor education more closely to industry and corporate goals (Klees, 2017; Easthope and Easthope, 2000; Watkins, 1994). Gewirtz (1997) remarked, "The labor process of teaching has always been intense ... what is different in this age of market forces and managerialism is the pattern and texture of intensification—the nature of the tasks that are absorbing increased quantities of teacher time and emotional labor." Decades later, this market-centered logic has continued to inform policy and constrain teachers' work (Klees, 2017). Sahlberg (2016) analyzed the global nature of these transformations, labeled the "Global Education Reform Movement," which is characterized by market-like competition within education systems, standardization and control over teaching and learning, and the use of corporate management models.

At the same time, Ballet and Kelchtermans (2009) emphasized how "teachers do not blindly accept [policy] changes but interpret them using their personal opinions about 'good education' and their sense of competence as a starting point. Such sense-making 'filters' the calls for change" (Ballet and Kelchtermans, 2009, p. 1154). Thus, teachers mediate work intensification through

their process of interpretation and sense-making, which are influenced by organizational working conditions (Ballet et al., 2006). Ballet, Kelchtermans and Loughran also posited that teacher work intensification stems from multiple sources, including both external and internal demands. Their analysis aligns with Coburn's (2001) work on how teachers co-construct understandings of policy and negotiate its implementation through their daily interactions. Fitzgerald et al. (2019) built on Ballet and Kelchtermans' (2008, 2009) scholarship to explore in "fine-grained" ways the differences in work intensification experiences, including distinctions in socio-educational school status, primary or secondary teacher orientation, and local context.

These calls for attention to how teachers have experienced intensification and lessened autonomy in different ways across contexts are useful for understanding teachers' varied experiences during COVID-19. However, when teachers faced these increased demands during the pandemic, they resisted them through creative pedagogical, relational, and personal strategies. Such adaptations allowed teachers to connect with students from a distance and protect their time (Bartlett et al., 2021). Teachers prioritized ethics of care for students despite the lack of physical proximity (Jones and Kessler, 2020), used inventive strategies to engage students even when learning configurations were designed without teacher input (Bartlett et al., 2021). Teachers also led sickouts and demonstrations to influence COVID-related decisions by administrators and elected officials.

In continuation with the history of teacher labor intensification and decreased autonomy, teachers' experiences during the pandemic help to illuminate potential longer-term shifts in their work. In what follows, we explore these implications across three domains: working conditions, teachers' pedagogy, and teachers' relational work.

Teachers' working conditions during COVID-19

The literature on teachers' working conditions has established the dynamic interplay of occupational norms, institutional policies, and social aspects of the workplace and profession that affect teachers' work engagement and career commitment (Merrill, 2021; DeAngelis and Presley, 2011; Ingersoll, 2001; Johnson et al., 2012). This section explores two essential aspects of the radical disruption of teachers' work and their implications for teachers' job satisfaction and commitment. First, it considers how the pandemic intensified teachers' work, the effect on working hours, increased teacher learning demands, and the burden of related stress and anxiety. In addition, it explores how teachers' experience of work intensification was amplified by their lack of voice as schools re-opened. It illuminates the manner in which educators navigated and contested these shifting circumstances. Specifically, it contrasts the early days of the pandemic, when teachers were granted increased autonomy and flexibility to identify and respond to students' needs, with the subsequent exclusion of teacher input over re-opening decisions and instructional modality structures. Teachers' experiences during the COVID re-opening correspond to past tensions over the control of teachers' work (Ingersoll, 2009; Apple, 2009).

Work intensification

The pandemic intensified teachers' work as they made rapid changes to their teaching practice in a dramatically altered work context charged with the associated stress of a world health crisis. Teacher work intensification is characterized by a lack of time, chronic and persistent work overload, the replacement of time spent teaching students with administrative demands, the enforced diversification of expertise, and the reduction in professional curriculum discretion (Hargreaves, 1994; Easthope and Easthope, 2000). One thread of pandemic-related research on teachers' experience focuses on determining the degree to which changes to school schedules and delivery models increased or decreased teachers' working time. Several reports found teachers working hours increased in the early days and continued through the pandemic (Kaufman and Diliberti, 2021; Kurtz, 2020; Long, 2020). Other reports showed a decrease in working hours at particular points in the pandemic (Gicheva, 2021; Herold and Kurtz, 2020). However, pertinent scholarship has further explored whether and how the quality of teachers' work was affected and intensified. Jones et al. (2022) contend that the pandemic changes dramatically redirected teacher work time in ways associated with lower job satisfaction. Notably, teacher time spent in instructional practice with their students decreased to 17% from 52% in the pre-pandemic. Time spent on administrative tasks, contact time with administrators, and communication with families increased. The SDRP data indicated a similar escalation of complexity, with 75% of teachers in November 2021 reporting overwork and exhaustion. These feelings were attributed to the intensity of the learning demands associated with instructional modality changes, the collapse of the boundaries between work and home, and the stress and anxiety of situational uncertainty and instability. In these ways, teachers' work was intensified in its complexity even when it did not increase in duration (Wilson et al., 2020; Fitzgerald et al., 2019).

The pandemic required many teachers to develop new technical skills and knowledge, often while physically removed from the regular structural supports for their work and under the duress of the pandemic. Changes in instructional delivery models required teachers to develop substantial new technical and related instructional skills. A 2021 OECD survey of teachers' experience in 37 countries reported that fewer than half (43%) felt "well prepared" or "very well prepared" to integrate technology into their teaching before the pandemic. Relatedly, in July 2020, more than half (54%) of 757 U.S. teachers surveyed reported having no previous experience with online teaching (Bartlett et al., 2020). Only 3% indicated they had substantial online teaching experience. And yet, in November 2020, 95% of SDRP teachers reported their schools were fully remote or operating in some form of hybrid instructional mode. Even the few schools that adopted in-person learning options reported supporting quarantined students through remote instruction. This rapid shift to simultaneous remote and online education represents a massive increase in teachers' technical

abilities during high stress and compromised institutional capacity. During the pandemic, experienced teachers found instruction challenges akin to being a “1st-year teacher” again.

In a survey of over 10,000 Australian teachers, respondents identified a work intensification resulting from a sharp escalation in the complexity of their work and an increase in administrative tasks, along with a sense of feeling undervalued by the government (Wilson et al., 2020). The research found that this complexity was compounded by the stress of working in “Covid-wary classrooms” and, noted the demands of up-skilling, the shift to teaching via remote learning. In March 2020, over 80% of responding SDRP teachers indicated they spent more time on lesson planning, changing instruction, communicating with students and families, attending school meetings, and grading.

These complexities were unfolding in a new stretched out workday that saw teachers working at all hours of the day to accommodate parents' working schedules and encourage student engagement. In the spring of 2020, a school day that formerly had a clear start and end time now extended into the evening with an elementary teacher describing “there was just no cut-off.” SDRP teachers of teenagers reported students often did not show up during regular school hours. One high school teacher described that after the official workday ended then:

Messages would start trickling in from the kids who were just waking up or just got off work—whatever it was—and I would just keep answering them. I mean everybody was so stressed and I wasn't sleeping well. I would get messages from kids at one or two in the morning and I'm already on my phone. The notification pops up and I'm like, “OK, I'll just answer you.”

The stretched-out workday was compounded for parenting teachers, especially mothers, who found themselves doing double duty. Teaching is, in most countries, gendered work historically considered compatible with family responsibilities (Lortie, 1975; Ingersoll et al., 2018). In the US, the teacher workforce is 77% female and in most parts of the world women comprise a large and increasing share of the profession (Le Nestour and Moscoviz, 2020; OECD, 2022). Thirty-one (41%) of the SDRP teachers had dependent children at home; four fathers and twenty-seven mothers. A survey by the Pew Research Center, found that working mothers assumed more childcare duties than working fathers during the pandemic, even when both parents were working from home (Igielnik, 2021). Other studies have corroborated the unequal division of domestic labor during the pandemic and the lasting effect this might have on women in the workforce (Aaronson et al., 2021; Bateman and Ross, 2021; Tedeschi, 2020).

Nineteen of the twenty-seven SDRF female teachers were the primary or sole domestic care providers. This was true even when these teachers had partners also working from home. As one elementary school teacher with school-aged children described, “My husband was also forced to stay home for work ... so he tried to help. But his schedule is not that helpful for my own children. And he doesn't know all the programs and the ins and outs. So, it was kind of left up to mom and teacher and that was a big struggle.” Another teacher, also the primary caregiver, described, “So it was one or the other. I could be a teacher to my fourth graders or I could be a parent to my kids, but I couldn't be both.” She explains the consequences of competing role demands on her sense of efficacy, “So that's another thing I think that's important. That's a piece that played into the emotion and the exhaustion as well. Feeling horrible as a parent and feeling crappy as a teacher. And then there's a pandemic and I don't want people to get sick.”

The pandemic also created generalized anxiety and stress that heightened work intensity. In March 2021, 82% of responding SDRP teachers (N = 71) indicated they were experiencing more or much more stress than they were pre-pandemic. They worried about student well-being, their capacity to meet student needs, and the health risks of everyone. Stress levels stayed high even into the third school year affected by the pandemic. In November of 2021, even after 85% of teachers (N = 62) had returned to teaching predominantly in-person, 47% indicated their work-related stress levels were higher than, and 31% reported the same levels as, during the 2020–21 school year. These same samples self-identified key aspects of affect associated with professional burnout: 52% indicating in November 2020 an increased cynicism about their work as a teacher, 44% a diminished capacity to support others, and 62% a diminished sense of success as a teacher. In a longitudinal study of four Texas high schools, Jabbar et al. (2021) found high degrees of increased work-related stress and signs of burnout reported by 80% of responding teachers. Based on a US sample survey, Pressley (2021) identified anxiety about COVID-19 and teaching demands as two of the four predictive variables for teacher burnout during the pandemic. Other research similarly found feelings of professional loss, uncertainty, and increased burnout risk exacerbated by the pandemic disruptions to teachers' understandings of what it means to be a teacher, the increased work demands, and the decline in the relational aspects of schooling that make teaching enjoyable (Reich et al., 2020; Steiner and Woo, 2021; Kauffman and Diliberti, 2021; Alves et al., 2021).

Teacher voice and control over working condition

As teachers experienced this intensification of work demands, they also navigated an eroding sense of control over their work. At the start of the COVID-19 pandemic, teachers were charged with swiftly adapting all aspects of their work to a remote context with little administrative oversight. One teacher from the SDRP study recalled, “We were given 2 h to switch to distance learning.” Teachers described a high degree of autonomy to shape teaching and learning conditions in this initial period. In a spring 2020 study by the Alberta Teachers' Association in Canada, 80% of teachers agreed or strongly agreed that they had “a great deal of professional autonomy with how to teach my students during the pandemic” (Hargreaves, 2021, p. 31). In another snapshot of this autonomy, 60% of teachers in the SDRP study reported complete flexibility to shape their daily and weekly schedules.

Yet as the pandemic progressed, teachers described heightened accountability measures and a lack of influence in shaping plans for the 2020–21 school year, particularly related to school re-opening and instructional modality structures. Apart from the role played by unions or teacher associations, local districts and schools established few mechanisms for teacher input into fall 2020 reopening plans. An SDRP teacher summed up her frustration:

There was no teacher input. So, when we started to hear the [fall reopening] plan, it was kind of like, “Why? What would make you think that was a good idea?”

Only about one-third of SDRP teachers felt that teachers had been solicited in any meaningful way for their input. While some were satisfied with reopening plans (24%), others reported mixed feelings and a quarter was so unhappy they were contemplating leaving teaching. These findings were mirrored in a RAND study of over 1000 U.S. teachers. Most teachers were dissatisfied or had mixed feelings about their districts’ decisions regarding remote or in-person learning ([Diliberti and Kaufman, 2020](#)).

An OECD report (2021) on how countries managed the response to COVID-19 in education revealed that, while more successful systems involved collaborative decision-making structures across all levels, less successful responses subjected educators to contradictory and constantly changing policy mandates. During the 2020–21 school year, 88% of SDRP teachers experienced some change to the schooling format throughout the year, and for almost one-third of teachers, the school format changed four or more times. Teachers were significantly affected when these decision-making processes were opaque or did not involve educators’ input. One SDRP teacher remarked, “Our district has continually changed protocols and decisions that were very much not founded in science and safety but instead appealed to the loud voices in the community.” Another SDRP teacher lamented, “I am deeply saddened and frustrated by my district’s response and honestly am very defeated by their lack of communication, planning, and response to teacher’s concerns.”

Teachers contested and creatively resisted these increased demands individually and collectively. Some teachers used strategies to circumvent particular policies they did not support. For example, one SDRP high school teacher disagreed with his district’s mandate that teachers report to and teach from their empty classrooms during remote learning. He contested the policy while avoiding disruption to student learning by drawing on his bank of sick days to remain at home while continuing to deliver lessons. He explained: “The kids are still learning and I’m not going to the building. So, it was a win-win for everybody, in my opinion.” Teachers also organized collectively to demand better working conditions and protections ([Mavhunga, 2021](#); [Will, 2020](#); [Mehéut, 2022](#)). Another SDRP teacher reflected:

I don’t feel that teachers are being heard. And I actually started some resistance ... I sat in on a meeting with teachers across the state and people were saying they’re scared and worried. And so, I ended up starting a group—we call ourselves The Motor Marches, and we organized all the marches, the motor parades that went on throughout the state.

In these ways, teachers both influenced and were influenced by the policy changes related to COVID-19, aligning with [Ballet et al. \(2006\)](#) analysis of how teachers make sense of and “filter” education policy change. Teachers’ experiences during the pandemic fit within a history of tensions regarding the control over teachers’ work. At the start of the pandemic, the autonomy granted to shape their daily schedules, lesson structure, and other working conditions departed from decades of tightened controls over teacher labor ([Ingersoll, 2009](#)). However, the subsequent erosion of teacher autonomy alongside teacher experiences of work intensification reinforced these broader patterns and exacerbated burnout conditions.

Teachers’ pedagogy

The substance and scope of the curriculum and the face of teaching both changed as teachers first accommodated the constraints of teaching remotely and then navigated the various instructional configurations that marked the 2020–2021 school year. In surveys and interviews, SDRP teachers highlight four aspects of teachers’ pedagogical experience during the period of school closure and then with the partial or full reopening of schools: the immediate pivot to and continuing reliance on digital technology; instructional adaptations and accommodations that rendered “active learning” difficult to plan for and organize in remote and hybrid arrangements; challenges associated with hybrid instructional configurations introduced by schools and districts as schools reopened; and the variation in available resources and in experiences associated with student demographics (race, ethnicity, and socio-economic levels) and community context. In these four respects, findings from the SDRP study parallel those from other COVID-related research.

Digital technology: a crucial resource and a learning curve

The pandemic swept the world at a time when online learning management systems and digital teaching and learning tools were increasingly common features of the schooling landscape. Many districts and schools, although certainly not all, had invested in

devices for students and teachers. Technology-related professional development opportunities had proliferated. Yet teachers brought to remote teaching very uneven levels of prior experience with online platforms and other digital tools. Some teachers—although a small minority—entered the shutdown with substantial experience using Google Classroom, other online platforms, or digital tools to support student learning in the classroom. However, even those teachers had not used such systems and tools to teach entirely from a distance.

Teachers reported some transition support from technology specialists, school librarians, grade-level and subject-specific colleagues, and from online sources such as Facebook teacher groups. However, they were still primarily on their own to master new learning platforms and digital tools. The experience of the SDRP teachers mirrors that of teachers elsewhere. An OECD survey of teachers' experience in 37 countries reported that fewer than half (43%) felt "well prepared" or "very well prepared" to incorporate technology in their teaching (OECD, 2021; see also Woltran et al., 2021). By the time schools attempted to re-open, many teachers had taken part in professional development using virtual platforms and digital tools. In one large-scale survey of a nationally representative sample of US teachers, 62% said they had received training on using the technology. However, fewer teachers found that professional development focused on promoting academic learning or engaging and motivating students from a distance (Hamilton et al., 2020).

Adapting curriculum and instruction

In the early stages of the pandemic, schools and districts concentrated on reviewing content already taught, noting students' uneven access to technology and other conditions that might affect student participation (Herold and Kurtz, 2020). Despite that policy stance, most teachers in the SDRP sample (75%) introduced some new content, and a few reported introducing all new content in the spring of 2020. Yet they also acknowledged that not all their students—and in some cases very few—accessed that content or completed assignments. A common refrain about new content was that it was "made available" but few students engaged with it. Over 90% of SDRP teachers, including those who had introduced "substantial" new content, taught less content than they would have otherwise and made selective decisions about what content to prioritize. One consideration was the perceived difficulty of teaching particular content without the support afforded by in-person instruction. RAND's nationally representative sample of teachers found that only 12% reported covering all or nearly all the curriculum they would have otherwise covered (Hamilton et al., 2020). No-harm policies adopted by districts in spring 2020—especially restrictions on grading and relaxed student attendance requirements—complicated teachers' efforts. In most cases, student participation plummeted—most drastically among students from lower income households, students with special needs, English Language Learners, and students who were already struggling academically.

Teachers at every level found themselves forced to modify the instructional strategies they had honed, often over decades, to support student learning. In every instance, the pedagogical shift represented a move away from what could be considered active learning—learning that entails substantial student-to-student collaboration, sense-making, and problem solving. A middle school math teacher, unable to organize students in small groups and foster student-to-student sense-making, and forced to supply explanations and teach explicit procedures, said, "I got really good at exactly the kind of teaching that I don't believe in. I figured out procedure-oriented goals because that's the only thing that I had any confidence in when there was no opportunity for [classroom] discourse." An elementary teacher complained that without rich classroom discussion, the quality of students' work went down, and added, "I think the quality of *my* work went down."

Other qualitative research yields similar portraits. Horn and McGugan (2020) interviewed 11 secondary mathematics teachers who had expressed and demonstrated a commitment to ambitious and equitable teaching: an emphasis on conceptual understanding, use of rich mathematical tasks to elicit and build on student thinking, opportunities for student collaboration and rich mathematical discourse, and efforts to foster students' cultural and academic identities. Teaching remotely, the teachers struggled to create lessons, assignments, and assessments consistent with those commitments, and to engage with students in ways that enabled collaboration and invited discussion. The teachers who attempted live synchronous instruction found themselves stymied when students kept cameras turned off, a common experience among the SDRP teachers as well. In the classroom, teachers observe body language and facial expressions to gauge a student's cognitive understanding or emotional state; with cameras off, teachers found themselves teaching to "a sea of black boxes," or a "blank screen."

Although many teachers discovered a range of new online resources and invented new ways of conducting instructional activities in a remote or virtual environment, they still found the experience far less satisfying than being in a classroom. The pandemic transformed teaching from a highly embodied, interactive experience in the intensive immediacy of the classroom to one that numerous teachers described as a "desk job," "chained to a computer." That transformation did not end for most teachers when schools sought to re-open in fall 2020, and more than half (61%) reported in March 2021 a diminished sense of success as a teacher.

The unique challenges of hybrid models

With the onset of the 2020–2021 school year, and with the pandemic not yet abated, teachers found themselves in one of several teaching modalities: still teaching remotely; teaching both remotely and virtually at different times; teaching remotely and virtually at the same time; or, in a relatively few cases, teaching entirely in person (Bartlett, 2022; Kraft, 2021). At the time of the November 2020 teacher survey (N = 73), nearly two-thirds (63%) of SDRP teachers reported working fully remotely, and more than one-

quarter (26%) were teaching in some hybrid configuration. Those required to teach both in-person and remote students simultaneously experienced the highest demands and voiced the greatest fatigue and frustration (see also Kraft et al., 2021).

By March of 2021 only about one-third of teachers were still fully remote (31%) and nearly half (47%) were teaching in some form of hybrid arrangement. Yet few teachers at that point—a year after the shutdown—reported that they could use most or all of their preferred instructional methods; student engagement remained a significant challenge for 80% of SDRP teachers and 78% were especially concerned about their most vulnerable students. By November of 2021, instructional configurations had shifted dramatically, with 88% of teachers teaching in person, a figure that corresponds to findings of a large-scale survey of teachers in 14 states (Conklin and Kraft, 2022). Three-quarters of SDRP teachers reported that they were able to use most or all of their preferred methods, although almost half (47%) still found student engagement to be a significant challenge and more than half (53%) still expressed concern about meeting the needs of their most vulnerable students. An interview- and observation-based case study of teachers in four high schools similarly found that hybrid configurations placed heavy burdens on teachers, with more than 80% rating student engagement a significant challenge under pandemic conditions (Jabbar et al., 2021).

Context mattered

The broad patterns outlined above mask significant contextual variations that shaped teachers' curricular decisions, instructional strategies, and interactions with students. Technology access and expertise varied at the teacher, student, and system levels. Teachers and students in rural areas were especially hampered by limited access to technology; veteran teachers faced a steeper learning curve and often turned to younger colleagues for help. Teachers in more affluent and predominantly white communities were more likely to secure student participation and engagement, while teachers in more diverse and low-income communities struggled to maintain contact with students and families and to supply adequate instructional support. Teachers of special needs students and English Language Learners similarly found sustained contact and support difficult. Teachers of young children found themselves dependent on parent support to engage the children in academic activity.

In contrast, secondary teachers contended with students caring for younger siblings, working a job, seeking instructional help at odd hours, or electing not to show up at all. Finally, teachers' instructional guidance, support, and leadership varied from local schools and districts. These same contextual variations surface in a range of survey-based and interview studies (Balfe et al., 2020; Hamilton et al., 2020; Jabbar et al., 2021; Kraft and Simon, 2020; Kraft et al., 2020, 2021; Reich et al., 2020). Throughout the various transitions, teachers with favorable working conditions and strong relationships with colleagues felt better able to manage the instructional challenges posed by teaching in a pandemic.

Teachers' relational work

Relationships are the heart of teaching (Hargreaves, 2001; Lortie, 1975; Nieto, 2003; Kim et al., 2021). Teachers' relationships with their students and their colleagues have been strongly linked to their resilience and wellbeing, and to feelings of professional fulfillment and success (Gu, 2014; Kelchtermans, 2017; Le Cornu, 2013). Until March 2020, teachers' relationships were typically forged and nurtured within the context of brick-and-mortar buildings. However, the disruption of in-person teaching profoundly affected these relationships in at least three ways. First, students' attendance and engagement with school plummeted and teachers found themselves missing their students and their classroom rituals. Second, the boundaries between home and school became blurry, with parents, especially those who shifted to working from home, getting front row access to their child's learning and teacher. Finally, pandemic teaching brought with it a number of unprecedented constraints (e.g., teaching from home in spring 2020 and social distancing and masking if teaching in person in 2020–2021) isolating teachers from their colleagues.

Students: disrupted participation, disrupted relationships

Student participation dropped precipitously in the spring of 2020 with 82% of SDRP teachers reporting a loss of student presence. The drop was attributed to insufficient technology infrastructure and challenges navigating the technology needed to access school for elementary teachers. For secondary teachers, it was fundamentally related to student absences, as many school districts eliminated attendance requirements and passed policies prohibiting failing grades (Bartlett et al., 2020). A high school science teacher in Massachusetts shared his struggles to stay connected with students,

Even that first quarter that ended the beginning of April [2020], I think, and to give you the harsh truth, I lost most of my kids. In the very beginning I heard from maybe 15% of them. I have 150 students. Then it went to 10%. And then I was literally just begging them to answer my messages and tell me that they and their families were OK.

Student engagement continued to be a challenge during the academic year 2020–2021. Teachers who taught online struggled to engage students who joined class with their audio muted and their cameras off. Teachers who taught in person had masked students in their classrooms and social distancing measures inhibiting movement. In a November 2020 survey, 75% of SDRP teachers cited

student engagement as a major challenge while also indicating their greatest source of satisfaction arose from their personal connections with students. Several studies (Kim et al., 2021; Niemi and Kousa, 2020; Reich et al., 2020; Tackie, 2022; Woltran et al., 2021), detailed the negative implications of this relational loss. In England, the importance of relationships with students emerged as a prominent theme in an interview-based study of 24 primary and secondary teachers from 20 schools: "In stories about teacher-pupil relationships, teachers expressed sadness and frustration about lack of contact with the pupils but seemed to feel a real sense of joy when they were in touch" (Kim et al., 2021, p. 252).

SDRP teachers acutely felt the loss of their daily, non-academic, interactions with students emphasizing the inherently relational nature of teaching. In referring to missing her students, one middle school teacher in Texas (T25 Aug. 5, 2020) expressed that "the reason I like my job is gone." The loss of student participation and engagement flattened relationships and pared down classrooms to spaces of content delivery, not the vibrant interactive environments that bring teachers pleasure in their work. Drawing from Noddings' (2013) concept of confirmation in relational care, Jones and Kessler (2020) discuss how changes to teacher student relationships from COVID conditions challenged teachers' identities and how they saw themselves as teachers. Indeed, SDRP teachers reported declines in feelings of efficacy as a teacher and increases in indicators of burnout, including increased cynicism and a decreased capacity to support others.

Families: support, scrutiny, and changing perceptions

Teacher-parent/caregiver relationships were also affected. The widespread shift to online schooling brought the work of teaching into homes like never before. In spring 2020, teachers, particularly elementary school teachers, provided parents with emotional and technological support. One SDRP elementary teacher (T24 Aug. 11, 2020) shared that half of her job involved supporting parents, showing them how to use Google Classroom and Zoom. An elementary teacher recalled her efforts to soothe parents, "I can't tell you how many parents I had to Zoom with or call just to talk them off the ledge because the stress level was just incredible." Recognizing that many parents were balancing their work demands and the home-based schooling of their young children, teachers made themselves available outside typical school hours. For example, a Kentucky elementary school teacher would take parent calls at 10 p.m. to accommodate work schedules.

Teachers spoke of parents' appreciation of their work and many felt that parents recognized their efforts and effectiveness. A Florida elementary school teacher expressed, "In the spring (2020) teachers were perceived as being miracle workers. We turned around on a dime and, you know, took our teaching remote and did the best we could with it." She experienced this appreciation, however, as short lived. Amid debates about fall 2020 returns to school, this teacher felt parental sentiment shifting, "And now I think teachers are being seen a little bit as the bad guy when they question the safety and whether or not they want to go into the classroom with kids in this situation ... they (parents) are not going to remember that they thought we (teachers) were heroes in the spring." Teachers felt demoralized by this noticeable shift in parents' perceptions. For example, an elementary school teacher in Massachusetts felt discouraged by the name calling from parents "I find it kind of discouraging because to look on social media ... as there is a lot of anger toward teachers and it's not simply parents saying, you know, 'I'm in a bind.' It's some really nasty name calling." Another elementary school teacher in Texas was upset to see that teachers' work merely involved looking after children while parents went to work "But I don't think that is the main point of our work. And so that's been. I think (it's) hard for me to just happen to be hearing a lot of that (parents saying teachers should go back to in person schooling) and seeing a lot of that."

Other studies have also indicated that teacher-parent relationships were affected by the pandemic. For example, in Norway, Bubb and Jones (2020), and in the UK, Kim et al. (2021) found that parents, especially those of young children, were very appreciative of teachers' support in spring 2020. In contrast, in Canada, the Alberta Teachers Association reported teachers feeling discouraged by parental frustration directed at them (Alberta Teachers Association, 2020; see also Fontenelle-Tereshchuk, 2021). Although additional research is needed to further understand whether the schools reopening discourse was as polarized in some countries as in the US, early reports from Canada and the UK (see Hargreaves, 2021; Marchant et al., 2021) suggest school reopening in fall 2020 was a subject of heated debate that caused schisms in some communities.

Colleagues: an important source of support

During school closures and throughout the 2020–2021 school year, teachers in the SDRP sample turned to other teachers for instructional and emotional support. More than two-thirds of the focal teachers interviewed relied on existing relationships in their own schools ("a very collaborative team," "good grade level teams," "a strong community of teachers"). Without a blueprint for pandemic teaching, teachers supported each other in multiple ways. They collaboratively learned and introduced each other to new technology, shared resources, planned lessons together, divided the workload, and helped one another reach out to students and families.

Those who collaborated closely with school-based colleagues credited that experience with easing their own stress and creating a more coordinated experience for students. A fifth grade teacher recalled the early days of the shutdown:

We kind of jumped into our own challenges and were each hosting our own lessons for our own classes. And then after three weeks, we were completely overwhelmed, and we got together and decided to collaborate. And it went a lot smoother after that.

Although teachers described technical and instructional support in some detail, they also highlighted the significant social and emotional support they enjoyed from colleagues. A high school teacher reported that “a lot of us have really banded together and have been some pretty strong emotional support for each other, so that’s been very powerful We were kind of being therapists for each other.”

Some teachers acknowledged stark contrasts between strongly or weakly collaborative contexts. An elementary teacher who has “always had really collaborative relationships” expressed gratitude for those relationships in part because “I talk to colleagues around the district ... who are not a part of collaborative teams, and they’re sinking, I mean, they’re just drowning.” Another speculated on the likely importance of the pre-pandemic school workplace culture:

I suppose if you were working well as a team before, then you are going to continue to work well as a team with Covid. ... If you were dysfunctional prior to Covid, you’re going to be even more dysfunctional now.

The experience of the SDRP teachers mirrors that of teachers in other studies. [Jabbar et al. \(2021\)](#), in an ongoing study of teacher retention, found teachers’ collegial interactions varied. Although teachers in that study also reported some crucial support from their colleagues, 61% reported that the pandemic had a negative effect on collaboration and collegial support. Distance deprived them of informal conversations and created a barrier to co-planning. [Johnson \(2020\)](#) explored the experience of teachers and administrators in one high-poverty district following the shutdown, with findings that underscore the importance of workplace culture and social organization. She writes:

In schools where teachers were accustomed to collaborating regularly on well-established teams, they swiftly moved their meetings online and continued to deliberate together as they made difficult decisions based on their best collective knowledge and judgment. ... By contrast, in compartmentalized schools, where professional norms of self-reliance and autonomy prevailed, teachers were on their own and students often paid the price (p. 2).

The teachers in SDRP, like participants in other studies, reaffirmed that despite the disruptions brought by distance learning, relationships remained the heart of teachers’ work.

Looking ahead: implications for teachers and teaching

In 2021, the United Nations named teachers the “driving force behind global education recovery from COVID-19” ([United Nations, 2021](#)). This calls upon educational policymakers, scholars, and all involved in educational change to reconsider what it means to support teachers’ work. The pandemic highlighted, as crises tend to do, the core satisfactions and stressors of the work. It also disrupted, albeit briefly, a narrow and technocratic framing of accountability as it related to teachers’ work. This moment has the potential to be a focusing event ([Birkland, 1998](#)), as it could provide the impetus and insight needed to rethink the structure and support of teachers’ work.

Experiences amid the COVID-19 pandemic underscore the extent to which teachers’ sense of success derives from daily interactions with students, discretion over instructional practice, and influence on their working conditions. Teachers struggled to find energy and stay optimistic when intrinsic rewards of teaching were disrupted: student relationships, community connections, and collegial collaboration. Their sense of efficacy was eroded when compelled by schooling modifications to develop practices contrary to their pedagogical philosophy. Foregrounding these priorities, they used inventive strategies to remain connected to students from a distance and contested learning configurations designed without their input.

In the spring of 2020, emergency conditions led many places to abandon a narrowly defined and externally managed version of accountability. Instead, school and community leaders asked teachers to use their professional knowledge and experience to customize educational programs to local circumstances. Teachers looked to one another to build capacity for the technical and pedagogical shifts needed to meet the moment. New professional knowledge networks emerged and established ones deepened as teachers determined student needs and developed tools for teaching and learning. Bureaucratic conceptions of accountability were replaced with a more broadly defined accountability rooted in the profession itself: one that required and drew on teachers’ professional expertise. This disruption to the narrow conception of teacher accountability was short-lived as fall 2020 efforts to resume schooling norms also reasserted technocratic control of teachers’ work.

A longer view of teachers’ work sheds light on how teachers’ experiences during the pandemic fits into broader patterns. If four decades ago marked the beginning of the “age of market forces and managerialism” in schooling that reshaped the “pattern and texture of [teacher work] intensification—the nature of the tasks that are absorbing increased quantities of teacher time and emotional labor” ([Gewirtz, 1997](#)), then the COVID-19 pandemic arrived at a time when these pressures were coalescing on the teacher workforce. Scholarship has documented a rise in teacher burnout during the 21st century tied to the technocratic accountability movement ([Dworkin, 2014, 2009](#); [Ryan et al., 2017](#)) and demonstrated that teacher shortages have worsened with particular

consequences for high-poverty schools (García and Weiss, 2019). Thus, teachers' reported discontent with the relational aspects that ground their work in conjunction with their experience of work intensification during the pandemic should compel attention.

Research is considering the pandemic's longer-term implications for teachers' work. One debate centers on educator recruitment and retention. Some reports identify growing dissatisfaction among teachers as an indication of a looming shortage (Walker, 2022), while others counter that longitudinal teacher attrition data do not support these claims (Goldhaber and Theobald, 2022). However, just as working hours alone are not an adequate measure of work intensification, attrition rates are just one of a range of effects on the teacher workforce. Teachers' experiences during the COVID-19 pandemic raise key concerns related to educator wellbeing, sense of success, and experience of burnout. Moreover, if educators' wellbeing is intimately tied to student wellbeing, then "we cannot serve our students by sacrificing our teachers" (Hargreaves, 2021, p. 1852).

The pandemic could be the impetus to shift policy accountability orientations toward teachers' work from one of tight control to one of increased voice and autonomy. Indeed, teachers themselves have led efforts to influence, reinterpret, and contest policies and to demand quality teaching and learning conditions throughout the pandemic. It remains to be seen which shifts related to teachers' work will be fleeting and which will lead to enduring change—and which direction change may take.

References

- Aaronson, D., Hu, L., Rajan, A., 2021. Did Covid-19 Disproportionately Affect Mothers' Labor Market Activity? Chicago Fed Letter Essays on Issues.
- Alberta Teachers Association, 2020. Alberta Teachers Responding to Coronavirus (COVID-19): Pandemic Research Study Initial Report. ATA, Alberta, Canada. Retrieved from. <https://www.teachers.ab.ca/SiteCollectionDocuments/ATA/News%20and%20Info/Issues/COVID-19/Alberta%20Teachers%20Responding%20to%20Coronavirus%20%28COVID-19%29%20-%20ATA%20Pandemic%20Research%20Study%20%28INITIAL%20REPORT%29-ExSum.pdf>.
- Alves, R., Lopes, T., Precioso, J., 2021. Teachers' well-being in times of Covid-19 pandemic: factors that explain professional well-being. *IJERI* (15), 203–217.
- Apple, M., 1986. Are teachers losing control of their skills and the curriculum? *J. Curric. Stud.* 18 (2), 177–184.
- Apple, M.W., 2009. Controlling the work of teachers. In: *The Curriculum Studies Reader*, vol. 3, pp. 199–213.
- Balfe, C., Harris, D.N., Liu, L., Mattei, N., Oliver, D., Slaughter, S., 2020. How America's Schools Responded to the COVID Crisis. REACH. <https://www.reachcentered.org/publications/how-americas-schools-responded-to-covid>.
- Ballet, K., Kelchtermans, G., Loughran, J., 2006. Beyond intensification towards a scholarship of practice: analysing changes in teachers' work lives. *Teach. Teach.* 12 (2), 209–229.
- Ballet, K., Kelchtermans, G., 2008. Workload and willingness to change: disentangling the experience of intensification. *J. Curric. Stud.* 40 (1), 47–67.
- Ballet, K., Kelchtermans, G., 2009. Struggling with workload: primary teachers' experience of intensification. *Teach. Teach. Educ.* 25 (8), 1150–1157.
- Bartlett, L., Thompson, A., Darwich, L., Little, J.W., Collins, R., Weaver, I.H., Harte, L., Ramirez, L., 2020. Suddenly Distant: Teachers' Work During COVID-19 in Spring 2020. Suddenly Distant Research Project. <https://sites.google.com/ucsc.edu/suddenlydistant/home>.
- Bartlett, L., 2022. Specifying hybrid models of teachers' work during COVID-19. *Educ. Res.* 51 (2), 152–155.
- Bartlett, L., Thompson, A., Darwich, L., Little, J.W., Collins, R., Weaver, I.H., Harte, L., Ramirez, L., 2021. Suddenly Distant & Still in Flux: Teachers' Work During COVID-19. Suddenly Distant Research Project. Retrieved January 18, 2022, from. https://drive.google.com/file/d/1o_BKC9Majv135b39V3nQ68-wwVAjPKha/view.
- Birkland, T., 1998. Focusing events, mobilization, and agenda setting. *J. Publ. Pol.* 18 (1), 53–74.
- Bateman, N., Ross, M., 2021. Why has covid-19 been especially harmful for working women? Brookings. Retrieved May 24, 2022, from. <https://www.brookings.edu/essay/why-has-covid-19-been-especially-harmful-for-working-women/>.
- Bubb, S., Jones, M.A., 2020. Learning from the COVID-19 home-schooling experience: listening to pupils, parents/carers and teachers. *Improv. Sch.* 23 (3), 209–222. <https://doi.org/10.1177/1365480220958797>.
- Chakraborti, S., 2013. Deskillings of the teaching profession. In: *Sociology of Education: An A–Z Guide*. SAGE Publications, United States.
- Coburn, C.E., 2001. Collective sensemaking about reading: how teachers mediate reading policy in their professional communities. *Educ. Eval. Pol. Anal.* 23 (2), 145–170.
- Conklin, M., Kraft, M.A., 2022. Navigating the New Normal: Teachers' Experiences & Perceptions During the Fall Semester of 2021 (rep.). Retrieved January 18, 2022, from. <https://share.hsforms.com/1LoKG7ei2R82rcYNto3bn0A1qgk0>.
- DeAngelis, K.J., Presley, J.B., 2011. Toward a more nuanced understanding of new teacher attrition. *Educ. Urban Soc.* 43 (5), 598–626.
- Diliberti, M., Kaufman, J.H., 2020. Will This School Year Be Another Casualty of the Pandemic? Key Findings From the American Educator Panels Fall 2020 COVID-19 Surveys. Data Note: Insights From the American Educator Panels. Research Report. RR-A168-4. RAND Corporation.
- Dworkin, A.G., 2009. Teacher burnout and teacher resilience: assessing the impacts of the school accountability movement. In: Saha, L.J., Dworkin, A.G. (Eds.), *New International Handbook of Teachers and Teaching*. Springer Publications, NY, pp. 491–509.
- Dworkin, A.G., Tobe, P.F., 2014. The effects of standards based school accountability on teacher burnout and trust relationships: a longitudinal analysis. In: Van Maele, D., Forsyth, P.B., Van Houtte, M. (Eds.), *Trust and School Life: The Role of Trust for Learning, Teaching, Leading, and Bridging*. Springer Science + Business Media, pp. 121–143.
- Easthope, C., Easthope, G., 2000. Intensification, extension and complexity of teachers' workload. *Br. J. Sociol. Educ.* 21 (1), 43–58.
- Fitzgerald, S., McGrath-Champ, S., Stacey, M., Wilson, R., Gavin, M., 2019. Intensification of teachers' work under devolution: a "tsunami" of paperwork. *J. Ind. Relat.* 61 (5), 613–636.
- Fontenelle-Tereshchuk, D., 2021. Homeschooling and the COVID-19 Crisis: The Insights of Parents on Curriculum and Remote Learning. *Interchange* 52 (2), 167–191.
- García, E., Weiss, E., 2022. The teacher shortage is real, large and growing, and worse than we thought: The first report in 'The perfect storm in the teacher labor market' series. Economic Policy Institute. Retrieved January 18, 2022, from. <https://www.epi.org/publication/the-teacher-shortage-is-real-large-and-growing-and-worse-than-we-thought-the-first-report-in-the-perfect-storm-in-the-teacher-labor-market-series/>.
- Gewirtz, S., 1997. Post-welfarism and the reconstruction of teachers' work in the UK. *J. Educ. Policy* 12 (4), 217–231.
- Gicheva, D., 2021. Teachers' working hours during the COVID-19 pandemic. *Educ. Res.* <https://doi.org/10.3102/0013189X211056897>.
- Goldhaber, D., Theobald, R., 2022. Teacher attrition and mobility over time. *Educ. Res.* 51 (3), 235–237. <https://doi.org/10.3102/0013189X211060840>.
- Grossmann, M., Reckhow, S., Strunk, K.O., Turner, M., 2021. All states close but red districts reopen: the politics of in-person schooling during the COVID-19 pandemic. *Educ. Res.* 50 (9), 637–648.
- Gu, Q., 2014. The role of relational resilience in teachers' career- long commitment and effectiveness. *Teach. Teach. Theor. Pract.* 20 (5), 502–529. <https://doi.org/10.1080/13540602.2014.937961>.
- Gudmundsdottir, G.B., Hathaway, D.M., 2020. "We always make it work": teachers' agency in the time of crisis. *J. Technol. Teach. Educ.* 28 (2), 239–250.
- Hamilton, L.S., Grant, D., Kaufman, J.H., Diliberti, M., Schwartz, H.L., Hunter, G.P., Young, C.J., 2020. COVID-19 and the State of K-12 Schools: Results and Technical Documentation From the Spring 2020 American Educator Panels COVID-19 Surveys. RAND Corporation.

- Hamilton, L.S., Kaufman, J.H., Diliberti, M., 2020. Teaching and Leading Through a Pandemic: Key Findings From the American Educator Panels Spring 2020 COVID-19 Surveys. The RAND Corporation. www.rand.org/RRAs/RR168-2.
- Hargreaves, A., 1989. American Educational Research Association Conference Papers. Teachers' Work and the Politics of Time and Space, vol. 150.
- Hargreaves, A., 1994. Changing Teachers, Changing Times: Teachers' Work and Culture in the Postmodern Age. Teachers College Press.
- Hargreaves, A., 2001. Emotional geographies of teaching. *Teach. Coll. Rec.* 103 (6), 1056–1080.
- Hargreaves, A., 2021. What the COVID-19 pandemic has taught us about teachers and teaching. *FACETS* 6, 1835–1863. <https://doi.org/10.1139/facets-2021-0084>.
- Hartney, M.T., Finger, L.K., 2020. Politics, markets, and pandemics: public education's response to COVID-19. *Perspect. Polit.* 1–17.
- Herold, B., 2021. The decline of hybrid learning for this school year, in 4 charts. *Educ. Week*. Retrieved from: <https://www.edweek.org/technology/the-decline-of-hybrid-learning-for-this-school-year-in-four-charts/2021/09>.
- Herold, B., Kurtz, H., 2020. Teachers work two hours less per day during COVID-19: 8 key EdWeek survey findings. *EdWeek*. <https://www.edweek.org/teaching-learning/teachers-work-two-hours-less-per-day-during-covid-19-8-key-edweek-survey-findings/2020/05>. (Accessed 12 January 2022).
- Horn, I.S., McGugan, K.S., 2020. Adaptive Expertise in Mathematics Teaching During a Crisis: How Highly-Committed Secondary U.S. Mathematics Teachers Adjust Their Instruction in the COVID-19 Pandemic.
- Igielnik, R., 2021. A rising share of working parents in the U.S. say it's been difficult to handle child care during the pandemic. Pew Research Center. Retrieved January 10, 2022. <https://www.pewresearch.org/fact-tank/2021/01/26/a-rising-share-of-working-parents-in-the-u-s-say-its-been-difficult-to-handle-child-care-during-the-pandemic/>.
- Ingersoll, R.M., Merrill, E., Stuckey, D., Collins, G., 2018. Seven Trends: The Transformation of the Teaching Force—Updated October 2018. University of Pennsylvania, p. 28.
- Ingersoll, R.M., 2001. Teacher turnover and teacher shortages: an organizational analysis. *Am. Educ. Res. J.* 38 (3), 499–534.
- Ingersoll, R.M., 2009. Who Controls Teachers' Work?: Power and Accountability in America's Schools. Harvard University Press.
- Jabbar, H., Jellison Holme, J., De La Garza, S., Trautmann, K., 2021. Repair the Fractures: Addressing the Impact of COVID-19 on Schools. The University of Texas at Austin College of Education.
- Johnson, S.M., Kraft, M.A., Papay, J.P., 2012. How context matters in high-need schools: the effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teach. Coll. Rec.* 114 (10), 1–39.
- Johnson, S.M., 2020. How COVID-19 Laid Bare the Strengths and Weaknesses of Schools. Harvard Education Publishing Group. Retrieved January 28, 2022, from: <https://www.hepg.org/blog/how-covid-19-laid-bare-the-strengths-and-weaknesses>.
- Jones, N.D., Camburn, E.M., Kelcey, B., Quintaro, E., 2022. Teachers' time use and affect before and after COVID-19 school closures. *AERA Open* 8 (1), 1–14.
- Jones, A.L., Kessler, M.A., 2020. Teachers' emotion and identity work during a pandemic. *Front. Educ.* 5, 195. *Frontiers*.
- Kim, L.E., Leary, R., Asbury, K., 2021. Teachers' narratives during COVID-19 partial school reopenings: an exploratory study. *Educ. Res.* 63 (2), 244–260. <https://doi.org/10.1080/00131881.2021.1918014>.
- Kaufman, J., Diliberti, M., 2021. Teachers are Not All Right: How the COVID-19 Pandemic is Taking a Toll on the Nation's Teachers. Center on Reinventing Public Education, University of Washington, Seattle, Wash.
- Kelchtermans, G., 2017. "Should I stay or should I go?": unpacking teacher attrition/retention as an educational issue. *Teach. Teach.* 23 (8), 961–977.
- Klees, S.J., 2017. Beyond neoliberalism: reflections on capitalism and education. *Policy Futures Educ.* 18 (1), 9–29. <https://doi.org/10.1177/1478210317715814>.
- Kraft, M.A., Simon, N.S., 2020. Teachers' Experiences Working From Home During the COVID-19 Pandemic. Upbeat.
- Kraft, M.A., Simon, N.S., Lyon, M.A., 2020. Sustaining a Sense of Success: The Importance of Teacher Working Conditions During the COVID-19 Pandemic. Annenberg EdWorkingPaper No. 20-279. Annenberg Institute, Brown University.
- Kraft, M.A., Simon, N.S., Donaldson, M.I., 2021. Teacher's Perceptions of Their Working Conditions During the Pandemic: Results From the Fall 2020 Teacher Engagement Survey. Upbeat.
- Kurtz, H., 2020. In-Person Learning Expands, Student Absences Up, Teachers Work Longer, Survey Shows. *Education Week*. <https://www.edweek.org/leadership/in-person-learning-expands-student-absences-up-teachers-work-longer-survey-shows/2020/10>.
- Larson, S.M., 1980. Proletarianization and educated labour. *Theory Soc.* 9 (1), 131–175.
- Le Cornu, R., 2013. Building early career teacher resilience: the role of relationships. *Aust. J. Teach. Educ.* 38 (4). <https://doi.org/10.14221/ajte.2013v38n4.4>.
- Le Nestor, A., Moscoviz, L., 2020. Six Things You Should Know About Female Teachers. Center for Global Development. Retrieved January 30, 2022, from: <https://www.cgdev.org/blog/six-things-you-should-know-about-female-teachers>.
- Long, C., 2020. Many Educators Buckling Under Pandemic Workload. National Education Association. Retrieved January 25, 2022, from: <https://www.nea.org/advocating-for-change/new-from-nea/many-educators-buckling-under-pandemic-workload>.
- Lortie, D., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago.
- Marchant, E., Todd, C., James, M., Crick, T., Dwyer, R., Brophy, S., 2021. Primary school staff perspectives of school closures due to COVID-19, experiences of schools reopening and recommendations for the future: a qualitative survey in Wales. *PLoS One* 16 (12). <https://doi.org/10.1371/journal.pone.0260396>.
- Mavhungu, C., 2021. Thousands of Zimbabwean Teachers Strike Over COVID-19 Concerns. VOA News. Retrieved from: https://www.voanews.com/a/africa_thousands-zimbabwean-teachers-strike-over-covid-19-concerns/6210232.html.
- Meheut, C., 2022. Teachers in France stage walkout over covid rules in schools. *N. Y. Times*. Retrieved from: <https://www.nytimes.com/2022/01/13/world/europe/france-teachers-walkout-covid-coronavirus.html>. (Accessed 2 February 2022).
- Merrill, B.C., 2021. Configuring a construct definition of teacher working conditions in the United States: a systematic narrative review of researcher concepts. *Rev. Educ. Res.* 91 (2), 163–203. <https://doi.org/10.3102/0034654320985611>.
- Niemi, H.M., Kousa, P., 2020. A case study of students' and teachers' perceptions in a Finnish high school during the COVID pandemic. *IJTES* 4 (4), 352–369.
- Nieto, S., 2003. Challenging current notions of "highly qualified teachers" through work in a teachers' inquiry group. *J. Teach. Educ.* 54 (5), 386–398.
- Noddings, N., 2013. *Caring: A Relational Approach to Ethics & Moral Education*, second ed. University of California, Berkeley, CA. <https://doi.org/10.1525/9780520957343>.
- OECD, 2021. The State of Global Education 18 Months Into the Pandemic. OECD.
- OECD, 2022. Education at a Glance: Distribution of Teachers by Age and Gender. OECD Education Statistics (database). <https://doi.org/10.1787/4ef9a105-en>. (Accessed 10 February 2022).
- Pressley, T., 2021. Factors contributing to teacher burnout during COVID-19. *Educ. Res.* <https://doi.org/10.3102/0013189X211004138>.
- Reich, J., Buttner, C.J., Coleman, D., Colwell, R.D., Faruqi, F., Larke, L.R., 2020. What's Lost, What's Left, What's Next: Lessons Learned From the Lived Experiences of Teachers During the 2020 Novel Coronavirus Pandemic.
- Ryan, S., von der Embse, N., Pendergast, L., Saeki, E., Segool, N., Schwing, S., 2017. Leaving the teaching profession: the role of teacher stress and educational accountability policies on turnover intent. *Teach. Teach. Educ.* 66, 1–11.
- Sahlberg, P., 2016. The Global Education Reform Movement and its Impact on Schooling. In: Wiley-Blackwell, pp. 128–144.
- Santibanez, L., Guarino, C., 2020. The Effects of Absenteeism on Cognitive and Social-Emotional Outcomes: Lessons for COVID-19. Annenberg Institute at Brown University. <https://doi.org/10.26300/yj9m-x430>. EdWorkingPaper:20-261.
- Shah, N., 2021. "I Feel Half as Successful": Teachers Push to Ban Hybrid Instruction, But Districts Want to Keep It. Politico, 29 December. Available at: <https://www.politico.com/news/agenda/2021/12/29/teachers-districts-hybrid-education-526214>. (Accessed 15 January 2022).
- Steiner, E.D., Woo, A., 2021. Job-Related Stress Threatens the Teacher Supply: Key Finding From the 2021 State of the US Teacher Survey. RAND Corporation. Technical Appendixes. Research Report. RR-A1108-1.

- Tackie, H.N., 2022. (Dis)Connected: establishing social presence and intimacy in teacher–student relationships during emergency remote learning. *AERA Open*. <https://doi.org/10.1177/23328584211069525>.
- Taie, S., Goldring, R., 2020. Characteristics of Public and Private Elementary and Secondary School Teachers in the United States: Results From the 2017–18 National Teacher and Principal Survey First Look (NCES 2020- 142). U.S. Department of Education. National Center for Education Statistics, Washington, DC.
- Tedeschi, E., 2020. COVID-19, school closures, and US parental labor outcomes. Working Paper. https://rebrand.ly/evt_parents.
- Thompson, G., 2021. The Global Report on the Status of Teachers 2021. Education International Research. <https://drive.google.com/file/d/18vEaj7Bu7SiOxFmMeJNqvQJvBwatwt/view?usp=sharing>.
- United Nations, 2020. Policy Brief: Education During COVID-19 and Beyond. Retrieved from. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf.
- United Nations, 2021. Policy Brief: Teachers are Driving Force Behind “Global Education Recovery” From COVID-19. Retrieved from. <https://news.un.org/en/story/2021/10/1102132>.
- Walker, T., 2022. Survey: Alarming number of educators may soon leave the profession. Retrieved February 7, 2022, from. <https://www.nea.org/advocating-for-change/new-from-nea/survey-alarming-number-educators-may-soon-leave-profession>.
- Watkins, P., 1994. The Fordist/post-Fordist debate: the educationalist implications. In: Kenway, J. (Ed.), *Economising Education: The Post-fordist Directions*. Geelong. Deakin University Press.
- Will, M., 2020. Across the country, teachers rally for safe reopenings and police-free schools. *Educ. Week*, 3 August. Available at. <https://www.edweek.org/teaching-learning/across-the-country-teachers-rally-for-safe-reopenings-and-police-free-schools/2020/08>. (Accessed 1 February 2022).
- Wilson, R., Stacey, M., McGrath-Champ, S., 2020. Teachers' Work During the COVID-19 Pandemic: Shifts, Challenges, and Opportunities. Centre for Strategic Education. Occasional Paper #169: Centre for Strategic Education.
- Winkler, A.M., Scull, J., Zeelandelaar, D., 2012. How Strong are US Teacher Unions? A State-By-State Comparison. Thomas B. Fordham Institute.
- Woltran, F., Chan, R., Lindner, K.-T., Schwab, S., 2021. Austrian elementary school teachers' perception of professional challenges during emergency distance teaching due to COVID-19. *Front. Educ.* 6, 759541. <https://doi.org/10.3389/educ.2021.759541>.
- Wong, J.L., 2006. Control and professional development: are teachers being deskilled or reskilled within the context of decentralization? *Educ. Stud.* 32 (1), 17–37.

Reclaiming the artistry of teaching

Gert Biesta^{a,b}, ^aThe Moray House School of Education and Sport, University of Edinburgh, Edinburgh, United Kingdom; and
^bCenter for Public Education and Pedagogy, National University of Ireland at Maynooth, Co. Kildare, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	648
The point and purpose of teaching	649
The double art of teaching: craft and wisdom	650
Lawrence Stenhouse on the artistry of teaching	652
Concluding comments: reclaiming the artistry of teaching	653
References	654

Psychology is a science, and teaching is an art; and sciences never generate arts directly out of themselves. An intermediary inventive mind must make the application, by using its originality.

James (1899, p. 14).

Introduction

In a chapter from 1984, republished as a journal article in 1988, the British curriculum scholar Lawrence Stenhouse wrote a sentence which, in hindsight, can only be characterized as visionary. It reads as follows: “Improving education is not about improving teaching as a delivery system [because] crucial is the desire of the artist to improve his or her art” (Stenhouse, 1988, p. 50). The sentence is visionary, because if there is one thing that has characterized educational policy, research, and practice over the past decades, it is precisely the ongoing effort to steer the whole educational enterprise toward becoming an effective delivery system of measurable learning outcomes (see, e.g. Ball, 2012). In this set up the teacher is identified as an important *factor*—there are numerous documents in which it is claimed that research has shown that the teacher is the most important “in school factor” in the production of measurable learning outcomes (see, e.g. King-Rice, 2003; OECD, 2005; McKinsey and Co., 2007)—but not as a thinking, judging and acting professional who is driven by an ongoing desire to improve his or her own “artistry” (I will return to this term below).

The key word in these developments is the notion of “effectiveness,” and the “double move” that has taken place in many countries over the past decades has been to make the operation of the education system more effective and to do so with regard to the production of an increasingly smaller number of measurable “learning outcomes” (see Biesta, 2010; Sellar et al., 2017; Hogan and Thompson, 2021). This line of thought has brought about a world-wide educational evidence “industry” that seeks to find out, through the conduct of large-scale experimental studies, what allegedly “works” in education. Moreover, the ambition of such research is to tell teachers what they should do in order to impact effectively on the production of particular “outcomes.” Some claim that the “extraordinary advances in medicine, agriculture and other fields” are the result of “the acceptance by practitioners of evidence as the basis for practice,” particularly evidence from the randomized controlled trial, and that it is about time that education embarks upon a similar path (see Slavin, 2002, p. 16). Others, however, have raised fundamental questions about this approach, asking both whether this is a *desirable* way forward for education, and whether it is even a *possible* way forward (see Thomas, 2021; see also; Biesta, 2015; D’Agnese, 2017; Davis, 2017; for an illuminating conversation see Hattie and Nepper Larsen, 2020; and for a critique from within the medical field see Holmes et al., 2006).

The question about the *possibility* of this vision has to do with how we envisage the dynamics of education. In the view of education as a delivery system, teaching is seen as an “intervention” which, under “ideal” circumstances, will bring about particular “outcomes.” We can characterize this as a causal conception of education in which teaching is seen as a cause and learning—or in the lingo of this approach: “learning outcomes”—as the effect of teaching interventions. In this set up the teacher does indeed appear as a factor in a production-line, and one could say that the only possible role for the teacher here is that of a *technician*, that is, the one who administers prescribed interventions. The question this line of thought raises is whether this is an adequate and accurate understanding of the dynamics of education and in what is to follow I will argue that this is *not* the case. Yet this also has to do with the question of the *desirability* of this vision. With regard to education itself, the question is whether we should indeed see education as a matter of control, that is, a matter of administering interventions which, under “ideal” circumstances, will bring about predetermined outcomes. This, in my view, is a problematic and impoverished conception of what education is and should be about. In addition to this, there is also the question of the desirability of the “figure” of the teacher that is entailed in this vision, that is, of the idea of the teacher as “factor” or technician who both is controlling his or her students and is being controlled by the prescriptions based on the alleged evidence about what does and what doesn’t “work.” Again, as I will argue, such a conception of the teacher amounts to a misrepresentation of the complex work of teaching and therefore doesn’t speak to what this work is about and what is at stake in it.

In several of my publications I have raised questions about the adequacy of a mechanistic or “quasi-causal” conception of education in which teaching is seen as an intervention, the teacher as a factor, and the focus is on increasing the effectiveness of teaching interventions for the production of learning outcomes. Instead, I have made a case for complexity, that is, for seeing education as an open, semiotic, and recursive system, that is, a system “populated” by thoughtful individuals who can make up their own mind and can act accordingly, and whose interaction is a matter of communication and interpretation, not of mechanical “push and pull” (see, for example, [Biesta, 2016](#)).

In this contribution I wish to focus on one particular aspect of this discussion, namely the suggestion that we should understand teaching and the work of the teacher as an *art*, first and foremost in the Aristotelian sense in which the arts are about human activity in the domain of the “variable.” Work in this domain requires “artistry”—the ability to make situated judgments about what are desirable ways of acting, which always entails judgment about the “how” and about the “what for.” In this regard, as I will suggest below, teaching can be characterized as a “double art.” In the three main sections of this chapter, I will discuss the question what teaching is, which, inevitably, has to engage with the question what teaching is *for*. I will then turn to Aristotle in order to look in more detail at the complexities of work in the domain of the variable. In a third step I will discuss ideas from Lawrence Stenhouse, being one of the authors who has made an explicit case for teaching as an art and for the artistry of teaching. I do this in order to show that a claim for the artistry of teaching does not have to start from scratch but can build upon ideas that may have been forgotten and definitely have become marginal in the educational conversation. Another author who has made significant contributions to the “case” for seeing teaching as an art and for focusing on the artistry of teaching is Eliot Eisner (see, e.g. [Eisner, 2002](#)). One reason for focusing on the work of Stenhouse rather than Eisner is that Eisner’s contribution stays more closely to the Aristotelian line of argument presented in this chapter, whereas Stenhouse, in a sense, locates the artistry of teaching closer to the (performing) arts. In the concluding section I will say a few words about the intellectual and political work needed for reclaiming a meaningful account of teaching.

The point and purpose of teaching

I wish to take my starting point in the observation that teaching is a relational, intentional, and purposeful activity (for a detailed discussion see [Biesta and Stengel, 2016](#)). Firstly, teaching is not something one can do on one’s own but implies a relationship between teachers and (their) students, which means that it takes at least two for teaching to take place. Secondly, teaching doesn’t happen by accident; those who teach, or try to teach, do so deliberately. And thirdly, teaching is done with a sense of purpose, which means that it is done for a reason. Teachers have, in other words, good reasons for what they teach and how they teach, and these reasons have something to do with what they hope that their activities will bring about on the side of their students, albeit that it remains open whether or not this will happen as the “success” of teaching is not just reliant on what teachers do but also one what students do with what they are being taught.

Over the past decades, the prevalent answer to the question what teaching is *for*, has become “learning.” This, however, is a rather impoverished idea of what teaching should aim for and actually also a rather unhelpful suggestion. This is not just because learning can happen anywhere, also without teaching. It is also because learning can go in any direction, and therefore just to say that teaching should bring about learning or that the work of teachers is to “facilitate” learning, says very little, if it says anything at all. The point of education rather is that students learning *something*, that they learn it *for a reason*, and that they learn it *from someone*. Education, in other words, is never about learning in general but always raises questions of content, purpose and relationships. And it is here that teaching comes in, because whereas students can learn all kind of things from being in educational settings—including how to cheat or how to pass an exam with minimal effort—the work of the teacher is to direct the attention of students to particular subject-matter, particular themes, particular issues, and particular tasks and, most importantly, to do so with particular reasons and purposes in mind. Teaching, in other words, is always conducted with an orientation.

Elsewhere (see for example [Biesta, 2009, 2010, 2020](#)) I have suggested that there are actually three purposes (or as I prefer: domains of purpose) that are always at stake when education takes place. One important reason why we engage in education is to make knowledge and skills available to our students. We can refer to this as knowledge acquisition, but it is perhaps better to say that one important purpose of education is *qualification*, that is, providing students with knowledge, skills, attitudes, and dispositions that makes it possible for them to do something. This “doing” can either be quite specific and precise, such as becoming qualified for a particular job or profession; but it can also be understood more broadly, such as the way in which schools seek to equip children and young people for their life in complex modern societies.

Some argue that qualification is the sole purpose of education because they are worried that anything else gets education into difficult normative questions, and these are better left to the family or community context. This may sound reasonable, but the problem is that since education is unable to provide children and young people with all existing knowledge and skills, there is inevitable selection going on in education. In everything we do we thus present our students with a particular “selection” of the world or, more positively formulated, with a particular *representation* of the world. The ways in which we do this inevitably influence our students in some way. Normative questions are therefore inescapable, even if education were confined to the domain of knowledge and skills.

In the literature the (re)presentation of the world, or the presentation of different representations of the world, is known as *socialization*. Some highlight the ways in which this goes on behind the backs of our students—an idea known as the hidden curriculum. Yet we can also think more positively about this and see socialization as an important second purpose of education, where we try to

provide our students with an orientation in the world, that is, with a sense of direction. Education as socialization is about providing our students with an orientation into existing cultures, traditions, and practices, with the invitation that they locate themselves within them.

Discussions about socialization, particularly strong(er) and (more) directive approaches, raise an important further question, which has been part of the modern educational conversation at least since the Enlightenment. The question here is whether education can and should approach students as “objects” that need to be (come) qualified and socialized—which seems to be the dominant “undertone” of the idea of teaching as a delivery system—or whether education always needs to ensure that children and young people can become subjects of their own life. This is partly a very complicated and deeply philosophical question, but it is also a very simple issue which many educators will immediately recognize. After all, in all education we want to make sure that students stop relying on our help and input and become able to think for themselves, make their own judgments, and be able to act and act *well*.

There are different ways in which we can refer to this third domain of educational purpose. I tend to refer to it with the word “*subjectification*,” which is perhaps a rather odd word in English, but precisely refers to the ambition that students end up as subjects of their own life and not just as objects of educational “interventions.” It therefore stands in sharp contrast to education that aims for objectification, that is, education which is only interested in controlling students and their acting, thinking and judgment. Of course, we cannot force our students to be subjects of their own life, but we can, in all kinds of ways, “remind” our students of this possibility to exist as subject of their own life. And we can provide them with opportunities to encounter and experiment with the complexities of what this means (see [Biesta, 2020](#) for more detail).

Dietrich Benner has suggested the phrase “*Aufforderung zur Selbsttätigkeit*” as a way to capture the special character of educational work in this domain (see [Benner, 2015](#)). This can be translated as “summoning to self-action,” although the “summoning” may sound a bit strict and a word like “encouragement” might be more helpful. This should not be understood as the encouragement to be yourself, and also not the encouragement just to become active. It is perhaps best to see this as the injunction to be a self, that is, to try to be a subject of your own life, with all the complexities and responsibilities that follow from it, rather than remain an object of influences from “elsewhere.”

Another set of concepts that is helpful here is Benner’s distinction between affirmative and non-affirmative education (see [Benner, 2015](#), pp. 146–155). Whereas qualification and socialization are, to a large degree, affirmative, in that they start from certain ideas about what education should achieve and where children and young people should end up, the domain of subjectification is precisely the opposite of this, because here it is not for educators to tell children and young people how they should be and become, but rather to provide opportunities for them to figure out for themselves how to live their own lives in the best way possible. That is why the educational work vis-à-vis this domain has to be non-affirmative and has to proceed with caution.

I wish to suggest that qualification, socialization, and subjectification are not only three *legitimate* purposes of education; in a sense they are also three *inevitable* purposes of all education. After all, in all instances of education there is always something for teachers to offer to students and for students to acquire to their benefit. Because qualification always represents (aspects of) the world in a particular way, there is always also socialization going on. And all this also has an impact on the student as subject, to begin with because becoming more knowledgeable or skilled (qualification) and gaining orientation in a particular domain or field (socialization) provides students with increased possibilities for thinking, judgment, and action.

The fact that these three domains of purpose are inevitable, first of all suggests that the three domains are always entangled with each other; they cannot exist separately, because every act of qualification is also an act of socialization and also impacts on the student’s subject-ness, positively or negatively. It suggests, secondly that in the design and enactment of education teachers should always consider what they seek to achieve—or what they seek their students to achieve—in relation to each of these domains. Thirdly, although the three domains are always “in play” in education, it doesn’t mean that they can exist in perfect harmony. There are always potential tensions between, say, what one seeks to achieve in the domain of qualification and what is possible in the other domains. There can be synergy—understanding subject-matter well also provides a degree of orientation and contributes to one’s agency—but there can also be conflicts—for example when a too strong push on the domain of qualification undermines students’ agency.

The challenge for teachers, therefore, is not just to think and act in a three-dimensional way, that is, with an eye on the three domains of educational purpose. The challenge is also to try to secure a meaningful balance and think carefully about the costs of emphasizing one domain to the detriment of the other domains. This, as I will argue in the next section, is one important reason why teaching needs to be understood as an art and why teachers need artistry rather than techniques or prescriptions.

The double art of teaching: craft and wisdom

One key message emerging from the discussion so far, is that teaching cannot and should not be enacted as a form of control or, to be more precise, as a kind of intervention that, under “ideal” circumstances and based upon the best “evidence” about what “works,” should be aimed at producing pre-specified learning outcomes. This is not to suggest that everything should be open, which is the mistake of those who denounce teaching in favor of learning. But it is to challenge the view that education is ultimately a causal system (an ontological claim) and that, once we have perfect knowledge about the mechanics of the system (an epistemological claim), teaching can become a matter of administering those interventions that produce the desired outcomes

(a praxeological claim). The whole point of teaching is precisely *not* the production of outcomes but a matter of educating human beings so that they become more qualified, that is, able to act thoughtfully, gain a sense of direction in the world, and, through this, take upon themselves the challenge of being subjects of their own lives.

If this is not a matter of implementing evidence-based intervention, how then should we understand the work of teachers, and what do they need to act thoughtfully in educational settings? These questions relate back to a rather old discussion which has been framed, correctly in my view, as the question whether teaching should be understood as a “science” or as an “art.” This question surfaces regularly in the educational conversation, including in the quote with which I opened this article. In my view, the most helpful engagement with this question can be found in the work of Aristotle, for whom the difference between “science” and “art” is not a matter of different kinds of knowledge but actually begins with a proper understanding of the reality we are acting “in” and “upon” in such practical endeavors as education.

With regard to this Aristotle makes a very helpful and, in my view, tremendously important distinction between the theoretical life (the “*bios theoretikos*”) and the practical life (the “*bios praktikos*”). The theoretical life is concerned with “the necessary and the eternal” (Aristotle, 1980, p. 140), that is, with those aspects of reality that do not change (Aristotle mentions such phenomena as the movement of the planets or the stars in the sky). He refers to knowledge about this reality as “*episteme*,” which is often translated as “science.” “*Episteme*” is knowledge about what is necessary and eternal and because the reality that this knowledge is about doesn’t change, the knowledge we have about this reality, once it is accurate, will not change either. From this we have an idea of true knowledge as being 100% stable, secure, and certain.

Aristotle’s main insight, however, is that most of what our lives are about takes place in the domain of the “variable” (for this term see Aristotle, 1980, p. 42), that is, the domain of change. It is the world in which we act and where our actions have consequences, but where there is no guarantee that our actions will always have the same consequences. It is, in other words, the domain of possibility, not the domain of certainty. Knowledge in this domain is about the relationships between our actions and the *possible* consequences of our actions. This is so for our actions in and interactions with the material world, our actions in and interactions with the living world, and our actions in and interactions with the social world. While we can gain much knowledge about possible relationships between actions and consequences in these domains, there is never a guarantee that when we act in the same way in the future, the same consequences will follow. And precisely there lies the difference between “science” and “art,” where art refers to our actions in the domain of the variable.

What is also interesting about Aristotle’s explorations is that he makes a further distinction between two “modes” of acting in the domain of the variable and two related forms of judgment. This is the distinction between *poiesis*, which Carr (1987) has helpfully translated as “making action,” and *praxis*, which Carr translates as “doing action.” *Poiesis* is the art of making things—such as, for example, a saddle or a ship—although I prefer to think of it slightly more widely, that is, as action that brings something into existence. We might also call it “productive action.” As Aristotle puts it, *poiesis* is about “how something may come into being which is capable of either being or not being” (which means that it is about the variable, not about what is eternal and necessary), and about things “whose origin is in the maker and not in the thing made” (which distinguishes *poiesis* from biological phenomena such as growth and development) (see Aristotle, 1980, p. 141). *Poiesis* is, in short, about bringing something in existence that did not exist before.

The kind of knowledge we need for *poiesis* is what Aristotle refers to as *techne*, which he defines as “knowledge of how to make things” (Aristotle, 1980, p. 141). *Techne* is about finding the means that will bring about what one seeks to bring about, to put it in general terms. *Techne* encompasses knowledge about the materials we work with—and we can take “materials” in the broad sense of the word—and about the techniques we can apply to work with those materials. Yet making something, such as a saddle, is never simply about following a recipe. It involves making judgments about the application of our previous knowledge and experience to *this* piece of leather, for a saddle to fit *this* particular horse, and for *this* particular person riding the horse. So we make judgments about application, production, and effectiveness in our attempts to bring something into existence. The best English word for *techne* is probably craftsmanship although in a slightly narrower translation we can also think of it as consisting of the combination of practical knowledge and practical judgment.

The domain of the variable is, however, not confined to the world of things and matters of making, but includes the social domain, that is, the world of human action and interaction. It is here that a second art is called for: the art of *praxis*. The orientation here is not toward the making of things but toward the promotion of the human good (the Greek term is *eudamonia*). *Praxis*, Aristotle writes, is “about what sort of things conduce to the good life in general” (Aristotle, 1980, p. 142). What we need in order to proceed here is not judgment about how to do things, but rather judgment “about what is to be done” (Aristotle, 1980). Aristotle refers to this kind of judgment as *phronesis*, which is usually translated as practical wisdom. Aristotle defines it as a “reasoned and true state of capacity to act with regard to human goods” (Aristotle, 1980, p. 143).

Aristotle thus provides a powerful argument for the idea that teaching is an art and not a science and also provides us with precise definitions of “art” and “science.” The key insight here is that teaching takes place in the domain of the variable, that is, the domain of actions and possible consequences, and the reason for this, to put it bluntly, is that in teaching we work with “living material,” that is, with human beings who are capable of their own thought and action. What is also interesting about Aristotle’s approach, is the distinction he makes between two different arts: the art of making, for which we need *techne*, which is the practical knowledge and practical judgment about how to bring something into existence, and the art of doing, for which we need *phronesis*, the practical wisdom we need to judge what is to be done. In this regard we could say that teaching is a “double art,” which requires both educational craftsmanship—the “*techne*” of teaching—and educational wisdom—the orientation of what teaching is *for*.

The final point to make here is that the “how” of teaching and the “what for” of teaching should not be seen as disconnected from each other. It is not that in education we can first set the goals and then just find the most effective and efficient way of getting there. The reason for this lies in the simple fact that the ways in which we proceed in education, the ways we teach, the ways we engage with our students, the ways we focus their attention, the ways we encourage them, are not just more or less effective interventions that happen behind the backs of our students. On the contrary, they are in full view of our students, and contain important messages for our students as well. This means that in addition to judgments about the purposes of our teaching, judgments about the way we try to balance the different domains of purpose, and judgments about possible trade offs in achieving a balance, we also need to judge the ways in which we teach. And this judgment is not just technical—which is the question of effectiveness—and also not just moral—“Are the ways in which we teach morally acceptable?”—but also needs to be *educational*, that is, to be judged in terms of the ways in which they may or may not contribute to the overall ambitions we have with our teaching.

If teaching is an art and, more specifically a “double art” of craft and wisdom, then it is important that teachers keep working on their own educational “artistry,” that is, their ability to make situated judgments about educationally desirable ways of acting in the always new situations they encounter. It is here that the whole question of the ongoing improvement of education finds its “home,” so to speak, because, to quote Lawrence Stenhouse once more, “improving education is not about improving teaching as a delivery system [because] crucial is the desire of the artist to improve his or her art” (Stenhouse, 1988, p. 50). And I wish to reiterate that to think of teaching as an art in the nuanced way in which I have discussed it in this section, provides a way to distinguish it from technicist accounts that seek to do away with the need for complex educational judgment and ongoing inventiveness in the educational situation.

Lawrence Stenhouse on the artistry of teaching

I have given this chapter the title “*reclaiming the artistry of teaching*” because I wish to highlight that the insight that teaching is an art (in the Aristotelian sense of the word) which requires artistry from its artisans, so to speak, is not a new insight. It partly goes back to the very origins of Western philosophy and, more specifically, to the work of Aristotle who, unlike Plato, can be seen as one of the first philosophers within the Western “canon” who had a keen interest in the practical life of change and possibility. The particular connection with the practice of teaching is also not new, and can, in a sense, be seen as an important “thread” in modern educational thought. Think, for example, of the idea of “tact” in the work of Herbart (for an excellent discussion see Van Manen, 1995; see also Van Manen, 1991), William James’s comments on the art of teaching quoted at the beginning of this chapter, or the work of Joseph Schwab on the idea of “the practical” (e.g. Schwab, 1991; see also Reid, 1999 on the “deliberative tradition”). We could even say that modern educational thought and practice are characterized by an ongoing tension between those who seek to push teaching toward the technicist end of the spectrum and those who make the case for teaching as thoughtful inventive situated action. One author whose work is clearly positioned within the latter approach and who not just has made the case for seeing teaching as an art but also has connected this explicitly to the idea of “artistry” is, as mentioned, Lawrence Stenhouse. In this section I wish to take a closer look at some of their ideas.

Lawrence Stenhouse locates the discussion about the artistry of teaching in the context of questions about “the relation of research to educational action” (Stenhouse, 1988, p. 43). Acknowledging that the contribution from disciplines such as history, philosophy, psychology, and sociology provide “a context in which to plan intelligent action” but that such disciplines “do not tell us how to act” (Stenhouse, 1988), he notes that “the yearning toward a form of research which might guide educational action [has] led educational researchers to look enviously at agricultural research” (Stenhouse, 1988). Such research is characterized by “field trials which [utilize] random sampling ... in order to recommend to farmers those strains of seed and crop treatments which would maximize yield” (Stenhouse, 1988). Stenhouse argues, however, that this approach—which nowadays is more commonly referred to in education as the “medical model”—is not appropriate for the field of education, both because of the idea of “random sampling” and because of the talk about “yield.” Not only does he make the interesting point that “it is the teacher’s job to work like a gardener rather than a farmer, differentiating the treatment of each subject and each learner as the gardener does each flower bed and each plant” (Stenhouse, 1988, p. 44). He also highlights that “the variability of educational situations is grossly underestimated” in the agricultural approach, which means that “sampling procedures cannot be related to educational action except on a survey basis rather than an experimental basis” (Stenhouse, 1988), and he points out that the intentionality of educational action is not a matter of having a goal that should be “produced,” but that “purpose in education is having an agenda” (Stenhouse, 1988).

Stenhouse also rejects the suggestion that for teachers to know what to do should be a matter of following instructions “laid down for us in the form of curricula and specifications of teaching methods” (Stenhouse, 1988). This is partly because of the totalitarian character of such a suggestion which, in his view, goes against the very point of education as “learning in the context of a search for truth” (Stenhouse, 1988). It is not only that “truth cannot be defined by the state even through democratic processes,” and if the state were to do this it would be as problematic as “the totalitarian control of art” (Stenhouse, 1988). Stenhouse also argues—and the point is entirely correct in my view—that teaching requires “situational professional judgment” because “prescriptions will vary according to cases” (Stenhouse, 1988). “We do not need doctors if all they are going to give us is a treatment laid down by the state or suggested by their professor without bothering to examine us and make a diagnosis” (Stenhouse, 1988).

For Stenhouse, then, the key argument against attempts to control teaching, either by state mandates or through particular forms of research that are supposedly able to tell teachers “what works,” lies in the simple statement that “good teaching is created by good teachers” (Stenhouse, 1988, p. 45). Stenhouse acknowledges that this statement may seem “self-evident to the point of absurdity” (Stenhouse, 1988), yet he also stresses that “the implications of this self-evident proposition do not seem to be widely grasped” (Stenhouse, 1988)—to which I wish to add that they are *still* not widely grasped, given the ongoing desire by many policy-makers, politicians, and others, including educational practitioners, to push the field of education and the work of teachers toward technicism. All this doesn’t mean that teachers should not welcome “access to ideas created by other people at other places or in other times,” but they do know—and so should those who generate such ideas—that all this is “not of much real use until [it is] digested to the point where [it is] subject to the teacher’s own judgment” (Stenhouse, 1988). And this is so because education cannot be done “behind the back” of teachers. As Stenhouse puts it: “only teachers are in the position to create good teaching” (Stenhouse, 1988).

It is at this point that Stenhouse brings forward the claim of teaching as art, and art as “an exercise of skill expressive of meaning” (Stenhouse, 1988). He explains: “The painter, the poet, the musician, the actor and the dancer all express meaning through skill.” Some are geniuses, but claiming teaching as art is not meant to elevate teachers inordinately, but rather to diagnose “the nature of their job in order to discern how performances might be improved” (Stenhouse, 1988, p. 46). What is distinctive about the art of teaching is that it “expresses in a form accessible to learners an understanding of the nature of that which is to be learned” (Stenhouse, 1988). Stenhouse makes this point in order to emphasize that teaching is not about the mere transmission of knowledge or skill, but always involves what he refers to as “the epistemology” of a particular discipline or field, that is, “the nature of its tenure on knowledge” (Stenhouse, 1988). “Teaching tennis is about understanding the logic and psychology and techniques of the game and about expressing that understanding through skill in teaching” (Stenhouse, 1988). Whether teaching is concerned with knowledge, the arts or with practical skills, makes no difference with regard to the requirement that “the teacher should aspire to give learners access to insight into the status of what they learn” (Stenhouse, 1988). This is as much a requirement for the teacher as it is for the student, as “good learning,” as Stenhouse puts it, is “not mere doing [but] is about constructing a view of the world” (Stenhouse, 1988, p. 47).

When, against this background, Stenhouse then (re)turns to the question of research, it is to highlight research as an inextricable dimension of art, not something that comes to it from the outside, tries to steer and direct it from the outside, and thus seeks to gain control over teaching from the outside. “All good art,” Stenhouse writes, “is an inquiry and an experiment” (Stenhouse, 1988, p. 47). Yet is it important to see that such inquiry and experiment is not the inquiry and experiment of the so-called “scientific method”—the agricultural approach which Stenhouse clearly rejects—but is the ongoing investigation of the complexities of one’s own practice and one’s own art. And the reason for this is not to increase control over education as a process of production but, as already mentioned before, to improve one’s own artistry. Stenhouse makes the interesting claim that “the purpose of the artist’s research is to improve the truth of his performance” (Stenhouse, 1988). Artists—Stenhouse mentions Leonardo, George Stubbs, Nureyev, Solti and the Chicago Symphony Orchestra, and Derek Jacobi—are all “engaged in inquiry, in research, in development of their own work” (Stenhouse, 1988). And while this will involve “improvement of technique, it is not for the sake of technique [but] for the sake of the expression of a truth in a performance” (Stenhouse, 1988). Teaching needs technique, so we might say, but the technique is never more than a means for enhancing one’s teacherly artistry. Improving teaching—improving one’s own teaching—is therefore a matter of “the development of understanding expressed in performance: understanding of the nature of knowledge expressed in the art form of teaching and learning” (Stenhouse, 1988, p. 50).

The final point Stenhouse makes, is for teachers to work together in the pursuit of enhancing their artistry. “I can exercise my art in secret, or even in a small group of consenting adults,” Stenhouse writes, but “if I want the support of a movement, I need to make alliances and develop some political power” (Stenhouse, 1988, p. 48). In this context Stenhouse mentions the “teacher-as-researcher” movement in Britain, suggesting that ideally universities and colleges of education should work together with teacher groups. What is required for this to happen, so Stenhouse argues (and we shouldn’t forget that he is writing in the beginning of the 1980s), “is that [universities] break the stranglehold of the ‘psycho-statistical and nomothetic paradigm’ of educational research” (Stenhouse, 1988). Stenhouse proposes case study as a much more appropriate “form” for enhancing one’s art, also because it “turns one toward one’s professional work rather than away from it”—a “gesture” he also recommends for teachers’ further study and development. His ideal—which, as he concedes, may be “so radical that I may not be communicating it” (Stenhouse, 1988, p. 50)—is for the emergence of communities of teachers “whose attention is primarily focused on the art of teaching,” just as “a company of actors does” (Stenhouse, 1988). This means that any inquiry that may take place should express itself “in performance of his or her art rather than (or as well as) in a research report” (Stenhouse, 1988, p. 51).

Concluding comments: reclaiming the artistry of teaching

In this chapter I have tried to make the case for seeing teaching as an art and for the importance of putting the artistry of teachers at the very center of our understanding of what teaching is about. I have tried to make this case in a clear and in a sense rather “basic” manner. I have done so by setting out the main characteristics of teaching, by reconstructing Aristotle’s insights about the domain of the “variable” and the roles of (the art of) *poiesis* and (the art of) *praxis* in this domain, and by exploring insights from Lawrence Stenhouse about the particular artistry of teaching. In one way I would say that the case is clear and straightforward and difficult to contest; but in another way I am acutely aware that many forces in contemporary educational theory, research, policy and practice

are strongly moving or have already strongly moved into the opposite direction, that is, constructing teaching as a “delivery system” and making sure that the delivery of the desired “outcomes” is done in the most effective and efficient way which, so it is claimed, requires that teaching becomes evidence-based.

While such arguments sound neat and attractive—also because they often come with the promise that they can “fix” education—they amount, in my view, to a misunderstanding but also a misrepresentation of the complexities of education and the work of teachers. Whereas Lawrence Stenhouse was still able to point a finger toward researchers based in universities as having a problematic conception of what “real” research was supposed to be, in our times I am worried that many teachers have actually adopted the agricultural logic of interventions and outcomes as the way to engage with and enhance their own practice. In this sense the ambition to reclaim the artistry of teaching is not just directed against particular strands of policy and research, but also entails a difficult but nonetheless important conversation with the teaching profession itself—and to speak in the singular here is perhaps no longer adequate.

I am also aware that my case for “reclaiming” refers back to older discussions and older scholarship and may, in this sense, be perceived as nostalgic move; a longing for a past that no longer exist and also no longer should exist because “the times have moved on.” In response I would like to suggest that “the times” may move on, but that there is always the need for us to judge whether “the times” are moving in a desirable or in a problematic direction. This point is not just political, but it is, in my view, also profoundly educational. After all, parents who just say “yes” to anything their children ask for, will most likely turn them into spoiled brats, rather than helping them to come into a relationship with their desires and acknowledge that not everything that “arrives” as a desire is automatically desirable. Education has, in other words, always also a “duty to resist,” as Philippe Meirieu has called it (see Meirieu, 2007). Rather than just keeping up with “the times,” there may be good reasons for teachers to see themselves as “untimely” professionals, as professionals who are “out of time” (on this idea see Kelchtermans, 2012), precisely so that they can keep a perspective on all the desires that policy makers, politicians, researchers, but also parents and students and society at large project onto education. In this regard, reclaiming the artistry of teaching may be the very act of resistance education may need in order to keep a sense of direction in turbulent times.

References

- Aristotle, 1980. *The Nicomachean Ethics*. Oxford University Press, Oxford.
- Ball, S.J., 2012. *Global Education Inc.* Routledge, London/New York.
- Benner, D., 2015. *Allgemeine Pädagogik. Eine systematisch-problemgeschichtliche Einführung in die Grundstruktur pädagogischen Denkens und Handelns*. 8. Auflage. Beltz/Juventa, Weinheim.
- Biesta, G., Stengel, B., 2016. Thinking philosophically about teaching. In: Gittomer, D.H., Bell, C.A. (Eds.), *Handbook of Research on Teaching*, fifth ed. AERA, Washington, DC, pp. 7–68.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21, 33–46.
- Biesta, G., 2010. *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Paradigm Publishers, Boulder, Co.
- Biesta, G.J.J., 2015. Resisting the seduction of the global education measurement industry: notes on the social psychology of PISA. *Ethics Educ.* 10, 348–360.
- Biesta, G.J.J., 2016. Improving education through research? From effectiveness, causality and technology, to purpose, complexity and culture. *Pol. Futures Educ. Internet* 14, 194–210.
- Biesta, G., 2020. Risking ourselves in education: qualification, socialization and subjectification revisited. *Educ. Theor.* 70, 89–104.
- Carr, W., 1987. What is an educational practice? *J. Philos. Educ.* 21, 163–175.
- D’Agnes, V., 2017. *Reclaiming Education in an Age of PISA. Challenging OECD’s Educational Order*. Routledge, London/New York.
- Davis, A., 2017. It worked there. Will it work here? Researching teaching methods. *Ethics Educ.* 12, 289–303.
- Eisner, E., 2002. From episteme to phronesis in the study and improvement of teaching. *Teach. Teach. Educ.* 18, 375–385.
- Hattie, J., Nepper Larsen, S., 2020. *The Purposes of Education: A Conversation Between John Hattie and Steen Nepper Larsen*. Routledge, London/New York.
- Hogan, A., Thompson, G. (Eds.), 2021. *Privatization and Commercialization in Public Education. How the Public Nature of Schooling is Changing*. Routledge, London/New York.
- Holmes, D., Murray, S.J., Perron, A., Rail, G., 2006. Deconstructing the evidence-based discourse in health science: truth, power and fascism. *Int. J. Evid. Base. Healthc.* 4, 160–186.
- James, W., 1899. *Talks to Teachers on Psychology: And to Students on Some of Life’s Ideals*. Henry Holt and Company, New York.
- Kelchtermans, G., 2012. *De leraar als (on)eigentijdse professional. Reflecties over de “moderne professionaliteit” van leerkrachten*. Onderwijsraad, Den Haag.
- King-Rice, J., 2003. *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. Economic Policy Institute, Washington, DC.
- McKinsey & Co, 2007. *McKinsey Report: How the World’s Best Performing School Systems Come Out on Top*. <https://www.bing.com/search?q=McKinsey+Report%3A+How+the+world%27s+best+performing+school+systems+come+out+on+top.&cid=ce9e2544fd98491082e83d5975a00c7b&aqs=edge..69i57.317j0j4&FORM=ANAB01&PC=U531>.
- Meirieu, P., 2007. *Pédagogie: le devoir de résister*. Issy-les-Moulineaux. Éditions Sociales Françaises (ESF).
- OECD, 2005. *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. OECD, Paris.
- Reid, W.A., 1999. *Curriculum as Institution and Practice. Essays in the Deliberative Tradition*. Lawrence Erlbaum, Mahwah, NJ & London.
- Schwab, J.J., 1969. The practical: a language for curriculum. *Sch. Rev.* 78, 1–23.
- Sellar, S., Thompkins, G., Rutkowski, D., 2017. *The Global Education Race: Taking the Measure of PISA and International Testing*. Brush Education, Edmonton.
- Slavin, R., 2002. Evidence-based educational policies: transforming educational practice and research. *Educ. Res.* 31, 15–21.
- Stenhouse, L., 1988. Artistry and teaching: the teacher as a focus of research and development. *J. Curric. Superv.* 4, 43–51.
- Thomas, G., 2021. Experiment’s persistent failure in education inquiry, and why it keeps failing. *Br. Educ. Res. J.* 47, 501–519.
- Van Manen, M., 1991. *The Tact of Teaching. The Meaning of Pedagogical Thoughtfulness*. SUNY Press, Albany, NY.
- Van Manen, M., 1995. Herbart und der Takt im Unterricht. In: Hopmann, S., Riquarts, K. (Eds.), *Didaktik und/oder Curriculum. Grundprobleme einer international vergleichenden Didaktik*. Beltz, Weinheim, pp. 61–80.

Tackling teacher shortages: doing the right thing at the right time and with the right evidence

Beng Huat See^a, Stephen Gorard^a, and Rebecca Morris^b, ^a Durham Evidence Centre for Education, Durham University, Durham, United Kingdom; and ^b Department for Education Studies, Warwick University, Coventry, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	655
Why don't we have enough teachers?	656
Lag time between pupil and teacher numbers	656
Limitations of the teacher supply model	657
Government policies that increase demand for teachers	657
Policies that suppress supply of teachers	658
Who wants to be a teacher?	660
Characteristics of potential teachers	661
Career motivations of intending teachers and non-intending teachers	662
Key factors differentiating the three groups	664
How to attract and retain teachers?	665
Promising approaches for attracting teachers to schools	665
Promising approaches for retaining teachers	667
Implications for research, policy and practice	668
References	668

Introduction

For decades countries across the world have reported shortages of teachers in some form or other, be it for certain subjects, in certain geographical locations or in some types of schools. There have been widespread media reports of teacher shortages in England (Bofey and Helm, 2015; Hazell, 2018; Sky News, 2017), the US and elsewhere (Caitlin, 2017; Passy, 2018; Strauss, 2016; Williams, 2018).

As pupil populations are projected to increase dramatically, international organizations, such as the OECD (Schleicher, 2012) and the European Cooperation in Education and Training, have called for the issue to be addressed urgently, and for a new approach to tackling teacher shortages (European Commission, 2015). Successive governments in many countries have invested in a raft of strategies and initiatives to address this issue, but the problems prevail. Much research has been funded and conducted around the world, but we are no closer to finding a long-term solution. One reason is because previous work has tended to look at teach shortages from a single perspective or only as a snapshot of the overall problem.

The demand for and supply of teachers are potentially influenced by a complex interaction of factors, but earlier work has tended to consider only a few factors in isolation. This kind of focus distorts the relative importance of some factors and so give misleading answers to the issue. Most studies of the supply pipeline also exclude through the research design, those individuals who do not apply for teaching or who have dropped out (e.g. Matthias, 2014; Heinz, 2015), ignoring the very people who can best explain the barriers to teaching (see Gorard, 2013). Relying on these kinds of studies for evidence, policy solutions consequently tend to be short-term and directed at economic and school environment factors, largely ignoring structural, societal and sociological ones. It is therefore not surprising that many of the current approaches to solving teacher shortages involve the use of monetary incentives (such as bonus incentives, higher wages, bursaries and scholarships), reduction in teachers' workload, and providing professional development, mentoring and induction for novice teachers. All of these are aimed at making teaching more appealing to attract and retain teachers, as reported by existing teachers themselves.

Given that current evidence on teacher demand and supply is based on partial understanding of the issue, initiatives to improve recruitment and retention based on monetary inducements are then not likely to be effective or appropriate. This explains why despite heavy investment in such programs the shortages persist, especially in subjects like maths and science. Most of these government and other initiatives have also never been robustly evaluated for independent evidence of their impact.

To address a complex issue like teacher supply would require a comprehensive approach that looks at the different factors together. In this article we describe our multi-pronged research approach, which includes policy documentary analysis looking into how policy decisions regarding planning on teacher numbers, curricular reforms in education and changes in admissions policies regarding teacher training can impact on teacher demand and supply. It also involves a time-series analysis to explore the complex interactions among the many variables associated with teacher demand and supply. To understand the issue better, we need also to consider individuals' decision-making process in their choice of teaching as a career. Therefore, we conducted a survey to understand undergraduates' career motivations. In addition, we synthesized existing international evidence to identify

potentially successful programs or interventions to attract and retain teachers. The combination of the four approaches provides a more complete picture for a targeted strategy to tackle the recurring teacher supply problem. This is, we believe, the first study that presents evidence from such a holistic approach.

This article summarizes the findings from the four strands of the study. We hope that this provides an example that can be developed for future research on the complex issue of teacher shortages. For this study, we used data from England and the focus is on the secondary sector where the shortage is most critical and chronic. We begin with an analysis of the supply issue and its possible causes, before moving on to examining why some people choose to go into teaching while others do not, or are hesitant. The latter is useful in designing policies to attract and retain teachers in the profession. Finally, we review international evidence to identify approaches that have the most promise in getting more people to teach in challenging areas, schools and shortage subjects.

Why don't we have enough teachers?

To get a better insight into the longstanding issue of teacher supply, we did a time-series analyses of official data from the DfE on number of schools and pupils, the School Workforce Census on teacher numbers, UCAS (Universities and Colleges Admissions Service) data on applications and acceptances into initial teacher training. Because education reforms can influence the demand and supply of teachers, we analyzed policy documents including reports from the School Teachers' Review Body, National College of Teaching, National Audit Office, Public Accounts Committee, the Houses of Commons and the Universities UK.

Lag time between pupil and teacher numbers

The number of teachers needed is assumed to be largely based on the projected pupil population, pupil:teacher ratios (PTR), and projected number of teachers entering and leaving. Such estimations are then likely to indicate a "shortage" because if the proportion of pupils enrolled is increasing, the rate of growth of the school population will exceed for a few years the number of graduates and teacher trainees from which potential teachers are drawn. Teachers take time to develop. For example, in England pupil numbers rose sharply from 2014 onwards (Fig. 1) but teacher numbers continued to fall from a peak in 2010. This was seen as the beginning of a teacher supply crisis. However, if we look at the numbers over time, the number of teachers available has tended to change with changes in pupil numbers although perhaps lagging behind pupil numbers by a year or so.

Similarly, the rise in PTR to an almost historic high in 2019 was a cause for concern, but this was not unexpected as pupil numbers had been increasing rapidly since 2014 (Fig. 2). Changes in the national PTR are largely due to the differential rates of growth in pupil and teacher numbers rather than simply a lack or surplus of teachers. In recent years, the PTR reached its highest point at 17.1 in 2004 before dropping to 15.5 in 2010. This was because teacher numbers continued to increase while pupil numbers were falling. The drop in PTR during this period was probably at least as much determined by the fall in pupil numbers from 2004, as the time lag before teacher numbers also began to fall. In summary, teacher numbers have historically kept pace with pupil numbers, and recruitment to postgraduate initial teacher training has generally responded to changes in intake targets albeit with a lag of 2 years or more.

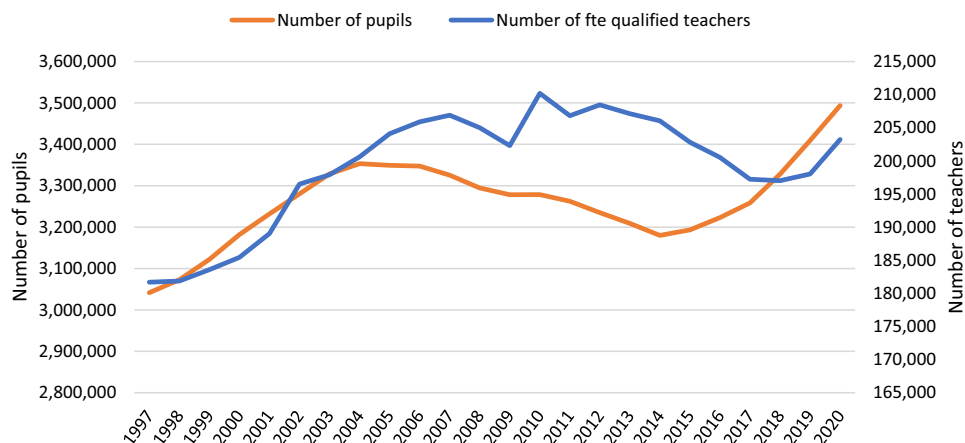


Fig. 1 Number of FTE qualified teachers and pupils in state-funded secondary schools (1971–2019).

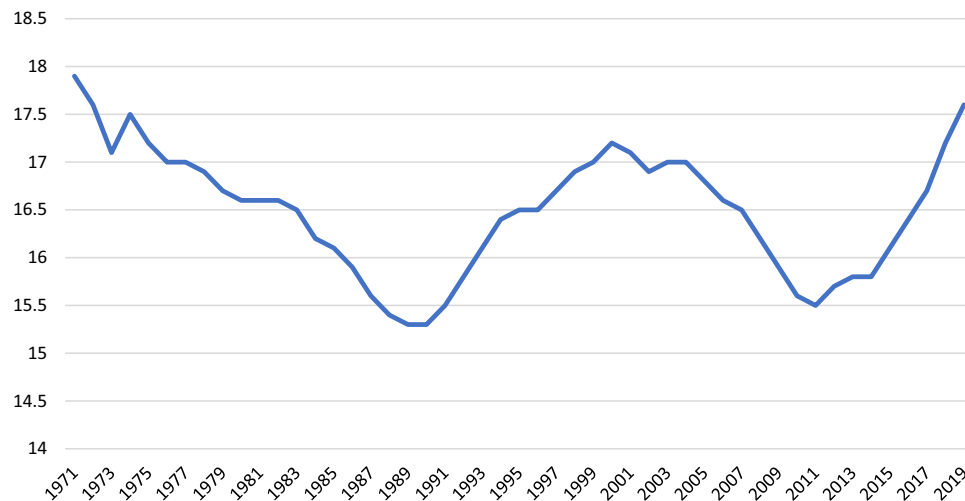


Fig. 2 Pupil:teacher ratios of qualified within school secondary teachers (1971–2019).

Limitations of the teacher supply model

In England, as in many countries, the number of teachers needed each year is estimated using a statistical model. This is known as the Teacher Supply Model (TSM). In the US, they have the TSDQ (Teacher Supply Demand and Quality) project models. TSM estimates the number of teachers needed based on the predicted number of pupils, the number of teachers leaving and the economic conditions, as well as how schools might react to these. Such factors are difficult to predict with certainty, and are usually beyond the control of the government. Information on the number of teachers is collected from the School Workforce Census, which is a snapshot measure carried out in November (from 2010 onwards), and therefore too late to influence recruitment for that academic year (House of Commons, 2017). This may underestimate the number of teachers needed as schools would have filled their September vacancies by using unqualified or supply teachers as a stopgap measure to ensure that classes are manned when term starts in September (House of Commons, 2017).

Government policies that increase demand for teachers

Sudden changes in government policy decisions can complicate planning for teacher supply. While TSM builds into its formula such policy changes, it can only do so when policy decisions have been confirmed. Where they are not confirmed, planning has to be based on a range of scenarios and a central estimate is often created and used (DfE, 2017). The NAO analysis suggests that these scenario-based estimations can vary between 25,000 and 38,000 in total number of teachers (National Audit Office, 2016a).

For example, from 2013/2014 onwards, the number of teachers recruited have consistently failed to meet the recruitment targets (the estimated number of new teachers needed). This shortfall became increasingly more apparent from 2015/16 onward where under 80% of the targets were met each year (Fig. 3). The number of subjects failing to meet targets had also increased. Many attributed this to the economic recovery, difficulties in getting new graduates into teaching because of competitions with other jobs, relatively poor pay of teachers and heavy workload of teaching (Borman and Dowling, 2017; Eteach, 2018; Foster, 2018; Hayes, 2017; Hobson et al., 2009; Public Accounts Committee, 2018).

But few have considered the influence of government policies. During the same period from 2013 onward, a number of education reforms were brought in. The leaving age for education and training was extended from 16 to 17 in 2013 and to 18 in 2015, the new¹ English Baccalaureate (EBacc) was introduced, and the GCSE curriculum was reformed (where the option to take Core Science was removed from 2016 so that some pupils would have to take two science subjects instead of one at GCSE). All these new policies have the effect of increasing the demand for teachers especially in traditional shortage subjects.

But the TSM is unable to plan ahead, and recruitment targets were only raised in the year when these policy changes came into being. This sudden and substantial increase of targets, for example, by 20% for maths and 30% for science within a year is likely to be problematic.

¹The EBacc is a school performance indicator where schools are measured on the number of pupils that take GCSEs in core subjects, including English, maths, the sciences, one humanities (history or geography) and a language, and on how well their pupils do in these subjects.

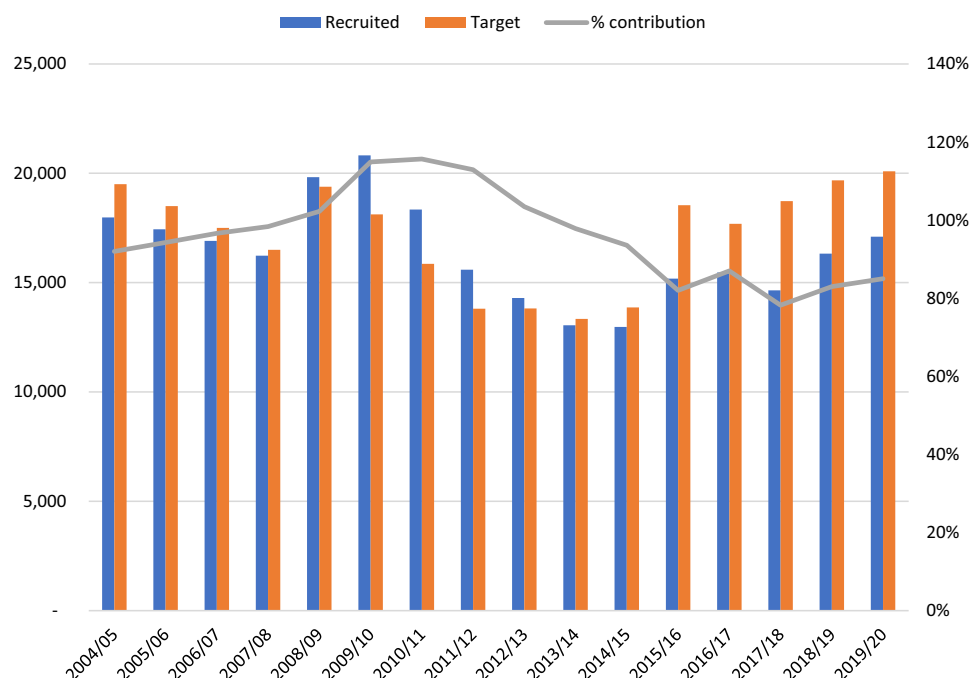


Fig. 3 Number of postgraduate ITT entrants and targets (total secondary).

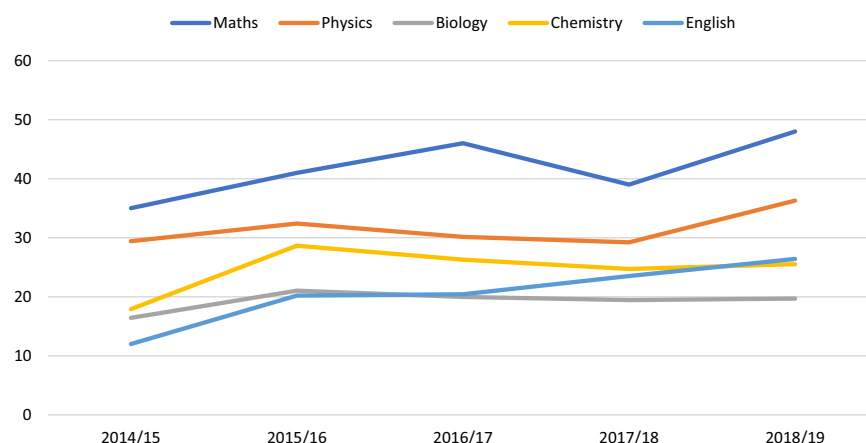


Fig. 4 Percentage of first degree graduates to be recruited into teaching to meet the ITT targets. Sources: <https://www.hesa.ac.uk/data-and-analysis/students/table-25#> and Initial Teacher Training: Trainee Number Census. <https://www.gov.uk/government/collections/statistics-teacher-training>; <https://www.hesa.ac.uk/data-and-analysis/students/table-25>.

And even if the number of new teachers required could be accurately predicted, it is still not possible to increase the number of graduates within a year since estimations are planned only 1 year in advance. If the targets for initial teacher training were to be met, we would need to recruit close to 50% of maths and 36% of physics graduates (Fig. 4). And these are subjects where there is reported unmet demand from industry and other employers as well.

Time is needed for the growth of graduates to meet the increasing demand. Teacher shortages are therefore, partly created by government policies as much as the mere increase in pupil population.

Policies that suppress supply of teachers

While some government policies have indirectly led to an increase in the demand for teachers, as illustrated above, other well-intended policies may actually have the effect of suppressing the supply of teachers. Although no one factor alone can be responsible for the teacher supply crisis, taken together they can exert considerable influence.

For example, in 2013 there were changes to the admissions policy to higher education institutions for teacher training. The then education secretary reduced the number of training places in higher education institutions, while at the same time increasing the

Table 1 Failure rate of skills test for 2012/13 to 2017/18.

Year	2012–2013	2013–2014	2014–2015	2015–2016	2016–2017	2017–2018
No. of trainees failed	3560	3639	4087	4202	4243	3748
Percentage failed	8%	9%	10.5%	10.5%	10.5%	9.5%

number of school-based training. As initial training for secondary school teachers (particularly for maths and science) was traditionally undertaken via university-led postgraduate courses, such policy decisions potentially reduced the number of people that can be trained. There were reports that School Direct (school-based training provider) was unable to fill a large number of places, leading to schools having to return their allocations to the² National College of Teaching and Leadership (NCTL). But NCTL did not allow allocations to be transferred between routes and subjects (National Audit Office, 2016a; Universities UK, 2013). So, while universities were oversubscribed they were not allowed to increase the number of training places allocated to them leading to many qualified applicants being turned away. We thus lost a large number of potential teachers even before the start of the pipeline.

At the same time, admissions into teacher training were made more difficult with the raising of the minimum entrance requirements and the Skills Test, which costs £24 each attempt. The Skills Test is a test of literacy and numeracy, which applicants had to pass before they can commence training. Applicants were allowed only three re-sits, and had to wait for 2 years before they could have another go (DfE, 2012). Around 10% of applicants failed the test every year (Table 1). In April 2019, it was discovered that over 200 potential trainees between September 2017 and 2018 were wrongly told that they had failed because of a marking error. Although passing the Skills Test is now no longer a requirement prior to training, the fact that thousands of people have applied to take the skills test indicates interest in teaching.

Government policies to raise the quality of teachers and thus teaching as a profession may have led to some risk-averse practices by teacher training providers, which have curtailed the numbers entering teaching. UCAS data show in the period from 2014 to 2017, the period when concerns of a teacher supply crisis were greatest, only under 60% of applicants were accepted into training (Fig. 5). In 2020, 15,080 applicants were rejected (see Fig. 11). Again, this shows that there is no lack of people interested in teaching.

Therefore, policies to increase the number of people going into teaching, such as bursaries, scholarships and bonus incentives may not be appropriate as the issue is not primarily a lack of numbers applying, but a mismatch between demand and supply. This mismatch, is to a certain extent, brought about government's own conflicting policies that have led to an increase in the demand for teachers while, at the same time, suppressing supply.

DfE's own analysis (Fig. 6) shows that those in receipt of higher bursary awards were less likely to end up teaching after qualifying than those awarded a smaller bursary (non-shortage subjects). Almost 20% of those awarded £25,000 or more (i.e. those training to teach maths and science) did not eventually go into teaching. This suggests that financial incentives alone are not the prime motivator for people to go into teaching.

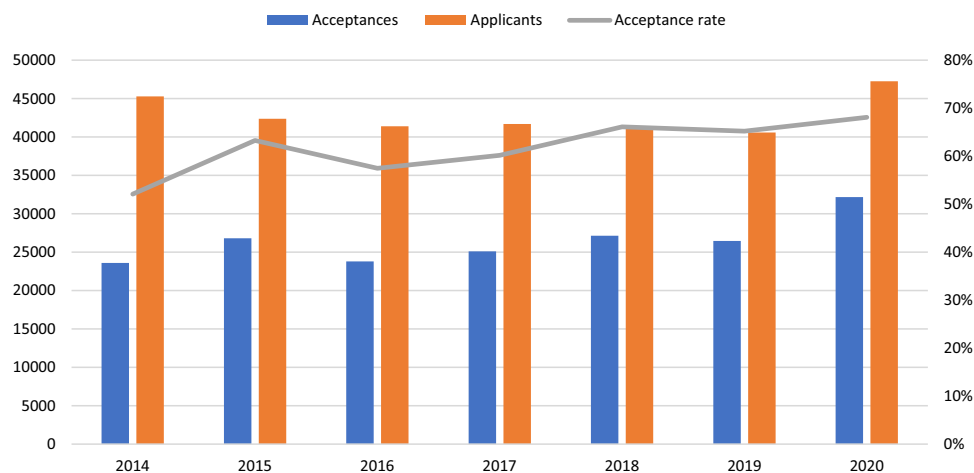


Fig. 5 Applicants and acceptances to all ITT courses (for England only). Source: UCAS Teacher Training statistics—applicants and acceptances by end of cycle 2014, 2015, 2016, 2017, 2018, 2019, 2020. *Acceptance figures include those who have accepted their conditional offers.

²NCTL was an agency for the DfE whose function was to ensure a provision of sufficient number of qualified teachers and maintain the quality of the teaching profession with oversight of the training of teachers. It has since been abolished.

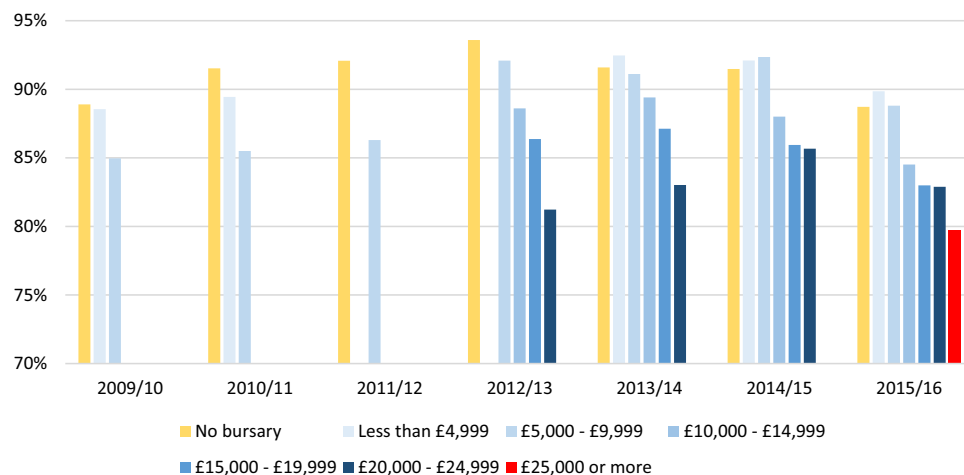


Fig. 6 Percentage of award recipients in teaching posts by the size of bursary award (2009/10–2015/16). Source: <https://www.gov.uk/government/statistics/teachers-analysis-compedium-4>.

Bursaries and other incentives might only attract people to train who have no intention to stay on as teachers (Higher Education Policy Institute, 2017). Doubts about the usefulness of bursaries and incentives to attract teachers have been expressed by the National Audit Office (National Audit Office, 2016b), and other commentators.

Yet as recently as October 2021 Boris Johnson, the British Prime Minister, announced a £3000 bonus incentive to entice new teachers in STEM subjects to teach in disadvantaged areas. Historically such schemes have been tried before in the UK and other countries, and have shown not to work (e.g. Clotfelter et al., 2008; Fitzgerald, 1986). This kind of policies are based on prior flawed research suggesting that teachers are motivated by financial gains, and offering such incentives can make the teaching profession and specific schools more attractive to current and potential teachers.

Individuals' choice of teaching as a career and their decisions to stay in teaching or not are generally not determined by government policy. There are deeper social and structural factors that account for people's behavior.

Who wants to be a teacher?

Looking at the data over time from 2009/10 to 2015/16 (Fig. 7), humanities students, maths, design & technology, biology and physical education teacher trainees were the most likely to continue into teaching (over 87%) after training while, physics, computing and modern languages trainees were less likely to go into teaching after qualification (c. 82%).

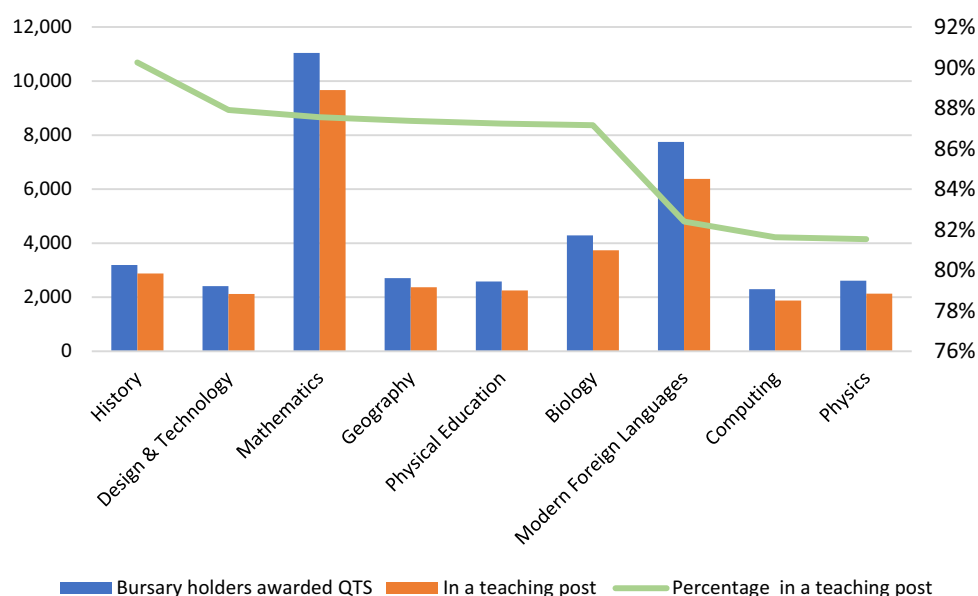


Fig. 7 Percentage of bursary holders in teaching in state-funded schools in England (between 2009/10 and 2015/16) by subjects.

This is consistent with our survey findings, which show that the strongest predictors of who is likely to have considered teaching are students' choice of subjects at university, their university entry qualification and their predicted degree classification (Gorard et al., 2021).

In our survey of c.5000 undergraduates in 53 universities in England across all disciplines, we asked students what they looked for in their career, the sources of information about careers they find useful, whether they have considered teaching, and which factors attract them to, or deter them from, teaching. We also collected information about parental occupation and education and the level of qualifications of respondents, and whether they felt that they could easily find a range of jobs. Crucially, our survey differed from many previous similar surveys in that we included undergraduates who have never considered teaching as a profession and those who had but decided not to be a teacher as well as those who have expressed firm intention to be teachers. Policies to attract people into teaching would do well to consider who is likely to be attracted to teaching. It would not be useful to try to recruit those who have never considered teaching. These are people studying subjects like law, medicine, dentistry or architecture, with clear professional outcomes. Likewise, the group who have firm intention to teach have already made the decision to go into teaching, so efforts to entice them into teaching will be redundant. It is the views of the middle group, those who have expressed interest but have decided against it, that any recruitment efforts should focus on. Findings of most previous work that focused only on those firmly intending to become teachers are misleading. For this group, the strongest predictors are their career concerns, such as pay, status and career prospects (Gorard et al., 2021) leading many researchers to conclude that addressing the relatively low pay, poor career progression and workload is a way to improving recruitment.

Our survey findings suggest that there are important differences between intending teachers and those who have expressed interest but do not intend to be teachers. It is this distinction that policy aimed at recruiting shortage subjects like maths, physics and chemistry should be focusing on.

Characteristics of potential teachers

Prospective teachers tend to be from less educated families with less prestigious occupational backgrounds and lower university entry qualifications (Table 2). They are more likely to enter university with a vocational qualification (e.g. a BTEC, GNVQ) and to expect lower degree results (Table 3) than their peers. Prospective teachers have lower tariff points on entry qualifications compared to those who have not considered teaching or had no intention to teach (Table 4). They are also more likely to study generic subjects like sport, English, classics, and history (Table 5). All of this suggests that higher attainers are less likely to be interested in teaching as a career.

However, student background characteristics, prior experiences and course choices are not malleable in the short-term, and so these differences do not help much in deciding how to attract more people into teaching. The kind of courses that students take is closely related to their career intention, suggesting that many have already made a decision prior to entry to university. To increase the number of teachers in some shortage subjects, like maths and science, might require an approach that targets students before they make their subject choice at university.

Table 2 Student background and their decision to teach.

<i>Gender</i>	<i>Considered %</i>	<i>Intend %</i>
Male	61.7	23.6
Female	54.8	14.1
Other	51.8	16.2
<i>Parental education</i>	<i>Considered %</i>	<i>Intend %</i>
Have a degree	54.6	15.4
Do not have a degree	65.5	26.4
Not known	54.7	18.7
Parental occupation		
Technical, health and education	63.4	21.4
Clerical, administrative assistant, secretary	60.0	20.4
Professional occupation (e.g. doctor, solicitor, scientist), university/college lecturer	52.6	14.7
Craft related jobs	66.7	28.9
Small employer (under 10 employees)	58.0	18.8
Not usually employed	67.3	30.6
Unknown	53.9	20.3

Table 3 Student prior qualification and their decision to teach.

<i>Prior qualification</i>	<i>Considered %</i>	<i>Intend %</i>
A-level	59.2	17.0
International baccalaureate	43.7	13.7
BTEC, GNVQ, other professional diploma	72.6	43.9
Access to higher education diploma	50.0	27.9
Scottish highers or advanced highers	63.7	5.0
A-level and BTEC/IB	76.4	49.6
Foundation year	66.7	8.3
Other or not known	49.6	15.8
<i>Expected degree results</i>	<i>Considered %</i>	<i>Intend %</i>
1st	54.5	16.8
2:1	61.2	21.3
2:2	69.9	30.8
3rd or pass	56.7	20.0
Not known	54.9	17.3

Table 4 University entry grade points and decision to teach.

<i>KSS tariff points</i>	<i>Mean</i>	<i>Standard deviation</i>	<i>Effect size</i>
Considered	133.80	23.68	
Not considered	137.19	20.55	−0.15
Intend to be a teacher	127.28	25.60	
Not intend	137.15	21.22	−0.44
Total	135.20	22.50	

Table 5 Course of study and decision to teach.

	<i>Considered %</i>	<i>Intend %</i>
Sports-related courses	75.5	42.8
Languages, English, classics	70.3	26.2
Other courses	66.7	16.7
Social, economic, political sciences and humanities, psychology	65.8	27.0
Creative arts and design, library and information science, media studies	65.1	22.9
Physical and mathematical sciences, computing, engineering and technology	52.7	11.8
Business, accountancy and administrative studies	41.7	14.2
Medicine, dentistry, biological sciences, veterinary sciences, agriculture and forestry	40.4	5.6
Law, architecture, building and planning	34.1	4.2

Career motivations of intending teachers and non-intending teachers

When students were asked to rate what they considered important in their choice of career (from 0—not important at all to 10—very important), the majority rated job satisfaction/enjoyment, subject interests, career prospects, job security and pay/salary (Table 6). Workload, incentives to train, introductory bonus and job status are least of their concerns.

But when the responses of those who have considered, intend to be teachers and those who have not considered teaching were compared, the differences between groups become apparent (Table 7). Potential teachers are more likely to report being motivated by having a chance to share their knowledge and give something back compared to their peers. They are less concerned with job status, career prospects and pay (column one). Perhaps this is why studies that focus only on teachers might downplay the importance of these extrinsic motivators in comparison to the more altruistic ones. For those who might otherwise have chosen teaching as a career, pay and career prospects seem to be more important to them. Crucial findings like this are lost when we consider only those who want to teach.

Table 6 Factors influencing students' career choice.

	<i>Mean</i>	<i>Standard deviation</i>
Job satisfaction, enjoyment	8.77	1.44
Interest in my subject area	7.66	2.33
Career prospects	7.60	1.92
Opportunity to develop skills	7.59	1.98
Job security	7.50	1.96
Pay, salary	7.26	2.04
Kinds of people I will be working with	7.03	2.35
Intellectual stimulation	6.95	2.26
Job that suits my temperament	6.85	2.37
Chance to give something back	6.84	2.56
Job responsibility	6.60	2.18
Autonomy, scope for initiative	6.51	2.18
Chance to share my knowledge	6.28	2.49
Chance to use academic knowledge	6.25	2.55
Ease of getting a job in that field	6.02	2.58
The workload required	5.89	2.43
A financial incentive to train	5.75	3.06
Length of working day, holidays	5.72	2.68
Status, public perception of the job	4.35	2.91
Introductory bonus when starting job	4.31	3.02

Table 7 Factors influencing potential teachers' career choice (effect sizes).

	<i>Considered</i>	<i>Intend</i>
Chance to give something back	+0.28	+0.44
Chance to share knowledge	+0.25	+0.42
Use academic knowledge	+0.12	+0.22
Length of working day	+0.12	+0.17
Ease of getting job	+0.09	+0.28
Job satisfaction	+0.07	+0.07
Workload	+0.05	+0.04
Job security	−0.01	+0.16
Introductory bonus	−0.04	+0.01
Autonomy	−0.06	−0.02
Job status	−0.12	−0.05
Career prospects	−0.13	−0.13
Pay	−0.23	−0.27

Note: Comparison is made between those who have considered teaching, those who intend to teach with those who have not considered teaching. The difference is expressed as effect sizes, calculated as the difference in means between the groups and the overall standard deviation.

Those who have considered teaching are also more likely to be influenced by financial incentives to train compared to those who have not considered teaching (Table 8). Financial incentives appear to be attractive to those who are already interested in teaching, but not those who have not considered teaching. To attract those who might have considered teaching to become teachers, policies would need to focus on job satisfaction, career prospects and subject interest, rather than money alone. This is where current government policies to increase the number of people into teaching (particularly those in STEM subjects) has failed to consider.

Table 8 Influence of financial incentives to teach (effect sizes).

<i>Incentives to teach</i>	<i>Considered</i>	<i>Intend</i>
Bursary for training	+0.48	+0.62
Training salary	+0.47	+0.58
Loan for tuition	+0.42	+0.58
Loan for maintenance	+0.40	+0.57

Key factors differentiating the three groups

To identify the important factors distinguishing the three groups, further analyses using logistic regression modeling were conducted. The first model compares those who had considered teaching and those who had not. The second model compares those who had considered teaching and indicated that they intended to teach with those who had considered but were not intending to teach.

The predictive factors included student family background, year of study at university, factors influencing their career motivation, sources of information about careers, perceptions of teaching as a career, and the role of financial incentives to become a teacher. These factors were added in chronological order with financial incentives added last to see how much each factor explains differences between groups.

The analyses show that the strongest predicting factors are those related to students' university years (Table 9). In other words, if we have information about factors related to students' subject choice at university, year of study, their university entry qualification and expected degree classification, we can predict with 66% accuracy who had considered teaching and who had not (model 1), and 71% accuracy (model 2) who had considered teaching and intend to teach or not. Adding these factors increases the accuracy of prediction by 4% points in model 1 and 5% points in model 2.

Among those who had considered teaching, the factors that most predict who would go into teaching or not are students' career motivations. Adding these factors increases the predictive accuracy by eight percentage points. Net of these factors, knowing students' perceptions of teaching as a career does not help much in predicting if they intend to teach or not. When comparisons are made with these two groups, financial incentives are not the deciding factors in students' intention to teach. Such incentives are important to those who are already considering teaching, but do not make a difference to their decision to teach or not.

These findings suggest that there are deeper, structural reasons for graduates' choice of teaching as a career, which may not be amenable to financial manipulations. Psychologists have long believed that such extrinsic motivators do not alter attitudes that underlie human behavior (Kohn, 1994). People are concerned about money only if they are not paid well. Therefore, research that asks people what matters to them in their choice of career, money/salary is often top of the list, but if they were asked what they care about, money becomes less important (Kohn, 1993). As Herzberg has argued, just because low pay is a disincentive, paying people more does not necessarily increase motivation to teach.

Most previous research on teacher shortages often conclude that the relatively low pay, poor career progression and workload as factors. These tended to be based on economic theory of reward and motivation. Sociological studies on choice of career and participation in training suggest that, to a certain extent, individual's choice of career is influenced by their socio-economic and cultural background (Gorard and Rees, 2002; See, 2011; Gorard et al., 2021). Therefore, a sociological model may provide a better explanation for individuals' choice of career thus offering a partial explanation for why the bursary schemes are not the ultimate solution. This has implications for future policy on teacher recruitment and retention.

Table 9 Regression models to predict who are likely to consider teaching or not and those who intend to teach.

<i>Block</i>	<i>Model 1</i>		<i>Model 2</i>	
	<i>% Predicted correctly—considered</i>	<i>Increase on previous figure</i>	<i>% Predicted correctly—intend</i>	<i>Increase on previous figure</i>
Base	60.6	—	65.1	—
Background	61.2	0.6	65.6	0.5
University	65.5	4.3	71.1	5.5
Career	68.1	2.6	79.4	8.3
Teacher factors	70.5	2.4	80.0	0.6
Incentives to teach	71.5	1.0	80.0	0

How to attract and retain teachers?

The survey result, the time-series analysis and policy analysis have suggested that current initiatives to increase teacher recruitment and retention through monetary incentives and workload reduction can only offer partial and short-term solutions to the teacher supply crisis.

To understand what approaches are most effective in attracting and retaining teachers, we reviewed and synthesized international research that evaluates strategies used to address teacher shortages. Knowing more about what “works” and what does not will allow policymakers and schools to make informed and targeted decisions on strategies to use. As the findings are meant to inform policies, it is essential that the evidence supporting these strategies is based on rigorous research. To ensure that the findings are reliable and free from bias we subjected each study to a quality appraisal using a “sieve” heuristic (see [Table 10](#)). This process rates each study based on its research design, scale, level of attrition, and outcome measures, which are key factors that can affect the validity of the results. Each piece of research is then assigned a score in the form of a padlock rating indicating the security of the findings. For more details on how to use the “sieve,” see [Gorard \(2021 p. 94\)](#).

The review identified 6731 potentially relevant titles from a search of the standard social, educational and psychological databases, e.g. JSTOR, Web of Science, Science Direct, ERIC, EBSCOhost (which covers the following databases: PsychINFO, British Education Index, PsycARTICLES, etc, ProQuest, International Bibliography of the Social Sciences, Google and Google Scholar. Another 347 studies from prior reviews or previously known work were added to the list. Of the 7078 studies, 120 deemed relevant to the research question were retained for data extraction, and their findings synthesized (see examples in [Table 11](#)). Details appear in [See et al. \(2020a\)](#).

Promising approaches for attracting teachers to schools

The only approach that appears to work in attracting teachers to hard-to-staff schools and areas (not into the profession) is the offer of financial incentives. However, these incentives do not offer long-term permanent solutions. A number of studies reported that bonus incentives can help increase the number of existing teachers in low performing schools ([Glazerman et al., 2013](#); [Steele et al., 2010](#)) and schools with a high proportion of disadvantaged pupils ([Fitzgerald, 1986](#); [Hough and Loeb, 2013](#)). But the impact is temporary, and lasts only while the incentive is available. Once the money stops (or even before), the teachers leave at the same rate as before.

Monetary incentives also work in attracting teachers to challenging schools, but only for relatively high performing schools with lower proportions of disadvantaged children ([Fulbeck and Richards, 2015](#); [Henry et al., 2012](#)). It is also the case that such incentives are effective only in attracting those who are already interested in teaching ([Rosen, 2012](#); [Dolan et al., 2012](#)). To get teachers to teach in challenging communities and school would require paying teachers substantially more to compensate for factors that might make the community or schools less attractive, such as costs of living and working conditions ([DeFeo et al., 2018](#); [Falch, 2017](#); [Goldhaber et al., 2010](#); [Zarkin, 1985](#)). Financial incentives are also only effective where there is a tie-in. In Norway, for example,

Table 10 The quality appraisal “sieve” for causal studies.

<i>Design</i>	<i>Scale</i>	<i>Dropout</i>	<i>Outcomes</i>	<i>Other threats</i>	<i>Rating</i>
Strong design for causal research question (e.g. RCT)	Large number of cases per comparison group	Minimal attrition with no evidence that it affects the outcomes	Standardized pre-specified independent outcome	No evidence of diffusion or other threat	4🔒
Good design for causal research question (comparison such as regression discontinuity, difference-in difference)	Medium number of cases per comparison group	Some initial imbalance or attrition	Pre-specified outcome, not standardized or not independent	Indication of diffusion or other threat, unintended variation in delivery	3🔒
Weak design for causal question (e.g. matched comparison using propensity score matching)	Small number of cases per comparison group	Initial imbalance or moderate attrition	Not pre-specified, but valid outcome	Evidence of experimenter effect, diffusion or variation in delivery	2🔒
Very weak design for causal question (poor comparison with no equivalence, e.g. comparing volunteers with non-volunteers)	Very small number of cases per comparison group	Substantial imbalance or high attrition	Outcomes with issues of validity and appropriateness	Strong indication of diffusion or poorly specified approach	1🔒
No consideration of design	A trivial scale of study (or N unclear)	Attrition not reported or too high for comparison	Too many outcomes, weak measures or poor reliability	No consideration of threats to validity	0

Table 11 Results of studies on financial incentives on recruitment.

<i>Positive outcome</i>	<i>Unclear/mixed outcome</i>	<i>Neutral/negative outcome</i>	<i>Security rating</i>
• DeFeo et al. (2018) • Fitzgerald (1986) • Glazerman et al. (2013) • Hough and Loeb (2013) • Steele et al. (2010) • Dolan et al. (2012) • Falch (2017) • Zarkin (1985)	Rosen (2012) Fulbeck and Richards (2015)	Bueno and Sass (2016) Gorard et al. (2021) Kraft et al. (2020)	3A 2A
	Goldhaber et al. (2010) Fowler (2003) Gordon and Vegas (2004)		1A

As this table portrays, there is no very strong research in this area at all.

teachers working in high-vacancy schools receive a wage premium while there, but they lose this once they move to a low-vacancy school (Falch, 2017). Other studies (e.g. Bueno and Sass, 2016) found that paying shortage subject teachers (e.g. maths and science teachers) higher salaries did not increase the number of maths or science teachers nor did it encourage people to switch to maths or science.

There are studies that include a range of other approaches (e.g. professional development and mentoring) alongside offers of financial incentives or bonuses, thus making it difficult to determine if it is the attraction of the money, the mentoring or a combination that drive the effects. For example, Bobronnikov and Price (2013) suggest that an incentive grant increased the number of recipients teaching maths and science in high need areas, but the program also came with a 5-year tie in and other mentoring and professional development support. Similarly, Scott et al. (2006) reported an increase in the number of people recruited to teach maths and science under a program that offered a scholarship. But the program also involved a 2-year commitment plus mentoring and tuition fee waivers. Others, like Rothstein (2015), suggest that bonus contracts were effective in attracting teachers of higher ability, but the impact was small. The tenure contract was more successful in increasing the number of high ability teachers as lower ability teachers would leave voluntarily because they know that their chances of tenureship are low. One problem with this research is that it assumes that teacher performance assessment is unbiased and that new teachers are recruited from the same population as current teachers ignoring the fact that there are potentially high ability teachers who would not consider teaching at all. One side effect of higher salary for tenureship is that class sizes or pupil:teacher ratios need to be increased to finance the higher teacher salary.

These incentives are not usually intended to attract more people into teaching, but merely to move existing specialist teachers to new areas. Unless there is a surplus of shortage subject teachers available to be deployed, enticing teachers to some areas might leave other areas understaffed. That could result in a fairer distribution of teaching resource, but it would be an equality of insufficiency. There are some suggestions from less robust correlational studies that financial incentives alone may not be enough to compensate for poor working conditions, issues with school leadership and school climate (Waters-Weller, 2009; Fulbeck, 2014).

Also, even if offering higher salaries can encourage more people into teaching, we need to ask if we really want people who are attracted by salaries alone. The preceding two sections have shown that those who are attracted by bursaries and scholarships are less likely to enter teaching after qualification and more likely to leave anyway. Money does not fundamentally change people's decision to teach or not. Individuals' career intentions are part of a longer-term "identity."

Other weaker studies in terms of research design for a causal question show mixed results. One study looked at offering highly qualified people a \$20,000 bonus incentive to switch careers to teach in high need areas (Fowler, 2003). Despite advertising across states, only seven candidates outside the state were recruited over 4 years. The program also failed to place all teachers in high-need schools. Dropout among bonus recipients was higher than the national average. Head teachers reported that bonus recipients were more attracted to the fast-track scheme than the bonus incentive (Churchill et al., 2002). Other studies also suggest that generally the effect of bonus incentives tends to be short-lived (Choi, 2011).

Overall, there is no strong evidence that monetary inducements work. Salary increases (within the usual ranges), bonuses and cash incentives do not attract people to teaching (Gorard et al., 2021). Bonus incentives, like the one that the UK Prime Minister promised at the 2021 party conference, are an apparent quick fix with no lasting effects. Moreover, incentives to try to attract teachers to specific local areas or schools could be at the expense of other schools, so may not benefit the system as a whole.

Apart from financial incentives, there is no good evidence that other approaches to attract teachers to teach in difficult-to-staff areas or shortage subjects work. This is largely because research on these approaches is not even as strong as those for initiatives involving monetary inducements (See et al., 2020b). Of course, it is possible that better research in some of these areas has been conducted since our review. Therefore, our understanding of what works or not and in what context will be revised accordingly as more and better studies are conducted.

Approaches, like Grow Your Own, for example, where teachers are trained and recruited from the local community, have not been robustly evaluated. The current evidence we have is that we do not know if such an approach works or not. Research in this area has tended to be small-scale and based on anecdotal reports of successful practice. It is, therefore, not possible to draw causal conclusions. We cannot say definitively that the strategy leads to increasing number of teachers, or if something else is happening—such as economic events which can encourage people to go into teaching, but have nothing to do with any concurrent initiative to attract teachers.

We also do not have convincing evidence that increasing the number of ways people can obtain qualification or accreditation helps improve recruitment, largely because the different routes are so varied in terms of the kind of people being targeted, and the extent to which they are actually different from the “traditional” routes on offer.

Teach for America (US) and its international equivalents, such as Teach First (England), Teach for Australia, and Teach for India, as well as other school-based training such as Schools Direct, Boston Residency program, School Centered Initial Teacher Training and the Troops to Teachers initiative are examples of alternative routes into teaching. The most promising study in this group is a Maths Immersion Program (MIP). [Boyd et al. \(2012\)](#) reported that the program was successful in attracting highly qualified teachers to teach in some of the most challenging schools. More teachers were trained through this route than the traditional college route and other alternative pathways, such as Teaching Fellowships (TF) and Teach for America. Similar studies also reported increases in the number trained via alternative routes (e.g. [Clewell and Villegas, 2001](#); [Harrell and Harris, 2006](#); [Papay et al., 2012](#)). While the results look promising, we have to consider that in most of these studies, participation in alternative routes was self-selected. Those who signed up for traditional programs were likely to be different from those who signed up for the online or alternative program. It is therefore not possible to rule out other exogenous factors, such as the graduate labor market conditions, which may have affected people’s decisions in changing careers. In some cases, candidates were pre-screened and selected for their potential success. None of the studies were able to effectively randomly assign recruits to the different types of program to test relative effectiveness. One study also suggests that the success or otherwise of alternative routes into teaching depends on the context within which they are implemented.

Promising approaches for retaining teachers

As with recruitment, the majority of the stronger studies on retention, in terms of research design, also involved some kind of financial incentives. However, the evidence is even weaker here. Although many studies do show positive results, the more robust studies which control for context suggest that money only keeps people in teaching or in the high need areas as long as the incentive is available. Such interventions are therefore not useful in addressing permanently the chronic shortage of teachers.

One such incentive is performance-related pay. Performance incentive pay may be effective in retaining effective teachers and removing less effective teachers. The assumption is that low-performing teachers would rather leave voluntarily than to be dismissed ([Dee and Wyckoff, 2015](#)). However, the effect is not clear cut. It depends on whether the incentive pay is awarded to individual teachers or to the whole school. [Springer and Taylor \(2016\)](#) found that individual awards are more effective for experienced teachers while school-level awards are more effective for beginning teachers. A randomized control trial also showed that school-level or group incentives may not be effective in retaining teachers ([Fryer, 2013](#)).

Other commentators have argued that financial incentives need to be large enough to neutralize the disadvantage of working in challenging schools or areas ([Booker and Glazerman, 2009](#)). However, the results are difficult to interpret. Dee & Wyckoff’s quasi-experimental study shows receipt of award improved retention, but regression discontinuity analyses ([Shifrer et al., 2017](#)) found that a large award had a negative effect compared to small award. This is perhaps because teachers in receipt of a large award are high performing teachers who can easily find better paid jobs elsewhere, and hence more likely to move. [Springer et al. \(2010\)](#) suggest that awards over \$6000 have no effect on teachers’ motivation to stay. So the amount and the implementation of such awards may be important considerations, as well as the fact that these incentive payments come with an agreement to commit to teaching targeted subjects or schools. Data from the Teacher Follow Up Survey showed that performance pay was not considered as the key reason for teachers’ decision to leave ([Jones, 2013](#)). In fact, the evidence suggests that the use of discriminatory incentives may even worsen overall retention.

In many cases, monetary incentives work only because teachers are required to commit to teach for a specified period or certain subjects in specified schools or areas as part of the contract agreement. These incentives often entail a penalty for breaking the contract, raising questions about the value of such an approach and the potential for a kind of enforced retention where teachers may feel “tied-in” to a role that they no longer wish to do.

Other financial incentives, such as loan forgiveness and one-off retention bonuses (e.g. [Feng and Sass, 2015, 2018](#)) targeted at beginning teachers teaching shortage subjects work only because there were strings attached to them. To be eligible for the loan forgiveness, beginning teachers had to agree to teach a shortage subject for 90 days and to be eligible for the retention bonus, teachers had to agree to continue teaching the shortage subject the following year.

Differentiated incentives to retain older teachers, with the amount paid varying by length of experience, have been not shown to work either ([Booker and Glazerman, 2009](#)).

Other approaches involving mentoring, induction programs, leadership support, flexible working, professional development and improved workload have not been robustly evaluated. And some studies show negative results. The jury is still out on most of these interventions.

Some correlational and survey-based studies suggest that improving school cultures and ethos can improve recruitment and retention. However, very few rigorous studies robustly evaluated such interventions related to areas such as accountability, teacher stress, working conditions, workload or leadership support. These are areas worth exploring further.

Implications for research, policy and practice

Our multi-strategy research approach combining survey data, time-series analysis of official data, policy documentary analysis and systematic review offers a comprehensive and holistic way of understanding the complex determinants of teacher demand and supply better. Tackling teacher shortages is a complex process involving multiple stakeholders, including policymakers, teacher education providers, schools and teachers. It requires longer-term, more joined-up planning than is usual, and well-resourced, evidence-informed policies, along with system-wide shifts in practice.

Our study finds that monetary incentives are not feasible long-term solutions to addressing the teacher supply problem. While they are important to those who are considering teaching, they do not make a difference in their decision to teach or not. Students' decisions to be teachers or not are largely mapped out by the time they are in university. Students' choice of subjects at university, their entry qualifications, their predicted degree classification and their parental occupation and education are strong determinants of who are likely to be teachers or not. However, these are not malleable, so knowing these differences is not very useful for policies to attract more people into teaching. But what we know is that teaching seems to be attracting lower performing students from less prestigious backgrounds. Perhaps, in the long-term, more thought could be put into how teaching could be made more attractive to those who would otherwise have chosen teaching as a profession. And this should be tackled much earlier in people's education so that by the time they reach university, whatever subjects they are studying, teaching will be seen as an attractive potential alternative for them.

Our analyses also show that there is no lack of people wanting to go into teaching. The fact that thousands have applied and been rejected every year suggests there is an artificial barrier to entering the teaching profession, unless of course, teaching is genuinely attracting tens of thousands of unsuitable candidates. We found that government policies on education can have a far more direct and immediate effect on teacher numbers than commonly cited reasons such as teacher salaries or workload. For example, government policies can increase or decrease teacher demand and supply almost overnight by changing mandatory class sizes, pupil:teacher ratios, entry requirements for teacher training, adjusting the number of training places available, reforming the way schools are funded and changing the number and types of schools (See and Gorard, 2020). Teacher shortages are therefore not simply the result of a growth in pupil numbers or the relative lack of attractiveness of teaching compared to other professions.

We have not found the "magic bullet" yet, but we are closer to understanding the determinants of the problem, and this is a positive step toward finding a solution. Almost any economic, legal, structural or cultural change over time could influence how many teachers are needed and/or how desirable teaching is as a career. A key finding is that policy interventions to address the challenge of teacher supply often only focus on the symptoms rather than the cause of teacher shortages. Our analyses of secondary data and government policies show that policies aimed at raising the quality of teachers can have the effect of reducing the number of teachers by raising the entry requirements. Therefore, balancing the quantity and quality of teachers needs to be carefully considered for their possible implications for teacher supply, and they need to be considered several years before actual implementation. The legislative programs of administrations in the UK are currently fragmentary and naïve in terms of timing.

Our new review highlights the gaps in many key areas of research (such as on workload). And most research in these areas is so poor (having no context, or such partial coverage) that it is of no use to inform policies. This is an issue for funders, universities, and governments to address. It is an ethical issue about the use of public funding and wasted opportunity costs as much as anything else.

More creative approaches to retention, such as those employed in business organizations, could be explored. This could include, for example, rewarding long-serving staff by offering sabbaticals for teachers after a certain number of years in post, to allow them to travel, conduct their own research or pursue further study. These ideas could then be tested properly to see whether they can help to reduce stress and burnout, and thus improve retention. Future research could focus much more on those young people who have considered teaching but rejected it to see how they could be enticed into teaching.

We could also consider the way in which school places are allocated, so that there are no clearly defined schools with such high levels of poverty as at present, meaning that certain schools might not be as hard to staff, even though some would remain geographically isolated. Of course, a wider and more effective solution would be to work beyond education to create better economic conditions in poorer areas anyway.

References

- Bobronnikov, E., Price, C., 2013. Use of a C-SITS Approach to Estimating the Impact of the Receipt of a Teacher Recruitment Incentive Grant on an IHE's Production of Certified STEM Teachers—Problems and Solutions. Paper Presented at the Society for Research on Educational Effectiveness Conference, Spring 2013.

- Boffey, D., Helm, T., 2015. Shortage of teachers set to spark new school crisis. *Guardian*. www.theguardian.com/education/2015/aug/29/shortage-teachers-new-schools-crisis-uk-trainee-shortfall.
- Booker, K., Glazerman, S., 2009. Effects of the Missouri Career Ladder Program on Teacher Mobility. Mathematica Policy Research, Inc, Washington, DC.
- Borman, G., Dowling, N., 2017. Teacher attrition and retention: a meta-analytic and narrative review of the research. *Rev. Educ. Res.* 78 (3), 367–409.
- Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., Wyckoff, J., 2012. Recruiting effective math teachers: evidence from New York City. *Am. Educ. Res. J.* 49 (6), 1008–1047.
- Bueno, C., Sass, T.R., 2016. The Effects of Differential Pay on Teacher Recruitment, Retention and Quality. Department of Economics Andrew Young School of Policy Studies, Georgia State University.
- Caitlin, O., 2017. Schools Throughout the Country are Grappling With Teacher Shortage. <https://edition.cnn.com/2017/08/21/health/teacher-shortage-data-trnd/index.html>.
- Choi, J., 2011. When are Signing Bonuses More than Just “Pay to Play”? An Experimental Investigation. Doctoral Dissertation. Emory University.
- Churchill, A., Berger, J., Brooks, C., Effrat, A., Griffin, L., Magouirk-Colbert, M., McDermott, K., Sharick, R., Sheehan, A., 2002. Revised Interim Report: The Massachusetts Institute for New Teachers and Master Teacher/NBPTS Programs. Center for Education Policy—University of Massachusetts.
- Clewell, B.C., Villegas, A.M., 2001. Absence Unexcused: Ending Teacher Shortages in High-Need Areas. Evaluating the Pathways to Teaching Careers Program. Urban Institute, Washington, DC.
- Cloffelter, C., Glennie, E., Ladd, H., Vigdor, J., 2008. Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *J. Publ. Econ.* 92 (5), 1352–1370.
- Dee, T.S., Wyckoff, J., 2015. Incentives, selection, and teacher performance: evidence from IMPACT. *J. Pol. Anal. Manag.* 34 (2), 267–297.
- DeFeo, D., Hirschberg, D., Hill, A., 2018. It's more than just dollars: problematizing salary as the sole mechanism for recruiting and retaining teachers in rural Alaska. In: *Wellness and Healing: Indigenous Innovation and Alaska Native Research*. University of Alaska Anchorage, Anchorage, AK, USA. Proceedings of the Alaska Native Studies Conference, Juneau, AK, USA, 13–15 April 2018.
- DfE, 2012. New Tougher Tests for Trainee Teachers. Department for Education, London, England. <https://www.gov.uk/government/news/new-tougher-tests-for-trainee-teachers>.
- DfE, 2017. Postgraduate Initial Teacher Training Places and the Teacher Supply Model, England (2017/18). Department for Education, London, England. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/655038/SFR42_2017_TSM_Main_Text.pdf.
- Dolan, P., Metcalfe, R., Navarro-Martinez, D., 2012. Financial Incentives and Working in the Education Sector; Research Report DfE-RR251. Department for Education, London.
- Eteach, 2018. Attract, Recruit, Retain: Teacher Recruitment for a Modern Education. The Eteach Report 2017–2018. Eteach, Camberley, Surrey.
- European Commission, 2015. Joint Report of the Council and the Commission on the Implementation of the Strategic Framework for European Cooperation in Education and Training (ET 2020). European Commission, Brussels. http://ec.europa.eu/dgs/education_culture/repository/education/documents/et-2020-draft-joint-report-408-2015_en.pdf. (Accessed 16 August 2018).
- Falch, T., 2017. Wages and recruitment: evidence from external wage changes. *ILR Rev.* 70 (2), 483–518.
- Feng, L., Sass, T.R., 2015. The Impact of Incentives to Recruit and Retain Teachers in “Hard-To-Staff” Subjects: An Analysis of the Florida Critical Teacher Shortage Program. Working Paper 141. National Center for Analysis of Longitudinal Data in Education Research (CALDER).
- Feng, L., Sass, T.R., 2018. The impact of incentives to recruit and retain teachers in “hard-to-staff” subjects. *J. Pol. Anal. Manag.* 37 (1), 112–135.
- Fitzgerald, C., 1986. Report on the High Priority Location Stipend Program. Dade County Public Schools Office of Educational Accountability, Miami, Florida. <https://eric.ed.gov/?id=ED283862>.
- Foster, D., 2018. Teacher Recruitment and Retention in England. Briefing Paper Number 7222. House of Commons Library, London, England.
- Fowler, R., 2003. The Massachusetts signing bonus program for new teachers: a model of teacher preparation worth? *Educ. Anal. Arch.* 11 (13), 1–24.
- Fryer, R.G., 2013. Teacher incentives and student achievement: evidence from New York City public schools. *J. Labor Econ.* 31 (2), 373–407.
- Fulbeck, E., 2014. Teacher mobility and financial incentives: a descriptive analysis of Denver's ProComp. *Educ. Eval. Pol. Anal.* 36 (1), 67–82.
- Fulbeck, E., Richards, M., 2015. The impact of school-based financial incentives on teachers' strategic moves: a descriptive analysis. *Teach. Coll. Rec.* 117 (9), 1–40.
- Glazerman, S., Protik, A., Teh, B., Bruch, J., Max, J., Warner, E., 2013. Transfer Incentives for High-Performing Teachers: Final Results From a Multisite Randomized Experiment. Executive Summary (NCEE 2014-4004). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education.
- Goldhaber, D., Destler, K., Player, D., 2010. Teacher labor markets and the perils of using hedonics to estimate compensating differentials in the public sector. *Econ. Educ. Rev.* 29 (1), 1–17. <https://doi.org/10.1016/j.econedurev.2009.07.010>.
- Gorard, S., Rees, G., 2002. Creating a Learning Society? Learning Careers and Politics for Lifelong Learning. Policy Press, Bristol.
- Gorard, S., Ventista, O., Morris, R., See, B.H., 2021. Who wants to be a teacher? Findings from a survey of undergraduates in England. *Educ. Stud.* <https://doi.org/10.1080/03055698.2021.1915751>.
- Gorard, S., 2013. Research Design: Creating Robust Approaches for the Social Sciences. SAGE, London.
- Gorard, S., 2021. How to Make Sense of Statistics. Sage, London.
- Gordon, N., Vegas, E., 2004. Education Finance Equalization, Spending, Teacher Quality and Student Outcomes: The Case of Brazil's FUNDEF. Education sector. Human Development Department, Latin America and the Caribbean Region, World Bank, Washington, DC.
- Harrell, P.E., Harris, M., 2006. Teacher preparation without boundaries: a two-year study of an online teacher certification program. *J. Technol. Educ.* 14 (4), 755–774.
- Hayes, P., 2017. Five Reasons for UK's Worst Ever Teacher Shortage. *Huffpost*. http://www.huffingtonpost.co.uk/patrick-hayes/teacher-shortage_b_9319692.html.
- Hazell, W., 2018. Senior DfE official: teacher supply problems “very severe” and getting worse. *Times Educ. Suppl.* <https://www.tes.com/news/senior-dfe-official-teachersupply-problems-very-severe-and-getting-worse>.
- Heinz, M., 2015. Why choose teaching? An international review of empirical studies exploring student teachers' career motivations and levels of commitment to teaching. *Educ. Res. Eval.* 21 (3), 258–297.
- Henry, G.T., Bastian, K.C., Smith, A.A., 2012. Scholarship to recruit the “best and brightest” into teaching: who is required, where do they teach, how effective are they and how long do they stay? *Educ. Res.* 41 (3), 83–92. <https://doi.org/10.3102/0013189X12437202>.
- Higher Education Policy Institute, 2017. Whither Teacher Education and Training. Higher Education Policy Institute, Oxford.
- Hobson, A., Giannakaki, M., Chambers, G., 2009. Who withdraws from initial teacher preparation programs and why? *Educ. Res.* 51 (3), 321–340. <https://doi.org/10.1080/00131880903156906>.
- Hough, H., Loeb, S., 2013. Can a District-Level Teacher Salary Incentive Policy Improve Teacher Recruitment and Retention? Policy Brief 13-4, Policy Analysis for California Education. Policy Analysis for California Education, Stanford University. https://cepa.stanford.edu/sites/default/files/PACE%20Policy%20Brief%2013-4_LowRes.pdf.
- House of Commons, 2017. Recruitment and Retention of Teachers. Fifth Report of Session 2016–17 (HC199). House of Commons, London, England.
- Jones, M.D., 2013. Teacher behavior under performance pay incentives. *Econ. Educ. Rev.* 37, 148–164.
- Kohn, A., 1993. Why incentive plans cannot work. *Harv. Bus. Rev.* Available <https://hbr.org/1993/09/why-incentive-plans-cannot-work>.
- Kohn, A., 1994. The Risks of Rewards. *ERIC Digest*. <https://www.vidyaonline.org/readings/pr42.pdf>.
- Kraft, M.A., Brunner, E.J., Dougherty, S.M., Schwegman, D.J., 2020. Teacher accountability reforms and the supply and quality of new teachers. *J. Public Econ.* 188, 104212.
- Matthias, C., 2014. Qualitative Research With Shortage Subject Teaching Candidates: The Journey to Teacher Training. Report for the National College for Teaching and Leadership. Department for Education, London, England.
- National Audit Office, 2016a. Training New Teachers. National Audit Office, London.
- National Audit Office, 2016b. Financial sustainability of schools. In: HC 850 Session 2016–2017. National Audit Office, London, England.

- Papay, J.P., West, M.R., Fullerton, J.B., Kane, T.J., 2012. Does an urban teacher residency increase student achievement? Early evidence from Boston. *Educ. Eval. Pol. Anal.* 34 (4), 413–434.
- Passy, J., 2018. Why America's teacher shortage is going to get worse. N. Y. Post. <https://nypost.com/2018/02/14/why-americas-teacher-shortage-is-going-to-get-worse/>.
- Public Accounts Committee, 2018. Retaining and Developing the Teaching Workforce, 17th Report of Session 2017–2019. HC 460. House of Commons, London, England.
- Rosen, R., 2012. Shortage-Field Incentives: Impacts in Teacher Retention and Recruitment. Ph.D. Thesis. Columbia University, New York, NY, USA.
- Rothstein, J., 2015. Teacher quality policy when supply matters. *Am. Econ. Rev.* 105 (1), 100–130.
- Schleicher, A., 2012. Preparing Teachers and Developing School Leaders for the 21st Century: Lessons From Around the World. OECD Publishing, Paris. <https://doi.org/10.1787/9789264xxxxx-en>.
- Scott, T.P., Milam, J.L., Stuessy, C.L., Blount, K.P., Bentz, A., 2006. Math and science scholars (MASS) program: a model program for the recruitment and retention of preservice mathematics and science teachers. *J. Sci. Teach. Educ.* 17 (4), 389–411.
- See, B.H., Gorard, S., 2020. Why don't we have enough teachers? A reconsideration of the available evidence. *Res. Pap. Educ.* 35 (4), 416–442.
- See, B.H., Morris, R., Gorard, S., Kokotsaki, D., Abdi, S., 2020a. Teacher recruitment and retention: a critical review international evidence of most promising interventions. *Educ. Sci.* 10, 262. <https://doi.org/10.3390/educsci10100262>.
- See, B.H., Morris, R., Gorard, S., El Soufi, N., 2020b. What works in attracting and retaining teachers in challenging schools and areas? *Oxford Rev.* 46 (6), 678–697. <https://doi.org/10.1080/03054985.2020.1775566>.
- See, B.H., 2011. Understanding Teacher Supply in England and Wales. LAP Lambert Academic Publishing, Saarbrücken.
- Shiffrer, D., Turley, R.L., Heard, H., 2017. Do teacher financial awards improve teacher retention and student achievement in an urban disadvantaged school district? *Am. Educ. Res. J.* 54 (6), 1117–1153.
- Sky News, 2017. Teacher Supply Crisis Hitting Schools in England. <https://news.sky.com/story/england-schools-suffering-from-teacher-supply-crisis-report-claims-10776066>.
- Springer, M.G., Taylor, L.L., 2016. Designing incentives for public school teachers: evidence from a Texas incentive pay program. *J. Educ. Finance* 41, 344–381.
- Springer, M.G., Lewis, J.L., Ehler, M.W., Podgursky, M.J., Crader, G.D., Taylor, L.L., Stuit, D.A., 2010. District Awards for Teacher Excellence (DATE) Program: Final Evaluation Report. Policy Evaluation Report. National Center on Performance Incentives, Nashville, TN, USA.
- Steele, J., Murnane, R., Willett, J., 2010. Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. *J. Pol. Anal. Manag.* 29 (3), 451–478. <https://doi.org/10.1002/pam.20505>.
- Strauss, V., 2016. The United States' growing teaching shortage: how it looks state by state. Washington Post. https://www.washingtonpost.com/news/answer-sheet/wp/2016/10/18/americas-growing-teaching-shortage-how-it-looks-state-by-state/?hpid=hp_hp-top-table-main-teachers%3Ahomepage%2Fstory&hpid=hp_hp-top-table-main-teachers%3Ahomepage%2Fstory&utm_term=.bb2d686bdb1e.
- Universities UK, 2013. Parliamentary Briefing: Initial Teacher Training. <http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2013/briefing-initial-teacher-training-nov-2013.pdf>.
- Waters-Weller, C.D., 2009. Attracting Veteran Teachers to Low Socioeconomic Status Schools: Initiatives and Considerations. Doctoral Dissertation. Walden University.
- Williams, A., 2018. Teacher Shortage in Mississippi [Web Log Post]. <http://www.wtok.com/content/news/Teacher-shortage-475850893.html>.
- Zarkin, G.A., 1985. The Importance of Economic Incentives in the Recruitment of Teachers. Final Report. Duke University, Durham, NC, USA.

Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms

A. Lin Goodwin^a, Seung Eun McDevitt^b, Crystal Chen Lee^c, and Sibel Akin-Sabuncu^{d,e}, ^aUniversity of Hong Kong, Hong Kong, HKSAR, China; ^bSt. John's University, Queens, NY, United States; ^cNorth Carolina State University, Raleigh, NC, United States; ^dTED University, Ankara, Turkey; and ^eTeachers College, Columbia University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Stories and studies of teachers of immigrant and refugee students: the case of Turkey	672
Facing a sudden influx of refugee students: the professional needs of teachers	672
Key learnings emerging from the findings	674
"Girl, you're smart": the story of one Latinx immigrant community-based educator's lived experiences	674
Lucia's case	674
"Girl, you're smart": cultivating and affirming assets of Latinx students	674
Informing her practices	675
"I am the first out of my family to go to college": using lived histories to prepare Latinx immigrant students for college and career readiness	675
Informing her practices	675
Key learnings of Lucia's lived experiences	676
From newcomers to welcomers: immigrant teachers in New York City	676
Choosing to be silent	676
Becoming "one of them"	676
Othered but empowered to unother	677
Key learnings	677
Discussion	677
360 degree belonging	678
Trans-culturally relevant pedagogy	678
Active unothering	679
Conclusion	679
References	679

The education of immigrant¹ students is a humanitarian crisis given forced displacement of 79.5 million people, 40% of whom are children (UNHCR, 2020a). In 2015, war drove a million plus refugees to Europe and Scandinavia, including thousands of unaccompanied minors (Connor, 2016). In the US, which "has more immigrants than any other country" (Radford, 2019, para. 1), one in four children under age 18 is an immigrant or a child of immigrants (Sugarman, 2019). Since 2011, 6.6 million people have fled Syria, escaping civil unrest (UNHCR, 2020a). Consequently, many countries are receiving newcomers who are challenging prevailing notions of teaching *well* and teaching *all*, in a world context where equitable education remains unrealized (OECD, 2018). Clearly educational interventions are needed.

These interventions are especially urgent given international data showing that immigrant students lag behind non-immigrant peers academically (APA and Presidential Task Force on Educational Disparities, 2012; OECD, 2016), report less sense of belonging, frequently attend segregated schools (OECD, 2015, 2018), and face numerous barriers, including insufficient language learning programs (Bartlett, 2015). We offer insight into teacher preparation for increasing diversity in classrooms, especially due to increasing newcomer students, through three mini case studies, each utilizing a distinct lens: teachers' lives; teachers' work; teachers' professional education. First is the case of Turkey which hosts the largest number of refugees worldwide (UNHCR, 2020d). The dramatic influx of refugee students into Turkish schools has accelerated local research on the issue, illuminating critical issues as well as teachers' professional development needs. The second case focuses on the work of community-based teachers of immigrant students. Here, both teachers and students are Latinx immigrants, affording a deeper understanding into immigrants from Latin America and Mexico who represent a sizable portion of immigrants globally (Noe-Bustamante and Lopez, 2019). Our third case examines the lives of immigrant teachers who originate from Asia, South America and the Caribbean but all teach in New York City, a destination city for newcomers to the US. These three cases tell a story of teacher preparation for immigrant students from the macro to the micro, from the vantage point of teachers, the students they serve, and the system in which they work. While each case offers recommendations for practice, our article concludes with a cross case analysis, enabling larger principles to emerge.

¹For simplicity, we use the term "immigrant" inclusively to encompass immigrants, migrants, and refugees.

In an effort to actively resist opportunity gaps, barriers to schooling, and teachers' pedagogical missteps that can sideline and other immigrant students, we frame our three mini cases within what we term, *Newcomer Unothering Pedagogy*. Schooling for immigrant children has long promoted conformity through culturally oppressive practices as opposed to liberatory, agentic transformation (Goodwin, 2010, 2020; Ladson-Billings, 2014; Yosso, 2005). "Whiteness as the 'invisibilized' norm against which difference is measured, constructed, and maintained" (Daniels and Varghese, 2019, p. 57) overtly and covertly enables the pathologizing and marginalizing, i.e., othering, of immigrant children by relegating them to the peripheral spaces in schools, obscuring their humanity, diverse histories and needs. *Unothering* is "critical action" (Goodwin, 2010, p. 3127) that consciously counters othering.

Newcomer Unothering Pedagogy (NUP) cultivates a sense of belonging as the essence of educating immigrant children, honoring the multiple identities they bring through their lived experiences and rich funds of knowledge (Moll et al., 1992). For newcomers, belonging is much more than simply fitting in or being included. Newcomers join an unrecognizable environment full of strangeness. For instance, newcomers may know the institution called school, yet it is likely that the norms–rules, regularities, organizational structures, languages and literacies, even the materialities of the school they enter is a version of which they have little familiarity. Teachers who enact NUP understand the multilayeredness of belonging, that newcomers are marked in countless ways as *alien*—from the clothes they wear to the way they comport themselves, and so aim consistently to redefine and deliberately structure and facilitate inclusion, community and belonging along those many dimensions.

NUP teachers direct their pedagogical practices toward the culturally relevant, sustaining and socially-just, centering immigrant children's "unique ways of knowing and integrating culturally meaningful content into the curriculum" (Goodwin, 2020, p. 11). These unique ways of knowing are culturally grounded, but also transnationally fluid as immigrant students begin to put down roots in their new community, but maintain—and sustain—ties to diasporic communities. Indeed, immigrant students often bridge disparate communities, not just "home" and school, but multiple homes and distant families, some of which may be in crisis. Dubois' concept of "double consciousness" (1903) is exponentially magnified, spanning place, time, histories, and familial ties, further adding complexity, trauma, and emotionality to any notion of identity or funds of knowledge.

NUP also exhorts teachers to become active resisters of othering pedagogies, to deliberately engage in unothering. Unothering "interrupt[s] colonizing discourses by abrogating or rejecting normativity in the (mis)representation, definition, and translation" of newcomer communities, perceived as different, deficient, and damaged (Goodwin, 2010, p. 3126). NUP exposes and counters othering that targets newcomers. A simple example of a common othering practice involves naming whereby newcomer names are often anglicized, mispronounced, or made the butt of jokes. NUP not only honors correct pronunciation of names—and works hard to learn them—but ensures all the other students and teachers in the school learn the same, while also speaking out about the importance of naming (respect, history, identity, belonging) for/of every child.

NUP transforms how immigrant children are viewed by seeing them as capable and powerful, by valuing all they bring, and by meeting them where they are academically, socially, and culturally through genuine relationships of care built among all members in their learning communities (Shah, 2013). Below, we examine stories of teachers of immigrant students in Turkey and the United States (North Carolina, and New York City) and their pedagogical stances through this lens of Newcomer Unothering Pedagogies to illuminate practices that invite newcomer students into learning such that they are "not passive bystanders ... but active participants" (Moll et al., 1992, p. 134), empowered to resist "the alienated condition of contemporary life" (Lave, 1991, p. 65) that seeks to discount them.

Stories and studies of teachers of immigrant and refugee students: the case of Turkey

Turkey, Syria's neighbor between Europe and Asia, has facilitated an open-door policy for refugees and asylum seekers (İçduygu, 2015), granting them legal "temporary protection" status (Kirişçi, 2014). However, as conditions continue to deteriorate and the number of arrivals increase tremendously, the positioning of Syrian refugees has changed from "guests" to "residents" (İçduygu, 2015; Karlı-Calamak and Kilinc, 2019; Kirişçi, 2014). Currently, Turkey hosts the largest number of refugees worldwide, almost 4.1 million, including 3.6 million Syrians and nearly 330,000 asylum-seekers and refugees of other nationalities (The UN Refugee Agency [UNHCR], 2020d). Specifically, over half of Syrian refugees are children (UNICEF, 2019). Of them, 684,919 Syrian children were enrolled in formal education as of March 2020 (UNHCR, 2020b), almost all of them in public schools (UNHCR, 2020c). Consequently, the Turkish education system is confronting massive flows of immigrant students, creating challenges for teachers and exerting pressure for policy changes to develop longer-term strategies that prioritize integration in education (İçduygu, 2015; Kirişçi, 2014).

Facing a sudden influx of refugee students: the professional needs of teachers

This study draws on journal articles published in Turkey between 2015 through 2019, a period of unprecedented global immigration (UNHCR, 2018) and the height of the refugee crisis in Syria, to offer a window into the knowledge and skills teachers need to meet the needs of immigrant students. A comprehensive search in both Turkish and English was conducted using relevant keywords, focusing on teaching and teacher education for immigrant students. To this end, 52 articles mentioning concepts related to migration, immigration, refugees, teachers, teacher candidates etc. were identified for a full read.

The 52 articles present a range of pedagogies for educating immigrant students in Turkey, along the continuum from othering to unothering practices. First, the findings from 35 articles suggested that a significant part of the pedagogies concerning the education of immigrant students involves certain instructional methods especially in bridging the language gap. Yet, most articles provided counter-examples of unothering pedagogies to teach immigrant students, rooted in a deficit-based approach. Accordingly, the most common challenges for teachers were their inability to communicate with refugee students/parents due to language barriers, their inability to provide psychological assistance to the students/parents facing depression and trauma due to the effects of war, migration, and economic difficulties (Aydin and Kaya, 2017; Aykırı, 2017; Erdem, 2017), lack of competence to employ linguistically-responsive instruction (Aydoğdu et al., 2019; Er and Bayındır, 2015), and limited access to relevant instructional materials (Biçer and Kılıç, 2017; Bulut et al., 2018). Moreover, teachers' incompetence in accommodating immigrant students' needs through relevant assessments, their inability to collaborate with parents of immigrant students (Morali, 2018; Saritaş et al., 2016), and their beliefs, dispositions, and conceptions of immigrants embedded in assimilationist perspectives (Başar et al., 2018) were among other serious issues. To illustrate, particularly highlighting language differences as a primary barrier for immigrant students to overcome, a teacher in the study by Aydin and Kaya (2017) reported that even when immigrant students are familiar with a given topic, their slow reading and comprehension puts them at a disadvantage in exams:

For example, we give exams. The student takes out his dictionary. He needs to look up the words, understand, and then answer the questions but by the time he does this the time is up. Thus, he leaves the dictionary. And answers the questions haphazardly. Then he fails. That means he needs to master Turkish first (p. 7).

Accordingly, the findings further showed that one issue related to supporting teachers to meet refugee students' needs was about the inadequacy and irrelevancy of pre-service teacher education. For instance, according to Arslangilay (2018), a newly graduated teacher candidate said "I don't have much knowledge, except that we read [about Syrian refugees] from newspapers," while another graduate highlighted the lack of relevant courses and practice-based opportunities in pre-service teacher education. On the other hand, some examples of the curricular pedagogies enacted by teachers to connect immigrant students to the curriculum included the integration of classroom materials and texts reflecting students' backgrounds and histories, employing specific second language instructional techniques (Aydoğdu et al., 2019), utilizing educational games or extracurricular activities to support immigrant students' language development, fostering language-focused instructional coaching, creating classroom interactions to support language learning (Kuzu-Jafari et al., 2018), developing professional collaboration between general education teachers and bilingual and/or immigrant teachers, fostering parent involvement (Bulut et al., 2018), and collaborating with others who can translate between Turkish and immigrant students'/parents' mother language (Aykırı, 2017). To illustrate, a teacher in the study conducted by Aykırı (2017) said:

We have been experiencing a lot of difficulties communicating with them [Syrian refugees]. Therefore, we try to communicate with parents [of Syrian refugee students] with the help of translators [other individuals at home/in school who can provide help in translation and build bridges between teachers and students/parents].

Translated by Akin-Sabuncu (p. 50).

Second, the findings from 15 articles addressed teachers' inability to extend beyond the teaching of individual learners in the classroom, to the ways in which teachers can build on what students bring to school with them, including their knowledge, interests, experiences, beliefs, and cultural and linguistic resources. Particularly, teachers struggled to adjust the curriculum and see immigrant students' transnational identities and experiences of living in multiple cultural worlds, families, communities as strengths (Başar et al., 2018; Solak and Celik, 2018; Tösten et al., 2017). The work of teachers is challenged by top-down and mandated K-12 curricula that do not directly include topics on immigrant students (Taskin and Erdemli, 2018), the pressure on teachers not to fall behind on the general curriculum (Börü and Boyacı, 2016), and lack of autonomy to implement alternative assessment and evaluation techniques in a pervasive testing culture. A teacher in the study by Börü and Boyacı (2016) remarked:

I teach this class, in which there are immigrant children, once a week. Since it is only one hour per week, I am obliged to implement the general curriculum and rush things until I finish everything. That's why I don't make any extra effort for those children.

As a result, in some cases, teachers advocated for assimilation following a top-down approach from education policies to "fix" the issues related to immigrant students. Although Başar et al. (2018) mention that there are also teachers who want to spend extra time with immigrant students in their classes, the challenges faced by teachers mostly result in their inability to educate immigrant students. Therefore, those students usually feel isolated in the class and lag behind their non-immigrant peers.

The remaining 2 articles offered little evidence for concrete steps for preparing teachers to privilege immigrant students' specific needs and making curriculum more culturally relevant (Zayimoğlu-Öztürk, 2018), or shift and develop thinking about immigrant students so that teachers would be able to face their own prejudices and change their perspectives (Gögebakan-Yıldız and Engin, 2019).

Key learnings emerging from the findings

The findings portray that Turkish teachers' current practices are not serving the needs of immigrants adequately and offer implications for pre-service and in-service teacher education to meet immigrant students' needs. Teachers feel unprepared and lack professional support to teach immigrant students and thereby, tend not to value their identities and backgrounds as a resource or strength to build on in the mainstream education context to accommodate their needs. Yet, this leads to the misrecognition of immigrant students' cultural and linguistic diversity in the classroom and perpetuates their position as passive and invisible "others," whose voices are not often heard or, at the best, lag behind their nonimmigrant peers. These problems are only aggravated when teachers are challenged by the unavailability of educational and economic resources and the presence of structural obstacles, political and social discourses in teaching immigrant students. Thus, given the expectation that classrooms of the future will be even more diverse, the findings call for pre-service and in-service teacher education to provide equal educational opportunities and employ pedagogies that would include immigrant students within the classroom and advocate against their marginalization and othering. To this end, teacher education needs to acknowledge that immigrant students bring a range of diverse cultures and experiences and thereby, must ensure that teachers develop various strategies to educate immigrants holistically by embracing their feelings, family, and community ties. The findings further signify that the education of a large number of immigrants, as in Turkey, requires systematic and organized collaboration among stakeholders to, first, bring in critical conversations about the educational needs of immigrant students to help teachers develop new perspectives. Thus, it is imperative to promote dialog, transformative actions for change, and shared responsibilities within and across state borders in today's global world to collectively dismantle unjust anti-immigrant practices and build on practices that are aligned with the tenets of social justice (Karsli-Calamak and Kilinc, 2019). Lastly, while Turkey's open-door policy has been universally appreciated, some criticisms have also been raised due to the legal uncertainties and administrative drawbacks; therefore, as Turkey does not have unlimited resources to accept and support further refugee flows (İçduygu, 2015), both Turkey and the international community should reevaluate their educational policies, including teacher education policy, utilizing a long-term perspective. To this end, they should consider immigrants' and refugees' protracted displacement and the likelihood of further arrivals and prioritize the educational policies for their integration while preparing for further refugee flows.

"Girl, you're smart": the story of one Latinx immigrant community-based educator's lived experiences

Unothering pedagogies are essential for formal schooling spaces in the United States as teachers rapidly pivot to meet the needs of growing numbers of immigrant students. Particularly, in the past decade, Latinx immigrants have represented a sizable portion of immigrants not just in the United States, but globally (Noe-Bustamante and Lopez, 2019). However, while schools are developing new pedagogies to teach immigrant students, immigrant educators in community-based organizations (CBOs) have long employed unothering practices within Latinx immigrant populations, such as centering Latinx students' lived histories and needs in education-based spaces. Specifically, Latinx immigrant community-based educators draw on their own immigration experiences and familial and community ties to inform how they engage and relate to immigrant students.

Lucia's case

The case of Lucia, a Latina community-based educator, elucidates what it means to teach immigrant students. The research summary draws from a 3-year in-depth study of 4 Latinx immigrant youth educators who work with adolescent Latinx immigrants in a CBO called Juntos NC² in North Carolina. Juntos NC works to unite community partners to provide Latinx 8–12th grade students and parents with the knowledge, skills, and resources needed for high school and college preparation. The data in this summary draw from Lucia's semi-structured interview, her published writing with Juntos NC students, and observations of her community-based work in Juntos NC. This research summary asks, "How do Lucia's lived experiences inform her educational community-based practices with immigrant students?"

"Girl, you're smart": cultivating and affirming assets of Latinx students

Lucia's case is one in which an immigrant educator's lived experiences informed how she centered immigrant students' lived histories and funds of knowledge. As a Mexican immigrant and a first-generation college graduate, Lucia's role as program director at Juntos NC was to create community and provide resources for Latinx high school students to belong, to succeed

²The study received institutional review approval to unblind the organization's name to honor the work of community-based staff. Lucia is a pseudonym.

academically, and to prepare for college-career readiness. Lucia's lived experiences of what she called "smartness" informed the ways in which she drew upon the assets of Latinx immigrant students. In her semi-structured interview, she told the story of her experiences as a high school student: "Honestly, I didn't really think that I would ever make it to college because I didn't think I was smart enough."

Lucia explained that she did not think she was "smart," because some teachers' expectations of Latinx students in her school were that they cannot go to college. Specifically, she remembered that teachers' impressions were either Latinx students were "not going to go" or "cannot go" to college. There were teachers who had not had much "experience working with [Latinx students]" and therefore did not "understand our culture." She also attributed a lack of understanding when it came to reading and writing. Lucia mentioned that as an ESL student, she really "struggled with writing and reading skills and in English." Even as Lucia improved, she lacked confidence in her literacy skills. This dearth of her teachers' understanding and ability to teach and encourage Latinx students who were ESL learners led her to believe that she was not "smart enough" for college. It was not until college that she began to realize, "Girl, you're smart."

Informing her practices

Lucia wishes that Latinx students could say, "you're smart" to themselves early on in high school. One way to do this is to build the confidence and assets of Latinx students. She modeled confidence building through cultivating the "smartness" in Juntos students by organizing and partnering with programs such as a literacy project that celebrated and developed students' voices and writing. Because she was "othered" by teachers as "not college ready," she used her own experiences to build bridges for teachers. For example, she traveled to high schools to host workshops for administrators and teachers about the assets of Latinx students. She regularly built relationships with school staff and community partners to develop valuable resources such as family nights, summer camps, and curriculum that informed Latinx students and parents about access to school resources and higher education. She has now made it her mission to cultivate, celebrate, and promote Latinx students' assets so that they could be "unothered" in formal schooling and moved from the margins of schooling spaces where she was once was.

"I am the first out of my family to go to college": using lived histories to prepare Latinx immigrant students for college and career readiness

Secondly, Lucia cultivated these assets because community educators had once cultivated hers. As the youngest of eight siblings, Lucia was the first to leave home and attend higher education: "Growing up, my parents did not understand the educational system ... I struggled the most when it came to the college process because I was not aware of the process." However, Lucia's trajectory changed when as a high school student, she connected with El Pueblo, a leadership development organization for both Latinx youth and adults. She went on a college tour to North Carolina State University (NC State) with El Pueblo and remembered seeing all the students and thinking, "Gosh, I'm just not smart. Honestly I thought college was not something that I could do and I wasn't smart enough to do it." However, after the tour, El Pueblo provided her with a mentor who helped her with the college admission process.

To her surprise, she got in. With a full ride. She was "shocked" and it felt "very surreal," but she still felt she was not "smart enough" and she was "not going to be smart enough." Lucia reflects on her experience by saying that "a lot of the students that I've worked with are in the same situation where they doubt themselves." Her reflection further emphasizes the need for teachers to encourage students and for community organizations to partner with students to provide them with the resources needed to pursue higher education.

Yet, Lucia's acceptance also came with hardships and her experience taught her that community-based educators must recognize the tensions of preparing students for college readiness. Though her admittance to NC State was filled with personal joy, her family felt that it was "almost like a hit to the heart; they just felt like I was betraying them because I was leaving." For her, the hidden curriculum of the college-going process was not only unknown to her, but in conflict with her parents' values and cultural strengths that were founded on familial ties. She worked through school to support her family while also building a career for herself, and over time, she was able to demonstrate to her family what her college education meant to her.

Informing her practices

Because a local community organization was able to bridge the gaps for Lucia to attend college, Lucia's experience motivated her to work in community organizations to help provide the resources and access to college for both Latinx students *and* their parents. Furthermore, her own college-going experience informed how she approached Latinx students *and* their families when preparing them for college-career readiness. Because she understood the cultural strengths and also the hesitations of college preparation, she was able to reach both families and students simultaneously to talk about the college-going process in culturally responsive ways. She unveiled the hidden curriculum of the college-going process by centering students' immigrant experiences and family values when discussing students' futures. Her approach reflects the culturally responsive practices that community educators and teachers must take to provide the resources "for them [students and parents] to understand the college process and to tell the students that [they] too can do it."

Key learnings of Lucia's lived experiences

Lucia's case is just one of many Latinx immigrant community-based educators like the three other stories within the larger study. Many, like Lucia, reflected on the power of knowing, believing, and hearing that they belonged. The importance of belonging—a central tenet of NUP—spoke to the necessity of “unothering” Latinx immigrant students. However, Lucia's lived experience as a Latinx immigrant student is profoundly impactful to her role as a program director who helps students and their families with the resources to be academically successful.

The deep look into Lucia's case demonstrates that immigrant teachers' lived experiences possess rich knowledge and experiences that influence how they perceive and enact educational practices (Goodwin and Kosnik, 2013). Lucia's case is particularly unique as she offers “funds of power” (Adair et al., 2012, p. 25) that can impact the assets of community-based organizations that support schooling institutions. Her unothering methods included (1) cultivating assets that honored and respected the multiple identities and voices of Latinx students, and (2) drawing upon her lived histories to be culturally responsive to her students' college-career readiness. Her experiences inform educators of the importance of being culturally responsive within classrooms, and even *within* communities. In doing so, immigrant stories from immigrant community educators have the power to humanize and to *unother* immigrant students' lives.

From newcomers to welcomers: immigrant teachers in New York City

New York City (NYC) is known as a destination city for many newcomer immigrants for centuries. Now, the city hosts newcomers mostly from Asia, and South and Central America, accounting for 44% of the entire city population, that includes 3 million immigrants in total (NYC Mayor's Office of Immigrant Affairs, 2020). Schools in NYC, thus, have a long history of educating immigrant children, boasting their diversities in their school profiles. More than half of the student demographic are either immigrants or children of immigrants who bring over 150 languages other than English to their learning environments (New York City Department of Education, n.d) with complex immigration histories and funds of knowledge. In responding to these incredible diversities, the presence of teachers who are culturally congruent and understand the multilayered contexts of immigrant students is crucial. However, little is known about teachers, who can be considered as “culturally-located insiders” (Goodwin, 2010, p. 3122), particularly those themselves who are immigrants.

In this research summary, immigrant teachers and their lived experiences are highlighted on issues around immigration and schooling. The following research questions are asked: What unique experiences do immigrant teachers bring to the education of immigrant students? How do their cultural, linguistic, and experiential competence shape their pedagogical approaches? In what follows, vignettes of three immigrant teachers³ are presented. Drawn from interviews and classroom observations, their stories illuminate the complexities and contradictions in their lives, meshed with their own immigration histories and schooling as newcomers and with their teaching of immigrant students as welcomers.

Choosing to be silent

Tuyet Nhi who immigrated with her parents from Vietnam described her initial feelings of being othered as a newcomer, “I was lonely, very lonely... and scared.” She shared how traumatic it was to be an immigrant child and learn the vulnerability of her parents in the foreign land. She mentioned witnessing her mother asking to change her grade level due to the interruption in her schooling but being completely ignored by school personnel. Since then and for a long time, Tuyet Nhi did not speak in school because she did not want to “bother [her] teachers” or perhaps she was afraid of being ignored by them.

Now entering her seventh year as a teacher of mostly Chinese immigrant children, she instinctively and intentionally uses pedagogical approaches of unothering. To prevent the perpetuation of harmful practices she once experienced as a newcomer, she pays extra attention to those who stay quiet and makes safe spaces to hear their interests and needs in both Chinese, “Chinglish (mix of Chinese and English),” and English. Tuyet Nhi stated that one of her teaching goals is to take bold and nontraditional actions to create a learning condition where immigrant children can be themselves for who they are and thrive while “having fun” in their schooling journey.

Becoming “one of them”

Al had layered immigration histories moving from Argentina to Israel as a child and then from Israel to the US as an adult. As an immigrant child in Israel, his way of survival was to become “more Israeli than Israelis.” He wanted to become “one of them” and to be accepted by his peers because he wanted to belong. Unlike his first immigration history, when he moved to NYC he struggled tremendously with not only his English language learning, but also deeper emotional issues, feeling as though he was “crippled.”

In his current classroom, he found his new sense of belonging, redefining his identity as a teacher of immigrant children. Using his immigration histories as a foundation to his pedagogy, he brings multimodal ways of learning, including music, arts, and film to provide multiple entry points to curriculum, enabling students who are new to the English dominant environment to express

³Tuyet Nhi, Al, and Gianna.

themselves through various modes of communication. Al emphasized that the core of working with immigrant students is to be able to “listen” to their funds of knowledge, to see them as “souls,” and to co-create an inclusive curriculum of unothering where belonging is a natural, fundamental part of their learning community.

Othered but empowered to unother

The struggle with learning a new language and not belonging were echoed by Gianna who had immigrated from the Dominican Republic as a mother to provide better living environments for her children. The transition was not an easy one. Her husband, who was an architect in the DR, became a taxi driver. Gianna also had to hold multiple jobs such as a gym locker room attendant and a maid to make ends meet. Later while working as an assistant teacher in her children’s school, she had an opportunity to attend graduate school to advance her career as a teacher. However, attending a university as a nontraditional immigrant woman came with many othering experiences. In addition to her language difference, she shared the many times she felt othered by professors who questioned her presence in the university. Yet, having worked in the classroom for many years already she was empowered to shine her strengths in her practicum courses by demonstrating her excellence in working with immigrant children and their families.

Gianna spoke about partnering with parents, who are low-income immigrants, and how they “repeat[ed] [her] stories.” She said, “Working with [immigrant] parents is my specialty” and noted that although they seem to be absent in their children’s education, they are involved in their education “at another level.” Reminiscing her early years as a newcomer working in multiple jobs to provide for her children, she confidently stated even though immigrant parents might not be able to “show up” to school events in traditional ways, they deeply care for their children’s education. And “the children know.”

Key learnings

In response to relatively little research on immigrant teachers, this research summary illustrated brief accounts of three immigrant teachers who were once newcomer immigrants and the ways in which their immigration and schooling stories shape their teaching of immigrant students. Privileging the voices of immigrant teachers this way enables us to see a glimpse of the complexity within their lives both as newcomer immigrants as well as teachers of immigrant students, bearing valuable insights into how schooling operates for and against immigrants (Elbaz-Luwisch, 2007b). It also provides a window into their teaching as compassionate welcomers of newcomer immigrant students and their families. Studying the lives of the immigrant teachers helps us understand their work from their own immigration and schooling histories, which laid the foundation for enacting a pedagogy that promotes inclusion, belongingness, and community for their immigrant students (Elbaz-Luwisch, 2007a,b).

In particular, the narratives of the immigrant teachers in this study made visible the nuances in teaching immigrant students. Entangled with their negotiation of identities and languages as well as the oppressive practices against immigrants in schooling, the three immigrant teachers teach us lessons in which understanding firsthand the experiences of immigrant students requires holistic and comprehensive views on who they are, their strengths as well as vulnerabilities. That is, teachers need to acknowledge and genuinely value immigrant students’ complex identities and their rich funds of knowledge, to assume them to be passionate knowers who are interested in learning and growing, and to consider their lived realities into curriculum in multiple and meaningful ways. Tuyet Nhi, for instance, demonstrated her pedagogical approaches by listening to her students through translanguaging practices. Al intentionally provided multiple ways for his students to express their learning and to co-create curriculum. For Gianna, teaching immigrant students went beyond the students themselves by respecting their families and valuing their nontraditional ways of being present in their children’s education. In doing so, the teachers were able to break the chain of the oppressive and exclusionary practices that have traditionally inhibited immigrant students from accessing inclusive and quality education (Goodwin, 2017). Rooted in their lived experiences of inclusion and exclusion and the trauma associated with being othered as newcomer immigrants, the teachers showed a strong desire to engage in undoing the othering processes by honoring students’ various language repertoires, creating learning conditions for them to be active knowledge contributors, and assuming the good in their families no matter how they appear to be superficially.

The method of inquiring into immigrant teachers’ lives offers important implications for teacher education for immigrant education. Adult learners like Gianna possess rich personal knowledge shaped by their own histories and narratives that influence how they perceive and enact teaching (Goodwin and Kosnik, 2013). Instead of rejecting their knowledge, teacher educators can employ autobiographical analysis (Genor and Goodwin, 2005) to incorporate their stories into teacher education curricula because their stories have value and are worth listening to, which in turn can inform teacher education about important social and political issues regarding immigration and schooling (Goodwin, 1997). Immigrant stories have the power to transform “the Curriculum of othering” (Goodwin, 2010, p. 3124) into the Curriculum of *unothering* in teacher education for the education of immigrants as a site of resistance and possibilities.

Discussion

The three cases we have presented each focuses on the education of immigrant students within a particular context. Turkey offers a macro view of a system overwhelmed by the sudden influx of large numbers of refugees who quickly re-position from temporary guest to more permanent resident. These “residents” bring with them thousands of school-aged children who require teachers to

educate them. NC and NYC both provide a more intimate look through the experience of teachers who are also immigrants—what unique knowledge do they hold about educating immigrant students given their insider perspective, either from within the formal educational system (NYC) or from outside it as community-based educators (NC)?

Our conceptualization of Newcomer Unothering Pedagogy provides an analytic lens for a cross-site analysis of immigrant student education across the three cases, enabling common themes and issues to come to the surface, but framed by a specific socio-political and country context. NUP centers on three tenets: (a) *360 degree belonging* that recognizes belonging as necessarily multi-dimensional and all-encompassing as newcomers add, shed, and transform identities as they move between realities; (b) *trans-culturally relevant pedagogy* that is culturally grounded but respects numerous and simultaneous transnational lived experiences across socio-historical and global boundaries; and (c) *active unothering* that names and combats any practice or policy designed to explicitly and implicitly marginalize and otherize immigrant students, and counters othering through deliberate unothering actions.

360 degree belonging

Engendering belonging for immigrant students requires fully comprehending their extent of loss—of the familiar, of home, of identity, of financial stability, of cultural comforts, of language and communication. Each of us wants to be able to stand out for reasons beyond our strangeness, as well as to be able to blend in with the everyday and with everyone. Belonging is a two-way process in which local inhabitants must be as prepared as newcomers to shift positions and make compromises, to bend and flexibly open up spaces so that *different* can find a niche, a comfortable space to enter, even as *typical* finds themselves transformed by their embrace of new.

Turkey tells a story of belonging where the two-way mutual process is still emerging, as newcomers continue to be expected to do the hard work of adapting and changing so as to become more like local residents and therefore less strange. Language difference is the key marker of (not) belonging, and governs whether teachers (can or choose to) communicate with students and their families, include immigrant children in the curriculum, and adapt instruction to ease their entry into Turkish schools. Ensuring Turkish teachers are able to engage NUP of belonging requires changes in teacher preparation, professional development and teacher education policy designed to enable those on the inside to see themselves as accountable for solving the “problem” of immigrant education. NC tells the story of learning from educators outside the formal educational system who have wider experience with NUP given their charge to build bridges of belonging for immigrant students using their lived experiences and cultural histories. As immigrants themselves with close ties to newcomer communities, their emic understanding of what it means—and what is needed—to create pathways for immigrant students to traverse the space from outsiders to a deep sense of belonging, is invaluable. However, the separations between school and *not*-school are wide, and the NC case urges us to embrace the wisdom already available in communities; community-based educators and families who hold funds of knowledge are waiting to be tapped and offer rich tools for bridge building. School insiders such as the immigrant teachers in the NYC case, also own rich funds of knowledge and are ready to be tapped. But they too remain invisible in plain sight. The NYC immigrant teachers’ practice is integrally informed by their personal stories of exclusion and each constructs their NUP of belonging from their own experiences, using their practice to counter contemporary othering that they also faced in their past as newcomers. Their stories reveal the complexity of exclusion, that marginalization is multi-layered and ongoing and the consequence of numerous actions, events, oversights that gain momentum over time. Thus NUP of belonging must be fully embracing so as to fully address 360 degrees of students’ lived experiences.

Trans-culturally relevant pedagogy

Identity and culture are never static, but fluid and ever-changing, states of being and thinking that, when privileged and honored, have the power to ignite and personalize learning, thus empowering disenfranchized learners (cf. Gay, 2018; Ladson-Billings, 1994; Pang, 2018). Who we are is shaped by our histories, familial ties, cultural roots, and when these are excluded from the formal curriculum, then our very spirit and essence is—we are—erased from the curriculum. The basic principles of culturally relevant pedagogy (CRP, Ladson-Billings, 1995) are salient in trans-culturally relevant pedagogy (TRP) with one significant difference: that the histories, familial ties, cultural roots, etc. of immigrant students are not just many and complex (as they are with each individual), but are further complicated and texturized by the simultaneous existence and influence, of inter-national, geo-political realities that are (sometimes viscerally) experienced in parallel in the lives of newcomers and embodied and enacted in their every day. TRP is, to put it simplistically, many-dimensional CRP.

Newcomer TRP requires this multi-dimensional understanding of “culture” even as it begins with coming to know newcomers as an imperative—and seeing them not just as different, but as worthwhile knowing. The case of Turkey demonstrates the need for preservice preparation and in-service PD to develop this understanding around knowing as a precursor to NUP through TRP, since the assimilationist stance remains apparent. But research reveals movement in this direction as teachers awaken to notions of teaching from the perspective of refugee children, vs. from the perspective of the state, set curriculum, or their own needs. Lucia in NC and the immigrant NYC teachers illuminate, through their stories, how country displacement, even by choice, results in adults/parents and their children walking the tightrope of who they were in “the old country”—educated, appropriately employed, facile communicators—and who they are forced to become in the new—students/children all over again as they relearn the basics and depend on their children to guide them, under-employed, often in menial jobs, silenced. TRP as one practice of NUP acknowledges and responds to the tension between identities that are trans-nationally located and span (often conflicting) socio-political boundaries, and how that tension affects agency, well-being and learning.

Active unothering

Active Unothering represents the DNA of NUP, because when teachers are able to see and interrupt unothering, they are able to enact the other essential tenets of NUP for immigrant students. Active Unothering requires the ability to see and unsee (Lewis, 2012), discern hidden realities that bring the wholeness, the “soul” of *Other* to the surface and in so doing resisting and rejecting the marginalizing identities they are assigned. Active Unothering is exemplified by Lucia’s insistence on unwrapping the mysteries of college admission to immigrant students, breaking through the pathologies that surround them and cause them to doubt their abilities; it is enacted in the “bold and non-traditional actions” Tuyet Nhi takes to ensure her immigrant students can “be themselves”; it is evident in general and ESL teachers in Turkey reaching out to one another to learn together about how to take positive action on behalf of their refugee students.

Conclusion

The “drastic increase in the number of immigrants ... caught many nations by surprise and left teachers poorly prepared for the changed composition of their classes” (Paine et al., 2016, p. 743; cf. European Commission, 2013; OECD, 2019). Educating teachers to successfully meet the needs of immigrant students in a world on the move is an imperative that cannot be ignored; the recent Afghanistan and Ukrainian crises heighten the issue. Teacher preparation must be designed to equalize educational opportunities and intervene in immigrant children’s “marginalization, denigration, segregation, and denial of competence” (Liu and Ball, 2019, p. 71). The three cases of immigrant students and teachers underscore the uniqueness of newcomer stories given differences associated with place, time, culture and immigrant histories, at the same time that these stories hold shared meanings across contexts. The US cases reveal the power of autobiography, that immigrant students’ lived experiences can become pathways into relationships, learning and belonging, especially when the teachers who nurture them can reach into their own newcomer stories to help create connections of-for knowing. The three cases together illustrate the complexity of the immigrant experience which in turn inform the nature and nuance of teacher preparation in response to specific needs. These cases are complementary, but are not comprehensive, and offer concrete enactments of Newcomer Unothering Pedagogies that take teacher preparation from understanding to action.

References

- Adair, J.K., Tobin, J., Arzubagi, A.E., 2012. The dilemma of cultural responsiveness and professionalization: listening close to immigrant teachers who teach children of recent immigrants. *Teach. Coll. Rec.* 114 (12), 1–37.
- APA (American Psychological Association), Presidential Task Force on Educational Disparities, 2012. Ethnic and Racial Disparities in Education: Psychology’s Contributions to Understanding and Reducing Disparities. Retrieved from: <http://www.apa.org/ed/resources/racial-disparities.aspx>.
- Arslangilay, S., 2018. Are Turkish teacher candidates ready for migrant students? *J. Educ. Learn.* 7 (2), 316–329.
- Aydin, H., Kaya, Y., 2017. The educational needs of and barriers faced by Syrian refugee students in Turkey: a qualitative case study. *Intercult. Educ.* 28 (5), 456–473.
- Aydoğdu, Z., Aydoğdu, Y., Asmaz, A., 2019. Yabancı dil olarak Türkçede kelime öğretimi üzerine bir durum tespiti: Suriyeli ortaokul öğrencilerine Türkçe öğreten öğretmenler örneği [A quantitative study on the vocabulary teaching of teachers teaching Turkish to Syrian secondary school students]. *Int. J. Teach. Turkish Foreign Lang.* 2 (2), 152–164.
- Aykın, K., 2017. Sınıf öğretmenlerinin sınıflarındaki Suriyeli öğrencilerin eğitim durumlarına ilişkin görüşleri [Determination of the experiences of social studies teacher candidates on the projects of community service practices]. *Turkish J. Prim. Educ.* 2 (1), 44–56.
- Bartlett, L., 2015. Access and Quality of Education for International Migrant Children. Paper Commissioned for the Education for All Global Monitoring Report 2015: Education for All 2000–2015: Achievements and Challenges. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000232474>.
- Başar, M., Akan, D., Çiftçi, M., 2018. Mülteci öğrencilerin bulunduğu sınıflarda öğrenme sürecinde karşılaşılan sorunlar [Learning-teaching process issues in classrooms with refugee students]. *Kastamonu Educ. J.* 26 (5), 1571–1578.
- Biçer, N., Kılıç, B.S., 2017. Suriyeli öğrencilere Türkçe öğretmek için kullanılan ders kitaplarının öğretmen görüşleri doğrultusunda değerlendirilmesi [Evaluation of textbooks which are used for teaching Turkish to Syrian students based on teacher opinions]. *Ana Dili Eğitimi Dergisi* 5 (4), 649–663.
- Börü, N., Boyacı, A., 2016. Göçmen öğrencilerin eğitim-öğretim ortamlarında karşılaştıkları sorunlar: Eskişehir ili örneği [Immigrant students’ problems in education—instruction processes (an example of the province of Eskişehir)]. *Turkish Stud.* 11 (14), 123–158.
- Bulut, S., Kanat-Soysal, Ö., Gülççek, D., 2018. Suriyeli öğrencilerin Türkçe öğretmeni olmak: Suriyeli öğrencilerin eğitiminde karşılaşılan sorunlar [Being a Turkish language teacher of Syrian refugee students: challenges and issues]. *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi* 7 (2), 1210–1238.
- Connor, P., 2016. Number of Refugees to Europe Surges to Record 1.3 Million in 2015. Retrieved from: <http://www.pewglobal.org/2016/08/02/number-of-refugees-to-europe-surges-to-record-1-3-million-in-2015/>.
- Daniels, J.R., Varghese, M., 2019. Troubling practice: exploring the relationship between whiteness and practice-based teacher education in considering a raciolinguized teacher subjectivity. *Educ. Res.* 49 (1), 56–63.
- Dubois, W.E.B., 1903. *Souls of Black Folks*. McClurg & Co, Chicago.
- Elbaz-Luwisch, F., 2007a. Immigrant teachers: stories of self and place. *Int. J. Qual. Stud. Educ.* 17 (3), 387–414.
- Elbaz-Luwisch, F., 2007b. Studying teachers’ lives and experience: narrative inquiry into K–12 teaching. In: Clandinin, D.J. (Ed.), *Handbook of Narrative Inquiry: Mapping a Methodology*. SAGE, pp. 357–382.
- Er, A.R., Bayındır, N., 2015. Pedagogical approaches of elementary teachers for primary refugee children [İlkokula giden mülteci çocuklara yönelik sınıf öğretmenlerinin pedagojik yaklaşımları]. *Int. J. Soc. Educ. Sci.* 2 (4), 175–185.
- Erdem, C., 2017. Sınıfta mülteci öğrenci bulunan sınıf öğretmenlerinin yaşadıkları öğretimsel sorunlar ve çözüme dair önerileri [Instructional problems experienced by primary school teachers who have refugee students in their classes and their solutions for problems]. *Medeniyet Eğitim Araştırmaları Dergisi* 1 (1), 26–42.
- European Commission, 2013. Supporting Teacher Competence Development for Better Learning Outcomes. Retrieved from: http://ec.europa.eu/dgs/education_culture/repository/education/policy/school/doc/teachercomp_en.pdf.
- Gay, G., 2018. *Culturally Responsive Teaching: Theory, Research and Practice*. Teachers College Press, NY.

- Genor, M., Goodwin, A.L., 2005. Confronting ourselves: using autobiographical analysis in teacher education. *N. Educ.* 1 (4), 31–331.
- Göğebakan-Yıldız, D., Engin, G., 2019. Migration education program for teacher candidates [Öğretmen adaylarına yönelik göç eğitim programı]. *Turkish Online J. Qual. Inq.* 10 (2), 180–214.
- Goodwin, A.L., Kosnik, C., 2013. Quality teacher educators = quality teachers? Conceptualizing essential domains of knowledge for those who teach teachers. *Teach. Dev.* 17 (3), 334–346.
- Goodwin, A.L., 1997. Multicultural stories: preservice teachers' conceptions of and responses to issues of diversity. *Urban Educ.* 32 (1), 117–145.
- Goodwin, A.L., 2010. Curriculum as colonizer: (Asian) American Education in the Current U.S. Context. *Teach. Coll. Rec.* 112 (12), 3102–3138.
- Goodwin, A.L., 2017. Who is in the classroom now? Teacher preparation and the education of immigrant children. *Educ. Stud.* 53 (5), 433–449.
- Goodwin, A.L., 2020. Learning to Teach Diverse Learners: Teachers and Teacher Preparation in the U.S. Oxford Research Encyclopedia of Education. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.913>.
- İçduygu, A., 2015. Syrian Refugees in Turkey: The Long Road Ahead. Migration Policy Institute, Washington, DC.
- Karslı-Calamak, E., Kilinc, S., 2019. Becoming the teacher of a refugee child: teachers' evolving experiences in Turkey. *Int. J. Incl. Educ.* 1–24.
- Kirişçi, K., 2014. Syrian Refugees and Turkey's Challenges: Going Beyond Hospitality. Brookings, Washington, DC.
- Kuzu-Jafari, K., Tonça, N., Kışla, H., 2018. Suriyeli öğrencilerin bulunduğu sınıflarda görev yapan sınıf öğretmenlerinin görüşleri ve uygulamalar [Practice and opinions of the teachers who work at the classes have Syrian students]. *Acad. J. Educ. Sci.* 2 (2), 134–146.
- Ladson-Billings, G., 1994. The Dream-Keepers: Successful Teachers of African American Children. Jossey Bass, San Francisco, CA.
- Ladson-Billings, G.J., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 35, 465–491.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: A.k.a. the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lave, J., 1991. Situating learning in communities of practice. In: Resnick, L., Teasley, S. (Eds.), *Perspectives on Socially Shared Cognition*. APA, Washington, DC, pp. 63–82.
- Lewis, T.E., 2012. The Aesthetics of Education: Theatre, Curiosity, and Politics in the Work of Jacques Ranciere and Paulo Freire. Bloomsbury, New York.
- Liu, K., Ball, A., 2019. Critical reflection and generativity: toward a framework for transformative teacher education for diverse learners. *Rev. Res. Educ.* 43 (1), 68–105.
- Moll, L., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theory Into Pract.* 31 (2), 132–141.
- Moralı, G., 2018. Suriyeli mülteci çocuklara Türkçenin yabancı dil olarak öğretiminde karşılaşılan sorunlar [The problems faced in teaching Turkish as a foreign language to Syrian refugee children]. *OPUS Uluslararası Toplum Araştırmaları Dergisi* 8 (15), 1426–1449.
- New York City Department of Education, (n.d). report2018–2019 English Language Learner Demographic Report: New York City Department of Education Division of Multilingual Learners. Retrieved from. <https://infohub.nyced.org/docs/default-source/default-document-library/ell-demographic-report.pdf>.
- New York City Mayor's Office of Immigrant Affairs, 2020. State of Our Immigrant City: Annual Report March 2020. Retrieved from. <https://www1.nyc.gov/assets/immigrants/downloads/pdf/MOIA-Annual-Report-for-2020.pdf>.
- Noe-Bustamante, L., Lopez, M., 2019. Latin America, Caribbean No Longer World's Fastest Growing Source of International Migrants. Pew Research Center. Retrieved from. <https://www.pewresearch.org/fact-tank/2019/01/25/latin-america-caribbean-no-longer-worlds-fastest-growing-source-of-international-migrants/>.
- OECD (Organisation for Economic Co-operation and Development), 2015. Immigrant Students at School: Easing the Journey Towards Integration. OECD Publishing, Paris. <https://doi.org/10.1787/9789264249509-en>. Retrieved from.
- OECD, 2016. Supporting Teacher Professionalism: Insights From TALIS 2013. OECD Publishing, Paris. <https://doi.org/10.1787/9789264248601-en>. Retrieved from.
- OECD, 2018. Equity in Education: Breaking Down Barriers to Social Mobility. OECD Publishing, Paris. <https://doi.org/10.1787/9789264073234-en>. Retrieved from.
- OECD, 2019. TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. OECD Publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>. Retrieved from.
- Paine, L., Blömeke, S., Aydarova, O., 2016. Teachers and teaching in the context of globalization. In: Gitomer, D., Bell, C. (Eds.), *Handbook of Research on Teaching*. AERA, Washington, DC, pp. 717–786.
- Pang, V.O., 2018. Diversity and Equity in the Classroom. Cengage Learning, Boston.
- Radford, J., 2019. Key Findings About U.S. Immigrants. Pew Research Center. Retrieved from. <http://www.pewresearch.org/fact-tank/2018/11/30/key-findings-about-u-s-immigrants/>.
- Sarıtaş, E., Şahin, Ü., Çatalbaş, G., 2016. İlkokullarda yabancı uyruklu öğrencilerle karşılaşılan sorunlar [Problems faced with foreign students in primary schools]. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 25 (1), 208–229.
- Shah, A.R., 2013. Exploring Hybrid Identities: South Asian American Women Pursue a Career in Teaching (Publication No. 3565344) [Doctoral Dissertation]. Columbia University, ProQuest Dissertations Publishing, Teachers College.
- Solak, E., Çelik, S., 2018. Türkiye'de eğitim gören mülteci öğrencilerin dilsel sorunlarının incelenmesi [Investigation of linguistic challenges faced by refugee students in Turkey]. *Uluslararası Sosyal Araştırmalar Dergisi* 11 (57), 425–432.
- Sugarman, J., 2019. Legal Protections for K–12 English Learner and Immigrant-Background Students. Migration Policy Institute, Washington, DC.
- Taskin, P., Erdemli, O., 2018. Education for Syrian refugees: problems faced by teachers in Turkey [Suriyeli öğrencilerin eğitimi: Türkiye'de öğretmenlerin karşılaştığı sorunlar]. *Eur. J. Educ. Res.* 75, 155–178.
- Tösten, R., Toprak, M., Kayan, M.S., 2017. An investigation of forcibly migrated Syrian refugee students at Turkish public schools. *Univ. J. Educ. Res.* 5 (7), 1149–1160.
- UNHCR (UN Refugee Agency), 2018. Syria Emergency. Retrieved from. <https://www.unhcr.org/syria-emergency.html>.
- UNHCR, 2020a. Figures at a Glance. Retrieved from. <https://www.unhcr.org/en-us/figures-at-a-glance.html>.
- UNHCR, 2020b. Turkey Education Sector Achievements as of May 2020. Retrieved from. <https://reliefweb.int/sites/reliefweb.int/files/resources/77119.pdf>.
- UNHCR, 2020c. Turkey Education Sector Dashboard. Retrieved from. <https://reliefweb.int/sites/reliefweb.int/files/resources/75200.pdf>.
- UNHCR, 2020d. Turkey Operational Update. Retrieved from. <https://reliefweb.int/sites/reliefweb.int/files/resources/UNHCR-Turkey-Operational-Update-November-FINAL.pdf>.
- UNICEF, 2019. UNICEF Turkey Humanitarian Situation Report #35. Retrieved from. <https://reliefweb.int/report/turkey/unicef-turkey-humanitarian-situation-report-35-july-sept-2019>.
- Yosso, T., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Zayımoğlu-Öztürk, F., 2018. Mülteci öğrencilere sunulan eğitim-öğretim hizmetinin sosyal bilgiler öğretmen görüşlerine göre değerlendirilmesi [Evaluation of educational service provided for refugee students according to social studies teachers' opinions]. *Anadolu J. Educ. Sci. Int.* 8 (1), 52–79.

Research for socially progressive teacher education programs

Aileen Kennedy, University of Strathclyde, School of Education, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	681
Conceptualizing socially progressive teacher education	681
Social progressivism as a political movement	682
Pedagogical progressivism as an educational movement	683
Social and environmental justice as ethical imperatives	684
Toward a working definition of socially progressive teacher education	686
In conclusion	687
References	688

Introduction

The idea of teacher education for social justice, or social justice-oriented teacher preparation, is common in both research literature and in program narratives across the globe. However, this has long been criticized as something so ill-defined that it is difficult to move beyond rhetoric and into reality. Cochran-Smith (2010) argues that “few who write about teacher education and social justice make the underpinning political and philosophical roots explicit, and that this increases the likelihood that it [social justice] exists in name only or that it is diluted, trivialized or co-opted” (p. 445). In this article, I want to propose that “socially progressive teacher education” might hold promise as an alternative, wider framing that is defensible conceptually and also capable of being applied in practice. My interest in seeking to conceptualize socially progressive teacher education is located ethically in my experience as what I would term a social justice educator, and practically in my current role as Director of Teacher Education within a university which promises “a relentless commitment to pursuing a globally socially progressive vision” (University of Strathclyde, 2020, p. 6).

Just as my own university has made its underpinning philosophy explicit, so too argues Darling-Hammond (2006), do what she calls “powerful teacher education programs, defined as:

Programs that prepare teachers to teach a wide range of students successfully, including those who struggle to learn from their first days in the classroom. These are programs whose graduates are sought out by principals and superintendents because they prove consistently capable of creating successful classrooms and helping lead to lead successful schools, even in circumstances where the deck is traditionally stacked against student success” (p. 27).

Here, “powerful” has a particular meaning: it is not simply powerful in terms of having capacity to influence, rather, that influence is specifically located in terms of a social justice agenda. So while “effective” teacher education programs might beg the question “effective at what?”, Darling-Hammond’s notion of powerful programs makes explicit the philosophical underpinning of the way she understands “powerful”, from a critical theory perspective. In the same vein, I draw on a range of research in seeking to make explicit the theoretical and philosophical underpinnings of socially progressive teacher education, while at the same time illuminating what this might look like in practice.

In striving to make philosophical underpinnings explicit, the article outlines three distinct, but linked ideas that influence how socially progressive teacher education might be understood and practiced: social progressivism as a political ideology; pedagogical progressivism as an educational movement; and social and environmental justice as ethical imperatives. The argument is founded on an acknowledgment that education in general, and teacher education in particular, cannot ever be neutral, despite some claims to the contrary.

Conceptualizing socially progressive teacher education

Like many terms in education, “socially progressive teacher education” may well mean different things to different people; it is not easy to define, just as social justice-oriented teacher education is similarly “undertheorized and vague” (Cochran-Smith et al., 2009, p. 347). The appeal to “progress”, however, is undoubtedly a persuasive one, yet progress is always a value-bound process, and one person’s progress may well be another’s idea of deterioration. I therefore want to draw on existing definitions of the three linked ideas mentioned above, as well as foregrounding issues of values, in exploring how the concept of socially progressive teacher education might serve to frame the development and delivery of ethically sound, contemporary teacher education programs.

It seems that there are three broad ideas which coalesce around the concept of socially progressive teacher education, namely, social progressivism as a political movement, pedagogical progressivism as an educational movement, and social and environmental justice as ethical imperatives (Fig. 1)

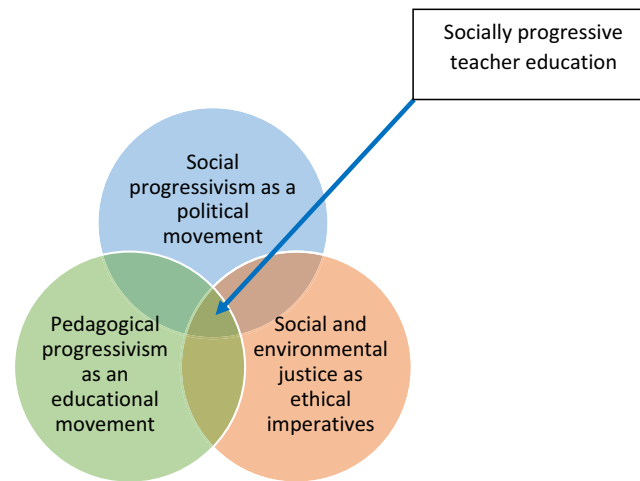


Fig. 1 Conceptualizing socially progressive teacher education—three overlapping ideas.

These three ideas loom large in any discussion of socially progressive teacher education, yet in many respects they stand quite separate from each other. In this **Introduction** section I want to outline how each idea individually can inform thinking about socially progressive teacher education, before exploring how the three ideas might interact.

Social progressivism as a political movement

Social progressivism is commonly considered to be a political philosophy which supports social reform for the good of society at large (as opposed to for particular groups or individuals). The “social” part of social progressivism acknowledges that change over time should be to the benefit of society, rather than simply change for change’s sake. The “progressivism” part connotes an acknowledgment that change is necessary in order to achieve social good, and that, therefore, reform is key to progressivism. This stands in opposition to social conservatism which relies on the status quo as a means of preserving what are deemed to be traditional social values.

Social progressivism rejects appeals to tradition and status quo, arguing that society should change and develop as new knowledge and understanding become available. The particular focus of socially progressive reform is in exposing root cause of things that are deemed to make aspects of society less than fair, and then to push through reform that will work against these identified barriers. There is, therefore, an innate commitment to what might be called social justice. However, it is not a radical or critical approach to social injustice as might be advocated by social justice proponents, rather, social justice is a driving force for incremental reform; it serves as a compass for decision-making.

Social progressivism does not locate its interest in any one particular cause, rather, it is about moving toward reform in a range of areas, with the common theme being that of improvements in society through challenging out-moded attitudes and practices particularly in relation to gender, race identity and power. Crucially, to be socially progressive means to be able to develop and change one’s own attitudes over time in light of new knowledge. This has clear implications for teacher education: it acknowledges explicitly the impossibility of an initial teacher education experience which could prepare a teacher for their whole career. Instead, it suggests an initial teacher education experience which prepares a teacher to be a career-long professional learner. Logically, then, it follows that part of socially progressive initial teacher education must be about supporting teacher candidates to be able to know how to continue their own learning, and to make autonomous choices about what and when to learn. This stands in direct contrast to what [Freire \(1970\)](#) calls the “banking model of education”, where the teacher “deposits” knowledge into the student, thereby rendering the student agency-less “in which the scope of action allowed to the students extends only as far as receiving, filling, and sorting the deposits” (p. 72). So, socially progressive teacher education requires a range of different types of knowledge, and the requisite skills of critical analysis and synthesis in order to make judicious choices about the rigor and appropriateness of what might be presented as “knowledge”.

While social progressivism relies on reform based on new knowledge, there is an inherent danger that such new knowledge, or evidence, might be adopted uncritically. For example, the current focus on “what works”, particularly prominent in England, with the concomitant growth in the promotion of randomized controlled trials (RCTs) runs the risk of only certain types of new knowledge being deemed reliable. While there remains considerable debate to be had about the relative merits of RCTs over other more qualitative research ([Connelly et al., 2018](#)), the fact remains that “knowledge” itself is not neutral, nor equally rigorous and reliable in its creation, and that socially progressive teacher education needs to support pre-service teachers to develop strong research literacy, particularly as consumers of research.

Key to social progressivism is that new insights should lead to structural and cultural reform that supports a fairer society: it is not sufficient simply to know that things might not be working in everyone’s interests, or that certain groups are

disadvantaged by particular practices, rather this new knowledge should lead to reform. Thinking about this in terms of teacher education it is possible to see how reform to the structures and governance of professional regulation might be used to work for the common good. However, governance of teacher education, which in Scotland is shared by the Scottish Government, the university sector and the General Teaching Council for Scotland (GTCS), is notoriously bureaucratic and conservative (Hulme and Kennedy, 2016). The homogenous nature of teacher education programs, all of which are university-based, and all of which are accredited by the GTCS following a fairly prescriptive set of criteria (GTCS, 2019a,b), means that responding to new knowledge or insights and amending programs accordingly is very difficult, except in relation to minor tweaks. While the existence of these sets of governance procedures arguably serves to maintain teacher education quality in Scotland (Hamilton, 2018), it can also result in teacher education programs having to make significant compromises in order to satisfy particular rules (Kennedy, 2018).

While Scotland's teacher education is governed centrally and is university-led (MacDonald and Rae, 2018), in other countries, for example nearby England, there is much more of an appeal to market-led governance and school-led initial teacher education. Mincu and Davies (2021) argue that the system in England is governed by a complex interaction of markets, networks and hierarchies, but these national systems of governance do not work in vacuums; it is important to recognize the increasingly powerful influence of global politics on national systems. For example, Cobb and Couch (2018) argue that the OECD's move to assessing "global competence" in PISA has had a significant impact on initial teacher education, serving to provide what they call "global pedagogic governance" (p. 41). What Cobb & Couch (Cobb and Couch, 2018) go on to propose is that pre-service teachers should be exposed to thinking about globalization, and how broader influences impact on schooling, and indeed on teacher education. This knowledge and understanding, they argue, "can provide pre-service teachers with the capacity to identify, and the ability to resist and transform, changes to education policy, structures and curricula" (p. 43). Thus, it seems fair to suggest that the centrality of a critical appreciation of new knowledge that is bound up with social progressivism, arguably requires teacher education students to develop skills of critical consumption of research together with an appreciation of the world beyond their immediate domain.

In summarizing the relationship between social progressivism as a political ideology, and socially progressive teacher education, there seem to be a number of key messages:

- The ultimate aim of social progressivism is to facilitate a fairer and more equal society;
- Our understanding of the barriers to achieving this aim, and the potential solutions at our disposal, develop over time, and therefore new knowledge and insights are central to driving socially progressive action;
- This focus on action is enacted through progressive and incremental reform which does not give time to sentimental appeal to tradition;
- In order to understand the potential contribution of new understanding to reform in teaching and teacher education, both teacher educators and pre-service teachers must develop their skills in critical literacy, particularly pertaining to the consumption of research;
- Socially progressive teacher education programs should eschew the "banking" notion of education and instead seek to empower teacher candidates as lifelong professional learners.

Pedagogical progressivism as an educational movement

It seems natural when confronted with the term socially progressive teacher education to think about the educational movement called "progressivism". In this **Introduction** section I want to explore how some of the key tenets of pedagogical progressivism might help us to define socially progressive teacher education, but also consider perhaps where some of the challenges might be in terms of mapping these two ideas on top of each other.

It is important, first, to be clear about what pedagogical progressivism is Labree (2005), writing from a US perspective, distinguishes usefully between what he terms "administrative progressivism", that is, the reform of structures and organization, and "pedagogical progressivism". He summarizes pedagogical progressivism succinctly, saying that "In the shorthand of educational jargon this adds up to 'child-centered instruction', 'discovery learning' and 'learning how to learn' (Labree, 2005, p. 277)." These elements, which draw deeply on the work of John Dewey, will be very familiar to educators in Scotland (and elsewhere), being the philosophical tenets of the current iteration of curriculum in Scotland (Thorburn, 2017). However, these ideas tend to be applied to the education of school pupils, rather than of teachers, or other professionals, perhaps due to the prescriptive nature of governance associated with professional education, as discussed earlier.

We might also wish to consider the extent to which pedagogical progressivism actually exists in practice, as opposed to merely in policy or academic texts. For example, Labree (2005) contends that while pedagogical progressivism might dominate education discourse in the US, it has yet to make significant impact on the classroom:

Instruction in American schools is overwhelmingly teacher-centered; classroom management is the teacher's top priority; traditional school subjects dominate the curriculum; textbooks and teacher talk are the primary means of delivering this curriculum; learning consists of recalling what texts and teachers say; and tests measure how much of this students have learned. (p. 278)

This issue of differences between rhetoric and reality are not, however, confined to the US. Some time ago, [Hartley \(1987\)](#) provided an articulate analysis of some of the reasons why progressive, child-centered curriculum policies of the 60s and 80s were never fully enacted in Scotland, citing as one reason, the economic imperatives which held dear to more traditional knowledge-focused understandings of children's learning. Similar to [Labree \(2005\)](#), [Hartley \(1987\)](#) too points to research which suggested that "the backlash against progressivism was largely misplaced: the schools had, throughout the period, remained places where didactic pedagogy had prevailed" ([Hartley, 1987](#), p. 119). [Hartley \(1987\)](#) went on to argue that progressivism requires a particular cultural and economic climate in order to thrive. It is therefore not enough for educationists to believe that progressive ideals provide the best way for children and young people to learn; educational philosophies and approaches do not exist in a vacuum and their successful integration into practice requires a favorable wider social, economic and, importantly, political context.

There are many critiques of the extent to which the discourse of pedagogical progressivism can be translated into practice, but there are also more fundamental philosophical critiques over its desirability. Such critiques center on the debate between valuing educational process over product, the role and the place of subject or disciplinary knowledge in education vis-à-vis interdisciplinary or place-based approaches ([Webber and Miller, 2016](#)). These arguments, while usually situated within the school education context, can also be applied to teacher education.

Concerning the argument that the progressivist emphasis on skills fails to value disciplinary knowledge, [Ashman \(2017\)](#) argues that there is little evidence that generic capacities to solve problems can be applied in all contexts, rather he argues that "the same person can think critically about an area where she possesses expertise and yet fail to think critically in an area where she lacks expertise" (no page number). This leads us to consider the merits of interdisciplinary learning, defined as integrating "knowledge derived from different disciplines to construct new perspectives on a particular problem or scenario" ([Stentoft, 2017](#), p. 53). Disciplinary knowledge is important in this context, and [Stentoft \(2017, p. 53\)](#) argues that the integration of disciplinary knowledge "may lead to new disciplinary insights". So, rather than watering down disciplinary knowledge, an interdisciplinary approach not only relies on discrete disciplinary knowledge but also has the capacity to elevate it to new heights. It seems reasonable to argue then that the potential criticism of pedagogical progressivism undervaluing disciplinary knowledge is not a well-founded position where a balance of disciplinary and interdisciplinary can be achieved.

There have also been concerns raised over the potential for the learner-centered nature of progressive education, with its emphasis on self-directed learning, to result in an over-emphasis on the individual at the expense of the collective, thereby failing to live up to its emancipatory promise ([Servant-Miklos and Noordegraaf-Eelens, 2019](#)). However, pedagogical progressivism is not always entirely positioned as an individualistic philosophy, and its emphasis on problem-solving in context suggests a need to work collaboratively with others. However, when juxtaposed alongside socially progressive political ideology and a concomitant emphasis on social and environmental justice, the focus on self-directed learning can be seen within the wider social context as opposed to something that is entirely individualistic and insular. Indeed, when seen as part of this three-pronged conceptualization, it seems that pedagogical progressivism has much to offer in terms of guiding a pedagogical approach that fits with socially progressive teacher education.

Social and environmental justice as ethical imperatives

If we agree that social progressivism focuses on the use of new knowledge and understanding to support reform for the purpose of social good, then there is a need to interrogate what might be meant by "social good". While it is difficult to see what arguments could be brought to bear against aspiring to do social good, the term itself is arguably rather vague, and, in common with the discussion around purposes of education, is open to multiple interpretations of what constitutes "good". Social good arguably attends to both human and environmental concerns, given that these are inextricably linked. It therefore makes sense that, notwithstanding the challenges in coming to agreed definitions and conceptualizations, social and environmental justice would feature as ethical imperatives within socially progressive teacher education.

A commitment to social justice is written into the aims of pretty much every teacher education program across the developed world, but the ambiguity of the concept increases the likelihood that it exists in the rhetoric only, and not in reality ([Cochran-Smith et al., 2009](#)). One of the difficulties here is that the term can be understood in many different ways, and this is not always fully explored and articulated across the development and delivery of teacher education programs. For example, social justice can be understood in terms of the redistribution of goods and resources, with a focus principally on economic concerns, and/or it can be understood as an issue of recognition ([Fraser, 2000](#)), which tends to focus on cognition struggles ([Benjamin and Emejulu, 2012](#)) rather than economic concerns. However, much of the conceptual work focusing on the recognition aspect of social justice tends to focus on the recognition of under-represented or minoritized groups, arguably sometimes missing the need for recognition of the individual, as highlighted by [Fleming \(2016\)](#) in his work on transformative learning theory. [Fleming \(2016, p. 6\)](#), drawing on Honneth, argues that "the struggle for recognition drives social change", thereby suggesting a link between social progressivism and the politics of recognition.

Notwithstanding the challenges in defining social justice, and articulating that conceptualization in program development, what is abundantly clear in the literature is that social justice cannot be taught effectively through discrete inputs, rather it has to permeate the program as a whole, through both on-campus and field-based elements ([Cochran-Smith et al., 2009](#)). Unfortunately, however, while research on social justice teacher education is a growing field, it has so far failed to make a significant impact on policy due to its fragmented and often small-scale nature ([Mills and Ballantyne, 2016](#)). This is of concern when we consider the earlier discussion

around the fairly inflexible and centralized governance of teacher education, and the concomitant challenges in seeking to progress social reform.

Cochran-Smith (2010) proposes that while activities in teacher preparation, such as fieldwork in diverse contexts and specific targeted readings, are important to the development of pre-service teachers' socially just practice, they are insufficient in themselves. Rather, she calls for "a coherent and intellectual approach to the preparation of teachers that acknowledges social and political contexts in which teaching, learning schooling, and ideas about justice have been located historically as well as acknowledging the tensions among competing goals" (p. 447). There is also some evidence to suggest that where programs seek to develop pre-service teachers' conceptual understanding, this work is best located within the specific classroom context for which they are preparing (Mills and Ballantyne, 2016).

As Mills and Ballantyne (2016) point out, though, much of the literature on teacher education for social justice focuses on changing the beliefs of pre-service teachers, "and 'blame' is, therefore, placed on the students for any deficiencies noted, rather than critically exploring the pedagogies and philosophies espoused by teacher educators" (p. 275). Goodwin and Darity (2019), however, argue that we need to go beyond teacher educators "espousing", and move instead to "enacting social justice in their own teaching and teacher educator practice" (p. 64). This, they argue, requires a concentrated focus on understanding what teacher educators need to know, thereby returning to fundamental underpinnings of social progressivism relating to the need to draw on knowledge in order to progress reform for social good.

Despite the plethora of writing which explicates concepts of social justice, and the smaller body of literature which interrogates social justice practice in teacher education, there somehow remains a chasm between theoretical understanding in the literature and the capacity to apply that theory systematically to practice in teacher education. Mills and Ballantyne (2016) point out that "much research in the field of social justice and teacher education focuses on an understanding of beliefs of preservice teachers and/or teacher educators, with limited exploration into what programs that prepare preservice teachers to engage with student diversity in socially just ways might look like in practice" (p. 263). While it is, of course, important to understand teachers' beliefs, this understanding needs to be put to use in terms of shaping and researching teacher education programs that make a difference. Aronson et al. (2020) contribute toward addressing this gap in understanding by identifying a number of specific theoretical perspectives that they have drawn on in their own program development work, namely: critical pedagogy; democratic education; critical race theory and critical whiteness studies; critical disability studies; and feminist theories. Importantly, they identify collectively, as a group of teacher educators, how they actually use these perspectives to inform practice; they claim to:

(1) Engage in self-reflection/praxis to promote ongoing contemplation and self-checking of personal biases and limited understandings based on our positionalities;

(2) Teach common theory vocabulary, language, and concepts throughout our courses so that students are scaffolded into higher learning;

(3) Engage theory and language into practice through assignments, projects, and outside classroom experiences;

(4) Emphasize how stereotyping and lack of critical understandings about the educational experiences of students of color can perpetuate structural inequalities in society.

Aronson et al. (2020, p. 27).

This seems to be a very helpful set of ideas for moving things forwards in a coherent way, although what is not so clear is how a group of colleagues might come to agreement about the key theoretical perspectives that they wish to base their program on. This part of the process is something that would be helpful to explore more explicitly.

Closely linked to social justice is the idea of environmental justice, another concept which evades clear and easy definition. It is defined by the US Environmental Protection Agency as "the fair treatment and meaningful involvement of all people, regardless of race, color, national origin, or income, with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies" (<https://www.epa.gov/environmentaljustice>), while the UK-based Environmental Justice Foundation talks about "protect[ing] the natural environment and the people and wildlife that depend on it" (www.ejfoundation.org). In Scotland, meanwhile, the discourse focuses on "learning for sustainability", a concept linked to, but not entirely consistent with, environmental justice. Scotland hosts a "regional center of excellence" as part of a global network focusing on education for sustainable development, linked to UNESCO's international Sustainable Development Goals. Here, at the national center, learning for sustainability is described as "the umbrella term for sustainable development education, outdoor learning, global citizenship and social welfare" (<https://learningforsustainabilityscotland.org>). Sustainability also features as a key aspect of the professional standards, described as "understanding and valuing environment, culture and heritage, developing a sense of place and belonging to the local, national and global community. It also means having a deep connection to the natural world and understanding the significance of the choices we make—now and in the future" (GTCS, 2021, p. 3). This expectation in the professional standards sits alongside

“professional values” and “leadership” as aspects which permeate each of the standards comprising the career-long suite. So the emphasis on sustainability in Scotland appears perhaps more environmentally-focused than justice-focused in terms of the discourse, but the links between these two perspectives are clear, as explored below. More fundamentally, However, [Christie et al. \(2019\)](#) suggest that learning for sustainability as a concept, while entrenched in the policy discourse in Scotland, is not widely understood by teachers or teacher educators.

There is, it seems, a continuum of definitions that run from being mostly planet-focused on the one hand, to a mixture of planet and people-focused on the other. Regardless of the definition we adopt, there is clearly a link between social and environmental justice, something more explicitly addressed in the social work education literature ([Teixeira and Krings, 2015](#); [Beltrán et al., 2016](#); [Fogel et al., 2016](#); [Erikson, 2018](#)) than in the general literature on teacher education at present, perhaps due to the more targeted focus of social work provision, whereas teaching provision is universal. Here it is acknowledged explicitly that those most likely to be marginalized in society are also more likely to bear the brunt of environmental problems, and that part of the solution is to adopt a much more overt place-based approach, which “emphasizes ‘hands-on, real-world learning experiences’, increases students’ connections to their own communities, the natural world, and encourages commitment to local civic engagement” ([Beltrán et al., 2016](#), p. 498). This place-based approach engages students with real-world problems, but in so doing, challenges traditional disciplinary-based approaches to education ([Webber and Miller, 2016](#)). A commitment to environmental justice, just as with a commitment to social justice, forces us to understand and challenge forms of hegemonic and overt power which serve to disadvantage some and privilege others.

In summary, while acknowledging the challenges in defining social and environmental justice, and in incorporating these perspectives in teacher education programs, it is clearly possible to agree a common set of theoretical perspectives upon which to base program design, development and delivery. This is necessary, as [Goodwin and Darity \(2019\)](#) stress, in order to ensure that teacher educator knowledge is appropriate for the job of educating for social justice. This, in turn, links to the imperative of accessing and acting upon new knowledge and insight inherent in social progressivism.

Toward a working definition of socially progressive teacher education

This chapter argues that socially progressive teacher education can be unpacked and understood by drawing on three key ideas of social progressivism, pedagogical progressivism and social/environmental justice. The foregoing discussion discusses each of these in turn, drawing out their relevance for teacher education. While these discussions touch on the links between the three ideas, this section aims to consider in more depth how key principles in the three ideas might integrate to inform both understanding of, and practice in, socially progressive teacher education. [Table 1](#) below summarizes key principles underpinning each of the three ideas:

The above summary of underpinning principles highlights a number of key themes prominent in socially progressive teacher education, namely:

1. That knowledge is important;
2. That context matters, and place-based approaches to learning are an important consideration;
3. That both the individual and the collective must be attended to explicitly;
4. That a commitment to social and environmental justice must be conceptualized, understood and operationalized (rather than simply appearing in the discourse around a program).

The centrality of knowledge is explicit across each of these three key ideas. In political terms, social progressivism is founded on reform as a result of new knowledge/insights. This is fundamental to teacher education both in the design and delivery of programs,

Table 1 Key principles underpinning the three elements of socially progressive teacher education.

<i>Social progressivism as a political ideology</i>	<i>Pedagogical progressivism as an educational movement</i>	<i>Social and environmental justice as ethical imperatives</i>
Belief in reform as a means of pursuing social good.	Learner-centered approach, starting from what the learner already knows and has experienced.	Importance of conceptual understanding of social and environmental justice (on the part of both teacher educators and pre-service teachers).
Reform based on ever-evolving understanding of society as a result of new knowledge/emerging insights.	Learner-centered approach, but not individualistic.	Permeation rather than discrete specialist teaching.
Considering both local and global insights from new knowledge.	Emphasis on discovery and problem-based learning.	Importance of teacher educators’ social justice knowledge.
Being open to changing attitudes and practices in light of new knowledge.	Includes secure disciplinary knowledge which can then be drawn on in interdisciplinary ways.	Connection with learners’ contexts key, suggesting place-based learning approaches
Having the capacity to interpret and critique knowledge, leading to understanding of research which can inform future practice (reform).	Inculcates awareness and deeper understanding of one’s own thinking and learning: metacognition.	

and in the content of what is taught. Knowledge is not static, and therefore socially progressive teacher education must be based on the premise that one can never know everything one will even need to know, and that to engage with new research is an important aspect of a teachers' work. Research literacy, involving the capacity to consume research in a critical and context-appropriate way is therefore fundamental.

This links to the second key theme—that context matters. It is clear that socially progressive reform cannot happen in a vacuum, and that context therefore matters. However, attention to key principles underpinning pedagogical progressivism also serve to focus the gaze on issues of place. Place-based education, with its clear alignment to interdisciplinary approaches, context-based and experiential learning, can serve to make socially progressive reform more immediately relevant. For pre-service teachers, the idea of place-based education, rather than “placement learning” implies greater engagement with the communities in which pupils live and learn, thereby potentially requiring a re-conceptualization of the field-based element of teacher education programs. However, while Smith and Sobel's (2010) ground-breaking work on place-based education provides clear exemplification of how this approach might work in schools, translating it into teacher education is a little more challenging. In this respect, it is worth drawing on the growing body of work on community-based teacher education (Guillen and Zeichner, 2018; Yuan, 2018), which promotes pre-service student learning that engages them fully with the communities in which their pupils live and learn.

This inescapable attention to context inevitably places emphasis on working collaboratively for the collective good. However, when set against the learner-centered aspect of pedagogical progressivism, any enactment of social progressive teacher education will inevitably need to attend to a potential tension between the individual and the collective. In this regard, it is helpful to draw on learning from the literature on social justice teacher education, which emphasizes the importance of knowing oneself, and locating this knowledge in relation to others. In their own definition of social justice teaching, Aronson et al. (2020) state a commitment to promoting “social awareness and an ongoing process of critical consciousness toward self in relation to others” (p. 23). It is clearly possible to both learner-centered and collaborative, but this requires skilled negotiation and acute awareness, and arguably starts with teacher educators modeling this way of working themselves.

While social progressivism urges us to draw on knowledge and understanding in support of the common good, it is worth drawing on Biesta (2009, p. 35) warning that “while it is always advisable to use factual information when making decisions about what ought to be done, what ought to be done can never be logically *derived* from what is.” He goes on to assert that all judgments on education are to an extent value-based, made in relation to what is deemed (often implicitly) to be desirable. However, if we simply accept that any version of “good” is acceptable, then we run the risk of allowing social and environmental injustice to continue unchecked. Particularly in the context of education, it is important to be able to accept that somethings are innately not good, regardless of respect for other people's values and beliefs. This is where the concept of social justice, rather than the slightly more nebulous notion of social good is key. That is, the justice element implies something more tangible than simply “good” does. In socially progressive teacher education it is crucial to delve more deeply though, and to engage in active conceptualization of what social and environmental justice mean in the specific local, national and global context in question. The key message here is that from a socially progressive perspective, decisions about reform, even if based clearly on new knowledge, should not be made without reference to context.

The drawing together of themes from across these three key ideas that inform this conceptualization of socially progressive teacher education serve to highlight overlaps between the ideas, as well as attending more explicitly to the importance of context (in its widest sense, including historical, cultural, social, political and geographical elements). This allows a little more detail to be added to the figure presented earlier in the article, thereby explicating the concept of socially progressive teacher education in a little more detail, allowing inter-relationships between elements to be seen more clearly (Fig. 2):

In summary, socially progressive teacher education draws on political, pedagogical and ethical ideas which inform both its conceptualization and operationalization. It is underpinned by thinking on social progressivism as a political ideology, on pedagogical progressivism as an educational movement, and on social and environmental justice as ethical imperatives. This conceptualization is explicitly not neutral, rather it is driven by a desire for knowledge-informed reform which works for social good through participative and contextually-relevant approaches. Social good, from this perspective, adopts an explicit critical lens and works to identify and breakdown barriers to equity, fairness and progress.

In conclusion

Kohn (2008) says that “a school that is culturally progressive is not necessarily educationally progressive”, and presumably the same holds true for teacher education. That is, while the values and aspirations espoused might be socially progressive, this does not necessarily mean that the program will adopt a pedagogically progressive approach. The three prongs of the conceptualization put forward here are therefore co-dependent and all equally important. The socially progressive political ideology serves to make values explicit and to position reform in favor of social good as a key tenet. Pedagogical progressivism then serves as a blueprint for the design of the learning approaches in the teacher education program(s) prioritizing learner-centered approaches and collective, place-based problem-solving. Driving all of this is an explicit privileging of social and environmental justice serving as an ethical, conceptual and practical compass.

While all of this might makes sense conceptually, the real challenge is in enacting the principles herein. This is a job bigger than university-based teacher educators alone can do, rather it will involve a collaborative and activist orientation toward genuinely socially progressive teacher education at institutional, local and national levels. Such an endeavor will require bravery and

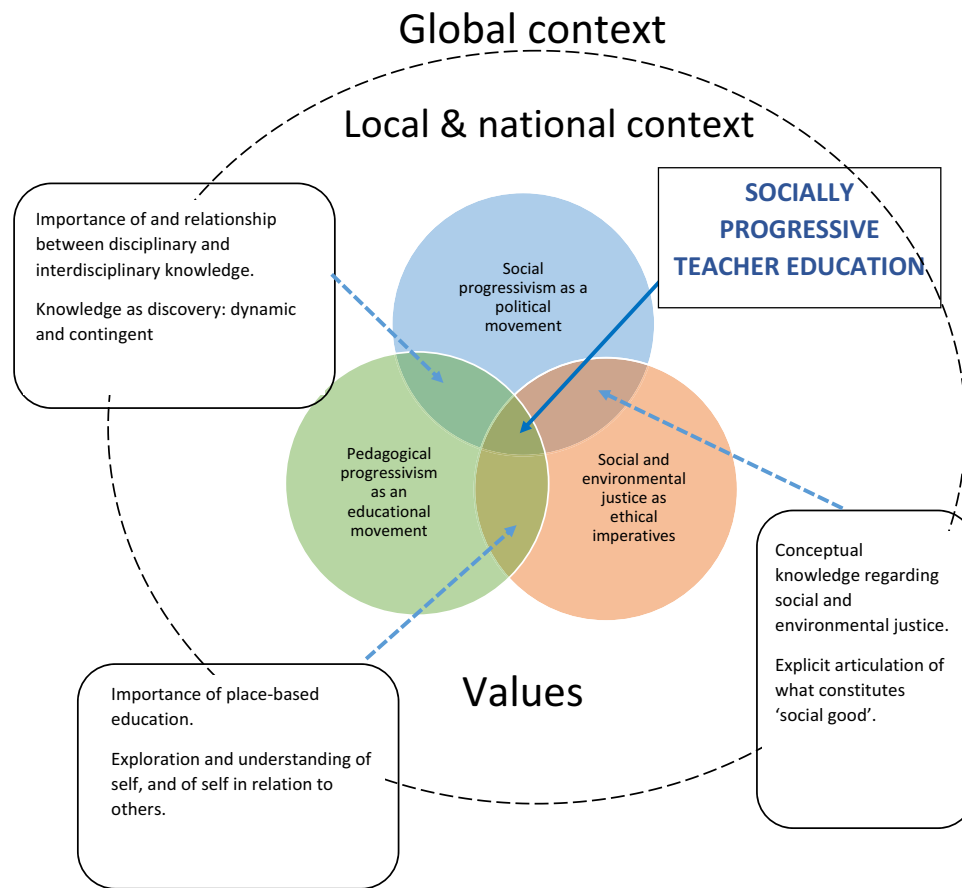


Fig. 2 Conceptualizing socially progressive teacher education and its inter-related elements.

a reconfiguring of control, but ultimately, to use socially progressive teacher education as a framing for this project seems not only to be defensible from a conceptual point of view, but also provides sufficient practical framing for program designers and leaders to make a genuinely socially progressive contribution to society.

References

- Aronson, B., Banda, R., Johnson, A., Kelly, M., Radina, R., Reyes, G., Sander, S., Wronowski, M., 2020. The social justice teaching collaborative: a collective turn towards critical teacher education. *J. Curr. Stud. Res.* 2 (2), 21–39. <https://doi.org/10.46303/jcsr.2020.8>.
- Ashman, G., 2017. Why progressivism matters. <https://gregashman.wordpress.com/2017/04/18/why-progressivism-matters/>.
- Beltrán, R., Hacker, A., Begun, S., 2016. Environmental justice is a social justice issue: incorporating environmental justice into social work practice curricula. *J. Soc. Work. Educ.* 52 (4), 493–502. <https://doi.org/10.1080/10437797.2016.1215277>.
- Benjamin, S., Emejulu, A., 2012. Learning about concepts, terminology and theories: from ambiguity to clarity. In: Arshad, R., Wrigley, T., Pratt, L., Menter, I. (Eds.), *Social Justice Re-examined: Dilemmas and Solutions for the Classroom Teacher*. Trentham Press, London, pp. 33–48.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21 (1), 33–46. <https://doi.org/10.1007/s11092-008-9064-9N09064>.
- Christie, B., Higgins, P., King, B., Collacott, M., Kirk, K., Smith, H., 2019. From rhetoric to reality: examining the policy vision and the professional process of enacting learning for sustainability in Scottish schools. *Scot. Educ. Rev.* 51 (1), 44–56.
- Cobb, D.J., Couch, D., 2018. Teacher education for an uncertain future: implications of PISA's global competence. In: Heck, D., Ambrosetti, A. (Eds.), *Teacher Education in and for Uncertain Times*. Springer, Singapore, pp. 35–48. https://doi.org/10.1007/978-981-10-8648-9_3.
- Cochran-Smith, M., Shakman, K., Jong, C., Terrell, D., Barnatt, J., McQuillan, P., 2009. Good and just teaching: the case for social justice in teacher education. *Am. J. Educ.* 115 (3), 347–377. <https://doi.org/10.1086/597493>.
- Cochran-Smith, M., 2010. Toward a theory of teacher education for social justice. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*. Springer International. https://doi.org/10.1007/978-90-481-2660-6_27.
- Connelly, P., Kennan, C., Urbanska, K., 2018. The trials of evidence-based practice in education: a systematic review of randomized controlled trials in education research 1980–2016. *Educ. Res.* 60 (3), 276–291. <https://doi.org/10.1080/00131881.2018.1493353>.
- Darling-Hammond, L., 2006. *Powerful Teacher Education: Lessons from Exemplary Programs*. John Wiley & Sons.
- Erikson, C.L., 2018. *Environmental Justice as Social Work Practice*. Oxford University Press.
- Fleming, T., 2016. Toward a Living Theory of Transformative Learning: Going Beyond Mezirow and Habermas to Honneth. Mezirow Memorial Lecture. Columbia University Teacher College. http://www.tedfleming.net/doc/Paper_for_TC_June_2016.pdf.

- Fogel, S., Barkdull, C., Weber, B.A. (Eds.), 2016. *Environmental Justice: An Issue for Social Work and Practice*. Routledge.
- Fraser, N., 2000. From redistribution to recognition? Dilemmas of social justice in a "post-socialist" age. *N. Left Rev.* 3, 68–93.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum.
- General Teaching Council for Scotland, 2019a. Guidelines for Accreditation of Initial Teacher Education Programs in Scotland. GTCS. <http://www.gtcs.org.uk/web/FILES/initial-teacher-education/ITE-Programme-Accreditation-Guidelines.pdf>.
- General Teaching Council for Scotland, 2019b. Memorandum of Entry Requirements to Programs of Initial Teacher Education in Scotland. GTCS. <http://www.gtcs.org.uk/web/FILES/about-gtcs/memorandum-on-entry-requirements-to-programmes-of-ite-in-scotland.pdf>.
- General Teaching Council for Scotland, 2021. The Standard for Provisional Registration. GTCS. <https://www.gtcs.org.uk/web/FILES/Professional-Standards/Standard-for-Provisional-Registration.pdf>.
- Goodwin, A.L., Darity, K., 2019. Social justice teacher educators: what kind of knowing is needed? *J. Educ. Teach.* 45 (1), 63–81. <https://doi.org/10.1080/02607476.2019.1550606>.
- Guillen, L., Zeichner, K., 2018. A university-community partnership in teacher education from the perspective of community-based teacher educators. *J. Teach. Educ.* 69 (2), 140–153. <https://doi.org/10.1177/002248711775113>.
- Hamilton, T., 2018. The General Teaching Council for Scotland as an independent body—but for how long? In: Bryce, T.G.K., Humes, W.M., Gillies, D., Kennedy, A. (Eds.), *Scottish Education*, fifth ed. Edinburgh University Press, pp. 873–883.
- Hartley, D., 1987. The convergence of learner-centred pedagogy in primary and further education in Scotland: 1965–1985. *Br. J. Educ. Stud.* 35 (2), 115–128. <https://doi.org/10.1080/00071005.1987.9973756>.
- Hulme, M., Kennedy, A., 2016. Teacher education in Scotland: consensus politics and "the Scottish policy style": responding to challenges across the UK & Ireland. In: *Teacher Education Group (Ed.), Teacher Education in Times of Change*. Policy Press, Bristol.
- Kennedy, A., 2018. Developing a new ITE program: a story of compliant and disruptive narratives across different cultural spaces. *Eur. J. Teach. Educ.* 41 (5), 638–653. <https://doi.org/10.1080/02619768.2018.1529753>.
- Kohn, A., 2008. *Progressive Education: Why It's Hard to Beta, But Also Hard to Find*. Independent School, Spring 2008. <https://www.alfiekohn.org/article/progressive-education/>.
- Labree, D., 2005. Progressivism, schools and schools of education: an American romance. *Paedagog. Hist.* 41 (1–2), 275–288. <https://doi.org/10.1080/0030923042000335583>.
- MacDonald, A., Rae, A., 2018. Initial teacher education in Scotland. In: fifth ed. Bryce, T.G.K., Humes, W.M., Gillies, D., Kennedy, A. (Eds.), *Scottish Education*. Edinburgh University Press, pp. 836–846.
- Mills, C., Ballantyne, J., 2016. Social justice and teacher education: a systematic review of empirical work in the field. *J. Teach. Educ.* 67 (4), 263–276. <https://doi.org/10.1177/0022487116660152>.
- Mincu, M., Davies, P., 2021. The governance of a school network and implications for initial teacher education. *J. Educ. Pol.* 36 (3), 436–453.
- Servant-Miklos, V., Noordegraaf-Eelens, L., 2019. Toward social-transformative education: an ontological critique of self-directed learning. *Crit. Stud. Educ.* 62 (2), 147–163.
- Smith, G.A., Sobel, D., 2010. *Place- and Community-Based Education in Schools*. Routledge, New York, NY.
- Stentoft, D., 2017. From saying to doing interdisciplinary learning: is problem-based learning the answer? *Act. Learn. High. Educ.* 18 (1), 51–61.
- Teixeira, S., Krings, A., 2015. Sustainable social work: an environmental justice framework for social work education. *Soc. Work. Educ.* 34 (5), 513–527. <https://doi.org/10.1080/02615479.2015.1063601>.
- Thorburn, M., 2017. Dewey, democracy, and interdisciplinary learning: a Scottish perspective. *Oxf. Rev. Educ.* 43 (2), 242–254. <https://doi.org/10.1080/03054985.2017.1284657>.
- University of Strathclyde, 2020. *Vision 2025*. University of Strathclyde. https://www.strath.ac.uk/media/1newwebsite/documents/Strategic_Plan_2025.pdf.
- Webber, G., Miller, D., 2016. Progressive pedagogies and teacher education: a review of the literature. *McGill J. Educ.* 51 (3), 1061–1079.
- Yuan, H., 2018. Preparing teachers for diversity: a literature review and implications from community-based teacher education. *High Educ. Stud.* 8 (1), 9–17.

The contribution of initial teacher education to social justice

Fiona R. Ell, Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Initial teacher education and social justice	690
Defining teacher education for social justice	691
Tension in teacher education for social justice	692
ITE's contribution to social justice	693
Contributions through preservice teachers	693
Contributions through teacher educators	694
Contributions through program structures, content and design	695
Making a greater contribution	695
Centering new voices, frameworks and knowledge	695
Criticality and inquiry	695
Place based, embodied, experiential connected teaching and learning	696
Be practical, not just conceptual	696
Conclusion	696
References	696

We can readily observe inequity in our world. Racism, sexism, ableism, discrimination based on religion or sexuality, poverty, classism and colonization are examples of social forces that play out through our institutions and social lives, differentially impacting people and causing poorer outcomes for some members of society. Often these forces intersect, meaning people are multiply disadvantaged. For those who are not impacted, and enjoy positions of privilege, these forces can be invisible. Those who cannot see these forces in play run the risk of contributing to them and perpetuating harm to others.

Most 21st century societies have education systems. Depending on the history and current social context of a country, education might be available freely to all as a public good or reserved for just some people. In some places and times, education is regarded as a dangerous commodity, liable to liberate people's minds and hearts, while in others it is seen as a form of indoctrination and control. In fact, this distinction can occur within a place, where the same education system functions in different ways for different groups of participants. For some it might be an opportunity for personal growth and advancement. For others, in the same system, it might be an oppressive and dehumanizing force. Education is simultaneously a highly personal process of growth, with the potential for transformation, and a systemic process, whereby society offers young people a chance to learn valued knowledge, skills and attitudes. What is valued will differ by political, social, cultural, and historical context. These differences might be national, or they might be between states or provinces within a nation, or they might vary by tribal areas or religious groupings. Teaching, and teacher education, are political, both at the level of the personal power wielded by teachers and their choices and at the level of policy making (Cochran-Smith et al., 2008).

As well as being both personal and systemic, education is often positioned as both the illness and the cure. Education systems, commonly manifest as schooling, reproduce inequities and marginalize groups of learners through their intentions, content, and processes (OECD, 2018). At the same time, education is often positioned in policy and in international studies and comparisons, as the "way out" or the "way up" for underserved groups (OECD, 2018). Education systems are fundamentally intersected by the social forces that produce and reproduce inequity (Walby, 2007), but are tasked with making social change (in politically valued directions) (Pantic and Florian, 2015).

Initial teacher education and social justice

Initial teacher education (ITE) is an integral part of the education system and faces the same paradoxes: it is a process that individuals undertake as a personal journey, but it is simultaneously systemic: ITE is the provision of a "fit for purpose" workforce for schools. It is also the illness and the cure: if young people do not achieve the outcomes that are valued by the education system, teachers are inadequate, and if teachers are inadequate then their selection and preparation was inadequate. The way to fix this is to change, or more commonly reform, ITE and have "better" selection and "better" preparation in order to develop "better" teachers. What is "better" depends, in the same way as what is valued in schooling, on the political, social, cultural, and historical context in which the ITE takes place. As part of the education system, ITE is also intersected by the social forces that prevail through all institutions and social life. ITE does not sit outside the schooling system it serves, nor can it avoid the workings of racism, sexism, ableism, discrimination based on religion or sexuality, poverty, classism, and colonization that underpin the education system as a whole.

Thus, ITE is an integral part of education systems, fundamentally intersected by social forces that create and perpetuate inequities and shaped by national and international policy actions. In response to this positioning, some ITE thinkers, researchers and practitioners advocate “teacher education for social justice” (Cochran-Smith, 2004; Sleeter, 2008). Teacher education for social justice is an idea (Cochran-Smith, 2004) and a “slogan” (Zeichner, 2010) that arises from plural theoretical bases (Chubbuck and Zembylas, 2016) and may mean different things in different times and places because of different histories. Writers on teacher education for social justice particularly point to the rise and dominance of neo-liberal political and economic agendas as antithetical to promoting social justice through teacher education (for example: Bennett, 2021; Kaur, 2012; Sleeter, 2008). In some times and places ITE may be concerned with gender inequity, in other times and places with racism, and in other times and places with the impacts of social class, poverty or disability. Sources of inequity tend to overlap and intersect, meaning that often the intersectional impacts of social forces such as poverty, class, race and colonization are the focus of social justice initiatives. Sometimes these underlying forces are named and sometimes they are hidden within terms such as diversity or multiculturalism (Chubbuck and Zembylas, 2016).

“Teacher education for social justice” is therefore often used as an umbrella term for a range of ideas and commitments, for example culturally relevant/responsive pedagogy (Gay, 2000; Ladson-Billings, 1995; Villegas and Lucas, 2002), equity (Cochran-Smith et al., 2016; Nieto, 2000), and anti-racist education (Au, 2009; Kumashiro, 2000). For this reason, “teacher education for social justice” is something that many teacher education programs can include in their list of aims and aspirations, without necessarily troubling some of their processes or assumptions (McDonald and Zeichner, 2008). Many different teacher education program structures, arrangements, foci and experiences can be identified as serving social justice purposes (Cochran-Smith et al., 2008). For example, the idea of core practices (Ball and Forzani, 2009; Grossman et al., 2009) focused teacher education on key practices that research suggested led to more equitable outcomes. In response, Philip et al. (2019) explain that centering teacher education on core practices can push equity and justice to the periphery. Proponents of both points of view are trying to address inequity and social injustice, but they have different conceptualizations of what perpetuates inequity and injustice and therefore how they might be addressed in teacher education. Cochran-Smith et al. (2008) label this the “ambiguity critique” of teacher education for social justice, where the multiplicity of meanings and uses of the term “teacher education for social justice” render it impotent and hollow. This critique remains salient (Reagan and Hambacher, 2021). A key problem is that “teacher education for social justice” has been used as a disguise, in two ways: people use the term as a slogan, label or signal to others without doing anything meaningful about social justice in their teacher education programs (Zeichner, 2010), or people use it as a more palatable substitute term for activities that others might find threatening or offensive if they were properly named (for example, anti-racist education). This second use is shifting as public discourse shifts, with terms such as “race-visible teacher education” (Hambacher and Ginn, 2021) naming key concerns more precisely, however the term “teacher education for social justice” remains ambiguous. The slippery dual potential of being a smokescreen or a Trojan horse is problematic for promoting social justice through teacher education.

Defining teacher education for social justice

Theorists and researchers of teacher education for social justice have therefore sought to clarify the field through recourse to theories of justice from other fields (for example, Cochran-Smith et al., 2016; Chubbuck and Zembylas, 2016; North, 2006; Reagan and Hambacher, 2021); and reviews of empirical studies and frameworks of social justice in teacher education (for example, Reagan and Hambacher, 2021; Mills and Ballantyne, 2016; Kaur, 2012; Goodwin and Darity, 2019).

Trying to define teacher education for social justice using discourse from other disciplines or other areas of education research is an attempt to specify that social justice is not just about equality of resources or equality of opportunity (Chubbuck and Zembylas, 2016; Reagan and Hambacher, 2021). Offering everyone “the same” might seem “fair” or “equal,” but to tackle inequity in outcomes different groups of people will need different support and resources. Socially just distribution of opportunities and resources is not even, it gives more to those who experience greater disadvantage. For this reason it is about redistribution of social resources, not even-handed distribution of social resources. Social justice is not just about economic processes, however, it is also about recognition and representation (North, 2006). Recognizing and valuing the contributions of all groups is necessary for social justice, as is ensuring that all voices are heard in decision making processes. This three-part conceptualization of social justice, arising from the work of Fraser (Chubbuck and Zembylas, 2016), brings new discourses under the social justice umbrella. Social justice becomes not just about “who has what,” but also who is seen, heard and valued, and who is allowed to participate in decision making. These ideas are the basis for a structural understanding of how inequity is perpetuated. North (2006) notes that these ideas arise from, and are discussed in, largely US terms in her summaries of the field (North, 2006, 2008). Local histories of oppression and injustice are significant in determining what “teacher education for social justice” might mean when it is used in different places and times. However, in order to consider ITE’s contribution to social justice meaningfully, a definition of social justice is needed. Cognizant of the complexities summarized above, Mills and Ballantyne (2016)’s definition of social justice as encompassing “... a fair redistribution of resources and opportunities, recognition of differences, the representation of interests in decision making and development of capabilities to live fulfilling lives” (p. 263) can provide a useful way forward and is taken up here to underpin consideration of what teacher education can contribute to social justice.

As well as examining the theoretical origins of social justice in order to provide clarity about what teacher education for social justice might mean, scholars have synthesized empirical research and conceptual work within teacher education to try to find themes and commonalities that might describe the concept (for example: Reagan and Hambacher, 2021; Mills and Ballantyne, 2016; Kaur, 2012; Goodwin and Darity, 2019). Many authors develop frameworks to describe or explain teacher education for

social justice (for example: [Cochran-Smith, 2004](#); [McDonald, 2005](#); [Pantic and Florian, 2015](#); [Sleeter, 2008](#)). These syntheses and frameworks highlight components of teacher education for social justice (for example, courses, fieldwork, community placements, admission processes) and tensions in teacher education for social justice (for example, between individual agency and institutional processes).

Tension in teacher education for social justice

Underpinning any consideration of ITE's contribution to social justice is a central tension. Understanding this duality is important for understanding why the project of teacher education for social justice is difficult and contested. This central tension was introduced above: the tension between agency and structure, the personal and the political, the local sphere of influence and the broader system, the micro and the macro ([North, 2006](#)). Inequity is perpetuated by systems of power and privilege that work against redistribution, recognition and representation. These systems are peopled by individuals. One response to this is to say "if individual people change what they do, social justice will result." This response aligns with the work that says that teachers make a difference ([Hattie, 2009](#)). In this framework, while teacher differences do not account for all the variance in student outcomes, there is a sizable "chunk" of influence that they can exert as individual teachers, and concentrating their efforts on this space will lead to more equitable outcomes for learners ([Hattie, 2009](#)). Another response is to say "All the actions taken by teachers are shaped and constrained by systemic forces and it is these forces, played out through teachers' choices and actions that create inequity." In this framework, while teachers do have choices, these are fundamentally shaped by systemic practices and ideologies that constrain their agency to improve equity. At the two extremes we have teachers positioned as responsible for perpetuating inequity if their individual efforts are to no avail, or as abdicating responsibility to forces beyond their control and being satisfied with inequitable outcomes because they cannot make a difference. These extreme positions serve to make clear the agency/structure tension; in reality both how teachers perceive their roles, and how they play out, are somewhere on a continuum between these two positions, with context an important contributing factor. Preparing teachers to work within this tension is a fundamental concern of ITE ([Cochran-Smith et al., 2016](#)).

This tension plays out at all levels of the ITE system. [Fig. 1](#) below illustrates this by showing how, at three levels of the ITE system, both personal actions and systemic actions are called for by a commitment to teacher education for social justice. In this conceptualization, teacher education programs as entities can be seen as having some agency, or "personal" choice about what they deliver and how, while also being shaped by the systems of which they are a part. In all the layers, while a useful distinction can be made between acting personally or acting systemically, action in these two spaces is related. Coming to understand one's cultural positioning, a highly personal task, can fuel action at the system level. Activism at the system level can cause re-examination of personal practices. The difficult part is that the knowledge and skills required to undertake the personal actions are different from the knowledge and skills required to undertake systemic actions ([North, 2006](#)). For example, knowing what inclusive pedagogies are and how to enact them with young learners is a different knowledge and skill set to knowing how to work with parents and community groups to advocate for changes to curriculum or policies. This poses a dilemma for ITE: should ITE prepare teachers to have the knowledge, skills and attitudes to work for social justice in both their personal sphere and at a systemic level? How do PSTs, teacher educators and teacher education programs navigate the tension between knowing that as a teacher you can do some things, but your work is shaped by broader social forces too ([Cochran-Smith et al., 2016](#))?

[Fig. 1](#) is an illustration of how the individual/systemic tension is present in multiple layers of teacher education for social justice. There are other layers, and other places, in which this tension plays out, but [Fig. 1](#) serves to show that a commitment to teacher education for social justice implicates both individual and systemic actions by more people and organizations than just individual

	Personal	Systemic
PSTs	Enacting culturally responsive pedagogies in classrooms Understanding their cultural positioning	Advocating against streaming or grouping by achievement level
Teacher educators	Using inclusive practices with PSTs Understanding their cultural positioning	Participate in designing admission processes that are equitable
Teacher education programs	Design programs with learning experiences that are known to enhance social justice commitments and competencies in PSTs.	Advocate with regulatory and approval bodies for socially just processes

Fig. 1 Examples of individual and systemic activities in teacher education for social justice.

PSTs. This suggests that ITE's contribution to social justice has to consist of more than just the provision of "socially-just teachers" for the education system.

ITE's contribution to social justice

ITE's contribution to social justice can be framed in two ways. These framings are simplifications, but they expand the tension described in the previous section, and broadly represent the arguments that are used to promote the idea that social justice has a place in teacher education. The first framing centers on the personal aspect of teaching for social justice: student outcomes from schooling are unjust, teachers have a significant influence on student outcomes (Hattie, 2009), teachers should teach in ways that promote social justice to produce more just outcomes, therefore teachers need ITE that prepares them with the knowledge, skills and attitudes to teach in ways that promote social justice. The second framing focuses on the systemic aspect of teaching for social justice: schooling is systemically disadvantaging groups of students, teachers are working within structures that shape and constrain their actions and individual efforts will not overcome these structures (and the structures are often invisible and unwittingly perpetuated by teachers), teachers should therefore be advocates for change to systemic inequities and processes that lead to unjust outcomes, therefore teachers need ITE that prepares them to be advocates and activists. These two framings are not necessarily incompatible, but they allow a wide scope for teacher education programs to claim they are providing teacher education for social justice (McDonald and Zeichner, 2008) and a program that attempts to do both has a very big agenda. For example, Sleeter (2008) conceptualizes social justice in teacher education using three strands: "supporting access for all students to high-quality, intellectually rich teaching that builds on their cultural and linguistic backgrounds, preparing teachers to foster democratic engagement among young people and preparing teachers to advocate for children and youth by situating inequities within a systemic sociopolitical analysis" (p. 611). The first of these strands covers the personal aspect described above, and the third strand is part of the systemic aspect described above, with the second strand combining the two: PSTs should be able to encourage engagement with social justice ideas among their students. This strand involves both knowing what and how to teach and being able to provide a systemic critique.

The following sections consider the contribution of ITE to social justice by taking the three levels in Fig. 1 and thinking about the contributions ITE can make to social justice at each of these levels. A three-level framework is used for this explanation in order to make it clear that there are more ways that ITE can contribute to social justice than through the PSTs it graduates. In a similar way to the key tension outlined above, seeing ITE's contribution to social justice as only in PSTs' teaching is ignoring the role ITE can play at a systems level. In practice, the three levels are interdependent with respect to promoting social justice. There is a point at which understanding the relationships among the levels becomes more explanatory than making a distinction between them. A final section considers these relationships and the ways in which ITE could make a greater contribution to social justice than it does at the moment.

Contributions through preservice teachers

One way in which ITE can contribute to social justice is through the PSTs that it works with. Most empirical work on teacher education for social justice centers on PSTs: their beliefs, attitudes, competencies and experiences (Mills and Ballantyne, 2016). In the US, and other colonial jurisdictions in particular, a key driver for social justice in teacher education is the "demographic imperative" (Cochran-Smith, 2004). The demographic imperative arises when the teaching workforce does not "match" the students they teach in terms of race, language and social class. There are two responses to this: recruit and admit PSTs who are more representative of the student population, or try to make the "mismatched" PSTs recognize their positioning and its implications for their students. Both of these approaches lead to difficulties for PSTs from non-dominant groups.

In actively recruiting more diverse students we may be using the logic "adding more teachers of Color in schools with majority students of Color from high poverty communities will change the schools, transform the educational experiences for the students and close the achievement gap" (Haddix, 2017). Inevitably, when this does not work because of systemic racism, the teachers of color will be blamed (Haddix, 2017). While diversifying the teaching workforce is a way to have new voices heard and knowledges shared, expecting people from oppressed groups to solve the problems created by the dominant system is unreasonable and unlikely to succeed.

Jackson et al. (2017) present an interesting perspective on the second approach—helping PSTs to recognize their positioning. PSTs of Color are the co-authors of this piece, which explores the experience of learning about teaching for social justice from a professor of Color in a predominantly white institution and teacher education program. Jackson et al. (2017) point out the ways that race intersects this experience: the professor of Color teaches the social justice course which creates a racialized dynamic with the white students that is further complicated for the PSTs of color who are positioned differently in relation to the material. Three assumptions are critiqued in this work: first, that PSTs of Color will automatically know how to teach for social justice, second that, being of Color or a member of a marginalized group in education automatically means that you know and understand about all marginalized groups and have no misconceptions, and third that the learning experiences in the course are shifting the positions of the white students (Jackson et al., 2017).

Studies of PSTs and teaching for social justice are often small, focused on one course or program and rely on one source of data, most commonly interviews (Mills and Ballantyne, 2016). This makes it difficult to know whether teacher education for social justice

is making a contribution to social justice in schooling through the knowledge, skills and attitudes of the PSTs it graduates. [Cochran-Smith et al. \(2009\)](#) describe the outcomes of a social-justice oriented teacher education program, emphasizing that the PSTs had not abandoned a focus on learning in order to teach in “good and just” ways. The PSTs in this study emphasized relationships, focused on learning for every student and took a critical stance toward their work. There was little evidence of them taking up the advocacy or activist elements of social justice. This study has a longer timeframe than many ([Mills and Ballantyne, 2016](#)), and tries to specify what social justice teaching might look like in PST practice. In doing so it suggests that it is possible for teacher education to shape PSTs to take up social justice commitments in their teaching and marry these with some practices that support their goals.

Contributions through teacher educators

One resource that teacher education has, which is often overlooked in the discussion of teacher education for social justice, is teacher educators. Studies often consider the impacts of program structures or practicum arrangements without acknowledging that it is educators who make these things happen, and that their attitudes, knowledge and skills impact teacher education for social justice. “Who teaches material relating to social justice?” is a key first question. Whose voices are heard in the teacher education space? What help and support are teacher educators offered to teach for social justice themselves? Do the teacher educators teach in ways that promote social justice in their own classes? Are the teacher educators critical and activist in their stance toward their work? Teacher educators in university settings could be advocates and models for teaching for social justice across disciplines as well as teaching PSTs in inclusive ways and helping to dismantle structures that perpetuate inequity in their institutions. Recent research work on teacher educators suggests that reaching that point will be a difficult and complex journey. Some of this work is described below to highlight how the key tension between the personal and the systemic elements of teacher education for social justice play out in the teacher educator space as well as with PSTs.

[Goodwin and Darity \(2019\)](#) reviewed literature from three jurisdictions that centered on teacher educators’ knowledge for teaching social justice in teacher education. Their findings echo the fundamental tension outlined above: teacher educators highlighted personal and contextual knowledge in teaching for social justice, and pedagogical, sociological and social knowledges were largely absent from the studies and from the discussion. This suggests that teaching for social justice is generally being taught by teacher educators as a personal and local task (which it is, in part), with the larger structural elements and critiques being absent. [Goodwin and Darity \(2019\)](#) call for formal preparation for teacher educators. Formal preparation would give space for prospective teacher educators to build their knowledge and understanding of teaching for social justice and to work out what it means for their own practice.

[Souto-Manning and Winn \(2019\)](#) list commitments that teacher educators need to take up in order to provide socially-just teacher education. Their first suggestion is “suspend harm as an ethical imperative,” pointing out the harm that teacher educators currently cause through the frames they communicate about students and schooling. [Jackson et al. \(2017\)](#) provide a valuable perspective on this when they report experiencing teacher educators as framing issues of race and equity around “... the failure of students of Color (Black students in particular) rather than the failure of the education system” (p. 104). Again, this links to the central tension above—locating social justice issues in personal rather than systemic spaces.

[Gorski and Parekh \(2020\)](#) provide data on institutional structures that may serve to maintain the status quo in teacher education for social justice. Noting that most North American teacher education programs put “equity-related skills” into one course, and that these courses vary from appreciating diversity and cultural competence to having a critical orientation and an activist stance (p. 265), [Gorski and Parekh \(2020\)](#) asked teacher educators for their perceptions of the criticality of their course and the support of their institution for the values they teach. They found a negative relationship between the two factors—teacher educators who took more critical approaches perceived that there was less institutional support for their work. Teacher educators, especially in universities, are often working in spaces that act as engines of privilege. This poses a challenge to teacher education for social justice, especially if teacher educators themselves do not recognize what is happening.

[Dover et al. \(2020\)](#) report on a promising practice for developing teacher educators’ social justice teaching. They formed a “faculty learning community” that created a space for teacher educators to critically evaluate their teaching for social justice, to understand the impacts of their practice and to consider how it was shaped by institutional and social forces. Like [Souto-Manning and Winn \(2019\)](#), [Dover et al. \(2020\)](#) point out that when teacher educators continue to do what they do they “either reinforce or actively perpetuate unjust, racist and unequal practices” (p. 62). Interrupting this, with collegial support, is necessary for teacher education for social justice to take place.

Teacher educators in many jurisdictions do not “match” diverse school populations, nor do they “match” the increasing diversity of PSTs. In Western colonies in particular, whiteness pervades teacher education ([Bennett, 2021](#); [Deckman and Ohito, 2019](#); [Souto-Manning and Winn, 2019](#)). However, as with the PSTs discussed above, teacher educators are not a homogeneous group. A response to social justice imperatives is to hire more diverse faculty members, but in the same way as recruiting more diverse PSTs, this is not a solution. [Deckman and Ohito \(2019\)](#) reflect on their experiences as Black women teacher educators in White teacher education spaces, revealing the multiple complexities of their positioning and their work. In particular they point out that while social justice as a term has connotations of healing and reconciliation, it is important to consider who benefits from teacher education for social justice and who might be harmed by it. For example, if we know that discussing issues of race is difficult work and that PSTs can be angry and anxious about it ([Jackson et al., 2017](#)), who are we asking to do that work? Is it people who have already been harmed by the racism they are being asked to work with explicitly?

The contribution of teacher educators to teacher education for social justice needs deeper consideration by researchers and practitioners. There are personal considerations: what do teacher educators do when they teach social justice? What are the experiences of teacher educators teaching for social justice? However, the greater considerations are systemic: who does the work of teacher education for social justice, and why? Why are some groups of people missing from teacher education? What institutional structures hold harmful practices in place? What supports are there for teacher educators to change their practice? Who does teacher education for social justice really serve?

Contributions through program structures, content and design

Discussion of teacher education for social justice often takes place at the program level. Programs of teacher education claim to be social justice oriented, or to promote social justice (McDonald and Zeichner, 2008). Discussions at the program level tend to center on the arrangements and activities in a program: courses, assignments, fieldwork/practicum and community or service learning, for example (Sleeter, 2008). Gorski and Parekh (2020) claim that in most North American programs, social justice content is confined to one or two courses on “multiculturalism” or “diversity,” rather than being taken up across courses in a program. Jackson et al. (2017) highlight a lack of coordination across programs, with students reporting that “everyone gives the same Ladson-Billings article but no one unpacks it” (p. 106). Chubbuck and Zembylas (2016) explain the need for holistic teacher education programs with a coherent and integrated focus on social justice if teacher education for social justice is going to be effective. Program-level structures need to be aligned with, and express, social justice aims, not just single courses or teacher educators (Cochran-Smith et al., 2016).

McDonald (2005) provides a case study of two teacher education programs with explicit commitments to social justice, and examines the nature of student teachers’ opportunities to learn about social justice in these programs. She offers a framework for thinking about programs on two dimensions: the elements of social justice that PSTs get the opportunity to learn about (from individual through to institutional) and whether these opportunities are conceptual or practical. Her analysis suggests that PSTs get more opportunities to learn about individual students and their needs than groups of students and recognition of oppression and its institutional causes. In addition, they get more opportunities to engage with conceptual tools than practical tools for teaching for social justice. This is echoed in Jackson et al. (2017) who report being given theoretical or conceptual work and being asked to work out for themselves what it might mean in practice.

All of the programs discussed above can be critiqued as centering Whiteness and European concerns (Bennett, 2021; Souto-Manning and Winn, 2019). Whitinui et al. (2018) present a range of promising practices in teacher education that are built on indigenous frameworks and concerns. Programs that are centered on indigenous world views present different opportunities, and meanings, for teacher education for social justice. Clarke et al. (2018) describe how a teacher education program was devised as a collaboration between Māori (the indigenous people of Aotearoa New Zealand) and a teacher education provider. This collaboration led to a program that was centered in Māori worldviews and approaches, with the metaphor of a braided river bringing together streams of knowledge at the heart of the program. Social justice in a program oriented in this way does not reinforce a deficit narrative about indigenous learners, nor is it approached from a Western point of view. Programs of this type suggest a way forward for teacher education to make a greater contribution to social justice.

Making a greater contribution

How, then, could ITE make a greater contribution to teacher education for social justice? Four ways forward can be seen from looking across the literature about frameworks and practice in social justice teacher education. Each of these four directions is considered briefly below.

Centering new voices, frameworks and knowledge

It is difficult for teacher education to contribute to social justice when it begins from unquestioned, dominant frameworks and practices. A greater contribution could be made if different voices and knowledges were centered in teacher education programs. Community involvement in designing teacher education (McDonald and Zeichner, 2008), beginning with indigenous frameworks (Clarke et al., 2018) or centering justice (Philip et al., 2019) are starting points for building social justice into the structure of teacher education, rather than making it a goal or outcome. Teacher education that is designed through a socially-just, power-sharing process is more likely to have arrangements, content and graduates that reduce inequity.

Criticality and inquiry

Two approaches to teaching and learning, and teacher education, that are associated with social justice practice are criticality and inquiry (Bennett, 2021; Cochran-Smith et al., 2016). Neither of these is a guarantee of deep learning about social justice as they can be undertaken superficially or without understanding (Kaur, 2012) but, when taught and used sincerely, criticality and inquiry are ways of thinking and approaching teaching that can promote social justice. Sustained use of criticality and inquiry throughout

a teacher education program, and by the teacher educators within a program supports PSTs to adopt these mindsets as habits, and apply them to their own practice.

Place based, embodied, experiential connected teaching and learning

Teacher education takes place in a context, and while there are international commonalities, responding to local conditions and histories is a way to contribute more to teacher education for social justice. Indigenous pedagogies or those from other community groups can make a significant contribution to how PSTs engage with social justice ideas and practices (Clarke et al., 2018; McDonald and Zeichner, 2008; Deckman and Ohito, 2019). Learning from these pedagogies and learning how to enact these pedagogies are both ways to increase teacher education's contribution to social justice. Embodied and experiential learning are useful approaches because they acknowledge the emotional and embodied nature of the learning that takes place when PSTs (and teacher educators) engage with social justice ideas and practices (Reagan and Hambacher, 2021).

Be practical, not just conceptual

Although a criticism of ITE at present could be that there is not enough consideration given to larger ideas about structural inequity and too much emphasis is placed on learning about catering to individual difference (Philip et al., 2019; Sleeter, 2008), a counter-intuitive problem also emerges from the literature: teaching about social justice is conceptual where it needs to also be practical (Goodwin and Darity, 2019; Jackson et al., 2017; McDonald, 2005). PSTs report being given assignments where they read about social justice and then are asked to create lesson plans that reflect what they have read. Goodwin and Darity (2019) note that there is rhetoric of social justice in the work they reviewed, but not necessarily practice. Being clear about how social justice ideas and commitments play out in schools and in teacher education spaces would contribute to teacher education for social justice. Practical examples from a range of settings, illustrating key principles and sparking focused discussion and reflection would help PSTs and teacher educators fully understand what teacher education for social justice means and what it can do.

Conclusion

Teacher education is part of education systems that reproduce inequities. For this reason teacher education needs to contribute to social justice. This effort is complicated by a lack of clarity about what social justice might mean and how teacher education might help. Social justice in schooling involves both individual actions that are just and systems that are just. ITE needs to prepare PSTs who can teach in ways that are socially just and contribute to reforming systems that perpetuate inequities. This can be hard to do when many teacher educators and PSTs are racially, socially and culturally "insiders" to the education system, where many of those they are working with have been positioned as "outsiders" to the education system. Institutional and broader social systems also hold processes and structures in place that reproduce inequity.

The contribution that ITE makes to social justice is limited by its position in the education system and its institutional nature. It is also limited by the attitudes, knowledge and skills of those involved in teacher education. These two limits reflect the core tension of teacher education for social justice: that it is about systemic influences and personal influences. Across both of these sets of influences, learning to teach for social justice involves dealing with difficult things. It is not a neutral activity. There is hard personal work to be done by many PSTs and teacher educators. Teaching for social justice makes demands of people that they may find hard to fulfill.

References

- Au, W., 2009. Rethinking Multicultural Education Teaching for Racial and Cultural Justice. A Rethinking Schools Publication, Milwaukee.
- Ball, D., Forzani, F., 2009. The work of teaching and the challenge for teacher education. *J. Teach. Educ.* 60 (5), 497–511.
- Bennett, J., 2021. Collective criticality: towards sustainable socially just teacher education programs. *Whiteness Educ.* 6 (1), 75–91.
- Chubbuck, S., Zembylas, M., 2016. Social justice and teacher education: context, theory and practice. In: Loughran, J., Hamilton, M. (Eds.), *International Handbook of Teacher Education*. Springer, Singapore, pp. 463–501.
- Clarke, T., Macfarlane, S., Macfarlane, A., 2018. Integrating indigenous Maori frameworks to ignite understandings—within initial teacher education—and beyond. In: Whitinui, P., Rodriguez de France, M., McIvor, O. (Eds.), *Promising Practices in Indigenous Teacher Education*. Springer, pp. 71–86.
- Cochran-Smith, M., Barnatt, J., Lahann, R., Shakman, K., Terrell, D., 2008. Teacher education for social justice: critiquing the critiques. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, NY, pp. 625–639.
- Cochran-Smith, M., Shakman, K., Jong, C., Terrell, D., Barnatt, J., McQuilla, P., 2009. Good and just teaching: the case for social justice in teacher education. *Am. J. Educ.* 115, 347–377.
- Cochran-Smith, M., Ell, F., Grudnoff, L., Haigh, M., Hill, M., Ludlow, L., 2016. Initial teacher education: what does it take to put equity at the center? *Teach. Teach. Educ.* 57, 1–12.
- Cochran-Smith, M., 2004. *Walking the Road: Race, Diversity and Social Justice in Teacher Education*. Teachers College Press, NY.
- Deckman, S., Ohito, E., 2019. Stirring vulnerability, (un)certainly and (dis)trust in humanizing research: duoethnographically remembering unsettling racialized encounters in social justice teacher education. *Int. J. Qual. Stud. Educ.* 33 (10), 1058–1076.
- Dover, A., Kressler, B., Lozano, M., 2020. Learning our way through: critical professional development for social justice in teacher education. *N. Educ.* 16 (1), 45–69.
- Gay, G., 2000. *Culturally Responsive Teaching: Theory, Research and Practice*. Teacher's College Press, New York.
- Goodwin, L., Darity, K., 2019. Social justice teacher educators: what kind of knowing is needed? *J. Educ. Teach.* 45 (1), 63–81.

- Gorski, P., Parekh, G., 2020. Supporting critical multicultural teacher educators: transformative teaching, social justice education, and perceptions of institutional support. *Intercult. Educ.* 31 (3), 265–285.
- Grossman, P., Hammerness, K., McDonald, M., 2009. Redefining teaching, re-imagining teacher education. *Teach. Teach.* 15 (2), 273–290.
- Haddix, M., 2017. Diversifying teaching and teacher education: beyond rhetoric and toward real change. *J. Lit. Res.* 49 (1), 141–149.
- Hambacher, E., Ginn, K., 2021. Race-visible teacher education: a review of the literature from 2002 to 2018. *J. Teach. Educ.* 72 (3), 329–341.
- Hattie, J., 2009. *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*, first ed. Routledge, UK.
- Jackson, T., Ballard, A., Drewery, M., Membres, B., Morgan, L., Nicholson, F., 2017. Black like me: female preservice teachers of color on learning to teach for social justice with a black female professor. *Adv. Race Ethn. Educ.* 6, 93–113.
- Kaur, B., 2012. Equity and social justice in teaching and teacher education. *Teach. Teach. Educ.* 28, 485–492.
- Kumashiro, K., 2000. Toward a theory of anti-oppressive education. *Rev. Educ. Res.* 70 (1), 25–53.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- McDonald, M., Zeichner, K., 2008. Social justice in teacher education. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, NY, pp. 595–611.
- McDonald, M., 2005. The integration of social justice in teacher education. *J. Teach. Educ.* 56 (5), 418–435.
- Mills, C., Ballantyne, J., 2016. Social justice and teacher education: a systematic review of empirical work in the field. *J. Teach. Educ.* 67 (4), 263–276.
- Nieto, S., 2000. Placing equity front and center: some thoughts on transforming teacher education for a new century. *J. Teach. Educ.* 51 (3), 180–187.
- North, C., 2006. More than words? Delving into the substantive meaning(s) of “social justice” in education. *Rev. Educ. Res.* 76 (4), 507–535.
- North, C., 2008. What’s all this talk about “social justice”? Mapping the terrain of education’s latest catchphrase. *Teach. Coll. Rec.* 110 (6), 1182–1206.
- OECD, 2018. *Equity in Education: Breaking Down Barriers to Social Mobility*. PISA, OECD Publishing, Paris. <https://doi.org/10.1787/9789264073234-en>.
- Pantic, N., Florian, L., 2015. Developing teachers as agents of inclusion and social justice. *Educ. Inq.* 6 (3), 333–351.
- Philip, T., Souto-Manning, M., Anderson, L., Horn, I., Andrews, D., Stillman, J., Varghese, M., 2019. Making justice peripheral by constructing practice as “core”: how the increasing prominence of core practices challenges teacher education. *J. Teach. Educ.* 70 (3), 251–264.
- Reagan, E., Hambacher, E., 2021. Teacher preparation for social justice: a synthesis of the literature, 1999–2019. *Teach. Teach. Educ.* 108, 2–11.
- Sleeter, C., 2008. Teacher education, neoliberalism and social justice. In: Ayers, W., Quinn, T., Stovall, D. (Eds.), *Handbook of Social Justice in Education*. Routledge, NY, pp. 611–624.
- Souto-Manning, M., Winn, L., 2019. Toward shared commitments for teacher education: transformative justice as an ethical imperative. *Theor. Pract.* 58 (4), 308–317.
- Villegas, A., Lucas, T., 2002. Preparing culturally responsive teachers: rethinking the curriculum. *J. Teach. Educ.* 53 (1), 20–32.
- Walby, S., 2007. Complexity theory, systems theory and multiple intersecting social inequalities. *Philos. Soc. Sci.* 37 (4), 449–470.
- Whitinui, P., Rodriguez de France, M., McIvor, O., 2018. In: *Promising Practices in Indigenous Teacher Education*. Springer.
- Zeichner, K., 2010. Competition, economic rationalization, increased surveillance, and attacks on diversity: neo-liberalism and the transformation of teacher education in the US. *Teach. Teach. Educ.* 20 (8), e1544–e1552.

Online teaching and learning practices in teacher education: past, present and future

Carmen Carrillo^a and Maria Assunção Flores^b, ^a Department of Music, Plastic Arts and Physical Education, Faculty of Educational Sciences, Autonomous University of Barcelona, Bellaterra, Spain; and ^b Research Centre for Child Studies, Institute of Education, University of Minho, Campus de Gualtar, Braga, Portugal

© 2023 Elsevier Ltd. All rights reserved.

Introduction	698
Methods	699
First study	699
Second study	699
Online teaching and learning practices driving impact in teacher education in the era pre-covid	700
Online teaching and learning practices and factors influencing impact in teacher education in the era post-covid	701
Practices influencing impact in online teacher education	701
Characteristics of teaching and learning practices	701
Reflexivity on practice in the practicum	702
Factors influencing impact in online teacher education	704
Personal and contextual elements	704
The culture of “care”	704
Challenges in the transition to online education	705
Online teacher education: implications for future practice and research	705
References	707

Introduction

Although online teaching and learning are not new, the outbreak of the COVID-19 pandemic in 2020 has indelibly marked teacher education practices due to the forced and global transition to online environments as a result of the closure of teacher education institutions and schools worldwide.

An “online emergency remote teaching” (Bozkurt and Sharma, 2020, p. ii) became a reality in many parts of the world in order to meet the challenges arising from the closure of the institutions. What became immediately evident was the digital gap visible in the inequalities in terms of access to adequate equipment, resources and interactivity for online teaching and learning but also in terms of digital literacy from the part of various stakeholders (Flores et al., 2021b). Moving from face-to-face to the online format has been seen by the more enthusiasts as a revolution in education while others are more cautious and alert to the exacerbation of inequalities and exclusion (Flores and Gago, 2020). Institutions around the world developed new strategies and scenarios for online teaching and learning and research has shown that a more positive adaptation has been associated with personal (e.g. self-regulatory and socio-emotional competences) and contextual factors (e.g. support, feedback) (Flores et al., 2021a).

Despite the variety of responses, online teaching and learning practices became not only necessary but also entailed a global response to a situation of crisis that was new and unexpected and that, at least, challenged existing practices. It is, therefore, possible to identify a before and after the COVID-19 pandemic for online teacher education that needs to be examined in all its dimensions. As Darling-Hammond and Hyler (2020) assert, there is a need to pay attention to teaching and teacher education during the pandemic and beyond and to invest in strategies aimed to support the social, emotional and academic needs of students through high-quality educator preparation, opportunities for professional development, mentoring, and the creation of the conditions for educators to collaborate. These issues raise a number of questions that need to be addressed, namely: To what extent have online teaching and learning practices in teacher education been changed (or at least challenged) during and in the post-covid time? What are their effects on the process of learning to teach? How has teacher educators and student teachers’ work been affected during the pandemic? What are the implications for the future of teacher education?

Situating COVID-19 as a turning point in online teacher education, this article aims to analyze online teaching and learning practices before and after the irruption of the pandemic to understand the nature and the kinds of practices and factors affecting impact in the online environment. As such, this article is organized around three parts. First, we look at the pre-covid time—the past—by summarizing what is known in the field. This section draws on previous work (Carrillo and Flores, 2020) and aims to provide an overview of what has been done in the pre-covid era in terms of online teaching and learning practices. Our first review included papers published between January 2000 and April 2020, leaving outside the more recent studies that examine issues on online teaching and learning in the exceptional circumstances of the COVID-19 pandemic. In such unexpected scenario, higher education institutions had to rapidly adapt to remote teaching in order to meet the requirements of the programs and to provide students with opportunities to learn and to meet the course goals. However, it is important to move beyond an emergency online

teaching and learning response and look at evidence-based practices driving impact. Thus, the second section is devoted to the COVID-19 time—the present—and, drawing also on previous work (Carrillo and Flores, *in review*), it explores the context and circumstances in which higher education institutions have had to develop online practices. Such a goal is important to understand the particularities of the moment and the decisions made in the transition from the face-to-face to the online format. It is, therefore, essential to examine how educators have adapted to remote teaching after the irruption of the pandemic, the conditions for operating teaching and learning and the strategies used in such challenging circumstances.

Finally, in the last section of this article, we examine the implications of the findings for online teaching and learning practices in teacher education—the future—and we discuss future trends and avenues for further research.

The three sections will be guided by the following research questions:

- What online practices have been demonstrated to drive impact on the teaching and learning process in the context of teacher education before the irruption of the COVID-19?
- What online practices and factors have been demonstrated to influence impact on the teaching and learning process in the context of teacher education after the outbreak of the virus?
- What are the implications of the findings derived from the two previous research questions for online teaching and learning practices in the post-pandemic time?

Methods

This article reports on data gathered from two searches of the literature on online teaching and learning practices in the context of teacher education. The first one includes publications between January 2000 and April 2020 and the second one focused on the period April 2020–April 2021. The selection process started with a search in the databases Web of Science (main collection) and Education Resources Information Center (ERIC).

Publications whose descriptors included the terms “online learning” (or “digital learning” or “e-learning” or “web-based learning” or “remote learning” or “distance learning” or “virtual learning”) and “teacher education” or “teacher training” or “teacher preparation” were sought. Albeit such terms and expressions entail different meanings and understandings, all of them were considered for the purpose of the study of teaching and learning practices in teacher education mediated by the use of technology in a remote scenario.

First study

In total, the search focusing on publications between January 2000 and April 2020 led to the identification of 134 papers which fulfilled the inclusion criteria established for this research in terms of the publication source (academic journals), the nature of the studies (empirical), and the relevance of the topic (main focus on online teaching and learning practices in the context of teacher education).

The 134 papers were subject to two types of analysis. First, a descriptive analysis was conducted. It implied the elaboration of a summary table for each of the papers specifying the focus of the study, characteristics of the sample, the methods and the main findings (due to the word limit, this table is not included in this article). Second, a content analysis (Ryan and Bernard, 2000) was undertaken following the Community of Inquiry Framework (CoI) (Garrison et al., 2000) to identify the categories. This framework is based on the idea that learning is the result of the interaction of three presences: (a) social presence, associated with the ability of participants to engage affectively with a community, communicate purposefully in a collaborative environment, and develop interpersonal relationships by projecting themselves as the people they are (Garrison et al., 2000); (b) cognitive presence, or the extent to which the participants are able to construct meaning through sustained reflection and communication in a community of inquiry (Garrison et al., 2000); and (c) teaching presence, related to the design, facilitation and direction of social and cognitive processes with the aim of achieving meaningful learning outcomes (Anderson et al., 2001). A cross-analysis of the data as related to each category (social, cognitive and teaching presence) was carried out in order to identify patterns that make sense beyond each specific case (Huberman and Miles, 1994) but also taking into consideration the peculiar features of each paper.

Peer reviews conducted through regular exchanges between the two researchers and co-authors, the second not directly involved in the data collection and analysis, provided an external check of the research process (Lincoln and Guba, 1985). This researcher (second author) regularly reviewed the material generated and contributed with comments on both the methodological process and the product. The insights provided by this researcher on the selection of the databases, the criteria established for the search, the types of analysis conducted and the accuracy of the analysis (process) as well as on the assessment of how the findings, interpretations and conclusions of this study are supported by the data (product) allowed to refine the research process.

Second study

As for the publications between April 2020 and April 2021 a total of 20 papers were considered taking into account the same above inclusion criteria. In addition, a manual search through the bibliography of some key papers (Bryman, 2012) was carried out in a second phase. This process was guided by the same inclusion criteria (see above) and generated eight additional papers.

Similar to the first study, a descriptive analysis was conducted which included the elaboration of a summary table for each of the papers specifying the focus of the study, characteristics of the sample, the methods and the main findings (due to the word limit, this table is not included in this article). Then, a systematic inductive analysis (Hatch, 2002) was carried out that took the researchers to independently read and identify salient domain connected with the second research question (practices and factors influencing impact in online teacher education); to reread the data to refine salient domains and confirm if these are supported by the data; to search for themes across domains, and; to discuss and agree on the final themes of analysis. These were established as follows: in relation to practices influencing impact, the characteristics of online teaching and learning practices and the reflexive element of the practicum; in relation to factors, the personal and contextual elements, the “caring” component and the challenges encountered in the transition to online education.

To warrant accuracy of the research process, two “verification” strategies (Creswell, 1998) were used. Member checking (Doyle, 2007) by the two researchers and co-authors was performed in order to verify the interpretation of the data. Also, detailed descriptions of the results are provided to allow readers to make decisions regarding transferability of the findings (Erlandson et al., 1993).

The main findings of the second study and a summary of the first study will be presented in the next two sections. Also, in light of the three main research questions identified in the introduction, a final section focusing on the implications for teacher education in the future will be included.

Online teaching and learning practices driving impact in teacher education in the era pre-covid

Out of the papers examined in the first study, only a few focused primarily on the *social presence* (e.g. Hramiak, 2010; Komninou, 2017; Li, 2011; Satar and Akcan, 2018; Yeh, 2010) although the vast majority included issues related to this presence and highlighted its centrality in terms of teaching and learning effectiveness. The ability of teachers and learners to engage affectively in relationships emerged as a key element in meaningful educational experiences. According to the papers reviewed, the factors influencing the establishment of a social presence included:

- Belongingness: shared values and interests; trusting human relationships.
- Cohesiveness: collaboration in small groups; high levels of support; affective responses.
- Optimal level of social presence:
 - By instructors: balance between providing support and leaving room for student participation.
 - By learners: developing an active and inclusive attitude.
- Interactivity: regular discussions; prompt communication; timely contributions; interpersonal interactions.
- Participation: authentic and practical activities; relevant commenting; prioritization of social interaction over the completion of the task.

The papers reviewed also suggest that the interplay of the social presence with cognitive and teaching components was key to ensure teaching and learning impact. The interactions that supported social presence were multiple, iterative and reciprocal (Hou, 2015) and they were fundamentally based on mutual respect, interdependence and trust to focus on the interpersonal connection between the members involved (Ham and Davey, 2005).

The *cognitive presence* was subject to actions of a social and pedagogical nature, although limited attention to the specific features supporting its development was sometimes detected. This may be explained by the difficulty to distinguish the elements of the learning environments that influenced knowledge development (Evens et al., 2017) or by the lack of a sound pedagogy of online education (Ham and Davey, 2005). Overall, the factors influencing the establishment of a cognitive presence according to the papers reviewed are related to:

- Concrete experience: sharing teaching experiences; observing the most practical videos within authentic classrooms situations.
- Contextualization and conceptualization: facilitating the active intervention of both instructors and learners in the process of constructing meaning; promoting interactions that go beyond mere description of field experiences; reflecting on each other's ideas; integrating different perspectives.
- Action development: developing critical personal teaching practices; addressing pedagogical challenges.

From the papers revised it was evident that the development of high-order cognitive skills required not only strong social presence, but also the ability of the teacher to guide students' collaborative efforts in the transition from content reflection to critical reflection (Kreber and Cranton, 2000). This occurred as a response to a triggering event (concrete experience), which was followed by a collective process that entailed making sense of this event (exploration), integrating different ideas into a more coherent framework (integration), and developing personal teaching practices (resolution) (Garrison et al., 2000).

In turn, the *teaching presence* was connected with the social and cognitive components in online teaching and learning practices. Examples of such presence were associated with the design and facilitation of online teaching and learning as well as to broader issues related to the pedagogical approach. The factors influencing the establishment of a teaching presence according to the papers reviewed include:

- Pedagogical approach: learner at the center; teacher as a facilitator; knowledge integration.
- Learning design: flexible; personalized; contextualized; social; formative; use of appropriate tools and technology.

Table 1 Issues leading to impact related to the social, cognitive and teaching presence (according to the papers reviewed).

<i>Presence</i>	<i>Key issues</i>
Social presence	• Belongingness • Cohesiveness • Optimal level of presence by instructors and learners • Interactivity
Cognitive presence	• Participation • Concrete experience • Contextualization • Conceptualization
Teaching presence	• Action development • Pedagogical approach • Learning design • Facilitation

Adapted from Carrillo and Flores (2020).

- Facilitation: clear expectations; appropriate questions; acknowledgment of cultural issues; timely, constructive, specific and detailed feedback; students' attitude and commitment.

Contextualized (e.g. based on everyday teaching situations), personalized (e.g. targeting individual differences), social (e.g. resulting from the interaction with others), formative (e.g. focusing on the process), and integrated (e.g. considering both formal and informal ways of learning) practices were influential to the effectiveness of the online teaching and learning process. Added to this was an active and thoughtful participation of all the members involved in the online learning environment in the achievement of effective practices. As such the need for a comprehensive and solid view of online pedagogy in the context of teacher education is advocated by acknowledging the complex interactions between the components involved in the online teaching and learning process (e.g. instructors, students, task and technology factors) (Borba et al., 2018; Mumford and Dikilitaş, 2020), including instructors' and students' previous experiences and dispositions toward online teaching and learning tools, to develop effective educational experiences that address their different (and changing) roles and evolving identities.

Table 1 synthesizes the main issues leading to impact connected with the three presences before the outbreak of the pandemic.

Online teaching and learning practices and factors influencing impact in teacher education in the era post-covid

Practices influencing impact in online teacher education

Characteristics of teaching and learning practices

The literature analyzed shows that online teaching and learning practices that used appropriate digital tools and were accessible, authentic, collaborative, democratic and personalized supported impact in online teacher education environments.

Use of appropriate digital tools

When remote teaching became the forced normal, technology tools allowed teachers and learners to continue their education. The literature analyzed provides examples of how digital resources impacted learning through both synchronous and asynchronous formats and with different purposes. Video conferencing was often undertaken using Zoom (Ghazi-Saidi et al., 2020; Kalloo et al., 2020; Quezada et al., 2020; Weaver et al., 2021), BlackBoard Collaborate (la Velle et al., 2020; Quezada et al., 2020), Microsoft Teams (la Velle et al., 2020) and Google Meets (la Velle et al., 2020; Sutiah et al., 2020), either through synchronous (live) or asynchronous (recorded) sessions. Additional digital resources used by the educators in the studies examined were mainly asynchronous in nature and provided opportunities for collaborative learning (e.g. Flipgrid for stimulating discussion around a given topic; Google Docs, Canvas or Padlet for working collaboratively and sharing documents) as well as for communication (e.g. Email and WhatsApp for communicating with trainee teachers) and self-directed development issues (e.g. mini-MOOC for developing specific skills; Google Classroom for sharing additional learning resources for the students) (Başal and Eryilmaz, 2021; Boltz et al., 2021; la Velle et al., 2020; Sutiah et al., 2020; Weaver et al., 2021).

Accessibility

In terms of characteristics of online teaching and learning practices influencing impact, accessibility showed to be an important prerequisite which encompassed both the design and delivery of the teaching materials and a flexible approach to the teaching and learning process (Scully et al., 2020). In regards to the former, Turchi et al. (2020) found that making sure that digital resources were accessible through institutional systems of the local public libraries in both university and school settings was important to achieve impact. This was also evident in Scully et al. (2020), who in addition highlighted that a flexible perspective that included team teaching, supported students' interaction and allowed teachers to respond to students' questions in real time in synchronous

activities. Within this second perspective, Coolican et al. (2020) also demonstrated how flexibility in terms of approaches to teaching and learning (e.g. teaching practices, theories underlying teachers' work) was relevant to effectively manage the virtual environment.

Authentic learning

When it comes to authentic learning, the literature examined showed how specific teaching dynamics and activities contributed to student teachers' engagement and meaningful learning. O'Brien et al. (2020) revealed how small group discussion in breakout rooms delivered meaningful episodes of experiential student learning and resulted in increased efficiency and engagement. Breakout rooms driving impact included clear expectations for timed student interactions and assignation of appropriate collaborative tasks (Kidd and Murray, 2020; Metscher et al., 2021). Quezada et al. (2020), Turchi et al. (2020), Weaver et al. (2021) also used small-scale discussion boards and digital tools to share ideas and build opportunities for sense making, which provided students with the opportunity to interact and engage with course content. This was often preceded by a triggered event such as the observation of videos (Ellis et al., 2020; Kidd and Murray, 2020; la Velle et al., 2020; Metscher et al., 2021), the engagement with readings about teaching (Metscher et al., 2021) or the use of authentic cases in online educational settings (Willems et al., 2021), which provided students with realistic opportunities and challenges to understand course content, support self-regulatory practices and enhance motivation.

Collaboration and collaborative learning

Small group discussion allowed student teachers to build learning in collaboration with the others (O'Brien et al., 2020; Turchi et al., 2020). Through breakout rooms and other collaborative digital tools (e.g. chat room features, Google docs), participants were able to share ideas, build on one another's responses and learn from each other (Başal and Eryilmaz, 2021; Turchi et al., 2020; Weaver et al., 2021). Collaboration in times of COVID-19 pandemic, however, was present across different educational stakeholders, beyond the boundaries of the classroom and with other purposes besides learning. In a college setting, collaboration increased opportunities to work and communicate within different departments (e.g. school management, administration and teachers) to address the challenges of the virtual environment (Coolican et al., 2020). In the context of teacher education programs, collaboration was also established through community-building activities aimed to promote a sense of belonging among students to deal with issues such as anxiety or loneliness (Metscher et al., 2021).

Democratic approach to teaching and learning

Online teaching and learning practices after the irruption of the pandemic derived toward a more democratic approach in which the dynamics of power between teacher and student blurred and more explicitly acknowledged the figure of the student as a key component for teachers' development and teaching and learning impact. Turchi et al. (2020), Quezada et al. (2020), for example, highlighted how teachers regularly asked their students for feedback after their lessons as an opportunity to voice their opinions and learn more to improve their practices. Cutri et al. (2020) also described classroom situations in which students' technological expertise surpassed educators', thereby contributing to revert traditionally assigned teaching and learning roles (see section **The culture of "care"** for a complementary perspective on the dynamics of power).

Personalized practices and additional characteristics contributing to impact

By building relationships and connecting more deeply with their students, teachers provided a more personalized approach to their teaching and made more explicit the tacit pedagogical practices from their face-to-face teaching (Scull et al., 2020). This was also evident in van der Spoel et al. (2020), who provided an example of how the use of technology guided and modeled a more differentiated and personalized education that increased students' motivation. Additional factors contributing to learning and engagement in online environments included just-in-time (i.e. that responds to learners' need in a timely manner) and relevant content (i.e. precise, concise and built on key essential concepts) (Boltz et al., 2021; Scull et al., 2020).

Reflexivity on practice in the practicum

Student teachers had to rapidly adapt their practicum to the new guidelines provided by their respective teacher education programs. The literature analyzed portrays cases of practicum that transitioned into a virtual mode following the practices developed by schools (e.g. Coolican et al., 2020; Quezada et al., 2020) and other situations in which the lack of possibilities for completing the practicum forced teacher educators to rethink the ways of compensating for the practical experience developed in educational institutions (Kalloo et al., 2020; la Velle et al., 2020). In relation to the former, Coolican et al. (2020) described how the teacher candidates at the secondary education levels were able to conduct their practicum virtually, but were required to adapt their strategies and activities (e.g. project design, teaching sequences, interviews, classroom observations, evaluation, and reflection during and post-practicum) to the new mode of instruction. This situation was similar to the one presented by Quezada et al. (2020) whose course projects adapted to the new circumstances through activities that included the delivery of a digital portfolio of learning artifacts targeting the required teaching standards. This was accompanied by increased flexibility in the process, both in terms of a more holistic approach to assessment and extended the deadlines for assignments.

With a focus on assessment, Ó Grádaigh et al. (2021) described how emergency virtual observation during the pandemic facilitated an authentic insight into classroom teaching and was the preferred assessment option in comparison to other means used (e.g. professional conversations, lesson planning, portfolio and written reflections). This method also allowed timely feedback, which provided student teachers with the opportunity to reflect on the class. It seems evident, then, that a practicum in the new virtual space also required a practice-based approach with a reflective component, idea that is further explored by Kidd and Murray (2020). The educators in this study accommodated students' learning about practice, even if they were no longer in practice and, in doing so, they demonstrated how the new digital places allowed reflexivity on practice. The new spaces emerged in the absence of the practicum allowed meaningful adaptations beyond the initial emergency responses which at times exceeded the expectations of the educators.

In regards to the cases in which the irruption of the pandemic made impossible the completion of a part of the practicum, alternative forms of assessment such as the development of a teaching video with an accompanied lesson plan (Kalloo et al., 2020) and the discussion of prospective teachers' trajectory against teacher standards through online tutorials (la Velle et al., 2020) were implemented. In both cases, however, the affectation of the lockdown on the practicum was limited (two-week practicum at the end of the second semester and one-third of the practicum, respectively).

Focusing on the opportunities and challenges of online teaching placement during the COVID-19, the prospective teachers depicted by Sepulveda-Escobar and Morrison (2020) perceived their experience as an occasion to learn to work with different online platforms and to reflect on the socio-cultural context surrounding school students and their families during the pandemic. However, they highlighted the lack of direct interaction with the learners, the absence of a "live" teaching experience, their distractions in the household and the limited internet access of their pupils as important disadvantages for their teaching practice. The absence of practicum "face-to-face" experiences in the school was also an important drawback pointed out in O'Brien et al. (2020) in a study focused on physical education teacher education. In a moment when the future of school closures due to COVID-19 remained uncertain, upholding the value of school practicum in the process of learning to teach and the need for suitable modes of learning and assessment in practical-based learning subjects like physical education were considered key factors that needed to be addressed.

Fig. 1 synthesizes the main characteristics of online teaching and learning practices, including the practicum, after the outbreak of the pandemic.

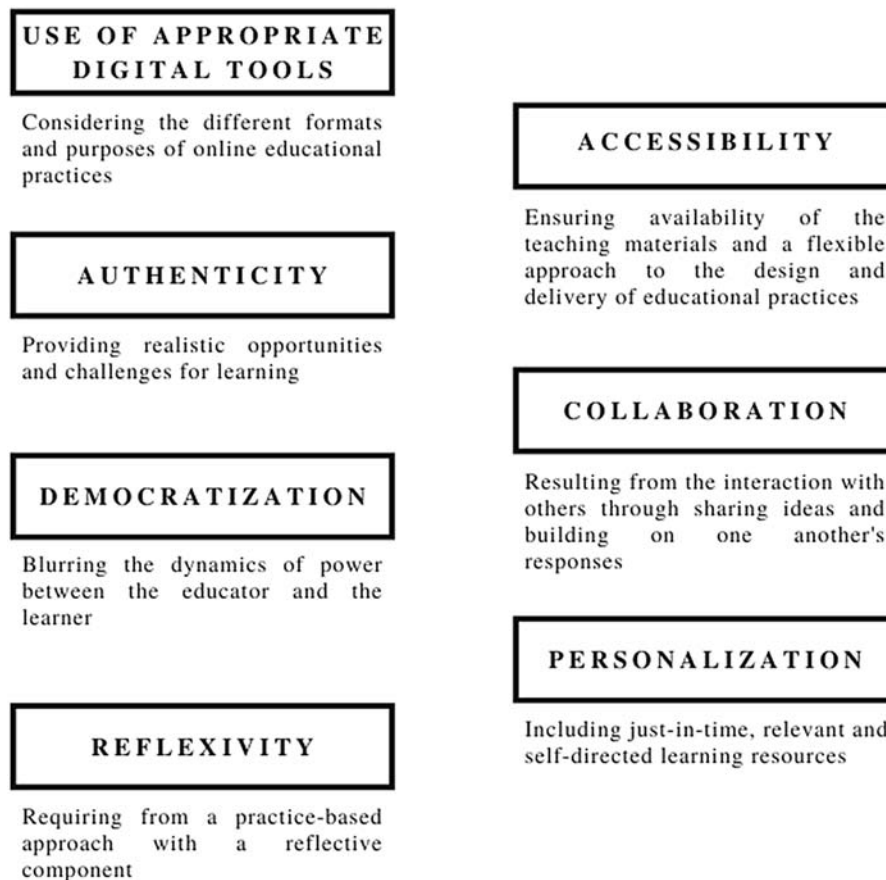


Fig. 1 Characteristics of online teaching and learning practices after the COVID-19 pandemic (according to the papers reviewed).

Factors influencing impact in online teacher education

Personal and contextual elements

The literature analyzed highlighted personal and contextual factors influencing the ways in which educators and students adapted to remote education after the irruption of the pandemic. From a personal perspective, previous knowledge, skills or experience in navigating digital tools and online teaching and learning processes impacted the effectiveness with which educators and learners transitioned into the virtual modality (Ghazi-Saidi et al., 2020; Turchi et al., 2020; van der Spoel et al., 2020). In addition, Cameron-Standerford et al. (2020) found how these personal factors affected faculty's emotional responses to the crisis situation and their ability to deliver a caring but rigorous course. Different yet complementary perspectives on the effect of personal factors in the adaption to online education were provided by Cutri et al. (2020), Bařal and Eryilmaz (2021), Turchi et al. (2020), respectively. On the one hand, limited referents in terms of previous knowledge or expertise related to remote education resulted in educators experiencing further difficulties in the adaption to the new reality and adopting more traditional or teacher-centered pedagogies (Bařal and Eryilmaz, 2021; Cutri et al., 2020), either derived from the replication of previous classroom teaching (Bařal and Eryilmaz, 2021) or from the strain that the crisis situation caused on teacher educators' philosophies and practices (Cutri et al., 2020). On the other hand, Turchi et al. (2020) pointed out how despite feeling confident in their knowledge about digital tools, the participating educators in this research were exhausted by the need to consider the "digitizing" factor to their planning (p. 10).

When it comes to contextual factors, appropriate infrastructure and support were recurrent themes that affected the success of online teaching and learning practices in the virtual medium. A proper infrastructure including good internet connection or access (Coolican et al., 2020; Ghazi-Saidi et al., 2020; Ozkaral and Bozyigit, 2020; Sutiah et al., 2020), equipment (Coolican et al., 2020; Ghazi-Saidi et al., 2020; la Velle et al., 2020; Ozkaral and Bozyigit, 2020; Sutiah et al., 2020; Zapata-Garibay et al., 2021), systems (König et al., 2020; la Velle et al., 2020) and space (Zapata-Garibay et al., 2021) showed to be important to cope with remote education at both the university and school contexts. Also, support from the university for teacher educators in both senses technological assistance and adequate training for their professional development (Coolican et al., 2020; Kalloo et al., 2020; O'Brien et al., 2020; Ozkaral and Bozyigit, 2020) was key for thriving and succeeding in the new online space. These digital gaps were especially evident in public universities (Zapata-Garibay et al., 2021), where limitations in terms of internet access or technological tools prevented them from working effectively in the online environment. In addition, the analyzed literature pointed out that certain assumptions by educational institutions (Cutri et al., 2020) or limited (or lack of) family's support with the younger children (Dvir and Schatz-Oppenheimer, 2020) contributed to the digital gap. Cutri et al. (2020), for example, showed how assumptions that university students and professors all have equitable access to the needs of online teaching and learning affected the effectiveness of the educational process. Dvir and Schatz-Oppenheimer (2020), in turn, highlighted how parents' ability to support their children in learning deepened the gaps in students' achievements.

The culture of "care"

During the pandemic, the socio-emotional needs of educators and students and the responses in terms of empathetic engagement, relationship building and emotional support came to the fore. Faculty members perceived the need to maintain the rigor of their already established course objectives; however, they felt even more important to deliver a caring yet rigorous course (Cameron-Standerford et al., 2020). Examples of the literature analyzed showed how the caring dimension of teaching, exacerbated for the situation that both educators and students were going through, was naturally embedded in the process of learning to teach and became a prerequisite for thriving also in the academic dimension. Some studies show how faculty set regular meetings in small groups and/or on an individual basis to support professional and academic progress and address general concerns around well-being (Kidd and Murray, 2020; Murray et al., 2020; Quezada et al., 2020). These aimed to know how their students were coping at both personal and professional levels with the consequences of the pandemic and to provide them with appropriate support (Quezada et al., 2020).

Additional common practices that addressed students' concerns such as anxiety or isolation and contributed to a sense of belonging and community included modeling a responsive classroom through community-building activities and building relationships and connecting with the students at a social level (Metscher et al., 2021; O'Brien et al., 2020; Scull et al., 2020). Through care-based pedagogical practices, teachers provided an empowering space for both students and educators and made a positive difference for their students that tackled issues beyond their socio-emotional needs (e.g. increased students' motivation and commitment to improving their sense of confidence for a subject; willingness to try challenging tasks; engagement in aspects of learning they do not like) (Sequeira and Dacey, 2020). From a different perspective, empathetic communication at the university level also allowed staff and students to keep informed, motivated and positive (Kalloo et al., 2020).

In times of COVID-19, shared vulnerability and uncertainty between the educator and the learner influenced the dynamics of care, control and power inherent in educational relationships (Murray et al., 2020). The teacher educators who participated in this research highlighted how, by respecting distance and under the exceptional circumstances imposed by the pandemic, they modeled caring relationships and did not feel the need to control. Instead of just trying to appear professional and in control in the new online space, the participants built human and collaborative relationships with their students. Similarly, Cutri et al. (2020) described experiences of teachers sharing power in their classrooms with students whose technological skills were above their own, thereby gaining a sense of empathy for them.

Challenges in the transition to online education

Issues related to increased teaching workload, poor communication and interactivity with and between the students, and assessment of students' performance were recurrent themes that challenged the normal functioning of online lessons.

The intensification of work by educators after the irruption of the pandemic was a common topic among the articles analyzed. The participants reported increased workload (Coolican et al., 2020; Kidd and Murray, 2020; van der Spoel et al., 2020; Zapata-Garibay et al., 2021) due to: the additional efforts required to prepare and adapt their lessons and methodologies for remote learning, and; the sense of having to be available to students and families at all times and through different means (Kidd and Murray, 2020), contributing to the blurring of work-home boundaries (Kidd and Murray, 2020; Turchi et al., 2020; Zapata-Garibay et al., 2021).

Limited communication and engagement was also identified as a primary challenge in the context of online education (e.g. Boltz et al., 2021). The students described in Ghazi-Saidi et al. (2020) found it challenging to stay engaged and reported a lack of interactive activities which affected their motivation levels. Kidd and Murray (2020) also highlighted little interaction as an obstacle for impact in online education, but particularly when large group work was used in the new space. In turn, van der Spoel et al. (2020) reported limitations in terms of interaction with and between students, but also pointed out unexpected positive experience by introvert students, who were more present compared to a classroom context. Despite the availability of better tools for communication, it was clear that relationships between educators and students resented from the forced adaptations due to the health crisis (Zapata-Garibay et al., 2021). This trend was confirmed by Scull et al. (2020), who also added that while their participating student teachers may possess adequate levels of digital skills, they may lack the competences required to effectively communicate in the online environment. This concurs with Turchi et al. (2020) findings, which highlighted the lack of students' competences to support meaning making through discussion even though they were fluent in digital tools, making it clear that the knowledge of such tools may not be enough for effectively using them in the online environment.

The challenges brought by limited communication and interaction often exacerbated when student teachers did not show their faces (Turchi et al., 2020). This raised issues of attendance and presence (Turchi et al., 2020) and other ethical and moral issues. Dvir and Schatz-Oppenheimer (2020), for example, showed how the anonymity and absenteeism derived from hiding their faces allowed some learners to be verbally abusive during the sessions, forcing teachers to show leadership and responsibility to avoid major consequences. In a different sense, the limitations in terms of interaction in the online environment brought to light worries about how to monitor and assess the learning that was happening in the discussions (Sutiah et al., 2020; Turchi et al., 2020). Educators reported adjustment to new assessment procedures (Kalloo et al., 2020; König et al., 2020) (e.g. involving or limiting final examinations, respectively); however, in the transition, they also raised concerns on the need to further learn and use different forms of formative assessments (Cutri et al., 2020; König et al., 2020).

Table 2 summarizes the most relevant factors impacting online teaching and learning practices after the pandemic.

Online teacher education: implications for future practice and research

This article has provided a review of the literature on online teaching and learning practices before and after the COVID-19 pandemic. Based on earlier work focusing on the practices prior to the pandemic, an analysis was undertaken in light of the social, cognitive and teaching presence according to the CoI. Findings revealed the blurring nature of the components driving to the three presences. The ability of educators and learners to engage in relationships showed to play a pivotal role to meaningful educational experiences. At the time, its interplay with components of teaching and cognitive nature was essential in achieving teaching and learning impact.

Table 2 Factors influencing impact in online teacher education (according to the papers reviewed).

Factors	Issues included
Personal factors	• Previous knowledge, skills or experience in navigating digital tools and online educational processes
Contextual factors	• Appropriate infrastructure (internet connection or access; equipment; systems; physical space) • Support from the university (technological assistance; adequate training) • Family's support with the younger children • Assumptions by educational institutions
Caring dimension	• Regular meetings in small groups or on an individual basis to support students' emotional needs or concerns around well-being • Community-building activities and connecting with the students at a social level to model a responsive classroom • Empathetic communication at the university level to keep staff and students informed, motivated and positive
Challenges encountered	• Increased teaching workload • Poor communication and interactivity with and between the students • Assessment of students' performance

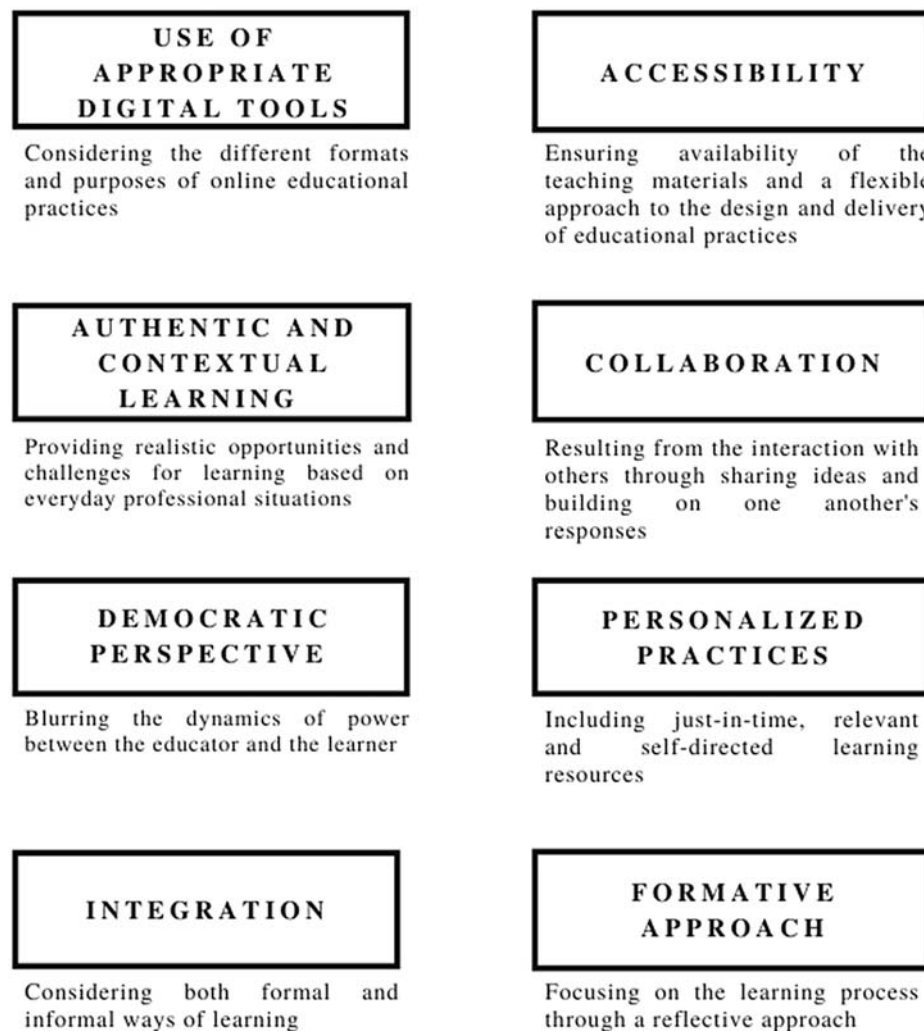


Fig. 2 Integration of the characteristics of online teaching and learning practices prior and after the pandemic (according to the papers reviewed derived).

The literature analyzed in both periods have pointed out that online teaching and learning practices that use appropriate digital tools and are accessible, authentic, collaborative, democratic, formative, personalized and integrated support impact in online teacher education environments (see Fig. 2).

The social dimension of teaching, which was found central in any online teaching and learning process before the COVID-19 crisis, was strongly impacted by the uncertainty and vulnerability generated by the pandemic. This exacerbated the caring dimension of teaching and opened the door to a more democratic approach. This caring dimension was embedded in the process of teaching and learning related to the socio-emotional needs of educators and students (Metscher et al., 2021; O'Brien et al., 2020; Scull et al., 2020), although it had an impact beyond this domain. At the time, the exceptional circumstances imposed by the pandemic and the role of the learner as a pillar for impact enhanced a more democratic perspective to teaching and learning (Cutri et al., 2020). Also, the pandemic supported the personalization of online teaching and learning, highlighting the relevance of practices that are accessible, just-in-time and relevant. Other studies have complemented this view by providing data that shows the important role of informal and self-directed learning when it comes to online educational practices (Robinson and Rusznyak, 2020; Trust and Whalen, 2020). The personalized approach to teaching and learning and its characteristics demonstrate the relevance of the online learning that is embedded in the day-to-day of educators' practices to help them tackle the challenges of the moment.

With reference to the practicum developed during the pandemic, findings from this review have emphasized the importance of providing a practice-based approach with a reflective component (Kidd and Murray, 2020), in line with Flores and Gago (2020). In addition, the perspective provided has highlighted its authentic and formative nature, and contributed with data on both the kinds of activities that facilitate reflexivity on practice (e.g. digital portfolio of learning artifacts accompanied by reflection in the process) and specific issues about how the process should be developed (e.g. timely feedback to enhance reflection after the observation of virtual classrooms). Besides activities and processes aiming to enhance reflexivity on practice, a number of options such as helping

associate teachers with technology; communicating and providing support to students virtually, and; learning new online digital platforms have been highlighted as useful practices carried out in the virtual practicum during the pandemic (Donitsa-Schmidt and Ramot, 2020; Sepulveda-Escobar and Morrison, 2020; van Nuland et al., 2020).

Although practical experience in schools is an essential component of initial teacher education programs (Allen and Wright, 2014), the above issues may be considered to address some of the constraints that many authors have associated with the emergency virtual practicum (e.g. limited contextualization; constraints of the mentoring by school teachers; absence of a live teaching experience; lack of direct interaction with the learners) (Flores and Gago, 2020; Prata-Linhares et al., 2020; Robinson and Rusznyak, 2020; Sepulveda-Escobar and Morrison, 2020). In this sense, the data provided in this article has supported the assumption that the appropriate technology and resources should facilitate a quality virtual practicum. However, more research is needed to understand how to take advantage of the affordances of technology (Moyo, 2020), as well as of the processes and resources needed to rethink traditional practicum practices and develop successful experiences in the online environment.

Contextual and personal factors impacting the development of online teaching and learning practices should not be ignored. The data derived from the analysis of the literature prior to the pandemic found that contextual factors such as limited access to technology and internet as well as differences in digital literacy that are deeply embedded in social, economic and cultural context (Beaunoyer et al., 2020), affected the normal development of teaching and learning practices and exacerbated inequalities and exclusion. The data from the post-pandemic time have confirmed these findings but have also pointed out additional issues impacting this process such as certain assumptions by educational institutions (e.g. that students and educators have the appropriate infrastructure to carry out their online learning and teaching), the support by the stakeholders involved in online education (e.g. educational institutions or families—the latter when there are young kids involved) and the limited communication and interaction derived from pupils hiding behind their cameras (see also Flores and Gago, 2020). Also, the latest findings have highlighted personal factors that need to be considered if effective online teaching and learning practices are pursued. The data analyzed has confirmed that appropriate knowledge, skills and experience in navigating digital tools and online processes by educators are needed to avoid teacher-centered pedagogies that fail to deliver an effective online education. However, it has been shown that the knowledge of such tools and processes may not be enough to ensure success in the online environment. This suggests the need to instill these practices in a comprehensive pedagogy of online education that takes into account the pedagogical possibilities associated with such digital tools and resources.

The data presented in this article has several implications for future online teaching and learning practices in the context of teacher education. First, it highlights the need to go beyond emergency online practices and provide an evidence-based approach to online teaching and learning that acknowledges the particularities of this pedagogy and its implications. This needs practices that move beyond a simplistic replication process of face-to-face format. Second, it emphasizes the need for a pedagogical approach that takes into consideration the suitability and affordances of the various technologies or digital resources to make sure that these help learners to effectively communicate and achieve the desired results in the virtual space. Also, this should acknowledge the different roles (e.g. teachers as facilitators; students as active and inclusive learners) and competences (e.g. emphasis on the social and caring components of learning) required to teach effectively in the online higher education environment (Ní Shé et al., 2019) and the nature of an online teaching and learning process that is inherently collaborative, formative, contextual, democratic and personalized. Third, it puts accent on the importance to support educators in meeting the emotional needs of their learners through acknowledging and supporting the development of the socio-emotional competences required for teaching (Hadar et al., 2020). This may include investing in high-quality teacher education programs, matching educators' current needs with professional learning opportunities, supporting mentoring, and enhancing collaboration between educators and other stakeholders involved in education, among other initiatives (Darling-Hammond and Hyler, 2020). Fourth, it points out the urgency to rely on support structures of higher education institutions in partnership with governments and organizations to achieve equity (Carrillo and Flores, 2020; Coolican et al., 2020; Sequeira and Dacey, 2020) and to move toward a more inclusive and democratic approach to learning “which empowers teachers to work on behalf of their rights and the rights of their students” (Kalloo et al., 2020, p. 461). Fifth, it highlights the importance to know how to best develop online educational experiences that are procedural in nature. For example, there is the need to understand how to take advantage of the affordances of technology (Moyo, 2020) as well of the possibilities of assessment tools in online environments (ElTahir Osman, 2020) to ensure a successful development of practical skills (e.g. those related to the practicum or other procedural areas such as Physical Education). And, finally there is a need to go beyond an instrumental approach to online teaching and learning and to include into the equation its ethical, political and pedagogical dimensions. For example, student teachers and teacher educators' roles and responsibilities in the online educational experiences and their implications for the process of learning to teach have to be considered. This also encompasses the consideration of issues of power and control over teaching and learning, i.e. elements connected with not only what and how but also who and why, which are questions that stand at the core of the development of professional knowledge and identity.

References

- Allen, J.M., Wright, S.E., 2014. Integrating theory and practice in the pre-service teacher education practicum. *Teach. Teach. Theor. Pract.* 20, 136–151. <https://doi.org/10.1080/13540602.2013.848568>.
- Anderson, T., Rourke, L., Garrison, R., Archer, W., 2001. Assessing teaching presence in a computer conferencing context. *J. Asyn. Learn.* 5 (2), 1–17. <https://doi.org/10.24059/olj.v5i2.1875>.
- Başal, A., Eryılmaz, A., 2021. Engagement and affection of pre-service teachers in online learning in the context of COVID 19: engagement-based instruction with web 2.0 technologies vs direct transmission instruction. *J. Educ. Teach.* 47 (1), 131–133. <https://doi.org/10.1080/02607476.2020.1841555>.

- Beaunoyer, E., Dupéré, S., Guitton, M.J., 2020. COVID-19 and digital inequalities: reciprocal impacts and mitigation strategies. *Comput. Hum. Behav.* 111, 106424. <https://doi.org/10.1016/j.chb.2020.106424>.
- Boltz, L.O., Yadav, A., Dillman, B., Robertson, C., 2021. Transitioning to remote learning: lessons from supporting K-12 teachers through a MOOC. *Br. J. Educ. Technol.* <https://doi.org/10.1111/bjet.13075>.
- Borba, M.C., Santana de Souza, A., Rangel, H., 2018. Interactions in virtual learning environments: new roles for digital technology. *Educ. Stud. Math.* 98, 269–286. <https://doi.org/10.1007/s10649-018-9812-9>.
- Bozkurt, A., Sharma, R.C., 2020. Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian J. Distance Educ.* 15 (1), i–vi. <https://doi.org/10.5281/zenodo.3778083>.
- Bryman, A., 2012. *Social Research Methods*, fourth ed. Oxford University Press, Oxford.
- Cameron-Standerford, A., Menard, K., Edge, C., Bergh, B., Shayter, A., Smith, K., VandenAvond, L., 2020. The phenomenon of moving to online/distance delivery as a result of COVID-19: exploring initial perceptions of higher education faculty at a Rural Midwestern University. *Front. Educ.* 5, 1–11. <https://doi.org/10.3389/educ.2020.583881>.
- Carrillo, C., Flores, M.A., 2020. COVID-19 and teacher education: a literature review of online teaching and learning practices. *Eur. J. Teach. Educ.* 43 (4), 466–487. <https://doi.org/10.1080/02619768.2020.1821184>.
- Carrillo, C., Flores, M.A., (in review). *Online Teaching and Learning Practices in Teacher Education in Times of COVID-19 Pandemic: What Can We Learn From the Research Literature?*
- Coolican, M., Borrás, J.C., Strong, M., 2020. Argentina and the COVID-19: lessons learned from education and technical colleges in Buenos Aires Province. *J. Educ. Teach.* 46 (4), 484–496. <https://doi.org/10.1080/02607476.2020.1802204>.
- Creswell, J.W., 1998. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage, Thousand Oaks.
- Cutri, R.M., Mena, J., Whiting, E.F., 2020. Faculty readiness for online crisis teaching: transitioning to online teaching during the COVID-19 pandemic. *Eur. J. Teach. Educ.* 43 (4), 523–541. <https://doi.org/10.1080/02619768.2020.1815702>.
- Darling-Hammond, L., Hyler, M.E., 2020. Preparing educators for the time of COVID ... and beyond. *Eur. J. Teach. Educ.* 43 (4), 457–465. <https://doi.org/10.1080/02619768.2020.1816961>.
- Donitsa-Schmidt, S., Ramot, R., 2020. Opportunities and challenges: teacher education in Israel in the Covid-19 pandemic. *J. Educ. Teach.* 46 (4), 586–595. <https://doi.org/10.1080/02607476.2020.1799708>.
- Doyle, S., 2007. Member checking with older women: a framework for negotiating meaning. *Health Care Women Int.* 28 (10), 888–908. <https://doi.org/10.1080/07399330701615325>.
- Dvir, N., Schatz-Oppenheimer, O., 2020. Novice teachers in a changing reality. *Eur. J. Teach. Educ.* 43 (4), 639–656. <https://doi.org/10.1080/02619768.2020.1821360>.
- Ellis, V., Steadman, S., Mao, Q., 2020. "Come to a screeching halt": can change in teacher education during the COVID-19 pandemic be seen as innovation? *Eur. J. Teach. Educ.* 43 (4), 559–572. <https://doi.org/10.1080/02619768.2020.1821186>.
- ElTahir Osman, M., 2020. Global impact of COVID-19 on education systems: the emergency remote teaching at Sultan Qaboos University. *J. Educ. Teach.* 46 (4), 463–471. <https://doi.org/10.1080/02607476.2020.1802583>.
- Erlandson, D.A., Harris, E.L., Skipper, B.L., Allen, S.D., 1993. *Doing Naturalistic Inquiry. A Guide to Methods*. Sage, London.
- Evens, M., Larmuseau, C.H., Dewaele, K., Craesbeek, L.V., Elen, J., Depaepe, F., 2017. The effects of a systematically designed online learning environment on preservice teachers' professional knowledge. *J. Digit. Learn. Teach. Educ.* 33 (3), 103–113. <https://doi.org/10.1080/21532974.2017.1314779>.
- Flores, M.A., Gago, M., 2020. Teacher education in times of COVID-19 pandemic in Portugal: national, institutional and pedagogical responses. *J. Educ. Teach.* 46 (4), 507–516. <https://doi.org/10.1080/02607476.2020.1799709>.
- Flores, M.A., Barros, A., Veiga Simão, A.M., Pereira, D., Flores, P., Fernandes, E., Costa, L., Ferreira, P.C., 2021a. Portuguese higher education students' adaptation to online teaching and learning in times of the COVID-19 pandemic: personal and contextual factors. *High. Educ.* 1–20.
- Flores, M.A., Machado, E.A., Alves, P., Vieira, D.A., 2021b. Ensinar em tempos de COVID-19: um estudo com professores dos ensinos básico e secundário em Portugal. *Rev. Port. Educ.* 34 (1), 5–27.
- Garrison, R., Anderson, T., Archer, W., 2000. Critical inquiry in a text-based environment: computer conferencing higher education. *Internet High Educ.* 2 (2–3), 87–105. <https://doi.org/10.1080/02607476.2020.1755205>.
- Ghazi-Saidi, L., Criffield, A., Kral, C.L., McKelvey, M., Obasi, S.N., Vu, P., 2020. Moving from face-to-face to remote instruction in a higher education institution during a pandemic: multiple case studies. *Int. J. Technol. Educ. Sci.* 4 (4), 370–383. <https://ijtes.net/index.php/ijtes/article/view/169>.
- Hadar, L.L., Ergas, O., Alpert, B., Ariav, T., 2020. Rethinking teacher education in a VUCA world: student teachers' social-emotional competencies during the Covid-19 crisis. *Eur. J. Teach. Educ.* 43 (4), 573–586. <https://doi.org/10.1080/02619768.2020.1807513>.
- Ham, V., Davey, R., 2005. Our first time: two higher education tutors reflect on becoming a "virtual teacher". *Innov. Educ. Teach. Int.* 42 (3), 257–264. <https://doi.org/10.1080/01587910500168017>.
- Hatch, J.A., 2002. *Doing Qualitative Research in Education Settings*. State University of New York Press, Albany, New York.
- Hou, H., 2015. What makes an online community of practice work? A situated study of Chinese student teachers' perceptions of online professional learning. *Teach. Teach. Educ.* 46, 6–16. <https://doi.org/10.1016/j.tate.2014.10.005>.
- Hrmiak, A., 2010. Online learning community development with teachers as a means of enhancing initial teacher training. *Technol. Pedagog. Educ.* 19 (1), 47–62. <https://doi.org/10.1080/14759390903579265>.
- Huberman, A.B., Miles, M.B., 1994. Data management and analysis methods. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, first ed. Sage, Thousand Oaks, pp. 228–244.
- Kaloo, R.C., Mitchell, B., Kamalodeen, V.J., 2020. Responding to the COVID-19 pandemic in Trinidad and Tobago: challenges and opportunities for teacher education. *J. Educ. Teach.* 46 (4), 452–462. <https://doi.org/10.1080/02607476.2020.1800407>.
- Kidd, W., Murray, J., 2020. The Covid-19 pandemic and its effects on teacher education in England: how teacher educators moved practicum learning online. *Eur. J. Teach. Educ.* 43 (4), 542–558. <https://doi.org/10.1080/02619768.2020.1820480>.
- Komninou, I., 2017. A case study of the implementation of social models of teaching in E-learning: "the social networks in education", online course of the Inter-orthodox Centre of the Church of Greece. *TechTrends* 62 (2), 146–151. <https://doi.org/10.1007/s11528-017-0247-4>.
- König, J., Jäger-Biela, D.J., Glutsch, N., 2020. Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *Eur. J. Teach. Educ.* 43 (4), 608–622. <https://doi.org/10.1080/02619768.2020.1809650>.
- Kreber, C., Cranton, P.A., 2000. Exploring the scholarship of teaching. *J. High Educ.* 71 (4), 476–496. <https://doi.org/10.2307/2649149>.
- la Velle, L., Newman, S., Montgomery, C., Hyatt, D., 2020. Initial teacher education in England and the Covid-19 pandemic: challenges and opportunities. *J. Educ. Teach.* 46 (4), 596–608. <https://doi.org/10.1080/02607476.2020.1803051>.
- Li, Z., 2011. Learners' reflexivity and the development of an E-learning community among students in China. *Res. Learn. Technol.* 19 (1), 5–17. <https://doi.org/10.1080/09687769.2010.548505>.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage Publications, Newbury Park.
- Metscher, S.E., Tramantano, J.S., Wong, K.M., 2021. Digital instructional practices to promote pedagogical content knowledge during COVID-19. *J. Educ. Teach.* 47 (1), 121–124. <https://doi.org/10.1080/02607476.2020.1842135>.
- Moyo, N., 2020. Covid- 19 and the future of practicum in teacher education in Zimbabwe: rethinking the "new normal" in quality assurance for teacher certification. *J. Educ. Teach.* 46 (4), 536–545. <https://doi.org/10.1080/02607476.2020.1802702>.

- Mumford, S., Dikilitaş, K., 2020. Pre-service language teachers reflection development through online interaction in a hybrid learning course. *Comput. Educ.* 144, 103706. <https://doi.org/10.1016/j.compedu.2019.103706>.
- Murray, C., Heinz, M., Munday, I., Keane, E., Flynn, N., Connolly, C., Hall, T., MacRuairc, G., 2020. Reconceptualizing relatedness in education in "distanced" times. *Eur. J. Teach. Educ.* 43 (4), 488–502. <https://doi.org/10.1080/02619768.2020.1806820>.
- Ni Shé, C., Farrell, O., Brunton, J., Costello, E., Donlon, E., Trevaskis, S., Eccles, S., 2019. Teaching Online is Different: Critical Perspectives from the Literature. Project Report. Dublin City University, Dublin. <https://doi.org/10.5281/zenodo.3479402>.
- Ó Grádaigh, S., Connolly, C., Mac Mahon, B., Agnew, A., Poole, W., 2021. An investigation of emergency virtual observation (EVO) in initial teacher education, in Australia and Ireland during the COVID-19 pandemic. *Ir. Educ. Stud.* <https://doi.org/10.1080/03323315.2021.1916561>.
- O'Brien, W., Adamakis, M., O'Brien, N., Onofre, M., Martins, J., Dania, A., Makopoulou, K., Herold, F., Ng, K., Costa, J., 2020. Implications for European Physical Education Teacher Education during the COVID-19 pandemic: a cross-institutional SWOT analysis. *Eur. J. Teach. Educ.* 43 (4), 503–522. <https://doi.org/10.1080/02619768.2020.1823963>.
- Ozkara, T.C., Bozyigit, R., 2020. Social studies and geography teacher candidates' views on coronavirus (COVID 19) and online education process. *Rev. Int. Geogr. Educ. Online* 10 (3), 467–484. <https://rigeo.org/volume-10-issue-3-summer-2020/>.
- Prata-Linhares, M.M., da Silva, T., Lopes-Jr, D.L., Zukowsky-Tavares, C., 2020. Social distancing effects on the teaching systems and teacher education programs in Brazil: reinventing without distorting teaching. *J. Educ. Teach.* 46 (4), 554–564. <https://doi.org/10.1080/02607476.2020.1800406>.
- Quezada, R.L., Talbot, C., Quezada-Parker, K.B., 2020. From bricks and mortar to remote teaching: a teacher education program's response to COVID-19. *J. Educ. Teach.* 46 (4), 472–483. <https://doi.org/10.1080/02607476.2020.1801330>.
- Robinson, M., Rusznyak, L., 2020. Learning to teach without school-based experience: conundrums and possibilities in a South African context. *J. Educ. Teach.* 46 (4), 517–527. <https://doi.org/10.1080/02607476.2020.1800408>.
- Ryan, G.W., Bernard, H.R., 2000. Data management and analysis methods. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, second ed. Sage, Thousand Oaks, pp. 769–802.
- Satar, H.M., Akcan, S., 2018. Pre-service EFL teachers' online participation, interaction, and social presence. *Lang. Learn. Technol.* 22 (1), 157–183 doi:10125/44586.
- Scull, J., Phillips, M., Sharma, U., Garnier, K., 2020. Innovations in teacher education at the time of COVID19: an Australian perspective. *J. Educ. Teach.* 46 (4), 497–506. <https://doi.org/10.1080/02607476.2020.1802701>.
- Sepulveda-Escobar, P., Morrison, A., 2020. Online teaching placement during the COVID-19 pandemic in Chile: challenges and opportunities. *Eur. J. Teach. Educ.* 43 (4), 587–607. <https://doi.org/10.1080/02619768.2020.1820981>.
- Sequeira, L., Dacey, C.M., 2020. The COVID-19 diaries: identity, teaching, and learning at a crossroads. *Front. Educ.* 5, 1–12. <https://doi.org/10.3389/feduc.2020.586123>.
- Sutiah, S., Slamet, S., Shafqat, A., Supriyono, S., 2020. Implementation of distance learning during the COVID-19 in faculty of education and teacher training. *Cypriot J. Educ. Sci.* 15 (5), 1204–1214. <https://un-pub.eu/ojs/index.php/cjes/article/view/5151>.
- Trust, T., Whalen, J., 2020. Should teachers be trained in emergency remote teaching? Lessons learned from the COVID-19 pandemic. *J. Technol. Teach. Educ.* 28 (2), 189–199. <https://www.learnedtechlib.org/p/215995>.
- Turchi, L.B., Bondar, N.A., Aguilar, L.L., 2020. What really changed? Environments, instruction, and 21st century tools in emergency online English language arts teaching in United States schools during the first pandemic response. *Front. Educ.* 5, 1–13. <https://doi.org/10.3389/feduc.2020.583963>.
- van der Spoel, I., Noroozi, O., Schuurink, E., van Ginkel, S., 2020. Teachers' online teaching expectations and experiences during the Covid-19-pandemic in The Netherlands. *Eur. J. Teach. Educ.* 43 (4), 623–638. <https://doi.org/10.1080/02619768.2020.1821185>.
- van Nuland, Mandzuk, D., Tucker, K., Cooper, T., 2020. COVID-19 and its effects on teacher education in Ontario: a complex adaptive systems perspective. *J. Educ. Teach.* 46 (4), 442–451. <https://doi.org/10.1080/02607476.2020.1803050>.
- Weaver, J.C., Matney, G., Goedde, A.M., Nadler, J.R., Patterson, N., 2021. Digital tools to promote remote lesson study. *Int. J. Lesson Learn. Stud.* 10 (2), 187–201. <https://doi.org/10.1108/JLLS-09-2020-0072>.
- Willems, P.P., Gonzalez-DeHass, A.R., Powers, J.R., Musgrove, A., 2021. The role of authentic teaching cases and mastery approach goals in online pre-service teachers' self-regulated learning. *Educ. Technol. Res. Dev.* 69, 1003–1023. <https://doi.org/10.1007/s11423-021-09972-9>.
- Yeh, Y., 2010. Integrating collaborative PBL with blended learning to explore preservice teachers' development of online learning communities. *Teach. Teach. Educ.* 26, 1630–1640. <https://doi.org/10.1016/j.tate.2010.06.014>.
- Zapata-Garibay, R., González-Fagoaga, J.E., González-Fagoaga, C.J., Cauich-García, J.R., Plascencia-López, I., 2021. Higher education teaching practices experience in Mexico, during the emergency remote teaching implementation due to COVID-19. *Front. Educ.* 6, 1–14. <https://doi.org/10.3389/feduc.2021.628158>.

Rethinking teacher education in pandemic times: lessons from and for Singapore

Ee-Ling Low, National Institute of Education, Nanyang Technological University, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	710
Singapore's education phases and initiatives	711
Learn for life: remaking pathways initiative (2019–present)	711
Comparisons with Shanghai, China and Estonia	712
Differentiated instruction (DI)	712
Assessments for and as learning	713
A national framework for teaching and learning	714
Balancing academic rigor with inspiring the joy in learning	715
Preparing future-ready teachers for future-ready learners	716
Main program components	716
Teacher professionalism	716
Rigorous programs	717
Evolving pedagogies	718
Multiple perspectives	718
Professional practice	718
Teacher education in pandemic times and beyond	718
Conclusion	718
References	719

Introduction

Education is a prime mover to help nations prepare their citizens for work and life in a globalized era where the future is fraught with uncertainty. In recent times, we have been confronted with multiple crises (e.g., the Covid-19 pandemic) and disruptions (e.g., Fourth Industrial Revolution [4IR]) forcing us to rethink and change the way we live and work (Gleason, 2018). The world has also become increasingly interconnected through the affordances brought about by global cyber networks (Schwab, 2016). While the Covid-19 pandemic has revealed that our societal structures (e.g., economics, education, social welfare, finance, health) are not as robust as we thought they were, the unprecedented pace of technological advancements (Ryder, 2018) is fast causing traditional jobs to become obsolete and irrelevant. Consequently, current jobs need to be redesigned, and new jobs need to be created, and these have multiple consequences which span economic, legal, ethical and social domains to name a few (Ryder, 2018). The pace of digital transformations needed to be further accelerated in order to keep learning and working going when most were confined to their homes due to imposed lockdowns. Though the pandemic has caused many disruptions, it has also opened up many opportunities to rethink and transform our current education systems in order to cater to the new realities confronting us.

Singapore's education system has been undergoing progressive enhancements in order to ensure that its citizens will be prepared to not just survive but also to thrive in the face of an uncertain future. This article will describe the country's significant education phases since the young nation state gained its independence in 1965, broadly known as the Survival-driven (1965–1978), Efficiency-driven (1979–1997), Ability-driven (1997–2011), and Values-driven, Student-centric education phases (2011–2018). These educational initiatives outline the underpinning philosophy of Singapore's education system as being in a constant state of evolution in order to ensure both local and global relevance. Two initiatives that have gained significant traction in local and international literature are the *Thinking School, Learning Nation* (henceforth TSLN launched in 1997) and the *Teach Less, Learn More* (henceforth TLLM launched in 2005) initiatives. In 2019, a new initiative known as *Learn for Life: Remaking Pathways* (Ministry of Education, Singapore [MOE] 2019) was introduced. This new phase aims at promoting lifelong learning for both teachers and students and to relook some of the longstanding education policies such as the over-emphasis on assessments and academic streaming.

A comparative study will be made with Shanghai in China and Estonia in order to provide an insight into how other jurisdictions are rethinking their teacher education policies and programs to gear up for the global challenges faced. Finally, the article will describe how Singapore's sole teacher education institute, the National Institute of Education (NIE), an autonomous institute housed within the renowned and research-intensive Nanyang Technological University (NTU) is rethinking and reviewing its teacher education programs in order to ensure that its teachers are capable of preparing their learners to be ready for the multiple challenges and complexities of their future workplace and society.

Singapore's education phases and initiatives

Singapore's educational phases are best viewed as a continuum (Goh, 1997), where one phase builds on the success of the previous, and each phase corresponds to a particular focus in the nation's developmental journey (see Table 1).

The TSLN initiative was launched by then Prime Minister Mr. Goh Chok Tong in 1997 during the Ability-Driven phase (1997–2011). TSLN aimed to create an education system that nurtured creativity, critical thinking and a passion for lifelong learning in all its learners (Goh, 1997). This initiative complemented Singapore's economic restructuring efforts, such as creating a knowledge-based economy in order to develop a thinking and inquiring workforce. Through the TSLN initiative, Ministry of Education, Singapore (MOE) reduced curriculum content by about 30%, reviewed its assessment system, enhanced teacher training, and encouraged the sharing of best practices and ideas (Ong, 2018b; The Straits Times, 2016). From 2000, teachers received at least 100 h of professional development (PD) annually to help them grow as teaching professionals career-long (The Straits Times, 2016).

Building on the TSLN initiative, the TLLM initiative was introduced in 2005 by Prime Minister Mr. Lee Hsien Loong. TLLM aimed to improve the quality of teaching and the teaching profession, enhance student learning through further curriculum content reduction, engage learners holistically through the use of innovative learning pedagogies and equip students with lifelong learning skills. TLLM's goal was to improve learning through the improved quality of interactions between teacher and students through pedagogies that would cater to the specific needs, abilities and interests of learners so that they would be better prepared to thrive in a dynamic 21st-century workplace and society. This initiative progressed into the Student-centric, Values-driven education phase (2012–2018), and, like any other national initiative, achieved success through a fidelity in its implementation by all education stakeholders (i.e., MOE, NIE and the schools).

Learn for life: remaking pathways initiative (2019–present)

In 2019, Singapore officially launched its next education phase as the country recognized that its citizens need to be developed into lifelong learners who are positioned to face an uncertain future that is fast evolving and filled with many disruptions. The *Learn for Life: Remaking Pathways* (LfL) (MOE, 2019) education phase builds on previous iterations of lifelong learning initiatives amid the massive changes in the local and global jobs sector, technological advancements and new ways of learning and working.

Table 1 First four education phases of the Singapore education system (Heng, 2015; Low, 2017; NIE/NTU, 2017; The Straits Times, 2016).

Education phases	Main features
Survival-driven (1965–1978)	- Aimed to reduce high school dropout rates and to raise school participation and enrollment rates by building primary and secondary schools, doubling enrollment rates from 1965 to 1970 - Developed a national curriculum and introduced the primary school leaving examination (PSLE), a national high-stakes examination that provides the academic pathway for entry into secondary education
Efficiency-driven (1979–1996)	- Carried out the recommendations made by the Goh Keng Swee report of 1979 to improve the education system - Aimed to reduce the high drop-out rates of students with differentiated academic streams being introduced at primary 4 and secondary 1 respectively - Developed the 6-4-2 education system (6-year primary, 4-year secondary and 2-year pre-university) as the default route with extended pathways possible - Instituted English as the medium of instruction (MOI) in all Singapore schools in 1987 - Established independent schools that gave schools greater autonomy - Standardized curriculum and assessment with all learners taking the PSLE, and the GCE O- and A-levels for entry into secondary, pre-university and tertiary education
Ability-driven (1997–2011)	- Focused on the development of every child to reach his/her fullest potential - Launched TSLN in 1997 - Launched TLLM in 2005
Values-driven, Student-centric (2011–2018)	- Integrated both the Ability-driven and the Values-driven, Student-centric phases to serve as the foundations for the Framework for 21st century competencies and student outcomes (MOE, 2018a) - Ensured that academic and life skills, values and character development in students were taught in schools while considering the diversity of learner abilities and talents - Envisioned all learners to be developed into confident persons, self-directed learners, active contributors and concerned citizens - Equipped students with the values, beliefs, character, knowledge, skills and competencies necessary for the 21st century

Then Minister for Education, Ong Ye Kung introduced LfL in 2018 and posited that lifelong learning is “a value, an attitude and a skill that our students need to possess, and it is fundamental in ensuring that education remains an uplifting force in society” (Ong, 2018a, para. 17). Skills and content are still important as Ryder (2018) observed, but they are only “the icing on the cake”; rather, the more important attribute is the aptitude for lifelong learning that must be underpinned by “strong cognitive skills, such as literacy and numeracy” (para. 9). Lifelong learning has become the principal consideration of the Singapore education system to help students to be “resilient, adaptable and global in their outlook”, who “must leave the education system still feeling curious and eager to learn, for the rest of their lives” (Ong, 2018a, para. 22). The system also recognizes that all learners go through different stages of learning, where each stage requires a significant transition and pupils need an adequate runway to adapt to the new demands. Thus, a further diversification of education pathways was introduced. This initiative also introduced subject-based banding (SBB), where academic streaming is abolished in favor of banding by subjects, a reduction of overall assessments across different levels of study and an urgent push to recalibrate the balance between upholding academic rigor while still igniting the joy of learning in all students.

Comparisons with Shanghai, China and Estonia

Singapore’s education transformation efforts are compared with Shanghai, China and Estonia. Shanghai, China was chosen because it is based in Asia, just like Singapore, and Estonia was selected as a jurisdiction that has had a sharp rise in its educational rankings as measured by the Organization for Economic Cooperation and Development (OECD)’s internationally benchmarked test of student achievement, the Program for International Student Assessment (PISA). All three jurisdictions have fared well in the PISA 2018 study with China (represented by Beijing, Shanghai, Zhejiang and Jiangsu) topping the league table for achievements in Reading, Mathematics and Science, Singapore emerging second in all three areas, and Estonia emerging top five for Reading and Science.

Differentiated instruction (DI)

The concept of DI will be explored via the various approaches taken by each of the three educational jurisdictions. It is important to first articulate what our understanding of DI is. Tomlinson (2005), credited with being a leading scholar in the field, defines DI as “a philosophy of teaching that is based on the premise that students learn best when their teachers accommodate the differences in their readiness levels, interests and learning profiles.” The main objective of DI is to ensure that every student’s learning potential is maximized. In terms of DI strategies, Singapore introduced full SBB, Shanghai China employs the Shanghai Mastery Approach, and Estonia’s education system is focused on providing equal opportunities for all and ensuring that an individual’s lifelong development is taken care of. Each of these approaches will be further elaborated upon.

Singapore’s SBB aims to meet the learning pace and needs of each student and to challenge them at academically appropriate levels (Chan, 2019). While all primary schools are already adopting the SBB system, secondary schools will embrace full SBB by 2024. SBB is meant to take the place of academic streaming that was introduced in the 1980s. In the early post-independent years, the academic streaming system was introduced to solve the issue of high student attrition rates by providing each student with an academic stream that is positioned to match their learning abilities and pace. Learners were placed in one of three academic streams as early as the upper primary level and continues to the secondary level of schooling. These academic streams do not take into account an individual’s talents or achievements in particular subjects since the stream determines the overall structure and duration of one’s academic program. It was, however, possible for a student who performed well to switch academic streams if they performed well in a particular year of study. Academic streaming has achieved its overall objective to decrease student attrition rates (Tan, 2019), however, the designation of an academic stream that is based solely on academic performance has the inadvertent effect of making less academically able students feel less confident about themselves.

SBB is introduced as an enhancement and a viable alternative to the longstanding academic streaming system as it focuses on students’ strength in individual subjects. Students are placed into one of three bands within each subject, namely, G1, G2, G3 which are customized to suit their learning abilities and aptitudes, thereby enhancing student performance through motivating each learner to learn according to their interest and aptitude in each subject (Wang, 2019). Chan (2019) expressed that SBB helps cater to the “Seven S” elements of education, namely, “specific abilities, suitability, self-concepts, stigma, student-centric schools, societal attitudes and subjective well-being.” The pedagogical customization of SBB considers the specific abilities of a learner, the student-subject match, improves students’ self-confidence, removes the social stigma associated with academic streaming, shifts societal attitudes from the ability-driven to a growth mindset (an essential element of lifelong learning), and protects the well-being of students from academic-related mental stress. While skeptics have terms SBB as placing “new wine in old bottles,” Chan argues that when students’ learning encounters are positive, “students experience high subjective well-being, which results in higher motivation, better performance and more pro-social behaviors.” This will eventually enable students to lead successful and fulfilling lives and allow them to function effectively both independently and interdependently as they enter the society and workforce in the future.

Being one of the four regions that participated in PISA 2018, Shanghai (with Beijing, Jiangsu and Zhejiang being the others) presents an interesting point of comparison as it is also considered the financial hub of China. The Shanghai Mastery Approach is one where a whole classroom learns a single concept in depth through the use of standardized textbooks and does not proceed to the next concept until every student has understood it (Neha, 2017). Classes are repetitive and progress slowly with each lesson, allowing teachers to correct learners’ mistakes and misunderstandings. While the approach primarily focuses on Mathematics, its

principles are also applied in the learning of languages, science and the humanities, especially in the early years to level up those who had a disadvantaged start to their education (Williams, 2019). Teachers possess the belief that all children are capable of learning anything if learning is presented in a precise manner. It therefore caters to disadvantaged students while reinforcing fundamental concepts in the academically more able students.

The constraint of such an approach is in the need for high-quality teachers to successfully implement this approach in a class of learners with mixed prior knowledge. All teachers need to be masters of their subjects and pedagogical practices, and they teach a specific subject and only that subject (Neha, 2017). The success of this approach is further hinged on the teacher's ability to simplify and scaffold learning and to constantly manage progression, assessments and maintain consistency while avoiding differentiation in instruction (Newell, 2019). This approach has its advantage by addressing the negative stigma of being a low progress learner (noted in Singapore's academic streaming system) as the whole class must understand the concept before lessons may progress to the next topic. A low progress learner is not isolated for differentiated instruction as the teacher has to ensure that the entire class has mastered a particular concept before moving on to the next topic. As the achievement gaps are tightened in the early years, the need for differentiated instruction becomes non-consequential. Interestingly, the Mastery Approach has been adopted by the United Kingdom in the teaching and learning of Mathematics in some primary and secondary schools (Boylan et al., 2018).

Described as one of the newest international education powerhouses (Butrymowicz, 2016), Estonia is on the upward trajectory, as evidenced in the PISA 2018 results. Estonia's education goal is embodied in their Lifelong Learning Strategy which provides everyone in Estonia "with learning opportunities that are tailored to their needs and capabilities throughout their whole lifespan, for them to maximize opportunities for dignified self-actualization within society" (Ministry of Education and Research, Estonia 2014, p. 3). The Estonian education system is rooted in its quest to provide equality for all, its philosophy of student readiness and individual development for life. Favoring a teacher-centered, learner-supported strategy, Estonia focuses on content more than soft skills (Hatch, 2017). In approaching customized learning and differentiated instruction, it adopts a two-pronged approach to support all learners through development-appropriate learning and individual development plans, thereby raising the achievement of all learners through equal learning opportunities. Moving from age-appropriate to development-appropriate learning aids, individual learners progress to a level that suits their pace of development.

In the Basic and Upper-Secondary Schools Act (Teataja, 2013), teachers are required to monitor the development of learners and adjust instruction and/or learning activities to help students cope with the demands of their academic progression and achievement levels. The Estonian Ministry of Education prescribed that all students must have an individual conversation with a teacher in each academic year to help them develop an individual developmental plan. The goals and means to get to the goals are discussed and the teacher serves to guide, coach and monitor their students' development. This purpose-driven conversation inculcates ownership toward learning and the emotional support rendered by teachers provides students with a safe and nurturing environment to cope with the rigorous demands of academic work. Since most teachers teach the same students for at least three years, they develop strong bonds and can cater their lessons to meet their student's ability without the need to academically stream them by classes. The sense of ownership is very strong within the teaching fraternity (European Agency for Special Needs and Inclusive Education, n.d.). The Estonian individually customized approach stands in contrast to Singapore's differentiated instruction by subjects and Shanghai's mass mastery approach.

Assessments for and as learning

Assessments for and as learning play important roles in teaching and learning by helping students to become self-directed learners who can engage in lifelong learning. Hence, assessments must be enacted with a clarity of purpose, thorough planning and careful implementation in order to provide teachers and learners the appropriate feedback to address gaps in learners' understanding.

In Singapore, the reduction in the emphasis on school-based assessments was introduced in 2018 by former Education Minister Mr. Ong Ye Kung at the MOE Schools Work Plan Seminar. He announced the removal of all weighted assessments for Primary One and Two to allow young learners time to experience the larger school environment, nurture their inquisitiveness, develop mastery in fundamental literacy and numeracy, and to encourage the holistic development of the youngest students in the formal school system (Ong, 2018a). This initiative is backed up by research evidence in child development where the reduction of testing in the lower primary years helps to build executive functioning skills, self-regulation and a growth mindset, which in turn can help develop learners' lifelong learning dispositions (Devine and Hughes, 2014; Portilla et al., 2014). The use of formative assessments rather than mid-year and end-year summative assessments helps to build confidence in young learners where the assessment rubrics and developmental goals guide learners toward attaining mastery in literacy and numeracy skills. Moreover, young learners develop an intrinsic motivation to learn through constant progression and mastery of developmentally appropriate literacy and numeracy skills (Reeve et al., 2018). The removal of mid-year examinations in the transitional schooling years of Primary Three and Five, and Secondary One and Three was also announced (Ong, 2018a). The reduction on the number of examinations is meant to recalibrate the over-emphasis on assessments and to allow students to discover the joy of learning and for teachers to adapt to the different learning dispositions of each student.

Echoing the reforms in Singapore, alternative assessments to develop students holistically (Tan and Ng, 2018) have been advocated by educators since 2014 in Shanghai, well-known for its stressful *Gao Kao* (university entrance examinations). Grounded on the principle to "uphold quality-oriented education and focus on students' moral, intellectual, aesthetic and all-round development" (Shanghai Municipal Education Commission, 2014), Shanghai's assessment reforms are well-received as educators see that the refinements will holistically equip students with an international outlook. This assessment reform may be seen in their "educational appraisal", a variant of formative assessment which includes more formative assessment elements and provides

teachers with the opportunities to provide more qualitative feedback. This stands in contrast to conventional summative assessments that provide limited quantitative feedback opportunities. To alleviate examination-related stress, Chinese students can sit for their foreign language papers twice. However, the adoption of alternative assessments is not uniform throughout Shanghai and will need time to be scaled up.

For a country that is a forerunner in Science, Technology, Engineering and Mathematics (STEM) education and has a high proportion of female graduates entering into technology-related sectors, Estonia adopts a fairly conventional assessment system. However, it is interesting to note that assessment results are seen as proxies of student readiness for progressing to the next stage in education. Student readiness is gauged by national examinations, sampled-based examinations, and regular classroom formative and summative assessments. Satisfactory levels of attainment must be obtained by students in curricular subjects and creative work in order to qualify for graduation. Every student undergoes a mandatory yearly developmental interview and schools are held accountable designing appropriate interventions to level up students with unsatisfactory year-end performance. With its robust and free pre-primary school program, 6-year-olds' readiness is assessed via a primary school readiness evaluation before they enter primary school. With the exception of school-leaving assessments that are highly stressful, students are tested only once every three years; other national assessments in elementary and lower secondary schools are strictly sampling tests that are to provide schools, districts and the Ministry a gauge to students' learning and to help guide policy formulation and planning (Ministry of Education and Research, Estonia 2019).

A national framework for teaching and learning

In conducting a literature survey of China and Estonia's education systems, a national-level framework for teaching and learning was not found equivalent to what has been articulated within Singapore and named the Singapore Teaching Practice (STP) (MOE, 2018b). It should be noted, however, that Singapore is a small country with a population of about 5.69 million (Department of Statistics Singapore (DOS), 2021) with a stable political system where the ruling party has been in power since 1965. This allows for policies to be enacted quickly system-wide. Even within Singapore, there exist a few different education frameworks developed. The STP, which is co-developed by MOE and NIE aims to serve as a framework to "develop our teachers, codify our practices, and document the perspectives and practices that support our educators in facilitating learning within and outside the classroom" (Liu and Lim, 2018, p. 6). It details how effective teaching and learning practices are achieved in Singapore classrooms. According to MOE's Director-General of Education Mr. Wong Siew Hoong, STP is "the reference point for us to examine our own professional practice, think deeply about it and improve it, in all the different aspects of the work that we do" (Chan, 2017, para. 2).

Schematically represented by the country's national flower the orchid, the STP (See Fig. 1) is anchored by the Singapore Curriculum Philosophy (SCP) and articulates Singapore educators' beliefs about teaching and learning, and the roles of learners. At its core are Pedagogical Practices (PP, see Fig. 2) which reflect key teaching processes that guide teaching actions to ensure that every student is an engaged learner. The teaching processes are divided into 24 teaching areas that are taken into consideration when

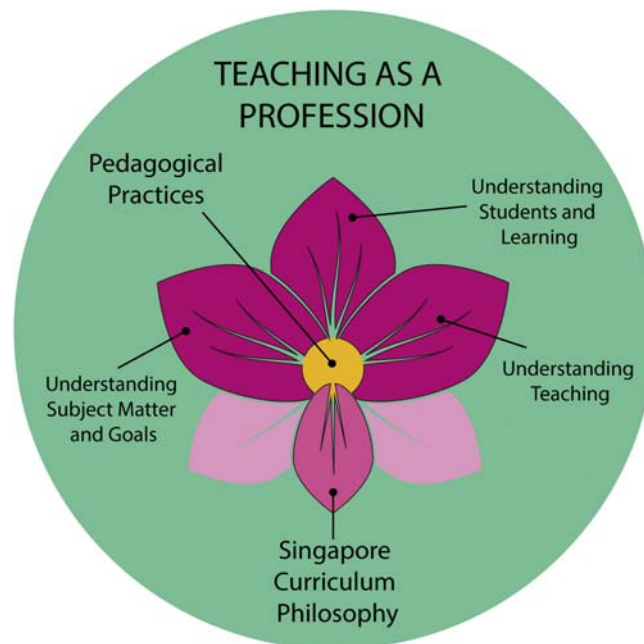


Fig. 1 The Singapore teaching practice (MOE, 2018b).

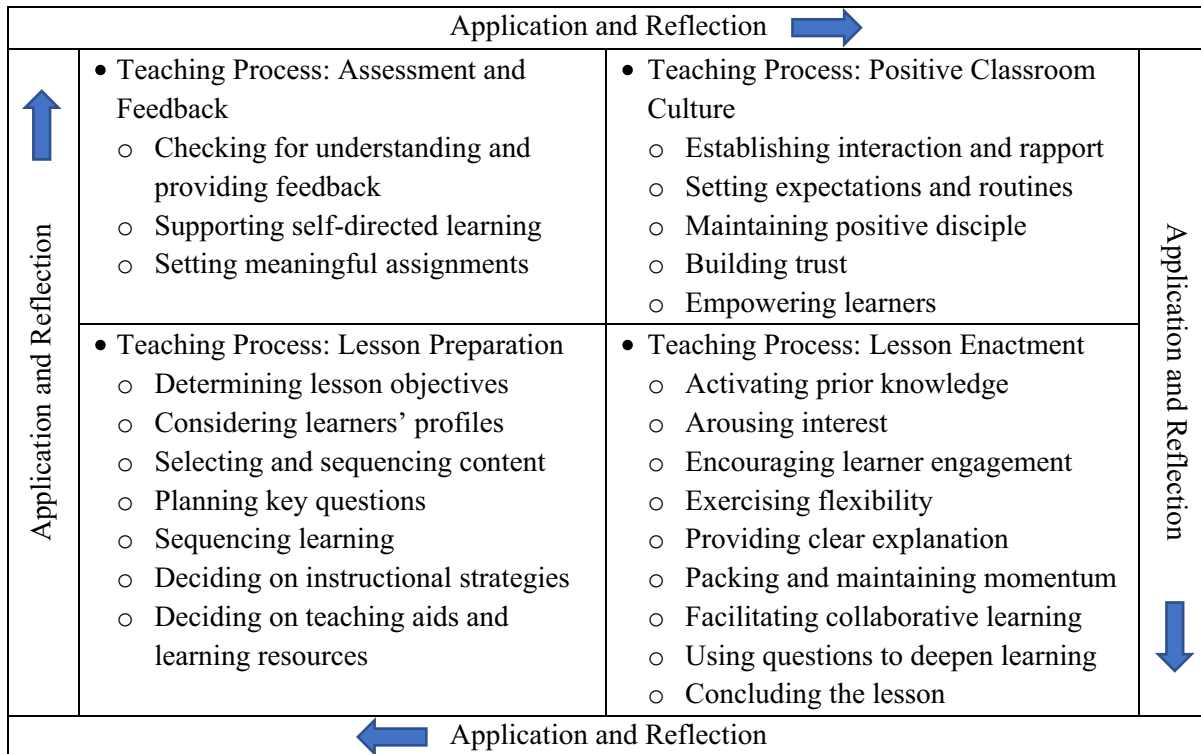


Fig. 2 Pedagogical practices (MOE, 2018b).

planning, enacting and reviewing effective lessons. PP is informed by three Knowledge Bases, namely, understanding subject matter and goals, understanding students and learning, and understanding teaching. The knowledge bases are drawn from both local and international research, which provide the evidence-base for contextualized local practice.

The components of STP “are situated within a larger ecosystem that supports teachers in developing and deepening their craft as competent professionals, collaborative learners, and transformative leaders” (Liu and Lim, 2018, p. 6). A course, entitled Pedagogical Practices (PP), introduces student teachers to the “why, what and how of teaching” and prepares student teachers for their clinical practice component, the teaching assistantship. The PP course provides pre-service teachers a clear link between theory and practice and takes place concurrently during their teaching assistantship where student teachers spend a day on campus and four days out in school. The four teaching processes in the STP are part of the assessable components of the clinical practice stints for pre-service teachers. In-service teachers are provided with PD workshops and courses to help them to mentor novice teachers using the same STP framework so that a common language exists across the continuum from pre-service to beginning teaching. This common framework is envisioned to help teachers across their career-span to “continually reflect on their roles, draw upon theories and research to deepen their understanding of learning, think systematically about their practice, and adapt their teaching to support their students’ learning” (Liu and Lim, 2018, p. 6). An online STP portal was also developed and provides a digital repository of relevant studies for each teaching area to reinforce and amplify good teaching practices that is backed by empirical local and international research evidence.

While having a national framework is an advantage in providing a clearer picture of how teaching and learning may be effectively achieved and in helping to provide a common pedagogical meta-language between key stakeholders, there is much to be learned from Shanghai’s mastery approach and Estonia’s personalized learner-centric approaches earlier described. Students from these systems are reported in newspaper interviews with school leaders and teachers (Blatchford, 2015; Hatch, 2017; Kangur, 2018) to be joyful, engaged learners and this intangible achievement is supported empirically by results of the PISA 2018 survey (OECD, 2020a,b).

Balancing academic rigor with inspiring the joy in learning

Based on the latest PISA 2018 findings, 59% of Chinese students surveyed and 70% of Estonian students reported being satisfied with their lives, with the OECD average being 67% (OECD, 2019b). 98% of Chinese students and 89% of Estonian students reported being happy sometimes or always happy. Interestingly, Shanghai students seem not to indicate pressure faced via the Mastery Approach (Blatchford, 2015). Once students can understand the lesson, the strive toward high performance is reported to be satisfying for both teachers and students. While Estonian students are reported in the PISA 2018 results to be happy with their lives, in a qualitative survey, a principal voiced concern that learners and teachers may not be happy (Kangur, 2018). To address this

concern, Hatch (2017) describes a project that gives learners autonomy and choice about how they want to spend their leisure time and to further develop their system of hobby schools and providing youth with work opportunities has been undertaken. The results about the well-being of the Singapore students were not listed in the PISA 2018 results. While it should be noted that well-being is a “multi-dimensional construct and that its measurement requires covering different domains, not just overall life satisfaction” (OECD, 2019b, p. 261), it should also be recognized that the different domains are interrelated and the PISA 2018 surveys “connect school life with the broader ecosystem in which students live—the family, their peers, the community—and provide information on the development of 15-year-old students” (OECD, 2019a, p. 47).

Singapore’s policymakers are keenly aware that academic success does not necessarily equate to positive holistic development. Then Minister of Education in 2018, Mr. Ong Ye Kung announced that “a balance [is needed] between rigor and joy, and there is a fairly strong consensus that we have tilted too much to the former” (Ong, 2018a, para. 31). He further commented that “education in schools is at risk of becoming too stressful and maybe some unwinding is in order” (Ong, 2018a, para. 8). Singapore is considering three initiatives in an attempt to rebalance the overemphasis on assessments with bringing out the joy of learning.

First, through the TLSN and TLLM initiatives curriculum content had already been reduced. What is currently the focus in the LfL initiative is to reduce the emphasis of studying to the test via the reduction in overall number of examinations in non-final years of study as mentioned earlier (Ong, 2018a, para. 36). Second, curriculum time should be focused on better teaching and learning activities, rather than more drill and practice (Ong, 2018a, para. 37). Singapore could study how Shanghai’s Mastery Approach manages to instil the joy of learning while upholding academic rigor. Third, the reduction of school-based assessment loads by 25% signals to Singaporean students and their parents more time is released to help nurture more creative and innovative thinkers. In tandem, pre-service teacher preparation programs have to be reviewed and enhanced to align with the national education initiative. This will be described in the next section of this article.

Preparing future-ready teachers for future-ready learners

The concept of future-readiness has taken on a whole new dimension in recent pandemic times because every education system is struggling to keep teaching and learning going in the midst of school closures and national lockdowns. Nurturing students who are future-ready requires equipping teachers with the attributes to help their students to prepare for a complex and uncertain future.

The envisioned attributes of future-ready teachers are boldly articulated in Singapore’s vision for teacher education as being creators of knowledge, facilitators of lifelong learning, architects of learning environments, shapers of characters and leaders of educational change. To elaborate, future-ready teachers should be:

- Active knowledge creators of content and pedagogy, adding new value to the professional teaching fraternity
- Facilitators of lifelong learning
- Architects of learning environments and active agents in curating the best physical and virtual learning spaces for their learners
- Shapers of characters who possess the values and moral fiber to cultivate sound values and develop their learners to be good and useful local and global citizens
- Leaders of educational change from within their classrooms emanating to the school and systems level

These attributes are pivotal in the uncertain era where the citizens of a nation must have “an aptitude to lifelong learning” (Ryder, 2018), combined with thinking minds that are “agile, flexible and cooperative” (Gleason, 2018, p. 165). The next section describes the main program components that enable Singapore to work toward its vision of providing a forward-looking teacher education program.

Main program components

Fig. 3 shows the main components of Singapore’s pre-service teacher education program viz. building teacher professionalism, designing and enacting rigorous programs, using innovative pedagogies, providing multiple perspectives and deepening professional practice.

Teacher professionalism

Singapore’s values-based education framework (Fig. 4) is premised upon the expanding environment framework which begins with the individual “myself” emanating to “my community”, “my nation”, ending with “my world”. Accordingly, the courses offered also fall within each of these concentric circles of influence. Of significance is the SPINE of the expanding environment values education framework which is about “character and citizenship education” which helps our students to be good and moral citizens of our nation and ultimately, active, ethical and positive contributors to the world as global citizens.

In each respective sphere of influence lies programs and initiatives that endeavor to build values in our students.

- **Myself:** Pre-service teachers undergo a personal and professional development journey via the Meranti project which is named after a strong and resilient tree symbolizing their deep commitment to teaching, learning and building the future of our nation.

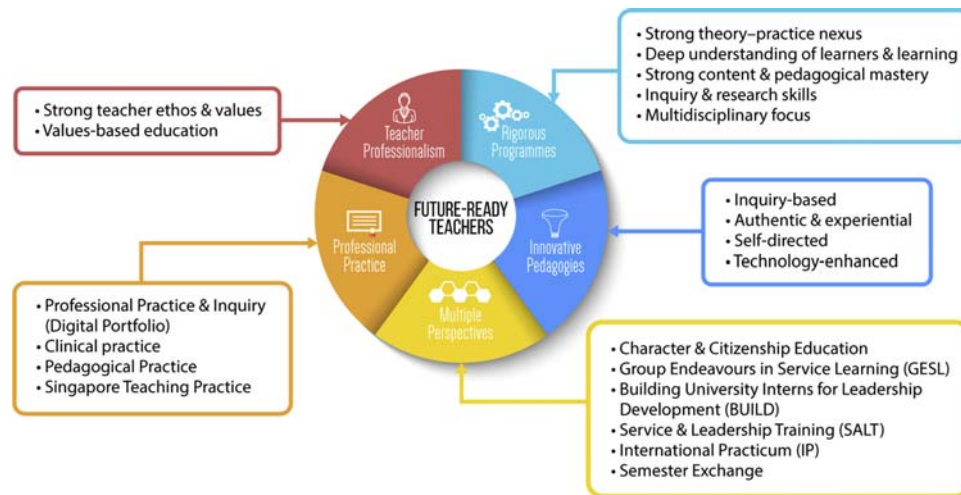


Fig. 3 Program components for preparing future-ready teachers.

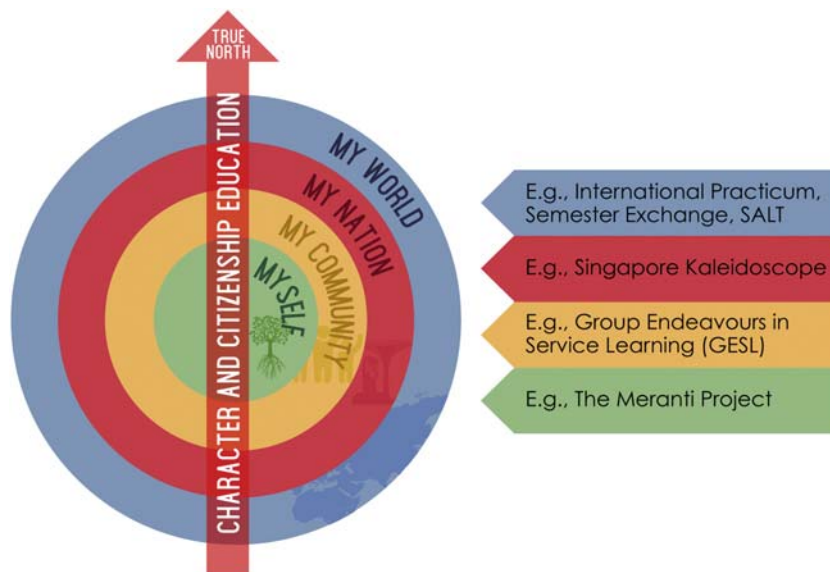


Fig. 4 Expanding environment approach in values-based education.

- **My Community:** All our student teachers undergo a mandatory community involvement project known as the “Group Endeavors in Service Learning” which helps them to work with a community partner of their choice and to serve that community through meaningful engagement.
- **My Nation:** Our students take the “Singapore Kaleidoscope” (SgK) course which uses appreciative inquiry as a pedagogical approach for our students to appreciate the diversity of our culture and heritage. The Multiculturalism course: “Appreciating and Valuing Diversity” also helps our students to appreciate Singapore as a multilingual, multi-racial and multi-cultural society.
- **My World:** Students are given multiple global experiences through semester exchanges and the international practicum components. Additionally, the post-pandemic world calls for a re-thinking of the archetype of the future-ready teacher who is able to thrive, survive and prosper in a highly volatile and uncertain world.

Rigorous programs

As a university-based teacher education program, Singapore needs to ensure a tight theory–practice nexus. This is done mainly by getting student teachers to do multiple reflections, via experiential learning, embarking on school-based research and inquiry-based projects focusing on problems of practice, enhancing the practicum stints such that there is coursework linkage via the PP course earlier described, and other pedagogical tools that allow classrooms to be virtually brought into the university. A program that provides a good showcase of the rigorous components of the pre-service program is the premier Teaching Scholars Program

(TSP) (NTU, 2019; OTE, n.d.) which endeavors to nurture teachers who will eventually have the potential to be leaders and movers of the education system, the core features of which will be described under the other program components. TSP scholars are also given the privilege of participating in the Undergraduate Research Experience on Campus (URECA) research projects where they supervised by any faculty within the university and aimed at inducting undergraduates to presenting their work at international conferences and to co-publish their work in reputable peer-reviewed journals.

Evolving pedagogies

The pre-service program adopts five pedagogical approaches to help nurture future-ready learning outcomes:

- (a) Active learning to foster a spirit of enterprise and discovery and to allow one to engage in lifelong learning through teaching and research
- (b) Appreciative inquiry where student teachers develop a greater appreciation for the subject matter at hand. For example, in the SgK course, student teachers explore and appreciate Singapore's nature and biodiversity, cultural heritage, and its position in the world. The course is assessed through producing an artifact that showcases student teachers' appreciation of local culture and heritage
- (c) Collaborative learning, where knowledge is co-constructed and facilitated by the collaborative classroom infrastructure in all tutorial rooms and virtual collaborative spaces
- (d) Flipped learning, where content is delivered asynchronously and synchronous sessions are to engage in discussion and to clarify any misconceptions
- (e) Mobile learning (or mLearning) both through the use of mobile devices and applications customized for their learning and supported by a digital library of resources for ready access anytime, anywhere

Multiple perspectives

Pre-service teachers are provided multiple local and global experiences via their local and international practicum stints, semester abroad exchange programs, local and international internships, service-learning, and leadership training opportunities.

Professional practice

The Professional Practice and Inquiry (PPI; NIE, n.d.) course is conceptualised as a digital portfolio and is designed as a learning portfolio whilst at pre-service, allowing student teachers to integrate and aggregate their learning. While serving as a showcasing portfolio about what they have learnt at pre-service during their clinical practice in schools, it is also meant to be their professional teaching portfolio when they step out into schools as novice teachers. The digital portfolio is a means to help student teachers track their growth, experiences, and developmental journey. The digital portfolio is learner-centered and gives student teachers agency and ownership over their learning. The latest iteration has learning analytic capabilities built in and offers more digital collaboration opportunities.

Teacher education in pandemic times and beyond

The Covid-19 pandemic has severely impacted the lives of individuals and societies around the world. Like many nations world-wide, there were school closures amid a national lockdown in Singapore. Singapore termed its lockdown the Circuit Breaker period from 7 April to 1 June 2020. All pre-service on-campus classes and laboratory sessions went 100% virtual and had to be delivered via online platforms like ZOOM or TEAMS. All final written examinations for non-final year student teachers were canceled with assessments re-scaled based on what had already been submitted during the course as part of formative assessments.

The Teaching Assistantship (TA) and the Final Teaching Practice (Practicum) for several hundred student teachers also had to take place in a remote, home-based learning (HBL) environment. When the circuit breaker was implemented, final year pre-service teachers had only completed five and a half weeks or 55% of their practicum on-site in schools. Their remaining four and a half weeks or 45% was completed virtually via HBL and this mixed-modality approach while proving to be highly challenging had also underscored the importance of tweaking the practicum assessment components to include the successful enactment of a virtual lesson.

Conclusion

This paper described the evolutions in Singapore's education system and attempted a comparison with two other jurisdictions with outstanding performances in the PISA 2018 surveys. It has also documented recent efforts to design and deliver a teacher preparation program that endeavors to prepare teachers capable of nurturing lifelong learners who are able to thrive in the context of a complex and uncertain future. Concrete examples of how the program had to respond to the pandemic during a national lockdown were also provided.

Reimers (2018) argues that education and schools must be designed to help individuals to develop their full potential, face the ever-evolving challenges of the future, and prepare our young to be equipped with the necessary and relevant attributes, skills, and competencies to live fruitful and fulfilling lives. As we find ourselves confronted with one of the most complex biological crises of

humankind that shows no sign of abating, it has become clear that the role of all educators around the world is to keep learning going no matter what, no matter when and no matter how. Such is the huge responsibility that rests with all teachers and the even more urgent mission confronting us is how we need to redesign our teacher education programs to adequately prepare our teachers to fulfill this essential role.

References

- Blatchford, R., 2015. Differentiation is out. Mastery is the new classroom buzzword. *Guardian*. www.theguardian.com/teacher-network/2015/oct/01/mastery-differentiation-new-classroom-buzzword.
- Boylan, M., Maxwell, B., Wolstenholme, C., Jay, T., Demack, S., 2018. The mathematics teacher exchange and "mastery" in England: the evidence for the efficacy of component practices. *Educ. Sci.* 8 (4), 202–233.
- Butrymowicz, S., 2016. Is Estonia the New Finland? With a focus on equity, the northern European country has quietly joined the ranks of the global education elite. *Atlantic*. www.theatlantic.com/education/archive/2016/06/is-estonia-the-new-finland/488351.
- Chan, A., 2017. Online platform for teachers to share and learn teaching practices. *Straits Times*. www.straitstimes.com/singapore/online-platform-for-teachers-to-share-and-learn-teaching-practices.
- Chan, D., 2019. The "Seven S" approach to subject-based banding in schools. *Straits Times A32*. www.straitstimes.com/opinion/the-seven-s-approach-to-subject-based-banding-in-schools.
- Department of Statistics, Singapore (DOS), 2021. Census of Population 2020 Statistical Release 1: Demographic Characteristics, Education, Language and Religion. Author, Singapore.
- Devine, R., Hughes, C., 2014. Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* 85 (5), 1777–1794. www.jstor.org/stable/24033017.
- European Agency for Special Needs and Inclusive Education, (n.d.). Raising the Achievement of All Learners in Inclusive Education. Brussels, European Agency for Special Needs and Inclusive Education, www.european-agency.org/sites/default/files/agency-projects/RaisingAchievement/CountryReports/Raising%20Achievement%20Estonia%20Country%20Report.pdf.
- Gleason, N.W., 2018. Singapore's higher education systems in the era of the Fourth Industrial Revolution: preparing lifelong learners. In: Gleason, N.W. (Ed.), *Higher Education in the Era of the Fourth Industrial Revolution*. Palgrave Macmillan, Singapore, pp. 145–170.
- Goh, C.T., 1997. Speech by Prime Minister Goh Chok Tong at the Opening of the 7th International Conference on Thinking on Monday, 2 June 1997, at 9.00 a.m. at the Suntec City Convention Centre Ballroom. National Archives of Singapore. www.nas.gov.sg/archivesonline/data/pdfdoc/19970602_0001.pdf.
- Hatch, T., 2017. 10 surprises in the high-performing Estonian education system. *Int. Educ. News*. internationalednews.com/2017/08/02/10-surprises-in-the-high-performing-estonian-education-system.
- Heng, S.K., 2015. Keynote Address by Mr. Heng Swee Keat, Minister for Education, at the Ministry of Education Work Plan Seminar 2015, on Tuesday, 22 September 2015 at 9.15 a.m. at Ngee Ann Polytechnic Convention Centre. Ministry of Education, Singapore. www.moe.gov.sg/news/speeches/keynote-address-by-mr-heng-swee-keat-minister-for-education-at-the-ministry-of-education-work-plan-seminar-2015-on-tuesday-22-september-2015-at-9-15am-at-nee-ann-polytechnic-convention-centre.
- Kangur, M., 2018. The Recipe for Success in the Estonian Basic Education System. *Estonian World*. estonianworld.com/knowledge/recipe-success-estonian-basic-education-system.
- Liu, W.C., Lim, Y.C., 2018. Understanding teaching, learning and learners: the Singapore teaching practice. In: Liu, W.C., Koh, C., Choy, D., Tay-Lim, J. (Eds.), *Understanding Teaching, Learning and Learners: A Guide for Singapore Teachers*. Cengage Learning Asia, Singapore, pp. 1–23.
- Low, E.-L., 2017. Singapore English. In: Low, E.-L., Hashim, A. (Eds.), *English in Southeast Asia: Features, Policy and Language in Use*. John Benjamins Publishing, Amsterdam, The Netherlands, pp. 35–54.
- Ministry of Education and Research, Estonia, 2014. Eesti elukestva õppe Strateegia 2020 [The Estonian Lifelong Learning Strategy 2020]. Ministry of Education and Research, Tallinn, Estonia.
- Ministry of Education and Research, Estonia, 2019. Pre-school, Basic and Secondary Education. Ministry of Education and Research, Estonia. www.hm.ee/en/activities/pre-school-basic-and-secondary-education.
- Ministry of Education, Singapore (MOE), 2018a. Framework for 21st Century Competencies and Student Outcomes. MOE. www.moe.gov.sg/education/education-system/21st-century-competencies.
- Ministry of Education, Singapore (MOE), 2018b. The Singapore Teaching Practice. Ministry of Education, Singapore, Singapore. www.moe.gov.sg/about/singapore-teaching-practice.
- Ministry of Education, Singapore (MOE), 2019. Learn for Life: Remaking Pathways. www.moe.gov.sg/microsites/cos2019/remaking-pathway.html.
- Nanyang Technological University (NTU), 2019. Teaching Scholars Program. National Institute of Education, Nanyang Technological University, Singapore. <https://ebook.ntu.edu.sg/20190312-tsp/full-view.html>.
- Neha, T.B., (2017). More children in the West are being taught math using China's fabled, slightly brutal "mastery" method. *Quartz*, March 24, qz.com/939208/more-children-in-the-west-are-being-taught-math-using-chinas-fabled-slightly-brutal-mastery-method.
- Newell, R., 2019. *Mastery Mathematics for Primary Teachers*. SAGE Publications, Newbury Park, CA.
- NIE/NTU, 2017. Presentation to ICSEI on Singapore's Education System, Presentation Slides.
- NIE/NTU, (n.d.). Professional Practice & Inquiry, NIE, www.nie.edu.sg/teacher-education/practicum/professional-practice-and-inquiry.
- Organisation for Economic Co-operation and Development (OECD), 2019a. PISA 2018 Assessment and Analytical Framework. Organisation for Economic Co-operation and Development, Paris, France. www.oecd-ilibrary.org/docserver/b25efab8-en.pdf?expires=1584952680&id=id&accname=guest&checksum=4FE61CC853B66EC4A1597B68AB2DEB85.
- OECD, 2019b. PISA 2018 Volume III: What School Life Means for Students' Lives. OECD, Paris, France. <http://www.oecd.org/education/pisa-2018-results-volume-iii-acd78851-en.htm>.
- OECD, 2020a. B-S-J-Z (China)—Country Note—PISA 2018 Results. OECD, Paris, France. www.oecd.org/education/2030-project/teaching-and-learning/learning/core-foundations/Core_Foundations_for_2030_concept_note.pdf.
- OECD, 2020b. Estonia—Country Note—PISA 2018 Result. OECD, Paris, France. www.oecd.org/pisa/publications/PISA2018_CN_EST.pdf.
- Ong, Y.K., 2018a. Opening Address by Mr. Ong Ye Kung, Minister for Education, at the Schools Work Plan Seminar. Ministry of Education, Singapore. www.moe.gov.sg/news/speeches/opening-address-by-mr-ong-ye-kung-minister-for-education-at-the-schools-work-plan-seminar.
- Ong, Y.K., 2018b. The next steps to learning for life. *Straits Times*. www.straitstimes.com/singapore/education/the-next-steps-to-learning-for-life.
- Office of Teacher Education (OTE), National Institute of Education, NTU, Singapore (NIE), (n.d.). Teaching Scholars Program, OTE, <http://tsp.nie.edu.sg>.
- Portilla, X., Ballard, P., Adler, N., Boyce, W., Obradović, J., 2014. An integrative view of school functioning: transactions between self-regulation, school engagement, and teacher–child relationship quality. *Child Dev.* 85 (5), 1915–1931. www.jstor.org/stable/24033029.
- Reeve, J., Ryan, R.M., Deci, E.L., 2018. Sociocultural influences on student motivation as viewed through the lens of self-determination theory. In: Liem, A.D., McInerney, D.M. (Eds.), *Big Theories Revisited*, vol. 2. Information Age Publishing, Charlotte, NC, pp. 15–40.

- Reimers, F.M., 2018. A study in how teachers learn to educate whole students and how schools build the capacity to support them. In: Reimers, F.M., Chung, C.K. (Eds.), *Preparing Teachers to Educate Whole Students: An International Comparative Study*. Harvard Education Press, Cambridge, MA, pp. 1–32.
- Ryder, G., 2018. As Tech Disrupts Our Jobs, It's Not Too Late to Turn Pain Into Gain. World Economic Forum. www.weforum.org/agenda/2018/10/tech-disrupt-jobs-fourth-industrial-revolution-ilo.
- Schwab, K., 2016. *The Fourth Industrial Revolution*. World Economic Forum, Cologny/Geneva, Switzerland.
- Shanghai Municipal Education Commission, 2014. *Shanghai shi Shenhua Gaodeng xuexiao kaoshi zhaosheng zonghe Gaige Shishi Fangan* [Implementation Plan to Deepen the Comprehensive Reform for College Entrance Exam in Shanghai City]. gaokao.eol.cn/zui_xin_dong_tai_2939/20140919/t20140919_1177783.shtml.
- Tan, C., Ng, C., 2018. Assessment reform in Shanghai: issues and challenges. *Int. J. Educ. Reform* 27 (3), 291–309. <https://doi.org/10.1177/105678791802700304>.
- Tan, J., 2019. Equity and meritocracy in Singapore. In: Teng, S.S., Manzoni, M., Poon, K.K. (Eds.), *Equity in Excellence: Experiences of East Asian High-Performing Education Systems*. Springer, Singapore, pp. 111–125.
- Teataja, R., 2013. Basic Schools and Upper Secondary Schools Act. Riigi Teataja. www.riigiteataja.ee/en/eli/ee/530102013042/consolide.
- The Straits Times, 2016. The Education System Over the Years. www.straitstimes.com/singapore/education/the-education-system-over-the-years.
- Tomlinson, C.A., 2005. Grading and differentiation: paradox or good practice? *Theory Into Pract.* 44 (3), 262–269.
- Wang, L.Y., 2019. Leveling up academically low-performing students in student-centric education in Singapore: global trend, local policies, and future directions. *Educ. Rev.* <https://doi.org/10.1080/00131911.2019.1642304>.
- Williams, H., 2019. What's the same and what's different: the use of the Mathematics Mastery program in a special school for pupils with moderate learning difficulties. *Proc. Br. Soc. Res. Learn. Math.* 39 (2), 1–6.

Reinventing teacher agency: shifting the narrative of teacher education

Poonam Batra, India Lead, TESF Network Plus; Formerly with Central Institute of Education, University of Delhi, Delhi, India

© 2023 Elsevier Ltd. All rights reserved.

Economic, social and political context of education reforms in India	722
Revisiting teacher agency	723
Critical spaces to achieve teacher agency	724
A teacher's personal narrative	724
Disrupting normative discourse and challenging socialization	726
Developing capacity to imagine and find voice	726
Experiencing the language of possibility and social justice	726
Conclusion	727
References	727

Education reforms instituted to meet neo-liberal agendas were adopted by several countries across the globe via processes of internationalization and policy borrowing. While policy texts in India of the mid-1980s (GoI, 1985a,b) underscored the agency of the teacher and the child; the sub-text of the discourse alluded toward reforms based on major shifts in economic policy. Initiated in the mid-1990s reforms were explicitly designed to meet the demands of shifting the aims of education from nation building and emancipation to global economic competitiveness and human resource development. A convergence between the neoliberal focus on skills and employability, and a behavioristic outcome-based model of education reshaped classroom practice. Cultures of performativity replaced the idea of teacher agency with prescriptive curricula and uncompromizing regimes of testing.

A decade later, the National Curriculum Framework (NCF) (NCERT, 2005) for school education in India, had the potential to reposition teachers as agents of educational change and social transformation (Batra, 2006). Even so, policy measures favored testing regimes over teacher support interventions, widening the gap between teachers and school curriculum. Despite a progressive school curricular discourse, precious little changed in the social ethos of preparing teachers. The view of teachers as “state implementing agency”—a role continuity since colonial India—continues to dominate policy thinking; and classroom practice is shaped by teachers whose work and worth is judged within a neo-liberal frame (Batra, 2014), that is now converging with a discourse around religio-cultural nationalism.

Shifts in curriculum policy in several countries, argue Priestley et al. (2015, p. 1) “heralds a [re]turn to the centrality of the teacher in school-based curriculum development.” In India the emphasis on teachers in addressing issues of quality can be seen in the rhetoric of redesigned schemes of universalizing elementary education (the *Samagra Shiksha Scheme* was initiated by the Government of India (2018) to improve quality of school education with an explicit emphasis on teachers, bringing under one umbrella, three central government schemes: SSA, RMSA and CSSTE). Scholars however, argue that there is very little latitude for teachers to exercise agency in real classrooms (Batra, 2005; Long et al., 2017). As the primary vehicle to implement politically mandated school-based reforms, teachers are often made partners in taking forward policy agendas and measures that undermine their individual and collective agency. The turn of policy discourse toward the centrality of teachers needs to be seen in this light and examined critically. This is especially so as several non-state and global actors step in to tell teachers what to do in classrooms. A recent study examining the impact of the pandemic and resultant school closure highlights how teachers are being left out of core decisions and processes of educational practice, undermining their potential role in ensuring quality education (Batra et al., 2021).

There is every reason to be skeptical about the renewed emphasis on teachers in India and elsewhere. The current policy focus on teachers is not very different from existing global discourses on teachers that have acknowledged that “teachers matter” (OECD, 2005) and the “centrality of effective teachers” (Barber and Mourshed, 2007). Global teacher governance frames, argues Robertson (2012) shift “sovereignty and authority away from the nation and the teacher, to the global and global actors” (p. 23).

Given this, the important questions are: whether the emphasis on teachers in the policy space can indeed be used to reinvent the idea and practice of teacher agency; and whether agentic teachers can disrupt status-quo and enhance their professional role and self-governing space?

In an attempt to examine the potentiality of teacher agency in “reclaiming” and “restoring” the rightful space for teachers in education and society, the paper engages with: the impact of educational reforms in India; key theoretical propositions of teacher agency; and how learning experiences can be designed to cultivate teacher agency. Education reforms around the teacher in India - driven by forces of liberalization and the post-Jomtien (EFA movement launched in 1990) pressures of an international community—and their impact are discussed first. Subsequently, key theoretical propositions of teacher agency are revisited to ascertain their feasibility in the contemporary Indian context. In the third section the paper presents the proposition and an argument that learning experiences can be designed to cultivate in teachers the potential to act individually and collectively. The argument developed is substantiated and authenticated by the processes identified in a teacher's personal narrative, a summary adaptation of which is presented as well.

Economic, social and political context of education reforms in India

A slew of reforms instituted by the Government of India (GoI) in the mid-1990s with World Bank support (District Primary Education Program) continued under the *Sarva Shiksha Abhiyan* (SSA) launched in 2003. With focus on textbook development and pedagogic renewal, nation-wide reforms led to isolated attempts at transforming the capacity and role of teachers. SSA enhanced school access to meet the goal of universalizing elementary education. Quality of education however, remained a concern. Several survey reports (ASER and NAS) indicated that precious little was achieved in terms of learning outcomes and in reducing gender and social inequities. This attracted a range of private players to occupy the educational space left vacuous by an unimaginative state school system.

Donor-driven reforms began with a stated focus on the child as active learner and the teacher as the object of reform. The key premise on which the process of pedagogical renewal was designed included addressing “student deficit” via remediation and compensatory education. Intervention focused on designing activity-based textbooks and training teachers to translate imported and abstract ideas of “child-centered education” in state schools. In the absence of adequate theoretical engagement due to impoverished teacher discourses, local interpretations of concepts related to “child-centered education” diluted most attempts to establish any serious discourse on praxis among teachers. It was no surprise that these new ideas could not take root in a school system that was highly bureaucratized with an emphasis on “delivery” in which the teacher was seen as mere implementer. Stagnant learning outcomes for over a decade could to some extent be attributed to misplaced ideas of educational reforms that in effect undermined teacher agency and failed to do justice to the agency of the child in the everyday classroom, in spite of mouthing child and teacher-centric platitudes.

In the absence of any noticeable change in classrooms and learning levels, several studies commissioned by the World Bank and other donor agencies built a national narrative around teacher absenteeism and teacher incompetence (Batra, 2017a). The need to examine and address the professional needs of teachers and their qualification was discounted. A simultaneous set of researches led to the overwhelming view that school teachers were highly politicized and large proportion (over 90%) of the education budget spent on teacher salaries brought very little returns (Kingdon and Muzammil, 2009). These studies became the basis of the policy trajectory during the third decade of liberalization in India. Compelling economic arguments resulted in crystallising the idea of “para teachers” (a generic term applied to school teachers appointed on contract basis, under varying service conditions) who were hired, sometimes at less than one-third the cost of a regular teacher’s salary. A spate of researches supported the argument that lowering teacher costs is necessary to achieve quality education for all (Muralidharan and Sundararaman, 2013; Atherton and Kingdon, 2010). Several reform measures emerged from the widely held view that elementary school teachers require minimal investment in terms of professional development and support.

This led to emasculating the school teacher; who was increasingly talked about as *the* object of reform rather than a partner in enhancing the quality of school education. Reform measures kept teachers away from critical decisions and processes of curriculum design, pedagogy and learner assessment. Viewed as “lacking” in the progressive discourse of “child-centered” education; teachers were made mere receptacles of imposed ideas that related neither to their pre-service education, nor to the prevailing realities of state classrooms. Divested of intellectual agency, school teachers felt increasingly alienated from their own work. As agents of the state they had little choice but to become complicit in bureaucracy-driven reforms. Public opinion about teachers reached an all-time low fomented by an anti-teacher discourse and state assault on the educational space and work of teachers.

While economic liberalization since the early 1990s pushed for opening up the educational space to the market; the curricular policy discourse via the NCF, 2005 brought renewed focus on the role of the state in maintaining pluralism and diversity of knowledge; seeking participation of teachers in designing textbooks and creating democratic spaces of learning. Though the NCF focused on altering the perspective and content of school curriculum, its expectations of the teacher’s role continued to be largely prescriptive and often contradictory. The NCF advocated constructivist pedagogic approaches, positioning the child at the center of curricular change. However, it undermined the criticality of the teacher in translating its vision in real classrooms. The principles that guide the perspective of the NCF—such as connecting knowledge to life outside school, enriching curriculum beyond the textbook and nurturing identities within the democratic polity of the country—could not be realized in the absence of proactive engagement of the school teacher (Batra, 2005, 2016). Empirical research bears testimony to this argument (Dalal, 2015).

In not addressing the teacher as central to learning, the NCF could do little to counter the anti-teacher discourse or to resist inflexible measures of accountability that defined the worth of teachers’ work in isolation of systemic context. It was no surprise that the Right to Education Act (2009) maintained a studied silence on who should be a teacher; but included the use of punitive measures against “teachers” who fail to “perform” (GoI, 2009, p. 8, 46).

Inadequate focus on teachers in instituting curricular reform; and the increasing wedge between the curricular discourse and neo-liberal policy measures of teacher accountability and large-scale testing took away the legitimacy of policy engagement with classroom processes. Deep penetration of the market in the education sector coincided with the entry of the corporates to alter the very purposes and processes of education. Elementary education began to be seen as the site for skill development and vocationalization to serve the needs of industry, now legitimized through the National Education Policy (GoI, 2020).

The most significant fall out was that “altered purposes” of education no longer required a teacher who could bring “professional judgment” to the task of teaching. Teachers were being “told” what to do and how. Edu-preneurs exploited the opportunity to market blue-prints of lesson plans—for profit, which schools implement with little regard to the contexts of learners, teachers or the school (Kamat et al., 2016).

Mirroring the colonial frame, schools and teachers as state agents appear to work together to perpetuate ideologies that preclude schools from becoming sites of contestation, debate and meaningful learning. With the gradual withdrawal of the state and the expansion of local and corporate schools, a new nexus, a new narrative and educational practice emerged. This is predicated on the lack of teacher agency and their complicity in maintaining status-quo. Relying on “normative assumptions” about learners in a diverse society, teachers tend to undermine the intellectual capacities of students from disadvantaged communities; exclude their narratives from spaces of learning (Avalos, 1986; Batra, 2017a) and rarely create learning experiences that foster critical consciousness about social inequities and systems of oppression as argued by several critical theorists (Freire, 1970; Giroux, 1986; McLaren, 2015).

Accounts of teachers’ work in select state schools reveal how the “diminished professional role of teachers has undermined their agency and ability to practice critical pedagogy for educating the subaltern classes” (Majumdar and Mooij, 2012, p. 33). Given the understanding that teachers’ voices are often silenced within the hegemonic discourse of reforms that undermine their agency, the question—whether teachers can indeed disrupt status-quo; reclaim and restore their place to empower students—becomes central to any advocacy for teacher agency. This question also becomes important in the context of the virtual capture of the teacher preparation space by private interests. In this light, how do we view the space, dominated by the “private”, as a space of debate and contestation required for developing critical thinking teachers? (Batra, 2022).

Revisiting teacher agency

Whether cultivating agency can enhance teachers’ professional and self-governing space becomes important in the light of the observation that structure-agency debates remain unresolved; though attempts have been made to synthesize this duality. Sayed and Novelli (2016, p. 18) argue that social systems are a “result of interaction between individuals (agency), who are aware of the ‘rules’ (structure) and are also capable of bringing about structural change by influencing the ‘rules’ that govern social action”. Given that the agency of teachers is related to their positions in society and in the education system, it needs to be examined in the current context of altered “structures” and “cultures” within which teachers work and which generate and shape educational practice.

At this point it would be useful to engage with some of the theoretical ideas around the construct of agency. The notion of “agency” as associated with “free” will and the capitalist notion of “enterprise”, central to the development discourse was rejected by social feminists to argue that agency is related to power and ideology (Wilson, 2008). In Butler’s (1990) view, agency is productive power that involves resistance to authority. Examining the discourse on racism, Baez (2000, p. 329) argues that the agency-structure dichotomy obscures how racism is produced and resisted at local sites of power, such as classrooms. Educators she argues, must engage in resistance-practices to challenge the “normalizing effects of institutional practices and relations” in (re)producing racism. Resistance therefore, needs to be seen in relation to agency of teachers working in an emancipatory framework to bring about change (Sannino, 2010). Barker et al. (2001, p. 46) view agency as “the socially constituted capacity to act”. Scholars have also distinguished between agency as discursively and socially constructed and agency as developing individual subjectivities and identities that are negotiated and constituted within the social (Eteläpelto et al., 2013).

Notwithstanding the structure-agency debates, more recent scholarship has addressed the fundamental question of: why agency? And whether teacher agency can even be achieved and how?

Within a more pragmatic framework, the idea of teacher agency can be linked to discursive shifts in the educational discourse as a consequence of neo-liberal reforms (Biesta, 2015). One such shift, as indicated earlier, is a focus on learning outcomes and evidence-based education. This, argues Biesta, undermines the subjective judgment (of the teacher); and assumes that teachers can and should be told what to do in classrooms. Teacher professional (subjective) judgment he asserts, is essential to the educational process as it is intimately tied to the question of purpose in education. It is therefore crucial that teachers be given the space and opportunity to exercise their professional judgment. Agency, in Biesta’s (2015) view therefore, is a key dimension of professional judgment; and professional judgment he argues, is crucial to achieve two critical functions of teacher education, namely, “socialization” and “subjectification”.

In the above argument, agency can be seen to be rooted in the philosophy of action developed by Dewey and Mead (Biesta, 2005, 2006); establishing the view that the purposes and processes of education are intertwined. Vygotsky’s cultural-historical framework and his emphasis on cultural tools also offer the scope to understand “how human agency plays into how we learn and act upon the world” (Edwards, 2017, p. 269). Agency in this frame is seen as an ecological phenomenon; and not “a capacity residing in individuals” but “the way actors critically shape their responses to problematic situations” (Priestley et al., 2013, p. 9). Agency emerges through the interplay of personal capacities and resources, enabling and constraining contexts, by means of which individuals act. Teachers are key to this process as they reflect and often act counter to societal constraints—sometimes enabled and often constrained by their social and material environments—each time bringing subjective judgment to the task of decision-making.

In theorizing agency as ecological phenomenon Priestley et al. (2015) assert the possibility of achieving “individual agency”. Individual teacher agency is reflected in processes of education comprising of interpretation, meaning-making and pedagogic communication. However, because teachers often work in isolation and in environments of competition rather than collaboration (Long and Lampen, 2014; Batra, 2017b) there are serious constraints in developing the collective agency of teachers. Also, cultivating individual agency does not resolve the structure-agency tangle. For teachers to feel empowered requires not only individual

agency to take professional decisions inside the classroom, but also political agency that comes from an understanding of the many facets that impact education, and a vision for a transformed society.

The concept of teacher agency would therefore entail a gamut of dimensions that equip teachers to act collectively and politically, in order to resist and impact policy and the structures within which they work. The journey of developing teachers may therefore need to include processes that augment “epistemic agency” and reflexivity; equip teachers with professional capacities to take a view; and the disposition and mettle to act collectively and politically.

Critical spaces to achieve teacher agency

In keeping with Biesta’s (2015) theorization, enabling discursive resources for teachers to develop agency through initial teacher education and ongoing professional development is critical and in the realm of possibility. Well-crafted, long duration initial teacher education programmes can enable “self-constitution” as a potent force via opportunities to interrogate teachers’ existing roles, identity and positions in society. Long et al. (2017, p. 17) observe that elements of the 4-year integrated Bachelor of Elementary Education (BEEd) program as envisaged and enacted (Batra, 2009), “where attention is given to the self, to conceptions of knowledge, and to socio-political structures, provide this pre-service teacher education focus”. The BEEd is acknowledged as an exemplar program by the National Council for Teacher Education (NCTE, 2009, p. 49). This initial teacher education offers the critical space where agency can be cultivated via emancipatory experiences that develop epistemic identities, reflexive dispositions, and critical professional repertoires. Examining teacher agency in the implementation of the new Scottish curriculum, Priestley et al. (2013) found evidence of agency within the collaborative and horizontal structures rather than hierarchical ones.

Specific principles and domains of knowledge that govern the structure of initial teacher education that aims to cultivate agency can be broadly categorized as follows: interdisciplinary and transdisciplinary enquiry; dialogical interplay between theory and practice; deconstruction and reconstruction of school knowledge; engagement with constructs of human relations, the self and the individual-social interface; and hands-on experience with creativity, professional capacities and pedagogic communication. Each of these provide wide-ranging entry points to engage with fundamental questions about the nature of knowledge and knowledge claims; nature of learner and learning; notions of selfhood, subjectivity and renegotiating identities - thus reconstructing the discourse and practice of education via critically designed components and pedagogic communication during the program. These enable the two critical functions of teacher education—“socialization” and “subjectification”—that Biesta (2015) proposes are necessary for developing teacher professional judgment.

The ecology of teacher agency hence comprises of a complex process of interaction of a range of dimensions—individual, social, cultural, material and structural. To enable agency via an “embedded process of social engagement” includes the unsettling of the normative discourse; developing the capacity to imagine, assert voice; and engaging with the language of possibility and social justice. Each of these discursive experiences are made possible through a combination of curricular components of the 4-year BEEd program, drawn from the conceptual base of a range of social science disciplines and the pedagogic experiences provided.

The next section of the paper presents a narrative that reveals how personal and professional aspects of a life are intertwined and how these find expression and become the medium of deep engagement during the journey of a developing teacher—leading her to be the person she is. This narrative is one of the several narratives—written by select graduates of the BEEd program—as part of a research study (undertaken by the author during the period 2009–2014), exploring connections between teacher education and social change. The original piece, written in 2009 by a graduate of the batch of 2002 is much longer. An edited and adapted version is presented here to illustrate the impact of the 4-year program. The original form of the narrative, as a coherent text of reflection and reasoning, has been maintained without subjecting it to reductionist analysis. In this form, the powerful impact of the program in providing opportunities and spaces is discernible. In the last section, some of the key processes, winnowed from the narrative, illustrate how these might have contributed in cultivating teacher agency.

A teacher’s personal narrative

I was brought up to become a ‘civilized, cultured girl and a good daughter’; and had internalized the roles ‘expected of a daughter, wife and mother’. During the four years of the programme, notions of the ‘good child’ started breaking, especially as I came face-to-face with myself through a series of self-development workshops and theoretical engagement with identity, caste, gender and social exclusion. I began to understand how my socialisation - at home and outside - and how our positions in society influenced the decisions we took in everyday life.

I started dialoguing with myself. My deep self that constantly sought recognition from others, gradually turned inwards. I felt compelled to question my feelings, limitations and potentials, even my existence. Discovering myself, I began to respect myself; and most importantly, to accept myself for the way I am. As I discovered my strengths, I burst into all sorts of creative expression - writing poetry, acting and the arts. These opportunities lent me a ‘voice’ - to assert and to fight for my rights as a woman and a Dalit.

(The term Dalit was used as a census classification of “Depressed Classes” prior to 1935; after which the term Scheduled Caste was used. The term is rooted in the historical and political struggle of the “Depressed Classes”, a term used by Dr. B.R. Ambedkar (1891–1956), himself a Dalit).

I chose to work on 'women and teaching', as part of my fourth year research project. Reading and discussing through the course on 'gender and schooling' led me to think about my assumptions and beliefs about gender. I began to understand subjective realities of women's oppression from the perspective of my own position in a patriarchal society. I ceased to view myself as a 'victim' of social prejudice.

The above is but an imperceptible peep into how I was to transform as a person. I could think critically but was confused. The world seemed diseased and wretched. I felt as if someone had stolen my tinted glasses that safeguarded/disallowed me from 'seeing' the world as it existed. Small injustices would trigger a chain of reactions, sometimes of awful proportion. I was seen as quarrelsome, a misfit in my family and society. Many felt I had lost my 'feminine charm' - of being compliant, forbearing, submissive and timid. I was often told how unfit I was for marriage. To my disbelief, my parents worried for my sister, as if I would sway her too. Friends outside college told me that 'textbook knowledge' is not about real life.

Today, I am far more resilient. Conflict coming out of constant questioning opened new ways of understanding the world and myself. I learnt that what we consider 'natural', 'inevitable', 'good' or even 'true', may not be so. Conflict forced me to transform myself - intellectually and emotionally. Even as I write, I continue to fight many battles in the everyday - as a woman and Dalit.

For years, I had nurtured the guilt of being a Dalit. I was afraid to reveal my caste identity. The thought of being labelled as belonging to a 'reserved' category' haunted me. I had seen many SC (scheduled caste) and ST (scheduled tribes) school children face blatant exclusion. In self-denial, I started criticizing the reservation policy. I thought people who availed reservation were considered incompetent; and took a vow never to avail benefits of reservation. Soon I felt alienated from my own community.
(As per the Indian Constitution, Scheduled Castes and Scheduled Tribes fall under the "reserved" category for the purpose of granting specific benefits including education and employment.)

Sociological concepts helped me understand my socialization and why the educational discourse is dominated by upper caste, urban and male thinking. In my enthusiasm to explore my milieu and the larger world, I dabbled with any activity that came my way. Learning took a different meaning for me. I celebrated this epistemic shift in all spheres of my life; it was exciting.

Engaging with 'curriculum studies' revealed how the dominant retain social control through education. I understood how teachers are influenced by political, economic and social forces that shape them to become mere carriers of textbook information. Analysing school textbooks revealed how ideology is used as a weapon to dominate.

In our batch students came from different cultures, taste, even dressing style. None of us were keen to become teachers. Differences were pronounced and sharp. Humiliating, criticising each other had become an everyday norm. By the end of the fourth year something changed - we bonded with love, care and concern for each other. The change started happening in the second year of the programme during self-development workshops. Many of my classmates opened up. Those who belittled others with such ease broke down in front of the whole group, reflecting on their childhood experiences. Workshops helped us develop insight into the prejudices we hold about people. Opportunities to reflect and question ourselves helped develop empathy and trust. I could now carry the inner strength gained, with me everywhere.

I came to terms with the fear of failure and isolation. I regained confidence and developed the courage to take risk. Theatre helped to imagine the possible and develop the conviction to alter reality. What I gained in the four years has been emancipatory and enlightening. I have gained the tools to think and explain; the language of critique and possibility; the capacity to imagine and innovate; and most significantly the courage to challenge hegemonies and speak against injustice; the conviction to fight for my rights and the rights of others.

The narrative reveals how opportunities for reflection creates dialog with the inner self—a core engagement of the 4-year program—which when interfaced with theory and the field, appears to have created a "unique chemistry of experiences and processes of engaging with social reality, questioning the self and social milieu, reflecting on one's position in society and developing a sense of agency" (Batra, 2015, p. 169). The narrative uncovers how voice and agency developed through an "embedded process of social engagement". Learning experiences led to a series of processes that disrupted the normative discourse and challenged socialization; developed the capacity to imagine and find voice; and made accessible the language and praxis of possibility and social justice. Each of these processes, discussed in detail below, are authenticated by the narrative, demonstrating how pedagogic experiences and communication can be designed to develop epistemic agency and a sense of social justice that inspires individual and collective action.

Disrupting normative discourse and challenging socialization

Normative discourse is intimately tied to subjective experiences of caste, gender, race, class, religion and community. These are reflected in the socio-cultural preconceptions that play out in everyday life in and outside school. Making a case for perceiving the school as counter-socialization, Kumar (2004, p. 87) argues that “we need to perceive the school in conflict with the community’s code of socialization”. To develop agency, student-teachers are given opportunities and well-crafted experiences to examine prevalent normative discourses using an inter-disciplinary lens; and how these determine social relationships and life trajectories. This facilitates engagement with questions of diversity - how diversities come into being and how stereotypical and discriminatory attitudes that buttress inequality are constructed. Inter-disciplinary engagement offers a framework to deconstruct and reconstruct school knowledge; challenge and verify knowledge claims. Examining the lived experiences of diverse communities enables contesting the perceived neutrality of knowledge, merit and objectivity. Contesting school curriculum, assumptions and cultural beliefs about knowledge and learning offer the opportunity to unsettle habitudes. Interrogating constructs embedded in society engenders new perspectives, enabling teachers to cognize the classroom as an imaginable site of confronting “power that dominates” and “educing power to transform reality”.

The narrative reveals the student-teacher’s struggle as she deconstructs perceptions about her “self” and understands her identity as intimately tied to issues of diversity and social justice. The careful integration of learning experiences creates the potential to challenge the “cultural self”, nudging it to question and aspire to understand the self and the other. The focus on and engagement with subjectivity creates opportunities to reconfigure personal identities. As student-teachers become aware of the social injustices around them, they learn to question and resist the projected image of education as apolitical. Components of the program bring teachers closer to their aspirations and identities, allowing them to grow as communal beings with shared purposes. Diverse pedagogic experiences provide them with conceptual tools that are emancipatory, allowing them to break the shackles of socialization and wrestle their way through structures that encumber agency. Agency is realized through the “socially constituted self”, wherein socialization is “creatively reconfigured in relation to hopes, fears and desires for the future”, substantiating Emirbayer and Mische’s (1998, p. 971) assertion that agency reflects the intent to transform the present.

Developing capacity to imagine and find voice

Theater in the BEEd provides the platform to imagine and create “possibilities” where none exist. It provides a context as well as a medium of learning, enabling epistemic shifts in the notion of pedagogy; and the “democratization of classroom space” (Shivapriya, 2004). In addressing issues deeply personal in nature, theater offers non-threatening spaces to engage with the theory-experience disconnect around gender, caste, identity, exclusion, injustice and the self. Probing the theory-experience dialectic forms the basis on which student-teachers learn to speak the language of possibility.

The de-contextualized psychological frame that forms the dominant sub-culture of teacher preparation is substituted by a deliberative discourse that breaks the false neutrality and “apolitical posture” within which teachers are prepared (Batra, 2014). Engagement with issues and concerns of contemporary society lead to major epistemological shifts in imagining the nation, where constructs of local knowledges, active citizenship, diversity and inclusion, attempt to redefine curriculum and establish teaching as socio-political practice. This prompts the developing teacher to view herself as central to a transformative agenda and not a mere implementer of a given curriculum.

The narrative highlights how teachers can be educated to engage with issues of inequity by being exposed to egalitarian methods of teaching and learning. Finding and asserting voice in the personal space of family and community is perhaps the most significant achievement for a young person preparing to become an educator. Learning to negotiate the familial space of authority and tradition creates a new sense of person-hood and the confidence to walk uncharted paths with courage and conviction. Drawing on Vygotsky’s cultural-historical perspective, it can be argued how student-teachers “push themselves to make sense and to connect their sense-making with the public meanings valued in teaching” (Edwards, 2017, p. 275).

Experiencing the language of possibility and social justice

The narrative demonstrates how traversing through her life journey created the possibility of interfacing the social with the individual. Engaging with the complexities of social reality develops capacities and the need to introspect, reason and call into question hierarchies and inequalities by voicing concern; expressing dissent and reaching out to make a difference. With the breakdown of power equations between the teacher and the taught, prospective teachers experience democratic spaces of learning. Developing notions of pedagogy that are non-hierarchical, fosters the idea that knowledge is not transmitted but actively constructed within the shared context of learning. Epistemic inquiry of social realities prompts critical understanding of concepts that have been taken for granted and cast in limited frames. For instance, questioning the popular notion of poverty as a mere economic indicator, prompts student-teachers to question related assumptions of poverty as an “individual condition” and a “state of mind” that is sought to explain children’s low performance in school. Foregrounding the importance of teachers’ conceptual understanding in

supporting agency, Greene and Yu (cited in MacLellan, 2017, p. 251) argue that taking the responsibility for “what we know and do not know”, “marks our epistemic agency”, “because knowledge arises from the choices we make individually to engage or not engage”.

Another graduate shares how initial teacher education “helped me develop perspectives with respect to various issues, whether they were related to children, society, or education ... teaching methodologies change a lot with one’s perspective with respect to knowledge and learning. As a professional teacher, I am able to appreciate children’s role in learning Various courses made me conscious of the fact that what is taught and how it is taught are deeply political questions. Much refinement came from the practicum, projects, field visits and the internship program” (excerpt from the personal narrative written by a BElEd graduate of 2004).

Examining the individual within the socio-political facilitates envisioning the role of the collective in creating change. Agency comes into being through the “socially constituted self”, emerging out of a focus on personhood and subjectivity that help reconfigure personal identities. In the words of the narrator, whose personal narrative is presented in this paper, “the malaise, conflict was an essential stage of understanding the world before we could undergo crucial shifts in perspective.” Conflict provided the impulse to struggle epistemically and emotionally. Positioning herself as a “catalyst of change”, the narrator substantiates Emibayer and Mische’s (1998) argument that agency is connected to the intended vision of a future different from the present. As Biesta (2015, p. 989) argues, the process of “continual imaginative reconstruction of the future involves drawing upon past experiences in order to clarify motives, goals, and intentions, to locate possible future constraints and to identify morally and practically appropriate courses of action.” Addressing teachers’ aspirations along with their identities as epistemic and social beings foregrounds the need, as argued by Belo (cited in Biesta, 2015), to critically examine their own assumptions, values, and beliefs.

Conclusion

This paper has argued that the operational linkage between a socially inclusive and humane education pivots around the teacher and teacher collectives. Several trans-national examples bear testimony to the socialist feminists’ assertion that agency is located at the collective as well as the individual level (Wilson, 2008). Analysis of educational reforms in Porto Alegre in Brazil highlights how democratic educational policies and practices have been sustained with the active participation of reflexive teachers (Gandin, 2011). Examples of teacher collectives offering resistance as in the Chicago teachers’ strike (Lipman, 2013) demonstrate the possibility and potential of collective teacher agency. Long et al. argue how teachers’ unions in South Africa successfully resisted administering the ANAs (Annual National Assessments) in 2015, because they “strongly believe that the ANA in its current form, is not in the best interest of our learners (and) the provision of quality education” (2017, p. 8).

Developing the agency of teachers requires deep engagement with structural reforms, and the disruption of teaching as “self-normalizing practice”. While the former prepares teachers for collective political action, the latter helps teachers recognize their own complicity and positioning in the conditions they contest. The personal narrative unveils that “the problem of teacher agency considers how power, identity, subjectivity and freedom intersect and inform each other” (Pignatelli, 1993, p. 412). Engaging teachers with complex dynamics between the self, the other, episteme and power is at the heart of the proposed shift in the narrative of educating teachers.

References

- Atherton, P., Kingdon, G., 2010. The Relative Effectiveness and Costs of Contract and Regular Teachers in India. Institute of Education, University of London, pp. 256–276.
- Avalos, B., 1986. Teaching Children of the Poor: An Ethnographic Study in Latin America. International Development Research Centre, Ottawa.
- Baez, B., 2000. Agency, structure, and power: an inquiry into racism and resistance for education. *Stud. Philos. Educ.* 19 (4), 329–348.
- Barber, M., Mourshed, M., 2007. How the World’s Best-Performing Schools Come Out on Top. McKinsey & Company, London.
- Barker, C., Galasinski, D., Galasinski, D.D., 2001. Cultural Studies and Discourse Analysis: A Dialogue on Language and Identity. Sage.
- Batra, P., 2005. Voice and agency of teachers: missing link in national curriculum framework 2005. *Econ. Polit. Wkly.* 4347–4356.
- Batra, P., 2006. Building on the National Curriculum Framework to enable the agency of teachers. *Contemp. Educ. Dialogue* 4 (1), 88–118.
- Batra, P., 2009. Teacher empowerment: the education entitlement–social transformation traverse. *Contemp. Educ. Dialogue* 6 (2), 121–156.
- Batra, P., 2014. Problematising teacher education practice in India: developing a research agenda. *Educ. Change* 18 (Suppl. 1), S5–S18.
- Batra, P., 2015. Curriculum in India. In: *Curriculum Studies in India*. Palgrave Macmillan, New York, pp. 35–63.
- Batra, P., 2016. Positioning teachers in the emerging education landscape of contemporary India. In: *India Infrastructure Report 2012*. Routledge India, pp. 257–269.
- Batra, P., 2017a. Quality of education and the poor: constraints on learning. In: *A Companion to Research in Teacher Education*. Springer, Singapore, pp. 417–433.
- Batra, P., 2017b. Reclaiming the space for teachers to address the UEE teaching-learning quality deficit. In: Govinda, R. (Ed.), *India Education Report: Progress of Basic Education*. Oxford University Press, New Delhi, pp. 130–150.
- Batra, P., 2022. Teacher education in India: Virtual capture of the “public” in Ian Menter (Ed). *The Palgrave Handbook of Teacher Education Research*. Palgrave Macmillan, London.
- Batra, P., Bazaz, A., Shanmugam, A., Ranjit, N., Kaur, H., Das, R., 2021. Impact of the COVID-19 Pandemic on Education, Livelihoods and Health. Bengaluru: TESF India, IHS. <https://doi.org/10.24943/ICPELH08>.
- Biesta, G., 2005. George Herbert Mead and the theory of schooling. In: Troehler, D. (Ed.), *Pragmatism and Education*. Sense Publishers, Rotterdam, pp. 117–132.
- Biesta, G., 2006. “Of All Affairs, Communication is the Most Wonderful.” *Education as Communicative Praxis*. John Dewey and Our Educational Prospect. A Critical Engagement With Dewey’s Democracy and Education, pp. 23–37.
- Biesta, G., 2015. What is education for? On good education, teacher judgement, and educational professionalism. *Eur. J. Educ.* 50 (1), 75–87.
- Butler, J., 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, London.
- Dalal, J., 2015. The indelible class identity: ethnographic examination of a school. *Econ. Polit. Wkly.* 36–39.

- Edwards, A., 2017. The Dialectic of Person and Practice: How Cultural-Historical Accounts of Agency Can Inform Teacher Education. *The Sage Handbook of Research on Teacher Education*, pp. 269–285.
- Emirbayer, M., Mische, A., 1998. What is agency? *Am. J. Sociol.* 103 (4), 962–1023.
- Eteläpelto, A., Vähäsantanen, K., Hökkä, P., Paloniemi, S., 2013. What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York.
- Gandin, L.A., 2011. Porto Alegre as a counter-hegemonic global city: building globalization from below in governance and education. *Discourse* 32 (2), 235–252.
- Giroux, H.A., 1986. Critical theory and the politics of culture and voice: rethinking the discourse of educational research. *J. Thought* 84–105.
- Gol, 2020. National Education Policy. Department of School Education and Literacy, MHRD, New Delhi. Accessed on 17th August, 2020. https://www.mhrd.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf.
- Government of India (Gol), 1985a. Challenge of Education: A Policy Perspective. MHRD, Government of India, New Delhi. Available at: <http://14.139.60.153/handle/123456789/1592?mode=full>. (Accessed 24 March 2022).
- Government of India (Gol), 1985b. *The Teacher and Society: Chattopadhyaya Committee Report (1983–85)*. MHRD, Government of India, New Delhi.
- Government of India (Gol), 2009. The Right of Children to Free and Compulsory Education Act. Gazette of India, Government of India, New Delhi.
- Government of India (Gol), 2018. Samagra Shiksha Abhiyan: An Integrated Scheme for School Education. Ministry of Human Resource Development, Government of India, New Delhi.
- Kamat, S., Spreen, C.A., Jonnalagadda, I., 2016. Profiting From the Poor: The Emergence of Multinational Edu-businesses in Hyderabad, India. A Report for Education International.
- Kingdon, G., Muzammil, M., 2009. A political economy of education in India: the case of Uttar Pradesh. *Oxf. Dev. Stud.* 37 (2), 123–144.
- Kumar, K., 2004. *What is Worth Teaching?* Orient Blackswan.
- Lipman, P., 2013. The rebirth of the Chicago Teachers Union and possibilities for a counter-hegemonic education movement. *Mon. Rev.* 65 (2), 1.
- Long, C., Lampen, E., 2014. Teacher agency and professional practice: developing and nurturing creativity in education. In: *Proceedings 20th Annual National Congress of the Association of Mathematics Education of South Africa*, vol. 1, pp. 275–284.
- Long, C., Graven, M., Sayed, Y., Lampen, E., 2017. Enabling and constraining conditions of professional teacher agency: the South African context. *Contemp. Educ. Dialogue* 14 (1), 5–21.
- MacLellan, E., 2017. Shaping Agency Through Theorizing and Practicing Teaching in Teacher Education. *The SAGE Handbook of Research on Teacher Education*, pp. 253–269.
- Majumdar, M., Mooij, J., 2012. *Education and Inequality in India: A Classroom View*. Routledge.
- McLaren, P., 2015. *Life in Schools: An Introduction to Critical Pedagogy in the Foundations of Education*. Routledge.
- Muralidharan, K., Sundararaman, V., 2013. Contract Teachers: Experimental Evidence From India (No. W19440). National Bureau of Economic Research.
- National Council for Educational Research and Training (NCERT), 2005. National Curriculum Framework, 2005. National Council for Educational Research and Training, New Delhi.
- National Council for Teacher Education (NCTE), 2009. National Curriculum Framework for Teacher Education: Towards Preparing Professional and Humane Teacher (NCFTE). National Council for Teacher Education (NCTE), New Delhi. Available at: https://ncte.gov.in/website/PDF/NCFTE_2009.pdf. (Accessed 24 March 2022).
- OECD, 2005. Teachers Matter: Attracting, Developing and Retaining Effective Teachers. Organisation for Economic Co-operation and Development, Paris. Available at: <https://www.oecd.org/education/school/34990905.pdf>. (Accessed 24 March 2022).
- Pignatelli, F., 1993. What can I do? Foucault on freedom and the question of teacher agency. *Educ. Theor.* 43 (4), 411–432.
- Priestley, M., Biesta, G., Robinson, S., 2013. Teachers as agents of change: teacher agency and emerging models of curriculum. In: *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*, pp. 187–206.
- Priestley, M., Priestley, M.R., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Publishing.
- Robertson, S.L., 2012. Placing teachers in global governance agendas. *Comp. Educ. Rev.* 56 (4), 584–607.
- Sannino, A., 2010. Teachers' talk of experiencing: conflict, resistance and agency. *Teach. Teach. Educ.* 26 (4), 838–844.
- Sayed, Y., Novelli, M., 2016. The Role of Teachers in Peacebuilding and Social Cohesion. The Research Consortium on Education and Peacebuilding. Universities of Amsterdam, Sussex and Ulster, pp. 1–100.
- Shivapriya, N., 2004. Drama in Teacher Education: An Evolving Perspective. Unpublished MEd Dissertation. Department of Education, University of Delhi, Delhi.
- Wilson, K., 2008. Reclaiming “agency”, reasserting resistance. *IDS Bull.* 39 (6), 83–91.

Mentoring “in the time of cholera”: perspectives, promises and challenges in the preparation of mentors for diversity in the new digital teacher learning space

Lily Orland-Barak, Faculty of Education, University of Haifa, Haifa, Israel

© 2023 Elsevier Ltd. All rights reserved.

Introduction and focus	729
Perspectives: from face to face to remote mentoring	730
The study of face-to-face mentoring	730
Mentoring in the new digital teacher learning space	730
Remote mentoring in diverse contexts: lessons from two local case studies	732
Mentoring student teachers in “the time of cholera” in public arab schools: impeding and facilitating socio-cultural, economic and organizational conditions of the context	732
Research-policy connections “in the time of cholera”: spheres of knowledge for mentoring in the digital teaching space	734
Challenges: conceptual, curricular and practical directions to mentoring in the digital teacher learning space	735
Conceptual directions	735
Curricular directions: integrating mentoring approaches in remote mediation	735
Promises: new channels of communication between mentoring research, policy and practice	736
Coda: back to the future	737
References	738

Human beings are not born once and for all on the day their mothers gave birth to them, but life obliges them over and over again to give birth to themselves.

Gabriel Garcia Marquez, “Love in the Time of Cholera” (2007).

Introduction and focus

The COVID-19 pandemic has forced educational leaders, researchers and policy makers worldwide to “give [new] birth to themselves” in the move to remote teaching and learning. The transition has been sharp, unexpected and created uncertainty and pressure (Reich et al., 2020), as well as new opportunities for educational leaders and systems to redefine themselves in this contested reality. Taken to the field of mentoring in teacher education (the focus of this article) we are witnessing a concerted effort to reinvent and redefine mentors’ roles and practices as agents of change and mediators of teacher learning in this new medium (Mutton, 2020; Floden et al., 2020). Metaphorically, it has led to the pressing need to “prepare mentors in the time of Cholera”, as the title of this article suggests, inspired by Gabriel Garcia Marquez’ best seller “Love in the Time of Cholera” (Garcia Márquez, 2007).

“Wisdom comes to us when it can no longer do any good”, Garcia Marquez writes. Indeed, his words resonate with a panelist’s comment participating in a recent virtual conference around the challenges of policy, research and practice ahead of us in teacher education: “Corona has motivated us to think outside the box in terms of how we can re-invent ourselves to access each other’s worlds in new ways. We have to become relevant to the on-going, day-to-day changes in the education arena, which, of course, directly influence the education of teachers’ [personal communication]. Addressing this call, this article discusses the perspectives, challenges and promises of mentoring in the new digital teacher learning space. Specifically, it focuses on recent research agendas and policy trends aimed at redefining and “reinventing” the preparation of mentors in diverse socio-cultural educational contexts across the globe. I then zoom in to the case of mentoring in Israeli teacher education. The socio-cultural and political complexity of such a context can serve as a rich illustrative case of how new conceptualization and operationalization of mentoring is dialectically constituted at the intersection between local settings of centralized politically-driven educational policy and the diverse and contested socio-cultural/ethnic contexts of teaching and teacher education within which such policy operates (Biesta et al., 2020; Nir, 2020). Within this background, I present two recent studies on mentoring in teacher education launched in Israel in response to the COVID-19 pandemic. These studies exemplify the challenges and concerted efforts to create new connections between research, policy and practice in order to inform on the preparation of mentors and teacher educators for supporting teacher learning in highly diverse and disputed socio-cultural educational contexts in the new digital/hybrid space.

The first part discusses global perspectives based on recently published research and conceptualization around the question: What characterizes the passage from synchronous to asynchronous/hybrid mentoring in terms of mentors’ roles, tasks, practices and interactions? The second part deepens into two recently launched local studies in the context of pre-service and in-service mentoring in Israel, with a focus on diverse teaching populations. This is followed by a discussion of the challenges that emerge from recent study initiatives for the conceptualization and preparation of mentors for diversity. Finally, I discuss the promises and potential opportunities afforded by an era of almost full digitalization in teacher education and mentoring.

Perspectives: from face to face to remote mentoring

Mentoring in teacher education, whether face-to-face, remote or hybrid, is essentially about the mediation of teacher professional learning. Resorting to their knowledge and expertise as experienced teachers, mentors constantly engage in mediating between persons and content in value-laden contexts of practice (Feiman-Nemser, 2001a,b; Orland-Barak, 2010; Clarke and Mena, 2020). Prior to pondering into studies that focus on the new teacher learning space of remote mentoring, let us consolidate what we know about face-to-face mentoring and the mediation of teacher learning.

The study of face-to-face mentoring

The study of face-to-face mentoring can be consolidated through several comprehensive reviews conducted on global perspectives on mentoring research over twenty years [1991–2014], (Wang and Odell, 2002; Orland-Barak, 2014, 2016). The reviews surface major shifts from technical, strategic, monological approaches to mentoring (predominant in the early 80s) to dialogical, socio-cultural, discursive views of mediation in mentoring and mentored learning (from the early 90s onwards) (Harrison et al., 2006; Roehrig et al., 2008; Richter et al., 2013; Orland-Barak, 2014; Orland-Barak and Wang, 2020). The latter view, still prevalent at the sprouts of the 21st century, emphasizes aspects of mediation related to mentors’ reasoning, competencies and skills (Feiman-Nemser, 2001a,b; Achinstein and Athanases, 2006) and/or contextual aspects of mentoring relationships related to cultural, ideological and identity aspects of mediation interactions (Zeichner, 2012, 2020). To this end, mentoring mediation is conceptualized as entrenched in multicultural realities, at the intersection between mentors and mentees’ values, belief systems and behaviors throughout the pre-service, induction, and in-service teacher life-long learning continuum.

The major conceptual shifts in the study of mentoring can be represented as moving from a focus on mentoring as transmitting knowledge to mentoring as mediating knowledge (Orland-Barak and Maskit, 2011), as illustrated in Fig. 1 below:

Recent studies on mentoring also stress the importance of integrating modeling and the acquisition of core practices with contextual aspects of mentoring (Forzani, 2014; Grossman, 2018; Kavanagh et al., 2020; Kavanagh and Danielson, 2020). This growing and often debatable trend (Philip et al., 2019), claims for the need to promote systematic reasoning through strategic mediation, geared at the acquisition of patterns and repertoires of practice that can be appropriated to the particularities of the context. Taken together, then, the mediation of teacher professional learning through mentoring is broadly conceptualized nowadays as a practice characterized by interconnectedness between cognitive, behavioral and socio-cultural factors and between forms of reasoning, prior knowledge and engagement in experiences, and patterns of best practice in activity (Hobson et al., 2009; Mena et al., 2017; Villegas and Lucas, 2002; Fenwick and Edwards, 2010). Such conceptualization extends Homer’s Mentor matriarchal and patriarchal functions of caring, support and guidance to include dynamic, collaborative and interactional mediating roles that develop within multifaceted networks of interactions in activity and in a variety of contexts (Edwards, 2017) as illustrated in Fig. 2 below:

In their article “Five waves of conceptualizing knowledge and learning for our future in a networked society” (2019), Hoadley and Kali contend that if we consider how technology has changed the relationship between individuals, learning, and society, one finds that technologies are creating the space for a fifth wave of technology mediated learning. Can we also think of mentoring, as the mediation of professional learning, as being redefined in this new wave? If so, how? To address this question, let us first consider recently published research on mentoring in the context of the new digital teacher learning space, which has developed worldwide during COVID-19.

Mentoring in the new digital teacher learning space

In his editorial for the Special Issue entitled Teacher education and COVID-19: Responses and opportunities for new pedagogical initiatives (Mutton, 2020), Trevor Mutton stresses “the fascinating way in which so many teacher educators internationally have seen the current crisis as an opportunity to consider new ways of working, to adopt innovative approaches to pedagogy and to re-conceptualize the nature of their teacher education programs” (p. 439). Indeed, the issue opens a pioneer window into a wealth of questions, challenges, pressing issues and innovative pedagogies that teacher educators are grappling with around the world. In

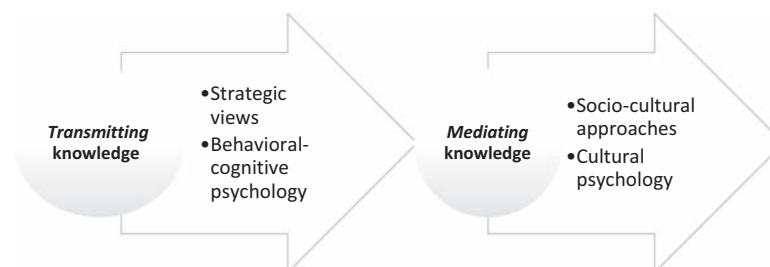


Fig. 1 Conceptual shifts in the study of mentoring.

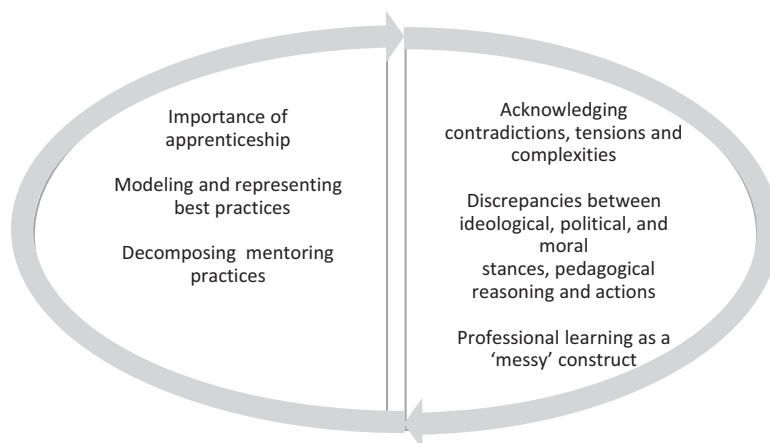


Fig. 2 Cognitive, behavioral and socio-cultural dimensions of mentoring roles.

the context of mentoring for teacher preparation, the various studies in the special issue as well as in other important recent publications, consider issues related to mentors' new roles and skills (Carrillo and Flores, 2020; Sepulveda-Escobar and Morrison, 2020; Hadar et al., 2020a,b), pedagogies and methodologies of mediation that are technologically enhanced (Philipsen et al., 2019; Locke, 2020; Mutton, 2020; Loose and Ryan, 2020; Hadar et al., 2020a); issues of equity and values that emerge in the digital learning space (Darling-Hammond and Hylar, 2020) and forms of adaptive expertise to the new teaching-learning environment, often positioning expert mentors as novices (Thompson et al., 2020; Donitsa-Schmidt and Ramot, 2020). These publications, indeed, take up Darling-Hammond et al.'s (2020) call to update and strengthen teaching and school leadership standards that reflect new emergent knowledge about social, emotional, and cognitive development; culturally responsive pedagogies; and trauma-informed practices. For example, Scull et al.'s study in Australia (Scull et al., 2020) suggests supporting teachers in the move to on-line teaching and learning through innovative pedagogies and indicators to develop and evaluate newly identified dimensions of access, participation and engagement. Robinson and Rusznyak's study (Robinson and Rusznyak's, 2020) in the South African context, consider other aspects of learning such as situational learning, relational learning and pedagogic reasoning that need to be redefined when mediating learning from and in practice in the new digital school-based experiences. Thompson et al. (2020), focus on how the abrupt move to remote teaching has disrupted traditional novice-expert mentoring relationships positioning mentors (as I have previously argued "sometimes as novices and sometimes as experts" in their reasoning and actions (Orland-Barak and Yinon, 2005). They describe situations in which the novice was in a position where s/he was more knowledgeable than the mentor, mostly as related to technical aspects of on-line teaching such as setting up zoom meetings, collecting school-work in virtual classrooms, or creating pre-recorded content for asynchronous instruction. As they contend, these new novice-expert positionings during mentoring interactions challenge traditional apprenticeship models whereby the mentor teacher functions as an expert and model to the novice teachers. I would further argue that it invites a fresh "restart" for redefining traditional apprenticeship top-down mentoring roles and mentored learning models, which are still strongly dominant in numerous teacher education programs. The need for such rethinking is strongly voiced in Leacock and Warrican's paper (2020), in the context of teacher education in the West Indies. They urge moving away from what they refer to as traditional colonial-based perspectives on teaching and teacher learning, in order to meet the new constructivist and digitally enhanced character of teaching and teacher support.

The above studies leave us with challenging questions for future thinking and research. For example, Barnes and Sax (2020) raise questions of whether traditional school-based assessment by mentors are appropriate for teacher certification given the new conditions, as well as questions of technological affordances such as online teaching simulation for assessing teacher competences. Thompson et al. (2020) wonder whether the uncertainty and vastly changed work context for which a high percentage of mentors are unprepared will require programs to screen mentors for a new set of skills and to revisit effective mentoring needs in a COVID-19 teaching context. Hadar et al.'s study on student teachers' social-emotional competencies during the COVID-19 crisis (2020), raises important questions such as how to enhance teachers' social-emotional competencies, how to redesign a teacher education curriculum that would balance between subject matter and pedagogical content knowledge and knowledge about coping resources, in order to strengthen students' emotional agency. In this vein, they suggest a possible additional direction in professional development programs, which focuses on new roles for mentors as agents of change through the dissemination of mindfulness practices to develop teachers' social-emotional competencies. Elaborating on new roles and pedagogies for mentors, Assunção Flores and Gago's study in the context of Portuguese teacher education (2020), highlights the need to prepare school mentors to mediate student teacher learning in areas that they identify as problematic in the digital teaching-learning space. Some of these are connecting between theory and practice, justifying pedagogical decisions, carrying out the observation-planning-evaluation cycle, conducting formative assessment and managing new conditions for learning to teach as related to time, space, forms of interaction with pupils and mentoring processes. In a parallel literature review co-authored by one of the same two authors, Carrillo and Assunção

Flores (2020) extend these aspects to difficulties related with poor online teaching infrastructure; inexperience in the use of digital instructional formats (quoting Huber and Helm, 2020) of both teachers and mentors (quoting Judd et al., 2020); limited information and resources, and the complex environment at home (quoting Zhang et al., 2020). These emergent problems, they contend, touch upon issues of agency, responsibility, flexibility, choice and student presence. They also foreground worrying issues of equity and inclusion, reinforced in Huber and Helm’s comparative study of German, Swiss and Austrian schooling evaluation, assessment and accountability in times of crises COVID-19 (Huber and Helm, 2020). One of their conclusions highlights the urgency of attending to student/teacher populations coming from socially weak families or educationally deprived families, those with low self-regulating abilities, and those who face difficulties in planning their day. School communities as networked systems, they contend, should collaborate to compensate for the socioeconomic disadvantages of students’ families, particularly in disadvantaged areas. They call on fostering coherent approaches with collegial support and networks with other partners within education and in the local and regional contexts. In the US context, the call to prepare teachers to work with families and communities in teacher education has been strongly and critically voiced by Ken Zeichner (2019, 2020), Cochran-Smith and Villegas (2016), and many others for the past three decades. Within these calls, Zeichner’s recent article “Preparing teachers as Democratic Professionals (2020) goes further to caution us about the dangers of trying to “save” students from their communities (Zeichner, 2019; Dzur, 2008), rather than preparing them for working with and for their students’ families and communities (Haddix, 2015; Murrell, 2001, quoted in Zeichner, 2020).

At the background of these concerns and challenges, Hill’s study (Hill et al., 2020) from British Columbia proposes interpreting these as opportunities for re-envisioning the preparation and support of teachers in managing equity and social justice, wellbeing and environmental protection in the new digital teacher learning environment. Likewise, Donitsa-Schmidt and Ramot’s study in the context of teacher education in Israel during the COVID-19 pandemic, stresses the “bright side” of the need to develop emergency remote teacher education platforms. Their study identified the emergence of numerous bottom-up professional learning communities that were initiated by the teachers themselves in collaboration with teacher colleges. Their findings highlight enhanced collegiality, solidarity and collaboration, which was built around trust and respect among the community members. This led to the development of safe and supportive environments for deepening understandings and trying new practices. Thus, they contend, paradoxically, the period of social distancing and self-quarantine led to increased peer collaboration, shared learning, mutual aid and growing partnership between academic institutions and training schools (p. 586–595). Similarly, Sepulveda and Escobars’ study on on-line teaching placement during the COVID-19 in the Chilean context suggests that, despite the impasses identified mainly in terms of lack of communication with supportive agents and mentors, which ultimately affected the quality of feedback student teachers received, participants, mentioned positive aspects of the experience and work with mentors and teacher educators. Specifically, they agreed that the experience enhanced their teacher preparation in terms of learning to adapt to unexpected situations, to learn how to work with different technological platforms and to design strategies for reaching their students in the virtual mode. Also taking a promising stance, Philipsen et al. (2019) combine goals, processes, strategies, context and outcomes to propose an integrative framework for teacher professional development for on line and blended learning through regular and just-in-time support and feedback.

The above said, Shelley’s aphorism “the rich get richer and the poor get poorer” (Shelley, 1845) resonates in my ears, especially when thinking about socio-cultural contexts where on-line affordances are scarce and often almost non-existent, as evidenced in some of the local contexts that are described in the next section.

Remote mentoring in diverse contexts: lessons from two local case studies

The following two local cases illustrate recent research initiatives in Israel to collect comprehensive evidence of how mentoring for teacher learning operates in the new learning digital space both in pre-service and in-service education. Case One is situated in the contested socio-cultural and economic context of pre-service teaching practice in Arab schools. Case Two demonstrates concerted efforts to create new collaborations between Ministry policy and local teacher education research to provide evidence-based foundations for developing remote mentoring and professional development in diverse school teaching settings.

Mentoring student teachers in “the time of cholera” in public arab schools: impeding and facilitating socio-cultural, economic and organizational conditions of the context

There is a growing acknowledgment that mentoring in a context of cultural diversity should address the voices and cultural practices of the various national and minority groups that are part of a society (Mullen, 2020; Mayer et al., 2017; Orland-Barak, 2010). Despite this, many teacher education and mentoring models worldwide still forward a rather uniform approach to learning to teach, taking little account of the diverse cultural, social and national affiliations and traditions that student teachers bring to teacher education from their home background (Cochran-Smith and Villegas, 2016). Israel is no different (Eilam, 2003). Despite the diverse social, religious and political tapestry of its society (Agbaria, 2016), Israeli teacher education institutions follow, for the most, a centralized uniform curricular policy. The new COVID-19 reality of social distancing has stressed the problem of uniformity even more acutely, given the unique features of the context that characterize mentoring and mentored learning in the Arab society, especially in Arab public schools in Israel (where student teachers do their practice teaching assisted by mentors). The Arab minority in Israel (including Muslims, Christians and Druze) form around 21.0% of the total population in Israel and around 55% live in a lower socioeconomic status than that of the Jewish population (Central Bureau of Statistics Israel, 2020). The Arab and Jewish

education systems in Israel are separate, and for the most part the two societies maintain a different way of life and culture (Arar, 2012). Despite positive developments that the Arab education system in Israel has undergone in recent years, it still faces challenges and barriers (Arar and Abu-Asbah, 2013; Arar and Massry-Herzllah, 2016). Many schools are characterized by low achievement levels, high dropout rates, inadequate budgets and resources, and low teaching quality with traditional norms and teaching methods in schools (Arar and Haj-Yehia, 2016). In addition, these schools are less technologically oriented, have limited access to technology and lack experience in virtual teaching skills (Arab Education Monitoring Committee, 2020). These constraints create unique challenges for pedagogical mentors working with these student teachers, and especially now, during the COVID-19 pandemic. At the background of this context, we launched a comprehensive two-phase research study (Daniel-Saad and Orland-Barak, *in progress*) aimed at exploring the pedagogical and personal experiences of around 200 pedagogical mentors from the Arab society in Israel working with student teachers in school practice teaching settings where in-person mentoring has changed almost fully into remote mentoring through a-synchronic (Moodle/mail and WhatsApp application) and synchronic mentoring (Zoom platform/Unicko). We combine various theories and models: Hofstede’s cultural dimensions theory (Hofstede, 1991, 2001) with Schwartz’ theory of basic values (2009) as complementary frameworks to examine how culture and values shape patterns of perception and processes of knowledge building as well as particular attributions to events and evaluations of behavior. We also examine these patterns in relation to different approaches to mentoring as conceptualized in Orland-Barak and Wang’s integrative mentoring model (2020), and to mentors’ adaptation to the new virtual, social and supportive organizational environments in terms of skills and dispositions through Moho’s model of human occupation (Kielhofner, 2009). The research questions revolved around mentors’ significant hardships and crisis events, successes and coping strategies, to gain access into their professional experiences in the passage from face to face to remote mentoring. We asked: What characterizes mentors’ transition experience in the Arab society to remote mentoring during the Corona crisis? How do pedagogical mentors perceive the crisis in their work as related to their beliefs, and approaches to mentoring? How do cultural characteristics shape mentors’ coping strategies in the virtual teacher learning space? How does the new virtual mentoring environment influence mentors’ functioning? How does the organizational environment operate as an enhancing/hindering condition for sustaining the passage? The research interview guide included questions about the following topics:

- Mentoring models and professional experiences: personal principles and values that helped manage emergent challenges.
- Mentoring roles and the practicum in an online environment: perceptions of role; ethical issues raised regarding remote mentoring (privacy, information security, developers’ copyrights); resources that mentors find as being most helpful in managing this period.
- Mentoring roles in the Arab Society in an online environment: environmental conditions that helped to adapt to the new emergency mentoring reality; inhibiting conditions; school possibilities for online work.
- Outcomes in the application of remote mentoring experiences: difficulties and challenges through accounts of particular events; strategies used to deal with the situation.
- Implications as to what is expected to happen.

The first phase of the study which was qualitative and has now been completed, gathered thick data from on-line semi-structured in-depth interviews with twenty selected Arab mentors, representing eleven major teacher education colleges and universities throughout the country (Tel Aviv, Haifa, Central, North Juda and Samaria and Jerusalem district) in different types of schools (regular and special education). Emergent findings from the data support recent studies described in earlier sections in regard to mentors being compelled to reinvent their roles, skills and practices “in action”, given the fact that they had no previous guiding models or knowledge to lean on. We also identified the blurring of formal traditional boundaries between mentors with their student teachers (where interactions used to take place in the stable school environment and had now moved to students’ homes, family cars or other areas outside the school). For example, mentors were bound to observe their student teachers teach in their homes (both for student teachers and pupils) and with their family members present. The blurring of boundaries was also felt in the move to an almost full time week work (although mentors’ work contracts were not extended) in which “matriarchal” mentor roles (which were less stressed before) were intensified: sustaining continued mentoring relationships, providing on-going emotional support, and assisting students to manage personal, familiar conflicts that emerged. To this end, mentors mentioned having to deal with challenges related to the particular collectivist and patriarchal features of the culture. For example, since mentoring meetings were conducted via zoom in the student teachers’ homes, there was a tendency to adopt norms of communication and interactions that were accepted by their family members. This often led to a more reserved and inhibiting discourse, avoiding confrontation or exposure of sensitive aspects that both student teachers and/or mentors might have wanted to raise. Some students coming from orthodox families also resisted the use of cameras, or appearing in video, which they reported as limiting the sharing of ideas and exchange of knowledge. Technologically, there were also reported challenges faced as related to lack of resources, and inadequate Internet connectivity. The socio-economic factor has also played a negative role in facilitating the online learning experience. One example is a case reported when a parent called to complain about the use of the camera during the lesson, since their house consisted of two rooms only, thus obliging the student’s mother to have her head cover on during the whole lesson since wherever she moved, she could be seen on the camera. Patriarchal orientations to education were also described. One example is the case in which the fiancé decided for the student teacher whether she would participate in a peer learning session or not, or cases of parents interfering with the college emergency instructions at times of lockdown. The dual role of the female student teacher, wife and mother obliged her to maneuver between the contradicting requirements of the family roles and those of the academic institution. This had consequences for the quality of their participation in mentoring sessions. All of these factors

constituted an additional challenging layer that hindered the passage to remote mentoring. Based on the findings from this phase, the third phase (2021) will entail the design of a questionnaire that will be distributed to 400 Arab mentors across the country. By the end of the project (December 2022) we hope to obtain a comprehensive evidence-based picture of what mentoring “looks like” and “feels like” for the Arab population of mentors working in Arab State schools.

Research-policy connections “in the time of cholera”: spheres of knowledge for mentoring in the digital teaching space

The demanding challenges of moving to online teacher learning prompted the Office of the Chief Scientist at the Ministry of Education in Israel to promote an initiative that brought together educational researchers and academics from universities and teacher colleges in order to provide support to the educational system in dealing with pressing areas related to remote learning and teaching. One of these focal areas was mentoring and teacher professional development for remote teaching in diverse contexts of practice. The committee for this area formulated five theoretical/evidence-based principles for teacher learning in the new digital environment: the need to cultivate technological skills in particular disciplinary contents (Mishra and Koehler, 2006); the need to cultivate technological and technological pedagogical teaching approaches with a focus on how these can address differences between student populations and contexts, while striving for equal opportunities for learning and assessment (Avime et al., 2020); the importance of attending to prospective and experienced teachers’ ingrained beliefs regarding the integration of technology in teaching and how these mis/align with principles of online learning (Blonder et al., 2013); the importance of including awareness of and tools for dealing with social, emotional and identity aspects of teaching and learning online (Garrison et al., 2001) and the need to create an integrative hybrid mentoring model that combines different approaches to teacher mentoring in terms of orientations to knowledge construction and development, content, context, experience and change processes, and that is sensitive to teachers’ developmental stages (Orland-Barak and Wang, 2020). At the background of this document, which was published and disseminated to the entire population of teacher educators, mentors and stakeholders involved in teacher education in the country, we launched a pilot study, which was funded by the Ministry of Education to investigate how these principles play out in the context of online practice teaching (Blonder and Orland-Barak, 2020). We assumed that the digital learning space has different characteristics from the physical space of the school (for example, distributed knowledge and cognition, activities that combine online interactions in a hybrid way with classroom teaching methods (hybrid learning) and involvement in new knowledge-building interactions that might raise issues of equity (Hoadley and Kali, 2019). We contended that these knowledge fields may challenge and expand traditional models of teacher knowledge defined generically (Shulman, 1986; Gess-Newsome, 2015) and disciplinarily (e.g. in EFL-Pang, 2016 and science teaching, Driets-Esser et al., 2017). We were curious to understand what happens in the encounter between theoretical knowledge about digital learning and the experiential knowledge that develops within the online practice school teaching setting. We aimed at gaining an initial glimpse at how mentor teachers in diverse settings deal with unplanned situations that arise in the online space, to encourage social interactions, to conduct discussions during a lesson, to conduct laboratory experiments in the online space, to identify evidence for learning in the online space, to develop online assessment, etc. A follow-up, large scale mixed-method study (in progress) is intended to gather additional data from multiple sources, toward the development of a theoretical and operative model of practice teaching in the digital environment, based on the bodies of experiential knowledge identified.

The initial study identified spheres of expert teacher knowledge (some of which reinforce recent studies described in previous section) that operate in the digital teaching practice space and have direct implications for mentoring roles and forms of assistance to address these spheres in diverse contexts (Singer, Doryoseph, Blonder, Orland-Barak, in progress).

- Reinventing teacher identity: This sphere of knowledge includes how a teacher identifies herself, her philosophy of teaching, her style and the many factors, which influence this perception of self, which may be transferred, transformed and/or disrupted in the digital space. This sphere connotes with previously described studies about how the shift to teaching in a digital space positioned experts as novices.
- Tracking Pupil Engagement: The teachers we interviewed each described their sense of tracking engagement in the face-to-face classroom and lacking it in the digital space, which ultimately created a reality of isolation for the students (Hargreaves and Fullan, 2020). Our expert teachers gave examples of seeking and finding ways to “feel” the class. The knowledge to be fostered in this area includes pedagogical awareness of the need to track student engagement, understanding of what constitutes evidence of engagement, and technological tools to enact activities, which provide evidence of engaging pupils.
- Assessing pupil knowledge: The sphere of knowledge for assessment includes identifying and tracking the development of pupil knowledge (Inbar-Lourie and Levi, 2020). In addition, it includes measures of the effectiveness of teaching methods through digital formative and summative assessment. It seems that this sphere of knowledge has room for new methods, suitable tools of testing and a new approach to testing.
- Identifying activity effectiveness: This category focuses on measuring the effectiveness of assessment methods in the digital space. The opportunities for outside help, even as simple as spellcheck, open the need for alternative methods of assessment. These new activities must be measured for effectiveness.
- Personalizing Communication: The direct face-to-face informal communication at school, in the hallway, before or after a lesson, when teachers watch the breaks, and during lessons, is absent in the digital space. Participants reported that during digital lessons they communicate in small windows of sixteen students at a time. In fact, when the students do not have a camera or have not turned it on, they only have a box with a name and maybe a voice, and there is a need to express personalized messages. So, this sphere of knowledge was found to be especially important. The knowledge needed in this sphere includes awareness of the importance of informal communication between teachers and students, understanding of the significance of

establishing personal connection between teachers and students in this new environment and designing activities which enable personalized communication through relevant technical tools for facilitating communication.

- Providing emotional support: Adapting known expressions of caring to the digital space was seen as very complex, and knowledge of how to provide emotional support was perceived as critical. The need for acquiring digital means of providing emotional support was strongly voiced in expressions such as “We can’t see them!”, conveying teachers’ frustration of the lack of resources for providing emotional support.
- Retooling content: This sphere of knowledge goes beyond knowing the subject matter being taught—it includes awareness of how the content can be adapted to fit in the digital space.
- Implementing technology-based pedagogy: The affordances of technology such as artificial intelligence, assistive technology, randomization, etc. were seen as a central aspects of new knowledge that needs to be developed. These also touch upon another complementary sphere identified, which was teaching students to become independent learners by maximizing the potential of the digital space.

How can the above initial identified spheres of expert teacher knowledge along with recent findings from local and global studies described in previous sections, inform on new directions for thinking about the practice of mentoring “in the time of cholera” and beyond? This is the focus of the following section.

Challenges: conceptual, curricular and practical directions to mentoring in the digital teacher learning space

Conceptual directions

In previous sections I posed the question of how/whether mentoring should be redefined in light of what [Hoadly and Kali \(2019\)](#) refer to as the fifth wave of technology mediated learning. The above synthesis of recently published and in progress research on mentoring, implies considering an additional layer of mediation to our understanding of mentored teacher learning. Conceptually, this new layer is grounded in the acknowledgment that the new digital teacher learning space grows out of socio-cultural views of teacher learning in the digital environment, including new norms and rituals of social interaction, of social and cognitive presence, new dimensions of equity, and of participation in the co-construction of knowledge in the digital space. Such a view invites ways of thinking about mentoring as operating in contexts of distributed cognition and knowledge in blended activity (i.e., across individuals, artifacts and tools which combine online educational materials and opportunities for interaction with traditional place-based classroom methods). The new layer of mediation necessitates preparing mentors to support teachers in using the features of the new digital context to redefine known pedagogical content and practical knowledge around aspects of organization and selection of content, classroom management (such as time, focus, voice, silence, participation, issues of equity, visibility, presence and language) while analyzing teaching situations and engaging in problem-solving mentoring interactions. It also invites a form of mentoring that combines distance learning conducted in a virtual learning environment with electronic study content designed for self-paced (asynchronous) or live web-conferencing (synchronous) online teaching, mentoring and tutoring. Furthermore, knowledge construction in community in hybrid learning spaces becomes, more than ever, central to teacher support and development. Finally (and most importantly), “context matters” ([Becher and Orland-Barak, 2018](#); [Van Den Berg, 2020](#)). This implies, as findings suggest, that mentoring in the new digital space should consider forms of mediation that are sensitive to the fact that experienced teachers are often positioned as novices (and vice-versa), and to the unique conditions of differing teaching contexts (i.e., socio-cultural, economic, political, physical, organizational) which can either promote or hinder the successful passage from face to face to remote teaching and mentoring.

Curricular directions: integrating mentoring approaches in remote mediation

The above conceptual directions push us to further refine recent conceptions of an integrated approach to mentoring ([Orland-Barak and Wang, 2020](#)), in order to attend to the unique features of the digital teacher learning space identified in recent research. The proposed refined model (illustrated in [Fig. 3](#) below) suggests connecting between the new identified areas of teacher learning and knowledge that need to be addressed in remote mentoring, and the kind of mentoring roles and approaches that can best facilitate the mediation of teacher learning in pre and in-service education in those areas: (1) Strategic and situated approaches to mentoring that stress mentor roles as models, as expert teachers and as co-thinkers can promote student teacher learning in the following areas of teacher learning identified in recent research: identifying activity effectiveness; personalizing communication; retooling content; evaluating participation; engagement and access; redefining the lesson cycle; tracking pupil engagement; assessing pupil knowledge; accessing technological affordances at the workplace; managing time and space; adapting to unexpected situations; working with different technological platforms; designing strategies for reaching students in the virtual mode. (2) Critical approaches whereby the mentor functions as critical pedagogue and agent of change to promote teacher awareness and agency can attend to the following areas of teacher learning (both novices and experienced) identified in recent research: Managing issues of ethics, equity and responsibility in the digital teaching space; attending to deprived socio-cultural-economic contexts with limited information, resources and access to technological platforms; questioning new practices in terms of the educational values promoted (3) Personal growth approaches in which the mentor assumes dialogical mentoring roles to promote teacher development (with a stress on in-service education) can facilitate in the following areas of growth identified in recent studies: being a novice again; reinventing teacher roles and identity; developing social-emotional competencies; developing adaptive expertise; promoting culturally-responsive teaching in the new digital

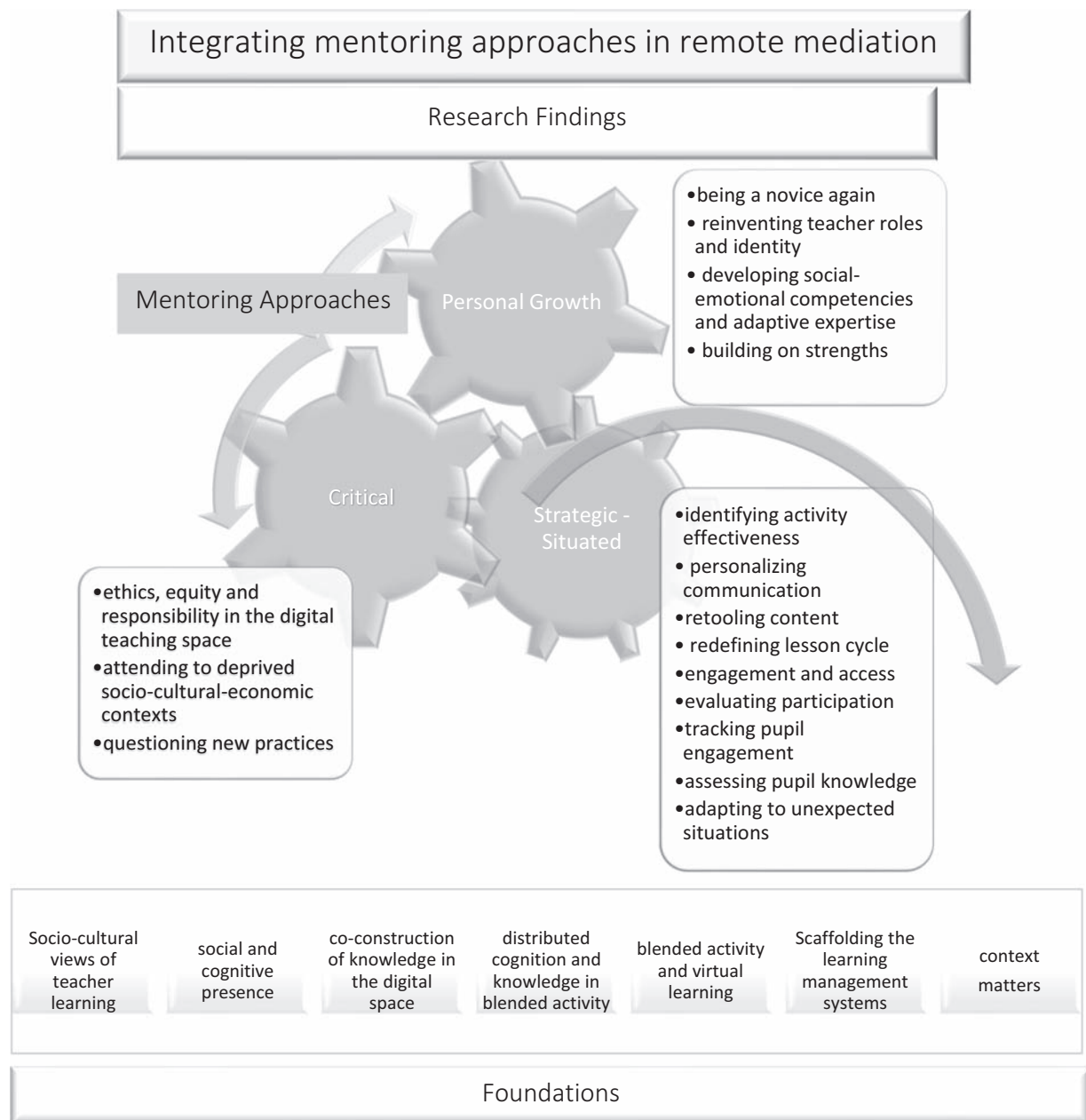


Fig. 3 Integrated mentoring approaches in remote mediation of teacher learning.

space; building on strengths to manage the new and unknown. Different mentoring roles, then, capitalize on the strengths of different approaches for developing particular areas of teacher learning in the digital space at different professional development stages. In any one mentoring interaction, these can operate in an integrative manner, combining diagnostic (i.e., mapping teachers' strengths, challenges and areas of learning to teach that need to be improved); deliberative (i.e., flexibly integrating relevant mentoring strategies in action) and inquiry (i.e., examining mentoring processes and outcomes) mentoring skills and competencies to support teacher learning (Orland-Barak and Wang, 2020). These integrative mentoring roles and practices require preparing mentors for thinking and performing adaptively, and for moving flexibly across different remote mentoring contexts. It also requires preparing them for managing tensions between teachers' needs, lacks and wants in the new digital teacher learning space, mentors' own tendencies to preserve familiar modes of practices, and professional demands on teacher learning.

Promises: new channels of communication between mentoring research, policy and practice

I conclude this article by discussing the promises of mentoring in the digital space, with a focus on new connections that have recently been established between the research-policy-practice discourse. It is widely recognized that the global and local discourses

of educational policy and research seldom engage in deep dialogue to improve the practice of teacher education (Biesta, 2007). Mentoring, as integral to teacher education processes, seems to suffer from a similar disconnection (Huber and Helm, 2020). True, the connection between research, policy and practice has always presented a challenge due to the different languages used, the different time frameworks in which they operate and the different principles that guide them. No doubt, each actor within this triangle is fully aware of the potential realization that these connections withhold for the field of education. However, the base for this relationship, which lies on a common ground of the perception of fidelity and truth creating trust, is not stable (Nelson et al., 2009). Research does not seem to be informing policy the way it should, and many would contend that policy changes usually occur detached from research agendas (Sleeter, 2014). It is also true that misalignments might be because research questions do not really attend to the interests and questions of potential users but rather solely to those of academic researchers (Levin, 2013).

The Israeli case is no different, especially given its political context, which is subject to constant instability and frequent changes of power. Within this context, we must mention the power of the two teacher organizations who are, in many cases, an obstacle to changes and improvements. Together, this environment has a drastic effect on the educational field as far as its priorities and ability to implement long-term changes are concerned. As a result, policy makers are forced into a tight schedule that does not enable them the leeway to perform research or even review up-to-date research to guide them in policy setting. Subsequently, like other aspects of educational policy, teacher professional development policy is subject to changes in the political sphere and to different emphases throughout the years. Turning to research, and although researchers aspire their work to have an impact on educational policy and practice, educational research in Israel values conceptualization and theoretical breakthrough, leaving much interpretation to overloaded and overwhelmed policy makers and practitioners. A well-known dispute between policy makers and researchers in conference panels in Israel is best illustrated in this excerpt (from one of these panels):

Policy maker: “I need answers to pressing problems of education “yesterday” but we only get them from you “late tomorrow” ... and then, when we get them, a couple of years later after the study is completed, we have already moved into something else”

Researcher: “And my biggest frustration is that you jump too fast from one reform to another and do not stop to build on gathered evidence in order to refine and improve initiatives that you started” [personal communication].

Recently, though, the above discourse seems to have undergone some welcome shifts. The new COVID-19 pandemic has created the need to open new channels of communication between the two discourses (Biesta et al., 2020) in order to successfully “survive” the challenges described in earlier sections. Thus, the common discourse of divide in Israel was compelled to move toward a more reconciling one, as policy makers and researchers joined forces in order to provide practical evidence-based solutions to the unprecedented and pressing reality of the COVID-19 uncertainty (see also Donitsa-Schmidt and Ramot’s study (2020) referenced in earlier section). This is evident in two major calls recently disseminated by the Ministry of Education (September and November 2020). Notice the opening paragraph of one of these calls (translated from Hebrew):

For many years we have witnessed the importance of establishing relevant connections between educational policy and research in order to improve educational practice. Despite this, we are constantly being engaged in dialogues that stress the disconnections between the three ... The Corona period has intensified the need to join forces and create better connections between the different participants in the educational arena. This has recently led to opening new fruitful communication channels between them. Following this emergent trend of collaboration, we invite you to submit research proposals that are geared to strengthening links between research, policy and practice in the following areas of education

Chief Scientist-Ministry of Education (November 2020).

In fact, one of the above local studies presented in this article is being financially supported by one of these calls. Connecting to Young et al.’s (2002) competing models of research and policy relations, COVID-19 has eventually prompted the shift in teacher education and mentoring research in Israel from a conception of educational research as being mostly guided by knowledge and enlightenment models of research (not necessarily committed to informing immediate policy concerns) toward more problem-solving and interactive models informed by policy issues, as well as by views of research and policy as being mutually influential. This has yielded recent agendas for both research and policy decision that are created within policy communities across a range of actors and within the entire spectrum of central policy, the laboratory and research teams.

Coda: back to the future

Amidst the above emergent shifts, it will be interesting to follow future research on mentoring to gain a deeper understanding of whether these winds of change are stable or fluctuating (especially post pandemics). Will research and policy go back to their divided roots or will they grow in new, entrenched directions? This invites following future research on mentoring with an eye on questions such as: What kind of questions and purposes guide future studies? What barriers and challenges of educational policy are taken into account? How are findings connected and/or suggested to be translated to practice? How do they connect to burning policy issues? These questions will help us understand future trends in mentoring in terms of the relations they exhibit between knowledge production and knowledge capacity (Levin, 2013) both in pre and in service education. For example, if a research question is prompted by a problem-solving model, how does it develop in the various stages of analysis, interpretation and representation of the focal study? How do the implications connect between findings and policy recommendations?

Back to the future, I return to Garcia Marquez’ novel ending:

He looked through the windows at the complete circle of the quadrant on the mariner’s compass, the clear horizon, the December sky, without a single cloud, the waters that could be navigated forever, and he said: “Let us keep going, going, going, back to La Dorada” ... “and how long do you think we can keep this goddamn coming and going?” ... “Forever, he said”.

(Garcia Marquez, 2007, p. 348).

References

- Achinstein, B., Athanases, S.Z., 2006. *Mentors in the Making: Developing New Leaders for New Teachers*. Teachers College Press, New York.
- Agbaria, A.K., 2016. Ethno-national politics of citizenship education in Israel and the counter-knowledge of Palestinian teachers. In: Banks, J.A., Suárez-Orozco, M., Ben-Peretz, M. (Eds.), *Global Migration, Diversity, and Civic Education: Improving Policy and Practice*. Columbia University Teachers College Press, New York, pp. 156–176.
- Arab Education Monitoring Committee, 2020. *Arab Education in the Shadow of the Corona Crisis* (Hebrew).
- Arar, K., Abu-Asbah, K., 2013. Not just location: attitudes and functioning of Arab local education administrators in Israel. *Int. J. Educ. Manag.* 27 (1), 54–73.
- Arar, K., Haj-Yehia, K., 2016. *Higher Education and the Palestinian Arab Minority in Israel*. Palgrave Macmillan, New York.
- Arar, K., Massry-Herzallah, A., 2016. Motivation to teach: the case of Arab teachers in Israel. *Educ. Stud.* 42 (1), 19–35.
- Arar, K., 2012. Israeli education policy since 1948 and the state of Arab education in Israel. *Ital. J. Sociol. Educ.* 4 (1), 113–145.
- Assunção Flores, M., Gago, M., 2020. Teacher education in times of COVID-19 pandemic in Portugal: national, institutional and pedagogical responses. *J. Educ. Teach.* 46 (4), 507–516.
- Aviran, E., Easa Livne, S., Blonder, R., 2020. Implementation of a personalized online learning system towards creating hybrid learning and teaching in chemistry classes. In: Glick, D., Cohen, A., Chang, C. (Eds.), *Early Warning Systems and Targeted Interventions for Student Success in Online Courses*. IGI Global, pp. 90–110. <https://doi.org/10.4018/978-1-7998-5074-8.ch005>.
- Barnes, M., Sax, P.E., 2020. Challenges of “return to work” in an ongoing pandemic. *N. Engl. J. Med.* 383 (8), 779–786.
- Becher, A., Orland-Barak, L., 2018. Context matters: contextual factors informing mentoring in art initial teacher education. *J. Teach. Educ.* 69 (5), 477–492.
- Biesta, G., Takayama, K., Kettle, M., Heimans, S., 2020. Teacher education between principle, politics and practice: a statement from the new editors of the Asia-Pacific journal of teacher education. *Asia Pac. J. Teach. Educ.* 48 (5), 455–459.
- Biesta, G.J.J., 2007. Bridging the gap between educational research and educational practice: the need for critical distance. *Educ. Res. Eval.* 13 (3), 295–301.
- Blonder, R., Jonatan, M., Bar-Dov, Z., Benny, N., Rap, S., Sakhnini, S., 2013. Can you tube it? Providing chemistry teachers with technological tools and enhancing their self-efficacy beliefs. *Chem. Educ. Res. Pract.* 14 (3), 269–285.
- Blonder and Orland-Barak, 2020.
- Carrillo, C., Flores, M.A., 2020. COVID-19 and teacher education: a literature review of online teaching and learning practices. *Eur. J. Teach. Educ.* 43 (4), 466–487.
- Central Bureau of Statistics Israel, 2020.
- Clarke, A., Mena, J., 2020. An international comparative study of practicum mentors: learning about ourselves by learning about others. *Teach. Teach. Educ.* 90, 103026.
- Cochran-Smith, M., Villegas, A.M., 2016. Preparing teachers for diversity and high-poverty schools: a research-based perspective. In: Lampert, J., Burnett, B. (Eds.), *Teacher Education for High Poverty Schools*. Springer International Publishing, Cham, pp. 9–31.
- Daniel-Saad and Orland-Barak (In Progress).
- Darling-Hammond, L., Hyler, M.E., 2020. Preparing educators for the time of COVID... and beyond. *Eur. J. Teach. Educ.* 43 (4), 457–465.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D., 2020. Implications for educational practice of the science of learning and development. *Appl. Dev. Sci.* 24 (2), 97–140.
- Donitsa-Schmidt, S., Ramot, R., 2020. Opportunities and challenges: teacher education in Israel in the Covid-19 pandemic. *J. Educ. Teach.* 46 (4), 586–595.
- Drits-Esser, D., Gess-Newsome, J., Stark, L.A., 2017. Examining the sustainability of teacher learning following a year-long science professional development program for inservice primary school teachers. *Prof. Dev. Educ.* 43 (3), 375–396.
- Dzur, A.W., 2008. *Democratic Professionalism: Citizen Participation and the Reconstruction of Professional Ethics, Identity, and Practice*. Penn State Press.
- Edwards, A., 2017. The dialectic of person and practice: how cultural-historical accounts of agency can inform teacher education. In: Clandinin, D.J., Husu, J. (Eds.), *The Sage Handbook of Research on Teacher Education*. Sage, Thousand Oaks, pp. 269–285.
- Eilam, B., 2003. Jewish and Arab teacher trainees’ orientations toward teaching-learning processes. *Teach. Educ.* 14 (2), 169–186.
- Feiman-Nemser, S., 2001a. From preparation to practice: designing a continuum to strengthen and sustain teaching. *Teach. Coll. Rec.* 103, 1013–1055.
- Feiman-Nemser, S., 2001b. Helping novices learn to teach: lessons from an exemplary support teacher. *J. Teach. Educ.* 52 (1), 17–30.
- Fenwick, T., Edwards, R., 2010. *Actor-Network Theory in Education*. Routledge, London, England.
- Floden, R., Stephens, A., Scherer, L. (Eds.), 2020. *Changing Expectations for the K-12 Teacher Workforce: Policies, Preservice Education, Professional Development, and the Workplace*. Consensus Study Report, The National Academies Press, Washington, DC.
- Forzani, F.M., 2014. Understanding “core practices” and “practice-based” teacher education: learning from the past. *J. Teach. Educ.* 65 (4), 357–368.
- Garrison, D.R., Anderson, T., Archer, W., 2001. Critical thinking, cognitive presence, and computer conferencing in distance education. *Am. J. Dist. Educ.* 15 (1), 7–23.
- Gess-Newsome, J., 2015. A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit. In: Berry, A., Friedrichsen, J., Loughran, J. (Eds.), *Re-examining pedagogical content knowledge in science education*. Routledge.
- Grossman, P., 2018. *Teaching Core Practices in Teacher Education*. Harvard Education Press, Cambridge, MA.
- Hadar, L.L., Ergas, O., Alpert, B., Ariav, T., 2020. Rethinking teacher education in a VUCA world: student teachers’ social-emotional competencies during the Covid-19 crisis. *Eur. J. Teach. Educ.* 43 (4), 573–586.
- Hadar, L.L., Alpert, B., Ariav, T., 2020. The response of clinical practice curriculum in teacher education to the Covid-19 breakout: a case study from Israel. *Prospects* 1–14.
- Haddix, M., 2015. Preparing community-engaged teachers. *Theory Into Pract.* 54 (1), 63–70.
- Hargreaves, A., Fullan, M., 2020. Professional capital after the pandemic: revisiting and revising classic understandings of teachers’ work. *J. Prof. Cap. Community* 5 (3–4), 327–336.
- Harrison, J., Dymoke, S., Pell, T., 2006. Mentoring beginning teachers in secondary schools: an analysis of practice. *Teach. Teach. Educ.* 22 (8), 1055–1067.
- Hill, C., Rosehart, P., St Helene, J., Sadhra, S., 2020. What kind of educator does the world need today? Reimagining teacher education in post-pandemic Canada. *J. Educ. Teach.* 46 (4), 565–575.
- Hoadley, C., Kali, Y., 2019. Five waves of conceptualizing knowledge and learning for our future in a networked society. In: Kali, Y., Baram-Tsabari, A., Schejter, A. (Eds.), *Learning in a Networked Society*. Computer-Supported Collaborative Learning Series, vol. 17. Springer, Cham, pp. 1–21.
- Hobson, A.J., Ashby, P., Malderez, A., Tomlinson, P.D., 2009. Mentoring beginning teachers: what we know and what we don’t. *Teach. Teach. Educ.* 25 (1), 207–216.
- Hofstede, G., 1991. Empirical models of cultural differences. In: Bleichrodt, N., Drenth, P.J.D. (Eds.), *Contemporary Issues in Cross-Cultural Psychology*. Swets & Zeitlinger Publishers, pp. 4–20.

- Hofstede, G., 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, second ed. Sage Publications, Thousand Oaks, CA.
- Huber, S.G., Helm, C., 2020. COVID-19 and schooling: evaluation, assessment and accountability in times of crises—reacting quickly to explore key issues for policy, practice and research with the school barometer. *Educ. Assess. Eval. Account.* 32 (2), 237–270.
- Inbar-Lourie, O., Levi, T., 2020. Assessment literacy as praxis: mediating teacher knowledge of assessment-for-learning practices. In: *Toward a Reconceptualization of Second Language Classroom Assessment*. Springer, Cham, pp. 241–259.
- Judd, J., Rember, B.A., Pellegrini, T., Ludlow, B., Meisner, J., 2020. “This is Not Teaching”: The Effects of COVID-19 on Teachers. https://www.socialpublishersfoundation.org/knowledge_base/this-is-not-teaching-the-effects-of-covid-19-on-teachers/.
- Kavanagh, S.S., Danielson, K.A., 2020. Practicing justice, justifying practice: toward critical practice teacher education. *Am. Educ. Res. J.* 57 (1), 69–105.
- Kavanagh, S.S., Metz, M., Hauser, M., Fogo, B., Taylor, M.W., Carlson, J., 2020. Practicing responsiveness: using approximations of teaching to develop teachers' responsiveness to students' ideas. *J. Teach. Educ.* 71 (1), 94–107.
- Kielhofner, G., 2009. The model of human occupation. In: *Conceptual Foundations of Occupational Therapy Practice*, fourth ed. FA Davis Company, Philadelphia, PA, pp. 147–174.
- Leacock, C.J., Warrican, S.J., 2020. Helping teachers to respond to COVID-19 in the Eastern Caribbean: issues of readiness, equity and care. *J. Educ. Teach.* 46 (4), 1–10.
- Levin, B., 2013. To know is not enough: research knowledge and its use. *Rev. Educ.* 1 (1), 2–31.
- Lockee, B.B., 2020. Shifting digital, shifting context: (re)considering teacher professional development for online and blended learning in the COVID-19 era. *Educ. Technol. Res. Dev.* 1–4.
- Loose, C.C., Ryan, M.G., 2020. Cultivating teachers when the school doors are shut: two teacher-educators reflect on supervision, instruction, change and opportunity during the Covid-19 pandemic. *Front. Educ.* 5, 231.
- Márquez, G.G., 2007. *Love in the Time of Cholera*, vol. 235. Everyman's Library.
- Mayer, D., Dixon, M., Kline, J., Kostogriz, A., Moss, J., Rowan, L., Walker-Gibbs, B., White, S., 2017. Studying the effectiveness of teacher education. In: Mayer, D., et al. (Eds.), *Studying the Effectiveness of Teacher Education*. Springer, Singapore, pp. 13–26.
- Mena, J., Hennissen, P., Loughran, J., 2017. Developmental learning approaches to teaching: stages of epistemological thinking and professional expertise. In: Clandinin, D.J., Husu, J. (Eds.), *The Sage Handbook of Research on Teacher Education*. Sage, Thousands Oaks, pp. 513–527.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach. Coll. Rec.* 108 (6), 1017–1054.
- Mullen, C.A., 2020. Mentoring epistemologies beyond Western modalities. In: *The Oxford Research Encyclopedia of Education*. Oxford University Press, New York, pp. 1–28.
- Murrell Jr., P.C., 2001. *The Community Teacher: A New Framework for Effective Urban Teaching*. Teachers College Press, Columbia University, New York, NY, p. 10027.
- Mutton, T., 2020. Teacher education and COVID-19: responses and opportunities for new pedagogical initiatives. *J. Educ. Teach.* 46 (4), 439–441.
- Nelson, S.R., Leffler, J.C., Hansen, B.A., 2009. *Toward a Research Agenda for Understanding and Improving the Use of Research Evidence*. Northwest Regional Educational Laboratory, Portland, Oregon.
- Nir, A.E., 2020. Educational centralization as a catalyst for coordination: myth or practice? *J. Educ. Adm.* 59 (1), 116–131.
- Orland-Barak, L., Maskit, D., 2011. Novices “in story”: what first-year teachers' narratives reveal about the shady corners of teaching. *Teach. Teach.* 17 (4), 435–450.
- Orland-Barak, L., Wang, J., 2020. Teacher mentoring in service of preservice teachers' learning to teach: conceptual bases, characteristics, and challenges for teacher education reform. *J. Teach. Educ.* <https://doi.org/10.1177/0022487119894230>.
- Orland-Barak, L., Yinon, H., 2005. Sometimes a novice and sometimes an expert: mentors' professional expertise as revealed through their stories of critical incidents. *Oxf. Rev. Educ.* 31 (4), 557–578.
- Orland-Barak, L., 2010. Introduction: learning to mentor-as-praxis foundations for a curriculum in teacher education. In: *Learning to Mentor-As-Praxis*. Springer, Boston, MA, pp. 1–13.
- Orland-Barak, L., 2014. Mediation in mentoring: a synthesis of studies in teaching and teacher education. *Teach. Teach. Educ.* 44, 180–188.
- Orland-Barak, L., 2016. Mentoring. In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education (Volume II)*. Springer Publications, Singapore, pp. 105–141.
- Pang, M., 2016. Pedagogical reasoning in EFL/ESL teaching: revisiting the importance of teaching lesson planning in second language teacher education. *Tesol Q.* 50 (1), 246–263.
- Philip, T.M., Souto-Manning, M., Anderson, L., Horn, I.J., Carter Andrews, D., Stillman, J., Varghese, M., 2019. Making justice peripheral by constructing practice as “core”: how the increasing prominence of core practices challenges teacher education. *J. Teach. Educ.* 70 (3), 251–264.
- Philipsen, B., Tondeur, J., Roblin, N.P., Vanslambrouck, S., Zhu, C., 2019. Improving teacher professional development for online and blended learning: a systematic meta-aggregative review. *Educ. Technol. Res. Dev.* 67 (5), 1145–1174.
- Reich, J., et al., 2020. Remote learning guidance from state education agencies during the COVID-19 pandemic: a first look. Retrieved from osf.io/k6zy/.
- Richter, D., Kunter, M., Lüdke, O., Klusmann, U., Anders, Y., Baumert, J., 2013. How different mentoring approaches affect beginning teachers' development in the first years of practice. *Teach. Teach. Educ.* 36, 166–177.
- Robinson, M., Rusznyak, L., 2020. Learning to teach without school-based experience: conundrums and possibilities in a South African context. *J. Educ. Teach.* 46 (4), 517–527.
- Roehrig, A.D., Bohn, C.M., Turner, J.E., Pressley, M., 2008. Mentoring beginning primary teachers for exemplary teaching practices. *Teach. Teach. Educ.* 24 (3), 684–702.
- Schwartz, S.H., 2009. Basic values: how they motivate and inhibit prosocial behavior. In: Mikulincer, M., Shaver, P.R. (Eds.), *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature*. American Psychological Association, Washington, DC, pp. 221–241.
- Scull, J., Phillips, M., Sharma, U., Garnier, K., 2020. Innovations in teacher education at the time of COVID-19: an Australian perspective. *J. Educ. Teach.* 46 (4), 497–506.
- Sepulveda-Escobar, P., Morrison, A., 2020. Online teaching placement during the COVID-19 pandemic in Chile: challenges and opportunities. *Eur. J. Teach. Educ.* 43 (4), 587–607.
- Shelley, P.B., 1845. *Essays, Letters From Abroad, Translations and Fragments*. William Smith.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15 (2), 4–14.
- Sleeter, C., 2014. Toward teacher education research that informs policy. *Educ. Res.* 43 (3), 146–153.
- Thompson, A., Darwich, L., Bartlett, L., 2020. Not remotely familiar: how COVID-19 is reshaping teachers' work and the implications for teacher education. *Northwest J. Teach. Educ.* 15 (2), 2.
- Van Den Berg, G., 2020. Context matters: student experiences of interaction in open distance learning. *Turk. Online J. Dist. Educ.* 21 (4), 223–236.
- Villegas, A.M., Lucas, T., 2002. *Educating Culturally Responsive Teachers: A Coherent Approach*. State University of New York Press, Albany.
- Wang, J., Odell, S.J., 2002. Mentored learning to teach according to standards-based reform: a critical review. *Rev. Educ. Res.* 72 (3), 481–546.
- Young, K., Ashby, D., Boaz, A., Grayson, L., 2002. Social science and the evidence-based policy movement. *Soc. Pol. Soc.* 1 (3), 215.
- Zeichner, K., 2012. The turn once again toward practice-based teacher education. *J. Teach. Educ.* 63 (5), 376–382.
- Zeichner, K., 2019. Moving beyond asset, equity, and justice- oriented teacher education. *Teach. Coll. Rec.* 121 (4), 1–5.
- Zeichner, K., 2020. Preparing teachers as democratic professionals. *Action Teach. Educ.* 42 (1), 38–48.
- Zhang, W., Wang, Y., Yang, L., Wang, C., 2020. Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak. *J. Risk Financ. Manag.* 13 (55), 2–6.

Comparative teacher education research in turbulent times

Ian Menter^{a,b} and Maria Teresa Tatto^{b,c}, ^a University of Oxford, Oxford, United Kingdom; ^b Kazan Federal University, Kazan, Russia; and ^c Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	740
Market models and performativity in teacher education: the impact of globalization	741
The market model, performativity and implications for education and teacher education	741
Marketization and performativity are not universally adopted	742
Summary	742
National populism: its influence on teacher education	742
Education and the nation-state	743
Influence on teacher education	743
Appealing to populist ideas and “evidence” to dismantle the profession	744
Responses from educationists	744
Summary	745
Global instability: climate change, pandemic, conflict, migration	745
Climate change	745
Pandemic	745
Conflict and migration	746
Summary	746
Resources of hope: the contribution of comparative studies	746
Conclusion	748
References	748

Introduction

The deployment of comparative approaches in the study of education now has a considerable history, stretching back to the 19th century. However, from the early days when the main task of comparative education was to describe national systems of education and to identify similarities and differences between them (e.g., King, 1967), we can now see a much greater range of perspectives within the field. We see for example much greater attention being given to the significance of “context” (Crossley and Watson, 2003; Phillips and Schweisfurth, 2008), we see greater understanding of “policy processes” (Bowe et al., 1992), and of the effects of “globalization” (Rizvi and Lingard, 2010). We can also see how comparative studies do not have to be undertaken only at the level of the national, or “macro,” we can also undertake comparative approaches to understanding policy and practice as realized at the institutional (meso) or indeed personal (micro) level. In other words, comparative studies may be undertaken at macro, meso and micro levels, each of which can reveal important insights (Tatto et al., 2018a,b).

In this chapter, while we fully acknowledge the pervasive influences in contemporary teacher education of “neoliberalism” and “performativity,” our purpose is to demonstrate the enormous potential benefits in taking comparative approaches. The argument developed in the chapter is that it is not only the research community that gains from comparative study, but that teachers themselves may enrich and enhance their own working lives and professional identities through connecting with such comparative investigations (Tatto, 2021). At a time when the impact of climate change is becoming more widely recognized as a threat to the very existence of life on this planet and a time when a global pandemic has reminded us how vulnerable every human being is, we also see many instances of increasing national populism (Eatwell and Goodwin, 2018) and “nativism,” as well as an escalation of associated phenomena including physical conflict and transnational migration.

In the sections that follow we move from considering the respective influences of economic factors, through to politics and then to ecology and health. In each of these we emphasize the importance of taking an historical approach to our comparative analysis. Thus, in the next section we explore the developing impacts of globalization on teacher education in more detail, seeking to show how these impacts are manifested in different experiences at different levels and how in any particular contexts they are influenced by the history and culture of that setting (Alexander, 2000). We then turn to consider the growth of nationalism and populism and how they have had an influence on policies and practices of teacher education, again in different ways and at various levels. Following that, we consider the relatively recently identified threat of climate change, how it is leading to a resetting of priorities in teacher education, a process which may have been accelerated by the occurrence of the COVID-19 pandemic, which commenced in early 2020. We move toward a conclusion by seeking to identify some “resources of hope” (Williams, 1989), which may offer positive ways forward, enabling teacher education to play a role in ensuring global stability, sustainability, peace and democracy in the future.

Market models and performativity in teacher education: the impact of globalization

The introduction of market models in education aided by globalization has had negative yet predictable consequences for the field of education and teacher education. The notion of predictability is important here as it underlines the purposeful character of these reforms.

The definition of the market model applied to education and its effects are well known. In the US for instance according to an insightful analysis by [Tucker \(2012\)](#), analysts' arguments were constructed as a direct attack against public education which they represented as a "government monopoly where teachers and administrators lacked incentives to control costs or deliver quality." He argues that following market models, the proposed reform would reallocate the public funds raised for the schools into the hands of parents by promoting school choice. The notion of school choice meant that in theory parents would be able to choose the best schools for their children as measured by higher achievement scores in international, national and local tests, with the idea that levels of achievement would rise and costs would go down. These proposals were rejected initially but continuing arguments by analysts and the increasing costs of education accompanied by the unwillingness of politicians to raise taxes to support schools created a strong rationale to engage in a market reform of the education system with the strong support from the U.S. Department of Education. The reform was backed by systems of rewards and punishments, the latter being school closures as a result of students' low achievement and poor teacher evaluations.

Diane [Ravitch \(2016\)](#), a US historian, has written extensively about the negative consequences of this turn to the market which includes privatization, the rise of charter schools, parental choice, and a trend toward performativity marked by increased standardization, testing, and teacher evaluations. These reforms have given rise to attacks against public schools, the teaching profession and teacher education. Contrary to what the reforms promised we have seen an increased lack of equitable access to quality education for disadvantaged students, a dramatic drop in the number of teachers entering the profession and an increase in the number of teachers leaving their posts. This has accelerated as a result of the COVID pandemic.

The market model, performativity and implications for education and teacher education

The turn to the market in education began in some countries such as England in the 1980s and in the US around the 1990s. Since that time evidence confirms the limited effect of this reform in achieving its alleged aim to improve equitable access to quality of education for the school-going population or an increase in learning as measured by achievement tests ([Waslander et al., 2010](#)). Market theory tells us that "market systems allow individuals to exchange goods and services voluntarily based on prices, without knowing one another" ([Kling, 2021](#)). This is the first of many problems with applying the model to education. The model goes against the culture of schooling. As anyone who has had a child in school or anyone who has taught or administered in schools knows, schools are intrinsically relational places. Children and their families get to know each other quite well, and the parent grapevine is quite effective in nominating the best teachers as per common perception (only rarely are decisions based on results of teacher evaluations). Most of the resulting decisions happen locally which (somewhat paradoxically) dramatically reduces the opportunities for choice. Another key problem of the theory is that education markets are characterized by the relative inelasticity of demand and supply. As [Waslander et al. \(2010:65\)](#) point out after examining several education systems across countries, "for markets to function efficiently, buyers need to respond to changes in the price/quality relation (elasticity of demand), and sellers need to respond to customer demand by either diminishing or extending production (elasticity of supply)" in other words parents are typically satisfied with local schools, and schools have difficulty "growing easily when increasing numbers of parents and pupils wish to attend," in other words, physical space may be limited, it is difficult to expand quickly or schools simply chose not to change/expand.

While international evidence shows that market reforms have been a failure in accomplishing their alleged intended objectives ([Waslander et al., 2010](#)), they have had dire consequences for the educational systems, notably in altering the way people relate to each other. According to [Fiske, 2004](#), as cited in [Kling, 2021](#)), an anthropologist who studies aspects of sociality across cultures, there are four basic models through which people interact, these help us understand how market reforms have altered relationships in educational systems. These models are: Communal Sharing (where the idea is to share such things as resources and responsibilities among the members of a group); Authority Ranking (where people have asymmetric positions in a linear hierarchy, such as administrators to teachers and teachers to students); Equality Matching (where people keep track of the balance or difference among participants and know what would be required to restore balance such as turn-taking); and Market Pricing (where relationships are oriented to socially meaningful ratios such as wages, or cost-benefit analyses). According to Fiske in this last model, money need not be the medium, and relationships need not be selfish, competitive, maximizing, or materialistic. All these four relational models can be found in schools but the turn to the market has served to destabilize these delicate relationships which indeed require a basic element that has been eroded in the process, namely trust in schools ([Bryk and Schneider, 2004](#)). Communal sharing where teachers collaborate with teachers, administrators and parents in developing school policies and curriculum standards and in supporting student learning has given way to regulatory mechanisms, where standards, curriculum and testing come from outside the schools and places teachers in the position of compliance lest they risk low achievement scores for their pupils. The capacity for Equality Matching among participants has been lost, as people (teachers, the school community) under the pressures of accountability and performativity, have lost the capacity to keep track of the balance or difference among participants and now lack the tools or the resources (such as time) that would have allowed them to restore balance. Under the Market Pricing model, the ideal that relationships need not be selfish, competitive, maximizing, or materialistic, has also been transformed by the introduction of parent

choice, testing, and teacher evaluations and the system of rewards and punishments in schools via inspections and accountability (Cochran-Smith et al., 2018; Tatto, 2007).

The turn to the market has also affected the university project of teacher education (Furlong, 2013). In the mid-1980s, in both England and the US, alternative routes to teacher certification emerged under the rationale that diversification would be a way to supply more teachers to schools and also as a response to growing criticism of university-based teacher education. These routes have been growing steadily in the US as well as in England and the quality of their graduates vary widely. Nevertheless, these routes have been legitimized and strongly supported by the U.S. Department of Education (U.S. Department of Education, Office of Innovation and Improvement, 2004), and England's Department for Education (DfE). In the US these programs have proliferated with the major initiatives being Teach For America, Academy for Urban School Leadership, New Teacher Project, American Board for Certification of Teacher Excellence (ABCTE), New York City Teaching Fellows, Mississippi Teacher Corps, and Educators of Change. Likewise, in England in addition to the university-led graduate teacher training program, known as the Post Graduate Certificate in Education (PGCE), other programs are School Direct (tuition fee or salaried), school-centered initial teacher training partnership (SCITTs), the Postgraduate Teaching Apprenticeship, the Teach First Leadership Development Program and Assessment Only, among others. Comparisons between those attending university-based teacher education and alternative programs show mixed results (Matsko et al., 2021; Sass, 2015; Stoddart and Floden, 1995; Brown, 2018; Sorensen, 2019).

Marketization and performativity are not universally adopted

While the market model has been adopted and defended by governments in several countries despite existing evidence that it does not work and is damaging for schools, teachers and students, other countries have been able to avoid partially or fully such a trend. Notably in Finland, Sahlberg (2012) has denounced the global education reform movement (GERM) because it has increased rather than reduced global inequality and has undermined the teaching profession. Japan has also avoided such trends from affecting its education system (LeTendre and Ikoma, 2019). Other authors have carefully documented the effects of these reforms and how the purposes of the reforms are misaligned with the purposes of education (values, human rights, democracy), and with the values and motives of teachers (Childs and Menter, 2013; Tatto et al., 2018).

When examining countries that have not adopted the market model in education and are ahead in educational outcomes, several features are noticeable, such as investing in children's wellbeing and early childhood education. Investment in teachers emerges as an important priority as well. Not only are teacher preparation programs more selective in enrolling high ability graduates from high school, but teachers are also prepared to master the subjects they are going to teach and pedagogy to teach them, teacher pay is higher, and teachers are specifically prepared to teach the required courses for their students (Darling-Hammond, 2017; Tucker, 2012). Other elements include a strong local or national curriculum and international level standards, and less testing.

Summary

The turn-to-the market in education has failed in creating a better, more equitable and resilient system. In the US investment in charter schools and school choice was the wrong type of investment (Berliner, 2020), it has benefitted the business interests of those that founded these schools at the expense of public schools (Tucker, 2012), and it has decimated the teaching profession. The emphasis on performativity through standards and teacher evaluation has not been less damaging. The results are as could be expected, teachers are leaving the profession and many who would have become teachers are not entering the profession. As of this writing schools in the US are lacking in teachers with the state of Florida, for instance, announcing nearly 5000 teacher vacancies, the situation is similar in other states including Arizona where one-quarter of teaching positions are vacant (Heyward, 2021; Khmara, 2020). The major impact of this reform is on the children of disadvantaged populations. The COVID pandemic has revealed with more clarity how much the education sector has been devastated by these reforms.

National populism: its influence on teacher education

Globalization and market-based reforms are not the only important movements that are affecting education and teacher education. The rise of national populism, another deliberate and purposeful movement, is an important and possibly permanent drive that has acquired much force in the last few years (exemplified by the rise of Trump in the US, and Johnson in Britain), marked by social movements in the US culminating with the march to the Capitol Building on January 6th 2021 (also called an insurrection led by Trump) in attempts to delegitimize and stop a cleanly elected president Biden from taking office). Across the Atlantic, the triumph of Brexit in Britain effectively separated the four UK nations from the European Union (EU). Populist parties in the EU are increasingly occupying parliamentary seats, most notably in Hungary and Poland. On the other side of the world in Brazil, a national populist candidate was elected president in 2018. The rise of far-right groups all over the world is another outcome of these populist movements. The concern with these movements is that they may turn into authoritarian governments—as indeed was shown by the actions of Presidents Trump (US) and Bolsonaro (Brazil) who, during the COVID pandemic, have contributed by their actions to the death of many by their lack of belief in science, which resulted in misinformation and lack of vaccinations for many citizens, a situation which is still true in Brazil at the time of writing (Pereira, 2020).

In their book, *National Populism: The Revolt Against Liberal Democracy*, Eatwell and Goodwin (2018) suggest that as much as the election of Trump and the passing of Brexit surprised many, it would be wrong to define national populism by these events alone. Rather they argue that national populism can be seen as a democratic movement that decries Western liberal hegemony with its emphasis on individual rights (rather than communal obligations and solidarity), ethnic change, economic and political globalization. To them national populism is

driven by a deeper desire [by the people who feel unempowered] to bring a broader set of values back onto the agenda and regain their voice; to reassert the primacy of the nation over distant and unaccountable international organizations; to reassert cherished and rooted national identities over rootless and diffused transnational ones; to reassert the importance of stability and conformity over the never-ending and disruptive instability that flows from globalization and ethnic change, and to reassert the will of the people over those of elitist liberal democrats who appear increasingly detached from the life experiences and outlooks of the average citizen.

Eatwell and Goodwin (2018: xxxii).

They argue that the movements we are witnessing could be the “beginning of an era of political fragmentation volatility and disruption” accelerated by a breakage of the “bonds between the people and the traditional parties [...] and unprecedented ethnic change and rising inequality.”

The same authors, Eatwell and Goodwin (2018: xxi ff.), direct us to four long-term trends or “historic shifts” that have fed into the increasing development of national populism and that are not always grounded on objective reality:

- *distrust* of politicians and institutions fueled by the perceived elitist nature of liberal democracy and lack of voice in the national conversation;
- *destruction* of the national group’s historic identity and established ways of life fueled by perceived threats caused by increased immigration and “hyper” ethnic change;
- *deprivation* or a sense of “relative” deprivation fueled by the perceived notion that certain groups are losing out relative to others as a result of globalization and market-based reforms, and rising inequalities of income and wealth in the West, which is also associated with a loss of faith in a better future;
- *de-alignment* or the weakening bonds between the traditional mainstream parties and the people, which makes political parties across the West more volatile, fragmented and unpredictable.

The authors argue that these trends need to be analyzed together as national populist leaders feed on these trends and have been able to persuade an increasing number of people to follow their lead, many toward authoritarian populism.

Education and the nation-state

A key question that emerges in the West as we witness these rapid changes and increased instability of historic forms of education is the role of the state in education and particularly in teacher education. Traditionally education systems were controlled by the state because, in great part, education was seen as indispensable for the development of healthy democracies. An important argument that supported the key role of education in the US came from Dewey (1916) who argued for the need to provide access to quality education for all citizens (universal public education) to support the advancement of the individual and society. His ideas as well as those of other philosophers and sociologists of education were disseminated widely and influenced the underlying philosophy of education for many years not only in the US but globally, and provided the rationale for the state to take most of the responsibility for such an important task.

As discussed above, the introduction to market models and the global influences on education gave rise to decentralization and privatization where intentionally the state devolved its responsibility for education to the regional departments of education or the private sector. Green (1997: 2–5) refers to this trend as the “post-national” era and wonders “how far can national states control education systems, in a world of global markets, supranational political organization [and] pressures for international convergence”? The case of England operating under market models illustrates how the actions of an authoritarian state is managing to destroy years of progress in the education sector using misleading arguments.

At this point, it would be important to once again turn to a comparative exploration and analyze developments in the East Asian countries and Central and Eastern Europe and their process of democratization where, as Green argues, the state tends to continue to see education as playing a central role in constructing a sense of national identity, language and culture. It may well be that in the West and after the COVID pandemic, education will be seen as a renewed priority and that the education and proper remuneration of teachers will be seen as essential in the development of modern citizens who, in Green’s words “would have a broader understanding of the interdependence of nations, the diversity of societies, and the global nature of solutions to the world’s problems” (Green, 1997).

Influence on teacher education

While globalization, marketization, and performativity have served to de-professionalize teaching, national populism has served to minimize the importance of the knowledge that teachers acquire in university-based programs. Many in the national populism

movement, including Johnson and Trump, have publicly attacked theory and science with serious consequences for the school curriculum and the teacher education curriculum. Currently, in England and the US, many individuals are entering teaching without the knowledge and skills needed for the job. The market reforms and the re-definition of what counts as a qualified teacher as well as the re-definition of the teacher role as performing to external standards guided by market-oriented purposes of education have discouraged many from entering the profession. Racism and discrimination which has been fueled by the national populist beliefs that increased during the Johnson and Trump administrations have prevented many students from entering the profession.

Appealing to populist ideas and “evidence” to dismantle the profession

In 2010 in England, a White Paper, *The Importance of Teaching* (DfE, 2010), produced under the leadership of Michael Gove, the then Secretary of State for Education, introduced a “revolutionary reform,” using as a justification England’s lower scores in the OECD PISA test on literacy and mathematics and a faulty analogy between teacher preparation and medical preparation, devolved the responsibility for teacher preparation to the schools which he saw as equivalent to hospitals! When doctors become interns and then residents in hospitals they are already *doctors* (they have a credential—a medical degree) and have undergone many years of study in schools of medicine in higher education institutions (in the US, four years, including rotations in different hospitals). Nevertheless, the English White Paper encouraged the ‘best schools’ to become independent, that is to convert to “academy” schools (the equivalent to charter schools in the US) and to group into supra-organizations called Multi Academy Trusts (MATs). This was meant to free schools and school personnel in order to “innovate and concentrate in education,” albeit under more strict regulations or accountability systems led by the Ofsted inspections regime.

Moving forward to 2021 also in England, the “Initial Teacher Training ITT Market Review Report” (Department for Education, 2021), reinforces and expands such reform by recommending a “new set of quality requirements to be implemented by all Initial Teacher Training providers of courses (programs) that lead to QTS [Qualified Teacher Status] and a robust accreditation process [...] to meet the exacting quality requirements in full upon accreditation and on a continuing basis” (p. 4). The review led by CEOs of various school trusts and directors of programs such as Teach First and one former curriculum research and design individual, recommended a re-design of the ITT curriculum, the identification of placement schools with trained mentors, a detailed assessment framework, a quality assurance requirement for accreditation, and partnerships (e.g., between HEI and SCITTs). The claim is that these recommendations are “fully aligned with the career-long, evidence-based professional development offer for all teachers” (p. 5). The introduction to the report only mentions higher education institutions in association with the SCITT and end by remarking the importance of “existing strong SCITTs and School Direct Lead Schools becoming part of the reshaped market.”

Responses from educationists

These sweeping reforms continue to be implemented without solid evidence of success despite the frequent mention of “evidence” in both reports. Helgetun and Menter (2020) analyzed how “evidence” is used in both of the above documents and urged educationists and analysts to carefully examine the nature of such evidence and question whether such evidence warrants the ever-evolving policies affecting teaching and teacher education. In other words, this evidence does not come from the systematic empirical evaluation of these reforms but from disparate pieces of work put together to advocate the reformers’ position.

The vulnerability of teacher education and the teaching profession more generally is not in most cases due to the quality of the provision but rather to the natural variability of the field. A recent comparative book examined the nature of professional knowledge in 12 countries, seeking to answer the larger question of how conceptions of knowledge across the countries shape teacher education and the knowledge and skills of teachers (Tatto and Menter, 2019). The review found that such knowledge is quite varied and while each system has the same purpose (to prepare high-quality teachers) the field as a whole is fragmented concerning the pedagogy and the actual knowledge teachers must possess. While this fragmentation is natural given the way teacher education programs must respond to cultures within and across settings, this fragmentation has allowed the market to intervene in many of these countries.

Given such fragmentation can we argue that there is a universal discipline of education such as in other areas of knowledge (e.g., medical knowledge)? Whitty and Furlong (2017) use a comparative analysis to explore what is understood as educational knowledge and whether such knowledge constitutes the discipline of education. They find that there are tensions that boil down to the purposes of education. In Germany for instance the definition of education as a discipline has benefitted from conceiving of education as deeply interdisciplinary which is seen as necessary to explain complex education issues—the insistence that this interdisciplinary composite or “science of education” is an autonomous discipline makes the German approach distinctive. In contrast in England, where the purposes of education are more practical—to prepare teachers for the job narrowly defined—allows more direct control from outside the profession and even from the markets. The conclusion is that education is constantly in tension between what national stakeholders require education to do such as inculcating values and attitudes or transmitting practical skills. The direction for England is currently dominated by the need to transmit practical skills which align with the dominance of the market model to education and is continuously reinforced by the OECD PISA test. Against this tendency toward convergence, however, the study highlights national variation as a source of strength, a finding that also emerged from an analysis of the U.S. decentralized system of education (Tatto and Clark, 2019).

Summary

The fragmentation of the education field in countries such as England and—less so in the decentralized system of the US—where the market has been allowed to operate, leaves the profession vulnerable to globalization trends and the influence of national populism movements that reject the value of theory and science, increase distrust in schools and teachers while introducing false narratives and divisiveness that seem directed at threatening democracy in possibly irreparable ways. The need to strengthen the teaching profession has never been more urgent.

Global instability: climate change, pandemic, conflict, migration

In parallel with the phenomena discussed in the previous two sections, the rise of neoliberal marketization and the growth of new manifestations of nationalism, we have seen evidence of growing instability around the world. So, in this section we consider the implications of climate change, the COVID-19 pandemic, social conflicts and migration. What impact do processes such as these have on schools and teacher education?

Climate change

The process of global warming has been linked to an increasing number of major environmental disasters, including floods, storms and fires. In many instances these have directly affected schools and other educational establishments. But the global effects of climate change pose a major threat to the existence of plant and animal species, including of course humans. What are the implications for teacher education and for the teaching profession? Much of the political action that has been undertaken in response to climate change has been led by young people in many parts of the world and the school strike actions, notably led by Greta Thunberg, have challenged schools, teachers and teacher educators in several countries. If children are absenting themselves from school to protest at the failure of their seniors to act on climate change, should teachers, student teachers and teacher educators offer their support? If school attendance is mandatory, then offering such support could be construed as an illegal act on the part of those adults (as well as the children). These are very real moral dilemmas for many professional educators.

Furthermore, what are the curriculum implications, both in schools and in teacher education programs, of global warming and climate change? And how can schools and universities reduce their carbon footprints to a minimum? To date there has been little evidence of significant educational responses to these challenges, with the exception of some environmental education initiatives and some significant developments in forms of outdoor and nature-based learning (Forest Schools for example, see [Knight, 2016](#); [Tiplady and Menter, 2021](#)). Even more radical arguments have been advanced at various times that schooling itself (and teacher education) is in some ways “anti-human” and destructive—ranging from Illich’s treatise on “deschooling society” ([Illich, 1971](#)) through to more recent accounts suggesting that children may be best educating themselves ([Gray, 2016](#)). Indeed, much contemporary conventional teacher education can be accused of leading to conformity and a kind of social stasis and/or reproduction ([Menter, 2016](#)).

Pandemic

The rapid spread of the COVID-19 virus around the world in 2020 had very direct impacts upon teacher education. Nearly every country around the world undertook various forms of “lockdown” in attempts to reduce the spread of the disease and to minimize the number of premature deaths. These lockdowns included the temporary closure of both schools and universities, the two key institutions involved in the provision of teacher education. Two aspects that arose are of particular significance in teacher education, the challenges to the “practicum” and the suspension of face to face teaching.

The practicum is universally seen as a core element of initial teacher education. Every ITE program in the world includes some element of school experience for the beginning teacher. Typically, the beginning teacher spends periods of time in school classrooms, taking increasing responsibility for teaching children, usually with the support of both teachers in school and of visiting staff from the university providing the program. The particular details and structures of the practicum vary enormously both between different national and state settings and even within these settings. However, it is a reasonable generalization to suggest that the beginning teacher’s performance in the school setting is assessed against a set of teaching standards or competences. That assessment is seen as a key indicator of that person’s preparedness to enter the profession as a qualified teacher. The challenge of completing a teacher education program when schools are closed was therefore considerable and led to some imaginative responses. In some settings even when schools reopened after a period of lockdown, there was great caution about allowing any “non-essential” personnel into the setting in the light of concerns about increasing the risk of the virus spreading.

The second key element of the effects of the pandemic was in relation to the mode of pedagogy deployed in teacher education. Students were not only unable to attend the practicum setting—the school experience—they were also unable to attend classes in their home university, as these institutions were also closed for lengthy periods. This led to a surge in the use of electronic, virtual means of communication. Various virtual learning environments (VLEs) were deployed, including Blackboard, Canvas, Webex and Minerva Collaborate. Other platforms were used for meetings and tutorials including Zoom, TEAMS and Skype. Such systems had

been in use in many contexts before the pandemic struck but the rapid expansion of their use during 2020 was quite phenomenal. The nature of the pedagogical relationships between teachers and learners was dramatically altered through the use of screens and the affective dimensions of interaction were shifted substantially with the reduced possibilities of effective non-verbal communication, usually a key element in learning, especially in one-to-one teaching situations.

Some of these issues arising from the pandemic are explored in more detail by teacher educators in special issues of two journals—see [Flores and Swennen \(2020\)](#) and [O'Meara and Hordatt Gentles \(2020\)](#). Reviewing the impact on the teaching profession in the UK of challenges such as those discussed above, [Breslin \(2021\)](#) suggested that:

In the main, school leaders, teachers and the wider school workforce, and those involved in initial and continuing teacher education, are emerging from lockdown with their professional status enhanced and a much richer public understanding of the complexity and range of the work that they do ...
[Breslin \(2021:152\)](#).

In similar vein and somewhat speculatively, in considering the provision of initial teacher education (ITE) “at scale,” Brooks concludes:

Much of the social media and educational press are speculating on what the impact of COVID-19 and the coronavirus are likely to be, including the impact on universities. Within England specifically, the status of teachers appears to be on the rise ..., and the demand for teacher education places is likely to grow as the effects on the wider economy are more broadly felt. Under these circumstances, questions of both quality and scale are likely to become more prominent.
[Brooks \(2021: 210\)](#).

Conflict and migration

Another aspect of global instability that may affect teacher education is the presence of political conflicts within societies. Thus, for example, we see that the disruptive effects of sectarian strife in Northern Ireland have meant that it has proved very difficult to reform and rationalize provision there, in spite of numerous reviews of teacher education and in spite of repeated efforts to integrate and bring harmony through curriculum reform ([Clarke and Magennis, 2016](#)). Physical violence directed toward teachers and teacher educators has been an issue in many divided societies, including South Africa, Syria and Afghanistan.

Conflict also frequently leads to forced migration, which can be another destabilizing global process. The movement of families and sometimes unaccompanied children across national boundaries is often a process fraught with difficulties, both physical and political. The arrival in schools of children who do not speak the national language can present particular challenges for teachers and for teacher educators. Furthermore, it is not uncommon for refugee communities to include teachers who are qualified in their country of origin. However, these qualifications are not always recognized in their country of arrival and so their potential contribution to the education service may be denied ([Kum et al., 2010](#)).

Summary

There is a continuing likelihood that processes such as those discussed in this section will lead to social inequality and injustice. While education cannot compensate for society ([Bernstein, 1970](#)) any more than teacher education can compensate for education, both schools and teacher education can play a part in addressing such social inequities. As Brooks has suggested in relation to teacher education:

Teacher education, with its complexity of stakeholders and interested parties, may be difficult to change, but teacher educators are adept at doing so, often navigating contradictory influences in order to provide the best possible experience for new teachers.
[Brooks \(2021:201\)](#).

Each of the global processes discussed here creates new demands on education systems and it is necessary for schools and teacher education institutions to respond positively to them. Some of the responses will be structural, for example expanding schools or teacher education programs, and some will be curricular, for example helping beginning teachers to understand the health risks associated with viral infections, helping them to model ecologically sound behavior for their students or developing skills in celebrating cultural and linguistic diversity.

Resources of hope: the contribution of comparative studies

One collection of essays written by the novelist and cultural theorist Raymond Williams was called *Resources of Hope* ([Williams, 1989](#)). Having previously set out early in the 1960s how he saw social progress in Britain as being a case of a “long revolution”

(Williams, 1961), Williams was becoming increasingly concerned about the growing impact of capitalism and the “free market” on culture and society (Williams, 1983). While he did not have a particular interest in teacher education, he did nevertheless see public education as a key plank in the development of democratic and equitable societies (see Menter, 2022a). In this section of our chapter, in the wake of our discussion of the impact of global turbulence on teacher education, as well as Brooks’ positive assessment just cited, of teacher educators’ ability to promote change, we turn to consider what resources of hope may be pertinent to the contribution that teacher education may make to social progress. We have already seen how a consideration of policies and practices in different geographical settings may enrich our understanding of teacher education and it is in the field of comparative studies that we may claim that many key insights are to be found.

The impact of neoliberalism has been pervasive in education systems around the world and this is no less the case in relation to teacher education (Tatto and Menter, *forthcoming*). Through studying the elements of professional knowledge that are deemed to be important we can detect the influence of national cultures and histories. This can be demonstrated in multi-nation states such as the United Kingdom (e.g., Teacher Education Group, 2016), in comparative studies of two countries (e.g., Tatto et al., 2018; Menter et al., 2017; Helgetun, 2022), as well as in comparative studies across many nations (Tatto et al., 2018, 2020; Tatto and Menter, 2019; Whitty and Furlong, 2017).

The foundations of comparative education were laid down in the 19th century and then developed in the 20th century by scholars such as Michael Sadler and Edmund King (refs). It was toward the end of the 20th century that significant developments took place in comparative research methodology (Crossley and Watson, 2013; Phillips and Schweisfurth, 2007) but also this was a time when there was increasing interest in teacher education, in part arising from the growing political priority being given to it. Mantras such as “the quality of an education system depend on the quality of the teachers working in that system” had great influence. Views of this kind were heightened by reports such as those undertaken by the McKinsey Corporation (Barber and Mourshed, 2007).

The “travel agent” of neoliberalism has been the process of globalization. Rizvi and Lingard (2010) and Spring (2015) have traced how the acceleration of communication processes and the influence of new forms of management have increasingly influenced policy and practice in education. But they also point out how this has not led to a simple convergence of approaches around the world. Rather, we can see distinctive forms of policy development emerging in a process that has been called vernacular globalization (Appadurai, 1996) or a coalescence of global and national characteristics—the “glonacal” (Marginson and Rhoades, 2002). A five-nation study of “culture and pedagogy” undertaken by Robin Alexander set new standards in the careful analysis of the relation between the two and made much reference to processes of teacher education (Alexander, 2000). Or in focusing on teacher education per se, a twelve-nation study examining knowledge, policy and practice in teacher education identified seven themes (one of which was added later) which could be used to delineate the particularities of each system within the wider global processes (Menter, 2019:272). These were:

1. The trajectory toward professionalization and “universitization” in teacher education—or its reverse.
2. The positioning of research in relation to policy and practice
3. Partnership and participants’ roles in teacher education.
4. Power and control, political structures, the relative influence of national/federal/state/regional/local institutions of government.
5. The impact of performativity and accountability and the promotion of standards in teaching.
6. The impact of digitization—technology and communication.
7. The links between pre-service, induction and continuing professional development—the teaching career continuum.

Each of these themes, when applied to a particular context, can tell us much about the condition of teacher education in that context.

However, it is not only in comparative work that we may gain important insights to the processes and politics of teacher education. Single nation studies can also demonstrate how developments and changes in one setting may have much wider significance. For example, Furlong’s analysis of “the university project” in teacher education in England (Furlong, 2013) shows how the connections between higher education and the education of schoolteachers is a critical factor—and this is relevant to any national setting. Or, in a review of teacher education in Russia, past, present and future (Menter 2022b), we can see how wider political changes such as the Russian Revolution of 1917 and the fall of communism late in the 20th century have had very significant impacts on policy and practice.

Reviewing policy processes in teacher education around the world, it is a challenge to find powerful examples of instances where teacher education has taken the lead in the development of democracy or the creation of more socially just societies. There are many more local examples of democratic practices being implemented in schools and local communities than in teacher education. And many of these examples have attracted opposition from the forces of the state. One might cite the work of Paulo Freire in rural settings in Brazil, which not only led to improved literacy but to campaigns for improved working conditions, leading to Freire being exiled (Collins, 1977). In the US, Apple and Beane (1999) documented a number of cases where radical innovation in the cause of “democratic schooling” was undertaken, frequently leading to ostracism from politicians. It is all too rare to find similarly radical approaches being introduced in teacher education, especially in the 21st century as we have seen in so many instances the impact of central or state government’s “tightening grip” (Childs and Menter, 2013).

Given the tendency toward conformity and social reproduction that is evident in so many systems of teacher education, the importance of comparative study cannot be underestimated. An analysis of the links between culture, politics and history is crucial

to developing our understanding and to informing policy development. There is a need for close links between the respective communities of policy, practice and research in teaching and teacher education. These links are currently much stronger in some places than in others and indeed the use of “evidence” remains a highly significant and often contested aspect of these processes (Helgetun and Menter, 2020).

Conclusion

In this chapter we have sought to understand how teacher education has been changing—and continues to change—during a period of considerable global instability. The economics, politics and cultural traditions within any particular context have shaped the patterns of provision and influenced policy and practice. We have seen how some influences occur at a global level, others are at national, institutional and even personal level (Tatto et al., 2018, explores this view more fully). However, a central message to be derived from our analysis is that investigation of any particular system will be greatly enhanced by ensuring that comparisons are made with other settings. It is through a comparative lens that we can begin to appreciate the interplay of structure and agency in the world of teacher education. We can begin to explore how it may be possible for those with specialist expertise and experience within teacher education to bring those assets to bear in the development of policy and of practice globally. In our own work we have found that a particular sociocultural research methodology may be especially powerful in highlighting the key processes that are at work in the field of teacher education (Burn and Menter, 2021).

If teacher education is to make a continuing contribution to global stability, peace and democracy around the world, then there is a great need, indeed it is essential that we see a growth of independent critical research, including much that is comparative in nature. In this rapidly changing world with the overbearing threats from climate change, the rise of new forms of populism and nationalism, continuing conflicts and strife, and associated migrations of populations, we will need policy and practice in teacher education to be informed by high quality research as well as by the professional judgment of experienced practitioners in the field. This will be as important in western, “advanced” nations, where much provision will continue to be at a large scale (Brooks, 2021), as it will be in recently democratized nations and those going through major economic, social and political change, as in the case of the BRICS nations (Brazil, Russia, India, China, South Africa), and many others. (An International Research Network of the World Educational Research Association was established in 2021 to coordinate research on teacher education in this group of nations.)

The overall impact of the neoliberal processes of marketization and performativity have undoubtedly caused considerable damage to the best features of teacher education systems in many parts of the world. It will be through a combination of professional agency and the achievement of continuing high-quality research that we may hope to see significant improvements in what is being offered and what is being realized in these systems. The global pandemic of 2020/21 has revealed not only the vulnerability of teacher education but also the great imagination and potential power of those teachers and teacher educators working within the field. Our resources of hope for the future lie with these colleagues and in the comparative research that must continue to be undertaken.

References

- Alexander, R., 2000. *Culture and Pedagogy*. Blackwell, Oxford.
- Appadurai, A., 1996. *Modernity at Large: Cultural Dimensions of Globalization*. University of Minnesota Press, Minneapolis.
- Apple, M., Beane, J. (Eds.), 1999. *Democratic Schools*. Open University Press, Buckingham.
- Barber, M., Mourshed, M., 2007. *The McKinsey Report: How the World's Best Performing Schools Come Out on Top*. McKinsey & Co, London.
- Berliner, D., 2020. How “successful” charter schools cull and skim students they don't want. D. Ravitch's Blog. <https://dianeravitch.net/2020/04/13/david-berliner-how-successful-charter-schools-cull-and-skim-students-they-dont-want/>.
- Bernstein, B., 1970. Education cannot compensate for society. *New Soc.* 38, 344–347.
- Bowe, R., Ball, S., Gold, A., 1992. *Reforming Education and Changing Schools*. Routledge, London.
- Breslin, T., 2021. *Lessons From Lockdown*. Routledge, London.
- Brooks, C., 2021. *Initial Teacher Education at Scale*. Routledge, London.
- Brown, T., 2018. *Teacher Education in England*. Routledge, London.
- Bryk, A., Schneider, B., 2004. *Trust in Schools: A Core Resource for Improvement*. Russell Sage Foundation, New York.
- Burn, K., Menter, I., 2021. Making sense of teacher education in a globalizing world: the distinctive contribution of a socio-cultural approach. *Comp. Educ. Rev.* 65, 4.
- Childs, A., Menter, I., 2013. Teacher education in the 21st Century in England: a case study in neo-liberal policy. *Revista Espanola de Educacion Comparada* 22, 93–116. ISSN 1137-8654.
- Clarke, L., Magennis, G., 2016. Teacher education policy in Northern Ireland: impediments, initiatives and influences. In: *Teacher Education Group, the (2016) Teacher Education in Times of Change*. Policy Press, Bristol.
- Cochran-Smith, M., et al., 2018. *Reclaiming Accountability in Teacher Education*. Teachers College Press.
- Collins, D., 1977. *Paulo Freire: His Life, Works & Thought*. Paulist Press, New York.
- Crossley, M., Watson, K., 2013. *Comparative and International Research in Education*. Routledge, London.
- Darling-Hammond, L., 2017. Teacher education around the world: what can we learn from international practice? *Eur. J. Teach. Educ.* 40 (3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>.
- Department for Education, 2010. *The Importance of Teaching*. TSO, London. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf.

- Department for Education, 2021. Initial Teacher Training (ITT) Market Review Report. TSO, London. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/999621/ITT_market_review_report.pdf.
- Dewey, J., 1916. Democracy and Education. <https://www.gutenberg.org/files/852/852-h/852-h.htm>.
- Eatwell, R., Goodwin, M., 2018. National Populism: The Revolt Against Liberal Democracy. Pelican Books, London.
- Fiske, A.P., 2004. Human sociality. In: Kling, A. (Ed.), *Encyclopaedia of Science, Technology, and Ethics* (Fall 2021). <https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/market-theory>.
- Special Issue: the COVID-19 pandemic and its effects on teacher education. In: Flores, M., Swennen, A. (Eds.), *Eur. J. Teach. Educ.* 43, 4.
- Furlong, J., 2013. Education—An Anatomy of the Discipline. Rescuing the University Project? Routledge.
- Gray, P., 2016. Mother nature's pedagogy: how children educate themselves. In: Lees, H., Noddings, N. (Eds.), *The Palgrave International Handbook of Alternative Education*. Palgrave, London, pp. 49–62.
- Green, A., 1997. Education, Globalization and the Nation State. MacMillan Press, London.
- Helgetun, J., Menter, I., 2020. From an age of measurement to an evidence era? Policy-making in teacher education. *Engl. J. Educ. Policy* 1 (18). <https://doi.org/10.1080/02680939.2020.1748722>.
- Helgetun, J., 2022, in press. The importance of context: teacher education policy in England and France compared. In: Menter, I. (Ed.), *The Palgrave International Handbook of Teacher Education Research*. Palgrave=Macmillan, London.
- Heyward, J., 2021. Substitute teachers never got much respect but now they are on demand. *N. Y. Times*. <https://www.nytimes.com/2021/11/11/us/substitute-teachers-demand.html?referringSource=articleShare>.
- Illich, I., 1971. *Deschooling Society*. Calder & Boyars, London.
- Khmara, D., 2020. One-quarter of Arizona teaching positions are vacant, report finds. *Ariz. Dly. Star*. https://tucson.com/news/local/one-quarter-of-arizona-teaching-positions-are-vacant-report-finds/article_edf27692-f940-5b56-99d1-3445be69e554.html.
- King, E., 1967. *Other Schools and Ours*. Holt, Rinehart & Winston, London.
- Kling, A., 2021. Market theory. In: *Encyclopaedia of Science, Technology, and Ethics*. <https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/market-theory>.
- Knight, S., 2016. Forest Schools: a model for learning holistically and outdoors. In: Lees, H., Noddings, N. (Eds.), *The Palgrave International Handbook of Alternative Education*. Palgrave, London, pp. 289–304.
- Kum, H., Menter, I., Smyth, G., 2010. Changing the face of the Scottish teaching profession—the experiences of refugee teachers. *Ir. Educ. Stud.* 29 (3), 321–338.
- LeTendre, G., Ikoma, S., 2019. How teachers learn to teach: the impact of teacher reforms in Japan. In: Tatto, M.T., Menter, I. (Eds.), *Knowledge, Policy and Practice in Learning to Teach: A Cross-National Study*. Bloomsbury Academic, London, pp. 153–172.
- Marginson, S., Rhoades, G., 2002. Beyond national states, markets and systems of higher education: a glonacal agency heuristic. *High Educ.* 43 (3), 281–304.
- Matsko, K.K., Ronfeldt, M., Nolan, G.N., 2021. How different are they? Comparing teacher preparation offered by traditional, alternative, and residency pathways. *J. Teach. Educ.* <https://doi.org/10.1177/00224871211015976>.
- Menter, I. (Ed.), 2022. *Teacher Education in Russia—Past, Present and Future*. Routledge, London.
- Menter, I., 2022a. *Raymond Williams and Education: History, Culture, Democracy*. Bloomsbury, London.
- Menter, I., Valeeva, R., Kalimullin, A., 2017. A tale of two countries—forty years on: politics and teacher education in Russia and England. *Eur. J. Teach. Educ.* 40 (5), 616–629.
- Menter, 2016. Teacher education: generator of change or a mechanism for conformity? In: Lees, H., Noddings, N. (Eds.), *The Palgrave International Handbook of Alternative Education*. Palgrave, London, pp. 257–272.
- Menter, I., 2019. The interaction of global and national influences. In: Tatto, M.T., Menter, I. (Eds.), *Knowledge, Policy and Practice in Teacher Education—A Cross-National Study*. Bloomsbury, London, pp. 268–279.
- Teacher education in the COVID-19 pandemic: a global snapshot [Special Issue]. In: O'Meara, J., Hordatt Gentles, C. (Eds.), *J. Educ. Teach.* 46 (4).
- Pereira, A., 2020. A Change in Brazil's National Populist Government. <https://blog.oup.com/2020/11/a-change-in-brazils-national-populist-government/>.
- Phillips, D., Schweisfurth, M., 2008. *Comparative and International Education*. Continuum, London.
- Ravitch, D., 2016. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, New York.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Sahlberg, P., 2012. How GERM is Infecting Schools Around the World? <https://pasisahlberg.com/text-test/>.
- Sass, T.R., 2015. Licensure and worker quality: a comparison of alternative routes to teaching. *J. Law Econ.* 58 (1). <https://doi.org/10.1086/682904>.
- Sorensen, N. (Ed.), 2019. *Diversity in Teacher Education*. UCL/IOE Press, London.
- Spring, J., 2015. *Globalization of Education*. Routledge, New York.
- Stoddart, T., Floden, R., 1995. Traditional and Alternative Routes to Teacher Certification: Issues, Assumptions, and Misconceptions. Michigan State University, National Center for Research on Teacher Learning. <https://files.eric.ed.gov/fulltext/ED383697.pdf>.
- Tatto, M.T., Clark, C.M., 2019. Institutional transformations, knowledge and research traditions in the USA. In: Tatto, M.T., Menter, I. (Eds.), *Knowledge, Policy and Practice in Learning to Teach: A Cross-National Study*. Bloomsbury Academic, London, pp. 233–256. <https://doi.org/10.5040/9781350068711>.
- Tatto, M.T., Menter, I. (Eds.), 2019. *Knowledge, Policy and Practice in Learning to Teach: A Cross-National Study*. Bloomsbury Academic, London. <https://doi.org/10.5040/9781350068711>.
- Tatto, M.T., Menter, I., (forthcoming). The turn to the market in teacher education: institutional and pedagogical consequences of market reforms and increased regulation in teacher education. In: Arnone, R., Torres, C.A., Misiaszek, L. (Eds.), *Comparative Education: The Dialectic of the Global and the Local*, pp. ###–###, fifth ed. Lanham, MD: Rowman & Littlefield.
- Tatto, M.T., Burn, K., Menter, I., Mutton, T., Thompson, I., 2018a. *Learning to Teach in England and the United States: The Evolution of Policy and Practice*. Routledge, Abingdon, England. ISBN: ISBN 978-1-138-93374-3.
- Tatto, M.T., Rodriguez, M., Smith, W., Reckase, M., Bankov, K. (Eds.), 2018b. *Exploring the Mathematics Education of Teachers Using TEDS-M Data*. Springer, Dordrecht, Netherlands. <https://doi.org/10.1007/978-3-319-92144-0>.
- Tatto, M.T., Rodriguez, M., Reckase, M., Smith, W., Bankov, K., Pippin, J., 2020. *The First Five Years of Teaching Mathematics (FIRSTMATH): Concepts, Methods and Strategies for Comparative International Research*. Springer, Dordrecht, Netherlands. <https://doi.org/10.1007/978-3-030-44047-3>.
- Tatto, M.T. (Ed.), 2007. *Reforming Teaching Globally*. Symposium Books, Oxford, UK (Reprinted in 2009 Information Age Publishers).
- Tatto, M.T., 2021. Comparative research on teachers and teacher education: global perspectives to inform UNESCO'S SDG 4 Agenda. *Oxf. Rev. Educ.* 47 (1), 25–44. <https://doi.org/10.1080/03054985.2020.1842183>.
- Teacher Education Group, 2016. *Teacher Education in Times of Change*. Policy Press, Bristol.
- Tiplady, L., Menter, H., 2021. The Breeze Project: supporting children and young people through Forest School. In: Jayman, M., Oihl, M., Jewett, L. (Eds.), *Supporting Digital Natives*. Policy Press, Bristol, pp. 78–96.
- Tucker, M., 2012. Why the “market theory” of education reform doesn't work. *Wash. Post*. <https://www.washingtonpost.com/news/answer-sheet/wp/2012/10/12/why-the-market-theory-of-education-reform-doesnt-work/>.
- U.S. Department of Education, Office of Innovation and Improvement, 2004. *Innovations in Education: Alternative Routes to Teacher Certification*. Washington, D.C. <https://www2.ed.gov/admins/tchrqual/recruit/altroutes/report.pdf>.

- Waslander, S., Pater, C., van der Weide, M., 2010. Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education. OECD Education Working Papers, No. 52. OECD Publishing. <https://doi.org/10.1787/5km4pskmkr27-en>. <https://www.oecd.org/education/ceri/47229500.pdf>.
- Whitty, G., Furlong, J., 2017. Knowledge and the Study of Education: An International Exploration. Symposium Books, Didcot.
- Williams, R., 1961. The Long Revolution. Penguin, Harmondsworth.
- Williams, R., 1983. Towards 2000. Penguin, Harmondsworth.
- Williams, R., 1989. Resources of Hope. Verso, London.

FOURTH EDITION

International Encyclopedia of Education

Learning, Cognition
and Human Development

VOLUME
6

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Ian Wilkinson
Judy Parr



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 6

Learning, Cognition, and Human Development

VOLUME 6 EDITORS

Ian Wilkinson and Judy Parr



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

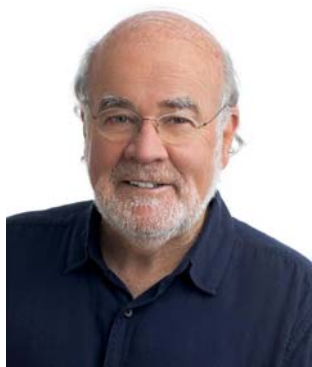
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



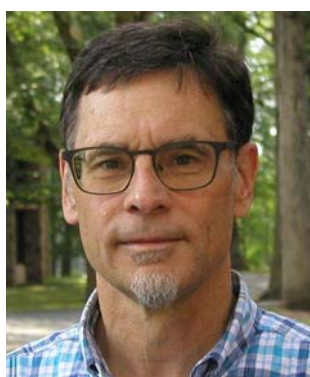
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 6: Learning, Cognition, and Human Development

Ian A G Wilkinson and Judy M Parr

© 2023 Elsevier Ltd. All rights reserved.

Introduction

Confronted with the task of assembling some 80 chapters for the volume on *Learning, Cognition, and Human Development*, we sought guidance (and solace) in the work of our predecessors, as well as others. We scoured the pages of the corresponding volume, titled *Learning and Cognition*, of the previous edition of the *IEE*. We pored over the contents of contemporary handbooks with related titles: handbooks in educational psychology, cognition, motivation and learning, cognitive processes, learning and instruction, and the learning sciences. We dutifully perused topics in related encyclopedias in education, educational psychology, and human development. What we were searching for was some form of conceptual framework to guide our selection and organization of topics. We did not find it or, if it was there, we did not see it.

A breakthrough, of sorts, came when we consulted the recently published report from the [National Academies of Sciences, Engineering, and Medicine \(2018\)](#), *How People Learn II: Learners, Contexts, and Cultures*. This report surveyed research on fundamental processes in the individual that impact learning (as did the earlier *How People Learn: Mind, Brain, Experience, and School*, [National Research Council, 2000](#)), but it also examined the influence of culture in shaping how people learn and the nature of learning over time. Internal processes involving knowledge, mental models, reasoning, and the like are important for learning, but they are shaped by contexts and cultures, and the way they function changes throughout a person's life span.

At the outset, then, we took learning as the primary conceptual construct to frame this volume (it is, after all, at the forefront in the title). Following the framework implicit in *How People Learn II*, we organized the content using the tripartite division of people, space, and time. From this, we derived our three core sections: *The Learner: Fundamental Processes*, *Learning in Contexts and Cultures*, and *Learning Over Time*. We bookended these core sections with a front section titled *Perspectives on Learning: Theory and Methods* and a back section titled *Optimizing Learning*. Our intention with the latter was to survey current interventions and pedagogical approaches that research suggests have the potential to optimize learning in formal (i.e., school) and informal settings, including digital environments.

Having arrived at a conceptual framework to organize this volume on *Learning, Cognition, and Human Development*, we spent considerable time identifying, discussing, and debating potential topics from both established and emerging areas of research. We made a deliberate decision to reflect contemporary movements in the field and to ensure as broad as possible a representation of topics and scholars. Within each of our five sections, we endeavored to secure, as the first chapter in the line-up, an overview that would serve as an introduction to the area. At the time of planning, we were not sure whether and how the overview chapters would serve our intentions and, indeed, whether it would even be possible to find someone to author these chapters.

The scholars we approached for the overviews exceeded our expectations. Their chapters serve to identify recent debates and developments within which the other chapters in the section are located. In this introduction, we highlight the contribution of the overview chapters in each section and draw connections between them and the respective constituent chapters. We conclude by reflecting on the content in relation to that of the previous edition and by considering ongoing research in the area, looking back and looking forward.

Perspective on Learning: Theory and Methods

Overview chapters at times challenge the reader, and the one leading the perspectives on learning section, by Danish, challenges us to think critically about the key theories of learning studied in the fields of learning, cognition, and human development, and particularly about the importance of recognizing the contexts within which these theories have been developed and are used. Emphasizing that theories of learning are situated in time and place, Danish initiates a conversation about how theories describe learning in the world that we live in now and in the wide range of learning contexts that reflect where learners are, including both formal and informal contexts. Acknowledging that this makes the conversation about what learning is more complicated, the author also notes that it increases the chance that we can engage in conversations about how these theories describe, and consequently shape, learning in the world and not simply function to prompt conversations between scholars. The title of the chapter encourages readers to ask of each theory "for whom?" and "to what end?" and discusses what our theories of learning do not yet address, in the hope of pushing us all to think more deeply about what theories of learning can and should do in the future.

Authors of the chapters in this section have a view about "to what end" the perspective they take is intended to function with respect to furthering the field of learning. The chapter on social cognitive theory emphasizes the theory's reflection of reciprocal interactions between personal, behavioral, and social/environmental variables, while that concerned with theories of embodied cognition discusses how concepts are essentially constituted by reinstatements of experiences related to perception and introspection as well as to bodily actions. The chapter offering a sociocultural perspective is clear about the role of theory in conceptualizing learning and development in the complex world that we live in. However, it is the author of the chapter on learning from a behaviorist perspective who asserts, perhaps controversially, that a science of practical use will be behavioristic in perspective. This section also reaches beyond more traditional theoretical perspectives to present a chapter considering the contribution of the interdisciplinary learning sciences perspective not only to advancing theoretical conceptions of learning but also to the act of

supporting learning, including new methodological approaches. Another interdisciplinary approach, as a response to complexity, is the chapter on educational neuroscience which illustrates how combining perspectives from psychology, neuroscience, and genetics helps to unravel mechanisms underlying learning. Further to the notion of the overview chapter that theories and perspectives are situated in time and place, there are two chapters relating to contemporary phenomena: one examines learning analytics, a recent interdisciplinary field encompassing learning (e.g., learning sciences), analytics (e.g., data science), and human-centered design perspectives, and the other discusses the perspectives of artificial intelligence and machine learning as they relate to education.

The Learner: Fundamental Processes

Reading Wiley and Guerrero's overview of fundamental processes of learning, one might be forgiven for thinking of it as an overview of the entire volume. Its coverage is both broad and deep. Consistent with the title, they adopt a largely cognitive perspective on many aspects of mental functioning: memory, attention, comprehension, reasoning, knowledge, and the like. They situate memory at the center of all cognitive processes and as fundamental for learning, but also give due consideration to comprehension and other higher-order cognitive processes. Comprehension, of course, fundamentally depends on memory but active engagement with the information is essential if the learner is to construct a rich mental representation. Under the heading of memory, comprehension, and other processes, they systematically consider conditions that may improve these processes and here they employ both cognitive and sociocultural theories of learning, touching on the importance of interaction with others be they peers or more knowledgeable others. They also consider skill acquisition over time, in reading and math, and in first- and second-language acquisition. The now taken-as-given mantra from constructivist theory of the learner as an active agent in their own learning is alive and well in Wiley and Guerrero's overview.

Other chapters in this section pick up on the cognitive perspective with due consideration given to attentional processes, processing limitations, expertise, knowledge, problem-solving, reasoning, and, of course, memory. But where these other chapters extend beyond the overview provided by Wiley and Guerrero is in consideration of more affective and social processes and constructs. Here we see discussion of creativity, curiosity and interest, emotion, grit and resilience, as well as self-concept and self-efficacy. The learner might indeed be an active agent in their own learning, and deploy various cognitive resources and processes as demanded by the task, but key affective and social influences support these cognitive resources and processes and, in turn, are fueled by these resources and processes. Learning is indeed much more than a cognitive endeavor.

Learning in Contexts and Cultures

The overview chapter that leads the contexts and cultures section, by Cope, Kalantzis and Tzirdes, provides one set of lenses to consider the conceptual underpinnings of this section. The first lens is the general distinction between learning that happens in the informal context of the lifeworld, contrasted with the formal or institutionalized learning of programs, curricula and syllabi—called “education.” The second lens is the social functions of both informal and formal learning, focusing on the role of education for working life, personal life, and public life. Then, cross-cutting these contexts, the authors discuss the increasing significance today of social diversity and digital tools for meaning-making. The overview finally moves from education to the broader society within which it functions to address the question, how and in what ways can learning be a force for educational justice and social change?

The other chapters in the section reflect, and refocus, some of these lenses. The two chapters that follow the overview broadly relate to the question the authors of the overview pose regarding educational justice and social change. The chapter on cross-cultural perspectives on learning offers an expansive understanding of culture not as static and homogenous, but rather as participation in cultural practices that include homogeneity and heterogeneity within and across time and spaces. Importantly, it provides research-based examples of educational efforts to recruit culturally rooted repertoires of practice to support learning. The chapter on Indigenous student identity development and school success discusses how Indigenous student academic identities are developed, experienced, and expressed, explaining how schools can create environments that enable Indigenous students to succeed alongside other students, and introducing a model that provides a framework for understanding the important conditions that Indigenous students need to succeed and thrive at school. In consideration of the importance of social context, a chapter considers social relationships at school, and another presents a social and cultural constructivist perspective on learning “problems.”

Then there is a grouping, beginning with a chapter discussing learning in the disciplines, clearly connected to the first lens the overview authors offer, namely, core curriculum learning (in history and social studies, math, reading and writing, the sciences) in formal settings. This group also contains chapters concerning phenomenon like learning an additional language and learning from multiple texts, both significant in our globalized, information-rich context. The second lens, informal learning, is represented through chapters on learning out of school and learning through work. Encompassing the formal, the informal, and the cultural contexts is a chapter which explores the intersection between spaces, materials, culture, and learning, while problematizing the need to look at past and present when designing for learning. It discusses materiality within the context of a Māori marae and within contemporary spaces for learning. Digital tools and contexts, the other phenomenon identified as “cross-cutting” contexts by the overview chapter authors, are well represented in chapters on networked learning, online learning, and social media.

Learning Over Time

In the overview chapter in our Time section, titled *Learning Across the Lifespan*, Alexander and Fusenig argue that much of learning that transpires in the domains and disciplines that form the backbone of formal educational systems is inextricably intertwined with cognition and development. The authors identify and discuss relevant issues pertaining to learning across the lifespan: that there are general principles about the nature of learning, common to diverse theories; that there are contrasting views of the relation between learning and development; that the notions of lifespan and lifelong learning converge on certain premises and practices while diverging on others; and that it is feasible to construct a model that brings together learning and development across the lifespan.

The other chapters in this section mostly concern age groupings commonly identified in educational literature. Integrative approaches characterize the descriptions, for example, efforts to clarify how distinct mechanisms work together to provide a comprehensive, parsimonious, description of infants' learning. In early childhood education, sociocultural theories are seen as significant, drawing attention to the cultural nature of learning and development, the diverse social and cultural contexts for learning and participation, the meaning of play, the cultural-historical significance of tools and materials, and children's agency and capabilities. In accord with the emphases on learning across the lifespan, chapters focus on examining the phenomena of lifelong learning as well as honing in on what characterizes learning, cognition, and development in certain periods in the lifespan. On the topic of neurocognitive aging, we have an interesting chapter on the relationship between educational level and cognitive functioning in aging. Chapters exploring learning progressions and the multiple and multidimensional nature of educational and life transitions conclude this section.

Optimizing Learning

In planning this section, we envisioned descriptions of research-based interventions and pedagogies that have been shown to be effective in promoting student learning, at least in relatively controlled situations—"what works" if you will. In the overview to this section, McNaughton problematizes such a notion. As is evident in the chapters in this section, we have good descriptions of various interventions and pedagogies for promoting learning and, in many cases, good theoretical explanations for their effects. However, scaling up such approaches—giving them wider application beyond the conditions under which they were studied—often presents problems. As McNaughton points out, the assumption in many instances is that, with good descriptions and "robust explanations," we have all that is needed to optimize learning. McNaughton rejects the idea of "what works" in a universal sense and acknowledges the contextual conditions and variations that lead to much variability in learning outcomes. Invoking implementation science (Bryk et al., 2015), he points out that, rather than asking "What works?," a better question is "What works for whom, under what contexts and conditions, and in a way that is sustainable/at scale?" Adopting a systems perspective, McNaughton interrogates the reasons for, sometimes extreme, variability in outcomes and offers suggestions for making progress in optimizing learning. Again from a systems perspective, he posits that differential targeting of interventions for children at risk and from marginalized communities, an emphasis on 21st century skills (inter- and intrapersonal skills, critical thinking, and critical literacy), and preserving identity, language, and culture may serve as potential "solutions" to the optimization dilemma.

One might expect some tension between the challenges posed by McNaughton and the content of the chapters that follow. True to form, they provide rich descriptions of interventions and pedagogies known to enhance learning along with theoretical explanations for their effects. Nonetheless, a close reading reveals that many attempt to address the (boundary) conditions for successful application of a particular pedagogical approach (e.g., inquiry-based learning). We also see attention to how instruction should be differentiated to (optimally) meet the needs of individual students. Issues of equity and diversity come to the fore in many of the chapters, as do the role of digital technologies to promote cooperation and collaboration for learning and development of higher-order thinking skills. In concert with developments in promoting mindfulness and wellbeing, it is here we see efforts to focus on so-called 21st century skills, both cognitive and noncognitive, and attention to learning in both formal and informal contexts. The challenges of scaling up and giving interventions more applicability is a theme in many of the chapters.

Reflections on the Content: Looking Back, Looking Forward

The context in which this Encyclopaedia has been produced is unparalleled. As a result of disruptions accompanying the COVID-19 pandemic, some authors we approached for this volume were unable to commit and some, in the end, were unable to produce their agreed chapters. So there are inevitable gaps. We would like to have included, for example, chapters on learning from a cognitive constructivist perspective, learning difficulties, motivation, self-regulation, 21st century skills, learning and development in middle childhood, and scaffolding and the gradual release of responsibility. Nonetheless, the vast majority of scholars we approached graciously accepted our invitation. We appreciate how fortunate we were to have had the opportunity to read, and learn from, the perspectives of the expert scholars who have written chapters for this volume.

If the previous edition of the *IEE* (2010) can be used as a benchmark, the fields of learning, cognition, and human development have greatly expanded. We were fortunate to secure many more chapters than in the earlier edition. In this volume, in the perspectives on learning section, a wider range of theoretical views on learning is presented. New topics to this edition include the learning sciences, embodied cognition, social and emotional learning, learning analytics, and machine learning. The section on the learner also shows expansion. Here new chapters cover the topics of knowledge, epistemic cognition, cognitive load, executive function, curiosity and interest, self-efficacy, grit and resilience, academic self-concept, and emotion. The factors and conditions considered

in the contexts and cultures section reflect the emphasis on culture and learning as well as the influence of the digital age with chapters on online learning, social media and learning, and learning spaces. Indeed, the influence and reach of ideas from sociocultural theories, in what has traditionally been a field dominated by cognitive perspectives, is apparent throughout this volume, as is the influence of technology. The section on learning over time and in different age groupings is entirely new to this edition, as is the final section on optimizing learning (of course, some topics within these sections received treatment in the previous edition).

As editors, a foremost aim for us was to secure expertise from a broad scholarly community—authors who would likely present a diversity of perspectives to reflect the globalization of our research world. Again in comparison to the previous volume, the number of countries represented by the authors in this volume has grown from 11 to 17 (though the United States still accounts for the largest percentage, one comparable to the last volume). We note that many more scholars from the European continent are now included.

We began this introduction with reference to *How People Learn II* as a framing device to guide the design and assembly of this volume. In that report, the committee identified two broad areas for a future research agenda. One was seeking a better understanding of variability in learning “by connecting research on internal mechanisms of learning with the shaping forces of contextual variation, including culture, social context, instruction, and time of life” (p. 227). The other was advancing understanding of the affordances (and constraints) of digital technologies for learning and ensuring design and developments in this area are well informed by the science of learning. We would like to think, and suggest, that chapters in this volume demonstrate movement in this agenda and that the work described herein goes some way towards realizing these goals.

References

- Bryk, A.S., Gomez, L.M., Grunow, A., LeMahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press, Cambridge, MA.
- National Academies of Sciences, Engineering, and Medicine, 2018. *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press. <https://doi.org/10.17226/24783>.
- National Research Council, 2000. *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. The National Academies Press. <https://doi.org/10.17226/9853>.

CONTRIBUTORS TO VOLUME 6

David Adamson

Turnitin, Pittsburgh, PA, United States

Patricia A Alexander

Department of Human Development and Quantitative Methodology, University of Maryland, College Park, College Park, MD, United States

Rosie Allen

Department of Psychology, School of Education and Psychology, University of Bolton, Bolton, United Kingdom

Yvonne Allsop

The Ohio State University, Department of Educational Studies, Columbus, OH, United States

Alicia C Alonzo

Department of Teacher Education, Michigan State University, East Lansing, MI, United States

Eric M Anderman

The Ohio State University, Department of Educational Studies, Columbus, OH, United States

Petra A Arndt

Transfer Center for Neuroscience and Education (ZNL), Ulm University, Ulm, Germany

K Aronsson

Stockholm University, Stockholm, Sweden

Keith C Barton

Department of Curriculum and Instruction, Indiana University, Bloomington, IN, United States

Geetanjali Basarkod

Institute for Positive Psychology and Education, Australian Catholic University, Sydney, NSW, Australia

Lisa M Baumgartner

Department of Counseling, Leadership, Adult Education and School Psychology, Texas State University, San Marcos, TX, United States

Stephen Billett

Griffith University, Queensland, QLD, Australia

Mimi Bong

Department of Education and the Brain & Motivation Research Institute (bMRI), Korea University, Seoul, South Korea

Nicole S Boucher

Arizona State University, Tempe, AZ, United States

Ivar Bråten

Department of Education, University of Oslo, Oslo, Norway

Tim Brennen

Department of Psychology, University of Oslo, Oslo, Norway

Victoria M Bryan

Department of Psychology, University of New Hampshire, Durham, NH, United States

Eric Burnheimer

Department of Education, School of Education and Information Studies, University of California, Los Angeles, CA, United States

Destany Calma-Birling

Institute of Child Development, University of Minnesota, Minneapolis, MN, United States

David Carless

Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong

Shana K Carpenter

Department of Psychology, Iowa State University, Ames, IA, United States

Jerome Carson

Department of Psychology, School of Education and Psychology, University of Bolton, Bolton, United Kingdom

Lucila Carvalho

Massey University, Auckland, New Zealand

Chih Ching Chang

Faculty of Education, University of Cambridge, Cambridge, United Kingdom

Amy L Chapman
Columbia University, Chester, NJ, United States

Micheline TH Chi
Arizona State University, Tempe, AZ, United States

Kimiko Ching
The Ohio State University, Department of Educational Studies, Columbus, OH, United States

Bill Cope
University of Illinois Urbana-Champaign, Champaign, IL, United States

G Curt Fulwider
Florida State University, Tallahassee, FL, United States

Joshua A Danish
Learning Sciences, Indiana University School of Education, Bloomington, IN, United States

Ton de Jong
University of Twente, Enschede, the Netherlands

Maria K DiBenedetto
The University of North Carolina at Greensboro, Greensboro, NC, United States

Curt Dudley-Marling
Boston College, Chestnut Hill, MA, United States

Darren L Dunning
Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom; and University of Hertfordshire, Hatfield, United Kingdom

Martin East
School of Cultures, Languages and Linguistics, Faculty of Arts, The University of Auckland, Auckland, New Zealand

Julia Eberle
Faculty of Philosophy and Education, Ruhr-University Bochum, Bochum, Germany

Karen Evans
UCL Institute of Education, Social Research Institute, London, United Kingdom

James Fiacco
Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, United States

Frank Fischer
Ludwig-Maximilians-Universität München, Department of Psychology, Munich Center of the Learning Sciences, Munich, Germany

Anne C Frenzel
Department of Psychology, University of Munich, Munich, Germany

Jannah Fusenig
Department of Human Development and Quantitative Methodology, University of Maryland, College Park, College Park, MD, United States

Sarah M Galvin
Michigan State University, East Lansing, MI, United States

Thomas Goetz
Department of Developmental and Educational Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria

Susan R Goldman
Learning Sciences Research Institute, University of Illinois Chicago, Chicago, IL, United States

Peter Goodyear
Centre for Research on Learning and Innovation, Sydney School of Education and Social Work, The University of Sydney, Sydney, NSW, Australia

Arthur C Graesser
University of Memphis, Memphis, TN, United States

Jenna W Gravel
CAST, Wakefield, MA, United States

Christine M Greenhow
Michigan State University, East Lansing, MI, United States

Tricia A Guerrero
Department of Psychology, University of Illinois at Chicago, Chicago, IL, United States

Jesse Ha
Arizona State University, Tempe, AZ, United States

Seung Yon Ha
The Ohio State University, Department of Educational Studies, Columbus, OH, United States

Päivi Häkkinen
University of Jyväskylä, Finnish Institute for Educational Research, Jyväskylä, Finland

Tina Hascher
University of Bern, Institute of Educational Science, Department of Research in School and Instruction, Bern, Switzerland

Mikael Heimann
Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Annette ME Henderson
School of Psychology, The University of Auckland, Auckland, New Zealand

Suzanne E Hidi

Department of Applied Psychology and Human Development, OISE, University of Toronto, Toronto, ON, Canada

Carol Hoare

Department of Counseling and Human Development, The George Washington University, Washington, DC, United States

Erika Hoff

Florida Atlantic University, Davie, FL, United States

Carmen Hofmann

Section for Cognitive Electrophysiology, Department of Psychiatry, Ulm University, Ulm, Germany; and Transfer Center for Neuroscience and Education (ZNL), Ulm University, Ulm, Germany

Tasos Hovardas

University of Cyprus, Nicosia, Cyprus

Glynda A Hull

University of California, Berkeley, School of Education, Berkeley, CA, United States

Kalypso Iordanou

University of Central Lancashire, Larnaka, Cyprus

Sanna Järvelä

University of Oulu, Department of Educational Sciences and Teacher Education, Oulu, Finland

Shiyan Jiang

School of Education, North Carolina State University, Raleigh, NC, United States

Divya Jindal-Snape

Professor of Education, Inclusion and Life Transitions, School of Education and Social Work, University of Dundee, Dundee, United Kingdom

Ithel Jones

Florida State University, Tallahassee, FL, United States

Mary Kalantzis

University of Illinois Urbana-Champaign, Champaign, IL, United States

Chathurika Kannangara

Department of Psychology, School of Education and Psychology, University of Bolton, Bolton, United Kingdom

Natasha Kersh

UCL Institute of Education, Social Research Institute, London, United Kingdom; and UCL Institute of Education, Department of Education, Practice and Society, London, United Kingdom

Anna Keune

Technische Universität München, TUM School of Social Sciences and Technology, München, Germany

Markus Kiefer

Section for Cognitive Electrophysiology, Department of Psychiatry, Ulm University, Ulm, Germany

Young-Suk Grace Kim

University of California Irvine, Irvine, CA, United States

Vickie E Lake

University of Oklahoma, Norman, OK, United States

Ard W Lazonder

Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands

Carol D Lee

School of Education and Social Policy, Northwestern University, Evanston, IL, United States

Hyun Ji Lee

Department of Education and the Brain & Motivation Research Institute (bMRI), Korea University, Seoul, South Korea

Haiying Li

Iowa State University, Ames, IA, United States

Anastasiya A Lipnevich

Queens College and the Graduate Center, The City University of New York, New York, NY, United States

Qian Liu

Faculty of Education, University of Cambridge, Cambridge, United Kingdom

Zhichun Liu

University of Massachusetts, Dartmouth, MA, United States

Yun Long

Faculty of Education, University of Cambridge, Cambridge, United Kingdom

Martin Lövdén

Department of Psychology, University of Gothenburg, Gothenburg, Sweden

Sofie MM Loyens

University College Roosevelt, Utrecht University, Utrecht, the Netherlands; and Department of Education, Psychology, and Child Studies (DPECS), Erasmus University Rotterdam, Rotterdam, the Netherlands

Jasmine Y Ma

New York University, New York, NY, United States

Armando Machado

University of Aveiro, Aveiro, Portugal

Abby Madden-Smith

Faculty of Education, University of Auckland, Auckland, New Zealand

Svein Magnussen

Department of Psychology, University of Oslo, Oslo, Norway

Lina Markauskaite

The University of Sydney, Sydney, NSW, Australia

Henry Markovits

Department of Psychology, Université du Québec à Montréal, Montréal, QC, Canada

Herbert W Marsh

Institute for Positive Psychology and Education, Australian Catholic University, Sydney, NSW, Australia

J Massonnié

School of Education and Sociology, Faculty of Humanities & Social Sciences, University of Portsmouth, Portsmouth, United Kingdom

John D Mayer

Department of Psychology, University of New Hampshire, Durham, NH, United States

Richard E Mayer

Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, United States

Kathryn S McCarthy

Georgia State University, College of Education and Human Development, Atlanta, GA, United States

Matthew T McCrudden

Department of Educational Psychology, Counseling, and Special Education, Pennsylvania State University, University Park, PA, United States

Danielle S McNamara

Arizona State University, Tempe, AZ, United States

Stuart McNaughton

The University of Auckland, Auckland, New Zealand

Alecia Moser

School of Psychology, The University of Auckland, Auckland, New Zealand

Piia Näykki

University of Jyväskylä, Faculty of Education and Psychology, Jyväskylä, Finland

Lars Nyberg

Department of Radiation Sciences, Radiology, Umeå University, Umeå, Sweden; Department of Integrative Medical Biology, Umeå University, Umeå, Sweden; Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden; and Center for Lifespan Changes in Brain and Cognition, University of Oslo, Oslo, Norway

Jonathan Osborne

Graduate School of Education, Stanford University, Stanford, CA, United States

Reinhard Pekrun

Department of Psychology, University of Essex, Colchester, United Kingdom; Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia; and Department of Psychology, University of Munich, Munich, Germany

Fanxiao Wani Qiu

Department of Educational Studies, Swarthmore College, Swarthmore, PA, United States

Seyedahmad Rahimi

Florida State University, Tallahassee, FL, United States

Kavita Rao

University of Hawai'i at Mānoa, College of Education, Honolulu, HI, United States

Chrysi Rapanta

Universidade Nova de Lisboa, Lisbon, Portugal

David Reilly

School of Applied Psychology, Griffith University, O'Connor, QLD, Australia

Peter Reimann

The University of Sydney, Sydney, NSW, Australia

Alexander Renkl

University of Freiburg, Freiburg, Germany

K Ann Renninger

Department of Educational Studies, Swarthmore College, Swarthmore, PA, United States

Remy MJ P Rikers

University College Roosevelt, Utrecht University, Utrecht, the Netherlands

Carolyn P Rosé

Language Technologies Institute, Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, United States

David H Rose
CAST, Wakefield, MA, United States

Ana Rubio-Jimenez
*Faculty of Education, University of Cambridge,
Cambridge, United Kingdom*

Mark A Runco
*Southern Oregon University, Ashland, OR, United
States*

Roger Säljö
*Department of Education, Communication and
Learning, University of Gothenburg, Gothenburg,
Sweden*

Ladislao Salmerón
ERI Lectura—University of Valencia, Valencia, Spain

William A Sandoval
*School of Education & Information Studies, University of
California, Los Angeles, CA, United States*

Alan H Schoenfeld
*School of Education, University of California, Berkeley,
CA, United States*

Kimberly A Schonert-Reichl
*University of Illinois at Chicago, Chicago, IL, United
States*

Katherine Schultz
*University of Colorado Boulder, School of Education,
Boulder, CO, United States*

Dale H Schunk
*The University of North Carolina at Greensboro,
Greensboro, NC, United States*

Andrei Semenov
*Institute of Child Development, University of Minnesota,
Minneapolis, MN, United States*

Peter Shea
*University at Albany, State University of New York,
Albany, NY, United States*

Shengpeng Shi
*Faculty of Education, University of Cambridge,
Cambridge, United Kingdom*

Valerie J Shute
Florida State University, Tallahassee, FL, United States

Daniel Sommerhoff
*Leibniz Institute for Science and Mathematics
Education, Department for Mathematics Education,
Kiel, Germany*

John Staddon
Duke University, Durham, NC, United States

K Bret Staudt Willet
Florida State University, Tallahassee, FL, United States

Lisa Stempfer
*Department of Developmental and Educational
Psychology, Faculty of Psychology, University of Vienna,
Vienna, Austria*

Kristina Stockinger
*Faculty of Philosophy and Social Sciences, University of
Augsburg, Augsburg, Germany*

Karin Strid
*Department of Psychology, University of Gothenburg,
Gothenburg, Sweden*

John Sweller
*School of Education, University of New South Wales,
Sydney, NSW, Australia*

Judith Ter Vrugte
University of Twente, Enschede, the Netherlands

MSC Thomas
*Center for Educational Neuroscience, Department of
Psychological Sciences, Birkbeck University of London,
London, United Kingdom*

T Nicole Tucker-Smith
Lessoncast, Baltimore, MD, United States

Michelle K Tulloch
Florida Atlantic University, Davie, FL, United States

Anastasia O Tzirides
*University of Illinois Urbana-Champaign, Champaign,
IL, United States*

Noreen M Webb
*Department of Education, School of Education and
Information Studies, University of California, Los
Angeles, CA, United States*

Melinda Webber
*Faculty of Education, University of Auckland, Auckland,
New Zealand*

Rupert Wegerif
*Faculty of Education, University of Cambridge,
Cambridge, United Kingdom*

Kathryn R Wentzel
*University of Maryland, College Park, MD, United
States*

Lisette Wijnia
*Open University of the Netherlands, Heerlen, the
Netherlands*

Jennifer Wiley
*Department of Psychology, University of Illinois at
Chicago, Chicago, IL, United States*

Christian Winterbottom
University of North Florida, Jacksonville, FL, United States

Elizabeth Ann Wood
School of Education, The University of Sheffield, Sheffield, United Kingdom

Zacharias C Zacharia
University of Cyprus, Nicosia, Cyprus

Philip David Zelazo
Institute of Child Development, University of Minnesota, Minneapolis, MN, United States

CONTENTS OF VOLUME 6

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 6: Learning, Cognition, and Human Development</i>	<i>xix</i>
<i>Contributors to Volume 6</i>	<i>xxiii</i>

PERSPECTIVES ON LEARNING: THEORY AND METHODS

What is learning, for whom, and to what end? An overview <i>Joshua A Danish and Jasmine Y Ma</i>	1
Learning from a behaviorist perspective <i>Armando Machado and John Staddon</i>	12
Learning from a social cognitive theory perspective <i>Dale H Schunk and Maria K DiBenedetto</i>	22
Learning from a sociocultural perspective <i>Roger Säljö</i>	36
Perspectives on learning from the learning sciences <i>Frank Fischer, Daniel Sommerhoff, and Anna Keune</i>	44
Perspectives on learning from neuroscience <i>J Massonnié and MSC Thomas</i>	57
Embodied cognition <i>Markus Kiefer, Carmen Hofmann, and Petra A Arndt</i>	67
Learning analytics <i>James Fiocco, Shiyan Jiang, David Adamson, and Carolyn P Rosé</i>	75
Machine learning <i>Valerie J Shute, G Curt Fulwider, Zhichun Liu, and Seyedahmad Rahimi</i>	83
Social and emotional learning <i>Kimberly A Schonert-Reichl</i>	92

THE LEARNER: FUNDAMENTAL PROCESSES

Fundamental processes of learning: an overview <i>Jennifer Wiley and Tricia A Guerrero</i>	107
Attention: a prerequisite for learning <i>Karin Strid and Mikael Heimann</i>	117
Cognitive load theory <i>John Sweller</i>	127
Creativity and learning <i>Mark A Runco</i>	135
Curiosity and interest: what should educators know and consider? <i>K Ann Renninger, Fanxiao Wani Qiu, and Suzanne E Hidi</i>	141
Emotions in education <i>Thomas Goetz, Anne C Frenzel, Kristina Stockinger, Anastasiya A Lipnevich, Lisa Stempfer, and Reinhard Pekrun</i>	149
Epistemic cognition <i>William A Sandoval</i>	162
Executive function and education <i>Destany Calma-Birling, Andrei Semenov, and Philip David Zelazo</i>	168
Teaching expertise: an activity system's lens <i>Peter Reimann and Lina Markauskaite</i>	179
From grit and resilience to academic tenacity <i>Chathurika Kannangara, Rosie Allen, and Jerome Carson</i>	190
Intelligences <i>John D Mayer and Victoria M Bryan</i>	198
Knowledge: a fundamental asset <i>Kathryn S McCarthy and Danielle S McNamara</i>	209
Memory <i>Tim Brennen and Svein Magnussen</i>	219
Problem solving <i>Richard E Mayer</i>	229
Reasoning and decision making <i>Henry Markovits</i>	235
Academic self-concept: a key predictor of achievement and learning <i>Geetanjali Basarkod and Herbert W Marsh</i>	243
Self-efficacy <i>Hyun Ji Lee and Mimi Bong</i>	250

LEARNING IN CONTEXTS AND CULTURES

Contemporary contexts for learning: an overview <i>Anastasia O Tzirides, Bill Cope, and Mary Kalantzis</i>	258
---	-----

Cross cultural lenses on learning <i>Carol D Lee</i>	267
Indigenous student identity development and school success: mitigating social identity and stereotype threats in schools <i>Melinda Webber and Abby Madden-Smith</i>	276
Students' social relationships at school <i>Kathryn R Wentzel</i>	288
A social/cultural constructivist perspective on learning "problems" <i>Curt Dudley-Marling</i>	298
Learning in the disciplines: a conceptual framework <i>Susan R Goldman</i>	305
Learning in history and social studies <i>Keith C Barton</i>	315
Learning in math(s) <i>Alan H Schoenfeld</i>	323
Learning to read and write <i>Young-Suk Grace Kim</i>	335
Learning in the sciences <i>Jonathan Osborne</i>	344
Learning from multiple texts <i>Matthew T McCrudden, Ivar Bråten, and Ladislao Salmerón</i>	353
Learning an additional language <i>Martin East</i>	364
Learning and literacy out of school: learning every day and everywhere <i>Katherine Schultz and Glynda A Hull</i>	374
Learning through work: premises, conceptions and practices <i>Stephen Billett</i>	381
Learning spaces and materiality: embracing past and present when designing for the future <i>Lucila Carvalho</i>	390
Gender differences in educational achievement and learning outcomes <i>David Reilly</i>	399
Networks and learning <i>Peter Goodyear</i>	409
Online learning <i>Peter Shea</i>	423
Social media and learning <i>Christine M Greenhow, Sarah M Galvin, K Bret Staudt Willet, and Amy L Chapman</i>	431

LEARNING OVER TIME

Learning across the lifespan: an overview <i>Patricia A Alexander and Jannah Fusenig</i>	443
---	-----

Learning and development in infancy: toward an integrated approach <i>Alecia Moser and Annette ME Henderson</i>	451
Early language development: major milestones and individual differences <i>Michelle K Tulloch and Erika Hoff</i>	467
Key developments in early childhood <i>Elizabeth Ann Wood</i>	477
Key developments during adolescence: implications for learning and achievement <i>Eric M Anderman, Yvonne Allsop, Kimiko Ching, and Seung Yon Ha</i>	486
Key developments in early and middle adulthood <i>Carol Hoare</i>	497
Older adults' development, learning, and education <i>Lisa M Baumgartner</i>	503
Education and neurocognitive aging - is there a relation? <i>Lars Nyberg and Martin Lövdén</i>	512
Lifelong learning beyond initial schooling <i>Karen Evans and Natasha Kersh</i>	520
Multiple and multi-dimensional educational and life transitions: conceptualization, theorization and XII pillars of transitions <i>Divya Jindal-Snape</i>	530
Learning progressions: affordances, critiques, and pathways toward equity <i>Alicia C Alonzo</i>	544

OPTIMISING LEARNING

Optimizing learning: an overview <i>Stuart McNaughton</i>	560
Apprenticeship learning <i>Julia Eberle</i>	568
Argumentation and critical thinking <i>Chrysi Rapanta and Kalypso Iordanou</i>	575
Computer supported collaborative learning <i>Sanna Järvelä, Päivi Häkkinen, and Piia Näykki</i>	588
Cooperative and collaborative learning <i>Noreen M Webb and Eric Burnheimer</i>	593
Dialogic education: tensions and dilemmas <i>Rupert Wegerif, Shengpeng Shi, Ana Rubio-Jimenez, Yun Long, Qian Liu, and Chih Ching Chang</i>	600
Exemplars <i>Alexander Renkl</i>	612
Feedback for student learning in higher education <i>David Carless</i>	623
Inquiry-based learning <i>Ard W Lazonder</i>	630

Intelligent tutoring systems and conversational agents <i>Arthur C Graesser and Haiying Li</i>	637
Mindfulness and learning <i>Darren L Dunning</i>	648
Learning through play-in-interaction <i>K Aronsson</i>	658
Optimizing learning through retrieval practice and spacing <i>Shana K Carpenter</i>	664
Service-learning and experiential education <i>Christian Winterbottom, Vickie E Lake, and Ithel Jones</i>	673
Simulations and games for learning <i>Tasos Hovardas, Judith Ter Vrugte, Zacharias C Zacharia, and Ton de Jong</i>	681
The efficacy of learning strategies from the ICAP perspective <i>Michelene TH Chi, Nicole S Boucher, and Jesse Ha</i>	689
Student-centered instruction: inquiry-, problem-, project-, and case-based learning <i>Sofie MM Loyens, Lisette Wijnia, and Remy MJ P Rikers</i>	701
Universal Design for Learning in its 3rd decade: a focus on equity, inclusion, and design <i>Kavita Rao, Jenna W Gravel, David H Rose, and T Nicole Tucker-Smith</i>	712
Well-being and learning <i>Tina Hascher</i>	721

This page intentionally left blank

What is learning, for whom, and to what end? An overview

Joshua A. Danish^a and Jasmine Y. Ma^b, ^a Learning Sciences, Indiana University School of Education, Bloomington, IN, United States; and ^b New York University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Our positionality	2
Learning in context (sociocultural theories of learning)	3
For whom?	3
To what end?	4
Related approaches	4
Individual learning (cognitive theories of learning)	5
For whom?	5
To what end?	6
Related approaches	6
What have we been missing?	7
For whom and to what end?	8
Final notes on reading theories of learning	8
Acknowledgments	9
References	9

Introduction

Writing this chapter in 2021, in the United States of America, it is hard not to reflect on what a complicated and messy world we live in. The world continues to struggle with COVID-19, political and social unrest including rampant racism, xenophobia, sexism, and other forms of discrimination, climate change, and uncertain economies, all at a time when technological innovations are increasing at a breakneck pace that has pundits both enthusiastic and frightened. Our personal histories with these different forces shape who we are, and how we understand and navigate the tensions that arise in our lives. All of these issues and opportunities have an undeniable impact upon learners of all ages, and thus on what it means to study learning. Recognizing these contextual issues, we believe that it is important when introducing the section of this volume that focuses on learning theories to situate those theories of learning in this time and place and try to begin a conversation about how they describe learning in the world that we live in now and in the wide range of learning contexts that reflect where learners are, including both formal and informal contexts. Thinking about all of this makes the conversation about what learning is quite a bit more complicated, but it also increases the chance that we can engage in conversations about how these theories describe, and consequently shape, learning in the world and not just within conversations between scholars.

In situating these theories in the current time and place, our intention is to attend to how they reflect, shape, and can support learning in the real world. We see this effort as grounded both in the present and the future that is informed and potentially transformed by our current efforts. That is, we want explicitly to recognize that teachers, learners, and theories of learning are all historical actors (Gutiérrez, 2018), meaning that they are shaped by their past and present contexts, while also having the potential to shape those contexts. Furthermore, the authors of these theories naturally bring their own histories and experiences with them to the act of generating and refining theories, and it is well documented that the academy is not as diverse as the populations that it often studies. The combination of these influences will, in turn, inform how theories continue to grow and evolve in the future. For example, theories that defined local funds of knowledge (Moll et al., 1992) or funds of identity (Esteban-Guitart and Moll, 2014) as resources that can be the basis of learning emerged in response to deficit-oriented approaches to learning that positioned minoritized learners as lacking key experiences or capabilities. Fortunately, the presence of these new theories brings attention to the resources that historically minoritized learners bring with them and have helped shape a whole new generation of theories and theorists (Esteban-Guitart, 2021).

Therefore, we recognize that learning contexts and learning theories are neither neutral, nor ahistorical. Rather, by characterizing how learning occurs, what counts as learning, and when, theories necessarily privilege certain visions of what it means to learn, or to be knowledgeable and capable, and who can do it well. These theories are grounded in and shaped by many forces and systems of power including the theories that came before them, which theories and whose voices get privileged in the field, the structures of the academy, the funding agencies that support research, the politics of local educational institutions, the politics of the nation states where education happens, and the ways they define their borders and those who cross those borders. Obviously, it's not possible for a single chapter to capture all of these important influences. Thus, our goal in this introductory chapter

is to encourage readers to reflect on these social and historical forces as they think about the learning theories that we introduce. We want to help readers to explore not only the basic tenets of these theories but to explicitly highlight and draw attention to how these different forces shaped the theories of learning, and impact the ways they describe learning and learners. In particular, we want to call attention to the idea that the key theories of learning that we study as a field differ not only in how they describe learning, but also in what they make salient and what they miss, or even hide, and the implications that these shifts in focus have for learners as people, and for learning as a lived, historically situated process. In making these differences visible, we encourage readers to interrogate these theories in an effort to think about how theories of learning can be complicit in defining and reifying the challenges that learners face, and also how theories of learning can support a more equitable, just, and positive future. Thus, we are both critical and optimistic in that we believe that theories, including our own, have not yet lived up to their potential to support all learners, but also that calling attention to this can help us move forward as a field to support more diverse learners in a broader range of contexts.

With these tensions and challenges in mind, we also note that the broad range of learning theories have already been reviewed in prior texts and chapters (e.g., [Greeno, 1998](#)) including some of our own work (e.g., [Danish and Gresalfi, 2018](#); [Ma et al., 2020](#)). Much of that work follows what has become a conventional chronological (based on Western scholarship) lens and sequence, building up from behaviorist accounts to cognitive and then ending with sociocultural theories. However, we fear that this common approach hides many of the challenges and motivations that lie within these theories, and also privileges certain theoretical traditions within the academy. For example, cognitive theory, with its focus on the individual, is often positioned in a privileged and central manner, as having overcome the challenges of behaviorism and served as a foundation for sociocultural theories which then get presented as having an additional focus on context and community. Many overview chapters are organized around cognitive notions of what it means to know, to represent knowledge, and to transfer knowledge despite the fact that sociocultural theories challenge the very notion of transfer (e.g., [Beach, 1999](#)), and the importance of seeing how knowledge is represented rather than how knowledge is used (e.g., [Lave, 1982](#)). Thus, while we agree that cognitive theories have much to offer, we also believe that this positioning means that cognitive theories are viewed in a simplistic light, and sociocultural theories are then viewed as an additional “layer” that can be tacked on to cognitive theory, helping to explore the role of the community and context in learning. This does a disservice to both theoretical traditions, ignoring how cognitive theories do position social contexts, and how sociocultural theories begin with fundamentally different assumptions about what it means for individuals, or groups of people to learn. We believe both of these theories have a lot to say about individuals, about social contexts, and about how each can be important for understanding learners and learning, and yet some of those ideas are lost in many of the traditionally organized discussions.

Our intention, therefore, is that this introductory chapter will help to disrupt these norms and challenge the community to reflect on why we typically review the literature in those ways, what that view hides, and what it reveals. We aim to shake things up a little and offer an alternative perspective to help us all think differently about our own theories and those of our colleagues as we look toward the future and reflect on what theory might do, and might become. We therefore present sociocultural theories first, followed by cognitive theories. We are also inspired by the work of [Philip et al. \(2018\)](#), to ask the following questions to help situate each of the theoretical traditions we will review: **How does this theory address learning, for whom, and to what end?** The intention of these questions is to recognize explicitly the way in which each theory and its chosen focus necessarily illuminates certain learning phenomena and simultaneously hides others. This necessarily shapes whose experiences, opportunities, and struggles might be highlighted or impacted by the theories, as well as the outcomes that the theories define and support as valuable. Using this lens helps us focus on how each theory may impact learners and learning within our societies. It is important to note that in attempting to answer this, we are necessarily offering our perspective on who is supported by any given analysis or study, and the goals of those studies. This will naturally reveal our own theoretical biases regarding the impact these theories have had. Whether a specific intervention does or does not actually feel inclusive to a given person or group of people is an empirical question that we might answer in the future. Answering the question “to what end” is even more potentially controversial as not all authors address this concern. Again, therefore, our goal is to present our perspective, and encourage each reader to develop their own. After discussing “for whom” and “to what end,” we highlight a few new directions that we believe have emerged in each tradition that we believe will be part of defining future scholarship.

Our positionality

Before we articulate our views of the key learning theory traditions, we believe it is important to note who we are as that shapes our approach to this exercise. We both consider ourselves “Learning Scientists” who have elements of “traditional” training in learning theories (e.g., the first author has a Ph.D. in Educational Psychology). However, we would both label ourselves as primarily sociocultural theorists. We are also deeply interested in pursuing issues of equity and justice through the study of learning. We also identify as theory builders and educators who seek to expand and transform the possibilities for how these theories of learning are understood. That is, like all authors, we have our biases. At the same time, our intent is not to label a single “correct” or objectively “better” theory because we do not believe that such a thing exists. Rather, we suggest that each theory is better for different people, goals, contexts, and tasks, and that these theories also have the potential to cause harm both directly and indirectly. We want to help all readers of learning theories to reflect on that, and on their own values as they consider learning.

Learning in context (sociocultural theories of learning)

Among the many flavors of sociocultural theories of learning, the commonality between them is that context is inextricably embedded in accounts of how people learn in messy real-world situations. However, the definition of context is defined in a range of ways, as is the relationship between context and learning—both of which, in turn, shape how learning is defined. One commonality, however, is that culture is considered a key element of context. Furthermore, as Lee et al. (2020) noted, “it is not simply that culture *impacts* development, but that development itself is cultural in nature” (p. xxii). For example, cultural historical activity theory (CHAT), which itself has many flavors, builds from Vygotsky (1978) and others’ arguments (e.g., Leontiev, 1978; Luria, 1976) that all cognition and therefore learning is shaped by the context of human activity, and the cultural tools that mediate all such activity (Cole and Engeström, 2007). Mediation refers to the ways that cultural tools can transform our perception of, and relationships with, people, tools, and goals, and all mediators are both material and semiotic, reflecting the fact that they have a physical impact on the world as well as a social one (Wertsch, 1981; 2017). Within a CHAT account, learning depends on changing interactions between people and tools (the mediators), developed over time by cultural communities, and these interactions in turn shape and potentially change those cultural communities. To understand how learning happens within these interactions, the key unit of analysis for CHAT is the activity system, which is defined by a shared object or set of goals that a community of people pursues together. To understand learning, we need to be able to see how relationships and tools are taken up in new ways through meaningful human activity (e.g., a shift in mediational means).

While some count only those accounts of learning based on activity theory in the category of sociocultural, we include all theories that take learning and development to be interdependent with socially and culturally constructed context. Legitimate peripheral participation (LPP; Lave and Wenger, 1991), for example, developed through a framework of social practice theory, treats learning as shifts in participation in activity. The focus for LPP is on what learners *do* in the context of activity, or how their practices change, rather than what they *know* (or come to know). To many, LPP and CHAT sound quite similar, though Sfard (1998) characterized the important difference between theories that use the metaphor of acquisition of knowledge to describe learning (e.g., the internalization or appropriation of mediational tools that Vygotsky referred to), and theories that instead use the metaphor of participation (e.g., the changing participation that Lave and Wenger referred to). Some sociocultural theorists bridge this divide by incorporating elements of both acquisition and participation metaphors in their accounts.

Others have also noted that the nuanced differences between these theories have implications for how educators may design learning experiences (Engeström, 1991). For example, a social practice theory like LPP invites analysis to foreground not just how individuals change as they participate in the ongoing work of their communities, but also attend to how the practices themselves are constituted through “situated negotiation and renegotiation of meaning in the world” (Lave and Wenger, 1991, p. 51). A Vygotskian (1978) lens, on the other hand, may focus on how individuals appropriate the use of individual mediational means. In contrast, more recent CHAT work appears to incorporate elements of both by focusing on how individuals are shaped by mediational means, but also how they participate in mediational actions in new ways that in turn can transform the activity system (Engeström, 2018).

Despite these differences, it becomes clear that sociocultural lenses view activity and social practice as more than simply a context *for* learning, but instead part and parcel *of* learning, and vice versa. In fact, many who use these theories understand “knowledge” as only taking on meaning *in use*, rather than a static system of skills and concepts acquired and carried around by learners. Therefore, learning is assumed to be *always* occurring, despite educators’ anxieties that their students are not learning and policy-makers’ war against “learning loss” (McKinney de Royston and Vossoughi, 2021). In other words, rather than *whether* something was learned or not, the questions that these theories help us ask are about *what was learned*, given the particularities of any given context.

For whom?

In order to discuss who learns in sociocultural studies of learning, it is valuable first to understand how sociocultural theories characterize identity. Sociocultural lenses treat identity as an integral part of learning. LPP, for example, is explicit that learning “implies becoming a full participant, a member, a kind of person” (Lave and Wenger, 1991, p. 53). Therefore, the study of an individual’s learning is incomplete without taking into account their shifting identities in relation to the learning of interest. While different ways of theorizing identity have been developed (e.g., Holland et al., 1998; Nasir, 2002; Stetsenko, 2009; Wenger, 1998), those that are compatible with (and build upon) sociocultural theories understand that identities are multiple, relational, formed and negotiated in context and in interaction, and shape learning even as they themselves develop (Philip and Gupta, 2020). This complicates any attempt to answer, “For whom?” because it challenges the very notion of categorizing people.

Furthermore, given how identities are implicated in sociocultural theories of learning, who learners are (or can be) must be investigated in relation to their contexts. Gee (1989) gave the example that it would not be very sensible to ask if the Pope was a bachelor, given his vow to remain unmarried. Similarly, in tracked schools it might make sense to discuss students as “fast” or “slow,” as the teachers in Horn (2007) did, whereas in an untracked school, where all students are assumed to be capable of success and rate of learning is less emphasized, these categories may be irrelevant. Additionally, theoretical lenses shape the possibilities for how analysts see different kinds of actors. Sociocultural theories illuminate not just the developing identities of learners, but also the social, cultural, and historical contexts that allow for and support or suppress particular kinds of identities. This stance makes it imperative to study diverse learners in the diverse contexts in which they live. For this reason, many scholars interested in supporting historically minoritized learners have chosen sociocultural theories as their underlying framework.

A number of theories have emerged from a sociocultural framework that focus on the challenges that arise for learners from non-dominant communities, particularly when the cultural history that they bring with them is not valued by the dominant culture of their school, context, or society (Moll et al., 1992). For example, Funds of Identity theory builds on Funds of Knowledge (FoK) and other sociocultural theories to highlight how individual learners are each unique and may connect to or find relevance in different aspects of their cultural identities, which can in turn serve as resource for their learning (Esteban-Guitart, 2021; Esteban-Guitart and Moll, 2014). Culturally responsive pedagogy inspired theories (Ladson-Billings, 2014) build on the idea that learners bring valuable cultural resources with them to formal learning contexts, and further note the importance of helping learners take a critical lens to their schooling so that they can appreciate and navigate the challenges that formal schooling often has for minoritized learners. Culturally sustaining theories go even further to promote recognizing, supporting, and lifting up learners' home cultural knowledge, explicitly arguing against assimilation as a goal of schooling (Paris and Alim, 2014). In short, while sociocultural theory does not require a focus on historically minoritized learners, it has proven a powerful tool for researchers to articulate and examine how learners' identities may conflict with those of the dominant cultural groups in a given context, and has begun to provide insights into how the field can support more diverse learners in these circumstances.

To what end?

As mentioned above, sociocultural theories have illuminated schools and classrooms as cultural communities in and of themselves. These studies have often been conducted with a goal of supporting all learners in being successful, though scholars have noted that this may require schools and other institutional factors to be fundamentally transformed (Sengupta-Irving, 2021). Sociocultural theories have proven useful tools for better understanding schooling as laden with cultural norms and practices that, at best, are a mismatch for some communities of learners, and, at worst, systematically marginalize and otherwise harm them (Nasir et al., 2008). In fact, some have gone further to point out how school-based learning is a relatively recent invention, and often at-odds with developing robust understanding of how to participate in society (e.g., Vakil, 2018). A result of these efforts is that the field has also taken up sociocultural theories to understand that schools are not the only place valuable learning happens; decentering schools has the potential to also decenter the dominant values and practices embedded there and, consequently, acknowledge and lift up others that have historically been erased. That is, many sociocultural studies of learning have an explicit goal of rethinking what education looks like, and for whom.

In addition to studying schooling, sociocultural theories have been taken up to study disciplinary practices, including mathematics, science, and literacies, in family, community, and workplace contexts (e.g., Booker and Goldman, 2016; Hall and Jurow, 2015), and how the organization of those contexts shapes disciplinary practices. This has served two important purposes. First, these studies have provided new ways for thinking about formal schooling, and how it might be reimagined to support learning of disciplinary norms. Second, however, these studies have shown how important it is to examine learning in non-schooling contexts as an important space for helping all learners to be successful in appropriating disciplinary expertise of value to them (Jackson, 2011).

Related approaches

In this section we briefly address some of the many promising approaches that have grown from sociocultural theory and gained momentum in the past decade. One long-standing concern of sociocultural theorists involves the many communities or activity systems of which a learner is a part. One way to take these into consideration is what has been called "third generation CHAT", where multiple intersecting activity systems constitute contexts for learning (Engeström, 2018). This notion has been addressed in a variety of additional ways, from Vossoughi and Gutiérrez (2014) arguing for multi-sited orientations to Barron's (2006) learning ecologies framework and Bell et al.'s (2021) cultural learning pathways. There are nuances differentiating these ideas, but crucial for all of them is the assembling of multiple contexts that learners navigate. This has had the general effect of bringing various contexts previously thought incompatible with, or irrelevant to, each other into analytic relationship, demonstrating how home and community practices can support school learning (e.g., Civil, 2002; Lee, 2001), or how the views of students constructed by processes of schooling can be narrow and harmful (e.g., Jackson, 2011). Further, better understanding of the various contexts of learners' activity is critical for researchers and educators to move beyond deficit views of marginalized communities and learn about and how to value their diverse practices and ways of knowing (Nasir et al., 2008). An important dimension of this work is that it also frequently treats the content or community as a learner, recognizing that clubs, schools, workplaces, and communities can all learn and grow (Cole and Engeström, 1993; Engeström, 1991, 2018). Furthermore, these communities grow in a dialectical relationship between the individual and the community, with individuals helping to transform the community, while also taking on new community practices. This fundamentally shifts the answer to "for whom?" to include groups of people in new and powerful ways.

Sociocultural theorists have also begun to consider more expansive notions of what constitutes a context for learning. For example, some theories of embodied cognition and embodiment consider learning to be located in the whole body, or multiple bodies in interaction, rather than simply the individual mind or brain, thereby implicating the body as both a cultural artifact that mediates learning as well as the learning subject, while continuing to highlight the role of social context and interaction in mediating this process (e.g., Danish et al., 2020; Gholson and Martin, 2019; Hall and Nemirovsky, 2012; Kelton and Ma, 2018; Takeuchi and Ishihara, 2020). Another aspect of context made increasingly relevant is spatiality, beyond space as static containers for activity (Leander et al., 2010). How spaces are constructed, oriented to, moved through, and understood as social sites of activity

and power have crucial consequences for what kinds of learning are made available, and for whom (e.g., [Marin, 2020](#); [Rahm, 2014](#); [Taylor, 2020](#)). Finally, we note the importance of phenomena that have begun to be investigated simultaneously as targets for learning as well as context for learning, such as emotion ([Vea, 2020](#)), dignity ([Espinoza et al., 2020](#)), ideology ([Philip et al., 2018a,b](#)), and axiological positionings ([Bang et al., 2016](#)). These bring an explicitly ethical dimension to our theories of learning, which we argue is urgently important in an effort for education research to work toward justice.

Individual learning (cognitive theories of learning)

There are many theories that fit under the “cognitive” umbrella including cognitive science, social cognitive theory, sociocognitive theory, constructivist theory, and information processing theories, but what they share in common is a fundamental focus on how thinking and learning of individuals can be modeled and understood in individual minds, with work in neuroscience going one step further to explore how knowledge is represented within the brain itself. By modeling these processes it becomes possible not only to understand them, but to support them more effectively. The majority of these approaches also build on computers and computational techniques as a sometimes implicit, underlying metaphor for how the mind and brain work, and many trace the early rise of cognitive science to the parallel rise in computing and computer science ([Block, 1990](#)). Building on the metaphor of the computer, many cognitive approaches focus on how information is processed, and stored in or retrieved from memory ([Bransford et al., 2000](#)). While more recent work does not always explicitly mention the computer as an inspiration, the focus on modeling how thinking, learning, memory, and perception are all related in one system is quite common. In part due to efforts to model cognition in this way, and in part due to a long tradition in psychology and related fields that have focused on individuals, cognitive approaches to thinking and learning typically privilege the individual person as the primary unit of analysis. Cognitive theorists then explore how information is perceived by the individual, represented in their mind and, ultimately, transformed or processed to support the kinds of activities that scholars are interested in.

Despite these similarities, cognitive theories focus on a broad range of phenomena, at grain sizes ranging from the process by which ideas are captured in our neurons to how we solve problems in complex tasks. For example, constructivists, building on the tradition of Piaget, tend to focus on the various kinds of information that individuals process (e.g., whether concrete or abstract), and document how this changes as children develop, with a particular focus on how it influences their performance on specific tasks ([Ginsburg and Oppen, 1988](#)). In contrast, cognitive scientists typically focus more explicitly on how important concepts are perceived and represented, while social constructivists explore the role of other people in one’s life in shaping one’s cognition ([Danish and Gresalfi, 2018](#)). Related theories such as theories of motivation can then build on this framework by exploring how cognitive processing is shaped by related variables including how learners value the content they are studying, and whether they believe that they can learn it—with research indicating that learners are more likely to persist in situations that they value and where they expect a reasonable chance at success ([Hidi and Harackiewicz, 2000](#)). Theories of meta-cognition also extend the focus of cognitive processing to consider how learners reflect on cognition itself, and how this can lead to strategies for improving one’s own learning and memory ([Finley et al., 2010](#)).

In contrasting cognitive theories with sociocultural theories, it is important to note that cognitive theorists do not deny the fundamental importance of culture and context. Rather, they view these as variables that might be modeled along with other important factors and viewed within the same fundamental model of cognition. After all, a key concept in cognitive theories of learning is how prior knowledge shapes one’s experiences of new situations, and how one connects new knowledge to existing structures ([Bransford et al., 2000](#)). From this perspective, experiences that are tied to the social world are modeled as aspects of perception, prior knowledge, and knowledge construction. For example, schema theory has long focused on how individuals remember, understand, and engage with cultural norms or scripts ([Rumelhart, 1980](#)). This difference in how the theories explore social interactions has led to a fundamental tension between theoretical frameworks, with cognitive theorists noting that many sociocultural frameworks are too vague in how they describe knowledge and knowledge construction because they do not explicitly model those interactions, while sociocultural scholars push back noting that treating context and culture as a variable diminishes their importance and allows one to treat thinking and learning as separable from context in ways that sociocultural theorists believe are inherently problematic. Some of these disagreements can be best seen in a now famous debate between [Anderson et al. \(1996\)](#), representing individual approaches, and [Greeno \(1997\)](#) representing sociocultural ones. This tension has led to some efforts to reconcile the disparate theoretical frameworks by viewing cognition as situated within sociocultural contexts, and by attempting to synthesize across them ([Danish and Gresalfi, 2018](#); [Sfard, 1998](#)). For example, the cognitive apprenticeship model explores how social relations of apprenticeship can support and help develop cognitive processes ([Collins et al., 1988](#)).

For whom?

Cognitive theories have historically been at the forefront of identifying how learners can engage with the kinds of content that have long been privileged in schools, helping to identify how learners come to be able to read and write, engage with mathematics, explore scientific concepts, and understand historical accounts ([Bransford et al., 2000](#)). Cognitive theories are particularly well-suited for the focus that is common in Western schooling traditions on understanding how individual students struggle, or excel when engaging with these core, disciplinary ideas as presented by schools. To accomplish this, cognitive theorists typically identify

a normative vision of what is valued as important for success within a discipline, often identified by disciplinary experts (Herrenkohl and Polman, 2018). They then model how individuals do or do not live up to that model. This can be quite powerful because it provides insights into the underlying mechanisms of learning, and thus provides clear suggestions for what might lead some learners to struggle. However, this approach also naturally centers those who are pursuing and/or interested in following these normative trajectories, with less attention to those who aim to resist or pursue alternative goals.

Nonetheless, within the bounds of school-based learning, cognitive approaches have proven useful in supporting both those who are traditionally successful, and those who struggle. For those who are traditionally successful in school, cognitive accounts provide both an explanation, and an opportunity to excel. In contrast, for those who struggle, cognitive accounts have proven crucial in helping identify underlying causes for perceived failure. For example, research that focuses on the relationship between metacognition and self-regulation has noted that successful learners are able to track their progress as they learn new content and challenge themselves to persevere in working with challenging ideas. Unsuccessful students struggle with this, but a wide range of scaffolds have been designed which can help learners to regulate their learning experience and know when it is best to revisit an idea or move on to a new one (Prather et al., 2020).

To what end?

The majority of the individual-focused, cognitive work we are most familiar with is aimed at achieving one of two ends, and often, but not always, both: (1) developing a more robust account of cognition and learning as a form of basic science from which more applied approaches can build, or (2) focusing on the applied dimension by improving practice that relies on cognition, for which education is but one practice, with others including professions, decision making, and building computer systems, to name a few. The first focus aligns well with the tendency of many cognitive studies to employ laboratory-based, experimental studies in that it is possible to build a robust and nuanced model of learning in those environments, without the kinds of complications and confounds that make some implementation-oriented studies challenging. In contrast, studies that share this second focus on improving practice must address many of these issues, though they often aim to do so by developing generalizable claims that cut across contexts and contextual variables because of the underlying assumption that minds operate in similar ways, and thus differences in how individuals act can be attributed to observable differences in their experiences, the context, or both.

In the first category, a key outcome of this effort is that cognitive models tend to be detailed and precise, and to predict outcomes in controlled environments. This focus on predictability and explainability is a hallmark of this approach and provides designers and researchers with clear empirically-testable design choices. For example, Mayer's theory of multimedia learning (Mayer, 2019) has evolved to include an explicit list of principles that can be employed when developing multimedia content to be viewed by individual learners. This includes the multimedia principle, which notes that the simple inclusion of pictures with words can help learners to develop a more elaborate understanding when the pictures add information that expands on and clarifies the words (redundant information is less productive).

If we look deeper into the second category that focuses on informing and improving practices including those in schools, it is important to ask what kinds of learning, or subjects are typically a focus? Here we see a distinct focus on disciplinary content that aligns with the current structure of school systems. For example, in their introduction to the Handbook of Research on Learning and Instruction, Mayer and Alexander (2010) note "The psychology of subject matter is the scientific study of how people learn school subjects such as reading ... and has been recognized as one of the major accomplishments of educational psychology" (p. 3). They rightly note that studies of these disciplinary areas go back over a century, and are oriented toward long-standing curricular boundaries, often those defined by experts within their field, or as is often the case, at least in part by politicians and political forces that shape many of our national education systems. There is a clear benefit to this in that it focuses attention on helping students to succeed within existing educational structures, with a specific goal of those students being able to perform successfully within those contexts (e.g., pass key tests). Within many societies, this process is also linked quite explicitly to future employability, and thus the goal is to train members of the society for employment within their local context.

Much of this work is also aimed at efficiency—how to support learners most effectively, in accomplishing as much as possible at the lowest possible cost in terms of time, resources, or both. This is certainly a goal that drives much of corporate-backed research. Even in purely academic contexts, however, increasing demands on teacher time, and decreasing budgets in many educational systems necessitate this kind of focus on efficiency. There are also motivational benefits, as students can spend their time engaging in productive learning opportunities rather than dedicating time to ineffective approaches.

Related approaches

While there are too many subfields to list here, a few key areas that have gained increasing attention and which each have their own chapter in this volume are educational explorations of neuroscience, advances in computational modeling of cognition and learning, and embodied cognition.

One prominent trend is attention to how innovations in neuroscience help the field to think more deeply about how cognition works on multiple levels. This is incredibly important given the value of viewing cognition at multiple levels simultaneously, as a way of appreciating if and how each level might be shaped by those which help to define it, or those that it helps to define (Marr, 1982). For example, many cognitive accounts are built on the assumption that cognition is both enabled and constrained

by various forms of memory (long and short term, conceptual or semantic, emotional and conceptual, etc.). Research into neuroscience has provided us with additional detail regarding how those memories are stored, and how they are related to each other in addition to how different parts of the brain contribute to this kind of activity given that findings indicate that the neurological architecture of thinking and learning is distributed (Just and Varma, 2007). This kind of research can provide simultaneously insights into prior observations about how we remember or recall important concepts, as well as the kinds of concepts that might be important to look for in studies of cognition. However, it is important to note that limits in how we can currently examine neurological processing necessarily limit the kinds of studies that can be conducted. For example, many measures such as fMRI rely upon focusing on simple, discrete actions such as decoding a word while remaining still (Gazzaniga et al., 2013), and struggle to document the kinds of long-term, continuous activity common in many classrooms. There is optimism that growth in neuroscience techniques and methods will overcome these challenges, but this necessarily limits the benefits of these findings in the short-term.

Advances in understanding neural architecture, in many cases, have paralleled advances in how we can model thinking and learning in computers. Continuous advances in computational power and tools have also fueled incredible growth in efforts that aim to model cognition and learning using computers. Broadly, these efforts can be divided into two overlapping categories: (1) efforts to reproduce intelligent behavior, and (2) efforts to document and understand cognition and learning. The first category includes efforts to produce Artificial Intelligence—tools that behave in ways that emulate or exceed what humans can do. Our lives are filled with such tools ranging from Apple’s Siri and Amazon’s Alexa to self-driving cars. The second category refers to those efforts to either model how humans are learning, such as cognitive tutors (Koedinger and Corbett, 2006) or pedagogical agents (Schroeder et al., 2013), and those efforts to identify patterns in learning and learning activity, often referred to under the umbrella term “Learning Analytics” (Wise et al., 2014). Naturally, innovations in these two areas often inform one another in powerful, synergistic ways. However, it is important to again note that the nature of these technologies often limits them. For example, many intelligent tutors are focused on math or computer science concepts that are easier to represent in a computer, with less attention to humanities content (for an exception see Easterday et al., 2012). Nonetheless, intelligent tutors have been shown to help thousands of learners to engage with previously challenging concepts (Koedinger and Alevan, 2015). However, it has been well documented how computational technologies, while promising, are not easily adopted and incorporated in all school systems and, in fact, are often adopted first in the privileged school districts that need them least (Apple, 1998). Therefore, we caution readers to examine these technological innovations with an eye toward who is currently benefiting from them, and what implications this might have for the future integration of these tools in schools and informal learning environments?

Embodied cognition is also studied quite heavily in the context of the individual and individual cognition, sometimes along with the neural basis of embodiment. The embodied turn in cognitive theories arose out of recognition that many of the same underlying cognitive/neurological systems that drive perception and action are also implicated in the kinds of reasoning that are often of value to scholars (Lindgren and Johnson-Glenberg, 2013). From this finding emerged recognition that a deeper understanding of this link could help us better understand how certain ideas are engaged, and how they can further be supported by attention to this overlap. For example, Alibali and Nathan (2012) demonstrated the importance of teachers’ gestures for helping learners to understand complex mathematical relationships. A range of theories have arisen that argue for various levels of the link between cognition and the body, some arguing that the body is always involved in cognition and learning, while others instead recognize only some crucial connections with key analogies that are exploited by learners and might be made more salient by educators. As with many other innovations, the focus on embodiment has largely been tied to existing disciplinary concepts that are already well-understood and are based on the kinds of embodiment that are familiar to the majority of learners. Fewer studies have focused on the kinds of concepts that embodiment might open up, that are not previously core to our learning environments. In addition, while the focus on embodiment has undeniable potential, we think it is worth asking what this means for learners who do not want to embody ideas, who have non-normative ways of embodying ideas, or who have specific traumatic experiences with embodiment (Fitzgerald, 2005)? Our hope is that future work in embodied cognition from both the individual and sociocultural perspectives will also incorporate these many learners.

What have we been missing?

As mentioned above, theories of learning are neither ahistorical nor neutral. They develop and get taken up in the context of social, political, and economic agendas (e.g., see Philip and Sengupta, 2020), and necessarily privilege certain viewpoints and background others. In doing so, they illuminate particular aspects of learning phenomena and obscure, or even distort, others. We acknowledge that, even with the best of intentions, in this overview chapter we have taken up and centered those clusters of theories dominant in the learning sciences. We hope that this final section may provide a developmental dimension to our overview—some insight into the field (and the authors of this chapter) as dynamic and itself (and ourselves) learning, with all the possibilities that may hold. Inspired and informed by less-commonly cited as well as more recent scholarship, we make note of what the dominant theories above have neglected, or decentered, or, in some cases, deliberately eradicated, and point to work that promises to shine the light anew on learning phenomena and the communities which steward them.

We begin by making note of two “common sense” (Garfinkel, 1967) assumptions and dimensions of dominant theories of learning to help bring into relief some of what the field has, for the most part, neglected. To begin, it is widely accepted that some activities count as learning, while (most) other activities do not. In fact, it is assumed that learning of value is so difficult

to achieve that most of our educational research, policy, and on-the-ground efforts are directed toward promoting it. However, research that decenters environments designed primarily for learning, like schools, helps us to decouple theories of learning from theories of teaching and thus recognize that important learning happens everywhere, all the time (Lave, 1982). Studies of previously unacknowledged or undervalued learning may contribute transformative dimensions for what matters for learning, or what counts for learning. Headrick Taylor's (2017) work, for example, highlights mobility and spatiality as a previously dismissed dimension of learning by exploring how youth who had access to bicycles began to think of their urban locale in new ways and allowing us to better understand emplacement and place-making as consequential for the health and survival of communities. A focus on mobility and spatiality not only illuminates neglected objects of learning, but also may expand our understandings of what counts as previously taken-for-granted disciplinary learning (e.g., mathematics; Kelton and Ma, 2020). Relatedly, scholarship that interrogates different forms of learning (e.g., Rogoff's work on learning by observing and pitching in, 2014) may contribute to theorizing that brings into relief previously unvalued activity, especially among non-dominant communities, as learning.

Additionally, we tend to see learning as an individual human phenomenon. To focus on individuals means our investigations of cultural and collective learning are limited to how they relate to individual learning, or perhaps how the individual and social are interrelated, but always with the individual at center. However, studies of consequential learning (Jurow et al., 2016) and social movements (Curnow and Jurow, 2021) demonstrate that units of analysis beyond the individual are valuable in their own right, not simply because of their interrelations with individual development. This learning-as-individual-and-human orientation also reproduces human exceptionalism, the idea that humans are fundamentally different from and more important than other forms of life (Bang, 2016). In CHAT, for example, learning is defined through the premise that "material objects are modified by human beings as a means of regulating their interactions with the world and each other" (Cole and Engeström, 2007, p. 485). Our theories of learning are therefore limited to treating the more than human, lands and water, and other phenomena as context for—or objects of—learning, rather than subjects of learning. Bang argued that this constitutes an epistemic erasure of Indigenous ways of knowing, and worse, has served to justify the denial of the most basic of rights to groups like enslaved peoples and communities that have been targets of genocide. In contrast, by recognizing the ways that Indigenous communities view the more-than-human world as part of their own experiences it becomes possible to study new ways of learning when humans explore the natural spaces around them (Marin and Bang, 2018).

For whom and to what end?

The contexts in which theory itself is constructed are inextricably embedded within the theory itself. It is not simply trivia, for example, that Vygotsky's ideas were developed (and then, along with other educational psychology-related work, suppressed) in postrevolutionary Russia (Cole and Scribner, 1978; Kozulin, 1986), or that Lave's situative theories developed in response to worrisome "political and ethical implications of cross-cultural psychological experimentation" (p. 1, Lave, 2011). We highlight this not to suggest that theories cannot be taken up across contexts, but instead to emphasize the importance of historicizing theory as situated tools for meaning-making. This is important for avoiding, for example, naive assumptions about theory as universal or neutral. At the same time, it is equally critical to avoid notions of portability—that Indigenous epistemologies, for example, can or should be used by anyone in any expedient manner. Instead, theory may be read and taken up with a deliberate eye toward voices and values that may be silenced or otherwise marginalized (Gholson, 2019; Philip and Sengupta, 2020), or co-opted (Tuck and Yang, 2012). This is a step toward more deliberately considering for whom we do our theorizing, and to what end.

The power of theory lies in its ability to represent the world and provide a "reality" to its users (Frye, 1983). In providing a reality, however, Frye noted the ease with which entire communities may be rendered non-existent. For theories of learning, this often entails assumptions about what normative learning processes or outcomes look like, with consequences for whom is viewed as able to learn and who is not. As a consequence, in some cases, communities or categories of learners get studied separately (e.g., learners with dis/abilities, emergent bilinguals, or racially marginalized learners). In others, categories become meaningless (e.g., LGBTQI identities are often treated as irrelevant in STEM learning; Velamuri et al., 2021). In some cases, the historical and socio-political contexts of learners necessitate theory that takes these issues into consideration (e.g., Annamma et al., 2013, take up Critical Race Theory in the context of dis/ability studies). We argue, however, that these concerns must not remain the domain of "specialized" sub-disciplines of the learning sciences, but instead constitute the purview of all learning scientists.

Final notes on reading theories of learning

Following Philip and Sengupta, we urge readers of learning theories to read contextually, consequentially, and contrapuntally. They explain:

To read contextually is to account for the political influences and forces, people, and openings that seed designs and theories and allow them to emerge, flourish, or languish. To read consequentially is to consider the possible ramifications of innovations, with an eye toward responsibility (Stilgoe et al., 2013), and with the courage for new social and political imaginations (Cutierrez, 2016). To read contrapuntally, as Said (1993) outlined, means "to draw out, extend, give emphasis and voice to [the logics of and resistances to empire and imperialism that] are silent or marginally present or ideologically represented" (p. 66) in our designs and theories (p. 335).

We add to this call the importance for scholars to read theory reflexively. Lather's (1986) points about reflexivity in terms of research design and methodology are relevant here as well. Our own positions in relation to current political configurations, the learners we study, the institutions in which they are entangled, the institutions in which *we* are entangled, are all implicated in how we read, and subsequently take up theory. In particular, we highlight the idea of reciprocity, reading dialogically to see how theory might help us negotiate meaning and value with the learners we study. In these ways readers of theory may construct what Gholson (2019) called a scholarly catchment, understanding theoretical narratives from multiple perspectives and positions, "flowing like water" (p. 419). In this scholarly catchment lies possible futures our theories of learning may lead us to.

Acknowledgments

We thank members of the Indiana University Representations, Activity, Play and Technology (RAPT) lab and New York University's Interaction, Analysis, and Learning Lab for their feedback on this chapter. We also thank Michael Kieffer for his feedback on several sections of this manuscript.

References

- Alibali, M.W., Nathan, M.J., 2012. Embodiment in mathematics teaching and learning: evidence from learners' and teachers' gestures. *J. Learn. Sci.* 21 (2), 247–286.
- Anderson, J.R., Reder, L.M., Simon, H.A., 1996. Situated learning and education. *Educ. Res.* 25 (4), 5–11. <https://doi.org/10.3102/0013189x025004005>.
- Annamma, S.A., Connor, D., Ferri, B., 2013. Dis/ability critical race studies (DisCrit): theorizing at the intersections of race and dis/ability. *Race Ethn. Educ.* 16 (1), 1–31. <https://doi.org/10.1080/13613324.2012.730511>.
- Apple, M., 1998. Teaching and technology: the hidden effects of computers on teachers and students. In: Beyer, L., Apple, M. (Eds.), *The Curriculum: Problems, Politics, and Possibilities*. State University of New York Press, New York, pp. 314–336.
- Bang, M., 2016. Towards an ethic of decolonial trans-ontologies in sociocultural theories of learning and development. In: Esmonde, I., Booker, A. (Eds.), *Power and Privilege in the Learning Sciences: Critical and Sociocultural Theories of Learning*. Routledge, pp. 115–138.
- Bang, M., Faber, L., Gurneau, J., Marin, A., Soto, C., 2016. Community-based design research: learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind Cult. Activ.* 23 (1), 28–41. <https://doi.org/10.1080/10749039.2015.1087572>.
- Barron, B., 2006. Interest and self-sustained learning as catalysts of development: a learning ecology perspective. *Hum. Dev.* 49 (4), 193–224. <https://doi.org/10.1159/000094368>.
- Beach, K., 1999. Consequential transitions: a sociocultural expedition beyond transfer in education. *Rev. Res. Educ.* 24 (1), 101–139.
- Bell, P., Tzou, C., Bricker, L., Baines, A.D., 2012. Learning in diversities of structures of social practice: accounting for how, why and where people learn science. *Hum. Dev.* 55 (5–6), 269–284. <https://doi.org/10.1159/000345315>.
- Block, N., 1990. The computer model of the mind. In: Osherson, D.N., Smith, E.E. (Eds.), *Thinking: An Invitation to Cognitive Science*, vol. 3. MIT Press, Cambridge, MA, pp. 247–289.
- Booker, A., Goldman, S., 2016. Participatory design research as a practice for systemic repair: doing hand-in-hand math research with families. *Cognit. Instruct.* 34 (3), 222–235. <https://doi.org/10.1080/07370008.2016.1179535>.
- Bransford, J., Brown, A.L., Cocking, R.R. (Eds.), 2000. *How People Learn: Brain, Mind, Experience, and School*. National Academies Press.
- Civil, M., 2002. Everyday mathematics, mathematicians' mathematics, and school mathematics: can we bring them together? *J. Res. Math. Educ. Monogr.* 11, 40–62.
- Cole, M., Engeström, Y., 1993. A cultural-historical approach to distributed cognition. In: Salomon, G. (Ed.), *Distributed Cognitions: Psychological and Educational Considerations*. Cambridge University Press, New York, pp. 47–87.
- Cole, M., Engeström, Y., 2007. Cultural-historical approaches to designing for development. In: Valsiner, J., Rosa, A. (Eds.), *The Cambridge Handbook of Sociocultural Psychology*. Cambridge University Press, pp. 484–507.
- Cole, M., Scribner, S., 1978. Introduction. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, pp. 1–14.
- Collins, A., Brown, J.S., Newman, S.E., 1988. Cognitive apprenticeship: teaching the craft of reading, writing and mathematics. *Thinking* 8 (1), 2–10.
- Curnow, J., Jurrow, A.S., 2021. Learning in and for collective action. *J. Learn. Sci.* 30 (1), 14–26.
- Danish, J.A., Gresalfi, M., 2018. Cognitive and sociocultural perspective on learning: tensions and synergy in the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, NY, pp. 34–43.
- Danish, J.A., Enyedy, N., Saleh, A., Humburg, M., 2020. Learning in embodied activity framework: a sociocultural framework for embodied cognition. *Int. J. Comput. Support. Collab. Learn.* 15 (1), 49–87. <https://doi.org/10.1007/s11412-020-09317-3>.
- Easterday, M.W., Pinkwart, N., McLaren, B., 2012. Policy World: a cognitive game for teaching deliberation. *Educ. Technol. Teach. Argument. Skills* 225–276.
- Engeström, Y., 1991. Non Scolae sed Vitae Discimus: toward overcoming the encapsulation of school learning. *Learn. Interact.* 1, 243–259.
- Engeström, Y., 2018. Expansive learning: towards an activity-theoretical reconceptualization. In: *Contemporary Theories of Learning*. Routledge, pp. 46–65.
- Espinoza, M.L., Vossoughi, S., Rose, M., Poza, L.E., 2020. Matters of participation: notes on the study of dignity and learning. *Mind Cult. Activ.* 27 (4), 325–347. <https://doi.org/10.1080/10749039.2020.1779304>.
- Esteban-Guitart, M., Moll, L.C., 2014. Funds of identity: a new concept based on the funds of knowledge approach. *Cult. Psychol.* 20 (1), 31–48.
- Esteban-Guitart, M., 2021. Advancing the funds of identity theory: a critical and unfinished dialogue. *Mind Cult. Activ.* 28 (2), 169–179. <https://doi.org/10.1080/10749039.2021.1913751>.
- Finley, J.R., Tullis, J.G., Benjamin, A.S., 2010. Metacognitive control of learning and remembering. In: Khine, M.S., Saleh, I.M. (Eds.), *New Science of Learning*. Springer, pp. 109–132.
- Fitzgerald, H., 2005. Still feeling like a spare piece of luggage? embodied experiences of (dis)ability in physical education and school sport. *Phys. Educ. Sport Pedagogy* 10 (1), 41–59.
- Frye, M., 1983. To be and be seen: the politics of reality. *The politics of reality: Essays in feminist theory*. The Crossing Press, pp. 152–174.
- Garfinkel, H., 1967. *Studies in Ethnomethodology*. Polity Press.
- Gazzaniga, M.S., Ivry, R.B., Mangun, G.R., 2013. *Cognitive Neuroscience: The Biology of the Mind*, fourth ed. W. W. Norton, New York.
- Gee, J.P., 1989. What do English teachers teach? (or why isn't the Pope a Bachelor?). *J. Educ.* 171 (1), 135–146.
- Gholson, M.L., Martin, D.B., 2019. Blackgirl face: racialized and gendered performativity in mathematical contexts. *ZDM* 51 (3), 391–404. <https://doi.org/10.1007/s11858-019-01051-x>.
- Gholson, M.L., 2019. Read me last: constructing a scholarly catchment through a Black feminist reading. *Cognit. Instruct.* 37 (3), 414–421. <https://doi.org/10.1080/07370008.2019.1624551>.

- Ginsburg, H.P., Oppen, S., 1988. *Piaget's Theory of Intellectual Development*, third ed. Prentice-Hall, Englewood Cliffs, NJ.
- Greeno, J.G., 1997. On claims that answer the wrong questions. *Educ. Res.* 26 (1), 5–17. <https://doi.org/10.3102/0013189x026001005>.
- Greeno, J.G., 1998. The Situativity of Knowing, Learning, and Research. *American Psychologist*.
- Gutiérrez, K.D., 2016. Designing resilient ecologies: social design experiments and a new social imagination. *Educ. Res.* 45 (3), 187–196.
- Gutiérrez, K.D., 2018. Social design-based experiments: a proleptic approach to literacy. *Literacy Res.* 34. <https://doi.org/10.1177/2381336918787823>, 238133691878782–23.
- Hall, R., Jurov, A.S., 2015. Changing concepts in activity: descriptive and design studies of consequential learning in conceptual practices. *Educ. Psychol.* 50 (3), 173–189. <https://doi.org/10.1080/00461520.2015.1075403>.
- Hall, R., Nemirovsky, R., 2012. Introduction to the special issue: modalities of body engagement in mathematical activity and learning. *J. Learn. Sci.* 21 (2), 207–215. <https://doi.org/10.1080/10508406.2011.611447>.
- Herrenkohl, L.R., Polman, J.L., 2018. Learning within and beyond the disciplines. In: *International Handbook of the Learning Sciences*. Routledge, pp. 106–115.
- Hidi, S., Harackiewicz, J.M., 2000. Motivating the academically unmotivated: a critical issue for the 21st century. *Rev. Educ. Res.* 70 (2), 151–179. <https://doi.org/10.3102/00346543070002151>.
- Holland, D., Lachicotte, W.J., Skinner, D., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press.
- Horn, I.S., 2007. Fast kids, slow kids, lazy kids: framing the mismatch problem in mathematics teachers' conversations. *J. Learn. Sci.* 16 (1), 37–79. <https://doi.org/10.1080/10508400709336942>.
- Jackson, K., 2011. Approaching participation in school-based mathematics as a cross-setting phenomenon. *J. Learn. Sci.* 20 (1), 111–150. <https://doi.org/10.1080/10508406.2011.528319>.
- Jurov, A.S., Teeters, L., Shea, M., Steenis, E.V., 2016. Extending the consequentiality of “invisible work” in the food justice movement. *Cognit. Instruct.* 34 (3), 210–221. <https://doi.org/10.1080/07370008.2016.1172833>.
- Just, M.A., Varma, S., 2007. The organization of thinking: what functional brain imaging reveals about the neuroarchitecture of cognition. *Cognit. Affect Behav. Neurosci.* 7, 153–191.
- Kelton, M.L., Ma, J.Y., 2018. Reconfiguring mathematical settings and activity through multi-party, whole-body collaboration. *Educ. Stud. Math.* 98 (2), 177–196. <https://doi.org/10.1007/s10649-018-9805-8>.
- Kelton, M.L., Ma, J.Y., 2020. Assembling a torus: family mobilities in an immersive mathematics exhibition. *Cognit. Instruct.* 38 (3), 318–347. <https://doi.org/10.1080/07370008.2020.1725013>.
- Koedinger, K.R., Alevan, V., 2015. An interview reflection on “intelligent tutoring goes to school in the big city”. *Int. J. Artif. Intell. Educ.* 26 (1), 13–24. <https://doi.org/10.1007/s40593-015-0082-8>.
- Koedinger, K.R., Corbett, A., 2006. Cognitive tutors: technology bringing learning sciences to the classroom. In: Sawyer, K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York, NY, pp. 61–78.
- Kozulin, A., 1986. The concept of activity in Soviet psychology: Vygotsky, his disciples and critics. *Am. Psychol.* 41 (3), 264.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: aka the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lather, P., 1986. Research as praxis. *Harv. Educ. Rev.* 56 (3), 257–277.
- Lave, J., 2011. Apprenticeship in critical ethnographic practice. In: *Apprenticeship in critical ethnographic practice*. University of Chicago Press, Chicago.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Lave, J., 1982. A comparative approach to educational forms and learning processes. *Anthropol. Educ. Q.* 13 (2), 181–187.
- Leander, K.M., Phillips, N.C., Taylor, K.H., 2010. The changing social spaces of learning: mapping new mobilities. *Rev. Res. Educ.* 34 (1), 329–394. <https://doi.org/10.3102/0091732x09358129>.
- Lee, C.D., Pea, R., de Royston, M.M., 2020. Introduction: reconceptualizing learning: a critical task for knowledge-building and teaching. In: *Handbook of the Cultural Foundations of Learning*. Routledge, pp. xvii–xxxv.
- Lee, C.D., 2001. Is October Brown Chinese? A cultural modeling activity system for underachieving students. *Am. Educ. Res. J.* 38 (1), 97.
- Leontiev, A.N., 1978. *Activity, Consciousness, and Personality*. Pergamon Press.
- Lindgren, R., Johnson-Glenberg, M., 2013. Emboldened by embodiment: six precepts for research on embodied learning and mixed reality. *Educ. Res.* 42 (8), 445–452. <https://doi.org/10.3102/0013189x13511661>.
- Luria, A.R., 1976. *Cognitive Development: Its Cultural and Social Foundations*. Harvard University Press.
- Ma, J., Kelton, M., Radke, S., Della Volpe, D., 2020. Toward conceptualizing resources for learning across settings. In: Gresalfi, M., Horn, I.S. (Eds.), *The Interdisciplinarity of the Learning Sciences*, 14th International Conference of the Learning Sciences (ICLS) 2020, vol. 1. International Society of the Learning Sciences, Nashville, Tennessee, pp. 254–261.
- Marin, A., Bang, M., 2018. “Look it, this is how you know:” family forest walks as a context for knowledge-building about the natural world. *Cognit. Instruct.* 36 (2), 89–118. <https://doi.org/10.1080/07370008.2018.1429443>.
- Marin, A.M., 2020. Ambulatory sequences: ecologies of learning by attending and observing on the move. *Cognit. Instruct.* 38 (3), 281–317. <https://doi.org/10.1080/07370008.2020.1767104>.
- Marr, D., 1982. *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. Henry Holt and Co., Inc., New York, NY, USA.
- Mayer, R.E., Alexander, P. (Eds.), 2010. *Handbook of Research on Learning and Instruction*. Routledge, New York.
- Mayer, R.E., 2019. *How Multimedia Can Improve Learning and Instruction*.
- McKinney de Royston, M., Vossoughi, S., 2021. Fixating on pandemic learning loss undermines the need to transform education. *Truthout*. <https://truthout.org/articles/fixating-on-pandemic-learning-loss-undermines-the-need-to-transform-education/>. (Accessed 23 April 2022).
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theory Into Pract.* 31 (2), 132–141.
- Nasir, N.I.S., Hand, V., Taylor, E.V., 2008. Culture and mathematics in school: boundaries between “cultural” and “domain” knowledge in the mathematics classroom and beyond. *Rev. Res. Educ.* 32 (1), 187–240.
- Nasir, N.S., 2002. Identity, goals, and learning: mathematics in cultural practice. *Math. Think. Learn.* 4 (2–3), 213–247. https://doi.org/10.1207/s15327833mtl04023_6.
- Paris, D., Alim, H.S., 2014. What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harv. Educ. Rev.* 84 (1), 85–100.
- Philip, T.M., Gupta, A., 2020. Emerging perspectives on the co-construction of power and learning in the learning sciences, mathematics education, and science education. *Rev. Res. Educ.* 44, 195–217. <https://doi.org/10.3102/0091732x20903309>.
- Philip, T.M., Sengupta, P., 2020. Theories of learning as theories of society: a contrapuntal approach to expanding disciplinary authenticity in computing. *J. Learn. Sci.* 30 (2), 330–349. <https://doi.org/10.1080/10508406.2020.1828089>.
- Philip, T.M., Bang, M., Jackson, K., 2018a. Articulating the “how,” the “for what,” the “for whom,” and the “with whom” in concert: a call to broaden the benchmarks of our scholarship. *Cognit. Instruct.* 36 (2), 83–88. <https://doi.org/10.1080/07370008.2018.1413530>.
- Philip, T.M., Gupta, A., Elby, A., Turpen, C., 2018b. Why ideology matters for learning: a case of ideological convergence in an engineering ethics classroom discussion on drone warfare. *J. Learn. Sci.* 27 (2), 183–223. <https://doi.org/10.1080/10508406.2017.1381964>.
- Prather, J., Becker, B.A., Craig, M., Denny, P., Loksa, D., Margulieux, L., 2020. What Do We Think We Think We are Doing? Paper Presented at the Proceedings of the 2020 ACM Conference on International Computing Education Research.

- Rahm, J., 2014. Space-time configurations of youth-voice drive science practices: insights into local and global mobilities. *Yearbk. Natl. Soc. Stud. Educ.* 113 (2), 445–464. <http://www.tcrecord.org/library/Issue.asp?volyear=2014&number=14&volume=116>.
- Rogoff, B., 2014. Learning by observing and pitching in to family and community endeavors: an orientation. *Hum. Dev.* 57 (2–3), 69–81. <https://doi.org/10.1159/000356757>.
- Rumelhart, D.E., 1980. Schemata: the building blocks of cognition. In: Spiro, R.J., Bruce, B.C., Brewer, W.F. (Eds.), *Theoretical Issues in Reading Comprehension*. Erlbaum, Hillsdale, NJ, pp. 33–58.
- Said, E., 1993. 23 The Politics of Knowledge. *Race, identity, and representation in education*, 306. Chicago.
- Schroeder, N.L., Adesope, O.O., Gilbert, R.B., 2013. How effective are pedagogical agents for learning? A meta-analytic review. *J. Educ. Comput. Res.* 49 (1), 1–39.
- Sengupta-Irving, T., 2021. Positioning and positioned apart: mathematics learning as becoming undesirable. *Anthropol. Educ. Q.* 52 (2), 187–208. <https://doi.org/10.1111/aeq.12378>.
- Stard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27 (2), 4–13.
- Stilgoe, J., Owen, R., Macnaghten, P., 2013. Developing a framework for responsible innovation. *Res. Policy*. 42 (9), 1568–1580. <https://doi.org/10.1016/j.respol.2013.05.008>. ISSN 0048-7333.
- Stetsenko, A., 2009. Personhood: an activist project of historical becoming through collaborative pursuits of social transformation. *New Ideas Psychol.* 30 (1), 144–153. <https://doi.org/10.1016/j.newideapsych.2009.11.008>.
- Takeuchi, M.A., Ishihara, V.A., 2020. Learning to assemble the hidden bodies: embodied and emplaced mathematical literacy in transnational migrant activism. *J. Learn. Sci.* 30 (1), 1–22. <https://doi.org/10.1080/10508406.2020.1820341>.
- Taylor, H.K., 2017. Learning along lines: locative literacies for reading and writing the city. *J. Learn. Sci.* 26 (4), 533–574.
- Taylor, K.H., 2020. Resuscitating (and refusing) Cartesian representations of daily life: when mobile and grid epistemologies of the city meet. *Cognit. Instruct.* 38 (3), 1–20. <https://doi.org/10.1080/07370008.2020.1766463>.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1, 1–40. <http://bussigel.com/communityart/wp-content/uploads/2017/08/18630-43263-1-PB.pdf>.
- Vakil, S., 2018. Ethics, identity, and political vision: toward a justice-centered approach to equity in computer science education. *Harv. Educ. Rev.* 88 (1), 26–52. <https://doi.org/10.17763/1943-5045-88.1.26>.
- Vea, T., 2020. The learning of emotion in/as sociocultural practice: the case of animal rights activism. *J. Learn. Sci.* 29 (3), 311–346. <https://doi.org/10.1080/10508406.2020.1748036>.
- Velamuri, A., Gholson, M., Mendick, H., Radke, S.C., Ma, J.Y., Berge, M., Ottemo, A., Silver, E., 2021. Researching experiences of mathematics: black/feminist and queer lenses. In: Kollasche, D. (Ed.), *Proceedings of the Eleventh International Mathematics Education and Society Conference*, vol. 1, pp. 24–27.
- Vossoughi, S., Gutiérrez, K.D., 2014. Studying movement, hybridity, and change: toward a multi-sited sensibility for research on learning across contexts and borders. *Teach. Coll. Rec.* 116 (14), 603–632.
- Vygotsky, L.S., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in society: the development of higher psychological processes*. Harvard University Press.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.
- Wertsch, J.V., 1981. *The Concept of Activity in Soviet Psychology*. M. E. Sharpe, Armonk, NY.
- Wertsch, J.V., 2017. Mediated action. In: Bechtel, W., Graham, G. (Eds.), *A Companion to Cognitive Science*. Blackwell Publishing Ltd, MA, USA, pp. 518–525.
- Wise, A., Zhao, Y., Hausknecht, S., 2014. Learning analytics for online discussions: embedded and extracted approaches. *J. Learning Anal.* 1 (2), 48–71.

Learning from a behaviorist perspective

Armando Machado^a and John Staddon^b, ^a University of Aveiro, Aveiro, Portugal; and ^b Duke University, Durham, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

The beginning	12
History	12
Distinctive features	13
Habituation	13
Classical conditioning	14
Acquisition	15
Extinction	15
Emitted behavior and the form of the conditioned response	16
Operant conditioning	17
Extinction and reinforcement schedules	17
Shaping	17
Aversive control: reinforcement and punishment	17
Stimulus control: discrimination and generalization	18
Conditioned reinforcement, chains, and delay-of-reinforcement gradient	18
Choice, matching and self-control	18
Behavioral psychology and education	19
The legacy of the behavioristic approach	20
References	20
Further reading	21

The beginning

In 1913, John B. Watson (1878–1958) gave voice to a set of ideas that upended the then-dominant approaches in psychology: structuralism and functionalism. Instead of consciousness and introspected mental states, inaccessible to a third party, hence impossible to verify, Watson argued that publicly accessible *behavior* should be the proper datum for psychology. Behavioristic psychology should follow the lead of the “hard” sciences, by sticking to measurable things. It should learn from biology and look to animals as well as humans for its principles.

Watson, like B. F. Skinner after him, intended that behavior should be studied not as a cue to the “mind” or even the nervous system but, “for its own sake.” By *behavior* they meant the activities of the whole organism in context: what it is doing, what the environment in which it acts is, what its history has been, and what the effects of its actions are. To do good science means controlling all these things—history, environment, past and present, and consequences—while measuring behavior in an objective way. The hope was that given an adequate set of experimental results, laws and principles might be derived inductively and then tested, following the scientific trajectory already so successful in chemistry, biology and other natural sciences.

Proper controlled experiments, especially if they involve substantial rewards and punishments, are difficult, practically and ethically, with human subjects but possible with animals. Hence, the study of animal behavior has always been integral to the behavioristic approach: behaviorism was a direct outgrowth of animal-learning studies during the last decades of the 19th and the first decades of the 20th century. Two directions became highly influential: the research of Russian physiologist Ivan P. Pavlov (1849–1936) on conditioned reflexes with dogs (later called *Pavlovian* or *classical* conditioning); and American psychologist Edward L. Thorndike’s (1874–1949) studies of trial-and-error learning with cats (later called *instrumental* or *operant* conditioning). What followed Pavlov, Watson, and Thorndike was laboratory animal research that greatly expanded the range of our knowledge of learning, and pointed to novel practical and clinical applications.

History

Watson extended the work of Pavlov and Thorndike, revising some of their assumptions and theoretical inclinations. But, after a scandal involving his student-and-then-wife Rosalie Rayner, Watson left his faculty job at Johns Hopkins University in 1920 for a career in advertising. Laboratory work in the behaviorist tradition was carried on by two schools that developed during subsequent decades: the neo-behaviorist Clark Hull (1884–1952), at Yale, and cognitive behaviorist Edward Tolman (1886–1959), at Berkeley. The main experimental method involved maze learning in rats. Hull’s group tended to focus on simple habits. Tolman devised novel mazes and emphasized his idea of cognitive maps. The schools debated about the nature of the processes mediating between stimuli and responses. Should we regard behavior as a product of neurophysiology—look for *hypothetical constructs*—or just

settle for concepts that summarize empirical observations—*intervening variables*? Hull and his disciple Kenneth Spence (1907–1967) at Iowa, also sought out mathematical laws for learned behavior.

A third school, which became the most influential, developed after 1938 from the work of *radical behaviorist*, B. F. Skinner (1904–1990), at Harvard. Skinner distinguished reflexive behavior from operant behavior: A reflex is *elicited* by stimuli and is amenable to classical conditioning, as Pavlov’s studies revealed; operant behavior is *emitted* spontaneously by the organism and is amenable to operant conditioning through response-reward contiguity, as Thorndike’s law-of-effect anticipated. Animals and people are spontaneously active; their actions have consequences. With one singular exception Skinner’s empirical research involved the “power” of operant conditioning to affect the environment, to change it, and in turn to be changed by it. Skinner and others would converge on a consensus, conceiving of operant conditioning as the selection of behavior by its consequences, much as Darwin conceived of evolution by means of natural selection.

Skinner contributed immensely to psychology, with technical innovations and refinements (the cumulative record, Skinner box), empirical discoveries (especially reinforcement schedules, and *shaping* as a way to train novel behavior—discovered in a “day of great illumination” (Peterson, 2004)), theoretical interpretations of human public and private behavior, epistemological writings on method and theory, philosophical articulations and defenses of his behavioristic approach, and proposals to apply the findings of the experimental analysis of behavior to all aspects of human behavior including the treatment of mental health problems, the education of children, language and the “design of a culture.” Many of these speculations were severely criticized but Skinner rarely responded.

The study of learning in animals has also led to numerous applications. For example, after Watson and Rayner demonstrated in 1920 the acquisition of a phobia through classical conditioning in an 11-month-old boy and, four years later, Mary Cover Jones demonstrated in another infant the undoing of fear conditioning, Joseph Wolpe proposed systematic desensitization in the 1950s as a classical conditioning-based technique to eliminate a variety of phobias. For the next several decades, other psychologists started to use operant conditioning principles to help people in a variety of ways: to teach a child with autism to wear prescription glasses; to stop ruminative vomiting in a child with an intellectual disability; or to teach nurses to modify patient behavior in psychiatric wards, for example. Thus, behavior therapy and behavior modification were born, the former owing more to classical conditioning, the latter to operant conditioning. During the 1970s and 1980s, as behavior modification was re-labeled Applied Behavior Analysis, other researchers proposed to merge cognitive interventions aimed to change dysfunctional thoughts with behavior interventions aimed to change response patterns, a hybrid approach known as Cognitive-Behavior Modification or Cognitive-Behavior Therapy.

During the last three or four decades of the 20th century, as behaviorism started to lose its prominence in psychology, a third wave of behavioral researchers became influential, conducting research on a variety of topics (e.g., choice and self-control, concept learning, the interactions between learning and species-typical behavior), advancing quantitative models of behavior and learning (e.g., of classical and operant conditioning, timing); and developing research at the intersection of different disciplines (e.g., behavioral economics, psychopharmacology). In addition to the methodological, purposive, and radical behaviorisms of Watson, Tolman and Skinner, new directions for the philosophy of behaviorism were followed: Most prominently, teleological behaviorism (Howard Rachlin, 1935–2021), theoretical behaviorism (John Staddon, 1937–), biological behaviorism (William Timberlake, 1942–2019), and the partially-correct *Premack Principle* of David Premack (1925–2015), the idea that reinforcement is relative: a behavior which is more frequent under free conditions can reinforce a less frequent behavior when access is restricted.

Distinctive features

Despite their differences, the various behaviorisms share a family resemblance. By conceiving of behavior as the proper subject matter of psychology, behaviorists have always accorded fundamental importance to the environment. Behavior cannot be understood apart from the environment in which it occurs or has occurred, so much so that this idea may be the central tenet of behaviorism: Nothing in behavior makes sense except in the light of the environment—both past and present. (It is also true, but less emphasized, that nothing in behavior makes sense without some understanding of the reacting organism. Individual organisms inevitably differ in the ways they react to some environments.)

The subject matter of psychology is the environment-behavior interactions of animals and people and how these interactions affect subsequent behavior. The study of interaction yields functional relations between behavior and environment: behavioral analysis. The study of the cumulative effects of recurrent interactions is the study of the organism’s history: behavioral synthesis.

Behavioristic research began with three main learning processes. We describe those at length, and then say a little about other directions.

Habituation

A crab runs away from a growing shadow; a rat startles at a sudden noise; a 3-month-old baby turns her head to and looks in the direction of clicking keys. These reflexes have static properties—louder noises elicit stronger startle responses, for example. They also have dynamic properties, such as *habituation*: if the noise (*stimulus*) is repeated at regular intervals, the magnitude or likelihood of

the startle response will diminish, *habituate*. Habituation is perhaps the most primitive form of learning—“[habituation] occurs for virtually all behavioral responses in virtually all organisms” (Thompson, 2009, p. 128).

Habituation itself has several properties. Consider, for example, the blink response of a young baby to the sound of hands clapping every 2 s behind the child’s back (Humphrey, 1933, p. 142): “blinking occurs several times, but has generally died down by the six or seventh stimulation. Habituation has set in.”

- If the clapping stimulus is now withheld, blinking to the stimulus tends to *spontaneously recover* over time: the longer the wait before a stimulus recurs, the closer the response to its original value.
- If we then repeat the original series of habituation training, the baby habituates faster, a long-term effect of habituation.
- Weaker or more frequent clapping tends to engender faster and more pronounced habituation than stronger or less frequent clapping.
- Although a more frequent stimulus produces faster and more complete habituation, habituation persists longer (recovers more slowly) after a series of less frequent stimuli, i.e., habituation is *rate-dependent*.
- Once the baby is no longer blinking to the 2 s apart clapping hands, a strong stimulus (e.g., shaking the cradle) will reinstate blinking to the hand clapping, a property known as *dishabituation*.
- Finally, within the same stimulus modality, habituation tends to be specific to the training stimulus—just how specific can be determined by varying the stimulus during test trials and recording which of them elicit the response; the result is a stimulus *generalization gradient*.

Habituation is generally adaptive. By ignoring stimuli that are unimportant an animal can attend selectively to new or informative stimuli and respond to them exclusively. But ceasing to respond to a stimulus may incur costs—the new stimulus may sometimes be informative: the sound of rustling leaves may warn of a predator. The properties of habituation seem to be a balance between the benefits and costs of treating a repeated stimulus as “the same and harmless” vs. “possibly important.” For example, since the environment is less likely to change after a short than a long interval, one would expect faster habituation in the presence of shorter interstimulus intervals, the observed result. Conversely, it takes longer to detect that stimulation has ceased after wide-spaced than close-spaced stimuli stimulus series. Hence, habituation recovers more slowly after a slow stimulus series than a rapid one. A similar argument suggests that the effects of habituation should be reversible, and that spontaneous recovery should be more pronounced the longer the no-stimulus interval.

The next two sections discuss classical and operant conditioning. Because classical conditioning was discovered by a physiologist, Ivan P. Pavlov, and operant conditioning by behaviorists, Thorndike, Skinner, et al. they have historically been studied as separate processes. As we point out later, this separation now looks artificial.

Classical conditioning

Not all “neutral” stimuli are irrelevant. When a neutral stimulus is followed by a positively or negatively valued stimulus, such as when a ringing bell announces food or a cracking twig announces a predator, another elementary form of learning takes place: classical conditioning. The organism learns to use the neutral stimulus to anticipate the next one: whether it will be good or bad, where and when it will occur, how strong it will be. In the first systematic laboratory study of the process by Ivan Pavlov, a hungry dog heard a buzzer, the neutral stimulus, and a few seconds later the dog received a piece of meat, a biologically significant and positively valued stimulus. After a relatively long interval (say, 3 min), the buzzer-food sequence repeated. Initially, the buzzer elicited an orienting response, whereas the meat elicited the inborn reflexes of salivation, chewing and swallowing. After a few buzzer-food pairings, the orienting response to the buzzer had habituated and a new response emerged—the dog salivated upon hearing the buzzer. More generally, this new form of environment-behavior interaction occurs when an insignificant stimulus (buzzer) comes to elicit a response (salivation) because of its temporally contiguous pairing with a highly valued stimulus (meat). We abbreviate the account by saying that the buzzer became a conditional (or conditioned) stimulus (CS) by virtue of its pairing with food, the unconditional stimulus (US). The salivation (and much else, as we will see) is the conditioned response (CR).

Classical conditioning is involved in the acquisition of emotional responses in humans. A child bitten by a dog or simply exposed to a barking dog may fear the animal, or dogs in general, subsequently. The event constitutes a conditioning trial: The dog is the CS; the biting or barking is the US; the fear elicited by the US is the unconditional response (UR) and the fear acquired to the CS is the CR. Fear may also be acquired simply by observing another person showing intense fear in the presence of, say, a snake, an ancient enemy.

Behavioral psychologists have studied classical conditioning extensively. Using a variety of laboratory preparations, they have examined the factors that determine whether, how well and how fast a CS-US association is learned or unlearned; the kinds of associations (e.g., excitatory, inhibitory) that can be learned (easily, with some difficulty, or hardly if at all); the form and strength of the CR; or the effects of learning on other stimuli. Psychologists have also developed theories and models to explain what is learned in classical conditioning, and to predict the effect of multiple stimulus and other more complex stimulus arrangements. These evolving theories have also inspired application of classical conditioning to socially relevant problems.

Acquisition

Whether an organism will learn a CS-US association depends on multiple factors including properties of the individual stimuli and their combination (type and salience of CS and type and strength of the US), the temporal and probabilistic relations between the stimuli, and the organism's history with those and similar stimuli. Classical conditioning is anything but simple!

To illustrate, humans and rats readily associate a new taste with a subsequent gastrointestinal illness, even if the illness, the US, is delayed by an hour or more from the taste CS. But rats have difficulty associating visual stimuli even with immediate illness. In contrast, crows readily learn to associate the visual cues of a poisonous monarch butterfly with the illness it causes. Taste-aversion learning (crows included) has two distinctive features: Only one CS-US pairing is required to learn the association, and the CS-US delay may be several hours. Other forms of classical conditioning require more pairings and shorter CS-US delays.

Within a preparation, the timing of the CS and US stimuli may have major effects on the strength of the CR, on the moment this response occurs, or on whether the CS will predict the presence or absence of the US. All else equal, it is easier to associate buzzer and food when a short interval separates them. When that interval is long, the subject may be unable to learn the association. However, if the delayed pairing occurs very intermittently, so that the US is indeed more likely after the CS than in its absence, some conditioning will occur. This is called a predictive *contingency*. Moreover, if we start with a short interval and increase it as the subject learns the association, by the end of training the dog may be salivating to the buzzer despite the long interval separating it from food. This is a critical but poorly understood learning principle (Pear, 2016).

With long CS-US delays, such as when a buzzer announces food 3 min later, the subject also learns to time its response to the buzzer: The dog will salivate during the last 20 or 30 s preceding the arrival of food. But if the buzzer-food order reverses (the food precedes the buzzer), the dog is likely to learn an inhibitory association, as the buzzer now announces the absence of food.

Similar studies with humans have examined how probabilistic relations between stimuli affect our judgments of their association. In a contingency judgment task, a participant reads scenarios in which events of type A and type B (akin to CS and US) may or may not co-occur. In the food allergy paradigm, for example, A may be the ingestion of a particular food and B an allergy. Afterward, the participant judges the strength of the relation between A and B. The goal is to understand how we form contingency judgments, how we set their positive or negative direction and their strength as a function of, among other factors, the stimulus occurrences and our prior experiences with them.

The learning of a particular CS-US association may depend on the learning taking place with other stimuli. For example, if food is already anticipated by a buzzer due to previous training, adding a light to the buzzer just before each new episode of food will not change the dog's behavior toward the light. This is called *blocking*. (There is a related effect called *overshadowing*: when two stimuli one highly salient, the other not, have the same predictive relation to the US, conditioning may occur only to the more salient stimulus, even though conditioning would have occurred to the less salient, had it been presented alone.)

Going one step further, a stimulus may become a CS (and elicit a new response) even without being paired with a US (e.g., food): pairing it with another CS may suffice. In the prototypical setting, a dog may learn to salivate to the sound of a metronome if the metronome is paired systematically with a previously conditioned, food-associated buzzer. In this case, the buzzer is a first-order conditional stimulus (CS1) because it is paired with food (US), and the metronome is a second-order conditional stimulus (CS2) because it is paired with a first-order CS. Although the CS2 is never paired with food, it nevertheless elicits salivation (*second-order conditioning*).

Second- and higher-order conditioning explain partly how we learn to like or dislike objects. Pairing an abstract, geometric drawing with an attractive picture (e.g., a pretty face) or a nice word (e.g., "honest"), for example, tends to change the drawing's valence from neutral to positive; pairing it with a threatening or disgusting picture, or with an unpleasant word such as "arrogant", tends to change its valence from neutral to negative. These examples of evaluative conditioning, used abundantly in advertising, seem to involve the transfer of value from first-order CSs (the positive or negative pictures or words) to second-order CS (the abstract drawing). Our impressions of a person may also depend on classical conditioning (e.g., when, on first meeting John, we hear a mutual friend describe him as "friendly"). First associations may even be more resistant to change than subsequently learned ones.

Extinction

Once an association is learned, how can it be unlearned? Let us return to the prototypical example. If the buzzer is presented without the food, the dog will salivate less and less to the buzzer and eventually will stop salivating to it. Presenting a CS without the US attenuates the CR and may even eliminate it, a procedure known as extinction. The procedure immediately suggests a way to treat nonadaptive emotional responses elicited by specific stimuli. Expose the subject to the phobic stimulus and ensure that no aversive event occurs during the exposure. In one variation of the procedure, *flooding*, the subject faces a strongly feared stimulus or a series of increasingly more fearful stimuli. In another, in-vivo systematic desensitization, the gradual exposure to more fearful stimuli occurs while the subject is in a relaxed state. Extinction-based exposure techniques are key ingredients of the most effective psychotherapies for anxiety.

But extinction of the CS-US association does not seem to be the simple undoing or reversing of the acquisition of the CS-US association. Part of the evidence is that the effects of extinction are often restricted to the extinction context—the buzzer that no longer elicits salivation in the extinction context may readily elicit salivation in a new context. Similarly, reencountering the US in the absence of the CS—the dog receives food in the absence of the buzzer—may suffice to reinstate salivation to the buzzer. Even the mere passage of time without further occurrences of buzzer or food may recover the response to the buzzer. These renewal,

reinstatement, and spontaneous recovery effects suggest (but the issue is far from settled) that extinction may involve the learning of inhibitory and context-dependent CS-US associations. Be that as it may, the study of extinction is helping applied psychologists understand the relapse of emotional reactions after their successful treatment and how to prevent them. As Vervliet et al. (2013) put it, learning theory “provides a continuous source of inspiration for new developments in the field of fear extinction and return” (p. 219).

Emitted behavior and the form of the conditioned response

Because classical conditioning was first studied in a physiology lab, the focus was on the autonomic response of salivation. But the procedure itself actually leads to much broader effects. In fact classical and operant conditioning are inseparable: classical conditioning, which has effects on both autonomic and somatic behaviors, provides the repertoire from which contingent *reinforcement* (i.e., reward produced by, and usually immediately following a response) can select.

Operant behavior is *emitted* according to processes never well defined, but presumably including the motivational state of the organism, its past history and current environment. Contingent reinforcement then selects from this emitted repertoire. Respondent behavior, on the other hand, is often spoken of as *elicited* by a neutral stimulus (the CS) that predicts an unconditioned stimulus (US) such as food (which can act as a reinforcer in an operant situation).

The distinction between “emitted” and “elicited” corresponds to no observable property: a rapid learned response may be as swift as an eyeblink conditioned reflex. But the prototypical reflex response was Pavlovian salivation, run by the autonomic nervous system, as opposed to operantly conditioned pecking, run by the somatic system. These differences are presumably what led Skinner to characterize behavior conditioned in the Pavlovian way (classical conditioning) as *respondent* and behavior conditioned by contingent consequences as *operant*—to treat them as two quite separate classes. In the same way, the field of animal learning as a whole was pretty clearly divided into operant vs. classical conditioners.

The distinction was challenged in 1968 when Brown and Jenkins published a paper on what they called *autoshaping*, although this was not clear at the time. Even though the procedure was purely Pavlovian, a 7 s light followed by food reward presented every minute or so, naïve pigeons swiftly learned to peck the light. Yet these anomalous results were at first seen in terms of the old paradigm, as the outcome of familiar operant processes: “[O]ur account essentially parallels the accounts offered by Skinner (1948) and by Morse and Skinner (1957) for the first and second types of superstitious behavior.”

The result was indeed Pavlovian conditioning, but not of salivation but *pecking* of the lighted key. These findings were soon corroborated by others, but none emphasized the obvious contradiction between the autoshaping result and the operant-respondent distinction. Here was pecking, the favored Skinnerian operant, behaving just like a respondent.

There had been earlier hints that a Pavlovian procedure can induce operant behavior. For example, a visitor to Pavlov’s lab noticed that if a trained dog is out of its usual harness, his first response to the conditioned stimulus was “the whole system of behavior patterns serving, in a number of *Canidae*, to beg food from a conspecific. It is, in fact, this whole system that is being conditioned in the classical experiment” i.e., operant not respondent behavior. The dog had developed an *expectation*, which implies a repertoire of food-related activities including, but not limited to, salivation.

The essential identity of the processes underlying operant and Pavlovian conditioning got additional support from a replication (Staddon and Simmelhag, 1971) of Skinner’s famous Pavlovian “superstition” experiment, which simply involved giving a hungry pigeon a bit of food every 15 s (*temporal conditioning*). Skinner reported the development of vigorous stereotyped behavior that Skinner interpreted as the result of *adventitious* (accidental) *reinforcement* of random behavior by the pigeon. When the experiment was repeated, but with more detailed temporal recording of the resulting behavior, the authors found at least two kinds of induced behavior: the *terminal* response, toward the end of the interfood interval, which in this case was just pecking, usually at the chamber wall above the feeder aperture, since there was no stimulus key—and what was called *interim* behavior, other activities that occurred earlier in the 12 s interfood interval. The interim behaviors were rarely contiguous with food, hence not operant in the conventional sense. And the terminal response was clearly operant.

A still-developing resolution of the problem is abandonment of both the operant-respondent dichotomy and the stimulus-response framework. For stimulus substitute *situation*, and for response substitute the *state* of the organism, where the state defines not a single response but a *repertoire* of potential actions. The range of behavioral variation depends on how “hot” (predictive of reinforcement) the situation is, and how the organism is constrained. At the high end, the repertoire is limited, and the organism displays rigid behavior typically associated with classical conditioning. In the middle we see a wide range of “emitted” behavior, available for selection by contiguous reinforcement. At the low end, the repertoire is again limited until the situation looks more promising. The shift from rigid stimulus-response links to the more Darwinian selection of a response from a repertoire associated with the state of the organism, controlled by its history and present situation, seems to represent a more realistic view of how learning works—in animals and perhaps in humans as well. The distinction between classical and operant conditioning remains, not as two separate processes but simply as labels for two different procedures.

Classical conditioning seems to be involved in problems of drug addiction. The direct drug effects (e.g., analgesia, hypothermia) constitute the US. The responses that the body produces to compensate for these effects constitute the URs. They have a regulatory or homeostatic function, to preserve the “internal milieu” in the presence of the severe disruption caused by the US. The cues accompanying the primary drug effects (e.g., the injection, the location) are the CSs; after a few pairings with the US, these CSs elicit the compensatory CRs (e.g., hyperalgesia, hyperthermia). These CRs contribute to chronic drug tolerance, the fact that a specific drug dose becomes less effective with repeated administrations of the drug, a kind of habituation. To experience the original drug effect,

the user increases the dose, but a higher dose also conditions stronger compensatory responses, which again reduce the drug effects. The process may spiral up and the user may reach a high dose that would have been fatal if consumed on the first day, without the benefit of the compensatory responses. In fact, at this moment the user may be at a higher risk of death if he consumes the drug in a new environment, one with few or no cues associated with the drug administration and therefore without the protective effect of the conditioned compensatory responses (a kind of dishabituation). Moreover, if drug-associated cues are present but the drug is not consumed, the compensatory responses themselves will not be attenuated by the drug effects. The full expression of the compensatory CRs contributes to the withdrawal symptoms (Siegel, 2008, 2016). Research on classical conditioning has direct implications for the treatment of addictions—finding ways to minimize conditioned reactions to a variety of cues, for example.

Operant conditioning

Operant conditioning is the way that most “higher” animals learn. The prototypical operant conditioning experiment is a hungry rat that is rewarded (reinforced) with food for pressing a lever. It may have learned by just emitting behavior at random, until it presses the lever and gets fed. Or it may have been *shaped* by being rewarded when it gets close to the lever. Or, like the autoshaped pigeon, the desired behavior may have been produced by classical conditioning, all in the service of facilitating a repertoire of behavior, a state, of which the target behavior is a member.

Extinction and reinforcement schedules

If a learned operant response is no longer reinforced—lever pressing no longer yields food—response rate decreases. This extinction process is often accompanied by emotional responses (e.g., biting the lever, pressing it harder). Between the extremes of continuous reinforcement, with each response reinforced, and extinction, with no response reinforced, lies partial or intermittent reinforcement, with only some responses reinforced. The rule specifying which responses will be reinforced defines a *reinforcement schedule*. Two examples are *variable ratio* schedules in which each response has a fixed probability of payoff, and *fixed interval* schedules, in which a response is reinforced provided a fixed interval of time has elapsed since the previous reinforcer.

Much operant conditioning research is devoted to the effects of reinforcement schedules on the rates and probabilities of responding, on the distribution of responses across time, and on the resistance of responses to extinction. For example, moderate-size variable ratio schedules (analogous to the gambler’s one-armed bandit device) generate vigorous, sustained rates of responding and great resistance to extinction.

Shaping

A “free” reinforcer may tell the organism it is in a food situation so it may emit food-related behaviors. But if one of those is reinforced, the animal’s repertoire is unlikely to increase further. How then do animals and people learn novel and sometimes complex behaviors (e.g., the choreographies performed by dogs with their trainers; the acrobatic moves of dolphins at water parks)? The usual method is *shaping*, the differential reinforcement of successive approximations to a target behavior. In each stage, a behavior closer to the target response than in the previous stage is reinforced while other behaviors are not.

Shaping may sometimes occur naturally. Field observations have shown that some rats living along the banks of the Po River in Italy dive for and feed on mollusks inhabiting the river bottom. This unusual behavior may have developed because wide, natural fluctuations in the water level first expose the mollusks and then gradually resubmerge them, creating the conditions for the rats to learn to dive for food by natural shaping (Galef, 1980). Another case of natural shaping may involve the female cowbird’s reaction (the reinforcer) to young male cowbirds singing. Wing stroking by females correlated with faster rate of song development in males; it was also related to differences in song potency. This non-vocal shaping process may be like the processes that influence phonological development in human infants (King et al., 2005; Goldstein et al., 2003).

Shaping is common in everyday human life, as children learn to dress, swim, write the alphabet, or read, for example. It has been used extensively in applied settings to teach rudimentary language skills to autistic children, appropriate behavior to severely psychotic patients, and complex performances to professional athletes. Shaping is also used to train cooperative behaviors in zoo animals.

Aversive control: reinforcement and punishment

Operant behavior increases following a reinforcement and decreases in probability after an *aversive stimulus: punishment*. There are two types of punishment: either an aversive stimulus is presented to the subject (e.g., a reprimand given to a child), or a reinforcer is taken away (e.g., the cell phone is taken away from a teenager). In the laboratory, by adjusting the parameters of the punisher (e.g., its intensity, delay, schedule), behavioral psychologists have identified several regularities (e.g., immediate is more effective than delayed; after every unwanted response is more effective than occasionally, etc.). B. F. Skinner often declared that punishment is an ineffective way to control human behavior. But the available animal research does not support this conclusion. More needs to be known, but unfortunately the laboratory study of punishment has almost ceased, for a combination of ethical and regulatory reasons.

Stimulus control: discrimination and generalization

Operant conditioning stresses behavior-outcome relations, but these relations typically hold only under specific conditions. Pecking a colored key produces food only if the key is this color in this Skinner box. A stimulus in the presence of which a response is likely to be reinforced is termed a *discriminative stimulus* (S^D). The fox is likely to obtain a meal by chasing the rabbit (the S^D) but not by chasing the hound; the child is likely to receive a smile by saying “daddy” in front of her father (the S^D) but not to a stranger. The fox and the child have learned a stimulus discrimination, to respond differently in the presence of different stimuli. Of course, the fox chases not just a single rabbit but a class of rabbits (e.g., white, gray, or black); chasing shows stimulus generalization. In fact, a finer analysis would certainly reveal that the likelihood of chasing the rabbit, or the vigor of the chase, vary with other stimulus attributes (e.g., color, size, and distance of the rabbit).

To study how a neutral stimulus may come to control a given response, that is, how an S becomes an S^D , researchers in the laboratory alternate the reinforcement of that response in the presence of S with the extinction of the response in the absence of S . In a Skinner box, a rat receives food pellets for pressing a lever when a light is on but not when the light is off, with each stimulus condition alternating every 3 min. If we peek into the Skinner box after a few days, we will see that when the light is turned off the rat engages in behaviors other than pressing the lever; when the light is turned on, the rat mainly presses the lever and eats the food.

By varying the stimulus, the response, the reinforcement schedules, and a host of other variables, behavioral psychologists have found many regularities concerning the acquisition and maintenance of stimulus control, how organisms attend and respond to selective features of their environment, how they generalize, and how they learn complex discriminations and discriminative repertoires. Their techniques and findings have been used to investigate a broad range of issues, ranging from animal psychophysics (e.g., do animals also have visual illusions?) and behavioral neuroscience (e.g., what is the role of neurotransmitter dopamine in temporal discriminations?) to problems of applied educational, organizational, and clinical psychology (e.g., how to decrease off-task behaviors in the classroom, train complex skills in a factory, or treat self-injurious behaviors).

Conditioned reinforcement, chains, and delay-of-reinforcement gradient

A reinforced response is always learned in a particular stimulus situation (the *discriminative stimulus*, S^D); this is Skinner’s *three-term contingency*. If the situation is changed, the response will not occur, or will be weaker. Not surprisingly, the S^D often becomes a conditioned or secondary reinforcer, a stimulus that can be used to strengthen behavior.

Chains of responses may thus be learned, with one and the same stimulus serving as the reinforcer of the response that precedes it and as the discriminative stimulus for the response that follows it. The actual causation is more complex, though: the reinforcing property of a stimulus weakens the farther back, in time or in the chain it is from the primary reinforcer. Chaining has been used in many applied settings. Recently, with a clicker in the role of conditioned stimulus, chaining was used successfully to teach complex procedural surgical skills (e.g., tying knots, making low angle drilling holes) to medical students (Levy et al., 2016).

The initial link of the chain announces a delayed reinforcer at the end of the chain. The teacher instructs the student to solve the first five homework problems before going to recess. Whether the child attends to the initial-link stimulus, the instruction, and whether she continues to respond throughout the delay as other stimuli occur, may depend on the child’s delay-of-reinforcement gradient, that is, on how rapidly reinforcement value, V , decreases with delay, D . Many studies have yielded results consistent with the function

$$V(D) = \frac{V_0}{1 + kD}, \quad k > 0$$

where V_0 is the value of the reinforcer without delay and k is a delay discounting parameter. A steep gradient (large k) may mean that the stimuli at the beginning of the chain may not reinforce adequately the behavior of attending to them, and the child’s attention will wander off task. Moreover, steep gradients may also select rapid responding over slower responding because only when behavior occurs rapidly will the final reinforcer affect the beginning of the action. In addition to attention problems, steep gradients may engender hyperactive behavior, as in children with attention-deficit hyperactivity disorder (Sagvolden et al., 2005).

Choice, matching and self-control

Reinforcers vary along multiple dimensions. A primary reinforcer such as food, for example, varies along quality, amount, probability, delay, and other dimensions. Discovering how each dimension affects behavior, separately and in combination, is a difficult but critical task if we are to understand naturally occurring behavior. Suppose we have a choice between two behaviors, B_1 and B_2 (e.g., dining at Chinese or Italian restaurants; in a classroom, engaging in on-task or off-task behaviors; calling John or Mary on the phone for a chat). Behaviors B_1 and B_2 produce reinforcers at rates R_1 and R_2 , respectively (how often we experience excellent meals at each restaurant, receive attention from classmates from each behavior, or have comforting conversations with each friend). What determines how we choose or, more specifically, how often we engage in one or the other behavior? Depends on the schedule. A simple choice amounts to a fixed-ratio schedule and most organisms respond optimally: they pick the best option exclusively. On the other hand, as economists and biologists well know, most reinforcers *satiates*: the more you get the less attractive each increment becomes. So, even when the schedule is FR1, subjects will often switch from the option initially favored after a while.

The same result can be obtained as a steady state if the two choices are available on variable-interval schedules (*concurrent VI VI*), since the probability of payoff for the disfavored choice, hence its attractiveness relative to the favored choice, increases with time away. Much research has been done with this rather complex procedure because the functional relations obtained by varying the values of the VIs associated with each choice are both simple and readily reproducible.

Given two choices, $B1$ and $B2$, the usual steady-state result is the so-called generalized matching law: The ratio of responses in a typical experimental session of an hour or two, $B1/B2$, is a power function of the ratio of obtained reinforcers, $R1/R2$:

$$\frac{B1}{B2} = b \left(\frac{R1}{R2} \right)^s, \quad 0 < s < 1$$

where s is a sensitivity parameter and b is a bias parameter. Given another procedural feature called a change-over delay, which prevents switches from being immediately reinforced, s and b tend to approach 1, which is Richard Herrnstein's initial formulation of his *matching law*.

But suppose the reinforcers for behaviors $B1$ and $B2$ vary along the dimensions of amount and delay, in fact, that they trade amount for delay: $B1$ yields more reinforcers than $B2$, but those reinforcers are also more delayed. All else equal, we know that more is better than less and immediate is better than delayed, but which is better, more/delayed or less/immediate? And how does value—and hence choice—change as the delays elapse? These questions are at the heart of research on self-control and impulsivity. In a typical laboratory situation, a pigeon may choose between 6 s of access to food (large reward) but delayed for 10 s, and 2 s access to food (small reward) delivered after 2 s. Similarly, a child may choose between two marshmallows after a 5 min waiting period, or one marshmallow immediately. Self-control is equated with the choice of the large-later reward, impulsivity with that of the small-sooner reward. Psychologists have studied empirically how choice depends on the magnitudes of the options, the choice context, prior learning experiences (e.g., learning what to do during the waiting period), and the species and developmental stage of the organism. They have also developed theories and models to predict preference for the self-control or the impulsive options, the reversal of preference from the former to the latter, and how to prevent such reversals in applied settings. The conc VI VI procedure is undoubtedly a useful tool for measuring the behavioral effects of these parameters. But because of the negative feedbacks built in to the concurrent VI VI procedure, the extent to which the beautiful regularities it yields are due to the organism and how much to the procedure, are still unclear.

Behavioral psychology and education

One of the distinctive features of behavioral psychology is its historical approach to the explanation of behavioral patterns. Instead of moving in space, from without to within the organism in search of inner cognitive or neurophysiological processes that generate behavior, the behavioral approach moves in time, from the present to past behavior-environment interactions in search of ways to understand how current behavior resulted from the integration of the effects of recurrent causes, each perhaps with a small effect, operating over long periods of time. The behavioral explanation is not unlike the historical explanation of an adaptive biological structure (e.g., the vertebrate eye) in terms of the cumulative effects of natural selection acting consistently over long stretches of time.

A correct behavioral account brings with it the immediate benefit of identifying how to implement specific kinds of behavior-environment interactions to promote desirable behaviors and reduce or prevent undesirable behaviors. A case in point is the Good Behavioral Game (GBG), a remarkably simple behavioral strategy that implements group-based contingencies to reduce disruptive classroom behaviors and promote academic engagement and prosocial behaviors. In its simplest version, teacher and students identify the rules that would make the setting a good place to learn, more enjoyable and pleasant, as well as the rules or "fouls" that would interfere with the desirable outcomes. The game is played at specified intervals (e.g., math class, reading class). The group is divided into two or more teams, and the team with the least number of fouls wins the game; alternatively, any team with the number of fouls below a criterion wins the game (this reinforcement rule is a group variant of a schedule known as differential reinforcement of low rates of behavior). A highly visible scorecard shows daily the fouls and wins of each team. Additional features of the game apply to different contexts and populations (e.g., students with emotional disorders, undergraduate students, preschool children).

Developed in 1969 by Barrish, Saunders, and Wolf, the GBG has been studied extensively since then. Its positive effects on student behavior were immediate. Other positive effects, however, were revealed only in long-term follow-up studies. In 2002, Embry remarked: "[The Game] worked pretty well," and became a behavior-modification "trick" most graduate students in behavior analysis or special education learned during the heyday of behavioral psychology. Neither Barrish, Saunders, or Wolf, nor the graduate students who learned to use the Game as a classroom strategy, had the slightest idea then how powerful the strategy might be for changing the future of children destined for multiple lifetime problems of substance abuse, violence, and school failure. In 2011, Kellam et al. summarized the results of a large-scale randomized control trial study that tested the game with first- and second-grade children with a follow-up at ages 19 to 21: "significantly lower rates of drug and alcohol use disorders, regular smoking, antisocial personality disorder, delinquency and incarceration for violent crimes, suicide ideation, and use of school-based services among students who had played the GBG. Several replications with shorter follow up periods have provided similar early results" (p. 73).

The legacy of the behavioristic approach

Behaviorism was the dominant school of psychology until about the last quarter of the 20th century. Its demise has been reported multiple times. On the other hand, no field that is still divided into “schools” can claim to be a mature science: there are no schools of physics or chemistry, or even biology. In fact all scientific psychologists are behaviorists of a sort, for they all share a commitment to objective methods and measurements—to draw inferences about learning and memory, for example. The difference is cumulativeness and the reality of theoretical entities. As we have described, behaviorism has progressed. We know more now about reinforcers and punishers and how they work, the complementary relation between classical and operant conditioning, why reinforcement schedules have the effects they do. Much more research is always needed, but the direction of advance is clear. Unfortunately the variants of cognitive psychology have not progressed in this way. Instead they have adopted questionable experimental methods (the replication crisis is an almost inevitable product of the almost-universal Null Hypothesis Significance Testing group-experimental method) and settled prematurely on the digital computer as a model for the mind. With a few exceptions, such as the early work of Jean Piaget, the nominal fallacy is rampant: experimental results in very restricted situations are said to measure things like self-efficacy, craving, self-enhancement, confirmation bias, etc., whereas real scientific concepts cannot just be labeled into existence but must be derivable from a theory with many testable predictions.

There is a deep irony here. On the one hand, psychology remains a fragmented science, more so with each passing decade. Currently, the American Psychological Association lists 54 divisions, a name befitting their effect. Anthropologist Geertz (1997) described the view from the outside: “[Psychology] looks like an assortment of disparate and disconnected inquiries classed together because they all make reference in some way or other to something or other called ‘mental functioning.’ Dozens of characters in search of a play” (p. 22).

On the other hand, the applications of behavioristic psychology remain unrivaled, even after four or five decades of cognitive psychology. De Houwer and Hughes (2020) summarize, after covering the main concepts of classical and operant conditioning:

[These concepts] (...) form the core of treatment programs that are dramatically improving the lives of millions of children and adults living with developmental and intellectual disabilities (...). They are shaping people into better parents, teachers, colleagues, and friends (...). They are being used to radically cut drug and alcohol use in people struggling with addictions (...). Concepts and procedures derived from the psychology of learning are the beating heart of many clinical therapies used to treat phobias, panic and anxiety disorders, and chronic pain—related issues, as well as psychosis, depression, and substance abuse (...). Those same ideas are deployed to train disobedient pets and improve the welfare of zoo animals and livestock (...), address health behaviors (...), help athletes improve their performances and pathological gamblers beat their addictions (...), promote recycling and reduce pollution (...), and much, much more (...). The takeaway message is that the procedures, effects, and principles [of classical and operant conditioning] ... drive human behavior and can help us save lives and improve well-being (p. 201).

If psychology is ever to be a science, if it is to be of practical use, it will be behavioristic.

References

- Barrish, H.H., Saunders, M., Wolf, M.M., 1969. Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *J. Appl. Behav. Anal.* 2 (2), 119–124. <https://doi.org/10.1901/jaba.1969.2-119>.
- Brown, P.L., Jenkins, H.M., 1968. Auto-shaping of the pigeon's key-peck. *J. Exp. Anal. Behav.* 11, 1–8.
- De Houwer, J., Hughes, S., 2020. *The Psychology of Learning: An Introduction From a Functional-Cognitive Perspective*. MIT Press.
- Embry, D.D., 2002. The Good Behavior Game: a best practice candidate as a universal behavioral vaccine. *Clin. Child Fam. Psychol. Rev.* 5, 273–297. <https://doi.org/10.1023/a:1020977107086>.
- Galef, B., 1980. Diving for food: analysis of a possible case of social learning in wild rats (*Rattus norvegicus*). *J. Comp. Physiol. Psychol.* 94, 416–425.
- Geertz, C., 1997. Learning with Bruner. *N. Y. Rev. Books* XLIV (6), 22–24.
- Goldstein, M.H., King, A.P., West, M.J., 2003. Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8030–8035.
- Humphrey, G., 1933. *The Nature of Learning in Its Relation to the Living System*. Kegan Paul, Trench, Trubner.
- Kellam, S.G., Mackenzie, A.C., Brown, C.H., Poduska, J.M., Wang, W., Petras, H., Wilcox, H.C., 2011. The good behavior game and the future of prevention and treatment. *Addiction Sci. Clin. Pract.* 6 (1), 73–84.
- King, A., West, M., Goldstein, M.H., 2005. Non-vocal shaping of Avian song development: parallels to human speech development. *Ethology* 111, 101–117.
- Levy, I.M., Pryor, K.W., McKeon, T.R., 2016. Is teaching simple surgical skills using an operant learning program more effective than teaching by demonstration? *Clin. Orthop. Relat. Res.* 474 (4), 945–955. <https://doi.org/10.1007/s11999-015-4555-8>.
- Morse, W.H., Skinner, B.F., 1957. A second type of superstition in the pigeon. *Am. J. Psychol.* 70, 308–311. <https://doi.org/10.2307/1419345>.
- Pear, J., 2016. *The Science of Learning*, second ed. Routledge.
- Peterson, G.B., 2004. A day of great illumination: B. F. Skinner's discovery of shaping. *J. Exp. Anal. Behav.* 82, 317–328. <https://doi.org/10.1901/jeab.2004.82-317>.
- Sagvolden, T., Johansen, E.B., Aase, H., Russell, V.A., 2005. A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav. Brain Sci.* 28 (3), 397–468. <https://doi.org/10.1017/S0140525X05000075>.
- Siegel, S., 2008. Learning and the wisdom of the body. *Learn. Behav.* 36, 242–252.
- Siegel, S., 2016. The heroin overdose mystery. *Curr. Dir. Psychol. Sci.* 25 (6), 375–379.
- Skinner, B.F., 1948. 'Superstition' in the pigeon. *J. Exp. Psychol.* 38 (2), 168–172. <https://doi.org/10.1037/h0055873>.
- Staddon, J.E.R., Simmelhag, V., 1971. The “superstition” experiment: a reexamination of its implications for the principles of adaptive behavior. *Psychol. Rev.* 78, 3–43.
- Thompson, R.F., 2009. Habituation: a history. *Neurobiol. Learn. Mem.* 92, 127–134. <https://doi.org/10.1016/j.nlm.2008.07.011>.
- Vervliet, B., Craske, M.G., Hermans, D., 2013. Fear extinction and relapse: State of the Art. *Annu. Rev. Clin. Psychol.* 9, 215–248.
- Watson, J.B., 1913. Psychology as the behaviorist views it. *Psychol. Rev.* 20, 158–177.

Further reading

Experimental science: introductory

Catania, C., 2006. *Learning*, fourth ed. Sloan Publishing.
Mazur, J., 2016. *Learning & Behavior*, eighth ed. Routledge.

Experimental science: advanced

Staddon, J.E.R., 2016. *Adaptive Behavior and Learning*, second ed. Cambridge.

Behaviorism

Rachlin, H., 2014. *The Escape of the Mind*. Oxford.
Staddon, J.E.R., 2021. *The New Behaviorism: Foundations of Behavioral Science*, third ed. Routledge.

Habituation

Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J., Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., McSweeney, F.K., Wilson, D.A., Wu, C.F., Thompson, R.F., 2009. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.* 92 (2), 135–138. <https://doi.org/10.1016/j.nlm.2008.09.012>.
Staddon, J.E.R., Higa, J.J., 1996. Multiple time scales in simple habituation. *Psychol. Rev.* 103, 720–733.

Education and good behavior game

Bowman-Perrott, L., Burke, M.D., Zaini, S., Zhang, N., Vannest, K., 2015. Promoting positive behavior using the good behavior game: a meta-analysis of single-case research. *J. Posit. Behav. Interv.* 18, 180–190. <https://doi.org/10.1177/1098300715592355>.
Joslyn, P.R., Donaldson, J.M., Austin, J.L., Vollmer, T.R., 2019. The good behavior game: a brief review. *J. Appl. Behav. Anal.* 52, 811–815. <https://doi.org/10.1002/jaba.572>.
Vargas, J.S., 2013. *Behavior Analysis for Effective Teaching*, second ed. Routledge.

Learning from a social cognitive theory perspective

Dale H. Schunk and Maria K. DiBenedetto, The University of North Carolina at Greensboro, Greensboro, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Background	23
Historical influences	23
Imitation	23
Latent learning	23
Social learning	23
Early modeling studies	23
Conceptual framework	24
Reciprocal interactions	24
Enactive and vicarious learning	25
Vicarious, symbolic, and self-regulatory processes	25
Vicarious processes	25
Functions of modeling	25
Observational learning	26
Educational research	26
Symbolic processes	27
Self-efficacy	27
Goals and self-evaluations of progress	28
Goal properties	28
Outcome expectations	28
Values	28
Social comparisons	29
Educational research	29
Self-regulation	29
Phases of self-regulation	30
Social-to-self progression	30
A self-regulated framework	30
Educational research	31
Contributions	31
Educational implications	32
Models	32
Goals and self-evaluations of progress	32
Self-efficacy for learning	32
Self-regulation	32
Future research directions	33
Methodological refinements	33
Cross-cultural relevance	33
Integration with technology	33
Conclusion	33
References	34

Contemporary theories of learning predict that learners are active seekers and processors of information (Mayer, 2012). They do not passively absorb information from the environment but rather proactively develop new knowledge by influencing what, where, why, when, and how they learn.

Social cognitive theory reflects this proactive self-regulatory emphasis. Social cognitive theory predicts that much human learning occurs in social environments. By interacting with others live or virtually, people learn knowledge, skills, strategies, beliefs, rules, and attitudes. They also learn about the appropriateness, usefulness, and consequences of behaviors. People act in accordance with their beliefs about their capabilities and the expected outcomes of actions.

Although there are different social cognitive perspectives, this chapter focuses on the theoretical perspective formulated and developed by Bandura (1977b, 1986, 1997, 2001). This theory highlights the key roles played by vicarious, symbolic, and self-regulatory processes, in the learning and performance of actions.

The next section describes historical influences on social cognitive theory and Bandura's early studies investigating social modeling. Then the conceptual framework of Bandura's theory is discussed. Subsequent sections describe the key roles played

by vicarious, symbolic, and self-regulatory processes. Some educational implications of the theory and research are given, and suggestions for future research are provided. It is hoped that this chapter will promote research on and development of social cognitive theory as a perspective on learning.

Background

Historical influences

There are several historical influences on Bandura's social cognitive theory. Three noteworthy ones are imitation, latent learning, and social learning.

Imitation

Historical perspectives on imitation postulated that people were naturally inclined to imitate the actions of others (James, 1890). McDougall (1926) believed that observing actions of others elicited an instinct in observers to copy those actions. Early behaviorists (e.g., Watson, 1924) dismissed this instinct notion because it assumed the existence of an inner drive that intervened between a stimulus (observed action) and response (copying that action).

Behaviorists initially postulated that imitation was conditioned. Skinner (1953) predicted that imitation represented a generalized response class that could be explained using a three-term contingency: discriminative stimulus → response → reinforcing stimulus. An observed action was the discriminative stimulus. Imitation occurred when an individual performed that action (response). If the response was reinforced (reinforcing stimulus), this reinforcement increased the future likelihood of imitating the actions of others. Because imitation was reinforced by different people under various conditions, the act of imitating became a generalized response class.

This view assumes that individuals can perform the imitated response. Yet much research shows that various types of behaviors can be learned by observing models (Rosenthal and Zimmerman, 1978). Another issue is that for imitation to occur it must be reinforced. Researchers have shown that modeling can occur without any external reinforcement (Bandura and Walters, 1963).

Latent learning

Latent learning is learning that occurs without reinforcement. In a seminal study, Tolman and Honzik (1930) let rats wander through a maze. Some rats always were fed in the maze and others were never fed. Rats fed in the maze quickly reduced their time and errors in running the maze, whereas time and errors for the other group did not diminish. Later, some rats from the nonreinforced group were given food for running the maze. Their time and number of errors quickly decreased to the levels of those who always had been fed, whereas the time and errors for rats who continued not to be reinforced did not change. Thus, rats in the nonreinforced group learned features of the maze by running through it without being reinforced. When food was given, latent learning was displayed.

Latent learning research showed that learning can occur when not displayed at the time of learning and that reinforcement affected performance of behaviors rather than learning. This distinction between learning and performance is central to Bandura's social cognitive theory. Because latent learning research involved animals, it could not explore such cognitive processes as goals and expectations, and not fully establish the potential of latent learning.

Social learning

Rotter (1954) developed a social learning theory of personality that was applied to learning. The theory included four primary variables: behavior potential, expectancy, reinforcement value, and the psychological situation. Behavior potential referred to the probability that an individual would act in a given way relative to alternative behaviors. This potential depended on the individual's expectancy (belief about the likelihood of reinforcement occurring as a consequence of behavior), and on reinforcement value (how much the individual valued a particular outcome relative to other potential outcomes). The psychological situation referred to the context of behavior. Through social interactions with others, people formed expectancies about the likely outcomes of actions, and they acted in accordance with their expectancies and the value they placed on the potential outcomes.

Rotter's theory represented an important advance over behavior theories because it showed that people act based on what they believe rather than in accordance with prior reinforcement. When beliefs contradict prior experience (e.g., a behavior has been reinforced but now people believe it will not be), people follow their beliefs.

Rotter's theory, however, only considered beliefs about potential outcomes of actions. Bandura's social cognitive theory includes these outcome expectations but also postulates influential roles for other cognitive beliefs including goals, self-efficacy, and self-evaluations of learning and goal progress (discussed later in this chapter).

Early modeling studies

Early studies by Bandura and his colleagues contributed to the development of contemporary social cognitive theory (Bandura, 1977b). In a classic demonstration of observational learning, Bandura et al. (1963) exposed children to live aggressive models, filmed aggressive models, aggression by cartoon characters, or no models. The models pummeled a Bobo doll by hitting, throwing, kicking, and sitting on it. Models were not rewarded or punished for their actions. Children subsequently played with a Bobo doll. Compared with no-model control children, children exposed to aggressive models displayed significantly higher levels of

aggression. Type of aggressive model (i.e., live, filmed, cartoon) made no difference in children's levels of aggression. The results suggested that children learned from observing models and that the consequences of modeled actions did not affect learning.

In addition to overt behaviors, researchers also demonstrated that people can learn cognitive rules from observing models. Bandura and Harris (1966) had models show children common nouns and sentences with each noun. Some children viewed models who demonstrated the desired construction (e.g., passive verbs) and who then took turns with them in constructing sentences. Reinforcement was given to some children for constructing passive voice sentences. Problem-solving set children were told to determine what type of construction earned them reinforcement.

Although the use of passive constructions by children was low following the experimental treatments, modeling combined with reinforcement and problem-solving set led to the highest number. Children did not imitate the models but rather acquired the grammatical rule and constructed their own sentences. Further evidence of rule learning was obtained in an early study by Rosenthal et al. (1971). Collectively, these studies showed that people learn from observing models and that reinforcement was not needed for learning.

Conceptual framework

The conceptual framework of Bandura's social cognitive theory is based on reciprocal interactions among personal, behavioral, and social/environmental factors. This section also discusses enactive and vicarious learning, as well as vicarious, symbolic, and self-regulatory processes.

Reciprocal interactions

A central tenet of Bandura's (1977b, 1986, 1997, 2001) social cognitive theory is that human psychological activity depends on *triadic reciprocity* involving reciprocal interactions among three sets of influences: personal (e.g., cognitions, beliefs, skills, affects); behavioral; and social/environmental factors. Individuals' cognitions and environments can affect their behaviors, which in turn can alter their environments and subsequent personal influences (Fig. 1).

These interacting influences can be exemplified using *self-efficacy*, which is a personal factor defined as one's perceived capabilities for learning or performing actions at designated levels (Bandura, 1977a, 1997). With respect to the interaction of self-efficacy and behavior, researchers have shown that self-efficacy influences achievement behaviors such as task choice, effort, persistence, and use of effective learning strategies (Schunk and DiBenedetto, 2016). These behaviors also affect self-efficacy. As students work on tasks and observe their learning progress, their self-efficacy for continued learning is enhanced.

The influence between personal and social/environmental factors can be illustrated with students with learning disabilities, many of whom hold low self-efficacy (Licht and Kistner, 1986). Individuals in their environments may react to them based on common attributes (e.g., low skills) rather than on actual capabilities. In turn, social/environmental feedback can affect students' self-efficacy, such as when a teacher tells a student "I know you can do this."

The link between behaviors and social/environmental factors is seen in many instructional sequences. Social/environmental factors direct behaviors when a teacher points to a display and says, "Look here," which students do without much conscious effort. Students' behaviors can alter their instructional environments. When teachers ask questions and students give incorrect answers, teachers re-teach the material rather than continuing the lesson.

Bandura's social cognitive theory reflects a view of human *agency* in which individuals proactively influence their lives (Schunk and Pajares, 2005). They hold beliefs that allow them to exert a large degree of control over important events in their lives. People who want to become musicians take music lessons and get mentoring from experienced musicians. In reciprocal fashion, people influence and are influenced by their actions and environments. Music students change their behaviors and environments in response to feedback from their mentors.

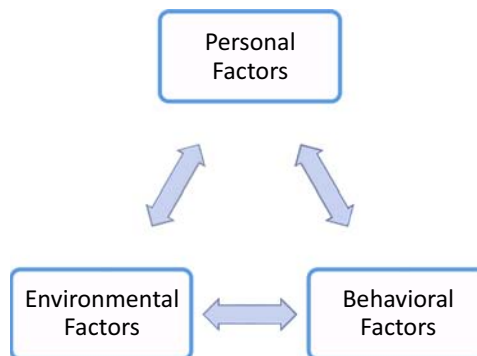


Fig. 1 Triadic reciprocity in social cognitive theory.

The scope of this reciprocal influence is broad because individuals live in social environments. *Collective agency* refers to people's shared beliefs about what they are capable of accomplishing as a group. Like individuals, groups also affect and are influenced by their actions and environments. A group of science teachers may collectively decide to use more project-based learning to deepen students' understanding of scientific concepts. They develop and implement activities and continue to refine them based on students' learning.

Enactive and vicarious learning

Social cognitive theory distinguishes *enactive learning* (doing) from *vicarious learning* (observing modeled performances—live, filmed, symbolic; Bandura, 1977b). Behaviors that result in success are retained, whereas those that are unsuccessful are discarded or modified. Social cognitive theory contends that behavioral consequences inform and motivate. They inform people of the accuracy or appropriateness of behaviors. People who are rewarded for their actions understand that they are performing well, whereas punishments convey inappropriateness of behaviors. Further, people are motivated to learn and perform behaviors that they believe will have desirable consequences and do not try to learn behaviors that they believe will be punished.

Much human learning occurs *vicariously*, or without learners performing at the time of learning. Vicarious learning accelerates learning over what would be possible if people had to perform every action, and it saves people from undesirable consequences. Observing or reading about safety techniques saves individuals from acting in potentially dangerous ways. As with enactive learning, observers are more motivated to learn actions leading to successes than those that are unsuccessful. People attend to successful models who demonstrate actions that they believe will benefit them (Bandura, 1969; Schunk, 1987).

Much human learning occurs through a blend of enactive and vicarious learning. Learners observe teachers explain and demonstrate skills and strategies. By observing teacher models, students may learn some aspects of skills. As students practice the skills, teachers provide feedback and corrective instruction as needed.

Vicarious, symbolic, and self-regulatory processes

Bandura's (1986) social cognitive theory postulates that human capabilities facilitate individuals attaining a sense of agency. Vicarious, symbolic, and self-regulatory processes are prominent capabilities.

The capability for learning vicariously allows individuals to acquire knowledge, beliefs, affects, and behaviors, by observing others. This capability also allows people to develop their lives because they proactively select environmental features (e.g., individuals, materials) to which they want to attend. People who want to learn about Mars seek out situations where they can learn vicariously, such as by reading books, watching televised shows, attending lectures, and so forth.

Symbolic capabilities involving language and mathematical and scientific notation allow people to adapt to and alter their environments (Bandura, 1986). They use symbolic processes to interpret actions and outcomes in their lives and guide their future actions. Symbolic processes allow people not simply to react to events but rather to solve problems and generate new actions. Symbolic processes also foster communications with others (e.g., in person, virtually), which leads to further learning.

Self-regulation processes play a major role in social cognitive theory (Bandura, 1986; Usher and Schunk, 2018). People regulate their behaviors to conform to their internal standards and goals. Prior to beginning a task, they determine their goals and strategies to use. As they work on tasks, they assess their goal progress and decide whether to continue or alter their strategies. During breaks and when tasks are complete, they reflect on their experiences, seek to make sense of them, and determine their next steps. As they reflect on what they have done, the belief that they have learned strengthens their self-efficacy and motivates them to continue learning. The following sections explain these important processes in greater depth.

Vicarious processes

Functions of modeling

Bandura (1986) distinguished three functions of modeling: response facilitation, inhibition and disinhibition, and observational learning. *Response facilitation* refers to modeled actions that act as social prompts for observers to perform the same way. Response facilitation effects are common in everyday life. An individual walking down a street who encounters a group of people looking in a store window may stop and look in the window.

Response facilitation effects do not represent learning because people know how to perform the actions. The behaviors of others serve as cues for observers' actions. Response facilitation effects even can occur without conscious awareness (Chartrand and Bargh, 1999). People walking past a store clerk handing out coupons may mindlessly take one and then examine it.

Inhibition and disinhibition effects occur when models strengthen or weaken inhibitions to perform behaviors. Inhibition may occur when models are punished for their actions; disinhibition can result when models perform threatening or prohibited actions without negative consequences. Classroom misbehavior may be disinhibited when students observe other students misbehaving without being reprimanded by the teacher; a sudden reprimand may inhibit further misbehavior.

Like response facilitation, inhibition and disinhibition effects represent performance of previously learned behaviors, not new learning. A difference between these two categories is that response facilitation involves behaviors that are socially acceptable, but

inhibition and disinhibition typically involve actions that have moral or legal implications (e.g., breaking rules) or involve strong emotions (e.g., fears).

Observational learning

Observational learning through modeling occurs when observers act in ways that they had not learned to do prior to exposure to the models (Bandura, 1969). Observational learning has four component processes: attention, retention, production, and motivation.

Observational learning requires that observers attend to relevant features so that they can be perceived. Certain features of models and situations command better attention. Observers are more likely to attend to models who have status and credibility. Task features can affect attention, such as when teachers use bright colors, oversized features, and interactive materials. Attention also is affected by observers' beliefs about the functional value of the modeled behaviors. Modeled activities that observers believe are important and likely to lead to desirable outcomes enhance attention. Students' attention should be raised when teachers announce that the material they are about to cover will be in a test.

Retention involves cognitively organizing, rehearsing, coding, and transforming information for storage in memory. This aspect of observational learning has not been extensively investigated by social cognitive researchers. Cognitive information processing theorists and researchers have addressed this aspect in depth (Mayer, 2012).

Production involves translating cognitive conceptions of modeled actions into behaviors. Especially with complex behaviors, it often is the case that observers will learn only some features. Learners refine their skills through practice and feedback that may include additional modeling.

Motivation is important because observers are more apt to attend to, retain, and produce modeled actions that they believe are important. People are selective; they do not perform everything that they observe. Rather, they attempt to learn those actions they believe will lead to desirable outcomes and help them attain their goals, and they avoid those actions they believe will result in dissatisfying outcomes. People form expectations about the anticipated outcomes of actions based on their observations of models and other experiences. Their learning and performance of actions is based in part on these expectations.

Educational research

Researchers have substantiated the preceding points regarding observational learning. In school contexts, peers are important modeled influences on students (Schunk and Usher, 2019).

Researchers have investigated characteristics of peers that make them more credible models. Students are likely to attend to peers whom they believe are competent. Perceived similarity between models and observers can affect peer modeling (Schunk, 1987). Similarity in important ways serves as a source of information for determining behavioral appropriateness, forming outcome expectations, and assessing one's self-efficacy. Age similarity between model and observer is important for gauging behavioral appropriateness but less so for actual learning, which is enhanced better by models whom observers believe are competent. When competence and age similarity do not match (e.g., younger model more competent than a same-age model), children are swayed more by the competent model. Peers can be effective models when children hold self-doubts about their learning or performance capabilities. Viewing a similar peer successfully perform a task can raise observers' self-efficacy when they believe that if the model could learn, they can as well (Schunk, 1987).

Model gender can influence modeling by conveying information about task appropriateness. In general, observing a same-gender peer model perform without negative consequences conveys that the action is appropriate. Model gender is less important in learning academic skills and strategies (Schunk, 1987).

Researchers have explored the effects of exposing students to mastery and coping models (Schunk, 1987). Mastery models demonstrate desired performance from the outset. Coping models initially have difficulty learning skills, but through effort gradually improve their skills and eventually perform as well as mastery models. In a study exploring coping skills on an athletic task, 60 female high school students observed either a coping female model demonstrating dart-throwing techniques, a mastery model demonstrating flawless techniques, or no model (Kitsantas et al., 2000). The girls in the coping model condition surpassed skill acquisition over the females in the other two groups. In addition, learners who observed the coping models exhibited greater self-regulatory skills, intrinsic interest, and better learning. Children who have experienced previous difficulties learning may benefit more from observing coping models (Schunk and Hanson, 1985), presumably because children perceive themselves as more similar to coping models.

Researchers have found that it is helpful to incorporate modeled explanations into demonstrations such that models verbalize their thoughts and reasons for performing given actions. Meichenbaum and Goodman (1971) used cognitive modeling as part of a self-instructional program with impulsive children. Schunk (1981) compared the effects of cognitive modeling with those of didactic instruction on children's long-division self-efficacy and achievement. Children who lacked division skills received instruction and practice over sessions. Cognitive-modeling children observed an adult model explain and demonstrate division solution strategies while applying them to problems. Didactic-instruction children received written instructional material than explained and demonstrated the operations. Cognitive modeling led to greater division achievement and accuracy of perceived capabilities (i.e., self-efficacy corresponded more closely with actual skills).

Symbolic processes

Researchers have investigated the operation of symbolic processes on motivation. Symbolic processes include cognitions, beliefs, and mathematical/scientific notation. Some influential symbolic processes are self-efficacy, goals and self-evaluations of progress, outcome expectations, values, and social comparisons (Fig. 2).

Self-efficacy

In the model of reciprocal interactions, self-efficacy is a personal variable hypothesized to influence behaviors and environments and in turn be affected by them (Bandura, 1986, 1997). Students who feel efficacious about learning should engage in actions that improve learning, such as setting goals, using effective learning strategies, monitoring comprehension, evaluating goal progress, and creating effective environments for learning. In turn, self-efficacy is influenced by the outcomes of behaviors (e.g., goal progress, achievement) and by input from the environment (e.g., feedback from teachers, social comparisons with peers; Schunk and Usher, 2019).

People acquire information to assess their self-efficacy primarily from four sources: actual performances; observations of others (vicarious experiences); forms of social persuasion; and physiological indexes (Bandura, 1997). Actual performances constitute the most reliable information because they are tangible indicators of one's capabilities (Schunk and DiBenedetto, 2016). Successful performances raise self-efficacy whereas failures may lower it, although an occasional failure/success after several successes/failures should not have much impact.

Individuals acquire self-efficacy information vicariously through knowledge of how others perform (Bandura, 1997). Similarity to others is a cue for gauging one's self-efficacy (Schunk, 1987). Observing similar others succeed can raise observers' self-efficacy and motivate them to try the task; they are apt to believe that if others can do it, they can as well. But a boost in self-efficacy through vicarious experiences can be negated by subsequent failure. Persons who observe similar peers fail may believe they lack the competence to succeed, which can dissuade them from attempting the task.

Individuals also assess self-efficacy based on social persuasion from others (e.g., "I know you can do this;" Bandura, 1997). Social persuasions must be credible for people to believe that success is attainable. Although positive feedback can raise individuals' self-efficacy, the increase will not endure if they subsequently perform poorly (Schunk and DiBenedetto, 2016).

People acquire self-efficacy information from physiological and emotional indicators such as anxiety and stress (Bandura, 1997). Strong emotional reactions provide cues about anticipated success or failure. When people experience negative thoughts and fears about their capabilities (e.g., feeling nervous when thinking about speaking in front of a large group), those affective reactions can lower self-efficacy. When people notice that they are become less stressful (e.g., not as anxious while taking a test), they may experience higher self-efficacy for performing well.

Sources of self-efficacy information do not affect self-efficacy automatically (Bandura, 1997). Individuals interpret the results of events, and these interpretations provide the information on which judgments are based (Schunk and Usher, 2019). People weigh and combine information from various sources to form self-efficacy beliefs.

Self-efficacy is not the only influence on behavior. No amount of self-efficacy will produce a competent performance when students lack the skills to succeed. Students' values (perceptions of importance and utility of learning) also can affect behavior (Wigfield et al., 2016). Even students who feel highly efficacious about their science skills will not become science majors in college if they do not value a career as a scientist. Also important are outcome expectations (beliefs about the anticipated outcomes of one's actions; Bandura, 1997). Students typically engage in activities that they believe will result in positive outcomes and avoid actions that they believe may lead to negative outcomes. Assuming requisite skills and positive values and outcome expectations, self-efficacy is a key determinant of individuals' motivation, learning, self-regulation, and achievement (Schunk and DiBenedetto, 2016).

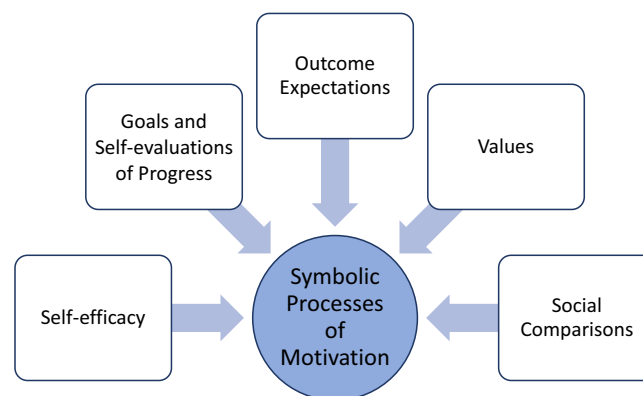


Fig. 2 Symbolic processes of motivation.

Self-efficacy can have diverse effects in achievement contexts (Bandura, 1997; Schunk and Usher, 2019). It can influence motivational outcomes such as choice of tasks, effort, persistence, and achievement. Individuals are apt to select tasks and activities at which they feel competent and avoid those at which they do not. Self-efficacy can affect how much effort people expend on an activity, how long they persist when they encounter difficulties, and their levels of learning and achievement. Students with high self-efficacy tend to set challenging goals, work diligently, persist in the face of failure, and recover their sense of self-efficacy after setbacks (Schunk and DiBenedetto, 2016). As discussed shortly, self-efficacy also can affect one's self-regulation activities (Usher and Schunk, 2018).

Goals and self-evaluations of progress

Goals are important symbolic processes that instigate and sustain actions. Goal commitment is critical because goals do not affect behavior without it (Locke and Latham, 2015). As people attempt to attain goals, they make progress judgments. Positive self-evaluations of progress strengthen self-efficacy, sustain motivation, and promote learning. Goals motivate people to expend effort and persist (Locke and Latham, 2015). A perceived discrepancy between performance and the goal may create dissatisfaction, which can raise effort and persistence and lead to better performance. Goals also help people attend to relevant task features, which can affect how people process information.

Goal properties

Goal properties of specificity, proximity, and difficulty affect motivation and learning (Zimmerman et al., 2015). Goals that include specific performance standards are more likely to lead to self-evaluations of progress and enhance motivation and learning than are general goals (e.g., "Do your best," Bandura, 1986). Specific goals indicate the amount of effort needed to succeed and evaluating progress toward specific goals is straightforward. Goals also differ by how far they project into the future. Proximal, short-term goals enhance motivation and learning better than do distant, long-term goals, because it is easier to determine progress toward goals that are closer at hand (Bandura and Schunk, 1981).

Difficulty, or the level of task proficiency required as assessed against a standard, can determine the amount of effort that people expend. Goals perceived as difficult but attainable are the best motivators. Goals viewed as too difficult do not motivate because people hold low self-efficacy for attaining them. Although people may feel efficacious for attaining goals perceived as very easy, these goals also may not motivate because people often procrastinate.

Learning and performance goals. Another distinction is between learning and performance goals. A *learning goal* refers to what knowledge, behavior, skill, or strategy students are to acquire; a *performance goal* refers to what task students are to complete. These goals can have differential effects on achievement behaviors (Anderman and Wolters, 2006). Learning goals focus students' attention on processes and strategies that help them acquire competence and improve their skills. The task focus motivates behavior and sustains attention on critical task aspects. Students who pursue a learning goal are apt to feel efficacious for attaining it and be motivated to expend effort, persist, and use effective learning strategies. Self-efficacy is substantiated as they work on the task and make positive self-evaluations of progress (Zimmerman et al., 2015).

In contrast, performance goals focus attention on completing tasks. They may not highlight the importance of the processes and strategies underlying task completion or raise self-efficacy for learning. Learners may not compare their present and past performances to determine progress. Performance goals can lead to social comparisons with others. These comparisons can lower self-efficacy among students who experience learning difficulties, which adversely affects motivation and learning (Anderman and Wolters, 2006).

Outcome expectations

Outcome expectations (e.g., "If I study hard, I should do well on the test.") are beliefs about the expected outcomes of actions. People form expectations based on personal experiences and observations of models (Bandura, 1986). Individuals act in ways they believe they will be successful and attend to models whom they believe will teach them valued skills. Outcome expectations can sustain goal-directed actions over lengthy periods when people believe their actions will eventually produce desired outcomes.

Self-efficacious students typically believe that their actions will result in positive outcomes, which keeps them motivated even when progress occurs slowly. Conversely, those whose self-efficacy is weaker may, when they encounter difficulties, work halfheartedly or quit. This situation can be demoralizing; students may believe that positive outcomes will result but lack the self-efficacy to continue. For example, they may believe that if they studied effectively they would do well on the test, but may doubt their self-efficacy to study effectively.

Values

Value refers to the perceived importance or usefulness of learning. Bandura's social cognitive theory predicts that actions reflect one's values (Bandura, 1986). Individuals strive to bring about what they desire and to avoid outcomes that are inconsistent with their values. They are motivated to learn when they deem that learning important. Students who value high grades are likely to engage in activities that should lead to high grades, such as attending class, keeping up with reading, turning in assignments on time, and studying for exams.

Achievement motivation researchers have shown that expectancies for success—a variable similar to self-efficacy—and values affect motivation (Wigfield et al., 2016). Expectancies predict achievement and values are better predictors of choices, such as intentions to take courses and enrollment in those courses (Wigfield et al., 2016).

Social comparisons

Social comparison refers to comparing ourselves with others (Wheeler and Suls, 2005). Although people often compare their performances with objective standards, they also evaluate their capabilities by comparing themselves with others. Social comparisons indicating that one is more competent than others can raise self-efficacy and motivation; social comparisons that result in negative self-evaluations can diminish these outcomes.

The most accurate self-evaluations arise from comparisons with others whom people believe are similar to themselves in the ability or characteristic being evaluated. The more alike observers are to models, the greater the probability that similar actions by observers will produce comparable results (Schunk, 1987). Model-observer perceived similarity in competence can improve learning (Braaksma et al., 2002). Observing similar others succeed can raise observers' self-efficacy and motivation. Similarity seems especially influential with persons who have experienced difficulties and possess self-doubts about performing well (Schunk and Usher, 2019).

Educational research

Researchers have explored the operation of social cognitive variables (e.g., self-efficacy, goals, social models) in various domains (e.g., education, business, health, computer applications). Participants in research studies have differed in age, developmental level (e.g., PreK-12, postsecondary), ethnicity, and abilities (e.g., regular, remedial, gifted) (DiBenedetto and Schunk, 2018). This research literature shows that self-efficacy influences motivation, achievement, and self-regulation, and in turn is affected by achievement outcomes (Bandura, 1997; Pajares, 1997; Schunk and DiBenedetto, 2016; Schunk and Usher, 2019).

Self-efficacy is especially influential on academic outcomes and performance accomplishments are the most reliable source as they provide evidence of what students can accomplish (Schunk and DiBenedetto, 2020). Self-efficacy correlates with motivation, learning, and achievement (Schunk and DiBenedetto, 2016). Both correlational and empirical research have demonstrated the effect of self-efficacy on academic self-regulation and achievement (Schunk and DiBenedetto, 2020). Cattellino et al. (2019) examined the mediation effects of self-efficacy on adolescent's self-regulation and performance in relation to the severity of parental and school rules on 804 Italian students. Students were administered a survey at the beginning of the school year, then again six months later. Path analysis revealed self-efficacy mediated the relationship between perceived severity of parental and school rules and school achievement six months after the initial survey was administered.

In another recent study, 273 twelfth-grade students' reading self-efficacy beliefs were collected using two instruments, one focused on reading self-efficacy and the other on reading comprehension (Dona et al., 2019). Findings showed a significant correlation between self-efficacy and performance. DiBenedetto and Bembenu (2013) examined changes in science self-efficacy for 113 college students over a semester in which students were enrolled in intermediate level science courses. Findings revealed self-efficacy beliefs at the end of the semester declined and yet were more closely related to final term averages than they were at the start of the semester, suggesting that students' beliefs about their performance became better calibrated as the semester progressed.

Experimental research studies have shown that instructional and social/environmental processes that convey information to students that they are becoming more competent raise self-efficacy, motivation, and achievement (Schunk and DiBenedetto, 2020). Falco and Summers (2019) conducted an intervention study incorporating the four sources of self-efficacy on high school girls' STEM (science, technology, engineering, mathematics) career self-efficacy beliefs. Eighty-eight ethnically diverse high school girls were split into intervention and control groups. The intervention group received nine 50-min counseling sessions targeted at building students' self-efficacy for making intentional career decisions and for building self-efficacy for careers in STEM. Results showed positive moderate-to-large effect sizes for the impact of the intervention on both students' career decision making self-efficacy and self-efficacy for careers in STEM.

Ramdass and Zimmerman (2011) examined the influence of modeling and social feedback on 76 sixth and seventh grade students' self-efficacy and mathematical achievement. Students were randomly assigned to one of four conditions: coping models with and without social feedback; and mastery models with and without social feedback. Findings revealed that students in the coping model conditions surpassed those in the mastery model conditions on the posttests and their self-efficacy was better associated with higher mathematics performance. These findings are similar to those cited earlier in the athletic study conducted by Kitasantas et al. (2000).

Self-regulation

An important expansion of social cognitive theory is into self-regulation (Usher and Schunk, 2018). This section discusses phases of self-regulation, the social-to-self progression in learning self-regulatory skills, and educational research.

Phases of self-regulation

Self-regulation refers to self-generated thoughts, feelings, and actions, which learners use to attain their goals (Zimmerman, 2000; Zimmerman and Schunk, 2004). Reciprocal interactions are evident in self-regulation. While working on tasks, learners may self-regulate personal factors, such as by setting goals, monitoring and evaluating progress, assessing self-efficacy for continued learning, and creating a positive emotional climate. They may exert behavioral self-regulation by employing learning strategies, verbalizing aloud or covertly steps to follow, and self-recording progress. They exert social/environmental control when they create productive work environments and seek assistance when needed. In turn, the result of their self-regulation affects self-efficacy, goals, strategy use, and other self-regulatory behaviors (Usher and Schunk, 2018).

Zimmerman (2000) developed a social cognitive cyclical model of self-regulation that includes reciprocal interactions and the changing nature of self-regulation. This model includes forethought, performance, and self-reflection phases. The forethought phase precedes learners' actions and includes processes that set the stage, such as setting goals and assessing self-efficacy for learning. During the performance phase, learners implement strategies and control their attention to focus on the learning context. The self-reflection phase, which occurs during breaks and after task completion, includes individuals' responses to their efforts, such as assessments of their learning progress and the effectiveness of their strategies.

Students' self-evaluations are critical and have cognitive, behavioral, and motivational effects. Especially important are learners' self-evaluations of their learning progress. Positive self-evaluations that they are making goal progress lead students to experience self-efficacy for continued learning and motivate them to continue to work diligently. Low self-evaluations of progress will not necessarily lower self-efficacy and motivation if learners believe that they can make better progress. Such learners may alter their self-regulatory processes by working harder, persisting longer, adopting what they believe is a better strategy, or seeking help from teachers, parents, coaches, or peers (Zimmerman, 2000). Research supports the use of the processes in the three phases of self-regulated learning among high performing students and demonstrates the processes to be highly correlated with one another (DiBenedetto and Zimmerman, 2010).

Social-to-self progression

Learners benefit from instruction on self-regulatory strategies, but it is critical that these strategies become *internalized*, or integrated into learners' self-regulatory systems. This internalization process occurs in a social-to-self progression comprising observation, emulation, self-control, and self-regulation (Fig. 3; Zimmerman and Schunk, 2004). Initially, learners acquire the rudiments of a skill by observing a model's performance. During emulation, students learn to duplicate a model's actions on the task, which involves emulating the model's general behavioral pattern or style of functioning.

At the self-control level, learners practice the skill apart from the model while monitoring their practice using cognitive and behavioral strategies. The fourth level (self-regulation) requires learners to make adjustments based on their outcomes. Learners can perform the skill with minimal self-monitoring because many elements have become internalized and may return to any of the levels as needed.

Self-regulatory skills can be learned without following these steps. For example, learners who have developed some self-regulatory skills may be able to refine new skills as they independently practice them. This progression, however, offers a sound basis for instruction and especially for students in the early stages of learning how to self-regulate their academic performances (Zimmerman, 2000).

A self-regulated framework

White and DiBenedetto (2015) presented an integrated framework whereby learners engage in the three phases of self-regulated learning within the four levels of self-regulatory competency (see Fig. 4). Pacing is critical in each of the phases as, initially, teachers are controlling the learning by modeling then gradually giving the students full control. As students begin to internalize self-efficacy

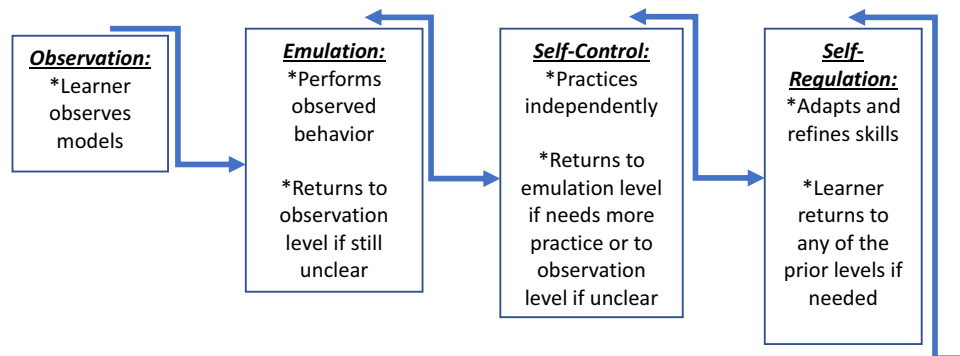


Fig. 3 Levels of self-regulatory competency.

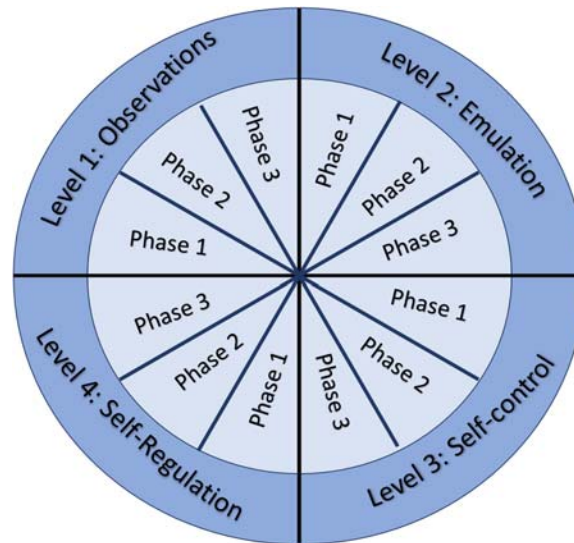


Fig. 4 Integrated framework: The three phases of self-regulated learning and the levels of self-regulatory competency.

and self-regulatory skills, they move from phase to phase within each level acquiring more independence. For example, during the *observation level* a fourth-grade teacher through cognitive modeling may verbalize how she would prepare for an upcoming mathematics exam by doing ten practice problems from the workbook (forethought phase: strategic planning). She models strategy use by completing the math practice problems and engages students to work on these practice problems with her (performance phase). The teacher then models evaluating her performance by making attributions and checking her progress on strategy use (self-reflection phase). The framework could be implemented as a form of instructional support at each of the four levels (observation, emulation, self-control, and self-regulation) to promote independent self-regulated learners (Bell and Pape, 2014). While research has been conducted on the three phases of self-regulated learning and the four levels of self-regulatory competency, research is needed to demonstrate the potency of this integrated framework.

Educational research

Researchers have identified the links between self-efficacy and self-regulation. In line with theoretical predictions, self-efficacy can influence such self-regulatory processes as goal setting, self-monitoring, self-evaluation, and strategy use (Usher and Schunk, 2018; Zimmerman, 2000). Self-efficacious students set challenging goals, persist, use effective learning strategies, and monitor their performances. These actions enhance self-efficacy and motivation for continued learning.

Self-evaluations of capabilities affect self-regulation and achievement (Schunk and Ertmer, 2000). Research by Schunk and Ertmer (1999) using college students showed that when opportunities for self-evaluation were infrequent, learning goals promoted achievement outcomes better than performance goals. However, when self-evaluation opportunities occurred regularly during instruction, the college self-evaluation the benefits of learning goals and led to enhanced self-efficacy and achievement regardless of learners' goals.

Researchers also have obtained support for the social-to-self progression. In teaching writing skills to college students, Zimmerman and Kitsantas (2002) found benefits when students observed models before they attempted performance. Other research has obtained support for the sequencing of the third and fourth levels (Zimmerman and Kitsantas, 1999). In this study, high school students initially received instruction and practice on a writing strategy (levels one and two). They then received a process goal (learning a revision strategy—level three), an outcome goal (decreasing the number of words in the revision—level four), or a process goal followed by an outcome goal. Shifting the goal resulted in better self-regulation than focusing exclusively on one type of goal.

Contributions

Although self-regulation has been addressed by other theoretical perspectives including behavioral, information processing, constructivist, and socio-cultural theories. There are noteworthy features of social cognitive theory. One is its emphasis on observational (vicarious) learning. This is especially evident in the social-to-self progression model. Vicarious learning speeds up the learning process over what would be possible if learners had to practice every action. Through observation of models, learners can acquire a sense of self-efficacy for learning that enhances their motivation and skill development.

Social cognitive theory also emphasizes self-efficacy. Learners initially may hold a sense of self-efficacy for learning, which motivates them to attend to models and practice skills. As they perceive that they are becoming more skillful and performing better, their self-efficacy for learning is strengthened, which leads to motivation for further improvement.

Finally, with its emphasis on reciprocal interactions between personal, behavioral, and social/environmental factors, social cognitive theory underscores the dynamic and changing nature self-regulation. Learners are proactive and seek ways to improve their skills, which can instate a sense of agency that they can exert much control over their academic learning.

Educational implications

Social cognitive theory and research suggest ways to improve teaching and learning through models, goals and self-evaluations of progress, self-efficacy for learning, and self-regulation.

Models

Instructional models inform and motivate learners. When they observe live or virtual models explain and demonstrate skills and strategies, learners acquire information about correct sequences of actions. When learners believe they understand what is being demonstrated, they likely will be motivated to demonstrate it. Peer models can be especially effective motivators. The perception of similarity raises observers' self-efficacy and motivates them to perform.

Students who often experience learning difficulties may benefit from observing peer coping models, that is, models who initially encounter difficulty learning but through effort, persistence, and correct strategy use, gradually improve their skills. Observers may perceive themselves as more similar in competence to coping models than to peer mastery models who demonstrate faultless performance from the start.

Goals and self-evaluations of progress

An effective means of developing learners' skills, self-efficacy, and motivation is for them to set learning goals and evaluate their learning progress. Goals that are close at hand, specific, and moderately difficult, are strong motivators. With long-term tasks (e.g., writing a paper), teachers can help students break the task into a series of subgoals (e.g., identify a topic, do research, write the first draft).

Students' learning and motivation benefit when they believe they are making progress in learning. Gauging progress is difficult for some tasks (e.g., becoming a better reader or writer). It is easier to gauge progress against subgoals than longer-term objectives. Learners can assess their progress in various ways such as by self-recording what aspects of the task they completed and judging their self-efficacy for performing subgoals. The perception of learning progress raises self-efficacy and motivation.

Self-efficacy for learning

For learners who lack skills, self-efficacy for performing them is apt to be low and thus not a motivator for students who lack the skills. For these students, self-efficacy for learning is a better motivator.

But if learners do not understand what is required to learn (i.e., what component skills are needed), they may judge self-efficacy for learning unrealistically low. In this case, teachers can point out what they have learned and how the new learning will draw on what they already know. Learners will feel efficacious about learning mathematical division when they understand that division requires estimating, subtracting, and multiplying skills they possess. Learners who engage in a task with a sense of self-efficacy for learning and who perceive that they are making learning progress will be motivated to continue.

Self-regulation

Social cognitive theory underscores learners becoming self-regulated. Self-regulatory skills such as strategy use and self-monitoring can be modeled. Setting goals and evaluating progress are self-regulatory processes that keep learners task-focused and maintain motivation. Having students evaluate self-efficacy at the outset of a task and during periods of self-reflection can convey that they are making progress, which strengthens self-efficacy and motivation.

DiBenedetto and Zimmerman (2010) showed that students who are self-regulated outperform students who engage in less self-regulatory processes. The cyclical nature of self-regulation suggests that teachers may need to assist learners to set new goals as they attain their current ones. Attaining goals denotes progress, especially when the goals represent subgoals in a larger task. Working on the next subgoal may require learning new skills, which teachers can facilitate through modeling.

Future research directions

There are several issues that merit future research. Discussed in this section are methodological refinements, cross-cultural relevance, and integration with technology.

Methodological refinements

Motivational variables stressed by social cognitive theory (e.g., outcome expectations, self-efficacy) are not static but can change quickly and frequently, but often motivation variables are assessed in a static fashion such as by using questionnaires that require aggregated and retrospective judgments. These measures capture changes between two time points but not changes that occur during task engagement.

Increasingly researchers are using real-time measures. For example, microanalytic methods collect assessments before, during, and after task engagement. Think-aloud protocols capture participants' thoughts as they engage in learning. Interviews ask participants about their thoughts at various times. Written diaries ask learners to record their thoughts as they work on tasks. With technology, traces show learners' progress through a learning episode. These and other types of assessments capture dynamic changes in motivation.

Because a variable such as self-efficacy can change rapidly, fine-grained measures are preferable because they show changes over time. Then researchers can map these changes against changes in instructional and social conditions, which will suggest future research directions and how to structure environments to promote beneficial motivation outcomes.

Cross-cultural relevance

Most social cognitive research has been conducted in Western societies, but this situation is changing as researchers are testing principles of social cognitive theory globally. The topics of self-efficacy and self-regulation have much international appeal. And cross-cultural research has yielded differences (McInerney, 2008). Klassen (2004) found that, compared with learners in collectivist cultures, those in individualistic (Western) cultures tend to judge self-efficacy higher. Interestingly, the alignment between students' self-efficacy and demonstrated skills is better for those in collectivist cultures. This is an important criterion. Students who overestimate their self-efficacy may attempt tasks beyond their means and perform poorly. Those who underestimate may be reluctant to engage in tasks and thereby preclude opportunities for learning (DiBenedetto and Schunk, 2018). These results suggest that collectivist cultures may promote modest self-efficacy judgments and that in some cultures collective self-efficacy (self-efficacy of what a group can accomplish) may better predict learning outcomes than will individual self-efficacy.

Overall, the principles espoused by social cognitive theory have shown themselves to be cross-culturally relevant. International researchers will continue to refine and test ideas in diverse cultural contexts.

Integration with technology

Social cognitive theory was largely developed prior to the advent of technological advances. Most research has been done in face-to-face settings. This does not mean that the theory needs major revisions because the principles are intended to be generic and apply across different contexts. But the role of technology requires some theoretical adaptations.

In particular, social cognitive research is needed with social media. These media offer ways for learners to have social contact with others, and we know little about how such interactions may influence self-efficacy and other variables. Learning from others is a source of self-efficacy information, and this should be true regardless of whether the interactions are live or virtual. Social media fit well with a theory with the "social" in its title.

Such research has implications for teaching and learning. There are educational uses using technologies such as Facebook and Zoom. How might these and other forms of media be used to help students set goals, monitor progress, assess self-efficacy for learning, and the like? How might instruction be designed to incorporate social media that take self-efficacy of learning and teachers into account? Research is needed to expand the generality of the theory beyond its original formulation.

Conclusion

Social cognitive theory is a general theory that stresses learning from the social environment. From its early focus on observational learning through modeling, social cognitive theory has expanded in scope to address such processes as motivation and self-regulation.

Bandura's social cognitive theory postulates reciprocal interactions among personal, behavioral, and social/environmental factors. Researchers in this tradition have investigated the operation of vicarious, symbolic, and self-regulatory processes, in the various ways that individuals interact with their environments and one another.

A key point is that persons seek to develop a sense of agency for being able to exert a large degree of control over important events in their lives. Among the influential variables affecting one's sense of agency are self-efficacy, outcome expectations, goals, and self-

evaluations of progress. We can expect further refinements in the theory as researchers apply this agentic perspective to human functioning in diverse areas in education and other domains.

References

- Anderman, E.M., Wolters, C.A., 2006. Goals, values, and affects: influences on student motivation. In: Alexander, P.A., Winne, P.H. (Eds.), *Handbook of Educational Psychology*, second ed. Erlbaum, Mahwah, NJ, pp. 369–389.
- Bandura, A., 1969. *Principles of Behavior Modification*. Holt, Rinehart & Winston, New York, NY.
- Bandura, A., 1977a. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215.
- Bandura, A., 1977b. *Social Learning Theory*. Prentice Hall, Englewood Cliffs, NJ.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, NJ.
- Bandura, A., 1997. Self-Efficacy: The Exercise of Control. Freeman, New York, NY.
- Bandura, A., 2001. Social cognitive theory: an agentic perspective. *Annu. Rev. Psychol.* 52, 1–26.
- Bandura, A., Harris, M.B., 1966. Modification of syntactic style. *J. Exp. Child Psychol.* 4, 341–352.
- Bandura, A., Ross, D., Ross, S.A., 1963. Imitation of film-mediated aggressive models. *J. Abnorm. Soc. Psychol.* 66, 3–11.
- Bandura, A., Schunk, D.H., 1981. Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *J. Pers. Soc. Psychol.* 41, 586–598.
- Bandura, A., Walters, R.H., 1963. *Social Learning and Personality Development*. Holt, Rinehart & Winston, New York, NY.
- Bell, C.V., Pape, S.J., 2014. Scaffolding the development of self-regulated learning in mathematics classrooms. *Middle Sch. J.* 45 (4), 23–32. Retrieved from: <https://login.libproxy.uncg.edu/login?url=https://www.proquest.com/scholarly-journals/scaffolding-development-self-regulated-learning/docview/1532768268/se-2?accountid=14604>.
- Braaksma, M.A.H., Rijlaarsdam, G., van den Bergh, H., 2002. Observational learning and the effects of model-observer similarity. *J. Educ. Psychol.* 94, 405–415.
- Cattellino, E., Morelli, M., Baiocco, R., Chirumbolo, A., 2019. From external regulation to school achievement: the mediation of self-efficacy at school. *J. Appl. Dev. Psychol.* 60, 127–133. <https://doi.org/10.1016/j.appdev.2018.09.007>.
- Chartrand, T.L., Bargh, J.A., 1999. The Chameleon Effect: the perception-behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910.
- DiBenedetto, M.K., Bembenutty, H., 2013. Within the science pipeline: self-regulated learning and academic achievement among college students in science classes. *Learn. Individ. Differ.* 2, 218–224.
- DiBenedetto, M.K., Schunk, D.H., 2018. Self-efficacy in education revisited through a sociocultural lens: current and future directions. In: Liem, G.A.D., McInerney, D.M. (Eds.), *Big Theories Revisited 2*. Information Age Press, Charlotte, NC, pp. 117–140.
- DiBenedetto, M.K., Zimmerman, B.J., 2010. Differences in self-regulatory processes among students studying science: a microanalytic investigation. *Int. J. Educ. Psychol.* 5 (1), 2–24.
- Dona, R., Dedi, S., Fernandita, F.G., 2019. The correlation between reading self-efficacy and reading comprehension. *Res. Teach. English* 3 (1), 1–13. <https://doi.org/10.33369/jeet.3.1.1-13>.
- Falco, L.D., Summers, J.J., 2019. Improving career decision self-efficacy and STEM self-efficacy in high school girls: evaluation of an intervention. *J. Career Dev.* 46 (1), 62–76. <https://doi.org/10.1177/0894845317721651>.
- James, W., 1890. *The Principles of Psychology*, vols. I & II. Henry Holt, New York, NY.
- Kitsantas, A., Zimmerman, B.J., Cleary, T., 2000. The role of observation and emulation in the development of athletic self-regulation. *J. Educ. Psychol.* 92 (4), 811–817.
- Klassen, R.M., 2004. Optimism and realism: a review of self-efficacy from a cross-cultural perspective. *Int. J. Psychol.* 39, 205–230.
- Licht, B.G., Kistner, J.A., 1986. Motivational problems of learning-disabled children: individual differences and their implications for treatment. In: Torgesen, J.K., Wong, B.W.L. (Eds.), *Psychological and Educational Perspectives on Learning Disabilities*. Academic Press, Orlando, pp. 225–255.
- Locke, E.A., Latham, G.P., 2015. Breaking the rules: a historical overview of goal-setting theory. In: Elliot, A.J. (Ed.), *Advances in Motivation Science*, vol. 2. Elsevier, Oxford, UK, pp. 99–126.
- Mayer, R.E., 2012. Information processing. In: Harris, K.R., Graham, S., Urdan, T. (Eds.), *APA Educational Psychology Handbook, Theories, Constructs, and Critical Issues*, vol. 1. American Psychological Association, Washington, DC, pp. 85–99.
- McDougall, W., 1926. In: *An Introduction to Social Psychology*, Rev. ed. John W. Luce, Boston, MA.
- McInerney, D.M., 2008. The motivational roles of cultural differences and cultural identity in self-regulated learning. In: Schunk, D.H., Zimmerman, B.J. (Eds.), *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Taylor & Francis, New York, NY, pp. 369–400.
- Meichenbaum, D., Goodman, J., 1971. Training impulsive children to talk to themselves: a means of developing self-control. *J. Abnorm. Psychol.* 77, 115–126.
- Pajares, F., 1997. Current directions in self-efficacy research. In: Maehr, M., Pintrich, P.R. (Eds.), *Motivation, Achievement, and Advanced*, vol. 10. JAI Press, Greenwich, CT, pp. 1–49.
- Rmadas, D., Zimmerman, B.J., 2011. The effects of modeling and social feedback on middle school's students' math performance and accuracy judgments. *Int. J. Educ. Psychol.* 7 (1), 4–23.
- Rosenthal, T.L., Moore, W.B., Dorfman, H., Nelson, B., 1971. Vicarious acquisition of a simple concept with experimenter as model. *Behav. Res. Ther.* 9, 217–227.
- Rosenthal, T.L., Zimmerman, B.J., 1978. *Social Learning and Cognition*. Academic Press, New York, NY.
- Rotter, J.B., 1954. *Social Learning and Clinical Psychology*. Prentice Hall, New York, NY.
- Schunk, D.H., 1981. Modeling and attributional effects on children's achievement: a self-efficacy analysis. *J. Educ. Psychol.* 73, 93–105.
- Schunk, D.H., 1987. Peer models and children's behavioral change. *Rev. Educ. Res.* 57, 149–174.
- Schunk, D.H., DiBenedetto, M.K., 2016. Self-efficacy theory in education. In: Wentzel, K.R., Miele, D.B. (Eds.), *Handbook of Motivation at School*, second ed. Routledge, New York, NY, pp. 34–54.
- Schunk, D.H., DiBenedetto, M.K., 2020. Motivation and social cognitive theory. *Contemp. Educ. Psychol.* 60. Article No. 101832.
- Schunk, D.H., Ertmer, P.A., 1999. Self-regulatory processes during computer skill acquisition: goal and self-evaluative influences. *J. Educ. Psychol.* 91, 251–260.
- Schunk, D.H., Ertmer, P.A., 2000. Self-regulation and academic learning: self-efficacy enhancing interventions. In: Boekaerts, M., Pintrich, P.R., Zeidner, M. (Eds.), *Handbook of Self-Regulation*. Academic Press, San Diego, pp. 631–649.
- Schunk, D.H., Hanson, A.R., 1985. Peer models: influence on children's self-efficacy and achievement. *J. Educ. Psychol.* 77, 313–322.
- Schunk, D.H., Pajares, F., 2005. Competence perceptions and academic functioning. In: Elliot, A.J., Dweck, C.S. (Eds.), *Handbook of Competence and Motivation*. Guilford Press, New York, NY, pp. 85–104.
- Schunk, D.H., Usher, E.L., 2019. Social cognitive theory and motivation. In: Ryan, R.M. (Ed.), *The Oxford Handbook of Human Motivation*, second ed. Oxford University Press, New York, NY, pp. 11–26.
- Skinner, B.F., 1953. *Science and Human Behavior*. Free Press, New York, NY.
- Tolman, E.C., Honzik, C.H., 1930. Introduction and removal of reward, and maze performance in rats. *Univ. Calif. Publ. Philos.* 4, 257–275.
- Usher, E.L., Schunk, D.H., 2018. Social cognitive theoretical perspective of self-regulation. In: Schunk, D.H., Greene, J.A. (Eds.), *Handbook of Self-Regulation of Learning and Performance*, second ed. Routledge, New York, NY, pp. 19–35.

- Watson, J.B., 1924. Behaviorism. Norton, New York, NY.
- Wheeler, L., Suls, J., 2005. Social comparison and self-evaluations of competence. In: Elliot, A.J., Dweck, C.S. (Eds.), *Handbook of Competence and Motivation*. Guilford Press, New York, NY, pp. 566–578.
- White, M.C., DiBenedetto, M.K., 2015. *Self-Regulation and the Common Core: Applications to ELA Standards*. Routledge, New York, NY.
- Wigfield, A., Tonks, S.M., Klauda, S.L., 2016. Expectancy-value theory. In: Wentzel, K.R., Miele, D.B. (Eds.), *Handbook of Motivation at School*, second ed. Routledge, New York, NY, pp. 55–74.
- Zimmerman, B.J., 2000. Attaining self-regulation: a social cognitive perspective. In: Boekaerts, M., Pintrich, P.R., Zeidner, M. (Eds.), *Handbook of Self-Regulation*. Academic Press, San Diego, pp. 13–39.
- Zimmerman, B.J., Kitsantas, A., 1999. Acquiring writing revision skill: shifting from process to outcome self-regulatory goals. *J. Educ. Psychol.* 91, 241–250.
- Zimmerman, B.J., Kitsantas, A., 2002. Acquiring writing revision and self-regulatory skill through observation and emulation. *J. Educ. Psychol.* 94, 660–668.
- Zimmerman, B.J., Schunk, D.H., 2004. Self-regulating intellectual processes and outcomes: a social cognitive perspective. In: Dai, D.Y., Sternberg, R.J. (Eds.), *Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development*. Erlbaum, Mahwah, NJ, pp. 323–349.
- Zimmerman, B.J., Schunk, D.H., DiBenedetto, M.K., 2015. A personal agency view of self-regulated learning: the role of goal setting. In: Guay, F., Marsh, H., McInerney, D.M., Craven, R.G. (Eds.), *Self-Concept, Motivation, and Identity: Underpinning Success With Research and Practice*. Information Age Publishing, Charlotte, NC, pp. 83–114.

Learning from a sociocultural perspective

Roger Säljö, Department of Education, Communication and Learning, University of Gothenburg, Gothenburg, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	36
The sociocultural perspective on learning	37
Learning and mediation	37
Tool mediated action	39
Learning, development and schooling	40
Learning in the zone of proximal development (ZPD)	41
Conclusion: the role of learning in the sociocultural perspective	41
References	42

Glossary

Mediation Zone of proximal development (ZPD)

Intermental Intramental psychological functions

Introduction

Humans have an incredible talent for learning. We are helpless at birth and completely dependent on external support for all our needs. During a period of a few years, we learn to walk, run and master our body, to play, talk and interact with other people, and a wide range of physical, intellectual and communicative skills. We are biological beings, primates, but our developmental trajectories during ontogeny differ, in intriguing ways, from those of our next of kin in the evolutionary sense. One of the features that stands out as particular about human learning and development during ontogeny is our flexibility. As a species, we are able to adapt to considerable variations in the cultural, material and societal conditions of the environment. Even though the capacity to learn a language is general, the language(s) we learn in infancy are those that are used in our immediate environment. A child in a monolingual family will learn the language of the parents, while a child growing up in a bi- or multilingual setting may learn two or even more languages, if the parents/guardians decide to expose the child to their respective mother tongues. Also, the material conditions in which children grow up testify to the flexibility of developmental trajectories. The material surroundings of children growing up in a megacity, beginning preschool/school at an early age and, in many cases, attending formal education well into their 20s, are different from those of a child growing up in a nomadic community in the desert. And, these differences will have consequences for socialization and what children learn and know. A rather dramatic illustration of the impact of cultural habits on children's developmental trajectories in contemporary society is the role that the Internet and digital tools currently play. In many societies, children as young as one or two years of age spend time on the Internet every day. They engage in practices such as watching video clips, listening to music, gaming and communicating with friends and relatives (Statens medieråd, 2019; UNICEF, 2019). These latter developments represent a significant change in the daily practices of children and young people that were inconceivable a few decades ago, and they are consequential for their learning trajectories and knowledge of the world.

However, and to add to the issue of flexibility, learning should not be seen as an individual activity only, as is commonly done. Throughout history, we have seen many changes in the living conditions of humans brought about by new technologies, inventions and insights. The wheel, the plow, the bow and arrow, writing, number systems, book printing, radio, television, cars, computers and an almost endless number of inventions and discoveries have changed our lifestyles, our mental and physical capacities and our knowledge of the world. This development of new technologies—artifacts—is an important outcome of the human capacity to learn. We are able to convert many of our insights into artifacts, which, in turn, transform how we live and learn.

Thus, we do not merely adapt to external conditions, we transform them. We live in a world that is designed—materially as well as intellectually. The transformation of the physical environment began with the development of material culture that goes back a long time before modern humans—*Homo Sapiens*—appeared. The oldest artifacts date back well over three million years (Harmand et al., 2015), while *Homo Sapiens* appeared as late as a few hundred thousand years ago (depending on definitions, see Callaway, 2017). So, learning is also about how collectives, societies and institutions learn and develop by continuously incorporating experiences, insights and technologies into the organization of daily practices.

This short summary provides a picture of the complexity and multilayered nature of the phenomenon of learning. In one sense, we are biological beings with limitation in terms of our physical strength and cognitive capacities. In another sense, we are tool-makers with infinite capacities for collaborating, inventing and expanding our knowledge and skills in ways in which no other species come close to. We build institutions and societies by accumulating knowledge—intellectual and material—over time; each new generation builds on the achievements of those that have gone before (Caldwell and Millen, 2008). This capacity for designing environments, in turn, implies that the conditions of learning and development continuously change. Societies become more diverse when knowledge expands. In the history of research on learning and development, there is a tension between conceptualizing learning as localized within individuals (in minds, brains, perceptual systems) versus conceiving it as processes that are embedded in social practices and in the co-ordination between people and between people and technologies and other external resources. This tension between perspectives divides research on learning (and development) in interesting ways.

The sociocultural perspective on learning

The sociocultural perspective, also referred to as the sociohistorical or cultural-historical research tradition, is best understood as a perspective rather than as a single and unified theory. There are important conceptual and methodological implications of adopting such a perspective in research and when designing educational activities, but all the research that is currently conducted under this conceptual umbrella in different fields cannot be reduced to one single theory.

What we now refer to as the sociocultural perspective originates in the work of the Russian psychologist Lev S. Vygotsky (1896–1934). As a student at the University of Moscow, Vygotsky had broad intellectual interests. His first degree was in law (1917), but his intellectual curiosity took him in other areas including linguistics, literature and philosophy. His dissertation, *The psychology of art* (Vygotsky, 1971), was published in 1925, but he never defended it owing to health problems (which continued to torment him throughout his short life). This is not the place to go into the history of the Vygotskian intellectual legacy (cf., e.g. Kozulin, 1986a,b), but it is interesting to note that his work did not receive much attention among Western scholars until the 1960s and 1970s. During the period of de-Stalinization in the Soviet Union, contacts between East and West became more frequent and academic exchange became possible.

In 1962, the first version of Vygotsky's influential work, *Thought and language*, appeared in a highly abridged version and with a foreword by Jerome Bruner (1915–2016). Later, extended editions appeared in 1986 (edited by A. Kozulin) and in 1987 as part of one of the volumes of Vygotsky's collected works. In the latter case, the title was changed to *Thinking and Speech*, and this reformulation is significant. Today, most scholars would agree that this translation better captures the interactional, processual, and communicative perspective on language and thinking that Vygotsky articulated. This issue of how to translate Vygotsky's concepts into English, in itself, has attracted intense scholarly debates over the past decades. Many researchers argue that the terms chosen in English in part misrepresent or even distort Vygotsky's ideas. There is an extensive literature on the reception and interpretation of Vygotsky's work by Western scholars (cf., e.g. van der Veer and Valsiner, 1991; Veraksa and Sheridan, 2018; Yasnitsky, 2012), which testifies to the significance of his achievements for contemporary research on learning, development and schooling. As is the case with many great scholars, his writings are inspiring and suggestive, and open to different interpretations and emphases.

When discussing the origins of the sociocultural perspective, the contributions by Alexander Luria (1902–1977) and Aleksei N. Leontiev (1903–1979) should be recognized. Luria was a young scholar at the Institute of Psychology in Moscow when Vygotsky joined during the 1920s. Luria was studying the correlates between psychoanalytical concepts and processes and more “objective” psychological observations. Luria was a truly interdisciplinary scholar and he finished a medical degree after the death of Vygotsky. He is now referred to as one of the founders of modern neuropsychology through his research on how brain injuries impact psychological functions (such as memory), issues that he studied in his capacity as clinician and researcher in hospital care of wounded soldiers during and after World War II. Leontiev is known as the founder of Activity Theory, a theoretical perspective that gained prominence in the Soviet Union during the 1970s and that grew out of Vygotskian ideas. But there is an interesting debate on the tensions between Activity theory and some of the basic Vygotskian assumptions (cf., e.g. Kozulin, 1986b). In the literature, Vygotsky, Luria and Leontiev are often referred to as the founding “troika” of the sociocultural perspective.

Learning and mediation

Vygotsky was a developmental psychologist with interests in language, learning, esthetic development, and education. One way of understanding the intellectual origins of the sociocultural perspective is to start by considering what Vygotsky saw as shortcomings in the psychology of his time. He repeatedly talked about psychology as in a state of “crisis”, as a result of a lack of proper theoretical and methodological foundations (cf. Kozulin, 1999, p. 73ff; Vygotsky, 1997). One central element of his critique was directed toward the dominant stimulus-response theorizing of this time, that is, what in the West was referred to as behaviorism and in the Russian context as reflexology (which preceded the behaviorist movement). Vygotsky's critique, shared with his contemporaries John Dewey (1859–1952) and Jean Piaget (1896–1980), is that the assumption that learning and development can be understood through stimulus-response connections and reinforcement in the Pavlovian sense is mistaken and far too simplistic. There are many elements of what people learn that cannot be explained in this way.

In this context, Vygotsky makes a distinction between what he refers to as “lower” and “higher psychological (or mental) processes.” Initially, this choice of terminology may seem a bit strange, but what is meant is that some of our behaviors are straightforward reactions and responses to stimuli in the environment, while others are of a very different nature. We have genetically determined, natural reactions such as sucking and closing our eyes when exposed to light. In modern parlance, these reactions are hardwired into us, and they require no thinking or conscious reflection. These are the lower psychological processes or functions that may be successfully studied by means of stimulus-response psychology in the Pavlovian, behaviorist tradition.

Higher psychological functions, on the other hand, are different (Podolskiy, 2012). They involve our capacities to think, to use signs and language, to act creatively and to control our own thinking and our behaviors. Such functions are unique to humans, and they emerge from social interaction, where we encounter and learn to use signs and categorize the world around us (Vygotsky, 1977). The concept of *sign*, thus, is central to Vygotsky’s thinking. In *Mind in society* (Vygotsky, 1978, p. 40) he uses a simple triangle to explain the role of signs for human thinking, cognitive development and communication more generally.

The idea here is that, in the stimulus-response connection (and in lower psychological processes), there is an immediate and direct link between a stimulus and the reaction (indicated by the upper line). However, only a very limited proportion of our behaviors builds on such direct connections. Instead, a peculiar feature of human thinking is that our relation to the world is *mediated* by signs. When we see something, our thinking involves the use of signs by means of which we perceive and categorize what we encounter. To take an example, when I look out my window, I admire the beautiful roses in my neighbor’s garden. I can see that they are very different in color (pink, red, yellow etc.), shape (some big, others small) and I know that the seasons when they blossom are different. Still, I see that they are roses. Why do I do this? The answer is that my perception is mediated by a sign that even an amateur is familiar with—I know what a rose looks like. In sociocultural terms, I have access to a term/sign that I use in a variety of actions: it mediates my seeing, remembering, communicating. Thus, my thinking is not a direct response to colors or shapes of the plants I observe, it is mediated by conceptual knowledge. And another important point is that this knowledge of what a rose is, is not my own invention, it is a concept that has a long history and that I have learned to use through social interaction. Thus, my knowledge is sociohistorical (has been used for a long period), social (shared with others) and individual (known by me).

Human thinking through sign systems thus implies that our thinking, metaphorically speaking, makes a detour via shared conceptual systems as indicated by the X in the triangle in Fig. 1. We perceive, talk and think about roses, clouds, cars, mammals and everything else that is around us through linguistic categories. And, much of our knowledge is conceptual or discursive to use a more recent term. We learn about the world largely by appropriating signs and discourses.

Thus, the sociocultural and neo-Vygotskian metaphor for learning is *appropriation* (Vygotsky referred to this as internalization but this term has some unfortunate connotations as has been pointed out by Rogoff, 1990, and others). Appropriation takes place through social interaction. Children (and adults) encounter concepts and knowledge in socially shared activities as “intermental (or interpersonal) categories”, and through this exposure, they may appropriate them as “intramental (intrapersonal) categories”, that is, as tools they can use for thinking. Thus, language (and sign systems) operate both between people and within them; “interpersonal use” leads to “intrapersonal use” (Vygotsky, 1978, p. 27ff). When I talk to my neighbor about her interest in roses, she can give detailed explanations of all the kinds of roses that exist (there are thousands, if you include hybrids), where they come from, what soil they require, how to cultivate them etc. Thus, she has appropriated a range of conceptual distinctions and physical knowledge that I am not familiar with but, if I were to spend time with her in her gardening, I could very well appropriate some of the relevant conceptual and physical skills. By engaging in joint activities and communication, the intermental categories that she uses when we talk may become intramental categories for me as a less knowledgeable partner in this particular interactional partnership (Wertsch, 1991). This is a central mechanism for communicating knowledge.

An important consequence of the capacity to use signs, emphasized by Vygotsky, is that they allow us to regulate our own actions and behaviors. Thus, concepts/signs are turned inwards in the sense that they allow us to remember, analyze and perform other activities that require sophisticated planning, logical reasoning and considerations of alternatives. Vygotsky (1978, p. 57) argues that the appropriation of “cultural forms of behavior involves the reconstruction of psychological activity on the basis of sign operations.” Through involvement in social interaction, children eventually “succeed in applying a social attitude to themselves” (p. 27) and consider their own actions and decisions (“I must be careful”, “I mustn’t not forget” etc.) in concrete situations.

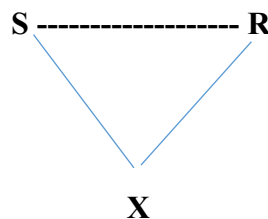


Fig. 1 Signs and mediation.

Tool mediated action

Language and sign systems, thus, are tools for thinking. The concept of tool is central to the Marxist thinking that Vygotsky was influenced by. We can think of language and other sign systems (writing, number systems and others) as in some sense analogous to physical tools that we use in activities such as gardening, digging and fishing. They are also integrated in social activities. The sophistication of the conceptual knowledge of my neighbor implies that she knows a lot about roses, their various appearances and their history, but part of her knowledge allows her to perform all the physical activities and use the full range of tools that are associated with cultivating roses. With physical, or technical, tools we carry out physical work, while conceptual tools are tools of the mind that allow us to think, communicate and plan our actions; in concrete actions they are interrelated.

Thus, in the sociocultural tradition, human actions, including thinking and physical action, are seen as tool mediated. Whatever we do, we use tools, and these tools have a social origin and sometimes a long history (sociogenesis). Vygotsky argues that language has a special status, it is the “tool of tools” and basic to human cognitive and intellectual functions. It is the technology of the intellect and an element of most of what we do. [Luria \(1981\)](#) points out that through the use of language as a sign system, we have a dual or double relationship to the world and what we do: we live *in* the world by performing complex activities (including talking), but we can also communicate *about* the world and what we do. We can discuss with others, borrow their insights and experiences, and engage in hypothetical thinking: what if we would do this instead of that? What would the consequences be? Again, this is something that is unique to humans.

The analogy between conceptual and physical tools should not make us think of them as totally separate. Rather, they coexist and presuppose each other. When I use a ruler or a scale to measure length and weight, the physical artifacts incorporate signs (numbers, classifications such as centimeters, inches, pounds), and the user orients to these signs while using them. Without the signs, artifacts of this kind would be useless or at least less helpful. As an alternative to using an ontology that separates intellectual (mental) and physical (material) tools, we may see them as cultural tools: human inventions that are socio-material and by means of which we interact with and transform our world.

Cultural tools extend human capacities far beyond the limitations of the resources provided to us by nature, as it were. With an axe, we can chop down a tree, and with a chain saw this is even quicker and requires less energy and strength. Analogously, with some knowledge of elementary geometry and a suitable measuring device, we can easily calculate the size of a room or a lawn when planning some concrete activity. Through history, there have been many inventions that have radically changed human intellectual capacities, our daily practices, and the societies we live in. One of the most obvious examples is the invention of writing about 5000–6000 years ago (depending on definition, see [Schmandt-Besserat, 1996](#)). Written language transformed how societies and individuals store and access information. The human memory at one level is quite limited. We have a limited attention span and we often forget and distort information. When writing and texts were introduced, the capacity for remembering was transformed for individuals as well as for societies. By documenting information and discourses, such as those used in science, religion, business and elsewhere, in texts, the human capacity for remembering and for conceptual thinking was radically transformed. In fact, when remembering is supported by texts, there are no absolute limits to how much information we can have access to. New insights and stories can always be documented and communicated, and this is precisely what we have seen over the millennia. This is a good illustration of higher psychological processes; no animal, no matter how sophisticated it is, will take a pencil and jot down something that has to be remembered.

New cultural tools thus imply a new division of labor between our minds, hands and artifacts. When new cultural tools are introduced and appropriated in society, our thinking, problem-solving, remembering, and other intellectual functions may change. Paper and pencil make it easier to multiply two four-digit number with two decimals than if are to perform these operations through mental arithmetic. An electronic calculator again changes the ways in which we perform calculations in daily practices. Using such a tool, multiplying large numbers with decimals is not more difficult than multiplying 2×2 for most people ([Säljö et al., 2006](#)). What happens in tool mediation is thus that some part of the cognitive process is achieved in collaboration with the tool. In the case of multiplying large numbers, the problem when not using paper and pencil or a calculator is that the calculation requires that we keep a lot of information in mind, that is, our short-term remembering is heavily taxed. Even if we know the principles of how to perform such calculations through mental arithmetic, it will be difficult for most people to manage this when the numbers are big and involve decimals ([Säljö et al., 2006](#)). With the tool (nowadays available on the mobile phone), the problem becomes almost trivial.

In a short paper, originally a lecture, [Vygotsky \(1981\)](#) makes a brief comment which illustrates the differences between traditional cognitivist and sociocultural perspectives on cognition and learning. In the former case, cognitive processes are seen as inherent to the mind/brain, while in the sociocultural perspective cognitive processes rely on cultural tools and are transformed when new cultural tools appear. Vygotsky thus argues that a new cultural tool “abolishes and makes unnecessary a number of natural processes, whose work is accomplished by the tool” (p. 139). The “natural processes” of remembering, calculating or whatever, will be transformed as the processing is accomplished in collaboration with tools. This is why, in a sociocultural and cultural-historical perspective, human thinking changes as the artifacts and technologies of the societies in which people grow up undergo change. To remember by using texts or databases, and to navigate at sea by using maps or GPS navigators, imply very different types of cognitive work in comparison to how such activities were handled earlier in history when such cultural tools were not available. Many examples of such changes in remembering, access to information and problem-solving can be observed as digitization spreads into the cognitive fabric of societies. Phone numbers do not have to be remembered any longer, search engines help us find what we look for and a specifically designed calculator from our bank helps us in calculating mortgages in ways that go far beyond what we would be capable of doing in the absence of such a cultural tool.

Learning, development and schooling

Tool mediated action, including the evolution of new cultural tools, is extremely interesting from the point of view of learning and human development. Tools are “cognitive amplifiers” (Nickerson, 2005) that extend our intellect far beyond what our “natural” capacities allow. The other side of the coin is that many cultural tools are not easy to understand and use. A mathematical formula, a grammatical or science concept (Vygotsky uses the example of Archimedes’ principle) or sophisticated software are powerful when used. At the same time, they introduce obstacles and can be hard to appropriate. Sometimes the process of appropriation may extend over years, such as when learning to be a proficient scholarly writer or understanding the details and implications of a complex theory.

Vygotsky makes an interesting distinction between different forms of mediation that is important to keep in mind when discussing learning in different contexts. When children develop, they learn “spontaneous (or “everyday”) concepts” by participating in communicative practices (Vygotsky, 1986). Terms such as sister, cousin and grandfather that mediate specific kinds of relationships to significant others are picked up as elements of daily concerns; they are in a sense learned from “below.” This mediation is largely “implicit” as Wertsch (2007) puts it. In the case of “scientific (or “academic”) concepts”, such as Archimedes principle in physics or debit and credit in bookkeeping, the situation is different. It does not matter how familiar we are with bathtubs, and what happens to the water level when we enter them, Archimedes principle about the relationships between the body and the force exerted on the body by the water will not occur to us. We notice that the water level rises as we enter the tub, but the scientifically formulated relationships between our weight and the force of the water displaced is not part of our spontaneous conceptual knowledge. To understand academic concepts of this kind, we need instruction and “explicit mediation” in Wertsch’s terminology.

This distinction points to something that is important in instruction in complex societies. Some cultural tools we use as part of everyday life and their functions become clear in contexts that are transparent for the individual; we experience the consequences of using them and we learn to generalize from “below.” For instance, and to return to the example above, I have a fair understanding of what a rose looks like just by living in a certain society where roses are talked about and presented in media. In the case of scientific concepts, the situation is different in the sense that the concepts will not be part of everyday interaction. Instead, they require explicit instruction and unpacking, sometimes over long periods and through systematic pedagogy. Learning about flowers at a scientific level takes a person into botany, classifications, Latin family names, Linnaeus’ sexual system and other areas of expertise. Sometimes, and even when we have learned academic concepts, they will often not be sustained through everyday interaction, since there are few settings in which they communicate well outside specialized communities. The botanist, the accountant, the nurse and the software engineer have appropriated a range of academic concepts, but these are primarily used in their professional communities, and not over a cup of coffee or tea friends.

But, academic or scientific concepts are very important for human learning and development, and one of the important responsibilities of schooling is to familiarize children (and adults) with such cultural tools. By appropriating central scientific concepts, relevant for understanding nature (greenhouse effect, energy, virus etc.) or society (inflation, democracy and how to calculate a mortgage), people are able to access and make use of the collective experiences and insights of their society. Thus, their knowledge and understanding will be broadened and no longer limited to the personal experiences they happen to have made in daily circumstances. When someone explains what a virus is and how it spreads, we are able to appropriate knowledge that we would never have discovered by ourselves. In the words of the evolutionary psychologist Merlin Donald (2018), through systematic instruction and education we are invited to share important insights of our “cultural memory.”

The learning process involved in appropriating academic concepts, thus, is different from the one characterizing spontaneous concepts. In the former case, we are introduced to a concept, a principle or a generalization, and we have to learn by realizing what they mean. Thus, we learn in a top-down fashion by first encountering a concept, a formula or an explanation, and then we embark on a journey attempting to learn how to use these cultural tools. In this sense, learning here goes from the general to the specific or particular. Once we have learned the principles of how to calculate the volume of a cylinder, we have to apply our knowledge to different kinds of cylinders in order to become efficient users, of the cultural tool. After a while, the concept or the procedure will be appropriated, and eventually fossilized in the sense that it becomes obvious to us how it is to be used. When we see a cylinder, our conceptual knowledge will be mobilized instantaneously.

The differences between spontaneous and scientific concepts, though, should not be understood in an absolute sense. This is an analytical distinction pointing to differences in the origin of our knowledge. Some concepts that were once used by specialized groups in society, for instance the greenhouse effect or inflation, have now entered everyday discourse and are encountered frequently in media and in public discussions. Many people, outside select groups of specialists, can give a rough account of what the greenhouse effect implies or what inflation is. In teaching and learning, there are also interesting connections between spontaneous and academic concepts. We begin learning academic concepts through our familiarity with spontaneous concepts or descriptions of phenomena. We learn about the concept of force or gravity by relying on experiences we have had in everyday life of things falling to the ground or sometimes going up into the air such as in the case of helium filled balloon. Instruction, at some level, has to build on children’s spontaneous concepts for the concepts and procedures to be intelligible and possible to appropriate as scientific concepts.

Thus, within a sociocultural perspective, generalizing and particularizing are significant elements of learning, and, in many cases, we oscillate between the general and the particular as we appropriate cultural tools. For instance, many people will think that they know what the difference between a fish and a mammal is at a general level, and they will be surprised when they hear that a whale or a dolphin is not a fish even though it swims and lives in the water. Thus, what is classified as fish and mammals in science will

lead to a reconsideration by the individual when encountering a scientific conceptual framework, and a new, more specific, understanding will emerge. Within a sociocultural perspective, appropriation of cultural tools is seldom, if ever, finished. Such tools will be refined and specified as we encounter new experiences and dive into specialized intellectual territories and activities. Generalization and particularization will continue when we engage in new practices.

But, an important conclusion is that many cultural tools are unlikely to be appropriated through everyday activities. They require teachers and explicit instruction. Writing, scientific discourses in many disciplines, mathematical practices, use of tabular representations, a foreign language and the principles of bookkeeping exemplify cultural tools where there are obvious thresholds, and where systematic pedagogy is a vital premise for appropriation.

Learning in the zone of proximal development (ZPD)

One of the most well-known and most discussed ideas in the Vygotskian framework for analyzing learning and development is the concept of ZPD. Vygotsky (1978, p. 90) argues that “learning creates a zone of proximal development; that is learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in their environment and in cooperation with his peers.” Thus, as the child appropriates concepts (round, square) and skills (knowing how to add one-digit numbers), there are neighboring forms of knowing that metaphorically speaking come within reach, but where the child temporarily will need the assistance of more “capable peers” to manage a task. ZPD is defined as

the distance between actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (1978, p. 86).

The concept of distance here implies that there are some insights and skills that the child (or the adult) already possesses (adding one-digit numbers), but where he or she is dependent on interactional support from more “capable peers” when encountering a more complex, but similar, task. In this example, we may think of the skill of adding a one-digit number with a number that has two digits (3 plus 12). If the child is unable to handle this, the adult support can take a variety of forms, for instance asking questions that are relevant and/or simplifying the problem (if you add 2 and 3, what do you get? What do you have to add then?). In the case of a physical skill, such as tying one’s shoelaces, the parent will simplify the appropriation by showing how to place the shoelaces when starting, how to cross them and so on (Nilholm and Säljö, 1996). Thus, for knowledge and skills that are within the ZPD, the child needs temporarily to “borrow” the skills of a more knowledgeable partner in the activity. After some practice, the child will have appropriated all the elements of the action and will be able to perform the task without support, or, as Vygotsky puts it, “independently.” Later on, the child will be able to handle such tasks while, at the same time, engaging in other activities such as talking to a friend or watching television. The skills needed have been appropriated to a level where they require almost no conscious attention.

All learning thus creates ZPDs. When we learn something, new ZPDs appear, and new knowledge comes within reach. In Vygotsky’s terminology, new “buds” of mental processes appear, that is, knowledge and skills that are still at an early, embryonic state, and which need to mature for the individual to be able to use them independently. From this perspective, there is a view of learning and development, where humans are seen as constantly engaged in appropriating new knowledge as they encounter new challenges. We do not reach final stages where our intellectual development is complete (as developmental theories often claim), rather we learn when we encounter new practices or cultural tools that may deepen or extend our knowledge. These ideas of new ZPDs appearing when we learn are not specific to children; they characterize learning throughout life.

Another reason why Vygotsky emphasized the role of ZPDs is that he argued that it is in the zone of proximal development that the child is sensitive to instruction. Thus, the child’s intellectual capacity in a particular setting has reached a level where he or she can understand and follow the explanations and support provided by the partner, thus making it possible to appropriate a skill or a concept. This means that knowing how to intervene in ZPDs is an essential constituent of the professional skills of teachers and other instructors. In the sociocultural literature, this kind of support is referred to as “scaffolding”, following a famous article published by Wood et al. (1976). There is a whole literature on the role of scaffolding in teaching and learning in many areas: mathematics, science, language learning etc.

Conclusion: the role of learning in the sociocultural perspective

Learning is a central concept within the sociocultural perspective. It is by appropriating cultural tools that people develop and are able to utilize significant elements of the knowledge and skills of the cultural memory of their society. This assumption differs from most theories of human development, which generally construe development as a process of maturation. The latter perspective, characteristic of the developmental psychology of Jean Piaget, describes development as essentially parallel to biological development. As the child grows older, his/her cognitive, communicative and physical capacities evolve through different stages until they reach those of adults. The sociocultural perspective represents an alternative view, where learning and appropriation are seen as preceding development. It is by appropriating cultural tools (language, sign systems and other cultural tools) that we develop

into competent participants in cultural practices. Simply put, learning comes first, development follows. And there is no final cognitive stage or level of knowledge; we continue to appropriate new knowledge and skills throughout our lives. In a sense, we are doomed to learn, and this is especially true if we live in societies where cultural tools change at a rapid pace.

The sociocultural perspective and the concepts of cultural tools and appropriation imply that learning and development are viewed in the context of the socialization practices of societies that are situated in time and history. Jerome Bruner (1997), in a famous article with the title “Celebrating diversity”, published on the occasion of the centennial of the birth of Vygotsky and Piaget, explicated the tensions between the two perspectives on learning and development. For Piaget, who is a universalist, development is a process of maturation and unfolding of intellectual capacities that reside within the individual mind. In such a tradition, the cognitive stages that characterize development are central, and learning plays a minor role. For Vygotsky, who is a contextualist and has an interest in understanding the relationships between cultural practices and human development, learning is the process by means of which individuals come to share the insights and skills of the society in which they are grow up. Thus, in the latter perspective, culture and cultural practices, including the cognitive and material artifacts, are formative of mind.

It is important to note that the sociocultural perspective does not imply that the human mind is influenced by culture, as is sometimes argued. Rather, the claim is much stronger and precise: culture is constitutive of mind. It is by appropriating cultural tools that we develop mental and practical skills relevant for social activities. In this sense, culture and cultural practices “hi-jack” the human mind as Donald (2001) puts it, and what we learn and know are the insights and skills of the cultural memory of our society. For Vygotsky, this process of cultural learning in its most powerful forms begins when the child is able to use language as a means of communication. When entering this phase at about one to two years, the child is able to live *in* the world and simultaneously increase his or her capacity to communicate *about* the world and his or her own activities. Vygotsky argued that it is almost as if we are born twice: once as biological beings (at birth), and then again when we are able to use language and appropriate the cultural tools of our society.

It is almost a hundred years ago that Vygotsky formulated his ideas. During a short and difficult life, he managed to articulate a perspective on learning, development and communication that continues to inspire scholars in a range of disciplines as well as practitioners in education and teacher education. His writings are fascinating, but sometimes they are sketches with original ideas and observations rather than finalized texts. There are several elements of his conceptual framework that he did not have time to develop in depth. This is one of the sources of disagreements among scholars as to what he meant. It is also obvious that the contexts in which we, as contemporary readers, understand his writings are very different from those that characterized his time. Relationships between adults and children have changed, the material circumstances are different as is the media situation and the ideals of education and teaching and learning. The Vygotskian account of learning and the sociocultural formation of mind has a lot to offer, and the theoretical and practical implications of this perspective in the context of technologically and socially complex societies inspire further research across the world.

References

- Bruner, J., 1997. Celebrating divergence: Piaget and Vygotsky. *Hum. Dev.* 40 (2), 63–73.
- Caldwell, C.C., Millen, E.A., 2008. Studying cumulative cultural evolution in the laboratory. *Philos. Trans. R. Soc. B* 363, 3529–3539. <https://doi.org/10.1098/rstb.2008.0133>.
- Callaway, E., 2017. Oldest *Homo sapiens* fossil claim rewrites our species' history. *Nature*. <https://doi.org/10.1038/nature.2017.22114>.
- Donald, M., 2001. *A Mind So Rare. The Evolution of Human Consciousness*. Norton.
- Donald, M., 2018. The evolutionary origins of human cultural memory. In: Wagoner, B. (Ed.), *Handbook of Culture and Memory*. Oxford University Press, pp. 19–40.
- Harmand, S., Lewis, J.E., Feibel, C.S., Lepre, C.J., Prat, S., Lenoble, A., et al., 2015. 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature* 521 (7552), 310–315. <https://doi.org/10.1038/nature14464>.
- Kozulin, A., 1986a. Vygotsky in context. In: Vygotsky, L.S. (Ed.), *Thought and Language*. The MIT-Press, pp. XI–LVI.
- Kozulin, A., 1986b. The concept of activity in Soviet psychology: Vygotsky, his disciples and critics. *Am. Psychol.* 41 (3), 264–274. <https://doi.org/10.1037/0003-066X.41.3.264>.
- Kozulin, A., 1999. *Vygotsky's Psychology. A Biography of Ideas*. Harvester/Wheatsheaf.
- Luria, A., 1981. *Language and Cognition*. Wiley.
- Nickerson, R.S., 2005. Technology and cognition amplification. In: Sternberg, R.J., Preiss, D.D. (Eds.), *Intelligence and Technology. The Impact of Tools on the Nature and Development of Human Abilities*. Erlbaum, pp. 3–27.
- Nilholm, C., Säljö, R., 1996. Co-action and situation definitions. An empirical study of problem solving in mother–child interaction. *Learn. Instruct.* 6 (4), 325–344.
- Podolskiy, A., 2012. Cultural-historical theory of development. In: Seel, N.M. (Ed.), *Encyclopedia of the Sciences of Learning*. Springer, pp. 879–881.
- Rogoff, B., 1990. *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press.
- Säljö, R., Eklund, A.-C., Mäkitalo, Å., 2006. Reasoning with mental tools and physical artefacts in everyday problem solving. In: Verschaffel, L., Dochy, F., Boekaerts, M., Vosniadou, S. (Eds.), *Instructional Psychology: Past, Present and Future Trends*. Pergamon, pp. 73–90.
- Schmandt-Besserat, D., 1996. *How Writing Came About*. University of Texas Press.
- Statens medieråd, S., 2019. *Småungar Och Medier [Kids and Media]*. Retrieved October 2019 from. <https://statensmedierad.se/publikationer/ungarochmedier/ungarochmedier2019.3347.html>.
- UNICEF, 2019. *Growing Up in a Connected World*. Retrieved October 2019 from. <https://www.unicef-irc.org/growing-up-connected>.
- van der Veer, R., Valsiner, J., 1991. *Understanding Vygotsky: A Quest for Synthesis*. Blackwell.
- Veraksa, N., Sheridan, S. (Eds.), 2018. *Vygotsky's Theory in Early Childhood Education. Russian and Western Values*. Routledge.
- Vygotsky, L.S., 1971. *The Psychology of Art*. The MIT-Press.
- Vygotsky, L.S., 1977. The development of higher psychological functions. *Sov. Psychol.* 15 (3), 60–73. <https://doi.org/10.2753/RPO1061-0405150360>.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Vygotsky, L.S., 1981. The instrumental method in psychology. In: Wertsch, J.V. (Ed.), *The Concept of Activity in Soviet Psychology*. M. E. Sharpe, pp. 134–143.
- Vygotsky, L.S., 1986. *Thought and Language*. Trans. In: Kozulin, A. (Ed.). MIT-Press.

- Vygotsky, L.S., 1997. The historical meaning of the crisis in psychology: a methodological investigation. Trans. In: van der Veer, R. (Ed.), *The Collected Works of L.S. Vygotsky*. In: Rieber, R.W., Wollock, J. (Eds.), *Problems of the Theory and History of Psychology*, vol. 3. Plenum, pp. 233–343.
- Wertsch, J.V., 1991. *Voices of the Mind: A Sociocultural Approach to Mediated Action*. Harvard University Press.
- Wertsch, J.V., 2007. Mediation. In: Daniels, H., Cole, M., Wertsch, J. (Eds.), *The Cambridge Companion to Vygotsky*. Cambridge University Press, pp. 178–192.
- Wood, D., Bruner, J.S., Ross, G., 1976. The role of tutoring in problem-solving. *J. Child Psychol. Psychiatr.* 17, 89–100.
- Yasnitsky, A., 2012. Lev Vygotsky: philologist and defectologist, a sociointellectual biography. In: Pickren, W.E., Dewsbury, D.A., Wertheimer, M. (Eds.), *Portraits of Pioneers in Developmental Psychology*. Psychology Press, pp. 109–133.

Perspectives on learning from the learning sciences

Frank Fischer^a, Daniel Sommerhoff^b, and Anna Keune^c, ^a Ludwig-Maximilians-Universität München, Department of Psychology, Munich Center of the Learning Sciences, Munich, Germany; ^b Leibniz Institute for Science and Mathematics Education, Department for Mathematics Education, Kiel, Germany; and ^c Technische Universität München, TUM School of Social Sciences and Technology, München, Germany

© 2023 Elsevier Ltd. All rights reserved.

A brief history of the learning sciences	44
The (self-)concept of the learning sciences	45
Theories of learning and guidance in the learning sciences	46
Theorizing learning	46
Supporting learning	47
The design of learning environments in a broad range of educational contexts	47
Pedagogies	47
Computer-Supported Collaborative Learning (CSCL) and its approaches to facilitating learning	49
Methods in the learning sciences	50
An outlook on advancing research on learning in the learning sciences	51
References	53
Relevant websites	56

A brief history of the learning sciences

Interest in human learning and effective ways of supporting learning through teaching and other forms of instruction can be traced back for millennia. The *learning sciences* as an emerging academic field originated in the 1980s (Hoadley, 2018). The emergence of the learning sciences was based on three groundbreaking shifts.

First, the idea of context and learning in situ moved to center stage. The shift toward context made it possible to consider the role of the social and physical world in and through which people learn as a basis for theorizing and understanding learning. Studying learning in context was a response to the cognitive paradigm of research on learning and instruction. As (socio-)constructivist and socio-cultural learning theories (e.g., Vygotsky and Cole, 1978) were increasingly adopted, the research focus shifted from studying the individual's cognition (cognitivism) in isolation to studying the learning of individuals and groups in their physical and social environments (e.g., Brown et al., 1989; Scardamalia and Bereiter, 1991). This development is mirrored in the influential 1993 special issue of the journal *Cognitive Science* which contained "a debate among proponents of two distinct approaches to the study of human cognition" (Norman, 1993, p. 1) and led to an increased interest in *situated action* or *situated cognition*, "emphasizing the role of the environment, the context, the social and cultural setting, and the situations in which actors [learners] find themselves" (p. 1). Viewing learning as driven by both internal and external factors led to a theoretical multiplicity, a tradition that the learning sciences continue to advance by staying open to new theoretical perspectives with utility for better understanding and supporting human learning.

Second, the understanding of learning as situated and influenced by socio-cultural and physical contexts required a new set of methodological approaches that allowed to capture (ongoing) learning as understood from these new perspectives. Established prior paradigms from (educational) psychology such as aptitude-treatment interaction (Cronbach and Snow, 1977) were questioned, as these did not adequately mirror the way learning happens within complex school and out-of-school settings, missing important aspects of ongoing learning. As a consequence, research in school and out-of-school learning environments became a dominant paradigm, necessitating the development and use of qualitative, quantitative, and mixed-methods approaches for capturing and facilitating learning (e.g., design-based research). Notably, the learning sciences are characterized by a methodological multiplicity that continues to develop in light of recent technological developments (e.g., Rosé, 2018; Shapiro and Garner, 2021; Sommerhoff et al., 2018).

Third, the learning sciences are also associated with a shift toward interdisciplinary research fueled by the rise of computer science, research on artificial intelligence, computer learning (i.e., how do computers learn?), and computerized learning (i.e., how do humans learn using computers or other technologies?) (e.g., Bobrow and Collins, 1975; Sleeman and Brown, 1982; Wenger, 1987). Trailblazing scholars, including Roger Schank and Seymour Papert, anticipated that computers would radically transform human learning and research on learning, also considering computers as tools that could advance mental structures through the production of personal (learning) designs (e.g., Papert, 1980). In particular, the extension of the phenomenon of *learning* from humans to computers and the comparison and combination of both agents (i.e., people and computers) brought together researchers from a diverse range of fields, including education, psychology, and computer science. For example, major interdisciplinary research lines were the implementation of learning in programming languages (e.g., LISP, LOGO), research on artificial intelligence and learning (e.g., Anderson et al., 1985), and the development of computer-supported learning environments

(e.g., [Cognition and Technology Group at Vanderbilt, 1992](#)). Corresponding interdisciplinary endeavors were, for example, nourished by first specific funding programs for the learning sciences, such as the “Cognitive Consequences of Programming” of the US National Institute of Education ([Pea and Linn, 2020](#)). The learning sciences have since participated in shaping a vision of computational innovations as mediating new and more advanced forms of learning (including computer-supported collaborative learning, see also [Lee, 2017](#)).

Together, these conceptual, methodological, and (inter-)disciplinary developments led to the formation of a community of researchers, which transcended the hitherto disciplinary boundaries and created the necessity for novel opportunities for meeting (e.g., conferences), publishing (e.g., journals), and funding (e.g., special programs). In 1989, the AERA special interest group “Education in Science and Technology” was founded and later also the sister special interest groups “Learning Sciences” and “Advanced Technologies for Learning” which became a central outlet for the developing learning sciences community. As a consequence of the significant growth of the learning sciences community in terms of impact and spread during the 1980s and 90s, higher education opportunities and separate scientific publication outlets for these areas of research became cogent for the further development of the field ([Kolodner, 2004](#)). The *Journal of the Learning Sciences* (JLS) was established in 1991 and the first doctoral program including the terms “Learning Sciences” in its title was launched at Northwestern University in 1992. The doctoral program evolved from the Institute for the Learning Sciences, which was founded at Northwestern in 1989. Parallel, the first International Conferences of the Learning Sciences (ICLS) and of Computer-Supported Collaborative Learning (CSCL) were held in the early 1990s ([Kolodner, 2004](#); [Sawyer, 2014](#)). Since 1995, both conferences have repeated regularly in a biannual format. To bring these conferences and the Journal of the Learning Sciences (JLS) under one institutionalized umbrella, the *International Society of the Learning Sciences* (ISLS) was founded as a non-profit professional society in 2002. In particular, ISLS brought together partially different communities clustering around the learning sciences and CSCL ([Hoadley, 2018](#)). To address the partially different interests of the CSCL community, the *International Journal of Computer Supported Collaborative Learning* (ijCSCL) was established in 2006 ([Stahl and Hesse, 2006](#)).

Research and community building in the context of the learning sciences and CSCL also emerged internationally beyond the North American focal point ([Pea and Linn, 2020](#)). For example, the European Research Network *Kaleidoscope* was established in 2004 based on a funding program by the European Commission that centered on Technologies for the Information Society. The network focused on *technology enhanced learning* and brought together over 1000 researchers from about 100 research units and approximately 20 countries, spanning the fields of computer and social sciences and education, with the goal to integrate different research communities in the context of technology-enhanced learning and to inform knowledge transfer between education, industry, and the wider society ([Laurillard et al., 2007](#)). Within Kaleidoscope, special interest groups on CSCL and on Inquiry Learning became focal points for new scholarly communities with strong conceptual and methodological relationships to the learning sciences. Many of these European CSCL researchers later joined the CSCL community within ISLS. Also in 2004, the *Asia-Pacific Society for Computers in Education* (APSCE), which aims to promote the conduct and communication of scientific research related to all aspects of the use of computers in education, was formed, evolving from the Asia-Pacific Chapter (APC) of the *Association for the Advancement of Computers in Education* (AACE), which itself was established in 1994.

Based on these national and international developments and its theoretical, methodological, and disciplinary pluralism, the learning sciences as an academic field has developed into a multidisciplinary “tent” ([Nathan et al., 2016](#)) that represents an international community and an emerging discipline (see also [Evans et al., 2016](#)). This is also reflected within the *Cambridge Handbook of the Learning Sciences* ([Sawyer, 2005, 2014](#)) and the *International Handbook of the Learning Sciences* ([Fischer et al., 2018b](#)), the contents of which comprise conceptual, methodological, and epistemological foundations of the learning sciences as well as an overview of current research. The handbooks serve as introductions to learning sciences as a community and field of research for students and researchers. The handbooks also capture the historical development of the learning sciences including proposals for future research trajectories for the academic field.

The (self-)concept of the learning sciences

Since its formation, the ISLS considers itself as a *community of practice* ([Wenger, 1999](#)) bringing together academics, professionals, and students who are seeking to advance the sciences and practices of learning. It investigates “fundamental inquiries on how people learn alone and in collaborative ways, as well as on how learning may be effectively facilitated by different social and organizational settings and new learning environment designs, particularly those incorporating information and communication technologies (ICT), as in computer-supported collaborative learning (CSCL)” ([ISLS, 2009](#), p. 1). It is a community that addresses authentic learning in contexts in which learning would naturally happen ([Kolodner, 2004](#)). The ISLS also has an educational mission, for example, to foster high quality learning sciences programs in higher education internationally. This educational aim led to the foundation of the *Network for Academic Programs in the Learning Sciences* (NAPLeS) at an ICLS conference in Sydney in 2012. NAPLeS’ first activities were the establishment of a world-wide directory of affiliated learning sciences master’s and PhD programs, a collection of syllabi for higher education introductory courses to the learning sciences, and a webinar series with a selection of learning scientists including pioneers who advanced groundbreaking insights that fundamentally shaped the learning sciences, including Carl Bereiter, Allan Collins, Jim Greeno, and Marlene Scardamalia. To date, NAPLeS has developed into an essential component of ISLS’s educational activities.

The unification of the biannual ICLS and CSCL conferences into the “ISLS Annual Meetings,” the first of them held in Bochum, Germany in 2021, reflects the co-development of both learning sciences and CSCL and the growing coalescence of the learning sciences as an emerging discipline. However, even with this convergent development, the learning sciences remain a community that distinguishes itself through theoretical, methodological, and disciplinary pluralism. Within this pluralistic understanding of studying learning as it unfolds within a social and physical context, certain core concepts and methods exist within the learning sciences that are orbited by other adjacent concepts and methods (e.g., [ISLS, 2009](#); [Nathan et al., 2016](#); [Sommerhoff et al., 2018](#)). This makes the learning sciences a dynamic society that aims to understand and foster human learning in authentic settings ranging from early childhood to adult education and informal learning over the life span.

Theories of learning and guidance in the learning sciences

Theorizing learning

From the perspective(s) of the learning sciences, learning is predominantly considered from cognitive and socio-cultural perspectives on learning, with embodied and neuroscientific learning theories expanding. However, what can be attended to as learning and what it means to learn in different contexts can be described differently depending on the theoretical perspective. Thus, a multitude of nuanced learning theories have evolved over time, which expand the broad theoretical approaches for more specific learning situations, in particular also combining learning theories with educational perspectives on how to support learning in these situations.

The cognitive revolution (see [Miller, 2003](#)), starting in the 1950s, led to a shift away from behavioral learning theories and was able to better account for many aspects of individual learning. However, subsequent research also highlighted multiple phenomena related to learning, for example individuals showing different performance in different contexts (e.g., [Carragher et al., 1985](#)), which could not be accounted for adequately by these theories. Corresponding findings led to socio-cultural learning theories, which extended current cognitivist conceptions of learning. The research of the Soviet psychologist Lev Vygotsky had a major influence (also outside of Russia), based on the re-publication (and partial revision) of his work “Mind in Society” in the late 1970s ([Vygotsky and Cole, 1978](#)). This conception of learning takes the perspective that all learning is socially embedded, leading to the focus on learning (within) communities and, for example, knowledge building ([Scardamalia and Bereiter, 1991](#); see section on Computer-Supported Collaborative Learning below) within these communities. This conception was further developed within the learning sciences context with concepts such as (i) “situated learning/cognition” ([Lave and Wenger, 1991](#)), positing that all knowledge is intrinsically linked to a specific social, cultural, and physical context, (ii) “distributed cognition” ([Hutchins, 1995](#)), focusing on how cognition is distributed over an environment rather than taking place solely in the individuals’ minds, and, (iii) “collective knowledge construction” ([Cress and Kimmerle, 2018](#)), which focuses on the joint creation of knowledge and content from relatively small groups to large communities, for example in wikis with hundreds of contributors.

For example, situated cognition draws on ideas from socio-cultural theories of learning and theorizes that learning is principally a social process ([Brown et al., 1989](#)) that emerges in interaction with others ([Vygotsky and Cole, 1978](#)). Within this theoretical approach, the notion of cognitive apprenticeship takes precedence, which refers to the enculturation of learners into the authentic, established, and valued practices of a community with the aim of becoming full members of that community ([Lave and Wenger, 1991](#)). Learning environments that are designed based on socio-cultural learning approaches support learners to interact with others toward engaging in ways of doing that are established and valued within the surrounding social context (e.g., [Engle and Conant, 2002](#)). This may mean identifying specific disciplinary practices and recognizing how learners use resources and approach content related to the discipline in ways that are recognizable and representative ([Engle and Conant, 2002](#)).

The shift from learning as something within an individual’s mind toward considering learning as situated within a community of practice means inviting learning to be theorized, captured, and designed for at multiple levels within the larger socio-cultural and historical system of relevant practices. This alludes to another central theoretical background of learning from a learning sciences perspective: Activity theory. While its origins go back to the Soviet psychologists Sergei Rubinstein and Alexei Leont’ev, it was taken up by several Western scholars (e.g., [Engeström, 2000](#)) and further developed, for example, through [Danish’s \(2014\)](#) work that employed an activity theoretical lens for educational design and analysis to analyze complex socio-cultural systems and their interactions. While activity theory aims at describing a single individual, it positions this individual in a broader context, a socio-technical system, that includes the community as a whole and, as crucial aspects of this community, their specific cultural tools, their rules, and the division of labor, all of which are necessary to describe the single individual within this context. Activity theorists often use the so-called activity triangle that can be applied as an analytical frame toward understanding an individual, a group, a community, or a system in context, their interactions, and possibly their changes over time. Besides this general necessity for context, activity theory also introduces a hierarchical structure of activity, which can be transferred to the notion of the multi-leveled nature of learning. [Roschelle \(1998\)](#) roughly characterizes these levels as “cultural, conscious, and automatic” (p. 243). While these three levels are often not considered explicitly enough, they inform why learning sciences research includes research approaches that range from the cultural level (e.g., using an ethnographic perspective) and the automatic level (e.g., using neuro-cognitive approaches). While less prominent in the learning sciences as a theoretical underpinning, a similar argument can be based on Newell’s unified theories of cognition ([1994](#)), which positions learning on different time scales from milliseconds (biological processes) to seconds (cognitive and rational processes) to hours, days, weeks, and beyond (socio-cultural, organizational) (see also [Nathan and Alibali, 2010](#)).

Each of these concepts for learning highlights that learning is situated (e.g., [Greeno, 2011](#)) in the sense that it is situated within a community of individuals, situated in certain situations, and situated within certain objects or artifacts, which might be used as means for learning. This view of learning can be characterized as core to the learning sciences; however, there is also a broad periphery of less radically social conceptions of learning around this core, which are used within the learning sciences and lead to tensions and synergies ([Danish and Gresalfi, 2018](#)).

An example of a specific theoretical approach trying to combine cognitive and socio-cultural perspectives is learning with multiple representations ([Ainsworth, 1999, 2018](#)). This approach gives a specific account on how multiple representations, such as texts, videos, or augmented/virtual reality impact learning and how they can (individually and in combination) be designed to support optimal learning. However, learning through multiple representations goes beyond purely cognitive theories of learning (e.g., [Mayer, 2014; Sweller et al., 1998](#)) by building on semiotic and sociological perspectives to better address the role of the context in the analysis and the design of representations for learning.

The learning sciences is also increasingly focusing on embodied ([Alibali and Nathan, 2012; Ma, 2016; Shapiro, 2019](#)) and neuro-cognitive learning theories ([Sousa, 2010; Varma et al., 2018](#)). The former point to the importance of the human body for cognition and learning, in particular regarding the sensory, perceptual, and the motor/kinesthetic systems as well as how bodies and a combination of bodies are positioned together to facilitate relationships to domain-specific ideas and concepts. It is rooted in part in theories from developmental psychology emphasizing that knowledge and thought are not only reflected in specific actions, but that actions, for example specific motions and gestures, can also support learning (e.g., [Goldin-Meadow, 2005](#)). Finger-counting or drawing geometric objects in midair to illustrate one's cognition when solving mathematical problems are two examples of embodied cognitive approaches to learning (e.g., [Alibali and Nathan, 2012](#)). Corresponding theories underline that effective learning environments consider possibilities for learners to engage in (bodily) actions, for example by using specific analog or digital manipulatives that can support these actions (e.g., [Abrahamson, 2009](#)).

While theories of embodied cognition and their implications are already permeating learning sciences research, neuro-cognitive learning theories are only starting to change current conceptions of learning as examined in the learning sciences (e.g., [Sousa, 2010](#)). Cognitive neuroscience has already made impressive progress on explaining cognition of individual learners focusing on short time-scales (partially at the level of fractions of seconds) in relatively controlled, laboratory settings (e.g., [Eden et al., 2004; Postle, 2020](#)). However, this is currently at odds with the learning sciences focus on learning in situ ([Varma et al., 2008](#)), which centralizes authentic, contextualized, social processes often related to timescales of days, weeks, and months.

Supporting learning

The design of learning environments in a broad range of educational contexts

Responding to the multilevel and multi-perspective view of learning, research that aims to support learning typically involves challenging situations that ask learners to grapple with tasks and problems “for which they do not have rote solutions or for which they cannot simply call upon a memorized factoid” ([Fischer et al., 2018a](#), p. 4). These challenging situations are typically embedded in information-rich environments that include interaction with other humans and technology.

Learning sciences research is addressing both the design of formal and informal learning environments. For example, formal learning environments have been investigated more systematically for primary and secondary mathematics and science education (e.g., [Belland et al., 2017](#)) and for argumentation and collaboration in different fields of higher education ([Jeong et al., 2019](#)). Informal learning environments have, for example, been studied in museums ([Lyons, 2018](#)) and in maker-centered learning spaces that can be situated within a range of educational institutions and that provide opportunities for people to engage with high- and low-technological tools and materials and can foster technological innovativeness ([Peppler et al., 2016](#)). Studying learning within out-of-school settings (rather than school settings) allows the advancement of knowledge about highly innovative designs for learning, as well as building inroads for later possible adaptation in school settings. Work on out-of-school settings also includes considering outdoor and indigenous knowledge spaces for broadening cultural relationships within STEM learning (e.g., [Litts et al., 2021; Tzou et al., 2019](#)).

In addition, learning sciences research considers the facilitation of learning and expertise across the lifespan. Although a majority of studies address K-12 learners, there is increasing interest in learning within early childhood (e.g., [van Horne and Bell, 2017](#)) and family settings (e.g., [Roque, 2016](#)). Research on expertise development has investigated environmental conditions and support for learners in a much broader age range, including learning at work and through professional development ([Ludvigsen and Nerland, 2018; Reimann and Markauskaite, 2018](#)).

Pedagogies

Learning scientists have systematically employed and explored pedagogies that acknowledge the multilevel nature of learning and address the design and use of challenging learning tasks (e.g., [Eberle, 2018](#)). They share the basic assumption that complex skills cannot be learned well solely through reading books, being exposed to teacher presentations, or solving massively simplified “toy” problems. Instead, situations are seen as more supportive of learning if they resemble real problem-solving situations in important dimensions of agency ([Brown et al., 1989](#)). Among these important dimensions are (a) agency in applying and transferring knowledge to new situations, cases, and problems; (b) agency in generating new knowledge (e.g., on natural and social phenomena), that is, to what extent learners take responsibility to find out about important mechanisms, effects, rules, or relations to explain natural

or social phenomena; and (c) agency in coordinating with, explaining to, and negotiating with others the processes and outcomes of knowledge generation and knowledge application (argumentation and knowledge building).

In the following, we briefly introduce three types of pedagogies, (1) learning through problem-solving, (2) learning through inquiry, and (3) learning through design. In addition, we sketch the concept of (4) scaffolding as it probably is the most prominent form of instructional support in complex learning environments that result when the three pedagogies are applied.

- (1) Learning through problem-solving: Environments for learning through problem-solving involve learners' engagement with complex problems with the goal of learning (Hmelo-Silver et al., 2018). Problem-solving pedagogies typically focus more on learners' agency with respect to applying and transferring of canonical knowledge (agency dimension a) than on learners' initial discovery of that or new knowledge. Highly productive lines of research have addressed an individual's learning through solving problems. One central research area has addressed the issue of cognitive load that arises through complex problems for learners with little prior knowledge, for example while trying to solve mathematical problems. Goal-free instruction and worked-out examples have been suggested as scaffolding associated with a reduced amount of cognitive load in comparison to solving the problem by oneself (Sweller et al., 1998). Four Component Instructional Design (4C/ID) is an instructional approach that builds on cognitive load theory and that has been applied in many contexts, including higher education (e.g., Sommerhoff et al., 2021) and vocational education (e.g., van Merriënboer and Kirschner, 2018).

Another longstanding and influential line of research that is associated with learning through problem-solving has been the investigation of intelligent tutoring systems (ITS or cognitive tutors, Graesser et al., 2018b). With ITS, learners try to solve problems presented to them on a computer, while ITSs use artificial intelligence to assess the process and the outcomes of a learner's attempts to solve a problem, to build a model of a learner's skills, and to offer feedback and scaffolding, for example, through more difficult or easier tasks, strategic hints, prompts, or examples. Beyond supporting individual learning through ITS, learning scientists began to explore the effects of collaborative learning within small group problem-solving that employ an ITS (Kumar and Kim, 2014; Rummel et al., 2016). Also, some approaches to game-based learning and simulations build on learning through problem-solving (Chernikova et al., 2020; Clark et al., 2016; Fields and Kafai, 2018).

There has been a long tradition of employing collaboration to enhance the effects of learning through problem-solving. The most prominent example is problem-based learning (PBL). In PBL scenarios, small groups of 5–12 learners collaborate around complex, domain-specific case problems (Dochy et al., 2003; Hmelo-Silver and Barrows, 2015). A main focus with respect to collaboration in this context is on pooling of resources: The group, supported by a tutor, identifies the knowledge needed to solve the case, for example to come up with an accurate diagnosis or a promising intervention for a patient case problem in medical education. The members of the group search for resources that offer access to the missing knowledge and share their findings in the group. In addition, further questions, including what makes tutors effective in PBL, have been addressed in research (Schmidt et al., 2011).

Recently, work within the learning sciences has started investigating the knowledge and skills necessary to engage in complex collaborative problem-solving (Graesser et al., 2018a) and has designed learning environments to facilitate them. For example, there have been approaches to use agent-based simulations where one or more human learners have interacted with collaboration partners that were simulated by a computer (e.g., Graesser and McNamara, 2010).

- (2) Learning through inquiry: Inquiry learning requires learners increasingly to take responsibility to find out about important mechanisms, effects, rules, and relations to explain natural or social phenomena. In inquiry-based learning environments, learners are supposed to explore phenomena, use disciplinary and generic tools and methods (e.g., measurement devices, lab protocols, spreadsheets), and come up with sound descriptions and explanations of the observed phenomena (see Pedaste et al., 2015). Many computer-supported inquiry learning environments have been developed for and employed in middle school science classrooms (e.g., Linn et al., 2003). Yet, increasingly, other domains (e.g., historical inquiry; Voet and de Wever, 2016) and other educational contexts (i.e., higher education; De Jong et al., 2013) have been explored in relation to inquiry-based pedagogies as well.

Building on research relating to the many problems that learners experience when engaging in knowledge generation through inquiry (e.g., De Jong and van Joolingen, 1998), several approaches to guided inquiry and discovery (in contrast to free discovery) have been developed and investigated (Donnelly et al., 2014). Typically, inquiry environments that learning scientists have developed include opportunities to interact with peers in small groups or whole classrooms in inquiry activities, such as formulating or evaluating hypotheses, generating evidence, or drawing conclusions and revising an explanation (Slotta et al., 2018).

- (3) Learning through designing: Pedagogies that build on learning through designing involve tasks in which learners engage in the process of utilizing disciplinary knowledge for generating ideas, devising a problem space, building prototypes, and evaluating and iteratively refining a product aimed at solving a challenging situation, often a provided problem. These pedagogies consider agency of the learners for engaging in inquiry activities by which they derive the disciplinary knowledge as well as the knowledge on the constraints and affordances of the problem, both of which are needed for designing, refining, and evaluating the product. Building on foundational work in the early 2000s (see Kafai, 2012; Kolodner et al., 2003), the field has substantially advanced by introducing more recent approaches to learning through designing digital games (Fields and Kafai, 2018), including augmented reality (Enyedy and Yoon, 2021), research on the maker movement (Halverson and Peppler, 2018), research on design thinking in different educational contexts (Svihla and Reeve, 2016; Goldman and Kabayadondo, 2016), and considering alternative tools and materials for the study of equitable STEM education through design activities (e.g.,

Peppler and Glosson, 2013). Design pedagogies typically include a focus on interaction among peers where collaboration has quite different functions, including peer feedback on initial ideas, inter-group competitions with respect to the performance of the designed product, and collaborative knowledge building through joint experimentation and discussion (Kolodner et al., 2003; see Fields and Kafai, 2018). Design pedagogies typically initiate and support peer communities using either digital contexts such as internet platforms (e.g., Scratch; see Fields and Kafai, 2018) or material contexts, for example in the context of science learning (e.g., Clegg and Kolodner, 2014).

Theoretical foundations of learning through designing pedagogies are often constructionist, that is assuming that learning happens best when learners come to know the world while creating personally meaningful projects that can be publicly shared (Papert, 1980). A focus lies on the creation and manipulation of tools and materials. For example, in the process of design, materials that are being manipulated turn into “objects-to-think-with” (Papert, 1980, p. 11) that make it possible for learners to become epistemologists who observe their own learning as well as personal and multiple ways of knowing (Papert and Harel, 1991). Studying particular material design situations thus allows research to expand on these theoretical ideas within constructionist approaches toward a richer understanding of ongoing learning within digital and physical learning through design situations.

- (4) Scaffolding: Referring to the concept of the Zone of Proximal Development, the zone between what learners can do on their own compared to what they can do with the support of more experienced others (Reiser and Tabak, 2014; Vygotsky and Cole, 1978), pedagogies preferred by learning scientists are centered around complex situations that are typically somewhat too challenging for learners themselves. Initially, the learners cannot master these problem-solving, design, or inquiry tasks without additional support. Of the various forms of potential additional instructional support that exist (e.g., feedback, additional information, additional incentives), the learning sciences community has put particular focus on the idea of scaffolding (Wood et al., 1976). Scaffolding refers to instructional support that enables learners to master tasks that are outside their reach without this support. Scaffolding is typically offered by more knowledgeable others or by (intelligent) computer systems. Wood et al. (1976) identified main functions of scaffolding which include sparking situational interest (*recruitment*, p. 98), reducing the complexity and difficulty of tasks (*reduction in degrees of freedom*, p. 98), keeping learners focused on their goal (*direction maintenance*, p. 98), (4) highlighting crucial features of a task (*marking critical features*, p. 98), (5) motivating disappointed learners (*frustration control*, p. 98), and (6) providing solutions and models of a task (*demonstration*, p. 98). Building on Bruner’s original work, several taxonomies of scaffolding functions emerged that differ mostly with respect to whether they address what is to be learned (e.g., scaffolding for joint sense-making or for building a hypothesis, e.g., Quintana et al., 2018) or whether they refer to how the learner is supported (e.g., prompting, hinting, channeling, Pea, 2004).

A joint assumption of scaffolding approaches is that the effectiveness of this kind of support depends on its contingency and time-liness. Scaffolding is considered to be temporary and needs to be removed (or faded) when the learner does not need it anymore because scaffolding can also be detrimental for learning (expertise reversal effect; Kalyuga, 2007). Thus, scaffolding requires a dynamic assessment of knowledge and skills, and consequential adaptive interventions—a complex interplay of guidance and self-regulation with respect to each and every individual learner (see Plass and Pawar, 2020). As this seems hardly possible in typical classrooms with only one or two teachers or tutors and 20 or more learners, technology and peers as sources of scaffolding increasingly became the focus of research (Tabak, 2004). Although theoretically straightforward, a recent meta-analysis found mixed evidence on the effects of instructionally adapting the scaffolding to the inferred needs of the learners (Belland et al., 2017).

There are several lines of research on scaffolding for individual learning (Belland et al., 2017). Scaffolding by using examples, so called example-based learning, is certainly one of the more systematically researched approaches to scaffolding (Renkl and Atkinson, 2003). In this approach, learners have the opportunity to follow someone else’s steps in solving a difficult problem, typically without further explanation of why these steps were taken. Then, the learners typically have to self-explain the steps by using domain knowledge (e.g., principles of probability calculus). This type of scaffolding example has been used effectively in several domains. In particular, studying and self-explaining examples was more effective than engaging in problem-solving activities for learners with little prior knowledge (see van Gog and Rummel, 2018).

Computer-Supported Collaborative Learning (CSCL) and its approaches to facilitating learning

CSCL is concerned with groups that coordinate, co-regulate, and engage in argumentation and knowledge building in problem-solving, inquiry, and design tasks in technology-rich environments. In CSCL research, technology conceptualizations include providing complex and dynamic contexts for these activities and associated pedagogies, as well as enabling and structuring peer collaboration of small and large groups. The *International Handbook of Computer-Supported Collaborative Learning* (Cress et al., 2021a,b) identified common ground of CSCL with other areas of the learning sciences but also differences with respect to phenomena, theories, and methods.

Collective knowledge construction within the learning sciences transcends a purely cognitivist perspective by building on key aspects of socio-cultural perspectives (e.g., Engeström, 2000; Vygotsky and Cole, 1978). Here, the group level as well as (group/cultural) artifacts, alongside individual learning are taken into consideration, with the aim to identify how to support both group and individual learning in the best way possible. Issues around *knowledge building* (see Scardamalia and Bereiter, 2021) have characterized CSCL research for many years. Conceptual and empirical research yielded design principles for learning environments in which learners in school create and share new knowledge for themselves, their class, or even the whole school through collaborative reasoning. The technology enabled the creation and linking of arguments, peer feedback on the arguments, and the revision of the

arguments. The approach has been taken up in many places around the world and further investigated through implementations in schools (see [Chan and van Aalst, 2018](#)) and detailed network analyses ([Oshima et al., 2012](#)). Some more recent theories of collective knowledge building ([Cress and Kimmerle, 2018](#); [Stahl, 2005](#)) describe and explain how groups of people, from small groups working on one project, to organizations developing several new projects, to large communities producing online artifacts, co-create new knowledge, both regarding their individual knowledge but, in particular, knowledge at the group or community level. The focus of collective knowledge construction theories lies in recognizing a multi-level and multi-agent structure that consists of individual learning and acting, as well as group learning, interaction processes, and artifacts (e.g., [Cress and Kimmerle, 2008](#)).

Often, learners' engagement in collaborative knowledge building and argumentation have been analyzed from different theoretical perspectives, reflecting the different disciplinary roots and multiple epistemological backgrounds of the learning sciences. These perspectives include socio-cognitive approaches which are mainly interested in individual learning and focus on discussion and argumentation as potentially facilitating contexts of learning (e.g., [Stegmann et al., 2012](#)); dialogic education which considers dialogic argumentation as a major cultural tool to shape individual thinking (e.g., [Philipson and Wegerif, 2016](#)), as well as approaches to collaborative argumentation mainly interested in understanding and advancing the human capacity to jointly generate knowledge and resolve conflicts (e.g., [Baker et al., 2021](#)).

An important discovery of research on collaborative learning has been that collaborative argumentation and problem solving among learners is not productive in a broad range of contexts because learners often fail to work together (e.g., [Barron, 2003](#); [Weinberger et al., 2010](#)). Two lines of research examine forms of support that make productive collaboration more likely.

- (1) *Awareness* research examines different forms of awareness support for self- and co-regulation in groups. For example, learners are provided with information about the prior knowledge of each collaboration partner or about the amount/quality of their own contributions compared to other learners in the group (cf. [Bodemer, 2018](#)). The effects of awareness support are mixed. It is likely that the extent to which this additional information can actually be used to improve collaborative problem-solving, inquiry, and design depends heavily on the self-regulation skills and learning preferences of the individuals and the group ([Jeong et al., 2019](#)). A promising new line of research uses dashboards to inform and support the teachers in their monitoring and scaffolding of computer-supported collaborative learners ([Tissenbaum and Slotta, 2019](#); [Van Leeuwen, 2015](#)).
- (2) *Scripting and orchestration* for groups: Research on CSCL has successfully broadened the concept of scaffolding individuals in groups and scaffolding groups as a whole (see [Kim et al., 2020](#)). Research on the design and the effects of collaboration scripts has emerged from an uptake and the extension of two complementary notions of the term script as a cognitive structure that enables understanding and acting ([Schank, 1999](#)) and as an instructional method to improve collaborative learning ([O'Donnell et al., 1987](#)). Through collaboration scripts, complex collaborative tasks are segmented into "scenes" and learners are supported by role specifications and, in some cases, engagement in individual subtasks such as formulating questions or arguments (see [Fischer et al., 2013](#)). Technology can be used for segmentation and coordination (e.g., [Manathunga and Hernández-Leo, 2018](#)) as well as to highlight a role, to provide just-in-time prompts and guidance, and to adapt the level of support (e.g., [Amarasinghe et al., 2019](#)). Collaboration scripts were found to be particularly effective for learning collaboration skills (e.g., engaging in dialogic argumentation in a discussion) and, to a somewhat lesser extent, for domain learning, and segmentation into scenes proved to be the most effective approach to scaffolding (see [Vogel et al., 2017](#)). A negative effect of structuring collaboration on motivation has not received support from the existing evidence (see [Radkowsitch et al., 2020](#)). An extension of the scripting idea was presented with the concept of orchestration. The concept of orchestration addresses the distribution and coordination of technology-supported activities on individual, group, and collective level ([Dillenbourg et al., 2018](#)), with orchestration graphs (e.g., [Prieto et al., 2018](#)) being a "notation system" for planning, enacting and analyzing multi-level, computer-supported activity systems.

Methods in the learning sciences

The methodologies and methods used in the learning sciences have been analyzed from several perspectives (e.g., [Sommerhoff et al., 2018](#); [Yoon and Hmelo-Silver, 2017](#)). According to these analyses, learning scientists have methodological roots in several disciplines, such as psychology, computer science, and science education. The different methods are often eclectically combined to address the analyses of complex processes in contexts like classrooms in schools and universities, laboratories, and informal learning contexts such as youth clubs, makerspaces ([Keune and Peppler, 2019](#)), and museums ([Lyons, 2018](#)).

The experimental paradigm has been expanded by an innovative combination with design methods coming from computer science. Design experiments ([Brown, 1992](#)) and later design research (see [McKenney and Reeves, 2021](#)) and design-based research ([Anderson and Shattuck, 2012](#); [Design-Based Research Collective, 2003](#); [Sandoval and Bell, 2004](#)) and design-based implementation research ([Fishman and Penuel, 2018](#)) have developed in response to the complexity of conditions that needed to be addressed when trying to investigate learning and teaching in authentic or real contexts. Design-based research is a cyclic methodology in which interventions are designed (and re-designed after testing their effects), mostly in collaboration with practitioners, to improve learning in real educational contexts with the equal goal to form and test theoretical advances toward understanding learning (see e.g., [Cobb et al., 2003](#); [DiSessa and Cobb, 2004](#)). One approach to advance theory on learning, in and through design research, is conjecture mapping ([Sandoval, 2014](#)), which originates from a higher-level idea about how learning could be facilitated in a particular context and moves across design conjectures (e.g., the design and how it is used) to theoretical conjectures (e.g., the outcomes

of learning and how they are observed with the design). Design-based Implementation Research (Fishman and Penuel, 2018) has developed as an extension to DBR to investigate and address the multiple questions arising from the implementation process of educational innovations in classrooms across longer periods of time and funding cycles.

In contrast to psychology and other disciplines investigating learning, units of analysis have been extended beyond the individual to include small groups or even classrooms, schools or large online communities (e.g., Stahl and Hakkarainen, 2021). These are addressed using a broad repertoire of methods, including cognitive, behavioral, and physiological measures, which either focus on individuals, interaction between individuals or their combination (see Sommerhoff et al., 2018). Besides DBR as the signature method of the learning sciences, diverse methods are employed. While case studies and classical experimentation are used frequently, methodological innovations are particularly widespread in the learning sciences. Qualitative and quantitative process analyses of “big data,” written and spoken language, of gestures, of artifacts, and of logfiles are used in these studies—and often so in combination in order to achieve a deep understanding of the mechanisms of learning and collaboration (Vogel and Weinberger, 2018) and the effects of technology for their support (Roschelle and Teasley, 1995; Koedinger and Aleven, 2007).

Think-aloud protocols and dialog analyses (see Vogel and Weinberger, 2018), as well as (social) network analyses (e.g., Oshima et al., 2012; Shaffer, 2017) are among the methods frequently used to analyze reasoning and learning in authentic settings. Logfile analyses have been used to more specifically target the interaction of learners with and within technology-enhanced learning environments. Increasingly, physiological measures like eye-tracking or heart rate variability are used to assess attentional processes and mental load in the learning process. Beyond these physiological measures, the learning sciences has been rather slow in making use of new methods from the neurosciences (see Varma et al., 2018).

Core and periphery. DBR is systematically taught in around 25% of learning sciences programs (Sommerhoff et al., 2018) and is the method of choice in an even smaller fraction of conference and journal contributions in the CSCL community within the learning sciences (Hmelo-Silver and Jeong, 2021a,b). However, the programs that teach DBR systematically seem to form a core of programs in the learning sciences community that are conceptually and methodologically more coherent, and relatively similar. Sommerhoff et al. (2018) suggest in their conclusions that this set of programs may be considered as the nucleus of an emerging discipline of the learning sciences. This nucleus is surrounded by orbiting programs that have strong roots in other disciplines. These programs in the periphery drive methodological innovation and shape the field of the learning sciences. Recent methodological developments with strong influence on the learning sciences include learning analytics (Lang et al., 2017), eye-tracking (Jarodzka et al., 2021), and multi-modal learning process analyses (Malmberg et al., 2019).

An outlook on advancing research on learning in the learning sciences

The learning sciences have produced and worked on a successful research program on individual and collaborative learning in authentic contexts in formal and informal settings over the past 40 years. This program includes new conceptual frameworks (e.g., collective knowledge construction, knowledge building, scaffolding) and new methods that were adopted from contributing disciplines and further developed to advance knowledge on learning (e.g., design-based research). Looking ahead, we highlight five fields of potential importance for the future of the learning sciences that can help to fill existing research gaps and to generate new insights to address pressing educational and societal issues. They span from (1) replicating prior findings, to (2) novel means of assessment and (3) conceptual and methodological innovation in future research, to (4) fundamental aspects related to educational equity within understanding learning and (5) personalized learning as a key focus of interest and goal within the learning sciences.

- (1) *Replicating research findings.* In its early years, research in the learning sciences was often descriptive with the aim to discover mechanisms and conditions rather than their validation through additional research. While the discoveries within the learning sciences contributed to novel perspectives on educational phenomena, topics such as replication and open science have hardly been addressed in the learning sciences. For example, there are only a few meta-analyses on learning sciences and CSCL research (e.g., Belland et al., 2017; Jeong et al., 2019) which leaves important programmatic papers (e.g., Wise and Schwarz, 2017) speculating on the state of research. The learning sciences have produced a wealth of new evidence through empirical studies. To advance the field, a more systematic examination of the robustness, replicability, and transferability of the discoveries is needed. It is currently still subject to discussion within the field whether and by which methods such investigations of robustness and transferability should be undertaken. It is conceivable that an expansion of research in the learning sciences that aims to confirm and evaluate findings across contexts is a prerequisite for systematic impact in related fields and practice (e.g., teacher education).
- (2) *Emphasis on assessment.* There is highly important work in the learning sciences that considers novel approaches to assessment that are aligned with educational perspectives central to the learning sciences (Hickey and Harris, 2021). Yet, it has not been a central focus of the learning sciences to develop new forms of assessment that are aligned with domain-specific and cross-domain knowledge and skills facilitated through problem-solving, inquiry, and design pedagogies. Better aligned assessment is needed to evaluate the effects of these pedagogies for different uses, including their formative use in classroom assessment to help students to learn, their summative use to determine individual achievement, and their use to evaluate specific interventions, programs, or institutions (Pellegrino, 2018). Given the strong focus on disciplinary knowledge and engagement (Herrenkohl and Polman, 2018) in the learning sciences, models of domain-specific competence development (Ufer and Neumann, 2018) and learning progressions (Alonzo, 2018) seem to provide good starting points for these new kinds

of assessment. Technology can play a new important role in assessment, beyond making the data collection more efficient. For example, in formative assessment, technology can be used to analyze log file data quickly to help teachers to give adaptive and process-related feedback to individuals and groups, and to adjust the level of scaffolding (e.g., van Leeuwen, 2015; Zhai et al., 2020). Ultimately, better means of assessment will be necessary for the learning sciences (i) to validate that the way learning is observed and supported within the learning sciences corresponds to their theoretical conceptions and (ii) to provide a basis for the further creation and confirmation of highly valid research findings.

- (3) *Appreciation of conceptual and methodological innovations from other disciplines.* Learning scientists differ on whether they believe they are a discipline with a common methodological core (Sommerhoff et al., 2018), an interdisciplinary group working on common projects requiring shared conceptual framework models (Wise and Schwarz, 2017), or a multidisciplinary community of researchers who mutually inspire each other through insights into their respective disciplinary research. It is likely that all three views apply, but to different groups of learning scientists (Sommerhoff et al., 2018) and that innovations from several related disciplines, which are being taken up and used, are an important driver of progress in the learning sciences. A clear example of this is the rapid development in the areas of data science in general, and educational data mining and learning analytics specifically (e.g., Romero and Ventura, 2020; Rosé, 2018; Teasley, 2019). Among others, similar innovations could be found, for example, in psychology with recent attempts to link the nomothetic (striving for general regularities) and the idiographic (explaining the individual case) research paradigms (e.g., Cook et al., 2018; Molenaar and Campbell, 2009). Taking up existing opportunities and generating new opportunities to monitor better the “innovations next door,” and to use them for learning sciences research on authentic contexts of learning and collaboration, will be of high importance for the learning sciences.
- (4) *Educational Equity.* Many learning scientists have grappled with questions of educational equity in recent years with the aim of understanding how forms of marginalization and representation relate to understanding learning and learning opportunities (e.g., Politics of Learning Writing Collective, 2017; Bang and Vossoughi, 2016; Esmonde and Booker, 2016; Uttamchandani et al., 2020). Within the work of educational equity, design has been considered as a foundational approach toward transforming social and community realities that, along the traditional aim of educational design research, contribute to theoretical approaches regarding equitable learning opportunities (Gutiérrez and Jurow, 2016). Among other things, researchers in this area ask to what extent learning materials, curricula, and pedagogies interact with learner characteristics associated with race, ethnicity, and gender. These interaction effects can systematically create and perpetuate disadvantage and the erasure of general learning effects. Moreover, knowledge about these interaction effects can also be used to support learning in specific conditions with specific measures. Empirical work related to educational equity within the learning sciences increasingly focuses on STEM and computer science learning, where diversified representation, restorative approaches, and justice-centered orientation promise more social sustainability also through the transformation of these domains (e.g., Barton et al., 2017; Vossoughi et al., 2013). Work in this area has focused on, for example, indigenous knowledges (e.g., Tzou et al., 2019; Kafai et al., 2014) as well as racial and ethnic representation (e.g., Tofel-Grehl and Searle, 2017; Holbert et al., 2020). Moreover, also LGBTQ+ issues (e.g., Uttamchandani et al., 2020) and gender equity (e.g., Buchholz et al., 2014; Keune et al., 2019) have been increasingly addressed. A crosscutting focus of this work has been transforming domain areas, inviting diverse cultural practices, and identifying relevance of cultural and community practices. For example, work with indigenous families has focused on opening up STEM practices to contribute to the socio-historical and community practices of nondominant and marginalized people in these fields through intergenerational design and storytelling (e.g., Tzou et al., 2019). While these issues have, to date, predominantly been identified within North American contexts, they have been increasingly taken as exemplary for raising pressing questions about educational equity in other countries and for addressing them with empirical research (e.g., Bang, 2019). The broader international interest in educational equity within the learning sciences has been substantiated with the formation of the ISLS Equity and Justice committee that is charged to identify and address “issues related to diverse representation, the systemic marginalization or erasure of particular voices, and justice-oriented scholarship” (ISLS, <https://www.isls.org/members/committees/>).
- (5) *Personalized learning.* Using technology to adapt the content to be learned and the instructional support to each individual learner’s current needs, both on micro and macro time scales, and to do this in meaningful social contexts is a vision in current approaches to personalized learning (see Plass and Pawar, 2020). Research projects are currently underway across the world that focus on the role of artificial intelligence in education more generally and in facilitating learning processes more specifically. The learning sciences have had a head start with decades of research on ITS, scaffolding, and CSCL in general. Strong theorizing about learning mechanisms and processes in individual and collaborative learning has made studies in the learning sciences amenable to data scientists, even beyond the often-employed well-structured domains. The algorithms used, and rapidly evolving, in learning analytics are often far superior in predicting learning outcomes based on multiple learning process measures than the mostly pure linear regression analyses that have been common in practice within the learning sciences and related educational research until now. Moreover, these approaches resonate, both with calls for educational equity and individual assessment that allow to capture processes and products as well as knowledge and skills. While there is initial evidence for positive effects of systematically employing these AI-based approaches in theory-based adaptive scaffolding environments, there are still many challenges waiting to be tackled by interdisciplinary teams of experts from the learning sciences and learning analytics (Rosé, 2018).

References

- Abrahamson, D., 2009. Embodied design: constructing means for constructing meaning. *Educ. Stud. Math.* 70 (1), 27–47.
- Ainsworth, S., 1999. The functions of multiple representations. *Comput. Educ.* 33 (2–3), 131–152.
- Ainsworth, S., 2018. Multiple representations and multimedia learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 96–105.
- Alibali, M.W., Nathan, M.J., 2012. Embodiment in mathematics teaching and learning: evidence from learners' and teachers' gestures. *J. Learn. Sci.* 21 (2), 247–286.
- Alonzo, A.C., 2018. An argument for formative assessment with science learning progressions. *Appl. Meas. Educ.* 31 (2), 104–112.
- Amarasinghe, I., Hernández-Leo, D., Jonsson, A., 2019. Data-informed design parameters for adaptive collaborative scripting in across-spaces learning situations. *User Model. User Adapt. Interact.* 29 (4), 869–892.
- Anderson, T., Shattuck, J., 2012. Design-based research: a decade of progress in education research? *Educ. Res.* 41 (1), 16–25.
- Anderson, J.R., Boyle, C.F., Reiser, B.J., 1985. Intelligent tutoring systems. *Science* 228 (4698), 456–462.
- Baker, M.J., Schwarz, B.B., Ludvigsen, S.R., 2021. Educational dialogues and computer supported collaborative learning: critical analysis and research perspectives. *Int. J. Comput. Support. Collab. Learn.* 16, 583–604.
- Bang, M., Vossoughi, S., 2016. Participatory design research and educational justice: studying learning and relations within social change making. *Cognit. Instr.* 34 (3), 73–193.
- Bang, M., 2019. *Storywork Across the Landscapes of Home and School: Towards Indigenous Futures in Thailand* (doctoral dissertation). University of Washington, Seattle. <http://hdl.handle.net/1773/45175>.
- Barron, B., 2003. When smart groups fail. *J. Learn. Sci.* 12 (3), 307–359.
- Barton, A.C., Tan, E., Greenberg, D., 2017. The makerspace movement: sites of possibilities for equitable opportunities to engage underrepresented youth in STEM. *Teach. Coll. Rec.* 119 (6), 1–44.
- Belland, B.R., Walker, A.E., Kim, N.J., Lefler, M., 2017. Synthesizing results from empirical research on computer-based scaffolding in STEM education: a meta-analysis. *Rev. Educ. Res.* 87 (2), 309–344.
- Bobrow, D.G., Collins, A., 1975. *Representation and Understanding: Studies in Cognitive Science*. Academic Press, New York.
- Bodemer, D., Janssen, J., Schnaubert, L., 2018. Group awareness tools for computer-supported collaborative learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 351–358.
- Brown, J.S., Collins, A., Duguid, P., 1989. Situated cognition and the culture of learning. *Educ. Res.* 18 (1), 32–42.
- Brown, A.L., 1992. Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2 (2), 141–178.
- Buchholz, B., Shively, K., Pepler, K., Wohlwend, K., 2014. Hands on, hands off: gendered access in crafting and electronics practices. *Mind Cult. Activ.* 21 (4), 278–297.
- Carraher, T.N., Carraher, D.W., Schliemann, A.D., 1985. Mathematics in the streets and in schools. *Br. J. Dev. Psychol.* 3 (1), 21–29.
- Chan, C.K., van Aalst, J., 2018. Knowledge building: theory, design, and analysis. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 295–307.
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., Fischer, F., 2020. Simulation-based learning in higher education: a meta-analysis. *Rev. Educ. Res.* 90 (4), 499–541.
- Clark, D.B., Tanner-Smith, E.E., Killingsworth, S.S., 2016. Digital games, design, and learning: a systematic review and meta-analysis. *Rev. Educ. Res.* 86 (1), 79–122.
- Clegg, T., Kolodner, J., 2014. Scientizing and cooking: helping middle-school learners develop scientific dispositions. *Sci. Educ.* 98 (1), 36–63.
- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., Schauble, L., 2003. Design experiments in educational research. *Educ. Res.* 32 (1), 9–13.
- Cognition and Technology Group at Vanderbilt, 1992. The Jasper series as an example of anchored instruction: theory, program description, and assessment data. *Educ. Psychol.* 27 (3), 291–315.
- Cook, C.R., Kilgus, S.P., Burns, M.K., 2018. Advancing the science and practice of precision education to enhance student outcomes. *J. Sch. Psychol.* 66, 4–10.
- Cress, U., Kimmerle, J., 2008. A systemic and cognitive view on collaborative knowledge building with wikis. *Int. J. Comput. Support. Collab. Learn.* 3 (2), 105–122.
- Cress, U., Kimmerle, J., 2018. Collective knowledge construction. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 137–146.
- Cress, U., Oshima, J., Rosé, C., Wise, A.F., 2021a. Foundations, processes, technologies, and methods: an overview of CSCL through its handbook. In: Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham, pp. 3–22.
- Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), 2021b. *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham.
- Cronbach, L.J., Snow, R.E., 1977. *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. Ardent Media US, Irvington.
- Danish, J.A., Gresalfi, M., 2018. Cognitive and sociocultural perspective on learning: tensions and synergy in the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 34–43.
- Danish, J.A., 2014. Applying an activity theory lens to designing instruction for learning about the structure, behavior, and function of a honeybee system. *J. Learn. Sci.* 23 (2), 100–148.
- De Jong, T., Van Joolingen, W.R., 1998. Scientific discovery learning with computer simulations of conceptual domains. *Rev. Educ. Res.* 68 (2), 79–201.
- De Jong, T., Linn, M.C., Zacharia, Z.C., 2013. Physical and virtual laboratories in science and engineering education. *Science* 340 (6130), 305–308.
- Design-Based Research Collective, 2003. Design-based research: an emerging paradigm for educational inquiry. *Educ. Res.* 32 (1), 5–8.
- Dillenbourg, P., Prieto, L.P., Olsen, J.K., 2018. Classroom orchestration. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 180–190.
- DiSessa, A.A., Cobb, P., 2004. Ontological innovation and the role of theory in design experiments. *J. Learn. Sci.* 13 (1), 77–103.
- Dochy, F., Segers, M., Van den Bossche, P., Gijbels, D., 2003. Effects of problem-based learning: a meta-analysis. *Learn. Instr.* 13 (5), 533–568.
- Donnelly, D.F., Linn, M.C., Ludvigsen, S., 2014. Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Rev. Educ. Res.* 84 (4), 572–608.
- Eberle, J., 2018. Apprenticeship learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 44–53.
- Eden, G.F., Jones, K.M., Cappell, K., Gareau, L., Wood, F.B., Zeffiro, T.A., Dietz, N.A., Agnew, J.A., Flowers, D.L., 2004. Neural changes following remediation in adult developmental dyslexia. *Neuron* 44 (3), 411–422.
- Engeström, Y., 2000. Activity theory and the social construction of knowledge: a story of four umpires. *Organization* 7 (2), 301–310.
- Engle, R.A., Conant, F.R., 2002. Guiding principles for fostering productive disciplinary engagement: explaining an emergent argument in a community of learners classroom. *Cognit. Instr.* 20 (4), 399–483.
- Enyedy, N., Yoon, S., 2021. Immersive environments: learning in augmented + virtual reality. In: Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham, pp. 389–405.
- Esmonde, I., Booker, A.N., 2016. *Power and Privilege in the Learning Sciences*. Routledge, New York.
- Evans, M.A., Packer, M.J., Sawyer, R.K. (Eds.), 2016. *Reflections on the Learning Sciences*. Cambridge University Press, Cambridge, UK.
- Fields, D.A., Kafai, Y.B., 2018. Games in the learning sciences: reviewing evidence from playing and making games for learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 276–284.

- Fischer, F., Kollar, I., Stegmann, K., Wecker, C., 2013. Toward a script theory of guidance in computer-supported collaborative learning. *Educ. Psychol.* 48 (1), 56–66.
- Fischer, F., Goldman, S.R., Hmelo-Silver, C.E., Reimann, P., 2018a. Introduction: evolution of research in the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 1–8.
- Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), 2018b. *International Handbook of the Learning Sciences*. Routledge, New York.
- Fishman, B., Penuel, W., 2018. Design-based implementation research. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 393–400.
- Goldin-Meadow, S., 2005. *Hearing Gesture: How Our Hands Help Us Think*. Harvard University Press, Cambridge, MA.
- Goldman, S.V., Kabayadondo, Z. (Eds.), 2016. *Taking Design Thinking to School*. Routledge, New York.
- Graesser, A., McNamara, D., 2010. Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educ. Psychol.* 45 (4), 234–244.
- Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W., Hesse, F.W., 2018a. Advancing the science of collaborative problem solving. *Psychol. Sci. Publ. Interest* 19 (2), 59–92.
- Graesser, A.C., Hu, X., Sottile, R., 2018b. Intelligent tutoring systems. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 246–255.
- Greeno, J.G., 2011. A situative perspective on cognition and learning in interaction. In: Koschmann, T. (Ed.), *Theories of Learning and Studies of Instructional Practice*. Springer, New York, pp. 41–71.
- Gutiérrez, K.D., Jurow, A.S., 2016. Social design experiments: toward equity by design. *J. Learn. Sci.* 25 (4), 565–598.
- Halverson, E., Peppler, K., 2018. The maker movement and learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 285–294.
- Herrenkohl, L.R., Polman, J.L., 2018. Learning within and beyond the disciplines. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 106–115.
- Hickey, D., Harris, T., 2021. Reimagining online grading, assessment, and testing using situated cognition. *Dist. Educ.* 42 (2), 290–309.
- Hmelo-Silver, C.E., Barrows, H.S., 2015. Problem-based learning: goals for learning and strategies for facilitating. In: Walker, A., Leary, H., Hmelo-Silver, C.E. (Eds.), *Essential Readings in Problem-Based Learning: Exploring and Extending the Legacy of Howard S. Barrows*. Purdue University Press, West Lafayette, pp. 69–84.
- Hmelo-Silver, C.E., Jeong, H., 2021a. An overview of CSCL methods. In: Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham, pp. 65–83.
- Hmelo-Silver, C.E., Jeong, H., 2021b. Benefits and challenges of interdisciplinarity in CSCL research: a view from the literature. *Front. Psychol.* 11.
- Hmelo-Silver, C.E., Kapur, M., Hamstra, M., 2018. Learning through problem solving. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 210–220.
- Hoadley, C., 2018. A short history of the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 11–23.
- Holbert, N., Dando, M., Correa, I., 2020. Afrofuturism as critical constructionist design: building futures from the past and present. *Learn. Media Technol.* 45 (4), 328–344.
- Hutchins, E., 1995. How a cockpit remembers its speeds. *Cognit. Sci.* 19, 265–288.
- International Society of the Learning Sciences, 2009. The International Society of the Learning Sciences. Retrieved from: https://www.isls.org/wp-content/uploads/2020/06/ISLS_Vision_2009.pdf.
- Jarodzka, H., Skuballa, I., Gruber, H., 2021. Eye-tracking in educational practice: investigating visual perception underlying teaching and learning in the classroom. *Educ. Psychol. Rev.* 33 (1), 1–10.
- Jeong, H., Hmelo-Silver, C.E., Jo, K., 2019. Ten years of computer-supported collaborative learning: a meta-analysis of CSCL in STEM education during 2005–2014. *Educ. Res. Rev.* 28, 100284.
- Kafai, Y., Searle, K., Martinez, C., Brayboy, B., 2014. Ethnocomputing with electronic textiles: culturally responsive open design to broaden participation in computing in American Indian youth and communities. In: Nagel, K., Dougherty, J.D. (Eds.), *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*. ACM, Atlanta, GA, pp. 241–246.
- Kafai, Y.B., 2012. *Minds in Play: Computer Game Design as a Context for Children's Learning*. Routledge, New York.
- Kalyuga, S., 2007. Expertise reversal effect and its implications for learner-tailored instruction. *Educ. Psychol. Rev.* 19 (4), 509–539.
- Keune, A., Peppler, K., 2019. Materials-to-develop-with: the making of a makerspace. *Br. J. Educ. Technol.* 50 (1), 280–293.
- Keune, A., Peppler, K., Wohliwend, K., 2019. Recognition in makerspaces: supporting opportunities for women to “make” a STEM career. *Comput. Hum. Behav.* 99, 368–380.
- Kim, N.J., Belland, B.R., Lefler, M., Andreasen, L., Walker, A., Axelrod, D., 2020. Computer-based scaffolding targeting individual versus groups in problem-centered instruction for STEM education: meta-analysis. *Educ. Psychol. Rev.* 32, 415–461.
- Koedinger, K.R., Aleven, V., 2007. Exploring the assistance dilemma in experiments with cognitive tutors. *Educ. Psychol. Rev.* 19 (3), 239–264.
- Kolodner, J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., Ryan, M., 2003. Problem-based learning meets case-based reasoning in the middle-school science classroom: putting learning by design (tm) into practice. *J. Learn. Sci.* 12, 495–547.
- Kolodner, J.L., 2004. The learning sciences: past, present, future. *Educ. Technol.* 44 (3), 34–40.
- Kumar, R., Kim, J., 2014. Special issue on intelligent support for learning in groups. *Int. J. Artif. Intell. Educ.* 24 (1), 1–7.
- Lang, C., Siemens, G., Wise, A., Gasevic, D. (Eds.), 2017. *Handbook of Learning Analytics*. Society for Learning Analytics and Research (SOLAR), New York, NY.
- Laurillard, D., Alexopoulou, E., James, B., Bottino, R.M., Bouhineau, D., Chiocciariello, A., Correia, S., Davey, P., Derry, J., Dettori, G., Dirckinck-Holmfeld, L., 2007. *The Kaleidoscope Scientific Vision for Research in Technology Enhanced Learning*. TEARN Open Archive. Retrieved from: <https://telearn.archives-ouvertes.fr/hal-00190011>.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, MA.
- Lee, V., 2017. A short history of the learning sciences. Foundations of learning and instructional design technology. In: West, R.E. (Ed.), *Foundations of Learning and Instructional Design Technology—Historical Roots and Current Trends*. Ch. 4. EdTech Books. Retrieved from: https://edtechbooks.org/lidtfoundations/history_of_learning_sciences.
- Linn, M.C., Clark, D., Slotta, J.D., 2003. WISE design for knowledge integration. *Sci. Educ.* 87 (4), 517–538.
- Litts, B.K., Searle, K.A., Brayboy, B.M., Kafai, Y.B., 2021. Computing for all?: examining critical biases in computational tools for learning. *Br. J. Educ. Technol.* 52 (2), 842–857.
- Ludvigsen, S., Nerland, M., 2018. Learning at work. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, NY, pp. 147–156.
- Lyons, L., 2018. Supporting informal STEM learning with technological exhibits: an ecosystemic approach. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, NY, pp. 234–245.
- Ma, J.Y., 2016. Designing disruptions for productive hybridity: the case of walking scale geometry. *J. Learn. Sci.* 25 (3), 335–371.
- Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., Siipio, A., 2019. Going beyond what is visible: what multichannel data can reveal about interaction in the context of collaborative learning? *Comput. Hum. Behav.* 96, 235–245.
- Manathunga, K., Hernández-Leo, D., 2018. Authoring and enactment of mobile pyramid-based collaborative learning activities. *Br. J. Educ. Technol.* 49 (2), 262–275.
- Mayer, R.E., 2014. Cognitive theory of multimedia learning. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, Cambridge, UK, pp. 43–71.
- McKenney, S., Reeves, T.C., 2021. Educational design research: portraying, conducting, and enhancing productive scholarship. *Med. Educ.* 55 (1), 82–92.
- Miller, G.A., 2003. The cognitive revolution: a historical perspective. *Trends Cognit. Sci.* 7 (3), 141–144.
- Molenaar, P.C., Campbell, C.G., 2009. The new person-specific paradigm in psychology. *Curr. Dir. Psychol. Sci.* 18 (2), 112–117.

- Nathan, M.J., Alibali, M., 2010. Learning sciences. *Wiley Interdiscip. Rev.* 1 (3), 329–345.
- Nathan, M.J., Rummel, N., Hay, K.E., 2016. Growing the learning sciences: brand or big tent. In: Evans, M.A., Packer, M.J., Sawyer, R.K. (Eds.), *Reflections on the Learning Sciences*. Cambridge University Press, Cambridge, UK, pp. 191–209.
- Newell, A., 1994. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA.
- Norman, D.A., 1993. Cognition in the head and in the world: an introduction to the special issue on situated action. *Cognit. Sci.* 17 (1), 1–6.
- O'Donnell, A.M., Dansereau, D.F., Hall, R.H., Rocklin, T.R., 1987. Cognitive, social/affective, and metacognitive outcomes of scripted cooperative learning. *J. Educ. Psychol.* 79 (4), 431–437.
- Oshima, J., Oshima, R., Matsuzawa, Y., 2012. Knowledge building discourse explorer: a social network analysis application for knowledge building discourse. *Educ. Technol. Res. Dev.* 60 (5), 903–921.
- Papert, S., Harel, I., 1991. Situating constructionism. *Constructionism* 36, 1–11.
- Papert, S.A., 1980. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic books, New York.
- Pea, R.D., Linn, M.C., 2020. Personal perspectives on the emergence of the learning sciences: 1970s–2005. *Front. Educ.* 5, 130.
- Pea, R.D., 2004. The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *J. Learn. Sci.* 13 (3), 423–451.
- Pedaste, M., Mäeots, M., Siiman, L.A., De Jong, T., Van Riesen, S.A., Kamp, E.T., Manoli, C.C., Zacharia, Z.C., Tsourlidaki, E., 2015. Phases of inquiry-based learning: definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61.
- Pellegrino, J.W., 2018. Assessment of and for learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 10–421.
- Peppler, K., Glosson, D., 2013. Stitching circuits: learning about circuitry through e-textile materials. *J. Sci. Educ. Technol.* 22 (5), 751–763.
- Peppler, K., Halverson, E., Kafai, Y.B. (Eds.), 2016. *Makeology: Makerspaces as Learning Environments*, vol. 1. Routledge, New York.
- Phillipson, N., Wegerif, R., 2016. *Dialogic Education: Mastering Core Concepts Through Thinking Together*. Routledge, New York.
- Plass, J.L., Pawar, S., 2020. Toward a taxonomy of adaptivity for learning. *J. Res. Technol. Educ.* 52 (3), 275–300.
- Politics of Learning Writing Collective, 2017. The learning sciences in a new era of US nationalism. *Cognit. Instr.* 35 (2), 91–102.
- Postle, B.R., 2020. *Essentials of Cognitive Neuroscience*. John Wiley & Sons, Hoboken, NJ.
- Prieto, L.P., Sharma, K., Kidzinski, E., Rodríguez-Triana, M.J., Dillenbourg, P., 2018. Multimodal teaching analytics: automated extraction of orchestration graphs from wearable sensor data. *J. Comput. Assist. Learn.* 34 (2), 193–203.
- Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G., Kyza, E., Edelson, D., Soloway, E., 2018. A scaffolding design framework for software to support science inquiry. *J. Learn. Sci.* 13 (3), 337–386.
- Radkowsch, A., Vogel, F., Fischer, F., 2020. Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *Int. J. Comput. Support. Collab. Learn.* 15 (1), 5–47.
- Reimann, P., Markauskaite, L., 2018. Expertise. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 54–63.
- Reiser, B.J., Tabak, I., 2014. Scaffolding. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, Cambridge UK, pp. 44–62.
- Renkl, A., Atkinson, R.K., 2003. Structuring the transition from example study to problem solving in cognitive skill acquisition: a cognitive load perspective. *Educ. Psychol.* 38 (1), 15–22.
- Romero, C., Ventura, S., 2020. Educational data mining and learning analytics: an updated survey. *Wiley Interdiscip. Rev.* 10 (3), e1355.
- Roque, R., 2016. Family creative learning. In: Peppler, K., Halverson, E., Kafai, Y.B. (Eds.), *Makeology: Makerspaces as Learning Environments*, vol. 1. Routledge, New York, pp. 47–63.
- Roschelle, J., Teasley, S.D., 1995. The construction of shared knowledge in collaborative problem solving. In: O'Malley, C. (Ed.), *Computer Supported Collaborative Learning*. Springer, Heidelberg, pp. 69–97.
- Roschelle, J., 1998. Activity theory: a foundation for designing learning technology? *J. Learn. Sci.* 7 (2), 241–255.
- Rosé, C.P., 2018. Learning analytics in the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 511–519.
- Rummel, N., Walker, E., Aleven, V., 2016. Different futures of adaptive collaborative learning support. *Int. J. Artif. Intell. Educ.* 26 (2), 784–795.
- Sandoval, W.A., Bell, P., 2004. Design-based research methods for studying learning in context: Introduction. *Educ. Psychol.* 39 (4), 199–201.
- Sandoval, W., 2014. Conjecture mapping: an approach to systematic educational design research. *J. Learn. Sci.* 23 (1), 18–36.
- Sawyer, R.K. (Ed.), 2005. *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, Cambridge, UK.
- Sawyer, R.K. (Ed.), 2014. *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, Cambridge, UK.
- Scardamalia, M., Bereiter, C., 1991. Higher levels of agency for children in knowledge building: a challenge for the design of new knowledge media. *J. Learn. Sci.* 1 (1), 37–68.
- Scardamalia, M., Bereiter, C., 2021. Knowledge building: advancing the state of community knowledge. In: Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham, pp. 261–279.
- Schank, R.C., 1999. *Dynamic Memory Revisited*. Cambridge University Press, Cambridge, UK.
- Schmidt, H.G., Rotgans, J.I., Yew, E.H., 2011. The process of problem-based learning: what works and why. *Med. Educ.* 45 (8), 792–806.
- Shaffer, D.W., 2017. *Quantitative Ethnography*. Lulu.com, Morrisville, NC.
- Shapiro, B.R., Garner, B., 2021. Classroom interaction geography: visualizing space & time in classroom interaction. *J. Res. Technol. Educ.* 1–15.
- Shapiro, L., 2019. *Embodied Cognition*. Routledge, New York.
- Sleeman, D., Brown, J.S. (Eds.), 1982. *Intelligent Tutoring Systems*. Academic Press, New York.
- Slotta, J.D., Quintana, R.M., Moher, T., 2018. Collective inquiry in communities of learners. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 308–317.
- Sommerhoff, D., Szameitat, A., Vogel, F., Chernikova, O., Loderer, K., Fischer, F., 2018. What do we teach when we teach the learning sciences? A document analysis of 75 graduate programs. *J. Learn. Sci.* 27 (2), 319–351.
- Sommerhoff, D., Kollar, I., Ufer, S., 2021. Supporting mathematical argumentation and proof skills: comparing the effectiveness of a sequential and a concurrent instructional approach to support resource-based cognitive skills. *Front. Psychol.* 11, 3664.
- Sousa, D.A. (Ed.), 2010. *Mind, Brain, & Education: Neuroscience Implications for the Classroom*. Solution Tree Press, Bloomington.
- Stahl, G., Hakkarainen, K., 2021. Theories of CSCL. In: Cress, U., Wise, A., Rosé, C., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, Cham, pp. 23–43.
- Stahl, G., Hesse, F., 2006. ijCSCL—a journal for research in CSCL. *Int. J. Comput. Support. Collab. Learn.* 1 (1), 3–7.
- Stahl, G., 2005. Group cognition in computer-assisted collaborative learning. *J. Comput. Assist. Learn.* 21 (2), 79–90.
- Stegmann, K., Wecker, C., Weinberger, A., Fischer, F., 2012. Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instr. Sci.* 40 (2), 297–323.
- Svihla, V., Reeve, R. (Eds.), 2016. *Design as Scholarship: Case Studies From the Learning Sciences*. Routledge, New York.
- Sweller, J., Van Merriënboer, J.J., Paas, F.G., 1998. Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10 (3), 251–296.
- Tabak, I., 2004. Synergy: a complement to emerging patterns of distributed scaffolding. *J. Learn. Sci.* 13 (3), 305–335.

- Teasley, S.D., 2019. Learning analytics: where information science and the learning sciences meet. *Inf. Learn. Sci.* 120 (1/2), 59–73.
- Tissenbaum, M., Slotta, J., 2019. Supporting classroom orchestration with real-time feedback: a role for teacher dashboards and real-time agents. *Int. J. Comput. Support. Collab. Learn.* 14 (3), 325–351.
- Tofel-Grehl, C., Searle, K., 2017. Critical reflections on teacher conceptions of race as related to the effectiveness of science learning. *J. Multicult. Aff.* 2 (1), 4.
- Tzou, C., Meixi, Suárez, E., Bell, P., LaBonte, D., Starks, E., Bang, M., 2019. Storywork in STEM-art: making, materiality and robotics within everyday acts of Indigenous presence and resurgence. *Cognit. Instr.* 37 (3), 306–326.
- Ufer, S., Neumann, K., 2018. Measuring competencies. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 433–443.
- Uttamchandani, S., Bhimdiwala, A., Hmelo-Silver, C.E., 2020. Finding a place for equity in CSCL: ambitious learning practices as a lever for sustained educational change. *Int. J. Comput. Support. Collab. Learn.* 15 (3), 373–382.
- van Gog, T., Rummel, N., 2018. Example-based learning. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 201–209.
- van Horne, K., Bell, P., 2017. Youth disciplinary identification during participation in contemporary project-based science investigations in school. *J. Learn. Sci.* 26 (3), 437–476.
- van Leeuwen, A., 2015. Learning analytics to support teachers during synchronous CSCL: balancing between overview and overload. *J. Learn. Anal.* 2 (2), 138–162.
- van Merriënboer, J.J., Kirschner, P.A., 2018. 4C/ID in the context of instructional design and the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 169–179.
- Varma, S., McCandliss, B.D., Schwartz, D.L., 2008. Scientific and pragmatic challenges for bridging education and neuroscience. *Educ. Res.* 37 (3), 140–152.
- Varma, S., Im, S.-H., Schmied, A., Michel, K., Varma, K., 2018. Cognitive neuroscience foundations for the learning sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 64–74.
- Voeet, M., De Wever, B., 2016. History teachers' conceptions of inquiry-based learning, beliefs about the nature of history, and their relation to the classroom context. *Teach. Teach. Educ.* 55, 57–67.
- Vogel, F., Weinberger, A., 2018. Quantifying qualities of collaborative learning processes. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 500–510.
- Vogel, F., Wecker, C., Kollar, I., Fischer, F., 2017. Socio-cognitive scaffolding with computer-supported collaboration scripts: a meta-analysis. *Educ. Psychol. Rev.* 29 (3), 477–511.
- Vossoughi, S., Escudé, M., Kong, F., Hooper, P., 2013. Tinkering, learning & equity in the after-school setting. In: *Proceedings of the Annual FabLearn Conference*. Stanford University, Palo Alto, CA.
- Vygotsky, L.S., Cole, M., 1978. *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Weinberger, A., Stegmann, K., Fischer, F., 2010. Learning to argue online: scripted groups surpass individuals (unscripted groups do not). *Comput. Hum. Behav.* 26 (4), 506–515.
- Wenger, E., 1987. *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge*. Morgan Kaufmann, San Francisco, CA.
- Wenger, E., 1999. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge UK.
- Wise, A.F., Schwarz, B.B., 2017. Visions of CSCL: eight provocations for the future of the field. *Int. J. Comput. Support. Collab. Learn.* 12 (4), 423–467.
- Wood, D., Bruner, J.S., Ross, G., 1976. The role of tutoring in problem solving. *JCPP* 17 (2), 89–100.
- Yoon, S.A., Hmelo-Silver, C.E., 2017. What do learning scientists do? A survey of the ISLS membership. *J. Learn. Sci.* 26 (2), 167–183.
- Zhai, X., Yin, Y., Pellegrino, J.W., Haudek, K.C., Shi, L., 2020. Applying machine learning in science assessment: a systematic review. *Stud. Sci. Educ.* 56 (1), 111–151.

Relevant websites

Homepage of the International Society of the Learning Sciences (ISLS), <https://www.isls.org/>.
 International Journal of Computer Supported Collaborative Learning, <https://www.springer.com/journal/11412pxn>.
 Committees of the ISLS, <https://www.isls.org/members/committees/>.
 Journal of the Learning Sciences, <https://www.tandfonline.com/toc/hlms20/current>.
 Network of Academic Programs in the Learning Sciences (NAPLeS), <https://www.isls.org/naples/>.

Perspectives on learning from neuroscience

J. Massonnié^a and M.S.C. Thomas^{b, a} School of Education and Sociology, Faculty of Humanities & Social Sciences, University of Portsmouth, Portsmouth, United Kingdom; and ^b Center for Educational Neuroscience, Department of Psychological Sciences, Birkbeck University of London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Defining educational neuroscience	57
Key research domains in educational neuroscience	59
Brain health fundamentals	59
Nutrition	59
Exercise	59
Sleep	59
Protection from stress	59
Social and emotional learning	59
Executive functions	60
Socioeconomic status and educational outcomes	61
Individual differences in educational outcomes: genetics	61
Translating educational neuroscience research to education	62
What does neuroscience bring above and beyond what educators and psychologists already know?	62
Educational values are not defined by educational neuroscience	62
Collaborative projects between scientists, educators and learners are essential	63
Where does the future of educational neuroscience lie?	63
Training researchers for interdisciplinarity	63
Involving learners in the research process	63
Conclusions	64
References	64

Defining educational neuroscience

The field of educational neuroscience has been growing since the “Decade of the brain” in the 90s, spurred by the progress in neuro-imaging methods (Pasquinelli, 2011). Educational neuroscience seeks to understand how we *learn* by looking at our behaviors, cognition, emotions and their underlying brain mechanisms. Some authors prefer to use the denomination “science of learning” (Howard-Jones et al., 2018; npj Science of Learning: <https://npjsciencelearningcommunity.nature.com/>) to highlight the evidence-based approach to learning that is at the core of educational neuroscience.

Educational neuroscience combines multiple disciplines such as biology, genetics, neuroscience, psychology, philosophy and ethology (depicted in Fig. 1). The denomination ‘Mind, Brain and Education’ particularly emphasizes the contribution of psychology and neuroscience to education. Current textbooks of educational neuroscience (Thomas et al., 2020), journals (Mind, Brain, and Education; Trends in Neuroscience and Education), research centers and conferences (Center for Educational Neuroscience; International, Mind, Brain and Education Society) embrace this interdisciplinarity. A wide diversity of methods and levels of investigation are being employed, from single-cell studies, to large-scale randomized controlled trials in classrooms. To cite but a few methods, behavioral geneticists employ genome analyses and twin studies to study genetic influences on educational achievement. Biologists and neuroscientists use Magnetic Resonance Imaging, Electroencephalography and Near-infrared Spectroscopy to study processes of change in the brain. Psychologists design behavioral studies, experiments and randomized controlled trials, to try to unpack which factors influence human learning. They also conduct observational and qualitative studies to enable an in-depth understanding of people’s attitudes, thoughts and emotions, all of which influencing the acquisition of skills and knowledge. Ethologists observe animal behaviors, because studies comparing humans, other primates and animal species allow us to understand learning in the context of evolution. Philosophers contribute to debates on how levels of description (i.e., cell, brain, individual, ecosystem) relate to each other. They propose thought experiments and challenging questioning on theoretical, methodological and ethical aspects of educational neuroscience.

Educational neuroscience can be seen as a multifaceted field where all these levels of description bring a unique perspective while contributing to a more global and complex understanding of the phenomenon of learning. Mirroring the diversity in the levels of description is the diversity in the scales of investigation. Some of the research does not involve any participants. That is the case when reflecting on the ethics of cognitive enhancement, or when debating the theoretical definition of “learning” or “thinking processes”. In contrast, some international studies, such as those investigating how phonics training aids literacy acquisition across different languages, involve thousands of participants. In between these two extremes, there are case studies with two or three participants, small scale qualitative focus groups, behavioral studies involving hundreds of participants, or large-scale randomized controlled trials. Each scale of study is informative.

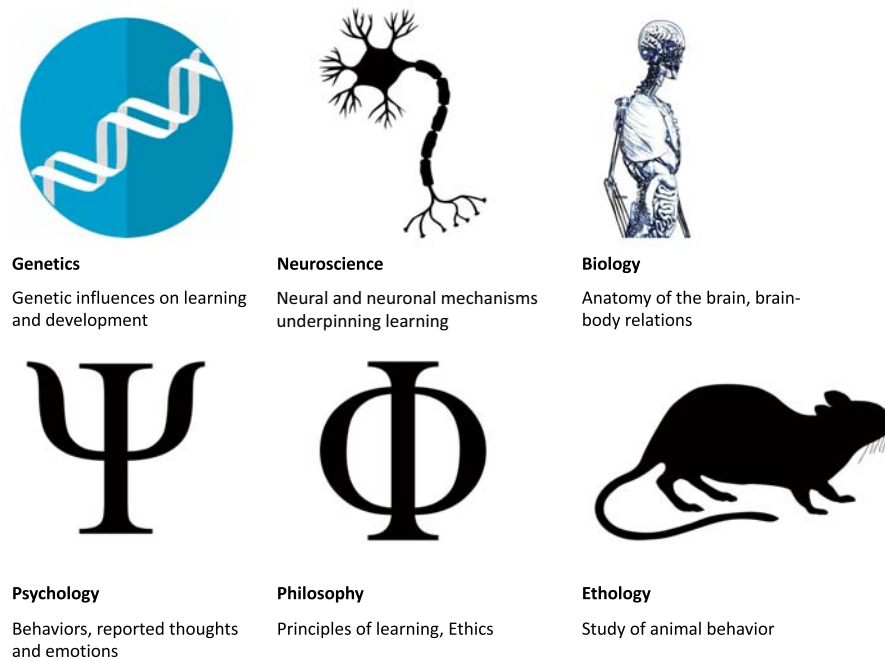


Fig. 1 The multiple disciplines comprising educational neuroscience.

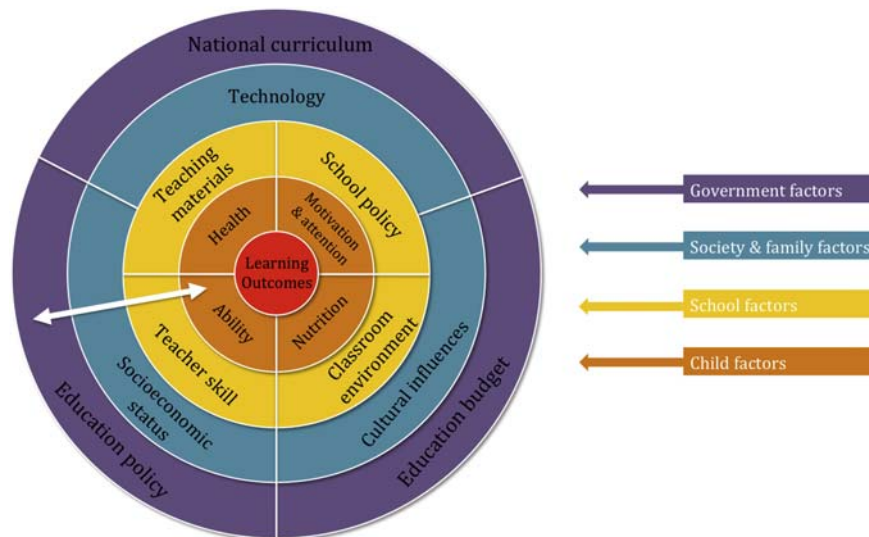


Fig. 2 Children's learning takes place within multiple circles of influence. The white arrow reflects bidirectional influences between circles. Taken from Thomas et al. (2019) [Open access article: <https://doi.org/10.1111/jcpp.12973>].

Using all these methods, educational neuroscience addresses the multitude of factors that influence learning. These factors can be seen as the interaction of different nested ecosystems, where the child is subject and actor within the school, family, society and governmental systems, as depicted in Fig. 2.

This article has three parts. In the first, we present some of the key research domains of educational neuroscience: brain health, social and emotional learning, executive functions, and individual differences in learning. We illustrate how they contribute to an understanding of learning processes within social contexts, and can inform the work of educators. For some, reviewing these domains of research questions the relevance of neuroscience as a discipline that would bring a new perspective above and beyond the professional knowledge of educators, and the well-established theories in psychology. In the second part, we give a flavor of this debate and present several arguments on the added value of neuroscience. We highlight the importance of collaborative projects involving interdisciplinary researchers and educators to generate knowledge that is both robust and impactful. In the third part, we suggest two core areas of development in educational neuroscience that could be taken forward by a new and talented generation of researchers and educators.

Key research domains in educational neuroscience

Brain health fundamentals

For students to be able to learn, they first need to have their basic physiological and social needs met. We will start by presenting some brain health fundamentals (nutrition, exercise, sleep, and protection from stress), whose mechanisms have been further unraveled by educational neuroscience.

Nutrition

Nutrition is fundamental to grow “a healthy mind in a healthy body.” There are several pathways through which food impacts brain functioning and development ([Center for Educational Neuroscience, 2016a](#)): (1) via an increase in blood glucose and energy available to the body, (2) via the release of neurotransmitters, related in part to the consumption of carbohydrates and proteins, and (3) via the provision of micronutrients. On a global scale, there is still progress to be made for children to have appropriate levels of nutrition. Twenty-two percent of the children under 5 in the world experience stunting, meaning that they suffer from growth impairment, in part due to poor nutrition ([UNICEF et al., 2021](#)). Stunting is associated with suboptimal social, cognitive and motor development ([Perkins et al., 2017](#)). At the other extreme, 5.7% of the children under 5 are overweight ([UNICEF et al., 2021](#)).

Exercise

Physical education lessons and physical games during recess time might not be the first activities we think about when it comes to learning at school. Beyond its intrinsic value as a joyful and relaxing activity (for some), physical exercise is associated with better cognitive and academic performance across the lifespan ([Hillman et al., 2008](#)). Exercise seems to stimulate the proliferation of neurons and communication between neurons (via an increase in the brain-derived neurotrophic factor) ([Stillman et al., 2020](#)). Comparisons between participants with different levels of physical activity also suggest differences in brain networks involved in the allocation of attentional resources, in the monitoring of errors and resolution of conflicts ([Hillman et al., 2008](#)). Exercise might also help children sleep at night, offering additional benefits for the learner.

Sleep

Sleep facilitates the learning and consolidation of new information. Fully rested learners are better able to pay attention and keep information in mind, as shown in an experimental study which voluntarily induced sleep deprivation in teenagers ([Lo et al., 2016](#)). Whereas sleep deprivation can negatively impact performance, small adjustments in the school system such as shifting the start of school by an hour are associated with positive changes in school performance, especially for teenagers ([Dunster et al., 2018](#)). Neuroscience and biology have allowed researchers to explore the fascinating mechanisms through which sleep helps the brain learn and memorize information. Information that has been newly encoded by learners is reactivated at night, following an active process of consolidation which helps to form long-term memories ([Rasch and Born, 2013](#)). Measures of brain activity have led researchers to investigate whether it is the total amount of sleep and/or the quality of sleep that is related to performance ([Hayes and Bainton, 2020](#)). Either way, protecting sleep from external stressors is important.

Protection from stress

Protection from chronic stress is essential for children to feel safe, and to free their emotional and cognitive resources for positive learning activities. Stressors can be physical, such as noise and pollution, or social (threat, conflict). Poverty, parental job instability, abuse, and trauma can all induce long-term stress with detrimental consequences for brain development. Stress increases blood pressure, heart rate, blood sugar and hormones such as cortisol ([Center on the Developing Child at Harvard University, 2021](#)). Animal studies have shown that increased cortisol is associated with alterations in the volume of the amygdala (involved in emotional processing), in the hippocampus (involved in memory processing) and with the atrophy of neurons that are essential for the communication between the prefrontal cortex and other brain regions ([Blair and Raver, 2016](#)). The prefrontal cortex is a key brain area involved in planning, attention, and self-regulation. Children exposed to prolonged stressors, whose energy is all focused on responding to stress, can have difficulties in regulating their attention and behavior in learning contexts. Stressors go beyond the school walls, and trauma and abuse require specific integrated interventions. With that in mind, several strategies have been identified to reduce stress in educational environments. These include: reducing physical stressors in the environment (e.g., noise), reappraising learning (e.g., presenting learning tasks in a way that plays to children's strengths), using breathing techniques, mindfulness, recovery times, and physical exercise ([Whiting et al., 2021](#)). Stable and supportive relationships also help mitigating the effect of stress on children.

Social and emotional learning

Educational neuroscience has reinforced the view that learning is a social and emotional process ([Immordino-Yang et al., 2018](#); [Immordino-Yang and Damasio, 2007](#)). In his seminal work, Antonio Damasio questioned the role of a “cold” and “emotionally neutral” reason that would guide our behaviors and decisions ([Damasio, 1994](#)). He showed that emotions are critical for decision making, because they allow us to process certain types of behaviors as rewards or punishments, and to weight risks and benefits. Complex emotions such as guilt and shame, for example, drive us to socially acceptable behaviors.

Damasio's studies in the 1990s were with neurologically impaired patients, who suffered from lesions in the frontal lobes and showed difficulties in social adaptation (Damasio et al., 1994; Eslinger et al., 1992). Neuroimaging studies with non-clinical patients further helped researchers to map out the brain regions that are involved in reasoning and decision making, confirming Damasio's findings. The frontal lobe, which often has been overly simplified as the area of "rationality," constantly interacts with the brain regions from the limbic system which are processing emotions, rewards and memories (Dumontheil and Mareschal, 2020). The orbitofrontal cortex maps out the associations between actions and their expected rewards. It is connected to the amygdala (which processes a wide range of emotions such as fear, disgust, love), and to the striatum (which processes rewards). The ventromedial prefrontal cortex, particularly important to making decisions in social contexts, is connected to the amygdala and to the hippocampus, which stores memories of our sensory experiences in different situations.

Understanding how social and emotional processes influence our reasoning and decision making has the potential to inform education by giving insights into what motivates and hampers students' learning. Adolescence is a key period of development for social cognition, including domains such as recognizing and expressing emotions, processing emotional rewards, understanding the perspective of others, planning, and controlling one's behaviors (Blakemore, 2010). This development is noticeable both in adolescent behaviors and in the protracted development of their prefrontal cortex. Adolescents are particularly sensitive to social approval from their peers (Tomova et al., 2021). This can have negative consequences, for example when substance abuse, or risky driving, are seen as valued within their peer group. However, a high sensitivity to social rewards also means that teenagers can be more prone to act prosocially if their peers do so (Tomova et al., 2021). In the context of classroom learning, organizing collaborative learning activities with structured, social feedback could be a way to recognize and harness adolescents' sensitivity to the opinion of their peers. Presenting adolescent role models could also favor students' engagement in learning activities.

Moving beyond a focus on performance *per se*, and considering the role of social and emotional processes in learning prompts us to create learning contexts that are socially and emotionally nurturing for students. It is not only about what students can do, and how fast they can do it, but how they learn, with whom, and what motivates them to engage in learning activities.

Executive functions

With that in mind, we can consider how individuals regulate their behaviors and their emotions during learning, and how they manage to focus and plan ahead to achieve their goals (Diamond, 2013). These skills are referred to as "executive functions" and include: (1) inhibition (i.e., the ability to control one's attention, behavior, thoughts, emotions, rather than acting on impulse or automatically in response to external stimuli); (2) working memory (i.e., the ability to keep information in mind and work with it); and (3) cognitive flexibility (i.e., the ability to change perspective and switch from one task to another) (Diamond, 2013; Miyake et al., 2000). They underpin academic achievement in a variety of domains (mathematics: Cragg and Gilmore, 2014; Lubin et al., 2013; science: Mareschal, 2016; reading: Cain et al., 2004). In the classroom, student's executive functions are at play when teachers implement behavioral management strategies, try to redirect attention on a task, or set up a debate.

Systematic research can help to better understand which tasks and contexts are particularly challenging for students' executive functions, and the age at which vulnerabilities are present, with the view to facilitate learning processes. Classrooms are filled with multisensory distractors, such as visual displays on the walls (Fisher et al., 2014) or background babble (Klatte et al., 2013; Shield and Dockrell, 2003), which place extra demands on students' executive functions. Children who report having greater difficulties in switching from one task to another report being more distracted and annoyed by noise in the classroom (Massonnié et al., 2020a). This shows the need to pay attention to the *physical* characteristics of the learning environments (Montazami et al., 2015). Distractions are not the only aspects of a classroom environment which are challenging for executive functions. Long instructions can also be difficult for students, who have to retain all the information provided by the teachers while also doing what is expected of them. Breaking down tasks and instructions into smaller components, simplifying the linguistic structure of verbal material, and using memory aids, can all be helpful (Gathercole and Alloway, 2007).

Younger children are particularly sensitive to the demands placed on executive functions, as these skills and their associated brain areas undergo protracted development during the school years. Younger children might need more scaffolding to succeed at specific tasks, such as a breakdown of instructions into smaller steps. They can also be particularly distracted by background speech and classroom noise, especially if they have difficulties inhibiting distractors (Massonnié et al., 2019). Paradoxically, classrooms in the early primary school years tend to be the noisiest (Picard and Bradley, 2001). Children are less able to inhibit the impulse to communicate at an age when they are also less able to screen out the distraction of other children's communication.

Research exploring executive functions has sometimes been associated with hype around training these skills via interventions such as computer training, mindfulness, or martial arts (Diamond and Ling, 2016; Diamond and Lee, 2011). However, there is currently little evidence that training on a given task requiring executive functions improves performance in other unrelated skills (so-called "far transfer"): children just tend to improve at the skills they are trained on (Kassai et al., 2019). If children are trained to hold more numbers in memory, they might be able to hold more numbers, perhaps also more words in memory, but they might not be better at all sorts of maths or science exercises. Promising interventions, such as the Unlocked project (Roy et al., 2019), look at *domain-specific* training of executive functions, that is, training that is embedded in the domain in which they are to be applied, such as science or mathematics. Another promising approach is for interventions to target: (a) children's overall conduct at school and/or (b) their confidence and motivation to exert self-regulation, which might have a snowball effect on their overall attentional focus in class' (Ng-Knight et al., 2021). When considering the impact of educational interventions, it is important to remember that children's ability to take advantage of specific pedagogical practices may, to some extent, depend on other properties of their environment, such as their socioeconomic background, and also on their genetic make-up. Educational neuroscience research has investigated both influences.

Socioeconomic status and educational outcomes

Policymakers are particularly concerned about gaps between children's educational outcomes, and one factor associated with these gaps is socioeconomic status (SES). SES reflects many facets of a child's environment, including material and financial resources, educational opportunities, power and social prestige (Hackman and Kraemer, 2020). Children coming from high SES backgrounds score more highly in tests of oral language, executive functions, reading and mathematical skills, which is likely to be reflected in their social interactions and educational success in the long run.

There are multiple factors explaining the associations between SES and children's outcomes (Evans, 2004; Farah, 2017). Children from low SES background are more likely to be exposed to adversity, which brings stress, threats and violence, and is detrimental to learning and emotional regulation. They also tend to be more exposed to noise and crowding within their home, which can disrupt sleep and induce distraction when learning. Access to health services and quality schools might be more difficult in areas of disadvantage, and family SES influences the types and number of educational resources available in the house, such as books and toys. Additionally, parents facing stressful financial or professional situations might find it difficult to be responsive when interacting with their children (Hackman and Kraemer, 2020).

However, it is important not to consider the relationships between early adversity and children's outcomes purely in terms of deficits and to take into account the strengths and adaptability these children might develop in non-academic areas of their life. Children experiencing adversity might be better able to detect threats, identify negative emotions, understand uncertainty and might be skilled at choosing whom to trust (Ellis et al., 2020). Children coming from low-SES families might also use different strategies to solve the same problems as their peers coming from a higher SES background (Hackman and Kraemer, 2020). Comparisons across SES environments and cultures suggest that the additional challenges in low SES environments may in some case stimulate advances in executive function skills (Howard et al., 2019).

Supporting families living in areas of disadvantage is essential to offer the best educational experience for children and to reduce SES disparities in achievement and learning opportunities. The multitude of pathways through which SES can influence child development allows consideration of various routes for support. Interventions can focus on financial resources with income support and cash transfer programs, which might help accessing schools, buying educational material and reducing income-related stress (Troller-Renfree et al., 2022). Free educational resources might support positive parent-child interactions (<https://wordsforlife.org.uk/>, Dockrell et al., 2020), and classrooms can work to provide a stress-free environment sensitive to children's individual needs and skills (Whiting et al., 2021). The combination of many small effects can lead to big effects on child development.

Individual differences in educational outcomes: genetics

Educational neuroscience has fed long-standing debates about how nature and nurture shape human learning experiences. The field of genetics has been the subject of both public interest and concerns, perhaps due to its potential to "profile" children and outline individual differences early on in life. Researchers have assessed the extent to which genetic make-up—whether indexed by degree of relatedness in identical versus fraternal twins or by direct measures of DNA variation—predict educational achievement (Lee et al., 2018). Genetic factors are indeed associated with school attainment (Donati et al., 2021; Lee et al., 2018). However, associations are complex: variations in multiple genes (perhaps thousands) are associated with a given competence, such as reading, and these genes are not entirely specific to reading, but are involved in other abilities. So, we cannot identify "the gene for intelligence" or "the gene for mathematics."

Furthermore, both genetic and environmental factors interact (and sometimes align) to influence children's learning outcomes. If a child from a family of passionate readers is thriving at school, getting their hands on every book available, it might be that they have inherited relevant genes from their parents. It might also be that they are raised in an environment where a lot of books are available, and where family members are encouraging reading practices by showing their own enthusiasm for books, or by organizing reading sessions.

The environments which children experience influence how, and to what extent, their genetic background is expressed. In that sense, genetic effects are not deterministic. Let us simplify the picture and establish that, based on their genetic profile, some children are more likely to be good readers than others. If their environment provides no reading opportunities (e.g., there are no books), this potential will not be expressed. None of the children will be an avid reader. In contrast to this group, children who do benefit from support from their environment (e.g., accessing books and libraries) would fare better, irrespective of their genetic background. It is only in an educational environment where a multitude of children are provided with learning opportunities (e.g., they are surrounded by books and introduced to reading) that individual differences based on genetic background will become salient. In sum, although children are born with a particular biological and genetic background, their environment determines how much of this potential is expressed. The goal of current research is, ultimately, to unpack the causal pathways by which environmental and genetic influences shape child development and educational achievement, to personalize education and offer the best support possible (Hart, 2016).

The five (non-exhaustive) areas of research we have reviewed in this section (brain health fundamentals, socio-emotional learning, executive functions, socioeconomic status and individual differences in genetic background) help to shed light on the complex ways by which the family, classroom and social environments influence children's learning experiences. We now move on to consider some of the debates that have been stimulated by the dialog between education and neuroscience.

Translating educational neuroscience research to education

What does neuroscience bring above and beyond what educators and psychologists already know?

There has been much debate as to whether neuroscience brings anything beyond what educators and psychologists already know. Some argue that neuroscience is merely re-establishing principles well known in psychology, or involving costly methods that do not bring new insight (Bowers, 2016; Dougherty and Robey, 2018). If teachers are dealing mainly with students' behaviors (i.e., what they say, what they demonstrate they know and do not know in assessments), do they need to know about brain processes?

First of all, re-establishing what educators already know is not a bad thing in itself. Converging evidence in educational neuroscience shows that sleep is important, and that chronic stress is deleterious. Neuroscience evidence has a "seductive allure" which can attract further attention from the media. When messages are backed up by robust neuroscience evidence, they can encourage decision-making in important areas such as the protection of sleep and the reduction of stressors in the environment. However, the "seductive allure" of neuroscience can also be a double-edged sword when neuroscience (or merely its trappings) are misapplied (e.g., when advertising products that are not evidence-based, Beck, 2010).

Importantly for educational practice, there are cases where what we think we already know is actually *not* backed-up by science. False beliefs about brain and learning processes, called "neuromyths," can have consequences for learning conditions. A seminal study from 2014 showed that more than 90% of the teachers from five different countries believed that individuals learn better when they receive information in their preferred learning style, an idea that is not backed-up by scientific research (Howard-Jones, 2014). Implementing teaching strategies based on "learning styles" might be unhelpful, categorizing students and reducing the range of learning opportunities available. Neuromyths about neurodevelopmental disorders (e.g., if a child responds to stimulant medication such as Ritalin, then they are likely to have Attention Deficit Hyperactivity Disorder) are also held by some educators and members of the general public (Gini et al., 2021). Tackling these myths would help to reduce stigma and to provide better support to learners. Another misconception is the idea of "multitasking." Research in educational neuroscience brings nuance to this idea, by suggesting that we are not deliberately paying attention to two things at the same time but instead constantly switching from one task to the other (The Learning Scientists, 2016). More awareness of this switching process can help to reduce cognitive load in the classroom by minimizing demands to "multitask." For more information on "neuro-myths" and "neuro-hits," see Center for Educational Neuroscience (2016b).

Neuroscience does more than proof-checking common sense when it helps to understand *how* we learn—investigating mechanisms. To take one example, teachers know that it can be challenging to teach some scientific concepts because they are counter-intuitive. Advances in scientific theories come from questioning our intuitions and commonly held theories (Kuhn, 2012), following a process called conceptual change. Teaching children that the world is round not flat, contrasts with their intuitive understanding from everyday experience (they walk on a flat floor). We know that scientific understanding does not completely replace or overwrite previous ideas. Instead, multiple ideas and theories are available to children who must acquire the relevant executive function skills to apply the new concepts in the contexts in which they are appropriate. For example, they will use the idea that the Earth is round when doing science tests or watching satellite images, but they can focus on the feeling of a flat surface when playing football. Neuroscience has revealed that learning new scientific concepts and inhibiting pre-existing beliefs is associated with brain areas involved in the detection of conflict (the anterior cingulate cortex), attention, inhibition, working memory and the integration of information (within the prefrontal cortex) (Mareschal, 2016). Houdé et al. (2000) suggested that a shift operates from perceptual to more analytical learning strategies when learning scientific concepts. This is based on a shift in cerebral activity from posterior to more frontal areas. An understanding of mechanism points to interventions that may help the learning of counterintuitive scientific concepts – namely improving inhibitory control skills (Roy et al., 2019).

Finally, beyond validating and refining theories from other fields, neuroscience helps generate new hypotheses to understand phenomena such as, for example, learning difficulties. Neuroscience can suggest where to look by giving an insight into impaired neural functions, and by comparing the neural circuits recruited by students who experience, and those who do not experience, difficulties—see Gabrieli (2016), Ramus (2014) for an illustration with dyslexia and Howard-Jones et al. (2016) for dyscalculia. A biological understanding of the brain and of neural mechanisms also constrains predictions that can be made about interventions (Howard-Jones et al., 2016). For example, if neural networks are domain specific, then training a specific ability will not generalize to another ability, constraining the possibility of far transfer (Thomas, 2019).

Educational values are not defined by educational neuroscience

Perhaps a reassuring stance for the skeptics is to remember that research findings in no way aim to be prescriptive (Willingham, 2009). Educational values are defined independently of science—what we decide to value in education is a choice of society, whether it is rote memorization, creativity, collaborative learning, or any other priority. To give an example, the international definition of 21st century skills, which include communication, collaboration, creativity, critical thinking, and learning to learn, echoes the need of learners to adapt to a fast-evolving job market and to be critical consumers of technologies (UNESCO Office Dakar and Regional Bureau for Education in Africa and Brookings Institution, 2020).

The precedence of the social system can be seen in sleep research. Research demonstrates changes in circadian rhythms from the primary school to the high school years. Adolescents often go to bed later but starting times at school are the same across grades, if not earlier in the higher grades (Carskadon, 2011). This means that, in effect, adolescents experience mild but cumulate sleep deprivation across the school week. There is promising research showing the benefits of starting school later for adolescents' sleep and

school results (Dunster et al., 2018). However, interventions are difficult to implement (Education Endowment Foundation, 2019) given that the start of school is tied to other social and economic factors such as parents' working hours.

Collaborations between educators and researchers are essential because educational neuroscience does not aim to be prescriptive. Rather, it aims to inform learning processes with the potential to be applied to the classroom, depending on the strength of evidence and on pedagogical priorities. It aims to provide a wider set of tools and methods for educators to fulfill their educational goals.

Collaborative projects between scientists, educators and learners are essential

Educators manifest interest in educational neuroscience and its potential application to the classroom (Simmonds, 2014), which can lead to the development of truly collaborative projects in which educators are participating in the design, implementation, and dissemination of research findings (Massonnié et al., 2020b). Collaborations with teachers have grown over the years with the idea that carrying out research *in* schools is not enough for it to inform practice: rather, research must be *with* schools. For translation to happen, the "why" and "how" of research must be sensitive to the reality and complexity of learning contexts, in which educators are the experts.

Teachers are shaping learning environments and have a key understanding of which factors influence learning in the classroom. Differences in terminology and methods between teachers and researchers might create challenges in collaborations (e.g., teachers might talk about "concentration" when researchers talk about "attention"). These can be overcome with open discussions, trust, and familiarization to the concepts and practices from the other field (Hobbiss et al., 2019). Researchers might not be aware of all the constraints faced by educators and truly collaborative projects might require extra flexibility. Some very useful practical guidance for school-based research has already been compiled (Plummer et al., 2014). Detailed and practical guidance on research methods has also been created for teachers (Churches and Dommett, 2016). Establishing a common ground goes beyond the "facilitation" of research projects, it can ultimately lead to more impactful projects answering social needs.

It is worth noting that not *all* neuroscience projects related to learning need involve educators or take place in naturalistic settings. In some senses, basic research remains a foundation upon which translation is built. Matusz et al. (2019) have proposed a model with three-stages of research, namely: (1) Classic laboratory research, with maximal control over stimuli and the environment to isolate very specific factors influencing learning, (2) Naturalist laboratory research, involving more dynamic and complex settings with some control on the environment, and (3) Fully naturalistic research. Each stage has its own advantages and inconveniences, and all three would ideally be implemented. Tools assessing the degree of ecological validity of a study might help in this endeavor (Naumann et al., 2021). We will now highlight two main ideas to move the field forward: a systematic training for researchers on interdisciplinarity and collaborations with educators, and the involvement of learners in the research process.

Where does the future of educational neuroscience lie?

Training researchers for interdisciplinarity

Educational neuroscience is a challenging field of research requiring both a strong expertise in its constituting disciplines and a capacity to gain perspective on their integration and application to education. This creates strong expectations for early career researchers working in the field of educational neuroscience, for previous generations of researchers who might have been specialized in one specific discipline, and for both generations to develop two-way collaborations with educators and educational stakeholders.

University programs have been created as early as in 2002 (Blake and Gardner, 2007; Harvard Graduate School of Education, 2021; Birkbeck University of London, 2021; Ecole Normale Supérieure, 2021; Universität Osnabrück, 2021) with the view to instigate interdisciplinarity by building expertise in multiple disciplines of educational neuroscience. In addition, mentoring schemes and peer support groups can help maintain and develop interdisciplinarity within the research community (Gotlieb et al., 2019).

However, collaboration needs to go beyond the walls of the university to truly inform learning in naturalistic contexts. Place-ments in schools and involvement in learning communities would provide invaluable opportunities to gain insight into the objectives, and working conditions of educators. Yet, it is still rarely done due to administrative and financial challenges as well as the flexibility required from both universities and schools to build synergies. Valuing time in school as part of the professional development of academics would be a great step forward to foster impact and knowledge exchange (Glennon et al., 2013).

Involving learners in the research process

So far, most of the endeavors related to knowledge exchange, dissemination and impact have focused on teachers as the gateways to educational settings and key providers of learning opportunities. However, one further step needs to be taken to involve the learners (Brookman-Byrne and Commissar, 2019). If we are to study learning processes, learners are the ultimate beneficiaries of the research. Dissemination outlets already exist for learners of different ages, as shown by initiatives such as the podcasts series from the Learning scientists (The Learning Scientists, 2022) or the collection "Everything you and your teacher need to know about the brain" published in Frontiers for Young Minds (Peters et al., 2020).

Introducing children to brain and learning processes can change their views on how the brain works, both in terms of its functions and its components (Chauvin et al., 2019; Marshall and Comalli, 2012). Explaining neuroplasticity (e.g., the fact that neuronal connections change in the process of learning) has the potential to increase students' motivation, although links to academic achievement are inconsistent (Sarrasin et al., 2018). Beyond sharing conceptual knowledge about learning and the brain, educational neuroscience has the potential to empower students in using evidence-based learning strategies (e.g., Dunlosky et al., 2013) and in taking care of their brain by considering, for example, their sleep patterns, nutrition, drug and alcohol usage. Students are in the premier position to enact learning processes, yet little research has been done so far on how advances in educational neuroscience influence their behavior, learning and well-being.

Furthermore, the direct involvement of learners in the design and creation of research projects is still rare. If educational neuroscience aims to support and improve learning, it is fundamental to incorporate the learners' views on the research project: Does the project have the potential to be impactful? Can it change learning conditions or attitudes towards learning? Can it support the learners' growth and development? (van Atteveldt et al., 2019). To go one step further, students can also suggest which research questions are to be investigated by academics (Tokuhami-Espinosa, personal communication, January 23, 2020). Some institutions such as the Center for Research and Interdisciplinarity (2021) in France are pairing classrooms with a scientific mentor to develop bespoke research projects. By combining innovative and impactful ideas with rigorous and ethical research methods, participatory research has the potential to energize the field of educational neuroscience while ensuring it serves the greater good (Kuriloff et al., 2009).

Conclusions

Educational neuroscience is a recent and growing field which combines the expertise from various disciplines to understand learning processes, and investigates how this developing knowledge might translate into educational practice. Educational neuroscience offers insight into many of the mechanisms influencing learning (including nutrition, exercise, sleep and stress, social interactions and emotional responses) and helps to explain individual differences between learners. Each discipline within educational neuroscience brings a unique perspective to shed light on the complexity of learning processes, and offers explanations of learning mechanisms at different levels (including the genetic, neural, and psychological levels). As such, collaborations between researchers of different fields are at the core of educational neuroscience. Collaborations between researchers, educators and the learners themselves are also critical to develop a comprehensive approach which includes not only laboratory studies but also controlled experiments and naturalistic studies in educational environments. Co-constructing research protocols with the learners themselves is relatively new and holds promise for the future development of the field. For learners are not only the recipients of educational opportunities, but full actors in their own development.

References

- Beck, D.M., 2010. The appeal of the brain in the popular press. *Perspect. Psychol. Sci.* 5 (6), 762–766. <https://doi.org/10.1177/1745691610388779>.
- Birkbeck University of London, 2021. Educational Neuroscience (MSC). https://www.bbk.ac.uk/study/2021/postgraduate/programmes/TMSDNRO_C/.
- Blair, C., Raver, C.C., 2016. Poverty, stress, and brain development: new directions for prevention and intervention. *Acad. Pediatr.* 16 (3), S30–S36. <https://doi.org/10.1016/j.acap.2016.01.010>.
- Blake, P.R., Gardner, H., 2007. A first course in mind, brain, and education. *Mind Brain Educ.* 1 (2), 61–65. <https://doi.org/10.1111/j.1751-228X.2007.00007.x>.
- Blakemore, S.-J., 2010. The developing social brain: implications for education. *Neuron* 65 (6), 744–747. <https://doi.org/10.1016/j.neuron.2010.03.004>.
- Bowers, J.S., 2016. The practical and principled problems with educational neuroscience. *Psychol. Rev.* 123 (5), 600–612. <https://doi.org/10.1037/rev0000025>.
- Brookman-Byrne, A., Commissar, L., 2019. Future avenues for educational neuroscience from the perspective of EARLI SIG 22 conference attendees. *Mind Brain Educ.* <https://doi.org/10.1111/mbe.12211>.
- Cain, K., Oakhill, J., Bryant, P., 2004. Children's reading comprehension ability: concurrent prediction by working memory, verbal ability, and component skills. *J. Educ. Psychol.* 96 (1), 31–42. <https://doi.org/10.1037/0022-0663.96.1.31>.
- Carskadon, M.A., 2011. Sleep in adolescents: the perfect storm. *Pediatr. Clin.* 58 (3), 637–647. <https://doi.org/10.1016/j.pcl.2011.03.003>.
- Center for Research and Interdisciplinarity, 2021. Savanturiers du Cerveau. <https://savanturiers-projects.cri-paris.org/repo/5d08fba7119a6fa0e7c46481>.
- Center on the Developing Child at Harvard University, 2021. Three Principles to Improve Outcomes for Children and Families, 2021 Update. <http://www.developingchild.harvard.edu>.
- Centre for Educational Neuroscience, 2016a. Diet Makes a Difference to Learning. <http://www.educationalneuroscience.org.uk/resources/neuromyth-or-neurofact/diet-makes-a-difference-to-learning/>.
- Centre for Educational Neuroscience, 2016b. Neuro-hit or Neuro-myth? <http://www.educationalneuroscience.org.uk/resources/neuromyth-or-neurofact/>.
- Chauvin, R., Lucchesi, A., Massonnié, J., Rodo, C., Malet, I., Lestage, H., Théro, H., Saunier, C., Bosch, E., Akkermans, S., Guimaraes, J., Przewdzik, I., 2019. Le conte comme moyen de changer des représentations des enfants sur le cerveau. *Grand N* 104, 5–21.
- Churches, R., Dommett, E., 2016. Teacher-Led Research: Designing and Implementing Randomised Controlled Trials and Other Forms of Experimental Research. Crown House Publishing.
- Cragg, L., Gilmore, C., 2014. Skills underlying mathematics: the role of executive function in the development of mathematics proficiency. *Trends Neurosci. Educ.* 3 (2), 63–68. <https://doi.org/10.1016/j.tine.2013.12.001>.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., Damasio, A.R., 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264 (5162), 1102–1105.
- Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. Avon Books.
- Diamond, A., Lee, K., 2011. Interventions shown to aid executive function development in children 4–12 Years old. *Science* 333 (6045), 959–964. <https://doi.org/10.1126/science.1204529>.

- Diamond, A., Ling, D.S., 2016. Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Dev. Cognit. Neurosci.* 18, 34–48. <https://doi.org/10.1016/j.dcn.2015.11.005>.
- Diamond, A., 2013. Executive functions. *Annu. Rev. Psychol.* 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>.
- Dockrell, J.E., Law, J., Mathers, S., Forrest, C., Charlton, J., 2020. Talking With Your Children: Activities for Parents to Help Nursery Children's Speaking and Listening Skills. <https://discovery.ucl.ac.uk/id/eprint/10111307/>.
- Donati, G., Dumontheil, I., Pain, O., Asbury, K., Meaburn, E.L., 2021. Evidence for specificity of polygenic contributions to attainment in English, Maths and Science during adolescence. *Sci. Rep.* 11 (1), 1–11. <https://doi.org/10.1038/s41598-021-82877-y>.
- Dougherty, M.R., Robey, A., 2018. Neuroscience and education: a bridge astray? *Curr. Dir. Psychol. Sci.* 27 (6), 401–406. <https://doi.org/10.1177/0963721418794495>.
- Dumontheil, I., Mareschal, D., 2020. An introduction to brain and cognitive development: the key concepts you need to know. In: *Educational Neuroscience—Development Across the Lifespan*. Routledge, pp. 23–49.
- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., Willingham, D.T., 2013. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Publ. Interest* 14 (1), 4–58. <https://doi.org/10.1177/1529100612453266>.
- Dunster, G.P., de la Iglesia, L., Ben-Hamo, M., Nave, C., Fleischer, J.G., Panda, S., Horacio, O., 2018. Sleepmore in Seattle: later school start times are associated with more sleep and better performance in high school students. *Sci. Adv.* 4 (12), eaau6200. <https://doi.org/10.1126/sciadv.aau6200>.
- Ecole Normale Supérieure, 2021. Cogmaster. <https://cogmaster.ens.psl.eu/en/presentation-13509>.
- Education Endowment Foundation, 2019. Teensleep. <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/teensleep/>.
- Ellis, B.J., Abrams, L.S., Masten, A.S., Sternberg, R.J., Tottenham, N., Frankenhuis, W.E., 2020. Hidden talents in harsh environments. *Dev. Psychopathol.* 1–19. <https://doi.org/10.1017/S0954579420000887>.
- Eslinger, P.J., Grattan, L.M., Damasio, H., Damasio, A.R., 1992. Developmental consequences of childhood frontal lobe damage. *Arch. Neurol.* 49 (7), 764–769.
- Evans, G.W., 2004. The environment of childhood poverty. *Am. Psychol.* 59 (2), 77–92. <https://doi.org/10.1037/0003-066X.59.2.77>.
- Farah, M.J., 2017. The neuroscience of socioeconomic status: correlates, causes, and consequences. *Neuron* 96 (1), 56–71. <https://doi.org/10.1016/j.neuron.2017.08.034>.
- Fisher, A.V., Godwin, K.E., Seltman, H., 2014. Visual environment, attention allocation, and learning in young children: when too much of a good thing may be bad. *Psychol. Sci.* 25 (7), 1362–1370. <https://doi.org/10.1177/0956797614533801>.
- Gabrieli, J.D.E., 2016. The promise of educational neuroscience: comment on Bowers (2016). *Psychol. Rev.* 123 (5), 613–619. <https://doi.org/10.1037/rev0000034>.
- Gathercole, S.E., Alloway, T.P., 2007. *Understanding Working Memory—A Classroom Guide*. Harcourt Assessment.
- Gini, S., Knowland, V., Thomas, M.S., Van Herwegen, J., 2021. Neuromyths about neurodevelopmental disorders: misconceptions by educators and the general public. *Mind Brain Educ.* 15 (4), 289–298. <https://doi.org/10.1111/mbe.12303>.
- Glennon, C., Hinton, C., Callahan, T., Fischer, K.W., 2013. School-based research. *Mind Brain Educ.* 7 (1), 30–34. <https://doi.org/10.1111/mbe.12004>.
- Gottlieb, R.J., Pollack, C., Younger, J.W., Toomarian, E.Y., Allaire-Duquette, G., Mariager, N.M., 2019. Next steps for mind, brain, and education: strengthening early-career development. *Mind Brain Educ.* 13 (3), 120–132. <https://doi.org/10.1111/mbe.12197>.
- Hackman, D.A., Kraemer, D.J., 2020. Socioeconomic disparities in achievement: insights on neurocognitive development and educational interventions. In: *Educational Neuroscience*. Routledge, pp. 88–119.
- Hart, S.A., 2016. Precision education initiative: moving toward personalized education. *Mind Brain Educ.* 10 (4), 209–211. <https://doi.org/10.1111/mbe.12109>.
- Harvard Graduate School of Education, 2021. *Mind, Brain and Education*. <https://www.gse.harvard.edu/masters/mbe>.
- Hayes, B., Bainton, J., 2020. The impact of reduced sleep on school related outcomes for typically developing children aged 11–19: a systematic review. *Sch. Psychol. Int.* 41 (6), 569–594. <https://doi.org/10.1177/0143034320961130>.
- Hillman, C.H., Erickson, K.I., Kramer, A.F., 2008. Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9 (1), 58–65. <https://doi.org/10.1038/nrn2298>.
- Hobiss, M., Massonnié, J., Tokuhama-Espinosa, T., Gittner, A., Arson de Sousa Lemos, M., Tovazzi, A., Hindley, C., Baker, S., Sumeracki, M.A., Wassenaar, T., Gous, I., 2019. “UNIFIED”: bridging the researcher-practitioner divide in mind, brain and education. *Mind Brain Educ.* 13 (4), 298–312. <https://doi.org/10.1111/mbe.12223>.
- Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., Tzourio-Mazoyer, N., 2000. Shifting from the perceptual brain to the logical brain: the neural impact of cognitive inhibition training. *J. Cognit. Neurosci.* 12 (5), 721–728. <https://doi.org/10.1162/089892900562525>.
- Howard, S.J., Cook, C.J., Everts, L., Melhuish, E., Scerif, G., Norris, S., Twine, R., Kahn, K., Draper, C.E., 2019. Challenging socioeconomic status: a cross-cultural comparison of early executive function. *Dev. Sci.* 23 (1), e12854. <https://doi.org/10.1111/desc.12854>.
- Howard-Jones, P.A., Varma, S., Ansari, D., Butterworth, B., De Smedt, B., Goswami, U., Laurillard, D., Thomas, M.S., 2016. The principles and practices of educational neuroscience: comment on Bowers (2016). *Psychol. Rev.* 123 (5), 620–627. <https://doi.org/10.1037/rev0000036>.
- Howard-Jones, P., Ioannou, K., Bailey, R., Prior, J., Yau, S.H., Jay, T., 2018. Applying the science of learning in the classroom. *Impact* 2, 9–12.
- Howard-Jones, P.A., 2014. Neuroscience and education: myths and messages. *Nat. Rev. Neurosci.* 15 (12), 817–824. <https://doi.org/10.1038/nrn3817>.
- Immordino-Yang, M.H., Damasio, A., 2007. We feel, therefore we learn: the relevance of affective and social neuroscience to education. *Mind Brain Educ.* 1 (1), 3–10. <https://doi.org/10.1111/j.1751-228X.2007.00004.x>.
- Immordino-Yang, M.H., Darling-Hammond, L., Krone, C., 2018. *The Brain Basis for Integrated Social, Emotional, and Academic Development: How Emotions and Social Relationships Drive Learning*. Aspen Institute.
- Kassai, R., Futo, J., Demetrovics, Z., Takacs, Z.K., 2019. A meta-analysis of the experimental evidence on the near-and far-transfer effects among children's executive function skills. *Psychol. Bull.* 145 (2), 165–188. <https://doi.org/10.1037/BUL0000180>.
- Klatte, M., Bergström, K., Lachmann, T., 2013. Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4, 578. <https://doi.org/10.3389/fpsyg.2013.00578>.
- Kuhn, T., 2012. *The Structure of Scientific Revolutions*. University of Chicago Press.
- Kurilloff, P., Reichert, M., Stoudt, B., Ravitch, S., 2009. Building research collaboratives among schools and universities: lessons from the field. *Mind Brain Educ.* 3 (1), 34–44. <https://doi.org/10.1111/j.1751-228X.2008.01051.x>.
- Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Linnér, R.K., Fonatana, M.A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P.N., Walters, R.K., Willoughby, E.A., Yengo, L., et al., 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* 50 (8), 1112–1121. <https://doi.org/10.1038/s41588-018-0147-3>.
- Lo, J.C., Ong, J.L., Leong, R.L., Gooley, J.J., Chee, M.W., 2016. Cognitive performance, sleepiness, and mood in partially sleep deprived adolescents: the need for sleep study. *Sleep* 39 (3), 687–698. <https://doi.org/10.5665/sleep.5552>.
- Lubin, A., Vidal, J., Lanø, C., Houdé, O., Borst, G., 2013. Inhibitory control is needed for the resolution of arithmetic word problems: a developmental negative priming study. *J. Educ. Psychol.* 105 (3), 701. <https://doi.org/10.1037/a0032625>.
- Mareschal, D., 2016. The neuroscience of conceptual learning in science and mathematics. *Curr. Opin. Behav. Sci.* 10, 114–118. <https://doi.org/10.1016/j.cobeha.2016.06.001>.
- Marshall, P.J., Comalli, C.E., 2012. Young children's changing conceptualizations of brain function: implications for teaching neuroscience in early elementary settings. *Early Educ. Dev.* 23 (1), 4–23. <https://doi.org/10.1080/10409289.2011.616134>.
- Massonnié, J., Rogers, C.J., Mareschal, D., Kirkham, N.Z., 2019. Is classroom noise always bad for children? The contribution of age and selective attention to creative performance in noise. *Front. Psychol.* 10 (381). <https://doi.org/10.3389/fpsyg.2019.00381>.
- Massonnié, J., Frassetto, P., Mareschal, D., Kirkham, N.Z., 2020a. Learning in noisy classrooms: children's reports of annoyance and distraction from noise are associated with individual differences in mind- wandering and switching skills. *Environ. Behav.* <https://doi.org/10.1177/0013916520950277>.

- Massonnié, J., Frassetto, P., Mareschal, D., Kirkham, N.Z., 2020b. Scientific collaboration with educators: practical insights from an in-class noise-reduction intervention. *Mind Brain Educ.* 14 (3), 303–316. <https://doi.org/10.1111/mbe.12240>.
- Matusz, P.J., Dikker, S., Huth, A.G., Perrodin, C., 2019. Are we ready for real-world neuroscience? *J. Cognit. Neurosci.* 31 (3), 327–338. https://doi.org/10.1162/jocn_e_01276.
- Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognit. Psychol.* 41 (1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>.
- Montazami, A., Gaterell, M., Nicol, F., 2015. A comprehensive review of environmental design in UK schools: history, conflicts and solutions. *Renew. Sustain. Energy Rev.* 46, 249–264. <https://doi.org/10.1016/j.rser.2015.02.012>.
- Naumann, S., Byrne, M.L., de la Fuente, L.A., Harrewijn, A., Nugiel, T., Rosen, M.L., van Atteveldt, N., Matusz, P., 2021. Assessing the Degree of Ecological Validity of Your Study: Introducing the Multidimensional Assessment of Research in Context (MARC) Tool. *PsyArXiv Preprint*. <https://doi.org/10.31234/osf.io/qb9tz>.
- Ng-Knight, T., Gilligan-Lee, K. A., Massonnié, J., Gaspard, H., Gooch, D., Querstret, D., Johnstone, N., 2021. Does Taekwondo improve children’s self-regulation? If so, how? a randomized field experiment. *Dev. Psychol.* <https://doi.org/10.1037/dev0001307>.
- Pasquinelli, E., 2011. Knowledge- and evidence-based education: reasons, trends, and contents. *Mind Brain Educ.* 5 (4), 186–195. <https://doi.org/10.1111/j.1751-228X.2011.01128.x>.
- Perkins, J.M., Kim, R., Krishna, A., McGovern, M., Aguayo, V.M., Subramanian, S.V., 2017. Understanding the association between stunting and child development in low-and middle-income countries: next steps for research and intervention. *Soc. Sci. Med.* 193, 101–109. <https://doi.org/10.1016/j.socscimed.2017.09.039>.
- Peters, S., van Atteveldt, N., Massonnié, J., Vogel, S.E., 2020. Everything You and Your Teachers Need to Know About the Learning Brain. *Frontiers Media SA*. <https://kids.frontiersin.org/collections/10207/everything-you-and-your-teachers-need-to-know-about-the-learning-brain>.
- Picard, M., Bradley, J.S., 2001. Revisiting Speech Interference in Classrooms: revisando la interferencia en el habla dentro del salón de clases. *Audiology* 40 (5), 221–244.
- Plummer, B.D., Galla, B.M., Finn, A.S., Patrick, S.D., Meketon, D., Leonard, J., Goetz, C., Fernandez-Vina, E., Bartolino, S., White, R.E., 2014. A behind-the-scenes guide to school-based research. *Mind Brain Educ.* 8 (1), 15–20. <https://doi.org/10.1111/mbe.12040>.
- Ramus, F., 2014. Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends Cognit. Sci.* 18 (6), 274–275. <https://doi.org/10.1016/j.tics.2014.01.009>.
- Rasch, B., Born, J., 2013. About sleep’s role in memory. *Physiol. Rev.* 93, 681–766. <https://doi.org/10.1152/physrev.00032.2012>.
- Roy, P., Rutt, S., Easton, C., Sims, D., Bradshaw, S., McNamara, S., 2019. Stop and Think: Learning Counterintuitive Concepts—Evaluation Report. Education Endowment Foundation.
- Sarrasin, J.B., Nenciovici, L., Foisy, L.-M.B., Allaire-Duquette, G., Riopel, M., Masson, S., 2018. Effects of teaching the concept of neuroplasticity to induce a growth mindset on motivation, achievement, and brain activity: a meta-analysis. *Trends Neurosci. Educ.* 12, 22–31. <https://doi.org/10.1016/j.tine.2018.07.003>.
- Shield, B.M., Dockrell, J.E., 2003. The effects of noise on children at school: a review. *Build. Acoust.* 10 (2), 97–116. <https://doi.org/10.1260/135101003768965960>.
- Simmonds, A., 2014. How Neuroscience is Affecting Education: Report of Teacher and Parent Surveys. Wellcome Trust.
- Stillman, C.M., Esteban-Cornejo, I., Brown, B., Bender, C.M., Erickson, K.I., 2020. Effects of exercise on brain and cognition across age groups and health states. *Trends Neurosci.* 43 (7), 533–543. <https://doi.org/10.1016/j.tins.2020.04.010>.
- The Learning Scientists, 2016. Weekly Digest #23: The Myth of Multitasking. <https://www.learningscientists.org/blog/2016/8/21/weekly-digest-23>.
- The Learning Scientists, 2022. Podcasts. <https://www.learningscientists.org/learning-scientists-podcast/>.
- Thomas, M.S.C., Ansari, D., Knowland, V.C.P., 2019. Annual research review: educational neuroscience: progress and prospects. *J. Child Psychol. Psychiatry* 60 (4), 477–492. <https://doi.org/10.1111/jcpp.12973>.
- Thomas, M.S., Mareschal, D., Dumontheil, I., 2020. Educational Neuroscience: Development Across the Life Span. Routledge.
- Thomas, M.S., 2019. Response to Dougherty and Robey (2018) on neuroscience and education: enough bridge metaphors—interdisciplinary research offers the best hope for progress. *Curr. Dir. Psychol. Sci.* 28 (4), 337–340. <https://doi.org/10.1177/0963721419838252>.
- Tomova, L., Andrews, J.L., Blakemore, S.-J., 2021. The importance of belonging and the avoidance of social risk taking in adolescence. *Dev. Rev.* 61 (100981). <https://doi.org/10.1016/j.dr.2021.100981>.
- Troller-Renfree, S.V., Costanzo, M.A., Duncan, G.J., Magnuson, K., Gennetian, L.A., Yoshikawa, H., Halpern-Meekin, S., Fox, N.A., Noble, K.G., 2022. The impact of a poverty reduction intervention on infant brain activity. *Proc. Natl. Acad. Sci.* 119 (5). <https://doi.org/10.1073/pnas.2115649119> e2115649119.
- UNESCO Office Dakar and Regional Bureau for Education in Africa, & Brookings Institution, 2020. Capturing 21st Century Skills: Analysis of Assessments in Selected Sub-saharan African Countries. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000374052>.
- UNICEF, World Health Organization, World Bank, 2021. Joint Child Malnutrition Estimates. UNICEF. <https://data.unicef.org/topic/nutrition/malnutrition/>.
- Universität Osnabrück, 2021. Cognitive Science B.Sc. https://www.ikw.uni-osnabrueck.de/en/courses_of_study/bachelor_program.html.
- van Atteveldt, N., Tijlma, G., Janssen, T., Kupper, F., 2019. Responsible research and innovation as a novel approach to guide educational impact of mind, brain, and education research. *Mind Brain Educ.* 13 (4), 279–287. <https://doi.org/10.1111/mbe.12213>.
- Whiting, S.B., Wass, S.V., Green, S., Thomas, M.S., 2021. Stress and learning in pupils: neuroscience evidence and its relevance for teachers. *Mind Brain Educ.* 15 (2), 177–188. <https://doi.org/10.1111/mbe.12282>.
- Willingham, D.T., 2009. Three problems in the marriage of neuroscience and education. *Cortex* 45 (4), 544–545. <https://doi.org/10.1016/j.cortex.2008.05.009>.

Embodied cognition

Markus Kiefer^a, Carmen Hofmann^{a,b}, and Petra A. Arndt^b, ^aSection for Cognitive Electrophysiology, Department of Psychiatry, Ulm University, Ulm, Germany; and ^bTransfer Center for Neuroscience and Education (ZNL), Ulm University, Ulm, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	67
From amodal over embodiment to hybrid theories of conceptual representations	68
Embodiment of conceptual memory for objects	69
Embodiment of conceptual memory for numbers	69
Embodiment of memory for abstract concepts	70
Embodiment of memory for events	71
Embodiment of reading and writing	71
Conclusion	72
Supplementary material	72
Acknowledgments	72
References	72

Introduction

Concepts held in human long-term memory are the basic building blocks of thought because they constitute the meaning of objects, events and abstract ideas (Kiefer and Pulvermüller, 2012; Kiefer and Harpaintner, 2020). Concepts are also crucial for language by establishing semantic meaning of words (Jackendoff, 1990). Concepts therefore play an important role in object recognition, problem solving and verbal communication (Humphreys et al., 1988; Levelt et al., 1999). Despite this general agreement about the function of concepts, the nature of human cognition has been subject of a debate in the current cognitive sciences, but also in philosophy for more than 2000 years. This debate concerns the fundamental question, whether human cognition, including concepts held in memory, is abstract and detached from perception or action, or whether it is essentially based on re-activations of perceptual and motor experiences, so-called modality-specific representations.

This debate dates back to the ancient Greek philosophers more than 2000 years ago (for an overview, see Runes, 1962; Kiefer and Pulvermüller, 2012): Since the early days of Platon, philosophers have speculated whether or not concepts are essentially grounded in our senses and in our actions with the environment (Markie, 2008). Platon and more modern philosophers such as Descartes, Leibnitz and Kant were very skeptical in this regard. According to their rationalist views, concepts are mental entities fundamentally distinct from sensory impressions. In fact, perception is even thought to require concepts as an a priori condition. According to this rationalist view, concepts are based upon innate categories or are formed by reasoning. Rationalist philosophers argued that concepts cannot reflect sensations because perception is too unstructured, in order to provide the basis for meaningful concepts. In contrast, Aristotle and more modern empiricist philosophers, including Locke and Hume, proposed that all concepts are derived from sensory experiences (impressions) and ideas, which are faint copies of impressions. Because any significant knowledge is to be gained ultimately from one's sense-based experience, concepts based upon reasoning cannot, in this view, provide valid information about the physical and social world.

A similar debate regarding the role of sensory and motor experiences for cognition has been taking place in the cognitive sciences. In the cognitive sciences, concepts are traditionally specified as abstract mental entities, different from the perceptual or motor brain systems, similarly to the rationalist philosophical stance (Quillian, 1969; Anderson, 1983; Tyler and Moss, 2001; Pylyshyn, 1984): Sensory or motor features of objects and events are transformed into a common amodal representation format, in which original modality-specific information is lost.

In contrast, similarly to the empiricist stance, embodied cognition theories assume that conceptual representations are essentially rooted in modal experiential information (Kiefer and Pulvermüller, 2012; Barsalou, 2008; Kiefer and Spitzer, 2001; Pulvermüller and Fadiga, 2010; Lakoff and Johnson, 1999; Bermeitinger and Kiefer, 2012; Kiefer and Harpaintner, 2020; Meteyard et al., 2012). According to embodied cognition theories, activity in brain systems involved in perception, introspection and action necessarily contributes to conceptual meaning. Although the terms "grounded cognition" and "embodied cognition" are frequently used interchangeably, these terms occasionally refer to slightly different classes of theories, depending on the precise theoretical stance (Kiefer and Harpaintner, 2020). While some embodied cognition theories confine experiential information to the sensory-motor systems, grounded cognition theories are typically broader and also consider the relevance of experiential information based on introspection and emotional states, that is perception of internal states.

Furthermore, strong embodied cognition theories assume a complete functional and neuroanatomical overlap between sensory-motor and conceptual processing, including mandatory activation of primary sensory-motor brain areas (Meteyard et al., 2012). In contrast, weaker embodied cognition theories and grounded cognition theories propose a partial overlap between sensory-motor and conceptual processing (for further differentiations between grounded and embodied cognition theories, see Kiefer and

Barsalou, 2013; Kiefer and Harpaintner, 2020). Accumulating evidence speaks against strong variants of embodied cognition theories but supports weaker embodied or grounded cognition theories (see discussions in Popp et al., 2019a,b; Meteyard et al., 2012). We return to this issue later in this article. For the purpose of this article, we use the terms “embodied cognition” and “grounded cognition” exchangeably to refer to theories, which emphasize the relevance of modal experiential information including internal states and emotion for cognition.

This debate about the role of sensory-motor experiences in the cognitive sciences is echoed by similar considerations in child development and education (Kiefer et al., 2022). According to Piaget’s theory of cognitive development, children acquire knowledge through sensory experiences and the manipulation of objects until the age of two. After learning to use symbols (e.g. words, pictures) to represent and think about objects and events, children are thought to develop logical thinking, which is still based on concrete events between the ages of 7 and 11 (Piaget, 1952). Formal logical thinking detached from sensory-motor experiences, the highest level of cognition, is assumed to develop later during adolescence. According to Piaget’s theory, formal logical thinking replaces thinking based on sensory and motor processes (Inhelder and Piaget, 1958). At this final stage of development, thought operations do not need to relate to concrete experiences and phenomena. The question whether this theory holds is of high relevance for the design of school lessons. At present, teachers often follow the implications of Piaget’s model: In primary schools, discovery learning and experiential learning activities are common instructional strategies for active learning arrangements. In contrast, in secondary schools these teaching strategies are less common. Often they are limited to active phases in natural science classes (if they occur at all). Sometimes, teachers even assume that experiential learning activities at secondary schools might be detrimental for an optimal development of abstract logical thinking in their students.

In order to illustrate the implications of these issues for education, consider a classroom situation, in which a teacher wants to introduce a certain type of tree, a beech, to her or his students. She or he has several possibilities for teaching the relevant information. For instance, (i) the teacher can verbally describe the shape of this tree, the color of the leaves, the texture of the bark, the use of the wood, etc. (ii) The teacher could show a movie of the physical properties and the use of a beech. (iii) The teacher could take the students to the forest where they can see a beech, can collect leaves or wood pieces and can touch the bark. They could additionally go to a carpenter’s shop, where they can observe the construction of furniture out of beech wood and where they can construct an own piece. In this latter case, perceptual and motor information elicited by the direct experience are the basis of knowledge building. The different methods to teach a beech may be differentially efficient in supporting learning. The question arises, whether research in the cognitive sciences can provide evidence, which methods has the most beneficial effects and why (Kiefer and Trumpp, 2012; Kiefer et al., 2022).

From amodal over embodiment to hybrid theories of conceptual representations

Psychological and neuroscientific research provides information on the relationship between sensory-motor processes and cognition. At the heart of the past and the current debates is the question, whether or not cognition is essentially grounded in our senses and in our actions with the environment (Markie, 2008). As outlined above, cognition has been thought for a long time to involve neuro-cognitive systems that are different from the perceptual or motor brain systems and code knowledge in an abstract amodal format, in which original modality-specific sensory-motor information is lost (Quillian, 1969; Anderson, 1983; Tyler and Moss, 2001; Pylyshyn, 1984). This theoretical account resembles the stage of formal logical thinking in Piaget’s classical theory (Inhelder and Piaget, 1958).

At a neuroanatomical level, amodal conceptual representations are assumed to be held in heteromodal association cortex such as the anterior temporal (McClelland and Rogers, 2003; Rogers et al., 2004) or posterior temporal cortex (Hoffman et al., 2012), so-called semantic hubs. Activation of modality-specific sensory and motor representations during language comprehension or conceptual thinking is considered in these class of theories only as an auxiliary process after the amodal concept has been accessed, due to imagery (Machery, 2007) or passive spreading of activation to input or output levels (Mahon, 2015).

Challenging this classical view, recent theories of embodied cognitions propose close links between the sensory and motor brain systems on the one hand and cognition on the other hand (Barsalou et al., 2018; Kiefer and Barsalou, 2013; Kiefer and Harpaintner, 2020; Pulvermüller and Fadiga, 2010; Lakoff and Johnson, 1999; Kiefer and Trumpp, 2012). Cognition and thinking is critically based on a reinstatement of external (perception) and internal states (proprioception, emotion and introspection) as well as bodily actions that produce simulations of previous experiences. These simulations of previous sensory-motor experiences (Kiefer and Barsalou, 2013) are often unconscious, but can be measured with behavioral or neuroscientific experimental techniques (Trumpp et al., 2013).

Most recent evidence suggests an interplay between modality-specific, bi- or trimodal, multimodal and amodal semantic hub regions, giving rise to the development of so-called hybrid theories (Kiefer and Harpaintner, 2020; Kuhnke et al., 2020; Patterson and Ralph, 2016; Popp et al., 2019a). Modality-specific and multimodal regions presumably represent conceptual feature content (Kuhnke et al., 2020), whereas semantic hubs code conceptual information in an overarching supramodal fashion (Binder, 2016). Most likely, a hierarchy of processing circuits ranging from lower-level modality-specific cortex over multimodal regions up to top-level amodal areas in heteromodal cortex, indexing increasing levels of abstraction, establishes conceptual representations (Kuhnke et al., 2020; Kiefer and Harpaintner, 2020).

In the next sections, we provide a comprehensive overview of the latest research on embodied cognition in several cognitive domains and discuss important implications for learning and teaching.

Embodiment of conceptual memory for objects

Concepts stored in semantic long-term memory (Tulving, 1972) code the sum of our sensory and motor experiences with the environment in a categorical fashion (Kiefer and Pulvermüller, 2012). For instance, the concept “beech” includes the information that a beech has a trunk, has branches, has leaves, the wood is used for carpentering, etc. Embodied cognition theories predict that concepts, the constituents of thought, are rooted in perception and action. Brain activation studies (for an overview see, Kiefer and Barsalou, 2013; Kiefer and Harpaintner, 2020) have provided converging evidence on the differential involvement of sensory-motor brain areas in the processing of words and concepts of different kinds (e.g. vision-related concepts vs. action-related concepts). They demonstrated activation in sensory-motor brain areas in a range of conceptual tasks (e.g. Hoenig et al., 2008; Simmons et al., 2005). In fact, conceptual and perceptual processing functionally and neuroanatomically overlap in sensory brain regions (Kiefer et al., 2008): Visual recognition of words denoting objects, for which acoustic features are highly relevant (e.g. sound-related concepts such as “telephone”), ignited cell assemblies in auditory brain regions that were also activated by sound perception. Processing of action words (e.g. “to throw”) elicited activity in motor areas (Hauk et al., 2004), partially overlapping with activity induced by real movements of the corresponding limb (e.g. hand-related movements). Similarly, picture of appetizing food or words related to smell (“cinnamon”) activated regions in the brain associated with gustation and olfaction (Simmons et al., 2005; Gonzalez et al., 2006).

More recent fMRI studies (Kuhnke et al., 2020; Popp et al., 2019a), however, indicated that not only modality-specific brain areas, as defined by localizer tasks (e.g. acoustic localizer: listening to sounds; motor localizer: moving the hands), but also adjacent higher-level multimodal regions responded to concepts with a high relevance of a given feature type (e.g. acoustic or action features). Activity in both modality-specific and multimodal regions was modulated by task demands indicating conceptual flexibility at various levels of the conceptual processing hierarchy. As already outlined at the beginning of this article, a hierarchy of processing circuits establishes conceptual representations (Kuhnke et al., 2020; Popp et al., 2019a; Kiefer and Harpaintner, 2020). These findings thus support hybrid models of conceptual representations combining assumptions of modal embodiment theories with those of amodal theories (Kiefer and Pulvermüller, 2012; Patterson and Ralph, 2016).

Embodied cognition theories predict that conceptual memory traces in sensory-motor areas are established through the learning-based formation of cortical cell assemblies as a direct consequence of the experience with the referent. One line of evidence comes from training studies on the experience-dependent acquisition of concepts for novel objects (James and Gauthier, 2003; Kiefer et al., 2007). For instance, human participants learned concepts of novel objects (“nobjects”) under different training conditions (Kiefer et al., 2007): Participants either made an action pantomime toward a feature of the novel object, which signaled a specific object function, or pointed to it. At test, only in the pantomime group, in which a meaningful action was performed toward the object during training, was early activation in frontal motor regions and later activation in occipito-parietal visual-motor regions found during conceptual processing. This indicates that action representations essentially constitute the concept. In the pointing training group, in which the action during training was not meaningfully related to the object, this sensory-motor activity was absent, suggesting that concepts were not grounded in action. Furthermore, at the performance level, initial learning was superior in the pantomime group compared to the pointing group (See also the video about the training study in the Video 1).

The second line of evidence comes from studies investigating experience-dependent formation of conceptual representations in experts, with real objects. For instance, only professional musicians, but not musical laypersons, activate auditory association cortex when accessing conceptual knowledge about musical instruments (Hoenig et al., 2011). Together with similar expertise studies (Beilock et al., 2008; Lyons et al., 2010), these findings confirm that the grounding of concepts in the sensory-motor circuits of the brain is the result of repeated meaningful interactions with the referent. If this experience is lacking, concepts are less rich and mainly based on verbal associations (Solomon and Barsalou, 2004). In fact, we found that deaf individuals, who could not rely on the auditory input channel since early childhood, recruited language brain systems more strongly than hearing individuals in a conceptual task (Trumpp and Kiefer, 2018). This study also showed that deaf individuals compensated the loss of the auditory channel by additional recruitment of visual and motor areas.

Returning to the beech example, when confronted with the name “beech,” for instance, a layperson may be able to retrieve other words typically co-occurring with the word “beech” such as “tree” or “forest” without having a clear idea how a beech really looks like and what it is used for. In contrast, forest rangers have profound experience and knowledge about the shape of a beech and its leaves, the texture of its bark as well as the smell and the use of its wood. This notion of experience-dependent plasticity of conceptual representation supports teaching approaches like scenic learning in foreign language teaching, where vocabulary training is accompanied by meaningful gestures and movements not only for action word but for nouns as well (Macedonia and Klimesch, 2014).

Embodiment of conceptual memory for numbers

Not only concrete object concepts, but also abstract concepts such as numbers seem to be related to perception and action. Processing number concepts (e.g. knowing that 6 is greater than 4) has been shown to involve the sensory-motor systems similar to concrete object concepts in several studies. Firstly, accessing number magnitude depends on a mental number line, which resembles visuo-spatial representations (Dehaene, 1992). Behavioral number comparison experiments (e.g. deciding, which digit is larger, 6 or 2)

provide objective evidence for the existence of an analog mental number line that has a logarithmic scale similar to the mental representation of the size or intensity of sensory stimuli (Nieder, 2005). Furthermore, neuroimaging studies consistently showed that number magnitude is represented in a parietal area (intraparietal sulcus), which is also involved in processing space (Nieder, 2005). Secondly, in addition to visuo-spatial representation, numbers are rooted in the motor system, particularly in hand actions (see also Lindemann et al., 2007). Most importantly, hand actions related to finger counting habits used in childhood to learn numbers are still activated in adults when they process numbers: Intercultural studies showed that reaction times in a number comparison experiment are strongly influenced by the finger counting habits typically used in a given culture (Domahs et al., 2010). Only in German participants with unimanual finger counting habits for numbers up to five and bimanual habits for numbers greater than five were number comparisons slower when they involved numbers both below and above five, that is numbers that require one versus two hands in the German finger counting system. In Chinese participants with an unimanual symbolic finger counting system for these numbers, this effect was absent. Moreover, finger recognition in childhood predicted later arithmetic abilities (Noel, 2005) demonstrating the relevance of sensory-motor systems for coding the meaning of numbers. In support of a grounding of numbers in the motor systems, processing number concepts enhanced activity in the motor cortex in adults (Tschemtscher et al., 2012). Hence, number concepts appear to be rooted in both visuo-spatial and action-related representations. Explicitly training children in finger counting as well as in spatial analogs of number magnitude accelerates number learning and has beneficial effects on subsequent mathematical performance even in their later school or professional career (see, Fischer et al., 2011).

Embodiment of memory for abstract concepts

Besides numbers, other abstract concepts such as abstract ideas (“justice”, “love”) or scientific terms (“gravitation”, “radiation”) also do not refer to physical objects that can be directly experienced by the senses. The representation of abstract concepts in general imposes challenges for all classes of theories of conceptual representation (see also, Dove, 2016), because their content is typically difficult to define (Kiefer and Harpaintner, 2020). Furthermore, abstract concepts are more complex and ambiguous than concrete concepts, because they apply to rather heterogeneous situations (Hoffman et al., 2013; Barsalou and Wiemer-Hastings, 2005). Therefore, all theories have to deal with a high degree of conceptual flexibility. Abstract concepts are a particular challenge for embodied cognition theories, because at the first glance, it is hard to imagine how concepts without a referent, which can be perceived or acted upon, could be grounded in the sensory and motor brain systems (Dove, 2009, 2016).

For a long time, it has been argued that abstract concepts require amodal, symbolic (Mahon and Caramazza, 2009) or verbal representations (Paivio, 1986). This view dates back to the influential work by Paivio on his Dual Coding Theory (Paivio, 1986). Paivio (1986) assumed that abstract concepts are stored in a verbal-symbolic code, while concrete concepts rely on both a visual imaginary and a verbal-symbolic code.

However, in the recent years, refined embodied cognition theories have been proposed, which are able to explain the representation of abstract concepts: We (Kiefer and Harpaintner, 2020) and others have suggested that abstract concepts might not only be grounded in the perception of external events such as situations, but also in the introspection of internal mental states and in mentalizing social constellations (Barsalou and Wiemer-Hastings, 2005; Harpaintner et al., 2018; Kiefer and Barsalou, 2013; Borghi and Binkofski, 2014) or in processing affective states (Kousta et al., 2011).

In support of refined embodied cognition theories, several studies indicate that abstract concepts not only depend on the verbal system, but also on a variety of modal systems including perception, action, emotion and introspection (for reviews, see Borghi et al., 2017; Kiefer and Harpaintner, 2020). For instance, in a property listing study (Harpaintner et al., 2018), participants generated a substantial proportion of introspective, emotional and social properties, in addition to verbal associations. In terms of quantity, however, sensory and motor properties played the most crucial role in this study. The broad diversity in participants’ listings is consistent with refined embodied cognition theories by showing that the semantic content of abstract concepts includes various semantic features. These results also suggest that only one relatively small subgroup of abstract concepts is predominantly related to verbal associations.

Converging evidence in support of refined embodied cognition theories is provided by brain activation studies. Neuroimaging results revealed activity in areas related to emotions (Vigliocco et al., 2014), mental states (Wilson-Mendenhall et al., 2013) and social interactions (Wilson-Mendenhall et al., 2013), when abstract concepts were processed. Furthermore, similar to concrete concepts, subgroups of abstract concepts have been shown to activate visual and motor areas, also involved in perception and action (Harpaintner et al., 2020). Abstract physical concepts related to periodicity (e.g. “frequency”) activated postcentral and parietal brain regions, regions found to be active when performing rhythmic movements (Mason and Just, 2016).

Hence, these findings demonstrate that even fairly abstract concepts are rooted in modal systems including emotions, introspection, perception and action in line with embodied cognition theories. This sort of evidence is incompatible with Piaget’s (Inhelder and Piaget, 1958) view that scientific or mathematical concepts exclusively build upon abstract formal logical reasoning. We argue that a grounding in experiences is necessary for a rich understanding of abstract concepts, whereas knowledge is superficial, when only based on verbal instruction. We therefore propose that abstract concepts should be taught by providing learners with meaningful visualizations or movements.

Embodiment of memory for events

Past events such as incidences happening at our summer vacation are stored in episodic memory, the long-term memory system for events (Tulving, 1972). When we recall these events, we reactivate stored sensory-motor experiences collected during the initial learning episode and not only abstract-symbolic verbal knowledge (Engelkamp and Jahn, 2003; Waldhauser et al., 2016). These reactivations of acquired sensory-motor memory traces are not epiphenomenal, but are essential for memory performance (Waldhauser et al., 2016). The so-called enactment effect nicely illustrates the importance of rich sensory-motor experiences for episodic memory (Engelkamp and Jahn, 2003): Participants remembered a list of action verbs better when they performed the corresponding actions in the learning phase compared with a condition when they simply read the words. Observing others, who performed the action, also improved subsequent memory compared with reading, but was inferior to self-performed actions (Senkfor et al., 2002). Neurophysiological recordings of brain activity during memory recall revealed an activation of motor areas only for self-performed actions during learning (Senkfor et al., 2002) suggesting that action representations established during word learning were reactivated and facilitated memory retrieval.

Reactivations of stored experiences in modality-specific brain areas (i.e. areas specifically engaged in perception or action) during memory retrieval are not only observed for self-performed actions, but also for sensory information associated with the learning episode such as vision or sound (Ranganath et al., 2004; Waldhauser et al., 2016). Hence, episodic memory is based on a reinstatement of sensory and motor experiences and in this sense rooted in the sensory-motor systems (Engelkamp and Jahn, 2003). Providing the relevant sensory and motor information during learning improves subsequent memory performance compared with pure verbal learning. The results suggest that teaching strategies such as (language) learning through drama, “experiments” and outdoor activities, e.g. in science learning, are suitable to support the building of memory by providing sensory and motor information (e.g. Uysal and Yavuz, 2018). Moreover, the enactment effect indicates that vocabulary training of action verbs can be improved by the corresponding meaning-related movements.

Embodiment of reading and writing

Writing is a manual sensory-motor skill, which depends on the acquisition, storage and execution of complex motor programs (Palmis et al., 2017; Heuer, 2016). Reading, in contrast, is typically considered to be purely based on visual information (e.g. McClelland and Rumelhart, 1981). However, embodiment theories predict that reading performance might be influenced by writing techniques, because the motor programs and sensory experiences during writing (e.g. forming specific letters and words with a pen) are assumed to be implicitly re-activated during reading. As a consequence, our habitual writing techniques could affect reading performance. This issue is particularly important, because nowadays digital writing devices associated with the use of mobile phones, tablets or computers have frequently replaced writing by hand (for an overview, see Kiefer and Velay, 2016). The sensory-motor experiences during handwriting (haptic, motor, visual etc.) are quite different from those during typewriting, mouse clicking or tapping on digital devices. In particular, handwriting requires carefully reproducing the shape of each letter, whereas in typewriting no such grapho-motor component is present. Given that children in our present days may learn writing by typing on a computer or mobile phone much before they master handwriting, it is important to know how this dramatic change in writing habits in the digital age affects reading performance (Mangen and Balsvik, 2016). Consistent with embodiment theories, several training studies in preschool children and adults showed that handwriting training of new letters resulted in better letter recognition in a subsequent test than typing training (e.g. Longcamp et al., 2005). This demonstrates that handwriting, which links rich sensory-motor representations to perceptual letter shapes, improves subsequent letter reading performance compared with typewriting. Furthermore, visual recognition of letters activated motor regions of the brain, only when letters were trained by handwriting, but not when they were trained by typewriting (James and Engelhardt, 2012). This supports the assumption that sensory-motor experiences must be meaningfully related to the learning target (here: shaping a letter by writing vs. pressing a key associated with a letter) to result in stronger sensory-motor memory traces that facilitate learning (Kiefer et al., 2007).

Although some intervention studies seem to suggest a superiority of handwriting training over typing training on subsequent reading and writing performance in young children, results are mixed. Improved letter recognition after handwriting training compared to typing training was not always replicated (Kiefer et al., 2015). Unfortunately, effects at the word level are also heterogeneous: A superiority of handwriting over typing training on word writing performance (Cunningham and Stanovich, 1990; Kiefer et al., 2015) was not found in other studies (Vaughn et al., 1992; Ouellette and Tims, 2014). Mayer et al. (2020) therefore investigated the influence of the writing tool on the acquisition of literacy skills at the letter and word level with various tests in a large sample of kindergarten children ($n = 147$). Using closely matched letter learning games, children were trained with 16 letters by handwriting with a pencil on a sheet of paper, by writing with a stylus on a tablet computer, or by typing letters using a virtual keyboard on a tablet across seven weeks. Training using a stylus on a touchscreen is an interesting comparison condition for traditional handwriting, because the slippery surface of a touchscreen has lower friction than paper and thus increases difficulty of motor control. Visuo-spatial skills were also assessed, in order to test, whether the different training regimens affected cognitive domains other than written language. Children of the pencil group showed superior performance in letter recognition and improved visuo-spatial skills compared with keyboard training. Keyboard training, however, resulted in superior performance in word writing and reading compared with handwriting training with a stylus on the tablet, but not compared with the pencil group. These findings indicate that handwriting with pencil fosters acquisition of letter knowledge and improves visuo-spatial skills compared to

keyboarding. At least given the current technological state, writing with a stylus on a touchscreen seems to be the least favorable writing tool, possibly because of increased demands on motor control. Writing training with a stylus on a tablet led to inferior reading and writing performance at the word level compared with keyboarding. At the same time, the beneficial effects of hand-writing training on letter recognition and visuo-spatial skills were less pronounced compared with writing with a pencil.

Conclusion

Accumulating evidence indicates that cognition is fundamentally rooted in the modal systems of perception, action and introspection. Recent studies suggest that even the most complex and abstract thoughts are grounded in these modal systems and are not abstract-symbolic. Findings from many cognitive domains thus demonstrate that human cognition depends on appropriate sensory-motor experiences to function at the highest level. Embodied cognition theories and associated empirical evidence is therefore naturally highly relevant for many issues associated with education (O'Loughlin, 2006; Kiefer and Trumpp, 2012; Kiefer et al., 2022): Embodied cognition theories highlight the relevance of experiential interactions with our environment during learning resulting in more enduring and—perhaps most importantly—in richer knowledge. These experiences frequently include perception and action, but may also be related to introspection of emotional and other mental states.

Embodied cognition theories thus confirm and considerably expand Piaget's theory of cognitive development, which has influenced education for many decades (Benz et al., 2015). Piaget's theory emphasized the role of sensory-motor experience for children before adolescence in line with embodied cognition theories (Piaget, 1952). Reception of Piaget's theory led to changes in school curricula and promoted integration of active, self-regulated learning phases into the lessons. However, according to Piaget's stage model (Inhelder and Piaget, 1958), secondary school students as well as adults acquire knowledge based on formal cognitive operations, detached from sensory-motor experiences. Embodied cognition theories and available evidence in this field, however, contradict Piaget's notion of formal logic operations being detached from modality-specific experiences and emphasize the importance of perception, action and introspection even for most abstract adult cognition. Hence, according to the embodied cognition view, discovery learning and experiential learning activities foster knowledge acquisition also in secondary schools and in adult education.

Consider the example of learning about a beech at the beginning: According to an amodal view of cognition, it would be sufficient for the teacher to verbally describe the shape of the trunk and the leaves, the texture of its bark, the use of beech wood for carpentry etc. in a written text. A direct experience with the object is not necessary. According to embodiment theories, rich knowledge about the unfamiliar concept beech can only be acquired when the students can see, touch and act upon the object. A pure verbal description should result in impoverished, less durable knowledge. As conceptual cognition is the basis for thought, language and problem solving, knowledge embodied in perception, action and introspection is crucial for the developing mind and for achieving educational success.

Supplementary material

Supplementary video related to this chapter can be found at <https://doi.org/10.1016/B978-0-12-818630-5.14009-6>

Acknowledgments

This research was supported by a grant of the German Research Foundation (DFG Ki 804/7-1) to MK.

References

- Anderson, J.R., 1983. *The Architecture of Cognition*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ.
- Barsalou, L.W., Wiemer-Hastings, K., 2005. Situating abstract concepts. In: Pecher, D., Zwaan, R. (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*. Cambridge University Press, New York.
- Barsalou, L.W., Dutriaux, L., Scheepers, C., 2018. Moving Beyond the Distinction Between Concrete and Abstract Concepts. *Philosophical Transactions of the Royal Society B-Biological Sciences*, p. 373.
- Barsalou, L.W., 2008. Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Beilock, S.L., Lyons, I.M., Mattarella-Micke, A., Nusbaum, H.C., Small, S.L., 2008. Sports experience changes the neural processing of action language. *Proc. Natl. Acad. Sci. U. S. A.* 105, 13269–13273.
- Benz, C., Peter-Koop, A., Grüßing, M., 2015. *Frühe mathematische Bildung: Mathematiklernen der Drei- bis Achtjährigen*. Springer, Wiesbaden.
- Bermeitinger, C., Kiefer, M., 2012. Embodied concepts. In: Koch, S.C., Fuchs, T., Summa, M., Müller, C. (Eds.), *Body Memory, Metaphor and Movement*. John Benjamins Publishing Company, Amsterdam.
- Binder, J.R., 2016. In defense of abstract conceptual representations. *Psychon. Bull. Rev.* 23, 1096–1108.
- Borghi, A.M., Binkofski, F., 2014. *Words as Social Tools: An Embodied View on Abstract Concepts*. Springer-Verlag, New York, USA.
- Borghi, A.M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., Tummolini, L., 2017. The challenge of abstract concepts. *Psychol. Bull.* 143, 263–292.
- Cunningham, A.E., Stanovich, K.E., 1990. Early spelling acquisition—writing beats the computer. *J. Educ. Psychol.* 82, 159–162.

- Dehaene, S., 1992. Varieties of numerical abilities. *Cognition* 44, 1–42.
- Domahs, F., Moeller, K., Huber, S., Willmes, K., Nuerk, H.C., 2010. Embodied numerosity: implicit hand-based representations influence symbolic number processing across cultures. *Cognition* 116, 251–266.
- Dove, G., 2009. Beyond perceptual symbols: a call for representational pluralism. *Cognition* 110, 412–431.
- Dove, G., 2016. Three symbol ungrounding problems: abstract concepts and the future of embodied cognition. *Psychon. Bull. Rev.* 23, 1109–1121.
- Engelkamp, J., Jahn, P., 2003. Lexical, conceptual and motor information in memory for action phrases: a multi-system account. *Acta Psychol.* 113, 147–165.
- Fischer, U., Moeller, K., Bientzle, M., Cress, U., Nuerk, H.C., 2011. Sensori-motor spatial training of number magnitude representation. *Psychon. Bull. Rev.* 18, 177–183.
- Gonzalez, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuan, A., Belloc, V., Avila, C., 2006. Reading cinnamon activates olfactory brain regions. *Neuroimage* 32, 906–912.
- Harpaintner, M., Trumpp, N.M., Kiefer, M., 2018. The semantic content of abstract concepts: a property listing study of 296 abstract words. *Front. Psychol.* 9, 1748.
- Harpaintner, M., Sim, E.-J., Trumpp, N.M., Ulrich, M., Kiefer, M., 2020. The grounding of abstract concepts in the motor and visual system: an fMRI study. *Cortex* 124, 1–22.
- Hauk, O., Johnsrude, I., Pulvermüller, F., 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.
- Heuer, H., 2016. Technologies shape sensorimotor skills and abilities. *Trends Neurosci. Educ.* 5, 121–129.
- Hoenig, K., Sim, E.-J., Bochev, V., Herrnberger, B., Kiefer, M., 2008. Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motion and motor-related areas. *J. Cognit. Neurosci.* 20, 1799–1814.
- Hoenig, K., Müller, C., Herrnberger, B., Spitzer, M., Ehret, G., Kiefer, M., 2011. Neuroplasticity of semantic maps for musical instruments in professional musicians. *Neuroimage* 56, 1714–1725.
- Hoffman, P., Pobric, G., Drakesmith, M., Ralph, M.A.L., 2012. Posterior middle temporal gyrus is involved in verbal and non-verbal semantic cognition: evidence from rTMS. *Aphasiology* 26, 1119–1130.
- Hoffman, P., Ralph, M.A.L., Rogers, T.T., 2013. Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words. *Behav. Res. Methods* 45, 718–730.
- Humphreys, G.W., Riddoch, M.J., Quinlan, P.T., 1988. Cascade processes in picture identification. *Cognit. Neuropsychol.* 5, 67–103.
- Inhelder, B., Piaget, J., 1958. *The Growth of Logical Thinking From Childhood to Adolescence: An Essay on the Construction of Formal Operational Structures*. Routledge & Kegan Paul, London.
- Jackendoff, R., 1990. *Semantic Structures*. MIT Press, Cambridge, Ma.
- James, K.H., Engelhardt, L., 2012. The effects of handwriting experience on functional brain development in pre-literate children. *Trends Neurosci. Educ.* 1, 32–42.
- James, T.W., Gauthier, I., 2003. Auditory and action semantic features activate sensory-specific perceptual brain regions. *Curr. Biol.* 13, 1792–1796.
- Kiefer, M., Barsalou, L.W., 2013. Grounding the human conceptual system in perception, action, and internal states. In: Prinz, W., Beisert, M., Herwig, A. (Eds.), *Action Science: Foundations of an Emerging Discipline*. MIT Press, Cambridge, pp. 381–407.
- Kiefer, M., Harpaintner, M., 2020. Varieties of abstract concepts and their grounding in perception or action. *Open Psychol.* 2, 119–137.
- Kiefer, M., Pulvermüller, F., 2012. Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex* 48, 805–825.
- Kiefer, M., Spitzer, M., 2001. The limits of a distributed account of conceptual knowledge. *Trends Cognit. Sci.* 5, 469–471.
- Kiefer, M., Trumpp, N.M., 2012. Embodiment theory and education: the foundations of cognition in perception and action. *Trends Neurosci. Educ.* 1, 15–20.
- Kiefer, M., Velay, J.L., 2016. Writing in the digital age. *Trends Neurosci. Educ.* 5, 77–81.
- Kiefer, M., Arndt, P., Mayer, C., 2022. Embodied concepts: basic mechanisms and their implications for learning and memory. In: Macrine, S.L., Fugate, J. (Eds.), *Movement Matters: How Embodied Cognition Informs Teaching and Learning*. MIT Press, Cambridge, pp. 25–38.
- Kiefer, M., Sim, E.-J., Liebich, S., Hauk, O., Tanaka, J.W., 2007. Experience-dependent plasticity of conceptual representations in human sensory-motor areas. *J. Cognit. Neurosci.* 19, 525–542.
- Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., Hoenig, K., 2008. The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J. Neurosci.* 28, 12224–12230.
- Kiefer, M., Schuler, S., Mayer, C., Trumpp, N.M., Hille, K., Sachse, S., 2015. Handwriting or typewriting? The influence of pen- or keyboard-based writing training on reading and writing performance in preschool children. *Adv. Cognit. Psychol.* 11, 136–146.
- Kousta, S.T., Vigliocco, G., Vinson, D.P., Andrews, M., DEL Campo, E., 2011. The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34.
- Kuhnke, P., Kiefer, M., Hartwigsen, G., 2020. Task-dependent recruitment of modality-specific and multimodal regions during conceptual processing. *Cereb. Cortex* 30 (7), 3938–3959.
- Lakoff, G., Johnson, M., 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, New York.
- Levelt, W.J., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. Discussion 38–75.
- Lindemann, O., Abolafia, J.M., Girardi, G., Bekkering, H., 2007. Getting a grip on numbers: numerical magnitude priming in object grasping. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1400–1409.
- Longcamp, M., Zerbato-Poudou, M.T., Velay, J.L., 2005. The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *Acta Psychol.* 119, 67–79.
- Lyons, I.M., Mattarella-Micke, A., Cieslak, M., Nusbaum, H.C., Small, S.L., Beilock, S.L., 2010. The role of personal experience in the neural processing of action-related language. *Brain Lang.* 112, 214–222.
- Macedonia, M., Klimesch, W., 2014. Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind Brain Educ.* 8, 74–88.
- Machery, E., 2007. Concept empiricism: a methodological critique. *Cognition* 104, 19–46.
- Mahon, B.Z., Caramazza, A., 2009. Concepts and categories: a cognitive neuropsychological perspective. *Annu. Rev. Psychol.* 60, 27–51.
- Mahon, B.Z., 2015. The burden of embodied cognition. *Can. J. Exp. Psychol.* 69, 172–178.
- Mangen, A., Balsvik, L., 2016. Pen or keyboard in beginning writing instruction? Some perspectives from embodied cognition. *Trends Neurosci. Educ.* 5 (3), 99–106.
- Markie, P., 2008. Rationalism vs. empiricism. In: Zalta, E.D. (Ed.), *Stanford Encyclopedia of Philosophy*, Fall 2008 ed. The Metaphysics Research Lab, Stanford.
- Mason, R.A., Just, M.A., 2016. Neural representations of physics concepts. *Psychol. Sci.* 27, 904–913.
- Mayer, C., Wallner, S., Budde-Spengler, N., Braunert, S., Arndt, P.A., Kiefer, M., 2020. Literacy training of kindergarten children with pencil, keyboard or tablet stylus: the influence of the writing tool on reading and writing performance at the letter and word level. *Front. Psychol.* 10.
- McClelland, J.L., Rogers, T.T., 2003. The parallel distributed processing approach to semantic cognition. *Nat. Rev. Neurosci.* 4, 310–322.
- McClelland, J.L., Rumelhart, D.E., 1981. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychol. Rev.* 88, 375–407.
- Mettey, L., Cuadrado, S.R., Bahrami, N., Vigliocco, G., 2012. Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804.
- Nieder, A., 2005. Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 6, 177–190.
- Noel, M.P., 2005. Finger gnosis: a predictor of numerical abilities in children? *Child Neuropsychol.* 11, 413–430.
- O’loughlin, M., 2006. *Embodiment and Education*. Springer, Heidelberg.
- Ouellette, G., Tims, T., 2014. The write way to spell: printing vs. typing effects on orthographic learning. *Front. Psychol.* 5.
- Paivio, A., 1986. *Mental Representations: A Dual Coding Approach*. Oxford University Press, New York.
- Palmis, S., Danna, J., Velay, J.L., Longcamp, M., 2017. Motor control of handwriting in the developing brain: a review. *Cognit. Neuropsychol.* 34, 187–204.
- Patterson, K., Ralph, M.A.L., 2016. The hub-and-spoke hypothesis of semantic memory. In: Hickok, G., Small, S.L. (Eds.), *Neurobiology of Language*. Academic Press, Burlington.
- Piaget, J., 1952. *The Origins of Intelligence in Children*. International Universities Press, New York.

- Popp, M., Trumpp, N.M., Kiefer, M., 2019a. Processing of action and sound verbs in context: an fMRI study. *Transl. Neurosci.* 10, 200–222.
- Popp, M., Trumpp, N.M., Sim, E.J., Kiefer, M., 2019b. Brain activation during conceptual processing of action and sound verbs. *Adv. Cognit. Psychol.* 15, 236–255.
- Pulvermüller, F., Fadiga, L., 2010. Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360.
- Pylyshyn, Z.W., 1984. *Computation and Cognition: Towards a Foundation for Cognitive Science*. MIT Press, Cambridge.
- Quillian, M.R., 1969. The teachable language comprehender. *Commun. ACM* 12, 459–476.
- Ranganath, C., Cohen, M.X., Dam, C., D'Esposito, M., 2004. Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* 24, 3917–3925.
- Rogers, T.T., Lambon Ralph, M.A., Garrard, P., Bozeat, S., McClelland, J.L., Hodges, J.R., Patterson, K., 2004. Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychol. Rev.* 111, 205–235.
- Runes, D.D., 1962. *Dictionary of Philosophy*. Littlefield, Adams, and Company, Totowa, NJ.
- Senkfor, A.J., Van Petten, C., Kutas, M., 2002. Episodic action for real objects: an ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *J. Cognit. Neurosci.* 14, 402–419.
- Simmons, W.K., Martin, A., Barsalou, L.W., 2005. Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cereb. Cortex* 15, 1602–1608.
- Solomon, K.O., Barsalou, L.W., 2004. Perceptual simulation in property verification. *Mem. Cognit.* 32, 244–259.
- Trumpp, N.M., Kiefer, M., 2018. Functional reorganization of the conceptual brain system after deafness in early childhood. *PLoS One* 13.
- Trumpp, N.M., Traub, F., Kiefer, M., 2013. Masked priming of conceptual features reveals differential brain activation during unconscious access to conceptual action and sound information. *PLoS One* 8, e65910. <https://doi.org/10.1371/journal.pone.0065910>.
- Tschentscher, N., Hauk, O., Fischer, M.H., Pulvermüller, F., 2012. You can count on the motor cortex: finger counting habits modulate motor cortex activation evoked by numbers. *NeuroImage* 59, 3139–3148.
- Tulving, E., 1972. Episodic and semantic memory. In: Tulving, E., Donaldson, W. (Eds.), *Organization of Memory*. Academic Press, New York.
- Tyler, L.K., Moss, H.E., 2001. Towards a distributed account of conceptual knowledge. *Trends Cognit. Sci.* 5, 244–252.
- Uysal, N.D., Yavuz, F., 2018. Language learning through drama. *Int. J. Learn. Teach.* 10, 376–380.
- Vaughn, S., Schumm, J.S., Gordon, J., 1992. Early spelling acquisition—does writing really beat the computer? *Learn. Disabil. Q.* 15, 223–228.
- Vigliocco, G., Kousta, S.T., Della Rosa, P.A., Vinson, D.P., Tettamanti, M., Devlin, J.T., Cappa, S.F., 2014. The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777.
- Waldhauser, G.T., Braun, V., Hanslmayr, S., 2016. Episodic memory retrieval functionally relies on very rapid reactivation of sensory information. *J. Neurosci.* 36, 251–260.
- Wilson-Mendenhall, C.D., Simmons, W.K., Martin, A., Barsalou, L.W., 2013. Contextual processing of abstract concepts reveals neural representations of nonlinguistic semantic content. *J. Cognit. Neurosci.* 25, 920–935.

Learning analytics

James Fiacco^a, Shiyang Jiang^b, David Adamson^c, and Carolyn P. Rosé^d, ^a Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, United States; ^b School of Education, North Carolina State University, Raleigh, NC, United States; ^c Turnitin, Pittsburgh, PA, United States; and ^d Language Technologies Institute, Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	75
Hot topics in recent work	75
Focus area: analytics of unstructured natural language data	77
Discourse analytics in computer-supported collaborative learning	77
Discourse analytics in automated analysis of student writing	77
Conclusions and vision for the field	79
Acknowledgments	79
References	80

Introduction

For more than two decades, the fields of Educational Data Mining (Baker and Yacef, 2009; Baker and Inventado, 2014) and Learning Analytics (Siemens and Baker, 2012; Siemens, 2013) have received growing prominence in research, policy and public practice literature. These fields have evolved from the applied disciplines of Machine Learning, Intelligent Tutoring Systems and Data Mining, which, in turn, have grown out of Applied Statistics, Education, Psychology, Cognitive Science, and Computational Linguistics. Two separate research societies formally focus on these areas: namely the International Educational Data Mining Society and the Society for Learning Analytics Research. However, the work has become so ubiquitous that it features prominently across the gamut of research societies focused on a wide range of areas under the broad heading of technology enhanced learning. Due to its expansive nature and the existence of earlier review articles recounting the history and scope, this chapter will not attempt to offer a comprehensive review of research in the areas of Learning Analytics or Educational Data Mining. Instead, the aim is to take a step back and consider some key lessons learned as well as potential steps forward.

At its core, a fundamental tenet of both Learning Analytics and Educational Data Mining is the idea that decision making regarding the administration of learning at all levels should be guided by data—both data about learners in general, and data about specific learners. In order to support design and decision making, causal models and appropriate generalizations are needed. Although these claims may come with strong caveats (Rosé et al., 2019). In particular, in embracing the value of empiricism, we must acknowledge the dangers of atheoretical empiricism (Wise and Shaffer, 2015). Theoretical frameworks are grounded in empiricism, but safeguarded through deep questioning and meaning making, which challenge us to separate spurious correlations from meaningful generalizations. Recent work on bias and fairness in Artificial Intelligence raise awareness of the harms that may arise when such consideration is glossed over (Gardner et al., 2019; Ferguson et al., 2016). With these tensions in mind—valuing empiricism while challenging total empiricism with theory grounded questioning—we embark upon this partial, but reflective, review.

The work of the field of Learning Sciences is to contribute to theories of how people learn, and then to work within those theoretical frameworks to effect positive change in the world. Computational tools, which include machine learning approaches, can serve as lenses through which researchers may make observations that contribute to theory, as machinery used to encode operationalizations of theoretical constructs, and as languages to build assessments that measure the world in terms of these operationalizations. In short, they are augmentations of human observation. They are more limited in the sense that, through the necessary discretization that takes place in applying them, the richness of signal is reduced from that in the real world. However, they are also more powerful in the sense of the speed and ubiquity of the observation that is possible. Thus, though the Learning Sciences community may include those who can fundamentally extend the abilities of computational approaches, the work of the community is in using these approaches to advance separate goals. On the other hand, that does not mean it is advisable for Learning Sciences researchers to be disinterested consumers of this technology. It would be more advantageous to strive to understand how to appropriate fully this technology as it grows and changes so that the work of the community can be accomplished more effectively.

Hot topics in recent work

In the past few years, capitalizing on its affordances, some of the most exciting work in Learning Analytics has aimed to scale up that technology in order to serve the large and diverse populations that recent trends in technology enhanced learning require.

The recent circumstances have revealed the need for such consideration, especially in times of crisis, as we have seen in the emergency move to online instruction necessitated by the pandemic. Taking a more expansive view of what can be learned about learning from data at scale—in terms of population size, temporal scope, geographic expanse, and subject matter diversity—some specific areas of focus including temporal analysis, multimodal analysis, and social network analysis have gained prominence. Bringing all these threads together, a reconsideration of the role of theories in rendering the technology safe and effective has again emerged.

Temporal analysis, also known as time-series analysis, has become an essential approach for investigating learning trajectories, styles, and strategies using process data, such as log data from digital platforms. The literature indicates that temporal analysis, combined with descriptive statistics, could enrich understanding of complex learning processes. In a recent study conducted by [Saint et al. \(2020\)](#), an online system captured computer engineering undergraduate students' interactions with learning materials, such as correctly solving assessment items and playing course videos. These interactions were categorized into self-regulated learning processes, such as setting goals and making plans. The study suggested that combining simple frequency measures and temporal analysis offered novel insights into patterns of self-regulated learning. While temporal analysis provided deep insights into learning processes, the field raised the issue of the choice of techniques affecting findings. [Matcha et al. \(2019\)](#) compared three different techniques (i.e., sequence, process, and network techniques) for identifying learning styles. As the three techniques generated different results, the authors argued that researchers and practitioners should interpret these results based on well-grounded learning theories and data analysis should also be based on such theories for generating meaningful learning strategies. Given the power of temporal analysis in expanding understanding of learning processes, researchers have devoted efforts to turn such insights into information that educators could use to monitor learning progress. Data storytelling (DS) principles have been applied to learning dashboards to support educators in interpreting findings from temporal analysis. An evaluation study with teachers demonstrated that integrating DS principles held the promise of helping teachers to interpret complex learning processes effectively ([Echeverría et al., 2018](#)). Furthermore, it has been noted that providing multiple layers of stories and engaging end users to select and combine layers to locate relevant information for feedback could support active exploration of learning progress ([Martinez-Maldonado et al., 2020](#)). Investigating strategies to engage end users to interpret results from temporal analysis is continuing to be a promising future research direction.

As learning activities are much broader than what digital platforms can currently provide, multimodal learning analytics has gained momentum with the aim of examining patterns based on discourse, gesture, and other modalities in a wide range of settings, especially collaborative learning settings. Based on the common understanding that communication and collaboration is naturally multimodal, [Cherubini et al. \(2008\)](#) collected multimodal data, including gaze movements, utterances, and gestures to build a model to detect misunderstanding during collaborative tasks. In this study, students worked in small groups to use a mapping tool to make a plan for organizing a festival on their campus. The study revealed that combining a linguistic model of utterances and eye-tracking data was effective for detecting misunderstandings. Contributing to examining collaborative learning with multimodal analytics, [Schneider et al. \(2020\)](#) captured moments when students focused on a common reference in identifying design principles for organizing warehouses in small groups. The study proposed that these moments provided nuanced information about how students negotiated ideas. In another study, researchers utilized multimodal analytics to derive deep insights about team-based, project-based learning activities in face-to-face settings ([Spikol et al., 2018](#)). It was a novel piece of work that was situated in a project-based learning setting that involved complex interactions among students and interactions between students and digital and physical materials. Collectively, these studies demonstrated that combining modes of data could uncover novel and valuable insights about collaborative learning. However, a single modality of data might be more suitable for real-time interventions ([Vrzakova et al., 2020](#)). There is need for more work in balancing trade-offs between the interpretability of complex learning processes such as collaborative learning and usability of multimodal analysis for real-time intervention.

Social network analysis has been widely used to probe collaborative learning. Addressing the gap within the theoretical framing in learning analytics research, [Kent and Cukurova \(2020\)](#) operationalized collaborative cognitive load theory ([Kirschner et al., 2018](#)) using social network analysis. They generated analytics of interactivity gains and coordination costs from students' interactions in online discussions. This study discussed the importance of visualizing these theoretical constructs for educators to support students' online collaboration. Meanwhile, studies in social network analysis have revealed that different methodological choices influenced the validity of the measurements, changed the insights obtained, and impacted the generalizability of results ([Saqr et al., 2020](#)). This implies that methodological choices should be presented in detail to ensure reproducibility and generalizability. Furthermore, intervention studies have been conducted to evaluate the potential of using social network analysis to monitor real-time collaborative learning. Through an experimental, observational repeated-measurement design, researchers found that social network analysis could be used to identify aspects related to teamwork in need of improvement, inform the design and implementation of intervention, and assess the efficacy of such intervention ([Saqr et al., 2018](#)). Research in real-time social network analysis is increasing and is worthy of study for supporting collaborative learning.

The scalability of learning analytics is receiving increasing attention as studies have shown evidence that these techniques enriched interpretation of learning in contexts and provided real-time feedback for end users. As an example, [Hutt et al. \(2019\)](#) used generic interaction features such as watching a video to model affect in math learning with around 70,000 students. The generic interaction features were used to ensure scalability and generalizability of the study findings. Students answered affect surveys after using a math learning platform. The authors explained that, in general, these generic interaction features could effectively predict affective states. Also, they suggested that some features were more closely related to affective states than others. This study implied

that interactions with platforms should be aligned with affective states so that these generic interaction features could be used for such prediction tasks. More work in scalability and critical evaluation on scaling learning analytics across contexts is needed to close the gap between gaining novel insights from learning analytics to translating these insights into concrete recommendations for practice at scale.

Focus area: analytics of unstructured natural language data

In the wake of the pandemic, we must acknowledge having been reminded of the importance of human contact and human interaction. In preparing for potential future events of a similar nature, it would benefit us to work toward more effective support for human-human interaction online. Human language is incredibly versatile and unique as a form of data in affording human interpretability, context-specificity, creative symbolization, levels of abstraction, and juxtaposition of frames of reference. Work in the area of natural language processing offers the opportunity to challenge the boundaries of what is possible to achieve within the neural revolution, the current paradigm of machine learning that drives recent innovation. Within this scope, we engage directly with the basis of what makes problems learnable: representations. We use these powerful representations to bridge to those of other forms of data, as highlighted in the previous section. Discourse Analytics is one area within the field of Learning Analytics (Buckingham-Shum et al., 2013, 2014), which focuses on natural language data as the target. It includes processing of open response questions in educational contexts, and a large proportion of research in the area focuses on assessment of writing, but it encompasses more than that, including analysis of discussions occurring in discussion forums, chat rooms, microblogs, blogs, and even wikis.

Discourse analytics in computer-supported collaborative learning

Research in the field of Computer Supported Collaborative Learning (CSCL) has produced technology for automating analysis of collaborative processes in real time (Rosé et al., 2008; Mu et al., 2012; Gweon et al., 2013; Fu et al., 2016) and using this analysis to trigger in process support that increased the effectiveness of collaboration and student learning (Kumar et al., 2011; Dyke et al., 2012; Adamson et al., 2014). Many frameworks have been developed for automated analysis of student discussion, as it has repeatedly been shown to be valuable for assessing student learning (McLaren et al., 2007; Dascalu et al., 2015; Rosé et al., 2008; McLaren et al., 2007; Ai et al., 2010; Gweon et al., 2013; Fiacco and Rosé, 2018), supporting group learning (Kumar et al., 2007), and enabling effective group assignments (Wen et al., 2016). Some work has explicitly addressed the issue of whether these frameworks generalize across domains (Mu et al., 2012; Fiacco and Rosé, 2018; Fiacco et al., 2021), which is critical to enabling educators in a variety of fields to leverage these tools. Consensus building and knowledge integration behaviors such as transactive conversational contributions have been a major focus of work in social analytics. Many of these conversational behaviors involve articulation of mental models and drawing connections between one's thinking and that of others. Successful collaborative learning is broadly believed to be fostered by discourse patterns characterized by the articulation of reasoning and idea co-construction (Chan, 2013; Chinn and Clark, 2013; van Aalst, 2009), which have theoretical connections with cognitive conflict and learning (de Lisi and Golbeck, 1999). On the other hand, building a shared understanding need not imply coming to an agreement, or avoiding disagreement (Holtz et al., 2018). Though much work takes a coding and counting approach, other work challenges whether this is enough, and proposes techniques that consider co-occurring codes within time windows to gain a richer sense of process (Csanadi et al., 2018). More recent work focuses on newer theoretical frameworks for evaluating collaborative learning processes, such as Social Regulation and Socially Shared Regulation of Learning (Su et al., 2018).

Discourse analytics in automated analysis of student writing

Automatic writing evaluation and feedback technologies have the potential to facilitate students' ability to learn to produce high-quality essays (McGarrell and Verbeem, 2007; Roscoe et al., 2013), thus developing writing skills that are important for their future success (Burstein et al., 2017; Cotos, 2011). While this aspect of learning analytics applied to human language is not formally focused on human-human communication processes, the skill of writing is essential for students to gain in order to be effective communicators in their adult professional lives. Writing is a key skill that has suffered alongside other areas of instruction in the pandemic (Kuhfeld et al., 2020; Engzell et al., 2021; Hamilton and Ercikan, 2022).

The field has developed Automated Writing Evaluation (AWE) and feedback technologies to facilitate students in writing well-organized essays (i.e., one idea flows logically to the next). While these technologies (e.g., Coh-Metrix (McNamara et al., 2010), Writing Pal (Roscoe and McNamara, 2013), Revision Assistant (West-Smith et al., 2018), and AcaWriter (Knight et al., 2020)) offer significant benefits for providing feedback for idea organization and flow, typically, they rely on local features, such as discourse markers and other syntactic indicators of attempts to signal rhetorical structure.

Writing feedback that is both adapted to an individual student's text and to the expectations of a given assessment may lead both to greater improvements in writing quality and a perception of more helpful feedback (Burstein et al., 2001). Even in recent published work, for example, it is reported that after using the AWE tool, MI Write, both students and teachers commented that the tool was able to provide general feedback but could not point out specific weaknesses of the, often ill-structured, essays (Palermo and Wilson, 2020). The need for offering advice on improvement of structure and flow is partially addressed in work grounded in

sequences of rhetorical moves and steps for the specific genre of scientific writing, where missing moves or steps can be suggested (Cotos, 2011; Fiocco et al., 2019). However, this is a schema-based approach grounded in a specific genre of writing and does not enable a generalizable form of evaluation of rhetorical structure.

Currently, automatic writing evaluation is effective at measuring the qualities already exhibited within an essay, however it remains an open challenge to be able to highlight missing qualities and provide insights on content to add. Even experienced human raters have considerable difficulty pinpointing specific and concrete weaknesses in free-form essays (Roscoe et al., 2018). The features targeted by current writing tutors are highly indicative of writing quality, serving as organizational features that relate to an abstract notion of “cohesion,” which has been observed to decrease with writing quality (Uccelli et al., 2013; Woods et al., 2017). However, identifying these features is not the same thing as analyzing the organization or architecture of ideas within an essay, for example, identifying what is the core idea and how the other parts serve to offer background, justification, or critiques.

Meanwhile, researchers have leveraged various discourse parsing techniques to generate automatic structural writing feedback that takes steps toward this more architectural notion. For instance, Cotos (2011) argued that automatic identification of steps and moves as discourse elements was essential for providing feedback for highly structured and generally well-written essays like multi-section academic papers. In another study, Burstein et al. (2001) showed evidence that a framework referred to as Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) could be used to detect discourse components such as thesis statement and conclusion, even with a relatively small corpus of human-annotated essays. They suggested that AWE systems should strive for providing specific structural writing feedback for free-form essays.

To add to this discussion, the author team of this article conducted a semi-structured interview study with teachers in order to build a vision toward next steps for the field, specifically in design for feedback on structure and flow of writing. In this interview study, the interviewer played the role of a student who wrote an essay, and the teachers walked the interviewer through the feedback that they would give to help the student improve the organization of the essay. From this we learned from teachers’ expressed perspectives that common organizational concerns in essays included: disconnected introduction (i.e., sentences in the introduction or first paragraph are not logically connected to present a clear thesis statement), body sections (i.e., paragraphs in the body do not connect to the main idea, and do not flow naturally from each to the next), and conclusion (i.e., final paragraph does not connect back to the ideas presented in the essay). These organizational concerns have not been fully addressed in the field of Automated Writing Evaluation (Wilson et al., 2017).

In addition, the process of giving feedback to address those organizational issues in the classroom usually involves helping students to reflect on the connections between sentences and paragraphs. In particular, teachers first identified weak spots, and then guided students to address missing or unproductive connections by iteratively reflecting on connections at the word, sentence, and paragraph level. For instance, after reading a student’s essay about positive values of sports, the teacher pointed out that “his major weakness is in his introduction, for the reader, by looking only at the introduction is clueless about the writer’s thesis about sports.” Specifically, the teacher first directed the student (the interviewer) to take a global perspective to summarize the thesis of the essay and then moved to local analysis of three sentences in the introduction, “what additional words and sentences need to be there (between sentences) so that anyone who picks this up would understand how much you love sports so that then that flows into there are a lot of positive values to playing sports (emphasizing the thesis of the essay).” Afterward, the teacher led the student to evaluate word choices in one of the three sentences, “I went back and glanced at your introduction again, and one thing that really threw me in your introduction was, when you were working on a hook, the word however.”

The iterative process in reflecting on idea organization across different levels indicates that we might need to present both a local and global view of writing structure for students to take meaningful actions in revision. One way of looking at the process we heard teachers discuss is that they first identify islands of cohesion within student writing and then seek to help students integrate those islands into a better structured whole. This might be accomplished through an automated process of segmentation, where the writing is deemed well-structured within segments, but not across segments.

To prompt continued work, we hypothesize that RST, a framework for connecting texts at multiple levels (e.g., sentences and paragraphs), might provide a standard for segmentation, which might then be automated, and used to provide automated feedback to improve structure and flow. RST is a popular approach to analysis of the structure of writing that leverages principles of the dependency-based hierarchical nature of text and is common across genres (Mann and Thompson, 1988). In RST, relations between segments of text called Elementary Discourse Units (EDUs) are represented as a tree, an example of which can be seen in Fig. 1. Unfortunately, though RST is one of the most popular theoretical foundations for analysis of hierarchical discourse structure, and though automated analysis of RST is a benchmark task in the field of Language Technologies, availability of labeled data for training automated models remains limited (Jiang et al., 2019). In particular, datasets for RST are time consuming to annotate and require a high degree of expertise to achieve reliability and this is especially true for student writing which often has unusual, or is lacking, structure (Jiang et al., 2019). Available public corpora of data with RST annotations are small relative to the magnitude of annotated data available for training syntactic parsers. Considering both the dearth of annotated data and the challenges of decision making for discourse relations based on local context, it is not surprising that RST parsing has remained a challenging task that has had only incremental improvements even since the deep learning revolution (Li et al., 2014, 2016; Ji and Eisenstein, 2014; Braud et al., 2017; Yu et al., 2018; Mabona et al., 2019). Improvement of automated RST parsing is a direction for future research in this important area. Future work in this area of computational linguistics would potentially have a transformative effect on this area of research.

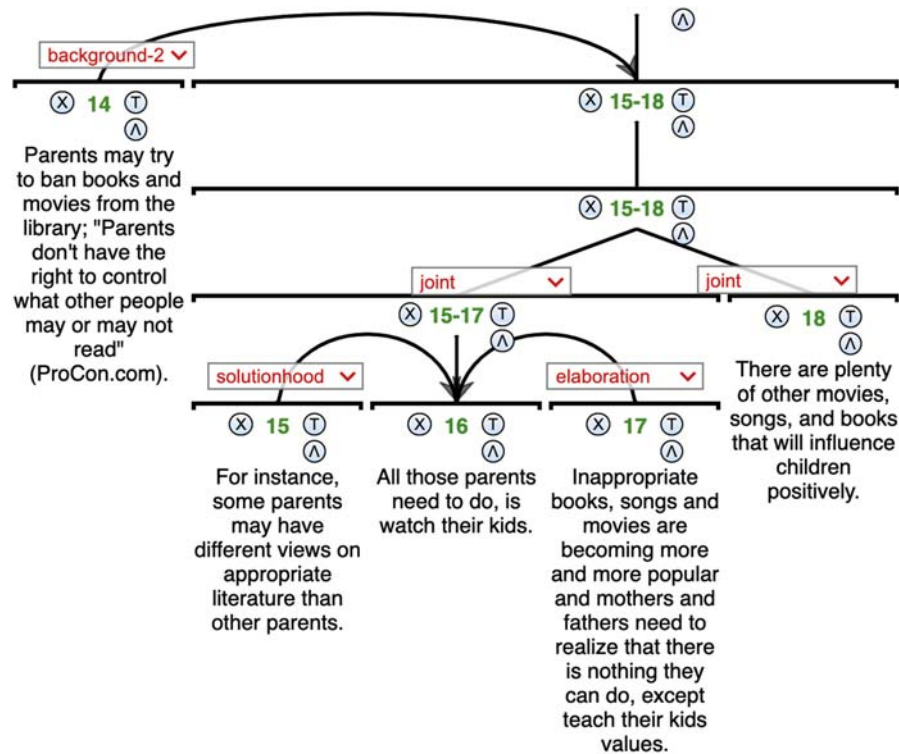


Fig. 1 Example RST tree of a fragment of student writing.

Conclusions and vision for the field

As we reflect on the field of Learning Analytics in light of our experiences in the pandemic, we have considered the connection between data and the human experience at a grand scale. We first considered emerging areas in the field that show promise for achievement of operation at scale. Recognizing where these technologies and this learning are contextualized within human experience, and in particular, human-human relationships and interconnectivity, we considered in depth the area of learning analytics applied to human language. Expanding out from this more specific focus, we are faced with the reality that, even in light of great advances in the field, there was a breakdown in bridging from research into practice that might have buffered some of the learning loss and disruption experienced in the pandemic.

A recent public engagement workshop bringing together the worlds of research, policy, and practice surfaced the need to consider more deeply the social and political context of the technology, taking care to provide scrutability, explainability, and mitigation of bias and unfairness alongside needed teacher training and support (Rosé, 2021). Some notable work provides a foundation on which to expand in the coming years. For example, though the problems are far from solved, a limited amount of research has carefully considered ethical issues related to bias and fairness specifically in the area of learning analytics (Gardner et al., 2019; Ferguson et al., 2016). The transition of research into practice is explored as a formal research endeavor on its own in work illuminating the issues surrounding adoption (Dawson et al., 2018; Macfadyen et al., 2014; Macfadyen and Dawson, 2012).

Though the focus of the field at large is on isolated studies and investigation of data sets and models, some notable work focuses specifically on the role of the teacher (Van Leeuwen, 2019; Van Leeuwen et al., 2019) or delves deeply into classroom culture, values, and orchestration (Chen and Zhu, 2019; Martinez-Maldonado, 2019; Tissenbaum and Slotta, 2019), and eventually views learning processes from the perspective of improvement of school systems (Krumm et al., 2018) especially regarding early detection and mitigation (Waddington et al., 2016). This bridging from the data sets and models we know how to work with, to the more unwieldy, larger, and largely untamed world of teaching in practice, is what we need going forward, especially as learning processes touch those aspects of learning most critical to human existence, learning, and sustainability.

Acknowledgments

This work was funded in part by NSF Grants IIS 1822831 and IIS 1546393.

References

- Adamson, D., Dyke, G., Jang, H., Rosé, C.P., 2014. Towards an agile approach to adapting dynamic collaboration support to student needs. *Int. J. Artif. Intell. Educ.* 24 (1), 92–124.
- Ai, H., Kumar, R., Nguyen, D., Nagasunder, A., Rosé, C.P., 2010. Exploring the effectiveness of social capabilities and goal alignment in computer supported collaborative learning. In: *International Conference on Intelligent Tutoring Systems*. Springer, Berlin, Heidelberg, pp. 134–143.
- Baker, R.S., Inventado, P.S., 2014. Educational data mining and learning analytics. In: Larusson, J.A., White, B. (Eds.), *Learning Analytics From Research to Practice: Methods, Tools, and Approaches*. Springer-Verlag, Berlin, pp. 103–119.
- Baker, R.S., Yacef, K., 2009. The state of educational data mining in 2009: a review and future visions. *J. Educ. Data Min.* 1 (1), 3–17.
- Braud, C., Lacroix, O., Søgaard, A., 2017. Cross-lingual and cross-domain discourse segmentation of entire documents. In: *Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 237–243.
- Buckingham Shum, S., de Laat, M., De Liddo, A., Ferguson, R., Kirschner, P., Ravenscroft, A., Sándor, Á., Whitelock, D., 2013. DCLA13: 1st International Workshop on Discourse-Centric Learning Analytics. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, p. 282.
- Buckingham Shum, S., De Liddo, A., Klein, M., 2014. DCLA meet CIDA: collective intelligence deliberation analytics. In: *2nd International Workshop on Discourse-Centric Learning Analytics, LAK14: 4th International Conference on Learning Analytics & Knowledge*.
- Burstein, J., Marcu, D., Andreyev, S., Chodorow, M., 2001. Towards automatic classification of discourse elements in essays. In: *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, pp. 98–105.
- Burstein, J., McCaffrey, D., Beigman Klebanov, B., Ling, G., 2017. Exploring Relationships Between Writing & Broader Outcomes With Automated Writing Evaluation. *Grantee Submission*.
- Chan, C.K., 2013. Collaborative knowledge building. Towards a Knowledge Creation Perspective. Teoksessa: Hmelo-Silver. In: Chinn, C., Chan, C., O'Donnell, A. (Eds.), *The International Handbook of Collaborative Learning*. Routledge, Taylor & Francis Group, New York.
- Chen, B., Zhu, H., 2019. Towards value-sensitive learning analytics design. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, pp. 343–352.
- Cherubini, M., Nüssli, M.A., Dillenbourg, P., 2008. Deixis and gaze in collaborative work at a distance (over a shared map) a computational model to detect misunderstandings. In: *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, pp. 173–180.
- Chinn, C.A., Clark, D.B., 2013. Learning through collaborative argumentation (Vanderbilt University). In: *The International Handbook of Collaborative Learning*. Routledge, pp. 326–344.
- Cotos, E., 2011. Potential of automated writing evaluation feedback. *Calico J.* 28 (2), 420–459.
- Csanadi, A., Eagan, B., Kollar, I., Shaffer, D.W., Fischer, F., 2018. When coding-and-counting is not enough: using Epistemic Network Analysis (ENA) to analyze verbal data in CSCL research. *Int. J. Comput. Support. Collab. Learn.* 13 (4), 419–438.
- Dascalu, M., Trausan-Matu, S., McNamara, D.S., Dessus, P., 2015. ReaderBench: automated evaluation of collaboration based on cohesion and dialogism. *Int. J. Comput. Support. Collab. Learn.* 10 (4), 395–423.
- Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., Gasevic, D., 2018. Rethinking learning analytics adoption through complexity leadership theory. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 236–244.
- de Lisi, R., Golbeck, S., 1999. Implications of Piagetian Theory for Peer Learning. Mahway.
- Dyke, G., Kumar, R., Ai, H., Rose, C., 2012. Challenging assumptions: using sliding window visualizations to reveal time-based irregularities in CSCL processes. In: *Proceedings of the International Conference of the Learning Sciences*. Summer, Sydney, Australia.
- Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiliza, K., Conati, C., Shum, S.B., 2018. Driving data storytelling from learning design. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 131–140.
- Engzell, P., Frey, A., Verhagen, M.D., 2021. Learning loss due to school closures during the COVID-19 pandemic. *Proc. Natl. Acad. Sci. U. S. A.* 118 (17).
- Ferguson, R., Hoel, T., Scheffel, M., Drachler, H., 2016. Guest editorial: ethics and privacy in learning analytics. *J. Learn. Anal.* 3 (1), 5–15. <https://learning-analytics.info/index.php/JLA/article/view/4912>.
- Fiacco, J., Rosé, C., 2018. Towards domain general detection of transactive knowledge building behavior. In: *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pp. 1–11.
- Fiacco, J., Cotos, E., Rosé, C., 2019. Towards enabling feedback on rhetorical structure with neural sequence models. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, pp. 310–319.
- Fiacco, J., Haan, K., Woolley, A., Rosé, C.P., 2021. Taking Transactivity Detection to a New Level. *Proceedings of the First Annual Meeting of the International Society of the Learning Sciences*, pp. 75–82.
- Fu, E.L.F., van Aalst, J., Chan, C.K.K., 2016. Toward a classification of discourse patterns in asynchronous online discussions. *Int. J. Comput. Support. Collab. Learn.* 11 (4), 441–478.
- Gardner, J., Brooks, C., Baker, R., 2019. Evaluating the fairness of predictive student models through slicing analysis. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, pp. 225–234.
- Gweon, G., Jain, M., McDonough, J., Raj, B., Rosé, C.P., 2013. Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *Int. J. Comput. Support. Collab. Learn.* 8 (2), 245–265.
- Hamilton, L.S., Ercikan, K., 2022. COVID-19 and US schools: using data to understand and mitigate inequities in instruction and learning. In: *Primary and Secondary Education During Covid-19*. Springer, Cham, pp. 327–351.
- Holtz, P., Kimmerle, J., Cress, U., 2018. Using big data techniques for measuring productive friction in mass collaboration online environments. *Int. J. Comput. Support. Collab. Learn.* 13 (4), 439–456.
- Hutt, S., Grafsgaard, J.F., D'Mello, S.K., 2019. Time to scale: generalizable affect detection for tens of thousands of students across an entire school year. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.
- Ji, Y., Eisenstein, J., 2014. Representation learning for text-level discourse parsing. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 13–24.
- Jiang, S., Yang, K., Suvama, C., Casula, P., Zhang, M., Rose, C., 2019. Applying Rhetorical Structure Theory to student essays for providing automated writing feedback. In: *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*, pp. 163–168.
- Kent, C., Cukurova, M., 2020. Investigating collaboration as a process with theory-driven learning analytics. *J. Learn. Anal.* 7 (1), 59–71.
- Kirschner, P.A., Sweller, J., Kirschner, F., Zambrano, R.J., 2018. From cognitive load theory to collaborative cognitive load theory. *Int. J. Comput.-Support. Collab. Learn.* 13 (2), 213–233.
- Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Mogarkar, R., Buckingham Shum, S., 2020. AcaWriter: a learning analytics tool for formative feedback on academic writing. *J. Writ. Res.* 12 (1), 141–186.
- Krumm, A., Means, B., Bienkowski, M., 2018. *Learning Analytics Goes to School: A Collaborative Approach to Improving Education*. Routledge.

- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., Liu, J., 2020. Projecting the potential impact of COVID-19 school closures on academic achievement. *Educ. Res.* 49 (8), 549–565. <https://doi.org/10.3102/0013189X20965918>.
- Kumar, R., Rosé, C.P., Wang, Y.C., Joshi, M., Robinson, A., 2007. Tutorial dialogue as adaptive collaborative learning support. *Front. Artif. Intell. Appl.* 158, 383–391.
- Kumar, R., Beuth, J., Rose, C., 2011. Conversational Strategies That Support Idea Generation Productivity in Groups. *Proceedings of AI in Education*.
- Li, J., Li, R., Hovy, E., 2014. Recursive deep models for discourse parsing. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 2061–2069.
- Li, Q., Li, T., Chang, B., 2016. Discourse parsing with attention-based hierarchical neural networks. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 362–371.
- Mabona, A., Rimell, L., Clark, S., Vlachos, A., 2019. Neural generative rhetorical structure parsing. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 2284–2295.
- Macfadyen, L.P., Dawson, S., 2012. Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *J. Educ. Technol. Soc.* 15 (3), 149–163.
- Macfadyen, L.P., Dawson, S., Pardo, A., Gašević, D., 2014. Embracing big data in complex educational systems: the learning analytics imperative and the policy challenge. *Res. Pract. Assess.* 9, 17–28.
- Mann, W.C., Thompson, S.A., 1988. Rhetorical structure theory: toward a functional theory of text organization. *Text* 8 (3), 243–281.
- Martínez-Maldonado, R., Echeverría, V., Fernández Nieto, G., Buckingham Shum, S., 2020. From data to insights: a layered storytelling approach for multimodal learning analytics. In: *Proceedings of the 2020 Chi Conference on Human Factors in Computing Systems*, pp. 1–15.
- Martínez-Maldonado, R., 2019. A handheld classroom dashboard: teachers' perspectives on the use of real-time collaborative learning analytics. *Int. J. Comput. Support. Collab. Learn.* 14 (3), 383–411.
- Matcha, W., Gašević, D., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., Pérez-Sanagustín, M., 2019. Detection of learning strategies: a comparison of process, sequence and network analytic approaches. In: *European Conference on Technology Enhanced Learning*. Springer, Cham, pp. 525–540.
- McGarrell, H., Verbeem, J., 2007. Motivating revision of drafts through formative feedback. *ELT J.* 61 (3), 228–236.
- McLaren, B.M., Scheuer, O., De Laat, M., Hever, R., De Groot, R., Rosé, C.P., 2007. Using machine learning techniques to analyze and support mediation of student e-discussions. *Front. Artif. Intell. Appl.* 158, 331–340.
- McNamara, D.S., Louwerse, M.M., McCarthy, P.M., Graesser, A.C., 2010. Coh-Metrix: capturing linguistic features of cohesion. *Discourse Process.* 47 (4), 292–330.
- Mu, J., Stegmann, K., Mayfield, E., Rosé, C., Fischer, F., 2012. The ACODEA framework: developing segmentation and classification schemes for fully automatic analysis of online discussions. *Int. J. Comput. Support. Collab. Learn.* 7 (2), 285–305.
- Palermo, C., Wilson, J., 2020. Implementing automated writing evaluation in different instructional contexts: a mixed-methods study. *J. Writ. Res.* 12 (1).
- Roscoe, R.D., McNamara, D.S., 2013. Writing Pal: feasibility of an intelligent writing strategy tutor in the high school classroom. *J. Educ. Psychol.* 105 (4), 1010.
- Roscoe, R.D., Varner, L.K., Crossley, S.A., McNamara, D.S., 2013. Developing pedagogically-guided algorithms for intelligent writing feedback. *Int. J. Learn. Technol.* 8 (4), 362–381.
- Roscoe, R.D., Allen, L.K., Johnson, A.C., McNamara, D.S., 2018. Automated writing instruction and feedback: instructional mode, attitudes, and revising. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62. SAGE Publications, Sage CA: Los Angeles, CA, pp. 2089–2093. No. 1.
- Rosé, C., Wang, Y.C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F., 2008. Analyzing collaborative learning processes automatically: exploiting the advances of computational linguistics in computer-supported collaborative learning. *Int. J. Comput. Support. Collab. Learn.* 3 (3), 237–271.
- Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R., 2019. Explanatory learner models: why machine learning (alone) is not the answer. *Br. J. Educ. Technol.* 50 (6), 2943–2958.
- Rosé, C.P., 2021. Enrichment: insights towards AI impact in education through a mycelial partnership between research, policy, and practice. *AI Matters* 7 (2), 10–11.
- Saint, J., Gašević, D., Matcha, W., Uzir, N.A.A., Pardo, A., 2020. Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. *Proceedings of the tenth international conference on learning analytics & knowledge* 402–411.
- Sagr, M., Fors, U., Tedre, M., Nouri, J., 2018. How social network analysis can be used to monitor online collaborative learning and guide an informed intervention. *PLoS One* 13 (3), 194–207.
- Sagr, M., Viberg, O., Vartiainen, H., 2020. Capturing the participation and social dimensions of computer-supported collaborative learning through social network analysis: which method and measures matter? *Int. J. Comput. Support. Collab. Learn.* 15 (2), 227–248.
- Schneider, B., Dich, Y., Radu, I., 2020. Unpacking the relationship between existing and new measures of physiological synchrony and collaborative learning: a mixed methods study. *Int. J. Comput. Support. Collab. Learn.* 15 (1), 89–113.
- Siemens, G., Baker, R.S.D., 2012. Learning analytics and educational data mining: towards communication and collaboration. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 252–254.
- Siemens, G., 2013. Learning analytics: the emergence of a discipline. *Am. Behav. Sci.* 57 (10), 1380–1400.
- Spikol, D., Ruffaldi, E., Dabisias, G., Cukurova, M., 2018. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *J. Comput. Assist. Learn.* 34 (4), 366–377.
- Su, Y., Li, Y., Hu, H., Rosé, C.P., 2018. Exploring college English language learners' self and social regulation of learning during wiki-supported collaborative reading activities. *Int. J. Comput. Support. Collab. Learn.* 13 (1), 35–60.
- Tissenbaum, M., Slotta, J., 2019. Supporting classroom orchestration with real-time feedback: a role for teacher dashboards and real-time agents. *Int. J. Comput. Support. Collab. Learn.* 14 (3), 325–351.
- Uccelli, P., Dobbs, C.L., Scott, J., 2013. Mastering academic language: organization and stance in the persuasive writing of high school students. *Writ. Commun.* 30 (1), 36–62.
- van Aalst, J., 2009. Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *Int. J. Comput. Support. Collab. Learn.* 4 (3), 259–287.
- van Leeuwen, A., Rummel, N., Van Gog, T., 2019. What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *Int. J. Comput. Support. Collab. Learn.* 14 (3), 261–289.
- van Leeuwen, A., 2019. Teachers' perceptions of the usability of learning analytics reports in a flipped university course: when and how does information become actionable knowledge? *Educ. Technol. Res. Dev.* 67 (5), 1043–1064.
- Vrzakova, H., Amon, M.J., Stewart, A., Duran, N.D., D'Mello, S.K., 2020. Focused or stuck together: multimodal patterns reveal triads' performance in collaborative problem solving. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 295–304.
- Waddington, R.J., Nam, S., Lonn, S., Teasley, S.D., 2016. Improving early warning systems with categorized course resource usage. *J. Learn. Anal.* 3 (3), 263–290. <https://doi.org/10.18608/jla.2016.33.13>.
- Wen, M., Maki, K., Wang, X., Dow, S.P., Herbsleb, J., Rose, C., 2016. Transactivity as a Predictor of Future Collaborative Knowledge Integration in Team-Based Learning in Online Courses. *International Educational Data Mining Society*.
- West-Smith, P., Butler, S., Mayfield, E., 2018. Trustworthy automated essay scoring without explicit construct validity. In: *2018 AAAI Spring Symposium Series*.

- Wilson, J., Roscoe, R., Ahmed, Y., 2017. Automated formative writing assessment using a levels of language framework. *Assess. Writ.* 34, 16–36.
- Wise, A.F., Shaffer, D.W., 2015. Why theory matters more than ever in the age of big data. *J. Learn. Anal.* 2 (2), 5–13. <https://doi.org/10.18608/jla.2015.22.2>.
- Woods, B., Adamson, D., Miel, S., Mayfield, E., 2017. Formative essay feedback using predictive scoring models. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2071–2080.
- Yu, N., Zhang, M., Fu, G., 2018. Transition-based neural RST parsing with implicit syntax features. In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 559–570.

Machine learning

Valerie J. Shute^a, G. Curt Fulwider^a, Zhichun Liu^b, and Seyedahmad Rahimi^a, ^aFlorida State University, Tallahassee, FL, United States; and ^bUniversity of Massachusetts, Dartmouth, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	83
What is machine learning?	84
Approaches to training	84
Supervised learning	84
Unsupervised learning	85
Reinforcement learning	85
A brief history of machine learning	85
Machine learning in education	86
Instructor aids	86
Student aids	87
Future of ML in education	88
Concerns moving forward	89
Conclusion	90
Acknowledgment	90
References	90

Introduction

Consider the following examples. Beverly, while starting her car, raises her wrist to her mouth and requests directions to her next appointment. Her request to the digital assistant is not understood at first, but it is after her second attempt. Since she first began using her smart watch, she has noticed a decrease in the times she must repeat herself. The watch is learning her voice and speech patterns. A few weeks later, she raises her wrist to find that her device is already offering to navigate to that location. It has recognized that every Thursday at 4:00 p.m. she travels to the same place. Beverly accepts the directions and heads down the road.

Across town, Jamal (a college sophomore) is figuring out which textbooks to purchase for his coming semester. Opening Amazon, he sees various college-related supplies being recommended; most of which are already monogrammed with his school's mascot. He selects his first textbook and notices that the site is suggesting another book commonly bought with the one he has selected. Checking his list again, he sees that the recommended book is also required. Both go into his shopping cart. After closing the browser, he sees a notification from his favorite social media platform. A "friend" has uploaded a somewhat embarrassing photo of them from the prior evening. The app has recognized Jamal's face, and is asking if he would like to be tagged in the photo. He declines (while also making a mental note to talk with his friend about posting photos online without permission).

These examples all fall under the umbrella of Artificial Intelligence (AI), which increasingly appears in the news because of the breadth of its reach. Part of this breadth comes from the generality of its definition. John McCarthy—who first coined the term in 1955—defined artificial intelligence as, "the science and engineering of making intelligent machines, especially intelligent computer programs" (Stanford University, n.d.). The purpose of AI is to model human thinking, and eventually surpass it. Because it is such a broad term, AI has found application in nearly all areas that make use of technology and data. The specific area within AI where the opening examples fall is called machine learning (ML).

ML has become a popular buzzword in recent years and is often incorrectly used interchangeably with AI. Rather, ML refers to the process by which a computer attempts to find regularities or patterns from a set of data—like speech patterns, purchasing habits, and facial features described in the examples above. Accurate recognition of patterns requires the algorithm to be exposed to large sets of examples. In this way, the algorithm "learns" to identify patterns that it can then recognize in novel data. Even then, the algorithm may, if provided the capacity to, improve itself by learning from feedback loops that inform it of cases which were mistakenly identified.

In the opening illustration, with Beverly's smart watch not recognizing her voice, the ML algorithm embedded within the device might take an initial sample of her speech patterns, but whenever Beverly must repeat herself, the algorithm identifies that as a mistaken interpretation and improves itself for future interactions. The watch may be artificially intelligent, but it is machine learning that allows the watch to improve over time. Another example of ML in use today is how Uber uses ML algorithms to predict how long it will take rides (or meals) to arrive, set pricing, and even determine where best to wait to give a driver the best odds of finding you for pick-up (GeekWire, 2016). These algorithms are specific to their purpose and likely trained using data they mined from their own apps. Finally, it is the ML technology behind driverless cars that can quickly adjust to diverse conditions while driving. The continual self-improvement of the algorithm, however, is what makes it truly "intelligent." ML is not the same as AI, but ML does provide AI with the ability to learn and improve.

What is machine learning?

Think of Pavlov's dogs. When he would ring a bell, they started to salivate. Pavlov taught them this pattern: every time the bell rings, food is on the way. The dogs soon learned the pattern and eventually began to respond to the bell *before* the food appeared. To cognitive psychologists, human brains are also natural pattern matchers. Since birth, we are constantly trying to find patterns in what we perceive, then associate new patterns with existing patterns in our memory. This pattern-matching process happens automatically. Through our everyday experiences, our brains are constantly collecting new patterns and adding them to memory. These patterns then influence the inferences and decisions we make in the future.

Inspired by an understanding of how the human brain works, researchers in the field of ML draw on connections to psychology, neuroscience, and the learning sciences. These connections are evident in the use of shared or similar terminology (e.g. reinforcement learning). Tom [Mitchell \(1997\)](#) wrote a widely quoted definition of ML: "A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E " (p. 2). Machine learning focuses on developing algorithms that allow machines to improve their performance on a certain task over time through accumulating experiences. Going back to Beverly, her watch implements an ML algorithm to understand her voiced instructions. Although the initial performance is not fully satisfying (i.e. she had to repeat herself several times), the watch becomes better over time as it experiences more instances of Beverly's voice. Modern companies that build these voice-recognition systems often pre-train the algorithm with massive amounts of voice recordings covering a wide range of accents, volumes, and acoustic environments prior to selling them to the customers. The initial accuracy of the watch's voice recognition should be acceptable, but it is further improved by specifically learning its owner's voice.

Approaches to training

Although all ML approaches share the goal of automatically improving performance on a task based on accumulated experiences, they differ largely on the target task, the nature and availability of training data, and the measure of performance. Given the complexity and variability of typical problems addressed with ML, there is no one-size-fits-all approach to training. Although not a comprehensive list, the three largest categories are: supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning

In supervised learning, the ML algorithms are trained using large collections of examples that are *labeled* with the correct "answer." Similar to how one studies for a test by completing endless practice problems, the algorithm "learns" how properly to identify or classify a given input by processing large, labeled datasets until it can do so with an elevated level of confidence—the larger the training set with rich and diverse examples, the more accurate the algorithm.

A simple example of supervised learning involves training an algorithm to differentiate between images of cats and dogs. The training dataset consists of a huge selection of different images showing either a cat or dog, along with its correct label. The algorithm "studies" these examples by identifying patterns within the pixels. Then, when provided with a novel (unlabeled) image, the algorithm will look for similar patterns and provide an answer—although "best guess" might be more appropriate—and a level of confidence. Now, what happens when the algorithm, trained on nothing but cats and dogs, is given an image of a rabbit? There are certainly similarities that could be drawn between rabbits, cats, and dogs, but the algorithm can only choose between the latter two animals. It will have to choose whichever category the rabbit image most closely resembles, but the confidence of the result will likely be low.

The training data are critical to a successful supervised learning approach, but reliance on the training data is also a limitation. If high-quality training datasets (i.e. containing accurate human coders' labels with unbiased and diverse training examples) are not available, it can take hundreds of hours for humans to collect and label a new training set. Moreover, if humans are doing the labeling, they must be very careful as each error will influence a dataset's quality and the overall performance of any ensuing predictive model.

Supervised learning primarily deals with problems of prediction and forecasting. Supervised learning algorithms look at existing data to make an inference about new instances or predict the outcome of a yet-to-be-seen example. And while identifying the species of pet photos on the Internet may be fun, supervised learning can be applied to data beyond images, such as using numerical data for predicting and forecasting. For example, stock prices or real estate trends, due to the potential financial gains and the ample amounts of historical data, are popular real-world examples of supervised learning. The prediction of a single house price, for instance, might use data that include several properties of individual houses in the dataset, called features. Each house might include features describing earlier sales prices, the neighborhood, square footage, number of bedrooms, and so forth. Another algorithm may be trained on patterns leading up to major shifts in the market to predict major trends in the future. Costly errors can be made when an algorithm, meant to predict stock prices, is trained on data that neglect the potential of dramatic swings in the market due to unique events (e.g. a natural disaster or war). Thus, supervised learning algorithms can have powerful predictive and discriminatory capabilities—provided the availability of enough quality training data that is also sufficiently representative of the targeted population.

Unsupervised learning

When well-labeled data is unavailable and the task of creating a labeled dataset is neither physically nor financially feasible, or when one is without preconceived categories (or unsure what to even look for), unsupervised learning may be the ideal approach. The unsupervised approach allows the algorithm to detect patterns that may be too complex for a human to recognize. The algorithm reads the unlabeled data and discovers hidden patterns within.

For instance, the construction of a commercial airliner involves millions of mechanical parts. Each part of the aircraft will have a degree of error in length, weight, or performance due to simple variation in quality control—all pieces meeting a minimum standard, but not perfect. Even though all parts individually may work without perceivable error, in a complex machine like an aircraft, the accumulated error may be catastrophic. Traditionally, technicians must manually conduct multiple systemic evaluations of the plane, which cost tremendous amounts of time and money. Using an unsupervised approach, such as Anomaly Detection, and data regarding the quality of each part, an algorithm can holistically assess the amount of accumulated error and determine to what extent the accumulated errors affect the quality of the plane's airworthiness before the plane even leaves the warehouse.

Reinforcement learning

Both supervised and unsupervised learning are powerful approaches, but sometimes real-world problems that are natural to humans can be too complex or dynamic for machines to label or model (e.g. driving a car on a highway). Compared to supervised learning, where the answers are provided upfront, reinforcement learning tackles a task with parameters for successful completion. One might draw a simplistic analogy comparing supervised learning to rote memorization and reinforcement learning to discovery learning. The algorithm works by making random attempts, each of which is provided immediate feedback, until it can successfully complete the task.

Self-driving cars are a well-known example of reinforcement learning. When training an algorithm to drive a car, one central task is to determine the optimal combination and sequence of gas pedal, brake, and angle of the steering wheel. Instead of labeling the behaviors of 1000 different drivers, reinforcement learning can happen in a *simulated* environment with a wide variety of driving situations.

The driving algorithm is given a straightforward goal: drive from point A to point B without accident and before time runs out. Any behaviors that progress toward meeting the goal are rewarded while the rest are penalized. The algorithm starts with random actions and, for a while, might not even move forward. However, after thousands of random trials, the algorithm begins to use the gas pedal to move forward (not backwards) because moving forward closes the gap between point A and point B, but it may not have learned yet how to use the steering wheel to turn. After running off the road, hitting stationary objects, making U-turns, and another million failed trials, the algorithm finally begins to guide the car to its destination. The cost of reinforcement learning is computational resources and time in simulated environments, which is relatively inexpensive and more efficient compared to crashing cars on a real road or collecting data on even 100 real drivers.

Supervised learning, unsupervised learning, and reinforcement learning are currently the most well-known approaches in ML. Other more complex approaches derive from these basic approaches. For example, semi-supervised learning and self-supervised learning combine both supervised and unsupervised learning techniques, such as using the results from unsupervised learning to inform labeling in supervised learning datasets. Additional ML approaches (e.g. ensemble learning, active learning, transfer learning) strive strategically to improve the performance of the basic approaches. These techniques do not have to be supervised, unsupervised or reinforcement learning, but they do employ strategies to improve the performance of any single model. For instance, ensemble learning combines multiple classifiers and then outputs more robust results compared to simpler approaches. Active learning can involve a human coder to resolve uncertain cases. And transfer learning, in some instances, can use pre-tuned model parameters to shorten the training time or deal with problems with fewer training examples needed.

In summary, ML is an applied field and different approaches are inspired by diverse practical problems across various disciplines. Therefore, these approaches do not necessarily all share the same theoretical foundation and historical root. Each approach to ML has developed a multitude of different algorithms for different purposes. This diversity began during the mid-20th Century and continues to evolve today.

A brief history of machine learning

As early as 1947, the British mathematician Alan Turing envisioned a machine that could, based on an initial set of instructions, modify its own operating instructions over time. The new instructions created by the machine were the result of responses to recently added information. He likened the process to students who have learned basic information from their teachers and went on to develop new knowledge on their own.

In such a case one would have to admit that the progress of the machine had not been foreseen when its original instructions were put in. It would be like a pupil who had learnt much from his master but had added much more by his own work.

Turing (1995, p. 13).

Turing's idea was too theoretical and abstract to implement in 1947, but today's development of ML is largely inspired by his vision. With modern technology, Turing's ideas have come to fruition.

One of the earliest practical examples of learning machines is often attributed to Arthur Samuel. In the 1950s, Samuel developed a computer program that could learn to play the game Checkers, a strategic board game where two players make diagonal moves and capture pieces by jumping over the opponent. Samuel's program evaluated each position by looking ahead at all the possible moves to determine the best next move. Later, he developed a simple learning method (i.e. rote learning) where, through repetitive self-play, the program memorized optimal moves given a set of conditions observed over thousands of games. Although the algorithms he used were relatively slow and limited by the computer memory at that time, the program was able to improve its performance continuously and reached the level of "better-than-average novice."

Around the same time, researchers came up with the idea for artificial neural networks (ANN) that imitate human neurons to deal with complex computing often involved in machine learning. This idea was initially discussed in the early days of machine learning (e.g. Perceptron invented by Frank Rosenblatt at Cornell Aeronautical Laboratory in 1957). In the 1970s and 1980s, classic structures of ANN were invented (e.g. Convolutional Neural Networks for image recognition and Long Short-Term Memory Recurrent Neural Networks for natural language processing). However, these efforts did not make a substantial impact outside of research labs due to the lack of computational power and limited data at the time. ANN has regained attention in recent years in the form of deep learning (i.e. multi-layered neural networks). Modern deep learning applications range from classic speech recognition and image recognition to complicated music composition, bioinformatics, drug discovery, and so on. Many applications reach the level of "experienced amateur" and some even surpass human performance. To illustrate, from 2015 to 2017, AlphaGo—a neural-network-based AI system—defeated every world champion in a strategic board game called Go (like chess). Originating in ancient China, Go is a vastly more complicated game with 361 potential opening moves compared to Checkers (four moves) or Chess (twenty moves). Brute-force assessment of every possible next move in Go, as is done in simpler algorithms used for Checkers and Chess, is beyond the computational power of any computer in existence. However, with the successful implementation of ML algorithms (e.g. Monte Carlo tree search and deep neural network), AlphaGo can play against itself just as Arthur Samuel envisioned, evolving over time in search of the optimal strategy.

Besides rote learning and artificial neural networks, ML has been influenced by many different disciplines and the field of ML has branched out to multiple lines of research through its history. For example, Email systems now routinely have spam filters to block junk emails. These filters are usually powered by algorithms like Naïve Bayes classifiers, based on Bayesian probabilistic inferences, invented in the 1970s. Later, around the 1990s, models like Support-vector machines (SVM; Cortes and Vapnik, 1995) were introduced to the ML field. They can be used for both regression and classification tasks and are typically used for classification purposes. SVMs provide an alternative, non-probabilistic method for classification compared to Naïve Bayes. Inspired by evolutionary theory, genetic algorithms were also invented around the same time as SVMs (Holland, 1992). Genetic algorithms were once widely used to solve complex and dynamic problems like scheduling optimization for flights and trains and later informed the development of reinforcement learning. The theoretical foundations and data structure of different algorithms may vary, but they all share a similar goal—identifying an optimal solution based on existing and incoming new data. Now, having laid a foundation about ML in general, we next describe how ML can be used in education.

Machine learning in education

At its core, the purpose of ML is to solve problems and offer insight into issues too complex for human analysis, or to automate routines too costly for human labor. Successful implementation of ML algorithms reduces the reliance on human effort by automating complicated or tedious tasks. In education, there are many such tasks that could benefit from ML. For instance, a report published by the Organization for Economic Cooperation and Development (OECD) in 2015 found that teachers spend, on average, about half as much time teaching as they do on non-teaching tasks including lesson planning, grading, administrative work, and collaboration with other teachers. Moreover, teachers struggle to provide equal attention to all students in addition to preparing high quality instructional activities to classrooms of 30+ students. Struggling students are often left alone to grapple with difficult subjects before the class moves on to the next topic. Such unaddressed gaps in learning may plague students for years.

The opportunities for ML to benefit the field of education are bountiful; however, change in education is often slow. Developing and implementing ML-powered systems requires significant financial investment, expertise in both education and machine learning techniques, and extra caution when dealing with vulnerable populations. For these reasons, intelligent machine aids are only slowly making their way into schools worldwide. We now describe some current ML applications that are helping both teachers and students.

Instructor aids

Teachers stand to benefit greatly from ML systems, such as in time saved from completing routine tasks that are easily automated. Such ML-powered systems can be employed to save teachers' time by grading homework assignments and exams that are in essay format. For example, the Educational Testing Service has been using an ML-based system called *c-rater* for years to score participants' essays (Sukkarieh and Blackmore, 2009).

Tools to enable facial recognition have already been shown to be helpful in simple daily tasks like taking attendance (e.g. [Okok-pujie et al., 2017](#); [Sanli and Ilgen, 2018](#)). Although checking which student is present may take between two to 10 min, depending on the class size, even just 2 min can add up to a total of 6 h across a 180-day school year. These systems, which have been tested in limited classrooms already, record the classroom at certain times throughout the day. Student faces are captured, and the facial recognition software identifies the student and marks them as present.

Attendance may seem trivial, but facial recognition algorithms can also help with the arduous task of noticing students' subtle behaviors during class. ML systems have demonstrated the ability to track facial expressions that can indicate students' engagement. Systems like these can notify teachers when students are not engaged, which in turn could prompt the teacher to change up a lesson, call for a break, or help a struggling student. For example, [Sümer et al. \(2021\)](#) used ML to analyze the audiovisual recordings from a classroom of students over a period of one and half months. They trained their engagement classifier models using two features of engagement (attention and affect). The results were promising. That is, their engagement classifier achieved an acceptable level of accuracy, which is notable because this study was the first in this area (i.e. multimodal engagement analysis in a physical classroom) and because of the relatively small amount of student-specific data used. Results from this type of research can help teachers identify whether their students are engaged or not so they can modify their teaching (e.g. slow down and elaborate, ask a thought-provoking question), as needed. Granted, facial recognition in classrooms raises issues of privacy and potential abuse. We address those concerns later.

Identifying real-time indicators of engagement is one benefit of ML, but long-term data collection can yield even more valuable insights. For example, researchers have used multiple ML algorithms to predict high-school students' risk of dropping-out based on records of coursework, attendance, and performance ([Coleman et al., 2019](#)). By recognizing patterns that lead to dropping out of courses, teachers can improve retention rates by offering interventions before the student has even seriously considered quitting, such as suggesting a student come for extra help after school prior to an important test. Moreover, ML has been used to model complicated socio-cognitive processes such as knowledge sharing and negotiation in computed-supported collaborative problem-solving contexts ([Stewart et al., 2019](#)). In general, receiving real-time data—especially that which has been accumulated over time—can provide recommendations so that teachers are able to make more informed decisions on subsequent learning activities. But ML can also benefit students directly.

Student aids

In addition to supporting teachers, ML has incredible potential for enhancing students' learning experiences. One early vision for implementing ML was to develop intelligent tutoring systems which were meant to emulate the role of a human teacher and respond to students thoughtfully as an expert tutor ([Self, 1990](#); [Shute and Zapata-Rivera, 2010](#)). In this case, ML can be used to (a) identify students' learning processes and behaviors, (b) diagnostically assess their weaknesses and strengths, and (c) provide individualized cognitive and affective supports as needed. As part of interacting within such digital learning environments, students demonstrate observable behaviors relative to targeted knowledge and skills. These behaviors could then be collected as information traces and used to train ML models, for example, identifying behavioral patterns that are successful and those that are not for various outcome measures. The models can then infer students' evolving knowledge and skill levels based on an accumulation of observed behaviors. Using these inferences, the intelligent learning environment can recommend the most appropriate learning supports or subsequent content to optimize students' learning experience. Diagnostic information and intervention recommendations can also be shared with the teacher.

Recently, the National Science Foundation granted \$20 M to an AI-institution project led by the University of Colorado, Boulder ([Strain, 2020](#)). This five-year project is targeted at middle- and high-school students and aims to improve teamwork. In the classroom envisioned by the project, ML algorithms will be used to recognize who is speaking, what is being said, and then react in real-time (e.g. provide feedback) as needed. The data collected over time will be used to create individual models of each student that can be used for formative and summative assessments. ML and similar techniques comprise the building blocks of such complex AI-based classrooms.

The costs of designing, developing, and testing intelligent systems are high, thus making them cost-prohibitive for most schools. However, many of the functions of such systems can be emulated, with the help of ML, by more familiar environments, like educational games. For instance, *Physics Playground* ([Shute et al., 2019](#)) is an educational game that incorporates stealth assessment ([Shute, 2011](#)) to estimate students' conceptual physics understanding in real-time. Stealth assessment is embedded into the game code to continually track a player's progress and support learning during gameplay using an automated statistical scoring and accumulation approach (Bayesian inference networks). Moreover, the game includes a quit-prediction model. This model can predict if a student is frustrated and is likely to quit playing. The game then can use this information (i.e. the probability of quitting for each student at any given time) to provide appropriate supports (e.g. affective supports such as motivational messages) to encourage persistence in solving game levels. The quit-prediction model uses 37 variables, also known as features, to determine the probability of a student quitting. These features were extracted from student-generated log files through an iterative ML process called feature engineering ([Shute et al., 2020](#); [Slater et al., 2017a,b](#)).

Thus far, we have looked at how a machine can assess and support student learning, both of which are dependent on ML approaches that focus on pattern recognition. Reinforcement learning provides another benefit that goes beyond monitoring and reporting on students' knowledge and skill levels. That is, reinforcement learning algorithms can provide individualized content and experiences that surpass what a single teacher could achieve. The ITSs mentioned earlier can reach a new level of sophistication

when paired with reinforcement learning. That is, earlier ITSs were dependent on rule-based, manually coded, and generalized responses to student interactions. The manual approach is limited by how many rules the designer can think of and integrate into the system. Using reinforcement learning, the ITS can adapt to learners' behaviors and select the optimum next step for learners based on their prior performance (Chi et al., 2011). Every student receives a unique learning experience, based on the insights gathered from the ML-based systems.

Taking ML's ability to generate tailored content in another direction via reinforcement learning to train and build simulated environments, individual students may interact with these environments to experiment and learn. What could only be explained with diagrams and hypotheticals can now be demonstrated in real time (or in accelerated time for slow-moving processes) as the student observes. Economic markets can be generated that mimic modern-day markets (e.g. Zheng et al., 2020). Students can adjust parameters to see how simple changes can have massive impacts. Tiny creatures can be modeled, created by students, and left to live out the perils of natural selection as students observe how they survive (or not) over long periods of time. Not only can students observe these complicated systems that have such powerful influences on our lives, but they can also create their own systems and see how they respond to different inputs. This is putting the world in students' hands and letting discovery learning take place on a whole new scale.

What does the future hold for ML in education? We believe that the influence of ML will continue to grow in educational systems around the world. Following are some ideas and examples of what we expect to see happen with ML and education in the future.

Future of ML in education

Although modern programming tools and programming libraries already make the workflow of structuring ML projects easier than before (Slater et al., 2017a,b), ML is far from accessible to the layperson. We do expect to see that, in the future, ML will become a fairly ubiquitous influence in the educational world by making insights and answers readily available to all stakeholders. For instance, educators will be able to actively use ML tools to support instructional decisions and students can use ML to diagnose their own performance without going through hundreds of hours of training only to get a single score. This accessible ML future first relies on the advancement of technology (e.g. increases in computational power). That is, although the main theoretical foundations of ML were established in the last century, applications in this field were largely restricted by the costs and limits of technology. Recent advances in computing—like quantum computing, Optical Neural Networks, and distributed computing—continue to increase computing performance and lower cost. Before long, computing resources required for more advanced ML applications, rather than being a luxury of well-funded computing labs, will be within the budgets of local school systems—making the benefits of ML more prevalent in education.

In addition to rising power and decreasing costs, this envisioned future depends on raising the bar for “basic” computer literacy. Just as the ability to read was once a luxury of the upper class, and not long ago creating content on the Internet was limited to only the technically savvy, the minimum threshold for what is considered “basic knowledge” of statistics, computer science, and engineering practices (e.g. simulation and modeling) must be raised. A transformation in society's computing literacy will take time, but we are optimistic. Again, just 30 years ago, when the Internet was young, creating a simple website required advanced programming skills. But today, anyone can create sophisticated and beautiful websites without any programming knowledge. As ML technology continues to increase in availability and become more prevalent in our everyday lives, we fully expect ML to become a basic programming skill and, one day, be a part of our children's core science curricula.

As more and more educational activities happen online, the data available to ML tools used in education will grow exponentially. The increase of data volume is predictable, and we expect the future will include the fusion of dynamic and diverse data being used for ML in education. ML will not have to focus on one model or source of data. Instead, data will come from many sources and channels. Students' academic performances (numerical data), facial expressions (visual data), classroom discourse with peers and instructors (verbal data), and engagement levels (physiological data) can all be merged to provide a holistic and more valid view of students' performance and understanding. We will be able automatically to identify that a student's difficulty in an English class is not because of language ability but due to test anxiety based on affect detection and other physiological data. In addition, data fusion can synthesize the conclusions from multiple sites (e.g. different school districts), which prevent the sharing of individual levels of identifiable digital traces. Ultimately, we can build up a comprehensive and multi-faceted data warehouse to benchmark educational ML performances and inform research and practices.

In the more distant future, we may see analytical systems, powered by ML, help individuals self-actualize by identifying dispositions and paths needed to achieve short- and long-term goals. Such ML-based systems would make the learning process more efficient, self-directed, individualized, and life-long. Through advanced personalized-learning environments, students could pursue their own interests instead of being constrained by one-size-fits-all curricula. Apart from individual support, ML-powered systems could be employed at the group level, for data-driven decision-making purposes. For instance, governments can use school, district, city, and state data to create policies and practices that can dramatically improve education quality with high precision. In this scenario, the unit of analysis is not a student; rather, it is a larger entity made up of students, teachers, educators, parents, administrators, and other stakeholders. Data from all these sources could be used to develop ML-based models to inform policy makers' decisions. Using ML algorithms, governments can build educational success models, starting with feature engineering (i.e. identifying all the important variables predictive of an outcome).

To work effectively, now and in the future, ML requires copious amounts of data, such as online accounts, learning management systems, and school databases. How to balance the use of all these data while protecting students' privacy is a concern as we move forward. The ethical issues of mass data collection and the overall use of ML and AI are current global concerns, but when dealing with vulnerable populations like young students, extra caution must be taken before ML can fully take off in education. Like any major advancement, the future of machine learning, and its benefits, will face hurdles we must surmount.

Concerns moving forward

As we move toward this new world of “smart” machines, the ethics of ML must remain at the forefront of progress. In general, computer scientists, researchers, educators, and educational policy makers who generate, put in use, and advance the technologies that are powered by ML must be mindful of the controversies to prevent their products from doing harm (Garrett et al., 2020). For instance, meticulous care must be given when both creating and training ML algorithms to ensure they remain unbiased. A misconception about AI and ML algorithms is that they are inherently unbiased because humans are not involved. However, humans have biases and humans are the ones writing the algorithms. Even well-intentioned people may possess implicit biases that unintentionally influence their designs (Bertrand and Mullainathan, 2004). More importantly, ML algorithms are trained on massive datasets that reflect the reality of human bias. Therefore, the algorithms can inherit those biases—e.g. racial or gender bias. Beyond that, human judgment is always involved in what is deemed worthy of being modeled by ML, and what is not. This may seem like a small issue, but even the underrepresentation of a single population can have a large impact during application. Consider the case of facial recognition algorithms powered by ML. Recent studies have shown that the accuracy of facial recognition software recognizing “darker” faces was lower compared to the accuracy of recognizing “lighter” faces (e.g. Buolamwini and Gebru, 2018). The reason for this discrepancy was due to the large number of lighter-skinned faces (i.e. about 85%) existing in the databases used to train the ML algorithm. Earlier, we provided two examples of using ML-powered facial recognition in classrooms (one for taking attendance and the other for building an engagement classifier). Such racial biases can sabotage those applications of ML in classrooms by mistakenly marking students of color as absent, or continually misidentifying a student's engagement due to idiosyncrasies associated with cultural differences.

ML algorithms can also be biased relative to gender. For example, Caliskan et al. (2017) found that natural language processing algorithms can inherit gender biases—either from the data that was used to train the algorithms or from ignoring such biases by the programmers. For instance, the word *nurse* is more frequently associated with the word *she* than with the word *he*. A biased model can perpetuate biases in those who interact with the biased environment; even the results of modern image searches can portray gender biases related to occupations (Kay et al., 2015). Allowing these skewed perspectives into the classroom can impair the struggle to reduce gender inequality.

There are some things we can do to reduce bias in data-driven, automated ML approaches. Ntoutsi et al. (2020) have suggested that we need to: (1) Understand bias relative to how it occurs in society and how it finds its way into the data used for training ML algorithms so that we can proactively address biases when creating algorithms and collecting data; and (2) Mitigate bias in each step of the ML decision making process. For example, we need to ensure that the ML algorithms are adequately trained on data from minority populations (i.e. gender- and race-balanced datasets) so that datasets provide a valid representation of all populations. Moreover, several researchers (e.g. Garrett et al., 2020; Saltz et al., 2019) have noted that it is in everyone's best interest to teach ethics in AI as part of computer science courses to make sure that future computer scientists (and others who will be writing the algorithms) are aware of ethical issues. But above all else, those who implement ML algorithms must retain “transparent classifiers which are interpretable on their own and exhibit predictive accuracy close to that of obscure models” (Ntoutsi et al., 2020, p. 8). Aggressive transparency, much like peer-review, helps establish a norm for public auditing of the algorithms that so intimately influence everyone's lives. And although there are ways to reduce the biases ML inherits from humans, we note that ML is not a panacea to all problems, particularly in educational systems. Thoughtful decisions about what to model (and why) need to be made prior to training ML algorithms.

Security (i.e. protection from hackers) and privacy (i.e. data anonymization) are also issues that need proper attention to ensure a successful future with ML (e.g. Mittelstadt and Floridi, 2016). When personal data on an individual are stored somewhere, people with harmful intentions may access the data (e.g. identity theft). Therefore, it is important to know how the data are stored, who owns the data, and how secure the data are. Regarding privacy, a quick resolution is to remove any identifiable information from the data, but this solution greatly diminishes an algorithm's ability to function as even just one's date of birth and zip code are enough to identify them (Mittelstadt and Floridi, 2016).

As mentioned previously, when facial recognition technology is used in classrooms and on campuses, concerns of privacy and abuse are quick to surface. However, we believe this trend is evident in the literature and the continued integration of ML (e.g. facial recognition) is a double-edged sword, offering great benefits alongside meaningful risks. If quick resolutions are non-solutions, then policies and legislature must ensure the security of students; for example, recommendations tend toward transparency and the ability to “opt-out” of having one's data collected. In short, legislation such as the General Data Protection Regulation passed in Europe and similar legislation passed in California possesses a spirit of security, privacy, and fairness that must extend to some extent to academia.

Finally, because new technologies appearing in the future may be more work and require additional learning for teachers who are already overwhelmed, we need research and development on user-friendly, easy-to-use tools that can attract teachers to use the tools. Moreover, professional education is needed to help teachers understand the strengths and weaknesses of the new tools. Any implementation of ML-powered systems in schools must be accompanied by comprehensive training for teachers to ensure they are fully in command of the power ML can give them, and aware of potential ethical issues in using them.

Conclusion

We first explained what AI and ML mean and how they relate to each other. Then, we focused on machine learning, its definition, history, and use in education, now and in the future. We concluded with a discussion of relevant concerns and issues related to the use of ML, especially in education, moving forward. What makes us confident that the future of education is bright is the amount of progress that has been made in the past several decades in terms of technology, learning sciences, computer science, psychology, psychometrics, and so on. No matter how far we eventually go in using ML in education, we need to remember that nothing can replace human-to-human and particularly teacher-student interactions.

The goal for ML should be to help—not replace—humans in classrooms. The algorithms that improve our daily online experiences can become extensions of the instructor, the administrator, and policymaker. ML can bring to one's fingertips answers to questions that just ten years ago required teams of researchers and decades of experience in the classroom. ML can automate and energize the classroom to the extent that the teacher is free to facilitate learning with a level of precision previously unattainable. Yes, caution must be taken, but with a careful hand, a watchful eye, and forethought informed by history, it is possible that we can press forward, propelled by technology, into a brave new world of education.

Acknowledgment

We would like to thank Ryan Baker and Angela Stewart for reviewing a draft of this paper.

References

- Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Am. Econ. Rev.* 94 (4), 991–1013.
- Buolamwini, J., Gebru, T., 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Conference on Fairness, Accountability, and Transparency*. PMLR, pp. 77–91.
- Caliskan, A., Bryson, J.J., Narayanan, A., 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* 356 (6334), 183–186.
- Chi, M., VanLehn, K., Litman, D., Jordan, P., 2011. Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Model. User Adapted Interact.* 21, 137–180. <https://doi.org/10.1007/s11257-010-9093-1>. Available at.
- Coleman, C., Baker, R.S., Stephenson, S., 2019. Proceedings of the 12th International Conference on Educational Data Mining July 2019, Montreal [online]. In: Lynch, C.F., Mercer, A., Desmarais, M., Nkambou, R. (Eds.), *A Better Cold-Start for Early Prediction of Student at-Risk Status in New School Districts*, pp. 732–737. Available from. <https://eric.ed.gov/?id=ED599170>.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>.
- Garrett, N., Beard, N., Fiesler, C., 2020. More Than “If Time Allows” the Role of Ethics in AI Education. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pp. 272–278.
- GeekWire, (2016). <https://www.geekwire.com/2016/uber-collapse-without-pattern-finding-computers-says-chief-machine-learning/>.
- Holland, J.H., 1992. Genetic algorithms. *Sci. Am.* 267 (1), 66–73.
- Kay, M., Matuszek, C., Munson, S.A., 2015. Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. *CHI'15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems April 2015, Seoul [online]*, pp. 3819–3828.
- Mitchell, T.M., 1997. *Machine Learning*. McGraw Hill.
- Mittelstadt, B.D., Floridi, L., 2016. The ethics of big data: current and foreseeable issues in biomedical contexts. In: *The Ethics of Biomedical Big Data*, pp. 445–480.
- Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.E., et al., 2020. Bias in data-driven artificial intelligence systems—an introductory survey. *Wiley Interdiscip. Rev.* 10 (3), e1356.
- Okokpuije, K., Noma-Osaghae, E., John, S., Grace, K.A., Okokpuije, I., 2017. A Face Recognition Attendance System with GSM Notification, 2017 IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON). IEEE, pp. 239–244.
- Saltz, J., Skirpan, M., Fiesler, C., Gorelick, M., Yeh, T., Heckman, R., et al., 2019. Integrating ethics within machine learning courses. *J. Educ. Resour. Comput.* 19 (4), 1–26.
- Sanli, O., Ilgen, B., 2018. Face Detection and Recognition for Automatic Attendance System. *Proceedings of SAI Intelligent Systems Conference*. Springer, Cham, pp. 237–245.
- Self, J.A., 1990. The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *Int. J. Artif. Intell. Educ.* 10, 350–364.
- Shute, V.J., Zapata-Rivera, D., 2010. Intelligent systems. In: Baker, E., Peterson, P., McGaw, B. (Eds.), *Third Edition of the International Encyclopedia of Education*. Elsevier Publishers, Oxford, UK, pp. 75–80.
- Shute, V., Almond, R., Rahimi, S., 2019. *Physics Playground (1.3) [Computer Software]*. <https://pluto.coe.fsu.edu/ppteam/pp-links/>.
- Shute, V.J., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C.-P., Kamikabeya, R., Liu, Z., Yang, X., Sun, C., 2020. Maximizing learning without sacrificing the fun: stealth assessment, adaptivity, and learning supports in Physics Playground. *J. Comput. Assist. Learn.* 37, 127–141. <https://doi.org/10.1111/jcal.12473>.
- Shute, V.J., 2011. Stealth assessment in computer-based games to support learning. In: Tobias, S., Fletcher, J.D. (Eds.), *Computer Games and Instruction*. Information Age Publishers, pp. 503–524.
- Slater, S., Bowers, A., Kai, S., Shute, V.J., 2017a. A Typology of Players in the Game Physics Playground. *Proceedings of the Digital Games Research Association Conference, Melbourne, Australia*, pp. 1–12. <http://www.digra.org/digital-library/publications/a-typology-of-players-in-the-game-physics-playground/>.
- Slater, S., Joksimović, S., Kovanovic, V., Baker, R.S., Gasevic, D., 2017b. Tools for educational data mining: a review. *J. Educ. Behav. Stat.* 42 (1), 85–106.

- Stanford University, (n.d.). Artificial Intelligence, [online]. Available from. <http://jmc.stanford.edu/artificial-intelligence/index.html>.
- Stewart, A.E.B., Vrzakova, H., Sun, C., Yonehiro, J., Stone, C.A., Duran, N.D., Shute, V., D'Mello, S.K., 2019. Proceedings of the ACM on Human-Computer Interaction. I Say, You Say, We Say: Using Spoken Language to Model Socio-cognitive Processes During Computer-Supported Collaborative Problem Solving, vol. 3. CSCW, pp. 1–19. <https://doi.org/10.1145/3359296>.
- Strain, D., 2020. New \$20 Million Center to Bring Artificial Intelligence Into the Classroom. CU Boulder Today. <https://www.colorado.edu/today/ai-education>.
- Sukkarieh, J.Z., Blackmore, J., 2009. c-Rater: automatic content scoring for short constructed responses. In: Lane, H.C., Guesgen, H.W. (Eds.), Proceedings of the Twenty-Second International Florida Artificial Intelligence Research Society Conference. Association for the Advancement of Artificial Intelligence Press, Menlo Park, CA, pp. 290–295.
- Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., Kasneci, E., 2021. Multimodal engagement analysis from facial videos in the classroom. IEEE Trans. Affect. Comput. <https://doi.org/10.1109/TAFFC.2021.3127692>.
- Turing, A.M., 1995. Lecture to the London Mathematical Society on 20 February 1947. M Comput. 12, 390–404.
- Zheng, S., Trott, A., Sunil, S., Nikhil, N., Melvin, G., Parkes, D.C., Socher, R., 2020. The AI economist: improving equality and productivity with AI-driven tax policies. ArXiv 1–46. Available from. <https://ui.adsabs.harvard.edu/abs/2020arXiv200413332Z>.

Social and emotional learning

Kimberly A. Schonert-Reichl, University of Illinois at Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

SEL: rationale and definitions	93
Why now? The need for SEL	93
SEL and neuroscience: we feel therefore we learn	93
Escalating mental health problems in children and youth	93
Widespread support for SEL	93
A brief history of SEL	94
What is SEL?	95
SEL frameworks	95
CASEL's SEL framework	95
What is transformative SEL?	96
What is systemic SEL?	97
SEL: the research evidence	97
SEL skills in childhood predict academic achievement and adult success	97
Social and emotional skills are malleable—they can be taught	98
SEL skills are durable	99
SEL: the essential ingredients	99
Ingredient 1: the learning context	99
Ingredient 2: SEL of students	100
Ingredient 3: SEL of teachers	101
Caring teachers: strategies for the classroom	101
Teacher stress and stress contagion	101
Promoting teachers' social and emotional competence and well-being	102
Future SEL efforts on the horizon	102
Acknowledgement	103
References	103
Further reading	105

Research has clearly shown that social-emotional competencies can be taught, that schools are appropriate places to teach them, and that SEL can make a positive difference in young people's lives.

Weissberg (2019, p. 65).

Recent years have witnessed amplified attention to the field of social and emotional learning (SEL) fueled by burgeoning research evidence indicating that promoting students' social and emotional competence in school improves not only their well-being and positive development, but such efforts lead to successful academic outcomes. Bolstered by a plethora of compelling empirical evidence showing that SEL programs and practices are successful in promoting students' social-emotional competence and deterring mental health problems and antisocial behaviors (e.g., Domitrovich et al., 2017; Weissberg et al., 2015) as well as cost-effective (Belfield et al., 2015), many school districts and jurisdictions in the US and internationally have strengthened their efforts to integrate SEL into the very fabric of schooling in efforts to educate the "whole child" and prepare young people for the 21st century (Darling-Hammond and Cook-Harvey, 2018).

Clearly, SEL is now a worldwide phenomenon and not just a passing fad, with SEL approaches and programs being implemented in countries throughout the world (Frydenberg et al., 2017; Humphrey, 2013; Torrente et al., 2015). Large scale organizations such as the World Bank, the World Health Organization (WHO), UNICEF, UNESCO, and the Organization for Economic and Cooperative Development (OECD) are joining in the call for a more explicit and intentional consideration of social and emotional competencies and social and emotional well-being into both education and health (e.g., OECD Education, 2018; Varela et al., 2013).

The overall aim of this chapter is to provide a summary of current theory and research in the field of SEL. The chapter begins by providing a rationale for the promotion of students' social-emotional competence in education, followed by a brief history of SEL, a definition of SEL, and a description of the social and emotional competencies that comprise SEL. Following a brief review of recent research that offers strong empirical support for an SEL approach is put forth. Next, essential ingredients of a systemic approach to SEL are delineated. The chapter ends by offering some future directions conclusions on how an understanding of SEL has implications for the transformation of education.

SEL: rationale and definitions

Why now? The need for SEL

My own work on SEL originated during my time as a middle school teacher and a teacher at an alternative high school for youth deemed “at risk” for high school completion during the 1980s. My students inspired me, taught me, and coached me to learn from them about what it takes to create an educational context in which they could learn and flourish. The phrase that resounded for me was “Students don’t care how much you know until they know how much you care.” It was at that time that I learned that my success as a teacher relied on not only educating my students’ minds, but educating their hearts as well. And to do so required that I pay explicit attention to cultivating their social-emotional competencies, including their ability to be aware of their strengths and challenges, identify and regulate their emotions, show empathy and compassion for others, and seek help when they needed it. Nonetheless, what was missing at that time was the field of SEL and the science behind it, so I had limited information—I worked on instinct and had little from which to draw from my teacher training or support from the educators in my school context. It was during those years that I also realized that so many of the social and emotional problems that I saw in my students were preventable. This inspired me to enroll in graduate school so that I could explore the research on prevention and the promotion of students’ social and emotional competencies to better understand the social and emotional dimensions of teaching and learning. This work inspired a professional career that has spanned over 30 years in which I have sought to answer one central question: What are the ways that we can create educational contexts that foster students’ social and emotional competencies, such as empathy, compassion, altruism, and resiliency so that all young people can thrive and reach their optimal potential?

SEL and neuroscience: we feel therefore we learn

Many of the recent discussions on education and educational reform have emphasized that young people’s social, emotional, and academic development are inextricably linked (Cantor et al., 2018; Jones and Kahn, 2017). These conjectures draw strong support from research in the field of social neuroscience. As Immordino-Yang and Damasio (2007) assert, “The aspects of cognition that are recruited most heavily in education, including learning, attention, memory, decision making, motivation, and social functioning, are both profoundly affected by emotion and in fact subsumed within the processes of emotion” (p. 7). Stated simply, how we feel affects how and what we learn.

Recent innovations in neuroscience that emphasize the vital role of social experiences on brain development and learning support the explicit integration of SEL into education (Immordino-Yang et al., 2019). Groundbreaking new research that has emerged during the last half of the 20th century challenges pre-existing beliefs about the brain’s malleability by demonstrating that the brain is changeable (or “plastic”) to experience across childhood and adolescence, and even into adulthood (Baltes et al., 2006; Pascual-Leone et al., 2005). Similarly, research from the field of developmental cognitive neuroscience (Diamond, 2012), indicates that social and emotional skills can be taught across the life span and are viewed as more malleable than IQ.

Escalating mental health problems in children and youth

An increased emphasis on the role of schools in promoting students’ social and emotional competence and well-being reflects, in part, growing concerns about the problems facing youth today, such as declining academic motivation (Roeser and Eccles, 2014), increasing school bullying and aggression (Hymel and Swearer, 2015), an escalating mental health crisis (Jones et al., 2022), and decreases in well-being in youth exacerbated by the Covid-19 pandemic (Gadermann et al., 2022; Luthar et al., 2021).

The crisis of children’s mental health and rising disparities in educational outcomes has risen to the forefront of national consciousness with the release of the most recent US Surgeon General’s Advisory in December 2021. The Advisory calls attention to the rising mental health issues of young people and disparities across subpopulations in the US. It then outlines a whole-society approach for tackling the problem and includes a set of recommendations designed to prevent these challenges and bolster children’s well-being. One of the primary recommendations outlined is to “support the mental health of children and youth in educational, community, and childcare settings,” which involves, “creating positive, safe, and affirming educational environments” and “expanding programming that promotes healthy development (such as social and emotional learning)” (p. 13). To support the well-being and mental health of students, the Advisory also encourages policymakers and schools to undertake efforts designed to promote the “mental health of all school personnel” (p. 20). For instance, the recommendations suggest regular assessments of staff well-being and the integration of wellness initiatives into professional development for school staff. The Advisory proports that these efforts will directly benefit school personnel, while, at the same time, help to foster teachers’ social and emotional competencies (SECs, e.g., empathy and compassion)—competencies that are needed to create positive learning environments that can promote student well-being and healthy development (Lever et al., 2017; US Dept. of Education, 2021). Furthermore, other policy makers and advocacy groups in the US have called for the increased training of teachers in SEL so that teachers are better equipped to support the development of students’ SECs and well-being throughout the COVID-19 crisis and beyond (Walker, 2020).

Widespread support for SEL

There now exists consensus among parents, teacher, principals, and the general public that SEL should be integrated into schooling. Indeed, recent research indicates that the myopic focus on academics as the sole purpose of education appears to be shifting, with support for SEL from parents (Edge Research, 2018; PDK International, 2017) and the general public (Bushaw and Lopez, 2013;). Results from the 2013 PDK/Gallup Poll of the Public’s Attitudes Toward Public Schools indicates that most Americans agree that

public schools should teach students a full range of social, emotional, and cognitive competencies, including how to set meaningful goals (89%), communication skills (94%), how to collaborate on projects (84%), and character (76%).

Similarly, teachers confirm that SEL is an essential part of education. For example, a nationally representative survey published by Civic Enterprises and Peter D. Hart Research Associates of more than 600 teachers (Bridgeland et al., 2013) illustrates this point. Their report showed that the vast majority of preschool to high school teachers believe that social and emotional skills are teachable (95%) and that promoting SEL will benefit students from both rich and poor backgrounds (97%) and will have positive effects on their school attendance and graduation (80%), standardized test scores and overall academic performance (77%), college preparation (78%), workforce readiness (87%), and citizenship (87%). Additionally, these same teachers reported that, in order to implement and promote SEL effectively in their classrooms and schools, they need strong support from district and school leaders. These findings are important because they demonstrate that, although there is a readiness among teachers to promote SEL, there is a need for systemic supports for implementation at the district level. Similarly, a 2019 Rand Cooperation report by Hamilton, Doss, and Steiner echoes strong support for SEL, notably among both principals and teachers. Their survey revealed that a large majority of principals identified SEL as a top priority and that both principals and teachers identified a wide range of SEL skills as important (Hamilton et al., 2019).

Despite resounding support for SEL, it is not without its critics. Most recently, prominent scholars and education leaders have voiced their concerns about the rapid interest and implantation of SEL throughout the US (e.g., Finn and Hess, 2019; Hamilton and Schwartz, 2019; Hess, 2019; McShane, 2019; National Commission on Social, Emotional, and Academic Development, 2018; Starr, 2019). One frequent point of criticism for SEL is in regard to its definition. Some argue (e.g., Finn and Hess, 2019; Starr, 2019) that the definition for SEL is ambiguous and can mean many different things to different people. In a thoughtful and comprehensive response to concerns about SEL, Shriver and Weissberg (2020) posit that “Just like every other widely used term in K-12 education—deeper learning, mastery learning, collaborative learning, instructional leadership, and so on—SEL will no doubt be misunderstood by some people and misused by others trying to cash in on its popularity” (p. 53). They note that while the organization in which they both are members—the Collaborative for Academic, Social, and Emotional Learning or CASEL—has worked for over 25 years to define SEL, no one organization owns a definition of SEL, therefore some definitional ambiguity is inevitable. Despite this, Shriver and Weissberg contend that there is widespread agreement among leaders in the SEL movement on a set of foundational principles. Specifically, any SEL Framework should include a focus on both intrapersonal and interpersonal skills and attitudes; be developmentally appropriate and culturally sensitive; be grounded in empirical research; and support implementation and evaluation programming by offering practitioners resources that are evidence-based (Blyth et al., 2019). Undoubtedly, debates about SEL will continue—what is needed for the field to grow and flourish are productive and open conversations in which different points of view can provide an impetus for innovation and new knowledge.

A brief history of SEL

Although the term “SEL” is relatively recent—dating back almost 30 years—the attention to children’s social and emotional development in relation to education dates back much longer. For instance, centuries ago Aristotle posited that, “Educating the mind without educating the heart is no education at all” emphasizing that the aim of education should not solely focus on the mastery of subject content areas such as reading, writing, math, science, and social studies, but should include an explicit and intentional focus on teaching students social and emotional competencies, including empathy and compassion. More recent instantiations of SEL can be traced back to the early 1900s with John Dewey’s (1916) progressive education efforts that focused on the promotion of students’ social responsibility, Maria Montessori (1912) schools that emphasized child-centered education and a focus on whole child education—cognitive, social, emotional, and physical development, and Jane Addams (2002) focus on engaged citizenship and student agency and autonomy which aligned with Dewey’s focus.

The theoretical and empirical work in SEL related fields in particular can be traced back to the beginning of the 20th century. Osher et al. (2016), in their recent review of the genesis of SEL, noted that “The roots of SEL are intellectually diverse and politically nonpartisan” (p. 647). In their review, they trace back the roots of SEL to educational reform efforts in the early part of the 20th century (including John Dewey and Jane Addams) to the “Mental Hygiene Movement” with its presage attention to prevention of mental illness, and to Thorndike’s (1920) conception of “social intelligence.” Over the course of the 20th century, the shift to understanding the contextual influences that shape child development began to emerge and researchers such as Bronfenbrenner and his “bioecological model of human development” (1979; Bronfenbrenner and Morris, 2006) espoused the importance of taking into account the nature and levels of context as influencers on child development. Bronfenbrenner’s “bioecological model of human development” led to increased attention to the implementation of systemic interventions in the 1980s and 1990s (Elias et al., 1997).

The growth of SEL work was gradual and somewhat intermittent across the century but expanded markedly in the 1990s when researchers such as Salovey and Mayer (1990) began to work in the new field of “emotional intelligence.” Their work both built upon and expanded the work of Howard Gardner’s (1983) “multiple intelligences” and Robert Sternberg’s (1985) practical intelligence, challenged typical definitions of intelligence by showing the critical roles of emotion and motivation in predicting life success.

In 1995 the publication of Daniel Goleman’s book *Emotional Intelligence: Why It Can Matter More than IQ* catalyzed the field by making SEL a term known by researchers and the larger public. At the same time that interest in SEL was growing, there was an accumulating research base regarding the effectiveness of programs to prevent problem behavior and promote healthy development

coupled with a growing knowledge base on the importance of school climate to promote students' healthy relationships, achievement, and school connectedness, and prevent dropout (e.g., [Thapa et al., 2013](#)). Similarly, growing concerns around problems such as bullying and school shootings spurred an urgency for the creation and implementation of school programs focused on promoting safe and supportive school environments ([Dwyer and Osher, 2000](#); [Dwyer et al., 1998](#)).

Today, SEL has been characterized in a variety of ways, often being used as an organizing framework for an array of promotion and prevention efforts in education and developmental science, including conflict resolution, cooperative learning, bullying prevention, and positive youth development ([Schonert-Reichl and Weissberg, 2015](#); [Weissberg et al., 2015](#)).

What is SEL?

Social and emotional learning, or SEL, involves evidence-based programs, practices, and policies through which children, youth, and adults acquire the competencies to recognize and manage emotions, develop caring and concern for others, establish positive relationships, make responsible decisions, and handle challenging situations effectively ([Weissberg, 2019](#); [Weissberg et al., 2015](#)). That is, SEL programs and practices teach the personal and interpersonal competencies we all need to handle ourselves, our relationships, and our work effectively and ethically. Accordingly, SEL is aimed at helping children and even adults develop fundamental skills for success in school and life.

SEL builds from work in child development, classroom management, prevention, and emerging knowledge about the role of the brain in self-awareness, empathy, social-cognitive growth (e.g., [Best and Miller, 2010](#); [Goleman, 2006](#); [Immordino-Yang et al., 2019](#)), and focuses on the skills that allow children to calm themselves when angry, make friends, resolve conflicts respectfully, and make ethical and safe choices. In short, SEL competencies comprise the foundational skills for positive health practices, engaged citizenship, and school success ([Weissberg et al., 2015](#)). SEL is sometimes called "the missing piece" in education because it represents a part of education that is inextricably linked to school success but has not been explicitly stated or given much attention until relatively recently. SEL emphasizes active learning approaches in which skills can be generalized across curriculum areas and contexts when opportunities are provided to practice the skills that foster positive attitudes, behaviors, and thinking processes.

In SEL, social-emotional competencies are viewed as "mastery skills" underlying virtually all aspects of human functioning. Moreover, SEL emphasizes active learning approaches in which skills can be generalized across curriculum areas and contexts when opportunities are provided to practice the skills that foster positive attitudes, behaviors, and thinking processes.

SEL frameworks

The SEL landscape is replete with SEL frameworks. [Berg et al. \(2017\)](#) conducted an extensive review of frameworks related to the social and emotional (SE) competencies among youth between the ages of 6–25 years. Through their examination of reports, papers, and discussions with experts across a range of fields, including economics and health, they identified 136 frameworks across 14 areas of study, including areas such as Positive Youth Development, Resilience, Character Education, School-Based Competency Development, Public Health, Mental Health, Mindfulness to name a few (see [Berg et al., 2017](#), for a full description of their approach and findings). One central finding that emerged in their review was "that different terms are used for competencies that have similar definitions, and that the same terms are used for competencies that have different definitions" (p. 40). Their work underlines the importance of conceptual clarity when identifying and analyzing appropriate SEL frameworks.

CASEL's SEL framework

The Collaborative for Academic, Social, and Emotional Learning (CASEL; www.casel.org) a nonprofit organization in Chicago, IL, is one of the organizations at the forefront in North American and international efforts to promote SEL. Founded in 1994 by Daniel Goleman and Eileen Rockefeller Growald, its mission is to advance the science of SEL and expand evidence-based, integrated SEL practices as an essential part of preschool through high school education.

CASEL's SEL framework is one of the most recognized SEL frameworks across the world (see [Fig. 1](#)). The framework takes a systems approach that includes a set of social-emotional competencies that underlie effective and successful performance for social roles and life tasks, drawing from extensive research in a wide range of areas, including social neuroscience and methods of learning and instruction. The concentric circles that surround the five social-emotional competencies—notably, communities, families and caregivers, schools, and classrooms—are the key settings in which SEL promotion occurs. These serve as an organizing framework for coordinating policies and practices that students experience both inside and outside of the classroom. The five social-emotional competencies identified by CASEL are described in more detail below ([Weissberg et al., 2015](#); www.CASEL.org):

The CASEL Five.

1. *Self-awareness*: The abilities to identify and recognize one's own emotions, thoughts, and influences on behavior, including recognizing one's own strength and challenges, being aware of one's own goals and values, possessing a well-grounded sense of self-efficacy and optimism, and having a growth mindset that one can learn through hard work. High levels of self-awareness require recognizing how thoughts, feelings, and actions are interconnected. Dimensions of self-awareness include: integrating personal and social identities, identifying personal, cultural, and linguistic assets, demonstrating honesty and integrity, linking feelings, values, and thoughts, examining prejudices and biases, and developing interests and a sense of purpose.
2. *Self-management*: The abilities to regulate one's emotions, thoughts, and behaviors effectively, including stress management, delaying gratification, impulse control, motivating oneself, and persevering through challenges to achieve personal and

- educational goals. It also includes self-management within social interactions. Other skills include: using planning and organizational skills, showing the courage to take initiative, and demonstrating personal and collective agency
3. *Social awareness*: The abilities to take the perspectives of others—including those who come from a different background and culture, to empathize with others, understand social and ethical norms, and to recognize resources and supports in family, school, and community. Other social awareness competencies include: recognizing strengths in others, demonstrating empathy and compassion, showing concern for the feelings of others, understanding and expressing gratitude, identifying diverse social norms, including unjust ones, recognizing situational demands and opportunities, and understanding the influences of organizations/systems on behavior.
 4. *Relationship skills*: The abilities to establish and maintain healthy and supportive relationships and to effectively navigate settings with diverse individuals and groups, communicate clearly, listen actively, cooperate, negotiate constructively during conflict, solve problems with others, navigate settings with differing social and cultural demands and opportunities effectively, demonstrate cultural competency, practice teamwork and collaborative problem-solving, resist social pressure, show leadership in groups, and to offer and seek help when needed, and stand up for the rights of others.
 5. *Responsible decision making*: This includes the knowledge, skills, and attitudes to make constructive choices regarding one's own behavior and social interactions, taking into account safety concerns, ethical standards, social and behavioral norms, consequences, and the well-being of self and others, and to evaluate the benefits and consequences of various actions for personal, social, and collective well-being. Additional examples include: demonstrating curiosity and open-mindedness, identifying solutions for personal and social problems, learning to make a reasoned judgment after analyzing information, data, facts, recognizing how critical thinking skills are useful both inside and outside of school, reflecting on one's role to promote personal, family, and community well-being, and valuating personal, interpersonal, community, and institutional impacts.

What is transformative SEL?

Until recently, the definition of SEL did not include an explicit focus on considering the cultural and racial differences, privilege, and inequities that exist in schools (Jagers et al., 2019). To address this and to help advance the research agenda for the next generation by focusing on SEL in the service of equity and excellence, an expanded definition of SEL has been put forth by Jagers et al. (2019). The concept “transformative SEL” is the process whereby young people and adults build strong, and respectful relationships that facilitate co-learning to critically examine root causes of inequity and to develop collaborative solutions that lead to personal, community, and societal well-being. This new definition sets a vision for what high-quality systemic SEL is and how it might be achieved. Moreover, this form of SEL is necessary to meet the growing political, economic, and health challenges we face in the US and around the world.

Transformative SEL competencies focus on identity, intersectionality, agency, belonging and engagement as central to furthering social-emotional development and achieving equity in education (Jagers et al., 2019). Strategies identified as promoting these



Fig. 1 The collaborative for academic, social, and emotional learning (CASEL) framework.

transformative social-emotional competencies are culturally infused SEL skill development, project-based learning and youth participatory action research.

Also relevant to a discussion of an equity in transformative SEL are disciplinary approaches. Restorative practices for resolving conflict and behavioral issues offer an inclusive approach to counteract exclusionary, punitive measures that disproportionately marginalize students of color (Gregory and Fergus, 2017). Research evidence suggests that restorative approaches teach students skills for conflict resolution and help students become better integrated into the school community (Morrison, 2013; Wearmouth et al., 2007).

What is systemic SEL?

Because virtually all children go to school during key developmental years, schools are important contexts for fostering children's positive social and emotional competencies. Widely acknowledged is that efforts to make SEL an enduring aspect of education, need to include SEL at every level of the school, including school leadership, teaching and learning, and with families and community (see Mahoney et al., 2020). As recently espoused by Mahoney et al. (2020, p. 1),

Systemic SEL is an approach to create equitable learning conditions that actively involve all Pre-K to Grade 12 students in learning and practicing social, emotional, and academic competencies. These conditions require aligned policies, resources, and actions at state and district levels that encourage local schools and communities to build the personal and professional capacities of adults to: implement and continuously improve evidence-based programs and practices; create an inclusive culture that fosters caring relationships and youth voice, agency, and character; and support coordinated school-family-community partnerships to enhance student development.

Especially noteworthy in the above description of systemic SEL is that any approach to promote social and emotional competence in students must consider the interpersonal and intrapersonal capacities of the adults in the education system.

CASEL has developed a set of guidelines and school and school district resources for implementing SEL at the level of the system (www.CASEL.org). For example, for schools, CASEL has put forth "10 Indicators of Schoolwide SEL" that are evidence-based, and include: (1) Explicit SEL instruction, (2) SEL integrated with academic instruction, (3) Youth voice and engagement, (4) Supportive school and classroom climate, (5) Focus on adult SEL and relationships, (6) Supportive discipline, (7) A continuum of integrated supports, (8) Systems for continuous improvement, (9) Authentic family partnerships, and (10) Aligned community partnerships. Full descriptions of these, along with a rubric for assessing them, and many other resources from CASEL for integration of SEL schoolwide, can be found at: <https://schoolguide.casel.org/>

SEL: the research evidence

SEL skills in childhood predict academic achievement and adult success

A growing body of literature supports the premise that children's social-emotional competence not only predicts success in school (e.g., Oberle et al., 2014; Wentzel, 1993), but also predicts a range of important outcomes in late adolescence and adulthood, including physical health, substance dependence, and overall well-being (Moffitt et al., 2011). Recognizing the interrelationship between social-emotional competencies and academic success, researchers have argued that fostering positive social and emotional development may be key to enhancing academic growth (see Greenberg et al., 2003; Zins et al., 2004).

One early study that was pioneering in its design and approach to demonstrating the critical role of social-emotional competencies in predicting students' academic success, is from Wentzel (1993) titled "Does Being Good Make the Grade?" In her study of 423 sixth and seventh graders, Wentzel (1993) found that students' prosocial classroom behaviors, such as helping, sharing, and cooperating, were better predictors of academic achievement than were their standardized test scores, even after taking into account academic behavior, teachers' preferences for students, IQ, family structure, sex, ethnicity, and days absent from school. Similarly, in a longitudinal study of 294 Italian children, Caprara et al. (2000) found that prosocial behavior in third grade (average age 8.5 years) as rated by self, peers, and teachers, significantly predicted both academic achievement (explaining 35% of the variance) and social preference (explaining 37% of the variance) 5 years later when children were in eighth grade. Most interestingly, this "prosocialness" score, which included cooperating, helping, sharing, and consoling behaviors, significantly predicted academic achievement five years later, even after controlling for third grade academic achievement. In contrast, early academic achievement did not contribute significantly to later achievement after controlling for effects of early prosocialness. Collectively, these correlational studies, demonstrate the importance of social-emotional competencies in predicting short-term and long-term success.

In a short-term, longitudinal study of 441 sixth grade Canadian students, Oberle et al. (2014) examined the association between social and emotional competence and academic achievement in early adolescents. Social-emotional competence in grade 6, operationalized in terms of both self-reports of social responsibility goals and teacher assessments of frustration tolerance, assertive social skills, task orientation, and peer interaction, were evaluated as predictors of student academic achievement test scores in math and reading grade 7. As hypothesized, teacher-reports of students' social-emotional competencies significantly predicted higher scores in math and reading in seventh grade. Self-reported social-emotional competence in grade 6 was a significant predictor of grade 7 reading scores for boys but not girls. Although more research is needed regarding the link between SEL and academic

achievement, there is a confluence of empirical evidence suggesting that, if we want students to be successful in school, efforts should be made to teach SEL intentionally and explicitly.

In addition to playing a crucial role in predicting academic success, recent longitudinal research also documents links between children's social and emotional skills and later success in adulthood. Jones et al. (2015), for instance, examined the degree to which late adolescent and early adult outcomes were predicted by teacher ratings of children's social-emotional competence measured many years earlier, when children were in kindergarten, following 753 kindergarten children longitudinally, 13–19 years later. Kindergarten teacher ratings of children's prosocial skills (getting along with others, sharing, cooperating) were found to be significant predictors of whether participants graduated from high school on time, completed a college degree, obtained stable employment in adulthood, and were employed full time in adulthood. Moreover, kindergarten children who were rated by their teachers as high in prosocial skills in kindergarten, were less likely as adults to receive public assistance, live in or seek public housing, be involved with police, placed in a juvenile detention facility, or arrested. Early social competence inversely predicted days of binge drinking in the last month and number of years on medication for emotional or behavioral problems during high school. Given these findings, the authors emphasize the importance of assessing young children's social and emotional competence early on and contend that these "softer" skills can be more malleable than IQ or other cognitive measures and, hence are important contenders for intervention.

In another notable longitudinal study Moffitt et al. (2011) followed a cohort of 1000 children from birth to 32 years in New Zealand, assessing children's self-control across the ages of 3, 5, 7, 9, and 11 years via reports from researcher-observers, teachers, parents, and the children themselves. Self-control in childhood was found to predict physical health, substance dependence, personal finances, and criminal offending in adulthood, even after taking into account intelligence, social class, and problems the children had in adolescence (e.g., smoking, school drop-out). The authors concluded that focusing on the promotion of children's self-control "might reduce a panoply of societal costs, save taxpayers money, and promote prosperity" (p. 1). Thus, results from several recent longitudinal studies examining the association between early SEL skills and later adult adjustment suggest that, in the long run, higher levels of social and emotional competence can increase the likelihood of high school graduation, financial success, mental and physical health, and reduce criminal behavior.

Social and emotional skills are malleable—they can be taught

One approach to promoting the well-being and success of our young people and to ensure that they grow into productive adults is to identify the cognitive and social-emotional knowledge, skills, attitudes, and values that are considered to be "malleable"—that is, those that can change and be promoted through education and other experiences. Although the previous terminology in psychology referred to many of these constructs as "traits," hence conveying a sense of immutability or permanence, current descriptions identify these constructs as "skills" because they can be shaped and changed over the life cycle.

As described earlier, SEL is grounded in research from developmental cognitive neuroscience (e.g., Diamond, 2012) that indicates that social and emotional skills can be taught across the life span and are more malleable than IQ. Immordino-Yang et al. (2019) posit that although the malleability of the brain can be changed by experience through the lifespan, there remain periods across human development for which the brain experiences is more actively changing. These periods include the prenatal period through childhood, early adolescence, transition to parenthood, and old age. This information is particularly important when making decisions about SEL programming during the elementary, middle, and high school years.

Perhaps some of the most compelling evidence for the malleability of students' social-emotional competencies comes from a meta-analysis conducted by Durlak et al. (2011) of 213 school-based, universal SEL programs involving 270,034 students from kindergarten through high school. Students in SEL programs, relative to students who did not receive an SEL program, were found to demonstrate significantly improved social-emotional competencies, attitudes, and behavioral adjustment (increased prosocial behavior and decreased conduct problems and internalizing problems). SEL students also outperformed non-SEL students on indices of academic achievement by 11-percentile points. Durlak et al. also found that classroom teachers and other school personnel effectively implemented SEL programs. Thus, SEL programs can be easily incorporated into routine school practices and do not require staff from outside the school for successful delivery. Taken together, these results provide strong empirical evidence for the "value-added" of SEL programs in fostering students' social and emotional skills, attitudes, and behaviors, and also counter the claim that taking time to promote students' SEL would be detrimental to academic achievement.

Similar results were obtained in another meta-analysis of 75 recently published studies of SEL programs conducted by Sklad et al. (2012). That is, Sklad et al. found that universal, school-based SEL programs had significant positive effects on seven outcomes: social-emotional skills, prosocial behavior, positive self-image, academic achievement, antisocial behavior, mental health problems, and substance abuse. Not surprisingly, the most positive effects were found for social-emotional skills, with an effect size of 0.70. In other words, students participating in SEL programs had social-emotional skills 0.7 of a standard deviation higher than comparison students, indicating that the average SEL program student had better social-emotional skills than 76% of non-SEL students. Moderate effect sizes (program effects of nearly a half of a standard deviation) emerged for four of the outcomes: academic achievement, positive self-image, prosocial behavior, and antisocial behavior. As for follow-up effects, the largest effects were found for academic achievement, followed by substance abuse.

SEL skills are durable

Do students maintain their SEL competencies after the SEL program has ended? Findings from [Durlak et al. \(2011\)](#) meta-analysis described above provide additional support for the durability of effects of SEL programming on students' social and emotional competencies. Durlak and colleagues examined a smaller group of 33 interventions that included follow-up data (an average follow-up period of 92 weeks) and found that the positive effects remained statistically significant, although the effect sizes were smaller.

Research by [Hawkins et al. \(2008\)](#) documented the long-term positive effects of multi-year SEL programming on student outcomes. Specifically, Hawkins et al. found significantly reduced diagnosable mental health disorders (e.g., major depression, generalized anxiety disorder) at age 24 and age 27, which was 12 and 15 years after their SEL intervention had ended. Their results also showed intervention effects indicating better educational and economic achievement among those individuals who received the SEL intervention in contrast to those who did not. Although more research is clearly needed, Hawkins et al.'s research provides important evidence about the potential long-term benefits of well-designed and well-implemented SEL interventions.

More recently, [Taylor et al. \(2017\)](#) conducted a systematic review to address the relative paucity of evidence on the long-term effectiveness of SEL programming in enhancing positive student outcomes. This review is important because it addresses a critical question regarding the cost-benefit of investment in SEL programs, informing allocation of resources for SEL in school budgets. A total of 82 research studies evaluating school-based, universal SEL programs involving 97,406 ethnically and socio-demographically diverse students, from kindergarten to high school and in urban and rural settings were reviewed. Results demonstrated that students who had received an SEL intervention continued to show increases in social-emotional skills, positive behaviors and academic achievement, and decreases in conduct problems, emotional distress, and drug use up to almost four years after program completion, in contrast to those students who did not receive an SEL intervention.

SEL: the essential ingredients

As noted earlier, several organizing frameworks have been proposed for SEL, each one outlining various components that influence SEL—such as school culture and climate, and teachers' pedagogical skills—and each framework identifies similar student outcomes, such as improved academic achievement and social and emotional competence. Hence, in addition to focusing on specific instruction in social and emotional skills, SEL is a process of creating a school and classroom community that is caring, supportive, and responsive to students' needs. Moreover, effective SEL interventions and skill development occurs when teachers possess the requisite social and emotional skills to create an environment that is safe, caring, supportive, and well-managed, and have the social-emotional competencies and knowledge to implement SEL programs effectively.

To create a parsimonious depiction of a systemic approach to SEL, I contend that there are three distinct and interrelated dimensions: (1) The learning context, (2) SEL of students, and (3) SEL of teachers. As illustrated in [Fig. 2](#), these three dimensions are portrayed in a circle to illustrate the interconnectedness of each dimension to each other and to highlight that each dimension is influenced by, and influences, the other dimensions. Each one is discussed in turn below.

Ingredient 1: the learning context

Children can learn and thrive when they are in school and classroom contexts in which they feel safe, secure, connected, and cared for—contexts in which their social emotional competence and academic growth is nurtured and cultivated ([Jones and Kahn, 2017](#)). Conversely, stressful classroom contexts can thwart children's learning and jeopardize their mental health—leading them on a trajectory toward educational failure and problem behaviors ([Domitrovich et al., 2017](#); [Sauve and Schonert-Reichl, 2019](#)). Classrooms and schools operate as systems, and decades of research suggest that the unique culture and climate of classrooms and schools affects how and what students learn (e.g., [Thapa et al., 2013](#)). School culture refers to a general set of norms, beliefs, and practices or "the way things are done around here" ([Hemmelgarn et al., 2006](#)), whereas school climate "reflects norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures" ([National School Climate Council, 2007](#), p. 4). Culture and climate, in combination, influence the interactions and relationships among administrators, teachers, school staff, and students and their approaches to teaching and learning ([Gottfredson et al., 2005](#)). Therefore, any approach to promoting SEL needs to take into account both school culture and climate and systematically and intentionally embed SEL into every dimension of a school.

SEL interventions and skill development should occur within supportive classroom and school environments, as well as help to create such a climate. Additionally, successful SEL-related school and classroom activities foster active student voice in decision-making, problem solving, and engagement in lifelong learning. Research also has shown that effective programs provide repeated opportunities to practice new skills and behaviors within the program structure and to apply them in real-life situations. That is, providing opportunities to practice within classroom lessons is important, but actual opportunities to practice in real-life situations are likely to have even more impact ([Durlak et al., 2011](#); [Weare and Nind, 2011](#)). Communication styles, high performance expectations, classroom structures and rules, school organizational climate, commitment to the academic success of all students, district policies, and parental and community involvement are all important components of an SEL approach in the learning context.

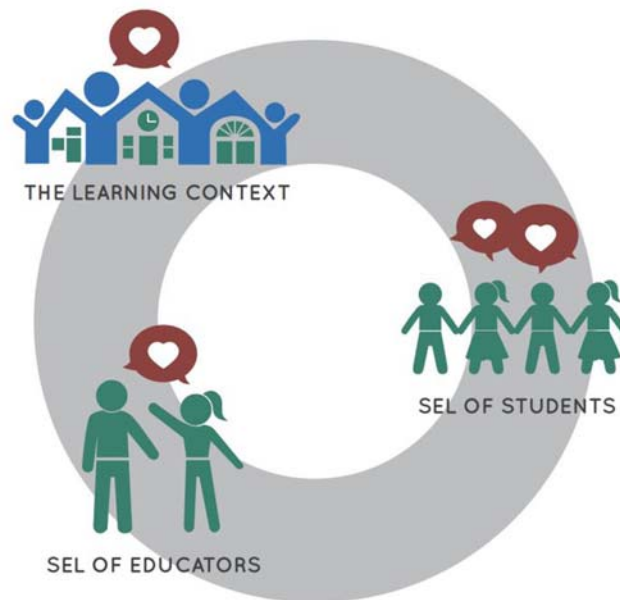


Fig. 2 The essential ingredients of SEL.

Ingredient 2: SEL of students

Promoting the SEL students includes several dimensions, including explicit SEL instruction through evidence-based SEL instruction, infusion of SEL into academic subjects, and providing students with opportunities for voice and agency. In 2003 CASEL published the results of a careful review of 80 SEL programs for kindergarten through to 12th grade entitled, *Safe and Sound: An Educational Leader's Guide to Evidence-Based Social and Emotional Learning (SEL) Programs*. This groundbreaking review provided educators with practical information on the features of different programs that could help them select programs most relevant and suited for their particular needs. In 2013, CASEL reviewed hundreds of SEL preschool and elementary school programs and identified dozens of programs that met high standards for program design, implementation supports, evidence of effectiveness, and applicability to specific grades. To be included in the 2013 CASEL Guide (CASEL, 2013) and designated as "SElect", the program had to meet three criteria: (a) be a well-designed classroom-based program that systematically promotes students' social and emotional competence, provides opportunities for practice, and offers multi-year programming, (b) deliver high-quality training and other implementation supports, including initial training and ongoing support to ensure sound implementation and, (c) be evidence-based with at least one carefully conducted evaluation that documents positive impacts on student behavior and/or academic performance. The 2013 CASEL Guide summarizes objective information about the characteristics of these nationally available, multi-year programs via a "consumer report" format that is clear and easy to read. Included in the Guide is information on SEL programs identified as CASEL SElect programs designed for preschool through elementary school students. A guide for middle school and high school programs was released by CASEL in 2015. Moreover, the CASEL Program Guide continues to be updated as new SEL programs are reviewed. Similar efforts to evaluate the effectiveness of various SEL programs are available on the University of Colorado Blueprints website (www.blueprintsprograms.com) and on the Public Health Agency of Canada's Violence Prevention Portal (<http://cbpp-pcpe.phac-aspc.gc.ca/public-health-topics/violence-prevention/>).

With regard to SEL program implementation, it is also critical to keep in mind the critical elements of effective SEL programs. In this vein, in their meta-analyses, Durlak et al. (2011) provided evidence that SEL programs promote better student outcomes when program implementers incorporate four elements represented by the acronym SAFE:

- (1) *Sequenced*: connected and coordinated set of activities to foster skills development;
- (2) *Active*: active forms of learning to help students master new skills;
- (3) *Focused*: a component that emphasizes developing personal and social skills; and
- (4) *Explicit*: targeting specific social and emotional skills.

The effective *implementation* of an SEL program plays a crucial role in influencing student outcomes. Unfortunately, some well-designed SEL programs do not promote positive student outcomes, often due to variability in way the program is implemented in the "real world" setting of a school and/or classroom. When implementing an established SEL program that has been shown to be effective, it is important for educators to recognize the importance of completing all lessons and activities in the program (dosage) and doing so as designed by the program developers (fidelity) in order to maximize the likelihood of success of the program in their own classroom environment.

Ingredient 3: SEL of teachers

Teachers are the engine that drives SEL programs and practices in classrooms and school (Schonert-Reichl, 2017) and teachers own SEL competence and well-being plays a critical role in influencing the learning context and the infusion of SEL into classrooms and schools (Jones et al., 2013). Researchers have shown that promoting the social and emotional competencies and of students is most effective when explicit attention is given to all levels of the system, including educator SEL (Mahoney et al., 2020).

Classrooms with warm teacher-child relationships facilitate deep learning among students (Merritt et al., 2012), and when children feel comfortable with their teachers and peers, they are more willing to grapple with challenging material and persist at difficult learning tasks. Conversely, when teachers poorly manage the social and emotional demands of teaching, students demonstrate lower levels of performance and on-task behavior (Marzano et al., 2003). Hence, it is essential that efforts are made to support the development of teachers' SEL competencies in order to optimize their classroom performance and their ability to promote SEL in their students (Jennings and Frank, 2015).

Caring teachers: strategies for the classroom

A caring teacher can transform the school experience, especially for students who face enormous difficulties, such as dropping out or dysfunctional home lives. The quality of teacher-student relationships is critical for children's academic achievement as illustrated by the work of Maldonado-Carreno and Votruba-Drzal (2011). Using data from the NICHD Study of Early Childcare of 1364 children from kindergarten through fifth grade, they found that increases in the quality of teacher-student relationships were associated with concomitant improvements in teacher-reported academic skills. Moreover, these relations remained unchanged as children progressed from kindergarten through fifth grade.

Some explicit strategies for fostering positive student-teacher relationships and caring classrooms include:

- Greeting students every day as they enter the classroom, with intentional efforts to have a brief *positive* conversation with them (e.g., noticing their new backpack or shoes) (Cook et al., 2018).
- Getting to know each student and the lives they live, learning about their strengths, challenges, interests and dreams. This could be done at the beginning of the school year, through individual interviews with each student. Teachers can also ask students about what they, as teachers, can do to help students learn and thrive in school.
- Actively listening to students to show you care. Authentic listening is demonstrated by hearing your students and then checking back with them to make sure you understand. Such interactions help to develop a trusting relationship between teachers and students.
- Asking students for advice and feedback as well as help when needed. For example, teachers can ask for help in setting up the classroom (e.g., what to put on the classroom bulletin boards, how to arrange the seating, how to organize activity centers), giving students a voice in the nature and organization of their physical environment. Through regular class meetings, teachers can engage students in developing the rules for the classroom and in creating a positive classroom environment. By considering student feedback, teachers demonstrate that student opinions and experiences are valued, and helps to create a classroom culture in which students feel safe to ask questions and take chances, enhancing the development of their SEL skills as well as their academic success.

Teacher stress and stress contagion

Clearly, teachers play a vital role in cultivating a positive and nurturing context that fosters students' social, emotional, and academic success (e.g., Braun et al., 2020). Nonetheless, teachers' stress and burnout can undermine their efforts to create an optimal context for healthy development and learning for their students. Decades of research have demonstrated that teaching is one of the most stressful professions in the human service industry (Montgomery and Rupp, 2005). Chronic work stress and exhaustion among teachers is associated with negative changes in biological indicators of stress as well with recent research showing that teachers who report chronic stress demonstrate atypical patterns of physiological stress reactivity as assessed via diurnal cortisol (Katz et al., 2016; Wolfram et al., 2013). Moreover, mounting evidence indicates that teachers' stress and burnout impedes the development of positive teacher-student relationships, effective classroom management, and student well-being (Jennings and Greenberg, 2009; Roeser et al., 2013; Schonert-Reichl, 2017).

In the context of education, teachers can experience stress when they appraise a situation as threatening and yet have limited ability to change or improve it. Take, for instance, the case of teacher autonomy. Among professional occupations, teachers rate lowest in feeling that they have a say in what happens in the workplace (Gallup, 2014). The percentage of teachers who report low job autonomy increased from 18% in 2004 to 26% in 2012 (Sparks and Malkus, 2015). The proportion of teachers who report significant levels of on-the-job stress is also rising. In a recent Gallup (2014) on occupational stress, 46% of teachers reported high daily stress—on par with nurses and just above doctors. Teachers and nurses had the highest levels of reported stress among all occupational groups.

Why does teacher stress matter for our understanding of SEL? High levels of chronic stress can lead to occupational burnout—characterized by emotional exhaustion, depersonalization and a low sense of accomplishment in one's work (Maslach et al., 2001). Teacher stress has also been linked to decreased job satisfaction, poor instructional practices and poor student outcomes (e.g., Schwarzer and Hallum, 2008). Taken together, there is evidence that the occupational stress of teachers can thwart efforts to promote students' social-emotional competence and well-being.

Chronic work stress and exhaustion among teachers is also associated with negative changes in biological indicators of stress. Recent research has found that teachers who report chronic stress demonstrate atypical patterns of physiological stress reactivity, as assessed via daytime levels of the stress hormone cortisol (Katz et al., 2016; Wolfram et al., 2013).

Moreover, research also shows that, like other emotions, stress is contagious and can spill over into the classroom. That is, when teachers are stressed, students are the collateral damage. Drawing from the stress-contagion framework, Oberle and Schonert-Reichl (2016) examined the link between teacher burnout and student stress in a sample of fourth and seventh grade children in Canada. To assess teacher burnout, teachers completed the Maslach Burnout Inventory modified for Teachers (Maslach et al., 1996). To assess student stress, students' salivary cortisol was collected as a biological indicator of students' stress reactivity. Biological stress reactivity is frequently assessed via the reactivity of the hypothalamic-pituitary-adrenal (HPA) axis, a homeostatic system that follows a circadian rhythm and is activated in response to cognitive (e.g., fear, excitement, anxiety) or non-cognitive (e.g., infections) stressors (Jessop and Turner-Cobb, 2008). Cortisol levels found in saliva or blood can be used as an indicator for HPA axis activity. The integrity of the HPA axis is essential to human health. In a typical diurnal HPA-axis regulation pattern, cortisol levels rise within 20–45 min after waking and then gradually decline across the day. Inappropriately low or elevated levels of cortisol can compromise HPA axis functioning (Jessop and Turner-Cobb, 2008). Students' salivary cortisol was collected from children at 9 a.m., 11.30 a.m., and 2 p.m. in the classroom setting. Analyses revealed that, after adjusting for differences in cortisol levels due to age, gender and time of waking, higher morning cortisol levels in students could be significantly predicted from higher levels of self-reported burnout of classroom teachers. Although these findings were correlational, the research conducted by Oberle and Schonert-Reichl (2016) was the first to show that teachers' occupational stress is linked to students' physiological stress regulation. What is not yet known is the direction of the stress contagion. That is, does teachers' burnout lead to higher levels of stress in students? Or do students who enter the classroom with higher levels of stress lead to increased teacher burnout? Only future research determining this causal relationship will lend further clarity to this relation.

Promoting teachers' social and emotional competence and well-being

As can be surmised, the majority of extant studies have focused on teacher well-being through a deficit lens by examining the teacher stress and burnout (Spilt et al., 2011; Collie et al., 2015). This research has been in concert in showing a general lack of teacher well-being among teachers across the globe (Kyriacou, 2011). Moreover, the research has shown that higher levels of teacher stress and burnout is associated with concomitant levels of student problem behavior (Hoglund et al., 2015) and lower student academic achievement (e.g., Herman et al., 2018). Although these findings are valuable, one criticism of examining teacher well-being through a deficit lens is that it fails to identify the influential factors within a school system that serve to strengthen teacher well-being (Collie et al., 2015).

Reviewing the evidence linking teachers own SEL competence and student outcomes, Jennings and Greenberg (2009) point to the importance of quality teacher–student relationships, and effective student and classroom management skills, (as well as implementation dosage and fidelity) in obtaining the best outcomes for students. Accordingly, they recommend the development and implementation of interventions designed to specifically address teachers' SEL competencies, reduce teacher stress and burnout, and improve teacher well-being. Although limited, the past few years have seen the emergence of interventions specifically targeted at improving teachers' SEL and stress management. For example, two programs designed to promote teachers' SEL competence by incorporating mindfulness-based approaches are CARE (Cultivating Awareness and Resilience in Education) and SMART-in-Education (Stress Management and Resiliency Training). Mindfulness is typically described as an attentive, nonjudgmental, and receptive awareness of present moment experience in terms of feelings, images, thoughts, and sensations/perceptions (e.g., Kabat-Zinn, 1990). Both programs aim to increase teachers' 'mindfulness, job satisfaction, compassion and empathy for students, efficacy for regulating emotions, and decrease stress and burnout. Initial research to date has supported the effectiveness of both the CARE (Cocchia and Greenberg, in press; Jennings et al., 2011) and SMART-in-Education (e.g., Benn et al., 2012; Roeser et al., 2013) programs in promoting teacher SEL competence and well-being. Nonetheless, further research is needed to examine whether such positive changes in teacher well-being spill over into the classroom and lead to improvements in students' SEL competence.

Future SEL efforts on the horizon

Now more than ever, SEL is needed to revolutionize and transform education across the globe in order to enable all children and youth to thrive. Today, SEL is no longer seen as an ancillary aspect of classrooms and schools—it is seen as an essential component. Simply stated, "SEL is not another thing on the plate, it is the plate."

Although much has been learned in the past decade about SEL and its effects on children's social-emotional competence and academic success, the field has further to go before firm conclusions can be made about the specific ways in which an SEL approach advances children's short-term and long-term school and life success. Indeed, many questions still remain regarding the ways in which programs and practices designed to promote children's SEL skills will best support children's future success. Many questions remain, such as: How does a systemic approach to SEL lead to successful improvements in children's social and emotional competence and short-term and long-term success? Which programs work best for which children and youth? What does a transformative approach to SEL look like at all levels of the system? What are the policies that need to support effective SEL implementation at federal and state levels? and how do we adequately prepare and support teachers and principals to implement SEL in classrooms and schools?

Another challenge that confronts the field of SEL is the translation of knowledge garnered from rigorous research on the effectiveness of programs into policy and widespread practice across countries (Greenberg, 2010; Shonkoff and Bales, 2011). Clearly, there is a need for greater efforts to translate science for practice and policy so that SEL approaches can be better integrated into schools and communities. Such efforts can help build the processes and structures needed to foster high-quality implementation and promote sustainability (see Elias et al., 2003).

Finally, greater collaboration between researchers and educators is needed to allow research to not only inform practice, but also that practice can inform research. To create a world characterized by the values and practices that illustrate caring and kindness among all people, it is essential that educators, parents, community members, and policy makers work in concert to achieve long-term change. It is therefore critical that we make intentional efforts to devise the most effective educational practices that promote SEL in all children. Such efforts must be based on strong conceptual models and sound research. Only then will we be in a position to advance the development of our world's children and youth.

Acknowledgement

To the life and legacy of Roger P. Weissberg, for his leadership, collaboration, inspiration, and endless tenacity in advancing SEL research, practice, and policy to promote the well-being of millions of school children.

References

- Addams, J., 2002. *Democracy and Social Ethics*. University of Illinois Press.
- Baltes, P.B., Reuter-Lorenz, P.A., Rosler, F. (Eds.), 2006. *Lifespan Development and the Brain: The Perspective of Biocultural Co-constructivism*. Cambridge University Press, New York, NY.
- Belfield, C., Bowden, B., Klapp, A., Levin, H., Shand, R., Zander, S., 2015. *The Economic Value of Social and Emotional Learning*. New York, NY.
- Benn, R., Akiva, T., Arel, S., Roeser, R.W., 2012. Mindfulness training effects for parents and educators of children with special needs. *Dev. Psychol.* 48 (5), 1476–1487.
- Berg, J., Osher, D., Same, M.R., Nolan, E., Benson, D., Jacobs, N., 2017. Identifying, Defining, and Measuring Social and Emotional Competencies - Final Report. American Institutes for Research, Washington, DC. Retrieved from: <https://www.air.org/resource/identifying-defining-and-measuring-social-and-emotional-competencies>.
- Best, J.R., Miller, P.H., 2010. A developmental perspective on executive function. *Child Dev.* 81 (6), 1641–1660.
- Blyth, D.A., Borowski, T., Farrington, C.A., Kyllonen, P., Weissberg, R.P., 2019. Ten Criteria for Describing and Selecting SEL Frameworks. Collaborative for Academic, Social, and Emotional Learning, Chicago, IL.
- Braun, S.S., Schonert-Reichl, K.A., Roeser, R.W., 2020. Effects of teachers' emotion regulation, burnout, and life satisfaction on student well-being. *J. Appl. Dev. Psychol.* 69, 101–151.
- Bridgeland, J., Bruce, M., Hariharan, A., 2013. *The Missing Piece: A National Teacher Survey on How Social and Emotional Learning Can Empower Children and Transform Schools*. Civic Enterprises with Peter D. Hart Research Associates, Washington, DC.
- Bronfenbrenner, U., Morris, P., 2006. The bioecological model of human development. In: Damon, W., Lerner, M. (Eds.), *Handbook of Child Psychology: Vol. 1. Theoretical Models of Human Development*, sixth ed. Wiley, Hoboken, NJ, pp. 793–828.
- Bushaw, W.J., Lopez, S.J., 2013. Which way do we go? *Phi Delta Kappan* 95 (1), 9–25.
- Cantor, P., Osher, D., Berg, J., Steyer, L., Rose, T., 2018. Malleability, plasticity, and individuality: how children learn and develop in context. *Appl. Dev. Sci.* 24 (4), 307–337.
- Caprara, G.V., Barbaranelli, C., Pastorelli, C., Bandura, A., Zimbardo, P.G., 2000. Prosocial foundations of children's academic achievement. *Psychol. Sci.* 11 (4), 302–306.
- CASEL, 2013. 2013 CASEL Guide: Effective Social and Emotional Learning Programs – Preschool and Elementary Edition. <https://casel.org/preschool-and-elementary-edition-casel-guide/>.
- Collie, R.J., Shapka, J.D., Perry, N.E., Martin, A.J., 2015. Teacher well-being: exploring its components and a practice-oriented scale. *J. Psychoeduc. Assess.* 33 (8), 744–756.
- Cook, C.R., Fiat, A., Larson, M., Daikos, C., Slemrod, T., Holland, E.A., Thayer, A.J., Renshaw, T., 2018. Positive greetings at the door: evaluation of a low-cost, high-yield proactive classroom management strategy. *J. Posit. Behav. Interv.* 20 (3), 149–159.
- Darling-Hammond, L., Cook-Harvey, C.M., 2018. *Educating the Whole Child: Improving School Climate to Support Student Success*. Learning Policy Institute, Palo Alto, CA.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. Free Press, New York, NY.
- Diamond, A., 2012. Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21 (5), 335–341.
- Domitrovich, C.E., Durlak, J., Staley, K.C., Weissberg, R.P., 2017. Social-emotional competence: an essential factor for promoting positive adjustment and reducing risk and school children. *Child Dev.* 88 (2), 408–416.
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., 2011. Enhancing students' social and emotional development promotes success in school: results of a meta-analysis. *Child Dev.* 82, 405–432.
- Dwyer, K., Osher, D., 2000. *Safeguarding Our Children: An Action Guide*. U.S. Departments of Education and Justice and American Institutes for Research, Washington, DC.
- Dwyer, K., Osher, D., Warger, C., 1998. *Early Warning, Timely Response: A Guide to Safe Schools*. U.S. Department of Education, Washington, DC.
- Edge Research, 2018. *Developing Life Skills in Children: A Roadmap for Communicating with Parents*. Learning Heroes, Alexandria, VA.
- Elias, M.J., Zins, J.E., Weissberg, R.P., Frey, K.S., Greenberg, M.T., Haynes, N.M., Shriver, T.P., 1997. *Promoting Social and Emotional Learning: Guidelines for Educators*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Finn, C.E., Hess, F.M., April 2019. *What Social and Emotional Learning Needs to Succeed and Survive*. American Enterprise Institute, Washington, DC.
- Frydenberg, E., Martin, A.J., Collie, R.J. (Eds.), 2017. *Social and Emotional Learning in Australia and the Asia-Pacific. Perspectives, Programs and Approaches*. Springer Social Sciences.
- Gademmann, A., Thomson, K., Gill, R., Schonert-Reichl, K.A., Gagne Petteni, M., Guhn, M., Warren, M., Oberle, E., 2022. Early adolescents' experiences during the COVID-19 Pandemic and changes in their well-being. *Front. Public Health* 10, 823303. <https://doi.org/10.3389/fpubh.2022.823303>.
- Gallup, P., 2014. *State of American Schools*.
- Gardner, H., 1983. *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books, New York.
- Gottfredson, G.D., Gottfredson, D.C., Payne, A.A., Gottfredson, N.C., 2005. School climate predictors of school disorder: results from a national study of delinquency prevention in schools. *J. Res. Crime Delinquen.* 42 (4), 412–444.
- Goleman, D., 2006. *Social Intelligence: The New Science of Human Relationships*. Bantam Dell, A Division of Random House, New York, NY.

- Greenberg, M.T., Weissberg, R.P., Utne O'Brien, M., Zins, J.E., Fredericks, L., Resnik, H., et al., 2003. Enhancing school-based prevention and youth development through coordinated social, emotional, and academic-learning. *Am. Psychol.* 58, 466–474.
- Greenberg, M.T., 2010. School-based prevention: current status and future challenges. *Effect Educ.* 2 (1), 27–52.
- Gregory, A., Fergus, E., 2017. Social-emotional learning and equity in school discipline. *Future Child.* 27, 117–136.
- Hamilton, L.S., Schwartz, H.L., August 2019. Get Smart About Social and Emotional Learning Measurement. American Enterprise Institute, Washington, DC.
- Hamilton, S., Doss, C.J., Steiner, E.D., 2019. Teacher and Principal Perspectives on Social and Emotional Learning in America's Schools: Findings from the American Educator Panels. RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR2991.html.
- Hawkins, J.D., Kosterman, R., Catalano, R.F., Hill, K.G., Abbott, R.D., 2008. Effects of a social development intervention in childhood 15 years later. *Arch. Pediatr. Adolesc. Med.* 162, 1133–1141.
- Hemmelgarn, A.L., Glisson, C., James, L.R., 2006. Organizational culture and climate: implications for services and interventions research. *Clin. Psychol. Sci. Pract.* 13 (1), 73–89.
- Herman, K.C., Hickmon-Rosa, J.E., Reinke, W.M., 2018. Empirically derived profiles of teacher stress, burnout, self-efficacy, and coping and associated student outcomes. *J. Posit. Behav. Interv.* 20 (2), 90–100.
- Hess, R., 2019. Foreword: What It'll Take for Social and Emotional Learning to Succeed and Survive. American Enterprise Institute, Washington, DC.
- Hoglund, W.L., Klinge, K.E., Hosan, N.E., 2015. Classroom risks and resources: teacher burnout, classroom quality and children's adjustment in high needs elementary schools. *J. Sch. Psychol.* 53 (5), 337–357.
- Humphrey, N. (Ed.), 2013. *Social and Emotional Learning: A Critical Appraisal*. SAGE Publications Limited.
- Hymel, S., Swearer, S.M., 2015. Four decades of research on school bullying: an introduction. *Am. Psychol.* 70 (4), 293–299.
- Immordino-Yang, M.H., Damasio, A., 2007. We feel, therefore we learn: the relevance of affective and social neuroscience to education. *Mind Brain Educ.* 1, 3–10.
- Immordino-Yang, M.H., Krone, C., Darling-Hammond, L., 2019. Nurturing nature: how brain development is inherently social and emotional, and what this means for education. *Educ. Psychol.* 54, 185–204.
- Jagers, R.J., Rivas-Drake, D., Williams, B., 2019. Transformative social and emotional learning (SEL): toward SEL in the service of equity and excellence. *Educ. Psychol.* 54 (3), 162–184.
- Jennings, P.A., Frank, J.L., 2015. Inservice preparation for educators. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford Press, New York, pp. 422–437.
- Jennings, P.A., Greenberg, M.T., 2009. The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev. Educ. Res.* 79 (1), 491–525.
- Jennings, P.A., Snowberg, K.E., Coccia, M.A., Greenberg, M.T., 2011. Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): results of two pilot studies. *J. Classr. Interact.* 46 (4), 37–48.
- Jessop, D.S., Turner-Cobb, J.M., 2008. Measurement and meaning of salivary cortisol: a focus on health and disease in children. *Stress* 11 (1), 1–14.
- Jones, S.M., Kahn, J., 2017. The Evidence Base for How We Learn: Supporting Students' Social, Emotional, and Academic Development – Consensus Statements of Evidence from the Council of Distinguished Scientists: National Commission on Social, Emotional, and Academic Development. The Aspen Institute, Washington, DC.
- Jones, D.E., Greenberg, M., Crowley, M., 2015. Early social-emotional functioning and public health: the relationship between kindergarten social competence and future wellness. *Am. J. Publ. Health* 105, 2283–2290.
- Jones, S.M., Bouffard, S.M., Weissbourd, R., 2013. Educators' social and emotional skills vital to learning. *Phi Delta Kappan* 94 (8), 62–65.
- Jones, S.E., Ethier, K.A., Hertz, M., et al., 2022. Mental health, suicidality, and connectedness among high school students during the COVID-19 pandemic—adolescent behaviors and experiences survey, United States, January–June 2021. *MMWR Suppl.* 71 (Suppl. 3), 16–21. <https://doi.org/10.15585/mmwr.su7103a3external.icon>.
- Kabat-Zinn, J., 1990. *Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness*. Bantam Doubleday Dell, New York, NY.
- Katz, D.A., Greenberg, M.T., Jennings, P.A., Klein, L.C., 2016. Associations between the awakening responses of salivary α -amylase and cortisol with self-report indicators of health and wellbeing among educators. *Teach. Teach. Educ.* 54, 98–106.
- Kyriacou, C., 2011. Teacher stress: from prevalence to resilience. In: Langan-Fox, J., Cooper, C.L. (Eds.), *Handbook of Stress in the Occupations*. Edward Elgar Publishing, pp. 161–176.
- Lever, N., Mathis, E., Mayworm, A., 2017. School mental health is not just for students: why teacher and school staff wellness matters. *Rep. Emot. Behav. Disord. Youth* 17 (1), 6–12.
- Luthar, S.S., Ebbert, A.M., Kumar, N.L., 2021. Risk and resilience during COVID-19: a new study in the Zigler paradigm of developmental science. *Dev. Psychopathol.* 33 (2), 565–580. <https://doi.org/10.1017/S0954579420001388>.
- Mahoney, J., Weissberg, R., Greenberg, M., Dusenbury, L., Jagers, R., Niemi, K., Schlinger, M., Schlund, J., Shriver, T., VanAusdal, K., Yoder, N., 2020. Systemic social and emotional learning: promoting educational success for all preschool to high school students. *Am. Psychol.* 1–16.
- Maldonado-Carreno, C., Votruba-Drzal, E., 2011. Teacher-child relationships and the development of academic and behavioral skills during elementary school: a within-and-between-child analysis. *Child Dev.* 82 (2), 601–616.
- Marzano, R.J., Marzano, J.S., Pickering, D.J., 2003. *Classroom Management that Works*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Maslach, C., Jackson, S.E., Leiter, M.P., 1996. *Maslach Burnout Inventory Manual*, third ed. Consulting Psychologist Press.
- Maslach, C., Schaufeli, W.B., Leiter, M.P., 2001. Job burnout. *Annu. Rev. Psychol.* 52, 397–422.
- McShane, M.Q., May 2019. What Social and Emotional Learning Advocates Can Learn from Common Core. American Enterprise Institute, Washington, DC.
- Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., et al., 2011. A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U. S. A.* 108 (7), 2693–2698.
- Montessori, M., 1912. *The Montessori Method: Scientific Pedagogy as Applied to Child Education in "The Children's House," With Additions and Revisions by the Author*. Translated from the Italian by Anne E. George; with an Introduction by Professor Henry W. Holmes. FA Stokes.
- Montgomery, C., Rupp, A.A., 2005. A meta-analysis for exploring the diverse causes and effects of stress in teachers. *Can. J. Educ.* 28 (3), 458–486.
- Morrison, B., 2013. Restorative justice in schools. In: *New Directions in Restorative Justice*. Willan, pp. 50–76.
- National Commission on Social, Emotional, and Academic Development, 2018. *From a Nation at Risk to a Nation of Hope: Recommendations From the National Commission on Social, Emotional, & Academic Development*. The Aspen Institute, Washington, DC.
- National School Climate Council, 2007. *The School Climate Challenge: Narrowing the Gap Between School Climate Research and School Climate Policy, Practice Guidelines and Teacher Education Policy*. Retrieved from. <http://www.schoolclimate.org/climate/advocacy.php>.
- Oberle, E., Schonert-Reichl, K.A., 2016. Stress contagion in the classroom? The link between classroom teachers' burnout and morning cortisol in elementary school students. *Soc. Sci. Med.* 159, 30–37.
- Oberle, E., Schonert-Reichl, K.A., Hertzman, C., Zumbo, B., 2014. Social-emotional competencies make the grade: predicting academic success in early adolescence. *J. Appl. Dev. Psychol.* 35, 138–147.
- Organization for Economic Co-operation and Development (OECD), 2018. *The Future of Education and Skills: Education 2030*. E2030 Position Paper. OECD Publishing, Paris.
- Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., Weissberg, R.P., 2016. Advancing the science and practice of social and emotional learning: looking back and moving forward. *Rev. Res. Educ.* 40 (1), 644–681.
- Pascual-Leone, A., Amedi, A., Fregni, F., Merabet, L.B., 2005. The plastic human brain cortex. *Annu. Rev. Neurosci.* 28, 377–401.
- PDK International, 2017. *The 49th Annual PDK Poll of the Public's Attitudes toward the Public Schools*. PDK International, Arlington, VA.

- Roeser, R.W., Eccles, J.S., 2014. School and the mental health of children and adolescents: a developmental systems approach. In: Lewis, M., Rudolph, K. (Eds.), *Handbook of Developmental Psychopathology*, third ed. Springer, pp. 163–184.
- Roeser, R., Schonert-Reichl, K.A., Jha, A., Wilensky, R., Oberle, E., Thomson, K., Taylor, C., Harrison, J., 2013. Mindfulness training and reductions in teacher stress and burnout: results from two randomized, waitlist-control field trials. *J. Educ. Psychol.* 105, 787–804.
- Salovey, P., Mayer, J.D., 1990. Emotional intelligence. *Imagin. Cognit. Pers.* 9 (3), 185–211.
- Saue, J.A., Schonert-Reichl, K.A., 2019. Creating caring classroom and school communities: lessons learned from social and emotional learning programs and practices. In: Fredricks, J., Reschly, A.L., Christenson, S.L. (Eds.), *The Handbook of Student Interventions: Working with Disengaged Youth*. Elsevier, New York, NY.
- Schonert-Reichl, K.A., 2017. Social and emotional learning and teachers. *Future Child*. 27, 137–155.
- Schonert-Reichl, K.A., Weissberg, R.P., 2015. Social and emotional learning during childhood. In: Gullotta, T.P., Bloom, M. (Eds.), *Encyclopedia of Primary Prevention and Health Promotion*, second ed. Springer Press, New York, NY, pp. 936–949.
- Schwarzer, R., Hallum, S., 2008. Perceived teacher self-efficacy as a predictor of job stress and burnout: mediation analysis. *Appl. Psychol.* 57 (s1), 152–171.
- Shonkoff, J.P., Bales, S.N., 2011. Science does not speak for itself: translating child development research for the public and its policymakers. *Child Dev.* 82 (1), 17–32.
- Shriver, T.P., Weissberg, R.P., 2020. A response to constructive criticism of social and emotional learning. *Phi Delta Kappan* 101 (7), 52–57.
- Sklad, M., Diekstra, R., De Ritter, M., Ben, J., 2012. Effectiveness of school-based universal social, emotional, and behavioural programs: do they enhance students' development in the area of skill, behaviour, and adjustment? *Psychol. Sch.* 49, 892–909.
- Sparks, D., Malkus, N., 2015. Public School Teacher Autonomy in the Classroom Across School Years 2003–2004, 2007–2008, 2011–2012. *Stats in Brief* (NCES 2015-089). US Dept of Ed. National Center for Education Statistics.
- Spilt, J.L., Koomen, H.M., Thijs, J.T., 2011. Teacher wellbeing: the importance of teacher–student relationships. *Educ. Psychol. Rev.* 23 (4), 457–477.
- Starr, J.P., March 21, 2019. Can we keep SEL on course? *Phi Delta Kappan* 100 (8), 70–71.
- Sternberg, J.P., 1985. *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge University Press, New York, NY.
- Taylor, R.D., Oberle, E., Durlak, J.A., Weissberg, R.P., 2017. Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. *Child Dev.* 88, 1156–1171. <https://doi.org/10.1111/cdev.12864>.
- Thapa, A., Cohen, J., Guffey, S., Higgins-D'Alessandro, A., 2013. A review of school climate research. *Rev. Educ. Res.* 83, 357–385.
- Thorndike, E.L., 1920. Intelligence and its uses. *Harpers Mag.* 140, 227–235.
- Torrente, C., Alimchandani, A., Aber, J.L., 2015. International perspectives on social and emotional learning. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford Publications, New York, NY, pp. 566–587.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services, 2021. Supporting Child and Student Social, Emotional, Behavioral, and Mental Health Needs. Washington, DC. Retrieved from: <https://www2.ed.gov/documents/students/supporting-child-student-social-emotional-behavioral-mental-health.pdf>.
- Varela, A.D., Kelcey, J., Reyes, J., Gould, M., Sklar, J., 2013. Learning and Resilience: The Crucial Role of Social and Emotional Well-Being in Contexts of Adversity. Education Notes. World Bank Group, Washington, DC.
- Walker, T., April 15, 2020. Social-emotional learning should be priority during COVID-19 crisis. NEA Today. Retrieved from: <http://neatoday.org/2020/04/15/social-emotional-learning-during-covid/>.
- Weare, K., Nind, M., 2011. Mental health promotion and problem prevention in schools: what does the evidence say? *Health Promot. Int.* 26 (1), i29–i69.
- Wearmouth, J., McKinney, R., Glynn, T., 2007. Restorative justice: two examples from New Zealand schools. *Br. J. Spec. Educ.* 34 (4), 196–203.
- Weissberg, R.P., Durlak, J.A., Domitrovich, C.E., Durlak, J.A., 2015. Social and emotional learning: past, present, and future. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford, pp. 3–19.
- Weissberg, R.P., 2019. Promoting the social and emotional learning of millions of school children. *Perspect. Psychol. Sci.* 14 (1), 65–69.
- Weissberg, R.P., Durlak, J.A., Domitrovich, C.E., Gullotta, T.P., 2015. Social and emotional learning: past, present, and future. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford, New York, NY, pp. 3–19.
- Wentzel, K.R., 1993. Does being good make the grade? Social behavior and academic competence in middle school. *J. Educ. Psychol.* 85 (2), 357–364.
- Wolfram, M., Bellingrath, S., Feuerhahn, N., Kudielka, B.M., 2013. Emotional exhaustion and overcommitment to work are differentially associated with hypothalamus-pituitary-adrenal (HPA) axis responses to a low-dose ACTH1–24 (Synacthen) and dexamethasone–CRH test in healthy school teachers. *Stress* 16 (1), 54–64.
- Zins, J.E., Weissberg, R.P., Wang, M.C., Walberg, H.J., 2004. *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* Teachers College Press, NY.

Further reading

Books

- Atwell, M.N., Bridgeland, J.M., 2019. *Ready to Lead: A 2019 Update of Principals' Perspectives on How Social and Emotional Learning Can Prepare Children and Transform Schools*. Civic with Hart Research Associates, Washington, DC.
- CASEL, 2015. 2015 CASEL Guide: Effective Social and Emotional Learning Programs – Middle and High School Edition. <https://casel.org/middle-and-high-school-edition-casel-guide/>.
- CASEL, 2019a. The CASEL Guide to Schoolwide Social and Emotional Learning. <https://schoolguide.casel.org/>.
- CASEL, 2019b. CASEL's District Resource Center. <https://drc.casel.org/>.
- CASEL, 2019c. Collaborating States Initiative Resources. <https://casel.org/csi-resources/>.
- Durlak, J.A., DuPre, E.P., 2008. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. <https://doi.org/10.1007/s10464-008-9165-0>.
- Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), 2015. *Handbook of Social and Emotional Learning: Research and Practice*. Guilford Press, New York.
- Elias, M.J., Zins, J.E., Weissberg, R.P., Frey, K.S., Greenberg, M.T., Haynes, N.M., Kessler, R., Schwab-Stone, M.E., Shriver, T.P., 1997. *Promoting Social and Emotional Learning: Guidelines for Educators*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Goleman, D., 2005. *Emotional Intelligence: Why it Can Matter More than IQ*. Bantam, New York.
- Greenberg, M.T., Brown, J.L., Abenavoli, R.M., 2016. *Teacher Stress and Health Effects on Teachers, Students, and Schools*. Edna Bennett Pierce Prevention Research Center, Pennsylvania State University. Robert Wood Johnson Foundation Brief.
- Greenberg, M.T., Domitrovich, C.E., Weissberg, R.P., Durlak, J.A., 2017. Social and emotional learning as a public health approach. *Future Child*. 27 (1), 13–32.
- Jagers, R., Rivas-Drake, D., Borowski, T., 2018. *Equity & Social and Emotional Learning: A Cultural Analysis*. CASEL, Chicago, IL.
- Jennings, P.A., Brown, J.L., Frank, J.L., Doyle, S., Oh, Y., Davis, R., Rasheed, D., DeWeese, A., DeMauro, A.A., Cham, H., Greenberg, M.T., 2017. Impacts of the CARE for Teachers program on teachers' social and emotional competence and classroom interactions. *J. Educ. Psychol.* 109 (7), 1010–1028.
- Lawlor, M.S., 2016. Mindfulness and social and emotional learning (SEL): a conceptual framework. In: Schonert-Reichl, K.A., Roeser, R.W. (Eds.), *Mindfulness in Education: Integrating Theory and into Practice*. Springer, New York, NY, pp. 3–16.
- Mahoney, J.L., Durlak, J.A., Weissberg, R.P., 2018. An update on social and emotional' learning outcome research. *Phi Delta Kappan* 100 (4), 18–23.

- Meherali, S., Punjani, N., Louie-Poon, S., Abdul Rahim, K., Das, J.K., Salam, R.A., Lassi, Z.S., 2021. Mental health of children and adolescents amidst COVID-19 and past pandemics: a rapid systematic review. *Int. J. Environ. Res. Publ. Health* 18 (7), 3432. <https://doi.org/10.3390/ijerph18073432>.
- Oliveira, S., Roberto, M.S., Salgado Pereira, N., Marques-Pinto, A., Veiga-Simão, A.M., 2021. Impacts of social and emotional learning interventions for teachers on teachers' outcomes: a systematic review with meta-analysis. *Front. Psychol.* 12, 677217. <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.677217/full>.
- Rimm-Kaufman, S.E., 2020. SEL from the Start: Building Skills in K-5 (Social and Emotional Learning Solutions). Norton.
- Roeser, R.W., Galla, B.M., Baelen, R.N., 2020. Mindfulness in Schools: Evidence on the Impact of School-Based Mindfulness Programs on Student Outcomes in P-12 Educational Settings. Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University, University Park, PA.
- Schonert-Reichl, K.A., Lawlor, M.S., 2010. The effects of a mindfulness-based education program on pre- and early adolescents' well-being and social and emotional competence. *Mindfulness* 1, 137–151.
- Schonert-Reichl, K.A., Oberle, E., Lawlor, M.S., Abbott, D., Thomson, K., Oberlander, T., Diamond, A., 2015. Enhancing cognitive and social-emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a Randomized Controlled Trial. *Dev. Psychol.* 51, 52–66.
- Schreier, H., Schonert-Reichl, K.A., Chen, E., 2013. Effect of volunteering on risk factors for cardiovascular disease in adolescents: a Randomized Controlled Trial. *JAMA Pediatr.* 167, 327–332.
- Wiglesworth, M., Lendrum, A., Oldfield, J., Scott, A., ten Bokkel, I., Tate, K., Emery, C., 2016. The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: a meta-analysis. *Camb. J. Educ.* 46, 347–376.

Reports

- Simmons, D.N., Brackett, M.A., Adler, N., 2018. Applying an Equity Lens to Social, Emotional, and Academic Development. Edna Bennett Pierce Prevention Research Center, Pennsylvania State University. Robert Wood Johnson Foundation Brief on SEL and Equity. On-line: <https://www.rwjf.org/en/library/research/2018/06/applying-an-equity-lens-to-social-emotional-and-academic-development.html>.

Articles

- Jones, S.M., Bailey, R., Jacob, R., 2014. Social-emotional learning is essential to classroom management. *Phi. Delta Kappan* 96 (2), 19–24.
- Jones, S.M., Doolittle, E.J., 2017. Social and emotional learning. *Future Child.* 27 (1). Princeton, NJ: Princeton-Brookings. www.futureofchildren.org.
- Jones, S., Bailey, R., Brush, K., Kahn, K., 2017. Kernels of Practice for SEL: Low-Cost, Low-Burden Strategies. EASEL Lab, Harvard Graduate School of Education. <https://www.wallacefoundation.org/knowledge-center/pages/kernels-of-practice-for-sel-low-cost-low-burden-strategies.aspx>.
- Immordino-Yang, M.H., Darling-Hammond, L., Krone, C., 2019. Nurturing Nature: How Brain Development Is Inherently Social and Emotional, and what This Means for Education (EdWorkingPaper: 19-106). Retrieved from Annenberg Institute at Brown University: <http://www.edworkingpapers.com/ai19-106>.
- Schonert-Reichl, K.A., 2020. Social and emotional learning: definitions, frameworks, and evidence. *Rethinking Learning: Mainstreaming Social and Emotional Learning in Education Systems* (Chapter 3). UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development (MGIEP), India. Available free on-line at: <https://mgiep.unesco.org/rethinking-learning>.
- Schonert-Reichl, K.A., Kitil, M.J., Sauve, J., Trach, J., 2020. Monitoring and evaluating social and emotional learning. In: *Rethinking Learning: Mainstreaming Social and Emotional Learning in Education Systems* (Chapter 4). UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development (MGIEP), India. Available free on-line at: <https://mgiep.unesco.org/rethinking-learning>.
- Weissberg, R.P., Cascarino, J., 2013. Academic learning + social-emotional learning = national priority. *Phi. Delta Kappan* 95 (2), 8–13.

Fundamental processes of learning: an overview

Jennifer Wiley and Tricia A. Guerrero, Department of Psychology, University of Illinois at Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Memory processes	108
The modal model	108
Sensory memory	108
Short-term memory	108
Immediate memory and working memory capacity	108
Long-term memory	109
Conditions that improve memory	109
Practice effects in memory and skill acquisition	109
Skill acquisition in language, reading, and math	110
Testing effects	110
Spaced practice effects	110
Generation effects	110
Semantic processing effects	111
Comprehension processes	111
Conditions that improve comprehension	111
Goal-directed or purposeful reading	112
Learning from questions and explanations	112
Learning from multiple sources	112
Learning from multiple representations	112
Conceptual change and restructuring processes	113
Other conditions that improve learning	113
Active and interactive learning with others	113
Learning by invention, learning from errors, preparation for future learning	113
Importance of feedback and guidance	114
Metacognitive monitoring and self-regulated learning	114
Conclusion	114
References	115

The fundamental processes that underlie learning can be seen as the processes that allow us to retain, update, and use information in new contexts. From a cognitive perspective, memory processes are essential components of learning that are constrained by attention and perception, and provide the basis for all higher-order cognitive functions as we recover and transform information to engage in acts of communication, comprehension, reasoning, and problem solving. In addition, learning can also be considered as the development or acquisition of skills, expertise, knowledge, or conceptual understanding.

Fueled by post World War II interests in advancing science and technology, the field of Artificial Intelligence flourished in the 1950s and 1960s and played a crucial role in the development of a cognitive approach to the understanding of thought that was built on an *information processing* framework (Broadbent, 1958). In this approach, the computer served as a powerful metaphor for the human mind. Broadbent is widely credited with developing the computer metaphor into an information processing model of human cognition with the guiding principles that the mind could be described in terms of input, internal representation or storage of information, processing, and output. Much of his work focused on exploring the limited capacity channels that constrain the input functions.

With the publication of *Cognitive Psychology* (1967), Neisser offered an initial definition of cognition as referring to all of the mental processes by which sensory input is transformed, reduced, elaborated, stored, recovered, and used. At the center of all these cognitive mental processes is memory. Memory is commonly thought of in terms of the processes that are responsible for acquiring information (*encoding*), retaining information (*storage*), and accessing stored information (*retrieval*). Just as memory serves as the fundamental basis for all cognitive processing, memory processes are also fundamental for learning. In many ways, all learning can be seen as the result of memory processes. All of the information and associations within our long-term memory store represent products of learning, and this memory store provides the basis for all other cognitive functions.

The goal for this chapter is to provide a broad overview of the fundamental cognitive processes involved in learning. Starting with early work using an information processing approach to explore different memory systems and processes, the chapter discusses

the critical constraints on learning that come from limitations in working memory capacity, as well as conditions that generally support better memory for information. In addition, a critical distinction is made between memory and comprehension outcomes, and an overview of conditions that support better understanding is provided. Learning is also considered as involving the development of skills, the acquisition of expertise, and engagement in knowledge revision or conceptual change. Finally, several other areas of research are highlighted that explore successful learning as part of a rich interplay between cognitive and metacognitive processes as well as those that allow the learner to pursue a deeper understanding of content including via interaction with others.

Memory processes

Given its prominence, a great deal of interest has been directed toward a better understanding of human memory. It is tempting to think of memory as simply a storage location that serves as the repository of all of our knowledge. However, information does not become a part of our memory via a singular path or process. Instead, there are a number of different types of processes associated with different memory functions. Although various taxonomic approaches have attempted to identify distinct types of memory stores and mechanisms (Atkinson and Shiffrin, 1968; Squire, 1987; Tulving, 1985), these are best seen as different perspectives on how we develop and use a rich, complicated interconnected network of associations. Nevertheless, it is helpful to consider some of these historical distinctions as they help to illustrate the factors that underlie learning and the many ways that learning can take place.

The modal model

Following from the information processing approach, one of the earliest models of memory was developed by Waugh and Norman (1965) and Atkinson and Shiffrin (1968), and is generally referred to as the *modal model*. The *modal model* of memory emphasized memory function in terms of how information is perceived and then moved into storage. The *modal model* attempted to capture the variability in the amount of time that information is available. It suggested distinctions among stores (or different *buffers*) that were defined by the duration that information is retained.

Sensory memory

An initial memory buffer, the *sensory memory store*, holds a large amount of incoming sensory information for only a few hundred milliseconds. This form of memory captures inputs in relatively unprocessed sensory form—iconic memory for visual inputs and echoic memory for auditory inputs. Although visual and auditory sensory memory have received the most research attention, there is thought to be a unique sensory store for each of the five senses. In general, any traces of sensory input that are represented in these stores are lost very quickly. Processes of selection and representation move information into the *short-term memory store*. Some of this information is then transferred into the much larger and more permanent *long-term memory store*. As information moves through these buffers or stages, perceptual and attentional processes serve as gatekeepers and determine what remains available for further cognitive processing.

Short-term memory

The *short-term memory store*, which receives input from sensory memory, holds by recent estimates only about four independent units of information for a few seconds (Cowan, 2014). Initial estimates based on Miller (1956) were that the capacity of this buffer might be closer to 7 units, plus or minus 2. But updated research and theorizing has suggested these were over-estimates that included some information that could be “chunked” or formed into meaningful units. In order for information to be maintained in short-term memory, it needs to be rehearsed, or it too will quickly be lost. Some of this information is then transferred into *long-term memory*. In contrast to sensory or short-term memory, long-term memory is characterized not just by its longer duration but also by its vast capacity. Yet, the ability to retrieve information at a later time depends on many factors, including how it is processed during encoding.

Immediate memory and working memory capacity

Modern approaches have advanced beyond the early *modal model* especially with respect to the notion of the short-term memory store. Current theories now tend to emphasize that short-term memory is not a storage place, buffer, or staging area outside of the long-term memory store, as much as it may represent information that is immediately available, currently active, currently being thought about, or information that one is currently processing or working on.

Baddeley's (2000) model of working memory characterizes *immediate memory* as a function of three distinct short-term memory subsystems, commanded by a *central executive*. One subsystem processes and stores visuospatial information, a second deals with phonological information (including speech), and a third integrates information from different modalities and temporal information into an episodic representation. Many studies using dual-task approaches have shown that sometimes we can do two tasks at once without a decrement in performance or without involving the central executive, specifically if a secondary task uses a different

subsystem than a primary task (Baddeley and Hitch, 1974). In contrast, sometimes attempting to do two tasks at once can cause performance to suffer, require involvement of the central executive, and put a *load* on working memory. Yet, other studies suggest that some single tasks may be so demanding that they load the central executive even without a secondary task.

An alternative approach characterizes *working memory capacity* from the perspective of individual differences in the general ability to control one's attention, or the ability to focus on a goal in the face of interference (Engle, 2002). This research has traditionally found that individual differences in attentional control affect the amount of information that people can consider in immediate memory, which in turn impacts their performance on many cognitive tasks. Regardless of whether a short-term or working memory approach is used to describe immediate memory, it is clear that there are severe processing limitations that exist. The ability to hold information in the focus of attention (or keep it active in working memory) is an important bottleneck that constrains human processing capacity (Cowan, 2014). These individual differences in capacity are generally found to influence how effectively or efficiently information may be encoded in long-term memory, and therefore often serve as strong predictors of learning outcomes and the ability to use information at a later time.

Long-term memory

Beyond the attentional bottleneck which constrains *whether* information may be encoded, the ability to use information at a later time also depends on *how* it was encoded. Thus, one set of distinctions among memory processes considers the way in which learning occurs. One main contrast is whether an individual intentionally attempts to learn information (*intentional learning*) or if the learning happens more incidentally by exposure (*incidental learning*, Postman, 1964). Learning can occur without explicit attention, as individuals are sensitive to patterns in their environments and can learn them implicitly (*implicit learning*, *statistical learning*). In contrast, when students are tasked with specific learning goals and they engage in an explicit attempt to memorize or learn information, this can be seen as intentional learning.

Other distinctions in long-term memory have been proposed according to content. Squire (1987) proposed a distinction between declarative and non-declarative memories, with *declarative memory* representing information that could be consciously accessed and declared (verbalized or stated), whereas non-declarative memory was comprised of non-conscious memories that are expressed through performance, such as *procedural memory* for actions or skills. A contrast that is typically used to illustrate this distinction is the difference between knowing the name of an object versus how you use it (what a bicycle is versus how to ride it). Similarly, in terms of language learning, vocabulary knowledge is often seen as declarative, whereas syntax is often seen as based in procedural memory.

Another common distinction that has been made is between episodic and semantic memory (Tulving, 1985). *Episodic memory* is thought to capture information from specific moments in time (autobiographical memory, memory for a specific event or episode). In contrast, *semantic memory* captures specific as well as abstracted, generalized concepts. Whereas episodic memory may be indexed by temporal context, semantic memory is thought of as a network of associations based in meaning. These different processes govern what information is encoded, how information is encoded, and whether and how information can be retrieved from long-term memory for use in higher-order processes such as communication, comprehension, reasoning, and problem solving.

Conditions that improve memory

Although all learners experience a bottleneck from the limited capacity of immediate memory, there are a variety of contexts or conditions that increase the likelihood that information will be remembered. These largely depend on two main mechanisms: repetition and meaningful associations. Engaging in practice for new associations, and integrating new associations with prior knowledge, are two ways that individuals can circumvent the limitations of the attentional bottleneck. Our ability to process information benefits from experience. Learners benefit from repetition, practice, and the automaticity that results from repeated experience with stimuli. Learners can also benefit from expectations based in prior knowledge and associations from long-term memory.

Practice effects in memory and skill acquisition

The extent to which information must be practiced, whether it must be explicitly processed, and how often information needs to be attended to during learning are important and well-studied topics (Pashler, 1998). Repeated exposure to information is one route to improving its likelihood of retrieval. Early in acquisition, novel skills or stimuli demand a great deal of attention. With repeated practice, processing becomes more automatic and requires less effort. Repetition benefits are often seen as *priming effects* (the amount of time it takes to recognize a stimulus decreases with repeated exposure). Benefits are also seen as *savings* during relearning attempts (it takes less time to re-learn the same list of words). Another typical finding is that the benefits from practice follow a *learning curve* with rapid improvement after initial practice followed by ever lesser improvements with further practice. Such negatively accelerated learning curves generally mimic power functions, thus, learning is often said to follow the *power law of practice*. This is typically the trajectory that is followed during skill acquisition (like learning motor sequences or procedures) as skills become more automated or proceduralized. Some examples of the benefits that are seen due to practice or repeated exposure include faster reading times for more common words (*word frequency effects*) and faster computation times for single digit number facts (e.g., $2 + 3$).

Skill acquisition in language, reading, and math

The fundamental processes involved in learning in several domains can also be considered from the perspective of developing skills over time. First-language acquisition is perhaps the most impressive of all cognitive feats. Language develops universally in some form in all humans, and development typically proceeds in a remarkably rapid fashion. Modern approaches emphasize both the biological predisposition to acquire language as well as the role of exposure to patterns in the environment (Gleitman and Newport, 1995). In contrast to first language acquisition, adult second language (L2) acquisition occurs under more variable learning conditions. Second language acquisition is typically assumed to occur in a more explicit, intentional, or strategic fashion than first language acquisition. Whereas most adults acquire their L2 at least partially through explicit instruction, as in a classroom setting, many others acquire their L2 primarily through implicit exposure, as is typical of an immersion environment. Although it has been widely believed that adults cannot learn a second language in the same way that children learn a first language, recent evidence suggests that adult learners of a second language can sometimes be indistinguishable from native speakers, and that this may be more likely under implicit learning conditions and with high quality input (VanPatten et al., 2019).

A great deal of cognitive research has also explored the development of reading skills. This research supports two important conclusions: first, that mastering the alphabetic principle (the idea that written symbols correspond to sounds, or phonics) is essential to becoming proficient, and second, that methods that teach the alphabetic decoding principles are more effective than methods that do not, especially for children who are having difficulty learning to read (Rayner et al., 2001).

Just as humans are thought to be predisposed to learn language, mathematical understanding is now thought to develop from an intuitive approximate number system. Beginning in early infancy, the approximate number system tracks differences in magnitude. It is thought to play a crucial role in development of other numerical abilities, such as the concept of exact number and simple arithmetic (Feigenson et al., 2004).

Testing effects

Just as repeated practice increases the ability to perform a skill well, repeatedly retrieving information stored in memory strengthens the memory trace and increases the likelihood that it can be retrieved again at a later instance. This phenomenon is known as the *testing effect* or *retrieval practice effect* (Roediger and Karpicke, 2006). Moreover, repeated retrieval also affects subsequent memory for associated information. Memory is reconstructive, and every time a memory is retrieved, it is altered. What information is tested may affect what is retained. Retrieving some information from memory can cause other non-retrieved information to be forgotten, known as *retrieval-induced forgetting*. Also, although retrieval practice can strengthen memory for retrieved information, it can increase false memories and cause other memory errors. For example, when distractors are presented as part of multiple-choice practice test questions, they can lead to memory for the incorrect answers.

Spaced practice effects

The benefits that come from practice become larger when it is spaced and repeated over time (Cepeda et al., 2006). Spaced practice schedules that involve either distributed or interleaved practice have been shown to lead to more robust learning than massed or blocked practice. *Distributed practice* occurs when students engage in repeated practice spaced out over time as opposed to practice being massed within a single session. *Interleaved practice* occurs when practice alternates between two or more concepts (such as subtraction and addition) as opposed to the two concepts being practiced separately. Practicing more than one concept at a time allows for comparisons to be made between the concepts, while spacing allows for some forgetting to occur between attempts. This means that to recover the information will require a more effortful retrieval process. The additional effort that is required ultimately strengthens the accessibility of the information in memory (Bjork, 1994). Another advantage from spaced learning is that the passage of time also changes the context of learning. With exposure to information in different contexts, the information becomes much more likely to be accessible in different situations and with different cues in the future.

Generation effects

When learners are asked to construct a response during learning, this has been shown to aid memory. In one classic demonstration of the *generation effect*, Slamenka and Graf (1978) found that words that are generated by learners are remembered better than when those same words are simply read. They showed this by asking learners to generate target words using a rule (opposite, synonym, or category member) from a cue word and a first letter (e.g., opposite: long—s _ _ _). Generation led to better memory for the generated response (short) than if the same words were presented to be read (long—short). In another demonstration, Jacoby (1978) showed that when learners were prompted to “solve” word fragment problems with associate words (e.g., foot—s _ _ e) they had better memory for the solution word (shoe) than when learners were simply shown the intact words (foot—shoe). One explanation for these benefits is that generating or constructing a response prompts the learner to take a more active role in the learning process, as opposed to passively receiving information. As with the benefits seen due to spaced practice, this greater effort during encoding leads to an increased chance of retrieval in the future.

Semantic processing effects

A long tradition of research has demonstrated more durable learning and memory from contexts that allow for the learner to generate meaningful connections or associations—when information is presented in a context that allows for elaboration or integration with prior knowledge. One of the seminal studies showing these effects is [Craik and Lockhart's \(1972\)](#) experiment on *levels of processing*. Individuals can process information on a number of different levels. When readers were exposed to word lists and asked to consider surface-level details such as whether the target words were written in capital letters (CHAIR), they found that memory was generally worse than when they were asked to consider the sound of the target words (e.g., Does this word rhyme with FLARE?). Further, people were best able to recall target words when they were prompted to consider their meaning during exposure (e.g., Is this a piece of FURNITURE?). Because the learner could integrate the target words more deeply into their long-term memory when they were prompted to consider the meaning of the words, this benefit is referred to as the *depth of processing effect*. Processing information in terms of its meaning leads to more associations, more integration with prior knowledge, and more elaborated memory traces. This helps the learner to leverage their prior knowledge when learning, to resist interference from competing pieces of information, and to have a richer set of retrieval cues which increases the chance of having access to the information in the future.

For similar reasons, new information is generally better recognized and recalled when it is supported by a meaningful context. For example, the *word superiority effect* results in individual letters being recognized more quickly when they appear in words as opposed to non-words ([Cattell, 1886](#)). Additionally, words are recognized more quickly when they appear in meaningful sentences ([Stanovich and West, 1979](#)). Similarly, experts have better memory for information that can be directly related to their domain knowledge. For example, chess experts are able to use their extensive knowledge of chess to remember the locations of pieces on a chess board that they are shown only very briefly ([Chase and Simon, 1973](#)). And, baseball fans have better memory for baseball games ([Chiesi et al., 1979](#)). Experts' prior knowledge is a richly connected network of associations, which provides a structure that allows them to easily encode and retrieve relevant information.

Text is also more memorable when it fits expectations and can be integrated into our prior knowledge. Individuals use their generic knowledge of typical event sequences and about everyday things (*schemas*) as an organizing framework for their memory. When we are given a context that allows us to use our prior knowledge, then our memory is greatly supported. There is a substantial body of work that shows how schemas can affect our memory. Sentences such as "First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do." are much easier to recall if you know they are about "Washing Clothes" ([Bransford and Johnson, 1972](#)).

In general, memory tends to be more robust as the learner engages more meaningfully or actively with information during encoding. Integrating the new information with prior knowledge, developing a richer memory trace with elaboration, or strengthening a memory trace through spaced or interleaved practice, are all conditions that generally will help improve the likelihood that information will be learned, remembered, and available in the future.

Comprehension processes

Although accumulating information in long-term memory is one goal of learning, another goal is for students to comprehend and understand new concepts. In contrast to simple retention of information, which can be strengthened through practice, comprehension requires processing information in terms of its meaning ([Kintsch, 1994](#)). The most common methods used to convey subject-matter understanding are reading expository text or listening to lectures. In these cases, learning depends on the comprehension of text or discourse. Theories of text and discourse processing propose that multiple levels of representation may be constructed from this information. For example, according to Kintsch's framework for comprehension based in the Construction-Integration model ([1998](#)), at the most basic level (the *surface level*) the learner forms a representation based on the exact words that they read. Beyond that, the learner may represent the meaning of each point made by the text in a slightly more abstracted, propositional, or gist form. This level of representation is known as the *textbase*, and still closely follows the information presented in the text. In contrast, as a learner attempts to understand what a text is about, it requires them to connect ideas across different parts of the text, and to integrate new information with relevant prior knowledge from long-term memory, to make inferences and generate relations that are only implied by the text. It is constructing this deeper level of representation, generally referred to as a *situation model* or mental model, that represents understanding or comprehension, and allows students to apply or transfer their knowledge to new situations ([Mayer, 1989](#)).

Conditions that improve comprehension

Because of the important role that integration with prior knowledge plays in the comprehension process, the benefits that come from possessing relevant general knowledge or expertise in a domain also extend to comprehension outcomes ([Bisanz and Voss, 1981](#)). Schematic processing which can be prompted by even just providing a title for a passage improves not only memory for

the content but also the ability to comprehend or understand it (Anderson, 1978). In addition, comprehension processes can be supported by a variety of instructional conditions that may help students with low prior knowledge also develop understanding. These instructional conditions help to encourage students to engage in the reasoning processes that are required to construct a situation or mental model of the content.

Goal-directed or purposeful reading

The processes that students engage in while learning depend on their perception of the goal or purpose for the activity, and it is important for students to be provided with specific goals or instructions prior to reading (Britt et al., 2017). These goals may direct the reader's attention to specific concepts in the text, serve as test-expectancies, or provide guidance about what might be needed for other post-reading tasks that may alter behaviors during reading. By clarifying the purpose for reading, it not only makes the task more meaningful, it also allows learners to make better decisions about how they are reading, and to better monitor whether they need to exert more effort to comprehend the information for a future application or examination.

Learning from questions and explanations

Activities such as elaborative interrogation or responding to adjunct questions can support construction of better situation models during learning. Generating questions and explanations can also encourage the activation and use of prior knowledge, and can aid monitoring of understanding and where gaps may exist in understanding (Graesser and Person, 1994). Although these activities can be effective for improving comprehension, the effects are highly dependent on the level of processing that the learner engages in during the task. To be most effective, the learner should be focused on constructive processing in which they are actively generating new inferences left implicit in the text as opposed to solely retrieving information already explicitly available in the text (Hinze et al., 2013).

Learning from multiple sources

With the emergence of the Internet, students are increasingly being tasked with learning from multiple sources, such as those obtained from web searches. Engaging in inquiry with multiple sources is a complex activity because it requires all the processes necessary for learning from individual texts, plus an additional set of processes that become particularly important when readers are confronted with information from more than one source (Wiley et al., 2018). Comprehending events or phenomena from multiple sources instead of a single source necessitates representation of information about the origin, nature, or quality of the various sources (who wrote it, for what purpose was it written, is it reliable), as well as relations between the sources (the presence of corroborating or conflicting statements, detecting consistency or inconsistency). With a single source, the construction of a good-enough situation model can often be achieved by representing the original text's intended purpose, structure, or argument in memory. However, when tasked with comprehending an event or phenomenon from multiple sources, the learner must actively impose selection, organization, and transformation to construct a mental model that integrates the information from different texts, rather than more passively creating distinct traces for each individual text. This provides an opportunity for richer student understanding. Learners may benefit from learning from multiple sources to the extent that they actively engage in constructive processing, attempt to build connections across ideas, and try to form a coherent, integrated model of the phenomena. Although different reading goals may affect learning in single-source contexts, they become even more critical to consider in multiple-source contexts. The likelihood of engaging in these activities depends on the learner's interpretation of the task, as well as individual differences such as epistemic or epistemological beliefs.

Learning from multiple representations

Exposure to multiple representations can also support better learning. When attempting to learn a new construct, adding alternative representations such as diagrams can provide other routes to understanding (Ainsworth, 2006). For example, while generating written explanations can support understanding, sketching can be thought of as a visual explanation process that can also support understanding.

Alternatively, gestures can be used to support spatial thinking, but also may be used to represent or express knowledge that is difficult to verbalize (Alibali, 2005). Gestures may reveal a student's preparedness for learning a concept even when their verbal expressions suggest poor understanding. Interestingly, although we may think of gestures as devices used to communicate with others, gestures can also be useful for helping students to clarify their own understanding.

Presenting multiple cases or adding an analogy can also provide a basis for understanding, allowing learners to compare or map relations across different contexts, and abstract important information (Catrambone and Holyoak, 1989). Research has also shown that presenting alternative representations or contrasting cases to the learner in a way that supports comparison and identification of distinguishing features can help the learner to form generalizations, and prepare the learner for future learning.

Conceptual change and restructuring processes

Although learning can sometimes proceed by simply adding to existing knowledge, there are times when learners need to overcome some of their intuitive ideas in order to comprehend accurate concepts of phenomena. Following [Piaget \(1977\)](#), learning can be characterized as resulting from two fundamental processes: adaptation and disequilibrium. *Adaptation* involves using existing knowledge as a basis for understanding new information by either assimilation or accommodation. *Assimilation* is the direct application of previously known concepts to the novel stimuli. *Accommodation* is the updating or fine-tuning of previously known concepts in the face of new information. Prior knowledge is refined by this process. *Disequilibrium* occurs when the new information cannot be simply added to or easily accommodated by previous concepts. In these cases, when new information or evidence in some way conflicts with prior understandings, then restructuring, *conceptual change*, or belief change is required ([Chinn and Brewer, 1993](#)).

Some of the most prominent ways to help students to engage in radical conceptual change or restructuring processes attempt to put them in situations where they must directly confront the inconsistency or conflict caused by the incorrect conception ([Vosniadou, 2013](#)). Some instructional contexts introduce anomalous evidence. Other instructional contexts employ refutational texts where an intuitive or incorrect conception is explicitly flagged and contrasted with the preferred conception ([Hynd and Alvermann, 1986](#)). Importantly, in order for conceptual change to occur, not only must some conflict or inconsistency exist and be noticed, but the learner also needs to actively engage in resolving it ([Diakidoy and Kendeou, 2001](#); [Limón, 2001](#)).

The process of constructing a correct representation is important not just for developing understanding when learning from text, it is also important for successful problem solving ([Koedinger and Nathan, 2004](#)). The first stage of the problem-solving process requires the construction of a problem representation, which is followed by the subsequent solution stage. When learners initially choose an incorrect problem representation, then they need to engage in representational change and restructure their interpretation of the problem in order to make progress. Starting with an incorrect representation is a common issue that arises for a number of reasons including misinterpreting linguistic expressions in word problems, misunderstanding of symbols, or depending on inappropriate concepts or prior knowledge (such as attempting to use arithmetic rather than algebraic approaches). In these cases, it is often useful for learners to reach an impasse or experience failure, as this can prompt reflection and revision as part of the representational change process.

In general, comprehension requires that the learner engages meaningfully and actively with information to construct a representation of what they are learning about. Integrating new information with prior knowledge, making connections across sentences or sources, mapping across information formats, detecting inconsistencies or conflicts and taking action to resolve them, all support better conceptual understanding and more coherent mental models.

Other conditions that improve learning

This final section highlights several areas of research that build on the key observations that successful learning depends on students engaging in meaningful, constructive, and reflective processing: the role of interaction with others; the role of an initial invention, exploration, or discovery phase on preparation for future learning; the role of feedback; and the role of monitoring in effective self-regulated learning.

Active and interactive learning with others

In contrast to the adaptation and restructuring processes that were emphasized as fundamental to learning by [Piaget \(1977\)](#), [Vygotsky and Cole \(1978\)](#) emphasized the processes of learning that come from interactions with others, and in particular from a more knowledgeable other. A common theme in educational research is exploring the ways that social interaction can support learning through small-group exercises or peer collaboration activities such as *reciprocal teaching* ([Palincsar and Brown, 1984](#)) or *learning by teaching* ([Guerrero and Wiley, 2021](#)).

[Chi's \(2009\)](#) ICAP (Interactive, Constructive, Active, Passive) framework places interactive processes as more effective than other constructive or active learning processes, which in turn are more effective than passive learning processes. Interacting with peers during learning can provide all members of a group with relevant knowledge that only few students may possess. Each member can contribute a new perspective or specific skills. Working with others can help to support planning and execution of complex tasks. Others can also provide feedback. Students can support each other by prompting evaluation, explanation, abstraction or reflection.

Learning by invention, learning from errors, preparation for future learning

Several lines of research have explored how the sequencing of learning activities influences the knowledge that learners construct, and their ability to use it in new contexts. Although many traditional approaches expose students to content via a reading or a lecture before they engage in quizzing or problem-solving activities, some alternative approaches have students engage in carefully designed inquiry, exploratory, or discovery activities before receiving direct instruction (*preparation for future learning*, [Schwartz and Martin, 2004](#); *learning by invention*, [Jarosz et al., 2017](#); *productive failure*, [Sinha and Kapur, 2021](#)). In an example

of a learning-by-invention paradigm, students work together in small groups to try to develop a method of quantifying variability before they engage in a follow-up discussion with the teacher in which they are taught the canonical formula for standard deviation. Although student groups are rarely successful at inventing the canonical formula on their own during the invention phase, the initial activity leads them to failures or impasses. The difficulties they experience during these attempts can help them to recognize gaps in their knowledge, which then greatly enhances the uptake of the subsequent lesson. In other words, by putting exploration and active learning first, these activities prepare students for future learning. Ultimately this helps the lesson to become much more meaningful, and students are much better prepared to transfer and use the information in new contexts (Schwartz and Bransford, 1998).

Importance of feedback and guidance

Opportunities that come from practice testing, initial exploration activities, or activities that involve interacting with peers also mean that students may be exposed to incorrect ideas. Feedback can play an important role in the active learning process by helping students to correct errors that may be generated. Feedback not only helps to direct attention toward correct information, and prompts students to repair and update incorrect representations, it provides guidance for future learning, and can help to lower cognitive load during learning (Shute, 2008).

Interacting with others can be particularly motivating, but there are also a number of pitfalls that can make collaborative learning less successful than individual work. Guidance for students on how to collaborate is critical, otherwise they may engage in superficial discussions that do not reach a high level of discourse. Invention activities are also highly motivating, but the guided discussions or reflection activities that follow exploratory or generative activities are critical to leverage the advantages of learning from invention, errors or failures. Intelligent tutors can guide students with an interactive experience that can offer timely feedback and expert knowledge (van Lehn, 2011). But if systems provide too much scaffolding or help, then the students may never actually meaningfully engage with the content.

Metacognitive monitoring and self-regulated learning

Metacognition is the act of monitoring one's own cognitive processes and includes assessing one's own learning, memory, and understanding (Brown, 1978; Flavell, 1979). These monitoring assessments then serve as input to the self-regulated learning process (Winne and Hadwin, 1998), and to the decision making that the learner engages in when choosing among behaviors and actions that can be taken. Accurate monitoring allows learners to make better decisions and engage in more effective self-regulated learning.

Much research on the topic of monitoring accuracy has used prediction paradigms for memory and comprehension tasks. *Metamemory* paradigms typically consist of tasks that include learning word pairs, such as a word in a foreign language and its English translation, or definitions for new vocabulary (Dunlosky and Metcalfe, 2008). When learners are asked to predict their ability to recall a translation or definition, metamemory accuracy has been found to be quite good, especially when judgments are made at a delay. When learners are asked to predict how well they will do on a test of comprehension following the reading of expository texts, *metacomprehension* accuracy is typically much lower. Yet, several contexts have now been found to promote accurate metacomprehension (Griffin et al., 2019). When learners are given clear comprehension goals as their purpose for reading, or are prompted to self-explain during learning, metacomprehension accuracy can be substantially improved. The ultimate value of supporting better monitoring accuracy among learners is that it then can help them to make effective decisions of what to re-study as they attempt to master material on their own.

Conclusion

This overview began with the consideration that memory is at the center of the fundamental cognitive processes that underlie learning. Much work has shown that human memory takes a variety of forms and may contribute to learning in different ways. Our ability to learn is affected by the attentional bottleneck which constrains our immediate capacity. Our learning can be improved by engaging in practice, repetition, and making meaningful associations based in memory. Comprehension fundamentally depends on memory, but in addition requires the learner to engage actively with the information to develop a more conceptual, complex mental representation. Many contexts that have been demonstrated to support effective learning outcomes involve engaging students in constructive, interactive, and reflective processes. Moving beyond these fundamental cognitive processes, broader approaches consider the rich interplay between cognition and metacognition for successful learning, as well as the critical role of motivation, curiosity, interest, and willingness to invest in learning and to explore what one does not know.

Some excellent overviews of how research on fundamental cognitive principles of learning can be applied to instruction and educational contexts can be found in Dunlosky et al. (2013), Roediger and Pyc (2012) and the *How People Learn* series (National Academies of Sciences, Engineering, and Medicine, 2018).

References

- Ainsworth, S., 2006. DeFT: a conceptual framework for considering learning with multiple representations. *Learn. Instr.* 16 (3), 183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>.
- Alibali, M.W., 2005. Gesture in spatial cognition: expressing, communicating and thinking about spatial information. *Spatial Cognit. Comput.* 5, 307–331. https://doi.org/10.1207/s15427633scc0504_2.
- Anderson, R.C., 1978. Schema-directed processes in language comprehension. In: Lesgold, A.M., Pellegrino, J.W., Fokkema, S.D., Glaser, R. (Eds.), *Cognitive Psychology and Instruction*. Springer, pp. 67–82. https://doi.org/10.1007/978-1-4684-2535-2_8.
- Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. *Psychol. Learn. Motiv.* 2, 89–195. [https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3).
- Baddeley, A.D., 2000. The episodic buffer: a new component of working memory? *Trends Cognit. Sci.* 4 (11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2).
- Baddeley, A.D., Hitch, G.J., 1974. Working memory. In: Bower, G.A. (Ed.), *Recent Advances in Learning and Motivation*, vol. 8. Academic Press, pp. 47–89. [https://doi.org/10.1016/s0079-7421\(08\)60452-1](https://doi.org/10.1016/s0079-7421(08)60452-1).
- Bisanz, G.L., Voss, J.F., 1981. Sources of knowledge in reading comprehension: cognitive development and expertise in a content domain. In: Lesgold, A.M., Perfetti, C.A. (Eds.), *Interactive Processes in Reading*. Lawrence Erlbaum, pp. 215–240.
- Bjork, R.A., 1994. Memory and metamemory considerations in the training of human beings. In: Metcalfe, J., Shimamura, A.P. (Eds.), *Metacognition: Knowing About Knowing*. The MIT Press, pp. 185–205.
- Bransford, J.D., Johnson, M.K., 1972. Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verb. Learn. Verb. Behav.* 11 (6), 717–726. [https://doi.org/10.1016/S0022-5371\(72\)80006-9](https://doi.org/10.1016/S0022-5371(72)80006-9).
- Britt, M.A., Rouet, J.-F., Durik, A.M., 2017. *Literacy Beyond Text Comprehension: A Theory of Purposeful Reading*. Routledge.
- Broadbent, D.E., 1958. *Perception and Communication*. Pergamon Press. <https://doi.org/10.1037/10037-000>.
- Brown, A.L., 1978. Knowing when, where, and how to remember: a problem of metacognition. In: Glaser, R. (Ed.), *Advances in Instructional Psychology*, vol. 1. Erlbaum, pp. 77–165.
- Catrambone, R., Holyoak, K.J., 1989. Overcoming contextual limitations on problem-solving transfer. *J. Exp. Psychol. Learn. Mem. Cognit.* 15 (6), 1147–1156. <https://doi.org/10.1037/0278-7393.15.6.1147>.
- Cattell, J.M., 1886. The time it takes to see and name objects. *Mind* 11 (41), 63–65.
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., Rohrer, D., 2006. Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132 (3), 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>.
- Chase, W.G., Simon, H.A., 1973. Perception in chess. *Cognit. Psychol.* 4 (1), 55–81. [https://doi.org/10.1016/0010-0285\(73\)90004-2](https://doi.org/10.1016/0010-0285(73)90004-2).
- Chi, M.T.H., 2009. Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Top. Cognit. Sci.* 1 (1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>.
- Chiesi, H.L., Spilich, G.J., Voss, J.F., 1979. Acquisition of domain-related information in relation to high and low domain knowledge. *J. Verb. Learn. Verb. Behav.* 18 (3), 257–273. [https://doi.org/10.1016/S0022-5371\(79\)90146-4](https://doi.org/10.1016/S0022-5371(79)90146-4).
- Chinn, C.A., Brewer, W.F., 1993. The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Rev. Educ. Res.* 63 (1), 1–49. <https://doi.org/10.3102/00346543063001001>.
- Cowan, N., 2014. Working memory underpins cognitive development, learning, and education. *Educ. Psychol. Rev.* 26 (2), 197–223. <https://doi.org/10.1007/s10648-013-9246-y>.
- Craik, F.I., Lockhart, R.S., 1972. Levels of processing: a framework for memory research. *J. Verb. Learn. Verb. Behav.* 11 (6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X).
- Diakidoy, I.-A.N., Kendeou, P., 2001. Facilitating conceptual change in astronomy: a comparison of the effectiveness of two instructional approaches. *Learn. Instr.* 11 (1), 1–20. [https://doi.org/10.1016/S0959-4752\(00\)00011-6](https://doi.org/10.1016/S0959-4752(00)00011-6).
- Dunlosky, J., Metcalfe, J., 2008. *Metacognition*. Sage Publications.
- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., Willingham, D.T., 2013. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Publ. Interest* 14 (1), 4–58. <https://doi.org/10.1177/1529100612453266>.
- Engle, R.W., 2002. Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11 (1), 19–23. <https://doi.org/10.1111/1467-8721.00160>.
- Flavell, J.H., 1979. Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. *Am. Psychol.* 34 (10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>.
- Feigenson, L., Dehaene, S., Spelke, E., 2004. Core systems of number. *Trends Cognit. Sci.* 8 (7), 307–314. <https://doi.org/10.1016/j.tics.2004.05.002>.
- Gleitman, L.R., Newport, E.L., 1995. The invention of language by children: environmental and biological influences on the acquisition of language. In: Gleitman, L., Liberman, M. (Eds.), *An Invitation to Cognitive Science: Language*. MIT Press, pp. 1–24.
- Graesser, A.C., Person, N.K., 1994. Question asking during tutoring. *Am. Educ. Res. J.* 31 (1), 104–137. <https://doi.org/10.3102/00028312031001104>.
- Griffin, T.D., Mielicki, M.K., Wiley, J., 2019. Improving students' metacomprehension accuracy. In: Dunlosky, J., Rawson, K. (Eds.), *Cambridge Handbook of Cognition and Education*. Cambridge University Press, pp. 619–646.
- Guerrero, T.A., Wiley, J., 2021. Expecting to teach affects learning during study of expository texts. *J. Educ. Psychol.* 113 (7), 1281–1303. <https://doi.org/10.1037/edu0000657>.
- Hinze, S.R., Wiley, J., Pellegrino, J.W., 2013. The importance of constructive comprehension processes in learning from tests. *J. Mem. Lang.* 69, 151–164. <https://doi.org/10.1016/j.jml.2013.03.002>.
- Hynd, C., Alvermann, D.E., 1986. The role of refutation text in overcoming difficulty with science concepts. *J. Read.* 29 (5), 440–446.
- Jacoby, L.L., 1978. On interpreting the effects of repetition: solving a problem versus remembering a solution. *J. Verb. Learn. Verb. Behav.* 17 (6), 649–667. [https://doi.org/10.1016/S0022-5371\(78\)90393-6](https://doi.org/10.1016/S0022-5371(78)90393-6).
- Jaros, A.F., Goldenberg, O., Wiley, J., 2017. Learning by invention: small group discussion activities that support learning in statistics. *Discourse Process.* 54 (4), 285–302. <https://doi.org/10.1080/0163853X.2015.1129593>.
- Kintsch, W., 1994. Text comprehension, memory, and learning. *Am. Psychol.* 49 (4), 294–303. <https://doi.org/10.1037/0003-066X.49.4.294>.
- Kintsch, W., 1998. *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Koedinger, K.R., Nathan, M.J., 2004. The real story behind story problems: effects of representations on quantitative reasoning. *J. Learn. Sci.* 13 (2), 129–164. https://doi.org/10.1207/s15327809jls1302_1.
- Limón, M., 2001. On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. *Learn. Instr.* 11 (4–5), 357–380. [https://doi.org/10.1016/S0959-4752\(00\)00037-2](https://doi.org/10.1016/S0959-4752(00)00037-2).
- Mayer, R.E., 1989. Models for understanding. *Rev. Educ. Res.* 59 (1), 43–64. <https://doi.org/10.3102/00346543059001043>.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63 (2), 81–97. <https://doi.org/10.1037/h0043158>.
- National Academies of Sciences, Engineering, and Medicine (NASEM), 2018. *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press. <https://doi.org/10.17226/24783>.
- Neisser, U., 1967. *Cognitive Psychology*. Appleton-Century-Crofts.

- Palinscar, A.S., Brown, A.L., 1984. Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognit. Instr.* 1 (2), 117–175. https://doi.org/10.1207/s1532690xci0102_1.
- Pashler, H., 1998. *The Psychology of Attention*. MIT Press.
- Piaget, J., 1977. *The Development of Thought: Equilibration of Cognitive Structures*. Viking.
- Postman, L., 1964. Short-term memory and incidental learning. In: Melton, A.W. (Ed.), *Categories of Human Learning*. Elsevier, pp. 145–201.
- Rayner, K., Foorman, B.R., Perfetti, C.A., Pesetsky, D., Seidenberg, M.S., 2001. How psychological science informs the teaching of reading. *Psychol. Sci. Publ. Interest* 2 (2), 31–74. <https://doi.org/10.1111/1529-1006.00004>.
- Roediger, H.L., Karpicke, J.D., 2006. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>.
- Roediger, H.L., Pyc, M.A., 2012. Inexpensive techniques to improve education: applying cognitive psychology to enhance educational practice. *J. Appl. Res. Mem. Cognit.* 1 (4), 242–248. <https://doi.org/10.1016/j.jarmac.2012.09.002>.
- Schwartz, D.L., Bransford, J.D., 1998. A time for telling. *Cognit. Instr.* 16 (4), 475–522. https://doi.org/10.1207/s1532690xci1604_4.
- Schwartz, D.L., Martin, T., 2004. Inventing to prepare for future learning: the hidden efficiency of encouraging original student production in statistics instruction. *Cognit. Instr.* 22 (2), 129–184. https://doi.org/10.1207/s1532690xci2202_1.
- Shute, V.J., 2008. Focus on formative feedback. *Rev. Educ. Res.* 78 (1), 153–189. <https://doi.org/10.3102/0034654307313795>.
- Sinha, T., Kapur, M., 2021. When problem solving followed by instruction works: evidence for productive failure. *Rev. Educ. Res.* 91 (5), 761–798. <https://doi.org/10.3102/00346543211019105>.
- Slamecka, N.J., Graf, P., 1978. The generation effect: delineation of a phenomenon. *J. Exp. Psychol. Hum. Learn. Mem.* 4 (6), 592–604. <https://doi.org/10.1037/0278-7393.4.6.592>.
- Squire, L.R., 1987. *Memory and Brain*. Oxford University Press.
- Stanovich, K.E., West, R.F., 1979. Mechanisms of sentence context effects in reading: automatic activation and conscious attention. *Mem. Cognit.* 7 (2), 77–85. <https://doi.org/10.3758/BF03197588>.
- Tulving, E., 1985. How many memory systems are there? *Am. Psychol.* 40 (4), 385–398. <https://doi.org/10.1037/0003-066X.40.4.385>.
- VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* 46 (4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>.
- VanPatten, B., Smith, M., Benati, A.G., 2019. *Key Questions in Second Language Acquisition: An Introduction*. Cambridge University Press.
- Vosniadou, S., 2013. *International Handbook of Research on Conceptual Change*. Routledge. <https://doi.org/10.4324/9780203154472>.
- Vygotsky, L.S., Cole, M., 1978. *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.
- Waugh, N.C., Norman, D.A., 1965. Primary memory. *Psychol. Rev.* 72 (2), 89–104. <https://doi.org/10.1037/h0021797>.
- Wiley, J., Jaeger, A.J., Griffin, T.D., 2018. Effects of instructional conditions on comprehension from multiple sources in history and science. In: Braasch, J.L.G., Braten, I., McCrudden, M.T. (Eds.), *Handbook of Multiple Source Use*. Routledge, pp. 341–361.
- Winne, P.H., Hadwin, A.F., 1998. Studying as self-regulated learning. In: Hacker, D.J., Dunlosky, J., Graesser, A.C. (Eds.), *Metacognition in Educational Theory and Practice*. Lawrence Erlbaum, pp. 277–304.

Attention: a prerequisite for learning

Karin Strid^a and Mikael Heimann^b, ^a Department of Psychology, University of Gothenburg, Gothenburg, Sweden; and ^b Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Attention in infancy	117
Selective attention	118
Attention to faces	118
Attention networks	119
Behavioral phases	119
Brain networks	120
Attention and developmental disorders	121
Genes and attention	122
Aspects of attention	122
Novelty attracts attention	122
Memory and attention	122
Pupil dilation and attention	122
Attention and learning	123
The directed-attention model	123
Dyadic attention	123
Joint attention	124
Intention requires attention	124
Digital media and attention	124
Compensatory systems	125
Final remarks	125
Acknowledgment	125
References	125
Further reading	126

The connections between the nodes of attention networks develops slowly over childhood and is influenced by the interaction of genes and the parental environment. The development of attention networks supports both self-regulation and the child's understanding of other minds.

Posner et al. (2020, p 220).

A child's ability to direct his or her attention drives awareness; only objects and events attended to will enter the child's mind. In this sense, attention processes are essential for learning and cognitive development. Attention influences and is, in turn, influenced by various brain systems, an interaction that creates priorities affecting both perception and action. Thus, alertness and the way attention regulates, or is regulated by, brain functions becomes highly relevant for all domains of learning. Without attention it becomes difficult, if not impossible, to take in new and important information. We, adults and children, are active in selecting what to attend to and what to ignore – a process that becomes more and more voluntary during development.

Attention in infancy

The view of an infant during the first months of life as passive, incapable of communicating and remembering, and unaware about self and others has changed radically during the last 50–60 years. Typically, developing infants have a sensory system, which makes them capable of perceiving and interacting with the environment promoting communication, emotional exchange, and signals for caretaking. In addition, they use their own body as a vehicle, not only for exploration of themselves, that is, building sensory motor schemas, but also for assembling information about other individuals. To learn from and understand other people and socially relate to them is crucial for a healthy development. In short, newborn children enter this world equipped with a couple of biases that helps them to initiate and maintain attention to other people (Bjorklund, 2021). Among those biases are (1) a probably innate ability to attend to biological motion; (2) an attractiveness to faces or face-like patterns; (3) a tendency to direct more attention to human speech sounds than other sounds in the immediate environment; and (4) an ability to respond in an imitative-like fashion already at birth.

Furthermore, by 4 months infants look longer if there is a mutual eye gaze with the adult compared to averted eye gaze. Also, auditory perception in 6-month-old infants is linked to sensitive detection of human voices, such as discrimination of phoneme segments, both in native and foreign languages. However, during the second half of the first year of life this ability declines showing

that early on our brain is prepared for a general language acquisition, but that exposure to one language with its specific prosodic and phonetic patterns reduces this capacity.

All sensory systems are functional at or before birth and of these, the visual system is probably both the best-understood and least-mature system. It is actually a difficult task for the newborn infant to respond to visual stimuli. When a newborn infant acts on complex visual information, as evident when the neonate mirrors facial movements, the infant responds in spite of the immaturity of the visual system: Jerky eye movements make it hard to control vision and it is difficult for the immature system to capture fast-moving objects. In short, the infant's control of fixation is mostly not under voluntary control at birth and, during the first months of life, infants cannot resist being drawn to certain patterns. Visually, the newborn infant is attracted to high spatial frequencies (e.g., checkerboard pattern and edges), to slowly moving objects (moving stimuli are more attractive than static ones), and to face-like patterns. The attraction to edges provides necessary input for the visual cortex to develop (input helps to organize cell columns in visual cortex) while the attraction to face-like patterns guides the infant toward the social world.

All nonvisual sensory systems are also functional at birth but their role in early attentional processes is less well studied, although auditory attention probably is as important as vision for early development. Hearing is essential both for language learning and for making the social world interesting. The newborn infant recognizes human voices at birth (in fact, a fetus can learn to identify the mother's voice several weeks post-partum) and the melody of the mother tongue is identified within the first weeks. The human voice, especially the female voice, attracts the infants' attention both to the social world and to language.

Selective attention

Selective attention develops rapidly over the first months of life. The newborn infant has less oculomotor control, less control over attentional shifts, and is attracted by salient details in perceptual displays. A neonate is not able to understand a partly occluded object as a unity and reacts only to direct visible information. It becomes a very difficult task if the neonate must identify a center-occluded object – the infant sees only the top and the bottom part of the object – in order to solve a task. This does not imply that a newborn child is not able to take in the relevant information, only that during the first months of life, perception and attention is more driven by environmental input (exogenous processes) coupled with biologically driven subcortical processes. Slowly, vision becomes more and more cortically controlled, acuity develops, the visual field increases, the eyes become better coordinated and inspection times decrease. A shift is usually seen around 2–3 months making the infant more visually competent when inspecting new objects, and, maybe most important, the child now becomes able to partake in prolonged face-to-face interactions with the caregiver.

The problem with perceptual completion, to see an object as a whole in spite of the fact that it is partly occluded, also changes during the first months of life. Some capacity is observable already at 2 months, but it is not until 4–5 months that the ability to solve perceptual-completion problems can be expected to be robust. Interestingly, children at 3 months seem to be in a transition phase. In an informative eye-tracking study, [Amos and Johnson \(2006\)](#) identified two groups of children, perceivers and non-perceivers: the children who were able to perceive unity (i.e., they solved the perception-completion task) used a more efficient strategy when solving a visual search task than the non-perceivers. These differences may stem from the possibility that the two groups reflect different stages in early brain development, the children are in the midst of a transition from exogenous to more endogenous oculomotor control, or, in the authors' own words, "...differences between groups of 3-month-olds on visual search performance may be related to different stages in the development of visual pathways underlying oculomotor control" ([Amos and Johnson, 2006](#), p. 1244). Selective attention is important since it makes it possible for the child to be an active participant in his/her development and early differences in this ability might have an impact on later cognitive development.

The capacity to attend selectively is also related to disengagement, an ability that develops rapidly over the first months of life. As depicted in [Fig. 1](#), infants younger than 2 months tend to focus on one stimulus at a time. If a second stimulus is added to the visual display while the first stimulus is still visible, the infant is more or less unable to make a gaze shift. However, by 3–4 months, this is no longer a problem.

Selective looking and habituation have dominated research on infant attention. Generally, the length of looking declines during infancy, especially from 3 to 4 months and onward. Younger infants need longer inspection time (familiarization) than older infants; studies indicate a negative correlation between visual recognition and the inspection time needed. Selective attention has also been used to study auditory attention (e.g., the ability to discriminate between specific sound signals or between the mother's voice and a stranger's voice).

Attention to faces

Human infants show a preference for faces right from birth, a tendency that seems to be especially strong for moving face-like stimuli ([Johnson et al., 2015](#)). This has been interpreted by some as indicating a pre-wired system molded by evolution. Such a system is described as being controlled by subcortical parts of the visual system functional at birth making the face a highly salient stimulus for the newborn. Evidence based on electrophysiological brain responses (ERP) show that cortical areas start to become more selectively attuned to faces around 3 months of age ([Conte et al., 2020](#)). Thus, the child is endowed with an in-built perceptual attention grabber for face-like patterns. Although the propensity for faces among newborns is not debated, the mechanism is. In a series of studies, Simion and co-workers ([Simion and Di Giorgio, 2015](#); [Buiatti et al., 2019](#)) have provided support for an alternative view. Instead of being born with a module sensitive specifically to faces, they argue that newborns are born with "pre-wired

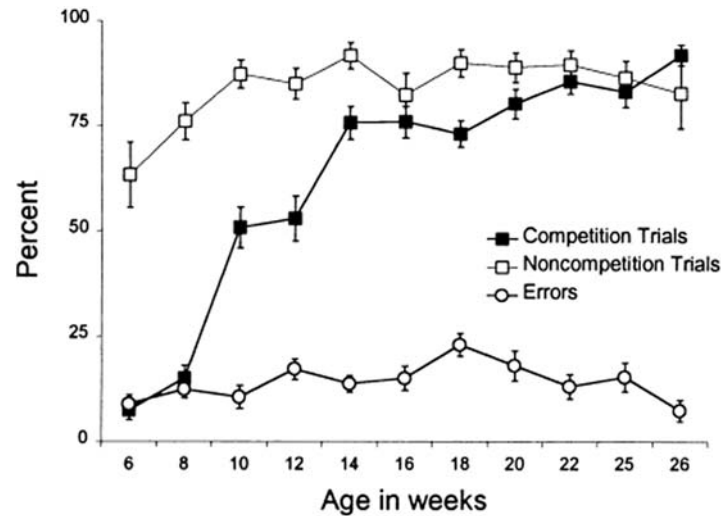


Fig. 1 The ability to disengage and shift gaze was studied in 16 infants observed longitudinally from 6 to 26 weeks. Gaze shift is indicated by the infants' responses to competition trials which required disengagement from one stimulus before a second (new) stimulus could be attended to. In noncompetition trials, the infant always saw only one target at a time. As can be seen from the graph, early development is characterized by a high degree of discontinuity. A rapid change is noted for the period 8–14 weeks, whereafter change continues but at a much slower pace. From [Butcher, Kalverboer, and Geuze \(2000, p 11\)](#), reprinted with permission.

domain relevant attentional biases toward faces" ([Simion and Di Giorgio, 2015, p 7](#)). These biases include sensitivity to top-heavy patterns, congruency and configural properties of the stimuli—biases that have the effect that faces becomes salient for the newborn in spite of the fact that the system is not specific to faces. Specific face systems build up with time as the infant gains experience and the relevant brain systems matures (e.g., the fusiform face area and the superior temporal sulcus) and becomes better integrated. The observation that infants early in life prefer own-race faces is a further indication of how experience molds our face recognition system. The conflict between the two contrasting views presented here is not resolved; we still do not know exactly how to describe the mechanism responsible for making newborn babies especially attracted to the human face or how a mature system for face representation develops.

Furthermore, observations also suggest that infants are better at processing female than male faces, a finding that probably stems from the fact that infants have much more experience in processing female faces. In addition, a child's temperament might also influence how faces are processed. Temperament factors such as orienting and surgency have been shown to affect how 6–10-month-old infants view faces according to race and/or gender ([Rennels et al., 2020](#)). Finally, over the course of the first year, infants begin to process naturally looking faces in the same fashion as adults do; that is, they analyze faces in a holistic gestalt-like fashion. Whereas 4-month-old children focus more on internal facial features, 10-month-old children use a holistic strategy. Six-month-old children seem to be in a transition phase using both strategies.

Attention networks

Behavioral phases

Attention in infancy is usually described as made up of separate phases representing different underlying processes. Arousal, a prerequisite for attention, reflects the physiological readiness of the organism to react to any stimuli. Usually, arousal varies from deep sleep to active wakefulness through several intermediate phases. A characteristic of the young infant is that state changes occur often and rapidly, which affects the manner in which early attention is modulated. Without adequate arousal, the child will not be alerted to an endogenous or exogenous stimulus and no attention can take place ([Colombo and Salley, 2015](#)). Behaviorally the process of attending can be parsed into three distinct phases ([Colombo et al., 2001](#); see [Fig. 2](#)). The first phase is an initial brief *orienting* (OR, sometimes also called selective attention) when the infant's attention is directed toward specific stimuli; that is, an interesting event is identified, and the system becomes prepared for further inspection. The second phase is *sustained attention* (SA), which reflects a period of active information processing or encoding. SA is often described as a voluntary process and it has been linked to anterior brain systems. In the newborn child SA is mostly governed by parents working actively to keep the child's attention. The final phase, *attention termination* (AT), on the other hand, is linked to disengagement processes reflecting continuous looking after cessation of any active information processing. AT has been proposed to reflect the posterior brain system and the disengagement processes have been observed to be a stronger prediction of recognition than sustained attention, at least for static stimuli.

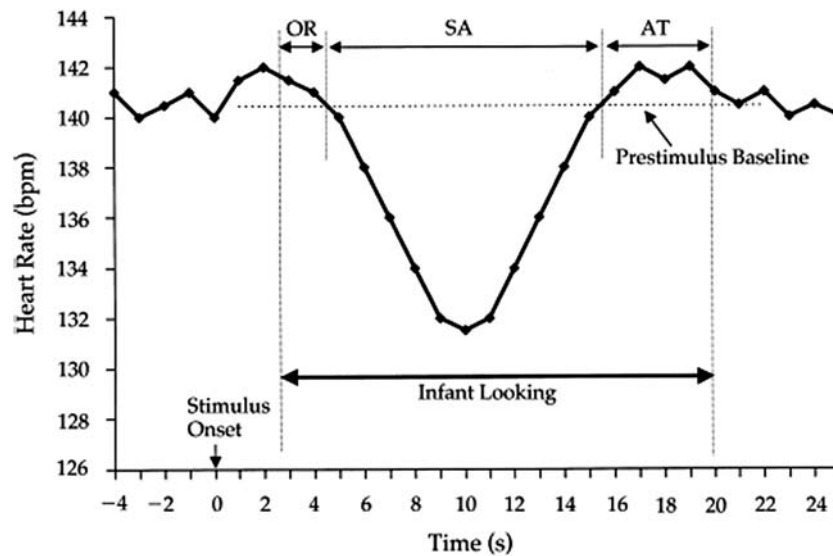


Fig. 2 The figure depicts how the three phases of attention described in the text might be observed when studying visual attention in infancy. In this particular study, 113 four-month-old infants participated; visual looking and heart rate (beats per minute) responses defined the phases: OR = Orienting; SA = Sustained attention; AT = Attention termination. Prestimulus baseline refers to baseline heart rate level. From Colombo et al. (2001, p. 1609), reprinted with permission.

Brain networks

In several studies, Posner and colleagues (e.g., Posner et al., 2013; Posner et al., 2020) have integrated behavioral and cognitive neuroscience studies in order to trace the development of attention from infancy to adulthood and to outline the three major brain-behavior networks (see Fig. 3) that underlie central aspects of attention: the alerting, orienting, and executive networks.

The first aspect, alerting, signifies that the child is in a state of readiness for reacting to new incoming stimuli. From a biological point of view, alerting has been linked to subcortical processes (locus coeruleus) as well as to right parietal and right frontal areas. Norepinephrine is the main neurochemical modulator.

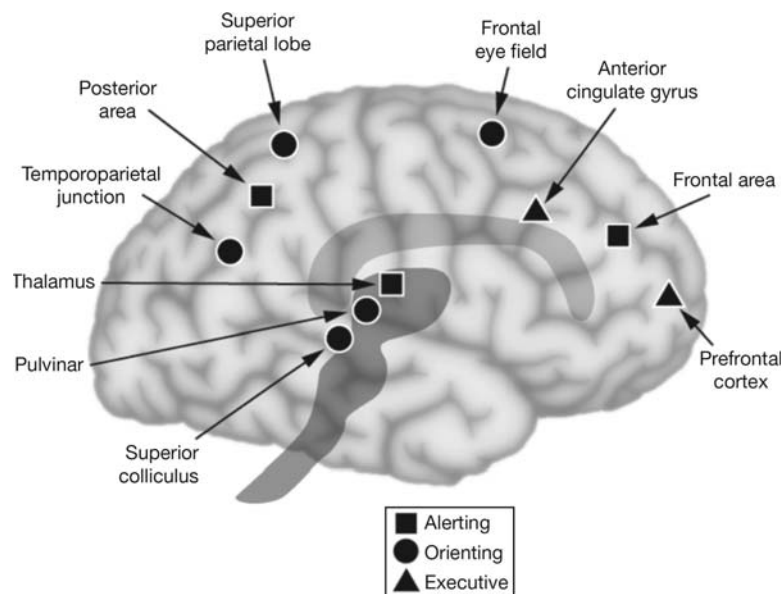


Fig. 3 The figure shows the different nodes of activation for the three main attention networks: alerting, orienting, and executive attention networks. From Posner et al. (2013, p. 396), reprinted with permission.

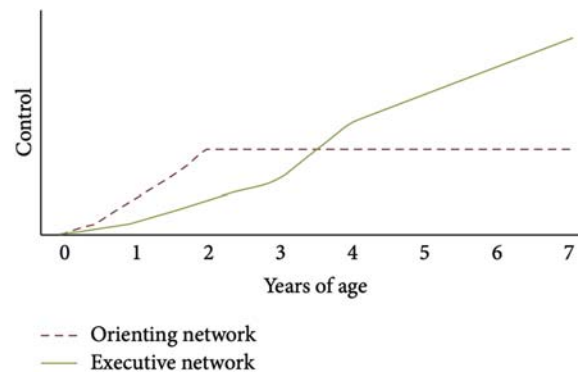


Fig. 4 The graph illustrates hypothetically how the orienting and the executive networks influence the development of self-control over the first seven years of life. From Posner, Rothbart, Sheese, and Voelker (2014, p. 2), reproduced under the Creative Commons Attribution License.

The second aspect, orienting, means that attention aligns with incoming sensory signals and that some information is selected. It is activated by posterior parts of the brain, mostly the parietal lobe (e.g., the superior parietal lobe, the temporal parietal junction, and the intraparietal sulcus) but also some frontal (frontal eye fields) and subcortical areas (pulvinar and superior colliculus). It is commonly assumed today that these areas represent two different control systems in the brain. One dorsal attention system involved in top-down visuospatial processes and one more ventral attention system more involved in bottom-up processes important for reorienting (Petersen and Posner, 2012). The main neurochemical component for the orienting systems is acetylcholine. “For newborn infants, the control of orienting is initially largely in the hands of caregiver presentations” (Posner et al., 2013, p. 400). By four months the infant has gained substantial voluntary control over his/her orienting responses and by late infancy self-regulatory capacities start to develop. A hypothetical relationship of how attention influences the development of self-control over the first seven years of life is illustrated in Fig. 4.

The final aspect, executive attention, indicates the ability to monitor responses, thoughts, and feelings, a process that not only involves mainly anterior brain areas (prefrontal cortex and the anterior cingulate gyrus) but also parts of the basal ganglia (e.g., striatum). Biologically, two separate brain systems have been identified that together seem to control executive attention (Petersen and Posner, 2012), one important for the moment-to-moment task requirements (the frontoparietal system) and one that is central for sustained attention (the cingulo-opercular system). It is hypothesized that executive attention plays an important role in the ability of the developing self to regulate positive and negative affect. Dopamine is the main neurochemical modulator although serotonin is also involved.

The more posterior parts of the attentional system develop relatively early. Thus, the young infant is more able to alert and orient than to show executive attention, although it takes time for these systems to become fully mature as well. The first signs of the executive system can be noted toward the end of the first year (e.g., as in the A-not-B task), but it will take about 8–10 years for it to reach adult-like levels although response time measures indicate continuous improvement into late adolescence. An additional finding relevant for understanding development is that connections between brain areas change over time. The young brain relies more on short connections in comparison to adults who have networks that includes long connections. Brain systems become more separate and functionally focused during development.

All three networks – especially the executive network – are essential for the infant to develop an ability to display effortful control, that is, to self-regulate one’s own behavior. In fact, the high-level temperamental construct effortful control, has been used as a proxy in studies on executive function in infancy. Effortful control reflects how well a child can suppress a major impulse in order to act on a second, less strong impulse. Temperamentally, effortful control builds on measures of the ability to control impulses, to shift attention and to focus attention (Putnam et al., 2006).

Attention and developmental disorders

Problems with attention are probably of importance for various neuropsychiatric disorders (Posner et al., 2019). As an example, the hyperactivity and concentration difficulties observed in ADHD have been found to be associated with the alerting network while disengagement deficits often noted in autism has been associated with the orienting network. Finally, the executive network has been associated with recurrent thoughts and repetitive behaviors often problematic for people with OCD. It is however important to keep in mind that attention is only one of many aspects of a disorder and also that, within a disorder, the relevant attention network can change over time. It has, for instance, been suggested that for ADHD, problems with alerting are most evident early in life while problems with the executive attention network become prominent later. It is also likely that attention problems might be important for reading difficulties or reading processes. A recent study showed that fluent reading in 8–12-year-old children was associated with functional connectivity between the orienting and executive networks. High fluency and high connectivity were found to be significantly associated. This held for both typical readers and for children with reading difficulties (Freedman et al., 2020).

Genes and attention

The observation that each of the three networks are associated with different neuromodulators has made it possible to investigate possible genetic links underlying each network. Genes found to influence norepinephrine seems to influence primarily the alerting network but not orienting, whereas genes affecting the cholinergic modulator are important for the orienting network. For the executive network, genes linked to dopamine and serotonin have been identified. Although several genes have been linked to attention processes there is no one-to-one correspondence between genes and attention. The genetic influence is complex due to genetic polymorphism, epigenetic factors, and environmental influences. The environment seems also to exert a strong influence, especially early in the development, through parenting practices, especially for the development of the executive network as shown through the child's ability to self-regulate or exert effortful control. Thus, the attentional abilities we measure probably rests on a strong gene $\times$ environment interaction.

The DRD4 gene, related to the dopamine system, is an illustrative example of how genes can influence attention, especially the executive network (Posner et al., 2013; Rothbart and Posner, 2015). A specific part of the DRD4 gene (the 7-repeat allele) has been linked to risk taking behavior and also to ADHD. This DRD4 7-repeat allele also interacts with environmental factors. In a study on parental training trying to decrease children's externalizing behavior, a positive effect was noted only for children carrying the 7-repeat allele.

Aspects of attention

Novelty attracts attention

Novel objects and novel locations affect attention from birth although the parallel process of familiarity also exerts a strong influence on attention during the first months of life. Behaviorally, novelty preference can be reliably observed from approximately 4–6 months of age. Identification of new information leads to faster and more reliable orienting, activates the alerting system, recruits available relevant brain resources (working memory and sustained attention), and prepares the system to encode new information that, in a second step, might be transferred to long-term memory.

The identification of a new object is governed by two processes. The first process to emerge is the ability to identify a new location. This ability is related to the control of eye movements and is, in some form, present at birth. It is observable through inhibition of return (IOR), which means that the visual system resists going back to a previous location; new locations are more attractive. The second process, the identification of a new object, is a slower-developing visual skill. It is usually established by 4–6 months and is related to the development of object recognition. This ability is observable through visual recognition memory.

Attention to faces and novel objects co-acts when, for instance, an adult demonstrates a new action with an object. In an eye-tracking study, Koch et al. (2018) found that 12- and 16-month-old infants shifted their attention from the presenter's face to the object as soon as the action was demonstrated for the first time. However, after repeated demonstrations, the infant's attention reverted to the presenter's face which suggest a dynamic relationship over time between attention to the action and the face. This reorientation to the face might "be driven by social interest and attempts to understand the presenter's intentions" (Koch et al., 2018, p 197).

Memory and attention

From 3 to 4 months and onward, the infant's ability to process new information has proved to be robust. This can be shown in paired-comparison experiments or tests when the child's preferential looking pattern is observed. Typically, a child prefers new information (a novel stimulus) in comparison with familiar information, which makes it possible for the researcher to calculate a novelty preference score. This score provides an index of the child's visual recognition memory, since the child must be able to compare the new information with the remembered familiar target in order to show a preference for the novel one (Morgan and Hayne, 2011). Measures of novelty preference (and therefore visual recognition memory) in early infancy have been found to be a significant predictor of intelligence, nonverbal communication, and language (e.g., Rose et al., 2012). As an example, it has been reported that visual recognition memory measured in infancy predicts receptive language at 3 years even when controlling for general IQ. These results suggest that attentional and memory capacities tapped by visual recognition memory as measured by novelty preference are important for later developing communication skills.

Visual recognition memory seems to tap into the very basic functions of our nervous system but exactly how these early attentional and memory processes affect later cognitive processes is still unexplained. One clue to an explanation has been proposed by Colombo and collaborators (2001) who found that, at 4 months, attention termination explained more of the observed variance in a novelty recognition task than sustained attention. The research group interpreted their findings as supporting the "hypothesis that individual differences in the disengagement underlie the relation between look duration and cognitive performance in early to mid-infancy" (Colombo et al., 2001, p 1605).

Pupil dilation and attention

Infants visual processing time is related to how well they learn and remember. This means that longer looking times should be related to increased memory for the information the infant is attending to, which is also supported by empirical research. However,

apart from time, infants' quality of attention is also related to learning, which can be assessed through pupil dilation. The human pupil dilates in response to changes in luminance, but an increase in pupil size is also evident when people experience internal arousal, conduct mentally challenging tasks, and when processing a stimulus that is novel and motivating. This makes pupillometry an appealing complementary measure in infancy research on visual processing. Looking time and pupil dilation were compared in an imitation study with infants (Sonne et al., 2016). The infants saw a three-step action on a screen and were tested on delayed recall two weeks later. Only pupil dilation showed a correlation with memory performance, suggesting that the quality of attention might be a better measure compared to looking time, if we want to understand the relation between attention and learning.

Attention and learning

The directed-attention model

The memory span for objects is around one item at 6 months and two to three items at 1 year. This means that infant learning and infants' complex social responses are carried out with a memory and attention system that is very limited compared to older children and adults. A model for how this can be achieved has been proposed by Reid and Striano (2007, see also Michel et al., 2015) who outline a five-stage directed-attention model of infant social cognition. The model describes five perceptual stages that infants typically will master within the first year of life. The model provides an attempt to integrate known perceptual and cognitive abilities and skills with the amazing social competence seen in the human infant. In brief, the proposed stages or phases are:

- 1 *Detection of socially relevant organisms.* Infants are born with a nervous system that directs them toward the social world. They are sensitive to distinctions between animate and inanimate objects, they prefer moving stimuli to static, biological motion to nonbiological motion, they are sensitive to the human voice, to rhythm, and they imitate (mimic) facial gestures.
- 2 *Identification of socially relevant organism.* The process to differentiate individual persons starts immediately at birth. The newborn infant rapidly learns to identify the mother's voice, smell, and face. More complex responses, such as imitation, are also used early on (from 6 weeks) to identify persons.
- 3 *Assessment of the locus of attention.* Once the socially relevant organism has been detected, infants start to "attend toward characteristic that index the locus of attention of the observed organism" (Reid and Striano, 2007, p 105). Eye and head movements provide important information to the infant in early social interactions. The infant's predisposition to enter into the social interactions with the caregiver is probably driven by intrinsic and biologically based motives to communicate with other humans.
- 4 *Detection of object-oriented attention.* Already at 4 months infants can use an adult's gaze to learn about objects and, by 8–9 months, objects become parts of highly motivating joint-attention encounters. Objects or an aspect of the environment that has been highlighted through joint attention with an adult will become more salient to the infant who will direct more of his or her attention to those objects/areas. The first sign of an emerging declarative memory is observed at 6 months through the child's ability to act on a memorized representation after a delay (deferred imitation).
- 5 *Inference of goals and/or prepare response.* Toward the end of the first year, infants begin to respond differently to accidental and intentional action. The child now starts to understand that people have goals that motivate their actions. The capacity of working memory increases, which makes it possible for the child to hold more than one piece of information online simultaneously.

Dyadic attention

The interaction between a parent and an infant is characterized by rhythm, intimacy, and emotional exchange, the so-called proto-conversations. This interaction is encouraged by the parent through smiles and increased gaze in such a way that the interaction is prolonged and infants are active partners in this interaction, creating turn-taking sequences. Trevarthen (2011) suggests that infants from an early age have the capacity to engage in communicative acts and to respond to the communication of others. In these early interactions, the infant and the parent are attuned to the subjective state of the other. Thus, the infant is aware of the parent's attention and is also responding to this attention. Trevarthen describes this awareness as primary intersubjectivity, which is supported by the parent by responding to the social initiatives from the infant. Both parent and infant are sensitive to the contingency and quality of this interaction and will soon develop expectancies of specific patterns of communication from each other. This is evidenced by using still-face conditions where the parent is either not responding to the infant's communication or by using a double-video-technique showing a pre-recorded interaction creating unsynchronized turn-taking from both partners. Furthermore, infants very early learn to anticipate, not only contingent communicative patterns, but also contingent behavior from their parents. When an infant's distress is accompanied by a parent's soothing, the association between arousal, the parent's response, and subsequent relief is easily learned – soothing is anticipated by the infant already at 4 months of age.

This intrinsic communication with human beings could not be enhanced without the infant exploring the physical environment. With an increasing ability of motor control, the child will experiment with its own actions on objects such as mobiles and thereby develop an early understanding of agency according to physical cause and effect. With the growth of contingent perception, the infant is able to accompany this physical agency into detection of cause and effect/reciprocity with partners into further developed nonverbal turn-taking sequences – both vocal and with objects (like pushing a ball back and forth). This means that infants are able to coordinate their attention with another person and understand the framework in which communication takes place between two persons.

Joint attention

Nine-month-old infants begin to understand that actors are pursuing goals and they combine the awareness of outside objects, events, and persons in order to share and coordinate their attention or perception of goal activities, the so-called joint attention. By performing joint activities such as building a tower of bricks or rolling a ball back and forth, infants start to understand the concept of sharing goals. This is accompanied by attention skills such as following the other persons' eye-gaze or pointing in order to direct others attention to objects in the surrounding environment or trying to modify a persons' behavior with gestures.

The earliest sign of triadic attention is seen when infants start to follow others' gaze toward external objects or events. Through gaze-following the infant can identify what other people are attending to and can begin to compare their own visual information with other person's attention. Gaze-following in early infancy has been suggested to lay the foundation for both referential communication and social cognitive development ([Brooks and Meltzoff, 2015](#)).

Attention skills that are initiated by the child (e.g., pointing) develop slightly later than behaviors that are responses to others (e.g., gaze-following). The capacity to direct another persons' gaze to objects by the infant's own interest develops between 9 and 12 months and infants can use the pointing gesture for different purposes, for both sharing attention and requesting. There is an interactive relationship between infant behavior and caregiver response in the development of pointing. Research shows that an infant's understanding of another's pointing gesture predicts their own production of pointing, suggesting that pointing is learned through observing and imitating the gesture in others. However, when infants begin to use pointing, their parents increase their own use of pointing, suggesting that the infant's ability to direct others' attention also influences their social and communicative environment ([Ger et al., 2018](#)).

It has been argued that declarative pointing, in contrast to imperative pointing, relies on the understanding of others as mental agents and is driven by a motivation to share attention and interest with other persons; it has therefore been suggested to be especially difficult. In typically developing infants, the motivation to share attention and interest is probably strong and declarative gestures are common, in contrast to children with social impairments like autism. An experiment with 12-month-olds has revealed that the social context is crucial for the number of pointings the infant makes ([Liszkowski et al., 2004](#)). Only when the adult was active in sharing the infant's attention to the event the infant pointed at, did the infants' pointing increase. The interpretation of this result was that the infant did not only want to direct the adult's attention, but also wanted to share this attention.

It has been shown in several studies that the capacity for joint attention is an important precursor to later-developing language and cognitive skills. Together with early memory measures (e.g., visual recognition memory and deferred imitation) joint attention probably lays the ground for later-emerging social cognition including intentional understanding. A study by [Brandone et al. \(2020\)](#) investigated this relationship longitudinally. They showed that the tendency to engage in joint attention episodes around 6 months of age predicted the infants' understanding of intentional actions three months later.

Intention requires attention

Humans are special in their capacity to understand others' intentions. It can be argued that infants, by the time they begin to follow and direct others' attention, have acquired some understanding of others as intentional agents and that other people act on the basis of their own view of the world. From 6 months of age, infants follow another person's gaze to objects in the surrounding environment, but they have been shown to pay more attention to where the head is turning, whereas older infants pay more attention to the eyes indicating understanding of the adult's intentions. When adults turn to an object with their eyes open or closed it was observed that infants by the age of 12 months looked at the target if the adult turned to it with open eyes but did not do so if the person turned to it with eyes closed ([Meltzoff and Brooks, 2007](#)).

Furthermore, studies have revealed that infants during their second year start to understand the intention behind an action and not only the action they actually have seen; when 18-month-old infants were observing an action that the adult failed to perform they did not imitate the failure, instead they performed the complete action ([Meltzoff, 1995](#)). Using anticipatory looking, studies have revealed that infants as young as 10 months differentiate between failed and successful actions ([Brandone et al., 2014](#)). The infants saw an actor repeatedly reach for a ball and either retrieved it or not. Ten-month-olds (but not 8-month-olds) looked at the goal of the action (the ball) both when the actor successfully reached it or when the actor failed, indicating that they interpreted the actor's intention and not the outcome.

Decades of research show that children are around 4 years of age when they understand that others' mental states can differ from their own and that another person's behavior is based on their beliefs, even if the child knows they are incorrect. More recent studies have suggested that some understanding of this "false belief" might be present already in infancy. These studies all use nonverbal methods, such as violation-of-expectation. According to the violation-of-expectation paradigm, infants pay more attention to situations that are surprising to them or contradict their expectations of how other people will act. Although numerous studies have found similar results using different nonverbal methods, the specific findings have also been difficult to replicate and, to date, there is still debate about the scope and limitations of the social cognitive understanding in early infancy (see [Poulin-Dubois et al., 2018](#)).

Digital media and attention

The fact that attention depends on different systems that develop over time raises obvious questions about how the new digital world influences children over the course of development ([Courage, 2017](#)). The newborn capacity for attention is limited and is to a large extent in the hands of the parents. However, some capacity to orient and attend do exist early which will help the infant in his/her socio-cognitive development. An important aspect is that the executive attention network develops relatively slow. That is,

the abilities important for controlling one's own attention, to switch attention between events, and to stop attending (attention termination) are difficult tasks for the very young child.

Even if digital media is attractive there is no scientific support for the idea that children below the age of three learn better or faster via digital media (see [Barr and Linebarger, 2017](#)). In fact, research has shown that children below that age actually need more exposure to take in new information compared with a real 3D-world situation when an adult can scaffold attention by creating an interaction that supports joint attention.

Another aspect of the new digital world relevant to attention processes are the popular action video games that often require fast switching between tasks, keeping strategies in mind, and online planning. These are tasks that, besides involving many important brain areas, to a large extent rely on the attention networks described in this chapter. Some studies have found that the type of multi-tasking required by action video games have positive effects on some vision skills (e.g., tracking peripheral objects) as well as on reaction times ([Rothbart and Posner, 2015](#)). However, the findings are not clear cut. It seems as if the positive findings are more related to the ability to orient whereas training effects of the executive network are less certain.

Compensatory systems

Questions have been raised as to whether the sensory system can compensate for a deficit in one area, since it is striking how blind people can use auditory and tactile cues for orienting themselves and getting information. In one study ([Forster et al., 2007](#)), it was shown that spatial tuning of tactile attention is more accurate in early blind individuals compared to sighted individuals when areas for Braille reading are stimulated, suggesting a compensatory scaffolding for individuals who have experienced visual deprivation from birth or early infancy. The plasticity of the brain, which makes it possible to overcome deficits, is underlined by the fact that the sensory systems work together, especially, according to the dimensions of space, time, and intensity. This so-called redundant, amodal information is seen as a cornerstone of perceptual development. For example, when an adult takes a child's hands and uses them for clapping, auditory, visual, and tactile systems are involved, discerning both rhythm and rate to the child.

Final remarks

With increasing age, looking and attention come more and more under voluntary control. The child becomes more able to decide what to focus on; he or she will, by 4–5 years of age, be able to choose to attend to information even if it is boring. However, attention never becomes completely voluntary. Processes like novelty preference and habituation influence attention throughout life.

An increasingly important area for future research is to create a better understanding of how digital media influence children's development from infancy to adolescence, to adulthood. To what degree is media helpful and to what degree might it create or hinder a child's attention abilities? This requires deeper understanding, not only of brain development, but also of psychological, sociological, and societal facts that all might influence the use or abuse of digital media.

In the future, we will hopefully also develop more specific knowledge about specific genes and their impact on attention and on gene $\times$ environment processes (see [Posner et al., 2014](#)). This might help us gain a better understanding of how attention develops, how individual differences in attention capacities evolve, how experience and biology co-act to create alertness and sustained attention, or the ability to initiate attention in another person – knowledge that will also be highly relevant for both our understanding of typical development as well as for children with various disabilities such as ADHD or autism.

Acknowledgment

The authors wish to acknowledge the late professor Tomas Tjus who co-authored a previous version of this paper (published in *IEE* third Ed). Although the text is thoroughly updated it still owes much in scope and rigor to his detailed and insightful work. The work of M.H. was supported by grants from the Marcus and Amalia Wallenberg Foundation, Sweden (MAW 2018.0084) and the Swedish Research Council (Grant no.: 2016-01033).

References

- Amos, D., Johnson, S.P., 2006. Learning by selection: visual search and object perception in young infants. *Dev. Psychol.* 42, 1236–1245.
- Barr, R., Linebarger, D.N. (Eds.), 2017. *Media Exposure during Infancy and Early Childhood: The Effects of Content and Context on Learning and Development*. Springer Nature, Cham.
- Bjorklund, D.F., 2021. *How Children Invented Humanity: The Role of Development in Human Evolution*. Oxford University Press, New York.
- Brandone, A.C., Horwitz, S.R., Aslin, R.N., Wellman, H.M., 2014. Infants' goal anticipation during failed and successful reaching actions. *Dev. Sci.* 17, 23–24.
- Brandone, A.C., Stout, W., Moty, K., 2020. Triadic interactions support infants' emerging understanding of intentional actions. *Dev. Sci.* 23, e12880.
- Brooks, R., Meltzoff, A.M., 2015. Connecting the dots from infancy to childhood: a longitudinal study connecting gaze following, language, and explicit theory of mind. *J. Exp. Child Psychol.* 130, 67–78.
- Buiatti, M., Di Giorgio, E., Piazza, M., Polloni, C., Menna, G., Taddei, F., Baldo, E., Vallortigara, G., 2019. Cortical route for facelike pattern processing in human newborns. *Proc. Natl. Acad. Sci. U. S. A.* 16, 4625–4630.
- Butcher, P.R., Kalverboer, A.F., Geuze, R.H., 2000. Infants' shifts of gaze from central to peripheral stimulus: a longitudinal study of development between 6 and 26 weeks. *Infant Behav. Dev.* 23, 3–21.

- Colombo, J., Salley, B., 2015. Biopsychosocial perspectives on the development of attention in infancy. In: Calkins, S.D. (Ed.), *Handbook of Infant Biopsychosocial Development*. Guilford Press, New York, pp. 71–96.
- Colombo, J., Richman, W.A., Shaddy, J., Greenhoot, A.F., Maikranz, J., 2001. Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Dev.* 72, 1605–1616.
- Courage, M.L., 2017. Screen media and the youngest viewers: implications for attention and learning. In: Blumberg, F.C., Brooks, P.J. (Eds.), *Cognitive Development in Digital Contexts*. Academic Press, London, UK, pp. 3–28.
- Conte, S., Richards, J.E., Guy, M.W., Xie, W., Roberts, J.E., 2020. Face-sensitive brain responses in the first year of life. *NeuroImage* 211, 116602.
- Forster, B., Eardley, A.F., Eimer, M., 2007. Altered tactile spatial attention in the early blind. *Brain Res.* 1131, 149–154.
- Freedman, L., Zivan, M., Farah, R., Horowitz-Kraus, T., 2020. Greater functional connectivity within the cingulo-opercular and ventral attention networks is related to better fluent reading: a resting-state functional connectivity study. *NeuroImage Clinical* 26, 102214.
- Ger, E., Altinok, N., Liszkowski, U., Küntay, A.C., 2018. Development of infant pointing from 10 to 12 months: the role of relevant caregiver responsiveness. *Infancy* 23, 708–729.
- Johnson, M.H., Senju, A., Tomalski, P., 2015. The two process theory of face processing: modifications based on two decades of data from infants and adults. *Neurosci. Biobehav. Rev.* 50, 169–179.
- Koch, F.-S., Sundqvist, A., Herbert, J., Tjus, T., Heimann, M., 2018. Changes in infant visual attention when observing repeated actions. *Infant Behav. Dev.* 50, 189–197.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., Tomasello, M., 2004. Twelve-month-olds point to share attention and interest. *Dev. Sci.* 7, 297–307.
- Meltzoff, A.M., 1995. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* 31, 838–850.
- Meltzoff, A.M., Brooks, R., 2007. Eyes wide shut: the importance of eyes in infant gaze-following and understanding other minds. In: Flom, R., Lee, K., Muir, D. (Eds.), *Gaze-following: Its Development and Significance*. Lawrence Erlbaum, Mahwah, pp. 217–241.
- Michel, C., Stets, M., Parise, E., Reid, V.M., Striano, T., Hoehl, S., 2015. Theta- and alpha-band EEG activity in response to eye gaze cues in early infancy. *NeuroImage* 118, 576–583.
- Morgan, K., Hayne, H., 2011. Age-related changes in visual recognition memory during infancy and early childhood. *Dev. Psychobiol.* 53, 157–165.
- Petersen, S.E., Posner, M.I., 2012. The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* 35, 73–89.
- Posner, M.I., Rothbart, M.K., Ghassemzadeh, H., 2019. Restoring attention networks. *Yale J. Biol. Med.* 92, 139–143.
- Posner, M.I., Rothbart, M.K., Ghassemzadeh, H., 2020. Developing attention in typical children related to disabilities. In: Gallagher, A., Bulteau, C., Cohen, D., Michaud, J.L. (Eds.), *Handbook of Clinical Neurology. Neurocognitive Development: Normative Development*. Elsevier, Amsterdam, Netherlands, pp. 215–223.
- Posner, M.I., Rothbart, M.K., Rueda, M.R., 2013. Developing attention and self-regulation in infancy and childhood. In: Rubenstein, J.L.R., Rakic, P. (Eds.), *Neural Circuit Development and Function in the Brain*. Academic Press, New York, pp. 395–411.
- Posner, M.I., Rothbart, M.K., Sheese, B.E., Voelker, P., 2014. Developing attention: behavioral and brain mechanisms. *Adv. Neurosci.* Article ID 405094.
- Poulin-Dubois, D., Rakoczy, H., Burnside, K., et al., 2018. Do infants understand false belief? We don't know yet – a commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognit. Dev.* 48, 302–308.
- Putnam, S.P., Gartstein, M.A., Rothbart, M.K., 2006. Measurement of fine-grained aspects of toddler temperament: the early childhood behavior questionnaire. *Infant Behav. Dev.* 29, 386–401.
- Reid, V.M., Striano, T., 2007. The directed attention model of infant social cognition. *Eur. J. Dev. Psychol.* 4, 100–110.
- Rennels, J.L., Kayl, A.J., Kulhanek, K.M., 2020. Individual differences in infants' temperament affect face processing. *Brain Sci.* 10, 474.
- Rose, S.A., Feldman, J.F., Jankowski, J.J., van Rossem, R., 2012. Information processing from infancy to 11 years: continuities and prediction of IQ. *Intelligence* 40, 445–457.
- Rothbart, M.K., Posner, M.I., 2015. The developing brain in a multitasking world. *Dev. Rev.* 35, 42–63.
- Simion, F., Di Giorgio, E., 2015. Face perception and processing in early infancy: inborn predispositions and developmental changes. *Front. Psychol.* 6, 969.
- Sonne, T., Kingo, O.S., Krøigaard, P., 2016. Empty looks or paying attention? Exploring infants' visual behavior during encoding of an elicited imitation task. *Infancy* 21, 728–750.
- Trevarthen, C., 2011. What is it like to be a person who knows nothing? Defining the active intersubjective mind of a newborn human being. *Infant Child Dev.* 20, 119–135.

Further reading

- Adamson, L.B., Bakeman, R., Suma, K., Robins, D.L., 2019. Sharing sounds: the development of auditory joint engagement during early parent-child interaction. *Dev. Psychol.* 55, 2491–2504.
- Bornstein, M.H., 2014. Human infancy...and the rest of the lifespan. *Annu. Rev. Psychol.* 65, 121–158.
- Burgoyne, A.P., Engle, R.W., 2020. Attention control: a cornerstone of higher-order cognition. *Curr. Dir. Psychol. Sci.* 29, 624–630.
- Colombo, J., 1993. *Infant Cognition: Predicting Later Intellectual Functioning*. Sage, Newbury Park.
- Cubit, L., Canale, R., Handsman, R., Kidd, C., Bennetto, L., 2021. Visual attention preference for intermediate predictability in young children. *Child Dev.* 92, 691–703.
- Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., Cowan, N., 2000. The development of auditory attention in children. *Front. Biosci.* 5, 108–120.
- Hoehl, S., Peykarjov, S., 2012. The early development of face processing - what makes faces special? *Neurosci. Bull.* 28, 765–788.
- Johnson, S.P., 2010. *Neuroconstructivism: The New Science of Cognitive Development*. Oxford University Press, New York.
- Johnson, M.J., Morton, J., 1991. *Biology and Cognitive Development: The Case of Face Recognition*. Blackwell, Oxford, UK.
- Posner, M.I., 2012. *Attention in a Social World*. Oxford University Press, New York.
- Posner, M.I., Rothbart, M.K., Voelker, P., 2016. Developing brain networks of attention. *Curr. Opin. Pediatr.* 28, 720–724.
- Rennels, J.L., Kayl, A.J., 2020. Infants and adults represent faces differently. *Dev. Psychol.* 56, 1–14.
- Reynolds, G.D., Romano, A.C., 2016. The development of attention systems and working memory in infancy. *Front. Syst. Neurosci.* 10, 15.
- Salley, B., Colombo, J., 2016. Conceptualizing social attention in developmental research. *Soc. Dev.* 25, 687–703.
- Spencer, J.P., 2020. The development of working memory. *Curr. Dir. Psychol. Sci.* 29, 545–553.
- Unsworth, N., Miller, A.L., 2021. Individual differences in the intensity and consistency of attention. *Curr. Dir. Psychol. Sci.* 30, 391–400.

Cognitive load theory

John Sweller, School of Education, University of New South Wales, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	127
Human cognitive architecture	127
Evolutionary educational psychology	127
Generic-cognitive knowledge and domain-specific knowledge	128
Instructional implications of the distinction between biologically primary, generic-cognitive skills and biologically secondary, domain-specific skills	128
Processing biologically secondary information	129
The borrowing and reorganizing principle	129
The randomness as genesis principle	129
The narrow limits of change principle	129
The information store principle	129
The environmental organizing and linking principle	129
Instructional consequences of human cognitive architecture	130
Some cognitive load effects	130
The worked example effect	130
The problem completion effect	130
The goal-free effect	130
The redundancy effect	130
The element interactivity effect	131
The expertize reversal effect	131
The guidance fading effect	131
The split-attention effect	131
The self-management effect	131
The modality effect	131
The transient information effect	132
The variability effect	132
The isolated elements effect	132
The imagination effect	132
The collective working memory effect	132
The human movement effect	132
The spacing effect	132
Conclusion	132
References	133

Introduction

Cognitive load theory can be subdivided two complementary aspects: A cognitive architecture based on (a) our knowledge of human cognition and evolutionary psychology and (b) a series of cognitive load effects based on instructional hypotheses generated by the cognitive architecture and tested using randomized, controlled experiments. The ultimate aim of both the cognitive architecture and the data is to provide prescriptive, instructional procedures. The cognitive architecture is discussed first.

Human cognitive architecture

While human cognitive architecture is immensely complex, only parts of our cognitive system are directly relevant to instructional issues and so it is those parts with which we are primarily concerned. Evolutionary psychology can be used to provide one of the dividing lines between categories of information.

Evolutionary educational psychology

Evolutionary educational psychology can be used to divide knowledge into biologically primary knowledge, which is a category that is less relevant to instructional issues, and biologically secondary knowledge, which is a category that is the subject of most instruction (Geary, 2008, 2012; Geary and Berch, 2016).

We have evolved to acquire biologically primary information. It is acquired effortlessly and unconsciously. Examples are learning to listen to and speak a native language, acquire general problem-solving and thinking skills, store knowledge for future use, and use that knowledge to plan future events. Because primary knowledge and skills are acquired unconsciously, they do not need to be taught and indeed, for that reason they cannot be taught because they will already have been acquired when needed. Primary knowledge and skills are modular with no relations between many of them. For example, our ability to generalize a skill from one set of conditions to another may be unrelated to our skill in planning future events. Different primary skills with distinct, unrelated functions are likely to have evolved during different evolutionary epochs.

Secondary knowledge is acquired for cultural reasons and differs from society to society. Unlike primary knowledge, it is not acquired automatically without conscious effort on the part of the learner. It also is unlikely to be acquired without deliberate instruction. Examples of biologically secondary information can be found in virtually all curricula of all educational institutions. Schools, as well as all other educational institutions, were developed precisely to teach biologically secondary information. Without such institutions, only tiny proportions of any general population will acquire the knowledge imparted by educational institutions. For example, reading and writing were invented several thousand years ago but only became common in some populations at the beginning of the last century with the advent of mass education. Without mass education, the majority of any general population would remain illiterate. The contrast with listening and speaking could not be more stark. Every developmentally regular child will learn the immensely complex tasks of listening and speaking their native language automatically, without formal instruction, because humans have evolved to do so since the advent of the human species. Indeed, most of us would have no idea how to teach children how to organize their tongues, lips, breath and voice in order to speak. We do not need to teach this skill because it is biologically primary and so acquired automatically. We do need to teach children how to write along with all of the other skills taught in education and training institutions because such skills are biologically secondary, considered important for our societies, and otherwise would not be acquired.

Generic-cognitive knowledge and domain-specific knowledge

There is another dividing line providing categories of information. That categorization scheme overlaps heavily with biologically primary and secondary information. Generic-cognitive knowledge leads to very general skills such as general problem-solving skills, thinking skills, or self-regulatory skills. They can be contrasted with domain-specific skills that apply to a very limited set of circumstances. For example, using means-ends analysis to solve a wide variety of diverse problems (Newell and Simon, 1972) provides an example of a generic-cognitive skill that can be contrasted with learning how to solve a limited range of algebra problems by multiplying out a denominator. (A means-ends problem-solving strategy involves considering a current problem state, comparing it with a problem goal or a sub-goal state, extracting differences between the two states, and lastly, attempting to find problem-solving operators that reduce differences between the two states.)

Because of their scope, generic-cognitive skills are much more important than domain-specific skills. Everyone needs to know how to solve problems using means-ends analysis in order to survive but can, of course, survive without knowing how to multiply out a denominator in order to solve a problem. Arguably, the importance of generic-cognitive skills is such that we have evolved to acquire them automatically and unconsciously—they are biologically primary. In contrast, we have not specifically evolved to acquire domain-specific skills such as multiplying out a denominator in an algebraic equation. Domain-specific skills must be explicitly taught and consciously learned as biologically secondary skills (Tricot and Sweller, 2014).

Instructional implications of the distinction between biologically primary, generic-cognitive skills and biologically secondary, domain-specific skills

As indicated above, biologically primary, generic-cognitive skills are central to human functioning. It is essential for us to be able to regulate our mental processes or solve novel problems that we encounter. In contrast, we can survive if we have not acquired biologically secondary, domain-specific information such as how to manipulate algebraic equations, or if we do not know the causes of World War 1. Because of this differential level of importance, there sometimes is a tendency to place a greater emphasis on the acquisition of generic-cognitive, primary skills than domain-specific, secondary skills (Tricot and Sweller, 2014). That emphasis can be seen in the labeling of generic-cognitive skills as “21st Century skills” with the implication that, for example, explicitly teaching students thinking skills should be part of modern curricula.

Nevertheless, despite the critical importance of generic-cognitive skills, if they are acquired automatically and unconsciously, then attempts to teach them are unnecessary and will inevitably fail in the sense that people taught such skills will show no greater ability to use them on novel, transfer tasks than people not provided instruction. (A novel transfer task is one that is completely different from the task used to teach the generic-cognitive skill. It needs to be different to provide evidence that any improvement really is due to the acquisition of the generic-cognitive skill and not due to the acquisition of domain-specific information.) The failure late last century and early this century of attempts to teach generic problem-solving skills is one of a long list of such failures. For example, the ground-breaking work of Newell and Simon (1972) used sophisticated computational models to establish means-ends analysis as a critical problem-solving strategy. Since that time, as far as I am aware, there has not been a single study demonstrating enhanced problem-solving skill following tuition in the use of the strategy. We all need to use the strategy in order to survive and so tuition in the use of the strategy is pointless. We have all evolved to acquire this complex, sophisticated, generic-cognitive

strategy without instruction. Of course, if a body of evidence indicating successful instruction of generic-cognitive skills becomes available, theory will need to follow the data.

Processing biologically secondary information

The biologically primary and secondary systems are closely connected with little or no secondary knowledge acquired without simultaneous use of the primary system. There is a specific cognitive architecture consisting of several biologically primary cognitive processes and structures associated with acquiring biologically secondary, domain-specific information. That architecture is closely analogous to the architecture that governs the processes of evolutionary biology. Both evolutionary biology and the structures and functions associated with the acquisition of biologically secondary information constitute examples of natural information processing systems (Sweller and Sweller, 2006). When applied to human cognition, that information processing system has instructional implications that are central to cognitive load theory (Sweller et al., 2011, 2019). The cognitive architecture required to process biologically secondary information can be described by five basic principles associated with (a) the acquisition of novel information represented by the borrowing and reorganizing principle and the randomness as genesis principle, (b) the organizing and storing of that information for later use represented by the narrow limits of change principle and the information store principle, and (c) the retrieval of that information when required, represented by the environmental organizing and linking principle.

The borrowing and reorganizing principle

In order to adequately function in most societies, humans need to obtain vast amounts of biologically secondary information. The vast bulk of that information is obtained from other people by listening to what they say or reading what they write. That information can be reorganized to accord with currently held information and stored for later use. Education and training systems were invented to facilitate this process. We have evolved to obtain information from others so while the information obtained can be biologically secondary, the principle itself is primary. We are one of the few species that naturally obtains information from other members of the species. Like all evolved skills, we vary in our ability to obtain information from other people.

The randomness as genesis principle

Information from others is not always available. When unavailable from others, information can be generated during problem solving using a random generate and test procedure. We do not need to teach people how to randomly generate moves, test them for effectiveness, and retain effective moves while rejecting ineffective ones. This general problem-solving ability is again, a biologically primary, generic-cognitive skill that does not need to be taught because it is acquired automatically. It is an essential process in the initial generation of knowledge. Without it, there would be no knowledge to acquire using the borrowing and reorganizing principle discussed above.

The narrow limits of change principle

All novel information initially must be processed by a limited capacity, limited duration working memory before it can be stored in long-term memory. We cannot hold more than about 7 items in working memory at a time (Miller, 1956), process more than about 3 or 4 items at a time (Cowan, 2001) where processing involves organizing, comparing or working on items, and cannot hold items in working memory for more than about 20 s without rehearsal (Peterson and Peterson, 1959).

The information store principle

To be subsequently useful, information acquired by the borrowing and reorganizing principle and the randomness as genesis principle must be stored. Very large amounts of information acquired by one of these two principles can be permanently stored in long-term memory during learning. There are no known limits to the amount of information that can be stored in long-term memory. The importance of this principle to activities such as problem solving became apparent following the work of De Groot (De Groot, 1965; De Groot and Gobet, 1996) who found that acquired problem-solving skill in chess came from the storage in long-term memory of many thousands of chess-board piece configurations taken from real games. Subsequent work demonstrated that the same storage of domain-specific information underlaid skill in a variety of academic areas (Chiesi et al., 1979; Egan and Schwartz, 1979; Jeffries et al., 1981; Sweller and Cooper, 1985).

The environmental organizing and linking principle

Once novel information has been processed by working memory and stored in long-term memory, it can be transferred back into working memory. Environmental signals inform us which organized information stored in long-term memory is required at any particular time. That information, transferred from long-term to working memory, can be used to govern actions appropriate to the environment that triggered the transfer. Unlike when dealing with novel information, working memory has neither capacity nor temporal restrictions when dealing with organized information transferred from long-term memory (Ericsson and Kintsch, 1995). In this way, learning transforms us. We can carry out actions that in the absence of the information stored in long-term memory would be impossible.

Instructional consequences of human cognitive architecture

Based on the above cognitive architecture, the purpose of instruction is to assist learners to transfer domain-specific, biologically secondary information from working to long-term memory. Once stored in long-term memory, the information can be transferred back to working memory to generate and govern appropriate action.

Some information naturally consists of numerous elements that need to be processed simultaneously because they interact. For such information, element interactivity is high resulting in a high intrinsic cognitive load. For example, most geometry materials consist of high element interactivity information. To understand a geometry theorem such as “the external angles of a triangle equal the addition of the opposite internal angles” and to use that theorem to solve problems, learners need to understand all of the elements that constitute a particular diagram and to simultaneously process the elements of any statements associated with the diagram and combine the elements of the diagram with the elements of the statements. The working memory load can be overwhelming due to a high intrinsic cognitive load. In contrast, learning the names of theorems and constructs also can be difficult because there are many to learn such as the “external angles of a triangle theorem” but the difficulty is not due to a heavy working memory load because the various names do not interact. The name of each theorem can be learned independently of the name of the other theorems. When learning the names of theorems, element interactivity is low resulting in a low intrinsic cognitive load (Chen et al., 2015, 2016a,b; Sweller, 2010).

Intrinsic cognitive load only can be changed by changing the nature of the information or changing the expertise of learners. Expertise is increased by altering the contents of long-term memory. The information store and the environmental organizing and linking principle alter what constitutes an element. With increasing expertise, information that previously consisted of multiple elements can be transformed into a single element that incorporates the element interactivity of sub-elements, thus reducing intrinsic working memory load.

One aim of instructional design is to optimize intrinsic cognitive load. If intrinsic cognitive load is too low, learners will not learn as much as they could while if it is too high, working memory may be overwhelmed resulting in a reduction of learning.

Extraneous cognitive load also can be changed by changing element interactivity (Sweller, 2010). Unlike intrinsic cognitive load that only can be changed by altering what is taught or by increasing expertise, extraneous cognitive load is imposed by instructional design procedures such as the way information is presented or the activities required of students and can be changed by altering those procedures. The aim is always to reduce element interactivity associated with extraneous cognitive load.

Some cognitive load effects

All of the empirical work associated with cognitive load theory concerns cognitive load theory effects that are demonstrated by using the theory to generate a new instructional design that either optimizes element interactivity associated with intrinsic cognitive load or reduces element interactivity associated with extraneous cognitive load. The effectiveness of a new design is tested against a more traditional procedure. Randomized, controlled experiments are used exclusively to demonstrate cognitive load theory effects. Several of the more important effects are discussed next.

The worked example effect

Learning can be facilitated more if learners study worked out problem solutions than if they attempt to solve the same problems themselves. The worked example effect is demonstrated when novice learners who are presented high element interactivity problems to solve obtain lower scores on a subsequent problem-solving test than learners presented the same problems as worked examples that they are asked to study (Cooper and Sweller, 1987). Conventional problem solving using a means-ends strategy (Newell and Simon, 1972) imposes a heavy, extraneous cognitive load that interferes with learning. Studying worked examples reduces that cognitive load.

The problem completion effect

The problem completion effect (van Merriënboer and Krammer, 1987) has a similar rationale to the worked example effect except that rather than presenting learners with a full worked example, they are presented with a partially completed worked example that they are asked to complete. Compared to full problems, learning is facilitated (van Merriënboer and Krammer, 1987).

The goal-free effect

Since conventional problem solving imposes a heavy extraneous cognitive load, alternative problem-solving strategies that reduce the cognitive load should be effective. Goal-free problems are structured so that rather than having a conventional goal such as, *What is the velocity of the object?* in a kinematics problem, the goal is changed to, *Calculate the value of as many variables as you can.* When a problem is worded in this way, it is impossible to use a means-ends strategy to reduce the difference between a current problem state and the goal state because the goal state is not specified, which reduces cognitive load. Goal-free problems facilitate learning (Sweller et al., 1983).

The redundancy effect

When worked examples or any other form of instruction includes essential information along with unnecessary information, learning is inhibited compared to the presentation of the essential information alone. A comparison of the same information

with and without redundancy yields the redundancy effect if less information improves test scores (Chandler and Sweller, 1991). There are different forms of redundancy. Some examples are presenting the same information in different modalities such as identical written and spoken text during PowerPoint presentations (Kalyuga et al., 2004), adding text to a self-explanatory diagram that does not require text to be fully understood (Chandler and Sweller, 1991), or adding cartoons to text. All have the same effect of depressing learning by having learners process additional, unnecessary information in limited working memory.

The element interactivity effect

All cognitive load effects rely on element interactivity being high. If element interactivity is low, an effect may disappear or even reverse. For example, the worked example effect only occurs when students are attempting to learn to solve difficult (high element interactivity) problems. The issues associated with element interactivity can be seen when the worked example effect appears to contradict another effect, the generation or testing effect. This effect occurs when learners presented a task such as learning a new vocabulary, learn more if they are asked to generate answers to questions on information with which they have some familiarity, rather than re-study the information. Chen et al. (2015, 2016a,b) obtained a conventional worked example effect with geometry learners studying worked examples performing at superior levels to learners solving the equivalent problems. In contrast, they obtained a reverse worked example effect (a generation effect) when tested on their knowledge of the names of the theorems and of the geometric shapes that they had to learn in order to solve the problems. Knowledge was increased more if they had to generate the names of the geometry theorems and shapes rather than re-studying those names. As indicated above, learning to solve a problem is high in element interactivity because each move is related to all of the other moves and so imposes a high working memory load. Each name of a geometry theorem or shape on the other hand, can be learned independently of all the other names.

The expertize reversal effect

This effect is a particular example of the element interactivity effect (Chen et al., 2017). As indicated previously, element interactivity can be altered by altering the information that is being processed in different tasks (Chen et al., 2015, 2016a,b). Element interactivity also can be altered by altering the levels of expertize of learners. The same information processed by someone with a higher level of expertize will have a lower level of element interactivity due to the environmental organizing and linking principle. According to that principle, large amounts of information that incorporate interacting elements can be transferred as a single element from long-term to working memory. As a consequence of these changing levels of element interactivity associated with expertize, effects such as the worked example effect can disappear or reverse with increasing expertize. With higher levels of expertize, solving problems may result in more learning than studying worked examples—a reverse worked example effect (Kalyuga et al., 2001).

This reversal associated with increased expertize can occur with all cognitive load effects. It is due to the redundancy effect. For a novice, studying a worked example is an important way of acquiring knowledge to be stored in long-term memory. For a more expert learner in a particular domain, studying a worked example is a redundant activity. More expert learners already know how to solve the problem. They only need to practice solving it so that they can do so automatically (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977) without imposing an unnecessary load on working memory.

The guidance fading effect

Based on the expertize reversal effect, it follows that an effective instructional sequence should begin with full worked examples followed by guidance fading using completion problems and ending with full problems. There is considerable evidence for the effectiveness of this sequence (Renkl and Atkinson, 2003).

The split-attention effect

This effect (Tarmizi and Sweller, 1988) occurs when learners must split their attention between multiple sources of information in order to understand it. If for example, information includes a diagram and associated text that are unintelligible in isolation and so must be mentally integrated to be understood, learning will be facilitated by physically integrating them. Note that the relations between the two sources of information are critical to the split-attention effect. If, for example, the diagram is fully intelligible in isolation without the associated text, then the text is redundant and should be eliminated in line with the redundancy effect, rather than being integrated in line with the split-attention effect.

The self-management effect

While the split-attention effect indicates that physically integrating information facilitates learning, students will frequently come across information that is presented in split-attention format. Roodenrys et al. (2012) hypothesized and demonstrated that teaching learners to physically integrate split-attention information would be beneficial.

The modality effect

Under split-attention conditions where two sources of information such as a diagram and text are unintelligible until they have been mentally or physically integrated, rather than physically integrating them, to reduce cognitive load, the text can be presented in spoken rather than written form. The presentation of both auditory and visual information increases available working memory resources by using both auditory and visual processors. The use of dual modalities in this fashion increases learning compared to the use of the visual modality alone (Mousavi et al., 1995).

The transient information effect

Spoken information is transient. What we hear now continually disappears to be replaced, most commonly, by new information. Writing was invented to convert transient, spoken information into permanent information. The modality effect assumes that spoken information will be short and easily processed in working memory, thus transferring some of the load from the visual to the auditory processor. If the textual material is long and complex rather than being short and simple, it will impose less of a cognitive load when presented in permanent, written rather than transient, spoken form. This result was demonstrated by [Leahy and Sweller \(2011\)](#). The modality effect was obtained with short, simple textual material but a reverse effect with superior visual only information was obtained with long, complex material that was difficult to process in transient, spoken form. Similar differences were obtained using animated information which is similarly transient in nature by [Wong et al. \(2012\)](#).

The variability effect

The same information can be presented under a higher rather than lower variety of conditions. With higher variability, students might learn that a problem such as, $a/b = c$, solve for a , is identical to, $x/y = z$, solve for x , and to, $c = a/b$, solve for a . The greater the range of such problems that learners are exposed to, the greater the variability and the cognitive load imposed. The increase in cognitive load is intrinsic because learners must not only learn the mechanics of solving this category of problems, but also must learn to distinguish which problems belong to the category and which do not. [Paas and van Merriënboer \(1994\)](#) found that this increase in intrinsic cognitive load facilitated learning. Intrinsic cognitive load needs to be optimized, unlike extraneous cognitive load that requires reduction.

The isolated elements effect

Similar to the variability effect, this effect is due to optimizing intrinsic cognitive load except that the effect relies on a reduction rather than an increase in intrinsic cognitive load. Some highly complex information cannot be appropriately processed in working memory due to high levels of element interactivity. If a task includes a large number of interacting elements, learning may be enhanced by ensuring that the constituent elements first are entrenched in working memory by presenting them to students in isolation before they are presented in full, interacting form ([Pollock et al., 2002](#)). Presenting interacting elements in isolated form initially reduces intrinsic cognitive load to a level that allows processing.

The imagination effect

If learners are asked to imagine a procedure rather than merely study it, they may learn more. This effect is particularly prone to expertize reversal. To obtain any advantage from imagining a procedure, learners must be sufficiently expert to be able to imagine the procedure. If they cannot imagine the procedure, studying is superior to imagining ([Cooper et al., 2001](#)).

The collective working memory effect

Collaborative learning can be effective because it allows differential knowledge between learners to be combined, using in effect, a single, collective working memory consisting of the limited working memory of all collaborators ([Kirschner et al., 2011, 2018](#)). The effect is not obtained in the absence of differential knowledge between participants. Collaboration tends to be ineffective if students have similar knowledge in long-term memory ([Retnowati et al., 2017, 2018](#)).

The human movement effect

Due to the transient information effect, animations can frequently increase cognitive load and decrease learning. One exception seems to occur when the animation depicts human movement ([Van Gog et al., 2009](#)). We seem to have evolved to readily assimilate depictions of human movement without an excessive working memory load.

The spacing effect

Recently, cognitive load theory has been used to explain the spacing effect that occurs when learning that is spaced with gaps between events enhances learning compared with the same events being massed without gaps ([Chen et al., 2018](#)). It was suggested that when working memory is overloaded, its resources are depleted resulting in less learning, with recovery after rest. Spaced learning allows that rest and recovery while massed learning does not, providing an explanation of the spacing effect.

Conclusion

The cognitive architecture used as a base for cognitive load theory has generated a range of novel instructional procedures. The assumptions that instructional design should be primarily concerned with domain-specific, biologically secondary information and that when dealing with such information, the information processing procedures follow those used by evolutionary biology, has resulted in instructional procedures that differ from many alternative treatments. The effectiveness of those instructional procedures has been supported by multiple, randomized, controlled studies. Through cognitive load theory, the known characteristics of human cognitive architecture can be used to generate novel instructional designs.

While cognitive load theory has generated many cognitive load effects, the manner in which it has done so should be noted. All cognitive load theory effects have limits imposed by the cognitive architecture that gave rise to them. No cognitive load effect should

be used without considering that architecture. Many new cognitive load effects were discovered as a consequence of the limits of older effects with cognitive architecture used to determine the nature of those limits. For example, the worked example effect was first demonstrated using algebra problems such as $(a + b)/c = d$, solve for a . A natural assumption was that the effect should be obtained using other problems such as geometry, in which a diagram is characteristically provided followed by separate statements concerning that diagram, or kinematics problems in which a statement is provided followed by algebraic procedures related to the statement. It took several years of replication failures to realize that the conventional manner in which worked examples in these areas were presented did not reduce extraneous cognitive load. The conventional way of demonstrating algebra worked examples did not incorporate split-attention. The conventional way of demonstrating geometry or kinematics problems did incorporate split-attention. If a worked example incorporated split-attention, it would not reduce working memory load compared to solving a problem. Once split-attention was eliminated from worked examples, the worked example effect could be readily demonstrated in new areas (Tarmizi and Sweller, 1988; Ward and Sweller, 1990).

This pattern of replication failure followed by the discovery of a new effect based on our knowledge of human cognition has been repeated on numerous occasions. Arguably, cognitive load theory has been developed more by replication failures than by any other factor.

References

- Chandler, P., Sweller, J., 1991. Cognitive load theory and the format of instruction. *Cogn. Instruct.* 8, 293–332.
- Chen, O., Kalyuga, S., Sweller, J., 2015. The worked example effect, the generation effect, and element interactivity. *J. Educ. Psychol.* 107, 689–704.
- Chen, O., Kalyuga, S., Sweller, J., 2016a. Relations between the worked example and generation effects on immediate and delayed tests. *Learn. Instruct.* 45, 20–30.
- Chen, O., Kalyuga, S., Sweller, J., 2016b. When instructional guidance is needed. *Educ. Dev. Psychol.* 33, 149–162.
- Chen, O., Kalyuga, S., Sweller, J., 2017. The expertise reversal effect is a variant of the more general element interactivity effect. *Educ. Psychol. Rev.* 29, 393–405. <https://doi.org/10.1007/s10648-016-9359-1>.
- Chen, O., Castro-Alonso, J.C., Paas, F., Sweller, J., 2018. Extending cognitive load theory to incorporate working memory resource depletion: evidence from the spacing effect. *Educ. Psychol. Rev.* 30, 483–501. <https://doi.org/10.1007/s10648-017-9426-2>.
- Chiesi, H., Spilich, G., Voss, J., 1979. Acquisition of domain-related information in relation to high and low domain knowledge. *J. Verb. Learn. Verb. Behav.* 18, 257–273.
- Cooper, G., Sweller, J., 1987. Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *J. Educ. Psychol.* 79, 347–362.
- Cooper, G., Tindall-Ford, S., Chandler, P., Sweller, J., 2001. Learning by imagining. *J. Exp. Psychol. Appl.* 7, 68–82.
- Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114.
- De Groot, A., Gobet, F., 1996. Perception and Memory in Chess: Heuristics of the Professional Eye. Van Gorcum, Assen, The Netherlands.
- De Groot, A., 1965. Thought and Choice in Chess. Mouton, The Hague, Netherlands (Original work published 1946).
- Egan, D.E., Schwartz, B.J., 1979. Chunking in recall of symbolic drawings. *Mem. Cogn.* 7, 149–158.
- Ericsson, K.A., Kintsch, W., 1995. Long-term working memory. *Psychol. Rev.* 102, 211–245.
- Geary, D., Berch, D., 2016. Evolution and children's cognitive and academic development. In: Geary, D., Berch, D. (Eds.), *Evolutionary Perspectives on Child Development and Education*. Springer, Switzerland, pp. 217–249.
- Geary, D., 2008. An evolutionarily informed education science. *Educ. Psychol.* 43, 179–195.
- Geary, D., 2012. Evolutionary educational psychology. In: Harris, K., Graham, S., Urdan, T. (Eds.), *APA Educational Psychology Handbook*, vol. 1. American Psychological Association, Washington, D.C., pp. 597–621.
- Jeffries, R., Turner, A., Polson, P., Atwood, M., 1981. Processes involved in designing software. In: Anderson, J.R. (Ed.), *Cognitive Skills and Their Acquisition*. Erlbaum, Hillsdale, NJ, pp. 255–283.
- Kalyuga, S., Chandler, P., Tuovinen, J., Sweller, J., 2001. When problem solving is superior to studying worked examples. *J. Educ. Psychol.* 93, 579–588.
- Kalyuga, S., Chandler, P., Sweller, J., 2004. When redundant on-screen text in multimedia technical instruction can interfere with learning. *Hum. Factors* 46, 567–581.
- Kirschner, F., Paas, F., Kirschner, P., 2011. Task complexity as a driver for collaborative learning efficiency: the collective working-memory effect. *Appl. Cogn. Psychol.* 25, 615–624.
- Kirschner, P., Sweller, J., Kirschner, F., Zambrano, J., 2018. From cognitive load theory to collaborative cognitive load theory. *Int. J. Comput. Support. Collab. Learn.* 13, 213–233.
- Leahy, W., Sweller, J., 2011. Cognitive load theory, modality of presentation and the transient information effect. *Appl. Cogn. Psychol.* 25, 943–951.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97.
- Mousavi, S.Y., Low, R., Sweller, J., 1995. Reducing cognitive load by mixing auditory and visual presentation modes. *J. Educ. Psychol.* 87, 319–334.
- Newell, A., Simon, H.A., 1972. *Human Problem Solving*. Prentice Hall, Englewood Cliffs, NJ.
- Paas, F., van Merriënboer, J., 1994. Variability of worked examples and transfer of geometrical problem-solving skills: a cognitive-load approach. *J. Educ. Psychol.* 86, 122–133.
- Peterson, L., Peterson, M.J., 1959. Short-term retention of individual verbal items. *J. Exp. Psychol.* 58, 193–198.
- Pollock, E., Chandler, P., Sweller, J., 2002. Assimilating complex information. *Learn. Instruct.* 12, 61–86.
- Renkl, A., Atkinson, R., 2003. Structuring the transition from example study to problem solving in cognitive skills acquisition: a cognitive load perspective. *Educ. Psychol.* 38, 15–22.
- Retnowati, E., Ayres, P., Sweller, J., 2017. Can collaborative learning improve the effectiveness of worked examples in learning mathematics? *J. Educ. Psychol.* 109, 666–679.
- Retnowati, E., Ayres, P., Sweller, J., 2018. Collaborative learning effects when students have complete or incomplete knowledge. *Appl. Cogn. Psychol.* 32, 681–692.
- Roodenrys, K., Agostinho, S., Roodenrys, S., Chandler, P., 2012. Managing one's own cognitive load when evidence of split attention is present. *Appl. Cogn. Psychol.* 26, 878–886.
- Schneider, W., Shiffrin, R.M., 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66.
- Shiffrin, R.M., Schneider, W., 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84, 127–190.
- Sweller, J., Cooper, G., 1985. The use of worked examples as a substitute for problem solving in learning algebra. *Cogn. Instruct.* 2, 59–89.
- Sweller, J., Sweller, S., 2006. Natural information processing systems. *Evol. Psychol.* 4, 434–458.
- Sweller, J., Mawer, R.F., Ward, M.R., 1983. Development of expertise in mathematical problem solving. *J. Exp. Psychol. Gen.* 112, 639–661.
- Sweller, J., Ayres, P., Kalyuga, S., 2011. *Cognitive Load Theory*. Springer, New York.
- Sweller, J., van Merriënboer, J., Paas, F., 2019. Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292.
- Sweller, J., 2010. Element interactivity and intrinsic, extraneous and germane cognitive load. *Educ. Psychol. Rev.* 22, 123–138.
- Tarmizi, R.A., Sweller, J., 1988. Guidance during mathematical problem solving. *J. Educ. Psychol.* 80, 424–436.

- Tricot, A., Sweller, J., 2014. Domain-specific knowledge and why teaching generic skills does not work. *Educ. Psychol. Rev.* 26, 265–283. <https://doi.org/10.1007/s10648-013-9243-1>.
- Van Gog, T., Paas, F., Marcus, N., Ayres, P., Sweller, J., 2009. The mirror neuron system and observational learning: implications for the effectiveness of dynamic visualizations. *Educ. Psychol. Rev.* 21, 21–30.
- van Merriënboer, J., Krammer, H.P., 1987. Instructional strategies and tactics for the design of introductory computer programming courses in high school. *Instr. Sci.* 16, 251–285.
- Ward, M., Sweller, J., 1990. Structuring effective worked examples. *Cogn. Instruct.* 7, 1–39.
- Wong, A., Leahy, W., Marcus, N., Sweller, J., 2012. Cognitive load theory, the transient information effect and e-learning. *Learn. Instruct.* 22, 449–457. <https://doi.org/10.1016/j.learninstruct.2012.05.004>.

Creativity and learning

Mark A. Runco, Southern Oregon University, Ashland, OR, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	135
Overview of creativity research	135
The relationship between creativity and learning	136
Operant theory and creativity	136
Piaget's constructionist theory and creativity	137
Discussion	139
References	139

Introduction

The benefits of creativity are numerous and varied. They include adaptability, flexibility, mindfulness, problem solving, health, and even quality of life (Runco, 2022). Creativity also contributes in several ways to learning and is an important topic for education. The relationship with learning is particularly important because it suggests that there are ways to fulfill creative potentials. The causal relationship between learning and creativity is reciprocal. Not only does learning sometimes support creativity, but creativity can also ensure that learning is meaningful.

This chapter explores the connections between creativity and learning. It begins with an overview of creativity research. It then reviews the learning theories that have been applied to the topic of creativity, and reviews the theoretical basis explaining how the creative process ensures that learning is authentic and meaningful. A particular focus is the 4P framework for studying creativity (i.e., creative person, creative product, creative process, and creative press or place). This chapter also reviews the operant view of creativity wherein insight may be the result of the spontaneous integration of previously learned responses. It explains how the intrinsic motivation so often involved in the creative process may be explained by reinforcement histories. It then goes into detail about how understanding experience depends on the creative process and how individuals create meaning as they interact with the world. Various implications for education are noted.

Overview of creativity research

Research on creativity has grown in a dramatic way in the last few decades. It has become an interdisciplinary field. There are neuroscientific, genetic, economic, developmental, clinical, historical, cognitive, educational, and cross-cultural studies of creativity (Runco, 2022). Perhaps the best way to examine the various perspectives on creativity involves the 4P framework for studying creativity, originally proposed by Rhodes (1961) but updated several times. Rhodes identified the four key strands of creativity research: investigations of the creative *person*, the creative *product*, the creative *process*, and the creative *press* (which is now usually described as the creative *place*). He used the term *press* because, at the point in time when he proposed the framework, it was common to describe situational pressures on human behavior in that manner. Importantly, some of these press factors have a universal impact, while others known as “beta press influences,” have only idiosyncratic impact, depending on how an individual interprets them. This is a very important point because interpretation is involved in the process whereby creativity is related to learning. This is examined in more detail below.

Studies of the creative person usually focuses on personality traits. There are several specific traits which seem to make up the core characteristics of creativity. These include openness, autonomy, flexibility, and intrinsic motivation. Personality tests are often used in this kind of research. The limitation of the personality approach is that traits are now viewed as interacting with situational factors, and thus not explanatory in and of themselves. Personality might provide creative potential but does not guarantee actual creative performance. This may explain why some students have several of the core characteristics but are not in fact manifestly creative. The concept of creative potential is also examined further below.

Studies of the creative product often focus on outcomes, and may therefore quantify paintings, musical scores, novels published, citations given to an individual scientist, or some other outcome of an individual's work. The focus is on the tangible, or at least objective, outcome. This approach has the advantage of being objective because products can be counted, and this provides objective data. The disadvantage is that products do not offer any information about the process that led to the product. Thus there is objectivity, but not much information provided about the mechanism or what students actually do when they are creative. In addition, product studies have a kind of sampling bias because they usually focus on mature individuals. The products studied tend to be socially recognized, including those outcomes listed above (citations, novels, etc.) and require time, expertise, and maturity to bring them to fruition. Children's products are occasionally studied (Rosenblatt and Winner, 1988; Runco, 1989) but they cannot be evaluated with the same professional standards as those from adults.

Studies of the creative process examine stages of the creative work. They may use something like Wallas' (1926) famous four-stage theory. It described how creative work starts with some sort of preparation (which may include problem discovery or problem identification), then moves to incubation, the result being an illumination (or a-ha!), which requires some sort of verification. Many contemporary views of the creative process use or extend Wallas' stage theory (see Runco, 1994). Theories of innovation, for example, often refer to the same four stages, but then they add a fifth stage, namely *implementation* (see Amabile, 2020 for recent discussion of this stage model).

There is also the divergent thinking approach to creativity, which is essentially a process theory as it describes how individual's thinking progresses and how people may associate one idea to the next. The process is called divergent because it moves in different directions. This may be contrasted with convergent thinking, where an individual takes several pieces of information and the thinking converges on a correct or conventional response or answer. Divergent thinking goes in different directions, some of which may be original. Divergent thinking is aligned with creativity because it is easier to find original ideas when the process is divergent rather than convergent.

Studies of the creative place examine conditions which may either support or inhibit creativity. Creativity is often inhibited, for example, when there are evaluations, so an individual is unable to follow his or her intrinsic interests, or when there are expectations of conventional behavior. Conventional behavior is by definition normative and thus nearly the opposite of the originality that is required for creativity. There are conditions that support creative behavior, including those which allow autonomy and independent thinking. There are a number of measures for assessing particular situations (e.g., the work setting or the classroom) to determine if creativity is supported or inhibited (Amabile, 1990; Rickards and Jones, 1991; Runco et al., 2016).

Runco (2007) reorganized the "4P" framework. This was done, in part, to acknowledge the idea presented above that some influences on creativity merely provide the potential for creativity but do not guarantee actual creative performance. Indeed, the main idea in this modified framework is that there is a distinction between potential and actual performance. Those two things are used as the most general categories of a hierarchy, with personality, press, and process all subsumed under *creative potential*, and products, along with "persuasion" (Simonton's (1988) idea that creative people and products change the way that others think) and creative achievement under the *creative performance* category. The significance of the hierarchical framework, which implies 6Ps instead of 4Ps, is in its suggesting that creative potential might be supported such that it eventually leads to actual creative performance. This is one way to think about education—as a method for fulfilling creative potentials so students will eventually actually earn socially recognized creative achievements. Creative behaviors that are beneficial but personal rather than socially recognized are also important.

Before turning to the relationship between creativity and learning, two other points about creativity research must be made. First, originality, mentioned several times above, is a prerequisite for creativity. It is necessary but not sufficient. There is more to creativity than originality. The "standard definition" of creativity requires originality (or novelty) and effectiveness (or utility or fit) (Runco and Jaeger, 2012). Without effectiveness of some sort, original things might be useless or just crazy. Simonton (2012) suggested that there should be a third requirement for creativity, namely surprise (as in non-obvious), and Kharkhurin (2014) pointed to aesthetic value and authenticity. The latter makes sense because many Eastern cultures actually value authenticity more than originality (e.g., Tan, 2016) and authenticity is a part of the self-actualization that Rogers (1959) and Maslow (1968) found to characterize highly creative individuals. "Value" has also been nominated as a requirement of creativity, but that suggestion has been met with criticisms of subjectivity (Harrington, 2018; Weisberg, 2018).

The second important point is that, when there is a need for precision, it is best to avoid the noun *creativity*. This is because the noun is used in so many ways, and not all of the meanings are interchangeable, as is obvious from the various perspectives captured by the 4P and hierarchical frameworks. What is said about the creative product may not apply well to the creative process, just to name one example. Note that this example suggests a useful alternative to using the noun: the adjective *creative*, as in creative process, creative achievement, creative activity, creative thinking, and so on. Using the adjective *creative* requires more precision and avoids confusion.

The relationship between creativity and learning

Theories of creativity have clear implications for learning, which of course is an important part of education. As a matter of fact, a case can be made that authentic learning depends on a creative process. This view is explained below and put into the context of the operant theory of learning. There are other theories of learning, but some only explain reflexive and involuntary behaviors, which means that they do not relate to creative behavior, while there are a number of connections between operant theory and the creative process.

Operant theory and creativity

At first blush, it may seem that traditional learning theories do not say much about creativity. This is because some of them (e.g., classical conditioning) only describe changes in reflexes, and those which focus on voluntary actions only describe changes in overt behavior. Skinner's (1974) theory of learning has sometimes been called *Behaviorism*, so named because its emphasis on maximal objectivity meant studying only behavior that was overt and observable, and therefore amenable to objective assessment. Many people prefer the label "operant theory" to Behaviorism. An operant is a voluntary behavior that is emitted in order to obtain

a reinforcer or to avoid a punisher. Whichever label is used, operant theory is distinct from other kinds of learning theories, most notably the Pavlovian school, also known as Classical conditioning, and Social Learning Theory.

Even with the emphasis on overt, maximally objective behavior, operant research has contributed to our understanding of the creative process. Some of this research uses Skinner's idea that behavior is displayed because of reinforcement history, not because there are immediate contingencies involved. Skinner thought that reinforcement histories can explain things like the creative efforts of the so-called *starving artist*. That refers to a stereotype of a creative individual who is starving in the sense that he or she is doing original things that are not being rewarded. Without rewards, in the form of money or similar support, the individual has little and is starving. Why would an artist invest in work that is not being supported? A personality psychologist may say that the artist is following intrinsic rather than extrinsic rewards. The operant view is that it is not intrinsic motivation but is instead a history of reinforcement that continues to support the artistic efforts or other expressions of originality (Runco, 2019). In this light a child might have been heavily rewarded for being original and continues to act in an original fashion, or even choose a line of work that allows originality, because of that history. This makes the most sense if, as a child, the individual received intermittent reinforcement, for that leads to behavior that is resistant to extinction.

This is an intriguing explanation for certain original behaviors, but a question arises because empirical research has demonstrated that extrinsic rewards tend to undermine creative work (Amabile, 1990; Eisenberger and Shannock, 2003). In fact, extrinsic rewards may eliminate the intrinsic motivation for creativity efforts. If individuals are intrinsically motivated, but then receive extrinsic rewards for that behavior, they may forget that they were initially intrinsically motivated (Hemmessey and Zbikowski, 1993). This is known as *overjustification*.

The operant idea of reinforcement history explaining the original behavior of starving artists is theoretical, but there is also empirical work using operant techniques to control thinking that is relevant to many creative achievements, namely *insight*. Insight is not the same thing as creativity, but many breakthroughs seem to have been the result of insight (Gruber, 1988). Insights involve a sudden breakthrough and a dramatic shift in one's thinking. Insights are sometimes called "a ha!" moments and are synonymous with the illumination stage mentioned earlier. One explanation for them is that there are changes in the individual's cognitive structures. Operant investigations have examined insight because it is operational and observable. In one example, Epstein (1990) started with a methodology that had been used in famous research conducted in the 1940s by Wolfgang Kohler. Kohler worked with an ape named Sultan. This research had a large impact at the time because it demonstrated that Sultan could be insightful. That was significant because, up to that point, scientists believed that apes were only capable of trial and error and were not capable of insight. Kohler disproved this with his demonstration that the ape Sultan was capable of insight and did not rely on trial-and-error thinking. Kohler demonstrated insight in a series of observational studies. In one, a banana was placed outside of Sultan's cage, beyond his reach. Sultan had various things in the cage with him, including a stick. The stick was not the only thing in the cage, but Sultan did not experiment with all available objects, which a theory of trial and error would predict. Instead, he stood at the bars of the cage, observed the banana, which he very much wanted, and looked back and forth from the banana to the stick. Without trial and error, Sultan eventually walked over to the stick, took it back to where he was standing at the bars, and used the stick to extend his reach and pull the banana toward him. He solved the problem in an insightful way and was rewarded with a banana.

In another experiment, a banana was hung from the ceiling of Sultan's cage, again out of reach. This time the cage contained various crates and boxes. Sultan did not use trial and error and experiment with all possible objects but instead looked at the boxes, then looked at the banana, looked back at the boxes, and so on. He eventually picked up one box moved it underneath the banana, went to another box and stacked it on top of the first box, climbed on top of both boxes, and was able to reach the banana.

Epstein (1990) conducted similar work with pigeons. This work is remarkable because it is one thing to accept that apes have thinking skills that are similar to that of humans, but pigeons? From the operant perspective what is remarkable is that the insightful thinking could be manipulated by controlling environmental conditions, including reinforcers. Epstein started by making a small banana out of pigeon chow. This banana was hung from ceiling of a cage. A pigeon was conditioned such that they learned that the banana shaped object was actually pigeon chow. Their wings were clipped, in a humane fashion, so they could not fly when they were placed in a cage with the "banana" hanging from the ceiling. Before they were placed in the cage they were also conditioned to push small boxes around, and, in separate experimental conditions, trained to stack the boxes. Importantly, the conditioning of these behaviors—pecking a banana shaped object, pushing boxes, and stacking boxes—was such that there was only one target behavior at a time. The learned behaviors were, then, discrete and independent. Yet when the pigeons were placed in the cage with the hanging banana, they seemed to do what Sultan had done. They looked at the banana, looked at the boxes, back and forth, and fairly quickly pushed the boxes, stacked them, hopped up, and pecked the banana. No trial and error. This led Epstein to conclude that seemingly insightful behaviors can in fact be explained in terms of reinforcement histories and that insightful solutions might reflect the "spontaneous integration of previously learned responses." According to Epstein, there was no reason to attribute any great intellectual ability to the pigeons. Instead it is reasonable to explain their behavior, which was very similar to Sultan's, as a result of their learning. Note that what was learned was insight, not creativity per se.

Piaget's constructionist theory and creativity

Like all work with subhuman species, the operant work just reviewed may not generalize to humans. Jean Piaget's work (1976), in contrast, involved children. Piaget's view of development is constructionistic. Useful learning, for Piaget, requires that the individual construct knowledge for him- or herself. This construction may be viewed as a creative process. New knowledge may be created as the individual interacts with the environment. The title of Piaget's (1976) monograph, *To Understand is to Invent*, captures this nicely.

Piaget (1976) argued that education is often superficial. In particular, he criticized learning that depends on memorized information. Useful learning, in contrast, involves the construction of knowledge. Piaget's ideas about learning and what he called adaptation were extensive, in part because he developed them over several decades. There are misunderstandings about his theory, and some of these reflect the fact that what he wrote early on was not identical to what he wrote later, after he had more data and his theory evolved. It is also important to note that his data were not from controlled research. He relied on a method called *empathic inference*. This involved carefully observing children under certain conditions and inferring what they must have been experiencing cognitively and how that led to the actions observed.

After observing quite a few children Piaget inferred that children develop through stages. The first is sensorimotor, the next is preoperational, followed by the concrete operational stage, and finally, formal operations. Sensorimotor thought is all about the body and movement. It may not even seem like thought, and it certainly is not like the thinking of older individuals, but the children are processing information and very early on form cognitive structures (schema, schemes, or schemata) which represent what they understand about their bodies. The preoperational stage occurs when a child develops the capacity to use symbols. This allows them to think using language, but also to think using various other symbol systems. Concrete operational thinking allows the child to understand classifications and to conserve quantities, while formal operations allows the individual to think in a way that is not unlike formal logic. The formal operational individual may also think deductively and about hypothetical things. Most important for the present purposes is that individuals develop structures—they structure their thinking—as they grow older. According to Piaget, the kind of structures and cognitive operations they can use are constrained by biological age, as well as experiences. This is of course where learning comes in, and where the connection of learning with creativity is apparent. Simply put, new structures are created. In other words, the individual uses a creative process to construct new structures, and those structures are indicative of learning.

Piaget had his own view of how individuals benefit from experience and preferred the label “adaptation.” A close examination of adaptation provides further detail about what is involved in creative thinking. The first important point, implied above, is that adaptation is very different from memorization. Adaptation leads to new structures, new meaning, and new understandings of the world. The second of those things, understanding, is special because it points to the reason why humans do in fact expend effort to adapt. Why do we allocate time and energy? So we can make sense of our experiences. This begins with the assimilation of information, and then an accommodation of that information. The first of these, assimilation, requires that an individual have new experiences, new in the sense that they are not understood. This does create a problem because how can the individual think about something that he or she does not understand? The answer is that the new information is altered, and some of it may be ignored. The individual alters information such that it can be brought into his or her cognitive system. Once it is in the system, it can be processed and the existing structures can accommodate to that new information. New structures may be formed or existing structures can be modified to differentiate so the new information is understood. The construction of new understandings is a creative process.

Importantly, individuals invest resources into adaptation because they are intrinsically motivated to do so. Piaget did not use the phrase “intrinsic motivation,” but he did propose that students did not need to be motivated to learn. They are genetically predisposed to maintain an understanding of their experiences, and this leads them to adapt to new experiences. No one likes to be in a situation that they cannot understand. For this reason everyone is motivated to develop and maintain an understanding of experience. That motivation is not a result of extrinsic factors such as bonuses, grades, salaries, points, or rewards of any sort. This is important and has obvious implications for education. As Piaget put it, the difficulty with education is to find things that challenge individual students. If they are challenged, and the challenge is slightly above their current level of functioning, they will invest the resources in order to adapt and create new understandings. Since those are new, they are original, and if they do in fact make sense, they are effective. Thus, they satisfy the standard definition of creativity, with its two requirements.

Elkind (1981) extended the Piagetian view of education and pointed to communication as the key educational issue. He described how children are in stages that are different from the educator (e.g., a concrete operational student with a formal operational teacher) and thus the student's thinking is a very different from that of the teachers. The students are “cognitive aliens.” This often makes it very difficult for the teacher to understand the student, which in turn means that communication is difficult. Runco and Beghetto (2018) examined similar issues but suggested that the creative process could be used by teachers to better communicate and understand. Their theory distinguished between *primary creativity*, defined as that which results from a student's constructing new knowledge via an original and effective interpretation of experience, and *secondary creativity*, which is the teacher's creative interpretation of the student's ideas.

Interestingly, teachers' thinking may differ from that of students but also from that of parents. Teachers' thinking (or “implicit theories”) of children's creativity have been systematically examined, with the results indicating that teachers give the most weight to creativity as expressed in social actions (e.g., creative behavior which is unusual and expressed by very few students). Parents, in contrast, tend to be more concerned about the cognitive bases of creativity (e.g., problem solving). These differences are important because it can be quite beneficial when teachers and parents are coordinated in their efforts to support children's development. Additionally, generalization is most likely when the efforts of teachers match the efforts of parents, so what is learned in school is also supported when students are at home.

Teachers generally value creativity and believe that it should be supported in the classroom, but at the same time, they very often do not realize that creativity depends on autonomy, divergent thinking, and intrinsic motivation (Westby and Dawson, 1995). This can make it very difficult for teachers to encourage creativity. It is understandable in that autonomy and these other qualities can be difficult to appreciate when teachers are working with groups of students. Each of these things would be much easier if the teacher could work with one child at a time. In a group, it is much easier for all students to follow a set curriculum and work at the same pace.

Discussion

This chapter has touched on many of the important concepts used in research on creativity. It covered the 4P framework and its extension into a hierarchy, summarized the standard definition of creativity and possible extensions of it, and used findings concerning intrinsic motivation, insight, and problem discovery as part of the review of possible relationships between learning and creativity.

Implications for education are various. The 4P and 6P theories suggest that classroom settings should allow for autonomy and avoid unnecessary evaluation. They also suggest that openness, autonomy, and flexibility are especially important characteristics, and as such should be supported, and that divergent thinking and insightful thinking are both related to the capacity for creative problem solving. The “press” research on creativity suggests that there are individual differences among students which mean that not all students will react the same way to classroom conditions. Interactions between a student’s personality and particular conditions should be expected. The ideas about intrinsic motivation, from both Amabile and Piaget, suggest that educators should find ways to allow students to follow their interests, when possible, and that intrinsic motivation is more likely to lead to creative thinking, compared with grades and other extrinsic rewards.

There are several ways of looking at the relationship between creativity and learning. There is learning, as in operant conditioning involving reinforcers and punishers, but there is also adaptation in the Piagetian sense. The latter is especially important because it explains how new understandings are actively constructed by individuals. The construction of meaning is a creative process. New understandings and meaning are created. These may eventually lead to creative products and socially recognized creativity, but the new understandings are of course important in and of themselves. They allow us to cope with challenges and deal with everyday experience. Creativity is often a personal thing. Social recognition is sometimes important, but not always.

One thing that should be clarified is that adaptation is not always creative. The two processes are not synonymous. Adaptation implies a reaction by the individual, and creativity is sometimes proactive rather than reactive. Also, recall here that the creative process often depends on problem discovery. This would occur in a proactive matter, before a problem has been identified. In addition, on a general level, adaptation could lead a student to conform rather than to originality. This kind of adaptation is behavioral and thus unlike the cognitive process Piaget (1976) described, the idea being that if an adaptive student finds him- or herself in a situation where there is pressure to conform, the adaptive response may be to go with the crowd. It is possible certainly that the individual has the ego strength to resist pressures to conform and does not conform, even if adaptable. It would probably be useful for the teacher to display creativity him- or herself and thus provide a model of creativity. Other descriptions of how adaptability may be related to creativity have been provided by Cohen (1989) and Russ (1998).

Teachers can themselves model creative behavior. This will convey the idea that creativity is a valuable thing, and students are likely to internalize those values. The modeling might demonstrate how to use various tactics that facilitate original thinking. Runco (2017) reviewed a number of such tactics. These included changing one’s perspective when faced with a problem, collecting more information, putting the issue or problem aside (to allow incubation), questioning assumptions, and changing the problem such that it more easily allows original options. It might be argued that, while tactics are helpful, they are not as important as the other direction of effect mentioned herein, whereby creativity leads to meaningful learning. This is why the idea of creating personal understanding and Piaget’s (1976) “to understand is to invent” was discussed in detail. This process of finding meaning and creative new understandings can help many times each day and can support the most important parts of life, including the recognition of one’s own meaning in life. Such meaning in life is one of the more important things that education may support. It can support it by encouraging creative thinking which in turn allows students to create meaning. This is just one of the numerous and varied personal and social benefits to the creative capacity. Much can and should be done to support the creativity of students.

References

- Amabile, T.M., 1990. Within you, without you: the social psychology of creativity, and beyond. In: Runco, M.A., Albert, R.S. (Eds.), *Theories of Creativity*. Sage, Newbury Park, CA, pp. 61–91.
- Amabile, T.M., 2020. On Wallas. In: Glavaneau, V. (Ed.), *Classic Readings in the Creativity Research*. Oxford University Press, pp. 409–430.
- Cohen, L.M., 1989. A continuum of adaptive creative behaviors. *Creativ. Res. J.* 2, 169–183.
- Eisenberger, R., Shanock, L., 2003. Reward, intrinsic motivation, and creativity: a case study of conceptual and methodological isolation. *Creat. Res. J.* 15, 121–130.
- Elkind, D., 1981. *Children and Adolescence*. Oxford University Press, New York.
- Epstein, R., 1990. Generativity theory. In: Runco, M.A., Albert, R.S. (Eds.), *Theories of Creativity*. Sage, Newbury Park, CA, pp. 116–140.
- Gruber, H.E., 1988. The evolving systems approach to creative work. *Creat. Res. J.* 1, 27–51.
- Harrington, D.M., 2018. On the usefulness of “value” in the definition of creativity: a commentary. *Creativ. Res. J.* 30 (1), 118–121. <https://doi.org/10.1080/10400419.2018.1411432>.
- Hennessey, B.A., Zbikowski, S.M., 1993. Immunizing children against the negative effects of reward: a further examination of intrinsic motivation training techniques. *Creat. Res. J.* 6, 297–307.
- Kharkhurin, A.V., 2014. Creativity.4in1: four-criterion construct of creativity. *Creat. Res. J.* 26, 338–352.
- Maslow, A., 1968. Creativity in self-actualizing people. In: *Toward a Psychology of Being*. Van Nostrand Reinhold, New York, pp. 135–145.
- Piaget, J., 1976. *To Understand is to Invent*. Norton, New York.
- Rhodes, M., 1961. An analysis of creativity. *Phi Delta Kappan* 42, 305–310.
- Rickards, T., Jones, L.J., 1991. Toward the identification of situational barriers to creative behaviors: the development of a self-report inventory. *Creativ. Res. J.* 4, 303–316.
- Rogers, C.R., 1959. Toward a theory of creativity. In: Anderson, H.H. (Ed.), *Creativity and its Cultivation*. Harper & Row, New York, pp. 69–82.
- Rosenblatt, E., Winner, E., 1988. The art of children’s drawing. *J. Aesthetic. Educ.* 22, 3–15.

- Runco, M.A., 1989. The creativity of children's art. *Child Stud. J.* 19, 177–189.
- Runco, M.A. (Ed.), 1994. *Problem Finding, Problem Solving, and Creativity*. Ablex, Norwood, NJ.
- Runco, M.A., 2007. A hierarchical framework for the study of creativity. *New Horizons in Education* 55 (3), 1–9.
- Runco, M.A., 2017. Tactics and strategies for creativity. In: *Reference Module in Neuroscience and Biobehavioral Psychology*. Elsevier, pp. 1–4.
- Runco, M.A., 2019. B. F. Skinner on poetry, fiction, and design: an implicit science of creativity. In: Glavaneau, V. (Ed.), *Classic Readings in the Creativity Research*. Oxford University Press, pp. 409–430.
- Runco, M.A., 2022. *Creativity: Theories and Themes: Research, Development, and Practice*, third ed. Academic Press, San Diego, CA.
- Runco, M.A., Jaeger, G., 2012. The standard definition of creativity. *Creat. Res. J.* 24, 92–96.
- Runco, M.A., Beghetto, R., 2018. Primary and secondary creativity. *Curr. Opin. Behav. Sci.* 27, 7–10.
- Runco, M.A., Acar, S., Campbell, W.K., Jaeger, G., McCain, J., Gentile, B., 2016. Comparisons of the Creative Class and Regional Creativity With Perceptions of Community Support and Community Barriers. *Business Creativity and the Creative Economy* 2, 83–92.
- Simonton, D.K., 1988. Creativity, leadership, and chance. In: Sternberg, R.J. (Ed.), *The Nature of Creativity: Contemporary Psychological Perspectives*. 386–426. Cambridge University Press, New York.
- Simonton, D.K., 2012. Taking the U.S. patent office criteria seriously: a quantitative three-criterion creativity definition and its implications. *Creat. Res. J.* 24, 96–107.
- Russ, S.W., 1998. Play, creativity, and adaptive functioning: implications for play interventions. *J. Clin. Child Psychol.* 27, 469–480.
- Skinner, B.F., 1974. *About Behaviorism*. Alfred Knopf., New York.
- Tan, C., 2016. Creativity and confucius. *Journal of Genius and Eminence (JGE)* 1 (1), 79–84.
- Wallas, G., 1926. *The Art of Thought*. Harcourt Brace & World, New York.
- Weisberg, R.W., 2018. Response to Harrington on the definition of creativity. *Creativ. Res. J.* 30 (4), 461–465. <https://doi.org/10.1080/10400419.2018.1537386>.
- Westby, E.L., Dawson, V.L., 1995. Creativity: asset or burden in the classroom? *Creat. Res. J.* 8 (1), 1–10.

Curiosity and interest: what should educators know and consider?

K. Ann Renninger^a, Fanxiao Wani Qiu^a, and Suzanne E. Hidi^b, ^a Department of Educational Studies, Swarthmore College, Swarthmore, PA, United States; and ^b Department of Applied Psychology and Human Development, OISE, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Definitions, characteristics, and case examples	141
Curiosity	142
Triggered situational interest that is sustained	142
Sustained and developing individual interest	143
Educational implications	143
Curiosity development	144
Interest development	144
Concluding thoughts and future study	146
Acknowledgments	146
References	146

Recent research and its discussion confirm that although curiosity and interest have similar characteristics, they are distinct constructs, and may involve different types of engagement (e.g., [Hidi and Renninger, 2019, 2020](#); [Murayama et al., 2019](#); [Shin and Kim, 2019](#); see also discussions in [Peterson and Hidi, 2019](#)). Nonetheless, the specific educational implications of these findings for practice have received relatively little attention. Given that basic research does not always map directly to practice, in this chapter we address what is understood about such applications. We also consider what still needs further study. We begin by defining and describing the characteristics of both curiosity and interest, and by highlighting the similarities and differences between them. To illustrate and consider these constructs, we draw on case materials. Following this, we discuss the educational implications of each, as well as the needed next steps in their research.

Definitions, characteristics, and case examples

Curiosity and interest describe psychological states, and share characteristics such as being universally motivating and not requiring extrinsic rewards. They are each experienced as energizing, and as directing information search—an activity that is rewarding because of its association with the reward circuitry in the brain (e.g., [Gottlieb et al., 2013](#)). However, curiosity and interest represent different types of information search. Information search related to curiosity is associated with a desire to resolve uncertainty, or to close a knowledge gap (e.g., [Berlyne, 1960](#); [Grossnickle, 2016](#); [Jirout et al., 2018](#); [Loewenstein, 1994](#); [van Schijndel et al., 2018](#)), whereas information search linked to interest is general and ongoing ([Ainley and Hidi, 2014](#); [Renninger and Hidi, 2019, 2020](#)). We note that curiosity, or openness, is also studied in personality theory as a trait of individuals ([Digman, 1990](#)); however, we do not address the study of curiosity as a trait in this chapter. Interest can be initiated by the triggering of situational or momentary interest that may continue to develop into a relatively stable individual interest, or by an aspect of the experience of engaging with an already existing interest ([Hidi and Renninger, 2006](#); see [Renninger and Hidi, 2016, 2022](#), for an update and overview). Thus, in addition to describing a psychological state during engagement, interest also refers to individuals' motivation to reengage with the content of interest over time.

Triggers commonly associated with curiosity have been uncertainty, novelty, surprise, or complexity; they characteristically create an information gap (e.g., [Berlyne, 1960](#); [Jirout et al., 2018](#); [Liquin and Lombrozo, 2020](#); [Smock and Holt, 1962](#)). In surveys, indicators of curiosity have also included confusion, and/or frustration as associated with the tension introduced by the knowledge gap before it is resolved (e.g., [Lamnina and Chase, 2019](#); [Litman and Jimerson, 2004](#)). Once the knowledge gap is resolved, the experience of curiosity becomes associated with positive affect (e.g., [Manz and Suárez, 2018](#); [Tang et al., 2020](#)).

Although surprise and complexity may also be triggers for interest, there are many other triggers, such as hands-on activity, group work, technology, self-related information processing, etc. ([Hidi and Renninger, 2019](#); [Renninger and Hidi, 2021](#); see [Renninger et al., 2019](#) for a detailed discussion of triggers in a classroom). Indicators of interest that together have been used to assess the development of interest in surveys, observational protocols, interviews, etc., include the frequency and depth of individuals' engagement with particular content, as well as whether their engagement is voluntary and independent when opportunity allows ([Renninger and Hidi, 2016](#); [Renninger and Wozniak, 1985](#)). It should be noted that this approach differs from assessment of interest based only on affect (e.g., How much do you like doing math?, [Eccles et al., 1993](#), p. 834).

As indicated in [Table 1](#), curiosity and interest differ not only with respect to their affective valence, but also in their duration, goals, and required level of knowledge (see extended discussions in [Grossnickle, 2016](#); [Hidi and Renninger, 2020](#); [Shin and](#)

Table 1 Characteristics of curiosity and interest.

	<i>Curiosity</i>	<i>Interest</i>
Most common triggers	Few: Uncertainty, novelty, surprise, and complexity	Many: Surprise, complexity, hands-on activity, group work, technology, self-related interventions such as personalization, utility-value, etc.
Duration	Short-lived	On-going
Goals	Information search to close a knowledge gap	General information search
Knowledge	Moderate level	All levels
Affect	Discomfort and tension prior to resolution of knowledge gap, followed by positive affect	Primarily positive

Kim, 2019). In the following, case materials from second grade students are provided in order to allow more nuanced consideration of the characteristics of curiosity and interest development, and their implications for practice.

Curiosity

Keisha's curiosity was piqued during a science lesson on simple machines, and later resolved. She rotated through different stations set up throughout the classroom that included opportunities to work with inclined planes, pulleys, and levers. At the pulley station, she stopped and was confused about why a bucket of water was hard to lift on its own, but felt easier to lift when attached to the pulley. Keisha wrote down her question on her worksheet as she had been told to, but must have been thinking about it even while she was supposed to be exploring inclined planes, as she was not really attentive to them.

During the whole class discussion, Keisha asked, "Why does the bucket feel lighter using the pulley?" Her teacher explained that pulleys redirect force and showed an instructional video on how pulleys work. This response resolved the question because she was given the information that she needed. She asked no further questions about pulleys, and did not pursue additional information or opportunities to work with pulleys after that class ended.

The trigger for Keisha's curiosity about pulleys was uncertainty (e.g., Engel, 2011; Jirout et al., 2018). However, for Keisha to seriously engage with the information and the video that her teacher provided, it appears that she may have needed more knowledge than she had (Grossnickle, 2016). Guidance in her work with the hands-on activities would have been very helpful to her (Furtak et al., 2012). If someone had picked up on and responded to her uncertainty about working with the bucket of water, Keisha might have continued to seek out information about pulleys, possibly triggering subsequently her interest development in pulleys or mechanics. In order to trigger her interest, her teacher might have offered information that it could feel easier to pull down on an object than to lift the same object up, and then could have invited her to try using some different objects on the pulley (see discussion in Luce and Hsi, 2014). In subsequent classwork, the teacher also could have included activities that would provide Keisha (and the other students) with additional opportunities to work with pulleys. For example, the teacher might have engaged all of the students in follow-on activities such as figuring out whether including many or fewer wheels changed how easy it was to lift an object, or assigning reading related to mechanics. Importantly, the teacher could make this kind of an addition to the curriculum, and may well trigger and sustain Keisha's interest in the pulleys.

Triggered situational interest that is sustained

On her walk to school one morning, Rosa saw a river otter. She was brimming with questions about the mammal she saw when she got to school. She asked her teacher what the animal was, and why it was on land. She said she thought that it could swim, and that it was really cute and playful. She wanted to know where it lived, and what it might eat. Her teacher was friendly and told Rosa that the animal she saw was likely to be a river otter.

She explained that there are a lot of otters out there, and asked Rosa to hang up her coat and take her seat. Rosa did as she was told, and joined her friends at their table. She began talking with them about the picture books they were looking at. It appeared that even though her interest in the river otter had been triggered, it would not be sustained, and that this was a missed opportunity to support her to develop her interest.

However, a few weeks later, when her teacher told the class that they would be working on the River Project, Rosa seemed to have returned to thinking about the otter. The teacher explained that each student would be researching a different mammal that lives along the river. The design of the project included each student reading several books on the mammal they had chosen, taking notes on its appearance, size, habitat, adaptations, as well as its prey and predators. The project would culminate with a written report and a presentation during an all-school assembly. Almost immediately, Rosa excitedly announced that she would be studying the river otter.

The River Project research provided an opportunity for Rosa to engage in information search that enabled her to answer her earlier questions, tackle those that were listed by her teacher, and to generate some additional questions to address on her own as well. Her project and presentation far exceeded the stated expectations for the project.

Rosa's interest in the river otter appears to have been triggered when she noticed that it was on the riverbank and was not in the water. It may have helped to trigger her interest that she thought the otter was cute and playful. She did not know what the mammal was. At first, it looked like her interest had only been triggered, and that she would not return to related information search. However, the River Project research provided an opportunity for her to re-engage in trying to learn about the otter.

Even though she did not know much at all about river otters, Rosa had enough information to ask questions and to begin work on the animal research project independently; there is no minimal amount of knowledge that is required for sustaining and developing interest once it is triggered (Grossnickle, 2016). It is likely that for her the most meaningful new information about the river otter would include answers to the questions that she had asked her teacher. Her questions addressed information that she wanted to understand and, as such, were self-related (Hidi et al., 2017, 2019; Renninger and Hidi, 2021). In answering those questions, as well as those that were assigned, Rosa had a base of knowledge from which to generate and pursue answers to additional questions for her report. The quality of the report signaled that her developing interest benefitted her learning, as well as her performance (e.g., Jansen et al., 2016).

Sustained and developing individual interest

Dom has been working on a collaborative writing project with his best buddy, and seemed to be willing to talk about writing with anyone who will listen. Early in the school year, he and his buddy had written a fictional story about a group of friends who went trick-or-treating at a haunted house. After working on "their book," Dom eagerly shared it with the rest of the class. He (and his buddy) reveled in the positive feedback that they received and the interest of their peers. Dom has continued to write more stories with his buddy. He has also begun writing stories on his own. Recently, he has been asking others to read them and to give him feedback, so that he can make them "even better."

There appear to be multiple triggers for Dom's interest in writing. (We note that even though triggers are typically studied individually in the laboratory, outside of the laboratory, multiple triggers for interest co-occur and provide support for sustaining triggered interest; Renninger et al., 2019; see discussion in Renninger and Su, 2012). One trigger appears to have been the opportunity to work together with his best buddy on a story (Bergin, 2016; Renninger and Hidi, 2016). Working together may also have been a chance for them to remember and describe their own experiences trick-or-treating and visiting a haunted house, as self-related information is also a trigger for interest (Renninger and Hidi, 2021). The positive feedback and interest from their peers that followed served as additional triggers for interest, and likely contributed to Dom's own sustained, and then developing interest in writing (Hidi et al., 2002; Hidi et al., in press). In working collaboratively, they were able to pick up on and stretch each other's ideas (Webb and Palincsar, 1996), which likely contributed to their developing knowledge about the process of writing. Moreover, as they were drawing on information that they already knew, no additional information was required (Hidi and Anderson, 1992; Grossnickle, 2016).

In summary, as Dom's and Rosa's cases both illustrate, when interest is triggered and sustained, learners readily begin to generate their own curiosity questions—questions that are novel to them (although they may not be novel to others) (Renninger, 2000). These questions may be verbalized or may just inform successive instances of their information search. Importantly, the individuals do not need support to ask and pursue answers to their questions, nor do they need a certain level of knowledge to continue meaningful and deepening interaction with the content. By contrast, once curiosity is resolved, as in Keisha's case, it is not likely to result in further engagement, as the information gap has been closed. If, however, the questions informing the triggering of curiosity were to provide a basis for continuing re-engagement, this subsequently could lead to the triggering of a situational interest, and in turn, ongoing information search and the development of interest (Hidi and Renninger, 2019; Luce and Hsi, 2014). In such situations, curiosity could be considered a type of trigger for interest.

Educational implications

Luce and Hsi (2014) suggested the potential of using students' spontaneously generated questions as guides for educators wishing to support the development of interest, as these provide information about the specific connections that learners are making to content. During discussions with sixth graders about their photojournalism essays, the researchers tracked the youths' curiosity expressions about causal mechanisms, teleological explanations, and inconsistencies in observations. They noted that the youths readily generated spontaneous curiosity expressions, and furthermore that the youths' curiosity expressions ranged in type and were idiosyncratic. Although they called these questions curiosity expressions, we suggest that their data provide yet another illustration of how individuals' information search informs the course of interest development (see related discussions in Alexander et al., 2019; Azevedo, 2013; Hecht et al., 2021). We also draw attention to the self-related nature of spontaneous expressions or questions of this sort, noting that because they are the individuals' "own" questions, and are based on the individuals' own information search, they are likely to be triggers for interest and ongoing learning (e.g., Harackiewicz et al., 2008; Bernacki and Walkington, 2018; see discussion in Renninger and Hidi, 2021). The self-related quality of interested engagement is critical to individuals voluntarily continuing to re-engage with content. Such reengagements provide opportunities for exploration, practice, and developing conceptual understanding.

Curiosity development

As depicted in [Table 2](#), the focus of curiosity research that has addressed educational practices drawing attention to uncertainty have ranged from more to less direct instruction in support of asking questions. Some have specifically prescribed curiosity-promoting instruction in schools, such as modeling comfort with uncertainty and question generation (see [Grossnickle, 2016](#); [Jirout et al., 2018](#); [Loewenstein, 1994](#)), whereas others have investigated the design of successive engagements with complexity (e.g., [Engel, 2011](#); [Manz and Suárez, 2018](#); [van Schijndel et al., 2018](#); [Wu et al., 2018](#)).

Studies of curiosity have generally suggested that promoting curiosity in the classroom requires optimal levels of uncertainty, which can be achieved by helping students to prefer and tolerate greater levels of uncertainty, even if negative feelings are triggered (e.g., [Engel, 2011](#); [Jirout, 2020](#); [Jirout et al., 2018](#)). Supporting curiosity development by teaching students to question has been shown to positively influence levels of curiosity and engagement in school (e.g., [Wu et al., 2018](#)), and also to have some limitations. [Van Schijndel et al. \(2018\)](#), for example, showed that although curious children learned more, they did not have a higher quality of exploration than those with less curiosity. In other words, the children closed their knowledge gaps and learned, but their search for information was not qualitatively better than that of those who had less curiosity.

[Clark et al. \(2019a,b\)](#) used [Rothstein and Santana's \(2011\)](#) Question Formulation Technique (QFT), a question-brainstorming intervention, in two studies designed to promote curiosity. The QFT includes six components: (1) question focus (the teacher or student introduces a prompt that is not a question), (2) question production (students generate as many questions as they can), (3) categorization (marking questions as either closed or open-ended and changing one into another as needed, discussing the advantages and disadvantages of each type), (4) question selection (students choose their top priority questions), (5) next steps (students plan how they want to use their questions to inform their learning), and (6) reflection (students reflect on their learning from the QFT). In the first study, [Clark et al. \(2019a\)](#) reported that teaching the adolescents they studied to ask questions positively influenced their curiosity. However, in a follow-up study, they further explained that the procedure of teaching question-asking had a negative impact on students' cognitive engagement and self-efficacy ([Clark et al., 2019b](#)).

Together with elementary science teachers, [Manz and Suárez \(2018\)](#) developed a method to support learners to ask their own self-related questions and, on the basis of these, to re-engage uncertainty, and search for additional information. The researchers described three strategies for supporting the second-grade students with whom they were working to develop what they termed "scientific uncertainty." These were: first, introduce a complex phenomenon; second, follow this with phases of experimentation with that phenomenon, and finally, leverage the variability that emerges in students' ideas. To this end, following the students' experimentation with the phenomenon, teachers then encouraged the students to reflect on what they had learned by asking them to say what they might change in their experiments the next time. The students were supported to formulate additional research questions and to choose what they wanted to study. This meant that they re-engaged the uncertainty that characterizes the scientific process and deepened their knowledge. Keisha, in the case described earlier, may have needed this type of support in order to continue work with the bucket and the pulley. We suggest that Manz and Suárez' strategies for promoting scientific uncertainty could be understood as supports for triggering and developing interest. Thus, when uncertainty is resolved, given that there is encouragement to sustain engagement, situational interest development can be triggered ([Renninger and Hidi, 2016](#)).

Interest development

Interest develops in relation to individuals' ongoing search for information that is not focused only on closing an information gap. As depicted in [Table 3](#), studies of interest development and classroom practice have generally focused on providing support for learners to engage in meaning-making about the content of their potential interest. Support to develop an interest may be facilitated

Table 2 Examples of studies suggesting educational approaches for promoting curiosity.

Approach	Context	Measures	Findings
Create optimal levels of uncertainty, teach tolerance for uncertainty (Jirout and Klahr, 2012 ; see also Jirout et al., 2018)	Laboratory	Task performance	A medium level of uncertainty leads to exploration.
Incorporate successive engagement in science instructional activities (Manz and Suárez, 2018)	Classroom	Teacher report	Successive engagement contributes to uncertainty and positive student experience.
Guide curiosity (Engel, 2011)	Laboratory	Task performance	Modeling curiosity promotes exploration.
Provide opportunities to invent in the curriculum (Lamnina and Chase, 2019 ; see also Glogger-Frey et al., 2015)	Classroom	Experiment, questionnaire, task performance	Engaging in invention leads to greater curiosity and transfer of learning.
Teach questioning (Clark et al., 2019a,b)	Classroom	QFT intervention, questionnaire, teacher fidelity check	Teaching questioning may positively influence curiosity, but also may negatively affect self-regulation and cognitive engagement.

Table 3 Examples of studies suggesting educational approaches for promoting interest.

<i>Approach</i>	<i>Context</i>	<i>Measures</i>	<i>Findings</i>
Include social support, autonomy, and choice (Nolen, 2007)	Classroom	Longitudinal, mixed methods	Higher levels of interest in writing.
Include text attributes, insert elaboration after main points (Hidi and Baird, 1986)	Classroom	Expository text	Increases interest, increases recall, less forgetting
Provide hands-on activity, computers, puzzles etc. (Swarat et al., 2012; see also Mitchell, 1993)	Classroom	Questionnaires, interviews	Triggers interest.
Incorporate multiple approaches to working with content (Xu et al., 2012)	Classroom	Interviews, observations	Provides new exposure, encourages independence, and promotes interest.
Encourage self-related reflection (Renninger et al., 2014; see also Chin and Chia, 2006)	Out-of-school workshop	Task-based interviews, observations, classroom artifacts	Increases in learning and interest.
Personalize context (Bernacki and Walkington, 2018)	Classroom	Math units, surveys, and tests	Triggers situational interest and increases performance.
Enable participation with others in inquiry-based, project-based or problem posing collaborations and inventions (Dohn, 2013)	Field trip	Observations, interviews	Triggers interest and positive affect
Include new information and meaningful context (Crouch et al., 2018; see also Hoffmann, 2002)	Classroom	Surveys, course performance, interviews	Increases in performance and in interest; differences based on more or less initial interest in physics.
Promote utility-value (Hecht et al., 2021)	Classroom	Utility-value intervention, questionnaires	Increases task interest and performance, specifically for students with less confidence.

or designed by other people and/or the nature of the tasks and activities the learners engage (e.g., Mitchell, 1993; Swarat et al., 2012). Learners also may be supported to self-generate connections to the content (e.g., Hecht et al., 2021).

Learning environments that promote interest development enable individuals to work with the content in multiple ways. For example, by providing tasks such as projects that allow different and repeated opportunities for working with content, as well as various formats for working with and thinking through content such as demonstrations, text, or group projects; see developed descriptions in Nolen, 2007; Pressick-Kilborn, 2015; Xu et al., 2012. They usually include richly textured tasks such as projects or activities, and social support from peers and/or educators. As Xu et al. (2012) showed, the benefit of these types of task is that they lead learners to continue to re-engage, and are beneficial for learners with varying levels of readiness.

Nolen (2007) described two 3-year longitudinal studies conducted in multi-age classrooms and the development of their interest in reading and writing. She demonstrated how situational factors were more likely to trigger and then sustain student interest when the classroom instruction took the form of a literate community—a context in which the community provides the basis for shared learning activity, rather than more traditional teacher-directed practices which highlighted the importance of learning reading and writing. The study of the literate community clearly demonstrated that supporting interest in that type of setting led students to develop their interest. It indicated that children whose interest in reading was initially triggered by social factors, such as peer excitement around popular book series, was then shown to be sustained through discussions and opportunities to share. As the students read more books that were similar to those they thought of as “favorites,” they began to develop a stable interest. Moreover, certain conditions, such as autonomy support and positive experiences with content, facilitated students’ situational interest to develop. Projects that involved choice were supportive of the students’ taking initiative; these were supportive of developing interest (Hidi and Renninger, 2006). It also appeared that time to work on collaborative projects led the youth to describe writing as a fun activity, similarly to the case of Dom and his buddy described above.

On the strength of her data, Nolen (2007) argued that environmental support for student interest is critical, especially in writing as writing is a challenging and effortful task. Nolen’s data also suggested that students are able to generate questions independently. This is especially the case when they are working with tasks that are complex or challenging enough to enable them to continue to stretch their understanding of the content of their interest (Renninger, 2010). In addition, Nolen’s data highlighted the importance of educators and their instructional process as a support for interest development, a point that Xu et al. (2012) also emphasized. In their study of the classrooms of exemplary African American teachers, Xu et al. noted the importance of the educators’ own interest in the subject that they teach, as well as their interest in their students as learners.

Research on interest regularly demonstrates that interest development benefits sustained engagement and performance (e.g., Crouch et al., 2018; Harackiewicz et al., 2008; Harackiewicz and Priniski, 2018), and also that interest is coordinated with and supports the development of both self-efficacy and self-regulation (e.g., Harackiewicz et al., 2008; Nuutila et al., 2020; Sansone et al., 2015). Findings also confirm that the development of interest is generally a positive process, and that even though interest may fall off because of lack of support or challenge, and can also be re-triggered (e.g., Alexander et al., 2019; Hecht et al., 2019; Renninger and Hidi, 2002).

Concluding thoughts and future study

Curiosity and interest describe two types of information search, and as such, they serve different but overlapping educational purposes. Understanding and studying their relation in practice is needed, although most studies focus only on one or the other construct. Not many of the abundant studies of curiosity have focused on classroom practice. Rather, the research conducted has primarily focused on promoting optimal levels of uncertainty, either through direct instruction to teach questioning and/or to provide experiences that involve recognizing and working with uncertainty. In promoting feelings of uncertainty, the expectation is that educators create a knowledge gap that learners want to close, and through this promote learning. However, closing a knowledge gap may negatively affect cognitive engagement and self-regulation as Clark et al. (2019b) findings suggested. Keisha's case and the Manz and Suárez (2018) study both illustrate that even when learners have generated their own uncertainty, and are provided with information intended to support them to resolve it, they are not likely to return to further think on the questions related to the uncertainty unless they are encouraged to do so.

By contrast, there are numerous studies of interest that have used instructional approaches which can be used in classrooms. These studies have pointed to a wide range of practices that educators may employ to enable learners to make meaningful connections to content. Moreover, as portrayed by both Rosa's and Dom's cases, once such connections are made, learners are then positioned to continue to seek information specific to the connections that they have been able to make, either by themselves, or along with others. Rosa and Dom benefited from continued support and opportunities to think and work with others on the content of their respective interest. Once their interest in that content began to develop, no one needed to tell them that writing was important, or to ask questions. As their case material shows, both children readily engaged in questioning; their developing interest made it rewarding to question and search for answers. The research on interest in the classroom confirms that these practices positively affect achievement, cognitive engagement, feelings of self-efficacy, and self-regulation for students of all ages, regardless of gender, race, or ethnicity (see reviews in Renninger and Hidi, 2016, 2020).

We have suggested that curiosity might be considered a particular type of trigger for interest. There are also advantages to supporting interest development that could lead to self-triggered curiosity questions. There are relatively few studies that have explicitly considered the relation between curiosity and interest development in the classroom. This situation means that there are open questions about the application of curiosity and interest in practice that need to be considered. It is recommended that future studies address how the relation between individuals' level of curiosity and interest development can be best supported. They also should consider whether the level of individuals' curiosity influences the design of curiosity-promoting activity and/or activity that promotes interest development. If so, this could suggest the need to consider whether and when it is useful to teach students to question, and/or to promote and reflect on uncertainty; it would also provide information about what type of follow-up activity, including reflection, might enable interest to be triggered and sustained.

Acknowledgments

We thank J. Melissa Running for her editorial support in preparing this chapter for publication, and gratefully acknowledge support from the Swarthmore College Faculty Research Fund.

References

- Ainley, M., Hidi, S., 2014. Interest and enjoyment. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *The International Handbook of Emotions in Education*. Taylor and Francis, New York and London, pp. 205–227. <https://doi.org/10.4324/9780203148211.ch11>.
- Alexander, J.M., Johnson, K.E., Neitzel, C., Renninger, K.A., Hidi, S.E., 2019. Multiple Points of Access for Supporting Interest in Science. In: *The Cambridge Handbook of Motivation and Learning*. Cambridge University Press, Cambridge, pp. 312–352. <https://doi.org/10.1017/9781316823279.015>.
- Azevedo, F.S., 2013. The tailored practice of hobbies and its implication for the design of interest-driven learning environments. *J. Learn. Sci.* 22 (3), 462–510. <https://doi.org/10.1080/10580406.2012.730082>.
- Bergin, D.A., 2016. Social influences on interest. *Educ. Psychol.* 51 (1), 7–22. <https://doi.org/10.1080/00461520.2015.1133306>.
- Berlyne, D.E., 1960. *Conflict, Arousal, and Curiosity*. McGraw-Hill Book Company, New York. <https://doi.org/10.1037/11164-000>.
- Bernacki, M.L., Walkington, C., 2018. The role of situational interest in personalized learning. *J. Educ. Psychol.* 110 (6). <https://doi.org/10.1037/edu0000250>.
- Chin, C., Chia, L., 2006. Problem-based learning: using ill-structured problems in biology project work. *Sci. Educ.* 90 (1), 44–67. <https://doi.org/10.1002/sce.20097>.
- Clark, S., Harbaugh, A.G., Seider, S., 2019a. Teaching questioning fosters adolescent curiosity: analyzing impact through multiple-group structural equation modeling. *Appl. Dev. Sci.* 1–20. <https://doi.org/10.1080/10888691.2019.1591956>.
- Clark, S., Harbaugh, A.G., Seider, S., 2019b. Fostering adolescent curiosity through a question brainstorming intervention. *J. Adolesc.* 75, 98–112. <https://doi.org/10.1016/j.adolescence.2019.07.007>.
- Crouch, C.H., Wisittanawat, P., Cai, M., Renninger, K.A., 2018. Life science students' attitudes, interest, and performance in introductory physics for life sciences: an exploratory study. *Phys. Rev. Phys. Educ. Res.* 14 (1). <https://doi.org/10.1103/PhysRevPhysEducRes.14.010111>.
- Digman, J.M., 1990. Personality structure: emergence of the five-factor model. *Annu. Rev. Psychol.* 41, 417–440. <https://doi.org/10.1146/annurev.ps.41.020190.002221>.
- Dohn, N.B., 2013. Situational interest in engineering design activities. *Int. J. Sci. Educ.* 35 (12), 2057–2078. <https://doi.org/10.1080/09500693.2012.757670>.
- Eccles, J., Wigfield, A., Harold, R.D., Blumenfeld, P., 1993. Age and gender differences in children's self- and task perceptions during elementary school. *Child Dev.* 64 (3), 830–847. <https://doi.org/10.2307/1131221>.
- Engel, S., 2011. Children's need to know: curiosity in schools. *Harv. Educ. Rev.* 81 (4), 625–645. <https://doi.org/10.17763/haer.81.4.h054131316473115>.
- Furtak, E.M., Seidel, T., Iverson, H., Briggs, D.C., 2012. Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev. Educ. Res.* 82 (3), 300–329. <https://doi.org/10.3102/0034654312457206>.

- Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., Renkl, A., 2015. Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learn. Instruct.* 39. <https://doi.org/10.1016/j.learninstruc.2015.05.001>.
- Gottlieb, J., Oudeyer, P.Y., Lopes, M., Baranes, A., 2013. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends Cognit. Sci.* 17 (11), 585–593. <https://doi.org/10.1016/j.tics.2013.09.001>.
- Grossnickle, E.M., 2016. Disentangling curiosity: dimensionality, definitions, and distinctions from interest in educational contexts. *Educ. Psychol. Rev.* 23, 23–60. <https://doi.org/10.1007/s10648-014-9294-y>.
- Harackiewicz, J.M., Durik, A.M., Barron, K.E., Linnenbrink-Garcia, L., Tauer, J.M., 2008. The role of achievement goals in the development of interest: reciprocal relations between achievement goals, interest, and performance. *J. Educ. Psychol.* 100 (1), 105–122. <https://doi.org/10.1037/0022-0663.100.1.105>.
- Harackiewicz, J.M., Priniski, S.J., 2018. Improving student outcomes in higher education: the science of targeted intervention. *Annu. Rev. Psychol.* 69 (1), 409–435. <https://doi.org/10.1146/annurev-psych-122216-011725>.
- Hecht, C.A., Grande, M.R., Harackiewicz, J.M., 2021. The role of utility value in promoting interest development. *Motiv. Sci.* 7 (1), 1–20. <https://doi.org/10.1037/mot0000182>.
- Hecht, M., Knutson, K., Crowley, K., 2019. Becoming a naturalist: interest development across the learning ecology. *Sci. Educ.* 103 (3), 691–713. <https://doi.org/10.1002/sce.21503>.
- Hidi, S., Anderson, V., 1992. Situational interest and its impact on reading and expository writing. In: Renninger, K.A., Krapp, A., Hidi, S. (Eds.), *The Role of Interest in Learning and Development*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, pp. 215–238.
- Hidi, S., Baird, W., 1986. Strategies for increasing text-based interest and students' recall of expository texts. *Read. Res. Q.* 23 (4), 465–483. <https://doi.org/10.2307/747644>.
- Hidi, S., Berndt, D., Ainley, M., 2002. Children's argument writing, interest and self-efficacy: an intervention study. *Learn. Instruct.* 12 (4), 429–446. [https://doi.org/10.1016/S0959-4752\(01\)00009-3](https://doi.org/10.1016/S0959-4752(01)00009-3).
- Hidi, S.E., Magnifico, A., Renninger, K.A. (in press). Students developing as writers: how and why interest makes a difference. In: Horowitz, R. (Ed.), *The Routledge Handbook of International Research on Writing: Human Development, Culture, and Digital Spaces*. New York: Routledge: Taylor & Francis.
- Hidi, S., Renninger, K.A., 2006. The four-phase model of interest development. *Educ. Psychol.* 41 (2), 111–127. https://doi.org/10.1207/s15326985ep4102_4.
- Hidi, S.E., Renninger, K.A., 2019. Interest development, curiosity, and needed neuroscientific research. *Educ. Psychol. Rev.* 31 (4), 833–852. <https://doi.org/10.1007/s10648-019-09491-3>.
- Hidi, S.E., Renninger, K.A., 2020. On educating, curiosity, and interest development. *Curr. Opin. Behav. Sci.* 35, 99–103. <https://doi.org/10.1016/j.cobeha.2020.08.002>.
- Hidi, S.E., Renninger, K.A., Northoff, G., 2017. The development of interest and self-related processing. In: Guay, F., Marsh, H.W., McInerney, D.M., Craven, R.G. (Eds.), *International Advances in Self Research. Vol. 6: SELF – Driving Positive Psychology and Well-Being*. Information Age Press, Charlotte, pp. 51–70.
- Hidi, S.E., Renninger, K.A., Northoff, G., 2019. The educational benefits of self-related information processing. In: Renninger, K.A., Hidi, S.E. (Eds.), *The Cambridge Handbook of Motivation and Learning*. Cambridge University Press, Cambridge, UK, pp. 15–35.
- Hoffmann, L., 2002. Promoting girls' interest and achievement in physics classes for beginners. *Learn. Instruct.* 12 (4), 447–465. [https://doi.org/10.1016/S0959-4752\(01\)00010-X](https://doi.org/10.1016/S0959-4752(01)00010-X).
- Jansen, M., Lüdtke, O., Schroeders, U., 2016. Evidence for a positive relation between interest and achievement: examining between-person and within-person variation in five domains. *Contemp. Educ. Psychol.* 46, 116–127. <https://doi.org/10.1016/j.cedpsych.2016.05.004>.
- Jirout, J.J., 2020. Supporting early scientific thinking through curiosity. *Front. Psychol.* 11, 1717. <https://doi.org/10.3389/fpsyg.2020.01717>.
- Jirout, J., Klahr, D., 2012. Children's scientific curiosity: in search of an operational definition of an elusive concept. *Dev. Rev.* 32 (2), 125–160. <https://doi.org/10.1016/j.dr.2012.04.002>.
- Jirout, J.J., Vitiello, V.E., Zumbro, S.K., 2018. Curiosity in schools. In: Gordon, G. (Ed.), *Psychology of Emotions, Motivations, and Actions. The New Science of Curiosity*. Nova Science, New York, pp. 243–265.
- Lammina, M., Chase, C.C., 2019. Developing a thirst for knowledge: how uncertainty in the classroom influences curiosity, affect, learning, and transfer. *Contemp. Educ. Psychol.* 59, 101785. <https://doi.org/10.1016/j.cedpsych.2019.101785>.
- Liquin, E.G., Lombrozo, T., 2020. A functional approach to explanation-seeking curiosity. *Cognit. Psychol.* 119, 101276. <https://doi.org/10.1016/j.cogpsych.2020.101276>.
- Litman, J.A., Jimerson, T.L., 2004. The measurement of curiosity as a feeling of deprivation. *J. Pers. Assess.* 82 (2). https://doi.org/10.1207/s15327752jpa8202_3.
- Loewenstein, G., 1994. The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* 116 (1), 75–98. <https://doi.org/10.1037/0033-2909.116.1.75>.
- Luce, M.R., Hsi, S., 2014. Science-relevant curiosity expression and interest in science: an exploratory study. *Sci. Educ.* 99 (1), 70–97. <https://doi.org/10.1002/sce.21144>.
- Manz, E., Suárez, E., 2018. Supporting teachers to negotiate uncertainty for science, students, and teaching. *Sci. Educ.* 102 (4). <https://doi.org/10.1002/sce.21343>.
- Mitchell, M., 1993. Situational interest: its multifaceted structure in the secondary school mathematics classroom. *J. Educ. Psychol.* 85 (3), 424–436. <https://doi.org/10.1037/0022-0663.85.3.424>.
- Murayama, K., FitzGibbon, L., Sakaki, M., 2019. Process account of curiosity and interest: a reward-learning perspective. *Educ. Psychol. Rev.* 31, 875–895. <https://doi.org/10.1007/s10648-019-09499-9>.
- Nolen, S.B., 2007. Young children's motivation to read and write: development in social contexts. *Cognit. Instruct.* 25 (2–3). <https://doi.org/10.1080/07370000701301174>.
- Nuutila, K., Tapola, A., Tuominen, H., Kupiainen, S., Pásztor, A., Niemivirta, M., 2020. Reciprocal predictions between interest, self-efficacy, and performance during a task. *Front. Educ.* 5. <https://doi.org/10.3389/educ.2020.00036>.
- Peterson, E.G., Hidi, S., 2019. Curiosity and interest: current perspectives. *Educ. Psychol. Rev.* 31 (4), 781–788. <https://doi.org/10.1007/s10648-019-09513-0>.
- Pressick-Kilborn, K., 2015. Canalization and connectedness in the development of science interest. In: Renninger, K.A., Nieswandt, M., Hidi, S. (Eds.), *Interest in Mathematics and Science Learning*. American Educational Research Association, Washington, DC, pp. 353–368. <http://www.jstor.org/stable/j.ctt1s474j0.26>.
- Renninger, K.A., 2000. Individual interest and intrinsic motivation. In: Sansone, C., Harackiewicz, J.M. (Eds.), *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*. Academic, New York, pp. 373–404.
- Renninger, K.A., 2010. Working with and cultivating the development of interest, self-efficacy, and self-regulation. In: Dai, D., Sternberg, R.M. (Eds.), *Innovations in Educational Psychology: Perspectives on Learning, Teaching, and Human Development*. Springer Publishing Company, New York, pp. 107–138.
- Renninger, K.A., Austin, L., Bachrach, J.E., Chau, A., Emmerson, M.S., King, B.R., Stevens, S.J., 2014. Going beyond the "whoa! that's cool!" of inquiry: achieving science interest and learning with the ICAN intervention. *Adv. Motivation Achiev.* 18. <https://doi.org/10.1108/S0749-742320140000018003>.
- Renninger, K.A., Bachrach, J.E., Hidi, S.E., 2019. Triggering and maintaining interest in early phases of interest development. *Learn. Cult. Soc. Interact.* 23, 100260. <https://doi.org/10.1016/j.lcsi.2018.11.007>.
- Renninger, K.A., Hidi, S., 2002. Student interest and achievement: developmental issues raised by a case study. In: Wigfield, A., Eccles, J.S. (Eds.), *Development of Achievement Motivation*. Academic Press, San Diego, CA, US, pp. 173–195. <https://doi.org/10.1016/B978-012750053-9/50009-7>.
- Renninger, K.A., Hidi, S.E., 2016. *The Power of Interest for Motivation and Engagement*. Routledge/Taylor & Francis Group, New York.
- Renninger, K.A., Hidi, S.E., 2019. *The Cambridge Handbook of Motivation and Learning*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781316823279>.
- Renninger, K.A., Hidi, S.E., 2020. To level the playing field, develop interest. *Policy Insights Behav. Brain Sci.* 7 (1), 1–9. <https://doi.org/10.1177/2372732219864705>.
- Renninger, K.A., Hidi, S.E., 2021. Interest development, self-related information processing, and practice. *Theory Into Pract.* <https://doi.org/10.1080/00405841.2021.1932159>.
- Renninger, K.A., Hidi, S.E., 2022. In: Elliot, A. (Ed.), *Adv. Motiv. Sci.*, 8. Elsevier, New York.
- Renninger, K.A., Su, S., 2012. Interest and its development. In: Ryan, R.M. (Ed.), *Oxford Library of Psychology. The Oxford Handbook of Human Motivation*. Oxford University Press, Oxford, pp. 167–187.
- Renninger, K.A., Wozniak, R.H., 1985. Effect of interest on attentional shift, recognition, and recall in young children. *Dev. Psychol.* 21 (4), 624–632. <https://doi.org/10.1037/0012-1649.21.4.624>.

- Rothstein, D., Santana, L., 2011. *Make Just One Change: Teach Students to Ask Their Own Questions*. Harvard Education Press, Cambridge, MA, US.
- Sansone, C., Thoman, D., Fraughton, T., 2015. The relation between interest and self-regulation in mathematics and science. In: Renninger, K.A., Nieswandt, M., Hidi, S. (Eds.), *Interest in Mathematics and Science Learning*. American Educational Research Association, Washington, DC, pp. 111–131.
- Shin, D.D., Kim, S., 2019. Homo curious: curious or interested? *Educ. Psychol. Rev.* 31 (4), 853–874. <https://doi.org/10.1007/s10648-019-09497-x>.
- Smock, C.D., Holt, B.G., 1962. Children's reactions to novelty: an experimental study of "curiosity motivation". *Child Dev.* 33 (3). <https://doi.org/10.2307/1126663>.
- Swarat, S., Ortony, A., Revelle, W., 2012. Activity matters: understanding student interest in school science. *J. Res. Sci. Teach.* 49 (4). <https://doi.org/10.1002/tea.21010>.
- van Schijndel, T.J.P., Jansen, B.R.J., Raijmakers, M.E.J., 2018. Do individual differences in children's curiosity relate to their inquiry-based learning? *Int. J. Sci. Educ.* 40 (9), 996–1015. <https://doi.org/10.1080/09500693.2018.1460772>.
- Tang, X., Renninger, K.A., Hidi, S., Murayama, K., Lavonen, J., Salmela-Aro, K., 2020. The Differences and Similarities between Curiosity and Interest: Meta-Analysis and Network Analyses. *PsyArXiv*. <https://doi.org/10.31234/osf.io/wfprn>.
- Webb, N.M., Palincsar, A.S., 1996. *Group Processes in the Classroom*. Prentice Hall International, London, England, pp. 841–873.
- Wu, P., Kuo, C., Wu, H., Jen, T., Hsu, Y., 2018. Learning benefits of secondary school students' inquiry-related curiosity: a cross-grade comparison of the relationships among learning experiences, curiosity, engagement, and inquiry abilities. *Sci. Educ.* 102 (5), 917–950. <https://doi.org/10.1002/sce.21456>.
- Xu, J., Coats, L.T., Davidson, M.L., 2012. Promoting student interest in science: the perspectives of exemplary African American teachers. *Am. Educ. Res. J.* 49 (1), 124–154. <https://doi.org/10.3102/0002831211426200>.

Emotions in education

Thomas Goetz^a, Anne C. Frenzel^b, Kristina Stockinger^c, Anastasiya A. Lipnevich^d, Lisa Stempfer^a, and Reinhard Pekrun^{e,f,b},

^a Department of Developmental and Educational Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria;

^b Department of Psychology, University of Munich, Munich, Germany; ^c Faculty of Philosophy and Social Sciences, University of Augsburg, Augsburg, Germany; ^d Queens College and the Graduate Center, The City University of New York, New York, NY, United States; ^e Department of Psychology, University of Essex, Colchester, United Kingdom; and ^f Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia

© 2023 Elsevier Ltd. All rights reserved.

Conceptualization and measurement of emotions in education	149
Conceptualization	149
Measurement	150
Emotions in education	150
Short history of emotions in education	150
Groups of academic emotions	151
Origins of emotions	152
General approaches to the origins of emotions	152
Origins of emotions in the context of education	152
Appraisal antecedents of achievement emotions	152
Social-cognitive antecedents of emotional appraisals	153
Relevance of emotions for learning and performance	153
Relevance of emotions for students	153
Relevance of emotions for teachers	154
Designing emotionally healthy learning environments	155
Influencing students' and teachers' emotional experiences during classroom instruction	156
Promoting enjoyment of learning in the classroom	156
Supporting students' feelings of competence, enhancing autonomy, emphasizing learning, and de-emphasizing social comparison	156
Optimizing learners' and teachers' emotional experiences by supporting mutual feelings of relatedness	156
Optimizing test-taking and feedback	156
Supporting and fostering adaptive emotion regulation	157
Conclusion	157
References	157

Conceptualization and measurement of emotions in education

Conceptualization

What do we mean, from a psychological perspective, when we speak of emotions? Despite certain nuanced differences in definitions, there is a general agreement in the field about the multi-componential nature of emotions. In line with Scherer (2000); see also Scherer and Moors (2019), we describe emotions as follows: An emotion is an evaluative reaction to an important external or internal stimulus (or mental representation thereof; e.g., a fond memory) that encompasses various psychological and physical subsystems (affective, cognitive, motivational, physiological, and expressive) and serves to be able to react as quickly and flexibly as possible to a given situation.

This multi-componential conceptualization makes an emotion a more or less clear-cut psychological phenomenon. It also implies that each emotional experience is characterized by a central nervous and peripheral-physiological reaction pattern that is reflected in subjective feelings, motor reactions (especially in facial expressions, but also gestures and posture), motivational tendencies (specifically, approach vs. avoidance tendencies), and cognitions (specifically, emotion-typical thought content).

Various approaches have been proposed to order and categorize different emotions above and beyond their constituent components. Most importantly, dimensional approaches presuppose placing the variety of emotional experiences in a space of a restricted range of meaningful descriptive dimensions. Probably the most widely used dimensional model is the circumplex model of affect that classifies emotional experiences along two orthogonal dimensions, namely valence (ranging from unpleasant to pleasant) and arousal (ranging from calm to alert; see Posner et al., 2005). For example, according to this approach, both anxiety and anger would be categorized as unpleasant, highly arousing emotions. Of course, there are also other ways to categorize emotions, for example, according to their timing (i.e., whether they primarily occur before, during, or after an event) or their object focus (e.g., self- or other-related; social, topic, or outcome-related; see below).

A further conceptual aspect worth considering in the context of emotions is that they can be understood either as states—that is, momentary experiences—or traits—that is, habitual experiences (e.g., trait test anxiety; [Spielberger et al., 1976](#)). Emotional states are often very short-lived. In contrast, when considering emotions as traits, the emphasis is on the fact that individuals differ in their tendency (or disposition, susceptibility, proneness) to experience certain emotional states more readily and frequently, and/or more intensely. In line with the definition of “trait” as a core concept in personality psychology, emotional traits are supposed to be relatively stable over time. As discussed in more depth below, differentiating between emotions as states versus traits represents an important conceptual distinction for scientific inquiry into emotions, and has important implications for their measurement.

Measurement

Due to the fact that emotions are characterized by individuals’ subjective experiences, it stands to reason that they can be assessed quite simply with self-report, that is, by asking individuals about these experiences. This can take the form of open-ended questions, with individuals reporting their emotional experiences through interviews, focus groups, or video-stimulated recall methods. Alternatively, it can take the form of closed-ended Likert-type items pertaining to verbal emotion descriptions, visual stimuli, or a combination of both (see [Schutz et al., 2016](#), for a discussion of tools and approaches to gauging emotions in education). In so doing, researchers need to decide whether to assess emotions as states, by asking, for instance, “How do you feel right now?”, or as traits, by asking “How do you generally feel?”. States are often assessed via so-called “experience sampling methods” ([Larson and Csikszentmihalyi, 2014](#); see also [Goetz et al., 2016a](#)) where participants are, for example, asked to rate their emotional states at multiple, typically randomly determined instances across lessons, days, or weeks. Digital self-reports are also used to explore learners’ state emotional experiences embedded in online learning environments ([D’Mello et al., 2017](#); [Harley, 2016](#)). Traits can be assessed by asking about how one typically feels with respect to specific types of situations (e.g., mathematics tests). An alternative way of indexing trait emotions is obtaining an average across various emotional state measures with respect to a specific type of situation. This way, individuals’ relative tendency to experience a certain emotion in a particular situation is captured.

A widely-used self-report instrument for the assessment of learners’ habitual, trait-like emotional experiences in achievement contexts is the Achievement Emotions Questionnaire (AEQ) developed by Pekrun and colleagues. The measure was originally created for the use with university students ([Pekrun et al., 2011](#)) and later adapted for the use with K-12 students, and contextualized for specific school subjects (e.g., AEQ-mathematics; [Frenzel et al., 2007](#)). The AEQ is intended to assess learners’ class-related, learning-related, and test-related experiences of enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom (sample items: “I enjoy being in class”, “I’m proud of how well I mastered the exam”, “While studying I drift off because it’s so boring”). A short version of the AEQ (i.e., four items for each scale) also exists ([Bieleke et al., 2021](#)). Additionally, drawing upon AEQ, Frenzel and colleagues developed a self-report instrument to assess teachers’ habitual experiences of enjoyment, anger, and anxiety during teaching (Teacher Emotions Scales, TES; [Frenzel et al., 2016](#)) with four items per emotion scale (sample items: “I enjoy teaching”, “I often have reasons to get angry when I teach”, “I am often worried that my teaching isn’t going so well”).

Although self-report approaches have face validity and allow for an efficient measurement of emotions, they also entail possible problems because they are prone to conscious and non-conscious biases. For example, a teacher may not want to admit that they are highly excited or agitated during teaching, or students might exaggerate their test anxiety hoping this can result in lenient grading ([Pekrun, 2020](#); [Pekrun and Bühner, 2014](#)). In order to mitigate these potential biases, objective methods of assessing emotions have been developed that are meant to bypass participants’ subjective responses. This might be achieved through qualitative methods of observation or in-depth ethnographic inquiry methods ([Schutz et al., 2016](#)). One can also make use of the expressive component of emotions by coding and quantifying typical emotion-related facial expressions ([Ekman et al., 2002](#)), which can be executed through artificial intelligence-based facial feature recognition (see e.g., [Frenzel et al., 2019](#)). Additionally, one can assess physiological processes associated with emotional experiences. These assessment approaches used to be solely lab-based, but recent technology in the form of so-called “wearables” allows for tracking a number of peripheral-physiological measures including skin-conductance, heart rate, or blood pressure, as individuals engage in certain tasks in educational settings (including classrooms; see e.g., [Donker et al., 2018](#)). Additionally, cortisol found in saliva has been used by researchers as an indicator of increased emotional arousal, and considered in the context of research on fear or stress, including test anxiety ([Mattarella-Micke et al., 2011](#)). Finally, scholars have also employed functional magnetic resonance imaging (fMRI) and electroencephalography (EEG; see e.g., [Dikker et al., 2017](#)) to examine brain-based processes underlying individuals’ emotional experiences. However, with the exception of analyzing facial expressions, self-report still constitutes the only suitable approach for assessing specific emotions (e.g., anxiety, anger, surprise) rather than valence or arousal more broadly, especially in instructional settings (see [Roos et al., 2021](#)).

Emotions in education

Short history of emotions in education

Existing research into emotions in education dates back to inaugural studies on test anxiety, conducted in the 1930s (e.g., [Brown, 1938](#)). Throughout the 20th century, research that focused on achievement emotions was largely restricted to anxiety, and test anxiety continued to be by far the most prominent construct in research on emotions in education. Test anxiety has attracted

researchers' attention since Mandler and Sarason (1952) proposed their habit interference model of anxiety and learning. Empirical evidence has confirmed the importance of anxiety for understanding human performance (Hembree, 1988; Zeidner, 1998; Richardson et al., 2012). Since the late 1960s, attributional research has broadened the perspective by considering cognitive antecedents of different emotions following success and failure, such as pride, gratitude, shame, guilt, and anger (Weiner, 1985). Researchers have further expanded the scope of inquiry during the past 20 years by exploring a broad range of achievement emotions, including, for example, enjoyment, hope, surprise, anger, boredom, disappointment, and hopelessness (Pekrun, 2006, 2018; Pekrun and Linnenbrink-Garcia, 2014; Pekrun et al., 2017b). In addition to the increasing number of different emotions that are being examined in education, the variety of contexts (e.g., achievement vs. non-achievement context), situations (social vs. non-social situation), and academic domains (e.g., mathematics, language domain), in which these emotions are being studied, has been growing exponentially during the last 20 years.

In the context of recent research on academic emotions, different groups of emotions have been identified with respect to activities, situations, and content areas to which emotions refer (i.e., their object focus). These groups include achievement emotions that are strongly related to experiences of success and failure, epistemic emotions that are caused by perceived gaps and changes in knowledge, social emotions (e.g., envy, contempt or affection for others in the classroom), as well as topic-related emotions that are directly triggered by the learning object (Pekrun et al., 2017b).

Groups of academic emotions

An overwhelming proportion of existing studies on emotions in education focuses on *achievement emotions*. Therefore, in this chapter, we pay particular attention to this group of emotions. These are emotions that occur in relation to achievement-related activities (i.e., solving a math task, or memorizing vocabulary; Camacho-Morles et al., 2021) as well as results of these activities (i.e., getting the task right, or failing to remember the correct translation), which, for students, often involve informal as well as formal achievement feedback in the form of grades or teacher comments (see Lipnevich and Smith, 2018). This feedback has been shown to elicit a range of emotions, which relate to improved or reduced motivation and performance. The ubiquity of feedback situations that inevitably elicit a wide variety of achievement emotions in students has become the focus of attention of researchers (Goetz et al., 2018; Lipnevich and Smith, 2009). Interestingly, links between specific emotions and feedback are not straightforward, and are often counter-intuitive. Specifically, positive emotions resulting from praise may tend to result in decreased motivation to invest further effort in a task (e.g., Brummelman, 2020; Lipnevich and Smith, 2009), whereas negative emotions resulting from low grades may enhance student motivation (Lipnevich et al., 2021; Turner and Schallert, 2001). Key examples of achievement emotions in relation to past achievement outcomes are pride, relief, shame, anger, and disappointment. With regard to impending achievement evaluation, anxiety, hopelessness, and hope may be experienced. Finally, joy, frustration, or boredom are typical achievement emotions that are related to achievement activities themselves (e.g., studying, completing in-class activities).

Epistemic emotions have recently started to play a major role in research (Muis et al., 2018; Pekrun and Stephens, 2012). The term "epistemic" is derived from ancient Greek and denotes thoughts and activities that aim to expand human knowledge, as exemplified by the pioneering thinking of the Greek philosophers. Thinking, however, is not just based on pure cognitive reasoning alone ("cold cognition"). Rather, it is closely tied to emotions such as surprise, curiosity, or confusion, which relate to the knowledge-generating qualities of cognitive tasks and drive the exploratory activities that are needed to expand knowledge. For these emotions, knowledge and the generation of knowledge are the objects of emotions (Pekrun, 2021a). A prototypical situation for the arousal of epistemic emotions is discrepant information and cognitive incongruity, which implies that different pieces of information are not compatible and do not fit together. A typical sequence of epistemic emotions induced by cognitive incongruity may involve (1) surprise; (2) curiosity and situational interest if the surprise is not resolved; (3) confusion when the incongruity cannot be resolved; (4) anxiety in case of severe incongruity and information that deeply disturbs existing beliefs; (5) enjoyment and delight experienced when the incongruity gets solved; or (6) frustration when solving the incongruity seems impossible.

There is also an increasing interest in *social emotions* and how they manifest in educational contexts (e.g., Immordino-Yang, 2010), although this area is less explored compared with work on achievement emotions and epistemic emotions. Learning is situated in social contexts. Even when studying alone, students never find themselves in a social vacuum; rather, the goals, contents, and outcomes of learning are socially constructed. By implication, academic settings induce a multitude of social emotions that include both social achievement emotions (such as admiration, envy, contempt, or empathy related to the success and failure of others) and non-achievement social emotions (such as love or hate in relationships with classmates and teachers). Similarly, teachers' professional activities are social in nature. Classroom teaching, interacting with colleagues, and giving advice to parents elicits a range of social emotions. Even if other people are not physically present, for example, during lesson preparation or marking exams at home, social emotions may still be prominent (e.g., getting angry at administrators for requiring detailed lesson plans).

Finally, there is the group of *topic emotions*, that is, emotions that are triggered by the contents of a learning task (Pekrun and Linnenbrink-Garcia, 2022). This group of emotions seems to be the least researched. Examples are empathic emotions pertaining to a protagonist's fate when reading a novel, emotions triggered by political events discussed during politics lessons, or emotions related to topics in science class, such as students' anxiety when told about predictions of rising temperatures related to global climate change. In contrast to achievement and epistemic emotions, topic emotions do not directly pertain to learning and problem solving, and are much narrower and more task-specific.

Origins of emotions

General approaches to the origins of emotions

Why do we experience emotions in the first place? How are emotions triggered? According to a biological perspective that considers basic physiological processes of emotions from the standpoint of the phylogenetic development of humans, emotions offer a survival advantage by preparing the organism to act adaptively. In contrast to instincts which imply highly automatic stimulus-response couplings, emotions provide the basic direction of behavior (e.g., approach or avoidance), which can then be adapted to the specific situation. Cognitive approaches to emotions focus on the extent to which cognitions (such as relevance judgments, attribution of causes of events) are involved in the genesis of emotions. Particularly influential are appraisal theories, which state that it is not situations themselves that evoke emotions, but rather, our interpretations and subjective evaluation of these situations. Finally, researchers who take a social-constructivist approach to emotions argue that they are social constructions strongly shaped by cultural and societal factors. For further reading on the origins of emotions, we encourage the reader to consult the Handbook of Emotions (Barrett et al., 2016; for further reading on appraisal theory, see Moors et al., 2013; Scherer and Moors, 2019).

Origins of emotions in the context of education

Appraisal antecedents of achievement emotions

A theory that addresses a broad range of emotions in education is Pekrun's control-value theory (CVT) of achievement emotions (Pekrun, 2006, 2018). This theory focuses primarily on achievement emotions (for a more generalized version of the CVT that also addresses social, epistemic, and other emotions, see Pekrun, 2021a). In addition to being grounded in appraisal theory, CVT integrates propositions from transactional theories of stress-related emotions (Lazarus and Folkman, 1984), attribution theory (Graham and Taylor, 2014; Weiner, 2007), and has a conceptual relation with expectancy-value theories of achievement motivation (Eccles and Wigfield, 2002; Pekrun, 1993; Wigfield and Cambria, 2010). The CVT posits that two appraisals are particularly important for the arousal of achievement emotions, namely, individuals' subjective control over learning- and performance-related activities and outcomes, and their subjective value of these activities and outcomes (see Fig. 1). Subjective control refers to one's perception of being able (or unable) to influence achievement activities and their results. Value appraisals have a categorical and a dimensional facet, in line with value concepts in other disciplines (e.g., economics). The categorical facet pertains to judgments of achievement activities and their results as desirable (positive value) or undesirable (negative value), whereas the dimensional facet pertains to how important or personally significant the activities or achievement results are (low vs. high positive or negative value).

According to the CVT, how one feels in a given achievement situation depends on the individual's combination of current control and value appraisals. Control determines our expectation of success or failure and, consequently, whether we experience positive (i.e., pleasant) or negative (i.e., unpleasant) emotions. For example, if perceived control over an upcoming exam is high and we expect success, we will tend to experience enjoyment. If, however, perceived control over achievement outcomes is very low, we may experience hopelessness. The value appraisal determines how important success or failure is in a given situation. Control and value appraisals jointly determine the intensity of the emotional experience. That is, the higher one's perceived control over the process or the outcome of a task, and the higher their positive value, the stronger the resulting positive emotions are.

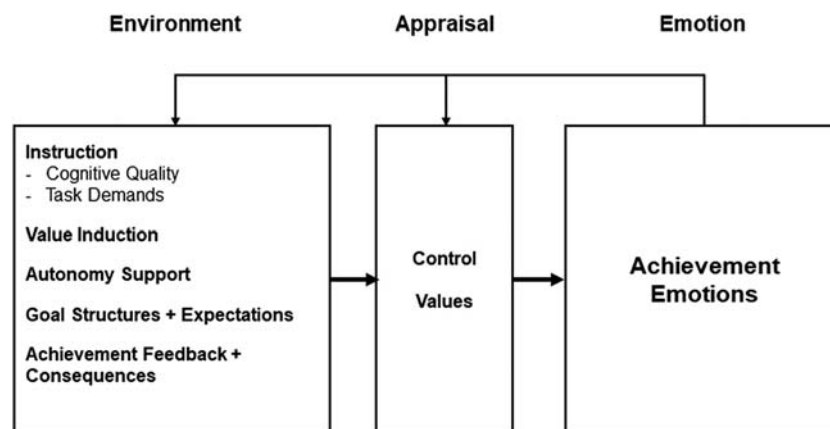


Fig. 1 The control-value theory related to the antecedents of achievement emotions: Reciprocal linkages between antecedents and emotions. Adapted from Pekrun (2006).

Conversely, from low control and high negative value, intense negative emotions will result. One exception with regard to the latter contingency is boredom, in that boredom is less pronounced when a learning task or a situation is considered important or personally relevant (Goetz et al., 2019).

Links between students' control and value appraisals, and a range of achievement emotions have long been supported by extensive empirical evidence (e.g., Forsblom et al., 2022; Pekrun and Perry, 2014; Loderer et al., 2020; Putwain et al., 2018a; Shao et al., 2020). Typically, studies addressing those links used students' self-concepts or perceived competence as proxies for control appraisals, and their beliefs about the importance of achievement, and/or the personal relevance of the material to be learned, as proxies for value appraisals, focusing on trait-level mechanisms. Studies addressing links between situation-specific judgments and state achievement emotions are relatively rare (e.g., Ahmed et al., 2010; Goetz et al., 2010, 2020). Interestingly, trait- and state-level approaches provide convergent evidence regarding the direction of relationships as proposed by the CVT (Goetz et al., 2016b), implying that emotional mechanisms operate in similar ways at both levels of analysis.

Social-cognitive antecedents of emotional appraisals

Assuming that cognitive appraisals are important antecedents of emotions, the question arises, what accounts for a situation being appraised in a certain way? In sum, appraisals reside in individuals and their personal beliefs (i.e., predispositions, biases, schemas, etc.). These personal beliefs provide a lens through which individuals perceive their environment, and thus strongly influence how they interpret a situation. The CVT's social-cognitive basis implies that the social environment, as perceived by individuals, plays an important role in shaping their control and value appraisals pertaining to a particular achievement situation. Specifically, Pekrun (2006, 2018) describes a number of environmental factors relevant for achievement emotion appraisals, including (1) facets of instruction, including cognitive quality and task demands, (2) value induction, (3) autonomy support, (4) goal structures and expectations, and (5) achievement, both actual levels and the corresponding feedback and consequences (see Fig. 1). There is extensive empirical evidence showing links between those environmental factors and students' emotions. The mediating role of control and value appraisals has also been consistently demonstrated. For example, Lazarides and Buchholz (2019) provided robust evidence on the links between facets of student-perceived instructional quality pertaining to teacher support and classroom management, and students' achievement emotions. Further, Flunger et al. (2019) provided compelling evidence from a field experiment, showing that if physics teachers taught a standardized teaching unit on heat transfer with significant autonomy support (by providing choices, rationales, and general information in a noncontrolling language) instead of using regular teacher-centered instruction, students experienced more positive achievement emotions, and fewer and less intense negative emotions. There is also consistent evidence that classroom goal structures are systematically linked to students' emotions (e.g., anger, anxiety, and shame showing positive links with performance goal structures, and boredom showing negative links with mastery goal structures; see e.g., Baudoin and Galand, 2017).

Fig. 1 summarizes the main propositions of the CVT related to antecedents of emotions. Importantly, even though this theory is rooted in appraisal theory and thus primarily proposes that environmental factors predict cognitive appraisals, which in turn predict emotional experiences, it also acknowledges that emotions are reciprocally linked to their cognitive antecedents (Goetz et al., 2021). This reasoning is reflected in the recursive arrows included in the model (Pekrun, 2006, 2018).

Relevance of emotions for learning and performance

Relevance of emotions for students

In applied, field-based educational and psychological research, emotions, motivation, behavior, and achievement outcomes are typically gauged together, often through self-reports, and often targeting specific domains, such as math or languages. There is substantial evidence demonstrating that emotions and achievement outcomes are systematically linked. Researchers have proposed that the link between achievement emotions and learning outcomes may be mediated by a range of cognitive, motivational, and self-regulatory processes (Pekrun, 2006, 2018).

Indeed, there is consistent evidence showing that negative emotions including anxiety, shame, anger, boredom, or hopelessness during learning are related to unfavorable cognitive processes that include task-irrelevant thinking, shallow learning strategies, or less sophisticated metacognitive strategies. Conversely, enjoyment of learning relates negatively to irrelevant thinking, and positively to a better task focus and more effective self-regulation of learning (see e.g., Ahmed et al., 2013; Obergrösser and Stoeger, 2020).

Research, describing links between achievement emotions and achievement motivation during learning, is also ample. For example, studies have demonstrated that positive emotions and mastery goals (specifically, mastery approach goals) were positively linked (see Huang, 2011, for a meta-analysis). Similarly, positive emotions, and enjoyment in particular, are key drivers and an inherent component of intrinsic motivation (Isen and Reeve, 2005). Conversely, negative emotions (e.g., anxiety, boredom, and anger) may focus students on consequences of their potential failure at the task or the unpleasant nature of the task, and are thus linked with less self-determined types of motivation. It is worth noting, however, that research on achievement emotions and on intrinsic versus extrinsic motivation has rarely been integrated (for exceptions, see e.g., González et al., 2012; Sutter-Brandenberger et al., 2018).

A number of studies has also explored the overall relations between achievement emotions and achievement. There is consistent evidence that these relations are generally positive for pleasant/positive emotions, and negative for unpleasant/negative emotions (for an overview, see Goetz and Hall, 2020; for meta-analytic evidence, see Camacho-Morles et al., 2021; Hembree, 1988; Loderer et al., 2020; Tze et al., 2016; Zhang et al., 2019). Recent studies examined these links by employing longitudinal designs, and their findings indicated that the observed correlations were driven by reciprocal effects between emotions and academic performance (Forsblom et al., 2022; Pekrun et al., 2017a; Putwain et al., 2018b; Putwain et al., 2022). Although there is evidence supporting contingencies of “pleasant = beneficial for performance” and “unpleasant = detrimental for performance”, these links may be considerably more nuanced. For example, Pekrun (2006, 2018) suggested a further differentiation between activating (i.e., anger, enjoyment) and deactivating achievement emotions (i.e., hopelessness and boredom), resulting in a 2-by-2 taxonomy. According to this reasoning, typically positive effects are expected for pleasant and activating emotions such as enjoyment and pride, and clear-cut negative effects are expected for negative-deactivating emotions. Effects on academic achievement are, in contrast, expected to be more variable “for both deactivating positive emotions, such as relaxation, and activating negative emotions, such as anxiety” (Pekrun, 2006, p. 326). However, for many emotions the evidence for these hypothesized emotion-achievement links is still scarce and in need of further exploration. Notable exceptions are enjoyment, anxiety, and boredom, with a substantial body of research accumulated for these emotions (Camacho-Morles et al., 2021; for confusion see D’Mello et al., 2014; for shame see Turner and Schaller, 2001). So, on average, students’ enjoyment of learning is positively correlated with their academic achievement, whereas correlations for anxiety and boredom are negative (Barroso et al., 2021; Camacho-Morles et al., 2021). Anxiety can be both beneficial and detrimental for performance; however, it remains controversial if the anxiety-performance relation follows an inverted U-shaped form for academic learning. For high levels of anxiety, negative linkages with achievement are expected (for evidence, see e.g., Sung et al., 2016).

In sum, there is compelling evidence demonstrating that emotions influence academic achievement and vice versa. In fact, emotions, cognitions, motivation, learning behaviors, and achievement mutually affect each other. More research on the complex links among these variables is needed (e.g., for links between emotions, appraisals, and achievement, see Forsblom et al., 2022).

Relevance of emotions for teachers

Historically, the focal point of research on emotions in the context of education has been on the learner. Although it has long been acknowledged that teaching is a social and highly emotional practice (D’Eon et al., 2000; Goldstein and Lake, 2000; Hargreaves, 1998; Nias, 1996), research on teachers’ emotions has been predominantly descriptive and far less systematic. Recently however, research focusing on teachers’ psychological characteristics, including their emotions, has been steadily growing. As with students’ emotions, appraisal theory has been proposed to be a meaningful framework for the study of teacher emotions (e.g., Chang, 2009; Fried et al., 2015; Tsouloupas et al., 2010). Appraisal theory is also at the heart of Frenzel and colleagues’ reciprocal model of causes and effects of teacher emotions (Frenzel, 2014; Jacob et al., 2017; see Fig. 2). In line with Butler (2007), who argued that “the school is an achievement arena not only for students but also for teachers who presumably strive to succeed at their job” (p. 242), Frenzel (2014) proposed that teachers’ appraisals concerning their successes or failures regarding their teaching goals were key to understanding the antecedents of teacher emotions. The researchers showed that those teachers who perceived themselves or their teaching as contributing to high student motivation, discipline, and high-quality teacher-student relationships also reported more enjoyment and less anxiety and anger, and that those links were also reflected when examined at the within-person level (Frenzel et al., 2020). That is, teachers’ enjoyment of teaching was higher when students performed better, were motivated and disciplined, and when they perceived they had better relationships with their students.

Further, it can be assumed that teachers’ emotions, due to their effects on motivation, decisions (e.g., choice of teaching strategy), and mental flexibility, in turn, affect their teaching. Studies revealed that teacher emotions were linked with how well teachers activated their students cognitively and motivationally, how well they succeeded in classroom management, and how well they supported their students socially and emotionally. As a result, recent findings showed that teachers’ emotions were systematically related to student outcomes (see Frenzel et al., 2021, for a review, and see Pekrun, 2021b, for discussion). Specifically, teachers’ pleasant emotions have been consistently linked with student motivation, enjoyment of learning, self-regulation, and performance (Babad, 2007; Banerjee et al., 2017; Beilock et al., 2010; Frenzel et al., 2016; Keller et al., 2016), and teachers’ unpleasant emotions have been linked with undesirable student outcomes including disruptive behavior, anxiety, and reduced performance (Aldrup et al., 2018; Arens and Morin, 2016; Chang, 2013; Klusmann et al., 2016).

In general, instructional behavior and teachers’ emotions have important effects on learners. This results in positive or negative cycles of teachers’ emotions, their teaching behavior, and student outcomes. When considering emotions, it becomes evident that educational psychological research can benefit from moving beyond seeing the teacher as the key effect-exerting factor driving student outcomes. As much as teachers affect students, the reverse is also true: Students’ individual or class-level characteristics, including their own emotional experiences, level of prior knowledge and performance, motivation, discipline, and their attitude toward the teacher, can also affect teachers and their emotional well-being (Frenzel et al., 2021; Nurmi and Kiuru, 2015).

When it comes to mutual influences between teachers’ and students’ emotions, research on social interaction has shown that people can “infect” each other emotionally (Elfenbein, 2014; Hatfield et al., 1994; Parkinson, 2011). Studies by Frenzel and colleagues (Frenzel et al., 2009, 2018) provided initial support for the transmission of enjoyment between mathematics teachers and their students. In a longitudinal study, the authors showed that emotional transmission was based on the mutual perception

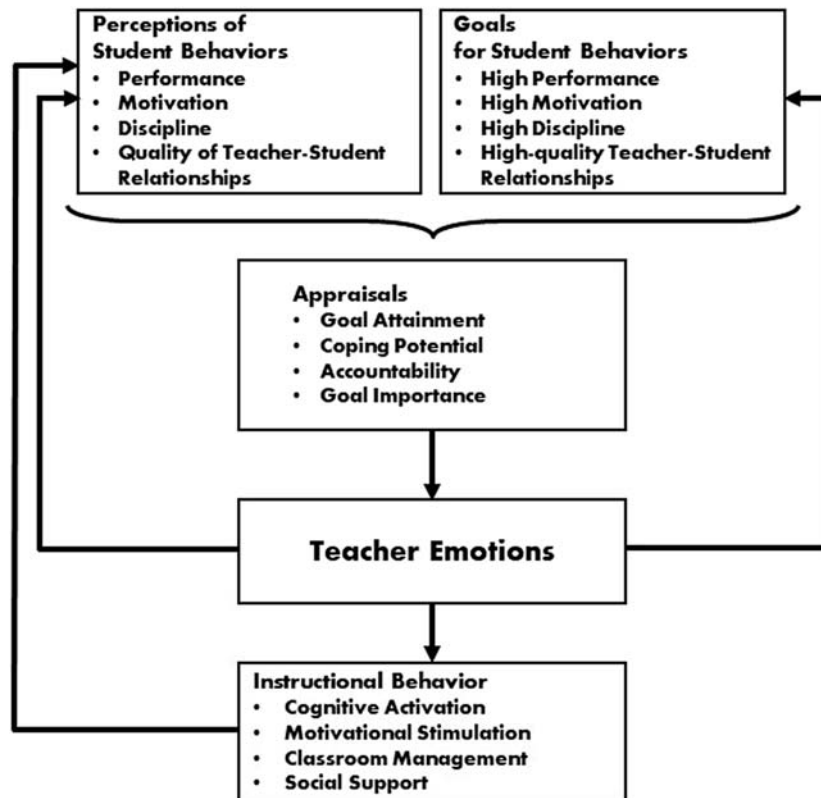


Fig. 2 The reciprocal model of causes and effects of teacher emotions. Figure available at: <https://doi.org/10.6084/m9.figshare.14493789.v2>, under CC license. (<https://creativecommons.org/licenses/by/4.0/>).

of behavior in the classroom (Frenzel et al., 2018). That is, positive effects of teachers' joy on students' joy were mediated by the perceived enthusiasm of teachers, and positive effects of students' joy on teachers' joy were mediated by the perceived motivational commitment of students. Similarly, Moskowitz and Dewaele (2019) showed that there are links among students' perceptions of their language, teachers' happiness, and motivation toward learning English. Using experience sampling methods, Poon et al. (2019) showed that teachers' nervousness is associated with higher subsequent levels of nervousness and irritability among students, which, in turn, impairs students' in-class satisfaction.

As such, it is imperative to consider not only students', but also teachers' emotional experiences in the context of learning and teaching. These variables represent important psychological outcomes in their own right, but they are also linked with academically meaningful cognitive outcomes, that include instructional quality and student achievement, and they are reciprocally linked with each other.

Designing emotionally healthy learning environments

A key implication of the empirical findings and theoretical contentions presented in this chapter is that it is not enough to consider only achievement outcomes, "cool cognitions", instructional ideals, or didactic techniques when designing learning environments. Educators should also tailor instructional design toward optimizing teachers' and learners' affective outcomes as critical preconditions for personal success and well-being. Such contentions have also been supported by advocates of social-emotional learning (SEL) programs who emphasize the importance of focusing on social and emotional development of students, in addition to their cognitive development (for a meta-analytic review showing that SEL programs also positively affect student achievement, see Corcoran et al., 2018, and Murano et al., 2020; for further reading, see Weissberg et al., 2015). Below we outline some ideas on how learning environments can be optimized, taking teachers' and students' emotional experiences into account. These ideas are inspired by the literature on achievement motivation and emotions (e.g., Pekrun's CVT, 2006, 2018, 2021a), emotion regulation, and SEL. These research traditions have operated in relative isolation, but they tend to converge in their implications as to how learning environments can be emotionally optimized.

Influencing students' and teachers' emotional experiences during classroom instruction

Promoting enjoyment of learning in the classroom

As emphasized by Linnenbrink-Garcia et al. (2016), in order to optimize the emotional tone in the classroom, students' enjoyment of learning should be deliberately fostered. An effective way to promote enjoyment and interest in the instructional content is to implement personally relevant and active tasks, for example, by situating lessons in the context of real-world applications of interest to students, and by allowing them to engage with study materials across various settings (Design Principle 3 suggested by Linnenbrink-Garcia et al., 2016). It is worth noting, however, that focusing on boosting students' mood and cheerful atmosphere is not enough to promote learning. It is far more important to realize that student attention and effort is more likely to be sustained when learning is enjoyable.

Supporting students' feelings of competence, enhancing autonomy, emphasizing learning, and de-emphasizing social comparison

These are design principles suggested by Linnenbrink-Garcia et al. (2016) that are of relevance to the control component of Pekrun's CVT. Students will tend to feel more in control over their learning activities and achievement outcomes when they feel competent in those tasks. The perception of mastery can be achieved by offering well-structured, reasonably paced, and clear instruction, and by implementing optimally challenging tasks at, or just slightly above, students' current skill levels. Students' feelings of autonomy can be supported with ample opportunities for student decision-making, and minimizing the use of extrinsic awards and practices that make students feel they are being controlled (e.g., avoiding telling them where to sit, or when to speak; Reeve, 2009). The potentially controlling nature of extrinsic rewards, however, may be employed to establish an initial level of motivation, but autonomy and agency should be subsequently fostered. Additionally, fostering mastery goals and taking focus away from performance goals is another key design principle for emotionally healthy classrooms. Despite the fact that many students enjoy competition, social comparison should not be sustained. In other words, games and activities that presuppose a single winner make success a scarce resource in the classroom, and suggest that success is limited to very few individuals (see Pekrun et al., 2019). Hence, this kind of instructional arrangement is not conducive to positive emotional classroom climate.

Optimizing learners' and teachers' emotional experiences by supporting mutual feelings of relatedness

As their fifth design principle for adaptive motivation and emotion in education, Linnenbrink-Garcia et al. (2016) suggested supporting students' feelings of relatedness and belonging. This notion can be extended to teachers' emotions and motivation as well. Students report enhanced interest and reduced negative affect when the perceived quality of teacher-student relationship is high (Goetz et al., 2020; Quin, 2017; Roorda et al., 2011). Similarly, the more connected teachers feel to their students, the higher their emotional well-being is shown to be (e.g., Klassen et al., 2012; Spilt et al., 2011). Feelings of relatedness between teachers and students can be promoted by spending nondirective one-on-one time (also referred to as "banking time"; cf. Driscoll et al., 2011; Pianta and Hamre, 2001) or by enhancing perceptions of mutual similarity (Gehlbach et al., 2012). The latter intervention involved having students and teachers complete a "get-to-know-you questionnaire" at the beginning of the school year, and then providing teachers with information about preferences or attitudes they shared with half of the students in their classes (randomly selected; the other half of the students were the control group). Teachers reported improved relationships with those students for whom they had received similarity information five weeks later.

Optimizing test-taking and feedback

The design of tests and assessments, the feedback provided after testing, and the consequences of individuals' performance influence students' emotions (Pekrun, 2006). If the structure, demands, and consequences of assessments are not clear to students, test anxiety is increased due to the possibility of failure. Further, tests that demand high attention can impair performance in students who are anxious about tests, because part of their cognitive resources are already consumed by worrying about failure. Anxiety can be further increased during such tests. This may be true, for example, when creative essay writing has to be carried out in a limited period of time.

Feedback about achievement on academic tests is one of the most powerful factors in the development of achievement emotions (Goetz et al., 2018). Repeated feedback on success can strengthen students' self-confidence and cultivate self-efficacy over time and increase their positive achievement emotions, such as hope for success and pride. In contrast, repeated feedback on failure undermines self-confidence and increases negative achievement emotions, such as anxiety, shame, and hopelessness. This is especially true if feedback on failure is coupled with the message that failure is due to lack of ability as opposed to students' lack of investing time and effort. Regarding feedback on achievement, teachers can strengthen students' self-confidence and positive emotions by using the following four principles (Pekrun et al., 2017b). (a) Whenever possible, mastery standards should be used and normative standards should be avoided for evaluating achievement. (b) Repeated feedback on success rather than on failure should be used by emphasizing improvement of performance, even if improvement is small. (c) Modeling adaptive approaches to dealing with failure has a vast instructional value. It is important to make clear to students that errors and mistakes should be regarded as opportunities to learn, and that abilities are malleable (e.g., by assigning the grade "not yet" instead of "fail"; Dweck, 2010). (d) Beyond evaluative feedback about success and failure, it is critical to provide formative feedback about how students can improve their competencies and attain mastery. Detailed formative feedback, combined with positive expectancies of mastery, will strengthen students' confidence in their abilities, and will support positive affective outcomes resulting from such confidence.

Supporting and fostering adaptive emotion regulation

Even if educators put forth their very best efforts to engage themselves and their learners and create emotionally sound environments for learning and teaching, negative emotions are bound to be experienced. These are inevitable because of heterogeneity in the classroom with respect to intelligence, prior knowledge and interest, as well as with respect to emotional dispositions of teachers and students alike (e.g., habitual enjoyment, test anxiety). As such, students and teachers need to have a range of strategies for regulating learning- and teaching-related emotions. There are numerous models and approaches describing the emotion regulation of students and teachers as well as effective coping strategies that enable individuals to downregulate their negative emotions (e.g., Burić et al., 2016; Donker et al., 2020; Harley et al., 2019; Jarrell and Lajoie, 2017; Lee et al., 2016).

Emotions can be regulated by directly changing their symptoms (emotion-oriented regulation) or by changing their antecedents (competence-, situation-, and appraisal-oriented regulation; Pekrun et al., 2017b). Successful emotion regulation presupposes abilities to recognize one's emotions and to select appropriate ways of managing them. These abilities are part of an individuals' emotional competencies, which consist of abilities to recognize, make use of, and regulate one's own emotions and the emotions of other persons (e.g., Goetz and Bieg, 2016). There are several ways in which emotion regulation can be trained in the classroom. For example, teachers can model emotion regulation by first labeling, identifying, and reframing their own emotions. SEL programs can be used to help students practice these skills. An example of a very popular SEL program, which focuses on fostering emotional intelligence, is the PATHS program (Promoting Alternative Thinking Strategies; Greenberg et al., 1995; for a description of the program, see also Brackett and Rivers, 2014; Parker et al., 2009). A model for the Promotion of Emotional Intelligence in Learning and Achievement Situations (the PEILAS Model) is outlined by Goetz and Bieg (2016).

When it comes to the regulation of specific emotions, the most prominent emotion is test anxiety, reflecting the long tradition of test-anxiety research (Zeidner, 1998). Programs focused on dealing with test anxiety target specific symptoms of anxiety (e.g., relaxation training), lack of self-confidence causing anxiety (e.g., cognitive therapy), or underlying competence deficits (behavioral skills training; e.g., training to improve learning strategies). Combinations of these approaches have been shown to be most effective for many students. In sum, theoretical and empirical findings suggest that there are numerous ways to design emotionally healthy learning environments for teachers and students alike. It is important to create instructional situations and contexts in such a way that students, who typically spend a large and important part of their lives in them, would experience a preponderance of positive emotions. After all, the joy of learning will undoubtedly contribute to positive academic outcomes and students' general sense of well-being. Hence, we implore researchers to continue on their path toward better understanding of links and contingencies between emotions and other meaningful variables.

Conclusion

In this chapter, we introduced emotions as multidimensional constructs with affective, cognitive, motivational, physiological, and expressive components. Based on traditional appraisal-theoretical approaches, it can be assumed that emotions arise from individuals' subjective evaluation of activities and outcomes. In the context of learning and performance, control and value appraisals are of particular importance. This evaluation is influenced by personal beliefs, but also by external circumstances. Thus, teachers have the opportunity to shape the emotional experience of students through targeted designs of learning environments and learning tasks. Emotions impact a range of cognitive, self-regulatory, and motivational processes, and, subsequently, student performance and learning. Moreover, they are important components of individuals' subjective well-being. Therefore, promoting positive emotions and reducing negative emotions in school and out-of-school learning should also be a meaningful goal and educational outcome. However, not only the emotions of students, but also those of teachers are critical to consider. After all, it's the teachers' emotions that may affect the quality of instructional processes and contribute to teachers' subjective well-being. Ultimately, it is rewarding for teachers to work with students who are filled with joy of learning and interest in the learning content, and schooling experience is more favorable for students who interact with enthusiastic and warm teachers.

References

- Ahmed, W., van der Werf, G., Minnaert, A., 2010. Emotional experiences of students in the classroom: a multimethod qualitative study. *Eur. Psychol.* 15 (2), 142–151. <https://doi.org/10.1027/1016-9040/a000014>.
- Ahmed, W., van der Werf, G., Kuyper, H., Minnaert, A., 2013. Emotions, self-regulated learning, and achievement in mathematics: a growth curve analysis. *J. Educ. Psychol.* 105 (1), 150–161. <https://doi.org/10.1037/a0030160>.
- Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., Trautwein, U., 2018. Student misbehavior and teacher well-being: testing the mediating role of the teacher-student relationship. *Learn. Instruct.* 58, 126–136. <https://doi.org/10.1016/j.learninstruc.2018.05.006>.
- Arens, A.K., Morin, A.J., 2016. Relations between teachers' emotional exhaustion and students' educational outcomes. *J. Educ. Psychol.* 108 (6), 800–813. <https://doi.org/10.1037/edu0000105>.
- Babad, E., 2007. Teachers' nonverbal behaviors and its effects on students. In: Perry, R., Smart, J.C. (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*. Springer, Dordrecht, pp. 201–261.
- Banerjee, N., Stearns, E., Moller, S., Mickelson, R.A., 2017. Teacher job satisfaction and student achievement: the roles of teacher professional community and teacher collaboration in schools. *Am. J. Educ.* 123 (2), 203–241. <https://doi.org/10.1086/689932>.
- Barrett, L., Lewis, M., Haviland-Jones, J.M. (Eds.), 2016. *Handbook of Emotions*, fourth ed. The Guilford Press, New York.

- Barroso, C., Ganley, C.M., McGraw, A.L., Geer, E.A., Hart, S.A., Daucourt, M.C., 2021. A meta-analysis of the relation between math anxiety and math achievement. *Psychol. Bull.* 147 (2), 134–168. <https://doi.org/10.1037/bul0000307>.
- Baudoin, N., Galand, B., 2017. Effects of classroom goal structures on student emotions at school. *Int. J. Educ. Res.* 86, 13–22. <https://doi.org/10.1016/j.ijer.2017.08.010>.
- Bellock, S.L., Gunderson, E.A., Ramirez, G., Levine, S.L., 2010. Female teachers' math anxiety affects girls' math achievement. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1860–1863. <https://doi.org/10.1073/pnas.0910967107>.
- Bieleke, M., Gogol, K., Goetz, T., Daniels, L., Pekrun, R., 2021. The AEQ-S: a short version of the achievement emotions questionnaire. *Contemp. Educ. Psychol.* 65, 101940. <https://doi.org/10.1016/j.cedpsych.2020.101940>.
- Brackett, M.A., Rivers, S.E., 2014. Transforming students' lives with social and emotional learning. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York and London, pp. 368–388.
- Brown, C.H., 1938. Emotional reactions before examinations: II. Results of a questionnaire. *J. Psychol.* 5, 11–26. <https://doi.org/10.1080/00223980.1938.9917549>.
- Brummelman, E. (Ed.), 2020. *Psychological Perspectives on Praise*. Routledge, Abingdon, UK.
- Burić, I., Sorić, I., Penezić, Z., 2016. Emotion regulation in academic domain: development and validation of the academic emotion regulation questionnaire (AERQ). *Pers. Individ. Differ.* 96, 138–147. <https://doi.org/10.1016/j.paid.2016.02.074>.
- Butler, R., 2007. Teachers' achievement goal orientations and associations with teachers' help seeking: examination of a novel approach to teacher motivation. *J. Educ. Psychol.* 99, 241–252. <https://doi.org/10.1037/0022-0663.99.2.241>.
- Camacho-Morles, J., Slemp, G.R., Pekrun, R., Loderer, K., Hou, H., Oades, L.G., 2021. Activity achievement emotions and academic performance: a meta-analysis. *Educ. Psychol. Rev.* 33, 1051–1095. <https://doi.org/10.1007/s10648-020-09585-3>.
- Chang, M.-L., 2009. An appraisal perspective of teacher burnout: examining the emotional work of teachers. *Educ. Psychol. Rev.* 21, 193–218. <https://doi.org/10.1007/s10648-009-9106-y>.
- Chang, M.-L., 2013. Toward a theoretical model to understand teacher emotions and teacher burnout in the context of student misbehavior: appraisal, regulation and coping. *Motiv. Emot.* 37, 799–817. <https://doi.org/10.1007/s11031-012-9335-0>.
- Corcoran, R.P., Cheung, A.C.K., Kim, E., Xie, C., 2018. Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* 25, 56–72. <https://doi.org/10.1016/j.edurev.2017.12.001>.
- D'Eon, M., Overgaard, V., Harding, S.R., 2000. Teaching as a social practice: implications for faculty development. *Adv. Health Sci. Educ.* 5 (2), 151–162. <https://doi.org/10.1023/A:1009898031033>.
- Dikter, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., Poeppel, D., 2017. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Curr. Biol.* 27 (9), 1375–1380. <https://doi.org/10.1016/j.cub.2017.04.002>.
- D'Mello, S., Lehman, B., Pekrun, R., Graesser, A., 2014. Confusion can be beneficial for learning. *Learn. Instr.* 29, 153–170. <https://doi.org/10.1016/j.learninstruc.2012.05.003>.
- D'Mello, S., Dieterle, E., Duckworth, A., 2017. Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educ. Psychol.* 52, 104–123. <https://doi.org/10.1080/00461520.2017.1281747>.
- Donker, M.H., van Gog, T., Mainhard, M.T., 2018. A quantitative exploration of two teachers with contrasting emotions: intra-individual process analyses of physiology and interpersonal behavior. *Frontline Learn. Res.* 6 (3), 162–184. <https://doi.org/10.14786/flr.v6i3.372>.
- Donker, M.H., Erisman, M.C., van Gog, T., Mainhard, T., 2020. Teachers' emotional exhaustion: associations with their typical use of and implicit attitudes toward emotion regulation strategies. *Front. Psychol.* 11, 126. <https://doi.org/10.3389/fpsyg.2020.00867>.
- Driscoll, K.C., Wang, L., Mashburn, A.J., Pianta, R.C., 2011. Fostering supportive teacher–child relationships: intervention implementation in a state-funded preschool program. *Early Educ. Dev.* 22 (4), 593619. <https://doi.org/10.1080/10409289.2010.502015>.
- Dweck, C.S., 2010. Even geniuses work hard. *Educ. Leader* 68 (1), 16–20.
- Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>.
- Ekman, P., Friesen, W.V., Hager, J.C., 2002. *Facial Action Coding System*. Research Nexus [E-book].
- Elfenbein, H.A., 2014. The many faces of emotional contagion: an affective process theory of affective linkage. *Organ. Psychol. Rev.* 4 (4), 326–362. <https://doi.org/10.1177/2041386614542889>.
- Flunger, B., Mayer, A., Umbach, N., 2019. Beneficial for some or for everyone? Exploring the effects of an autonomy-supportive intervention in the real-life classroom. *J. Educ. Psychol.* 111 (2), 210–234. <https://doi.org/10.1037/edu0000284>.
- Forsblom, L., Pekrun, R., Loderer, K., Peixoto, F., 2022. Cognitive appraisals, achievement emotions, and students' math achievement: a longitudinal analysis. *J. Educ. Psychol.* 114 (2), 346–367. <https://doi.org/10.1037/edu0000671>.
- Frenzel, A.C., Thrash, T.M., Pekrun, R., Goetz, T., 2007. Achievement emotions in Germany and China: a cross-cultural validation of the Academic Emotions Questionnaire–Mathematics (AEQ-M). *J. Cross Cult. Psychol.* 38 (3), 302–309. <https://doi.org/10.1177/0022022107300276>.
- Frenzel, A.C., Goetz, T., Lüdtke, O., Pekrun, R., Sutton, R.E., 2009. Emotional transmission in the classroom: exploring the relationship between teacher and student enjoyment. *J. Educ. Psychol.* 101, 705–716. <https://doi.org/10.1037/a0014695>.
- Frenzel, A.C., Pekrun, R., Goetz, T., Daniels, L.M., Durksen, T.L., Becker-Kurz, B., Klassen, R., 2016. Measuring enjoyment, anger, and anxiety during teaching: the teacher emotions scales (TES). *Contemp. Educ. Psychol.* 46, 148–163. <https://doi.org/10.1016/j.cedpsych.2016.05.003>.
- Frenzel, A.C., Becker-Kurz, B., Pekrun, R., Goetz, T., Lüdtke, O., 2018. Emotion transmission in the classroom revisited: a reciprocal effects model of teacher and student enjoyment. *J. Educ. Psychol.* 110 (5), 628–639. <https://doi.org/10.1037/edu0000228>.
- Frenzel, A.C., Marx, A.K.G., Pekrun, R., Reck, C., Müller, M., 2019. Teachers' and Learners' Emotional Experiences in Class: A Field-Based Video Study. *American Educational Research Association Annual Conference*, Toronto, Canada.
- Frenzel, A.C., Fiedler, D., Marx, A.K.G., Reck, C., Pekrun, R., 2020. Who enjoys teaching, and when? Between-and within-person evidence on teachers' appraisal-emotion links. *Front. Psychol.* 11, 1092. <https://doi.org/10.3389/fpsyg.2020.01092>.
- Frenzel, A.C., Daniels, L., Burić, I., 2021. Teacher emotions in the classroom and their implications for students. *Educ. Psychol.* 56 (4), 1–15. <https://doi.org/10.1080/00461520.2021.1985501>.
- Frenzel, A.C., 2014. Teacher emotions. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York and London, pp. 494–519.
- Fried, L., Mansfield, C., Dobozy, E., 2015. Teacher emotion research: introducing a conceptual model to guide future research. *Issues Educ. Res.* 25 (4), 415.
- Gehlbach, H., Brinkworth, M.E., Harris, A.D., 2012. Changes in teacher-student relationships. *Br. J. Educ. Psychol.* 82 (4), 690–704. <https://doi.org/10.1111/j.2044-8279.2011.02058.x>.
- Goetz, T., Bieg, M., 2016. Academic emotions and their regulation via emotional intelligence. In: Lipnevich, A.A., Preckel, F., Roberts, R.D. (Eds.), *Psychosocial Skills and School Systems in the 21st Century: Theory, Research and Practice*. Springer, Switzerland, pp. 279–298.
- Goetz, T., Hall, N.C., 2020. Emotion and achievement in the classroom. In: Hattie, J., Anderman, E.M. (Eds.), *International Guide to Student Achievement*. Routledge, New York and London, pp. 192–196.
- Goetz, T., Frenzel, A.C., Stoeger, H., Hall, N.C., 2010. Antecedents of everyday positive emotions: an experience sampling analysis. *Motiv. Emot.* 34, 49–62. <https://doi.org/10.1007/s11031-009-9152-2>.
- Goetz, T., Bieg, M., Hall, N.C., 2016a. Assessing academic emotions via the experience sampling method. In: Zembylas, M., Schutz, P.A. (Eds.), *Methodological Advances in Research on Emotion in Education*. Springer, Switzerland, pp. 245–258.

- Goetz, T., Sticca, F., Pekrun, R., Murayama, K., Elliot, A.J., 2016b. Intraindividual relations between achievement goals and discrete achievement emotions: an experience sampling approach. *Learn. Instruct.* 41, 115–125. <https://doi.org/10.1016/j.learninstruc.2015.10.007>.
- Goetz, T., Lipnevich, A.A., Krannich, M., Gogol, K., 2018. Performance feedback and emotions. In: Lipnevich, A.A., Smith, J.K. (Eds.), *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press, pp. 554–574.
- Goetz, T., Hall, N.C., Krannich, M., 2019. Boredom. In: Renninger, K.A., Hidi, S.E. (Eds.), *The Cambridge Handbook on Motivation and Learning*. Cambridge University Press, pp. 465–486.
- Goetz, T., Keller, M.M., Lütke, O., Nett, U.E., Lipnevich, A.A., 2020. The dynamics of real-time classroom emotions: appraisals mediate the relation between students' perceptions of teaching and their emotions. *J. Educ. Psychol.* 112 (6), 1243–1260. <https://doi.org/10.1037/edu0000415>.
- Goetz, T., Bieleke, M., Gogol, K., van Tartwijk, J., Mainhard, T., Lipnevich, A.A., Pekrun, R., 2021. Getting along and feeling good: reciprocal associations between student-teacher relationship quality and students' emotions. *Learn. Instruct.* 71, 101349. <https://doi.org/10.1016/j.learninstruc.2020.101349>.
- Goldstein, L.S., Lake, V.E., 2000. "Love, love, and more love for children": exploring preservice teachers' understanding of caring. *Teach. Teach. Educ.* 16, 861–872. [https://doi.org/10.1016/S0742-051X\(00\)00031-7](https://doi.org/10.1016/S0742-051X(00)00031-7).
- González, A., Paoloni, V., Donolo, D., Rinaudo, C., 2012. Motivational and emotional profiles in university undergraduates: a self-determination theory perspective. *Spanish J. Psychol.* 15 (3), 1069–1080. https://doi.org/10.5209/rev_SJOP.2012.v15.n3.39397.
- Graham, S., Taylor, A.Z., 2014. An attributional approach to emotional life in the classroom. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York and London, pp. 96–119.
- Greenberg, M.T., Kusche, C.A., Cook, E.T., Quamma, J.P., 1995. Promoting emotional competence in school-aged children: the effects of the PATHS curriculum. *Dev. Psychopathol.* 7, 117–136. <https://doi.org/10.1017/S0954579400006374>.
- Hargreaves, A., 1998. The emotional practice of teaching. *Teach. Teach. Educ.* 14, 835–854. [https://doi.org/10.1016/S0742-051X\(98\)00025-0](https://doi.org/10.1016/S0742-051X(98)00025-0).
- Harley, J.M., Pekrun, R., Taxer, J.L., Gross, J.J., 2019. Emotion regulation in achievement situations: an integrated model. *Educ. Psychol.* 54, 106–126.
- Harley, J.M., 2016. Measuring emotions: a survey of cutting edge methodologies used in computer-based learning environment research. In: Tettegah, S.Y., Gartmeier, M. (Eds.), *Emotions, Technology, Design, and Learning*. Elsevier Academic Press, pp. 87–114.
- Hatfield, E., Cacioppo, J.T., Rapson, R.L., 1994. *Emotional Contagion*. Cambridge University Press.
- Hembree, R., 1988. Correlates, causes, effects, and treatment of test anxiety. *Rev. Educ. Res.* 58, 47–77. <https://doi.org/10.3102/00346543058001047>.
- Huang, C., 2011. Achievement goals and achievement emotions: a meta-analysis. *Educ. Psychol. Rev.* 23, 359–388. <https://doi.org/10.1007/s10648-011-9155-x>.
- Immordino-Yang, M.H., 2010. Toward a microdevelopmental, interdisciplinary approach to social emotion. *Emot. Rev.* 2 (3), 217–220. <https://doi.org/10.1177/1754073910361985>.
- Isen, A.M., Reeve, J., 2005. The influence of positive affect on intrinsic and extrinsic motivation: facilitating enjoyment of play, responsible work behavior, and self-control. *Motiv. Emot.* 29 (4), 295–323. <https://doi.org/10.1007/s11031-006-9019-8>.
- Jacob, B., Frenzel, A.C., Stephens, E.J., 2017. Good teaching feels good—but what is "good teaching"? Exploring teachers' definitions of teaching success in mathematics. *ZDM Math. Educ.* 49 (3), 461–473. <https://doi.org/10.1007/s11858-017-0848-6>.
- Jarrell, A., Lajoie, S.P., 2017. The regulation of achievements emotions: implications for research and practice. *Can. Psychol.* 58 (3), 276–287. <https://doi.org/10.1037/cap0000119>.
- Keller, M.M., Woolfolk Hoy, A., Goetz, T., Frenzel, A.C., 2016. Teacher enthusiasm: reviewing and redefining a complex construct. *Educ. Psychol. Rev.* 28 (4), 743–769. <https://doi.org/10.1007/s10648-015-9354-y>.
- Klassen, R.M., Perry, N.E., Frenzel, A.C., 2012. Teachers' relatedness with students: an underemphasized component of teachers' basic psychological needs. *J. Educ. Psychol.* 104, 150–165. <https://doi.org/10.1037/a0026253>.
- Klusmann, U., Richter, D., Lütke, O., 2016. Teachers' emotional exhaustion is negatively related to students' achievement: evidence from a large-scale assessment study. *J. Educ. Psychol.* 108 (8), 1193–1203. <https://doi.org/10.1037/edu0000125>.
- Larson, R., Csikszentmihalyi, M., 2014. The experience sampling method. In: Csikszentmihalyi, M. (Ed.), *Flow and the Foundations of Positive Psychology*. Springer, Dordrecht, pp. 21–34. https://doi.org/10.1007/978-94-017-9088-8_2.
- Lazarides, R., Buchholz, J., 2019. Student-perceived teaching quality: how is it related to different achievement emotions in mathematics classrooms? *Learn. Instruct.* 61, 45–59. <https://doi.org/10.1016/j.learninstruc.2019.01.001>.
- Lazarus, R.S., Folkman, S., 1984. *Stress, Appraisal, and Coping*. Springer, New York.
- Lee, M., Pekrun, R., Taxer, J.L., Schutz, P.A., Vogl, E., Xie, X., 2016. Teachers' emotions and emotion management: integrating emotion regulation theory with emotional labor research. *Soc. Psychol. Educ.* 19 (4), 843–863. <https://doi.org/10.1007/s11218-016-9359-5>.
- Linnenbrink-Garcia, L., Patail, E.A., Pekrun, R., 2016. Adaptive motivation and emotion in education: research and principles for instructional design. *Policy Insights Behav. Brain Sci.* 3 (2), 228–236. <https://doi.org/10.1177/2372732216644450>.
- Lipnevich, A.A., Smith, J.K., 2009. The effects of feedback on student examination performance. *J. Exp. Psychol. Appl.* 15, 319–333. <https://doi.org/10.1007/s11092-009-9082-2>.
- Lipnevich, A.A., Smith, J.K., 2018. *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press.
- Lipnevich, A.A., Murano, D., Krannich, M., Goetz, T., 2021. Should I grade or should I comment: links among feedback, emotions, and performance. *Learn. Individ. Differ.* 89, 102020. <https://doi.org/10.1016/j.lindif.2021.102020>.
- Loderer, K., Pekrun, R., Lester, J.C., 2020. Beyond cold technology: a systematic review and meta-analysis on emotions in technology-based learning environments. *Learn. Instruct.* 70, 101162. <https://doi.org/10.1016/j.learninstruc.2018.08.002>.
- Mandler, G., Sarason, S.B., 1952. A study of anxiety and learning. *J. Abnorm. Soc. Psychol.* 47, 166–173. <https://doi.org/10.1037/h0062855>.
- Mattarella-Micke, A., Mateo, J., Kozak, M.N., Foster, K., Beilock, S.L., 2011. Choke or thrive? The relation between salivary cortisol and math performance depends on individual differences in working memory and math-anxiety. *Emotion* 11 (4), 1000–1005. <https://doi.org/10.1037/a0023224>.
- Moors, A., Ellsworth, P., Scherer, K.R., Frijda, N., 2013. Appraisal theories of emotion: state of the art and future development. *Emot. Rev.* 5, 119–124. <https://doi.org/10.1177/1754073912468165>.
- Moskowitz, S., Dewaele, J.-M., 2019. Is teacher happiness contagious? A study of the link between perceptions of language teacher happiness and student attitudes. *Innovat. Lang. Learn. Teach.* 1–14. <https://doi.org/10.1080/17501229.2019.1707205>.
- Muis, R.R., Chevrier, M., Singh, C., 2018. The role of epistemic emotions in personal epistemology and self-regulated learning. *Educ. Psychol.* 53, 145–165. <https://doi.org/10.1080/00461520.2017.1421465>.
- Murano, D., Sawyer, J., Lipnevich, A.A., 2020. A meta-analytic review of preschool social and emotional learning interventions. *Rev. Educ. Res.* 90 (2), 227–263. <https://doi.org/10.3102/0034654320914743>.
- Nias, J., 1996. Thinking about feeling: the emotions in teaching. *Camb. J. Educ.* 26, 293–306. <https://doi.org/10.1080/0305764960260301>.
- Nurmi, J.E., Kiuru, N., 2015. Students' evocative impact on teacher instruction and teacher-child relationships: theoretical background and an overview of previous research. *Int. J. Behav. Dev.* 39, 445–457. <https://doi.org/10.1177/0165025415592514>.
- Obergriesser, S., Stoeger, H., 2020. Students' emotions of enjoyment and boredom and their use of cognitive learning strategies—how do they affect one another? *Learn. Instruct.* 66, 101285. <https://doi.org/10.1016/j.learninstruc.2019.101285>.
- Parker, J.D.A., Saklofske, D.H., Wood, L.M., Collin, T., 2009. The role of emotional intelligence in education. In: Stough, C., Saklofske, D.H., Parker, J.D.A. (Eds.), *Assessing Emotional Intelligence. Theory, Research and Applications*. Springer, New York, pp. 239–255.

- Parkinson, B., 2011. Interpersonal emotion transfer: contagion and social appraisal. *Soc. Personal. Psychol. Compass* 5 (7), 428–439. <https://doi.org/10.1111/j.1751-9004.2011.00365.x>.
- Pekrun, R., Bühner, M., 2014. Self-report measures of academic emotions. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York and London, pp. 561–579.
- Pekrun, R., Linnenbrink-Garcia, L. (Eds.), 2014. *International Handbook of Emotions in Education*. Routledge, New York and London.
- Pekrun, R., Linnenbrink-Garcia, L., 2022. Academic emotions and student engagement. In: Reschly, A.L., Christenson, S.L. (Eds.), *The Handbook of Research on Student Engagement*, second ed. Springer, New York.
- Pekrun, R., Perry, R.P., 2014. Control-value theory of achievement emotions. In: Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York and London, pp. 120–141.
- Pekrun, R., Stephens, E.J., 2012. Academic emotions. In: Harris, K.R., Graham, S., Urdan, T., Royer, J.M., Zeidner, M. (Eds.), *APA Educ. Psychol. Handb.*, vol. 2. American Psychological Association, Washington, DC, pp. 3–31.
- Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., Perry, R.P., 2011. Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemp. Educ. Psychol.* 36, 36–48. <https://doi.org/10.1016/j.cedpsych.2010.10.002>.
- Pekrun, R., Lichtenfeld, S., Marsh, H.W., Murayama, K., Goetz, T., 2017a. Achievement emotions and academic performance: longitudinal models of reciprocal effects. *Child Dev.* 88, 1653–1670. <https://doi.org/10.1111/cdev.12704>.
- Pekrun, R., Muis, K.R., Frenzel, A.C., Goetz, T., 2017b. *Emotions at School*. Routledge, New York and London.
- Pekrun, R., Murayama, K., Marsh, H.W., Goetz, T., Frenzel, A.C., 2019. Happy fish in little ponds: testing a reference group model of achievement and emotion. *J. Pers. Soc. Psychol.* 117, 166–185. <https://doi.org/10.1037/pspp0000230>.
- Pekrun, R., 1993. Facets of students' academic motivation: a longitudinal expectancy-value approach. In: Maehr, M., Pintrich, P. (Eds.), *Adv. Motiv. Achiev.*, vol. 8. JAI Press, Greenwich, CT, pp. 139–189.
- Pekrun, R., 2006. The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. <https://doi.org/10.1007/s10648-006-9029-9>.
- Pekrun, R., 2018. Control-value theory: a social-cognitive approach to achievement emotions. In: Liem, G.A.D., McInerney, D.M. (Eds.), *Big Theories Revisited 2: A Volume of Research on Sociocultural Influences on Motivation and Learning*. Information Age Publishing, Charlotte, NC, pp. 162–190.
- Pekrun, R., 2020. Self-report is indispensable to assess students' learning. *Frontline Learn. Res.* 8, 185–193.
- Pekrun, R., 2021a. Self-appraisals and emotions: a control-value approach. In: Dicke, T., Guay, F., Marsh, H.W., Craven, R.G., McInerney, D.M. (Eds.), *Self—A Multidisciplinary Concept*. Information Age Publishing, Charlotte, NC, pp. 1–30.
- Pekrun, R., 2021b. Teachers need more than knowledge: why motivation, emotion, and self-regulation are indispensable. *Educ. Psychol.* 56 (4), 312–322. <https://doi.org/10.1080/00461520.2021.1991356>.
- Pianta, R., Hamre, B.K., 2001. Students, Teachers, and Relationship Support (STARS). Psychological Assessment Resources.
- Poon, C.Y.-S., Hui, V.K.-Y., Yuen, G.W.-C., Kwong, V.W.-Y., Chan, C.S., 2019. A well-slept teacher is a better teacher: a multi-respondent experience-sampling study on sleep, stress, and emotional transmission in the classroom. *Psych J.* 8 (3), 280–292. <https://doi.org/10.1002/pchj.282>.
- Posner, J., Russell, J.A., Peterson, B.S., 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17 (3), 715–734. <https://doi.org/10.1017/S0954579405050340>.
- Putwain, D.W., Becker, S., Symes, W., Pekrun, R., 2018a. Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learn. Instruct.* 54, 73–81. <https://doi.org/10.1016/j.learninstruc.2017.08.004>.
- Putwain, D.W., Pekrun, R., Nicholson, L.J., Symes, W., Becker, S., Marsh, H.W., 2018b. Control-value appraisals, enjoyment, and boredom in mathematics: a longitudinal latent interaction analysis. *Am. Educ. Res. J.* 55, 1339–1368. <https://doi.org/10.3102/0002831218786689>.
- Putwain, D., Wood, P., Pekrun, R., 2022. Achievement emotions and academic achievement: reciprocal relations and the moderating influence of academic buoyancy. *J. Educ. Psychol.* 114 (1), 108–126. <https://doi.org/10.1037/edu0000637>.
- Quin, D., 2017. Longitudinal and contextual associations between teacher-student relationships and student engagement: a systematic review. *Rev. Educ. Res.* 87, 345–387. <https://doi.org/10.3102/0034654316669434>.
- Reeve, J., 2009. Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educ. Psychol.* 44, 159–175. <https://doi.org/10.1080/00461520903028990>.
- Richardson, M., Abraham, C., Bond, R., 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138 (2), 353–387. <https://doi.org/10.1037/a0026838>.
- Roorda, D., Koomen, H., Spilt, J., Oort, F., 2011. The influence of affective teacher-student relationships on students' school engagement and achievement: a meta-analytic approach. *Rev. Educ. Res.* 81 (4), 493–529. <https://doi.org/10.3102/0034654311421793>.
- Roos, A.-L., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., Pekrun, R., 2021. Test anxiety and physiological arousal: a systematic review and meta-analysis. *Educ. Psychol. Rev.* 33, 579–618. <https://doi.org/10.1007/s10648-020-09543-z>.
- Scherer, K.R., Moors, A., 2019. The emotion process: event appraisal and component differentiation. *Annu. Rev. Psychol.* 70, 719–745. <https://doi.org/10.1146/annurev-psych-122216-011854>.
- Scherer, K.R., 2000. Emotions as episodes of subsystems synchronization driven by nonlinear appraisal processes. In: Lewis, M.D., Granic, I. (Eds.), *Emotion, Development, and Self-Organization*. Cambridge University Press, pp. 70–99.
- Schutz, P.A., DeCuir-Gunby, J.T., Williams-Johnson, M.R., 2016. Using multiple and mixed methods to investigate emotions in educational contexts. In: Zembylas, M., Schutz, P.A. (Eds.), *Methodological Advances in Research on Emotion and Education*. Springer, Switzerland, pp. 217–229.
- Shao, K., Pekrun, R., Marsh, H.W., Loderer, K., 2020. Control-value appraisals, achievement emotions, and foreign language performance: a latent interaction analysis. *Learn. Instruct.* 69, 101356. <https://doi.org/10.1016/j.learninstruc.2020.101356>.
- Spielberger, C.D., Anton, W.D., Bedell, J., 1976. The nature and treatment of test anxiety. In: Zuckerman, M., Spielberger, C.D. (Eds.), *Emotions and Anxiety: New Concepts, Methods, and Applications*. Erlbaum, Hillsdale, NJ, pp. 317–344.
- Spilt, J.L., Koomen, H.M.Y., Thijs, J.T., 2011. Teacher wellbeing: the importance of teacher-student relationships. *Educ. Psychol. Rev.* 23 (4), 457–477. <https://doi.org/10.1007/s10648-011-9170-y>.
- Sung, Y.-T., Chao, T.-Y., Tseng, F.-L., 2016. Reexamining the relationship between test anxiety and learning achievement: an individual-differences perspective. *Contemp. Educ. Psychol.* 46, 241–252. <https://doi.org/10.1016/j.cedpsych.2016.07.001>.
- Sutter-Brandenberger, C.C., Hagenauer, G., Hascher, T., 2018. Students' self-determined motivation and negative emotions in mathematics in lower secondary education—investigating reciprocal relations. *Contemp. Educ. Psychol.* 55, 166–175. <https://doi.org/10.1016/j.tate.2018.05.008>.
- Tsoulopas, C.N., Carson, R.L., Matthews, R., Grawitch, M.J., Barber, L.K., 2010. Exploring the association between teachers' perceived student misbehaviour and emotional exhaustion: the importance of teacher efficacy beliefs and emotion regulation. *Educ. Psychol.* 30 (2), 173–189. <https://doi.org/10.1080/01443410903494460>.
- Turner, J.E., Schallert, D.L., 2001. Expectancy-value relationships of shame reactions and shame resiliency. *J. Educ. Psychol.* 93, 320–329. <https://doi.org/10.1037/0022-0663.93.2.320>.
- Tze, V.M.C., Daniels, L.M., Klassen, R.M., 2016. Evaluating the relationship between boredom and academic outcomes: a meta-analysis. *Educ. Psychol. Rev.* 28 (1), 119–144. <https://doi.org/10.1007/s10648-015-9301-y>.
- Weiner, B., 1985. An attributional theory of achievement motivation and emotion. *Psychol. Rev.* 92, 548–573.

- Weiner, B., 2007. Examining emotional diversity in the classroom: an attribution theorist considers the moral emotions. In: Schutz, P.A., Pekrun, R. (Eds.), *Emotion in Education*. Elsevier Academic Press, pp. 75–88.
- Weissberg, R.P., Durlak, J.A., Domitrovich, C.E., Gullotta, T.P., 2015. Social and emotional learning: past, present, and future. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford Press, New York, pp. 3–19.
- Wigfield, A., Cambria, J.M., 2010. Expectancy-value theory: retrospective and prospective. In: Urdan, T.C., Karabenick, S.A. (Eds.), *The Decade Ahead: Theoretical Perspectives on Motivation and Achievement*, vol. 16. Emerald Group, UK, pp. 35–70.
- Zeidner, M., 1998. *Test Anxiety: The State of the Art*. Plenum Press, New York.
- Zhang, J., Zhao, N., Kong, Q.P., 2019. The relationship between math anxiety and math performance: a meta-analytic investigation. *Front. Psychol.* 10. <https://doi.org/10.3389/fpsyg.2019.01613>.

Epistemic cognition

William A. Sandoval, School of Education & Information Studies, University of California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	162
Strands of research related to epistemic cognition	162
Issues in epistemic cognition research	164
Emerging trends in epistemic cognition research	165
References	166
Further reading	167

Glossary

Epistemic Of or relating to knowledge or the conditions for acquiring it

Epistemology A branch of philosophy that investigates the origin, nature, methods, and limits of human knowledge

Introduction

Epistemic cognition refers to how people think about what they know and how they know it, including ideas about the nature of knowledge and reliable methods for constructing and evaluating knowledge claims. The study of knowledge itself, epistemology, has been concerned with defining knowledge, distinguishing it from opinion and other sorts of beliefs, and characterizing processes of knowledge construction for millennia. An educational interest in people becoming competent judges of knowledge and knowledge-making processes similarly can be traced back at least to Plato. The systematic study of how individuals develop ideas about knowledge and knowing, and how those ideas influence learning and reasoning, extends only to approximately the last 50 years, from Perry's (1970) seminal study of Harvard students' development of epistemic beliefs. This created a specialization in educational psychology known as the study of personal epistemology, but there are related strands of developmental and cognitive research with clear relations to how people think about what and how they know. Consequently, epistemic cognition is best thought of as a topic of interest to scholars in many fields, rather than as a field of its own.

Strands of research related to epistemic cognition

Several sub-fields of cognitive developmental research and cognitive, educational, and social psychology have pursued various lines of research that can all be viewed as focused on or related to epistemic cognition.

The field of personal epistemology originated from Perry's (1970) study of male Harvard undergraduates and their "intellectual development" through college. Perry described a developmental stage model, with nine sequential positions grouped into four main stages, each of which included predictions about how college students would interpret all of their experiences, regardless of domain or discipline. People in the first stage of Perry's model were termed dualists, and they were posited to have an objective view of the world, believing that knowledge is factual, unchanging, and the property of experts who pass it along to others. Such views meant that dualists would not engage in critical thinking about knowledge or its sources. Through cognitive disequilibrium resulting from experiences in education and elsewhere, Perry argued that people eventually move into a multipist view, believing there are no objective standards for knowledge, making all views equally valid, a sort of radical subjectivity. Some people, if they continue to encounter experiences fostering cognitive disequilibrium, were thought to move into the relativist category, acknowledging the contingent and contextual nature of knowledge and the necessity of using one's own logic and reasoning (i.e., critical thinking) to sort knowledge propositions from other types of claims. Perry's fourth stage, commitment within relativism, largely concerns how individuals adopt particular values and identities that shape their reasoning.

Perry's work initiated a line of scholarship aimed at modeling individual's beliefs about knowledge, their variation, and their development. These models can be grouped into two classes. *Developmental models* characterize differences in epistemic cognition in terms of phases or stages, positing sequential movement from one stage to the next with concomitant changes in people's views of knowledge and knowing. *Dimensional models*, on the other hand, characterize people's views of knowledge and knowing in terms of multiple beliefs or theories, which can change independently from one another. Perry's work directly inspired a number of other developmental models of epistemic cognition (e.g., Baxter Magolda, 1992; King and Kitchener, 1994).

Kuhn et al.'s (2000) work on informal reasoning about everyday problems resulted in a model that paralleled Perry's trajectory through objective and subjective views of knowledge. Absolutists view claims as facts that describe or explain a directly knowable, certain, and objective reality. These people do not see a need to think critically beyond determining which sources are expert and which knowledge claims match reality. Multiplists deny that reality can be directly knowable and view claims as subjective opinions, thus making critical thinking irrelevant. Finally, evaluativists share multiplists' view that reality is not directly knowable but, like Perry's committed relativists, they believe that critical thinking can be used to determine which claims are more justifiable than others. One of Kuhn's major contributions to the field of epistemic cognition was to distinguish domains of value judgments (e.g., personal taste, esthetics) from truth judgments (e.g., social and natural sciences) and provide evidence that developmental levels might vary across these domains.

In a major shift in conceptualization of epistemic cognition, Schommer (1990) proposed a model that comprised multiple, somewhat independent belief dimensions on which individuals might vary. These dimensions have naïve and sophisticated poles, with the assumption that more sophisticated beliefs are associated with better performance on various learning outcomes. The dimensional model most widespread in educational psychology was proposed by Hofer and Pintrich (1997) and included four related dimensions: the simplicity of knowledge, the certainty of knowledge, the source of knowledge, and knowledge justification. Such dimensions map on to features of developmental models. For example, Kuhn's (2000) absolutist would be said to believe that knowledge is simple (rather than complex), certain, is external to the knower, and is justified by authority. A key difference between developmental and dimensional models is that the former posit that these beliefs change in a systematic way, whereas the latter posit that people might develop asynchronously across dimensions.

A feature of early models of epistemic cognition is that all assumed that any conceptions about the nature of knowledge did not develop until adolescence. Over the last few decades, developmental research has made clear that such ideas begin to develop quite early in life and continue to develop throughout life (see Sandoval et al., 2016). Research on theory of mind has shown that children develop ideas about the state of other people's thinking by the age of 2 years (Montgomery, 1992). A crucial milestone is the development of false belief: the understanding that other people can believe something different from you, even when you know it to be wrong. Such an idea requires at least a tacit concept of knowledge as something independent of a knower. Developmental research has convincingly demonstrated that from infancy children are adept at drawing causal inferences from empirical evidence. By the time children reach school age (5 years) they can explicitly evaluate their own and others' causal claims and evaluate both the means of obtaining evidence and the quality of such evidence to discriminate between competing claims (Bullock et al., 2009). These abilities appear limited primarily by the knowledge children have of the particular situations being reasoned about, although it is not entirely clear how maturational constraints may operate. Developmental research demonstrates the social nature of such thinking, as much of children's early learning about the world is derived from other, usually older, people's testimony of the world. Children's trust in what other people tell them is not blind; rather, it is based on basic evaluations of trustworthiness referring to such features as the accuracy, competence, coherence, audience reception, and benevolence of sources (Clément, 2016).

Finally, research on learning in the disciplines has been concerned with relationships between how learners think about knowledge in the discipline and their efforts to learn that discipline. This work has been especially prominent in science, both in studies of learners' conceptions of the nature of science (i.e., its epistemology) and their practices of scientific investigation. A range of work has also been conducted in mathematics, history, and more recently in literature (Depaepe et al., 2016; Elby et al., 2016; Lee et al., 2016; Vansledright and Maggioni, 2016). One feature of epistemic cognition research in the disciplines is that it has been framed within discipline-specific epistemological concerns. Sometimes these overlap with the more general conceptions from personal epistemology research; for example, the concern in science education that students understand the tentativeness of science knowledge maps onto the general dimension of certainty of knowledge. Yet, within disciplinary education, the epistemologies of the disciplines are fleshed out more deeply and specifically than is typically the case in the psychology or personal epistemology literatures. Science educators link tentativeness to the theory-laden nature of observation, while asserting that some claims are more certain than others. Historians locate uncertainty in the distance between the present and the past and the ambiguity of historical objects. In mathematics, the issue of certainty presents as whether one believes mathematical procedures are given or constructed. Disciplinary research has been strongly influenced by the conception of epistemic cognition as originally coined by Kitchener (1983), that epistemic cognition is a more specified form of metacognition. Kitchener argued that epistemic cognition is used in situations where some idea of the nature of the answer being sought and possible methods for obtaining it are necessary for successful problem solving. For instance, to evaluate a scientific theory requires some understanding of what makes such a theory credible. That is, such evaluations are discipline-specific.

The last half-century has seen research on epistemic cognition grow immensely in a number of fields. The disparate roots of research into how people generate knowledge, interpret knowledge, and think about the nature of knowledge have begun, since the turn of the millennium, to converge in some ways. Current models are explicitly informed by relevant work in the field of philosophical epistemology. Contemporary models of epistemic cognition also pay more attention to the antecedents and early forms of epistemic cognition suggested by developmental psychology. Earlier debates about the specificity or generality of epistemic beliefs have mostly given way to a recognition that epistemic cognition has contextual and disciplinary aspects that are only somewhat understood, creating a need for better explanations of both the extent and mechanisms of generalization. These trends are converging to create a broader conceptualization of epistemic cognition as a process invoked in any of a myriad of situations in which people may have need to consider how or what they know. These issues are treated in depth in the further reading suggested at the end of this chapter.

Issues in epistemic cognition research

The expansion of research on epistemic cognition has raised a number of outstanding issues in the study of epistemic cognition. They include both how scholars theorize the nature of cognition itself and the consequences for modeling epistemic cognition, and how epistemic cognition can be observed and measured.

A fundamental issue in making sense of research on epistemic cognition is making sense of the cultural turn in psychology and other social sciences during the last decades of the 20th and early decades of the 21st centuries. In psychology, the emergence of Vygotskian and other sociocultural perspectives on development have co-occurred with the emergence of social perspectives in epistemology and in science studies (Duschl, 2008). The cultural turn includes a shift from seeing cognition as a fundamentally individual activity to a fundamentally social one, and from knowledge as an individual to a fundamentally social entity. This shift has been taken up to different extents, and differences in where scholars, and even fields, see the locus of explanation of epistemic cognition on this individual-social continuum creates a tangle of empirical results and interpretations that have yet to be clarified.

The cognitive constructivism underlying Piaget's genetic epistemology (Piaget, 1972) abandons the idea that knowledge is an object in the world. Instead, "*knowledge* refers to conceptual structures that epistemic agents, given the range of present experience within their tradition of thought and language, consider *viable*" (Von Glasersfeld, 1989, p. 124). That is, while cultural traditions of thought and language may shape experience, individual epistemic agents build conceptual structures that can viably be considered knowledge. This is not to say that constructivists necessarily abandon the idea of an objective reality (see Phillips, 1995, for a discussion of variants of constructivism on this point), only that individuals construct their knowledge of reality in their minds. This view of knowledge construction as an individual achievement continues to underlie much of the psychology research pertinent to epistemic cognition. The term *personal epistemology* as a moniker for this research itself conveys the assumption that persons construct epistemologies as part of their efforts to interpret the world around them. Both the developmental and dimensional models of epistemic cognition developed from personal epistemology research maintain this cognitive constructivism as the underlying definition of knowledge. Developmental models propose progressive trajectories that vary mainly in the generality they posit for the stages along the way. Dimensional models articulate dimensions of belief and their possible influences and interactions. All of these models share an assumption that cognitive representations (i.e., knowledge and belief) are constructed by individuals, via some level of reflection on experience, and then applied to subsequent experiences.

Sociocultural theories of cognition and development derived from Vygotsky (1962, 1978) locate the individual knower within a specific cultural and historical context, thereby framing knowledge as a cultural product and knowing as a cultural activity. Models of epistemic cognition that have been influenced by sociocultural perspectives vary in where they locate knowledge and knowing. What can be called a cognitive contextualist view attempts to account for how social factors influence individual knowledge construction, including individuals' ideas about knowledge and knowing. Such models attempt to account for how environmental (i.e., social) factors and features of specific domains influence individual epistemic cognition. Some propose that general epistemic beliefs develop within the social and cultural context in which people grow up (Muis et al., 2006). Elby and Hammer (2010) suggest that participation in culturally shaped patterns of activity lead to the construction of epistemic "resources," fine-grained cognitive structures tied closely to contexts of use. These contexts of use become bound into epistemic "frames" that organize the conditional activation of epistemic resources. These models differ in their proposed mechanisms of generalization of epistemic conceptions, and the extent of such generalization, but share the notion that social factors are influences on individual cognitive construction.

A stronger social perspective on knowledge and knowing derives from situative perspectives on cognition (Engeström, 1987; Lave and Wenger, 1991), allied with the social epistemology (Fuller, 1988) view of knowledge as fundamentally a social construct and epistemology as a social phenomenon. The situative perspective takes as its base unit of analysis the individual-in-interaction, emphasizing features of interaction external to individuals and their minds. This interactional view on cognition sees knowledge as distributed among social and material resources and deployed within cultural activity. Thus, individuals' and groups' interpretations of the purpose of an activity are seen as shaping the epistemic cognition that might occur. Chinn et al.'s (2014) AIR model highlights processes of epistemic cognition as tied to the epistemic aims an individual might pursue during activity, how those aims relate to non-epistemic aims, and how the aims are related to reliable processes of knowledge development. These reliable processes are social practices developed within particular communities to solve specific epistemic problems. The situative perspective on epistemic cognition thus sees knowledge as reflected in practice—in how people accomplish activity—rather than as a mental representation of some aspect of reality that is then applied.

The divergent theoretical perspectives on knowledge and cognition generally are reflected in current models of epistemic cognition, not surprisingly. The crucial difference among these perspectives with respect to epistemic cognition research is that they theorize knowledge itself, its development, and related concepts such as belief and the role of context, rather differently. Such differences have implications for the unit of analysis in epistemic cognition research (e.g., individuals' construction of knowledge propositions vs. the interplay of cultural forms of knowledge and individuals' practice) as well as for perspectives on how epistemic cognition develops (e.g., a focus on changes in individuals' conceptions of knowledge vs. changes in a community's norms and practices and individual participation in those). Whether such differences are incommensurate is perhaps an ideological question. What is clear is that broad differences in how knowledge and knowing are theorized produce differences in how issues of contextualization and generality are addressed in research on epistemic cognition. They lead to different explanations of the nature of developmental pathways for epistemic cognition. They also, naturally, produce differences in research methods for identifying, defining, and studying phenomena of epistemic cognition.

Early models of epistemic cognition, both developmental and dimensional, that proposed epistemic beliefs as general and applied consistently across domains have fallen out of favor. The idea that epistemic cognition occurs at different levels of specificity, as well as the idea that there are discipline-specific forms of epistemic cognition, has gained substantial empirical support and more recent models attempt to account for both of these ideas (see [Sandoval et al., 2016](#)). At the same time, evidence that at least some conceptions of knowledge and knowing appear at least somewhat general has left scholars of epistemic cognition with something of a dilemma. On the one hand, mechanisms of generalization are not yet clearly articulated, leaving a conceptual gap in explaining how any individual's experiences might produce particular forms of generalized epistemic conceptions, and what the limits of such generalization might be. On the other hand, methods for empirically observing or measuring epistemic conceptions may produce specificity or generality as an unintended artifact of instrumentation. For example, a number of studies have demonstrated that students' epistemic conceptions vary by domain, and even topic. Yet, given that the measures used presume, *a priori*, the existence of a common set of dimensions or beliefs across domains, it could be argued that what has been captured within such research is not really domain- or discipline-specific epistemic cognition but merely the extent to which scores on general belief dimensions vary across domains. Consequently, both the extent to which particular aspects of epistemic cognition may be general (applied across a range of contexts) and the mechanisms by which such generalizations develop remain unclear. The variability of epistemic cognition seen in research grounded in the disciplines (science, mathematics, etc.) implies that general models of epistemic belief cannot capture important differences in the epistemic norms and practices across the disciplines. Developing models grounded within specific disciplines and comparing them may provide a path for scholars to consider the aspects of epistemic cognition that potentially generalize and how they do so.

[Sandoval et al. \(2016\)](#) review of research designs and methods in epistemic cognition revealed a number of trends concerning the assessment methods used, the contexts in which research is conducted, and the kinds of epistemic cognition under study. Psychometric problems with self-report questionnaires and similar methods of assessing epistemic cognition, coupled with the often decontextualized or inauthentic contexts in which these data are gathered, and their presumptions of domain-consistency or trait-like aspects of epistemic cognition, cast doubt on research findings from such studies. The growing acceptance that epistemic cognition is largely tacit, and that people's enactment of epistemic cognition can often be different from their expressed reconstruction of that cognition, has led an increasing number of researchers to pursue concurrent assessment methods (e.g., observations, discourse analysis, think-aloud protocols) that can be implemented in contexts where participants are asked to think about knowledge and knowing (e.g., inquiry activities, searching the Internet for information on controversial issues). Researchers' models of epistemic cognition seem to predict their choice of context, with adherents to dimensional or developmental models often investigating how people enact epistemic cognition within clinical research settings, and adherents to situative models of epistemic cognition more often using various kinds of inquiry contexts as research settings. Given the wide variety of both conceptual models and concomitant research methods, there is a clear need for research involving multiple, mixed methods of gathering data, across a variety of contexts, to understand how epistemic cognition happens and develops. It would be especially salutary if competing theories and models were tested against one another.

Emerging trends in epistemic cognition research

Two recent trends in epistemic cognition research warrant mention. One is a general expansion of contexts in which scholars are studying epistemic phenomena. There is a general acknowledgment that how one thinks about what it means to know, what counts as knowledge, and how knowledge can be evaluated is relevant to a great many contexts in everyday life, in addition to formal school environments. Consequently, scholars are exploring how epistemic cognition occurs in various out-of-school settings, including museums, after-school clubs, and even everyday family interactions in the home or on walks in nature (e.g., [Bricker and Bell, 2012](#); [Marin and Bang, 2018](#)). This work reveals how epistemic cognition is sensitive to the contexts in which people interact and learn. How people think about knowledge seems to depend highly on what sort of knowledge they see themselves as trying to produce or evaluate, and for what purposes they are doing so. Scholars therefore are exploring how non-epistemic features of situations, such as family and other social relations, influence how people reason about epistemic issues. This has led to more expansive concepts related to epistemic cognition, such as epistemic climate ([Muis and Duffy, 2013](#)) and epistemic affect ([Jaber, 2015](#)), in efforts to account for how knowledge construction and evaluation are intertwined with most aspects of human activity and interaction. This expansion of the domains considered relevant to epistemic cognition and constructs to account for it is related to the broader cultural turn in psychology and the social sciences discussed above, reflecting that "knowing" is an inherent aspect of activity. At the same time, this expansion of the scope of epistemic cognition research has rendered the boundaries of the phenomena and constructs rather fuzzy. On the one hand, there is a salutary effect on research in acknowledging that human activity is nearly always knowledge-rich in multiple ways, so that studying how it is so across a broad range of contexts is likely to help us understand more clearly how people's ideas about what and how they know develop. On the other hand, a profusion of constructs and theoretical conceptualizations makes the production of a coherent account of epistemic cognition seem distant.

A second important trend is a push, driven mainly by Indigenous scholars, to recognize that Western views of knowledge constitute only one epistemological system, one that arguably does harm to people who do not share it. Within science education, in particular, the last decade has seen increasing attention to framing epistemic features of scientific disciplines as one particular sort of epistemology, and comparing that to other ways of knowing, especially Indigenous knowledge systems of nature (Marin and Bang, 2018; Marin et al., 2017). One interest in such scholarship is how members from non-dominant groups negotiate between the epistemologies of their home communities and traditions and those of institutional settings, like school, that embody Western epistemology. Another interest is how people navigate between multiple epistemologies as they become involved in projects related to equity and justice and thus make efforts to understand knowledge systems from other cultures. This work draws more directly on feminist and other standpoint epistemologies, Indigenous epistemologies, and more. It is very clearly located within situative perspectives on knowledge and cognition, theorizing knowledge itself as a cultural construct and cultural variations in forms and functions of knowledge as central foci of research (Marin et al., 2017).

Today research on epistemic cognition is being conducted across a broad array of fields and theoretical perspectives. There is a broad movement away from efforts to assess epistemic cognition using self-report measures, increased respect for the epistemological variations among disciplines and between cultures, and an increasing reliance on more direct methods of observation of epistemic cognition in action. As cultural differences in worldviews become not just increasingly apparent but increasingly hostile to each other, there is an urgent awareness that formal schooling must help young people to become reflective about their own views of knowledge and open to and tolerant of differing views. The complex issues confronting the world now and over the coming decades require citizens who can effectively engage in complex deliberations about how to resolve them in equitable ways. Research on epistemic cognition is likely to contribute in important ways to that educational project.

References

- Baxter Magolda, M.B., 1992. *Knowing and Reasoning in College*. Jossey-Bass, San Francisco.
- Bricker, L., Bell, P., 2012. Argumentation and Reasoning in Life and in School: Implications for the Design of School Science Learning Environments. In: Khine, M. (Ed.), *Perspectives on Scientific Argumentation: Theory, Practice, and Research*. Springer, Dordrecht, pp. 117–133.
- Bullock, M., Sodian, B., Koerber, S., 2009. Doing experiments and understanding science: development of scientific reasoning from childhood to adulthood. In: Schneider, W., Bullock, M. (Eds.), *Human Development from Early Childhood to Early Adulthood: Findings From a 20 Year Longitudinal Study*. Psychology Press, New York, NY, pp. 173–197.
- Chinn, C.A., Rinehart, R.W., Buckland, L.A., 2014. Epistemic cognition and evaluating information: applying the AIR model of epistemic cognition. In: Rapp, D.N., Braasch, L.G. (Eds.), *Processing Inaccurate Information*. MIT Press, Cambridge, MA, pp. 425–453.
- Clément, F., 2016. Social cognition. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 98–111.
- Depaepe, F., DE Corte, E., Verschaffel, L., 2016. Mathematical epistemological beliefs: a review of the research literature. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 147–164.
- Duschl, R.A., 2008. Science education in 3 part harmony: balancing conceptual, epistemic and social goals. *Rev. Res. Educ.* 32, 268–291.
- Engeström, Y., 1987. *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Orienta-Konsultit, Helsinki.
- Elby, A., Hammer, D., 2010. Epistemological resources and framing: a cognitive framework for helping teachers interpret and respond to their students' epistemologies. In: *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*, pp. 409–434.
- Elby, A., Macrander, C., Hammer, D., 2016. Epistemic cognition in science: uncovering old roots to turn over new leaves. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 113–127.
- Fuller, S., 1988. *Social Epistemology*. Indiana University Press, Bloomington.
- Hofer, B.K., Pintrich, P.R., 1997. The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67, 88–140.
- Jaber, L., 2015. Attending to students' epistemic affect. In: Robertson, A.D., Scherr, R., Hammer, D. (Eds.), *Responsive Teaching in Science and Mathematics*. Routledge, New York.
- King, P.M., Kitchener, K.S., 1994. *Developing Reflective Judgment*. Jossey-Bass, San Francisco.
- Kitchener, K.S., 1983. Cognition, metacognition, and epistemic cognition. *Hum. Dev.* 26, 222–232.
- Kuhn, D., Cheney, R., Weinstock, M., 2000. The development of epistemological understanding. *Cognit. Dev.* 15, 309–328.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, UK.
- Lee, C.D., Goldman, S.R., Levine, S., Magliano, J., 2016. Epistemic cognition in literary reasoning. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 165–183.
- Marin, A., Bang, M., 2018. "Look it, this is how you know:" family forest walks as a context for knowledge-building about the natural world. *Cognit. Instr.* 36, 89–118.
- Marin, A., Medin, D., Ojalehto, B., 2017. Conceptual change, relationships, and cultural epistemologies. In: Amin, T., Levri, O. (Eds.), *Converging Perspectives on Conceptual Change: Mapping an Emerging Paradigm in the Learning Sciences*. Routledge, New York, NY.
- Montgomery, D.E., 1992. Young children's theory of knowing: the development of a folk epistemology. *Dev. Rev.* 12, 410–430.
- Muis, K.R., Bendixen, L.D., Haerle, F.C., 2006. Domain-general and domain-specificity in personal epistemology research: philosophical and empirical reflections in the development of a theoretical framework. *Educ. Psychol. Rev.* 18, 3–54.
- Muis, K.R., Duffy, M.C., 2013. Epistemic climate and epistemic change: instruction designed to change students' beliefs and learning strategies and improve achievement. *J. Educ. Psychol.* 105, 213–225.
- Perry Jr., W.G., 1970. *Forms of Intellectual and Ethical Development in the College Years*. Holt, Rinehart & Winston, New York.
- Phillips, D., 1995. The good, the bad, and the ugly: the many faces of constructivism. *Educ. Res.* 25 (5), 5–12.
- Piaget, J., 1972. *Principles of Genetic Epistemology: Selected Works*. Routledge, New York.
- Sandoval, W.A., Greene, J.A., Bråten, I., 2016. Understanding and promoting thinking about knowledge: origins, issues, and future directions of research on epistemic cognition. *Rev. Res. Educ.* 40, 457–496.
- Schommer, M., 1990. The effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psychol.* 82, 498–504.
- Vansledright, B., Maggioni, L., 2016. Epistemic cognition in history. In: Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, pp. 128–146.

- Von Glasersfeld, E., 1989. Cognition, construction of knowledge, and teaching. *Synthese* 80 (1), 121–140.
- Vygotsky, L.S., 1962. *Thought and Language*. MIT Press, Cambridge, MA.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.

Further reading

- Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), 2016. *Handbook of Epistemic Cognition*. Routledge, New York.

Executive function and education

Destany Calma-Birling, Andrei Semenov, and Philip David Zelazo, Institute of Child Development, University of Minnesota, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Classroom conditions supportive of EF	169
Training students' EF skills	172
Conclusion	174
References	174

Executive function (EF) skills are a set of attention-regulation skills that modulate attention in the service of a goal, and they are usually measured behaviorally as three inter-related skills: cognitive flexibility, working memory, and inhibitory control (Miyake et al., 2000; Zelazo et al., 2013). Cognitive flexibility involves thinking about something in multiple ways—for example, considering someone else's perspective on a situation. Working memory refers to keeping information in mind and manipulating it in some way. You might keep the beginning of a paragraph in mind while you read the rest of it, for example. Inhibitory control involves deliberately suppressing attention (and other responses) to something (e.g., ignoring a distraction while taking a test). Respectively, these three EF skills involve using attention flexibly, maintaining attention over time, and focusing attention selectively. Although these skills can be distinguished, they typically work together when we are pursuing a goal, and they are all important for characterizing problems, reasoning (making inferences), planning, keeping goals and plans in mind, and being flexible in the face of feedback.

These EF skills are often measured using abstract, decontextualized tasks that lack a significant affective or motivational component, and as such they can be considered relatively “cool” EF skills. In contrast, more “hot” EF skills are needed in situations that people really care about because they stand to win or lose something (Zelazo and Müller, 2002; Zelazo and Carlson, 2012). These situations require us to govern the motivational tendencies to approach or to avoid a meaningful stimulus. A classic example is delay of gratification as measured by the famous “marshmallow test” (Mischel et al., 1989), in which children are told they can take a small treat now or wait for a larger one. This task requires children to avoid the impulse to take the tempting reward immediately, however compelling it may be, and instead make the prudent choice to wait for the larger delayed reward.

Both cool and hot EF skills provide a necessary foundation for learning in a classroom setting: They make it possible for children to sit still, pay attention, remember and follow rules, and manage their emotions and behavior (Zelazo et al., 2016). By enabling students to hold information in mind and flexibly adopt new perspectives, cool EF skills also allow for more fully engaged, active, and reflective forms of learning (Marcovitch et al., 2008; Zimmerman, 2008), which leads to deeper and longer lasting understanding. Indeed, children with better cool EF skills actually learn more (i.e., retain more information) from a given amount of instruction and practice than do their peers with worse cool EF skills (Benson et al., 2013). For example, children with better cool EF skills showed a larger gain in math achievement between kindergarten and first grade, especially on applied problems (Hassinger-Das et al., 2014), and children with better cool EF skills were more likely to learn new concepts when trained in biology (Grenell and Carlson, 2021; Zaitchik et al., 2014).

Whereas cool EF is more strongly associated with academic outcomes, including math and reading, poor hot EF is more strongly associated with problem behaviors in school (e.g., inattentive and overactive behavior; Brock et al., 2009; Willoughby et al., 2011). Hot EF skills are important for helping children to manage their emotions and social interactions, and problem behaviors may interfere with school success. Hot EF skills in early childhood also predict later academic outcomes, even controlling for variables such as vocabulary and maternal education. For example, children's performance on a measure of delay of gratification predicted school performance in adolescence (e.g., Mischel et al., 1989). Similar long-term predictions are found for hot EF skills. For example, children's attention span-persistence, required for attending to relevant information and persisting on a task, measured at age 4 years, not only predicted math and reading achievement at age 21 years but also predicted the odds of completing college by age 25 (McClelland et al., 2013). The predictive power of EF is often greater than that of IQ, and long-term predictions of academic success and other positive outcomes are found even after controlling for IQ and childhood SES (e.g., Duckworth and Seligman, 2005; Moffitt et al., 2011; Purpura et al., 2017; Richmond-Rakerd et al., 2021; Viterbori et al., 2015).

Although EF skills are related to many aspects of academic achievement, including reading and science (e.g., Gropen et al., 2011; Kieffer et al., 2013; Nayfeld et al., 2013; St Clair-Thompson and Gathercole, 2006), the relations between EF skills and math appear to be especially robust (e.g., Blair and Razza, 2007; Blair et al., 2015; Fuhs et al., 2014; Hassinger-Das et al., 2014; Schmitt et al., 2015; Zelazo and Carlson, 2020). Positive correlations are seen even when controlling for baseline math skills, crystallized intelligence, verbal ability, general processing speed, and family background (e.g., Blair et al., 2015; Bull et al., 2011; Espy et al., 2004; McClelland et al., 2007; Monette et al., 2011; Prager et al., 2016; Verdine et al., 2014).

Several longitudinal studies have found prospective links from EF skills in early childhood to later math abilities (Blair et al., 2015; Bull et al., 2008; Clark et al., 2010, 2013; Fuhs et al., 2014; McClelland et al., 2007; Viterbori et al., 2015; Welsh et al., 2010). Studies have also found that early EF skills are a stronger predictor of later math abilities in elementary school than is IQ

or processing speed. In fact, EF skills have been found to be even better than math instruction in predicting growth in math abilities (Ribner, 2020), and serve as a protective factor for students at risk for math learning difficulties (Ribner et al., 2017).

Good EF skills also protect children against the academic and psychosocial risks associated with poverty and disadvantage. For example, although children who are (or have been) homeless and highly mobile are, on average, more likely than other children to perform poorly in school, this does not apply to children with relatively good EF skills; homeless children with good EF skills tend to succeed in school despite being homeless, thus showing evidence of resilience (Masten et al., 2012).

Research like this suggests that EF skills may play a key role in the academic achievement gap between children from lower versus higher socioeconomic status (SES) backgrounds. On average, children from lower SES families show lower levels of educational achievement than children from higher SES families, a gap that has widened over recent decades (Reardon, 2013). This gap emerges early (it is evident in kindergarten) and grows or remains stable over time (Masten, 2013). Lower SES is reliably associated with lower EF skills, even when controlling for general cognitive ability (for a meta-analysis, see Lawson et al., 2018), and EF skills account for more of the correlation between SES and academic outcomes than do measures of IQ or language ability (Fitzpatrick et al., 2014; Lawson and Farah, 2017; Nesbitt et al., 2013). Research comparing children with better and worse EF skills (and social competence) shows a gap in reading and math that starts in kindergarten, becomes larger until second grade, and then persists at least through sixth grade (e.g., Cutuli et al., 2013; Obradovic et al., 2009; McClelland et al., 2006).

Children from lower-SES families have been found to receive less environmental support for EF skills (e.g., fewer opportunities to play games that require attention regulation; Korucu et al., 2019), and they are also more likely to be exposed to more chaos, including inconsistent schedules and the lack of family routines (e.g., Matheny et al., 1995). This can make it difficult for children to predict and anticipate sequences of events (essential for goal-directed behavior such as planning and delay of gratification), and these lifestyle factors often disrupt sleep (e.g., Koopman-Verhoeff et al., 2019) and elicit high levels of chronic and recurring stress (e.g., Lupien et al., 2009)—effects that are likely to impair EF development over time due to the toxic influence of high levels of cortisol (a stress-related hormone) on relevant brain structures (e.g., Kolb et al., 2017; Shonkoff et al., 2012). Lower SES in children and youth has been found to be related to higher cortisol (Vliegthart et al., 2016), which in turn is related to smaller regional brain volumes (Merz et al., 2019). In a cross-sectional study of 1099 typically developing individuals (3–20 years), cortical surface area partially mediated the relation between SES and cool EF skills, but not the relations between SES and vocabulary or reading (Noble et al., 2015).

Children who are at risk for difficulties with EF skills include not only children growing up in poverty and exposed to more stress (e.g., Hanson et al., 2010), but also children with a wide range of clinical conditions with childhood onset (Zelazo, 2020). Indeed, difficulties with EF skills are pervasive characteristics of atypical development, and are seen in learning difficulties and learning disorders, as well as other conditions that can interfere with classroom learning, such as ADHD and disruptive behavior problems. Atypical development of hot and/or cool EF skills is also implicated in many other conditions and appears to be a transdiagnostic indicator of psychopathology (Zelazo, 2020).

The healthy development of EF skills during childhood is influenced importantly by caregivers, including teachers, who support the acquisition of these skills in specific ways. Research suggests that the development of EF skills depends importantly on autonomy-supportive mentorship, which includes providing guidance consistent with children's current level of competence (i.e., scaffolding), encouraging children to take initiative on challenging tasks, taking children's perspectives, respecting their pace, and providing them with choices, as opposed to taking control or letting them struggle too much on their own (Whipple et al., 2011). These behaviors are believed to foster independent problem solving, agency, and intrinsic motivation, thereby increasing children's opportunities to engage in self-initiated goal-directed activities and gradually strengthen their EF skills through practice, especially when combined with appropriate environmental structure (Grolnick and Pomerantz, 2009). Importantly, the sensitive calibration of challenges to children's developing skill level (providing problems that are not too easy and not too hard) allows children to obtain more practice actually solving problems, and thereby to build their EF skills—to learn by doing.

The effects of experience on brain structure and brain function, including EF skills, are seen across the lifespan. Collectively, this research shows that the activation of specific neural pathways, repeated over time, changes those pathways, rendering them more efficient (Hebb, 1949). For example, school-age children with abacus training showed evidence of greater myelination in parts of corpus callosum as well as other fiber tracts involved in motor and visuospatial function (Hu et al., 2011). These practice-related differences in myelination were correlated with behavioral performance. Similarly, children given a brief opportunity to practice EF skills in a scaffolded, supportive context showed improvements not only in their EF skills, but also in their social understanding, and these changes were associated with predictable changes in neural function (Espinete et al., 2013).

Classroom conditions supportive of EF

Similar to the home environment, the school environment can promote the healthy development of children's and adolescents' EF skills. For instance, recent research indicates that children's working memory grows significantly more during the school year months compared to the summer months, suggesting that the school environment provides children with unique opportunities to improve and practice their EF skills (Finch, 2019). Over the last several decades, research examining different aspects of the classroom environment in relation to students' EF skills have revealed a number of classroom conditions that can influence those skills, including classroom visual displays, teacher-student relationships, classroom management behaviors, the type of instructional support, and teachers' EF skills and well-being.

The ability to focus and sustain attention to relevant information are important for learning in the classroom. Growing evidence indicates that visual features of the classroom environment, specifically wall displays, can influence students' cognitive abilities and academic outcomes. For example, kindergarteners randomly assigned to a decorated classroom (i.e., many items displayed on the wall) spent significantly more instructional time off task and performed significantly worse on the post-lesson learning assessments compared to kindergarteners assigned to a non-decorated classroom (i.e., no items displayed on the wall; [Fisher et al., 2014](#)). Moreover, time off task mediated the effect of the classroom condition on kindergarteners' learning; being in a decorated classroom led to greater time off task which in turn predicted worse performance on the learning assessments.

Subsequent studies using different study designs, learning assessments, ecological contexts, and study populations have reported similar findings, namely that too many visual displays in primary school students' surrounding environment is associated with poor learning outcomes ([Barrett et al., 2015](#); [Hanley et al., 2017](#)). The influence of visually rich environments on EF skills has also been studied more directly. Children's and adolescents' performance on measures of working memory was significantly worse when randomly assigned to a testing room with many pictures compared to one with no pictures ([Rodrigues and Pandeirada, 2018, 2019](#)).

Building rapport with students is an important aspect of teachers' job as it creates a classroom environment that is more supportive of students' academic learning and the healthy development of their EF skills. Teacher-student relationships are associated with students' adjustment to school, their classroom engagement, and their academic achievement ([Berry, 2012](#); [Berry and O'Connor, 2010](#); [Cadima et al., 2016b](#); [Kořir and Tement, 2014](#); [Rimm-Kaufman et al., 2014](#)). Teacher-student relationships characterized by high conflict (i.e., high degree of discordance, negativity, and unpredictability in the teacher-student relationship) predict worse EF skills in elementary school students ([Berry, 2012](#); [de Wilde et al., 2016](#); [McKinnon and Blair, 2019](#)), whereas teacher-student relationships characterized by high closeness (i.e., high degree of warmth, positive affection, and emotional security between the teacher and student) predicts better EF skills in both elementary and middle school students ([Cadima et al., 2016b](#); [Vandenbroucke et al., 2017](#)).

Preliminary research indicates that positive teacher-student relationships protect against the negative effects of negative parent-child relationships and cumulative risk on children's EF skills. For instance, in an experimental study examining the effect of teacher support on elementary and middle school students' working memory, researchers found that when students reported a negative relationship with their parents, hearing an emotionally supportive message from a teacher after a stress induction task led to students' performing better on a measure of working memory, whereas hearing a supportive message from their parent was associated with worse performance ([Vandenbroucke et al., 2017](#)). In a more recent study, the negative correlation between cumulative risk and working memory in kindergarten students was found to be attenuated by teacher-child closeness ([Suntheimer and Wolf, 2020](#)). Together, these findings suggest not only that teacher-student relationships promote (or hinder) the healthy development of EF skills in children and adolescents, but also that positive teacher-student relationships can serve as a protective factor for students' experiencing stress at home and in other areas of their life.

Regarding negative teacher-student relationships, the association between high conflict teacher-student relationships and low levels of students' EF skills might be due to children's maladaptive behaviors in the classroom (e.g., increased aggression and inattention problems) and elevated stress levels. For example, one study showed that children with low levels of inhibitory control prior to kindergarten entry were more likely to experience higher levels of conflict with their teacher, on average; and this relationship was partially explained by children's greater inattentive and aggressive behaviors, as reported by teachers ([Berry, 2012](#)). Further, the findings linking poor inhibitory control to higher levels of teacher-student conflict were observed at multiple time points during the elementary school years, suggesting that aggression and inattention problems may set the stage for the dynamic interplay between teacher-student conflict and children's EF development across elementary school.

Because stress is known to undermine children's EF skills (e.g., [Blair et al., 2017](#)), and conflict with teachers has been shown to increase students' physiological stress levels ([Lisonbee et al., 2008](#)), the negative association between teacher-student conflict and students' EF might also be a result of students' elevated stress levels. Although research in the context of teacher-student relationships has yet to directly examine the associations between teacher-student conflict, stress, and students' EF skills, there is some evidence that shows high conflict relationships are negatively related to children's EF skills via elevated cortisol ([Martin et al., 2017](#)).

Positive teacher-student relationships may promote students' EF skills through increased engagement and exploration. When students have positive interactions and a close relationship with their teachers, they are more likely to enjoy school ([Gest et al., 2005](#)), have higher levels of confidence ([Cui et al., 2020](#)), and feel safe in the classroom ([Ben-Gal Dahan and Mikulincer, 2020](#)), leading them to engage in more challenging activities and show greater persistence when encountering difficulties during task completion ([Kikas and Tang, 2019](#); [Roorda et al., 2017](#)). Engaging in more challenging activities may provide students with more opportunities to practice using their EF skills, and thereby to grow them.

When teachers demonstrate effective classroom management—by clarifying rules and expectations, using proactive behavior management strategies, minimizing transition time between classroom activities, and establishing consistent classroom routines—they create conditions that support students' academic outcomes and cognitive skills. Studies have shown that young children's EF skills exhibit greater growth over the school year in classrooms where teachers use effective classroom management behaviors ([Choi et al., 2016](#); [Raver et al., 2011](#); [Rimm-Kaufman et al., 2009](#)), and effective classroom management has been linked to greater academic achievement, classroom engagement, and on-task behavior in kindergarteners ([Brock et al., 2009](#); [Rimm-Kaufman et al., 2009](#)). In well-managed classrooms, it is likely that children know what to do when transitioning from one activity to another due to consistent routines and clear expectations set by their teacher; whereas in poorly managed classrooms, children may not know what to do because of unclear expectations and inconsistencies in the organization of daily activities, resulting in more

off-task and counterproductive behaviors. Thus, it is likely that well-managed classrooms provide external guidance for students to practice organizing their behavior in a goal-directed manner, which subsequently strengthens their EF skills through the internalization of regulation strategies.

In addition to effective classroom management, an autonomy-supportive teaching style has been associated with a number of positive academic outcomes in students, including higher levels of academic intrinsic motivation, greater classroom engagement and use of self-regulated learning strategies, and academic achievement (Li et al., 2021; Reyes et al., 2012; Schuitema et al., 2016; Sierens et al., 2009). Teachers who use autonomy-supportive instructional practices structure classroom activities and communicate in ways that motivate students to become engaged in their learning because it reflects the students' own personal values, preferences, interests, and goals (Patall and Zambrano, 2019). Some examples of teachers' autonomy-supportive behaviors and instructional practices include: providing choice and seeking student input in learning activities, taking students' perspectives, acknowledging and accepting students' negative affect and resistance rather than suppressing or ignoring them, providing explanatory rationales of the personal relevance of learning topics and activities, using invitational language as opposed to controlling language, offering hints when students seem stuck, encouraging and praising effort and persistence, and displaying patience (Reeve and Cheon, 2021).

Research examining the associations between teachers' autonomy-supportive behaviors and students' EF skills has yielded similar results to research relating maternal autonomy-supportive behaviors to children's EF development (e.g., Bernier et al., 2010). Teachers' use of autonomy-supportive behaviors is positively associated with students' EF skills in childhood and adolescence (Merritt et al., 2012; Piccolo et al., 2019; Schuitema et al., 2016; Sosic-Vasic et al., 2015). Autonomy-supportive teachers likely influence students' EF skills through multiple pathways, including their instructional techniques, classroom management behaviors, and relationships with their students.

Regarding instructional techniques, autonomy-supportive teachers might support students' EF skills by teaching students how to become self-regulated learners. Self-regulated learners are active agents in their learning process, meaning they control their learning by setting goals, using effective study strategies, planning and using their time wisely, organizing and prioritizing materials and information, monitoring their learning and shifting approaches as needed, and seeking feedback on their performance (Zimmerman, 2008). There is growing evidence that students' use of self-regulated learning strategies is associated with their EF skills (Davis et al., 2021; Effeney et al., 2013; Follmer and Sperling, 2016; Rutherford et al., 2018). Observational research of teachers' use of self-regulated learning strategies in the classroom showed that teachers' scaffolding of students' planning and organization behaviors and flexible thinking (i.e., teaching students how to consider different approaches, perspectives, and answers to the same question or problem and providing strategies, resources, and tools to help students with planning and organization) positively predicted students' EF skills across the school year (Bardack and Obradović, 2019).

Additionally, viewing students as active agents in their own learning and allowing students to make choices in their learning activities might also help to explain the positive correlation between an autonomy-supportive teaching style and students' EF skills. For instance, in contrast to teacher-directed activities, child-directed activities are rooted in educational philosophies (e.g., the Montessori method) that perceive children as active, rather than passive, learners (van Hout-Wolters et al., 2000). In child-directed activities, children are allowed to choose what activities to be involved in and receive less direct input from their teachers. The active nature of child-directed activities is thought to provide students with opportunities for practicing and developing their EF skills. Indeed, in one study examining classroom time use and children's EF, researchers found that children who spent more time in (child-directed) free choice activities exhibited greater gains in inhibitory control across the school year (Goble and Pianta, 2017). Similarly, kindergarteners enrolled in classrooms that spent the most time in child-directed activities (about a third of the school day) demonstrated larger gains in cognitive flexibility across the school year (Ansari and Purtell, 2017). Other research comparing EF skills between children attending Montessori schools and children attending traditional public schools revealed similar findings, namely that students attending Montessori schools perform significantly better on measures of EF compared to children attending traditional schools (Courtier et al., 2021; Denervaud et al., 2019; Lillard and Else-Quest, 2006).

The work of teachers is dynamic, requiring constant attention and awareness to the shifting needs of students and the changing demands of classroom interactions. As a result, effective teaching requires EF skills. Primary and secondary school teachers' with better EF skills are more likely to engage in exemplary classroom practices, namely cultivating positive student-teacher relationships; creating a classroom climate where students feel valued, supported, and welcomed; making their lessons engaging, accessible, and coherent for their students; and creating classroom conditions that allow students to stay focused on achieving their full potential (Hiver et al., 2021). Lastly, teachers who displayed greater EF difficulties (e.g., impulsive, distracted, and disorganized classroom behaviors) were less effective at managing their time and routines in the classroom, less likely to teach their students strategies and practices that support their planning and organizational skills, and less likely to engage in autonomy-supportive teaching practices (Bardack and Obradovic, 2019). Although more research examining the associations between teachers' EF skills and classroom practices is needed, these findings suggest that teachers' EF skills might play an important role in their use of effective classroom practices.

Teaching often involves long hours without meaningful breaks or interactions with colleagues, low pay and appreciation, overcrowded classrooms and class discipline problems, and inadequate training in classroom management (Allegretto and Mishel, 2016; Stevenson et al., 2020). Not surprisingly, these demands take a psychological toll on teachers, resulting in higher levels of

perceived stress and burnout compared to other professions (Johnson et al., 2005; Steiner and Woo, 2021). The high levels of teacher stress and burnout are concerning because they are related not only to teachers' health and well-being (Burke and Green-glass, 1996; Collie et al., 2012; Poormahmood et al., 2017; Steinhardt et al., 2011; Suleman et al., 2018), but also to teacher attrition rates (Diliberti et al., 2021; Madigan and Kim, 2021); and teacher attrition has been shown to hinder students' academic progress (Ronfeldt et al., 2013).

Teachers experiencing high levels of stress and burnout likely have a reduced capacity to engage in the work of cultivating positive relationships with their students, using effective teaching methods, and creating a supportive and inclusive classroom environment for students. Indeed, teachers who experience stress and exhaustion tend to withdraw from relationships with students and report more conflict with their students (Chang, 2009; Whitaker et al., 2015). They also report feeling less effective and confident in their teaching and in providing support for their students (Grayson and Alvarez, 2008; Hoglund et al., 2015). Findings from observational and experimental studies indicate that stress impairs teachers' ability to implement effective classroom practices (Braun et al., 2019; Larson et al., 2018; Reinke et al., 2013), resulting in more negative social-emotional and academic outcomes among students (Hoglund et al., 2015; McLean and Connor, 2015). There is even evidence showing that teacher stress is negatively related to kindergartners' EF skills (Neuenschwander et al., 2017). Thus, the findings from these studies underscore the importance of reducing teacher stress to promote classroom environments where both teachers and students can flourish.

Training students' EF skills

Given the predictive importance of EF skills, there is a strong desire to improve or train students' EF skills. Early intervention is particularly important because improvements can have cascading positive effects on a child's life (e.g., Moffitt et al., 2011). Early interventions take advantage of periods of heightened neuroplasticity (often called "sensitive periods") when certain regions of the brain are particularly sensitive to environmental influence (both promotive and adverse influences; Hutten-locher, 2009).

However, EF training isn't relegated exclusively to early childhood; rather, EF training is effective across the lifespan (e.g., Diamond and Ling, 2019). Indeed, adolescents may benefit more than younger children from specific types of EF training. For example, adolescents respond better to task-relevant feedback on EF tasks than younger children (Peters and Crone, 2017). Adolescents are also more responsive to rewards and incentives used to motivate EF training (Geier et al., 2010; Payne, 2012; Steinberg, 2008). Thus, successful EF training depends on identifying developmentally appropriate practices that leverage the neural mechanisms and the environment of the trainee.

EF training has been found to be effective, but there are questions about whether training in a specific context (or on a particular task) generalizes to untrained behaviors, also known as "far transfer." In other words, does training EF skills through a game or set of activities translate to improvements in self-regulation behavior in the classroom? The research on EF training suggests that EF skills are similar to other acquired skills (e.g., cooking), and that we improve EF skills the same way we improve our cooking skills—by practicing them (e.g., Weicker et al., 2016; Zelazo, 2015).

Similar to training other skills, training EF skills can be separated into two separate domains: training the process and training the strategy. For example, cooking processes such as knife skills are trained through repetition until they become relatively automatic procedural routines, whereas cooking strategies concern declarative knowledge regarding what techniques are required in a given situation. This might be knowing which ingredients pair well together or recognizing whether a vegetable needs to be diced, chopped, or julienned. Strategies are typically taught explicitly. Similarly, EF *processes* such as working memory capacity can be improved through repetition. Meanwhile, EF-supportive strategies, such as chunking information together to effectively extend the amount of information you can hold in working memory, or psychologically distancing yourself from a situation to improve inhibitory control (White and Carlson, 2016), are usually learned through more direct instruction. In order to effectively improve someone's cooking or EF skills, both the technique and the strategy need to be trained.

An illustration of the distinction between process and strategy can be seen in working memory training. Adaptive working memory training, where task difficulty is increased as performance improves, is associated with increases in frontal and parietal brain activity (indicating ongoing EF challenges) as well as larger improvements in working memory skills than those that result from non-adaptive training (Olesen et al., 2004). In contrast, performance following non-adaptive working memory training is associated with reductions in frontal and parietal activity, compared to activity before training (Landau et al., 2004). This pattern of results suggests that training working memory skills through repetitive practice causes these skills to become automatized. In contrast, when working memory training focuses on promoting new strategies (e.g., rehearsal, mnemonics, and chunking), then frontal and parietal brain activity remains high (Kelly and Garavan, 2005). A combination of both process and strategy training yields the greatest results (Chenault et al., 2006; Jolles and Crone, 2012).

The distinction between process and strategy can help explain why so many programs that purport to support EF skills typically find evidence for near transfer, but not far transfer, because they focus exclusively on training processes. In addition to acquiring more proficient EF skills, however, children also need to understand these skills, what they are good for, and when and how to apply them. Programs that integrate both process and strategy are particularly successful and are more likely to show evidence of far transfer (e.g., Espinet et al., 2013; Hadley et al., 2020; Pozuelos et al., 2019).

One successful method for promoting EF skills is integrating EF training into structured play activities, including video games. Although this has been studied almost exclusively in early childhood, the basic principles are the same for school-age children more generally. During play (or imaginative exploration), and in a safe, low-stakes environment, children are more likely to be motivated to practice EF skills. The structured aspect of structured play can be used by educators to train the process, by encouraging repeated practice at ever higher levels of challenge, while the social interactions in play (e.g., planning) can be used to help develop strategy. For example, an educator can set up a dramatic play scenario where children pretend to be Buckingham Palace Guards and therefore must remain very still no matter what distractions appear. By recognizing what level of support children need in this task, educators can shorten or lengthen the duration of the task in order to appropriately challenge a student's inhibitory control. When children repeatedly practice these skills at an appropriate (and increasing) level of challenge, the skills (and underlying brain networks) adapt and become stronger and more efficient. But educators can also explicitly link behaviors practiced during this play activity to other, real-world scenarios, helping children learn how to apply their newly learned skills in those scenarios. These types of structured play interactions occur in the program *Tools of the Mind* (Bodrova and Leong, 2006).

A comprehensive review of EF training programs concluded that successful programs feature adaptive challenges, are sustained for long durations (e.g., 8 weeks or year-long school curricula), involve high doses of training (e.g., 30 min or more at a time), and are spaced out over time (training occurs regularly across multiple weeks rather than rapidly in consecutive days; Diamond and Ling, 2019). Furthermore, training benefits those with lower EF skills the most, and in order to see benefits on EF, assessment should challenge the limits of children's skills (i.e., they should be relatively difficult). These conclusions apply to other skills, too. For example, one wouldn't expect a one-off cooking class in a decontextualized environment to translate to improvements in home cooking.

One way to promote strategy learning and increase the likelihood of far transfer is to include a component of meta-cognitive reflection training. Reflection is a crucial component of intentional self-regulation (Cunningham and Zelazo, 2007; Semenov et al., 2020; Zelazo, 2015) during which a mental construal of a situation or problem is considered and reconsidered. Reflection helps identify what aspects of a stimulus are important relative to the problem; it can help identify which strategies to recruit in a given problem and align those strategies in a goal directed way. Scaffolded reflection can help children succeed in EF tasks they would otherwise fail (Espinete et al., 2013).

One particularly promising method of promoting reflection in children, adolescents and adults is mindfulness. Mindfulness is a particular way of paying attention on purpose, in the present moment, and nonjudgmentally (Kabat-Zinn, 1994). During mindfulness practice, participants train their attention to focus on a particular mental object (e.g., the breath) while nonjudgmentally noting other thoughts or distractions that might arise. Focusing attention in such a way, and calmly bringing it back when it wanders, is associated with improvements in EF skills and reductions in stress in children and adults (Jha et al., 2010; Klingbeil et al., 2017; Zelazo et al., 2018; Zenner et al., 2014; Zoogman et al., 2015). It also produces increased activation in key brain regions associated with EF skills (Brefczynski-Lewis et al., 2007; Hölzel et al., 2007). Mindfulness practice also reduces the impact of emotional processes that interfere with successful EF recruitment (Semenov et al., 2020). For example, adults enrolled in a mindfulness class (versus a relaxation meditation class) showed a reduction in the extent to which negative pictures (e.g., a car crash) slowed performance on a simple cognitive task (Ortner et al., 2007). By reducing reactivity to potentially distracting emotional stimuli and also training focused attention and a sense of acceptance, mindfulness practice helps provide both the necessary processes and strategies for EF skills to develop.

Two promising contexts within which EF training is most likely to be successful are school programs and family routines. Training EF skills in the classroom, specifically integrating EF practice into structured curricular activities (e.g., *Tools of the Mind*) can increase the likelihood that improvements in EF skills will be seen in the classroom setting (Blair and Raver, 2014). The structured play activities within *Tools of the Mind* keep children motivated in their EF practice while providing them with the applied practice necessary to train EF skills. Furthermore, within the context of play, children often perform at a level above their typical ability (Vygotsky, 1967). Another program that integrates EF training along with mindfulness practice and emotion regulation is *MindUP* (Hawn Foundation, 2011). In *MindUP* classrooms, students start by learning about the science behind brain development (e.g., experience changes the brain), then engage in mindfulness practices designed to reduce stress. Next, students practice EF and emotion regulation through exercises including role play and writing. Both *Tools of the Mind* and *MindUP* have been found to produce improvements in children's EF skills on standard laboratory tasks (Diamond et al., 2007; Schonert-Reichl et al., 2015).

Another promising approach is integrating EF training into daily family routines. Family routines provide structure and predictability to a child's daily schedule, making it easier for children to engage their EF skills proactively. Furthermore, the content of routines makes EF training during routines more ecologically valid—EF skills are being used in real-world situations like getting to school on time or preparing for a family meal (for review see Semenov and Zelazo, 2019; Semenov, 2021).

Given the importance of EF skills and their capacity to be trained, there is a pressing need for more research into effective training programs, especially in childhood and adolescence. Integrating programs into schools in the form of structured play (e.g., dramatic play and challenging games) and using practices such as metacognitive reflection and mindfulness are among the most effective strategies. Similar approaches can be undertaken in at-home training programs that leverage family routines to deliver sustainable EF training; however, more research is needed to identify the best practices for parents to apply research from the lab (or the school) to the living room.

Conclusion

EF skills provide a foundation for learning and predict academic success. Hot and cool EF skills are called upon throughout the school day to solve difficult math problems, remember instructions, and manage classroom interactions. Students with better EF skills learn more from a given amount of instruction. EF skills are malleable, and they can be influenced by classroom factors, such as visual aids, and teacher behaviors, such as autonomy supportive teaching practices. Effective training of EF skills requires that these skills be continually challenged by adapting the difficulty of the exercises as the skills grow. Classroom interventions have been shown to improve students' EF skills, and an emerging body of research has started to examine the types of activities outside of school (e.g., at home) that may play a role in child EF development (e.g., Barker et al., 2014; Stucke et al., 2022). Parent autonomy support behaviors, similar to teacher autonomy support behaviors, appear to promote the development of EF skills (Distefano et al., 2018; Meuwissen and Carlson, 2019). School-family partnerships have shown to be critical for promoting other non-academic skills (e.g., social emotional learning [SEL] skills; Garbacz et al., 2015) and future research might usefully investigate intervention programs that combine school and home-based approaches.

References

- Allegretto, S.A., Mishel, L., 2016. The Teacher Pay Gap is Wider Than Ever: Teachers' Pay Continues to Fall Further Behind Pay of Comparable Workers. Economic Policy Institute.
- Ansari, A., Purcell, K.M., 2017. Activity settings in full-day kindergarten classrooms and children's early learning. *Early Child. Res. Q.* 38, 23–32.
- Bardack, S., Obradović, J., 2019. Observing teachers' displays and scaffolding of executive functioning in the classroom context. *J. Appl. Dev. Psychol.* 62, 205–219.
- Barker, J.E., Semenov, A.D., Michaelson, L., Provan, L.S., Snyder, H.R., Munakata, Y., 2014. Less-structured time in children's daily lives predicts self-directed executive functioning. *Frontiers in Psychology* 5, 593.
- Barrett, P., Davies, F., Zhang, Y., Barrett, L., 2015. The impact of classroom design on pupils' learning: final results of a holistic, multi-level analysis. *Build. Environ.* 89, 118–133.
- Ben-Gal Dahan, A., Mikulincer, M., 2020. Attachment and task persistence: attachment orientations, perception of teacher's responsiveness, and adolescents' persistence in academic tasks. *Am. J. Bioeth.* 1–22.
- Benson, J.E., Sabbagh, M.A., Carlson, S.M., Zelazo, P.D., 2013. Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Dev. Psychol.* 49 (9), 1615.
- Bernier, A., Carlson, S.M., Whipple, N., 2010. From external regulation to self-regulation: early parenting precursors of young children's executive functioning. *Child Dev.* 81 (1), 326–339.
- Berry, D., O'Connor, E., 2010. Behavioral risk, teacher–child relationships, and social skill development across middle childhood: a child-by-environment analysis of change. *J. Appl. Dev. Psychol.* 31 (1), 1–14.
- Berry, D., 2012. Inhibitory control and teacher–child conflict: reciprocal associations across the elementary-school years. *J. Appl. Dev. Psychol.* 33 (1), 66–76.
- Blair, C., Raver, C.C., 2014. Closing the achievement gap through modification of neurocognitive and neuroendocrine function: results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLoS One* 9 (11), e112393. <https://doi.org/10.1371/journal.pone.0112393>.
- Blair, C., Razza, R.P., 2007. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78 (2), 647–663.
- Blair, C., Ursache, A., Greenberg, M., Vernon-Feagans, L., 2015. Multiple aspects of self-regulation uniquely predict mathematics but not letter–word knowledge in the early elementary grades. *Dev. Psychol.* 51 (4), 459.
- Blair, C., Berry, D.J., FLP Investigators, 2017. Moderate within-person variability in cortisol is related to executive function in early childhood. *Psychoneuroendocrinology* 81, 88–95.
- Bodrova, E., Leong, D.J., 2006. *Tools of the Mind: The Vygotskian Approach to Early Childhood Education*, second ed. Pearson, Upper Saddle River, NJ.
- Braun, S.S., Zadza, K.M., Miller, A.M., Gest, S.D., 2019. Predicting elementary teachers' efforts to manage social dynamics from classroom composition, teacher characteristics, and the early year peer ecology. *Soc. Psychol. Educ.* 22 (4), 795–817.
- Brefczynski-Lewis, J.A., Lutz, A., Schaefer, H.S., Levinson, D.B., Davidson, R.J., 2007. Neural correlates of attentional expertise in long-term meditation practitioners. *Proc. Natl. Acad. Sci. U. S. A.* 104 (27), 11483–11488.
- Brock, L.L., Rimm-Kaufman, S.E., Nathanson, L., Grimm, K.J., 2009. The contributions of “hot” and “cool” executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Child. Res. Q.* 24 (3), 337–349.
- Bull, R., Espy, K.A., Wiebe, S.A., 2008. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology* 33 (3), 205–228. <https://doi.org/10.1080/87565640801982312>.
- Bull, R., Espy, K.A., Wiebe, S.A., Sheffield, T.D., Nelson, J.M., 2011. Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematics achievement. *Dev. Sci.* 14 (4), 679–692.
- Burke, R.J., Greenglass, E., 1996. Work stress, social support, psychological burnout and emotional and physical well-being among teachers. *Psychol. Health Med.* 1 (2), 193–205.
- Cadima, J., Verschueren, K., Leal, T., Guedes, C., 2016. Classroom interactions, dyadic teacher–child relationships, and self-regulation in socially disadvantaged young children. *J. Abnorm. Child Psychol.* 44 (1), 7–17.
- Chang, M.L., 2009. An appraisal perspective of teacher burnout: examining the emotional work of teachers. *Educ. Psychol. Rev.* 21 (3), 193–218.
- Chenault, B., Thomson, J., Abbott, R.D., Berninger, V.W., 2006. Effects of prior attention training on child dyslexics' response to composition instruction. *Dev. Neuropsychol.* 29, 243–260.
- Choi, J.Y., Castle, S., Williamson, A.C., Young, E., Worley, L., Long, M., Horm, D.M., 2016. Teacher–child interactions and the development of executive function in preschool-age children attending head start. *Early Educ. Dev.* 27 (6), 751–769.
- Clark, C.A., Pritchard, V.E., Woodward, L.J., 2010. Preschool executive functioning abilities predict early mathematics achievement. *Dev. Psychol.* 46 (5), 1176.
- Clark, C.A., Sheffield, T.D., Wiebe, S.A., Espy, K.A., 2013. Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Dev.* 84 (2), 662–677.
- Collie, R.J., Shapka, J.D., Perry, N.E., 2012. School climate and social–emotional learning: predicting teacher stress, job satisfaction, and teaching efficacy. *J. Educ. Psychol.* 104 (4), 1189.
- Courtier, P., Gardes, M.L., Van der Henst, J.B., Noveck, I.A., Croset, M.C., Epinat-Duclos, J., Léone, J., Prado, J., 2021. Effects of Montessori education on the academic, cognitive, and social development of disadvantaged preschoolers: a randomized controlled study in the French public-school system. *Child Dev.* 92 (5).
- Cui, T., Liu, Q., Liu, J., Wang, C., 2020. The relationship between teacher–student relationships, self-confidence, and academic achievement in the Chinese context. *Educ. New Dev.* 24–28.
- Cunningham, W.A., Zelazo, P.D., 2007. Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends Cognit. Sci.* 11 (3), 97–104.

- Cutuli, J.J., Desjardins, C.D., Herbers, J.E., Long, J.D., Heistad, D., Chan, C.-K., Hinz, E., Masten, A.S., 2013. Academic achievement trajectories of homeless and highly mobile students: resilience in the context of chronic and acute risk. *Child Dev* 84, 841–857. <https://doi.org/10.1111/cdev.12013>.
- Davis, H., Valcan, D.S., Pino-Pasternak, D., 2021. The relationship between executive functioning and self-regulated learning in Australian children. *The British Journal of Developmental Psychology* 39 (4), 625–652. <https://doi.org/10.1111/bjdp.12391>.
- de Wilde, A., Koot, H.M., van Lier, P.A., 2016. Developmental links between children's working memory and their social relations with teachers and peers in the early school years. *J. Abnorm. Child Psychol.* 44 (1), 19–30.
- Denervaud, S., Knebel, J.F., Hagmann, P., Gentaz, E., 2019. Beyond executive functions, creativity skills benefit academic outcomes: insights from Montessori education. *PLoS One* 14 (11), e0225319.
- Diamond, A., Ling, D.S., 2019. Review of the evidence on, and fundamental questions about, efforts to improve executive functions, including working memory. In: Bunting, M., Novick, J., Dougherty, M., Engle, R.W. (Eds.), *Cognitive and Working Memory Training: Perspectives From Psychology, Neuroscience, and Human Development*. Oxford University Press, New York. <https://doi.org/10.1093/oso/9780199974467.001.0001>.
- Diamond, A., Barnett, W.S., Thomas, J., Munro, S., 2007. Preschool program improves cognitive control. *Science* 318 (5855), 1387.
- Diliberti, M.K., Schwartz, H.L., Grant, D.M., 2021. Stress Topped the Reasons Why Public School Teachers Quit, Even Before COVID-19. RAND.
- Distefano, R., Galinsky, E., McClelland, M.M., Zelazo, P.D., Carlson, S.M., 2018. Autonomy-supportive parenting and associations with child and parent executive function. *J. Appl. Dev. Psychol.* 58, 77–85.
- Duckworth, A.L., Seligman, M.E., 2005. Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychol. Sci.* 16 (12), 939–944.
- Efenev, G., Carroll, A., Bahr, N., 2013. Self-regulated learning and executive function: Exploring the relationships in a sample of adolescent males. *Educational Psychology* 33 (7), 773–796. <https://doi.org/10.1080/01443410.2013.785054>.
- Espinet, S.D., Anderson, J.E., Zelazo, P.D., 2013. Reflection training improves executive function in preschool-age children: behavioral and neural effects. *Dev. Cognit. Neurosci.* 4, 3–15.
- Espy, K.A., McDiarmid, M.M., Cwik, M.F., Stalets, M.M., Hamby, A., Senn, T.E., 2004. The contribution of executive functions to emergent mathematic skills in preschool children. *Dev. Neuropsychol.* 26 (1), 465–486.
- Finch, J.E., 2019. Do schools promote executive functions? Differential working memory growth across school-year and summer months. *AERA Open* 5 (2).
- Fisher, A.V., Godwin, K.E., Seltman, H., 2014. Visual environment, attention allocation, and learning in young children: when too much of a good thing may be bad. *Psychol. Sci.* 25 (7), 1362–1370.
- Fitzpatrick, C., McKinnon, R.D., Blair, C.B., Willoughby, M.T., 2014. Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learn. Instruct.* 30, 25–31.
- Follmer, D.J., Sperling, R.A., 2016. The mediating role of metacognition in the relationship between executive function and selfregulated learning. *Br. J. Educ. Psychol.* 86 (4), 559–575.
- Fuhs, M.W., Nesbitt, K.T., Farran, D.C., Dong, N., 2014. Longitudinal associations between executive functioning and academic skills across content areas. *Dev. Psychol.* 50 (6), 1698.
- Garbacz, S.A., Swanger-Gagné, M.S., Sheridan, S.M., 2015. The role of school-family partnership programs for promoting student SEL. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), *Handbook of Social and Emotional Learning: Research and Practice*. Guilford Press, pp. 244–259.
- Geier, C.F., Tervilliger, R., Teslovich, T., Velanova, K., Luna, B., 2010. Immaturities in reward processing and its influence on inhibitory control in adolescence. *Cerebr. Cortex* 20 (7), 1613–1629.
- Gest, S.D., Welsh, J.A., Domitrovich, C.E., 2005. Behavioral predictors of changes in social relatedness and liking school in elementary school. *J. Sch. Psychol.* 43 (4), 281–301.
- Goble, P., Pianta, R.C., 2017. Teacher-child interactions in free choice and teacher-directed activity settings: Prediction to school readiness. *Early Education and Development* 28 (8), 1035–1051. <https://doi.org/10.1080/10409289.2017.1322449>.
- Grayson, J.L., Alvarez, H.K., 2008. School climate factors relating to teacher burnout: a mediator model. *Teach. Teach. Educ.* 24 (5), 1349–1363.
- Grenell, A., Carlson, S.M., 2021. Individual differences in executive function and learning: the role of knowledge type and conflict with prior knowledge. *J. Exp. Child Psychol.* 206, 105079.
- Grolnick, W.S., Pomerantz, E.M., 2009. Issues and challenges in studying parental control: toward a new conceptualization. *Child Dev. Perspect.* 3 (3), 165–170.
- Gropen, J., Clark-Chiarelli, N., Hoisington, C., Ehrlich, S.B., 2011. The importance of executive function in early science education. *Child Dev. Perspect.* 5 (4), 298–304.
- Hadley, L.V., Acluche, F., Chevalier, N., 2020. Encouraging performance monitoring promotes proactive control in children. *Developmental Science* 23, e12861. <https://doi.org/10.1111/desc.12861>.
- Hanley, M., Khairat, M., Taylor, K., Wilson, R., Cole-Fletcher, R., Riby, D.M., 2017. Classroom displays—attraction or distraction? Evidence of impact on attention and learning from children with and without autism. *Dev. Psychol.* 53 (7), 1265.
- Hanson, J.L., Chung, M.K., Avants, B.B., Shitcliff, E.A., Gee, J.C., Davidson, R.J., Pollak, S.D., 2010. Early stress is associated with alterations in the orbitofrontal cortex: a tensor-based morphometry investigation of brain structure and behavioral risk. *J. Neurosci.* 30 (22), 7466–7472.
- Hassinger-Das, B., Jordan, N.C., Glutting, J., Irwin, C., Dyson, N., 2014. Domain-general mediators of the relation between kindergarten number sense and first-grade mathematics achievement. *Journal of Experimental Child Psychology* 118, 78–92. <https://doi.org/10.1016/j.jecp.2013.09.008>.
- Hawn Foundation, 2011. *The MindUp Curriculum: Brain-Focused Strategies for Learning—and Living*. Scholastic Teaching Resources, New York.
- Hebb, D.O., 1949. *The Organization of Behavior: a Neuropsychological Theory*. Psychology Press.
- Hiver, P., Whiteside, Z., Sánchez Solarte, A.C., Kim, C.J., 2021. Language teacher metacognition: beyond the mirror. *Innovat. Lang. Learn. Teach.* 15 (1), 52–65.
- Hoglund, W.L., Klinge, K.E., Hosan, N.E., 2015. Classroom risks and resources: teacher burnout, classroom quality and children's adjustment in high needs elementary schools. *J. Sch. Psychol.* 53 (5), 337–357.
- Hölzel, B.K., Ott, U., Hempel, H., Hackl, A., Wolf, K., Stark, R., Vaitl, D., 2007. Differential engagement of anterior cingulate and adjacent medial frontal cortex in adept meditators and non-meditators. *Neurosci. Lett.* 421 (1), 16–21.
- Hu, Y., Geng, F., Tao, L., Hu, N., Du, F., Fu, K., Chen, F., 2011. Enhanced white matter tracts integrity in children with abacus training. *Hum. Brain Mapp.* 32 (1), 10–21.
- Huttenlocher, P.R., 2009. *Neural Plasticity*. Harvard University Press.
- Jha, A.P., Stanley, E.A., Kiyonaga, A., Wong, L., Gelfand, L., 2010. Examining the protective effects of mindfulness training on working memory capacity and affective experience. *Emotion* 10 (1), 54.
- Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., Millet, C., 2005. The experience of work-related stress across occupations. *Journal of Managerial Psychology* 20 (2), 178–187. <https://doi.org/10.1108/02683940510579803>.
- Jolles, D., Crone, E.A., 2012. Training the developing brain: a neurocognitive perspective. *Front. Hum. Neurosci.* 6, 76.
- Kabat-Zinn, J., 1994. *Wherever You Go, There You Are*. Hyperion, New York.
- Kelly, A.C., Garavan, H., 2005. Human functional neuroimaging of brain changes associated with practice. *Cerebr. Cortex* 15 (8), 1089–1102.
- Kieffer, M.J., Vukovic, R.K., Berry, D., 2013. Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Read. Res. Q.* 48 (4), 333–348.
- Kikas, E., Tang, X., 2019. Child-perceived teacher emotional support, its relations with teaching practices, and task persistence. *Eur. J. Psychol. Educ.* 34 (2), 359–374.
- Klingbeil, D.A., Fischer, A.J., Renshaw, T.L., Bloomfield, B.S., Polakoff, B., Willenbrink, J.B., et al., 2017. Effects of mindfulness-based interventions on disruptive behavior: a meta-analysis of single-case research. *Psychol. Sch.* 54 (1), 70–87.
- Kolb, B., Harker, A., Gibb, R., 2017. Principles of plasticity in the developing brain. *Dev. Med. Child Neurol.* 59, 1218–1223. <https://doi.org/10.1111/dmcn.13546>.

- Koopman-Verhoeff, M.E., Serdarevic, F., Kocevskaja, D., Bodrij, F.F., Mileva-Seitz, V.R., Reiss, I., et al., 2019. Preschool family irregularity and the development of sleep problems in childhood: a longitudinal study. *JCPP* 60 (8), 857–865.
- Korucu, I., Rolan, E., Napoli, A.R., Purpura, D.J., Schmitt, S.A., 2019. Development of the Home Executive Function Environment (HEFE) scale: assessing its relation to preschoolers' executive function. *Early Child. Res. Q.* 47, 9–19.
- Košir, K., Tement, S., 2014. Teacher–student relationship and academic achievement: a cross-lagged longitudinal study on three different age groups. *Eur. J. Psychol. Educ.* 29 (3), 409–428.
- Landau, S.M., Schumacher, E.H., Garavan, H., Druzgal, T.J., D'Esposito, M., 2004. A functional MRI study of the influence of practice on component processes of working memory. *NeuroImage* 22 (1), 211–221.
- Larson, M., Cook, C.R., Fiat, A., Lyon, A.R., 2018. Stressed teachers don't make good implementers: examining the interplay between stress reduction and intervention fidelity. *Sch. Ment. Health* 10 (1), 61–76.
- Lawson, G.M., Farah, M.J., 2017. Executive function as a mediator between SES and academic achievement throughout childhood. *IJBD* 41 (1), 94–104.
- Lawson, G.M., Hook, C.J., Farah, M.J., 2018. A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Dev. Sci.* 21 (2), e12529.
- Li, J.B., Leung, I.T.Y., Li, Z., 2021. The pathways from self-control at school to performance at work among novice kindergarten teachers: the mediation of work engagement and work stress. *Child. Youth Serv. Rev.* 121, 105881.
- Lillard, A., Else-Quest, N., 2006. The early years: evaluating Montessori education. *Science* 313 (5795), 1893–1894.
- Lisonbee, J.A., Mize, J., Payne, A.L., Granger, D.A., 2008. Children's cortisol and the quality of teacher–child relationships in child care. *Child Dev.* 79 (6), 1818–1832.
- Lupien, S.J., McEwen, B.S., Gunnar, M.R., Heim, C., 2009. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat. Rev. Neurosci.* 10 (6), 434–445.
- Madigan, D.J., Kim, L.E., 2021. Does teacher burnout affect students? A systematic review of its association with academic achievement and student-reported outcomes. *Int. J. Educ. Res.* 105, 101714.
- Marcovitch, S., Jacques, S., Boseovski, J.J., Zelazo, P.D., 2008. Self-reflection and the cognitive control of behavior: implications for learning. *Mind Brain Educ.* 2 (3), 136–141.
- Martin, M.J., Davies, P.T., Cummings, E.M., Cicchetti, D., 2017. The mediating roles of cortisol reactivity and executive functioning difficulties in the pathways between childhood histories of emotional insecurity and adolescent school problems. *Dev. Psychopathol.* 29 (4), 1483–1498.
- Masten, A.S., Herbers, J.E., Desjardins, C.D., et al., 2012. Executive function skills and school success in young children experiencing homelessness. *Rev. Educ. Res.* 41 (9), 93–124. <https://doi.org/10.3102/0034654318815707>.
- Masten, A.S., 2013. Risk and Resilience in Development.
- Matheny Jr., A.P., Wachs, T.D., Ludwig, J.L., Phillips, K., 1995. Bringing order out of chaos: psychometric characteristics of the confusion, hubbub, and order scale. *J. Appl. Dev. Psychol.* 16 (3), 429–444.
- McClelland, M.M., Acocck, A.C., Morrison, F.J., 2006. The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly* 21 (4), 471–490. <https://doi.org/10.1016/j.ecresq.2006.09.003>.
- McClelland, M.M., Cameron, C.E., Connor, C.M., Farris, C.L., Jewkes, A.M., Morrison, F.J., 2007. Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Dev. Psychol.* 43 (4), 947.
- McClelland, M.M., Acocck, A.C., Piccinin, A., Rhea, S.A., Stallings, M.C., 2013. Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Child. Res. Q.* 28 (2), 314–324.
- McKinnon, R.D., Blair, C., 2019. Bidirectional relations among executive function, teacher–child relationships, and early reading and math achievement: a cross-lagged panel analysis. *Early Child. Res. Q.* 46, 152–165.
- McLean, L., Connor, C.M., 2015. Depressive symptoms in third-grade teachers: relations to classroom quality and student achievement. *Child Dev.* 86 (3), 945–954.
- Merritt, E.G., Wanless, S.B., Rimm-Kaufman, S.E., Cameron, C., Peugh, J.L., 2012. The contribution of teachers' emotional support to children's social behaviors and self-regulatory skills in first grade. *Sch. Psychol. Rev.* 41 (2), 141–159.
- Merz, E.C., Desai, P.M., Maskus, E.A., Melvin, S.A., Rehman, R., Torres, S.D., et al., 2019. Socioeconomic disparities in chronic physiologic stress are associated with brain structure in children. *Biol. Psychiatr.* 86 (12), 921–929.
- Meuwissen, A.S., Carlson, S.M., 2019. An experimental study of the effects of autonomy support on preschoolers' self-regulation. *J. Appl. Dev. Psychol.* 60, 11–23.
- Mischel, W., Shoda, Y., Rodriguez, M.I., 1989. Delay of gratification in children. *Science* 244 (4907), 933–938.
- Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology* 41 (1), 49–100.
- Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., et al., 2011. A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U. S. A.* 108 (7), 2693–2698.
- Monette, S., Bigras, M., Guay, M.C., 2011. The role of the executive functions in school achievement at the end of Grade 1. *J. Exp. Child Psychol.* 109 (2), 158–173.
- Nayfeld, I., Fuccillo, J., Greenfield, D.B., 2013. Executive functions in early learning: extending the relationship between executive functions and school readiness to science. *Learn. Indiv. Differ.* 26, 81–88.
- Nesbitt, K.T., Baker-Ward, L., Willoughby, M.T., 2013. Executive function mediates socio-economic and racial differences in early academic achievement. *Early Child. Res. Q.* 28 (4), 774–783.
- Neuenschwander, R., Friedman-Krauss, A., Raver, C., Blair, C., 2017. Teacher stress predicts child executive function: moderation by school poverty. *Early Educ. Dev.* 28 (7), 880–900.
- Noble, K.G., Houston, S.M., Brito, N.H., Bartsch, H., Kan, E., Kuperman, J.M., et al., 2015. Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.* 18 (5), 773–778.
- Obradović, J., Long, J.D., Cutuli, J.J., Chan, C.K., Hinz, E., Heistad, D., Masten, A.S., 2009. Academic achievement of homeless and highly mobile children in an urban school district: longitudinal evidence on risk, growth, and resilience. *Dev. Psychopathol.* 21 (2), 493–518.
- Olesen, P.J., Westerberg, H., Klingberg, T., 2004. Increased prefrontal and parietal activity after training of working memory. *Nat. Neurosci.* 7 (1), 75–79.
- Ortner, C.N., Kilner, S.J., Zelazo, P.D., 2007. Mindfulness meditation and reduced emotional interference on a cognitive task. *Motiv. Emot.* 31 (4), 271–283.
- Patall, E.A., Zambrano, J., 2019. Facilitating student outcomes by supporting autonomy: implications for practice and policy. *Pol. Insights Behav. Brain Sci.* 6 (2), 115–122.
- Payne, M.A., 2012. “All gas and no brakes!” Helpful metaphor or harmful stereotype? *J. Adolesc. Res.* 27 (1), 3–17. <https://doi.org/10.1177/0743558411412956>.
- Peters, S., Crone, E.A., 2017. Increased striatal activity in adolescence benefits learning. *Nat. Commun.* 8 (1), 1–9.
- Piccolo, L.R., Merz, E.C., Noble, K.G., Pediatric Imaging, Neurocognition, and Genetics Study, 2019. School climate is associated with cortical thickness and executive function in children and adolescents. *Dev. Sci.* 22 (1), e12719.
- Poormahmood, A., Moayedi, F., Alizadeh, K.H., 2017. Relationships between psychological well-being, happiness and perceived occupational stress among primary school teachers. *Arch. Hell. Med./Arheia Ellenikes Iatrikes* 34 (4).
- Pozuelos, J.P., Combata, L.M., Abundis, A., Paz-Alonso, P.M., Conejero, A., Guerra, S., Rueda, M.R., 2019. Metacognitive scaffolding boosts cognitive and neural benefits following executive attention training in children. *Developmental Science* 22, e12756. <https://doi.org/10.1111/desc.12756>.
- Prager, E.O., Sera, M.D., Carlson, S.M., 2016. Executive function and magnitude skills in preschool children. *J. Exp. Child Psychol.* 147, 126–139.
- Purpura, D.J., Schmitt, S.A., Ganley, C.M., 2017. Foundations of mathematics and literacy: the role of executive functioning components. *J. Exp. Child Psychol.* 153, 15–34.
- Raver, C.C., Jones, S.M., Li-Grining, C., Zhai, F., Bub, K., Pressler, E., 2011. CSRP's impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child Dev.* 82 (1), 362–378.

- Reardon, S.F., 2013. The widening income achievement gap. *Educ. Leader*. 70 (8), 10–16.
- Reeve, J., Cheon, S.H., 2021. Autonomy-supportive teaching: its malleability, benefits, and potential to improve educational practice. *Educ. Psychol.* 56 (1), 54–77.
- Reinke, W.M., Herman, K.C., Stormont, M., 2013. Classroom-level positive behavior supports in schools implementing SW-PBIS: identifying areas for enhancement. *J. Posit. Behav. Interv.* 15 (1), 39–50.
- Reyes, M.R., Brackett, M.A., Rivers, S.E., White, M., Salovey, P., 2012. Classroom emotional climate, student engagement, and academic achievement. *J. Educ. Psychol.* 104 (3), 700.
- Ribner, A.D., Willoughby, M.T., Blair, C.B., Family Life Project Key Investigators, 2017. Executive function buffers the association between early math and later academic skills. *Front. Psychol.* 8, 869.
- Ribner, A.D., 2020. Executive function facilitates learning from math instruction in kindergarten: evidence from the ECLS-K. *Learn. Instruct.* 65, 101251.
- Richmond-Rakerd, L.S., Caspi, A., Ambler, A., d'Arbeloff, T., de Bruine, M., Elliott, M., et al., 2021. Childhood self-control forecasts the pace of midlife aging and preparedness for old age. *Proc. Natl. Acad. Sci. U. S. A.* 118 (3). <https://doi.org/10.1073/pnas.2010211118> e2010211118.
- Rimm-Kaufman, S.E., Curby, T.W., Grimm, K.J., Nathanson, L., Brock, L.L., 2009. The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Dev. Psychol.* 45 (4), 958.
- Rimm-Kaufman, S.E., Larsen, R.A., Baroody, A.E., Curby, T.W., Ko, M., Thomas, J.B., Merritt, E.G., Abry, T., DeCoster, J., 2014. Efficacy of the responsive classroom approach: results from a 3-year, longitudinal randomized controlled trial. *Am. Educ. Res. J.* 51 (3), 567–603.
- Rodrigues, P.F., Pandeirada, J.N., 2018. When visual stimulation of the surrounding environment affects children's cognitive performance. *J. Exp. Child Psychol.* 176, 140–149.
- Rodrigues, P.F., Pandeirada, J.N., 2019. The influence of a visually-rich surrounding environment in visuospatial cognitive performance: a study with adolescents. *J. Cognit. Dev.* 20 (3), 399–410.
- Ronfeldt, M., Loeb, S., Wyckoff, J., 2013. How teacher turnover harms student achievement. *Am. Educ. Res. J.* 50 (1), 4–36.
- Roorda, D.L., Jak, S., Zee, M., Oort, F.J., Koomen, H.M., 2017. Affective teacher–student relationships and students' engagement and achievement: a meta-analytic update and test of the mediating role of engagement. *Sch. Psychol. Rev.* 46 (3), 239–261.
- Rutherford, T., Buschkuhl, M., Jaeggi, S.M., Farkas, G., 2018. Links between achievement, executive functions, and self-regulated learning. *Applied Cognitive Psychology* 32 (6), 763–774. <https://doi.org/10.1002/acp.3462>.
- Schmitt, S.A., McClelland, M.M., Tominney, S.L., Acock, A.C., 2015. Strengthening school readiness for head start children: evaluation of a self-regulation intervention. *Early Child. Res. Q.* 30, 20–31.
- Schonert-Reichl, K.A., Oberle, E., Lawlor, M.S., Abbott, D., Thomson, K., Oberlander, T.F., Diamond, A., 2015. Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a randomized controlled trial. *Dev. Psychol.* 51 (1), 52.
- Schuitema, J., Peetsma, T., van der Veen, I., 2016. Longitudinal relations between perceived autonomy and social support from teachers and students' self-regulated learning and achievement. *Learn. Individ. Differ.* 49, 32–45.
- Semenov, A.D., Zelazo, P.D., 2019. Mindful family routines and the cultivation of executive function skills in childhood. *Hum. Dev.* 63 (2), 112–131.
- Semenov, A.D., Kennedy, D., Zelazo, P.D., 2020. Mindfulness and executive function: implications for learning and early childhood education. In: Thomas, M.S.C., Dumontheil, I., Mareschal, D. (Eds.), *Educational Neuroscience*. Routledge, New York, pp. 298–331.
- Semenov, A.D., 2021. Ready4Routines: Improving Child Executive Function Skills Through Autonomy-Supportive Parent-Child Reflective Routines. Doctoral dissertation, University of Minnesota. ProQuest Dissertation Publishing.
- Shonkoff, J.P., Garner, A.S., Siegel, B.S., Dobbins, M.I., Earls, M.F., McGuinn, L., et al., 2012. The lifelong effects of early childhood adversity and toxic stress. *Pediatrics* 129 (1).
- Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., Dochy, F., 2009. The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *Br. J. Educ. Psychol.* 79 (1), 57–68.
- Sosic-Vasic, Z., Keis, O., Lau, M., Spitzer, M., Streb, J., 2015. The impact of motivation and teachers' autonomy support on children's executive functions. *Front. Psychol.* 6, 146.
- St Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology* 59 (4), 745–759. <https://doi.org/10.1080/17470210500162854>.
- Steiner, E.D., Woo, A., 2021. Job-related stress threatens the teacher supply: Key findings from the 2021 State of the U.S. Teacher Survey. Research Report. RAND ©. www.rand.org/pubs/research_reports/RRA1108-1.html.
- Steinberg, L., 2008. A social neuroscience perspective on adolescent risk-taking. *Dev. Rev.* 28 (1), 78–106.
- Steinhardt, M.A., Smith Jaggars, S.E., Faulk, K.E., Gloria, C.T., 2011. Chronic work stress and depressive symptoms: assessing the mediating role of teacher burnout. *Stress Health* 27 (5), 420–429.
- Stevenson, N.A., VanLone, J., Barber, B.R., 2020. A commentary on the misalignment of teacher education and the need for classroom behavior management skills. *Educ. Treat. Child.* 43 (4), 393–404.
- Stucke, N.J., Stoet, G., Doebel, S., 2022. What are the kids doing? Exploring young children's activities at home and relations with externally cued executive function and child temperament. *Dev. Sci.* <https://doi.org/10.1111/desc.13226>.
- Suleman, Q., Hussain, I., Shehzad, S., Syed, M.A., Raja, S.A., 2018. Relationship between perceived occupational stress and psychological well-being among secondary school heads in Khyber Pakhtunkhwa, Pakistan. *PLoS One* 13 (12), e0208143.
- Suntheimer, N.M., Wolf, S., 2020. Cumulative risk, teacher-child closeness, executive function and early academic skills in kindergarten children. *Journal of School Psychology* 78, 23–37. <https://doi.org/10.1016/j.jsp.2019.11.005>.
- van Hout-Wolters, B., Simons, R.J., Volet, S., 2000. Active learning: self-directed learning and independent work. In: *New Learning*. Springer, Dordrecht, pp. 21–36.
- Vandenbroucke, L., Spilt, J., Verschueren, K., Baeyens, D., 2017. Keeping the spirits up: the effect of teachers' and parents' emotional support on children's working memory performance. *Front. Psychol.* 8, 512.
- Verdine, B.N., Irwin, C.M., Golinkoff, R.M., Hirsh-Pasek, K., 2014. Contributions of executive function and spatial skills to preschool mathematics achievement. *J. Exp. Child Psychol.* 126, 37–51.
- Viterbori, P., Usai, M.C., Traverso, L., De Franchis, V., 2015. How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: a longitudinal study. *J. Exp. Child Psychol.* 140, 38–55.
- Vliegenthart, J., Noppe, G., Van Rossum, E.F.C., Koper, J.W., Raat, H., Van den Akker, E.L.T., 2016. Socioeconomic status in children is associated with hair cortisol levels as a biological measure of chronic stress. *Psychoneuroendocrinology* 65, 9–14.
- Vygotsky, L.S., 1967. Play and its role in the mental development of the child. *Sov. Psychol.* 5 (3), 6–18.
- Weicker, J., Villringer, A., Thöne-Otto, A., 2016. Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology* 30 (2), 190.
- Welsh, J.A., Nix, R.L., Blair, C., Bierman, K.L., Nelson, K.E., 2010. The development of cognitive skills and gains in academic school readiness for children from low-income families. *J. Educ. Psychol.* 102 (1), 43.
- Whipple, N., Bernier, A., Mageau, G.A., 2011. Broadening the study of infant security of attachment: maternal autonomy-support in the context of infant exploration. *Soc. Dev.* 20 (1), 17–32.
- Whitaker, R.C., Dearth-Wesley, T., Gooze, R.A., 2015. Workplace stress and the quality of teacher–children relationships in head start. *Early Child. Res. Q.* 30, 57–69.
- White, R.E., Carlson, S.M., 2016. What would Batman do? Self-distancing improves executive function in young children. *Dev. Sci.* 19 (3), 419–426.
- Willoughby, M., Kupersmidt, J., Voegler-Lee, M., Bryant, D., 2011. Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Dev. Neuropsychol.* 36 (2), 162–180.

- Zaitchik, D., Iqbal, Y., Carey, S., 2014. The effect of executive function on biological reasoning in young children: an individual differences study. *Child Dev.* 85 (1), 160–175.
- Zelazo, P.D., Carlson, S.M., 2012. Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, pp. 354–360. <https://doi.org/10.1111/j.1750-8606.2012.00246.x>.
- Zelazo, P.D., Carlson, S.M., 2020. The neurodevelopment of executive function skills: implications for academic achievement gaps. *Psychol. Neurosci.* 13 (3), 273.
- Zelazo, P.D., Müller, U., 2002. Executive functions in typical and atypical development. In: Goswami, U. (Ed.), *Handbook of childhood cognitive development*. Blackwell, Oxford, pp. 445–469.
- Zelazo, P.D., Anderson, J.E., Richler, J., Wallner-Allen, K., Beaumont, J.L., Weintraub, S., 2013. II. NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *Monogr. Soc. Res. Child Dev.* 78 (4), 16–33.
- Zelazo, P.D., Blair, C.B., Willoughby, M.T., 2016. Executive function: Implications for education. U.S. Department of Education, pp. 1–148. Retrieved from. <https://ies.ed.gov/ncer/pubs/20172000/pdf/20172000.pdf>.
- Zelazo, P.D., Forston, J.L., Masten, A.S., Carlson, S.M., 2018. Mindfulness plus reflection training: effects on executive function in early childhood. *Front. Psychol.* 9, 208.
- Zelazo, P.D., 2015. Executive function: reflection, iterative reprocessing, complexity, and the developing brain. *Dev. Rev.* 38, 55–68.
- Zelazo, P.D., 2020. Executive function and psychopathology: a neurodevelopmental perspective. *Annu. Rev. Clin. Psychol.* 16, 431–454.
- Zenner, C., Hermleben-Kurz, S., Walach, H., 2014. Mindfulness-based interventions in schools—a systematic review and meta-analysis. *Front. Psychol.* 5, 603.
- Zimmerman, B.J., 2008. Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *Am. Educ. Res. J.* 45 (1), 166–183.
- Zoogman, S., Goldberg, S.B., Hoyt, W.T., Miller, L., 2015. Mindfulness interventions with youth: a meta-analysis. *Mindfulness* 6 (2), 290–302.

Teaching expertise: an activity system's lens

Peter Reimann and Lina Markauskaite, The University of Sydney, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	179
Exceptional performance: from experience and routine expertise to deliberate and adaptive development of personal capabilities	180
Deliberate practice: how to become an expert	180
Adaptive expertise: how to stay an expert	181
Knowledge building: from acquisition to construction of expert knowledge	182
Collective expertise: from individual excellence to networked and relational expertise	183
Networked expertise	183
Relational expertise	183
The ecologies of expertise: from fragmented to integrated expertise	184
The ecologies of individual expertise	184
Multilevel conceptualizations of expertise	185
Multiplicity of expertises	185
Synthesis and methodological implications: expertise through an activity system's lens	185
Acknowledgment	187
References	187

Introduction

If the experience of teaching were a valid indicator of expertise, then everybody who went through 12 years of schooling would be an expert in teaching: Most of us have experienced many hours of teaching as students in classrooms. Of course, the experience of teaching is not identical with *practicing* teaching. The 10,000 h of practice deemed necessary to develop expertise (based on early studies on chess playing; Gobet and Charness, 2006) would need to take the form of active teaching and deliberate learning. The number is important for two other reasons. Firstly, anybody who studies to become a teacher enters teacher education equipped with a 12-year experience of being taught, which is well above 10,000 h. Secondly, many who went through K-12 assume that they know what teaching—good and bad—entails. Nobody has had anything even close to this experience before starting their working life for many other professions or vocations, such as medicine, accounting, law, plumbing, hairdressing, or even cooking.

Further, many adults and even kids have at least some experience of *doing* teaching: parents teach their children, kids teach their peers, working adults teach their work colleagues, etc. Informal, often spontaneous, teaching experiences permeate human lives (Sterelny, 2012). Thus, teaching is also a universal cultural practice rather than just a specialist area of expertise: It is widely available to everybody as a part of the culture, implicitly learned, and often accomplished independently.

As a cultural activity, teaching comes with a set of routines that are recognized by members of a culture as appropriate and accountable and which tend to be relatively stable over time. This has the consequence that one can be a “good teacher” by sticking to teaching routines that are not necessarily the best. It also restricts teachers’ space for changing their routines. Parents may become worried when their children are taught in ways that they have not experienced themselves; they may not recognize the new practice as constituting appropriate teaching.

Research on teaching and teacher expertise suggests that the image of the 21st-century teacher needs to be revised. There are at least four clear reasons. First, good teaching is not identical with the mastery of stable teaching routines. Expert teachers are adaptive experts who deliberately and continuously develop their professional knowledge and skills (Bohle Carbonell et al., 2014; Ericsson et al., 1993). Secondly, teaching is not an individual practice; instead, it is a practice accomplished by a complex activity system. This activity system comprises a range of social agents (e.g., parents, students, teachers, principals, community stakeholders) and symbolic and material resources (e.g., curriculum documents, policies, teaching and learning resources, physical and digital learning spaces), all interacting with each other on multiple levels—individual, organizational, and societal (Davis, 2008). Thirdly, the teaching profession is also changing along with societal trends: professional knowledge in teaching is becoming more specialized, intertwined with theory, requiring abilities to produce various professional knowledge products and engage in a range of professional knowledge practices. In short, teaching is an expert knowledge culture (Markauskaite and Goodyear, 2017; Nerland, 2012). Finally, like other professionals, teachers increasingly face the need to solve problems that require expertise from different domains, such as teaching students with special needs and using technological tools. Therefore, they need the expertise to work with people outside their expert knowledge culture (e.g., school counselors, technology experts, parents, future employers, and other community members). Taken together, an expert teacher needs to be highly skilled, deliberate, adaptive, capable of contributing to the advancement of professional knowledge, and relationally competent.

A typical handbook chapter on expertise would focus on the question: What makes an expert? (Reimann and Markauskaite, 2018). The question directs attention to the personal capabilities that an individual possesses. Encyclopedias agree that an expert is “a person with a high level of knowledge or skill in a field” (Merriam-Webster Thesaurus, emphasis added). In this chapter, we want to avoid this possessive person-centered framing of *teacher expertise*. Building on Stigler and Miller (2018), we prefer to speak of *teaching expertise*, thereby affording the view that expert teaching is a practice associated with a *system* that is larger than a single teacher and involves *activity*—what teachers do.

In primary and secondary classrooms, as Stigler and Miller (2018) describe, teaching can be defined as consisting of “the interactions of teachers, students, and content, in classrooms, that are intended to achieve some goal or goals for students, within a specified time (e.g., a classroom lesson), together with the planning that takes place before, and the analysis that takes place after, the lesson” (p. 433). This definition draws on Lampert’s (2003) suggestion that classroom teaching is mainly relational work among the three core elements of a lesson: (1) teacher, (2) students, and (3) content. The teacher’s primary task is to manage the relationship between students and content over an extended time. Even this core system is a complex system as it is composed of multiple actors, dynamic and emergent. It becomes even more complex when we consider all the other necessary elements: (4) learning goals, (5) lesson activities, (6) learning tools, and (7) the timing and sequencing of activities. Further, teacher work is not confined to classroom teaching alone; activities such as curriculum and lesson planning involve many other “actors” who are each their own “system”. While classroom teaching is often seen as largely individual work, it is to the extreme dependent on factors outside of the individual teacher’s control and involves activities beyond the classroom. Our review of teaching expertise in this chapter embraces this expansive, activity-oriented notion of teaching.

The rest of the chapter is structured as follows: section **Exceptional performance: from experience and routine expertise to deliberate and adaptive development of personal capabilities** focuses on what teachers bring to the activity system and how they develop expert capabilities for classroom teaching. In sections **Knowledge building: from acquisition to construction of expert knowledge** and **Collective expertise: from individual excellence to networked and relational expertise**, the focus switches from teaching as the central object of expert teacher activity to the creation of knowledge relevant for teaching. While section **Knowledge building: from acquisition to construction of expert knowledge** concentrates on the kinds of knowledge that expert teachers produce, section **Collective expertise: from individual excellence to networked and relational expertise** extends this by focusing on the relational aspects of teacher expertise. Section **The ecologies of expertise: from fragmented to integrated expertise** expands this further by reviewing some ecological qualities of expertise. The final section considers how these ideas could be brought together and discusses methodological implications.

Exceptional performance: from experience and routine expertise to deliberate and adaptive development of personal capabilities

Building on the definition of classroom teaching provided above, a teacher needs to be proficient in four categories of activities: (1) formulation of learning goals, (2) selection and implementation of instructional strategies, (3) creation of learning opportunities, and (4) assessment of student progress (Stigler and Miller, 2018). All of these require knowledge, skill, and judgment and need to be performed under more or less strict time limits. The body of knowledge necessary for proficient teaching is multidimensional, comprising content knowledge, general and curriculum-specific pedagogical knowledge, and knowledge about learners and their characteristics (Shulman, 1987).

As is the case for other domains of expertise, teaching expertise requires integrating these bodies of knowledge into a performance-oriented structure (Chi et al., 1981). The differences in knowledge, skills, and judgment between novices and experts have been the subject of by now classical expertise research, documented in seminal works (Ericsson et al., 2006, 2018). The gist of this research is that experts do not only have more knowledge, but they also have it organized differently from novices, and the routine parts of it are partially or fully automatized. That is, some of the experts’ knowledge has become *tacit*; it enables perception and action without rising to the level of conscious control (Leinhardt, 1989).

Deliberate practice: how to become an expert

Looking at teacher expertise through the lens of an activity system means that we are looking at an *activity* and *system* that are different from classroom teaching: the object of which is teachers’ learning and expertise. The adage that “practice makes perfect” is scientifically supported by the observed relation between practice time and time for a particular task (the speed by which the task is executed.) The power law of practice states that the time needed for a particular task decreases logarithmically with the number of practice trials taken. This means, the improvement of performance (measured in time needed) is fast initially but slows down over time. Most explanations for this relation build on mechanisms of automatization (Fitts and Posner, 1967) and the idea that something is slowing down the learning process. For instance, we can conceive of learning from practice by initially mastering isolated sub-skills—just the serve in tennis—and later requiring the integration of sub-skills.

Although such theories can explain how somebody becomes proficient, it does not account for expertise. Experience is not identical with expertise; it may be necessary but not sufficient to achieve top levels of performance. Somewhat counter to intuition, experts resist routines (Crawford and Brophy, 2006). This has been known for years, since Ericsson et al. (1993) published their

results on the most effective forms of training in learning to play an instrument at a top level of performance. They, drawing on their results, called these most effective forms that resist routine “deliberate practice.” Ericsson and colleagues described:

At music academies master teachers provide students with individualized instruction and help them identify goals and methods for their practice sessions between meetings—this form of solitary practice was named deliberate practice

Ericsson and Harwell (2019, p. 1).

This concept engendered a considerable amount of research, which, however, has been mostly on forms of practice other than the deliberate one (i.e., single practice guided by an expert teacher). Nevertheless, the key idea remains important: To gain expertise, one must deliberately design one’s learning environment: find the right mentor, engage in the evaluative judgment of own performance and outcomes, and create the right level of challenge. In contrast to the theories of the development of everyday (including work) skills, the deliberate learner does not aim only for automatization to minimize effort but also strives to increase the ability to control performance and refine it (Ericsson and Harwell, 2019, p. 4).

Despite its importance for developing exceptional individual expertise, deliberate practice is rarely used to develop teacher expertise in its canonical form, as introduced by Ericsson. Most of the research has been conducted in pre-service teacher education (e.g., Bronkhorst et al., 2013). For in-service teachers, an older survey study (Dunn and Shriner, 1999) found that teachers’ evaluation and planning best parallels deliberate practice activities. Stigler and Miller (2018) suggest the (originally Japanese) practice of lesson study is a promising form for teachers’ deliberate practice. With the rise of video-supported lesson studies (Goldman et al., 2014) this is increasingly becoming a reality. However, at this point in time, reflection on practice is the most frequently used form of deliberate learning by far in the teaching profession (e.g., Saylor and Johnson, 2014).

Adaptive expertise: how to stay an expert

The notion of adaptive expertise complements deliberate practice and is used to describe the continuous improvement of one’s performance, particularly in changing contexts (Bohle Carbonell et al., 2014). Introduced by Hatano and Iganaki (1986), adaptive expertise is best conceptualized in comparison to *routine expertise*: unlike routine experts, adaptive experts are able to adapt to new situations quickly and to regain a high level of proficiency after a short phase of re-orientation and learning. Research into adaptive expertise focuses on the developmental dimension of professional expertise, considering not only the individual factors but also the social and organizational environment. Adaptive expertise is conceptualized as having all of the characteristics of routine expertise and deliberate practice but going beyond those by drawing to a greater extent on meta-cognitive knowledge (Bohle Carbonell et al., 2014). Specifically, adaptive expertise is seen as richer in contextual knowledge: knowing when to apply specific methods and skills and knowing how to adapt them to changes in situations and task demands.

Research on adaptive expertise in teaching is only just emerging. Outside teaching, a recent scoping literature review of adaptive expertise that primarily focuses on health education shows that “Factors influencing development include: *predisposing factors* such as knowledge (ability to integrate knowledge and innovate), beliefs and attitudes (high motivation and humility), *enabling factors* such as skills (people skills, implementing reflection and scholarly activities), resources such as curricular enablers (providing variability of cases, allowing flexibility to generate solutions, critical appraisal of textbooks) and *reinforcing factors* such as mentor-guided feedback and constant curricular review” (Kua et al., 2021, p. 347). Studies that investigate approaches for developing adaptive expertise show the importance of exploration, discovery, and access to information needed for learning. Training with low structure leads to more flexible skills when compared with procedural instruction. For example, Bell and Kozłowski (2008) compared unguided exploration with detailed step-by-step instructions in a game simulation and found that exploration yielded higher scores in an adaptive expertise task. In the same study, Bell and Kozłowski varied how learning from errors was framed: comparing error encouragement with error avoidance. Participants in the error-encouragement condition showed marginally significant higher adaptive expertise. Hughes et al. (2013) compared the same difference in error framing, but different from Bell and Kozłowski (2008), the participants did not receive information about how their errors relate to the knowledge and skills being practiced. They reported no significant correlations between error framing and adaptive expertise. This can be explained by stipulating that to learn from errors, one must attend to the error with the mindset of learning from it and gain sufficient information from the error analysis to improve knowledge and skill.

In teaching, research shows the importance of adaptive expertise in embracing innovations. For example, a case study by Yoon et al. (2015) applied an adaptive expertise lens to a science lesson enacted by three high school biology teachers. Analyzing teachers’ flexibility, deep level of understanding, and deliberate practice, they concluded that adaptive expertise could help explain teachers’ practices during major curriculum reforms. Bowers et al. (2019) used the difference between teacher adaptive expertise and teacher routine expertise to describe differences in how five science teachers went about a lesson in an elementary class. They also concluded teachers’ enactment of new science standards was associated with the moves that characterize adaptive expertise. What is interesting in both studies from an activity system’s point of view is that adaptive expertise is set directly in relation to the central object of the teacher activity: “classroom science teaching.” That is, teachers’ adaptive expertise was studied and seen as important not during teachers’ introductory learning but directly during their teaching. The activity system’s view suggests that teachers’ enactment of new tools (e.g., moving to teach online from a face-to-face classroom) or new ways of organizing an activity (e.g., joint curriculum

planning) also could be associated with their adaptive expertise, but empirical studies exploring the role and functioning of teacher adaptive expertise are still rare.

In the next section, we take an expansive view of teacher expertise and focus on the forms of knowledge underpinning teaching expertise and expert teachers' roles in constructing this knowledge.

Knowledge building: from acquisition to construction of expert knowledge

Teacher expertise is tied to exceptional performance in classrooms and also to particular roles advancing the profession. Experienced teachers often find themselves mentoring more junior colleagues. In addition, a set of new professional roles is emerging that expert teachers increasingly are expected to take on, such as innovators (Penuel et al., 2007), researchers (Darling-Hammond and Bransford, 2005), and knowledge brokers. Expert teachers as *knowledge brokers* translate research knowledge into actionable and useable knowledge for practitioners (e.g., by reducing jargon, creating templates) and facilitate the transfer from research context into local practice context (Joram et al., 2020). This is valuable because, as Cain's (2016) recent review suggests, teachers primarily attend to research when its applications to their practice are highly evident.

Knowledge management expands on the notion of knowledge brokering by focusing on teachers being in charge of facilitating an expert knowledge culture in their school and potentially across schools by working with colleagues in the same role in other schools. *Knowledge cultures* are specific to professions, comprising "... the forms of knowledge in use, the artifacts and tools provided for professional practice, the traditions and methods of knowledge production, and the collective models for knowledge application serve to give communities an integrative power" (Nerland, 2012, p. 28) (For a more extensive discussion of the concept of epistemic cultures and knowledge cultures, see Knorr Cetina (1999) and Markauskaite and Goodyear (2017)). An expert teacher is enculturated into their profession's knowledge culture and therefore in a position to enact—to adopt, adapt and extend—the values and processes in the local context. Expert teachers also are able to enculturate other teachers into how teachers accumulate, distribute and use knowledge. They do so by shaping the processes of *how* knowledge gets created and *what kind* of knowledge gets created.

Knowledge-building processes start from and end in knowledge. Hence, it is essential to ask what kinds of knowledge get created. What is it a profession or a community or practice wants to produce that counts as knowledge? Building on Blackler's (1995) taxonomy of modes of knowledge relevant for studying professional expertise, we can distinguish between (1) knowledge that is closely tied to the individual ("embrained" and "embodied" in Blackler's words), (2) knowledge that is included in social practices and tools ("encultured" and "embedded"), and (3) knowledge that is "encoded" in symbols and signs.

In teaching, the first two forms of knowledge dominate. For example, Klette and Carlsen's (2012) review shows that research concentrates on individualistic and inter-personal forms of knowledge. They reinforce in their study that teachers (in Norway in this case) engage little with research-based—that is, encoded—knowledge, nor do they produce much of external knowledge objects—that is, they are also not occupying themselves with the encoding of knowledge. This is also supported in a review on teachers' use of informal online communities and networks, where none of the use patterns point to the co-construction of knowledge objects that would go beyond narrative accounts of personal experiences or descriptive accounts of learning activities and learning resources (Macià and García, 2016).

When we look for more normative recommendations—what kinds of knowledge should teachers use if not produce?—examples from the literature comprise: (1) shared instructional products (Morris and Hiebert, 2011), (2) theories of action (Brown and Flood, 2018), and (3) principled practical knowledge (Bereiter, 2014). These are all "public, changeable knowledge products" (Morris and Hiebert, 2011, p. 6), ordered along a dimension from instrumental to conceptual.

Morris and Hiebert (2011) ground their suggestion for *shared instructional products* in improvement science, which leads them to a pragmatic-instrumental view of useful knowledge products for teachers: they need to be created around particular learning goals, of a grain size that can guide classroom instruction, be testable and improvable, and accessible to teachers when needed. Although this view of useful knowledge is based on the assumption of direct use by practitioners, Brown and Flood (2018) demonstrate in their literature review that teachers rarely apply research knowledge directly. They experiment with adapting (rather than adopting) an intervention developed elsewhere by developing a "theory of action" for the intervention. "Theories of action represent the systematic exposition of why it is believed strategies or interventions have led or will lead, to change" (Brown and Flood, 2018, p. 144). This necessitates that teachers engage with questions of why and how an intervention works, not only to what extent it works "on average."

This brings us to the third kind of knowledge: principled, practical knowledge (PPK) has characteristics of both practical know-how and scientific theory (Bereiter, 2014). Like basic scientific theory, PPK meets standards of being explicitly encoded in symbols and signs and of having explanatory coherence. However, its primary function is not explanation or prediction but practical guidance: it is "know how" intertwined with "know why." Bereiter includes the guidance property that we find also suggested by Morris and Hiebert (2011), but he makes more explicit what makes knowledge provide guidance: It is the coherence attribute.

We argue that, at this point in time, teacher expertise is limited less by the interest and ability of individual practitioners but by a lack of a knowledge culture that advances powerful knowledge objects. Those need to go beyond the instrumental objects and include the conceptual knowledge objects. For example, the knowledge that makes modern professions, such as engineering, so powerful is "what-if" knowledge (Ylikoski and Kuorikoski, 2010): "What would happen if ... we introduced <x> into our teaching

(in this school)”? In contrast, the knowledge produced in education is overly backward-oriented, too much reliant on evidence and experience. Or more precisely: it lacks the means by which evidence can inform envisioning possible futures. The production of “what-if” and other generative kinds of knowledge are missing from teachers’ practices.

The following section will zoom in on the social processes at play when teachers produce and share knowledge.

Collective expertise: from individual excellence to networked and relational expertise

The need to adapt rapidly to unforeseen and quickly changing circumstances and solve wicked challenges that cannot be understood without considering different perspectives is seen as a key feature that characterizes contemporary professional work (Goodyear and Markauskaite, 2019). Expertise to engage in joint knowledge work, distributed among people from the same or different professional domains (e.g., school counselors, social workers, and teachers), as well as professionals and non-professionals (e.g., parents), are seen as critical in present-day society. To reflect this, some contemporary notions of expertise go beyond the instrumentalist view of what counts as expertise and how it develops, to approaches that acknowledge the context-dependent and relational character of expertise. Such conceptualizations of expertise have roots in different theoretical traditions and disciplinary fields, such as cognitive sciences (Hakkarainen et al., 2017, 2004), social psychology (Edwards, 2010, 2017; Mieg, 2012), philosophy (Boon et al., 2019; Collins and Evans, 2007) and organizational studies (Treem and Leonardi, 2016). These redefinitions of expertise come under different names, such as “networked expertise” (Hakkarainen et al., 2017), “interactional expertise” (Collins and Evans, 2007), “interdisciplinary expertise” (Boon et al., 2019) and “relational expertise” (Edwards, 2017). These approaches inevitably have their distinct features—and some scholars may even disagree with our attempt to put them into the same group—but all of them acknowledge the communicative, collectively accomplished, dynamic, and, therefore, relational nature of expertise.

It is beyond our scope to review all these approaches. To illustrate this broader perspective, we focus on two theoretical lines that have been used in education most extensively: networked expertise and relational expertise.

Networked expertise

Networked expertise is an extension of the established cognitive approaches to the social-constructivist views of knowledge creation (Hakkarainen et al., 2004, 2017). According to this perspective, networked expertise is a higher-level competence that arises within networked communities from sustained engagement in knowledge sharing, joint problem-solving, and knowledge-creation practices. Such expertise is a characteristic of collectives as it emerges from “the fine-tuning of individual competencies to specific requirements of collaborative activity” (Hakkarainen et al., 2017, p. 133). The empirical studies that take this theoretical line primarily focus on how professionals share and gain their knowledge and in other ways create and cultivate their professional networks. Most of these studies rely on the methods and tools offered by social network analysis to get insight into how professionals shape their cognitive socialization.

Tuomainen et al.’s (2012) study of special needs education teachers’ networks within and beyond their schools offers an example of this perspective. Using egocentric network interviews, they asked participants to indicate from whom they get information, ask advice, collaborate, and with whom they interact informally within and beyond their school. Their results showed that these special education teachers worked as networked experts with different teachers within their schools and across the school’s boundaries and other special education-related organizations. These teachers had central positions as collaborators, boundary brokers, and knowledge-sharing agents even if they played marginal roles in the informal networks of their schools.

Overall, studies of networked expertise usually focus on understanding what this expertise looks like within collectives and what kinds of conditions enable it to emerge. These studies conceptualize networked expertise as a generic emergent quality of a group. However, they usually say less about what kind of individual capabilities underpin this expertise, how these capabilities develop, or how exactly this expertise plays out when professionals work on specific challenges or other joint objects.

Relational expertise

Relational expertise has its roots within cultural-historical activity theory (Edwards, 2010, 2017). It primarily focuses on the processes and capacities that enable professionals from two or more different fields to work across professional practice boundaries on specific joint challenges. According to this view, relational expertise is seen as arising from two interrelated processes. Firstly, it involves working together on a joint challenge or object by recognizing what kinds of professional standpoints, motives, and resources others bring to the task. Secondly, it involves aligning one’s interpretations with the responses to the interpretations and responses of the others.

According to this view, relational expertise involves a set of personal capabilities: “Knowing how to recognize the expertise of others and to be able to make one’s own expertise explicit” (Edwards, 2017, p. 8). Such capacity is described as being “professionally multilingual, recognizing the meanings that different practices give to words and their importance in each practice” (Edwards, 2017, p. 8). These capabilities enable teachers to construct shared knowledge. This common knowledge is not a simple extension of what one knows to others, but the integration of professional motives, knowledge, and ways of doing with the motives, knowledge, and ways of doing of others. Relational expertise is, thus, primarily a transformational capability.

According to this perspective, capacities for relational expertise can be learned. For example, “giving and asking for reasons” is seen as a productive way of creating common knowledge (Derry, 2013). This perspective has been often used to examine interprofessional work and collaborations across professional and community boundaries. Studies often draw on research methods and techniques for examining sociocultural practices, such as observations of joint meaning-making activities. For example, Rai (2017) has applied this perspective to examine practices through which schools in rural India have engaged parents and the local community with their children’s education and wellbeing. His specific focus was on how schools, by giving and asking for reasons in meetings with parents and other community members, most of whom had not attended schools, were creating common knowledge about what matters for each party and aligning their responses. According to this perspective, relational expertise does not substitute for contributory specialist knowledge. Instead, it is an expertise that augments the specialist expertise in ways that enables collective work.

Some relational approaches to expertise try to look deeper at the relationship between the specialist (professional or disciplinary) capabilities and the capabilities that allow working relationally across professional or disciplinary boundaries. For example, one such approach conceptualizes interdisciplinary expertise as involving “the *cognitive* ability to connect, translate and establish links between disciplinary knowledge, as well as the *metacognitive* ability to explain and understand the role of *disciplinary perspective*—consisting of, e.g., basic concepts, theories, models, methodologies, technologies, and specific ways of measuring, reasoning and modeling in a discipline—in how knowledge is used and produced” (Boon et al., 2019, p. 668). However, the empirical work lags behind corresponding theoretical advances.

The ecologies of expertise: from fragmented to integrated expertise

Expertise increasingly attracts more diverse and expansive conceptualizations. Too inclusive accounts could be unhelpful. For example, what we see in the emergence of post-truth society or denial of climate change could be regarded as an effect of making boundaries between credible expertise and plain opinion too thin. However, narrow elitist accounts that emerged from early studies of expertise (see **Exceptional performance: from experience and routine expertise to deliberate and adaptive development of personal capabilities** section) are also unhelpful. They provide too little guidance on how high levels of performance in complex domains of human life could be achieved by a broader range of people.

In this section, we focus on three aspects that are important for education: (1) the ecological notions of individual expertise; (2) multilevel conceptualizations of expertise; and (3) multiplicity of expertise.

The ecologies of individual expertise

One of the biggest challenges is that resourceful professional practice—teaching and other professions for which we prepare students—is inseparable from embodied, enacted, and distributed performance and situated within the cultural and material settings thereof (Markauskaite and Goodyear, 2017; Reimann and Markauskaite, 2018). Expertise, once conceptualized as an individual mental phenomenon that could be neatly accounted for by the theories of cognitive architecture and skill development, now increasingly is scrutinized from other theoretical perspectives, some of which we have reviewed above (Collins, 2013; Collins and Evans, 2017; Edwards, 2010). These perspectives acknowledge that human expertise cannot be understood by examining human cognition, including higher-level metacognitive processes, in isolation from the material and social environments and embodied performance. Neither can it be understood without acknowledging that expertise has a tacit, enacted dimension (Markauskaite and Goodyear, 2014). In short, the human mind, body, culture, and the world form an ecological system in which expertise develops and comes into life.

Theories of expertise development, while they sometimes acknowledge this ecology, tend to focus on individual aspects, such as skill development (e.g., deliberate practice), metacognitive ability (e.g., adaptive expertise), or communication (e.g., relational expertise). Cognitive ecology aims to bring these diverse conceptualizations together into one ontologically coherent theoretical account (Hutchins, 2010; Nersessian, 2005). Accounts of professional learning and expertise development that draw on the ecological views are just emerging. These accounts are inevitably tailored for specific domains of expertise, such as scientific discovery (Nersessian, 2012) or professional innovation (Markauskaite and Goodyear, 2017). For example, the account of professional epistemic fluency (Markauskaite and Goodyear, 2017) focuses on the capabilities that underpin professional expertise to take knowledgeable action in solving complex professional problems in human-service professions, such as teaching or school counseling. This account draws on a grounded, ecological view of human cognition (Barsalou, 2008; Hutchins, 2010) and emphasizes that such expertise involves four main capabilities: (1) to combine and integrate *different kinds of knowledge*, such as theoretical knowledge from different disciplines, practical know-how, and just-in-time developed situated understanding; (2) to coordinate and weave *different ways of knowing*, such as theory-informed and case-based reasoning; (3) to create *epistemic tools* and assemble *epistemic environments* that support one’s own or group’s joint knowledge work; and (4) to construct *conscious and conscientious-self*, that is, to be attuned to one’s own capabilities and performance as well as of capabilities and performance of those who work together. Empirical studies of expertise and learning in the human service professions often draw on ethnographic methods. Simultaneously they employ diverse analytical techniques that offer a sharp insight into the diversity of resources that people bring to the solutions of complex tasks and actions that they take (Hutchins et al., 2013; Markauskaite et al., 2020; Nersessian, 2019).

Multilevel conceptualizations of expertise

Why do teams assembled from smart students fail (Barron, 2003)? Why can the Finnish “miracle” of a successful educational system not be explained solely by the expertise of individual teachers (Simola, 2005)? The relationship between individual expertise and expertise of collectives is one of the most critical aspects in contemporary workplaces where complex tasks are done, decisions are made by multiple people, and success relies not only on individual expert performance but on the expert performance of groups and organizations. The expertise of individuals does not guarantee collective expertise, and collective expertise might not be an outcome of star individual performances. For example, an expert teacher may struggle to engage students in deep thinking if school or system assessment regimes do not promote this; similarly, a school may not need to be composed of star teachers to achieve students’ success. An essential question for education is how to develop not only capacities that underpin solo expertise but also collective expertise.

Educational research still grapples with conceptualizing the multi-layered nature of expertise and empirically studying this phenomenon. The emerging multilevel accounts of expertise that have their roots in human ecology (Fulk, 2016; Hawley, 1986) offer some tools for exploring relationships from which collective expertise emerges. Such productive relationships include not only cooperative links—usually seen as key for collective success—but also other kinds of relationships. For example, one critical factor for the successful functioning of software developers’ forums within which expertise is shared is the presence of two types of experts: those who answer questions and those who ask them. “Question askers” are not just free riders but essential actors that allow collective expertise in such networks to emerge (Fulk, 2016).

Similarly, competition in crowdsourcing, during which the expertise of multiple unaffiliated people is mobilized for a short time to work on a particular challenge, proves to be an effective way to promote rapid learning. For example, Lakhani et al. (2013) report that an online competition for solving an immunogenetic problem that lasted two weeks and involved 600 experts resulted in 1000 times faster and more accurate solutions than the benchmark. We still know little about productive configurations for expert learning.

Multiplicity of expertises

Last but not least, expertise is usually seen as a one-dimensional construct that develops through attaining increasingly higher levels of “a mastery of knowledge in a particular domain or subject area” (Fulk, 2016, p. 253) or “an objectively measurable individual problem-solving capacity” (Hakkarainen et al., 2017, p. 133, citing Ericsson, 2009). This conceptualization, however, unhelpfully hides some productive diversity that should be taken into account when we talk about learning in domains where neither the problems can be easily fitted into a particular subject area nor knowledge to solve them can be objectively measured using conventional methods. For example, what kind of expertise do different parties bring to joint problem-solving when a school counselor, a teacher, and a parent work together to develop a special learning strategy for a child with a special learning need? Seeing all three kinds of expertise as equal—thereby placing them on one continuum—or discounting some as merely experience or competence is not helpful. The very idea of expertise, thus, needs to be more flexible and ecological to account for different kinds of expertise.

Researchers have been proposing various typologies for expertise (Collins and Evans, 2007; Mieg, 2012). For example, Collins (2013) has suggested examining the trajectories through which expertise develops by locating them along three continuums, that we introduced in the **Introduction** section: (1) degree of exposure to tacit knowledge; (2) how ubiquitous or rare the expertise is (i.e., esotericism); and (3) whether it is an individual or group accomplishment. For example, expertise gained through years of parenting is rich in tacit knowledge and it is often developed through engagement in a community with other parents. Nevertheless, such expertise is quite ubiquitous within the population. Expertise in teaching is rarer but less tacit than expertise in parenting. It will be learned through the combination of exposure to formal knowledge as well as participation in collective activities with teachers, students, and parents. Counselor’s knowledge is likely to be the rarest but least tacit and developed through engagement in a specialist community rather than in the community of teachers or parents. None of these expertises are just plain opinions, and all require time to develop. Nevertheless, each of them is learned differently.

Synthesis and methodological implications: expertise through an activity system’s lens

For many years expertise has been conceptualized primarily as a property of individuals who achieve exceptional performance levels. It has been an almost exclusive research domain of cognitive psychologists. The last 20 years have brought significant changes in the studies of expertise, acknowledging that expertise as a concept applies simultaneously to individuals and collectives (Fulk, 2016). These collectives include individuals and larger units, such as groups, organizations, and systems. Further, expertise is not only about their accumulated capital and other properties but about performance, which includes learning. Therefore, what counts as expertise is also inevitably not only a cognitive question but also a philosophical, social, and political matter, where the issues of knowledge, culture, and power come into play. What counts as an expert teacher in various accreditation standards, for example, is rarely defined by explicitly referring to research on teacher expertise. Emerging studies of expertise, therefore, increasingly come from different domains, making it sometimes hard to understand how different conceptualizations of expertise interrelate and how their findings contribute to understanding the overall phenomena.

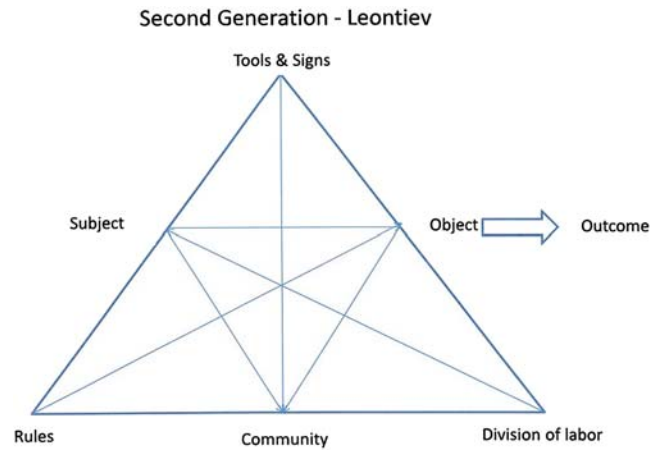


Fig. 1 A representation of second generation CHAT: collective activity, mediational means, division of labor. After Cole and Engeström (1993).

To theorize this dependence, and to motivate the conclusion that teaching expertise should be reconceptualized not as personal property but as an activity of a system, we could build on Cultural Historical Activity Theory (CHAT) that is broadly used in education (for early work see Roth and Lee, 2007). The basic CHAT framework, depicted in Fig. 1, provides a simple conceptual tool that helps view teaching (and related activities such as learning design) as an *activity system* (for detailed description and history of CHAT see Engeström, 2001; Engeström and Pyörälä, 2021).

Taking the example of a teacher conducting a lesson in a classroom, the *subject* (actor) is the teacher, the *object* (a focal entity or desired outcome) is students' learning (i.e., a change in the relation between students and lesson content) and the *tool(s)* is whatever the teacher employs to pursue students' learning and achieve the desired outcomes. A tool can be conceptual—such as a good analogy—or material, such as a paper model or worksheet. The *community of significant others* is "... the people who share with the subject interest in and involvement with the same object" (Foot, 2014, p. 331). These are primarily other teachers and students in the classroom. Relations between the subject and the community are mediated by rules and by the division of labor. *Rules* regulate the subjects' actions (e.g., a teacher cannot leave the classroom during a lesson). The *division of labor* guides what is being done by whom and how during activity, including divisions of tasks and power among teachers and students.

In CHAT, an activity involves a human collective; it is not an individual "behavior" or "action." "Activity" refers to the organized whole engaged in a process that brings about a change to the object. This process consists typically of several steps or phases, which are constituted by the subject's and participants' actions and operations. *Actions* are conscious, goal-oriented, and facilitated by tools. *Operations* are components of actions that are automated or routinized and, hence, fully or partially unconscious. Activities rise and disappear, reacting to the needs of the subject, directed toward the objects of these needs. Hence, activities can be distinguished from each other primarily by their object.

The heuristic value of CHAT is that it suggests a holistic, relational way of thinking about human activities: A change in any of the nodes will have knock-on effects on all other nodes. For instance, a change as simple as a teacher using a new app in her class to teach fractions may change not only classroom routines (the division of labor) but may change what it means to learn/do fractions (the object).

The above example representing a single activity in a classroom obviously oversimplifies teaching. For instance, there will be more than one activity instantiated in any one lesson, such as when the teacher switches between individual and group work or needs to address a disruptive student. And teachers' work is not confined to teaching alone; activities such as lesson planning, preparation for teaching, curriculum planning with peers, and professional learning during curriculum change demand their own "systems."

Nevertheless, CHAT provides some insights into a holistic picture of the metaphorical elephant that the different approaches to (teacher) expertise explore:

- The research showing that expert performance is dependent on *deliberate practice* and partial automatization (Anderson, 1987) is reflected in the distinction between (conscious, intentional) actions and (unconscious) operations.
- *Adaptive expertise* can be modeled primarily through the sensitivity of changes in the object to changes in subjects' needs and also through flexibility in tool use and rules.
- *Knowledge-building practices* and the dependence of what counts as teaching expertise on cultural and social norms (Stigler and Miller, 2018) are reflected in the tools and rules.
- *Relational expertise* and *networking expertise* are partially reflected in divisions of labor between communities and, more generally, in the network of different activity systems.
- *The multilevel, ecological views* of expertise inevitably invite us to look at the dependencies among multiple elements and activity systems.

In addition, to help organize and synthesize findings from the multi-layered and multi-disciplinary (teaching) expertise research, CHAT also provides suggestions for how to study professional expertise development with a method for intervention in activity systems called *expansive learning* (Postholm, 2014; Sannino et al., 2016). This is only an indirect contribution, though, as it is designed as an organizational change method rather than an expertise research method (Engeström, 2011). It is a welcome contribution nevertheless, because the experimental method is largely absent in expertise research where, for obvious reasons, observational and psychometric methods are much more prevalent (e.g., Ackerman and Beier, 2018; Ericsson, 2018).

Without intervening—without experimenting—scientific explanations are difficult to develop (Woodward, 2003). However, any experimental-interventionist approach to studying the development of professional expertise has to address the complexities that come with learning in open systems such as at the workplace (Campbell et al., 2001). Not only does it mean that controlling—or even just measuring—all “factors” that could interfere with the intervention is impossible, it also means conceptually that one has to accept multi-causality: A single intervention can have multiple effects and any of the effects may be produced by one or a combination of other factors (Ragin, 2014). In consequence, the unit of analysis for interventionist research on teacher expertise is the expert teaching system, not the teacher solo. This resonates with the view of teaching and teacher development as an emergent property of complex systems such as articulated by Opfer and Pedder (2011).

A second, more specific contribution of Activity Theory to the study of complex systems by intervening builds on Vygotsky's method of *double stimulation* (Vygotsky, 1999). In modern terminology, in a double simulation intervention the subject is provided with resources new to them that might potentially be useful for solving a problem even though the subject may be able to solve the problem with familiar means. The point of a double simulation experiment is “... the study of those means and devices that the subject used to organize his behavior in concrete forms most adequate for each given task” (Vygotsky, 1999, p. 59). By providing the subject with resources rather than methods and tools in ready-made form, the analysis focusses on the creative and innovative potential of the subject, thus acknowledging that the subject may have a different—and in the case of experts likely deeper—understanding of the activity than the “experimenter.” Put differently, the focus of the analysis is method or tool appropriation rather than learning to use a method or tool (Ritella and Hakkarainen, 2012). This resource-view of change also forms the core of realist accounts of causation of human behavior (Pawson and Tilley, 1997).

Overall, studies of expertise as an activity system invite researchers to combine multiple disciplines, theoretical perspectives, and methods (Nicolini, 2009): to “zoom in” and investigate very specific aspects of teacher practice using methods that give the best insight into what is happening and why, and “zooming out” and investigate relationships, using methods that give the best insight into connections between individual aspects and emerging phenomena that we as a community call “teacher expertise.”

Acknowledgment

Lina Markauskaite's work on this chapter was part-funded by the Australian Research Council through Discovery Project grant DP200100376 (Developing interdisciplinary expertise in universities) and NSW Department of Education through Strategic Research grant G212673 (Developing teachers' interdisciplinary expertise).

References

- Ackerman, P.L., Beier, M.E., 2018. Methods for studying the structure of expertise: psychometric approaches. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge, MA, pp. 213–232.
- Anderson, J.R., 1987. Skill acquisition: compilation of weak-method problem solutions. *Psychol. Rev.* 94, 192–210.
- Barron, B., 2003. When smart groups fail. *J. Learn. Sci.* 12 (3), 307–359.
- Barsalou, L.W., 2008. Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Bell, B.S., Kozlowski, S.W., 2008. Active learning: effects of core training design elements on self-regulatory processes, learning, and adaptability. *J. Appl. Psychol.* 93 (2), 296–316.
- Bereiter, C., 2014. Principled practical knowledge: not a bridge but a ladder. *J. Learn. Sci.* 23, 4–17.
- Blackler, F., 1995. Knowledge, knowledge work and organizations: an overview and interpretation. *Organ. Stud.* 16 (6), 1021–1046.
- Bohle Carbonell, K., Stalmeijer, R.E., Könings, K.D., Segers, M., van Merriënboer, J.J.G., 2014. How experts deal with novel situations: a review of adaptive expertise. *Educ. Res. Rev.* 12, 14–29.
- Boon, M., van Baalen, S., Groenier, M., 2019. Interdisciplinary expertise in medical practice: challenges of using and producing knowledge in complex problem-solving. *Med. Teach.* 41 (6), 668–677.
- Bowers, N., Merritt, E., Rimm-Kaufman, S., 2019. Exploring teacher adaptive expertise in the context of elementary school science reforms. *J. Sci. Teach. Educ.* 31 (1), 34–55.
- Bronkhorst, L.H., Meijer, P.C., Koster, B., Vermunt, J.D., 2013. Deliberate practice in teacher education. *Eur. J. Teach. Educ.* 37 (1), 18–34.
- Brown, C., Flood, J., 2018. Lost in translation? Can the use of theories of action be effective in helping teachers develop and scale up research-informed practices? *Teach. Teach. Educ.* 72, 144–154.
- Cain, T., 2016. Research utilisation and the struggle for the teacher's soul: a narrative review. *Eur. J. Teach. Educ.* 39, 616–629.
- Campbell, T.C., Cook, T.D., Shadish, W.R., 2001. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, second ed. Cengage Learning.
- Chi, M.T.H., Feltovich, P., Glaser, P., 1981. Categorization and representation of physics problems by experts and novices. *Cognit. Sci.* 5, 121–152.
- Cole, M., Engeström, Y., 1993. A cultural-historical approach to distributed cognition. In: Salomon, G. (Ed.), *Distributed Cognitions: Psychological and Educational Considerations*. Cambridge University Press, New York, NY, pp. 1–46.
- Collins, H., Evans, R., 2007. *Rethinking Expertise*. University of Chicago Press, Chicago.
- Collins, H., 2013. Three dimensions of expertise. *Phenomenol. Cognit. Sci.* 12 (2), 253–273.
- Crawford, V.M., Brophy, S., 2006. Adaptive Expertise: Methods, Findings, and Emerging Issues. Retrieved from: <http://ctl.sri.com/publications/downloads/AESymposiumReportOct06.pdf>.

- Darling-Hammond, L., Bransford, J. (Eds.), 2005. *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, New York.
- Davis, B., 2008. Complexity and education: vital simultaneities. *Educ. Philos. Theor.* 40 (1), 50–65.
- Derry, J., 2013. Can inferentialism contribute to social epistemology? *J. Philos. Educ.* 47 (2), 222–235.
- Dunn, T.G., Shriner, C., 1999. Deliberate practice in teaching: what teachers do for self-improvement. *Teach. Teach. Educ.* 15 (6), 631–651.
- Edwards, A., 2010. *Being an Expert Practitioner: The Relational Turn to Expertise*. Springer, Dordrecht.
- Edwards, A., 2017. Revealing relational work. In: Edwards, A. (Ed.), *Working Relationally in and Across Practices: A Cultural-Historical Approach to Collaboration*. Cambridge University Press, Cambridge, pp. 1–22.
- Engeström, Y., Pyörälä, E., 2021. Using activity theory to transform medical work and learning. *Med. Teach.* 43 (1), 7–13.
- Engeström, Y., 2001. Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14, 133–156.
- Engeström, Y., 2011. From design experiments to formative interventions. *Theor. Psychol.* 21 (5), 598–628.
- Ericsson, K.A., Harwell, K.W., 2019. Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: why the original definition matters and recommendations for future research. *Front. Psychol.* 10, 2396.
- Ericsson, K.A., Krampe, R.T., Teschroer, C., 1993. The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100 (3), 363–406.
- Ericsson, K.A., Charness, M.N., Feltovich, P.J., Hoffman, R.R. (Eds.), 2006. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge Univ. Press, New York.
- Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), 2018. *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge, MA.
- Ericsson, K.A. (Ed.), 2009. *Development of Professional Expertise: Toward measurement of Expert Performance and Design of Optimal Learning Environments*. Cambridge University Press, Cambridge, MA.
- Ericsson, K.A., 2018. Capturing expert thought with protocol analysis: concurrent verbalizations of thinking during experts' performance on representative tasks. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A., Williams, A.M. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge, MA, pp. 192–212.
- Fitts, P.M., Posner, M.I., 1967. *Human Performance*. Brooks Cole, Belmont, CA.
- Foot, K.A., 2014. Cultural-historical activity theory: exploring a theory to inform practice and research. *J. Hum. Behav. Soc. Environ.* 24, 329–347.
- Fulk, J., 2016. Conceptualizing multilevel expertise. In: Treem, J.W., Leonardi, P.M. (Eds.), *Expertise, Communication, and Organizing*. Oxford University Press, Oxford, pp. 253–270.
- Gobet, F., Charness, N., 2006. Expertise in chess. In: Ericsson, K.A., Charness, M., Feltovich, P.J., Hoffman, R.R. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge Univ. Press, New York, pp. 523–538.
- Goldman, R., Zahn, C., Derry, S., 2014. Frontiers of digital video research in the learning sciences: mapping the terrain. In: Sawyer, R.K. (Ed.), *Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, New York, pp. 213–232.
- Goodyear, P., Markauskaite, L., 2019. The impact of practice on wicked problems and unpredictable futures. In: Higgs, J., Cork, S., Horsfall, D. (Eds.), *Challenging Future Practice Possibilities*. Brill-Sense, Rotterdam, The Netherlands, pp. 41–52.
- Hakkarainen, K., Palonen, T., Paavola, S., Lehtinen, E., 2004. Communities of Networked Expertise: Professional and Educational Perspectives. Elsevier, Amsterdam.
- Hakkarainen, K., Hytönen, K., Vekkila, J., Palonen, T., 2017. Networked expertise, relational agency, and collective creativity. In: Edwards, A. (Ed.), *Working Relationally in and Across Practices: A Cultural-Historical Approach to Collaboration*. Cambridge University Press, Cambridge, pp. 133–152.
- Hatano, G., Inagaki, K., 1986. Two courses of expertise. In: Stevenson, H.A.H., Hakuta, K. (Eds.), *Child Development and Education in Japan*. Freeman, New York, pp. 262–272.
- Hawley, A.H., 1986. *Human Ecology: A Theoretical Essay*. University of Chicago Press, Chicago.
- Hughes, M.G., Day, E.A., Wang, X., Schuelke, M.J., Arsenault, M.L., Harkrider, L.N., Cooper, O.D., 2013. Learner-controlled practice difficulty in the training of a complex task: cognitive and motivational mechanisms. *J. Appl. Psychol.* 98 (1), 80–98.
- Hutchins, E., Weibel, N., Emmenegger, C., Fouse, A., Holder, B., 2013. An integrative approach to understanding flight crew activity. *J. Cognit. Eng. Decis. Mak.* 7 (4), 353–376.
- Hutchins, E., 2010. Cognitive ecology. *Top. Cognit. Sci.* 2 (4), 705–715.
- Joram, E., Gabriele, A.J., Walton, K., 2020. What influences teachers' "buy-in" of research? Teachers' beliefs about the applicability of educational research to their practice. *Teach. Teach. Educ.* 88, 102980.
- Klette, K., Carlsten, T.C., 2012. Knowledge in teacher learning: new professional challenges. In: Jensen, K., Leif, C.L., Nerland, M. (Eds.), *Professional Learning in the Knowledge Society*. Sense Publishers, Rotterdam, pp. 69–84.
- Knorr Cetina, K., 1999. *Epistemic Cultures*. Harvard University Press, Cambridge, MA.
- Kua, J., Lim, W.-S., Teo, W., Edwards, R.A., 2021. A scoping review of adaptive expertise in education. *Med. Teach.* 43 (3), 347–355.
- Lakhani, K.R., Boudreau, K.J., Loh, P.-R., Backstrom, L., Baldwin, C., Lonstein, E., et al., 2013. Prize-based contests can provide solutions to computational biology problems. *Nat. Biotechnol.* 31 (2), 108–111.
- Lampert, M., 2003. *Teaching problems and the problems of teaching*. Yale University Press, New Haven, CT.
- Leinhardt, G., 1989. Math lessons: a contrast of novice and expert competence. *J. Res. Math. Educ.* 20 (1), 52–75.
- Macià, M., García, I., 2016. Informal online communities and networks as a source of teacher professional development: a review. *Teach. Teach. Educ.* 55, 291–307.
- Markauskaite, L., Goodyear, P., 2014. Tapping into the mental resources of teachers' working knowledge: insights into the generative power of intuitive pedagogy. *Learn. Cult. Soc. Interact.* 3 (4), 237–251.
- Markauskaite, L., Goodyear, P., 2017. *Epistemic Fluency and Professional Education: Innovation, Knowledgeable Action and Actionable Knowledge*. Springer, Dordrecht.
- Markauskaite, L., Goodyear, P., Sutherland, L., 2020. Learning for knowledgeable action: the construction of actionable conceptualisations as a unit of analysis in researching professional learning. *Learn. Cult. Soc. Interact.* 100382.
- Merriam-Webster Thesaurus. Retrieved from <https://www.merriam-webster.com/thesaurus/expert>.
- Mieg, H.A., 2012. *The Social Psychology of Expertise: Case Studies in Research, Professional Domains, and Expert Roles*. Psychology Press, New York.
- Morris, A.K., Hiebert, J., 2011. Creating shared instructional products: an alternative approach to improving teaching. *Educ. Res.* 40 (1), 5–14.
- Nerland, M., 2012. Professions as knowledge cultures. In: Jensen, K., Leif, L., Nerland, M. (Eds.), *Professional Learning in the Knowledge Society*. Sense Publishers, Rotterdam, pp. 27–48.
- Nersessian, N.J., 2005. Interpreting scientific and engineering practices: integrating the cognitive, social, and cultural dimensions. In: Gorman, M.E., Tweney, R.D., Gooding, D.C., Kincannon, A.P. (Eds.), *Scientific and Technological Thinking*. LEA, Mahwah, pp. 17–56.
- Nersessian, N.J., 2012. Engineering concepts: the interplay between concept formation and modeling practices in bioengineering sciences. *Mind Cult. Activ.* 19 (3), 222–239.
- Nersessian, N.J., 2019. Interdisciplinary in action: cognitive ethnography of bioengineering sciences research laboratories. *Perspect. Sci.* 27 (4), 553–581.
- Nicolini, D., 2009. Zooming in and out: studying practices by switching theoretical lenses and trailing connections. *Organ. Stud.* 30 (12), 1391–1418.
- Opfer, V.D., Pedder, D., 2011. Conceptualizing teacher professional learning. *Rev. Educ. Res.* 81 (3), 376–407.
- Pawson, R., Tilley, N., 1997. *Realistic Evaluation*. Sage, Los Angeles.
- Penuel, W.R., Roschelle, J., Shechtman, N., 2007. Designing formative assessment software with teachers: an analysis of the co-design process. *Res. Pract. Technol. Enhanc. Learn.* 2 (1), 51–74.
- Postholm, M.B., 2014. Methodologies in cultural–historical activity theory: the example of school-based development. *Educ. Res.* 57 (1), 43–58.
- Ragin, C.C., 2014. *The Comparative Method*, second ed. The University of California Press, Los Angeles, CA.

- Rai, P.C., 2017. Building and using common knowledge for developing school–community links. In: Edwards, A. (Ed.), *Working Relationally in and Across Practices: A Cultural-Historical Approach to Collaboration*. Cambridge University Press, Cambridge, pp. 96–112.
- Reimann, P., Markauskaite, L., 2018. Expertise. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 54–63.
- Ritella, G., Hakkarainen, K., 2012. Instrumental genesis in technology-mediated learning: from double stimulation to expansive knowledge practices. *Int. J. Comput. Support. Collab. Learn.* 7, 239–258.
- Roth, W., Lee, Y., 2007. Vygotsky's neglected legacy: cultural-historical activity theory. *Rev. Educ. Res.* 77 (2), 186–232.
- Sannino, A., Engeström, Y., Lemos, M., 2016. Formative interventions for expansive learning and transformative agency. *J. Learn. Sci.* 25 (4), 599–633.
- Saylor, L.L., Johnson, C.C., 2014. The role of reflection in elementary mathematics and science teachers' training and development: a meta-synthesis. *Sch. Sci. Math.* 114 (1), 30–39.
- Shulman, L.S., 1987. Knowledge and teaching: foundations of the new reform. *Harv. Educ. Rev.* 57 (1), 1–22.
- Simola, H., 2005. The Finnish miracle of PISA: historical and sociological remarks on teaching and teacher education. *Comp. Educ.* 41 (4), 455–470.
- Sterelny, K., 2012. *The Evolved Apprentice: How Evolution Made Humans Unique*. MIT Press, Cambridge, MA.
- Stigler, J.W., Miller, K.F., 2018. Expertise and expert performance in teaching. In: Williams, A.M., Kozbelt, A., Ericsson, K.A., Hoffman, R.R. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, Cambridge, MA.
- Treem, J.W., Leonardi, P.M. (Eds.), 2016. *Expertise, Communication, and Organizing*. Oxford University Press, Oxford.
- Tuomainen, J., Palonen, T., Hakkarainen, K., 2012. Special educators' social networks: a multiple case study in a Finnish part-time special education context. *Scand. J. Educ. Res.* 56 (1), 21–38.
- Vygotsky, L.S., 1999. *Tool and Sign in the Development of the Child* (Vol. 6 Scientific Legacy). Kluwer/Plenum, New York.
- Woodward, J., 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, New York.
- Ylikoski, P., Kuorikoski, J., 2010. Dissecting explanatory power. *Phil. Stud.* 148, 201–219.
- Yoon, S.A., Koehler-Yom, J., Anderson, E., Lin, J., Klopfer, E., 2015. Using an adaptive expertise lens to understand the quality of teachers' classroom implementation of computer-supported complex systems curricula in high school science. *Res. Sci. Technol. Educ.* 33 (2), 237–251.

From grit and resilience to academic tenacity

Chathurika Kannangara, Rosie Allen, and Jerome Carson, Department of Psychology, School of Education and Psychology, University of Bolton, Bolton, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

An overview of resilience	190
Relevance of resilience for education	190
A short review of grit	191
Relevance of grit for education	191
Investigating the construct of grit and its relationship to other factors	192
Study 1	192
Study 2	192
Study 3	192
A critical overview of grit and resilience as independent factors	193
Beyond grit and resilience: the relevance of academic tenacity	193
Study 1	193
Study 2	194
International validation of the Bolton Uni-Stride Scale	194
Does academic tenacity actually matter?	194
Conclusion	195
References	195

An overview of resilience

Despite the popularity of research into psychological resilience and the ever-growing use of the term “resilience,” there is no universal definition (Aburn et al., 2016). One definition, provided by Abiola and Udofia (2011), states that resilience is “a psychological process developed in response to intense life stressors that facilitates healthy functioning” and further points out that “resilience connotes inner strength, competence, optimism, flexibility, and the ability to cope effectively when faced with adversity” (p. 2). However, it remains uncertain what is truly meant by the term “resilience” in an academic context and there is little agreement on the terminology used by researchers to ascribe what resources and factors contribute toward resiliency. For example, social support networks (Grant and Kinman, 2012; Friborg et al., 2003) and making connections, are detailed as integral elements in building resilience. Despite the use of different terminology, all substantiate the same claim. That is, having access to healthy social support networks increases the likelihood of developing resilience. Yet, the inconsistency in definitions and terminology hinders the progression of research, knowledge and understanding of such an important construct (Turner et al., 2017). Although a variety of definitions exist for resilience, the definitions are largely re-iterating the same concept (Cassidy, 2015). That is, resilience provides an individual with an increased probability to overcome challenges and cope with stressful events, while retaining or regaining their mental health (Herman et al., 2011).

In the absence of an agreed definition, the measurement of resilience must also be evaluated. As expected, there are several resilience scales, none of which are seen as the most preferable (Windle et al., 2011). Only a handful of resilience measures have been developed specifically for use in a university setting. For instance, the Academic Resilience Scale (ARS-30), focuses on self-regulated learning and self-efficacy beliefs related to a university context (Cassidy, 2016) and its focus is linked to academic success. At a similar time, the Resilience at University (RAU) scale was developed (Turner et al., 2017), which importantly recognized resilience as a skill that could be trained or developed. However, there is still a need for a university specific resilience scale that is reliable, valid, structurally sound and acknowledges the vital role of wellbeing.

Relevance of resilience for education

An evident inconsistency in definitions, terminology and measures, has left resilience research lacking in its ability to progress toward an understanding of its importance in university students. The sheer complexity and multi-dimensional nature of resilience as a construct, warrants a deeper investigation into the qualities and resources that underpin it. The development of certain internal characteristics that influence resilience are apparent, but it is essential to recognize the external factors within the environment that aid in nurturing these characteristics. Understanding what factors must be present for resilience to be developed and maintained, is crucial when attempting to cultivate a resilient student population. It is those very factors that will aid in boosting resiliency and therefore improve the overall psychological well-being and educational success of students. Regardless of this, research has demonstrated the role resilience plays in the education and wellbeing of students.

Resilience provides students with the ability to cope and manage with the stressors that often accompany studying at university (Stallman, 2010). More specifically, resilience differentiates between students who can cope with the pressures of university life and those who struggle to cope (Pidgeon et al., 2014). A huge part of studying at university demands an adaptation to the environment. Many students come from far away and it may be their first time living away from home, without the immediate help and support from their family and friends. Resilient students are increasingly likely to adjust to the university environment (Wagnild and Young, 1993; Ahern et al., 2006) and are better able to adapt under pressure (Wang, 2009; Edwards et al., 2001). Much research has demonstrated that academic resilience is a strong predictor of academic achievement in general (Beauvais et al., 2014; Mwangi and Okatcha, 2015; Solomon, 2013) and it has been further identified as a key aspect of student engagement and persistence on tasks (Jowkar et al., 2014).

Not only does resilience convey educational advantages, it is also evident that university students who have the characteristics and resources necessary to cope with the demands of university life, have better mental health (DeRosier et al., 2013; Vago and Silbersweig, 2012; Epstein and Krasner, 2013; Carver et al., 2010). Furthermore, empirical evidence shows that resilience is a key characteristic that aids an individual to cope with psychological pressures (Waxman et al., 2003) and allows students to maintain their mental health as they have the ability to quickly recover from difficult situations (Souri and Hasanirad, 2011; Ho et al., 2010). Students who believe they have the ability to overcome challenges and bounce back from adversity, are typically more likely to report feeling happier, optimistic and more confident (McIntosh and Shaw, 2017). Similarly, resilient students have greater self-esteem, self-efficacy, life satisfaction and quality of life (Abiola and Udofia, 2011; McIntosh and Shaw, 2017). Resilient students are less likely to report feeling stressed, lonely, anxious or depressed (Felsten, 2004; McIntosh and Shaw, 2017; Abiola and Udofia, 2011), and are less likely to experience psychological distress (Klohenen et al., 1996; Pidgeon et al., 2014; Byrd et al., 2012; Storrie et al., 2010).

A short review of grit

While traditional theories in psychology advocated the importance of intelligence testing when predicting achievement and success, research surfaced later that contradicted this perspective. Indeed, research indicated that neither intelligence nor IQ were the most significant predictors of success, rather it was an individual's perseverance of effort and consistency of interests (Duckworth et al., 2007). "Grit" has been defined as expressing a passion and perseverance for pursuing long-term goals. It provides individuals with the ability to strive toward achieving their long-term goals, while persisting in the face of adversity (Duckworth, 2017). Initially articulated by Angela Duckworth in 2007, the concept of grit has expanded in terms of attracting the attention of education and psychology researchers from across the globe (Duckworth et al., 2007; Duckworth, 2017). Subsequently, investigations into the potential applications of grit began, including the usefulness of grit in predicting academic outcomes such as engagement, motivation, performance, and achievement. As a result, grit was shown to be a characteristic that was strongly associated with academic outcomes. For instance, students who expressed a passion toward their schoolwork and persevered with their studies, despite academic and social challenges, were most likely to experience academic success (Lee and Sohn, 2017).

Relevance of grit for education

Since Duckworth's early research, other researchers from across the world have investigated grit and reported several useful applications, including in an educational context. Research shows the central role of grit in determining which students will excel in their academic studies and identified grit as a significant predictor of academic achievement at all levels (Wolters and Hussain, 2014; Egalite et al., 2015; Rimmelfeld et al., 2016). Moreover, grittier students were shown to spend more time studying (Cross, 2013), exemplify increased academic motivation (Eskreis-Winkler et al., 2014) and expressed a higher sense of self-efficacy (Oriol et al., 2017). It has also been documented that "grittier students" are more likely to be engaged in their academic studies (Datu et al., 2015; Hodge et al., 2017) and are more likely to obtain higher grades (Lee and Sohn, 2017; Cross, 2013; Pate et al., 2017; Montas et al., 2021).

The retention and dropout rates of students have long been a concern for educational institutions and organizations all over the world (Brown, 2012). Grit has been shown to predict the retention of military cadets in an extensive training program (Duckworth et al., 2007; Eskreis-Winkler et al., 2014; Maddi et al., 2012) and of novice teachers in challenging education environments (Robertson-Kraft and Duckworth, 2014; Duckworth and Quinn, 2009). More importantly, research has suggested that grit can greatly impact the retention of students in schools, colleges, and universities (Bazelaïs et al., 2016; Crede et al., 2016; Eskreis-Winkler et al., 2014). The identification and awareness of long-term goals and the early identification of their course of interest, are critical for student retention (Hagedorn et al., 2001).

Recently, research has suggested that applications of grit are not confined to academic outcomes, but also extend to well-being. The student mental health crisis has become of increasing concern, especially to the United Kingdom government and we know that grit is a strong contributor toward higher well-being in university students (Vainio and Daukantaite, 2015; Sharkey and Gamwell, 2017; Kannangara et al., 2018; Salles et al., 2014; Jin and Kim, 2017; Goodman et al., 2017). Further, grittier individuals are happier (Singh and Jha, 2008), have better life satisfaction (Singh and Jha, 2008) and a greater sense of belonging (Bowman et al., 2015), with further analysis showing that emotional stability during stressful or negative life events is buffered

by higher levels of grit (Blalock et al., 2015). Higher levels of grit predict positive affect (Singh and Jha, 2008; Hill et al., 2014), a greater sense of purpose (Hill et al., 2014), better self-esteem (Weisskirch, 2016), more meaning in life (Hill et al., 2014) and enhance the likelihood of pursuing engagement and pleasure in life (Culin et al., 2014). Moreover, a negative correlation has been revealed between grit and negative affect (Singh and Jha, 2008; Datu et al., 2015), depression (Jin and Kim, 2017) and stress (Lee, 2017; Kannangara et al., 2018). So, not only does grit help to determine which students are more likely to achieve and perform to a high standard, it also has an influential effect on psychological well-being. Identifying students who present low levels of grit, enables educators to direct them to support services that provide the opportunity for personal development and growth, leading to better academic and well-being outcomes.

Investigating the construct of grit and its relationship to other factors

Kannangara (2015), first articulated a languishing thriving model for dyslexia. She commented, "... a thriving dyslexic predominantly displays characteristics such as positive acceptance toward challenges, embraces them, uses signature strengths when they come across obstacles, learns from criticisms, perseveres, withstands and finds alternative approaches when they face failures," (Kannangara, 2015, p. 3). Languishing dyslexics do the reverse. The three empirical studies we describe below, show that the languishing thriving model can also be applied to university students. In these studies, Kannangara et al. (2018) investigated the validity of grit and its relationship to resilience. The three studies involved an examination of a range of factors such as resilience, self-control, well-being and mindsets in relation to grit. High grit groups and low grit groups were also compared.

Study 1

The aim of Study 1 was to investigate the construct of grit and its relationship with other non-cognitive constructs such as self-control, well-being and perceived stress. The study was carried out at the University of Bolton with a total of 440 foundation, undergraduate and postgraduate students. There were 245 female participants and 170 males. All participants were asked to complete a series of questionnaires including, the 12-item Grit Scale (Duckworth et al., 2007), the Perceived Stress Scale (PSS) (Cohen et al., 1983), the Mental Well-being Scale (Short Warwick-Edinburgh Mental Well-being Scale), the ONS Personal Well-being items (Office for National Statistics, 2011) and the Self-Control Scale (SCS) (Tangney et al., 2004).

As there were more female participants, a randomized gender sample of 150 females was compared with 150 males. Women were found to have significantly higher grit scores, mean = 41.54, versus 39.26 for men ($p < 0.01$, effect size = 0.33). Older participants scored significantly higher than younger participants, as did postgraduate students. Grit correlated highest with self-control, $r = 0.52$, followed by mental well-being, $r = 0.38$.

Study 2

Study 2 was aimed at further examining the construct of grit and its relationship to other known predictors of success. A similar method to Study 1 was used in Study 2, including measures to investigate self-control and well-being. Additionally, resilience and mindsets were also measured to further validate the meaningfulness of grit as a construct. The construct validity of grit was investigated by comparing low grit and high grit scoring participants. A total of 340 foundation, undergraduate and postgraduate students were recruited. In addition to the measures included in Study 1, resilience was measured using the Connor-Davidson Resilience Scale (CD-RISC 10) and mindsets were measured using the Mindsets Quiz.

The second study confirmed the earlier finding of gender differences in grit with women again scoring higher 40.93, versus 38.81 ($p < 0.05$, effect size = 0.31). Similar differences were also found with age and postgraduate study, as in Study 1. Older students again had significantly higher grit scores, as did postgraduate students. A multiple regression analysis showed self-control and resilience were the best predictors of grit.

Study 3

This study was designed to examine the construct of grit and its links with success and academic outcomes, from a qualitative perspective. Using a snow-ball sampling technique, a total of 10 successful graduates from the university were interviewed. Thematic analysis was used to analyze the interview transcripts. Those invited for an interview were students who graduated with a first-class honors degree and those who scored high on the Grit Scale (scores above 4, the highest score on the Grit Scale is 5).

Three main themes were identified from the interview transcripts. These were passion and perseverance, self-control and positive mindsets. Two of the 10 students interviewed started their degrees with only limited knowledge of written and spoken English and yet ended up with first class honors degrees. The students reported numerous personal challenges over the 3-year course of their degrees, including financial problems, bereavements, childcare and health problems. One student had four surgeries over the 3 years. These successful students also had excellent time management skills. In addition, their knowledge of their personal strengths and weaknesses helped them increase their success. These students also had a growth mindset and used constructive criticism from their assignments to improve their work.

Taken together these three studies show the importance of grit for academic success. The “thriving” students embrace challenges and have a growth mindset. They demonstrate learned optimism and persistence in the face of difficulties. They used their personal strengths to get the best out of themselves. Grit shows itself in their passion and determination to achieve their goals. Finally, self-control helps them bring all of the above together, with good time management and organizational skills. These studies showed that, in addition to grit and resilience, a growth mindset, strengths use, self-control and mental well-being, were also important factors in student success (Kannangara et al., 2018). This led us to think about developing a composite measure, the Bolton Uni-Stride Scale (BUSS), as a way of capturing this group of factors which were important for students to “thrive” at university (Kannangara et al., 2020).

A critical overview of grit and resilience as independent factors

The constructs of resilience and grit are often researched and discussed interchangeably. After all, there is some overlap with their definitions and applications in educational and psychological research (Stoffel and Cain, 2018). As expected, resilience and grit are closely associated (Kannangara et al., 2018; Musso et al., 2019; Montas et al., 2021; Meyer et al., 2020). More specifically, perseverance, a key concept of grit, can predict academic resilience (Ali and Rahaman, 2012). Also, goal setting, a core component of grit, is a characteristic that strongly influences the development of resilience (McIntosh and Shaw, 2017). For instance, setting relevant and meaningful goals that are aligned with their interests develops a student’s sense of motivation and commitment, that compels them to persist in the face of adversity. Indeed, recent research posited students who score high on grit are significantly more resilient than students who score low on grit (Kannangara et al., 2018). Clearly, these constructs are related to some degree. Even Angela Duckworth herself, suggests that resilience is a component of grit (Duckworth et al., 2007; Duckworth and Quinn, 2009), and research advocates the strong interrelated nature of the two constructs.

Nonetheless, evidence has mounted to suggest that resilience does not necessarily equate to grittiness, and vice versa (Hardeman, 2016). One study directly compared grit and resilience and found a negative correlation, whereby particularly gritty individuals showed lower levels of resilience (Hardeman, 2016). Although grit and resilience are highly correlated, an individual can be resilient, while that individual may not necessarily be “gritty.” Resilience reflects an ability to bounce back after challenges and engenders a positive response to failure or adversity. On the other hand, grit is considerably more goal orientated and reflects a consistent effort toward reaching short and long-term goals, despite encountering adversities. They are undoubtedly related, as part of being “gritty” is to be resilient in the face of adversity (Perkins-Gough, 2013). After all, resilience is an important part of being “gritty,” as you would be less likely to persist with long-term goal pursuit if you gave up when you encountered challenges (Stoffel and Cain, 2018). Yet, being “gritty” also requires a consistency of interests that drives passion in an individual to stay focused for a long period of time and commit to plans to achieve long term goals, which is not necessarily a characteristic of resilience. Likewise, grit is a contributor toward resilience, but there are many personal characteristics, behavioral, societal and genetic influences, affecting an individual’s capacity to overcome stressful life events (Matthews, 2020). Similarly, grit is correlated with, but distinct from, many concepts that constitute resilience, such as coping strategies (Musso et al., 2019). Both resilience and grit play an important role for the others existence, but remain distinctly separate constructs (Meyer et al., 2020).

Beyond grit and resilience: the relevance of academic tenacity

In our earlier work, we had identified six factors as being linked to academic achievement. These were grit, resilience, strengths use, self-control, mindsets, and mental well-being. Although many researchers have examined these factors separately, we have argued that they are strongly inter-related (Kannangara et al., 2020). To measure all six constructs individually would require a lengthy survey. We therefore decided to construct a short scale, which we named the Bolton Uni-Stride Scale (BUSS), to measure what we referred to as “academic tenacity.” Items were chosen from existing scales, but re-worded. We chose items which correlated the highest with the total score on the scale of origin. The final scale had 12 items. Three were adapted from the Grit Scale (Duckworth et al., 2007), two from the Self-Control Scale (Tagney et al., 2004), two from the Warwick-Edinburgh Mental Well-Being Scale (Tennant et al., 2007), two from the Mindsets Quiz, one from the CD-RISC Resilience Scale (Connor and Davidson, 2003), and two on strengths use were generated from a review of the strengths’ literature. Kannangara et al. (2020) conducted two studies to examine the validity of the BUSS.

Study 1

The BUSS was given to 1117 1st year students at the University of Bolton. Principal Axis factoring using Varimax rotation was used to identify the least number of factors. Two Factors were obtained. Factor 1 “Tenacity” had seven items and Factor 2 “Self-Composure” had five items. The Tenacity Factor accounted for 30.51% of the total variance, while Self-Composure accounted for 14.20%. Together both factors accounted for 44.71% of the total variance in BUSS scores.

Study 2

In the second study, the BUSS was given to 340 students, who were also asked to complete the short Warwick-Edinburgh Mental Well-Being Scale, the CD-RISC 10, the Grit Scale, the Self-Control Scale and the Mind-Sets Quiz. A second Factor Analysis of scores on the BUSS, confirmed the original two factor solution found in Study 1. Test-retest reliability over a three-week period was found to be 0.70 for Tenacity and 0.77 for Self-Composure. Tenacity correlated highest with resilience, $r = 0.61$, followed by mental well-being at 0.54, and then grit at 0.46. Self-Composure correlated highest with grit at $r = 0.59$, followed by self-control at 0.55.

Both studies showed that the BUSS had good reliability and validity. However, the scale had been piloted and developed at only one English university. There was therefore a need to test the scale with different populations of university students.

International validation of the Bolton Uni-Stride Scale

As noted above the problem with the further development of the BUSS was that it had been developed solely at the University of Bolton in the North-West of England. With the encouragement of our colleague, Professor Patrick McGhee, [Kannangara et al. \(2022\)](#) conducted an international validation of the BUSS. We designed a web-based survey, which was delivered over the Prolific platform. Prolific offers access to survey participants worldwide, who are paid for completing surveys ([Palan and Schitter, 2018](#)). Prolific enables researchers to specify the samples they wish to survey. In our case we wanted to ensure that we surveyed university students, and participants who were fluent in English. We recruited a sample of 1043 individuals from across 30 different countries. The largest group of students was from the United States ($n = 399$), followed by Portugal ($n = 119$), Poland ($n = 102$) and Mexico ($n = 68$). Participants were asked to complete the following standardized questionnaires:

1. The Bolton Uni-Stride Scale ([Kannangara et al., 2020](#)).
2. The Self-Control Scale ([Tagney et al., 2004](#)).
3. The Future Work Self Scale ([Strauss et al., 2012](#)).
4. The Connor Davidson Resilience Scale, CD-RISC 10 ([Connor and Davidson, 2003](#)).
5. The 14-item version of the Warwick Edinburgh Mental Wellbeing Scale ([Tennant et al., 2007](#)).
6. The Short Grit Scale ([Duckworth and Quinn, 2009](#)).

Although participation was voluntary, respondents received a small cash payment via the Prolific website for completing the questionnaires.

A confirmatory factor analysis on the BUSS international data again confirmed the existence of a two-factor structure of the BUSS, with seven items loading on the Tenacity Factor and five on the Self-Composure Factor. Item 8, *"I set goals, but after a while I decide on a new set of goals,"* loaded only weakly onto the Self-Composure Factor. The internal consistency of the BUSS was good at 0.80, with Tenacity being 0.83 and Self-Composure, 0.57. Encouragingly Total BUSS scores correlated 0.73 with the Short Grit Scale and 0.69 with the CD-RISC Resilience Scale. Item discrimination analysis showed that all 12 BUSS items discriminated between high and low scorers on the scale. A linear regression with Tenacity as the dependent variable, showed that the most significant predictors were Grit, Resilience and Self-Control, in that order. The international validation study showed that the BUSS was a reliable and valid measure of academic tenacity with this sample of international university students.

Does academic tenacity actually matter?

To answer this question, the international sample was divided into High scorers ($n = 466$), versus Low scorers on the BUSS ($n = 492$). High scorers were defined as those individuals with scores between 41 and 57 and Low scorers those with scores between 20 and 40. [Table 1](#) clearly demonstrates that Low scorers on the BUSS, had significantly lower scores on each of the study dependent

Table 1 High versus low BUSS scorers on the study dependent variables.

Measure	High BUSS 41–57, $n = 466$	Low BUSS 20–40, $n = 492$	Value of t , p value	Effect size, Hedges' g
Self-control	38.73	28.35	$t = -18.87$, $p = 0.001$	1.22
Future work self	18.12	13.27	$t = -15.64$, $p = 0.001$	1.01
CD RISC resilience	38.50	30.99	$t = -20.74$, $p = 0.001$	1.35
Warwick Edinburgh Mental wellbeing	50.05	40.97	$t = -17.06$, $p = 0.001$	1.12
Short grit scale	28.54	22.36	$t = -21.34$, $p = 0.001$	1.37

variables. The largest effect sizes were for the Short Grit Scale and the CD-RISC 10 Resilience measure, again showing the importance of the concepts of Grit and Resilience. The Table demonstrates that Academic Tenacity, as measured by the BUSS, does actually matter. The biggest effect sizes are for grit and resilience.

Conclusion

This chapter has considered the importance of the concepts of grit and resilience, specifically for university students. Three studies conducted by the authors examined the concept of grit. Although these studies confirmed the importance of grit for academic success, they also suggested that other factors were important. The authors put forward a model of “thriving versus languishing students.” They argued that six factors were critical for thriving. In addition to grit, these were resilience, self-control, mindsets, mental well-being, and strengths use. Rather than assessing each of these attributes individually, the authors devised a scale, the BUSS, to tap into all six factors in a short 12-item scale. Two studies established a two-factor structure for the scale, “Tenacity” and “Self-Composure.” The scale possesses good internal reliability and test re-test reliability, has good face and content validity, as well as good convergent and concurrent correlational validity. This has been confirmed in an international validation. The BUSS is now being translated into different languages and is being used by other researchers.

It is of course one thing to measure the concepts of grit, resilience and tenacity, but what can be done to improve them? The Hummingbird Project (Platt et al., 2020), offered an opportunity to deliver a six-week psycho-education project based on Positive Psychology (Rashid and Seligman, 2018; Carr, 2020), to improve wellbeing, hope and tenacity in high school students. The intervention was delivered to over 1000 high school students in its 1st year and has now been running for 3 years. The researchers have also developed Hummingbird Primary and, along with the Open University, have developed a “Bounce Back” program for parents, to share positive psychology techniques with them. Additionally, there are plans to devise a similar multicomponent Positive Psychology Intervention for university students.

References

- Abiola, T., Udofia, O., 2011. Psychometric assessment of the Wagnild and Young's resilience scale in Kano, Nigeria. *BMC Res. Notes* 4 (1), 105. <https://doi.org/10.1186/1756-0500-4-509>.
- Aburn, G., Gott, M., Hoare, K., 2016. What is resilience? An integrative review of the empirical literature. *J. Adv. Nurs.* 72 (5), 980–1000.
- Ahern, N., Kiehl, E., Sole, M., Byers, J., 2006. A review of instruments measuring resilience. *Issues Compr. Pediatr. Nurs.* 29 (2), 103–125.
- Ali, J., Rahaman, A., 2012. A comparative study of grit between male and female fencers of Manipur. *Shield* 7, 32–36.
- Bazelais, P., Lemay, D.J., Doleck, T., 2016. How does grit impact college students' academic achievement in science? *Eur. J. Sci. Math. Educ.* 4 (1), 33–43.
- Beauvais, A.M., Stewart, J.G., Denisco, S., Beauvais, J.E., 2014. Factors related to academic success among nursing students: a descriptive correlational research study. *Nurse Educ. Today* 34 (6), 918–923. <https://doi.org/10.1016/j.nedt.2013.12.005> PMID: 24380623.
- Blalock, D.V., Young, K.C., Kleiman, E.M., 2015. Stability amidst turmoil: grit buffers the effects of negative life events on suicidal ideation. *Psychiatr. Res.* 228 (3), 781–784. <https://doi.org/10.1016/j.psychres.2015.04.041>.
- Bowman, N.A., Hill, P.L., Denson, N., Bronkema, R., 2015. Keep on truckin' or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Soc. Psychol. Personal. Sci.* 6 (6), 639–645. <https://doi.org/10.1177/1948550615574300>.
- Brown, J., 2012. Developing a freshman orientation survey to improve student retention within a college. *Coll. Student J.* 46 (4), 834–851.
- Byrd, D.R., 2012. Peer reviewed: race/ethnicity and self-reported levels of discrimination and psychological distress, California, 2005. *Prev. Chronic Dis.* 9, E156. <https://doi.org/10.5588/pcd9.120042>.
- Carr, A., 2020. *Positive Psychology and You: A Self-Development Guide*. Routledge.
- Carver, C.S., Scheier, M.F., Segerstrom, S.C., 2010. Optimism. *Clin. Psychol. Rev.* 30 (7), 879–889.
- Cassidy, S., 2015. Resilience building in students: the role of academic self-efficacy. *Front. Psychol.* 6, 1781. <https://doi.org/10.3389/fpsyg.2015.01781>.
- Cassidy, S., 2016. The Academic Resilience Scale (ARS-30): a new multidimensional construct measure. *Front. Psychol.* 7. <https://doi.org/10.3389/fpsyg.2016.01787> PMID: 27917137.
- Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. *J. Health Soc. Behav.* 24, 386–396.
- Connor, K., Davidson, J., 2003. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depress. Anxiety* 18 (2), 76–82. <https://doi.org/10.1002/da.10113> PMID: 12964174.
- Crede, M., Tynan, M.C., Harms, P.D., 2016. Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Pers. Soc. Psychol.* 113, 492–511. <https://doi.org/10.1037/pspp0000102>.
- Cross, T.M., 2013. *Staying the Course: Grit, Academic Success, and Non-traditional Doctoral Students*. Doctoral Dissertation. Pepperdine University.
- Culin, K.R., Tsukayama, E., Duckworth, A.L., 2014. Unpacking grit: motivational correlates of perseverance and passion for long-term goals. *J. Posit. Psychol.* 9 (4), 306–312. <https://doi.org/10.1080/17439760.2014.898320>.
- Datu, J.A., Valdez, J.P., King, R.B., 2015. Perseverance counts but consistency does not! Validating the Short Grit Scale in a collectivist setting. *Curr. Psychol.* 35 (1), 121–130. <https://doi.org/10.1007/s12144-015-9374-2>.
- DeRosier, M.E.P., Frank, E.P., Schwartz, V.M.D., Leary, K.A.P., 2013. The potential role of resilience education for preventing mental health problems for college students. *Psychiatr. Ann.* 43 (12), 538–544.
- Duckworth, A., 2017. *Grit: The Power of Passion and Perseverance*. Vermilion, London.
- Duckworth, A.L., Quinn, P.D., 2009. Development and validation of the short grit scale (grit-S). *J. Pers. Assess.* 91, 166–174. <https://doi.org/10.1080/00223890802634290>.
- Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: perseverance and passion for long-term goals. *J. Pers. Soc. Psychol.* 92, 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>.
- Edwards, K.J., Hershberger, P.J., Russell, R.K., Markert, R.J., 2001. Stress, negative social exchange, and health symptoms in university students. *J. Am. Coll. Health* 50, 75–79. <https://doi.org/10.1080/07448480109596010>.
- Egalite, A.J., Kisida, B., Winters, M.A., 2015. Representation in the classroom: the effect of own-race teachers on student achievement. *Econ. Educ. Rev.* 45, 44–52. <https://doi.org/10.1016/j.econedurev.2015.01.007>.

- Epstein, R.M., Krasner, M.S., 2013. Physician resilience: what it means, why it matters, and how to promote it? *Acad. Med.* 88 (3), 301–303. <https://doi.org/10.1097/ACM.0b013e318280cfff>.
- Eskreis-Winkler, L., Shulman, E.P., Beal, S.A., Duckworth, A.L., 2014. The grit effect: predicting retention in the military, the workplace, school and marriage. *Front. Psychol.* 5, 36. <https://doi.org/10.3389/fpsyg.2014.00036>.
- Felsten, G., 2004. Stress reactivity and vulnerability to depressed mood in college students. *Pers. Individ. Differ.* 36, 789–800. [https://doi.org/10.1016/S0191-8869\(03\)00152-1](https://doi.org/10.1016/S0191-8869(03)00152-1).
- Friborg, O., Hjemdal, O., Rosenvinge, J.H., Martinussen, M., 2003. A new rating scale for adult resilience: what are the central protective resources behind healthy adjustment? *Int. J. Methods Psychiatr. Res.* 12, 65–76. <https://doi.org/10.1002/mpr.143>.
- Goodman, F.R., Disabato, D.J., Kashdan, T.B., Machell, K.A., 2017. Personality strengths as resilience: a one-year multi-wave study. *J. Pers.* 85 (3), 423–434. <https://doi.org/10.1111/jopy.12250>.
- Grant, L., Kinman, G., 2012. Enhancing wellbeing in social work students: building resilience in the next generation. *Soc. Work. Educ.* 31 (5), 605–621.
- Hagedorn, F., Schleppei, P., Bucher, J., Flühler, H., 2001. Retention and leaching of elevated N deposition in a forest ecosystem with gleysols. *Water Air Soil Pollut.* 129 (1), 119–142.
- Hardeman, J., 2016. Comparing Resilience and Grit: An Empirical Examination. Master's Thesis. Drexel University, Philadelphia, PA.
- Herman, H., Stewart, D.E., Diaz-Granados, N., Berger, E.L., Jackson, B., Yuen, T., 2011. What is resilience? *Can. J. Psychiatr.* 56 (5), 258–265. <https://doi.org/10.1177/070674371105600504>.
- Hill, P.L., Burrow, A.L., Bronk, K.C., 2014. Persevering with positivity and purpose: an examination of purpose commitment and positive affect as predictors of grit. *J. Happiness Stud.* 17, 257–269. <https://doi.org/10.1007/s10902-014-9593-5>.
- Ho, M.Y., Cheung, F.M., Cheung, S.F., 2010. The role of meaning in life and optimism in promoting well-being. *Pers. Individ. Differ.* 48 (5), 658–663.
- Hodge, B., Wright, B., Bennett, P., 2017. The role of grit in determining engagement and academic outcomes for university students. *Res. High. Educ.* <https://doi.org/10.1007/s11162-016-9435-x>.
- Jin, B., Kim, J., 2017. Grit, basic needs satisfaction, and subjective well-being. *J. Individ. Differ.* 38 (1), 29–35. <https://doi.org/10.1027/1614-0001/a000219>.
- Jowkar, B., Kojuri, J., Hayat, A.A., 2014. Academic resilience in education: the role of achievement goal orientations. *J. Adv. Med. Educ. Prof.* 2 (1), 33–38.
- Kannangara, C., 2015. From languishing dyslexia to thriving dyslexia: developing a new conceptual approach to working with people with dyslexia. *Front. Psychol.* 6, 1976. <https://doi.org/10.3389/fpsyg.2015.01976>.
- Kannangara, C.S., Allen, R.E., Waugh, G., Nahar, N., Khan, S., Rogerson, S., Carson, J., 2018. All that glitters is not grit: three studies of grit in university students. *Front. Psychol.* 9. <https://doi.org/10.3389/fpsyg.2018.01539>.
- Kannangara, C., Allen, R., Carson, J., Khan, S., Waugh, G., Kandadi, K., 2020. Onwards and upwards: the development, piloting and validation of a new measure of academic tenacity. *The Bolton Uni-Stride Scale (BUSS)*. *PLoS One* 15 (7), e0235157.
- Kannangara, C.S., Allen, R., Hochard, K., Carson, J., 2022. An international validation of the Bolton Unistride Scale (BUSS) of tenacity. *PLoS One* 17 (3), e0264889. <https://doi.org/10.1371/journal.pone.0264889>.
- Klohn, E.C., Vandewater, E.A., Young, A., 1996. Negotiating the middle years: ego-resiliency and successful midlife adjustment in women. *Psychol. Aging* 11, 431–442. <https://doi.org/10.1037/0882-7974.11.3.431>.
- Lee, W.W., 2017. Relationships among grit, academic performance, perceived academic failure, and stress in associate degree students. *J. Adolesc.* 60, 148–152. <https://doi.org/10.1016/j.adolescence.2017.08.006>.
- Maddi, S.R., Matthews, M.D., Kelly, D.R., Villarreal, B., White, M., 2012. The role of hardiness and grit in predicting performance and retention of USMA cadets. *Mil. Psychol.* 24 (1), 19–28. <https://doi.org/10.1080/08995605.2012.639672>.
- Matthews, M.D., 2020. The link between grit and resilience. *Psychology Today*. <https://www.psychologytoday.com/us/blog/head-strong/202012/the-link-between-grit-and-resilience>.
- McIntosh, E., Shaw, J., 2017. Student Resilience: Exploring the Positive Case for Resilience. Unite Students, London.
- Meyer, G., Shatto, B., Kuljeerung, O., Nuccio, L., Bergen, A., Wilson, C.R., 2020. Exploring the relationship between resilience and grit among nursing students: a correlational research study. *Nurse Educ. Today* 84, 104246.
- Montas, M., Rao, S.R., Atassi, H.A., Shapiro, M.C., Dean, J., Salama, A.R., 2021. Relationship of grit and resilience to dental students' academic success. *J. Dent. Educ.* 85 (2), 176–186.
- Musso, M., Tatum, D., Hamer, D., Hammarlund, R., Son, L., McMahon, P., 2019. The relationship between grit and resilience in emergency medical service personnel. *Ochsner J.* 19 (3), 199–203.
- Mwangi, C.N., Okatcha, F.M., 2015. Relationship between academic resilience and academic achievement among secondary school students in Kiambu County, Kenya. *Int. J. Sch. Cogn. Psychol.* 2 (S2). <https://doi.org/10.4172/2469-9837.1000s2003>.
- Oriol, X., Miranda, R., Oyanedel, J.C., Torres, J., 2017. The role of self-control and grit in domains of school success in students of primary and secondary school. *Front. Psychol.* 8, 1716.
- Palen, S., Schitter, C., 2018. Prolific.ac: A subject pool for online experiments. *Behav. Exp. Finance* 17, 22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>.
- Pate, A.N., Payakachat, N., Harrell, T.K., Pate, K.A., Caldwell, D.J., Franks, A.M., 2017. Measurement of grit and correlation to student pharmacist academic performance. *Am. J. Pharmaceut. Educ.* 81, 105. <https://doi.org/10.5688/ajpe816105>.
- Perkins-Gough, D., 2013. The significance of grit: a conversation with Angela Lee Duckworth. *Educ. Leader.* 71 (1), 14–20.
- Pidgeon, A.M., Rowe, N.F., Stapleton, P., Magyar, H.B., Lo, B.C.Y., 2014. Examining characteristics of resilience among university students: an international study. *Open J. Soc. Sci.* 2, 14–22.
- Platt, I., Kannangara, C., Tytherleigh, M., Carson, J., 2020. The Hummingbird Project: a positive psychology intervention for secondary school students. *Front. Psychol.* 11. <https://doi.org/10.3389/fpsyg.2020.02012>.
- Rashid, T., Seligman, M., 2018. Positive Psychotherapy: Clinician Manual. Oxford University Press.
- Rimfeld, K., Kovas, Y., Dale, P.S., Plomin, R., 2016. True grit and genetics: predicting academic achievement from personality. *J. Pers. Soc. Psychol.* 111 (5), 780–789. <https://doi.org/10.1037/pspp0000089>.
- Robertson-Kraft, C., Duckworth, A.L., 2014. True grit: trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teach. Coll. Rec.* 116 (3).
- Salles, A., Cohen, G.L., Mueller, C.M., 2014. The relationship between grit and resident well-being. *Am. J. Surg.* 207 (2), 251–254. <https://doi.org/10.1016/j.amjsurg.2013.09.006>.
- Sharkey, C.M., Bakula, D.M., Gamwell, K.L., Mullins, A.J., Chaney, J.M., Mullins, L.L., 2017. The role of grit in college student health care management skills and health-related quality of life. *J. Pediatr. Psychol.* 42 (9), 952–961. <https://doi.org/10.1093/jpepsy/jsx073>.
- Singh, K., Jha, S.D., 2008. Positive and negative affect, and grit as predictors of happiness and life satisfaction. *J. Indian Acad. Appl. Psychol.* 34 (Special Issue), 40–45.
- Solomon, O., 2013. Exploring the Relationship Between Resilience, Perceived Stress and Academic Achievement. Retrieved. [https://espace.mmu.ac.uk/576570/1/Olivia%20SOLOMON%20\(MMU\).pdf](https://espace.mmu.ac.uk/576570/1/Olivia%20SOLOMON%20(MMU).pdf). (Accessed 8 February 2018).
- Souri, H., Hasanirad, T., 2011. Relationship between resilience, optimism and psychological well-being in students of medicine. *Procedia Soc. Behav. Sci.* 30, 1541–1544.
- Stallman, H.M., 2010. Psychological distress in university students: a comparison with general population data. *Aust. Psychol.* 45, 249–257. <https://doi.org/10.1080/00050067.2010.482109>.
- Stoffel, J.M., Cain, J., 2018. Review of grit and resilience literature within health professions education. *Am. J. Pharmaceut. Educ.* 82 (2), 6150. <https://doi.org/10.5688/ajpe6150>.

- Storrie, K., Ahern, K., Tuckett, A., 2010. A systematic review: students with mental health problems—a growing problem. *Int. J. Nurs. Pract.* 16, 1–6. <https://doi.org/10.1111/j.1440-172X.2009.01813.x>.
- Strauss, K., Griffin, M., Parker, S., 2012. Future work selves: how hoped for identities motivate proactive career behaviours. *J. Appl. Psychol.* 97 (3), 580–589.
- Tagney, J., Baumeister, R., Boone, A., 2004. High self-control predicts good adjustment, less pathology, better grades and interpersonal success. *J. Pers.* 72 (2), 271–324. <https://doi.org/10.1111/j.0022-3506.2004.00263.x> PMID: 15016066.
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Welch, S., Parkinson, J., Secker, J., Stewart-Brown, S., 2007. The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): development and validation. *Health Qual. Life Outcomes* 5, 63.
- Turner, M., Scott-Young, C.M., Holdsworth, S., 2017. Promoting wellbeing at university: the role of resilience for students of the built environment. *Construct. Manag. Econ.* 35 (11–12), 707–718.
- Vago, D.R., Silbersweig, D.A., 2012. Self-awareness, self-regulation, and self-transcendence (S-ART): a framework for understanding the neurobiological mechanisms of mindfulness. *Front. Hum. Neurosci.* 6, 1–30. <https://doi.org/10.3389/fnhum.2012.00296>.
- Vainio, M.M., Daukantaite, D., 2015. Grit and different aspects of well-being: direct and indirect relationships via sense of coherence and authenticity. *J. Happiness Stud.* 17 (5), 2119–2147. <https://doi.org/10.1007/s10902-015-9688-7>.
- Wagnild, G.M., Young, H.M., 1993. Development and psychometric evaluation of the resilience scale. *J. Nurs. Meas.* 1 (2), 165–178.
- Wang, J., 2009. A study of resiliency characteristics in the adjustment of international graduate students at American universities. *J. Stud. Int. Educ.* 13 (1), 22–45.
- Waxman, H.C., Gray, J.P., Padron, Y.N., 2003. Review of Research on Educational Resilience. Research Report. Institute of Educational Sciences, Washington, DC.
- Weisskirch, R.S., 2016. Grit, self-esteem, learning strategies and attitudes and estimated and achieved course grades among college students. *Curr. Psychol.* <https://doi.org/10.1007/s12144-016-9485-4>.
- Windle, G., Bennett, K., Noyes, J., 2011. A methodological review of resilience measurement scales. *Health Qual. Life Outcomes* 9 (8), 1–18.
- Wolters, C.A., Hussain, M., 2014. Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacogn. Learn.* 10 (3), 293–311. <https://doi.org/10.1007/s11409-014-9128-9>.

Intelligences

John D. Mayer and Victoria M. Bryan, Department of Psychology, University of New Hampshire, Durham, NH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	198
Our coverage	199
One general intelligence or many intelligences? The origins of the debate: 1990–1949	199
Spearman's <i>g</i>	199
Binet's contribution	199
Thurstone's primary mental abilities and the advent of factor analysis	200
One more intelligence ... Thorndike's social intelligence	200
More extremely opposed than warranted?	201
The role of early mathematical models of intelligence	201
Basic models	201
Understanding a factor model of <i>g</i>	201
Primary mental abilities (intelligences in the plural)	201
Hierarchical (second order) models	202
Further developments, debate, and integration: 1950–1999	203
Cattell and Horn's Gf-Gc model: multiple manifestations of <i>g</i>	203
Howard Gardner's multiple intelligences	203
Mathematical models renewed	204
Comparative model fitting	204
John Carroll and a key outcome of comparative model fitting	204
What do we know of these multiple intelligences today?	205
What are the recognized broad intelligences of today?	205
Still newer intelligences?	205
Relations among broad intelligences	205
Organizing the broad intelligences	205
Implications for educators and for education	206
Conclusions and reflections	206
Consistencies across the past years	206
Considerations regarding education	206
References	207

Introduction

The capacity of people to learn, to be educated, and to think intelligently became a focus of attention during the 19th century as nations industrialized and became more dependent upon the skills of their citizenry. National education systems emerged, and with them, decisions were needed as to who could best benefit from educational opportunities at schools and universities (Anderson, 1982). Alfred Binet's first intelligence tests were introduced in response to the French government's call for help regarding the need to identify students with intellectual disabilities in a fair and impartial manner (Binet and Simon, 1916; Fancher, 1987). Governments wanted to know who could best take advantage of higher education and could use their learning to contribute to society.

Binet and others recognized the need for impartiality in using newly-developed indices of mental ability (Devonis, 2016; Fancher, 1987), concerns that extend to the present day (Camilli, 2013; Deary, 2001). Scientists used their new intelligence measures, and the theories around them, to understand people's potential at solving problems. *Intelligence*, as used here, refers to people's capacity to problem solve. The title of this entry is "*Intelligences*", in the plural, and we take as our role explaining how experts gradually came to accept and understand the multifaceted aspects of human reasoning recognized today.

During much of the 20th and early 21st centuries, anyone hoping to discuss intelligences in the plural would need to defend their conception that intelligence was multifaceted rather than unitary. As recently as 1994, in a Wall Street Journal editorial entitled "Mainstream Science on Intelligence," Linda Gottfredson and 52 co-signers emphasized that intelligence was a single, unitary and "... very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (reprinted as Gottfredson, 1997, p. 13). However, mainstream science on the nature of intelligence was shifting even as the editorial was written.

Our coverage

The debate over the existence of a single general intelligence versus multiple intelligences extends over a century of theoretical and empirical work. The approach here is to trace that debate and to explain some of the research challenges that made it so difficult to resolve. Our history begins with the section “One General Intelligence or Many ...?”, and covers the first half of the 20th century, followed by “The Role of Early Mathematical Models of Intelligence”, a consideration of how research findings were modeled and interpreted at the time by members of one side of the debate or the other. We then continue with the section “Further Developments, Debate, and Integration” which covers from 1950 to 1999, and is followed by a second section of “Mathematical Models Renewed” that explains how, in the early 1990s, mathematical advances helped integrate the two competing perspectives.

We conclude by covering the contemporary understanding of broad intelligences and, very briefly, their implications for education. Although increasing evidence exists for intelligences in the plural, most school curricula already are organized somewhat by subjects that correspond to those specialized mental abilities. And, although creative innovations have been recommended to address people’s individual strengths and weaknesses, relatively little well-controlled empirical evidence exists on the effectiveness of such programs. As some theoretical positions point out, intelligences can be thought of as a system of parts that operate together (Detterman, 1982).

One general intelligence or many intelligences? The origins of the debate: 1990–1949

Does human intelligence properly consist of one general intelligence to understand the world, or multiple, partially-distinct intelligences—that is, “intelligences” in the plural (Lubinski, 2004; Spearman, 1904; Thurstone, 1938). If there were one overall general intelligence, or *g*, then a single IQ could sum up a person’s intellectual prowess; if there were many, somewhat distinct mental abilities, then a more nuanced approach would be needed.

Spearman’s *g*

Working at the start of the 20th century, Charles Spearman initiated the debate over one-versus-many intelligences—pointing to evidence for the unitary quality of mental abilities. Spearman’s studies were inspired by the work of Sir Francis Galton. Galton had persuaded nine-thousand visitors to the 1884 International Health Exhibition at the South Kensington Museum in London, to pay 3 pence each to take a battery of measurements he believed might indicate intelligence—from head size and reaction time, to sensory acuity (color sense, good eyesight; Fancher, 1987, pp. 41–42).

Spearman took a somewhat different starting point. Visiting local schools, he indexed students by their class rank in such subjects as classics, language, and math; he also administered music and other sensory tasks. He next computed correlation coefficients among the students’ various class ranks and noticed that all the correlations were positive—a phenomenon that later became a signature feature of the relations among mental abilities, termed the *positive manifold*; there also were positive correlations with the sensory measures he administered. These correlations indicated that an individual’s mental abilities rose and fell together. That is, if a child were high on one mental ability, they tended to be high on all the others.

Spearman (1904) further suggested that any matrix of correlated test scores could be divided into two parts “the one great common Intellective Function” —the pure, general, rise and fall of ability across tasks, which he termed *g*—and the specific variability due to individual tasks (Lovie and Lovie, 1993, p. 310). In essence, Spearman developed a model in which he argued for the existence of two “levels”—a general intelligence at the peak to account for the common scores being high for some students and low for others, and, beneath that general capacity, the specific skills assessed by each of the individual tasks themselves.

Objections to the idea of a general intelligence were raised almost immediately, however, and the idea of one single “engine” behind problem solving (Spearman likened intelligence to a steam engine) seemed potentially objectionable.

Binet’s contribution

Across the English Channel in Paris, Spearman’s contemporary, Alfred Binet, set to work on how to improve the measurement of intelligence of school children (Binet and Simon, 1915). Public education in France had expanded greatly in the 1800s (Anderson, 1982); that expansion introduced questions such as which students could reasonably be educated in public schools and which students, owing to intellectual disabilities, might require special education. Working at the behest of an educational commission established by the French government, Binet constructed the first valid measures of school aptitude. His new tests, developed with Theophile Simon and Victor Henri, focused on children’s capacity to conceptualize, to reason, to compare objects, and similar tasks expected of them at school (Binet and Simon, 1915). Binet’s team refocused intelligence measurement away from features that Sir Francis Galton examined, such as head size and reaction time, to what students actually did in school. The teams’ measures included memory for concepts, verbal knowledge of objects, repetition of (i.e., memory for) numbers, placing objects in order of their weight, comparison of weights, and the ability to define abstract terms.

Although Binet conceptualized intelligence as focused on higher-level abstract reasoning, he believed that it might be both multifaceted and malleable (Beaujean and Benson, 2019, p. 201; Haney, 1984, p. 602). But Binet had a job to do: to develop a measure simple enough to be used by non-psychologists in school settings. And so he and his colleagues designed his tests—despite their varied subscales—to produce one easy-to-use overall score. The variety of scales employed on the test, however, inspired other researchers to look more closely at whether the diverse tasks might plausibly indicate more than one ability (Boake, 2002).

Thurstone's primary mental abilities and the advent of factor analysis

L.L. Thurstone, then at the University of Chicago, expanded on the tasks laid out by Binet, Henri, and Simon. Thurstone developed a group of diverse intelligence tasks, administering them to large samples of test-takers. To analyze his data, Thurstone introduced new approaches to *factor analysis*, a family of mathematical modeling techniques that could be used for evaluating correlation matrices. Factor analysis was to prove foundational not only to understanding the relations among intelligences, but to measurement in psychology more generally. We will turn to factor analysis shortly.

Whereas Spearman interpreted his matrices of correlations to reflect one general intelligence factor and many specific task factors, Thurstone proposed a set of seven independent intelligences he referred to as “primary mental abilities,” which included Reasoning, Perceptual Speed, and Word Fluency, among others listed in Table 1 (left-side column). Thurstone found evidence in support of his proposed primary mental abilities from a sample of college students who were relatively comparable in their IQ scores but who nonetheless differed in their respective scores on the primary mental abilities (Thurstone, 1938). Today, multiple, related mental abilities such as Thurstone examined are referred to as “broad intelligences.”

One more intelligence ... Thorndike's social intelligence

Beyond the central figures of the one-versus-multiple intelligence debate, other thinkers proposed additional candidates for intelligences. In 1920, Edward L. Thorndike proposed that there existed an intelligence for understanding people that he called *social intelligence* that involved “the ability to understand and manage men and women, boys and girls, and to act wisely in human

Table 1 Proposed broad intelligences.

Generic classes of broad intelligences	Thurstone's (1938) primary mental abilities	Gardner's (1983) multiple intelligences ^a	Regularly-studied intelligences ^b from Carroll's model (1993)	Potential additions and/or pending (various)
Abilities commonly regarded as more fluid				
Fluid intelligence	—	—	Fluid intelligence	—
General reasoning	Reasoning	—	—	—
Speededness	Perceptual speed	—	Processing speed	—
Working memory	—	—	Short-term memory	—
Abilities commonly regarded as more crystallized				
Verbal comprehension	Word fluency	Linguistic-verbal	Comprehension knowledge	—
	Verbal comprehension		Reading and writing	
Visuo-spatial	Spatial visualization	Visual-spatial	Visuo-spatial processing	—
Mathematical-quantitative	Number facility	Logical-mathematical	Quantitative reasoning	—
Long-term memory retrieval	Associative memory		Long-term retrieval	—
Other sensory	—	Bodily-kinesthetic	—	Tactile-kinesthetic ^c
		Musical	Auditory intelligence	—
Creativity			Creative intelligence ^d	
Practical	—	—	—	Practical
People-centered ^d	—	Intrapersonal	Emotional intelligence	Personal intelligence ^e
	—	Interpersonal	—	Social intelligence ^f

^aAlthough most the other classifications are based on psychometric tests, Gardner's group were determined by his logical argument that each intelligence met eight criteria, such as having plausible bases in human evolution, an associated brain structure (isolation by brain damage), savants in the area, and also including psychometric criteria.

^bThe most regularly studied broad intelligences were compiled more recently (Bryan and Mayer, 2020). Although not a part of the original review because few studies employed it, creative intelligence also is included owing to its theoretical place in the original CHC and other models. Emotional intelligence was confirmed as a member more recently (MacCann et al., 2014).

^cBodily-kinesthetic (Roberts et al., 1997).

^dPeople-centered is a class suggested by Mayer to encompass emotional, personal, and social intelligences.

^ePersonal intelligence (an intelligence about personality) (Mayer et al., 2012, 2019).

^fSocial intelligence (e.g., Thorndike, 1920; Weis and Süß, 2007).

^gIncluded under retrieval fluency in Carroll (1993).

relations" (Thorndike, 1920; Thorndike and Stein, 1937). Several attempts at measuring this intelligence were made, but with equivocal success—the ability appeared to blend imperceptibly into general intelligence (Hunt, 1928; Thorndike and Stein, 1937; Walker and Foley, 1973). Surprisingly, a good measure of social understanding has yet to be developed although progress continues (Weis and Süß, 2007; Conzelmann et al., 2013), and it therefore appears in the "Proposed/Pending" column to the right in Table 1.

More extremely opposed than warranted?

With the advantage of hindsight, we observe that both Spearman and Thurstone made unintentional errors in their work that exaggerated their differences. Spearman, for example, had derived a (correct) formula to adjust correlations upward so as to compensate for errors in measurement—now known as the *correction for attenuation*. The formula was a useful one, however, Spearman appeared to have misapplied it initially, leading him to believe that the many positive correlations he witnessed would be near $r = 1.0$ once error was corrected. This position was widely criticized at the time and in fact overstated the case (Fancher, 1987, pp. 90–92). Thurstone, on the other hand, used samples of students who were so similar in their ability as to reduce the magnitude of the correlations he observed—a phenomenon now known as restriction of range. This led him to believe that the correlations among mental abilities were lower than they actually were. He swiftly corrected his viewpoint once the error became clearer to him (Haier, 2016, p. 34; Thurstone and Thurstone, 1941). By the 1940s, the correlations among mental abilities were generally understood to be moderate—residing in a gray area between the one and many—and making the resolution of the issue challenging.

The role of early mathematical models of intelligence

Basic models

In the debate over single versus multiple intelligence in the early 20th century, psychologists generally argued on the basis of (a) mental test data that they or others had collected and (b) applied mathematical models from a family of techniques called "factor analysis" to the data. To understand the debate over one versus many intelligences, therefore, it helps to have some introductory knowledge of how factor analysis operates, and we therefore pause our historical review of the debate here to describe some of the basic techniques used through the 1950s.

Factor analysis is a method developed to examine and interpret the many correlations among intelligence tasks at once (and its use has broadened to many other areas of psychology since). The technique groups together variables that are (relatively) highly correlated with one another and separates other variables that are less related. The key protagonists in this debate generally assembled a group of mental ability tests, administered it to a sample of research participants, and then analyzed the data.

As a first step in such a data analysis, a researcher will calculate a matrix of correlations among people's scores on the mental ability tasks. Here, we present a correlation matrix employing just four mental tasks so as to keep our example simple, but more such tasks are desirable when conducting actual research. The starting matrix is shown toward the top of Fig. 2 in the "Step 1" column, where the correlations among the tasks range between $r = 0.35$ and 0.48 . (The Fig. 1 data provide a reasonable facsimile of actual findings). Factor analysis is then employed to evaluate the matrix and to sort test scores into groups called *factors*.

Understanding a factor model of g

Because the four mental tasks represented in the first correlation matrix in Fig. 1 (under the "One-Factor Case") exhibit moderate, positive correlations with one another, their relations suggest there is just one general intelligence—just as Spearman had speculated. We have applied factor analysis to the correlation matrix to carry out the second step. This step uses matrix algebra to reorganize the correlations in a different form, indicating which tasks might cluster together, and which are distinct. The results of the analysis show a single general factor, indicated under the Roman numeral I (Fig. 1, Step 2 column). Each number beneath the Roman numeral represents the correlation between the original test scores of T1 through T4 with the obtained factor for that particular sample of people. This same result can be represented pictorially and is indicated in the final panel to the right. There, the observed four tasks are represented as squares (directly measured scores) and the general factor is represented as an oval (a latent, i.e., inferred, variable). The lines connecting them and the numbers beside them represent the four loadings (i.e., correlations) between the task and the factor.

Primary mental abilities (intelligences in the plural)

The next row down of Fig. 1 proceeds through the same steps. This time, however, the starting correlation matrix shows that Tasks 1 and 2 are correlated $r = 0.48$, and Tasks 3 and 4 are correlated $r = 0.45$, but Tasks 1 and 2 are relatively uncorrelated with Tasks 3 and 4. If Thurstone saw this, he would say there were two factors: The first representing Tasks 1 and 2, and the second representing Tasks 3 and 4. And, indeed, the factor analysis yields two independent factors (Fig. 1, middle). Tasks 1 and 2 correlate with (load on) Factor I, $r = 0.69$ and 0.69 , but not appreciably on Factor II. Tasks 3 and 4 correlate with (load on) Factor II, $r = 0.67$ and 0.67 , but not on Factor I. A simple pictorial version of the model is indicated to the right. These simplifications omit some information, including the possible correlations between Factors I and II as well as error terms that indicate degrees of measurement precision.

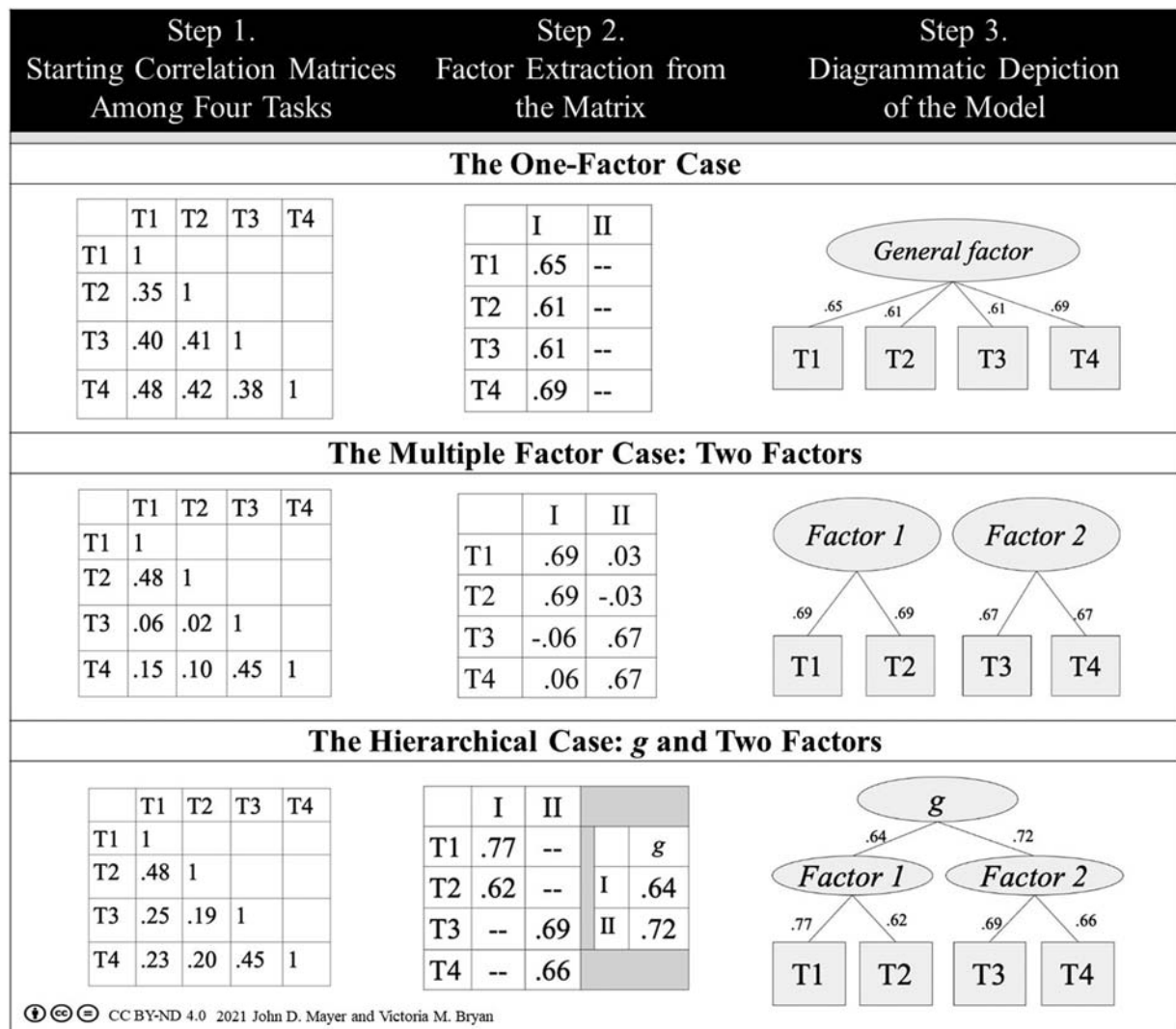


Fig. 1 From correlation matrices to factor models The variables T1 through T4 of this example represent four mental tasks administered to 1000 participants. Each row begins with a different (hypothetical) correlation matrix that illustrates a given relationship among the mental tasks: A one-factor, two-factor, and two-factor hierarchical case are indicated beneath each of the row headings in the first column. Each of the correlation matrices were subjected to a factor analysis. In the Step 2, “Factor Extraction” column (middle column), the roman numerals designate the first and second factors extracted for the one- and two-factor cases. The numbers beneath the roman numerals represent the estimated correlations between the four individual tasks and the two factor(s); those correlations are referred to as factor loadings. The final “Diagrammatic” column to the right represents a pictorial depiction of the results. Several further details for those interested are that the one- and two-factor solutions were both exploratory factor analyses employing principal factor extractions, the two-factor version employed an oblimin rotation, and the hierarchical model was a confirmatory factor analysis. All analyses were carried out in R. The figure is reproduced from an online version that is accompanied by further details of the analyses (Mayer and Bryan, 2021).

To carry out such analyses in the first part of the 20th century often required extensive investment of time and resources; in the better-funded laboratories, graduate students employed mechanical calculators to obtain the correlations they and their professors needed, and then the research team carried out transformations of the matrices to come up with a particular solution; sometimes a mechanical geometric apparatus was constructed to further transform the correlation matrices.

Hierarchical (second order) models

Perhaps the above tedious work discouraged most laboratories from considering anything more complex. There were, however, more adventurous souls who suggested that one could “stack” factors in a hierarchy. [Thurstone \(1944\)](#) developed an idea of “second order factors”, which [Schmid and Leiman \(1957\)](#) further refined and relabeled a *hierarchical* model—the name that stuck. Hierarchical models permitted intelligence researchers to obtain a general factor that corresponded to general intelligence

(Spearman's g), and then to subdivide it into a second set of factors that represented broad intelligences. These broad intelligences were defined at a lower level by the specific indicators (mental ability tasks), T1 to T4 in this example, of each broad intelligence. This hierarchical situation is depicted in Fig. 1, bottom row.

In the starting matrix, Tasks 1 and 2 still correlate highly with one another, and Tasks 3 and 4 also correlate with each other; notice, however, that now we have introduced moderate positive correlations among the tasks throughout the matrix. Factor analyzing that matrix yields evidence for both an overall g and two broad intelligences. The model is specified in some detail in advance, which includes setting loadings to zero where none are expected. The factor solution now shows the loadings for Factors I and II (as before), and also the loadings of the two factors on a more general factor. The diagram representing the result in the far-right panel indicates that there exists a general intelligence factor (g), under which is subsumed two broad factors (I and II). The problem remained, however, that some researchers obtained robust one-factor solutions, others, multiple factors, and rather few researchers tried the more demanding second-order, hierarchical solutions.

Further developments, debate, and integration: 1950–1999

Cattell and Horn's Gf-Gc model: multiple manifestations of g

Raymond B. Cattell, a student of Spearman, was well-versed in applying factor analysis to data of personality generally, and mental abilities specifically. Cattell speculated that, subtly embedded in many correlational patterns among intelligence tasks might be evidence that g divided into two subsidiary, highly correlated factors that were, "so highly cooperative, i.e., so similar in most loadings" that earlier factor analytic methods had missed them (Cattell, 1961, p. 742).

One of the two factors, Cattell argued, loaded on tasks like inductive reasoning and numerical ability, and appeared to represent a flexibility or fluidness in people's reasoning, independent of a person's education or culture; Cattell labeled this *fluid general ability* (Gf). The other factor, which loaded on tasks pertaining to verbal ability and stored knowledge acquired through education and culture, Cattell labeled as *crystallized general ability* (Gc). John Horn, a student of Cattell, expanded Gf-Gc theory to include additional aspects of general reasoning that emerged from intelligence tests, such as general fluency, which pertained to recall and recognition, general speediness (Gs), which involved the ability to rapidly perceive relations or solve problems, and general visualization (Gv) which involved mentally manipulating and transforming the spatial patterns among objects (Cattell and Horn, 1978; Horn and Cattell, 1966). Cattell and Horn's view was influential in the field and would be raised again in the 1990s.

Howard Gardner's multiple intelligences

The debate over multiple intelligences was a matter chiefly carried out by intelligence researchers studying intelligence tasks and using factor analysis. But in the early 1980s, Howard Gardner introduced a novel theoretical view of intelligence with his book entitled "Frames of Mind" (Gardner, 1983), in which he employed eight criteria to determine what was (and was not) an intelligence. Among the criteria were that symbolization was involved, that evolutionary pressures would require the intelligence for survival, and that brain areas could be isolated for the mental ability as a consequence of brain damage, or in cases of savants. Although intelligence testing was on Gardner's list as well, it appeared last (and in Gardner's view, perhaps, least).

To argue for the existence of linguistic-verbal intelligence, for example, Gardner described poetry (and language more generally) as operating with symbols; he detailed the often challenging, trial-and-error process that the choice of the right word can require; and he wrote of the need to understand semantics, syntax, phonology, and the core operations of language in communication. Gardner profiled the work that poets sometimes engaged in to compose their writings, drawing in part on the correspondence and notes of T. S. Eliot (Gardner, 1983, p. 77). He used examples of hyperlexic children, who learn to read at two or three years of age, and the localization of syntax understanding in part to Broca's area of the brain. Collectively, these and other topics allowed him to build a case for the intelligence in an argument that is largely independent of psychometric findings.

The remainder of Gardner's intelligences included visuo-spatial, logical mathematical, bodily-kinesthetic, musical, and intra- and interpersonal (see Table 1). Although he used substantially different criteria from those of other psychologists, Gardner arrived at a list much the same as others had proposed (other than, perhaps, his bodily-kinesthetic). More recently, some researchers have continued to localize brain functions relevant to Gardner's (and other intelligence researchers') lists (e.g., Shearer, 2020).

Yet Gardner made two arguments that were viewed as untenable to those who regularly worked in the area. First, although Gardner acknowledged that a general intelligence might exist, he regarded his multiple intelligences as "not subsumed under g "—implying that what he had identified was not a part of (or necessarily correlated with) the rest of human intelligence (Gardner, 1999, p. 87)—a claim that flew in the face of a large and consistent database of findings (both then and now) that showed otherwise. Second, Gardner contended that "having a battery of ... tests is not consistent with the major tenets of the theory," thereby actively discouraging any empirical tests of his approach. Although Gardner had his reasons for staking this anti-testing position (concerns over the potential misuse of his theory; Gardner, 1999, pp. 80–81), it rendered his proposals unverifiable by any methods of the day. But science requires collecting data relevant to theoretical positions and interpreting that data (e.g., Strevens, 2020), so we return to such work next, examining the continued progress of research into general intelligence and its constituent intelligences.

Mathematical models renewed

Comparative model fitting

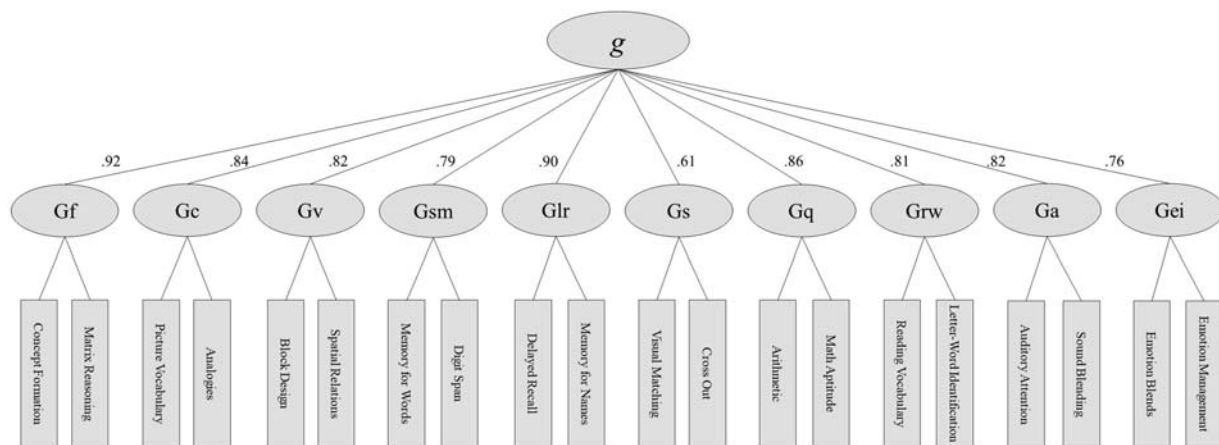
With factor models fairly-well envisioned, and the major models of intelligence relevant to 20th-century research in place, a further issue remained: The initial techniques of factor analysis could not settle the one-versus-many debate. Different intelligence tests often correlated at a moderate level of about $r = 0.25\text{--}0.45$. Such correlational levels were neither perfectly correlated nor entirely independent of one another.

Given the mathematical techniques of the day, researchers of the 1950s and 1960s often were forced into one of two positions—which they could find evidence for depending on the kind of factor analysis they used and the nature of their data set. The first position was that mental ability tests reflected a pure general intelligence accompanied by a kind of residual sludge of error due to specific tasks. The second position was that each distinct class of intelligence tasks represented a truly independent intelligence, and that these intelligences appeared to correlate among themselves somewhat, perhaps due to some common artifact such as the test-taker's understanding of written or spoken instructions. [Gardner \(1993, p. 17\)](#) commented, "... given the same set of data ... using one set of factor-analytic procedures ... supports the idea of a 'g' factor; using another ... it is possible to support ... relatively discrete mental abilities," and no resolution was in sight. But change would arrive faster than Gardner had anticipated.

An effort had emerged across several disciplines beginning in the 1910s to build mathematical models that could describe the relations among multiple, correlated variables; these early models, which included path analysis, evolved into a broader set of techniques that ultimately would be integrated with the work of factor analysts ([Matsueda, 2012](#)). Those who developed the techniques hoped not only to model complex relations but also to develop statistical tests for *how well* a given model fit the data. In the late 1960s, several psychometricians solved the basic problem of assessing the fit of a given factor model ([Jennrich and Robinson, 1969](#); [Joreskog, 1969](#)). Their approach, under the label Structural Equation Modeling (SEM), provided powerful new methods for modeling data ([Boomsma, 2000](#); [Matsueda, 2012](#)). Factor analysts now could compare factor models to select the models that fit data the best, and discard the rest. Although these innovations predated Gardner's work by 10-plus years, the complexity of the techniques, their unfamiliarity, and the relative unfriendliness of the initial software for carrying out the analyses, meant that the use of SEM filtered slowly through the field at first. Nonetheless, intelligence researchers and researchers in many other fields began to employ the technique regularly, and as the software improved, ever-more sophisticated models could be tested on data.

John Carroll and a key outcome of comparative model fitting

In 1993, John Carroll published a review and reanalysis of 400 datasets from intelligence research throughout the 20th century using SEM. For a number of years, Carroll had visited intelligence laboratories at many universities and asked the researchers there for any data sets they might share that included correlations among specific intelligence tasks for large samples of test-takers (The first author was present at one such in-person visit to Douglas Detterman's laboratory at Case Western Reserve University around 1980). Carroll compiled the correlational tables he collected and subjected them to a series of factor models, testing the fit of the models using confirmatory factor analysis techniques from SEM. The result was a three-tiered model later termed the Cattell-Horn-Carroll (CHC), or three-stratum model, that arranged general intelligence at the top, eight broad intelligences (akin to Thurstone's primary abilities and including Cattell and Horn's Gf and Gc) at the second tier, and specific tasks below. Fairly quickly, researchers regarded the CHC model as having ended the debate over whether there were multiple intelligences, or just one, and tied it up in a rather elegant bow: Both positions were correct, and they could be integrated in a single hierarchical model. We illustrate one such version of the model in [Fig. 2](#), drawing on a recent meta-analysis of correlations among intelligence tests ([Bryan and Mayer, 2020](#)).



© CC BY-ND 4.0 2021 Victoria M. Bryan and John D. Mayer

Fig. 2 The relations of broad intelligences to g *Figure Note:* Gf = fluid intelligence; Gc = Comprehension Knowledge; Gv = Visuospatial Processing; Gsm = Short-Term Memory; Glr = Long-Term Retrieval; Gs = Processing Speed; Gq = Quantitative Reasoning; Grw = Reading and Writing Ability; Ga = Auditory Intelligence; Gei = Emotional Intelligence. Reproduced from [Bryan and Mayer \(2021a\)](#).

What do we know of these multiple intelligences today?

With the growing awareness of Carroll's and others' contributions to model-fitting, the field changed markedly. By 2001, Ian Deary was able to conclude regarding the CHC model:

... there is consensus that there is a general cognitive factor that accounts for about 50% or so of the variance in a broad assembly of mental tests given to a large sample of the population ... there is still variance attributable to separable "group" factors of intelligence. Beyond the general factor and group factors some test performance variance ... is manifest in very specific mental abilities. This so-called three-stratum model affords freedom from otiose arguments about being "for or against the general factor."

Deary (2001, p. 127).

Intelligences—or *broad intelligences*, using today's terminology—are a now well-established part of understanding how people solve problems.

What are the recognized broad intelligences of today?

Bryan and Mayer (2020) reviewed 61 studies that contained intelligence test data from 20,498 participants. From those studies, they concluded that there had been 10 widely-studied broad intelligences from the decades surrounding Carroll's integrative model (see Fig. 2). These 10 included the now familiar instances of fluid intelligence, comprehension knowledge (i.e., verbal-comprehension), visuospatial processing, short-term memory retrieval, processing speed, quantitative reasoning, and more recently-proposed intelligences described next.

Still newer intelligences?

Carroll's integration of broad intelligences with *g* did not stop the search for further areas of intelligence. For example, Mayer and Salovey proposed an ability model of emotional intelligence (Mayer et al., 2016; Salovey and Mayer, 1990) and recent research provides evidence in favor of its status as a broad intelligence (MacCann et al., 2014). More provisionally, there may also exist a *personal* intelligence—an intelligence about personality—that appears partially distinct from emotional intelligence (Mayer, 2008; Mayer et al., 2019); and other researchers have picked up on Gardner's idea of a kinesthetic intelligence, finding some supportive evidence (Roberts et al., 1997).

Another group of intelligences suggested by Robert Sternberg included the analytical, creative, and practical (Sternberg, 1985; Wagner, 2011)—of which analytical and creative already appeared in the CHC model, but not the practical. The kinesthetic, personal, and practical intelligences appear in the "Potential Additions" right-most column of Table 1 toward the bottom (Mayer et al., 2012, 2019; Roberts et al., 1997).

Relations among broad intelligences

While the number of mental abilities considered at the second stratum of the CHC model paints a nuanced picture as to the breadth of human intelligence, some researchers have expressed concern over their large number. Hedlund and Sternberg (2000) asked whether there might be "Too many intelligences?" That surely would be the case if the broad intelligences correlated too highly among themselves. But a recent meta-analysis of the area determined that the average estimated correlations among the broad intelligences across two alternative factor models used in the literature were $r = 0.58$ and $r = 0.62$, or about $r = 0.60$ overall (Bryan and Mayer, 2020). This level allowed for distinctions among multiple, meaningfully different intelligences. The exact number of intelligences—awkwardly large though it sometimes may seem—appears to be more a neuropsychological consequence of human evolution than of researcher preference.

Organizing the broad intelligences

But with new intelligences proposed, is there some way to organize or simplify these broad intelligences? In the same meta-analysis discussed above, Bryan and Mayer (2020) refactored one of their composite correlation tables and re-encountered an older finding: the intelligences did indeed seem to divide according to their more-fluid versus crystallized qualities, as pointed out earlier by Cattell and Horn (Cattell, 1963). The fluid-intelligence factor loaded on fluid intelligence itself, with visuospatial processing, and processing speed; the crystallized abilities included long-term retrieval, reading and writing, and emotional intelligences.

Additional divisions have been proposed as well. Mayer suggested a possible continuum of content-focused broad intelligences from a group of "thing-centered" intelligences typified by visuo-spatial and quantitative reasoning abilities, to a group of "people-centered" intelligences, typified by emotional, personal, and social intelligences (Mayer, 2018). A second meta-analysis covering a set of 87 studies of thing- and people-centered intelligences indicated the empirical viability of a people-versus-thing continuum of reasoning as applied to the broad intelligences of the CHC model (Bryan and Mayer, 2021b).

Another proposed organization divided such intelligences into those related to perceptual processing (e.g., auditory, kinesthetic), those related to controlled attention (e.g., fluid reasoning and working memory), and those related to acquired knowledge (e.g., comprehension-verbal, quantitative) with an additional set of speed and fluency abilities (e.g., processing speed and speed of memory retrieval). This elaboration on the original fluid versus crystallized distinction also may prove viable (Schneider et al., 2016).

Implications for educators and for education

Research to-date suggests that education has a healthy effect on students' intellects: that each year of schooling a student undergoes will raise their IQ from 1 to 5 IQ points overall (Brinch and Galloway, 2012; Ritchie and Tucker-Drob, 2018). Like a rising tide, such education may float all boats together, for the most part (Gustafsson and Undheim, 1992; Undheim, 1976). This general improvement may be a consequence of the requirement by most educational systems that younger students' studies are distributed across academic subjects. The general increase in children's mental ability acknowledged, adult intellectual development may be different, as education and intellectually-demanding work often require further specialization in one or two broad intelligences relative to the rest (Ackerman, 2000).

Given the correlations among broad intelligences, the possible differences in ability level across intelligences are such that most students likely will benefit from many forms of educational approaches. The attempt to match a student to a particular form of instruction based on their constellation of broad intelligences may be unneeded, although exceptions exist where ability in a particular area is impaired, for example, by a specific learning disability.

That said, two recent meta-analyses indicated that the applied use of multiple intelligence theory in the classroom may enhance classroom achievement (Aydin, 2019; Bas, 2016). Both meta-analyses were motivated by the implementation of a Turkish Ministry of Education primary school program that promoted sensitivity to "multiple intelligences and the individual differences of students" (Aydin, 2019, p. 2). The appendices to the Ministry's program included two examples of the kinds of lessons that might be carried out. We will describe the Bas work as representative of the two.

Bas (2016) examined 75 studies of the Turkish program's outcomes and found substantial positive effective sizes for student achievement, comparing such multiple-intelligence focused classrooms to control classrooms. Although carefully completed, Bas' work left open whether such improvements might be due to research artifacts included under the broad umbrella of *Hawthorne Effects* (Adair, 1984; Chiesa and Hobbs, 2008), which can include student improvements due to better-thought-out lesson plans, communicating to students that one is using special teaching techniques, or simply disrupting routine in a positive way. As Bas puts it (2016, p. 1836) "even the planning stages of these [multiple intelligence] activities are full of fun, colorful, and creative works" (p. 1836).

In such applied work, tracking how a curriculum is implemented across different settings, and whether assessments of multiple intelligences in students actually were carried out validly, as well as the methods used to match students to their given intellectual strength(s) is challenging. All this is not to deny the creative contributions made by advocates of such approaches (see, for example, Kornhaber et al., 1990; Kornhaber and Gardner, 2006). Further, fun activities that promote student learning are surely worth cheering on. But the research does still leave open the question of the exact causes of educational improvements in the area.

Conclusions and reflections

Consistencies across the past years

The drama and the debate of the early 20th century over whether intelligence is singular or plural has given way to researchers' development of better statistical and measurement techniques to study intelligence. Collectively, the evidence indicates that multiple, broad intelligences do exist, partially independent of general intelligence. The list of those broad intelligences has been remarkably stable over the past 40 or so years, with the exception that emotional intelligence has been added to the group (Mayer et al., 2016), and personal and social intelligences may be further reasonable candidates (Bryan and Mayer, 2021b; Mayer et al., 2019).

With progress over the years gratefully acknowledged, it remains of interest that if one compares Thurstone's Primary Mental Abilities from 1939 to those of Gardner in the 1980s, who recall, employed entirely different criteria to identify intelligences, the continuity is fairly striking. The Cattell-Horn-Carroll (CHC) model, in turn, updates and integrates both conceptions. New approaches seek to organize the broad intelligences themselves into conceptual groups, although this undertaking is just beginning (Bryan and Mayer, 2021b; Mayer, 2018; Schneider and Newman, 2015).

Considerations regarding education

It seems likely that students find some classroom topics more appealing than others owing to their ability and interest in a given subject. Science, math, and art all potentially draw on spatial reasoning, composition on verbal skills, and literature, perhaps, on verbal skills and emotional and personal intelligences. By surrounding students with subjects, some of which they may excel at, we are naturally enhancing thinking in these areas. That makes sense because the broad intelligences—as Gardner himself pointed

out—have bases in our evolutionary past, and for the most part they remain relevant to mental development today, as people grapple with diverse kinds of problem solving in their everyday lives.

Societies are becoming more information dependent rather than less, and the ability to think well and clearly is more important for citizens than ever before. For that reason, education—and the increases in intelligence it brings about—are crucial contributions to people's broad and general intelligences—and to our future.

References

- Ackerman, P.L., 2000. Domain-specific knowledge as the "dark matter" of adult intelligence: *GI/Gc*, personality and interest correlates. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 55B (2), P84.
- Adair, J.G., 1984. The Hawthorne effect: a reconsideration of the methodological artifact. *J. Appl. Psychol.* 69 (2), 334–345. <https://doi.org/10.1037/0021-9010.69.2.334>.
- Anderson, R.D., 1982. New light on French secondary education in the Nineteenth Century. *Soc. Hist.* 7 (2), 147–165.
- Aydin, H., 2019. The effect of multiple intelligence(s) on academic success: a systematic review and meta-analysis. *Eurasia J. Math. Sci. Technol. Educ.* 15 (12). <https://unh.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1265558&site=ehost-live>.
- Bas, G., 2016. The effect of multiple intelligences theory-based education on academic achievement: a meta-analytic review. *Educ. Sci. Theor. Pract.* 16 (6), 1833–1864.
- Beaujean, A.A., Benson, N.F., 2019. The one and the many: enduring legacies of Spearman and Thurstone on intelligence test score interpretation. *Appl. Meas. Educ.* 32 (3), 198–215. <https://doi.org/10.1080/08957347.2019.1619560>.
- Binet, A., Simon, T., 1915. A method of measuring the development of the intelligence of young children. In: *A Method of Measuring the Development of Intelligence of Young Children*. Chicago Medical Book Company, pp. 1–107. <https://doi.org/10.1037/11068-001>.
- Binet, A., Simon, T., 1916. Upon the necessity of establishing a scientific diagnosis of inferior states of intelligence. *L'Année Psych.* 1905, 163–191. <https://doi.org/10.1037/11069-001>.
- Boake, C., 2002. From the Binet-Simon to the Wechsler-Bellevue: tracing the history of intelligence testing. *J. Clin. Exp. Neuropsychol.* 24 (3), 383–405. <https://doi.org/10.1076/jcen.24.3.383.981>.
- Boomsma, A., 2000. Reporting analyses of covariance structures. *Struct. Equ. Model.* 7 (3), 461–483.
- Brinch, C.N., Galloway, T.A., 2012. Schooling in adolescence raises IQ scores. *Proc. Natl. Acad. Sci. U. S. A.* 109 (2), 425–430. <https://doi.org/10.1073/pnas.1106077109>.
- Bryan, V.M., Mayer, J.D., 2020. A meta-analysis of the correlations among broad intelligences: understanding their relations. *Intelligence* 81. <https://doi.org/10.1016/j.intell.2020.101469>.
- Bryan, V.M., Mayer, J.D., 2021a. Ten Broad Intelligences and Their Relations to General Intelligence: An Ancillary Analysis of the Open-Source Data From Bryan and Mayer, 2020. Personality Laboratory at the University of New Hampshire. <https://mypages.unh.edu/jdmayer/graphics>.
- Bryan, V.M., Mayer, J.D., 2021b. Are people-centered intelligences psychometrically distinct from thing-centered intelligences? A meta-analysis. *J. Intell.* 9 (4), 48. <https://doi.org/10.3390/jintelligence9040048>.
- Camilli, G., 2013. Ongoing issues in test fairness. *Educ. Res. Eval.* 19 (2–3), 104–120. <https://doi.org/10.1080/13803611.2013.767602>.
- Carroll, J.B., 1993. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press.
- Cattell, R.B., 1961. Fluid and Crystallized Intelligence. In: Jenkins, J.J., Paterson, D.G. (Eds.), *Studies in individual differences: The search for intelligence*, pp. 738–746. <https://doi.org/10.1037/11491-064>. Appleton-Century-Crofts.
- Cattell, R.B., 1963. Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.* 54 (1), 1–22. <https://doi.org/10.1037/h0046743>.
- Cattell, R.B., Horn, J.L., 1978. A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *J. Educ. Meas.* 15 (3), 139–164. <https://doi.org/10.1111/j.1745-3984.1978.tb00065.x>. APA PsycInfo.
- Chiesa, M., Hobbs, S., 2008. Making sense of social research: how useful is the Hawthorne effect? *Eur. J. Soc. Psychol.* 38 (1), 67–74. <https://doi.org/10.1002/ejsp.401>.
- Conzelmann, K., Weis, S., Süß, H.-M., 2013. New findings about social intelligence: development and application of the Magdeburg Test of Social Intelligence (MTSI). *J. Individ. Differ.* 34 (3), 119–137. <https://doi.org/10.1027/1614-0001/a000106>.
- Deary, I.J., 2001. Human intelligence differences: a recent history. *Trends Cognit. Sci.* 5 (3), 127–130. [https://doi.org/10.1016/S1364-6613\(00\)01621-1](https://doi.org/10.1016/S1364-6613(00)01621-1).
- Detterman, D.K., 1982. Does "g" exist? *Intelligence* 6 (2), 99–108. [https://doi.org/10.1016/0160-2896\(82\)90008-3](https://doi.org/10.1016/0160-2896(82)90008-3).
- Devonis, D.C., 2016. Intelligence testing in historical context. In: Woody, W.D., Miller, R.L., Wozniak, W.J. (Eds.), *Psychological Specialties in Historical Context: Enriching the Classroom Experience for Teachers and Students*, pp. 177–188. <https://unh.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2018-41336-012&site=ehost-live>.
- Fancher, R.E., 1987. *The Intelligence Men: Makers of the I.Q. Controversy*. W. W. Norton.
- Gardner, H., 1983. *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2004-18831-000&site=ehost-live>.
- Gardner, H., 1993. *Frames of Mind: The Theory of Multiple Intelligences*, tenth-anniversary ed. Basic Books. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2004-18831-000&site=ehost-live>.
- Gardner, H., 1999. *Intelligence Reframed: Multiple Intelligences for the 21st Century*. Basic Books. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1999-04335-000&site=ehost-live>.
- Gottfredson, L.S., 1997. Mainstream science on intelligence: an editorial with 52 signatories, history and bibliography. *Intelligence* 24 (1), 13–23. [https://doi.org/10.1016/S0160-2896\(97\)90011-8](https://doi.org/10.1016/S0160-2896(97)90011-8).
- Gustafsson, J.-E., Undheim, J.O., 1992. Stability and change in broad and narrow factors of intelligence from ages 12 to 15 years. *J. Educ. Psychol.* 84 (2), 141–149.
- Haier, R.J., 2016. *The Neuroscience of Intelligence*, illustrated ed. Cambridge University Press.
- Haney, W., 1984. Testing reasoning and reasoning about testing. *Rev. Educ. Res.* 54 (4), 597–654. <https://doi.org/10.2307/1170177>.
- Hedlund, J., Sternberg, R.J., 2000. Too many intelligences? Integrating social, emotional, and practical intelligence. In: Parker, J.D.A. (Ed.), *The Handbook of Emotional Intelligence: Theory, Development, Assessment, and Application at Home, School, and in the Workplace*, vol. 1. Jossey-Bass, pp. 136–167. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-00355-007&site=ehost-live>.
- Horn, J.L., Cattell, R.B., 1966. Refinement and test of the theory of fluid and crystallized general intelligences. *J. Educ. Psychol.* 57 (5), 253–270. <https://doi.org/10.1037/h0023816>. APA PsycInfo.
- Hunt, T., 1928. The measurement of social intelligence. *J. Appl. Psychol.* 12 (3), 317–334. <https://doi.org/10.1037/h0075832>.
- Jennrich, R.I., Robinson, S.M., 1969. A Newton-Raphson algorithm for maximum likelihood factor analysis. *Psychometrika* 34 (1), 111–123. <https://doi.org/10.1007/BF02290176>.
- Joreskog, K.G., 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34 (2), 183–202. <https://doi.org/10.1007/BF02289343>.
- Kornhaber, M.L., Gardner, H., 2006. Multiple intelligences: developments in implementation and theory. In: Constanas, M.A., Sternberg, R.J. (Eds.), *Translating Theory and Research Into Educational Practice: Developments in Content Domains, Large-Scale Reform, and Intellectual Capacity*. Lawrence Erlbaum Associates Publishers, pp. 255–276. <https://unh.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2006-04726-011&site=ehost-live>.
- Kornhaber, M.L., Krechevsky, M., Gardner, H., 1990. Engaging intelligence. *Educ. Psychol.* 25 (3–4), 177–199. https://doi.org/10.1207/s15326985ep2503&4_3.

- Lovie, A.D., Lovie, P., 1993. Charles spearman, cyril burt, and the origins of factor analysis. *J. Hist. Behav. Sci.* 29 (4), 308–321. [https://doi.org/10.1002/1520-6696\(199310\)29:4<308::AID-JHBS2300290402>3.0.CO;2-P](https://doi.org/10.1002/1520-6696(199310)29:4<308::AID-JHBS2300290402>3.0.CO;2-P).
- Lubinski, D., 2004. Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "General intelligence," objectively determined and measured. *J. Pers. Soc. Psychol.* 86 (1), 96–111. <https://doi.org/10.1037/0022-3514.86.1.96>.
- MacCann, C., Joseph, D.L., Newman, D.A., Roberts, R.D., 2014. Emotional intelligence is a second-stratum factor of intelligence: evidence from hierarchical and bifactor models. *Emotion* 14 (2), 358–374.
- Matsueda, R.L., 2012. Key advances in the history of structural equation modeling. In: Hoyle, R.H. (Ed.), *Handbook of Structural Equation Modeling*. The Guilford Press, pp. 17–42. <https://unh.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2012-16551-002&site=ehost-live&scope=site>.
- Mayer, J.D., Bryan, V.M., 2021. From Correlation Matrices to Factor Models. Personality Laboratory at the University of New Hampshire. <https://mypages.unh.edu/jdmayer/graphics>.
- Mayer, J.D., Panter, A.T., Caruso, D.R., 2012. Does personal intelligence exist? Evidence from a new ability-based measure. *J. Pers. Assess.* 94, 124–140. <https://doi.org/10.1080/00223891.2011.646108>.
- Mayer, J.D., Caruso, D.R., Salovey, P., 2016. The ability model of emotional intelligence: principles and updates. *Emot. Rev.* 8, 1–11.
- Mayer, J.D., Caruso, D.R., Panter, A.T., 2019. Advancing the measurement of personal intelligence with the test of personal intelligence, version 5 (TOPI 5). *J. Intell.* 7 (4), 1–17. <https://doi.org/10.3390/jintelligence7010004>.
- Mayer, J.D., 2008. Personal intelligence. *Imagin. Cognit. Pers.* 27 (3), 209–232. <https://doi.org/10.2190/IC.27.3.b>.
- Mayer, J.D., 2018. Intelligences about things and intelligences about people. In: Sternberg, R.J. (Ed.), *The Nature of Human Intelligence*. Cambridge University Press, pp. 270–286. <https://doi.org/10.1017/9781316817049.018>.
- Ritchie, S.J., Tucker-Drob, E.M., 2018. How much does education improve intelligence? A meta-analysis. *Psychol. Sci.* 29 (8), 1358–1369. <https://doi.org/10.1177/0956797618774253>.
- Roberts, R.D., Stankov, L., Pallier, G., Dolph, B., 1997. Charting the cognitive sphere: tactile-kinesthetic performance within the structure of intelligence. *Intelligence* 25 (2), 111–148.
- Salovey, P., Mayer, J.D., 1990. Emotional intelligence. *Imagin. Cognit. Pers.* 9 (3), 185–211.
- Schmid, J., Leiman, J.M., 1957. The development of hierarchical factor solutions. *Psychometrika* 22, 53–61. <https://doi.org/10.1007/BF02289209>.
- Schneider, W.J., Newman, D.A., 2015. Intelligence is multidimensional: theoretical review and implications of specific cognitive abilities. *Hum. Resour. Manag. Rev.* 25 (1), 12–27.
- Schneider, W.J., Mayer, J.D., Newman, D.A., 2016. Integrating hot and cool intelligences: thinking broadly about broad abilities. *J. Intell.* 4 (1), 1–25.
- Shearer, C.B., 2020. A resting state functional connectivity analysis of human intelligence: broad theoretical and practical implications for multiple intelligences theory. *Psychol. Neurosci.* 13 (2), 127–148. <https://doi.org/10.1037/pne0000200>.
- Spearman, C., 1904. "General intelligence," objectively determined and measured. *Am. J. Psychol.* 15 (2), 201–293. <https://doi.org/10.2307/1412107>.
- Sternberg, R.J., 1985. *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge University Press. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1985-97046-000&site=ehost-live>.
- Stevens, M., 2020. *The Knowledge Machine*. Liveright/W.W.Norton.
- Thorndike, R.L., Stein, S., 1937. An evaluation of the attempts to measure social intelligence. *Psychol. Bull.* 34 (5), 275–285. <https://doi.org/10.1037/h0053850>.
- Thorndike, E.L., 1920. Intelligence and its uses. *Harpers Mag.* 140, 227–235.
- Thurstone, L.L., Thurstone, T.G., 1941. Factorial studies of intelligence. *Psychometr. Monogr.* 2, 94.
- Thurstone, L.L., 1938. *Primary Mental Abilities*. University of Chicago Press, Chicago. <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1938-15070-000&site=ehost-live>.
- Thurstone, L.L., 1944. Second-order factors. *Psychometrika* 9, 71–100. <https://doi.org/10.1007/BF02288715>.
- Undheim, J.O., 1976. Ability structure in 10–11 year-old children and the theory of fluid and crystallized intelligence. *J. Educ. Psychol.* 68 (4), 411.
- Wagner, R.K., 2011. In: Sternberg, R.J., Kaufman, S.B. (Eds.), *Practical intelligence*, vol. 1. Cambridge University Press, pp. 550–563. <https://doi.org/10.1017/CBO9780511977244.028>.
- Walker, R.E., Foley, J.M., 1973. Social intelligence: its history and measurement. *Psychol. Rep.* 33 (3), 839–864. <https://doi.org/10.2466/pr0.1973.33.3.839>.
- Weis, S., Süß, H.-M., 2007. Revisiting the search for social intelligence—a multitrait-multimethod study of its structure and construct validity. *Pers. Individ. Differ.* 42 (1), 3–14. <https://doi.org/10.1016/j.paid.2006.04.027>.

Knowledge: a fundamental asset

Kathryn S. McCarthy^a and **Danielle S. McNamara^b**, ^a Georgia State University, College of Education and Human Development, Atlanta, GA, Unites States; and ^b Arizona State University, Tempe, AZ, Unites States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	209
Defining knowledge	209
Knowledge: deficit vs. asset	210
Measuring knowledge	210
Types of knowledge	211
Dimensions of knowledge	212
Building knowledge: acquisition and revision	213
Leveraging existing knowledge	213
Updating and revising knowledge	214
Implications and conclusions	215
Acknowledgments	215
References	216

Introduction

Knowledge is a fundamental concept in education. Indeed, we typically think of learning as the acquisition of knowledge. This knowledge is then used to solve new problems and to support future learning. Knowledge begets knowledge. Hence, understanding what knowledge is, how it can be measured, and how it can be used in future learning is a critical foundation for educators. In this article, we first define knowledge in the context of information processing and mental model building with emphasis on knowledge as an asset. We then describe the multifaceted nature of knowledge and how different types and dimensions of knowledge can be assessed and how these qualities of knowledge relate to learning. Finally, we explore how an asset-based approach to knowledge building can enhance learning and provide recommendations for how educators can help students to leverage what they already know to build their knowledge base.

Defining knowledge

We intuitively think of our knowledge as all of the information that we have learned in the past, our past experiences, our skills (e.g., reading, writing, math, riding a bike). It is the stuff that is in *there*, and the stuff that we try to find when we need to complete a task, answer a question, or simply think about an event. In the field of education, knowledge is the objective: educators strive to increase students' knowledge of various topics (e.g., history, science) and their ability to complete various tasks (e.g., solve math problems, write an essay, read a complex science text). The objective is to enhance students' ability to demonstrate their knowledge and use and apply their knowledge flexibly across multiple situations.

The study of knowledge has a rich history across a variety of fields. Indeed, there are few fields of study that do not contemplate the nature of knowledge.¹ For our purposes, we define knowledge as "all of the information stored in one's memory" (McCarthy and McNamara, 2021; see also Greene et al., 2016a). This information may be correct or incorrect, coherent or incoherent, and the information may be more or less accessible in a given situation. Knowledge is not the process or success of accessing information, but rather it is all of the information that is potentially available from long-term memory.

In educational research and practice, a central objective is to increase learners' knowledge and to make that knowledge more accurate and more coherent. Newly learned knowledge becomes prior knowledge, which can then be leveraged in future situations to create new knowledge. Thus, in addition to measuring what a student has learned in a given task or unit, researchers and instructors are also interested in what a learner already knows before the activity begins, or their *prior knowledge*. Prior knowledge is a strong predictor of comprehension and learning (Dochy et al., 1999). Prior knowledge plays a role in a variety of processes involved in learning. Prior knowledge affects *how* people remember textual information (Bartlett, 1932) and how much information they recall (Bransford and Johnson, 1972; Langer and Nicolich, 1981). Prior knowledge affects how readers make use of comprehension strategies (Cromley and Azevedo, 2007; McNamara, 2004). Prior knowledge also allows learners to more effectively process information and recall it more quickly (Graesser and Clark, 1985; Trabasso and van den Broek, 1985). Those with more knowledge are also better able to identify critical information (Afflerbach, 1990; Chi et al., 1981) and ignore irrelevant information (McNamara and McDaniel, 2004; Voss et al., 1980). Thus, knowledge is not only an outcome, but also a critical predictor of future learning.

¹For a comprehensive review of the philosophical and educational history of the study of knowledge, we recommend Murphy et al. (2012).

Considerations of knowledge are crucial to understanding student outcomes across the spectrum of disciplines and contexts (e.g., Goldman et al., 2016; Kendeou et al., 2011).

Knowledge: deficit vs. asset

When we consider knowledge, it is important to consider whether we conceptualize it as a glass half empty or a glass half full. A deficit (half empty) model of knowledge dominates education and educational research. One considers what a student lacks in knowledge in order to fill those gaps—*because* the objective is to remedy knowledge gaps or deficits by *giving* the students the knowledge that they lack. Alternatively, an asset model of knowledge attempts to discern the knowledge that a student has in order to capitalize and build upon it.

No one comes into the picture completely lacking knowledge. A student who knows very little about science, still knows some things about the world, people, events, and relations. In a deficit framing, the goal of instruction would be to *give* student knowledge about science. By contrast, an asset-based approach aims to capitalize on the student's knowledge about the world in order to build knowledge about science (Scardamalia and Bereiter, 2005). The notion that we can give a student knowledge is a misconception about learning that has been reinforced by assumptions that *giving* lectures or having students read a textbook are (by themselves) sufficient for learning (McNamara, 2010). Unfortunately, that is not the case—students cannot merely consume knowledge. Knowledge is an asset to be built upon, capitalized, used, and enhanced.

Asset-based approaches also emphasize the need to recognize culturally-bound knowledge. Work in culturally-responsive pedagogy highlights that minoritized students do not often see themselves or their lived experiences presented in the classroom and in the curriculum. This can suggest to students that the knowledge that they have is not important or relevant to the classroom context and can artificially suggest to instructors that these students *lack* knowledge and overlooking the *cultural funds of knowledge* that they do possess (Gay, 2018; Ladson-Billings, 2009; Moll et al., 1992). Culturally-responsive pedagogies argue for the recognition and leveraging of the knowledge that students have, rather than focusing on what they do not have. For example, the *Cultural Modeling Framework* (Lee et al., 2004; Lee, 2007) capitalizes on students' informal knowledge and helps them to translate this knowledge to formal coursework (Risko and Walker-Dalhouse, 2007).

Conceptualizing knowledge as an asset rather than a deficit has the potential to radically change education. Considering knowledge as an asset means that educators must consider multiple aspects of knowledge rather than just the amount of knowledge specifically related to a target domain. It means that educators must consider how to build upon knowledge, rather than simply add information. Here, we discuss how to measure knowledge from multiple angles as well as methods to prompt students to use the knowledge that they have in order to construct new knowledge.

Measuring knowledge

Measuring or assessing what a student knows involves not only what concepts are known, but also how those concepts are connected and organized in memory. In addition to understanding *what* information or *how much* information is available in long-term memory, researchers who study knowledge and learning are most interested in how information is organized or structured. How researchers measure knowledge reflects assumptions about how information is stored and represented in memory. Cognitive theories of comprehension and learning (e.g., Anderson and Lebiere, 1998; Johnson-Laird, 1989; Kintsch, 1988) suggest that knowledge can be represented in a connectionist network as a series of nodes (concepts) and links (relations). This network is conceptualized as a multi-layered, multi-dimensional *mental model* or *mental representation*. If a student were to read the sentence "The dog was asleep on the couch," information such as "cat" and "living room" are likely to also be activated in memory through *spreading activation* because these terms are strongly connected to the nodes of "dog" and "couch," respectively. The more and more connected information is, the easier it is for learners to activate and integrate that knowledge in future learning. Mental models that have more densely connected nodes are said to be more *elaborated* and more *coherent*. Thus, in order to measure knowledge, researchers explore the quality of the learner's mental model. Of course, these mental models are not directly observable. Researchers use a variety of tasks and measures to make inferences about the nature of the learner's mental model.

Learning assessments are used to evaluate what a student has learned over time (e.g., a post-reading quiz and end-of-semester final exam). These assessments often use a variety of types of questions to assess different features of the learner's mental model. Surface level questions (e.g., recall, definitions) are designed to evaluate what information is remembered to assess which nodes are present and available. Other tasks, such as inference verification, transfer, or application questions require the learner to integrate information. Thus, they are designed to target the coherence of the learner's mental model. While researchers often carefully develop post-activity knowledge assessment, prior knowledge tests often consist of self-report items, vocabulary recognition, or multiple-choice tests to gauge general familiarity with the topic at hand. While these types of tests are easy to develop and administer, they are often less rigorous and are often not clearly aligned to measuring the various qualities of the learner's existing mental model (Dochy et al., 1999; McCarthy and McNamara, 2021).

In order to evaluate the quality of a mental model, researchers must be sensitive to the fact that knowledge is multi-faceted and multi-dimensional. We borrow from de Jong and Ferguson-Hessler (1996) who argue that knowledge is comprised of different *types* of knowledge and that there are multiple *qualities* or *dimensions* of each of these types of knowledge. These different aspects of knowledge contribute to learning in different ways. Thus, a richer understanding of the nature of knowledge can help educators to consider

which types or dimensions of knowledge are most critically relevant to a given learning task—again, not necessarily what is missing, but perhaps more importantly, what the learner and the learning environment brings to the table.

Types of knowledge

Researchers have identified and distinguished between various *types* of knowledge (e.g., Alexander et al., 1994b; Anderson, 1989; de Jong and Ferguson-Hessler, 1996; Dochy and Alexander, 1995; Kendeou and O'Brien, 2015; Tulving, 1972). Rather than drawing from a specific typology or taxonomy or providing an exhaustive list of the variety of terms and definitions, we describe the most commonly referenced types of knowledge. The boundaries between these types of knowledge can often be relatively fuzzy and the different types of knowledge are also not mutually exclusive.

One common distinction is whether the knowledge is explicit or implicit. **Explicit** knowledge is information that is in conscious awareness, whereas **implicit** or *tacit*, knowledge is assumed to be acquired and used without conscious awareness (Jacoby and Witherspoon, 1982). For example, a person might learn the words to a popular song without intentionally studying the lyrics or may not be able to consciously articulate the steps they take to stay upright on a bike. While explicit knowledge is easy to probe for, implicit knowledge is relatively challenging to assess because the learner may not recognize that they have the knowledge or they may be unable to clearly articulate or express that knowledge.

Perhaps the most familiar type of knowledge is declarative knowledge (Squire, 1992). **Declarative**, also sometimes called conceptual knowledge, refers to predominantly explicit knowledge about facts and concepts (*semantic* memory) as well as knowledge of personal and world events (*episodic* memory; Tulving, 1972). In the context of learning in formal settings, declarative knowledge is often considered the “content knowledge” that one possesses relevant to the current task. Declarative knowledge is “knowing what” or “knowing that.” By contrast, **procedural** knowledge refers to “knowing how” (Ryle, 1949). Procedural knowledge refers to knowing the sequence of steps needed to complete a task (e.g., riding a bike; long division). Often procedural knowledge is either gained implicitly or is taught explicitly and becomes more implicit, or *proceduralized*, over time. Procedural knowledge has largely been studied in the context of problem solving and mathematics (e.g., Anderson and Fincham, 1994; McNamara, 1995; Rittle-Johnson and Siegler, 1998). Relatedly, **embodied** knowledge is intrinsically tied to bodily perceptions and actions. Embodiment of concepts by linking them to perceptual and motoric movements can be used to facilitate meaning making and learning (Glenberg, 2011; Glenberg et al., 2007).

Metacognitive knowledge refers to the “awareness and knowledge about one’s own cognition” (Pintrich, 2002, p. 219; see also Flavell, 1979). Metacognitive knowledge is the information about one’s own knowledge, such as whether you know more or less about one topic compared to another or how much you know as compared to the average person or a peer. Metacognitive knowledge also refers to knowledge of particular learning strategies and the conditions under which one would deploy the strategy or strategies. Knowledge of what one knows, or perhaps more importantly, what one does not know, can help guide learning. Thus, metacognitive knowledge and metacognitive monitoring play critical roles in theories of self-regulated learning (Griffin et al., 2013; Thiede et al., 2003; Winne, 1996, 2018).

Epistemological knowledge refers to the knowledge of “ways of knowing.” Measures of epistemological knowledge (see also epistemic cognition, epistemic thinking, epistemological beliefs, epistemic frames) evaluate the degree to which a person views knowledge as certain or uncertain, simple or complex, whether knowledge or truth is derived internally by the individual or externally through authority (e.g., Greene et al., 2008, 2016b; Hofer and Pintrich, 1997; Schommer, 1990; Wood and Kardash, 2002). A student’s epistemological beliefs can vary across topics and disciplines (Bråten et al., 2009; Buehl et al., 2002). Further epistemic knowledge is predictive of performance across a variety of domains (see Greene et al., 2018).

In some traditions, *beliefs* are considered to be a type of knowledge. A belief is generally defined as an individual’s acceptance that a statement or information is true or represents something that exists. As such, researchers and educators often view knowledge and beliefs as separate constructs; beliefs reflect a greater level of subjectivity and affect, whereas knowledge is often assumed to be more objective in nature (see Murphy and Mason, 2006). Research suggests that learners’ beliefs (including their general and topic-specific epistemologies and attitudes) influence the ways in which knowledge is accessed and applied (e.g., Bråten et al., 2011; List and Alexander, 2017; Schiefele, 1990). Thus, beliefs are undoubtedly an important part of the learning process, but reflect, at least to some extent, a distinct construct to consider *in addition* to knowledge.

In the context of education, another critical type of knowledge is **disciplinary** knowledge, or knowledge related to learning in a particular field, discipline, or domain. Disciplinary knowledge intersects with other types of knowledge in the sense that learners can possess knowledge within the discipline (e.g., declarative facts about a topic, procedural steps to complete a process), but also knowledge about the discipline or domain itself (e.g., epistemic frames, learning goals). For example, literary experts are not only more familiar with common literary themes and conventions, but they also have the knowledge that literary reading encourages nonliteral interpretation and exploration of how the language of the text works to invite that interpretation (McCarthy, 2015; McCarthy and Goldman, 2019). Learning can be particularly challenging when students’ domain-specific epistemological beliefs are mis-matched with the underlying assumptions in domains. For example, students may think of science as inarguable facts. Science experts, on the other hand, recognize that science is subject to multiple interpretations and theories as well as revisions based on new empirical findings (Bråten et al., 2011; Elby et al., 2016; Goldman et al., 2016; Lederman, 2007). Students whose knowledge does not align with this conception of the nature of science may have difficulty synthesizing across conflicting findings or reasoning between competing hypotheses.

Another important, but commonly overlooked distinction is the context in which the knowledge is gained. Knowledge can be acquired in **formal** learning environments (i.e., “school knowledge”). However, learners also acquire sophisticated knowledge from **informal** learning in everyday activities (Saxe, 1988; Scribner and Cole, 1973). Thus, this distinction is not a feature of the information itself, but rather meta-information about the source of the knowledge. Informal knowledge is often more implicit in nature. Researchers are interested in how informal spaces can be used to enhance learning (e.g., museum science) as well as how educators can invite students to make their implicitly learned informal knowledge more explicit in formal educational settings.

Different assessments are more or less likely to tap into variations in each type of knowledge. For example, declarative knowledge is commonly assessed with multiple-choice questions with the assumption that answering correctly means that the learner knows this fact. By contrast, procedural knowledge is assessed as the ability to complete a task (e.g., ride a bicycle, complete a math problem, fly a plane, etc.). Both metacognitive knowledge and epistemic knowledge can be evaluated through self-report surveys (e.g., Mokhtari and Reichard, 2002; Schommer, 1990; Wood and Kardash, 2002), but these aspects of knowledge can also be assessed through interviews, verbal protocol tasks, as well as tasks that require the use of knowledge (McNamara, 2011; Millis et al., 2018).

Dimensions of knowledge

Although the majority of research in knowledge has been on these various types of knowledge described above, more recent work has emphasized the need to more closely examine different *dimensions* of these types of knowledge. The Multidimensional Knowledge in Text Comprehension Framework (MDK-C) Framework (McCarthy and McNamara, 2021) draws from notions of mental model building to argue that prior content knowledge varies along four intersecting dimensions: amount, accuracy, specificity, and coherence. It is of note that the framework was developed with respect to declarative content knowledge, but it is likely that these dimensions are also evident in other types of knowledge.

The **amount** of knowledge refers to the number of concepts in memory. More simply put, this dimension answers the question “how *much* knowledge does the learner have?” Most measures of knowledge in research and education focus on estimating the amount of knowledge across various types of knowledge. Assuming that the items on a knowledge assessment are varied and well-rounded, the score on the test can serve as a relative measure of prior knowledge. That is, students who answer more questions on the knowledge assessment are assumed to know more than those who score less well. However, the MDK-C framework emphasizes that a focus on only how much a learner knows overlooks important other dimensions of that knowledge that can contribute to learning.

A second dimension is the **accuracy** of knowledge. In some philosophical views, knowledge refers only to information that is true in the world. However, more cognitive and education-focused approaches argue that what one “knows” may not be accurate (e.g., vaccines cause autism). Inaccurate knowledge may reflect inaccurate information that was intentionally passed along to the learner (i.e., misinformation, disinformation), but it also includes information that reflects a naive conception that was more implicitly drawn from misunderstanding everyday experiences (Vosniadou and Skopeliti, 2017). In order to measure the accuracy of knowledge, researchers often include distractor items that reflect common misconceptions or naive understandings. A common example of this is in multiple choice math problems. In addition to the correct answer, one of the choices is often a solution that the learner would arrive at if the learner did a procedure out of order. For example, in the problem “ $6 \times 3 = ?$ ”, one of the distractors might be “9.” If a student selects this answer, it suggests that the student might not know arithmetic signs or may not understand the difference between multiplication and addition. Inaccurate knowledge can impair comprehension and learning and can further perpetuate misunderstandings (Kendeou and van den Broek, 2005, 2007; Rapp and Braasch, 2014; van den Broek, 2010). Indeed, inaccurate information can be well embedded in individuals’ knowledge structures, and challenging to overcome or correct (Guzzetti et al., 1992; Vosniadou, 2009). Just like the amount of knowledge, the accuracy of ideas in knowledge is assumed to be on a continuum. Knowledge of ideas may be more or less accurate and the connections between ideas may be more or less accurate. When a learner has many ideas and connections between ideas that contradict known truths, the misconceptions are more challenging to overcome.

The third dimension of knowledge is its **specificity** in relation to the current task. Knowledge about the targeted topic is more specific, whereas general questions about the domain or world knowledge is less specific (e.g., Alexander et al., 1994a,b; McCarthy et al., 2018a,b,c). The specificity of knowledge is not a static feature of the information. That is, a given node in the mental representation is not inherently general or specific, but rather that particular information might be more or less related to the topic at hand. For example, if a student is learning about the digestive system, questions regarding the digestive organs would tap into *topic* specific knowledge. The student may also have subject knowledge about anatomy more generally that may not be immediately relevant to the text, but can support the activation of relevant related information. It may also be important to consider the degree to which the students have more general science (domain) knowledge and also very broad *general* world knowledge as this more distal information can support additional knowledge activation that can facilitate processing as well as the integration and organization of new information.

The fourth dimension of knowledge is **coherence**, which most closely aligns with the notions of structure or organization of knowledge. We borrow the term coherence from theories of text and discourse to highlight the importance of *interconnectedness* of the information in the construction of an elaborated mental model. More connections facilitate knowledge activation, which in turn increases comprehension (e.g., Gernsbacher, 1991; Kintsch, 1988). The importance of coherence is also reflected in theories of expertise which posit that expertise is not defined by mere quantity of knowledge, but rather by how the information is organized

(Alexander, 2004; Bedard and Chi, 1992; Bereiter and Scardamalia, 1986; Langer and Nicolich, 1981; Langer, 1984). Although coherence plays a critical role in learning, many assessments do not include high quality items that can probe for the coherence of the learner's mental model. This is, in part, because assessing the connections between ideas is more complex than measuring whether or not a particular node of information exists in memory. In comprehension tests, researchers often use inference questions to assess the coherence of the learner's mental model. Answers to inference items cannot be found directly in the text. Instead, the reader must connect information from different parts of the text or take the information from the text and apply it to a new scenario. The question "Which cell organelle makes proteins?" asks the learner to search their memory for a particular concept. By contrast, an inference question might ask "What would happen to a cell if there were no ribosomes?" This question is more likely to require the learner to draw upon what they know about the ribosome's function in order to infer what might happen if they were removed. Other methods of evaluating coherence include card sorting tasks (which allow researchers to understand how a learner thinks that different concepts "go together") and evaluating students' open-ended responses (e.g., summaries, explanations, essays) and scoring them for organization or for the completeness or coherence of the written product (see McCarthy et al., 2018c).

The MDK-C framework highlights that these different dimensions of knowledge exist in tandem and depend on one another, with complex interactive effects. For example, broad general knowledge may compensate for a lack of topic-specific knowledge (e.g., McCarthy et al., 2018a) or a coherent, but ultimately inaccurate understanding of a topic (i.e., a misconception) might be more difficult to overcome than a single inaccurate idea among a disparate network of loosely connected ideas (Chi, 2008). Considering multiple dimensions or aspects of knowledge will yield a more complete picture of what a student knows and does not know, and the potential effects on comprehension, learning, and problem solving. Different types of assessment measures are needed to capture the multidimensional nature of knowledge. In turn, more work is needed to better understand how these dimensions manifest across the various types of knowledge (e.g., explicit vs. implicit, declarative vs. procedural, metacognitive, epistemological, disciplinary, formal vs. informal). Considering these four dimensions potentially enhances the ability to more comprehensively evaluate the quality of learners' mental models and make more nuanced predictions regarding how knowledge can or will be used in future tasks.

Building knowledge: acquisition and revision

Assessing what a learner knows before or after a learning activity is only one piece of the puzzle. These assessments serve as formative feedback that can be used to tailor future learning. However, this tailoring can be challenging for educators when they face a classroom of students who each possess varying aspects of prior knowledge or who may have inaccurate knowledge that could impede learning of future content.

Instructors often use lectures or background readings in an attempt to attenuate the effects of students' varying levels of prior knowledge. Essentially, many assume that "giving" students information renders their knowledge of the topic equivalent. However, such an approach is not particularly effective. For example, if a student is provided with background information that is not familiar, then the student must comprehend and learn the new material - it is, in itself, a learning activity: a student with little knowledge of the topic must first learn the information in the text. By contrast, this background reading will serve to activate a knowledgeable readers' existing knowledge base. Thus, while the background reading *may* help a less-knowledgeable reader in the task, the background reading will also facilitate processing and learning from the more knowledgeable reader. Although enhancing learning for all learners is not inherently bad, this approach does not actually "level the playing field" as is usually the intent. Moreover, such an approach can disadvantage students with less knowledge if they do not understand or learn the background material (McNamara and Kintsch, 1996). In lieu of this approach, which leans upon assumptions of *knowledge-as-consumption* and *knowledge-as-deficit*, research in knowledge building and knowledge revision suggests that it is more effective to develop activities and interventions that help learners to leverage and manipulate the knowledge they already possess.

Leveraging existing knowledge

Rather than attempting to quickly introduce new information prior to a lesson, it is more effective to meet the students where they are. This can happen in two ways. The first is by providing individualized scaffolding. For example, students who demonstrate a deep misconception may need to engage in a different task than those who have little understanding of the topic at all. Different activities and interventions may be better suited to particular student profiles. A second complementary way to support learning is helping students to leverage the knowledge that they already possess. This is particularly important because information that exists in memory is not always immediately available, active, or used in a given learning task. Prompting students with more active learning strategies can encourage a deeper search of long-term memory. For example, students with more prior knowledge are more resistant to accepting inaccurate information (Rapp and Braasch, 2014; Richter et al., 2009). Explicitly prompting students to draw upon what they already know by correcting information as they read questionable new content does not fully inoculate them against the inaccurate information, but it can reduce the negative effects (Rapp et al., 2014).

Similarly, tutoring programs such as AutoTutor (Graesser et al., 2004) use questions and hints to help students to bring to bear their existing knowledge. AutoTutor incorporates automated, adaptive hints and prompts that encourage activation of knowledge as well as metacognitive processing, reasoning, and problem-solving strategies. The pedagogical agent moves the student toward the answer, allowing them to generate the solution themselves, rather than giving it to the student.

In addition to specific active learning activities, educators can also help students to learn and effectively use comprehension strategies that facilitate the activation and integration of existing knowledge during learning. For example, interventions like Self-Explanation Reading Training (SERT, McNamara, 2004, 2017) and reciprocal teaching (Palincsar and Brown, 1986; Rosenshine and Meister, 1994) provide instruction and practice to support comprehension strategies such as paraphrasing, predicting, and inference. These strategies encourage learners to try to draw upon their existing prior knowledge to connect information from across different parts of a text and to think about how their current reading relates to things that they already know.

Likewise, iSTART (Interactive Strategy Training for Active Reading and Thinking; McNamara et al., 2004) is an adaptive tutoring system that provides instruction on how to use self-explanation to build upon the knowledge that students have about the topic by using various comprehension strategies. Students learn to paraphrase and self-explain challenging content and then make links, either to prior knowledge or to prior text. iSTART has been shown to be effective across a wide spectrum of learner ages and abilities (Jackson and McNamara, 2013; McCarthy et al., 2018b; McCarthy et al., 2020a,b). Low-knowledge students in particular gain from this type of instruction because they learn that they can use the knowledge that they have to make sense of challenging text (Jackson et al., 2010).

Instructors can also enhance learning by encouraging students to draw upon different types of knowledge. Intervention work shows that having students gesture during learning encourages them to draw upon embodied knowledge that can help them to leverage additional knowledge that they possess about the world (Glenberg et al., 2007; Hostetter and Alibali, 2019; Johnson-Glenberg et al., 2014). Encouraging students to gesture during reading (Glenberg et al., 2011), while solving math equations (Broaders et al., 2007; Cook et al., 2008; Goldin-Meadow et al., 2009), or during science activities (Johnson-Glenberg and Megowan-Romanowicz, 2017) can support deeper understanding and longer retention.

In formal education settings, many students can struggle because of a perceived distinction between “school knowledge” and everyday experience. For example, English Language Arts (ELA) students are often described as lacking the relevant knowledge to engage in discipline-appropriate interpretive reasoning. Levine (2019) argues that the issue is not that students lack this knowledge, but rather educators tend to frame interpretive reasoning as a school-bound, or formal learning activity devoid of connection to everyday experience. This framing prevents students from drawing upon the wealth of knowledge that they have. To demonstrate the potential of “everyday interpretation,” Levine asked students in an ELA course to read a literary short story. In one condition, students were given a common ELA classroom prompt asking about interpretation of themes. Students who were given this prompt tended to struggle with generating interpretations and story-appropriate themes. In another condition, students were asked to complete a sentence stem that encouraged the students to draw on their own everyday experience (“Reading this story suggests that the world can be a place where ...”). These students not only generated more interpretations, but these interpretations were more aligned with the textual information.

Educational research suggests that students are likely to have relevant knowledge that they may not recognize as relevant to the classroom context or to a particular learning task. Educators can help their students bring to bear their own knowledge through activities and prompts that encourage active generation of content that can draw upon multiple types of knowledge. This may involve extended instruction and practice so that students can readily deploy effective learning strategies, but they may also be short prompts or bite-sized changes to the wording of task instructions that can more explicitly guide students to see and use what they already know.

Updating and revising knowledge

It is impossible to learn new information without encountering false information or incorrectly interpreting information. Thus, learning also depends on processes of knowledge updating and revision. Once inaccurate information has been encoded and consolidated in long-term memory, it cannot simply be deleted or replaced (Dole and Sinatra, 1998; Kendeou and O’Brien, 2014). Indeed, researchers in areas of conceptual change argue that positioning learners’ misunderstandings or naive conceptions as misconceptions² that must be replaced, denies both the complexity of knowledge and its constructive nature (Smith et al., 1994). Active generation during learning necessarily includes generating both right and wrong answers, and having incorrect or naive assumptions and understandings. Learning involves gradually developing a coherent understanding of concepts, with the general objective for knowledge to be *accurate*, or aligned with claims asserted by domain experts. Unfortunately, updating and revising knowledge can be challenging.

In the era of “fake news” and misinformation, researchers have been working to understand how we can protect learners from encoding inaccurate information (Rapp et al., 2014; Rapp, 2016). While these steps for preventing inaccurate information from entering long-term memory are important, it is also critical to recognize that not all inaccurate information is derived through direct instruction. Students are not empty vessels; at the least, they bring to the table everyday experiences and observations and informal theories about the world (Carey, 1985). These naive conceptions and experiential knowledge often run counter to scientific explanations (Chinn and Brewer, 1993; Vosniadou and Brewer, 1992, 1994; Vosniadou and Skopeliti, 2017). For example, students experience gravity by seeing objects fall to the ground. These observations lead students to the conclusion that gravity’s force is exerted in an up and down motion along a flat surface. This makes it difficult for students to reckon with the scientific understanding that

²The term “misconception” remains hotly debated (see Vosniadou and Skopeliti, 2017). Researchers also use terms such as “pre-conception,” “naive conception,” “alternative conception” and “misconceived knowledge” to more accurately capture the developmental and often informal nature of this knowledge.

earth's gravity is a centripetal force upon a spherical object. When experience runs counter or slightly askew to scientific explanations, students tend to develop fragmented mental models that allow for maintenance of counterfactuals or they construct coherent, but ultimately inaccurate understandings of the concepts (Disessa et al., 2004). Inaccurate information and mental representations of concepts can render learning more difficult. Discrepant information makes it more difficult to construct a coherent mental model (Alvermann et al., 1985; Kendeou and van den Broek, 2007). Learners can also use their inaccurate information in mental model building, further misdirecting their understanding (Guzzetti et al., 1993; Kendeou and van den Broek, 2005). As this inaccurate understanding becomes more robust, the activation of inaccurate information becomes easier, making it even more challenging to revise the inaccurate nodes and links in the mental model.

A common method of supporting conceptual change is through refutation text (Tippett, 2010; Zengilowski et al., 2021). In refutation texts, the common misconception is mentioned and then immediately refuted with the correct information. Intuitively, it may seem counterproductive to restate incorrect information. However, it is this *co-activation* of information that drives revision. Having both the correct and incorrect information active in memory at once increases the likelihood that the student will recognize and reconcile the discrepancy (Kendeou et al., 2014; van den Broek and Kendeou, 2008). Of course, knowledge revision is not likely to occur in a single instance, but rather refutations offer an opportunity to begin to cultivate a more accurate mental model through additional instruction and support.

Knowledge revision faces further challenges when new information clashes with inaccurate knowledge that is tied to strong attitudes, beliefs, and emotions (Trevors et al., 2017b; Murphy and Mason, 2006; Thacker et al., 2020). For example, students often struggle to understand natural selection because the process is more complicated than common agent-based cause-and-effect relations (Chi et al., 2012; McLure et al., 2020). In addition to the issue of the process itself, ideas of natural selection and evolution are also at odds with many religious beliefs. Attitudes and beliefs tend to bias processing in that students willingly accept attitude-consistent information while rejecting and scrutinizing attitude-inconsistent content (Kunda, 1990; Lord et al., 1979; Maier and Richter, 2013). Beyond topic-specific beliefs, learners' epistemic knowledge and epistemic beliefs about the nature of knowing can also influence the extent to which learners engage with content that is inconsistent with their current understanding (Sinatra et al., 2014; Trevors et al., 2017a,b; Wiley et al., 2020). Interacting with conflicting information can also give rise to negative affective responses which can impede knowledge revision (Trevors et al., 2017b).

These findings suggest that knowledge revision and conceptual change are more complex than simply correcting or replacing individual pieces of inaccurate information. Supporting students' ability to engage in conceptual change and knowledge building requires not only co-activation and updating of declarative knowledge, but also attending to the learners' attitudes and affects and supporting long-term changes in epistemic knowledge and beliefs so that students can develop discipline-appropriate knowledge and epistemologies (e.g., Goldman et al., 2016; Lombardi and Sinatra, 2017). These approaches enable students to build upon their existing knowledge in ways that help them to construct more accurate and coherent mental models.

Implications and conclusions

Helping students to develop their knowledge is the cornerstone of education. However, acquiring knowledge is only beneficial insofar that learners can use that knowledge in new situations. Thus, knowledge is not merely an outcome, but also a key predictor in future learning. Existing research highlights that effective learning comes from capitalizing on what a student already knows and engaging the student in active, knowledge *building*. Learning activities that help students to draw upon the many types of knowledge that they already possess support more meaningful and long-term learning. In addition to understanding the many types of knowledge a learner has, it is also critical to consider multiple dimensions of that knowledge. Such a conceptualization of knowledge further pushes researchers and educators to move beyond diagnoses of the "haves" and "have nots" toward a richer understanding of the nature of what a student knows.

High quality evaluation of knowledge should include summative and formative assessments that tap the multiple types and dimensions of knowledge. It would be impossible to evaluate the whole of a learner's knowledge in a single assessment. Instead, the aim would be to more carefully consider which aspects of knowledge are most clearly related with the current learning task or learning objective and to find or develop measures that align with that knowledge. Through more nuanced understanding of student knowledge, educators can develop instruction and practice that can help their students to more successfully build upon what they already know.

Acknowledgments

Research related to the content of this article is supported by grants from the Department of Education (R305A150176, R305A180144, R305A190063), the Office of Naval Research (N00014-14-1-0343, N00014-17-1-2300, N00014-19-1-2424), the American Psychological Association Division 15, and the Spencer Foundation (201900217). The views expressed are those of the authors and do not necessarily reflect the views of the IES, ONR, APA, or the Spencer Foundation.

References

- Afflerbach, P.P., 1990. The influence of prior knowledge on expert readers' main idea construction strategies. *Read. Res. Q.* 25 (1), 31–46. <https://doi.org/10.2307/747986>.
- Alexander, P.A., Kulikowich, J.M., Jetton, T.L., 1994a. The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Rev. Educ. Res.* 64 (2), 201–252. <https://doi.org/10.3102/00346543064002201>.
- Alexander, P.A., Kulikowich, J.M., Schulze, S.K., 1994b. The influence of topic knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learn. Indiv. Differ.* 6 (4), 379–397. [https://doi.org/10.1016/1041-6080\(94\)90001-9](https://doi.org/10.1016/1041-6080(94)90001-9).
- Alexander, P.A., 2004. A model of domain learning: reinterpreting expertise as a multidimensional, multistage process. In: *Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development*. Lawrence Erlbaum Associates Publishers (The Educational Psychology Series), Mahwah, NJ, US, pp. 273–298.
- Alvermann, D.E., Smith, L.C., Readence, J.E., 1985. Prior knowledge activation and the comprehension of compatible and incompatible text. *Read. Res. Q.* 20 (4), 420–436. <https://doi.org/10.2307/747852>.
- Anderson, J.R., Fincham, J.M., 1994. Acquisition of procedural skills from examples. *J. Exp. Psychol. Learn. Mem. Cognit.* 20 (6), 1322–1340. <https://doi.org/10.1037/0278-7393.20.6.1322>.
- Anderson, J.R., Lebiere, C., 1998. *The Atomic Components of Thought*.
- Anderson, J.R., 1989. A theory of the origins of human knowledge. *Artif. Intell.* 40 (1), 313–351. [https://doi.org/10.1016/0004-3702\(89\)90052-0](https://doi.org/10.1016/0004-3702(89)90052-0).
- Bartlett, F.A., 1932. *A Study in Experimental and Social Psychology*. Cambridge University Press, New York.
- Bedard, J., Chi, M.T., 1992. Expertise. *Curr. Dir. Psychol. Sci.* 1 (4), 135–139.
- Bereiter, C., Scardamalia, M., 1986. Educational relevance of the study of expertise. *Interchange* 17 (2), 10–19. <https://doi.org/10.1007/BF01807464>.
- Bransford, J.D., Johnson, M.K., 1972. Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verb. Learn. Verb. Behav.* 11 (6), 717–726. [https://doi.org/10.1016/S0022-5371\(72\)80006-9](https://doi.org/10.1016/S0022-5371(72)80006-9).
- Bråten, I., et al., 2009. Personal epistemology across cultures: exploring Norwegian and Spanish university students' epistemic beliefs about climate change. *Soc. Psychol. Educ.* 12 (4), 529–560.
- Bråten, I., et al., 2011. The role of epistemic beliefs in the comprehension of multiple expository texts: toward an integrated model. *Educ. Psychol.* 46 (1), 48–70.
- Broaders, S.C., et al., 2007. Making children gesture brings out implicit knowledge and leads to learning. *J. Exp. Psychol. Gen.* 136 (4), 539.
- Buehl, M.M., Alexander, P.A., Murphy, P.K., 2002. Beliefs about schooled knowledge: domain specific or domain general? *Contemp. Educ. Psychol.* 27 (3), 415–449. <https://doi.org/10.1006/ceps.2001.1103>.
- Carey, S., 1985. *Conceptual Change in Childhood*. MIT press.
- Chi, M.T.H., Feltovich, P.J., Glaser, R., 1981. Categorization and representation of physics problems by experts and novices. *Cognit. Sci.* 5 (2), 121–152.
- Chi, M.T., et al., 2012. Misconceived causal explanations for emergent processes. *Cognit. Sci.* 36 (1), 1–61.
- Chi, M.T., 2008. Three types of conceptual change: belief revision, mental model transformation, and categorical shift. *Int. Handb. Res. Concept. Change* 61, 82.
- Chinn, C.A., Brewer, W.F., 1993. The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Rev. Educ. Res.* 63 (1), 1–49.
- Cook, S.W., Mitchell, Z., Goldin-Meadow, S., 2008. Gesturing makes learning last. *Cognition* 106 (2), 1047–1058.
- Cromley, J.G., Azevedo, R., 2007. Testing and refining the direct and inferential mediation model of reading comprehension. *J. Educ. Psychol.* 99 (2), 311–325. <https://doi.org/10.1037/0022-0663.99.2.311>.
- De Jong, T., Ferguson-Hessler, M.G.M., 1996. Types and qualities of knowledge. *Educ. Psychol.* 31 (2), 105–113. https://doi.org/10.1207/s15326985ep3102_2.
- Disessa, A.A., Gillespie, N.M., Esterly, J.B., 2004. Coherence vs. fragmentation in the development of the concept of force. *Cognit. Sci.* 28 (6), 843–900.
- Dochy, F.J.R.C., Alexander, P.A., 1995. Mapping prior knowledge: a framework for discussion among researchers. *Eur. J. Psychol. Educ.* 10 (3), 225–242. <https://doi.org/10.1007/BF03172918>.
- Dochy, F., Segers, M., Buehl, M.M., 1999. The relation between assessment practices and outcomes of studies: the case of research on prior knowledge. *Rev. Educ. Res.* 69 (2), 145–186. <https://doi.org/10.3102/00346543069002145>.
- Dole, J.A., Sinatra, G.M., 1998. Reconceptualizing change in the cognitive construction of knowledge. *Educ. Psychol.* 33 (2–3), 109–128.
- Elby, A., Macrander, C., Hammer, D., 2016. Epistemic cognition in science. In: *Handbook of Epistemic Cognition*, pp. 113–127.
- Flavell, J.H., 1979. Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. *Am. Psychol.* 34 (10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>.
- Gay, G., 2018. *Culturally Responsive Teaching: Theory, Research, and Practice*. Teachers College Press.
- Gernsbacher, M.A., 1991. Cognitive processes and mechanisms in language comprehension: the structure building framework. In: Bower, G.H. (Ed.), *Psychology of Learning and Motivation*. Academic Press, pp. 217–263. [https://doi.org/10.1016/S0079-7421\(08\)60125-5](https://doi.org/10.1016/S0079-7421(08)60125-5).
- Glenberg, A.M., et al., 2007. What brains are for: action, meaning, and reading comprehension. In: *Reading Comprehension Strategies: Theories, Interventions, and Technologies*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 221–240.
- Glenberg, A.M., Goldberg, A.B., Zhu, X., 2011. Improving early reading comprehension using embodied CAI. *Instr. Sci.* 39 (1), 27–39.
- Glenberg, A.M., 2011. How reading comprehension is embodied and why that matters. *Int. Electron. J. Environ. Educ.* 4 (1), 5–18.
- Goldin-Meadow, S., Cook, S.W., Mitchell, Z.A., 2009. Gesturing gives children new ideas about math. *Psychol. Sci.* 20 (3), 267–272.
- Goldman, S.R., et al., 2016. *Disciplinary Literacies and Learning to Read for Understanding: A Conceptual Framework of Core Processes and Constructs*, Grantee Submission. Available at: <https://eric.ed.gov/?id=ED565384>. (Accessed 5 September 2020).
- Graesser, A.C., Clark, L.F., 1985. The generation of knowledge-based inferences during narrative comprehension. In: *Advances in Psychology*. Elsevier, pp. 53–94.
- Graesser, A.C., et al., 2004. Autotutor: a tutor with dialogue in natural language. *Behav. Res. Methods Instrum. Comput.* 36 (2), 180–192. <https://doi.org/10.3758/BF03195563>.
- Greene, J.A., Azevedo, R., Torney-Purta, J., 2008. Modeling epistemic and ontological cognition: philosophical perspectives and methodological directions. *Educ. Psychol.* 43 (3), 142–160. <https://doi.org/10.1080/00461520802178458>.
- Greene, J.A., Sandoval, W.A., Bråten, I., 2016a. An introduction to epistemic cognition. In: *Handbook of Epistemic Cognition*. Routledge, pp. 13–28.
- Greene, J.A., Sandoval, W.A., Bråten, I., 2016b. *Handbook of Epistemic Cognition*. Routledge.
- Greene, J.A., Cartliff, B.M., Duke, R.F., 2018. A meta-analytic review of the relationship between epistemic cognition and academic achievement. *J. Educ. Psychol.* 110 (8), 1084.
- Griffin, T.D., Wiley, J., Salas, C.R., 2013. Supporting effective self-regulated learning: the critical role of monitoring. In: Azevedo, R., Aleven, V. (Eds.), *International Handbook of Metacognition and Learning Technologies*. Springer (Springer International Handbooks of Education), New York, NY, pp. 19–34. https://doi.org/10.1007/978-1-4419-5546-3_2.
- Guzzetti, B.J., Snyder, T.E., Glass, G.V., 1992. Promoting conceptual change in science: can texts be used effectively? *J. Read.* 35 (8), 642–649.
- Guzzetti, B.J., et al., 1993. Promoting conceptual change in science: a comparative meta-analysis of instructional interventions from reading education and science education. *Read. Res. Q.* 117–159.
- Hofer, B.K., Pintrich, P.R., 1997. The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67 (1), 88–140.
- Hostetter, A.B., Alibali, M.W., 2019. Gesture as simulated action: revisiting the framework. *Psychon. Bull. Rev.* 26 (3), 721–752.
- Jackson, G.T., McNamara, D.S., 2013. Motivation and performance in a game-based intelligent tutoring system. *J. Educ. Psychol.* 105 (4), 1036.
- Jackson, G.T., Boonthum, C., McNamara, D.S., 2010. The efficacy of iSTART extended practice: low ability students catch up. In: *International Conference on Intelligent Tutoring Systems*. Springer, pp. 349–351.

- Jacoby, L.L., Witherspoon, D., 1982. Remembering without awareness. *Can. J. Psychol.* 36 (2), 300.
- Johnson-Glenberg, M.C., Megowan-Romanowicz, C., 2017. Embodied science and mixed reality: how gesture and motion capture affect physics education. *Cogn. Res. Princ. Implic.* 2 (1), 24.
- Johnson-Glenberg, M.C., et al., 2014. Collaborative embodied learning in mixed reality motion-capture environments: two science studies. *J. Educ. Psychol.* 106 (1), 86–104. <https://doi.org/10.1037/a0034008>.
- Johnson-Laird, P.N., 1989. Mental models. In: *Foundations of Cognitive Science*. The MIT Press, Cambridge, MA, US, pp. 469–499.
- Kendeou, P., O'Brien, E.J., 2014. The Knowledge Revision Components (KReC) Framework: Processes and Mechanisms.
- Kendeou, P., O'Brien, E.J., 2015. Prior knowledge: acquisition and revision. In: *Handbook of Individual Differences in Reading: Reader, Text, and Context: Reader, Text, and Context*, pp. 151–163. <https://doi.org/10.4324/9780203075562>.
- Kendeou, P., van den Broek, P., 2005. The effects of readers' misconceptions on comprehension of scientific text. *J. Educ. Psychol.* 97 (2), 235–245. <https://doi.org/10.1037/0022-0663.97.2.235>.
- Kendeou, P., van den Broek, P., 2007. The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Mem. Cognit.* 35 (7), 1567–1577. <https://doi.org/10.3758/BF03193491>.
- Kendeou, P., Muis, K.R., Fulton, S., 2011. Reader and text factors in reading comprehension processes. *J. Res. Read.* 34 (4), 365–383.
- Kendeou, P., et al., 2014. Knowledge revision processes in refutation texts. *Discourse Process* 51 (5–6), 374–397.
- Kintsch, W., 1988. The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95 (2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>.
- Kunda, Z., 1990. The case for motivated reasoning. *Psychol. Bull.* 108 (3), 480.
- Ladson-Billings, G., 2009. *The Dreamkeepers: Successful Teachers of African American Children*. John Wiley & Sons.
- Langer, J.A., Nicolic, M., 1981. Prior knowledge and its relationship to comprehension. *J. Read. Behav.* 13 (4), 373–379. <https://doi.org/10.1080/10862968109547426>.
- Langer, J.A., 1984. Examining background knowledge and text comprehension. *Read. Res. Q.* 19 (4), 468–481. <https://doi.org/10.2307/747918>.
- Lederman, N.G., 2007. Nature of science: past, present, and future. *Handb. Res. Sci. Educ.* 2, 831–879.
- Lee, C.D., et al., 2004. Cultural modeling as a frame for narrative analysis. In: *Narrative Analysis: Studying the Development of Individuals in Society*, pp. 39–62.
- Lee, C.D., 2007. *Culture, Literacy, & Learning: Taking Bloom in the Midst of the Whirlwind*. Teachers College Pr.
- Levine, S., 2019. Using everyday language to support students in constructing thematic interpretations. *J. Learn. Sci.* 28 (1), 1–31.
- List, A., Alexander, P.A., 2017. Cognitive affective engagement model of multiple source use. *Educ. Psychol.* 52 (3), 182–199. <https://doi.org/10.1080/00461520.2017.1329014>.
- Lombardi, D., Sinatra, G.M., 2017. Don't believe everything you think. In: *Converging Perspectives on Conceptual Change: Mapping an Emerging Paradigm in the Learning Sciences*, vol. 129.
- Lord, C.G., Ross, L., Lepper, M.R., 1979. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.* 37 (11), 2098.
- Maier, J., Richter, T., 2013. Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cogn. Instr.* 31 (2), 151–175.
- McCarthy, K.S., Goldman, S.R., 2019. Constructing interpretive inferences about literary text: the role of domain-specific knowledge. *Learn. Instr.* 60, 245–251. <https://doi.org/10.1016/j.learninstruc.2017.12.004>.
- McCarthy, K.S., McNamara, D.S., 2021. The multidimensional knowledge in text comprehension framework. *Educ. Psychol.* 56 (3), 196–214. <https://doi.org/10.1080/00461520.2021.1872379>.
- McCarthy, K.S., Guerrero, T.A., et al., 2018a. Comprehension in a scenario-based assessment: domain and topic-specific background knowledge. *Discourse Process* 55 (5–6), 510–524. <https://doi.org/10.1080/0163853X.2018.1460159>.
- McCarthy, K.S., Likens, A.D., et al., 2018b. Metacognitive overload! Positive and negative effects of metacognitive prompts in an intelligent tutoring system. *Int. J. Artif. Intell. Educ.* 28 (3), 420–438. <https://doi.org/10.1007/s40593-018-0164-5>.
- McCarthy, K.S., Kopp, K.J., et al., 2018c. Methods of studying text: memory, comprehension, and learning. In: Otani, H., Schwartz, B.L. (Eds.), *Handbook of Research Methods in Human Memory*, pp. 104–124.
- McCarthy, K.S., et al., 2020a. Personalized learning in iSTART: past modifications and future design. *J. Res. Technol. Educ.* 52 (3), 301–321. <https://doi.org/10.1080/15391523.2020.1716201>.
- McCarthy, K.S., Soto, C.M., de Blume, A.P.G., Palma, D., González, J.I., McNamara, D.S., 2020b. Improving reading comprehension in Spanish using iSTART-E: a pilot study. *Int. J. Comput. Assist. Lang. Learn. Teach.* 10 (4), 66–82.
- McCarthy, K.S., 2015. Reading beyond the lines: a critical review of cognitive approaches to literary interpretation and comprehension. *Sci. Study Lit.* 5 (1), 99–128. <https://doi.org/10.1075/ssol.5.1.05mcc>.
- McLure, F., Won, M., Treagust, D.F., 2020. Students' understanding of the emergent processes of natural selection: the need for ontological conceptual change. *Int. J. Sci. Educ.* 1–18.
- McNamara, D.S., Kintsch, W., 1996. Learning from texts: effects of prior knowledge and text coherence. *Discourse Process* 22 (3), 247–288. <https://doi.org/10.1080/01638539609544975>.
- McNamara, D.S., McDaniel, M.A., 2004. Suppressing irrelevant information: knowledge activation or inhibition? *J. Exp. Psychol. Learn. Mem. Cognit.* 30 (2), 465–482. <https://doi.org/10.1037/0278-7393.30.2.465>.
- McNamara, D.S., Levinstein, I.B., Boonthum, C., 2004. iSTART: interactive strategy training for active reading and thinking. *Behav. Res. Methods Instrum. Comput.* 36 (2), 222–233.
- McNamara, D.S., 1995. Effects of prior knowledge on the generation advantage: calculators vs. calculation to learn simple multiplication. *J. Educ. Psychol.* 87 (2), 307–318. <https://doi.org/10.1037/0022-0663.87.2.307>.
- McNamara, D.S., 2004. SERT: self-explanation reading training. *Discourse Process* 38 (1), 1–30. https://doi.org/10.1207/s15326950dp3801_1.
- McNamara, D.S., 2010. Strategies to read and learn: overcoming learning by consumption. *Med. Educ.* 44 (4), 340–346. <https://doi.org/10.1111/j.1365-2923.2009.03550.x>.
- McNamara, D.S., 2011. Measuring deep, reflective comprehension and learning strategies: challenges and successes. *Metacogn. Learn.* 6 (2), 195–203.
- McNamara, D.S., 2017. Self-explanation and reading strategy training (SERT) improves low-knowledge students' science course performance. *Discourse Process* 54 (7), 479–492.
- Millis, K.K., et al., 2018. *Deep Comprehension: Multi-disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension*. Routledge.
- Mokhtari, K., Reichard, C.A., 2002. Assessing students' metacognitive awareness of reading strategies. *J. Educ. Psychol.* 94 (2), 249–259. <https://doi.org/10.1037/0022-0663.94.2.249>.
- Moll, L.C., et al., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theory Pract.* 31 (2), 132–141.
- Murphy, P.K., Mason, L., 2006. Changing knowledge and changing beliefs. In: Alexander, P.A., Winne, P. (Eds.), *Handbook of Educational Psychology*, second ed. Erlbaum, Mahwah, NJ, pp. 305–324.
- Murphy, P.K., Alexander, P.A., Muis, K.R., 2012. Knowledge and knowing: the journey from philosophy and psychology to human learning. In: *APA Educational Psychology Handbook, Vol. 1: Theories, Constructs, and Critical Issues*. American Psychological Association, pp. 189–226.
- Palincsar, A.S., Brown, A.L., 1986. Interactive teaching to promote independent learning from text. *Read. Teach.* 39 (8), 771–777.
- Pintrich, P.R., 2002. The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Pract.* 41 (4), 219–225.
- Rapp, D.N., Braasch, J.L., 2014. *Processing Inaccurate Information: Theoretical and Applied Perspectives From Cognitive Science and the Educational Sciences*. MIT Press.
- Rapp, D.N., et al., 2014. Reducing reliance on inaccurate information. *Mem. Cognit.* 42 (1), 11–26.

- Rapp, D.N., 2016. The consequences of reading inaccurate information. *Curr. Dir. Psychol. Sci.* 25 (4), 281–285.
- Richter, T., Schroeder, S., Wöhrmann, B., 2009. You don't have to believe everything you read: background knowledge permits fast and efficient validation of information. *J. Pers. Soc. Psychol.* 96 (3), 538.
- Risko, V.J., Walker-Dalhouse, D., 2007. Tapping students' cultural funds of knowledge to address the achievement gap. *Read. Teach.* 61 (1), 98–100.
- Rittle-Johnson, B., Siegler, R.S., 1998. The relation between conceptual and procedural knowledge in learning mathematics: a review. In: *The Development of Mathematical Skill*, pp. 75–110.
- Rosenshine, B., Meister, C., 1994. Reciprocal teaching: a review of the research. *Rev. Educ. Res.* 64 (4), 479–530.
- Ryle, G., 1949. *The Concept of Mind*. Barnes and Noble Inc, New York.
- Saxe, G.B., 1988. The mathematics of child street vendors. *Child Dev.* 1415–1425.
- Scardamalia, M., Bereiter, C., 2005. Does education for the knowledge age need a new science. *Eur. J. Sch. Psychol.* 3 (1), 21–40.
- Schiefele, U., 1990. The influence of topic interest, prior knowledge, and cognitive capabilities on text comprehension. In: *Learning Environments*. Springer, pp. 323–338.
- Schommer, M., 1990. Effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psychol.* 82 (3), 498.
- Scribner, S., Cole, M., 1973. Cognitive consequences of formal and informal education. *Science* 182 (4112), 553–559.
- Sinatra, G.M., Kienhues, D., Hofer, B.K., 2014. Addressing challenges to public understanding of science: epistemic cognition, motivated reasoning, and conceptual change. *Educ. Psychol.* 49 (2), 123–138.
- Smith III, J.P., Disessa, A.A., Roschelle, J., 1994. Misconceptions reconceived: a constructivist analysis of knowledge in transition. *J. Learn. Sci.* 3 (2), 115–163.
- Squire, L.R., 1992. Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *J. Cognit. Neurosci.* 4 (3), 232–243.
- Thacker, I., et al., 2020. Using persuasive refutation texts to prompt attitudinal and conceptual change. *J. Educ. Psychol.* 112 (6), 1085.
- Thiede, K.W., Anderson, M., Theriault, D., 2003. Accuracy of metacognitive monitoring affects learning of texts. *J. Educ. Psychol.* 95 (1), 66.
- Tippett, C.D., 2010. Refutation text in science education: a review of two decades of research. *Int. J. Sci. Math. Educ.* 8 (6), 951–970. <https://doi.org/10.1007/s10763-010-9203-x>.
- Trabasso, T., van den Broek, P., 1985. Causal thinking and the representation of narrative events. *J. Mem. Lang.* 24 (5), 612–630.
- Trevors, G.J., et al., 2017a. Adolescents' epistemic profiles in the service of knowledge revision. *Contemp. Educ. Psychol.* 49, 107–120.
- Trevors, G.J., Kendeou, P., Butterfuss, R., 2017b. Emotion processes in knowledge revision. *Discourse Process* 54 (5–6), 406–426.
- Tulving, E., 1972. *Episodic and Semantic Memory*, vol. 1. Organization of Memory, pp. 381–403.
- van den Broek, P., Kendeou, P., 2008. Cognitive processes in comprehension of science texts: the role of co-activation in confronting misconceptions. *Appl. Cognit. Psychol.* 22 (3), 335–351.
- van den Broek, P., 2010. Using texts in science education: cognitive processes and knowledge representation. *Science* 328 (5977), 453–456.
- Vosniadou, S., Brewer, W.F., 1992. Mental models of the earth: a study of conceptual change in childhood. *Cognit. Psychol.* 24 (4), 535–585.
- Vosniadou, S., Brewer, W.F., 1994. Mental models of the day/night cycle. *Cognit. Sci.* 18 (1), 123–183.
- Vosniadou, S., Skopeliti, I., 2017. Is it the Earth that turns or the Sun that goes behind the mountains? Students' misconceptions about the day/night cycle after reading a science text. *Int. J. Sci. Educ.* 39 (15), 2027–2051. <https://doi.org/10.1080/09500693.2017.1361557>.
- Vosniadou, S., 2009. *International Handbook of Research on Conceptual Change*. Routledge.
- Voss, J.F., Vesonder, G.T., Spilich, G.J., 1980. Text generation and recall by high-knowledge and low-knowledge individuals. *J. Verb. Learn. Verb. Behav.* 19 (6), 651–667.
- Wiley, J., et al., 2020. Epistemic beliefs about the value of integrating information across multiple documents in history. *Learn. Instr.* 65, 101266.
- Winne, P.H., 1996. A metacognitive view of individual differences in self-regulated learning. *Learn. Individ. Differ.* 8 (4), 327–353. [https://doi.org/10.1016/S1041-6080\(96\)90022-9](https://doi.org/10.1016/S1041-6080(96)90022-9).
- Winne, P.H., 2018. Cognition and metacognition within self-regulated learning. In: *Handbook of Self-Regulation of Learning and Performance*, second ed. Routledge/Taylor & Francis Group (Educational Psychology Handbook Series), New York, NY, US, pp. 36–48.
- Wood, P., Kardash, C., 2002. Critical elements in the design and analysis of studies of epistemology. In: *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 231–260.
- Zengilowski, A., Schuetze, B.A., Nash, B.L., Schallert, D.L., 2021. A critical review of the refutation text literature: methodological confounds, theoretical problems, and possible solutions. *Educ. Psychol.* 56 (3), 175–195.

Memory[☆]

Tim Brennen and Svein Magnussen, Department of Psychology, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

This is an update of S. Magnussen, T. Brennen, Memory, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 197–203, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00481-4>.

Introduction	220
Varieties of human memory	220
Declarative memory	220
Nondeclarative memory	221
Short-term memory	221
Working memory	221
Prospective memory	221
Conceptualizing memory and its failures	222
Metaphors	222
Factors that make memory inaccurate	222
False autobiographical memories	223
Memory in everyday contexts	224
Memory for people	224
Long-term autobiographical memory	224
Memory training	225
Memory in the classroom	225
Testing improves retention	225
Long-term retention of knowledge from school	225
The future of memory science	226
Cognitive neuroscience	226
The seductive allure of neuroscience explanations	226
Dissemination and implementation of the established science	227
References	227
Further reading	228

Glossary

Autobiographical memory The sum total of one's memory for one's life makes up one's autobiographical memory. In addition to memory for specific episodes, it contains knowledge of the phases of one's life, of repeated experiences and of pieces of information like telephone numbers and addresses, past and present.

Declarative memory An essential property of declarative memory is that one is conscious of what one is remembering. Such memories are, by their very nature, verbal. For example, one can talk about one's experiences, and about new words that one has learned.

Episodic memory Personal experiences and events are stored in episodic memory (see declarative memory).

False memory A memory that appears real to the person remembering, but is not based on veridical experience is referred to as a false memory. They can be induced in the laboratory, but they are also prevalent in everyday situations, sometimes with drastic legal and interpersonal consequences.

Forgetting It is the inability to recall an episode or a skill that was previously learned. In forgetting of declarative memory (*quod vide*), it is suspected that many failures to recall are in fact due to a failure to access a still-existent memory trace.

Memory systems In contrast to previous schools of thought in psychology, where memory was considered to be unitary, in cognitive psychology one considers memory to be a multifaceted set of modules. There is a host of evidence to suggest that we possess many different types of memory.

[☆]*Change History:* October 2020. The article has been substantially revised and expanded. TJB changed the order of the authors. Opening anecdote about the plot of *The Bourne Identity*, reduced the "Dodo" one. Wrote about Neuroscience and Seductive explanations. SJM wrote about prospective memory, memory training, updated the face recognition paragraph, updated the section on testing in the classroom, and added a description of the [Baddeley \(2012\)](#) study in the section on long-term autobiographical memory. TJB added Autobiographical memory to the Glossary, wrote the Abstract, added a paragraph about number of years spent in education worldwide and the Sustainable Development Goals, rewrote the Memory for People section, updated the False memory section, revised the Dissemination section, and updated the Further Reading. TJB & SM revised according to the feedback, added some relevant recent references, expanded on long-term autobiographical memory.

Nondeclarative memory We are not aware of everything that affects our behavior. These nonconscious effects of previous experience are evidence of nondeclarative memory. Also, our memory for motor skills, such as walking, typing, and skiing, is nondeclarative because motor skills are essentially nonverbal.

Perceptual representation system Conceived of as the interface between perception and memory, it is a basic memory system storing perceptual information that enables us to perceive the world as meaningful.

Priming Things that you have just seen, heard or thought about will facilitate thoughts about related concepts. This expression of memory is referred to as priming.

Semantic memory Believed to be stored separately from personal experiences or episodes, semantic memory refers to a person's general knowledge of the world (see episodic memory).

Working memory A coordinated set of short-term memory mechanisms that allow us to consciously manipulate what we are thinking about at any given moment.

Introduction

Memory is a central part of the brain's attempt to make sense of experience, and to tell coherent stories about it. These tales are all we have of our past, so they are potent determinants of how we view ourselves and what we do. Yet our stories are built from many ingredients. Snippets of what actually happened, thoughts about what might have happened, and beliefs that guide us as we attempt to remember. Our memories are the powerful but fragile products of what we recall from the past, believe about the present, and imagine about the future (Schacter, 1996, p. 308).

The 2002 film *The Bourne Identity* starts with Matt Damon's character being rescued by fishermen off the coast of Marseille. Once aboard it soon transpires that although he can still speak English and several other languages he does not know who he is, how he came to be in the sea, what music he likes, or indeed what his name is. He says "I can read, I can write, I can add, subtract, I can make coffee, I can shuffle cards, I can set up a chessboard" and yet he has no recollection of his prior life. Contrary to what you might think, this Hollywood script is not ridiculously implausible: Conditions similar to Jason Bourne's do in fact crop up. For instance, in January 2003, Dodo, a young man of Asian origin who, in January 2003, woke up outdoors in a small village in Switzerland with a well-equipped rucksack stuffed with expensive clothes and a money belt containing US\$5000, but no identity papers or tickets and with absolutely no personal memory. Dodo wandered around Europe for several weeks, and somehow managed to get to Oslo, Norway. His memory loss of the time before he woke up in Switzerland was massive, he had no idea who he was, and he did not recognize his own face in the mirror. He had even lost his native language – he spoke heavily accented English but not any Asian language; the only things that he knew about himself were that he smoked Camel cigarettes and liked pop music.

Jason Bourne and Dodo suffered from the condition of retrograde amnesia. That is, they lost their memory in the sense people usually use the term, the recollection of private experiences and facts we have learned about the world. But as mentioned, Bourne was still able to draw on past experience in order to read and to shuffle cards, and Dodo knew how to operate photographic equipment and understood the value of money. This type of selective memory impairment is a main argument cited by memory researchers in support of the idea that memory, rather than being a single cognitive process or system, is a collective term for a family of neurocognitive systems that stores information in different formats (Schacter et al., 2000; Tulving, 2002).

Varieties of human memory

Modern taxonomies distinguish between several forms of memory, or memory systems. An important distinction is drawn between declarative memory, which refers to the conscious recollection of facts and personal experiences, and nondeclarative memory, which refers to behavior changes resulting from previous experiences that may or may not be accompanied by conscious recollection (see Fig. 1).

Declarative memory

The story of Dodo illustrates the central role of declarative memory in human life. This young man had lost not only his personal past – his autobiographical memories – but also large parts of his general knowledge of the world and even his ability to speak his native language. Thus, the systems or forms of memory that we term episodic and semantic memory are heavily affected. Episodic memory is assigned a special role in human life: it is unique in the sense that memories are associated with a place and a time, and is, according to Tulving (2002), the only known example of a process where the arrow of time is turned back and the past can be re-experienced. Without episodic memory, the mental representation that psychologists call the self – the organization of personal memories in a historical context – is lost. Obviously, this is what had happened to Dodo though his nondeclarative memory seemed largely intact.

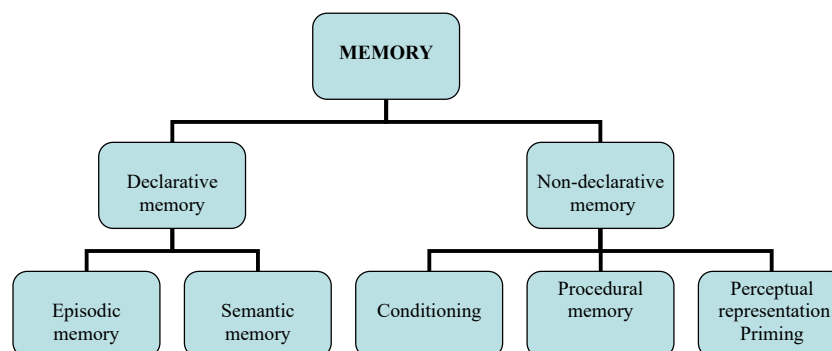


Fig. 1 A common textbook taxonomy of human memory.

Nondeclarative memory

In nondeclarative memory, the effects of previous experiences and actions one has carried out manifest themselves directly in behavior; the individual learning sessions may only be vaguely remembered or even be completely inaccessible to conscious recollection. Procedural memory is responsible for the maintenance of the cognitive and motor skills that we have acquired throughout life, from knowing how to eat with a knife and fork to the mastery of swimming, cycling, or driving a car, as well as the advanced skills of playing billiards or playing a saxophone. Conditioning is a basic memory system that in humans is particularly important in tying emotional reactions to external stimuli or situations. For example, in phobic reactions, anxiety is triggered by the phobic stimuli but the person is typically not aware of the learning episodes in which the connection between the emotion and stimulus was established. Similarly, a piece of music or a specific odor may evoke romantic feelings without an accompanying experienced-memory episode. Perceptual learning, or the perceptual representation system (PRS), enables us to perceive the world as consisting of meaningful entities, because in order to produce a perceptual experience, online sensory signals must reach stored representations. To see is to recognize, or to realize that you do not recognize. [Tulving and Schacter \(1990\)](#) also identified priming as an expression of nondeclarative memory, coupled to the PRS. In priming, the person's performance on a specific task is facilitated (or inhibited) by the previous presentation of related information. For example, perceptual identification of a picture of Melania Trump is speeded by the prior presentation of a picture of Donald Trump, compared to the prior presentation of a picture of, say, Tom Hanks.

An important feature of memory systems is that they store information in different formats, so that the information in one store is not directly translatable into another. However, because the memory systems operate independently and in parallel, most experiences would be recorded and stored in different formats, and an act of remembering would be executed by several memory systems working in concert ([Tulving, 2002](#)).

Short-term memory

Another basic distinction in memory theory is the division between a short-term memory mechanism that is the seat of consciousness and active processing, and which is only able to store limited quantities of information for a limited period of time, and a long-term memory mechanism that stores unlimited amounts of information for unlimited periods of time ([Atkinson and Shiffrin, 1968](#)).

Working memory

A well-developed model of short-term memory is [Baddeley's \(1986, 2003\)](#) concept of working memory, a coordinated set of mechanisms that combine incoming information with information retrieved from long-term memory, assisted by three support systems that actively manipulate verbal, visuo-spatial, and episodic information. The concept of working memory is closely associated with the concept of attention; only those pieces of information that are actively attended to, enter memory. Since the capacity of attention and working memory is limited, it follows that only part of the potentially available information triggering our senses at any one time is registered, and much of the information that is registered is immediately forgotten. In a sense, most of the things we believe we have forgotten, never really entered a durable memory system.

Prospective memory

We usually think of memory as the recollection of things from the past, but in addition to retrospective memory, prospective memory tasks fill our daily lives. Prospective memory involves remembering future events, things we have to do or have planned to do. Some of these memory tasks represent automatic routines, such as turning off the lights when leaving the house; others are one-off events, such as buying theater tickets for the kids' Christmas show. For the latter sort of tasks, many of us nowadays keep some sort of digital diary to remind us to perform the action, but less so for the things that need carrying out daily and it is these

automatic, prospective memory tasks that challenge us in daily life as they are easily interrupted by simultaneously ongoing tasks. Failures of prospective memory may cause severely ill patients to forget to take their medication (Zogg et al., 2012; Avci et al., 2018), and may be responsible for a number of tragic aircraft accidents (Eysenck, 2020a).

Conceptualizing memory and its failures

Metaphors

In the attempt to understand memory, scientists as well as laymen are forced to try to describe something that is not directly observable. We may be aware of some of the processes that take place during learning, and similarly of the processes that take place when we attempt to search for a piece of information in memory, but we have little access to the processes that mediate between encoding and retrieval. In order to make sense of the phenomena to which we have little access, both laymen and researchers tend to use metaphors and analogies, many of which are borrowed from the physical world (Magnussen et al., 2007).

Koriat et al. (2000) pointed out that the study of memory following Ebbinghaus (1885) has been dominated by a storehouse metaphor, in which discrete items of information are stored and later retrieved, and which tends to evaluate memory in terms of the number of items that are retained (or lost). The storehouse metaphor invites an interest in the number of stored items. An effective store is one that contains many items, retains these items for long periods of time, and allows easy access to them. Memory then is evaluated in terms of its quantity – how much is retained, how much is lost. Quantitative aspects of memory are frequently of great interest in everyday life, and research relevant to the classroom and academic achievement typically uses this metaphor.

An alternative approach to the study of memory derives from the correspondence metaphor that can be traced to the work of Bartlett (1932) who had British participants listen to a Native American folk tale called *The War of the Ghosts*. When recalling the story repeatedly over weeks, the memories became shorter, and details dropped out and were replaced by others that were more familiar to the participants; for example, “seal hunt” would be remembered as “fishing expedition” and “canoe” as “boat”. The narratives often became coherent and well-polished, but also unrecognizable as the original. Bartlett’s claim that these results demonstrate the constructive nature of memory is still valid today. In this type of paradigm memory performance is evaluated in terms of the correspondence between the original event and the memory of it: The accuracy of the memory report. A high-quality memory is more faithful to the remembered event than an inferior memory, and contains fewer distortions and errors. Sometimes people remember events that never happened. The assessment of memory correspondence should therefore start with the output – what the person reports – rather than with the input – what actually happened. An output-bound assessment reflects the accuracy of what is remembered – how much of what the person reports did in fact occur (Koriat et al., 2000). In many real-life situations, such as in court, there is much more emphasis on the accuracy of the report than with the amount of information reported.

Factors that make memory inaccurate

The modern study of the accuracy of episodic memory has convincingly disproved another popular metaphor, the idea that our memory is like digital technology where you press “Play” and get the same performance each time. Episodic memory does not reproduce, it constructs, and the reconstruction of previous episodes is based on information from many sources with the assistance of many neural systems (Rubin, 2006). As some of the sources of information used in the construction of episodic memories are external to the original event, memory accuracy suffers. In a widely read book, Schacter (2001) identified seven factors that are behind the tricks that episodic memory plays on us, and called them the “seven sins of memory”. One of the sins is *absent-mindedness*: To be encoded into episodic memory, a piece of information or an event must be attended to. When focal attention is diverted or attention is switched to auto-pilot, information may be missed, and the event is not properly remembered because the information was never registered in the first place. The sin of *transience* refers to fact that memories fade. Some memories fade rapidly, others fade more slowly, depending on a number of factors, but left to itself no memory gets better over time. The general course of memory decay, first described by Ebbinghaus (1885), is the negatively accelerated function shown in Fig. 2A, which represents the long-term forgetting of ordinary life events.

However, as illustrated by Fig. 2B, frequently studied or over-learned material shows very little decay, and some extraordinary life events, in particular dramatic and traumatic experiences, appear to be very resistant to fading and rather obey the sin of *persistence*. In fact, memories we might like to forget can haunt us for the rest of our lives. And sometimes information that we have stored is unavailable to memory because of a failure of retrieval. It feels close – it might be on the tip of your tongue – but the sin of *blocking* renders it temporarily irretrievable. Memories not only fade, they are subject to influences from several sources. The human mind is *suggestible*, our memories of events and facts are influenced by information from external sources, from what we read, and from what other people tell us. Memories are subject to *bias* in various ways. Our memories of the past depend on and are colored by the present. That is, your current attitude to your boss or partner influences how you now recall your feelings about them many years ago. The sin of *misattribution* refers to the confusions and mistakes we often make concerning the source of a particular piece of information, confusions of times, people, and places, and in the extreme case, the confusion of genuine memory and fantasy that gives rise to false memories.

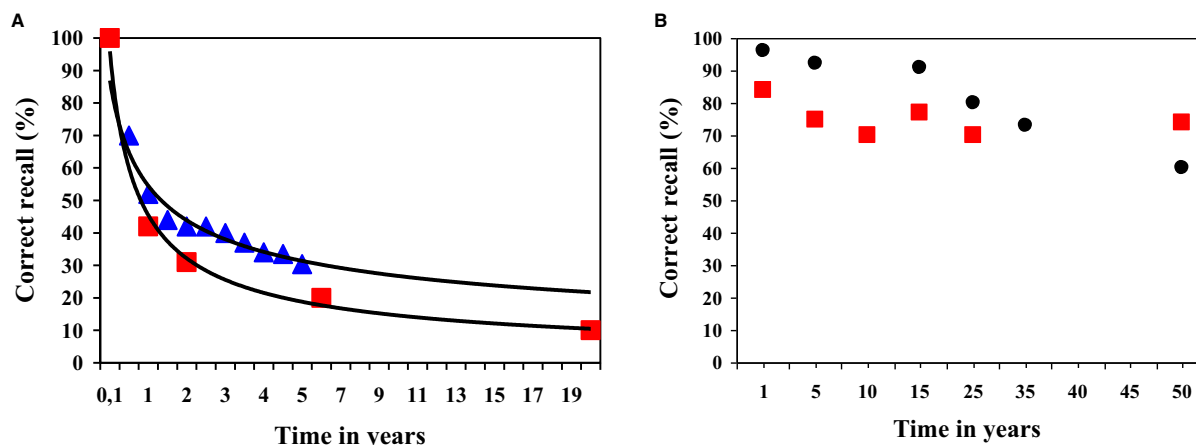


Fig. 2 (A) Memory for personally experienced everyday events, Wagenaar (1986) plotted with triangles and White (2002) plotted with squares. (B) Very long-term memory for over-learned material, obtained by asking participants to match pictures of high-school classmates “faces to their names, (circles), and testing participants’ ability to translate unused Spanish vocabulary learned in school to English, (squares). Despite the fact that we treasure our autobiographical memories and they play an integral role in our personal identity, we become incapable of bringing to mind the specifics of most of what we experience within a year or so of it happening. By contrast, multiple repeated exposures to something that is quite unimportant for one in later life and never practiced or recalled can allow access to memories decades later. This may be seen as paradoxical until one realizes that motivation to remember is not always as important as the amount of practice one has put in to learning something. (A) Adapted from Wagenaar, W., 1986. My memory: a study of autobiographical memory over six years. *Cognit. Psychol.* 18, 225–252; White, R., 2002. Memory for events after twenty years. *Appl. Cognit. Psychol.* 16, 603–612. (B) Adapted from Bahrnick, H.P., Bahrnick, P.O., Wittlinger, R.P., 1975. Fifty years of memories for names and faces: a cross-sectional approach. *J. Exp. Psychol. General* 104, 54–75; Bahrnick, H.P., 1984. Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *J. Exp. Psychol. General* 113, 1–29.

False autobiographical memories

In the 1990s, the science of memory was the subject of a fierce debate in scientific journals, newspapers, and courts of law. A phenomenon that was first observed in the United States but that has since spread to other countries is that cases of alleged child abuse were being brought to court by the victims, years and even decades later, because the victims had previously been unable to retrieve the memories. This was a powerful and deeply shocking claim that split psychology into those who thought of these memories as recovered, and those who thought of them as false, the so-called memory wars.

Since Bartlett’s seminal research on how our memory for narrative introduces errors that fit in with the rest of our knowledge, researchers have known about the malleability of our memories for events. Due to the memory wars, such memory illusions have been much studied over the past decade, showing that under certain circumstances, we can become convinced that, several years previously, one had been up in a hot air balloon, or had knocked over a punch bowl at a wedding when these events simply had not happened (Loftus, 2003). The key ingredients appear to be that one believes the authority telling one the information, and that one is given time to try to retrieve the information: over time, the false memories come to feel real to the person remembering. There is then a danger in the therapy situation where a clinician meets a patient, if the clinician believes that symptoms of depression and anxiety are in fact the manifestation of experienced child abuse that the patient has forgotten. There are documented cases where entirely false accounts of abuse have been induced in patients. These “recovered memories” may be highly implausible but nevertheless genuinely believed by the patient. They are produced by the combination of an authority figure, in this case the therapist, and several practice sessions in which the patient tries to “recover” the memories. A number of documented cases show that false memories may be quite dramatic, as for example, remembering being witness to a murder that never happened (Goodman et al., 2007). Modern research also fails to support the idea that early traumatic childhood experiences can be stored in a detailed fashion but repressed or blocked from consciousness, only later to be exhumed by therapeutic exercises. Rather, well-controlled studies show that such memories obey the sin of persistence (Goodman et al., 2003).

Knowledge of the constructive nature of autobiographical memory is important for many professions, especially those dealing with psychiatric patients, children, and other groups with little power. The vulnerability of autobiographical or episodic memory is also the focus of another important area of applied memory research, the study of factors affecting the reliability of eyewitness testimony (Howe and Knott, 2015). In the interrogation room and in court, failures to remember are less damaging to the judicial process than are memory errors; absence of memory points nowhere, memory errors may point in the wrong direction.

In memory implantation experiments it is challenging for researchers to reliably categorize participants as having experienced a false memory or not. Different ways of coding the data give different estimates of what proportion of participants fully accepted the false memory (Brewin and Andrews, 2017; Scoboria et al., 2017). However, even those researchers urging caution accept that a nontrivial proportion do end up convinced that the false event in fact occurred.

Memory in everyday contexts

Ebbinghaus (1885) pioneering studies investigated his own memory for nonsense syllables, and a century later, as cognitive psychology became the dominant school in psychology in the 1970s, the typical experiment on memory still clearly bore Ebbinghaus' imprint: university undergraduates sat in a darkened lab and learned long lists of words. A degree of discomfort with the potential lack of relevance of these experiments for understanding life outside the lab, led to a push for memory research to be more ecologically valid, and the study of applied memory has since grown into a very active research field. Drawing on the correspondence metaphor of memory, a key underlying question in this field has been how accurate our memory is, as it is used in everyday life. Another change has been in the types of material used: breaking out of the verbal learning tradition (learning lists of words), a concerted research effort over the past 30 years has, for instance, tried to shed light on our ability to remember more real-life stimuli, like people, and events we have experienced in our lives (a skill that Dodo apparently no longer possesses).

Memory for people

A person represents an extremely complex and multifaceted biological stimulus: While we often recognize a person from their face, we can also judge who a person is from their name, their voice, their gait, besides many other potential routes. Not all types of person information are equally memorable. People sometimes say "I forget names but never faces"; the scientific literature shows that while the first part of the claim is likely to be true, the second is only partly so. People's names do pose a problem when it comes to recall. All readers will have experienced the tip-of-the-tongue phenomenon for a name, when one is unable to recall a name but feels that one might recall it at any time, and may be able to give partial information about the missing word. This phenomenon occurs particularly for people's names, and its frequency increases with age (Valentine et al., 1996).

When it comes to remembering faces, it is true that we have good memory for people we have spent lots of time with or frequently viewed in public media. Humans are social animals heavily reliant on vision and there is a strong need to distinguish between family, friends, and enemies, for instance. However, our memory for faces that we have only seen once is very poor. For briefly-viewed unfamiliar faces, there is little evidence of any memory after one year (Deffenbacher et al., 2008). Even comparing a briefly-seen face to a series of pictures taken the same day yields 10% errors, and when the line-up pictures were taken 3 months previously 30% of the identifications were incorrect (Megreya et al., 2013). Thus we are not gifted at matching a person to their photo ID (which may have been taken several years previously), and indeed shop assistants are terrible at detecting that the face on the card does not match the person proffering the card (Kemp et al., 1997). Similarly, this work raises questions about how reliable lineup identifications are in criminal cases. The evidence suggests they are indeed not very good; results from a recent analysis of lineups showed that about 30% of the identifications by witnesses are of a person in the lineup who is known to be innocent (Wells et al., 2020).

Long-term autobiographical memory

How much of our lives do we remember? Cognitive psychologists tend to agree with the sentiments of the novelist Milan Kundera (2002) who pointed out that in fact we are only able to recall a vanishingly small proportion of the events in our lives. While we do of course remember many individual events from our lives, much of our autobiographical memory appears to be based on generalized averages of similar events, so that while you may only be able to specifically recall a few of the times you have been to the Odeon cinema, you may know very well where you usually park, what sort of films you see, who you go with, and what snacks you buy. A systematic approach to answering the question was taken by Wagenaar (1986) who wrote down one incident from his life each day over a period of 4 years. When subsequently tested, he found that being cued by the who, what, or where of an event was more powerful a cue than when, and that a single piece of information (who, what, or where) triggered recall of the incident on only about 50% of occasions for the first year and declined to 20% after 4 years. Multiple cues triggered over 60% of memories, even after 4 years, and even cases where he had apparently forgotten the incident completely could be recalled when the event involved another person who could be asked to provide more cues. Thus, it is a case of the glass being half-full or half-empty: Even events which were thought to be completely forgotten turned out to be retrievable, but on the other hand, normal cues failed to trigger a majority of memories. Results from Wagenaar's (1986) report together with results from a similar study carried out across a 20-year interval by White (2002) are shown in Fig. 2A and illustrate the dramatic decline of ordinary everyday memories. White (2002) tested his memory by having an assistant read event descriptions picked randomly, and he decided if he remembered the event vividly, vaguely, or not at all. Most of the remembered events were only vaguely remembered, and he encountered not a single case where an event forgotten at one point in time was later remembered.

Do we eventually forget almost everything from our past? Of course not. These studies tested memory for isolated events, short episodes that were not embedded in a broader context. But autobiographical memory - what we remember from our lives - is more than a collection of isolated events. We remember events integrated as part of more extended memory schemata that form our coherent life story and concept of self (Courage and Howe, 2010). Episodes are understood and remembered in the context of longer life periods, which can last for several years and tend to be personally significant, such as our years in college, our romantic relationship(s), holidays abroad and professional positions. White also kept a diary from January 1978 to December 1980 when he among other things worked temporarily at another university and participated in several scientific conferences, and found that, forty years after the original events, he spontaneously remembered a lot of events and single episodes when cued by these extended

events. On the other hand, reminiscent of the question of the glass being half-full or half-empty mentioned above, two-thirds of the events he remembered 20 years after they occurred were unavailable 40 years after (White, 2020).

In 2012 Alan Baddeley, a major figure in memory research, performed an informal test of his own autobiographical long-term memory via letters sent to his mother during a year spent in the US 40 years previously. Already known as a prolific writer, we thereby discovered that he was also a devoted son, and because his mother kept the letters it was possible to identify events from his distant past. Of 62 episodes he was able to remember 23, but only half of them vividly. It may sound disappointing that only 20% of our past is remembered, but, as suggested by Kundera, the truth may be even more sobering: These long-term studies may vastly overestimate our everyday memory performance for two reasons. Firstly, the selection of episodes to be remembered is probably not random so that trivial events such as “I had Chinese take-away for supper” are unlikely to be recorded and secondly the memory is strengthened by the very fact that it is chosen and written down. Classical studies where participants are required to record what they are doing at random intervals have shown that only one-fifth of such episodes are remembered after six weeks, which may also be an overestimation due to the rehearsal effect. Our memory for details of daily events and episodes is even worse. In a recent study (Misra et al., 2018), participants wearing a head-mounted camera and eye-view recorder walked a pre-determined route in a city for about an hour. Their memory for the walk was tested a day later by having the participant choose between 2-s recordings of their own walk and walks through the same route by other participants, who met other people, other cars, and walked under different light conditions. The results showed that participants scored close to chance. Most of the details we experience in daily life disappear very quickly indeed.

Memory training

Can memory be trained? This is a hot topic, and quick-fix courses flourish in the popular media. Is it possible through mental exercises to increase one's memory capacity and make your memory performance more reliable or even spectacular? The answer is both yes and no. There are well-known ancient memory techniques that are effective for most of us when trying to remember things such as a shopping list or the names of a new group of people, and memory experts who excel on memory tasks that require the repetition of vast numbers of single items rely on such techniques. It has been known for many years, however, that the effects of memory training are restricted to the tasks one has spent time perfecting (typically list learning) or very similar tasks, with little or no transfer to memory performance in general (De Beni et al., 2007). Thus, practicing learning of word or digit lists does not make you better at remembering books you read or events you have witnessed: There is no way of enhancing your general memory performance (Eysenck, 2020b).

Memory in the classroom

The vast majority of children in the West spend upwards of a decade of their lives going to school. Sadly, on the poorer continents many children, especially girls, spend fewer years than this in formal schooling. In Niger, which in 2018 was the lowest ranked country on the Human Development Indices and Indicators, a child can only expect to spend five years in school (UNDP, 2019). The UN's fourth Sustainable Development Goal is to “Ensure inclusive and equitable quality education and promote life-long learning opportunities for all”. Delivering a quality education includes documenting that pupils have managed to learn, and memory science can be a source of advice as to how to achieve this.

Testing improves retention

It might be surprising to some that a body of evidence has built up recently to suggest that the best way to study for a test is to test oneself. For instance, Roediger and Karpicke (2006) asked participants to read and learn a short essay. After all participants had read the passage through, one group was then given a test on what they had remembered, whereas the other group reread the passage. There was then a final memory test given, either 5 min, 2 days, or a week subsequent to the reading. After 5 min, those participants who had reread the passage performed better than the other group. However those tested after 2 days and those tested after a week showed the opposite pattern: an extra opportunity to study the material led to worse memory compared to those who were tested in the first phase (see Fig. 3).

The testing effect has proved to be very robust, confirmed by a number of subsequent studies conducted in the laboratory (Rowland, 2014) and real-life contexts (Schwieren et al., 2017), as well as by a recent review (Yang et al., 2021). The message from this literature is clear: the very act of testing improves performance. It is to be hoped that educational systems will confront the challenge to incorporate the consequences of this finding into their classrooms: testing is not only a way of assessing students' performance, it also a way of facilitating their learning.

Long-term retention of knowledge from school

How much of the knowledge one acquires in school becomes lifelong? One of the secrets to creating robust long-term memories, known to science since at least the 19th century and known to centuries of teachers, is repetition or over-learning. Schoolchildren spend a lot of time in classrooms learning from teachers, books and digital media, but the exams that test if the requisite knowledge

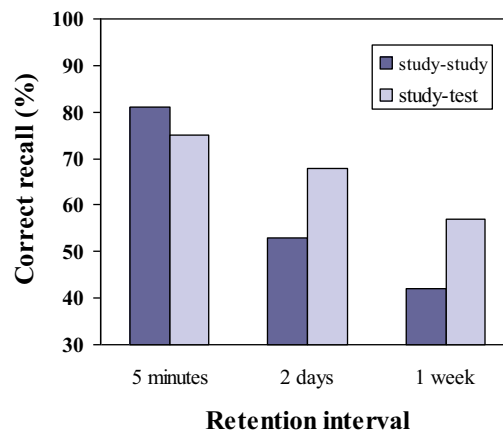


Fig. 3 Having heard a story they were trying to learn, some participants were immediately tested on what they could recall, while others had a second chance to learn the story. Five minutes later, the second-chance group (*the dark column*) had better recall, but of potential importance to educationalists, the immediate-test group (*the light column*) did better on tests 2 and 7 days later. Adapted from Roediger, H.L., Karpicke, J.D., 2006. Test-enhanced learning – taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255.

has been acquired usually take place within a few months of the teaching, which begs the question of the fate of this knowledge in the long term. How much of the Spanish vocabulary that you learned in high school can you still recall? What about the fate of your university degree that you haven't used since your graduation exams? Seminal research by Harry Bahrick has shown that we are able to recall surprising amounts of knowledge even several decades after learning (see for example, [Bahrick, 1984](#)). Lots of material seems to become unavailable within the first year of learning. About 2 years after studying a topic, people are able to perform at around 50% of the level they were at when they stopped studying, but the surprising finding is that up to 50 years later, one performs at about the same level as one did 2 years after: very little forgetting appears to occur after an initial period. Similar results have been reported for mathematics, Spanish, and for cognitive psychology, so that this function appears to reflect general properties of the long-term semantic memory system, or *permastore* as Bahrick termed it (see [Fig. 2B](#)).

The future of memory science

Memory research finds itself at a touchpoint between basic and applied research. It is becoming more integrated into neuroscience, with studies using cognitive tasks in conjunction with brain imaging methods, in particular functional magnetic resonance imaging (fMRI) techniques, by which brain processes may be visualized online concurrently with ongoing cognitive tasks. Has this integration of memory research and fMRI advanced our understanding of memory?

Cognitive neuroscience

While brain science still holds the promise of one day providing an integrated multilevel account of memory function that is applicable in practical contexts like schools, the current stage of cognitive neuroscience is rather distant from this. On a pessimistic view, it is somewhat bound by its laboratory methods where, typically, participants are placed in physically restricting postures performing esoteric yet trivial intellectual tasks; small effects are observed in complex patterns of data, the interpretations of which have almost limitless degrees of freedom, and which invariably end up confirming the original cognitive research anyway. These challenges come in addition to the finding that frequently-used brain scanning measures have low reliability ([Elliott et al., 2020](#)).

The seductive allure of neuroscience explanations

This would perhaps not need pointing out were it not for the generally overoptimistic attitudes to neuroscience in society at large, referred to as the “Seductive Allure of Neuroscience Explanations” ([Weisberg et al., 2008](#)). In 2014, a large number of psychologists and neuroscientists signed a document decrying the claims of the “brain training industry” which, for a fee, promises to increase your IQ or to stave off dementia by means of some sessions of concentrated practice on some cognitive tasks ([A Consensus on the Brain Training Industry from the Scientific Community, 2014](#)). Another reason for a sober approach to the ability of neuroscience to deliver is the effect reported by Weisberg et al. showing that when judging how good explanations of psychological phenomena are, nonexperts mistakenly give positive weight to irrelevant neuroscientific information. [Im et al. \(2017\)](#) demonstrated the same effect when members of the public read articles about educational issues; in particular, they found that articles were rated as more credible when the spurious neuroscience information was accompanied by a picture of a brain.

Dissemination and implementation of the established science

On the other hand, a reason for optimism is that the emphasis on an understanding of memory in everyday contexts is strengthening. This is a welcome development because the focus in recent decades on more applied sides of memory – how it is actually used in daily settings – has produced results with important real-world implications. For all in the legal system, from the police to judges to jury members, the literature showing the fallibility of eyewitnesses' memory for people and events should be required reading. The application of the findings from the field of memory (and cognitive psychology more generally) to educational settings is beset with controversy (see for example Claxton, 2021). Regardless of one's position on, for example, the dichotomy between "progressive" and "traditional" teaching methods, one would hope that educational planners have, in addition to specific knowledge about phenomena like the positive gains obtained by the very act of testing someone's knowledge, an appreciation of the limits and character of memory.

References

- A Consensus on the Brain Training Industry from the Scientific Community, 2014. Max Planck Institute for Human Development and Stanford Center on Longevity. Available at: <http://longevity.stanford.edu/a-consensus-on-the-brain-training-industry-from-the-scientific-community-2/>. (Accessed 9 September 2020).
- Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. In: Spence, K.W., Spence, J.T. (Eds.), *The Psychology of Learning and Motivation*, vol. 1. Academic Press, New York, pp. 89–195.
- Avci, G., Sheppard, D.P., Tierney, S.M., Kordovski, V.M., Sullivan, K.L., Woods, S.P., 2018. A systematic review of prospective memory in HIV disease: from the laboratory to daily life. *Clin. Neuropsychol.* 32, 858–890. <https://doi.org/10.1080/13854046.2017.1373860>.
- Baddeley, A.D., 1986. *Working Memory*. Oxford University Press, Oxford.
- Baddeley, A.D., 2003. Working memory. Looking back and looking forward. *Nat. Neurosci. Rev.* 4, 829–839. <https://doi.org/10.1038/nrn1201>.
- Baddeley, A.D., 2012. Reflections on autobiographical memory. In: Berntsen, D., Rubin, D.C. (Eds.), *Understanding Autobiographical Memory: Theories and Approaches*. Cambridge University Press, Cambridge, pp. 70–88.
- Bahrack, H.P., 1984. Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *J. Exp. Psychol. Gen.* 113, 1–29. <https://doi.org/10.1037//0096-3445.113.1.1>.
- Bartlett, F.C., 1932. *Remembering*. Cambridge University Press, Cambridge.
- Brewin, C.R., Andrews, B., 2017. Creating memories for false autobiographical events in childhood: a systematic review. *Appl. Cognit. Psychol.* 31, 2–23. <https://doi.org/10.1002/acp.3220>.
- Claxton, G., 2021. *The Future of Teaching and the Myths That Hold It Back*. Routledge, London.
- Courage, M.L., Howe, M.L., 2010. Autobiographical memory: individual differences and developmental course. In: Gruszka, A., Matthews, G., Szymura, B. (Eds.), *Handbook of Individual Differences in Cognition*. Springer, New York, NY, pp. 403–417. https://doi.org/10.1007/978-1-4419-1210-7_24.
- De Beni, R., Magnussen, S., Cornoldi, C., Larsson, M., Rönnerberg, J., 2007. Memory experts. Visual learning, wine tasting, orienteering and speech reading. In: Magnussen, S., Helstrup, T. (Eds.), *Everyday Memory*. Psychology Press, Hove, UK, pp. 201–227.
- Deffenbacher, K.A., Bornstein, B.H., McGorty, E.K., Penrod, S.D., 2008. Forgetting the once-seen face: estimating the strength of an eyewitness's memory representation. *J. Exp. Psychol. Appl.* 14, 139–150. <https://doi.org/10.1037/1076-898X.14.2.139>.
- Ebbinghaus, H., 1885. *Über das Gedächtnis*. Dunker, Leipzig.
- Elliott, M.L., Knott, A.R., Ireland, D., Morris, M.L., Poulton, R., Ramrakha, S., Sison, M.L., Moffitt, T.E., Caspi, A., Hariri, A.R., 2020. What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychol. Sci.* 31, 792–806. <https://doi.org/10.1177/0956797620916786>.
- Eysenck, M.W., 2020a. Prospective memory. In: Baddeley, A., Eysenck, M.W., Anderson, M.C. (Eds.), *Memory*. Routledge, New York, pp. 425–445.
- Eysenck, M.W., 2020b. Improving your memory. In: Baddeley, A., Eysenck, M.W., Anderson, M.C. (Eds.), *Memory*. Routledge, New York, pp. 538–562.
- Goodman, G.S., Gheiti, S., Quas, J.A., et al., 2003. A prospective study of memory for child sexual abuse: new findings relevant to the repressed memory controversy. *Psychol. Sci.* 14, 113–118. <https://doi.org/10.1111/1467-9280.01428>.
- Goodman, G.S., Magnussen, S., Andersson, J., et al., 2007. Memory illusions and false memories in real life. In: Magnussen, S., Helstrup, T. (Eds.), *Everyday Memory*. Psychology Press, Hove, pp. 157–182.
- Howe, M.L., Knott, L.M., 2015. The fallibility of memory in judicial processes: lessons from the past and their modern consequences. *Memory* 23 (5), 633–656. <https://doi.org/10.1080/09658211.2015.1010709>.
- Im, S., Varma, K., Varma, S., 2017. Extending the seductive allure of neuroscience explanations effect to popular articles about educational topics. *Br. J. Educ. Psychol.* 87, 518–534. <https://doi.org/10.1111/bjep.12162>.
- Kemp, R., Towell, N., Pike, G., 1997. When seeing should not be believing: photographs, credit cards and fraud. *Appl. Cognit. Psychol.* 11, 211–222. [https://doi.org/10.1002/\(SICI\)1099-0720\(199706\)11:3<211::AID-ACP430>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1099-0720(199706)11:3<211::AID-ACP430>3.0.CO;2-O).
- Koriat, A., Goldsmith, M., Pansky, A., 2000. Toward a psychology of memory accuracy. *Annu. Rev. Psychol.* 51, 481–537. <https://doi.org/10.1146/annurev.psych.51.1.481>.
- Kundera, M., 2002. *Ignorance*. Faber & Faber, London.
- Loftus, E., 2003. Our changeable memories: legal and practical implications. *Nat. Neurosci. Rev.* 4, 231–234. <https://doi.org/10.1038/nrn1054>.
- Magnussen, S., Endestad, T., Koriat, A., Helstrup, P., 2007. What do people believe about memory, and how do they talk about memory? In: Magnussen, S., Helstrup, T. (Eds.), *Everyday Memory*. Psychology Press, Hove, pp. 5–25.
- Megreya, A.M., Sanford, A., Burton, A.M., 2013. Matching face images taken on the same day or months apart: the limitations of photo ID. *Appl. Cognit. Psychol.* 27, 700–706. <https://doi.org/10.1002/acp.2965>.
- Misra, P., Marconi, A., Peterson, M., Kreiman, G., 2018. Minimal memory for details in real life events. *Sci. Rep.* 8, 16701. <https://doi.org/10.1038/s41598-018-33792-2>.
- Roediger, H.L., Karpicke, J.D., 2006. Test-enhanced learning – taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>.
- Rowland, C.A., 2014. The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463. <https://doi.org/10.1037/a0037559>.
- Rubin, D.C., 2006. The basic-systems model of episodic memory. *Perspect. Psychol. Sci.* 1, 277–311. <https://doi.org/10.1111/j.1745-6916.2006.00017.x>.
- Schacter, D.L., Wagner, A.D., Buckner, R.L., 2000. Memory systems of 1999. In: Tulving, E., Craik, F.I.M. (Eds.), *The Oxford Handbook of Memory*. Oxford University Press, Oxford, pp. 627–643.
- Schacter, D.L., 1996. *Searching for Memory: The Brain, the Mind, and the Past*. Basic Books, New York.
- Schacter, D.L., 2001. *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin, Boston, MA.

- Schwieren, J., Barenberg, J., Duke, S., 2017. The testing effect in the psychology classroom: a meta-analytic perspective. *Psychol. Learn. Teach.* 16, 179–196. <https://doi.org/10.1177/1475725717695149>.
- Scoboria, A., Wade, K.A., Lindsay, D.S., Azad, T., Strange, D., Ost, J., Hyman, I.E., 2017. A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory* 25, 146–163. <https://doi.org/10.1080/09658211.2016.1260747>.
- Tulving, E., Schacter, D.L., 1990. Priming and human memory systems. *Science* 247, 301–306. <https://doi.org/10.1126/science.2296719>.
- Tulving, E., 2002. Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. <https://doi.org/10.1146/annurev.psych.53.100901.135114>.
- United Nations Development Programme, 2019. Human Development Reports. Available at: <http://hdr.undp.org/en/2018-update>. (Accessed 26 October 2020).
- Valentine, T., Brennen, T., Brédart, S., 1996. *The Cognitive Psychology of Proper Names: The Importance of Being Ernest*. Routledge, London.
- Wagenaar, W., 1986. My memory: a study of autobiographical memory over six years. *Cognit. Psychol.* 18, 225–252. [https://doi.org/10.1016/0010-0285\(86\)90013-7](https://doi.org/10.1016/0010-0285(86)90013-7).
- Weisberg, D.S., Keil, F.C., Goodstein, J., Rawson, E., Gray, J.R., 2008. The seductive allure of neuroscience explanations. *J. Cognit. Neurosci.* 20, 470–477. <https://doi.org/10.1162/jocn.2008.20040>.
- Wells, G., Kovera, M.B., Douglass, A.B., Brewer, N., Meissner, C.A., Wixted, J.T., 2020. Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law Hum. Behav.* 44, 3–36. <https://doi.org/10.1037/lhb0000359>.
- White, R., 2002. Memory for events after twenty years. *Appl. Cognit. Psychol.* 16, 603–612. <https://doi.org/10.1002/acp.819>.
- White, R.T., 2020. Autobiographical memory after 40 years. *Appl. Cognit. Psychol.* 34, 776–781. <https://doi.org/10.1002/acp.366>.
- Yang, C., Luo, L., Vadillo, M.A., Yu, R., Shanks, D.R., 2021. Testing (quizzing) boosts classroom learning: a systematic and meta-analytic review. *Psychol. Bull.* 147, 399–435. <https://doi.org/10.1037/bul0000309>.
- Zogg, J.B., Woods, S.P., Saucedo, J.A., Wiebe, J.S., Simoni, J.M., 2012. The role of prospective memory in medication adherence: a review of an emerging literature. *J. Behav. Med.* 35, 47–62. <https://doi.org/10.1007/s10865-011-9341-9>.

Further reading

- Baddeley, A.D., Eysenck, M.W., Anderson, M.C., 2020. *Memory*, third ed. Routledge, New York.
- Cleary, A.M., Schwartz, B.L., 2020. *Memory Quirks: The Study of Odd Phenomena in Memory*. Routledge.
- Magnussen, S., Helstrup, T. (Eds.), 2007. *Everyday Memory*. Psychology Press, Hove.
- Schwartz, B.L., 2017. *Memory: Foundations and Applications*, third ed. Sage, Thousand Oaks, CA.

Problem solving

Richard E. Mayer, Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Definitions	229
Types of problems	230
Cognitive processes and types of knowledge in problem solving	230
Rigidity in thinking	231
Problem-solving transfer	231
The distinction between productive and reproductive thinking	232
The nature of insight	232
Problem space and search processes	233
Problem solving in realistic situations	233
Teaching of thinking skills	233
Acknowledgment	234
References	234

Problem solving is frequently in the spotlight of educational reform, particularly in light of calls for graduates to be able to be innovative (Pelligrino and Hilton, 2012). In addition to remembering information, students must be able to use what they have learned to solve new problems. A major goal of education is to help students become effective problem solvers, that is, people who can generate useful and original solutions when they are confronted with problems they have never seen before. This entry examines aspects of the large body of research on problem solving (Ball and Thompson, 2019; Holyoak and Morrison, 2012; Mayer, 1992) that are most relevant for education.

Definitions

A *problem* consists of a given state (i.e., a description of the current situation), a goal state (i.e., a description of the desired situation), and a set of operators (i.e., rules for moving from one state to another). A problem occurs when a situation is in one state, the problem solver wants it to be in another state, and there are obstacles preventing a smooth transition from the given state to the goal state. Duncker (1945: 1) defined a problem in the following way: “A problem arises when a living creature has a goal but does not know how this goal is to be reached.” Although Duncker’s definition is still valid, it must be updated to include the possibility of problem solving by machines, and so the term problem solver can refer to both living creatures and machines.

Problem solving occurs when a problem solver engages in cognitive activity aimed at overcoming a problem. Duncker (1945: 1) noted that “when one cannot go from the given situation to the desired situation simply by action, then there has to recourse to thinking” and “such thinking has the task of devising some action which may mediate between the existing and desired situations.” Similarly, Polya (1981: ix) defined problem solving as “finding a way out of a difficulty, a way around an obstacle.” In developing computer simulations of problem solving, Newell and Simon (1972) defined problem solving as a search for a path between the given and goal states of a problem. Mayer (1992) summarized three major elements in a definition of problem solving: (1) problem solving is cognitive because it occurs internally within the problem solver’s cognitive system, (2) problem solving is a process, because it involves manipulating or performing operations on the problem solver’s cognitive representations, and (3) problem solving is directed, because the problem solver is attempting to achieve some goal. In short, problem solving is directed, cognitive processing aimed at finding a way to achieve a goal.

What is the relation between problem solving and other high-level cognitive processes such as thinking and reasoning? Thinking can be broken down into two types—directed and nondirected thinking. Problem solving is a common and pervasive type of thinking, namely, directed thinking in which the thinker engages in cognitive processing aimed at achieving some goal. On the contrary, in nondirected thinking, the thinker engages in cognitive processing that is not aimed at achieving some goal, such as day-dreaming or the abnormal thinking of people with schizophrenia. In general, the terms problem solving and thinking can be used interchangeably, with the recognition that nondirected thinking is excluded.

Reasoning can be viewed as a type of problem solving that involves inference in deductive reasoning tasks and inductive reasoning tasks. In deductive reasoning, the problem solver is given premises and must apply the rules of logic to derive a conclusion. For example, if you know all four-sided polygons are quadrilaterals and all squares are four-sided polygons, you may logically conclude, all squares are quadrilaterals. In inductive reasoning, the problem solver is given a series of instances or events or examples and must infer a rule. For example, after learning the Spanish words *la casa*, *el libro*, *la escuela*, *el perro*, *la muchacha*, and *el muchacho*, you may conclude that the article *la* goes with words ending in *a* and the *el* goes with words ending in *o*, a grammatical rule that is not without exceptions in Spanish.

Creative thinking occurs when a problem solver generates ideas and critical thinking occurs when a problem solver evaluates them. Two important criteria in creative and critical thinking are that the ideas must be original and useful.

Finally, intelligence is sometimes defined as the ability to adapt to new situations or, more simply, as the ability to learn (Hunt, 2010; Sternberg, 2020). Viewed in this way, individual differences in intelligence can be considered as individual differences in the ability to engage in problem solving.

Types of problems

An important distinction—based on the clarity of the problem statement—can be made between well-defined problems and ill-defined problems. A well-defined problem has a clear given state, a clear goal state, and a clear set of allowable operators. For example, finding the value of x in an algebraic equation such as $2X + 5 = 8$ is a well-defined problem because the given state is the equation, the goal state is a value for $X = \underline{\hspace{1cm}}$, and the operators are defined by the rules of algebra and arithmetic. In contrast, an ill-defined problem has a poorly specified given state, goal state, and/or operators. For example, choosing an appropriate education for a career path is an ill-defined problem because the goal and allowable operators are not clearly specified. Most problems encountered in school are well-defined problems, whereas most crucial problems in everyday life are ill-defined.

Another important distinction—based on the knowledge of the problem solver—can be made between routine and nonroutine problems. Routine problems are identical or very similar to problems that the problem solver already knows how to solve, and therefore require what Wertheimer (1959) called *reproductive thinking*—reproducing responses that have been produced previously. For example, a routine problem for most high school students is “ $55 \times 5 = \underline{\hspace{1cm}}$ ”. In the strictest sense, routine problems do not conform to the definition of problem, since they do not include an obstacle between the given and goal states. In contrast, nonroutine problems are different from problems that the problem solver already knows how to solve and therefore require what Wertheimer (1959) called *productive thinking*—creating a novel solution. Examples for most high school students include writing a computer program to compute the mean and standard deviation of a sample, or writing a short essay answer on why Spanish explorers waited several centuries before colonizing California. In school, students often work on routine problems called exercises; however, most important problems in everyday life are nonroutine.

A third distinction can be made between problems requiring convergent and divergent thinking. Convergent thinking problems have a single correct answer that can be determined by applying a procedure or retrieving a fact from memory. Examples include arithmetic computation problems. Divergent thinking problems have many possible answers, and so the problem solver’s job is to create as many solutions as possible (Guilford, 1967). Classic examples include uses problems, such as “List all the possible uses of a brick,” and consequences problems, such as “List all the consequences of humans having six rather than five fingers.” Creativity can be measured in terms of the originality and usefulness of the answers, and divergent thinking skills, which underlie creativity, are taught in courses on creative thinking (Kaufman and Sternberg, 2019; Kaufman et al., 2017). Although divergent thinking is regarded as the hallmark of creativity, solving insight problems (i.e., problems that require devising a novel solution method) has been found to be most strongly related to convergent thinking (Webb et al., 2017).

Cognitive processes and types of knowledge in problem solving

Problem solving can be divided into two major phases: problem representation and problem solution. Problem representation involves building a mental representation of the problem, and includes the cognitive process of representing (i.e., building a situation model, that is, a mental representation of the situation described in the problem). Problem solution involves devising and carrying out a plan for solving the problem, and includes the cognitive processes of planning (i.e., devising a plan), executing (i.e., carrying out the plan), and monitoring (i.e., tracking the effectiveness of the plan).

The cognitive processes involved in problem representation and problem solution may interact, rather than occur in linear order. For example, a student may be given the following word problem: “Sarah has three marbles. David has two more marbles than Sarah. How many marbles does David have?” In representing the problem, the student must translate each sentence into an internal mental representation, such as “Sarah’s marbles = 3” and “David’s marbles = Sarah’s marbles + 2,” and mentally integrate them into a situation model, such as a spatial representation consisting of a bar for Sarah’s marbles (3 units high), a bar on top of it for the difference between Sarah’s and David’s marbles (2 units high), and a bar next to these for David’s marbles (indicating that Sarah’s marbles and the difference set of marbles are subsets of David’s marbles). Planning involves determining the operations to be performed, such as determining that 3 and 2 must be added together. Executing involves carrying out the operation(s), such as computing that 5 is the sum of 3 and 2. Monitoring involves detecting when a plan is not working, a step was not executed correctly, or an answer is questionable. Although school instruction tends to emphasize execution of basic skills, students’ major difficulties are in learning how to represent problems, devise plans, and monitor problem-solving processes.

Several types of knowledge are required for successful problem solving: facts, concepts, procedures, strategies, and beliefs (Anderson et al., 2001; Mayer, 2011). Facts are elements of a factual knowledge about the world, such as, “There are 1000 mL in a liter.” Concepts are principles or models (which are elements of conceptual knowledge) such as, “In the number 567, 6 refers to the number of tens,” or categories or schemas (which are elements of schematic knowledge), such as knowing that “What is the probability of flipping a fair coin three times and getting heads all three times?” is a joint probability problem. Procedures are step-

by-step processes, such as knowledge of the procedure for long division used for 252 divided by 12. Strategies are general methods for problem solving, such as knowing how to break a problem into smaller parts. Meta-strategies are methods for controlling and adjusting the use of all the five kinds of knowledge during a task. Beliefs are thoughts about one's cognitive processing, such as believing, "I am good at solving statistics problems." Facts and concepts are useful in the process of representing; procedures are useful for the process of executing; and strategies and beliefs are useful for planning and monitoring. Although instruction may tend to emphasize facts and procedures, all five kinds of knowledge are needed to support problem solving. This means that problem solving can depend on having learned appropriate knowledge.

Rigidity in thinking

A major obstacle to effective problem solving is rigidity in thinking. For example, in some problem-solving situations, the problem solver must use an object in a new way, such as using a brick as a doorstop or using a pencil as a bookmark. When a problem solver can only conceive of using an object in its most common function, then the problem cannot be solved. Duncker (1945) used the term *functional fixedness* to refer to a situation in which a problem solver cannot think of using an object in a new function that is required to solve the problem. Another example of rigidity occurs when a problem solver uses a well-learned procedure on a problem for which the procedure is inappropriate. For example, if a student solved a long series of arithmetic story problems that all contain the word "more" and require adding the numbers together, the student may incorrectly carry out this same addition procedure for a new problem that actually requires subtracting the numbers from one another. Luchins (1942) used the term *Einstellung* (or problem-solving set) to refer to this phenomenon. A goal of instruction in problem solving is to help students avoid rigid thinking.

Problem-solving transfer

Transfer is the effect of prior learning on new learning. When the new learning task is a problem to solve, we can use the term problem-solving transfer to refer to the effect of prior learning on solving a new problem (Mayer and Wittrock, 2006). Positive transfer occurs when previous learning helps you on a new task, whereas negative transfer occurs when previous learning hurts you on a new task. For example, if you have automated your use of arithmetic procedures, this should make it easier for you to solve an arithmetic word problem—which would indicate positive transfer. If you learned to drive on the right side of the road in the United States, you may experience negative transfer in trying to learn to drive on the left side of the road in Australia.

How does transfer work? This has been a central research question in psychology and education since Thorndike's (1931; Thorndike and Woodworth, 1901) pioneering work in the early 1900s, and has generated three alternative explanations: *general transfer*, *specific transfer*, and *mixed transfer*. General transfer is the idea that learning task *A* can help you on task *B*, even if *A* and *B* have nothing specifically in common. For example, the doctrine of formal discipline (which is a classic theory of general transfer) posited that learning certain school subjects, such as Latin and geometry, would improve students' minds in general and thereby help them on unrelated tasks in the future. The doctrine of formal discipline was used to justify the establishment of Latin schools, in which the curriculum consisted of learning Latin, Greek, geometry, and similar subjects. In one of the first experiments in the field of educational psychology, Thorndike (1931; Thorndike and Woodworth, 1901) was able to show that students who learned Latin did not perform any better in learning bookkeeping than did students who had not learned Latin. Subsequent research (Singley and Anderson, 1989) also found little support for the idea of general transfer. In short, learning to be a good problem solver may require different kinds of training for each academic discipline or subject area.

In contrast, specific transfer is the idea that learning task *A* will help you in task *B* only to the degree that *A* and *B* have identical elements in common. Thus, learning Latin may help you learn Spanish because some of the verb conjugations are similar and many of the words are similar. The theory of specific transfer is problematic for educators because it suggests that students need to learn every specific piece of knowledge they will ever need.

Mixed transfer is a compromise between general and specific transfer that involves specific transfer of a general principle, that is, mixed transfer occurs when a learner abstracts a general principle from learning *A* and is able to apply it to solving a new problem *B*. The transfer is specific because both *A* and *B* can be solved by the same general principle but what is being transferred is a general principle rather than specific behaviors. For example, if students learn to make sense out of text passages by producing summaries, this general method can be transferred to the task of making sense out of a new text passage. Research on teaching of cognitive strategies shows that students can benefit from learning general strategies such as summarizing in reading comprehension (Fiorella and Mayer, 2015). The implication for education is to focus on general principles, concepts, and models that can transfer from one situation that requires the principle to another, such as a mental number line for understanding addition and subtraction or the nature of air pressure for answering questions about how a tire pump works.

Promoting positive transfer is a fundamental goal of education, that is, educators seek to help students learn in ways so that they will be able to use what they have learned to solve new problems. Mayer and Wittrock (2006) have described seven ways to promote problem-solving transfer: load-reducing methods, such as helping students build automaticity in basic skills; structure-based methods, such as using concrete models; schema-based methods, such as using advance organizers or pre-training; generative methods, such as encouraging learners to engage in elaboration or self-explanation; guided discovery methods, such as providing

hints as someone solves a problem; modeling methods, such as providing worked examples; and teaching thinking skills, such as training people to use effective methods and strategies.

The distinction between productive and reproductive thinking

Why is it that some people invent clever solutions when confronted with a problem, whereas others do not? The Gestalt psychologist, [Wertheimer \(1959\)](#) attempted to answer this question by distinguishing between two kinds of thinking: *productive thinking* and *reproductive thinking*. Productive thinking involves producing a novel solution when confronted with a problem, whereas reproductive thinking occurs when problem solvers use solution procedures that they already know as a result of solving previous problems.

For example, [Wertheimer \(1959\)](#) described two ways of learning how to find the area of a parallelogram—*learning by rote* and *learning by understanding*. In learning by rote (or rote learning), the student is taught to measure the height, measure the base, and then multiply height times base. According to Wertheimer, students who learn by rote perform well on retention tests, such as finding the area of similar parallelograms, and poorly on transfer problems, such as finding the area of an unusually shaped parallelogram. In contrast, students who learn by understanding (or by meaningful learning) are encouraged to discover that the parallelogram can be converted into a rectangle by cutting the triangle off one end and moving it to the other end. Students who learn by understanding are expected to perform well on both retention and transfer tests. Thus, rote learning leads to reproductive thinking (as measured by retention tests) whereas meaningful learning leads to productive thinking (as measured by transfer tests). Most educators seek to develop students who can engage in productive thinking, although there is a place for reproductive thinking when students can automate common tasks such as how to make arithmetic computations.

The nature of insight

Insight is the cognitive process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve a problem ([Mayer, 1995](#); [van Steenburgh et al., 2012](#)). Insight plays a crucial role in creative thinking ([Kaufman and Sternberg, 2019](#)), in which a problem solver invents novel solutions to a problem. How does insight work? Gestalt psychologists and others have offered five somewhat interrelated explanations ([Mayer, 1995](#)): insight as completing a schema, insight as sudden visual reorganization, insight as reformulation of a problem, insight as removing mental blocks, and insight as finding a problem analog. Selz, working in the early 1900s in the Netherlands, produced psychology's first explanation of insight as completing a schema ([Frijda and De Groot, 1982](#)). For example, when given a problem such as "What is a co-ordinate of baseball?" a problem solver may say, "Let's see. Baseball is a sport. Another sport is football, so football is the answer." In this case, the problem solver is not following a chain of associations, but rather is trying to build a cognitive structure that has a superset (sport) linked to two subsets that are co-ordinates (baseball and one more), so coming up with an answer amounts to completing a schema. The idea that meaningful learning requires active construction by the learner is the fundamental theme of many current theories of academic learning ([Mayer, 2011, 2014a](#)).

Kohler, also working in the early 1900s, provided evidence that insight is a process of sudden visual reorganization in which the problem solver literally sees how all the parts of the problem fit together ([Kohler, 1925](#)). For example, when an ape was put in an area that had stackable crates on the floor and bananas hanging overhead out of reach, the ape looked around and then in an apparent flash of insight, suddenly stacked the crates to form a sort of ladder leading to the bananas. This approach is consistent with current interest in using computer-assisted visualizations and concrete representations to help people understand how various systems work ([Mayer, 2020](#)).

[Duncker \(1945\)](#) described insight as a reformulation of the problem, particularly a restatement of the givens or the goal in a new way. For example, in the tumor problem, you are asked to free a person of an inoperable stomach tumor by using "rays which destroy organic tissue at sufficient intensity" (p. 1). In order to solve the problem, the problem solver must restate the goal as, "lower the intensity of the rays as they pass through the healthy tissue," which leads to the solution of having many weak rays all converge on the tumor. This approach is consistent with the current idea that the most difficult aspect of problem solving is mentally representing the problem in a productive way.

[Duncker \(1945\)](#) also described insight as a process of removing mental blocks, that is, of being able to use an object in way that is different from its conventional use. For example, in the candle problem, the problem solver is given a box containing candles, a box containing tacks, and a box containing matches, and is asked to mount a lighted candle on a wall. The solution—involving using a box as the base, which is tacked into the wall—is much more difficult if the objects are in the boxes rather than next to them. According to Duncker, presenting the objects in the boxes creates functional fixedness—the tendency to be able to conceive of only one use of an object even though a problem solution requires using an object in a new way. Removing mental blocks is a key focus of current programs aimed at teaching thinking skills ([Behar-Horenstein and Niu, 2011](#); [Wegerif et al., 2015](#)).

Finally, [Wertheimer \(1959\)](#) offered a fifth explanation of insight—finding a problem analog—in which the problem solver abstracts a general principle from one problem and applies it to a new one. Thinking by analogy is still an important theme in cognitive science, and is the basis for more current views of problem-solving transfer ([Holyoak, 2012](#)).

A sixth explanation—insight as nothing new—holds that solving insight problems is no different from solving other problems (Weisberg and Suls, 1973), although Metcalfe and Wiebe (1987) have shown that problem solvers use qualitatively different thinking processes for insight problems and routine problems. More recently, Chuderski and Jastrzebski (2018) reported that solving insight problems was highly correlated to measures of working memory, so that “little above and beyond WM is unique about insight” (p. 257).

Problem space and search processes

Information-processing theories of problem solving focus on constructing a problem space and finding a path through the problem space (Newell and Simon, 1972; Novick and Bassok, 2005). A problem space consists of a representation of the initial state, goal state, and all intervening states. For example, the problem space for solving the equation, $2X - 5 = X$, has this equation as the initial state, and $X = ___$ as the goal state. Two of the intervening states, directly after the initial state are $2X = X + 5$ and $2X - X - 5 = 0$, which were created by applying legal operators such as add 5 to both sides or subtract X from both sides. Similarly, other states are created by applying operators to these states, and so on.

Once a problem is represented as a problem space, the problem solver’s task is to search for a path from the initial state to the goal state. Means-ends analysis is a search strategy in which the problem solver works on one goal at a time; if that goal cannot be achieved directly, the problem solver sets a new goal of removing barriers, and so on. This search strategy is commonly used in computer simulations of problem solving and is consistent with the way that beginners solve problems.

Computer simulation models of human problem solving have been used to design online cognitive tutoring systems in academic disciplines (Anderson et al., 1995; Nye et al., 2014). These tutoring systems are based on a specification of the knowledge, procedures, and strategies needed to support competent problem solving.

Problem solving in realistic situations

Although classic research focused mainly on solving artificial puzzles or formal syllogisms, cognitive science research on problem solving has been shifting toward realistic situations including everyday problem solving, expert problem solving, and problem solving in subject areas (Ericsson et al., 2018; Holyoak and Morrison, 2012).

Research on everyday thinking shows that people rarely use school-taught methods to solve problems encountered outside of school (Lave, 1988; Nunes et al., 1993). For example, to determine the best buy in a supermarket—such as 90 cents for a 10-ounce can of peanuts or 45 cents for a 4-ounce can—the school-taught procedure is to compute the unit cost of each item (i.e., 9 cents vs. 11.25 cents, respectively). However, Lave (1988) found that people almost never used the school-taught procedure; instead, they invented arithmetic procedures suited to the situation, such as the ratio strategy in which the problem solver notes that the larger one is a better buy because it costs twice as much and gives you more than twice as many ounces.

Research on expert problem solving compares differences in how novices and experts solve problems in domains such as physics, medical diagnosis, computer programming, and the game of chess. For example, when Larkin (1983) asked experts and novices to think aloud as they solved physics problems, she found that experts were more likely to focus on underlying physics concepts (such as forces and weights) whereas novices focused on surface characteristics (such as pulleys and ropes). Similarly, when Chi et al. (1981) asked experts and novices to sort physics problems into groups, experts sorted problems based on their underlying physics principle (such as conservation of energy), whereas novices sorted the problems based on their surface characteristics (such as inclined planes and springs). Results of expert–novice studies suggest that experts represent and solve problems differently from novices, and so instruction can focus on helping novices think more like experts.

Another example of problem solving in realistic situations involves psychologies of subject matter, that is, research on problem solving in subject areas, such as reading, writing, mathematics, science, and history (Mayer, 2008). Instead of asking, how do people think in general, researchers ask, how do people think about testing a scientific theory, solving a mathematics word problem, or explaining why a historical event happened, or how do people think as they create an essay or make sense out of a printed passage. This approach suggests that instruction in subject matter areas should focus on helping students learn the cognitive processes and strategies required for successful problem solving.

Teaching of thinking skills

Can people improve their problem solving or thinking skills? Mayer (2008) summarized four evidence-based principles for the design: (1) Concerning what to teach, thinking is based on a collection of smaller component skills rather than improving the mind as a single intellectual ability. (2) Concerning how to teach, focus on the process of thinking through modeling rather than only on the product of thinking (i.e., getting the right answer). (3) Concerning where to teach, embed thinking skill instruction in domain specific subject areas rather than as a general all-purpose course. (4) Concerning when to teach, it is possible to teach higher order thinking skills before lower-level skills are completely automated.

One recent attempt to teach thinking skills involves computer-based games and simulations. In spite of strong claims for brain-training games, rigorous research has not provided convincing evidence to support the claims (Bainbridge and Mayer, 2018; Mayer, 2014b). However, games that are designed to teach to a specific thinking skill based on principles of skill learning have been successful (Anguera et al., 2013; Parong et al., 2017).

Acknowledgment

Preparation of this entry was supported by Grant N00014612046 from the Office of Naval Research.

References

- Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R., 1995. Cognitive tutors: lessons learned. *J. Learn. Sci.* 4, 167–207.
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., et al., 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, New York.
- Anguera, J.A., Boccanfuso, J., Rintoul, J.L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolfe, C., Johnston, E., Gazzaley, A., 2013. Video game training enhances cognitive control in older adults. *Nature* 501 (7465), 97–101.
- Bainbridge, K., Mayer, R.E., 2018. Shining the light of research on Lumosity. *J. Cogn. Enhanc.* 2, 43–62.
- Ball, L.J., Thompson, V.A. (Eds.), 2019. *International Handbook of Thinking and Reasoning*. Routledge, New York.
- Behar-Horenstein, L.S., Niu, L., 2011. Teaching critical thinking skills in higher education: a review of the literature. *J. Coll. Teach. Learn.* 8 (2).
- Chi, M.T.H., Feltovich, P.J., Glaser, R., 1981. Categorization and representation of physics problems by experts and novices. *Cogn. Sci.* 5, 121–152.
- Chuderski, A., Jastrzebski, J., 2018. Much ado about aga! Insight problem solving is strongly related to working memory capacity and reasoning ability. *J. Exp. Psychol. Gen.* 147, 257–281.
- Duncker, K., 1945. On problem solving. *Psychol. Monogr.* 58 (270), 3.
- Ericsson, K.A., Hoffman, R.R., Kozbelt, A. (Eds.), 2018. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, New York.
- Fiorella, L., Mayer, R.E., 2015. *Learning as a Generative Activity*. Cambridge University Press, New York.
- Frijda, N.H., de Groot, A.D., 1982. Otto Selz: His Contribution to Psychology. Mouton, The Hague.
- Guilford, J.P., 1967. *The Nature of Human Intelligence*. McGraw-Hill, New York.
- Holyoak, K.J., Morrison, R.G. (Eds.), 2012. *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, New York.
- Holyoak, K.J., 2012. Analogy and relational reasoning. In: Holyoak, K.J., Morrison, R.G. (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, New York, pp. 234–259.
- Hunt, E., 2010. *Human Intelligence*. Cambridge University Press.
- Kaufman, J.C., Sternberg, R.J. (Eds.), 2019. *The Cambridge Handbook of Creativity*, second ed. Cambridge University Press, New York.
- Kaufman, J.C., Glaveanu, V.P., Baer, J. (Eds.), 2017. *The Cambridge Handbook of Creativity Across Domains*. Cambridge University Press, New York.
- Kohler, W., 1925. *The Mentality of Apes*. Liveright, New York.
- Larkin, J.H., 1983. The role of problem representation in physics. In: Gentner, D., Stevens, A.L. (Eds.), *Mental Models*. Erlbaum, Mahwah, NJ, pp. 75–98.
- Lave, J., 1988. *Cognition in Practice*. Cambridge University Press, Cambridge, UK.
- Luchins, A., 1942. Mechanization in problem solving. *Psychol. Monogr.* 54 (248), 6.
- Mayer, R.E., Wittrock, M.C., 2006. Problem solving. In: Alexander, P.A., Winne, P.H. (Eds.), *Handbook of Educational Psychology*, second ed. Erlbaum, Mahwah, NJ, pp. 287–304.
- Mayer, R.E., 1992. *Thinking, Problem Solving, Cognition*, second ed. Freeman, New York.
- Mayer, R.E., 1995. The search for insight: grappling with Gestalt psychology's unanswered questions. In: Sternberg, R.J., Davidson, J.E. (Eds.), *The Nature of Insight*. MIT Press, Cambridge, MA, pp. 3–32.
- Mayer, R.E., 2008. *Learning and Instruction*, second ed. Pearson Merrill Prentice Hall, Upper Saddle River, NJ.
- Mayer, R.E., 2011. *Applying the Science of Learning*. Pearson Merrill Prentice Hall, Upper Saddle River, NJ.
- Mayer, R.E. (Ed.), 2014a. *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, New York.
- Mayer, R.E., 2014b. *Computer Games for Learning*. MIT Press, Cambridge, MA.
- Mayer, R.E., 2020. *Multimedia Learning*, third ed. Cambridge University Press, New York.
- Metcalfe, J., Wiebe, D., 1987. Intuition in insight and non-insight problem solving. *Mem. Cogn.* 15, 238–246.
- Newell, A., Simon, H.A., 1972. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.
- Novick, L.R., Bassok, M., 2005. Problem solving. In: Holyoak, K.J., Morrison, R.G. (Eds.), *The Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, New York, pp. 321–350.
- Nunes, T., Schliemann, A.D., Carraher, D.W., 1993. *Street Mathematics and School Mathematics*. Cambridge University Press, Cambridge, UK.
- Nye, B.D., Graesser, A.C., Hu, X., 2014. Multimedia learning with intelligent tutoring systems. In: *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, New York, pp. 705–728.
- Parong, J., Mayer, R.E., Fiorella, L., MacNamara, A., Plass, J., Homer, B., 2017. Learning executive function skills by playing focused video games. *Contemp. Educ. Psychol.* 51, 141–151.
- Pellegrino, J.W., Hilton, M.L., 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. National Academies Press, Washington, DC.
- Polya, G., 1981. *Mathematical Discovery*. Wiley, New York.
- Singley, M.K., Anderson, J.R., 1989. *The Transfer of Cognitive Skill*. Harvard University Press, Cambridge, MA.
- Sternberg, R.J. (Ed.), 2020. *The Cambridge Handbook of Intelligence*. Cambridge University Press, New York.
- Thorndike, E.L., Woodworth, R.S., 1901. The influence of improvement in one mental function upon the efficiency of other functions. *Psychol. Rev.* 8, 247–261.
- Thorndike, E.L., 1931. *Human Learning*. Century, New York.
- van Steenburgh, J.J., Fleck, J.I., Beeman, M., Kouniois, J., 2012. Insight. In: Holyoak, K.J., Morrison, R.G. (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, New York, pp. 475–491.
- Webb, M.E., Little, D.R., Cropper, S.J., Roze, K., 2017. The contributions of convergent thinking, divergent thinking, and schizotypy to solving insight and non-insight problems. *Think. Reas.* 23, 235–258.
- Wegerif, R., Li, L., Kaufman, J.C. (Eds.), 2015. *The Routledge International Handbook of Research on Teaching Thinking*. Routledge, New York.
- Weisberg, R.W., Suls, J., 1973. An information processing model of Duncker's candle problem. *Cogn. Psychol.* 4, 255–276.
- Wertheimer, M., 1959. *Productive Thinking*. Harper and Collins, New York.

Reasoning and decision making

Henry Markovits, Department of Psychology, Université du Québec à Montréal, Montréal, QC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Development of logical capacities	235
Content and development	236
Reasoning development as a form of divergent thinking	237
Understanding individual differences in reasoning	238
Dual process models	238
Dual strategy model	239
Conclusion	241
References	241

Reasoning and decision making form the conscious corner stone of the way that people decide what to do in difficult circumstances. There are some commonly accepted logical norms for these sorts of decisions. However, many decades of research have clearly shown that many (if not most) people do not follow these norms in any systematic way (Evans, 2008; Tversky and Kahneman, 2004). Certainly, one important question is understanding what factors determine people's tendencies to make decisions that are not normative. But, before examining these, there is another underlying question that must be addressed. The clear evidence that people show a great deal of variability in most forms of reasoning and decision-making does not preclude the possibility that there is some underlying capacity for normative, logical reasoning. In fact, there is evidence of a clear developmental progression in people's logical abilities between early childhood and adulthood (Markovits, 2017a). Thus, understanding how people arrive at conclusions and how they make decisions requires examining both aspects. In the following, I will concentrate on two major themes (1) how logical capacities develop and (2) how to characterize individual differences in reasoning and decision making.

Development of logical capacities

In the following I will concentrate on research examining developmental patterns in conditional (if-then) reasoning. This form of reasoning is critical to both logic and a variety of scientific and mathematical domains. It is accordingly the most-studied kind of reasoning and a key to understanding developmental patterns. Conditional statements take the general form "If antecedent, then consequent." Understanding conditional logic must be done by examining the four basic inferences that can be made based on assuming the truth of a single condition major premise combined with a minor premise that is the affirmation or the negation of the antecedent and consequent terms. The Modus ponens inference (which we will refer to as MP) takes the form: "If P then Q. P is true." and leads to the logically necessary conclusion that "Q is true." The Modus tollens inference (which we will refer to as MT) takes the form: "If P then Q. Q is false." and leads to the logically necessary conclusion that "P is false." The two remaining conditional inferences do not lead to a single necessary conclusion but are indeterminate, since they do not allow for a single necessary conclusion. The first of these is the Affirmation of the consequent inference (referred to as AC), which takes the form: "If P then Q. Q is true." and is compatible with either the invited conclusion that "P is true" or its negation "P is false," and thus has no single necessary conclusion. Similarly, the Denial of the antecedent inference (referred to as DA), which takes the form: "If P then Q. P is false." and is compatible with either the invited conclusion that "Q is false" or its negation "Q is true" and thus has no single necessary conclusion.

Empirical studies have given a very wide variety of results about the way that children and adolescents reason and correspondingly different interpretations of the development of reasoning. Some researchers have concluded both that logical reasoning is fairly primitive and well within the capacities of even young children (Hawkins et al., 1984; Romain et al., 1983) and others, that such reasoning is beyond the abilities of even well-educated adults (Evans et al., 2005; Oaksford and Chater, 2009). Despite what appears to be irreconcilable differences in both results and interpretation, there is a clear developmental pattern that can be discerned if one examines the way that reasoning is affected by content. In fact, one of the clearest factors that have been shown to modify reasoning is related to the specific content of the premises. The most important of these is related to what have been called "alternative antecedents." In the context of a given major premise of the form "If P then Q," an alternative antecedent is any case A, such that "If A then Q" is also true. For example, alternative antecedents for the following premise "If a rock is thrown at a window, the window will break" are the category of things that can break windows (hard objects, storms, etc.). Several studies have shown that responses to the AC and DA inferences with premises that refer to concrete and familiar content are strongly affected by the extent to which consequent terms allow spontaneous retrieval of alternative antecedents from memory (Janveau-Brennan and Markovits, 1999; Cummins et al., 1991; Cummins, 1995; Thompson, 1994). This has been measured by an alternatives generation task (Cummins et al., 1991). This simply requires asking participants to generate as many potential alternatives as possible to a given

conditional premise in a given period of time. Most studies have used a dichotomous categorization of the results of alternatives generation tasks. For example, when given a premise such as “If a rock is thrown at a window, the window will break” people will generally be able to produce a great many alternative antecedents (e.g., throwing a chair, storm, etc. are all ways that could lead to broken windows). By contrast, premises such as “If a finger is cut, it will bleed” produce relatively few alternative antecedents, since there are few alternative ways of making fingers bleed. One of the most robust empirical phenomena in conditional reasoning is the observation that when people are given major premises for which they can retrieve relatively more alternative antecedents, they will respond more frequently with the logically correct answer of uncertainty to the AC and DA inferences, while performance on these same inferences is much lower with major premises for which relatively few alternative antecedents can be produced (Cummins et al., 1991; Cummins, 1995; Markovits and Vachon, 1990). To illustrate the size of this effect, Janveau-Brennan and Markovits (1999) found that Grade 5 children (10 years of age) produced an average of 82% correct responses on the AC inference with many alternatives premises while producing only 55% correct responses on the same inference with premises for which there were relatively few alternatives. In fact, the difference between these two kinds of premises corresponds to a 1 or 2 year developmental difference. This same factor has been shown to influence reasoning in the same way among very young children (Markovits, 2000; Janveau-Brennan and Markovits, 1999), adolescents (Markovits and Vachon, 1990), and adults (Cummins et al., 1991; Cummins, 1995; Thompson, 1994).

What makes this factor important to understanding developmental patterns is that content-related differences are large enough to mask developmental effects when only the nature of inferences are examined. For example, one of the more important differences observed when only performance on specific logical forms are examined is that between the valid (MP, MT) and invalid (AC, DA) forms. Generally, performance on the latter forms of reasoning leads to lower rates of correct responding even in educated adults (Markovits, 1985; Cummins et al., 1991). If only the form of an inference were important in understanding development, then it would certainly be the case that young children should find it very difficult to respond correctly to the AC and DA forms. However, with the right content, levels of correct performance on these two logical forms can actually be greater than those typically observed with adults (Markovits and Thompson, 2008; Markovits, 2000), and greater than sometimes observed with the MP inference (which is generally accepted as the easiest of the four inferences). Thus, in order to understand how logical reasoning develops it is vital to understand the origins of the effect of content on reasoning.

Content and development

Understanding how content can influence conditional reasoning is thus a key component of the interactions between development, reasoning and content. There are different possible approaches to this problem. One such approach considers that context affects the way that conditional statements are interpreted. This acknowledges the fact that conditionals are pragmatically contextualized in a way that can alter underlying implicit inferences, and thus modify the underlying logic. There is evidence that conditional promises, of the general form: If you do X, then you will get Y are not intended as standard conditionals, but are meant to be interpreted as biconditional statements of the general form: You will get Y if and only if you do X (Markovits and Lesage, 1990; Fillenbaum, 1975; Rumain et al., 1983). In these cases, there is a strong tendency to accept the invited inference to all four inferences, even when the conditional is expressed as an if-then statement. For example, given a conditional promise such as: “If you mow the lawn, I will give you \$5,” people will conclude that if you received \$5, then you must have mowed the lawn. There is also evidence that statements expressing permissions (If you want to drink, then you must be older than 21) or obligations (If you are over 18, then you have to register for the draft) generate different forms of pragmatic structures (Cheng and Holyoak, 1985; Cheng et al., 1986; Holyoak and Cheng, 1995). However, while pragmatic factors certainly have an impact on the way that conditionals are interpreted, they do not show any developmental variability (Light et al., 1989) and cannot explain the observed variability within a given kind of conditional, when pragmatic factors are not in play.

Understanding how content affects reasoning irrespective of pragmatics requires understanding the interplay between reasoning and memory (Markovits, 2014a). One theory that makes such an explicit claim is a variant of mental model theory (Johnson-Laird, 2001, 2012; Johnson-Laird and Byrne, 1991) proposed by Markovits and Barrouillet (2002). Mental model theory supposes that when people reason with a conditional of the form If P then Q, they start by creating a very simple, initial, representation of the major premise:

P Q

Under certain circumstances, this initial model will be fleshed out by the addition of other relations corresponding to possibly true combinations of the P and Q terms. Thus, the representation corresponding to the full conditional is:

P Q

Not-P not-Q

Not-P Q

Potential conclusions are evaluated by scanning models for possible counterexamples. Thus, when evaluating the AC inference (Q is true), people with the initial model will conclude that if “Q is true” then P must be true. People with the complete model will conclude that there is no valid conclusion since “Q is true” is compatible both with P is true and not-P. Critically, the fleshing out process is essentially semantic, since it uses an abstract representation of the conditional to determine which combinations are possible. By contrast, Markovits and Barrouillet (2002) proposed that while the basic process is semantic, fleshing out is done by associative access to semantic memory, at least in the case of conditionals with concrete content.

The critical component here is the way that potential counterexamples, which in the case of the AC and DA inferences are cases of “not-P and Q” are integrated or not into the mental model used for reasoning. Available evidence clearly suggests that, for premises with familiar content, such counterexamples must be activated and incorporated from semantic memory (see [Markovits, 2014a](#) for a review of this evidence). Importantly, these must be generated online during the reasoning process. This is relatively difficult since it must be done while working memory capacity is occupied by the retention and processing of the premises. Thus, one key component that determines the relative difficulty of generating counterexamples is the relative difficulty of the search process that is required.

More specifically, there are two types of premises that require different levels of access to memory, and that determine an important developmental difference. Category-based premises are major premises that refer to a category and some properties of this category, for example: “If an animal is a dog, then it has four legs.” Counterexamples for this form of premise can be generated by searching memory for examples of other category types that share the same property (e.g., cats have four legs). The second premise type refers to causal conditionals, that is, conditionals of the form “If cause P then effect Q.” Generating counterexamples for these kinds of premises requires a more complex process, since the reasoner must be able to generate the class of causes that can lead to the effect. This kind of information is not stored directly in memory, but must be constructed on-line, which is a form of ad hoc category ([Barsalou, 1983](#)). For example, for the premise “If a rock is thrown at a window, the window will break” generating counterexamples requires generating the class of “ways to break windows,” which can give a wide variety of very different potential causes (e.g., storm, tree branch, etc.). This in turn suggests that reasoning with category-based conditionals should be less difficult than reasoning with causal conditionals, for exactly the same basic inferences ([Markovits, 2017b](#)). In addition, the distinction between category-based and causal conditionals corresponds to an important developmental period. If reasoning “logically” is defined by the ability to distinguish between the necessary MP inference and the uncertain AC and DA inferences, then even very young children 6–7-years of age can reason “logically” with category-based premises ([Markovits and Thompson, 2008](#); [Markovits, 2000](#)). However, it is not until early adolescence that children can reliably distinguish between the certain and uncertain inferences with causal conditionals ([Janveau-Brennan and Markovits, 1999](#)). Thus, the transition between the early ability to reason well with category-based premises and the ability to do so with causal conditionals determines children’s developmental trajectories during the period from initial schooling through pre-adolescence. The key factor in this period is the increasing ability of children to retain premises in working memory and the efficiency of their retrieval processes.

However, while reasoning with concrete content is certainly a major component of how people reason, another important dimension is the ability to reason with more abstract content, which mirrors the kind of reasoning found in mathematics and science. This concerns reasoning well with premises that do not have any evident real-world knowledge available in long term memory, of which the most obvious is reasoning with meaningless content, such as “If X then Y.” Such abstract forms of reasoning are very difficult, and even college students have great difficulty with these kinds of conditionals ([Venet and Markovits, 2001](#); [Markovits and Vachon, 1990](#)). One reason for this is that the counterexample generation process cannot rely on any access to memory but requires a very indeterminate form of counterexample (maybe Z can lead to Y). However, an important question is how the transition between reasoning with concrete premises to reasoning with abstract premises occurs. There is evidence that this transition is modulated by reasoning with contrary-to-fact premises, which refers to premises for which the relation in the major premise is empirically false (such as: “If a shirt is put into mud, it will become clean”). With such premises, generating a counterexample allows using memory-based information but in a complex manner, so that the kinds of counterexamples generated will correspond to an imaginary world in which mud makes things clean—leading to counterexamples of the form “dirt also makes things clean,” etc.

The ability to reason well with contrary-to-fact premises starts to be seen at the beginning of adolescence ([Markovits and Vachon, 1989](#)), well before people can reason well with fully abstract premises ([Markovits and Vachon, 1990](#)). The transitional nature of such reasoning is indicated by clear evidence that contrary-to-fact reasoning has a direct effect on abstract reasoning ([Markovits and Lortie-Forgues, 2011](#)), much more so than reasoning with familiar premises. Thus, the transition between the beginning of the ability to reason well with these kinds of premises to the ability to reason well with abstract premises represents a second developmental level, although one which is less uniformly attained than the previous one (see [Markovits, 2018](#) for a more complete synthesis).

Reasoning development as a form of divergent thinking

The model of reasoning that has been presented places great emphasis on the ability to actively generate counterexamples to given premises. During the initial period of reasoning with familiar concrete premises, this requires fairly direct access to information stored in memory. There are however clear individual differences in the way that people can access such information. One of the distinctions that is important in this context is between information which is activated directly from a given cue and information which is less directly related ([Rosen and Engle, 1996](#)). When reasoning with a given if-then premise, potential antecedents that are more directly associated with the given consequent term are more easily accessed and require relatively little cognitive effort. However, when the antecedent term is one which is highly associated to the consequent, generating alternatives becomes more difficult. For example, when people are asked to generate potential causes for the consequent “a dog scratches constantly,” they tend to produce the possibility that “the dog has fleas.” When asked to reason with the premise “If a dog has fleas, then it will scratch constantly” people tend to have greater difficulty in accepting the uncertainty of the AC inference, compared to another premise with the same consequent but with a different antecedent, such as “If a dog is ill, then it will scratch constantly” ([Quinn and](#)

Markovits, 1998). This is due to the fact that retrieving an alternative antecedent that is highly associated to the consequent requires less cognitive effort, compared to retrieving an alternative antecedent which is less highly associated. Surprisingly, this analysis is very similar to the way that divergent thinking, which is a component of creativity is conceptualized. Without going into too much detail, divergent thinking refers to the capacity to produce ideas that are conceptually distant from the prevailing context (Cropley, 2006; Guilford, 1967). For example, one of the tasks that is used to evaluate divergent thinking is the Instances task, which asks children to generate as many different and original things that can make noise as possible (Wallach and Kogan, 1965). Such tasks are very similar to the alternatives generation task that is used to categorize if-then premises as having many or few potential alternatives (Cummins et al., 1991). Individual differences in this latter task have been shown to be related to the ability to generate logical responses to causal conditionals in preadolescents (Janveau-Brennan and Markovits, 1999). In the same vein, when people are asked to rapidly produce potential causes of effects where a single cause is very strongly related to the effect (e.g., a dog scratches constantly is strongly related to having fleas), the time required to produce the second cause is also related to logical reasoning (Markovits and Quinn, 2002). These results suggest that the alternatives generation process that is assumed to underlie conditional reasoning is a form of divergent thinking, if it is assumed that the major premise generates a semantic context within which generating potential alternatives must occur. Thus, the ability to go beyond the suggested relation between the antecedent and consequent to produce other alternatives is very similar to the way that divergent thinking is conceptualized. Consistent with this idea are results which suggest that simply giving an alternatives generation task before reasoning increases the extent to which logical responses are given (Markovits and Brunet, 2012; Markovits, 2014b). More recently, studies examining very young children's ability to reason logically by accepting the MP inference and rejecting the AC inference for the same major premise have provided direct evidence for the role of divergent thinking in the early stages of conditional reasoning (de Chantal et al., 2019b; de Chantal and Markovits, 2017).

The development of logical reasoning capacities has often been conceptualized as an increase in the understanding of the basic rules of logic. The perspective that the semantic retrieval model instead suggests is that this developmental sequence is better understood as a consequence both of this understanding and of the increasing ability to generate out-of-the-box ideas that allow distinguishing necessary (e.g., MP) inferences from uncertain ones (e.g., AC). This basic role of a form of divergent thinking would also underlie the ability to reason logically with increasingly abstract content, which determines a lengthy developmental progression.

Understanding individual differences in reasoning

Overall, what the preceding model suggests is that people do develop an increasingly sophisticated ability to reason "logically." This in turn would imply that this process would reach an end-point at which everyone should be able to reason more or less logically. Unfortunately, there is a great deal of evidence that this is not the case. There are a variety of sources of the resulting variability that are generally encompassed by what are referred to as dual process models.

Dual process models

One of the problems that bedevils attempts to reason uniformly logically is the existence of several forms of heuristic factors that have a clear effect on the inferences that people actually make. One of the most studied of these is the belief-bias effect (Evans et al., 1983; Markovits and Nantel, 1989; Oakhill et al., 1989). When given inferences of the same logical form, but with putative conclusions that vary in the level of belief that people have that the conclusion is empirically true, there is a clear tendency to confound logical validity with belief in the conclusion. For example, people will tend to judge the following AC inference as valid: "All flowers have petals. Roses have petals. Therefore, roses are flowers." By contrast, they will tend to judge the inference: "All cars have motors. Airplanes have motors. Therefore, airplanes are cars" as invalid, despite the fact that both have the identical logical structure. Thus, belief in the conclusion clearly interferes with people's understanding of the logic behind these inferences (although the extent of this interference is an open question). There are other forms of what can be considered to be logical short-cuts, or heuristics, that have been documented, most extensively by the heuristics and biases program of Kahneman and Tversky (Kahneman, 2011; Tversky and Kahneman, 2004). For example, base-rate neglect leads people to misjudge the actual diagnostic value of a test. Imagine a test that can detect whether a person really has a given disease 80% of the time, and misdiagnoses 20% of cases. However, the disease is only found among 5% of the population (base-rate). People tend to think that if a person receives a positive diagnosis, the chances that they have the disease is 80%. But in actual fact, if one considers that the test will misdiagnose 20% of the 95% of people who do not have the disease, the actual probability that a positive test is associated with having the disease is much lower than 80% (it is closer to 10%). Clearly, these heuristics influence how people think. Generally, they lead to inferences that are logically inaccurate, although it has been claimed that in some cases heuristics are useful short-cuts when reasoning is complex (Gigerenzer and Selten, 2001).

A further question is how to conceptualize the interactions between heuristics and logical reasoning. Although there are some exceptions, the most common way of doing this are what are known as dual process theories (Sloman, 1996; Evans and Stanovich, 2013; Stanovich and West, 2000; Evans, 2008; Pacini and Epstein, 1999). These suppose the existence of two competing systems of reasoning which have been given a variety of names. System 1 uses rapid associatively based inferences and relies heavily on people's stored beliefs and short-cuts. This system operates at high speed and allows generating conclusions that are very rapid

and that reflect the nature of information directly stored in memory. System 2 by contrast uses more conscious forms of processing which require cognitive resources in order to generate more “analytic” conclusions, which in many cases can be “logical.” Although once again there are a variety of models, the general idea is that people’s inferences are the product of an interaction between the two systems. System 1 responses are generated very rapidly and have temporal precedence. System 2 processes must intervene before System 1 response is produced, by recognizing that System 1 response is not adequate, inhibiting this response so that it is not actually produced and generating an alternative response using more complex processes.

These models see reasoning as a sort of internal debate between two very different ways of understanding problems. There is an overall tendency to consider that the intuitive, System 1 response is often faulty, while the more analytic, System 2 response is more logical. However, this analysis is incomplete. For example, there is evidence that at least among more expert reasoners, intuitive responses can be “logical” (de Neys, 2012; de Neys and Glumicic, 2008; Thompson et al., 2018). Second, there are some good reasons for the role that intuitive, fast and heuristic reasoning plays. This kind of reasoning has the advantage of using a person’s knowledge and experience in order to generate the kinds of very rapid inferences required for active, fluid participation in a complex social and physical environment. It could in fact be argued that without these kinds of inferences people would have great difficulty in interacting, since every decision would require the kind of intensive cognitive effort that characterizes System 2 functioning. There is indeed evidence that much of human activity is underpinned by unconscious, intuitive processing (Bargh and Morsella, 2008). Thus, it is important to understand the important role that intuitive processes play, and not simply to consider them to be a form of faulty reasoning. System 2 processes require increased use of cognitive resources, and do not appear to be deployed as regularly. However, the ability to more fully understand the difference between suggested and logical judgments when the two are in conflict does mostly reside in use of System 2 processes with this basic model.

One of the problems with the dual process framework is how to understand the interaction between the two systems. Typically, these models suppose that an initial, rapid and intuitive response is produced, which must then be inhibited by the more working memory intensive form of reasoning used by System 2 processes. However, this supposes that even when the initial response is actually the one that is overtly produced, it must in many cases have been evaluated in comparison to the analytic response. This in turn implies that even when the “non-logical” response is produced, there should be some trace of the alternative. There is indeed evidence that people do have some intuition that when a System 1 response is produced, that this is not necessarily the only option (de Neys, 2012). Thus, there is some indication that people conserve the intuition of what the “logical” response is, which is consistent with results indicating some level of logical intuitions.

One of the important questions within the context of dual process theories concerns the determinants of individual differences in the relative use of the two systems. The processes required to undo an initial System 1 response require cognitive capacity, and it is generally supposed that the ability to do so is related to individual differences in cognitive capacity, such as IQ and working memory, along with other dispositional factors (Stanovich and Toplak, 2012; Stanovich and West, 1997, 2000; Toplak et al., 2011; Quayle and Ball, 2000; Pennycook et al., 2012). Thus, the chances that someone will produce the “logical” response increase as a function of measures of cognitive capacity and motivation. Of course, this assumes that slower, System 2 responses are generally more logical. This is however an assumption that has really not been tested. There is even some evidence that at least in some contexts slower reasoning can be less logical (Newman et al., 2017).

Nonetheless, the distinction between more intuitive, associative processes and more consciously analytic and working memory intensive processes seems to be well-accepted. This model has been applied to a variety of reasoning and judgment paradigms, and has been used to explain the difficulty that people have in making logical judgments. Interestingly, within a developmental context, there is some evidence that older children and adults are sometimes less “logical” than younger ones due to lesser impact of some heuristics (Artman et al., 2006; Noveck, 2001). Some of the sources of the heuristics that are used by intuitive responding come from the increasing capacity to understand the usual pragmatic forms of inference that characterize speech acts. For example, if-then premises are often intended as effective biconditionals even when they are couched in the same terms as typical if-the statements. When a parent says to a child “If you finish your meal, you’ll get dessert” the parent implicitly means that dessert will come if, and only if, the meal is finished (Fillenbaum, 1975). The tendency to interpret statements within their pragmatic context clearly is learned and appropriately increases with age, since this is necessary for social functioning. However, this illustrates one of the important paradoxes of development and reasoning. On the one hand, there is a clear increase in the ability of children to reason “logically,” that is without the constraints of assumptions due to empirical understanding. However, coupled with this is the increasing ability of children to understand the implicit underpinnings of social exchange, which increases their tendency to use these implicit rules in order to make inferences. One way of conceptualizing this developmental pattern is that the influence of System 1 intuitive inferences increases with experience, which allows people to function in increasingly fluid ways that reflect the patterns of interactions found in their individual environments. Interrupting this form of processing in order to engage System 2 processing, which is generally required to overcome intuitive responses becomes increasingly difficult and engages more of people’s cognitive resources. Thus, while there is a clear increase in reasoning ability, this is not sufficient to overcome the increasing importance of intuitive heuristics, which would account for the enduring impact of the latter on the thinking processes of even well-educated adults.

Dual strategy model

Dual process models assume that both intuitive and analytic processes co-exist in all reasoners, and that it is the interaction and potential conflicts between them that determine the extent to which a person is “logical.” However, these models also acknowledge

the presence of individual differences that might well affect how logical people are at some overall level, by modulating the ability or the motivation of the cognitive system to override an intuitive response. However, any discussion of such differences is complicated by the fact that the nature of the basic processes postulated by dual process theories remains undetermined at a very fundamental level. This is particularly the case for analytic processes which are only defined by some general characteristics (e.g., use of working memory). This is a critical problem, since the nature of analytic processes must clearly impact how this system works, and could clearly alter the general framework of dual process theories. In fact, there are at least two current and very different models of the nature of the inferential processes that all people might deploy. Mental model theories (Johnson-Laird, 2012; Johnson-Laird and Byrne, 1991) propose that when people make inferences, they first construct an internal representation of important problem parameters. This representation is used to suggest a potential conclusion and can subsequently be scanned for possible counterexamples to this conclusion. Probabilistic theories suggest that people make inferences by accessing available information about the problem context and constructing an estimation of the likelihood of a possible conclusion (Evans et al., 2007; Oaksford and Chater, 2009). Both kinds of theory present models of analytic reasoning, but although they do often make similar predictions, they suggest very different underlying processes, and are meant to represent competing models of reasoning. However, although these theories have been suggested as unique models for the way that people reason, there is some clear similarity to the analytic/intuitive dimension suggested by dual process models, with mental model theories presenting a working memory intensive process, while probabilistic theories suggest processes that are more associative and rely on rapid access to existing knowledge, something that is similar to the description of intuitive processing.

There is evidence that, as opposed to dual process accounts, these two models may represent two different modes of processing information. The recent dual strategy model of reasoning explicitly proposes that people will preferentially use one of two different strategies that roughly correspond to these two models of reasoning. This model was based on an initial proposal by Verschueren et al. (2005) as a version of dual process theory. However, Markovits et al. (2012) developed a diagnostic instrument in the context of if-then reasoning that allowed distinguishing the two different reasoning strategies. This relies on giving people access to frequency information about events related to a given inference of the form: If P then Q, Q is true. Is “P is true” a logically valid conclusion? The key version of this presents 1000 observations, of which 950 find that P is true and Q is true, and 50 find that P is false and Q is true. This suggests that if “Q is true,” the probability that “P is true” is very high, but also provides counterexamples for the same conclusion. Using this form of diagnostic allows distinguishing two forms of reasoning strategy. *Statistical* reasoners will use the probabilistic pattern to conclude that “P is true” is a valid conclusion, while *Counterexample* reasoners will use the same information to conclude that this is not a valid conclusion (although about 30% of subjects fall into an Other category, which represents people who switch strategies). Subsequent studies examined the relationship between this diagnostic and other forms of reasoning. These have found a correlation between reasoning strategy and subsequent performance on other forms of reasoning (Brisson and Markovits, 2020; Markovits et al., 2013, 2020). Although counterexample reasoners generally have been found to be more “logical” than statistical reasoners, this is not always the case (Markovits et al., 2016). More specifically, studies examining the belief-bias effect have found that strategy use is an important predictor of “logical” reasoning when there is a conflict between logic and belief (de Chantal et al., 2019a; Markovits et al., 2017). These have shown that people who are categorized as counterexample reasoners are more able to produce the valid response on these kinds of problem. This difference is even more pronounced when reasoning must be done with important time constraints (Markovits et al., 2020). Importantly, results using eye-tracking have shown that part of the difference between counterexample and statistical reasoners is attentional (de Chantal et al., 2019a). Specifically, when faced with a belief-biased inference, counterexample reasoners spend relatively more time focused on the problem premises and less on the conclusion. This along with the pattern of some previous results suggests that reasoning strategies reflect a basic difference in the way that information is processed, one that applies to more general forms of judgment. More specifically, statistical reasoners use a more intuitive form of processing in which a broader range of information is processed by accessing information which is highly associated with problem parameters. Counterexample reasoners use more focused processing related to what are considered to be important parameters of a given problem and are more sensitive to potential counterexamples to any given judgment.

This broader analysis suggests that reasoning strategy should be a determinant of individual differences in a much broader range of judgments than are involved in the inferential tasks that led to its development. Corresponding to this, reasoning strategies have been found to modulate well-known gender differences in both the processing of negative emotions (Markovits et al., 2018) and mental rotation judgments (Markovits, 2019), with gender differences accentuated among statistical reasoners and attenuated among counterexample reasoners. More recently, reasoning strategy has been found to predict the extent to which people are prone to a variety of social stereotypes (Gagnon-St-Pierre et al., 2020).

In addition to these, recent studies have examined the extent to which reasoning strategies can predict the interaction between heuristic and logical responding compared to the most common explanatory mechanisms derived from dual process theories (i.e., working memory). Thompson and Markovits (2021) found that reasoning strategy alone explained variation in the extent to which logical reasoning is used in an array of heuristic judgments over and above the contribution of working memory, and other factors. In fact, studies which have introduced time constraints on reasoning have found that the relative contribution of reasoning strategy increases when time is reduced (Markovits, 2019; Markovits et al., 2020).

Overall, these studies suggest that there is a clear individual difference in the way that people process information that is critical to any form of reasoning. Statistical reasoners tend to use a more intuitive, probabilistic form of processing and are accordingly more influenced by heuristics and belief. Counterexample reasoners use a more focused form of representation of key problem parameters, and are correspondingly less influenced by heuristics and belief. Although most studies examining this model have

looked at adults, there is some evidence that this same distinction might apply to children (Markovits et al., 2019). This in turn suggests that at least some of the variation in learning and understanding basic concepts is related to the way that children are processing information in different ways.

Conclusion

Both the developmental patterns found with conditional reasoning and the individual differences highlighted by the dual strategy model suggest that reasoning and decision making cannot be understood as the result of applying abstract rules. They are better understood as forms of problem solving which are defined by a variety of different kinds of information. Whereas some of this information is explicitly described, other kinds of information will be accessed during the reasoning process. This latter reflects active use of people's store of knowledge and beliefs. In some cases, such as the production of potential alternative antecedents when reasoning with conditionals, this information is critical to producing the "logical" response. In other cases, such as evaluating the believability of a given conclusion, accessed information can result in producing the "nonlogical" response. One of the key components of reasoning ability is understanding the way that information that is actively accessed by a reasoner can have multiple effects. The fact that even well-educated adults, who are able to reason with abstract content at a reasonably high level, are still prone to the various heuristic effects shows the difficulty of differentiating these effects. One of the goals of education must therefore be the development of understanding of the way that people use implicitly generated information and how to distinguish contexts in which this information is useful from those in which it is harmful.

References

- Artman, L., Cahan, S., Avni-Babad, D., 2006. Age, schooling and conditional reasoning. *Cognit. Dev.* 21, 131–145.
- Bargh, J.A., Morsella, E., 2008. The unconscious mind. *Perspect. Psychol. Sci.* 3, 73–79.
- Barsalou, L.W., 1983. Ad hoc categories. *Mem. Cognit.* 11, 211–227.
- Brisson, J., Markovits, H., 2020. Reasoning strategies and semantic memory effects in deductive reasoning. *Mem. Cognit.* 1–11.
- Cheng, P.W., Holyoak, K.J., 1985. Pragmatic reasoning schemas. *Cognit. Psychol.* 17, 391–416.
- Cheng, P.W., Holyoak, K.J., Nisbett, R.E., Oliver, L.M., 1986. Pragmatic versus syntactic approaches to training deductive reasoning. *Cognit. Psychol.* 18, 293–328.
- Cropley, A., 2006. In praise of convergent thinking. *Creativ. Res. J.* 18, 391–404.
- Cummins, D.D., Lubart, T., Alksnis, O., Rist, R., 1991. Conditional reasoning and causation. *Mem. Cognit.* 19, 274–282.
- Cummins, D.D., 1995. Naive theories and causal deduction. *Mem. Cognit.* 23, 646–658.
- de Chantal, P.L., Markovits, H., 2017. The capacity to generate alternative ideas is more important than inhibition for logical reasoning in preschool-age children. *Mem. Cognit.* 45, 208–220.
- de Chantal, P.L., Newman, I.R., Thompson, V., Markovits, H., 2019a. Who resists belief-biased inferences? The role of individual differences in reasoning strategies, working memory, and attentional focus. *Mem. Cognit.* 1–17.
- de Chantal, P.L., Gagnon-St-Pierre, É., Markovits, H., 2019b. Divergent thinking promotes deductive reasoning in preschoolers. *Child Dev.* 91 (4), 1081–1097.
- de Neys, W., Glumicic, T., 2008. Conflict monitoring in dual process theories of thinking. *Cognition* 106, 1248–1299.
- de Neys, W., 2012. Bias and conflict a case for logical intuitions. *Perspect. Psychol. Sci.* 7, 28–38.
- Evans, J.S.B., Stanovich, K.E., 2013. Dual-process theories of higher cognition advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241.
- Evans, J.S.B., Barston, J.L., Pollard, P., 1983. On the conflict between logic and belief in syllogistic reasoning. *Mem. Cognit.* 11, 295–306.
- Evans, J.S.B., Over, D.E., Handley, S.J., 2005. Suppositionals, extensionality, and conditionals: a critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychol. Rev.* 112, 1040–1052.
- Evans, J.S.B.T., Over, D.E., Handley, S.J., 2007. Rethinking the model theory of conditionals. In: Schaeken, W., Vandierendonck, A., Schroyens, W., D'yevalle, G. (Eds.), *The Mental Models Theory of Reasoning: Refinements and Extensions*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Evans, J.S.B.T., 2008. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* 59, 255–278.
- Fillenbaum, S., 1975. If: some uses. *Psychol. Res.* 37, 245–260.
- Gagnon-St-Pierre, É., Doucerein, M.M., Markovits, H., 2020. Reasoning strategies explain individual differences in social reasoning. *J. Exp. Psychol. Gen.* 150 (2), 340–353.
- Gigerenzer, G., Selten, R., 2001. The adaptive toolbox. In: *Bounded Rationality: The Adaptive Toolbox*, pp. 37–50.
- Guilford, J.P., 1967. Creativity: yesterday, today and tomorrow. *J. Creativ. Behav.* 1, 3–14.
- Hawkins, J., Pea, R.D., Glick, J., Scribner, S., 1984. "Merds that laugh don't like mushrooms": evidence for deductive reasoning by preschoolers. *Dev. Psychol.* 20, 584–594.
- Holyoak, K.J., Cheng, P.W., 1995. Pragmatic reasoning with a point of view. *Think. Reas.* 1, 289–313.
- Janveau-Brennan, G., Markovits, H., 1999. The development of reasoning with causal conditionals. *Dev. Psychol.* 35, 904–911.
- Johnson-Laird, P.N., Byrne, R.M.J., 1991. *Deduction*. Erlbaum, Hove, London.
- Johnson-Laird, P.N., 2001. Mental models and deduction. *Trends Cognit. Sci.* 5, 434–442.
- Johnson-Laird, P.N., 2012. Inference with mental models. In: *The Oxford Handbook of Thinking and Reasoning*, pp. 134–145.
- Kahneman, D., 2011. *Thinking, Fast and Slow*. Macmillan.
- Light, P., Blaye, A., Gilly, M., Girotto, V., 1989. Pragmatic schemas and logical reasoning in 6-to 8-year-old children. *Cognit. Dev.* 4, 49–64.
- Markovits, H.B., Barrouillet, P., 2002. The development of conditional reasoning: a mental model account. *Dev. Rev.* 22, 5–36.
- Markovits, H., Brunet, M.-L., 2012. Priming divergent thinking promotes logical reasoning in 6-to 8-year olds: but more for high than low SES students. *J. Cognit. Psychol.* 24, 991–1001.
- Markovits, H., Lesage, C., 1990. Pragmatic reasoning schemas for conditional promises: context and representation. In: Caverni, J.-P., Fabre, J.M., Gonzalez, M. (Eds.), *Cognitive Biases. Advances in Psychology*. Oxford, England: North-Holland.
- Markovits, H., Lortie Forgues, H., 2011. Conditional reasoning with false premises facilitates the transition between familiar and abstract reasoning. *Child Dev.* 82, 646–660.
- Markovits, H., Nantel, G., 1989. The belief-bias effect in the production and evaluation of logical conclusions. *Mem. Cognit.* 17, 11–17.
- Markovits, H., Quinn, S., 2002. Efficiency of retrieval correlates with "logical" reasoning from causal conditional premises. *Mem. Cognit.* 30, 696–706.
- Markovits, H., Thompson, V., 2008. Different developmental patterns of simple deductive and probabilistic inferential reasoning. *Mem. Cognit.* 36, 1066–1078.
- Markovits, H., Vachon, R., 1989. Reasoning with contrary-to-fact propositions. *J. Exp. Child Psychol.* 47, 398–412.

- Markovits, H., Vachon, R., 1990. Conditional reasoning, representation, and level of abstraction. *Dev. Psychol.* 26, 942–951.
- Markovits, H., Forgues, H.L., Brunet, M.-L., 2012. More evidence for a dual-process model of conditional reasoning. *Mem. Cognit.* 40, 736–747.
- Markovits, H., Brunet, M.-L., Thompson, V., Brisson, J., 2013. Direct evidence for a dual process model of deductive inference. *J. Exp. Psychol. Learn. Mem. Cognit.* 49, 1213–1222.
- Markovits, H., Brisson, J., De Chantal, P., 2016. Logical reasoning versus information processing in the dual-strategy model of reasoning. *J. Exp. Psychol. Learn. Mem. Cognit.* 43 (1).
- Markovits, H., Brisson, J., De Chantal, P.-L., Thompson, V.A., 2017. Interactions between inferential strategies and belief bias. *Mem. Cognit.* 1–11.
- Markovits, H., Trémolière, B., Blanchette, I., 2018. Reasoning strategies modulate gender differences in emotion processing. *Cognition* 170, 76–82.
- Markovits, H., De Chantal, P.-L., Brisson, J., Gagnon-St-Pierre, É., 2019. The development of fast and slow inferential responding: evidence for a parallel development of rule-based and belief-based intuitions. *Mem. Cognit.* 47, 1188–1200.
- Markovits, H., De Chantal, P.-L., Brisson, J., Dubé, É., Thompson, V., Newman, I., 2020. Reasoning strategies predict use of very fast logical reasoning. *Mem. Cognit.* 1–12.
- Markovits, H., 1985. Incorrect conditional reasoning among adults: competence or performance? *Br. J. Psychol.* 76, 241–247.
- Markovits, H., 2000. A mental model analysis of young children's conditional reasoning with meaningful premises. *Think. Reas.* 6, 335–347.
- Markovits, H., 2014a. Conditional reasoning and semantic memory retrieval. In: *Reasoning as Memory*, vol. 53.
- Markovits, H., 2014b. On the road toward formal reasoning: reasoning with factual causal and contrary-to-fact causal premises during early adolescence. *J. Exp. Child Psychol.* 128, 37–51.
- Markovits, H., 2017a. The development of logical reasoning. In: *International Handbook of Thinking and Reasoning*. Routledge.
- Markovits, H., 2017b. In the beginning stages: conditional reasoning with category based and causal premises in 8-to 10-year olds. *Cognit. Dev.* 41, 1–9.
- Markovits, H., 2018. The development of logical reasoning. In: Thompson, V., Ball, L. (Eds.), *International Handbook on Thinking and Reasoning*. Psychology Press, Hove, UK.
- Markovits, H., 2019. Reasoning strategy modulates gender differences in performance on a spatial rotation task. *Q. J. Exp. Psychol.* 72, 2870–2876.
- Newman, I.R., Gibb, M., Thompson, V.A., 2017. Rule-based reasoning is fast and belief-based reasoning can be slow: challenging current explanations of belief-bias and base-rate neglect. *J. Exp. Psychol. Learn. Mem. Cognit.* 43, 1154.
- Noveck, I.A., 2001. When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78, 165–188.
- Oakhill, J., Johnson-Laird, P.N., Garnham, A., 1989. Believability and syllogistic reasoning. *Cognition* 31, 117–140.
- Oaksford, M., Chater, N., 2009. Précis of Bayesian rationality: the probabilistic approach to human reasoning. *Behav. Brain Sci.* 32, 69–84.
- Pacini, R., Epstein, S., 1999. The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *J. Pers. Soc. Psychol.* 76, 972.
- Pennycook, G., Cheyne, J.A., Seli, P., Koehler, D.J., Fugelsang, J.A., 2012. Analytic cognitive style predicts religious and paranormal belief. *Cognition* 123, 335–346.
- Quayle, J.D., Ball, L.J., 2000. Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *Q. J. Exp. Psychol.* 53A, 1202–1223.
- Quinn, S., Markovits, H., 1998. Conditional reasoning, causality, and the structure of semantic memory: strength of association as a predictive factor for content effects. *Cognition* 68, B93–B101.
- Rosen, V.M., Engle, R.W., 1996. The role of working memory capacity in retrieval. *J. Exp. Psychol. Gen.* 126, 211–227.
- Rumain, B., Connell, J., Braine, M.D., 1983. Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: if is not the biconditional. *Dev. Psychol.* 19, 471–481.
- Sloman, S.A., 1996. The empirical case for two systems of reasoning. *Psychol. Bull.* 119, 3–22.
- Stanovich, K.E., Toplak, M.E., 2012. Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind Soc.* 11, 3–13.
- Stanovich, K.E., West, R.F., 1997. Reasoning independently of prior belief and individual differences in actively open-minded thinking. *J. Educ. Psychol.* 89, 342.
- Stanovich, K.E., West, R.F., 2000. Individual differences in reasoning: implications for the rationality debate? *Behav. Brain Sci.* 23, 645–726.
- Thompson, V.A., Markovits, H., 2021. Reasoning strategies versus cognitive capacity as predictors of reasoning performance. *Cognition* 217.
- Thompson, V.A., Pennycook, G., Trippas, D., Evans, J.S.B., 2018. Do smart people have better intuitions? *J. Exp. Psychol. Gen.* 147, 945.
- Thompson, V.A., 1994. Interpretational factors in conditional reasoning. *Mem. Cognit.* 22, 742–758.
- Toplak, M.E., West, R.F., Stanovich, K.E., 2011. The Cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Mem. Cognit.* 39, 1275.
- Tversky, A., Kahneman, D., 2004. *Judgment Under Uncertainty: Heuristics and Biases*. Psychology Press, New York, NY, US.
- Venet, M., Markovits, H., 2001. Understanding uncertainty with abstract conditional premises. *Merrill-Palmer Q.* 47, 74–99.
- Verschueren, N., Schaeken, W., D'yevalle, G., 2005. Everyday conditional reasoning: a working memory-dependent tradeoff between counterexample and likelihood use. *Mem. Cognit.* 33, 107–119.
- Wallach, M.A., Kogan, N., 1965. A new look at the creativity-intelligence distinction 1. *J. Pers.* 33, 348–369.

Academic self-concept: a key predictor of achievement and learning

Geetanjali Basarkod and Herbert W. Marsh, Institute for Positive Psychology and Education, Australian Catholic University, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

The structure of self-concept	243
Impact of multidimensional structure on self-concept research	243
The link between academic self-concept and educational outcomes	244
Academic self-concept and academic achievement	244
Correlational evidence	244
Longitudinal evidence	244
Generalizability of evidence	245
Other learning outcomes	247
Summary and implications	247
References	248

Self-concept is a person's perceptions of themselves, formed through their experiences with and interpretations of their environment, and impacted by others' evaluations of them (Shavelson et al., 1976). It includes feelings of self-confidence, self-worth, self-acceptance, and perceptions of one's competence and ability. It affects how we act, feel, and adjust to a changing environment. Indeed, past research has shown positive self-concepts to be important for a variety of outcomes such as social and emotional health (Harter, 2012; Marsh et al., 2004), academic functioning (Chen et al., 2013; Marsh and Craven, 2006; Marsh and Yeung, 1997a), social and personal adjustment of children and adolescents (Harter, 1990; Parker et al., 2020), approaches to learning (Burnett et al., 2003), emotional adjustment and socialization (Donahue et al., 1993), bullying (Marsh et al., 2004), elite athlete success (Marsh and Perry, 2005), and parent-adolescent relations (Barber et al., 2003). Therefore, the enhancement of self-concept is a major goal in many fields such as education, child development, and sports.

In this chapter, we focus on self-concept within the field of education and the construct of academic self-concept (ASC). ASC is an individual's perceptions of their academic performance, and it sits within the multidimensional and hierarchical structure of self-concept. We first provide an overview of self-concept and its structure, to allow for a fuller conceptualization of ASC. We then discuss research on the link between ASC and educational outcomes.

The structure of self-concept

Self-concept is multifaceted and hierarchically organized, with global, general self-concept (also known as self-esteem) at the apex, inferences about the self in broad domains (e.g., social, physical, and academic) in the middle of the hierarchy, and perceptions of personal behavior in specific situations at the base of the hierarchy (Shavelson et al., 1976). Although contested in the 60s and 70s by proponents of a unidimensional structure of self-concept (Coopersmith, 1965; Marx and Winne, 1978), this multidimensional structure of self-concept was thoroughly assessed and confirmed via the Self-Description Questionnaire (SDQ; Marsh et al., 1983) by Marsh and Shavelson (1985).

Marsh and Shavelson (1985) conducted a series of studies with preadolescent and older adolescent children. They demonstrated that self-concept grew more and more multifaceted as children aged. While the self-concepts of preadolescent children comprised seven dimensions (i.e., physical appearance, physical abilities, peer relationships, parent relationships, reading, mathematics, and other school subjects), the self-concepts of older adolescent children was best explained by 13 dimensions (general, same-sex, opposite sex, parent relationship, emotional stability, problem-solving, religion/spirituality, honesty/dependability, physical ability, physical appearance, reading, math, and other school subjects). Providing evidence for the robustness of these factors, Marsh and Shavelson (1985) found that the correlations between specific self-concepts were small and that these correlations decreased with age. In the preadolescent samples, the average correlation between factors declined from $r = 0.27$ in Grade 2 to $r = 0.14$ in Grade 4, and in the older adolescent samples, all r s were lower than 0.10. Further support for this trend was provided by Marsh et al. (2005), with the development of a version of the SDQ for preschool children. Although children as young as four were able to distinguish between multiple dimensions of self-concept, the correlations between the facets were much higher at this young age ($r = 0.73$) compared to the typically small correlations observed with older children and adolescents.

Impact of multidimensional structure on self-concept research

Before we proceed to discussing the link between ASC and learning and other educational constructs, it is important to note how the multidimensionality of self-concept impacts research using global versus domain-specific measures. Although global self-concept

(also referred to as self-esteem) measures have stronger associations with global outcomes than domain-specific outcomes, the opposite is true for domain-specific self-concept measures. Therefore, using self-esteem measures in studies with a focus on broad, global outcomes is appropriate. However, ASC is more relevant to academic outcomes, and specific components of ASC are more relevant to outcomes concerning specific school subjects. For instance, [Valentine et al. \(2004\)](#) showed that stronger effects of self-belief measures are found when assessing self-beliefs specific to the academic domain and when self-beliefs and achievement are matched by the domain (i.e., same subject area).

The link between academic self-concept and educational outcomes

Past studies have consistently shown ASC to positively impact a host of learning and other educational outcomes, such as academic achievement ([Marsh and Craven, 2006](#)), interest and satisfaction in school ([Marsh et al., 2005](#)), course selection ([Marsh et al., 2019](#); [Marsh and Yeung, 1997a](#); [Parker et al., 2014](#)) persistence, and long-term attainment ([Guo et al., 2015a,b](#); [Marsh and O'Mara, 2008](#)). These effects hold even after accounting for prior achievement and IQ, suggesting that a positive ASC is highly desirable not only as a goal, but also to facilitate subsequent learning and academic accomplishments. In this section, we first outline the literature with regard to ASC and academic achievement, and then highlight how ASC is beneficial for other learning outcomes.

Academic self-concept and academic achievement

Correlational evidence

Studies have shown that ASC is substantially correlated to numerous academic achievement measures (e.g., test scores, school grades), and across domains. Evidence for the association between ASC and test scores has been garnered from numerous cross-national databases. For instance, using data from 15-year-old students from the 26 countries participating in the Programme for International Students Assessment (PISA) 2000 cohort, [Marsh and Hau \(2003\)](#) showed that the overall effect (in terms of standardized beta coefficients—for this and for the following estimates) of individual student academic achievement (non-subject specific) on individual student ASC was 0.38. [Seaton et al. \(2009\)](#); 41 countries in PISA 2003) and [Marsh et al. \(2020\)](#); 68 countries in PISA 2012) showed that the overall effects of math achievement on math self-concept were 0.52 and 0.91, respectively. [Nagengast and Marsh \(2011\)](#); 56 countries in PISA 2006) showed that the overall effect of science achievement on science self-concept was 0.17. Lastly, [Basarkod et al. \(2020\)](#); 72 countries in PISA 2018) showed that the overall effect of individual student reading achievement on individual student reading self-concept of competence was 0.35 and on individual student reading self-concept of difficulty was -0.32 . Note that all these effects were from three-level models (students nested within schools nested within countries) that accounted for a host of independent (e.g., school- and country-average achievement) and control variables (e.g., gender, socio-economic status).

Using another international dataset—the combined Trends in Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) dataset from 2011—[Guo et al. \(2018\)](#) also showed cross-national evidence for the robust association between ASC and achievement for 4th grade students (from the 15 Organization for Economic Cooperation and Development [OECD] countries that participated in these surveys). Specifically, they showed that the overall zero-order correlations between self-concept and individual student achievement in the same domain were 0.31 for math, 0.20 for science, and 0.40 for reading. The average beta estimates from two-level models in each country (controlling for other independent variables and covariates) were 0.42 for math, 0.19 for science, and 0.42 for reading.

Importantly, correlational studies have also shown that the link between ASC and achievement is stronger than that for other learning measures with achievement ([Marsh and Martin, 2011](#)). This is best highlighted by results of the OECD-PISA's Students' Approaches to Learning instrument (SAL; [Marsh et al., 2006](#)). The SAL measures 14 factors that assess self-regulated learning strategies, self-beliefs, motivation, and learning preferences. Marsh et al. evaluated SAL responses from nationally representative samples of approximately 4000 15-year-olds from each of 25 countries ($N = 107,899$). Across this set of 14 factors, mathematics and verbal achievement were consistently more strongly correlated with the three (math, verbal, and academic) self-concept measures than they were with any other SAL constructs. Furthermore, formal tests of factorial invariance showed that the pattern, direction, and relative sizes of these correlations were relatively invariant across the 26 countries.

Longitudinal evidence

Evidence from cross-national datasets attest to the robustness of the association between self-concept and achievement measures. However, these datasets have been cross-sectional, and therefore, do not provide clarity on the causal relationship between the two constructs. For this, we need to turn to longitudinal studies. [Byrne \(1984\)](#) proposed three criteria that studies addressing causal relationships must satisfy: (a) a statistical relationship must be established, (b) a clearly established time precedence must be evident, and (c) a causal model must be correctly specified.

Early studies addressed the question of what comes first, achievement or self-concept, and tested two competing hypotheses: the self-enhancement model and the skill-development model. According to the self-enhancement model, self-concept is a primary determinant of academic achievement. That is, individuals who have high academic achievement levels do so because of their high levels of ASC. In contrast, the skill-development model posits the opposite causal ordering; ASC emerges principally as a consequence of academic achievement so that it is enhanced by developing stronger academic skills. These early studies (e.g., [Calsyn and](#)

Kenny, 1977), however, aimed to determine which of these paths was the predominant one, rather than whether paths in both directions were significant. That is, self-concept research did not claim that it was solely self-concept that caused achievement, or vice versa, but that both these constructs played some role in predicting the other. Much like the nature versus nurture debate that has plagued psychology, the question was not *which* of these two contrasting hypotheses was “true,” rather *whether* both of these acted in conjunction. In addition, these early studies were not methodologically strong as they were largely based on correlational evidence and did not use more methodologically appropriate techniques such as SEMs (see Marsh, 1990 for a detailed review of these arguments).

It was not until the 1990s that both the self-enhancement and skill-development models were tested simultaneously in the reciprocal effects model (REM). This was first done by Marsh (1990) with data from the large, nationally representative US Youth in Transition study. He considered data from Times 1 (early 10th Grade), 2 (late 11th Grade), 3 (late 12th Grade), and 4 (1 year after normal high school graduation). Three latent constructs were considered: academic ability inferred on the basis of four standardized test scores, ASC inferred from self-report responses, and school grades. Of particular importance were the effects of latent constructs in one wave on latent constructs in subsequent waves. At T2, ASC was influenced by academic ability and T1 ASC, but not by T1 grades. At T2, school grades were influenced both by T1 ASC and by T1 school grades. Similarly, school grades at T3 were influenced significantly both by T2 ASC and by T2 grades. ASC at T4 was influenced significantly by ASC at T2 (there was no T3 ASC measure), but not by T3 school grades. The findings provide strong support for the effect of prior self-concept on subsequent school grades as they spanned the two time intervals.

Since the Marsh (1990) seminal study, there have been plenty of empirical studies, cross-cultural studies, as well as meta-analyses that lend support to the REM predictions (Huang, 2011; Marsh and Martin, 2011; Marsh et al., 2018; Valentine et al., 2004). Marsh et al. (2018) demonstrated the reciprocal effects of math self-concept and math achievement over 6 years in a sample of 3370 German students from 42 secondary schools. They used two measures of math achievement—test scores and grades. For models run separately for each of the achievement measures, and with time-invariant cross-lags, the path estimates (standardized beta weights) from prior math self-concept to subsequent math test scores was 0.05 ($SE = 0.01$) and to subsequent school grades was 0.10 ($SE = 0.01$). The path from prior test scores to subsequent math self-concept was 0.13 ($SE = 0.02$) and from prior school grades to subsequent math self-concept was 0.16 ($SE = 0.01$). Thus, although the estimates for achievement predicting self-concept were larger, both directions were significant and positive (results from a model simultaneously accounting for both school grades and test scores were similar and significant, albeit marginally smaller).

Generalizability of evidence

Across cultures

The REM of ASC and achievement has been shown in different cultural settings, such as in a large nationally representative sample of Hong Kong students (Marsh et al., 2002) and large samples of East and West German students at the time of the fall of the Berlin Wall (Marsh and Köller, 2003; Marsh et al., 2001). Support for the generalizability also comes from research based on French-speaking Canadian primary students (Guay et al., 2003), German high school students (Marsh et al., 2005), and 487 high school students (Grades 7, 8, and 9) from mainland China (Yeung and Lee, 1999).

More generally, in their meta-analysis of REM studies, Valentine et al. (2004) considered the country where the sample was drawn from as a moderator variable. However, they found that support for the REM did not differ as a function of country. Although this research provides cross-national and cross-cultural support for the REM, we note that the majority of the research comes from Western and industrialized countries, so it is premature to claim that support for the REM is universal.

Across achievement tracks

Seaton et al. (2015) tested the equivalence of the REM across a sample of high school students who attended academically selective ($n = 738$) and mixed-ability comprehensive ($n = 2048$) schools in Australia. The average path estimate (standardized beta weight) for prior mathematics self-concept predicting subsequent mathematics achievement was 0.06, and that for prior mathematics achievement predicting subsequent mathematics self-concept was 0.07. All cross-paths were significant. Further, their multigroup analyses revealed that these reciprocal effects were present for students attending both academically selective and mixed-ability comprehensive schools. In addition, there were no differences between the groups in either the size or the direction of the paths.

Similarly, Arens et al. (2017) examined the REM in a sample of German students ($N = 3425$) from grades 5 through 9, who attended different achievement tracks. The average path estimates (standardized beta weights) across all groups and time points were as follows: math self-concept to math test scores, 0.056 (range = 0.054–0.059); math self-concept to math school grades, 0.047 (range = 0.045–0.049); math test scores to math self-concept, 0.082 (range = 0.075–0.093); and math school grades to math self-concept, 0.087 (range = 0.083–0.093). Thus, the paths were consistent across time intervals, types of achievement measures, and different achievement tracks. Although the results presented here were from a combined model, they were also replicated when separate models were used. In addition, as the students were tracked from when they were in Grade 4 to Grade 9, this study also provided evidence for the REM to be robust across age.

Across age groups

A study by Guay et al. (2003) also addressed the issue of whether the REM held across age groups. They used a multicohort, multi-occasion design, a methodological approach that is especially well-suited to address this issue. In particular, they considered responses of students at T1 in Grades 2–4 (i.e., three age cohorts 8–10 years of age). Responses for all three cohorts were then

collected annually for the next 3 years (i.e., the three measurement occasions). They found strong support for the REM for all three age cohorts, and these results were reasonably invariant when rigorously tested with multigroup tests of invariance across the three age cohort groups.

Across subject areas

Although most studies examining the REM have focused on math, research in the domains of reading and science has also provided evidence for the model (e.g., Chiu, 2012; Sewasew and Koester, 2019). For instance, Sewasew and Koester (2019) analyzed the REM for reading using a sample of primary school children in the United States from the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) dataset. The paths from prior reading self-concept to subsequent reading achievement ($\beta = 0.08$), and from prior reading achievement to subsequent reading self-concept ($\beta = 0.25$), were both significantly positive. Indeed, as is the case with research in the math domain (e.g., Marsh, 1990), they also found that the path from achievement to self-concept was stronger than the path in the opposite direction. Evidence for the REM for reading has been observed in multiple languages such as English, German, and Chinese (Chen et al., 2013; Möller et al., 2014; Niepel et al., 2014).

Controlling for other variables

The bidirectional causal relationship between self-concept and achievement holds even after controlling for other related constructs. For instance, Seaton et al. (2014) used a sample of 2786 Australian high school students (ages 11–17) who participated at four time points, each 6 months apart. Separate models indicated that there were reciprocal associations between mathematics self-concept and achievement, as well as between mathematics performance goal orientation and achievement. There was little evidence of such reciprocal relations between mastery goal orientation and achievement. Important to our discussion here, only self-concept had significant reciprocal relationships with achievement when all variables were included in a single model. Similarly, using a sample of 702 8th grade students in China, Xu (2018) showed that the positive effect of math homework self-concept on math achievement remained even after controlling for students' levels of interest and effort for the homework.

Meta-analytic evidence for moderators

In a theoretical review and meta-analysis of studies on the REM between self-beliefs (including self-concept, self-efficacy, self-esteem), Valentine et al. (2004) reviewed 55 publications on the REM. This meta-analysis included 60 independent samples and 282 different effect sizes. The effect of prior self-beliefs on subsequent achievement, controlling for the effects of prior achievement, was highly significant overall and positive in 90% of the studies in their meta-analysis (average $\beta = 0.08$, 95% CI = 0.06–0.10). The same was true for achievement predicting self-beliefs (average $\beta = 0.15$, 95% CI = 0.12–0.18).

Valentine et al. (2004) considered a wide variety of potential moderators of the REM effects: year the study was published/reported; base year of data collection; sample size of the study; stability of the achievement measure (i.e., T1–T2 stability coefficient); reliability of the self-measure; the number of variables used as controls in the analysis; whether the effect size was from an analysis of manifest or latent variables; use of a convenience sample versus random selection from a known population; age of students; type of achievement; time interval between the collection of T1 and T2 measures; and country from which the sample of students came. However, none of these potential moderators had a significant effect on the size of the REM. These meta-analysis results provide compelling support for the robustness and generalizability of the REM in a way that could not be achieved on the basis of any single primary study.

The only two variables that had a substantial moderating effect on the REM was the globality (versus specificity) of the self-belief measure and school transition. In support of a multidimensional perspective of self-concept (and indeed, self-belief measures in general), Valentine et al. (2004) found that the effects of prior self-beliefs were significantly stronger when the self-belief measure was based on academic self-beliefs rather than on global measures such as self-esteem. This was also the case when the self-belief and achievement measures were matched in terms of subject area (e.g., mathematics achievement and math self-concept). In particular, they reported little evidence of any effects of global or generalized self-beliefs (e.g., self-esteem) on academic achievement. In terms of school transition, they found that the strength of the REM was weaker in studies where students experienced school transition (e.g., from elementary school to middle school). In particular, the effects of prior self-beliefs (collected prior to the transition between school stages) on subsequent achievement (collected after the transition) were smaller than in studies where there was no transition.

A comprehensive meta-analysis on the reciprocal relationship between self-concept (rather than all self-beliefs) and achievement by Huang (2011) examined 39 independent and longitudinal samples by integrating meta-analytic and path analysis procedures. Huang (2011) found similar results to Valentine et al. (2004). Specifically, the mean observed correlation between prior self-concept and subsequent achievement ranged from 0.20 to 0.27, whereas that for the opposite path ranged from 0.19 to 0.25. The only significant moderator of the REM was the globality (versus specificity) of the self-concept measures used. The other moderators they tested (i.e., globality of academic achievement measure, the indicator of academic achievement, delay between time waves, mean age of sample at baseline) were not significant.

These meta-analyses provide clear support for the REM, the robustness and generalizability of the effects, the theoretical focus on a multidimensional perspective, and the methodological focus on a construct validity approach that has been central to REM studies and self-concept research more generally.

Other learning outcomes

Besides being a positive and significant predictor of academic achievement, ASC is also important for a host of other learning and educational outcomes. For instance, Marsh and Yeung (1997a,b) found that, whereas self-concept, academic achievement, and coursework selection were all highly correlated, prior self-concept was a much better predictor of subsequent coursework choice than was prior academic achievement. Indeed, self-concept predicted course selection in many different academic subjects (i.e., English, math, science, computer science, history, economics, geography, foreign language, and industrial arts). In another study, Marsh et al. (2005) tested the REM between self-concept and achievement (both school grades and test scores) and included academic interest in the model. They conducted these SEMs using longitudinal data based on two large, nationally representative samples of German high school students. For both samples, they found clear support for the REM based on ASC and achievement, demonstrating that the effect of prior math self-concept was substantial for subsequent math school performance as well as for math test scores. Important for our discussion here, prior self-concept also significantly influenced subsequent academic interest measures beyond the effects of earlier measures of school performance, achievement test scores, and academic interest. In contrast, prior academic interest had only a small effect on subsequent self-concept and little or no effect on either school performance or test scores beyond what could be explained by self-concept.

Self-concept is also a positive predictor of academic aspirations and attitudes toward school. For example, Marsh (1991) investigated the effect of school-average ability on a wide range of educational outcomes including academic aspirations. When controlling for ASC, the previously confirmed negative effects of school-average ability were reduced and, as such, ASC showed considerable total and direct effects on academic aspirations. In a study by Green et al. (2012), self-concept was shown to positively predict attitudes toward school ($\beta = 0.22$ and 0.28) over and above the effect of academic motivation constructs. In turn, attitudes toward school positively predicted class participation, homework completion, test performance, and negatively predicted absenteeism.

Wouters et al. (2011) demonstrated that self-concept also predicts ($\beta = 0.28$) academic adjustment (i.e., the extent to which students can meet their academic demands) and success in the 1st year of higher education ($\beta = 0.20$), after accounting for high school academic achievement. They conducted this study with a sample of 536 students in Belgium, a country where more than half of the students who enroll in university do not graduate (Groenez et al., 2003). Indeed, the OECD has reported that the average university completion rate across 24 OECD countries is 36%. Thus, the finding that self-concept positively predicts meeting university demands and success at university, in the crucial adjustment period of 1st year, is an important one.

Summary and implications

This chapter emphasizes the importance of ASC for learning and other educational outcomes. We provide clear and robust evidence regarding the bidirectional and reciprocal relationship between ASC and achievement. This causal ordering is consistent across cultures, achievement tracks, ages, and subject areas. It also holds when controlling for other important educational constructs, and is only moderated by the globality versus specificity of the measures used, and the presence versus absence of school transition. In addition, we also outline studies that show ASC positively predicts other learning outcomes such as coursework selection, academic interest, academic aspirations, attitudes toward school, academic adjustment, and success in higher education, while controlling for prior achievement. Although the sheer number of studies that show the REM is promising, it should be noted that the majority of the studies utilize the same methodology (i.e., cross-lag panel models; CLPM). Future research is required to supplement these studies by testing the causal relationship between self-concept and achievement with newly developed longitudinal methods such as the random-intercept cross lag panel model (RI-CLPM) or the bivariate STARTS (Stable Trait, Autoregressive Trait, and State) model (see Usami et al., 2019 for a description of these methods).

It is important to note, however, that the cross-lag estimates from these models would need to be interpreted differently (Orth et al., 2020). For instance, a significant positive cross-lag from self-concept to achievement using the CLPM would indicate that individuals who have higher self-concept at Time 1 are more likely to have higher achievement levels at Time 2, compared to individuals who have lower self-concept at Time 1. In contrast, the same cross-lag in a RI-CLPM would imply that individuals who have higher self-concept than usual at Time 1 will show subsequent increases in achievement, while individuals who have lower self-concept than usual at Time 1 will show subsequent decreases in achievement.

The fact that self-concept and achievement are mutually reinforcing has important implications for policy and practice. On the one hand, the finding that self-concept promotes achievement implies that programs and teaching practices should be designed to protect self-concept (self-enhancement interventions). On the other hand, the finding that achievement promotes self-concept provides a basis for increasing standards and accountability for student learning (e.g., mandatory promotion/graduation requirements; skill development interventions). Based on their meta-analytic findings, however, Huang (2011) concluded that “as high self-concept is related to high academic performance and vice versa, intervention programs that combine self-enhancement and skill development should be integrated” (p. 505). Thus, the ideal interventions would not aim to improve self-concept or achievement solely, but both concurrently.

It is important to note that although there is ample evidence for positive and significant paths in both directions, the path from achievement to self-concept is generally stronger than the path from self-concept to achievement. Still, one cannot be improved at the detriment of the other. For example, if teachers enhance students' ASCs without improving achievement, then self-concept gains

are likely to be short-lived. However, if teachers improve students' academic achievement without also fostering students' self-beliefs in their academic capabilities, then the achievement gains are also unlikely to be long lasting. If teachers focus on either one of these constructs to the exclusion of the other, both are likely to suffer. In addition, as evidenced by the studies discussed in this chapter, especially through the moderators examined in the reported meta-analyses, the REM effects tend to be stronger in specific domains, rather than in academia in general. It follows then that interventions that aim at increasing achievement may want to focus on specific domains rather than achievement in general. In this way, intervening to increase achievement and self-concept within the same domain may multiply the benefits.

References

- Arens, A.K., Marsh, H.W., Pekrun, R., Lichtenfeld, S., Murayama, K., vom Hofe, R., 2017. Math self-concept, grades, and achievement test scores: long-term reciprocal effects across five waves and three achievement tracks. *J. Educ. Psychol.* 109 (5), 621. <https://doi.org/10.1037/edu0000163>.
- Barber, C.N., Ball, J., Armistead, L., 2003. Parent-adolescent relationship and adolescent psychological functioning among African-American female adolescents: self-esteem as a mediator. *J. Child Fam. Stud.* 12 (3), 361–374. <https://doi.org/10.1023/A:1023948029266>.
- Basarkod, G., Marsh, H.W., Guo, J., Dicke, T., Xu, K.M., Parker, P., 2020. The Big-Fish-Little-Pond Effect for Reading Self-Beliefs: A Cross-National Exploration With PISA 2018. *EdArXiv*. <https://doi.org/10.35542/osf.io/7wbxj>.
- Burnett, P.C., Pillay, H., Dart, B.C., 2003. The influences of conceptions of learning and learner self-concept on high school students' approaches to learning. *Sch. Psychol. Int.* 24 (1), 54–66. <https://doi.org/10.1177/0143034303024001621>.
- Byrne, B.M., 1984. The general/academic self-concept nomological network: a review of construct validation research. *Rev. Educ. Res.* 54 (3), 427–456. <https://doi.org/10.3102/00346543054003427>.
- Calsyn, R., Kenny, D., 1977. Self-concept of ability and perceived evaluations by others: cause or effect of academic achievement? *J. Educ. Psychol.* 69, 136–145. <https://doi.org/10.1037/0022-0663.69.1.136>.
- Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., Lin, S.S., 2013. The relationship between academic self-concept and achievement: a multicohort–multioccasion study. *Learn. Indiv. Differ.* 23, 172–178. <https://doi.org/10.1016/j.lindif.2012.07.021>.
- Chiu, M.-S., 2012. Differential psychological processes underlying the skill-development model and self-enhancement model across mathematics and science in 28 countries. *Int. J. Sci. Math. Educ.* 10 (3), 611–642. <https://doi.org/10.1007/s10763-011-9309-9>.
- Coopersmith, S., 1965. *The Antecedents of Self-Esteem*. Princeton.
- Donahue, E.M., Robins, R.W., Roberts, B.W., John, O.P., 1993. The divided self: concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *J. Pers. Soc. Psychol.* 64 (5), 834–846. <https://doi.org/10.1037/0022-3514.64.5.834>.
- Green, J., Liem, G.A.D., Martin, A.J., Colmar, S., Marsh, H.W., McInerney, D., 2012. Academic motivation, self-concept, engagement, and performance in high school: key processes from a longitudinal perspective. *J. Adolesc.* 35 (5), 1111–1122. <https://doi.org/10.1016/j.adolescence.2012.02.016>.
- Groenez, S., Van den Brande, I., Nicaise, I., 2003. *Cijferboek sociale ongelijkheid in het Vlaamse onderwijs. Een verkennend onderzoek op de Panelstudie van Belgische huishoudens*.
- Guay, F., Marsh, H.W., Boivin, M., 2003. Academic self-concept and academic achievement: developmental perspectives on their causal ordering. *J. Educ. Psychol.* 95 (1), 124–136. <https://doi.org/10.1037/0022-0663.95.1.124>.
- Guo, J., Marsh, H.W., Morin, A.J., Parker, P.D., Kaur, G., 2015a. Directionality of the associations of high school expectancy-value, aspirations, and attainment: a longitudinal study. *Am. Educ. Res. J.* 52 (2), 371–402. <https://doi.org/10.3102/0002831214565786>.
- Guo, J., Parker, P.D., Marsh, H.W., Morin, A.J., 2015b. Achievement, motivation, and educational choices: a longitudinal study of expectancy and value using a multiplicative perspective. *Dev. Psychol.* 51 (8), 1163–1176. <https://doi.org/10.1037/a0039440>.
- Guo, J., Marsh, H.W., Parker, P.D., Dicke, T., 2018. Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learn. Instr.* 58, 210–219. <https://doi.org/10.1016/j.learninstruc.2018.07.007>.
- Harter, S., 1990. Processes underlying adolescent self-concept formation. In: Montemayor, R., Adams, G.R., Gullotta, T.P. (Eds.), *Advances in Adolescent Development: An Annual Book Series*, Vol. 2. From Childhood to Adolescence: A Transitional Period? Sage Publications, Inc, pp. 205–239.
- Harter, S., 2012. *Self-Perception Profile for Adolescents: Manual and Questionnaires*. University of Denver, Department of Psychology. <https://doi.org/10.1037/t05703-000>.
- Huang, C., 2011. Self-concept and academic achievement: a meta-analysis of longitudinal relations. *J. Sch. Psychol.* 49 (5), 505–528. <https://doi.org/10.1016/j.jsp.2011.07.001>.
- Marsh, H.W., 1990. Causal ordering of academic self-concept and academic achievement: a multiwave, longitudinal panel analysis. *J. Educ. Psychol.* 82 (4), 646–656. <https://doi.org/10.1037/0022-0663.82.4.646>.
- Marsh, H.W., 1991. Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *Am. Educ. Res. J.* 28 (2), 445–480. <https://doi.org/10.3102/00028312028002445>.
- Marsh, H.W., Craven, R.G., 2006. Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1 (2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>.
- Marsh, H.W., Hau, K.-T., 2003. Big-Fish–Little-Pond effect on academic self-concept: a cross-cultural (26-country) test of the negative effects of academically selective schools. *Am. Psychol.* 58 (5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>.
- Marsh, H.W., Köller, O., 2003. Bringing together two theoretical models of relations between academic self-concept and achievement. *Int. Adv. Self Res.* 1, 17–45.
- Marsh, H.W., Martin, A.J., 2011. Academic self-concept and academic achievement: relations and causal ordering. *Br. J. Educ. Psychol.* 81 (1), 59–77. <https://doi.org/10.1348/000709910X503501>.
- Marsh, H.W., O'Mara, A., 2008. Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. *Pers. Soc. Psychol. Bull.* 34 (4), 542–552. <https://doi.org/10.1177/0146167207312313>.
- Marsh, H.W., Perry, C., 2005. Self-concept contributes to winning gold medals: causal ordering of self-concept and elite swimming performance. *J. Sport Exerc. Psychol.* 27 (1), 71–91. <https://doi.org/10.1123/jsep.27.1.71>.
- Marsh, H.W., Shavelson, R., 1985. Self-concept: its multifaceted, hierarchical structure. *Educ. Psychol.* 20 (3), 107–123. https://doi.org/10.1207/s15326985sep2003_1.
- Marsh, H.W., Yeung, A.S., 1997a. Causal effects of academic self-concept on academic achievement: structural equation models of longitudinal data. *J. Educ. Psychol.* 89 (1), 41–54. <https://doi.org/10.1037/0022-0663.89.1.41>.
- Marsh, H.W., Yeung, A.S., 1997b. Coursework selection: relations to academic self-concept and achievement. *Am. Educ. Res. J.* 34 (4), 691–720. <https://doi.org/10.3102/00028312034004691>.
- Marsh, H.W., Relich, J.D., Smith, I.D., 1983. Self-concept: the construct validity of interpretations based upon the SDQ. *J. Pers. Soc. Psychol.* 45 (1), 173–187. <https://doi.org/10.1037/0022-3514.45.1.173>.
- Marsh, H.W., Köller, O., Baumert, J., 2001. Reunification of East and West German school systems: longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *Am. Educ. Res. J.* 38 (2), 321–350. <https://doi.org/10.3102/00028312038002321>.

- Marsh, H.W., Hau, K.-T., Kong, C.-K., 2002. Multilevel causal ordering of academic self-concept and achievement: influence of language of instruction (English compared with Chinese) for Hong Kong students. *Am. Educ. Res. J.* 39 (3), 727–763. <https://doi.org/10.3102/00028312039003727>.
- Marsh, H.W., Parada, R.H., Craven, R.G., Finger, L., 2004. In the looking glass: a reciprocal effects model elucidating the complex nature of bullying, psychological determinants, and the central role of self-concept. In: Sanders, C.E., Pyle, G.D. (Eds.), *Bullying*. Elsevier, pp. 63–106. <https://doi.org/10.1016/B978-012617955-2/50009-6>.
- Marsh, H.W., Trautwein, U., Lüdtke, O., Köller, O., Baumert, J., 2005. Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76 (2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>.
- Marsh, H.W., Hau, K.-T., Artelt, C., Baumert, J., Peschar, J.L., 2006. OECD's brief self-report measure of educational psychology's most useful affective constructs: cross-cultural, psychometric comparisons across 25 countries. *Int. J. Test.* 6 (4), 311–360. https://doi.org/10.1207/s15327574ijt0604_1.
- Marsh, H.W., Pekrun, R., Murayama, K., Arens, A.K., Parker, P.D., Guo, J., Dicke, T., 2018. An integrated model of academic self-concept development: academic self-concept, grades, test scores, and tracking over 6 years. *Dev. Psychol.* 54 (2), 263. <https://doi.org/10.1037/dev0000393>.
- Marsh, H.W., Parker, P.D., Guo, J., Pekrun, R., Basarkod, G., 2020. Psychological comparison processes and self-concept in relation to five distinct frame-of-reference effects: Pan-human cross-cultural generalizability over 68 countries. *Eur. J. Pers.* 34 (2), 180–202. <https://doi.org/10.1002/per.2232>.
- Marx, R.W., Winne, P.H., 1978. Construct interpretations of three self-concept inventories. *Am. Educ. Res. J.* 15 (1), 99–109. <https://doi.org/10.3102/00028312015001099>.
- Möller, J., Zimmermann, F., Köller, O., 2014. The reciprocal internal/external frame of reference model using grades and test scores. *Br. J. Educ. Psychol.* 84 (4), 591–611. <https://doi.org/10.1111/bjep.12047>.
- Nagengast, B., Marsh, H.W., 2011. The negative effect of school-average ability on science self-concept in the UK, the UK countries and the world: the Big-Fish-Little-Pond-Effect for PISA 2006. *Educ. Psychol.* 31 (5), 629–656. <https://doi.org/10.1080/01443410.2011.586416>.
- Niepel, C., Brunner, M., Preckel, F., 2014. The longitudinal interplay of students' academic self-concepts and achievements within and across domains: replicating and extending the reciprocal internal/external frame of reference model. *J. Educ. Psychol.* 106 (4), 1170–1191. <https://doi.org/10.1037/a0036307>.
- Orth, U., Clark, D.A., Donnellan, M.B., Robins, R.W., 2020. Testing prospective effects in longitudinal research: comparing seven competing cross-lagged models. *J. Pers. Soc. Psychol.* <https://doi.org/10.1037/pspp0000358>.
- Parker, P.D., Trautwein, U., Marsh, H.W., Basarkod, G., Dicke, T., 2020. Development in relationship self-concept from high school to university predicts adjustment. *Dev. Psychol.* 56 (8), 1547–1555. <https://doi.org/10.1037/dev0000992>.
- Seaton, M., Marsh, H.W., Craven, R.G., 2009. Earning its place as a pan-human theory: universality of the big-fish-little-pond effect across 41 culturally and economically diverse countries. *J. Educ. Psychol.* 101 (2), 403–419. <https://doi.org/10.1037/a0013838>.
- Seaton, M., Parker, P., Marsh, H.W., Craven, R.G., Yeung, A.S., 2014. The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educ. Psychol.* 34 (1), 49–72. <https://doi.org/10.1080/01443410.2013.825232>.
- Seaton, M., Marsh, H.W., Parker, P.D., Craven, R.G., Yeung, A.S., 2015. The reciprocal effects model revisited: extending its reach to gifted students attending academically selective schools. *Gift. Child. Q.* 59 (3), 143–156. <https://doi.org/10.1177/0016986215583870>.
- Sewasew, D., Koester, L.S., 2019. The developmental dynamics of students' reading self-concept and reading competence: examining reciprocal relations and ethnic-background patterns. *Learn. Indiv. Differ.* 73, 102–111. <https://doi.org/10.1016/j.lindif.2019.05.010>.
- Shavelson, R.J., Hubner, J.J., Stanton, G.C., 1976. Self-concept: validation of construct interpretations. *Rev. Educ. Res.* 46 (3), 407–441. <https://doi.org/10.3102/00346543046003407>.
- Usami, S., Todo, N., Murayama, K., 2019. Modeling reciprocal effects in medical research: critical discussion on the current practices and potential alternative models. *PLoS One* 14 (9), e0209133. <https://doi.org/10.1371/journal.pone.0209133>.
- Valentine, J.C., DuBois, D.L., Cooper, H., 2004. The relation between self-beliefs and academic achievement: a meta-analytic review. *Educ. Psychol.* 39 (2), 111–133. https://doi.org/10.1207/s15326985ep3902_3.
- Wouters, S., Germeijs, V., Colpin, H., Verschueren, K., 2011. Academic self-concept in high school: predictors and effects on adjustment in higher education. *Scand. J. Psychol.* 52 (6), 586–594. <https://doi.org/10.1111/j.1467-9450.2011.00905.x>.
- Xu, J., 2018. Reciprocal effects of homework self-concept, interest, effort, and math achievement. *Contemp. Educ. Psychol.* 55, 42–52. <https://doi.org/10.1016/j.cedpsych.2018.09.002>.
- Yeung, A.S., Lee, F.L., 1999. Self-concept of high school students in China: confirmatory factor analysis of longitudinal data. *Educ. Psychol. Meas.* 59 (3), 431–450. <https://doi.org/10.1177/00131649921969965>.

Self-efficacy

Hyun Ji Lee and Mimi Bong, Department of Education and the Brain & Motivation Research Institute (bMRI), Korea University, Seoul, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Introduction	250
Definition of self-efficacy	250
Sources of self-efficacy	251
Roles of age, gender, and culture in self-efficacy formation	252
Motivational, behavioral, and achievement consequences of self-efficacy	252
Measuring self-efficacy	253
Promoting self-efficacy of learners	254
Concluding remarks	255
References	255

Introduction

In the social cognitive theory of human functioning, [Bandura \(1989, 1997\)](#) posits the reciprocal interactions among person, behavior, and environment. This theory emphasizes human agency, considering humans as active agents who pursue a goal and exercise control over the course of actions toward the goal. Humans can interpret their behaviors and environment, reflect upon themselves, and regulate their behaviors based on expectations, which in turn can influence their environment. Self-efficacy, the personal conviction that one can perform a given task successfully, lies at the center of this whole self-regulatory process. Without strong self-efficacy, individuals are not able to feel a sense of agency in a situation and are less willing to engage in a task even for potentially positive outcomes. In academic contexts, self-efficacy (i.e., academic self-efficacy; [Schunk, 1991](#)) plays an important role in affecting a vast array of learners' academic motivation and behaviors, including their goal pursuits, academic emotions, choice, engagement, self-regulation, and achievement.

Definition of self-efficacy

Self-efficacy represents one's personal conviction about her or his capability to successfully perform a specific behavior or a specific task at the desired level ([Bandura, 1997](#); [Schunk, 1991](#)). It is a subjective judgment and, as such, does not necessarily reflect one's actual capabilities. Rather, individuals collect relevant information from diverse sources about themselves, the task, and surrounding environments and estimate the possibility that they could successfully accomplish the task under the given circumstances. Self-efficacy perceptions are thus highly context-specific. For example, students may hold different self-efficacy beliefs for performing well depending on the specific problem, subject (e.g., algebra, biology, history), domain (e.g., verbal, quantitative), test format (e.g., multiple-choice, short answer, essay), and testing environment (e.g., timed, open-book, online, take-home).

Self-efficacy should be distinguished from outcome expectancy, which is one's expectations about the response-outcome contingency ([Bandura, 1997](#)). Whereas self-efficacy represents the conviction about one's own capabilities to execute a certain behavior, outcome expectancy represents the belief that particular behaviors will lead to a certain outcome. Both self-efficacy and outcome expectancy are required to initiate and maintain behaviors. However, even with a high outcome expectancy, people who do not believe in their own effectiveness are not likely to initiate the task or actively cope with challenges ([Bandura, 1977](#)). Self-efficacy lies at the core of human agency because it enables people to engage in the course of actions to achieve desired outcomes.

Another similar yet separable concept from self-efficacy is perceived control ([Bandura, 1997](#); [Schunk, 1991](#)). Closely related to outcome expectancy, perceived control is a causal belief regarding the pursued outcome rather than the behavior itself. Whereas self-efficacy is a belief in whether one can execute the requisite behavior, perceived control is a belief in whether one can exert control over an outcome and is therefore similar to the perceived locus of control ([Rotter, 1966](#)). Students may have strong self-efficacy for getting a good score on the upcoming test but have low perceived control over the actual grade because of the normative evaluation system. The perceived uncontrollability of the desired outcome can cause learned helplessness ([Miller and Seligman, 1975](#)). Regardless of the level of self-efficacy, people may feel helpless when they see the outcome to be independent of their own behaviors.

Self-efficacy is also different from other evaluations about self, such as self-esteem or self-concept ([Bong and Clark, 1999](#); [Bong and Skaalvik, 2003](#)). Self-esteem is an evaluation of oneself in general, while self-concept is a conception of oneself in a specific domain. Both self-esteem (e.g., "I feel proud of myself") and self-concept (e.g., "I'm good at math") are relatively stable and include emotional as well as cognitive reactions to the self and its attributes. In comparison, self-efficacy is more malleable and largely

cognitive, since it is a judgment of one's capability to perform a specific task at a designated level (e.g., "I'm highly certain that I can learn to use computers"; Bandura, 2006). Emotional reactions are considered as correlates and consequences of efficacy judgments rather than components of self-efficacy.

Although both self-concept and self-efficacy tap competence perceptions, they can be further distinguished. Self-concept is an aggregated evaluation of one's competence in a given domain, which depends heavily on past achievements and intra- and inter-individual comparisons (Bong and Skaalvik, 2003; Möller et al., 2009; Zimmerman, 2000). Compared to self-efficacy, self-concept is relatively past-oriented and normative. In contrast, self-efficacy is a situation-specific expectation of successful performance on a given task and, as such, more future-oriented than self-concept. While self-efficacy judgments also rely heavily on one's past performances and use social comparative information, the most critical question is whether individuals feel confident that they can accomplish what they set out to achieve. In this sense, self-efficacy should be viewed as a goal-referenced construct than a normative construct (Bong and Skaalvik, 2003).

Sources of self-efficacy

Self-efficacy is shaped through information coming from diverse sources. Four major sources that have been established to influence self-efficacy formation are mastery experience, vicarious experience, social persuasion, and physiological and affective states (Bandura, 1997; Usher and Pajares, 2008).

The first and the most influential source that individuals turn to when gauging their self-efficacy is mastery experience—one's past experiences of successes and failures (Bandura, 1997; Byars-Winston et al., 2017). Mastery experience consists of one's previous encounters with the same or similar tasks and provides individuals with authentic and the most reliable information for estimating their confidence. When individuals' attempts end up with success or failure, their self-efficacy toward the same or similar tasks moves upward or downward, respectively. The experience of overcoming challenges and eventually succeeding on a difficult task is powerful evidence of one's competence and noticeably boosts one's self-efficacy. Yet, it should be noted that it is individuals' own interpretations and attributions of these success and failure experiences that wield influence on subsequent self-efficacy. The same success or failure can bring about varying degrees of changes in self-efficacy beliefs across individuals, depending on their perceptions of contexts and conditions that may moderate the influence of mastery experience on efficacy appraisals (Bandura, 1997). For example, even when a student receives an A on a writing assignment, her self-efficacy for writing may not improve much if she thinks that is because the assignment was quite easy (i.e., task difficulty) or benefitted from the help from a writing advisor (i.e., external aid).

Vicarious experience serves as the second source of self-efficacy information (Bandura, 1997; Usher and Pajares, 2008). Individuals can gauge the likelihood of their success by consulting their own past performances as well as their observations of others' performance on the same task. This source is useful especially when direct experience is not available or there is no absolute barometer of competence (Bandura, 1997). For instance, even with no prior experience, observing others' success and failure on a given task could lead individuals to expect a similar level of performance when they try it themselves. The performance of others can also serve as a standard against which one's own performance is socially compared. Individuals' self-efficacy will be enhanced if they performed better than others did but diminished if they performed worse than others did. Perceived similarity of the model to oneself increases the personal relevance of the information conveyed by the model and increases the power of this self-efficacy source (Schunk, 1987). Performances of peers who share similar backgrounds and characteristics with oneself can be more informative than those of peers who do not. In a related vein, between the coping model who displays gradual improvement in performance and the mastery model who demonstrates proficient performance from the start, the coping model is generally more helpful to struggling students because of the similarity the students perceive in the model (Schunk et al., 1987; Schunk, 1987).

The third source of self-efficacy is social persuasion by significant others (Bandura, 1997; Usher and Pajares, 2008). For example, verbal and nonverbal messages from family members, teachers, and peers, which can be either supportive or destructive, convey information to learners regarding the possibility of their success in a certain domain or task (Ahn et al., 2017). Although its influence may be weaker and short-lived compared to that of direct mastery and indirect vicarious experiences, persuasions could provide an effective means to strengthen one's self-efficacy when combined with multifaceted strategies to develop actual skills (Bandura, 1997). Self-efficacy information from this source is weighted differently depending on both the perceived credibility of persuaders and the contents of persuasion. Social persuasion by individuals who are judged to be trustworthy, knowledgeable, and competent enough to deliver accurate diagnostic information about one's ability is more powerful in shaping the listener's self-efficacy than social persuasion by individuals who lack these qualities (Won et al., 2017). Social persuasion also becomes more reliable when the contents of persuasion are plausible to the listener when considering the temporal proximity of upcoming performance and the nature of a task (Bandura, 1997).

Physiological and affective states are the fourth and the least influential source of self-efficacy information (Bandura, 1997; Usher and Pajares, 2008). Individuals can sense their physical reactions (e.g., fatigue, pains) and unpleasant emotional symptoms (e.g., negative mood states, arousal, and stress reactions toward a task) and interpret them as signs of their vulnerability and lack of competence (Bandura, 1997). However, information from this source is quite ambiguous to interpret accurately and thus often biased due to various personal and environmental factors such as one's past experiences, pre-existing self-efficacy beliefs, and situational cues that affect attributions. For instance, rapid heartbeats before embarking on a task can be interpreted as excitement by

successful and confident learners but anxiety by less successful and insecure learners. Likewise, if learners engage in a task right after running, the same rapid heartbeats would be attributed to running rather than anxiety and hence may not influence beliefs of self-efficacy.

Roles of age, gender, and culture in self-efficacy formation

Research shows that the age, gender, and culture of individuals can affect their self-efficacy development (DiBenedetto and Schunk, 2018; Schunk and DiBenedetto, 2020; Schunk and Meece, 2006; Usher and Weidner, 2018). Self-efficacy tends to decrease and become more accurate as students get older (Schunk and Meece, 2006). Whereas young children are often overly confident about their ability, older students who are equipped with superior metacognitive capability form more realistic assessments of their ability by integrating self-efficacy information from diverse sources (Muenks et al., 2018). Older students are also more sensitive to social comparative information and start to have differentiated conceptions of ability (Muenks and Miele, 2017; Nicholls, 1978, 1984). To young children, those who spend greater effort are also the ones with superior ability. To older children, having to spend greater effort than others to achieve the same level of success is evidence of a lack of ability. These developmental differences suggest the importance of using age-appropriate goals and evaluation strategies for enhancing self-efficacy. While providing metacognitive guides to show progress and task mastery would be effective for young children, more deliberate approaches using norm-referenced evaluations and effort attributions may be necessary for older adolescents (cf. Amemiya and Wang, 2018). Recent evidence suggests that mastery experiences are particularly crucial in the self-efficacy formation of pre-college students (elementary, middle, high school) than that of college samples (Byars-Winston et al., 2017).

Self-efficacy development differs by gender as well. Gender differences have been identified in the mean-level of self-efficacy beliefs in several gender-typed domains such as math (Else-Quest et al., 2010; Huang, 2013; Jacobs et al., 2002). Despite the comparable and sometimes even superior levels of performance by women to those of men, ability perceptions in male-typed domains such as math, computer science, and sports remain gendered such that male students tend to express stronger self-efficacy than female students (Eccles et al., 1990; Else-Quest et al., 2010; Huang, 2013). These mean-level gender differences suggest that self-efficacy perceptions are under the heavy influence of prevalent gender stereotypes and point to the necessity of dealing with stereotype issues when designing self-efficacy interventions in gender-typed domains (e.g., Lee et al., 2021, 2020). Gender differences have also emerged in the relative weight assigned to the sources of self-efficacy information (Butz and Usher, 2015; Zeldin et al., 2008). Compared to male students, female students are more strongly influenced by vicarious information and social persuasion. Modeling and encouragement by trusted family members and teachers have proved to be particularly effective in strengthening the self-efficacy of women in male-typed domains. Capitalizing on the information from social agents, therefore, can be useful for improving female students' self-efficacy (Lee et al., 2020).

The culture in which individuals reside is closely associated with self-construal (Markus and Kitayama, 1991) and thus affects the way they shape their self-efficacy. For instance, compared to learners in Western countries who tend to overestimate their confidence, learners in Eastern countries report relatively modest levels of self-efficacy (Klassen, 2004). However, this lower level of self-efficacy by East Asian learners does not signify a correspondingly low level of performance; their self-efficacy is more accurate and predicts actual performance better than the self-efficacy of Western students. As was the case with gender, the relative weight given to self-efficacy information sources also differs across cultures (Usher and Weidner, 2018). In general, Western countries have individualistic cultures, featuring an independent view of the self, and Eastern countries have relatively more collectivistic cultures, espousing an *interdependent* view of the self (Markus and Kitayama, 1991). While personal mastery experiences are consistently the most influential source of self-efficacy information across cultures, learners from collectivistic cultures, compared to those from individualistic cultures, depend relatively more heavily upon social factors such as group mastery experiences and social support (Ahn et al., 2016; Klassen, 2004; Usher and Weidner, 2018). Further differences in the relative importance of different socializers even within similar cultures have also been witnessed (Ahn et al., 2016).

Motivational, behavioral, and achievement consequences of self-efficacy

Self-efficacy operates as a core contributor to motivation and a facilitator of learning and performance. It is not merely a "passive foreteller" but an "active producer" (Bandura, 1997, p. 38) of performance such that firm assurance or self-doubt about one's ability alters how people think of, feel about, and approach a given task, easily overriding the existing skills (Bandura, 1989; Zimmerman, 2000). To date, significant associations between self-efficacy and diverse academic outcomes have been established across a wide range of learners and domains (e.g., Multon et al., 1991; Rottinghaus et al., 2003). Among diverse self-constructs, self-efficacy is arguably the strongest predictor of subsequent motivation, engagement, and performance in the cyclical self-regulatory processes (e.g., Schunk, 1991; Zimmerman, 2000). Below we describe some of the most representative outcomes of self-efficacy in academic situations.

First, self-efficacy predicts goal pursuits. It can lead learners to set different levels of goals and adopt different types of achievement goals (Bong, 2005). Learners tend to set more challenging goals and also pursue approach-oriented achievement goals when they feel confident about their capability to perform tasks successfully. Self-efficacy predicts higher goal-settings, which in turn lead to better self-regulation and performance (e.g., Brown et al., 2011; Lee et al., 2014; Zimmerman et al., 1992). According to meta-

analyses examining the relations between self-efficacy and achievement goals (e.g., [Cellar et al., 2011](#); [Huang, 2016](#); [Senko and Dawson, 2017](#)), self-efficacy also shows moderately positive relationships ($0.29 \leq \bar{r} \leq 0.45$) with mastery-approach goals (i.e., aiming to improve competence and to master the learning tasks), which are the most adaptive form of achievement goals. Although the effect size is smaller than that observed with mastery-approach goals, self-efficacy also relates positively ($0.04 \leq \bar{r} \leq 0.30$) to performance-approach goals (i.e., aiming to demonstrate competence and to do better than others), which are usually associated with higher achievement. Self-efficacy generally demonstrates negative or non-significant relationships ($-0.18 \leq \bar{r} \leq 0.00$) with performance-avoidance goals (i.e., aiming to hide incompetence and to avoid doing worse than others), which are the most maladaptive form of achievement goals in terms of motivation and academic performance. Therefore, feeling self-efficacious helps learners to set and pursue adaptive goals when embarking on the self-regulatory process of learning and performance.

Second, self-efficacy is strongly associated with academic emotions experienced in achievement settings. While perceived physiological and affective states function as a source of self-efficacy ([Bandura, 1997](#)), self-efficacy can further evoke positive or negative emotions relevant to achievement pursuits such as interest ([Bandura and Schunk, 1981](#); [Bandura, 1982](#); [Rottinghaus et al., 2003](#)) and anxiety ([Bandura, 1988, 1989](#); [Nie et al., 2011](#)). Learners who are confident in performing a task are likely to experience feelings of interest and actively engage in the task with enjoyment; learners who doubt their capability to perform the task are likely to experience feelings of anxiety and the urge to avoid the task. Strong self-efficacy increases interest through a sense of causal agency in bringing about the desired outcome and experiencing anticipated satisfaction from the success ([Bandura, 1997](#)); weak self-efficacy induces high levels of anxiety and stress through apprehensive thoughts about their deficiencies to cope with the task demands and the negative outcome should they fail ([Bandura, 1988](#)). Other emotions may also accompany self-efficacy ([Pekrun, 2006](#); [Shao et al., 2020](#)). For example, individuals with strong self-efficacy to achieve positive outcomes or avoid negative outcomes can experience anticipatory joy or anticipatory relief, respectively, through the perceived control over the outcome. Individuals with weak self-efficacy may experience hopelessness and frustration due to the perceived ineffectiveness of their effort.

Third, self-efficacy activates motivated behaviors such as choice, engagement, and use of self-regulatory strategies, consistent with what [Bandura \(1997\)](#) originally postulated in his social cognitive theory. Along with value in the expectancy-value theory ([Wigfield and Eccles, 2000](#)), self-efficacy functions as one of the two proximal determinants of academic choice and behaviors ([Bong, 2001a](#)). Learners choose to pursue their education and career paths in the domain in which they feel competent ([Hackett and Betz, 1989](#)). Even when individuals perceive high value in a certain domain, their perceived inefficacy to succeed in that domain can deter them from entering the domain. Self-efficacy also energizes individuals to invest behavioral and cognitive effort and persist in the face of difficulties and obstacles. Numerous studies have shown the significant effect of self-efficacy beliefs on academic engagement and persistence (e.g., [Linnenbrink and Pintrich, 2003](#); [Komarraju and Nadler, 2013](#); [Schunk, 1989](#)), with moderate effect sizes ($0.17 \leq \bar{r} \leq 0.48$) between self-efficacy and persistence in meta-analyses ([Multon et al., 1991](#); [Robbins et al., 2004](#)).

Regarding strategy use, self-efficacy positively predicts the use of adaptive strategies and negatively predicts that of maladaptive self-regulatory strategies. For example, learners with high self-efficacy tend to engage more in adaptive help-seeking and less in avoidant help-seeking because they worry less about showing their lack of ability by asking for help from others ([Ryan and Shin, 2011](#)). Besides, learners with lower self-efficacy are more likely to expect the worst outcomes in achievement contexts, making excuses for their potential or actual failures, cheating on academic tasks and tests to appear able, and putting off important academic tasks with no good reason ([Lee et al., 2014](#)). Meta-analyses have shown negative correlations of medium size ($-0.44 \leq \bar{r} \leq -0.38$) between self-efficacy and procrastination, indicating that people often procrastinate simply because they do not have the confidence to carry out the task. Self-efficacy enhancement can provide one effective way of treating this maladaptive procrastination ([Steel, 2007](#); [van Eerde, 2003](#)).

Fourth and most important, self-efficacy is a powerful predictor of eventual performance. Learners who firmly believe that they can succeed at a designated task perform better than those who do not ([Bandura, 1997](#)). As “a generative system of human competence” (p. 37), self-efficacy enables individuals to recruit and integrate the necessary subskills to perform a given task effectively. As reviewed earlier, it affects important cognitive, emotional, and behavioral outcomes that in turn contribute to higher achievement. Out of 50 antecedents that include cognitive ability and prior performance measures, self-efficacy has emerged as the strongest predictor of college students’ achievement ([Richardson et al., 2012](#)). Meta-analyses of studies on the association between self-efficacy beliefs and performance have covered extensive breadths and types of achievement, including academic achievement ([Honicke and Broadbent, 2016](#); [Multon et al., 1991](#); [Richardson et al., 2012](#); [Robbins et al., 2004](#); [Stajkovic et al., 2018](#)), memory performance ([Beaudoin and Desrichard, 2011](#)), language proficiency ([Wang and Sun, 2020](#)), creativity ([Haase et al., 2018](#)), sports performance ([Moritz et al., 2000](#)), and work-related learning ([Sitzmann and Ely, 2011](#)) and performance ([Brown et al., 2011](#); [Stajkovic and Luthans, 1998](#)). All meta-analyses have shown positive relationships of small to medium effect sizes ($0.10 \leq \bar{r} \leq 0.59$) between self-efficacy and performance, supporting the exceptional predictive utility of self-efficacy.

Measuring self-efficacy

Among many distinctive qualities that self-efficacy beliefs possess, the hallmark is arguably the highly context-specific nature of this construct ([Bandura, 1997](#); [Pajares, 1996](#)). Context-specificity encompasses domain-specificity, content-specificity, or measurement specificity but is not equal to any one of them. Self-efficacy beliefs can be assessed at varying levels of specificity that are most consistent with the investigators’ explanatory and predictive goals ([Bandura, 1997](#)). Learners hold self-efficacy beliefs that are skill-specific (e.g., self-efficacy for solving two-digit addition problems), task-specific (e.g., self-efficacy to perform well on the next algebra quiz),

or domain-specific (e.g., math self-efficacy). Academic self-efficacy is frequently assessed at the skill-specific level not because it is useful only at those minute levels but because researchers are often interested in predicting students' impending performance on those particular classes of tasks (Pajares, 1996). Depending on what it is that the researcher is most interested in forecasting, self-efficacy may be assessed at increasingly more general or increasingly more specific levels (Bandura, 1997; Bong, 2001b; Bong and Hocevar, 2002).

When self-efficacy items are developed to predict some concrete outcome, the assessment must emulate the content, scope, and difficulty levels of the prediction target to a sufficient degree. Otherwise, the accuracy of prediction suffers. For instance, Pajares and Miller (1995) asked a group of college students to report their self-efficacy for solving a set of specific math problems, completing everyday math tasks, or performing successfully in math-related courses. Students' math problem-solving performance and their choice of math-related majors comprised the two prediction outcomes. Although the three self-efficacy scores correlated positively with one another, students' self-efficacy ratings made against the specific math problems were the strongest predictor of their performance on the same set of math problems; students' self-efficacy ratings toward the math courses were the strongest predictor of their choice of math-related majors. The predictive utility of self-efficacy beliefs is maximized when they are estimated in reference to the tasks and contexts that best correspond to the criterial variable (Bandura, 1997; Bong, 2002; Pajares, 1996).

The key in self-efficacy assessment, therefore, is to capture the competence belief that individuals recruit as they enter the performance situations (Pajares, 1996). To measure self-efficacy in a context-specific fashion, important features of the tasks and achievement settings that could affect the quality of outcomes should be represented in the respective items and scales (Bandura, 1997). Successful performance at a broader level (e.g., writing an essay) requires more than a simple sum of specific requisite skills (e.g., spelling, grammar, punctuation, development of topic sentences, organization of ideas) and accompanies a unique set of situational affordances and constraints (e.g., themes, timeline, word limit, individual vs. collaborative work). Even when individuals feel highly confident about their capability to successfully perform each of the subskills, they could nonetheless remain skeptical about their ability to effectively execute the ultimate performance because doing so depends on the dexterous integration of all required subskills. For this reason, the correlation coefficient between any two self-efficacy measures decreases as the difference in the levels of specificity associated with these beliefs increases, even within a single domain (Bong and Hocevar, 2002; Bong, 2001b). Not only are individuals able to discriminate between their self-efficacy beliefs across multiple levels of assessment specificity, but they also rely on different cues and apply distinct self-evaluative schemes when appraising their sense of efficacy at different levels. To conserve the context-specific and generative nature of self-efficacy judgments, the complexities of performance situations should be closely mirrored in the corresponding assessments.

Promoting self-efficacy of learners

As self-efficacy plays a central role in motivating learners and actualizing their potentials, strategies to improve learners' academic confidence in their capability to learn and perform well have received much attention. In particular, various attempts have been made to promote learners' self-efficacy by manipulating the sources of self-efficacy information. Self-efficacy interventions could be classified into those involving goal setting; modeling and strategy training; evaluations and reward structures; and attributional feedback (Lee et al., 2020). We introduce several self-efficacy enhancing strategies with demonstrated utility within this categorization.

First, setting certain types of goals is beneficial for bolstering self-efficacy. Specific and proximal goals are more effective than general and distal goals because they appear relatively easily attainable. Specific goals also function as clear standards for evaluating performance, making progress and mastery along the way visible to learners (Schunk, 1990). Bandura and Schunk (1981) have demonstrated the benefits of having a specific goal with proximal subgoals in promoting self-efficacy toward solving math problems for children who initially suffered from poor arithmetic performance and a lack of confidence (see also Schunk, 1983a). The children displayed not only stronger self-efficacy but also enhanced interest and problem-solving performance. Self-set goals can also promote self-efficacy. When learners set the goal for themselves, they tend to establish a more feasible goal, consider it more achievable, and commit to and engage in it more strongly, which can further enhance their self-efficacy (Schunk, 1985). Goals that are attainable yet slightly challenging can lead to greater mastery experiences once they are accomplished. Scaffolding can contribute to self-efficacy enhancement by helping learners experience success in achieving challenging goals (Sun and Hsu, 2019). Trying to reach goals that are specific, proximal, and realistic yet demanding helps learners appreciate their progress and build positive beliefs in their capability.

Second, modeling and strategy training can boost self-efficacy by teaching learners how to succeed on a task. When learners lack requisite skills and knowledge, efficacy-building strategies such as social persuasion that do not involve actual improvement in their competence would prove futile. Learners will become confident when they receive specific instructions on concrete learning and self-regulatory strategies that work. Schunk and Cox (1986) illustrated how training children with learning difficulties in the strategy of explicit verbalization increased their self-efficacy and math performance. Combining modeling with strategy training can be even more effective because it shows how to use the strategies and simultaneously conveys the indirect message to learners that they can do it too (Schunk, 1986). Models that are perceived as similar to themselves can amplify the positive influence of this vicarious message. Schunk and Hanson (1985) compared the effects of different types of models on the improvement in math problem-solving self-efficacy and skills among children who initially exhibited low performance. Observing a peer model resulted in the greatest gain in their self-efficacy and performance. Observing a teacher model was the next most effective, which was, in turn,

more effective than not observing any model. These results corroborate the importance of model similarity as a partial determinant of the modeling effect. Direct teaching and modeling of effective strategies are thus one of the surest ways to develop a stable sense of confidence among learners.

Third, types of evaluations and reward structures are other important considerations when trying to raise learners' self-efficacy. Criterion-referenced and self-referenced evaluations offer a clear advantage over norm-referenced evaluations. The standard in norm-referenced evaluations is the performance of others (Popham and Husek, 1969). Such a standard is unpredictable, not consistently tied to one's performance, and perceived as out of one's control, making learners frustrated and disoriented particularly when they fail. In contrast, the standard in criterion- and self-referenced evaluations is unambiguous, directly tied to one's performance, and hence within one's control. In criterion-referenced evaluations, learners understand what objectives to meet and to what extent; in self-referenced evaluations, learners strive to perform better than their previous performance. Both types of evaluations provide learners with information on their current levels of mastery and progress so they can focus on improving the needed skills and competence. The two evaluations can thus function as progress feedback, communicating to learners the degree of their mastery and improvement that directly feeds into their perceptions of self-efficacy (Bandura, 1997; Schunk, 1982). The relative effectiveness of a criterion-based, self-referenced evaluation to a norm-referenced evaluation has been observed reliably (e.g., Chan and Lam, 2010; Shih and Alexander, 2000). The manner with which rewards are given can also make a difference in self-efficacy development. Rewards are most effective when they are tied to one's actual performance, thus signifying the mastery or progress on a task (Cameron et al., 2005; Schunk, 1989). When the effects of rewards for performance, rewards for participation, and unexpected rewards were compared, those for performance resulted in the largest increases in the engagement, actual skills, and self-efficacy of learners (Schunk, 1983b).

Last, healthy attributional feedback improves learners' self-efficacy. Aside from direct evaluation, teachers and parents often provide learners with various attributional feedback for their achievement outcomes, which conveys efficacy information. According to attribution theory (Weiner, 1979, 1985), success and failure can be attributed to diverse reasons, most notably ability, effort, difficulty, or luck. Attributing their academic outcomes to external factors such as task difficulty and luck would not affect learners' self-efficacy as much as would attributing the same outcomes to internal factors. By comparison, making success and failure attributions to ability, an internal factor, increases and decreases learners' self-efficacy, respectively, because ability is often considered stable in nature and hence expected to wield a continued influence on future outcomes.

Attributing failures to a lack of effort instead of a lack of ability can help maintain learners' sense of efficacy. Yet, caution is required when encouraging effort attributions for academic outcomes. When past accomplishments are attributed to effort, suggesting that learners have done well through their hard work, it can increase not only their self-efficacy but also their engagement in subsequent learning and actual skill acquisition (Schunk and Cox, 1986; Schunk, 1981, 1982). Effort attribution also promotes a growth mindset, a belief that ability can be developed through effort and the use of effective strategies, which protects learners' motivation upon failures (Yeager and Dweck, 2012). However, when failures and difficulties are blindly attributed to effort, learners can question the validity of such attribution and start questioning their ability (Schunk, 1981). This scenario is especially likely when learners believe they have invested enough effort but still could not experience satisfactory progress or task mastery. For children who are old enough to believe that high effort compensates for low ability (Nicholls and Miller, 1984; Nicholls, 1978), blaming the lack of effort for their failures might be interpreted to mean that they do not possess the ability to succeed. The potential backlash of effort attribution can intensify in contexts where one's performance is being compared to that of others (Nicholls, 1984). Under these circumstances, attributing failures to the use of inadequate strategies rather than a lack of effort could help maintain learners' self-efficacy beliefs.

Concluding remarks

Bandura's social cognitive theory features an agentic view of human functioning. From this viewpoint, individuals are not passive creatures that are under the control of environmental forces but active organisms that exercise control over their behaviors and environments in pursuit of their goals. At the core of these self-regulatory processes lies self-efficacy. The subjective conviction that one can successfully execute the course of actions required to achieve the desired goal enables the person to regulate their own cognition, motivation, and behavior. During the past four decades since Bandura proposed the concept of self-efficacy, more than enough evidence has accumulated on the critical role of self-efficacy in diverse spheres of life. While available findings have already established self-efficacy as a key foreteller of individuals' motivation, learning, and performance, continued research on this important construct will deepen our understanding of effective human functioning even further.

References

- Ahn, H.S., Usher, E.L., Butz, A., Bong, M., 2016. Cultural differences in the understanding of modeling and feedback as sources of self-efficacy information. *Br. J. Educ. Psychol.* 86 (1), 112–136.
- Ahn, H.S., Bong, M., Kim, S., 2017. Social models in the cognitive appraisal of self-efficacy information. *Contemp. Educ. Psychol.* 48, 149–166.
- Amemiya, J., Wang, M., 2018. Why effort praise can backfire in adolescence. *Child Dev. Perspect.* 12 (3), 199–203.
- Bandura, A., Schunk, D.H., 1981. Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *J. Pers. Soc. Psychol.* 41 (3), 586–598.

- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84 (2), 191–215.
- Bandura, A., 1982. Self-efficacy mechanism in human agency. *Am. Psychol.* 37 (2), 122–147.
- Bandura, A., 1988. Self-efficacy conception of anxiety. *Anxiety. Res.* 1, 77–98.
- Bandura, A., 1989. Human agency in social cognitive theory. *Am. Psychol.* 44 (9), 1175–1184.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman, New York, NY.
- Bandura, A., 2006. Guide for constructing self-efficacy scales. In: Pajares, F., Urdan, T. (Eds.), *Self-Efficacy Beliefs of Adolescents*. Information Age, Greenwich, CT, pp. 307–337.
- Beaudoin, M., Desrichard, O., 2011. Are memory self-efficacy and memory performance related? A meta-analysis. *Psychol. Bull.* 137 (2), 211–241.
- Bong, M., Clark, R.E., 1999. Comparison between self-concept and self-efficacy in academic motivation research. *Educ. Psychol.* 34 (3), 139–153.
- Bong, M., Hocevar, D., 2002. Measuring self-efficacy: multitrait-multimethod comparison of scaling procedures. *Appl. Meas. Educ.* 15 (2), 143–171.
- Bong, M., Skaalvik, E.M., 2003. Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15 (1), 1–40.
- Bong, M., 2001a. Between- and within-domain relations of academic motivation among middle and high school students: self-efficacy, task-value, and achievement goals. *J. Educ. Psychol.* 93 (1), 23–34.
- Bong, M., 2001b. Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemp. Educ. Psychol.* 26 (4), 553–570.
- Bong, M., 2002. Predictive utility of subject-, task-, and problem-specific self-efficacy judgments for immediate and delayed academic performances. *J. Exp. Educ.* 70 (2), 133–162.
- Bong, M., 2005. Within-grade changes in Korean girls' motivation and perceptions of the learning environment across domains and achievement levels. *J. Educ. Psychol.* 97 (4), 656–672.
- Brown, S.D., Lent, R.W., Telander, K., Tramayne, S., 2011. Social cognitive career theory, conscientiousness, and work performance: a meta-analytic path analysis. *J. Vocat. Behav.* 79, 81–90.
- Butz, A.R., Usher, E.L., 2015. Salient sources of early adolescents' self-efficacy in two domains. *Contemp. Educ. Psychol.* 42, 49–61.
- Byars-Winston, A., Diestelmann, J., Savoy, J.N., Hoyt, W.T., 2017. Unique effects and moderators of effects of sources on self-efficacy: a model-based meta-analysis. *J. Counsel. Psychol.* 64 (6), 645–658.
- Cameron, J., Pierce, W.D., Banko, K.M., Gear, A., 2005. Achievement-based rewards and intrinsic motivation: a test of cognitive mediators. *J. Educ. Psychol.* 97 (4), 641–655.
- Cellar, D.F., Stuhlmacher, A.F., Young, S.K., Fisher, D.M., Adair, C.K., Haynes, S., Twichell, E., Arnold, K.A., Royer, K., Denning, B.L., Riester, D., 2011. Trait goal orientation, self-regulation, and performance: a meta-analysis. *J. Bus. Psychol.* 26, 467–483.
- Chan, J.C.Y., Lam, S., 2010. Effects of different evaluative feedback on students' self-efficacy in learning. *Instr. Sci.* 38, 37–58.
- DiBenedetto, M.K., Schunk, D.H., 2018. Self-efficacy in education revisited through a sociocultural lens. In: Liem, G.A.D., McInerney, D.M. (Eds.), *Big Theories Revisited 2*. Information age, Charlotte, NC, pp. 117–139.
- Eccles, J.S., Jacobs, J.E., Harold, R.D., 1990. Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *J. Soc. Issues* 46 (2), 183–201.
- Else-Quest, N.M., Hyde, J.S., Linn, M.C., 2010. Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136 (1), 103–127.
- Haase, J., Hoff, E.V., Hanel, P.H.P., Innes-Ker, Å., 2018. A meta-analysis of the relation between creative self-efficacy and different creativity measurements. *Creativ. Res. J.* 30 (1), 1–16.
- Hackett, G., Betz, N.E., 1989. An exploration of the mathematics self-efficacy/mathematics performance correspondence. *J. Res. Math. Educ.* 20 (3), 261–273.
- Honicke, T., Broadbent, J., 2016. The influence of academic self-efficacy on academic performance: a systematic review. *Educ. Res. Rev.* 17, 63–84.
- Huang, C., 2013. Gender differences in academic self-efficacy: a meta-analysis. *Eur. J. Psychol. Educ.* 28, 1–35.
- Huang, C., 2016. Achievement goals and self-efficacy: a meta-analysis. *Educ. Res. Rev.* 19, 119–137.
- Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S., Wigfield, A., 2002. Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73 (2), 509–527.
- Klassen, R.M., 2004. Optimism and realism: a review of self-efficacy from a cross-cultural perspective. *Int. J. Psychol.* 39 (3), 205–230.
- Komarraju, M., Nadler, D., 2013. Self-efficacy and academic achievement: why do implicit beliefs, goals, and effort regulation matter? *Learn. Indiv. Differ.* 25, 67–72.
- Lee, J., Bong, M., Kim, S., 2014. Interaction between task values and self-efficacy on maladaptive achievement strategy use. *Educ. Psychol.* 34 (5), 538–560.
- Lee, W., Lee, M.-J., Bong, M., 2014. Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemp. Educ. Psychol.* 39, 86–99.
- Lee, H.J., Shin, D.D., Bong, M., 2020. Strengthening adolescents' confidence to learn. In: Liem, G.A.D., McInerney, D.M. (Eds.), *Promoting Motivation and Learning in Contexts: Sociocultural Perspectives on Educational Interventions*. Information age, Charlotte, NC, pp. 37–61.
- Lee, J., Lee, H.J., Song, J., Bong, M., 2021. Enhancing children's math motivation with a joint intervention on mindset and gender stereotypes. *Learn. Instr.* 73, 101416.
- Linnenbrink, E.A., Pintrich, P.R., 2003. The role of self-efficacy beliefs in student engagement and learning in the classroom. *Read. Writ. Q.* 19, 119–137.
- Markus, H.R., Kitayama, S., 1991. Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98 (2), 224–253.
- Miller, W.R., Seligman, M.E.P., 1975. Depression and learned helplessness in man. *J. Abnorm. Psychol.* 84 (3), 228–238.
- Möller, J., Pohlmann, B., Köller, O., Marsh, H.W., 2009. A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Rev. Educ. Res.* 79 (3), 1129–1167.
- Moritz, S.E., Feltz, D.L., Fahrbach, K.R., Mack, D.E., 2000. The relation of self-efficacy measures to sport performance: a meta-analytic review. *Res. Q. Exerc. Sport* 71 (3), 280–294.
- Muenks, K., Miele, D.B., 2017. Students' thinking about effort and ability: the role of developmental, contextual, and individual difference factors. *Rev. Educ. Res.* 87 (4), 707–735.
- Muenks, K., Wigfield, A., Eccles, J.S., 2018. I can do this! The development and calibration of children's expectations for success and competence beliefs. *Dev. Rev.* 48, 24–39.
- Multon, K.D., Brown, S.D., Lent, R.W., 1991. Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *J. Counsel. Psychol.* 38 (1), 30–38.
- Nicholls, J.G., Miller, A.T., 1984. Reasoning about the ability of self and others: a developmental study. *Child Dev.* 55 (6), 1990–1999.
- Nicholls, J.G., 1978. The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Dev.* 49, 800–814.
- Nicholls, J.G., 1984. Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychol. Rev.* 91 (3), 328–346.
- Nie, Y., Lau, S., Liao, A.K., 2011. Role of academic self-efficacy in moderating the relation between task importance and test anxiety. *Learn. Indiv. Differ.* 21, 736–741.
- Pajares, F., Miller, M.D., 1995. Mathematics self-efficacy and mathematics performances: the need for specificity of assessment. *J. Counsel. Psychol.* 42, 190–198.
- Pajares, F., 1996. Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578.
- Pekrun, R., 2006. The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341.
- Popham, W.J., Husek, T.R., 1969. Implications of criterion-referenced measurement. *J. Educ. Meas.* 6 (1), 1–9.
- Richardson, M., Abraham, C., Bond, R., 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138 (2), 353–387.
- Robbins, S.B., Lauver, K., Le, H., Davis, D., Langley, R., Carlstrom, A., 2004. Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol. Bull.* 130 (2), 261–288.
- Rotter, J.B., 1966. Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.* 80 (1), 1–28.
- Rottinghaus, P.J., Larson, L.M., Borgen, F.H., 2003. The relation of self-efficacy and interests: a meta-analysis of 60 samples. *J. Vocat. Behav.* 62, 221–236.

- Ryan, A.M., Shin, H., 2011. Help-seeking tendencies during early adolescence: an examination of motivational correlates and consequences for achievement. *Learn. Instr.* 21, 247–256.
- Schunk, D.H., Cox, P.D., 1986. Strategy training and attributional feedback with learning disabled students. *J. Educ. Psychol.* 78 (3), 201–209.
- Schunk, D.H., DiBenedetto, M.K., 2020. Motivation and social cognitive theory. *Contemp. Educ. Psychol.* 60, 101832.
- Schunk, D.H., Hanson, A.R., 1985. Peer models: influence on children's self-efficacy and achievement. *J. Educ. Psychol.* 77 (3), 313–322.
- Schunk, D.H., Meece, J.L., 2006. Self-efficacy development in adolescences. In: Pajares, F., Urdan, T. (Eds.), *Self-Efficacy Beliefs of Adolescents*. Information Age, Greenwich, CT, pp. 71–96.
- Schunk, D.H., Hanson, A.R., Cox, P.D., 1987. Peer-model attributes and children's achievement behaviors. *J. Educ. Psychol.* 79 (1), 54–61.
- Schunk, D.H., 1981. Modeling and attributional effects on children's achievement: a self-efficacy analysis. *J. Educ. Psychol.* 73 (1), 93–105.
- Schunk, D.H., 1982. Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *J. Educ. Psychol.* 74 (4), 548–556.
- Schunk, D.H., 1983a. Developing children's self-efficacy and skills: the roles of social comparative information and goal setting. *Contemp. Educ. Psychol.* 8, 76–86.
- Schunk, D.H., 1983b. Reward contingencies and the development of children's skills and self-efficacy. *J. Educ. Psychol.* 75 (4), 511–518.
- Schunk, D.H., 1985. Participation in goal setting: effects on self-efficacy and skills of learning-disabled children. *J. Spec. Educ.* 19 (3), 307–317.
- Schunk, D.H., 1986. Vicarious influences on self-efficacy for cognitive skill learning. *J. Soc. Clin. Psychol.* 4 (3), 316–327.
- Schunk, D.H., 1987. Peer models and children's behavioral change. *Rev. Educ. Res.* 57 (2), 149–174.
- Schunk, D.H., 1989. Self-efficacy and achievement behaviors. *Educ. Psychol. Rev.* 1 (3), 173–208.
- Schunk, D.H., 1990. Goal setting and self-efficacy during self-regulated learning. *Educ. Psychol.* 25 (1), 71–86.
- Schunk, D.H., 1991. Self-efficacy and academic motivation. *Educ. Psychol.* 26 (3–4), 207–231.
- Senko, C., Dawson, B., 2017. Performance-approach goal effects depend on how they are defined: meta-analytic evidence from multiple educational outcomes. *J. Educ. Psychol.* 109 (4), 574–598.
- Shao, K., Pekrun, R., Marsh, H.W., Loderer, K., 2020. Control-value appraisals, achievement emotions, and foreign language performance: a latent interaction analysis. *Learn. Instr.* 69, 101356.
- Shih, S.S., Alexander, J.M., 2000. Interacting effects of goal setting and self- or other-referenced feedback on children's development of self-efficacy and cognitive skill within the Taiwanese classroom. *J. Educ. Psychol.* 92 (3), 536–543.
- Sitzmann, T., Ely, K., 2011. A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychol. Bull.* 137 (3), 421–442.
- Stajkovic, A.D., Luthans, F., 1998. Self-efficacy and work-related performance: a meta-analysis. *Psychol. Bull.* 124 (2), 240–261.
- Stajkovic, A.D., Bandura, A., Locke, E.A., Lee, D., Sergeant, K., 2018. Test of three conceptual models of influence of the big five personality traits and self-efficacy on academic performance: a meta-analytic path-analysis. *Pers. Indiv. Differ.* 120, 238–245.
- Steel, P., 2007. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol. Bull.* 133 (1), 65–94.
- Sun, J.C.Y., Hsu, K.Y.C., 2019. A smart eye-tracking feedback scaffolding approach to improving students' learning self-efficacy and performance in a C programming course. *Comput. Hum. Behav.* 95, 66–72.
- Usher, E.L., Pajares, F., 2008. Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78 (4), 751–796.
- Usher, E.L., Weidner, B.L., 2018. Sociocultural influences on self-efficacy development. In: Liem, G.A.D., McInerney, D.M. (Eds.), *Big Theories Revisited 2*. Information age, Charlotte, NC, pp. 141–164.
- van Eerde, W., 2003. A meta-analytically derived nomological network of procrastination. *Pers. Indiv. Differ.* 35 (6), 1401–1418.
- Wang, C., Sun, T., 2020. Relationship between self-efficacy and language proficiency: a meta-analysis. *System* 95, 102366.
- Weiner, B., 1979. A theory of motivation for some classroom experiences. *J. Educ. Psychol.* 71 (1), 3–25.
- Weiner, B., 1985. An attributional theory of achievement motivation and emotion. *Psychol. Rev.* 92 (4), 548–573.
- Wigfield, A., Eccles, J.S., 2000. Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81.
- Won, S., Lee, S., Bong, M., 2017. Social persuasions by teachers as a source of student self-efficacy: the moderating role of perceived teacher credibility. *J. Adolesc.* 54 (5), 532–547.
- Yeager, D.S., Dweck, C.S., 2012. Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47 (4), 302–314.
- Zeldin, A.L., Britner, S.L., Pajares, F., 2008. A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *J. Res. Sci. Teach.* 45 (9), 1036–1058.
- Zimmerman, B.J., Bandura, A., Martinez-Pons, M., 1992. Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *Am. Educ. Res. J.* 29 (3), 663–676.
- Zimmerman, B.J., 2000. Self-efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25, 82–91.

Contemporary contexts for learning: an overview

Anastasia O. Tzirides, Bill Cope, and Mary Kalantzis, University of Illinois Urbana-Champaign, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

On the contexts of learning	258
Learning in formal and informal settings	258
The social functions of education in context	259
Learner diversity as a major contemporary contextual factor	260
Digital media as a major contemporary contextual factor	262
Addressing changing contexts in formal education	264
References	265

On the contexts of learning

This article views the context of learning through several different lenses. The first is the general distinction between learning that happens in the informal context of the lifeworld contrasted with the formal or institutionalized learning of programs, curricula and syllabi—called “education.” The second lens is the social functions of both informal and formal learning, focusing on the role of education for working life, personal life and public life. Cross-cutting these three contexts, we discuss the increasing significance today of social diversity and digital tools for meaning-making. In the section that follows, we explore the dimensions of learner diversity, and the connections between lifeworld differences and the learning that happens in formal education. A fourth area of focus is the changes that have occurred in recent decades in the media applied to education, with particular attention to the rise of e-learning ecologies. In a final section, we move in the other direction, from education back into its context to address the question with, how and in what ways can learning be a force for educational justice and social change?

Learning in formal and informal settings

“Context of learning” refers to the environment in which learning occurs. Education is a process and profession that must be understood by the influences of context. Its work is oriented to that context. It serves purposes located beyond itself, in context.

In this article, we reflect upon contexts of education as we move into the middle decades of the 21st century, when formal education seems more important than ever before. Politicians talk about the knowledge society; business people discuss the new economy; and parents worry about the role a good education will play in their children’s futures. Despite this, educators struggle with the resources available to them as they attempt to meet increasing expectations of the social context in which they work.

Going back to the fundamentals, what is this process, “education,” that is expected to do so much? In its most visible manifestation, it consists of institutions: schools, colleges and universities. Education also occurs in workplaces, museums, communities, outdoors and a myriad of other places, both formally and informally. Broadly conceived, education is a social process, a relationship of teaching and learning, where one person supports another to learn, and to learn how to learn.

Learning refers to the process of familiarizing with and understanding new things. This can include knowing about a specific idea or object and knowing how to act under particular circumstances. Humans are in their nature highly capable of learning at all stages in their lives. They constantly learn in an informal way simply by immersion in their lifeworld. This happens through the rich surroundings and the people (peers, older children, adults) who have an innate social inclination to teach those around them in an informal way, as well as to learn from them (Kalantzis and Cope, 2012).

In the formal domain of learning, since the rise of mass-institutional education the focus of education has primarily been cognitive, and more specifically the ability to retain empirical knowledge and application of routines for the application of mental procedure in long-term memory (Kirschner et al., 2006). However, by focusing on the mind of the individual, we are disregarding the knowledge that exists in the social environment. John B. Watson, the father of behaviorism, had suggested to bring the focus of learning on the activity that occurs in the context of an environment (Cope and Kalantzis, 2020b). Nevertheless, behaviorism has been the focus of much criticism for the mechanical and disciplinary character of its core stimulus-response notion.

Formal education consists of the knowledge acquired through an educational institution (from preschool to graduate school), which is organized, structured and goal oriented (Kalantzis and Cope, 2012; Raikou and Karalis, 2010; Schugurensky, 2000; Melnic and Botez, 2014). Formal education follows a concrete curriculum, with specific objectives, content and methodology (Melnic and Botez, 2014). Its learning is systematic and deliberate (Kalantzis and Cope, 2012) and usually results in the acquisition of quantifiable credit, a certificate or a diploma.

However, much of our knowledge has been acquired by interacting with people other than our formal instructors and outside of school. Informal learning is not restricted to a specific place, time or age. It happens in our everyday life, at home, with our family, at work, at a library or a market, when we are with our friends. Most of the time we do not even realize that we are learning (Kalantzis

and Cope, 2012). Informal learning does not follow a specific curriculum and it is usually characterized as an experiential learning process (Strimel et al., 2014). It can happen individually or when being part of a group (Schugurensky, 2000). Some examples of informal learning activities are: visiting a museum or an exhibit, listening to the radio or watching TV, reading magazines and newspapers, using the social media (Facebook, Twitter, Instagram, etc.), attending conferences, traveling, meeting new people, living in a foreign country, etc. Informal learning is usually spontaneous and incidental (Melnic and Botez, 2014). It addresses the world around us by immersion in the experiences that are literally in front of our eyes (Kalantzis and Cope, 2012) (Fig. 1).

Schugurensky (2000) categorizes informal learning into three types, based on forms of intentionality and awareness. The first type is self-directed learning, which happens when individuals learn something intentionally, but without the presence of an instructor, and they are conscious of their learning result. The second type is incidental learning, and it happens when an individual learns something from an experience they had, without doing it on purpose, but by realizing in the end that they had acquired some new knowledge. The third category is socialization, which occurs through interaction with other people in our everyday life, without doing it on purpose and without understanding that we are learning something. Sometimes, we may become aware of such social learning later in life, when thinking of our previous acts and our beliefs.

The social functions of education in context

Formal learning or education prepares people for productive working lives. Education also has an important role to play in the development of responsible citizenship. Together, these two foundational functions link education to its context by developing learners' capacities for social participation.

To take the first of these contextual objectives, work has changed enormously since the introduction of mass-institutionalized education in the 19th century. These changes have been characterized by transitions from "Fordism," "post-Fordism" to the "information economy." In transition, the nature of work has changed: technology, management, workers' education and skills, and markets and society. Workers of today need to have acquired quite different skills and sensibilities from those of the past (Kalantzis and Cope, 2012).

In the area of citizenship, the modern state takes considerable responsibility for education, much of which is public or at least partly supported by the state. Education plays an important role in shaping certain kinds of citizens. The nature of citizenship is changing just as dramatically as that of work over the past century. Modern, mass, institutionalized schooling emerged in an era characterized by nationalism in the public sphere. Large and strong states established command relationships with their citizens. These states promoted the idea of a homogeneous community with clear borders. Schools played an important role in this process as they became one of the primary places for the socialization of children—using didactic teaching and taking an exclusionary or assimilationist approach to learners' differences.

Since the last quarter of 20th century, we have entered a phase referred to as "neoliberalism," during which the size of the state has been shrunk and an ethics of the free market and self-reliance revived from its 19th-century roots. This has translated into a partial withdrawal of the state from education and increasing pressures for the privatization of education. We have also reached the point of a global environmental and climate crisis which is becoming harder to ignore.

"Civic pluralism" is a strategically optimistic reading of contemporary trends. It describes and prescribes a system of distributed governance and responsible citizenship—in workplaces, community organizations and schools, for instance. Civic pluralism involves continually negotiating differences and problem solving, frequently supported these days by digital media. For this social context, education for civics fosters capacities to participate based on an ethic of self-governance and participation. It is environmentally aware, striving for sustainable solutions to the challenge of meeting human needs. Today, an important new set of capacities is captured by the notion of "digital citizenship," covering everything from participation in classrooms and self-organizing online communities, to digital democracy more widely conceived as a formal system of social governance (Kalantzis and Cope, 2012).

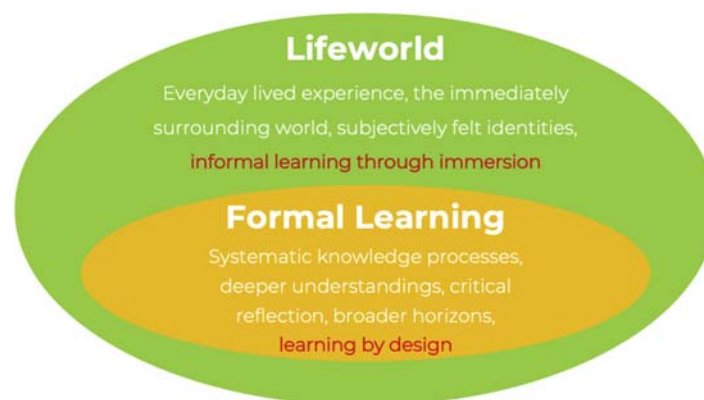


Fig. 1 The lifeworld context of learning.

The challenges of these changes in social context require that schools broaden and deepen their educational objectives. Among the touchstones for the next generation of educational environments are pedagogies that inspire critical, creative and design thinking. In such environments educators will strive to achieve equality without demanding sameness. These are ambitious goals, but the collaborative and intelligent tools of our era of e-learning make these large educational aspirations at last practicable.

Among the two most important social transformations of recent times have been the increasing recognition of lifeworld diversity and the rise of digitally-mediated communications (Fig. 2). These have had substantial impacts upon all aspects of working, personal and public life.

In the following sections of this article, we are going to focus on these enormously important, complex and frequently vexed dimensions of the contemporary educational context: learner diversity and digital technologies.

Learner diversity as a major contemporary contextual factor

In its earlier phases, mass-institutionalized schooling attempted to avoid, ignore or over-write one of the key aspects of its context: learner diversity. This encompasses: material conditions (social class, locale and family); corporeal attributes (age, race, sex/sexuality, physical and mental abilities); and symbolic representations (language, ethnos, communities of commitment). These factors shape the learner's personality. They also have an enormous impact on their engagement with learning and their educational and social outcomes.

Historically, education has managed its context of human diversity by exclusion (e.g., separate schools), assimilation (e.g., one-size-fits-all, sink-or-swim education), or more recently and aspirationally, practices of inclusion. Inclusive educators will negotiate the uniqueness of every learner's lifeworld context, and the fluidity of life in a time of intense social change. Inclusive schools ensure that their institutional organization and pedagogical practices make it possible for all members of the learning community to belong and achieve, in their difference.

Each of these dimensions of diversity has over the past decade been in a process of transformation, both of our understandings and the way in which schools and society negotiate these differences (Kalantzis and Cope, 2016b). Age differences are being transformed as the nature of childhood, adolescence, adulthood, and aging are shifting, and education is expanding its scope across the lifespan. "Race" is coming to prominence with increasing consciousness of the impacts of racism on educational opportunities. Sex and sexuality are recognized to be non-binary, with a large range of variation across LGBTQ, transgender/cisgender, and other alternatives, raising issues that schools have traditionally avoided or ignored. Physical and mental abilities are no longer considered straightforwardly as dis-abilities, but as embodied potentials and neurodiversity. In the era of mainstreaming, educators need to ensure equitable access by means of differentiated strategies. In the area of culture, students increasingly live their everyday lives in affinity communities, whose distinctiveness is frequently magnified by social media. Racism and sexism take new forms in the era of cyberbullying and pervasive microaggressions. Language differences and pedagogies are addressed in new ways with ubiquitous translation tools. First Nations' perspectives also need to be an integral part of contemporary education, including Indigenous pedagogies that contain insights of value to all learners in times of ecological crisis and social uncertainty. Meanwhile, social and regional inequalities are becoming more pronounced, complicated by the intersection of inequality with the other dimensions of diversity.

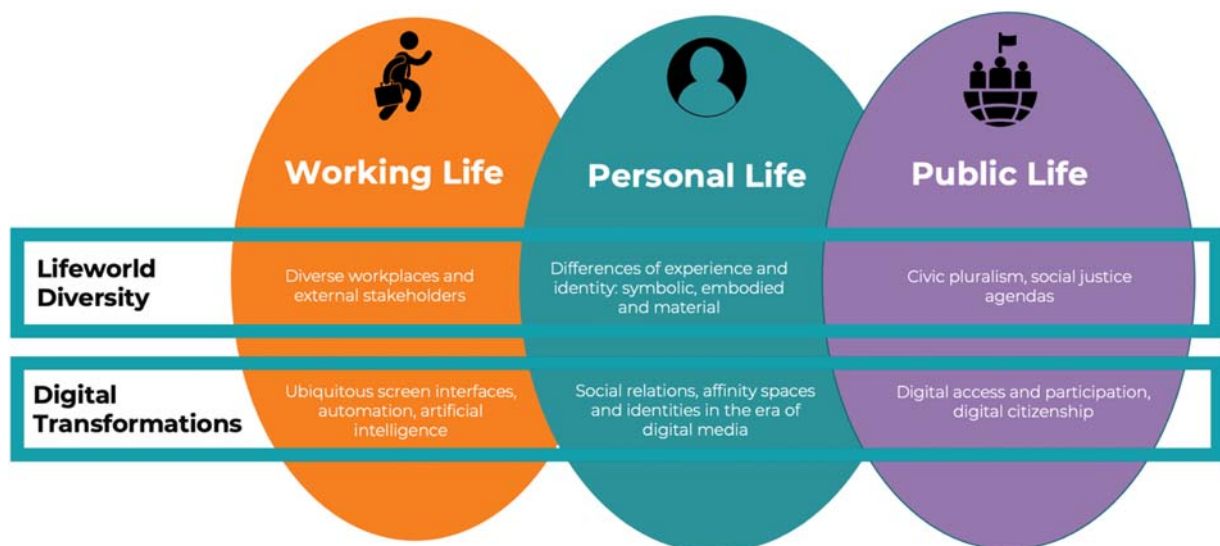


Fig. 2 Contemporary transformations in the contexts of learning.

“Diversity” is a neutral-sounding word, but any such neutrality hides a multitude of evils. Diversity—or in the case of education, the lifeworld differences that learners bring to school and then encounter in their schooling—is riven with inequalities. After decades of incremental redress through social programs of redistribution, by the early 1990s in almost every country of the world material inequalities had begun to expand again, returning us by the third decade of the 21st century to disparities of wealth that are as severe as those of unrestrained industrial capitalism at the end of the 19th century (Piketty, 2014; Sassen, 2014). Meanwhile, embodied and symbolic differences are as often markers of inequality as they are innocent differences, worthy of celebration. A pyramid visualization (Fig. 3) captures the contextual layers of material, embodied and symbolic differences that are always present. For every learner the whole that constitutes their identity is uniquely configured, albeit with cross-cutting historical commonalities with others. This is the “lifeworld” that the learner brings to school from their social context.

What role can education play? What impacts can it have on its context? On past experience we might, with full justification, claim that these and any other agendas of education reform can have little or no impact. What, after all, can education do to influence its social context? What is the potential of education to redress inequality?

Any such project must start by addressing what we have termed “material differences.” More egregious than any of the differences in societies that profess democracy and to offer equal opportunities (at least, if not equal outcomes), are differences in resourcing for education. In the United States, there can be a four-fold difference in the per student funding for public schools, and private schools can cost parents as much as ten times the revenue spent per student on public schools. This is without the differential access to family-funded learning opportunities, from books, to smart toys, to digital devices, to private lessons, and such like. This inevitably produces a spectrum of embodied “abilities,” measured less these days by intelligence tests than standardized tests of domains of knowledge. In the realm of symbolic capital, this translates into a series of capacities to survive the rigors of the modern world captured in the notion of “episteme.” But the differences start—prosaically, and incontrovertibly in the relation of resource inputs to educational outcomes. Along the way, the differences are refracted through specific histories of systemic discrimination such as racism and sexism, but level of education stands out above all other demographic variables as a predictor of social outcomes.

In their book, *Deaths of Despair and the Future of Capitalism*, Princeton University economist Anne Case and Nobel Prize winner Angus Deaton argue that having a bachelor’s degree or not, is a more precipitous dividing line than race or gender when measured in terms of social outcomes. Their measures are not only the 80% “earnings premium” for those with a bachelor’s degree or more, but dramatically different unemployment rates (those without bachelor’s degrees are twice as likely to be unemployed), labor force participation rates (2.5 times as likely to have dropped out of the workforce), life satisfaction (according to national surveys), health outcomes, and tragically rising adult mortality rates—what they call “deaths of despair” related to suicide, the opioid epidemic, smoking, obesity and lack of access to medical insurance and quality care (Case and Deaton, 2020).

The COVID-19 crisis magnified this disparity where, apart from age, level of education was the principal predictor of likely death—whether that was the result of underlying health conditions, or working in essential services where the luxury of shelter-in-place and work from home was not feasible (Chen et al., 2021). And while disproportionately larger numbers of Blacks and Latinx are imprisoned, by far the most significant factor predicting imprisonment is level of education. Of course, if you are Black, or Latinx, or a woman, or disabled, you are less likely to have a bachelor’s degree, and so histories of institutional racism, sexism and exclusion factor into the underlying educational injustice. However, the many whites without bachelor’s degrees do about as badly on these measures of social outcome, at times worse even. In a society and economy that today places a premium on epistemic capital, education trumps other demographics as a variable that will predict the most fundamental of social outcomes.

Speaking to the US context, Blacks, Latinx, women, and people with disabilities can do well if they have higher educational qualifications. Whites, men, and people without disabilities often do badly when their education ends at school. We must join the rage

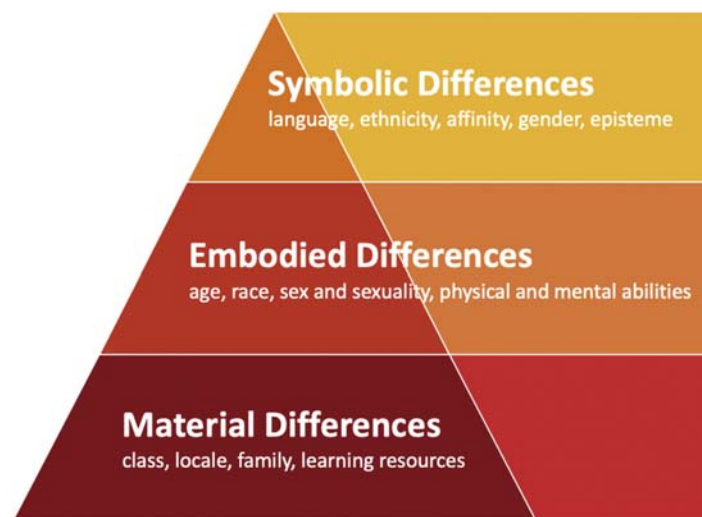


Fig. 3 “Lifeworld” as the social context of education.

at egregious practices of racial injustice that motivate the Black Lives Matter movement, and the oppressive sexism of unequal power exposed by #metoo. Redressing these abominations is not merely symbolic matter, as if there were easy fix of recalibrating our attitudes or nodding in recognition. Nor can these agendas remain a narrowly personalized and or neatly compartmentalized into a politics of “identity.” This is because specific injustices of embodiment and symbolic representation are always sourced in and refracted through the material injustices that they re-articulate and exacerbate. In the inseparable unity of our lifeworlds, the layers can’t be peeled apart, nor can their circumstantial intersections straightforwardly be identified.

Digital media as a major contemporary contextual factor

Educators today are working in a time of great change, where the contexts of learning and conditions of social interaction are changing radically. One major factor that redefines the relation of learning to its contexts is the development of digital technologies. These have already had a significant impact on the way informal learning occurs. What will be their likely impacts on formal learning?

The current era is characterized by the abundance of information and the easy access we can have to it through the Internet. This contextual change—the ubiquity of digital media—is a major key contextual factor that is today transforming education. Indeed, applied to education, digital technologies point to perhaps the largest shift in the systems of modern education since the 19th century. Among the major dimensions of change are: configurations of space, learner-to-teacher and learner-to-learner relationships, the textual forms of knowledge to which learners are exposed, the kinds of knowledge artifacts that students create, and the way the outcomes of their learning are measured.

As a consequence of ubiquitous access to information in the era of digital media, what contemporary students (of all ages) require from their education is to teach them “how to learn” and navigate an excess of information, and not “what to learn,” since the content of the information is easily accessible. Lifelong learning and skill development are necessary to succeed in a changing society, more than the ability to remember things from school or long-term memory (Cope and Kalantzis, 2020a). This challenge could be addressed in part by integrating and/or recognizing informal learning in formal educational settings. Below are some methods in which, according to Strimel et al. (2014), informal knowledge sources and process of learning can be incorporated in formal education:

- **Use of technology:** Smartboards, smartphones, tablets and iPads, in classrooms or in the workplace can support learners to be more engaged and also can create a more personalized learning experience.
- **Social media:** Blogs, wikis, Facebook, Google, Twitter, Instagram, Skype and other media can be used by students to search information and document their learning.
- **Open educational resources (OER):** These are easily accessed and free of charge resources that can assist individuals to advance their skills and knowledge.
- **MOOCs (Massive Open Online Courses):** Open and free-of-charge courses offered to anyone worldwide who is interested in an area of new knowledge or learning a new skill.
- **Formalized coaching:** This process includes using live case studies in the classroom or workplace, during which the teacher can meet personally with each student.
- **Electronic portfolios:** This tool can be used by students to demonstrate learning outcomes.
- **Work-based learning:** Includes practicing and applying skills that are closely related to workplace.

Kraemer (2008) has defined distance learning as the process that occurs when learners and educators are in a different space or time. This idea has evolved over the years considering the influx of digital technologies in learning contexts. New terms have appeared, such as online learning, e-learning, open learning, blended or hybrid learning (Blake, 2009). Oblinger and Hawkins (2005), Dublin (2003), Holmes and Gardner (2006) point out that there is not a universal way of defining e-learning. Arkorful and Abaidoo (2015) claim that e-learning refers to the facilitation of learning through the utilization of digital technologies. Therefore, the term can include distant, fully or hybrid online courses that are realized through the use of the Internet (Gotschall, 2000; Maltz and Deblois, 2005) or learning through the use of information and communication technologies as a complement to traditional, online or hybrid learning (OECD, 2005). In our definition, online learning can be considered any learning interactions supported by connected digital devices. Thus, online learning is not just a synonym for remote learning; it can also involve both in-person learning with online support and remote learning online.

In the phrase “e-learning ecologies,” the metaphor “ecology” is used for learning contexts that could be considered as ecosystems, which involve the complex interaction of human, textual, discursive and spatial dynamics (Cope and Kalantzis, 2017). Such e-learning ecologies can play a critical role in the evolution of modern education. However, digital technologies can also reproduce traditional pedagogical modes, resulting in minimal transformational impact on the substance of pedagogy. What changes are possible in the relation of education to social context as a consequence of new educational technologies?

Applied to education, digital technologies can be used as a force for reform. On the other hand, we may introduce a whole lot of technology into schools, and yet nothing changes in institutional or pedagogical senses. Educational technologies may bring to life again, and reproduce, traditional relations of learning—sometimes with justification, but other times representing a step back. Technology is pedagogically neutral. It can be a force for change, but it can also be used to reproduce old practices.

The affordances of the digital, however, can be harnessed to create learning ecologies that are more engaging for learners, more effective and resource efficient, as well as a more equitable response to learner diversity. Here are seven affordances of the digital that can be utilized in support of learning, be that in-person, blended or remote.

1. **Ubiquitous Learning** refers to learning anytime, anywhere, anyhow. This affordance connects with the concept of informal learning, which can happen in every aspect of our life, no matter the time or the location. The use of new digital and web-based media has enabled the realization of this affordance, by giving the opportunity to learners to learn without any physical or time restrictions. Especially during the COVID-19 pandemic, this approach suddenly became the new normal, as in-person contacts were limited and thus, for almost 2 years, intermittently in different places around the world, people were learning remotely.
2. **Recursive Feedback** is a version of formative assessment, or continuous feedback that is constructive to learning. Based on the concept of reflexive pedagogy, students nowadays should get feedback on their performance while they are still in the process of learning. This will help them to target any areas that need more effort, and they will all end up achieving “mastery” (Bloom, 1968), thus achieving better learning outcomes. This feedback can be in the form of peer comments, machine responses, teacher comments, as well as self-reflection (Pinheiro, 2020).
3. **Multimodal Meaning:** This affordance addresses the use of various forms of meaning, like images, videos, audio, text, diagrams, data, etc. at the same time and in various combinations to enhance the meaning of the topic addressed and create richer knowledge representations and communication. This new type of multimodal meaning making encompasses genres observed in social media, websites, apps, infographics and data visualizations. Its benefit from the educational perspective is that it prompts learners to think about the form of their text and also diversify and broaden options for knowledge representation for a broader range of learners.
4. **Active Knowledge Making:** This refers to the design of meanings and the creation of new knowledge in the learning process. The idea encompasses the process where learners are positioned as knowledge producers and meaning makers. They co-create knowledge through a scaffolded process guided by the instructor. This concept introduces a change in the balance of agency in the learning context, where students are positioned as active participants of the learning process.
5. **Collaborative Intelligence:** This affordance concerns the process where learners draw down and acknowledge social knowledge, instead of memorizing it as if it were straightforwardly and individually theirs. In our everyday life, we have access to millions of sources that provide information to our queries. Thus, instead of trying to memorize all these sources, we only have to source them when needed. This practice leads us to the idea of knowledge collaboration and social learning. Since knowledge can be reached through a few steps in our digital devices, we should aim in developing skills about how to collaborate in this new setting, by creating knowledge, making it available to other learners and also, using appropriately the work of others. This practice represents a shift in contemporary ways of dealing with knowledge in the digital era and evolving learning contexts.
6. **Metacognition:** This refers to thinking about thinking, a learner thinking about what they have learned, in what way, what steps or methodology they followed, what they have acquired in comparison to their previous knowledge, and what goals they have achieved. This process corresponds in some sense with self-reflection, while it also focuses on the learning process and the disciplinary practice. The literature suggests that this process improves learner performance (Bransford et al., 2000) and that is why it is considered an attribute that needs to be valued in a contemporary learning context.
7. **Differentiated Learning:** This notion targets the different needs and interests that the students bring from the context of their lives to a learning situation. Digital technology systems can easily address these context-dependent factors, because of the adaptability and flexibility they offer. In recognition of contextual variation, learners can work at their own pace, dealing with content personalized to their level, without restrictions of space and time. Each student learns according to their interests and needs, and they bring to the learning process diverse perspectives that prompt deeper discussions. In the process of learning through digital tools, learners can work on different things at the same time or on the same thing, but at a different pace. In an equitable and inclusive education, they no longer need to be doing the same work at the same time, on the same metaphorical page.

Looking forward to a new wave of digital technologies, what might be the impact of might “artificial intelligence” (Cope and Kalantzis, 2016; Cope et al., 2020)? Artificial intelligence is profoundly different from human intelligence but offers the potential to transform learning contexts in ways that, paradoxically, can make them more human. Artificial intelligence is often identified by its methods, and it is assumed that the outcomes of these calculations can become more human, through machine learning, deep learning in neural networks, and perhaps, quantum computing in the future (Cope and Kalantzis, 2020a). However, these methods are only calculations and the question that rises is how they can support human meanings?

Since digital technologies can represent and communicate sound, text, speech and image, digital tools can act as “cognitive prostheses” and expand our capacities to make meaning. They function as a substitution for long-term memory and provide supplementary tools for reasoning. There is no value in repeating cognitive processes from memory when apps can work on them for us. However, there is value now on the ability to use these apps and digital devices to access social memory. Therefore, the learning contexts today should be designed on the basis of the sociability of knowledge.

To give a practical sense in this theoretical notion of cognitive prostheses, we mention one of the recent projects we are working on at the University of Illinois within the domain of language learning. We have created an app within the social learning platform CGScholar, where pairs of native speakers (Greek–English and Spanish–English) work together on each other’s written language, creating and offering each other scaffolded, machine-supported formative feedback on multiform representations and communication (Cope et al., 2020). The research aims for further improvement of the digital tools and their application in language learning

classrooms. At the same time, this is a practical example of putting “cognitive prostheses” to work, where human-technology interaction has become an integral part of the communicative and learning repertoire of humans. Students in this context are encouraged to learn how to use digital tools effectively in their everyday lives, in this case to facilitate communication in peer language learning. In this learning environment, considering the needs of contemporary learners of additional languages, transposition of meanings on the level of languages and multimodality has central place in all the dimensions of the environment. Practically, this occurs by supporting hybrid language use between participants and using the semantic prostheses of machine translation tools as an integral part of the learning practice.

We mention these practices around the use of artificial intelligence in education in order to highlight the potential it provides for learning, as well as the necessity to be considered in a social context currently in a process of transformation as a consequence of these developments. Machine intelligence has become an integral part of our daily lives. Therefore, from the educational point of view, it is critical to do some forward thinking and consider the needs and interests of contemporary and future learners when designing and implementing a new generation of learning environments.

Addressing changing contexts in formal education

The integration of digital technologies in all aspects of society is a phenomenon that is affecting our everyday lives. However, education today, even when it uses digital media, is frequently repackaged in technically outdated and pedagogically narrow learning management systems that were designed to replicate engineers’ recall of their own school days but not to realize contemporary educational values. Among recent innovations is “personalized learning,” allowing learners to work at their own pace but at the cost of tethering them to their screens and isolating them from learning community interaction. Behavioral surveillance systems might even make old-school behaviorist educational psychologists blush. Games try to make children think that rote learning can be fun, but the kids know that “real” games are much more fun. Relentlessly frequent online quizzes and tests are narrowly cognitivist in their orientation, mostly measuring memory of facts and correct replication of operations. Applications of artificial intelligence to education bring the perils of inhuman automation, surveillance and profiling that produces a new and often invisible tracking, along with commercial ownership of personal education data and digital identities (Cope and Kalantzis, 2015a,b).

Beneath these practices of pedagogy, schooling continues to serve and even to amplify social inequality. The school insists on inequality so long as its systems rely on summative and norm-referenced assessments. Learners are brought into the class with the expectation that all can perform, but summative assessment requires that only a few excel while many do moderately, or not very, well. Aptly, the word used to evaluate the value of deceptively nearly correct, trick items in select response assessment theory is “discrimination.” With a cognitivist and individualist focus, the mediocre and poor performers are left to blame themselves for what they might suppose to be a lack of intelligence or ability. Schooling legitimates a dishonest meritocracy, because in “real” life those who succeed on measures of wealth and power are not necessarily terribly meritorious. The few who succeed do so because the system—including the education system—is rigged in their favor. This is how the outcomes of schooling come to be as unfairly unequal as the larger society, and even magnify that unfairness. So, what is to be done?

The outcomes of schooling, we might term “epistemic capital.” If we navigate our way around the dangers of back-to-the-future educational media and harness their innovative and inclusive potentials, this may be our new best chance to redress the longstanding failures of schooling to achieve its social objectives of equity (a level playing field of opportunity) or progress toward equality (of outcomes). In schools, e-learning is the first big change in educational media since the widespread application in early modern times of the teacher lecture, the textbook, and the test. These discursive forms were subsequently mass institutionalized in the industrial era. Of course, things may go from bad to worse: if in discursive terms the e-textbook is just a digital version print textbook where all the students are on the same page at the same time; if the learning management system is just an online replicant of the lock-step syllabus; if the flipped classroom video is just a version of the lecture that the student quiescently watches; if the video discussion is just a version of the in-class discussion where one domineering student second-guesses the correct answer to the teacher’s question; and if the computerized test is just a reincarnation of old, summative select response and supply response tests.

Things do not have to be this way. A purposeful shift to learning analytics could offer formative feedback, where every student can succeed albeit at a pace that works for them (Dawson and George, 2014; van Haren and Harroun, 2019). Differentiated instruction could be oriented to equivalent outcomes. Scaffolded peer feedback and artificial intelligence could be designed into digital knowledge ecologies, offering extensive and timely support on a scale previously impracticable for teachers on anything more than one-to-one teaching (Cope et al., 2020). Where the old communicative artifacts of formal learning were biased toward reading and listening, textbooks and teachers, e-learning ecologies can rebalance learner agency, so every student spends as much time engaged in active meaning making as “reading” those meanings, including those of their peers (Montebello et al., 2018a,b). Individual intelligence now becomes inseparable from collaborative intelligence, and instead of penalizing learners for “cheating,” we might offer “help credits” for their contributions to the learning of others and the learning community. As soon as epistemic agency is rebalanced, learner differences become visible, and these become valued and a resource for learning. An architecture of templated sameness (watching the video lecture, being on the same page of the e-textbook, marching lock-step through the learning management system, and standardized online tests) is replaced by a learning ecology of productive diversity where then differences between learners are leveraged in scaffolded peer feedback as a resource for learning (Kalantzis and Cope, 2016a).

Today’s education realities are confronting and, as educators, our responsibility is to confront them. Here, in the relation of education to its wider social context, we want to apply Nancy Fraser’s argument that justice is an intersecting politics of

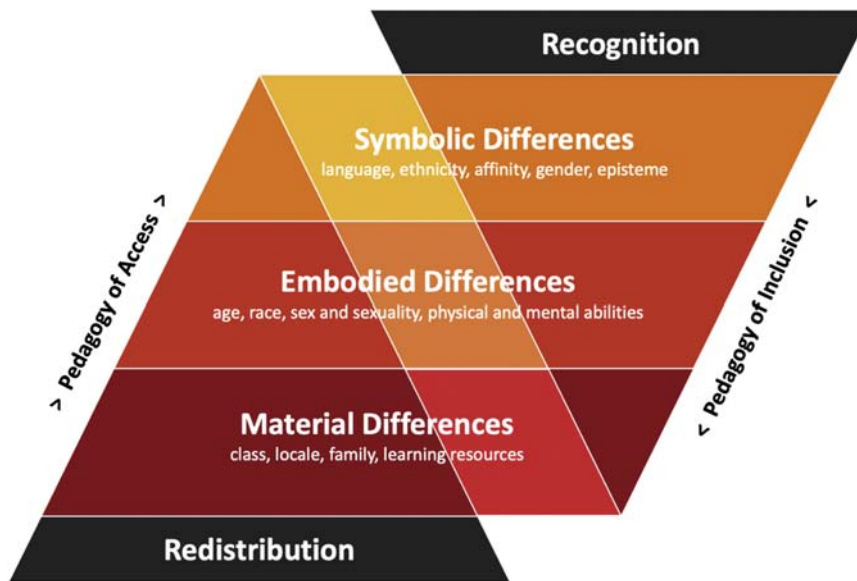


Fig. 4 Toward justice pedagogy: parallel programs of redistribution (for access) and recognition (for inclusion).

redistribution and recognition (Fig. 4). Epistemic capital is a powerful determinant of social outcomes, transcending race, gender, ethnicity and other markers of privilege. Of course, race, gender, and ethnicity are crucial determinants of educational outcomes, but the demographics of outcomes tells us that education makes a crucial difference whatever the other variables. There can be no justice in education so long as the already-privileged get many more resources spent on their education than the less-privileged. Democracy promises equality of opportunity, if not necessarily equal outcomes. This is called “equity,” a project which we fall back upon when the project of equality for the moment at least seems impossible. More than ever in the era of digitization, automation and artificial intelligence, education is the pathway to epistemic capital. But its promise of opportunity for all might be classed “fake news” when the resources attributed to the development of epistemic capital are so unevenly distributed. We describe the meaning-making essence of this epistemic capital as multiliteracies and its pedagogy of access.

Of parallel significance is a politics of recognition and a pedagogy of inclusion—where the exclusionary discourse and institutional practices of racism, sexism, homophobia, and anti-elite, anti-science intellectualism are confronted in schools before they can become baked into adult consciousness. These require the development of cultures of schooling that afford dignity to diverse personhood, extend a sense of belonging to all members of the learning community in equal measure, and treat every person with respect for the unique conjunction of their lifeworld experience and sensibilities. To learn, students need to feel they belong in that learning. The learning must also create space for agency that offers opportunities for alternative paths of self-development, the outcomes of which may be different but equivalent in the scope they offer for social action and personal autonomy.

A politics of redistribution without recognition fails to the extent that it frequently misunderstands and alienates the subjects whose lives it aims to improve—this was the error of the welfare state and its one-size-fits-all, comprehensive curriculum. On the other hand, a politics of recognition without a politics of redistribution can ironically, tragically, and incorrectly be read by some as a politics of exclusion, when for instance poor whites mistakenly feel they are not getting the attention afforded blacks or immigrants. The politics of identity without a politics of material distribution has produced destructive ethnonationalist rage whose shrill voices, amplified by social media, have at times bordered on unhinged.

As for the equity agenda, in Jim Gee’s words, this may simply be a matter of “learning to ‘play by the rules’.” Equity in this sense means successful adaptation to context. However, at the same time it may also be “a way station on the route to people being pro-social agents able to innovate and transform ... for their own purposes and for the creation of a better world” (Gee, 2009: 27). Now the conversation turns to how education can be a force that can change its context.

References

- Arkorkful, V., Abaidoo, N., 2015. The role of e-learning, advantages and disadvantages of its adoption in higher education. *Int. J. Instruct. Technol. Dist. Learn.* 12 (1), 29–42.
- Blake, R.J., 2009. The use of technology for second language distance learning. *Mod. Lang. J.* 93 (Suppl. S1), 822–835. <https://doi.org/10.1111/j.1540-4781.2009.00975.x>.
- Bloom, B.S., 1968. Learning for mastery. *Eval. Comment* 1 (2).
- Bransford, J.D., Brown, A.L., Cocking, R.R., 2000. In: N. R. C. Commission on Behavioral and Social Sciences and Education (Ed.), *How People Learn: Brain, Mind, Experience and School*. National Academy Press, Washington D.C.
- Case, A., Deaton, A., 2020. *Deaths of Despair and the Future of Capitalism*. Princeton University Press, Princeton NJ.
- Chen, J.T., Testa, C., Waterman, P., Krieger, N., 2021. Intersectional Inequities in Covid-19 Mortality by Race/Ethnicity and Education in the United States, vol. 23. Harvard Center for Population and Development Studies, Cambridge MA.

- Cope, B., Kalantzis, M., 2015a. Assessment and pedagogy in the era of machine-mediated learning. In: Dragonas, T., Gergen, K.J., McNamee, S., Tseliou, E. (Eds.), *Education as Social Construction: Contributions to Theory, Research, and Practice*. Worldshare Books, Chagrin Falls, OH, pp. 350–374.
- Cope, B., Kalantzis, M., 2015b. The things you due to know: an introduction to the pedagogy of multiliteracies. In: Cope, B., Kalantzis, M. (Eds.), *A Pedagogy of Multiliteracies: Learning by Design*. Palgrave, pp. 1–36.
- Cope, B., Kalantzis, M., 2016. Big Data comes to school: implications for learning, assessment and research. *AERA Open* 2 (2). <https://doi.org/10.1177/2332858416641907>.
- Cope, B., Kalantzis, M., 2017. Conceptualizing e-learning. In: Cope, B., Kalantzis, M. (Eds.), *E-Learning Ecologies: Principles for New Learning and Assessment*. Routledge, pp. 1–45.
- Cope, B., Kalantzis, M., 2020a. Futures for research in education. *Educ. Philos. Theor.* 1–8. <https://doi.org/10.1080/00131857.2020.1824781>.
- Cope, B., Kalantzis, M., 2020b. *Making Sense: Reference, Agency and Structure in a Grammar of Multimodal Meaning*. Cambridge University Press.
- Cope, B., Kalantzis, M., Sears Smith, D., 2020. Artificial intelligence for education: knowledge and its assessment in AI-enabled learning ecologies. *Educ. Philos. Theor.* 52 (16), 1–17. <https://doi.org/10.1080/00131857.2020.1728732>.
- Dawson, S., George, S., 2014. Analytics to literacies: the development of a learning analytics framework for multiliteracies assessment. *Int. Rev. Res. Open Dist. Learn.* 15 (4).
- Dublin, L., 2003. If you only look under the street lamps ... or nine e-learning myths. *eLearn. Dev. J.* Retrieved from <https://www.elearningguild.com/pdf/2/061603man.pdf>.
- Gee, J.P., 2009. Reflections on reading Cope and Kalantzis' "Multiliteracies: new literacies, new learning". *Pedagogies* 4 (3), 196–204. <https://doi.org/10.1080/15544800903076077>.
- Gotschall, M., 2000. E-learning strategies for executive education and corporate training. *Fortune* 141 (10), 5–59.
- Holmes, B., Gardner, J., 2006. *E-Learning: Concepts and Practice*. SAGE Publications, London.
- Kalantzis, M., Cope, B., 2012. *New Learning: Elements of a Science of Education*, second ed. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CB09781139248532>.
- Kalantzis, M., Cope, B., 2016a. New media and productive diversity in learning. In: Barsch, S., Gluttsch, N. (Eds.), *Diversity in Der Lehrerinnenbildung*. Waxmann, Münster DE, pp. 310–325.
- Kalantzis, M., Cope, B., 2016b. Learner differences in theory and practice. *Open Rev. Educ. Res.* 3 (1), 85–132. <https://doi.org/10.1080/23265507.2016.1164616>.
- Kirschner, P.A., Sweller, J., Clark, R.E., 2006. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41 (2), 75–86. https://doi.org/10.1207/s15326985ep4102_1.
- Kraemer, A., 2008. Formats of distance learning. In: Goertler, S., Winke, P.P. (Eds.), *Opening Doors Through Distance Language Education: Principles, Perspectives and Practices*. CALICO, San Marcos, TX, pp. 11–42.
- Maltz, L., Deblois, P., 2005. Top ten IT issues. *Educ. Rev.* 40 (1), 15–28.
- Melnic, A., Botez, N., 2014. Formal, non-formal and informal interdependence in education. *Econ. Transdiscip. Cogn.* 17 (1), 113–118.
- Montebello, M., Cope, B., Kalantzis, M., Sears Smith, D., Amina, T., Olga Tzirides, A., Zhao, N., Chen, M., Haniya, S., 2018a. Deepening E-Learning Through Social Collaborative Intelligence. Proceedings of the 48th IEEE Annual Frontiers in Education (FIE) Conference.
- Montebello, M., Pinheiro, P., Cope, B., Kalantzis, M., Amina, T., Sears Smith, D., Cao, D., 2018b. The Impact of the Peer Review Process Evolution on Learner Performance in E-Learning Environments [Paper Presentation]. Proceedings of the Fifth Annual ACM Conference on Learning at Scale (L@S 2018), B. In London UK. <https://doi.org/10.1145/3231644.3231693>.
- Oblinger, D.G., Hawkins, B.L., 2005. The myth about e-learning. *Educ. Rev.* 40 (4), 14–15.
- OECD, 2005. *E-Learning in Tertiary Education*. Available at. <http://www.oecd.org/edu/eri/35991871.pdf>.
- Piketty, T., 2014. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge MA.
- Pinheiro, P., 2020. Text revision practices in an e-learning environment: fostering the learning by design perspective. *Innovat. Lang. Learn. Teach.* 14 (1), 37–50. <https://doi.org/10.1080/17501229.2018.1482902>.
- Raikou, N., Karalis, T., 2010. Non-formal and informal education processes of European lifelong learning programs for higher education: the case of the Erasmus program in a Greek Peripheral University. *Int. J. Interdiscip. Soc. Sci.* 5 (1), 103–113.
- Sassen, S., 2014. *Expulsions: Brutality and Complexity in the Global Economy*. Harvard University Press, Cambridge MA.
- Schugurensky, D., 2000. *The Forms of Informal Learning: Towards a Conceptualization of the Field*. Centre for the Study of Education and Work, OISE/UT. <https://tspace.library.utoronto.ca/handle/1807/2733>.
- Strimel, G., Reed, P., Dooley, G., Bolling, J., Phillips, M., Cantu, D.V., 2014. Integrating and monitoring informal learning in education and training. *Techniques* 89 (3), 48–54.
- van Haren, R., Harroun, J., 2019. Cgscholar's analytics: progress to mastery learning. *Ubiquitous Learn.* 12 (2), 1–24. <https://doi.org/10.18848/1835-9795/CGP/v12i02/1-24>.

Cross cultural lenses on learning

Carol D. Lee, School of Education and Social Policy, Northwestern University, Evanston, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Understanding culture	267
Why understanding cultural communities of practice is important	268
Studies of differences in pathways of learning and development across cultural communities of practice	271
References	273

In order to examine cross cultural perspectives on learning, three fundamental propositions must be interrogated. The first is how we understand culture. The second is understanding the functionality of culture in the evolution of humans as a species. And the third is understanding the role of diversity in cultural communities in terms of our evolution and survival as a species. These foundational propositions allow us to place historical studies of diversity in pathways of learning in context and to argue why such diversity is functional and important.

Understanding culture

Across fields in psychology and even anthropology, culture has been ascribed to groups of people presumed to be homogenous. Such groupings have typically included nationality, religion, ethnicity and most recently—in the last 500 years—constructions of race. There is a distinction between nationality and ethnicity. National borders historically shift based on warfare and political shifts. For example, Spain did not become a unified nation until 1492. The movements among the Catalan peoples even today around maintaining their language and even calls for separating from the Spanish state are just one example of ethnicity preceding nationality in many cases. Ethnic groups typically have longer histories than nation states. This is evident in the many different ethnic groups that either make up most nation states or that cross national borders, such as the Roma people who can be found across Europe. And it is equally important to recognize the heterogeneity within the Roma ethnic group. Race as a social and political construction is relatively recent in human history, taken up during the period of the European Enlightenment by philosophers and scientists in the physical sciences. From Emmanuel Kant and J.F. Blumenbach in the 1770s to French writer Gobineau's *Essay on the Inequality of the Human Race* in 1853 to Ripley's 1899 *The Races of Europe*, the argument that human beings could be divided into hierarchies based on skin color had become deeply established in European and US intellectual thinking. These philosophical arguments buttressed and provided moral warrants for the political and economic efforts to colonize and enslave populations from Africa and indigenous populations in the Americas and other parts of the world (Mills, 1997). Among the ironies of this system of racial classification has been the internal presumptions of social, cultural and biological hierarchies among populations presumed to be "white." For example, the English demonized the Scots and Irish even prior to their colonization efforts in the Americas, with a history of lynching that precedes such horrendous acts during the enslavement of Africans in the Americas (Dunbar-Ortiz, 2014). Ripley in his 1899 *The Races of Europe* distinguished between the Nordic races which he presumed to be superior to the "feeble-minded" lower Mediterraneans. By the 1800s scientific racism emerges with the field of craniology where scientists use measures of presumed differences in brain size among racial groupings as scientific evidence of natural superiority (Gould, 1981). In 1892 the American Psychology Association was formed with its first president, renowned psychologist G. Stanley Hall who believed in racial hierarchies. The Eugenics movement evolves in the early part of the 20th century, largely in the US, where psychology as a field takes up these assumptions of racial superiority in natural intelligence, advocated by preeminent educational psychologists such as Lewis Terman, Robert Yerkes, and Edward Thorndyke. By 1926 The American Eugenics Society was formed. By 1912 Goddard completes a study of a family of poor whites whom he argues are inferior and therefore should not be allowed to reproduce. Ironically, the Nazi regime of Germany studies the US Eugenics movement as a model for their own system of racial hierarchies. In the US who gets classified as white has shifted (Ignatiev, 1996; Williamson et al., 2007). In the US the Eugenics movement has had a profound impact on the organization of public schooling, leading to differentiated instruction or tracking associated with race and class (Tyack, 1974; Winfield, 2007). By 1969 Jensen publishes "How Much Can We Boost IQ and Scholastic Achievement" in the *Harvard Education Review* and in 1994 Hernstein and Murray publish *The Bell Curve* and in 1999 James Traub publishes "What New School Can Do" in the *New York Times Magazine*. The boundaries of who is positioned as the other, shifts. More recently, propositions around inherent deficits has generalized to what has come to be called "a culture of poverty" (Jensen, 2009; Payne, 1999). This abbreviated history of conceptions of human communities illustrates how these ideas have been taken up in philosophy, in the sciences, and indeed in the world of practice.

The point here is that constructions of cultural communities of humans is long, complex, highly contested, and its contestations deeply embedded in formal intellectual disciplines. It is important to note that contestations over presumptions of racial hierarchies and even ethnic hierarchies have been made consistently by the communities that have been demonized. DuBois (1897), Hancock (2005) as early as 1897 critiqued the very idea of race. Historians and sociologists from Carter G. Woodson to Mary McLeod

Bethune, historians from Anna Julia Cooper to Lerone Bennett and John Henrik Clarke have critiqued the construct of race as ideological and political. In psychology as a domain, for example, the field of Black Psychology has examined the impact of cultural practices among peoples of African descent as evidence of long histories of African ethnicity rather than race associated with enslavement (Akbar, 1979; Azibo, 2003; Cross, 1991; Jones, 1980).

Culture is defined in terms of routine communities of practice (Kris Gutierrez and Rogoff, 2003; Rogoff, 2003). As such, people belong to multiple communities of practice. Some communities of practice are historical and inter-generational. Such inter-generational practices often transcend national boundaries. For example, pan-ethnic groups often sustain belief systems, routinized forms of social interaction, patterns of social networking, and ways of using language that move across national borders and times. Boykin (1985) describes this phenomenon as the deep structure of culture. Peoples of African descent—on the continent of Africa and dispersed across the diaspora in Europe and the Americas—have sustained myriad shared cultural practices (Akbar, 1979; Asante and Asante, 1990; Blakey, 1998; Champion, 2003; Karenga, 2003; King, 1976; Nobles, 1972). These include respect for elders, the value of the extended family, beliefs around connections among ancestors, the living and the unborn, the centrality of rhythm in communication and expression of emotion, among others. These practices can be viewed among African descent communities around the world. At the same time, it is equally important to recognize heterogeneity within these diverse African descent communities. This is because routine inter-generational cultural practices must simultaneously adapt to changing physical, historical, economic and political circumstances.

Cultural communities of practice may be historical and inter-generational to include nationality, ethnicity, religion. People are often born into such communities, influencing how they grow up, but also may elect to disengage from those communities over time. Some cultural communities of practices are ones that you can elect to join. Others do not have porous boundaries and are ones that can only enter through birth or some defined process of initiation. There are also cultural communities that entail practices of personal interest and often require specialized skills. These may include members of sports communities, communities of practice within the arts (e.g., hip hop, European influenced classical music, readers of Harry Potter novels), or other routinely organized practices. One can also think of communities of practice associated with historical periods. For example, Elder (2018) analyzes the impact of growing up during the Great Depression in the US on generations. Today we talk about the Millennium generation. The point here is that an important starting point for conceptualizing learning and development across cultures is that any such theorizing must be rooted in an appreciation for the very complexity of the idea itself. People belong to multiple communities of practice. Such communities of practice are both homogenous and heterogenous at the same time. All such communities of practice support and socialize repertoires, tools for thinking and acting, that serve diverse functions and goals. Vossoughi and Gutierrez (2014) conceptualize engagement in and across communities of practice as movement and not as simple stable positioning.

Why understanding cultural communities of practice is important

In education, we attend to issues of culture and cultural practices as they are potentially useful in the design of environments that support ideally robust learning (Gutiérrez, 2016; Lee, 2010). Ideally the robust learning includes the holistic development of people, particularly young people. Holistic development includes cognitive, social, emotional and physical development that supports well-being. This conception of goals of human learning extends beyond simple technocratic goals of learning academic content in preparation for work as an adult. A wide range of social settings contribute to such holistic development of children and adolescents, including the immediate and extended family, routine social networks, community settings, schools, churches, and communal resources available in both the neighborhood, the town, city, or rural region and the nation (Bronfenbrenner and Morris, 1998, 2006), and indeed resources that may be available from across the world. For example, while internet access is supported or not by local and national entities, the intellectual and social resources available on the internet come from around the globe.

The foundational question is, what role does participation in cultural practices play in human learning and development? The most basic answer is that participation in cultural practices is essential to human development (Gopnic et al., 1999; Kuhl and Meltzoff, 1996; Lee et al., 2020; Meltzoff, 1988, 2013). At birth and from early infancy, infants pay more attention to human faces than objects. Babies and young children learn their most essential competencies through imitation and exploration of the physical and social environments into which they are born and grow up. These competencies include learning language, learning to read the internal states of others—human and animate—learning to use their bodies to walk and manipulate objects. They navigate learning to read the internal states of others in relation to their evolving perceptions of their own ego focused goals and needs. Infants at birth can hear all the sounds/phonemes of all human languages. However, through imitation—hearing and testing the functionality of understanding and re-producing the sounds of the language or languages they hear—their neural networks prune to home in on the language or languages of their social environment. The documentary film *Babies* provides a powerful visual representation of the universal efforts to learn the basic human expectations for the first year of life. It documents the first year of life of four babies—from Namibia, Mongolia, Japan and the US. There is no talking during the film. We only see the babies doing what human babies do. That is they engage in routine cultural practices, embedded in the social, physical and cultural environments into which they are born, to learn to walk, to manipulate, to connect with other human beings and animals (for those growing up in rural environments). It is arbitrary to argue that one form of cultural socialization is superior to another. Rather the film vividly illustrates the diversity of pathways of routine cultural practices to achieve the foundational tasks of human development, tasks inherited from our evolution as a species. It should be noted, also, although not documented in the film, that diversity of pathways for achieving foundational

competencies we need as humans includes how children born with what are typically called disabilities also adapt through participation in routine cultural practices. And indeed our adaptability as a species contributes to navigating difference. Children who are blind typically develop enhanced hearing and touch, for example. And this adaptability is an outgrowth of both biological adaptability in tandem with the development of artifacts in human communities at particular moments in historical time—for example, invention of sign language, of hearing aids, of prosthetics for walking and manipulating. The development of artifacts is embedded in cultural practices and they serve as partners, if you will, for thinking and acting. Cole (1996), drawing on Wartofsky (1973), defines this as the mediational role of artifacts.

There are now well-developed programs of research that document and illustrate the essential dialogical relationships among physiological processes that we inherit from our evolution as a species (some of which we share to different degrees with our close animal relatives), our participation in routine cultural practices (Cacioppo et al., 2000; Chiao and Ambady, 2007; Cole, 2007; Han et al., 2013; Tomasello, 1999). These dialogic relationships also include individualized characteristics that contribute to individual personality as well as individual acuties that, if exposed to appropriate cultural supports, can lead to the development of highly developed skills (e.g., musicians, artists, experts in the sports, writers, dancers, scientists, etc.). One interesting emerging area is the study of epigenetics (Russo et al., 1996). These studies illustrate how the expression of genetic markers can be moderated by participation in cultural practices, shifts in genetic expressions that can cross generations.

We understand from these studies in the neurosciences how thinking, feeling, perceptions—of the self, of others, of settings, of tasks—interact in dynamic and dialogic ways to shape what we do, what goals we set, how motivated we are to engage, and how persistent we are in pursuing goals—short and long term (Immordino-Yang and Damasio, 2007; Zajonc and Marcus, 1984). These processes—thinking, feeling, perceiving—in the mature adult have some relatively stable dimensions, but are also highly malleable depending on the processes entailed in the routine cultural practices in which we participate (Rudman et al., 2001). This means that as we consider how participation in routine cultural practices can contribute to robust learning environments, we want to understand the following (Cantor et al., 2018; Osher et al., 2018): how micro-level interactional processes, along with the availability and functionality of artifacts in such settings, support a complex understanding of cognitive processes required for the task(s) at hand; how they support feelings of belonging and safety; how they recruit perceptions that people bring around who they are, what they are capable of, how relevant they view the tasks and the demands of participation, and how they relate to people in the environment. It is equally important to understand or at least consider what people learn around these psychological processes in one setting as they are relevant to other settings, related to questions of transfer.

These complex relationships entail identity development and contribute in important ways to how people may or may not be resilient in face of challenge (Nasir and Peele-Eady, 2012). As we consider cultural diversity within and across communities of practice, it is important to note that there are many cultural communities who face persistent and often inter-generational structural barriers that must be navigated and resisted in order to work toward well-being (American Psychological Association Task Force on Resilience and Strength in Black Children and Adolescents, 2008; Berry et al., 2006; Boykin, 1986; Grande et al., 2015; Lee, 2017). These possibilities of resilience in the face of challenge, indeed in the face of trauma, are extremely important to understand. In education and other fields, there has been a tendency to either assume only victimhood in marginalized communities, to assume that only through internalization of what are presumed to be mainstream cultural practices can persons from marginalized communities achieve and be resilient in the face of structural challenges and barriers. In many ways, these assumptions of hierarchies of cultural practices mirror the same assumptions of social and cultural hierarchies discussed earlier around some of the same minoritized communities. It is important to also recognize that such othering has also historically been ascribed to women as a gender.

Included here are two theoretical frameworks that embody these complex propositions about resilience in light of diversity and of the dialogic relations among physiological and psychological processes.

The first is the Phenomenological Variant of Ecological Systems Theory (PVEST) developed by Margaret Beale Spencer (Spencer, 2006; Margaret Beale Spencer, 2019; Spencer et al., 2006). Supported by several decades of empirical research, PVEST identifies a set of processes and relationships that contribute to developmental processes. PVEST is a consequential tool for theorizing about diverse pathways of learning and development across cultural communities because it articulates both foundational human processes of meaning making but requires examination of the specific sets of cultural practices taken up in these processes. Fig. 1 below captures the PVEST model. The first component asks us to identify net vulnerability. Net vulnerability is not simply what we think of as the objective reality of an individual or a group's lived experience. Rather net vulnerability is the relationship between risks faced and protective factors available. For example, poverty is not an undifferentiated risk. In countries that have societal protective factors (e.g., free health care, free pre-school, supports for housing, universal access to healthy food, enriched education for all), the experience of poverty is different from countries that do not provide these societal supports. In the same vein, growing up in a single parent household does not hold the same risks as growing up in a single parent household that includes inter-generational family and deep social networks that include friends, church, etc. Net stress refers to how the individual experiences that net vulnerability. How one experiences exposure to vulnerability depends on relationships between the nature of the challenges experienced and the supports available. Sometimes supports are available, but are not perceived by the person or the community as actually relevant or useful. This second component of the model is important because it pushes against the idea that everyone exposed to vulnerabilities will experience them the same way. This component stresses the importance of perceptions to experience. The third component is reactive coping processes. One important underlying proposition here is that to be human is to be vulnerable. To be human is to face risks. It is not simply the risks we face, but equally important the resources that are available to us to wrestle with those risks, which influence how we cope, how we make sense of such risks. Such coping, Spencer (2006) argues, may be adaptive or maladaptive. It is important to note that these coping responses are malleable, and malleable across the

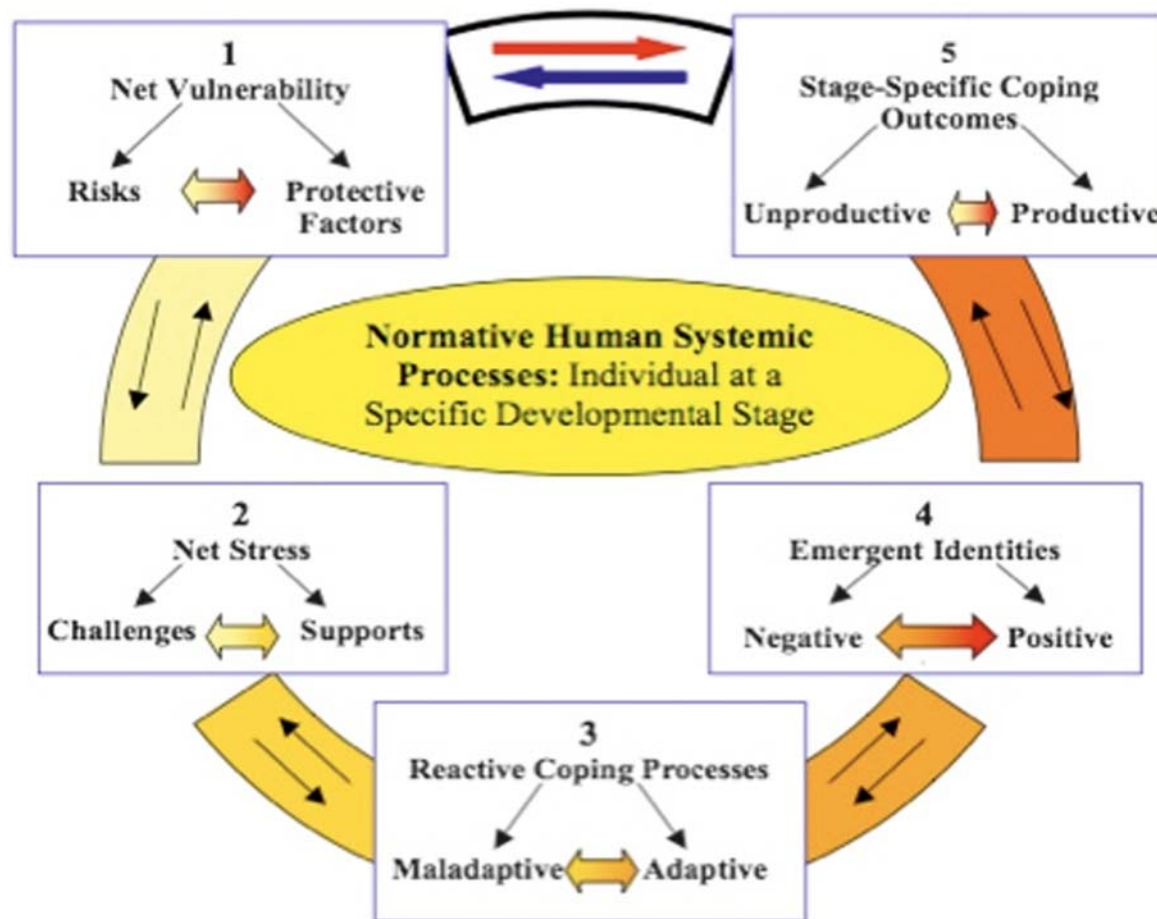


Fig. 1 PVEST. Source: [Spencer \(2006\)](#).

life course. These coping processes that we learn over time become internalized as what she calls emergent identities. Emergent is important for several reasons. One is because such identities are malleable. There is no point in the life course when we are incapable of seeing ourselves differently. The fifth component is what she calls stage-specific coping outcomes that can be productive or unproductive. Stage specific takes into account how identity processes and life course outcomes shift from early childhood through middle childhood through adolescence into adulthood, and indeed include the stages of adulthood.

PVEST is offered here as a framework that supports how we can think about both homogeneity and heterogeneity within cultural communities, about both groups and individuals within groups.

The second framework is the R.I.S.E. principles. This framework evolved from the syntheses of studies of the role of culture in human learning and development, captured in the handbook by [Nasir et al. \(2020\)](#). Drawing from research in the neurosciences, cognitive and social psychology, the learning sciences, and human development, these principles capture the foundational conception of human learning and development as outgrowths of complex dynamic systems that include the following: physiological processes rooted in our evolution as a species as these physiological processes are taken up in culturally organized processes; processes that develop and unfold over time across the life course; whose intersections unfold and are embodied in social interactions, which in tandem entail and influence identity, cognition and emotions. This is indeed a complicated space whose complexity has and continues to pose challenges to how we interrogate human learning and development.

- (1) Rooted in evolutionary, biological, and neurological systems,
- (2) Integrated with other developmental processes—emotion, identity, cognition,
- (3) Shaped in culturally organized practices across people's lives,
- (4) Experienced as embodied and coordinated through social interaction

At the same time, we ignore the dynamic relations among these processes at our peril. They exist and are at work whether or not we have the tools to study them. At the same time, the evidence of individuals and communities who persist, resist, and are resilient in the face of challenge and trauma is real life evidence that these complex processes are at work; and that people in everyday settings intuitively understand them. Families understand the importance of food and rest for healthy development, the importance of loving relationships where we seek to make children feel loved and safe, value routine, often inter-generational practices for

socializing children; and as children become adolescents and adults, parents come to learn how these supports must be adaptive over the life course. We also know there are families who are middle-class and belong to communities presumed by the broader society as privileged because of ascriptions of whiteness who also experience trauma without adequate social and emotional supports to be resilient. Spencer, drawing on [Anthony \(1987\)](#), argues that there are people in privileged communities who are deeply vulnerable specifically because they have not learned to wrestle with challenge as well as people in minoritized communities and communities living in persistent inter-generational poverty who have learned to be resilient specifically because of the supports they have for wrestling with adversity. Cultural communities are both homogeneous and heterogeneous.

Because these principles are central to understanding human functioning, having been validated by decades of research converging across multiple disciplines—cognitive and social psychology, the neurosciences, human development, and learning sciences—they provide useful lenses for examinations of how these foundational principles unfold along diverse pathways, serving diverse functions within and across diverse cultural communities. We understand that diversity among cultural communities not only exists but is essential for the adaptability of our species to survive in light of challenge and changing circumstances. The warrants that undergird this body of scientific research stand in stark contrast to the earlier periods of research on human differences, such as traditions of scientific racism, eugenics, and more recent propositions about a culture of poverty.

Studies of differences in pathways of learning and development across cultural communities of practice

Two major 20th century figures influencing our understanding of human learning are Piaget and Vygotsky ([Pass, 2004](#)). [Piaget \(1952, 1975\)](#) focused on cognitive development of children across multiple stages as outgrowths of biological maturation and processes and exploration of the physical and social world around them. Rather than focusing on discrete knowledge, Piaget was concerned with the development of underlying mental structures or schemas for logical reasoning. Piaget documented these stages of cognitive development largely through experiments with European children. While Piaget attended to children's interaction with their environments, he still largely focused on cognitive development as internal processes within individuals and presumed universal stages of development. During the same period of Piaget's early research, Lev Vygotsky in Russia was studying children's development. While these two are often pitted against one another, they actually did correspond and connect to one another's work ([Pass, 2007](#)). [Vygotsky \(1978, 1981, 1977\)](#) argued for the centrality of social interactions as the medium of development, with social interactions as the way children learn to intuit cultural norms and practices. This proposition means that developmental processes are not universal but vary by virtue of diversity in the range of cultural practices into which children are socialized. Rather than stages of development as Piaget proposed, Vygotsky identified what he called the zone of proximal development or the ZPD. Rather than learning as an outgrowth only of the individual child's exploration, Vygotsky argued the usefulness of focusing on what children are able to learn as tasks just beyond what they can do on their own, scaffolded by the support of more knowledgeable others, along with the possibilities that different cultural artifacts make available to the learner. Such cultural artifacts include language. [Cole \(1996\)](#), drawing on earlier work by [Warkofsky \(1973\)](#), identifies different categories of cultural artifacts: primary artifacts (used directly in acts of production), secondary artifacts (representations of primary artifacts and modes of action, such as recipes and beliefs), and tertiary artifacts (examples include imaginary worlds such as works of art). These artifacts are cultural in that they are created by human beings working in coordination with one another for the purposes of carrying out socially constructed tasks. This brief history is offered because these two bodies of research influenced a powerful family of research on cultural foundations of learning across different cultural communities.

By the 1970s, a number of cognitive psychologists (Michael Cole, Vera John-Steiner, James Wertsch) discovered Vygotsky and translated much of his seminar publications into English. A number of these psychologists decided to test the idea of universal stages of logical reasoning by examining how communities of practice within different cultural communities outside of Europe engaged in logical reasoning, including looking at adults and children. They were faced with the challenge of disentangling the role of formal schooling and interrogating the assumptions that informed the instruments being used to measure logical reasoning. Interesting, as a result of these cross-cultural studies, [Piaget \(1966/1974\)](#) began to think differently about his final stage of formal logical reasoning. These cross-cultural studies often found that non-schooled populations would classify not on the basis of taxonomic categories but rather on the basis of functionality. On tasks of syllogistic reasoning, some studies, such as [Luria's \(1928\)](#) interviews with rural non-schooled and schooled Tadzhikistan and Uzbekistan populations in Central Asia found that non-schooled people did not accept the fundamental assumptions identified in the syllogistic task, but rather responded that the reasoning depended on the circumstances. Circumstances often involved first-hand knowledge as the criteria for truth rather than an abstract de-contextualized claim. In cross cultural studies of memory, researchers found that how objects subject to recall were situated in culturally familiar contexts mattered. As [Rogoff and Chavajay \(1995\)](#) note:

The conclusion from the memory studies, as with the classification and logic studies, is that researchers cannot just assume, as we did before cross-cultural research, that a cognitive test reveals a general ability across tasks unrelated to people's experience ... It is a very simple conclusion but one that transformed how researchers think about the generality of cognitive tests and about their relation to people's experience with particular institutions such as schooling (p. 863).

Rogoff and Chavajay offer a comprehensive overview of these foundational studies. [Scribner and Cole \(1981\)](#) studied how differences in literate practices among the Vai people of Liberia supported different skills, concluding that literate skills were not simply generic. Vai literacy included their indigenous phonemic-syllabic language, Arabic learned as part of Islam, and English learned in school. Each contributed to different literate skill sets. Another body of research documented the learning of arithmetic operations through the practices of tailoring and merchant activities in countries in Africa and South America ([Carraher et al., 1985](#); [Lave, 1977](#); [Nunes et al., 1993](#)). As opposed to general strategies, they studied differences in strategies for arithmetic problem solving associated with the routine everyday practices rather than simply invoking canonical strategies associated with schooling ([Lave, 1988](#)). In the studies of developing and using everyday numeric reasoning strategies, researchers were able to show how social relationships, scaffolded supports from others and artifacts structured problem solving, rather than viewing such reasoning as solely a process of internalization by individuals. For example, [Saxe \(1988a,b\)](#) studied children in Brazil who were poor, often unschooled, who sold candy on the streets of an urban city. He documents how these children are able to mentally calculate sums and cash due back as they sold packets of candy, largely to foreigners; buy candy in bulk and determine the best pricing per unit for re-sale, influenced in part by daily differences in the rate of exchange for foreign currency. He carefully compares the calculation strategies these candy sellers learn on the streets with the calculation strategies taught in schools, also showing how these children were more efficient in calculating in problems that mirror the tasks they tackle in selling on the streets. Another important dimension of Saxe's studies here was his documentation of the social processes and interactions through which children developed different levels of expertise in the practice. These included older children who were more experienced as candy sellers scaffolding the efforts of younger children, along with supports by merchants from whom they purchased candy in bulk. This was a unique example of scaffolding among children.

These studies together support the idea that displays of competence are not universal. Some competencies that have characterized measures of intelligence have presumed that tasks associated with formal western schooling capture undifferentiated universal indicators of ability. However, decades of cross-cultural studies of performance have shown how skill in performance is associated with participation in routine cultural practices which do not have to be limited to the practices of schooling. These performances around memory, classification, literacy and numeracy clearly are functional in that people in particular communities are able to carry out consequential tasks using culturally specific skills.

Another interesting dimension of these cross-cultural studies has been differences in community based shared assumptions about what it means to be intelligent, what it means to be skilled. These cultural values help to shape people's identities in terms of goals to pursue. For example, [Marcus and Kitayama \(1991\)](#) have demonstrated differences in perceptions of the self among Pan Asian communities and European descent communities. Asian communities tend to value the idea of an inter-dependent self where relationships with others are privileged. At the same time, they document that these community values are not necessarily held by every individual in the community. In a similar vein, earlier studies among ethnic communities in Africa document a similar valuing of inter-dependence—responsibilities to family and elders, of social obligation ([Beattie, 1980](#); [Diop, 1978](#); [Serpell, 1974, 1982](#); [Super and Harkness, 1986](#)). Other studies have shown how this valuing of inter-dependence can play out in formal schooling where the valued forms of social interaction privilege independence. [Philips \(1983\)](#) has shown how children from indigenous communities in the US may be reticent in drawing attention to themselves in classrooms that value competition. [Boykin \(1997\)](#) has shown how many African-American children prefer learning experiences that are collaborative versus ones that based on individual competition, and experiences that embody physical verve. [Rogoff et al. \(1995, 2014\)](#) have shown how children from indigenous communities of the Americas learn by intent participation. Here, as opposed to age-based segregation, children are incorporated into routine productivity activities of the family and learn by observing and participating.

One of the most informative studies to capture the breadth of complexity entailed in how people learn and the crucial role of participation in cultural practices within and across time is Saxe's ([Saxe and Esmonde, 2005](#)) longitudinal study of mathematical everyday reasoning among Oksapmin of Papua New Guinea. The Oksapmin historically developed a system for counting using a system of 27 body parts. Originally, saying Fu with a raised fist meant that all 27 body parts have been counted. However, in later years Fu came to mean double the last body part identified. Saxe showed how the meaning of Fu shifted over historical periods, influenced in part by the system of economic exchange involving cash exchanges using pounds and shillings brought by western colonial powers, and then shifts by the Australian government to dollars and cents, and the eventual independence of Papua New Guinea with currency in kina and toea, all requiring abilities to calculate amounts greater than the body counting system supported. That, in conjunction with bringing in western schooling, introduced school-based strategies for counting. Saxe developed a series of tasks asking people of different ages to explain the meaning of Fu. He found differences in the understanding of Fu and its shifting meaning based on age cohorts—elders, unschooled adults, schooled adults, schooled adolescents, and children. The goal was to identify persons who experienced both the shift in forms of economic exchange, some formal schooling but earlier training in the traditional methods, and young children who experienced only western formal schooling. That longitudinal study shows the functional role of cultural artifacts (Fu as a traditional construct, systems of economic exchange based on shifts in currency) in helping to shape how people reason, and how such influences are historically situated and taken up in specific cultural practices.

There are more recent cross-cultural studies that have introduced an important idea, namely, to reconsider the ontologies of the academic disciplines that have presumed to have been the primary purview of formal schooling and European and Euro-American intellectual traditions and practices. In some cases they offer fundamentally different epistemological dimensions to the knowledge domain, as in the work on indigenous knowledge systems in reasoning about the natural world by Bang and colleagues ([Bang and Marin, 2015](#); [Bang et al., 2013](#); [Bang and Medin, 2010](#)). Bang argues that indigenous knowledge systems do not position humans at the top of a hierarchical ladder. Rather they privilege the inter-dependence across all elements of the natural world. She further

argues that despite formal western schooling privileging humans at the top of a natural order, more recent basic research in the sciences reflects an understanding of our inter-dependence (Chief et al., 2018; Kimmerer, 2013; Makondo and Thomas, 2018). In other cases, studies document how foundational concepts and reasoning processes are embedded in everyday cultural practices that are not directly connected to or typically taken up in schooling. These include Lee's (1995, 2001, 2007) work in Cultural Modeling on sensemaking around narrative and figuration in everyday language practices among speakers of African American English, in hip hop and other popular music genres; Taylor's (2009, 2012) work documenting the mathematical reasoning around number sense among young children outside of schooling; Brown's (Brown, 2004; Brown and Kloser, 2006; Brown and Kelly, 2007; Brown and Kloser, 2009) research documenting the intuitive understandings of physics among school aged baseball pitchers; Nasir (2000, 2002, 2005) research on mathematical reasoning of children learning to play dominoes and high school basketball players; Major's (2001, 2003, 2015) research on narrative sensemaking and moral reasoning in what she calls Shop Talk in African-American beauty salons; Fisher-Winn's (Fisher, 2003, 2005, 2007; Winn, 2010) research on literacy learning, particularly around narrative and figuration in African-American community based arts institutions and sites of incarceration; Emdin's (2010a,b) work connecting hip hop and science learning (Orellana, 2009; Orellana et al., 2003; Orellana, 2001; Valdes, 1996, 2001, 2002; de los Ríos, 2019; de los Ríos and Molina, 2020; de Los Ríos and Seltzer, 2017) research examining the meta-linguistic knowledge embedded in culturally rooted and translanguaging practices among young people who are English Language learners (using structural knowledge of the first language to inform comprehension processes in the second language; translating issues of point of view and positioning from English into the home language for family members); Gutierrez's (Gutierrez, 2014; Gutiérrez and Jurow, 2016; Gutiérrez et al., 2009) design research with children from migrant communities in California using cultural practices around story telling (testimonios) as tools for creating external representations of their critical social and political analyses of ideologies and structures that create barriers for their communities.

The point is that all of these studies are centered in what young people learn through socialization in community specific cultural practices outside of formal schooling, and where the researchers are able to document the underlying epistemological and logical relationships between these community-based practices and those within the disciplines of literary analyses, science, and mathematics as formally represented in the traditions of formal schooling. Equally important, these bodies of research include how these settings of everyday learning in cultural contexts entail attention to social relationships among actors, position novices as efficacious, create spaces in which learners feel safe and perceive what they are learning as relevant, and engage with culturally specific artifacts. In other words, learning in these diverse everyday settings captures the complexity of the underlying principles of human learning and development discussed in the opening sections of this chapter. The underlying principles are universal. However, their enactment and embodiment differ by virtue of differences in routine cultural practices and the differences in culturally informed purposes such practices are intended to support.

References

- Akbar, N., 1979. African roots of Black personality. In: Smith, W.D., Burlew, K.H., Mosley, M.H., Whitney, W.M. (Eds.), *Reflections on Black Psychology*. University Press of America, Washington, DC, pp. 79–98.
- American Psychological Association Task Force on Resilience and Strength in Black Children and Adolescents, 2008. Resilience in African American Children and Adolescents: A Vision for Optimal Development. Retrieved from: <http://www.apa.org/pi/cyfl/resilience.html>.
- Anthony, E.J., Cohler, B.J., 1987. *The Invulnerable Child*. Guilford Press.
- Asante, M.K., Asante, K.W., 1990. *African Culture: The Rhythms of Unity*. Africa World Press, Trenton, NJ.
- Azibo, D.A. (Ed.), 2003. *African Centered Psychology*. Carolina Academic Press, Durham, NC.
- Bang, M., Marin, A., 2015. Nature–culture constructs in science learning: human/non-human agency and intentionality. *J. Res. Sci. Teach.* 52 (4), 530–544.
- Bang, M., Marin, A., Faber, L., Suzukovich III, E.S., 2013. Repatriating indigenous technologies in an urban Indian community. *Urban Educ.* 48 (5), 705–733.
- Bang, M., Medin, D., 2010. Cultural processes in science education: supporting the navigation of multiple epistemologies. *Sci. Educ.* 94 (6), 1008–1026.
- Beattie, J., 1980. Representations of the self in traditional Africa. *Africa* 50, 313–320.
- Berry, J.W., Phinney, J.S., Sam, D.L., Vedder, P., 2006. Immigrant youth: acculturation, identity, and adaptation. *Appl. Psychol.* 55 (3), 303–332.
- Blakey, M.L., 1998. The New York African Burial Ground Project: an examination of enslaved lives, a construction of ancestral ties. *Transform. Anthropol.* 7 (1), 53–58.
- Boykin, A.W., 1986. The triple quandary and the schooling of Afro-American children. In: Neisser, U. (Ed.), *The School Achievement of Minority Children*. Lawrence Erlbaum, Hillsdale, NJ, pp. 57–92.
- Boykin, A.W., Jagers, R.J., Ellison, C.M., Albury, A., 1997. Communalism: conceptualization and measurement of an Afrocultural social orientation. *J. Black Stud.* 409–418.
- Boykin, A.W., Toms, F.D., 1985. Black child socialization: a conceptual framework. In: McAdoo, H.P., McAdoo, J.L. (Eds.), *Black Children: Social, Educational, and Parental Environments*. Sage, Newbury Park, CA, pp. 33–52.
- Bronfenbrenner, U., Morris, P.A., 1998. The ecology of developmental processes. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology: Theoretical Models of Human Development*, fifth ed., vol. 1. Wiley & Sons, NY, pp. 993–1028.
- Bronfenbrenner, U., Morris, P.A., 2006. The bioecological model of human development. *Handb. Child Psychol.* 793–828.
- Brown, B., 2004. Discursive Identity: assimilation into the culture of science and its implications for minority students. *J. Res. Sci. Teach.* 41 (8), 810–834.
- Brown, B., Kloser, M., 2006. Conceptual Continuity & Accessing Everyday Scientific Understandings. *Cultural Studies in Science*. Paper presented at the Annual Meeting of the American Educational Research Association.
- Brown, B.A., Kelly, G.J., 2007. When clarity and style meet substance: language, identity, and the appropriation of science discourse. In: *Science, Learning, Identity*. Brill Sense, pp. 283–299.
- Brown, B.A., Kloser, M., 2009. A view of the tip of the iceberg: revisiting conceptual continuities and their implications for science learning. *Cult. Stud. Sci. Educ.* 4 (4), 921–928.
- Cacioppo, J.T., Berntson, G.G., Sheridan, J.F., McClintock, M.K., 2000. Multilevel integrative analyses of human behavior: social neuroscience and the complementing nature of social and biological approaches. *Psychol. Bull.* 126 (6), 829.
- Cantor, P., Osher, D., Berg, J., Steyer, L., Rose, T., 2018. Malleability, plasticity, and individuality: how children learn and develop in context. *Appl. Dev. Sci.* 1–31.
- Carraher, T.N., Carraher, D.W., Schliemann, A.D., 1985. Mathematics in the street and in school. *Br. J. Dev. Psychol.* 3, 21–29.

- Champion, T., 2003. Understanding Storytelling Among African American Children: A Journey From Africa to America. Lawrence Erlbaum Associates, Mahwah, NJ.
- Chiao, J., Ambady, N., 2007. Cultural neuroscience: parsing universality and diversity across levels of analysis. In: Kitayama, S., Cohen, D. (Eds.), *Handbook of Cultural Psychology*. Guilford, NY, pp. 237–254.
- Chief, K., Emanuel, R.E., Conroy-Ben, O., Torres, R., Pete, S., 2018. Water in the Native World: An Intersection of Indigenous Water Knowledge and Hydrologic Sciences. AGU Fall Meeting Abstracts.
- Cole, M., 1996. Cultural Psychology, A Once and Future Discipline. The Belknap Press of Harvard University Press, Cambridge, MA.
- Cole, M., 2007. Phylogeny and cultural history in ontogeny. *J. Physiol.* 101, 236–246.
- Cross, W., 1991. *Shades of Black: Diversity in African American Identity*. Temple University Press, Philadelphia.
- de los Ríos, C.V., 2019. “los Músicos”: Mexican corridos, the aural border, and the evocative musical renderings of transnational youth. *Harv. Educ. Rev.* 89 (2), 177–200.
- de los Ríos, C.V., Molina, A., 2020. Literacies of refuge: “Pidiendo posada” as ritual of justice. *J. Literacy Res.* 52 (1), 32–54.
- de Los Ríos, C.V., Seltzer, K., 2017. Translanguaging, coloniality, and English classrooms: an exploration of two bicoastal urban classrooms. *Res. Teach. Engl.* 55–76.
- Diop, C.A., 1978. *The Cultural Unity of Black Africa*. Third World Press, Chicago, IL.
- DuBois, W.E.B., 1897. *The Conservation of Races*, vol. 2. American Negro Academy Occasional Papers.
- Dunbar-Ortiz, R., 2014. *An Indigenous Peoples’ History of the United States*, vol. 3. Beacon Press.
- Elder, G.H., 2018. *Children of the Great Depression: Social Change in Life Experience*. Routledge.
- Emdin, C., 2010a. Affiliation and alienation: hip-hop, rap, and urban science education. *J. Curric. Stud.* 42 (1), 1–25.
- Emdin, C., 2010b. *Urban Science Education for the Hip-Hop Generation*. Sense Publishers, Rotterdam.
- Fisher, M.T., 2003. Open mics and open minds: spoken word poetry in African diaspora participatory literacy communities. *Harv. Educ. Rev.* 73 (3), 362–389.
- Fisher, M.T., 2005. From the coffee house to the school house: the promise and potential of spoken word poetry in school contexts. *Engl. Educ.* 37 (2), 115–131.
- Fisher, M.T., 2007. *Writing in Rhythm: Spoken Word Poetry in Urban Classrooms*. Teachers College Press, NY.
- Gobineau, A., 1853. *Didot frères. Essai sur l’inégalité des races humaines*, vol. 1. Rumpler, Hanovre.
- Goddard, H.H., 1912. *The Kallikak Family*. Macmillan Company, NY.
- Gopnic, A., Meltzoff, A., Kuhl, P., 1999. *The Scientist in the Crib: What Early Learning Tells Us About the Mind*. Harper Collins, NY.
- Gould, S.J., 1981. *The Mismeasure of Man*. W.W. Norton, NY.
- Grande, S., San Pedro, T., Windchief, S., Salett, E., Koslow, D., 2015. Indigenous Peoples and Identity in the 21st Century: Remembering, Reclaiming, and Regenerating. In: *Multicultural Perspectives on Race, Ethnicity, and Identity*, pp. 105–122.
- Gutiérrez, K., 2014. The science of justice: migrant youth and sociocritical literacies in science. In: Harper, S. (Ed.), *The Elusive Quest for Civil Rights in Education*. Center for the Study of Race and Equity: University of Pennsylvania, Philadelphia, PA, pp. 12–13.
- Gutiérrez, K., Rogoff, B., 2003. Cultural ways of learning: individual traits or repertoires of practice. *Educ. Res.* 32 (5), 19–25.
- Gutiérrez, K.D., 2016. 2011 AERA presidential address: designing resilient ecologies: social design experiments and a new social imagination. *Educ. Res.* 45 (3), 187–196.
- Gutiérrez, K.D., Jurow, A.S., 2016. Social design experiments: toward equity by design. *J. Learn. Sci.* 25 (4), 565–598.
- Gutiérrez, K.D., Morales, P.Z., Martinez, D.C., 2009. Re-mediating literacy: culture, difference, and learning for students from nondominant communities. *Rev. Res. Educ.* 33 (1), 212–245.
- Han, S., Northoff, G., Vogeley, K., Wexler, B.E., Kitayama, S., Varnum, M.E., 2013. A cultural neuroscience approach to the biosocial nature of the human brain. *Annu. Rev. Psychol.* 64, 335–359.
- Hancock, A.M., 2005. DuBois and the scientific study of race: W.E.B. DuBois intellectual forefather of intersectionality. *Souls* 7 (3–4), 74–84.
- Hernstein, R., Murray, C., 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press.
- Ignatiev, N., 1996. *How the Irish Became White*. Routledge, NY.
- Immordino-Yang, M.H., Damasio, A., 2007. We feel, therefore we learn: the relevance of affective and social neuroscience to education. *Mind Brain Educ.* 1 (1), 3–10.
- Jensen, A., 1969. How much can we boost IQ and scholastic achievement. *Harv. Educ. Rev.* 39, 1–123.
- Jensen, E., 2009. *Teaching With Poverty in Mind: What Being Poor Does to Kids’ Brains and What Schools Can Do About It*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Jones, R.L., 1980. *Black Psychology*, second ed. Harper & Row Publishers, New York.
- Karenga, M., 2003. The pan-African initiative in the Americas: culture, common struggle and the Odu Ifa. *Race Democr. Am.* 156–172.
- Kimmerer, R.W., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions, Minneapolis, MN.
- King, J.R., 1976. African survivals in the Black community: key factors in stability. *J. Afro-Am. Issues* 4 (2), 153–167.
- Kuhl, P.K., Meltzoff, A.N., 1996. Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100 (4), 2425–2438.
- Lave, J., 1977. Cognitive consequences of traditional apprenticeship training in West Africa. *Anthropol. Educ. Q.* 8 (3), 177–189.
- Lave, J., 1988. *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge University Press, Cambridge.
- Lee, C.D., 1995. A culturally based cognitive apprenticeship: teaching African American high school students’ skills in literary interpretation. *Read. Res. Q.* 30 (4), 608–631.
- Lee, C.D., 2001. Is October Brown Chinese: a cultural modeling activity system for underachieving students. *Am. Educ. Res. J.* 38 (1), 97–142.
- Lee, C.D., 2007. *Culture, Literacy and Learning: Taking Bloom in the Midst of the Whirlwind*. Teachers College Press, NY.
- Lee, C.D., 2010. Soaring above the clouds, delving the ocean’s depths understanding the ecologies of human learning and the challenge for education science. *Educ. Res.* 39 (9), 643–655.
- Lee, C.D., 2017. Expanding visions of how people learn: the centrality of identity repertoires. *J. Learn. Sci.* 26 (3), 517–524.
- Lee, C., Meltzoff, A., Kuhl, P., 2020. The braid of human learning and development: neuro-physiological processes and participation in cultural practices. In: Nasir, N.S., Lee, C.D., Pea, R., de Royston, M.K. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, New York, NY, pp. 24–43.
- Luria, A.R., 1928. Psychological expedition to central Asia. *Sci. Educ.* 74, 383–384.
- Majors, Y., 2001. Passing mirrors: subjectivity in a mid-western hair salon. *Anthropol. Educ. Q.* 32, 116–130.
- Majors, Y., 2003. Shoptalk: teaching and learning in an African American hair salon. *Mind Cult. Activ.* 10 (4), 289–310.
- Majors, Y., 2015. *Shop Talk*. Teachers College Record, NY.
- Makondo, C.C., Thomas, D.S., 2018. Climate change adaptation: linking indigenous knowledge with western science for effective adaptation. *Environ. Sci. Pol.* 88, 83–91.
- Markus, H., Kitayama, S., 1991. Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253.
- Meltzoff, A.N., 1988. Imitation, objects, tools, and the rudiments of language in human ontogeny. *Hum. Evol.* 3.
- Meltzoff, A.N., 2013. Origins of social cognition: bidirectional self-other mapping and the “Like-Me” hypothesis. In: Banaji, M.R., Gelman, S.A. (Eds.), *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*. Oxford University Press, NY, pp. 139–144.
- Mills, C.W., 1997. *The Racial Contract*. Cornell University Press, Ithaca, NY.
- Nasir, N., 2000. “Points ain’t everything”: emergent goals and average and percent understandings in the play of basketball among African American students. *Anthropol. Educ.* 31 (1), 283–305.
- Nasir, N., 2002. Identity, goals, and learning: mathematics in cultural practice. In: Nasir, N., Cobb, P. (Eds.), *Mathematical Thinking and Learning: Special Issue on Diversity, Equity and Mathematics Learning*, vol. 4, pp. 211–247.
- Nasir, N., 2005. Individual cognitive structuring and the sociocultural context: strategy shifts in the game of dominoes. *J. Learn. Sci.* 14, 5–34.
- Nasir, N.I.S., Peele-Eady, T.B., 2012. Identity and learning. *Encycl. Sci. Learn.* 1482–1484.

- Nasir, N.i., Lee, C.D., Pea, R.D., McKinney de Royston, M. (Eds.), 2020. *Handbook of Cultural Foundations of Learning*. Routledge, NY.
- Nobles, W., 1972. African philosophy: foundations for Black psychology. In: Jones, R. (Ed.), *Black Psychology*. Harper & Row, NY, pp. 18–32.
- Nunes, T., Schliemann, A.D., Carraher, D.W., 1993. *Street Mathematics and School Mathematics*. Cambridge University Press, New York.
- Orellana, M.F., 2001. The work kids do: Mexican and Central American immigrant children's contributions to households and schools in California. *Harv. Educ. Rev.* 71 (3), 366–390.
- Orellana, M., 2009. *Translating Immigrant Childhoods: Children's Work as Culture and Language Brokers*. Rutgers University Press, New Brunswick, NJ.
- Orellana, M., Reynolds, J., Dorner, L., Meza, M., 2003. In other words: translating or "para-phrasing" as a family literacy practice in immigrant households. *Read. Res. Q.* 38 (1), 12–34.
- Osher, D., Cantor, P., Berg, J., Steyer, L., Rose, T., 2018. Drivers of human development: how relationships and context shape learning and development. *Appl. Dev. Sci.* 1–31.
- Pass, S., 2004. *Parallel Paths to Constructivism: Jean Piaget and Lev Vygotsky*. Information Age Publishing, Greenwich, Conn.
- Pass, S., 2007. When constructivists Jean Piaget and Lev Vygotsky were pedagogical collaborators: a viewpoint from a study of their communications. *J. Constr. Psychol.* 20 (3), 277–282.
- Payne, R.K., 1999. *A Framework for Understanding and Working With Students and Adults From Poverty*. RFT Publishing.
- Phillips, S.U., 1983. *The Invisible Culture: Communication in Classroom and Community on the Warm Springs Indian Reservation*. Longman, NY.
- Piaget, J., 1952. *The Origins of Intelligence in Children*. M. Cook, Trans. International Universities Press, NY.
- Piaget, J., 1966/1974. Need and significance of cross-cultural studies in genetic psychology. In: Berry, J.W., Dasen, P.R. (Eds.), *Culture and Cognition: Readings in Cross-Cultural Psychology*. Methuen, London.
- Piaget, J., 1975. *The Child's Conception of the World*. Littlefield Adams, Totowa, NJ.
- Rogoff, B., 1995. Observing sociocultural activity and three planes: participatory appropriation, guided participation, and apprenticeship. In: Wertsch, J., del Rio, P., Alvarez, A. (Eds.), *Sociocultural Studies of Mind*. Cambridge University Press, New York, pp. 139–164.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press, NY.
- Rogoff, B., 2014. Learning by observing and pitching in to family and community endeavors: an orientation. *Hum. Dev.* 57 (2–3), 69–81.
- Rogoff, B., Chavajay, P., 1995. What's become of research on the cultural basis of cognitive development. *Am. Psychol.* 50 (10), 859–877.
- Rudman, L.A., Ashmore, R.D., Gary, M.L., 2001. "Unlearning" automatic biases: the malleability of implicit prejudice and stereotypes. *J. Pers. Soc. Psychol.* 81 (5), 856.
- Russo, E.A., Martienssen, R.A., Riggs, A.D. (Eds.), 1996. *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Saxe, G.B., 1988a. Candy selling and math learning. *Educ. Res.* 17 (6), 14–21.
- Saxe, G.B., 1988b. The mathematics of child street vendors. *Child Dev.* 1415–1425.
- Saxe, G., Esmonde, I., 2005. Studying cognition in the flux: a historical treatment of FU in the shifting structure of Oksapmin mathematics. *Mind Cult. Activ.* 12 (3 & 4), 171–225.
- Scribner, S., Cole, M., 1981. *The Psychology of Literacy*. Harvard University Press, Cambridge, MA.
- Serpell, R., 1974. Aspects of intelligence in a developing country. *Afr. Soc. Res.* 17, 578–596.
- Serpell, R., 1982. Measures of perception, skills, and intelligence. *Rev. Child Dev. Res.* 6, 392–440.
- Spencer, M.B., 2006. Phenomenology and ecological systems theory: development of diverse groups. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology*, sixth ed., vol. 1. Wiley, NY, pp. 829–893.
- Spencer, M.B., 2019. Developmental and Intersectional Insights About Diverse Children's Identity. *Fla. L. Rev. Forum*.
- Spencer, M.B., Harpalani, V., Cassidy, E., Jacobs, C., Donde, S., Goss, T.N., 2006. Understanding vulnerability and resilience from a normative development perspective: implications for racially and ethnically diverse youth. In: Chicchetti, D., Cohen, E. (Eds.), *Handbook of Developmental Psychopathology*, vol. 1. Wiley, Hoboken, NJ.
- Super, C., Harkness, S., 1986. The developmental niche: a conceptualization at the interface of child and culture. *Int. J. Behav. Dev.* 9, 545–569.
- Taylor, E.V., 2009. The purchasing practice of low-income students: the relationship to mathematical development. *J. Learn. Sci.* 18 (3), 370–415.
- Taylor, E.V., 2012. Supporting children's mathematical understanding: professional development focused on out-of-school practices. *J. Math. Teach. Educ.* 15 (4), 271–291.
- Tomasello, M., 1999. *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.
- Traub, J., 1999. What no school can do. *The New York Times Magazine*, January 16, 52–56.
- Tyack, D., 1974. *The One Best System: A History of American Urban Education*. Harvard University Press, Cambridge, MA.
- Valdes, G., 1996. *Con Respeto: Bridging the Distances Between Culturally Diverse Families and Schools*. Teachers College Press, NY.
- Valdes, G., 2001. *Learning and Not Learning English: Latino Students in American Schools*. Teachers College Press, NY.
- Valdes, G., 2002. Expanding the Definitions of Giftedness: The Case of Young Interpreters From Immigrant Countries. Lawrence Erlbaum, Mahwah, NJ.
- Vossoughi, S., Gutiérrez, K.D., 2014. Studying movement, hybridity, and change: toward a multi-sited sensibility for research on learning across contexts and borders. *Teach. Coll. Res.* 116 (14), 603–632.
- Vygotsky, L.S., 1977. The development of higher psychological functions. *Sov. Psychol.* 15 (3), 60–73.
- Vygotsky, L., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Vygotsky, L., 1981. The genesis of higher mental functions. In: Wertsch, J. (Ed.), *The Concept of Activity in Soviet Psychology*. M.E. Sharpe, Armonk, NY.
- Wartofsky, M., 1973. *Models*. D. Reidel, Dordrecht.
- Williamson, J.A., Rhodes, L., Dunson, M., 2007. A selected history of social justice in education. *Rev. Res. Educ.* 31, 195–224.
- Winfield, A.G., 2007. *Eugenics and Education in America: Institutionalized Racism and the Implications of History, Ideology, and Memory*, vol. 18. Peter Lang.
- Winn, M.T., 2010. "Our side of the story": moving incarcerated youth voices from margins to center. *Race Ethn. Educ.* 13 (3), 313–325.
- Zajonc, R.B., Marcus, H., 1984. Affect and cognition. In: Izard, C.E., Kagan, J., Zajonc, R.B. (Eds.), *Emotions, Cognition and Behavior*. Cambridge University Press, Cambridge, England, pp. 73–102.

Indigenous student identity development and school success: mitigating social identity and stereotype threats in schools

Melinda Webber and Abby Madden-Smith, Faculty of Education, University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	276
Social identity theory	276
SIT main influences for Indigenous students	277
Stereotype threat	278
Stereotype threat and Indigenous students	278
Indigenous social-psychological theories	279
Māori student success and the mana model	280
Mana whānau—connectedness to other and collective agency	281
Mana ūkaipō—belonging and relationship to place	281
Mana motuhake—embedded achievement and self concept	282
Mana tū—social psychological competence	282
Mana tangatarua—diverse knowledge base and skillset	283
Conclusion	284
References	284

Introduction

Indigenous perspectives of identity development stress that identity formation is situational, and impacted by one's sense of belonging to a collective and connectedness to other factors in the context, including people, places, and spaces (McIntosh, 2006; Webber, 2015). In this chapter we argue that Indigenous student academic identities are developed, experienced, and expressed in unique ways by Indigenous students. Drawing from Social Identity Theory, Stereotype Threat, and Indigenous Social-Psychological Theory this chapter explains how schools can create environments that enable Indigenous students to succeed alongside other students. Social Identity Theory (SIT) (Tajfel, 1981) posits that a student's sense of who they are is highly influenced by their group membership(s) and student motivation at school relates to whether those groups are negatively or positively represented in society more broadly. This links to what Steele (1997, 2010; Steele and Aronson, 1995) and others refer to as "stereotype threat" where generalized societal perceptions (stereotypes) of social groups manifest as a social-psychological threat that can hinder a negatively stereotyped student's performance and engagement at school. Both SIT and Stereotype Threat are explored in this chapter, along with their influence on the engagement of Indigenous students in schooling contexts.

Indigenous social-psychological theory suggests that in order for Indigenous students to succeed in education, more attention must be paid to their identities, particularly their ethnic and cultural identities and what this means for them, their whānau and wider communities (González et al., 2022; Houkamau and Sibley, 2015; Webber and Macfarlane, 2020). This perspective is framed within research that argues for a holistic and culturally rich approach to teaching and learning. The Mana Model presented in the second half of this chapter, developed from an Indigenous Māori worldview, provides a framework for understanding the important conditions that Indigenous students need to succeed and thrive at school. These conditions include connectedness to others and collective agency (mana whānau); belonging and relationship to place (mana ūkaipō); embedded achievement and self-concept (mana motuhake); social psychological competence (mana tū); and a diverse knowledge base and skill set (mana tangatarua). The Mana Model suggests that Indigenous student thinking, behavior and mental wellbeing is motivated by the desire to achieve a sense of mana—self-efficacy, purpose, pride and belonging. This chapter argues that when schools focus on these five conditions for mana development for all students, they are able to learn in a way that supports their wellbeing, academic persistence, and confidence. The Mana model, drawn from Indigenous social psychology, is underpinned by the belief that all students possess valuable knowledge, skills, and attributes that they can utilize to succeed at school and contribute positively to the world around them.

Social identity theory

Social identity theory (SIT) is a theory that explores intergroup relations and social identity. The work of Tajfel (1969, 1974, 1978, 1981; Tajfel et al., 1971) and his graduate student Turner (Tajfel and Turner, 1979; Turner, 1975) added a valuable social dimension to social psychology, which predominantly focused on individual and inter-individual processes at the time. Tajfel (1972) defined social identity as an "individual's knowledge that he belongs to certain social groups together with some emotional and value

significance to him of this group membership" (p. 272). SIT was predicated on the idea that the wider social context, and an individual's social identities, are significant factors in intergroup behavior and self-concepts. SIT seeks to understand how one's social identity is constructed, internalized and expressed.

A key element of this process involves social categorization, a process whereby individuals categorize both themselves and others into groups. Factors that might influence categorizations include ethnicity, age, and gender as well as occupation and socioeconomic status. [Tajfel and Turner \(1979\)](#) proposed that social categorizations are a way to organize and stratify social contexts and help individuals to define and place oneself within social settings. Additionally, being a member of a social group comes with ingroup defining expectations of how to think and act. [Tajfel \(1981\)](#) states,

in an infinite variety of situations throughout his life, an individual feels, thinks and behaves in terms of his social identity created by the various groups of which he is a member and in terms of his relation to the social identity of others, as individuals or *en masse*.

[Tajfel \(1981, p. 31, emphasis in original\).](#)

[Turner \(1987\)](#) and colleagues ([Onorato and Turner, 2004](#); [Turner and Reynolds, 2012](#)) later clarified the categorization process as Social Categorization Theory (SCT). SCT posits self-concept as situational, in that all social contexts compel individuals to either emphasize or de-emphasize particular social identities in order to construct a salient "ingroup" identity relevant to the context ([Onorato and Turner, 2004](#); [Turner and Reynolds, 2012](#)). This ingroup identification consequently influences an individual's thinking and behavior in that space ([Turner and Reynolds, 2012](#)).

SIT proposes that individuals are motivated by a desire for a positive self-esteem and that this drives group membership. A central idea within SIT is that individuals are inherently motivated to achieve a positive sense of social identity ([Tajfel, 1974](#); [Tajfel and Turner, 1979](#)). A positive sense of social identity develops when an individual belongs to certain social group(s) that are distinguishable and invoke a feeling of positive membership ([Tajfel, 1978](#)). Based on a large and continual body of research SIT posits that group categorization, coupled with the desire for a positive social identity, is an explanation for intergroup conflict, discrimination, and prejudice ([Hogg, 2016](#); [Turner, 1975](#)) and the prevalence and reproduction of unequal group power relations in society ([Abrams and Hogg, 2010](#)). [Hogg \(2000\)](#) also argues that social identity is motivated by the human desire for subjective certainty—understanding how to think and feel, knowing how to behave and what to expect in different situations and relationships. This argument posits that individuals are motivated to self-categorize with a clear and distinct group in order to alleviate identity uncertainty and view and understand the self in comparison to others ([Abrams and Hogg, 2010](#); [Hogg, 2000](#)).

The stereotypes associated with one's social group can impact whether membership in that group contributes positively or negatively to an individual's social identity and self-esteem. [Tajfel \(1969\)](#) theorized that stereotypes are assigned to social groups to create clear distinctions between groups, and that they are culturally specific and evolve throughout history ([Tajfel, 1969](#)). [Tajfel \(1969\)](#) stated:

through personal and cultural experience, dimensions such as "intelligent," "lazy," or "honest" are subjectively associated with classifications of people into groups. As long as we have little specific knowledge about an individual, we shall tend to ascribe to him the characteristics which we derive from our knowledge of his class membership, be it a class of trade unionists, undergraduates, animal lovers, or Patagonians (p. 178).

He goes on to say that at both an individual and group level, once these stereotypes are assigned, finding evidence to support the stereotype is far easier than noticing their contradictions. Stereotypes contribute to the overall social perception of a group, influencing the status of different groups and the positive or negative influence being a member of a group has on an individual's self-perception. For Indigenous students the stereotypes associated with their social group can be negative given they are often ascribed by the dominant group that, as posited by SIT, is motivated to maintain their own positive self-identity by negatively stereotyping others.

SIT main influences for Indigenous students

In educational contexts, stereotypes and in-group and out-group categorizations can impact Indigenous students' academic motivation and engagement. It does so by disrupting the identity-motivation process and influencing the parameters within which they can imagine their possible selves and any strategies needed to attain success. [Oyserman \(2007\)](#) has asserted that motivation is fueled by a desire to (a) progress from a current-self state to a future-self goal (or possible self), and (b) away from possible negative selves. [Oyserman \(2007\)](#) argued that this is a culturally located process as culture influences desirable goals, behaviors, and strategies to succeed. Furthermore, research has shown that behaviors conducive to academic achievement (such as studying and in-class questioning) have become infused with the dominant group due to the association between success at school and success and power in wider society ([Oyserman, 2007](#); [Ogbu, 2004](#); [Burnaford, 2015](#)). The educational behaviors accepted in wider society may not align with Indigenous student's social identities (including racial-ethnic identity, gender and class-based identities), reducing the

likelihood that Indigenous students will be motivated to participate fully in these behaviors (Oyserman and Lewis, 2017). Certain behaviors are associated with what “people like me” do and when these identities are salient it inspires group-defining behavior, that is education engagement or dis-engagement (Burnaford, 2015; Oyserman, 2007; Oyserman et al., 1995; Oyserman and Lewis, 2017).

Indigenous students face additional challenges as they may be less likely to imagine their possible academic selves (defined by the education system) while still remaining central members of their social groups (Burnaford, 2015). This is because an Indigenous student’s academic self may conflict with their social identity, particularly if their social identity has been constructed in opposition to the dominant group (Fordham and Ogbu, 1986). An individual can have multiple social identities and when they are not compatible, certain social identities may dominate and “override” others (Oyserman, 2007). Verkuyten et al. (2019) argued that ethnic-racial discrimination (stereotypes and prejudice based on ethnic or racial group) threatens core components of identity (belonging, esteem and control). These authors have proposed that when identity is threatened students can respond with increased or decreased academic engagement.

Indigenous students who are aware of negative societal attitudes toward their group (such as being stereotyped as low achievers) may attempt to “leave” or distance themselves from their social group or categorizations, and join or be accepted by another group through certain behaviors such as academic engagement (Tajfel and Turner, 1979). This “strategy of racelessness” (Fordham, 1988, p. 54) means that Indigenous students feel as though they must distance from their social groups in order to succeed in an individualistic education system. For many Indigenous students this strategy is difficult because (a) physical features may render groups impermeable (Abrams and Hogg, 2010) and (b) disconnecting from one’s ethnic or cultural group and engaging in out-group behaviors may result in discrimination from in-group members (Fordham and Ogbu, 1986; Ogbu, 2004; Carter, 2006). In essence, Indigenous students who experience negative stereotypes and/or self-concept in educational contexts, or who feel as though their identities are not congruent with academic behaviors and success (Oyserman and Lewis, 2017), are more likely to withdraw from school in search of a different positive social identity (Abraham, 1989; Kelly, 2007; Tajfel and Turner, 1979; Webber, 2015). Constructing a group identity with oppositional references and behaviors to the dominant group is a minority identity protective factor (Fordham and Ogbu’s, 1986). Within educational contexts, the temporary reprieve from one’s negative social identity and its influences on self-esteem has long term consequences for Indigenous students’ futures and contributes to the perpetuation of underachievement for certain groups.

Stereotype threat

Despite many Indigenous students and their families valuing education and having high academic aspirations for themselves or their children (Carter, 2006; Graham et al., 2010; Macfarlane et al., 2014; Mello, 2009; Rudolphi and Salikutluk, 2021; Webber and Macfarlane, 2020) there remains an inconsistency between Indigenous educational aspirations and academic outcomes. Identity motivation theory posits that many students envisage academic possible selves, the issue is whether these possible future identities are salient and connected to effective strategies and help students persevere through difficulty (Oyserman and Lewis, 2017). Stereotype threat is helpful to understand how stereotypes about one’s identity may disrupt this process and be an obstacle for Indigenous students learning and motivation. Steele and Aronson (1995) first coined the term stereotype threat, defining it as “being at risk of confirming, as self-characteristic, a negative stereotype about one’s group” (p. 797) “both to one’s self and to others who know the stereotype” (p. 807). Steele (2010) stated that “stereotype threat is a standard predicament of life” (p. 5) because humans are critically aware of each other, we know how our identities correlate to stereotypical ideas, including the assumptions surrounding gender, age, wealth and ethnicity. Thus, when Indigenous students are presented with a situation at school in which they might be judged because of who they are or are at risk of confirming existing stereotypes about element(s) of their identity, there is a social-psychological threat at play which can hinder their performance (Steele, 1997). Appel and Kronberger (2012) also add that stereotype threat can hinder individuals “from building knowledge and abilities in the first place” (p. 630), making stereotype threat a key issue in education.

Stereotype threat and Indigenous students

The achievement gap between Indigenous and other ethnic minority students and their dominant (often White) peers is evident globally, for instance: Black and Hispanic students in the United States (De Brey et al., 2019); students from Black backgrounds in England (Hutchingson et al., 2020); and Indigenous and Aboriginal students in Australia (Fahey, 2021). In Aotearoa New Zealand, while there have been considerable improvements in the past 10 years, the academic achievement gap between Māori and Pacific students and their European peers persists (NZQA, 2020). Despite achievement gaps being influenced by a myriad of socioeconomic factors (Reardon et al., 2019), as well as being influenced by the values within the education system itself (Webber, 2012), many subgroups are stereotyped as groups that will underachieve “in certain cognitive tasks” (Appel and Kronberger, 2012, p. 610). Persistent stereotypes within the education system in Aotearoa New Zealand frame Māori and Pacific students in deficit ways and make assumptions about the families and home lives of these students to explain underachievement (Turner et al., 2015).

Stereotyped students are aware of societal assumptions about their group. Webber et al. (2013) study found that Māori students were able to clearly identify academic stereotypes related to Māori, including “the perceived notion that Māori are ‘not very smart’, ‘dumb’ and ‘drop out’” (p. 50). When Indigenous students encounter spaces where they feel their stereotypes are highlighted, academic achievement and engagement can be negatively impacted (Webber and Macfarlane, 2020). Steele and Aronson (1995) demonstrated that African American students noticeably underperformed their White peers when they sat test under conditions that elevated “negative stereotypes about their group’s intellectual ability” (p. 807). They showed that when students were told that their scores were indicative of their intellectual ability or were asked to record their ethnicity, stereotype threat was heightened to the detriment of the outcomes in the test. Conversely, when conditions alleviated stereotype threat (i.e., tests were not presented as being diagnostic of ability or did not ask students to record their ethnicity) African American students performed to the same standard as White students. This demonstrates that stereotype threat has a marked impact on test performance and can result in an achievement gap between negatively stereotyped students and those who are not. As well as ethnic minority groups, research has illustrated the impact of stereotype threat on the cognitive performance of physically disabled students (Desombre et al., 2018) and girls’ and women’s achievement and engagement in mathematics (Steele, 1997; Good et al., 2008), highlighting how stereotype threat can affect many different groups in an array of contexts where stereotypes apply.

Stereotype threat impacts the personal, social and institutional lives of Indigenous students, creating an added obstacle to personal thriving. A further consequence of stereotype threat is that Indigenous students may feel they must compromise or minimize elements of themselves (Webber and Macfarlane, 2020) or redirect their education because of the threat that existing stereotypes have for their self-esteem and identities (Steele, 1997). Steele (1997) speaks to the difficulty of stereotype threat, pointing out that the more students are invested in a particular domain, and the greater influence this domain has on their identity, the higher their risk of stereotype threat. Thus, stereotype threat predominantly impacts Indigenous students that have both the interest and skills to achieve at school. The risk is that these students will disengage from domains where they feel stereotyped, tragically impacting the “academic identities and performance outcomes of large segments of society” (Steele, 1997, 614). An example of this is the disproportionate number of Indigenous scientists (McAllister, 2022) and academics working in universities (McAllister et al., 2019). In essence, Indigenous students are more likely to disidentify and disengage from domains in which they experience a threat and remove the influence these domains have on their identities (Good et al., 2007).

This perspective on the obstacles facing stereotyped students places emphasis on social-cultural change, deeming that school environments (including teacher expectations) and implicit and explicit messaging within these environments are crucial factors that can help and hinder student engagement and achievement. Steele (1997) offers the hopeful perspective that alleviating stereotype threat in education settings will reduce its impact on student decision making and achievement.

Indigenous social-psychological theories

Māori, the Indigenous people (tangata whenua) of Aotearoa understand identity formation as a process underpinned by the concepts *mana* and *mauri*. Mana is defined as honor, pride and esteem (Royal, 2006). Mana has a both an individual and social component. It involves an individual’s sense of self, and the degree in which they feel confident, empowered, and well. Mana is also attributed to an individual through the perceptions others have of them based on their contributions to their communities, and the ways they positively impact the lives of others around them (Royal, 2006; Dell, 2017). Royal (2006) explains that due to colonization Māori have experienced painful forced separation from their land and disenfranchisement throughout history that has resulted in a collective dulling of mana Māori. Thus, he argues, mana restoration must occur at an individual and collective level. For individuals, this process is heterogeneous with some developing their Māori identity through connection with Māori communities (iwi, hapu), Māori knowledge, history and whakapapa (genealogy), whereas for others this process may involve fulfilling relationships with non-Māori.

This process of developing mana is linked to another key Māori concept called “mauri” (Royal, 2006). Mauri is a spiritual “active force of life” (Mead, 2006, p. 83) that empowers the body, mind and spirit of all individuals. According to a Māori worldview, mauri is infused within all living beings, including the natural environment (land, trees, water, and animals) and people. Durie (2001) explains that mauri flows between these elements mediating their relationship and continuing their mauri force. The intrinsic mauri force energizes the pursuit of mana as well as the journey to belonging, wellbeing, and a sense of purpose. These concepts are central to Māori identity construction and align with theories of becoming.

From an Indigenous perspective, identity is forged through a negotiation between past, present, and future. For Indigenous populations this positioning is defined and shaped, to varying degrees, by colonization (McInerney and Flowerday, 2016). Thus, identity construction for many Indigenous populations is influenced by the struggle to survive as a minority ethnic group in their own lands, often subjected to marginalization and direct attempts at assimilation (Alfred and Cornthassel, 2005). McIntosh (2006) explored the complexity of Māori identity, considering the “fixed,” “fluid,” and “forced” ways in which contemporary Māori identities are constructed. While she states that “Māori identity is a marginal identity in Aotearoa/New Zealand” (p. 70, emphasis in original) she argues that marginality is expressed in many forms and that there is no universal process of Māori identity construction nor homogenous Māori identity.

Firstly, Māori identity can be constructed through traditional characteristics. “Traditional” indicators of Māori identity such as kin, Māori community, cultural participation, knowledge of whakapapa and Māori history, and use of te reo Māori are examples of characteristics from a Māori ecology (Moeke-Pickering, 1996) also referred to as “primordial” constructs of Māori identity (Webber, 2007). Whakapapa—the layering of generations, tribal stories and traditions that link Māori to the past and to each other (Connor, 2019; Te Rito, 2007), is “generally agreed to be the key characteristic” of Māori identity (Webber, 2012, p. 22). Te Rito (2007) writes, “[h]aving knowledge of whakapapa helps ground us to the earth. We have a sense of belonging here, a sense of purpose, a *raison d’être* which extends beyond the sense of merely existing on this planet” (p. 4). Traditional Māori informed ideas about what constitutes a Māori identity is important to assert positive Māori-centric perspectives and progress political activism (Hoskins, 2012; McIntosh, 2006). However, the classification of Māori identity according to “fixed” characteristics (Metge, 1964; Ritchie, 1963) have been criticized for being exclusory, overly simplistic, and restrictive (McIntosh, 2006; Walker, 1989). Durie (1995) argued that these traditional characteristics of Māori ethnic identity will be of varying importance to modern Māori, and do not influence any individual’s status as Māori.

Māori scholars have also argued that Māori identity construction is contextually influenced (Durie, 1995; Webber, 2007; Houkamau, 2011). Kukutai (2004) points out that defining Indigenous populations across the world for policy purposes is difficult, given the variety of expressions of Indigenous identity and the diverse lifestyles, social economic status, experiences, and attitudes within Indigenous communities. “Fluid” Indigenous identities emerge in response to context and are generally inclusive of diverse expressions of Māori identity. As McIntosh (2006) stated “[t]he fluid identity ‘plays’ with cultural markers such as language, custom, and place and reconfigures them in a way that gives both voice and currency to their social environment” (p. 83). The current political and social climate, socio-economic status, lifestyle choices and experiences influence what being Māori looks and feels like at an individual and group level. Factors such as land alienation and political struggle for resources have impacted Māori identity and potentially restricted access to some primordial characteristics (Moeke-Pickering, 1996; Greaves, 2017). Thus, this form of identity construction establishes Māori “roots” in the diverse lives of Māori people, establishing community and culture in more flexible ways.

Lastly, McIntosh (2006) speaks to “forced” identity construction whereby outsider group perceptions dominate identity formation. In the case of Māori, forced identities are shaped by the lack of agency Māori have to define themselves and further constrained by the reality of disadvantage that many Māori experience. Māori experience much higher rates of racial discrimination than Pākehā (Cormack et al., 2018), have lower educational attainment (Ministry of Education, 2020a,b,c), suffer poorer health outcomes (Ministry of Health, 2019; Statistics NZ, 2021), including mental health (Fleming et al., 2020; Paterson et al., 2018), child poverty (Duncanson et al., 2021), unemployment (Ministry of Business, Innovation and Employment, 2017), and prison rates (Department of Corrections, 2021). These statistics contribute to a societal narrative about what it means to be Māori, and the reality is that many Māori inherit these negative beliefs about their own identity (McIntosh, 2006).

Evidently, Indigenous theorization of identity construction includes multiple perspectives, forms and expressions. There is no singular way to “be” Indigenous and thus instead literature explores how Indigenous identities interact with tradition, context, and with the multitude of other identities that individuals hold. Through a Māori worldview all individuals have an innate energy force (*mauri*) to pursue honor, pride, and esteem (*mana*) (Royal, 2006). This chapter asserts that educators, educational researchers, and policymakers must come to understand the integral nature of *mana* and *mauri* to fully meet the educational needs to Indigenous Māori students. They must also reflect on how they can be activated and developed in classrooms and school contexts. A focus on the development of *mana*, and the activation of *mauri*, may go some way to helping Indigenous Māori students to feel positive about their social and academic identities—while keeping their ethnic and cultural identities intact.

Māori student success and the *mana* model

Māori perspectives take a holistic view of student success recognizing that most definitions of success are often defined by the dominant group (Milne, 2009). Māori students must be seen as capable of succeeding and as having the *mauri* (energy), and *mana* (intrinsic interests and strengths) that enable them to thrive at school and in classroom environments. Durie (2011) states,

Indigenous success, often the product of resilience, has several dimensions but essentially encompasses individuals and groups, a capacity for positive engagement, and a level of autonomy. Much of the literature on resilience centres on the potential of individuals to overcome personal trauma and succeed. However, resilience is also about the achievements of collectives: families, *whānau*, communities, tribes, mobs, and whole populations. Success in that sense is a shared experience which reflects an ability to adapt and a propensity for turning adversity into accomplishment (p. 30).

The *Mana Model* (Webber and Macfarlane, 2020) is an example of an Indigenous model of student success that places emphasis on Māori student identities, effective classroom conditions for learning, enduring teacher-student-family relationships, and localized curriculum. The *Mana Model* was developed through research with successful Māori students, their *whānau*, and community to identify the personal and contextual factors that contributed to their success (Macfarlane et al., 2014). The model asserts that students are “motivated by the desire to achieve a sense of *mana*” (Webber and Macfarlane, 2020, p. 35) which lies at the heart of student identity, self-concept, and their sense of belonging. The *Mana Model* includes five *mana*-based conditions that enable

student success from an Indigenous Māori perspective: Mana Whānau (connectedness to others and collective agency), Mana Ūkaipō (belonging and relationship to place), Mana Motuhake (embedded achievement and self-concept), Mana Tū (social psychological competence), and Mana Tangatarua (diverse knowledge base and skillset).

Mana whānau—connectedness to other and collective agency

Mana whānau is an important condition for Māori student success. The word *whānau* refers to familial relationships in the Māori language. Mana whānau is derived from connectedness to others and a sense of collective agency. This dimension aligns with Indigenous research that expresses the foundational nature of whānau and close personal relationships with others for Māori identity development (Houkamau and Sibley, 2010) in both traditional and contemporary times (Moeke-Pickering, 1996). Research shows that whānau and social relationships are central to a range of Māori identities such as those of women (Houkamau, 2011), men (Rua et al., 2017), youth in South Auckland (Borell, 2005), and students (Macfarlane et al., 2014). The importance of social well-being for student self-concept, motivation, and behavior is also supported by social psychological theories. These theories state that humans have a fundamental “need-to-belong” and that self-esteem develops according to levels of belongingness and the degree to which one believes they deserve to belong (Baumeister and Leary, 1995; Leary and Baumeister, 2000; Allen et al., 2021). Elliot et al. (2004) also argue that a sense of mattering—feeling significant in their relationships with others—is important for self-concept. Thus, students benefit from reciprocal relationships where they feel they have the skills to help and uplift others. Within this, relationships where students feel connected to others (relational) and appreciated as unique individuals (autonomy) are crucial (Barber and Schleiterman, 2008). Mana whānau asserts that students have an inherent need to feel safe, valued, and important in their social networks, from this they are able to develop valuable connections with those around them and a sense of belonging to their families, school and community.

Mana whānau is fostered when whānau (including school whānau) create child-centric environments that are invested in the holistic wellbeing of students and committed to helping students succeed (Webber and Macfarlane, 2020). Having whānau at home that are engaged with students development, learning and achievements establishes connectedness and belonging (Flores-Verduzco et al., 2020). Through community-school relationships, this child-centric attitude can align with the school environment and create consistent social support in the lives of students. When students have this crucial foundation, they are able to develop fundamental social skills and a secure sense of identity. Students are then motivated to succeed and contribute positively to their social groups, be it their school, family, or community (Webber and Macfarlane, 2020).

Mana ūkaipō—belonging and relationship to place

Mana ūkaipō refers to a sense of belonging and relationship to place. The word *ūkaipō* refers to a sense of place and belonging in the Māori language. Mana ūkaipō involves being connected to the school environment and being socially connected to others within the school environment. A relationship to place develops through meaningful positive experiences in a space, experiences which create place attachment (Reih, 2020) and establish a “meaningful location” (Lewicka, 2011). This is important for student development and wellbeing (Reih, 2020). Furthermore, school belonging has been defined by Goodenow (1993) as “the extent to which students feel personally accepted, respected, included, and supported by others in the school social environment” (p. 80). Supportive relationships with peers, teachers, and parents contribute to school belonging (Allen et al., 2016). School belonging has been linked to improved mental health and wellbeing (Arslan, 2018; Allen et al., 2016; Pikulski et al., 2020) as well as academic outcomes such as achievement, engagement, and behavior (Allen et al., 2016; Blum, 2005; Furrer and Skinner, 2003; Korperschok et al., 2019). Indigenous knowledge systems develop within the unique ecology of a place and thus contain fundamental resources for learning (Barnhardt, 2005; Highfield and Webber, 2021). Harnessing the language and culture “Indigenous to a particular place is a fundamental prerequisite for the development of culturally-healthy students and communities associated with that place” (Alaska Native Knowledge Network, 1998, p. 2). From this perspective, establishing school belonging and attachment necessarily involves learning that is situated in the surrounding environment. This sentiment is shared by Place Based Education (PBE) which broadly focuses on two core questions: “what is this place?” and “what is our relationship with it?” (Penitito, 2009). The intention of PBE is to learn *through* culture as opposed to *about* culture; PBE provides a lens in which to do so and combines Indigenous and Western perspectives to the benefit of all students (Penitito, 2009; Webber and Macfarlane, 2018).

Learning through (Indigenous) culture, language, and history fosters mana ūkaipō (Webber and Macfarlane, 2020) and strengthens Māori students’ academic and social identities. This is a view shared by culturally sustaining schooling—also referred to as culturally relevant pedagogy and culturally responsive teaching (Highfield and Webber, 2021). Researchers in this field argue that a mismatch between student’s home lives and school contributes to the achievement gap, and that creating a synergy between home and school culture can benefit students’ learning and engagement (particularly Indigenous students) (Castagno and Brayboy, 2008). Webber and Macfarlane (2018) show how education centered around iwi (tribal) knowledge and genealogy can support Māori success, achievement, and cultural identity as Māori historical icons demonstrate important cultural values and inspire motivation and commitment to success. Importantly, Indigenous and cultural knowledge is held by the communities themselves

requiring that schools establish genuine and reciprocal relationships with their wider communities (Webber and Macfarlane, 2018). It is through these relationships and the meaningful incorporation of cultural knowledge that student's cultural identities are recognized, fostering a sense of connection and belonging.

Mana motuhake—embedded achievement and self concept

Mana motuhake develops when students have a positive sense of identity and self-concept as this is central to learning and wellbeing. The Māori word *motuhake* refers to a person's distinctiveness, self-concept, or sense of personal authority in the Māori language. Self-concept is the thoughts and feelings one holds about oneself (Rosenberg, 1979) that are derived from past experiences, the perceptions of others, and how an individual thinks about their success, failures, skills, and strengths (Bong and Skaalvik, 2003). Research shows that having a positive self-concept is linked to self-esteem, wellbeing, and resilience in the face of negative stereotypes and racial discrimination (Bong and Skaalvik, 2003; Steele, 1997; Webber et al., 2013). Self-efficacy is a concept also important for mana motuhake. Self-efficacy is the degree to which an individual believes they have the skills and capabilities to achieve certain goals, which "influence[s] how people think, feel, motivate themselves and act" (Bandura, 1995, p. 2). Positive self-efficacy fosters motivation and achievement (Bong and Skaalvik, 2003). These concepts are important as academic self-concept and academic self-efficacy are derived from how an individual sees their academic selves generally (Diseth et al., 2014) and how they perceive their ability and capacity for growth in an academic context (Vestad et al., 2021).

The mana model posits that students' social identities, particularly ethnic and cultural identities, are central to developing mana motuhake. This is because self-concept and self-efficacy are bolstered by positive ethnic and cultural identity. As discussed previously in this chapter, individuals are aware of the societal view (including stereotypes) of their social categorizations (such as ethnic and cultural groups) and this can contribute positively or negatively to their self-concepts. Establishing strong and positive affiliations with one's social identity can reduce the negative impact of discrimination and racism and the subsequent impact on self-concept (Neblett et al., 2012; Sellers et al., 2006), and academic engagement and achievement (Webber, 2012). For instance, Macfarlane et al. (2014) found that successful Māori students had positive Māori identities and that families and community played a significant role in fostering this sense of identity. Whakapapa (knowledge and engagement in genealogy) and tribal knowledge were foundational to Māori identity and developed cultural efficacy for students (Webber and Macfarlane, 2020). Houkamau and Sibley (2011) suggest that cultural efficacy improves wellbeing because it indicates affinity with Māori others, resulting in social support and a sense of belonging. This aligns with international research that links positive ethnic identity with desirable psychological outcomes for Indigenous and minority groups (Dockery, 2010; Martinez and Dukes, 1997; Williams et al., 2018). Martinez and Dukes (1997) argue that education focused on strengthening student's ethnic identities and understanding race relations will benefit the wellbeing, self-esteem, self-confidence, and purpose in life for all students.

Connectedness, awareness of racism and embedded achievement are considered three vital components of a strong racial-ethnic identity which supports student success despite contextual circumstances (Oyserman et al., 1995; Altschul et al., 2006). Within this, it is important student know that they are capable of achieving *because* of their ethnic and cultural identity and knowledge. Embedded achievement is the belief that (ethnic) group membership is associated with valuing and succeeding in education (Oyserman and Lewis, 2017, p. 162). Having a sense of embedded achievement fosters resilience by pushing against the stigma that many Indigenous groups face. Thus, Indigenous perspectives posit that motivation and resilience at school is rooted in students' self-concept and self-efficacy, which crucially involves their ethnic and cultural identities (Webber and Macfarlane, 2020). As a result, strong ethnic and cultural role models are vital for inspiring embedded achievement and aspirational academic and personal goals for Māori students. It is important that education environments engage students by highlighting that "people like them" belong in certain domains (Highfield and Webber, 2021). The Mana Model is based on the assertion that all students are capable, and that academic, social, and cultural efficacy can provide them with the confidence and self-belief to thrive.

Mana tū—social psychological competence

Mana tū, highlights the social psychological competencies that students need in order to persevere through the challenges they may face. The Maori word *tū* refers to an individual's firm stance or ability to stand strong. From a Mana tū perspective, success involves social and mental wellbeing so that individuals can drive their own success. Mana tū describes the environmental and personal characteristics that strengthen students' capacity to work toward success. Holistic familial support is also an important environmental factor because family play a critical role in terms of modeling resilience strategies to students. Some students also utilized cultural supports (traditions, beliefs and family) to respond to challenges (Webber and Macfarlane, 2020).

Successful Māori students demonstrate personal characteristics such as resilience, academic persistence, grit, tenacity, and perseverance. These characteristics align with previous research that has shown that characteristics such as grit, perseverance, and consistency can aid academic achievement (Akos and Kretchmar, 2017; Bowman et al., 2015; Strayhorn, 2014; Credé et al., 2017; Duckworth et al., 2007; Wolters and Hussain, 2015) because they boost motivation, strategy use, and through avoiding

procrastination (Wolters and Hussain, 2015). In addition, resilience enables Māori students to overcome challenges and achieve positive outcomes (Masten, 2001, 2014). Resilience is beneficial for all students but particularly vital for Indigenous students who “are more likely to experience high levels of cumulative and co-occurring risks that lower their resilience, engagement, and participation in education” (McCalman and Bainbridge, 2021, p. 199). As Durie (2011) states,

Indigenous resilience is a reflection of an innate determination by Indigenous peoples to succeed. Resilience is the polar opposite of rigidity. It provides an alternate perspective to the more usual scenarios that emphasize Indigenous disadvantage, and allows the Indigenous challenge to be reconfigured as a search for success rather than an explanation of failure (p. 30).

Resilience offers a strengths-based approach to Indigenous education (McGuire-Kishebakabaykwe, 2010) and a positive lens for Indigenous students to see themselves.

Research shows that students benefit from being able to self-reflect on their thoughts feelings and behaviors in order to regulate and direct their own learning and learn from their past successes and failures (Grant et al., 2002; Voskamp et al., 2020; Colomer et al., 2021; Zimmerman, 2015). Self-reflection is also linked to insight (Hixon and Swann, 1993). When students have insight they understand themselves well and have clarity in who they are and what they need to do to succeed (Grant et al., 2002). Insight in particular is linked to positive wellbeing (Harrington and Loffredo, 2010; Lyke, 2008). This is mediated by the ability to maintain a positive outlook (Stein and Grant, 2014).

Mana tangatarua—diverse knowledge base and skillset

Indigenous perspectives of success recognize that many students traverse two or more worlds, each with differing sets of expectations, norms, and values. Mana tangatarua makes reference to the idea that many Indigenous students have the ability to maintain mana in two or more worlds. This concept is underpinned by the assertion that “academic success should never come at the expense of Māori identity” (Webber and Macfarlane, 2020, p. 41). This echoes the work of Durie (2011) who stated, “[i]t is no longer acceptable that education and training should lead to the abandonment of an Indigenous identity” (p. 33). Cultural identity and academic identity are both considered important facets of student’s identities, neither more important than the other. Schools tend to focus on individualized educative goals, which are important, albeit not the entirety of success from an Indigenous perspective.

Findings from the Ka Awatea study found that Māori students were grounded and led by their Māori identity (Macfarlane et al., 2014; Webber and Macfarlane, 2020), indicating that cultural/ethnic identity is central to mana tangatarua. Research in Aotearoa New Zealand and overseas has also shown that racial-ethnic identity is strengthened by cultural connectedness (Webber et al., 2013; Oyserman et al., 1995; Altschul et al., 2006) and that cultural connectedness positively impacts mental health outcomes for youth (Snowshoe et al., 2015; Fleming et al., 2020). Furthermore, cultural connectedness develops cultural efficacy (Houkamau and Sibley, 2011) which is “a person’s perception of his/her own capability to function effectively in situations characterized by cultural diversity” (Briones et al., 2009, p. 301). This capacity is rooted in knowledge of one’s own culture and history which enables students to appreciate and learn from the cultures of others. Since wider society and education typically reflects the dominant culture, many Indigenous students do not have their cultural identities developed at school (Milne, 2009; Zdrenka et al., 2015). Student’s families and communities are therefore vital and rich cultural resources for students to develop cultural connectedness, knowledge, and efficacy and for schools to support this (Webber et al., 2013; Webber and Macfarlane, 2020).

To walk in two (or more worlds), students need to be able to navigate the educational expectations and school culture while maintaining a sense of belonging and pride to their respective cultural groups. Importantly, a strong racial-ethnic identity is aided by an awareness of racism (Oyserman et al., 1995; Altschul et al., 2006). A large body of research exists pertaining to the obstacles Indigenous and other minority students face such as systemic and interpersonal racism, discrimination, and stereotypes (Benner et al., 2018; Cormack et al., 2018; Fisher et al., 2000; Turner et al., 2015). These factors create unique personal, social, economic, and educational barriers for Indigenous students. An awareness of how racism is a symptom of (colonial) history and wider societal issues can help to depersonalize negative racial experiences (Neblett et al., 2013). For Māori, an awareness of racism can help build resilience (Webber, 2015) and protect mental health (Houkamau and Sibley, 2011). Cultural connectedness coupled with an awareness of racism can improve student achievement (Altschul et al., 2006) and protect against the negative influence teacher discrimination can have on achievement (Thomas et al., 2009).

Conversely, dominant students can also benefit from understanding their own culture, particularly as many students do not see themselves as having a culture due to dominant culture being presented as the norm (Milne, 2009; Webber et al., 2013). Furthermore, through learning from the culture of others and about history and racism, dominant students, teachers, and wider community can gain perspective on the assumptions in education and beyond and the privileges that many benefit from as a result that (Small, 2019). Therefore, from an Indigenous perspective, it is important that student’s families, school environment, and local community provide support and learning opportunities with both cultural and educational foci (Webber and Macfarlane, 2020).

Conclusion

SIT suggests that Indigenous students' social identities are influential in how they see themselves and interact with others. Importantly, SIT posits that Indigenous students are motivated to achieve a positive sense of social identity which is largely established through comparison with other groups. Thus, while being a part of a socially valued or dominant group can bolster self-esteem, membership in a less socially valued or non-dominant group threatens self-esteem. This negatively impacts on Indigenous students' behavior and engagement at school. Similarly, stereotype threats are the persistent generalized, and often negative, assumptions and perceptions of one's social group that are an additional obstacle for Indigenous students. Stereotype threat impacts Indigenous student achievement and can be ruinous in terms of Indigenous students disengaging from school, or education as a whole, for fear of being confirming negative stereotypes.

This chapter has considered how Indigenous social-psychological theory, and research conducted in Aotearoa New Zealand with Māori students, might offer a useful perspective on identity formation and how schools and educators might best support students' social identities and develop student resilience against stereotype threat. The chapter presents the Mana Model, and a number of ways schools might better support the diverse social identities of Indigenous students and ensure that they do not need to choose between their cultural identity and their academic identity to achieve at school. When schools work to support students' social identities and foster their connection to their culture, community and peers, there is a positive impact on students individually as well as on their academic engagement and outcomes. Furthermore, when students develop mana in various facets of their lives as outlined by the mana model, resilience is fostered, and stereotype threat is mitigated.

References

- Abraham, J., 1989. Testing Hargreaves' and Lacey's differentiation-polarisation theory in a setted comprehensive. *Br. J. Sociol.* 46–81.
- Abrams, D., Hogg, M.A., 2010. Social identity and self-categorization. In: Dovidio, J.F., Hewstone, M., Glick, P., Esses, V.M. (Eds.), *The SAGE Handbook of Prejudice, Stereotyping and Discrimination*, pp. 179–193. <https://doi.org/10.4135/9781446200919.n11>.
- Akos, P., Kretchmar, J., 2017. Investigating grit at a non-cognitive predictor of college success. *Rev. High. Educ.* 40 (2), 163–186. <https://doi.org/10.1353/rhe.2017.0000>.
- Alaska Native Knowledge Network, 1998. Alaska Standards for Culturally Responsive Schools. Adopted by the Assembly of Alaska Native Educators. <http://www.ankn.uaf.edu/publications/standards.html>.
- Alfred, T., Cornthassel, J., 2005. Being indigenous: resurgences against contemporary colonialism. *Gov. Oppos.* 40 (4), 597–614. <https://doi.org/10.1111/j.1477-7053.2005.00166.x>.
- Allen, K., Kern, M.L., Vella-Brodick, D., Hattie, J., Waters, L., 2016. What schools need to know about fostering school belonging: a meta-analysis. *Educ. Psychol. Rev.* 30 (1), 1–34.
- Allen, K.A., Gray, D.L., Baumeister, R.F., Leary, M.R., 2021. The need to belong: a deep dive into the origins, implications, and future of a foundational construct. *Educ. Psychol. Rev.* 34 (2), 1133–1156. <https://doi.org/10.1007/s10648-021-09633-6>.
- Altshul, I., Oyserman, D., Bybee, D., 2006. Racial-ethnic identity in mid-adolescence: content and change as predictors of academic achievement. *Child Dev.* 77 (5), 1155–1169.
- Appel, M., Kronberger, N., 2012. Stereotypes and the achievement gap: stereotype threat prior to test taking. *Educ. Psychol. Rev.* 24 (4), 609–635. <https://doi.org/10.1007/s10648-012-9200-4>.
- Arslan, G., 2018. Exploring the association between school belonging and emotional health among adolescents. *Int. J. Educ. Psychol.* 7 (1), 21–41. <https://doi.org/10.17583/ijep.2018.3117>.
- Bandura, A., 1995. *Self-Efficacy in Changing Societies*. Cambridge University Press.
- Barber, B.K., Schluterman, J.M., 2008. Connectedness in the lives of children and adolescents: a call for greater conceptual clarity. *J. Adolesc. Health* 43, 209–216.
- Barnhardt, R., 2005. Indigenous knowledge systems and Alaska native ways of knowing. *Anthropol. Educ. Q.* 36 (1), 8–23.
- Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117 (3), 497–529.
- Benner, A.D., Wang, Y., Shen, Y., Boyle, A.E., Polk, R., Cheng, Y., 2018. Racial/ethnic discrimination and well-being during adolescence: a meta-analytic review. *Am. Psychol.* 73 (7), 855–883. <https://doi.org/10.1037/amp0000204>.
- Blum, R.W., 2005. A case for school connectedness. *Educ. Leader.* 62 (7), 16–20.
- Bong, M., Skaalvik, E.M., 2003. Academic self-concept and self-efficacy: how different are they? *Educ. Psychol. Rev.* 15, 1–40.
- Borell, B., 2005. Living in the City Ain't So Bad: Cultural Diversity of South Auckland Rangatahi. Master's Thesis, Massey University. Massey Research Online. <https://mro.massey.ac.nz/xmlui/handle/10179/5812>.
- Bowman, N.A., Hill, P.L., Denson, N., Bronkema, R., 2015. Keep on truckin' or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Soc. Psychol. Personal. Sci.* 6 (6), 639–645. <https://doi.org/10.1177/1948550615574300>.
- Briones, E., Tabernero, C., Tramontano, C., Caprara, G.V., Arenas, A., 2009. Development of a cultural self-efficacy scale for adolescents (CSES-A). *Int. J. Intercult. Relat.* 33 (4), 301–312.
- Burnaford, R.M., 2015. Countervailing forces in minority identity. In: Rubie-Davies, C., Stephens, J.M., Watson, P. (Eds.), *The Routledge International Handbook of Social Psychology of the Classroom*. Routledge, pp. 122–130.
- Carter, P.L., 2006. Straddling boundaries. *Sociol. Educ.* 79 (4), 304–328.
- Castagno, A., Brayboy, B., 2008. Culturally responsive schooling for Indigenous youth: a review of the literature. *Rev. Educ. Res.* 78 (4), 941–993. <https://doi.org/10.3102/0034654308323036>.
- Colomer, J., Teresa Serra, T., Gras, M.E., Cañabate, D., 2021. Longitudinal self-directed competence development of university students through self-reflection. *Reflective Pract.* 22 (5), 727–740. <https://doi.org/10.1080/14623943.2021.1964947>.
- Connor, H.D., 2019. Whakapapa back: mixed Indigenous Māori and Pākehā genealogy and Heritage in Aotearoa/New Zealand. *Genealogy* 3 (4), 73. <https://doi.org/10.3390/genealogy3040073>.
- Cormack, D., Stanley, J., Harris, R., 2018. Multiple forms of discrimination and relationships with health and wellbeing: findings from national cross-sectional surveys in Aotearoa/New Zealand. *Int. J. Equity Health* 17 (26). <https://doi.org/10.1186/s12939-018-0735-y>.
- Credé, M., Tynan, M.C., Harms, P.D., 2017. Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Pers. Soc. Psychol.* 113 (3), 492–511. <https://doi.org/10.1037/pspp0000102>.
- De Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., Branstetter, C., Wang, X., 2019. Status and Trends in the Education of Racial and Ethnic Groups 2018 (NCES 2019-038). U.S. Department of Education. National Center for Education Statistics, Washington, DC. <https://nces.ed.gov/pubs2019/2019038.pdf>.

- Dell, K., 2017. Te hokinga ki Te Ūkaipō: Disrupted Māori Management Theory; Harmonising whānau Conflict in the Māori Land Trust. Unpublished Doctor of Philosophy in Management and International Business. University of Auckland, New Zealand.
- Department of Corrections, 2021. Prison Facts and Statistics—September 2021. [https://www.corrections.govt.nz/resources/statistics/quarterly_prison_stats_september_2021](https://www.corrections.govt.nz/resources/statistics/quarterly_prison_statistics/prison_stats_september_2021).
- Desombre, C., Aneymar, S., Delelis, G., 2018. Stereotype threat among students with disabilities: the importance of the evaluative context on their cognitive performance. *Eur. J. Psychol. Educ.* 33 (2), 201–214.
- Diseth, Å., Meland, E., Breidablik, H.J., 2014. Self-beliefs among students: grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learn. Individ. Differ.* 35, 1–8. <https://doi.org/10.1016/j.lindi.2014.06.003>.
- Dockery, A.M., 2010. Culture and wellbeing: the case of indigenous Australians. *Soc. Indic. Res.* 99 (2), 315–332. <https://doi.org/10.1007/s11205-010-9582-y>.
- Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: perseverance and passion for long-term goals. *J. Pers. Soc. Psychol.* 92 (6), 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>.
- Duncanson, M., van Asten, H., Adams, J., McAnally, H., Zhang, X., Wicken, A., Oben, G., 2021. Child Poverty Monitor 2021 Technical Report. NZ Child and Youth Epidemiology Service, University of Otago, Dunedin. <https://ourarchive.otago.ac.nz/handle/10523/12540>.
- Durie, M.H., 1995. Te hoe nuku roa framework a Māori identity perspective. *J. Polyn. Soc.* 104 (4), 461–470. <https://www.jstor.org/stable/20706636>.
- Durie, M., 2001. *Mauri Ora: The Dynamics of Māori Health*. Oxford University Press.
- Durie, M., 2011. Indigenous resilience: from disease and disadvantage to the realisation of potential. In: Durie, M. (Ed.), *Nga tini whetu: Navigating Māori Futures*. Huia (NZ) Ltd, pp. 29–40.
- Elliott, G., Kao, S., Grant, A.M., 2004. Materring: empirical validation of a social-psychological concept. *Self Ident.* 3 (4), 339–354.
- Fahey, G., 2021. Mind the Gap: Understanding the Indigenous Education Gap and How to Close It. Research Report 41. The Centre for Independent Studies. <https://www.cis.org.au/publications/research-reports/mind-the-gap-understanding-the-Indigenous-education-gap-and-how-to-close-it/>.
- Fisher, C.B., Wallace, S.A., Fenton, R.E., 2000. Discrimination distress during adolescence. *J. Youth Adolesc.* 29 (6), 679–695.
- Fleming, T., Tiatia-Seath, J., Peiris-John, R., Sutcliffe, K., Archer, D., Bavin, L., Crengle, S., Clark, T., 2020. Youth19 Rangatahi Smart Survey, Initial Findings: hauora hinengaro/Emotional and Mental Health. The Youth19 Research Group, The University of Auckland, Victoria University of Wellington, Auckland, New Zealand.
- Flores-Verduzco, G.R., Fraijo-Sing, B.S., Tapia-Fonlle, C.O., 2020. School environment as a mediating variable between family support and social wellbeing in high school students. *Int. J. Emot. Educ.* 12 (2), 26–40.
- Fordham, S., 1988. Racelessness as a factor in black students' school success: pragmatic strategy or pyrrhic victory? *Harv. Educ. Rev.* 58 (1), 54–84.
- Fordham, S., Ogbu, J.U., 1986. Black students' school success: coping with the "burden of acting white". *Urban Rev.* 18 (3), 176–206. <https://doi.org/10.1007/BF01112192>.
- Furrer, C., Skinner, E., 2003. Sense of relatedness as a factor in children's academic engagement and performance. *J. Educ. Psychol.* 95 (1), 148–162.
- González, R., Carvacho, H., Jiménez-Moya, G., 2022. Psychology and indigenous people. *Annu. Rev. Psychol.* 73, 431–459.
- Good, C., Dweck, C.S., Aronson, J., 2007. Social identity, stereotype threat, and self-theories. In: Fuligni, A.J. (Ed.), *Contesting Stereotypes and Creating Identities: Social Categories, Social Identities, and Education Participation*. Russell Sage Foundation.
- Good, C., Aronson, J., Harder, J.A., 2008. Problems in the pipeline: stereotype threat and women's achievement in high-level math courses. *J. Appl. Dev. Psychol.* 29 (1), 17–28. <https://doi.org/10.1007/s10212-016-0327-4>.
- Goodenow, C., 1993. The psychological sense of school membership among adolescents: scale development and educational correlates. *Psychol. Sch.* 30 (1), 79–90.
- Graham, J., Meyer, L.H., McKenzie, L., McClure, J., Weir, K.F., 2010. Māori and Pacific secondary student and parent perspectives on achievement, motivation and NCEA. *Assess. Matters* 2, 156.
- Grant, A.M., Franklin, J., Langford, P., 2002. The self-reflection and insight scale: a new measure of private self-consciousness. *SBP* 30 (8), 821–835. <https://doi.org/10.2224/sbp.2002.30.8.821>.
- Greaves, L.M., 2017. He Pōkākē Uenuku I Tū Ai: A Quantitative Exploration of Māori Identity, Political Attitudes, and Behaviour. PhD Thesis, The University of Auckland. ResearchSpace. <https://researchspace.auckland.ac.nz/handle/2292/37064>.
- Harrington, R., Loffredo, D.A., 2010. Insight, rumination, and self-reflection as predictors of well-being. *J. Psychol.* 145 (1), 39–57. <https://doi.org/10.1080/00223980.2010.528072>.
- Highfield, C., Webber, M., 2021. Mana Ūkaipō: Māori student connection, belonging and engagement at school. *N. Z. J. Educ. Stud.* 56 (2), 145–164.
- Hixon, J., Swann, W., 1993. When does introspection bear fruit? Self-reflection, self-insight, and interpersonal choices. *J. Pers. Soc. Psychol.* 64 (1), 35–43.
- Hogg, M.A., 2000. Subjective uncertainty reduction through self-categorization: a motivational theory of social identity processes. *Eur. Rev. Soc. Psychol.* 11, 223–255. <https://doi.org/10.1080/14792772043000040>.
- Hogg, M.A., 2016. Social identity theory. *Peace Psychology Book Series*. In: McKeown, S., Haji, R., Ferguson, N. (Eds.), *Understanding Peace and Conflict Through Social Identity Theory*. Springer International Publishing, pp. 3–17. https://doi.org/10.1007/978-3-319-29869-6_1.
- Hoskins, T.K., 2012. A fine risk: ethics in Kaupapa Māori politics. *N. Z. J. Educ. Stud.* 47 (2), 85–99.
- Houkamau, C.A., 2011. Transformations in Māori women's identity: some things change, some stay the same. *Qual. Res. Psychol.* 8 (3), 292–313. <https://doi.org/10.1080/14780880903304501>.
- Houkamau, C.A., Sibley, C.G., 2010. The multi-dimensional model of Māori identity and cultural engagement. *N. Z. J. Psychol.* 39 (1), 8–28.
- Houkamau, C.A., Sibley, C.G., 2011. Māori cultural efficacy and subjective wellbeing: a psychological model and research agenda. *Soc. Indic. Res.* 103 (3), 379–398.
- Houkamau, C.A., Sibley, C.G., 2015. The revised multidimensional model of Māori identity and cultural engagement (MMM-ICE2). *Soc. Indic. Res.* 122 (1), 279–296. <https://doi.org/10.1007/s11205-014-0686-7>.
- Hutchingson, J., Reader, M., Akhal, A., 2020. Education in England: Annual Report 2020. Education Policy Institute. https://epi.org.uk/wp-content/uploads/2020/09/EPI_2020_Annual_Report.pdf.
- Kelly, S., 2007. Race, social class, and student engagement in middle school English classrooms. *Soc. Sci. Res.* 37 (2), 434–448. <https://doi.org/10.1016/j.ssresearch.2007.08.003>.
- Korpershoek, H., Canrinus, E.T., Fokkens-Bruinsma, M., de Boer, H., 2019. The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: a meta-analytic review. *Res. Pap. Educ.* 35 (6), 641–680. <https://doi.org/10.1080/02671522.2019.1615116>.
- Kukutai, T., 2004. The problem of defining an ethnic group for public policy: who is Māori and why does it matter? *Soc. Pol. J. N. Z.* 23, 86–108.
- Leary, M.R., Baumeister, R.F., 2000. The nature and function of self-esteem: sociometer theory. In: Zanna, M.P. (Ed.), *Advances in Experimental Social Psychology*, vol. 32. Academic Press, pp. 1–62. [https://doi.org/10.1016/S0065-2601\(00\)80003-9](https://doi.org/10.1016/S0065-2601(00)80003-9).
- Lewicka, M., 2011. Place attachment: how far have we come in the last 40 years? *J. Environ. Psychol.* 31 (3), 207–230.
- Lyke, J.A., 2008. Insight, but not self-reflection, is related to subjective well-being. *Pers. Individ. Differ.* 46 (1), 66–70. <https://doi.org/10.1016/j.paid.2008.09.010>.
- Macfarlane, A.H., Webber, M., Cookson-Cox, C., McRae, H., 2014. Ka Awatea: an iwi Case Study of Māori Students' Success. Te Rū Rangahau.
- Martinez, R., Dukes, R., 1997. The effects of ethnic identity, ethnicity and gender on adolescent well-being. *J. Youth Adolesc.* 26 (5), 503–516.
- Masten, A.S., 2001. Ordinary magic. *Am. Psychol.* 56 (3), 227–238.
- Masten, A.S., 2014. Global perspectives on resilience in children and youth. *Child Dev.* 85 (1), 6–20.
- McAllister, T.G., Kidman, J., Rowley, O., Theodore, R.F., 2019. Why isn't my professor Māori? *MAI J.* 8 (2), 235–249.
- McAllister, T.G., Naepi, S., Wilson, E., Hikuroa, D., Walker, L.A., 2022. Under-represented and overlooked: Māori and Pasifika scientists in Aotearoa New Zealand's universities and crown-research institutes. *J. Roy. Soc. N. Z.* 52 (1), 38–53.

- McCalman, J., Bainbridge, R., 2021. Indigenous education, well-being, and resilience—a systemic approach. In: Ungar, M. (Ed.), *Multisystemic Resilience: Adaptation and Transformation in Contexts of Change*. Oxford University Press, pp. 199–219.
- McGuire-Kishebakabaykwe, P.D., 2010. Exploring resilience and Indigenous ways of knowing. *Pimatisiwin* 8 (2), 117–131.
- McInerney, D.M., Flowerday, T., 2016. Indigenous issues in education and research: looking forward. *Contemp. Educ. Psychol.* 100 (47), 1–3.
- McIntosh, T., 2006. Māori identities: fixed, fluid, forced. In: Liu, J.H., McCreanor, T., McIntosh, T., Teaiwa, T. (Eds.), *New Zealand Identities: Departures and Destinations*. Victoria University Press, Wellington, pp. 67–94.
- Mead, H.M., 2006. *Tikanga Māori: Living by Māori Values*. Huia (NZ) Ltd.
- Mello, Z.R., 2009. Racial/ethnic group and socioeconomic status variation in educational and occupational expectations from adolescence to adulthood. *J. Appl. Dev. Psychol.* 30 (4), 494–504. <https://doi.org/10.1016/j.appdev.2008.12.029>.
- Metge, J., 1964. *A New Māori Migration: Rural and Urban Relations in Northern New Zealand*. Athlone Press, London.
- Milne, A., 2009. *Colouring in the White Spaces: Cultural Identity and Learning in School: A Research Report Prepared for the ASB/APPA Travelling Fellowship Trust Following a Study Tour in 2009*. ASB/APPA Travelling Fellowship Trust, Auckland.
- Ministry of Business, Innovation and Employment, 2017. Māori in the Labour Market. <https://www.mbie.govt.nz/business-and-employment/employment-and-skills/labour-market-reports-data-and-analysis/other-labour-market-reports/Māori-labour-market-trends/>.
- Ministry of Education, 2020a. Education Indicator: Education and Learning Outcomes: School Leavers With NCEA Level 1 or Above. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0017/208070/Indicator-NCEA-Level-1-or-Above-2020.pdf.
- Ministry of Education, 2020b. Education Indicator: Education and Learning Outcomes: School Leavers With NCEA Level 2 or Above. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0018/208071/Indicator-NCEA-Level-2-or-Above-2020.pdf.
- Ministry of Education, 2020c. Education Indicator: Education and Learning Outcomes: School Leavers With NCEA Level 3 or University Entrance Standard. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0019/208072/Indicator-NCEA-Level-3-and-UE-2020.pdf.
- Ministry of Health, 2019. *Wai 2575 Māori Health Trends Report*. Ministry of Health, Wellington. <https://www.health.govt.nz/publication/wai-2575-Māori-health-trends-report>.
- Moeke-Pickering, T., 1996. Māori Identity Within whānau: A Review of Literature. University of Waikato. <https://researchcommons.waikato.ac.nz/handle/10289/464>.
- Neblett Jr., E.W., Rivas-Drake, D., Umaña-Taylor, A.J., 2012. The promise of racial and ethnic protective factors in promoting ethnic minority youth development. *Child Dev. Perspect.* 6 (3), 295–303. <https://doi.org/10.1111/j.1750-8606.2012.00239.x>.
- Neblett Jr., E.W., Banks, K.H., Cooper, S.M., Smalls-Glover, C., 2013. Racial identity mediates the association between ethnic-racial socialization and depressive symptoms. *Cult. Divers. Ethn. Minor. Psychol.* 19 (2), 200.
- New Zealand Qualification Authority (NZQA), 2020. Annual Report: NCEA, University Entrance and NZ Scholarship Data and Statistics. <https://www.nzqa.govt.nz/assets/About-us/Publications/stats-reports/NCEA-Annual-Report-2020.pdf>.
- Ogbu, J.U., 2004. Collective identity and the burden of “acting white” in black history, community, and education. *Urban Rev.* 36 (1), 1–35. <https://doi.org/10.1023/B:URRE.0000042734.83194.f6>.
- Onorato, R.S., Turner, J.C., 2004. Fluidity in the self-concept: the shift from personal to social identity. *Eur. J. Soc. Psychol.* 34 (3), 257–278. <https://doi.org/10.1002/ejsp.195>.
- Oyserman, D., 2007. Social identity and self-regulation. In: Kruglanski, A.W., Higgins, E.T. (Eds.), *Social Psychology: Handbook of Basic Principles*. The Guilford Press, pp. 432–453.
- Oyserman, D., Lewis, N., 2017. Seeing the destination and the path: using identity-based motivation to understand and reduce racial disparities in academic achievement. *Soc. Issues Pol. Rev.* 11 (1), 159–194. <https://doi.org/10.1111/sipr.12030>.
- Oyserman, D., Gant, L., Ager, J., 1995. A socially contextualized model of African American identity: possible selves and school persistence. *J. Pers. Soc. Psychol.* 69, 1216–1232.
- Paterson, R., Durie, M., Disley, B., Rangihuna, D., Tiatia-Seath, J., Tualamali'i, J., 2018. *He Ara Oranga: Report of the Government Inquiry Into Mental Health and Addiction*. Government inquiry into mental health and addiction. <https://mentalhealth.inquiry.govt.nz/inquiry-report/he-ara-oranga/>.
- Penetito, W., 2009. Place-based education: catering for curriculum, culture, and community. *N. Z. Annu. Rev. Educ.* 18, 5–29.
- Pikulski, P., Pella, J., Casline, E., Hale, A., Drake, K., Ginsburg, G., 2020. School connectedness and child anxiety. *J. Psychol. Couns. Sch.* 30 (1), 13–24. <https://doi.org/10.1017/jgc.2020.3>.
- Reardon, S.F., Kalogrides, D., Shores, K., 2019. The geography of racial/ethnic test score gaps. *Am. J. Sociol.* 124 (4), 1164–1221.
- Rieh, S., 2020. *Creating a Sense of Place in School Environments: How Young Children Construct Place Attachment*. Routledge. <https://doi.org/10.4324/9780429441332>.
- Ritchie, J., 1963. *The Making of a Māori*. A.H. & A.W. Reed, Wellington.
- Rosenberg, M., 1979. *Conceiving the Self*. Basic Books, New York.
- Royal, C., 2006. A Modern View of Mana. *New Zealand Psychological Society*, pp. 8–13. *Bulletin* 107.
- Rua, M., Hodgetts, D., Stolte, O., 2017. Māori men: an Indigenous psychological perspective on the interconnected self. *N. Z. J. Psychol.* 46 (3), 55–63.
- Rudolph, F., Salikutluk, Z., 2021. Aiming high, no matter what?: educational aspirations of ethnic minority and ethnic majority students in England, Germany and the Netherlands. *Comp. Sociol.* 20 (1), 70–100. <https://doi.org/10.1163/15691330-BJA10025>.
- Sellers, R.M., Copeland-Linder, N., Martin, P.P., Lewis, R.L., 2006. Racial identity matters: the relationship between racial discrimination and psychological functioning in african american adolescents. *J. Res. Adolesc.* 16 (2), 187–216.
- Small, D., 2019. Education and history. Who do we think we are? In: Kamp, A. (Ed.), *Education Studies in Aotearoa: Key Disciplines and Emerging Directions*. NZCER, Wellington.
- Snowshoe, A., Crooks, C.V., Tremblay, P.F., Craig, W.M., Hinson, R.E., 2015. Development of a cultural connectedness scale for First Nations youth. *Psychol. Assess.* 27, 249–259. <https://doi.org/10.1037/a0037867>.
- Statistics New Zealand, 2021. *Wellbeing Statistics—March 2021 Quarter*. <https://www.stats.govt.nz/information-releases/wellbeing-statistics-march-2021-quarter>.
- Steele, C., 1997. A threat in the air: how stereotypes shape the intellectual identities and performance. *Am. Psychol.* 52 (6), 613–629.
- Steele, C.M., 2010. *Whistling Vivaldi: And Other Clues to How Stereotypes Affect Us (Issues of Our Time)*. Norton, New York, NY.
- Steele, C.M., Aronson, J., 1995. Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>.
- Stein, D., Grant, A.M., 2014. Disentangling the relationships among self-reflection, insight, and subjective well-being: the role of dysfunctional attitudes and core self-evaluations. *J. Psychol.* 148 (5), 505–522. <https://doi.org/10.1080/00223980.2013.810128>.
- Strayhorn, T.L., 2014. What role does grit play in the academic success of black male collegians at predominantly white institutions? *J. Afr. Am. Stud.* 18 (1), 1–10. <https://doi.org/10.1007/s12111-012-9243>.
- Tajfel, H., 1969. Cognitive aspects of prejudice. *J. Soc. Issues* 25, 79–97. <https://doi.org/10.1017/S0021932000023336>.
- Tajfel, H., 1972. Social categorization (English translation of “La categorisation sociale”). In: Moscovici, S. (Ed.), *Introduction à la psychologie sociale*, vol. 1. Larousse, Paris, pp. 272–302.
- Tajfel, H., 1974. Social identity and intergroup behaviour. *Soc. Sci. Inf.* 13 (2), 65–93. <https://doi.org/10.1177/053901847401300204>.
- Tajfel, H., 1978. The achievement of inter-group differentiation. In: Tajfel, H. (Ed.), *Differentiation Between Social Groups*. Academic Press, London, pp. 77–100.
- Tajfel, H., 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge University Press, Cambridge; New York.
- Tajfel, H., Turner, J.C., 1979. An integrative theory of intergroup conflict. In: Austin, W.G., Worchel, S. (Eds.), *The Social Psychology of Intergroup Relations*. Brooks/Cole, Monterey, CA, pp. 33–47.
- Tajfel, H., Billig, M., Bundy, R.P., Flament, C., 1971. Social categorization and intergroup behaviour. *Eur. J. Soc. Psychol.* 1, 149–177.
- Te Rito, J., 2007. Whakapapa: a framework for understanding identity. *MAI Rev.* 2, 1–10.

- Thomas, O.N., Caldwell, C.H., Faison, N., Jackson, J.S., 2009. Promoting academic achievement: the role of racial identity in buffering perceptions of teacher discrimination on academic achievement among African American and Caribbean Black adolescents. *J. Educ. Psychol.* 101 (2), 420–431. <https://doi.org/10.1037/a0014578>.
- Turner, H., Rubie-Davies, C.M., Webber, M., 2015. Teacher expectations, ethnicity and the achievement gap. *N. Z. J. Educ. Stud.* 50 (1), 55–69.
- Turner, J.C., 1975. Social comparison and social identity: some prospects for intergroup behaviour. *Eur. J. Soc. Psychol.* 5, 5–34.
- Turner, J.C., 1987. *Rediscovering the Social Group: Self-Categorization Theory*. Basil Blackwell.
- Turner, J.C., Reynolds, K.J., 2012. Self-Categorization theory. In: Van, L.P.A.M., Kruglanski, A.W., Higgins, E.T. (Eds.), *Handbook of Theories of Social Psychology*, vol. 2. SAGE Publications, London, pp. 399–417.
- Verkuyten, M., Thijs, J., Gharaei, N., 2019. Discrimination and academic (dis)engagement of ethnic-racial minority students: a social identity threat perspective. *Soc. Psychol. Educ.* 22 (2), 267–290. <https://doi.org/10.1007/s11218-018-09476-0>.
- Vestad, L., Bru, E., Virtanen, T.E., Stallard, P.N., 2021. Associations of social and emotional competencies, academic efficacy beliefs, and emotional distress among students in lower secondary school. *Soc. Psychol. Educ.* 24, 413–439. <https://doi.org/10.1007/s11218-021-09624-z>.
- Voskamp, A., Kuiper, E., Volman, M., 2020. Teaching practices for self-directed and self-regulated learning: case studies in Dutch innovative secondary schools. *Educ. Stud.* 1–18. <https://doi.org/10.1080/03055698.2020.1814699>.
- Walker, R., 1989. Māori identity. In: Novitz, D., Willmott, B. (Eds.), *Culture and Identity in New Zealand*. Government Printer, Wellington.
- Webber, M., 2007. *Hybrid Māori/Pakeha: Explorations of Identity for People of Mixed Māori/Pakeha Descent*. Master's Thesis, The University of Auckland. The University of Auckland Research Repositories, ResearchSpace. <http://hdl.handle.net/2292/21670>.
- Webber, M., 2012. Identity matters: the role of racial-ethnic identity for Māori students in multiethnic secondary schools. *SET Res. Inf. Teach.* 2, 20–25. <https://doi.org/10.18296/set.0370>.
- Webber, M., 2015. Optimizing Māori student success with the other three “R”s. In: Rubie-Davies, C., Stephens, J., Watson, P. (Eds.), *The Routledge International Handbook of Social Psychology of the Classroom*, pp. 102–111. <https://doi.org/10.4324/9781315716923.ch9>.
- Webber, M., Macfarlane, A., 2018. The transformative role of tribal knowledge and genealogy in indigenous student success. In: Smith, L., McKinley, E. (Eds.), *Indigenous Handbook of Education*. Springer, Singapore, pp. 1049–1074.
- Webber, M., Macfarlane, A., 2020. Mana tangata: the five optimal cultural conditions for Māori student success. *J. Am. Indian Educ.* 59 (1), 26–49.
- Webber, M., McKinley, E., Hattie, J., 2013. The importance of race and ethnicity: an exploration of New Zealand Pākehā, Māori, Samoan and Chinese adolescent identity. *N. Z. J. Psychol.* 42 (2), 17–28.
- Williams, A.D., Clark, T.C., Lewycka, S., 2018. The associations between cultural identity and mental health outcomes for indigenous Māori youth in New Zealand. *Front. Publ. Health* 6, 319. <https://doi.org/10.3389/fpubh.2018.00319>.
- Wolters, C.A., Hussain, M., 2015. Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacogn. Learn.* 10 (3), 293–311. <https://doi.org/10.1007/s11409-014-9128-9>.
- Zdrenka, M., Yogeewaran, K., Stronge, S., Sibley, C.G., 2015. Ethnic and national attachment as predictors of wellbeing among New Zealand Europeans, Māori, Asians, and Pacific nations peoples. *Int. J. Intercult. Relat.* 49, 114–120. <https://doi.org/10.1016/j.ijintrel.2015.07.003>.
- Zimmerman, B., 2015. Self-Regulated learning: theories, measures, and outcomes. *Int. Encycl. Soc. Behav. Sci.* 21, 541–546. <https://doi.org/10.1016/B978-0-08-097086-8.26060-1>.

Students' social relationships at school

Kathryn R. Wentzel, University of Maryland, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Interpersonal relationships with peers and teachers	289
Relationships with teachers	289
Teacher–student relationships and motivation	289
Teacher–student relationships and achievement	289
Teacher–student relationships and social-behavioral outcomes	289
Relationships with peers	289
Peer relationships and motivation	290
Peer relationships and academic achievement	290
Peer relationships and social-behavioral outcomes	291
Relationship provisions and affordances	291
Communicating norms and expectations	291
Teachers	291
Peers	291
Instrumental aid: provisions of help, advice, and instruction	292
Teachers	292
Peers	292
Social and emotional support and safety	293
Teachers	293
Peers	293
Rewards and reinforcements	294
Teachers	294
Peers	294
Summary	294
Conclusions	294
References	295

There is growing consensus that the nature and quality of children's relationships with their teachers play a critical and central role in motivating and engaging students to learn and to behave in socially-acceptable ways (see [Wentzel, 2016](#)). Effective teachers are typically described as those who develop relationships with students that are emotionally close, safe, and trusting, that provide access to instrumental help, and that foster a more general ethos of community and caring in classrooms. In doing so, they support the development of students' emotional well-being and positive sense of self, motivational orientations for social and academic outcomes, and actual social and academic skills. They also provide a context for communicating positive and high expectations for performance and teaching students what they need to know to become knowledgeable and productive citizens. Likewise, children who enjoy positive relationships with their peers also tend to be engaged in, and even excel at, academic tasks and display positive forms of behavior more than those who have peer relationship problems (see [Wentzel and Muenks, 2016](#)). As with teachers, positive relationships with peers also appear to afford opportunities for developing levels of trust and safety, receiving help and information concerning what is expected of them academically and socially.

A focus on the social ecologies of schooling reflects the notion that the ability to be competent in any social setting, including classrooms and schools, is contingent on opportunities and affordances of the social environment. [Bronfenbrenner \(1989\)](#) argued that competence is achieved, in part, when contexts provide opportunities for the growth and development of personal skills and attributes as well as scaffolding for learning what is expected by the social group. Expanding on this notion of situated competence, [Ford \(1992\)](#) argued that several contextual supports can promote positive social engagement, including information concerning what is expected and valued, help and instruction that support attempts to achieve these valued outcomes, a safe, non-threatening environment; and emotional involvement such that individuals are made to feel like a valued member of the group.

Using this ecological approach as a guiding framework, I first review the literature on teacher–student and peer relationships as it pertains to children's school-related competencies: specifically, motivation to engage socially and academically, academic achievement, and social-behavioral outcomes. I then describe the benefits and affordances of these relationships in competence development, including communications of norms and expectations, instrumental aid, emotional support and safety, and rewards and reinforcements.

Interpersonal relationships with peers and teachers

Within school contexts, positive interpersonal relationships are valued outcomes in and of themselves; they also are necessary for successful group functioning and appear to be significant motivators of student outcomes. In the following sections, literature on teacher–student and peer relationships and their associations with student motivation, academic achievement, and social-behavioral outcomes is described.

Relationships with teachers

Evidence to support the importance of teacher–student relationships in students' lives has increased tremendously in the past ten years. Research has focused primarily on attachment-like qualities of dyadic teacher–student relationships, and students' perceptions of supports from teachers (see [Wentzel, 2016](#) for a review). In young children, this typically takes the form of attachments to teachers and relationship qualities reflective of warmth and emotional caring (e.g., [Pianta, 2006](#)). In older children, researchers tend to focus on students' subjective appraisals of their relationships with teachers (e.g., teacher emotional involvement) and teachers' emotional support (e.g., [Wentzel et al., 2016](#)). In contrast, not all teacher–student relationships are positive. They also can be marked by conflict, criticism, and distrust ([Brinkworth et al., 2018](#)), and characterized by ineffective teaching, lack of help, and inattention to students' needs for autonomy and relatedness ([Liu et al., 2018](#)).

The literature on teacher–student relationships and student outcomes is described next.

Teacher–student relationships and motivation

Teacher–student relationships marked by emotional closeness and emotional support has been related significantly to students' motivation throughout the school-aged years, including school liking, students' mastery and performance goal orientations, academic values, interest, engagement, and self-efficacy (see [Wentzel, 2016](#) for a review). In addition, when students perceive their relationships with teachers as caring and emotionally supportive, they are more likely to report satisfaction of needs for autonomy, relatedness, and competence (see, Self-Determination Theory, [Deci and Ryan, 2002](#); [Froiland et al., 2019](#)). Over the course of a school year, perceived emotional support from teachers has predicted a positive sense of autonomy, engagement and mastery motivation ([Dietrich et al., 2015](#); [Ruzek et al., 2016](#)), a sense of community and collective efficacy (a belief in the ability to achieve common goals; [Bandura, 1986](#)), and a sense of positive well-being ([Capone et al., 2018](#)).

Other researchers have documented a cumulative effect of teacher–student relationships such that the greater number of positive relationships students have with teachers, the more likely they are to participate in academics, enjoy their work, and aspire to higher levels of education ([Martin and Collie, 2019](#)), and less likely to experience emotional distress and anxiety at school ([Connor et al., 2014](#)). Finally, in a review of 46 studies of teacher–student relationships, [Quin \(2017\)](#) concluded that positive relationships with teachers were related to students' attendance, and grades, even when taking into account other personal, family school and teacher characteristics.

Teacher–student relationships and achievement

Meta-analyses have documented significant relations between positive teacher–student relationships and academic outcomes in young children (e.g., [Kincaid et al., 2020](#)). [Hamre and Pianta \(2001\)](#) found kindergartners' relationships with teachers marked by conflict and dependency predicted lower grades and standardized test scores over the course of eight years. Meta-analytic studies also have linked perceived emotional support and positive relationships with teachers consistently to academic achievement in older students ([Quin, 2017](#); [Roorda et al., 2011](#)).

Teacher–student relationships and social-behavioral outcomes

Positive teacher–student relationships also have been related to students' socially competent behavior, whereas relationships marked by conflict have been related positively to children's aggressive and socially incompetent behavior and peer rejection (e.g., [Ladd and Burgess, 2001](#)). [Hamre and Pianta \(2001\)](#) documented associations between teacher–student conflict in kindergarten and fewer positive work-habits and increased numbers of disciplinary infractions through eighth grade, especially for boys. Other longitudinal studies have documented that qualities of teacher–student relationships in Kindergarten predict social-emotional outcomes in first and second grade (e.g., [Peisner-Feinberg et al., 2001](#); [Pianta and Stuhlman, 2004](#)). In older samples, positive teacher–student relationships have been related to students' social-emotional functioning, including prosocial and responsible behavior ([Wentzel, 2013](#)). Students' appraisals of teacher emotional support also have been related to students' pursuit of goals to behave in prosocial and responsible ways (see [Wentzel, 2016](#) for a review of this research).

Relationships with peers

Positive relationships with peers also have been related consistently to a range of positive academically-related accomplishments ([Wentzel, 2017](#)). However, unlike teacher–student relationships, peer relationships have been studied in a variety of ways,

including dyadic friendships, the degree of peer acceptance by the larger peer group, and membership in specific peer groups (see, [Juvonen et al., 2012](#); [Wentzel, 2017](#)). Peer relationships also are examined within structured interactions related to instruction (e.g., cooperative and collaborative learning; see, [Wentzel et al., 2021](#); [Wentzel and Watkins, 2011](#)).

To assess dyadic friendships, researchers ask students to nominate their best friends at school; often, nominations are then matched to determine reciprocity, or best friendships. Peer acceptance and social (e.g., sociometric) status is typically determined by unilateral assessments of a child's relative standing or reputation within a larger group of peers, such as a group of classmates. Based on these assessments, students are assigned to a sociometric status group (i.e., popular, rejected, neglected, controversial, and average status; see [Asher and Dodge, 1986](#)), or described in terms of overall acceptance or rejection by peers. As such, peer acceptance or social status of a student is determined by a diverse set of peers who are not necessarily friends with the student and with whom interactions might be infrequent.

Membership in peer crowds and groups is typically determined by identifying clusters of students who are friends with each other using statistical procedures (e.g., [Kindermann and Gest, 2009](#)) or by asking students to identify groups characterized by activities (e.g., sports) or behavioral characteristics (e.g., substance use) or, more simply, those who spend time together (e.g., who hangs out with whom) ([Brown, 1989](#)). Students' membership in specific peer crowds and groups has been studied most frequently in adolescent samples (see [Brown, 1989](#); [Brown and Dietz, 2009](#)). Adolescent crowds often include "Populars" ("Popular" crowds are often described in negative terms such as being dominant and exclusionary ([Parkhurst and Hopmeyer, 1998](#)), or in terms of students who engage in positive forms of academic as well as social behavior), "Jocks" (characterized by athletic accomplishments), "Druggies" (engaged in delinquent and other illicit activities), and "Normals" (fairly average students). Research on peer crowds has been mostly descriptive, identifying the central norms and values that uniquely characterize each crowd. A related construct, peer networks, reflects groups of students formed on the basis of mutual friendships.

The central distinction between having friends and involvement with larger peer groups is that friendships reflect relatively private, egalitarian relationships often formed on the basis of idiosyncratic criteria. In contrast, peer groups are defined by publicly acknowledged and therefore easily identified and predictable characteristics that are valued by the group. In addition, whereas friendships are enduring aspects of children's peer relationships at all ages, peer groups and crowds emerge primarily during middle school, peak at the beginning of high school, and then diminish in prevalence as well as influence by the end of high school ([Brown, 1989](#)).

Each of these peer relationship "types" has been studied in association with students' school-based functioning, including students' motivation, achievement, and social-behavioral outcomes. This literature is described in the following sections.

Peer relationships and motivation

A broad range of motivation outcomes have been studied with reference to peers (see [Kindermann and Skinner, 2012](#); [Wentzel, 2017](#); [Wentzel and Muenks, 2016](#), for reviews). Sociometric status and peer acceptance have been related to pursuit of goals to learn, interest in school, and perceived academic competence across the school-age years. In kindergarten, children who are rejected by their peers tend to engage in class activities less often than those who enjoy high levels of peer acceptance, and they report strong desires to avoid school; having friends is associated with positive school affect and school liking. In older elementary school students, peer acceptance has been related positively to academic self-concept over time ([Flook et al., 2005](#)), while peer rejection has been linked negatively to children's self-concept and school engagement ([Buhs, 2005](#)).

During middle school, having friends has been related to positive aspects of academic and social motivation and engagement in school-related activities. Peer group membership has been associated with liking and enjoyment of school, and pursuit of learning-related goals over time. During high school, peer acceptance and group membership also have been related to a range of motivation outcomes (e.g., goals, efficacy, interest, effort). In addition, elementary-age students tend to self-select into groups of peers that have motivational orientations to school similar to their own and these orientations often become stronger and more similar within groups over the course of the school year (e.g., [Kindermann, 1993](#)). Similar findings have been reported in middle school samples (e.g., [Wentzel and Caldwell, 1997](#)).

Peer relationships and academic achievement

Peer relationships also have been related to academic achievement. Classroom grades and IQ have been related consistently to peer status and acceptance for elementary-aged children as well as adolescents (see [Cillessen and van den Berg, 2012](#); [Wentzel, 2017](#) for reviews). In kindergarten and elementary school, children who are rejected by their peers tend to perform less well on tests of achievement in comparison to those with positive peer relationships. Longitudinal studies have documented the stability of relations between peer acceptance and academic accomplishments over time.

Friendship also has been associated with academically-related outcomes (e.g., [Wentzel et al., 2021](#)). Simply having friends has been related positively to grades and test scores at all ages, whereas students without a reciprocated friend tend to adjust less well to academic demands. The power of peer groups and crowds to influence academic achievement is supported by findings that adolescent peer groups differ in the degree to which they pressure members to become involved in academic activities ([Brown and Dietz, 2009](#)).

Peer relationships and social-behavioral outcomes

An extensive body of work has documented associations between peer acceptance and social behavioral outcomes (see [Wentzel, 2013](#); [Wentzel and Muenks, 2016](#) for reviews). In general, socially accepted and popular students at all ages tend to be more pro-social and sociable and less aggressive, and rejected students less compliant, less self-assured, less sociable and more aggressive and withdrawn. Students' friendships and peer groups also have been associated with social-behavioral outcomes. Children with friends tend to be more sociable, cooperative, and self-confident when compared to their peers without friends; children with reciprocated friendships also tend to be more independent, emotionally supportive, altruistic and prosocial, and less aggressive than those who do not have such friendships ([Newcomb and Bagwell, 1995](#)). As with peer acceptance, these findings appear to be robust for students of all ages.

Relationship provisions and affordances

As summarized above, aspects of students' relationships with teachers and peers are related to motivation, achievement, and social-behavioral outcomes, and these findings appear to be robust for school-aged children from Kindergarten through high school. Of additional interest in the field is why these relationships and interactions predict positive outcomes at school. One approach to this issue has been to the benefits of and the social supports afforded by interactions and ongoing relationships with teachers and peers.

Relationship provisions or supports that are most likely to occur at school are: communications concerning norms and expectations, instrumental help, emotional support and safety, and rewards and reinforcements. Specifically, students will come to value and subsequently pursue academic and social goals valued by teachers and peers when they perceive their interactions and relationships with them as providing clear direction concerning goals that should be achieved; as facilitating the achievement of their goals by providing help, advice, and instruction; as being safe and responsive to their goal strivings; and as being emotionally supportive and nurturing (see, [Ford, 1992](#); [Wentzel, 2018](#)). In this manner, students' school-based competencies are a product of social reciprocity between a student and her teachers and classmates. Just as students must behave in ways that meet the expectations of others, so must teachers and peers provide support for the positive cognitive and social-emotional functioning of students. These supports are described in the following sections.

Communicating norms and expectations

Norms and expectations are communicated primarily by way of input from teachers and peers. In general, they concern socially-valued goals and expectations for academic performance and social behavior, and standards against which judgments of personal efficacy can be made. In other words, they define, clarify, maintain, and enforce academic and social norms.

Teachers

Teachers who create a dependable classroom structure provide students with the information and assistance they need to learn successfully. Structure is created by communicating expectations clearly; by responding consistently, predictably, and contingently to students' behavior; by offering help and support to those who are struggling; and by adjusting teaching strategies to individual differences. These practices promote a climate of interpersonal trust and fairness that promotes students' willingness to listen to teacher communications and adopt desired behavioral and learning goals and values. Students who experience optimal levels of this kind of structure are likely to be the most effortful, persistent, and highly engaged in classroom activities ([Skinner and Belmont, 1993](#)).

Teachers also convey expectations about ability and performance to individual students. As part of ongoing interpersonal interactions, these communications have the potential to influence a student's beliefs about her own ability and goals to achieve academically, especially when teachers' expectations result in their differential treatment of students ([Weinstein, 2002](#)). These expectations can reflect beliefs that students are able to achieve more than previously demonstrated, or reflect underestimations of student ability. The effects of these latter negative expectations appear to be fairly weak and short-lived, although they can serve as powerful self-fulfilling prophecies that can have stronger effects on students who are African-American, from low socio-economic backgrounds, and low achievers (see [Jussim et al., 2009](#)). In contrast, teachers' overestimations of ability seem to have a somewhat stronger effect in raising levels of achievement, especially for low performing students ([Madon et al., 1997](#)). Therefore, teachers who communicate high expectations for individual students can bring about positive changes in academic accomplishments.

Peers

Peers can be especially influential in communicating norms and expectations. Peers who convey expectations that academic engagement and positive social interactions are important and enjoyable are likely to lead others to form similar values and goals ([Bandura, 1986](#)). Empirical evidence that peers communicate expectations and opinions is suggested in findings of associations between

peer group membership and student outcomes (e.g., Kindermann, 1993; Kiuru et al., 2007). To illustrate, Hamm et al. (2013) found that middle school students who had sustained membership in a social network with norms supportive of effort and achievement were more likely to value school. With respect to behavioral outcomes, perceived expectations from peers for behaving prosocially are significant predictors of internalized values for and displays of prosocial behavior (Wentzel et al., 2007); peer communications concerning antisocial behavior appear to increase levels of aggression and misconduct at school (Ehrenreich et al., 2014).

Modeling is also a powerful process by which information is communicated. For example, students might develop specific behavioral styles or interests because they are desirable characteristics modeled by their peers (e.g., Barry and Wentzel, 2006). Children utilize their peers for comparative purposes as early as four years of age (Butler, 2005). Elementary school students have been found to choose difficult academic goals for themselves only when they observe a peer who prefers difficult academic goals (Sagot-sky and Lepper, 1982). Peer modeling is also particularly effective for supporting positive academic self-efficacy. As children work on academic tasks that require fairly specific skills and are evaluated with respect to clearly defined standards, they use each other to monitor and evaluate their own abilities (Schunk and Pajares, 2009).

Modeling effects seem to be especially likely to occur when students are friends (e.g., Ricciardelli and Mellor, 2012). Although evidence is indirect, changes in young children's competence perceptions from fall to spring have been associated with the competence perceptions of their very best friends. Adolescents' perceptions of their friends' achievement-related behavior, engagement, and academic performance have predicted their own self-concept, achievement-related choices, goal pursuit, engagement, and academic performance (see, Wentzel, 2018).

Finally, the influence of peer crowds on adolescent functioning is illuminated in ethnographic studies that describe how peer crowds facilitate the formation of students' identity and self-concept and structure their ongoing social interactions (Brown, 1989). With respect to identity formation, crowds are believed to provide adolescents with values, norms, and interaction styles that are sanctioned and commonly displayed. Behaviors that are characteristic of a crowd are modeled frequently so that they can be learned easily and adopted by individuals. In this manner, crowds provide prototypical examples of various identities (including that of being a good student) for those who wish to "try out" different lifestyles and, in doing so, can affirm an adolescent's sense of self.

It should be noted that peer expectations concerning appropriate behavior will contribute to the development of competencies valued by teachers and other adults only insofar as the peer group believes that adult standards for achievement and norms for conduct are important and legitimate. As children enter adolescence, however, they are less likely to acknowledge the legitimacy of adult-imposed norms (Smetana and Bitz, 1996) or automatically enforce classroom rules (Eccles and Midgley, 1989). Therefore, dependence on peers to communicate and enforce adult-generated rules might not be appropriate for many older students.

Instrumental aid: provisions of help, advice, and instruction

Instrumental aid in the form of help-giving, advice and direct instruction is a powerful support that facilitates students' academic and social adjustment to school. Social cognitive theory (Bandura, 1986) suggests that direct instruction and modeling of skills are especially powerful mechanisms whereby students learn the skills that enable them to go about meeting teacher and peer expectations.

Teachers

Teachers play the central role of transmitting knowledge and training students in academic subject areas by providing information and advice, modeled behavior, or specific experiences that facilitate learning as well as interpersonal competence. Brophy and Good (1974) documented the relevance of teachers' relationships with elementary-aged students for gaining access to academic resources. Teachers in their study reported that they were more appreciative and positive toward students who were cooperative and persistent than toward students who were less cooperative but displayed high levels of creativity and achievement. Teachers also responded with help and encouragement to students about whom they were concerned when these students sought them out for help. In contrast, students toward whom they felt rejection were treated most often with criticism and typically were refused help.

Peers

Students also communicate important information, teach valuable skills, and provide instrumental help to each other. Help giving is perhaps the most explicit and obvious way in which peers can have a direct influence on students' academic and social competence. During adolescence, students report that their peers are as or more important sources of instrumental aid than their teachers (Lempers and Clark-Lempers, 1992). Peers also provide input concerning reasons for engaging in academic tasks (e.g., because it is important or fun; Wentzel, 2004). Finally, advice and feedback from peers following success or failure can lead to adjustments in perceived competence and expectations for future success (e.g., Altermatt et al., 2002).

Social and emotional support and safety

A number of theoretical perspectives, including attachment theory (e.g., Bretherton, 1987), social support perspectives (Sarason et al., 1990) and self-determination theory (Ryan and Deci, 2000) propose that receiving emotional support conveys a message that one is a valued and integral member of a relationship or of a larger social group (Connell and Wellborn, 1991). In turn, feelings of emotional security and being socially connected are believed to facilitate the adoption of goals and interests valued by others, including goals to contribute in positive ways to the overall functioning of the social group (e.g., Deci and Ryan, 2002). This appears to be the case both for teachers and peers.

Teachers

In conjunction with communicating clear expectations and providing instrumental help, teachers also create contexts characterized by levels of emotional support and personal safety (Connell and Wellborn, 1991). As noted earlier, emotionally supportive relationships have the potential to influence student motivation, achievement and social-behavioral outcomes. An additional aspect of teachers' emotional support is reflected in their efforts to protect students' physical well-being. Most frequently, issues of student safety are discussed, with regard to peer interactions, in the form of aggression and bullying. However, teachers can play a central role in creating classrooms that are free of peer harassment and in alleviating its negative effects once it has occurred (Olweus and Limber, 2009). Of special interest are findings that students are more likely to enjoy affectively positive relationships with teachers when they feel safe at school (e.g., Crosnoe et al., 2004).

Teachers also can promote positive classroom climates by providing students with help in forming positive relationships with each other. Although not much research has been done in this area, several conclusions can be drawn. First, teacher communications and instructional practices can influence peer relationships. For example, young students appear to be aware of the ability and behavioral expectations their teachers hold for their fellow classmates and tend to reject or accept their peers based on these perceived expectations (e.g., Donohue et al., 2003). In addition, students tend to have more close friends and a greater number of friends, and are less likely to experience peer rejection when teachers use learner-centered practices (e.g., involving students in decision making) as opposed to teacher-centered practices (e.g., focusing on rote learning, norm-referenced evaluation; Donohue et al., 2003).

Evidence also suggests that teachers who are more attuned to the peer dynamics of their classrooms tend to have greater influence over the social cultures of their classrooms than those who are not attuned. Teacher attunement to peer groups has been related to students' sense of school belonging, levels of peer acceptance, and with greater intolerance and disapproval of bullying and aggression among students (Hamm et al., 2011). Finally, perhaps the most straightforward method of promoting positive relationships among students is to employ cooperative learning activities as part of instruction. Cooperative learning activities offer motivational benefits because they respond to students' relatedness needs, as well as potential learning benefits because they engage students in the social construction of knowledge (see, Slavin et al., 2003).

Peers

Also noted earlier, students who perceive that their peers care about them tend to be interested and engaged in academic pursuits and positive forms of behavior, whereas students who do not perceive their relationships with peers as positive and caring tend to be at risk for motivational and academic problems and displaying antisocial behavior. The impact on school-related functioning likely comes from the positive opportunities and supports afforded to children with accepted peers, such as effective role models, direct instruction concerning normative behavior and skill development, and sources of positive reinforcements for social and academic behavior. The importance of social acceptance and rejection for understanding student motivation is underscored by findings that strong affective bonds and perceived support from others can serve as buffers from stress and anxiety and contribute to a positive sense of emotional well-being (Sarason et al., 1990). In turn, these positive feelings are believed to facilitate the adoption of goals and interests valued by others (e.g., Deci and Ryan, 2002).

In addition, students who are accepted by their peers and who have established friendships with classmates also are more likely to enjoy a relatively safe school environment and less likely to be the targets of peer-directed violence and harassment than their peers who do not have friends (e.g., Schwartz et al., 2000). Theorists also have argued that goals to sustain positive group identity and cohesion can result in exclusionary practices within and between groups (e.g., Bennett, 2014), including the formation of stereotypes and discriminatory practices (e.g., Dovidio et al., 2005). These group-level processes can have a profound effect on students' emotional well-being (Brown et al., 2008), their access to opportunities and supports (Haslam et al., 2012), and motivational beliefs, including a sense of group efficacy (Bandura, 1986).

The general effects of peer harassment on student motivation and academic competence have not been studied frequently. However, students who are frequently victimized also tend to report higher levels of distress and depression than those who are not (e.g., Kochenderfer-Ladd and Wardrop, 2001). Few studies have identified pathways whereby peer victimization affect academic outcomes. However, as with perceived support, peer abuse and exclusion is likely to be associated with academic achievement by way of emotional distress (Buhs, 2005; Flook et al., 2005). Therefore, although indirect, having supportive peers in these negatively charged situations can have positive effects on a wide range of social, motivational, and academic outcomes.

Rewards and reinforcements

Teachers and peers can also exert influence by way of reinforcements and rewards. These reinforcements can be extrinsic, conveyed through tangible benefits (e.g., help, social inclusion, or exclusion), feedback and instruction (e.g., praise or criticism), and coercive strategies (e.g., bullying, persuasion). They also can be intrinsic in that students might engage in behavior valued by teachers and peers because they experience it as enjoyable and of personal value.

Teachers

An extensive literature has documented ways in which teachers administer rewards to students through instruction, feedback, and behavioral management strategies (see [Wentzel, 2020](#), for a review). For example, instructional strategies that involve competition, salient rewards (e.g., food, stickers, free time), or public feedback reflect extrinsically-motivating strategies that can lead to perceptions by students that engaging in the activity is controlling and a threat to autonomy and choice. These effects are especially powerful when students perceive themselves to be engaged in academic tasks in order to please an authority figure, obtain a reward, or escape punishment. However, extrinsic rewards can also be valuable because they send a message concerning what students are supposed to do to achieve success in school, whether or not they find activities interesting or enjoyable.

Peers

Peers have been studied most often with respect to negative reinforcements and reward systems. In this literature, peers' use of intimidation, power assertion, and negative reinforcement has been the focus of discussions concerning peer pressure and how peer groups function to establish normative standards and power hierarchies (e.g., [Cairns et al., 1989](#)). To illustrate, popular children often exert power over others by using social acceptance and status as a reward for compliance and conformity (e.g., [Sandstrom, 2011](#)). Simultaneous use of prosocial (positive rewards) as well as coercive and aggressive (negative reinforcement) tactics appears to be a particularly effective strategy to gain social control and power among peers (e.g., [Hawley et al., 2002](#)), and to solidify and maintain group norms and practices ([Cairns et al., 1989](#)). An extensive literature has also demonstrated how negative reinforcement in the form of physical and relational aggression results in a broad range of outcomes related to motivation, emotional functioning, and social behavior, for perpetrators as well as victims (e.g., [Crick et al., 2009](#)). With respect to academics, evidence suggests that students can experience negative reinforcement for doing well academically through labeling and stigmatization ([Boehnke, 2008](#)).

Summary

Support for a focus on relationship provisions at school is reflected in findings that expectations for academic performance, help, emotional support and safety, and rewards and reinforcements from teachers and peers have the potential to contribute to a broad range of motivational, achievement and social behavioral outcomes at all ages (see [Wentzel, 2013](#)). Clear communication and instrumental help in the form of direct instruction and modeling are powerful mechanisms whereby students learn what is expected of them and how to go about meeting those expectations. These mechanisms also can result in the development of skills and abilities that bolster one's sense of efficacy and positive beliefs about control over the learning process. Reinforcements and rewards from teachers and peers are additional and important sources of information concerning socially-valued goals, values, and perceptions of ability and control. These social cues convey to students the potential affective and interpersonal consequences if they do not behave or perform as expected. Finally, levels of emotional closeness, security and safety associated with teacher-student relationships can play a causal role in the development of a range of positive motivational outcomes in children.

Conclusions

There is no doubt that teachers and peers can have a powerful impact on a student's day-to-day life at school. Relationships and interactions that demonstrate essential supports and caring are important facilitators of student motivation, achievement, and a range of social competencies. It is troubling, therefore, that some researchers have reported that up to 25% of elementary-age students believe that school is an uncaring place ([Watson et al., 1989](#)). In middle school, these percentages can be even greater (e.g., [Wentzel, 1997](#)). In these cases, students feel alienated from the life of the classroom and become disengaged and unmotivated ([Legault et al., 2006](#)). In addition, research suggests that supportive relationships might be more important for some students than for others. For example, the literature on teacher-student relationships suggests that students from lower SES backgrounds and members of minority groups tend to benefit more from close relationships than do other students; school-level factors such as safety, racial homogeneity, SES of the student body, and composition of instructional teams also appear to impact the degree to which perceived support is related to student outcomes (see [Wentzel, 2016](#)).

These findings suggest that implementing programs to improve teacher–student and peer relationships should be a central component of school policy. Although intervention efforts have been infrequent, several programs appear to be promising and noteworthy. For example, the Responsive Classroom (RC) approach has been associated with improved teacher–student relationships and classroom behavior in addition to academic gains (Rimm-Kaufman and Chiu, 2007; Rimm-Kaufman et al., 2007). Designed to increase teachers' positive interactions with students and promote sensitive instructional practices, RC also has demonstrated improvement in positive peer interactions of students high in disruptive behavior (Mikami et al., 2012). The REAL program, which includes a focus on teachers' understanding of classroom social dynamics, also has demonstrated positive gains in students' motivation and academic outcomes, along with changes in peer norms for academic effort and achievement (Hamm et al., 2011). The Child Development Project (CDP; Developmental Studies Center) has been successful at improving classroom management practices and interpersonal relationships.

In addition, teachers and school administrators can play a central role in creating schools that are free of peer harassment and in alleviating the negative effects of harassment once it has occurred (e.g., Olweus and Limber, 2009). Interventions designed to offset the often negative influence of peer groups and gangs on behavior and school attendance are especially successful if students have access to adults who provide them with warmth and strong guidance (e.g., Chaskin, 2010). Schools that stress intergenerational bonding also support the development of positive teacher–student relationships that can act as buffers against the potentially negative effects of aggressive peers on behavior (Crosnoe and Needham, 2004). Utilizing these tools to improve teacher–student and peer relationships at school has the potential to support the development of school-based competencies for all children.

References

- Altermatt, E.R., Pomerantz, E.M., Ruble, D.N., Frey, K.S., Greulich, F.K., 2002. Predicting changes in children's self-perceptions of academic competence: a naturalistic examination of evaluative discourse among classmates. *Dev. Psychol.* 38, 903–917.
- Asher, S.R., Dodge, K.A., 1986. Identifying children who are rejected by their peers. *Dev. Psychol.* 22, 444–449.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Barry, C.M., Wentzel, K.R., 2006. Friend influence on prosocial behavior: the role of motivational factors and friendship characteristics. *Dev. Psychol.* 42, 153–163.
- Bennett, M., 2014. Intergroup social exclusion in childhood: forms, norms, context, and social identity. *J. Soc. Issues* 70, 183–195.
- Boehnke, K., 2008. Peer pressure: a cause of scholastic underachievement? A cross-cultural study of mathematical achievement among German, Canadian, and Israeli middle school students. *Soc. Psychol. Educ.* 11, 149–160.
- Bretherton, I., 1987. New perspectives on attachment relations: security, communication and internal working models. In: Osofsky, J. (Ed.), *Handbook of Infant Development*. John Wiley, New York, pp. 1061–1100.
- Brinkworth, M.E., McIntyre, J., Juraschek, A.D., Gehlbach, H., 2018. Teacher–student relationships: the positives and negatives of assessing both perspectives. *J. Appl. Dev. Psychol.* 55, 24–38.
- Bronfenbrenner, U., 1989. Ecological systems theory. In: Vasta, R. (Ed.), *Annals of Child Development*, vol. 6. JAI, Greenwich, CT, pp. 187–250.
- Brophy, J.E., Good, T.L., 1974. *Teacher–Student Relationships: Causes and Consequences*. Holt, Rinehart & Winston, Inc, New York.
- Brown, B.B., Dietz, E.L., 2009. Informal peer groups in middle childhood and adolescence. In: Rubin, K.H., Bukowski, W.M., Laursen, B. (Eds.), *Handbook of Peer Interactions, Relationships and Groups*. Guilford Press, New York, NY, pp. 361–376.
- Brown, B.B., Bakken, J.P., Ameringer, S.W., Mahon, S.D., 2008. A comprehensive conceptualization of the peer influence process in adolescence. In: Prinstein, M., Dodge, K. (Eds.), *Understanding Peer Influence in Children and Adolescents*. Guilford, New York, pp. 17–44.
- Brown, B.B., 1989. The role of peer groups in adolescents' adjustment to secondary school. In: Berndt, T.J., Ladd, G.W. (Eds.), *Peer Relationships in Child Development*. Wiley, New York, NY, pp. 188–215.
- Buhs, E.S., 2005. Peer rejection, negative peer treatment, and school adjustment: self-concept and classroom engagement as mediating processes. *J. Sch. Psychol.* 43, 407–424.
- Butler, R., 2005. Competence assessment, competence, and motivation between early and middle childhood. In: Elliot, A., Dweck, C. (Eds.), *Handbook of Competence and Motivation*. Guilford, New York, pp. 202–221.
- Cairns, R.B., Neckerman, H.J., Cairns, B.D., 1989. Social networks and shadows of synchrony. In: Adams, G.R., Gullota, T.P., Montemayor, R. (Eds.), *Advances in Adolescent Development*. Sage, Beverly Hills, CA.
- Capone, V., Donizzetti, A.R., Petrillo, G., 2018. Classroom relationships, sense of community, perceptions of justice, and collective efficacy for students' social well-being. *J. Community Psychol.* 46 (3), 374–382.
- Chaskin, R.J., 2010. *Youth Gangs and Community Intervention: Research, Practice, and Evidence*. Columbia University Press, New York, NY.
- Cillessen, A., van den Berg, Y., 2012. Popularity and school adjustment. In: Ryan, A., Ladd, G. (Eds.), *Peer Relationships and Adjustment at School*. IAP Information Age, Charlotte, NC, pp. 135–164.
- Connell, J.P., Wellborn, J.G., 1991. Competence, autonomy, and relatedness: a motivational analysis of self-system processes. In: Gunnar, M.R., Sroufe, L.A. (Eds.), *Self Processes and Development: The Minnesota Symposia on Child Development*, vol. 23. Erlbaum, Hillsdale, NJ, pp. 43–78.
- Conner, J.O., Miles, S.B., Pope, D., 2014. How many teachers does it take to support a student? Examining the relationship between teacher support and adverse health outcomes in high-performing, pressure-cooker high schools. *High Sch. J.* 98 (1), 22–42.
- Crick, N., Murray-Close, D., Marks, P., Moharjeri-Nelson, S., 2009. Aggression and peer relationships in school-age children: relational and physical aggression in group and dyadic contexts. In: Rubin, K., Bukowski, W., Laursen, B. (Eds.), *Handbook on Peer Interactions, Relationships, and Groups*. Guilford, New York, NY, pp. 287–302.
- Crosnoe, R., Needham, B., 2004. Holism, contextual variability, and the study of friendships in adolescent development. *Child Dev.* 75, 264–279.
- Crosnoe, R., Johnson, M.K., Elder Jr., G.H., 2004. Intergenerational bonding in school: the behavioral and contextual correlates of student–teacher relationships. *Sociol. Educ.* 77, 60–81.
- Deci, E., Ryan, R., 2002. *Handbook of Self-Determination Research*. University of Rochester Press, Rochester, NY.
- Developmental Studies Center, (n.d.). Retrieved August 2, 2007, from www.devstu.org/.
- Dietrich, J., Dicke, A., Kracke, B., Noack, P., 2015. Teacher support and its influence on students' intrinsic value and effort: dimensional comparison effects across subjects. *Learn. Instruct.* 39, 45–54.
- Donohue, K., Perry, K., Weinstein, R., 2003. Teachers' classroom practices and children's rejection by their peers. *J. Appl. Dev. Psychol.* 24, 91–118.
- Dovidio, J.F., Gaertner, S.L., Hodson, G., Houlette, M.A., Johnson, K.M., 2005. Social inclusion and exclusion: recategorization and the perception of intergroup boundaries. In: Abrams, D., Hogg, M., Marques, J. (Eds.), *The Social Psychology of Inclusion and Exclusion*. Psychology Press, New York, NY, US, pp. 245–264.

- Eccles, J.S., Midgley, C., 1989. Stage-environment fit: developmentally appropriate classrooms for young adolescents. In: Ames, C., Ames, R. (Eds.), *Research on Motivation in Education*, vol. 3. Academic Press, New York, pp. 139–186.
- Ehrenreich, S., Underwood, M., Ackerman, R., 2014. Adolescents' text message communication and growth in antisocial behavior across the first year of high school. *J. Abnorm. Child Psychol.* 42, 251–264.
- Flook, L., Repetti, R.L., Ullman, J.B., 2005. Classroom social experiences as predictors of academic performance. *Dev. Psychol.* 41, 319–327.
- Ford, M., 1992. *Motivating Humans: Goals, Emotions, and Personal Agency Beliefs*. SAGE, Newbury Park, CA.
- Froiland, J.M., Worrell, F.C., Oh, H., 2019. Teacher–student relationships, psychological need satisfaction, and happiness among diverse students. *Psychol. Sch.* 56 (5), 856–870. <https://doi-org.proxy-um.researchport.umd.edu/10.1002/pits.22245>.
- Hamm, J.V., Farmer, T.W., Dadisman, K., Gravelle, M., Murray, A.R., 2011. Teachers' attunement to students' peer group affiliations as a source of improved student experiences of the school social-affective context following the middle school transition. *J. Appl. Dev. Psychol.* 32 (5), 267–277.
- Hamm, J.V., Lambert, K., Agger, C.A., Farmer, T.W., 2013. Promotive peer contexts of academic and social adjustment among rural African American early adolescent boys. *Am. J. Orthopsychiatr.* 83, 278–288.
- Hamre, B.K., Pianta, R.C., 2001. Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Dev.* 72, 625–638.
- Haslam, S., Reicher, S., Levine, M., 2012. When other people are heaven, when other people are hell: how social identity determines the nature and impact of social support. In: Jetten, J., Haslam, C., Haslam, S. (Eds.), *The Social Cure: Identity, Health and Well-Being*. Psychology Press, New York, NY, US, pp. 157–174.
- Hawley, P.H., Little, T.D., Pasupathi, M., 2002. Winning friends and influencing peers: strategies of peer influence in late childhood. *Int. J. Behav. Dev.* 26, 466–474.
- Jussim, L., Robustelli, S., Cain, T., 2009. Teacher expectations and self-fulfilling prophecies. In: Wentzel, K.R., Wigfield, A. (Eds.), *Handbook of Motivation at School*. Taylor Francis, New York, NY, pp. 349–380.
- Juvonen, J., Espinoza, E., Knifsend, C., 2012. The role of peer relationships in student academic and extracurricular engagement. In: Christenson, S.L., Reschly, A.L., Wylie, C. (Eds.), *Handbook on Student Engagement*. Springer, New York, NY, pp. 387–401.
- Kincaid, L., Cook, C., Goerdt, A., 2020. Meta-analysis and common practice elements of universal approaches to improving student–teacher relationships. *Rev. Educ. Res.* 90 (5), 710–748.
- Kindermann, T.A., Gest, S.D., 2009. Assessment of peer group: identifying naturally occurring social networks and capturing their effects. In: Rubin, K.H., Bukowski, W., Laursen, B. (Eds.), *Handbook of Peer Interactions, Relationships, and Groups*. Guilford, New York, pp. 100–117.
- Kindermann, T.A., Skinner, E.A., 2012. Will the real peer group please stand up? A “tensegrity” approach to examining the synergistic influences of peer groups and friendship networks on academic development. In: Ryan, A., Ladd, G. (Eds.), *Peer Relationships and Adjustment at School*. IAP Information Age, Charlotte, NC, pp. 51–77.
- Kindermann, T.A., 1993. Natural peer groups as contexts for individual development: the case of children's motivation in school. *Dev. Psychol.* 29, 970–977.
- Kiuru, N., Aunola, K., Vuori, J., Nurmi, J.E., 2007. The role of peer groups in adolescents' educational expectations and adjustment. *J. Youth Adolesc.* 36, 995–1009.
- Kochenderfer-Ladd, B., Waldrop, J.L., 2001. Chronicity and instability of children's peer victimization experiences as predictors of loneliness and social satisfaction trajectories. *Child Dev.* 72, 134–151.
- Ladd, G.W., Burgess, K.B., 2001. Do relational risks and protective factors moderate the linkages between childhood aggression and early psychological and school adjustment? *Child Dev.* 72, 1579–1601.
- Legault, L., Green-Demers, I., Pelletier, L., 2006. Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *J. Educ. Psychol.* 98 (3), 567–582.
- Lempers, J.D., Clark-Lempers, D.S., 1992. Young, middle, and late adolescents' comparisons of the functional importance of five significant relationships. *J. Youth Adolesc.* 21, 53–96.
- Liu, P.P., Savitz-Romer, M., Perella, J., Hill, N., Liang, B., 2018. Student representations of dyadic and global teacher–student relationships: perceived caring, negativity, affinity, and differences across gender and race/ethnicity. *Contemp. Educ. Psychol.* 54, 281–296.
- Madon, S., Jussim, L., Eccles, J., 1997. In search of self-fulfilling prophecy. *J. Pers. Soc. Psychol.* 72, 791–809.
- Martin, A.J., Collie, R., 2019. Teacher–student relationships and students' engagement in high school: does the number of negative and positive relationships with teachers matter? *J. Educ. Psychol.* 111 (5), 861–876.
- Mikami, A.Y., Griggs, M.S., Reuland, M.M., Gregory, A., 2012. Teacher practices as predictors of children's classroom social preference. *J. Sch. Psychol.* 50, 95–111.
- Newcomb, A.F., Bagwell, C.L., 1995. Children's friendship relations: a meta-analytic review. *Psychol. Bull.* 117, 306–347.
- Olweus, D., Limber, S.P., 2009. The Olweus bullying prevention program: implementation and evaluation over two decades. In: Jimerson, S.R., Swearer, S.M., Espelage, D.L. (Eds.), *Handbook of Bullying in Schools: An International Perspective*. Routledge, New York, NY, pp. 377–402.
- Parkhurst, J.T., Hopmeyer, A., 1998. Sociometric popularity and peer-perceived popularity: two distinct dimensions of peer status. *J. Early Adolesc.* 18 (2), 125–144. <https://doi-org.proxy-um.researchport.umd.edu/10.1177/0272431698018002001>.
- Peisner-Feinberg, E.S., Burchinal, M.R., Clifford, R.M., Culkin, M.L., Howes, C., Kagan, S.L., et al., 2001. The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Dev.* 72, 1534–1553.
- Pianta, R.C., Stuhlman, M.W., 2004. Teacher–child relationships and children's success in the first years of school. *Sch. Psychol. Rev.* 33, 444–458.
- Pianta, R., 2006. Classroom management and relationships between children and teachers: implications for research and practice. In: Evertson, C., Weinstein, C. (Eds.), *Handbook of Classroom Management—Research, Practice, and Contemporary Issues*. Erlbaum, Mahwah, NJ, pp. 685–710.
- Quin, D., 2017. Longitudinal and contextual associations between teacher–student relationships and student engagement: a systematic review. *Rev. Educ. Res.* 87 (2), 345–387.
- Ricciardelli, L.A., Mellor, D., 2012. Influence of peers. In: Rumsey, N., Harcourt, D. (Eds.), *The Oxford Handbook of the Psychology of Appearance*. Oxford University Press, New York, pp. 253–272.
- Rimm-Kaufman, S.E., Chiu, Y.J.I., 2007. Promoting social and academic competence in the classroom: an intervention study examining the contribution of the responsive classroom approach. *Psychol. Sch.* 44 (4), 397–413.
- Rimm-Kaufman, S.E., Fan, X., Chiu, Y.J., You, W., 2007. The contribution of the Responsive Classroom Approach on children's academic achievement: results from a three year longitudinal study. *J. Sch. Psychol.* 45 (4), 401–421.
- Roorda, D.L., Koomen, H.M., Spilt, J.L., Oort, F.J., 2011. The influence of affective teacher–student relationships on students' school engagement and achievement: a meta-analytic approach. *Rev. Educ. Res.* 81, 493–529.
- Ruzek, E.A., Hafen, C.A., Allen, J.P., Gregory, A., Mikami, A.Y., Pianta, R.C., 2016. How teacher emotional support motivates students: the mediating roles of perceived peer relatedness, autonomy support, and competence. *Learn. Instr.* 42, 95–103.
- Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78.
- Sagotsky, G., Lepper, M.R., 1982. Generalization of changes in children's preferences for easy or difficult goals induced through peer modeling. *Child Dev.* 53, 372–375.
- Sandstrom, M., 2011. The Power of popularity: influence processes in childhood and adolescence. In: Cillessen, A., Schwartz, D., Mayeux, L. (Eds.), *Popularity in the Peer System*. Guilford, New York, NY, pp. 219–244.
- Sarason, B.R., Sarason, I.G., Pierce, G.R., 1990. Traditional views of social support and their impact on assessment. In: Sarason, B.R., Sarason, I.G., Sarason, G.R. (Eds.), *Social Support: An Interactional View*. Wiley, New York, pp. 9–25.
- Schunk, D.H., Pajares, F., 2009. Self-efficacy theory. In: Wentzel, K.R., Wigfield, A. (Eds.), *Handbook of Motivation at School*. Routledge, New York, pp. 35–54.

- Schwartz, D., Dodge, K.A., Pettit, G.S., Bates, J.E., The Conduct Problems Prevention Research Group, 2000. Friendship as a moderating factor in the pathway between early harsh home environment and later victimization in the peer group. *Dev. Psychol.* 36, 646–662.
- Skinner, E., Belmont, M., 1993. Motivation in the classroom: reciprocal effects of teacher behavior and student engagement across the school year. *J. Educ. Psychol.* 85, 571–581.
- Slavin, R.E., Hurley, E.A., Chamberlain, A., 2003. Cooperative learning and achievement: theory and research. In: Reynolds, W., Miller, G. (Eds.), *Handbook of Psychology, Educational Psychology*, vol. 7. Wiley, New York, pp. 177–198.
- Smetana, J., Blitz, B., 1996. Adolescents' conceptions of teachers' authority and their relations to rule violations in school. *Child Dev.* 67, 1153–1172.
- Watson, M., Solomon, D., Battistich, V., Schaps, E., Solomon, J., 1989. The child development project: combining traditional and developmental approaches to values education. In: Nucci, L. (Ed.), *Moral Development and Character Education: A Dialogue*. McCutchan, Berkeley, pp. 51–92.
- Weinstein, R., 2002. *Reaching Higher: The Power of Expectations in Schooling*. Harvard University Press, Cambridge.
- Wentzel, K.R., Caldwell, K., 1997. Friendships, peer acceptance, and group membership: relations to academic achievement in middle school. *Child Dev.* 68, 1198–1209.
- Wentzel, K.R., Muenks, K., 2016. Peer influence on students' motivation, academic achievement and social behavior. In: Wentzel, K., Ramani, G. (Eds.), *Handbook of Social Influences in School Contexts: Social-Emotional, Motivation, and Cognitive Outcomes*. Taylor Francis, New York, NY, pp. 13–30.
- Wentzel, K.R., Watkins, D.E., 2011. Peer relationships and learning: implications for instruction. In: Mayer, R., Alexander, P. (Eds.), *Handbook of Research on Learning and Instruction*. Routledge, New York, NY, pp. 322–343.
- Wentzel, K.R., Filisetti, L., Looney, L., 2007. Adolescent prosocial behavior: the role of self-processes and contextual cues. *Child Dev.* 78, 895–910.
- Wentzel, K.R., Russell, S., Baker, S.A., 2016. Emotional support and expectations for behavior from peers, teachers, and parents: predictors of adolescent competence at school. *J. Educ. Psychol.* 108, 242–255.
- Wentzel, K.R., Jablansky, S., Scalise, N., 2021. Peer social acceptance and academic achievement: a meta-analytic study. *J. Educ. Psychol.* 157–180.
- Wentzel, K.R., 1997. Student motivation in middle school: the role of perceived pedagogical caring. *J. Educ. Psychol.* 89, 411–419.
- Wentzel, K.R., 2004. Understanding classroom competence: the role of social-motivational and self-processes. In: Kail, R. (Ed.), *Advances in Child Development and Behavior*, vol. 32. Elsevier, New York, NY, pp. 213–241.
- Wentzel, K.R., 2013. School adjustment. In: Reynolds, W., Miller, G. (Eds.), *Handbook of Psychology, Educational Psychology*, vol. 7. Wiley, New York.
- Wentzel, K.R., 2016. Students' relationships with teachers as motivational contexts. In: Wentzel, K., Miele, D. (Eds.), *Handbook of Motivation at School*, second ed. LEA, Mahwah, NJ.
- Wentzel, K.R., 2017. Peer relationships, motivation, and academic performance at school. In: Elliot, A., Dweck, C., Yeager, D. (Eds.), *Handbook of Competence and Motivation*, second ed. Guilford, New York, NY, pp. 586–603.
- Wentzel, K.R., 2018. A competence-in-context approach to understanding motivation at school. In: McInerney, D., Liem, G. (Eds.), *Big Theories Revisited*. IAP, Charlotte, NC, pp. 193–212.
- Wentzel, K.R., 2020. *Motivating Students to Learn*, fourth ed. Taylor Francis, New York, NY.

A social/cultural constructivist perspective on learning “problems”

Curt Dudley-Marling, Boston College, Chestnut Hill, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Social construction of learning	298
A social/cultural constructivist perspective on learning “problems”	300
Implications: a presumption of competence	302
Conclusion	303
References	303

Formal education is highly valued commodity in nearly all cultures and societies. In western, capitalist counties in particular, school success has been linked to a variety of positive outcomes including higher incomes, enhanced social standing, higher self-esteem and even healthier and longer lives (Chowdhury, 2018; OECD, 2013). Conversely, school failure has been associated with poor health, low self-esteem and higher levels of unemployment and incarceration (OECD, 2013; Harlow, 2003; Marks, 2006). Given the stakes, it is no wonder that public education is the subject of fierce debates in the popular press, academic journals and the halls of government.

Despite the desire to leave “no child behind” (No Child Left Behind [NCLB], 2002), failure may be a built-in feature of modern schooling. Two taken-for-granted assumptions underlying schooling, for example, are that educational achievement distributes more or less “normally,” along the lines of a bell curve—meaning that fully half of students will achieve that status of “below average” (Dudley-Marling and Gurn, 2010; Annamma et al., 2013)—and that every student is expected to do better than everyone else (McDermott et al., 2006). Failure is a given in this situation.

Other analyses also implicate the institution of schooling and its employees for widespread school failure. Some economists, for example, have argued that the failures of public schooling can be traced to the lack of economic competition among schools (Friedman, 1955; Chubb and Moe, 1990). Public schooling is a monopoly and, therefore, indifferent to reform. From this perspective, school choice, in the form of vouchers and charter schools, is the remedy for failing schools (see Hursh, 2007 for a review of free market proposals for reforming public schooling). Other critics of public education have focused on the quality of teachers and teacher education. The influential National Council on Teacher Quality (NCTQ), for example, has repeatedly questioned teachers’ preparation for teaching reading, mathematics and content area subjects (NCTQ, 2021).

Whatever the inadequacies of modern public schooling, when individual children struggle to achieve at acceptable levels the tendency is to situate the *problem* in the minds and bodies of children who are presumed to be deficient in motivation, effort or abilities required for academic success. Children who struggle in school are assigned a variety of formal and informal labels by school officials ranging from learning disabled or intellectually impaired to at-risk, slow or struggling learners or simply “lazy.” Whatever the label, the presumption is that the *problem* is intrinsic to the child who must be remediated or “fixed.” An alternative, social constructivist narrative situates school failure outside minds and bodies of students in the complex web of human relations (Gergen, 1990). From this point of view, no one can fail (or succeed) in school on their own. As Collins and Ferri (2016) put it, learning difficulties are “not seen as a problem located in particular bodies or brains ... but rather as a social artifact created and maintained by a lack of fit between a particular learner and his or her context” (p. 8).

This paper begins with an overview of a social constructivist view of learning before turning to a social constructivist perspective on school failure. The final section considers the implications of a shift from an individualistic, deficit-oriented take on learning failure in schools to a socio-cultural perspective.

Social construction of learning

Conventionally, learning is understood to be an individual accomplishment mediated by parents, teachers and other more skilled persons. From this perspective, children learn to talk, read, and write, for example, by acquiring specific sets of skills associated with effective reading, writing, speaking and so on with the help of parents, teachers, siblings and other members of the community. Indeed, learning to speak or read, do math or just find your way to the store, is understood as the mastery of a set of specific skills. And, again, while no one learns completely on their own, learning is generally assumed to be largely the responsibility of the individual.

A social construction of learning offers a radically different perspective both on what is learned and how it is learned. To begin with, a social constructivist perspective holds that people do not learn to read, write, or speak “once and for all.” Instead, they learn to read, write, and speak in particular ways, for particular purposes and audiences, in particular social settings. (Gee, 2012). Specifically, reading, writing, and speaking (and other learning) are sets of social practices that enable participation in particular cultural activities in various cultural communities. Learning language, for instance, involves more than mastering a set of linguistic skills (vocabulary, syntax, etc.). Language practices always involve ways of behaving, thinking, interacting, valuing, believing, and

speaking, in short, becoming particular kinds of people (Gee, 2012). The use of language as part of a religious ritual, for instance, includes saying and doing just the right things in just the right way and the right time in the appropriate setting. Moreover, the right sort of clothing, demeanor, and attitude and beliefs will also be necessary for this social practice to “count” as a religious expression. In other words, participating in this social practice involves being a particular kind of (religious) person. Other social practices, entailing different purposes, settings, participants and so on, involve learning to become different sorts of people. In this sense, learning implies becoming a certain kind of person, indeed, becoming different sorts of people (i.e., different identities) in various social and cultural settings. In other words, what is learned are different relationships among and between people in various contexts; therefore, learning always involves the construction of identities (Lave and Wenger, 1991). Ultimately, the community will adjudicate learning, that is, whether people demonstrate they have learned by showing that they have learned to be the “right sort of person” in particular social/cultural settings.

A social construction of learning indicates that what people learn is largely about participating in various communities—including school. This is also where meaning resides since people always make sense from a particular (cultural) perspective. But this is not just about learning in (a social) context. The sociocultural context is not just where learning takes place or where knowledge is constructed; it is part of what is learned (Cole, 1996). What is learned is always social and, therefore, all knowledge carries social and cultural meanings. Even what appears to be solitary learning has social character since people learn as members of social and cultural groups and what they learn is related to social practices that are enacted in communities. Bakhtin famously argued that language—perhaps the most powerful of cultural tools—is far from a neutral medium since it is “overpopulated with the intentions of others” (Bakhtin, 1981, p. 294). Language, the learning tool extraordinaire, is saturated with social, cultural and historical meanings; therefore, learning mediated by language—as is usually the case for higher order learning—will always have a social character. Given that the primary tool for learning, language, is imbued with a social and cultural quality even what appears to be solitary learning has a social character.

From a social constructivist point of view, learning does not manifest (merely) in the minds and bodies of autonomous individuals. A social construction of learning emphasizes the profoundly social/cultural character of learning. Learning is a process that takes place in participation frameworks and it is in the context of engaging with these frameworks that people learn cultural activities (social practices) that make them into particular kinds of people (Lave and Wenger, 1991). However, people cannot participate or learn to participate in cultural activities or practices on their own. Effective participation in cultural practices always requires coordinated activity among members of the cultural community and, indeed, it is this coordinated activity that is learned. Moreover, coordinated action is *how* learning occurs and *where* it occurs.

The metaphor of a dance is a useful, if somewhat oversimplified, way to explicate this coordinated action in which various members of the community work (dance) together to perform social practices that enact particular (learning) identities—ways of being in the world—for each of the participants (dancers). Consider the context of formal school learning. The teaching-learning interaction can be likened to an intricate dance to which both students and teachers contribute. This dance is mediated by the curriculum, school policies, the culture of the school and classroom (and schooling more generally), the language of students and teachers, and so on. Take, for example, the classic *initiation* (teacher asks a question)—*response* (student answers the question)—*evaluation* (teacher evaluates the correctness of the student’s response) (IRE) paradigm. In the context of the IRE paradigm learning has a very particular meaning but for students to succeed—or fail—requires the coordination of a particular set of moves by the teacher and students. It also takes someone with the authority to adjudicate the student’s response, in this case the teacher. Importantly, different moves by the teacher (or the student) will alter the shape of the dance, potentially transforming students’ and teachers’ identities. If, instead of responding to an unexpected response by a student with an evaluation (the “E” in I-R-E), the teacher asks, “what do you mean?” she (potentially) transforms her relationship with her student and, ultimately, the meaning of learning. In either case, what students and teachers learn is a socializing routine, a particular set of moves or practices that “count” as learning in the classroom. They also learn an identity. In the case of I-R-E, students are positioned as “getters of the answer the teacher has in mind.” In the alternative example where the teacher responds to a “wrong” answer with “what do you mean?”, students are re-positioned as thinkers or explainers.

The meaning of coordinated actions—in this case school learning—also relies on just the right people doing just the right moves in a particular time and place. The I-R-E paradigm, for example, would signal a very different meaning performed outside the context of a classroom (in say, a bar) or if the evaluative move was performed by someone who had not been granted to the authority to adjudicate the performance of the other party.

The teaching-learning “dance” in a classroom has a particular character that sets it apart from other kinds of learning. But the concept of learning as coordinated action among particular people performing just the right action at the right time and place applies to all forms of learning, from learning to drive a car, learning to cook, learning to read, write, or speak, or learning to make love, to learning to solve a math problem. These are all social/cultural practices that are learned through the process of coordinated actions in contexts where these practices are being performed. From the perspective of a social construction of learning, learning is very much context specific, guided by others, and mediated by particular cultural tools and artifacts. Returning to the dance metaphor, learning the dance depends on people working together to accomplish a particular set of moves that will be recognized as dancing. These moves cannot be accomplished by a person acting on their own or without the aid of music and, perhaps, props (certain dances require specific clothing, spaces, types of music, etc.). Again, learning is never a solitary activity accomplished by autonomous individuals, nor is learning “failure” a solo accomplishment. Things (i.e., neurological activity) do go on in peoples’ heads, but those processes are inextricably linked to social, cultural, and historical factors. Learning is situated in relations among

people in coordinated activity in, with, and emerging from the socially and culturally constructed world (Lave and Wenger, 1991). This is what is learned, how it is learned, and where learning occurs.

If social/cultural practices are what, how and where people learn, it follows that learning resides in the practice and not just in people’s head; it is distributed across cultural practices in the actions of people as they participate in (cultural) activities, in the cultural tools (like language) that have been created as part of the practices, the social setting, and so on. Returning again to the dance metaphor, the “memory” of the dance simultaneously inhabits the dancers, the music, the venue created for the performance of the dance, even the dancers’ shoes and clothing. In other words, the memory of the dance is *in* the dance. Learning to participate in the school practice known as Initiation-Response-Evaluation—and any other school practice—resides in teachers, students, curriculum and testing materials, the physical organization of the classroom (teacher standing in front of students seated in rows) and, most of all, in the language teachers and students draw on to enact this practice. Ultimately, learning does not belong to individual persons, but to the various cultural activities or practices of which they are a part (McDermott, 1993). In this way, learning is distributed across people as they participate in culturally relevant activities (Robbins, 2005).

A social/cultural constructivist perspective on learning “problems”

Considerable attention has been given to the seemingly high proportion of school children who fail to achieve at acceptable levels (i.e., above average). These students are referred to by a range of formal and informal designations such as at-risk, struggling learners, learning disabled, intellectually challenged or just “below average.” Whatever the label, the popular assumption is that academic failure resides in the minds of individual students and, therefore, remediation focuses on “fixing” students’ deficiencies. An alternative discourse focuses on the role that the structures of modern schooling play in producing high levels of school failure. James Gee (2004), for instance, asks: what is there about schooling that transforms so many children who have repeatedly demonstrated that they are good at learning outside school into children who are not good at learning within the walls of schools? Cambourne (1995) puts it this way: “how could a brain which could master such complex learning in the world outside school be considered deficient with respect to the kinds of learning that were supposed to occur inside school?” (p. 183). McDermott et al. (2006) offer the following as an answer. “The American classroom is well organized for the production and display of failure, one child at a time, if possible, but group by group if necessary” (p. 13). Varenne and McDermott (1999) discuss how this process works:

A half-century of ethnographic studies has shown that American education is compulsively competitive. In American classrooms, every child not only has to learn, but to learn better or faster than his or her neighbors. As such, American education is well organized to make hierarchy out of any differences that can be claimed, however falsely, as natural, inherent, and potentially consequential in school. (p. 4)

The proof that some students have learned “better and faster” is the presence of students who fail. In other words, the success of some is determined in relation to the failure of others. Indeed, there is a sense in which school failure is inevitable, a *natural* outcome of the *normal* distribution of human behavior including academic achievement. Indeed, the elimination of school failure may be unacceptable to those who rely on the failure (of some) to affirm their own successes.

A social constructivist perspective challenges the school-failure-as-deficiency position by asserting that the performance of school failure requires the coordination of just the right people doing just the right things in just the right time and place. In the rest of this section, I offer a more detailed discussion of a social constructivist theory of school failure before considering what comes from that perspective.

Traditional psychology defines identity in terms of essential, often immutable traits that are presumed to be the possession of autonomous individuals (Shweder, 1990). It is generally assumed, for example, that people are happy or sad, outgoing or shy, smart or dull, engaged or distractible, calm or excitable and so on. It is never quite so simple as this, however. The happiest person is sometimes sad and even the dullest person can, occasionally, have brilliant insights. And, of course, a student labeled with a reading problem, for example, may have extraordinary mathematical abilities. Still, the popular assumption remains that, over time, there are more or less immutable, essential qualities that define who we are.

An alternative perspective rejects essentialist, in-the-head notions of identity in favor of the idea that identities, including learning identities, are always constructed in relation to other people engaged in coordinated activity in particular social and cultural settings. At the most basic level, “identity emerges in discourse through the temporary roles and orientations assumed by participants, such as evaluator, joke teller, or engaged listener” (Bucholtz and Hall, 2005, p. 591). Consider, for example, the quality of “smart,” which is generally considered to be an essential identity trait in Western cultures. The performance of *smartness* requires the production of certain behaviors valued by the culture (e.g., high academic achievement), performed at the right time and place (e.g., schools, certain workplaces), the presence of others authorized to adjudicate *smartness* (e.g., teachers, psychologists, supervisors) and, crucially, the presence of others deemed not to be smart. Ultimately, no one can perform the identities of smart, slow, at-risk or struggling learner on their own. People do bring various traits to the different contexts in which they live but it is the sociocultural context—full of actors, actions, institutional norms and so on—that gives meaning to these traits (Gee, 2000).

Consider the following example, which illustrates how the identity struggling learner emerged in an instructional interaction between a special education teacher, Mrs. Stroh, and Regis, a student for whom school was a struggle. In this excerpt Mrs. Stroh and Regis were discussing a picture he brought to school.

S: What are you doing in the picture?

R: Skating.

S: Will you tell me a sentence about it?

R: I go skating.

S: Where?

R: In the Mill Pond.

S: In the Mill Pond? Are you right in the pond?

R: Yeah. No.

S: No [laughs]. Where are you?

R: Skating.

S: Yes, but you're not in the Mill Pond. You'd be all wet if you were in it. Where are you?

R: Skating, I think. I don't know. Icing, I think.

S: Sometimes we say, “Skating at the Mill Pond.”

R: Yeah.

Through the lens of deficiency, this confused and confusing interaction lends credence to the school's assessment that there is something wrong with Regis, that he has a learning problem. Certainly, that's how Mrs. Stroh saw it. In a separate interview she indicated that this lesson demonstrated Regis's difficulties “expressing ideas.” Regis did seem to be having trouble expressing himself in relation to Mrs. Stroh's prompts, and even made up a word (“icing”) to describe where he was and what he was doing in the picture. However, from Regis's point of view, this miscommunication may have occurred because one conversational partner (the teacher) treated Regis's utterances as nonsense (Gergen, 1990) and Regis seemed more than a little mystified by Mrs. Stroh's interrogations. By treating Regis's language as problematic, Mrs. Stroh limited his possibilities, highlighting what it appeared he could not do and limiting opportunities to show what he could do. Paraphrasing Ladson-Billings (1994), “when students are treated as [in]competent they are likely to demonstrate [in]competence” (p. 123). This was, arguably, the case with Regis.

It is not that there was not something *different* about Regis, but these differences were not the source of the problems he experienced at school. The problem did not consist in his having learning difficulties—being deficient in verbal skills—as much as in *his living in* “a world well organized to label and disable him” (McDermott and Varenne, 1999, p. 42). As McDermott and Varenne (1999) argue,

The problems many people have in the American School stem only incidentally from what they can or cannot do and much more radically from the way they are treated by others in relation to the designation, assignment, and distribution of more or less temporary or partial difficulties interpreted as success or failure and responded to in the terms of the testing world. (pp. 134–135)

Regis's “struggle” was located in the interaction between him and

features of learning environment[s], including text genres, patterns of talk, and modes of literate expression. Struggle is not a characteristic of individual learners; a struggling reader or writer is one who is experiencing a mismatch between their preferred literate mode and the one they are being asked to communicate with in school.

Collins and Ferri (2016, p. 11).

To be clear Mrs. Stroh is not the villain here. Her interactions with Regis constructed her as a particular kind of person, too. As a primary teacher she has been deputized to police children's behavior, always on the lookout for deficiencies to be remediated. She invoked a familiar instructional script in response to Regis's perceived struggles with verbal expression made salient by his identity as a low achiever. Together Regis and Mrs. Stroh engaged in a perverse, improvisational dance mediated by institutional scripts of schooling and, in this way, they performed particular identities that made them into particular kinds of people, in Regis's case, a student with a learning problem.

But this performance goes beyond just Regis and Mrs. Stroh. Lots of people participated in the choreography of their interactions. The work of schooling is “embedded in the concerted activities of millions of people engaging a surveillance system of professionals—doctors, psychologists, lawyers, educators, and parents are all involved and at the ready before the children show up” (McDermott et al., 2006, p. 6), looking for and producing evidence of school failure. If it takes a village to raise children like Regis, it takes a culture to disable them.

Social constructivism offers a theoretical rationale for rejecting deficit views of school failure. But this is not merely a theoretical matter. Like deficit models of school failure, a social constructivist lens has implications for how schools are organized and how teachers teach. The following section takes up the implications of a social constructivist position on both instruction and how schools might be organized to accommodate the incredible diversity that exists among the student populations.

Implications: a presumption of competence

When it comes to education, theory matters, affecting both how teachers perceive their students and how students perceive themselves. The presumption that poor school achievement is an individual failure residing in the minds and bodies of failing students, for example, leads to practices designed to remediate or “fix” students. Paraphrasing Rist (1970), there is a greater tragedy than being labeled a poor learner and that is being treated like one. From a deficit perspective, being treated like a poor learner typically translates into a reductive, low-level curricula that creates failure by limiting students’ affordances for learning (Annamma et al., 2013; Horn, 2007; White, 2020; Wong, 2010). Students learn less because they are taught less, reinforcing a cycle of failure, a self-fulfilling prophecy that reinforces the initial claim that it is the student who is the problem.

As has been argued above, a social constructivist perspective rejects the position that school failure resides solely in the minds and bodies of students. The student is not the problem. Rather, the problem is the problem (see Collins and Ferri, 2016). When students produce perplexing responses, as in the example of Regis and Mrs. Stroh, the question is not, “what’s wrong with this student?” but “what’s going on here?” How is it that a child who is good at learning in other contexts is not good at learning inside the walls of schooling? (Gee, 2004). Teachers working from a social constructivist perspective begin with a presumption of competence (Biklen and Burke, 2006), “that each person, regardless of age, gender, economic circumstance, or geographic location, is constantly in the process of constructing meanings based on her or his own life experiences” (Miller, 1993, p. 56).

Two instructional principles follow from a social constructivist stance on learning failure: All learners are smart (Miller, 1993) and students learn what they are taught. If the instructional focus for low achievers is on low-level skills, that is what they will learn. They will not learn the rich, engaging practices afforded high-achieving students—unless they are taught them. Therefore, from a social constructivist perspective, what low-achieving students need most is to be afforded the same opportunities and quality of instruction—what has been referred to as high-expectation curricula (Dudley-Marling and Michaels, 2012)—as that afforded the most academically successful students. Recalling the dance metaphor used earlier in this article, the emphasis is on altering the shape of the dance—the interaction between teachers, students, the curriculum and so on—and on not merely “fixing” the dancers. As Gutiérrez (1994) has shown, even a small change in the pattern of interactions between students and teachers can have a significant impact.

A particularly striking example of high-expectation curricula with low-achieving students comes from Dudley-Marling and Michaels (2012) who describe the implementation of a high-expectation literacy curricula, Shared Inquiry, in a high-poverty, low-achieving school in New York City. Students in Lexington Elementary engaged in routine, lengthy (53 min on average), evidence-based discussions of challenging texts. Over the course of a school year, students learned to work together to make claims, cite textual evidence and, if needed, make explicit the evidentiary warrant for their claims in response to open-ended questions posed by their teachers. This practice stood in stark contrast to the district mandated literacy curriculum that focused almost exclusively on low-level skills like phonemic awareness and sound blending. After a single academic year, Lexington Elementary students showed dramatic gains on district-mandated achievement tests. More specifically, Shared Inquiry discussions created affordances that enabled students in a low-achieving, high-poverty school to display a high level of academic competence as students constructed sophisticated academic arguments while discussing challenging texts (Dudley-Marling and Michaels, 2012), contradicting the widely-held belief that cognitively demanding modes of instruction are inappropriate for low-achievers (Horn, 2007).

Other examples of high-expectation curricula with low-achieving students include Project Challenge (Chapin and O’Connor, 2012) and the Optimum Learning Environment (OLE) Project (Ruiz et al., 1996). Project Challenge is a demanding, inquiry-based mathematics curriculum emphasizing mathematical reasoning skills and implemented in low-achieving, high-poverty schools. The OLE Project, created curricula for special education students in low-achieving, bilingual schools, is informed by progressive pedagogical principles that underpin gifted education including student choice, active student engagement in learning, meaningful and authentic purposes for learning, and learning in social contexts. Both the OLE Project and Project Challenge report significant improvement in student achievement on standardized assessments (Chapin and O’Connor, 2012; Ruiz et al., 1996).

There is also a substantial body of literature supporting the efficacy of heterogeneous, de-tracked classrooms in which low-achieving students are educated alongside their higher achieving peers (Oakes, 1985/2005; Thompson, 2001). The Preuss School, for example, is a single track, college-preparatory public charter school that provides a rigorous curriculum to low-income students, with high potential but undeveloped skills, selected through lottery (Mehan, 2012). Preuss classes emphasize project-based learning and a portfolio of assessments fulfilling or exceeding the entry requirements of the University of California and California State University systems. An average of 84% of the graduates of the first classes at Preuss (2004–2009) attended four-year colleges. Preuss was also among the highest ranked high schools in the district in 2009, 2010, and 2011, outpacing many schools with a much lower proportion of students living in poverty.

Arguably, the presumption of competence that underpins de-tracking and high-expectation curricula like OLE and Project Challenge transform student learning identities—and enhance student achievement—by altering the shape of the *pedagogical dance* that makes teachers and students into particular kinds of people (see Dudley-Marling and Michaels, 2012 for other examples of high-expectation curricula). Such practices also contribute to student success by transfiguring students’ beliefs about themselves and their identities as learners (Brodeur, 2021).

Re-constructing learning problems will require more than the simple abnegation of deficit thinking in favor of challenging, transformative curricula. Low-achieving students, like all learners, will require some measure of explicit, individualized and intensive instruction (Dudley-Marling and Paugh, 2009). Low achievers, even those who have acquired formal labels like “learning disabled,” do not, however, require unique, narrowly defined, evidence-based practices based on a mythical average student (Collins and Ferri,

2016). As Collins and Ferri (2016) put it, “If someone is having difficulty reading a particular text, finding what works in a large, randomly assigned clinical trial matters less than finding what works for that person” (p. 9). Ultimately, schools must create inclusive spaces that are congenial to the range of student learning and cultural identities they bring with them to school (see Dudley-Marling and Burns, 2014 for a comprehensive discussion of inclusive classrooms).

Conclusion

Deficit thinking is a pernicious practice that situates the blame for low academic achievement solely in the minds and bodies of individual students. Believing that the problem of school failure is an individual accomplishment, the solution for learning difficulties has tended to focus on “fixing” students through curricular practices that limit students’ affordances for learning, thereby exacerbating the problem. To the degree that deficit thinking does consider broader influences, the tendency has been to implicate the language and culture of families and communities as the root causes of school failure, especially among the poor (e.g., Hart and Risley, 1995; Payne, 2005). This deficit stance simultaneously blames poor students for school failure while deflecting attention from systemic explanations for disproportionately high levels of failure in high poverty schools, places overpopulated by students of color (NAEP, 2019). Like the case for individual students outlined above, it is not the deficit perspective per se that harms poor children as much as what comes from this stance. And there is an abundance of evidence that students living in poverty are much more likely to experience a circumscribed curriculum—what Haberman (1991) referred to as a “pedagogy of poverty”—that severely limits their educational and vocational attainments (Anyon, 1980; Kozol, 2005), reinforcing racist and classist beliefs about the poor and children of color. Arguably, such disparities in curricula account for the widespread finding that, as they pass through the grades, “children living in high-poverty areas tend to fall further behind, regardless of their initial ... skill level” (Snow et al., 1998, p. 98).

A social constructivist lens exposes the systemic factors that create so much failure in schools, especially among students already burdened by discrimination and poverty. It begins with the presumption that all students are competent and a skepticism toward claims that individual or groups of children lack the innate abilities to succeed in school.

References

- Annamma, S.A., Boelé, A.L., Moore, B.A., Klingner, J., 2013. Challenging the ideology of normal in schools. *Int. J. Incl. Educ.* 17 (12), 1278–1294.
- Anyon, J., 1980. Social class and the hidden curriculum of work. *J. Educ.* 162, 67–92.
- Bakhtin, M.M., 1981. Trans. Caryl Emerson and Michael Holquist. In: Holquist, M. (Ed.), *The Dialogic Imagination: Four Essays*. University of Texas Press, Austin and London.
- Biklen, D., Burke, J., 2006. Presuming competence. *Equity Excell. Educ.* 39 (2), 166–175.
- Brodeur, K., 2021. “A type of student they actually want”: narratives from and about a sixth-grade student with a specific learning disability. *Read. Writ. Q.* 37 (1), 45–64.
- Bucholtz, M., Hall, K., 2005. Identity and interaction: a sociocultural linguistic approach. *Discourse Stud.* 7 (4–5), 585–614.
- Cambourne, B., 1995. Toward an educationally relevant theory of literacy learning: twenty years of inquiry. *Read. Teach.* 49, 181–190.
- Chapin, S., O’Connor, C., 2012. Project challenge: using challenging curriculum and mathematical discourse to help all students learn. In: Dudley-Marling, C., Michaels, S. (Eds.), *High-expectation Curricula: Helping All Students Succeed With Powerful Learning*. Teachers College Press, New York, NY, pp. 113–127.
- Chowdhury, M., 2018. Top 10 Benefits of Education Include Poverty Reduction. *Borgen Magazine*. <https://www.borgenmagazine.com/top-10-benefits-of-education/>. (Accessed 24 May 2021).
- Chubb, J.E., Moe, T.M., 1990. *Politics, Markets, and America’s Schools*. The Brookings Institution, Washington, DC.
- Cole, M., 1996. *Cultural Psychology: A Once and Future Discipline*. Belknap Press, Cambridge, MA.
- Collins, K., Ferri, B., 2016. Literacy education and disability studies: reenvisioning struggling students. *J. Adolesc. Adult Lit.* 60 (1), 7–12.
- Dudley-Marling, C., Paugh, P., 2009. *A Classroom Teacher’s Guide to Struggling Readers*. Heinemann, Portsmouth, NH.
- Dudley-Marling, C., Gurn, A., 2010. Troubling the foundations of special education: examining the myth of the normal curve. In: Dudley-Marling, C., Gurn, A. (Eds.), *The Myth of the Normal Curve*. Peter Lang, New York, pp. 9–23.
- Dudley-Marling, C., Michaels, S., 2012. Shared inquiry: making students smart. In: Dudley-Marling, C., Michaels, S. (Eds.), *The High-Expectation Curriculum: Helping All Students Succeed With Powerful Learning*. Teachers College Press, New York, pp. 99–110.
- Dudley-Marling, C., Burns, M.B., 2014. Two perspectives on inclusion in the United States. *Glob. Educ. Rev.* 1 (1), 14–31. <http://ger.mercy.edu/index.php/ger/article/view/10>.
- Friedman, M., 1955. The role of government in education. In: Solo, R.A. (Ed.), *Economics and the Public Interest*. Rutgers University Press, New Brunswick, NJ.
- Gee, J.P., 2000. Identity as an analytic lens for research in education. In: Secada, W.G. (Ed.), *Review of Research in Education*, vol. 25. American Educational Research Association, Washington, DC, pp. 99–125.
- Gee, J.P., 2004. *Situated Language and Learning: A Critique of Traditional Schooling*. Routledge, New York.
- Gee, J.P., 2012. *Social linguistics and literacies: Ideology in discourses*, 4th ed. Routledge, New York, NY.
- Gergen, K.J., 1990. Social understanding and the inscription of self. In: Sigler, J.W., Shweder, R.A., Herdt, G. (Eds.), *Cultural Psychology: Essays on Comparative Human Development*. Cambridge University Press, New York, NY, pp. 569–606.
- Gutiérrez, K.D., 1994. How talk, context, and script shape contexts for learning: a cross-case comparison of journal sharing. *Linguist. Educ.* 5, 335–365.
- Haberman, M., 1991. The pedagogy of poverty versus good teaching. *Phi Delta Kappan* 73, 290–294.
- Harlow, C.W., 2003. *Education and Correctional Populations*. Bureau of Justice Statistics Special Report. <https://www.bjs.gov/content/pub/pdf/ecp.pdf>. (Accessed 24 May 2021).
- Hart, B., Risley, T.R., 1995. *Meaningful Differences in the Everyday Experiences of Young American Children*. Brookes, Baltimore, MD.
- Horn, I.S., 2007. Fast kids, slow kids, lazy kids: framing the mismatch problem in mathematics teachers’ conversations. *J. Learn. Sci.* 16 (1), 37–79.
- Hursh, D.W., 2007. Marketing education: the rise of standardized testing, accountability, competition, and markets in public education. In: Ross, E.W., Gibson, R. (Eds.), *Neoliberalism and Education Reform*. Hampton Press, Cresskill, NJ, pp. 15–34.
- Kozol, J., 2005. *The Shame of the Nation: The Restoration of Apartheid Schooling in America*. Crown, New York.
- Ladson-Billings, G., 1994. *The Dreamkeepers: Successful Teachers of African American Children*. Jossey-Bass, San Francisco, CA.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, New York.

- Marks, G.N., 2006. Influences on, and the consequences of, low achievement. *Aust. Educ. Res.* 33 (1), 95–115.
- McDermott, R., Varenne, H., 1999. Adam, Adam, Adam and Adam: the cultural construction of a learning disability. In: Varenne, H., McDermott, R. (Eds.), *Successful Failure: The School America Builds*. Westview, Boulder, CO, pp. 25–44.
- McDermott, R., Goldman, S., Varenne, H., 2006. The cultural work of learning disabilities. *Educ. Res.* 35 (6), 12–17.
- McDermott, R.P., 1993. The acquisition of a child by a learning disability. In: Chaiklin, C., Lave, J. (Eds.), *Understanding Practice: Perspectives on Activity and Context*. Cambridge University Press, New York, pp. 269–305.
- Mehan, H., 2012. Detracking: re-forming schools to provide students with equitable Access to college and career. In: Dudley-Marling, C., Michaels, S. (Eds.), *The High-Expectation Curriculum: Helping All Students Succeed With Powerful Learning*. Teachers College Press, New York, pp. 15–27.
- Miller, L., 1993. *What We Call Smart: A New Narrative for Intelligence and Learning*. Singular, San Diego, CA.
- National Assessment of Educational Progress (NAEP), 2019. *The Nation's Report Card*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, Washington, DC.
- NCTQ, 2021. National Council on Teacher Quality. <https://www.nctq.org>.
- No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319, 2002.
- Oakes, J., 1985/2005. *Keeping Track: How Schools Structure Inequality*, second ed. Yale University Press, New Haven, CT.
- OECD, 2013. *What are the Social Benefits of Education?*. In: *Education Indicators in Focus* OECD Publishing, Paris. <https://doi.org/10.1787/5k4ddxn139vk-en>. No. 10.
- Payne, R.K., 2005. *A Framework for Understanding Poverty*, fourth ed. Aha! Process, Inc, Highlands, TX.
- Rist, R.C., 1970. Student social class and teacher expectations: the self-fulfilling prophecy in ghetto education. *Harv. Educ. Rev.* 40, 72–73.
- Robbins, J., 2005. Contexts, Collaboration, and Cultural Tools: a sociocultural perspective on researching children's thinking. *Contemp. Issues Early Child.* 6 (2), 140–149.
- Ruiz, N.T., García, E., Figueroa, R.A., 1996. *The OLE Curriculum Guide*. California Department of Education, Sacramento, CA.
- Shweder, R.A., 1990. Cultural psychology—what is it? In: Stigler, J.W., Shweder, R.A., Herdt, G. (Eds.), *Cultural Psychology: Essays on Human Development*. Cambridge University Press, New York, NY, pp. 1–46.
- Snow, C.E., Burns, M.S., Griffin, P., 1998. *Preventing Reading Difficulties in Young Children*. National Research Council, Washington, D.C.
- Thompson, S., 2001. The authentic standards movement and its evil twin. *Phi Delta Kappan* 82 (5), 358–362.
- Varenne, H., McDermott, R., 1999. Introduction. In: Varenne, H., McDermott, R. (Eds.), *Successful Failure: The School America Builds*. Westview, New York, NY.
- White, K.L., 2020. The social construction of “at-risk” status: the intersection of policy, practice, and the white gaze in a first grade. *Classr. Perspect. Provocat.* 9, 1–37.
- Wong, J., 2010. Exploring the construction/deconstruction of learning disabilities in an urban school: revisiting Sleeter's essay. *Disabil. Stud. Q.* 30 (2). Retrieved April 27, 2021 from <http://dsq-sds.org/article/view/1242/1286>.

Learning in the disciplines: a conceptual framework

Susan R. Goldman, Learning Sciences Research Institute, University of Illinois Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	305
Assumptions underlying the conceptual framework	305
Development of the conceptual framework for disciplinary literacies	306
Research on expertise	306
Core constructs for learning in the disciplines	307
Core constructs: a conceptual tool for emergent disciplines	311
Value-added of the conceptual framework for learning in the disciplines for designing adaptive instruction in the 21st century	312
Conclusion	312
References	313

Introduction

Learning in the Disciplines reflects a convergence of research on how people learn and develop and the competencies required of an educated citizenry in the 21st century. Knowing what is no longer sufficient for effective functioning in a society in which information must be vetted and evaluated rather than merely being accepted as fact. People need to understand how knowledge is generated; how it is represented and communicated; and how to process, interpret, and evaluate new information in light of their existing knowledge and belief systems. Doing so allows people to take a critical stance toward information. These 21st century competencies align well with what is referred to as disciplinary literacy - the reading, reasoning, and argumentation practices of specialists in disciplinary content areas (e.g., [Goldman, 2012, 2015](#); [Moje, 2008, 2015](#)). However, surveys of industry employers, college instructors, and healthcare providers, as well as national and international assessments, unfortunately indicate that most citizens lack these 21st century competencies ([American College Testing, 2006](#); [Ananiadou and Claro, 2009](#); [National Alliance of Business, 2002](#); [National Assessment of Educational Progress, 2009, 2022](#); [National Center for Education Statistics, 2013, 2022](#); [National Research Council, 2012](#); [Organization for Economic Co-operation and Development, 2013](#)). These findings should not be surprising given that typical subject matter learning involves students reading to learn content rather than engaging in practices that experts in the field used to generate the content (e.g., [Goldman et al., 2016](#); [Greenleaf et al., 2001](#); [Greenleaf and Valencia, 2017](#); [Rainey and Moje, 2012](#); [Vaughn et al., 2013](#)). The result is that students do not have opportunities to engage in the literacy practices appropriate to specific disciplines or be exposed to considerations of how similar and different particular practices are across different disciplines. Nor do they have experiences learning from one another to collaboratively solve problems and innovate on traditional ways of thinking and approaching problems.

The topic of this chapter, Learning in the Disciplines, reflects an emphasis among those concerned with defining the goals of schooling to better equip individuals with the competencies 21st century society demands. This includes not only redefining what learners learn but how they come to learn it. In this chapter, a conceptual framework is presented that has proven valuable for organizing the varied kinds of disciplinary knowledge and practices that constitute what disciplinary experts know in their domain of expertise. The framework was the result of a large collaborative effort over several years to characterize the kinds of knowledge disciplinary experts bring to bear when they reason, argue, and generate new knowledge within their disciplinary communities ([Goldman et al., 2016](#)).

Assumptions underlying the conceptual framework

Two foundational assumptions underlie the framework. The first is that learning is situated in interactional spaces in which the individual and community are changing or evolving together through their joint participation ([Rogoff, 1997](#)). Knowledge is generated or constructed through interactions with others, material resources, and tools (e.g., [Vygotsky, 1978](#)). Learning happens through and with others. Learners take up practices of those they come into contact with through observation or participation in those practices ([Rogoff, 1997](#)). These interactions and observations become the basis of internalized representations and include cognitive (memory, perceptual, reasoning processes), as well as social and affective dimensions. In turn, what has become internalized shapes how learners participate in, experience, and observe subsequent interactions: how and what learners think and feel arises from complex interactions that reflect learners' cultural and contextual circumstances. Learning then is defined, in part, as the transformation of an individual's ability to participate in valued social and cultural activities. The processes by which an individual engages in practice involve emotional, motivational, and relational aspects of self, not just knowing ([Holland and Lave, 2009](#)).

Consistent with the first assumption about the nature of learning is the second: Disciplines can be conceptualized as communities of practice (Lave and Wenger, 1991) wherein the members negotiate the norms, conventions, and criteria for proposing, arguing for, establishing, and evaluating knowledge claims and the arguments put forth for them. Criteria are established for what constitutes valid and reliable inquiry practices, including patterns of logic and reasoning for connecting evidence to the claims it is intended to support (Toulmin et al., 1984). Evidence that does not meet these criteria compromises whatever claims are being made on the basis of that evidence. Members of a disciplinary community also share common epistemic commitments to the aims, goals, and purposes of argument and knowledge generation within their discipline (e.g., explanation, evaluating alternatives, proposing policies) and the representational forms used to communicate their ideas with one another (e.g., Bazerman, 1985; Shanahan et al., 2011). However, what is shared among members of a community may not be known by those outside the community. Becoming a member of a disciplinary community involves learning these norms and conventions in addition to learning specific content. Indeed, “outsiders” to the disciplinary community who do not have access to disciplinary norms and conventions often read and attempt to make sense of disciplinary information differently than intended by the disciplinary community, resulting in different interpretations (Bazerman, 2004; Goldman and Bisanz, 2002; Latour and Woolgar, 1986; Wineburg, 1991). The “insider—outsider” dichotomy is in part related to the rise in misinformation and disinformation occurring in present day society. The insider-outsider issue is one of the challenges adolescents face when they are tasked with reading specialized disciplinary texts in content area classes in order to perform discipline specific tasks (Goldman and Bisanz, 2002; Moje and O’Brien, 2001). However, instruction that focuses on the specialized ways of reading, thinking, and conveying information within a discipline is relatively rare (Alvermann and Moore, 1991). Rather, equipped with only generic reading strategies, learners often encounter unfamiliar and technical vocabulary, complex syntax, and a variety of genres within and across disciplines (Lee and Spratley, 2010). Thus, in addition to building knowledge in different content areas, they are grappling with disciplinary literacies—the oral and written reasoning, argumentation, and communication practices of the disciplines (Moje, 2008; Carnegie Council on Advancing Adolescent Literacy, 2011). In developing a conceptual framework for learning in the disciplines, our team endeavored to capture the various kinds of knowledge that disciplinary “insiders” bring to their disciplinary work, including knowledge of how disciplinary information is represented and communicated as important to the construction and generation of disciplinary knowledge. All of these define what it means to be literate and learn in a discipline.

Development of the conceptual framework for disciplinary literacies

The initial phase of building the conceptual framework for Learning in the Disciplines focused on identifying within three disciplines—science, history, and literary reading—the various kinds of knowledge that members of a disciplinary community negotiate and share in their efforts to generate new knowledge. As reported in Goldman et al. (2016), the team conducted a conceptual metanalysis of theory and research on experts in each of these three disciplinary areas (science, history, and literary reading), identified differences between experts and those engaged in learning these disciplines, and reviewed efforts that had been made to create learning environments and interventions intended to support learners in becoming more expert in these disciplines. In a second phase, we looked across the findings from the conceptual metanalyses for each discipline and asked how the disciplines were similar and how they were different. This process enabled us to generate a higher order characterization of the kinds of knowledge that we labeled Core Constructs. Disciplinary learning goals are developed by instantiating the core constructs for specific disciplines. Thus the Core Constructs constitute a conceptual framework for the disciplinary literacies involved in learning in the disciplines.

In the present context, we first summarize the major findings of the conceptual metanalyses focusing especially on characteristics of domain experts. While there are similarities in the general characteristics of experts and expertise across disciplines, instantiating how these are manifest in specific disciplinary contexts is necessary to determine what it takes to support learners in engaging in learning a discipline.

Research on expertise

Cognitive research on the nature of expertise indicates that expertise is domain specific. In their domains of expertise, experts have extensive but highly organized knowledge and strategies that are not present in those new to the discipline (See for discussion Chi et al. (1988) and Goldman et al. (1999)). It is well-established that learning is an active process where the knowledge, skills and dispositions that a learner brings to a learning task guide how they perceive and engage with new information and new tasks (See Bransford et al. (2002) for elaboration). Given the differences in the knowledge bases of domain experts and non-experts, it is not surprising that they approach domain tasks differently. The knowledge bases of experts enable them to see meaningful patterns in data, store and effectively retrieve information in response to task demands and develop more efficient problem solving strategies as compared to novices in the domain (See Goldman et al. (1999) for review). In one classic study, Sabers et al. (1991) compared expert and novice teachers’ responses to a classroom video. There were differences between them in what they noticed and how they interpreted what they did notice. The novice teachers provided simple descriptions of isolated events and made few if any

inferences about the relationship between the instruction and the students' behavior. The experts, on the other hand, paid attention to a great range of activities and their interrelationships and to what the verbalizations implied for interpreting the events shown on the video. Differences in professional perception lead to differences in problem definitions and approaches to their solutions (Seidel and Stürmer, 2014; Wolff et al., 2016).

Domain experts also display metacognitive knowledge of their learning processes, particularly with respect to detecting inconsistencies. They can provide explanations for the strategies they employ in their problem-solving efforts. For example, in one study, history experts indicated that they began reading historical documents by first ascertaining sourcing information, including when the document was produced, by whom and for what purpose. This information was critical to how they related the information to their prior knowledge and interpreted it (Wineburg, 1991, 1994). Students jumped right into reading the artifacts provided to them. In other studies, students have been found to simply summarize the content of documents rather than submitting them to critical examination or evaluating the trustworthiness and potential biases of a document (Greene, 1994). Similar differences have been observed in the sciences. Across a variety of science fields, scientists evaluate information against established criteria for reliable and valid data (e.g., Chinn and Brewer, 1996; Dunbar, 1995; Koslowski, 1996; Schauble et al., 1995). Experts' knowledge representations have also been shown to reflect deeper, principled understandings of the information in comparison to those of novices. This enables them to discern meaningful patterns in situations as diverse as chess board arrays (Chase and Simon, 1973) and physics representations (Chi et al., 1981) for chess and physics experts respectively.

Some might argue that domain experts are just smart people and smarter than the non-experts to whom they have been compared in the research studies. This is not the case. Outside of their domain of expertise, they display noticing, problem solving, and metacognitive processes different from those they display within their area of expertise. For example, experts reading outside their fields display different approaches than when reading within their area of expertise (Bazerman, 2004). In other words, they perform like non-experts in domains beyond their own.

From a situated, sociocultural perspective, experts are members of specific communities of practice that have evolved and share physical, social, and cognitive resources that constitute the cultural context of the community (Danish and Gresalfi, 2018; Greeno and Engstrom, 2014). Experts are very good at organizing and structuring their time and space to capitalize on the cultural resources and tools at their disposal as members of their communities of practice (Hutchins, 1995; Kirsh, 1995). It is also the case that no two situations of knowledge application are ever identical. Experts are often adept at adapting what they know to the particularities of the performance situation. This is especially the case for expert teachers (Markauskaite and Goodyear, 2017).

Core constructs for learning in the disciplines

What emerged from the analyses of expertise in each of the three disciplines were discipline-specific clusters of the kinds of knowledge that experts relied on when reading, reasoning, solving problems, constructing and conveying evidence-based arguments in their fields. We organized the clusters for each discipline into five higher-order categories of what we called disciplinary core constructs:

- Epistemology: Beliefs about the nature of knowledge and the nature of knowing. What counts as knowledge? How do we know what we know?
- Inquiry Practices, Reasoning Strategies: Ways in which claims and evidence are established, related, and validated
- Overarching concepts, themes, principles, frameworks: Foundational concepts, ideas, reasoning principles, and assumptions. These serve as a basis for warranting, justifying, legitimizing connections between evidence and claims.
- Forms of information representation/types of texts: Types of texts and media (e.g., traditional print, oral, video, digital) in which information is represented and expressed.
- Discourse and language structures: The oral and written language forms in which information is expressed.

Over the course of our work with these constructs, we came to see epistemology as central, providing purpose and motivation to the ways in which inquiry was conducted; the reasoning principles that were invoked; and the forms in which information was represented, expressed, examined, and critiqued, and negotiated in and through oral and written discourse and language structures. Differences in the norms and conventions of different disciplinary communities arise in part from epistemological commitments about the nature and purposes of the knowledge generated, that is what constitutes the disciplinary knowledge base. Furthermore, the epistemic aims, values, and goals of a discipline have implications for inquiry questions and the practices associated with knowledge generation, reasoning, and representation (Andriessen, 2006; Braaten and Windschitl, 2011; Chinn et al., 2011; Gee, 1992; Hillocks, 2011; Megill, 1989; Osborne, 2002; Pearson et al., 2010; Wineburg, 2001). In literature for example, a central epistemic aim of literary interpretation is exploration of the human condition through multiple perspectives, including philosophical, moral, social, political, economic, and cultural contexts of the human experience. Literature invites dialogic interaction among readers; it is open to multiple interpretations even as readers seek to understand authorial intent. Accordingly, literary epistemology values attention to both content and form as critical windows into authorial messages. Indeed, much of literary criticism is concerned with the

relationship between meaning and intentional decisions that authors make about character types, plot structures, and specific language choices as cues to authorial intent (e.g., [Rabinowitz, 1987](#)). For example, patterns of word or phrase repetitions or the occurrence of the unexpected are features of literary works that suggest possible figurative uses of language and the need to go beyond the literal meaning to entertain symbolic or metaphoric interpretations (For discussion see [Lee et al., 2016](#)).

On the other hand, in science, the epistemic aims and associated inquiry questions are often about explaining causes of phenomena in the natural and designed worlds ([Rutherford and Ahlgren, 1990](#)). Scientists accrue evidence for and against potential explanations of science phenomena, putting forward databased arguments for peer critique and justifying the viability of their explanatory models by appealing to existing bodies of scientific principles ([Berland and Reiser, 2009](#); [Ford and Wargo, 2012](#)). The nature of scientific knowledge is tentative and subject to revision in the face of new empirical evidence that may reflect changes and advances in technology and measurement methods. Science inquiry practices are consistent with this epistemological orientation. New investigations are often generated through careful reading of published data sets and methods and are scrutinized and evaluated for reliability and convergence with other findings (e.g., [Chinn and Malhotra, 2002](#); [Duschl and Osborne, 2002](#)).

In contrast, virtually all historians agree that the epistemic purpose of history is the generation of interpretations of the past based on admittedly incomplete traces, also known as primary sources (e.g., [Megill, 1989](#)). Both the traces of the past and interpretations of them are acknowledged as human constructions that can be questioned, critiqued, and contested ([Lee, 2005](#); [Seixas, 2010](#)). Inquiry into the past therefore uses practices for comparing across primary source traces and contextualizing the information in these sources in what else was going on at the time. These inquiry practices help establish the credibility and reliability of primary sources (e.g., [Leinhardt and Young, 1996](#); [Wineburg, 1991](#)). Inquiries into the past are limited by the incompleteness of the historical record and an important question is whose voice(s) are missing. Historical knowledge is provisional and subject to revision based on new evidence. These characteristics of historical knowledge often result in competing and contested interpretations. Historical inquiry requires reading and reasoning practices that closely attend to metadata related to particular sources as well as points of convergence and divergence across multiple representational levels of meaning in service of generating a coherent interpretive model of the past, a goal that cannot always be reached (e.g., [Britt et al., 2018](#); [Reisman, 2012](#); [Rouet and Britt, 2011](#)).

Differences among disciplines in epistemic aims and goals produce differences in the essential questions of each discipline; in the practices that are followed with respect to gathering, vetting, and selecting information to address the questions; and in the forms and content that constitute products of the disciplinary inquiry ([Goldman, 2018](#)). More specifically, differences in the products of inquiry imply differences in the questioning, data gathering, and reading and reasoning practices needed for problem solving and inquiry in each discipline. The inquiry processes must be situated and attuned to the ways in which information is represented, the goals of disciplines, and the discourse forms of specific disciplines so that comprehension, sense-making and problem-solving focus on what matters and what counts for addressing disciplinary inquiry questions. Across disciplines however, the practices can be described at a general level: careful attention to data—closely reading the data; synthesis and analysis within and across different data sets that were gathered; creation of claim–evidence–reasoning structures; use of interpretive frameworks, principles, and criteria; evaluation of reliability and justification; and appreciation for the epistemology of the discipline. What it looks like to engage in these practices is what varies from discipline to discipline.

To illustrate the nature of differences from discipline to discipline in how the core constructs are instantiated, [Table 1](#) shows the results of the specification of core constructs for literature/literary reading, science, and history. [Goldman et al. \(2016\)](#) provide the theoretical and empirical grounding for these instantiations. Note that one limitation of the effort in science was that we focused primarily on text-based inquiry rather than encompassing expertise in bench science. However, we intend the phrase text to refer to the broad array of forms in which data and models are generated and conveyed in the sciences, including diagrams, schematics, data visualizations (e.g., figures, tables), photographs, and so forth. Expertise in science entails using these various “text” forms to make sense of observational, historical, and experimentally-generated information. The specifics of how the core constructs were manifest within each discipline reflects the differences among the disciplines in their epistemological aims and values; inquiry practices; big ideas, principles, and interpretive outcomes; and how they are represented and communicated.

Clearly, the instantiation of the core constructs is a heuristic for approaching the analysis of a discipline in terms of its aims and goals, inquiry methods, and so forth. Variations on the specifications provided in [Table 1](#) would be expected for any of the various subdisciplines within each of the very large disciplinary categories shown in [Table 1](#). Furthermore, the instantiations in [Table 1](#) reflect a largely Western perspective on these three disciplines, especially regarding the epistemic goals and aims and the inquiry methods. The process of interrogating each of these core constructs is itself a situated learning activity occurring in the cultural contexts of those engaged in the analysis. Thus, alternative instantiations of these three disciplinary areas can be anticipated. The dialogic that results from considering such alternatives can be expected to yield fruitful understandings of the ways in which established norms and conventions of disciplinary communities are themselves time and place bound and should be expected to evolve. The community not only shapes but is shaped by the individual members. As membership in disciplinary communities becomes more diverse, the perspectives they bring should naturally result in the re-negotiation of norms and conventions for generating new disciplinary knowledge and specifications of core constructs. In addition, learners come to the disciplinary canon with ways of making sense in their everyday worlds (e.g., [Moll et al., 1992](#)). These everyday sensemaking practices are resources that can act as pathways into disciplinary learning practices using approaches such as cultural modeling ([Lee, 2007](#)).

Table 1 Highlights of instantiations of core constructs literature, science, and history.

Epistemological Aims and Values		
Literature/Literary Reading	Science: Text-based Inquiry	History
Literature as a window into interrogating the human experience including the social, political economic, and cultural contexts of the world. Literary texts are open to dialogue among communities of readers within and across time. Literary critique values attention to both content and form, and investigates how decisions by authors regarding structure and language use influence meaning.	Explanations of the natural and engineered worlds expressed as models and theories that are • approximations and have limitations. • based on sound empirical data • socially constructed • meet criteria of parsimony, and logical cohesion. • subject to revisions with successive empirical efforts that reflect changes in technology, theories and paradigms, and cultural norms.	Claims about the past are based on a historical record that is inherently incomplete, conflicting, and obscure, and therefore provisional and contestable. Historians make interpretive arguments about the past based on the historical record but from their own perspectives.
Inquiry Practices, Reasoning Strategies		
Literature/Literary Reading	Science: Text-based Inquiry	History
Infer from details within texts • Plot sequence and causal links between events, actions • Characters motivations and psychological states, and relationships among of characters. • Structural generalizations about how inferred meanings are achieved rhetorically • Generalizations the reader argues the author is making about the world beyond the text itself Evaluate how rhetorical strategies are employed to help shape reader's response. To construct interpretations, draw on prior knowledge of • Social milieu of the text • Human intentionality • Moral and philosophical	Scientific knowledge is built by: • Developing coherent, logical explanations, models or arguments from evidence • Advancing and challenging explanations • Converging/corroboration of evidence • Comparing/integrating across sources and representations • Evaluating of sources and evidence in terms of scope, reliability, and extent to which it accounts for evidence.	Historians investigate questions, interrogate the historical record, or analyze historical interpretations through • Sourcing • Contextualization • Corroboration • Questioning inclusiveness (what perspectives, sources, data are included or omitted) • Questioning coherence • Avoiding logical and reasoning fallacies

(Continued)

Table 1 Highlights of instantiations of core constructs literature, science, and history.—cont'd

precepts • Related literary and non literary texts • Other works by the author		
Overarching Concepts, Themes, Principles		
Literature/Literary Reading	Science: Text-based Inquiry	History
Readers construct interpretations of literary texts using belief systems and principles that reflect • Morality and philosophy • Historical contexts • Traditions of critical theory (e.g., Reader response, feminist, New Criticism, Post Structuralism) • Intertextuality, including relations among literary texts, between literary texts as well across literary and historical, sociological, philosophical texts	Scientists connect evidence to claims using • Cross-cutting concepts (patterns; cause and effect; scale, proportion and quantity; systems and system models; energy and matter in systems; structure and function; stability and change of systems). • Disciplinary core ideas in the physical sciences, earth and space sciences; life sciences; and engineering, technology, and applications of science	Historians draw on a variety of interpretive frameworks: • Societal systems (political, social, economic, technological, geographic) • Relational (chronological, contingent, causal, coincidental, chance) • Change over time • Thematic (migratory processes, industrialization, power relations) • Complex systems
Forms of information representation/types of texts		
Literature/Literary Reading	Science: Text-based Inquiry	History
Prototypical genres and ways of structuring plots; kinds of prototypical protagonists. • Generic genres (e.g. stories, poems, dramas) • Specialized poetic genres (e.g. haiku, ballad, sonnet) • Exemplar plot structures (e.g., coming of age, science fiction, fable, satirical works, myth, magical realism) • Exemplar Protagonists (tragic hero, antihero, mythic hero, trickster,)	Scientific texts may have different explanatory purposes (e.g., cause effect, correlation, comparison, process sequence, chronology, enumeration, description). Science texts convey meaning with multiple representations (e.g., verbal,digital, equations, graphs, tables, simulations, flowcharts, schematics, videos)types of sources (e.g., bench notes, refereed journal articles, textbooks, websites, blogs).	Texts types in history are • Primary sources that include all texts from the period of study (e.g., autobiography, memoir, fiction, news story , editorial, political cartoon, art, graphic novel, video, legal documents). • Secondary and tertiary sources (e.g., biography, historical fiction, data tables, textbook, editorial, polemic essay, political map) History texts appear in a range of media, including traditional print, radio, TV, video, and hypermedia.

Table 1 Highlights of instantiations of core constructs literature, science, and history.—cont'd

Discourse and language structures		
Literature/Literary Reading	Science: Text-based Inquiry	History
How author's selection and sequence of action, dialogue and description create an imaginary world into which the reader is invited through the manipulation of language. Major ways of using language are • Imagery: Language used to create a visual representation and invite emotional response (e.g., use of descriptions, metaphor, simile) • Figuration: Language used to invite a figurative interpretation beyond the literal (e.g., symbolism, irony, satire) • Problems of point of view: Who is speaking; reliability; authorial versus narrative audience; relation of narrator's point of view to the author (e.g., omniscient narrator, unreliable narrator, multiple narrators) • Rhetorical strategies and patterns (e.g., parallelism, contrast, repetition, understatement, exaggeration, allusion, dialogue typography, privileged placement)	Science texts contain • distinctive grammatical structures (e.g., nominalizations, passive voice). • technical and specialized expressions. • signals to the degree of certainty, generalizability, and precision of statements. Argumentation is a scientific discourse practice in which evidence is used to support knowledge claims, and scientific principles and methods are used as warrants.	Structure of historical arguments can be • descriptive (<i>What was the case?</i>) • explanatory (<i>Why was it the case?</i>) • narrative (both descriptive and explanatory structures) Linguistic features of history texts mark conventions of chronology, periodization, perspective, theme and topic. History texts include accurate and thorough documentation of sources, usually in the form of footnotes or notes. Conventions for claim and evidence presentation in oral and written forms include • One-sided, two-sided arguments, multi-sided • Two-sided, multi-sided refutational arguments • Implicit arguments (embedded in descriptive and narrative structure) • Oral arguments (debates, discussions, conversations)

Core constructs: a conceptual tool for emergent disciplines

The complexity of the 21st century is giving rise to the emergence of new disciplinary areas, often beginning in multi- or inter-disciplinary collaborations. The Core Constructs Disciplinary Literacies framework can serve as a tool for explicitly attending to the epistemological commitments of constituent disciplines in multi- and inter-disciplinary work and for negotiating the aims and goals of emergent, integrated disciplines. Such an effort would lay the groundwork for considering the inquiry methods, big ideas and principles, representational forms and rhetorical structures that characterize the knowledge base of emergent disciplines. There is a clear need for this in that there is already a robust movement for integrated STEM as an approach to address “conditions that require the application of knowledge and practices from multiple STEM disciplines to learn about or solve transdisciplinary problems” (Nadelson and Seifert, 2017, p. 221). Bioinformatics (biological sciences and informatics/data sciences) and earthquake engineering (seismology and engineering) (Cavlazoglu and Tuessy, 2017) are two examples. Attention to the types of knowledge being brought together in such integrations is critical to understanding what new learning is required on the parts of experts in the “combining” disciplines in order to work together productively.

It is important to attend to these issues because knowledge gaps between the home disciplines and the integrated disciplines can threaten the success of integrated disciplines. Representing the core constructs of the home disciplines can be one way to mitigate unproductive, and facilitate fruitful, interactions and the emergence of new disciplinary communities by explicating the knowledge demands of the integrated discipline in comparison to those of the origin/home disciplines. Specifying the inquiry practices, representational literacies, and epistemic commitments of interdisciplinary areas and integrated disciplines is a necessary first step toward preparing professional learning experiences that enable teachers to teach these disciplines. For example, in the case of bioinformatics, knowledge gaps between biological sciences and bioinformatics include environmental science, public health, statistics, and computer science content along with practices for working with real-world messy data, a type of data not typically covered in high school biology courses (Kelley and Knowles, 2016; Kjølvik and Schultheis, 2019). Yoon et al. (2022) document the challenges they confronted when they focused solely on preparing teachers to implement a bioinformatics curriculum without also providing the biology teachers with opportunities to delve in-depth into the gap areas between biology and bioinformatics.

In sum, integrated disciplines and interdisciplinary work in general reflect new communities of practice in which the members need to explore and negotiate their epistemic commitments, consider compatibilities and incompatibilities in the inquiry methods of the originating fields, and learn each other's representational forms and rhetorical structures. This is especially so for those tasked with interdisciplinary teaching and the teaching of integrated disciplines to high school and college students.

Value-added of the conceptual framework for learning in the disciplines for designing adaptive instruction in the 21st century

Periodically, and especially over the last dozen years, frameworks and standards have emerged in efforts to redefine the goals of schooling to reflect the needs of 21st century society, especially to engage learners in authentic disciplinary practices of knowledge generation. Examples of these include the Common Core State Standards (Council of Chief State School Officers, 2010), Next Generation Science Standards (Next Generation Science Standards Lead States, 2013), the 3C Framework in Social Studies (National Council for the Social Studies, 2013), and PISA (Organization of Economic and Cultural Development, 2013). The substance of such standards and frameworks drew heavily on identifying the important ideas and principles in specific disciplinary domains as well as the practices disciplinary experts engaged in to generate this content. However, in and of themselves standards are inadequate to accomplish the transformation in classroom instruction that is needed to prepare 21st century citizens. This is because to be useful in guiding instruction, standards need to be unpacked and translated into learning processes and outcomes that can provide visible evidence relevant to learning goals (Shephard et al., 2009). Goals and evidence need to be calibrated to reflect developmentally appropriate performances and trajectories of such performances over time. Too frequently, insufficient attention is paid to this unpacking process and standards become the learning goals; instructional activities are keyed to specific standards and the purpose of the activity is to teach the standard, with little attention devoted to how different standards connect or are related to the overarching goals of the discipline. The result is often disconnected and disjointed with students having no greater grasp of the practices of the discipline and how knowledge is generated than when instruction focused on reading and memorizing "known" disciplinary content from a textbook for purposes of reproducing that information on outcome assessments.

The Conceptual Framework for Learning in the Disciplines outlined in this chapter has proven to be a useful tool for engaging educational practitioners in a process of unpacking standards in English Language Arts, Science and History. It provided them with an organizational structure that allowed them to move from teaching to the standards to building instructional units that were motivated by essential questions of the discipline—that addressed the aims and goals of the disciplinary inquiry using practices authentic to the discipline, albeit at developmentally tractable levels (e.g., Goldman et al., 2019; Levine et al., 2019). The instructional designs explicitly introduced students to the purposes of historical, literary, or scientific inquiry by making visible to middle and high school students the essential questions each discipline aimed to address as well as the reading, thinking, and reasoning inquiry practices that could be employed to address the essential questions. Example units can be accessed under publications and case library at www.projectreadi.org. Collaboration and discussion became functional participation structures for engaging in the complexities that arose in the inquiry, complexities that could easily overwhelm an individual learner. As students wrestled with multiple perspectives, conundrums, and inconsistencies they argued to make sense and learned important disciplinary content, principles, and practices in the process. Standards were integrated and incorporated into a context of meaningful, motivating and intellectually challenging disciplinary work. Teachers consistently reported that exploring their discipline using the Core Constructs of the framework gave them opportunities to both access their own inquiry practices and deepen their understanding of the epistemic commitments of the discipline (e.g., Goldman et al., 2022; Goldman and Popp, in press; Hall and Goldman, 2022; Hall et al., 2022; Ko et al., 2022). They were then able to re-purpose classroom teaching and learning to more closely reflect disciplinary purposes and dialogic practices, both critical to supporting the development of citizens who can critically analyze and evaluate information.

Conclusion

Now in the midst of the third decade of the 21st century, the need for citizens to be able to critically analyze and evaluate information is obvious. We no longer live in the world that existed in 1897 when Adolph Ochs adopted the slogan for the New York Times "All the news that's fit to print" as a commitment to impartial news reporting. In contemporary times, the increasing

dominance of social media as a main source of information underscores the need to educate citizens who can adopt a critical stance toward the information that confronts them in their personal, professional, and civic lives. The Conceptual Framework for Learning in the Disciplines provides a basis for building classroom instruction that supports the development of the competencies needed to do so.

References

- Alvermann, D.E., Moore, D.W., 1991. Secondary schools. In: Barr, R., Kamil, M.L., Pearson, P.D. (Eds.), *Handbook of Reading Research*, vol. 2. Longman, New York, pp. 951–983.
- American College Testing, 2006. *Reading Between the Lines: What the ACT Reveals About College Readiness in Reading*. American College Testing, Iowa City, IA.
- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries. OECD Education Working Papers, No. 41. OECD Publishing, Paris. <https://doi.org/10.1787/218525261154>.
- Andriessen, J.E., 2006. Arguing to learn. In: Sawyer, K.L. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge, Oxford, pp. 443–459.
- Bazerman, C., 1985. Physicists reading physics: schema-laden purposes and purpose-laden schema. *Writ. Commun.* 2, 3–23.
- Bazerman, C., 2004. Speech acts, genres, and activity systems: how texts organize activity and people. In: Bazerman, C., Prior, P.A. (Eds.), *What Writing Does and How It Does It: An Introduction to Analyzing Texts and Textual Practices*. Erlbaum, Mahwah, NJ, pp. 309–339.
- Berland, L., Reiser, B.J., 2009. Making sense of argumentation and explanation. *Sci. Educ.* 93 (1), 26–55.
- Braaten, M., Windschitl, M., 2011. Working toward a stronger conceptualization of scientific explanation for science education. *Sci. Educ.* 95, 639–669.
- Bransford, J.D., Brown, A.L., Cocking, R., 2002. *How People Learn: Brain, Mind, Experience, and School*. National Academy of Sciences, Washington, DC.
- Britt, M.A., Rouet, J.-F., Durik, A.M., 2018. *Literacy Beyond Text Comprehension: A Theory of Purposeful Reading*. Routledge, New York, NY.
- Carnegie Council on Advancing Adolescent Literacy (CCAAL), 2011. *Time to Act: An Agenda for Advancing Adolescent Literacy for College and Career Success*. Carnegie Corporation of New York, New York, NY. <https://eric.ed.gov/?id=ED535318>.
- Caviazoglu, B., Stuessy, C., 2017. Changes in science teachers' conceptions and connections of STEM concepts and earthquake engineering. *J. Educ. Res.* 110 (3), 219–254. <https://doi.org/10.1080/00220671.2016.1273176>.
- Chase, W.G., Simon, H.A., 1973. Perception in chess. *Cognit. Psychol.* 1 (1), 33–81.
- Chi, M.T.H., Feltovich, P., Glaser, R., 1981. Categorization and representation of physics problems by experts and novices. *Cognit. Sci.* 5, 121–152.
- Chi, M.T.H., Glaser, R., Farr, M. (Eds.), 1988. *The Nature of Expertise*. Erlbaum, Hillsdale, NJ.
- Chinn, C.A., Brewer, W., 1996. Mental models in data interpretation. *Philos. Sci.* 63, S211–S219.
- Chinn, C.A., Malhotra, B.A., 2002. Epistemologically authentic reasoning in schools: a theoretical framework for evaluating inquiry tasks. *Sci. Educ.* 86, 175–218.
- Chinn, C.A., Buckland, L.A., Samarapungavan, A., 2011. Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology. *Educ. Psychol.* 46, 141–167.
- Council of Chief State School Officers (CCSSO), 2010. *Common Core State Standards for. National Governors Association Center for Best Practices*, Washington, DC.
- Danish, J., Gresalfi, M., 2018. Cognitive and sociocultural perspective on learning: tensions and synergy in the Learning Sciences. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 34–43.
- Dunbar, K., 1995. How scientists really reason: scientific reasoning in real-world laboratories. In: Sternberg, R.J., Davidson, J. (Eds.), *Mechanisms of Insight*. MIT Press, Cambridge, MA, pp. 365–395.
- Duschl, R.A., Osborne, J., 2002. Supporting and promoting argumentation discourse in science education. *Stud. Sci. Educ.* 38, 39–72.
- Ford, M.J., Wargo, B.M., 2012. Dialogic framing of scientific content for conceptual and epistemic understanding. *Sci. Educ.* 96, 369–391.
- Gee, J.P., 1992. *The Social Mind: Language, Ideology, and Social Practice*. Bergin and Garvey, New York, NY.
- Goldman, S.R., 2012. Adolescent literacy: learning and understanding content. *Future Child.* 22, 89–116.
- Goldman, S.R., 2015. Reading and the web: broadening the need for complex comprehension. In: Spiro, R.J., DeSchryver, M., Hagerman, M.S., Morsink, P., Thompson, P. (Eds.), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices*. Routledge, New York, NY, pp. 89–103.
- Goldman, S.R., 2018. Discourse of learning and the learning of discourse. *Discourse Process.* 55, 434–453. <https://doi.org/10.1080/0163853X.2018.1440866>.
- Goldman, S.R., Bisanz, G.L., 2002. Toward a functional analysis of scientific genres: implications for understanding and learning processes. In: Otero, J., Leon, J.A., Graesser, A.C. (Eds.), *The Psychology of Science Text Comprehension*. Erlbaum, Mahwah, NJ, pp. 19–50.
- Goldman, S.R., Popp, J.S., (in press). Transforming sixth grade social studies from “just the facts” to historical inquiry: a case study of teacher learning. *Hist. Teach.*
- Goldman, S.R., Petrosino, A.J., The Cognition and Technology Group at Vanderbilt, 1999. Design principles for instruction in content domains: lessons from research on expertise and learning. In: Dursio, F.T. (Ed.), *Handbook of Applied Cognition*. Wiley, Chichester, England, pp. 595–627.
- Goldman, S.R., Britt, M.A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C.D., Shanahan, C., Project READI, 2016. *Disciplinary literacies and learning to read for understanding: a conceptual framework of core processes and constructs*. *Educ. Psychol.* 51, 219–246.
- Goldman, S.R., Greenleaf, C., Yukhymenko-Lescroart, M., with Brown, W., Ko, M.-L.M., Emig, J.M., George, M., Wallace, P., Blaum, D., Britt, M.A., Project READI, 2019. Explanatory modeling in science through text-based investigation: testing the efficacy of the Project READI intervention approach. *Am. Educ. Res. J.* 56 (4), 1148–1216. <https://doi.org/10.3102/0002831219831041>.
- Goldman, S.R., Hall, A., Ko, M.-L.M., 2022. Co-design as an interactive context for teacher learning. In: Castro Superfine, A.M., Goldman, S.R., Ko, M.-L.M. (Eds.), *Teacher Learning in Changing Contexts: Perspectives From the Learning Sciences (xxx-xxx)*. Routledge.
- Greene, S., 1994. The problems of learning to think like a historian: writing history in the culture of the classroom. *Educ. Psychol.* 29, 89–96.
- Greenleaf, C., Valencia, S., 2017. Missing in action: learning from texts in subject-matter classrooms. In: Appleman, D., Hinchman, K. (Eds.), *Adolescent Literacy: A Handbook of Practice-Based Research*. Guilford Press, New York, NY, pp. 235–256.
- Greenleaf, C.L., Schoenbach, R., Cziko, C., Mueller, F.L., 2001. Apprenticing adolescent readers to academic literacy. *Harv. Educ. Rev.* 71, 1–51.
- Greeno, J.G., Engeström, Y., 2014. Learning in activity. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, New York, NY, pp. 128–147.
- Hall, A.H., Goldman, S.R., 2022. Toward Dialogic Disciplinary Discourse in the Classroom: Teacher and Researcher Collaborative Reflection on Classroom Video. In *Proceedings of the CSCL Conference 2022*.
- Hall, A.H., George, M., Goldman, S.R., 2022. Productive Disciplinary Discussion in Literary Reading: Shifting Classroom Practice. Paper Presented at American Educational Research Association. San Diego, CA.
- Hillocks, G., 2011. *Teaching Argument Writing, Grades 6–12: Supporting Claims With Relevant Evidence and Clear Reasoning*. Heinemann, New York, NY.
- Holland, D., Lave, J., 2009. Social practice theory and the historical production of persons. *Action* 2, 1–15.
- Hutchins, E., 1995. *Cognition in the Wild*. Cambridge University Press, Cambridge, MA.
- Kelley, T.R., Knowles, J.G., Holland, J.D., Jung, H., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (1), 1–11.
- Kjelvik, M.K., Schultheis, E.H., 2019. Getting messy with authentic data: exploring the potential of using data from scientific research to support student data literacy. *CBE Life Sci. Educ.* 18 (2), es2. <https://doi.org/10.1187/cbe.18-02-0023>.

- Kirsh, D., 1995. The intelligent use of space. *Artif. Intell.* 73, 31–68.
- Ko, M.-L.M., Hall, A., Goldman, S.R., 2022. Making teacher and researcher learning visible: collaborative design as a context for professional growth. *Cognit. Instr.* 40 (1), 27–54.
- Koslowski, B., 1996. *Theory and Evidence: The Development of Scientific Reasoning*. MIT Press, Cambridge, MA.
- Latour, B., Woolgar, S., 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton, NJ.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, New York, NY.
- Lee, P.J., 2005. Putting principles into practice: understanding history. In: Donovan, M.S., Bransford, J.D. (Eds.), *How Students Learn: History, Mathematics, and Science in the Classroom*. The National Academies Press, Washington, DC, pp. 31–77.
- Lee, C.D., 2007. *Culture, Literacy and Learning: Taking Bloom in the Midst of the Whirlwind*. Teachers College Press, New York, NY.
- Lee, C.D., Sprattley, A., 2010. *Reading in the Disciplines: The Challenges of Adolescent Literacy*. Carnegie Corporation of New York, New York, NY.
- Lee, C.D., Goldman, S.R., Levine, S., Magliano, J.P., 2016. Epistemic cognition in literary reasoning. In: Green, J., Sandoval, W., Bråten, I. (Eds.), *Handbook of Epistemic Cognition*. Routledge, New York, NY, pp. 165–183.
- Leinhardt, G., Young, K.M., 1996. Two texts, three readers: distance and expertise in reading history. *Cognit. Instr.* 14, 441–486.
- Levine, S., Hall, A.H., Goldman, S.R., Lee, C.D., 2019. A design architecture for engaging middle and high school students in epistemic practices of literary interpretation. In: Nachowitz, M., Wilcox, K.C. (Eds.), *High Literacy in Secondary English Language Arts: Bridging the Gap to College and Career*. Lexington, Lanham, MD, pp. 105–132.
- Markauskaite, L., Goodyear, P., 2017. *Epistemic Fluency and Professional Education: Innovation, Knowledgeable Action and Actionable Knowledge*. Springer, Dordrecht, Netherlands.
- Megill, A., 1989. Recounting the past: “description,” explanation, and narrative in historiography. *Am. Hist. Rev.* 627–653.
- Moje, E.B., 2008. Foregrounding the disciplines in secondary literacy teaching and learning: a call for change. *J. Adolesc. Adult Lit.* 52, 96–107.
- Moje, E.B., 2015. Doing and teaching disciplinary literacy with adolescent learners: a social and cultural enterprise. *Harv. Educ. Rev.* 85, 254–278.
- Moje, E.B., O’Brien, D.G., 2001. *Constructions of Literacy: Studies of Teaching and Learning in and Out of Secondary Schools*. Lawrence Erlbaum Associates, Mahwah, N.J.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141.
- Nadelson, L.S., Seifert, A.L., 2017. Integrated STEM defined: contexts, challenges, and the future. *J. Educ. Res.* 110 (3), 221–223.
- National Alliance of Business, 2002. *A Nation of Opportunity: Building America’s 21st Century Workforce*. Department of Labor, Washington, DC.
- National Assessment of Educational Progress, 2009. *NAEP 2008 Trends in Academic Progress*. National Center for Educational Statistics, Washington, DC.
- National Assessment of Educational Progress, 2022. *Explore NAEP Long Term Trends in Reading and Mathematics*. Online. <https://www.nationsreportcard.gov/ltr/?age=9>. (Accessed 23 May 2022).
- National Center for Education Statistics, 2013. *The Nation’s Report Card: A First Look: 2013 Mathematics and Reading*. NCES 2014-451. Institute for Education Sciences, Washington, DC.
- National Center for Education Statistics, 2022. *U.S. Adults With Low Literacy and Numeracy Skills: 2012/14 to 2017*. Program for the International Assessment of Adult Competencies (PIAAC). <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022004>.
- National Council for the Social Studies, 2013. *The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History*. NCSS, Silver Spring, MD.
- National Research Council, 2012. In: Pellegrino, J.W., Hilton, M.L. (Eds.), *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*, Committee on Defining Deeper Learning and 21st Century Skills. The National Academies Press, Washington, DC.
- Next Generation Science Standards Lead States, 2013. *Next Generation Science Standards: For States, by States*. National Academies Press, Washington, DC.
- Organization for Economic Co-operation and Development, 2013. *PISA 2012: Results in Focus*. OECD, Paris.
- Organization of Economic and Cultural Development, 2013. *PISA 2015 Draft Frameworks*. Retrieved from. <http://www.oecd.org/pisa/pisaproducts/pisa2015>.
- Osborne, J.F., 2002. Science without literacy: a ship without a sail? *Camb. J. Educ.* 32, 203–215.
- Pearson, P.D., Moje, E., Greenleaf, C., 2010. Literacy and science: each in the service of the other. *Science* 328, 459–463.
- Rabinowitz, P.J., 1987. *Before Reading: Narrative Conventions and the Politics of Interpretation*. Cornell University Press, Ithaca, NY.
- Rainey, E., Moje, E.B., 2012. Building insider knowledge: teaching students to read, write, and think within ELA and across the disciplines. *Engl. Educ.* 45, 71–90.
- Reisman, A., 2012. Reading like a historian: a document-based history curriculum intervention in urban high schools. *Cognit. Instr.* 30, 86–112.
- Rogoff, B., 1997. Evaluating development in the process of participation: theory, methods, and practice building on each other. In: Amsel, E., Renninger, A. (Eds.), *Change and Development: Issues of Theory, Application, and Method*. Erlbaum, Hillsdale, NJ, pp. 265–285.
- Rouet, J.-F., Britt, M.A., 2011. Relevance processes in multiple document comprehension. In: McCrudden, M.T., Magliano, J.P., Schraw, G. (Eds.), *Relevance Instructions and Goal-Focusing in Text Learning*. Information Age Publishing, Greenwich, CT, pp. 19–52.
- Rutherford, F., Ahlgren, A., 1990. *Science for All Americans*. Oxford University, New York, NY.
- Sabers, D.S., Cushing, S., Berliner, D.C., 1991. Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *Am. Educ. Res. J.* 28 (1), 63–88.
- Schauble, L., Glaser, R., Duschl, R.A., Schulze, S., John, J., 1995. Students’ understanding of the objectives and procedures of experimentation in the science classroom. *J. Learn. Sci.* 4, 131–166.
- Seidel, T., Sturmer, K., 2014. Modeling and measuring the structure of professional vision in preservice teachers. *Am. Educ. Res. J.* 51 (4), 739–771.
- Seixas, P., 2010. A modest proposal for change in Canadian history education. *Int. Rev. Hist. Educ.* 6, 11–26.
- Shanahan, C., Shanahan, T., Misischia, C., 2011. Analysis of expert readers in three disciplines : history, mathematics, and chemistry. *J. Lit. Res.* 43, 393–429.
- Education Policy White Paper. In: Shepard, L., Hannaway, J., Baker, E. (Eds.), 2009. *Standards, Assessments, and Accountability*. National Academy of Education, Washington, DC.
- Toulmin, S., Rieke, R., Janik, A., 1984. *An Introduction to Reasoning*. Macmillan Publishing Company, New York, NY.
- Vaughn, S., Swanson, E.A., Roberts, G., Wanzek, J., Stillman-Spisak, S.J., Solis, M., Simmons, D., 2013. Improving reading comprehension and social studies knowledge in middle school. *Read. Res. Q.* 48, 77–93.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Wineburg, S.S., 1991. Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *J. Educ. Psychol.* 83 (1), 73–87.
- Wineburg, S., 1994. The cognitive representation of historical texts. In: Leinhardt, G., Beck, I., Stinton, C. (Eds.), *Teaching and Learning in History*. Erlbaum, Hillsdale, NJ, pp. 85–135.
- Wineburg, S.S., 2001. *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past*. Temple University Press, Philadelphia, PA.
- Wolff, C.E., Jarodzka, H., van den Bogert, N., Boshuizen, H.P.A., 2016. Teacher vision: expert and novice teachers’ perception of problematic classroom management scenes. *Instr. Sci.* 44 (3), 243–265. <https://doi.org/10.1007/s11251-016-9367-z>.
- Yoon, S.A., Shim, J., Miller, K., Cottone, A., Noushad, N., Yoo, J.-U., Gonzalez, M.V., Urbanowicz, R., Himes, B.E., 2022. Professional development for STEM integration: analyzing bioinformatics teaching by examining teachers’ qualities of adaptive expertise. In: Castro, A.M., Goldman, S.R., Ko, M.-L.M. (Eds.), *Teacher Learning in Changing Contexts: Perspectives From the Learning Sciences*. Routledge, pp. xx–xxx.

Learning in history and social studies

Keith C. Barton, Department of Curriculum and Instruction, Indiana University, Bloomington, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Social studies as a school subject	315
Purposes of learning social studies	315
Constraints on research in social studies	317
Learning in social studies	317
Learning in and out of school	318
The affective nature of learning social studies	319
Influence of instruction	319
Conclusions	320
References	320

Conceptualizing, researching, and teaching social studies—including history and the social sciences—is a daunting endeavor. The field varies significantly across national contexts, often faces significant social and political constraints on its content, and is subjected to ongoing and frequently rancorous debates about its nature and purpose. This chapter describes the social and academic contexts that shape social studies curriculum and instruction, including diverse and conflicting perspectives on its content, but it also outlines convergent findings on teaching and learning the subject. In particular, the chapter describes how research has shown that students’ in-school learning is influenced by their social backgrounds and experiences, and that these often have an affective component that cannot be ignored in school. To promote deeper learning and engagement, meanwhile, educators not only must attend to students’ prior ideas but also should engage them in classroom discussion, alongside active investigation of meaningful issues.

Social studies as a school subject

Social Studies refers to a group of school subjects aimed primarily at developing elementary and secondary students’ understanding of society and social relations. Use of the term *social studies*, however, varies across national settings. In some countries, *social studies* is an umbrella term that includes separate and specific school subjects such as history, human geography, economics, and government. Less commonly taught subjects such as psychology, sociology, and anthropology also are usually considered part of social studies. In other countries, *social studies* is a single subject that focuses primarily on contemporary social issues and is taught separately from history, geography, and other social sciences, although it usually incorporates content from those subjects as well. In some countries, *social studies* is the elementary-level equivalent of separate secondary subjects and features separate lessons or units on each of these, with varying levels of integration across subjects. And in still other countries, the term is not used in schools, and the constellation of subjects dealing with society are simply referred to with a designation such as *history and social sciences*. Although the term *social education* is a more comprehensive way of referring to these various configurations of subject matter, and is used by some scholars and professional organizations, that term has not been widely adopted.

Regardless of how these subjects are referred to or organized in particular settings, students’ learning can be understood only by considering the societal purposes that the subjects serve. Because social studies touches on fundamental social and political processes—including national identity, civic participation, international relations, and racial and religious conflict—the subject can be highly controversial, and in many countries the curriculum is a continuing object of cultural and ideological dispute (e.g., [Baildon et al., 2014](#); [Bentrovato et al., 2016](#); [Taylor and Guyver, 2012](#)). Moreover, social studies in schools is influenced by a variety of forces, including not only curriculum documents and instructional materials but also cultural expectations, community pressure, administrative guidelines, and teachers’ own knowledge and beliefs ([Fitchett et al., 2014](#); [Ho et al., 2014](#); [Marker and Mehlinger, 1992](#); [Thornton, 2005](#)). As a result, official curricula and classroom practices often are only loosely connected, and the perspectives and recommendations of scholars in the field typically differ from both of these. Understanding students’ learning in social studies, then, requires consideration of what students are expected to learn and the opportunities they have to do so.

Purposes of learning social studies

The most basic distinction of perspectives in the field is between those that accept and embrace social and civic purposes for social studies, and those that downplay these purposes in favor of individual development or purportedly academic and disciplinary goals. In addition, within the large group of teachers, scholars, and educational agencies that aim to orient the subject toward social and civic life, a further distinction is crucial: Between those whose primary goal is to socialize students into the existing social

structure, and those who see social studies as a means for developing critical capacities that can lead to societal transformation (Thornton, 1994). Although there are important areas of overlap among each of these perspectives, the conceptual distinctions among them profoundly influence the knowledge, skills, and dispositions promoted in schools. Analyzing these differences helps clarify why discussions of social studies are contentious, why existing curricula contain inconsistent and even conflicting content, and why research on learning has been difficult to systematize.

Socialization usually involves developing a sense of national identification and pride (Barton and Levstik, 2004; Thornton, 1994; VanSledright, 2008). This may be attempted through history lessons, patriotic exercises, and holiday celebrations that emphasize national origins and significant political developments in the national past, all of which are typically presented as objects of veneration rather than of critical analysis. In some settings, this approach involves stories of national triumph or suffering and may aim to develop explicit support for ruling regimes, or even to evoke national or ethnic supremacy (Carretero, 2011; Lall, 2008; Papadakis, 2008). In other settings, socialization takes more inclusive forms, by celebrating racial, ethnic, or religious diversity, or aims to develop national identification in more subtle ways, such as through repeated emphasis on national achievements and exclusion of other countries, radical viewpoints, and controversial events or policies (Barton and Levstik, 2004; Ho, 2017; Torney-Purta et al., 1999). These latter approaches are particularly common at the primary level, where apart from occasional celebrations of national heroes, the curriculum largely focuses on simplified and idealized depictions of how local communities are organized, including community services, forms of employment, and geographical features (Brophy and Alleman, 2009; Halvorsen, 2017). The overall goal of this perspective is to enable students to fit into existing society, or at least into an officially promoted view of that society.

A second group of approaches is equally concerned with students' relation to society but aims to develop a more critical, even oppositional, approach to existing social relations (e.g., Busey and Walker, 2017; Engebretson, 2018; Ross, 2017). This usually involves guiding students in identifying societal issues and problems, examining their causes, and making individual or group decisions about how to respond (and often, taking direct action to do so). Toward that end, students study (depending on setting) topics such as racism, sexism, homophobia, economic inequality, Indigenous rights, political oppression, warfare, global inequality, and other pressing societal concerns, whether situated in the past or present. (At the primary level, issues are sometimes more local and immediate, such as food insecurity or the needs of the elderly in the local community). To the extent that individuals or events are celebrated in this approach, examples are typically drawn from activists or protest leaders and their work, although these too are most often chosen from a national framework. Often, this approach is oriented toward achieving superordinate societal goals, such as democracy (Parker, 2003), harmony and justice (Barton and Ho, 2022), or intercultural understanding (Nordgren and Johansson, 2015).

Whereas socialization has the sanction of both tradition and official requirements, critical perspectives have a less pervasive and secure place in the curriculum. Although widely promoted by scholars, in practice they usually depend on efforts of individual educators at the classroom or school level, and implementation takes a wide variety of forms. Especially at the secondary level, critical approaches often are associated with deliberation of controversial issues or with direct social and political action (Wood et al., 2013; Ho et al., 2017; Levinson, 2012). Other times, including but not limited to the primary grades, schools may avoid political topics in favor of transforming relationships among individuals and groups, with the goal of developing more tolerant, supportive, and just interactions, with or without a corresponding focus on the systemic forces that reinforce social divisions (Cremin and Bevington, 2017; Parker and Bickmore, 2020; Zembylas, 2016). In some settings, critical perspectives are part of civics or history, but where those subjects are more constrained, they may be incorporated in other courses or promoted as part of programs such as *human rights education*, *peace education*, *sustainability education*, *multicultural education*, or *global education*.

The third perspective on social studies education eschews patriotic nationalism and other forms of political indoctrination but also holds critical perspectives at arm's length. The overarching goal of this perspective is to teach "academic" content, particularly that which is drawn from university-based disciplines such as history, geography, and economics (e.g., Lambert, 2016; Reisman, 2012; Shanahan and Shanahan, 2016). In some settings, this leads to an emphasis on procedural knowledge, such as how to engage in geographic inquiry or how to develop supportable conclusions from historical evidence. In other settings, the emphasis is not on the procedures of these disciplines but on the substantive conclusions of scholars. Economics education, for example, is strongly oriented toward developing students' understanding of economic concepts and principles (Council for Economic Education, 2010), while in some countries the chief concern of history educators is to align content with current scholarship, so that students are not learning outdated or mythic interpretations (e.g., Kang, 2016).

Civics or citizenship education has a more fraught relationship with this perspective, in part because there is no university discipline called "civics," and the closest analog, political science, differs substantially from the content of most school courses. Nonetheless, a major goal of many civics courses is for students to learn about the structure and function of government (including principles such as popular sovereignty, the rule of law, separation of powers, and so on, as well as methods for influencing politics), and in this form it mirrors the emphasis on learning a body of substantive and procedural knowledge. Across all social studies subjects, particularly at the secondary level, the most common manifestation of the academic perspective consists of this kind of attempt to cover the content of textbooks and official curricula, with little thought given to its implicit or explicit purposes. Many educators—and the wider public—simply equate such official documents with academic disciplines, even though much of the content derives solely from traditions of school subject matter with little connection to current scholarship.

There are important areas of overlap between critical and academic approaches, but also important differences. Both emphasize conceptual understanding rather than memorization or routine application of knowledge; with some exceptions, both focus on procedures of investigation and the use of evidence; and both typically include consideration of multiple perspectives and

interactions or negotiation among competing interests. Academic approaches differ from more critical ones, however, in devoting less attention to students' attitudes, such as commitments to equality, human rights, or political participation. Although such attitudes may be an implicit goal of disciplinary approaches, these are usually considered indirect or secondary outcomes. For example, learning to use evidence to reach conclusions in history or geography should enable students to engage in the same practices when they consider contemporary affairs, and learning about the history of racism, colonialism, or other crucial social issues should not only alert students to the importance of these topics but provide alternatives to popular (and often divisive) understandings. Similarly, learning about concepts such as due process or human rights is frequently assumed to lead to a commitment to such principles. There is, however, limited evidence to support the claim of indirect effects of such academic knowledge (Hicks et al., 2012), and the field is not characterized by any coherent theory of how such outcomes would materialize.

Constraints on research in social studies

The diverse, even conflicting and contradictory, set of purposes for social studies constrains research that seeks to identify a set of straightforward principles of learning in the subject. In the most obvious sense, without agreement on the goals to be achieved, it is difficult to develop a systematic set of research-based recommendations for promoting learning. With some educators prioritizing socialization, others critical perspectives, and others academic ones, comparable analysis of differing settings becomes nearly impossible. As a result, research programs in social studies frequently are difficult to compare and synthesize. Learning to be patriotic, to confront systemic racism, and to analyze data on settlement patterns are likely to make use of very different cognitive processes, and it is difficult to imagine a coherent and useful perspective on learning that would encompass all of them, except at the most general level. This multiplicity of approaches is further complicated by national differences in both scholarship and school curricula, because the very nature of socialization and political action differs across countries (Hahn, 2010). Depending on setting, for example, national ideals may focus on progress, freedom, harmony, grievance, or ethnic identity. Similarly, social action in different settings might be oriented toward community reconciliation, participation in mainstream politics, or popular resistance and civil disobedience. Even academic subjects differ across countries, and so researchers sometimes investigate purportedly disciplinary practices that seem marginal or irrelevant to those in other settings.

The diversity of curricular purposes has other and perhaps more significant consequences for research on learning. One of the most important is that educational practice rarely displays a systematic concern with any one goal. Instead, most official and enacted curricula contain a miscellaneous assortment of socializing exercises, academic content, and attention to societal issues, combined with directives from politicians, recommendations by professional lobbying groups, and the strong hand of tradition. As a result, researchers have difficulty identifying natural settings in which they can investigate any consistent set of curricular goals. Often, in order to secure study sites, researchers in social studies must design instructional interventions that are sufficiently brief and structured that teachers are willing to implement them without disrupting the rest of their curriculum. Such interventions, however, may be so simplified that they fail to address either substantive curricular goals or the daily practice of educational institutions. Some investigators have engaged in designed-based research (e.g., Parker et al., 2018; Rubin et al., 2019) in order to enhance the possibility of collecting more focused data, as well as to promote more ethical and collaborative research, but the number of educators who are willing to take part in such projects—which often represent a major departure from common practice—may be limited.

Contributing to this divide between scholarship and practice is the disparity between the most prevalent curricular goals and the interests of scholars. On the one hand, few researchers are interested in investigating the socializing function of social studies, particularly in its common patriotic or nationalistic forms. Although research on political socialization was once common, most educational scholars today reject the importance of studying how to develop students' allegiance to the nation-state through formal education. To the extent that they are interested in this topic, their research focuses on how schools aim to blunt critical perspectives (e.g., Heilig et al., 2012; Sabzalian et al., 2021); how students' social identities develop outside formal schooling (e.g., Lévesque and Croteau, 2022; Levy, 2014); and how students may ignore or resist the socializing function of schooling (e.g., Choi et al., 2011; El-Haj, 2007; Harris and Reynolds, 2014). On the other hand, many teachers are just as reluctant to engage in practices that involve critically examining societal issues, particularly since these inevitably are controversial and thus spark fears of community resistance, uncontrolled emotional reactions from students, or teachers' own pedagogical limitations (reviewed in Ho et al., 2017; Pace, 2019). Although academic approaches are often considered "safer," and thus easier to investigate in naturalistic settings, these too frequently can lead to controversial perspectives that teachers would rather avoid, or that they may find threatening to their own worldviews (James, 2008).

Learning in social studies

Despite obstacles to research on the effectiveness of instructional methods, a substantial body of knowledge exists that is relevant to understanding how students' ideas develop. A number of handbooks on social studies education and its constituent subjects (Arthur et al., 2008; Carretero et al., 2017; Davies et al., 2018; Levstik and Tyson, 2008; Manfra and Bolick, 2017; Metzger and Harris, 2018), along with synthetic reviews of research in the field (Aitken and Sinnema, 2008; Barton and Avery, 2016; Hicks et al., 2012), provide extensive and specific insight into findings related to learning, particularly with regard to history and civic education.

This research illustrates that students from a young age are active constructors of meaning; that their ideas have arisen in a variety of contexts; that ideas developed outside school interact in complicated ways with curriculum content; and that the social purposes of learning influence students' understandings. The following sections draw primarily from these prior reviews, supplemented by more recent research.

Learning in and out of school

One of the most basic findings of research on students' understanding of social studies is that they develop powerful ideas about the social world outside of formal schooling, due to their participation from an early age in civic activities, economic transactions, engagements with the past, and navigation of their physical and social environments. Each of these is mediated by social context—family, community members, institutions, and the media (Hauver, 2019; Kang, 2020; Woodson, 2019). From these sources, students not only acquire discrete facts but also develop more complex conceptual structures that influence their understanding of social interactions, spatial patterns, and change over time. Students in the early years of schooling, for example, already know that fashion, technology, and the built environment have changed over time, and they have at least rough ideas about the nature of those changes and their relative chronology.

Information and conceptual structures such as these can provide building blocks for formal instruction, provided that teachers systematically elicit and activate students' prior knowledge and make direct connections to new content. Educators must be aware, however, of the ways in which prior knowledge varies among individuals and groups. Because knowledge of the social world is closely tied to social context, students with differing experiences are likely to bring diverse understandings to the classroom. Students from dominant, majority, or privileged backgrounds typically have greater access to the knowledge prioritized at school, and this may leave students from other backgrounds at a disadvantage. All students, however, have relevant experiences that can provide the basis for learning, if educators are encouraged to elicit and incorporate their knowledge. Students whose parents are migrant agricultural laborers, for example, may have rich concepts of geography and economics that other students lack.

Although students' prior understanding can provide an important foundation for later learning, these ideas may also differ from formal curricula and, in some cases, present obstacles to instruction. In history, for example, younger students are more likely to have learned about how social and material culture have changed over time than about the kinds of legal and political history that often frames the subject in school. Similarly, outside school younger students typically encounter social relations in terms of interpersonal relationships, and they have limited opportunities to develop complex understandings of societal structures such as government or the economy (Barton, 2020). Older students may demonstrate some of these same patterns of thinking, but they may have existing insights into societal power relations when these are linked to their own experiences (Colley, 2019). Educators must be careful, then, not to assume that students uniformly have the foundational knowledge needed to make sense of curriculum content. Concepts related to societal institutions in particular are ideas that must be developed as part of formal instruction—through careful links to the prior knowledge that students do have—rather than starting points for learning.

Two areas in which students may have incomplete or misleading ideas are related to the use of evidence to reach conclusions, and consideration of perspectives that differ from their own. Findings related to both these areas come primarily from history education, but these patterns are likely to be representative of other aspects of social studies as well. Outside school, students typically have not developed an awareness of how historical knowledge is constructed from evidence such as documents, artifacts, and other original sources; instead, they often assume that knowledge of the past is handed down through word of mouth. Similarly, students may struggle to evaluate authors, sources, and evidence when engaging with online information about social and political topics (McGrew et al., 2018). Students can learn to evaluate sources through instruction, but breaking through their initial ideas is easier than developing deeper understanding. Students easily come to see that history depends on a variety of sources that are incomplete and inconsistent, for example, and that knowledge of the past is uncertain, but they have more trouble seeing how to use such sources to reach supportable conclusions. Rather than engaging in a careful synthesis of evidence, they sometimes dismiss all interpretations as "just an opinion," and consider one account as good as another.

Students often are aware that people in the past thought and acted differently than people today, but they sometimes attribute this to deficiencies of character or intelligence, rather than to the influence of differing historical contexts. Students who have greater historical knowledge, however, are better able to engage in such contextualization (Huijgen et al., 2017). Similarly, although younger students are aware of superficial cultural differences in the present, they may harbor simplified and stereotypical understandings that show little awareness of the reasons behind differing cultural patterns (Torres, 2019). As with learning about historical evidence, students can easily come to recognize that people in the past were not simply deficient versions of modern people, or that people from other cultures have differing beliefs and assumptions, but they have more difficulty imagining how the perspectives of people today may be superseded in the future, or how their own cultures may look to others. Present-day and local patterns of thought and action, that is, still serve as the norm to which other perspectives are compared.

Students' understanding of history also is heavily influenced by narrative structure, and this influence occurs at multiple levels. First, students often over-narrativize historical content, by simplifying it into a limited number of people, settings, and events, and by organizing that content into the problem-solution structure commonly found in fictional stories. In addition, students often assimilate new historical content not just into a general narrative structure but into specific narrative templates, such as "freedom and progress" or "struggle against hardship," even when these have limited applicability (Barton, 2021; Kim, 2018b). And finally,

students may rely on specific narratives with which they are already familiar as a basis for interpreting—and misunderstanding—other historical events and settings (Kim, 2018a; Santiago, 2017). These patterns, however, differ across countries and illustrate the pervasive impact of societal discourses on students' learning. Common historical narratives differ between and within countries, as do the templates that underlay specific narratives. Even the prevalence of narrative itself, particularly in school settings, varies across countries, and this influences students' understanding of the past.

The affective nature of learning social studies

The most pervasive historical narratives almost always are connected to contemporary social and political events, and this connection illustrates an inescapable aspect of learning social studies: the affective salience of much of the content (Garrett et al., 2020). Current political issues often inspire deeply emotional reactions, and history often touches on topics that involve violence, tragedy, and trauma experienced by communities with which students identify. As a result, students' reactions to both contemporary and historical issues often have a strong affective component, and this may either motivate civic action or cause students to turn away from difficult topics (Abowitz and Mamlok, 2019; Miles, 2019). Students' affective connection to issues and identities also influences how they evaluate historical and contemporary sources, as they often dismiss those that do not align with their prior commitments (Jacobsen et al., 2018; Stoddard and Chen, 2018).

In addition, because much of social studies aims, implicitly or explicitly, to develop a sense of national identity among students—and frequently to develop allegiance to particular governmental structures or even specific leaders—students' social identities, and the relationship of their communities to the nation or the state, intersect in important ways with formal learning. Historical narratives that students learn in their communities may be unrelated to, or in opposition to, those they encounter in schools, which typically ignore or downplay forms of systematic oppression, and often omit or misrepresent the experience of minorities, as well as women and other groups who have historically lacked power (Crocco, 2001; King and Brown, 2014; Rodríguez, 2018; Sabzalian, 2019; Thornton, 2002). Even the seemingly uncontroversial content of most elementary curricula in social studies, which usually emphasizes local neighborhoods and communities, often presents a simplified and idealized version of society that has little connection to many students' lives.

As a result of this potential disparity between identities promoted within and outside school, students who identify with oppressed or marginalized groups, or who belong to immigrant or undocumented groups, may ignore or resist much of the content of social studies, or simply find it irrelevant to their own lives. Other times, they may develop fluid, shifting, and complex identities with multiple allegiances (Bondy, 2016; Hahn, 2020), or draw on historical and contemporary content learned both in and out of school as they develop their own interpretive frameworks (Barton, 2012; Johnson, 2019; Yoder, 2020). Similarly, although students from all backgrounds have experience with civic practices, these may reflect forms of public participation that have limited connection to the mainstream ideas they are expected to learn at school (Clay and Rubin, 2020), and their own ideas about civic identity may be grounded in affective attachments and boundaries that transcend a given nation-state (Keegan, 2019). And because social studies often touches on conflicting perspectives, differing political opinions, and controversial issues, students' ideas may be influenced by their own, socially-situated ideological commitments, interpretative frames, and emotional reactions (Colley, 2019). Not only students' ideas but also their willingness to engage in civic activities and discussions is influenced by affective factors, including the relationships they have with others, their assessment of the trust they can feel toward others, their feelings of empathy, their sense of agency, and their judgments of risk (Bellino, 2017; Hauver, 2019).

Influence of instruction

Despite the numerous ways in which factors external to schooling contribute to and intersect with students' learning, formal instruction does have an important influence on students' understanding, including for the youngest students. This seems an unremarkable observation, but it stands in contrast to two long-standing assumptions of some scholars and educators: (1) That historical reasoning is too difficult for any but the most advanced students; (2) That civic and political attitudes are largely impervious to specific interventions at school. Decades of research, though, have now undermined these misconceptions and have provided insight into the impact of instruction. Most of this research has focused on the pedagogical tasks that lead to "deep processing" of content (Hicks et al., 2012) and that are associated with cognitive and affective gains.

Numerous studies from a variety of settings, and using differing methodologies, have demonstrated that classroom discussion (in which students engage in in-depth, substantive exchange of ideas) and an open classroom climate (in which students are encouraged to share their opinions freely) are associated with a variety of positive outcomes. These include greater knowledge, more empathy and tolerance, better argumentation, higher levels of perceived political efficacy, and greater intention of voting and engaging in other aspects of public participation. Some studies have separated the effect of discussion and climate from other aspects of instruction, but most interventions include these in combination with other tasks, so their discrete impact is not always obvious. Most social studies educators assume that discussion and an open climate are required for any effective instructional procedure, and conversely, that the exchange of ideas without substantive content would be pointless.

Another body of research investigates the impact of activities that engage students in analyzing and drawing conclusion from evidence, such as geographic data or historical documents and artifacts. These studies have shown that such investigation-oriented tasks, particularly when combined with products such as writing essays or developing presentations, usually are associated with greater knowledge construction by students; longer, more accurate, and more thesis-based products; better understanding of

causal relations, and increases in cultural awareness and empathy (Schlemper et al., 2018). Similarly, engaging in simulations of political processes such as the United Nations or national legislative assemblies—which also require investigation, synthesis, and presentation of information—usually are associated with greater student engagement, as well as gains in knowledge, political efficacy, and perspective-taking (Levy, 2018). Notably, however, many of these studies show mixed findings, with some outcomes better supported than others, or with results that intersect with other variables such as reading level. In addition, this body of research consistently points to the importance of carefully structured instructional supports, even when problems themselves are unstructured.

One of the most troubling findings from research on instructional practices in social studies is that discussion- and investigation-oriented pedagogies are more likely to be offered to students from higher income backgrounds, of higher academic achievement, and (in the United States) who are White. Students in lower academic tracks, and immigrant or transnational students (or those who are assumed to be), can also experience negative and limiting messages from schools about their roles in society, along with lower expectations for future political participation (Dabach, 2014; Salinas, 2006). These experiences are not uniform, however, and high expectations and high-quality instructional practices can encourage young people from marginalized backgrounds to engage productively with social studies content and to develop positive civic identities (Callahan and Obenchain, 2013).

Conclusions

In order for social studies to be a meaningful subject of study, educators must attend to the ideas students bring with them. These ideas have been developed through students' own life experiences, as they participate in local, national, and even global communities. Without attention to these ideas, instruction at school will be both ineffective (because it does not build on students' prior knowledge or attend to gaps in their understanding) and irrelevant (because it fails to address perspectives and practices that students consider important). Schools cannot, that is, attempt to deliver a social studies curriculum that fails to engage with the social experiences of students themselves. Only by connecting with students' backgrounds—in all their complexity and diversity—can social studies deepen and enrich their ideas, so that students better understand the social world.

Connecting with students also means developing clarity about the subject's purpose. Students live in a world with immense social problems—hunger, poverty, racism, homophobia, gender discrimination, warfare, and other persistent issues. Many students and their communities experience these problems themselves, and all must play a role in solving them. Social studies that aims to indoctrinate students or develop an uncritical acceptance of the status quo can never be relevant to the issues that have meaning for students, and can only be a tool of further oppression. Academic or disciplinary approaches are hardly any better, for they too distance the subject from topics that matter in the real world, with little more than an unfounded hope that students may someday find their content useful. Social studies and its constituent subjects must seek a new path forward, one that directly engages students in using evidence to deliberate contemporary social issues. Educators may find it threatening to confront difficult and controversial questions, but only by doing so can social studies contribute to a more just and harmonious future.

References

- Abowitz, K.K., Mamlok, D., 2019. The case of #NeverAgainMSD: when proceduralist civics becomes public work by way of political emotion. *Theor. Res. Soc. Educ.* 47, 155–175. <https://doi.org/10.1080/00933104.2019.1586611>.
- Aitken, G., Sinnema, C., 2008. *Effective Pedagogy in Social Sciences/tikanga ā iwi: Best Evidence Synthesis Iteration*. New Zealand Ministry of Education, Wellington, New Zealand.
- Arthur, J., Davies, I., Hahn, C. (Eds.), 2008. *The SAGE Handbook of Education for Citizenship and Democracy*. Sage, Los Angeles, CA.
- Baidon, M., Seng, L.K., Lim, I.M., Inanç, G., Jaffar, J. (Eds.), 2014. *Controversial History Education in Asian Contexts*. Routledge, New York.
- Barton, K.C., Avery, P.G., 2016. Research on social studies education: diverse students, settings, and methods. In: Bell, C.A., Gitomer, D. (Eds.), *Handbook of Research on Teaching*, fifth ed. American Educational Research Association, Washington, DC, pp. 985–1038.
- Barton, K.C., Ho, L.C., 2022. *Curriculum for Justice and Harmony: Deliberation, Knowledge, and Action in Social and Civic Education*. Routledge, New York.
- Barton, K.C., Levstik, L.S., 2004. *Teaching History for the Common Good*. Routledge, New York.
- Barton, K.C., 2012. School history as a resource for constructing identities: implications of research from the United States, Northern Ireland, and New Zealand. In: Carretero, M., Asensio, M., Rodríguez-Moneo, M. (Eds.), *History Education and the Construction of National Identities*. Information Age, Charlotte, NC, pp. 93–107.
- Barton, K.C., 2020. Students' understanding of institutional practices: the missing dimension in human rights education. *Am. Educ. Res. J.* 57, 188–217. <https://doi.org/10.3102/0002831219849871>.
- Barton, K.C., 2021. Schematic templates and diverse populations in the United States: narrative limitations in young people's understanding. In: de Luna, I.B., van Alphen, F. (Eds.), *Reproducing, Rethinking, Resisting National Narratives: A Sociocultural Approach to Schematic Narrative Templates in Times of Nationalism*. Information Age, Charlotte, NC, pp. 79–95.
- Bellino, M., 2017. *Youth in Postwar Guatemala: Education and Civic Identity in Transition*. Rutgers University Press, New Brunswick, NJ.
- Bentrovato, D., Korostelina, K.V., Schulze, M., 2016. History Can Bite: History Education in Divided and Postwar Societies. V&R Academic, Göttingen, Germany.
- Bondy, J.M., 2016. Latina youth, education, and citizenship: a feminist transnational analysis. *Theor. Res. Soc. Educ.* 44, 212–243. <https://doi.org/10.1080/00933104.2016.1170644>.
- Brophy, J., Allman, J., 2009. Meaningful social studies for elementary students. *Teach. Teach.* 15, 357–376. <https://doi.org/10.1080/13540600903056700>.
- Bussey, C.L., Walker, I., 2017. A dream and a bus: black critical patriotism in elementary social studies standards. *Theor. Res. Soc. Educ.* 45, 456–488. <https://doi.org/10.1080/00933104.2017.1320251>.
- Callahan, R.M., Obenchain, K.M., 2013. Bridging worlds in the social studies classroom: teachers' practices and Latino immigrant youths' civic and political development. In: Kawecka Nenga, S., Taft, J.K. (Eds.), *Sociological Studies of Children and Youth 16: Youth Engagement: The Civic-Political Lives of Children and Youth*. Emerald Group Publishing, Bingley, United Kingdom, pp. 97–123.

- Carretero, M., Grever, M., Berger, S. (Eds.), 2017. *The Palgrave Handbook of Research in Historical Culture and Education*. Palgrave Macmillan, New York.
- Carretero, M., 2011. *Constructing Patriotism: Teaching History and Memories in Global Worlds*. Information Age Publishing, Charlotte, NC.
- Choi, Y., Lim, J.H., An, S., 2011. Marginalized students' uneasy learning: Korean immigrant students' experiences of learning social studies. *Soc. Stud. Res. Pract.* 6, 1–17.
- Clay, K.L., Rubin, B.C., 2020. "I look deep into this stuff because it's a part of me": toward a critically relevant civics education. *Theor. Res. Soc. Educ.* 48, 161–181. <https://doi.org/10.1080/00933104.2019.1680466>.
- Colley, L., 2019. (Un)Restricting feminism: high school students' definitions of gender and feminism in the context of the historic struggle for women's rights. *Theor. Res. Soc. Educ.* 47, 426–455. <https://doi.org/10.1080/00933104.2019.1593268>.
- Council for Economic Education, 2010. *Voluntary National Content Standards in Economics*, second ed. Council for Economic Education, New York.
- Cremin, H., Bevington, T., 2017. *Positive Peace in Schools: Tackling Conflict and Creating a Culture of Peace in the Classroom*. Routledge, New York.
- Crocco, M.S., 2001. The missing discourse about gender and sexuality in the social studies. *Theor. Pract.* 40, 65–71. https://doi.org/10.1207/s15430421tip4001_10.
- Dabach, D.B., 2014. "You can't vote, right?": when language proficiency is a proxy for citizenship in a civics classroom. *J. Int. Soc. Stud.* 4, 37–56.
- Davies, I., Ho, L.C., Kiwan, D., Peck, C.L., Peterson, A., Sant, E., Waghid, Y., 2018. *The Palgrave Handbook of Global Citizenship and Education*. Palgrave Macmillan, New York.
- El-Haj, T., 2007. "I was born here, but my home, it's not here": educating for democratic citizenship in an era of transnational migration and global conflict. *Harv. Educ. Rev.* 77, 285–316. <https://doi.org/10.17763/haer.77.3.41217m737q114h5m>.
- Engelbreton, K., 2018. Toward a gender-inclusive vision for powerful and authentic social studies. In: Martell, C.C. (Ed.), *Social Studies Teacher Education: Critical Issues and Current Perspective*. Information Age, Charlotte, NC, pp. 113–132.
- Fitchett, P.G., Heafner, T.L., Lambert, R.G., 2014. Examining elementary social studies marginalization: a multilevel model. *Educ. Pol.* 28, 40–68. <https://doi.org/10.1177/0895904812453998>.
- Garrett, H.J., Segall, A., Crocco, M.S., 2020. Accommodating emotion and affect in political discussions in classrooms. *Soc. Stud.* 111, 312–323. <https://doi.org/10.1080/00377996.2020.1758015>.
- Hahn, C.L., 2010. Comparative civic education research: what we know and what we need to know. *Citizen. Teach. Learn.* 6, 5–23. https://doi.org/10.1386/ctl.6.1.5_1.
- Hahn, C.L., 2020. Educating citizens in an age of globalization, migration, and transnationalism: a study in four European democracies. *Theor. Res. Soc. Educ.* 48, 244–284. <https://doi.org/10.1080/00933104.2019.1707139>.
- Halvorsen, A., 2017. Children's learning and understanding in their social world. In: Manfra, M.M., Bolick, C.M. (Eds.), *The Wiley Handbook of Social Studies Research*. Wiley-Blackwell, Malden, MA, pp. 385–413.
- Harris, R., Reynolds, R., 2014. The history curriculum and its personal connection to students from minority ethnic backgrounds. *J. Curric. Stud.* 46, 464–486. <https://doi.org/10.1080/00220272.2014.881925>.
- Hauver, J., 2019. *Young Children's Civic Mindedness: Democratic Living and Learning in an Unequal World*. Routledge, New York.
- Heilig, J.V., Brown, K., Brown, A., 2012. The illusion of inclusion: a critical race theory textual analysis of race and standards. *Harv. Educ. Rev.* 82, 403–424. <https://doi.org/10.17763/haer.82.3.84p8228670j24650>.
- Hicks, D., van Hove, S., Doolittle, P.E., VanFossen, P., 2012. Learning social studies: an evidence-based approach. In: Harris, K.R., Graham, S., Urden, T., Bus, A.G., Major, S., Swanson, H.L. (Eds.), *APA Educational Psychology Handbook 3: Application to Learning and Teaching*. American Psychological Association, Washington, DC, pp. 283–307.
- Ho, L.C., Alviar-Martin, T., Leviste, E.N., 2014. "There is space, and there are limits": the challenge of teaching controversial topics in an illiberal democracy. *Teach. Coll. Rec.* 116, 1–28. <https://doi.org/10.1177/016146811411600402>.
- Ho, L.C., McAvoy, P., Hess, D., Gibbs, B., 2017. Teaching and learning about controversial issues and topics in the social studies. In: Manfra, M.M., Bolick, C.M. (Eds.), *The Wiley Handbook of Social Studies Research*. Wiley-Blackwell, Malden, MA, pp. 319–335.
- Ho, L.C., 2017. Freedom can only exist in an ordered state': harmony and civic education in Singapore. *J. Curric. Stud.* 49, 476–496. <https://doi.org/10.1080/00220272.2016.1155648>.
- Huijgen, T., van Bortel, C., van de Grift, W., Holthuis, P., 2017. Toward historical perspective taking: students' reasoning when contextualizing the actions of people in the past. *Theor. Res. Soc. Educ.* 45, 110–144. <https://doi.org/10.1080/00933104.2016.1208597>.
- Jacobsen, R., Halvorsen, A., Frasier, A.S., Schmitt, A., Crocco, M., Segall, A., 2018. Thinking deeply, thinking emotionally: how high school students make sense of evidence. *Theor. Res. Soc. Educ.* 46, 232–276. <https://doi.org/10.1080/00933104.2018.1425170>.
- James, J.H., 2008. Teachers as protectors: making sense of preservice teachers' resistance to interpretation in elementary history teaching. *Theor. Res. Soc. Educ.* 36, 172–205. <https://doi.org/10.1080/00933104.2008.10473372>.
- Johnson, M.W., 2019. Trump, Kaepernick, and MLK as "maybe citizens": early elementary African American males' analysis of citizenship. *Theor. Res. Soc. Educ.* 47, 374–395. <https://doi.org/10.1080/00933104.2019.1582381>.
- Kang, S.J., 2016. Postcolonial discourses and teaching national history: the history educators' attempts to overcome colonialism in the Republic of Korea. In: Carretero, M., Berger, S., Grever, M. (Eds.), *Palgrave Handbook of Research in Historical Culture and Education*. Palgrave Macmillan, New York, pp. 331–353.
- Kang, J., 2020. Victims or competitors: Korean students' competing understandings of migrants. *Multicult. Educ. Rev.* 12, 195–214. <https://doi.org/10.1080/2005615X.2020.1808931>.
- Keegan, P., 2019. Migrant youth from West African countries enacting affective citizenship. *Theor. Res. Soc. Educ.* 47, 347–373.
- Kim, G., 2018a. "Because the United States is a Great Melting Pot": The Influence of Students' Sociocultural Backgrounds on Understanding World History. Unpublished Doctoral Dissertation. Indiana University, Bloomington, IN.
- Kim, G., 2018b. Holding the severed finger: Korean students' understanding of historical significance. *J. Curric. Stud.* 50, 508–534.
- King, L.J., Brown, K., 2014. Once a year to be Black: fighting against typical Black history month pedagogies. *Negro Educ. Rev.* 65, 23–43.
- Lall, M., 2008. Educate to hate: the use of education in the creation of antagonistic national identities in India and Pakistan. *Compare* 38, 103–119.
- Lambert, D., 2016. Geography. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The Sage Handbook of Curriculum, Pedagogy and Assessment*. Sage, Los Angeles, CA, pp. 391–408.
- Lévesque, S., Croteau, J., 2022. "We will continue our struggle for success": French Canadian students, narrative, and historical consciousness. *Theor. Res. Soc. Educ.* 50, 101–124. <https://doi.org/10.1080/00933104.2021.2019153>.
- Levinson, M., 2012. *No Citizen Left Behind*. Harvard University Press, Cambridge, MA.
- Levstik, L.S., Tyson, C.A., 2008. *Handbook of Research in Social Studies Education*. Routledge, New York.
- Levy, S.A., 2014. Heritage, history, and identity. *Teach. Coll. Rec.* 116, 1–34. <https://doi.org/10.1177/016146811411600605>.
- Levy, B.L.M., 2018. Youth developing political efficacy through social learning experiences: becoming active participants in a supportive Model United Nations Club. *Theor. Res. Soc. Educ.* 46, 410–448.
- Manfra, M.M., Bolick, C.M., 2017. *The Wiley Handbook of Social Studies Research*. Wiley-Blackwell, Maley, MA.
- Marker, G., Mehlinger, H., 1992. Social studies. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 830–851.
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., Wineburg, S., 2018. Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theor. Res. Soc. Educ.* 46, 165–193. <https://doi.org/10.1080/00933104.2017.1416320>.
- Metzger, S.A., Harris, L.A., 2018. *The Wiley Handbook of History Teaching and Learning*. Wiley-Blackwell, Malden, MA.
- Miles, J., 2019. Seeing and feeling difficult history: a case study of how Canadian students make sense of photographs of Indian residential schools. *Theor. Res. Soc. Educ.* 47, 472–496. <https://doi.org/10.1080/00933104.2019.1626783>.
- Nordgren, K., Johansson, M., 2015. Intercultural historical learning: a conceptual framework. *J. Curric. Stud.* 47, 1–25. <https://doi.org/10.1080/00220272.2014.956795>.

- Pace, J., 2019. Contained risk-taking: preparing preservice teachers to teach controversial issues in three countries. *Theor. Res. Soc. Educ.* 47, 228–260. <https://doi.org/10.1080/00933104.2019.1595240>.
- Papadakis, Y., 2008. Narrative, memory and history education in divided Cyprus: a comparison of schoolbooks on the “History of Cyprus”. *Hist. Mem.* 20, 128–148.
- Parker, C., Bickmore, K., 2020. Classroom peace circles: teachers’ professional learning and implementation of restorative dialogue. *Teach. Teach. Educ.* 95. <https://doi.org/10.1016/j.tate.2020.103129>.
- Parker, W.C., Valencia, S.W., Lo, J.C., 2018. Teaching for deeper political learning: a design experiment. *J. Curric. Stud.* 50, 252–277. <https://doi.org/10.1080/00220272.2017.1343386>.
- Parker, W.C., 2003. *Teaching Democracy: United and Diversity in Public Life*. Teachers College Press.
- Reisman, A., 2012. Reading like a historian: a document-based history curriculum intervention in urban high schools. *Cognit. Instr.* 30, 86–112. <https://doi.org/10.1080/07370008.2011.634081>.
- Rodríguez, N.N., 2018. From margins to center: developing cultural citizenship education through the teaching of Asian American history. *Theor. Res. Soc. Educ.* 46, 528–573. <https://doi.org/10.1080/00933104.2018.1432432>.
- Ross, E.W., 2017. *Rethinking Social Studies: Critical Pedagogy in Pursuit of Dangerous Citizenship*. Information Age, Charlotte, NC.
- Rubin, B.C., Freedman, E.B., Kim, J., 2019. *Design Research in Social Studies Education: Critical Lessons From an Emerging Field*. Routledge, New York.
- Sabzalian, L., Shear, S.B., Snyder, J., 2021. Standardizing Indigenous erasure: a TribalCrit and QuantCrit analysis of K–12 U.S. civics and government standards. *Theor. Res. Soc. Educ.* 49, 321–359. <https://doi.org/10.1080/00933104.2021.1922322>.
- Sabzalian, L., 2019. *Indigenous Children's Survivance in Public Schools*. Routledge, New York.
- Salinas, C., 2006. Educating late arrival high school immigrant students: a call for a more democratic curriculum. *Multicult. Perspect.* 8, 20–27. https://doi.org/10.1207/s15327892mcp0801_4.
- Santiago, M., 2017. Erasing differences for the sake of inclusion: how Mexican/Mexican American students construct historical narratives. *Theor. Res. Soc. Educ.* 45, 43–74. <https://doi.org/10.1080/00933104.2016.1211971>.
- Schlemper, M.B., Stewart, V.C., Shetty, S., Czajkowski, K., 2018. Including students’ geographies in geography education: spatial narratives, citizen mapping, and social justice. *Theor. Res. Soc. Educ.* 4, 603–641. <https://doi.org/10.1080/00933104.2018.1427164>.
- Shanahan, T., Shanahan, C., 2016. Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harv. Educ. Rev.* 78, 40–59. <https://doi.org/10.17763/haer.78.1.v62444321p602101>.
- Stoddard, J., Chen, J., 2018. The impact of political identity, grouping, and discussion on young people’s views of political documentaries. *Learn. Media Technol.* 43, 418–433. <https://doi.org/10.1080/17439884.2018.1504790>.
- Taylor, T., Guyver, R., 2012. *History Wars and the Classroom: Global Perspectives*. Information Age, Charlotte, NC.
- Thornton, S.J., 1994. The social studies near century’s end: reconsidering patterns of curriculum and instruction. *Rev. Res. Educ.* 20, 223–254. <https://doi.org/10.3102/0091732X020001223>.
- Thornton, S.J., 2002. Does everybody count as human? *Theor. Res. Soc. Educ.* 30, 178–189. <https://doi.org/10.1080/00933104.2002.10473189>.
- Thornton, S.J., 2005. *Teaching Social Studies That Matters: Curriculum for Active Learning*. Teachers College Press, New York.
- Torney-Purta, J., Schwill, J., Amadeo, J. (Eds.), 1999. *Civic Education Across Countries: Twenty-Four National Case Studies From the IEA Civic Education Project*. The International Association for the Evaluation of Educational Achievement, Amsterdam, Netherlands.
- Torres, H.J., 2019. “They have their own way, and you should respect that”: investigating the outcomes of an elementary world cultures curriculum. *Theor. Res. Soc. Educ.* 47, 548–576. <https://doi.org/10.1080/00933104.2019.1653803>.
- VanSledright, B., 2008. Narratives of nation-state, historical knowledge, and school history education. *Rev. Res. Educ.* 32, 109–146. <https://doi.org/10.3102/0091732X07311065>.
- Wood, B.E., Taylor, R.M., Atkins, R.A., 2013. Fostering active citizenship through the New Zealand social studies curriculum: teachers’ perceptions and practices of social action. *N. Z. J. Educ. Stud.* 48, 84–98.
- Woodson, A.N., 2019. Racial code words, re-memberings and Black kids’ civic imaginations: a critical race ethnography of a post-civil rights leader. *Anthropol. Educ. Q.* 50, 26–47. <https://doi.org/10.1111/aeq.12277>.
- Yoder, P.J., 2020. “He wants to get rid of all the Muslims”: Mexican American and Muslim students’ use of history regarding candidate Trump. *Theor. Res. Soc. Educ.* 48, 346–374. <https://doi.org/10.1080/00933104.2020.1773364>.
- Zembylas, M., 2016. Toward a critical-sentimental orientation in human rights education. *Educ. Philos. Theor.* 48, 1151–1167. <https://doi.org/10.1080/00131857.2015.1118612>.

Learning in math(s)

Alan H. Schoenfeld, School of Education, University of California, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction and overview	323
Characterizing learning in general, and specifically in mathematics	324
Learning goals and learning outcomes	326
Mathematical thinking and learning	326
Antecedents	327
Current understandings	329
Mathematical content	329
Mathematical processes and practices	329
Metacognitive behavior	330
Orientations and identity	330
Productive learning environments	331
Dimension 1, classroom affordances for mathematics learning	331
Dimension 2, cognitive demand and the accessibility of ideas	332
Dimension 3, equitable access to key ideas for all students	332
Dimension 4, opportunities for students to develop productive belief systems, a sense of mathematical agency, and positive mathematical identities	333
Dimension 5, formative assessment	333
Implications and prospects	333
Acknowledgments	334
References	334
Websites of interest	334

Introduction and overview

A comprehensive characterization of learning in mathematics calls for the careful examination of (a) the nature of learning, (b) what is intended to be learned, and (c) the nature of learning environments and the impact they have on learning. Researchers' understandings of these aspects of learning have evolved over the 20th and early 21st centuries, with significant growth in recent decades as mathematics education coalesced and matured as a discipline.

Characterizing learning in general, and specifically in mathematics provides and elaborates on characterizations of learning in general and mathematics learning in particular, the latter being:

Mathematics learning is change in a person's understanding of mathematical content, practices, habits of mind, orientations and mathematical identity that results from that person's classroom and other experiences.

Central ideas in **Characterizing learning in general, and specifically in mathematics** include a broad characterization of mathematical thinking (the goal of instruction) that encompasses knowing mathematical content and practices, developing powerful mathematical habits of mind, and developing productive mathematical orientations (including beliefs and dispositions) and a positive mathematical identity. Mathematics learning involves not only mathematical content and practices, but also misconceptions, changes in one's sense of the discipline, and changes in one's mathematical identity. It includes not only what students take from classroom experiences but, in addition, what they come to understand from the social surround, for example, that mathematical proficiency might be innate or learnable, or gender- or race-linked. All these understandings are consequential because they shape if and how one engages in mathematics.

Learning goals and learning outcomes addresses issues related to the mathematics curriculum and learning outcomes. Issues of standards and the formal curriculum—characterizations of what should be learned, when, and by whom—are in part social questions. They are typically addressed by governmental panels, with input from various stakeholder groups such as mathematicians, education researchers, teachers, and parents or political groups. Such goals (e.g., statistical or mathematical literacy) evolve as a function of societal needs and because of deepening understandings of learning processes (e.g., problem solving).

In addition to the formal curriculum, there are also the “hidden curriculum” and the unintended consequences of learning. The hidden curriculum involves the tacit but often strong messages students pick up about the nature of schooling—for example, how they should interact with peers, teachers, and other adults; how they should perceive different races, groups, or classes of people; or

what ideas and behaviors are considered acceptable or unacceptable (<https://www.edglossary.org/hidden-curriculum/>).” What is learned may include beliefs about what mathematics is, how mathematics is to be learned and done, and who is capable of doing mathematics.

Mathematical thinking and learning addresses the nature of knowing and learning, with a focus on mathematics. **Antecedents** describes philosophical and psychological antecedents to the current era, which began in the 1960s. As discussed in **Current understandings**, mathematics education began to coalesce as a discipline in the 1960s; the “cognitive revolution” began in the 1970s and mathematics education turned to a focus on cognitive processes. Studies of mathematical learning indicated that mathematical thinking, learning, and understanding involved not only acquiring knowledge but the development of problem solving skills, and that aspects of thinking such as metacognition and belief systems are central to the processes of learning and doing mathematics. Attention to such processes and their social origins evolved into a focus on socio-cognitive issues, grounded in the understanding that learning is a function of the practices in which one is engaged, and thus a function of interactions with others. This, of course, has implications both for learning goals (See **Learning goals and learning outcomes**) and learning environments (See **Productive learning environments**). Learning goals must include the development of metacognitive skills and robust mathematical beliefs as well as content knowledge and problem solving skills, and learning environments must be reconceptualized in ways that support the development of such understandings.

Productive learning environments addresses the nature of learning environments and their impact on mathematical thinking and learning. The key point, representing a change in the field’s understanding, is a change in focus from “teacher-centered” instruction to examining what instruction looks and feels like from each student’s perspective. Traditionally one would ask, “what kinds of teaching help students learn most effectively?” However, a main point of the section **Productive learning environments** is that learning is a sociocultural phenomenon—that what one learns (not only content, but processes and beliefs) is a function of interactions with others. Hence, while the teacher’s role remains central, equally critical issues concern the nature of the learning environment that the teacher creates, and how each student experiences it.

Once this shift in perspective is made, one can explore the attributes of learning environments that support productive engagement—not simply with regard to content learning, but with regard to the full range of learning described in **Mathematical thinking and learning**: the development of problem solving skills, powerful mathematical habits of mind, and productive mathematical orientations (beliefs and dispositions) including a sense of mathematical agency and positive mathematical identities.

Implications and prospects provides a brief discussion of the implications of the main ideas discussed in this entry and prospects for change.

Characterizing learning in general, and specifically in mathematics

This entry will use the following definition of learning:

Learning is a change in a person’s understanding or behavior that results from that person’s experiences.

Readers may be struck by the breadth of this characterization. This section explains why the definition is as encompassing as it is, and some of the consequences of having such a definition.

The phrase “that person’s experiences” encompasses the interpersonal aspects of interactions. This definition is inherently socio-cultural in character, for the reasons discussed below. There are other traditions, some of which are discussed in **Antecedents**. For an extensive review, see [Greeno et al. \(1996\)](#).

The term “understanding” refers to what an individual takes to be true, even if it is not true. This usage differs from everyday usage. The reason is that in any work that people do, what matters is what they take to be true. For example, the student who considers the incorrect equation $(a + b)^2 = a^2 + b^2$ to be correct will employ that understanding when solving algebraic problems.

Accordingly, the use of the term “change” (as opposed to “growth”) in the definition is consequential. In everyday usage learning is considered to be a gain in knowledge or a productive change in behavior, typically the result of deliberate instruction. This perspective reflects the traditional epistemological view of knowledge as fixed and immutable, and of learning processes as contributing to the growth of what an individual knows. Standard metaphors for this view include thinking of the learner as a blank slate to be written upon, or an empty jug to be filled with knowledge. In such a view, only the correct knowledge counts and the growth of knowledge is unidirectional.

However, as indicated by the algebra example given above, what an individual learns is not necessarily correct. Musicians or athletes may learn counterproductive techniques, and retraining to unlearn or overcome them can be a significant challenge. The same is the case with regard to learning a discipline. Extensive literatures on misconceptions in mathematics and science indicate that learners consistently develop misunderstandings that are, for at least some time, stable and predictable, for example, misunderstandings of the subtraction procedure that are so robust that researchers can predict the incorrect answers to problems that students will produce, before the students work the problems.

It is thus essential for any definition of learning to encompass the incorrect as well as correct understandings and behaviors that result from instruction. It is also essential to note that generalization is a natural human trait that underpins much learning. The

incorrect algebraic expression given above represents a natural, albeit incorrect, use of the distributive law, where the exponent 2 is “distributed” across the expression $(a + b)$. This kind of inappropriate generalization is a very common occurrence. Another is students’ thinking that “multiplication makes things bigger”—a truth when dealing with positive whole numbers that no longer holds when one is working with fractions.

These examples point to the fact that what is learned goes far beyond the intended content and process goals of instruction. Whether or not they are stated explicitly, students’ consistent mathematics-related experiences in classrooms (technically called sociomathematical norms) define for students what is considered legitimate mathematical behavior. Such understandings shape the ways in which students do mathematics. For example, students whose daily classroom experience consists of being shown how to solve certain kinds of tasks, after which they are given multiple practice exercises and subsequently tested on tasks that resemble the practice exercises, tend to develop the understanding that “all problems can be solved in just a few minutes.” The behavioral consequence of this learning is that many students stop working after just a few minutes when given richer and more substantial problems, when they might well solve them if they persevere. Similarly, students in a classroom where mathematical hypotheses or claims are expected to be accompanied by explanations of the type “I think/claim that this is true because ...” will come to incorporate the explanatory norms of doing mathematics as part of their understanding of the discipline. Even more broadly, there is the “hidden curriculum”: from their engagement in routine practices of schooling, students learn the (local, contextual) definitions of what it means to be a good student, how to behave appropriately, etc.

Some of what students learn in mathematics classes is deeply personal. Students are positioned not only by test results but by social interactions to see themselves as good or bad at mathematics, as fast or slow, to have particular strengths or weaknesses, and so on. These understandings about themselves have profound effects on the ways students engage in mathematics. Students’ mathematical identities—how they see themselves vis-à-vis mathematics—are fundamentally shaped by classroom experiences. This is consequential. Students are likely to engage in and with mathematics to the degree that they see themselves as mathematically proficient.

The mechanisms of learning are neither simple nor cumulative. People build complex networks of understandings, and unstable knowledge is often misremembered or forgotten. What a person currently understands shapes how new information is perceived. Information that fits into an individual’s knowledge network is more likely to be established and retained. Indeed, new and correct information may not take hold if enough connections to it are not built, and older, incorrect knowledge may regenerate. The more ways in which new information is connected to established knowledge, the more robustly it is retained. For example, research indicates that many students, when first studying the graphs of linear functions, develop the understanding that three pieces of information (the slope of the line, its x-intercept, and its y-intercept) are necessary to produce the graph of a line. This understanding persists for weeks and sometimes months. However, it tends to fade with experience. Ultimately almost all students come to understand the normative slope-intercept formula, in which only two pieces of information are necessary (Schoenfeld et al., 1993).

For all the preceding reasons, the definition of learning given above is necessarily broad. What must also be broad is the conception of mathematics to be learned. The question is, what does it mean to think mathematically? This encompasses far more than simply knowing established facts, procedures, and techniques. The goal of mathematics instruction is for students to become productive mathematical thinkers. That means having a solid knowledge base. In addition, it means having productive mathematical habits of mind—habits of inquiry and problem solving that enable the individual to make and verify mathematical connections. Being an effective mathematical thinker entails being metacognitive while solving problems—monitoring the process and making effective, reflective use of what one knows. It means having a positive mathematical identity, orientations, and a sense of agency—a willingness to engage with mathematical challenges in productive ways (See [Mathematical thinking and learning](#) for detail.). Given the fact that memory fades, a goal for mathematics learning cannot be that students will remember all the mathematics they have learned. Rather, it must be that students should remember and understand enough to regenerate the essential mathematics that they need to know and use.

Finally, it should be noted that mathematics learning is the result of activities inside and outside the classroom. Just as most students “know their ABCs” before entering the classroom, students develop some sense of number and geometry before they enter school. Similarly, some aspects of identity (e.g., “is mathematics for me?”) are shaped in significant ways by both in-school and out-of-school experiences.

Thus, an elaboration of the definition that began this section, with a focus on for mathematics learning is as follows:

Mathematics learning is change in a person’s understanding of mathematical content, practices, habits of mind, orientations and mathematical identity that result from that person’s classroom and other experiences.

The primary focus of this entry is on classroom learning, with an emphasis on the nature of learning environments from which students emerge as knowledgeable and resourceful mathematical thinkers and problem solvers.

Learning goals and learning outcomes

Nations across the globe employ varied mechanisms to specify the mathematics that students are expected to learn. Some educational systems are “top down.” For example, in the mid-to-late 20th century it was said, only partly in jest, that all 10th grade mathematics students in France could be found working the same problem at the same time. In Japan, curriculum specifications were developed at the national level, with a small number of textbook publishers developing and marketing curricula that complied with those specifications. In Germany, curricular goals were established by the sixteen German states. In the US, prior to the year 2000, a small number of the 50 independent states had legally mandated curricula, sets of accredited textbooks, and assessments. Across the US more than 15,000 local and regional school districts had the ostensible freedom to develop and implement curricula at will.

The phrase “ostensible freedom” is important because external constraints limited the degrees of freedom that school districts appeared to have. Expectations that students would enroll in calculus at the university level meant that there was a consistent “algebra to geometry to advanced algebra to trigonometry to pre-calculus” sequence across the US; consolidation in the textbook industry meant that school districts had a limited number of text options to choose from. (Elementary school math texts are typically sold in sets spanning the elementary grades, so that students in a district are committed to a particular instructional sequence. However, students who moved to other districts often found that their background did not mesh with their new district’s content sequence.)

Within the US, a strong trend toward homogenization began when the National Council of Teachers of Mathematics issued a volume of curriculum and evaluation standards in 1989. It continued through the issuance of Common Core State Standards (<http://www.corestandards.org/>) and aligned assessments in the 2010s. The standards and assessments were voluntary in that states could decide whether or not to adopt them; but with federal funding incentives for adoption, the year-by-year content sequences specified in the CCSS resulted in relatively consistent content specifications across the US.

Every nation has seen an evolution of curricular focus over time. Across the globe, curricula condensed over the course of the 20th century. In the US, Geometry, once a first-year college course, became a 10th grade course; calculus, once studied by advanced undergraduates, became a secondary school curricular option. Although there is some consistency in names and grade levels of mathematics courses across nations, there are curricular differences. In the Netherlands, for example, the curriculum is much more “applied” than in the US, a difference in emphasis reflected by the nature of the questions in two international mathematics examinations, PISA (Programme for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study). Nonetheless, the two examinations have a great deal in common, not only with one another but with almost all national exams and curricular standards documents. Universally, the general focus of mathematics curriculum and assessment is to determine what mathematics content a student has learned, with scant attention given to the practices and processes of doing mathematics.

As discussed in **Mathematical thinking and learning**, this focus on content means that a large part of mathematics proficiency—problem solving, metacognition, and students’ beliefs about self and the discipline—goes largely unaddressed in instruction. It is similarly unaddressed in testing, which is typically considered a measure of mathematics learning. Indeed, what has been called “mathematical literacy” (Burkhardt and Schoenfeld, this encyclopedia), the ability to use one’s mathematical knowledge to make consequential decisions in one’s daily life, gets very little meaningful attention. Before turning directly to those issues, however, we point to aspects of learning that are fundamentally unaddressed in curriculum documents. These pertain to the hidden curriculum and the role of mathematics in situating students, individually and collectively, in social and racial hierarchies.

The phrase “hidden curriculum” refers to the unwritten and often unacknowledged lessons conveyed to students by virtue of their participation in schooling. These lessons are often behavioral, concerning appropriate ways to comport oneself in school. Other lessons are learned as well. Practices such as “streaming” (alternatively, “setting” or “tracking”), which place students in different academic trajectories based in their perceived ability, send very strong messages about students’ perceived talents and futures. In the US, tracking correlates strongly with race—to the point where one research study (Shah, 2013) reports that an African-American student was asked “Oh you’re in this class? For real?” as he entered a high school precalculus class that was otherwise populated by whites and Asian-Americans. At the other end of the academic spectrum, remedial or low track mathematics classes tend to be populated predominantly by students of color. Being socially and academically positioned as not mathematically proficient shapes students’ identities and, in the case of tracking or streaming, can limit their potential learning opportunities in mathematics (see, e.g., Oakes, 2005).

Mathematical thinking and learning

Although mathematical thought has been the object of reflection by philosophers, mathematicians, and scientists for millennia and empirical studies of children engaged in mathematical tasks date to the early 20th century, mathematics education itself did not coalesce as a discipline until the 1960s. *Educational Studies in Mathematics* appeared in 1968, for example; *ZDM Mathematics Education* was first published (as *Zentralblatt für Didaktik der Mathematik*) in 1969; and the first issue of the *Journal for Research in Mathematics Education* appeared in 1970. Prior to that time, mathematics was the object of philosophical reflection and psychological experimentation, but learning to think mathematically was not an explicit focus of study.

Antecedents

Theories of mathematical thinking, teaching, and (tacitly) learning date to the ancient Greeks. The most notable example is the Meno dialog (Plato, 2005), in which Socrates uses a series of leading questions to demonstrate that a slave “knows” a special case of the Pythagorean theorem and that bringing that knowledge to mind is simply a matter of asking the right questions. This approach is now enshrined as the Socratic Method.

There is a long history of philosophical inquiry into the nature of thinking, but not of learning. Through history, mathematicians and scientists have reflected on their thinking (more than learning, as with philosophy), with little empirical substantiation. Thus, for example, Descartes wrote three volumes of *Rules for the direction of the mind*, on methods for solving problems; but these volumes were the product of introspection. The same is the case for Poincaré’s fabled description of his inspired discovery of the structure of Fuchsian functions. Poincaré described having worked deeply on the problem (the “saturation” phase of the Gestalt psychologists’ model of problem solving), taking a break from it (the incubation phase), having the answer come to him in a flash (the inspiration phase), and confirming the solution (verification). The challenge, however, was that such reports were anecdotal and not replicable. Other than providing vague guidance of the type, “work on the problem until you are completely immersed in it, then take a break, hoping for inspiration,” they hardly served as a basis for a theory of learning or instruction.

The Gestaltists were deeply concerned with mathematical structure, meaning, and understanding. Wertheimer’s (1945) critique of contemporary pedagogy emphasized that mathematics should make sense to students. He decried pure rote learning, noting, for example, that many students, trained by their experience, would laboriously calculate

$$\frac{274 + 274 + 274 + 274 + 274}{5}$$

without seeing that they were being asked to add a number five times and then divide by five, so the result had to be the number they started with. Wertheimer’s most famous example recounts his experience in a classroom where students were shown how to find the area of a parallelogram in “standard position,” as in Fig. 1A, but were nonplussed and considered it unfair when he asked them about the area of the identically shaped parallelogram in Fig. 1B; they had not been shown how to compute the areas of such figures.

Wertheimer’s (1945/1959) volume *Productive thinking* has many such examples, and suggestions for experiences students might have to help them develop deep understandings. As mathematically and pedagogically rich as it is, it has no theory of learning—in fact, the word “learning” is not in the index of the volume.

In sharp contrast, the behaviorists had both theories of learning and empirical data to back them up. The key idea was a theory of bonds, which characterized the strength of the relationship between stimuli and responses to them. Just as Pavlov’s dog learned to salivate at the sound of a bell when the bell ringing had been accompanied by Pavlov giving the dog treats, Thorndike’s law of effect posited that humans would learn to engage more frequently in behaviors that were associated with pleasure. Positive reinforcement would thus encourage particular behaviors, and negative reinforcement discourage others; in this way both non-human animals and people could be trained to perform complex acts. Thorndike, Skinner, and others applied the work done in animal training to human training, with educational theorists such as Gagné elaborating on how to break tasks into component parts that could be mastered through reinforcement.

Behaviorist theory captured aspects of skill acquisition well. When skills were carefully decomposed into learnable pieces and students were given appropriate practice and reinforcement, students learned them. However, the behaviorist approach is fundamentally limited. Behaviorists explicitly renounced “mentalism,” the idea that people have minds. The only things that counted as data in behaviorist theory were people’s objectively observable behaviors, the underlying theoretical idea being that all human actions can be explained without recourse to concepts such as “mind” or “thinking.” In consequence, important processes such as problem solving and creativity are not meaningfully encompassed within the theory. Behaviorism only addresses a small part of mathematical thinking and understanding, as elaborated in [Current understandings](#).



Fig. 1A A parallelogram in standard position.

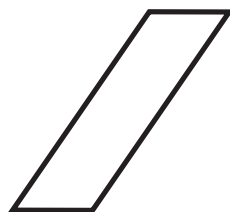


Fig. 1B An identically shaped parallelogram in non-standard position.

The contemporary counterpoint to behaviorism was constructivism. Piaget's theory of genetic epistemology—where “genetic” is the adjectival form of “genesis,” meaning developmental, so the theory is a developmental theory of knowledge and understanding—centered on the development of people's understanding, starting with their neonatal understandings and proceeding through the point where, in adulthood, people became formal thinkers. Foundational in Piaget's work was his focus on understanding the world as seen by the child. For example, in *The child's conception of number*, Piaget and Szeminska (1941/1952) showed that many children younger than 7 years old would, when asked “does this (top) row (of coins or checkers) have more circles, does this (bottom) row have more circles, or do they have the same number of circles?” would answer “the same” if the objects were placed in two identical arrays as in Fig. 2A; but if they watched Piaget spread the objects in the top row, resulting in Fig. 2B, they would say that the top row had “more.”

Being able to understand that the two rows of objects in Fig. 2B contain the same number of objects is known as “conservation of number.” Piaget's observational/experimental research documented that coming to understand conservation of number is a developmental phenomenon: young children do not have this understanding, but (somewhat homogeneously) they do develop it at some point. Conservation of number was one of hundreds of developing understandings elaborated by Piaget.

Most important in Piaget's theorizing are the notions of constructivism and of genetic epistemology. Constructivism posits that people do not perceive the world directly, but that they build interpretive filters through which they interpret the information conveyed by their senses. (At the most basic level, this is obvious; if it were not the case, optical illusions would be impossible.) Piaget's work on genetic epistemology indicated that there was a great deal of consistency in the ways that humans develop those understandings (knowledge schemas) over time. Moreover, his theorizing included mechanisms of change—the notions of equilibration (roughly, settling into a stable set of knowledge schemas) and disequilibration (a disturbance of equilibrium), which will either result in assimilation (the incorporation of new information into established knowledge schemas) or accommodation (restructuring current schemas so that they better fit the new information). These mechanisms offer one way of characterizing the underlying mechanisms of learning.

Piaget was explicit about the fact that his was a theory of development, not a theory of learning. He was cautious about prescribing instruction except at a general level (e.g., that the learning experiences people are given should fit with their level of development). Much of Piaget's theorizing has been problematized, for example, the homogeneity of development (he did not adequately take into account variation in experience) and a failure to address the socio-cultural dimensions of learning. For that one turns to Vygotsky.

Vygotsky (1978) specifically addressed learning, arguing in essence that learning (whether in formal or informal settings, often the latter) resulted from the internalization of originally external operations. His best known example is a description of how an infant's reaching unsuccessfully for an object ultimately evolved into the infant's pointing to the object, after that child's reaching attempts were interpreted by the parent as calls for that object. This notion of “interiorizing” external interactions into internal thoughts is central to a socio-cultural perspective on learning. A major contribution of Vygotsky's theory is the concept of the Zone of Proximal Development, or ZPD—the space of understandings that is beyond an individual's competence when acting alone, but within reach when acting with guidance from or in partnership with a more knowledgeable other. The implication for instruction is that learning is optimal when students are working within their ZPD.

Two final antecedents paved the way for the phenomenal growth of mathematics education since the 1960s. The first is George Pólya. Writing in the spirit of Descartes, Pólya—one of the preeminent mathematicians of the 20th century—reflected on his own problem solving and wrote a series of volumes elaborating on problem solving heuristics, strategies typically used by mathematicians for making progress on challenging problems. The titles of Pólya's major books on problem solving—*How to solve it*, *Mathematical discovery*, and *Patterns of plausible reasoning*—all point to the dynamic, process-oriented ways in which mathematics is done. Pólya's books provide detailed elaborations of strategies, and the problems in them are sequenced so that working through the problems engages readers deeply with mathematics. Pólya's oeuvre contributed immeasurably to the teaching and learning of mathematical thinking, but, being practically rather than theoretically oriented, did not approach learning as a scientific endeavor.



Fig. 2A Equally spaced objects.



Fig. 2B Unequally spaced objects.

The field that did focus on learning as both a theoretical and empirical enterprise was cognitive science (now the “learning sciences”), which emerged as a multidisciplinary amalgam of many fields concerned with learning—anthropology, artificial intelligence, computer science, education, linguistics, philosophy, psychology, and sociology among others. The mix of disciplines challenged established paradigms and methods, bringing multiple perspectives and techniques to the study of learning. The notion of human as “information processor” was introduced; scientific models of thinking and learning were developed; and “rich, thick” observations of classrooms and the ways they shaped learning complemented previous statistical methods of measuring learning. Concurrently, mathematics education coalesced to the point where it would study the specifics of mathematics teaching and learning. This was necessary for significant progress, in that it is essential to wed a deep understanding of the nature of the discipline to be learned with the rigorous study of thinking, learning, and learning environments.

Current understandings

This section characterizes the spectrum of mathematical understandings that constitute mathematical proficiency and, more broadly, learning outcomes from schooling. They are, in order: mathematical content; mathematical processes and practices; meta-cognitive behavior; and belief systems and orientations.

Mathematical content

General issues of content specification were discussed in [Learning goals and learning outcomes](#). Brief descriptions of many nations’ educational systems, as well as characterizations of mathematics standards, frameworks, and curricula can be found at <http://timssandpirls.bc.edu/timss2015/encyclopedia/countries/>.

Mathematical processes and practices

The essence of mathematical proficiency consists of using what one has learned in the service of thinking mathematically (alternatively, mathematical sense making). Thinking mathematically involves fundamental processes such as abstracting, generalizing, proving, reasoning, representing, modeling and mathematizing, and solving problems. These processes and practices (mathematical habits of mind) should be central objectives for mathematics learning ([McDougal, 2017](#); [Schoenfeld, 1985](#)).

This is a major arena in which mathematics curricula fall short. Once again, there is significant variation from nation to nation: one can say, for example, that Japanese curricula attend to issues of problem solving more than curricula in many other nations, and that curricula in the Netherlands emphasize mathematical modeling more than in many other nations (see, e.g., <http://timssandpirls.bc.edu/timss2015/encyclopedia/countries/japan/>). What follows is broadly the case, within such bounds of variation.

The two processes that have received the most curricular attention are problem solving and modeling.

There has, historically, been some confusion about problem solving. The broadly accepted definition of “problem,” for some decades, is that a problem is a task that a person wants to solve, and for which they do not have a solution method at hand. If a student has been shown how to solve tasks that are closely related in type to the task currently assigned, then that task is an exercise, not a problem. The goal of problem solving instruction is to prepare students to grapple successfully with tasks that they have not been shown how to solve. Research on problem solving was motivated by [Pólya \(1945/57\)](#). Building on these ideas, [Schoenfeld \(1985\)](#) and others elaborated on problem solving strategies to the point where it would be possible at all grade levels for students to learn problem solving strategies and, as is done in Japan, to develop their understandings of mathematics content by solving problems that engage the students with the nuances and complexities of the content.

This is not the current reality. In the US, the *Common Core State Standards for Mathematics* devote 87 pages to delineating content, but 3 pages to listing practices—of which one is problem solving. In Great Britain, a governmental research review document produced as this entry was being written (May 25, 2021 Research and analysis document, *Research Review Series: Mathematics—Gov.UK, 2021*) that was intended to serve as the underpinning for curriculum and assessment development used the term “problem” only in the sense of exercise, thus denying problem solving a place in the curriculum. More generally, testing and assessment focus on the content that is specified in the curriculum; under time pressure, teachers tend to focus on what is to be tested. As a result, students do not generally have opportunities to learn to become mathematical problem solvers. Students are unlikely to learn what they are not exposed to, and what is not valued (as indicated by its absence on tests.)

The absence in most nations of curricular support for mathematical modeling and mathematical literacy, the application of mathematical thinking to “real world” issues, is comparably problematic ([Burkhardt and Schoenfeld, 2022](#)). A purported virtue of mathematics is that students not only learn to think logically, but to use their mathematical understandings to make thoughtful decisions about issues they encounter in their lives outside of school. As in the case of problem solving, these skills are well understood: there is a well-documented “modeling cycle,” and empirical evidence that students can learn, when provided appropriate instruction, to mathematize—to represent “real world” phenomena and make mathematically sound judgments about them. As in the case of problem solving, students are not likely to learn what they are not taught.

The other productive mathematical habits of mind listed above have received less attention, especially at the K-12 level. Evidence indicates that grade-level-appropriate forms of reasoning and proving can be taught and learned from the early grades. However, such experiences are rare, with the result that learning to prove—whether in secondary school geometry classes or in undergraduate

courses requiring a facility with proof—is encountered by students as a major roadblock. Encountering such ideas late in the curriculum is like learning a second language without prior experience, as opposed to having picked up bits and pieces of it through the years. Opportunities for students to abstract and generalize are extremely rare.

Metacognitive behavior

Fig. 3 shows a timeline graph of two students working a mathematics task that was a problem in the sense described above: the task was unfamiliar, but it was known that they possessed the mathematical knowledge to solve it. The students read the problem, decided to try a particular solution approach, and pursued that approach until they ran out of time.

Over years and multiple studies, more than 50% of the timeline graphs of students working unfamiliar problems resembled Fig. 3. The key point is that, having chosen an inappropriate direction for their work, the students were guaranteed to fail to solve the problem unless they changed direction. The failure to take stock—to assess how successful they were being, and to consider alternatives—was a failure of what is called monitoring and self regulation, an aspect of metacognition.

Fig. 4 shows a timeline graph of two students working an unfamiliar problem, subsequent to their having taken a problem solving course that placed a significant emphasis on monitoring and self-regulation.

Each of the triangles in Fig. 4 represents a time when the problem solvers paused to assess how well their solution was going. Roughly 3 min into the solution, they questioned the efficacy of the approach they had taken, and they soon abandoned that approach. Roughly 6 and 7 min into the solution, they again took stock and decided to persevere; but about 11 min into their work they abandoned their second approach. There are never guarantees that one will find a solution, as they did, about 13 min into their attempt. However, what is guaranteed is that they would have failed to solve the problem had they not assessed and abandoned their first two attempts.

There is, then, clear evidence that monitoring and assessing—fundamentally important aspects of mathematical (and other) problem solving—can be learned. However, these skills go almost completely unaddressed in mathematics classrooms.

Orientations and identity

An individual's orientations (beliefs, dispositions, preferences, and values) are shaped both by personal experience and in what that person learns from or is told by others. In some countries, for example the US, people tend to hold the belief that mathematical talent is innate, that is, a person's level of mathematical skill and potential is fixed. In other countries, for example Japan, people tend to hold the belief that how well one does in mathematics depends on how hard one works at it—the harder one works, the more successful one will be. Such beliefs are consequential: those who believe they are untalented are much less likely to strive to do well than those who believe that hard work pays off. This is an example of how cultural understandings are internalized and shape personal behavior. In the US, cultural lore includes the widely held stereotypes that some Asians and Whites are good at mathematics, while people of Black and Latin heritage are not.

People's understandings regarding the nature of mathematics and their relationship to it (their mathematical identities) come in large measure from their personal experiences. In the case of mathematics, the primary source of such experiences is in the classroom. Patterns of classroom interactions over mathematics, with teachers and other students, become abstracted as individuals' belief systems. They shape the individuals' dispositions regarding mathematics, including their willingness to grapple with new challenges (their sense of mathematical agency), what they think matters in doing mathematics (e.g., how important is it to be able to explain why something is true?), and thus how they engage in mathematics.

In the US, for many years the predominant instructional mode was “demonstrate and practice”: the teacher showed students how to solve a particular kind of task and gave them closely related tasks to practice, both in class and as homework. Typically, such exercises were expected to be completed in just a few minutes. Exams often contained a large number of tasks that students were expected to know how to solve. (In one research study a 50 min test contained 20 tasks; shortly before the exam the teacher told the students that they needed to have practiced their work carefully, “because you won't have time to think.”) The Graduate

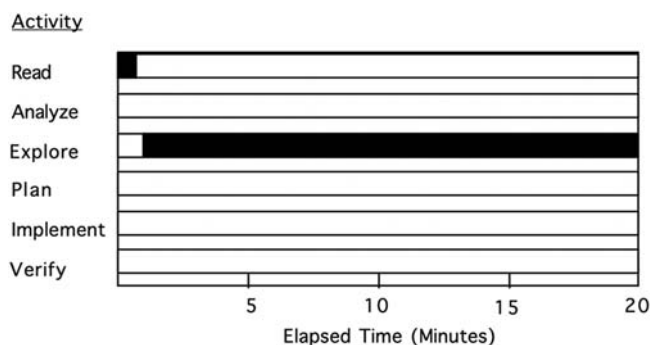


Fig. 3 A timeline graph of a typical student attempt to solve a non-standard problem. Reproduced with permission from Schoenfeld (1985).

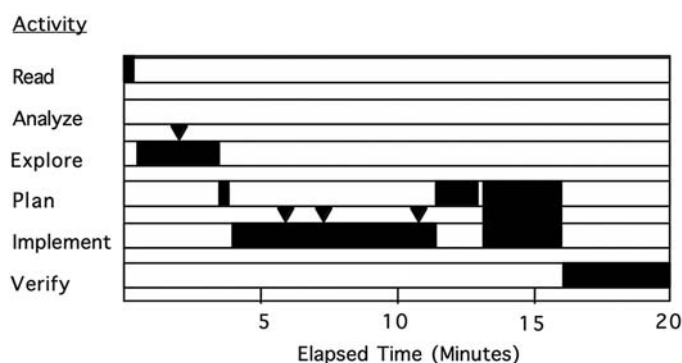


Fig. 4 A timeline graph of two students working a problem after having finished a course in problem solving. Reproduced with permission from Schoenfeld (1985).

Record Exam in mathematics, taken by university students who hoped to enroll in graduate school, contained 65 “problems” to be worked on for at most 3 h. Years of such experience led students to develop expectations and understandings such as the following: mathematics is learned by internalizing the rules and procedures given to you by experts; problems will either be solvable in a few minutes or, if you have forgotten the method, not at all; speed matters; your innate talent is revealed by how well you do on tests and how you are seen (and positioned) by others. In contrast, there are classrooms where the sociomathematical norms differ from these—students are given opportunities to engage in problem solving, and to work collaboratively in generating and testing ideas and conjectures, and to provide explanations of their reasoning. Here students tend to develop much more productive beliefs about how one goes about doing mathematics.

People’s mathematical identities are shaped by cultural assumptions. These include: assumptions about who is likely to be good at mathematics, whether or not mathematical ability is fixed, and what math is and how useful it is for one’s future. Similarly, identities are shaped by the ways that schools position students via sorting mechanisms such as tracking or streaming; by classroom discourse, in which students are positioned by the teacher and other students, either explicitly or tacitly (under what circumstances does the teacher call on them, or fellow students ask for or listen to their contributions?); and by performance measures such as tests. As noted at the beginning of this subsection, people’s mathematical identities are consequential. People who see themselves as effective mathematical thinkers and problem solvers are more likely to engage with mathematics tasks and persevere at them than those who do not.

Productive learning environments

The preceding sections discussed key aspects of mathematical thinking and learning, and some of the ways in which schooling and the larger social environment shape learning outcomes. In considering learning and learning environments, it is essential to consider the complete spectrum of outcomes, ranging from mathematical content, practices, and habits of mind to belief systems, orientations, and the shaping of individuals’ mathematical identities. This section discusses a comprehensive framework for characterizing learning environments regarding their potential to support mathematics learning, the Teaching for Robust Understanding (TRU) framework (Schoenfeld and the Teaching for Robust Understanding Project, 2016; <http://truframework.org>).

A distillation of the framework is given in Fig. 5, which outlines five key dimensions of all learning environments, including mathematics classrooms. Research indicates that student outcomes correspond to the quality of the classroom environment, when characterized according to the framework. That is, the richer the learning environment is along the five dimensions, the more likely it is that students will emerge from instruction as knowledgeable and resourceful mathematical thinkers and problem solvers. As discussed below, this entails the shaping of metacognitive behavior, belief systems, orientations, and mathematical identities as well. The five dimensions of the framework are discussed in **Dimension 1, classroom affordances for mathematics learning** through **Dimension 5, formative assessment. Implications and Prospects** offers a brief summing-up.

Dimension 1, classroom affordances for mathematics learning

Dimension 1 focuses on the opportunities students have to develop the deep sense of mathematics described in **Mathematical thinking and learning**. This includes engagement with rich mathematical content—content organized in ways that, in the words of the Common Core, is focused and coherent. Students should be presented with opportunities to make connections, learn and use

The Five Dimensions of Powerful Mathematics Classrooms				
The Mathematics	Cognitive Demand	Equitable Access to Mathematics	Agency, Ownership, and Identity	Formative Assessment
<i>The extent to which classroom activity structures provide opportunities for students to become knowledgeable, flexible, and resourceful mathematical thinkers. Discussions are focused and coherent, providing opportunities to learn mathematical ideas, techniques, and perspectives, make connections, and develop productive mathematical habits of mind.</i>	<i>The extent to which students have opportunities to grapple with and make sense of important mathematical ideas and their use. Students learn best when they are challenged in ways that provide room and support for growth, with task difficulty ranging from moderate to demanding. The level of challenge should be conducive to what has been called “productive struggle.”</i>	<i>The extent to which classroom activity structures invite and support the active engagement of all of the students in the classroom with the core mathematical content being addressed by the class. Classrooms in which a small number of students get most of the “air time” are not equitable, no matter how rich the content: all students need to be involved in meaningful ways.</i>	<i>The extent to which students are provided opportunities to “walk the walk and talk the talk” – to contribute to conversations about mathematical ideas, to build on others’ ideas and have others build on theirs – in ways that contribute to their development of agency (the willingness to engage), their ownership over the content, and the development of positive identities as thinkers and learners.</i>	<i>The extent to which classroom activities elicit student thinking and subsequent interactions respond to those ideas, building on productive beginnings and addressing emerging misunderstandings. Powerful instruction “meets students where they are” and gives them opportunities to deepen their understandings.</i>

Fig. 5 The teaching for robust understanding framework. Reproduced with permission from Schoenfeld and the Teaching for Robust Understanding Project (2016).

multiple representations, etc. They should have opportunities to engage in mathematics in ways that support the development of powerful mathematical habits of mind, e.g., engaging with problems and learning problem solving techniques, in ways that support the development of metacognitive awareness and performance; explicitly focusing on modeling and issues of mathematical literacy; and being supported in age-appropriate engagement with abstraction, generalization, and producing coherent chains of reasoning, argumentation, and proof.

In the absence of such opportunities for engagement, students are unlikely to develop the relevant understandings. However, opportunities for engagement are hardly enough to result in powerful learning. The nature of the engagement (See **Dimension 2, cognitive demand and the accessibility of ideas** through **Dimension 5, formative assessment** of the framework) is critically important.

Dimension 2, cognitive demand and the accessibility of ideas

When students are given work that they find too easy, there is little “cognitive demand” and little room for them to grow. On the other hand, mathematically rich presentations or task sets will do learners little good if the students do not find the ideas in them to be within reach. The instructional challenge is to arrange instruction so that the level of challenge supports productive struggle. More formally, the idea is for students to be working within their zones of proximal development, being able to build on what they know.

This is a significant instructional challenge. Mechanisms for easing the challenge include the use of problems that have multiple entry points and a range of possible solution methods. Individual students have a greater likelihood of gaining a foothold on such problems, meaning their engagement is more likely to be productive. Moreover, challenging problems with such properties can be profitably used for group work. When students try to understand and support each other’s approaches, they engage more deeply with the material. Comparing and contrasting approaches taken by different students leads to making mathematical connections, and student-to-student conversations provide opportunities for students to produce explanations. This does not diminish the teacher’s role in instruction but expands it. More students are involved in generating and reacting to ideas, giving the teacher more “reach.” (See **Dimension 5, formative assessment**).

Dimension 3, equitable access to key ideas for all students

A central goal for powerful instruction is that each and every student has the opportunity to engage meaningfully and powerfully with important ideas. There may be rich mathematical conversations taking place in a classroom, but if some students are excluded from them, the instruction cannot be considered powerful. The classic report *How schools shortchange girls* (American Association of University Women, 1992) documented, for example, that in STEM classrooms girls were called on less frequently, and that when

they were called on, the kinds of questions they were asked tended to be lower level than the questions posed to boys. There are comparable issues regarding all demographic groups. Although there are no surefire approaches to equitable participation, a range of techniques can be used to enfranchise increasing numbers of students. These include the choice of problems, as described in **Dimension 2, cognitive demand and the accessibility of ideas**; creating a context where students feel safe and ideas are refined and built upon rather than simply being critiqued or declared wrong; “revoicing” and supporting the use and refinement of various spoken registers, from everyday to the formal mathematics register; similarly supporting second language use; using tasks that may connect to and build on students’ lives; and carefully employing randomizing devices or “equity tools” that support equitable participation. The social dimensions of equitable participation are addressed in **Dimension 4, opportunities for students to develop productive belief systems, a sense of mathematical agency, and positive mathematical identities**.

Dimension 4, opportunities for students to develop productive belief systems, a sense of mathematical agency, and positive mathematical identities

The curriculum and its implementation play large roles in shaping students’ beliefs regarding the nature of mathematics. Students who experience mathematics in a purely procedural way are likely to develop the understanding that mathematics is dull and mechanical, while students who get to explore and engage in sensemaking will come to see the discipline in rather different ways. But the curriculum is only a small part of what shapes identity and beliefs about oneself. The issue is, what opportunities do students have to engage in mathematical activity, and how are their efforts greeted in the classroom? Can students relate the curriculum to their own lives? Do they have opportunities to venture ideas? Are those ideas taken up and built upon, and do the students have opportunities to build similarly on others’ ideas? If students are grappling with formal mathematical language, can they express themselves in ways that feel comfortable? Are their ideas revoiced in ways that they see themselves as contributors while at the same time their capacity to produce mathematical arguments, and to use formal mathematical language, is being enhanced?

All classrooms are fundamentally social environments. Whom the teacher calls on, in what ways, and how students in small groups position each other, are the mirrors through which students see themselves reflected. Those who are treated as meaningful contributors to the class’s mathematical work are much more likely to see themselves in a mathematically positive light than those who do not. They are more likely to develop productive mathematical dispositions, including a sense of agency (the willingness to engage with mathematical tasks and to contribute to mathematical discussions), and to develop positive mathematical identities.

Dimension 5, formative assessment

Learning is optimized when activities proceed well along the first four dimensions of the framework: the mathematics in which students engage is rich and powerful; students are engaged in productive struggle; each and every student is engaged in meaningful ways with core content and practices; and engagement with the mathematics, the teacher, and other students provides opportunities for the development of productive beliefs about the discipline and oneself. Working toward these ends demands constant adjustments on the part of the teacher. If a group finds a task too easy, is there an extension or an ancillary task that will challenge them productively? If they are stuck, is there a way of providing scaffolding that enables the students to make progress but does not make things so easy that there is no longer productive struggle? What can be done to re-engage students who have disengaged? What sociomathematical norms are established for productive engagement, and when do they need repair (e.g., when a student’s contributions are rejected or demeaned), and in what ways?

Making such judgments during instruction and modifying tasks or classroom dynamics accordingly is called formative assessment. Formative assessment calls for making student thinking public; it needs to be seen and heard if adjustments are to be made. It calls for discretion in judgment—how long does one let students work unproductively (recovering from mistakes is a good learning experience), and how and when does one intervene, to adjust the level of challenge or the social dynamics? Research indicates that it is not easy for teachers to learn to use the techniques of formative assessment, but that when it is used, student learning increases dramatically (Black and Wiliam, 1998). There do exist support tools for teachers. For example the Formative Assessment Lessons downloadable from the Mathematics Assessment Project website at <https://www.map.mathshell.org/> provide extensive guidance for teachers including detailed lesson plans, lists of misconceptions and other issues students are likely to run into when working the lessons, and suggestions for questions and activities when they do. One should not underestimate the challenges teachers face when they try to move from the more traditional form of presenting students with information to the more open form of supporting students when they are grappling with challenging problems. The research indicates that classroom management can be a challenge at first, and monitoring students can feel overwhelming. It also indicates, however, that teachers can settle into new pedagogical habits, with significant learning gains for their students.

Implications and prospects

The past half-century has witnessed tremendous leaps in researchers’ understandings of mathematical thinking, teaching, and learning. The full spectrum of mathematical understandings (content, practices, and habits of mind; metacognitive behavior; beliefs about mathematics and self, including issues of identity) has been identified, as have the dimensions of learning environments that

support mathematical learning along those lines. Unfortunately, instructional practice lags significantly behind these understandings. While some progress has been made, curricula focus for the most part on content and a small subset of practices: some problem solving, some reasoning, and some collaborative group work. However, many central mathematical practices including abstraction, generalization, and effective metacognitive behavior go largely unaddressed, despite evidence that they can be learned. Assessments tend to reify this problem, in that they too focus largely on content. Structurally problematic practices such as tracking and streaming perpetuate structural inequalities, as do the most subtle but equally pernicious ways in which students are positioned by others, with negative impact on identity and potential opportunity.

Addressing such issues is partly a matter of research, and partly a matter of political will and fiscal investment. Researchers and developers need to work on deeper understandings and more useable tools for teachers—the five dimensions outlined in **Productive learning environments** have yet to be fleshed out in adequate detail. But, enough is known about the issues in this article to provide teachers with much more instructional support than is currently the case. Administrators need to understand these issues, and to implement support structures that help teachers work toward these goals. These are all significant challenges. Notable progress has been made, however, and one hopes to see it continue.

Acknowledgments

The author gratefully acknowledges substantive and editorial suggestions from Abraham Arcavi, Hugh Burkhardt, Heather Fink and members of the Functions research group at the University of California, Berkeley.

References

- American Association of University Women, 1992. *How Schools Shortchange Girls*. AAUW and NEA, Washington, DC.
- Black, P.J., William, D., 1998. Assessment and classroom learning. *Assess. Educ.* 5, 7–74.
- Burkhardt, H., Schoenfeld, A.H., 2022. Assessment and mathematical literacy: a brief introduction, fourth ed. *International Encyclopedia of Education*. Gov.UK, 2021. Research Review Series: Mathematics. <https://www.gov.uk/government/publications/research-review-series-mathematics>.
- Greeno, J., Collins, A., Resnick, L., 1996. Cognition and learning. In: Berliner, D., Calfee, R. (Eds.), *Handbook of Educational Psychology*. Simon & Schuster McMillan, New York, pp. 15–46.
- McDougal, T. (Ed.), 2017. *Essential Mathematics for the Next Generation: What and How Students Should Learn*. Tokyo Gakuhei University, Tokyo, Japan, pp. 104–129.
- National Council of Teachers of Mathematics, 1989. *Curriculum and Evaluation Standards for School Mathematics*. NCTM, Reston, VA.
- Oakes, J., 2005. *Keeping Track: How Schools Structure Inequality*. Yale University Press, New Haven, CT.
- Piaget, J., Szeminska, A., 1941/1952. *The Child's Conception of Number*. Routledge & Kegan Paul, London.
- Plato, 2005. *Meno and Other Dialogues*. Trans. In: Waterfield, R. (Ed.). Oxford University Press, Oxford.
- Pólya, G., 1945. *How to Solve It*, 1957, second ed. Princeton University Press, Princeton.
- Schoenfeld, A.H., The Teaching for Robust Understanding Project, 2016. *An Introduction to the Teaching for Robust Understanding (TRU) Framework*. Graduate School of Education, Berkeley, CA. Retrieved from: <http://truframework.org>.
- Schoenfeld, A.H., Smith, J.P., Arcavi, A.A., 1993. Learning. In: Glaser, R. (Ed.), *Advances in Instructional Psychology*, vol. 4, pp. 55–175.
- Schoenfeld, A.H., 1985. *Mathematical Problem Solving*. Academic Press, Orlando, FL.
- Shah, N., 2013. *Racial Discourse in Mathematics and Its Impact on Student Learning, Identity, and Participation*. Ph.D. Dissertation. University of California, Berkeley, p. 3593980. ProQuest Dissertations Publishing.
- Vygotsky, L., 1978. *Mind in Society*. Cambridge University Press, Cambridge.

Websites of interest

- <http://www.corestandards.org/>, The Common Core Standards.
- www.map.mathshell.org/, The Mathematics Assessment Project Website.
- <https://parenting.kars4kids.org/conservation-tasks-what-piaget-taught-me-about-children/>, A Number Conservation Video.
- <http://timssandpirls.bc.edu/timss2015/encyclopedia/countries/japan/>, An Overview of the Japanese Educational System.
- <http://truframework.org>, The Teaching for Robust Understanding Framework.

Learning to read and write

Young-Suk Grace Kim, University of California Irvine, Irvine, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	335
Knowledge and skills that influence reading development	335
Knowledge and skills that influence writing development	336
Reading and writing relations	337
Implications	340
Acknowledgments	341
References	341

Introduction

The ability to read and write is the backbone of daily functions and learning in contemporary society. Reading and writing involve highly complex workings of the human mind (Huey, 1968) and develop through interactions between individual characteristics (e.g., language and cognitive skills and knowledge) and multiple layers of environment such as the language and literacy environment in the home, formal instruction, and larger systems (e.g., state or national policy; Joshi and Aaron, 2012; Kim, 2020a). In this chapter, I review the knowledge and skills that influence development of reading and writing skills, respectively. This is followed by a review of reading-writing relations and the nature of structural relations among contributing skills, reading, and writing skills, such as hierarchical, interactive, and dynamic relations. Practical implications are also discussed.

Knowledge and skills that influence reading development

Reading comprehension involves visually perceiving letters and words, converting them into speech sounds (decoding), parsing linguistic information, developing initial understanding of propositions based on the given information, and integrating propositions for an accurate and coherent mental representation (see Kintsch, 1988; Snow, 2010 for different levels of mental representations and comprehension). Decoding or word reading skills are absolutely necessary for reading comprehension as reading involves written texts. Word reading processes draw on the knowledge of orthographic symbols (orthographic awareness; e.g., knowledge of alphabet letters and their sounds), speech sounds represented in words (phonological awareness), and morphological structure (morphological awareness; Foorman et al., 2016; National Institute of Child Health and Human Development [NICHD], 2000; Snow et al., 1998).

Writing systems represent oral language in various ways where orthographic symbols represent either speech sounds or meaning (see Kim, 2022a). The vast majority of the writing systems in the world employ an alphabetic writing system where orthographic symbols represent sounds. English is an example of an alphabetic writing system. In English, the Roman alphabet letters are used as orthographic symbols that represent the smallest unit of speech sounds, phonemes. For example, the word *dog* has three letters, *d*, *o*, and *g*, each of which represents associated sounds /d/ /ɔ/ and /g/. The three letters, *dog*, are not a single unit that represents the syllable of /dɔg/ or the meaning of a dog. Various orthographic symbols are used in alphabetic writing systems (e.g., Hangul in Korean, Cyrillic in Russia and other countries), but the principle is the same. Orthographic symbols do not always represent phonemes. In Japanese Hiragana and Fidel in Ethiopia, orthographic symbols represent syllables. For example, < in Hiragana represents the /ku/ syllable, and is not a combination of two parts that respectively represent /k/ and /u/ phonemes. In contrast to alphabetic writing systems, Chinese has a meaning-based writing system, called a morphosyllabic writing system where each character represents a meaning unit, morpheme, and a sound unit, syllable, simultaneously (see DeFrancis, 1984 for details about the Chinese writing system).

In languages with alphabetic writing systems, learners need to know orthographic symbols and their associated sounds (orthographic awareness; e.g., letter *d* represents /d/). Learners also need to develop an understanding of the sound structure of a language, phonological awareness, because sounds map onto graphemes (letters or groups of letters). Essentially, learners need to develop an understanding of the alphabetic principle (orthographic symbols represent speech sounds) and grapheme-sound correspondences (mapping of letters or combinations of letters to sounds). Learners also need to understand morphemes and their representation in spelling. Morphemes are the smallest unit of meaning, such as plural *-s* (e.g., *dogs*) and affixes (e.g., *un-* [*undo*]). In many alphabetic writing systems, spellings of words do not represent speech sounds alone, but represent morphological information as well (this is called a morphophonological writing system). For instance, the word *sign* has *g* although it is not pronounced because the word *sign* is morphologically related to *signal*. Meaning information, in addition to phonological information, is preserved in word's spelling. Overlay of meaning on speech sound information frequently causes mismatch between

word's spelling and speech sounds, creating difficulty in decoding and spelling. For example, the words *electric* and *electricity* retain the same spelling *c* despite changes in the sound (/k/ and /s/). Therefore, an understanding of morphological composition of words, morphological awareness, helps word reading and spelling especially for multimorphemic words (Joshi et al., 2008–2009; Kim et al., 2013).

The consistency between graphemes and sounds influences the learning rate of word reading. Word reading acquisition is faster in languages with consistent mappings of graphemes and sounds (transparent orthography; e.g., Finnish, Italian, Spanish) than in languages with inconsistent mappings (opaque orthography; e.g., English). For instance, students learning to read in transparent orthographies reach over 90% accuracy of word reading after 1 year of reading instruction whereas students learning to read in opaque orthographies take 2.5 times longer to reach a similar level of proficiency (Seymour et al., 2003). Writing systems also influence the extent to which word reading and spelling draw on phonology, orthography, and morphology (Katz and Frost, 1992; Kim, 2020a; McBride-Chang et al., 2008; Ziegler and Goswami, 2005). In transparent orthographies, orthographic awareness and phonological awareness are key determiners of word reading and spelling skills whereas in opaque orthographies, morphological awareness plays an important role (e.g., McBride-Chang et al., 2008). Furthermore, in orthographies with high visual complexity (e.g., Chinese), visual processing is of great importance (Zhou et al., 2014).

During the initial phase of word reading development, decoding is a slow process of mapping graphemes to sounds and assembling sounds to pronounce a word. With practice, word reading becomes accurate and rapid, i.e., fluent word reading. Fluent word reading is instantaneous recognition of words as a whole without a slow decomposition process (recognizing a single word instantaneously; Ehri, 2005). It is also important to develop fluent reading of connected texts. This is referred to as text reading fluency or oral reading fluency—the skill to read connected texts accurately and rapidly with expression (NICHD, 2000). Text reading fluency facilitates reading comprehension as the reader is no longer limited or constrained by word reading and thus can focus on understanding the meaning of the text (Jenkins et al., 2003; Kim, 2015a; LaBerge and Samuels, 1974).

While word reading skill is necessary for reading comprehension, it is certainly not sufficient. Comprehension is a highly complex construct and requires an array of knowledge and skills: knowledge of the meanings of individual words (vocabulary) and chunks of words (syntactic knowledge); the ability to reason and infer unstated information (inference), monitor one's own understanding (comprehension monitoring), and employ appropriate comprehension strategies; knowledge of a given topic (topic knowledge); knowledge of discourse features (e.g., text structure); and domain-general cognitive abilities such as working memory and attentional control (see Cain and Oakhill, 2007; Kim, 2020a,b). Reading skills also interact with social-emotional aspects toward reading (e.g., attitude, motivation, anxiety; Kim, 2020a,b), and text features (Francis et al., 2018; Kim, 2020a,b). Reasoning and inference are broad constructs that include multiple aspects and types. For example, inference includes inferring spatial relations among entities, referents of nouns and pronouns, causal structure of propositions, and understanding perspectives (characters', authors', and the reader's own perspective; Graesser et al., 1994; Kim, 2020a,b; Shanahan et al., 2010). Texts in general do not explicitly state all necessary information for comprehension, and therefore, inferring unstated information is part and parcel of establishing coherent understanding of text.

A vital point that is often not well recognized is that comprehension skills initially develop in the oral language context (listening comprehension) and continue to develop in the written language context (reading comprehension). Listening comprehension is a skill to comprehend oral texts, such as stories, multi-utterance conversations, and oral informational texts (e.g., watching or listening to a segment on life cycles of an animal; Kim, 2015b, 2016; Kim and Pilcher, 2016). The above noted skills and knowledge such as vocabulary, background knowledge, and inference are all utilized in listening comprehension (e.g., Florit et al., 2014; Kendeou et al., 2008; Kim, 2016; Kim and Phillips, 2014; Tompkins et al., 2013) and continue to be used in reading comprehension (e.g., Cain and Oakhill, 2007). Put simply, reading comprehension is listening comprehension with additional word reading processes (Gough and Tunmer, 1986), and word reading and listening comprehension, respectively, involve multiple skills and knowledge described above (see the direct and indirect effects model of reading; Kim, 2017, 2020a,b).

Word reading and its component skills (orthographic, phonological, and morphological knowledge) have a large influence on reading comprehension in the beginning phase of reading development whereas listening comprehension and its associated skills and knowledge (e.g., background knowledge, vocabulary, inference) exert increasing influence as word reading skill develops (Kim, 2020a,b). For example, on average, in lower elementary or primary grades, one's word reading skill primarily determines their reading comprehension (Kim, 2015a; Kim and Wagner, 2015). In contrast, by middle school, listening comprehension skill primarily determines reading comprehension performance for the majority of students (e.g., Adlof et al., 2006). Regardless of grade level, however, for students who struggle with word reading, their inadequate word reading skill limits reading comprehension.

Knowledge and skills that influence writing development

Writing requires something to write about, i.e., idea generation. Generated ideas need to be translated or expressed into words and sentences. Translated ideas then need to be converted to printed words and sentences (encoding or transcription). Drafts are revised for expression, cohesion, and coherence, and edited for writing conventions (Hayes and Flower, 1980). These writing

processes draw on a number of skills and knowledge. Idea generation requires world, topic or content knowledge or reading skills. Translation into words and sentences draws on oral language skills. Appropriate and precise use of words and sentence structures is necessary for conveying thoughts accurately, and this requires oral language skills such as vocabulary, morphosyntactic, and syntactic knowledge (assembling words to construct phrases and sentences). The encoding process requires transcription skills such as spelling and handwriting or keyboarding, which place a large constraint on the writing process (Abbott and Berninger, 1993; Graham et al., 1997; Santangelo and Graham, 2016). Spelling draws on the same skills as does word reading—that is, orthographic awareness, phonological awareness, and morphological awareness (Apel et al., 2012; Kim, 2020a; Kim et al., 2013; NICHD, 2000; Snow et al., 1998). Handwriting or keyboarding requires motor skills and orthographic awareness (Berninger et al., 1992).

A quality written composition is more than an assembly of accurately encoded words and sentences. Sentences and ideas should be semantically related and be arranged in a logical order, and the text as a whole should be globally coherent and aligned with the writer's intended meaning. The writer needs to decide what information to present explicitly versus suggest implicitly, arrange sentences for a logical flow of ideas, and determine how to express and present ideas for the presumed audience and their needs. The writer also needs to monitor one's own writing process and be able to employ necessary strategies (e.g., using graphic organizers for planning, or using a writing checklist; Graham et al., 2012, 2016; Kim, 2020c). All these require higher order cognitive skills, such as reasoning, inferencing, and perspective taking (Kim, 2020c; Kim and Park, 2019). In addition, the writer needs to know and employ genre conventions such as text structure and associated linguistic features. Narrative texts typically include characters, a setting, a series of events, problems, and resolutions. Informational texts have different subtypes such as persuasion, opinion, compare-contrast, and description, and each has its own prototypical structure. Therefore, the writer needs to select the appropriate text structure for a given goal using their discourse knowledge to effectively convey ideas (Kim, 2020c; Kim and Park, 2019) because readers find it difficult to comprehend texts that do not follow conventional structures (Meyer, 1975). The writing process and the component skills that contribute to the writing process also require domain-general cognitive skills such as working memory and attentional control (Kellogg, 1996; Kim and Park, 2019; Kim and Schatschneider, 2017). Furthermore, writing skills interact with socio-emotions toward writing (e.g., motivation, attitude, interest, anxiety, self-concept; Graham et al., 2007).

The relations of component skills to writing change depending on the phase of writing development. In the beginning phase of writing development, transcription skills (spelling, handwriting, keyboarding) severely constrain the writing process, whereas other skills and knowledge (background knowledge, higher order cognitions, oral language skills) increasingly influence written composition as transcription skills develop (Kim and Graham, 2022; Kim and Park, 2019). Furthermore, writing develops from knowledge-telling to knowledge-transforming to knowledge-crafting phases (Bereiter and Scardamalia, 1987; Kellogg, 2008). In the beginning phase of writing development, writers tell what they know (hence, knowledge-telling). Writing during this phase is essentially a process of retrieving knowledge from long-term memory and transcribing it into a written product. In the knowledge-transforming phase, writing reflects the writer's thought process and "becomes a way of actively constituting knowledge representations in long-term memory" (Kellogg, 2008, p. 7). The knowledge-crafting phase is when professional expertise in writing is achieved. In this phase, the writer coordinates author, text, and reader representations and crafts writing with the potential reader in mind.

Reading and writing relations

A robust body of literature indicates that reading and writing are not independent, but instead they are related. A recent meta-analysis found that word reading and spelling are strongly related ($r = 0.82$) and that reading comprehension and written composition are moderately related ($r = 0.48$; Kim et al., 2022). One potential reason for reading-writing relations is that many tasks and activities require use of reading and writing skills simultaneously (e.g., writing a book report, writing a grocery list and reading it), and therefore, reading and writing skills are functionally related (Langer and Applebee, 1987; Shanahan, 2016). Another explanation for reading-writing relations is that reading and writing draw on shared skills and knowledge. In other words, highly similar skills and knowledge underlie reading and writing, and consequently reading and writing are related. A heuristic representation of this idea is illustrated in Fig. 1: Reading and writing skills may appear distinct at the surface; however, they are products supported by shared underlying language and cognitive skills that may not be readily visible (Kim, 2020d).

The idea that reading and writing are built on highly similar skills and knowledge was described in the shared knowledge hypothesis (Fitzgerald and Shanahan, 2000). According to Fitzgerald and Shanahan (2000), shared knowledge between reading and writing include (a) metaknowledge about functions and purposes of reading and writing, monitoring one's own meaning-making and word identification or production strategies, and monitoring one's own knowledge; (b) domain knowledge about substance and content, which includes vocabulary and meaning created through context of connected text; (c) knowledge about universal text attributes, including graphophonics (phonological awareness, grapheme awareness, and morphology), syntax (syntax of sentences and punctuation), text format (text structure, and organization); and (d) procedural knowledge and skill to negotiate reading and writing, such as knowing how to access, use, and generate knowledge, and smoothly instantiate various processes.

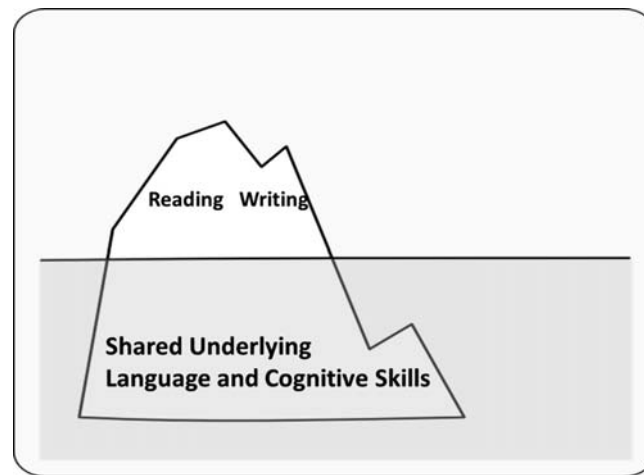


Fig. 1 A heuristic for reading-writing relations and their shared language and cognitive skills. Reproduced from Kim (2020d).

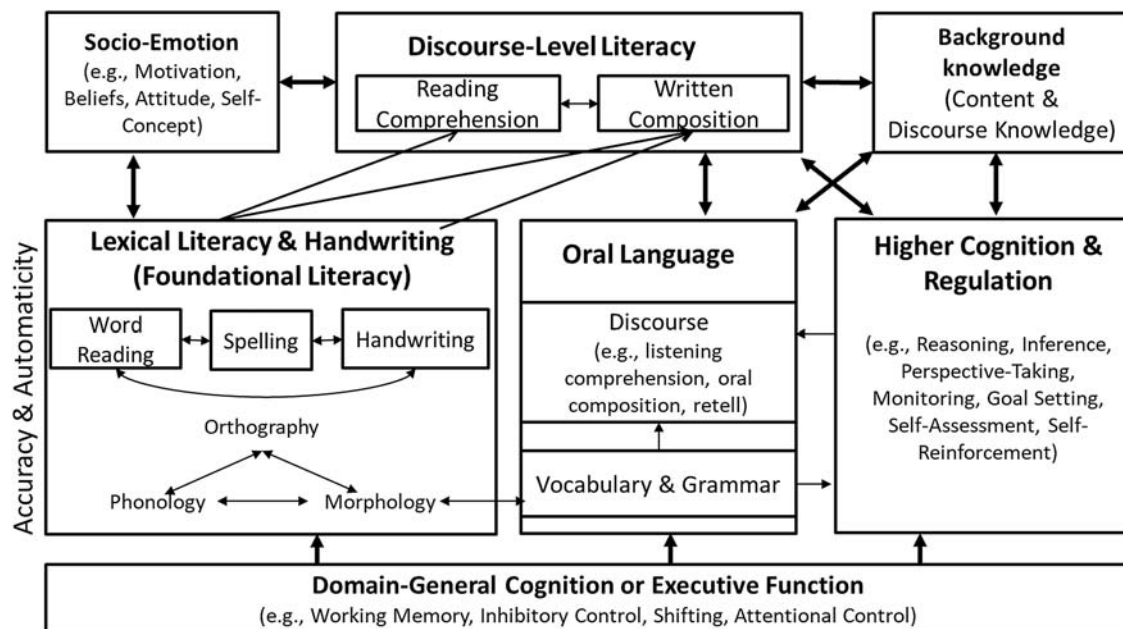


Fig. 2 Interactive dynamic literacy model. Reproduced from Kim (2020d).

The shared knowledge hypothesis was recently expanded in the interactive dynamic literacy model (Kim, 2020d). Fig. 2 shows details on the shared language and cognitive skills and knowledge as well as the nature of their relations according to the interactive dynamic literacy model (Kim, 2020d). Reading comprehension and written composition are both discourse-level literacy skills, built on lexical literacy skills (word reading and spelling) and handwriting/keyboarding, oral language skills (vocabulary, grammatical knowledge, and discourse language skills), higher order cognition and regulation, and background knowledge. All these are built on domain-general cognition or executive function. Reading and writing skills also interact with socio-emotions toward reading and writing.

The interactive dynamic literacy model has four hypotheses regarding the nature of relations. The first hypothesis is that the skills and knowledge that contribute to reading and writing skills (component skills) have hierarchical relations. According to this hypothesis, the component skills do not have a flat structure where every skill and knowledge is directly related to reading comprehension and written composition. Instead, there are multiple layers of higher order and lower order skills, and lower order skills support higher order skills in a systematic and cascaded way, resulting in multiple chains of relations. For example, domain-general cognition supports phonological, orthographic, and morphological awareness, which, in turn, support lexical literacy skills (word reading and spelling) and handwriting/keyboarding skills, which support reading comprehension and written composition. Domain-general cognition also supports vocabulary and grammatical knowledge (morphosyntactic and syntactic knowledge),

which, in turn, support higher order cognition and regulation, which support discourse oral language skills (e.g., listening comprehension, oral composition, retell), which, in turn, support reading comprehension and written composition (see Fig. 2). The hierarchical relations hypothesis entails a series of interconnected relations among skills such that the numerous language and cognitive skills and knowledge are directly “and” indirectly related to reading comprehension and written composition. Therefore, once higher order skills are accounted for, the direct relations of lower order skills to reading comprehension and written composition are reduced (in the case of partial mediation by higher order skills) or disappear (in the case of full mediation by higher order skills). For example, if a statistical model includes working memory, vocabulary, and listening comprehension as predictors of reading comprehension, working memory and vocabulary would not directly predict reading comprehension because their contributions are captured by listening comprehension, a higher order skill, as long as listening comprehension is adequately measured (i.e., the vocabulary and working memory demands of listening comprehension tasks and reading comprehension tasks do not vastly differ; see Kim, 2020b, 2022b for further details).

The second hypothesis of the interactive dynamic literacy model is that reading and writing skills as well as their component skills develop interactively, influencing each other (i.e., interactive relations hypothesis; Kim, 2020d). The interactive relations are mediated by learning experience and practice. Reading and writing have bidirectional relations (Ahmed et al., 2014; Shanahan and Lomax, 1986), and reading experience supports writing skills, and writing experience supports reading skills (Graham et al., 2018). For example, reading texts helps individuals recognize and learn text structures, linguistic features, and authors’ purpose and intent, all of which support writing texts; the experience of writing texts also helps individuals understand these aspects, which, in turn, supports reading comprehension (Graham and Hebert, 2010). Word reading experience leaves orthographic, phonological, and morphological traces in memory, which supports spelling development, and the same applies to the experience of spelling words for word reading development. Similarly, vocabulary knowledge supports reading comprehension, and students with better reading comprehension tends to read a greater amount of texts than students who struggle with reading, which leads to greater exposure of vocabulary, which leads to better reading comprehension. Successes or failures of reading and writing experiences trigger one’s socio-emotions toward reading and writing, which, in turn, support or inhibit the development of reading and writing skills, which further leads to socio-emotions toward reading and writing.

Interactive relations implicate virtuous and vicious cycles of development of reading and writing skills, and their component skills. This hypothesis is akin to the Matthew Effects hypothesis in learning—the rich get richer, and the poor get poorer (Stanovich, 1986). Initial skill level is critical to setting off the cycle, and therefore, supporting development of foundational skills is crucial for successful reading and writing development. Also critical is continuous support for reading and writing development and their component skills with sustained instruction and opportunities for practice as learning experiences and practice are key to promoting the virtuous cycle and breaking the vicious cycle.

The third hypothesis of the interactive dynamic literacy model is that reading and writing difficulties co-occur (Kim, 2020d). If reading and writing are related and they rely on largely the same resources, success or difficulties with reading and writing development likely co-occur. Those who experience difficulties in reading will likely experience difficulties in writing, and vice versa (e.g., Graham et al., 2021). Furthermore, aligned with the hierarchical relations hypothesis, difficulties with reading and writing co-occur with weaknesses in oral language and cognitive skills. That is, those who have a weak foundation in oral language skills (including those with developmental language disorder) likely experience weak reading and writing skills (Berninger et al., 2015; Carretti et al., 2016; Catts et al., 1999; Dockrell and Connelly, 2015; Fey et al., 2004; Graham et al., 2020). Those who have weak skills in lexical literacy skills (word reading and spelling), including those with dyslexia, also experience difficulties in reading comprehension and written composition (Berninger and May, 2011; Berninger et al., 2008; Botting, 2007; Catts et al., 2005). Those who have weak higher order cognitions (e.g., inference or perspective taking) also experience difficulties in comprehension (Cain and Oakhill, 2007) and composition (Kim, 2020c; Kim and Park, 2019). Those who have weaknesses in working memory or attentional control are vulnerable to weaknesses in language and cognitive skills, and consequently in reading comprehension and written composition (Berninger, 2008; Kim, 2020b,c).

The last hypothesis is dynamic relations: reading-writing relations differ due to the development of reading and writing skills, grain size of reading and writing, and tasks and measurement (Kim, 2020d). Dynamic relations as a function of development refer to differential roles of component skills to reading and writing skills depending on the developmental phase of lexical literacy skills (word reading, transcription). In the beginning phase, lexical literacy skills have a large influence on discourse literacy skills. Development of lexical literacy skills reduces its constraining role, and permits oral language, background knowledge, and higher order cognitions to play greater roles in reading comprehension and written composition (Kim, 2020b; Kim and Graham, 2022).

The strength of reading-writing relations also differs as a function of grain size: The relation is stronger between word reading and spelling than between reading comprehension and written composition. Indeed, evidence indicates that word reading and spelling are very strongly related whereas reading comprehension and written composition are moderately related (Kim et al., 2022). The differential strength of reading-writing relations has implications for the occurrence of reading-writing difficulties. The strong relation between word reading and spelling implies that individuals who struggle with word reading are highly likely to experience difficulties with spelling, and those who struggle with spelling are highly likely to experience difficulties with word reading. The moderate relation between reading comprehension and written composition implies that although strong readers tend to be strong writers, there will be more divergent cases—skilled readers may not necessarily be strong writers (Kim, 2020d).

Word reading and spelling are strongly related because they essentially involve the same processes of retrieval of graphemes and their associated phonological (and morphological) information. Although the sequence of processes involved in word reading and spelling is opposite—word reading involves retrieving letters and graphemes and their associated phonological (and morphological) information, selecting appropriate grapheme-sound correspondences, and assembling them, whereas spelling involves retrieving accurate phonological (and morphological) information, mapping them with graphemes, and assembling and writing graphemes in correct order (Ehri, 1997; Kim, 2022a)—word reading and spelling rely on the same constrained set of skills: phonological, orthographic, and morphological awareness.

Compared to word reading and spelling, reading comprehension and written composition involve considerably divergent processes although they both involve meaning-making processes and draw on highly similar skills and knowledge. In reading comprehension, the text is given to readers and therefore, the extent of meaning-making is delimited by the given text (Kim, 2020d; Langer, 1986; Langer and Flihan, 2000). In written composition, the writer has to create a text, and therefore, the idea generation delimits the subsequent writing process. Let's take an example of reading versus writing a recipe. When reading a recipe, one needs to decode, understand each proposition, and develop a mental model of the cooking sequence. When writing a recipe, one has to know and select a recipe to write about, think about necessary materials (ingredients, tools) and the organization of each step, think about how to express each step by selecting vocabulary and sentence structures, and consider readers' needs (what units of measurement to use, whether to use images such as pictures for which steps, how much detail to provide, and what style of language and organization to use). This illustration clearly shows that writing a recipe requires more than knowing or comprehending a recipe. It requires extensive planning and critical thinking to a much greater extent than reading a recipe. In other words, although both reading comprehension and written composition draw on highly similar skills and knowledge, the skills and knowledge are differentially tapped during reading and writing processes.

The relation between reading and writing, particularly between reading comprehension and written composition, is also expected to differ depending on the dimension of writing—reading comprehension is more strongly related to writing quality than writing productivity (Kim and Graham, 2022; Kim et al., 2022). Written composition is evaluated in multiple ways. For instance, quality of overall writing, the presumed outcome in written composition, typically evaluates quality of ideas, organization, word choice, sentence structures, voice, spelling, and punctuation. Measures beyond writing quality are also widely used in the field, particularly for developing writers, including writing productivity (amount of written text; Anderson, 1988; Kim et al., 2015), writing fluency ("efficiency and automaticity in writing connected texts"; Kim et al., 2018, p. 322), or curriculum-based writing indicators (CBM writing; e.g., correct minus incorrect word sequence; Graham et al., 2011; Kim et al., 2015; McMaster and Espin, 2007). While these indicators are not the ultimate goal of written composition, they are used as indicators of writing quality for developing writers as they are reliable, have high utility for progress monitoring purposes, and are moderately to strongly related to writing quality for developing writers (Kim et al., 2015; Kim and Graham, 2022). However, the extent to which writing quality, writing productivity, writing fluency, and CBM writing draw on language and cognitive skills differs. Writing quality draws on all the skills shown in Fig. 2 whereas writing productivity and writing fluency are strongly influenced by transcription skills, and CBM writing is strongly influenced by one's grammatical and spelling skills (Kim et al., 2014, 2015; Kim and Graham, 2022). Given the multi-dimensional nature of written composition, the magnitude of the relation between written composition and reading comprehension varies depending on the focal dimension of written composition (Kim and Graham, 2022; also see Kim, 2022b; Francis et al., 2006 for a discussion of measurement of reading comprehension).

Reading-writing relations are also expected to vary depending on tasks. Reading and writing are used in a variety of tasks for different goals, and different tasks differentially tap skills. For example, reading a recipe to prepare dinner versus reading about World War II for a class project would not require the same extent of language and cognitive skills. Similarly, higher order cognitions are differentially needed for writing a short recount of a weekend activity versus writing a book report elaborating on motivations and behaviors of multiple characters. Depending on the nature of the tasks and goals, readers and writers draw on skills and knowledge differentially (Kim, 2020a,b; Kim and Graham, 2022).

Implications

Reading and writing skills as well as their component skills develop interacting with environments. One prominent environment that has a direct impact on reading and writing development is formal instruction. Variation in quality of instruction makes a difference in students' reading and writing development (e.g., Connor et al., 2013; Gersten et al., 2010; Kelcey and Carlisle, 2013). Theoretical models and evidence reviewed above imply that component skills of reading and writing development need to be systematically and explicitly taught. For example, for word reading and spelling development, orthographic symbols and their associated sounds, phonological compositions of words, and morphological structures need to be taught. Unlike oral language skills, reading and writing skills do not develop "naturally" (Rayner et al., 2001), and cracking the code requires explicit and systematic instruction for the majority of learners (e.g., Foorman et al., 2016; Graham and Santangelo, 2014; NICHD, 2000; Snow et al., 1998). Instruction should also target oral language skills and higher order cognitions, and support building background knowledge as early as possible. "Natural" development does not discount the importance of environmental input—although individuals develop oral language skills as long as language experience is present, linguistic environment has considerable influence on language development (e.g., Hart and Risley, 1995). Vocabulary tends to receive instructional attention when it comes to oral language skill, but oral language skill is a much larger construct, which includes syntax or sentence structures and discourse skills (e.g., listening

comprehension, oral retell and composition; Kim, 2016, 2020a). Higher order cognitions such as inference are often taught in the context of reading comprehension when students have developed word reading skills. However, higher order cognitive skills develop and should be also fostered in the oral language context before reading skills develop.

The hierarchical nature of relations indicates that instruction should be systematic considering the chain of interconnections (Kim, 2020b, 2022b). For instance, fluent word reading and spelling skills are needed to free cognitive resources for higher order thinking processes. Interactive relations indicate that integrated instruction of reading and writing, in addition to instruction that focuses on reading and writing respectively, will result in a synergistic effect that maximizes reading and writing development (Graham, 2021; Kim, 2020d). The hypothesis of co-occurrence of reading-writing difficulties implies a need for attending to reading and writing skills together in identifying and addressing students' needs. For example, for individuals with dyslexia, it is important to consider and address difficulties not only in reading, but also in writing. Similarly, for individuals with writing difficulties (including those with dysgraphia), difficulties in reading should be assessed and addressed (Kim, 2020d).

Lastly, the complex and multi-dimensional nature of reading comprehension and written composition suggests a need for a thoughtful approach to measuring reading comprehension and written composition, and a need for taking this into consideration in assessment and instruction. For example, it is important to use multiple tasks to measure reading comprehension and written composition skills precisely, especially for high-stakes decisions (e.g., determining one's eligibility to school-based services; e.g., Francis et al., 2006; Kim et al., 2017).

Acknowledgments

This chapter was supported by Grants R305A130131, R305A200312, R305A180055, and R305A170113 from the Institute of Education Sciences, US Department of Education as well as P50 HD052120 from the National Institute of Child Health and Human Development. The content is solely the responsibility of the author and does not necessarily represent the official views of the funding agencies.

References

- Abbott, R.D., Berninger, V.W., 1993. Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *J. Educ. Psychol.* 85, 478–508.
- Adlof, S.M., Catts, H.W., Little, T.D., 2006. Should the simple view of reading include a fluency component? *Read. Writ.* 19, 933–958. <https://doi.org/10.1007/s11145-006-9024-z>.
- Ahmed, Y., Wagner, R.K., Lopez, D., 2014. Developmental relations between reading and writing at the word, sentence, and text levels: a latent change score analysis. *J. Educ. Psychol.* 106, 419–434. <https://doi.org/10.1037/a0035692>.
- Anderson, P., 1988. Productivity, syntax, and ideation in the written expression of remedial and achieving readers. *J. Read. Writ. Learn. Disabil. Int.* 4 (2), 115–124.
- Apel, K., Wilson-Fowler, E.B., Brimo, D., Perrin, N.A., 2012. Metalinguistic contributions to reading and spelling in second and third grade students. *Read. Writ.* 25, 1283–1305. <https://doi.org/10.1007/s11145-011-9317-8>.
- Bereiter, C., Scardamalia, M., 1987. *The Psychology of Written Composition*. Erlbaum.
- Berninger, V.W., May, M.O., 2011. Evidence-based diagnosis and treatment for specific learning disabilities involving impairments in written and/or oral language. *J. Learn. Disabil.* 44 (2), 167–183. <https://doi.org/10.1177/0022219410391189>.
- Berninger, V.W., Yates, C.W., Cartwright, A., Rutberg, J., Remy, E., Abbott, R., 1992. Lower-level developmental skills in beginning writing. *Read. Writ.* 4, 257–280.
- Berninger, V.W., Nielsen, K.H., Abbott, R.D., Wijsman, E., Raskind, W., 2008. Writing problems in developmental dyslexia: under-recognized and under-treated. *J. Sch. Psychol.* 46, 1–21. <https://doi.org/10.1016/j.jsp.2006.11.008>.
- Berninger, V.W., Richards, T.L., Abbott, R.D., 2015. Differential diagnosis of dysgraphia, dyslexia, and OWL LD: behavioral and neuroimaging evidence. *Read. Writ.* 28, 1119–1153. <https://doi.org/10.1007/s11145-015-9565-0>.
- Berninger, V.W., 2008. Defining and differentiating dysgraphia, dyslexia, and language learning disability within a working memory model. In: Silliman, E., Mody, M. (Eds.), *Language Impairment and Reading Disability-Interactions Among Brain, Behavior, and Experience*. Guilford, pp. 103–134.
- Botting, N., 2007. Comprehension difficulties in children with specific language impairment and pragmatic language impairment. In: Cain, K., Oakhill, J. (Eds.), *Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective*. Guilford Press, pp. 81–103.
- Cain, K., Oakhill, J.V., 2007. Reading comprehension difficulties: correlates, causes, and consequences. In: Cain, K., Oakhill, J. (Eds.), *Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective*. Guilford Press, pp. 41–75.
- Carretti, B., Motta, E., Re, A., 2016. Oral and written expression in children with reading comprehension difficulties. *J. Learn. Disabil.* 49 (1), 65–76. <https://doi.org/10.1177/0022219414528539>.
- Catts, H.W., Fey, M.E., Zhang, X., Tomblin, J.B., 1999. Language basis of reading and reading disabilities: evidence from a longitudinal investigation. *Sci. Stud. Read.* 3 (4), 331–361. https://doi.org/10.1207/s1532799xssr0304_2.
- Catts, H.W., Adlof, S.M., Hogan, T.P., Weismer, S.E., 2005. Are specific language impairment and dyslexia distinct disorders? *J. Speech Lang. Hear. Res.* 48, 1378–1396.
- Connor, C.M., Morrison, F.J., Fishman, B., Crowe, E.C., Al Otaiba, S., Schatschneider, C., 2013. A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychol. Sci.* 24, 1408–1419. <https://doi.org/10.1177/0956797612472204>.
- DeFrancis, J., 1984. *The Chinese Language: Fact and Fantasy*. University of Hawaii Press.
- Dockrell, J., Connelly, V., 2015. The role of oral language in underpinning the text generation difficulties in children with specific language impairment. *J. Res. Read.* 38 (1), 18–34. <https://doi.org/10.1111/j.1467-9817.2012.01550.x>.
- Ehri, L.C., 1997. Learning to read and learning to spell are one and the same, almost. In: Perfetti, C.A., Rieben, L., Fayol, M. (Eds.), *Learning to Spell: Research, Theory, and Practice Across Languages*. Lawrence Erlbaum Associates Publishers, pp. 237–269.
- Ehri, L.C., 2005. Development of sight word reading: phases and findings. In: Snowling, M.J., Hulme, C. (Eds.), *Blackwell Handbooks of Developmental Psychology. The Science of Reading: A Handbook*. Blackwell Publishing, pp. 135–154.
- Fey, M.E., Catts, H.W., Proctor-Williams, K., Tomblin, J., Zhang, X.Y., 2004. Oral and written story composition skills of children with language impairment. *J. Speech Lang. Hear. Res.* 47 (6), 1301–1318.

- Fitzgerald, J., Shanahan, T., 2000. Reading and writing relations and their development. *Educ. Psychol.* 35, 39–50. https://doi.org/10.1207/S15326985EP3501_5.
- Florit, E., Roch, M., Levorato, M.C., 2014. Listening text comprehension in preschoolers: a longitudinal study on the role of semantic components. *Read. Writ.* 27, 793–817. <https://doi.org/10.1007/s11445-013-9464-1>.
- Forman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C.A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., Wissel, S., 2016. Foundational Skills to Support Reading for Understanding in Kindergarten Through 3rd Grade (NCEE 2016-4008). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education, Washington, DC. <http://whatworks.ed.gov>.
- Francis, D.J., Snow, C.E., August, D., Carlson, C.D., Miller, J., Iglesias, A., 2006. Measures of reading comprehension: a latent variable analysis of the diagnostic assessment of reading comprehension. *Sci. Stud. Read.* 10, 301–322. <https://doi.org/10.1177/0014402918795830>.
- Francis, D.J., Kulesz, P.A., Benoit, J.S., 2018. Extending the simple view of reading to account for variation within readers and across texts: the Complete View of Reading (CVR). *Remedial Spec. Educ.* 39, 274–288. <https://doi.org/10.1177/0741932518772904>.
- Gersten, R., Dimino, J., Jayanthi, M., Kim, J.S., Santoro, L.E., 2010. Teacher study group: impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *Am. Educ. Res. J.* 47 (3), 694–739. <https://doi.org/10.3102/0002831209361208>.
- Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10.
- Graesser, A.C., Singer, M., Trabasso, T., 1994. Constructing inferences during narrative text comprehension. *Psychol. Rev.* 101, 371–395.
- Graham, S., Hebert, M.A., 2010. Writing to Read: Evidence for How Writing Can Improve Reading. A Carnegie Corporation Time to Act Report. Carnegie Corporation of New York.
- Graham, S., Santangelo, T., 2014. Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Read. Writ.* 27, 1703–1743. <https://doi.org/10.1007/s11445-014-9517-0>.
- Graham, S., Berninger, V.W., Abbott, R.D., Abbott, S.P., Whitaker, D., 1997. Role of mechanics in composing of elementary school students: a new methodological approach. *J. Educ. Psychol.* 89, 170–182. <https://doi.org/10.1037/0022-0663.89.1.170>.
- Graham, S., Berninger, V.W., Fan, W., 2007. The structural relationship between writing attitude and writing achievement in first and third grade students. *Contemp. Educ. Psychol.* 32, 516–536. <https://doi.org/10.1016/j.cedpsych.2007.01.002>.
- Graham, S., Harris, K.R., Hebert, M., 2011. Informing Writing: The Benefits of Formative Assessment. Alliance for Excellent Education, Washington, DC.
- Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., Olinghouse, N., 2012. Teaching Elementary School Students to Be Effective Writers: A Practice Guide (NCEE 2012-4058). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, DC. http://ies.ed.gov/ncee/www/publications_reviews.aspx#pubsearch.
- Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Olson, C.B., Smither Wulsin, C., 2016. Teaching Secondary Students to Write Effectively (NCEE 2017-4002). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education, Washington, DC. <http://whatworks.ed.gov>.
- Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K.R., Aitken, A., Barkel, A., Kavanaugh, C., Talukdar, J., 2018. Reading for writing: a meta-analysis of the impact of reading interventions on writing. *Rev. Educ. Res.* 88, 243–284. <https://doi.org/10.3102/0034654317746927>.
- Graham, S., Hebert, M., Fishman, E., Ray, A.B., Rouse, A.G., 2020. Do children classified with specific language impairment have a learning disability in writing? A meta-analysis. *J. Learn. Disabil.* <https://doi.org/10.1177/0022219420917338>.
- Graham, S., Aiken, A.A., Hebert, A., Camping, A., Santangelo, T., Harris, K.R., Eustice, K., Sweet, J.D., Ng, C., 2021. Do children with reading difficulties experience writing difficulties? A meta-analysis. *J. Educ. Psychol.* <https://doi.org/10.1037/edu0000643>.
- Graham, S., 2021. The sciences of reading and writing must become more fully integrated. *Read. Res. Q.* 55 (S1), S35–S44. <https://doi.org/10.1002/rq.332>.
- Hart, B., Risley, T., 1995. Meaningful Differences in the Everyday Experience of Young American Children. Paul H. Brookes.
- Hayes, J.R., Flower, L.S., 1980. Identifying the organization of writing processes. In: Gregg, L.W., Steinberg, E.R. (Eds.), *Cognitive Processes in Writing*. Erlbaum, pp. 3–29.
- Huey, E., 1968. *The Psychology and Pedagogy of Reading*. MIT Press.
- Jenkins, J.R., Fuchs, L.S., van den Broek, P., Espin, C., Deno, S.L., 2003. Sources of individual differences in reading comprehension and reading fluency. *J. Educ. Psychol.* 95 (4), 719–729. <https://doi.org/10.1037/0022-0663.95.4.719>.
- Joshi, R.M., Aaron, P.G., 2012. Componential Model of Reading (CMR): validation studies. *J. Learn. Disabil.* 45, 387–390. <https://doi.org/10.1177/0022219411431240>.
- Joshi, R.M., Treiman, R., Carreker, S., Moats, L.C., 2008-2009. How Words Cast Their Spell: Spelling is an Integral Part of Learning the Language, Not a Matter of Memorization. *American Educator*, pp. 6–43.
- Katz, L., Frost, R., 1992. The reading process is different for different orthographies: the orthographic depth hypothesis. In: Frost, R., Katz, L. (Eds.), *Advances in Psychology. Orthography, Phonology, Morphology, and Meaning*, vol. 94. North-Holland, pp. 67–84.
- Kelcey, B., Carlisle, J.F., 2013. Learning about teachers' literacy instruction from classroom observations. *Read. Res. Q.* 48 (3), 301–317. <https://doi.org/10.1002/rq.51>.
- Kellogg, R.T., 1996. A model of working memory in writing. In: Levy, C.M., Ransdell, S.E. (Eds.), *The Science of Writing*. Erlbaum, pp. 57–71.
- Kellogg, R.T., 2008. Training writing skills: a cognitive developmental perspective. *J. Writ. Res.* 1 (1), 1–26.
- Kendeou, P., Bohn-Gottler, C.M., White, M.J., van den Broek, P., 2008. Children's inference generation across different media. *J. Res. Read.* 31, 259–272. <https://doi.org/10.1111/j.1467-9817.2008.00370.x>.
- Kim, Y.-S.G., Graham, S., 2022. Expanding the Direct and Indirect Effects model of Writing (DIEW): dynamic relations of component skills to various writing outcomes. *J. Educ. Psychol.* 114 (2), 215–238. <https://doi.org/10.1037/edu0000564>.
- Kim, Y.-S.G., Park, S., 2019. Unpacking pathways using the Direct and Indirect Effects model of Writing (DIEW) and the contributions of higher order cognitive skills to writing. *Read. Writ.* 32 (5), 1319–1343. <https://doi.org/10.1007/s11445-018-9913-y>.
- Kim, Y.-S., Phillips, B., 2014. Cognitive correlates of listening comprehension. *Read. Res. Q.* 49 (3), 269–281. <https://doi.org/10.1002/rq.74>.
- Kim, Y.-S.G., Pilcher, H., 2016. What is listening comprehension and what does it take to improve listening comprehension? In: Schiff, R., Joshi, M. (Eds.), *Handbook of Interventions in Learning Disabilities*. Springer, pp. 159–174.
- Kim, Y.-S.G., Schatschneider, C., 2017. Expanding the developmental models of writing: a direct and indirect effects model of developmental writing (DIEW). *J. Educ. Psychol.* 109, 35–50. <https://doi.org/10.1037/edu0000129>.
- Kim, Y.-S.G., Wagner, R.K., 2015. Text (oral) reading fluency as a construct in reading development: an investigation of its mediating role for children from grades 1 to 4. *Sci. Stud. Read.* 19, 224–242. <https://doi.org/10.1080/1088438.2015.1007375>.
- Kim, Y.-S., Apel, K., Al Otaiba, S., 2013. The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to instruction. *Lang. Speech Hear. Serv. Sch.* 44, 1–11. [https://doi.org/10.1044/0161-1461\(2013\)12-0013](https://doi.org/10.1044/0161-1461(2013)12-0013).
- Kim, Y.-S., Al Otaiba, S., Folsom, J.S., Greulich, L., Puranik, C., 2014. Evaluating the dimensionality of first-grade written composition. *J. Speech Lang. Hear. Res.* 57, 199–211. [https://doi.org/10.1044/1092-4388\(2013\)12-0152](https://doi.org/10.1044/1092-4388(2013)12-0152).
- Kim, Y.-S., Al Otaiba, S., Wanzek, J., Gattin, B., 2015. Towards an understanding of dimension, predictors, and gender gaps in written composition. *J. Educ. Psychol.* 107, 79–95. <https://doi.org/10.1037/a0037210>.
- Kim, Y.-S.G., Schatschneider, C., Wanzek, J., Gattin, B., Al Otaiba, S., 2017. Writing evaluation: rater and task effects on the reliability of writing scores for children in grades 3 and 4. *Read. Writ.* 30, 1287–1310. <https://doi.org/10.1007/s11445-017-9724-6>.
- Kim, Y.-S.G., Gattin, B., Al Otaiba, S., Wanzek, J., 2018. Theorization and an empirical investigation of the component-based and developmental writing fluency construct. *J. Learn. Disabil.* 51 (4), 320–335. <https://doi.org/10.1177/0022219417712016>.
- Kim, Y.-S.G., 2015a. Developmental, component-based model of reading fluency: an investigation of word-reading fluency, text-reading fluency, and reading comprehension. *Read. Res. Q.* 50, 459–481. <https://doi.org/10.1002/rq.107>.

- Kim, Y.-S., 2015b. Language and cognitive predictors of text comprehension: evidence from multivariate analysis. *Child Dev.* 86, 128–144. <https://doi.org/10.1111/cdev.12293>.
- Kim, Y.-S.G., 2016. Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *J. Exp. Child Psychol.* 141, 101–120. <https://doi.org/10.1016/j.jecp.2015.08.003>.
- Kim, Y.-S.G., 2017. Why the simple view of reading is not simplistic: unpacking the simple view of reading using a direct and indirect effect model of reading (DIER). *Sci. Stud. Read.* 21 (4), 310–333. <https://doi.org/10.1080/10888438.2017.1291643>.
- Kim, Y.-S.G., 2020a. Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: testing the direct and indirect effects model of reading (DIER). *J. Educ. Psychol.* 112 (4), 667–684. <https://doi.org/10.1037/edu0000407>.
- Kim, Y.-S.G., 2020b. Toward integrative reading science: the direct and indirect effects model of reading (DIER). *J. Learn. Disabil.* 53 (6), 469–491. <https://doi.org/10.1177/0022219420908239>.
- Kim, Y.-S.G., 2020c. Structural relations of language, cognitive skills, and topic knowledge to written composition: a test of the direct and indirect effects model of writing (DIEW). *Br. J. Educ. Psychol.* 90, 910–932. <https://doi.org/10.1111/bjep.12330>.
- Kim, Y.-S.G., 2020d. Interactive dynamic literacy model: an integrative theoretical framework for reading and writing relations. In: Alves, R., Limpo, T., Joshi, M. (Eds.), *Reading-Writing Connections: Towards Integrative Literacy Science*. Springer, pp. 11–34. https://doi.org/10.1007/978-3-030-38811-9_2.
- Kim, Y.-S.G., 2022a. A tale of two closely related skills: word reading and spelling development and instruction. In: Philippakos, Z.A., Graham, S. (Eds.), *Reading and Writing Connections*. Guilford Press.
- Kim, Y.-S.G., 2022b. Simplicity meets complexity: expanding the simple view of reading with the direct and indirect effects model of reading. In: Cabell, S., Neuman, S., Patton-Terry, N. (Eds.), *Handbook on the Science of Early Literacy*. Guilford Press.
- Kim, Y.-S.G., Wolters, A., Lee, J.-W., 2022. Reading and Writing are Differentially Related Depending on Grain Size and Dimensions of Written Composition: A Meta-analysis. Manuscript submitted for publication.
- Kintsch, W., 1988. The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95 (2), 163–182.
- LaBerge, D., Samuels, S.J., 1974. Toward a theory of automatic information processing in reading. *Cognit. Psychol.* 6, 293–323.
- Langer, J.A., Applebee, A.N., 1987. *How Writing Shapes Thinking: A Study of Teaching and Learning*, second ed. National Council of Teachers of English, Urbana, IL.
- Langer, J.A., Flihan, S., 2000. Writing and reading relationships: constructive tasks. In: Indrisano, R., Squire, J.R. (Eds.), *Writing: Research/Theory/Practice*. International Reading Association, pp. 112–129.
- Langer, J.A., 1986. Reading, writing, and understanding: an analysis of the construction of meaning. *Writ. Commun.* 3 (2), 219–267.
- McBride-Chang, C., Tardif, T., Cho, J.-R., Shu, H., Fletcher, P., Strokes, S.F., Wong, A., Leung, K., 2008. What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Appl. Psycholinguist.* 29, 437–462. <https://doi.org/10.1017/S014271640808020X>.
- McMaster, K., Espin, C., 2007. Technical features of curriculum-based measurement in writing: a literature review. *J. Spec. Educ.* 41, 68–84.
- Meyer, B.J.F., 1975. *The Organization of Prose and Its Effects on Memory*. North-Holland.
- National Institute of Child Health and Human Development, 2000. Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction (NIH Publication No. 00-4769). <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>.
- Rayner, K., Foorman, B., Perfetti, C.A., Pesetsky, D., Seidenberg, M.S., 2001. How psychological science informs the teaching of reading. *Psychol. Sci. Publ. Interest* 2 (2), 31–74. <http://www.jstor.org/stable/40062357>.
- Santangelo, T., Graham, S., 2016. A comprehensive meta-analysis of handwriting instruction. *Educ. Psychol. Rev.* 28 (2), 225–265. <https://doi.org/10.1007/s10648-015-9335-1>.
- Seymour, P.H.K., Aro, M., Erskine, J.M., 2003. Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. <https://doi.org/10.1348/000712603321661859>.
- Shanahan, T., Lomax, R.G., 1986. An analysis and comparison of theoretical models of the reading-writing relationship. *J. Educ. Psychol.* 78, 116–123.
- Shanahan, T., Callison, K., Carriere, C., Duke, N.K., Pearson, P.D., Schatschneider, C., Torgesen, J., 2010. Improving Reading Comprehension in Kindergarten Through 3rd Grade: A Practice Guide (NCEE 2010-4038). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from whatworks.ed.gov/publications/practiceguides.
- Shanahan, T., 2016. Relationships between reading and writing development. In: MacArthur, C.A., Graham, S., Fitzgerald, J. (Eds.), *Handbook of Writing Research*. Guilford Press, pp. 194–207.
- Snow, C.E., Burns, M.S., Griffin, P. (Eds.), 1998. *Preventing Reading Difficulties in Young Children*. National Academy Press.
- Snow, C.E., 2010. Reading comprehension: reading for learning. In: McGaw, B., Peterson, P., Baker, E. (Eds.), *The International Encyclopedia of Education*, third ed. Elsevier, pp. 413–418.
- Stanovich, K.E., 1986. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21 (4), 360–407. <https://doi.org/10.1598/RRQ.21.4.1>.
- Tompkins, V., Guo, Y., Justice, L.M., 2013. Inference generation, story comprehension, and language in the preschool years. *Read. Writ.* 26, 403–429. <https://doi.org/10.1007/s11145-012-9374-7>.
- Zhou, Y., McBride-Chang, C., Wong, N., 2014. What is the role of visual skills in learning to read? *Front. Psychol.* 5. <https://doi.org/10.3389/fpsyg.2014.00776>. Article 776.
- Ziegler, J.C., Goswami, U., 2005. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131 (1), 3–29. <https://doi.org/10.1037/0033-2909.131.1.3>.

Learning in the sciences

Jonathan Osborne, Graduate School of Education, Stanford University, Stanford, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	344
Challenge 1: the failure of common sense	345
Challenge 2: abstraction	346
Challenge 3: learning the languages of science	346
Lexical density	347
Vocabulary & Academic language	347
Logical connectives	347
Polysemy or Ambiguity of Meaning	348
Nominalizations	348
Multimodality	348
Visualizations	348
Challenge 4: the nature of the curriculum	349
Conclusion	350
References	350

Introduction

Learning science is hard. Such a statement automatically invites the question of why science is any harder than any other subject—a not unreasonable response. Yet, I will argue that there are somethings about the sciences that make developing any understanding harder than most, if not all other subjects. Some of these reasons are intrinsic to the nature of the sciences and some of them are extrinsic to the way the sciences are framed and the way they are taught.

At a very basic level, there is the commonsense test. If you were to wonder into an undergraduate chemistry lecture, confronted with symbols, mathematics and charts, most of us would be at a loss to understand what was being discussed. In contrast, wondering into a humanities lecture, most of us would at least be able to understand the gist of what was being discussed. What this points to is that one of the challenges of learning science is learning a language—essentially the challenge of acquiring the languages of the discipline and becoming disciplinary literate (Shanahan and Shanahan, 2008). And the plural “languages” is used because words alone are insufficient to communicate meaning in science. Rather, as well as words, science relies on charts, symbols, mathematics, diagrams and pictures to communicate meaning.

Why this is so is obvious if you think for a moment about how difficult it would be to explain the workings of the heart using words alone. Hence, the language of science is multi-semiotic or multi-modal (Kress et al., 1998; Kress and Van Leeuwen, 2001; Lemke, 1998; Martin, 1998).

To start, it helps to define what I mean by “science”. First, as the reader will note from the title, I would argue that there is no singular science. Rather there is a diversity of activities which share a common goal of seeking to answer three questions about the material world (Ogborn, 1988; Osborne, 2011). These are:

1. What exists? What philosophers call the ontic question
2. What causes this event/phenomenon to happen? The causal question
3. How do we know? The epistemic question

In addition, scientific knowledge invites two other questions:

4. How do we talk about this phenomenon/explanation? The communicative question
5. What can we do with this knowledge? The technological question

The sciences are also defined by their use of a set of distinctive “styles of reasoning” (Crombie, 1994; Hacking, 1992; Kind and Osborne, 2017). These are Mathematical Deduction, Experimental Exploration, Hypothetical Modeling, Categorization and Classification, Probabilistic Reasoning and Historical-Based Evolutionary Reasoning. In doing so the sciences use a mix of deductive mathematical and logical arguments such as those used in the derivation of Newton’s Laws; inductive generalizations represented by the laws of conservation of energy and conservation of momentum; and abductive arguments which are inferences to the best explanation (e.g., theory of evolution, origin of the continents). However, which of these is the focus depends on which science you are engaged in. The plant taxonomist is clearly concerned with categorization and classification while the astrophysicist working in cosmology relies much more on developing theoretical arguments for the evolution of the universe, stars and other celestial objects. In short, all the sciences really share is a common epistemic focus represented by the three questions above.

The fact that there is such diversity in the sciences points to the first difficulty and challenge of learning any science—there is simply an enormous amount of knowledge to be learned or acquired, more than any mere mortal can acquire in a lifetime. This invites the question of which science to start with and, more importantly, which end of the telescope? Do you, for instance, begin by building the foundational bricks and let the picture of the edifice emerge over time or, alternatively, do you present a picture of the edifice and worry less about presenting a picture of how it was built? As for which science should be taught in school, in reality there is little choice as what is offered is a product of tradition and embedded institutionalization. For instance, a good argument could be made that environmental science is more important than physics, chemistry, or biology. However, the facilities and the teachers to teach it do not exist so—the extant choices persist.

So, what are the challenges to learning any science? In what follows, I will argue that scientific ideas are not simple and straightforward. Rather, there are four specific challenges. First, they are often extraordinary and difficult to believe. Second, the goal of science is to construct models or representations of reality. Doing this often means producing representations of things that we cannot see or are too large to imagine. Third, because of the complexity, it is communicated in a form that is reliant on not just words but a diversity of forms of language—all of which have to be taught and learned but often are not. The final challenge lies in the way the sciences present themselves. As a consequence they do not show themselves in their best light.

Challenge 1: the failure of common sense

To seek to explain is only human. Explanation is supported by a mix of observation and our social interaction with each other. The problem is that relying on observation and common-sense is a poor way to generate scientific understanding. Observation would suggest that day and night are caused by a moving Sun, that the Moon is bigger than the Sun, and that the Earth is flat. Common sense tells you that heavier things fall faster, that you need a force to keep an object in motion, and that air is weightless. School science tells you that the opposite may be true. And observation itself is pretty useless at explaining the structure and function of living organisms, why we get infected with diseases, or why we look like our parents. Common sense ideas are embedded in our language too, as we worry about “letting the cold in” rather than letting the heat out, or we “cast a glance” or “stare daggers” at others which suggests that vision is an active process. And humans are definitely not animals as science would have you believe.

Moreover, common sense and our everyday experience reinforces these ideas such that by the time you arrive in a science classroom as a neophyte student you have already formed a good body of ideas about how the world works. The problem is that most of them are wrong. Scientific ideas, it turns out, are unnatural or uncommon sense (Cromer, 1993; Wolpert, 1992). Rather, the ideas developed by science could be categorized as crazy. Physics, for instance, wishes to tell us that heavier things do not fall faster, that air has mass and that it is possible for constant motion to be sustained without a constant force. When it comes to the atomic scale, science wants to argue that most of the atom is empty space and, at the macroscopic scale, that the Universe is 13 billion years old and created in a big bang.

Biology tells us that diseases are spread by tiny, invisible, living organisms that are pervasive in the air we breathe; that we look like our parents because every cell in our body carries a chemically coded blueprint of how to reproduce ourselves; and that plants are capable of taking carbon dioxide and water and turning them into sugars in their leaves. As for the geological sciences they would have you believe that continents can move! In short, the teacher of science is a teacher of crazy ideas. Few teachers of science recognize this though as so many of these ideas have become scientific commonplaces. Lost is the intellectual achievement they represent. And, as many a scientist might say, for good reasons. Science is a prospective discipline. What matters is creating new knowledge not celebrating the achievements of yesteryear. Once consensus has been achieved, science moves on and the struggles that were required to achieve this hard won knowledge are too often forgotten and remain untold (Matthews, 2015; Monk and Osborne, 1997). Yet, the lack of history, the lack of a story about the intellectual and creative struggle robs the teaching of science of any account of the cultural achievement it represents. It removes the human dimension, the social context and an important element of awe and wonder at the minds that were able to solve great intellectual challenges. Such awe and wonder are an important element of engagement for many young people (Gottlieb et al., 2018; Hadzigeorgiou, 2012).

More fundamentally for the learning science it also poses another problem. The extensive research carried out from a constructivist perspective on young children’s understanding of science shows conclusively that young people do not arrive in the classroom as *tabula rasa* (Black and Lucas, 1993; Driver et al., 1985; Osborne and Freyberg, 1985; Pfundt and Duit, 1994). And, given that it is well-established that we make sense of new ideas drawing on our existing background knowledge (Ausubel, 1968; Chi, 2009; Osborne and Wittrock, 1985), when the young student is confronted with an idea that differs significantly from their pre-existing conceptions, learning the new idea requires a substantial act of mental accommodation and transformation. In short, it takes time, reflection and exploration of the new idea and direct instruction is, not surprisingly, highly ineffective (Chi, 2009; Hake, 1998; Kuhn, 2007). Rather, the teaching of science must be seen as a conversation which begins with the teacher asking the student how they see the world, and then the teacher responding with their strange and unfamiliar version. “So”, you can imagine, the teacher saying, “you think the Earth is flat. I can understand why. It looks flat. But I have a strange idea to tell you. It only looks flat at your scale. If you go out, it is actually curved. Here is my argument to convince you ...”. Research shows that teachers who begin teaching this way have students who learn more effectively (Sadler et al., 2013). Likewise, texts that explain why the common misconception is flawed, as well as explaining why the scientific idea is correct, produce significantly better

learning outcomes (Guzetti et al., 1997; Sinatra and Broughton, 2011). Ausubel's famous mantra that "If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly" is as true today as when it was first written (Ausubel, 1968).

Challenge 2: abstraction

However, there is another challenge embedded in the conversation required to learn science. This is, as Rom Harré has pointed out, is that a scientific conversation is really about three types of entities. These are:

- The macroscopic things that surround us, both living and non-living
- The microscopic or invisible elements that we cannot see such as particles, germs or the insides of our bodies and the macroscopic entities which are too large to imagine such as a continent, the solar system, or the Universe
- The theoretical which are unobservable even with instrumentation—Quarks, Higgs Boson, strings (Harré, 1986).

At the macroscopic level, the referents are concrete and easily observable. This is why so much of the focus of elementary science is on living things and on introducing a way of categorizing and classifying such entities. What is the difference between living and non-living? What objects float and which ones sink? What are the phases of the Moon and how do they vary? Until such entities have been identified and labeled, we have no way of talking about them (Bowker and Leigh Star, 1999).

The challenge begins when the conversation moves to discussing anything that is not observable such as the workings of the human body or, alternatively, a new abstract concept such as force, speed, or an electric current. The only way we can talk about things we cannot see is by constructing models (Gilbert and Justi, 2016; Johnson-Laird, 1985; Nercessian, 2008). Some of these are physical models which bear a close similarity to the actual thing—such as skeletal or anatomical models of the heart, eye and other organs. Some, however, are more removed such as the models we enact to explain how day and night occurs, particle models of the air to explain how scent crosses a room, the various models of how electricity flows, or the Bohr model of the atom.

Models are not veridical accounts of nature. A moment's consideration of the Bohr model of the atom identifies numerous flaws (the electrons are far too large, the electrons are much further away, the nuclei are not colored and so on). Rather they are tools which are "true enough" that enable us to reason about the material worlds. They are essentially "felicitous falsehoods" (Elgin, 2017) whose aim is "to come up with a representation that affords an understanding of the phenomena, not one that replicates the phenomena."

However, such models are mentally challenging to the young mind which prefers concrete referents. For, despite the many arguments that have surrounded the work of Piaget (Adey and Shayer, 1990; Metz, 1995), it is impossible to escape the conclusion that the young child's brain is different from the adult, and that the young child thinks differently (Dehaene, 2020). During adolescence, we are in the process of building neural networks and strengthening the neural network that is the basis of our cognitive capabilities. Learning to think with a model requires intellectual effort and active processing. It is challenging. Challenging tasks need support and time to practice. The British project "Cognitive Acceleration through Science Education" was one approach that took this idea seriously with results that showed significant gains for the experimental group (Adey, 1992; Adey and Shayer, 1994).

In addition to models, there is also the fact that many of the abstractions employed by science such as force, acceleration, density, heat, and power have both no concrete referent and are often ratios. For instance, explaining why it is important to keep young babies wrapped up well requires an understanding of the ratio of surface area to volume of any object. And while velocity or speed are a ratio of distance per unit of time, acceleration is a ratio of a ratio! Moreover, many of the models constructed by science idealize the entities so we talk of point masses, ideal gases and genes in a manner that belies the reality that surrounds us. Getting your mind around these is hard, intellectual work. Not for nothing did Sylvia Plath write of her early experience of physics education that:

The day I went into the physics class it was death. A short dark man with a high lisping voice, named Mr. Manzi, stood in front of the class in a tight blue suit holding a little wooden ball. He put the ball on a steep grooved slide and let it run down to the bottom. Then he started talking about let a equal acceleration and let t equal time and suddenly he was scribbling letters and numbers and equals signs all over the blackboard and my mind went dead.

Plath (1966, p. 36).

Indeed, the mathematics Sylvia Plath experienced is an example of the Harré's third realm—the theoretical which we represent with mathematics. Mathematics is an integral part of science and unavoidable. However, its abstract nature and distance from any common macroscopic referent makes it difficult for many students to understand.

Challenge 3: learning the languages of science

One of the major challenges of science, however, is that the sciences are reliant on a multi-semiotic mode of communication. As Lemke argues:

Science does not speak of the world in the language of words alone, and in many cases it simply *cannot* do so. The natural language of science is a synergistic integration of words, diagrams, pictures, graphs, maps, equations, tables, charts, and other forms of visual mathematical expression.

Lemke (1998).

Clearly the sciences make extensive use of technical vocabulary. Moreover, sentences are written in a dense prose often using the passive voice and they are accompanied by many diagrams, charts and figures that are essential to understanding the meaning of the text. But why is this? Must it be so? The answer is an unequivocal yes. In all contexts, language evolves to be as efficient as possible in communicating meaning. The consequence is a development of a professional language that is only understood by those who have been inducted into the meaning of the terms and how to use them appropriately. This is as true for a group of lawyers, doctors or architects as it is for scientists. This view of language, which goes under the name of systemic functional linguistics, is that:

Technical language has evolved in order to classify, decompose and explain. The major scientific genres—report, explanation and experiment—have evolved to structure texts which document a scientist's world view. The functionality of these genres and the technicality they contain cannot be avoided; it has to be dealt with. To deal with it, teachers need an understanding of the structure of the genres and the grammar of technicality Ways must be devised to provide access to this technology. And the answer must not be watering the technology down.

Halliday and Martin (1993).

What then are the features of this technical language that make it so difficult? Knowing what those challenges are is a vital element of the knowledge of the professional science educator if they are to help their students learn from such texts. In summary, they are:

Lexical density

The most recognized feature of informational science texts is their sheer quantity of information or lexical density (Halliday and Martin, 1993). Lexical density refers to the number of content words within a text and how much information is in the text. The consequence is that the cognitive demand of constructing meaning from science texts is such that it cannot be comprehended in one pass. Rather, the reader must break, annotate, and reflect to construct meaning (Davies and Greene, 1984). More than anything else, lexical density is the major distinction between narrative and informational texts.

Vocabulary & Academic language

Informational science texts are lexically dense using both technical vocabulary and academic language extensively (Pearson et al., 2010; Snow, 2010; Wellington and Osborne, 2001). Technical vocabulary are words that have a specific meaning within a discipline, while academic language refers to words that are encountered in academic discourse but are not specific to any discipline. Indeed, the quantity of words encountered in a high school science course can exceed that introduced in learning a foreign language (Merzyn, 1987). One of the fundamental difficulties in teaching the sciences is the common perception that it is reliant on a technical vocabulary of words which are unfamiliar to the uninitiated. Yet as study after study has shown, while the technical words do naturally cause difficulty with comprehension, it is the non-technical but unfamiliar words—in short words that are used in academic English or academic words like “abundant”, “linear”, “diversity”, “stimulate” and more that students find just as difficult to comprehend (Farrell and Ventura, 1998; Marshall et al., 1991; Pickersgill and Lock, 1991; Snow, 2010).

Another feature of academic language is the use of the passive tense. As Snow points out, “in everyday language most sentences begin with pronouns or animate subjects; verbs refer to actions rather than relations; and long sentences are characterized by sequencing of information rather than embeddings” (2010, p. 450). But what of the sciences? Here, the personal is excised and students are encouraged to write in the passive voice. So rather than writing “we took the Bunsen burner and heated the copper sulfate”, the standard genre of science would use the wording “the copper sulfate was heated” resulting in the excision of any sense of an actor or the personal. Similarly, reports or explanations in the sciences tend to remove the agents, the scene, the motives, and any sense of temporality. The argument for this approach in the sciences is that it is the object that is the primary feature of interest rather than what is done to it (Wellington and Osborne, 2001). Maintaining the impersonal authoritative stance, however, creates a distanced tone that is often puzzling to adolescent readers and is extremely difficult for adolescents to emulate in writing.

Logical connectives

Logical connectives are words that combine two sentences or statements (e.g., therefore, so, thus). Logical connectives are essential to the process of constructing an argument, generating the relationships between claims, warrants, and data, and a means of contrasting and comparing phenomena (Fang, 2008; Gardner, 1977; Halliday and Martin, 1993). In science they are vital to express causality—how one thing follows from another. Commonly, however, connectives are excised from science texts to improve readability by shortening sentences even though struggling readers benefit from more explicit texts (i.e., the ones that contain more connectives) and readers benefit from adding explicit connectives to increase inferential processing (Ainsworth and Burcham, 2007; O'Reilly and McNamara, 2007; Roman et al., 2006).

Polysemy or Ambiguity of Meaning

Polysemy refers to the existence of multiple meanings for a single word or phrase. Many research studies have shown that all students have considerable difficulty determining the appropriate meaning of a word in different contexts (Cassels and Johnstone, 1985; Pickersgill and Lock, 1991). Scientists would like their words to be a purely technical sign with no index of meanings. However, this can never occur because words can never be detached from their context of use (Wittgenstein, 1953). As Montgomery (1996) points out, “perceiving any word as exclusively technical—any word, above all that has long been used in other forms of discourse—can only come about from a socialized practice of reading, one that relies in the end on a kind of trained blindness (p. 46).” This is particularly true of words which are used in the sciences with one meaning and in everyday discourse with another meaning—words such as force, power, theory, excite and more.

Nominalizations

A nominalization is a verbal description of a process that is converted into a noun (e.g., the process of cutting down the forest is referred to as “deforestation”). Nominalizations allow complex ideas to be expressed with the minimal use of language (Halliday and Martin, 1993). In scientific journals, one in seven nouns is a nominalization (Holtz, 2009). Nominalization removes human agency, as in the example above, we no longer know who conducted the deforestation; and the process becomes an abstraction, making it harder to conceptualize (Schlepppegrell, 2004; Unsworth, 1999). Typical examples might be “the rate of engine failure” rather than writing “the rate at which engines fail” or “glass crack rate” rather than “the rate at which a glass crack grows”. In both of these examples, the verb has been converted into a noun. The sciences abound with such words—condensation, absorption, scattering, reduction, transmission are examples of words which are derived from nouns and used because they are functionally effective at capturing a process with the minimum number of words. But while a scientist might say “The grid supports the transmission of electricity” for a 9-year-old child, we would write “The grid is how electricity is transmitted”. The use of the verb reduces the grammatical complexity and is more akin to how young children learn to talk. Thus, the common use of nominalizations in science texts makes them unfamiliar and difficult for young children to read.

Multimodality

Natural language is very limited in its ability to describe continuous variation, shape, and the interrelationships of structure, form, and function. Indeed, as Lemke argues, often it *cannot* do so (Lemke, 1998). To achieve this goal, the sciences depends on a synergy of semiotic signs—symbols to represent elements; quantities and units to represent amounts; graphs and charts to summarize relationships, frequencies, and patterns; tables to summarize numerical data; and mathematics to express relationships (Kress et al., 2001; Lemke, 1998; O’Halloran, 2011). That these are all interdependent can be seen from a cursory examination of any contemporary scientific paper. Temporarily excising one of these components makes the process of constructing its meaning significantly harder, if not impossible. Thus, multimodality is the combination of two or more symbolic systems of meanings and the sciences do not speak of the world in the language of words alone, and in many cases they simply cannot do so.

Visualizations

Increasingly, ideas in the sciences are communicated through static or dynamic visualizations. That is, visualizations are representations of an object, situation, or set of information as a chart or other image. Contrary to the common perception that these facilitate communication, students often have difficulty in understanding and interpreting such representations (O’Toole, 2004; Phillips et al., 2010; Zhang and Linn, 2011). Thus, contrary to the idea that a diagram, chart or computer visualization might make it easier to understand, the opposite is often true.

For instance, consider the common standard diagram of the heart. Imagine if you will, how difficult it would be to describe this organ to another without the use of diagrammatic representation. Likewise, many other phenomena and their patterns of interaction are best described in the language of mathematics, which becomes a bridge between verbal language and the meaning scientists seek to express. So complex are some of the concepts and ideas that the sciences wish to communicate that they cannot be communicated with words. The modern theoretical physicist often speaks a language heavily dependent on mathematics. Thus, the task confronting the student is not one of learning the language of science but one of learning the *languages* of the sciences.

Moreover, the language of the sciences is exceptional in that its discourse is cumulative. This does not mean that it contains some inner hierarchy but merely that each conversation in any given scientific domain builds on ones that have gone before; the science thus progresses in a fundamental way that most other disciplines do not (Tallis, 1995). The consequence is that the discourse of the sciences increasingly deviates from that of other discourses. Compare, for instance, the writings of any 19th century introduction to gravitational theory with any contemporary text on gravity with its matrix mathematics, tensor calculus and more. The two are virtually incommensurable. For the neophyte student the multimodality of the sciences is an additional barrier to entry, extending the period of apprenticeship and learning. The issue is that students must be taught how to read and interpret these modes of semiotic communication which are often not straightforward.

Challenge 4: the nature of the curriculum

The final challenge in learning science is the way the sciences present themselves and the nature of the curriculum that frames how it is perceived by many a student. Across the globe, the standard approach is to take a foundationalist approach to curriculum design. Thus, the basic concepts of any topic are introduced first. Students studying physics will learn about force, velocity, speed, and Newton's Laws of Motion without much rationale as to why these might be important or any explanation of the cultural significance. Likewise, in Chemistry, the periodic table is presented as if it were commonplace, with no significance attached to the fact that it is one of the great achievements of the discipline. Photosynthesis will be studied in biology in a mechanistic manner but there will be little recognition of how amazing it is that plants can produce complex sugars from light, carbon dioxide and water in a chemical reaction that we cannot replicate.

Learning science means learning to understand a body of consensually agreed, well-established knowledge. There is no uncertainty about such knowledge. This can be seen in the fact that there is no meta-discourse about the degree of certainty of the nature of the knowledge that is presented. Any "may" or "possibly", as in Rutherford's initial interpretation of the implications of his experiment investigating how alpha particles were scattered by a thin gold foil, is absent. His finding suggested to him that the atom was mainly empty space with a very small dense nucleus. These tentative statements are simply transformed into statements of the form "the atom consist of a small dense nucleus of neutron and protons with electrons orbiting at a distance". In such a manner, hard-won knowledge is transformed into a set of taken-for-granted-facts.

Yet so much of science education has lost sight of that awe. As Dawkins so elegantly puts in his own homage to the sciences, *Unweaving the Rainbow* (Dawkins, 2000):

There is an anaesthetic of familiarity, a sedative of ordinariness which dulls the senses and hides the wonder of existence. For those of us not gifted in poetry, it is at least worthwhile from time to time making an effort to shake off the anaesthetic (p. 6).

Or as Pinker puts it, "[s]cience has provided the world with images of sublime beauty: stroboscopically frozen motion, ... graceful spiral galaxies and diaphanous nebulae ..., and a luminous Planet Earth rising above the moon's horizon" (Pinker, 2018). Moreover, as he points out, it has "granted us the gifts of life, health and wealth" and knowledge which has eradicated diseases such as smallpox which killed 300 million people in the 20th century alone, making the current Covid-19 pandemic look like chump change. Likewise, the development of penicillin reduced the deaths from infectious diseases per year from 90,000 in 1900 to 2300 in 1980 in the UK alone. And if, such knowledge is commonplace, why should it be of interest to the neophyte student?

Thus, laying out such unequivocal and unquestionable bricks of the scientific edifice makes the learning of science hard. As Claude Bernard famously said—"Science is a hall full of awe and wonder, the problem is the long dark kitchen you have to go through to get there." (Bernard, 1957) Students are given no sense of the scientific edifice, the intellectual achievement it represents and the awe and wonder that should be an inherent feature of such a great cultural achievement is simply lost. Associated with this is a lack of any discussion of the epistemology of science. Presented in this manner, students are given the impression that this is the way the world is. There is no acknowledgment that the goal of science is the construction of explanatory models (Lehrer and Schauble, 2006) and that these are not accurate representations. If there is any treatment of the epistemic basis of science it is often reduced to an introductory chapter on "the scientific method" which is (a) philosophically laughable as the methods of science are numerous and diverse (Allchin, 2012; Bauer, 1992; Crombie, 1994), and (b) suggests that all creativity in science is reduceable to the application of a singular algorithm. If this is the nature of scientific work, what is the value in making the effort to learn a body of knowledge which seems to have little relevance to their everyday life or practice?

Yet, science education has a duty to prepare students to evaluate scientifically based arguments. At a minimum that requires the ability to evaluate expertise in science. Essential to this competency is an understanding that science is a communal enterprise, that science is not infallible. In many ways, error is absolutely the norm in the sciences (Allchin, 2012). However, while individual scientists or teams might be mistaken, the community is more trustworthy and that the goal of the sciences is closure which is attained when there is overwhelming consensus—consensus which is established through the fundamental mechanism of peer review. Moreover, that consensus is a consensus of experts (Collins, 2014; Crease and Selinger, 2006; Oreskes, 2019). Science is not a democracy where one scientist's views are as valuable as another. The views that matter are from those who have relevant background knowledge, skills in interpreting particular results, and awareness of potential flaws in reasoning. Thus, a petition claiming that vaccines are dangerous is meaningless if the signatories are not scientific experts on vaccination. Then there is a set of data literacy skills—the knowledge that correlations are not necessarily causal, knowledge of the common manipulations used to massage data to support an inappropriate interpretation, and of reasoning based on flawed samples. Finally, there is a need to develop the generic skills that are applicable in all disciplines advocated by those who are concerned about civic reasoning (Hertwig, 2017; Lee et al., 2021; Wineburg and McGrew, 2017).

In short, there is a need to take an externalist view of the sciences and ask what are the capabilities and knowledge necessary to become a competent outsider—somebody who knows that they stand at an epistemic distance to science itself but yet is capable of interrogating the work of scientists in a valid and justified manner (Feinstein, 2011). What Feinstein argued is that the focus of

traditional science on the canon of concepts and theories leaves them as “marginal insiders”—that is people with a repertoire of poorly understood ideas that are of limited utility. As he points out, a knowledge of fluid dynamics is hardly helpful when it comes to fixing your brakes. Rather, competent outsiders are “people who have learned to recognize the moments when scientific knowledge has some bearing on their needs and interests and to interact with sources of scientific expertise in ways that help them achieve their own goals” (Feinstein, 2011). Moreover, competent outsiders have had an education which has developed the basic tools required to interrogate expertise. The failure of school science to present something which is seen as useful to their lives simply leaves the student with the impression that they are sitting on a train with blacked out windows. They know they are going somewhere but only the train driver knows where and too many decide that the journey is not worthwhile and the effort too much.

Conclusion

So, if the sciences are hard, what do we know about how learning in the sciences can be supported? As previously indicated, it is important to understand that learning has to begin with what the students already know, or do not know. Eliciting this information is the beginning of a conversation. However, beyond this, while good quality explanation is significant, learning science must be seen as learning a new language. As Postman and Weingartner (1971) argue:

Almost all of what we customarily call “knowledge” is language, which means that the *key to understanding a subject is to understand its language* [emphasis added]. A discipline is a way of knowing, and whatever is known is inseparable from the symbols (mostly words) in which the knowing is codified. What is biology (for example) other than words? If all the words that biologists use were subtracted from the language, there would be no biology This means, of course, that every teacher is a language teacher: teachers, quite literally, have little else to teach, but a way of talking and therefore seeing the world (p. 102).

Thus, a more productive approach to learning science is to recognize what language teachers understand—that language is acquired by interaction and participation in the disciplinary practices that require the use of that language. In the case of science these are asking and framing questions, designing and evaluating experimental designs, analyzing and interpreting data, engaging in argument from evidence, and communicating and obtaining scientific information. Not for nothing did Chi find, in her extensive review of the evidence on learning, that there was a hierarchy to the efficacy of learning activities. The least effective are those where the recipient is passive and does little or no processing of information—copying notes being the archetypal example. “A device”, as one wit pointed out, “by which the notes of the lecturer become the notes of the student without going through the minds of either.” Then there are those activities which require the student to be active and cognitively process information, such as the demands made by a teacher question. Better are those which require the student to produce something—the solution to a problem, a written report or a Freyer model defining a concept such as “physical change” with examples and non-examples. Best, however, are those which require the student to be interactive, requiring the student to engage in dialog with another in the production of an artifact—be it a response to a clicker question, the production of a powerpoint presentation, turning to talk with a partner about a question posed by a teacher, or working collaboratively discussing what makes a good concept map (Chi, 2009). Hake’s evidence about the failure of direct instruction is unequivocal (Hake, 1998). In short, research of the past 50 years had clearly defined the nature of the challenge and the best ways of addressing it. Yet too many continue to do the same thing over and over again, expecting different results. As Einstein allegedly said—the definition of insanity.

References

- Adey, P., Shayer, M., 1990. Accelerating the development of formal thinking in middle and high school students. *J. Res. Sci. Teach.* 27 (3), 267–285. <https://doi.org/10.1002/tea.3660270309>.
- Adey, P., Shayer, M., 1994. *Really Raising Standards*. Routledge.
- Adey, P., 1992. The CASE results: implications for science teaching. *Int. J. Sci. Educ.* 14 (2), 137–146.
- Ainsworth, S., Burcham, S., 2007. The impact of text coherence on learning by self-explanation. *Learn. Instr.* 17, 286–303.
- Allchin, D., 2012. Teaching the nature of science through scientific errors. *Sci. Educ.* 96 (5), 904–926. <https://doi.org/10.1002/sce.21019>.
- Ausubel, D.P., 1968. *Educational Psychology: A Cognitive View*. Holt, Rinehart and Winston, New York.
- Bauer, H.H., 1992. *Scientific Literacy and the Myth of the Scientific Method*. University of Illinois Press.
- Bernard, C., 1957. *An Introduction to the Study of Experimental Medicine*. Dover Publication (1865 Paris).
- Black, P., Lucas, A. (Eds.), 1993. *Children’s Informal Ideas About Science*. Routledge.
- Bowker, G.C., Leigh Star, S., 1999. *Sorting Things Out: Classification and Its Consequences*. MIT Press.
- Cassels, J.R.T., Johnstone, A.H., 1985. *Words That Matter in Science*.
- Chi, M., 2009. Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Topics Cognit. Sci.* 1, 73–105.
- Collins, H., 2014. *Are We All Scientific Experts Now?* Polity Press.
- Crease, R.P., Selinger, E., 2006. *The Philosophy of Expertise*. Columbia University Press.

- Crombie, A.C., 1994. *Styles of Scientific Thinking in the European Tradition: The History of Argument and Explanation Especially in the Mathematical and Biomedical Sciences and Arts*, vol. 1. Duckworth London.
- Cromer, A., 1993. *Uncommon Sense: The Heretical Nature of Science*. Oxford University Press.
- Davies, F., Greene, T., 1984. *Reading for Learning in the Sciences*. Oliver & Boyd.
- Dawkins, R., 2000. *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder*. Mariner Books.
- Dehaene, S., 2020. *How We Learn: The New Science of Education and the Brain*. Penguin UK.
- Driver, R., Guesne, E., Tiberghien, A., 1985. Children's ideas and the learning of science. In: Driver, R., Guesne, E., Tiberghien, A. (Eds.), *Children's Ideas in Science*. Open University Press, pp. 1–9.
- Elgin, C.Z., 2017. *True Enough*. Harvard University Press.
- Fang, Z., 2008. Going beyond the Fab Five: helping students cope with the unique linguistic challenges of expository reading in intermediate grades. *J. Adolesc. Adult Lit.* 51 (6), 476–487.
- Farrell, M.P., Ventura, F., 1998. Words and understanding in physics. *Lang. Educ.* 12 (4), 243–253.
- Feinstein, N., 2011. Salvaging science literacy. *Sci. Educ.* 95 (1), 168–185. <https://doi.org/10.1002/sce.20414>.
- Gardner, P.L., 1977. Logical connectives in science—a summary of the findings. *Res. Sci. Educ.* 7, 9–25.
- Gilbert, J.K., Justi, R., 2016. *Modelling-Based Teaching in Science Education*. Springer.
- Gottlieb, S., Keltner, D., Lombrozo, T., 2018. Awe as a scientific emotion. *Cognit. Sci.* 42 (6), 2081–2094.
- Guzetti, B., Williams, W.O., Skeels, S.A., Ming Wu, S., 1997. Influence of text structure on learning counterintuitive physics concepts. *J. Res. Sci. Teach.* 34 (7), 701–719.
- Hacking, I., 1992. "Style" for historians and philosophers. *Stud. Hist. Philos. Sci.* 23 (1), 1–20.
- Hadzigeorgiou, Y.P., 2012. Fostering a sense of wonder in the science classroom. *Res. Sci. Educ.* 42 (5), 985–1005.
- Hake, R.R., 1998. Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66 (1), 64–74. <http://link.aip.org/link/?AJP/66/64/1>.
- Halliday, M.A.K., Martin, J.R., 1993. *Writing Science: Literacy and Discursive Power*. Falmer Press.
- Harré, R., 1986. *Varieties of Realism: A Rationale for the Natural Sciences*. Basil Blackwell.
- Hertwig, R., 2017. When to consider boosting: some rules for policy-makers. *Behav. Publ. Pol.* 1 (2), 143–161. <https://doi.org/10.1017/bpp.2016.14>.
- Holtz, M., 2009. Nominalisation in scientific discourse: a corpus-based study of abstract and research articles. In: Mahlbert, M., González-Díaz, V., Smith, C. (Eds.), *Proceedings of Corpus Linguistics Conference*. University of Liverpool, pp. 341–359.
- Johnson-Laird (Ed.), 1985. *Mental Models*.
- Kind, P.M., Osborne, J.F., 2017. Styles of scientific reasoning: a cultural rationale for science education? *Sci. Educ.* 101 (1), 8–31. <https://doi.org/10.1002/sce.21251>.
- Kress, G.R., Van Leeuwen, T., 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Hodder Arnold.
- Kress, G.R., Ogborn, J., Jewitt, C., Tsatsarelis, B., 1998. *Rhetorics of the Science Classroom: A Multimodal Approach [Research]*.
- Kress, G.R., Jewitt, C., Ogborn, J., Tsatsarelis, C., 2001. *Multimodal Teaching and Learning: The Rhetorics of the Science Classroom*. Continuum Books.
- Kuhn, D., 2007. Is direct instruction an answer to the right question? *Educ. Psychol.* 42 (2), 109–113.
- Lee, C.D., White, G., Dong, D., 2021. *Educating for Civic Reasoning and Discourse*.
- Lehrer, R., Schauble, L., 2006. Scientific thinking and science literacy. In: Damon, W., Lerner, R.M., Eisenberg, N. (Eds.), *Handbook of Child Psychology*. Wiley, pp. 153–196.
- Lemke, J., 1998. *Teaching All the Languages of Science: Words, Symbols, Images and Actions*. <http://academic.brooklyn.cuny.edu/education/jlemke/papers/barcelon.htm>.
- Marshall, S., Gilmour, M., Lewis, D., 1991. Words that matter in science and technology. *Res. Sci. Technol. Educ.* 9 (1), 5–16.
- Martin, J.R., 1998. Discourses of science. In: Martin, J.R., Veal, R. (Eds.), *Reading Science*. Routledge, pp. 3–14.
- Matthews, M.R., 2015. *Science Teaching: The Role of History and Philosophy of Science, 20th Anniversary Revised and Expanded ed.* Routledge.
- Merzyn, G., 1987. The language of school science. *Int. J. Sci. Educ.* 9 (4), 483–489.
- Metz, K.E., 1995. Reassessment of developmental constraints on children's science instruction. *Rev. Educ. Res.* 65, 93–127.
- Monk, M., Osborne, J.F., 1997. Placing the history and philosophy of science on the curriculum: a model for the development of pedagogy. *Sci. Educ.* 81 (4), 405–424.
- Montgomery, S.L., 1996. *The Scientific Voice*. Guilford Press.
- Nercessian, N., 2008. Model-based reasoning in scientific practice. In: Duschl, R.A., Grandy, R.E. (Eds.), *Teaching Scientific Inquiry: Recommendations for Research and Implementation*. Sense, pp. 57–79.
- Ogborn, J., 1988. *A Map of Science: Personal Submission to the National Curriculum Working Group on Science Education*. London.
- O'Halloran, K., 2011. Multimodal discourse analysis. In: Hyland, K., Paltridge, B. (Eds.), *Companion to Discourse*. Continuum, pp. 120–137.
- O'Reilly, T., McNamara, D., 2007. Reversing the reverse cohesion effect: good texts can be better for strategic, high-knowledge readers. *Discourse Process.* 43, 121–152.
- Oreskes, N., 2019. *Why Trust Science?* Princeton University Press.
- Osborne, R.J., Freyberg, P. (Eds.), 1985. *Learning in Science*. Heinemann.
- Osborne, R.J., Wittrock, M., 1985. The generative learning model and its implications for science education. *Stud. Sci. Educ.* 12, 59–87.
- Osborne, J.F., 2011. Science teaching methods: a rationale for practices. *Sch. Sci. Rev.* 93 (343), 93–103.
- O'Toole, M., 2004. What is difficult to read, why might this be so, and what could, or should, be done about it? An overview of section two. In: P. A., Cleghorn, A. (Eds.), *Missing the Meaning*. Palgrave Macmillan, pp. 161–178.
- Pearson, D., Moje, E.B., Greenleaf, C., 2010. Literacy and science: each in the service of the other. *Science* 328, 459–463.
- Prundt, H., Duit, R., 1994. *Bibliography: Students' Alternative Frameworks and Science Education*, fourth ed. IPN.
- Phillips, L., Norris, S., Macnab, J., 2010. *Visualization in Mathematics, Reading and Science Education*. Springer.
- Pickersgill, S., Lock, R., 1991. Students' understanding of selected non-technical words in science. *Res. Sci. Technol. Educ.* 9 (1), 71–79.
- Pinker, S., 2018. *Enlightenment NOW: The Case for Reason, Science, Humanism and Progress*. Allen Lane.
- Plath, S., 1966. *The Bell Jar*. Faber and Faber.
- Postman, N., Weingartner, C., 1971. *Teaching as a Subversive Activity*. Penguin, Harmondsworth.
- Roman, D.X., Briceño, A., Rohde, H., Hironaka, S., 2006. Linguistic cohesion in middle-school texts: a comparison of logical connectives usage in science and social studies textbooks. *Electron. J. Sci. Educ.* 20 (1–19).
- Sadler, P.M., Sonnett, G., Coyle, H.P., Cook-Smith, N., Miller, J.L., 2013. The influence of teachers' knowledge on student learning in middle school physical science classrooms. *Am. Educ. Res. J.* 50 (5), 1020–1049.
- Schleppegrell, M., 2004. *The Language of Schooling: A Functional Linguistics Perspective*. Erlbaum.
- Shanahan, T., Shanahan, C., 2008. Teaching disciplinary literacy to adolescents: rethinking content area literacy. *Harv. Educ. Rev.* 78 (1), 40–59.
- Sinatra, G.M., Broughton, S.H., 2011. Bridging reading comprehension and conceptual change in science education: the promise of refutation text. *Read. Res. Q.* 46 (4), 374–393.
- Snow, C., 2010. Academic language and the challenge of reading for learning about science. *Science* 328, 450–452.
- Tallis, T., 1995. *Newton's Sleep: Two Cultures and Two Kingdoms*. St Martin's Press.

- Unsworth, L., 1999. Developing critical understanding of the specialised language of school science and history texts: a functional grammatical perspective. *J. Adolesc. Adult Lit.* 42, 508–521.
- Wellington, J., Osborne, J.F., 2001. *Language and Literacy in Science Education*. Open University Press.
- Wineburg, S., McGrew, S., 2017. *Lateral Reading: Reading Less and Learning More When Evaluating Digital Information*.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Blackwell.
- Wolpert, L., 1992. *The Unnatural Nature of Science*. Faber and Faber.
- Zhang, Z., Linn, M.C., 2011. Can generating representations enhance learning with dynamic visualizations? *J. Res. Sci. Teach.* 48 (10), 1177–1198.

Learning from multiple texts

Matthew T. McCrudden^a, Ivar Bråten^b, and Ladislao Salmerón^c, ^a Department of Educational Psychology, Counseling, and Special Education, Pennsylvania State University, University Park, PA, United States; ^b Department of Education, University of Oslo, Oslo, Norway; and ^c ERI Lectura—University of Valencia, Valencia, Spain

© 2023 Elsevier Ltd. All rights reserved.

Models of learning from multiple texts	354
Construction-Integration (C-I) model	354
Document models framework (DMF)	354
Multiple-document task-based relevance assessment and content extraction (MD-TRACE) model	354
Reading as problem solving (RESOLV) model	355
Discrepancy-induced source comprehension (D-ISC) model	355
Content-source integration (CSI) model	355
Integrated framework of multiple texts (IF-MT)	356
The role of individual differences	356
Comprehension skills	356
Prior knowledge	356
Beliefs	356
Motivation and engagement	357
Strategic performance	357
Measurement of multiple source comprehension	357
Teaching integration and sourcing skills in multiple text contexts	358
Teaching content integration	358
Teaching source evaluation	359
Reading on the internet	359
Conclusion	361
References	361

Individuals read for many purposes. When individuals have questions, are attempting to solve problems, or seek to improve their understanding, they commonly access text. These purposes can be self-generated or assigned by others (e.g., teachers) and can be part of informal contexts (e.g., reading to understand the nutritional value of different foods) and formal contexts (e.g., reading as part of a class to complete an assessment task). For this article, we define *multiple texts* (or multiple textual sources) as different text-based sources about the same topic. *Learning from multiple texts* can involve selecting, comprehending, evaluating, and integrating information from two or more textual sources to carry out an assigned task or a self-selected reading goal (Braasch et al., 2018).

To illustrate, primary school students may be asked to read to determine whether a dolphin is a fish or a mammal. The students could be provided different texts about fish, mammals, and dolphins. They could identify the defining characteristics of fish and mammals, compare the characteristics of dolphins with the defining characteristics of fish and mammals, and subsequently infer that a dolphin is a mammal. Each text relates to a larger whole that is not specified by any single text. In this example, the teacher may provide students with texts that are deemed credible and provide information consistent with biologists' accepted understandings of fish and mammals. However, in other cases, readers may need to consider characteristics of the sources. An information source can be defined as the person or entity to whom the information is attributed (Macedo-Rouet et al., 2013). Suppose students are asked to evaluate the impact of progressive tax rates on different segments of the economy (e.g., job creation, capital investment, wage rates, foreign investment). This example differs from the previous example (dolphins) not only in the number and complexity of the concepts, but some of the texts may provide conflicting views. Thus, readers need to consider not only the content of a text, but also the source to whom the content is attributed.

The purpose of this article is to provide an overview of learning from multiple texts. The article consists of seven main sections. In the first section, we provide an overview of models that describe learning from multiple texts. In the second section, we review research on the role of individual differences in learning from multiple texts. In the third section, we describe how researchers measure readers' comprehension of multiple texts and their representation of information about sources. In the fourth section, we discuss approaches to teaching students to integrate content information across texts and to critically evaluate source information. In the fifth section, we discuss reading on the Internet (including social media) and the use of multimedia documents and sources. Finally, we end with general comments about learning from multiple texts.

Models of learning from multiple texts

Models are useful for trying to describe the processes and products of learning from multiple texts. Processes refer to the moment-by-moment activities that occur before, during, and after a reading experience. Products refer to the mental representations and external artifacts (e.g., notes, essays) that readers generate from a reading experience. In this section, we review seven models. The first model, the Construction-Integration model (Kintsch, 1988, 1998) was developed to describe how readers construct meaning from text, implicitly a single text. However, this model is foundational to subsequent models developed to describe learning from multiple texts. In particular, the Construction-Integration model introduces different levels of mental representation that inform subsequent models. These later models build upon and/or complement each other by specifying additional processes, variables, or contexts for describing how readers learn from multiple texts.

Construction-Integration (C-I) model

According to the Construction-Integration (C-I) model (Kintsch, 1988, 1998), two main processes facilitate comprehension. The first is *construction*, which begins when text information and prior knowledge are active in the reader's focus of attention. The text information can be the sentence currently being read or sentences read earlier. Sentences from the text can be fused with each other and with prior knowledge, forming a propositional network. The second is *integration*, which involves the continuous spread of activation throughout this network until activation settles on the overlapping nodes in the network, which decreases activation of isolated nodes or nodes that have less overlap.

Over the course of reading, the reader constructs a mental model of the text. The reader's mental model can be represented at three levels: surface code, textbase, and situation model (Kintsch, 1998). The *surface code* represents the exact wording and syntax of sentences. Readers typically only retain the surface code of the most recent clause. The *textbase* represents the gist meaning of the propositions stated explicitly in the text, rather than the exact wording and syntax. The *situation model* represents the reader's overall understanding of the information in the text. The situation model is the fusion of information in the text with the reader's prior knowledge. Thus, skilled readers generally create similar textbases, which reflect the content and organization as provided by the author. However, situation models may vary as a function of readers' goals, interests, and background knowledge.

Document models framework (DMF)

The C-I model (Kintsch, 1988, 1998) was conceived of as a framework for describing learning from a single text. Building upon the C-I model, Perfetti et al. (1999) introduced the Documents Model Framework (DMF) to describe the mental representations that are created when learning from multiple texts. According to the DMF, readers generate a documents model that includes a situations model and an intertext model. The *situations model* (or integrated mental model) reflects the individual situation model for each text, but also the inter-connections between/among the individual situation models. Thus, the situations model is a composite of the situation models for each text and represents the reader's overall understanding of the topic or phenomenon described in the texts. The *intertext model* contains information about the individual texts themselves and how they are related, such as each text's style (e.g., informational vs. narrative) and intent (e.g., to persuade or inform), the author's status, perceptions of text quality, and whether the authors of the different texts agree or disagree.

Multiple-document task-based relevance assessment and content extraction (MD-TRACE) model

The MD-TRACE model describes the resources and processes readers use when learning from multiple texts (Rouet and Britt, 2011). As with the Documents Model Framework, the reader generates a documents model (situations model/integrated mental model and intertext model) based on the texts. However, the MD-TRACE model extends learning from multiple texts to include a task model. A *task model* is the reader's understanding of the task requirements, which contributes to goal formation and suggests the means necessary to achieve that goal. Thus, when reading multiple texts, the reader not only constructs a documents model (situations model and intertext model), but also generates a task model that directs the reading activity.

The MD-TRACE model includes five core processes that occur *iteratively* over the course of learning from multiple texts. For simplicity, we will describe the processes as steps, but it is important to recognize that the process is iterative. The first step involves constructing a task model. It includes the goals readers create and actions that should be taken to reach the goals. The second step involves assessing information needs. These needs are clarity about the task to be done (e.g., write an essay vs. prepare for a multiple-choice test), familiarity with the task (e.g., knowing how to write an argument essay), prior knowledge about the content of the text, and what resources are needed (e.g., books, journal articles) and where to find them. The third step involves selecting, processing, and integrating document information. *Selection* involves determining the relevance of information in relation to the goal for reading. Information that more effectively contribute to a reader's goals is more relevant than information that less effectively contributes to a reader's goals (McCrudden and Schraw, 2007). Processing involves attempts to create meaning. It involves cognitive activities to make sense of ideas, connecting ideas from the text with each other, and connecting ideas from the text with prior knowledge. It can also include evaluations of the author/source as well as the texts themselves. The fourth step is the construction of a task product. The reader engages with multiple texts to reach some goal, which is generally reflected in the task model. The task product involves the utilization of the documents model and the external resources, such as the texts themselves, to create some

output. This can be simple, such as an oral response, or complex, such as a paper that synthesizes information from several sources. The last step involves assessing product quality. The reader assesses whether the product meets the task goals, and whether improvements are needed. All five of these steps draw heavily on readers' use of self-regulation.

Reading as problem solving (RESOLV) model

The MD-TRACE model (Rouet and Britt, 2011) highlights the decision-based nature of reading. The RESOLV model (REading as problem SOLVing; Britt et al., 2018; Rouet et al., 2017), extends this focus further. It includes three additional assumptions. First, dimensions of the reading context other than the task instructions influence reader goals. Second, readers' interpretations of contextual cues influence their goals. Third, motivation, interest, and value influence reading decisions, such as whether to read and what to read.

Readers generate goals for reading in context. As such, according to the RESOLV model readers generate a context model. A *context model* can include information about the request (what is the task statement), the requester (who is making the request), the audience (who will receive the completed request), supports or obstacles (what external resources are available and how can they support or impede request completion), and the self (the readers' assessment of their own skill, knowledge, and interest related to the request and their assessment of the perceived costs and benefits of the activity). Thus, readers who receive the same task instructions may generate different interpretations of the task context and/or have different levels of motivation for initiating, sustaining, and completing activities needed to achieve the task.

In both the MD-TRACE and RESOLV models, the reader creates a task model. In both, the task model includes the goal (goal state for the final product, subgoals, constraints) and how to accomplish it (plans, strategies). However, the RESOLV model also includes the values associated with subgoals (importance of desired states and methods). In other words, readers assess the value of engaging in effortful behaviors and strategies when deciding what to do while reading. Thus, the task model complements the context model.

Like previous models, RESOLV assumes that the mental product of a reading activity is a documents model (situations model/integrated mental model and intertext model). Rouet et al. (2017) further specify the nature of the intertext model. The *intertext model* represents source information (e.g., who wrote the document and in what context) and connections between content and sources. Source-to-content connections are links between a source and specific information or ideas from a particular document (e.g., who said what). Source-to-source connections are links between sources that represent readers' interpretations about whether information and ideas from different documents corroborate or conflict with each other.

Discrepancy-induced source comprehension (D-ISC) model

The discrepancy-induced source comprehension (D-ISC) model, which stems from the DMF, describes the cognitive processes readers use when reading to understand controversial messages (Braasch et al., 2012; Braasch and Bråten, 2017; Braasch and Kessler, 2021). The DMF is relevant to situations in which information from multiple sources is consistent, complementary, or conflicting. The D-ISC model is particularly relevant to situations in which multiple sources present conflicting information about the same issue (e.g., sun exposure is harmful vs. healthy). In these situations, readers need to understand the conflict and attempt to reconcile the different perspectives so they can build an integrated model of the situation or issue. One way to do this is to shift attention strategically toward the sources of conflicting information in order to resolve the conflict in light of the sources (e.g., by acknowledging that different authors may have different expertise or motives).

Content-source integration (CSI) model

The Content-Source Integration (CSI) model (Stadtler and Bromme, 2014) also describes how readers may use a document's source to draw conclusions about conflicting information. The CSI model consists of three stages. The first stage is *conflict detection*, which occurs when a reader perceives a contradiction between information from different sources. In this stage, readers notice when information they are currently reading conflict with information they have read previously. This introduces a coherence break. To restore coherence, readers engage in the second step, which is *conflict regulation*. There are three common ways that readers can restore coherence: (a) ignore the conflict (decide not to attempt to do anything to interpret the conflict) (b) reconcile the conflict using text (e.g., the text provides a satisfactory explanation for the conflict), or prior knowledge (e.g., use prior knowledge to explain the conflict), or (c) accept the conflict as due to different sources or points of view (i.e., interpret the disruption of coherence as due to different perspectives rather than a genuine conflict). The third stage is *conflict resolution*. In this stage, the reader needs to develop a personal stance on the conflict. They need to evaluate the truthfulness of conflicting claims, which entails making a first-hand or second-hand evaluation (Bromme and Goldman, 2014). When readers make first-hand evaluations, they use their understanding of the subject to assess the validity of the competing claims. The question to be addressed is: What is true? However, when readers have limited understanding of the subject, they may need to make a second-hand evaluation, which entails evaluating the sources themselves. The question then becomes: Whom to believe? Thus, when a topic is more familiar, readers may rely to a greater extent on first-hand evaluation, but when a topic is less familiar, readers may rely to a greater extent on second-hand evaluation (McCrudden et al., 2016).

Integrated framework of multiple texts (IF-MT)

The integrated framework of multiple texts (IF-MT; List and Alexander, 2019) describes the process of learning from multiple texts as a three-stage process. The first stage is *preparation*, in which the reader evaluates the reading task and develops a stance for how to complete it. Two dimensions characterize the reader's stance: affective engagement and behavioral dispositions toward the evaluation of sources and verification of information the sources provide (List and Alexander, 2017). *Affective engagement* refers to the degree to which students are motivated to complete a task. Motivational variables include topic interest, attitudes toward the topic, self-efficacy for the reading task, and value placed on doing the task. *Behavioral dispositions* refer to the extent to which readers have mastered the skills needed to evaluate sources and the content provided by sources, and their ability to integrate texts. Further, this dimension includes the extent to which the reader commonly uses these skills.

The second stage is *execution*, which describes the behaviors and strategies readers use while engaging with multiple texts. There are three core strategies. *Behavioral strategies* refer to the observable actions that readers take when they engage with the texts, including information search, text selection, and note taking activities. *Cognitive strategies* refer to the mental processes that readers use when they engage with the texts including strategies used to comprehend individual texts and strategies used to integrate ideas between/across texts. *Metacognitive strategies* refer to the regulatory actions that readers use when they engage with the texts, such as comprehension monitoring (i.e., self-evaluations of understanding).

The third stage is *production*, which involves the generation of an external product, such as an essay, to address the reading task. This usually reflects selective elements of the reader's mental representation of the texts, based on the specific task expectations. For instance, when writing a summary, the learner might include main ideas, but exclude some details that are retained in memory.

The role of individual differences

Working with multiple texts may result in a more complex and deeper understanding of a topic, issue, or phenomenon than working with one single text (Wolfe and Goldman, 2005). Further, working with multiple texts may lead to greater flexibility because the acquired knowledge is less tied to one specific text and therefore more accessible across different contexts (Spiro et al., 2015). However, there is no reason to believe that all learners will profit equally from working with multiple sources. Among the individual differences likely to play a role are learners' comprehension skills, prior knowledge, beliefs, motivation and engagement, and strategic processing.

Comprehension skills

Several recent studies have linked learners' single text comprehension skills to aspects of multiple text comprehension (e.g., Florit et al., 2019; Mahlow et al., 2020; Mason et al., 2020; Salmerón et al., 2020b). For example, Florit et al. (2019), in a study of Italian fourth-grade students, found that comprehension skills measured with a single informational text uniquely predicted multiple text comprehension measured with post-reading essay tasks. Likewise, Mahlow et al. (2020), in a study of German university students, found that comprehension skills measured with a single text predicted multiple text comprehension when controlling for other relevant variables. Such findings are consistent with the idea that comprehension of each single text indeed matters when learners try to build integrated understanding across multiple texts.

Prior knowledge

There is evidence that learners' prior knowledge about the topic discussed across multiple texts contributes to the comprehension of those texts (e.g., Bråten et al., 2014; Gil et al., 2010; Stang Lund et al., 2019). For example, Stang Lund et al. (2019) found that Norwegian upper-secondary students' prior knowledge both directly and indirectly predicted their ability to integrate source and content information when reading about the issue of sun exposure and health. Presumably, prior knowledge contributes to multiple source comprehension because it facilitates bridging inferences that create interconnection and coherence in complex, divergent textual materials (Bråten et al., 2020).

Results are mixed regarding the relationship between prior knowledge and attention to source information, such as the authors, when working with multiple texts. Notably, quite a few studies have found no relationship between prior knowledge and sourcing (for review, see Anmarkrud et al., 2021). This inconsistency seems to reflect theoretical divergence. For instance, some models (List and Alexander, 2017, 2019) imply that more prior knowledge will increase attention to source information through adoption of a critical analytic stance. However, others (Stadtler and Bromme, 2014) imply that less prior knowledge will increase attention to source information because learners do not feel competent in validating knowledge claims themselves.

Beliefs

Two kinds of beliefs have been studied in relation to multiple text comprehension. The first type of beliefs, called topic beliefs, concerns positions people take about the truth value of statements about some aspect of reality (Wolfe and Griffin, 2018). The second type, called epistemic beliefs, concerns positions people take about the truth value of statements about knowledge and knowing (Bråten and Strømsø, 2020).

Research has indicated that learners' prior topic beliefs may influence the comprehension of multiple texts, especially when those beliefs are inconsistent with information presented in the texts (for reviews, see Bråten and Strømsø, 2020; Richter and Maier, 2017). In such instances, learners tend to be biased toward their own prior beliefs when evaluating arguments and drawing conclusions, which does not bode well for creating a balanced mental representation of controversial issues discussed across multiple sources (Richter and Maier, 2017). However, learners may sometimes decouple their beliefs from the meaning making process and consider information independently of their prior beliefs, which increases the likelihood of belief revision and construction of a more balanced mental representation of the issue in question (McCrudden, 2020).

Much research has also been directed toward understanding the role of learners' beliefs about the nature of knowledge and the process of knowing when working with multiple texts (for reviews, see Bråten et al., 2016; Bråten and Strømsø, 2020). This research has indicated that believing that knowledge is tentative and complex and that the process of knowing involves justifying claims by corroborating and integrating information across sources may promote multiple text processing and comprehension (e.g., Barzilai and Eshet-Alkalai, 2015; Bråten et al., 2014; Wiley et al., 2020). Presumably, this is because such beliefs about knowledge and knowing are well aligned with the requirements of multiple text learning contexts, which involve open, complex problem spaces characterized by claims and arguments distributed across diverse, often conflicting textual sources (Bråten et al., 2020).

Motivation and engagement

Because comprehending multiple texts may be considered more challenging than comprehending a single text on a topic, it is reasonable to assume that comprehending multiple texts requires higher levels of learner motivation and engagement. Thus far, the motivation construct most researched in relation to multiple text comprehension is topic interest (e.g., Bråten et al., 2014; List, 2020; List et al., 2019b; Salmerón et al., 2018b; Stang Lund et al., 2019), although this research provides little support for a direct relationship between topic interest and multiple text comprehension. However, some research (Bråten et al., 2014; List, 2020) has suggested that topic interest may indirectly affect multiple text comprehension through behavioral engagement, that is, learners' active, observable involvement in tasks as typified by time, effort, persistence, and productivity (Eccles and Wang, 2012).

In a range of studies, behavioral engagement has been shown to be a proximal contributor to multiple source comprehension. In some of these studies (Bråten et al., 2018a; Bråten et al., 2014; Du and List, 2020; List, 2020; List et al., 2019b), the time learners invest in reading and comprehension tasks has uniquely predicted their comprehension performance. In other studies (Bråten et al., 2018a; Kammerer et al., 2021; Latini et al., 2019), learners' productivity (length of written responses) when working on post-reading essay tasks used to measure multiple text comprehension has been shown to be positively related to comprehension performance.

Strategic performance

At the same time, it is important to ensure that learners' engagement is channeled into adaptive strategic processing of the text materials. In particular, learners' use of cross-text elaboration strategies has been linked to multiple text comprehension (e.g., Bråten et al., 2014; Cho et al., 2017; List, 2020; List and Alexander, 2020; List et al., 2019a; McCrudden et al., 2021). These strategies involve intentional intertextual processing, such as connecting, comparing, and contrasting information across texts, to build an integrated understanding of the topic in question. Conversely, strategic processing that involves an accumulation of piecemeal information from single texts does not seem to promote integrated understanding of multiple texts (Bråten and Strømsø, 2011).

Strategic attention to source information, such as authors, publications, and date and purpose of creation, to evaluate information in light of its source, has been given special status within this area of research (Scharrer and Salmerón, 2016). Accordingly, a range of studies have indicated that learners' critical evaluation of source information is related to adaptive meaning making in multiple text contexts (e.g., Bråten et al., 2009; Cho et al., 2017; Florit et al., 2019; Salmerón et al., 2018a). Presumably, this is because it allows learners to prioritize information from competent, credible, and less biased sources, as well as to understand reasons for conflicting views on the same issue (e.g., because authors have different competencies or motives).

Measurement of multiple source comprehension

In reviewing the literature on multiple source comprehension, Primor and Katzir (2018) found that researchers had either used expressive tasks in the form of essays and open-ended questions or receptive tasks in the form of verification tasks to measure content integration across multiple texts. Essay tasks have sometimes been scored holistically (e.g., Anmarkrud et al., 2014; List et al., 2019b) and sometimes by segmenting them into idea units and identifying the types of inferences made (i.e., intratext vs. intertext inferences; Gil et al., 2010; Salmerón et al., 2018b). When open-ended questions have been used, they have directly or indirectly required learners to integrate different views across sources, capturing how well learners are able to reason about the issue in terms of the claims, reasons, and evidence provided across texts (Barzilai and Ka'adan, 2017; Bråten et al., 2014). In some instances (Le Bigot and Rouet, 2007; Latini et al., 2019), content integration has also been measured by taking learners' use of connective words into account. For example, Latini et al. (2019) scored undergraduates' written responses to open-ended questions by coding their use of causal (e.g., because, therefore) and adversative (e.g., however, whereas) connectives in combining, comparing, and contrasting information across sources.

Regarding verification tasks, Strømsø et al. (2008) introduced the intertextual inference verification task. This task requires learners to verify (yes or no) whether presented statements can be reasonably inferred by combining information across two or more texts recently read. Importantly, the statements included in an intertextual inference verification task should not just be combinations of statements included in the texts, but statements requiring the use of bridging inferences between information from those texts.

These different ways of measuring integration across multiple texts have different strengths and weaknesses. For example, whereas essays and open-ended questions seem well suited to assess deeper level, integrated understanding of multiple texts, such tasks involve additional writing skills. Ideally, writing skills should therefore be measured independently and their effects partialled out when researchers use essays and open-ended question tasks in correlational or experimental designs. Correspondingly, when learners respond orally instead of in writing to open-ended questions (e.g., Andresen et al., 2019), expressive oral language skills come into play and should preferably be accounted for. Whereas intertextual inference verification tasks do not require writing or expressive oral language skills, such tasks often have lower reliabilities unless they include a large number of items, and they involve a 50% chance of guessing correctly. Further, intertextual inference verification tasks typically include quite complex items (i.e., statements) that may be difficult to understand. As such, these tasks may require comprehension skills in and of themselves.

With respect to learners' representation of source information, an important distinction is between measurement of spontaneous (i.e., unprompted) and prompted representation of sources. For example, spontaneous representation of sources is measured through learners' references to source information or inclusion of source-content links when trying to explain conflicting positions on the same issue in an essay task, that is, without being given any specific instructions to refer to sources (e.g., Anmarkrud et al., 2014; Salmerón et al., 2018b). On the other hand, prompted representation of sources may be measured by probing learners' memory for source features after reading (e.g., Kammerer et al., 2016; Salmerón et al., 2018b), by asking them to rank or rate sources according to credibility (e.g., Barzilai et al., 2020b; Bråten et al., 2009), and by asking them to justify their ranking or rating of the sources (e.g., Braasch et al., 2014; Mason et al., 2018). As highlighted by Anmarkrud et al. (2021), learners generally have been found to perform rather poorly on measures of spontaneous source representation. On measures of prompted source representation, more variation among learners are typically found, which also makes it easier to discover relationships between learners' representation of source information and other variables, such as individual differences (Anmarkrud et al., 2021).

Noteworthy is also the development of scenario-based digital assessment tools within multiple source comprehension, such as the Online Research and Comprehension Assessment (ORCA; Leu et al., 2015) and the Global Integrated Scenario-based Assessment (GISA; Sabatini et al., 2018). In essence, scenario-based assessments involve purposeful reading of multiple thematically related sources, with built-in guidance provided throughout task completion and with both content integration and critical source evaluation targeted by the assessment (Sabatini et al., 2018). Wang et al. (2021) showed that a scenario-based assessment may be more difficult than a traditional single text comprehension assessment, with learners' prior topic knowledge also contributing more to performance on the former than on the latter. However, the scenario-based assessment that Wang et al. (2021) evaluated was based on only multiple-choice questions and did not seem to include guidance or assess critical evaluation of source information.

Teaching integration and sourcing skills in multiple text contexts

Given the great challenges that multiple text comprehension represents across educational levels, direct instruction and guidance seem necessary. This means that making multiple texts available to students, or even providing them with a selection of relevant sources, is hardly sufficient. In such situations, students may still perceive the task as one of gathering and reproducing as much information as possible from different texts, without constructing a coherent representation of the topic by integrating information across texts and without critically evaluating the credibility of the content in light of the source information. Students must therefore learn *how* diverse and complex text materials should be handled. In the following, we discuss approaches to teaching integration of content information across texts and critical evaluation of source information, respectively.

Teaching content integration

In a review of interventions targeting integration of information across multiple texts, Barzilai et al. (2018) showed that many of these interventions have emphasized collaboration and discussion among students. In such interventions, students have typically worked in pairs or small groups in which they have discussed texts and engaged in various text-based writing activities. As an example, Lundstrom et al. (2015) had students work with multiple texts individually before they shared what they had learned with their group. Students then organized information from the texts collaboratively and created an outline for writing a synthesis based on the texts.

Other interventions targeting content integration have emphasized various forms of strategy instruction, that is, guidance in using strategies to integrate content across texts (Barzilai et al., 2018). These interventions include guidance in how to annotate or summarize single texts (e.g., Britt and Sommer, 2004; Hagerman, 2017), as well as guidance in how to create and use graphic representations (e.g., tables or maps) that explicate how information from different texts hang together (e.g., Barzilai and Ka'adan, 2017).

In an 8 h intervention where Israeli ninth-graders worked with multiple texts on a historical topic that included both complementary and contradictory information, Barzilai et al. (2020a) combined student discussions and graphic representations. Students

learned to represent relationships between the contents of multiple texts graphically by means of a map, in which they also mapped connections between content and source information and between sources of information. Students used this graphic representation when they produced their own texts on the topic. They worked in pairs both when they mapped the connections between the texts and when they produced their own texts, and they also discussed their experiences with writing with support from the map in class. The results of this intervention showed that students who created and used the graphic representation collaboratively clearly outperformed students in a control group with respect to integrating information across texts on an essay task.

Teaching source evaluation

A number of interventions targeting sourcing skills have been implemented and evaluated during the last decades (for review, see Brante and Strømso, 2018). However, many of these interventions have been quite brief, implemented by researchers rather than teachers, and failed to provide evidence for long-term effects (Brante and Strømso, 2018).

As an example of an extensive intervention implemented by teachers as part of regular classroom activities and demonstrating not only immediate but also long-term effects, Bråten et al. (2019) conducted a study in which high school students were taught to consider relevant source features (i.e., author, author competence, author affiliation, type of document, publisher, publication date) when selecting texts, interpreting and evaluating content information, and writing from multiple texts. This six-week intervention was integrated into regular classroom activities in language arts and implemented by regular teachers who participated in professional development seminars led by the research team. The intervention built on a contrasting cases approach (Schwartz and Bransford, 1998) in which students were presented with two fictitious students who varied with respect to source evaluation skills, and they were tasked to compare and contrast these students' strategic attention to source information when working on authentic multiple text tasks framed by the curriculum. In this way, students were presented with a model of how they could approach multiple text tasks effectively, that is, by trying to emulate the sourcing skills illustrated by the "adaptive sourcer" case.

Results showed that students who participated in the intervention put more emphasis on source information when selecting texts, processed the selected texts more thoroughly, and more frequently attributed textual ideas to their respective sources during text-based writing than did students in the control group. Importantly, these effects were observed on far transfer tasks and were maintained over a 5.5-week period (Bråten et al., 2019).

Another example of a successful approach to teaching source evaluation is an extensive Finnish intervention study in which sixth-graders were taught to evaluate the credibility of multiple Web pages (Hämäläinen et al., 2020). This intervention, which also was implemented by the students' regular teachers, shared some important features with the Bråten et al. (2019) study. Thus, the younger Finnish students also were presented with and asked to compare contrasting cases that varied with respect to source evaluation, discussed source evaluation with each other and with the teacher, and practiced source evaluation when working with authentic multiple text tasks integrated in subject matter teaching. The results of this intervention showed that students who had participated in the intervention, as compared with a control group, to a larger extent based their evaluation of the credibility of Web pages on relevant source features, such as author expertise. However, no follow-up data were provided in this study.

Further research directed toward the design, implementation, and evaluation of long-term, classroom-based interventions is highly needed in this area. Preferably, teachers and researchers should collaborate to conduct such intervention research. Further, it should take place within the framework of regular subject teaching, where the taught sourcing skills can be practiced while working with authentic tasks that involve the evaluation of students' academic performance. Finally, although the teaching of source evaluation and the teaching of content integration have been discussed in two separate sections, it seems essential that these aspects of multiple text comprehension, in line with theoretical accounts, are integrated in educational interventions. The effects of such interventions also need to be evaluated on different tasks and on delayed post-tests to investigate the transferability and sustainability of those effects.

Reading on the internet

Understanding how students read on the Internet requires an understanding of Internet sources. First, all Internet sources are linked to other sources by means of hyperlinks. Reading a printed book involves collecting a book from a library or book store, whereas reading from the Internet often involves inspecting a search engine results page (SERP) to click and access a particular web page. Accordingly, reading on the Internet demands specific navigation competences, or the ability to select texts to read and the sequence in which to read those texts (Cho and Afflerbach, 2017).

Second, the Internet was created with an open editorial philosophy, which means that anybody with Internet access can share content with the rest of the World. The quality of printed books is often maintained by editorial gatekeepers (editors, librarians), whereas content on the Internet is not subject to the same editorial standards and can range from high-quality content written by authors with content-area expertise to low-quality content provided by novice or biased sources, such as bots spreading fake news (Braasch and Graesser, 2019). In a survey conducted in 2018, approximately two-thirds of EU citizens reported that they encountered fake news at least once a week (European Commission, 2018). Accordingly, reading on the Internet requires that readers regularly use active sourcing processes to filter, select, and interpret multiple texts (Bråten et al., 2020).

Third, the open nature of the Internet enables easy access to vast amounts of texts on virtually any topic. Thus, Internet readers have easy access to complementary information, alternative (or opposing) perspectives, or information in different formats (e.g., text, animated pictures, videos). But as Internet sources are created as single entities, and not as part of a coherent document, chances are high that documents referring to the same topic are not easily comparable. For example, they may lack contextual information, use different terminology, or use different types of arguments (e.g., scientific claims vs. personal accounts). Thus, reading multiple texts via the Internet may require continuous effort to integrate multiple sources across different formats.

Empirical evidence for the existence of the three major competencies that can be assumed to constitute the core of Internet reading (i.e., navigation, evaluation and integration), comes from a study by Kiili et al. (2018), who modeled the factor structure of scores on an Internet reading assessment in a large sample of sixth grade students (ages 12 and 13). They identified six factors that could be further categorized within the three major competencies: Navigation (1 factor: locating information with a search engine), Evaluation (3 factors: questioning credibility of information, confirming credibility of information, and communicating a justified, source-based position), and Integration (2 factors: identifying main ideas from a single online source, and synthesizing information across multiple online sources). Yet, empirical studies consistently show that young adolescents frequently display limited Internet reading competencies. For example, young readers frequently struggle to navigate, evaluate, and integrate digital information (e.g., Keil and Kominsky, 2013), and they often have difficulty evaluating the credibility of the sources they encounter (e.g., on YouTube or Twitter; e.g., Breakstone et al., 2019). Because such difficulties create particular challenges for younger or less-skilled readers, it is important to understand how Internet reading competences develop during school. Although some cross-sectional studies have investigated how navigation and evaluation competencies develop during school (Breakstone et al., 2019; Keil and Kominsky, 2013; Potocki et al., 2020; Rouet et al., 2011; Salmerón et al., 2018a), less is known about how the ability to integrate Internet sources develops. A widely used indicator of navigation efficiency is the percentage of relevant pages visited (e.g., OECD, 2009). This index takes into consideration that some Internet pages include relevant information for completing learning tasks, whereas other pages include irrelevant or distracting information. Navigation is positively related to students' reading comprehension abilities (Salmerón et al., 2018a). The challenge is that readers must use a few words in a hyperlink to infer the extent to which it contains semantically relevant information.

Like reading comprehension, navigation competencies increase with grade level across adolescence. Salmerón et al. (2018a) analyzed navigation (log-files) from a sample of adolescents that completed a digital reading assessment constructed on the basis of the PISA framework (OECD, 2009). Results revealed that navigation efficiency tended to be rather high, with average percentages of navigation efficiency between 80 and 85%, meaning that for every 10 pages visited, 8–8.5 contained relevant information for the students' goal. In terms of development, seventh and eighth graders showed similar navigation efficiency, but were less efficient than ninth and 10th graders, who also showed comparable navigation efficiency. Such development resembles the pattern obtained by Keil and Kominsky (2013), who asked participants to rate the relevance of pages from a list of four pages, where only two were semantically relevant to the learning goal. Participants' ability to detect relevant links substantially improved between the 6–7th grade and 9–10th grade samples. Interestingly, scores did not improve beyond 10th grade (11–12th grade and undergraduates). In a similar study, Rouet et al. (2011) found that young adolescents (5th grade) were highly influenced by superficial cues, such as using uppercase keywords in the hyperlink, regardless of their topic relevance. The influence of superficial cues on navigation gradually decreased and eventually disappeared across 7–11th grade students.

The assessment of evaluation competence usually includes the ability to identify sources, interpret source expertise and intentions, and to use such knowledge to judge the quality of information (Bråten et al., 2018b). Potocki et al. (2020) used an inventory to investigate the development of evaluation competence in a sample of 5th, 7th, 9th-graders and undergraduates. Overall, students across all age groups were proficient at identifying source elements (e.g., author) when requested to do so. Regarding source interpretation, the ability to discriminate competent from less competent authors increased gradually across the grade levels assessed, but the ability to discriminate benevolent from biased sources was more challenging and did not progress across grades. Finally, the ability to justify the selection of pages from a SERP based on source characteristics was challenging for fifth graders, with less than 10% citing source features to justify their selections, but increased across grades (approx. 45% in seventh grade, and approx. 65% in undergraduates).

Breakstone et al. (2019) assessed evaluation competence, specifically the ability to use source features to judge the quality of information. A representative sample of US high school students from grades 9th to 12th were presented with a set of Internet documents (Facebook posts, social media posts, or webpages), and they were asked to judge their credibility (e.g., Evaluate whether a video posted on Facebook is good evidence of voter fraud). Responses were coded using a rubric with three levels: beginning (incorrect responses), emerging (in the right direction but without proper reasoning), and mastery (effective evaluation by investigating sources and evidence). Overall, task performance was rather low, with participants obtaining emerging or mastery levels ranging across different tasks from 1% (min) to 39% (max). Regarding the development of evaluation competence, a small but continuous progression was observed across high school grades. Nevertheless, even 12th grade students scored well below mastery performance.

In sum, Internet reading competencies should be explicitly taught and practiced throughout the high school years. Several programs have been implemented to improve students' navigation, evaluation, and integration of Internet documents (for reviews see Barzilai et al., 2018; Brante and Strømsø, 2018; Salmerón and Llorens, 2019). In general, such programs have successfully increased high school students' specific competences. Still, current instructional programs do not fully consider the interrelated nature of Internet reading competencies. The three competencies are closely related and, potentially, students' engagement in

any one of them may support or trigger the other two (Salmerón et al., 2018a). Recent research efforts have identified possible types of adaptive coordination between the competencies (e.g., Delgado et al., 2020; Kammerer et al., 2016; Salmerón et al., 2020a; Sampietro and Salmerón, 2021), and there is a clear need to move from instruction addressing each competence in isolation toward instruction teaching students to self-regulate their knowledge construction during Internet reading by coordinating the required competencies.

Conclusion

As the nature of reading contexts in the 21st century changes, models of learning from multiple texts will need to be developed to provide a basis for representing and understanding how readers create meaning. Several individual difference variables play a key role in understanding how readers learn from multiple texts. The variables include (but are not limited to) comprehension skills, prior knowledge, beliefs, motivation and engagement, and strategic processing. Measurement of reading processes and the products of reading play a key role in understanding how readers learn from multiple texts and how individual differences impact learning. Similarly, by understanding how readers learn from multiple texts, interventions and instruction can be developed to promote engagement and learning. Importantly, readers not only need to be able to locate, comprehend, evaluate, and integrate texts, they need to be able to critique sources to whom texts are attributed (sourcing). Readers in the 21st century are consumers of information in an age dominated by rapid access to vast amounts of information. Thus, the challenge moving forward is to consider how readers can complement their roles as consumers of information to builders of knowledge.

References

- Andresen, A., Anmarkrud, Ø., Bråten, I., 2019. Investigating multiple source use among students with and without dyslexia. *Read. Writ.* 32 (5), 1149–1174. <https://doi.org/10.1007/s11145-018-9904-z>.
- Anmarkrud, Ø., Bråten, I., Strømso, H.I., 2014. Multiple-documents literacy: strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learn. Individ. Differ.* 30, 64–76. <https://doi.org/10.1016/j.lindif.2013.01.007>.
- Anmarkrud, Ø., Bråten, I., Florit, E., Mason, L., 2021. The role of individual differences in sourcing: a systematic review. *Educ. Psychol. Rev.* <https://doi.org/10.1007/s10648-021-09640-7>.
- Barzilai, S., Eshet-Alkalai, Y., 2015. The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learn. Instruct.* 36, 86–103. <https://doi.org/10.1016/j.learninstruc.2014.12.003>.
- Barzilai, S., Ka'adan, I., 2017. Learning to integrate divergent information sources: the interplay of epistemic cognition and epistemic metacognition. *Metacognition Learn.* 12 (2), 193–232. <https://doi.org/10.1007/s11409-016-9165-7>.
- Barzilai, S., Zohar, A.R., Mor-Hagani, S., 2018. Promoting integration of multiple texts: a review of instructional approaches and practices. *Educ. Psychol. Rev.* 30 (3), 973–999. <https://doi.org/10.1007/s10648-018-9436-8>.
- Barzilai, S., Mor-Hagani, S., Zohar, A.R., Shlomi-Elooz, T., Ben-Yishai, R., 2020a. Making sources visible: promoting multiple document literacy with digital epistemic scaffolds. *Comput. Educ.* 157. <https://doi.org/10.1016/j.compedu.2020.103980>. Article 103980.
- Barzilai, S., Thomm, E., Shlomi-Elooz, T., 2020b. Dealing with disagreement: the roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learn. Instruct.* 69. <https://doi.org/10.1016/j.learninstruc.2020.101367>. Article 101367.
- Braasch, J.L.G., Bråten, I., 2017. The discrepancy-induced source comprehension (D-ISC) model: basic assumptions and preliminary evidence. *Educ. Psychol.* 52 (3), 167–181. <https://doi.org/10.1080/00461520.2017.1323219>.
- Braasch, J.L.G., Graesser, A.C., 2019. Avoiding and overcoming misinformation on the Internet. In: Sternberg, R.J., Halpern, D.F. (Eds.), *Critical Thinking in Psychology*, second ed. Cambridge University Press, pp. 125–151.
- Braasch, J.L.G., Kessler, E.D., 2021. Working toward a theoretical model for source comprehension in everyday discourse. *Discourse Process.* <https://doi.org/10.1080/0163853X.2021.1905393>.
- Braasch, J.L.G., Rouet, J.-F., Vibert, N., Britt, M.A., 2012. Readers' use of source information in text comprehension. *Mem. Cognit.* 40, 450–465. <https://doi.org/10.3758/s13421-011-0160-6>.
- Braasch, J.L.G., Bråten, I., Strømso, H.I., Anmarkrud, Ø., 2014. Incremental theories of intelligence predict multiple document comprehension. *Learn. Individ. Differ.* 31, 11–20. <https://doi.org/10.1016/j.lindif.2013.12.012>.
- Braasch, J.L.G., Bråten, I., McCrudden, M.T. (Eds.), 2018. *Handbook of Multiple Source Use*. Routledge.
- Brante, E.W., Strømso, H.I., 2018. Sourcing in text comprehension: a review of interventions targeting sourcing skills. *Educ. Psychol. Rev.* 30 (3), 773–799. <https://doi.org/10.1007/s10648-017-9421-7>.
- Bråten, I., Strømso, H.I., 2011. Measuring strategic processing when students read multiple texts. *Metacognition Learn.* 6 (2), 111–130. <https://doi.org/10.1007/s11409-011-9075-7>.
- Bråten, I., Strømso, H.I., 2020. On the roles of dispositions and beliefs in learning from multiple perspectives. In: Van Meter, P., List, A., Lombardi, D., Kendeou, P. (Eds.), *Handbook of Learning From Multiple Representations and Perspectives*. Routledge, pp. 141–163.
- Bråten, I., Strømso, H.I., Britt, M.A., 2009. Trust matters: examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Read. Res. Q.* 44 (1), 6–28. <https://doi.org/10.1598/RRQ.44.1.1>.
- Bråten, I., Anmarkrud, Ø., Brandmo, C., Strømso, H.I., 2014. Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learn. Instruct.* 30, 9–24. <https://doi.org/10.1016/j.learninstruc.2013.11.002>.
- Bråten, I., Strømso, H.I., Ferguson, L.E., 2016. The role of epistemic beliefs in the comprehension of single and multiple texts. In: Afflerbach, P. (Ed.), *Handbook of Individual Differences in Reading: Text and Context*. Routledge, pp. 67–79.
- Bråten, I., Brante, E.W., Strømso, H.I., 2018a. What really matters: the role of behavioral engagement in multiple document literacy tasks. *J. Res. Read.* 41 (4), 680–699. <https://doi.org/10.1111/1467-9817.12247>.
- Bråten, I., Stadler, M., Salmerón, L., 2018b. The role of sourcing in discourse comprehension. In: Schober, M., Rapp, D.N., Britt, M.A. (Eds.), *Handbook of Discourse Processes*. Routledge, pp. 141–166.
- Bråten, I., Brante, E.W., Strømso, H.I., 2019. Teaching sourcing in upper-secondary school: a comprehensive sourcing intervention with follow-up data. *Read. Res. Q.* 54 (4), 481–505. <https://doi.org/10.1002/rrq.253>.

- Bråten, I., Braasch, J.L.G., Salmerón, L., 2020. Reading multiple and non-traditional texts: new opportunities and new challenges. In: Moje, E.B., Afflerbach, P., Enciso, P., Lesaux, N.K. (Eds.), *Handbook of Reading Research*, vol. V. Routledge, pp. 79–98.
- Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., Saavedra, A., 2019. Students' Civic Online Reasoning. A National Portrait. <https://stacks.stanford.edu/file/druid:gf151tb4868/Civic%20Online%20Reasoning%20National%20Portrait.pdf>.
- Britt, M.A., Sommer, J., 2004. Facilitating textual integration with macro-structure focusing tasks. *Read. Psychol.* 25 (4), 313–339. <https://doi.org/10.1080/02702710490522658>.
- Britt, M.A., Rouet, J.-F., Durik, A.M., 2018. Representations and processes in multiple source use. In: Braasch, J.L.G., Bråten, I., McCrudden, M.T. (Eds.), *Handbook of Multiple Source Use*. Routledge, pp. 17–33.
- Bromme, R., Goldman, S.R., 2014. The public's bounded understanding of science. *Educ. Psychol.* 49 (2), 59–69. <https://doi.org/10.1080/00461520.2014.921572>.
- Cho, B.-Y., Afflerbach, P., 2017. An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In: Israel, S.E. (Ed.), *Handbook of Research on Reading Comprehension*, second ed. Guilford, pp. 109–134.
- Cho, B.-Y., Woodward, L., Li, D., Barlow, W., 2017. Examining adolescents' strategic processing during online reading with a question-generating task. *Am. Educ. Res. J.* 54 (4), 691–724. <https://doi.org/10.3102/0002831217701694>.
- Delgado, P., Stang Lund, E., Salmerón, L., Bråten, I., 2020. To click or not to click: investigating conflict detection and sourcing in a multiple document hypertext environment. *Read. Writ.* 33, 2049–2072. <https://doi.org/10.1007/s11145-020-10030-8>.
- Du, H., List, A., 2020. Researching and writing based on multiple texts. *Learn. Instruct.* 66. <https://doi.org/10.1016/j.learninstruc.2019.101297>. Article 101297.
- Eccles, J., Wang, M., 2012. Part I commentary: so what is student engagement anyway? In: Christenson, S.L., Reschly, A.L., Wylie, C. (Eds.), *Handbook of Research on Student Engagement*. Springer, pp. 133–145.
- European Commission, 2018. Fake News and Disinformation Online. <https://europa.eu/eurobarometer/surveys/detail/2183>.
- Florit, E., Cain, K., Mason, L., 2019. Going beyond children's single-text comprehension: the role of word reading, working memory, comprehension monitoring and source use in 4th graders' multiple-document comprehension. *Br. J. Educ. Psychol.* 90 (2), 449–472. <https://doi.org/10.1111/bjep.12288>.
- Gil, L., Bråten, I., Vidal-Abarca, E., Strømsø, H.I., 2010. Summary versus argument tasks when working with multiple documents: which is better for whom? *Contemp. Educ. Psychol.* 35 (3), 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>.
- Hagerman, M.S., 2017. Disrupting students' online reading and research habits: the LINKS intervention and its impact on multiple Internet text integration skills. *J. Literacy Technol.* 18 (1), 105–156.
- Hämäläinen, E.K., Kili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., Leppänen, P.H.T., 2020. Promoting sixth graders' credibility evaluation of Web pages: an intervention study. *Comput. Hum. Behav.* 110. <https://doi.org/10.1016/j.chbh.2021.101985>. Article 106372.
- Kammerer, Y., Meier, N., Stahl, E., 2016. Fostering secondary-school students' intertext model formation when reading a set of websites: the effectiveness of source prompts. *Comput. Educ.* 102, 52–64. <https://doi.org/10.1016/j.compedu.2016.07.001>.
- Kammerer, Y., Gottschling, S., Bråten, I., 2021. The role of Internet-specific justification beliefs in source evaluation and corroboration during Web search on an unsettled socio-scientific issue. *J. Educ. Comput. Res.* 59 (2), 342–378. <https://doi.org/10.1177/0735633120952731>.
- Keil, F.C., Kominsky, J.F., 2013. Missing links in middle school: developing use of disciplinary relatedness in evaluating Internet search results. *PLoS One* 8 (6), e67777. <https://doi.org/10.1371/journal.pone.0067777>.
- Kili, C., Leu, D.J., Utraiainen, J., Coiro, J., Kanninen, L., Tolvanen, A., Lohvansuu, K., Leppänen, P.H.T., 2018. Reading to learn from online information: modeling the factor structure. *J. Literacy Res.* 50 (3), 304–334. <https://doi.org/10.1177/1086296x18784640>.
- Kintsch, W., 1988. The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95 (2), 163–182. <https://doi.org/10.1037/0033-295x.95.2.163>.
- Kintsch, W., 1998. *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Latini, N., Bråten, I., Anmarkrud, Ø., Salmerón, L., 2019. Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple document context. *Contemp. Educ. Psychol.* 59. <https://doi.org/10.1016/j.cedpsych.2019.101797>. Article 101797.
- Le Bigot, L., Rouet, J.-F., 2007. The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *J. Literacy Res.* 39 (4), 445–470. <https://doi.org/10.1080/10862960701675317>.
- Leu, D.J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., Timbrell, N., 2015. The new literacies of online research and comprehension: rethinking the reading achievement gap. *Read. Res. Q.* 50 (1), 37–59. <https://doi.org/10.1002/rq.85>.
- List, A., Alexander, P.A., 2017. Cognitive affective engagement model of multiple source use. *Educ. Psychol.* 52 (3), 182–199. <https://doi.org/10.1080/00461520.2017.1329014>.
- List, A., Alexander, P.A., 2019. Toward an integrated framework of multiple text use. *Educ. Psychol.* 54 (1), 20–39. <https://doi.org/10.1080/00461520.2018.1505514>.
- List, A., Alexander, P.A., 2020. Strategy use in learning from multiple texts: an investigation of the integrative framework of learning from multiple texts. *Front. Educ.* 5. <https://doi.org/10.3389/feduc.2020.578062>. Article 578062.
- List, A., Du, H., Wang, Y., Lee, H.Y., 2019a. Toward a typology of integration: examining the documents model framework. *Contemp. Educ. Psychol.* 58, 228–242. <https://doi.org/10.1016/j.cedpsych.2019.03.003>.
- List, A., Stephens, L.A., Alexander, P.A., 2019b. Examining interest throughout multiple text use. *Read. Writ.* 32 (2), 307–333. <https://doi.org/10.1007/s11145-020-10041-5>.
- List, A., 2020. Investigating the cognitive affective engagement model of learning from multiple texts: a structural equation modeling approach. *Read. Res. Q.* <https://doi.org/10.1002/rq.361>.
- Lundström, K., Diekema, A.R., Leary, H., Haderlie, S., Holliday, W., 2015. Teaching and learning information synthesis: an intervention and rubric based assessment. *Commun. Inf. Lit.* 9 (1), 60–82. <https://doi.org/10.15760/comminfolit.2015.9.1.176>.
- Macedo-Rouet, M., Braasch, J.L.G., Britt, A.M., Rouet, J.-F., 2013. Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cogn. Instr.* 31 (2), 204–226.
- Mahlow, N., Hahnel, C., Kroehne, U., Artelt, C., Goldhammer, F., Schoor, C., 2020. More than (single) text comprehension? On university students' understanding of multiple documents. *Front. Psychol.* 11. <https://doi.org/10.3389/fpsyg.2020.562450>. Article 562450.
- Mason, L., Scrimin, S., Tornatora, M.C., Suitner, C., Moé, A., 2018. Internet source evaluation: the role of implicit associations and psychophysiological self-regulation. *Comput. Educ.* 119, 59–75. <https://doi.org/10.1016/j.compedu.2017.12.009>.
- Mason, L., Zaccolletti, S., Scrimin, S., Tornatora, M.C., Florit, E., Goetz, T., 2020. Reading with the eyes and under the skin: comprehending conflicting digital texts. *J. Comput. Assist. Learn.* 36 (1), 89–101. <https://doi.org/10.1111/jcal.12399>.
- McCrudden, M.T., Schraw, G., 2007. Relevance and goal-focusing in text processing. *Educ. Psychol. Rev.* 19 (2), 113–139. <https://doi.org/10.1007/s10648-006-9010-7>.
- McCrudden, M.T., Stenseth, T., Bråten, I., Strømsø, H.I., 2016. The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: a mixed methods study. *J. Educ. Psychol.* 108 (2), 147–162. <https://doi.org/10.1037/edu0000057>.
- McCrudden, M.T., Huynh, L., Lyu, B., Kulikowich, J.M., 2021. Bridging inferences and learning from multiple complementary texts. *Discourse Process.* <https://doi.org/10.1080/0163853X.2021.1924586>.
- McCrudden, M.T., 2020. Processes and products of encountering belief-related information. In: Van Meter, P., List, A., Lombardi, D., Kendeou, P. (Eds.), *Handbook of Learning From Multiple Representations and Perspectives*. Routledge, pp. 191–204.
- Organization for Economic Co-operation and Development (OECD), 2009. PISA 2009 Assessment Framework—Key Competencies in Reading, Mathematics and Science. <http://www.oecd.org/dataoecd/11/40/44455820.pdf>.
- Perfetti, C.A., Rouet, J.-F., Britt, M.A., 1999. Towards a theory of documents representation. In: van Oostendorp, H., Goldman, S.R. (Eds.), *The Construction of Mental Representations during Reading*. Erlbaum, pp. 88–108.

- Potocki, A., De Pereyra, G., Ros, C., Macedo-Rouet, M., Stadler, M., Salmerón, L., Rouet, J.-F., 2020. The development of source evaluation skills during adolescence: exploring different levels of source processing and their relationships. *Infancia Y Aprendizaje* 43 (1), 19–59. <https://doi.org/10.1080/02103702.2019.1690848>.
- Primor, L., Katzir, T., 2018. Measuring multiple text integration: a review. *Front. Psychol.* 9, 2294. <https://doi.org/10.3389/fpsyg.2018.02294>.
- Richter, T., Maier, J., 2017. Comprehension of multiple documents with conflicting information: a two-step model of validation. *Educ. Psychol.* 52 (3), 148–166. <https://doi.org/10.1080/00461520.2017.1322968>.
- Rouet, J.-F., Britt, M.A., 2011. Relevance processes in multiple document comprehension. In: McCrudden, M.T., Magliano, J.P., Schraw, G. (Eds.), *Text Relevance and Learning From Text*. Information Age, pp. 19–52.
- Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M., Dinot, J., 2011. The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learn. Instruct.* 21 (2), 205–219. <https://doi.org/10.1016/j.learninstruc.2010.02.007>.
- Rouet, J.-F., Britt, M.A., Durik, A.M., 2017. RESOLV: readers' representation of reading contexts and tasks. *Educ. Psychol.* 52 (3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>.
- Sabatini, J., O'Reilly, T., Wang, Z., Dreier, K., 2018. Scenario-based assessment of multiple source use. In: Braasch, J.L.G., Bråten, I., McCrudden, M.T. (Eds.), *Handbook of Multiple Source Use*. Routledge, pp. 447–465.
- Salmerón, L., Llorens, A., 2019. Instruction of digital reading strategies based on eye-movements modeling examples. *J. Educ. Comput. Res.* 57, 343–359. <https://doi.org/10.1177/0735633117751605>.
- Salmerón, L., García, A., Vidal-Abarca, E., 2018a. The development of adolescents' comprehension-based Internet reading skills. *Learn. Individ. Differ.* 61, 31–39. <https://doi.org/10.1016/j.lindif.2017.11.006>.
- Salmerón, L., Gil, L., Bråten, I., 2018b. Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemp. Educ. Psychol.* 52, 25–35. <https://doi.org/10.1016/j.cedpsych.2017.12.002>.
- Salmerón, L., Delgado, P., Mason, L., 2020a. Using eye-movement modeling examples to improve critical reading of multiple webpages on a conflicting topic. *J. Comput. Assist. Learn.* 36, 1038–1051. <https://doi.org/10.1111/jcal.12458>.
- Salmerón, L., Sampietro, A., Delgado, P., 2020b. Using Internet videos to learn about controversies: evaluation and integration of multiple and multimodal documents by primary school students. *Comput. Educ.* 148. <https://doi.org/10.1016/j.compedu.2019.103796>. Article 103796.
- Sampietro, A., Salmerón, L., 2021. Incivility in online news and Twitter: effects on attitudes toward scientific topics when reading in a second language. *Lang. Sci.* 85, 101385. <https://doi.org/10.1016/j.langsci.2021.101385>.
- Scharrer, L., Salmerón, L., 2016. Sourcing in the reading process: introduction to the special issue. *Read. Writ.* 29 (8), 1539–1548. <https://doi.org/10.1007/s11145-016-9676-2>.
- Schwartz, D.L., Bransford, J.D., 1998. A time for telling. *Cognit. Instruct.* 16 (4), 475–522. https://doi.org/10.1207/s1532690xci1604_4.
- Spiro, R.J., Klautke, H., Johnson, A.K., 2015. How certain kinds of reading to learn on the Web are totally different from what we learned from research on traditional text comprehension and learning from text. In: Spiro, R.J., DeSchraver, M., Hagerman, M.S., Morsink, P.M., Thompson, P. (Eds.), *Reading at a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices*. Routledge, pp. 45–50.
- Stadler, M., Bromme, R., 2014. The content-source integration model: a taxonomic description of how readers comprehend conflicting scientific information. In: Rapp, D.N., Braasch, J.L.G. (Eds.), *Processing Inaccurate Information: Theoretical and Applied Perspectives From Cognitive Science and the Educational Sciences*. The MIT Press, pp. 379–402.
- Stang Lund, E., Bråten, I., Brandmo, C., Brante, E.W., Strømso, H.I., 2019. Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Read. Writ.* 32 (2), 335–356. <https://doi.org/10.1007/s11145-018-9868-z>.
- Strømso, H.I., Bråten, I., Samuelstuen, M.S., 2008. Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learn. Instruct.* 18 (6), 513–527. <https://doi.org/10.1016/j.learninstruc.2007.11.001>.
- Wang, Z., O'Reilly, T., Sabatini, J., McCarthy, K.S., McNamara, D.S., 2021. A tale of two tests: the role of topic and general academic knowledge in traditional versus contemporary scenario-based reading. *Learn. Instruct.* 73. <https://doi.org/10.1016/j.learninstruc.2021.101462>. Article 101462.
- Wiley, J., Griffin, T.D., Steffens, B., Britt, M.A., 2020. Epistemic beliefs about the value of integrating information across multiple documents in history. *Learn. Instruct.* 65. <https://doi.org/10.1016/j.learninstruc.2019.101266>. Article 101266.
- Wolfe, M.B.W., Goldman, S.R., 2005. Relations between adolescents' text processing and reasoning. *Cognit. Instruct.* 23 (4), 467–502. https://doi.org/10.1207/s1532690xci2304_2.
- Wolfe, M.B.W., Griffin, T.D., 2018. Beliefs and discourse processing. In: Schober, M.F., Rapp, D.N., Britt, M.A. (Eds.), *The Routledge Handbook of Discourse Processes*. Routledge, pp. 295–314.

Learning an additional language

Martin East, School of Cultures, Languages and Linguistics, Faculty of Arts, The University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	364
Contrasting theories of learning	364
Theories of learning and L1 acquisition	364
Translating theories of learning to the L2 context	365
A teacher-led model of CLT	366
A learner-centered model of CLT	367
The emergence of more balanced CLT models	368
Drawing conclusions for optimal learning in the classroom	369
Putting a communicative approach into practice	370
Conclusion	371
References	371

Introduction

There is broad consensus across the world that the essential purpose of learning an additional language (hereafter L2), at whatever level of the educational endeavor (primary, secondary, or tertiary), is for learners to be able to communicate with first language (L1) speakers in the target language (TL). In the late 1960s, this consensus began to give rise to the phenomenon that became known as Communicative Language Teaching (CLT), or the “communicative approach.” CLT represented an alternative to strongly grammar-focused or literature-oriented “academic” programs that had hitherto dominated L2 teaching and learning for many years (such as so-called grammar-translation). CLT is now well established worldwide as a leading paradigm for L2 pedagogy.

In the context of wider debates about human learning and development, CLT does not represent just one approach or method. Rather, it acts as an umbrella label for a range of approaches. These share in common the goal of enhancing learners’ communicative competence in the L2. However, over the years these approaches have developed different emphases and taken several contrasting directions. Summarizing significant issues reported in some detail in [East \(2021\)](#), the purpose of this article is to trace key aspects of the development of CLT, in theory and in practice, and thereby to provide a historical account of its emergence alongside contemporary implications. The article considers L2 learning in both naturalistic environments, or so-called *second* language contexts, where the TL and the language of the majority of the community may be the same (e.g., learning English as L2 in Australia), as well as so-called *foreign* language contexts where the TL is not the same as the language of the majority of the community (e.g., learning French as L2 in the UK). In both contexts, the focus is on what goes on to support language learning in classrooms (instructed L2 learning). The article also explores some of the tensions that have emerged as theorists, researchers, and pedagogues have sought to pinpoint optimal conditions for the learning of an L2 from a communicative perspective.

Contrasting theories of learning

A fundamental question for all those who have a stake in the contemporary L2 educational endeavor is “how do students best acquire an L2 for communicative purposes?” Attempts to answer this question have taken stakeholders in several directions as they have explored not only theories of second language acquisition (SLA) but also what has become known as the phenomenon of *instructed* SLA, or ISLA (see [Loewen, 2020](#), for a useful overview). [Mitchell et al. \(2019\)](#) argued that explorations of effective SLA sit within broader discussions about human learning in general. As they put it, “[h]ow much of human learning derives from innate predispositions, that is, some form of genetic pre-programming, and how much of it derives from social and cultural influences as we grow up?” (p. 11). The so-called *nature-nurture* debate that lies at the core of Mitchell et al.’s two questions has influenced (and continues to influence) pedagogical thinking and pedagogical practice in classrooms, both more generally and with specific reference to classroom-based L2 teaching and learning. Before considering the relevance of nature versus nurture for learning an L2 in the classroom, the next section explores how the debate has influenced our understanding of first language (L1) acquisition.

Theories of learning and L1 acquisition

The *nurture* side of the learning debate finds its starkest expression in the psychological theoretical standpoint referred to as behaviorism—a perspective on learning that has a very long history but was particularly pervasive in the US during the 1940s

and 1950s, and influenced, among others, by the work of the American psychologist Burrhus Frederic Skinner (1904–1990). Skinner proposed the theory of *operant conditioning*—the notion that behavior is determined by consequences: when a particular behavior is positively reinforced, it is likely to occur again; when it is punished, its re-occurrence is less likely.

When it comes to young children learning their L1, a typical behaviorist perspective would be that children come into the world as *blank slates* (*tabula rasa*) and all the knowledge they need to grow and develop can be written on those slates. From this standpoint, L1 acquisition occurs as a result of children imitating and practicing the sounds and language patterns they encounter in their environments. Successful instances of attempted communication (positive reinforcement) lead to habitual and ingrained patterns of language use that enable children to develop their L1 proficiency. Thus, from a behaviorist perspective, what is needed is for young children to be consistently exposed to, and imitate, accurate samples of language.

In the 1960s, the behaviorist stance on L1 acquisition was challenged by a contrasting theoretical perspective—innatism—which owed much to the work of Noam Chomsky (1928–date), an American theoretical linguist. Chomsky's theorizing began what Atkinson (2011) described as a “cognitive revolution” (p. 6) that represented not only “a direct response to American behaviorism” (p. 8), but also the *nature* side of the debate about L1 acquisition.

Critiquing Skinner's (1957) book on verbal behavior, Chomsky (1959) argued that behaviorism, as Lightbown and Spada (2013) explained it, “failed to account for ‘the logical problem of language acquisition’—the fact that children come to know more about the structure of their language than they could reasonably be expected to learn on the basis of the samples of language they hear” (p. 20).

That is, the language samples that children are exposed to in their environments are often incomplete, sometimes inaccurate, and do not represent all the permutations in which segments of language can be put together. Nevertheless, children do develop the skill to put isolated components of language together successfully and correctly—they can form complete chunks of language even if they have not necessarily encountered every linguistic permutation. Chomsky speculated on this basis that children's minds are *not* blank slates, and L1 acquisition cannot simply be explained by imitation, repetition, and positive reinforcement. Rather, in his view, the child's innate ability explained the acquisition—children can reason things out for themselves and construct correct patterns from individual pieces of language they have encountered.

Children, Chomsky theorized, have been biologically programmed to acquire language because they possess what he termed a Language Acquisition Device (LAD). Furthermore, they have been born with an innate ability to work out the rules of language for themselves because they can align the rules in some way to an internal template that contains the framework that governs all language systems—what Chomsky labeled Universal Grammar (UG). Thus, from an innatist perspective, the child's inborn capability is fundamental to L1 development.

Behaviorism and innatism represent two fundamentally different and radically oppositional stances on learning. Mitchell et al.'s (2019) “how much?” questions suggest, however, that effective learning is not necessarily down to an exclusive either-or; rather, it is a matter of degree. A third theoretical stance on human learning, which we could label *interactionist*, attempts to reconcile purely behaviorist or purely innatist standpoints.

An interactionist perspective on L1 acquisition considers the linguistic environment *in interaction with* the child's innate ability. It is built on the assumption that both components are necessary and suggests that language proficiency develops as a consequence of the complex interaction between the environments to which the child is exposed (home, school, etc.) and the child's unique human characteristics.

An interactionist stance on early learning in fact owes much to the theorizing of two influential psychologists of the very early 20th century—Jean Piaget (1896–1980) from Switzerland and Lev Vygotsky (1896–1934) from Russia. The work of these two illustrates somewhat different positions on the nature-nurture interface as far as L1 development is concerned: Piaget seemed to place greater emphasis on children's internal cognitive processing ability, and, for him, a crucial element was children's active role in the learning process through self-discovery (see, e.g., Piaget, 1926). Vygotsky seemed to be more concerned with the influence of the environment, and his sociocultural-interactionist theory of learning (see, e.g., Vygotsky, 1978) put learning down to a social process and collaborative interaction with more capable peers who would provide the child with *scaffolding* which supports the child's evolving L1 development.

Translating theories of learning to the L2 context

It might be reasonable to assume at this point that an interactionist perspective on learning is a more reasonable or logical proposition than a separatist either-or standpoint (and this proposition will be revisited later in this article). Nevertheless, and turning now to the historical development of classroom-based L2 learning approaches, a range of perceptions demonstrates that, in both theory and practice, the situation has not been clear-cut. The next section traces developments to the CLT paradigm, highlights their strengths and limitations, and draws conclusions about optimal learning conditions in the L2 classroom. This historical overview is designed to illustrate the theoretical drivers for the diverse ways in which CLT has been operationalized over the years.

Benson and Voller's (1997) succinct summarization of the emergence of CLT underscored its significance at the time. They noted:

From time to time, a new concept enters the field of language education as an alternative method or approach, but rapidly grows in significance to the point where it comes fundamentally to condition thinking throughout the field. Such was the case with Communicative Language Teaching ... which began life in the late 1960s as an alternative to "structural" and "grammar-translation" models of teaching, but rapidly became an axiom of language teaching methodology (p. 10).

Benson and Voller (1997) continued that, against a backdrop of increasing recognition that L2 learners needed to acquire skills in communicating with others across the globe, CLT turned its back on "traditional anglocentric assumptions" that the overriding purpose of learning an L2 was "to broaden the mind" (as realized, e.g., through the study of literature in the TL). Instead, CLT "focused attention on learners who were learning languages because they needed to *use* them in an ever-shrinking world" (p. 11). As East (2008) put it, CLT represented "a distinct move away from artificiality of language with its emphasis on frequently decontextualized grammatical structures and vocabulary learning." It represented, as an alternative, "a move towards an understanding that language exists for purposes of *real* communication with *real* individuals in *real* contexts" (p. 14). The genesis of the communicative movement was traceable to simultaneous developments in both Europe and the US (Savignon, 1991), thereby becoming informed by a broad range of views about the nature of both language and language learning. It soon became an *international* initiative due to its attempt to respond to the perceived needs of L2 learners in a range of contexts.

Quite apart from substantial increases in travel and migration for which language in use was becoming important, the 1960s represented a turbulent time in history for many other reasons. Societally, this period witnessed significant challenges to the status quo from several quarters (Christison and Murray, 2014). In the context of education, at least in the US, so-called "Romantic Radicalism" challenged and denigrated teacher-dominated approaches and experimented with child-centered schools and programs, and experiential learning (Glatthorn et al., 2018). Chomsky's ideas represented arguably the most radical counter-influence to behaviorist stances on learning to arise in the 1960s, at least with regard to acquiring a language. As CLT began to emerge, Chomsky's theorizing around language acquisition became influential in the critiquing of strongly teacher-led pedagogical models. However, as the CLT movement began to build in momentum during the 1970s, there were inevitably tensions as more traditionally-oriented teachers, who may or may not have wanted to step onto the CLT bandwagon, grappled with what CLT might look like in their thus far grammar-oriented classrooms. Thus, in light of all else that was happening at the time, CLT had the potential to become a radical phenomenon for language education, heralding sweeping changes to practice. The reality as CLT began to unfold was more subtle and more complex (see, e.g., Medgyes, 1986; Rose, 2019).

Part of the early process of attempts to build a level of cohesion for the communicative movement was the articulation of *communicative competence* as the underlying theoretical framework. One early influential model, based on CLT principles that were beginning to become apparent and accepted, was presented by the Canada-based researchers Canale and Swain at the start of the 1980s. Although their model has been developed by others (the introductory chapter in Walker et al., 2018, presents a useful historical overview), the model represents a concise formulation of what Canale and Swain (1980) described, at this formative time in CLT's history, as "the content and boundaries of communicative competence" which, they maintained, would lead to "more useful and effective second language teaching" (p. 1). There were four dimensions (Canale, 1983):

1. *Grammatical or formal competence*, or knowing the rules;
2. *Sociolinguistic competence*, or knowing the language that is appropriate for the context;
3. *Discourse competence*, or knowing how to engage with extensive use of L2 in context;
4. *Strategic competence*, or knowing what to do when you do not know something in the L2.

This useful theoretical framework was, however, differentially realized as CLT began to become more widely influential in classrooms.

A teacher-led model of CLT

By the 1960s, strongly behaviorist-influenced pedagogical approaches to classroom-based L2 learning had become the status quo (Richards, 2006). In these approaches, the teacher as expert and leader would impart knowledge and explain principles (e.g., how grammar works); the learners would sit passively and, as *tabula rasa*, soak up the imparted knowledge; they would then practice what they had been taught (e.g., through grammar exercises). Drilling, repeating, and memorizing would be key components, and feedback on language use would focus on accuracy, with a view to positive reinforcement (i.e., praising correct language use and correcting language errors so that patterns of correct language use could become habitual).

Theorizing the behaviorist-informed stance from the perspective of grammar learning, Long (e.g., 2000) described the approach in the top-down, teacher-led classroom as *focus on forms* (FonFS). FonFS, as the name suggests, pays specific attention to the forms or rules of the language. In the grammar-translation classroom, this would have been to the exclusion of any communicative language practice. The grammatical forms to be taught would be presented in a textbook or synthetic syllabus, and ordered in a hierarchy from simple to more complex. Decisions around sequencing were often made on intuitive or arbitrary rather than objectively

defensible grounds (i.e., teachers or textbook writers making assumptions about what is grammatically straightforward and what is less so). The rules would be presented one at a time and practiced through grammar drills, exercises, and translation, alongside memorization of chunks of language.

A form of CLT that emerged early in the development of the communicative paradigm became known as the “weak” model. Weak CLT essentially reflected many of the teacher-led principles that had informed earlier approaches. In essence, the ostensible focus was communicative, with learners being introduced to situations or contexts for communicative language use (e.g., what to say in a café/hotel/airport) and/or different functions for which the L2 might be required (e.g., asking for something, saying sorry, negotiating). Weak CLT tended to foreground memorization of the correct language for the circumstances; it also foregrounded grammar teaching (as in grammar-translation). This was often through the “classic lesson structure” of Presentation-Practice-Production or PPP (Klapper, 2003)—first, teachers *presented* the grammatical rule; second, learners *practiced* the grammatical rule through grammar exercises; third, learners *produced* the grammatical rule, often through some kind of structured role-play undertaken in pairs or groups (e.g., ordering food and drink in the café; checking into the hotel). The communicative activity was designed to practice the rule. Weak CLT thus favored accuracy over fluency. Although dimensions that were later to find their way into Canale and Swain’s articulation of communicative competence were ostensibly attended to (e.g., grammatical competence, sociolinguistic competence), the pedagogical approach was essentially top-down, teacher-led, and behaviorist-informed.

In Long’s (2000) words, lessons based on an FonFS model, which focused in a linear way on linguistic items which students were required to learn and use accurately, tended to be “rather dry” (p. 181) and to “produce boring lessons, with resulting declines in motivation, attention, and student enrollments” (p. 182). Crucially, from a communicative perspective, there was little, if any, opportunity to use language creatively or spontaneously beyond the confines of the practiced language, and the development of higher levels of communicative competence was thereby compromised.

A learner-centered model of CLT

Strong interest in an innatist-informed learner-centered and experiential approach to L2 learning was fueled by the theorizing of the American linguist Stephen Krashen (1941–date), who, beginning in the 1970s, essentially toyed with Chomsky’s theories concerning the LAD and UG, but translated these ideas from the L1 to the L2 context. Krashen (1981, 1982, 1985) proposed that acquiring an L2 for communicative purposes was an entirely separate process from learning the rules of the L2. Furthermore, he theorized that *only* acquired language can contribute to communicative competence, and what is learned cannot be turned into acquired language. Krashen also theorized that L2 learners acquired the L2 through input alone, or, more precisely, through *understanding* the language they were receiving and processing in the classroom (the input hypothesis). *Using* the acquired language was a subsequent step.

Following Krashen’s theorizing, direct instruction about the L2 would be seen as unnecessary. Rather than sitting passively in class, learners would need to play a central role in acquiring the L2. Acquisition would be dependent on learners’ ability to think through and construct for themselves the rules underpinning the input. Fundamental here would be a great deal of contact with (or immersion in) the L2. All that would be necessary for L2 acquisition to occur would be exposure to input and opportunities for L2 learners to notice the rules for themselves.

Krashen’s ideas were foundational to the development of the Natural Approach (Krashen and Terrell, 1983; Terrell, 1977, 1982), so-called because it assumed that acquisition of an L2 followed the “natural acquisition processes” (as theorized by Chomsky) that were required to acquire the L1 (Terrell, 1982, p. 122). Within the Natural Approach, actually *using* the language communicatively was important. However, language use was not forced; rather, room was given for language use to emerge spontaneously after learners had had ample opportunities to process comprehensible language input.

Long (2000) described an innatist-informed approach to grammar in the bottom-up, student-centered classroom as focus on meaning (FonM). For Long, FonM, in stark contrast to FonFS, represented “a radical pendulum swing” (p. 182), and a stance taken by those who perceived significant limitations in FonFS, at least as far as communicative competence was concerned. The notion assumed (as Krashen did) that learners could decipher the grammar for themselves, just by being exposed to sufficient input in the TL.

So-called “strong” CLT emerged as a fiercely learner-centered reaction against teacher-dominant approaches, which were perceived as inadequate due to their emphases on repetition and memorization, grammatical accuracy, and controlled and limited samples of language. In its strongest form, emphasis would be placed exclusively on communication and group work. There would be no overt attention to the grammatical rules, based on the principle of immersion in the L2. Learners would be exposed to a range of authentic language input, with huge reliance placed on learners’ own cognitive processing ability. Consequently, strong CLT could be described as wholly experiential, and favoring fluency over accuracy. Although broader dimensions of Canale and Swain’s articulation of communicative competence were supposedly being addressed, the development of grammatical competence was largely considered to be implicit—a stance that, at its worst, could be construed as an “‘anything-goes-as-long-as-you-get-the-message-across’ approach to second language teaching” (Savignon, 1983, p. 1).

Long (2000) commented that FonM lessons are “purely communicative,” and conceded that they are “often interesting, relevant, and relatively successful” (p. 183). Furthermore, it may be argued that an FonM approach may lend itself better to second language contexts where opportunities are available to use the TL outside the classroom. However, Long went on to cite Swain (1991), whose investigations into the impact of learning French in the immersion context of bilingual Canada revealed that, even after 12 years of language immersion in the classroom, students’ command of grammatical principles remained sub-optimal—presumably because

they had simply failed to notice the correct forms in the input for themselves. Also, in classroom contexts where the L2 being taught is not the same as the L1 shared by many in the general population (i.e., *foreign* language contexts), there is arguably simply not enough time to expose learners to sufficient input, to “immerse” them in the language.

In summary, behaviorism and innatism (or nurture versus nature), as realized, for example, in the weakest or strongest forms of CLT, represented “two very contrasting approaches that inform teaching and learning—with the behaviourist approach being top-down, teacher-led and expository (teach students everything they need to know) and the innatist approach being bottom-up, learner-centred and experiential (let students work it all out for themselves)” (East, 2021, p. 8). However, it seems that both approaches are limited when it comes to determining the optimal conditions for classroom-based L2 learning from a communicative perspective to occur. Long (2015b) captured the essence of the problem, stating the “two major traditional language teaching orthodoxies” represented by CLT in its weakest and strongest forms were “excessively interventionist, on the one hand, and irresponsibly, wholly non-interventionist, on the other” (p. 20).

The emergence of more balanced CLT models

Earlier in this article, the suggestion was put forward that interactionism provides a more balanced and inclusive perspective on the components that are necessary for effective learning than the either-or proposals represented in behaviorism or innatism. Furthermore, interactionist perspectives as theorized, for example, by Piaget and Vygotsky, although somewhat different in emphasis, have contributed to the emergence of *constructivism* as a dominant paradigm for education in instructed contexts.

Constructivism emphasizes learners’ *active construction* of their own knowledge, in contrast to passively receiving information from teachers or textbooks. From this standpoint, primarily instructing learners (as in the weak CLT classroom) is not sufficient. Learners require opportunities to construct their own meanings, whether individually or working with others, and “raise their own questions, generate their own hypotheses and models as possibilities and test them for validity” (Fosnet, 1996, p. 29). Nonetheless, in the constructivist classroom, teachers have facilitative and supportive roles to play. From this standpoint, simply letting students get on with it (as in the strong CLT classroom) is also not sufficient. As Weimer (2013) noted, constructivist-oriented teachers do impart knowledge directly to students. The difference, however, is that usually “this direct transfer of information occurs *after* students have grappled with the issue, *after* they have a sense of what it is they need to know” (p. 23, my emphasis). The purpose of holding back from an initial period of instruction is that “once students realize they need to know something, they listen attentively for the answers” (p. 23).

How, then, has an interactionist/constructivist perspective on learning influenced CLT? This article has already presented the early and innatist-influenced ideas of Krashen. As non-interventionist ideas became subject to critique, developments to Krashen’s input hypothesis began to emerge that increasingly saw crucial roles for interactions between learners and interventions by teachers. Swain’s output hypothesis (Swain, 1985, 1995, 2005) was a developed notion that represented “an eloquent assault upon the input hypothesis” (Nunan, 2004, p. 80). Swain argued that learners of an L2 required opportunities to *produce* language through speaking or writing—that is, create output. When L2 users try to communicate something, they often come across barriers—items of language that they do not understand or do not know. When they become aware that they do not know something, they may amend what they are trying to communicate to get their message across more effectively and make it more understandable. In this sense, they “modify output.” Furthermore, when L2 users are *pushed* to create more accurate output, they may receive feedback on that output which helps them to notice what they do not know and take steps to overcome the gap. This process, according to Swain, helps learners to acquire more language and enhance the accuracy of their language because, as a consequence of feedback, they learn something new about how the L2 works. Implicit in Swain’s conceptualization was interaction—the means through which feedback and instruction could be offered.

Long took the matter further in his proposal of the interaction hypothesis (Long, 1981, 1983a,b, 1985, 1996). Long’s work in this space was so crucial that he has justifiably been described as “one of the founders of the modern Cognitive-Interactionist Tradition” (Markee, 2015, p. 23). Long argued that conversational interaction was an essential, or even sufficient, condition of L2 acquisition. The cognitive-interactionist perspective was built on the argument that communicative competence was essentially developed through face-to-face communication, and that learners learn how to communicate more effectively in the L2 as they seek to make their meaning clear to others (i.e., they do not just *modify* output—which might suggest a monologic activity; they *negotiate meaning*—implying at least two interlocutors). In this process, Long (1996) saw the need for corrective feedback that might be given by a teacher or more capable peer. This feedback would enable L2 users to undertake output modifications that would help them to negotiate meaning more successfully.

Long (2000) also theorized how grammar might be attended to from a cognitive-interactionist perspective, putting forward the notion of *focus on form* (FonF). FonF provides the space for teachers (or peers) to draw L2 learners’ attention to aspects of language, for example, language they are using incorrectly in their attempts to communicate. Crucial for Long, however, was that the attention paid to form must be in the context of L2 learners’ attempts to use language communicatively. That is, the grammar is not pre-selected, ordered, and taught in advance (as in FonFS); nor are learners left entirely to their own devices to work out the rules (as in FonM). Rather, teachers and learners both have roles to play in the co-construction of knowledge—in this case, the grammatical knowledge required for successful communication to take place. The aim is to induce noticing (Schmidt, 1990, 1993, 2001). The rules, whether noticed directly by the learner or indirectly through teacher or peer instruction/feedback, become available for acquisition. That is, “what learners notice in input is what becomes intake for learning” (Schmidt, 1995, p. 20).

An interactionist perspective on classroom-based L2 learning has also been influenced by the sociocultural-interactionist stance proposed by Vygotsky. From Vygotsky's socially-oriented perspective, learners must be enabled to work within their individual Zone of Proximal Development (ZPD), where the ZPD represents the space in which learners can move from what they initially cannot do on their own to what they can do independently as they collaborate with other, more capable peers and are scaffolded by them.

Thus, both cognitive-interactionist and sociocultural-interactionist perspectives place emphasis on meaningful interactions between L2 learners. In both cases, group work between learners would be a dominant feature of practice, but there would also be room for teachers (or peers) to intervene with instruction, explanations, or feedback. In other words, a constructivist approach to L2 teaching and learning is anticipated.

As communicatively-oriented theorists and practitioners began to wrestle with the crucial significance of *interaction* for purposes of SLA, the importance of interaction in the L2 became increasingly embraced in classrooms. Brown (2014), for example, spoke of a "turn-of-the-century wave of interest," influenced by a social-constructivist viewpoint, that saw increasing emphasis on three dimensions: language came to be viewed as "interactive communication among individuals, each with a sociocultural identity;" teachers were seen as "treating the language classroom as a locus of meaningful, authentic exchanges among users of language;" and L2 learning was perceived as "the creation of meaning through interpersonal negotiation among learners" (p. 206). In some instances, this has seen the rather limited "communicative" activities of the weak CLT classroom (e.g., role-playing buying food and drink in a café) develop into more open activities that allow greater room for creative language use.

One form of CLT to emerge around the early 1980s that has been heavily influenced by both cognitive-interactionist and sociocultural-interactionist perspectives is referred to as task-based language teaching or TBLT (see, e.g., Ellis et al., 2019). Essentially, TBLT sees "important roles for holism, experiential learning, and learner-centered pedagogy" alongside "the interactive roles of the social and linguistic environment in providing learning opportunities, and scaffolding learners into them" (Norris et al., 2009, p. 15). The construct of "task" represents "the hundred and one things that people *do* in everyday life at work, at play, and in between" (Long, 1985, p. 89) or "the real-world activities people think of when planning, conducting, or recalling their day" (Long, 2015a, p. 6). In TBLT, tasks are fundamental to interactive classroom activities and create greater freedom for learners to experiment with real language than the more structured role plays of the weak CLT classroom. Additionally, *task-based* suggests that these communicative tasks play a central role. Consequently, FonF has become a strongly advocated model for attention to the more formal aspects of language within TBLT.

It must be acknowledged that, over the years, TBLT has grown and developed in different directions that have emphasized elements of either weaker or stronger forms of CLT; it is beyond the scope of this article to explore these developments but see East (2021) for a more detailed discussion.

Drawing conclusions for optimal learning in the classroom

The above historical presentation of the CLT paradigm may have left some readers with the impression that developments to the paradigm have followed a logical and linear progression from a strongly behaviorist-informed teacher-dominant model (weak CLT) to a radically different innatist-influenced learner-centered and experiential model (strong CLT), and settling on a balanced model that fits within a constructivist view of learning and that consequently sees important roles for both teachers and learners in the L2 learning endeavor. However, the purpose of presenting the historical developments has in fact been to reach the conclusions that different historical emphases are still influential, and that there is value in considering dimensions of each of the stances on learning that have been outlined. As Van den Branden et al. (2009) observed, the pedagogical reality in many L2 classrooms is that teachers operate within "gradations of pedagogical choice," rather than within "two mutually exclusive extremes" (p. 3). These gradations of choice have been influenced as teachers have wrestled with messages about, and interpretations and applications of, sometimes very conflicting theoretical standpoints. The gradations are also influenced by the *context* for instructed L2 learning—where the TL and the language of the majority of the community may be the same (second language), or where the TL is not the same as the language of the majority of the community (foreign language).

Looking at educational approaches more generally, constructivism has secured a strong place as a theory of learning supported by research evidence that offers "a convincing commendation of learner-centered approaches" (Weimer, 2013, p. 54). It has nonetheless not been beyond critique. Coe et al. (2014), for example, asserted a counterclaim that "research evidence ... broadly favours direct instruction" (p. 23). Furthermore, it seems that behaviorist-informed principles continue to find a place in discussions about effective teaching and learning (see, e.g., Fischer and Hänze, 2019). These broader tensions in education are also reflected in the literature on effective approaches to L2 pedagogy. Adamson (2004), for example, argued that language *teaching* is "a complex undertaking," precisely because it is "shaped by views of the nature of language, of teaching and learning a language specifically, and of teaching and learning in general" (p. 604). Mitchell et al. (2019) maintained that L2 *learning* is "an immensely complex phenomenon" (p. 2), precisely because it is "coloured by debates on fundamental issues in human learning more generally" (p. 11).

It must also be acknowledged that the weak CLT model that was forcefully critiqued by those who advocated for learner-centered, experiential practices in light of the Chomskyan revolution of the 1960s in fact became, by the 1980s, "more or less standard practice" (Howatt, 1984, p. 279). By contrast, strong CLT did not really make huge inroads (see, e.g., Medgyes, 1986). Notwithstanding the emergence of TBLT as one means to address this imbalance, behaviorist-informed principles do continue

to be used in L2 classrooms. The direct teaching of grammar, alongside formal practice of grammatical rules, remains commonplace in many contexts (Larsen-Freeman, 2015). So too do vocabulary learning, repetition of phrases and sentences, pronunciation drilling, and memorized dialogues. These elements of L2 pedagogy cannot necessarily be dismissed as outmoded, irrelevant, or ineffective. Nonetheless, they must be viewed as part of a broader package.

Richards (2006) provided several perspectives on CLT that still hold true today. Although acknowledging that CLT finds its locus in “the notion of communicative competence as the goal of second and foreign language teaching,” Richards spoke of the wide and varied implementation of CLT approaches in attempting to reach that goal. Furthermore, as Richards made clear, CLT is influenced by a range of different educational paradigms and traditions. This means in practice that “there is no single or agreed upon set of practices that characterize current communicative language teaching.” Rather, CLT “refers to a set of generally agreed upon principles that can be applied in different ways, depending on the teaching context, the age of the learners, their level, their learning goals, and so on” (p. 22). This eclecticism in practice is reflected more recently in Mitchell et al.’s (2019) assertion that there can be “no ‘one best method’, however much research evidence supports it, which applies at all times and in all situations, with every type of learner” (p. 406). Indeed, and as acknowledged at the start of this article, CLT represents a broadly-based *approach*, rather than a specific *method*. In fact, in an era that some have described as “post-methods,” principled eclecticism which incorporates selecting methods deemed to be appropriate to the teaching and learning context is encouraged (Richards and Rodgers, 2014).

Putting a communicative approach into practice

Taking all of the above into account, which principles might help guide teachers to create optimal conditions for learning in instructed L2 contexts, whether second or foreign language?

Fundamentally, and drawing on the different theoretical perspectives presented above, one thing is clear: if L2 learners are to develop their communicative competence, there needs to be plenty of exposure to authentic TL input, and plentiful opportunities for students to interact with others, in the TL, around that input. It is in this respect that the L2 context arguably makes a difference.

In second language contexts, opportunities for L2 learners to interact authentically with TL speakers beyond the classroom are plentiful (in theory at least), and learners can be immersed genuinely in language and can try things out for themselves. In practice, although outside-of-class opportunities may be there, learners, for a variety of reasons, may not necessarily take full advantage of these. However, interactions in the classroom can build on the context, perhaps even linking classroom-based work with specific outside-of-class language use goals so as to help learners to take full advantage of authentic communicative opportunities (e.g., learners might be supported to undertake, whether as individuals, pairs, or groups, specific outside-of-class projects that require negotiating meaning with L1 speakers). In foreign language contexts, by contrast, there may be minimal or no opportunities for L2 learners to interact authentically with TL speakers beyond the classroom. In these cases, emphasis will need to be placed on creating ample opportunities for peer-to-peer interactions in the classroom. In both contexts, interactional opportunities present means through which learners can develop (among other things) their sociolinguistic, discourse, and strategic competence.

Attention to form (i.e., the development of grammatical competence) has also been established as an important component of L2 learning. Following on from ample opportunities to use the TL communicatively, the L2 context will also make a difference when it comes to the attention paid to the rules of language. In second language classrooms, there may be less necessity for TL interactions to dominate, and grammar instruction and practice may become more of a focus. In foreign language classrooms, grammar instruction, although important, may need to become more peripheral so as not to detract from interactional opportunities. In both contexts, however, where an important emphasis of classroom activity is on interactions between pairs and groups, teachers have crucial roles to play with regard to grammar. That is, instruction in, and feedback on, language in use are important components of L2 learning, whatever the context. The purpose of instruction and feedback is to enable learners to move from one level of competence to a higher level of competence in using the L2. The question becomes how to cater for direct instruction and corrective feedback in ways that support opportunities for learners to communicate authentically but that also help learners to learn how to handle language better (i.e., more accurately).

There would appear to be no advantage in dealing with grammar abstractly as a specific phenomenon of study (as in FonFS). That is not to say, however, that there is no place for the direct teaching of grammar rules or discrete practice of the rules. It is to say that grammar should be approached in the context of its contribution to effective (or accurate) communication.

A *communicatively*-oriented approach to grammar may well arise from what learners notice about language as they engage with TL input and/or interact with others in communicative contexts (FonF). In this sense, the focus on grammar is integrated into communicative language use. This approach might include both feedback on language use, either during or after the interactive activity, and direct instruction regarding a communicatively-relevant grammar feature after learners have attempted to use the language. This may perhaps involve drawing explicit attention to a feature that appeared to have been consistently or largely misapplied when learners engaged in interactive classroom-based activities, and using this feature in subsequent isolation from the communicative interaction as a focus for whole-class teacher input. Valeo and Spada (2016), for example, argued that each approach (integrated and isolated) is valuable for L2 acquisition, asserting that “recognition of the value of both approaches is consistent with research and theory claiming that each option may play a different role in L2 learning” (p. 329). This does not preclude the judicious use of grammar practice exercises, which may serve the function of consolidation.

Teachers may also perceive the need, from time to time, to provide direct instruction and practice in particular TL features in *preparation* for communicative activity. Thus, grammar might also be attended to explicitly through some form of direct teaching or instruction, alongside discrete practice. This replicates some dimensions of an FonFS approach or the PPP sequence. However, provided that the grammar being attended to is directly relevant to the communicative goals the teacher has in mind, there may be value, at least for some learners, in this initial top-down, teacher-led strategy.

East (2018) suggested that one way of achieving direct instruction that does not compromise plentiful in-class opportunities for L2 learners to use the TL creatively, meaningfully, and collaboratively is to “flip the classroom.” In a flipped classroom model, direct teaching can be facilitated by, for example, “finding or creating short instructional videos that carefully introduce relevant grammatical principles, coupled with practice material that helps learners to consolidate their learning” (p. 114). East concluded that getting the most benefit out of learning in the classroom will require a balance between language instruction and language use (or, teacher-led and learner-centered moments). A flipped classroom model “does not negate the need, from time to time, for explicit instruction, but capitalises on opportunities for meaningful interaction by relegating much of the preparatory work to outside the class” (p. 116). Also, giving learners a good deal of ownership for their own learning means that opportunities are provided for them to work both collaboratively and independently. (Certainly, as teachers across the globe become more familiar with online formats for language instruction as a result of COVID-19, teacher expertise in online asynchronous [not in real time] teaching is growing, and more imaginative and creative ways of providing learners with crucial input are emerging.)

When it comes to fluency versus accuracy, the reality remains that finding the most appropriate balance between meaning and form is challenging. Long (2015a), for example, argued, “[t]he jury is still out on optimal uses and timing of various kinds and combinations of instruction (explicit, implicit, focus on form, focus on forms, etc.)” (p. 26). Certainly, theoretical frameworks for optimal learning (as outlined earlier in this article) are important and should be taken into account. At the end of the day, however, the context will be a crucial element in teachers’ decision-making. This reality may be unsettling for teachers who may wish to have more robust theoretical guidance on optimal conditions for learning. The teacher remains nonetheless “the expert on the local classroom situation, after all” (Long, 2000, p. 188). Teachers know their contexts, their classes, and their students, and, as Thomas and Brereton (2019) expressed it, often just “prefer to *get on with it*” (p. 276) and “go with their own instincts regarding what works” (p. 278). Eclecticism in practice, albeit informed by the theoretical standpoints and arguments presented here, provides the space for this to happen.

Conclusion

The purpose of this article has been to trace the historical development of CLT in both theory and practice, with a view to highlighting some of the tensions that have emerged. Eclecticism in practice is the current reality in many L2 classrooms. Nevertheless, certain principles for maximizing learning have been drawn out from a consideration of sometimes contrasting theories of effective pedagogical practice.

In a recent article that acknowledged the considerable evolution and expansion of CLT since its inception, Spada (2018) argued that “21st century CLT” represents “a greater balance, scope and depth” (p. 12). This includes the interface between direct instruction and communicative practice in learner-centered classrooms that have been explored in this article. It must also be acknowledged that, even though principles such as input, output, and interaction appear to focus principally on the conversations that students might undertake in the communicative classroom, CLT is not just about speaking. CLT embraces a range of skills, which may well be integrated (e.g., responding in speaking to listening or reading input). Furthermore, there are some activities (such as writing) that can be undertaken collaboratively, but that also lend themselves particularly well to independent work, both in and beyond the classroom.

CLT also increasingly embraces the need for students to develop skills in negotiating cultural differences, or so-called *intercultural communicative competence*—a matter that is beyond the scope of this article (but see Walker et al., 2018, for some discussion of this). However, communicative interactions, particularly between L2 learners and L1 speakers, provide significant opportunities for noticing and negotiating cultural differences that may inform not only language choice (i.e., sociolinguistic competence) but also appropriate behavior, or taking into account in communication “the culture-specific context embedding the norms, values, beliefs, and behaviour patterns of a culture” (Alptekin, 2002, p. 58).

Spada (2018) concluded that L2 teaching and learning in today’s globalized world remains “important, indeed crucial work” (p. 12). As such, it is important to consider the factors and elements that are likely to optimize the learning of an L2 in the classroom, whatever the context.

References

- Adamson, B., 2004. Fashions in language teaching methodology. In: Davies, A., Elder, C. (Eds.), *The Handbook of Applied Linguistics*. Wiley-Blackwell, pp. 604–622.
- Alptekin, C., 2002. Towards intercultural communicative competence in ELT. *ELT J.* 56 (1), 57–64.
- Atkinson, D., 2011. Cognitivism and second language acquisition. In: Atkinson, D. (Ed.), *Alternative Approaches to Second Language Acquisition*. Routledge, pp. 1–23.

- Benson, P., Voller, P. (Eds.), 1997. *Autonomy and Independence in Language Learning*. Longman.
- Brown, H.D., 2014. *Principles of Language Learning and Teaching*, sixth ed. Pearson.
- Canale, M., Swain, M., 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Appl. Linguist.* 1 (1), 1–47.
- Canale, M., 1983. On some dimensions of language proficiency. In: Oller, J.W.J. (Ed.), *Issues in Language Testing Research*. Newbury House, pp. 333–342.
- Chomsky, N., 1959. Review of Skinner's verbal behavior. *Language* 35 (1), 26–58.
- Christison, M.A., Murray, D., 2014. *What English Language Teachers Need to Know*, vol. III. Routledge.
- Coe, R., Aloisi, C., Higgins, S., Elliot Major, L., 2014. *What Makes Great Teaching? Review of the Underpinning Research*. The Sutton Trust.
- East, M., 2008. *Dictionary Use in Foreign Language Writing Exams: Impact and Implications*. John Benjamins.
- East, M., 2018. Learning in the classroom. In: Burns, A., Richards, J.C. (Eds.), *The Cambridge Guide to Learning English as a Second Language*. Cambridge University Press, pp. 110–117.
- East, M., 2021. *Foundational Principles of Task-Based Language Teaching*. Routledge.
- Ellis, R., Skehan, P., Li, S., Shintani, N., Lambert, C., 2019. *Task-Based Language Teaching: Theory and Practice*. Cambridge University Press.
- Fischer, E., Hânze, M., 2019. Back from “guide on the side” to “sage on the stage”? Effects of teacher-guided and student-activated teaching methods on student learning in higher education. *Int. J. Educ. Res.* 95, 26–35.
- Fosnot, C.T. (Ed.), 1996. *Constructivism: Theory, Perspectives, and Practice*. Teachers College Press.
- Glatthorn, A., Boschee, F., Whitehead, B.M., Boschee, B.F., 2018. *Curriculum Leadership: Strategies for Development and Implementation*, fifth ed. Sage.
- Howatt, A.P.R., 1984. *A History of English Language Teaching*. Oxford University Press.
- Klapper, J., 2003. Taking communication to task? A critical review of recent trends in language teaching. *Lang. Learn.* J. 27, 33–42.
- Krashen, S., Terrell, T.D., 1983. *The Natural Approach: Language Acquisition in the Classroom*. Pergamon.
- Krashen, S., 1981. *Second Language Acquisition and Second Language Learning*. Pergamon Press.
- Krashen, S., 1982. *Principles and Practice in Second Language Acquisition*. Pergamon Press.
- Krashen, S., 1985. *The Input Hypothesis: Issues and Implications*. Longman.
- Larsen-Freeman, D., 2015. Research into practice: grammar learning and teaching. *Lang. Teach.* 48 (2), 263–280.
- Lightbown, P., Spada, N., 2013. *How Languages Are Learned*, fourth ed. Oxford University Press.
- Loewen, S., 2020. *Introduction to Instructed Second Language Acquisition*, second ed. Routledge.
- Long, M., 1981. Input, interaction, and second language acquisition. *Ann. N. Y. Acad. Sci.* 379 (1), 259–278.
- Long, M., 1983a. Linguistic and conversational adjustments to nonnative speakers. *Stud. Sec. Lang. Acquis.* 5 (2), 177–193.
- Long, M., 1983b. Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Appl. Linguist.* 4 (2), 126–141.
- Long, M., 1985. A role for instruction in second language acquisition: task-based language teaching. In: Hyltenstam, K., Pienemann, M. (Eds.), *Modeling and Assessing Second Language Acquisition*. Multilingual Matters, pp. 77–99.
- Long, M., 1996. The role of the linguistic environment in second language acquisition. In: Ritchie, W., Bhatia, T. (Eds.), *Handbook of Second Language Acquisition*. Academic Press, pp. 413–468.
- Long, M., 2000. Focus on form in task-based language teaching. In: Lambert, R.D., Shohamy, E. (Eds.), *Language Policy and Pedagogy: Essays in Honor of A. Ronald Walton*. John Benjamins, pp. 179–192.
- Long, M., 2015a. *Second Language Acquisition and Task-Based Language Teaching*. Wiley-Blackwell.
- Long, M., 2015b. TBLT: building the road as we travel. In: Bygate, M. (Ed.), *Domains and Directions in the Development of TBLT: A Decade of Plenaries from the International Conference*. John Benjamins, pp. 1–26.
- Markee, N. (Ed.), 2015. *The Handbook of Classroom Discourse and Interaction*. John Wiley & Sons.
- Medgyes, P., 1986. Queries from a communicative teacher. *ELT J.* 40 (2), 107–112.
- Mitchell, R., Myles, F., Marsden, E., 2019. *Second Language Learning Theories*, fourth ed. Routledge.
- Norris, J., Bygate, M., Van den Branden, K., 2009. Introducing task-based language teaching. In: Van den Branden, K., Bygate, M., Norris, J. (Eds.), *Task-based Language Teaching: A Reader*. John Benjamins, pp. 15–19.
- Nunan, D., 2004. *Task-Based Language Teaching*. Cambridge University Press.
- Piaget, J., 1926. *The Language and Thought of the Child*. Brace.
- Richards, J.C., Rodgers, T.S., 2014. *Approaches and Methods in Language Teaching*, third ed. Cambridge University Press.
- Richards, J.C., 2006. *Communicative Language Teaching Today*. Cambridge University Press.
- Rose, H., 2019. Dismantling the ivory tower in TESOL: a renewed call for teaching-informed research. *Tesol Q.* 53 (3), 895–905.
- Savignon, S., 1983. *Communicative Competence*. Addison-Wesley.
- Savignon, S., 1991. Communicative language teaching: state of the art. *Tesol Q.* 25 (2), 261–277.
- Schmidt, R., 1990. The role of consciousness in second language learning. *Appl. Linguist.* 11 (1), 17–46.
- Schmidt, R., 1993. Consciousness, learning and interlanguage pragmatics. In: Kasper, G., Blum-Kulka, S. (Eds.), *Interlanguage Pragmatics*. Oxford University Press, pp. 21–42.
- Schmidt, R., 1995. Consciousness and foreign language learning: a tutorial on the role of attention and awareness. Technical Report No. 9. In: Schmidt, R. (Ed.), *Attention and Awareness in Foreign Language Teaching and Learning*. University of Hawaii at Manoa, pp. 1–64.
- Schmidt, R., 2001. Attention. In: Robinson, P. (Ed.), *Cognition and Second Language Instruction*. Cambridge University Press, pp. 3–32.
- Skinner, B.F., 1957. *Verbal Behavior*. Copley Publishing Group.
- Spada, N., 2018. Isolating or integrating attention to form in communicative instruction: a dilemma? *Babel* 53 (1), 7–12.
- Swain, M., 1985. Communicative competence: some roles for comprehensible input and comprehensible output in its development. In: Gass, S., Madden, C. (Eds.), *Input in Second Language Acquisition*. Newbury House, pp. 235–256.
- Swain, M., 1991. French immersion and its offshoots: getting two for one. In: Freed, B. (Ed.), *Foreign Language Acquisition: Research and the Classroom*. Heath, pp. 91–103.
- Swain, M., 1995. Three functions of output in second language learning. In: Cook, G., Seidlhofer, B. (Eds.), *Principles and Practice in Applied Linguistics: Papers in Honor of H. G. Widdowson*. Oxford University Press, pp. 125–144.
- Swain, M., 2005. The output hypothesis: theory and research. In: Hinkel, E. (Ed.), *Handbook of Research in Second Language Teaching and Learning*. Routledge, pp. 471–484.
- Terrell, T.D., 1977. A natural approach to second language acquisition and learning. *Mod. Lang. J.* 61 (7), 325–336.
- Terrell, T.D., 1982. The natural approach to language teaching: an update. *Mod. Lang. J.* 66 (2), 121–132.
- Thomas, N., Brereton, P., 2019. Practitioners respond to Michael Swan's “Applied Linguistics: a consumer's view”. *Lang. Teach.* 52 (2), 275–278.
- Valeo, A., Spada, N., 2016. Is there a better time to focus on form? Teacher and learner views. *Tesol Q.* 50 (2), 314–339.
- Van den Branden, K., Bygate, M., Norris, J., 2009. Task-based language teaching: introducing the reader. In: Van den Branden, K., Bygate, M., Norris, J. (Eds.), *Task-Based Language Teaching: A Reader*. John Benjamins, pp. 1–13.

Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Walker, I., Chan, D., Nagami, M., Bourguignon, C. (Eds.), 2018. *New Perspectives on the Development of Communicative and Related Competence in Foreign Language Education*. De Gruyter.

Weimer, M., 2013. *Learner-Centered Teaching: Five Key Changes to Practice*, second ed. Jossey Bass. **This page intentionally left blank**

Learning and literacy out of school: learning every day and everywhere

Katherine Schultz^a and Glynda A. Hull^{b, a} University of Colorado Boulder, School of Education, Boulder, CO, United States; and
^b University of California, Berkeley, School of Education, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	374
New methodologies	375
Foregrounding racial justice and power	376
Learning on the move	376
Ethical dimensions of technology	377
Conclusion	378
References	379

Introduction

Youthful candy sellers, teenage Tik-tokers, online activists, aspiring artists, aficionados of writing, reading, and sense making in a plethora of community-based programs: these are the able learners and the thriving spaces, then and now, of learning and literacy out of school. Beginning three decades ago, and continuing to this moment, educational researchers have looked beyond the schoolhouse door in order to document, theorize, and celebrate learning (Hull and Schultz, 2002). The early studies showed how youth and adults regularly displayed competencies and motivations to learn and flourish in out-of-school settings, achievements that were not as visible in formal schoolrooms. These researchers asked why this was so, and they also wondered hopefully whether out-of-school practices, particularly digital ones, might be transported to schools with transformational results. The newer studies, especially those from the learning sciences that we feature here, likewise feature the accomplishments of learners out of school. But the recent work also tacks in different directions. And notably, this newer work in the learning sciences focuses on documenting learning everywhere, rather than focusing on whether and how these practices might inform learning in schools. In this chapter we chart some of the new conceptual pathways on learning every day and everywhere in out-of-school contexts that seem to us the most promising.

The early days of research on learning out of school yielded striking and important conceptual advances. In studies of literacy, for example, examinations of reading and writing in specific communities revealed a plurality of meaning-making practices and literacy's situated nature, shaped by socio-political, cultural, and institutional contexts (i.e., Heath, 1983; Scribner and Cole, 1981; Street, 1984). It turned out, notably, that school-based literacy is just one kind of reading and writing in a universe of many kinds, and this theorization itself represented a dramatic conceptual breakthrough, one that beckoned an ethnographic flowering, as literacy researchers embraced the invitation to characterize reading and writing as particular practices in multiple community, activity, and identity contexts (e.g., Barton and Hamilton, 1998; Heller, 1997; Kalman, 1999; Moss, 1994). "Out of school" was therefore a salient category in the early research, even as the kinds of learning done in such out-of-school contexts continued to play second fiddle to the kinds of learning done in school, where academic literacy and school-based knowledge, demonstrated through testing and accountability regimes, continued to hold sway.

The intervening years, however, have seen important shifts in this regard. Due in large part to the digital revolution, especially the rise of social media and the proliferation of mobile electronic devices, opportunities to socialize, create, communicate, teach, and learn have proliferated, as have digital platforms and spaces that host and afford such activities (Mills, 2010). Gone are the days when schools, even institutions of higher learning (Christensen and Eyring, 2011), had a monopoly on the transmission of knowledge and the socialization of learners. Learning happens all the time and everywhere, and increasingly learners have options; they are able to pick and choose their knowledge purveyors. What this means for research on learning out of school is that it is more important than ever before to shift our focus from schools as the only—or sometimes even the primary—site of learning, to recognizing and supporting the learning that takes place in and across a wide range of contexts, those carefully crafted by educators and researchers, and those that are more spontaneously constructed by children, youth, and adults. In addition, this new understanding of where, when, and how learning occurs provides new theoretical insights into literacy practices and education more broadly.

In this chapter we feature research on out of school learning that is informed by the learning sciences (cf., Nasir et al., 2020a). Doing so allows us to foreground insights about learning that arise from methodological innovations and theoretical directions, considerations that are especially salient in out-of-school learning. However, it is important to note that research on learning and literacy out of school is alive and well in various fields. As they have for many years, youthful desires and capacities to engage with digital tools continue to fascinate media researchers, who find fertile ground in such activities for exploring the development of creativity through, and criticality toward, popular media (e.g., Kraxenberger and Lauer, 2022). With the ubiquity of platforms like Twitter, new kinds of social (e.g., Marwick and Boyd, 2011) and political activity have emerged, such as hashtag activism (Jackson et al., 2020), which has proven of keen interest to researchers in multiple fields. Whole new industries have grown up around the provision of online learning, which promises the certified outcomes of formal schooling minus traditional time, space, and

attendance requirements (Impey, 2020). Such forms of non-school or extra-school learning have attracted the interest of university researchers, but also the corporate world of ed tech, which has designs on doing educating better than schools (Houghtelling, 2021). In the realm of play, identity experimentation, fantasy, and role play, which afford learning and development of myriad kinds, virtual worlds lead the parade for researchers intent on exploring the newest forms of technological meditation (i.e., *Journal of Virtual Worlds Research*). It is no accident that many of our examples connect to the digital, for this modality has claimed the lion's share of interest from out-of-school researchers for several years. Nonetheless, we also note that there remains abundant interest in forms of learning and development out of school that are not primarily mediated by the internet or digital tools: think dance, sports, and other kinds of embodied engagement (Katz, 2013) and imagine them in the context of after-school programs, a space that was an early focus of research on out-of-school learning and one that continues to thrive.

We are curious about how the current research in the learning sciences has brought new ideas into the discussions and research, in the past designed primarily by anthropologists and literacy scholars, that have documented the plethora of literacy practices across a wide range of contexts. We begin this chapter with a brief discussion of social design experiments, a central methodological approach used by researchers in the studies we review. As with current research on literacy in out-of-school contexts (e.g., Schultz et al., 2016), much of the latest research in the learning sciences is grounded in social and cultural theories. In addition, there is a focus on learning across time, place, and context, and on the movement that propels the learner across these contexts. Nasir et al. (2020b) remind us that learning takes place over time and across a lifespan, and through participation in a range of activities in both formal and informal settings. At the same time, they argue that people are differently positioned to have access to this range of opportunities for learning (Nasir et al., 2020b). Introducing the concept of learning pathways, they explain:

We conceptualize learning as occurring along culturally organized learning pathways—sequences of consequential participations and transitions (Beach, 1999) in learning activities that move (or do not move) one towards greater social recognition as competent in particular learning domains and situations (p. 195).

Nasir and her colleagues focus, as well, on the ways that learning pathways are intertwined with identities that affect how the pathways are determined, accessed, and resisted.

New methodologies

This emphasis on examining everyday activities as sites of learning—taking a walk, cooking a meal, playing a game—leads to the use of different research methods from those often used to document the learning that takes place in out-of-school programs, which has characterized much of the literature on out-of-school literacies and learning (e.g., Hull and Schultz, 2002). Rather than viewing programs or communities as sites for rich descriptions of youth's accomplishments in out-of-school activities, the more recent research in the learning sciences often leans on design-based research or research focused on the design of opportunities for sustained learning. Gutiérrez and Jurow (2016) introduced the method of social design experimentation that emphasizes the importance of co-designing projects with participants, focusing on justice and equitable learning opportunities that are available for youth from marginalized or non-dominant communities (See also Gutiérrez and Vossoughi 2010; Bang et al., 2016 provide a discussion of community-based research design.) Importantly, social design experiments seek to transform social institutions (Gutiérrez and Jurow, 2016) and re-imagine possible futures (Gutiérrez et al., 2020). Gutiérrez and Jurow explain:

Of significance is that social design experiments' explicit focus on disrupting educational, structural, and historical inequities through the design of transformative learning activity provides openings for learning, a context of critique for resisting and challenging the conditions that create and sustain inequities, and a space for generating their possible solutions. (p. 4)

As a result, social design experiments highlight the notion that participants are partners in the research, who have agency to design opportunities for their consequential learning in dynamic contexts, while participating in changing their futures. More recently Gutiérrez et al. (2020) have envisioned this design-based research along utopian lines, emphasizing the importance of seeing promise in everyday practices, in order to design new and more equitable and just forms of education for the most vulnerable communities.

The use of design-based research in learning sciences opens up new ways of describing and understanding learning. In what follows, we describe a few of the nascent themes we have observed in this literature that add to our understandings of learning in informal and out-of-school spaces. We begin with a discussion of the centrality of racial justice, power, and politics in these research studies. Equity and justice are at the heart of these studies of learning. Next, we explore the ways learning is documented as movement across space and time. Finally, we describe the ethical turn recently taken in studies of technology and learning.

Foregrounding racial justice and power

In our review of the recent studies of out-of-school learning in the learning sciences, we are struck that racial justice and power are foregrounded in new ways in many of the studies. While in the past, researchers often located their investigations in marginalized communities and thereby called attention to inequities as well as the creativity and potential of “underserved” youth, the newer work in the learning sciences takes an explicit stance toward advancing justice and social activism, highlighting BIPOC communities. For instance, [Barajas-López and Bang \(2018\)](#) situate their study of an Indigenous STEAM program in “the history of colonialism and the intended erasure of indigeneity through the world” with a goal of “creating equitable and transformative learning (p. 7).” Their work underscores what [hooks, 1990](#) describes as the possibilities for imagining learning from a position of marginality and oppression. hooks introduces the concept of radical openness as “a generative space for freedom, a space alive with continual engagement, play, and the re-making of cultural and historical relations that can create flourishing ways of being and becoming,” explaining that this type of learning is fundamentally about acting in the world with care and compassion (p. 150). Putting these ideas into practice, [Jurow and Shea \(2015\)](#) describe their project connected to the food justice movement as designed for radical openness and possibility ([Gutiérrez et al., 2020](#)) in order to make visible the ways people make decisions shaped by power and authority. [Gutiérrez, Jurow, and Vakil \(2020\)](#) write, “The shifting nature of our research and designs was a mirrored response to the shifting nature of injustices facing the community (p. 341).” Politics is explicitly situated in this research, which documents a political movement. Its form, message, and findings are all political.

In their introduction to a special issue of the *Journal of the Learning Sciences* focused on learning and collective action, [Curnow and Jurow \(2021\)](#) explain: “The kind of work we are advocating calls for a distinctive set of interventions, embedded in collective action that is anti-racist, anti-colonial, queer, feminist, anti-ableist, abolitionist, and that makes justice possible and inevitable.” Their identification of this wide range of groups and their explicit framing of this work around advocacy signals a shift in both theory and practice in the learning sciences and adds to our knowledge of informal learning. Their centering of justice as method as well as context leads to new practices and new understandings about learning and learners. This discussion of learning through political engagement is explicitly tied to learning in everyday life and seems less committed, especially in comparison to earlier studies, to finding a way to bring this way of learning and knowing explicitly into the classroom.

An example of work that centers justice and politics is a study conducted by [Tivaringe and Kirshner \(2021\)](#) that traces the learning that occurred during the growth of a local activist group in South Africa as it became a social movement. The authors describe how members of this group—Equal Education (EE), an organization committed to expanding access to quality education—participated in a range of activities, including reading circles and participating—or in some cases apprenticing—in political activism such as campaigns. One youth described how her participation in a campaign helped her to learn that there were a range of ways to engage in electoral politics, from voting to civil disobedience. The authors conclude that their study illustrates the ways that youth learn how to navigate contentious political moments and wield political power through the participation in multigenerational groups. Their research illustrates a particular form of learning that includes organizing and leads to political change. The authors clarify that the youth participants likely gained this knowledge by apprenticing with more experienced participants through the Equal Education group. They further explain the ways that this learning occurred not only in the courthouses and on the streets, but also through social media and other online spaces. At the same time, the learning that took place in these meetings and actions does not resemble school learning nor could it take place in that context.

Learning on the move

A second theme salient to conceptual advances in learning is the recent interest in studying “learning on the move” (e.g., [Marin, 2020; Marin et al., 2020](#)). This work emphasizes the spatial and temporal dimensions of learning in everyday contexts, including relationships between humans and nature. In their study of learning through walking together, [Marin et al. \(2020\)](#) explain that a focus on mobility in learning, or learning on the move, allows us to reconceptualize nature-culture relationships. Rather than studying learning in specified places or programs, whether in classrooms, after school programs, or in homes and communities, these authors examine how learning occurs in specified locations or during particular events, as well as during the movement between them. In contrast to the typical understanding of context as place, [Marin and Bang \(2018\)](#) describe context as temporal. In addition, they illustrate how the process of walking, reading the land, and exchanging stories (or storying) leads to deep learning about the natural world as well as understanding how a person (child or adult) is situated in the world. They explain further that this way of knowing reflects Indigenous epistemologies that have been practiced for generations (see also [Brayboy, 2005](#), for an early discussion of TribalCrit or Tribal Critical Race Theory.)

In their introduction to the *Handbook of the Cultural Foundations of Learning*, [Lee et al. \(2020\)](#) elaborate the Indigenous epistemologies that animate current theories of learning:

Operating from epistemological and ontological belief systems that center the well-being of the self, family, the community, the land and waters, the spirits, and the ancestors, these traditions view human learning and development within a framework of connectedness, embeddedness, and reciprocal relationships with all things past, present, and future (p. xxv).

In a study that illustrates this work, drawing on indigenous epistemologies, [Marin and Bang \(2018\)](#) describe the ways a mother and her son “story” the land, while walking in an urban setting and investigating through talk whether the water they observe is indeed a river. Marin and Bang describe this process of learning, explaining:

[I]t is another example of epistemological navigation where mother and son use questions and directives, as well as embodied moves, to open and close spaces of wonder and inquiry. They also rely on evidence both from the field and past experiences, knitting together landscapes to consider systems that are present in multiple places. Also, the mother’s use of tentative language provides an opportunity for Jason to be a coconstructor in knowledge building and provides openings for him to offer both his disagreement and his own ideas (p. 111).

The authors explain that the learning occurs between people in relationship to the land and water or between humans and more than humans.

A recent study by Taylor provides another example of learning on the move. Introducing the term “locative literacies”, [Taylor \(2017\)](#) traces the learning of youth as they tracked their pathways through a designed task that used global positioning system software (GPS) in order for youth to produce their own maps to use in conversation with policymakers (See also [Taylor and Hall, 2013](#)). As [Taylor \(2017\)](#) explains,

Learning along lines is a kind of participation framework ([Goffman, 1981](#)) or “fluid structure of mutual engagement and disengagement” ([Jordan and Henderson, 1995](#), p. 67) that organized bodies, technologies, maps, and places along lines that youth read, talked about, walked, drew, gestured, and examined (p. 535).

Through her research, Taylor expands notions of embodied learning. She explains that the notion of learning along lines—or learning through moving across the city, reading and writing its contours, and recording insights—highlights how people can be part of the text rather than simply recipients or producers of that text. In order to capture this embodied learning, Taylor produced multimodal transcripts that weave together talk, body position, gaze, gesture, and the larger context, including tools, in order to glean new insights about learning that is not bound by time and place, but instead traces the movement of youth across a city. Taylor writes about how this process of learning along the lines created opportunities to engage in new forms of digital literacies with mobile and location-aware technologies, such as GPS. This process was both embodied and understood through the youth’s own narratives and experience with the place where this learning occurred. Rather than seeing the youth as either consumers or producers of texts, Taylor describes the ways that they are a part of the text, thus affording new understandings and learning. Ultimately, the youth were able to use their knowledge gained from this project to inform the urban community mapping process, sharing their insights with urban planners, cartographers, and city government officials.

What strikes us about these examples of political activism and learning on the move, is that they illustrate opportunities for learning that do not exist in school or school-like programs. Furthermore, they are not examples of learning in service of schools (or for school-learning) nor are they examples of the kinds of learning schools might replicate or count. Each of these examples has equity and justice at its center. They are examples of how opportunities for learning can be designed to engage learners in new ways, through new modalities, as well as in everyday activities.

Ethical dimensions of technology

Much learning and doing out of school, both past and present, is mediated by digital tools, platforms, and socio-technical systems. Famously, such learning and doing have privileged—during the last two decades—creative, productive activities. Whereas decoding or viewing or comprehending text was the quintessential activity that distinguished literate participation during the last century, it is the creation and production and circulation of texts, including multimodal ones, that is increasingly today’s literate calling card, particularly out of school (cf. [Brandt, 2014](#)). And even more expansively, creative production fueled by the affordances of digital media also increasingly includes the design of platforms and applications through learning to code, instead of just textual artifacts ([Peppler and Kafai, 2007](#)). The rallying cry has thus become youth as makers, not only consumers of technology, and so we witness youth-created video-games and robots, as well as artistic remixes, mashups, and memes. This creative, productive, digitally-mediated activity and learning typically lives in community centers, hacker spaces, homes, media labs, and on the move as youth and everyone carry and deploy their devices in intersection with geographic and social spaces. Documenting, analyzing, theorizing, and supporting such activities are still important foci for research on out-of-school learning and doing.

Still, our review of recent studies of out of school learning associated with the learning sciences has also revealed a new and important focus around youth’s associations with technology that we wish to highlight here, and it takes an ethical turn. Specifically, there is growing interest in creating learning environments and harnessing digital tools in support of youth’s engagement with the moral and ethical dimensions of technology. [Vakil and McKinney de Royston \(2022\)](#) helpfully provide three broad categories for characterizing educational motives for using technology. The first promotes digital and media literacy by emphasizing youth as sophisticated users of technology, able to comprehend and consume digital content. The second privileges making, whether designing technologies or creating digital artifacts, as we discussed above and identified as the major approach toward digitally

mediated learning out of school during the last two decades. The third category prioritizes critique and positions youth as “critical agents of technology” (p. 4). While this approach appreciates the importance of youthful identities as makers, it has a critical edge, whereby technology is mastered toward social justice ends. That is, “learners are supported to use, manipulate, and create technologies that subvert power” (p. 4), and in “decoding technology’s relationship with power” (p. 1).

For a long time, many have felt a moral compulsion to put technology in the hands of minoritized and underserved youth, contributing thereby to leveling the playing field and addressing material inequities. There has been a commitment as well to giving youth access to powerful digital media practices, whether digital storytelling or coding, and thereby supporting their agency as makers and communicators. But often, this work has not explicitly foregrounded the ethical implications of technology use—that is, “learning how to decode and unmake tech’s relationship with power through artistic, moral and humanistic inquiry” (Vakil and McKinney de Royston, p. 2). This is the approach that Vakil and McKinney de Royston advocate and that they illustrate through the Young People’s Race Power and Tech Project (YPRPT), a community-university partnership in Chicago where high school students, community educators, undergraduates, and computer scientists jointly investigated how artificial intelligence and big data technologies were affecting minoritized communities in Chicago. During a summer workshop and a year-long after-school program, students explored facial recognition technology and other advanced technologies with an eye toward how they were operating in their own communities. They also produced documentaries that explored the ethical and technical dimensions of the technologies. In their article, Vakil and McKinney de Royston feature their analysis of a documentary film produced by two youth to explore how ICE used facial recognition software to monitor social media accounts and to surveil public areas in their communities, with the end result of detaining, deporting, and criminalizing immigrants. The authors refer to the students’ work as “philosophical sensemaking” (p. 9) and demonstrate how they combined personal narration and political inquiry to artistic and rhetorical advantage.

Among the windows opened by this provocative effort to position youth as philosophers of technology is a fresh look at what constitutes generative outcomes where learning about and with technology is concerned. Historically, out-of-school or afterschool programs were taken to task if they were not able to demonstrate gains in school-based skills as demonstrated by conventional measures, such as reading comprehension scores. Vakil and McKinney de Royston, and their collaborators at YPRPT throw off these shackles and make a compelling argument through their data, following [Bang and Vossoughi \(2016\)](#), about the importance of growth in relationality as an important goal and outcome. It turned out that the young people who are featured in their research and the YPRPT program built strong relationships with each other over the course of project activities, which in turn revealed the interconnections between relationality and the understandings the youth came to about technology, race, and power. In fact, they argue that a pedagogical design that centers the development of racial and political solidarity is a prerequisite for the kind of ethical sensemaking around technology they hoped to foster. But, on the flip side, Vakil and McKinney de Royston also show that relationality can itself be an outcome of ethical sensemaking. Grappling with content around the interconnections among race, power, and technology, contextualizing their work in the current historical moment of a national reckoning around systemic racism, bringing their burgeoning understandings to bear in the creation of a documentary—all of these activities likewise deepened their relationships with each other, the strength of their connection to social issues, and their nascent theorizations of the impacts of technology in today’s world.

Another study that foregrounds the ethical dimensions of technology, this one set in India, is [Hull et al. \(2021\)](#) design-based exploration of a virtual reality storytelling exchange. Drawing from a critical account of cosmopolitanism, their essay theorizes young people’s rights and capacities to speak, listen, and be heard in a global world and, in tandem, it explores the communicative power and human reach of new forms of storytelling. The study demonstrates that it is indeed possible to design learning spaces that mitigate communicative and representational inequalities and where young people learn to communicate and imagine across semiotic, cultural, and geographic divides. It also examines the challenges surrounding such acts of trust and imagination and emphasizes the importance for youth of developing and exercising capacities to “world the world”—to represent their own lived experiences in relation to understanding the lived realities of others. These acts of perspective-offering and perspective-taking, the authors claim, are quintessential critical cosmopolitan capacities for a global era. Especially important are the linguistic and semiotic aspects of these imaginings, seen in the ways that youth world their worlds by marshaling symbolic resources that cross and blend languages, modes, and media ([Canagarajah, 2013](#)) and that transcend logocentric notions of texts ([Lizárraga et al., 2015](#)).

Conclusion

As we write this chapter, a global pandemic, that has lasted several years and continues to persist, has shaped education and lives. The experience of loss is top of mind as we enter a post-Covid time (though that term remains contested), the loss of lives uppermost, but also the loss of jobs, mobility, sociality, and, for educators and families most poignantly, the loss of opportunities to learn. Indeed, “learning loss” is now a ubiquitous term at every level of schooling, as analyses reveal that a consequence of a world-wide pandemic is that children fell behind in mathematics and reading, gaps widened in achievement and opportunity, and long-standing inequities increased. This is a picture of an apocalyptic educational landscape, the contours of which we acknowledge, even as we also aspired in this essay to offer a different vocabulary for describing it and some distinctive strategies for addressing it. If, as research suggests, we understand that people are engaged in learning all of the time and in a wide variety of contexts, then we should clarify that what was lost for some children was specifically school learning. Learning certainly occurred during the

pandemic when students in the US and other countries were confined to their homes. However, it was learning that is not typically accounted for in schools. Rather than a focus on the learning losses during Covid, we urge educators, parents, and caregivers to notice the learning that did happen and build on that knowledge. This might also lead to funding community-based, youth-initiated learning in response to “learning loss”, as well as summer schools that replicate school learning. The new work on learning, such as the research we feature in this article, helps us to notice and understand these new and potentially powerful types of learning.

In their outgoing editors’ note for *The Journal of Learning Sciences*, the signature journal for the field, Radinsky and Tabak, 2017 recounted their accomplishments during their term as editors. Among other changes to the journal, they mentioned that: “[t]he retirement of the Learning Outside of School strand at *JLS* corresponds with an increasing diversity of contexts in which learning scientists study learning—such that outside of school is no longer seen as a necessary descriptor.” While we understand the direction they point us toward—that learning occurs everywhere and the previous containers of “school” and “out-of-school” may be limiting in our current times—we argue that it is important to pay attention to out-of-school spaces, as well as everyday learning in order to deepen our understanding of the ubiquity and variation of learning.

The research we reviewed in this chapter reveals that learning outside of school—across space, time, and contexts—is flourishing. The studies we describe, as well as countless others, illustrate the promise and power of new contexts for learning. In this research, we have witnessed youth learning together in programs connected to schools or embedded in communities. We also offer research that documents the explorations of youth learning with family members and on their own. We were particularly taken by the multi-generational examples where youth learned as they participated in political campaigns and other forms of activism. And while the current research hones our vision and allows us to see learning everywhere and as a part of daily life, it also suggests new questions and research directions about other kinds of learning we should look for in the future and document as researchers, educators, and community members.

We conclude with some of these questions that might guide our inquiries into how learning transpires in the future. We know that many youth across the world are enraged about global climate change and the deteriorating planet and are seeking to make their voices and opinions heard. As youth engage in activism, in multigenerational groups, and on their own terms, as well as through their own invented routines and rituals, what are the contours of their learning? We know that the learning connected to social and political movements is generally not connected with school (and often students explicitly leave school to do this work), we wonder whether and how this learning is connected with school knowledge, while at the same time asking whether that question is even important? A second question that comes to mind concerns the connections between belonging and learning. As preK-12 schools, universities, and organizations seek to diversify their institutions, the notion of belonging, especially for minoritized students, has become critical. How do we understand belonging and identity as intrinsic to learning, and how do we construct educational environments, including online spaces, where learners feel they belong? We wonder, as well, how our understandings of learning will change as new technological advances offer new opportunities for exploration, experimentation, and the development of new knowledge. In tandem, what guardrails can we build as misinformation flourishes via social media and threats to privacy and equity abound through algorithmic bias in an AI world? What might and what should learning every day and everywhere look like in the future, when that future includes virtual worlds? We live in a precarious time, characterized by distrust and, throughout much of the United States, a loss of faith in the future of the democratic foundation of the country. What does this sense of despair and hopelessness mean for how we understand needs and opportunities for learning? And how might a focus on learning as it occurs every day and everywhere lead us toward hope and possibility?

References

- Bang, M., Vossoughi, S., 2016. Participatory design research and educational justice: studying learning and relations within social change making. *Cognit. Instruct.* 34 (3), 173–193.
- Bang, M., Faber, L., Gurneau, J., Marin, A., Soto, C., 2016. Community-based design research: learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind Cult. Activ.* 23 (1), 28–41.
- Barajas-López, F., Bang, M., 2018. Indigenous making and sharing: claywork in an indigenous STEAM program. *Equity Excell. Educ.* 51 (1), 7–20.
- Barton, D., Hamilton, M., 1998. *Local Literacies*. Routledge, London.
- Brandt, D., 2014. *The Rise of Writing: Redefining Mass Literacy*. Cambridge University Press, Cambridge.
- Beach, K., 1999. Consequential transitions: a sociocultural expedition beyond transfer in education. *Rev. Res. Educ.* 24 (1), 101–139.
- Brayboy, B.M.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37 (5), 425–446.
- Canagarajah, S., 2013. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge, London, UK.
- Christensen, C., Eyring, H.J., 2011. *The Innovative University: Changing the DNA of Higher Education From the Inside Out*. Jossey-Bass, San Francisco.
- Curnow, J., Jurow, A.S., 2021. Learning in and for collective action. *J. Learn. Sci.* 30 (1), 14–26.
- Goffman, E., 1981. *Forms of Talk*. University of Pennsylvania Press, Philadelphia, PA.
- Gutiérrez, K.D., Jurow, A.S., 2016. Social design experiments: toward equity by design. *J. Learn. Sci.* 25 (4), 565–598.
- Gutiérrez, K.D., Jurow, A.S., Vakil, S., 2020. Social design-based experiments: a utopian methodology for understanding new possibilities for learning. In: Nasir, N.S., Lee, C.D., Pea, R., McKinney de Royston, M. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, New York, pp. 330–347.
- Gutiérrez, K.D., Vossoughi, S., 2010. Lifting off the Ground to return a new” documenting and designing for equity and transformation through social design experiments. *J. Teach. Educ.* 61 (1–2), 100–117.
- Heath, S., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.
- Heller, C.E., 1997. *Until We Are Strong Together: Women Writers in the Tenderloin*. Teachers College Press, New York.
- hooks, b., 1990. Choosing the margin as a space of radical openness. In: hooks, B. (Ed.), *Yearnings: Race, Gender, and Cultural Politics*. South End, Boston, MA, pp. 145–153.

- Houghtelling, W., 2021. The Past, Present and Future of Learning and Development. <https://willhoughtelling.medium.com/the-past-present-and-future-of-learning-development-daeea4e39af1>. (Accessed 30 June 2022).
- Hull, G., Schultz, K. (Eds.), 2002. *School's Out!: Bridging Out-Of-School Literacy With Classroom Practices*. Teachers College Press, New York, NY.
- Hull, G.A., Unackat, D., Adams, J., 2021. Subaltern voices and social imaginaries. In: Schecter, S.R., James, C. (Eds.), *Critical Approaches toward a Cosmopolitan Education*. Routledge, New York.
- Impey, C., 2020. Massive Online Open Courses See Exponential Growth During COVID-19 Pandemic. <https://theconversation.com/massive-online-open-courses-see-exponential-growth-during-covid-19-pandemic-141859#:~:text=Massive%20open%20online%20classes%2C%20or,from%201.6%20to%2010.3%20million>. (Accessed 20 June 2022).
- Jackson, S., Bailey, M., Welles, B.F., 2020. #HashtagActivism: Networks of Race and Gender Justice. The MIT Press, Boston.
- Jordan, B., Henderson, A., 1995. Interaction analysis: foundations and practice. *J. Learn. Sci.* 4 (1), 39–103.
- Jurow, A.S., Shea, M., 2015. Learning in equity-oriented scale-making projects. *J. Learn. Sci.* 24 (2), 286–307.
- Kalman, J., 1999. *Writing on the Plaza: Mediated Literacy Practices Among Scribes and Clients in Mexico City*. Hampton Press, Cresskill, NJ.
- Katz, M. (Ed.), 2013. *Moving Ideas: Multimodality and Embodied Learning in Communities and Schools*. Peter Lang, New York.
- Kraxenberger, M., Lauer, G., 2022. Wreading on Online Literature Platforms. *Written Communication*, pp. 1–35.
- Lee, C.D., Nasir, N.S., Pea, R., McKinney de Royston, M., 2020. Introduction: reconceptualizing learning: a critical task for knowledge-building and teaching. In: Nasir, N.S., Lee, C.D., Pea, R., McKinney de Royston, M. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, New York, pp. xvii–xxxv.
- Lizárraga, J.R., Hull, G., Scott, J.M., 2015. Translingual practices in a social media age: lessons learned from youth's transnational communication online. In: Mole, D., Sato, E., Boals, T., Hedgspeth, C. (Eds.), *Multilingual Learners and Academic Literacies: Sociocultural Contexts of Literacy Development in Adolescents*. Routledge, New York, NY, pp. 105–132.
- Marin, A.M., 2020. Ambulatory sequences: ecologies of learning by attending and observing on the move. *Cognit. InStruct.* 38 (3), 281–317.
- Marin, A., Bang, M., 2018. "Look it, this is how you know:" family forest walks as a context for knowledge-building about the natural world. *Cognit. InStruct.* 36 (2), 89–118.
- Marin, A., Taylor, K.H., Shapiro, B.R., Hall, R., 2020. Why learning on the move: intersecting research pathways for mobility, learning and teaching. *Cognit. InStruct.* 38 (3), 265–280.
- Marwick, A., Boyd, D., 2011. To see and be seen: celebrity practice on twitter. *Convergence* 17 (2), 139–158.
- Mills, K.A., 2010. A review of the "digital turn" in the new literacy studies. *Rev. Educ. Res.* 80 (2), 246–271.
- Moss, B. (Ed.), 1994. *Literacy across Communities*. Hampton Press, Cresskill, NJ.
- Nasir, N.S., Lee, C.D., Pea, R., McKinney de Royston, M. (Eds.), 2020a. *Handbook of the Cultural Foundations of Learning*. Routledge, New York.
- Nasir, N.S., McKinney de Royston, M., Barron, B., Bell, P., Pea, R., Stevens, R., Goldman, S., 2020b. Learning pathways: how learning is culturally organized. In: Nasir, N.S., Lee, C.D., Pea, R., McKinney de Royston, M. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, New York, pp. 195–211.
- Peppler, K.A., Kafai, Y.B., 2007. From SuperGoo to scratch: exploring creative digital media production in informal learning. *Learn. Media Technol.* 32 (2), 149–166.
- Radinsky, J., Tabak, I., 2017. Outgoing editors' note: the *journal of the learning sciences* as a mirror of trends in the field. *J. Learn. Sci.* 26 (1), 1–6.
- Schultz, K., Hull, G., Higgs, J., 2016. After writing, after school. In: MacArthur, C., Graham, S., Fitzgerald, J. (Eds.), *Handbook of Writing Research*. Guilford Press, New York, NY, pp. 102–115.
- Scribner, S., Cole, M., 1981. *The Psychology of Literacy*. Harvard University Press, Cambridge, MA.
- Street, B.V., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Taylor, K.H., 2017. Learning along lines: locative literacies for reading and writing the city. *J. Learn. Sci.* 26 (4), 533–574.
- Taylor, K.H., Hall, R., 2013. Counter-mapping the neighborhood on bicycles: mobilizing youth to reimagine the city. *Technol. Knowl. Learn.* 18, 65–93.
- Tivaringe, T., Kirshner, B., 2021. Learning to claim power in a contentious public sphere: a study of youth movement formation in South Africa. *J. Learn. Sci.* 30 (1), 125–150.
- Vakil, S., McKinney de Royston, M., 2022. Youth as philosophers of technology. *Mind Cult. Activ.* <https://doi.org/10.1080/10749039.2022.2066134>. Available at. (Accessed 29 June 2022).

Learning through work: premises, conceptions and practices

Stephen Billett, Griffith University, Queensland, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Learning through work	381
Explaining learning through work: premises	382
Practice curriculum, pedagogies, and epistemologies	383
Practice or workplace curriculum	384
Ordering of experiences: curriculum pathways of work	384
Practice pedagogies	385
Personal epistemological practice	386
Securing occupational capacities through practice	388
References	388

Learning through work

Throughout human history, across continents and cultures, most occupational skills have been learned through experiences in workplaces, in ways quite distinct from what occurs in educational institutions (Dorman, 2012; Eraut, 2007; Roodhouse, 2007). These capacities have been learned by workers, rather than being taught by others (Rogoff, 1990). Understanding how learning arises through work requires explanatory premises about, and conceptions of, this phenomenon. Yet, these explanations need to be premised on workplaces as sites of learning, not those used to explain the practices and outcomes of educational institutions (Berryman, 1993; Billett et al., 2018b; Tynjala, 2008). This distinction is critical as there are important lessons here for addressing both the initial preparation of occupational capacities and their ongoing development across lengthening working lives. A fundamental difference in these explanations is that this learning and advancing of occupational practices have arisen through processes primarily based on individuals' active processes of learning, not through teaching by others (Webb, 1999). These processes have been essential for human survival and the progress of human cultures in the past (Donald, 1991; Jordan, 1989), and are likely to remain so in the future.

Across human history, local workplaces have been the sites for this learning: in Europe (Greinert, 2002), India (Menon and Varma, 2010), Japan (Singleton, 1989), China (Ebrey, 1996), Greece (Lodge, 1947), Mesopotamia (Finch and Crunkilton, 1992), and, likely, elsewhere. These sites and processes have been central to meeting the needs and wants of humans across diverse cultures and continents. In Europe, family-based occupational preparation was displaced only by industrialization (Greinert, 2002), a situation repeated elsewhere such as in China (Gowlland, 2012). The evidence suggests that this learning has arisen largely through individuals' participation in occupational practice, their active mediation of what they experienced, and processes referred to as mimesis (Byrne and Russon, 1998) or mimetic learning (Billett, 2014). The learning is that arising through individuals' active engagement in observation, imitation, and practice, not through being taught (Marchand, 2008; Singleton, 1989). This situation can be difficult to comprehend in contemporary schooled societies, where experiences in and through educational institutions are taken as the most accepted and effective sites for learning, and with that learning being premised upon teaching (i.e. the passage of knowledge from a more informed to a less informed person). This sometimes leads to teaching and learning being seen as synonymous. Such a view would have been inconceivable prior to the era of modernity and industrialization and the development of schooled societies. Before industrialization, only a tiny percentage of the population would have attended schools or other educational institutions and the focuses of much of those educational programs would have been far removed from developing occupational expertise, nor would they have been attended by those who would be interested in that kind of practice.

Consequently, it is important to identify and articulate premises and concepts that can explain how this learning occurs. Proposing premises and concepts that are distinct from those that are orthodox makes it even more important to be clear about how we should understand and appraise workplaces as learning environments, through carefully introducing, discussing, and advancing these precepts. These elaborations also offer ways to consider how the securing of occupational capacities can progress in contemporary times and are introduced below. Then, drawing on anthropological, historical, psychological, and sociological disciplines, considerations of the kinds of curriculum, pedagogic, and personal practices supporting learning through work are advanced. Firstly, models of workplace curriculum are explained and illustrated as pathways of work activities along which novices and newcomers can progress. Then, some examples of practice pedagogies are offered as a means of augmenting experiences in work settings. These are nothing more or less than affordances (invitations to engage), and their effectiveness in supporting learning is premised on the personal epistemological practices of the workers and students who participate in and learn in the circumstances of work.

Explaining learning through work: premises

Some of the concepts advanced above go against much orthodox understanding about learning through educational programs. This includes the privileging of accessing and learning the knowledge required for work through practice, rather than being taught (Dornan, 2012; Eraut, 2004; Tynjala, 2008). So, it is important that some explanations are advanced. These are four-fold. Firstly, the knowledge required to be learned for performance at work comprises the occupation's domain-specific conceptual, procedural (Stevenson, 2002), and dispositional knowledge (Perkins et al., 1993) and these have:

- canonical (occupational) and
- situational (workplace) levels (Billett et al., 2018a,b).

The canonical knowledge is what would be expected to be possessed by all individuals who practice a specific occupation. Accounts of that knowledge are often captured as statements of professional competence and goals and content for educational programs. Beyond the canonical occupational knowledge are the specific requirements of workplaces that require a situated manifestation of that occupational knowledge. The canonical occupational knowledge is that associated with studies of expertise, which is often viewed as the individual possessing comprehensive, well-structured and organized knowledge of that domain (Glaser, 1984). However, it is also noted that the key quality of expertise—being able to resolve nonroutine problems—also has a situational requirement (Billett, 2001a,b,c). The tasks a mechanic does in an inner-city vehicle dealership may be quite different from the work of a mechanic in a small country town or regional center. From their engagement in work activities such as these, individuals construct the instances of knowledge they encounter as personal domains of occupational knowledge (Billett, 1999). Their personal structuring of that knowledge will be shaped as occupational domain-specific:

- conceptual knowledge—“knowing that” (Ryle, 1949) i.e. occupational concepts, facts, propositions—surface-to-deep conceptual knowledge (Glaser, 1989);
- procedural knowledge—“knowing how” (Ryle, 1949; i.e. how we achieve goals through thinking and acting [both specific and strategic occupational procedures]; (Anderson, 1993; Stevenson, 2002; Sun et al., 2001); and
- dispositional knowledge—“knowing for” (i.e. values, attitudes) related to occupational and situated instances of practice (Perkins et al., 1993), including criticality.

Respectively, these three forms of knowledge represent what we know (conceptual), can do (procedural), and value (dispositional).

Secondly, there is no separation between engaging in activities—such as work—and learning (Lave, 1993). When we engage in everyday goal-directed activities we are also learning micro-genetically, or moment-by-moment (Rogoff and Lave, 1984). When we undertake tasks that are familiar to us, we refine and hone what we already know, can do, and value. When we engage in activities that are new or novel to us, we generate new knowledge and, potentially, ways of knowing. Moreover, the work activities and interactions in which we engage change constantly (Darrach, 1996). Across all kinds of social settings our activities are generative of learning, albeit of different kinds and innovations as we revise them. Importantly, that process also comprises the remaking of occupational tasks, as individuals engage in work activities in particular settings and respond to specific problems (Billett et al., 2005). Most often, that remaking is about adapting what we know to changing circumstances; however, it can also be transformational when novel responses are required for emerging problems or workplace tasks. So, as we engage in work activities, there are two legacies: learning, and the remaking of occupational practices.

Thirdly, the knowledge that needs to be learned arises from these interactions, yet is itself a product of history, culture, and situation (Scribner, 1985). The process of accessing, engaging, and remaking that knowledge occurs through inter-psychological processes: between the person and the social world (Wertsch and Tulviste, 1992). What is suggested here is that these processes are characterized by interdependence (i.e. the reciprocal engagement between individuals and the world in which they engage).

Fourthly, the richness of these learning processes is premised on the duality between the kinds of interactions and activities that individuals are afforded (Fuller and Unwin, 2004), and on how individuals elect to engage with them (Billett, 2001b). Consequently, it is more than the provision of experiences that is important, but also the interdependence between what experiences are afforded, and the individual engaging in and with those experiences. This is the duality which comprises workplace performance and individuals' engagement.

A program of research has identified these affordances as being four-fold, and each makes contributions to learning through everyday work activities that have been identified to include:

- (i) engagement in work tasks (“just doing it”)—legacies of goal-directed activities (cognitive and socio-cultural constructivist accounts);
- (ii) indirect guidance provided by the setting (“just being there”)—observation and imitation (cognitive and neuro-science accounts);
- (iii) practice within that setting—practicing, rehearsing, refining, and associating (cognitive accounts of procedural and conceptual development); and
- (iv) close guidance (proximal) by other practitioners and experts—assisting the development of knowledge that cannot be learned through discovery (Billett, 2001a).

Evident in this list is the key role played by the learners. Three of the four contributions are largely mediated by the learners, with only one—close guidance—being premised upon others' mediation of individuals' learning. Noteworthy is that it is incorrect to view affordances as always being constructive and helpful in shaping the required occupational practice: The same program of research also identified limitations of learning through practice through what has been afforded by workplaces. These include:

- learning that is inappropriate (i.e. bad, unhelpful knowledge and perilous ways of doing things);
- lack of workers' access to work activities and workplace guidance;
- workers not understanding the goals for performing tasks;
- reluctance of experienced co-workers/experts/supervisors to provide guidance to workers;
- absence of expert guidance in the workplace; and
- reluctance of workers to participate effortfully in learning (Billett, 2001a).

So, what is afforded workers—what is provided for them and they can experience—variously either supports or inhibits what can be learned through everyday work activities.

This section has sought to set out some of these premises to understand learning in and through work, and their similarities and distinctions with participating in experiences provided through educational settings. It follows that improving learning in the workplace necessitates trying to secure or enhance the contributions to learning that comprises the affordances of workplace experiences listed above and then attempting also to redress the limitations listed here. In the next section, consideration is given to how workplaces can be effective learning environments through curriculum and pedagogic practices, and learners' personal epistemologies.

Practice curriculum, pedagogies, and epistemologies

The practice curriculum comprises sets and sequencing of experiences through which individuals learn about their occupational practice. This relates back to the original meaning of the word "curriculum" which refers to the "course" or "pathway to progress along" (Marsh, 2004), as it comprises the sequence of activities that learners progressively engage in as they develop their occupational capacities. Pedagogic practices are activities or interactions that can enrich and augment learning arising from experiences in work settings. Note that the term "pedagogic practices" (Portisch, 2010; Tomasello, 2004) describes ways of supporting learning that go beyond teacherly processes (Fuller and Unwin, 2002). It includes those that arise through engagement in work activities and interactions and with artifacts such as partially completed jobs. An important element of workplace pedagogic practice is that it enables learners to access knowledge that they would not be able to discover on their own. Yet, these pedagogic practices are quite distinct from those enacted in classrooms. Then, there are learners' epistemological practices which comprise the basis on which individuals engage with, construe, and construct knowledge through practice (Billett, 2009). These are a product of individuals' personal histories and ontogenetic development (Scribner, 1985)—their repertoire of engagement in activities and interactions across their lives, and how these practices come to shape the ways in which they engage with what they have learned in their workplaces and how they learn from that process.

These three foundational components of learning through practice are likely to be enacted in tandem in most work-related activities. To illustrate these foundations and their interplay, consider this example of Sue, a third-year medical student (Billett and Sweet, 2015). Sue engages in a longitudinal clinical placement in a medical practice in a small town in regional Australia. For the first 6 weeks, she works alongside a doctor in her consulting room, largely observing how the doctor goes about her work. Along the way, Sue asks questions and the doctor provides explanations and sometimes questions Sue about her knowledge. Throughout this period the learning is based on mimetic processes of observation, engagement, and some practice. Being assigned to this doctor also means that Sue visits a hospital where the doctor works, and she comes to observe medical work in that environment. After 6 weeks, Sue is given a consulting room next to the doctor. Then, with the consent of patients, she engages in history taking and examination and arrives at conclusions about patients' health and likely remedies. At this point, the doctor engages to check the information as soon as it is gathered and to check the basis of Sue's prognosis. These are elements of practice curriculum: the progression along the sequence of activities that incrementally assist in developing learners' capacities to enact their occupation. Later, Sue wanted to know how to take blood from patients. From her experiences at the local hospital, she was able to identify a nurse who was competent in this. Sue then observed this nurse and engaged in episodes of practice so that she became competent in taking blood. In this way she exercised the processes that she had been subjected to, observed, imitated, and then practiced: all elements of mimetic learning.

As can be seen, Sue exercised her personal epistemologies in constructive ways: She asked questions, engaged intensively, and then was able to participate in procedures that allowed her to gain information and make clinical decisions. Along the way, she developed a range of procedural capacities; for instance, her decision to approach the hospital indicates the importance of agency interest. Beyond that, Sue actively engaged with the local community: Rather than returning to her home on weekends, she elected to stay and live in this town and to join the local church and the tennis club. Thus, she engaged with the community and her health care colleagues, which then allowed her to seek advice from the hospital and engage with hospital staff. Indeed, this interaction led to staff members phoning her when interesting cases were being considered at the hospital.

This example illustrates how the curriculum—the sequence of experiences that soon progressed, the use of pedagogic practices such as questioning, observation, and personal engagement—is generative of rich learning. In the next two sections, a brief outlining of practice curriculum and pedagogies, and learners' epistemological practices, is provided.

Practice or workplace curriculum

The literature on how individuals have learned through their work practices across history and diverse cultures suggests that there are two broad categories of approach. Firstly, there is what might be referred to as an apprenticeship as a way of life: learning through the lived experience of being in a community and engaging in its practices (Bunn, 1999; Jordan, 1989; Rogoff, 1990); that is, learning as part of growing up, all being embedded in cultural circumstances (e.g. the lived experience of work, such as family business, farm, or nomadic community). Secondly, there is the deliberate structuring of learning experiences to achieve certain learning outcomes (Bunn, 1999; Chan, 2013; Fuller and Unwin, 2003). The former is likely to be the model that has been dominant throughout human history up until the modern era, the advent of schooling, and preparation for occupations outside the circumstances of their enactment. Even today, in countries other than those with advanced industrial economies, these are the models that continue to exist for skill formation. However, within developed countries, the kinds of immersion in occupational practice that featured in earlier times (i.e. by living and working in the master's house and in a street full of similar workplaces) now rarely exist as a model for learning the capacities required for effective work. So, it is the second category of approach—the deliberate ordering of experiences—that is relevant to much of contemporary work. Helpfully, this ordering of experiences is very consonant with the original meaning of the word curriculum. What is evident in the literature is a range of models of workplace curriculum based upon different kinds of pathways.

Ordering of experiences: curriculum pathways of work

Perhaps the most well-known and acknowledged pathway is the sequencing of activities that move from novices engaging in activities in which there are few consequences if errors are made (low error risk) to those activities in which the consequences of errors are greater (high error risk) (Lave, 1990). Lave's study identified the progressions of tailors' apprentices based on this principle. Her work also identified the pedagogic qualities and intents in this sequencing, in many ways analogous to what occurs in educational settings (Billett, 2006). In my own work investigating the learning pathways of hairdressers, production workers, and room attendants, similar pathways were evident. Equally, a study of junior doctors indicated progression of activities and learning premised on this principle (Sinclair, 1997). However, not all curriculum pathways are based upon considerations of error alone. For instance, the stages or pathways associated with learning to be a potter are premised on access to artifacts (i.e. the potter's wheel) and materials (Singleton, 1989). Table 1 sets out different kinds of sequencing, including the approach that Sue experienced in her parallel

Table 1 Workplace curriculum practices—qualities and ordering of experiences for effective work and learning.

Practice	Description	Purposes
Apprenticeship as a way of life	Engaging individuals in the lived experience of workplaces by participating in their everyday activities and interactions (Bunn, 1999; Jordan, 1989) to understand the practices and requirements for performance (Lave, 1990; Makovichy, 2010).	Identifying goals for workplace performance and learning procedural and conceptual knowledge, including honing procedures, forming propositional associations, and developing strategic procedures.
Ordering of experiences	Providing access to and ordering of experiences required to learn occupational capacities not acquired through everyday work; might entail skill acquisition in stepwise manner from those easy to learn to those which are more difficult (Chan, 2013).	Offering a pathway of experiences through which occupational capacities (conceptual, procedural, and dispositional) can be developed.
Learning curriculum	Organizing access to work activities from those having low consequences when errors occur to those where costs of errors are high; starting with observations to understand goal states and progressing through increasingly demanding work requirements (Lave, 1990). In work where all components carry the same salience (e.g. midwifery, (Jordan, 1989) junior doctors, (Sinclair, 1997), learning might be acquired in a linear order.	Organizing a pathway of experiences whose ordering is based on pedagogic and practice-based considerations, such as increased complexity or error cost.
Learning activities as work conditions permit	Learning staged around workplace imperatives e.g. learning pottery premised on access to potter's wheel: (i) pre-practice observation with apprentice engaged in menial work activities, (ii) tentative experiments with wheel (when not used for productive purposes), (iii) assigned regular practice at wheel, (iv) assigned production tasks at wheel, and (v) a period of work to repay training (Singleton, 1989).	A pathway of experiences based upon access to other workers, equipment, resources, that bring together the imperatives of work, workplace, and learning.
Parallel practice	Individuals engaging in an occupational practice and being monitored and checked by a more experienced partner at key points in task completion (Billett and Sweet, 2015), e.g. doctors seeing patients, taking histories, and conducting examinations, often in parallel with what has been done by a registrar (Sinclair, 1997).	Opportunities to engage in authentic activities while being monitored and checked by a qualified practitioner.

practice. It presents models of workplace curriculum, a description of their characteristics, and the purposes which they can realize. Although far from comprehensive, the range of models here indicates that there are options for pathways depending upon the kind of work being learned and also the kinds of circumstances in which that work and learning co-occurs.

An elaboration of these models can be found elsewhere (e.g. Billett, 2014). Here, they are outlined to indicate the diversity of curriculum models that are a product of exigencies of work and workplaces and, by degree, that offer pathways of learning for workers. Yet, beyond a consideration of these curriculum models as pathways, there are other ways by which experiences in workplaces can be augmented to enrich learning; that is, practice pedagogies.

Practice pedagogies

As proposed earlier in this article, much of the learning required for work arises through individuals' everyday engagement in it and their work practices. These are what permit the process of ongoing development and expose workers to new challenges, from which novel learning arises. Nevertheless, the quality of that learning can be supported and assisted by interactions with others and by the clues and cues provided by artifacts and other sources of information (Fuller and Unwin, 2002; Ingold, 2000; Lave et al., 1984). These are referred to as pedagogic practices and, in so far as they are embedded within the conduct of occupational practice, they are referred to as practice pedagogies. The term "pedagogies" is used to capture more broadly the contributions, that assist and augment learning, and that go beyond those provided directly by more informed others (teachers, experienced workers, supervisors, etc.) through instruction, story-telling, and collaborative working. They extend to artifacts, such as tools and completed (Lave et al., 1984) or half-completed jobs (Makovichy, 2010), and to what can be secured through observing, hearing, and other forms of sensory engagement (Rice, 2010; Tomasello, 1998). These pedagogies are important for augmenting learning through work; in particular, for going beyond the circumstances that might otherwise be limited by individual discovery alone. There are limitations to learning through everyday work activities because not all knowledge is easy to access, or to comprehend when accessed. Certainly, a consideration for contemporary work is that much of the knowledge required for effective work practice is conceptual and symbolic (Barley, 1996) and, therefore, requires being accessed and learned in ways which make it more explicit and accessible (Harteis, 2018). With the increase in the use of electronic technologies, for instance, the kind of knowledge that used to be learned through sensory processes and engagement is becoming a lesser requirement as work tasks become less dependent on manual manipulation (Harteis, 2018). Instead, less accessible forms of conceptual knowledge are now becoming important to learn and need to be engaged with through symbols and other forms of representation that can assist individuals' engagement in and building representations of that knowledge. Hence, there is a requirement for pedagogies that can enrich experiences and support that learning. The word "pedagogic" is used here to encompass a range of contributions that can mediate individuals' thinking and acting in support of that learning. Often, there is a need to go beyond teaching or direct instruction, to identify practices which make knowledge accessible and engaging, particularly when related to the kinds of activities which can be learned and applied. The anthropological literature offers a range of pedagogic practices. For instance, the use of available artifacts (Pelissier, 1991) and the role of story-telling (Jordan, 1989) as a means by which a narrative provides a context, and a way in which individuals can engage in, remember, and recall what they have learned, are important contributions to individuals' mimetic learning. The process of verbalization (Gowlland, 2012) has a similar quality, as it makes the thinking of experts accessible to novices who can engage with what is being stated, then appraise it and relate it to their knowledge and ways of knowing.

Then, there are certain work activities that are pedagogically rich (Billett et al., 2018a,b); that is, they have inherent qualities that make them potentially potent as learning experiences. The handovers held by nurses and other health care workers are representative of these experiences. The ability to discuss, compare, and evaluate patients, their condition(s), treatment(s), and progress, and then to permit prognoses to be constructed, represents rich learning experiences because of the contextual nature of the activities, individuals' engagement in them, and the responsibility drawn from them. In addition, there is the close guidance that can be provided by more experienced workers when what needs to be learned is unlikely to be achieved through individuals' discovery alone (Billett, 2001a; Gowlland, 2012; Rogoff, 1995). Consequently, in those circumstances, engaging with more experienced co-workers, gaining more knowledge, and making that knowledge accessible can be crucial for learning "hard to learn" knowledge. Beyond direct guidance is that which occurs indirectly through observing, listening to, and otherwise indirectly engaging with more experienced co-workers and their actions and outcomes (Gowlland, 2012). There are also specific procedures such as heuristics or tricks of the trade (Billett and Rose, 1997; Sinclair, 1997) which are helpful procedurally for securing outcomes. In terms of conceptual development, mnemonics are helpful for providing reminders, cues, and devices to recall demanding and interrelated knowledge (Sinclair, 1997). Physical contributions can also come from the layout and qualities of the setting (Barker, 1978; Lave et al., 1984) and from being able to access cues and clues, such as partially worked or completed examples of the tasks that need to be accomplished (Makovichy, 2010).

In Table 2, some of these kinds of pedagogic practices are set out. They are listed in the left-hand column, described in the middle column, and their potential purposes are provided in the right-hand column. Again, what is evident is that these pedagogic practices are of different kinds, have distinct qualities, and can be directed toward different kinds of purposes; they are elaborated in other publications (e.g. Billett, 2014).

It will be noted that these pedagogic practices are quite different from those used in classrooms. This is an important distinction: Many teaching strategies and classroom pedagogies are founded on the exigencies of the physical environments of classrooms and their norms and practices. These environments are premised on and designed for instruction for large numbers of individuals

Table 2 Workplace pedagogic practices—activities and interactions that permit performance and enrich learning.

<i>Practice</i>	<i>Description</i>	<i>Purposes</i>
Storytelling	Telling stories about work events and incidents (Jordan, 1989).	Illustrating or capturing concepts or hypothetical formulations to assist in decision-making; legitimate practitioner as expert.
Verbalization	Talking aloud while performing a work task as a form of direct guidance, linked to “hands on” engagement (Gowlland, 2012).	Explaining the thinking and acting being used while performing work tasks.
Pedagogically rich activities	Workplace activities that are inherently pedagogically rich, e.g. handovers (Billett et al., 2018a,b) or case discussion, moderation, or planning meetings.	Developing conceptual, specific, and strategic procedural capacities.
Guided learning (proximal guidance)	Direct interaction between more and less experienced co-workers to promote learning in work settings. Use of modeling, demonstrating, guided practice, monitoring progress, and gradual withdrawal of direct guidance (Billett, 2000; Collins et al., 1989; Rogoff, 1995).	Extending what individuals can learn through discovery alone by modeling of activities to be learned, guidance to assist achieving modeled performance, and providing opportunities to refine and hone.
Partially worked example/direct instruction and hands on	Combination of guidance and using a worked example, e.g. experienced lace-maker producing a small piece of simple lace, showing novice how bobbins are held, and placing hands on those of a novice to assist learning hand movements to use the bobbins; also asking novice which bobbin they should use for the next stitch before novice makes movement, leading to confidence in action; questioning gradually ceasing as competence demonstrated (Makovichy, 2010).	Provision of models for performance, ability to engage in sub-skills associated with that performance and build understanding about procedural capacities.
Heuristics	Tricks of the trade (i.e. procedures that will give you certainty; Billett and Rose, 1997).	Developing procedural capacities through providing devices that assist in securing workplace goals.
Mnemonics	Developing and using mnemonics (doctors’ use of 5 Fs, DANISH, to remember about cerebellar lesions) and actual patients (“remember Mr Leeming and you will remember about duodenal ulcers”, Freddie Mercury and missed seroconversion) (Sinclair, 1997) for procedural efficiency (Rice, 2008, 2010).	A means to remember and recall propositions and secure procedural efficacy.
Artifacts	Embedding the knowledge required in a localized context to assist skill and proficiency (Makovichy, 2010).	Providing clues and cues on how to proceed.

positioned as students, probably modeled on church pulpits. Practice pedagogies, however, are associated with the social and physical circumstances of workplaces and the practices within them and, importantly, are embedded within the work practices to be learned.

These considerations of curriculum and pedagogic practices are what can be afforded to individuals through their workplaces and work activities. However, learning through social practices is premised on the duality between (a) what individuals have been afforded, and (b) how they elect to engage with what has been afforded. This suggests that the contributions, invitations, and affordances are nothing other than potential contributions to individuals’ learning and development. Ultimately, it is how individuals engage with that learning. Central here are individuals’ personal epistemologies that shape how they engage with what has been afforded to them.

Personal epistemological practice

As has been stated, it is insufficient to explain the contributions to and processes of learning through work by referring only to what is afforded by the physical and social setting in which workers engage (i.e. workplaces). Instead, understanding and accounting for this process of learning needs to go beyond what has been afforded and to consider how individuals come to engage with the ways in which they are invited to participate in, and learn through, work activities and interactions (Bauer et al., 2004). As noted earlier, throughout human history the vast majority of learning and development is likely to have arisen through individuals’ personally mediated learning: how they use what they know, can do, and value to construe and construct what they are experiencing. So, when considering individuals’ learning through work, it is important to acknowledge that it is shaped by individuals’ personal epistemologies (their bases for knowing, engaging, and learning) (Bauer et al., 2004; Billett, 2009).

The literature identifies a range of ways in which this personal mediation of experience and the legacies from it (i.e. learning and development) arise. These include what is referred to as imitation, which is held to be an important process for most kinds of

animals but, in particular, for humans, who have taken this process to a high level (Gardner, 2004; Jordan, 1989; Marchand, 2008; Tomasello, 2004). This process is consonant with the process of *mimesis* (i.e. observation, imitation, and action) and with what has been referred to above as mimetic learning, which goes beyond a reliance on observation and proposes that all sensory inputs play a role in human sense-making and learning. There are a range of strategies associated with the deployment of *mimesis* and one of these is referred to as ontogenetic ritualization (Tomasello, 2004). This concept refers to how two interlocutors engage with, and secure access to, what one needs from the other. Based on observations of orangutans, where babies negotiated with their mothers to gain access to the breast, this process has parallels in the way that humans negotiate with each other to gain access to partnerships, to shared activities, and to each other's knowledge.

However, given the personally mediated nature of this form of learning and development, the importance of learner readiness (Bunn, 1999; Singleton, 1989) has come to the fore. Readiness here refers to the level or capacity at which individuals engage in activities and learn from them. For instance, if individuals have no knowledge of a specific language or its grammatical structure, merely being immersed in a situation where that language is being used may not be a productive learning experience, because they are not ready to engage effectively. Furthermore, individuals need to assent to that engagement (Mishler, 2004). Only they will decide on the ways in which they will come to engage in the effortful process of constructing meaning, developing procedural capacities and developing further their values and interests. This is aligned with the process referred to as deliberate practice (Ericsson, 2006; Gardner, 2004; Sinclair, 1997): the effortful, intentional, and focused engagement by individuals to learn, hone, and further develop their capacities. Such activities are most likely to occur when individuals sense that this learning is important and see it as part of sustaining their sense of self. Table 3 sets out a range of these practices, with a description provided and those purposes associated with learning in and through work.

Table 3 Personal epistemological practices—premised on individuals' capacities, agency, intentionality, and sense of self.

Practice	Description	Purposes
Mimetic learning	Ability to imitate and interest to do so at work (Billett, 2014); a fundamental process of human learning (Jordan, 1989) by imitating implicitly or explicitly how experienced workers practice (Marchand, 2008), including asking questions or listening to them (Gardner, 2004)—a higher order activity.	Process through which goals, procedures (both specific and strategic), propositional links, and associations can be learned through personal engagement.
Ontogenetic ritualization	A system of communication and engagement between social partners negotiated through repeated social interactions (Tomasello, 2004); supports close guidance and reciprocal interactions in workplaces permitting co-workers to engage in learning reciprocally.	Basis through which to engage with others from whom guidance is derived.
Embodied knowledge	Ability to do rather than talk about it, arising through practice (Jordan, 1989); renders unconscious much of what is required for performance (Lakoff and Johnson, 1999; Reber, 1992).	Procedural learning derived through practice and rehearsal that becomes able to be applied without conscious recall.
Deliberate practice	Practicing skills and monitoring their improvement (Gardner, 2004) as the means by which agency, focused intentionality, and the act of rehearsal come together to secure effective performance (Anderson, 1982).	Process of rehearsing, honing skills, refining, monitoring personal performance, and making conscious efforts to improve performance.
Guided re-discovery	Novices engaging in situations where they can practice, hone skills, and gain experience by themselves, yet can still access direct guidance (Ingold, 2000; Makovichy, 2010).	Process of remaking knowledge through opportunities for observation, listening, and other forms of social engagement and practice.
Active engagement & construction	Importance of active engagement (observation, listening, and <i>mimesis</i>); apprenticeship means <i>apprehendere</i> in French (to seize, lay hold of, apprehend, and apprise with the mind—to inform, give form, to shape; Webb, 1999).	Learning arises through seeking out, actively constructing, and appropriating the knowledge able to be accessed.
Observation	Japanese word for apprentice is <i>minarai</i> —one who learns by observation; unobtrusive observation <i>minarai kyooiku</i> —serious learning progresses without didactic instruction (Singleton, 1989).	Provides access to goals, models for procedures, and bases for developing procedural capacities.
Averting gaze	Mediating the external (i.e. immediately social) contributions—actively ignoring erroneous suggestions to secure an immediate goal to effect productive activities (Glenberg et al., 1998).	Intentional engagement with what needs to be learned.
Readiness	The level and qualities of what individuals want to know and can do.	Mediates the capacities for learning as a result of earlier experiences (and learning).
Assent	Individuals' willingness to assent to effortful learning (Mishler, 2004).	Mediates how individuals engage with external suggestions, supports.

The kinds, breadth and diversity of these personal practices emphasize the centrality of the active engagement of individuals as learners. This includes the role that they play in initiating (Bauer et al., 2004), directing the intentionality (Malle et al., 2001) and focus of their thinking and acting (Rausch et al., 2017) associated with learning. Much of these practices are premised upon their existing knowledge and ways of knowing, based upon how they have reconciled earlier experiences (i.e. pre-mediate), which is quite central to understanding learning through and for work. Indeed, as with the learning of any other socially derived practice, it is insufficient simply to consider an educational experience, either organized in or through an educational institution or workplace setting. Instead, it is necessary to consider how people come to engage with that experience, which includes the kinds of societal sentiment which informs them whether such an activity is worthwhile and important.

From the above, it can be seen that understanding the processes of learning through work and enhancing those processes is likely premised upon three distinct sets of practices: (i) practice curriculum, (ii) practice pedagogies and (iii) individuals' personal practices, including what they know, can do and value.

Securing occupational capacities through practice

It has been proposed in this article that the growing interest in promoting learning through work, both for the initial development of occupational capacities and for sustaining and developing them further across working life, needs to be supported by understanding the particular ways how learning of occupational knowledge arises through work settings. Consequently, it is also important to identify and articulate the premises that assist in understanding and generating sets of concepts that can explain this phenomenon. These are essential in understanding and appraising workplaces as environments in schooled societies in which there is a privileging of learning processes and outcomes that arise through experiences in educational settings. The conceptual and explanatory elaborations advanced here offer bases to consider how the securing of occupational capacities can progress in contemporary times. Moreover, to achieve the outcomes which are now being sought by government, workplaces, workers, and educational institutions, it is necessary to advance ways in which these can be realized. Proposed here within the broad framework of a duality between workplace affordances and individual engagement are the key tenets of practice curriculum and pedagogies, and personal epistemologies. Listed above are a range of proposed models of practice curriculum that are, in different ways, applicable to specific kinds of workplaces and occupations. The listing of pedagogic practices has particular processes and purposes that they are likely to promote. And at least as important are the learners' personal epistemologies with their attendant processes and qualities. It is proposed that, together, these three elements help explain not only how learning through work can progress, but also the ways in which learning through work can be promoted to achieve the kinds of outcomes required for contemporary workplaces. Hopefully, these suggestions might assist with this important task.

References

- Anderson, J.R., 1982. Acquisition of cognitive skill. *Psychol. Rev.* 89 (4), 369–406.
- Anderson, J.R., 1993. Problem solving and learning. *Am. Psychol.* 48 (1), 35–44.
- Barker, R.G., 1978. *Habitats, Environments and Human Behavior*. Jossey-Bass, San Francisco.
- Barley, S., 1996. Technicians in the workplace: ethnographic evidence for bringing work into organizational studies. *Adm. Sci. Q.* 41, 404–441.
- Bauer, J., Festner, D., Gruber, H., Harteis, C., Heid, H., 2004. The effects of epistemological beliefs on workplace learning. *J. Workplace Learn.* 16 (5), 284–292.
- Berryman, S., 1993. Learning for the workplace. *Rev. Res. Educ.* 19, 343–401.
- Billett, S., Rose, J., 1997. Developing conceptual knowledge in the workplace. *Aust. J. Adult Community Educ.* 37 (1), 12–26.
- Billett, S., Sweet, L., 2015. Understanding and appraising healthcare students' learning through workplace experiences: participatory practices at work. In: Cleland, J., Durning, S. (Eds.), *Researching Medical Education*. Wiley, Oxford, UK, pp. 117–127.
- Billett, S., Smith, R., Barker, M., 2005. Understanding work, learning and the remaking of cultural practices. *Stud. Cont. Educ.* 27 (3), 219–237.
- Billett, S., Noble, C., Sweet, L., 2018a. Pedagogically-rich activities in hospital work: handovers, ward rounds and team meetings. In: Delany, C., Molloy, L. (Eds.), *A Practical Guide for Learning and Teaching in a Clinical Context*. Elsevier, Australia, pp. 207–220.
- Billett, S., Harteis, C., Gruber, H., 2018b. Developing occupational expertise through everyday work activities and interaction. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A. (Eds.), *Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, New York, pp. 105–126.
- Billett, S., 1999. Experts' ways of knowing. *Aust. Vocat. Educ. Rev.* 6 (2), 25–36.
- Billett, S., 2000. Guided learning at work. *J. Workplace Learn.* 12 (7), 272–285.
- Billett, S., 2001a. *Learning in the Workplace: Strategies for Effective Practice*. Allen and Unwin, Sydney.
- Billett, S., 2001b. Learning through work: workplace affordances and individual engagement. *J. Workplace Learn.* 13 (5), 209–214.
- Billett, S., 2001c. Knowing in practice: re-conceptualizing vocational expertise. *Learn. Instruct.* 11 (6), 431–452.
- Billett, S., 2006. Constituting the workplace curriculum. *J. Curric. Stud.* 38 (1), 31–48.
- Billett, S., 2009. Personal epistemologies, work and learning. *Educ. Res. Rev.* 4, 210–219.
- Billett, S., 2014. *Mimetic Learning at Work: Learning in the Circumstances of Practice*. Springer, Dordrecht, The Netherlands.
- Bunn, S., 1999. The nomad's apprentice: different kinds of apprenticeship among Kyrgyz nomads in Central Asia. In: Ainley, P., Rainbird, H. (Eds.), *Apprenticeship: Towards a New Paradigm of Learning*. Kogan Page, London, pp. 74–85.
- Byrne, R.W., Russon, A., 1998. Learning by imitation: a hierarchical approach. *Behav. Brain Sci.* 21 (5), 667–721.
- Chan, S., 2013. Learning through apprenticeship: belonging to a workplace, becoming and being. *Vocat. Learn.* 6 (3), 367–383.
- Collins, A., Brown, J.S., Newman, S.E., 1989. Cognitive apprenticeship: teaching the crafts of reading, writing and mathematics. In: Resnick, L.B. (Ed.), *Knowing, Learning and Instruction: Essays in Honor of Robert Glaser*. Erlbaum & Associates, Hillsdale, NJ, pp. 453–494.
- Darrah, C.N., 1996. *Learning and Work: An Exploration in Industrial Ethnography*. Garland Publishing, New York.
- Donald, M., 1991. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press, Cambridge, MA.

- Dornan, T., 2012. Workplace learning. *Perspect. Med. Educ.* 1 (1), 15–23.
- Ebrey, P.B., 1996. China: Illustrated History. Cambridge University Press, Cambridge, UK.
- Eraut, M., 2004. Informal learning in the workplace. *Stud. Cont. Educ.* 26 (2), 247–273.
- Eraut, M., 2007. Learning from other people in the workplace. *Oxf. Rev. Educ.* 33 (4), 403–422.
- Ericsson, K.A., 2006. The influence of experience and deliberate practice on the development of superior expert performance. In: Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffmann, R.R. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, Cambridge, pp. 685–705.
- Finch, C.R., Crunkilton, J.R., 1992. Curriculum Development in Vocational and Technical Education: Planning, Content and Implementation, fifth ed. Allyn and Bacon, Boston.
- Fuller, A., Unwin, L., 2002. Developing pedagogies for the workplace. In: Evans, K., Hodgkinson, P., Unwin, L. (Eds.), *Working to Learn: Transforming Workplace Learning*. Kogan-Page, London.
- Fuller, A., Unwin, A., 2003. Fostering workplace learning: looking through the lens of apprenticeships. *Eur. Educ. Res. J.* 2 (1), 41–55.
- Fuller, A., Unwin, L., 2004. Expansive learning environments: integrating organizational and personal development. In: Rainbird, H., Fuller, A., Munroe, A. (Eds.), *Workplace Learning in Context*. Routledge, London, pp. 126–144.
- Gardner, H., 2004. What we do & don't know about learning. *Daedalus* 133 (1), 5–12.
- Glaser, R., 1984. Education and thinking—the role of knowledge. *Am. Psychol.* 39 (2), 93–104.
- Glaser, R., 1989. Expertise and learning: how do we think about instructional processes now that we have discovered knowledge structures? In: Klahr, D., Kotovsky, K. (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon*. Erlbaum & Associates, Hillsdale, NJ, pp. 289–317.
- Glenberg, A.M., Schroeder, J.L., Robertson, D.A., 1998. Averting the gaze disengages the environment and facilitates remembering. *Mem. Cognit.* 26 (4), 651–658.
- Gowland, G., 2012. Learning craft skills in China: apprenticeship and social capital in an artisan community of practice. *Anthropol. Educ. Q.* 43 (4), 358–371.
- Greiner, W.-D., 2002. (October). European and vocational training systems: the theoretical context of historical development. In: *Towards a History of Vocational Education and Training (VET) in Europe in a Comparative Perspective*. Retrieved from: www.cedefop.europa.eu/files/5153_1_en.pdf.
- Harteis, C. (Ed.), 2018. *The Impact of Digitalization in the Workplace*, vol. 21. Springer, Dordrecht.
- Ingold, T., 2000. *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. Routledge, London.
- Jordan, B., 1989. Cosmopolitical obstetrics: some insights from the training of traditional midwives. *Soc. Sci. Med.* 28 (9), 925–944.
- Lakoff, G., Johnson, M., 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, New York.
- Lave, J., Murtaugh, M., de la Roche, O., 1984. The dialectic of arithmetic in grocery shopping. In: Rogoff, B., Lave, J. (Eds.), *Everyday Cognition: Its Development in Social Context*. Harvard University Press, Cambridge, Mass, pp. 76–94.
- Lave, J., 1990. The culture of acquisition and the practice of understanding. In: Stigler, J.W., Shweder, R.A., Herdt, G. (Eds.), *Cultural Psychology*. Cambridge University Press, Cambridge, UK, pp. 259–286.
- Lave, J., 1993. The practice of learning. In: Chaiklin, S., Lave, J. (Eds.), *Understanding Practice: Perspectives on Activity and Context*. Cambridge University Press, Cambridge, UK, pp. 3–32.
- Lodge, R.C., 1947. *Plato's Theory of Education*. Kegan Paul, Trench, Trubner, London.
- Makovichy, N., 2010. "Something to talk about": notation and knowledge-making among Central Slovak lace-makers. *J. Roy. Anthropol. Inst.* 16 (Suppl. S1), 80–99.
- Malle, B.F., Moses, L.J., Baldwin, D.A., 2001. Introduction: the significance of intentionality. In: Malle, B.F., Moses, L.J., Baldwin, D.A. (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. The MIT Press, Cambridge, Massachusetts, pp. 1–26.
- Marchand, T.H.J., 2008. Muscles, morals and mind: craft apprenticeship and the formation of person. *Br. J. Educ. Stud.* 56 (3), 245–271.
- Marsh, C.J., 2004. *Key Concepts for Understanding Curriculum*. RoutledgeFalmer, London.
- Menon, J., Varma, S., 2010. Children playing and learning: crafting ceramics in ancient Indor Khera. *Asian Perspect.* 49 (1), 85–109.
- Mishler, E.G., 2004. *Storylines: Craftartists' Narratives of Identity*. Harvard University Press, Cambridge, MA.
- Pelissier, C., 1991. The anthropology of teaching and learning. *Annu. Rev. Anthropol.* 20, 75–95.
- Perkins, D., Jay, E., Tishman, S., 1993. Beyond abilities: a dispositional theory of thinking. *Merrill-Palmer Q.* 39 (1), 1–21.
- Portisch, A.O., 2010. The craft of skillful learning: Kazakh women's everyday craft practices in western Mongolia. *J. Roy. Anthropol. Inst.* 16 (Suppl. 1), 62–79.
- Rausch, A., Seifried, J., Harteis, C., 2017. Emotions, coping and learning in error situations at work. *J. Workplace Learn.* 29 (Five), 374–393.
- Reber, A.S., 1992. An evolutionary context for the cognitive unconscious. *Phil. Psychol.* 5 (1), 33–51.
- Rice, T., 2008. Beautiful murmurs: stethoscopic listening and acoustic objectification. *Senses Soc.* 3 (3), 293–306.
- Rice, T., 2010. Learning to listen: auscultation and the transmission of auditory knowledge. *J. Roy. Anthropol. Inst.* 16, S41–S61.
- Rogoff, B., Lave, J. (Eds.), 1984. *Everyday Cognition: Its Development in Social Context*. Harvard University Press, Cambridge, MA.
- Rogoff, B., 1990. *Apprenticeship in Thinking—Cognitive Development in Social Context*. Oxford University Press, New York.
- Rogoff, B., 1995. Observing sociocultural activity on three planes: participatory appropriation, guided participation, apprenticeship. In: Wertsch, J.W., Alvarez, A., del Rio, P. (Eds.), *Sociocultural Studies of Mind*. Cambridge University Press, Cambridge, UK, pp. 139–164.
- Roodhouse, S., 2007. Special issue introduction. *Educ. Train.* 49 (3), 161–169.
- Ryle, G., 1949. *The Concept of Mind*. Hutchinson University Library, London.
- Scribner, S., 1985. Vygotsky's use of history. In: Wertsch, J.V. (Ed.), *Culture, Communication and Cognition: Vygotskian Perspectives*. Cambridge University Press, Cambridge, UK, pp. 119–145.
- Sinclair, S., 1997. *Making Doctors: An Institutional Apprenticeship*. Berg, Oxford.
- Singleton, J., 1989. The Japanese folkcraft pottery apprenticeship: cultural patterns of an educational institution. In: Coy, M.W. (Ed.), *Apprenticeship: From Theory to Method and Back Again*. SUNY, New York, pp. 13–30.
- Stevenson, J., 2002. Concepts of workplace knowledge. *Int. J. Educ. Res.* 37, 1–15.
- Sun, R., Merrill, E., Peterson, T., 2001. From implicit skills to explicit knowledge: a bottom-up model of skill development. *Cognit. Sci.* 25, 203–244.
- Tomasello, M., 1998. Emulation learning and cultural learning. *Behav. Brain Sci.* 21, 703–704.
- Tomasello, M., 2004. Learning through others. *Daedalus* 133 (1), 51–58.
- Tynjälä, P., 2008. Perspectives into learning in the workplace. *Educ. Res. Rev.* 3 (2), 130–154.
- Webb, E., 1999. Making meaning: language for learning. In: Ainley, P., Rainbird, H. (Eds.), *Apprenticeship: Towards a New Paradigm of Learning*. Kogan Page, London, pp. 100–110.
- Wertsch, J.V., Tulviste, P., 1992. L. S. Vygotsky and contemporary developmental psychology. *Dev. Psychol.* 28 (4), 548–557.

Learning spaces and materiality: embracing past and present when designing for the future

Lucila Carvalho, Massey University, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	390
Embracing the past—the materiality of a marae and its long-lasting values	391
Embracing the present—shaping spaces for networked learning in contemporary societies	393
Framing the analysis of complex learning environments	394
Applying the ACAD lens to micro, meso and macro levels	395
ACAD and a Māori marae	395
ACAD and innovative or learning environments	395
Conclusion	396
References	397

Introduction

Archaeologists, urban planners, and designers are constantly examining connections between materials, the built environment and human behavior. Their respective fields have accumulated extensive bodies of theoretical, empirical, and practical work, with practitioners especially trained to reflect on the nature of human activity as it relates to culture and built forms. Conversely, in education, the influence of materiality in learning has gone largely unnoticed for many years. Traditionally, educators have focused on curriculum design, pedagogical practices and social interactions, while spaces, built forms, digital and material elements, are usually seen as a backdrop for where the “real” action happens—they have too often been neglected aspects of the complex learning entanglement in education (Carvalho and Yeoman, 2021).

This chapter explores the intersection between spaces, materials, culture and learning, and problematizes the need to look at past and present when designing for learning. The chapter suggests that paying attention to the qualities of materials, and the ways these might influence learning, is best considered within an ecology of elements—through searching part-whole relationships and noticing how multiple elements relate to one another and to learning activity (Carvalho and Yeoman, 2018). In other words, learning activity is seen as enmeshed in an assemblage of elements that includes physical spaces, artifacts, digital technologies, people, ideas, all coming together in a learning entanglement. Drawing on literature about a Māori *marae* in the context of Aotearoa New Zealand, the chapter makes a case for the materiality of a marae and its pedagogical function, as a space that reflects long-standing social values of positive interdependence and collaboration—values that have been around for thousands of years (Addis et al., 2011; Austin, 2011). The chapter argues that many of these long-standing values are being reframed, reformed, and reinstated in both material and digital incarnations (in learning entanglements), where their fundamental essence is still present and still influencing and shaping the ways students are being prepared for the future.

In addressing the present times, the chapter discusses how learning spaces have increased in complexity and are becoming more and more ubiquitous (Beetham and Sharpe, 2019). The chapter delves into the qualities of materials and learning spaces in modern society, discussing these in connection to the challenges of preparing our students for an evolving digitally mediated future, and therefore, the challenges associated with design for learning. Drawing on theories that foreground socio-materiality in learning (Fenwick, 2015; Sørensen, 2009) and networked learning (Goodyear et al., 2004; Networked Learning Editorial Collective, 2020), the chapter calls for a deeper understanding of the interplay between the physical (digital and material), conceptual, and social aspects of learning, and their combined influence on emergent activity.

Themes related to spaces, materials, culture and learning are explored as follows. First, the chapter brings a reflection about the materiality of a Māori marae, discussing its pedagogical function and the ways “learning,” “knowing” and “being” unfold in these spaces (Addis et al., 2011; Austin, 2011). Second, the chapter problematizes learning as part of a networked society, considering the role of material and digital elements, within learning ecologies in education (Carvalho and Yeoman, 2018, 2021; Damşa et al., 2019). The need to nurture skills connected to collaboration and critical participation in contemporary society is also foregrounded (Ananiadou and Claro, 2009; Sawyer, 2009). Third, the chapter discusses an architecturally inspired analytical framework that helps illuminate the structure of complex learning environments (Goodyear and Carvalho, 2014; Goodyear et al., 2021). And, finally, the chapter illustrates how ideas in this framework can be used to map relations between spaces, materials and learning within educational environments.

Overall, the chapter argues that materials matter, and that in order to best design for learning, educators need to notice the qualities of materials, while making sure these are not seen as isolated entities. Instead, it is important to account for multiple elements and to consider carefully how they influence the learning whole, searching for key designable components and understanding what

works well, and how. Only then, can we use this knowledge to (re)design our learning environments. As educators, it is important that we acknowledge and build on the past and present, if we are to figure out how best to prepare our students for an uncertain future.

Embracing the past—the materiality of a marae and its long-lasting values

Aotearoa New Zealand is a bi-cultural country in the south pacific, which recently reached a population of 5 million. History books tell us that when Captain Cook and his crew first arrived in 1769, Māori people had already been living in the land for nearly a thousand years (Austin, 2011). Over time, many other migrants also came to these shores. The 2018 census shows 70% of Aotearoa New Zealand population as European descendants, and 16% as Māori (Statistics New Zealand, 2020). Since the 1700s, there has been an increase in population diversity, Māori becoming a minority according to the census data. But even so, one could argue that the traditions of the Māori culture are still strongly infused in Aotearoa New Zealand society, in everyday living—deeply reflected on the nations’ identity, cuisine, language, and customs (Te Ara, 2020).

Māori studies have focused on social relations, art, culture and traditional knowledge, but some scholars have also specifically explored the architecture of Māori settlements and the interaction of Māori people with the built environment (Austin, 2011). Traditionally, Māori settlements were formed around a community building (or a meeting house) and its associated open space (the marae) (Austin, 2011). A marae therefore consists of “a cluster of buildings and spaces, usually incorporating a wharenui (meeting house), a marae ātea (an open space in front of the wharenui), a kitchen and dining area, and an amenities block” (Statistics New Zealand, 2014, p. 6). A marae may also include other spaces such as an urupā (cemetery). Over the years, as Māori migrated from rural areas to the cities, the phenomenon of an ‘urban marae’ also emerged, referring to their presence outside the context of the traditional tribal structures often in rural areas (Addis et al., 2011). Urban maraes provide a cultural space for Māori, particularly those who might have been distanced from their closest ones, after relocating to larger cities. These urban maraes allow for interaction with others who are perhaps experiencing a similar situation, so that people can connect, explore and understand their identity. Interestingly, maraes are not necessarily found only in connection to Māori in Aotearoa New Zealand, but there are also maraes throughout Polynesia in various forms (Addis et al., 2011).

A marae is essentially a space where a community gathers together for weddings, funerals, for community discussions and other social events. A marae belongs to a particular iwi (or a tribe), or to a whānau (family), and Māori see their marae as their place “to stand and belong” (Statistics New Zealand, 2014, p. 6). As such, maraes are places infused with symbolism and meaning, where the Māori community stands together but, at the same time, these are also physical spaces, with a meeting house, which has a specific architectural characteristic and layout, which often displays intricately ornated carvings, paintings and patterns (Fig. 1).

Māori traditions and knowledge have, for many years, been orally passed on via stories and songs, through many generations. As such, a marae plays a crucial role for the Māori community. Ka’ai (2008) refers to the symbolic cultural pedagogical significance of a marae, as an environment that allows teaching and learning and practice of traditional Māori pedagogies such as ako. Ako means both to teach and to learn—a single word that embodies learning as reciprocal, through the notion that knowledge and new understandings emerge through learning, but also through teaching (Macfarlane, 2015; Napan et al., 2020). In considering the pedagogical function of a marae, one positions learning activity as happening beyond schools or formal contexts, as occurring, for example, while one works alongside elders, and watches, observes and learns behaviors under their guidance (Addis et al., 2011). As a pedagogical tool, the marae is seen as an environment that nurtures positive interdependence and collaboration, helping to create a sense of whānau (or a sense of family), an environment where essential long-standing Māori values are passed on. These include core



Fig. 1 Meeting house at Te Whaiti. Original photographic prints and postcards from file print collection, Box 1. Ref: PAColl-5471-013. Alexander Turnbull Library, Wellington, New Zealand <http://www.records/22524369>.

cultural values such as having a sense of shared responsibility, learning to work with one another, showing commitment, nurturing positive interaction, connectedness, and solidarity (Addis et al., 2011).

As a community gathers together in a marae, they are encouraged to engage with particular ways of learning, knowing and being, where elders might tell stories and sing songs and, as the younger generation experiences these traditions, they do so, within a particular physical setting. This physical setting also arguably plays a significant role in the passing on of Māori traditions. As Addis et al. (2011) explain, the communal large space and the intricate architectural features of a meeting house—represented through carvings and paintings that decorate the space—are often used as mnemonic aids, helping the younger and old to memorize genealogies and the values of the culture. Surface patterns and paintings function as references, and through touch and sight might help people capture and share their stories (Fig. 2). A marae meeting house reflects an atmosphere that seems created to nurture cultural values, where the types of interaction give shape to a type of “pedagogy” infused with a mission of passing on the traditions of a culture. But the patterns, carvings and decorations themselves are also symbolically holding pieces of Māori stories. Features in the walls and ceilings act as reminders to those who are listening and to those who are telling these stories.

In thinking about a marae and the types of activities that unfold in these spaces, a beautiful example emerges to illustrate the power of the performativity of material things—where materials and the surroundings subtly suggest specific types of interaction and learning. A socio-materialist stance often challenges us to move away from an isolated focus on either entities or individuals, toward a more nuanced understanding of material relationality in learning (Fenwick and Edwards, 2011, 2014; Gourlay and Oliver, 2018; Sørensen, 2009). Socio-materialism speaks of knowledge, people, objects and systems, as influenced by connections and activity, a perspective where performativity comes into action through complex webs of relations. As Fenwick (2015) stresses, things “act together with other types of things and forces to exclude, invite, and regulate particular forms of participation” (p. 670). And it is easy to imagine a marae subtly inviting specific forms of participation: the open plan of a meeting house “speaks” of gathering together in a community hall, the carving and paintings “support” the oral stories, helping the passing of the traditions and values of this culture.

The essence of Māori values such as the shared responsibility, working with one another, commitment, positive interaction, connectedness, and solidarity (Addis et al., 2011), can also be seen as infused broadly in Aotearoa New Zealand’s education system. From 2015 onwards, Aotearoa New Zealand’s policies, standards and guidelines (Benade, 2017; Education Review Office, 2018; New Zealand Ministry of Education, 2015) started to reflect global trends in education, which encouraged schools and universities to transition into new spatial configurations, or into flexible learning spaces (Cardellino et al., 2017; Daniels et al., 2018; Wood, 2019; Woolner et al., 2018). As such, schools and university spaces have been (re)designed not only to accommodate technology-rich activity, but also to enable collaborative teaching and learning in ways that actively engage students in scaffolded inquiry (Mulcahy et al., 2015). Te Kete Ipurangi (2020), an educational portal created by the New Zealand Ministry of Education, suggests that working in an innovative environment involves considerations about the physical, social, and pedagogical aspects of learning. These types of environment support strengths-based teaching and learning and promote flexibility, agency, ubiquity, and connectedness. In these environments “reflections and inquiries are shared, and communities engaged [which often] leads to a more robust, continuously improving community of practice” (Te Kete Ipurangi, 2020, online). These views of an innovative learning environment seem connected to positive interdependence and collaboration, and often acknowledge that navigating these changes from traditional to flexible spaces in formal education settings requires experimentation, adaptation and deep consideration about what is valued and what is possible in newly configured spaces. Te Kete Ipurangi (2020) refers to the role of emerging technologies and the development of teaching practices that support team teaching and collaboration. Similar emphasis on *connections* is also reflected through the notion of networked learning, a field of study that examines the use of technology in the context of learning and as a tool to connect to others, to learning resources and to a learning community (Goodyear et al., 2004).



Fig. 2 Interior of Hinenuitepo meeting house at Te Whaiti, 1930. Godber, Albert Percy, 1875–1949: Collection of albums, prints and negatives. Ref: APG-1670-1/2-G. Alexander Turnbull Library, Wellington, New Zealand <http://www.records/23070749>.

Embracing the present—shaping spaces for networked learning in contemporary societies

A recent definition of networked learning explains that:

Networked learning involves processes of collaborative, co-operative and collective inquiry, knowledge-creation and knowledgeable action, underpinned by trusting relationships, motivated by a sense of shared challenge and enabled by convivial technologies. Networked learning promotes connections: between people, between sites of learning and action, between ideas, resources and solutions, across time, space and media.

Networked Learning Editorial Collective (2020, p. 8).

Indeed, people are living in increasingly connected societies. Some may use the term *connections* to refer to the wider access to digital technologies or the ability to communicate to one another in an instant (Dohn, 2018). There is no doubt, it is now possible to experience simultaneously multiple places and situations (Gourlay and Oliver, 2018), like being in a café in New York while chatting to a friend in Bangladesh, or watching a teleconference seminar presentation broadcast in London, while stroking a cat in your home lounge in Sydney. But the term *connections* can also be used to reflect how what happens to one of us at one side of the world, often reverberates elsewhere through a chain of interdependent links, and in a range of situations. Essentially, no matter where one is in this world, there are common issues that affect all of us as global citizens. Famine, global warming, the displacement of refugees, human rights violation and recently the COVID-19 pandemic are just some examples. When young people participated in protests led by climate activist Greta Thunberg (2019), they did so to raise collective awareness of the need to take action as a community and engage with climate issues in a more proactive way. The pandemic arrived and reminded us again about the need to think as a community, the need to look after one another, if we are to eliminate this threatening virus. These are highly complex global issues and, arguably, would be best approached if people from different disciplinary areas worked co-operatively, jointly searching for effective responses. Often called “wicked problems”—these are problems that are ill-defined, for which there is more than one possible explanation and for which there might be multiple potential solutions (Buchanan, 1992). They demand that people work with one another, sharing ideas, exchanging expertise, and it is in trying to reach a shared understanding of the issue at hand, that collective action can successfully take shape, allowing society to move forward.

Now more than ever educators need to look beyond preparing students for future employability only, and instead help students develop the capabilities so desperately needed, so that they can contribute to the shaping of sustainable ways of living. Educators need to support young people to be empowered, to feel that they can work together to create and build a better world for all of us.

A core challenge for educators is indeed how to create spaces for learning in ways that support young people to learn how to collaborate, how to think creatively and critically, how to facilitate students’ engagement with problem-solving activity or how to help young people engage with the use of technology in productive ways (Ananiadou and Claro, 2009; Sawyer, 2009). As educators we need to make sure young people are harnessing the potential or the positive side of technologies, while also still being mindful about its constraints and wary of the unethical uses of digital data (Pangrazio and Selwyn, 2019). Educators need to shape spaces for learning in ways that will successfully support students to feel connected, to feel that they can productively work with one another, to encourage their curiosity and experimentation, to foster their creativity and critical thinking. In sum, this is essentially an educational design problem, and more specifically, a “wicked problem.”

The Organization for Economic Co-operation and Development (OECD) started exploring innovative ways of encouraging young people’s learning through the Innovative Learning Environments (ILE) project (OECD, 2013, 2015a,b). Their main focus was on generating forward-thinking insights that would help devise productive guiding principles to support learning and innovation in the new millennium (Blackmore et al., 2012). The project emphasized learner-centered approaches to education, looking at how teachers might support inquiry and autonomous learning while acknowledging individual and group differences. Social aspects of learning, involving cooperation and group arrangements, were also highlighted. More recently, the OECD’s Learning Environments Evaluation Program (LEEP) (OECD, 2017, 2019) further developed ideas connected to learning environments, which were then described as “the result of interactions between physical resources (i.e., learning spaces, material and technology), learners, educators, content, learning leadership, society and policy” (OECD, 2017, p. 12). It is interesting to notice that in this definition a learning environment is seen as influenced by multiple elements, with a special reference to “learning spaces, material and technology”. In so doing, LEEP re-positions the significance of material and digital elements in learning, bringing these to the fore. Such a perspective is also in line with developments in learning theories and ecological approaches which see learning activity as part of an ecosystem of multiple elements (e.g., learners, educators, families, communities, knowledge, digital and material resources) (Barron, 2006; Damşa et al., 2019; Ellis and Goodyear, 2018; Vartiainen et al., 2018). Not surprisingly, the OECD policy developments have also been influencing multi-million dollar investments in (re)building new infra-structure, at schools and universities across the globe, and these policies also recognize the need to address learning in a digital world.

As such, learning spaces in schools and universities have been transitioning into new configurations, which often include open plan layouts, with bright flexible furnishings that are light and easily moveable, for example with desks that have wheels. In addition, these spaces are often set to accommodate multiple digital devices in use. The flexibility of the materials and furnishings allows teachers and students to quickly reconfigure the space according to the type of tasks students are asked to engage with. Activity in these newly designed spaces is rarely homogenous, as students might be constantly switching between periods of individual reflection, to a group work activity, they might be watching a one-to-many presentation, or the broadcasting of a video, or they might be using digital devices either individually or in groups, while completing a collaborative task and so on. These open spaces also

require that teachers work with others, exchanging ideas and developing collaborative skills (Carvalho, 2021; Carvalho et al., 2020). Not only are they preparing our students for learning how to collaborate and tackle the complex problems of our societies, but teachers also need to engage in collaborative problem solving themselves. And within a school or university setting, collaborative activity often involves people's ability to navigate mental, social, and spatial interactions. People need to learn how to negotiate shared meaning and so, to be successful in these types of situations, they need to understand how best to orchestrate tools and materials available to them, but also how to employ social strategies that might help sustain their (learning) activity until a resolution is reached.

As the complexity of learning situations around us increases, and with more demands for collaboration and participation, educators are having to rethink their pedagogical strategies, but also to (re)consider in depth how to design for learning in a digital era (Beetham and Sharpe, 2019).

The next section introduces an approach used to frame the analysis and design of complex learning environments.

Framing the analysis of complex learning environments

The Activity-Centered Analysis and Design (ACAD) framework (Goodyear and Carvalho, 2014; Goodyear et al., 2021) is an analytical tool that builds on ideas from architecture and design, to foreground the structural elements that come together in a learning network. The framework is concerned with educational design, and it acknowledges the need to consider the complexity of the diverse assemblages of tools, tasks and people, as one engages in orchestrating and navigating knowledgeable action in contemporary society.

ACAD allows educational researchers to break down the complexity of learning situations and, in so doing, to consider relationships between designable forms and emergent learning activity. In placing activity at the center (Fig. 3), ACAD expresses that what learners actually do is of utmost importance—their actions, thoughts and feelings. Learning outcomes are only indirectly reached, as learners have agency to co-create whatever is designed and proposed by a teacher. As such, the framework lays out three main “designable” components and a fourth element that is characterized as “emergent”. The three designable structures are:

- *set design*—includes digital and material tools or resources which are made available to learners at a learning situation (e.g., a pen and paper, a smartphone, a special configuration of furniture in a room);
- *social design*—this is about the types of social arrangements of learners, which can include scripted roles or divisions of labor (e.g., working in pairs or groups) and
- *epistemic design*—proposed tasks, ways of knowing (e.g., what are learners asked to do, pacing and sequencing of information).

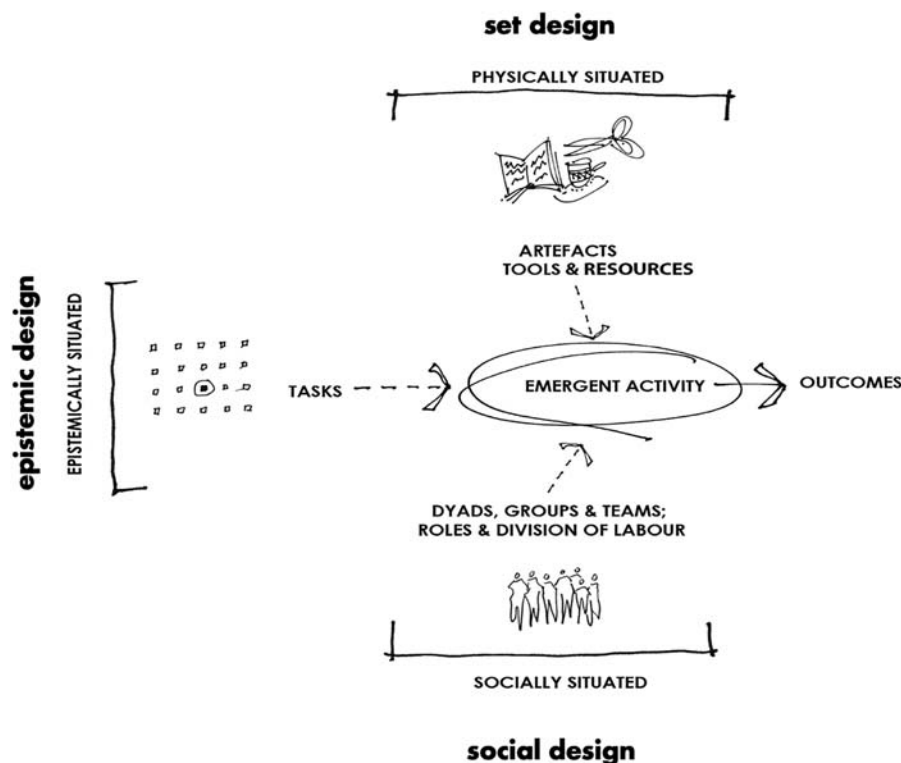


Fig. 3 ACAD framework. Adapted from Goodyear and Carvalho (2014, p. 59).

The fourth element in the framework is the *co-creation and co-configuration activity*, which is characterized as emergent and not designable. This element relates to learners' agency to re-configure and co-create what has been proposed by a teacher. Underlying the ACAD framework is a view of learning that considers learners' activity as epistemically, socially and physically situated (Goodyear et al., 2021). It is the set design dimension that most acutely brings the material and digital elements to the fore but, as the framework highlights, elements should be seen as part of an assemblage. One may consider what it means to use a certain material or to have access to flexible furniture, particularities of the layout of a space and how it accommodates the use of technologies, but it is also important to consider connections to social and epistemic elements. The ACAD framework emphasizes learning activity as an emergent phenomenon, which is influenced and shaped by various design elements, which can be noted at the micro, meso and macro levels (Carvalho and Yeoman, 2018; Goodyear et al., 2021).

Applying the ACAD lens to micro, meso and macro levels

In reflecting on the question “*how do we shape our spaces for learning?*” different perspectives and levels of granularity might be considered. For example, one might look at this very question while focusing on micro, meso and/or macro levels (Carvalho and Yeoman, 2018). To illustrate how the ACAD lenses may support the analysis of part-whole relationships at these different levels, the chapter now presents two scenarios. The first scenario explores the structural dimensions based on the earlier description of a Māori marae, in relation to designable components and the types of activities that unfold within the space of a meeting house. The second illustration foregrounds the application of the ACAD lenses in relation to innovative learning environments in schools and universities.

ACAD and a Māori marae

In the Māori marae example, at the *micro level* of *set design*, one could highlight elements such as the carvings in the walls and the meeting house open space itself. This is a space where the community gathers to share stories or responsibilities. The opportunities for young and old people being together can be seen as part of *micro social design*. The young generations watch and observe elders to learn about specific knowledge or cultural stories, or to learn how to share responsibilities, at the *micro epistemic design*.

At the *meso-epistemic design*, there are the traditional Māori pedagogies, such as *ako*, or principles such as positive interdependence. *Meso-social design* would include arrangements such as being from a particular *iwi* (or tribe) or *whanau* (family) and *meso set design* could signal the wider setting of the Māori marae. At the *macro-set design*, one could refer to the geographical location of Aotearoa New Zealand as a Pacific Island, as a bi-cultural nation (*macro-social design*), where long standing Māori values (*macro-epistemic design*) are all connected to the types of activities that unfold within these settings (Table 1).

ACAD and innovative or learning environments

Similarly, ACAD can be used to illuminate core design components in educational research or when undertaking educational design work to shape spaces for learning. Again, there are different perspectives and levels of granularity that might be considered (Table 2). At the macro level of educational design, one might focus on the development of government policies, or on specific learning theories and best practices (*macro-epistemic design*). It might involve considerations about the built environment and core infrastructure, including access to the Internet, bandwidth and WIFI. This is about whether the appropriate infrastructure exists or not, or what provisions are needed to be in place for our learners, so that some are not missing out on access to core resources (*macro-set design*).

At the meso level, one might consider particularities of an educational institution, looking at what might be the university or school's educational mission, how their strategies relate to learning, how curriculum is organized or how courses relate to one another (*meso-epistemic design*). It might involve noticing what types of learning spaces are available and how space is allocated and occupied by educators and students in a given institution (*meso-set design*). And at the micro level, we might consider the shape of spaces for learning while looking at more specific and local issues at the classroom level. We might notice the tools that are brought to our learners, or what students are working with, the digital and material resources that students might have access to, or how furniture is going to be (re)arranged in the space (*micro-set design*). We might look at specifics of learning tasks, their sequencing and pacing, or what students are asked to do (*micro-epistemic design*). Importantly, however, these scale levels are

Table 1 Māori marae mapped through the ACAD dimensions.

	<i>Set</i>	<i>Social</i>	<i>Epistemic</i>
Macro	Aotearoa New Zealand	A bi-cultural nation	Long standing Māori values
Meso	Marae	A <i>iwi</i> (or a tribe), or a <i>whanau</i> (family)	Traditional Māori pedagogies Ako (reciprocity in teaching and learning)
Micro	Meeting house, carvings in the walls	Elders, adults and children coming together	Opportunities to share stories, share responsibilities, watch, observe behaviors under elder's guidance

Table 2 Innovative learning environments and the ACAD dimensions.

	<i>Set</i>	<i>Social</i>	<i>Epistemic</i>
Macro	Wider infra-structure WiFi, access to Internet	Social and political structures	OECD policy Government policies Government curriculum Standards and best practices
Meso	Allocation of space in a school	Community orientation School culture	School vision and mission Discipline curriculum
Micro	Open plan classroom flexible spaces Breakout rooms Light furniture Digital devices	Students working in groups Specific scripted roles	A learning task involving creating a knowledge artifact (e.g., a video about the habits of an animal of interest)

also often interconnected. For example, policy (macro) is likely to influence an institution's mission and curricula (meso), and without the appropriate technological infrastructure (macro), it may be difficult to use a laptop in a lesson (micro) and so on.

An awareness of these connections is also important for educational research even when a researcher's interest might be on specific learning processes. For example, as a researcher explores students' engagement in a classroom activity, the focus may be at the micro level. The researcher is perhaps capturing students' interactions, or interactions between teachers and students, but emergent activity can also be seen in relation to other multiple structural components. It may include aspects related to the task itself and what students are being asked to do (*micro epistemic design*); how are students organized, whether they are working individually or in groups (*micro social design*); and how is the learning setting organized, what tools and resources are being used, what devices students are working on (*micro-set design*). All these elements are likely to influence what students actually do in the learning activity. And again, what happens at the micro level is influenced by meso and macro structures. For example, a classroom task that involves a group of students searching the Internet about a given topic, will only work if students have access to devices and to a wider infrastructure, such as WiFi or satellite reception. The particular task that a teacher designs for her students, takes into account a discipline curriculum (*meso epistemic design*), which in turn might be based on government policies, OECD standards and best practices in education (*macro epistemic design*). As an ecological approach, the ACAD framework may help educators break down the complexity of learning environments, to render visible the connections between multiple key designable elements in set, social and epistemic design, at different levels of granularity (micro, meso and macro), which are all important when considering learning activity in innovative learning environments.

Conclusion

In many parts of the world, conventional teaching models are evolving in accordance with new developments in the learning sciences, but also through a more marked presence of emerging technologies in classrooms, overall requiring that formal learning environments are upgraded to reflect 21st-century practices (Deppeler et al., 2021; Freeman et al., 2017). These open and flexible learning spaces have been on the rise in schools in Australia (Yeoman, 2016), Aotearoa New Zealand (Benade, 2017; Fletcher and Everatt, 2021), in the UK (Daniels et al., 2018; Woolner et al., 2018), in Finland (Niemi, 2020; Kokko and Hirsto, 2020), the Netherlands (Könings et al., 2017), Austria (Schabmann et al., 2016), and many other places. As such, researchers, educators, government and policymakers worldwide are trying to figure out what valuable elements should be included as part of contemporary design for learning, in order to best support processes of educational transformation.

Learning theorists have been calling for more holistic ways of seeing learning, or for considering learning activity as part of larger ecologies. Ecological perspectives are usually grounded on sociocultural and socio-material approaches to learning. They portray the mutually constitutive relations between individuals and larger educational and societal contexts, as also influencing learning processes. Hutchins (2010, 2014) offers a systemic perspective, where cognitive phenomena are not seen as individualized or isolated, but instead, as emerging through distributed activity. And according to Barron (2006) a learning context often involves "a unique configuration of activities, material resources, relationships, and the interactions that emerge from them" (p. 195). Vartiainen et al. (2018) refer to learning networks as co-evolving and connected to multiple contexts, social practices and tools. Essentially, while it is crucial to recognize the role of materials in learning, materials and spaces themselves cannot be seen in isolation. As Woolner (2010) reminds us, although a learning space might help or hinder particular ways of using materials, or invite some forms of social interactions, the space itself does not determine the teaching and learning that takes place there. A more holistic way of understanding the influence of tools, tasks and people on learning activity is needed, one that includes sensitivity to the materials (Ingold, 2013).

As the chapter illustrates, ecological perspectives are also reflected in international policy developments related to innovative learning environments, seen as involving learners, teachers, the community, pedagogical practices, and tools and resources (OECD, 2013; 2015a,b, 2017). Across the globe, as teachers and students transition into open and flexible learning spaces, educationalists are pondering what is needed when embracing new practices, for example as teachers learn how to find optimal ways of

working in teams, or how to prepare their students for more sustainable ways of living. Indeed, educators have reported experiences that involve having to learn to team teach and collaborate, having to adapt, and being in a period of transition, which is ongoing and evolving (Carvalho, 2021; Carvalho et al., 2020). However, as the chapter reveals, some of the values enacted within these innovative learning environments do not appear entirely new, and instead seem nicely related to the long-standing values of Māori traditions—namely the importance of nurturing positive interdependence, learning to work together with one another, learning how to collaborate, feeling connected and respected by others (Addis et al., 2011). One could perhaps argue that some of these educational practices are not new but, instead, education systems are somewhat reframing and reforming some of these long-standing values, trying to figure out how these can be reinstated within contemporary innovative learning environments, where material and digital are also to be considered as important aspects of the complex learning entanglement. Transition into new flexible configurations may require honoring valuable past experiences, as well as considering current contemporary needs, so that learning environments can support what we value in learning going forward, in evolving networked societies.

References

- Addis, P., Hall, M., Higgins, R., Higgins, T.R., 2011. Ask the posts of our house: using cultural spaces to encourage quality learning in higher education. *Teach. High. Educ.* 16 (5), 541–551. <https://doi.org/10.1080/13562517.2011.570440>.
- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries. OECD Education Working Papers, No. 41. OECD Publishing, Paris. <https://doi.org/10.1787/218525261154>.
- Austin, M., 2011. A description of the Māori marae. In: Rapoport, A. (Ed.), *The Mutual Interaction of People and Their Built Environment*. Mouton Publishers, Paris, pp. 229–242.
- Barron, B., 2006. Interest and self-sustained learning as catalysts of development: a learning ecology perspective. *Hum. Dev.* 49, 193–224. <https://doi.org/10.1159/000094368>.
- Beetham, H., Sharpe, R., 2019. *Rethinking Pedagogy for a Digital Age: Principles and Practices of Design*. Routledge, New York.
- Benade, L., 2017. The evolution of policy: a critical examination of school property under the National-led Government. *Waikato J. Educ.* 22 (1), 97–112.
- Blackmore, J., Bateman, D., Cloonan, A., Dixon, M., Loughlin, J., O'Mara, J., Senior, K., 2012. *Innovative Learning Environments Research Study*. Department of Education and Early Childhood, Victoria/OECD, Melbourne.
- Buchanan, R., 1992. Wicked problems in design thinking. *Des. Issues* 8 (2), 5–21. <https://www.jstor.org/stable/pdf/1511637.pdf>.
- Cardellino, P., Araneda, C., Alvarado, R.G., 2017. Classroom environments: an experiential analysis of the pupil-teacher visual interaction in Uruguay. *Learn. Environ. Res.* 20, 417–431.
- Carvalho, L., Nicholson, T., Yeoman, P., Thibaut, P., 2020. Space matters: framing the New Zealand learning landscape. *Learn. Environ. Res.* 23, 307–329. <https://doi.org/10.1007/s10984-020-09311-4>.
- Carvalho, L., Yeoman, P., 2018. Framing learning entanglement in innovative learning spaces: connecting theory, design, and practice. *Br. Educ. Res. J.* 44 (6), 1120–1137.
- Carvalho, L., Yeoman, P., 2021. Performativity of materials in learning: the learning-whole in action. *J. N. Approaches Educ. Res.* 10 (1), 1–15. <https://doi.org/10.7821/naer.2021.1.627>.
- Carvalho, L., 2021. (Re)Shaping spaces for learning: experiences from Aotearoa New Zealand. *N. Z. Annu. Rev. Educ.* 26, 52–59.
- Damsa, C., Nerland, M., Andreiadakis, Z., 2019. An ecological perspective on learner-constructed learning spaces. *Br. J. Educ. Technol.* <https://doi.org/10.1111/bjet.12855>.
- Daniels, H., Tse, H.M., Stables, A., Cox, S., 2018. Design as a social practice: the experience of newbuild schools. *Camb. J. Educ.* 49 (2), 215–233.
- Deppeler, J., Thompson, I., Corrigan, D., 2021. International studies of school design—lessons for Europe on innovation and risk. *Eur. Educ. Res. J.* 1–7. <https://doi.org/10.1177/14749041211014604>.
- Dohn, N. (Ed.), 2018. *Designing for Learning in a Networked World*. Routledge, Abingdon.
- Education Review Office, 2018. *Leading Innovative Learning in New Zealand Schools*. Education Review Office. <https://www.ero.govt.nz/publications/leading-innovative-learning-in-new-zealand-schools-april-2018/>.
- Ellis, R., Goodyear, P. (Eds.), 2018. *Spaces of Teaching and Learning: Integrating Perspectives on Research and Practice*. Springer, Singapore.
- Fenwick, T., Edwards, R., 2011. Considering materiality in educational policy: messy objects and multiple reals. *Educ. Theor.* 6 (6), 709–726.
- Fenwick, T., Edwards, R., 2014. Networks of knowledge, matters of learning, and criticality in higher education. *High Educ.* 64, 35–50. <https://doi.org/10.1007/s10734-013-9639-3>.
- Fenwick, T., 2015. Sociomateriality and learning: a critical approach. In: Scott, D., Hargreaves, E. (Eds.), *The Sage Handbook of Learning*. Sage Publications, London, pp. 83–93.
- Fletcher, J., Everatt, J., 2021. Innovative learning environments in New Zealand: student teachers' perceptions. *N. Z. J. Educ. Stud.* 1–21. <https://doi.org/10.1007/s40841-021-00195-3>.
- Freeman, A., Adams Becker, S., Cummins, M., Davis, A., Hall Giesinger, C., 2017. *NMC/CoSN Horizon Report: 2017 K–12 Edition*. The New Media Consortium, Austin, Texas.
- Goodyear, P., Carvalho, L., 2014. Framing the analysis of learning network architectures. In: Carvalho, L., Goodyear, P. (Eds.), *The Architecture of Productive Learning Networks*. Routledge, New York, pp. 48–70.
- Goodyear, P., Banks, S., Hodgson, V., McConnell, D., 2004. *Advances in Research in Networked Learning*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Goodyear, P., Carvalho, L., Yeoman, P., 2021. *Activity-Centred Analysis and Design (ACAD): Core Purposes, Distinctive Qualities and Current Developments*. Educational Technology Research & Development.
- Gourlay, L., Oliver, M., 2018. *Student Engagement in the Digital University: Sociomaterial Assemblages*. Routledge, New York.
- Hutchins, E., 2010. Cognitive ecology. *Topics Cogn. Sci.* 2, 705–715.
- Hutchins, E., 2014. The cultural ecosystem of human cognition. *Phil. Psychol.* 27 (1), 34–49. <https://doi.org/10.1080/09515089.2013.830548>.
- Ingold, T., 2013. *Making: Anthropology, Archaeology, Art and Architecture*. Routledge, Oxford.
- Ka'ai, T., 2008. The role of marae in tertiary education institutions. *Te Kaharoa* 1, 193–202.
- Kokko, A., Hirsto, L., 2020. From physical spaces to learning environments: processes in which physical spaces are transformed into learning environments. *Learn. Environ. Res.* 24, 71–85.
- Könings, K., Bovill, C., Woolner, P., 2017. Towards an interdisciplinary model of practice for participatory building design in education. *Eur. J. Educ.* 52 (3), 306–317.
- Macfarlane, A.H., 2015. Restlessness, resoluteness and reason: looking back at 50 years of Māori education. *N. Z. J. Educ. Stud.* 50, 177–193.
- Mulcahy, D., Cleveland, B., Aberton, H., 2015. Learning spaces and pedagogic change: envisioned, enacted and experienced. *Pedagog. Cult. Soc.* 3 (4), 575–595.
- Napan, K., Connor, H., Toki, L., 2020. Cultural pedagogy and transformative learning: reflections on teaching in a Māori environment in Aotearoa/New Zealand. *J. Transformative Educ.* 18 (1), 59–77. <https://doi.org/10.1177/1541344619858978>.
- Networked Learning Editorial Collective, 2020. *Networked learning: inviting a redefinition*. Postdigit. Sci. Educ. <https://doi.org/10.1007/s42438-020-00167-8>.
- New Zealand Ministry of Education, 2015. *Designing Schools in New Zealand: Requirements and Guidelines*. New Zealand Ministry of Education. <https://education.govt.nz/>.

- Niemi, K., 2020. "The best guess for the future?" Teachers' adaptation to open and flexible learning environments in Finland. *Educ. Inq.* <https://doi.org/10.1080/20004508.2020.1816371>.
- OECD, 2013. *Innovative Learning Environments*. OECD Publishing, Paris. <http://www.oecd.org/education/ceri/innovativelearningenvironmentspublication.htm>.
- OECD, 2015a. *Schooling Redesign: Towards Innovative Learning Systems*. OECD Publishing, Paris. <https://www.oecd.org/education/schooling-redesigned-9789264245914-en.htm>.
- OECD, 2015b. *Students, Computers and Learning*. OECD Publishing, Paris.
- OECD, 2017. *Learning Environments Evaluation Program*. https://www.oecd.org/education/LEEP_Brochure_OECD.pdf.
- OECD, 2019. *Effective Learning Environments (ELE)*. <https://www.oecd.org/education/effective-learning-environments/>.
- Pangrazio, L., Selwyn, N., 2019. Personal data literacies: a critical literacies approach to enhancing understandings of personal digital data. *New Media Soc.* 21 (2), 419–437. <https://doi.org/10.1177/1461444818799523>.
- Sawyer, K., 2009. *The Future of Learning in the Age of Innovation: Review for the Beyond Current Horizons Programme*. Futurelab, Bristol. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.504.541&rep=rep1&type=pdf>.
- Schabmann, A., Popper, V., Schmidt, B., et al., 2016. The relevance of innovative school architecture for school principals. *Sch. Leader. Manag.* 36 (2), 184–203.
- Sørensen, E., 2009. *The Materiality of Learning: Technology and Knowledge in Educational Practice*. Cambridge University Press, Cambridge.
- Statistics New Zealand, 2014. *Taku marae e: Connecting to ancestral marae 2013*. <https://www.stats.govt.nz/>.
- Statistics New Zealand, 2020. 2018 Census. <https://www.stats.govt.nz/2018-census/>.
- Te Ara, 2020. *Story: Ngā Mahi Tika*. <https://teara.govt.nz/en/nga-mahi-tika>.
- Te Kete Ipurangi, 2020. *Innovative Learning Environments*. <https://elearning.tki.org.nz/Teaching/Innovative-learning-environments>.
- The Guardian, 2019. Greta Thunberg to World Leaders: "How Dare You—You Have Stolen My Dreams and My Childhood"—Video. <https://www.theguardian.com/environment/video/2019/sep/23/greta-thunberg-to-world-leaders-how-dare-you-you-have-stolen-my-dreams-and-my-childhood-video>.
- Vartiainen, H., Nissinen, S., Pöllänen, S., Vanninen, P., 2018. Teachers' insights into connected learning networks: emerging activities and forms of participation. *AERA Open*. <https://doi.org/10.1177/2332858418799694>.
- Wood, A., 2019. Built policy: school-building and architecture as policy instrument. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2019.1578901>.
- Woolner, P., Thomas, U., Tiplady, L., 2018. Structural change from physical foundations: the role of the environment in enacting school change. *J. Educ. Change* 19, 223–242. <https://doi.org/10.1007/s10833-018-9317-4>.
- Woolner, P., 2010. *The Design of Learning Spaces*. Continuum, London.
- Yeoman, P., 2016. A study of correspondence, dissonance, and improvisation in the design and use of a school-based networked learning environment. In: Carvalho, L., Goodyear, P., De Laat, M. (Eds.), *Place-Based Spaces for Networked Learning*. Routledge, New York, pp. 41–58.

Gender differences in educational achievement and learning outcomes

David Reilly, School of Applied Psychology, Griffith University, O'Connor, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Gender differences in intelligence	400
Gender differences in specific cognitive abilities	400
Interpreting the magnitude of gender gaps	401
Gender differences in verbal and language abilities	402
Reading	402
Reading in an international context	403
Writing	403
Spelling	403
Grammar and punctuation usage	403
Summary	403
Gender differences in mathematics and science	404
Mathematics and science achievement in the United States	404
Mathematics and science achievement in an international context	404
Psychological factors associated with gender differences in mathematics and science	405
Summary	406
Educational implications of gender gaps in educational outcomes	406
Conclusions	406
Acknowledgment	407
References	407
Further reading	408

One of the most vexing but also most important research questions in the intersection of education and psychology is whether there are gender differences in intelligence and learning. For the answers to this question has important, nay, even profound implications for how we educate boys and girls. Do boys and girls have different purported “learning styles,” and are we doing either gender a disservice by presenting the material the same way, in the same classes, with co-educational schooling? Some readers will likely have firmly entrenched views on the subject; others may be undecided but curious to know more. Often the subjective impressions we draw from our own teaching experiences can differ from those of other educators, and the field of education is a broad profession that accepts many diverse perspectives on teaching. Endorsement of support for single-sex schooling is also guided by our political and ideological worldview: both educators and parents are more likely to believe single-sex schooling is beneficial if they hold less liberal, more conservative outlooks. Indeed, writes Mills (2004), single-sex schooling is explicitly marketed to that demographic and often accompanied by an appeal to traditional gender norms. Regardless though of which camp you may fall under or, if indeed, like many educators you remain undecided, the founding tenet on which single-sex schooling rests is that boys and girls think, behave, and learn differently to one other.

Such claims should be put to the test, and there is a large body of scientific and educational literature that does just that. There are many myths and misconceptions amongst educators and parents about the nature of gender differences in intelligence; these gender stereotypes can serve to undermine student performance and magnify gender gaps in learning outcomes. Beliefs that there is a “smarter sex” are often propagated in educational literature or that males and females are radically different (aka the “men are from mars, women from venus” perspective), as well as in popular culture. However, neuroscientific evidence is unequivocal that males and females are equal in general intelligence, and that there are also gender-specific differences for more applied cognitive abilities. The purpose of this chapter is to examine how gender affects learning outcomes. Firstly, it examines gender differences in verbal and language abilities, where females typically outperform males. Secondly, it examines gender differences in quantitative reasoning (mathematics and science). In this domain the picture is less clear, as students in Western countries perform quite differently to those in other cultures.

Rather than cherry-picking certain studies to present an argument in support or against the existence of gender differences, this review examines evidence from large-sized nationally representative samples to determine whether there are different patterns of educational achievement between males and females. This represents the gold standard of evidence for examining gender differences (Becker and Hedges, 1988), as most studies recruit highly-selective and non-representative samples to estimate the extent (if any) of gender differences. Selection bias is a serious problem in educational and psychological research, as we know that there are important demographic factors that also impact educational outcomes and interact with gender. Older reviews on this subject have presented highly selective evidence which may lead to erroneous conclusions about gender and mathematics (for a critique see Spelke, 2005), or verbal ability (Wallentin, 2009). A further limitation of much of the literature is that it is highly US-centric, and ignores gender and learning in an international context. As the reader will see, there is considerable variation cross-culturally. This

suggests that sociocultural factors and educational systems contribute to the presence or absence of gender differences. Gender differences in some learning outcomes are quite large and impactful, while in other areas are relatively small. Educational implications are then discussed.

Gender differences in intelligence

Psychology is the discipline that studies human thought and cognition, and one of its aims is the accurate measurement of human intelligence. Historically, the fields of education and psychology have always been intertwined. Indeed, the very first intelligence tests used in psychology were developed by Alfred Binet under the commission of the French Ministry of Education, to devise objective procedures for measuring student aptitude and classroom placements. From this, sprang forth the intelligence and educational assessment movement; prior to this, decisions made about intellectual aptitude and educational abilities of an individual relied, instead, on subjective impressions, racial/gender stereotypes, and often socioeconomic class. More than a century later, after the measurement of millions of children and adults, we have a firm bedrock of data to answer the question of whether there are actually gender differences in intelligence, and if so, how large is their magnitude. The answers are more nuanced than is commonly assumed, and depends on both how and where one looks.

From the outset, let the reader be clear in their mind that there are no gender differences in general intelligence (IQ). When representative samples of men and women are compared at the population level, their psychometrically assessed general intelligence is equivalent (Jensen, 1998; Neisser et al., 1996). There is near-unanimous consensus in psychology and neuroscience that men are just as smart as women (or conversely, women are just as smart as men!). Sternberg (2014, p. 178) concluded that “there is no evidence, overall, of sex differences in levels of intelligence.” However, a large number of studies have documented that the performance of males shows greater statistical variance than females (i.e., there are a greater number of both high-scoring and low-scoring males at the extreme tails of the ability distribution), termed the great male variability (GMV) hypothesis. For the non-mathematician, the practical effect of this is that even though there are no gender differences in overall IQ, recruiting from highly select samples that are high in ability (such as groups of gifted and talented) may result in a greater number of high-achieving males, while selecting from a lower ability sample may see boys and men overrepresented with learning difficulties. This can lead to a distorted impression sometimes among educators working with such populations: either that boys are “smarter” than girls if they work with gifted and talented, or that boys are “less capable” by educators dealing primarily with remedial students.

Subjective impressions of gender differences in intelligence are also affected by the books and literature that parents, educators, and policy-makers read. There is no shortage of authors arguing there are large biological differences between males and female brains (Elliot, 2009), who talk about the effect of “raging hormones” on learning, or rely on fallacies around gender differences in brain size and intelligence (Lynn, 1994). These neuromythologies largely ignore peer-reviewed scientific evidence, and reflect, to a large degree, implicit beliefs about human intelligence and gender (Halpern, 2010). When asked, most adults will agree that men and women are equal in intelligence but when researchers ask large representative samples to estimate their children’s intelligence, both mothers and fathers rate the intelligence of sons as being higher than daughters (Furnham et al., 2002). Despite there being no gender difference in objectively measured IQ, when asked to provide an estimate of their own intelligence, men on average rate their intelligence as higher than do women (Rammstedt and Rammesayer, 2002). This finding is robust and has been reproduced cross-culturally (Szymanowicz and Furnham, 2011), and has been termed the male-hubris/female humility effect (Furnham et al., 2001). So it seems we have an innate predisposition to believe quasi-scientific arguments that since men’s heads and brains are slightly larger than females, males might just possibly be brighter and more capable than females,¹ particularly if they suit our ideological worldview or implicit beliefs about gender (Nosek et al., 2002). This is why it is important to consider evidence on gender and intelligence carefully: as Carl Sagan so elegantly quipped, extraordinary claims require extraordinary evidence. We should hold literature on gender and learning to the same standard.

Gender differences in specific cognitive abilities

While it is true that, overall, men and women are equally intelligent, we do also see gender-specific patterns for certain types of cognitive tasks (i.e., that some types of tasks show a female advantage and others a pattern of higher male performance). For example, it may come as no great surprise to educators that there is a female advantage in certain verbal and language tasks—the most common of which encountered in educational settings is reading skills (Logan and Johnston, 2009). Boys and men, as a group, tend to show greater performance for visual-spatial tasks such as mental rotation or navigation (Voyer et al., 1995). Visual-spatial reasoning also lays down a foundation for quantitative reasoning (mathematical and scientific aptitude). Males do show slightly higher mean scores on standardized tests of mathematics and science achievement (Reilly et al., 2015), but the extent of this is not nearly as large as the gender gap in reading. This may contribute to the underrepresentation of women in certain fields of science, technology, engineering and mathematics (STEM). But we should also be mindful that when psychologists make such

¹This argument plays on the neuroscientific fallacy that there is a linear correlation between brain volume and intelligence. As I jokingly remind my students when teaching about intelligence, following this line of reasoning the elephant must have a vastly superior intellect to us all!

statements, we are talking about group differences at the population level: some males score extremely high on reading and language tasks just as some females score low on them. A child's gender, in and of itself, should never be used as a proxy for measuring ability when making decisions about educational placements, just as an applicant's gender should not determine their suitability for job-hiring.

The three domains outlined above (verbal and language abilities, visual-spatial ability, and quantitative reasoning) are the most clearly identified and actively researched areas where we may see gender differences in cognitive abilities. There are other areas such as moral reasoning (You et al., 2011), nonverbal perception (Hall, 2005), and emotional intelligence (Petrides and Furnham, 2000) which are more contentious, but as they also involve non-cognitive processes (e.g., social and emotional) they are somewhat impure and not strictly "cognitive" abilities. Furthermore, on many—perhaps even most—cognitive tasks, males and females show similar levels of performance. So, it is important when reading the literature to keep this in perspective: there are gender differences for certain types of cognitive tasks that may impact learning, but for most types of reasoning males and females as a group show equal performance.

Out of the three domains, verbal and language abilities have the most impact on educational achievement, however, the gender gap in visual-spatial ability contributes to later gender gaps in mathematics and science. Therefore, one must be careful not to make too much of gender differences in cognition ("males and females are worlds apart"), but neither should one make too little of its educational impact either—particularly in the case of verbal and language abilities which scaffold educational achievement in secondary and tertiary education. In some areas, there are important effects of gender on learning that lead to later inequality of outcomes (Riegle-Crumb et al., 2018; Buchmann et al., 2008). For example, across most Western nations, more women than men enroll in college/university education after leaving high school, and when they do, women have quite a higher completion rate (OECD, 2020). A growing number of educators have expressed concern that gender gaps in literacy are leaving a generation of young men unprepared for future learning (Buchmann et al., 2008). We also live in a time of technological disruption, when traditionally male-dominated professions (e.g., manufacturing, assembly, transport) face the risk of automation and elimination of the need for unskilled or lowly-skilled labor (Muro et al., 2019). The ability to read and learn new skills will be vital to prepare students for the future. We also see gender inequality in the other direction, with a strong underrepresentation of women in many fields of science, technology, engineering and mathematics (STEM). While progress has been made in some fields of science (particularly biology and life sciences), women remain greatly underrepresented in most others (National Science Foundation, 2019). High-tech professions have also been more resilient against disruption during the COVID-19 pandemic, and STEM careers offer greater economic security—boosting female representation could make much-needed progress in addressing the gender-wage gap. Even small differences in equality of educational outcomes at the population level translate to large economic ones, making it important for all educators to understand how gender may affect learning.

Interpreting the magnitude of gender gaps

A difficulty most non-psychologists and statisticians encounter when reading literature in this area is how to evaluate the magnitude of gender gaps (i.e., just how large is the difference between the typical male and female student, and whether we are making headway in addressing these gaps). Most studies in psychology and the social sciences describe evidence in terms of "statistical significance." Statistical significance is a way of evaluating the likelihood that a difference observed between two groups (such as boys and girls) is genuine and not illusory. It indicates whether or not if we were to repeat the assessment with a different group of randomly selected students, we would likely see the same results. We use the probability value, also known as the p -value, or sometimes even just referred to as p (italicized). It is a statistical safeguard that the evidence is solid—but the term statistically significant often gets truncated to "significant," which has a very different meaning in everyday language. It also privileges studies with larger sample sizes: gather a sufficiently large enough sample and even a very small difference in means may be statistically significant, even though the magnitude of the gender gap may be trivial in size. As I often remind my statistical students, it is not the size of your p that is impressive! Rather, you should be evaluating the magnitude of the difference between two groups, using a metric like an effect size. This practice allows the reader to evaluate whether a gender gap is *practically* significant instead of merely statistically significant.

Effect sizes provide a quantitative (i.e., numerical) index of relationships between variables like gender. These provide an objective way of describing the magnitude of an observed effect in a way that non-statisticians can evaluate without formal training (for a gentle introduction to the topic, see Hedges, 2008). We use them to communicate the strength of an observed difference in a format that will be accessible to anyone: from educators to parents to policy-makers. The difference between two groups, such as boys and girls, can be expressed as Cohen's d . Understanding the statistical formula behind it is not necessary, but for the genuinely curious, it is the difference between two means divided by their pooled standard deviation which expresses it in standardized units. By convention, a positive effect size ($d > 0$) indicates higher male performance while a negative effect size ($d < 0$) indicates higher female performance.

Cohen (1988) offered some initial guidelines for interpreting effect sizes, applying the following labels to guide interpretation: $d = 0.20$ would be regarded as a relatively small difference between groups, while around $d = 0.50$ or so could be regarded as medium and $d = 0.80$ as large. Most effects studied in the social sciences and education are "small," with some medium and very few being large. These labels were intended as a starting point for a non-statistician to use, but Cohen also advised interpreting effect sizes relative to other (known) effects. For example, when a study reports a gender gap in reading as being "small," by Cohen's (1988) guidelines, we might also compare it to the gender gap in mathematics so as to allow for a comparative interpretation. By

way of example, two separate studies examined gender differences in large-scale national assessments of American school students in reading and mathematics. For reading, Reilly et al. (2019) report an effect size of $d = -0.32$ for students in the 12th grade which would be regarded as “small to medium” under Cohen’s guidelines, but is actually three times the size of the gender gap found in mathematics, $d = 0.10$ (Reilly et al., 2015). It is often helpful then to relate the size of gender differences in one area to another. So while “small” and “medium” might be useful yardsticks to go by, the reader is reminded that these were small only in a statistical sense, and that even small mean differences have profound educational and occupational implications for those students affected. Such differences also affect a student’s attitude towards reading, mathematics and science, as well as self-efficacy beliefs; gender gaps in these domains are typically far larger than actual abilities.

Gender differences in verbal and language abilities

Verbal and language abilities is a term that encompasses any task involving words and language, covering both *language reception* (listening and reading) as well as *language production* (orally or in a written form). In what was for its time the first truly comprehensive review of the psychological and educational literature on gender differences, Maccoby and Jacklin (1974, p. 351) concluded that “girls score higher on tasks involving both receptive and productive language, and on ‘high-level’ verbal tasks (analogies, comprehension of difficult written material, creative writing) as well as upon the ‘lower-level’ measures (fluency).” This view is supported by more contemporary researchers, with a female advantage seen in verbal fluency, spelling, grammar, reading and writing, and oral comprehension. Indeed, the only area of language where we do not see a pronounced female advantage is in word vocabulary: for most students, vocabulary is linked to the amount and variety of reading a student engages in as well as their general intellectual ability. But on all other tasks, there appears to be a pronounced female advantage though the magnitude of this gender gap varies across the type of task.

Reading

In an educational setting, the most prominently tested language skill will be reading ability. There is an emphasis on phonetic skill and whole-word reading at earlier grades, but as a child progresses through successive years of schooling, emphasis is placed less on the mechanics of reading and more on the speed of reading and quality of reading comprehension. Thus, the skills being assessed in early years are more remedial, while educational assessments in high school focus not just on basic reading skills but also reading comprehension and understanding. Large-scale educational assessments of reading achievement typically measure both students’ overall reading ability and whether they reach a minimum literacy benchmark. While boys are already greatly overrepresented in the number of children formally diagnosed with reading impairment or dyslexia (Berninger et al., 2008), there is a greater number still that fall at or somewhat above the clinical threshold for a formal diagnosis (Shaywitz et al., 1992). While educators will already be familiar with a higher number of boys than girls being classified as dyslexic, what sometimes comes as a surprise is that there are also mean gender differences in reading (i.e., the typical boy scores much lower than the typical girl).

Poor readers find themselves at a profound educational disadvantage compared to their peers. Not only may they struggle with reading tasks, but there are flow-on effects to other domains of learning given the amount of independent reading expected of students in high school. Reading ability is also correlated with reading interest and motivation. When reading flows easily for a student, it becomes an easy (or at least achievable) task, while it may be effortful and fatiguing for a poorer reader. We also know that the amount of time devoted to reading is also associated with reading improvements. Students with a higher starting point tend to receive greater practice over time, while poor readers tend to read much less than their peers (Harmon et al., 2016). For this very reason, early reading intervention is widely acknowledged by educators as critical (Foorman et al., 1997), and is typically offered to those identified as very poor readers. The long-term outcomes for students receiving early intervention are some degree of improvement (Hurry and Sylva, 2007), though they may still qualify for a diagnosis of dyslexia later in life. However, a considerable number of students still struggle with reading but at a sub-clinical level that prevents them from qualifying for diagnosis and intervention. This, in turn, holds back language development and acquisition of new learning in other domains; children who are functionally literate but who struggle to read and comprehend a textbook are then ill-equipped for high school and college. The distribution of sub-clinical poor readers appears to fall along gendered lines, but concerted efforts have been made in recent years to try to address this with often innovative suggestions. An important research question remains whether these efforts are actually productive, and whether educators are making headway over time.

Such answers may be determined by examining data from large, nationally representative samples of educational achievement that track student performance over a lengthy timeframe. For students in the United States, the National Assessment of Educational Progress (NAEP) has evaluated student progress in reading for over three decades using a national reading framework. It makes for one of the largest available datasets tracking reading over time, but until recently had seen little secondary analysis beyond annual reports noting statistical significance. Reilly et al. (2019) examined data from over 3.9 million students from the period 1988 to 2015. Due to the large sample size, all assessment waves were statistically significant, but the study also evaluated the magnitude of the gender gap averaged across all assessments. There were relatively small gender differences observed for students in Grade 4 ($d = -0.19$), but these grow progressively larger with additional years of schooling. By eighth grade, the gap between boys and girls has grown to small-to-medium in magnitude ($d = -0.30$), and remains so as students exit high school at Grade 12 ($d = -0.32$). As mentioned earlier, this is over three times the gender gap in mathematics for children in the United States. Furthermore, there was no substantive change across assessment waves indicating that no progress had been made closing the gap over time.

Reading in an international context

When we look to data from other countries, we see similar patterns to that found in the United States. There are two large-scale international assessments of reading achievement: firstly, the Progress in International Reading Literacy (PIRLS) measuring students in Grades 4 and 8, and secondly, the Program for International Student Assessment (PISA) measuring students near the end of compulsory schooling at age 15. In every country surveyed, the mean score for males was considerably lower than that of their female peers (i.e., the effect was universal); in no country did boys reach the standard of girls in reading (Reilly, 2021; Guiso et al., 2008). There was also a similar developmental trend to that reported in US testing data. The gender gap may start as small ($d = -0.22$ in Grade 4), but widens considerably with successive years of schooling ($d = -0.44$ by age 15). These gaps are not being driven just by a higher prevalence of dyslexia, but rather reflect poorer overall reading comprehension in boys (occurring in the middle of the ability distribution). Furthermore, despite concerted efforts made by educators to address this issue, there is no evidence that these were effective in reducing the disparity between the typical male and female student.

If, on average, boys perform considerably lower in reading in all countries, and if there appears to be no reduction in the gender gap over time, then there may be the temptation for educators and policy-makers to consider this to be an intractable and unsolvable problem. There may even be some that use this to advocate for single-sex schooling or for reducing funding to educational programs aimed at addressing the gender gap. That would be a flawed—and dangerous—line of reasoning, though. When we look at the data cross-culturally, there is actually considerable variation across countries in the magnitude of the gap (Reilly, 2012; Guiso et al., 2008). This variation highlights the contribution of structural factors that differ between countries, such as their educational practices and provision of early education programs that focus on literacy. Some countries do this better than others. Additionally, there are likely sociocultural factors contributing to this, such as parental involvement in reading to children and cultural norms around reading. Further, research is needed though to develop new educational interventions that can raise the performance of sub-clinical poor readers. A rising tide lifts all boats, and may in time reduce the disparity between male and female reading.

Writing

While the divide between boys and girls in reading has taken on high prominence, gender gaps are also present for other verbal and language abilities. Though less visible and less reported, the size of the gender gap in writing ability is actually more substantial. Part of the reason this has been previously overlooked is that there are relatively few large-scale writing educational assessments that recruit representative samples. For example, the international PIRLS and PISA assessments measure reading literacy quite well but do not seek to measure writing ability (or indeed any other verbal component, such as spelling, grammar, or language conventions). Within the United States, nationally representative samples of writing achievement have been collected by the NAEP. A recent analysis by Reilly et al. (2019) of several decades of NAEP performance found that the gender gap in writing achievement was substantial and followed a developmental trend towards a larger divide with successive years of schooling. The gender gap in writing is quite substantial when first measured in Grade 4 ($d = -0.42$) and increases for older students in Grade 12 ($d = -0.55$). By comparison, the gender gap in writing ability for students at the end of their compulsory schooling is about five times the size of the mathematics gap—yet substantially more time and resources are invested in reducing the gender gap in mathematics. Similar results are found in nationally representative samples for Australia (Thomas, 2020) and New Zealand (Parr, 2010).

Spelling

Another area where we see quite substantive gender gaps is in spelling, with girls on average scoring significantly higher than boys. Feingold (1988) examined archives of educational testing data from between 1940 and 1980, finding significant gender differences and a similar developmental trend to that observed in reading achievement ($d = -0.50$ in Grade 12). This may reflect a common underlying deficit in phonological processing which also affects reading ability. Stanley et al. (1992) reported similarly sized gender differences in large representative samples on the Differential Aptitude Test (DAT) Spelling subtest for US students, while Lynn (1992) replicated this for British students with the UK English version. Other standardized spelling tests demonstrate similar effects: for example, Reynolds et al. (2015) found similar sized gender differences on the Kaufmann Test of Educational Achievement KTEA-II Spelling subtest ($d = -0.45$).

Grammar and punctuation usage

There are also considerably-sized gender differences in understanding and usage of grammar and punctuation. Going back to earlier decades, Feingold (1988) reported girls scored significantly higher in the understanding of grammar from assessments conducted in the 1940s until testing in the 1980s (average $d = -0.43$). Stanley et al. (1992) found this as well with nationally representative samples on the Differential Aptitude Test (Grade 12 $d = -0.50$). All of these studies investigated grammar in US students, as there are no large-scale international assessments (such as PIRLS or PISA) that measure grammatical usage in a student's native language. However, it is quite likely that this effect is found globally, mirroring the pattern for reading.

Summary

In summary, there are modest but highly impactful gender differences in reading found globally, and relatively little progress has been made at narrowing the gender gap. However, these are dwarfed by gender differences in writing, as well as the components of writing (spelling and grammar/punctuation). While modest in the early years of elementary school, the gap widens considerably

with successive schooling years. This places poor writers at a substantial disadvantage in high school, particularly for courses that require students to write essays or lab reports (Zhang et al., 2019). It may also jeopardize both their prospects of entry into college/university and, once enrolled, their subsequent educational achievement given the heavy use of written assignments and short answer questions in exams. These issues have received less attention than reading and mathematics, in part because of the difficulty of measuring these other outcomes as few large-scale assessments include them. This also means many educators are unaware of the true extent of gender differences in these areas: knowing is a necessary prerequisite for tackling them in the classroom.

With an ever-crowded curriculum, it can be tempting to prioritize just reading and mathematics—particularly if other skills like writing, spelling, and grammar are not measured in state-based and national educational assessments. We know though that these skills improve with practice, and even more so when students are given guided feedback on their writing. Early writing instruction and feedback in primary/elementary school shows substantial improvement (Graham et al., 2012), but there are educationally meaningful improvements in high-school-aged students as well (Graham and Perin, 2007). The use of word processing software may highlight rudimentary spelling and grammar errors, but are not a replacement for the mechanics of grammar. Students also need guidance on high-level writing tasks like paragraphing, linking sentences, establishing an argument and writing persuasively. However, technological based writing aids can be used to supplement good writing pedagogy in the classroom, and can provide additional assistance to students with their writing outside of classroom time (Little et al., 2018). While there are many males who demonstrate great proficiency in writing ability (and some females that do not), large-scale assessments of educational achievement show that many males lag behind the typical female student. Redressing this inequality in achievement must surely be an important educational priority.

Gender differences in mathematics and science

While verbal and language abilities show a considerable female advantage, educators have long been aware that there is a male advantage in quantitative reasoning (a collective term for mathematics and science). Again, the reader is reminded that when we talk about group differences between males and females, this refers to differences on average rather than individuals. Many would have taught male students who struggled with mathematical and scientific content, and female students who excelled in these domains. Though “statistically significant” gender differences were frequently observed in educational assessments, there was little guidance for educators and policy-makers on just how *large* a difference this was, and whether this had a practical impact educationally. Additionally, it was unclear whether progress was being made over time at reducing the gender gap, as had been claimed by some researchers (e.g. Feingold, 1988).

Mathematics and science achievement in the United States

Hedges and Nowell (1995) were the first to aggregate results across multiple waves of nationally representative educational assessments to try to estimate the magnitude of gender differences in mathematics and science. Examining data NAEP data from 1978 to 1992, they observed that the gender gap in mathematics was relatively small (d 's ranging from 0.11 to 0.19), but that the gap in general science was more substantial ($d = 0.22$ to 0.33). The authors wrote that gender differences in mathematics and science “are of concern, because ability and achievement in science and mathematics may be necessary to excel in scientific and technical occupations” (p. 45). Furthermore, the authors found no evidence for any change in the extent of the gender gap over time, despite significant educational reform and social changes in gender attitudes. Numerous successive waves of PISA have been conducted since, and a recent study by Reilly, Neumann and Andrews (Reilly et al., 2015) found that there are still small gender differences in mathematics ($d = 0.10$) and science ($d = 0.13$) in modern samples. Additionally, the gender ratio of students achieving the highest level of proficiency in mathematics and science favors males, by a ratio of 2 males to 1 female. This is a far more problematic outcome, because it is from the pool of high-achievers that we would be recruiting students for STEM-degrees in college and university, and ultimately into the STEM-pipeline for the workforce. Thus, while mean gender differences are still observed and an imbalance in the gender ratio of high achievers, neither should the impression be taken that the gender gap is vast either or that it should be used as an argument in debates about single-sex schooling.

Mathematics and science achievement in an international context

This original high profile study by Hedges and Nowell (1995) helped galvanize educators, policy-makers, and researchers to make concerted efforts to address gender inequalities in mathematics and science outcomes, both in the United States and abroad. This could be evaluated periodically using large-scale international assessments of student ability such as the Trends in Mathematics and Science Study (TIMSS) measuring in Grades 4 and 8, and the Program for International Student Assessment (PISA) measuring students at age 15 near the end of their compulsory schooling. These assessments measure student achievement in mathematics and science, but also psychological and social factors such as student attitudes as well as student self-efficacy beliefs about their mathematical and scientific ability.

Results from TIMSS show that there are relatively small gender differences in mathematics and science achievement favoring males, but that these are not universal. In many countries boys and girls achieve parity, and in many countries girls greatly outperform their same-aged male peers (Reilly et al., 2017). Earlier waves of TIMSS assessments tended to find a pattern of higher male performance overall but reflected information from predominantly Western countries where gender-science stereotypes are greater (Nosek et al., 2009). This completely negates any claims that males always outperform females in mathematical and scientific achievement, that they reflect “innate” biological differences between boys and girls, or that “males are just better suited for

STEM.” Rather, it shows that sociocultural factors (such as endorsement of gender-STEM stereotypes), as well as the educational environment play a pivotal role in shaping mathematics and science (Reilly and Hurem, *in press*). As a society, how we raise and teach boys and girls shapes their worldview, and how they see themselves in relation to science.

Even at a young age, in TIMSS (Grade 8) we can start to see quite pronounced gender differences in mathematics and science attitudes (males report more positive attitudes to these subjects) and self-efficacy beliefs (males rate their mathematical and scientific ability higher than females). These presage a widening of the gender gap in mathematics and science achievement for older students, and may reflect the pernicious nature of gender-science stereotypes and self-fulfilling prophecies (Miller et al., 2018; Reilly and Hurem, *in press*). Girls who believe that STEM is more for boys begin to doubt their abilities, lose interest and motivation, and start to disengage from these courses. In later years of high school students generally have voluntary course selection of either basic or more advanced mathematics and science coursework. As noted in Eccles’ (2013) expectancy-value model of achievement motivation, students tend to select courses that they either like or that they expect they will do well in. Thus attitudinal factors like self-efficacy serve as a bottleneck limiting the number of girls pursuing STEM in the final years of high school.

We can see the direct effect of this with a widening of the gender gap in PISA results for students examined at age 15. In most waves of PISA testing, boys demonstrate significantly greater performance in mathematics (Guiso et al., 2008; Else-Quest et al., 2010) and science achievement (Reilly, 2012), especially in Western countries. Importantly though, the effect is not universal across all countries—many non-OECD countries showed significantly higher female performance in mathematics and science while the United States held one of the largest male-female differences (Reilly, 2012). As more non-Western countries participate in PISA assessments, though, we have seen global mathematics and science shift away from a pattern of superior male performance and closer to parity (OECD, 2019).

Psychological factors associated with gender differences in mathematics and science

So, while relatively little progress has been made at reducing the gender gap in STEM achievement over time in countries like the United States, United Kingdom, Canada and Australia, we now have a much clearer picture of why these gaps exist. Rather than reflecting innate gender differences in aptitude as had been claimed by some researchers (e.g. Geary, 1996), we see greater interest and, more importantly, self-confidence in males about STEM compared to females. Particularly, in countries where there is higher endorsement of gender-STEM stereotypes (Nosek et al., 2002), these limit the motivation and self-confidence of girls and women. Educational psychologists have developed brief and cost-effective interventions that can address these self-efficacy beliefs (for a review see Reilly and Hurem, *in press*), and these translate into educationally meaningful improvements in grades for girls and women (Miller and Halpern, 2013; Miyake et al., 2010; Mix et al., 2020). While the results these interventions deliver are impressive in just how large a difference they can make to female student grades, these studies are typically published in psychological rather than educational journals. Currently, educational psychologists need to make greater efforts in communicating these findings to educators and policy-makers, so as to increase their usage in the field.

Before concluding this critique of gender differences in mathematics and science achievement, there is one other issue that educators (and perhaps students) should be mindful of. The majority of evidence evaluating gender differences in STEM looks at performance on standardized tests. When we observe that males as a group have higher test scores, we draw the inference that this reflects actual differences in ability for quantitative reasoning. The underlying assumption is that test content is representative of the types of mathematical and scientific reasoning tasks students would encounter, and that it is, to the greatest extent possible, free of systematic error. But what is less frequently talked about by educational researchers is the *test and grade discrepancy* in mathematics and science (Reilly, 2019). While males generally tend to perform better than females on standardized tests, we often see a different pattern for student grades. Recall that there were small gender differences in mathematics and science achievement for US students (as measured by the NAEP). In an analysis of high school transcript data from 1990 to 2005, Shettle et al. (2007) found that girls earned significantly higher grades than boys in mathematics and all fields of science. A more recent meta-analysis by Voyer and Voyer (2014) included a broader complement of samples, including elementary/middle-school students, high school and college students. Overall across ages, females earned slightly higher grades in mathematics ($d = -0.07$) and in science ($d = -0.15$). This contrasts to the relatively small pattern of higher male performance documented earlier on standardized tests.

The exact reasons for the test and grade discrepancy in mathematics and science are complex and require further educational research. However, we cannot rule out the possibility that, especially on high-stakes standardized tests (where test anxiety are higher and all those self-efficacy doubts come in to play), girls show reduced performance but are actually equal or even greater in ability when alternate forms of assessment (such as assignments, essays, lab reports) are used. Unlike the gender difference observed for verbal and language ability addressed earlier, when it comes to science and mathematics there remains a degree of uncertainty about what the data from these standardized tests represent. Quite often in debates about the merit of single-sex or co-educational schooling, we see a selective cherry-picking of the literature to suit the worldview of the author. It is with both bemusement and some concern that, in citation searches of my own work, I find myself cited as both validating and disputing the existence of meaningful gender differences in mathematics and science. But it would be remiss not to present the entire picture with all its limitations: data from standardized tests and from grades differ. And when we make sweeping statements about boys and girls abilities in STEM, we must also remember that any observations are culturally-bound. Students raised in different countries may not hold the same endorsement of gender-STEM stereotypes, and the direction of gender gaps typically vary cross-culturally. TIMSS and PISA data present a very different view of gender differences than are found in US-centric studies.

Summary

Data from large-scale educational assessments show that significant and meaningful gender differences in mathematics and science achievement are found in students from the United States and do not appear to be diminishing over time despite educational reform and changes in attitudes to the role of women in society. Further progress must, and can, be made. Psychological factors contribute to these differences, such as endorsement of gender-science stereotypes, attitudes toward mathematics and science, as well as self-efficacy beliefs. In younger students, gender differences in mathematics and science ability are minimal, but by adolescence we can see a divide in girls' and boys' attitudes and beliefs about STEM. These presage later gender differences in objectively measured ability with standardized tests. However, we also see a discrepancy between performance on standardized tests and student grades; girls earn significantly higher grades than boys for mathematics and all science domains. This is encouraging because it suggests we may have been underestimating girls' abilities all along and parallels their own lack of confidence in their abilities. Effective evidence-based interventions to boost female performance in STEM have been developed, but have found minimal adoption in the field. Additionally, we see a completely different pattern of results when we look at mathematics achievement internationally in large-scale assessments like TIMSS and PIRLS. In many countries, female students achieve parity or even greatly exceed their male peers in mathematics and science. This invalidates claims made by some researchers that males are better suited to STEM for biological reasons, and places on educators and policy-makers an obligation to act.

Educational implications of gender gaps in educational outcomes

Given the complexities involved and the need for specialized statistical training, it is no wonder that cognitive psychology studies investigating gender differences in cognitive abilities are read by very few outside of this discipline or that when they are, they are frequently misinterpreted. Many studies are limited by small samples and selection bias (Becker and Hedges, 1988), which limits the validity of any conclusions drawn. This becomes a problem when studies are cherry-picked rather than looking at the consensus, or drawing on studies that recruit demographically representative samples. However, the educational and public-policy implications of such research are also great, and touch on important issues such as pedagogical practices, child-rearing practices, early education, affirmative action in college/university placements, and the practice of single-sex or co-educational schooling (Halpern, 2014). While we may be living in a post-truth, "fake news" world where gut instinct or ideological faith sometimes drives our decision-making, the weighty implications for educational practice and public policy behoove us to evaluate the evidence carefully. For some types of cognitive tasks, gender differences are comparatively small and for others the gender gap is quite large. In fact, for tasks like writing and grammar, the effect sizes are so large that it may prompt us to reconsider the way in which these skills are taught and practiced inside and outside of the classroom. Good writers are not born but rather made: writing is a skill that improves with guided practice. This is especially true when direct feedback to the writing is given in a timely fashion, and students are then asked and supported to revise and edit their work (Parr and Timperley, 2010). While one gender may be predisposed to higher achievement, evidence-based interventions are available to mitigate this and reduce the gender gap, including the use of technology (Graham et al., 2018; Little et al., 2018). We also need to better support readers who come in at sub-clinical levels of reading impairment. This is a problem that should be approached collaboratively, with parents and early educators working hand-in-hand.

Prevailing gender stereotypes that mathematics and science are for men or that boys have a natural aptitude for them are very much alive in the minds of students, their parents, and their teachers. Yet, when we have a look at the actual evidence on the subject, the gender gap in mathematics and science is actually comparatively small in size when related to reading and writing. It is certainly not a trivial difference (especially when you look at the gender ratio of higher achievers, which falls roughly 2:1), but neither is it so large that gender parity is unattainable. When we look to international studies, in many countries girls meet or even greatly exceed boys in mathematics and science. But at least in Western nations, at some time during early high school, we start to see a decline in girls' general attitudes toward science while boys remain high. Data from self-reports tells us that girls start to doubt their mathematical and scientific aptitude as well. This directly impacts subject selection and serves as a bottleneck, leading to an underrepresentation of women pursuing STEM coursework in college and entering the STEM workforce. Educationally, we need to better support and nurture girls' interest in science and their self-confidence. The test and grade discrepancy shows us girls and women have the natural aptitude for STEM—even though they may sometimes doubt so themselves.

Conclusions

Reasonable people may reasonably disagree on how best to educate children, adolescents and young adults. Still, we should at least keep an open mind and model those same critical reasoning skills we wish to imbue our students with. Educators should be wary about reviews that "cherry-pick" the literature, or cite only very small and highly select samples. Evidence from large-scale, nationally representative samples show there are small but non-trivially-sized gender differences in mathematics and science favoring males. There are also far more pronounced differences in verbal and language abilities favoring females. Biology may well contribute to gender differences in reading and writing (as the literature shows moderately large differences are found cross-culturally), but it is a poor explanation at best when it comes to gender differences in mathematics and science since there are many countries where females outperform males. I have attempted to outline the evidence base for gender differences in cognitive abilities without wading too heavily into the partisan debate on single-sex schooling; on this matter it is difficult to persuade a mind

that is already fixed while a more open one can find the issue debated elsewhere. However, the review does present (at the time of writing) the most up-to-date view of gender differences in educational outcomes for verbal and quantitative reasoning. There is little evidence of any decrease in the magnitude of gender gaps in educational achievement at present, but I remain hopeful that this will not always be the case.

Acknowledgment

I thank my colleague Dr. Aida Hurem for her helpful suggestions and proofing of this manuscript.

References

- Becker, B.J., Hedges, L.V., 1988. The effects of selection and variability in studies of gender differences. *Behav. Brain Sci.* 11, 183–184. <https://doi.org/10.1017/S0140525X00049256>.
- Berninger, V.W., Nielsen, K.H., Abbott, R.D., Wijsman, E., Raskind, W., 2008. Gender differences in severity of writing and reading disabilities. *J. Sch. Psychol.* 46, 151–172. <https://doi.org/10.1016/j.jsp.2007.02.007>.
- Buchmann, C., DiPrete, T.A., McDaniel, A., 2008. Gender inequalities in education. *Annu. Rev. Sociol.* 34, 319–337. <https://doi.org/10.1146/annurev.soc.34.040507.134719>.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Eccles, J.S., 2013. Gender and achievement choices. In: Gershoff, E.T., Mistry, R.S., Crosby, D. (Eds.), *Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy*. Oxford University Press, New York.
- Elliot, L., 2009. *Pink Brain, Blue Brain*. Houghton Mifflin Harcourt, Boston, MA.
- Else-Quest, N.M., Hyde, J.S., Linn, M.C., 2010. Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. <https://doi.org/10.1037/a0018053>.
- Feingold, A., 1988. Cognitive gender differences are disappearing. *Am. Psychol.* 43, 95–103. <https://doi.org/10.1037/0003-066X.43.2.95>.
- Forman, B.R., Francis, D.J., Shaywitz, S.E., Shaywitz, B.A., Fletcher, J.M., 1997. The case for early reading intervention. In: Blachman, B.A. (Ed.), *Foundations of Reading Acquisition and Dyslexia: Implications for Early Intervention*. Routledge, New York.
- Furnham, A., Hosoe, T., Tang, T., 2001. Male hubris and female humility? A cross-cultural study of ratings of self, parental, and sibling multiple intelligence in America, Britain, and Japan. *Intelligence* 30, 101–115. [https://doi.org/10.1016/S0160-2896\(01\)00080-0](https://doi.org/10.1016/S0160-2896(01)00080-0).
- Furnham, A., Reeves, E., Budhani, S., 2002. Parents think their sons are brighter than their daughters: sex differences in parental self-estimations and estimations of their children's multiple intelligences. *J. Genet. Psychol.* 163, 24–39. <https://doi.org/10.1080/00221320209597966>.
- Geary, D.C., 1996. Sexual selection and sex differences in mathematical abilities. *Behav. Brain Sci.* 19, 229–247. <https://doi.org/10.1017/S0140525X00042400>.
- Graham, S., McKeown, D., Kihara, S., Harris, K.R., 2012. A meta-analysis of writing instruction for students in the elementary grades. *J. Educ. Psychol.* 104, 879–896. <https://doi.org/10.1037/a0029185>.
- Graham, S., Perin, D., 2007. A meta-analysis of writing instruction for adolescent students. *J. Educ. Psychol.* 99, 445–476. <https://doi.org/10.1037/0022-0663.99.3.445>.
- Graham, S., Rouse, A.G., Harris, K.R., 2018. Scientifically supported writing practices. In: O'Donnell, A. (Ed.), *The Oxford Handbook of Educational Psychology*. Oxford.
- Guiso, L., Monte, F., Sapienza, P., Zingales, L., 2008. Culture, gender, and math. *Science* 320, 1164–1165. <https://doi.org/10.1126/science.1154094>.
- Hall, J.A., 2005. Meta-analysis of nonverbal behavior. In: Manusov, V. (Ed.), *The Sourcebook of Nonverbal Measures: Going Beyond Words*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Halpern, D.F., 2010. How neuromythologies support sex role stereotypes. *Science* 330, 1320.
- Halpern, D.F., 2014. It gets crowded with an elephant and an ape in the room: teaching about female and male cognitive differences and similarities. *Teach. Psychol.* 41, 88–93. <https://doi.org/10.1177/0098628313514187>.
- Harmon, J., Wood, K., Smith, K., Zakaria, N., Ramadan, K., Sykes, M., 2016. Teaching and learning in high school reading classes: perspectives of teachers and students. *Read. Psychol.* 37, 962–994.
- Hedges, L.V., 2008. What are effect sizes and why do we need them? *Child Dev. Perspect.* 2, 167–171. <https://doi.org/10.1111/j.1750-8606.2008.00060.x>.
- Hedges, L.V., Nowell, A., 1995. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. <https://doi.org/10.1126/science.7604277>.
- Hurry, J., Sylva, K., 2007. Long-term outcomes of early reading intervention. *J. Res. Read.* 30, 227–248.
- Jensen, A.R., 1998. *The G Factor: The Science of Mental Ability*. Praeger, London.
- Little, C.W., Clark, J.C., Tani, N.E., Connor, C.M., 2018. Improving writing skills through technology-based instruction: a meta-analysis. *Rev. Educ.* 6, 183–201.
- Logan, S., Johnston, R., 2009. Gender differences in reading ability and attitudes: examining where these differences lie. *J. Res. Read.* 32, 199–214.
- Lynn, R., 1992. Sex differences on the differential aptitude test in British and American adolescents. *Educ. Psychol.* 12, 101–102. <https://doi.org/10.1080/0144341920120201>.
- Lynn, R., 1994. Sex differences in intelligence and brain size: a paradox resolved. *Pers. Individ. Differ.* 17, 257–271. [https://doi.org/10.1016/0191-8869\(94\)90030-2](https://doi.org/10.1016/0191-8869(94)90030-2).
- Maccoby, E.E., Jacklin, C.N., 1974. *The Psychology of Sex Differences*. Stanford University Press, Stanford.
- Miller, D.I., Halpern, D.F., 2013. Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learn. Individ. Differ.* 26, 141–152. <https://doi.org/10.1016/j.lindif.2012.03.012>.
- Miller, D.I., Nolla, K.M., Eagly, A.H., Uttal, D.H., 2018. The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US Draw-a-Scientist studies. *Child Dev.* 89, 1943–1955. <https://doi.org/10.1111/cdev.13039>.
- Mills, M., 2004. The media, marketing, and single sex schooling. *J. Educ. Pol.* 19, 343–360.
- Mix, K.S., Levine, S.C., Cheng, Y.-L., Stockton, J.D., Bower, C., 2020. Effects of spatial training on mathematics in first and sixth grade children. *J. Educ. Psychol.* 113, 304–314. <https://doi.org/10.1037/edu0000494>.
- Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., Ito, T.A., 2010. Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science* 330, 1234–1237. <https://doi.org/10.1126/science.1195996>.
- Muro, M., Maxim, R., Whiton, J., 2019. *Automation and Artificial Intelligence: How Machines Are Affecting People and Places*. Brookings Institution, Washington, DC.
- National Science Foundation, 2019. *Women, Minorities, and Persons With Disabilities in Science and Engineering: 2019*. National Science Foundation, Arlington, VA.
- Neisser, U.C., Boodoo, G., Bouchard Jr., T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., Urbina, S., 1996. Intelligence: knowns and unknowns. *Am. Psychol.* 51, 77–101. <https://doi.org/10.1037/0003-066X.51.2.77>.
- Nosek, B.A., Banaji, M.R., Greenwald, A.G., 2002. Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. <https://doi.org/10.1037/0022-3514.83.1.44>.

- Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N.M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., 2009. National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10593–10597. <https://doi.org/10.1073/pnas.0809921106>.
- OECD, 2019. PISA 2018 Results (Volume II): Where All Students Can Succeed. OECD Publishing, Paris.
- OECD, 2020. Education at a Glance 2020: OECD Indicators. Organisation for Economic Co-Operation and Development, Paris. <https://doi.org/10.1787/69096873-en>.
- Parr, J.M., 2010. A dual purpose data base for research and diagnostic assessment of student writing. *J. Writ. Res.* 2, 129–150. <https://doi.org/10.17239/jowr-2010.02.02.3>.
- Parr, J.M., Timperley, H.S., 2010. Feedback to writing, assessment for teaching and learning and student progress. *Assess. Writ.* 15, 68–85. <https://doi.org/10.1016/j.asw.2010.05.004>.
- Petrides, K., Furnham, A., 2000. Gender differences in measured and self-estimated trait emotional intelligence. *Sex. Roles* 42, 449–461. <https://doi.org/10.1023/a:1007006523133>.
- Rammstedt, B., Rammsayer, T.H., 2002. Self-estimated intelligence: gender differences, relationship to psychometric intelligence and moderating effects of level of education. *Eur. Psychol.* 7, 275–284. <https://doi.org/10.1027/1016-9040.7.4.275>.
- Reilly, D., 2012. Gender, culture and sex-typed cognitive abilities. *PLoS One* 7, e39904. <https://doi.org/10.1371/journal.pone.0039904>.
- Reilly, D., 2019. Sex Differences in Cognitive Abilities and Educational Outcomes: Examining the Contribution of Sex-Role Identification. Doctoral Dissertation. Griffith University, Australia.
- Reilly, D., Hurem, A., in press. Designing equitable STEM education: guidelines for parents, educators, and policy-makers to reduce gender/racial achievement gaps. In: Liu, X., Wang, L. (Eds.), *International Encyclopedia of Education*. vol. 13: STEM and Beyond. fourth ed. Oxford, UK: Elsevier.
- Reilly, D., 2021. Gender differences in reading, writing and language development. In: Zhang, L. (Ed.), *Oxford Research Encyclopedia of Education*. Oxford University Press, New York, pp. 520–542.
- Reilly, D., Neumann, D.L., Andrews, G., 2015. Sex differences in mathematics and science: a meta-analysis of National Assessment of Educational Progress assessments. *J. Educ. Psychol.* 107, 645–662. <https://doi.org/10.1037/edu0000012>.
- Reilly, D., Neumann, D.L., Andrews, G., 2017. Investigating gender differences in mathematics and science: results from the 2011 trends in mathematics and science survey. *Res. Sci. Educ.* 1–26. <https://doi.org/10.1007/s11165-017-9630-6>.
- Reilly, D., Neumann, D.L., Andrews, G., 2019. Gender differences in reading and writing achievement: evidence from the National Assessment of Educational Progress (NAEP). *Am. Psychol.* 74, 445–458. <https://doi.org/10.1037/amp0000356>.
- Reynolds, M.R., Scheiber, C., Hajoovsky, D.B., Schwartz, B., Kaufman, A.S., 2015. Gender differences in academic achievement: is writing an exception to the gender similarities hypothesis? *J. Genet. Psychol.* 176, 211–234. <https://doi.org/10.1080/00221325.2015.1036833>.
- Riegle-Crumb, C., Kyte, S.B., Morton, K., 2018. Gender and racial/ethnic differences in educational outcomes: examining patterns, explanations, and new directions for research. In: Schneider, B. (Ed.), *Handbook of the Sociology of Education in the 21st Century*. Springer International Publishing, Cham.
- Shaywitz, S.E., Escobar, M.D., Shaywitz, B.A., Fletcher, J.M., Makuch, R., 1992. Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *N. Engl. J. Med.* 326, 145–150. <https://doi.org/10.1056/NEJM199201163260301>.
- Shettle, C., Rhee, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., Brown, J., 2007. The Nation's Report Card: America's High School Graduates—Results From the 2005 NAEP High School Transcript Study. National Center for Education Statistics, Washington, DC.
- Spelke, E.S., 2005. Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *Am. Psychol.* 60, 950–958. <https://doi.org/10.1037/0003-066X.60.9.950>.
- Stanley, J.C., Benbow, C.P., Brody, L., Dauber, S., Lupkowski, A., 1992. Gender differences on eighty-six nationally standardized aptitude and achievement tests. In: Colangelo, N., Assouline, S.G., Ambrosio, D.L. (Eds.), *Talent Development: Proceedings From the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development*. Trillium Press, Unionville, NY.
- Sternberg, R.J., 2014. Teaching about the nature of intelligence. *Intelligence* 42, 176–179. <https://doi.org/10.1016/j.intell.2013.08.010>.
- Szymanowicz, A., Furnham, A., 2011. Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: four meta analyses. *Learn. Indiv. Differ* 21, 493–504. <https://doi.org/10.1016/j.lindif.2011.07.001>.
- Thomas, D.P., 2020. Rapid decline and gender disparities in the NAPLAN writing data. *Aust. Educ. Res.* 47. <https://doi.org/10.1007/s13384-019-00366-8>.
- Voyer, D., Voyer, S., 2014. Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. <https://doi.org/10.1037/a0036620>.
- Voyer, D., Voyer, S., Bryden, M.P., 1995. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117, 250–270. <https://doi.org/10.1037//0033-2909.117.2.250>.
- Wallentin, M., 2009. Putative sex differences in verbal abilities and language cortex: a critical review. *Brain Lang.* 108, 175–183. <https://doi.org/10.1016/j.bandl.2008.07.001>.
- You, D., Maeda, Y., Bebeau, M.J., 2011. Gender differences in moral sensitivity: a meta-analysis. *Ethics Behav.* 21, 263–282. <https://doi.org/10.1080/10508422.2011.585591>.
- Zhang, M., Bennett, R.E., Deane, P., van Rijn, P.W., 2019. Are there gender differences in how students write their essays? An analysis of writing processes. *Educ. Meas.* 38, 14–26.

Further reading

- Buchmann, C., DiPrete, T.A., McDaniel, A., 2008. Gender inequalities in education. *Annu. Rev. Sociol.* 34, 319–337. <https://doi.org/10.1146/annurev.soc.34.040507.134719>.
- Halpern, D.F., 2014. It gets crowded with an elephant and an ape in the room: teaching about female and male cognitive differences and similarities. *Teach. Psychol.* 41, 88–93. <https://doi.org/10.1177/0098628313514187>.
- Halpern, D.F., Eliot, L., Bigler, R.S., Fabes, R.A., Hanish, L.D., Hyde, J.S., Liben, L.S., Martin, C.L., 2011. The pseudoscience of single-sex schooling. *Science* 333, 1706–1707. <https://doi.org/10.1126/science.1205031>.
- Voyer, D., Voyer, S., 2014. Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. <https://doi.org/10.1037/a0036620>.

Networks and learning

Peter Goodyear, Centre for Research on Learning and Innovation, Sydney School of Education and Social Work, The University of Sydney, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Orientation	409
Taking a network perspective	409
Learning networks and networked learning	410
Productive learning networks	411
Learning network architectures	411
Personal and organisational perspectives	411
Communities and networks	411
A brief history of networked learning	411
Origins	411
Research and development in networked learning	412
The transition to social media and to a postdigital world	413
Approaches to research and development	413
Analyzing online conversations	413
Understanding participants' experiences	413
(Social) network analysis	414
Mapping participation	416
Activity-centred analysis and design (ACAD): design research and the creation of design knowledge	417
Practical applications: examples from teachers' professional learning	418
Personal learning networks (PLNs)	418
Externally organised/managed networks	418
Heterogeneous networks	419
References	420

Orientation

This chapter provides an overview of research and practice in the field of networked learning. It describes some of the core characteristics of the field, summarizes the history of its development and identifies connections with related work that goes under other names, such as online or blended learning. The main part of the article introduces and discusses key concepts and identifies some foci and approaches to research. Examples of practical applications are provided.

Taking a network perspective

For a researcher interested in learning, taking a network perspective is a deliberate choice. It is possible to examine learning and related educational phenomena from other perspectives – spatial, systems, ecological, architectural, market and hierarchical, for example. Different framings and combinations of framings make it easier to see some kinds of patterns, and harder to see others. The choice of perspective can also affect the kinds of relationships and explanations that may be readily advanced. For example, the use of a spatial perspective foregrounds location, presence, proximity, place-making and containment. It is common, though often implicit, in accounts of classroom processes. Similarly, a systems perspective foregrounds dynamic change, positive and negative feedback, and emergence. Those choosing a network perspective are foregrounding connections.

Exactly what is connected in a network can also be open to definition by the researcher. Networks consist of nodes and connections between nodes. The connections are also referred to as links, ties or edges. Each node may represent a person, an organization, a device, an idea – indeed, almost anything can be represented as a node and almost any of its relations to other nodes can be represented as a link. The question for the researcher is whether this way of looking at a particular set of phenomena of interest provides insights that would be hard to obtain in other ways. As Jones (2004, 82) puts it: “when we analyze in terms of networks, we also help to perform networks into being.”

Some networks can be said to have an existence independent of a researcher's gaze. Obvious examples are artificial networks such as those used for electricity distribution, for transport (e.g., road, rail and airline networks) and in digital communications (e.g., the Internet). Everyday use of network terminology can also lead us to think that networks exist in nature – as with a spider's web or among the synapses and neurons of an animal's brain. But even with these examples, there is an element of researcher choice about how, exactly, to carve up reality.

It is not just researchers who can bring networks into being. For example, over the last 25 years, a number of educational practitioners have chosen to describe their work in terms of “networked learning”. Part of their professional activity helps bring into being various kinds of “learning networks”, which are then open to researchers to study.

It is often said that we live in a networked world (Castells, 1996, 2001; Rainie and Wellman, 2012). In this regard, Dohn (2018) points to six different senses with which “networked” is used:

1. Linking geographically dispersed places through an infrastructure of connected computers,
2. Communication that increasingly relies on the mediation of the Internet,
3. Information processing – including retrieval, exchange and archiving – that is carried out between machines, with little human involvement,
4. Denoting the way that “our lives play out in hybrid spaces where physical and virtual resources integrate in complex ways to support us in our activities” (op. Cit., p. 5),
5. Acknowledging the importance of connections between people – that we have personal networks and engage with a range of networks relevant to different areas of our lives, and
6. To note that events and situations are linked to other events and situations by technologies and other people.

The fifth and sixth of these senses are not intrinsically dependent on computers and the digital. Though the rapid development and take-up of digital technologies have changed how people and events connect, networks predate the digital. Indeed, there is a growing body of research in areas such as history, archaeology and economics using network techniques to understand trade, communication, and kinship networks, running back into antiquity (Hodder, 2012; Knappett, 2013). Some of these networks have/had an explicit learning function, as in networks of scientists and scholars (Collar, 2013; Pietsch, 2013). And one might argue that virtually all networks involve some significant incidental learning, such as when traders learn to build trust and to share information up and down supply chains about the quality of products and about market preferences (Duguid, 2005; Hancock, 2005). Goodyear (2014) notes resemblances between the networks based on London’s 18th century coffee shops, with their libraries and magazines, and the functioning of some contemporary learning and innovation networks based on social media.

Social scientists interested in networked organisational forms have observed that networks can be said to possess some advantages over the two best known alternatives: markets and hierarchies. Networks combine some of the flexibility of markets with some of the predictability of hierarchies.

We define a network form of organization as any collection of actors ($N \geq 2$) that pursue repeated, enduring exchange relations with one another and, at the same time, lack a legitimate organizational authority to arbitrate and resolve disputes that may arise during the exchange. In a pure market, relations are not enduring, but episodic, formed only for the purpose of a well-specified transfer of goods and resources and ending after the transfer. In hierarchies, relations may endure for longer than a brief episode, but a clearly recognized, legitimate authority exists to resolve disputes that arise among actors.

Podolny and Page (1998, p. 59).

Many contemporary learning networks, as with some other kinds of digitally mediated networks and online affinity groups, have an organizing authority of some sort, though that authority may be quite distributed, fragile and limited in its ability to resolve disputes. Authority is more centralized and visible in learning networks that are organized as part of formal education provision, where networks take on some of the characteristics of hierarchies. Authority is more diffuse in informal self-organizing learning networks.

As with the early (predigital) networks found in areas such as trade and scholarship, the predictability of life in contemporary digital or hybrid networks depends heavily on reciprocity and trust. Indeed, network members typically spend a good deal of time maintaining mutually supportive relationships and most long-lasting networks have a repertoire of methods for building consensus and resolving differences through discussion. Building trust and creating shared understandings involves a good deal of information exchange: much more so than we find with markets. The boundaries of networks are easier to modify than is usually the case with hierarchical organisations. This combination of rich communication and flexible boundaries gives networks advantages over markets and hierarchies in times of change. One can argue that part of the attraction to networked ways of thinking and organizing comes from a view that we are living in dynamic, unpredictable times.

Learning networks and networked learning

What distinguishes a learning network from networks in general? Linda Harasim was among the first educators to popularize the term (e.g., Harasim et al., 1997) but neither she nor her collaborators came up with a satisfactory definition (Carvalho and Goodyear, 2014). The best approach to characterizing a learning network may be to start with “networked learning”. There is now a well-established field of research and development on networked learning, with biennial conferences and a book series. Until recently, the customary definition of networked learning, used in calls for papers and introductory texts ran as follows:

We define ‘networked learning’ as learning in which information and communications technologies are used to promote *connections*: between one learner and other learners, between learners and tutors; between a learning community and its learning resources. Some of the richest examples of networked learning involve interaction with on-line materials and with other people. But use of on-line materials is not a sufficient characteristic to define networked learning.

Goodyear et al. (2004, pp. 1–2).

One might then say that a learning network is the organisational form in which networked learning takes place. An updated version of the customary definition is given later in this article, but the fundamental idea of a learning network remains intact.

Productive learning networks

Carvalho and Goodyear (2014) extended the concept, in order to speak of productive learning networks. Here “productive” is not used as an evaluative term – the antithesis of unproductive or ineffective. It is used to highlight the creative and constructive aspects of networked learning, with an emphasis on learning through the co-creation of knowledge (production rather than consumption), and on self-realization and identity formation through participation in shared work.

Learning network architectures

Carvalho and Goodyear’s (2014) architectural approach to the study of learning networks is motivated by a goal of analyzing working examples of learning networks in order to extract reusable design knowledge: knowledge that is useful to people who have responsibility for helping other people learn together. They argue that architecture provides ways of acting and reasoning that are characterized by indirection, invitation, epistemic fluency and an ability to move fluidly across scale levels.

Architectural design works indirectly in that it creates built forms that are intended to have effects on how people act, think and feel. Like other areas of design, architecture knows how to work through invitation – it offers suggestions – rather than through compulsion. This designerly combination of invitation (e.g., suggesting worthwhile tasks) and indirect action (e.g., creating supportive learning environments) can help educators find ways of working skilfully with dynamically changing configurations of learner agency and structure.

Architectural practice also involves multiple forms of knowledge and combines diverse ways of knowing (e.g., from materials science, finance and art history). By extension, people who are designing to support networked learning, and to help a learning network flourish, need to know how to combine diverse kinds of knowledge, created through diverse epistemic practices, from across the learning sciences (Goodyear et al., 2016). Finally, an architect’s attention moves quickly back and forth between grand plans and the resolution of fine details. In hierarchical organisations, educational and otherwise, divisions of labor can render this fluid movement between high-level goals and implementational details very difficult. In contrast, architects know how important it is to get the devil out of the details.

The design knowledge produced by architecturally-informed analysis can be formalized and shared in a number of ways, including through design principles (e.g., Dohn et al., 2020) and design patterns (Goodyear, 2005; Goodyear and Retalis, 2010; Mor et al., 2014.) Design principles are general statements aimed at supporting reasoning about connections between desired outcomes and designable artifacts. Design patterns describe well-tested design solutions, and help explain why they work in specific contexts.

Personal and organisational perspectives

Learning networks can be analyzed and understood from at least two perspectives: a first-person perspective, in which the experience of a single person is foregrounded, and where their view of the network(s) in which they are involved is given primacy, and a holistic or organisational perspective, in which structural and/or aggregate data are gathered about an entire network. Sociological studies of networks refer to the first kind in terms of “ego networks”. In networked learning research, “personal learning network” and “professional learning network” refer to ego-networks used for learning. (See Practical Applications in the Section on **Practical applications: examples from teachers’ professional learning**, below).

Communities and networks

The terms “community” and “network” are sometimes used as synonyms, as with “online learning community” and “learning network”. But sometimes the broader connotations of the terms make one preferable to the other. For example, “community” can suggest stability, closure and an inward-looking orientation, whereas “network” can imply a degree of openness and flux. “Network” implies travel – of ideas, people, goods or messages. “Community” does not.

A brief history of networked learning

Origins

The origins of networked learning can be traced back to experimental work by Murray Turoff for the US Office of Emergency Preparedness. Turoff and colleagues at the New Jersey Institute of Technology developed a system called EMISARI (the Emergency Management Information System and Reference Index) in 1970–71, based on earlier use of Delphi techniques to solicit advice from

geographically distributed teams of experts. EMISARI offered many of the functions that we now associate with online discussion systems, or computer-mediated communication systems: sharing messages and other texts, discussion threads, tagging texts, search tools and notifications of updates. Until the early 1980s, EMISARI and its successor system (EIES: the Electronic Information Exchange System) were seen mainly as tools for expert users, including networks of scientists working on fast-moving topics. In 1982 came the first use of EIES in teaching, creating what [Hiltz \(1986\)](#) dubbed a “Virtual Classroom”. A number of researcher-practitioners pioneered various pedagogical and technological arrangements through the 1980s, mainly in the United States, though sometimes connecting up with participants in other countries. For example, [Riel and Levin \(1990\)](#) describe several pilot experiments using email to connect teachers and schools, including in international networks. [Quinn et al. \(1983\)](#) provide some models for analyzing differences in discourse among students in networked and face-to-face classes.

By the early 1990s, the use of digital networks to connect students, teachers and learning resources was becoming more widespread, especially at the university level and for postgraduate and professional development programs of various kinds. Experimental work was reported under a number of different labels, often anchored in technological rather than educational choices (e.g., [Mason and Kaye, 1989](#); [Steeple et al., 1994, 1996](#), writing about “computer-mediated communications” or [Hardy et al., 1994](#), on “computer conferencing”). Ten years later, the educational lexicon had grown to accommodate “online learning” and “online teaching”, “e-learning” and “blended learning” ([Clark and Mayer, 2002](#); [Harasim et al., 1997](#); [Salmon, 2000](#); [Thorne, 2003](#)). These particular streams of educational innovation and associated research did not necessarily attend to the forms of social learning that characterize networked learning, an exception being the subfield of computer-supported collaborative (or co-operative) learning – known widely as CSCL ([Dillenbourg et al., 1995](#); [Koschmann, 1996](#); [O’Malley, 1995](#)). Within CSCL, the vast majority of research has focused on the dynamics of small group learning. Research at the scale of communities or extended networks has been rare ([Goodyear et al., 2013](#)). It is this combination of (a) insistence on the importance of social learning and (b) social learning at scale, that distinguishes the area of networked learning research and practice.

Networked learning crystallized in the late 1990s by distinguishing itself from developments in digital education that were undermining human connectivity— developments that threatened to reduce education to the production, delivery and consumption of ‘content’ (‘online materials’).

Networked Learning Editorial Collective (NLEC) (2021, p. 315).

Research and development in networked learning

In the UK and the rest of Europe, a network of networked learning researchers began to take shape in the late 1990s, brought together around a series of biennial conferences and resourced, in part, by some linked research project teams ([Goodyear, 2014](#); [Hodgson and McConnell, 2019](#); [McConnell et al., 2012](#)). Many of the researchers shaping this field were/are also educators, and so accounts of networked learning often mix normative and descriptive statements. In other words, the teacher-researchers involved have strong views about pedagogy and educational relations. The texts they produce combine statements about what is, and what ought to be. For example, in analyzing the distinguishing characteristics of an MED program in Networked Learning at the University of Sheffield, [Hodgson & McConnell \(2019\)](#) identified six key features:

1. A commitment to democracy and openness in the educational process.
2. Self-managed learning, allowing freedom to learn.
3. A real purpose to cooperative activities.
4. A supportive learning environment.
5. Collaborative assessment of learning.
6. Ongoing review of learning activities.

They see these as being actualized through work on four design components: a learning community, action research, learning sets and workshops.

Abstracting from their experiences as networked learning teachers and researchers, [Hodgson and McConnell \(2019\)](#) offer eight principles to underpin networked learning.

1. The focus is on learning which has a perceived value to the learners.
2. Responsibility for the learning process should be shared (between all actors in the network).
3. Time has to be allowed to build relationships.
4. Learning is situated and context dependent.
5. Learning is supported by collaborative or group settings.
6. Dialogue and social interaction support the co-construction of knowledge, identity and learning.
7. Critical reflexivity is an important part of the learning process and knowing.
8. The role of the facilitator/animator is important in networked learning.

From the outset, networked learning has embraced a strong, though not always explicit, critical and emancipatory tradition. The customary definition, reproduced above, worked well as a way of providing focus for the first 20 years, but the rationale for its emphasis on connections was not explained; it was silent on the question of purpose. In an effort to update the definition and address this omission, a team of networked learning researchers jointly created the following aspirational statement.

Networked learning involves processes of collaborative, co-operative and collective inquiry, knowledge-creation and knowledgeable action, underpinned by trusting relationships, motivated by a sense of shared challenge and enabled by convivial technologies. Networked learning promotes connections: between people, between sites of learning and action, between ideas, resources and solutions, across time, space and media.

NLEC (2021, p. 320).

It remains to be seen whether this statement, or something like it, will be adopted as a description of the character and purposes of networked learning, going forward. A broad range of views on the matter can be found in NLEC et al. (2021).

The transition to social media and to a postdigital world

It is perhaps surprising that some of the preoccupations to be found in research and descriptions of innovative practice in the early days of networked learning are still visible in contemporary literature. Interest remains strong in such topics as the role of teachers, trust in peer learning, the nature of discourse and argumentation, identity formation, digital exclusion, and social justice. This steadiness of gaze has been maintained during a 20–30 year period in which the technologies used have changed almost beyond recognition, in terms of performance, usability, access and familiarity. Writing about networked learning in the 1990s was work from the margins. Now, widespread use of social media – Twitter, Facebook, Instagram, Tik-Tok and the like – means that learning from other people over digital connections is almost ubiquitous and goes almost unremarked. This raises interesting questions about differences between intentional and incidental learning (and teaching) and between formal, non-formal and informal educational arrangements.

Accelerating availability of the technologies underpinning networked learning has also been changing people's expectations and habits, such that much of what people do, and expect to be able to do, involves quite complex arrangements of digital and material things. The rapid interpenetration of the digital and material worlds has caused some researchers to argue that we are now in a post-digital age: one in which digitalization remains important and worth continuing scrutiny, but in which the binary distinctions between "real" and "virtual" are no longer helpful (Fawns, 2019; Macgilchrist, 2021; Sinclair and Hayes, 2019).

Approaches to research and development

Those who research networks use a variety of methodologies, with a range that includes approaches that use sophisticated mathematical techniques to analyze network structures and approaches that elicit nuanced accounts of participants' experiences (see e.g., Barabási, 2002, 2016; Jones, 2015). The fact that much of the research in networked learning is carried out by educators inquiring into their own praxis means that a good deal of the work is relatively small scale, locally focused and sensitive to context (Goodyear et al., 2006; Jones, 2015; Levy, 2003). This section provides a brief summary and follow-up references to five approaches: analyzing online conversations, understanding participant experiences, social network analysis, mapping the quality of participation and activity-centred analysis and design.

Analyzing online conversations

The persistence and searchability of digital texts afford research on the structure of educational discussions. This opportunity was seized quite rapidly by researchers already interested in (face-to-face) classroom discourse, so there is now a 40-year corpus of research examining what is said by whom, with what frequency and what results (Häkkinen, 2013; Hara et al., 2000; Hardy et al., 1994; Henri, 1992; Howell-Richardson and Mellar, 1996; Ouyang and Chang, 2019; Quinn et al., 1983; Schrire, 2006).

It is impossible to summarize briefly the findings emerging from this literature, especially given the diversity of pedagogical strategies, educational designs, aims and purposes, age ranges and technological devices and infrastructures involved. That said, there are some questions to which research frequently returns, often deriving similar answers. These include (a) the importance of shared understandings of the intentions of teachers and students in networked learning discussions, (b) the value of appropriate modeling and scaffolding for students' participation in such discussions, (c) the significance of ethics of participation, particularly in avoiding the marginalization of some subsets of participants (Dohn et al., 2018; Jones, 2015).

Understanding participants' experiences

Within research on networked learning, there is also a well-established and strong line of research into the experiences of participants, including the experiences of those in formal education who have recognized roles as students or teachers. Much of this research, especially in higher education, has used a phenomenographic approach, based on the ideas and methods of Marton and colleagues (Marton, 1984; Marton and Booth, 1997). For example, Jones and Asensio (2002) studied the networked learning design experiences of university teachers. Roberts (2003), in the United Kingdom, and Gonzalez (2010), in Australia, identified a hierarchy of conceptions of online teaching held by university lecturers. Cutajar (2017) and Ellis and Goodyear (2010) have similarly identified more and less sophisticated conceptions of networked learning among university students. Related research has

made innovative use of digital ethnography to create detailed accounts of participant experiences in a range of networked learning situations (Hodgson and Watland, 2004; Jones et al., 2008).

(Social) network analysis

There is much more to be said about a network than can be inferred from the experiences of participants. Nor can all its characteristics be defined as aggregates of the actions or wishes of participants.

Network-based analysis and reconstruction methods can offer insight into local and global properties of systems of interconnected objects, which can neither be discovered by studying the interacting agents individually or in pairs, nor by studying the average properties of the system as a whole ... network models have the power to facilitate structural comparison across different scales and source materials and to formulate hypotheses and models for further exploration.

Sindbæk (2013, p. 72).

Sindbæk was speaking of network analysis in archaeology, but his claim is also sustainable in relation to learning networks. A number of educational researchers have applied mathematical methods for analyzing and modeling learning networks, and indeed social network analysis (SNA) and related methods have become an important part of the field of learning analytics (De Laat et al., 2007a, b; Gunawardena et al., 2016).

Figs. 1–3 have been constructed to help explain the kinds of analysis and visualization work that can be done, and to give examples of some of the sorts of insights that can be generated. The illustration created for this purpose is set in a formally organized learning network and offers a simplified representation of interactions between a class of 32 students who are using an online discussion board.

Each student is shown as a node. A tie is shown between a pair of nodes if, and only if, the students concerned have engaged in an exchange of messages. An exchange of messages is defined in this example as: (a) a student posts an online discussion item, (b) another student posts a response to that item, building on what the first student has written, (c) the first student posts an acknowledgment, extending the line of discussion further. Fig. 1 shows the ties (exchanges of messages, as just defined) at the end of the first week. It reveals that most students have been involved in at least one exchange. (Students A and Z are the only exceptions.) A few students have been involved in exchanges with more than one other student. (Students C and G are examples.) Fig. 1 is typical of classes in which the first week's set task is to post, respond and acknowledge and where most students have (a) satisfied the task requirement, but no more and (b) have done this reciprocally, in a pair.

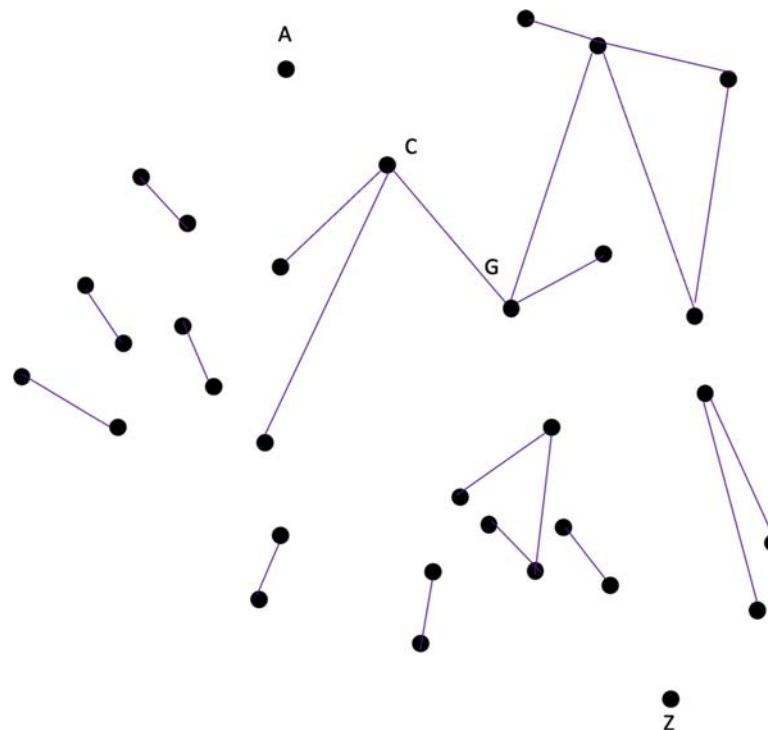


Fig. 1 Network connections established through online discussion after one week.

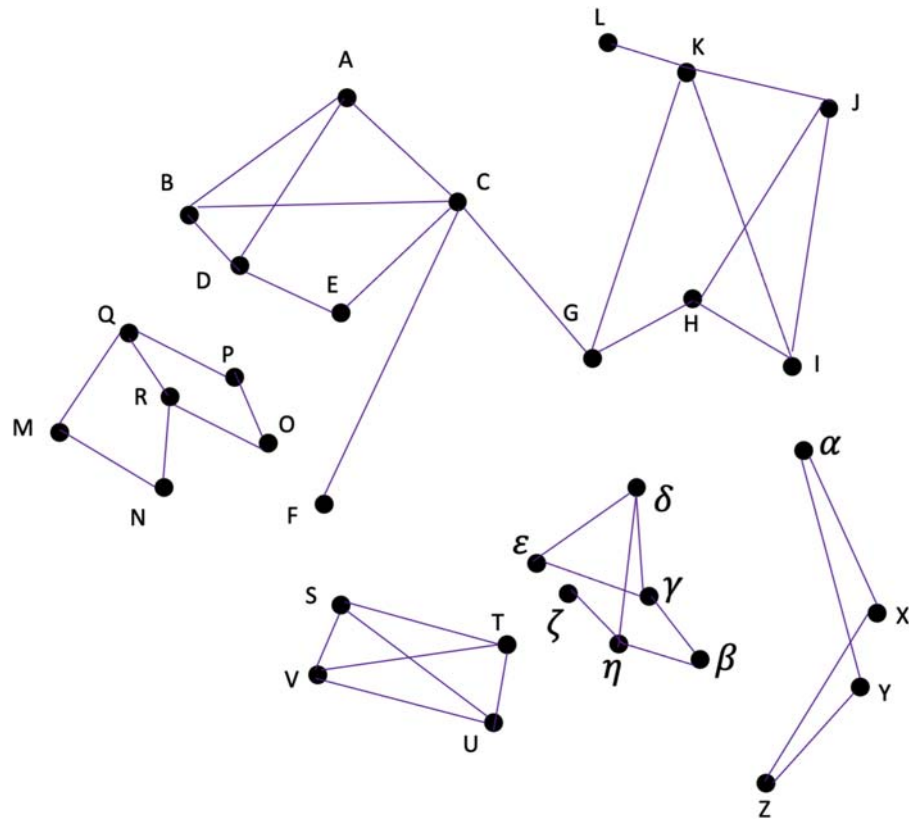


Fig. 2 Network connections after two weeks.

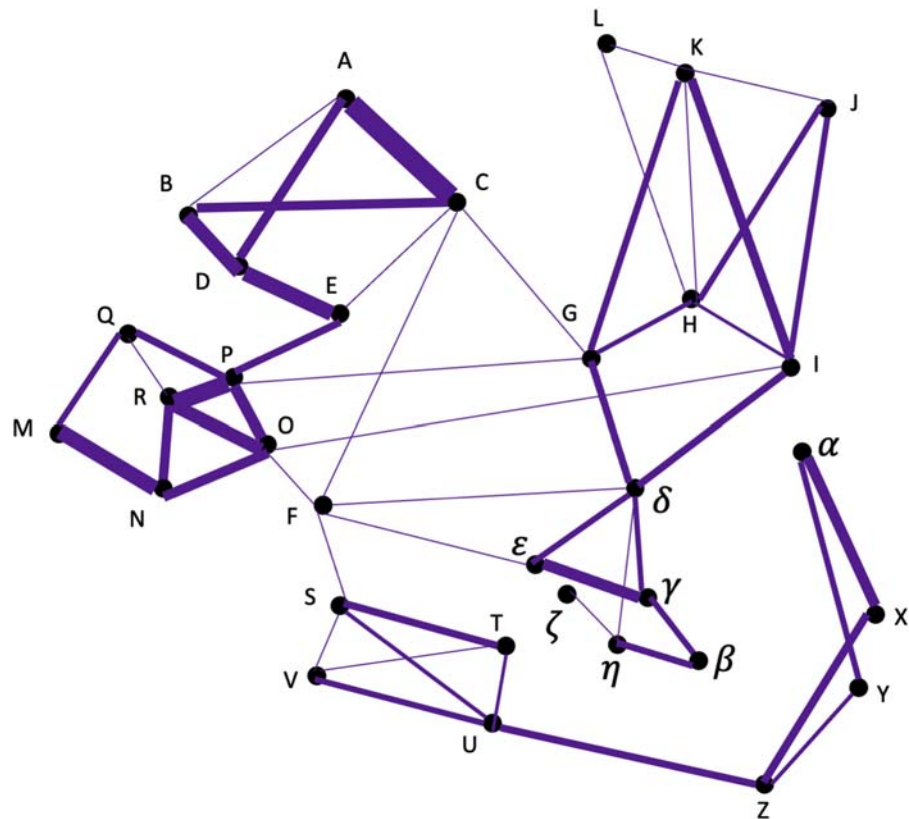


Fig. 3 Network connections toward the end of the course.

Fig. 2 shows the situation at the end of the second week. It reflects a suggestion from the class teacher to the students that they should work in small groups, of four to six members. So, for example, students A to E appear to have formed a group of five; M to R a group of six; S, T, U and V a group of four. All students have now engaged in at least one exchange of messages. It is not clear, at this point, whether student F is a member of a group.

Fig. 3 shows connections between members of the class after several weeks of discussion. The width of the line for each tie is proportional to the number of exchanges between the pair of students linked. So, for example, there have been a substantial number of exchanges between A and C, and between R and P. Fig. 3 shows that the small group structure that commenced in week 2 has persisted: many of the exchanges have been within groups. But there have also been exchanges between students in different groups. For example, student F has been involved in a single exchange with five other students, in four different groups. Network analysts refer to the connections represented by thick lines in Fig. 3 as “strong ties”, whereas those characterized by few exchanges are “weak ties” (Granovetter, 1973; Jones et al., 2008; Ryberg and Larsen, 2008). It is important to note that weak ties can be useful in passing information between different groups, so F may well be playing an important role in this network.

Of course, there are strict limits to what can be inferred by an outsider solely from the presence and absence of connections. Student F may have been tasked by the teacher to act as a rapporteur, harvesting and sharing information between groups. Or it could be the case that student F’s messages are banal or offensive and nobody wants to engage in an ongoing correspondence with them. Finding out more about the nature of the discussion, exchanges and relationships requires one to look at the texts exchanged, not just the pattern of connections. For this reason, (social) network analysis of this kind is often complemented with analysis of the content of messages (De Laat et al., 2007b).

Mapping participation

This approach to researching learning networks focuses on characterizing the nature of collaborative encounters – where people choose to collaborate to accomplish some shared aim, which may be, for example, professional development, or planning joint action of some kind. It is particularly useful in tracking the evolution of a learning network over time. The original ideas about mapping intentional collaboration come from Manzini (2015) and were applied to professional learning networks in Goodyear (2019). The approach works with four dimensions: degree of active involvement, degree of collaborative involvement, strength of social ties and relational intensity. These are explained in Table 1.

Dimensions 1 and 2 allow one to map participant involvement (PI), while dimensions 3 and 4 capture interaction quality (IQ). Figs. 4 and 5 show how PI and IQ can be mapped, for any one network, at any stage in its history. The explanation is from Goodyear (2019).

Fig. 4 maps the space of participant involvement (PI): who does what, with whom and how. A high level of active involvement coupled with a low level of collaboration can be characterized as a DIY (do-it-yourself) arrangement. Self-drive car-sharing schemes are a typical example: high user input to the service/activity itself, but little or no contact with other people. In the networked learning area, we could think of self-directed learning from online repositories of instructional videos as an analogous case. Low levels of both active involvement and collaboration are typical of mainstream service delivery. Using a “ride sharing” service or getting a quick answer to a question via an online chat-based helpdesk are examples here. Higher levels of collaboration, coupled with low levels of personal active involvement in the service itself, are commonly found with co-managed services, as when a group of people work together on policies or high-level management issues but do not get directly involved in actual service delivery. A housing co-op which employs maintenance and cleaning staff is an example, or a university management team whose members are

Table 1 Four dimensions for mapping collaborative encounters.

(1) Degree of active involvement	Participants may take quite a passive role – partaking of a service provided by another person – or they may more actively co-produce the service, fielding personal resources as they do so (time, energy, attention, skills, expertise, etc).
(2) Degree of collaborative involvement	This may range from close to zero (doing almost everything alone or virtually alone) to intense (engaging closely with others).
(3) Strength of social ties	May vary from weak to strong. Weak ties can be created quite quickly and may not persist; strong ties take time and commitment.
(4) Relational intensity	Characterizes the affective and empathetic qualities of the encounter, the depth of the relationship, the degree to which people treat each other as fellow human beings rather than (say) seller and buyer.

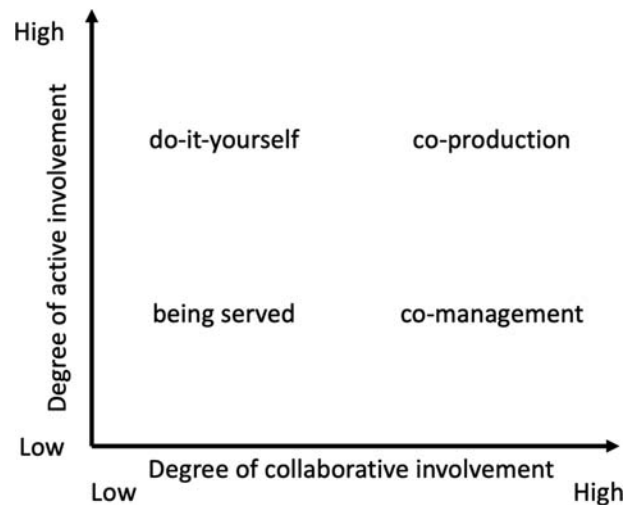


Fig. 4 Mapping collaborative encounters – participant involvement (PI).

no longer involved in teaching. Finally, **Fig. 4** situates co-production as involving high levels of active involvement in production of the service itself and of collaborating with others. People who both organize and do voluntary work in a community garden are in this zone, as are people who collaborate within a self-managed learning network. It is important to point out that high levels of active or collaborative involvement are not in themselves virtuous. It would be exhausting to live one's life entirely in that way. The key issue is to make, and help others make, considered choices about collaborative engagements.

Fig. 5 maps a similar space but for interaction quality (IQ). This is a less elaborated construct in Manzini's work: the space is rendered simply in terms of the strength of social ties and the quality of the relationship (deep, warm and personal or formal and rule-defined). A reason for this may be that the dimensions of PI have the capacity to frame deliberate action in advance, whereas those of IQ reflect something more emergent: qualities that become clearer in and after the fact. Indeed, Manzini's examples say more about the warmth or depth of connection in an encounter as experienced, and rather less about characteristics of the arrangements that might be recognisable or clearly specifiable *a priori*. (One can make a deep and warm connection with someone even in a highly regulated, formal situation.) As with PI, none of these arrangements is intrinsically better. For example, being able to collaborate in weak-tie situations makes life more open and less tribal.

Activity-centred analysis and design (ACAD): design research and the creation of design knowledge

The kinds of research methods and tools sketched above can be used for a variety of purposes. This final sub-section links them into a higher level framework known as ACAD. ACAD is a meta-theoretical framework that can be applied to understand and improve the functioning of a learning network (Goodyear et al., 2021). There are two distinctive features to ACAD. First, the framework

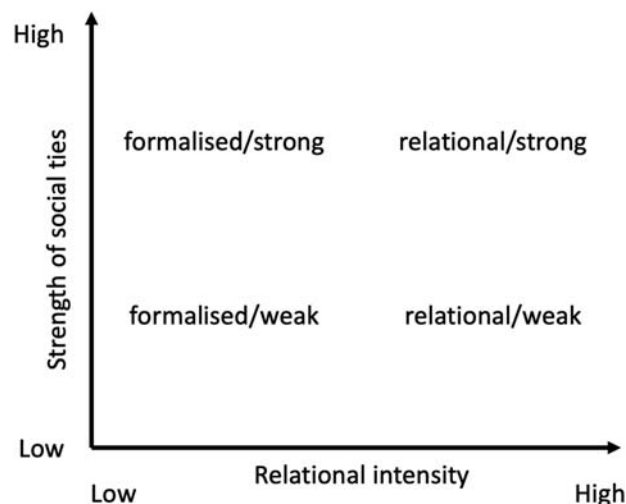


Fig. 5 Mapping collaborative encounters – interaction quality (IQ).

insists on a clear separation between (a) “design time” or a forward-looking design perspective and (b) “learning time” or a retrospective examination of actual, emerging experiences. The design perspective is taken by people who have a stake in supporting and improving the functioning of a learning network. It looks at a set of representations of the network and involves discussions about what can and should be changed (added, removed, fine-tuned, and so on). Any discussions held, and commitments made, need to pay sharp attention to pragmatics, to what can be changed through upfront design work. The retrospective analysis attends to the actuality of what has been occurring in the network – real-world activities that have real consequences for participants, including what they have learned and how. Ideally, this analysis and design work is cyclical, paced by the rhythms of the network. The second distinctive feature of ACAD is that it divides up the entities that situate learning activity into three kinds of design components: tasks (epistemic design), physical structures (set design, which includes material, digital and hybrid) and social structures (social design). At “design time”, some changes can be made to the entities in each of these groups. For example, the specification for a task may be modified, a new set of online resources may be added, and recommendations about group-structures or divisions of labor may be revised. At “learning time”, these all become entangled in the activities of the network. It makes sense to divide up the three different kinds of design component, because different kinds of reasoning are needed in each case. But, in the actuality of networked learning activities, they become harder to separate.

Practical applications: examples from teachers’ professional learning

This section provides some examples of networked learning in practice, focusing on teacher learning in professional networks. Recent surveys of this rapidly developing but still under-explored area of work can be found in [Kelly et al. \(2016\)](#), [Macia and Garcia \(2016\)](#), [Lantz-Andersson et al. \(2018\)](#) and [Kelly \(2019\)](#). Teachers’ engagement with others in networks for learning and professional development has been studied in a range of ways, using perspectives that foreground ego-centred personal networks, informal or self-organized networks and externally organized and managed networks. [Lantz-Andersson et al. \(2018\)](#) find a number of common themes in both formally organized (“top down”) and informally organized (“bottom up”) networks. For example, the value teachers report from network participation often resolves into one or more of: sharing information relevant to professional practice; benefitting from the filtering and curating work of peers (avoiding information overload); feeling collegial support – emotional and as a prompt to reflection – and encouragement to experiment and take risks.

Personal learning networks (PLNs)

[Trust et al. \(2016\)](#) report a survey of 732 schoolteachers, in which teachers were asked about their personal, professional learning networks (PLN). In this case, what is being asked about is each teacher’s personal (ego-centred) network: its nature, the purposes it serves and its effects, if any, on the educational activity and outcomes of the teachers’ students. [Trust et al. \(2016\)](#) argue that PLNs align with a “whole teacher” approach to professional development (PD) – in contrast with other forms of professional development that (can) only address a subset of any one teacher’s needs. Other forms of PD can suffer from problems of decontextualization, especially when the PD is offered outside school, and to meet the assumed needs of “typical” teachers, rather than the specific needs of actual teachers. Such forms of PD are necessarily piecemeal, unevenly and uncertainly distributed in time and space. In contrast, and by their very nature, PLNs are assembled and maintained by each teacher, in ways that match their own sense of the full range of their needs. This is not to say that teachers always succeed in managing their PLNs efficiently and effectively. Nor are they always able to learn what they need, or what their students need. The key distinguishing feature of PLNs, however, is that they adjust dynamically to changing needs and they integrate around each individual teacher. Given this personal focus, it is not surprising that PLNs are themselves diverse, typically involving multiple platforms, tools and media; often including both online and face-to-face interactions, and having multiple purposes. For example, in the [Trust et al. \(2016\)](#) survey, teachers revealed that the other people in their networks were drawn from many different areas of work and life – not just education – and from other school systems, including school systems in other countries. Teachers reported that their PLNs gave them a sense of being (internationally) connected educators, able to identify with communities of other enthusiastic teachers.

Teachers’ use of PLNs is widespread internationally, quite likely mirroring social media habits and Internet usage more generally. For example, [Le et al. \(2021\)](#) report a study of how English as a Foreign Language (EFL) teachers in Vietnam use social media for professional learning purposes. As with studies from other countries, the main uses include searching for, acquiring, and sharing teaching resources and building professional relationships.

Externally organised/managed networks

[Kelly et al. \(2018\)](#) describe *TeachConnect*, a platform created collaboratively by eight Australian universities to support newly qualified teachers entering the teaching profession. These first years in schools are well-documented as problematic, for many teachers. Poor experiences in classrooms and staffrooms lead significant numbers of early career teachers to leave the profession. *TeachConnect* was designed, in part, to help address this problem through mentoring, community support, help with induction and other forms of knowledge sharing (see also [Kelly et al., 2016](#)). [Kelly et al. \(2018\)](#) used the ACAD framework, described in the Section on **Activity-centred analysis and design (ACAD): design research and the creation of design knowledge** above, to structure their design and evaluation work. A major challenge identified during the early design work was the difficulty of creating trust within very

large networks. After some experience with the platform, Kelly et al. (2018) also noted the difficulty of securing ongoing participation from the majority of those who signed up. (Only about 25% of those registering were ongoing participants.) ACAD helped the team designing and running the network to identify some priorities for the next rounds of redesign, focusing particularly on aspects of social design.

In the space between PLNs and national networks, one can also find learning networks organized around action on a smaller scale. For example, van Amersfoort et al. (2019) report a study of two teacher learning and development networks, each based around a cluster of schools and each involving teachers, school managers and researchers. The networking arrangements were informed by a number of principles, such as that the teachers' networked learning should be embedded in their daily practices, promoting the sharing of problems of professional practice, and candidate solutions; the teachers should shape their own learning experiences, deciding on topics, timing and relationships, and networks should be established around particular problematic areas, but be free to evolve new interests and purposes over time. A special focus in this study was to identify different forms of value created within the network, using the model elaborated by Wenger et al. (2011). Notably, the authors found that teacher participants benefited through reframing their understandings of professional learning, including their sense of the value of learning from peers.

Heterogeneous networks

The discussion of networks thus far, in this chapter, has been restricted to homogeneous networks in which the nodes are of the same kind: people, or organisations, for example. Researchers also study heterogeneous networks, in which the nodes vary in kind. Notably, researchers may represent, as a network, the mix of people, devices, tools and artifacts involved in some selected event, process or practice. Purposes vary, but such networks can be used to explain causal connections, the propagation of effects, lines of influence and/or dependence, and so on.

The following example shows what we mean, and also serves as a further illustration of learning networks in action. Fig. 6 is based on materials from a study reported in greater detail in Goodyear and Carvalho (2019). The study involved an ethnographic approach to researching the role of iPads in the education of students who were learning to be paramedics. The initial focus of the study was on the specific contribution that use of the iPads made to the students' learning during a simulated mountain-rescue exercise. The exercise was held in a relatively remote upland area, over two days. For many students, this was their first experience of such an environment. Other participants in the exercise included some very experienced mountain rescue personnel and actors playing casualties.

The researcher was briefed to take an exploratory approach, roving around while the exercise unfolded, making field notes, and trying to identify and describe as carefully as possible the networks of things that seemed to influence the activities and their outcomes. The researcher's field notes mentioned that the iPads were used, though not often. Table 2 captures other tools and devices that featured in the exercise.

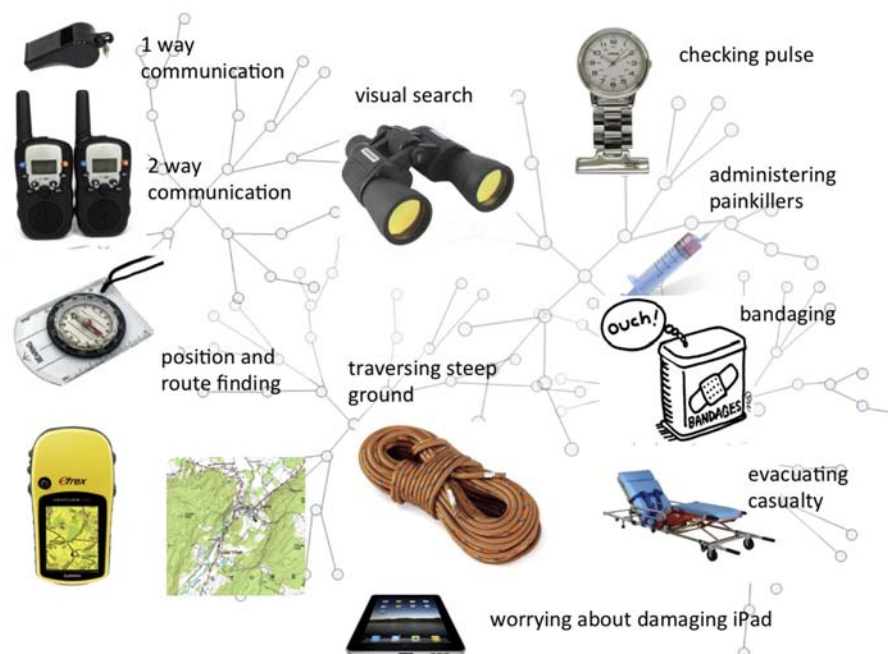


Fig. 6 Heterogeneous network: mountain-rescue exercise.

Table 2 Tools and people involved in the mountain rescue exercise.

<i>Tools for navigation and communication</i>	<i>Tools for “treatment” of simulated casualties</i>	<i>People</i>
Compasses, maps, GPS devices, torches, whistles, ropes, binoculars and walkie-talkies	Stretchers, bandages, scissors, watches (for measuring a pulse), stethoscopes, medications, and syringes	Actor/casualties (semi-skilled), the mountain rescue volunteers (very skilled), the tutors (semi-skilled) and the students (often lost, cold, and confused; worried about breaking the iPads)

Of course, the design and evolution of this exercise also necessitated being in the mountains.

The difficulties of traversing rough terrain, locating a casualty when hidden in a valley or by vegetation, coping with poor visibility and communicating without mobile phones all played a substantial role in the exercise. Proper clothing is also important. A conventionally-minded instructional designer might be forgiven for thinking that good boots and a warm, waterproof coat are things for the students to provide. But those who forgot these important items were unable to complete the exercise. And which instructional design guideline tells you that fingerless gloves are useful when trying to use an iPad on a cold mountain?

Goodyear and Carvalho (2019, pp. 58–9).

Two points can be made here. First, the conventional educational technologist’s framing of researching this exercise caused the initial focus to be on a single kind of device (the iPad) and on gathering evidence of whether or not it made a positive contribution to the students’ learning. This narrow aperture would have rendered invisible all of the other devices and people involved in the exercise. It is also far from clear that any simple measure of learning outcome would have shed light on the role of the iPads. Second, the opportunity to pay close attention to the activities enabled the research team to (re-) construct a depiction of the outcomes of the exercise: outcomes that were not expressed in these terms in any of the course’s documentation, and which largely remained implicit or (at best) were caught up in fragments of briefing and debriefing discourse. In short,

Successful participation in the exercise involved:

- Learning to use each of the tools listed in [Table 2](#), at least with sufficient fluency to be able to act according to the established protocols, but ideally with a level of automaticity that binds tool and action in a smooth flow
- Weaving the use of the tools into an integrated set of activities, involving smooth effective action, co-ordination with others, focusing on the priority goals, etc.
- Turning the individual and aggregate experiences of the exercise into learning that lasts.

The combination of observant participation and a network sensibility eventually enabled the research team to map a heterogeneous set of interdependent actions, devices, people, roles and places, and to draw out some otherwise implicit high-level goals:

The point of the exercise, for each student, is not just to master the individual tools but to participate in the construction of a co-ordinated web of activity that can result in a successful rescue, minimising danger to participants, and leaving traces, in some kinds of memory, that mean doing something like this again will not feel entirely new.

Goodyear and Carvalho (2019, p. 60)

Learning in networks can be rich and rewarding, but it is not straightforward to analyze or to promote. Patterns, experiences and outcomes can be very diverse. Learning networks can exist for a variety of purposes and not all of them are beneficial ([Castells, 2015](#)). As we have seen in recent years, digital networks are being used to spread misinformation and deepen mistrust. They can be used to organize peaceful protests but have also played a role in provoking and coordinating violence. Research needs to acknowledge that social learning in online networks can have a darker side. Researchers must be prepared to render more visible the multiplicity of ways in which networked learning unfolds, for good and ill.

References

- Barabási, A.-L., 2002. *Linked: The New Science of Networks*. Perseus Publishing, Cambridge MA.
- Barabási, A.-L., 2016. *Network Science*. Cambridge University Press, Cambridge.
- Carvalho, L., Goodyear, P. (Eds.), 2014. *The Architecture of Productive Learning Networks*. Routledge, New York.
- Castells, M., 1996. *The Rise of the Network Society*. Blackwell, Oxford.
- Castells, M., 2001. *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford University Press, Oxford.
- Castells, M., 2015. *Networks of Outrage and Hope: Social Movements in the Internet Age*, second ed. Polity Press, Cambridge.
- Clark, R., Mayer, R., 2002. *E-learning and the Science of Instruction*. Pfeiffer, San Francisco.

- Collar, A., 2013. Re-thinking Jewish ethnicity through social network analysis. In: Knappett, C. (Ed.), *Network Analysis in Archaeology: New Approaches to Regional Interaction*. Oxford University Press, Oxford, pp. 223–245.
- Cutajar, M., 2017. The student experience of learning using networked technologies: an emergent progression of expanding awareness. *Technol. Pedagog. Educ.* 26 (4), 485–499. <https://doi.org/10.1080/1475939X.2017.1327451>.
- De Laat, M., Lally, V., Lipponen, L., Simons, R.-J., 2007a. Investigating patterns of interaction in networked learning and computer-supported collaborative learning: a role for social network analysis. *Int. J. Comput. Support. Collaborat. Learn.* 2 (1), 87–103.
- De Laat, M., Lally, V., Lipponen, L., Simons, R.-J., 2007b. Online teaching in networked learning communities: a multi-method approach to studying the role of the teacher. *Instr. Sci.* 35 (3), 257–286.
- Dillenbourg, P., Baker, M., Blaye, A., O'Malley, C., 1995. The evolution of research on collaborative learning. In: Reimann, P., Spada, H. (Eds.), *Learning in Humans and Machines: Towards an Interdisciplinary Learning Science*. Pergamon, London, pp. 189–211.
- Dohn, N. (Ed.), 2018. *Designing for Learning in a Networked World*. Routledge, Abingdon.
- Dohn, N., Cranmer, S., Sime, J.-A., De Laat, M., Ryberg, T. (Eds.), 2018. *Networked Learning: Reflections and Challenges*. Springer, Cham.
- Dohn, N., Hansen, J.J., Goodyear, P., 2020. Basic design principles for learning designs to support knowledge transformation. In: Dohn, N.B., Hansen, S.B., Hansen, J.J. (Eds.), *Designing for Situated Knowledge Transformation*. Routledge, Abingdon, pp. 160–179.
- Duguid, P., 2005. Networks and knowledge: the beginning and end of the port commodity chain, 1703–1860. *Bus. Hist. Rev.* 21 (2), 109–118.
- Ellis, R., Goodyear, P., 2010. Students' Experiences of E-Learning in Higher Education: The Ecology of Sustainable Innovation. RoutledgeFalmer, New York.
- Fawns, T., 2019. Postdigital education in design and practice. *Postdigital Sci. Educ.* 1, 132–145.
- Gonzalez, C., 2010. What do university teachers think eLearning is good for in their teaching? *Stud. High Educ.* 35 (1), 61–78.
- Goodyear, P., 2005. Educational design and networked learning: patterns, pattern languages and design practice. *Australas. J. Educ. Technol.* 21 (1), 82–101.
- Goodyear, P., 2014. Productive learning networks: the evolution of research and practice. In: Carvalho, L., Goodyear, P. (Eds.), *The Architecture of Productive Learning Networks*. Routledge, New York, pp. 23–47.
- Goodyear, P., 2019. Networked professional learning, design research and social innovation. In: Littlejohn, A., Jaldemark, J., Vrieling, E., Nijland, F. (Eds.), *Networked Professional Learning: Emerging and Equitable Discourses for Professional Development*. Springer, Cham, pp. 239–256.
- Goodyear, P., Carvalho, L., 2019. The analysis of complex learning environments. In: Beetham, H., Sharpe, R. (Eds.), *Rethinking Pedagogy for a Digital Age: Principles and Practices of Design*, third ed. Routledge, Abingdon, pp. 49–65.
- Goodyear, P., Retalis, S. (Eds.), 2010. *Technology-Enhanced Learning: Design Patterns and Pattern Languages*. Sense Publishers, Rotterdam.
- Goodyear, P., Banks, S., Hodgson, V., McConnell, D. (Eds.), 2004. *Advances in Research on Networked Learning*. Kluwer Academic Publishers, Dordrecht.
- Goodyear, P., De Laat, M., Lally, V., 2006. Using pattern languages to mediate theory-praxis conversations in design for networked learning. *J. Assoc. Learn. Technol.* 14 (3), 211–223.
- Goodyear, P., Jones, C., Thompson, K., 2013. Computer-supported collaborative learning: instructional approaches, group processes and educational design. In: Spector, J.M., Merrill, D., Elen, J., Bishop, M. (Eds.), *Handbook of Research on Educational Communications and Technology*. Springer, New York, pp. 439–451.
- Goodyear, P., Carvalho, L., Dohn, N., 2016. Artefacts and activities in the analysis of learning networks. In: Ryberg, T., Sinclair, C., Bayne, S., De Laat, M. (Eds.), *Research, Boundaries and Policy in Networked Learning*. Springer, New York, pp. 93–110.
- Goodyear, P., Carvalho, L., Yeoman, P., 2021. Activity-Centred Analysis and Design (ACAD): core purposes, distinctive qualities and current developments. *Educ. Technol. Res. Dev.* 69 (2), 445–464. <https://doi.org/10.1007/s11423-020-09926-7>.
- Granovetter, M., 1973. The strength of weak ties. *Am. J. Sociol.* 78 (6), 1360–1380.
- Gunawardena, C.N., Flor, N.V., Gomez, D., Sanchez, D., 2016. Analyzing social construction of knowledge online by employing interaction analysis, learning analytics, and social network analysis. *Q. Rev. Dist. Educ.* 17 (3), 35.
- Häkkinen, P., 2013. Multiphase method for analysing online discussions. *J. Comput. Assist. Learn.* 29 (6), 547–555. <https://doi.org/10.1111/jcal.12015>.
- Hancock, D., 2005. The trouble with networks: managing the Scots' early-modern Madeira trade. *Bus. Hist. Rev.* 79 (3), 467–491.
- Hara, N., Bonk, C.J., Angeli, C., 2000. Content analysis of online discussion in an applied educational psychology course. *Instr. Sci.* 28 (2), 115–152.
- Harasim, L., Hiltz, S., Teles, L., Turoff, M., 1997. *Learning Networks: A Field Guide to Learning and Teaching Online*. MIT Press, Cambridge Mass.
- Hardy, V., Hodgson, V., McConnell, D., 1994. Computer conferencing: a new vehicle for investigating issues in gender and learning. *High Educ.* 28, 403–418.
- Henri, F., 1992. Computer conferencing and content analysis. In: Kaye, A. (Ed.), *Collaborative Learning through Computer Conferencing: The Najaden Papers*. Springer Verlag, Berlin, pp. 115–136.
- Hiltz, S., 1986. The "virtual classroom": using computer-mediated communication for university teaching. *J. Commun.* 36 (2), 95–104. <https://doi.org/10.1111/j.1460-2466.1986.tb01427.x>.
- Hodder, I., 2012. *Entangled: An Archaeology of the Relationships between Humans and Things*. Wiley-Blackwell, Chichester.
- Hodgson, V., McConnell, D., 2019. Networked learning and postdigital education. *Postdigital Sci. Educ.* 1 (1), 43–64.
- Hodgson, V., Watland, P., 2004. Researching networked management learning. *Manag. Learn.* 35 (2), 99–116. <https://doi.org/10.1177/1350507604043027>.
- Howell-Richardson, C., Mellar, H., 1996. A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instr. Sci.* 24 (1), 47–69. <https://doi.org/10.1007/bf00156003>.
- Jones, C., 2004. Networks and learning: communities, practices and the metaphor of networks. *J. Assoc. Learn. Technol.* 12 (1), 81–93. <https://doi.org/10.1080/0968776042000211548>.
- Jones, C., 2015. *Networked Learning: An Educational Paradigm for the Age of Digital Networks*. Springer, Dordrecht.
- Jones, C., Asensio, M., 2002. Designs for networked learning in higher education: a phenomenographic investigation of practitioners' accounts of design. In: Steeples, C., Jones, C. (Eds.), *Networked Learning: Perspectives and Issues*. Springer Verlag, London, pp. 253–278.
- Jones, C.R., Ferreday, D., Hodgson, V., 2008. Networked learning a relational approach: weak and strong ties. *J. Comput. Assist. Learn.* 24 (2), 90–102.
- Kelly, N., 2019. Online Networks in Teacher Education. *Oxford research encyclopedia of education*. <https://doi.org/10.1093/acrefore/9780190264093.013.416>.
- Kelly, N., Clarà, M., Kehrwald, B., Danaher, P.A., 2016. Online Learning Networks for Pre-service and Early Career Teachers. Palgrave, London.
- Kelly, N., Russell, N., Kickbusch, S., Barros, A., Dawes, L., Rasmussen, R., 2018. Online communities of teachers to support situational knowledge: a design-based study. *Australas. J. Educ. Technol.* 34 (5), 150–166. <https://doi.org/10.14742/ajet.3867>.
- Knappett, C. (Ed.), 2013. *Network Analysis in Archaeology: New Approaches to Regional Interaction*. Oxford University Press, Oxford.
- Koschmann, T. (Ed.), 1996. *CSCL: Theory and Practice of an Emerging Paradigm*. Lawrence Erlbaum Associates, Mahwah New Jersey.
- Lantz-Andersson, A., Lundin, M., Selwyn, N., 2018. Twenty years of online teacher communities: a systematic review of formally-organized and informally-developed professional learning groups. *Teach. Teach. Educ.* 75, 302–315. <https://doi.org/10.1016/j.tate.2018.07.008>.
- Le, V.H., Maor, D., McConney, A., 2021. The potential of social networked sites for continuing professional learning: investigating the experiences of teachers with limited resources. *Stud. Cont. Educ.* 1–17. <https://doi.org/10.1080/0158037X.2021.1932453>.
- Levy, P., 2003. A methodological framework for practice-based research in networked learning. *Instr. Sci.* 31 (1–2), 87–109.
- Macgillchrist, F., 2021. Theories of postdigital heterogeneity: implications for research on education and datafication. *Postdigital Sci. Educ.* <https://doi.org/10.1007/s42438-021-00232-w>.
- Macia, M., García, I., 2016. Informal online communities and networks as a source of teacher professional development. *Teach. Teach. Educ.* 55, 291–307. <https://doi.org/10.1016/j.tate.2016.01.021>.

- Manzini, E., 2015. *Design, when Everybody Designs: An Introduction to Design for Social Innovation*. MIT Press, Cambridge MA.
- Marton, F. (Ed.), 1984. *The Experience of Learning*. Scottish Academic Press, Edinburgh.
- Marton, F., Booth, S., 1997. *Learning and Awareness*. Lawrence Erlbaum Associates, Mahwah New Jersey.
- Mason, R., Kaye, A. (Eds.), 1989. *Mindweave: Communication, Computers and Distance Education*. Pergamon, Oxford.
- McConnell, D., Hodgson, V., Dirckinck-Holmfeld, L., 2012. Networked learning: a brief history and new trends. In: Dirckinck-Holmfeld, L., Hodgson, V., McConnell, D. (Eds.), *Exploring the Theory, Pedagogy and Practice of Networked Learning*. Springer, New York, pp. 3–24.
- Mor, Y., Mellar, H., Warburton, S., Winters, N., 2014. *Practical Design Patterns for Teaching and Learning with Technology*. Springer, Dordrecht.
- Networked Learning Editorial Collective (NLEC), 2021. Networked learning: inviting redefinition. *Postdigital Sci. Educ.* 3 (2), 312–325. <https://doi.org/10.1007/s42438-020-00167-8>.
- Networked Learning Editorial Collective, Gourlay, L., Rodríguez-Illera, J.L., Barberà, E., Bali, M., Gachago, D., Knox, J., 2021. Networked learning in 2021: a community definition. *Postdigital Sci. Educ.* 3 (2), 326–369. <https://doi.org/10.1007/s42438-021-00222-y>.
- O'Malley, C. (Ed.), 1995. *Computer Supported Collaborative Learning*. Springer, Berlin.
- Ouyang, F., Chang, Y.-H., 2019. The relationships between social participatory roles and cognitive engagement levels in online discussions. *Br. J. Educ. Technol.* 50 (3), 1396–1414. <https://doi.org/10.1111/bjet.12647>.
- Pietsch, T., 2013. *Empire of Scholars: Universities, Networks and the British Academic World 1850-1939*. Manchester University Press, Manchester.
- Podolny, J., Page, K., 1998. Network forms of organization. *Annu. Rev. Sociol.* 24, 57–76.
- Quinn, C., Mehan, H., Levin, J., Black, S., 1983. Real education in non-real time: the use of electronic messaging systems for instruction. *Instr. Sci.* 11 (4), 313–327.
- Rainie, L., Wellman, B., 2012. *Networked: The New Social Operating System*. MIT Press, Cambridge MA.
- Riel, M.M., Levin, J.A., 1990. Building electronic communities: success and failure in computer networking. *Instr. Sci.* 19 (2), 145–169.
- Roberts, G., 2003. Teaching using the web: conceptions and approaches from a phenomenographic perspective. *Instr. Sci.* 31 (1/2), 127–150.
- Ryberg, T., Larsen, M.C., 2008. Networked identities: understanding relationships between strong and weak ties in networked environments. *J. Comput. Assist. Learn.* 24 (2), 103–115. <https://doi.org/10.1111/j.1365-2729.2007.00272.x>.
- Salmon, G., 2000. *E-moderating: The Key to Teaching and Learning Online*. Kogan Page, London.
- Schrire, S., 2006. Knowledge building in asynchronous discussion groups: going beyond quantitative analysis. *Comput. Educ.* 46, 49–70.
- Sinclair, C., Hayes, S., 2019. Between the post and the com-post: examining the postdigital 'work' of a prefix. *Postdigital Sci. Educ.* 1 (1), 119–131. <https://doi.org/10.1007/s42438-018-0017-4>.
- Sindbæk, S., 2013. Broken links and black boxes: material affiliations and contextual network synthesis in the Viking world. In: Knappett, C. (Ed.), *Network Analysis in Archaeology: New Approaches to Regional Interaction*. Oxford University Press, Oxford, pp. 71–94.
- Steeple, C., Goodyear, P., Mellar, H., 1994. Flexible learning in higher education: the use of computer-mediated communications. *Comput. Educ.* 22, 83–90.
- Steeple, C., Unsworth, C., Bryson, M., Goodyear, P., Riding, P., Fowell, S., Duffy, C., 1996. Technological support for teaching and learning: computer-mediated communications in higher education (CMC in HE). *Comput. Educ.* 26 (1–3), 71–80.
- Thorne, K., 2003. *Blended Learning: How to Integrate Online and Traditional Learning*. Kogan Page, London.
- Trust, T., Krutka, D.G., Carpenter, J.P., 2016. "Together we are better": professional learning networks for teachers. *Comput. Educ.* 102, 15–34.
- van Amersfoort, D.I., Korenhof, M., Nijland, F., De Laat, M., Vermeulen, M., 2019. Value creation in teacher learning networks. In: Littlejohn, A., Jaldemark, J., Vrieling, E., Nijland, F. (Eds.), *Networked Professional Learning: Emerging and Equitable Discourses for Professional Development*. Springer, Cham, pp. 187–205.
- Wenger, E., Trayner, B., De Laat, M., 2011. Promoting and Assessing Value Creation in Communities and Networks: A Conceptual Framework. Retrieved from Heerlen: https://wenger-trayner.com/documents/Wenger_Trainer_DeLaat_Value_creation.pdf.

Online learning

Peter Shea, University at Albany, State University of New York, Albany, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Development and growth of online learning	423
The COVID-19 pandemic and online learning enrollment growth	424
Pedagogy of online learning	425
Emergency remote instruction	426
Blended learning	426
Massive Open Online Courses (MOOCs)	426
Research and effectiveness of online and blended learning	427
The future of online learning research	428
References	428

For the purpose of this article, online learning refers to a form of distance education which depends on information and communication technologies for its design and delivery. Online learning assumes that students and faculty engage in instruction while separated spatially, temporally or both. This distinguishes online learning from more generic descriptors such as “e-learning” or “educational technology”. Unlike older forms of distance education developed prior to the digital age and dependent on analog technologies and physical media, online learning is defined by its networked digital nature. Whereas previous systems of distance education used forms of physical correspondence including the postal service, printed course materials, physical media such as analog videotape, and one-way broadcast media such as radio and television, current forms of online learning depends on digital materials, digital networks and the internet to foster two-way, asynchronous and synchronous interaction. This distinction has significant consequences for how online learning has grown, the forms of pedagogy it enables, and its effectiveness. This article focuses on online learning as a form of course and program delivery and credentialing, mostly in higher education settings. The organizing principle for what to include in the article centers on the role of online learning in increasing access to higher education and degree programs. The article also covers aspects of blended learning that combine face-to-face and online instruction where the goal is to increase access and improve pedagogy in higher education settings. Additionally, the recent massive shift to distance education made necessary by the COVID-19 pandemic is discussed. Finally, the article also touches on Massive Open Online Courses (MOOCs) in their role to increase access to higher education.

While online learning has been implemented in pre-college settings its impact and reach has been more limited due to the custodial functions of education setting designed for younger learners. Primary and secondary schools are accountable for the students enrolled in them including establishing physical proximity, taking attendance, reporting absences etc., so forms of distance education are less compatible with their overall function. This is not to say that online learning has not been important to primary and secondary students, especially in the transition to forms of distance education as a result of the COVID-19 pandemic, but the coordinated development of online learning has primarily occurred in post-secondary settings, where issues related to access, pedagogy, students, and faculty have shaped online education.

Development and growth of online learning

While beyond the scope of this article, it should be noted that the history of distance education dates back to the 1800s or earlier and modern distance methods and scholarship to the early 1900s. Pioneers such as Charles Wedemeyer utilized radio in the 1930s and 1940s; developed course books implemented for distance learning for military personnel during World War two; devised methods for combining correspondence education with television broadcasting and team-based instruction in the 1950s; and led the way with the use of satellite-based teaching methods in the 1960s. His work was highly influential on the founding of the Open University of the UK and shaped the work of later scholars such as Otto Peters, Börje Holmberg, and Michael Graham Moore (Diehl and Cano, 2019). Distance education has a very rich history that precedes and informs digital networked-based online learning.

Early examples of online learning in the US date back to 1970s and the first online courses and programs to the 1980s (Harasim, 2000). However, significant coordination of resources to support the development of online learning for the purpose of credentialing in the United States began in 1992 with the Alfred P. Sloan Foundation's Anytime, Anyplace Learning grant program (Sloan Foundation, N.D) for the development of what were then called “asynchronous learning networks.” This grant program distributed more than \$70 million to develop software, platforms, courses, programs, and sustaining infrastructure. An overarching goal was to ensure quality and develop a community of practitioners through convening grant recipients which ultimately led to the creation of the Online Learning Consortium (OLC, formerly the Sloan Consortium). This funding and the community the Sloan Foundation established has had a lasting legacy.

OLC also developed a series of reports documenting online learning enrollment growth in the US. Starting in 2002, the numbers of students taking at least one online course from a US higher education institution were estimated to exceed 1.4 million (Allen and Seaman, 2007). Those estimates continued to climb in the US and, by the 14th version of the report in 2017, by then relying on analyses from the National Center for Education Statistics' Integrated Postsecondary Education Data System (IPEDS), online enrollments stood at more than 6.3 million (Seaman et al., 2018). In 2019 the number of students enrolled in some or all online courses was nearly 7 million (National Center for Education Statistics, 2019). Tracking these figures after 2019 has become complicated by the COVID pandemic which resulted in a much more massive shift to forms of distance education.

The growth of online learning in the US has also been characterized as proceeding through four stages (Dziuban et al., 2016). These include the 1990s and internet propelled distance education; 2000–2007 categorized by the increasing use of Learning Management Systems or LMS (specialized software for digital course design and delivery); 2008–2012 characterized by the growth of Massive Open Online Courses—MOOCs; and the period since 2012 with growth of online higher education enrollments outpacing growth in traditional higher education enrollments. The standardization made possible during earlier periods, especially the development of the LMS, set the stage for ongoing growth in later years. Use of an LMS is nearly universal in higher education (Dahlstrom et al., 2014). This format of digitalization and networked delivery represents the advanced “industrialized learning” as described by Otto Peters (2007) in that it is characterized by professional planning, formalization, standardization, automation, division of work, and centralization.

Growth in online learning globally has been more difficult to estimate. According to a report by Qayyum and Zawacki-Richter, as of 2016, enrollments for a majority of countries providing distance and open education was 23,257,100 students (Qayyum and Zawacki-Richter, 2018). This included estimates for Australia, Brazil, Canada, China, Germany, India, Russia, South Africa, South Korea, Turkey, United Kingdom and the United States. This count included over 6.4 million enrollments in China and 4.6 million enrollments in India. However, these estimates are difficult to compare across countries. In the case of India, for example, more recent survey data from September 2018 from the Indian Ministry of Education shows 3,972,068 students enrolled in distance education in all higher education institutions (Ministry of Education of India, 2019). Qayyum and Zawacki-Richter comment on this challenge, noting that it is difficult to tally exact numbers of students enrolled in open and distance education as more precise enrollment counts reflect different populations (e.g., some students enrolled in self-paced courses in which they interact only with content rather than with instructors and other students). The key takeaway is that we lack reliable comparative higher education online enrollment estimates for most countries worldwide.

Regardless of exact comparative global enrollment data, observers acknowledge that online learning has been growing in most countries over the last two decades (Qayyum and Zawacki-Richter, 2018). The US, for example, has seen continuous, though varying, levels of online enrollment growth for 19 years from 2002 to 2020. The growth of online learning in the US bucks the downward trend in overall higher education in the US (Alan and Seaman, 2017). Similar patterns are apparent in other countries such as Brazil where distance education enrollment grew by 60% between 2009 and 2014 (Litto, 2018), though it should be noted that much of the infrastructure for distance learning is not fully digital in Brazil, especially outside large cities. Another example of online enrollment growth is China where the Ministry of Education reports (2016), online enrollments in higher education in China increased from 2.37 million to 6.45 million between 2004 and 2016. This steady growth in enrollments in countries worldwide reflects student demand for flexible options in the provision of higher education.

It should be noted that the growth of online learning did not occur without controversy, especially in the US. Early critics decried the end of education (Postman, 1992) and the rise of digital diploma mills (e.g., Noble, 1998). In reality, numerous for-profit higher education institutions were investigated by the US federal government. These investigations were prompted by concerns that many for-profit institutions were recipients of significant levels of federal funding, while comparatively underperforming. For example, in 2009, students at for-profit colleges received over \$4 billion in Pell Grants and more than \$20 billion in federal loans funded by the US Department of Education (United States GAO, 2010).

In the decade leading up to 2010, enrollment at for-profit colleges grew four fold, with US grants and loans accounting for over 90% of revenue at many institutions (Shireman, 2017). For-profit colleges accounted for more than a third of all student loan defaults, while these colleges enrolled just a little more than 10% of postsecondary students. For profit institutions have been accused of widespread fraud including distorting enrollment benefits, job placements and program offerings. Investigators concluded that, in many cases, these predatory practices left students with large debts and without college credentials. Investigations, subsequent regulation, and competition from non-profit providers ultimately resulted in lower enrollment growth among for profit institutions, though regulations have been relaxed in recent years (Ivory et al., 2018).

The COVID-19 pandemic and online learning enrollment growth

The pandemic forced institutions around the world to transition to distance education with much of that transition featuring some form of online remote instruction. Approximately 1.5 billion learners of all ages across the world were affected by school and university closures owing to the COVID-19 pandemic (UNESCO, 2020; UNICEF, 2020). This shift can be seen in enrollment figures in, for example, the US where total higher education enrollments were down (from 19.7 million to 19.1 million) but the number of students taking only online courses or some online courses, rose significantly. For Fall 2020, 44.7% of students took all of their courses online, 28.1% took some of their courses online, and 27.2% took no online courses. By way of contrast, for Fall 2019, those percentages were 17.3%, 19.3%, and 63.4% (National Center for Education Statistics, 2019, 2021). The pandemic thus has

a significant impact on the number of faculty teaching in digital distance environments, the number of students experiencing some form of digital distance education, and the number of support staff assisting with these tasks.

It has been noted that the form of digital distance education implemented in response to the global COVID-19 pandemic was not the same as the predominant form of online learning developed and implemented prior to the pandemic (Hodges et al., 2020). Many have referred to the pandemic-initiated shift as “emergency remote instruction” or similar terms to distinguish pandemic-response forms of distance learning from “traditional” online learning. The long term impact of the COVID-related shift to emergency remote instruction is still unclear, but a heightened institutional priority for online learning is evident from recent surveys (e.g., CHLOE 6, 2021).

Pedagogy of online learning

While online learning enables a broad range of approaches to teaching and learning, primary differences between older forms of distance education and online learning relate to assumptions about how people learn and what digital technologies enable. Online learning development roughly coincided with a relatively recent emphasis on social dimensions of learning in education research (e.g., Hmelo-Silver, 2003; Palincsar, 1998; Vygotsky, 1978) and away from strictly behavioral and cognitive paradigms. Analog forms of distance education were more likely to depend on independent learning and less on group learning, reflecting both technological constraints and views of how people learn. Correspondence methods of instruction necessarily relied heavily on student-content and student-teacher interactions. Online learning can more easily include greater student-student interaction, consistent with socio-cultural theories of learning, as well as student interactions with content and instructors in both asynchronous and synchronous digital formats. The use of digital networks and materials facilitates content delivery, interaction, learning activities, and assessments in ways that are also less costly and cumbersome than older analog methods. So, while two-way video conferencing, and the pedagogy it enables was possible before the digital networks such approaches were not scalable and thus not as ubiquitous as digital online learning has become.

The most frequently cited framework for online pedagogy in higher education, The Community of Inquiry (CoI) model (Garrison et al., 2001), reflects this socio-cultural focus on facilitating content and group interaction in the service of promoting learning. The CoI framework assumes that teaching and learning in higher education is not simply a matter of transferring knowledge from teacher to student, but falls within the conceptual domain of “collaborative constructivist” (Garrison, 2011) pedagogical approaches. This approach to online pedagogy prioritizes the facilitation of social as well as instructional and cognitive processes. Central to the CoI approach are modified replications of the discursive processes of science and other disciplinary communities’ epistemic practices that feature the engagement with arguments with reference to various forms of evidence valued by the discipline.

The assumptions that teaching and learning are inherently interactive, rather than based in transmission and reception of information, are reflected in the terms “community” and “inquiry” which describe online teaching and learning as both dialogic and active. This is not to say that all online learning follows this model, but to indicate that online learning does differ from older forms of distance education in light of its pedagogical potential. However, it should be noted that the concepts of conversation in distance learning, a goal of the community of inquiry model, was developed and described by earlier authors such as Holmberg (1983) who advocated for supportive and congenial dialog between students and instructors. The difference between Holmberg’s earlier conception of guided didactic instruction and the CoI model is in who does the talking, when, and under what conditions.

Holmberg and other earlier distance education scholars (e.g., Keegan, 1980) assumed that the dialog would occur between the instructor and individual students via mail, telephone or another medium (Diehl and Cato, 2019), whereas CoI-guided pedagogy articulates the benefits of instructor-facilitated discussion between students and their peers, enabled by digital asynchronous threaded discussion and other information and communication technologies that promote learner-to-learner dialog.

This distinction was highlighted by early researchers of online learning such as Gunawardena and McLissac (2003),

Asynchronous or time-delayed computer conferencing has shown the capability to network groups of learners over a period of time thereby challenging Keegan’s 1980 definition that learners need to be taught as individuals rather than in groups.

Gunawardena and McLissac (2003, p. 358).

The two-way nature of online group interaction is a defining feature of online learning environments that distinguish them from analog forms of distance education. The use of asynchronous threaded discussion with groups or the whole class, for example, allows educators to design activities that go beyond earlier correspondence methods. Many other digital learning technologies can be integrated into interactive online learning including simulations, video, multimedia-based asynchronous discussion, presentations, quizzes, tests and other educational technologies.

The adoption of learning management systems led to standardization (for better or worse) that enabled this asynchronous model to expand from course offerings to program offerings. Utilizing the LMS as a container for all course materials, assignments, assessments, and learning activities was foundational to this growth (as well as some constraints on creativity some have complained, e.g., Hill, 2014). Institutions found that there was a large market for this form of interactive asynchronous online learning

leading to college degrees. In turn, offering fully online programs brought online education to the mainstream of many academic departments, rather than residing off in separate distance education units.

Emergency remote instruction

As noted above, the global pandemic forced all institutions of higher education to move to distance education, largely based on digital networks. This move was conducted in days or weeks and the typical work of designing and building online courses, a process that can take months or longer, was not carried out. The resulting courses were quite different from those employing the interactive asynchronous pedagogical model described above. Most institutions relied much more heavily on synchronous modes of instruction through video conferencing and indicate that future online program offerings would likely rely more heavily on synchronous modes (CHLOE 6, 2021, p. 20). Prior to the pandemic fully synchronous online programs were somewhat rare. They may become more common given the prevalence of synchronous approaches across a broad swath of higher education as a result of emergency remote instruction.

A recent national study (Means and Neisler, 2021) examined student responses to pandemic-initiated emergency remote instruction and found a disconcerting set of issues. These included course satisfaction levels that were significantly lower after courses moved online, frequent technical issues, student difficulty staying motivated, a lack of immediate instructor feedback, and reduced collaboration between students. In addition, non-white students and poorer students encountered more challenges than did non-Hispanic White and more affluent students. Similar issues reflecting social inequities were found across the globe (Bozkurt et al., 2020).

Blended learning

Blended learning appears to be a straightforward concept that is a mix of online and classroom instruction. However, how online and classroom instruction are combined differs significantly across varying contexts. Some argue that blended learning implies a reduction of classroom seat time and that without this reduction the result is web-enhanced instruction (Mayadas et al., 2016). Others use the term hybrid instruction to refer to mixed online and classroom instruction (Graham, 2019). Attempts have been made to calculate precise levels of online and classroom instruction that qualify for the term blended including the definitions used by Allen and Seaman (2003) that blended learning combines at least 30% and not more than 79% online instruction with classroom instruction; below 30% is considered web-enhanced, 80% or more is considered fully online.

To reduce confusion a loose but parsimonious definition of blended learning has been proposed as the *strategic* combination of online and in-person learning (Graham, 2022), leaving aside ratios of online to classroom instruction and instead focusing on modality, media, and methods. Modality in this model refers to the physical setting (online or classroom); media are the digital and non-digital tools employed, and methods are the general and content specific pedagogies used.

The strategic combination of online and in-person instruction thus articulates how these “three M’s” are put to use in the service of containing or lowering cost, and or increasing access and student learning. The active ingredient in blended learning is not the modality, nor is it the media—it is the methods. For these reasons Graham advocates that researchers investigating blended learning clearly identify what is meant by blended learning and specify the physical and pedagogical dimensions of the blended model that is being researched.

Massive Open Online Courses (MOOCs)

Massive Open Online Courses (MOOCs) extend a longstanding goal of democratizing education through distance learning (Larreamendy-Joerns and Leinhardt, 2006). The first use of the term MOOC was applied to a course offered by George Siemens and Stephen Downes, *Connectivism and Connective Knowledge*, and was coined by Dave Cormier of Prince Edward Island University in 2008 (Cormier, 2008). The course was offered to both fee paying students of the University of Manitoba and to more than 2000 students who attended the course for free. This tied MOOCs to the “open” movement in higher education that focused on making education more accessible.

The course represented what came to be known as a cMOOC (Connectivist MOOC) to distinguish it from other large, open, online courses that developed several years later and would be known as xMOOCs. The original MOOCs were founded on the learning theory of Connectivism which highlights the power of networking with other individuals, learning from diverse perspectives, and the value in the reuse and remixing of resources. In essence, connectivism is the proposition that knowledge is dispersed across a network of connections, and learning is the ability to build and navigate those networks (Downes, 2011).

The term MOOC would later be applied to other free and open online courses that were founded on a different view of learning. These xMOOCs (or extended MOOCs) were typified by the Stanford course, *Introduction Into AI*, launched by Sebastian Thrun and Peter Norvig in 2011 (Pérez-Peña, 2012). The xMOOCs, frequently offered by elite institutions, enrolled tens and sometimes hundreds of thousands of students in online versions of courses taught by the same faculty who taught in the classroom. The courses

were organized around recorded video lectures, exercises, and quizzes, a significant departure from the more free-wheeling pedagogy of cMOOCs. Consortia of MOOC providers developed including Coursera, edX, and FutureLearn.

Given that enrollment in these kinds of courses was normally limited to the small numbers of students who were accepted into Ivy League and other highly selective (and expensive) colleges, offering them for free to anyone created tremendous interest. For a period of time after 2011 it was common to find discussions of technology in the media that predicted that MOOCs were either the future of, or end of higher education. The media touted MOOC developments, frequently conflating online learning itself with MOOCs.

The key difference between MOOCs and traditional online learning is in cost, credentialing and open enrollment. While offering learners free access to online courses, MOOC providers did not initially offer college credit or degrees to students. This would change for a small number of institutions in years to come, but with the initial hype (e.g., Pappano, 2012), “traditional” online learning, offered by thousands of colleges, was overshadowed by the MOOC phenomena. As noted elsewhere in this article, prior to the emergence of MOOCs many millions of students were already studying online and earning college degrees. Many of these students graduated from open enrollment institutions such as open universities and community colleges. But traditional online degrees were, of course, not free.

Two visions animated discussion of MOOCs—the original open access view that higher education, of the highest quality, could be provided to all learners, everywhere, via the internet for free (Koller, 2012) and, alternatively, a focus on a business model that would make MOOCs either sustainable or profitable. At this point the two visions continue to inform the development of MOOCs, but the awarding of credentials remains a sticking point. Students generally have to pay to access MOOCs in which they receive official credentials, either certificates of completion or, in some cases, complete degrees. The costs, it should be noted, are considerably lower than traditional college alternatives, but the open access policy has been removed for most MOOCs that award college degrees—students must qualify for admission to these programs. As of 2021 there were 72 MOOC based degree programs, 7 of which are at the undergraduate level (Ledwon, 2021) compared to thousands of traditional online degree programs, and tens of thousands of classroom-based degree programs. Most MOOC-based degrees are priced between \$10,000 and \$22,000, with some costing more (Ledwon, 2021).

Another persistent issue with MOOCs has to do with completion rates, which generally vary between 5% and 15% (Jordan, 2013). It has been long acknowledged that few students complete an entire MOOC course of study. Some argued that this was an indicator of quality or utility issues, with an early MOOC pioneer, Sebastian Thrun, stating that MOOCs were “a lousy product” (Chaffkin, 2013). Others argue that self-directed learners can gain as much as they need from a MOOC without necessarily completing an entire course and that specific learner demographics and motivations predict MOOC completion rates. Still others (e.g., Jordan, 2015) argue that MOOCs are not like traditional online courses and that comparisons are misguided. Course completion rates in MOOCs increase for those who pay for a certificate or degree. Additionally, completion rates fluctuate significantly according to course length, with courses of longer duration having lower completion rates. Dropout varies by when the course was first offered with more recent courses having higher completion rates (Jordan, 2015).

MOOCs, despite their initial hype, have clearly not replaced higher education, but neither have they been relegated to obscurity. In 2020, more than 950 universities offered more than 16,000 MOOCs to 180 million registered MOOC users, the largest growth in MOOC registrants in a single year ever (Shah, 2020).

Research and effectiveness of online and blended learning

Online learning has been the subject of thousands of studies examining how and under what conditions students learn most effectively in online settings, who learns, what content, and under what conditions. Several peer-reviewed scholarly journals are dedicated to understanding questions about learning in technology-mediated environments including online higher education. This includes journals such as the *International Review of Research in Open and Distributed Learning*, *The Internet and Higher Education*, and *Online Learning*, among many others.

In addition to the many individual studies presented in such journals, numerous meta-analytic reviews of online learning have also been undertaken on the effectiveness of distance education, online and blended learning. This section includes more recent reviews of distance education as the meta-analyses that cover this topic include online learning by the early years of the 21st century.

The emerging answer about online learning effectiveness is that there is no significant difference between online and classroom learning. For example, in 2004 a meta-analysis by Bernard and his colleagues (Bernard et al., 2004) included an analysis of achievement, attitude, and retention outcomes across 232 studies. Results indicated effect sizes of essentially zero on all three measures indicating no significant differences between online and classroom environments studied. A year later Zhao and his colleagues reviewed 51 studies examining 99 outcomes (Zhao et al., 2005) and found no significant difference between distance education and face-to-face education. The following year Tallent-Runnels and her colleagues (Tallent-Runnels et al., 2006) conducted a narrative (rather than meta-analytic) review focusing on online learning specifically. In this study the authors reported that asynchronous communication appeared to support in-depth communication (equivalent to traditional classes), that students liked to move at their own pace; that learning outcomes appeared to be the same as in traditional courses, and that students with prior training in computers were more satisfied with online courses.

Each of the meta-analyses reports wide variability within the online learning literature with some studies reporting outcomes that favor classroom or traditional settings and some studies favoring online or distance settings. It is rarely the case that the

individual studies report no significant difference between distance and face-to-face instruction. However, the differences fade when the studies are considered in the aggregate. Zhao et al. (2005) took note of this and asked why some studies find distance education students have better achievement than their counterparts in traditional classrooms and some studies find the opposite. What factors account for the differences? The authors investigated 51 of the most rigorous studies of online and distance education and examined the many factors that might account for differences. While concluding, as others had earlier, that overall there were no significant differences in the aggregate, they also found that many more studies reported distance/online learning to be superior on outcome measures. About two-thirds of studies favored distance/online learning and one-third favored classroom instruction. The authors also found that more recent studies favored distance/online learning, as did those in which there were higher levels interaction in the online condition.

In 2009 with the large growth of online learning, the US Department of Education sponsored what was to be a gold-standard, highly controlled study of web-based distance education. The overarching goal of the study (Means et al., 2010) was to provide policymakers, administrators, and educators with direction about how to implement online learning for K–12 education and teacher preparation. However, very few studies of k-12 populations existed and the purposes of the research had to expand to include higher education studies. Despite finding more than a thousand such studies, as with the Zhao study, a limited number of these met rigorous standards for quality that would allow inclusion in the meta-analysis. This study differed from previous meta-analyses in that it focused solely on web-based online learning, exclusively focused on objective measures of learning, and included only studies with random assignment or controlled quasi-experimental designs. Of the 45 studies that met inclusion criteria the authors calculated effects of online instruction for 50 contrasts or comparisons with classroom instruction.

The meta-analysis concluded that students in online conditions performed modestly better, on average, than those learning the same material through traditional face-to-face instruction (Means et al., 2010). In addition to this general finding, the authors also found that combining online and face-to-face elements through blended or hybrid instruction was even more effective, compared to purely face-to-face instruction, than was purely online instruction without some elements of blended or hybrid instruction. Further, the authors concluded that online instruction that was collaborative or instructor-directed was more effective than in conditions where online learners worked independently without much interaction with instructors or peers.

The authors warn that despite what might be construed as strong endorsement for blended learning, the studies reviewed and analyzed in this meta-analysis do not demonstrate that online learning is superior as a medium of instruction per se. Studies that showed better outcomes with blended learning had differences in the online and classroom conditions with regard to time, curriculum, and pedagogy. It was the combination of these elements, which likely included additional learning time, different content, and additional opportunities for collaboration, that resulted in better outcomes rather than some feature of the online medium (Means et al., 2010). The final conclusion was that studies contrasting blends of online and face-to-face instruction with conventional face-to-face classes, find the former to be more effective, justifying resources required to design and implement blended approaches. On its own, online learning is as effective as classroom instruction, that is, not significantly different from traditional instruction in terms of learning outcomes reported in the literature.

In summarizing the meta-analytic literature on the effectiveness of distance, online, and blended instruction compared to classroom instruction, Bernard and his colleagues indicated that the two can be equivalent in terms of student achievement (Bernard et al., 2019). These authors further argue that a small number of variables seem to matter in making a difference between forms of distance education and classroom instruction. Higher levels of interaction, whether occurring between students, students and instructors, or students and content, lead to better outcomes for distanced, online, and blended instruction, relative to classroom instruction. Further, as found by Means et al. (2010), Bernard found that blended learning has a more pronounced effect size in terms of student achievement than does pure online learning, when compared to classroom instruction. Finally, these authors argue that intentional design provides additional advantages with regard to student-student or collaborative interaction above and beyond the mere presence of collaborative opportunities made possible by collaborative tools or affordances (Bernard et al., 2019). Enabling the learning benefits associated with collaborative learning is enhanced through intentional design.

The future of online learning research

Consensus exists that additional studies comparing online and classroom instruction are not needed and that the field will only move forward if new research comparing different interventions in online and blended environments is completed (Clark, 2000; Cook, 2009; Bernard et al., 2019). According to Bernard et al. (2019) what is needed are more primary research studies that compare one distance education, online learning, or blended learning condition to another, rather than new studies comparing these to classroom instruction. Rigorous studies that are grounded in theory and which focus on different instructional methods, rather than modality, represent the most productive path forward for the field. Altering the target of research to “important questions of design, interaction, student and teacher involvement, feedback, etc is critical ...” (Bernard et al., 2019, p. 102).

References

- Allen, I.E., Seaman, J., 2003. Sizing the Opportunity: The Quality and Extent of Online Education in the United States, 2002 and 2003. Retrieved from. <http://files.eric.ed.gov/fulltext/ED530060.pdf>.

- Allen, I.E., Seaman, J., 2007. Online Nation: Five Years of Growth in Online Learning. Sloan Consortium, Needham, MA.
- Allen, I.E., Seaman, J., 2017. Digital Learning Compass: Distance Education Enrollment Report 2017. The Babson Survey Research Group, Wellesley.
- Bernard, R.M., Abrami, P.C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., et al., 2004. How does distance education compare to classroom instruction? A meta-analysis of the empirical literature. *Rev. Educ. Res.* 74, 379–439.
- Bernard, R.M., Borokhovski, E., Tamim, R.M., 2019. The state of research on distance, online, and blended learning from the perspectives of meta-analyses and qualitative systematic reviews. In: Moore, M.G., Diehl, W.C. (Eds.), *Handbook of Distance Education*, fourth ed. Routledge, New York, NY.
- Bozkurt, A., Jung, I., Xiao, J., Vladimirci, V., Schuwer, R., Egorov, G., Lambert, S., Al-Freih, M., Pete, J., Olcott Jr., D., Rodes, V., Aranciaga, I., Bali, M., Alvarez, A.J., Roberts, J., Pazurek, A., Raffaghelli, J.E., Panagiotou, N., de Coëtlogon, P., Shahadu, S., Brown, M., Asino, T.I., Tumwesige, J., Ramirez Reyes, T., Barrios Ipenza, E., Ossiannilsson, E., Bond, M., Belhamel, K., Irvine, V., Sharma, R.C., Adam, T., Janssen, B., Sklyarova, T., Olcott, N., Ambrosino, A., Lazou, C., Mocquet, B., Mano, M., Paskevicius, M., 2020. A global outlook to the interruption of education due to COVID-19 pandemic: navigating in a time of uncertainty and crisis. *Asian J. Dist. Educ.* 15 (1), 1–126. Retrieved from <http://www.asianjde.com/ojs/index.php/AsianJDE/article/view/462>.
- Chafkin, M., 2013. Udacity's Sebastian Thrun, Godfather of Free Online Education, Changes Course. Fast Company. <https://www.fastcompany.com/3021473/udacity-sebastian-thrun-uphill-climb>.
- CHLOE 6, 2021. CHLOE 6: Online Learning Leaders Adapt for a Post-pandemic World, the Changing Landscape of Online Education, 2021. Quality Matters and Eduventures Research.
- Clark, R.E., 2000. Evaluating distance education: strategies and cautions. *Q. Rev. Dist. Educ.* 1 (1), 3–16.
- Cook, D., 2009. The failure of e-learning research to inform educational practice, and what we can do about it, 2009 *Med. Teach.* 31 (2), 158–162. <https://doi.org/10.1080/01421590802691393>.
- Cormier, D., 2008. The CCK08 MOOC—Connectivism Course, 1/4 Way. Dave's Educational Blog. <http://davecormier.com/edblog/2008/10/02/the-cck08-mooc-connectivism-course-14-way/>.
- Dahlstrom, E., Brooks, D., Bichsel, J., 2014. The Current Ecosystem of Learning Management Systems in Higher Education: Student, Faculty, and IT Perspectives: Research Report.
- Diehl, W., Cano, L., 2019. Charles Wedemeyer and the first theorists. In: Moore, M.G., Diehl, W.C. (Eds.), *Handbook of Distance Education*, fourth ed. Routledge, New York, pp. 19–31.
- Downes, S., 2011. "Connectivism" and connective knowledge. *Huffpost Educ.* https://www.huffpost.com/entry/connectivism-and-connecti_b_804653.
- Dziuban, C., Picciano, A.G., Graham, C.R., Moskal, P.D., 2016. Conducting research in online and blended learning environments: new pedagogical frontiers. Routledge, Taylor & Francis Group, New York.
- Garrison, D.R., Anderson, T., Archer, W., 2001. Critical thinking, cognitive presence, and computer conferencing in distance education. *Am. J. Dist. Educ.* 15 (1), 7–23. <https://doi.org/10.1080/08923640109527071>.
- Garrison, D.R., 2011. E-Learning in the 21st Century: A Framework for Research and Practice. Taylor & Francis, London, UK.
- Graham, C.R., 2019. Current research in blended learning. In: Moore, M.G., Diehl, W.C. (Eds.), *Handbook of Distance Education*, fourth ed. Routledge, New York, pp. 173–188.
- Graham, C., 2022. Exploring definitions, models, frameworks. In: Picciano, A., Dziuban, C., Graham, C., Moskal, P. (Eds.), *Blended Learning: Research Perspectives*, vol. 3. Routledge, New York.
- Gunawardena, C.N., McLissac, M.S., 2003. Distance education. In: Jonassen, D.H. (Ed.), *Handbook of Research for Educational Communications and Technology*. Lawrence Erlbaum, Mahwah, NJ, pp. 355–395.
- Harasim, L., 2000. Shift happens: online education as a new paradigm in learning. *Internet High. Educ.* 3 (1–2), 41–61.
- Hill, P., 2014. Opening up the LMS walled garden. *Eliterate Blog.* <https://eliterate.us/opening-lms-walled-garden/>.
- Hmelo-Silver, C.E., 2003. Analyzing collaborative knowledge construction: multiple methods for integrated understanding. *Comput. Educ.* 41, 397–420.
- Hodges, C., Moore, S., Lockee, B., Trust, T., Bond, A., 2020. The difference between emergency remote teaching and online learning. *Educ. Rev.* 27. <https://er.educase.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>.
- Holmberg, B., 1983. Guided didactic conversation in distance education. In: *Distance Education: International Perspectives*. Croom Helm, London, pp. 114–122.
- Ivory, D., Green, E., Eder, S., 2018. Education department unwinds unit investigating fraud at for-profits. *N. Y. Times.* <https://www.nytimes.com/2018/05/13/business/education-department-for-profit-colleges.html>.
- Jordan, K., 2013. MOOC Completion Rates: the Data. In: *Massive Open Online Courses (MOOCs) Database.* <http://www.katyjordan.com/MOOCproject.html>.
- Jordan, K., 2015. Massive open online course completion rates revisited assessment, length and attrition *International Review of Research. Open Dist. Learn.* 16 (3), 341–358.
- Keegan, D., 1980. On defining distance education. *Dist. Educ.* 1 (1), 13–36.
- Koller, D., 2012. What we're learning from online education. In: TED Talk. June 2012. https://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.
- Larremendy-Joens, J., Leinhardt, G., 2006. Going the distance with online education. *Rev. Educ. Res.* 76 (4), 567–605.
- Ledwon, H., 2021. [2021] 70+ Legit Master's Degrees You Can Now Earn Completely Online. The Report by Class Central. <https://www.classcentral.com/report/mooc-based-masters-degree/>.
- Litto, F., 2018. Brazil. In: Qayyum, A., Zawacki-Richter, O. (Eds.), *Open and Distance Education in Australia, Europe and the Americas: National Perspectives in a Digital Age*, vol. 1. Springer, Singapore.
- Mayadas, F., Miller, G., Sener, J., 2016. Updated E-Learning Definitions. Online Learning Consortium. <https://onlinelearningconsortium.org/updated-e-learning-definitions-2/>.
- Means, B., Neisler, J., 2021. Teaching and learning in the time of COVID: the student perspective. *Online Learn.* 25 (1), 8–27. <https://doi.org/10.24059/olj.v25i1.2496>.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., Jones, K., 2010. Evaluation of Evidence-Based Practices in Online Learning: A Meta-analysis and Review of Online Learning Studies. U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Washington, DC.
- Ministry of Education of China, 2016. Education Statistics 2016. Retrieved from http://www.moe.edu.cn/s78/A03/moe_560/jytsj_2016/.
- Ministry of Education of India, 2019. All India Survey on Higher Education (AISHE). Ministry of Human Resource Development. Department of Higher Education. <https://bit.ly/3BgZGTO>.
- National Center for Education Statistics, 2019. Number and Percentage Distribution of Students Enrolled at Title IV Institutions, by Control of Institution, Student Level, Level of Institution, Distance Education Status of Student, and Distance Education Status of Institution. United States, fall 2018. <https://bit.ly/3uqnPVH>.
- National Center for Education Statistics, 2021. Number and Percentage Distribution of Students Enrolled at Title IV Institutions, by Control of Institution, Student Level, Level of Institution, Distance Education Status of Student, and Distance Education Status of Institution. United States, fall 2020. <https://bit.ly/3F8Uyng>.
- Noble, D., 1998. Digital Diploma Mills: The Automation of Higher Education. Monthly Review Press, New York, NY.
- Palincsar, A.S., 1998. Social constructivist perspectives on teaching and learning. *Annu. Rev. Psychol.* 45, 345–375.
- Pappano, L., 2012. The year of the MOOC. *N. Y. Times.* <https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html>.
- Pérez-Peña, R., 2012. Top universities test the online appeal of free. *N. Y. Times.* <https://www.nytimes.com/2012/07/18/education/top-universities-test-the-online-appeal-of-free.html>.
- Peters, O., 2007. The most industrialized form of education. In: Moore, M.G. (Ed.), *Handbook of Distance Education*, second ed. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 57–68.
- Postman, N., 1992. The End of Education: Redefining the Value of School. Random House Digital, Inc, New York, NY.
- Qayyum, A., Zawacki-Richter, O. (Eds.), 2018. Open and Distance Education in Australia, Europe and the Americas: National Perspectives in a Digital Age, vol. 1. Springer, Singapore. <https://www.springer.com/la/book/9789811302978>.
- Seaman, J.E., Allen, I.E., Seaman, J., 2018. Grade Increase: Tracking Distance Education in the United States. The Babson Survey Research Group, Wellesley.

- Shah, D., 2020. The Second Year of the MOOC: A Review of MOOC Stats and Trends in 2020. The Report by Class Central. <https://www.classcentral.com/report/the-second-year-of-the-mooc/>.
- Shireman, R., 2017. The For-Profit College Story: Scandal, Regulate, Forget, Repeat. <https://tcf.org/content/report/profit-college-story-scandal-regulate-forget-repeat/>.
- Sloan Foundation, (N.D.). <https://sloan.org/programs/completed-programs/anytime-anyplace-learning#strategy>.
- Tallent-Runnels, M.K., Thomas, J.A., Lan, W.Y., Cooper, S., Ahern, T.C., Shaw, S.M., Liu, X., 2006. Teaching courses online: a review of research. *Rev. Educ. Res.* 76, 93–135.
- UNESCO, 2020. COVID-19 Education Response. <https://en.unesco.org/covid19/educationresponse/globalcoalition>.
- UNICEF, 2020. UNICEF and Microsoft Launch Global Learning Platform to Help Address COVID-19 Education Crisis. <https://www.unicef.org/press-releases/unicef-and-microsoft-launch-global-learning-platform-help-address-covid-19-education>.
- United States Government Accountability Office (GAO), 2010. For-Profit Colleges: Undercover Testing Finds Colleges Encouraged Fraud and Engaged in Deceptive and Questionable Marketing Practices. In: Highlights of GAO-10-948T, a Testimony Before the Committee on Health, Education, Labor, and Pensions, U.S. Senate. <http://www.gao.gov/products/GAO-10-948T>.
- Vygotsky, L.S., 1978. *Mind in society*, Fourteenth Edition. Harvard University Press, Cambridge, MA.
- Zhao, Y., Lei, J., Yan, B., Lai, C., Tan, H.S., 2005. What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teach. Coll. Rec.* 107 (8), 1836–1884.

Social media and learning

Christine M. Greenhow^a, Sarah M. Galvin^a, K. Bret Staudt Willet^b, and Amy L. Chapman^c, ^a Michigan State University, East Lansing, MI, United States; ^b Florida State University, Tallahassee, FL, United States; and ^c Columbia University, Chester, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to social media and learning	431
Social media and learning	432
Definitions and brief history	432
Themes in the research on social media and learning	433
Literacies for a changing world	433
Features and descriptions of social media literacies	434
Social media composition as literacy	434
Affordances and constraints of social media as spaces for literacies	434
Social media facilitate identity exploration and self-expression	434
Social media offer connections to large, interactive audiences	435
Emerging issues in social media literacies	435
Civic learning with social media	435
Civic engagement on social media	435
Understanding civic learning and engagement on social media	436
The goals of civic education	436
How civics is frequently taught	436
What is missing from civic education today?	436
Affordances and constraints of social media for civic engagement	436
Emerging issues in education	437
Networked professional learning spanning social media	437
Definitions	437
Understanding networked professional learning on social media	437
Affordances and constraints of social media as spaces for networked professional learning	438
Emerging issues in networked professional learning	438
Social media and learning: challenges and opportunities for the future of education	439
References	439

Glossary

Gifs A *gif* is an image, or series of images for an animated gif, saved in the graphics interchange format. This format compresses files to allow for the quick sending and sharing of images. Gifs are commonly used to send, post, or share images (or short video clips) via various social media, messaging or texting platforms, and emailing. Images of key moments from pop culture, movies, or tv shows are often turned into gifs and used to express humor or represent the sharer's emotional state or response when posting online. For example, after seeing a colleague's tweet about a recent accomplishment at work, a friend reply by posting a gif of an audience applauding or characters from a well-known tv show cheering

Subreddit A *subreddit* is one of many distinct, theme-based discussion forums hosted on Reddit's platform ([reddit.com](https://www.reddit.com)), typically denoted by a "r/" in the title of the forum. For example, "r/Teachers" is a discussion forum on Reddit organized and moderated by teachers, providing a space for teachers to ask questions and share stories related to their profession

Introduction to social media and learning

Approximately half the world's population is now using social media (Statista, 2021a,b), a trend that accelerated during the COVID-19 pandemic (Koeze and Popper, 2020). Human learning has moved increasingly online as individuals now have more choices about when, where, how and with whom they learn (Greenhow et al., 2020a,b; Greenhow and Lewin, 2021). Researchers in various disciplines have documented the rising use of social media for educational purposes (Rehm et al., 2019). Indeed, scholars around the world have theorized and examined social media's varied and complex impacts on student and teacher learning within and out of schooling (Greenhow et al., 2009). For instance, social media such as social networking sites and microblogging can expand learning opportunities (Manca and Ranieri, 2013, 2016); enhance students' well-being, identity, and literacies (Greenhow et al., 2011; Greenhow and

Gleason, 2012; Greenhow and Robelia, 2009a,b; Marich et al., 2021); and deepen student–teacher relations (Gao et al., 2012). The integration of social media in education can foster students’ active learning, collaboration skills and connections with communities beyond the classroom (Greenhow et al., 2020a,b, 2019a). Social media are also impacting the learning and communication of teacher-scholars, including how scholarship is reviewed, disseminated or taken up by various audiences (e.g., Greenhow et al., 2017; Greenhow and Gleason, 2014; Greenhow et al., 2019b,c; Li and Greenhow, 2015).

However, social media also pose ethical challenges and risks; marketing efforts present social media algorithms to the public as if they are objective and value-neutral tools. In essence, algorithms reflect the values of their builders who hold positions of power (Hrastinski et al., 2019). Whenever people create algorithms, they also create a set of data that represents society’s historical and systemic biases, which ultimately transform into algorithmic bias (Krutka et al., 2019a,b); we can see various gender and racial biases in different social media platforms (Stahl and Wright, 2018). Thus, ethical issues such as privacy, surveillance, autonomy, bias, and discrimination must be considered when social media are implicated in learning, and educators will face different ethical concerns and challenges depending on their students’ grade and age of development (Akgun and Greenhow, 2021; Greenhow et al., 2016).

Given the possibilities and perils of social media and learning today, in this article we draw from published relevant research and literature reviews of over 15 years of scholarship to provide a brief overview of the state-of-the-field. Although critical engagement with all relevant topics related to social media and learning is outside the scope of this article, we highlight three topics on the cutting edge of research and practice: (1) *literacies* for a changing world; (2) *civic engagement* on social media; and (3) *professional learning networks* spanning social media.

Overall, our goal is to catalyze and contribute to global conversations about the diverse purposes, practices, and policies of our educational future, spanning its formal-informal learning continuum and with the internet and social media playing a pivotal role (Greenhow and Lewin, 2016). The COVID-19 global health emergency upended educational systems as countries were forced to rapidly switch to online, blended, and remote forms of learning. Nations watched as a temporary solution to the problem of delivering education during the COVID-19 pandemic played out on the internet: “The primary objective was to quickly provide temporary, reliable access to instruction and support during a crisis, not to re-create a robust educational ecosystem” (Hodges et al., 2020, para. 13). As the world moves forward, beyond the pandemic, we have the opportunity to rethink and restructure educational ecosystems and advocate for visionary scholarship. This article is intended as a step in this direction.

Social media and learning

Definitions and brief history

In 2019, Merriam-Webster defined *social media* as “forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos).” Hallmarks of social media are that they are *Web 2.0 Internet-based* applications that feature *user-generated content*, *profiles* for the site or app created by users, and the development of *online social networks* by connecting a user’s profile with those of other individuals or groups within the system (Obar and Wildman, 2015). This last hallmark feature: the facilitation and development of online social networks is more prominent in some social media types than others, such as in social network sites like Facebook which spotlight people’s list of connections and feature the ability to view and traverse the connections of others. Some scholars have originated terms like *educational social networking* to characterize the combination of the networking features on social media with education-based goals or content (Galvin and Greenhow, 2020). Educational social networking occurs when teachers, students, or others gain or create knowledge together through social networking activities (e.g., teachers using #EdChat to ask questions and gain support from peers (Greenhow et al., 2021)). Next, we identify central themes and conceptualizations related to social media in education.

Although social media did not become mainstream until the mid 2000s, scholars generally agree that the first social media site arose in 1997 with the creation of *Sixdegrees.com*, which allowed users to: (1) create profiles, (2) list a set of other users with whom they shared a connection, and (3) view and traverse others’ connections (Boyd and Ellison, 2007). Although these features existed in various forms independently, *Sixdegrees.com* was the first to combine them (Boyd and Ellison, 2007). The late 1990s and early 2000s saw the rise of blogging, on platforms such as Blogger and LiveJournal, which featured the composition and maintenance of multimedia web articles posted in reverse chronological order and often having images and links to other websites. Around this same time, social media called wikis, used for collaborative web-authoring on platforms such as wikispaces, gained in popularity. Prior to 2001 wikis were virtually unknown outside closed circles of computer programmers. In the early 2000s Friendster, considered one of the first social network sites, was growing in prominence, mostly in Asia. From these beginnings there are now thousands of platforms across various types of social media. Experts generally agree that social media include 13 types with some of the most prominent types being: blogs (e.g., Blogger), wikis (e.g., Wikipedia), social network sites (e.g., Facebook), microblogs (e.g., Twitter), photo-sharing services (e.g., Instagram and Pinterest), video-sharing services (e.g., YouTube), products/services review (e.g., Yelp), social bookmarking and social gaming (Obar and Wildman, 2015). Furthermore, mobile social media are social media on mobile devices such as smartphones and tablets and typically incorporate the location of the users. According to global social media statistics reported in July 2021, approximately 57% of the world’s population use social media or 4.48 billion users, an increase of 13% over the past year (DataReportal, 2021). Almost all users (99.0%) are mobile users, spending approximately 2.5 h each day on their chosen social media. Across the globe, the 16 social media with the most active monthly users, listed in

order are: Facebook (FB), YouTube, WhatsApp, Instagram, FB Messenger, WeChat, TikTok, QQ, Douyin, Telegram, Sina Weibo, Snapchat, Kuaishou, Pinterest, Reddit, Twitter. Interestingly, of these top social media, 4 are owned by Facebook: Facebook, FB Messenger, WhatsApp, and Instagram (DataReportal, 2021). Many people are mobile, multi-platform social media users. Among young people (ages 18–29) in the United States, for instance, the most used social media platforms (as of February 2021) were YouTube (95%), Instagram (71%), Facebook (70%), Snapchat (65%), Tik Tok (48%) and Twitter (42%) (Pew Research Center, 2021).

Themes in the research on social media and learning

Participatory, constructivist and social learning frameworks have been some of the most widely used frameworks to study social media and learning (Greenhow et al., 2020a,b; Macia and Garcia, 2016). Other commonly used theoretical frameworks are Professional Learning Networks (PLNs) and Communities of Practice (CoP) theories. Recent literature reviews reveal that a large proportion of studies dealing with social media and learning address the continuum of formal-informal learning contexts (Greenhow and Lewin, 2016) and the rise of spontaneous, agentic learning possibilities for changing contexts/times (Greenhow et al., 2021, 2020a,b; Manca and Ranieri, 2016).

Existing published literature reviews related to social media in education have focused mainly on learning or teaching with a particular social media type (e.g., social network sites) (Greenhow and Askari, 2017; Rodríguez-Hoyos et al., 2015) or platform, most commonly Facebook (Aydin, 2012; Hew, 2011; Manca and Ranieri, 2013, 2016; Wilson et al., 2012) or Twitter (Alias et al., 2013; Gao et al., 2012; Tang and Hew, 2017). Most of the studies in these reviews focus on the perceptions and experiences of college students (Aydin, 2012; Manca and Ranieri, 2013) and higher education faculty (Forkosh-Baruch and Hershovitz, 2012). Moreover, the benefits of appropriating social media into teaching and learning contexts are contested in the research literature. Some studies on the integration of social network sites like Facebook in higher education suggest their affordances for collaboration, information and resource sharing (Maxman and Usluel, 2010); encouraging participation and critical thinking (Ajjan and Hartshorne, 2008; Mason and Rennie, 2006); and increased peer support and communication about course content and assessment (DiVall and Kirwin, 2012). Others warn against exploiting social media for learning because of its potentially negative impact on student outcomes such as college GPA (Junco and Cotton, 2013; Kirschner and Karpinski, 2010) or ethical dilemmas that must be understood and navigated by students and teachers (Krutka et al., 2019a,b).

In a recent literature review of over a decade of research on social media in education (i.e., studies published in 2004–2017) (Greenhow et al., 2020a,b) we focused on teaching and learning with social media in elementary and secondary level education—among learners ages 5–18. We found that this topic has been studied in 17 different countries and that such studies had been increasing in number in the last 5–7 years (Greenhow et al., 2020a,b; Rehm et al., 2019). Facebook and Twitter were the most studied platforms, with many articles covering teachers' use of social media for informal professional development (PD) (i.e., PD not structured or required by the teacher's school or district) (Greenhow et al., 2020a,b). Classroom-based studies indicated social media's potential to improve student learning experiences, such as through increased critical thinking, active learning, and engagement. However, most research did not address implementations tied to student outcomes (Greenhow et al., 2020a,b, 2019a). Challenges to teaching with social media included navigating the complications of maintaining teacher–student interactions on social media and the balance between professional boundaries and being accessible to and supportive of students. Through using social media for informal PD, teachers were able to self-direct their learning, build networks and communities that may not be available in face-to-face settings, and make connections to others that satisfied their emotional needs and alleviate isolation (Greenhow et al., 2020a,b).

Next, we highlight three topics on the cutting edge of research and practice in relation to social media: (1) *literacies* for a changing world; (2) *civic engagement* on social media; and (3) *professional learning networks* spanning social media. Each topical section begins with an overview or introduction to the topic, its significance, and what we currently know from contemporary educational research. The section continues with the affordances and constraints of social media related to that topic (i.e., literacies, civic engagement, professional learning) and concludes with emerging issues in the topic and why they are important to forward-looking research and/or practice.

Literacies for a changing world

In today's 21st-century society, people often take on the role of *creator* or *author* through various online forms and modalities. Beyond putting pen to paper, individuals go online to design and share content made of combinations of text, images, videos, and audio recordings. On social media billions of users compose content ("Most used social media 2021," 2021), especially young people (Perrin and Anderson, 2019). The wide range of skills—or *literacies* (Lankshear and Knobel, 2008)—now commonly utilized in digital spaces to create and share content suggests that conceptions of literacy are rapidly evolving alongside technological advancements.

Unsurprisingly, digital literacies, such as those used on social media, have become some of the most in-demand skills for workplace success (Adams Becker et al., 2017; Oberländer et al., 2020). According to Adams Becker et al. (2017, p. 6), university graduates surveyed indicated the largest digital literacy skills gap they faced was the ability to create, produce, and develop content with digital tools. Furthermore, the researchers found a positive correlation between digital literacy training and occupational success.

Employers' expectations about the digital literacy skills of their workers continue to increase, yet potential applicants remain under-prepared (Murray and Perez, 2014; Oberländer et al., 2020; Raish and Rimland, 2016). More attention to how social media function as spaces for important literacy practices and how these are integrated into formal education is needed.

Features and descriptions of social media literacies

As described above (see **Definitions and brief history** section), social media platforms have distinct features that distinguish them from other online spaces, including the opportunity to create an individual profile, to establish and view others' network connections (e.g., friend list), and to view and participate in a stream of user-generated content (e.g., Instagram feed) (Ellison and Boyd, 2013, p. 158). These features translate to digital literacy skills by requiring users to learn and understand how to engage with different platforms' unique elements. For example, on Facebook—the most-used social media platform globally as of July 2021 (“Most used social media 2021,” 2021)—users design a profile page, where personal information, pictures, and media can be posted. They can also create story posts: short video clips users can post for others to view for 24 h. Users can find and add Facebook Friends, creating lists of people that can view various aspects of their profile and certain posts. Posts can be shared with their entire Friend network (i.e., standard post on one's profile); they can be shared publicly but toward a particular friend (i.e., posting on a friend's profile or tagging them personally), or they can be shared privately to an individual through Facebook's direct messaging feature. Individuals can reply with text and images (e.g., memes, gifs) or react to posts by clicking a “like,” “love,” “laughing,” “amazed,” or “angry” emoji. The intricacies of how communication and interaction take place across platforms can be surprisingly complex and require knowledge not only of how to access and implement platform features but also of social norms practiced on that platform. The following section elaborates further how these social media practices are understood as a form of digital literacy.

Social media composition as literacy

The current paradigm in literacy research, New Literacy Studies (NLS), is a large umbrella topic that defines literacy as sociocultural practices (i.e., the social and cultural context of literacy use define its meaning and significance). In this way, literacies are both expressions of identity and social practices. Literacy happens differently depending on one's audience and how the author/composer seeks to present themselves to that audience (e.g., Gee, 2008). For instance, the literacies students use to present certain identities when interacting with their instructors will be different from the literacies that students choose to use when presenting different identities during interactions with their friends.

Within NLS, the *new literacies studies* movement pluralizes literacies in order to represent the inclusion of proficiencies beyond reading and writing as types of literacy practices (e.g., media literacy, information literacy, digital literacy). Digital literacies are one example of proficiencies intended to represent all proficiencies necessary for effective use of any digital tools or resources (e.g., social media). Understanding these specific proficiencies continues to challenge researchers as technology continually develops and, therefore, the savvy needed to utilize digital affordances constantly evolves (e.g., Pangrazio, 2016). It is common for social media, for example, to rise or fall in popularity quickly, meaning the literacies users need to navigate new or changing platforms must be developed just as rapidly.

Affordances and constraints of social media as spaces for literacies

Social media offer various affordances and constraints as spaces for literacies. This section highlights two of the most significant and unique aspects of social media that can function as both affordances and constraints: (1) social media as spaces for identity exploration and self-expression, and (2) social media as a connection to large, interactive audiences.

Social media facilitate identity exploration and self-expression

One important affordance of social media is their ability to facilitate creative freedom and opportunities for users to explore and express new identities. For example, Gleason (2018) followed three youth as they explored, developed, refined, and solidified their own unique feminist identities on Twitter over the course of multiple years. The participants created personalized tweeting styles, voices, and interests that reflected their beliefs. Relatedly, in Wargo's (2017, pp. 575–576) case study of LGBTQ youth, he found the participants' created content to express who they were “on their own terms,” giving them the power to control the narrative communicated to others about themselves and to “be known differently.” Social media welcomes and empowers all types of individuals to be themselves and is designed to prompt personal expression through features that allow users to customize their profile and content presentations (e.g., pictures, music, videos, other formatting features) (Gallant et al., 2007).

Although social media's emphasis on creativity affords self-expression and identity development that is often valued in formal educational settings, its openness to creative freedom can also act as a constraint that limits how effectively it can be used in school. Research suggests that youth seek out social media as outlets for their passions and interests that are not accepted elsewhere, such as in school (e.g., Bickerstaff, 2012; Greenhow et al., 2015; Padgett and Curwood, 2016). The separation of social media from school settings, which are often associated with restrictive standards and expectations, seems to add to social media's appeal. Attempts to bring social media literacies into the classroom can be problematic and confusing to students who may want to keep some of their online identities or practices private (e.g., Chromey et al., 2016; Whitney, 2011) or who are unsure how to navigate a traditionally out-of-school space within formal education structures (e.g., Galvin and Greenhow, 2019; Lantz-Andersson et al., 2013).

Social media offer connections to large, interactive audiences

Another important affordance of social media is the access platforms offer to unique audiences that are often larger and more interactive than can be found in offline settings. In school, students are taught to think of their audience as a receptive reader: someone who will read what they write but not necessarily someone who will respond and interact with them in a timely manner. Lammers and Marsh (2015) found that teen Laura used social media to share her fanfiction specifically because she sought a social audience—peers who genuinely appreciated her work and communicated with her about it. While on social media, instead of relying on social connections to those only in close geographical proximity, “people now exist as networked individuals, able to call on a large and diverse set of [online] connections when they really need them” (Humphreys, 2016, p. 161). This increase in social capital is especially valued by youth for supporting their relationships with peers (Ellison et al., 2007; Greenhow et al., 2011).

Importantly, it is not just the interactivity nor larger size of a networked audience that makes it an affordance for social media users, it is the perceived authenticity of users’ connections with their audience (e.g., Padgett and Curwood, 2016). Having an authentic audience can be motivating and inspiring; however, this type of audience can also pose challenges or even dangers to online content creators. Important restrictions schools and teachers place on social media access or allowed practices to protect students can limit the achievable audience authenticity when using social media in the classroom, reducing student interest or engagement (e.g., Galvin and Greenhow, 2019). Furthermore, users’ obsessions over audience feedback and positive reinforcement can become an addition, negatively impact their self-esteem and self-worth (e.g., Balakrishnan and Griffiths, 2017; Burrow and Rainone, 2017). While the connection to audiences available on social media can be enriching and exciting, learning to understand and mitigate risks related to sharing content online must be a priority.

Emerging issues in social media literacies

Social media literacies are ubiquitous for many people worldwide yet, as described in this section, questions and challenges remain. Considering how digital literacies used on social media may be effectively integrated into classroom learning continues to be an important issue. Bridging in-school literacies with out-of-school literacies does not mean a rejection of traditionally academic writing nor a new expectation of writing Facebook posts during class. Rather it is a notion meant to inspire a new open-mindedness and a reevaluation of which literacies are privileged in schools compared to the literacies that are most valuable and meaningful to students. Now that students are spending significant time developing literacy skills online (Anderson and Jiang, 2018), potentially more than they are during school, and the future of our 21-century job market is founded on writing and other digital literacies (Adams Becker et al., 2016), social media have emerged as spaces integral to our daily literacy practices. This raises concerns for user safety, and calls for education that prepares users to think critically about audience and the interactive nature of social media in general (Krutka et al., 2019b). To continue accessing and elevating the affordances of social media, users must know how to engage in online spaces as safe community members, not just as content creators. In the next section, we elaborate on one aspect of how to be community members on social media: civic learning and civic engagement.

Civic learning with social media

Civic education, or the teaching of young people to be informed and active citizens, must improve if it is to prepare young people to inherit democracy (EAD, 2021; Luntz, 2020). Civic education in the United States, for instance, is often misaligned with established best teaching practices or the ways in which students participate in civic life (Chapman and Greenhow, 2021). While the need for improved civic education is now being emphasized globally in response to various world events, few studies have explored how social media functions in the civic sphere and how to support young people’s civic engagement via social media. Nearly half of people aged 18–29, for instance, used social media for civic engagement in the 2020 U.S. election cycle, by creating their own civic content, describing their experiences as members of society, or by designing and sharing media (CIRCLE, 2021). This trend has prompted scholars to theorize the ways in which young people are engaging in civic life via social media, as well as how social media can support civic learning.

Civic engagement on social media

While the types of media used to create and distribute information, to engage in boycotts or #hashtag movements, or to connect with other citizens is varied, the use of social media for civic education is understudied in the educational research literature (Chapman, 2019). For example, people took to Twitter to connect to, participate in, or critique the demonstrations, protests, and uprisings known as the Arab Spring, the Occupy Wall Street movement, the Black Lives Matter movement, the #MeToo movement, and the #NeverAgain movement. In 2011 in Egypt, young people used Twitter to protest and organize activists who brought down the regime of President Hosni Mubarak (Pew, 2012). Also in 2011, young adults used Twitter to propel a suggestion in a magazine into the Occupy Wall Street movement (Preston, 2011). American high school students (ages 14–18) who survived the school shooting in Parkland, Florida in February 2018 began the #NeverAgain movement, which prompted the March for Our Lives in cities across the United States (Shear, 2018). Such hashtag movements can also be longstanding: while the #BlackLivesMatter hashtag first emerged in 2013 in response to the murder of the African American man, Trayvon Martin, its use and reach

have increased dramatically since the murder of another African American man, George Floyd, in 2020, with the hashtag being tagged on over 25 million posts and garnering 444 billion likes (Wirtschafter, 2021). Today, Twitter is actively used by young people as both a method and a tool of civic engagement.

Understanding civic learning and engagement on social media

The goals of civic education

Civic education, and the preparation of young people to be active citizens, has been an aspect of public education in many countries for centuries (Mann, 1842; Krutka, 2014). Recently, the goals of civic education in the United States were outlined in a report entitled: "The Republic is (Still) at Risk," (Levine and Kawashima-Ginsberg, 2017). The report outlined ten aspects of US civic education that should be included in order to prepare students to be active, engaged, participatory citizens:

- Classroom instruction in civics, government, history, law, economics, and geography
- Discussion of current events
- Service learning
- Extracurricular activities
- Student participation in school governance
- Simulations of democratic processes and procedures
- News media literacy
- Action civics
- Social-emotional learning (SEL)
- School climate reform

How civics is frequently taught

In spite of calls for these ten components of American civic education, and research which supports the importance of their inclusion for effective civic education, only two of these approaches (knowledge transmission and discussion) are consistently incorporated (EAD, 2021; Hanson et al., 2018). Further, the content taught by US civics teachers is narrow; most civic education focuses only on the American form of government or wars and military heroes (Lopez et al., 2006). Only 11% of students surveyed reported that civics classroom topics included discussion of contemporary problems or issues for the US today (Lopez et al., 2006). While in the last decade, some states have required discussion as part of their civics requirements, it is concerning that most civic education is focused on content acquisition rather than on developing civic skills or dispositions (Hanson et al., 2018).

What is missing from civic education today?

Despite the increased use of social media for civic engagement among young people, traditional forms of civic engagement, such as writing letters to newspapers and elected officials or joining a political party, are emphasized in formal civics education (Torney-Purta et al., 2001). While young people frequently do not engage via these approaches, they remain the focus of civic education, often relying on ineffective methods focused on skills that are infrequently used among youth (Hanson et al., 2018; Torney-Purta et al., 2001). Social media features that can facilitate new forms of civic engagement for young people are currently underexplored and undertheorized. The use of social media in civic education allows learners to construct their own knowledge in social environments that are realistic, addresses several of the areas where current civic education practices fall short of research-supported pedagogies, and taps into ways in which youth prefer to engage in civic life (Chapman, 2019; Hanson et al., 2018; Robelia et al., 2011; Torney-Purta et al., 2001).

Affordances and constraints of social media for civic engagement

Social media has numerous affordances for the development of civic literacy. In particular, social media allows users to create and share their own content; the inherently participatory nature of social media provides users with an opportunity for civic engagement (Chapman, 2019; Chapman and Greenhow, 2019; Roberson, 2018). When used in educational settings, social media can increase students' participation and engagement in content area learning (Ebner et al., 2010; Gao et al., 2012). Junco et al. (2011) further found that the use of Twitter in class enabled students who would not normally participate, to do so.

Gao et al. (2012) found that classes which used Twitter were able to connect with Twitter users beyond the classroom, who brought lived experiences to learning and added to the authenticity of the learning experience; students reported being excited by these interactions (Gao et al., 2012). Interacting with people and groups outside the classroom can facilitate students' application of in-school learning to the real world, and research has also shown that students who use Twitter are more likely to become involved in the wider community (Rinaldo et al., 2011). Students learning about civics can find people on social media they would not otherwise interact with, including government representatives, student activists, and leaders of social movements.

Another affordance of social media is that it can expand the pool of learners and instructors, as well as the roles that they play. Tweeting about topics that are brought up in class promotes interaction between students, the instructor and other Twitter users (Gao et al., 2012). While traditionally instructors are suppliers (or sometimes producers) of knowledge, and students are consumers of knowledge, using social media allows those roles to become more fluid. Students can consume information from a variety of

sources; they can also co-produce and curate knowledge in ways that are not usually seen in traditional classrooms (Gao et al., 2012) but which can promote civic engagement. When students were able to co-create and contribute information and materials, they were more likely to be active students (Dunlap and Lowenthal, 2010; Perifanou, 2009).

There are also constraints to using social media to develop civic literacy which need to be addressed. First, some teachers perceive that students do not need to learn about technology in order to use it; this has been shown to be inaccurate (Brown and Czerniewicz, 2010; Chapman, 2019). Secondly, in addition to teaching students how to use social media, teachers must teach *about* social media, particularly the ways in which users of it can be manipulated or subject to misinformation (Chapman and Greenhow, 2021).

Emerging issues in education

To address these constraints, we need to understand how teachers can better support students' civic participation via social media through civic education. Students need to learn how to use social media for civic participation and how social media platforms work. The very ways in which social media platforms are designed can create echo chambers, spread misinformation and manipulate users (Chapman and Greenhow, 2021). The future of civics teaching with social media must include preparing students to combat these more nefarious aspects of these platforms. Only then will students be prepared to use the affordances of social media for civic participation.

Networked professional learning spanning social media

As the field of education continually changes and evolves through various reform movements and innovations, education professionals must adapt in kind. This necessitates ongoing professional learning among educators that cannot be fully met by formal education (e.g., teacher preparation program, graduate programs in education) or traditional professional development offerings (e.g., district workshops and in-service days). Increasingly, educators have turned to digital technologies to connect with each other, gain knowledge and secure resources needed to address challenges in their professional contexts. Social media platforms, in particular, have created new opportunities for educators' ongoing professional learning (Trust and Prestridge, 2021). However, the number of social media platforms continues to increase, leaving educators to navigate an increasingly complex landscape of professional learning.

Definitions

Networked learning can be understood as a shift in emphasis from purely internal aspects of learning from an educational psychology perspective, to including external factors (e.g., social influences, environment) made possible through the internet (Natriello, 2016). In other words, the internet has allowed people to build and sustain networks of interpersonal connections that "increasingly and more obviously contribute to the learning opportunities of individuals in postindustrial societies" (Natriello, 2016, p. 337). Of particular interest in this article is *networked professional learning*—specifically, educators' learning supported by the novel social and environmental conditions created by social media.

Understanding networked professional learning on social media

Scholars have framed networked professional learning in numerous ways, including broad, general conceptions such as "grassroots professional development" (Forte et al., 2012) and "open online learning" (Haythornthwaite et al., 2018). Others have emphasized learning theories of learning, such as by focusing on knowledge acquisition through socialized knowledge communities (Hu et al., 2018) or networks of practice (Brown and Duguid, 2017; Ranieri et al., 2012).

The framing of networked professional learning is important because each framework emphasizes different aspects (Staudt Willet et al., 2017) of what is occurring when educators connect with each other on social media for professional purposes. Here, we consider four perspectives on networked professional learning that offer distinct views: putting the social media platform at the center, placing the individual learner at the center, centering the group, and taking a system-wide view.

The *affinity spaces* framework focuses attention on the social media platform. Affinity spaces can be understood as locations—including virtual ones—where shared interests bring people together (Gee, 2004). The common affinity is reflected in the content of the space and in how users of the space interact with each other. Affinity spaces are a general and neutral approach to understanding social media, and this framework has been used by a broad range of social media researchers. For instance, affinity spaces are a way of understanding the more than 900 distinct education-related Twitter hashtags (Participate Inc., 2019), which include #Edchat for broad education conversation (Staudt Willet, 2019; Xing and Gao, 2018), #michED for educational concerns in the US state of Michigan (Greenhalgh et al., 2020), #NGSSchat for discussions of science teaching standards (Rosenberg et al., 2020), and #SocialMediaSyllabus and #MarginalSyllabus for collaboratively developing teaching syllabi (Greenhow et al., 2019b). Affinity spaces have also been used to describe the distinct, teaching-related discussion forums (i.e., subreddits) comprising the Reddit Education Network (Carpenter and Staudt Willet, 2021; Staudt Willet and Carpenter, 2020, 2021), such as r/Teachers, r/education, and r/teachingresources.

In contrast to a spatial frame, *professional learning networks* (PLN) focus attention on the individual learner. PLNs can be understood as a support system of interconnected tools, resources, people, and spaces that an educator builds around themselves to help improve their professional practice (Krutka et al., 2017; Trust et al., 2016; Trust and Prestridge, 2021). An educator's PLN can span both online and local contexts. For instance, a new teacher may consider their PLN to include peer and mentor teachers in the school where they teach, a professional learning community organized by the school district for teachers of the same subject, a major professor from their teacher preparation program, a Facebook group for teachers, and a handful of Twitter accounts and hashtags to follow.

The *communities of practice* framework focuses attention on a group of people as a site for learning, rather than the individual learner. The emphasis of communities of practice is on learning as a form of social participation (Wenger, 1998), an apprenticeship model where learning occurs as people become members of the group (Lave and Wenger, 1991). Communities of practice are marked by mutual engagement (i.e., reciprocity), joint enterprise, and a shared repertoire of artifacts or concepts that are collected or created by the group (Wenger, 1998). Past research has examined identity formation (Kimmons and Veletsianos, 2014) and sustained participation (Xing and Gao, 2018) in communities of practice hosted on social media platforms. Social media may be especially useful sites for potential participants to observe the norms of behavior and membership before engaging in more active and public ways (Bozkurt et al., 2020)—what Lave and Wenger (1991) termed legitimate peripheral participation en route to membership in the community of practice.

Finally, the *learning ecology* framework takes a system-wide view of learning to understand “the set of contexts found in physical or virtual spaces that provide opportunities for learning” (Barron, 2006, p. 195). An ecological perspective on networked professional learning is concerned with how all the pieces fit together to make a whole (Stevenson et al., 2019; Veletsianos et al., 2019). For instance, an educator may ask teaching-related questions in the r/Teachers subreddit (i.e., one of many distinct, theme-based discussion forums hosted on Reddit's platform) and learn about educational policy matters in the r/education subreddit (Carpenter and Staudt Willet, 2021; Staudt Willet and Carpenter, 2021) as a self-directed follow up to issues raised in a professional development workshop.

Affordances and constraints of social media as spaces for networked professional learning

Past research has shown that educators regularly use social media to connect with people beyond their local school, often around the world (Beach, 2017; Macià and García, 2016; Trust, 2012). Educators have voiced their appreciation for the ways social media frees them from temporal and geographical constraints, offering perspectives more diverse than would be available locally (Hur and Brush, 2009; Prestridge, 2017; Trust et al., 2016). Networked professional learning is particularly helpful for educators struggling with professional isolation (Carpenter and Krutka, 2014; Trust and Horrocks, 2017), such as those working in a small elementary school with only one teacher per grade level (Hur and Brush, 2009) or teaching a specialized content area like computer science (Yadav et al., 2016). In addition, Rehm and Notten (2016) found that educators continued to form new social connections as they spent more time on the social media platform Twitter. Other examples of social media affordances include new teachers developing a mentor network through Twitter (Smith Risser, 2013) and educators finding allies for justice-oriented teaching through social media movements like #EduColor (Bartell et al., 2019).

However, with these affordances come constraints that must be considered. For instance, although social media can connect educators to more voices offering purported expertise, these many voices present new challenges of reconciling conflicting messages about what and how to teach (Thompson et al., 2013). The ease of creating and disseminating content through social media also means that opportunities for networked professional learning can be lost in large amounts of spam (i.e., undesired or unrelated) messages (Carpenter et al., 2020). In addition, because of the sheer volume of resources available through social media, educators may have difficulty finding content that is useful and relevant. Problematic resources are also disseminated online alongside legitimate ones (Sawyer et al., 2019), and the global scope of resources raises a local-global tension where content needs to be adapted for a specific educational context (Jones and Preece, 2006). Finally, there are social constraints related to networked professional learning. For instance, past studies have reported a prevalence of self-serving behaviors among educators on social media (e.g., Prestridge, 2019; Staudt Willet, 2019), and Lantz-Andersson et al. (2018) speculated whether social media participation fosters collaboration or competition. Additionally, increased connectivity also increases expectations that educators should always be available and accessible to students (Fox and Bird, 2017; Selwyn et al., 2017).

Emerging issues in networked professional learning

Following the affordances and constraints of using social media for networked professional learning, there are numerous issues still in need of investigation. We conclude this section naming three issues, organized by the group of stakeholders, including teachers, administrators, and researchers. First, teachers need to continue to learn about the potential technoethical issues raised by seeking professional learning on social media. Krutka et al. (2019a) suggested six questions to answer when considering the use of a technology for teaching and learning:

1. Was this technology designed ethically and is it used ethically?
2. Are laws that apply to our use of this technology just?
3. Does this technology afford or constrain democracy and justice for all people and groups?

4. Are the ways the developers profit from this technology ethical?
5. What are the unintended and unobvious problems to which this technology might contribute?
6. In what ways does this technology afford and constrain learning opportunities about technologies?

Teachers and teacher educators should work to answer these questions with social media in mind when considering engaging in or recommended networked professional learning. Second, administrators and education leaders should be mindful of the identity-agency—that is, taking responsibility for and actively investing in one’s own self-development as a professional (Ruohotie-Lyhty and Moate, 2016)—required to initiate and sustain networked professional learning, which often occurs outside of traditional work hours. Past research has characterized this in a positive light, as a type of informal learning. Future studies should take a more critical approach (Selwyn, 2017; Watters, 2014) to explore networked professional learning as a new form of invisible labor. Through all these considerations, it is evident that networked professional learning spanning social media offers numerous new possibilities but also raises significant new perils.

Social media and learning: challenges and opportunities for the future of education

Looking across the sections above, we conclude with two challenges and opportunities for educational researchers and practitioners interested in the future of social media and learning in educational contexts. First, we have yet to fully understand the *role of social media in the classroom*, especially the multidimensional ways in which social media can be used effectively in classroom learning and associated challenges that must be navigated (e.g., ethical issues such as students and parents’ privacy concerns, forms of bias and discrimination propagated on social media). Whether considering how social media composition fits into literacy instruction or implementing social media activities as opportunities for relevant civic learning and engagement, further research is needed. Social media are recognized as important spaces where students spend significant time exploring, growing, and connecting with their communities outside of school, but the practices and pedagogies necessary to embed social media (or education about social media) into formal classroom settings are still being developed.

Second, we have yet to fully understand the relationship between *teachers’ use of social media for their own professional learning and its impact on their pedagogy*. Teachers’ perspectives on students’ social media use (e.g., valuing student writing on social media, recognizing social media as important platforms for civic engagement) directly impacts the role these media can play classroom learning. Teachers’ own use of social media, particularly for networked professional learning, also influence how their instruction *with* or *about* social media is taken up in school. Therefore, attention to educators’ social media use and how they integrate social media as part of their classroom content is critical to supporting effective social media learning practices. Teachers have the opportunity to acknowledge both technoethical issues and the importance of critical media literacy within their own networked professional learning and pass these skills and perspectives on to their students. Relatedly, just as students are spending significant time on social media, research shows teachers are also spending significant time learning and pursuing professional development on social media. Further research to understand this phenomenon and establish the value of teachers’ time spent on networked professional learning (e.g., recognition of effort or labor required, evaluation of content quality) is needed. Just as social media cannot be ignored as spaces integral to students’ literacy development and civic engagement—despite being commonly accessed outside of the classroom—so too must social media be acknowledged and explored as integral to educators’ teaching and learning.

References

- Adams Becker, S., Pasquini, L.A., Zentner, A., Alexander, B., 2016. Digital Literacy: An NMC Horizon Project Strategic Brief. 3.3. The New Media Consortium, Austin.
- Adams Becker, S., Pasquini, L.A., Zentner, A., 2017. 2017 Digital Literacy Impact Study: An NMC Horizon Project Strategic Brief. 3.5. The New Media Consortium, Austin.
- Ajjan, H., Hartshorne, R., 2008. Investigating faculty decisions to adopt Web 2.0 technologies: theory and empirical tests. *Internet High Educ.* 11, 71–30.
- Akgun, S., Greenhow, C., 2021. Artificial intelligence in education: addressing ethical challenges in K-12 settings. *AI Ethics*. <https://doi.org/10.1007/s43681-021-00096-7>.
- Alias, N., Sabdan, M.S., Aziz, K.A., Mohammed, M., Hamidon, I.S., Jomhari, N., 2013. Research trends and issues in the studies of Twitter: a content analysis of publications in selected journals (2007–2012). *Proc. Soc. Behav. Sci.* 103, 773–780.
- Anderson, M., Jiang, J., 2018. Teens, social media & technology 2018. Pew Res. Center. Available at: <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- Aydin, S., 2012. A review of research on Facebook as an educational environment. *Educ. Technol. Res. Dev.* 60 (6), 1093–1106.
- Balakrishnan, J., Griffiths, M., 2017. Social media addiction: what is the role of content in YouTube? *J. Behav. Addict.* 6, 364–377.
- Barron, B., 2006. Interest and self-sustained learning as catalysts of development: a learning ecology perspective. *Hum. Dev.* 49 (4), 193–224. <https://doi.org/10.1159/000094368>.
- Bartell, T., Cho, C., Drake, C., Petchauer, E., Richmond, G., 2019. Teacher agency and resilience in the age of neoliberalism. *J. Teach. Educ.* 70 (4), 302–305. <https://doi.org/10.1177/0022487119865216>.
- Beach, P., 2017. Self-directed online learning: a theoretical model for understanding elementary teachers’ online learning experiences. *Teach. Teach. Educ.* 61, 60–72. <https://doi.org/10.1016/j.tate.2016.10.007>.
- Bickerstaff, S., 2012. “I am the rock goddess of lyrics”: writerly identities of adolescents returning to school. *J. Adolesc. Adult Literacy* 56, 56–66.
- Boyd, D., Ellison, N., 2007. Social network sites: definition, history, and scholarship. *J. Comput. Mediat. Commun.* 13 (1), 210–230.
- Bozkurt, A., Koutropoulos, A., Singh, L., Honeychurch, S., 2020. On lurking: multiple perspectives on lurking within an educational community. *Internet High Educ.* 44, 100709. <https://doi.org/10.1016/j.iheduc.2019.100709>.
- Brown, C., Czemiewicz, L., 2010. Debunking the “digital native”: beyond digital apartheid, towards digital democracy. *J. Comput. Assist. Learn.* 26 (5), 357–369.
- Brown, J.S., Duguid, P., 2017. *The Social Life of Information: Updated, With a New Preface*. Harvard Business Review Press, Boston, MA.

- Burrow, A.L., Rainone, N., 2017. How many likes did I get? Purpose moderates links between positive social media feedback and self-esteem. *J. Exp. Soc. Psychol.* 69, 232–236.
- Carpenter, J.P., Krutka, D.G., 2014. How and why educators use Twitter: a survey of the field. *J. Res. Technol. Educ.* 46, 414–434. <https://doi.org/10.1080/15391523.2014.925701>.
- Carpenter, J.P., Staudt Willet, K.B., 2021. The teachers' lounge and the debate hall: anonymous self-directed learning in two teaching-related subreddits. *Teach. Teach. Educ.* 104, 103371. <https://doi.org/10.1016/j.tate.2021.103371>.
- Carpenter, J.P., Staudt Willet, K.B., Koehler, M.J., Greenhalgh, S.P., 2020. Spam and educators' Twitter use: methodological considerations and challenges. *TechTrends* 64, 460–469. <https://doi.org/10.1007/s11528-019-00466-3>.
- Chapman, A., Greenhow, C., 2019. Citizen-scholars: social media and the changing nature of scholarship. *Publications* 7 (1), 11–20.
- Chapman, A.L., Greenhow, C., 2021. Applying a critical lens to teachers' use of social media for civic education. *Contemp. Issues Technol. Teach. Educ.* 21 (2). <https://citejournal.org/volume-21/issue-2-21/social-studies/applying-a-critical-lens-to-teachers-use-of-social-media-for-civic-education/>.
- Chapman, A.L., 2019. Meeting Them Where They Are: The Use of Twitter in Youth Civic Education. Michigan State University, East Lansing, Michigan.
- Chromey, K.J., Duchsherer, A., Pruett, J., Vareberg, K., 2016. Double-edged sword: social media use in the classroom. *Educ. Media Int.* 53, 1–12.
- Circle, 2021. Young people created media to uplift their voices in 2020. Center Inf. Res. Civic Learn. Engag. <https://circle.tufts.edu/latest-research/young-people-created-media-uplift-their-voices-2020>.
- DataReportal, 2021. Global Social Media Stats. Retrieved on August 30, 2021 from. <https://datareportal.com/social-media-users>.
- DiVall, M.V., Kirwin, J.L., 2012. Using Facebook to facilitate course-related discussion between students and faculty members. *Am. J. Pharmaceut. Educ.* 76 (2), 1–5.
- Dunlap, J.C., Lowenthal, P.R., 2010. Tweeting the night away: using Twitter to enhance social presence. *J. Inf. Syst. Educ.* 20 (2), 129–135.
- Ebner, M., Lienhardt, C., Rohs, M., Meyer, I., 2010. Microblogs in higher education—a chance to facilitate informal and process-oriented learning? *Comput. Educ.* 55 (1), 92–100.
- Educating for American Democracy (EAD), 2021. Educating for American Democracy: Excellence in History and Civics for All Learners. iCivics, Cambridge, MA. www.educatingforamericandemocracy.org.
- Ellison, N.B., Boyd, D.M., 2013. Sociality through social network sites. In: Dutton, W. (Ed.), *The Oxford Handbook of Internet Studies*. Oxford University Press, Oxford.
- Ellison, N.B., Steinfield, C., Lampe, C., 2007. The benefits of Facebook “friends”: social capital and college students' use of online social network sites. *J. Comput. Mediat. Commun.* 12, 1143–1168.
- Forkosh-Baruch, A., Hershovitz, A., 2012. A case study of Israeli higher-education institutes sharing scholarly information with the community via social networks. *Internet High Educ.* 15, 58–68. <https://doi.org/10.1016/j.iheduc.2011.08.003>.
- Forte, A., Humphreys, M., Park, T., 2012. Grassroots professional development: how teachers use Twitter. In: *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM)*. Association for the Advancement of Artificial Intelligence, pp. 106–113. Retrieved from. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12>.
- Fox, A., Bird, T., 2017. The challenge to professionals of using social media: teachers in England negotiating personal-professional identities. *Educ. Inf. Technol.* 22 (2), 647–675. <https://doi.org/10.1007/s10639-015-9442-0>.
- Gallant, L.M., Boone, G.M., Heap, A., 2007. Five heuristics for designing and evaluating web-based communities. *Clin. Hemorheol. and Microcirc.* 12.
- Galvin, S., Greenhow, C., 2019. Writing on social media: a review of research in the high school classroom. *TechTrends* 64, 57–69.
- Galvin, S., Greenhow, C., 2020. Educational networking: a novel discipline for improved K-12 learning based on social networks. In: Peña-Ayala, A. (Ed.), *Educational Networking: A Novel Discipline for Improved Learning Based on Social Networks*. Springer, New York.
- Gao, F., Luo, T., Zhang, K., 2012. Tweeting for learning: a critical analysis of research on microblogging in education published in 2008–2011. *Br. J. Educ. Technol.* 43 (5), 783–801.
- Gee, J.P., 2004. *Situated Language and Learning: A Critique of Traditional Schooling*. Routledge, New York, NY.
- Gee, J.P., 2008. *Social Linguistics and Literacies: Ideology in Discourses*, third ed. Routledge, London.
- Gleason, B., 2018. Thinking in hashtags: exploring teenagers' new literacies practices on Twitter. *Learn. Media Technol.* 43, 165–180.
- Greenhalgh, S.P., Rosenberg, J.M., Staudt Willet, K.B., Koehler, M.J., Akcaoglu, M., 2020. Identifying multiple learning spaces within a single teacher-focused Twitter hashtag. *Comput. Educ.* 148, 103809. <https://doi.org/10.1016/j.compedu.2020.103809>.
- Greenhow, C., Askari, A., 2017. Learning and teaching with social network sites: a decade of research in K-12 related education. *Educ. Inf. Technol.* 22 (2), 623–645.
- Greenhow, C., Gleason, B., 2012. Twitteracy: tweeting as a new literacy practice. *Educ. Forum* 76, 463–477.
- Greenhow, C., Gleason, B., 2014. Social scholarship: reconsidering scholarly practices in the age of social media. *Br. J. Educ. Technol.* 45 (3), 392–402.
- Greenhow, C., Lewin, C., 2016. Recrafting formal education: shifting the boundaries of formal and informal learning. *Learn. Media Technol.* 41 (1), 6–30. <https://doi.org/10.1080/17439884.2015.1064954>.
- Greenhow, C., Lewin, C., 2021. Online and blended learning: contexts and conditions for education in an emergency. *Br. J. Educ. Technol.* 52 (4).
- Greenhow, C., Robelia, E., 2009a. Old communication, new literacies: social network sites as social learning resources. *J. Comput. Mediat. Commun.* 14, 1130–1161.
- Greenhow, C., Robelia, E., 2009b. Informal learning and identity formation in online social networks. *Learn. Media Technol.* 34 (2), 119–140.
- Greenhow, C., Robelia, E., Hughes, J., 2009. Web 2.0 and classroom research: what path should we take now? *Educ. Res.* 38 (4), 246–259. <https://doi.org/10.3102/0013189X09336671>.
- Greenhow, C., Burton, L., Robelia, B., 2011. Help from my “Friends”: social capital in the social network sites of low-income high school students. *J. Educ. Comput. Res.* 45 (2), 223–245.
- Greenhow, C., Menzer, M., Gibbins, T., 2015. Re-thinking scientific literacy: arguing science issues in a niche Facebook application. *Comput. Hum. Behav.* 53, 593–604.
- Greenhow, C., Sonnevend, J., Agur, C. (Eds.), 2016. *Education and Social Media: Toward a Digital Future*. The MIT Press, Digital Media & Learning Series, Cambridge, MA.
- Greenhow, C., Gleason, B., Marich, H., Staudt Willet, K.B., 2017. Educating social scholars: examining novice researchers' practices with social media. *Qwerty* 12 (2), 30–45.
- Greenhow, C., Galvin, S., Staudt Willet, K.B., 2019a. What should be the role of social media in education? *Policy Insights Behav. Brain Sci.* 6 (2), 178–185.
- Greenhow, C.M., Gleason, B., Staudt Willet, K.B., 2019b. Social scholarship revisited: changing scholarly practices in the age of social media. *Br. J. Educ. Technol.* 50 (3), 987–1004. <https://doi.org/10.1111/bjet.12772>.
- Greenhow, C., Li, J., Mai, M., 2019c. From tweeting to meeting: expansive professional learning in the academic conference backchannel. *Br. J. Educ. Technol.* 50 (4), 1656–1672. <https://doi.org/10.1111/bjet.12817>.
- Greenhow, C., Galvin, S., Brandon, D., Askari, E., 2020a. A decade of research on K-12 teaching and teacher learning with social media: insights on the state of the field. *Teach. Coll. Rec.* 122 (6). <https://www.tcrecord.org>. ID Number: 23303.
- Greenhow, C.M., Lewin, C., Staudt Willet, K.B., 2020b. The educational response to COVID-19 across two countries: a critical examination of digital pedagogy adoption. *Technol. Pedagog. Educ.* 30 (1), 7–15. <https://doi.org/10.1080/1475939X.2020.1866654>.
- Greenhow, C., Staudt Willet, K.B., Galvin, S., 2021. Inquiring tweets want to know: #Edchat supports for #remoteteaching during COVID-19. *Br. J. Educ. Technol.* 52 (4). <https://doi.org/10.1111/bjet.13097>.
- Hanson, M., Levesque, E.M., Valant, E., Quintero, D., 2018. 2018 Brown Center Report on American Education: An Inventory of State Civics Requirement. The Brookings Institution, Washington, DC. <https://www.brookings.edu/research/2018-brown-center-report-on-american-education-an-inventory-of-state-civics-requirements/>.
- Haythornthwaite, C., Kumar, P., Grudz, A., Gilbert, S., Esteve del Valle, M., Paulin, D., 2018. Learning in the wild: coding for learning and practice on Reddit. *Learn. Media Technol.* 43, 219–235. <https://doi.org/10.1177/1461444816662933>.
- Hew, K.F., 2011. Students' and teachers' use of Facebook. *Comput. Hum. Behav.* 27, 662–676.
- Hodges, C., Moore, S., Locke, B., Trust, T., Bond, A., 2020. The difference between emergency remote teaching and online learning. *Educause* 1–12. Retrieved from. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>.

- Hrastinski, S., Olofsson, A.D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., Ryberg, T., Öberg, L., Fuentes, A., Gustafsson, U., Humble, N., Mozelius, P., Sundgren, M., Utterberg, M., 2019. Critical imaginaries and reflections on artificial intelligence and robots in postdigital K-12 education. *Postdigit. Sci. Educ.* 1, 427–445.
- Hu, S., Torphy, K.T., Opperman, A., Jansen, K., Lo, Y.J., 2018. What do teachers share within socialized knowledge communities: a case of Pinterest. *J. Prof. Cap. Community* 3 (2), 97–122. <https://doi.org/10.1108/JPC-11-2017-0025>.
- Humphreys, A., 2016. *Social Media*. Oxford University Press, New York.
- Hur, J.W., Brush, T.A., 2009. Teacher participation in online communities: why do teachers want to participate in self-generated online communities of K–12 teachers? *J. Res. Technol. Educ.* 41, 279–303. <https://doi.org/10.1080/15391523.2009.10782532>.
- Jones, A., Preece, J., 2006. Online communities for teachers and lifelong learners: a framework for comparing similarities and identifying differences in communities of practice and communities of interest. *Int. J. Learn. Technol.* 2 (2), 112–137. <https://doi.org/10.1504/IJLT.2006.010615>.
- Junco, R., Cotton, S.R., 2013. No A 4 U: the relationship between multitasking and academic performance. *Comput. Educ.* 59, 505–514.
- Junco, R., Heiberger, G., Loken, E., 2011. The effect of Twitter on college student engagement and grades. *J. Comput. Assist. Learn.* 27 (2), 119–132.
- Kimmons, R., Veletsianos, G., 2014. The fragmented educator 2.0: social networking sites, acceptable identity fragments, and the identity constellation. *Comput. Educ.* 72, 292–301. <https://doi.org/10.1016/j.compedu.2013.12.001>.
- Kirschner, A.P., Karpinski, A.C., 2010. Facebook and academic performance. *Comput. Hum. Behav.* 26, 1237–1245.
- Koeze, E., Popper, N., 2020. The virus changed the way we internet. *N. Y. Times*. <https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html>.
- Krutka, D.G., Carpenter, J.P., Trust, T., 2017. Enriching professional learning networks: a framework for identification, reflection, and intention. *TechTrends* 61, 246–252. <https://doi.org/10.1007/s11528-016-0141-5>.
- Krutka, D.G., Heath, M.K., Staudt Willet, K.B., 2019a. Foregrounding technoethics: toward critical perspectives in technology and teacher education. *J. Technol. Teach. Educ.* 27 (4), 555–574. <http://learntechlib.org/p/208235/>.
- Krutka, D.G., Manca, S., Galvin, S., Greenhow, C., Koehler, M.J., Askari, E., 2019b. Teaching “against” social media: confronting problems of profit in the curriculum. *Teach. Coll. Rec.* 121.
- Krutka, D., 2014. Democratic twittering: microblogging for a more participatory social studies. *Soc. Educ.* 78 (2), 86–89.
- Lammers, J.C., Marsh, V.L., 2015. Going public: an adolescent’s networked writing on Fanfiction.net. *J. Adolesc. Adult Literacy* 59, 277–285.
- Lankeshear, C., Knobel, M. (Eds.), 2008. *Digital Literacies: Concepts Policies and Practices*. Peter Lang, New York.
- Lantz-Andersson, A., Vigmo, S., Bowen, R., 2013. Crossing boundaries in Facebook: students’ framing of language learning activities as extended spaces. *Int. J. Comput. Support. Collab. Learn.* 8, 293–312.
- Lantz-Andersson, A., Lundin, M., Selwyn, N., 2018. Twenty years of online teacher communities: a systematic review of formally-organized and informally-developed professional learning groups. *Teach. Teach. Educ.* 75, 302–315. <https://doi.org/10.1016/j.tate.2018.07.008>.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Levine, P., Kawashima-Ginsberg, K., 2017. The Republic is (still) at risk-and civics is part of the solution. In: *Proceedings of Crossroads National Summit*. Jonathan M. Tisch College of Civic Life, Tufts University, Medford, MA. <https://www.civxnow.org/sites/default/files/resources/SummitWhitePaper.pdf>.
- Li, J., Greenhow, C., 2015. Scholars and social media: tweeting in the conference backchannel for professional learning. *Educ. Media Int.* 52 (1), 1–14.
- Lopez, M.H., Levine, P., Both, D., Kies, A., Kirby, E., Marcelo, K., Williams, D., 2006. *The 2006 Civic and Political Health of the Nation*. Center for Information and Research on Civic Learning and Engagement (CIRCLE), Washington, DC.
- Luntz, F., 2020. Data From a Representative Sample of 1000 Likely Voters in a Nationwide Survey of Policy Priorities. CivXNow, Cambridge, MA. <https://www.civxnow.org/sites/default/files/resources/CivXNow%20infographic%20-%20Luntz%20polling%20-%20FINAL.pdf>.
- Macià, M., García, I., 2016. Informal online communities and networks as a source of teacher professional development: a review. *Teach. Teach. Educ.* 55, 291–307. <https://doi.org/10.1016/j.tate.2016.01.021>.
- Manca, S., Ranieri, M., 2013. Is it a tool suitable for learning? A critical review of the literature on Facebook as a technology-enhanced learning environment. *J. Comput. Assist. Learn.* 29 (6), 487–504.
- Manca, S., Ranieri, M., 2016. Is Facebook still a suitable technology-enhanced learning environment? An updated critical review of the literature from 2012 to 2015. *J. Comput. Assist. Learn.* 32 (6), 503–528. <https://doi-org.proxy2.ci.msu.edu/10.1111/jcal.12154>.
- Mann, H., 1842. *An Oration, Delivered Before the Authorities of the City of Boston, July 4, 1842*. WB Fowle and N. Capen, Boston.
- Marich, H., Brandon, D., Greenhow, C., Hartman, D.K., 2021. Eight tweeters tweeting: the writing processes of second graders’ composing with social media. *Elem. Sch. J.* 122 (1). <https://doi-org.proxy1.ci.msu.edu/10.1086/715481>.
- Mason, R., Rennie, F., 2006. Using web 2.0 for learning in the community. *Internet High Educ.* 10 (3), 196–203.
- Mazman, S.G., Usuel, Y.K., 2010. Modeling educational uses of Facebook. *Comput. Educ.* 55 (2), 444–453.
- Murray, M., Perez, J., 2014. Unraveling the digital literacy paradox: how higher education fails at the fourth literacy. *Issues Inf. Sci. Inf. Technol.* 11, 85–100.
- Natriello, G., 2016. Networked learning. In: Corno, L., Anderman, E.M. (Eds.), *Handbook of Educational Psychology*, third ed. Routledge, New York, NY, pp. 337–348.
- Obar, J.A., Wildman, S., 2015. Social media definition and the governance challenge: an introduction to the special issue. *Telecommun. Pol.* 39 (9), 745–750.
- Oberländer, M., Beinicke, A., Bipp, T., 2020. Digital competencies: a review of the literature and applications in the workplace. *Comput. Educ.* 146, 103752.
- Padgett, E.R., Curwood, J.S., 2016. A figment of their imagination: adolescent poetic literacy in an online affinity space. *J. Adolesc. Adult Literacy* 59, 397–407.
- Pangrazio, L., 2016. Reconceptualizing critical digital literacy. *Discourse* 37, 163–174.
- Participate Inc, 2019. Update Regarding Twitter Chats on Participate! [Web archive]. <https://explore.participate.com/twitter-chats>.
- Perifanou, M.A., 2009. Language micro-gaming: fun and informal microblogging activities for language learning. In: Lytras, M.D., Ordonez de Pablos, P., Damiani, E., Avison, D., Naeve, A., Horner, D.G. (Eds.), *Best Practices for the Knowledge Society. Knowledge, Learning, Development and Technology for All, Communications in Computer and Information Science*, vol. 49. Springer, Berlin, pp. 1–14.
- Perrin, A., Anderson, M., 2019. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. *Pew Res. Center*. Available at. <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/>.
- Pew Research Center, 2012. The Role of Social Media in the Arab Uprisings. <http://www.journalism.org/2012/11/28/role-social-media-arab-uprisings/>.
- Pew Research Center, 2021. Social Media Use in 2021. <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>.
- Preston, J., 2011. Protesters look for ways to feed the web. *N. Y. Times*. <https://www.nytimes.com/2011/11/25/business/media/occupy-movement-focuses-on-staying-current-on-social-networks.html?mtref=www.google.com&auth=login-smartlock>.
- Prestridge, S., 2017. Conceptualizing self-generating online teacher professional development. *Technol. Pedagog. Educ.* 26 (1), 85–104. <https://doi.org/10.1080/1475939X.2016.1167113>.
- Prestridge, S., 2019. Categorizing teachers’ use of social media for their professional learning: a self-generating professional learning paradigm. *Comput. Educ.* 129, 143–158. <https://doi.org/10.1016/j.compedu.2018.11.003>.
- Raish, V., Rimland, E., 2016. Employer perceptions of critical information literacy skills and digital badges. *Coll. Res. Libr.* 77, 87–113.
- Ranieri, M., Manca, S., Fini, A., 2012. Why (and how) do teachers engage in social networks? An exploratory study of professional use of Facebook and its implications for lifelong learning. *Br. J. Educ. Technol.* 43, 754–769. <https://doi.org/10.1111/j.1467-8535.2012.01356.x>.
- Rehm, M., Notten, A., 2016. Twitter as an informal learning space for teachers? The role of social capital in Twitter conversations among teachers. *Teach. Teach. Educ.* 60, 215–223. <https://doi.org/10.1016/j.tate.2016.08.015>.
- Rehm, M., Manca, S., Brandon, D., Greenhow, C., 2019. Beyond disciplinary boundaries: mapping educational science in the discourse on social media. *Teach. Coll. Rec.* 121 (14).

- Rinaldo, S.B., Tapp, S., Laverie, D.A., 2011. Learning by tweeting: using Twitter as a pedagogical tool. *J. Market. Educ.* 33 (2), 193–203.
- Robelia, B., Greenhow, C., Burton, L., 2011. Adopting environmentally responsible behaviors: how learning within a social networking application motivated students to act for the environment. *Environ. Educ. Res.* 17 (4), 553–575.
- Roberson, R., 2018. To create media literate students, we need to start making media. *Soc. Educ.* 82 (4), 238–241.
- Rodríguez-Hoyos, C., Salmón, I.H., Fernández-Díaz, E., 2015. Research on SNS and education: the state of the art and its challenges. *Australas. J. Educ. Technol.* 31 (1), 100–111.
- Rosenberg, J.M., Reid, J.W., Dyer, E.B., Koehler, M.J., Fischer, C., McKenna, T.J., 2020. Idle chatter or compelling conversation? The potential of the social media-based #NGSSchat network for supporting science education reform efforts. *J. Res. Sci. Teach.* 57 (9), 1322–1355. <https://doi.org/10.1002/tea.21660>.
- Ruohotie-Lyhty, M., Moate, J., 2016. Who and how? Preservice teachers as active agents developing professional identities. *Teach. Teach. Educ.* 55, 318–327. <https://doi.org/10.1016/j.tate.2016.01.022>.
- Sawyer, A., Dick, L., Shapiro, E., Wismer, T., 2019. The top 500 mathematics pins: an analysis of elementary mathematics activities on Pinterest. *J. Technol. Teach. Educ.* 27 (2), 235–263. <https://www.learnlib.org/p/208241/>.
- Selwyn, N., Memolin, S., Johnson, N., 2017. High-tech, hard work: an investigation of teachers' work in the digital age. *Learn. Media Technol.* 42, 390–405. <https://doi.org/10.1080/17439884.2016.1252770>.
- Selwyn, N., 2017. *Education and Technology: Key Issues and Debates*, second ed. Bloomsbury, London, UK.
- Shear, M.D., 2018. Students lead huge rallies for gun control across the U.S. *N. Y. Times*. <https://www.nytimes.com/2018/03/24/us/politics/students-lead-huge-rallies-for-gun-control-across-the-us.html>.
- Smith Risser, H., 2013. Virtual induction: a novice teacher's use of Twitter to form an informal mentoring network. *Teach. Teach. Educ.* 35, 25–33. <https://doi.org/10.1016/j.tate.2013.05.001>.
- Stahl, B.C., Wright, D., 2018. Ethics and privacy in AI and big data: implementing responsible research and innovation. *IEEE Secur. Priv.* 16 (3), 26–33.
- Statista, 2021a. Most Used Social Media in 2021. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
- Statista, 2021b. Number of social network site users worldwide from 2017 to 2025. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>.
- Staudt Willet, K.B., Carpenter, J.P., 2020. Teachers on Reddit? Exploring contributions and interactions in four teaching-related subreddits. *J. Res. Technol. Educ.* 52 (2), 216–233. <https://doi.org/10.1080/15391523.2020.1722978>.
- Staudt Willet, K.B., Carpenter, J.P., 2021. A tale of two subreddits: change and continuity in teaching-related online spaces. *Br. J. Educ. Technol.* 52 (2), 714–733. <https://doi.org/10.1111/bjet.13051>.
- Staudt Willet, K.B., Koehler, M.J., Greenhalgh, S.P., 2017. A tweet by any other frame: comparing three theoretical frameworks for studying educator interactions on Twitter. In: Liu, L., Gibson, D. (Eds.), *Research Highlights in Technology and Teacher Education 2017*. Association for the Advancement of Computing in Education (AACE), pp. 63–70. <http://www.learnlib.org/p/180960/>.
- Staudt Willet, K.B., 2019. Revisiting how and why educators use Twitter: tweet types and purposes in #Edchat. *J. Res. Technol. Educ.* 51 (3), 273–289. <https://doi.org/10.1080/15391523.2019.1611507>.
- Stevenson, M., Bower, M., Falloon, G., Forbes, A., Hatzigianni, M., 2019. By design: professional learning ecologies to develop primary school teachers' makerspaces pedagogical capabilities. *Br. J. Educ. Technol.* 50 (3), 1260–1274. <https://doi.org/10.1111/bjet.12743>.
- Tang, Y., Hew, K.F., 2017. Using Twitter for education. Beneficial or a waste of time? *Comput. Educ.* 106, 97–118.
- Thompson, J., Windschitl, M., Braaten, M., 2013. Developing a theory of ambitious early-career teacher practice. *Am. Educ. Res. J.* 50, 574–615. <https://doi.org/10.3102/0002831213476334>.
- Torney-Purta, J., Lehmann, R., Oswald, H., Schulz, W., 2001. *Citizenship and Education in Twenty-Eight Countries: Civic Knowledge and Engagement at Age Fourteen*. IEA Secretariat, Amsterdam, The Netherlands.
- Trust, T., Horrocks, B., 2017. "I never feel alone in my classroom": teacher professional growth within a blended community of practice. *Prof. Dev. Educ.* 43 (4), 645–665. <https://doi.org/10.1080/19415257.2016.1233507>.
- Trust, T., Prestridge, S., 2021. The interplay of five elements of influence on educators' PLN actions. *Teach. Teach. Educ.* 97, 103195. <https://doi.org/10.1016/j.tate.2020.103195>.
- Trust, T., Krukka, D.G., Carpenter, J.P., 2016. "Together we are better": professional learning networks for teachers. *Comput. Educ.* 102, 15–34. <https://doi.org/10.1016/j.compedu.2016.06.007>.
- Trust, T., 2012. Professional learning networks designed for teacher learning. *J. Digit. Learn. Teacher Educ.* 28 (4), 133–138. <https://doi.org/10.1080/21532974.2012.10784693>.
- Velestianos, G., Johnson, N., Belikov, O., 2019. Academics' social media use over time is associated with individual, relational, cultural and political factors. *Br. J. Educ. Technol.* 50 (4), 1713–1728. <https://doi.org/10.1111/bjet.12788>.
- Wargo, J.M., 2017. "Every selfie tells a story ...": LGBTQ youth livestreams and new media narratives as connective identity texts. *New Media Soc.* 19, 560–578.
- Watters, A., 2014. *The Monsters of Education Technology*. Retrieved from: <http://monsters.hackeducation.com/>.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, New York, NY.
- Whitney, A.E., 2011. In search of the authentic English classroom: facing the schoolishness of school. *Engl. Educ.* 44, 51–62.
- Wilson, R.E., Gosling, S.D., Graham, L.T., 2012. A review of Facebook research in the social sciences. *Perspect. Psychol. Sci.* 7 (3), 203–220.
- Wirtschafter, V., 2021. How George Floyd Change the Conversation About BLM. The Brookings Institution, Washington, DC. <https://www.brookings.edu/techstream/how-george-floyd-changed-the-online-conversation-around-black-lives-matter/>.
- Xing, W., Gao, F., 2018. Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Comput. Educ.* 126, 388–398. <https://doi.org/10.1016/j.compedu.2018.08.010>.
- Yadav, A., Gretter, S., Hambrusch, S., Sands, P., 2016. Expanding computer science education in schools: understanding teacher experiences and challenges. *Comput. Sci. Educ.* 26 (4), 235–254. <https://doi.org/10.1080/08993408.2016.1257418>.

Learning across the lifespan: an overview

Patricia A. Alexander and Jannah Fusenig, Department of Human Development and Quantitative Methodology, University of Maryland, College Park, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	443
Cross-theoretical principles of learning	443
Learning and development	445
Learning across the lifespan	445
Conceptualizations	445
Commonalities	446
Distinctions	446
The Model of Domain Learning	447
The MDL and learning across the lifespan	447
The MDL and traditional models of expertise	448
Concluding thoughts	449
References	449

Introduction

Throughout this section of the “International Encyclopedia of Education” covering “Learning, Cognition and Development,” contributors have offered insights into learning within specific fields of study, such as cognitive and neurological processes that underlie human learning and performance, developmental characteristics of children, adolescents, and adults, and trends in cognitive, biophysical, social-emotional, and moral development. Our goal in this article is to offer a global and integrated perspective that weds learning, cognition, and development under the banner of “learning across the lifespan.”

Moreover, we posit that marked changes in any of these critical areas will invariably affect other areas of study. As we will endeavor to document, much of learning that transpires in the domains and disciplines, which form the backbone of educational systems, is inextricably intertwined with cognition and development (Alexander, 2017). As the primary source of evidence for this integrated and lifelong view of learning, we will focus on the past decades of theory and research on the Model of Domain Learning (MDL)—a multidimensional model of development in academic domains and disciplines based on systematic changes in learners’ knowledge, strategic processing, and interest (Alexander, 1997, 2003). Before delving into the MDL, we briefly discuss relevant issues pertaining to learning across the lifespan, namely:

- Cross-theoretical principles of learning;
- The relation of learning and development; and
- Models of lifelong learning.

Cross-theoretical principles of learning

Numerous theories of learning exist within the extant literature—many of which are described in this volume—including behaviorism, social-cognition, cognitive constructivism, embodied cognition, and socioculturalism. Given such a wealth of theories that all seek to explain the nature of learning, it is easy to assume that there is no common ground shared by these diverse perspectives. However, Alexander et al. (2009), whose own theoretical orientations toward learning were appreciably different, set about the task of determining whether such common ground could be identified. What emerged from their in-depth exploration were general principles about the nature of learning that are shared by many disparate theories:

- Learning is change
- Learning is inevitable, essential, and ubiquitous
- Learning can be resisted
- Learning may be disadvantageous
- Learning can be tacit and incidental as well as conscious and intentional
- Learning is framed by our humanness
- Learning refers to both a process and a product
- Learning is different at different points in time
- Learning is interactional (p. 178).

Whichever theory one espouses, there is the underlying assumption that “learning brings about change.” Whether learning is neurophysiological, social, emotional, or cultural in nature; whether it comes about because of formal instruction or just from living in the world; or whether the learning occurs in specific content domains or is more generic, it will result in change. Whether that change may be imperceptible or dramatic, learning leaves its mark, nonetheless (Mahoney et al., 2018; Whitlock et al., 2006). What diverse theories also share is the view that “learning is inevitable, essential, and ubiquitous.” This means that learning happens to us all in some form and that human existence is predicated on the changes learning brings (Bargh and Chartrand, 1999). Moreover, the opportunities for and the sources of learning exist everywhere—in the world or in the mind.

Largely unvoiced but still underlying diverse theories is the acknowledgment that humans have the capacity to “resist learning” particular ideas or experiences (Alcorn, 2013; Finelli and Borrego, 2020). Perhaps the focus of that resisted learning conflicts with individuals’ prior understandings or strongly held beliefs or it is simply too distant from what they have experienced in their lives (Murphy and Mason, 2006; Sinatra and Hofer, 2021). Whatever the reasons, individuals may close their minds or divert their attention from such ideas or experiences. Yet, learners’ resistance may not be wholly successful in that some trace of the encounter may find its way into individuals’ or groups’ thoughts or awareness. It is also the case that learning cannot be presumed to be positive in all cases. There is simply too much evidence that individuals are exposed to incorrect or misleading information frequently or they misunderstand or misperceive what they see or hear (Lombardi et al., 2016). Inaccurate or misleading information can also unknowingly distort what is learned subsequently (Vosniadou and Mason, 2012). In that way, “learning can be disadvantageous.” Further, although it is recognized that schools are institutions that are formed, in part, to facilitate learning for children, youth, and adults within a community or society, learning cannot be confined solely to classrooms or to formal education (Dabbagh and Kitsantas, 2012; Rogoff et al., 2016). As we discussed, learning can occur at any place humans move in the world. Even more to the point, learning can transpire without individuals’ awareness that it is happening and without any explicit goal on their part. Thus, learning can be both “tacit and incidental” as well as “conscious and intentional” (Javis, 2009).

Another shared premise about learning, not expressly considered but tacitly accepted by most, in many theories of learning—with the exception of embodied cognition (Lakoff and Johnson, 1980)—is that “learning is framed by our humanness.” It is easy to overlook the fact that, as humans, we enter the world with certain neurobiological capabilities and constraints that distinguish us from other living creatures and shape how and what we learn. For example, our ability to see, hear, smell, touch, our sensory faculties, afford us access to the empirical world and become the raw materials of our learning (Martin and Martin, 2011). Similarly, the configuration and maturation of our brains and the inner workings of our minds give us the means to perceive, think, reason, feel, and imagine (Baruch et al., 2019). It is also evident that variations in these human capabilities and their affordances can influence the “how” and “what” of learning. Even relatively minor problems with sensory systems, for example, can translate into differences in information access, perceptions, or interpretations.

Theories of learning are essentially explanatory frameworks meant to capture how learning unfolds—that is, the process of learning. These theories may well differ in whether they are expressly interested in what this process looks like within the mind of one person or within an entire sociocultural group or community. For instance, theories represented in this volume may set out to explain how rewards and punishments bring about change, or how individuals’ participation within social or cultural communities shapes their understanding and view of the world. Yet, the process per se is but a facet of what learning signifies. Learning also refers to the outcomes that arise from the process. For cognitivists, those products of learning may take the form of knowledge stored and organized in the mind, whereas to socioculturalists the products of importance are found in the demonstration of practices valued within the community. In this way, “learning is both a process and a product.”

As may be evident through this discussion, learning can occur at any time and any place and can be markedly affected by conditions that exist at that moment in time. However, it is important to keep in mind that learning, as a process, also unfolds over time and over many contexts. Consequently, “learning is different at different points in time.” As individuals or groups develop with the passage of time, or as the conditions and contexts of learning shift, the processes and products of learning will correspondingly manifest change (Vygotsky, 1978). Contributors to this volume have offered rich descriptions of the physical, neurocognitive, social, and emotional characteristics of those in infancy, childhood, adolescence, and early and late adulthood. The developmental differences across these age ranges have direct implications for the nature of learning and the types of learning that are regarded as “typical” (Bjorklund and Pellegrini, 2002). Similarly, important distinctions in the contexts and the content of learning are associated with these age ranges. We will delve deeper into the alignments between development and learning later in this article.

According to Alexander et al. (2009), “learning is interactional,” in that “learning engages an intermutual sequence of operations that are shaped by human culture and biology ... and by how humans act and react to a dynamically changing world” (p. 180). Such interactivity pertains not only to the person and their environment, and the person and their culture or biology, but also pertains to the interpersonal. Whether through an exchange of ideas by means of discussion and dialog (Murphy et al., 2009) or by observing and modeling the actions of others (Bandura, 1989), persons who populate the environments in which individuals live and learn are as critical to learning interactionally as the cultural artifacts and educational resources that are present in the environment (Bandura, 1989).

With this overview of the cross-theoretical principles of learning complete, we want to put forward the claim that learning, and especially learning across the lifespan, is a form of development. In the following section, we will attempt to make a compelling case to that effect.

Learning and development

In 1978, Vygotsky addressed the provocative question of the relation between learning and development. His intention was to cast doubt on the researchers and theorists who espoused what he regarded as extreme positions on this foundational but unsettled issue. At one extreme were experimental studies in development, like those conducted by Piaget, that were designed under the assumption that thinking, reasoning, and causal determinations unfolded in the mind of the individual in the absence of any external influences such as schooling or the support of others. At the other extreme was the work of psychologists like Binet, who assumed “that development is always a prerequisite for learning” (p. 35). Thus, if individuals’ minds had not matured sufficiently, instruction dealing with content or mental processes would be futile.

Vygotsky’s alternative to those extreme perspectives on learning and development was his conceptualization of the Zone of Proximal Development (ZPD). Through the ZPD, Vygotsky laid the groundwork for his claim that there were two developmental levels that had to be taken into consideration if researchers and theorists wanted to grasp the true relation between learning and development. The first level, which he labeled the “actual developmental level” (Vygotsky, 1978, p. 37), refers to those mental abilities or cognitive processes individuals demonstrate on measures or tasks that are decontextualized and that occur in the absence of feedback or support. The second level, identified as the “potential developmental level,” pertains to the mental abilities or cognitive processes individuals manifest when supports or guidance are made available. ZPD signifies the difference between these two developmental levels—actual and potential.

Consistent with Vygotsky’s contention, many contemporary theorists and researchers recognize the difference between what children, adolescents, or adults can “typically” and “independently” demonstrate on assessments or tasks, and what they may demonstrate under more favorable conditions (White and Alexander, 1986; Gelman, 1979; Goswami and Bryant, 2007). For instance, stage theories of development regard the cognitive functions required of analogical reasoning to be beyond the capabilities of young children (Case, 1972; Piaget, 1952). However, White and Alexander (1986) set out to determine whether 4-year-old children would manifest analogical reasoning when presented with analogy problems of the form A:B::C:D. These analogy problems were composed of attribute blocks common in their preschool environments and cast as a game. Similar to what Gelman (1979) and others (Goswami and Bryant, 2007; Veenman and Spaans, 2005) had reported, White and Alexander (1986) found that these young children were indeed capable of reasoning analogically when materials were familiar, the context was comfortable and supportive, and the task was engaging.

Therefore, when developmental profiles of young children, adolescents, or adults are forwarded, as they are in this volume, it is important to keep in mind that researchers are likely portraying typical capabilities, characteristics, and tendencies. These authors are not, by default, discounting the possibility that individuals might demonstrate higher levels of performance on certain tasks presented under particularly favorable conditions or with the guidance of a more knowledgeable other, as was exhibited in Vygotsky’s (1978) ZPD experiments. Further, this discussion is a reminder that learning and development are not wholly separate phenomena in humans of any age or background. Moreover, rich environments and appropriate instructional practices can promote “both” learning and development (Darling-Hammond et al., 2020). In this way, learning impacts development, just as development impacts learning. Further, this reciprocity is not only limited to cognitive capabilities but extends to many areas of human functioning, including social, emotional, and moral development.

While much of the theoretical and empirical research that addresses the interplay between learning and development has focused on very young children or students in primary through secondary school, their joint influence is evident well beyond the classroom context and well into the later stages of life. In the ensuing section, we will overview models of lifelong learning that make this continuing influence apparent.

Learning across the lifespan

For educational researchers, questions about learning are most often directed toward school-aged children and adolescents; that is, preschool through higher education. Of course, learning begins well before young children enter nursery or preschool and continues long after compulsory education is over. In effect, learning is on-going from “cradle to grave.” Similarly, the span of human development has been graphically described as “womb to tomb.” Humans are always in the process of change in terms of their learning and their development. Also, understanding the changes that unfold across the entire lifespan is crucial for those who seek to promote learning within individuals or for the enhancement of societies. Therefore, we will survey the literatures in two areas that consider such expansive changes over time; learning across the lifespan and lifelong learning.

Conceptualizations

“Learning across the lifespan” and “lifelong learning” are ideas to which many laypersons can relate; that is to say, many individuals are comfortable with what these terms generally mean. Yet, whether those laypersons could offer a very clear or precise definition of either term or discuss possible differences between them is quite another story. Even more intriguing for the purpose of this article is that there is a similar lack of conceptual clarity or specificity surrounding each of these popular concepts in the expanding scholarly literature (MacLachlan and Osborne, 2009; Schuetze and Casey, 2006). Fleming (2011) echoed this very concern in the introduction to the “Oxford handbook of lifelong learning.” Specifically, he observed that: “As lifelong learning does not have great conceptual

clarity this may explain how it has become a highly accepted concept; like apple pie, it is difficult to be against it!" (p. 32). The same could be said for learning across the lifespan (Istance, 2003; Tuijnman and van der Kamp, 1992). While there do appear to be relevant differences between these two bodies of literature that require our consideration, such as their disciplinary roots and their overarching aims, we will first identify areas of convergence. That is, what characteristics do these two critical domains of inquiry have in common?

Commonalities

The themes and foci that learning across the lifespan and lifelong learning seemingly hold in common are the following:

- Learning cannot be constrained to specific years of life or to particular environments.

As we noted at the outset of this section, learning occurs at each stage of life and happens in every space that humans occupy. It is not only cradle to grave but also in-school, out-of-school, and in the workplace (Javis, 2009; OECD, 2021). On that point, theorists and researchers in the domains of learning across the lifespan and lifelong learning concur.

- The well-being of individuals and societies encompasses many to all facets of human existence including cognitive, social, motivational, socioeconomic and biophysiological.

A theme of this article has been the multidimensionality of human learning and development. That theme applies to theory and research related to learning across the lifespan and lifelong learning. Thus, attention in these literatures is paid not only to cognitive abilities or academic learning, but also to physical and mental health, and to socioeconomic wellness (Elfert, 2015; Sheffler et al., 2021).

- The developmental changes that occur for individuals, especially those who are among the most senior members of society, are significantly tied to their well-being, independent functioning, and their contributions to their social groups, communities, and society.

Change is inevitable for all humans who populate this planet. Still, the nature of the changes associated with aging may be especially relevant to those concerned with learning across the lifespan and lifelong learning (National Academies of Science, Engineering, and Medicine [NAS] 2018; OECD, 2021). There are certain decrements commonly associated with advanced age, such as declines in physical strength and mobility, memory, or mental agility (Sheffler et al., 2021; Tuijnman and van der Kamp, 1992). Yet, those invested in learning across the lifespan and lifelong learning appreciate that there are actions and programs that ameliorate such declines (e.g., reduce cognitive decline associated with dementia) and enrich the lives of all members of society, regardless of age (Fleming, 2011; London, 2021).

- Learning across the lifespan and lifelong learning are influenced by conditions and forces that are internal and external to the individual.

Whether individuals struggle or thrive over their lifetime is dependent on a confluence of factors. Some of those factors are specific to the person, such as their mental acuity, personal interests, or acquired knowledge (Javis, 2009; Volles, 2016). Equally important, however, are conditions external to individuals which might include healthy living conditions, educational opportunities, and professional development resources (Schuetze and Casey, 2006). How these internal and external factors shift over time becomes consequential for the continued learning and development under both frameworks.

- There are actions that can be taken or policies that can be instituted to promote learning across the lifespan and lifelong learning.

Another commonality in these two literatures is their shared focus on enhancing learning in a manner that improves the lives of individuals or contributes to the betterment of society (Darling-Hammond et al., 2020; Fleming, 2011). These ends may be pursued through specific programs that offer training or support in areas of need such as professional development, counseling, or medical assistance (Brown et al., 2015). Alternatively, the avenue for improvement may take the form of competencies or standards articulated for fields of study or professional domains, or state or national policies that are meant to ensure that adequate attention or fiscal resources are made available (Takayama, 2013).

Distinctions

Beyond the aforementioned goals, emphases, or actions shared by those identified with learning across the lifespan and lifelong learning, there are important ways in which these lines of inquiry diverge. For one, works that speak to learning across the lifespan more often represent an educational, psychological, or developmental perspective. Thus, the emphasis in these writings is typically on how to promote the overall cognitive, social, emotional well-being of individuals; facilitate their ability to develop to their fullest potential; and to acquire new knowledge and skills regardless of age, background, or physical limitations.

Because of the educational orientation represented in this literature, there is also more attention paid to systematic changes in specific domains of learning or arenas of professional practice. Such a framing was evident in "How People Learn II: Learners, Contexts, and Cultures," a volume published by NAS (2018). The contributors intended this book to serve as a tool that could

“enrich discussions about research and practice in education and learning for people of all ages” (p. ix). In keeping with that goal, the contributors dedicated an entire chapter to this subject.

Although there was no definitive definition proffered, contributors to this article noted that learning across the lifespan operates principally outside of compulsory education and is motivated largely by individuals’ personal choices and life circumstances and by shifts in their cognitive, physical, and socioeconomic resources and capabilities. The rich discussion of learning across the lifespan in this article was synthesized into two major conclusions:

CONCLUSION 9-1: People continue to learn and grow throughout the life span, and their choices, motivation, and capacity for self-regulation, as well as their circumstances, influence how much and how well they learn and transfer their learning to new situations.

CONCLUSION 9-2: People learn continually through active engagement across many settings in their environments; learning that occurs outside of compulsory educational environments is a function of the learner’s motivation, interests, and opportunities. Engagement with work (especially complex work that involves both intellectual and social demands), social engagement, physical exercise, and adequate sleep are all associated with lifelong learning and healthy aging (p. 223).

In comparison to this more educational and developmental perspective, the literature on lifelong learning has a more competency-based, sociopolitical, or economic focus, as well as a greater input from those working in adult education. Key international organizations, such as the United Nations Educational, Scientific and Cultural Organization (UNESCO) and the Organization for Economic Co-operation and Development (OECD), have become the driving forces for the rising attention to lifelong learning for many countries. The rationale for this growing interest globally can be seen in the introduction to “OECD Skills Outlook 2021: Learning for Life:”

Individuals’ ability to adapt and thrive in a fast-evolving world rests on them having acquired strong foundation skills, the willingness to learn and a habit of learning (otherwise known as lifelong learning attitudes). These skills and attitudes are vital for them to absorb and expand the knowledge and skills required to navigate new labour-market needs and life circumstances. However, walking the talk of lifelong learning requires a change in mindset: from a vision of learning that is compartmentalised in different phases of life to a process that evolves during all of life (Foreword, p. 1).

The economic implications of lifelong learning and the availability of a competent workforce are central themes of this report and others issued from world organizations.

According to Volles (2016) and Bynner (2017), critical analysis of policy documents from UNESCO and OECD reveals a progression away from a more humanistic orientation to a perception that lifelong can be used “as a cure for a wide range of maladies, ranging from high unemployment to low innovation rates and a lack of entrepreneurship” (Volles, 2016, p. 343). Further, the emphasis on essential competencies and skills required for success in a knowledge-rich and technology-based economy has resulted in efforts to specify what the educational agenda should be, especially for adult populations. Technology-specific competencies and skills, not surprisingly, figure heavily among those essentials for future success.

The importance of adult education or re-education was clearly apparent in the contributions to the first Oxford handbook of lifelong learning (London, 2011). For instance, the programs, technologies, and techniques detailed in this volume were expressly targeted to the workplace, and notably to employees or executives in the corporate and business world. Those thrusts are still evident in the second edition of the handbook (London, 2021) but with more attention being given to continuing education, professional development, and to career survival shaped by technological advancements.

The Model of Domain Learning

In this closing section of the article, we want to overview the Model of Domain Learning (MDL; Alexander, 1997, 2003). We have chosen to devote space to the MDL for two reasons. First, it illustrates how one theoretical model portrays the interplay of learning and development, capturing this interplay by explicating how learning and learners systematically change over time as a result of maturation and experience (Alexander et al., 2009). Second, as discussed in the preceding section, the MDL serves as an example of learning across the lifespan. There are several reasons why the MDL fits best within the literature on learning across the lifespan rather than lifelong learning.

The MDL and learning across the lifespan

Several features of the MDL place it more squarely within learning across the lifespan. For one, the MDL describes change at the level of the person and not society (NAS, 2018). For another, the realm under consideration is educational instead of economic or sociopolitical. Moreover, the goal of the MDL is the optimization of human potential and not the redress of global issues such as migration, underemployment, or economic sustainability (Darling-Hammond et al., 2020). Also, the theoretical and empirical literature grounding the MDL comes from psychology, development, academic domains, and fields of professional practice with limited attention to adult education, professional development, or geo-economics (Alexander, in press). Finally, there is recognition of the importance placed on learners’ motivations and on their investment in self-regulation and goal

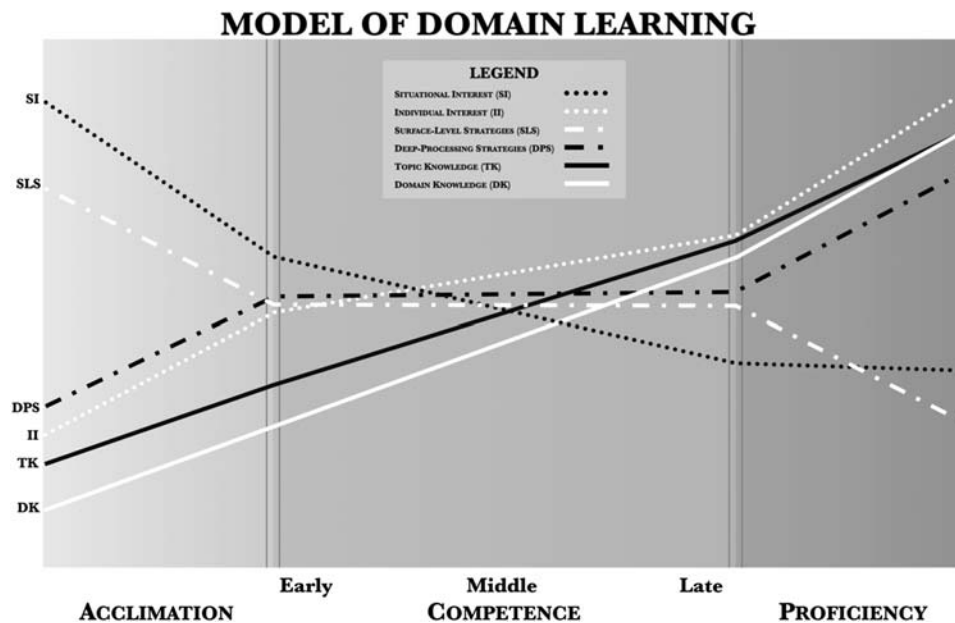


Fig. 1 The components and stages of the Model of Domain Learning.

attainment (Dabbagh and Kitsantas, 2012; Sheffler et al., 2021). At its core, the MDL seeks to answer the question of how individuals change in their knowledge, strategic processing, and their interests as they become more and more competent in a given field of study, such as mathematics, history, or computer science. In this way, the goal of mapping those changes reflects learning across the lifespan (Alexander, 2003).

Specifically, the MDL is a theoretical model of expertise development comprising three stages: acclimation, competence, and proficiency/expertise. The characterization of each stage comes from the shifting relations among individuals' breadth (domain) and depth (topic) of subject-matter knowledge, their reliance on surface-level or deep-processing strategies, and their enduring (individual) or fleeting (situational) interests (see Fig. 1). Although it is a model of expertise development, the MDL is not intended only for those seeking to become experts, particularly since only a very small percentage of individuals actually achieve expertise in any domain (Alexander, 1997). Rather, from the onset of "acclimation"—the moment of earliest exposure to a domain such as reading or science—individuals unknowingly embark on a journey that should carry them toward greater competence or capability in that domain. How long and how far students progress in this journey and, concomitantly, how competent they ultimately become, is highly dependent on both internal forces (e.g., underlying capabilities, goals, and efforts) and external factors (e.g., opportunities and resources) that arise during this developmental journey. Among the influential factors that can foster or frustrate progress in any academic domain is the quality and support students receive while engaged in formal schooling (Darling-Hammond et al., 2020). A quality education that is also well matched to learners' personal goals, sociocultural background, and specific strengths and needs can do much to propel them forward in their academic development.

The MDL and traditional models of expertise

Unlike more traditional models of expertise (Ericsson and Charness, 1994), the MDL does not simply categorize the populace as either a novice or an expert nor assume that hours of deliberate practice alone will lead to expertise. Instead, the model represents the movement toward expertise as beginning in acclimation, the stage where individuals initially become oriented to the domain in question. Once they have acquired a foundation of domain-specific knowledge, problem-solving strategies, and manifest personal interest in that field of study, learners cross over into the stage of "competence." Competent individuals operate with a growing and more integrated body of subject-matter knowledge, use both surface-level and deep-processing strategies when encountering domain problems, and form personal connections to the domain that are less reliant on features of the environment that may prove stimulating or intriguing but potentially volatile and fleeting. Competence, in the MDL, is the stage of development that most students reach for fields that form the backbone of school curricula in primary, secondary, and tertiary education.

Because of its expansive and diverse nature, competence is subdivided into early, middle, and late periods. As seen in Fig. 1, noticeable differences in individuals' knowledge, strategic processing, and interest are evident in these subdivisions. Of particular importance to learning across the lifespan is mid-competence. According to Alexander (2003, in press), it is at mid-competence that those who identify strongly with a domain and are committed to the pursuit of expertise seek out experiences that can contribute to

their continued development. These experiences will typically extend well beyond the years of compulsory and tertiary education. Those who elect to continue their journey toward expertise may succeed in reaching the proficiency or expertise stage of the MDL. As captured in Fig. 1, individuals at this stage are predicted to manifest an abiding interest in and self-identify within the domain. Further, because experts are expected to contribute new knowledge to their domain, they would necessarily rely on the deep-processing strategies required to re-envision and transform that field.

From the standpoint of learning across the lifespan, we want to emphasize that achieving expertise is not the culmination of learning and development according to the MDL. Domains of learning are never stagnant. They continue to expand or shift over time. Thus, to maintain their expertise, those in proficiency must continue to learn, grow in their understandings, and perhaps even shift their views or perspectives. Moreover, as we have suggested, the journey toward expertise can consume decades of one's life. During those ensuing years, those pursuing expertise have undergone all manners of change—cognitive, biophysical, social-emotional, or socioeconomic. Those developmental changes will naturally influence individuals' capacity and desire to meet the challenges that inevitably will arise on this lifelong journey.

Concluding thoughts

Our overarching goal in this chapter was to demonstrate how learning and development are intertwined across individuals' lifetime. By surveying the principles of learning from Alexander et al. (2009), for example, we wanted to illustrate that even theories of learning that differ in significant ways still largely adhere to these principles. We then looked more specifically at contrasting views of the relation between development and learning, turning to Vygotsky's (1978) Zone of Proximal Development as a bridge between these contrasting perspectives on learning and development. With these foundational elements in place, we wanted to examine two viewpoints on how learning may unfold from cradle to grave—learning across the lifespan and lifelong learning. Through this discussion, we sought to illustrate how these two approaches converge on certain premises and practices while diverging in other important ways. The final topic in this article was the Model of Domain Learning (Alexander, 1997, 2003). In this overview of the MDL, we sought to highlight one model that brings together learning and development by describing individuals' journeys from acclimation through competence and, potentially, to proficiency or expertise. This journey, which falls under the banner of learning across the lifespan, brings together cognitive, strategic, and motivational dimensions of learning and development. By understanding the complexity of learning across the stages of human development, individuals and those guiding their journey at each developmental stage can be better prepared to navigate the challenges that will undoubtedly be encountered and to relish the experience of a lifetime.

References

- Alcorn, M., 2013. Resistance to Learning: Overcoming the Desire Not to Know in Classroom Teaching. Springer, New York.
- Alexander, P.A., Schallert, D.L., Reynolds, R.E., 2009. What is learning anyway? A topographical perspective considered. *Educ. Psychol.* 44 (3), 209–214. <https://doi.org/10.1080/00461520903029006>.
- Alexander, P.A., 1997. Mapping the multidimensional nature of domain learning: the interplay of cognitive, motivational, and strategic forces. In: Maehr, M.L., Pintrich, P.R. (Eds.), *Advances in Motivation and Achievement*, vol. 10. JAI Press, Stamford, CT, pp. 213–250.
- Alexander, P.A., 2003. The development of expertise: the journey from acclimation to proficiency. *Educ. Res.* 32 (8), 10–14. <https://doi.org/10.3102/0013189X032008010>.
- Alexander, P.A., 2017. Relational reasoning in STEM domains: a foundation for academic development. *Educ. Psychol. Rev.* 29, 1–10. <https://doi.org/10.1007/s10648-016-9383-1>.
- Alexander, P.A. (in press). The interplay of knowledge, strategies, and interest in the development of expertise within professions. In: Glückler, J., Winch, C., Punstein, A.M. (Eds.), *Professions and Proficiency (Knowledge and Space)*, vol. 18. New York: Springer.
- Bandura, A., 1989. Human agency in social cognitive theory. *Am. Psychol.* 44 (9), 1175–1184. <https://doi.org/10.1037/0003-066X.44.9.1175>.
- Bargh, J.A., Chartrand, T.L., 1999. The unbearable automaticity of being. *Am. Psychol.* 54 (7), 462–479. <https://doi.org/10.1037/0003-066X.54.7.462>.
- Barutcu, A., Toohey, S., Shivasani, M.N., Fifer, J.M., Crewther, S.G., Grayden, D.B., Paolini, A.G., 2019. Multisensory perception and attention in school-age children. *J. Exp. Child Psychol.* 180, 141–155. <https://doi.org/10.1016/j.jecp.2018.11.021>.
- Bjorklund, D.F., Pellegrini, A.D., 2002. The Origins of Human Nature: Evolutionary Developmental Psychology. American Psychological Association, Washington, DC. <https://doi.org/10.1037/10425-000>.
- Brown, R.I., Cobigo, V., Taylor, W.D., 2015. Quality of life and social inclusion across the lifespan: challenges and recommendations. *Int. J. Dev. Disabil.* 61 (2), 93–100. <https://doi.org/10.1179/2047386914Z.000000000092>.
- Bynner, J., 2017. Whatever happened to lifelong learning? And does it matter? *J. Br. Acad.* 5, 61–89. <https://doi.org/10.5871/jba/005.061>.
- Case, R., 1972. Validation of a neo-Piagetian mental capacity construct. *J. Exp. Child Psychol.* 14 (2), 287–302. [https://doi.org/10.1016/0022-0965\(72\)90051-3](https://doi.org/10.1016/0022-0965(72)90051-3).
- Dabbagh, N., Kitsantas, A., 2012. Personal learning environments, social media, and self-regulated learning: a natural formula for connecting formal and informal learning. *Internet High Educ.* 15 (1), 3–8. <https://doi.org/10.1016/j.iheduc.2011.06.002>.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D., 2020. Implications for educational practice of the science of learning and development. *Appl. Dev. Sci.* 24 (2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>.
- Elfert, M., 2015. UNESCO, the Faure report, the Delors report, and the political utopia of lifelong learning. *Eur. J. Educ.* 50 (1), 88–100. <https://doi.org/10.1111/ejed.12104>.
- Ericsson, K.A., Charness, N., 1994. Expert performance: its structure and acquisition. *Am. Psychol.* 49 (8), 725–747. <https://doi.org/10.1037/0003-066X.49.8.725>.
- Finelli, C.J., Borrego, M., 2020. Evidence-based strategies to reduce student resistance to active learning. In: Mintzes, J., Walter, E. (Eds.), *Active Learning in College Science*. Springer, New York, pp. 943–952. https://doi.org/10.1007/978-3-030-33600-4_58.
- Fleming, T., 2011. Models of lifelong learning: an overview. In: London, M. (Ed.), *The Oxford Handbook of Lifelong Learning*. Oxford University Press, pp. 29–39.
- Gelman, R., 1979. Preschool thought. *Am. Psychol.* 34 (10), 900–905. <https://doi.org/10.1037/0003-066X.34.10.900>.

- Goswami, U., Bryant, P., 2007. Children's Cognitive Development and Learning (Primary Review Research Survey 2/1a). The Cambridge Primary Review Research Surveys, pp. 161–189.
- Istance, D., 2003. Schooling and lifelong learning: insights from OECD analyses. *Eur. J. Educ.* 38 (1), 85–98. <https://www.jstor.org/stable/1503484>.
- Javis, P., 2009. Learning from everyday life. In: Javis, P. (Ed.), *The Routledge International Handbook of Lifelong Learning*. Routledge, New York, pp. 49–60.
- Lakoff, G., Johnson, M., 1980. *Metaphors We Live By*. University of Chicago Press, Chicago.
- Lombardi, D., Nussbaum, E.M., Sinatra, G.M., 2016. Plausibility judgments in conceptual change and epistemic cognition. *Educ. Psychol.* 51 (1), 35–56. <https://doi.org/10.1080/00461520.2015.1113134>.
- London, M., 2011. *The Oxford Handbook of Lifelong Learning*. Oxford University Press.
- London, M., 2021. *The Oxford Handbook of Lifelong Learning*, second ed. Oxford University Press.
- MacLachlan, K., Osborne, M., 2009. Lifelong learning, development, knowledge and identity. *J. Comp. Int. Educ.* 39 (5), 575–583. <https://doi.org/10.1080/03057920903138506>.
- Mahoney, J.L., Durlak, J.A., Weissberg, R.P., 2018. An update on social and emotional learning outcome research. *Phi Delta Kappan* 100 (4), 18–23. <https://doi.org/10.1177/0031721718815668>.
- Martin, P., Martin, V., 2011. Learning and the senses. In: Javis, P. (Ed.), *The Routledge International Handbook of Learning*. Routledge, New York, pp. 7–17.
- Murphy, P.K., Mason, L., 2006. Changing knowledge and beliefs. In: Alexander, P.A., Winne, P.H. (Eds.), *Handbook of Educational Psychology*. Taylor & Francis, New York, pp. 305–324.
- Murphy, P.K., Wilkinson, I.A., Soter, A.O., Hennessey, M.N., Alexander, J.F., 2009. Examining the effects of classroom discussion on students' comprehension of text: a meta-analysis. *J. Educ. Psychol.* 101 (3), 740–764. <https://doi.org/10.1037/a0015576>.
- National Academies of Sciences, Engineering, and Medicine, 2018. *How people learn II: Learners, contexts, and cultures*. National Academies Press.
- OECD, 2021. *OECD Skills Outlook 2021: Learning for Life*. OECD Publishing, Paris. Accessed online. <https://www.oecd.org/education/oecd-skills-outlook-e11c1c2d-en.htm>.
- Piaget, J., 1952. *The Origins of Intelligence in Children*. International Universities Press, New York.
- Rogoff, B., Callanan, M., Gutiérrez, K.D., Erickson, F., 2016. The organization of informal learning. *Rev. Res. Educ.* 40 (1), 356–401. <https://doi.org/10.3102/0091732X16680994>.
- Schuetze, H.G., Casey, C., 2006. Models and meanings of lifelong learning: progress and barriers on the road to a learning society. *Compare* 36 (3), 279–287. <https://doi.org/10.1080/03057920600872365>.
- Sheffler, P., Rodriguez, T.M., Cheung, C.S., Wu, R., 2021. Cognitive and metacognitive, motivational, and resource considerations for learning new skills across the lifespan. *Wiley Interdiscip. Rev.* 1585, 1–26. <https://doi.org/10.1002/wcs.1585>.
- Sinatra, G., Hofer, B., 2021. *Science Denial: Why It Happens and What to Do About It*. Oxford University Press, New York.
- Takayama, K., 2013. OECD, key competencies and the new challenges of educational inequality. *J. Curric. Stud.* 45 (1), 67–80. <https://doi.org/10.1080/00220272.2012.755711>.
- Tuijnman, A.C., van der Kamp, M., 1992. *Learning Across the Lifespan: Theories, Research, Policies*. Pergamon, Oxford, UK.
- Veenman, M.V., Spaans, M.A., 2005. Relation between intellectual and metacognitive skills: age and task differences. *Learn. Individ. Differ.* 15 (2), 159–176. <https://doi.org/10.1016/j.lindif.2004.12.001>.
- Volles, N., 2016. Lifelong learning in the EU: changing conceptualizations, actors, and policies. *Stud. High Educ.* 41 (2), 343–363. <https://doi.org/10.1080/03075079.2014.927852>.
- Vosniadou, S., Mason, L., 2012. Conceptual change induced by instruction: a complex interplay of multiple factors. In: Graham, S., Harris, K., Urdan, T. (Eds.), *Educational Psychology Handbook, Individual Differences and Cultural and Contextual Factors*, vol. 2. American Psychological Association, Washington, DC, pp. 221–246.
- Vygotsky, L., 1978. Interaction between learning and development. In: Gauvin, M., Cole, M. (Eds.), *Readings on the Development of Children*. Scientific American Books, New York, pp. 34–40.
- White, C.S., Alexander, P.A., 1986. Effects of training on four-year-olds' ability to solve geometric analogy problems. *Cognit. Instruct.* 3 (3), 261–268. https://doi.org/10.1207/s1532690xci0303_6.
- Whitlock, J.R., Heynen, A.J., Shuler, M.G., Bear, M.F., 2006. Learning induces long-term potentiation in the hippocampus. *Science* 313 (5790), 1093–1097. <https://doi.org/10.1126/science.1128134>.

Learning and development in infancy: toward an integrated approach

Alecia Moser and Annette M.E. Henderson, School of Psychology, The University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	451
Learning in infancy	451
Associative learning	451
Statistical learning	452
Social learning	453
Imitation and emulation	453
Foundations of social learning	454
Selective social learning	455
Summary	457
Moving beyond separate accounts toward integrated frameworks for infant learning	457
Attention	458
Social partners and social contingency	459
Incentive value	460
Summary	461
Moving forward: research directions, considerations, and conclusions	461
References	462

Introduction

Infancy is a period marked by rapid learning about infants' physical and social worlds. With time, and experience, infants meet this formidable challenge by becoming robust, efficient, and active learners. Decades of research examining learning through infancy point to several learning mechanisms, including those that promote learned associations, statistical learning, and social learning. To date, most of the prior studies on infants' learning have investigated how different learning mechanisms facilitate learning achievements in isolation and under a high degree of experimental control. These studies have provided a wealth of information and confidence about the existence and the strength of the learning tools available to infants. Despite these important and extensive insights, many of infants' learning achievements have not been explained and comprehensive models of infants' learning across diverse domains remain elusive. Having set the foundation for the development of comprehensive frameworks, we argue that the critical next step is to consider the combined influence of multiple mechanisms and how they function together to support the rapid knowledge and skill development evident in infancy. In this contribution, we bridge the gap between seminal and recent research on three key individual mechanisms by describing approaches that have sought to integrate two or more types of learning and reflect on how integrative approaches will be key to achieving a comprehensive understanding of the mechanisms and processes that support infants' learning and development.

Learning in infancy

Associative learning

General learning mechanisms (GLMs) provide infants with the ability to learn by association from the moment that they are born. Learning by *association* occurs when learners make connections between pieces of information that are perceptually salient, co-occur either at the same time, or within a constrained temporal window. Infants' early abilities to rapidly associate the words that they hear with the appropriate aspect of the world around them (i.e., the referent) is often lauded as showing the power of learning by association (Plunkett, 1997; Smith, 2000). Within associative accounts of word learning, infants assign a novel word, such as a new object label, onto the object or aspect of their environment that is the most salient to them when they hear the word (Hollich et al., 2000; Smith et al., 2003; Werker et al., 1998). Associative learning has been argued to get infants' word learning off the ground and thus, has been linked with infants' word learning throughout the first year and a half of their lives (Byers-Heinlein et al., 2013; Colunga and Smith, 2005; Hollich et al., 2000; Smith, 2000).

Beyond word learning, associative learning within and across modalities is imperative across diverse domains such as the development of concepts (Sloutsky and Deng, 2019), object categories (Goldstone and Barsalou, 1998; Johnson et al., 2003; Rakison et al., 2008), body representations (de Klerk et al., 2021) and, even, fear (Rakison, 2022). To illustrate the role of associative mechanisms in infants' categorical learning, we draw on work examining an early emerging category distinction that infants as young as

9 months make—the distinction between animate (i.e., people, insects, animals) and inanimate (e.g., plants, vehicles, tools) things in their world (Poulin-Dubois et al., 1996). There is little doubt that perceptual aspects of objects, such as shape, color, and patterns of movement, are employed to make such category decisions (Goldstone and Barsalou, 1998; Quinn and Eimas, 2000). Yet, infants also distinguish between members of divergent categories that are perceptually similar (e.g., plane versus outstretched bird). One explanation is that infants become familiar with the co-occurring physical attributes of objects through associative learning. In the case of the plane versus outstretched bird example raised above, infants learn relevant part-level perceptual details of objects, such as vehicles often have wheels whereas animals often have legs, as they begin to appreciate the distinctions between object categories, such as animate versus inanimate objects.

While evidence clearly suggests that associative mechanisms catalyze infants' early learning, associative learning cannot account for all of the learning that infants demonstrate for two key reasons: (1) learning by association can be time intensive and (2) the world of an infant is often filled with ambiguity in which there is a vast array of incoming signals that may often be equally salient and thus, relying on associative learning as the sole learning mechanism would often produce learning errors. Within the word learning domain, the limitations of associative learning accounts are best illustrated by Quine's problem of reference (Quine, 2013; Quine and Van, 1960): for any novel word one encounters, there is often an infinite number of potential referents for the word, many of which do not always "stand out" of the realm of possibilities (see also Blythe et al., 2016; Hollich et al., 2000). Indeed, learning solely through associations would be a slow process filled with errors, which does not adequately capture infants' learning.

Statistical learning

Fortunately, infants' learning is boosted by an early-developing, remarkable ability to detect patterns across a diverse range of stimuli and contexts (for reviews see Erickson and Thiessen, 2015; Krogh et al., 2013; Romberg and Saffran, 2010; Saffran, 2020; Saffran and Kirkham, 2018; Thiessen et al., 2013). Studies focusing on infants' language learning have been critical for our understanding of how infants harness the statistical regularities in their environments for learning; *statistical learning* (Ellis et al., 2021; Saffran, 2020). The results of this work demonstrate that infants are sensitive to the statistical regularities within the languages to which they are exposed (Johnson et al., 2003), which are informative based on their predictability and redundancy (Romberg and Saffran, 2010; Tenenbaum et al., 2011). In their pioneer study, Saffran et al. (1996) demonstrated that infants as young as 8 months of age detect the statistical regularities (i.e., transitional probability of sounds) present within and between word boundaries using a familiarization-preference language perception task (Jusczyk and Aslin, 1995). During an initial familiarization period (~2 min), infants were presented with an uninterrupted stream of an artificial language comprised of four nonsense words each with three syllables (e.g., bidakupadotigolabubidaku ...). Infants' attention toward test trials where the speech was either consistent or inconsistent with the statistical properties of the familiarization stimulus was recorded. Infants looked reliably longer during the inconsistent test trials providing evidence that they could detect the statistical regularities present in a language, even when that language is new to them. Since this study, numerous investigations have documented the role that infants' sensitivity to statistical regularities plays in their language development (for reviews see Erickson and Thiessen, 2015; Krogh et al., 2013; Romberg and Saffran, 2010; Saffran, 2020; Saffran and Kirkham, 2018; Thiessen et al., 2013). Beyond language learning, there is a broad recognition of the potential power of statistical learning to underpin infant learning across various domains and modalities (Smith et al., 2018), including laying the foundation for infants' learning about objects (Colunga and Smith, 2005; Fiser and Aslin, 2002; Johnson et al., 2003), object sequences (Kirkham et al., 2002), and even infants' socio-cognitive understanding (Ruffman et al., 2012).

Statistical learning studies have revealed at least three types of statistical learning that infants demonstrate including: conditional, distributional, and cue-based statistics (Thiessen et al., 2013). *Conditional* statistics are regularities defined by how often two (or more) events, objects, or object features co-occur conditionally upon one another. Infants are sensitive to events that co-occur frequently (A and B), and as noted above, events or objects that are presented together frequently are typically learned together via association. However, in conditional statistical learning, the association between A and B is strongest when A and B occur together more often than when they occur individually. Infants become attuned to the co-occurrences of these pairings over time as indicated by their learning to expect B when A is presented. The study by Saffran et al. (1996) described above provides evidence that infants are sensitive to the transitional probabilities of recently learned sound pairings within and across word boundaries.

Infants notice conditional regularities in stimuli other than speech stimuli, including stimuli that is nonadjacent (i.e., learning that A always precedes C if exposed to sequences like ABC, ADC, BAC etc.) (e.g., Creel et al., 2004) and infants' own and others' action sequences (e.g., Baldwin et al., 2001). To illustrate, Fiser and Aslin (2002) demonstrated that infants 8 months of age who were repeatedly shown scenes with elements that occurred together according to different conditional probabilities found that infants preferred to look at scenes that contained the elements that were consistent with the conditional probabilities to which they had previously been exposed. Together, evidence illustrates that before their first birthdays, infants demonstrate a robust sensitivity to the conditional probabilities present in the input they receive across multiple modalities.

A second type of statistical learning reviewed by Thiessen et al. (2013) is *distributional* statistical learning, which is characterized by a sensitivity to differences in how often different sets of events or stimuli occur together relative to how often other sets of events or stimuli occur. For example, an infant's category of "bird" constructed through distributional statistics is likely to include

exemplars that all have properties such as has wings, a beak, and can fly. This category would help infants successfully categorize new birds they encounter into the category so long as the new bird has properties that are shared with the exemplar(s) (e.g., flamingo), but less likely to do so for an atypical bird (e.g., penguin or kiwi). Infants' ability to use distributional statistics to learn category boundaries has been demonstrated in the domains of objects (e.g., Rakison, 2004; Younger and Cohen, 1986) and language (e.g., Maye et al., 2002; Thiessen and Yee, 2010; Werker et al., 2007).

Given a very noisy environment, statistical learning of conditional and distributional regularities would require a great deal of exposure before consistent patterns are detected. This is because the learning context for infants is expansive and usually contains a diverse set of events and objects simultaneously (consistent with the critique of associative learning in **Associative learning** section). A third form of statistical learning, *cue-based* statistics, helps overcome this limitation. Cue-based statistics occur when an obvious perceptual property acts as a cue for a less obvious, conceptual property with which it is associated. An example of this is a speaker's use of pauses or different speech stress patterns, which are common cues that signal boundaries between words. In the non-linguistic domain, cue-based learning has been linked to an individuals' ability to use biological motion to distinguish between animate versus inanimate stimuli. Learners learn the cue through a sensitivity to the weighting of a particular cue with respect to how likely the cue is present with a particular property. Importantly, cue-based statistics are not restricted to a single instance or pairing but are applied (i.e., generalized) to other, unfamiliar, situations that involve the same cue, suggesting that once learned, the cue can fuel subsequent learning (for a review see Thiessen et al., 2013).

We have summarized three forms of statistical learning that have been demonstrated by infants across linguistic and non-linguistic domains of development. Infants' sensitivity to a range of statistical regularities in their environment clearly provides them with the tools they need to develop a broad base of knowledge, which has been supported by recent evidence linking 18-month-old infants' statistical learning skills in a language perception task with their subsequent vocabulary development (Ellis et al., 2021). While statistical learning clearly supports infants' learning in important ways, focusing on statistical learning mechanisms overlooks another form of learning that defines infancy—social learning.

Social learning

Decades of research has shown that infants exhibit a remarkable ability to learn through observations of and interactions with the individuals in their social world. This type of learning, aptly named *social learning* (e.g., Heyes, 1994), is argued to be particularly important for infants' learning of knowledge that is difficult, if not impossible, to uncover on their own, such as how to use novel objects (Barr et al., 1996; Meltzoff, 2010; Nielsen, 2006; Want and Harris, 2002) or cultural conventions (Csibra and Gergely, 2006; Henderson et al., 2008; Legare, 2017; Tomasello et al., 2010). To illustrate the power of social learning, consider how an infant might learn how to eat applesauce in a bowl that has been placed beside a shiny object that they have never used (i.e., a spoon). The infant could learn this skill through trial-and-error by experimenting with diverse ways to get the sauce into their mouth (e.g., try to use their hands, lift their bowl to their face, become distracted by making noise with the spoon), all of which are likely to be ineffective and messy! However, it is more likely that the infant would recruit their keen social learning skills and learn how to eat the applesauce by watching their people around them use their own spoons to successfully move the applesauce from their bowls to their mouths with minimal mess or loss of the delicious food. Social learning emerges early in infancy and encompasses learning that is enhanced or achieved via the presence or actions of a social partner across various domains (Heyes, 1994; Rendell et al., 2011).

Infants demonstrate their social learning in diverse ways, including by copying (i.e., imitating) others' actions (Barr et al., 1996; Elsner, 2007; Hayne et al., 2000) and using others' emotional expressions to guide their behavior (i.e., social referencing; Feinman, 1982; Stenberg, 2009; Walden and Ogan, 1988). We highlight infants' exceptional social learning skills by reviewing the development of imitation through infancy.

Imitation and emulation

Imitation studies have been used extensively to study generalized learning across infancy as imitation paradigms capitalize on infants' propensity to copy the gestures and/or outcomes produced by others while also having low language requirements, making such paradigms suitable for preverbal infants. In imitation tasks, a demonstrator shows the infant how to perform an action or sequence of actions on a novel object to complete a goal (e.g., puzzle as per Dickerson et al., 2013; Jones, 2009 for a review). This process is often repeated several times, particularly with younger infants who show memory benefits from multiple repetitions (Barr et al., 2007). Then the infant is invited to act on the object and their reproduction of target action(s) and goal(s) are recorded as an indicator of learning. Multiple types of social learning strategies around behavioral copying of the actions of others have been identified including mimicry, imitation, and emulation. *Mimicry* is defined as behavioral copying of another's actions without the knowledge of whether the copying was an intentional behavior elicited by the infant, or a reflexive response (Davis et al., 2021; Huang and Charman, 2005; Slaughter, 2021). Once infants develop an understanding of others' intentions, they engage in *imitation*, the copying of demonstrated actions and goals, and *emulation*, a term coined to describe the copying of the goals of the demonstrator without direct copying of the actions produced (Huang and Charman, 2005; Nielsen, 2006).

Early investigations of infants' imitation suggested that infants just days old could imitate others' facial expressions (Meltzoff and Moore, 1977). However, mixed evidence from subsequent work sparked debate about neonates' abilities to reliably imitate (Davis et al., 2021; Slaughter, 2021). More recent evidence suggests the earliest signs of imitation emerge between 4 and 7 months of age as indicated by infants' abilities to reliably mimic social partners' facial expressions and gestures (Hashiya et al., 2018; Iso-mura and Nakano, 2016).

Although early ontogeny is rife with learning statistical regularities, the prioritization of social information and the emergence of imitation as a robust learning tool emerges by 14 months (Meltzoff, 1985). In one study, Barr et al. (1996) investigating immediate and *deferred imitation*, where infants' imitation is tested after a delay (of hours to days), infants aged 14 months and 24 months showed evidence of learning a one-step action sequence and retaining that information after a 24 h delay. Age-related effects were evident, such that 24-month-olds performed better than 14-month-olds. Several studies have shown age-related increases in imitative learning, with the earliest accounts showing that infants as young as 6 months can imitate a one-step action immediately and after a 24 h delay when the action is demonstrated repeatedly during the learning phase (Barr et al., 1996). With increases in memory flexibility and information processing capabilities, older infants can tolerate greater delays between learning and test and can imitate more complex action sequences and goals (Barr and Hayne, 2015; Hayne et al., 2000).

How do researchers know when infants are engaging in "true" imitation where they understand the demonstrator's goal? Two examples of infants' social understanding underlying their imitation come from the findings of emulation studies. In these studies, infants are presented with a task for which there are multiple ways to achieve the task goal. For example, when completing a puzzle, the demonstrator may slide the puzzle pieces together using two fingers, but the goal (i.e., to complete the puzzle) can be achieved using any number of other actions (e.g., grasp puzzle pieces and slide, pick up and move puzzle pieces). When infants understand the goal and that actions are goal-directed, they may emulate, that is, they complete the task in a more efficient manner than that which was demonstrated (Huang and Charman, 2005; Matheson et al., 2013; Moser et al., 2015). Studies investigating the conditions under which infants use an emulation strategy reveal factors such as age and task complexity influence whether infants choose an imitation or emulation strategy (Huang and Charman, 2005; Nielsen, 2006).

Another social learning strategy that has been identified is *overimitation* where infants faithfully imitate a demonstrator's actions even when they are not related to attaining the goal, or when a more efficient strategy is available (Keupp et al., 2013, 2018; Lyons et al., 2007). Theories about the phenomenon of overimitation propose that infants may lack the understanding that the demonstrator's actions are not casually related to or relevant to the goal (Flynn, 2008; Lyons et al., 2007), recognize actions as goals in and of themselves when the demonstrator increases their salient (Keupp et al., 2018), or possess motivations to be affiliative (Over and Carpenter, 2012). The last two accounts suggest that, around their second birthday, infants emphasize the social aspects of social learning at the expense of more efficient learning, demonstrating the power of social learning.

The imitation work described above reveals that information conveyed socially becomes an important means of learning new skills in infancy, particularly after the first year following improvements in memory, information processing, and physical abilities (e.g., dexterity during object manipulation).

Foundations of social learning

Infants' social learning is supported by several early emerging behavioral patterns, such as a sensitivity and responsivity to social and communicative cues, social contingency, and behavioral mimicking, all of which lay the foundation for social learning mechanisms to develop. Sensitivity to social stimuli, such as faces (Guellai and Streri, 2011), gaze direction (Batki et al., 2000) and biological motion (Simion et al., 2011) emerges within the first year. Newborns, for instance, not only show a preference for face-like stimuli (Rosa Salva et al., 2011; Simion et al., 2011), but also exhibit greater recognition, as indexed through longer gaze time, for faces that have previously spoken to them with direct eye gaze compared to averted faces (Guellai and Streri, 2011). Infants seem particularly attuned to caregiver's *socio-communicative cues*, which include *ostensive cues* designed to grab attention (e.g., direct eye gaze, infant-directed speech), and *referential cues* that function to shift attention (gaze shifts, pointing) to objects and events in their shared physical space. Sensitivity and use of socio-communicative cues are theorized to be the building blocks of social learning and social cognition (Csibra and Gergely, 2006).

Within the first year of their postnatal lives, infants rapidly become sensitive and responsive to socio-communicative cues. Sensitivity to a caregiver's contingent responsivity to infant behaviors, termed *social contingency*, has been observed in infants as young as 3–4 months (Bigelow and DeCoste, 2003; Striano et al., 2005). Between 3 and 5 months of age, infants look longer at objects that their social partner had previously looked at relative to objects that were present, but had not been looked at by their social partner (Striano and Stahl, 2005). By 5–7 months of age, infants shift their attention in response to a social partner's gaze shifts that are accompanied by head turns (Hood et al., 1998; Senju and Csibra, 2008; Woodward, 2003) and then by the end of the first year of their lives, infants reliably follow gaze independently of head movement (Corkum and Moore, 1995; Tomasello et al., 2007). Reddy (2003) argued that a sensitivity to an adult's direct gaze brings about a specific kind of awareness in an infant—that they themselves are the object of another person's attention. In addition, gaze-following has been argued to be a critical component to understanding other people's immediate focus of attention, intentions, emotions, and aids in predicting others' future action and goals (D'Entremont et al., 1997; Farroni et al., 2002; Hood et al., 1998).

By 12 months, infants follow gaze reliably, but their patterns of attention are different from younger infants. For instance, older infants allocate more attention than younger infants to triadic interactions where the object of attention shifts (Woodward, 2003). Woodward (2003) revealed that 12-month-olds, but not younger infants, were sensitive to when an actor changed her focus of

attention to a different object than what she had repeatedly looked at previously. This effect was not observed in 7- to 9-month-old infants. Together, these patterns suggest that 12-month-olds are aware of the relation between a person's gaze and the object of their gaze (i.e., that gaze may reveal information about a person's object preferences). This learned behavior is likely to be a prerequisite to understanding the intentions and mental state of social agents (Woodward, 2003), a prominent theory in the study of social attention and theory of mind (Baron-Cohen, 1991; Charman et al., 2001).

Infants' developing sensitivity and responsivity to caregivers' social cues through the first year of their postnatal lives is argued to be key to the development of joint attention in the second year of life (Corkum and Moore, 1998; Gattis et al., 2020; Moore and Corkum, 1998). *Joint attention* is the coordination of attentional focus of two people on an object or event in their shared space. First, a social partner establishes mutual gaze with the infant, then proceeds to shift their gaze toward an object, and the infant responds by following their partner's gaze to the object (Carpenter et al., 1998; Mundy and Newell, 2007; Scaife and Bruner, 1975). Joint attention episodes are ubiquitous in observation and imitation learning experiments (Mundy and van Hecke, 2008; Slaughter and McConnell, 2003).

In sum, across the first year of their postnatal lives, infants become sensitive to socio-communicative cues from others and the presence of these cues impact infants' attention to, processing of, and learning about their world. These cues also inform infants about their social world—from *when* (e.g., ostensive signals) and *what* (e.g., referential signals) to attend to, to *what to avoid* (see social referencing; Vaillant-Molina and Bahrick, 2012). Through active engagement in social interactions (e.g., joint attention), infants gain an appreciation for social attention and also social cognition, such as the understanding of the intentionality of other's behaviors (Carpenter et al., 1998; Meltzoff and Brooks, 2008). As we discuss in the following section, as infants' experience and knowledge of others grows, so does their awareness that not all social partners provide accurate and reliable information.

Selective social learning

Infants are skilled at learning information provided by social partners and in so doing gain knowledge that they may not have been able to attain on their own, knowledge that is likely to be particularly relevant to the contexts and communities in which they are being raised. However, there are instances in which relying too much on others to acquire information can lead to learning errors or inaccurate assumptions. For example, if an infant places too much weight on learning from their caregivers, they may end up learning incorrect information particularly in contexts in which their caregiver is uncertain or does not possess the most accurate knowledge (e.g., a parent might mis-label an unknown object which could result in a word learning error). Another instance would be infants making the incorrect assumption that certain types of information, such as information about a person's preferences, reflect the preferences of all individuals in their social world. In so doing, infants might be less likely to attend to and learn about the unique aspects of individuals that make them who they are, which may have downstream consequences on their relationships. Thus, not only do infants need to be skilled social learners to acquire knowledge, but they also need to be skilled *selective* social learners.

Contexts of selective learning in infancy

Building on decades of research demonstrating that preschool-aged children are skilled at identifying the situations in which they should prioritize learning information from one individual over another, a substantial body of recent evidence suggests that infants are also selective social learners (for a review see Poulin-Dubois and Brosseau-Liard, 2016). Infants have been shown to adjust their learning in response to informant characteristics that signal when an informant may not be providing information likely to be correct, or relevant to the infant such as an informant's accuracy, competence, and language spoken.

Studies that assess selective learning tend to use a two-informant paradigm modeled from the pioneer study with preschool-aged children conducted by Koenig et al. (2004). Within this paradigm, participants are familiarised with two informants who act on the same object, but in divergent ways. For example, in the original study one informant provided correct labels for familiar objects (e.g., labeled a toy car "a car") while the other informant provided an incorrect label for the same object (e.g., labeled a toy car "a brush"). After familiarization, there was a label training phase in which the informants used different labels to label an object that was unfamiliar to the participant. When participants were later presented with the same unfamiliar object, the key question was whether they would be more likely to attach the label provided by the previously accurate informant to the unfamiliar object than they would the label provided by the previously inaccurate informant. The original findings revealed that, preschool-aged children were more likely to learn the word from the accurate informant suggesting that they used the past information about the informants' accuracy versus inaccuracy to inform their subsequent learning (Koenig et al., 2004). The original paradigm has been adapted for use with infants and, like preschool-aged children, 24-month-olds are more likely to learn novel words from an informant who labeled objects correctly in the past than from an informant who previously labeled objects incorrectly (Koenig and Woodward, 2010; Krogh-Jespersen and Echols, 2012), or used known objects in incorrect ways such as using a toy car to brush their hair (Henderson et al., 2015).

Infants' selective learning has been shown beyond word learning contexts. Within the action imitation domain, 18-month-old infants reproduced non-goal related actions when they were completed by an informant who accurately labeled familiar objects, but not when they were completed by an inaccurate informant (Brooker and Poulin-Dubois, 2013). In this study, infants were more likely to imitate an action that was not directly tied to attaining a goal (e.g., moving an object in a particular manner to its end goal) that had been provided by an informant who had previously labeled objects correctly than they were actions provided by an informant who labeled familiar objects incorrectly. Interestingly, infants were also more likely to help the accurate informant

over the inaccurate informant when both informants were in need, suggesting that selective learning preferences may also extend to selectivity in prosocial behavior (Brooker and Poulin-Dubois, 2013). In addition to labels, infants are also attentive to other behaviors of informants. For instance, infants as young as 14 months of age selectively imitate actions provided by an informant who has used familiar objects correctly (Zmyj et al., 2010) and informants who have been shown to speak the same language as the infant (Buttelmann et al., 2013; but see Howard et al., 2015).

An informant's emotional signals also influence infants' social learning. This was first shown by Chow et al. (2008) using a between-subjects adaptation of the unreliable informant paradigm described above (i.e., Koenig et al., 2004). In this study, 14-month-old infants were familiarized to a reliable informant who had emoted positively toward a box containing a toy or an unreliable informant who had emoted positively toward an empty box. A subsequent gaze following task was used to test how quickly infants in each condition moved to examine the object that the informant had looked toward. The results revealed that infants who had been exposed to a reliable informant were quicker to follow the informant's gaze and explore the object than were infants who had been exposed to an unreliable informant. These findings suggest that infants had learned about the reliability of the informant in the familiarization phase and infants in the unreliable condition had learned that the person was not reliable at signaling when there was something interesting for them to see. In addition to demonstrating another context that highlights infants' selective social learning, these findings contribute to the debate surrounding the nature of infants' gaze following abilities. Rather than supporting infants' gaze following because of an automatic mechanism (e.g., Langton et al., 2000) or a developed sensitivity to contingency (e.g., Moore, 1996), Chow and colleagues' findings suggest that, by 14 months of age, infants' gaze following is supported by their understanding of the referential nature of gaze (i.e., that a person's gaze is goal-directed toward a particular referent and, in turn, socially meaningful).

Building from 14-month-olds' tendency to selectively examine objects attended to by reliable over unreliable informants, by 18 months, infants preferentially imitate actions completed by informants who previously provided reliable emotional signals over actions provided by informants who previously provided unreliable emotional signals (see Chiarella and Poulin-Dubois, 2013 for studies examining infants' selective use of emotional signals).

Even more impressive are the nuances in infants' selective social learning. For example, in a study conducted by Henderson et al. (2015), 24-month-olds learned atypical labels from informants who used familiar objects in atypical ways, but not informants who used familiar objects in typical ways, showing that infants do not simply acquire any information from informants who have been shown to be accurate in the past. Another example of nuanced selective learning in infants comes from a study in which infants showed that they would learn new actions from informants who did not speak the infant's own language in a live face-to-face context, but not a video context (Howard et al., 2015). Interestingly, a follow-up study using the same data revealed that infants' everyday experiences influence their social learning with infants who were being raised in more ethnically diverse (implying linguistic diversity) neighbourhoods being more likely to imitate informants from diverse linguistic groups (Howard et al., 2014). The findings from studies examining selective social learning in infancy reveal that, between 14 and 18 months of age, infants can selectively apply their social learning skills toward informants and contexts in which they are most likely to acquire accurate and relevant information.

Foundations of selective social learning

While selective social learning has yet to be reliably demonstrated in infants younger than 14 months of age, there is robust evidence suggesting that some key foundations likely to support infants' selective social learning are in place before infants' first birthdays. To date, several studies have demonstrated infants' early sensitivity to the varied informant characteristics that have been presented to older infants in selective learning studies. Infants as young as 8 months of age have been shown to be sensitive to information that signals an informant's accuracy. In this study, 8-month-old infants were more likely to visually examine objects that were looked at by an informant who had previously been reliable at looking toward where an object would appear on a screen than they were to examine objects that had been looked at by an unreliable informant (Tummeltshammer et al., 2014). In another series of experiments, 12-month-old infants were more likely to pay attention to informants who played with a set of toys competently (e.g., put bricks in the correct holes and objects away in the correct places) than they were an incompetent informant (e.g., did not put the bricks in the correct holes) (Stenberg, 2013). Informant competence also influenced infants' object-directed play with infants who had been introduced to an unfamiliar and ambiguous toy by the competent informant were more likely to interact with it than were infants who had been introduced to it by the incompetent informant (Stenberg, 2013).

Recently, Bazhydai et al. (2020) developed a paradigm conceptually like the inaccurate informant paradigm described above (Koenig et al., 2004) to test whether 12-month-old infants would actively seek knowledge from an accurate informant over an informant who had explicitly expressed a lack of knowledge (e.g., said "I don't know") in response to the presentation of an object familiar to the infant. Of interest was what infants would later do in an ambiguous context when they were shown two novel objects and asked to find one of them using a novel word (i.e., "Which one is the modi?"). Specifically, the researchers were interested in whether infants would show more signals of social referencing, such as increased visual attention or information-seeking points, toward the knowledgeable informant than toward the ignorant non-informant during this "test phase" of the experiment than during the other phases of the study. The results revealed that infants were more likely to look toward the previously accurate informant than they were the individual who had explicitly declared their lack of knowledge. Young infants' early emerging selective social referencing behavior provides evidence that pre-verbal infants are sensitive to key characteristics of informants that signal whether, or not, they are likely to be useful providers of knowledge. An important question for future research will be to test whether infants' early selective social referencing maps onto their later developing selective social learning.

Another important remaining question involves the psychological mechanisms that support infants' selective social learning abilities, which is a topic of much debate. The potential mechanisms that have been proposed to underlie selective social learning range from higher-order, domain-specific capacities, such as an understanding of theory of mind (e.g., Poulin-Dubois and Brosseau-Liard, 2016), to lower-level non-specialized general learning mechanisms like associative (Heyes, 2017) or statistical (Sobel and Kushnir, 2013) learning. The above findings of the study conducted by Tummeltshammer et al. (2014) reveal that infants' sensitivity to statistical regularities plays a role in infants' statistical learning. However, to date, only one study has directly tested the extent to which infants' statistical learning or developing theory of mind abilities contribute to infants' selective social learning (Crivello et al., 2018). In this study, associations between infants' selective word learning from an accurate or an inaccurate informant and infants' theory of mind and statistical learning abilities were examined. The results supported links between infants' selective learning and their developing theory of mind, but not their statistical learning abilities. Specifically, for infants who were taught novel words by a previously unreliable informant, infants who had more advanced social cognitive skills (i.e., were able to accurately attribute knowledge, or lack thereof, to others) were less likely to learn the words than were infants who had less advanced social cognitive skills (Crivello et al., 2018).

The findings reported by Crivello et al. (2018) were followed-up in a similar study involving a within-subjects design (rather than a between-subjects design) in which 18-month-old infants were exposed to both an accurate and an inaccurate informant and an associative learning task (rather than a statistical learning task) (Crivello et al., 2021). Consistent with their prior work, this study revealed that, when exposed to two informants of differing levels of reliability, infants selectively learned from the previously accurate informant. Further, they revealed that infants with more advanced theory of mind abilities were less likely to learn novel words from an inaccurate speaker. Thus, consistent with prior evidence linking more advanced theory of mind skills and selective social learning in older children (e.g., Brosseau-Liard et al., 2015; DiYanni and Kelemen, 2008; Mills and Elashi, 2014), the findings reported by Crivello and colleagues (Crivello et al., 2018, 2021) support the possibility that the ability to avoid learning new knowledge from informants who may not be providing correct, or socially relevant information, is supported by infants' domain-specific abilities, namely infants' developing theory of mind, and not the domain-general abilities of statistical or associative learning. There is, however, recent evidence suggesting that two other domain-general abilities - infants' causal reasoning and pre-metacognitive awareness—are associated with infants' selective social learning (Kuzyk et al., 2020). Thus, the extent to which infants' selective social learning is driven by domain-general versus domain-specific abilities remains a topic for debate with further research clearly needed to resolve.

Summary

Infants clearly have a variety of tools at their disposal to acquire the foundational knowledge that will support their learning and development beyond infancy. The above findings suggest that learning within the first year of infants' postnatal lives is primarily supported by general learning mechanisms, involving associative and statistical learning. Throughout the second year of life, infants' social learning abilities take center stage. Infants begin to recognize the value of socio-communicative cues, such as a social partner's eye gaze and pointing gestures. As signals they use to acquire information and prioritize such signals over other signals for learning, such as temporal contiguity or salience reviewed in [Associative learning](#) section. This social learning revolution enables infants to expand their knowledge in efficient, yet flexible ways as demonstrated by infants' selective social learning. While existing evidence clearly suggests that infants recruit a range of sources in their everyday lives to learn, what is less clear is how these mechanisms work together to support learning. This is because the decades of research into infants' associative, statistical, and social learning have been quite disparate. As such, the findings supporting different mechanisms tend to suggest that learning is very island based—separate and not connected. However, this is not likely to be the case and does not present a parsimonious view of development. Others have recognized this disparity and have begun to work toward developing unified, integrated frameworks for learning. In the following section we describe some of these approaches.

Moving beyond separate accounts toward integrated frameworks for infant learning

One of the main motivators of research in infant learning is the pursuit of explanations that are necessary and sufficient to explain observed patterns of behavior. Perspectives that emphasize or stress the importance of only one learning mechanism over others do not adequately capture infants' performance across diverse domains of learning, which includes learning words, actions, and about objects which we reviewed in [Learning in infancy](#) section. Further, although researchers often allude to an infant's ability to detect and use statistical regularities that are present in both their physical and social environments, few studies have systematically investigated how associative, statistical, and social learning mechanisms work together to support learning within a single study. In this section, we consider attempts that have been directed toward examining the joint influence of general learning mechanisms, statistical learning, and social learning in infancy.

One noteworthy empirical investigation of the complementary roles that associative and social learning processes play in infants' word learning was conducted by Hollich et al. (2000). In this pioneering, eight-experiment study, word learning was examined in 12- to 24-month-old infants to examine the extent to which infants prioritize different cues and sources of information to identify the correct referent (i.e., meaning) of a novel word across development. The main strength of the experiments described by Hollich et al. was that the same basic paradigm (i.e., the Interactive Intermodal Preferential Looking Paradigm (IIPLP)) was used

systematically to test the age at which infants used associative versus social cues during word learning. While a detailed review of the experiments and their findings is beyond the scope of this article, the results can be summarized easily given the literature reviewed in **Learning in infancy** section. During word learning, infants at 12 months of age placed greater emphasis on cues that support associative learning such as perceptual salience, whereas older infants (between 18- and 24-months of age) placed greater emphasis on using social cues (e.g., informant's eye gaze). However, the results also demonstrated that, while infants placed differential emphasis on associative versus social cues across development, other cues are not ignored. Twelve-month-old infants' word learning was influenced by social cues and 24-month-olds did not disregard perceptual salience as a relevant cue for uncovering word meanings. By demonstrating how multiple learning strategies work in coordination to shape early word learning, this pioneer work opened the discussion for researchers to develop paradigms designed to gain a comprehensive understanding of the ways in which multiple learning mechanisms work together to support learning across the first two years of postnatal life.

While the work of [Hollich et al. \(2000\)](#) enhanced understanding of how different learning strategies come together to support early word learning, their framework does not easily generalize to other learning contexts. However, the work of [Goldstein et al. \(2010\)](#) provides an example of an integrated approach to learning that does. In their article, Goldstein and colleagues recognize that infants outperform models of language learning that are based entirely on statistical learning mechanisms and propose a framework, called ACCESS (Align Candidates, Compare, Evaluate Statistical/Social Significance), which emphasizes how an infant's social context may support statistical learning. Within this framework, infants are viewed not only as excellent pattern recognisers, but as active learners who make inferences and extract meaning from the regularities within their environment. This is where the social context becomes most important.

In the ACCESS model, social partners influence infants' *attention*, which functions to alleviate the processing required through statistical learning, and the *temporal structure*, which binds statistical events across various time windows though consistent, temporally constrained social patterns. Social interactions are behaviorally reinforcing when presented in close temporal proximity to learning experiences, particularly due to their social contingency and high incentive value compared to other non-social reinforcers ([Thiele et al., 2021](#); [Vermetti et al., 2017](#)). For example, caregivers who respond *contingently* to their infant, and whose responses are salient, relevant, and meaningful have an elevated *incentive value* compared to non-social rewards (e.g., dynamic object movement). In the following sections, we elaborate on how social information, present at temporally critical learning experiences, promote learning of associations ([Thiele et al., 2021](#)), hierarchical rule-based learning ([Werchan and Amso, 2020](#)), statistical learning ([Wu et al., 2011](#)) and generalization of learning across contexts ([Werchan and Amso, 2020](#)). In addition, we consider how this framework might apply to domains outside of language learning from which it originated.

Attention

Infants' social partners use of socio-communicative cues, such as eye gaze, function to direct and sustain infants' attention for longer than what would happen in the absence of these cues ([Suarez-Rivera et al., 2019](#)). With age, infants become increasingly sensitive to their social partners' eye gaze. At approximately 14 months of age, infants show sustained attention, particularly during joint attention episodes when caregivers' and infants' attention is coordinated on the same object ([Suarez-Rivera et al., 2019](#); [Yu and Smith, 2016](#)). By 24 months, infants reliably use eye gaze to an object, paired with an object label, to successfully map the label to the object ([Graham et al., 2010](#)). Even more revealing is that [Graham et al. \(2010\)](#) found that performance was dramatically impaired when gaze direction and object labeling were inconsistent. Critically, these cues become more beneficial when the learning environment contains visual distractors that increase ambiguity as to what is relevant to attend to ([Wu et al., 2011](#)), a common occurrence in naturalistic learning settings ([Medina et al., 2011](#)). This observation is echoed in the ACCESS framework, which suggests that social cues from a reliable and knowledgeable partner support learning in distracting environments by narrowing infants' visual attention to that which is most relevant for learning; that is, social attention cues restrict infants' statistical processing to a more focused sample of their environment.

[Wu et al. \(2011\)](#) provide further evidence that social cues enhance infants' statistical learning by influencing infants' attention. In this four-experiment study, 9-month-old infants were repeatedly shown different patterns depicting three novel objects. Critically, two of the objects always appeared together while the third differed across trials. Of interest was how infants used statistical regularities and social cues to learn which two of three shapes co-occurred with one another across contexts that varied in complexity. To test if infants learned the co-occurrence of the two objects during the learning phase, infants were shown two types of test trials: the consistent test trials preserved the association between the two objects that previously co-occurred whereas the inconsistent test trials did not (the two co-occurring objects were shown moving away from one another). The authors hypothesized that if infants learned which objects co-occurred, they would look longer toward the inconsistent test trials that violated their expectations about the objects. The findings of their first experiment were consistent with this expectation confirming that 9-month-olds in their study were sensitive to the co-occurrence of two novel "objects". In subsequent experiments, infants were shown the test trials at the same time (rather than in alternation) thereby introducing distractions to make the task more difficult by increasing the visual attention and processing demands placed on infants. While infants' performance was influenced by the presence of the distractors during the test trials (i.e., infants did not show looking time differences during the test phase with the distractors present), the introduction of an informative social cue (i.e., a face turned toward the pattern showing the co-occurring objects) enhanced infants' performance in the challenging context. Thus, these findings demonstrate the joint utility of statistical and social cues in infants' learning; while 9-month-old infants are sensitive to co-occurring regularities, social cues are helpful in cueing infants' attention in challenging learning contexts.

While informative social cues can support infants' attention and learning, particularly in distracting contexts (Wu et al., 2011), information provided by social sources is not always helpful (as described in **Social learning** section). Xiao et al. (2018) examined whether a social informant's reliability impacts infants' ability to learn cue-object pairings in a statistical learning task. Building on prior work on infants' selective social learning (see Poulin-Dubois and Brosseau-Liard, 2016 for review) and evidence that infants show a same-race-bias (e.g., Quinn et al., 2016; Xiao et al., 2014), Xiao et al. (2018) examined whether 7-month-old infants would use a social informant's race as a cue to guide their learning in uncertain contexts. In a series of experiments, infants were exposed to one of three learning phases where either a same-race or other-race informant provided gaze cues to a location that predicted a target location on 25%, 50% or 100% of the trials. During the test phase, infants' tendency to follow the gaze of each informant without a target appearing was measured. Infants followed the gaze of both informants when their gaze was highly predictable (100%) of target location and neither informants' gaze in low reliability (25%) contexts. However, when an informant's reliability was at 50%, the same-race effect emerged such that infants were more likely to follow the gaze of the same-race informant than they were the other-race informant. This same-race bias suggests that, under challenging learning situations where the probability of the social partner being correct is statistically at chance, infants use prior experience and knowledge to guide their attention and learning.

The above findings demonstrate that in contexts in which it is challenging for infants to extract the statistical patterns present, or if the patterns are unreliable, infants use social information to guide their learning (even information that may not be useful for that specific learning context—e.g., race). Interestingly though, not all social information aids learning. For example, Monroy et al. (2017) investigated statistical learning of action sequences in a sample of 18-month-olds and adults. Participants were presented with a video of two actors who verbally established they were playing a game either cooperatively or in parallel. In both conditions, the actors sat opposite one another, and performed an identical action sequence on a unique set of objects in front of them. The action sequence contained a statistical structure based on transitional probabilities that, if learned, would predict upcoming actions. Participants' anticipatory eye gaze during this phase provided evidence of their initial learning of the sequence. Afterward, participants' memory for the action sequence was tested as they were invited to interact with the objects from the learning phase, both independently and cooperatively with an experimenter. Of interest was whether participants learned better (i.e., were more sensitive to the statistical regularities) in a highly social context (i.e., when the actors were cooperating). Anticipatory eye gaze results revealed that infants and adults correctly predicted action sequences during the learning phase, illustrating they learned the statistical structure of the sequence. While the cooperative context boosted adults' learning relative to the parallel context, infants did not show a learning advantage in the cooperation context. One explanation for the lack of influence of the social context was that the statistical learning of the action sequence was easy to achieve without any social support, like the 100% certainty condition in the Xiao et al. (2018) study. It is also worth noting that the social context did not directly involve the infant, through ostensive (e.g., direct eye gaze) communicative cues, as did other studies described that found an effect of social information on statistical learning (Wu et al., 2011; Xiao et al., 2018). Like the social context in the Monroy et al. (2017), a study by Wu et al. (2014) featured a social but non-ostensive cue condition which failed to facilitate multimodal learning in an attention cuing paradigm. Ostensive cues, however, did promote multimodal learning (Wu et al., 2014), and may be necessary to capture attention and enable the extraction of complex statistical patterns.

This section demonstrates how infants' statistical learning is enhanced by ostensive signals. When statistical learning is made challenging through distractors (Wu et al., 2011) or conditions of uncertainty (Xiao et al., 2018), infants use social cues to enhance their learning of the relevant statistical regularities. In these situations, social cues may resolve ambiguity about who is a relevant source of information and what is to be learned. Further research is needed to better understand how and when infants exploit social information to aid in learning the associations, rules, and statistics present in the environmental inputs they encounter.

Social partners and social contingency

In addition to providing, at a minimum, greater attentional focus and differential processing of to-be-learned objects and events, social cues are embedded within social interactions that have predictable (i.e., statistically reliable) structures. Caregivers have predictable patterns of interacting with their infants across a variety of tasks, such as in their contingent communication, how they teach their infants about objects (Suarez-Rivera et al., 2019; Tamis-LeMonda et al., 2014), and in playing cooperative games such as block building and peekaboo (Bakeman and Adamson, 1984; Duncan and Farley, 1990). Goldstein et al. (2010) argued that these types of social interactions, which are contingent and coordinated with the infant's behavior, serve learning by binding together multiple repetitive patterns across the learning episode and across learning sessions (i.e., over a delayed period). Memory for and detection of the behavioral patterns is thought to aid in the learning of the statistical regularities associated with such social interactions. When considering peekaboo games, caregivers play several rounds of peekaboo in a single play session. During these rounds, caregivers contingently respond to their infant's behavior, which serves to increase infants' attention and engagement in the task. Although caregivers' behaviors have consistency across rounds (e.g., exaggerated speech and infant-directed action aka "motionese"), slight variability in their behavior (e.g., the way they say peekaboo, how they hide themselves, or their accompanying gestures) is theorized to facilitate extraction of regularities across multiple repetitions (Goldstein et al., 2010).

Expanding on the predictability of social partners' behavior during everyday interactions described by Goldstein et al. (2010), caregivers also scaffold learning in predictable ways (Brand et al., 2013; Levine et al., 2019; Yu and Smith, 2012). In a study conducted by Yu and Smith (2012) caregivers labeled objects for 18-month-old infants in a naturalistic setting while infants wore a head-mounted camera and manually explored the objects. Infants then performed a test where the same objects were presented to the infant and named by the experimenter. Better learning, as indicated by infants' looking toward the named objects at test, was associated with episodes during which the caregiver named the object when the infant held the object in the center of their visual field. Interestingly, caregivers tended to name objects when infants were visually focused on the object, suggesting that caregivers are sensitive to this optimal window for infants' learning; that is, when there is most likely to be a one-to-one mapping of label to object.

Given caregiver's sensitivity to optimal learning conditions, it should not come as a surprise that they modify their behaviors when interacting with their infant to draw attention to meaningful boundaries in action sequences (e.g., infant-directed action/ "motionese"), functionally like the effect of infant-directed speech on language learning (Levine et al., 2019). In one study, Brand et al. (2013) examined mothers' scaffolding behaviors while demonstrating a series of action sequences to their 7- and 12-month-old infants. Of particular interest was whether mothers' ostensive, infant-directed gaze occurred more frequently at event boundaries during two object demonstrations. Mothers were given two objects to demonstrate, one object had a set of ordered actions that resulted in a goal (enabling-sequence) while the other did not have a specific order to the action sequence nor an end goal (arbitrary-sequence). Results revealed that mothers conveyed infant-directed gaze cues more regularly at action boundaries. Given the mounting evidence that patterns of reliable social information in naturalistic settings predict upcoming actions (see Levine et al., 2019 for review) and that the availability of social cues in experimentally controlled settings aid in detecting statistical regularities (Wu et al., 2011), more research on the systematic behaviors of caregivers could illuminate *if and how* their predictability benefits learning of statistical input patterns.

Incentive value

In addition to the attentional and temporal cues that social interactions provide which can guide infants' associative and statistical learning, infants also appreciate that social interactions, and the objects associated with social interactions, are inherently meaningful and rewarding. This possibility was recently examined by Thiele et al. (2021) with 13-month-old infants. In this study, infants were presented with a non-social cue (i.e., a green circle or a blue triangle) before being shown a video clip that contained two experimenters who were interacting socially or independently (i.e., completing actions when they were either facing toward or away from one another). The social versus non-social videos were always paired with the same cue and played on the same side of the display, which allowed the researchers to test whether infants were more likely to learn the cue-target association (i.e., side of video) in the social interaction condition. As hypothesized, infants showed enhanced learning of the cue-target association, demonstrated by reduced latencies and anticipatory gaze shifts toward the side on which the video would play after being presented with the cue, for the social interaction condition than they were the nonsocial condition. Critically, the findings were not simply a result of a general tendency for infants to allocate greater attention toward the side of the display that regularly displayed the social interaction. Infants' enhanced cue-target learning in the social condition provides support for the priority that infants place on social interactions as a source of potentially relevant meaningful information, even when infants are simply observing video recordings of a third-party social interaction.

In addition to cue-target associations, infants as young as 8 months use social stimuli to learn hierarchical rule-based structures, which are sensory input patterns that follow specific rules that are generalizable and can be organized into separate rule sets (Werchan et al., 2015). To demonstrate this ability, Werchan and Amso (2021) examined 9-month-old infants' learning of cue-target pairings. In each learning trial, a social attention cue (i.e., a female face presented on a colored background) was paired with a cartoon animation that appeared in one of four locations on a display. The specific face (dark-haired female, light-haired female) and background color (blue, green) combination of the cue corresponded to specific location where the cartoon would play. This design allowed the infant to learn through association (out of four possible person-background color combinations) or by using a hierarchical rule-based structure using either the faces or the background color as the top tier of the hierarchy. Following learning, infants performed one of two generalization conditions, featuring a new set of faces or a new set of colors. The generalization phase was designed to test: (1) whether infants would transfer their initial learning onto a separate set of stimuli and (2) which learning technique (hierarchical with face or color dominant versus associative learning) they employed. The results revealed that infants who were given a new set of faces maintained faster reaction times (visual attention) to target locations compared to infants in the colour-change generalization condition. These findings show that infants formed associations between object properties (such as color) and social information (such as face information) when learning. More importantly, the findings suggest that infants use social information to organize these associations into hierarchical rule-based structures that can be flexibly generalized to new learning sets. Consistent with the findings discussed in Attention section, the authors interpreted their findings to demonstrate that social information guides infants' attention. Critically though, they propose that social information also reinforces learning through the reward system by acquiring incentive value, which has previously been shown to support generalized learning in infants (e.g., Tummeltshammer et al., 2019) and reinforcement learning in preschool-aged children (e.g., Verneti et al., 2017).

Summary

Prior to 12 months, associative, statistical, and social learning mechanisms are available to help infants make sense of the stream of sensory information they encounter. Although infants may be capable of learning solely through a statistical learning approach, for instance, this may not accurately reflect real-world conditions or accurately predict the developmental trajectories of more advanced learning achievements. And in fact, there is now a growing body of evidence of the validity of an integrated approach to understand infant learning (Goldstein et al., 2010; Wu et al., 2011; Xiao et al., 2018).

Following from the concepts outlined in Goldstein et al. (2010) model of learning across time and space, we took a close look at recent studies that tested whether social stimuli would moderate the learning of associations (Thiele et al., 2021), rule-based (Werchan and Amso, 2021) and statistical structures (Wu et al., 2011; Xiao et al., 2018). Empirical evidence suggests that by 9 months, socio-communicative cues capture and direct infants' visual attention to and aid in their learning of hierarchical rule-based structures (Werchan and Amso, 2021) and statistical regularities (Wu et al., 2011). Social cues appear to function to resolve uncertainty about *what* information is most relevant (Wu et al., 2011) or when social agents are unpredictable (Xiao et al., 2018). However, not all social information is equally prioritized. Social but non-ostensive information, as opposed to socio-communicative cues, may support associative learning at 13 months, as a secondary reinforcer with high incentive value (Thiele et al., 2021), but there is no evidence of such information's involvement in more advanced statistical and transfer learning performance (Monroy et al., 2017).

Social cues support attentional systems in their pursuit of patterns of input, but the delivery of social cues is not random or arbitrary. Social partners provide infants with socially contingent, predictable social attention and scaffolding cues (Brand et al., 2013; Levine et al., 2019). Infants are sensitive to these behavioral patterns, and they have been shown to support mapping of labels to objects (Yu and Smith, 2012). Infants may also appreciate and be driven to learn associations and rules based on the reinforcing qualities of social information (Thiele et al., 2021; Werchan and Amso, 2021). However, what we know about the temporal characteristics of naturalistic behavioral patterns of caregivers and their incentive value make investigations into their potential facilitatory effect of statistical or rule-based learning alluring.

Moving forward: research directions, considerations, and conclusions

To date, most of the research surrounding the mechanisms of early learning has focused on studies designed to test infants' learning through isolated mechanisms, or pitting individual mechanisms against one another, rather than looking at how the distinct types of learning work together to support infants' learning across contexts and domains. In Section **Moving beyond separate accounts towards integrated frameworks for infant learning** we introduced two approaches that defy this norm by providing integrative accounts of infant learning that strive to explain infants' learning via general learning mechanisms, such as associative and statistical learning, and social learning mechanisms and processes. One of the most alluring aspects of integrated approaches, such as that proposed by Goldstein et al. (2010), is the promise of capturing infants' learning across diverse contexts in a parsimonious way. With parsimonious integrative frameworks in tow, researchers can design empirical studies to test how infants recruit their diverse set of learning tools to acquire knowledge across domains and modalities. We end by suggesting four broad research areas that we believe would benefit from more integrative frameworks.

First, we propose that researchers could use the Goldstein et al. (2010) framework as a starting point to design studies to examine systematically the statistical regularities present in infants' everyday environments and how these regularities differ across diverse cultural contexts (see Rogoff et al., 2018 for a relevant discussion). The work of Linda Smith and colleagues offers promising avenues to examine these questions (Clerkin et al., 2017; Schroer et al., 2019; Smith et al., 2018; Suanda et al., 2019; Suarez-Rivera et al., 2019; Yu and Smith, 2012, 2016). In their studies, Smith and colleagues attach head-mounted cameras onto infants to capture the regularities in infants' environments from the perspective of the infant. In many of these studies, the infant view is enhanced by recordings from a second head-mounted camera on the caregiver, which allows for caregivers' attentional focus and multi-modal behaviors to also be captured. While this work has been conducted primarily in the USA thus far, recruiting this methodology to study caregiver-infant interactions across the globe will provide invaluable foundational knowledge about the ways in which statistical information combines with social signals and cues in infants' everyday lives to support infants' learning. Given that to date, much of the research on infant learning and the strategies still comes primarily from research groups in Western, Educated, Industrialized, Rich and Democratic (i.e., WEIRD) countries with WEIRD populations (Henrich et al., 2010; Singh et al., 2021), integrated approaches may be the key to moving the field beyond WEIRD studies on infant learning and development.

Second, integrative approaches will also allow for systematic investigations of how different interactive partners in infants' social worlds shape their learning environments and opportunities for learning. As described above, caregivers and infants develop predictable patterns of interacting, which subsequently shape infants' learning (Brand et al., 2013; Levine et al., 2019; Yu and Smith, 2012). One major limitation of much of the work on caregiver behavior and infant-caregiver interactions is the continued focus on mothers even though infants routinely interact with many other social partners in their everyday lives such as a second caregiver at home or caregivers at the early childcare center they attend. Without systematic investigations of the statistical regularities, social cues, and interactive patterns in the range of caregiving contexts infants experience, a complete

understanding of infant learning will elude the field. There has been some prior work demonstrating some ways in which infants (and caregivers) respond differently to different interactive social partners (e.g., Morrison et al., 2021), future work from an integrative framework can examine how variation in expected statistical regularities, or social cues, across caregivers and other social partners influences infants' learning.

Third, one consistently challenging learning task faced by infants is how they generalize or transfer learning from one context to another. Contextual changes are broadly defined; they may refer to a temporal delay between learning and test, or situations in which the physical features of the stimuli or physical learning environment change from learning to test (Barnett and Ceci, 2002). For example, in one study Moser et al. (2015) showed toddlers how to assemble a virtual three-piece puzzle on a 2D touchscreen tablet. Toddlers were then immediately given one of two tests, the same virtual puzzle, or a 3D replica of the virtual puzzle. Toddlers who were tested with the same puzzle pieces that they encountered during the test were significantly more likely to reproduce the goal (to assemble the puzzle) than were toddlers who were given the real puzzle pieces. These generalization or transfer deficits have been studied extensively in screen media learning (dubbed the "video deficit"; Anderson and Pempek, 2005; Barr, 2010; Hipp et al., 2017; Kirkorian et al., 2016), as well as in other domains relating to learning about objects (Woodward and Hoyne, 1999; Wu et al., 2011), language (Maye et al., 2008) and hierarchical rules (Werchan et al., 2015), to name a few. Investigations with integrated approaches may be able to shed a light on how infants may accomplish the transfer problem.

Finally, we encourage researchers to adopt integrative frameworks to describe infants' learning and development beyond the cognitive and language domains focused on in this contribution. In addition to cognitive development, infants make significant strides in their social development throughout the first two years of their lives. By 24 months of age, infants demonstrate at least a basic understanding of goal-directed actions across individual- and shared-goal contexts (for a review see Woodward et al., 2009) and engage in a range of prosocial behaviors (Brownell, 2011, 2016; Hay, 2009). We propose that integrative frameworks that consider the joint contributions of social learning processes and the powerful general learning mechanisms of associative and statistical learning would allow for the generation of complete theories and models of social development. For example, we know from prior work that infants imitate prosocial actions modeled by a social partner (Schuhmacher et al., 2019). Following up on these findings, integrative frameworks could be used to design studies to examine how statistical information of prosocial behaviors might interact with social learning mechanisms to support development of prosocial actions. Or, they can be used to examine whether individual differences in infants' prosocial competencies can be described by variation in the distribution of different exemplars of prosocial actions and the people who engage in such actions in infants' everyday lives.

Here, we have introduced four ways in which integrated approaches will be key to extending understanding of infants' learning and development although there are likely to be many more. We note though that the design of studies taking integrated approaches must be thoughtful and diligent. For instance, researchers need to be mindful of the different behavioral responses that infants might have to ostensive versus non-ostensive cues, or a combination of cues (e.g., infant-directed speech + gaze shift to target; Szufnarowska et al., 2014). It would also be beneficial for adopted paradigms to be appropriate for a wide age-range to enable robust examinations of age-related changes. This is particularly important as we know that infants' attention to, and use of social cues for, learning change across the infancy period (Brooks and Meltzoff, 2005; Mundy et al., 2007; Striano and Stahl, 2005). While designing studies within an integrative framework may have its challenges, for the field to move forward, the questions driving research should no longer be *if* associative, statistical, and social learning shapes learning but should be *how* they come together to promote infants' learning of the knowledge foundations for optimal development.

References

- Anderson, D.R., Pempek, T.A., 2005. Television and very young children. *Am. Behav. Sci.* 48 (5), 505–522. <https://doi.org/10.1177/0002764204271506>.
- Bakeman, R., Adamson, L.B., 1984. Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Dev.* 55, 1278–1289.
- Baldwin, D.A., Baird, J.A., Saylor, M.M., Clark, M.A., 2001. Infants parse dynamic action. *Child Dev.* 72 (3), 708–717. <https://doi.org/10.1111/1467-8624.00310>.
- Barnett, S.M., Ceci, S.J., 2002. When and where do we apply what we learn? A taxonomy for far transfer. *Psychol. Bull.* 128 (4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>.
- Baron-Cohen, S., 1991. Precursors to a theory of mind: understanding attention in others. In: *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*. Basil Blackwell, pp. 233–251.
- Barr, R., 2010. Transfer of learning between 2D and 3D sources during infancy: informing theory and practice. *Dev. Rev.* 30 (2), 128–154. <https://doi.org/10.1016/j.dr.2010.03.001>.
- Barr, R., Dowden, A., Hayne, H., 1996. Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behav. Dev.* 19 (2), 159–170. [https://doi.org/10.1016/S0163-6383\(96\)90015-6](https://doi.org/10.1016/S0163-6383(96)90015-6).
- Barr, R., Hayne, H., 2015. Developmental changes in imitation from television during infancy. *Child Dev.* 70 (5), 1067–1081. <https://doi.org/10.1111/1467-8624.00079>.
- Barr, R., Muentener, P., Garcia, A., Fujimoto, M., Chávez, V., 2007. The effect of repetition on imitation from television during infancy. *Dev. Psychobiol.* 49 (2), 196–207. <https://doi.org/10.1002/dev.20208>.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., Ahluwalia, J., 2000. Is there an innate gaze module? Evidence from human neonates. *Infant Behav. Dev.* 23 (2), 223–229. [https://doi.org/10.1016/S0163-6383\(01\)00037-6](https://doi.org/10.1016/S0163-6383(01)00037-6).

- Bazhydai, M., Westermann, G., Parise, E., 2020. "I don't know but I know who to ask": 12-month-olds actively seek information from knowledgeable adults. *Dev. Sci.* 23 (5), 1–10. <https://doi.org/10.1111/desc.12938>.
- Bigelow, A.E., DeCoste, C., 2003. Sensitivity to social contingency from mothers and strangers in 2-, 4-, and 6-month-old infants. *Infancy* 4 (1), 111–140. https://doi.org/10.1207/S15327078IN0401_6.
- Blythe, R.A., Smith, A.D.M., Smith, K., 2016. Word learning under infinite uncertainty [physics and society; computation and language]. *Cognition* 151, 18–27. <https://doi.org/10.1016/j.cognition.2016.02.017>.
- Brand, R.J., Hollenbeck, E., Kominsky, J.F., 2013. Mothers' infant-directed gaze during object demonstration highlights action boundaries and goals. *IEEE Trans. Auton. Mental Dev.* 5 (3), 192–201. <https://doi.org/10.1109/TAMD.2013.2273057>.
- Brooker, I., Poulin-Dubois, D., 2013. Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy* 18, E46–E68. <https://doi.org/10.1111/inf.12027>.
- Brooks, R., Meltzoff, A.N., 2005. The development of gaze following and its relation to language. *Dev. Sci.* 8 (6). <https://doi.org/10.1111/j.1467-7687.2005.00445.x>.
- Brosseau-Liard, P., Penney, D., Poulin-Dubois, D., 2015. Theory of mind selectively predicts preschoolers' knowledge-based selective word learning. *Br. J. Dev. Psychol.* 33 (4), 464–475. <https://doi.org/10.1111/bjdp.12107>.
- Brownell, C.A., 2011. Early developments in joint action. *Rev. Philos. Psychol.* 2 (2), 193–211. <https://doi.org/10.1007/s13164-011-0056-1>.
- Brownell, C.A., 2016. Prosocial behavior in infancy: the role of socialization. *Child Dev. Perspect.* 10 (4), 222–227. <https://doi.org/10.1111/cdep.12189>.
- Buttelmann, D., Zmyj, N., Daum, M., Carpenter, M., 2013. Selective imitation of in-group over out-group members in 14-month-old infants. *Child Dev.* 84 (2), 422–428. <https://doi.org/10.1111/j.1467-8624.2012.01860.x>.
- Byers-Heinlein, K., Fennell, C.T., Werker, J.F., 2013. The development of associative word learning in monolingual and bilingual infants. *Biling. Lang. Cognit.* 16 (1), 198–205. <https://doi.org/10.1017/S1366728912000417>.
- Carpenter, M., Nagell, K., Tomasello, M., 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63 (4), 1–143. <https://doi.org/10.2307/1166214>.
- Charman, T., Carroll, F., Sturge, C., 2001. Theory of mind, executive function and social competence in boys with ADHD. *Emot. Behav. Difficulties* 6 (1), 31–49. <https://doi.org/10.1080/13632750100507654>.
- Chiarella, S.S., Poulin-Dubois, D., 2013. Cry babies and Pollyannas: infants can detect unjustified emotional reactions. *Infancy* 18 (Suppl. 1), 1–16. <https://doi.org/10.1111/inf.12028>.
- Chow, V., Poulin-Dubois, D., Lewis, J., 2008. To see or not to see: infants prefer to follow the gaze of a reliable looker: paper. *Dev. Sci.* 11 (5), 761–770. <https://doi.org/10.1111/j.1467-7687.2008.00726.x>.
- Clerkin, E.M., Hart, E., Rehg, J.M., Yu, C., Smith, L.B., 2017. Real-world visual statistics and infants' first-learned object names. *Phil. Trans. Biol. Sci.* 372 (1711). <https://doi.org/10.1098/rstb.2016.0055>.
- Colunga, E., Smith, L.B., 2005. From the lexicon to expectations about kinds: a role for associative learning. *Psychol. Rev.* 112 (2), 347–382. <https://doi.org/10.1037/0033-295X.112.2.347>.
- Corkum, V., Moore, C., 1995. Development of joint visual. In: *Joint Attention: Its Origins and Role in Development*, vol. 61.
- Corkum, V., Moore, C., 1998. The origins of joint visual attention in infants. *Dev. Psychol.* 34 (1), 28–38. <https://doi.org/10.1111/j.1540-5834.1998.tb02782.x>.
- Creel, S.C., Newport, E.L., Aslin, R.N., 2004. Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 30 (5), 1119–1130. <https://doi.org/10.1037/0278-7393.30.5.1119>.
- Crivello, C., Grossman, S., Poulin-Dubois, D., 2021. Specifying links between infants' theory of mind, associative learning, and selective trust. *Infancy* 26 (5), 664–685. <https://doi.org/10.1111/inf.12407>.
- Crivello, C., Phillips, S., Poulin-Dubois, D., 2018. Selective social learning in infancy: looking for mechanisms. *Dev. Sci.* 21 (3), 1–10. <https://doi.org/10.1111/desc.12592>.
- Csibra, G., Gergely, G., 2006. Social learning and social cognition: the case for pedagogy. In: *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*.
- Davis, J., Redshaw, J., Suddendorf, T., Nielsen, M., Kennedy-Costantini, S., Oostenbroek, J., Slaughter, V., 2021. Does neonatal imitation exist? Insights from a meta-analysis of 336 effect sizes. *Perspect. Psychol. Sci.* 16 (6), 1373–1397. <https://doi.org/10.1177/1745691620959834>.
- de Klerk, C.C.J.M., Filippetti, M.L., Rigato, S., 2021. The development of body representations: an associative learning account. *Proc. Biol. Sci.* 288 (1949). <https://doi.org/10.1098/rspb.2021.0070>.
- D'Entremont, B., Hains, S.M.J., Muir, D.W., 1997. A demonstration of gaze following in 3- to 6-month-olds. *Infant Behav. Dev.* 20 (4), 569–572. [https://doi.org/10.1016/S0163-6383\(97\)90048-5](https://doi.org/10.1016/S0163-6383(97)90048-5).
- Dickerson, K., Gerhardtstein, P., Zack, E., Barr, R., 2013. Age-related changes in learning across early childhood: a new imitation task. *Dev. Psychobiol.* 55 (7), 719–732. <https://doi.org/10.1002/dev.21068>.
- DiYanni, C., Kelemen, D., 2008. Using a bad tool with good intention: young children's imitation of adults' questionable choices. *J. Exp. Child Psychol.* 101 (4), 241–261. <https://doi.org/10.1016/j.jecp.2008.05.002>.
- Duncan, S., Farley, A.M., 1990. Achieving parent-child coordination through conventions. *Child Dev.* 61 (3), 742–753.
- Ellis, E.M., Borovsky, A., Elman, J.L., Evans, J.L., 2021. Toddlers' ability to leverage statistical information to support word learning. *Front. Psychol.* 12 (April), 1–17. <https://doi.org/10.3389/fpsyg.2021.600694>.
- Elsner, B., 2007. Infants' imitation of goal-directed actions: the role of movements and action effects. *Acta Psychol.* 124 (1), 44–59. <https://doi.org/10.1016/j.actpsy.2006.09.006>.
- Erickson, L.C., Thiessen, E.D., 2015. Statistical learning of language: theory, validity, and predictions of a statistical learning account of language acquisition. *Dev. Rev.* 37, 66–108. <https://doi.org/10.1016/j.dr.2015.05.002>.
- Farroni, T., Csibra, G., Simion, F., Johnson, M.H., 2002. Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U. S. A.* 99 (14), 9602–9605. <https://doi.org/10.1073/pnas.152159999>.
- Feinman, S., 1982. Social referencing in infancy. *Merrill-Palmer Q.* 28 (4), 445–470.
- Fiser, J., Aslin, R.N., 2002. Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U. S. A.* 99 (24), 15822–15826. <https://doi.org/10.1073/pnas.232472899>.
- Flynn, E., 2008. Investigating children as cultural magnets: do young children transmit redundant information along diffusion chains? *Phil. Trans. Biol. Sci.* 363 (1509), 3541–3551. <https://doi.org/10.1098/rstb.2008.0136>.
- Gattis, M., Winstanley, A., Sperotto, R., Putnick, D.L., Bornstein, M.H., 2020. Foundations of attention sharing: orienting and responding to attention in term and preterm 5-month-old infants. *Infant Behav. Dev.* 61, 101466. <https://doi.org/10.1016/j.infbeh.2020.101466>.
- Goldstein, M.H., Waterfall, H.R., Lotem, A., Halpern, J.Y., Schwade, J.A., Onnis, L., Edelman, S., 2010. General cognitive principles for learning structure in time and space. *Trends Cognit. Sci.* 14 (6), 249–258. <https://doi.org/10.1016/j.tics.2010.02.004>.
- Goldstone, R.L., Barsalou, L.W., 1998. Reuniting perception and conception. *Cognition* 65 (2–3), 231–262. [https://doi.org/10.1016/S0010-0277\(97\)00047-4](https://doi.org/10.1016/S0010-0277(97)00047-4).
- Graham, S.A., Nilsen, E.S., Collins, S., Olineck, K., 2010. The role of gaze direction and mutual exclusivity in guiding 24-month-olds' word mappings. *Br. J. Dev. Psychol.* 28 (Pt 2), 449–465. <https://doi.org/10.1348/026151009x424565>.
- Guellai, B., Streri, A., 2011. Cues for early social skills: direct gaze modulates newborns' recognition of talking faces. *PLoS One* 6 (4), e18610. <https://doi.org/10.1371/journal.pone.0018610>.

- Hashiya, K., Meng, X., Uto, Y., Tajiri, K., 2018. Overt congruent facial reaction to dynamic emotional expressions in 9-10-month-old infants. *Infant Behav. Dev.* 54, 48–56. <https://doi.org/10.1016/j.infbeh.2018.12.002>.
- Hay, D.F., 2009. The roots and branches of human altruism. *Br. J. Psychol.* 100 (Pt 3), 473–479. <https://doi.org/10.1348/000712609X442096>.
- Hayne, H., Boniface, J., Barr, R., 2000. The development of declarative memory in human infants: age-related changes in deferred imitation. *Behav. Neurosci.* 114 (1), 77–83. <https://doi.org/10.1037//0735-7044.114.1.77>.
- Henderson, A.M.E., Gerson, S., Woodward, A.L., 2008. The birth of social intelligence. *Zero Three* 28 (5), 13–20.
- Henderson, A.M.E., Graham, S.A., Schell, V., 2015. 24-Month-Olds' selective learning is not an all-or-none phenomenon. *PLoS One* 10 (6), e0131215. <https://doi.org/10.1371/journal.pone.0131215>.
- Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? *Behav. Brain Sci.* 33 (2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>.
- Heyes, C., 2017. Rattling the cage and opening the door. *Dev. Sci.* 20 (2). <https://doi.org/10.1111/desc.12416>.
- Heyes, C.M., 1994. Social learning in animals: categories and mechanisms. *Biol. Rev. Camb. Phil. Soc.* 69 (2), 207–231. <http://www.ncbi.nlm.nih.gov/pubmed/8054445>.
- Hipp, D., Gerhardstein, P., Zimmermann, L., Moser, A., Taylor, G., Barr, R., 2017. The dimensional divide: learning from TV and touchscreens during early childhood. In: *Media exposure during infancy and early childhood*. Springer, Cham, pp. 33–54.
- Hollich, G.J., Hirsh-pasek, K., Golinkoff, R.M., 2000. Breaking the Language Barrier: An Emergentist Coalition Model for the Origins of Word Learning, vol. 65. Monographs of the Society for Research in Child Development.
- Hood, B.M., Willen, J.D., Driver, J., 1998. Adult's eyes trigger shifts of visual attention in human infants. *Psychol. Sci.* 9 (2), 131–134. <https://doi.org/10.1111/1467-9280.00024>.
- Howard, L.H., Carrazza, C., Woodward, A.L., 2014. Neighborhood linguistic diversity predicts infants' social learning. *Cognition* 133 (2), 474–479. <https://doi.org/10.1016/j.cognition.2014.08.002>.
- Howard, L.H., Henderson, A.M.E., Carrazza, C., Woodward, A.L., 2015. Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Dev.* 86 (1), 259–275. <https://doi.org/10.1111/cdev.12299>.
- Huang, C.-T., Charman, T., 2005. Gradations of emulation learning in infants' imitation of actions on objects. *J. Exp. Child Psychol.* 92 (3), 276–302. <https://doi.org/10.1016/j.jecp.2005.06.003>.
- Isomura, T., Nakano, T., 2016. Automatic facial mimicry in response to dynamic emotional stimuli in five-month-old infants. *Proc. Biol. Sci.* 283 (1844). <https://doi.org/10.1098/rspb.2016.1948>.
- Johnson, S.P., Amso, D., Slemmer, J.A., 2003. Development of object concepts in infancy: evidence for early learning in an eye-tracking paradigm. *Proc. Natl. Acad. Sci. U. S. A.* 100 (18), 10568–10573. <https://doi.org/10.1073/pnas.1630655100>.
- Jones, S.S., 2009. The development of imitation in infancy. *Phil. Trans. Biol. Sci.* 364 (1528), 2325–2335. <https://doi.org/10.1098/rstb.2009.0045>.
- Jusczyk, P.W., Aslin, R.N., 1995. Infants' detection of the sound patterns of words in fluent speech. *Cognit. Psychol.* 29 (1), 1–23. <https://doi.org/10.1006/cogp.1995.1010>.
- Keupp, S., Behne, T., Rakoczy, H., 2013. Why do children overimitate? Normativity is crucial. *J. Exp. Child Psychol.* 116 (2), 392–406. <https://doi.org/10.1016/j.jecp.2013.07.002>.
- Keupp, S., Behne, T., Rakoczy, H., 2018. The Rationality of (Over) Imitation. <https://doi.org/10.1177/1745691618794921>.
- Kirkham, N.Z., Slemmer, J.A., Johnson, S.P., 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83 (2), B35–B42. [https://doi.org/10.1016/s0010-0277\(02\)00004-5](https://doi.org/10.1016/s0010-0277(02)00004-5).
- Kirkorian, H.L., Choi, K., Pempek, T.A., 2016. Toddlers' word learning from contingent and noncontingent video on touch screens. *Child Dev.* 87 (2), 405–413. <https://doi.org/10.1111/cdev.12508>.
- Koenig, M. A., Clément, F., Harris, P.L., Clément, F., 2004. Trust in testimony: children's use of true and false statements. *Psychol. Sci.* 15 (10), 694–698. <https://doi.org/10.1111/j.0956-7976.2004.00742.x>.
- Koenig, M.A., Woodward, A.L., 2010. Sensitivity of 24-month-olds to the prior inaccuracy of the source: possible mechanisms. *Dev. Psychol.* 46 (4), 815–826. <https://doi.org/10.1037/a0019664>.
- Krogh, L., Vlach, H.A., Johnson, S.P., 2013. Statistical learning across development: flexible yet constrained. *Front. Psychol.* 3 (JAN), 1–11. <https://doi.org/10.3389/fpsyg.2012.00598>.
- Krogh-Jespersen, S., Echols, C.H., 2012. The influence of speaker reliability on first versus second label learning. *Child Dev.* 83 (2), 581–590. <https://doi.org/10.1111/j.1467-8624.2011.01713.x>.
- Kuzyk, O., Grossman, S., Poulin-Dubois, D., 2020. Knowing who knows: metacognitive and causal learning abilities guide infants' selective social learning. *Dev. Sci.* 23 (3), 1–14. <https://doi.org/10.1111/desc.12904>.
- Langton, S.R.H., Watt, R.J., Bruce, V., 2000. Do the eyes have it? Cues to the direction of social attention. *Trends Cognit. Sci.* 4 (2), 50–59. [https://doi.org/10.1016/S1364-6613\(99\)01436-9](https://doi.org/10.1016/S1364-6613(99)01436-9).
- Legare, C.H., 2017. Cumulative cultural learning: development and diversity. *Proc. Natl. Acad. Sci. U. S. A.* 114 (30), 7877–7883. <https://doi.org/10.1073/pnas.1620743114>.
- Levine, D., Buchsbaum, D., Hirsh-Pasek, K., Golinkoff, R.M., 2019. Finding events in a continuous world: a developmental account. *Dev. Psychobiol.* 61 (3), 376–389. <https://doi.org/10.1002/dev.21804>.
- Lyons, D.E., Young, A.G., Keil, F.C., 2007. The hidden structure of overimitation. *Proc. Natl. Acad. Sci. U. S. A.* 104 (50), 19751–19756. <https://doi.org/10.1073/pnas.0704452104>.
- Matheson, H., Moore, C., Akhtar, N., 2013. The development of social learning in interactive and observational contexts. *J. Exp. Child Psychol.* 114 (2), 161–172. <https://doi.org/10.1016/j.jecp.2012.09.003>.
- Maye, J., Werker, J.F., Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82 (3), B101–B111. [https://doi.org/10.1016/s0010-0277\(01\)00157-3](https://doi.org/10.1016/s0010-0277(01)00157-3).
- Maye, J., Weiss, D.J., Aslin, R.N., 2008. Statistical phonetic learning in infants: facilitation and feature generalization. *Dev. Sci.* 11 (1), 122–134. <https://doi.org/10.1111/j.1467-7687.2007.00653.x>.
- Medina, T.N., Snedeker, J., Trueswell, J., Gleitman, L., 2011. How words can and cannot be learned by observation. *Proc. Natl. Acad. Sci. U. S. A.* 108 (22), 9014–9019. <https://doi.org/10.1073/pnas.1105040108>.
- Meltzoff, A.N., 1985. Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Dev.* 56 (1), 62–72. <https://doi.org/10.1111/j.1467-8624.1985.tb00086.x>.
- Meltzoff, A.N., 2010. Social cognition and the origins of imitation, empathy, and theory of mind. In: *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, second ed. <https://doi.org/10.1002/9781444325485.ch2> Issue July 2010.
- Meltzoff, A.N., Brooks, R., 2008. Self-experience as a mechanism for learning about others: a training study in social cognition. *Dev. Psychol.* 44 (5), 1257–1265. <https://doi.org/10.1037/a0012888>.
- Meltzoff, A.N., Moore, M.K., 1977. Imitation of facial and manual gestures by human neonates. *Science* 198 (4312), 75–78. <https://doi.org/10.1126/science.198.4312.75>.
- Mills, C.M., Elashi, F.B., 2014. Children's skepticism: developmental and individual differences in children's ability to detect and explain distorted claims. *J. Exp. Child Psychol.* 124, 1–17. <https://doi.org/10.1016/j.jecp.2014.01.015>.
- Monroy, C., Meyer, M., Gerson, S., Hunnius, S., 2017. Statistical learning in social action contexts. *PLoS One* 1–20. <https://doi.org/10.1371/journal.pone.0177261>.
- Moore, C., 1996. Theories of mind in infancy. *Br. J. Dev. Psychol.* 14 (1), 19–40. <https://doi.org/10.1111/j.2044-835X.1996.tb00691.x>.
- Moore, C., Corkum, V., 1998. Infant gaze following based on eye direction. *Br. J. Dev. Psychol.* 16 (4), 495–503. <https://doi.org/10.1111/j.2044-835X.1998.tb00767.x>.
- Morrison, S., Henderson, A.M.E., Sagar, M., Kennedy-costantini, S., Adams, J., 2021. Infant behavior and development peek-a-who? Exploring the dynamics of early communication with an interactive partner swap paradigm and state space grid visualization. *Infant Behav. Dev.* 64 (July 2020), 101576. <https://doi.org/10.1016/j.infbeh.2021.101576>.

- Moser, A., Zimmermann, L., Dickerson, K., Grenell, A., Barr, R., Gerhardstein, P., 2015. They can interact, but can they learn? Toddlers' transfer learning from touchscreens and television. *J. Exp. Child Psychol.* 137, 137–155. <https://doi.org/10.1016/j.jecp.2015.04.002>.
- Mundy, P., Newell, L., 2007. Attention, joint attention, and social cognition. *Curr. Dir. Psychol. Sci.* 16 (5), 269–274.
- Mundy, P., van Hecke, A., 2008. Neural systems, gaze following, and the development of joint attention. In: *Handbook of Developmental Cognitive Neuroscience*, second ed. MIT Press, pp. 819–837.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Vaughan, A., Hecke, V., Parlade, M.V., 2007. Individual differences and the development of joint attention in infancy. *Child Dev* 78 (3), 938–954. <https://doi.org/10.1111/j.1467-8624.2007.01042.x>.
- Nielsen, M., 2006. Copying actions and copying outcomes: social learning through the second year. *Dev. Psychol.* 42 (3), 555–565. <https://doi.org/10.1037/0012-1649.42.3.555>.
- Over, H., Carpenter, M., 2012. Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *J. Comp. Psychol.* 126 (2), 182–192. <https://doi.org/10.1037/a0024555>.
- Plunkett, K., 1997. Theories of early language acquisition. *Trends Cognit. Sci.* 1 (4), 146–153. [https://doi.org/10.1016/S1364-6613\(97\)01039-5](https://doi.org/10.1016/S1364-6613(97)01039-5).
- Poulin-dubois, D., Brosseau-liard, P., 2016. The developmental origins of selective social learning. *Curr. Dir. Psychol. Sci.* 25 (1), 60–64. <https://doi.org/10.1177/0963721415613962>.
- Poulin-Dubois, D., Lepage, A., Ferland, D., 1996. Infants' concept of animacy. *Cognit. Dev.* 11 (1), 19–36. [https://doi.org/10.1016/S0885-2014\(96\)90026-X](https://doi.org/10.1016/S0885-2014(96)90026-X).
- Quine, W.V.O., 2013. *Word and Object*. MIT Press.
- Quine, W., Van, O., 1960. Word and object: an inquiry into the linguistic mechanisms of objective reference. In: *Word and Object: An Inquiry Into the Linguistic Mechanisms of Objective Reference*. John Wiley.
- Quinn, P.C., Eimas, P.D., 2000. The emergence of category representations during infancy: are separate perceptual and conceptual processes required? *J. Cognit. Dev.* 1 (1), 55–61. https://doi.org/10.1207/S15327647JCD0101N_6.
- Quinn, P.C., Lee, K., Pascalis, O., Tanaka, J.W., 2016. Narrowing in categorical responding to other-race face classes by infants. *Dev. Sci.* 19 (3), 362–371. <https://doi.org/10.1111/desc.12301>.
- Rakison, D., 2022. Fear Learning in Infancy: An Evolutionary Developmental Perspective, pp. 303–323. https://doi.org/10.1007/978-3-030-76000-7_14.
- Rakison, D.H., 2004. Infants' sensitivity to correlations between static and dynamic features in a category context. *J. Exp. Child Psychol.* 89 (1), 1–30. <https://doi.org/10.1016/j.jecp.2004.06.001>.
- Rakison, D.H., Lupyan, G., Oakes, L.M., Walker-Andrews, A.S., 2008. Developing object concepts in infancy: an associative learning perspective. *Monogr. Soc. Res. Child Dev.* 73 (1), 1–127.
- Reddy, V., 2003. On being the object of attention: implications for self-other consciousness. *Trends Cognit. Sci.* 7 (9), 397–402. [https://doi.org/10.1016/s1364-6613\(03\)00191-8](https://doi.org/10.1016/s1364-6613(03)00191-8).
- Rendell, L., Fogarty, L., Hoppitt, W.J.E., Morgan, T.J.H., Webster, M.M., Laland, K.N., 2011. Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cognit. Sci.* 15 (2), 68–76. <https://doi.org/10.1016/j.tics.2010.12.002>.
- Rogoff, B., Dahl, A., Callanan, M., 2018. The importance of understanding children's lived experience. *Dev. Rev.* 50 (June), 5–15. <https://doi.org/10.1016/j.dr.2018.05.006>.
- Romberg, A.R., Saffran, J.R., 2010. Statistical learning and language acquisition. *Wiley Interdiscip. Rev.* 1 (6), 906–914. <https://doi.org/10.1002/wcs.78>.
- Rosa Salva, O., Farroni, T., Regolin, L., Vallortigara, G., Johnson, M.H., 2011. The evolution of social orienting: evidence from chicks (*Gallus gallus*) and human newborns. *PLoS One* 6 (4), e18802.
- Ruffman, T., Taumoepeau, M., Perkins, C., 2012. Statistical learning as a basis for social understanding in children. *Br. J. Dev. Psychol.* 30 (1), 87–104. <https://doi.org/10.1111/j.2044-835X.2011.02045.x>.
- Saffran, J.R., 2020. Statistical language learning in infancy. *Child Dev. Perspect.* 14 (1), 49–54. <https://doi.org/10.1111/cdev.12355>.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274 (5294), 1926–1928.
- Saffran, J.R., Kirkham, N.Z., 2018. Infant statistical learning. *Annu. Rev. Psychol.* 69, 181–203. <https://doi.org/10.1146/annurev-psych-122216-011805>.
- Scaife, M., Bruner, J.S., 1975. The capacity for joint visual attention in the infant. *Nature* 253 (5489), 265–266. <https://doi.org/10.1038/253265a0>.
- Schroer, S.E., Smith, L.B., Yu, C., 2019. Examining the multimodal effects of parent speech in parent-infant interactions. *CogSci* 1015–1021.
- Schuhmacher, N., Köster, M., Kärtnner, J., 2019. Modeling prosocial behavior increases helping in 16-month-olds. *Child Dev.* 90 (5), 1789–1801. <https://doi.org/10.1111/cdev.13054>.
- Senju, A., Csibra, G., 2008. Gaze following in human infants depends on communicative signals. *Curr. Biol.* 18 (9), 668–671. <https://doi.org/10.1016/J.CUB.2008.03.059>.
- Simion, F., di Giorgio, E., Leo, I., Bardi, L., 2011. The processing of social stimuli in early infancy: from faces to biological motion perception. *Prog. Brain Res.* 189, 173–193. <https://doi.org/10.1016/B978-0-444-53884-0.00024-5>.
- Singh, L., Cristia, A., Karasik, L.B., Oakes, L., 2021. Mainstream Research in Early Development Remains Non-diverse: Barriers and Bridges Towards a Diversified Science. <https://doi.org/10.31234/osf.io/hgukc>.
- Slaughter, V., 2021. Do newborns have the ability to imitate? *Trends Cognit. Sci.* 25 (5), 377–387. <https://doi.org/10.1016/j.tics.2021.02.006>.
- Slaughter, V., McConnell, D., 2003. Emergence of joint attention: relationships between gaze following, social referencing, imitation, and naming in infancy. *J. Genet. Psychol.* 164 (1), 54–71. <https://doi.org/10.1080/00221320309597503>.
- Sloutsky, V.M., Deng, W., 2019. Categories, concepts, and conceptual development. *Lang. Cogn. Neurosci.* 34 (10), 1284–1297. <https://doi.org/10.1080/23273798.2017.1391398>.
- Smith, L., 2000. How to learn words: an associative crane. In: *Becoming a Word Learner: A Debate on Lexical Acquisition*.
- Smith, L.B., Jayaraman, S., Clerkin, E., Yu, C., 2018. The developing infant creates a curriculum for statistical learning. *Trends Cognit. Sci.* 22 (4), 325–336. <https://doi.org/10.1016/j.tics.2018.02.004>.
- Smith, L.B., Jones, S.S., Yoshida, H., Colunga, E., 2003. Whose DAM account? Attentional learning explains booth and waxman. *Cognition* 87 (3), 209–213. [https://doi.org/10.1016/s00100277\(02\)00236-6](https://doi.org/10.1016/s00100277(02)00236-6).
- Sobel, D.M., Kushnir, T., 2013. Knowledge matters: how children evaluate the reliability of testimony as a process of rational inference. *Psychol. Rev.* 120 (4), 779–797. <https://doi.org/10.1037/a0034191>.
- Stenberg, G., 2009. Selectivity in infant social referencing. *Infancy* 14 (4), 457–473. <https://doi.org/10.1080/15250000902994115>.
- Stenberg, G., 2013. Do 12-month-old infants trust a competent adult? *Infancy* 18 (5), 873–904. <https://doi.org/10.1111/inf.12011>.
- Striano, T., Henning, A., Stahl, D., 2005. Sensitivity to social contingencies between 1 and 3 months of age. *Dev. Sci.* 8 (6), 509–518. <https://doi.org/10.1111/j.1467-7687.2005.00442.x>.
- Striano, T., Stahl, D., 2005. Sensitivity to triadic attention in early infancy. *Dev. Sci.* 8 (4), 333–343. <https://doi.org/10.1111/j.1467-7687.2005.00421.x>.
- Suanda, S.H., Barnhart, M., Smith, L.B., Yu, C., 2019. The signal in the noise: the visual ecology of parents' object naming. *Infancy* 24 (3), 455–476. <https://doi.org/10.1111/inf.12278>.
- Suarez-Rivera, C., Smith, L.B., Yu, C., 2019. Multimodal parent behaviors within joint attention support sustained attention in infants. *Dev. Psychol.* 55 (1), 96–109. <https://doi.org/10.1037/dev0000628>.
- Szafnarowska, J., Rohlfing, K.J., Fawcett, C., Gredebäck, G., 2014. Is ostension any more than attention? *Sci. Rep.* 4. <https://doi.org/10.1038/srep05304>.
- Tamis-LeMonda, C.S., Kuchirko, Y., Song, L., 2014. Why is infant language learning facilitated by parental responsiveness? *Curr. Dir. Psychol. Sci.* 23 (2), 121–126. <https://doi.org/10.1177/0963721414522813>.
- Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D., 2011. How to grow a mind: statistics, structure, and abstraction. *Science* 331 (6022), 1279–1285. <https://doi.org/10.1126/science.1192788>.

- Thiele, M., Hepach, R., Michel, C., Gredebäck, G., Haun, D.B.M., 2021. Social interaction targets enhance 13-month-old infants' associative learning. *Infancy* 26 (3), 409–422. <https://doi.org/10.1111/inf.12393>.
- Thiessen, E.D., Kronstein, A.T., Hufnagle, D.G., 2013. The extraction and integration framework: a two-process account of statistical learning. *Psychol. Bull.* 139 (4), 792–814. <https://doi.org/10.1037/a0030801>.
- Thiessen, E.D., Yee, M.N., 2010. Dogs, bogs, labs, and lads: what phonemic generalizations indicate about the nature of children's early word-form representations. *Child Dev.* 81 (4), 1287–1303. <http://www.jstor.org/stable/40801474>.
- Tomasello, M., Hare, B., Lehmann, H., Call, J., 2007. Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *J. Hum. Evol.* 52 (3), 314–320. <https://doi.org/10.1016/j.jhevol.2006.10.001>.
- Tomasello, M., Kruger, A.C., Ratner, H.H., 2010. Cultural learning. *Behav. Brain Sci.* 16 (03), 495. <https://doi.org/10.1017/S0140525X0003123X>.
- Tummeltshammer, K., Feldman, E.C.H., Amso, D., 2019. Using pupil dilation, eye-blink rate, and the value of mother to investigate reward learning mechanisms in infancy. *Dev. Cogn. Neurosci.* 36, 100608. <https://doi.org/10.1016/j.dcn.2018.12.006>.
- Tummeltshammer, K.S., Wu, R., Sobel, D.M., Kirkham, N.Z., 2014. Infants track the reliability of potential informants. *Psychol. Sci.* <https://doi.org/10.1177/0956797614540178>.
- Vaillant-Molina, M., Bahrick, L.E., 2012. The role of intersensory redundancy in the emergence of social referencing in 5½-month-old infants. *Dev. Psychol.* 48 (1), 1–9. <https://doi.org/10.1037/a0025263>.
- Vernetti, A., Smith, T.J., Senju, A., 2017. Gaze-contingent reinforcement learning reveals incentive value of social signals in young children and adults. *Proc. Biol. Sci.* 284 (1850). <https://doi.org/10.1098/rspb.2016.2747>.
- Walden, T.A., Ogan, T.A., 1988. The development of social referencing. *Child Dev.* 59 (5), 1230–1240.
- Want, S.C., Harris, P.L., 2002. How do children ape? Applying concepts from the study of non-human primates to the developmental study of "imitation" in children. *Dev. Sci.* 5 (1), 1–14. <https://doi.org/10.1111/1467-7687.00194>.
- Werchan, D.M., Amso, D., 2020. Top-down knowledge rapidly acquired through abstract rule learning biases subsequent visual attention in 9-month-old infants. *Dev. Cogn. Neurosci.* 42, 100761. <https://doi.org/10.1016/j.dcn.2020.100761>.
- Werchan, D.M., Amso, D., 2021. All contexts are not created equal: social stimuli win the competition for organizing reinforcement learning in 9-month-old infants. *Dev. Sci.* 24 (5), 1–10. <https://doi.org/10.1111/desc.13088>.
- Werchan, D.M., Collins, A.G.E., Frank, M.J., Amso, D., 2015. 8-month-old infants spontaneously learn and generalize hierarchical rules. *Psychol. Sci.* 26 (6), 805–815. <https://doi.org/10.1177/0956797615571442>.
- Werker, J.F., Cohen, L.B., Lloyd, V.L., Casasola, M., Stager, C.L., 1998. Acquisition of word-object associations by 14-month-old infants. *Dev. Psychol.* 34 (6), 1289–1309. <https://doi.org/10.1037/0012-1649.34.6.1289>.
- Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., Amano, S., 2007. Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* 103 (1), 147–162. <https://doi.org/10.1016/j.cognition.2006.03.006>.
- Woodward, A., Hoeyne, K., 1999. Infants' learning about words and sounds in relation to objects. *Child Dev.* 70 (1), 65–77. <https://doi.org/10.1111/1467-8624.00006>.
- Woodward, A.L., 2003. Infants' developing understanding of the link between looker and object. *Dev. Sci.* 6 (3), 297–311. <https://doi.org/10.1111/1467-7687.00286>.
- Woodward, A.L., Sommerville, J.A., Gerson, S., Henderson, A.M.E., Buresh, J., 2009. The emergence of intention attribution in infancy. *Psychol. Learn. Motiv. - Adv. Res. Theory* 51, 187–222.
- Wu, R., Gopnik, A., Richardson, D.C., Kirkham, N.Z., 2011. Infants learn about objects from statistics and people. *Dev. Psychol.* 47 (5), 1220–1229. <https://doi.org/10.1037/a0024023>.
- Wu, R., Tummeltshammer, K.S., Gliga, T., Kirkham, N.Z., 2014. Ostensive signals support learning from novel attention cues during infancy. *Front. Psychol.* 5, 251. <https://doi.org/10.3389/fpsyg.2014.00251>.
- Xiao, N.G., Wu, R., Quinn, P.C., Liu, S., Tummeltshammer, K.S., Kirkham, N.Z., Ge, L., Pascalis, O., Lee, K., 2018. Infants rely more on gaze cues from own-race than other-race adults for learning under uncertainty. *Child Dev.* 89 (3), e229–e244. <https://doi.org/10.1111/cdev.12798>.
- Xiao, W.S., Quinn, P.C., Pascalis, O., Lee, K., 2014. Own- and other-race face scanning in infants: implications for perceptual narrowing. *Dev. Psychobiol.* 56 (2), 262–273. <https://doi.org/10.1002/dev.21196>.
- Younger, B.A., Cohen, L.B., 1986. Developmental change in infants' perception of correlations among attributes. *Child Dev.* 57 (3), 803–815.
- Yu, C., Smith, L.B., 2012. Embodied attention and word learning by toddlers. *Cognition* 125 (2), 244–262. <https://doi.org/10.1016/j.cognition.2012.06.016>.
- Yu, C., Smith, L.B., 2016. The social origins of sustained attention in one-year-old human infants. *Curr. Biol.* 26 (9), 1235–1240. <https://doi.org/10.1016/j.cub.2016.03.026>.
- Zmyj, N., Buttelmann, D., Carpenter, M., Daum, M.M., 2010. The reliability of a model influences 14-month-olds' imitation. *J. Exp. Child Psychol.* 106 (4), 208–220. <https://doi.org/10.1016/j.jecp.2010.03.002>.

Early language development: major milestones and individual differences

Michelle K. Tulloch and Erika Hoff, Florida Atlantic University, Davie, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	467
Major milestones in early language development	468
The first year: discriminating speech sounds, finding words, and communicating	468
Discriminating speech sounds	468
Producing speech sounds	468
Word segmentation	468
Beginnings of intentional communication	469
Second year: the vocabulary spurt and the onset of combinatorial speech	469
The first word and the word spurt	469
Building semantic networks	470
Combining words	470
Expanding communicative purposes	470
From 24 to 36 months: mastering grammar and developing discourse skills	470
First sentences	470
The development of conversation and narrative skill	471
After the third birthday: further developing skills	471
Complex, multiclausal utterances	471
Continued growth and influences from schooling	471
Individual differences in early language development	471
The nature of individual differences	471
Internal sources of individual differences	471
External sources of individual differences	471
The social context and language development	472
Socioeconomic status and language development	472
Multilingual exposure and language development	472
Summary and conclusion	473
Acknowledgment	473
References	473
Further reading	476

Introduction

In four short years, newborn infants go from no apparent linguistic competence to mastery of the basics of the language or, in some cases more than one language, that is spoken around them. Apart from extreme atypicality of endowment or environment, all children acquire language. One aim of the study of language development is to understand this universal human achievement. However, children do not all learn at the same rate or reach the same level of proficiency. A second aim of research on language development is to understand the sources of individual differences. Because language is central to most human activity, the individual differences in children's language skills translate to individual differences in school success (Von Stumm et al., 2020) and later life outcomes (Chetty et al., 2010). Thus, the study of language development is both a scientific exploration of human nature and a practical search for the factors that support the fulfillment of human potential.

The core areas of knowledge and skill that children acquire in learning to understand and produce language have been usefully divided by the field of linguistics into four domains: phonology, the lexicon, morphology, and syntax. Phonology is the sound system: each language makes use of a repertoire of speech sounds, with regularities in which differences between sounds signal differences in meaning and regularities in how speech sounds can be combined to form words. The lexicon is the mental repository of word forms and their associated meanings. Morphology is word structure and the regularities in forming words by combining word stems with small, meaningful word pieces such as the sounds that mark a noun as a plural or a verb as past tense. Syntax comprises the system for combining words into meaningful and well-formed sentences. Grammar is typically considered the combination of morphology and syntax, although strictly speaking the regularities in phonological systems are part of grammar as well. When children have mastered these four components (phonology, lexicon, morphology, and syntax), they are considered linguistically competent, but they must also learn to be communicatively competent, which means communicating in socially appropriate ways. Communicative competence includes pragmatic and sociolinguistic competence. Pragmatics is the knowledge of how to

communicate effectively in order to “do things with words” (Austin, 1962; Searle, 1969) such as to make requests, to comment, or to recount narratives. Sociolinguistics concerns the social rules of language use.

This article provides an introduction to the research on language development in the first four years of life that has described the major milestones children achieve in acquiring language and some of the factors that create individual differences in the rate and outcome of those processes.

Major milestones in early language development

The first year: discriminating speech sounds, finding words, and communicating

Discriminating speech sounds

From birth, infants are able to discriminate (i.e., tell the difference between) all speech sounds, or phones, in any language. With experience, infants learn to ignore differences that do not matter. For example, speech sounds differ depending on the other sounds produced before and after them (this is known as coarticulation effects), and sounds are also different when produced by different speakers. These small variations in naturally occurring speech do not disrupt the infant’s ability to recognize the sounds in speech. In laboratory settings, by about 10.5 months of age, infants ignore some aspects of speech variation from unfamiliar speakers to learn new words (Houston and Jusczyk, 2000; Singh et al., 2004). But, in the cases where variation is greater, children’s ability to recognize words is affected. For example, an American English speaker may hear “bowl” when an Australian English speaker produces “ball” (van Heugten and Johnson, 2014). In experimental studies, at 12 months of age but not at 9 months of age, infants are able to recognize words produced in an unfamiliar, regional dialect (Schmale et al., 2010). Only by 13–15 months of age can infants recognize words produced in an unfamiliar, foreign accent (i.e., Spanish-accented English; Best et al., 2009; Schmale and Seidl, 2009; van Heugten and Johnson, 2014). This ability to recognize words in unfamiliar, accented speech increases with age and experience (van Heugten et al., 2015).

The particular sound contrasts that mark meaning distinctions differ somewhat from language to language. As infants gain experience hearing one language, they improve their abilities to discriminate the sound contrasts that language uses and lose the ability to hear sound contrasts that are not used in their language (Kuhl et al., 1992, 2006; Werker and Tees, 1984). This results in the difficulty adult learners of a new language often have in hearing distinctions that are not used in the language they know.

Producing speech sounds

Before infants produce adult-like speech sounds, their vocal productions follow a fairly regular course of pre-speech development. In the first few months of life, infants begin cooing. Coos are vowel-like speech sounds that start as single sounds and eventually develop into multiple vowel-like sounds strung together into one coo. In the next stage, infants incorporate consonant-like sounds to produce marginal babbling. This period of time is sometimes referred to as the period of vocal play (Stark, 1986), as infants try out a variety of sounds that they can produce in combination. Between 6 and 9 months of age, infants produce canonical babbles, which consist of distinct, repeated syllables. This stage of babbling is also described as reduplicated babbling because infants often continuously produce the same consonant-vowel syllables in a series (*mama*). Following this period of reduplicated babbling, non-reduplicated or variegated babbling emerges. In this phase, infants produce sequences of different consonant and vowel pairs—often in long strings that mimic the intonation of adult speech, also termed prosody. This type of babbling is often called jargon, and infants vary widely in how much and for how long they produce jargon (Dore, 1975).

The ability to produce recognizable speech sounds depends, in part, on physical growth of the organs in the mouth and throat that are involved in speech sound production and also on the development of motor control over these articulators. The influence of these biologically paced developments makes some speech sounds earlier developments than others across languages (e.g., /m/ and /d/ are early sounds; /s/ and /θ/ “th” are later acquired sounds). However, the order in which speech sounds are produced is also influenced by the language the children are acquiring (Stokes and Surendran, 2005; van Severen et al., 2013). Children master the production of speech sounds earlier if those sounds are involved in contrasts that distinguish different words in their language (Pye et al., 1987). Most children produce adult-like speech sounds by around 8 years (Hulit and Howard, 2006; Owens, 2006).

Word segmentation

One early challenge in language acquisition is identifying individual words within the speech stream. Even for adults, it is nearly impossible to identify when one word ends and the next begins when listening to natural speech in a completely unfamiliar language. This challenge is made somewhat easier for infants because of the way adults speak to children. Infant- and child-directed speech, sometimes termed motherese or parentese, is a style of speech adopted by parents (and others) when speaking directly to infants and young children. This speech differs from speech addressed to adults. It is slower, clearer, with shorter utterances and longer pauses between utterances, and the pitch is higher (Fernald et al., 1989; Fernald and Simon, 1984; Gerken, 1994). Infants prefer to listen to this speech style over adult-directed speech when given a choice in experimental settings (Carey, 1978; Markman et al., 2003; Mervis and Bertrand, 1994) and are better able to isolate individual words in parentese than in adult-directed speech (Ma et al., 2011). Also, in child-directed speech words are often presented in isolation (Lew-Williams et al., 2011) or with clear intonation differences (Thiessen et al., 2005), and this helps infants find the word boundaries. Once an infant can identify recurring sequences as words, they can start to understand what those words refer to. As early as 6 months of age, children can recognize their names and a few common words (Bergelson and Swingle, 2012; Tincoff and Jusczyk, 1999).

In addition to benefitting from the support provided by features of infant-directed speech, infants make use of statistical learning to identify words. From as young as 5 to 8 months of age, infants are able to learn the statistical regularities in the speech they hear and apply that information to new streams of speech, allowing infants to distinguish frequently occurring sound sequences from infrequent sound sequences. Frequent sound sequences are likely to be single words, whereas infrequent sequences are more likely to include a boundary between different words (Saffran et al., 1996; see Saffran and Kirkham, 2018 for a review of the topic). In addition, infants learn language-specific cues that aid in finding word boundaries. For example, in English, stressed syllables are typically on the onset of a word. With enough experience, by about 9 months of age, English learning infants will learn this regularity and use this as a cue of a new word (Thiessen and Saffran, 2003, 2007).

Beginnings of intentional communication

Infants' first sounds are not considered intentional communication (Bates et al., 1975). Initially, when infants cry, fuss, or smile, these behaviors are reflexive and do not arise from an intent to communicate with others (Oller, 2000). An important milestone in early communicative development is the onset of joint attention. Early in development, infants will attend to objects or to people in their environment, but, by around 9–12 months of age, infants develop the ability to point, gesture, and to share attention with another person. This sharing of attention to a third entity with another person is termed joint attention (Brooks and Meltzoff, 2005), and it is a capacity that sets the stage for language learning. Infants who engage in more joint attention with adults advance faster in developing language (Camaioni and Perucchini, 2003; Carpenter et al., 1998; Mundy et al., 2007). Importantly, children who engage in less joint attention may also hear less speech addressed directly to them, which also affects their language development (Hoff, 2006). We will discuss differences in the home environment leading to individual differences in **The nature of individual differences** section.

By 12 months of age, children point with intention, expecting that an adult will look and engage with the item or direction the infant points (Tomasello et al., 2007). Infants will even point in order to help an adult who is searching for something the infant can see (Tomasello, 2007). The frequency of infants' communicative pointing also predicts their subsequent language development (Camaioni and Perucchini, 2003; Iverson and Goldin-Meadow, 2005; Rowe et al., 2008).

Second year: the vocabulary spurt and the onset of combinatorial speech

The first word and the word spurt

Infants typically have identified the word boundaries in speech, have started to understand the meaning of words, and have used gestures with communicative intent before they produce recognizable speech. Children tend to produce their first word around their first birthday, although the range of variability among children is wide and it is not unusual for first words to appear as early as 10 months and as late as 14 months. The words that children first produce are different from words produced by adults. Adults' words are label for categories (i.e., *duck* refers to all ducks, not just one specific toy duck). For children, the words they first produce may relate to a specific event, game, or item. This label used by children does not necessarily encompass other events, games, or items that would have the same label, and therefore it is considered context-bound (Barrett, 1986; Bloom, 1973; Caselli et al., 1995). With time and experience, words become decontextualized and children start to use the same label for other items or events that share similarities (Barrett, 1986; Bates et al., 1979; Dore et al., 1976). When decontextualizing newly learned words, children may overextend and underextend word meanings. For example, a child may overextend the meaning of "kitty" to mean any four legged, furry pet or they may underextend the meaning of "doggy" to just include their pet but not any dog that passes by them in the park. While these types of speech errors may be very obvious to an adult, they actually do not occur too often in children's speech (Rescorla, 1980) and they may not be due to an inability to understand the meaning of words (Hoek et al., 1986; Hudson and Nelson, 1984; Huttenlocher and Smiley, 1987). Instead, it seems that the child may not know or remember the appropriate label for a referent, and they select the next closest thing in their mental lexicon. In general, each word within the lexicon is an entry that contains some semantic, phonologic, and grammatical information. As a child hears a word for the first time, they store some information about that word. With repeated exposure to that word, in different contexts and with different speakers, children learn the meaning associated with the word as well as the way to produce the word accurately. This description is an oversimplification of a complex process, and there are many theories that discuss how the mental lexicon works, how words are stored, and how words are retrieved for production (see de Sousa and Gabriel, 2015 for a review). The next section discusses in more detail how semantic information is stored within a semantic network.

After a child has produced their first word, more vocabulary follows. By around 18 months of age, children typically have a productive vocabulary of around 50 words (Nelson, 1973). The particular words in a child's vocabulary vary depending on their own experiences (Clark, 1979), but across children, early productive vocabularies tend to be dominated by nouns, with fewer words from other grammatical categories (Caselli et al., 1995; Nelson, 1973). For many, but not all children, the 50-word milestone is followed by a period of rapid lexical growth, termed the word spurt. During this period, children, on average, increase their rate of word learning from around 8–11 words each month to 22–37 words per month (Benedict, 1979). Learning the meaning of a new word often involves an initial fast, but partial, mapping of a meaning onto the word based on the context (Carey, 1978; Markman et al., 2003; Mervis and Bertrand, 1994). With experience, as this new word is heard over and over again in multiple contexts, the meaning becomes more clearly defined (Medina et al., 2011).

Productive vocabulary is not the only type of vocabulary. Children have receptive or comprehension vocabularies, which differ from their expressive or productive vocabularies. As discussed earlier, children start to understand words months before they

produce any and at every age comprehension abilities exceed production abilities. One study of 1803 children found that at 16 months of age, the children understood between 92 and 321 words, but they could produce fewer than 50 words (Fenson et al., 1994).

Building semantic networks

While children are still learning new words, they are also building semantic networks, that is, mental representations in which the words they know are connected based on their shared features. For example, within a category of animals, a golden retriever and a Siamese cat are stored close together due to the meaningful, shared features of those animals: they are both four-legged, household pets, that are covered in fur, and make noises. In contrast, a golden retriever and a parrot, while both animals, share fewer features and therefore are stored farther from each other in the network. In laboratory settings, children at 18 months are quicker to recognize a word when it is preceded by a word from the same category (Arias-Trejo and Plunkett, 2010; Wojcik and Saffran, 2013). This task, called semantic priming, is one piece of evidence of children's early development of semantic networks. Additional evidence comes from children's use of decontextualized nouns, as they tend to overextend to referents that share features (i.e., calling a horse a dog; Moskowitz, 1978). Due to the limited procedures available to measure semantic networks early in development, little is known about these structures in infancy. But by 2 years of age, children are able to demonstrate their ability to use their existing semantic network to learn new words. Two-year-old children learn new words more accurately if the words come from a category in which the children already know more words (i.e., animals) compared to new words from categories in which children know fewer words (i.e., vehicles; Borovsky et al., 2016). At this age, children also learn new words better if those words belong to a category that they are interested in, with interest measured through pupil dilation (Ackermann et al., 2020).

Combining words

After children have produced their first word, their vocabulary grows and they begin to combine those newly learned words. Between 18 and 24 months, children start to put words together in multi-word utterances and eventually, complex sentences (see **First sentences** and **Complex, mult clause utterances** for more details on advanced language structures). First word combinations are termed telegraphic speech because they are typically combinations of nouns or nouns and verbs and do not include grammatical morphemes, a category of word forms including determiners, prepositions, auxiliary verbs, and bound morphemes at the end of nouns and verbs (Brown, 1973). Children may have these words and bound morphemes in their mental lexicons but fail to produce them because of production limitations (Atkinson, 1992; Gerken et al., 1990; Radford, 1990). Children begin to add these grammatical morphemes to their sentences after their first word combinations. Grammatical morphemes are acquired in a similar order across children, and they are reliable components of children's longer and more complex sentences (Brown, 1973).

One feature of children's early use of grammatical morphemes is overregularizations. These errors occur when children overextend a regular pattern (or rule, depending on the theoretical approach one adopts) to words that are exceptions to the pattern or rule (Berman, 1981; Marcus et al., 1992). For example, children may produce *goed*, *runned*, or *eated*. Importantly, the adults that children are receiving input from in their environments are not producing these speech errors. Instead, children are demonstrating that they have learned something about the regularities in language (Marcus et al., 1992).

Expanding communicative purposes

By 2 years of age, most children produce multi-word utterances to serve a variety of communicative purposes. As vocabulary and morphosyntactic knowledge continue to grow, the range of purposes to which children put their speech and the range of topics children talk about also develop. Children learn the pragmatics of conversations: how turn taking works, how to initiate a new topic of conversation, and how to repair a miscommunication (Bloom et al., 1976). Children start by talking mostly about the here-and-now, and conversations are sustained largely by the efforts of the adults involved (Snow, 1977). Early conversations are typically short, due to children's limited attention, but frequent, as children find new topics to discuss easily (Nelson, 1996).

From 24 to 36 months: mastering grammar and developing discourse skills

First sentences

Children's first sentences tend to be very simple sentences that involve one verb and take on three general forms: an imperative sentence, (a command), a declarative sentence (a statement or explanation), or a question. At first, children use mostly imperative sentence structures, followed by declarative sentences, and then questions, which are used more frequently as children develop (Vasilyeva et al., 2008).

Children also produce negative statements from an early age, but the form of those statements is not initially adult-like. In English, it is typical to use an auxiliary verb in order to express a negative statement (e.g., *I do not want to go*), and that auxiliary (*do*) is inserted between the subject and verb. When children first produce negatives they tend to put the negative marker outside the sentence (e.g., *No want go*) followed by putting the negative marker inside the sentence (e.g., *I no want go*; Klima and Bellugi, 1966). Children's first question forms also differ from the adult forms. Children will initially ask questions through intonation alone. With time, they will start to add *wh*-words to their questions (*What that is?*). By the time children are around 4 years of age, negative statements and questions contain the appropriate auxiliaries and children sound adult-like in their sentence structure.

The development of conversation and narrative skill

Between the ages of 2 and 3 years, children's conversation skills develop. Children start to talk about past events (Adamson and Bakeman, 2006) and they start to produce their first narratives (Eisenberg, 1985; Reese and Fivush, 1993). These are typically prompted by adults who may ask follow-up questions to encourage the continuation of the narrative, a technique called scaffolding (Clarke-Stewart and Beck, 1999). While narrative topics vary depending on the interests and experiences of the child (Miller et al., 1993; Nelson, 2004), these narratives are just a first step in their development. With time and experience, children come to produce more complex, cohesive narratives (Karmiloff-Smith, 1986) and their ability to produce narratives will aid them in their later academic language (Snow and Uccelli, 2009) and literacy development (Snow and Dickinson, 1990).

After the third birthday: further developing skills

Complex, mult clause utterances

After children produce their first multiword utterances, their utterances grow in length and complexity. Sentences may include multiple verbs, or clauses. Children begin using these more complex structures once they start producing utterances with four words (Bowerman, 1976). Between 2 and 3 years of age, the use of these sentences increases (Bloom, 1991), and children produce an adult-like range of complex sentence forms by about 4 years of age (Bowerman, 1979; Limber, 1973).

Continued growth and influences from schooling

By the age of 4 years, children have acquired the foundational skills necessary for linguistic proficiency. They will continue to develop their competencies in all domains of language. When children begin formal schooling, they will face a period of transition as they learn to adapt to a new environment. This new environment includes influences from social interactions with their teachers and peers, and also the start of literacy instruction. Many of the skills children learned before school will aid them in their later reading achievement.

Individual differences in early language development

The nature of individual differences

In broad outline, typically developing children follow similar paths in acquiring language. In production, infants first coo then babble, and then produce their first words, which tend to be produced in single-word utterances. Next, children put words together, and then their utterances grow longer, and the grammatical complexity of those utterances increases. Both the sequence and timing of these major milestones of language development—canonical babbling, first words, and first word combinations—appears fairly constant across children and environments (Casillas et al., 2020), and all typically developing children achieve these milestones. In contrast, the rate at which children add new words to their vocabularies and increase the grammatical complexity of their speech varies widely, and even among adults there is substantial variability in vocabulary knowledge and the grammatical complexity of speech (Hoff, 2006).

Internal sources of individual differences

Some of the variability among children in the rate of language growth is a function of individual differences in language learning abilities. Abilities that have been found to predict language growth include the following: the gaze of speakers and sustaining joint attention (Brooks and Meltzoff, 2008), short term phonological memory (Gathercole, 2006), a variety of attention and related skills that fall under the umbrella of executive function (Gandolfi and Viterbori, 2020; Kapa and Colombo, 2014; Ronfard et al., 2021), and statistical learning (Kidd, 2012; Kidd and Arciuli, 2016). These language learning skills may themselves develop with time and experience, but there is good evidence of a strong genetic component to language acquisition as well (Dale et al., 2000; Stromswold, 2001).

External sources of individual differences

The rate at which children grow in the size of their vocabularies and the range of grammatical structures that they use and understand depends substantially on the quantity and quality of their language exposure. The effects of quantity are easy to describe: children who are talked to more develop language more rapidly than children who are talked to less (Hart and Risley, 1995; Huttenlocher et al., 1991). But not all speech that children hear is equally useful to them in building language knowledge. Children who hear many different words build vocabularies faster than children who hear fewer words (Hoff, 2003; Hoff and Naigles, 2002). This effect holds within grammatical categories: children who hear many different verbs as the main verb in sentences subsequently use a larger verb vocabulary in their own speech (Hsu et al., 2017).

Hearing a varied vocabulary also helps children learn aspects of grammar. Children who hear many different nouns used as the subject of sentences learn that verbs take different endings depending on whether the subject is singular or plural (e.g., *the girl walks* but *the girls walk*) at an earlier age than children who hear fewer nouns as sentence subjects (Hadley et al., 2017). Children also benefit from syntactic complexity in input. The length of the utterances mothers address to their children predicts their children's

subsequent vocabulary growth (Hoff, 2003), and the diversity of syntactic structures in the speech addressed to children predicts the diversity of structures children come to use in their own speech (Huttenlocher et al., 2010). Apart from properties of the speech children hear, children's language development also benefits when their conversational partners are responsive to their speech and early vocalizations (Tamis-LeMonda et al., 2001), and expressive language skills also develop more rapidly when the children, themselves, contribute to conversations (Donnelly and Kidd, 2021).

The social context and language development

Children who grow up in different circumstances may have very different language learning experiences and their language development may also differ as a result (Hoff, 2006). One important field of research is aimed at describing the both the variability and the commonalities in children's language learning experiences worldwide, and the relation of the variability that exists to children's language development (e.g., Hoff, 2006; Rowe and Weisleder, 2020). Here we review the impact of two social context variables that shape language experience and language growth for many children: socioeconomic status and exposure to multilingual environments.

Socioeconomic status and language development

The quantity and quality of speech children hear varies as a function of their families' socioeconomic status (SES; Hoff and Laursen, 2019), with consequences for the children's language development (Hoff, 2006, 2013). This was dramatically and famously documented by Hart and Risley (1995), who recorded the speech addressed to children from different socioeconomic strata on a monthly basis from the time the children were 9 months old through the age of 3 years. Based on the number of words children heard per hour in those recordings, Hart and Risley (1995) calculated that by the age of 4 years, children from low SES homes will have heard 30 million fewer words than children from high SES homes. This gap in linguistic experience, often called the word gap, is present not only in the quantity of speech that children hear, but also in the quality of speech they hear. Many other studies have found SES-related differences in the richness of the vocabulary and diversity of grammatical structures children hear in the speech addressed to them (Golinkoff et al., 2019; Hoff, 2013; Huttenlocher et al., 2010). The purposes of adult talk have also been found to differ as a function of SES. Compared to higher SES mothers, lower SES mothers more frequently address their children in order to direct their behavior and less frequently talk to their children for the purpose of engaging them in conversation (Hoff, 2006; Hoff and Laursen, 2019).

Children from low SES households differ from children from higher SES households in having smaller vocabularies, shorter utterances, and a more restricted range of grammatical structures (Arriaga et al., 1998; Hart and Risley, 1995; Huttenlocher et al., 2010; Snow, 1999), although the findings with respect to grammar are less robust than the gap in vocabulary skill (Hoff and Tian, 2005; Noble et al., 2005). The differences in children's language skill associated with SES can be linked to the differences in child-directed speech that are associated with SES (Hoff, 2003; Huttenlocher et al., 2010). By school entry, children from low-income households are less able to produce effective communicative messages (Lloyd et al., 1998). In addition, children from low-income households have less sophisticated narratives (Vernon-Feagan et al., 2001), limited phonological awareness skill (Lonigan et al., 1998; McDowell et al., 2007), and are slower in accessing known words (Fernald and Marchman, 2012).

Multilingual exposure and language development

Many children are exposed to more than one language in early childhood, and multilingual exposure creates additional variability in the quantity and quality of children's language exposure. The most obvious effect on experience is that the quantity of children's exposure to each language is, on average, less than it would be if they were hearing only one language. In the early stages of development, bilingually developing children acquire each of their languages at a slower pace than monolingual children acquire their single language, and the relative amount of time children hear each language is a significant predictor of their rates of growth in each language (Hoff et al., 2012; Hoff and Ribot, 2017). When knowledge in their two languages is counted together, it is clear that bilingual children are acquiring language at a comparable pace to monolingual children (Hoff et al., 2012). The lag in the rate of development compared to monolinguals reflects the fact that their experience and therefore their growth is distributed across two languages. If children continue to receive high levels of high-quality exposure to two languages, the evidence suggests bilingual children catch up to monolingual children for all practical purposes. However, many of the real world circumstances in which children are exposed to two languages advantage one language over the other.

For example, when one language is used only at home, the range of vocabulary and grammar the children are exposed to may be less than for the language that the children also use in school. As a result, children in immigrant families often end up with stronger skills in the language of their new country and school than in the language they first heard from their parents (Giguere and Hoff, 2020). Children in bilingual or multilingual environments also often hear a great deal of speech from nonnative speakers (Place and Hoff, 2011, 2016). The proficiency of nonnative speakers influences the quality of their child-directed speech (Hoff et al., 2020). The difference in quality of speech by non-native parents with lower proficiencies impacts children's vocabulary development (Buac et al., 2014). This impact can be lessened when one parent is a native speaker of the target language (Hoff et al., 2014, 2018; Place and Hoff, 2011).

Maternal education, which is one of the indicators of SES that is reliably associated with both the quantity and quality of child-directed speech and with children's language development in monolingual samples (Hoff, 2003; Rowe, 2008, 2012) is a more complicated variable for bilingual mothers—and fathers. Many bilingual parents are immigrants who attended school both before immigration—in their native language—and after immigration—in the language of their new country. Some research suggests that maternal education has language-specific effects on language use in the home and on children's language growth. In a sample of bilingual mothers in Canada who varied in their first language, those who were more educated in English (their second language) spoke English more often to their children (Duncan and Paradis, 2020). Among Spanish-English bilinguals, maternal education in Spanish predicted children's expressive vocabulary growth in Spanish, but not English, whereas the education level they achieved in English predicted the children's English vocabulary and not their Spanish vocabulary (Hoff et al., 2018).

Because language learning experiences can be so varied in multilingual environments, bilingual and multilingual speakers vary widely in their language skills—in childhood (Hoff et al., 2021) and in adulthood as well (Luk and Bialystok, 2013). Bilingual children who attend school in one of their languages tend to achieve monolingual-like levels of proficiency in that language (Giguere and Hoff, 2020), but even early exposure does not guarantee development of a language that is only a home language (Ardila et al., 2019; DeHouwer, 2007; Giguere and Hoff, 2020).

Summary and conclusion

Language is a universal characteristic of the human species and language acquisition a universal aspect of early development. Language development follows a fairly regular course across children. In this article we have provided a broad overview of that course of development in first four years of life. We have described the changes in what children are able to produce and the communicate purposes to which they put their growing linguistic abilities. We have described some of the individual differences observed in language development, and we have said something about the sources of those individual differences.

The list of topics that we have not covered in this article is long. We have said little about the processes underlying the child's acquisition of language, and, accordingly, little about the theoretical controversies in the field. We have ended our description at 4 years and not discussed later developments in oral language nor the acquisition of literacy. Our description is based largely on studies of typically developing, monolingual children acquiring English in mainstream North American families. We have said little about children with atypical abilities and little about the acquisition of other languages or language acquisition in other cultures. We invite the interested reader to pursue these and other topics in the in the suggested further readings.

Acknowledgment

Erika Hoff was supported by the US National Institutes of Health grant #HD068421.

References

- Ackermann, L., Hepach, R., Mani, N., 2020. Children learn words easier when they are interested in the category to which the word belongs. *Dev. Sci.* 23, 1–12.
- Adamson, L.B., Bakeman, R., 2006. Development of displaced speech in early mother–child conversations. *Child Dev.* 77, 186–200.
- Ardila, A., Rosselli, M., Ortega, A., Lang, M., Torres, V.L., 2019. Oral and written language abilities in young Spanish/English bilinguals. *Int. J. BiLing.* 23, 296–312.
- Arias-Trejo, N., Plunkett, K., 2010. The effects of perceptual similarity and category membership on early word-referent identification. *J. Exp. Child Psychol.* 105, 63–80.
- Arriaga, R.I., Fenton, L., Cronan, T., Pethick, S.J., 1998. Scores on the MacArthur Communicative Development Inventory of children from low-and middle-income families. *Appl. Psycholinguist.* 19, 209–224.
- Atkinson, M., 1992. *Children's Syntax*. Cambridge University Press, Cambridge, MA.
- Austin, J.L., 1962. *How to do things with words*. Oxford University Press, Oxford, UK.
- Barrett, M.D., 1986. Early semantic representations and early word usage. In: Kuczaj, S.A., Barrett, M.D. (Eds.), *The Development of Word Meaning*. Springer-Verlag, New York, pp. 39–67.
- Bates, E., Camaioni, L., Volterra, V., 1975. The acquisition of performatives prior to speech. *Merrill-Palmer Q. Behav. Dev.* 21, 205–226.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., Volterra, V., 1979. *The Emergence of Symbols: Cognition and Communication in Infancy*. Academic Press, New York.
- Benedict, H., 1979. Early lexical development: comprehension and production. *J. Child Lang.* 6, 183–200.
- Bergelson, E., Swingle, D., 2012. At 6–9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3253–3258.
- Berman, R.A., 1981. Children's regularizations of plural forms. *Pap. Rep. Child Lang. Dev.* 21, 34–43.
- Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B., Quann, C.A., 2009. Development of phonological constancy: Toddlers' perception of native-and Jamaican-accented words. *Psychol. Sci.* 20, 539–542.
- Bloom, L., Rocissano, L., Hood, L., 1976. Adult child discourse: developmental interaction between information processing and linguistic knowledge. *Cognit. Psychol.* 8, 521–552.
- Bloom, L., 1973. *One Word at a Time*. Mouton, The Hague, The Netherlands.
- Bloom, L., 1991. *Language Development From Two to Three*. Cambridge University Press, Cambridge, UK.
- Borovsky, A., Ellis, E.M., Evans, J.L., Elman, J.L., 2016. Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Dev. Sci.* 19, 918–932.
- Bowerman, M., 1976. Semantic factors in the acquisition of rules for word use and sentence construction. In: Morehead, D.M., Morehead, A.E. (Eds.), *Normal and Deficient Child Language*. University Park Press, Baltimore, pp. 99–179.
- Bowerman, M., 1979. The acquisition of complex sentences. In: Fletcher, P., Garman, M. (Eds.), *Language Acquisition*. Cambridge University Press, Cambridge, UK, pp. 285–306.
- Brooks, R., Meltzoff, A.N., 2005. The development of gaze following and its relation to language. *Dev. Sci.* 8, 535–543.

- Brooks, R., Meltzoff, A.N., 2008. Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *J. Child Lang.* 35, 207–220.
- Brown, H.D., 1973. Affective variables in second language acquisition. *Lang. Learn.* 23, 231–244.
- Buac, M., Gross, M., Kaushanskaya, M., 2014. The role of primary caregiver vocabulary knowledge in the development of bilingual children's vocabulary skills. *J. Speech Lang. Hear. Res.* 57, 1804–1816.
- Camaioni, L., Perucchini, P., 2003. Profiles in Declarative/Imperative Pointing and Early Word Production. Paper Presented at the European Conference for Developmental Psychology. Milan, Italy.
- Carey, S., 1978. The child as a word learner. In: Halle, M., Bresnan, J., Miller, G. (Eds.), *Linguistic Theory and Psychological Reality*. MIT Press, Cambridge, MA, pp. 264–293.
- Carpenter, M., Nagell, K., Tomasello, M., 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–174.
- Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanders, L., Weir, J., 1995. A cross-linguistic study of early lexical development. *Cognit. Dev.* 10, 159–199.
- Casillas, M., Brown, P., Levinson, S.C., 2020. Early language experience in a Tzeltal Mayan village. *Child Dev.* 91, 1819–1835.
- Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D., 2010. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Q. J. Econ.* 126, 1593–1660.
- Clark, H.H., 1979. Responding to indirect speech acts. *Cognit. Psychol.* 11, 430–477.
- Clarke-Stewart, K.A., Beck, R.J., 1999. Maternal scaffolding and children's narrative retelling of a movie story. *Early Child. Res. Q.* 14, 409–434.
- Dale, P.S., Dionne, G., Eley, T.C., Plomin, R., 2000. Lexical and grammatical development: a behavioral genetic perspective. *J. Child Lang.* 27, 619–642.
- De Houwer, A., 2007. Parental language input patterns and children's bilingual use. *Appl. Psycholinguist.* 28, 411–424.
- de Sousa, L.B., Gabriel, R., 2015. Does the mental lexicon exist? *Rev. Estud. Ling.* 23, 335–361.
- Donnelly, S., Kidd, E., 2021. The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Dev.* 92, 609–625.
- Dore, J., Franklin, M.B., Miller, R.T., Ramer, A.L., 1976. Transitional phenomena in early language acquisition. *J. Child Lang.* 3, 13–28.
- Dore, J., 1975. Holophrases, speech acts and language universals. *J. Child Lang.* 2, 21–40.
- Duncan, T.S., Paradis, J., 2020. How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English? *Int. J. Biling.* 24, 46–61.
- Eisenberg, A., 1985. Learning to describe past experiences in conversation. *Discourse Process.* 8, 177–204.
- Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., 1994. Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59, 85–123.
- Fernald, A., Marchman, V.A., 2012. Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Dev.* 83, 203–222.
- Fernald, A., Simon, T., 1984. Expanded intonation contours in mothers' speech to newborns. *Dev. Psychol.* 20, 104–113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., Fukui, I., 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.* 16, 477–501.
- Gandolfi, E., Viterbori, P., 2020. Inhibitory control skills and language acquisition in toddlers and preschool children. *Lang. Learn.* 70, 604–642.
- Gathercole, S.E., 2006. Nonword repetition and word learning: the nature of the relationship. *Appl. Psycholinguist.* 27, 513–543.
- Gerken, L., Landau, B., Remez, R.E., 1990. Function morphemes in young children's speech perception and production. *Dev. Psychol.* 26, 204–216.
- Gerken, L., 1994. Child phonology: past research, present questions, future directions. In: *Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics*. Academic Press, New York, pp. 781–820.
- Giguere, D., Hoff, E., 2020. Home language and societal language skills in second-generation bilingual adults. *Int. J. Biling.* 24, 1071–1087.
- Golinkoff, R.M., Hoff, E., Rowe, M.L., Tamis-LeMonda, C.S., Hirsh-Pasek, K., 2019. Language matters: denying the existence of the 30-million-word gap has serious consequences. *Child Dev.* 90, 985–992.
- Hadley, P.A., Rispoli, M., Holt, J.K., 2017. Input subject diversity accelerates the growth of tense and agreement: indirect benefits from a parent-implemented intervention. *J. Speech Lang. Hear. Res.* 60, 2619–2635.
- Hart, B., Risley, T.R., 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Brookes, Baltimore.
- Hoek, D., Ingram, D., Gibson, D., 1986. Some possible causes of children's early word overextensions. *J. Child Lang.* 13, 477–494.
- Hoff, E., Laursen, B., 2019. Socioeconomic status and parenting. In: *Handbook of Parenting*. Routledge, Milton Park, Oxfordshire, pp. 421–447.
- Hoff, E., Naigles, L., 2002. How children use input to acquire a lexicon. *Child Dev.* 73, 418–433.
- Hoff, E., Ribot, K.M., 2017. Language growth in English monolingual and Spanish-English bilingual children from 2.5 to 5 years. *J. Pediatr.* 190, 241–245.
- Hoff, E., Tian, C., 2005. Socioeconomic status and cultural influences on language. *J. Commun. Disord.* 38, 271–278.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., Parra, M., 2012. Dual language exposure and early bilingual development. *J. Child Lang.* 39, 1–27.
- Hoff, E., Rumiche, R., Burridge, A., Ribot, K.M., Welsh, S., 2014. Expressive vocabulary development in children from bilingual homes: a longitudinal study from two to four years. *Early Child. Res. Q.* 29, 433–444.
- Hoff, E., Burridge, A., Ribot, K.M., Giguere, D., 2018. Language specificity in the relation of maternal education to bilingual children's vocabulary growth. *Dev. Psychol.* 54, 1011–1019.
- Hoff, E., Core, C., Shanks, K.F., 2020. The quality of child-directed speech depends on the speaker's language proficiency. *J. Child Lang.* 47, 132–145.
- Hoff, E., Tuloach, M.K., Core, C., 2021. Profiles of minority-majority language proficiency in 5-year-olds. *Child Dev.* 1–16.
- Hoff, E., 2003. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Dev.* 74, 1368–1378.
- Hoff, E., 2006. How social contexts support and shape language development. *Dev. Rev.* 26, 55–88.
- Hoff, E., 2013. Interpreting the early language trajectories of children from low SES and language minority homes: implications for closing achievement gaps. *Dev. Psychol.* 49, 4–14.
- Houston, D.M., Jusczyk, P.W., 2000. The role of talker-specific information in word segmentation by infants. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1570–1582.
- Hsu, N., Hadley, P.A., Rispoli, M., 2017. Diversity matters: parent input predicts toddler verb production. *J. Child Lang.* 44, 63–86.
- Hudson, J., Nelson, K., 1984. Play with language: overextensions as analogies. *J. Child Lang.* 11, 337–346.
- Hulit, L.M., Howard, M.R., 2006. *Language acquisition: a theoretical journey*. In: *Born to Talk: An Introduction to Speech and Language Development*, fourth ed. Allyn & Bacon, Boston, pp. 16–54.
- Huttenlocher, J., Smiley, P., 1987. Early word meanings: the case of object names. *Cognit. Psychol.* 19, 63–89.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., Lyons, T., 1991. Early vocabulary growth: relation to language input and gender. *Dev. Psychol.* 27, 236–248.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., Hedges, L.V., 2010. Sources of variability in children's language growth. *Cognit. Psychol.* 61, 343–365.
- Iverson, J.M., Goldin-Meadow, S., 2005. Gesture paves the way for language development. *Psychol. Sci.* 16, 368–371.
- Kapa, L.L., Colombo, J., 2014. Executive function predicts artificial language learning. *J. Mem. Lang.* 76, 237–252.
- Karmiloff-Smith, A., 1986. From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition* 23, 95–147.
- Kidd, E., Arciuli, J., 2016. Individual differences in statistical learning predict children's comprehension of syntax. *Child Dev.* 87, 184–193.
- Kidd, E., 2012. Implicit statistical learning is directly associated with the acquisition of syntax. *Dev. Psychol.* 48, 171–184.
- Klima, E., Bellugi, U., 1966. Syntactic regularities in children's speech. In: *Psycholinguistic Papers*, pp. 183–208.
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.

- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., Iverson, P., 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21.
- Lew-Williams, C., Pelucchi, B., Saffran, J.R., 2011. Isolated words enhance statistical language learning in infancy. *Dev. Sci.* 14, 1323–1329.
- Limber, J., 1973. The genesis of complex sentences. In: Moore, T. (Ed.), *Cognitive Development and the Acquisition of Language*. Academic Press, New York, pp. 169–186.
- Lloyd, P., Mann, S., Peers, I., 1998. The growth of speaker and listener skills from five to eleven years. *First Lang.* 18, 81–104.
- Lonigan, C.J., Burgess, S.R., Anthony, J.L., Barker, T.A., 1998. Development of phonological sensitivity in 2- to 5-year-old children. *J. Educ. Psychol.* 90, 294.
- Luk, G., Bialystok, E., 2013. Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cognit. Psychol.* 25, 605–621.
- Ma, W., Golinkoff, R.M., Houston, D.M., Hirsh-Pasek, K., 2011. Word learning in infant- and adult-directed speech. *Lang. Learn. Dev.* 7, 185–201.
- Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J., Xu, F., 1992. Overregularization in language acquisition. *Monogr. Soc. Res. Child Dev.* 57, i–178.
- Markman, E.M., Wasow, J.L., Hansen, M.B., 2003. Use of the mutual exclusivity assumption by young word learners. *Cognit. Psychol.* 47, 241–275.
- McDowell, K.D., Lonigan, C.J., Goldstein, H., 2007. Relations among socioeconomic status, age, and predictors of phonological awareness. *J. Speech Lang. Hear. Res.* 50, 1079–1092.
- Medina, T.N., Snedeker, J., Trueswell, J.C., Gleitman, L.R., 2011. How words can and cannot be learned by observation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9014–9019.
- Mervis, C.B., Bertrand, J., 1994. Acquisition of the novel name—nameless category principle. *Child Dev.* 65, 1646–1662.
- Miller, P.J., Hoogstra, L., Mintz, J., Fung, H., Williams, K., 1993. Troubles in the garden and how they get resolved: a young child's transformation of his favorite story. In: Nelson, C.A. (Ed.), *Memory and Affect in Development*. Lawrence Erlbaum, Hillsdale, NJ, pp. 87–114.
- Moskowitz, B.A., 1978. The acquisition of language. *Sci. Am.* 239, 92–109.
- Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A.V., Parlade, M.V., 2007. Individual differences and the development of joint attention in infancy. *Child Dev.* 78, 938–954.
- Nelson, K., 1973. Structure and strategy in learning to talk. *Monogr. Soc. Res. Child Dev.* 38, 1–135.
- Nelson, K., 1996. *Language in Cognitive Development: Emergence of the Mediated Mind*. Cambridge University Press, Cambridge, MA.
- Nelson, K., 2004. Construction of the cultural self in early narratives. In: Daiute, C., Lightfoot, C. (Eds.), *Narrative Analysis: Studying the Development of Individuals in Society*. Sage Productions, London, UK, pp. 87–107.
- Noble, K.G., Norman, M.F., Farah, M.J., 2005. Neurocognitive correlates of socioeconomic status in kindergarten children. *Dev. Sci.* 8, 74–87.
- Oller, D.K., 2000. *The Emergence of the Speech Capacity*. Lawrence Erlbaum, Mahwah, NJ.
- Owens, R.E., 2006. Development of communication, language, and speech. In: *Human Communication Disorders: An Introduction*, seventh ed. Pearson, Allyn & Bacon, Boston, pp. 226–270.
- Place, S., Hoff, E., 2011. Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Dev.* 82, 1834–1849.
- Place, S., Hoff, E., 2016. Effects and noneffects of input in bilingual environments on dual language skills in 2½-year-olds. *Bilingualism* 19, 1023–1041.
- Pye, C., Ingram, D., List, H., 1987. A comparison of initial consonant acquisition in English and Quiche. In: Nelson, K.E., van Kleeck, A. (Eds.), *Children's Language*. Erlbaum, Mahwah, NJ, pp. 175–190.
- Radford, A., 1990. *Synthetic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Blackwell, Oxford, U.K.
- Reese, E., Fivush, R., 1993. Parental styles of talking about the past. *Dev. Psychol.* 29, 596–606.
- Rescorla, L.A., 1980. Overextension in early language development. *J. Child Lang.* 7, 321–335.
- Ronfard, S., Wei, R., Rowe, M.L., 2021. Exploring the linguistic, cognitive, and social skills underlying lexical processing efficiency as measured by the looking-while-listening paradigm. *J. Child Lang.* 1–24.
- Rowe, M.L., Weisleder, A., 2020. Language development in context. *Annu. Rev. Dev. Psychol.* 2, 201–223.
- Rowe, M.L., Özçaliskan, S., Goldin-Meadow, S., 2008. Learning words by hand: gesture's role in predicting vocabulary development. *First Lang.* 28, 182–199.
- Rowe, M.L., 2008. Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *J. Child Lang.* 35, 185–205.
- Rowe, M.L., 2012. A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Dev.* 83, 1762–1774.
- Saffran, J.R., Kirkham, N.Z., 2018. Infant statistical learning. *Annu. Rev. Psychol.* 69, 181–203.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Schmale, R., Seidl, A., 2009. Accommodating variability in voice and foreign accent: flexibility of early word representations. *Dev. Sci.* 12, 583–601.
- Schmale, R., Cristia, A., Seidl, A., Johnson, E.K., 2010. Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy* 15, 650–662.
- Searle, J., 1969. *Speech acts*. Cambridge University Press, Cambridge, UK.
- Singh, L., Morgan, J.L., White, K.S., 2004. Preference and processing: the role of speech affect in early spoken word recognition. *J. Mem. Lang.* 51, 173–189.
- Snow, C.E., Dickinson, D.K., 1990. Social sources of narrative skills at home and at school. *First Lang.* 10, 87–103.
- Snow, C.E., Uccelli, P., 2009. The challenge of academic language. In: Olson, D.R., Torrance, N. (Eds.), *The Cambridge Handbook of Literacy*. Cambridge University Press, Cambridge, MA, pp. 112–133.
- Snow, C.E., 1977. The development of conversation between mothers and babies. *J. Child Lang.* 4, 1–22.
- Snow, C.E., 1999. Social perspectives on the emergence of language. In: MacWhinney, B. (Ed.), *The Emergence of Language*. Erlbaum, Mahwah, NJ, pp. 257–276.
- Stark, R.E., 1986. Prespeech segmental feature development. In: Fletcher, P., Garman, M. (Eds.), *Language Acquisition*, second ed. Cambridge University Press, Cambridge, NY, pp. 149–173.
- Stokes, S.F., Surendran, D., 2005. Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *J. Speech Lang. Hear. Res.* 48, 577–591.
- Stromswold, K., 2001. The heritability of language: a review and metaanalysis of twin, adoption, and linkage studies. *Language* 77, 647–723.
- Tamis-LeMonda, C.S., Bornstein, M.H., Baumwell, L., 2001. Maternal responsiveness and children's achievement of language milestones. *Child Dev.* 72, 748–767.
- Thiessen, E.D., Saffran, J.R., 2003. When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev. Psychol.* 39, 706–716.
- Thiessen, E.D., Saffran, J.R., 2007. Learning to learn: infants' acquisition of stress-based strategies for word segmentation. *Lang. Learn. Dev.* 3, 73–100.
- Thiessen, E.D., Hill, E.A., Saffran, J.R., 2005. Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71.
- Tincoff, R., Jusczyk, P.W., 1999. Some beginnings of word comprehension in 6-month-olds. *Psychol. Sci.* 10, 172–175.
- Tomasello, M., Carpenter, M., Liszkowski, U., 2007. A new look at infant pointing. *Child Dev.* 78, 705–722.
- Tomasello, M., 2007. Cooperation and communication in the 2nd year of life. *Child Dev. Perspect.* 1, 8–12.
- van Heugten, M., Johnson, E.K., 2014. Learning to contend with accents in infancy: benefits of brief speaker exposure. *J. Exp. Psychol. Gen.* 143, 340–350.
- van Heugten, M., Krieger, D.R., Johnson, E.K., 2015. The developmental trajectory of toddlers' comprehension of unfamiliar regional accents. *Lang. Learn. Dev.* 11, 41–65.
- van Severen, L., Gillis, J.J.M., Molemans, I., van den Berg, R., de Maeyer, S., Gillis, S., 2013. The relation between order of acquisition, segmental frequency, and function: the case of word-initial consonants in Dutch. *J. Child Lang.* 40, 703–740.
- Vasilyeva, M., Waterfall, H., Huttenlocher, J., 2008. Emergence of syntax: commonalities and differences across children. *Dev. Sci.* 11, 84–97.
- Vernon-Feagans, L., Hammer, C.S., Miccio, A., Manlove, E., 2001. Early language and literacy skills in low-income African American and Hispanic children. *Handb. Early Literacy Res.* 1, 192–210.

- von Stumm, S., Rimfeld, K., Dale, P.S., Plomin, R., 2020. Preschool verbal and nonverbal ability mediate the association between socioeconomic status and school performance. *Child Dev.* 91, 705–714.
- Werker, J.F., Tees, R.C., 1984. Phonemic and phonetic factors in adult cross-language speech perception. *J. Acoust. Soc. Am.* 75, 1866–1878.
- Wojcik, E.H., Saffran, J.R., 2013. The ontogeny of lexical networks: toddlers encode the relationships among referents when learning novel words. *Psychol. Sci.* 24, 1898–1905.

Further reading

- Bavin, E., Naigles, L. (Eds.), *The Cambridge Handbook of Child Language*, second ed., pp. 483–503. Cambridge University Press.
- Eigsti, I.M., de Marchena, A.B., Schuh, J.M., Kelley, E., 2011. Language acquisition in autism spectrum disorders: a developmental review. *Res. Autism Spectr. Disord.* 5, 681–691.
- Hoff, E., 2021. Why bilingual development is not easy. In: Lockman, J. (Ed.), *Advances in Child Development and Behavior.*, pp. 129–167. Elsevier.
- Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D.J., Padden, C., Rathmann, C., 2014. Ensuring language acquisition for deaf children: what linguists can do. *Language* 90, e31–e52.

Key developments in early childhood

Elizabeth Ann Wood, School of Education, The University of Sheffield, Sheffield, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	477
Section 1: How have socio-cultural theories developed over time?	478
Section 2: What key concepts are informing early childhood education?	479
Learning leads development	479
The zone of proximal development	479
Cultural tools and symbol systems (tools for use/tools and use)	480
Children's agency and motivation	481
Play as social practice	481
Section 3: What are the implications of socio-cultural theories for early childhood education?	483
References	484

Introduction

Early childhood education (ECE) is informed by a range of ideas, philosophies and theories, which have developed over time and in different national contexts (Gunn and Nuttall, 2019). These diverse influences have been mystical, spiritual, ideological and pragmatic, based on contrasting ways of understanding children and childhood. From the middle of the nineteenth century the early childhood "tradition" or "inheritance" created a blueprint for the dominance of Euro-American theories in driving changes in society's perceptions of children's need for free play, education and care, and understanding childhood as a distinctive period of human development. The different approaches that emerged from this tradition combined beliefs about how children should be treated in society, and how they learn and develop. Governmental and societal concerns for children's health, welfare and education required a scientific approach to justifying public investment. Hence the philosophical and ideological orientations of the "early pioneers" of the nineteenth century gave way to the new discipline of child study, with educational and child psychology bringing methods and theories from the learning sciences. How these theories were interpreted during the twentieth century emphasized the child's natural/biological development, and created a maturational narrative of learning and development progressing through ages and stages, with children needing to become ready to progress from one stage to another. In the field of practice, the work of influential psychologists such as Susan Isaacs and Jean Piaget, and the educational philosopher, John Dewey, was aligned with progressive principles of child-centered education, which emphasized hands-on experience, and active learning through play, exploration, inquiry and discovery. The expansion of provision for children before compulsory education, and guidance for practice, were thus informed by a complex inter-relationship between the learning sciences and educational reform. Developmental and educational psychology subsequently provided the dominant theoretical orientation to educating ECE teachers and practitioners, and informing practice, alongside enduring but often uncritical commitments to the sedimented knowledge from the early childhood "tradition."

Shaping children's development through home, community and education contexts reflected socio-economic shifts toward shared responsibilities between private (family) life, and public (state) investment in ECE. As Wood (2020) argued, child development theories appealed to this dual economic-education rationale because they provided a scientific basis for child-rearing, child-care and health, and presented development as hierarchical, regular, and linear. The influence of the learning sciences on ECE led to the adoption of the Piagetian theoretical orientation that development leads or precedes learning through biological maturation, and social interaction. Drawing on research conducted primarily in the USA and western European countries, early childhood educators responded to the powerful discourse of "developmentally appropriate practice" (DAP), which, from the 1980s, secured international reach and impact (NAEYC.org/DAP) (National Association for the Education of Young Children, 2020). Educational programs informed by DAP were based on the child's observed trajectory, and mapped against developmental goals, organized into the developmental domains or areas of learning.

When integrated into educational approaches, and ECE policy frameworks, child development theories have created a regime of truth (Bartholomaeus, 2016) based on typical expectations, goals or milestones. The tendency toward "normalization" results in children being positioned as normal/abnormal, achieving/underachieving, being typical or ahead, or lagging behind (Wood, 2020). These forms of labeling influence how children (as groups or individuals) are positioned as being "at risk," "developmentally delayed" and "in need" of interventions. A core problem of conceiving development in an "ages and stages" model is that developmental processes are not consistently hierarchical, causal or deterministic: transitions vary in their timing, and different approaches to learning are situated in different socio-cultural-historical contexts (Chen et al., 2011; Gutiérrez and Rogoff, 2003). Contemporary interpretations of socio-cultural theories provide contrasting perspectives on learning and development, with potential for responding to changing societal and global challenges.

The work of the Russian psychologist, Lev Vygotsky was revolutionary in its own time, and laid the foundations for new ways of thinking about human development that contest "ages and stages," and illuminate complexity. It is important to emphasize that

Vygotsky laid the foundations, because his theories and research were still in development when he died in 1934 at the age of 37. Following translations of his work into English, Vygotsky's theories were subsequently developed posthumously, and have been contested and reinterpreted by different people, through a range of disciplines, and in different socio-cultural contexts, from early childhood, across the lifespan. Contemporary socio-cultural theories may be post-Vygotskian, or neo-Vygotskian, but generally trace their origins, and subsequent interpretations to his original works, and identify what omissions have been addressed (Bruner, 1996; Daniels, 2008, 2016; Newman and Holzman, 2013; Rogoff, 2003). New insights have been developed through empirical research, drawing on inter-disciplinary perspectives, for example in play (Holzman, 2017; Sutton-Smith, 1997; Wood, 2013), and learning (Chesworth and Hedges, 2023; Hedges, 2021; Lenz-Taguchi, 2010). A key theoretical shift, as argued by Hatch (2010, 2012) and Holzman (2017), is that learning leads development, with learning and development being in dialectical relationship. This shift contests the epistemological hierarchy of child development theories, informed by global perspectives and contemporary critique and interpretations of socio-cultural theories (Hedges, 2022; Stagg Peterson and Friedrich, 2022; Vintimilla and Pacini-Ketchabaw, 2020).

The aim of this chapter is not to critique or analyze different interpretations of socio-cultural theories, but to examine the implications of selected key concepts for the policies, purposes and practices of early childhood education. This chapter addresses three key questions.

1. How have socio-cultural theories developed over time?
2. What key concepts are informing early childhood education?
3. What are the implications for early childhood education: from developmentally to culturally appropriate practice

Section 1: How have socio-cultural theories developed over time? provides an overview of the development of socio-cultural theories, drawing on the work of Lev Vygotsky. **Section 2: What key concepts are informing early childhood education?** examines three key concepts that derive from Vygotsky's original work, and the different ways in which these have been developed and interpreted in ECE. The three concepts are learning leading development; cultural tools and symbol systems (tools for use/tools and use) and children's agency and motivation. **Section 2: What key concepts are informing early childhood education?** elaborates the selected concepts in relation to contemporary understanding of play as a leading activity in ECE, and play as social practice. **Play as social practice** considers the implications for early childhood education of contemporary interpretations of socio-cultural theories, specifically the conceptual reorientation from developmentally to culturally appropriate practice, and to learning leading development.

Section 1: How have socio-cultural theories developed over time?

Socio-cultural theories provided a conceptual shift toward understanding human development and learning as dialectical, socially- and culturally-situated, drawing on contemporary practices, and those of previous generations, hence the cultural-historical focus (Rogoff, 2003; Rogoff et al., 2018). In ECE, socio-cultural theories have moved attention from what is developmentally appropriate to what is culturally appropriate, so that diversity and difference are understood within children's everyday social and cultural practices in their homes, families and communities. Contemporary interpretations are linked to wider debates about global knowledge and policy flows, and the global politics of knowing, learning and teaching. The focus on the cultural nature of human development provokes critical engagement with inclusion, equity and equality, and the policies, purposes and practices of ECE.

For Vygotsky (1978), learning is a social process, situated in socio-cultural/cultural-historical contexts. This foundational concept informs the contemporary emphasis on co-construction as peers and adults engage and participate in shared practices. Children's learning happens within developmental environments, or interrelated systems (Darling-Hammond et al., 2020) that are organized to promote learning and teaching. In ECE, the developmental domains are interrelated—physical/emodied, social/relational, emotion/affect, psychological/cognitive. The subject areas or disciplines are understood as cultural tools and symbol systems, including their distinctive skills and concepts, and ways of thinking, inquiring, creating and problem solving. Engagement with the subject disciplines, whether in everyday/emergent, or scientific forms, is transformative in that children participate with increased competence, agency and sense of identity. The social situation for development promotes engagement and participation with humans and non-humans, cultural tools, materials, symbols, technologies, the natural world, and valued knowledge. The term “valued knowledge” is understood from a socio-cultural perspective as contextually situated, drawing on children's cultural repertoires, languages, interests and shared knowledge in their homes, families, communities and education settings, including popular culture and digital media. For example, in children's social play, knowledge that is valued and useful includes knowing how to join a play episode, how to contribute imaginative ideas that develop the play theme and events, and how to keep the play going by taking turns, sharing and listening. Different forms of “valued knowledge” are also inscribed in ECE policy and curriculum frameworks.

Knowledge is not acquired “in the brain,” but is co-constructed in social contexts through distributed social practices that involve using concepts and cultural tools with purpose and meaning. Learning over time is not simply a matter of “acquisition”—learning, remembering or knowing more. In their everyday activities, children are capable of co-creating their own interests and purposes, drawing on available cultural tools and resources, co-constructing meaning, and producing outcomes that are relevant in the moment, and build toward further learning. These processes lead to the transformation of participation (Rogoff, 2003; Rogoff et al., 2018): as learners become more skillful, competent, knowledgeable and confident, they take responsibility for assuming

active roles in the management and development of activities. They progress toward deeper engagement in communities of practice as they develop their identities, and their agency.

Section 2: What key concepts are informing early childhood education?

Learning leads development

How we understand learning and development has implications for children's education, care, health and well-being, and their rights. As noted earlier, a key socio-cultural perspective is that learning leads development, with learning and development being in dialectical relationship (Hatch, 2010, 2012; Hedges, 2022; Holzman, 2017). Children are not "adults-in-waiting"; they have their own complex ways of being, becoming and belonging, making meaning, and coming to know in their homes and communities, and the distinctive social contexts of ECE settings. Taking a learning leads development perspective respects diversity and difference, and responds to children's agency and participation in social contexts. Key themes (such as play) and concepts (such as the zone of proximal development) have been particularly influential in ECE, again with contrasting interpretations and emphases.

The zone of proximal development

Vygotsky's original writing about the zone of proximal development (ZPD) has been re-interpreted, including different translations of his theories and meanings, with different implications for research and practice (Daniels, 2008, 2016; Hill and Wood, 2019). Although Vygotsky (1978) proposed critical periods and crises in development, these were not structured into defined ages and stages. For Vygotsky (1978) learning and development are interrelated from birth, but learning leads development within the zone of proximal development.

It is worth remembering that the ZPD was a conceptual tool through which educators and psychologists could understand the internal course of development as a social-cultural-historical process. The ZPD is also a metaphor, to conceptualize the distance between actual and potential development, and how development occurs collectively through guidance or collaboration with more or differently capable humans—adults or peers, with the support of cultural tools. Vygotsky proposed a dialectical relationship between a child's actual and potential level of development, as evident in what a child can do unaided, and with support. He used the metaphor of the "buds" or flowers' and the "fruits" of development (1978, p. 86) to explain dynamic processes of change:

... an essential feature of learning is that it creates the zone of proximal development, that is, learning awakens a variety of internal developmental processes that are able to operate only when a child is interacting with people in his environment and in cooperation with his peers (1978, p. 90).

The concept of the ZPD draws attention to the quality of the social situation for development, specifically the developmental potential of the environment. The processes of mediation involve children making meaning through cultural tools, symbol systems and materials, drawing on their own motives and agency. Taking a learning-leading-development perspective, Holzman (2017, p. 29) understands the ZPD as a collective form of working together and argues for emotionality to be considered along with the traditional emphasis on cognition. Holzman (2017, p. 121) further argues that if we accept the ZPD as a "life space in which and how we all live, inseparable from the we who produce it, then explanations for how child and world are "connected" are not needed." From contemporary perspectives, the ZPD might equally be understood as a zone for potential learning.

In ECE, the ZPD been interpreted as a pedagogical zone where the teacher/adult (often positioned as the more knowledgeable other) supports a child, or children to learn specific skills, behaviors or concepts. Pedagogical strategies include direct, systematic instruction, as well as guided participation, scaffolding, intersubjectivity, and co-playing. Activities may be formal/structured or informal and open-ended, and are typically organized via learning centers to direct children's engagement with specific materials and concepts (for example in Literacy, Mathematics, Science), or providing space, resources and time for freely-chosen play. However, understanding the ZPD as a pedagogical zone demands attention to who, and what is in the ZPD, and raises questions about who decides what is "proximal" for an individual child, or group, and what is the further potential or direction of development or learning in the activity? In an adult-directed pedagogical activity, the adult may decide the object of the activity - what is proximal for the child or group. In many countries, early childhood educators are likely to be using curriculum frameworks that act as cultural tools to describe or prescribe learning goals, and structure the course of learning and development over time. In child-led activities, children co-construct events and experiences that draw on different proximities and motives, creating a mash-up of knowledges, skills, languages and cultural repertoires. Contemporary interpretations of the ZPD question the much-used mantra that learning must always start from the child, and must be in alignment with the current level of development (which is usually determined at an individual-developmental level). Such a starting point does not acknowledge the potential for collective creativity, or children's collective agency in influencing their learning and development, even where the next steps may lie at the limits of their capabilities. Learning that "starts from the child" but without challenge is likely to leave children where they started, and may not provide the participation, engagement and motivation needed to realize potential. In the context of play, "starting from the child" demands attention to the imaginative qualities of play where children may be drawing on diverse sources of knowledge and interests (including popular culture, home and family practices) that are not always visible to adults in ECE settings.

A key theoretical challenge in ECE is to understand how children progress in their knowledge, skills and understanding. From a Piagetian perspective, progression is linear - from simple to complex, or “mature” forms of thinking and action. As noted previously, theories about children’s development progressing through “ages and stages” have proved appealing to policy makers in designing curriculum frameworks and organizing goals in linear sequences. In contrast, socio-cultural theories problematize progression as continuous emergence from everyday to scientific (discipline-based) concepts, with pedagogical support, as children engage in contexts and shared activities with different forms of complexity. Contemporary understanding of emergence supports the proposal that learning leads development, which, as [Hatch \(2010\)](#) argues, provokes different theoretical questions, and has different implications for practice. Everyday concepts can be seen as emergent, naive or intuitive, and scientific concepts are more highly structured and developed. For example, mathematics provides the conceptual tools for children to count, order, organize, estimate, measure, share and divide. However, the power of those concepts is realized as children use and apply them in different contexts, showing a co-relationship between everyday and scientific concepts. Being and becoming a mathematician is about the inter- and intra-subjective experience of being in a world that can be understood, organized and expressed in mathematical ways, and enables children to create and solve problems with a variety of tools.

Whether we look at the creative arts, sciences or humanities, within developmental environments, or interrelated systems, there is no “end point” of development, and no fixed definition of “mature” thought or specification of “higher mental processes.” For [Rogoff \(2003\)](#) development occurs when the balance of power shifts toward the learner as they take increasing management of, and responsibility for, activities. It is here that the transformative potential of learning is experienced as an ongoing and cumulative process. If there is no end point of learning or development, then specifying “outcomes” at a certain age (such as the transition from pre-school to compulsory schooling) is more to do with the socio-political framing of education, and the specification of curriculum goals in policy frameworks.

If the ZPD is interpreted as a metaphor for understanding the potential of human development ([Holzman, 2017](#)), there are implications for understanding the dialectical relationship between cultural tools and symbol systems, materials and resources, relationships and interactions, and the pedagogical or relational strategies that are used by more or differently knowledgeable others (not just “teachers”). These “others” may be human—peers and adults, and non-human, including material, digital, spatial, physical and virtual resources.

Cultural tools and symbol systems (tools for use/tools and use)

Understanding learning as a socio-culturally mediated process, [Holzman \(2017\)](#) discusses the Vygotskian concept of tools for use/tools and use, and the role of tools in creative and transformative activity. Children learn to think and act with and through cultural tools and symbol systems, with language/languages at the center of understanding and making meaning. In contemporary contexts, the “more” or “differently” knowledgeable other can be physical or virtual, including cultural tools such as the internet, Apps and social media. The meditational role of tools (tools for use), and the meanings children ascribe to them (tools and use), motivate activity, including how they communicate and participate in different contexts, often with imagination and creativity.

Vygotsky proposed a dynamic interaction between thought and language (or thinking and speech) that combines inter- and intra-individual cognitive and social processes. For example, as children collectively play and learn together, they combine individual and shared communicative repertoires. They remember, plan, organize and reflect, so that cognitive and meta-cognitive development is distributed across people, places, experiences, events and materials. These are the means by which cognition in use and in social interaction becomes internalized, appearing, in Vygotskian terms, first on the social and then on an individual plane. Contemporary research understands language, and languages, as more than grammar (syntax, morphology, phonology and semantics). Children’s communicative practices are multi-modal, including speech, physical gestures, musicality, movement, facial expressions, drawings, constructions, signs and symbols (including sign languages). In multi-diverse societies, many children are growing up as bi- and multilingual, which shifts attention from the dominant culture and language of the classroom, school and wider society, to shared knowledge of many cultures and languages.

Socio-cultural theories have been re-contextualized in the context of super-diverse, multi-lingual societies, with attention to the power relationships in children’s lives. Questions about what is culturally appropriate include concerns with power relationships, and critical engagement with social structures and systems. Research by [Tatham-Fashanu \(2021\)](#) focused on mono- and multi-lingual children, aged 4–5 in a primary school in England, to examine how they established and experimented with their individual identities through their language choices. Combining socio-cultural and post-structural perspectives, Fashanu challenged dominant language discourses, specifically the way in which “English-only” was promoted as a means of engaging the children in language learning, de-valuing their home and community languages. Using ethnographic methods, many instances of children subverting or contesting this dominant discourse were recorded, notably during lunchtime, and free play and free choice activities ([Fashanu et al., 2020](#)). Children were observed using their home languages, and explaining the meanings of words or phrases to their peers. The children were skilfully navigating the space, subverting the rules, and establishing their own linguistic identities and repertoires. The research reveals how dialectical relationships between the individual/social are infused with power dynamics, where children find ways of resisting rules imposed by adults, and find spaces in which to exercise agency and participate in social practices that are transformative on their own terms.

Technological innovations have brought about transformative changes in play, work and leisure, with implications for understanding children’s motives and agency in learning and development, and for disrupting institutional and inter-personal power relationships ([Erstad et al., 2017](#)). The uptake and use of digital technologies has been rapid, motivated by touchscreen technologies,

which were designed from the outset to be accessible for young children, and to offer new potential for creativity and multi-modality. There are no “developmental pathways” to prescribe children’s engagement with digital technologies. The concept of convergence (including converged or connected play) explains how digital technologies have blurred the boundaries between formal and informal contexts for learning, physical/virtual, traditional/digital, online/offline, place/space, play/work, and between children as novices and adults as experts (Marsh, 2017; Edwards et al., 2020). From the perspective of convergence, children’s engagement with digital technologies (as tools for use) exemplifies learning leading development: children acting with technologies (tools and use) are motivated to create different possibilities and trajectories for learning and development, based on multi-modal communicative and digital literacy practices.

Interpreted as a collective potential space, the concept of the ZPD opens possibilities for understanding children’s emergent ways of thinking, making meaning and coming to know. In contemporary interpretations of the ZPD, children are active participants in shaping and mediating their own developmental environments, in collaboration with cultural tools and symbols systems. Respecting children as active participants raises concerns with power relationships, and children’s agency and motivation.

Children’s agency and motivation

From a socio-cultural perspective, human beings are agentic because they are constantly

Reshaping the very environments that determine them, performing who they are becoming, creating culture and transforming the world. The “products” of this activity are not outcomes, but part of the unity that is “process and product” or “tool and result.”

Holzman (2009, p. 20).

Holzman’s concepts of reshaping, performing and transforming indicate that children are capable citizens who participate in the lives and activities of their communities. They do not wait for adults to organize knowledge, skills and concepts into manageable or simplified stages. Cognition, affect, relationships and imagination are part of this unity as children co-create their cultural identities and capabilities through their interests and inquiries (Hedges, 2022), and extended repertoires of participation. Progression is understood as children participating collectively in more complex social situations for development, using different forms of cultural knowledge, and performing with increasing agency, expertise and control. Children find their own ways of pitching in, and contributing to the development of the activity, with movement and direction building over time, and not just in one event or episode. From this perspective, the established concepts of “active learning” through “hands-on experiences” require further elaboration because the focus has shifted from agency as an individual/developmental trait, to agency being collective, relational and distributed. Acting with agency, and cultural tools, children drive their own and each other’s development in ways that may, over time, align with curriculum goals, but are likely to extend beyond those boundaries to reflect the complexity of play and of children’s life experiences. Socio-cultural theories understand play as a social practice, fundamental to children’s agency and motivation.

Play as social practice

Referring to the canon of knowledge that underpinned ECE from the 19th and early 20th centuries, play is valued as a natural developmental process, the main focus of the child’s activity and the child’s world. The concepts of child-initiated activity, freedom, discovery and exploration, have remained central to the powerful discourse that reifies play as the predominant way in which children learn. However, within global policy flows government investment in pre-school education is framed by discourses such as raising quality, improving children’s outcomes, and addressing social inequalities (Hunkin, 2021). Play has been drawn into socio-political discourses of educational and economic effectiveness, with many regions and countries creating curriculum guidance documents with varying degrees of prescription and description of desired outcomes (Gunn and Nuttall, 2019; Yang and Li, 2019). Within those frameworks, play occupies an ambivalent position (Wood, 2014), especially within the education-economic rationale for government investment. The emphasis in many policy frameworks has shifted toward a technical or utilitarian perspective, where play must pay into curriculum outcomes and goals (Hedges et al., 2018; Wood, 2020; Yang and Li, 2019). There are debates about the relative merits of child-initiated and adult-led play, and recommendations about play-based pedagogy and play-based curriculum can be in tension with children’s peer cultures, choices and interests (Chesworth, 2019). The move from pre-school to compulsory school varies internationally from age 4/5 to age 6/7 years old, with play being phased out in the transition year to ensure that children are “school ready.” The policy focus on “educational play” should not detract from the importance of play for its own sake, regardless of the age at which children make this transition.

Debates about the relative merits of child-initiated and adult-led play foreground key socio-cultural principles about the role and value of play for children’s meanings and purposes, and not necessarily for producing defined learning outcomes. It is worth returning to the concept of “play for its own sake,” because play connects children’s home, family and community life to the distinctive culture of education settings. Being in a pre-school or kindergarten offers new and varied opportunities for play, with different adults, peers, materials, space, tools and resources. Although Vygotsky wrote little about play, his indicative theories have been foundational to further research on understanding play across time, contexts and cultures. Vygotsky (1978) understood play as

a leading activity in early childhood. This does not mean that play is the most important or the predominant activity, rather that play creates zones for proximal development that are child-initiated and child-led. Play integrates social and symbolic complexity, expressed through language and multi-modal communicative practices. Cultural tools and resources may be imbued with pre-determined meanings (a pencil for writing, a buggy for a baby), but meanings can be imaginatively reconfigured in the context of play (a pencil becomes a light saber, a buggy becomes a forklift truck).

For [Vygotsky \(1978\)](#) and subsequent theorists, freedom of choice, imitation, imagination, and activity with imagination, are defining characteristics of play. [Vygotsky \(1978\)](#) identified the powerful role of imitation in play. Imitation is not merely “copying” what children see and experience in their everyday social worlds. Rather imitation is a highly creative process that involves mediation, creative interpretation, and recontextualization, with the addition of imagination and pretense to create and convey meaning. Imitation is informed by diverse sources of knowledge as well as everyday roles and practices in children’s homes, communities and families. Within play, children’s interests, and their cultural and linguistic repertoires become visible, because they can use their environment to establish and experiment with their identities through their language choices ([Fashanu et al., 2020](#)). Cultural tools and resources can be re-mediated and re-contextualized according to the nature of the activity, or the play theme, and children’s purposes or intentions.

Different interpretations of the ZPD can be found in a substantial body of work that focuses on children’s funds of knowledge and interests in play ([Chesworth and Hedges, 2023](#); [Hedges, 2022](#)). Play creates zones for proximal development that are child-led, collective and relational, and draw on the cultural resources of the setting as well as children’s home and community lives. [Chesworth \(2019\)](#) draws on sociological theories to focus on children’s peer cultures, and understands play as a shared social practice that supports the development of intersecting interests, identities, languages and knowledge/ways of knowing. [Hill and Wood \(2019\)](#) report an ethnographic study of children, aged 4–5, building working theories in their play-based activities. The children contributed collectively to creating interests as inquiries, and to co-constructing knowledge. They drew on working theories about human nature, the social world, and the physical and natural world, as well as existential matters of birth, life and death, which derived from their home and family experiences, including popular culture, religious beliefs and everyday practices.

Play remains a complex space for research, theory and practice, with the addition of policy as a fourth dimension. Analyses of early childhood policy frameworks indicate that play has earned its place within the broad recommendations for play-based learning, play-based curriculum and play-based pedagogy ([Hedges et al., 2018](#)). However, the capture of play in policy discourses and curriculum frameworks is a mixed blessing because there are varying interpretations of what counts as play, and what play is expected to produce. Finding a “best fit” between the apparent orderliness of curriculum frameworks and the complexities of play remains problematic ([Ruscoe and Barratt-Pugh, 2021](#); [Wood, 2013](#)), and the urge toward “structured” and “guided” play leans more toward apparent orderliness. With the play/education debates, there are different positions. Free play is seen as relevant for achieving developmental goals, and adult-guided/structured play being more appropriate for achieving academic goals ([Pyle and Danniels, 2017](#)). This bifurcation indicates that biological/maturational theories of play retain power within the field, where development leads learning. At the same time, curriculum frameworks in different countries exert different degrees of pressure on educators, according to the levels of prescription of defined outcomes (for example in England) and description of indicative goals (for example in New Zealand) ([Barblett et al., 2016](#); [Gunn and Nuttall, 2019](#); [Wood and Hedges, 2016](#); [Wood, 2020](#)). Connecting play with curriculum and pedagogical approaches, [Pyle and Danniels \(2017\)](#) propose that “purposeful play” sits on a continuum between developmental goals and academic goals, but this remains a problematic construct depending on whose purposes are privileged—those of the children, or those of the educators. [Fleer \(2015\)](#) focused on the pedagogical roles of educators inside and outside imaginary play situations. The educators in the study acted in five different ways according to the physical positioning or proximity to the play, the levels of intersubjectivity between adults/children, resourcing and supporting the play theme, being engaged with the play, and being inside the play, in imaginative but restricted ways. Fleer’s research exemplifies the range of possible pedagogical decisions and actions that educators take in relation to play, and the challenges of sustaining imaginative play against more instrumental discourses within ECE policy frameworks. For Vygotsky, school instruction and work are “compulsory activity based on rules” (1978, p. 104). This is not to downplay how children engage with the knowledge around which curriculum content is framed, as they are being/becoming mathematicians, geographers, scientists, artists, technologists, readers, writers, designers and players. However, children also draw on sources and funds of knowledge from their own cultural repertoires to inform how they experience and build curriculum ([Chesworth, 2016](#); [Hedges, 2022](#); [Hill and Wood, 2019](#)). Extending these concepts, [Stagg-Peterson and Friedrich \(2022\)](#) elaborate the interrelationship between play and place as “playce” to foreground Indigenous philosophies and ontologies, in which the nature of being/becoming is embodied in concepts of Land, place and space.

The role of early childhood educators is critical to interpreting policies in ways that honor the significance of truly free play in ECE settings, but at the same time ensure that all children can be involved in creatively shaping their own developmental environments. From the perspective of play, curriculum is not understood as hierarchical or linear goals. Rather curriculum design and enactment are dynamic and co-constructive processes, informed by children’s identities, interests, heritages, languages, knowledges and practices. For early childhood educators, the pedagogical implications are complex: understanding what lies beyond the boundaries of curriculum goals and outcomes informs the design of culturally responsive curricula that includes and values children’s diverse home and community knowledge.

In light of the substantial investment in ECE in many regions of the world, it is not surprising that play has been drawn into debates about the different purposes and practices of education. Play occupies a tenuous position in practice and is perhaps most vulnerable to being marginalized in favor of adult-led activities, or guided play, as a means of ensuring children achieve defined goals, and are “ready” to make the transition to compulsory education. Debates about the role and value of play are

enmeshed with wider debates about what and whose knowledges are valued, and how ECE curricula should be designed to reflect dimensions of diversity. Play is thus implicated in the global politics of knowing, learning, teaching, and governing.

Section 3: What are the implications of socio-cultural theories for early childhood education?

This chapter has focused on some of Vygotsky's original ideas, and their subsequent interpretations over time, along with critical perspectives from post-structural theories. The continued working out of socio-cultural theories remains an important endeavor in ECE practice, and in policy frameworks. The two orientations discussed here—development leads learning, and learning leads development—are not mutually exclusive, and can be understood in dialectical relationship. However, an emphasis on one at the expense of the other produces different ways of understanding children's agency, with different implications for creating education spaces, and designing curriculum, pedagogical and assessment practices. Moreover, embedding socio-cultural theories as the foundational knowledge for pedagogy, curriculum and assessment practices remains challenging.

Debates about the nature of pedagogy and curriculum, and the place of play in ECE have intensified in light of trends toward policy frameworks that prescribe or describe desired learning goals and outcomes. Focusing on the Early Years Foundation Stage (EYFS), the statutory framework for children from birth to 5 in England (DfE, 2017), Wood (2020) uses critical discourse analysis to reveal how the prescribed learning outcomes exemplify a global preoccupation with policy technologies to determine linear notions of progress, as a means of shaping children's developmental pathways toward "school readiness." The framing of curriculum and pedagogy is informed by a standardized conceptualization of knowledge in which predetermined learning outcomes are assumed to have universal relevance for all children (Chesworth, 2019), and is aligned with accountability and regulation. Pedagogy and assessment practices are organized to provide comparative data about the extent to which children have acquired, and can demonstrate, developmental goals as learning outcomes. The EYFS exemplifies how selected child development theories remain influential in the field, and have been interpreted to sustain a development leads learning orientation. In contrast, the ECE policy frameworks in Australia (the Early Years Learning Framework) and New Zealand (Te Whāriki) are informed explicitly by contemporary interpretations of socio-cultural theories, as well as critical post-structural perspectives (Gunn and Nuttall, 2019). However, there remain challenges for initial and continuing professional development programs in engaging teachers with more complex conceptualizations of children's learning over time. Chesworth and Hedges (2023) identify theoretical and methodological heterogeneity in ECE, incorporating developmental psychology, sociocultural theory, new materialism, and post-humanism. This heterogeneity is informing the urgency of undoing normalization (Vintimilla and Pacini-Ketchabaw, 2020) by challenging normative child development theories, their ethnocentricity (Rogoff, 2003), and their various instantiations in ECE policy frameworks (Wood and Hedges, 2016; Hedges et al., 2018). In combination with contemporary theories, the continuing endeavor to understand changing socio-cultural/cultural-historical contexts for learning and development is informed by critical perspectives. Different interpretations of socio-cultural theories continue to emerge, in ongoing efforts to address new social challenges that play out at local and global levels.

Vygotsky has been described as a "revolutionary scientist" because he was concerned with profound and progressive social change (Holzman, 2017). A century later, it seems that we need to sustain a revolutionary approach to understand children's learning in the context of the profound and progressive changes of the 21st century. Global movements which foreground ethical commitments to children's rights, decolonization, and postcolonial practices highlight the structural inequalities that exist for particular communities. Socio-cultural theories have been reinterpreted to foreground the significance of knowledges and practices within Indigenous communities, including cultural identity and sense of belonging, and how these can inform the contemporary focus on the environment, sustainability and climate change (Stagg Peterson and Friedrich, 2022). Post-developmental theories draw critical attention to the discursive effects of "normal" or "typical" patterns of development (Hatch, 2010, 2012; Vintimilla and Pacini-Ketchabaw, 2020) to address concerns with diversities, equity and inclusion, and generate new understandings of identity and culture. A growing field of critical policy analysis foregrounds concerns with power, and power relationships in governance and policy formulation, in the wider contexts of global policy flows (Kay, 2022; Hunkin, 2021; Wood, 2020). New materialism and post-human theories challenge the primacy of the human in theories of development, and draw attention to relationships between humans and non-humans, place/land, space, materials, tools, and the natural world (Pacini-Ketchabaw and Taylor, 2015; Stagg Peterson and Friedrich, 2022).

Contemporary interpretations of socio-cultural theories remain integral to these challenges. As Rogoff et al. (2018) have argued more attention is needed to the cultural paradigms of children's lived experiences, including how they learn to navigate across, and participate in the distinct cultural settings of their everyday lives. In light of these complexities, a continuing project for all involved in ECE is to engage critically across the intersections of theory-policy-practice to articulate the purposes and practices of education. This is an important point because debates no longer center mainly on the theory-practice relationship in ECE. Rather debates center on policy, theory and practice, ethics and politics, and on critical engagement with the selected research and evidence that underpins policy (Wood, 2020). Research within the learning sciences continues to be relevant to ECE (Goswami, 2015), but it is important to remain critically aware of what methods are used, what forms of knowledge are produced in different disciplinary orientations, and how findings are used selectively in the contexts of policy and practice. Ellefson et al. (2019) draw critical attention to the structural and cultural differences between educational and psychological research, and the complexity of using research findings to inform policy making and decisions. They argue that

Findings can be much more influential if they account for this complexity from its earliest conception through shared agenda setting and the cocreation of more appropriate research designs that benefit both education and psychology (2019, p. 254).

For post-structural researchers, methods and methodology are part of the challenge, specifically the notion of research as the application of specific methods or tools, in a linear process, to produce results that inform policy ideologies, often in piecemeal or selective ways. Hunkin (2021) argues for a new advocacy focus on questioning edu-policies and politics on ontological and epistemological grounds, particularly the different conditions of workforce and governance that mark ECE as distinct from compulsory schooling. Hunkin's work foregrounds the necessity for political engagement with policy discourses, and their promotion of particular truths, including early learning inputs and outcomes (p. 203).

Different theoretical perspectives have brought about methodological innovations, which were not used in Vygotsky's original research, as well as ontological and epistemological shifts in understanding learning and development. In particular, an ethical focus on children as participants in research respects them as informed interpreters of their own lives and experiences (Robson and Flannery Quinn, 2015). What we understand as social, cultural, and material has expanded with rapid shifts in technologies, their uptake by young children, the convergence between physical and digital play and everyday activities, and the global influence of popular culture. Post-colonial attention to inclusion, equity and diversities in families and communities challenges ethnocentric assumptions in policy, theory and practice. Ethical-political concerns regarding the decolonization of theory and practice locate ECE within the politics of knowing, learning and teaching (Gunn and Nuttall, 2019) and draw attention to new modes of governing, and their impacts, in the context of global/local reform movements (Hunkin, 2021; Lee, 2021).

Socio-cultural/cultural-historical theories continue to be relevant for understanding human development in contemporary contexts, where rapid changes and new social challenges are provoking new relationships between human beings and society, between human- and non-human beings, and between humans and the natural and made environments. These issues imply that the education and care of young children needs to acknowledge their agency and capacities for engaging in and with their social worlds, specifically matters that interest and concern them. Socio-cultural theories remain relevant for articulating and critiquing the diverse policies, purposes and practices of early childhood education. However, a key challenge for the field is to further explore the theoretical heterogeneity identified in this chapter, and the implications for informing ECE in the context of social and technological change, and global influences.

References

- Barblett, L., Knaus, M., Barratt-Pugh, C., 2016. The pushes and pulls of pedagogy in the early years: competing knowledge and the erosion of play-based learning. *Austral. J. Early Childhood* 41 (4), 36–43. <https://doi.org/10.1177/183693911604100405>.
- Bartholomaeus, C., 2016. Developmental discourses as a regime of truth in research with primary school students. *Int. J. Qual. Stud. Educ.* 29 (7), 911–924. <https://doi.org/10.1080/09518398.2016.1174896>.
- Bruner, J., 1996. *The Culture of Education*. Harvard University Press, Cambridge, Mass.
- Chen, J.-Q., Masur, A., McNamee, G., 2011. Young children's approaches to learning: a socio-cultural perspective. *Early Child. Dev. Care* 181 (8), 1137–1152. <https://doi.org/10.1080/03004430.2010.53004430>.
- Chesworth, L., 2016. A funds of knowledge approach to examining play interests: listening to children's and parents' perspectives. *Int. J. Early Years Educ.* 24 (3), 294–308.
- Chesworth, L., 2019. Theorising young children's interests: making connections and in-the-moment happenings. *Learn. Cult. Soc. Interact.* 23, 100263. <https://doi.org/10.1016/j.lcsi.2018.11.010>.
- Chesworth, L., Hedges, H., 2023. Children's interests and curriculum making in early childhood. In: Tierney, R.J., Rizvi, F., Ercikan, K. (Eds.), *International Encyclopaedia of Education*, 4th. Elsevier. In press.
- Daniels, H., 2008. *Vygotsky and Research*. Routledge, England and New York.
- Daniels, H., 2016. *Vygotsky and Pedagogy*. Routledge, England and New York (Classic Edition).
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D., 2020. Implications for educational practice of the science of learning and development. *Appl. Dev. Sci.* 24 (2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>.
- DfE (Department for Education), 2017. *Statutory Framework for the Early Years Foundation Stage: Setting the Standards for Learning, Development and Care for Children From Birth to Five*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596629/EYFS_STATUTORY_FRAMEWORK_2017.pdf. (Accessed 1 April 2021).
- Edwards, S., Mantilla, A., Grieshaber, S., Nuttall, J., Wood, E., 2020. Converged play characteristics for early childhood education: multi-modal, global-local, and traditional-digital. *Oxf. Rev. Educ.* 46 (5), 637–660. <https://doi.org/10.1080/03054985.2020.1750358>.
- Ellefson, M.R., Baker, S.T., Gibson, J.L., 2019. Lessons for successful cognitive developmental science in educational settings: the case of executive Functions. *J. Cognit. Dev.* 20 (2), 253–277. <https://doi.org/10.1080/15248372.2018.1551219>.
- Erstad, O., Flewitt, R., Kümmerling-Meibauer, B., Pereira, S., 2017. *The Routledge Handbook of Digital Literacies in Early Childhood*. Routledge, London and New York.
- Fashanu, C., Wood, E., Payne, M., 2020. Multilingual communication under the radar: how multilingual children challenge the dominant monolingual discourse in a super-diverse, Early Years educational setting in England. *Engl. Educ.* 54 (1), 93–112. <https://doi.org/10.1080/04250494.2019.1688657>.
- Fleer, M., 2015. Pedagogical positioning in play – teachers being inside and outside of children's imaginary play. *Early Child. Dev. Care* 185 (11–12), 1801–1814. <https://doi.org/10.1080/03004430.2015.1028393>.
- Goswami, U., 2015. *Children's Cognitive Development and Learning*. Cambridge Primary Review Trust, York.
- Gunn, A.C., Nuttall, J., 2019. *Weaving Te Whāriki*. Aotearoa New Zealand's Early Childhood Curriculum Document in Theory and Practice, third ed. New Zealand Council for Educational Research, Wellington, New Zealand.
- Gutiérrez, K.D., Rogoff, B., 2003. Cultural ways of learning: individual traits or repertoires of practice. *Educ. Res.* 32 (5), 19–25.
- Hatch, J.A., 2010. Rethinking the relationship between learning and development: teaching for learning in the early childhood classroom. *Educ. Forum* 74 (3), 258.
- Hatch, J.A., 2012. From theory to curriculum: developmental theory and its relationship to curriculum and instruction in early childhood education. In: File, N., Mueller, J., Wisneski, D. (Eds.), *Curriculum in Early Childhood Education: Re-examined, Rediscovered, Renewed*. Routledge, London, pp. 42–53.

- Hedges, H., 2021. The place of interests, agency and imagination in funds of identity theory. *Mind Cult. Activ.* 28 (2), 111–124. <https://doi.org/10.1080/10749039.2020.1833931>.
- Hedges, H., 2022. *Children's Interests, Inquiries and Identities: Curriculum, Pedagogy, Learning and Outcomes in the Early Years*. Routledge, London.
- Hedges, H., Stagg Peterson, S., Wajskop, G., 2018. Modes of play in early childhood curricular documents in Brazil, New Zealand and Ontario. *Int. J. Play* 7 (1), 11–26. <https://doi.org/10.1080/21594937.2018.1437379>.
- Hill, M., Wood, E., 2019. 'Dead Forever': an ethnographic study of young children's interests, funds of knowledge and working theories in free play. *Learn. Cult. Soc. Interact.* 23, 100292. <https://doi.org/10.1016/j.lcsi.2019.02.017>.
- Holzman, L., 2009. *Vygotsky at Work and Play*. Routledge, New York and Sussex.
- Holzman, L., 2017. *Vygotsky at Work and Play*, second ed. Routledge, London and New York.
- Hunkin, E., 2021. Devastating impacts? Investigating 'edu-quality' discourse in early childhood policy and its implications. *J. Educ. Pol.* 36 (2), 196–210. <https://doi.org/10.1080/02680939.2019.1690709>.
- Kay, L., 2022. 'What works' and for whom? Bold Beginnings and the construction of the school ready child. *J. Early Child. Res.* 20 (2), 172–184. <https://doi.org/10.1177/1476718X211052791>.
- Lee, S.F., 2021. Governing 'disadvantage' through funded early years places and reconfigured spaces. *J. Early Child. Res.* 19 (2), 267–280.
- Lenz-Taguchi, H., 2010. Going Beyond the Theory/Practice Divide in Early Childhood Education. Introducing an Intra-active Pedagogy. Routledge, Abingdon.
- Marsh, J.A., 2017. The internet of toys: a posthuman and multimodal analysis of connected play. *Teach. Coll. Rec.* 119.
- National Association for the Education of Young Children, 2020. Position Statement: Developmentally Appropriate Practice. www.naeyc.org/DAP. (Accessed 5 January 2022).
- Newman, F., Holzman, L., 2013. *Lev Vygotsky: Revolutionary Scientist*. Psychology Press, New York.
- Pacini-Ketchabaw, V., Taylor, A., 2015. *Unsettling the Colonial Places and Spaces of Early Childhood Education*. Routledge, Abingdon.
- Pyle, A., Danniels, E., 2017. A continuum of play-based learning: the role of the teacher in play-based pedagogy and the fear of hijacking play. *Early Educ. Dev.* 28 (3), 274–289.
- Robson, S., Flannery Quinn, S., 2015. *The Routledge International Handbook of Young Children's Thinking and Understanding*. Routledge, London and New York.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press, Oxford.
- Rogoff, B., Dahl, A., Callanan, M., 2018. The importance of understanding children's lived experience. *Dev. Rev.* 50 (a), 5–15. <https://doi.org/10.1016/j.dr.2018.05.006>.
- Ruscoe, B., Barratt-Pugh, 2021. Discourses in power: policy and curriculum demands in the first year of compulsory school. *Contemp. Issues Early Child.* <https://doi.org/10.1177/1463949121990905>.
- Stagg Peterson, S., Friedrich, N., 2022. *The Role of Place and Play in Young Children's Language and Literacy*. University of Toronto Press, Toronto.
- Sutton-Smith, B., 1997. *The Ambiguity of Play*. Harvard University Press.
- Tatham-Fashanu, C., 2021. A third space pedagogy: embracing complexity in a super-diverse, early childhood education setting. *Pedagog. Cult. Soc.* <https://doi.org/10.1080/14681366.2021.1952295>.
- Vintimilla, C.D., Pacini-Ketchabaw, V., 2020. Weaving pedagogy in early childhood education: on openings and their foreclosure. *Eur. Early Child. Educ. Res. J.* 28 (5), 628–641.
- Vygotsky, L.S., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society. The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Wood, E., 2013. *Play, Learning and the Early Childhood Curriculum*. Sage, London.
- Wood, E., 2014. The play-pedagogy interface in contemporary debates. In: Brooker, L., Edwards, S., Blaise, M. (Eds.), *The SAGE International Handbook of Play and Learning*. Sage, London, England, pp. 145–156.
- Wood, E., 2020. Learning, development and the early childhood curriculum: a critical discourse analysis of the Early Years Foundation Stage in England. *J. Early Child. Res.* 18 (3), 321–336.
- Wood, E., Hedges, H., 2016. Curriculum in early childhood education: asking critical questions about content, coherence and control. *Curric. J.* 27 (3), 387–405. <https://doi.org/10.1080/09585176.2015.1129981>.
- Yang, W., Li, H., 2019. Changing culture, changing curriculum: a case study of early childhood curriculum innovations in two Chinese kindergartens. *Curric. J.* 30 (3), 279–297. <https://doi.org/10.1080/09585176.2019.1568269>.

Key developments during adolescence: implications for learning and achievement

Eric M. Anderman, Yvonne Allsop, Kimiko Ching, and Seung Yon Ha, The Ohio State University, Department of Educational Studies, Columbus, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Key developments in adolescence	486
Cognitive development and adolescent learning	487
Piagetian perspective of adolescent learning and development	487
Sensorimotor stage	487
Preoperational stage	487
Concrete operations stage	487
Formal operations stage	487
Limitations of the Piagetian perspective	488
Vygotskian perspective of adolescent learning and development	488
Metacognitive and information processing perspectives	488
Summary	488
Social/moral development	489
Theoretical perspectives on social/moral development	489
The roles of adolescent social relationships and experiences	489
Expanding social networks	490
Summary	490
Identity development and adolescent learning	490
Exploring and changing identities	490
Possible selves and self-concept	491
Gender identity	491
Racial and ethnic identity	491
Summary	492
Academic motivation development	492
Expectancies and values	492
Achievement goal theory	492
Why do motivational beliefs change during adolescence?	493
Summary	493
Discussion	493
Conclusion	494
References	494

Key developments in adolescence

Adolescence is a pivotal developmental period. During the adolescent years, youth are developing rapidly across a wide range of domains. They are experiencing changes in the ways that they think, in their beliefs about who they are and whom they might become, and in their relationships with peers and adults. The numerous changes that occur during this time period greatly impact adolescent learning.

We focus in this chapter in particular on developmental trajectories and milestones that affect academic learning during adolescence. When educators, parents, and policymakers are aware of the complexity and inter-relatedness of development across a variety of domains during adolescence, they are better able to provide developmentally appropriate instructional practices to support adolescents' success in their academic pursuits. This is important, because research clearly indicates that when instructional practices and contexts are aligned with adolescents' developmental needs, adolescents are likely to achieve at high levels and maintain optimal motivation (Eccles and Roeser, 2009).

We discuss four domains of development: cognitive, social/moral, identity, and motivation. Whereas development within each of these domains can be discussed in isolation, it is essential for educators and policymakers to acknowledge that adolescent development occurs across these four domains simultaneously; thus changes in one domain (e.g., cognitive development) are not independent of changes in another domain (e.g., moral development). These changes, though largely predictable and generalizable to adolescents across different cultures, may naturally occur at stages unique to each individual as he or she progresses from childhood to adolescence and into adulthood.

We first briefly review the major developmental milestones associated with each of these domains. We then provide a more holistic overview of the interrelations of development across these domains.

Cognitive development and adolescent learning

Of the fundamental changes that occur during adolescence, cognitive changes have far-reaching implications for learning and the processes adolescents can use to enhance their learning. Generally, there is agreement that adolescents think differently than children. However, the processes that underlie the cognitive differences in learning between children and adolescents are less universally accepted. Various periods of growth experienced by an individual may be influenced by both biological (i.e., genetic) and environmental (i.e., physical and social/environmental) factors. The debate of heredity versus environmental influences on cognitive development was once seen as an “either-or” debate (Bradley et al., 1989; Greenfield and Bruner, 1966; Newman et al., 1937). However, scholars today readily agree that heredity and environment are intermingled in such a way that one’s influence cannot be separated from the other (e.g., Petrill et al., 2004; Plomin and Asbury, 2005; Tucker-Drob et al., 2013).

Cognitive development prior to and during adolescence is influenced by a variety of factors (e.g., genetics, nutrition, physical and social environments), and development is reflected by ongoing physiological changes in the brain (Bronfenbrenner, 1992). For this chapter, we focus on three psychological perspectives that describe cognitive development during adolescence, rather than focusing on neurological development in the brain. Specifically, we focus on Piaget’s theory, Vygotsky’s theory, and metacognitive or information processing perspectives of learning and development.

Piagetian perspective of adolescent learning and development

Jean Piaget’s theory of cognitive development (Piaget, 1929, 1952, 1964) has been highly influential in understanding the cognitive changes that occur throughout childhood and adolescence. In his observation of children and adolescents, Piaget noticed various changes in logical thinking patterns. He asserted that these cognitive shifts are both sequential and distinct, and that the type of thinking employed by children and adolescents is different. Known as a constructivist view of learning and development, Piaget noted that children are active, motivated learners who construct knowledge rather than absorb it. Children learn through assimilation (modifying new information to fit what one already knows) and accommodation (restructuring what one already knows to integrate new information). They also interact with their surroundings; these interactions may lead to disequilibrium in thought, ultimately promoting progress toward more complex thought.

According to Piaget (1964), the process of cognitive development proceeds through four separate stages, hierarchical in nature, and each characterized by a particular type of thinking: *sensorimotor* (birth through age 2), *preoperational* (generally ages 2–7), *concrete operational* (generally ages 7–11), and *formal operational* (approximately age 11 onward). It should be noted that individuals transitioning from one stage to the next may exhibit characteristics of two adjoining stages simultaneously. Also, as children progress from one stage to the next, they do not lose the abilities gained in prior stages while gaining newer, more advanced thinking abilities. Furthermore, Piaget’s stages of cognitive development describe how individuals can think, rather than how they actually, or always, do think. We very briefly describe the milestones associated with each of these stages.

Sensorimotor stage

Early in life, one’s actions and thoughts (known as schemas) are based mostly on exploration of an infant’s surroundings. Infants and toddlers use their senses and engage in actions (e.g., crawling and walking) that allow them to experiment through looking, hearing, touching, mouthing, and grasping. Very young children cannot think about things that are not directly in front of them, until object permanence and symbolic thought emerge.

Preoperational stage

During this stage, symbolic thinking abilities increase dramatically, allowing for greater ability to represent and refer to objects outside of immediate experience. This results in a great expansion of vocabulary and language structure, increased pretend play, and intuitive thought.

Concrete operations stage

Throughout this stage, children begin to think logically and understand analogies, but only concretely (i.e., pertaining to real-life situations). Abilities acquired during this stage include distinction between thoughts of self and others, categorization of objects in simultaneous categories, and conservation.

Formal operations stage

Young adolescents begin to apply logical thinking and reasoning processes to both abstract and concrete ideas and objects. Adolescents develop the ability to engage in logical and deductive reasoning, to consider hypotheticals, and to think about abstract topics. Moreover, metacognitive abilities develop further, thus enabling adolescents to think about their own cognitive and self-regulatory processes.

Limitations of the Piagetian perspective

Piaget's idea that children construct their own knowledge is widely accepted. However, Piaget's stages of development have limitations. For example, children have been found to engage in formal operations at younger ages (e.g., Kuhn, 2008; Siegler, 1998). Moreover, aspects of cognition may develop more sporadically, rather than through highly predictable stages (Keating, 2004, 2012; Kramer, 1983; Kuhn, 2008). In addition, the concepts of assimilation and accommodation, which are critical to Piaget's theory, remain quite vague, and interactions with the physical environment are not necessarily limited to direct sensory experiences (Bebko et al., 1992; Brainerd, 2003).

Vygotskian perspective of adolescent learning and development

Rather than children independently constructing their own knowledge, Lev Vygotsky believed that adults systematically influence and foster children's cognitive development. This is known as a sociocultural theory of cognitive development (Vygotsky, 1978, 1987). Vygotsky's sociocultural theory of development, first introduced nearly 90 years ago, framed cognitive development around the central notion that children completing tasks with adult assistance promotes cognitive development.

According to Vygotsky (1987), learning is a social activity. Learning occurs formally and informally through interactions with more experienced others. Cognitive tools (general learning techniques or processes) passed on to children and adolescents by more experienced peers, teachers, parents, or other influential adults, enhance cognitive abilities and eventually are internalized as individuals engage in various activities. Internalization of these cognitive tools leads children and adolescents to gradually learn to independently direct their own mental activities, rather than needing the support and guidance of others.

The zone of proximal development (ZPD), fundamental to Vygotsky's theory, is represented by the range of tasks or activities which a child or adolescent cannot yet perform independently but can perform with the guidance and help of a more experienced other (Vygotsky, 1978, 1987). A child can typically perform certain tasks successfully without the assistance of another; however, learning is typically not stimulated by independent completion of tasks one can already perform. In Vygotsky's view, learning occurs within the ZPD. The ZPD is limited however, as there are tasks a child cannot perform, even with aid and support of another. Naturally, an individual's ZPD will change over time as various tasks are mastered and more complex, difficult tasks are encountered. The ZPD is representative of one's actual (independent) development level and their level of potential development (Vygotsky, 1978, 1987).

Metacognitive and information processing perspectives

Thinking about thinking—or metacognition, is representative of an individual's own knowledge about his or her own cognitive processes. The characterization of thinking abilities that change from childhood into adolescence is encompassed by the information-processing perspective. Between childhood and adolescence, cognitive development is evident in youth's attention, memory, speed, organization, and metacognition (Demetriou et al., 2002; Keating, 2004). Improvements in each of these areas occur as individuals transition from childhood into adolescence, evidenced by the emergence of advanced cognitive abilities, such as abstract thought and hypothetical thinking (Keating, 2004).

The ability to attend to selective stimuli improves over time. Both selective (focus on one stimulus while turning out other stimuli) and divided (focus on two stimuli simultaneously) attention improve as children mature. Thus, adolescents are better able to focus on one stimulus while tuning out other stimuli, as well as pay attention to two or more stimuli simultaneously (Higgins and Turnure, 1984; Kuhn, 2009). Memory encompasses both working memory (short-term) and long-term memory. Adolescents are better able to temporarily store information in short-term memory than are younger children. Moreover, adolescents can more easily recall information stored in long-term memory (Kuhn, 2008). Adolescents are also better able to quickly and efficiently process and solve problems than are younger children (Kail and Ferrer, 2007). Additionally, adolescents are more adept at organization than are younger children (i.e., they are more planful; Siegler, 1988). Finally, metacognition refers to one's knowledge about their own cognitive processes. Adolescents can more easily think about their own thoughts, and experience greater self-consciousness, self-reflection, and self-awareness. Improvements in metacognition lead to an increase in the ability to monitor and reflect upon learning processes and strategies (Kuhn, 2009).

Summary

The three perspectives that we have reviewed each contribute to our understanding of cognitive development. Each perspective offers unique insights that are relevant to adolescent learning. For example, the Piagetian perspective reminds us that adolescents are more able than are younger children to understand and engage with abstract, complex ideas (e.g., debating about the causes of poverty). The Vygotskian perspective reminds us that many topics are best learned when adolescents can partner with more experienced peers or adults. Thus, an adolescent who is struggling with a complex trigonometry problem might be better able to understand and learn the strategies necessary to solve the problem when working with a partner who has a deeper understanding of trigonometry. The information processing perspective reminds us that there are multiple cognitive processes that develop throughout childhood and adolescence. For example, educators often can teach organizational strategies to adolescents, and expect that those adolescents will be able to learn and eventually apply those strategies in novel contexts.

Social/moral development

During childhood, the critical tasks of social/moral development and learning include the acquisition of an understanding of various norms and expectations that apply in different social/moral contexts, as well as an understanding of the beliefs and feelings of others (i.e., theory of mind). As children transition toward adolescence, multiple internal and external factors impact ongoing social/moral development and learning. For example, with enhanced cognitive ability (e.g., abstract and hypothetical thinking), adolescents are able to understand and interpret social/moral issues they encounter. Concomitantly, the social situations and relationships that adolescents experience expand and become increasingly complex and multifaceted, requiring them to think critically about the complex and sometimes conflicting issues which arise (e.g., equity, justice). Adolescents are more able than younger children to make sense of, evaluate, and reconcile various perspectives, in order to thoughtfully participate in society.

We briefly review several central issues in the study of adolescent social and moral development. First, we discuss several theoretical approaches regarding the development of individual social/moral reasoning during adolescence. Second, we discuss the roles that adolescent social relationships and experiences play in social/moral development and learning.

Theoretical perspectives on social/moral development

Current research and theories on social/moral development can be traced back to the work of [Piaget \(1932\)](#) and [Kohlberg \(1973\)](#), which suggested that humans use multiple social perspectives to make sense of social events. Piaget used the term “social-arbitrary knowledge” to describe the type of knowledge needed in the understanding of humans and social interactions. Social-arbitrary knowledge consists of the manners, rules, and signs accepted and agreed upon by members of a society. One can learn this kind of knowledge through various types of interaction with others, such as communication and instruction. [Kohlberg \(1964\)](#) viewed moral competence as “the capacity to make decisions and judgments, which are moral (based on internal principles) and to act in accordance with such judgment” (p. 425). According to Kohlberg, moral competence develops with age as individuals become increasingly able to distinguish morality from social conventions, and to mentally coordinate these different perspectives. [Kohlberg \(1973\)](#) asserted that most adolescents are capable of making social/moral decisions based on external rewards or punishments and self-interest, and that they begin to recognize the value of maintaining, supporting, and justifying interpersonal and societal rules and expectations (c.f., conventional level of moral development). However, Kohlberg also posited that in general, adolescents are not yet capable of making autonomous and conscious decisions based on abstract and ethical principles that are consistent with universal principles of justice and rights.

Extending Piaget and Kohlberg’s framework on social/moral development, more recent approaches have focused on how children and adolescents coordinate and apply different social and moral beliefs to different social/moral situations (e.g., peer interactions, social interactions involving different cultures). Social domain theory ([Turiel, 1983](#)) posits children’s (even young children’s) ability to construct three conceptually distinct domains of social understanding—moral, societal, and psychological—through social interactions and experiences. The moral domain pertains to understanding issues like justice, fairness, and rights. The societal domain involves the understanding of authority, tradition, norms, and expectations within social systems. The psychological domain includes matters that primarily affect the self, such as privacy, personal choice, and preferences. Additional domains, such as pragmatic considerations (i.e., understanding of practical issues such as convenience) and prudence (i.e., understanding issues related to one’s own safety, comfort, and health) were later identified and have been examined in more recent research ([Smetana and Turiel, 2003](#); [Tisak, 1993](#)). Developmental research based on the social domain perspective (e.g., [Nucci and Turiel, 2009](#)) suggests that in ambiguous social/moral situations, early adolescents tend to emphasize personal rights, whereas older adolescents exhibit a more sophisticated ability to consider and coordinate diverse aspects of social/moral situations (e.g., government has the right to constrain individuals’ rights for a larger public good). Thus during adolescence, the ability to understand social/moral situations extends beyond a focus on familiar and concrete social/moral events (e.g., physical harm); indeed, adolescents become more able to appreciate the complexity of the ambiguity that is inherent in social/moral discussions.

The roles of adolescent social relationships and experiences

There is a general consensus that interpersonal experiences, particularly with parents and peers, play critical roles in social/moral development ([Dodge et al., 1990](#); [Eisenberg et al., 2015](#)). Parents’ use of inductive discipline methods, including the use of explanations, rationales, and reasoning, as well as having an authoritative style of parenting, help facilitate the development of adolescents’ advanced moral reasoning and prosocial behavior. These parental practices provide information about the nature of social/moral transgressions, and encourage adolescents to feel empathy toward others, and reflect on their actions, contributing to prosocial development. One of the key processes that occurs during adolescence is the renegotiation of parent-child relationships. Parent-adolescent relationships that acknowledge adolescent needs for independence and autonomy and that enable open communication—rather than trying to control adolescent behavior and feeling—have been associated with positive adolescent outcomes, such as better self-control, better academic performance, and fewer school behavior problems (e.g., [Barber, 1996](#); [Lam and Ducreux, 2013](#)). On the other hand, adolescents living with hostile and rejecting parents are more likely to display hostile attribution bias, aggressive behavior, and fewer positive social skills, all of which are detrimental to successful interactions with peers ([Dodge et al., 1990](#); [Repetti et al., 2002](#)).

Younger children's notions of friends and friendship often refer to acquaintances and focus on simple, shared activities. Adolescents, however, begin to form more intimate relationships with their peers. Adolescents within peer groups tend to be similar to one another in a variety of attributes such as attitudes, values, and behavior. On one hand, adolescents select peers who are similar to them; on the other hand, friends who spend time together shape each others' attitudes, values, and behaviors. These two processes both play roles at different points in the formation and maintenance processes of peer relationships (Kandel, 1978).

Peer relationships can serve both positive and negative social/moral functions during adolescence. Positive peer relationships often are egalitarian and characterized by intimacy, self-disclosure, and mutual responsiveness. Positive peer relationships can provide opportunities for role taking, negotiation, and cooperation, as well as developing a more empathetic psychological understanding of peers (Kohlberg, 1964). However, peer relationships also can be problematic. For example, perceived pressure from peers can lead adolescents to make risky decisions or engage in problematic behaviors (e.g., alcohol and substance use, criminal activity). The perceived pressure to conform to peers generally rises in early adolescence, peaks at about age 14, and then declines through age 18 and beyond. Moreover, being rejected by one's peers may lead to subsequent unfavorable outcomes, such as biases in social information processing, which may result in aggressive responses and behaviors (Dodge et al., 2003).

Expanding social networks

Adolescents are exposed to larger social networks than are younger children, and these larger networks increasingly affect social/moral development. Adolescents' micro-level relationships (e.g., relationships with family members) evolve within these larger social networks (e.g., culture, society). The norms and expectations of other members of these larger communities (including adult role models) are more likely to be internalized by adolescents than by younger children. Research suggests that neighborhood contexts influence adolescents' attitudes toward risky behavior and deviance (e.g., violence, teenage sex, pregnancy, and drug use; Bingenheimer et al., 2005; Warner et al., 2011; Warner, 2016). In addition, adolescents today are among the most avid users of digital technology and social media. In 2018, 45% of American teenagers (13–17 years) were reportedly connected to the Internet "almost constantly," a nearly twofold increase of adolescents' near-constant online presence reported in 2014–15 (Anderson and Jiang, 2018). The uptake of social media can provide more opportunities for adolescents to communicate with others, establish relationships including romantic relationships, and civic engagement.

Summary

Both children and adolescents' beliefs, behaviors, and attitudes within the social and moral domains change over time. As youth develop more advanced cognitive capabilities (e.g., the ability to understand abstract concepts), they are better equipped to learn from peers and adults in their expanding social networks. Educators can support adolescents' social and cognitive development by providing opportunities for them to engage with peers and adults while pursuing prosocial goals. For example, educators might encourage adolescents to engage in community service initiatives; indeed, research suggests that engagement in community service supports positive development of adolescents' competence and attitudes in the academic, social, and civic domains (e.g., social responsibility, appreciation of diversity, social responsibility; Van Goethem et al., 2014). Lastly, the development of wireless communications technology has enabled adolescents to connect with others without geographical constraints, and form and maintain social relationships virtually, drastically changing the way they socialize and redefining the realm of social/moral development.

Identity development and adolescent learning

Identity can be understood as an individual's perception and understanding of themselves and how they believe they are perceived by others. Many aspects of identity develop during adolescence. In adolescence, youth define who they are and see changes in all aspects of identity. They may be trying to answer the questions: "Who am I? And who do I want to be?" They are attempting to define their biological, social, and psychological identities. Adolescents may feel pressured to conform or assimilate to a societal norm among peers and with adult influences. If they differ from this majority group identity, they may struggle in school and face adversity.

Identity development may be the most extended developing aspect of adolescence beyond cognitive and physiological development. Identity development can begin in late childhood, around age ten, and continue into emerging adulthood around age 23 or 24. However, even after adolescence, identity can continue to change and develop well into adulthood (Arnett, 2000).

Exploring and changing identities

Erik Erikson proposed that the only way to form a strong identity is through exploration and experimentation with various identities (Erikson, 1968, 1995). James Marcia expanded on these ideas, arguing that adolescents may take various pathways in developing their sense of self (Marcia, 1980). Marcia put forth a four-stage model of identity development. These stages are categorized by whether the individual is still in exploration and whether the individual feels a commitment to this identity. The four stages are (1) identity achievement, (2) identity moratorium, (3) identity foreclosure, and (4) identity diffusion.

In Marcia's model, *identity achievement* is the most advanced stage of identity development. In identity achievement, the individual has gone through exploration and has now committed to one identity path. Identity moratorium and foreclosure are

both intermediate stages. In *identity moratorium*, the individual is exploring many identities but has not committed to any. In *identity foreclosure*, the individual has committed to an identity, but did not go through exploration. Finally, *identity diffusion* is the least advanced stage. In identity diffusion, an individual is neither exploring nor committing to any paths.

Marcia's model allows for explanations of different ways adolescents develop identities. However, building on this framework, we must remember that identity is fluid and not unidimensional. As adolescents age, they become more aware of how different groups of people are similar or different from themselves, and how they may act differently across different contexts with different groups of people. Adolescents' identities are dynamic and ever-changing (Arnett, 2000).

Possible selves and self-concept

The fluidity of identity adds to how adolescents view themselves, and their perceptions of their future identities will shape their identity development. This idea of *possible selves* allows adolescents to consider who they are now and who they may be in the future (Markus and Nurius, 1986). These possible selves include their hoped-for self (i.e., the self they most hope will describe them in the future), their feared-self (i.e., the self they most fear will describe them in the future), and their actual self (i.e., the self they feel they are actually most likely to be).

These possible selves contribute to adolescents' identities and how they view themselves. Adolescents' *self-concept* (i.e., how they feel about themselves in different contexts) is influenced by who they see themselves as now and who they see themselves as in the future. In turn, adolescents' self-concept will influence their *self-efficacy* (i.e., their confidence in their own abilities with specific tasks), their *self-esteem* (i.e., their global feelings about themselves), and their *self-worth* (i.e., their global feelings about their own value; Baumeister, 2013; Bong and Clark, 1999).

Gender identity

Before discussing gender identity, it is important to understand the difference between gender and sex. Sex is typically assigned as male and female; it is determined by chromosomes and is assigned at birth (Helgeson, 2020). Gender is an individual's assessment of themselves, and it may or may not align with their sex. Gender identity often becomes particularly salient during adolescence (Guz et al., 2021). Adolescents' gender identity refers to how they view themselves and their gender; it is how they define their gender for themselves. Adolescents' gender roles refer to the expectations they hold for their gender and the societal roles they expect to fill. Adolescents' social interactions and influences in their everyday experiences will shape how they form and view their identities.

Individuals begin to understand gender during childhood but further define their own gender identity beginning during adolescence (Ferrer-Wreder et al., 2020). Bussey (2011) put forth a model of gender identity in which adolescents' gender identity is co-constructed through individual factors (i.e., personal understanding of gender, sex, and other processes such as self-efficacy), behavioral factors (i.e., individuals' performance of gender in day-to-day life and different contexts), and environmental factors (i.e., how others see an individual's gender socially, culturally, and historically).

All the factors suggested by Bussey (2011) are brought to the forefront in educational environments where adolescents interact with peers and teachers. Adolescents may feel pressure to perform to societal norms of "masculinity" or "femininity." However, adolescents may become social outcasts if they counter societal norms. Indeed, gender nonconforming youth face high levels of distress, mental health crisis, and bullying (Hatchel et al., 2021). In looking specifically at STEM fields, women and girls are less likely to study STEM subjects in school and avoid fields dominated by men (Dasgupta and Stout, 2014).

Racial and ethnic identity

Race and ethnicity are complex identities. In early adolescence, children may not have acquired the developmental capability to think about their racial or ethnic identities in complex ways. They may know about their race from their family, and they may even know their ethnicity. However, distinguishing between the two and developing their sense of identity is more likely to happen as they age and become more able to engage with complexity in later adolescence. This is especially true for adolescents in homogeneous environments where they do not encounter many people with races/ethnicities different from their own.

Racial identity is a more biological identity incorporating the physical characteristics of a group. *Ethnic identity* refers to the cultural, historical, and social identities of a group. Thus, an individual may be Asian (racial identity) and Japanese (ethnic identity). Applying Marcia's (1980) stages to racial and ethnic identity, an adolescent may explore the history, beliefs, and traditions of their ethnic group; then, commitment refers to how closely they feel their racial or ethnic group is to who they are.

Adolescents with a more positive ethnic identity who feel more closely tied to their ethnicity and culture often are better able to psychologically adjust in school (Umaña-Taylor et al., 2014). As adolescents begin to understand their race and ethnicity more, they will also begin to understand the prejudices and stereotypes held about their race. Black and Latinx students may not see the same benefits from a positive ethnic identity as their White counterparts due to stereotype threat. *Stereotype threat* is a psychological threat in which a student worries about acting in such a way that confirms a stereotype about a certain identity (e.g., the stereotype that Black students do not perform as well in school or that women do not do as well in STEM fields), which in turn often results in the students fulfilling this stereotype (e.g., scoring lower on tests or devaluing education/certain fields as a whole; Spencer et al., 2016). Of course, this is not the only factor at play that may negatively impact adolescents in school as they struggle with identity

development. Overall, adolescents who have not yet committed to their identity, or who are in identity crisis, will struggle more with academics (Jensen, 2010). Adolescents who feel more secure in their identities are better able to foster adaptive academic habits and positive academic self-concepts.

Summary

Adolescents' conceptions of who they are, and whom they might become in the future, are likely to change during adolescence. As students' self-perceptions change, their interests in various school subjects and the amount of time that they are willing to devote to academic pursuits also may change. Moreover, adolescent identity develops across a range of domains, including their present identities with their peers, the possible future identities that they are considering (e.g., in terms of a profession), and their identities with regard to personal characteristics (e.g., gender, racial, and ethnic identities). Educators who work with adolescents and are attuned to changes in adolescents' self-perceptions can support adolescent learning by allowing and gently encouraging adolescents to explore diverse identities.

Academic motivation development

Adolescence is often associated with a sharp decline in students' motivation. The media, in particular, often portrays adolescents as being bored and disenchanted with school. Research does in fact demonstrate that motivational beliefs change during adolescence; however, these shifts in motivation are not by any means universally detrimental (Wigfield and Wagner, 2005). Adolescents' motivational beliefs are largely determined by the social contexts in which they interact (e.g., Yeager et al., 2017).

In reviewing the research on adolescent motivation, it is important to distinguish between different types of motivational beliefs. In the present chapter, we focus on two specific frameworks that have been examined in longitudinal studies including adolescents. The first framework for examining academic motivation is referred to as Expectancy-Value Theory (and more recently as Situated Expectancy Value Theory [SEVT]; Eccles and Wigfield, 2020; Wigfield and Eccles, 2000). This theory posits that there are two overarching motivational beliefs: students' expectancies for success, and students' valuing of academic content.

Expectancies and values

Expectancy beliefs refer to students' beliefs about whether or not they can be successful in a given domain (e.g., math) or at a specific task (e.g., playing tennis). Students' valuing of academic tasks encompasses four related beliefs. These include (a) attainment value (i.e., the perceived importance of the subject, particularly as it pertains to the student's identity), (b) intrinsic value (i.e., how much the student likes/enjoys the subject), (c) utility value (i.e., the perceived usefulness of the subject area, particularly as it pertains to future goals), and (d) cost (i.e., considerations of what the student might have to give up doing in order to be engaged with the task/subject; Rosenzweig et al., 2019).

Numerous longitudinal studies have examined the ways that students' expectancies and values change during adolescence (see Wigfield et al., 2015, for a review). When children are in the early grades in school, they tend to have high expectancies – they often believe that they can accomplish whatever they set out to do, and that they can learn how to become good at just about anything. But as students move into higher grades, they begin to doubt their abilities in some areas. During early adolescence in particular, students' expectancy beliefs may decline in one or more areas. For example, a third grader might feel very confident at both reading and math, but by the time that third grader reaches the seventh grade, they might still feel confident as a reader, but less so as a mathematician.

Simultaneously, students' valuing of various subject areas also often declines as young children move into adolescence (Jacobs et al., 2002). A young child who deeply enjoyed doing science experiments as a second grader may not find science interesting anymore when that same child is in the ninth grade. Moreover, students' expectancies for success and valuing of academic content are positively related: students who maintain positive beliefs about their competencies in a domain (e.g., reading) are more likely to continue to value that domain, compared to students who do not maintain confidence in their abilities to succeed (Eccles and Wigfield, 2020).

Achievement goal theory

Achievement goal theory (AGT) focuses on the goals that students have when they engage with academic tasks. Although adolescents have many different goals, AGT focuses on two types of goals, each of which can be further dichotomized into an approach-type goal and an avoidance-type goal (Elliot and Harackiewicz, 1996). A mastery goal is the goal of deeply learning about a particular subject or task (Urdan and Kaplan, 2020). Students who hold mastery goals want to “master” the content. They are willing to take the necessary time to understand the topic in depth, they exert a great deal of effort, and they often see mistakes or misconceptions as opportunities to learn, rather than as embarrassing representations of their lack of skill or knowledge. Mastery goals are often dichotomized as mastery approach goals (i.e., the goal is to gain mastery of a particular subject or task), and mastery avoid goals (i.e., the goal is to avoid misunderstanding a topic/subject).

A performance goal is a goal that is focused on a student's perceived ability at completing a task or learning in a particular subject domain (Urdu and Kaplan, 2020). Students who hold performance goals are concerned with how their capabilities to engage with a task compared to other students. A student who is focused on demonstrating their ability in outperforming others holds performance-approach goals, whereas a student who is focused on avoiding appearing "dumb" or unable to complete a task holds performance-avoid goals.

Longitudinal studies also have examined the ways that achievement goals change during adolescence (e.g., Anderman and Midgley, 1997). Similar to expectancies and values, students' achievement goals often shift as students move into higher grades. Students often become more focused on performance (i.e., how they compare to others) and less focused on mastery (i.e., taking the time to learn deeply about a topic) as they progress into higher grades. In particular, adolescents tend to become more focused on performance (as well as on extrinsic outcomes) and less focused on mastery as they transition from elementary school into middle school (Anderman and Maehr, 1994).

Why do motivational beliefs change during adolescence?

Whereas stereotypes in the media often characterize adolescents as being lazy and often attribute declines in motivation to physiological development, current scholarship suggests that negative shifts in motivational beliefs are strongly attributable to changes in social contexts. Studies conducted both using SEVT and AGT indicate that the academic climates in many elementary school classrooms are focused on mastery, exploration, and individual progress, whereas the climates in many (but certainly not all) middle schools and high schools become more focused on grades, ability difference, and social comparison (e.g., Anderman and Midgley, 2004).

For example, when students transition from elementary school into middle school, they often encounter more grouping by ability than they experienced in elementary school. When students are placed into a seemingly low-achieving class in a given subject area (e.g., science), those students may begin to doubt their abilities to succeed in those subject domains, and consequently may come to value those subjects less. Relatedly, negative shifts in ability perceptions and values can potentially affect long-term career considerations (Anderman and Patrick, 2012). Indeed, a child who was considering a career in science might have a change of heart when that child is suddenly "told" by their middle teachers that they are not as good at science as many of their peers (because they were placed into a lower-achieving classroom). Studies from an AGT perspective also mirror these observations, with students who have recently transitioned into middle schools often reporting increases in performance goals and decreases in mastery goals (e.g., Anderman and Midgley, 1997).

Summary

Students' motivation toward academic pursuits changes during adolescence. Whereas the quality of motivation does decline for some youth, downward turns in motivation are not inevitable. Adolescent motivation is not a unidimensional construct. Adolescent motivation can be conceptualized in a variety of ways. Indeed, motivation during adolescence is a function of students' expectancies for success, their valuing of various academic (and non-academic) subjects, and their goals while engaged with academic tasks. Adolescent motivation is particularly sensitive to contextual influences. When schools emphasize grades and ability differences, adolescents may focus on their grades and performance, more so than on the intrinsic value of what they are learning.

Discussion

Adolescence is a period of dramatic change. Whereas the physical growth and maturation associated with adolescence is prominent, equally important are the changes that are occurring in psychological domains. Adolescents differ from younger children with regard to their cognitive capabilities, their comprehension and understanding of social and moral issues, their beliefs about who they are and who they hope to become, and their motivation toward academic work and school success (Anderman, 2012).

Adolescence is a broad period, often inclusive of a wide range of ages (e.g., sometimes youth as young as 10 or 11 are considered adolescents, and at times young adults in their early twenties are still referred to as adolescents). It is imperative for educators to carefully and critically consider the ongoing effects of all of these developmental shifts on learning (Eunice Kennedy Shriver National Institute of Child Health and Human Development NIH DHHS, 2007). Nevertheless, development in all of the psychological domains reviewed in this chapter transpires via varying trajectories. It is particularly important for teachers of adolescents to understand and acknowledge that their students will be at different stages of development both within and across these domains.

To illustrate, consider a language arts teacher who instructs students who are 14–15 years old. In the United States, these students would probably be in the 9th grade. Let us suppose that the class is reading a short story, and that one of the central characters in the story works at a fast food establishment, and is caught stealing some food at the end of their shift. The language arts teacher might decide to ask the class to discuss the incident, and to consider whether the character was justified in stealing the food. The teacher would probably hope that the students would consider other circumstances surrounding the incident that had been described in the story (e.g., peer pressure; food insecurity; the student's ability to understand how their actions might impact others; etc.).

A productive discussion around this topic would draw upon many of the topics addressed in this chapter. Students might respond differently to the scenario based on their own abilities to understand the complexities of the situation (i.e., cognitive domain), to consider the reasons as to why the food was stolen (i.e., social/moral domain), to empathize with the character and consider how this action might affect one's future goals (i.e., identity domain), and to consider the student's motives that precipitated the action (i.e., motivation domain). Nevertheless, adolescents who are 14–15 years old inevitably will be at different places with regard to their development in these four domains. Students are likely to analyze the character's actions (i.e., stealing) based on their own developmental levels. Consequently, some students are likely to disagree with each other with regard to their interpretations of the incident. Some students may empathize with the character; for example, some students might consider the character's home life and whether or not the character's family had enough money to buy food. Other students might focus on the potential detrimental effects that stealing could have on the character's future (e.g., if caught, the character would have a criminal record, which could affect subsequent educational and employment opportunities). These differing responses and interpretations are not "right" or "wrong." They all have merit, and reflect the different levels of development among these students.

This variability in development can be viewed by teachers as either an instructional problem to be overcome, or an opportunity to educate and promote development across these domains. Most current views of psychological development acknowledge the social nature of development, and embrace the notion that development can be enhanced when youth are exposed to ideas, beliefs, and opinions of others that differ from their personal beliefs and experiences. Consequently, a teacher who understands, acknowledges, and embraces the diversity in development among their students is likely to treat this diversity in development across the four domains as an asset, and to support their students in understanding and learning from each others' viewpoints.

Conclusion

In addition to physiological development, adolescents are developing across an array of psychological domains. In this chapter, we briefly described the major developmental shifts that occur across these domains. We noted in particular that changes in each of these domains do not occur simultaneously. A student may be able to consider and discuss complex, abstract issues (i.e., the student might be in the advanced stages of cognitive development), but that same student might have no idea about what they want to do in the future (i.e., in the early stages of identity development). Moreover, in addition to individual differences in development within each adolescent, educators also must acknowledge that there is likely to be vast variation in development between their students.

We can support adolescents' learning and achievement when we are aware of the diversity in the trajectories at which adolescents develop across these domains. When educators understand these issues and can truly support all of their students by embracing and promoting development across the four domains, adolescents will benefit in countless ways.

References

- Anderman, E.M., Maehr, M.L., 1994. Motivation and schooling in the middle grades. *Rev. Educ. Res.* 64 (2), 287–309.
- Anderman, E.M., Midgley, C., 1997. Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemp. Educ. Psychol.* 22 (3), 269–298.
- Anderman, E.M., Midgley, C., 2004. Changes in self-reported academic cheating across the transition from middle school to high school. *Contemp. Educ. Psychol.* 29 (4), 499–517.
- Anderman, E.M., Patrick, H., 2012. Achievement goal theory, conceptualization of ability/Intelligence, and classroom climate. In: Christenson, S., Reschly, A.L., Wylie, C. (Eds.), *The Handbook of Research on Student Engagement*. Springer Science, pp. 173–191.
- Anderman, E.M., 2012. Adolescence. In: Harris, K.R., Graham, S., Urdan, T. (Eds.), *APA Handbook of Educational Psychology, Volume 3: Applications to Learning and Teaching*. American Psychological Association, Washington D.C., pp. 43–61.
- Anderson, M., Jiang, J., 2018. Teens, Social Media & Technology 2018. Pew Research Center. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- Arnett, J.J., 2000. Emerging adulthood: a theory of development from the late teens through the twenties. *Am. Psychol.* 55 (5), 469–480. <https://doi.org/10.1037/0003-066x.55.5.469>.
- Barber, B.K., 1996. Parental psychological control: revisiting a neglected construct. *Child Dev.* 67, 3296–3319. <https://doi.org/10.1111/j.1467-8624.1996.tb01915.x>.
- Baumeister, R.F. (Ed.), 2013. *Self-Esteem: The Puzzle of Low Self-Regard*. Springer Science & Business Media.
- Bebko, J.M., Burke, L., Craven, J., Sarlo, N., 1992. The importance of motor activity in sensorimotor development: a perspective from children with physical handicaps. *Hum. Dev.* 35 (4), 226–240.
- Bingenheimer, J.B., Brennan, R.T., Earls, F.J., 2005. Firearm violence exposure and serious violent behavior. *Science* 308 (5726), 1323–1326. <https://doi.org/10.1126/science.1110096>.
- Bong, M., Clark, R.E., 1999. Comparison between self-concept and self-efficacy in academic motivation research. *Educ. Psychol.* 34 (3), 139–153.
- Bradley, R.H., Caldwell, B.M., Rock, S.L., Ramey, C.T., Barnard, K.E., Gray, C., Gottfried, A.W., Siegel, L., Johnson, D.L., 1989. Home environment and cognitive development in the first 3 years of life: a collaborative study involving six sites and three ethnic groups in North America. *Dev. Psychol.* 25 (2), 217.
- Brainerd, C.J., 2003. Jean Piaget, learning research, and American education. In: Zimmerman, B.J., Schunk, D.H. (Eds.), *Educational Psychology: A Century of Contributions*. Erlbaum, Mahwah, NJ, pp. 225–287.
- Bronfenbrenner, U., 1992. *Ecological Systems Theory*. Jessica Kingsley Publishers.
- Bussey, K., 2011. Gender identity development. In: *Handbook of Identity Theory and Research*, pp. 603–628. https://doi.org/10.1007/978-1-4419-7988-9_25.
- Dasgupta, N., Stout, J.G., 2014. Girls and women in science, technology, engineering, and mathematics. *Policy Insights Behav. Brain Sci.* 1 (1), 21–29. <https://doi.org/10.1177/2372732214549471>.
- Demetriou, A., Christou, C., Spanoudis, G., Platsidou, M., Fischer, K.W., Dawson, T.L., 2002. The Development of Mental Processing: Efficiency, Working Memory, and Thinking. *Monographs of The Society for Research in Child Development*, pp. i–167.

- Dodge, K.A., Bates, J.E., Pettit, G.S., 1990. Mechanisms in the cycle of violence. *Science* 250, 1678–1683. <https://doi.org/10.1126/science.2270481>.
- Dodge, K.A., Lansford, J.E., Burks, V.S., Bates, J.E., Pettit, G.S., Fontaine, R., Price, J.M., 2003. Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Dev.* 74 (2), 374–393. <https://doi.org/10.1111/1467-8624.7402004>.
- Eccles, J.S., Roeser, R.W., 2009. Schools, academic motivation, and stage-environment fit. In: Lerner, R.M., Steinberg, L. (Eds.), *Handbook of Adolescent Psychology: Individual Bases of Adolescent Development*, third ed. John Wiley & Sons, Inc, Hoboken, NJ, pp. 404–434.
- Eccles, J.S., Wigfield, A., 2020. From expectancy-value theory to situated expectancy-value theory: a developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* 61. <https://doi.org/10.1016/j.cedpsych.2020.101859>.
- Eisenberg, N., VanSchyndel, S.K., Hofer, C., 2015. The association of maternal socialization in childhood and adolescence with adult offspring's sympathy/caring. *Dev. Psychol.* 51, 7–16. <https://doi.org/10.1037/a0038137>.
- Elliot, A.J., Harackiewicz, J.M., 1996. Approach and avoidance achievement goals and intrinsic motivation: a mediational analysis. *J. Pers. Soc. Psychol.* 70 (3), 461.
- Erikson, E.H., 1968. *Identity: Youth and Crisis*.
- Erikson, E.H., 1995. *Childhood and Society*. Vintage.
- Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, DHHS, 2007. *Child and Adolescent Development Research and Teacher Education: Evidence-Based Pedagogy, Policy, and Practice* (NA). U.S. Government Printing Office, Washington, DC.
- Ferrer-Wreder, L., Kroger, J., Kroger, J., 2020. *Identity in Adolescence: The Balance Between Self and Other*. Routledge.
- Greenfield, P.M., Bruner, J.S., 1966. Culture and cognitive growth. *Int. J. Psychol.* 1 (2), 89–107.
- Guz, S., Kattari, S.K., Atteberry-Ash, B., Klemmer, C.L., Call, J., Kattari, L., 2021. Depression and suicide risk at the cross-section of sexual orientation and gender identity for youth. *J. Adolesc. Health* 68 (2), 317–323. <https://doi.org/10.1016/j.jadohealth.2020.06.008>.
- Hatchel, T., Polanin, J.R., Espelage, D.L., 2021. Suicidal thoughts and behaviors among LGBTQ youth: meta-analyses and a systematic review. *Arch. Suicide Res.* 25 (1), 1–37.
- Helgeson, V.S., 2020. *The Psychology of Gender*. Routledge, Taylor & Francis Group.
- Higgins, A.T., Tummure, J.E., 1984. Distractibility and concentration of attention in children's development. *Child Dev.* 1799–1810.
- Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S., Wigfield, A., 2002. Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73 (2), 509–527.
- Jensen, L.A., 2010. The cultural-developmental theory of moral psychology. In: *Bridging Cultural and Developmental Approaches to Psychology*, pp. 3–25. <https://doi.org/10.1093/acprof:oso/9780195383430.003.0001>.
- Kail, R.V., Ferrer, E., 2007. Processing speed in childhood and adolescence: longitudinal models for examining developmental change. *Child Dev.* 78 (6), 1760–1770.
- Kandel, D.B., 1978. Homophily, selection, and socialization in adolescent friendships. *Am. J. Sociol.* 84 (2), 427–436. <https://doi.org/10.1086/226792>.
- Keating, D.P., 2004. Cognitive and brain development. In: Lerner, R., Steinberg, L. (Eds.), *Handbook of Adolescent Psychology*, second ed. Wiley, New York.
- Keating, D.P., 2012. Cognitive and brain development in adolescence. *Enfance* (3), 267–279.
- Kohlberg, L., 1964. Development of moral character and moral ideology. In: Hoffman, L.W., Hoffman, M.L. (Eds.), *Review of Child Development Research*. Russel Sage Foundation, pp. 381–431.
- Kohlberg, L., 1973. The claim to moral adequacy of a highest stage of moral judgment. *J. Philos.* 70 (18), 630–646. <https://doi.org/10.2307/2025030>.
- Kramer, D.A., 1983. Post-formal operations? A need for further conceptualization. *Hum. Dev.* 26 (2), 91–105.
- Kuhn, D., 2008. Formal operations from a twenty-first century perspective. *Hum. Dev.* 51 (1), 48–55.
- Kuhn, D., 2009. Adolescent thinking. In: Lerner, R., Steinberg, L. (Eds.), *Handbook of Adolescent Psychology*, third ed. Wiley, New York.
- Lam, B.T., Ducreux, E., 2013. Parental influence and academic achievement among middle school students: parent perspective. *J. Hum. Behav. Soc. Environ.* 23 (5), 579–590. <https://doi.org/10.1080/10911359.2013.765823>.
- Marcia, J.E., 1980. Identity in adolescence. *Handb. Adolesc. Psychol.* 9 (11), 159–187.
- Markus, H., Nurius, P., 1986. Possible selves. *Am. Psychol.* 41 (9), 954–969. <https://doi.org/10.1037/0003-066x.41.9.954>.
- Newman, H.H., Freeman, F.N., Holzinger, K.J., 1937. *Twins: A Study of Heredity and Environment*. Univ. Chicago Press.
- Nucci, L., Turiel, E., 2009. Capturing the complexity of moral development and education. *Mind Brain Educ.* 3 (3), 151–159. <https://doi.org/10.1111/j.1751-228X.2009.01065.x>.
- Petrill, S.A., Lipton, P.A., Hewitt, J.K., Plomin, R., Cherny, S.S., Corley, R., DeFries, J.C., 2004. Genetic and environmental contributions to general cognitive ability through the first 16 years of life. *Dev. Psychol.* 40 (5), 805.
- Piaget, J., 1929. *The Child's Conception of the World*. Harcourt, Brace, New York, NY.
- Piaget, J., 1932. *The Moral Judgment of Children*. Routledge & Kegan-Paul.
- Piaget, J., 1952. In: Cook, M. (Ed.), *The Origins of Intelligence in Children*. Norton, New York, NY.
- Piaget, J., 1964. Part I: cognitive development in children: Piaget development and learning. *J. Res. Sci. Teach.* 2 (3), 176–186.
- Plomin, R., Asbury, K., 2005. Nature and nurture: genetic and environmental influences on behavior. *Ann. Am. Acad. Polit. Soc. Sci.* 600 (1), 86–98.
- Repetti, R.L., Taylor, S., Seeman, T.E., 2002. Risky families: family social environments and the mental and physical health of offspring. *Psychol. Bull.* 128 (2), 330–366. <https://doi.org/10.1037/0033-2909.128.2.330>.
- Rosenzweig, E.Q., Wigfield, A., Eccles, J.S., 2019. Expectancy-value theory and its relevance for student motivation and learning. In: Renninger, K.A., Hidi, S.E. (Eds.), *The Cambridge Handbook of Motivation and Learning*. Cambridge University Press, pp. 617–644. <https://doi.org/10.1017/9781316823279.026>.
- Siegler, R.S., 1988. Individual differences in strategy choices: good students, not-so-good students, and perfectionists. *Child Dev.* 833–851.
- Siegler, R.S., 1998. *Emerging Minds: The Process of Change in Children's Thinking*. Oxford University Press.
- Smetana, J.G., Turiel, E., 2003. Morality during adolescence. In: Adams, G.R., Berzonsky, M. (Eds.), *The Blackwell Handbook of Adolescence*. Blackwell, pp. 247–268.
- Spencer, S.J., Logel, C., Davies, P.G., 2016. Stereotype threat. *Annu. Rev. Psychol.* 67 (1), 415–437.
- Tisak, M., 1993. Preschool children's judgments of moral and personal events involving physical harm and property damage. *Merrill-Palmer Q.* 39, 375–390.
- Tucker-Drob, E.M., Briley, D.A., Harden, K.P., 2013. Genetic and environmental influences on cognition across development and context. *Curr. Dir. Psychol. Sci.* 22 (5), 349–355.
- Turiel, E., 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge University Press.
- Umaña-Taylor, A.J., Quintana, S.M., Lee, R.M., Cross, W.E., Rivas-Drake, D., Schwartz, S.J., Syed, M., Yip, T., Seaton, E., 2014. Ethnic and racial identity during adolescence and into young adulthood: an integrated conceptualization. *Child Dev.* 85 (1), 21–39. <https://doi.org/10.1111/cdev.12196>.
- Urdan, T., Kaplan, A., 2020. The origins, evolution, and future directions of achievement goal theory. *Contemp. Educ. Psychol.* 61. <https://doi.org/10.1016/j.cedpsych.2020.101862>.
- Van Goethem, A., Van Hoof, A., Orbio de Castro, B., Van Aken, M., Hart, D., 2014. The role of reflection in the effects of community service on adolescent development: a meta-analysis. *Child Dev.* 85 (6), 2114–2130. <https://doi.org/10.1111/cdev.12274>.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Vygotsky, L.S., 1987. In: Rieber, R.W., Carton, A.S. (Eds.), *The Collected Works of L.S. Vygotsky*, vol. 3. Plenum Press, New York, NY.
- Warner, T.D., Giordano, P.C., Manning, W.D., Longmore, M.A., 2011. Everybody's doin' it (right?): neighborhood norms and sexual activity in adolescence. *Soc. Sci. Res.* 40 (6), 1676–1690. <https://doi.org/10.1016/j.ssresearch.2011.06.009>.
- Warner, T.D., 2016. Up in smoke: neighborhood contexts of marijuana use from adolescence through young adulthood. *J. Youth Adolesc.* 45 (1), 35–53. <https://doi.org/10.1007/s10964-015-0370-5>.
- Wigfield, A., Eccles, J.S., 2000. Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25 (1), 68–81.

- Wigfield, A., Wagner, A.L., 2005. Competence, motivation, and identity development during adolescence. In: Elliot, A.J., Dweck, C.S. (Eds.), *Handbook of Competence and Motivation*. Guilford Publications, pp. 222–239.
- Wigfield, A., Eccles, J.S., Fredricks, J.A., Simpkins, S., Roeser, R.W., Schiefele, U., 2015. Development of achievement motivation and engagement. In: Lamb, M.E., Lerner, R.M. (Eds.), *Handbook of Child Psychology and Developmental Science: Socioemotional Processes*, , seventh ed.vol. 3. John Wiley & Sons, Inc, pp. 657–700. <https://doi.org/10.1002/9781118963418.childpsy31>.
- Yeager, D.S., Lee, H.Y., Dahl, R.E., 2017. Competence and motivation during adolescence. In: Elliot, A.J., Dweck, C.S., Yeager, D.S. (Eds.), *Handbook of Competence and Motivation: Theory and Application*, second ed. The Guilford Press, pp. 431–448.

Key developments in early and middle adulthood

Carol Hoare, Department of Counseling and Human Development, The George Washington University, Washington, DC, United States

© 2023 Elsevier Ltd. All rights reserved.

Young adulthood	497
Emerging adulthood	497
Young adulthood	497
Personality stability versus development	498
Identity and work commitment	498
Sustained intimacy	498
Middle adulthood	498
Midlife: a bridge period of positives and negatives	499
Generativity	500
Cognitive developments in midlife	500
The myths of middle age: the mid-life crisis, empty nest stress, and menopausal difficulties	501
References	501

Young adulthood

In recent decades, the path from adolescence to mature adulthood has lengthened. Increasing numbers of high school graduates attending college and graduate school have postponed career entry, marriage, and parenthood until the mid to late twenties or early thirties. Settling into mature roles and responsibilities has thus been delayed. These changes led [Arnett \(2000, 2015\)](#) to conceptualize the period of 18–25 years as emerging adulthood. This leaves the years of 25–40 as young adulthood proper, even though [Mehta et al. \(2020\)](#) recently conceived of the age range of 30–45 years as “established adulthood.” We consider these two divisions of young adulthood separately.

Emerging adulthood

Developmentally, the period of emerging adulthood is qualitatively unique, with its subjects engaging in more extensive identity excursions. Thus, emerging adults tend to remain self-focused as they delay commitments to careers and permanent responsibilities. [Arnett \(2000\)](#) notes that 40% of such adults move back into and then eventually leave their parents’ homes. Along with the approximately 60% of youngsters now attending college, of those who complete their baccalaureate education, about 33% then pursue graduate degrees. In the age range of 18–25 years, most subjects in Arnett’s study claimed that they were somewhere between adolescence and adulthood ([Arnett, 2001](#)). Developmentally, emerging adults believe that autonomy in decision making, financial independence, and the acceptance of personal responsibility for the self are paramount. As Arnett claims, these characteristics of self-reliance are assumed prior to making serious work commitments or assuming responsibility for others through enduring intimacy with a partner and parenthood (see also [Nelson, 2005](#)).

Extending [Erikson’s \(1968\)](#) depiction of identity development during adolescence, emerging adulthood is now the scene of identity exploration and development. In their quest for a vocational commitment, young adults in this period remain unstable while they seek to determine the nexus of their interests, talents, and values. Marking their exploration, job changes in the age range of 18–29 years average eight in number ([U.S. Department of Labor, 2012](#)). In recent decades, professional opportunities for young women now extend beyond the options of teaching or nursing, the dominant choices available during the middle of the twentieth century. Thus, for both genders in contemporary society, full identity development is deferred until professional development is complete.

[Arnett \(2000\)](#) has noted that the epoch of emerging adulthood applies best to societies that delay the beginning of mature adult roles. For example, the median age of marriage for women in industrialized societies is 26 while for women in developing countries it is 21 ([Arnett, 2000](#)). Parenthood in industrialized countries is often delayed until 30 years of age while it remains far earlier in less developed nations ([Arnett et al., 2014](#)).

Young adulthood

Once into young adulthood proper, several developmental markers appear. Self-concept and personality become more stable, commitments to a career or work role are realized, sequential intimacies mature into deep intimacy with a loved partner, and parenthood begins. Thus, young adulthood sees the culmination and stabilization of various forms of development that were previously in flux. We consider the key developmental attributes of personality, work engagement, and sustained intimacy with a partner.

Personality stability versus development

In recent decades, some personality theorists have held that just as the self-concept stabilizes during young adulthood, personality is fixed, typically, by age 30 (McCrae and Costa, 2008). This “plaster” position is countered by the premise of personality “plasticity.” This latter position accepts that certain aspects of genetics shape personality to some extent; however, adaptation to changing contexts alters personality over time. Indeed, changing roles and events (e.g., marriage, parenthood) have been shown to result in personality change. Denissen et al. (2019) demonstrated greater emotional stability, conscientiousness, and agreeableness during young adulthood. Others have noted that work requirements increase conscientiousness (Wagner et al., 2019).

Sociocultural changes have also been found to alter personality characteristics. For example, Roberts and Helson (1997) studied women over a span of 31 years and found that individualism and autonomy increased, particularly between 1960 and 1985. Autonomy increased as did national trends of heightened individualism and expectations of equality in work roles. A reciprocal association appeared in that assertiveness and individualism emerged from the feminist movement which then fostered increased autonomy in work environments.

It is important to note that the preponderance of personality development research employs the Five Factor Model (FFM) of McCrae and Costa (2008), specifically the attributes of extraversion, conscientiousness, openness to experience, agreeableness, and neuroticism. However, Srivastava et al. (2003) have claimed that the FFM is too “tidy” to comprise the entire range of personality attributes. In fact, a meta-analysis by Ackerman and Heggestad (1997) found that 19 personality characteristics were examined in 135 studies. Furthermore, in their content analysis of key personality journals, Cribb et al. (1992) discovered that more than 500 attributes were in evidence.

Identity and work commitment

Erikson (1950, 1968) established that identity is a lifelong process that is never fully complete. However, the beginning consolidation of identity occurs in young adulthood, largely due to commitments to the values, talents, and interests that coalesce into a work commitment. As Arnett (2015) noted, work commitments that previously took place in a person’s early twenties now occur in the thirties.

Identity development and evolution include learning about the self in work functions, roles, and contexts, just as one learns about the job itself. Meaningful work validates the self, connotes worth, and provides inclusion. Furthermore, some aspects of development and learning occur unconsciously. At best, adults who are immersed in work functions and the identity that these serve, tend to lose consciousness about themselves as they contribute effort in work roles. Thus engaged, persons exist in a time-out-of mind zone. As Erikson (1968) quoted from a letter of William James to his wife:

A man’s character (identity) is discernible in the mental or moral attitude in which, when it came upon him, he felt himself most deeply and intensely active and alive. At such moments there is a voice inside which speaks and says: “This is the real me!” (p. 19).

Sustained intimacy

As Alperin (2001) has held, the human’s desire for intimacy is innate. Intimacy sustains the self and exists in many relationships (e.g., relationships with a partner, with parents, and with friends). Young adults who develop close relationships with others are shielded from developing a sense of isolation or uncertainty about their place in society (see, e.g., Tanner and Arnett, 2016). The expression of intimacy in a sustained partnership with a loved other is the profound intimacy that begins in the years of young adulthood. According to Erikson (1968), beyond the mere sexual experiences that may precede actual genital maturity, intimacy depends, in part, on an established identity. This allows for unification with another and a deep commitment to the other. Together, these result in a blend of mutual commitment and a mutually fulfilling relationship despite the concessions and self-denial required. Such “mutual devotion” develops during young adulthood because psychosocial and physical readiness, together with acceptance of societal norms, prepare the way for a committed life together (Erikson, 1968, p. 137). In the best of circumstances, this supports species continuity through parenthood.

In young adults who are incapable of developing sustained intimacy with a loved partner, a “fear of fusion” and of engulfment by another inhibit the development of intimacy (Alperin, 2001, p. 137). In others, fear of abandonment by the loved other may preclude intimacy. In such circumstances, self-absorption with isolation is the prevailing resolution (Erikson, 1980).

Middle adulthood

At the outset, four points about middle adulthood bear emphasizing. First, it is important to note that, for 2019, the US Census Bureau estimated that 25% of the population was in the 40–59-year age group, roughly considered middle adulthood. At the far end of this spectrum, however, the term middle age is a misnomer. No one can reasonably expect to live until 118 years of age.

Second, there is no universal set of experiences that typify adult development during the midlife period. Because persons become more heterogeneous as they age, it is far easier to say what a group of 20 year-old emerging adults are like than it is to stipulate patterns of developmental change among a set of middle aged adults. Individuals interact with varying experiences and characteristics that are altered by, for example, socioeconomic status, health manifestations, culture, ethnicity, sexual orientation, diverse

work experiences, and unique relationships (see, e.g., Lachman, 2004). The only common experiences are women's menopause and, in both genders at midlife, a growing awareness of life's approaching end.

Third, middle adulthood is the least studied period of life (Lachman et al., 2015). This is partly because it is far easier to study children, young adults, and the elderly than to study middle aged persons who are actively engaged in their work and personal lives and are not seen in institutional environments such as schools or elderly care settings. Furthermore, Freud influenced students of psychological development to see the adult period of life as a barren stretch of time between genital maturity and death in which adults were expected to merely love and to work. To Erikson (1977), in Freud's thought, all one could expect of adults was the comparative absence of childhood "fixations," "guilt," "drives," and "repressions." The situation was compounded by G. Stanley Hall who, with his 1907 volume *"adolescence"* and his 1922 volume *"senescence,"* implied that there was little worthy of mention in the middle years. In his book titles alone, Hall excluded all of adulthood between adolescence and the deterioration of old age. Hall (1922) even held that "senescence starts in the early forties, or before in women ..." (p. vii).

It is also the case that only since the middle of the twentieth century have increases in longevity made it possible to conceive of a middle period of life prior to old age. For example, average US life expectancy was only 39 years in 1900. Only since the beginning of the current century has life expectancy exceeded 70 years of age.

However, even by the middle of the twentieth century when life expectancy had increased substantially, investigations into middle adulthood were rare. Looking to Psychological Abstracts, only in 1973 did the Abstracts first use the subject heading "middle aged." That was also the 1st year in which the Abstracts distinguished between young and middle adulthood as distinct periods of adult life.

Even when some notions of adult development were entertained, for a long period of time ideas about the adult were characterized by a product-like notion in which adults were conceived of as merely physically complete children. Development, when it was considered, charted mere chronology and event dependence. Collected years and marker occasions such as marriage, childbirth, and retirement were substitutes for developmental change.

Fourth, the term *development* is misleading for it implies only positive change. Departing from this, the psychological literature holds that human development is bi-directional throughout life. That is, growth and decline are seen across multiple domains (e.g., Lachman, 2004).

Recently, several large-scale studies have begun to explore this previously neglected period of life. For example, the Midlife in the United States (MIDUS) survey in 1995–1996 and 2011–2016 studied more than 7000 adults. It focused on psychosocial and physical health, and the ways in which middle aged persons work through midlife concerns (see Brim et al., 2004) and the Survey of Health Aging and Retirement in Europe (SHARE) randomly selected 27,500 adults over the age of 50. Its concern was with relationships between adult children and their parents (Dykstra and Fokkema, 2011). Surveys such as Health and Retirement Study (HRS), the English Longitudinal Study of Aging (ELSA), the German Socioeconomic Panel Study (GSOEP), Household Income and Labor Dynamics (HILDA), and the Midlife in Japan (MIDJA) study have also included middle aged adults (Infurna et al., 2020).

Midlife: a bridge period of positives and negatives

In many respects, midlife stands as a connecting period of life. It bridges early adulthood with old age, is a link between positive development and eventual decline, and connects the generations. It is also a period of heightened responsibilities and control, typically one in which earnings crest, control is exerted in the work realm, and self-confidence is high (Lachman et al., 2015).

Arnett (2018) described this period of life as one in which adults are "happily stressed." In his study of 834 US middle agers, most reported a peak in well-being compared with earlier adulthood. The top sources of pleasure were those of relationships with adult children (88%), hobbies or leisure pursuits (86%), and travel or holidays (82%). The primary sources of stress were financial issues (62%), work concerns (51%), and personal health (42%). Based on data from the Midlife in the US (MIDUS) survey, Lachman et al. (2015), study participants reported similar experiences. They related more positive than negative emotions, heightened autonomy, good relationships, and satisfaction with life. Gains in control were held as positive, knowledge and experience peaked, and, in general, difficulties were largely offset by growth and mastery.

While middle age is largely reported as a period of significant life satisfaction, anxiety and depression peak during this epoch (Infurna et al., 2020). This is especially the case for women, for those who claim Lesbian, Gay, Bisexual, and Transgender status (Brody et al., 2018), and for those of lower socioeconomic position (Infurna et al., 2020). Unsurprisingly, middle aged adults of lower socioeconomic position report more difficulties such as unemployment and financial difficulties. Low socioeconomic status has been associated with heightened psychological and medical problems and with more disabilities (Fingerman et al., 2018; Conger and Donnellan, 2007; Reiss et al., 2019). Many of those who had experienced childhood adversity (e.g., poverty, abuse, sub-par parenting, parental death) also show difficulties in middle age (Lachman et al., 2015). Such childhood difficulties are correlated with diminished mental and physical health at midlife as well as cognitive and interpersonal problems (Conger and Donnellan, 2007; Infurna et al., 2020; Reiss et al., 2019).

In recent decades, emotional and monetary support of adult children and aging parents has led middle agers to experience stress. Middle-aged adults are caught in the middle. Well-being is adversely affected by the multiple responsibilities entailed, particularly when aging parents or adult children of middle agers experience financial, emotional, and physical problems (Fingerman et al., 2018). Infurna et al. (2020) have termed the phenomenon "double- and triple-duty caregiving" (p. 474). In 2018, Mason held that 40 million US caregivers provide such care and support. This estimate may be low, in that the National Alliance for Caregiving has estimated that 65.7 million Americans currently serve as caregivers for disabled or ill family members.

Generativity

In his second edition of *Childhood and Society*, Erikson (1963) first wrote about the polarity of generativity versus stagnation, self-absorption, reactivity, and narcissism. Erikson held that generative adults focus their efforts on concern for others and for the products they create in personal and work domains. Generativity includes “procreativity, productivity, and creativity,” actions directed toward fostering the positive development of those who will follow (Erikson, 1982, p. 67). In caring for children and grandchildren, mentoring others, and engaging in civic commitments, middle aged adults sublimate their own needs in service to others, thus showing faith in species continuity. By “cultivating strength in the next generation,” they create a legacy that will survive them (Erikson, 1982, p. 67).

In the nearly 60 years since Erikson first coined the term, he and others have depicted generativity as a “universal human need and strength” (Erikson, 1969, p. 395). Erikson’s (1969) volume *Gandhi’s*.

Truth described the origins of militant nonviolence and generative concern for others in the life and work of Mahatma Gandhi. Since then, research has demonstrated that generativity is associated with a variety of positive characteristics and is a hallmark of positive psychosocial development at midlife.

Longitudinal studies have shown that generativity is correlated with positive psychosocial adaptation and with the use of developmentally mature coping mechanisms. Studies have also found an association between generativity and self-esteem (Blatny et al., 2019), with life satisfaction and subjective well-being, and with self-efficacy. Generativity is negatively correlated with neuroticism (Newton et al., 2014).

Cognitive developments in midlife

As for other characteristics, it is once again important to emphasize that there is substantial variability in cognitive changes among middle aged adults. That said, most investigators have found that cognitive ability is stable until 60 years of age (Diehl et al., 2020; Schaie, 2013). This includes vocabulary, accumulated knowledge, and semantic memory (Irwin et al., 2018; Salthouse, 2009). The exceptions are speed of processing and performance, reaction time, and numerical ability, all of which show some decline during the middle years of life, largely in the sixth decade (Diehl et al., 2020; Irwin et al., 2018). Greater distractibility also tends to manifest in some during the middle years.

Some studies have found that cognition is either stable or enhanced during the middle years. For example, cognitive functions that depend on experience often improve (Lachman, 2004). Furthermore, although some middle-aged adults express concerns about memory declines, research findings dispute this tendency (Lachman, 2004). In fact, any cognitive declines that occur are gradual and are offset by improvements in reasoning and vocabulary (Willis and Schaie, 1999).

Many experts associate high socioeconomic circumstances in childhood, high cognitive ability in childhood, and early education as protective against cognitive decline during mid-life (Brayne, 2007; Richards et al., 2004; Wang et al., 2019). Importantly, in 50% of middle aged and elderly adults cognitive ability has been portrayed as being due to childhood intellectual ability (Gow et al., 2011). More extensive education in youth is particularly important in the protection against mid-life cognitive deterioration.

Studies have found that engaging in work positions of high complexity (with data and people) maintains intellectual capacity (Staudinger, 2020; Wang et al., 2019; Singh-Manoux et al., 2005). In addition, cognitively demanding leisure activities (e.g., doing crossword and other word puzzles, playing chess, playing musical instruments) and ongoing learning foster better cognitive performance throughout middle and older adulthood (Singh-Manoux et al., 2005; Wang et al., 2019). Engaging in challenging tasks and experiencing novelty also maintain vitality and cognitive maintenance (Staudinger, 2020). As well, openness to experience, a key personality characteristic, buffers against intellectual decline (Staudinger, 2020). It is also the case that those with more education and continual learning fare better cognitively than those with limited education who may also restrict their learning (Diehl et al., 2020; Hughes et al., 2018).

Middle aged adults of higher socioeconomic position, due in part to their expansive social networks, and who also engage in complex jobs, show limited, if any, cognitive decline (Singh-Manoux et al., 2005; Wang et al., 2019). Of course, it may also be true that middle aged persons whose cognitive performance is high engage in activities that are more intellectually demanding. In general, among those who actively engage in stimulating activities, semantic memory (i.e., general knowledge, vocabulary, and word fluency) tends to decline far later in life (Irwin et al., 2018). Increases in semantic memory among such subjects typically increase until 60 years of age. Furthermore, although episodic memory (i.e., autobiographical material, statement recall, and the recollection of personally important events) tends to decline earlier than semantic memory, both forms of memory are preserved longer in persons who are cognitively able and who exhibit greater personal control in their lives (Robinson and Lachman, 2018; Tanner and Arnett, 2016).

Physical exercise also fosters the maintenance of intellectual abilities. Experts have suggested that physical fitness may lead to neurochemical changes that buffer against cognitive decline (Hultsch et al., 1999; Lachman et al., 2018). It is now known that neurogenesis (new neuron development), angiogenesis (new blood vessel development), and synaptogenesis (new synapse development) occur throughout life. It is also the case that aerobic exercise fosters these processes (Lista and Sorrentino, 2009).

Experts have also posed that, in some, cognitive decline is the result of disease processes and sedentary lifestyles. Hypertension, diabetes, high cholesterol, and obesity have all been implicated in intellectual deterioration (Singh-Manoux et al., 2012). Thus, some factors that persons can influence may result in cognitive maintenance or decline.

The myths of middle age: the mid-life crisis, empty nest stress, and menopausal difficulties

Jacques (1965) coined the term “midlife crisis” to depict what he observed in his analysis of the creations of artists and in patients’ changed productivity and their awareness of life’s finiteness. A classic portrayal of such a crisis is that of a middle-aged man who buys a sports car, divorces his wife, and has Botox injections to look younger. It appears that an actual crisis is largely a myth, in that studies have found that only 10%–20% of people report experiencing it (Infurna et al., 2020; Lachman, 2004; Wethington et al., 2004). It may be that the term “crisis” exaggerates something that occurs, but in a less traumatic manner. Some time ago, Neugarten (1968) held that middle-aged persons experience a change in their time perspective, moving from an assessment of “life since birth” to “time remaining to be lived.” The future is no longer limitless (Freund and Ritter, 2009). Earlier goals may have been reached and the midlife adult asks: “Is that all there is?” Or, goals now reached, the adult questions those goals and wishes he or she had pursued other dreams.

Another myth is that of stressful experiences due to the empty nest, of adaptations that must be made because children have left home. Presumably, mothers are affected more than fathers. It appears that many women consider this an opportune time to initiate new interests and to forge renewed bonds with their spouses, thus turning a potential trauma into advantages (Mitchell, 2006).

A third myth held sway for decades, the belief that menopause caused psychological difficulties due to the loss of fertility and because the phenomenon was but a preamble to death (Rossi, 2004). Currently, menopause is considered as a “benign event,” one that is not associated with eventual death (Rossi, 2004, p. 154). In fact, after menopause most women today still have one-third of their lives ahead of them. Some women have recently reported relief when menopause occurs, possibly because the childbearing years are now behind them. Furthermore, altered menopausal status has not been found to predict stress (Hedgeman et al., 2018). Wilk and Kirk (1995) have held that the menopause should be considered a developmental process and not a deficiency.

References

- Ackerman, P.L., Heggestad, E.D., 1997. Intelligence, personality, and interests: evidence for overlapping traits. *Psychol. Bull.* 121, 219–245.
- Alperin, R.M., 2001. Barriers to intimacy. *Psychoanal. Psychol.* 18, 137–156.
- Arnett, J.J., Zukauskienė, R., Kazuma, S., 2014. The new life stage of emerging adulthood at ages 18–29 years: implications for mental health. *Lancet* 1, 569–576.
- Arnett, J.J., 2000. Emerging adulthood. *Am. Psychol.* 55, 469–480.
- Arnett, J.J., 2001. Conceptions of the transition to adulthood: perspectives from adolescence to midlife. *J. Adult Dev.* 8, 133–143.
- Arnett, J.J., 2015. *Emerging Adulthood: The Winding Road From the Late Teens Through the Twenties*, second ed. Oxford University Press, New York.
- Arnett, J.J., 2018. Happily stressed: the complexity of well-being in midlife. *J. Adult Dev.* 25, 270–278.
- Blatny, M., Milova, K., Martin, J., Romanakova, M., 2019. Personality predictors of midlife generativity: a longitudinal study. *J. Adult Dev.* 26, 213–231.
- Brayne, C., 2007. The elephant in the room—healthy brains in later life, epidemiology and public health. *Neuroscience* 8, 233–239.
- Brim, O.G., Ryff, C.D., Kessler, R.C. (Eds.), 2004. *How Healthy are We? A National Study of Well-Being at Midlife*. The University of Chicago Press, Chicago, IL.
- Brody, D.J., Pratt, L.A., Hughes, J., 2018. Prevalence of Depression Among Adults Aged 20 and Over: United States, 2013–2016. NCHS Data Brief no.303. National Center for Health Statistics, Hyattsville, MD.
- Conger, R.D., Donnellan, M.B., 2007. An interactionist perspective on the socioeconomic context of human development. *Annu. Rev. Psychol.* 58, 175–199.
- Cribb, D.A., Ozone, S.J., Pipes, R.B., 1992. A compilation of personality instruments. In: Paper Presented at the Annual Meeting of the American Psychological Association.
- Dennissen, J.J.A., Chung, J.M., Luhmann, M., Bleidorn, W., 2019. Transactions between life events and personality traits across the adult lifespan. *J. Pers. Soc. Psychol.* 116, 612–633.
- Diehl, M., Smyer, M.A., Mehrotra, C.M., 2020. Optimizing aging: a call for a new narrative. *Am. Psychol.* 75, 577–589.
- Dykstra, P.A., Fokkema, T., 2011. Relationships between parents and their adult children: a West European typology of late-life families. *Ageing Soc.* 31, 545–569.
- Erikson, E.H., 1950. *Childhood and Society*. Norton, New York.
- Erikson, E.H., 1963. *Childhood and Society*, second ed. Norton, New York.
- Erikson, E.H., 1968. *Identity: Youth and Crisis*. Norton, New York.
- Erikson, E.H., 1969. *Gandhi’s Truth*. Norton, New York.
- Erikson, E.H., 1977. Erikson on Erikson. *Parents Magazine Films*, Chicago, IL.
- Erikson, E.H., 1980. *Identity and the Life Cycle*. Norton, New York.
- Erikson, E.H., 1982. *The Life Cycle Completed*. Norton, New York.
- Fingerman, K.L., Huo, M., Graham, J.L., Kim, K., Birditt, K.S., 2018. A family affair: family typologies of problems and midlife well-being. *Gerontologist* 58, 1054–1064.
- Freund, A.M., Ritter, J.O., 2009. Midlife crisis: a debate. *Gerontology* 55, 582–591.
- Gow, A.J., Johnson, W., Pattie, A., Brett, C.E., Roberts, B., Starr, J.M., Deary, I.J., 2011. Stability and change in intelligence from age 11 to ages, 70, 79, and 87: the Lothian birth cohorts of 1921 and 1936. *Psychol. Aging* 26, 232–240.
- Hall, G.S., 1922. *Senescence: The Last Half of Life*. Appleton, New York.
- Hedgeman, E., Hasson, R.E., Karvonen-Gutierrez, C.A., Herman, W.H., Harlow, S.D., 2018. Perceived stress across the midlife: longitudinal changes among a diverse sample of women, the Study of Women’s Health Across the Nation (SWAN). *Women’s Midlife Health* 4, 1–11.
- Hughes, M.L., Agrigoroaei, S., Jeon, M., Lachman, M., 2018. Change in cognitive performance from midlife into old age: findings from the Midlife in the United States (MIDUS) study. *J. Int. Neuropsychol. Soc.* 24, 805–820.
- Hultsch, D.F., Hertzog, C., Small, B.J., Dixon, R.A., 1999. Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging? *Psychol. Aging* 14, 245–263.
- Infurna, F.J., Gerstorf, D., Lachman, M.E., 2020. Midlife in the 2020s: opportunities and challenges. *Am. Psychol.* 75, 470–485.
- Invin, K., Sexton, C., Daniel, T., Lawlor, B., Naci, 2018. Healthy aging and dementia: two roads diverging in midlife? *Front. Aging Neurosci.* 10, 1–12.
- Jacques, E., 1965. Death and the mid-life crisis. *Int. J. Psychoanal.* 46, 502–514.
- Lachman, M.E., Teshale, S., Agrigoroaei, S.W., 2015. Midlife as a pivotal period in the life course: balancing growth and decline at the crossroads of youth and old age. *IJBD* 39, 20–31.
- Lachman, M.E., Lipsitz, L., Lubben, J., Castaneda-Scept, C., Jette, A.M.J., 2018. When adults don’t exercise: behavioral strategies to increase physical activity in sedentary middle-aged and older adults. *Innov. Aging* 2, 1–12.
- Lachman, M.E., 2004. Development in midlife. *Annu. Rev. Psychol.* 55, 305–331.
- Lista, I., Sorrentino, G., 2009. Biological mechanisms of physical activity in preventing cognitive decline. *Cell. Mol. Neurobiol.* 30, 493–503.

- Mason, D., 2018. The many faces of family caregivers. *Innov. Aging* 2, 239–240.
- McCrae, R.R., Costa, P.P., 2008. The five-factor theory of personality. In: John, O.P., Robins, R.W., Pervin, L.A. (Eds.), *Handbook of Personality: Theory and Research*, 3rd. ed. Guilford Press, New York, pp. 159–181.
- Mehta, C.J., Arnett, J.J., Palmer, C.G., Nelson, L.J., 2020. Established adulthood: a new conception of ages 30 to 45. *Am. Psychol.* 75, 431–444.
- Mitchell, B.A., 2006. *The Boomerang Age: Transitions to Adulthood in Families*. Aldine, New Brunswick, NJ.
- Nelson, L.J., 2005. Distinguishing features of emerging adulthood: the role of self-classification as an adult. *J. Adolesc. Res.* 20, 242–262.
- Neugarten, B.L. (Ed.), 1968. *Middle Age and Aging: A Reader in Social Psychology*. University of Chicago Press, Chicago, IL.
- Newton, J.J., Herr, J.M., Pollack, J.I., McAdams, D.P., 2014. Selfish or selfless? Generativity and narcissism as components of legacy. *J. Adult Dev.* 21, 59–68.
- Reiss, D., Nielsen, L., Godfrey, Y.K., McEwen, B., Power, C., Sseman, S., 2019. Midlife reversibility of early-established risk factors: a research agenda. *Dev. Psychol.* 55, 2203–2218.
- Richards, M., Shipley, B., Fuhrer, R., Wadsworth, M.E.J., 2004. Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study. *Br. Med. J.* 328, 552–554.
- Roberts, B.W., Helson, R., 1997. Changes in culture, changes in personality: the influence of individualism in a longitudinal study of women. *J. Pers. Soc. Psychol.* 72, 641–651.
- Robinson, S.A., Lachman, M.E., 2018. Perceived control and cognition in adulthood: the mediating role of physical activity. *Personality Aging* 33, 769–781.
- Rossi, A.S., 2004. The menopause transition and aging processes. In: Brim, O.G., Ryff, C.D., Kessler, R.C. (Eds.), *How Healthy are We? A National Study of Well-Being at Midlife*. The University of Chicago Press, Chicago, IL, pp. 153–201.
- Salthouse, T.A., 2009. When does age-related cognitive decline begin? *Neurobiol. Aging* 30, 507–514.
- Schaie, K.W., 2013. *Developmental Influences on Adult Intelligence: The Seattle Longitudinal Study*, second ed. Oxford University Press, New York.
- Singh-Manoux, A., Richards, M., Marmot, M., 2005. Socioeconomic position across the lifecourse: how does it relate to cognitive function in mid-life? *Assoc. Educ. Psychol.* 15, 572–578.
- Singh-Manoux, S., Kivimäki, M., Glymour, A.M., Elbaz, A., Berr, C., Ebmeier, K.P., Ferrie, J.E., Dugravot, A., 2012. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *Br. Med. J.* 10, 1–8.
- Srivastava, S., John, O.P., Gosling, S.D., Potter, J., 2003. Development of personality in early and middle adulthood: set like plaster or persistent change? *J. Pers. Soc. Psychol.* 5, 1041–1053.
- Staudinger, U.M., 2020. The positive plasticity of adult development: potential for the 21st century. *Am. Psychol.* 75, 540–553.
- Tanner, J.L., Arnett, J.J., 2016. Emerging adult clinical psychology. In: Norcross, J.C., VandenBos, G.R., Freedheim, D.K. (Eds.), *APA Handbook of Clinical Psychology: Roots and Branches*, vol. 1. American Psychological Association, Washington, D.C., pp. 127–138.
- U.S. Department of Labor, 2012. Number of Jobs Held, Labor Market Activity, and Earnings Growth Among the Youngest Baby Boomers: Results From a Longitudinal Survey Summary. Economic News Release.
- Wagner, J., Ludtke, O., Robitzsch, A., 2019. Does personality become more stable with age: disentangling state and trait effects for the big five across the life span using local structural equation modeling. *J. Pers. Soc. Psychol.* 116, 666–680.
- Wang, Y., Du, Y., Li, J., Qiu, 2019. Lifespan intellectual factors, genetic susceptibility, and cognitive phenotypes in aging: implications for interventions. *Front. Aging Neurosci.* 11, 1–11.
- Wethington, E., Kessler, R.C., Pixley, J.E., 2004. Turning points in adulthood. In: Brim, O.G., Ryff, C.D., Kessler, R.C. (Eds.), *How Healthy are We? A National Study of Well-Being at Midlife*. The University of Chicago Press, Chicago, IL, pp. 586–613.
- Wilk, C.A., Kirk, M.A., 1995. Menopause: a developmental stage, not a deficiency disease. *Psychotherapy* 32, 233–241.
- Willis, S.L., Schaie, K.W., 1999. Intellectual functioning in midlife. In: Willis, S.L., Reid, J.D. (Eds.), *Life in the Middle: Psychological and Social Development in Middle Age*. Academic, San Diego, CA, pp. 233–247.

Older adults' development, learning, and education

Lisa M. Baumgartner, Department of Counseling, Leadership, Adult Education and School Psychology, Texas State University, San Marcos, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Physical changes	503
Vision	503
Hearing	504
Brain changes	504
Studies on the effect of physical training on cognition	504
Studies on cognitive training and improved cognition	505
Studies on the effects of cognitive and physical training on the brain	505
Older adult learning and education	505
Informal learning	506
Informal learning and technology	506
Men's sheds	507
Older adults' informal learning through volunteering and play	507
Nonformal learning settings: universities of the Third Age, Osher Lifelong Learning Institutes, and Road Scholar	508
Demographics, participation and participants' interests at OLLIs and U3As	509
Summary and conclusions	509
References	510
Relevant website	511

The world's population is aging. Globally, there are "727 million persons aged 65 years or over in 2020" ([United Nations Department of Economic and Social Affairs, Population Division, 2020](#), p. 1). In 2020, 9.3% of the population was over age 65 (p. 1). By 2050, the "number of older persons worldwide is projected to be more than double, reaching over 1.5 billion" (p. 1).

Because the older adult population is increasing, and learning is a lifelong endeavor, it is important to understand older learners and their learning. In this chapter, I discuss the definitions of "older adult." I touch on physical changes, namely differences in sight and hearing because these changes can profoundly affect learning. Additionally, I explore brain changes and the effects of exercise and cognitive training on the brain. Next, I investigate older adults' informal learning in daily activities and nonformal settings. Last, I conclude with observations about the literature on older adults in these areas.

The minimum age used to define an older adult varies. The age of 65 is typically chosen as a minimum age for later adulthood because "In 1889, the German Chancellor Otto von Bismark decided to set this as the age when people could receive social insurance payments" ([Whitbourne and Whitbourne, 2014](#), p. 8). Gerontologists divide older adults into three groups. Individuals ages 65–74 are "young-old," while those 75–84 are "middle-old" with age 85 and older considered "old-old" (p. 8). Some scholars use ages 50 or 60 as minimum ages to define "older adults." Most of the information presented in this chapter will concern adults aged 50 or older.

Physical changes

Individuals' appearances change in the skin, hair, body build, mobility, changes in vital functions, and diminishment in senses, including sight, hearing, taste, touch, and smell. Information about vision and hearing will be featured as a diminishment in these senses may significantly affect learning.

Vision

Although vision loss can occur at any age, most individuals with vision impairments are age 50 or older ([WHO, 2021](#)). Worldwide, in 2020, people aged 50 or older accounted for 78% of those who were blind (33.6 million of 43.3 million) and 70% of those with "moderate or severe vision impairment" (143 million of 265 million) (e133). Declining vision affects how people access and process information. The primary causes for vision loss in older adults include cataracts, glaucoma, macular degeneration, and diabetes-related vision complications ([Pelletier et al., 2016](#)). Cataracts occur when the eye lens becomes clouded ([Tyler, 2020](#)). Glaucoma causes irreversible damage to the optic nerve, which can result in blindness ([Pelletier et al., 2016](#)). Symptoms for age-related macular degeneration can include difficulties adapting to the dark as this disease affects the retina ([Pelletier et al., 2016](#)).

Fortunately, those with access to treatment for eye conditions can regain or maintain some sight. Researchers have shown that vitamin supplements may slow the progression of macular degeneration while eye drops and laser surgery can treat the progression of vision loss due to glaucoma (Pelletier et al., 2016). Cataract surgery does not require hospitalization, and the procedure takes less than an hour (Whitbourne and Whitbourne, 2014).

Hearing

In addition to having vision impairments, many older adults have hearing loss. Approximately 25% of those aged 60 or over experience “disabling” hearing loss worldwide (WHO, 2021, paragraph 6). Disabling hearing loss “refers to hearing loss greater than 35 dB (dB) in the better hearing ear” (WHO, 2021, paragraph 6). For comparison, bird calls are typically about 40 dB, and rustling leaves are about 20 dB (IAC Acoustics, 2021, Table 1). Common conditions related to hearing loss include age-related hearing loss (presbycusis) and tinnitus or ringing in the ears (Whitbourne and Whitbourne, 2014). Hearing aids can ameliorate some hearing issues and those who are hard of hearing use techniques such as looking directly at the person and turning down background noise (Whitbourne and Whitebourne, 2014).

Brain changes

In addition to changes in the senses, age-related brain changes affect memory, planning, learning, and decision-making. Older adults may have difficulty recalling names, find multitasking more challenging, and experience an inability to pay attention (National Institute on Aging, 2020). Brain changes include the shrinkage of the hippocampus, which affects one’s memory (Tyler, 2020). In addition, complex tasks take longer with age because there is a loss of myelin which connects brain areas (Tyler, 2020). However, new neurons continue to form brain-training exercises that help older adults “compensate for any age-related changes” (Erber and Szuchman, 2015, cited in Tyler, 2020, p. 1165). Following are studies on physical and cognitive training effects on the brain.

Studies on the effect of physical training on cognition

In general, physical training improves scores on tests measuring cognition. Data from two extensive studies have been analyzed: The Mexican Health and Aging Study (MHAS) ($n = 5853$) and the Korean Longitudinal Study of Aging (KLoSA) ($n = 5188$) to understand “the aging process in community-dwelling adults” aged 50 and older (Aarsland et al., 2020). Researchers conducted three (Mexico) and four-year (Korea) follow-up studies. In the MHAS, researchers asked people if they engaged in “vigorous exercise” 3 or more times a week, such as “heavy housework or sports” (p. 2). Participants in the MHAS were considered physically active if they participated in 150 min of this type of activity a week, and KLoSA participants were asked to report their physical activities. They were coded as “physically active” if they reported 150 or more minutes of exercise a week and “inactive” if they did less than 150 min a week. Cognitive functioning was assessed using the Cross-Cultural Cognitive Examination test (CCCE) for the MHAS and the Mini-mental state examination (MMSE) for the KLoSA. Physically active Korean participants scored better on these cognitive exams than their inactive peers. However, physically active Mexican participants did not score higher than their inactive peers. Researchers cited Mexicans’ poorer diet and lower average education level (less than six years) for lower cognitive test scores.

Scholars conducted a meta-analysis of 39 studies that had randomized controlled trials of physical exercise interventions in community-dwelling adults with an outcome measure of cognitive function (Northey et al., 2018). Participants were aged 50 or older. Thirty-six studies demonstrated that physical exercise, including “aerobic exercise, resistance training, multicomponent training, and tai chi” improved cognition (p. 154). Based on the findings, moderate aerobic and resistance training exercise is recommended for 45 min a day.

Similarly, Netz (2019) conducted a historical review of highly cited research studies on the effect of exercise on cognition. He found physical training, such as aerobic and strength training, improved brain blood flow, which improved cognition, while motor training, such as balance and coordination activities, directly affected cognition. The intensity of the physical exercise enhances the brain’s ability to adapt to change by forming more neural connections and ultimately improves cognition, whereas, with motor training, the complexity of the exercise improves cognition. Dual-task training that included cognitive demands in addition to physical or motor activity was more effective in preserving or improving cognitive functioning than a single task.

Kramer and Colcombe (2018) revisited their meta-analytic study of fitness effects on cognitive function. The original study (Colcombe and Kramer, 2003) reviewed 18 studies published from 1966 to 2001 that examined exercise effects on cognition in randomized controlled human trials with adults aged 55 or older. Exercise had the most significant impact on “executive control functions such as planning, scheduling, dealing with ambiguity, working memory, and multitasking” (p. 214). Since their review, more studies have examined how exercise results in brain changes. These studies have found that cardio fitness increases hippocampal

volume (Erickson et al., 2011, cited in Kramer and Colcombe, 2018, p. 214). In addition, aerobic training has improved parietal regions of an attention network that corresponded to improved performance on a selective attention task (Colcombe et al., 2004, cited in Kramer and Colcombe, 2018, p. 214). The authors note that most studies of change in brain structure are cross-sectional and observational, and there is a need for studies that establish causality between brain health and fitness.

Studies on cognitive training and improved cognition

Like physical training, some studies have shown brain training improves cognitive scores. For example, a 12-week chess training improved older adults' cognitive state, mood, and quality of life (Cibeira et al., 2021). Researchers recruited participants from day-care centers and nursing homes. Twenty-two participants over age 60 completed the study. Eleven participants spent 2 hours a week in a chess-training program, and the 11 participants in the control group had no intervention. Results showed "improved general cognitive status, attention processing speed, executive functions and QoL (quality of life)" for those trained in chess (p. 899).

Nguyen et al. (2019) conducted a meta-analysis of articles that examined the immediate and long-term effects of cognitive training on cognitive flexibility, inhibition, or working memory. Qualifying studies included a pretest and post-test and involved participants aged 60 or older with no neurological impairments, mild cognitive impairment, or diagnoses of dementia. Cognitive training occurred on a technological device, including a computer, gaming device, or mobile device. Researchers measured changes in attention, fluid intelligence, global executive function, language, memory, processing speed, and visuospatial ability. Sixty-one of the studies examined the immediate effect of training. Sixteen studies focused on the long-term effects of cognitive training (from three weeks to eighteen months). Cognitive training gains were significant in executive functioning, memory, and visuospatial ability from pretest to posttest.

Another meta-analysis of experimental studies on cognitive training showed mixed results (Sala and Gobet, 2019). Training improved working memory but did not improve other outcomes, such as verbal and non-verbal intelligence (Sala and Gobet, 2019) or enhance generalized cognitive function (Guye and von Bastian, 2017). The authors of the meta-analysis cautioned that many of the studies involved poor study designs and analysis and that placebo effects may have been responsible for some of the positive effects of brain training. They concluded that, at best, training on a particular task (working memory) may improve performance on that task but does not improve general intelligence (Sala and Gobet, 2019).

Studies on the effects of cognitive and physical training on the brain

Researchers have examined the effects of physical and cognitive training on the brain. Chapman et al. (2016) studied the differential effects of cognitive and physical training on the brain and behavior. Respondents completed cognitive training ($n = 18$) or physical training ($n = 18$) for 12 weeks. Participants learned strategies to improve their attention, reasoning abilities, and innovative thinking skills. The physical training program consisted of aerobic activity for 150 min a week. Exercise increased performance on memory tasks, whereas cognitive training enhanced executive function. Likewise, in a review of 19 articles on the effects of physical and cognitive training, researchers found that dual exercise training (e.g., aerobic training and weight training) interventions or physical-cognitive training interventions improved physical and cognitive test scores and that this type of training was the best intervention strategy (Levin et al., 2017). The authors noted that because the exercise interventions and test protocols were so diverse, it was impossible to see how these types of training affected individuals' performances.

In summary, physical training, including aerobic and resistance training and motor activities, has improved cognitive function (Aarsland et al., 2020; Netz, 2019; Northey et al., 2018). Cognitive training enhances attention processing and executive functions (Cibeira et al., 2021; Nguyen et al., 2019) and learning and memory. Researchers have found that cognitive training improves working memory but not other areas such as verbal or non-verbal intelligence (Guye and von Bastian, 2017) or general cognition (Sala and Gobet, 2019). Last, physical-cognitive training appears to produce the best results, but there needs to be improved testing protocols and interventions to understand better the associations between training and performance.

Cognitive changes affect learning, but how and where does adult learning occur? In the next section, I review the different types of learning and settings in which this learning occurs for older adults.

Older adult learning and education

Learning occurs in different settings. Formal education "refers to educational institutions including all levels of schools both private and public as well as specialized programs offering technical and professional training" (Merriam and Brockett, 2011, p. 15). Nonformal education is "any organized educational activity outside the established formal system ... that is intended to serve identifiable learning clientele and learning objectives" (Coombs, 1989, p. 11, cited in Merriam and Brockett, 2011, p. 15). Finally, informal learning is "generally unplanned, experience-based, incidental learning that occurs in the process of people's daily lives" (Merriam and Brockett, 2011, p. 14). Examples of this type of learning include reading a blog, viewing a TikTok video, or conversing with a friend. Since most older adults' learning experiences occur in nonformal settings and informal learning, I focus on these areas.

Informal learning

Informal learning occurs as adults navigate daily life. Examples of informal learning of older adults could include internet searches on leisure activities such as pickleball, crafts, or conducting genealogical research. In addition, informal learning includes self-directed learning. Individuals decide what they want to learn, which materials they want to use, how they want to engage in learning, and their self-evaluation (Knowles, 1975). An example of a self-directed learning project may be learning how to build a birdhouse. The learner decides whether to consult others on how to build a birdhouse. They may view YouTube videos, read books, or use any other methods they choose. The final product is a functional birdhouse. This section on informal learning discusses older adults' interaction with technology in their informal learning as technology is part of everyday life. Next, I focus on research on "Men's Sheds," which provide a casual learning environment for men, who typically come from a background in the trades, to gather and exchange information and work on projects informally. Last, I conclude the section with a discussion on older adult learning through volunteering and play.

Informal learning and technology

Technology permeates daily life. Its adoption for older adults' learning purposes is an area of research interest (Barrantes-Caceres and Cozzubo-Chaparro, 2019; Lai, 2020). Lai (2020) surveyed 238 older adult students, 55 and older, from 11 computer classes in central Taiwan. All participants used a smartphone or tablet for at least 30 min a day. Social influence (e.g., the opinions of their children and grandchildren), facilitating conditions (environmental factors that made use of technology easier), and performance expectancy (how the technology will help people achieve their goals), affected older adults' adoption of smartphones and tablets for learning. Older adults living with children were more likely to access the internet as the younger people can provide computer support and instruction (Barrantes-Caceres and Chaparro-Cozzubo, 2019). In addition, having a spouse or partner increases the amount of *time* older people spend on the internet (Barrantes-Caceres and Chaparro-Cozzubo, 2019).

Several studies concerned older adults' use of technology in their informal learning. Morrison and McCutcheon (2019) investigated older Canadians' online personal learning networks (oPLNs). oPLNs are "a collection of individuals and/or groups with whom you communicate, via the Internet, to engage in discussion, exchange information, and/or share resources to learn more about a hobby or interest" (Morrison and McCutcheon, 2019, p. 4). These researchers wanted to know who engages in online self-directed learning? What do oPLNs "look" like for these individuals (e.g., access, type, purpose)? How do these individuals use their oPLNs for informal self-directed learning? What is the perceived value of self-directed learning in their oPLNs? They surveyed 385 retired older adults who had oPLNs, and were 55 or older. The sample demographics included 57% male, most lived in Ontario, Canada, and had a household income of \$60,000 or more. Most respondents had used the internet for more than 15 years, with no significant difference in age and internet use.

Participants reported communicating with their oPLNs via email, Facebook, blogs, Google plus, and video conferencing, with email being the most common way to communicate (Morrison and McCutcheon, 2019). The purposes for using oPLNs included wanting to learn new things (90.9%), learning from others (73%), sharing knowledge (28.8%), "to learn by interacting with others" (45.7%), (p. 8), and "talking with others about interests or hobbies" (39.5%). (p. 8). Respondents were most likely to access and share websites for information. Nearly 73% strongly agreed or agreed that they "felt a sense of community with other learners in their oPLN" (p. 9). In addition, 86% said that members of their oPLNs provided expert knowledge. Younger participants (age 55–64) were more likely than the older participants (age 75 and above) to seek information online actively.

Older adults may also access the internet via a mobile device. In a literature review that examined the informal learning of older adults using mobile devices, Jin et al. (2019) reviewed 28 studies that met the following criteria: (a) appeared in a peer-reviewed journal, (b) written in English, and (c) published between 2005 and 2017. The 2005 date was chosen because wireless mobile communication became more prevalent, and data speeds increased. Several themes emerged. People engaged in health-related self-directed learning on their mobile devices. Individuals shared information about their health concerns and used mobile applications to monitor their fluid intake, weight, exercise, memory improvement and meditation. Respondents generally had positive attitudes toward using technology, and they saw the benefits of using mobile educational applications. However, they wanted to be assured of privacy, needed help setting up the devices, and wanted "easy access to continuous tutoring and instructional manuals" (p. 130). Older adults use mobile devices to communicate with others, including intergenerational communication. Studies showed the potential for collaborative learning by seniors using mobile devices. Participants shared information and commented on video content. The authors encouraged more research on specific cohorts of older adults (e.g., young-old (65–74), middle old (75–84), and old-old (85 and up) (Jin et al., 2019).

Scholars have also studied the informal learning of older adults on social media. Ebardo et al. (2020), for example, wanted to know why older adults used Facebook. The researchers collected 1707 posts between February and June 2020 from a Facebook community of 71,133 Filipino older adults in the Philippines (Ebardo et al., 2020). The most prominent theme was staying healthy. People discussed their routines for staying healthy and solicited advice when they had a health crisis. Second, individuals discussed staying safe from weather-related events. Third, they discussed caring for family members and maintaining relationships through COVID-19.

Older adults also find health-related support groups online. Litchman et al. (2018) conducted in-depth interviews with 20 United States Baby Boomers from a diabetes online community (DOC). Nine participants were male, with 11 were female; 14 participants possessed a bachelor's degree or higher, and 19 of the 20 identified as White. Most participants had Type 1 diabetes.

Participants used the DOC to learn about and manage diabetes, obtained emotional support from DOC members, and DOC "elders" served as role models and resources for less experienced members. Older members enjoyed giving back to the community. In summary, the DOC filled knowledge gaps and helped with the day-to-day management of diabetes.

Men's sheds

Men's Sheds originated in Australia in the late 1990s (Golding, 2015). These settings provide informal learning in a social, community setting for "older, typically retired, and sometimes isolated men" (Golding, 2011, p. 105). Activities at each Men's Shed vary according to men's interests, but generally there are places for hands-on activities and socialization (Carragher and Golding, 2015). Technology can be informal learning in Men's Sheds, although typical projects include woodworking and other trade-related activities. While Men's Sheds are open to all, this setting has been shown to appeal to older men with vocational schooling and less formal education (Golding et al., 2007). In this setting, men are not "treated by the organization, or by the women involved as customers, patients, students or clients from negative, deficit or ageist models" (Golding and Carragher, 2015, p. 104). Instead, they make friends and socialize, gain a sense of belonging (Carragher and Golding, 2015), feel increased self-worth, increased physical fitness (Crabtree et al., 2018) and feel less isolated (Nurmi et al., 2018). Attendees are both teachers and learners (Nurmi et al., 2018).

Recently, scholars have explored the effect of Men's Sheds on men's health and wellbeing. Kelly et al. (2019) provided a logic model of health and well-being outcomes of Men's Shed's activities. Their model had five components: interventions, mediating variables, intermediate outcomes, long-term outcomes, and results in physical health and well-being. They reported interventions found in studies, including providing skills and education, socialization, and the inclusivity of Men's Sheds spaces. A second model component, mediating variables, included increased skills, physical movement, improved social networks, improved health literacy, and decreased alcohol and drug use in favor of shed activities. Third, intermediate outcomes included increased fitness, motivation to leave the house, better social bonds, health-seeking behaviors, improved family relationships, and increased feelings of security and safety. Long-term outcomes were listed as increased confidence, decreased socialization, increased feelings of belonging, happiness, enjoyment, with less anxiety and feelings of depression. The last component of the model resulted from improved physical health, mental health, and social well-being. The authors provided an observed pathway to physical health: (1) The men's shed providing opportunities for physical activity increased physical movement—improved fitness—improved physical health. The authors developed "a logic model showing how Men's Sheds may influence health and well-being" but state, "There is a lack of reliable and systematic evidence of the health and well-being outcomes of shed activities" (p. 1155). Few studies included demographic information, and "Associative relationships between factors such as employment status of marital status and health and well-being outcomes are under-explored" (p. 1155). Areas of concern include a lack of exploration of contextual factors such as a rural location or levels of deprivation and subsequent links to health inequalities within communities.

Further, many studies were qualitative and had small sample sizes. The authors call for more quantitative and mixed-methods studies and longitudinal approaches. Subsequent studies have confirmed these health-related findings, such as increased mobility, decreased alcohol use, and improved health literacy (Kelly and Steiner, 2021). Some authors have answered the call for an expansion of research on the men's shed idea by detailing the development of women's sheds (e.g., Golding et al., 2021).

Intergenerational learning can also occur at Men's Sheds. Misan et al. (2018) examined the intergenerational learning between women college students majoring in Occupational Therapy and older men. College students conducted informal information sessions on cooking, exercise, and diet. Men participated in lawn bowling and a health check. The older men reciprocated by informally discussing the use of hand tools, portable power tools, and basic wood construction and finishing techniques. The men were happy to share their skills with the much younger women, building rapport.

In a study that examined how Men's Sheds could be marketed, researchers conducted focus groups with 64 men with 57 of them being 55 years or older (Nurmi et al., 2018). Twenty-two men previously had attended Men's Sheds. Additional participants were recruited from community organizations, coffee shops, and through a one-day advertisement in community newspapers. Researchers conducted four focus groups: two with non-Men's Shed members and two groups of Men's Shed members. Men's Shed's members discussed the activities men are interested in and how marketing is done. Non-members provided opinions about the branding images they saw about Men's Sheds. Men recommended advertisements be placed at men's places of employment, in newspapers, and in the community to attract new members. Another idea was having a Men's Shed member invite others to the Men's Shed. Creating a welcoming atmosphere would also attract members (Nurmi et al., 2018). Men mentioned barriers to attendance, including traveling outside their own neighborhood, not speaking English as a first language, and health issues.

Older adults' informal learning through volunteering and play

Older adults also engage in informal learning in other venues. Researchers wanted to understand how older men learned through volunteer activities (Chen, 2016). Participants were retired men over sixty who volunteered for at least two years. The researcher reached out to the non-profit organization and used personal networks to gain participants. Seventeen participants between the ages of 63 and 93 were interviewed. Researchers used the constant comparative method to uncover themes.

The results indicated that volunteers became more knowledgeable, skilled, and wiser (Chen, 2016). The knowledge learned depended on the type of volunteering. A nursing home volunteer gained medical knowledge while a museum volunteer learned about local history, astronomy, and botany. Second, learners used their life experiences were a foundation for learning. They

learned new things or deepened learning from what they already knew. They learned by watching others, doing the activities, engaging in individual reflection, talking with other volunteers, and participating in formal training provided by the institution. For example, one man learned better interpersonal skills by reflecting on interactions with others. Individuals learned skills through formal courses or by sharing information with other volunteers. A third theme included participants believing that they learned differently through volunteering than in the classroom. They found the volunteer type of learning more practical than classroom learning.

Another area of interest was how older adult motorcyclists learned through play (Zacharakis, 2019). He used grounded theory to uncover themes. He interviewed six individuals over age 60 (5 men and a woman). Participants developed their motorcycle skills through reading magazines, participating in workshops and classes, and talking with other motorcyclists. A participant liked teaching his children and grandchildren to motorcycle. Another worked on his agility and balance in an obstacle course. He enjoyed competing on obstacle courses, socializing with others, and repairing bikes. Others looked at websites and read magazines to learn more about motorcycling. Zacharakis (2019) found that people seek additional information to improve their knowledge and skills about motorcycling. Learning through play was intrinsically motivated. In addition, the social nature of the hobby was essential to people. The author noted that this study confirmed Bandura's (1977) theory that people learn from observation and enjoy the social part of learning. Second, this is an example of informal learning where the learner oversees the teaching (Marsick and Watkins, 1990, cited in Zacharakis, 2019, p. 181).

In summary, informal learning occurs in our daily lives. First, an investigation of the benefits of informal learning in Men's Sheds shows men socialize, learn and exchange information, make friends, and feel valued (Carragher and Golding, 2015; Crabtree et al., 2018; Golding et al., 2007; Nurmi et al., 2018). Second, much informal learning involves technology. Older adults adapt to technology if they are encouraged by family members to do so, environmental factors are conducive to adapting technology, and if there is support for technical issues (Barrantes-Caceres and Chappro-Cozzubo, 2019). Third, the use of technology in informal learning shows that adults use a variety of technologies (e.g., Google, Email, blogs, videoconferencing) to share information to learn new things, learn from others, gain a sense of community, and share their hobbies (Morrison and McCutcheon, 2019). Last, older people learn information as volunteers at various organizations (Chen, 2016), and they learn specific skills and knowledge like how to motorcycle and the benefits of motorcycling.

Nonformal learning settings: universities of the Third Age, Osher Lifelong Learning Institutes, and Road Scholar

Informal learning can occur in nonformal settings. Although nonformal settings can include book clubs, knitting groups, and sports clubs, I focus on centers where adults can engage in nonformal learning, as this has been an area of interest for researchers. For example, universities of the Third Age (U3As), primarily found in Europe and Asia, offer older adults the opportunity to engage in largely non-credit courses of their choosing. Individuals in North America have Osher Lifelong Learning Institutes (OLLIs) and Road Scholar (formerly Elderhostel) that provide similar programming (Formosa, 2019). The term "Third Age" refers to the third life stage where individuals are retired (Vellas, 2019, p. 19).

The first U3A program originated at the University of Toulouse, France, in 1968, and its curriculum initially focused on "gerontological subjects," later focusing on "humanities" (Formosa, 2019, p. 8). Teachers favored the lecture approach, although students also participated in debates, field trips, and recreational opportunities. Nowadays, lecture offerings are diverse and include Tai Chi, history, philosophy, art, religion, law, economics, and natural sciences (Vellas, 2019). The French model is linked with universities, and experts make programming decisions rather than class participants (Vellas, 2019).

The U3A in the United Kingdom (UK) was founded in 1981 (Percey, 2019) and has approximately 450,000 members with 1057 U3As (U3Aa, 2021, paragraph 2). Unlike the top-down French model, the UK model has local groups engage in peer learning (non-credit). There is a myriad of courses. Subjects include astronomy, boating, chess, etymology, learning to be retired, mindfulness and meditation, and sociology, among others (U3Ab, 2021, paragraph 5). U3As around the world use some variation of the French and British models. Countries with U3As include Iceland, Italy, Germany, Malta, Poland, the Russian Federation, Spain, Sweden, Australia, China, Lebanon, Malaysia, New Zealand, South Korea, Taiwan, and Thailand (Formosa, 2019).

In the United States, Osher Lifelong Learning Institutes and Road Scholar provide nonformal education for older adults. Osher Lifelong Institutes are associated with colleges and universities offering non-credit courses to adults aged 50 or older (Bernard Osher Foundation, 2005, paragraph 2). There are 124 OLLIs in the US (Bernard Osher Foundation, 2005 paragraph 2). Course offerings vary by Institute and include lectures by faculty, community experts, and OLLI members (University of Oregon, 2021). https://osher.uoregon.edu/about/types_of_courses.php (paragraph 2).

In the United States, Road Scholar (formerly Elderhostel) was founded in 1975 as a "not for profit education organization" (Road Scholar, 2021a, paragraph 1), with the first programs occurring at college campuses in New Hampshire (Road Scholar, 2021a, paragraph 2). Initially, the Elderhostel program offered courses to people aged 60 or older at colleges in New England. Road Scholar, 2021c, paragraph 4). By 1980, all 50 US States had programs, and by 1981 the organization had its first programs outside the United States (Road Scholar, 2021a paragraph 4). Currently, Road Scholar offers "5500 learning adventures in 150 countries and all 50 [US] states, serving more than 100,000 participants per year," and the program is not tied to college campuses (Road Scholar, 2021b, para. 5). Programs are open to all ages but typically attended by those aged 50 and older.

Demographics, participation and participants' interests at OLLIs and U3As

Researchers investigated various aspects of OLLIs, Road Scholar programs, and other Learning in Retirement centers. An area of focus has been on member profile differences (Lee et al., 2018), participation patterns (Truluck et al., 2010), members' experiences at OLLIs (Talmadge et al., 2019), barriers to age-friendly universities (Hansen et al., 2019), and membership characteristics (Hansen et al., 2016). Membership characteristics remain similar across studies. The typical OLLI or Learning in Retirement Center participant is a highly educated White woman. At least two-thirds of participants are women (Hansen et al., 2019; Truluck et al., 2010). At least 79% possess a bachelor's degree or higher (Hansen et al., 2019; Truluck et al., 2010). Those aged 60–74 comprise the most significant membership (Lee et al., 2018). Liberal arts, public policy, and current affairs courses are popular (Hansen et al., 2016; Lee et al., 2018; Truluck et al., 2010). Barriers to full participation in OLLI programs related to time, cost, transportation, and issues related to health, hearing, and physical mobility (Hansen et al., 2019). Racially, only 5.24% of OLLI participants identify as “non-White,” with “2.66% identifying as African American” (Hansen et al., 2019, p. 233). Race and the percentage of individuals with college or graduate degrees in these programs may be barriers to less-educated people and people of color. The authors recommend having OLLIs reach out to those who may not come to OLLIs because of unpleasant educational experiences (Hansen et al., 2019).

Some of the research and scholarship on U3As concerning older adult learning and education focuses on programmatic issues and/or benefits of U3As (de Maio Nascimento and Giannouli, 2019; Hachem and Vulopala, 2016; Ricardo and Porcarelli, 2019) and identity issues such as participants' concepts of old age (Camargo et al., 2018). Scholars analyzed U3A program models in various countries. For example, Brazil blends the French and British U3A models and offers formal courses through universities and extension courses on art, literature, health education, tourism and leisure, general knowledge, and foreign languages (de Maio Nascimento and Giannouli, 2019). Given Brazil's “intense body culture,” the U3A's focus on physical activities for older adults stands out. These courses benefit older adult participants and contribute to older adults' physical activity knowledge. In addition, older adults are part of university-conducted research studies that focus on action research and empowerment.

Hachem and Vuopala (2016) surveyed 461 older adults at a University of the Third Age in Lebanon to uncover the benefits and challenges faced by participants. Eighty-three percent were women, with 71.5% being between 50 and 69 years old. Cognitive benefits included gaining knowledge and intellectual stimulation. Social benefits included making friends, learning from each other, and belonging to a community. Psychological benefits included experiencing positive emotions. Participants' challenges included the use of lectures and the timing of sessions.

In summary, membership characteristics for OLLIs and Learning in Retirement centers tend to be middle-class, highly educated White women. There is a plethora of course offerings and adventures in these settings. Barriers to participation in classes include time (Hansen et al., 2019), and fewer people of color or those without a college education attend courses at these institutions (Lee et al., 2018; Truluck et al., 2010).

Summary and conclusions

Physical changes in older adults may affect how they process information. Changes to vision and hearing may be especially challenging. Glaucoma, presbyopia, and cataracts are common eye diseases that affect adults worldwide (WHO, 2021). However, treatments for glaucoma, including eye drops and laser surgery, may be options for those with glaucoma and, similarly, hearing loss is common in older adults, including presbycusis and tinnitus (Whitbourne and Whitbourne, 2014). Hearing aids can help with hearing loss.

In addition to vision and hearing changes, brain changes also occur. Multitasking becomes more challenging, and brain areas, including the hippocampus, may shrink (Tyler, 2020). However, physical training has improved cognitive function (Aarsland et al., 2020). Mental training has been shown to improve executive functioning, memory, and visuospatial ability (Nguyen et al., 2019). Cognitive training improves attention processing and executive functions (Nguyen et al., 2019) but not other areas such as verbal or non-verbal intelligence (Guye and von Bastian, 2017) or general cognition (Sala and Gobet, 2019).

There will be advancements in medical science. More will be discovered about the older adults' brain and maintaining and improving cognition through physical and cognitive exercises and diet. Additionally, scientific advancements in treating eye diseases and hearing loss may allow older adults to participate in learning activities more fully in the future.

Most adults engage in informal learning. This learning can occur in everyday life or nonformal settings such as knitting groups or organizations like Universities of the Third Age. (U3As). Recently, scholarship has focused on participation in OLLIs or U3As or Learning in Retirement Centers (Truluck et al., 2010), membership profiles (Hansen et al., 2016; Lee et al., 2018), and demographics of organizations (Lee et al., 2018; Truluck et al., 2010) that cater to older adults such as U3As, Road Scholar and Osher Lifelong Learning Institutes. People of color and individuals with lower levels of education are less likely to attend these learning opportunities. However, the older, less educated man may be interested in the informal learning opportunities in Men's Sheds (Carragher and Golding, 2015), where men informally gather to work on hands-on projects and exchange information about various concerns.

In addition to continuing research on the informal learning of older adults in nonformal settings, researchers should focus on more clearly defining the term “older adult” in their research. They should also seek to understand the learning experiences of older adults marginalized by gender, sexual orientation, ability status, class, and race/ethnicity. Finally, taking a more critical look at the influence of social capital on the learning experiences of older adults may be in order. Learning is lifelong, and as the population of older adults increases worldwide, more researchers need to be interested in various aspects of this population.

References

- Aarsland, V., Borda, M.G., Aarsland, D., Garcia-Cifuentes, E., Anderssen, S.A., Tovar-Rois, D.A., Gomez-Arteaga, C., Perez-Zapeda, M.U., 2020. Association between physical activity and cognition in Mexican and Korean older adults. *Arch. Gerontol. Geriatr.* 89, 104047.
- Bandura, A., 1977. *Social Learning Theory*. Prentice Hall, Englewood Cliffs.
- Barrantes-Caceres, R., Cozzubo-Chaparro, A., 2019. Age for learning, age for teaching: the role of intergenerational, intra-household learning in internet use by Internet services Constanza Bianchi *Journal of Services Marketing older adults in Latin America. Inf. Commun. Soc.* 22, 250–266.
- Bernard Osher Foundation, 2005. The Programs. <https://www.osherfoundation.org/programs.html>.
- Camargo, M.C., Lima-Silva, T.B., Ordonez, T.N., Batistoni, S.S.T., Yassuda, M.S., de Melo, R.C., Lopes, A., Domingues, M.A.R.D.C., Cachioni, M., 2018. Beliefs, perceptions, and concepts of old age among participants of a University of the Third Age. *Psychol. Neurosci.* 11, 417.
- Carragher, L., Golding, B., 2015. Older men as learners: Irish men's sheds as an intervention. *Adult Educ. Q.* 65, 152–168.
- Chapman, S.B., Aslan, S., Spence, J.S., Keebler, M.W., DeFina, L.F., Didehban, N., Perez, A.M., Lu, H., D-Esposito, M., 2016. Distinct brain and behavioral benefits from cognitive vs. physical training: a randomized trial in aging adults. *Front. Hum. Neurosci.* 10, 338.
- Chen, L.K., 2016. Not just helping: what and how older men learn when they volunteer. *Educ. Gerontol.* 42, 175–185.
- Cibeira, N., Lorenzo-López, L., Maseda, A., Blanco-Fandiño, J., López-López, R., Millán-Calenti, J.C., 2021. Effectiveness of a chess-training program for improving cognition, mood, and quality of life in older adults: a pilot study. *Geriatr. Nurs.* 42, 894–900.
- Colcombe, S., Kramer, J., 2003. Fitness effects on the cognitive function of older adults. A metaanalytic study. *Psychol. Sci.* 14, 125–130.
- Colcombe, S.J., Kramer, A.F., Erickson, K.I., Scalf, P., McAuley, E., Cohen, N.J., Webb, A., Marquez, D., et al., 2004. Cardiovascular fitness, cortical plasticity, and aging. *Proc. Natl. Acad. Sci.* 101, 3316–3321.
- Coombs, P.H., 1989. Formal and nonformal education: future strategies. In: Titmus, C.J. (Ed.), *Lifelong Education For Adults: An International Handbook*. Pergamon Press, New York, pp. 57–60.
- Crabtree, L., Tinker, A., Glaser, K., 2018. Men's sheds: the perceived health and wellbeing benefits. *Work. Older People* 22, 101–110.
- de Maio Nascimento, M., Giannouli, E., 2019. Active aging through the University of the Third Age: the Brazilian model. *Educ. Gerontol.* 45, 11–21.
- Ebarido, R., Tuazon, J.B., Suarez, M.T., 2020. We learn from each other: informal learning in a Facebook community of older adults. In: *Proceedings of the 28th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education. Center for Science and Technology of Learning, Taiwan*, pp. 594–601.
- Erber, J., Szuchman, L., 2015. *Great Myths of Aging*. John Wiley and Sons, West Sussex, UK.
- Erickson, K., Voss, M., Prakash, R., Basak, C., Szabo, A., Chaddock, L., et al., 2011. Exercise training increases size of hippocampus and improves memory. *Proc. Natl. Acad. Sci.* 108, 3017–3022.
- Formosa, M., 2019. Active ageing through lifelong learning: a University of the Third Age. In: Formosa, M. (Ed.), *The University of the Third Age and Active Ageing: European and Asian-Pacific Perspectives. International Perspectives on Aging Series*, vol. 23. Springer, Berlin, Germany, pp. 3–18.
- Golding, B., Carragher, L., 2015. Community men's sheds and informal learning: an exploration of their gendered roles. In: *Private World(s): Gender and Informal Learning of Adults*. Brill Sense, Boston, MA, pp. 103–118.
- Golding, B., Brown, M., Foley, A., Harvey, J., Gleeson, L., 2007. Men's Sheds in Australia: Learning Through Community Contexts. NCVER, Adelaide. Available online at: www.ncver.edu.au/publications/1780.html.
- Golding, B., Carragher, L., Foley, A., 2021. The women's shed movement: scoping the field internationally. *Aust. J. Adult Learn.* 61, 150–174.
- Golding, B., 2011. Older men's wellbeing through community participation in Australia. *Int. J. Men's Health* 10, 26–44.
- Golding, B., 2015. *The Men's Sheds Movement: The Company of Men*. Common Ground Publishing, Chicago, IL.
- Guye, S., Von Bastian, C.C., 2017. Working memory training in older adults: Bayesian evidence supporting the absence of transfer. *Psychol. Aging* 32, 732–746.
- Hachem, H., Vuopala, E., 2016. Older adults, in Lebanon, committed to learning: contextualizing the challenges and the benefits of their learning experience. *Educ. Gerontol.* 42, 686–697.
- Hansen, R.J., Brady, E.M., Thaxton, S.P., 2016. Demographic and behavioral characteristics of Osher lifelong learning institute members. *J. Cont. High Educ.* 64, 42–50.
- Hansen, R.J., Talmage, C.A., Thaxton, S.P., Knopf, R.C., 2019. Barriers to age-friendly universities (AFU): lessons from Osher Lifelong Learning Institute demographics and perceptions. *Gerontol. Geriatr. Educ.* 40, 221–243.
- IAC Acoustics, 2021. Comparative Examples of Noise Levels. <https://www.iacacoustics.com/blog-full/comparative-examples-of-noise-levels.html>.
- Jin, B., Kim, J., Baumgartner, L.M., 2019. Informal learning of older adults in using mobile devices: a review of the literature. *Adult Educ. Q.* 69, 120–141.
- Kelly, D., Steiner, A., 2021. The impact of community men's sheds on the physical health of their users. *Health Place* 71, 102649.
- Kelly, D., Steiner, A., Mason, H., Teasdale, S., 2019. Men's Sheds: a conceptual exploration of the causal pathways for health and well-being. *Health Soc. Care Community* 27, 1147–1157.
- Knowles, K., 1975. *Self-Directed Learning*. Association Press, New York, NY.
- Kramer, A.F., Colcombe, S., 2018. Fitness effects on the cognitive function of older adults: a meta-analytic study—revisited. *Perspect. Psychol. Sci.* 13, 213–217.
- Lai, H.J., 2020. Investigating older adults' decisions to use mobile devices for learning, based on the unified theory of acceptance and use of technology. *Interact. Learn. Environ.* 28, 890–901.
- Lee, J.M., Rogers, A., Young, W., 2018. Member profile differences between Florida and National OLLI Institutes. *Adult Learn.* 29, 141–149.
- Levin, O., Netz, Y., Ziv, G., 2017. The beneficial effects of different types of exercise interventions on motor and cognitive functions in older age: a systematic review. *Eur. Rev. Aging Phys. Act.* 14, 20. <https://doi.org/10.1186/s11556-017-0189>.
- Litchman, M.L., Rothwell, E., Edelman, L.S., 2018. The diabetes online community: older adults supporting self-care through peer health. *Patient Educ. Counsel.* 101, 518–523.
- Marsick, V.J., Watkins, K.E., 1990. *Informal and Incidental Learning in the Workplace*. Routledge, London.
- Merriam, S.B., Brockett, R.G., 2011. *The Profession and Practice of Adult Education: An Introduction*. John Wiley & Sons, San Francisco, CA.
- Misan, G., Ellis, B.J., Hutchings, O.M., Beech, A.K., Moyle, C., Thiele, N.J., 2018. Teaching old dogs new tricks: health promotion through intergenerational learning in a regional men's shed. *Aust. Int. J. Rural Educ.* 28, 3–16.
- Morrison, D., McCutcheon, J., 2019. Empowering older adults' informal, self-directed learning: harnessing the potential of online personal learning networks. *Res. Pract. Technol. Enhanc. Learn.* 14, 1–16.
- National Institute on Aging, 2020. How the Aging Brain Affects Thinking. <https://www.nia.nih.gov/health/how-aging-brain-affects-thinking>.
- Netz, Y., 2019. Is there a preferred mode of exercise for cognition enhancement in older age?—a narrative review. *Front. Med.* 6, 57.
- Nguyen, L., Murphy, K., Andrews, G., 2019. Immediate and long-term efficacy of executive functions cognitive training in older adults: a systematic review and meta-analysis. *Psychol. Bull.* 145, 698–733.
- Northey, J.M., Cherbuin, N., Pumpa, K.L., Smee, D.J., Rattray, B., 2018. Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. *Br. J. Sports Med.* 52, 154–160.
- Nurmi, M.A., Mackenzie, C.S., Roger, K., Reynolds, K., Urquhart, J., 2018. Older men's perceptions of the need for and access to male-focused community programmes such as men's sheds. *Ageing Soc.* 38, 794–816.
- Pelletier, A.L., Rojas-Roldan, L., Coffin, J., 2016. Vision loss in older adults. *Am. Fam. Physician* 94, 219–226.

- Percey, K., 2019. An alternative ageing experience: an account and assessment of the University of the Third Age in the United Kingdom. In: Formosa, M. (Ed.), *The University of the Third Age and Active Ageing: European and Asian-Pacific Perspectives, International Perspectives on Aging Series*, vol. 23. Springer, Berlin, Germany, pp. 33–43.
- Ricardo, R., Porcarelli, A., 2019. Education and socialisation in later life: the case of a University of Third Age in Portugal. *Eur. J. Res. Educ. Learn. Adults* 10, 247–260.
- Road Scholar, 2021a. Think of the World as Your Learning Laboratory. <https://www.roadscholar.org/roadscholar-experience/mission-history/>.
- Road Scholar, 2021b. We Are Explorers, Adventurers, and Students of the World. <https://www.roadscholar.org/about/>.
- Road Scholar, 2021c. Our Story. <https://www.roadscholar.org/about/our-story/>.
- Sala, G., Gobet, F., 2019. Cognitive training does not enhance general cognition. *Trends Cognit. Sci.* 23, 9–20.
- Talmage, C.A., Hansen, R.J., Knopf, R.C., Thaxton, S.P., McTague, R., Moore, D.B., 2019. Unleashing the value of lifelong learning institutes: research and practice insights from a National Survey of Osher Lifelong Learning Institutes. *Adult Educ. Q.* 69, 184–206.
- Truluck, J., Kim, Y.S., Valentine, T., 2010. Participation patterns and learning selections of senior members of a learning in retirement organization. *Adult Learn.* 21, 31–36.
- Tyler, S., 2020. Human Behavior and the Social Environment I. University of Arkansas Libraries, Fayetteville. <https://uark.pressbooks.pub/hbse1/>.
- U3Aa, 2021. About. <https://www.u3a.org.uk/about>. (Accessed 1 November 2021).
- U3Ab, 2021. Subject Advice. <https://www.u3a.org.uk/learning/subjects>. (Accessed 18 June 2021).
- United Nations Department of Economic and Social Affairs, Population Division, 2020. World Population Ageing 2020 Highlights: Living Arrangements of Older Persons (ST/ESA/SER.A/451). viewed 12 June 2021. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd-2020_world_population_ageing_highlights.pdf.
- University of Oregon, 2021. Osher Lifelong Learning Institute Continuing and Professional Education. https://osher.uoregon.edu/eugene_springfield/course_descriptions.ph.
- Vellas, F., 2019. Origins and development: the francophone model of Universities of the Third Age. In: Formosa, M. (Ed.), *The University of the Third Age and Active Ageing: European and Asian-Pacific Perspectives, International Perspectives on Aging Series*, vol. 23. Springer, Berlin, Germany, pp. 19–30.
- Whitbourne, S.K., Whitbourne, S.B., 2014. *Adult Development and Aging: Biopsychosocial Perspectives*, fifth ed. Wiley Publishers, San Francisco, CA.
- World Health Organization, 2021. Deafness and Hearing Loss. Viewed August 8, 2021. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>.
- Zacharakis, J., 2019. Elder play: preliminary research results on how older adults learn through motorcycling. *Adult Learn.* 30, 176–182.

Relevant website

- World Health Organization, Blindness and Vision Impairment, 2021. March. Viewed 12 June 2021. <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>.

Education and neurocognitive aging - is there a relation?

Lars Nyberg^{a,b,c,d} and Martin Lövdén^e, ^a Department of Radiation Sciences, Radiology, Umeå University, Umeå, Sweden;

^b Department of Integrative Medical Biology, Umeå University, Umeå, Sweden; ^c Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden; ^d Center for Lifespan Changes in Brain and Cognition, University of Oslo, Oslo, Norway; and ^e Department of Psychology, University of Gothenburg, Gothenburg, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Education and brain aging	514
Education and cognitive aging	515
Conclusion and outlook	518
Acknowledgments	518
References	518

Education has been linked to many advantageous health outcomes (e.g., Ross and Wu, 2006). In the present chapter we will address the question of whether formal education, which typically occurs in childhood and early adulthood and is indexed by years of completed education or educational qualifications, influences “neurocognitive aging”. That is, will education impact the decline in integrity of the brain (e.g., shrinkage of the brain; *atrophy*) and cognitive functioning (e.g., *episodic or working memory*) that are characteristic for the normal aging process and become magnified in pathological aging, such as in neurodegenerative diseases that result in dementia? This question is at the heart of the broader issue of *modifiable risk factors*, an issue which currently is attracting considerable interest in the scientific community in relation to global population aging (Livingston et al., 2017). There is convergence from many longitudinal studies of cognitive aging for marked heterogeneity in the rate of decline in cognitive functions such as episodic memory (Nyberg and Pudas, 2019). Similarly, there are marked individual differences in brain aging, such that some older adults manifest major decline whereas other present with more “youth-like” brains, and the influential *brain-maintenance* theory states that relative lack of brain pathology constitutes the primary determinant of well-maintained cognition in aging (Nyberg et al., 2012). The question of which factors promote maintained brain and cognition in aging is currently intensively researched, and some consistent findings have been reported for both genetic factors (e.g., Josefsson et al., 2012; Lin et al., 2017) and lifestyle activities such as physical activity (Dunas et al., 2021). Critically, theoretical models also include educational attainment as a predictor of rate of neurocognitive change in aging (see e.g., Cabeza et al., 2018).

The educational attainment of people is, in theory, modifiable, but to qualify as such a factor the underlying assumption is that education has a *causal* effect on aging of brain and cognition. However, educational attainment may also be viewed as an outcome of preexistent differences in brain and cognition between people. In addition, associations between educational attainment and aging of brain and cognition may stem from external factors, such as parental socioeconomic status, that have independent effects on both aging and education. In the course of the chapter, we will highlight evidence (or the absence of evidence) for these alternative accounts.

How could educational attainment – in theory – impact brain and cognitive aging? This “million-dollar question” is often answered by arguments that education contributes to building a “reserve”. Based on such a reserve, some individuals are assumed to be better than others at tolerating or compensating for brain insults conferred by aging, damage, or disease. The reserve concept has a quite long history (Blessed et al., 1968; Rothschild, 1937), but more modern conceptualizations distinguish between a more passive *brain reserve* versus a more active form of *cognitive reserve*.

A higher *brain reserve* could mean that some individuals from childhood and onwards in life have bigger brains and/or a higher proportion of large neurons in the brain (cf. Katzman et al., 1988). Thereby, compared to individuals with less brain reserve, it will take longer for the high-reserve individuals to reach a threshold for functional impairment – even if their change functions (slopes) are similar (Fig. 1A; Y-axis displays a metric of brain reserve). If education has an effect on building a brain reserve, then people with higher educational attainment can be predicted to have bigger brain (and better cognitive performance) after the end of formal education, but thereafter show similar rates of change of brain and cognition as those with lower education. In other words, by having more brain tissue (higher intercept), a certain amount of brain atrophy is better tolerated in aging in high-reserve individuals as compared to low-reserve individuals before the atrophy translates into behavioral problems. (the threshold for cognitive impairment is reached at an older age for individuals with higher brain reserve).

The active form of reserve, *cognitive reserve* (Stern, 2009), emphasizes functional attempts at *compensating* for negative brain changes. By this view, individuals with a higher cognitive reserve can use alternative cognitive strategies to cope with the challenges induced by a certain level of brain atrophy, whereas individuals of the same age but with lower cognitive reserve will show cognitive impairment for a similar level of atrophy (Fig. 1B, left panel). Thus, high cognitive reserve is not considered to protect against age-related brain atrophy, but instead reduce the negative influence of such brain changes on cognition as compared to individuals with less cognitive reserve. It should be noted that it remains a possibility that higher education and, by inference, higher cognitive reserve will confer a cognitive advantage across age and atrophy levels (Fig. 1B, right panel). That is, in particular for more challenging cognitive tasks, it may well be that the advantage of high cognitive reserve is not only manifested when a certain atrophy

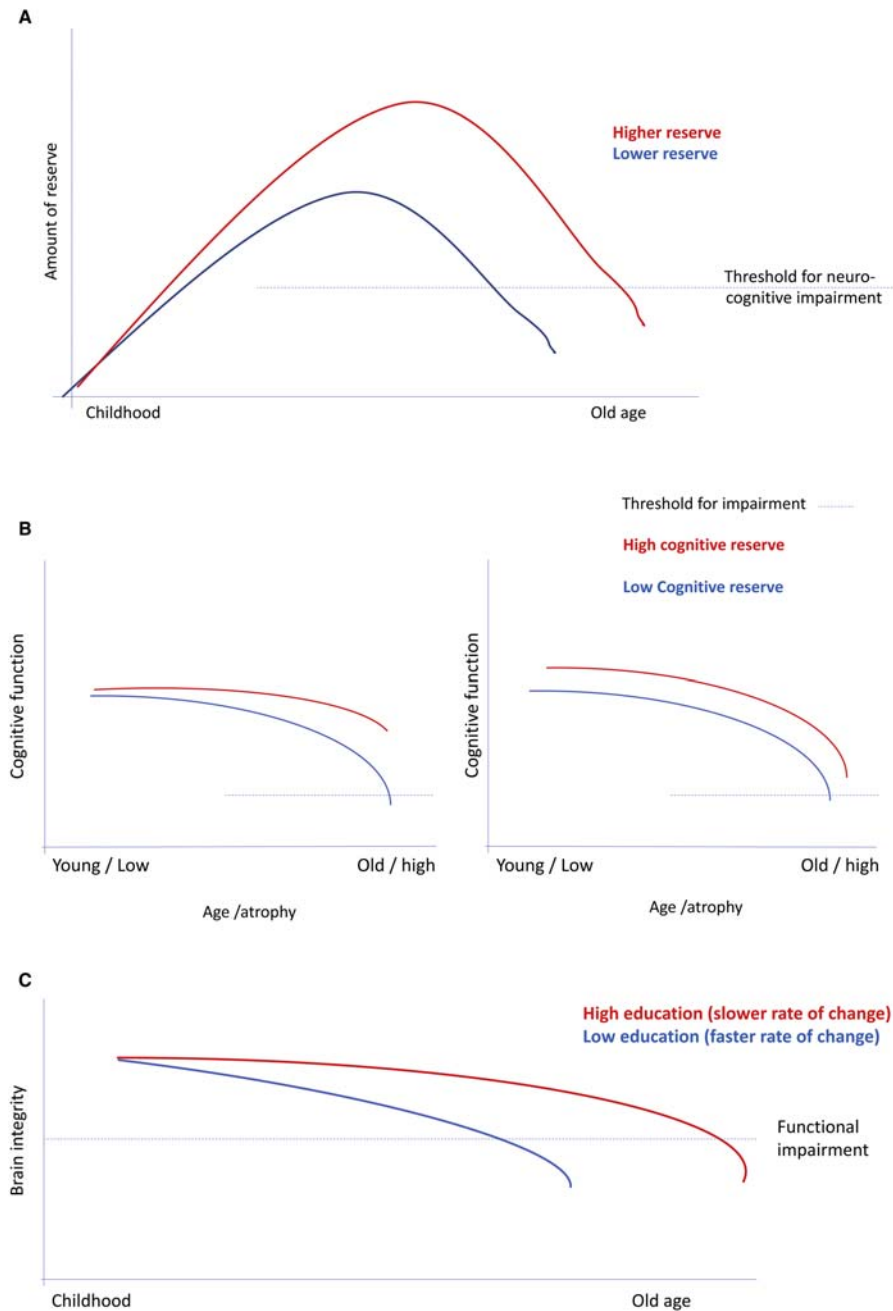


Fig. 1 Theoretical models of how education could relate to brain and cognitive status in older age (see text for details).

level has been reached. Still, the net result is the same – a given level of atrophy translates earlier into cognitive impairment for individuals with low education and reserve (i.e., the blue line crosses the threshold for impairment earlier in both the left and right panel of Fig. 1B).

The question of how education could impact brain and cognitive aging could – in theory – also reflect a situation where higher education actually slows the *rate* of brain aging. This scenario is hypothetically illustrated in Fig. 1C, where individuals with higher education (red line) have a slower rate of brain change than individuals with lower levels of education (blue line). Thus, this scenario, in contrast to those in Fig. 1A and B, holds that brain changes in aging may actually be slower for individuals with higher education. Thereby, at a given age, there will be less brain atrophy for individuals with higher education, which translates into a cognitive advantage.

Next, we turn to reviewing the state of the art regarding if and how educational attainment is associated with neurocognitive aging. The different scenarios outlined in Fig. 1A–C will feature in discussions of brain aging as well as cognitive aging, and in a concluding section we offer a synthesis and outline some issues for future studies.

Education and brain aging

Brain aging is an umbrella term that spans several changes within the central nervous system that tend to increase in frequency and magnitude as individuals age. Brain aging is examined at many different scales, ranging from the macroscopic whole-brain level to the microscopic sub-cellular level. Influences of education on human brain aging, in principle, could occur at any of these levels and they could also be expressed at multiple levels simultaneously. The majority of studies related individual differences in educational attainment to different kinds of data acquired by means of magnetic resonance imaging (MRI). MRI is a highly versatile tool that offers detailed examination of structural and functional brain integrity. Age-effects on brain structure and function are found for most MRI measures and regions, but with substantial heterogeneity with regard to region and shape of trajectories (Nyberg et al., 2022). In addition, there is growing consensus for marked heterogeneity within the older population, such that some individuals show a faster-than-average rate of brain aging whereas others present with a relatively well-maintained brain in older age (Nyberg et al., 2012). In view of such heterogeneity, numerous studies have explored genetic and lifestyle factors that may predict individual differences in brain aging (Nyberg et al., 2020).

Studies of educational attainment have yielded mixed results regarding if and how it relates to individual differences in brain aging. For example, Rzezak et al. (2015) compared older adults (around 70 years of age), who had less than 4 years of education, with a group with an average of 10 years of education. The results identified one frontal region with significantly greater gray matter volume in those with higher educational attainment, and another region showing a similar trend. Some other studies reported the opposite direction of results, with thinner cortex in high-compared to low-educated older adults (Pillai et al., 2012), other studies found that education is associated with higher cortical thickness in some areas but lower in other areas (e.g., Habeck et al., 2020), and still other studies found no relation of educational level on individuals' brain age (Dunas et al., 2021; but see also Steffener et al., 2016). Nevertheless, several studies have reported a similar results pattern as in Rzezak et al. (2015), with larger brain metrics in individuals with higher levels of education. A large-scale example of the latter comes from a study of 19,793 individuals in the UK Biobank (UKB) study (Nobis et al., 2019), which reported that higher education levels were associated with larger hippocampus volumes, albeit with a very small effect size (0.07 Standard Deviations; SD).

The studies mentioned above, as well as the majority of studies in the field more generally, based their conclusions on *cross-sectional comparisons*. That is, brain integrity of a group of older adults with high education was compared with that in a group of older adults with lower education. While a finding of a larger brain volume in the group with higher education might be interpreted to reflect reduced rates of brain aging, an alternative interpretation is that the difference is not specific to older age and aging *per se*. That is, it is quite possible that the brains of the high-education individuals were larger already in childhood and adulthood (cf. Cox et al., 2016). To be able to say more conclusively that individuals with higher education display slower brain aging, *longitudinal designs* (i.e., repeated measurements of the same individuals over time) is vital. Such designs provide information on within-person changes over time, and thus offer a stronger test of whether a factor like education has any influence on brain aging. To date, there are only limited relevant longitudinal data on this topic.

Raz et al. (2005) published a seminal study in the field in which five-year change in the regional volumes of healthy adults were analyzed. The statistical approach was latent difference score modeling, and the analyses involved 72 participants. The results revealed significant volume reductions in all examined brain regions (except the primary visual cortex, where little change was expected). In most of these regions, there were significant individual differences in rates of change. To some degree, those individual differences related to hypertension, but no support was obtained that higher levels of educational attainment were associated with reduced brain shrinkage.

In another study, Pillai et al. (2012) investigated 91 healthy controls and 140 individuals with Mild Cognitive Impairment (MCI) from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Analyses of one-year atrophy rates revealed a significant effect of diagnostic group (more atrophy for the MCI individuals), but the analyses provided no support for a relation between regional shrinkage and higher (18 or more years) or lower (13 or less years) educational attainment.

A final example of a longitudinal study of the influence of educational attainment on brain aging (MRI measures of brain atrophy) comes from a recently published paper (Nyberg et al., 2021). In a first step, regional samples from the Lifebrain project (N = 735; 1844 scans) were examined with education coded as a continuous variable. In good agreement with prior observations, longitudinal analyses revealed significant atrophy in several cortical regions and also in the hippocampus. Critically, however, level of education was not associated with rates of cortical or hippocampal change. In a second, replication step, similar analyses were conducted on data from the UKB (N = 1289; 2578 scans). As in the Lifebrain analyses, longitudinal analyses of the UKB data revealed marked atrophy in the cortex and in the hippocampus, but again educational attainment (college/university vs. non-university) was not associated with rates of change. Thus, based on 4422 observations and a longitudinal coverage up to 11.2 years, it was concluded that higher education does not slow down brain aging (see Fig. 2).

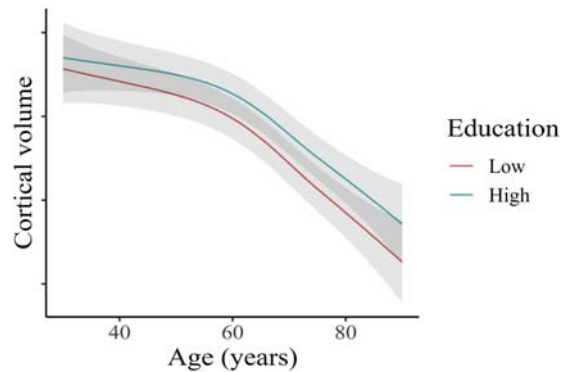


Fig. 2 Cortical volume decreases with age in both high and low educated people. Although cortical volume is slightly higher in those with high education, the rate of change is the same. Data source: Lifebrain and UK Biobank. Credit: Fredrik Magnussen.

Despite the lack of influence of educational attainment on brain change, it still remained a possibility that education would be associated with volume itself (i.e., an intercept but not slope effect). In both analyses of data from Lifebrain and UKB, some limited support for this hypothesis was obtained in that higher levels of education were found to relate to a larger cortical volume (Fig. 2). At the regional level, the highest between study overlap was located near the left central sulcus. In addition, some study-specific cross-sectional associations were observed in both studies. As suggested by the above mentioned Nobis et al. (2019) UKB study, analyses based on a greater number of observations will likely reveal additional intercept effects (e.g., in the hippocampus), but with limited effect sizes. Thus, modest education effects on brain volume may exist, but such a relation does not necessarily reflect a direct and causal influence of education on brain size but may instead be driven by shared genetic effects (cf. Grasby et al., 2020).

Regarding other metrics of brain aging, some of the above-mentioned studies included MRI measures of *white-matter integrity*. Raz et al. (2005) included assessment of changes in frontal and parietal white-matter integrity in their analyses, and no evidence was found for a relation between rates of change and education. Relatedly, Dunas et al. (2021) included various MRI white-matter images in their computation of “brain age” but found no relation to education. A study by Dufouil et al. (2003) examined MRI measures of small brain lesions (white-matter hyperintensities) and found a relation of higher white-matter lesion and lower cognitive functioning on some tests in individuals with low, but not high, education. The authors concluded that their findings indicate that education modulates the consequences of lesions on cognitive performance. However, as discussed above, as this was a cross-sectional study, it cannot be ruled out that the low-educated individuals with high lesion load and lower cognitive performance actually had a lower cognition already in younger age (for example due to genetic and/or lifestyle factors underlying vascular risk).

The structural brain measures considered so far, reflecting gray- and white-matter structural integrity, may most closely relate to the notion of passive brain reserve as discussed above (Fig. 1). Functional MRI (fMRI) offers a means to examine task-related brain activity, indirectly indexed by brain blood flow, while participants are given various cognitive challenges. In addition, fMRI allows measurement of resting-state connectivity, which might be seen as reflecting the status of the *functional connectome* in aging and disease (for a review of fMRI at rest and task, (see e.g., Raichle, 2010)). As such, fMRI may more directly relate to active cognitive reserve, and several relevant cross-sectional studies have been published (see e.g., Anthony and Lin, 2018; Steffener and Stern, 2012). Some studies specifically examined educational attainment in relation to fMRI responses, and some support has been found that education relates to fMRI measures (e.g., Chan et al., 2021; Kim et al., 2021). However, at present, convincing longitudinal evidence is lacking that individuals with higher education have altered fMRI response changes in aging compared with individuals with lower levels of education. Therefore, in view of demonstrations of fundamental differences between cross-sectional and longitudinal estimations of age-related fMRI responses (Nyberg et al., 2010), we conclude that the current state of knowledge remains unclear regarding if and how education impacts functional brain activity changes in aging.

Education and cognitive aging

A positive association between educational attainment and cognitive performance is remarkably robust across age, gender, historical times, and societies (Lovden et al., 2020; Opdebeeck et al., 2015; Strenze, 2007). Focusing on cognitive performance in older age, Opdebeeck et al. (2015) reported an average association of around $r = 0.3$ between general cognitive ability and educational attainment. In other words, an individual with a university degree is expected to have an advantage of roughly 4–5 IQ points relative to an individual with a high school diploma (i.e., 5 fewer years of education). It might be expected that education is more closely related to so called crystallized cognitive abilities, such as vocabulary and academic knowledge, than to fluid cognitive abilities, such as

abstract reasoning and processing speed. However, the available meta-analyses do not support this prediction (Opdebeeck et al., 2015). It could also be imagined that the association between educational attainment and cognitive ability is not linear. For example, further education may have diminishing effects on performance at higher levels of education. Although a few studies have reported trends for such non-linear effects, the association remains positive also at higher levels of educational attainment (Kobayashi et al., 2017; Myerson et al., 1998).

In contrast to the lack of evidence that would allow for causal interpretations of any effects of education on brain integrity, there are solid results that speak to this issue for cognitive performance as an outcome (Ceci, 1991; Ritchie and Tucker-Drob, 2018). For example, researchers in several countries have used policy changes (e.g., increasing years of mandatory schooling) as an instrument for estimating causal effects of early schooling. This is possible because there is often regional variation in the timing of when such policy changes have been implemented. Typically, cognitive performance is analyzed in large samples of individuals that are affected by the policy changes and who have taken part in assessments of performance during military conscription. A meta-analysis of such studies indicate that effects of schooling on performance are sizable (about 1–2 IQ points per year of additional schooling) and do not differ in magnitude between crystallized and fluid cognitive abilities (Ritchie and Tucker-Drob, 2018). Because a main purpose of schooling is to convey forms of declarative and procedural knowledge, effects on crystallized cognitive abilities are clearly expected. However, schooling may also improve cognitive strategies and test-taking skills (Ceci, 1991) or support more and better abstract thinking (Flynn, 1984), and through these pathways it could affect fluid cognitive abilities. In addition, effects of education on cognitive performance could emerge if education acts to protect young individuals from hazards of not being in school, such as harmful effects (e.g., stress) of work environment (Lager et al., 2017).

Nevertheless, a causal effect of education is unlikely to be the only origin of the association between educational attainment and cognitive performance. There is ample evidence to indicate that an individual's educational attainment is partly the outcome of socioeconomic factors (e.g., parental resources) that operate throughout development in childhood. It is quite possible that these factors also influence development of cognitive performance. In addition to being an outcome of chances in life, educational attainment is also influenced by an individual's own behaviors (e.g., cognitive performance, personality) and by how the individual interacts with the sociocontextual chances that are available (Tucker-Drob et al., 2013). Cognitive performance is perhaps the most important of these individual factors: Reported correlations between general cognitive ability and scholastic achievement average around $r \approx 0.50$ (Johnson et al., 2009; Mackintosh, 2011). Many of the abilities and traits (e.g., cognitive performance) that are important for performance in school, and therefore also for educational attainment, are heritable. This helps to clarify the background to why scholastic performance (Krapohl et al., 2014), as well as educational attainment (Cesarini and Visscher, 2017; Davies et al., 2016), also are heritable. That is, parts of the association between cognitive performance and education are of genetic origin.

The complex interactions among abilities, sociocontextual factors, and education during development lead to an association between educational attainment and cognitive performance that remains into older age. Therefore, educational attainment is an important factor that influences when critical thresholds (e.g., functional impairments) are reached in older (see Fig. 1). It is not surprising, then, that educational attainment also is a powerful risk factor for dementia diagnosis (Sharp and Gatz, 2011; Xu et al., 2016). But is educational attainment also associated with change (i.e., the slopes) in cognitive performance in aging? The evidence for such an association is much more unconvincing than that observed for associations between educational attainment and levels of performance. In one study, the authors (Lipnicki et al., 2017) analyzed harmonized data from several longitudinal studies conducted in 12 countries (Australia, Brazil, France, Greece, Hong Kong, Italy, Japan, Singapore, Spain, South Korea, the United Kingdom, and the United States: total $n = 42,170$). The average estimates across the studies indicated that educational attainment had no noteworthy associations with changes in cognitive function. A weak association between education and change in memory (one year of education comes with less than 0.01 SD per decade of memory decline) was reported in another major recent analysis of over 11 000 older individuals taking part in the Survey of Health, Aging, and Retirement in Europe (Cadar et al., 2017). A comprehensive meta-analysis by Seblova et al. (2020) confirmed the general impression from these major studies. The point estimates of the average relationship between one additional year of education and cognitive changes were not statistically significant and they were very small: less than 0.001 SD per decade for episodic memory and reasoning, and less than 0.002 SD per decade for general intelligence and processing speed. Even if we take these effect sizes at face value we must conclude that they are of an unimportant size: Over two decades of decline (e.g., from age 60 to 80 years), the general magnitude of these associations comes with an effect on level of cognitive performance in older age (at age 80 years) that is of less than 1/10 of the importance as the association between education and level of cognitive performance in younger adulthood that was discussed earlier (Lovden et al., 2020). Reviews of the literature on changes in cognitive performance in individuals that are later diagnosed with dementia come to the same conclusion – there is no substantial evidence linking educational attainment to changes in cognitive performance (Lovden et al., 2020) (Fig. 3).

It is however noteworthy that there is a substantial heterogeneity across studies in the association between educational attainment and change in cognitive performance (Seblova et al., 2020). It may be that some substantial association may appear in some populations or societies. For example, effects of educational attainment on cognitive change may appear in samples of individuals who are poorly educated. No such pattern was however apparent in the reviewed meta-analysis. The roles of factors such as retirement, birth cohort effects, gender, and retest effects, that differ as a function of educational attainment, are also not fully understood. Studies also differ greatly in their treatment (inclusion vs. exclusion) of individuals with dementia, but studies aiming to address the effect of this treatment find no noteworthy differences (Lipnicki et al., 2017). It also possible that associations between educational attainment and cognitive change depend on equality in access to higher education. However, the meta-analysis by Seblova et al. (2020) showed no association between effect sizes and the Gini coefficient (a common measure of a country's

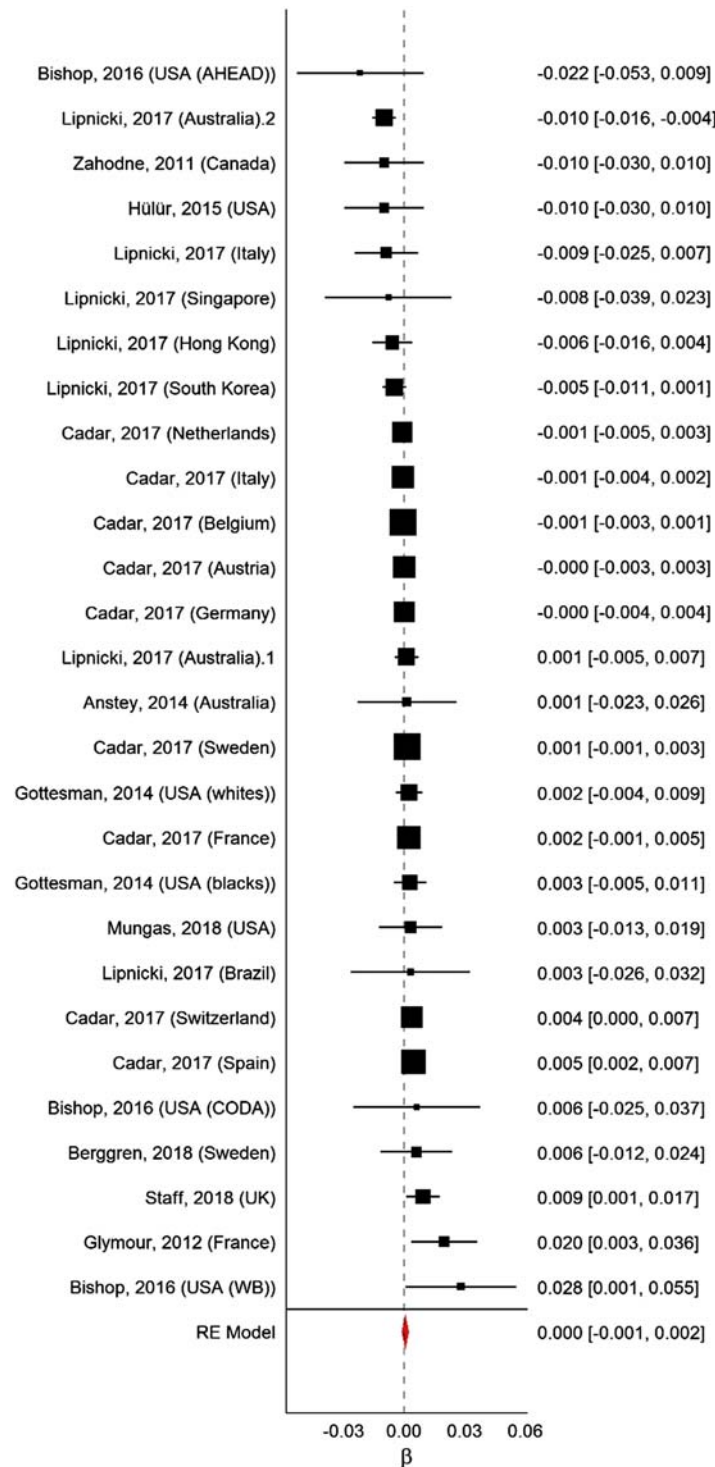


Fig. 3 Outcomes from studies examining the association between age-related change in episodic memory and education. The figure shows correlations (squares) reported for individual studies. The area of each square is proportional to the sample's weight in the meta-analysis, and the horizontal lines represent confidence intervals (see also the right-hand column for exact numbers). The red diamond shows the weighted effect of the meta-analysis. A positive correlation indicates that higher educational attainment is associated with slower cognitive decline. Adapted from Seblova, D., Berggren, R., Lovden, M., 2020. Education and age-related decline in cognitive performance: systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* 58, 101005. <https://doi.org/10.1016/j.arr.2019.101005>.

inequality). Thus, if the heterogeneity of results across studies is genuine, then the sources of it remain unknown. In addition, any substantial association between education and change in cognitive performance could only be an exception, and not the rule.

In summary, the evidence available to date reveals that the association between educational attainment and cognitive changes in older age is typically unreliable and much less important for life in older age than the association between education and level of cognitive performance that arises during childhood.

Conclusion and outlook

Educational attainment displays a similar relationship to both aging of the brain and aging of cognitive performance. Education is robustly associated with level of cognitive performance, and to some extent with brain integrity, but there is no substantial evidence that neurocognitive aging, in terms of individual changes over time, is much affected by peoples' formal education in early life. The empirical pattern is not in line with theoretical ideas suggesting that education conveys a cognitive reserve (e.g., individual resources such as knowledge and cognitive strategies) that can be mobilized to uphold high levels of cognitive performance in aging. The pattern is also inconsistent with theoretical speculations suggesting that the health benefits that education is related to (e.g., better health care and loss risk behaviors such as smoking) significantly impact aging of the brain (and therefore cognition). Rather, the pattern of results is more consistent with a passive reserve account: Educational attainment is primarily related to late-life cognitive functioning by playing a role in the developmental processes that lead to differences in brain and cognition during early adulthood, and these differences persist into older age. Having said this, we note that the influence of education attainment on the association between aging-related changes in brain and cognition remains largely unaddressed so far. Estimating this effect in the future will be the critical test that can distinguish between the validity of alternative theories that have been proposed, with the cognitive reserve theory predicting such an effect, while the brain maintenance theory would predict an absence of such an effect. Recent efforts to harmonize large-scale longitudinal studies of changes in brain and cognition may allow, for the first time, well-powered tests of this fundamental question.

Acknowledgments

L.N. is supported by a *scholar* grant from Knut and Alice Wallenberg's (KAW) Foundation. M.L. was supported by a grant from FORTE (2018–00791).

References

- Anthony, M., Lin, F., 2018. A systematic review for functional neuroimaging studies of cognitive reserve across the cognitive aging spectrum. *Arch. Clin. Neuropsychol.* 33 (8), 937–948. <https://doi.org/10.1093/arclin/acx125>.
- Blessed, G., Tomlinson, B.E., Roth, M., 1968. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br. J. Psychiatry* 114, 797–811.
- Cabeza, R., Albert, M., Belleville, S., Craik, F.I.M., Duarte, A., Grady, C.L., et al., 2018. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat. Rev. Neurosci.* 19 (11), 701–710. <https://doi.org/10.1038/s41583-018-0068-2>.
- Cadar, D., Robitaille, A., Clouston, S., Hofer, S.M., Piccinin, A.M., Muniz-Terrera, G., 2017. An international evaluation of cognitive reserve and memory changes in early old age in 10 European countries. *Neuroepidemiology* 48 (1–2), 9–20. <https://doi.org/10.1159/000452276>.
- Ceci, S.J., 1991. How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Dev. Psychol.* 27, 703.
- Cesarini, D., Visscher, P.M., 2017. Genetics and educational attainment. *Npj Sci. Learn.* 2 (1). <https://doi.org/10.1038/s41539-017-0005-6>.
- Chan, M.Y., Han, L., Carreno, C.A., Zhang, Z., Rodriguez, R.M., LaRose, M., et al., 2021. Long-term prognosis and educational determinants of brain network decline in older adult individuals. *Nat. Aging* 1 (11), 1053–1067. <https://doi.org/10.1038/s43587-021-00125-4>.
- Cox, S.R., Dickie, D.A., Ritchie, S.J., Royle, N.A., Valdés Hernández, M.d.C., Maniega, S.M., et al., 2016. Associations between education and brain structure at age 73 years, adjusted for age 11 IQ. *Neurology* 87, 1820–1826.
- Davies, G., Marioni, R.E., Liewald, D.C., Hill, W.D., Hagenaars, S.P., Harris, S.E., Deary, I.J., 2016. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol. Psychiatr.* 21 (6), 758–767. <https://doi.org/10.1038/mp.2016.45>.
- Dufouil, C., Alperovitch, A., Tzourio, C., 2003. Influence of education on the relationship between white matter lesions and cognition. *Neurology* 60 (5), 831–836. <https://doi.org/10.1212/01.wnl.0000049456.33231.96>.
- Dunas, T., Wahlén, A., Nyberg, L., Boraxbekk, C.J., 2021. Multimodal image analysis of apparent brain age identifies physical fitness as predictor of brain maintenance. *Cerebr. Cortex* 31 (7), 3393–3407. <https://doi.org/10.1093/cercor/bhab019>.
- Flynn, J.R., 1984. The mean IQ of Americans: massive gains from 1932 to 1978. *Psychol. Bull.* 95, 29–51.
- Grasby, K.L., Jahanshad, N., Painter, J.N., Colodro-Conde, L., Bralten, J., Hibar, D.P., et al., 2020. The genetic architecture of the human cerebral cortex. *Science* 367 (6484). <https://doi.org/10.1126/science.aay6690>.
- Habeck, C., Gazes, Y., Razlighi, Q., Stern, Y., 2020. Cortical thickness and its associations with age, total cognition and education across the adult lifespan. *PLoS One* 15 (3), e0230298. <https://doi.org/10.1371/journal.pone.0230298>.
- Johnson, W., Deary, I.J., Iacono, W.G., 2009. Genetic and environmental transactions underlying educational attainment. *Intelligence* 37 (5), 466–478. <https://doi.org/10.1016/j.intell.2009.05.006>.
- Josefsson, M., de Luna, X., Pudas, S., Nilsson, L.G., Nyberg, L., 2012. Genetic and lifestyle predictors of 15-year longitudinal change in episodic memory. *J. Am. Geriatr. Soc.* 60 (12), 2308–2312. <https://doi.org/10.1111/jgs.12000>.
- Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., et al., 1988. Clinical, pathological, and neurochemical changes in dementia - a subgroup with preserved mental status and numerous neocortical plaques. *Ann. Neurol.* 23 (2), 138–144. Retrieved from <Go to ISI>://A1988M308400005.

- Kim, Y., Kim, S.W., Seo, S.W., Jang, H., Kim, K.W., Cho, S.H., et al., 2021. Effect of education on functional network edge efficiency in Alzheimer's disease. *Sci. Rep.* 11 (1), 17255. <https://doi.org/10.1038/s41598-021-96361-0>.
- Kobayashi, L.C., Glymour, M.M., Kahn, K., Payne, C.F., Wagner, R.G., Montana, L., et al., 2017. Childhood deprivation and later-life cognitive function in a population-based study of older rural South Africans. *Soc. Sci. Med.* 190, 20–28. <https://doi.org/10.1016/j.socscimed.2017.08.009>.
- Krapohl, E., Rimfeld, K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.B., et al., 2014. The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proc. Natl. Acad. Sci. U. S. A.* 111 (42), 15273–15278. <https://doi.org/10.1073/pnas.1408777111>.
- Lager, A., Seblova, D., Falkstedt, D., Lovden, M., 2017. Cognitive and emotional outcomes after prolonged education: a quasi-experiment on 320 182 Swedish boys. *Int. J. Epidemiol.* 46 (1), 303–311. <https://doi.org/10.1093/ije/dyw093>.
- Lin, F.V., Wang, X., Wu, R., Rebok, G.W., Chapman, B.P., Alzheimer's Disease Neuroimaging, Initiative, 2017. Identification of successful cognitive aging in the Alzheimer's disease neuroimaging initiative study. *J. Alzheim. Dis.* 59 (1), 101–111. <https://doi.org/10.3233/JAD-161278>.
- Lipnicki, D.M., Crawford, J.D., Dutta, R., Thalamuthu, A., Kochan, N.A., Andrews, G., et al., 2017. Age-related cognitive decline and associations with sex, education and apolipoprotein E genotype across ethnic groups and geographic regions: a collaborative cohort study. *PLoS Med.* 14 (3), e1002261. <https://doi.org/10.1371/journal.pmed.1002261>.
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S.G., Huntley, J., Ames, D., et al., 2017. Dementia prevention, intervention, and care. *Lancet*. [https://doi.org/10.1016/S0140-6736\(17\)31363-6](https://doi.org/10.1016/S0140-6736(17)31363-6).
- Lovden, M., Fratiglioni, L., Glymour, M.M., Lindenberg, U., Tucker-Drob, E.M., 2020. Education and cognitive functioning across the life span. *Psychol. Sci. Publ. Interest* 21 (1), 6–41. <https://doi.org/10.1177/1529100620920576>.
- Mackintosh, N., 2011. *IQ and Human Intelligence*. Oxford University Press, Oxford, UK.
- Myerson, J., Rank, M.R., Raines, F., Schnitzler, M.A., 1998. Race and general cognitive ability: the myth of diminishing returns to education. *Psychol. Sci.* 9, 139–142.
- Nobis, L., Manohar, S.G., Smith, S.M., Alfaro-Almagro, F., Jenkinson, M., Mackay, C.E., Husain, M., 2019. Hippocampal volume across age: nomograms derived from over 19,700 people in UK Biobank. *Neuroimage Clin.* 23, 101904. <https://doi.org/10.1016/j.nicl.2019.101904>.
- Nyberg, L., Boraxbekk, C.J., Sorman, D.E., Hansson, P., Herlitz, A., Kauppi, K., et al., 2020. Biological and environmental predictors of heterogeneity in neurocognitive ageing: evidence from Betula and other longitudinal studies. *Ageing Res. Rev.* 64, 101184. <https://doi.org/10.1016/j.arr.2020.101184>.
- Nyberg, L., Lovden, M., Riklund, K., Lindenberg, U., Backman, L., 2012. Memory aging and brain maintenance. *Trends Cognit. Sci.* 16 (5), 292–305. <https://doi.org/10.1016/j.tics.2012.04.005>.
- Nyberg, L., Magnussen, F., Lundquist, A., Baare, W., Bartres-Faz, D., Bertram, L., et al., 2021. Educational attainment does not influence brain aging. *Proc. Natl. Acad. Sci. U. S. A.* 118 (18). <https://doi.org/10.1073/pnas.2101644118>.
- Nyberg, L., Pudas, S., 2019. Successful memory aging. *Annu. Rev. Psychol.* 70, 219–243. <https://doi.org/10.1146/annurev-psych-010418-103052>.
- Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., et al., 2010. Longitudinal evidence for diminished frontal cortex function in aging. *Proc. Natl. Acad. Sci. U. S. A.* 107 (52), 22682–22686. <https://doi.org/10.1073/pnas.1012651108>.
- Nyberg, L., Walhovd, K.B., Fjell, A.M., 2022. Neurobiological aging and memory: average patterns and individual differences. In: Kahana, M.J., Wagner, A.D. (Eds.), *Oxford Handbook of Human Memory*. Oxford University Press, Oxford.
- Opdebeeck, C., Martyr, A., Clare, L., 2015. Cognitive reserve and cognitive function in healthy older people: a meta-analysis. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 1–21. <https://doi.org/10.1080/13825585.2015.1041450>.
- Pillai, J.A., McEvoy, L.K., Hagler Jr., D.J., Holland, D., Dale, A.M., Salmon, D.P., et al., 2012. Higher education is not associated with greater cortical thickness in brain areas related to literacy or intelligence in normal aging or mild cognitive impairment. *J. Clin. Exp. Neuropsychol.* 34 (9), 925–935. <https://doi.org/10.1080/13803395.2012.702733>.
- Raichle, M.E., 2010. Two views of brain function. *Trends Cognit. Sci.* 14 (4), 180–190. <https://doi.org/10.1016/j.tics.2010.01.008>.
- Raz, N., Lindenberg, U., Rodrigue, K.M., Kennedy, K.M., Head, D., Williamson, A., et al., 2005. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebr. Cortex* 15 (11), 1676–1689. <https://doi.org/10.1093/cercor/bhi044>.
- Ritchie, S.J., Tucker-Drob, E.M., 2018. How much does education improve intelligence? A meta-analysis. *Psychol. Sci.* 29 (8), 1358–1369. <https://doi.org/10.1177/0956797618774253>.
- Ross, C.E., Wu, C., 2006. The links between education and health. *Am. Socio. Rev.* 60, 719–745.
- Rothschild, D., 1937. Pathological changes in senile psychoses and their psychobiologic significance. *Am. J. Psychiatr.* 93, 757–788.
- Rzezak, P., Squarzon, P., Duran, F.L., de Toledo Ferraz Alves, T., Tamashiro-Duran, J., Bottino, C.M., et al., 2015. Relationship between brain age-related reduction in gray matter and educational attainment. *PLoS One* 10 (10), e0140945. <https://doi.org/10.1371/journal.pone.0140945>.
- Seblova, D., Berggren, R., Lovden, M., 2020. Education and age-related decline in cognitive performance: systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* 58, 101005. <https://doi.org/10.1016/j.arr.2019.101005>.
- Sharp, E.S., Gatz, M., 2011. Relationship between education and dementia: an updated systematic review. *Alzheimer Dis. Assoc. Disord.* 25, 289–304.
- Steffener, J., Habeck, C., O'Shea, D., Razlighi, Q., Bherer, L., Stern, Y., 2016. Differences between chronological and brain age are related to education and self-reported physical activity. *Neurobiol. Aging* 40, 138–144. <https://doi.org/10.1016/j.neurobiolaging.2016.01.014>.
- Steffener, J., Stern, Y., 2012. Exploring the neural basis of cognitive reserve in aging. *Biochim. Biophys. Acta* 1822 (3), 467–473. <https://doi.org/10.1016/j.bbadis.2011.09.012>.
- Stern, Y., 2009. Cognitive reserve. *Neuropsychologia* 47 (10), 2015–2028. <https://doi.org/10.1016/j.neuropsychologia.2009.03.004>.
- Strenze, T., 2007. Intelligence and socioeconomic success: a meta-analytic review of longitudinal research. *Intelligence* 35 (5), 401–426. <https://doi.org/10.1016/j.intell.2006.09.004>.
- Tucker-Drob, E.M., Briley, D.A., Harden, K.P., 2013. Genetic and environmental influences on cognition across development and context. *Curr. Dir. Psychol. Sci.* 22 (5), 349–355. <https://doi.org/10.1177/0963721413485087>.
- Xu, W., Tan, L., Wang, H.F., Tan, M.S., Tan, L., Li, J.Q., et al., 2016. Education and risk of dementia: dose-response meta-analysis of prospective cohort studies. *Mol. Neurobiol.* 53 (5), 3113–3123. <https://doi.org/10.1007/s12035-015-9211-5>.

Lifelong learning beyond initial schooling

Karen Evans^a and Natasha Kersh^{a,b}, ^a UCL Institute of Education, Social Research Institute, London, United Kingdom; and ^b UCL Institute of Education, Department of Education, Practice and Society, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	520
European contexts for lifelong learning development	522
The interplay between international and national agendas for change	522
“Informal learning” recognition is pivotal for lifelong learning strategies	523
Analytic perspectives and theoretical approaches question dominant discourse	524
Human capital development that focuses on the learning individual	524
Human capital development that focuses on the social organization of learning	524
Learning as social process that focuses on the learning individual	525
Learning as social practice, focusing on the social organization of learning	525
Reframing debate and action beyond “formal, non-formal or informal” learning	526
Conclusions	527
References	527

Introduction

Issues and perspectives in learning beyond initial schooling can be understood in terms of a triadic conception of lifelong learning, which is animated by the interplay between three dimensions: (1) educating for a more highly skilled and knowledgeable working population; (2) supporting personal development that activates inquiry and explores what life has to offer; (3) creating conditions for a stronger and more inclusive society (Aspin et al., 2012).

Furthermore, two analytical perspectives are pertinent in considering issues and perspectives. The first focuses on the social organization of learning. This perspective emphasizes adjustments to changed conditions in work and society and the need to address dysfunctionality in existing educational provision; the need for “flexibility,” and recognition of the prior and informal learning of adults. It recognizes the workplace and community as significant settings for learning, debating the characteristics of “learning organizations” or “communities of practice.” The second analytical perspective focuses on learning as a biographical process—the learning individual. It is predicated upon the multiple sources of evidence of the cognitive social, psychological and cultural dimensions of adult learning that have challenged deficit models that equate aging with cognitive losses (Formosa, 2019) and generated new insights into the expansion of adults’ capabilities through purposeful activity at all ages. This perspective recognizes the heterogeneity of adults’ learning, the conditions and opportunities that shape learning; the social structuring of the life course and the accumulation of risk and inequalities. At the same time, it keeps in view personal agency (Evans et al., 2013), reflexivity in learning, the significance of social networks and role of multiple forms of capital (O’Tuama, 2020, 2022).

The triadic conception is holistic in nature and represents a departure from approaches that translate the aims of personal development, higher skills and knowledge and enhanced social participation into competing modes and models. Interdependencies of the three dimensions of higher skills and knowledge, personal development and social participation are critical to the triadic conception; “if any one of the dimensions fails, the center cannot hold” (Evans, 2021, p. 350). A triadic conception is evidence-informed. We know, for example, that if education for higher skills and knowledge in the workforce is also accompanied by growing inequalities between the knowledge rich and knowledge poor, the conditions for the strong and inclusive society are undermined (Green, 2011). For the center to hold, wider distribution of higher skills and knowledge in the adult population is necessary, but is that sufficient for sustaining a stronger society? According to the triadic conception, not so; for a stronger and more inclusive society higher average levels of knowledge and skills have also to be accompanied by personal attributes of inquiry that support long-term learning engagement in changing circumstances: “strategies for workforce development and channels for enhanced social participation will not be sustainable in the longer term unless they are powered by curious and capable people who have been supported in their personal development” (Evans, 2021, p. 350).

Keeping in view the two analytical perspectives, this chapter explores the potential for fresh thinking on lifelong learning according to new realities in the interplay of the learning individual and the social organization of learning. Questions of how lifelong learning systems can recognize and support the learning of adults of all ages throughout, and beyond, extended working lives are central to international developments in the field and the possible answers vary according to the dominant discourse that is framing the debate at any particular time. Competing discourses co-exist, and the fact that one is more dominant at any time does not mean that the others have lost their power to explain or persuade.

A discourse that has been dominant in earlier times has taken as its point of departure that the exercise and preservation of democracy is of predominant importance in real life (we go to war in its name), that real life is about active and engaged citizenry, and the proper relationship between education and real life has to be understood in terms of lifelong engagement with the ideas and

practices of democracy. It is, for example, about educating ourselves collectively to exercise control over social impact and advances of technology. This version was important in western Europe in the early part of the 20th Century, at the time of the foundation of many civic universities and the international sweep of emancipatory discourses which are equal and opposite to the globalization discourses which are sweeping the world we inhabit a century on. If we take a different starting point for our analysis, in which the flowering of human potential is of predominant importance, then the logic of the learning individual begins to define the proper relationship between education and real life. *Learning to Be* (Fauré, 1972) defined the humanistic aim of “Lifelong Education” as “the fulfillment of man, through the flexible organization of different stages of education”, a process which should last the whole life for individuals (Fauré, 1972) and not just be tacked onto school or university for a privileged or specialized few. Husen (1974), in his well-known articulation of the concept of a learning society, emphasized not only “second chance” conceptions of lifelong learning but also the process by which untapped potential comes to be seen as the norm, not the exception, in the population. International bodies such as UNESCO used the humanistic version of lifelong learning to provide master concepts for educational planning, but their use represented worthy ideals with relatively little to show in the way of tangible advances in access to lifelong learning opportunities.

The global political consensus on lifelong learning which was constructed at the end of the 20 century was apparent in the Delors Report (Delors, 1996) the EU Competitiveness and Economic Growth Agenda and in statements of intent of OECD, the G8 countries and the Delhi Declaration. The EU Memorandum on Lifelong Learning, 2000 set out an agenda for lifelong learning for the first two decades of the 21st Century, based on analyses that emphasized the ways in which the nature of work and working lives are changing; the need to reform dysfunctional systems rooted in the norms of the past; the search for innovation, and “new skills for new jobs.”

Take up of the ideas of learning through the life-course as necessary to sustain livelihoods in times of critical social change fundamentally reshaped the discourses of lifelong learning. The basis for analysis of the “proper relationship between education and real life” came to be defined by the knowledge-driven economy, in which the relationship between education and “real life” is that “the more we learn, the more we earn.” We are engaged in a lifelong competition for livelihoods, learning for a living.

The role of government is recast as providing the means and encouragement for people to take responsibility for their own “employability,” as it is claimed that national governments can no longer guarantee employment in a global competitive environment, as governments seek to maximize their share of high skill high wage jobs in a global economy. The political consensus did not hold for long. The projections of OECD in 1997 showed growing awareness of the consequences of reducing learning, specifically lifelong learning, to investment in exploitable economic capital. Contrasts were drawn between the sense of control “over their own lives and their society” which comes from successful experience of education, with the isolation of the excluded. An equal investment in “social capital” was called for, but is the attempt to balance these two types of capital, social and economic, inherently contradictory? UNESCO, in an attempt to regain its role of values-based leadership in lifelong learning, has established the Global Monitoring of adult learning, with periodic Global Reports on Adult Learning and Education (GRALE, 2016, 2019) and international conferences to advance the Bélem Action Framework for adult education as a human right and as a means for transforming individuals, communities and societies as well as a means for economic development. According to UNESCO, although the perspective of economic development has been dominant, both the empowering and the instrumental rationales are needed in our societies (UNESCO, 2009a). Since then, lifelong learning has been incorporated in Sustainable Development Goals (SDGs), optimistically viewed as the beginning of a fourth generation in which lifelong learning is actively promoted in the context of international development (Elfert, 2019).

The analyses of governmental organizations and international agencies have concerned themselves principally with the social organization of institutions of learning, and the ways in which they need to be reorganized around new priorities to respond to social and economic changes. In this process the discontinuities of change are often exaggerated, while the continuities of disadvantage are under-estimated. New forms of learning are superimposed on old forms that nevertheless retain much of their original power. Social institutions, old and new, continue to interlock to shape typical life courses, that biographies that may still be typified according to social position. As Alheit and Dausien (2002) have argued, policies which aim to balance economic and social capital must take the “learning individual” much more seriously to avoid becoming trapped in the contradictions which lie at their heart. As well as continual reorganization of the institutions of learning, we need to try to understand better the reflexive ways in which people’s lives are shaped, bounded or change direction as they engage with these education, labor market and workplace institutions.

Much attention has been given to the impact of globalization and to the forms of modernization that are connected to globalizing tendencies. Yet modernization is neither uni-linear nor inevitable. There are strong differences rooted in local and national cultures which engage in different ways with globalizing influences.

The argument concerning the effects of globalization on education and training is that governments everywhere are conceding control to consumers and markets (leading to commodification of education). They are also being forced to concede control upwards to international organizations, and downwards to regional organizations. In this process, national education systems can start to lose their distinctive features as they begin to converge on a common instrumental norm. Social pluralism generates a growing diversity of needs which are met, increasingly, in privatized and individualistic ways. The logical conclusion is that it could eventually become impossible to transmit national cultures through education, as skill formation progressively displaces citizen formation as an educational goal, and “the collective project [of national education] is eclipsed by individualistic aspirations and norms” (Green, 1997, p. 4). Yet in moving from international discourses on lifelong learning to implementable policies at the national level, a view of learning as culture and lifestyle, a way of life, remains paramount in many societies (Fadzil, 2013).

European contexts for lifelong learning development

At least four models, analytically defined, can be identified in Europe. While all societies contain elements of all four models, it can be argued that one, or at most two, of the models have characterized, in broad terms, the context for lifelong learning development in any specific country.

The first model is characterized by strength of the “non-formal” sector. Where early school leaving generates inequalities in adult life, and aggregate skills levels and productivity levels are relatively low, the non-formal sector often plays an important role in supporting employment and families, potentially offering learning opportunities for young people and adults who have experienced barriers in formal settings of teaching and training. A strong culture of self-employment, especially in the countryside, will also often provide family support. Lifelong learning policies are directed toward mitigating problems of exclusion and disadvantage while putting in place supporting structures for adult learning that can ensure consistent access to learning opportunities and entitlements.

The second model is market-led, with learning arrangements that are close to labor markets, making work-integrated pathways possible. In such models, diversity in modes of schooling generates inequalities. Adult learning participation rates are often high, with measures designed to keep adults close to the changing labor markets and learning is often seen as ideally situated in the “real world” of work. Weaknesses include insecure pathways, in which private and public providers compete and initiatives come and go. New programmes and bridging schemes can be easily inserted into such a structure because it is so flexible. The flexibility also means that the pathways to be taken are insecure and, for those with the least advantages, the possibility of dropping out again remains a constant risk. In the market-driven structure the spread of incomes is large and tends to widen. This means that many become trapped in low-paid jobs, some of which require only minimal competences and provide little job satisfaction. Entitlements to wider forms of learning in adult life prioritize basic skills development.

“Dual” enterprise-school structures characterize the third model, built around secure structured education and training pathways that are associated with high aggregate skill and productivity. But inflexible pathways are associated with lower levels of participation in employment-related lifelong learning. The system does not easily allow for deviating pathways, which means that adults are constantly directed back to “square one” if they want to pursue a new occupational direction. High levels of participation in adult education build on traditions of supporting general education and integrating migrant populations.

The fourth model is characterized by the dominance of strong school-based structures. Egalitarian schooling, together with strong labor protection and regulation policies and redistributive effects from welfare systems, ensure relative income equality. High skills and productivity are further promoted through active labor market policies. High levels of participation in adult education build on a tradition of supporting general education. But barriers encountered by young people in making school-to-work transitions and weak education/industry links are often identified as impediments to effective implementation of lifelong learning policies. Educational “drop-out” becomes a problem under unfavorable conditions of the labor market, requiring the introduction of special initiatives to bring people back into the system in early adult life.

In practice most countries exhibit increasing hybridity in their approaches (Verdier, 2018). Societal trajectories are complex and even more so when abrupt changes have taken place whose ramifications take place over decades. This has applied most notably in the Eastern European and post-communist states in which multiple hybrids, built on organized markets and educational universalism, are being re-shaped through the interplay of internal socio-political logics and supra-national influences.

The interplay between international and national agendas for change

The key messages in lifelong learning policy espoused globally through international organizations (such as EU, UNESCO, OECD, ILO) emphasize over-arching priorities such as new basic skills for all, as primary skills for changing labor markets; increased investment in human resources; innovation, including shifts toward “user-oriented” or learner-centered learning systems; valuing learning, particularly recognizing informal and non-formal learning; the need for improved guidance and information for both young people and adults. Migrations of refugees have further highlighted the potential value of lifelong learning in recognizing previously acquired capabilities and integrating newcomers (Slowey and Schuetze, 2020).

In Europe, these international priorities have been interpreted and translated in EU lifelong learning policies in ways that place a strong accent on different forms of adults’ participation, political, social and economic.

Recent research on adult education (Kersh et al., 2021) compared country-specific approaches in policy development across Europe and neighboring countries. The research revealed both convergences and divergences in the development of adult education and lifelong learning policies and practices cross-nationally. Common trends in the provision of adult education that aims to foster different forms of participation and inclusion have been influenced by some common global challenges as well as EU policies and research in the area of adult education and lifelong learning. Specifically, policies and practices that aim to tackle social exclusion and promote social, political and economic participation of adults became centrally important to the European lifelong learning agenda. An economic rationale for lifelong learning as a means to secure and sustain livelihoods in times of critical social change has driven inclusion strategies focusing on employability and labor market skills. Tackling unemployment and equipping adults with the skills required by the contemporary labor market have been among most influential trends of lifelong learning, defining policy and practice responses across Europe and its neighboring countries. Strong interdependencies between economic, social and political forms of participation have also characterized some country-specific responses to contemporary challenges (Evans, 2009; Kersh et al., 2021). Policies and practices with a strong focus on the socio-economic dimension, that is, integration through entering

employment, promote versions of active citizenship implicitly rather than explicitly. According to this view, equipping adults with employability skills facilitates their chances to take a more active role in their social settings through their engagement with employment.

In practice governments everywhere pick and choose what to focus on and, more importantly, how to interpret concepts and priorities generated elsewhere, whether in the international literature or international organizations, in the context of their own systems, cultures and political stances. Governments implement measures that are consistent with their dominant socio-economic models and assumptions. For example, many Nordic interpretations see lifelong learning through a social welfare lens. Indeed, it was Nordic influences that secured a broadening of the initial [European Commission \(2000: 3\)](#) definition of Lifelong Learning (LLL) as “all purposeful learning activity undertaken on an ongoing basis with the aim of improving knowledge, skills and competence” to emphasize the range of perspectives, beyond employment, that can be taken on these activities: “all learning activity undertaken throughout life, with the aim of improving knowledge skills and competences within a personal, civic, social and/or employment-related perspective” ([European Commission, 2001](#), p. 33).

Moreover, the ways in which priorities are interpreted and acted upon in member (nation) states, depend crucially on relationships with existing forms of education and social provision. Socio-economic inclusion has also underpinned various education schemes set up by national governments, with the aim of integrating newcomers, such as refugees and migrants. However, as highlighted by Kersh et al.’s research ([2021](#)), the approaches toward inclusion are often driven by national political priorities rather than by the multiple needs of adults, specifically those, in vulnerable situations.

Relationships with existing traditions of adult education are always an issue, as attention has focused progressively, at level of international organizations, on how the outcomes of “informal” lifelong learning can be better recognized and certified to support adults across lengthening (working) lives and shift educational systems toward lifelong learning objectives.

“Informal learning” recognition is pivotal for lifelong learning strategies

Informal learning has been described as not organized, unintentional, and is often referred to as “learning by experience” or simply as “experience” ([OECD, 2010](#), p. 22). It relates to the capabilities that people can learn from a range of experiences, such as education, workplace activities and learning, family experience, community work ([Eraut, 2004](#)). The approaches toward conceptualizing informal learning (and informal skills) have been commonly based on the contention that significant capabilities have been acquired as a result of prior learning and, more generally, prior experiences. The development of knowledge and skill is often tacit in nature and thus individuals do not necessarily recognize to what extent they use their previously acquired skills in their current practices or draw on their previously acquired knowledge ([Evans, Kersh, and Sakamoto, 2004](#)). The benefits of capturing such skills for the purpose of recognition and accreditation have been recognized as significant measures of progress for different kinds of learners, including those who find themselves in vulnerable positions, such as refugees and migrants. As [Hart and Chakroun \(2019, p. 13\)](#) put it “the volume of migration has created a need for mechanisms and agreements to assist both those making claims for recognition and those responding to these claims by giving employment or granting admission to further learning” (2019, p. 13). Recognition of learning outcomes irrespective of how and where they originate thus potentially contributes to fostering equity and strengthening access to education and to the labor market for various disadvantaged groups, including young people and older learners ([OECD, 2010](#)).

The strategic processes by which lifelong learning can be promoted at international level, for take up in widely differing systems and cultures, are most comprehensively exemplified by the European Union development process. Validation (accreditation) of forms of learning characterized as informal and non-formal became cornerstones to plans for overcoming barriers and resistances to the realization of adequate lifelong learning systems across Europe. The 1995 Commission *White Paper on Teaching and Learning* ([European Commission, 1995](#); [Dukekot et al., 2005](#)) is often cited as a formative paper in the development of policies to promote validation of learning outside and beyond formal systems at the European level. Another influential document, the 2001 Communication of the European Commission: *Making the European area of lifelong learning a reality* ([European Commission, 2001](#)) further emphasized the significance of validation of informal learning as well as recognition of learning through the life course. This policy development was strongly supported by the 2002 Copenhagen declaration, which defined the European strategy for enhanced cooperation in vocational education and training ([European Council and European Commission, 2002](#)). Subsequently, many lifelong learning policies and initiatives have been characterized by their focus on validation of “informal learning” and accreditation of outcomes of learning irrespective of where and how that learning has occurred. As noted by [Villalba and Bjørnåvold \(2017\)](#) various European programmes (e.g. Leonardo da Vinci, Grundtvig and Socrates) played a key role in testing solutions and disseminating experiences of validation and further supporting the setting up of validation arrangements in various European countries.

At national level, policies on validation and accreditation have exhibited both some common and diverging trends in setting out and implementing these initiatives. The 2008 adoption of the European qualifications framework (EQF) has been highlighted as one of key developments, facilitating the process of validation across European countries. As observed by [Villalba and Bjørnåvold \(2017, p. 75\)](#) “the long-term effect of this, directly influencing validation, is the more systematic promotion of learning outcomes at national level.” Following this, The European Council 2012 Recommendation on validation of non-formal and informal learning ([Council of the European Union, 2012](#)) has become an important landmark in the development of policies and practices of validation of non-formal and informal learning in Europe. The recommendation stressed the importance of “offering individuals the opportunity to demonstrate what they have learned outside formal education,” where recognition of prior learning means the

validation of learning outcomes, "whether from formal education or non-formal or informal learning, acquired before requesting validation" (Council of the European Union, 2012). Within this broad interpretation, European countries have translated and adapted this concept to their specific contexts. Although the specific features of the process of validation of informal and prior learning vary between different country contexts and the extent to which they have been grafted onto existing approaches and long-standing cultural commitments, some common broad principles have been identified. As summarized by Villalba-Garcia (2019, pp. 45–46) the processes of validation of informal/non-formal learning across different countries, share the features that they are voluntary processes based on the identification of learning outcomes irrespective of how and where the occurred. Their validation can lead to different types of qualification, distinguished according to criteria of demonstrated knowledge, skill, attitudes and competence.

The many classifications of approaches reveal multiple dimensions of variation. The topology developed by Damesin et al. (2014) focuses on the interplay between public learning validation policies and the mode of involvement of the social players, specifically identifying countries where validation of learning, irrespective of how and where the individual has learned, are integral to development of lifelong learning system (Denmark, Finland, France and Portugal); countries where validation of learning outside formal systems is primarily a matter for local and regional initiatives, and targets particular sectors and/or professions (Italy, Spain) countries where more or less active trade union intervention in terms of training goes hand in hand with an incomplete institutional system for validation of learning outside formal systems, whose development is not considered to be a priority: Germany, Romania and Poland.

Examples of ways in which international defined priorities and policy drivers play out at national and local levels illustrate how path dependencies generate contrasting results "on the ground," often reinforcing structural inequalities. The discussion so far yields insights that are in some respects specific to variations in the European context and in other respects resonate with experiences captured in global reports of UNESCO, ILO. The Global Report on Adult Learning and Education (GRALE 4) report (GRALE, 2019) has compiled evidence of the persistence of deep-rooted inequalities in participation in adult learning and education, and calls for major changes in approach, backed by adequate investment, in the context of the Sustainable Development Goals.

As alternative approaches and structures become embedded and institutionalized, quality frameworks have become important for their long-term sustainability. The international organizations have sought to promote transparency and to create the conditions for co-operation and exchange across borders, in recognition of diversity and changing landscapes at national and local levels (EU Renewed agenda for lifelong learning 2012–2014). EU priorities have a labor market and stakeholder orientation, contrasting with that of the UNESCO (2010, p. 8) Bélem framework which emphasizes the need to "foster a culture of quality in adult learning" including the enrichment of learning environments, motivation and empowerment of individuals and communities and the role of public authorities in ensuring effectiveness and the rights of users of services.

Analytic perspectives and theoretical approaches question dominant discourse

Beyond, and independently from, the concerted efforts of international organizations to overcome obstacles to systemic development of lifelong learning by concrete means, the two analytic perspectives—on the social organization of learning and the learning individual, respectively—reveal the fundamental need for much more inclusive accounting of the capabilities of adults and the circumstances in which learning arises, across and beyond extended working lives.

These analytic perspectives can be considered in tandem with two broad avenues for inquiry into lifelong learning and human development. The first theoretical approach focuses on the individual knowledge and skills that are part of human capital, and on learning and development as human capital accumulation. The second theoretical approach sees human capabilities as shaped by context and social practices. These lenses focus attention, in different ways, in which the learning and development is both embedded in and stimulated by personal and work situations in adult life. In four quadrants, Fig. 1 outlines the assumptions and beliefs in these different analytic perspectives from the perspective of an individual focus on learning and a social perspective on learning.

Human capital development that focuses on the learning individual

In the context of lifelong learning this perspective translates into policy arguments which assume that developing competences will bring the greater likelihood for an economically sustainable life. In this view, the absence of competences portrayed as key, core or foundational, can impact negatively on an individual's economic and social opportunities. Inquiries into the efficacy of lifelong learning have to track individual journeys over time to demonstrate sustainability of working life and they also have to address the complexity of the relationship between adult learning and adult experiences (Reder and Bynner, 2009).

Human capital development that focuses on the social organization of learning

Although most of the literature on human capital examines the relationship between individual human capital and individual earnings and life-experiences, human capital development is recognized as having broader social dimensions in health, welfare, social participation and poverty reduction (McMahon and Oketch, 2013). It is argued that higher levels of human capital have spill-over effects. In other words, a society with large numbers of educated and skilled people is believed to benefit in ways which go beyond

	Human capital approaches	Social process approaches
Focus on the learning individual	A Learning clearly defined sets of skills and competences, the absence of which can impact negatively on an individual's economic and social opportunities.	B Emphasises socially positioned life course and biographical narrative; embedded and embodied learning.
Focus on social organisation of learning	C Emphasises shaping and organising education and adult learning 'provision' for socio-economic ends such as increased productivity, social mobility.	D Emphasises environments and affordances for learning; communities of social practice; informal and 'everyday' learning.

Fig. 1 Broad avenues for inquiry into lifelong learning and human development. Adapted from [Evans \(2011\)](#).

the higher incomes and better life-chances of the skilled individuals themselves. The people who work alongside skilled people will also be more productive, and there will organization-wide changes which would not occur without the presence of skilled individuals. This insight applies across all forms of organization from the company to the voluntary organization, from the firm to the local community group, orchestra, choir or athletics club. Conversely the spiraling down of skills and abilities in successive low-quality assignments may also have negative spillover effects, as yet unresearched.

Learning as social process that focuses on the learning individual

In this perspective learning journeys in adult life are defined in terms of biography and socially embedded learning, recognizing that adult learning takes on a genuinely active, unpredictable quality when it draws on resources in "borderless" ways, in critical life situations and turning points in the life-course ([Chisholm, 2008](#); [Biasin and Evans, 2019](#)). Furthermore, people's beliefs in their ability to change their situation by their own efforts, individually or collectively, are significant for the development of their capabilities. These beliefs change and develop over time and according to life experiences. Notions of borderlessness view learning as becoming more extensive (lifelong), more specialized (life-near), more differentiated (life-wide), more flexible, more individualized and more contingent ([Chisholm 2008:142](#)).

Learning as social practice, focusing on the social organization of learning

Learning as a social practice emphasizes the social context of knowledge and capabilities that are learnt and developed over time, with attention to important milieus for development of capabilities and immersion in cultures. Situated learning theories enrich the concept of learning space; emphasizing socialization into a wider community of practice that involves membership, identity formation and expertise development through participation in the activities of social practices ([Lave and Wenger, 1991](#); [Kersh, 2015](#)). In this view the group is seen as important rather than the individual, and people learn through shared practices of the community, with individuals actively influencing one another's learning and enculturation through a process of co-participation ([Billett, 2004](#)). Learning thus becomes the expansion of human capabilities through the lived experiences and social relations of activity.

Researching the learning and development of adults takes each of these analytic and theoretical perspectives to the limits of explanatory and exploratory power. New conceptual approaches are showing how adult learning and the exercise of contextualized preferences in adult life involve much more than a series of rational choice calculations ([Evans et al., 2013](#); [Evans, 2021](#)), while cultural signs and tools such as "qualifications" and "standards" ([Wertsch 1998](#)) mediate conceptualization of skills, intelligence and capability and what people believe is possible for them and for others to achieve.

As argued previously, recognizing and certifying individuals who have demonstrable capabilities but are not in possession of the official forms of recognition that are needed for them to practice or progress has been a longstanding priority in the development of lifelong learning systems. The search for workable mechanisms that can connect the learning individual with the social organization of learning has been driven internationally by the need for concrete instruments that can translate lifelong learning principles into action. The definition and recognition of "non-formal and informal learning" have been central to the process.

To function effectively, new approaches to recognizing and certifying learning and the concepts, cultural signs and tools associated with them have to become embedded in the wider eco-systems of opportunities. Of course, concepts and associated mechanisms for recognizing and certifying adult learning already existed in different forms in the dominant model(s) outlined above

and, in many cases, fundamental questions have been generated as approaches reliant on definitions of informal and non-formal learning and driven by the international organization priorities entered the field and started to change the rules of the game. Multiple contextual and cultural factors come into play, as the comparisons of the next section reveal.

Reframing debate and action beyond “formal, non-formal or informal” learning

The debates about formal, informal and non-formal learning which have structured development at the international level, are connected in complicated ways with conceptual understandings of adult learning and also with culturally embedded educational development at national and local levels. A team of scholars (Harring et al., 2016) set out to explore different understandings and approaches internationally, considering variations in context and concept as well as implementation efforts through cases exemplifying the diversity of Asian approaches (focusing on Japan and Bangladesh) and North American perspectives (from US and Canada), as well as variations in Europe. The results reveal not only contextual variations but are also indicative of social-ecological interdependencies of learning and practice.

There is no Japanese language term to distinguish informal learning from non-formal and formal education, although the idea of learning throughout life and from all experiences is culturally embedded. Informal learning was imported as a foreign term into the educational lexicon from the late 1970s, as a precursor to the Japanese Government’s phased implementation, from the late 1980s, of policy measures to construct a lifelong learning society. The policy process adopted UNESCO definitions incorporating informal learning as the unintentional learning of knowledge skills attitudes and perspectives occurring in family, community and workplace spaces as well as through multi-media and the internet (Sawano, 2015).

In the contrasting Asian context of Bangladesh, Shohel (2015) has also observed that there was no place for acknowledging informal learning in mainstream education, a scenario that international organizations have explicitly challenged in the promotion of lifelong learning systems. Informal, non-formal and formal learning are portrayed holistically as “three types of currents flowing in same ocean of education – overlapping and very hard to separate” (Shohel, 2015, p. 277). When learning is viewed as a holistic process, elders, parents, employers and peers are seen to help individuals in analyzing incidental everyday experiences in ways that can make tacit knowledge explicit, with entertainment, gossip and chats, fables and storytelling, indigenous knowledge transfer, religious congregation all identified in a rich cultural appreciation of the scope and nature of learning beyond schooling.

Japan and Bangladesh, despite their obvious socio-economic and political differences, have thus both adopted terminologies from international organizations for policy purposes. These terminologies have been superimposed, in both examples, on culturally embedded and pervasive notions of the lifelong importance of unintentional forms of learning.

North American contributions trace a shift in thinking toward more elaborated understandings of the nature of adult learning (Watkins and Marsick, 2015), moving on from preoccupations with place and setting to an increasing focus on the open-ended, collaborative, social and contextually dependent learning that evolves and transforms over time and adapts according to resources and technology. From a Swedish perspective, Norqvist et al. (2015) have embraced a plurality of approaches, arguing that this is the last generation to experience a clear definition of when learning is “school” or is “not school.” In an age of online and digital learning, it is important to ask whether and to whom these distinctions matter in life worlds of the learning individual; Norqvist et al. (2015) advance the case for an integrative view of learning in a digital media world, based on the notion of pedagogical attitude. Further to this, (Singh, 2015), p. 20) has argued that learning must not be seen as “dichotomous and discrete categories, but rather as continuous elements within the ‘learning continuum’.” Less attention to formal/informal definitional debates and more to the combination and interplay of the “currents” of learning, potentially advances deeper insights into the multiple ways in which systems of representation are built up in ways which govern modes of thinking and acting in the world.

The definitional debate about formal, non-formal and informal learning is arguably becoming redundant. The definitions have been comprehensively used to challenge dominant assumptions that equated worthwhile learning with formal education. They have been used to erode “artificial and outmoded barriers” (Faure, 1972, p. 189) and have, in that respect, been fundamental to the mission of international organizations to create conditions more conducive to the development of lifelong learning systems. This process has often been linked, in a rather generalized way, with neo-liberal reforms.

The development of the French system provides a counter example to prevailing assumptions that systems for recognition at national levels are unequivocally driven and shaped by neo-liberal reform agendas. Lessons from the French model are of considerable interest at the present time, since they represent successful introduction of the validation of experiential learning outcomes into a national context and culture of relative academic elitism. To remedy perceived “social justice deficit” and weakness in equality of opportunity, legislation in 2002 established the right for all working individuals to earn a recognized diploma or professional qualification through the accreditation of prior experience. According to Rémy and Merle (2014) the implementation of APL can now be seen as a “small revolution” in France. The 2002 Validation of Acquired Experiences (VAE) allowed “informal routes to qualifications that are formal,” assessed against officially recognized standards used throughout the educational systems. The concept of informal learning had little currency in France and was not influential in this VAE development, according to Werquin (2015) who argues that it is rooted in 200 year-old commitments to the principle that for workers to be productive and to remain so they have to be trained throughout life. While there is body of literature on learning that is not formal, it is “relatively scarce” (Werquin 2015, p. 185) compared to an extensive literature on experience, experiential learning and “*acquis d’apprentissage*” which is distinctively French in origin and has been pioneering in approach. Moreover, the concepts of informal, non-formal and formal learning are most often used to by international organizations rather than researchers, leading Werquin to comment that these are not

airtight categories and to observe that only institutions that have an immediate interest in using the concepts for concrete actions are involved in promoting the definitions.

The VAE system targets the less qualified yet competent adults, aiming to make visible the learning outcomes that they have achieved through any type of activities that involve learning and that they can demonstrate in practice. The VAE system requires preparation of two portfolios: one portfolio to support candidature by establishing eligibility; the second portfolio for the main assessment, against criteria, by a panel. Accrediting institutions have to set out how candidates should demonstrate skills and knowledge, how they should be guided in building solid “candidature” and how juries will be chosen to assess. This removes barriers to candidates accessing institutions and comes to rely on how adults can gain the quality of experiences. The candidates are assessed by interview and practical tests. The candidates prepare their narratives with guidance, analyzing and selecting from their work history and informally acquired knowledge which might have been gained through community activities, experience in other cultures, the pursuit of personal interests or other means. The challenges that face the system, which has substantially increased take up since it started in 2002, are securing parity of esteem for adults who have taken this route. Even though they have achieved exactly the same qualification and standards as those who have qualified through the conventional routes there is, nevertheless, an enduring, culturally embedded lack of belief in “true” equivalence of alternatives to conventional modes that resonates in many other contexts. Many people who are already qualified add to their qualifications through this system, but the fact that half of all beneficiaries have been at the first level of the French NQF indicates success in reaching previously unqualified adults. The qualification they gain has dual currency—both for further progression to the next level of education and in securing positions in the labor force.

The example of France has been chosen here as it contains insights for systems and societies seeking means and modes for supporting and recognizing adult learning at a national level. The system provides an incomplete solution that works partially for a proportion of adults. It is both instrumental and empowering, within its own limits. It is home grown, neither neo-liberal in conception nor universal in its applicability. It is both enabled and limited by the its embeddedness in tradition and is set to evolve in the context of changing ecologies of learning and practice.

Beyond working life applications, it is important to note that the broader recognition of adult learning includes action to promote environments and learning spaces which are rich in opportunities. Public non-profit cultural spaces may not have overt educational purposes—but have powerful effects on the cultural languages of life and people’s understandings of themselves in the world (Halttunen, Koivisto, and Billett, 2014).

Pedagogies centered on fresh understanding of how and why adults learn (Carré, 2020; Norqvist et al., 2015) can enhance the ways adults make meaning and are activated in changing social landscapes. They are part of the ecologies of learning and practice (Barnett and Jackson, 2020) that are interdependent not only with technological developments but also with local responses to global issues such as inequality, migration, climate change and pandemic responses.

Conclusions

The search for ways of recognizing the learning of adults beyond initial schooling lies at the heart of lifelong learning and its long-term evolution as the organizing concept for education systems of the future. The definitional debate about formal, non-formal, and informal learning has probably served its purposes, challenging dominant assumptions that equated worthwhile learning with formal education. The concepts became central to the efforts of international organizations to create conditions more conducive to the development of lifelong learning systems. They have endured for this reason but are superseded, we argue, both conceptually and in a context in which upcoming generations are unlikely to equate learning with formal education in the way of previous generations and will not make the same assumptions about how, why and where people learn.

In the modern world, ideas for better connecting system worlds and the life-worlds of learning individuals can now focus compellingly on learning as a lifelong process which links rather than separates generations. Adults moving in changing social landscapes have strivings, aspirations and scope for action to change lives for the better, individually and collectively. An holistic, triadic vision for lifelong learning that calls for people to look for new options beyond the choices laid before them also has to recognize that human development is profoundly affected by macro-economic conditions, institutional structures, social background, gender and ethnicity as well as acquired attributes and resources available through family, communities networks. Bringing the lifelong perspective of the learning individual more centrally into debates about the social organization of education and learning means recognizing changing ecologies of learning and social practice and rethinking connections between system worlds and life worlds, over time. This process also highlights the challenges for triadic conceptions of lifelong learning in supporting the development of educated attributes, abilities to read the world critically and protecting environments that enable young people and adults to learn mutual responsibility in the collective free spaces of life.

References

- Alheit, P., Dausien, L., 2002. The double face of lifelong learning: two analytical perspectives in a learning revolution. *Stud. Educ. Adults* 34, 3–22.
- Aspin, D., Chapman, J., Evans, K., Bagnall, A., 2012. *Second International Handbook of Lifelong Learning*, vol. 1. Springer, Dordrecht.
- Barnett, R., Jackson, N., 2020. *Ecologies for Learning and Practice*. Routledge, Abingdon.

- Biasin, C., Evans, K., 2019. Agency, identity and learning at turning points in women's lives: a comparative UK-Italian analysis. *Eur. J. Res. Educ. Learn. Adults* 10 (1), 47–63. <https://doi.org/10.3384/rela.2000-7426.rela915>. ISSN 2000-7426.
- Billett, S., 2004. Co-participation at work: learning through work and throughout working lives. *Stud. Educ. Adults* 36 (2), 190–205. <https://doi.org/10.1080/02660830.2004.11661496>.
- Carré, P., 2020. Pourquoi et comment les adultes apprenent. Dunod, Paris.
- Chisholm, L., 2008. Re-contextualising learning in second modernity. *Res. Post-Compul. Educ.* 13, 139–147. <https://doi.org/10.1080/13596740802141253>.
- Council of the European Union, 2012. Council recommendation on 20 December 2012 on the validation of non-formal and informal learning. Official Journal of the European Union, C 398/01. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2012:398:0001:0005:EN:PDF>. (Accessed 9 January 2022).
- Damesin, R., Fayolle, J., Fleury, N., Malaquin, M., Rode, N., 2014. Challenges, Actors and Practices of Non-formal and Informal Learning and its Validation in Europe. European Trade Unions Institute, Brussels.
- Delors, J., 1996. The Treasure within: Report to UNESCO of the International Commission on Education for the 21st Century. UNESCO, Paris.
- Duvekot, R., Schuur, K., Pulusse, J., 2005. VPL it is! Valuation and validation of prior learning as a new perspective for life-long learning. In: The Unfinished Story of VPL: Valuation and Validation of Prior Learning in Europe's Learning Cultures. Foundation EC-VPL & Kenniscentrum EV, Utrecht, pp. 1–10.
- Elfert, M., 2019. Lifelong learning in Sustainable Development Goal 4: what does it mean for UNESCO's rights-based approach to adult learning and education? *Int. Rev. Educ.* 65, 537–556. <https://doi.org/10.1007/s11159-019-09788-z>.
- Eraut, M., 2004. Informal learning in the workplace. *Stud. Cont. Educ.* 26 (2), 247–273.
- European Commission, 1995. Teaching and Learning: Towards the Learning Society. Office for Official Publications of the European Commission, Luxembourg.
- European Commission, 2000. A Memorandum on Lifelong Learning [Commission Staff Working Paper, 30.10.2000 SEC (2000) 1832]. European Commission, Brussels. https://arhiv.acs.si/dokumenti/Memorandum_on_Lifelong_Learning.pdf. (Accessed 9 January 2022).
- European Commission, 2001. Making a European Area of Lifelong Learning a Reality. Communication from the Commission, 21.11.01 [COM (2001) 678 Final]. European Commission, Brussels. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0678:FIN:EN:PDF>. (Accessed 9 January 2022).
- European Council and European Commission, 2002. The Copenhagen declaration. https://www.cedefop.europa.eu/files/copenhagen_declaration_en.pdf. (Accessed 9 January 2022).
- Evans, K., 2009. Learning, Work and Social Responsibility: Challenges for Lifelong Learning in Global World. Springer, Dordrecht.
- Evans, K., 2011. Perspectives and lenses. In: Wolf, A., Evans, K. (Eds.), *Improving Literacy at Work*. Routledge, Abingdon.
- Evans, K., 2021. What's new about lifelong learning – and what does it mean for working lives. In: Malloch, M., Cairns, L., Evans, K., O'Connor, B.N. (Eds.), *The Sage Handbook of Learning and Work*. Sage, London.
- Evans, K., Kersh, N., Sakamoto, A., 2004. Learner biographies: exploring tacit dimensions of knowledge and skills. In: Rainbird, H., Fuller, A., Munro, A. (Eds.), *Workplace Learning in Context*. Routledge, London/New York, pp. 222–241.
- Evans, K., Schoon, I., Weale, M., 2013. Can lifelong learning reshape life chances? *Br. J. Educ. Stud.* 61 (1), 25–47. <http://www.tandfonline.com/doi/abs/10.1080/00071005.2012.756163#.VsJN2n2LTnl>.
- Fadzil, M., 2013. Lifelong learning in Malaysia: from Buzzword to Blueprint. In: *Lifelong Learning and Employability*, ASEM Lifelong Learning No. 03, pp. 22–23.
- Fauré, E., 1972. *Learning to Be: The World of Education Today and Tomorrow*. UNESCO, Paris.
- Formosa, M., 2019. *The University of the Third Age and Active Ageing*. Springer, Cham.
- GRALE III, 2016. Global Report on Adult Learning and Education 2016. UNESCO Institute for Lifelong Learning. <https://uil.unesco.org/adult-education/global-report/third-global-report-adult-learning-and-education-grale-3>. (Accessed 9 January 2022).
- GRALE IV, 2019. Global Report on Adult Learning and Education 2016. UNESCO Institute for Lifelong Learning. <https://unesdoc.unesco.org/ark:/48223/pf0000372274>. (Accessed 9 January 2022).
- Green, A., 1997. *Education, Globalization and the Nation State*. Macmillan, London.
- Green, A., 2011. Lifelong learning, equality and social cohesion1. *Eur. J. Educ.* 46, 228–243. <https://doi.org/10.1111/j.1465-3435.2011.01478.x>.
- Halttunen, T., Koivisto, M., Billett, S. (Eds.), 2014. *Promoting, Assessing, Recognizing and Certifying Lifelong Learning: International Perspectives and Practices*. Springer, Dordrecht.
- Harring, M., Witte, M.D., Burger, T. (Eds.), 2016. *Handbuch Informelles Lernen*. Beltz/Juventa, Weinheim/Basel.
- Hart, J., Chakroun, B., 2019. UNESCO world reference levels for lifelong learning: a tool for comparison and recognition of learning outcomes. In: Cedefop (Ed.), *Global Inventory of Regional and National Qualifications Frameworks 2019 Volume I: Thematic Chapters*. https://www.cedefop.europa.eu/files/2224_en_0.pdf.
- Husen, T., 1974. *The Learning Society*. Methuen, London.
- Kersh, N., 2015. Rethinking the learning space at work and beyond: the achievement of agency across the boundaries of work-related spaces and environments. *Int. Rev. Educ.* 61 (6), 835–851. <https://doi.org/10.1007/s11159-015-9529-2>.
- Kersh, N., Toivainen, H., Pitkänen, P., Zarifis, G.K. (Eds.), 2021. *Young Adults and Active Citizenship: Towards Social Inclusion through Adult Education*, vol. 26. Springer, Lifelong Learning Book Series. <https://doi.org/10.1007/978-3-030-65002-5>.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation (Learning in Doing: Social, Cognitive and Computational Perspectives)*. Cambridge University Press.
- McMahon, W.W., Oketch, M., 2013. Education's effects on individual life chances and on development: an overview. *Br. J. Educ. Stud.* 61 (1), 79–107. <https://doi.org/10.1080/00071005.2012.756170>.
- Norqvist, L., Leffler, E., Jahnke, I., 2015. Sweden and informal learning -towards integrated views of learning in a digital media world. In: Hanning, M., Witte, M.D., Burger, T. (Eds.), *Handbuch Informelles Lernen*. Beltz/Juventa, Weinheim/Basel.
- Ó Tuama, S., 2020. Learning neighbourhoods. *Andragoška Spoznanja* 26 (1), 53. <https://doi.org/10.4312/as.26.1.53-65>.
- OECD, 1997. *Overcoming Exclusion through Adult Learning*. Organisation for Economic Co-operation and Development (OECD), Paris.
- OECD, 2010. *Recognising Non-formal and Informal Learning: Outcomes, Policies and Practices*. OECD, ISBN 978-92-64-06385-3.
- O'Tuama, S., 2022. Five capitals in lifelong learning. In: Evans, K., et al. (Eds.), *The Third International Handbook of Lifelong Learning*. Springer, Cham.
- Reder, S., Bynner, J., 2009. *Tracking Literacy and Numeracy Skills: Findings from Longitudinal Research*. Routledge, New York.
- Rémery, V., Merle, V., 2014. French approaches to accreditation of prior learning: practices and research. In: Halttunen, T., et al. (Eds.), *Promoting, Assessing, Recognizing and Certifying Lifelong Learning: International Perspectives and Practices*. Springer, Dordrecht, pp. 265–280.
- Sawano, Y., 2015. Informal learning in Japan. In: Hanning, M., Witte, M.D., Burger, T. (Eds.), *Handbuch Informelles Lernen*. Beltz/Juventa, Weinheim/Basel.
- Shohel, M.M.C., 2015. Informal learning in Bangladesh – context, challenges and prospects. In: Hanning, M., Witte, M.D., Burger, T. (Eds.), *Handbuch Informelles Lernen*. Beltz/Juventa, Weinheim/Basel.
- Singh, M., 2015. *Global perspectives on recognising non-formal and informal learning why recognition matters*. Springer, Cham.
- Slowey, M., Schuetze, H., 2020. *Inequality, Innovation and Reform in Higher Education*. Lifelong Learning Series 25. Springer, Cham.
- UNESCO Institute for Lifelong Learning, 2009. *Global Report on Adult Learning (GRALE Report)*. UIL, Hamburg.
- UNESCO Institute for Lifelong Learning, 2010. *Harnessing the Power and Potential of Adult Learning and Education for a Viable Future. Bélem Framework for Action*. UIL, Hamburg. <https://unesdoc.unesco.org/ark:/48223/pf0000187789/PDF/187789qaa.pdf.multi>.
- Verdier, E., 2018. Europe: Comparing lifelong learning systems. In: Milana, M., Webb, S., Holford, J., Waller, R., Jarvis, P. (Eds.), *The Palgrave International Handbook on Adult and Lifelong Education and Learning* London Palgrave Macmillan. Springer Nature, pp. 461–483.

- Villalba, E., Bjørnåvold, J., 2017. Validation of non-formal and informal learning: a reality in europe? European centre for the development of vocational education and training (Cedefop). In: Global Inventory of Regional and National Qualifications Frameworks. CEDEFOP. <https://uil.unesco.org/lifelong-learning/qualification-frameworks/global-inventory-regional-and-national-qualifications-0>. (Accessed 9 January 2022).
- Watkins, K.E., Marsick, V.J., 2015. Informal learning in the U.S. In: Hanning, M., Witte, M.D., Burger, T. (Eds.), *Handbuch Informelles Lernen*. Beltz/Juventa, Weinheim/Basel.
- Werquin, P., 2015. Informal learning in France – recognition of informal learning outcomes for lowly qualified individuals. In: Hanning, M., Witte, M.D., Burger, T. (Eds.), *Handbuch Informelles Lernen*. Beltz Juventa, Weinheim/Basel.
- Wertsch, J., 1998. *Mind as Action*. Oxford University Press, New York.

Multiple and multi-dimensional educational and life transitions: conceptualization, theorization and XII pillars of transitions

Divya Jindal-Snape, Professor of Education, Inclusion and Life Transitions, School of Education and Social Work, University of Dundee, Dundee, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	530
Conceptualization of transitions	530
Defining successful transitions	532
Theoretical perspectives	532
Adapting theories borrowed from other contexts and disciplines	532
Developing research informed transitions theories	534
Educational and life transitions (ELT) theory	534
Multiple and multi-dimensional transitions (MMT) theory	534
Transition-to-trained-doctor (T3D) model	537
The 12 pillars of successful multiple and multi-dimensional transitions	539
Conclusion	541
References	541

Introduction

In this article, the focus is on conceptualization and theorization of transitions as there is a clear gap in literature about these (Jindal-Snape et al., 2020, 2021). Given that these form the very basis of good research, policy and practice, it is important to start a debate and discussion about transitions based on our 15 years plus transitions research across educational stages and life course. As educational transitions do not happen in a vacuum and are accompanied by, or trigger and are triggered by, other transitions, it is important to understand them as multiple transitions. For example, a student moving to higher education would experience a transition related to academic level of studies, relationship transitions such as leaving existing friends and creating new friendships, geographical transitions such as moving to a new city, developmental transitions and accompanying identity transitions from a young person to a young adult, living independently, to name but a few (Jindal-Snape and Rienties, 2016). Therefore, in the following sections we will consider educational transitions in the context of all these multiple transitions. The article starts with the conceptualization of transitions, followed by defining successful transitions, and theoretical frameworks used in transitions research. The educational stages and sectors covered include primary and secondary schools, transitions to and through university (international students, students identifying as LGBTQ+, and those with life limiting health conditions), and continuing professional development of higher stage medical trainees. At the end, some new concepts have been introduced and provide an insight into how successful transitions can be facilitated.

Conceptualization of transitions

Transitions can be defined as an ongoing process of psychological, social and educational adaptations due to changes in context, interpersonal relationships and/or identity, which can be simultaneously exciting and worrying for an individual and others in their lives, and which can require ongoing additional support (Jindal-Snape, 2016, 2018). The theories underpinning this conceptualization are Multiple and Multi-dimensional Transitions Theory (MMT Theory, Jindal-Snape, 2016), and Educational and Life Transitions Theory (ELT Theory, Jindal-Snape, 2016; Jindal-Snape and Ingram, 2013) which will be discussed later.

However, if you consider transitions literature, it is clear that there is no shared understanding among researchers, professionals and policy makers about the definition of transitions (Jindal-Snape et al., 2020). The way that researchers conceptualize educational and life transitions has important implications for their research designs, study findings and their interpretation, and implications emerging from that which are used to inform future research, policy and practice.

This view of the lack of conceptualization, in part, is based on a recent systematic literature review of primary-secondary transitions research articles written between 2008 and 2018. Most authors of the 96 articles included in the review had not clearly defined or operationalized transitions (Jindal-Snape et al., 2020, 2021). A team of four researchers had to establish their conceptualization through inferring from either words they had used, any theory they had cited or their methodology; there were multiple conceptualizations in some articles for that reason; a rigorous cross-check was undertaken to avoid inaccuracies in categorization. A total of 54 papers seemed to consider transition to be change with a focus on changes in pedagogical approach (e.g., Mackenzie et al., 2012), organization (e.g., Arens et al., 2013) or systems (e.g., Strnadova et al., 2016). This seems to be similar to health

education literature, in which authors have conceptualized transitions as change, for instance, in daily routines and social networks (Draus et al., 2014). Forty one papers conceptualized primary-secondary transitions as a normative life transition, within which 13 articles presented it as a normative life event (Neal et al., 2016); 20 articles conceptualized it as a normative period in educational journey (e.g., Brewin and Statham, 2011; Burchinal et al., 2008; Weiss and Baker-Smith, 2010); 11 papers referred to multiple transitions with multiple changes being experienced at the same time (e.g., Knesting et al., 2008; Lofgran et al., 2015; Serbin et al., 2013). Eight articles referred to transition as disruptive and highlighted the risk factors for children (Keay et al., 2015; Mackenzie et al., 2012; Maher and Maher, 2010), with six highlighting it as a time of curricular and relational discontinuity (e.g., Makin and Pellicano, 2017; Rainer and Cropley, 2015). The latter could be seen to be similar to the conceptualization of transition as change. Three articles used a life course conceptualization (Benner and Wang, 2014; Fortuna, 2014; Witherspoon and Ennett, 2011), three a rite of passage (Bailey and Baines, 2012; Rainer and Cropley, 2015; West et al., 2010), and three conceptualized it as a turning point (Andreas and Jackson, 2015; Langenkamp, 2010; Scanlon et al., 2016).

From most studies it is not clear whether the authors asked their participants about their conceptualization of transitions (Jindal-Snape et al., 2020). In one study, also in the context of primary-secondary transitions, professionals were asked about their conceptualization and it was found that they did not have a shared understanding within the same school or group of schools, with conceptualization ranging from “change” to “adaptation to change” (Jindal-Snape and Cantali, 2019). In another longitudinal study, the participants, who were higher stage medical trainees, were asked what transitions meant to them, and notably their views changed over time from transitions being about a change and being singular (at the start of the study), to a more complex, dynamic and multiple process (after 6–9 months).

Another interesting aspect that emerged from studying conceptualizations in existing primary-secondary transitions literature (Jindal-Snape et al., 2020) is that the discourse around transitions is primarily negative (60/96 articles; e.g., McIntosh et al., 2008; Mudaly and Sukhdeo, 2015; Peters and Brooks, 2016; Poorthuis et al., 2014; Serbin et al., 2013; Tso and Strnadova, 2017; Vasquez-Salgado and Chavira, 2014), with 25 that had a mixed discourse (but 15 more negative than positive), nine neutral and only two predominantly positive (Madjar and Chohat, 2017; Neal and Frederickson, 2016).

The articles with a negative or mixed discourse had used previous literature in their introduction to provide the context of transitions and their findings matched the previous literature. It is not clear whether the purpose of undertaking these studies was due to difficulties that had been noticed in schools and schools inviting researchers to undertake the study, leading to a negative discourse or whether the researchers’ thinking had been informed by prior research and had designed the study and questions in a way that they elicited primarily negative transition experiences. Interestingly, most authors had published a one-off study about transitions and therefore it was not possible to see any trends, including change, in the authors’ conceptualizations over the years. Further, it is possible that they had therefore relied on previous authors’ conceptualizations of transitions which had not been clearly articulated and/or used a primarily negative discourse.

This is concerning as the negative discourse in research literature might influence the views of professionals, families and learners, to the extent that they might anticipate a negative experience leading to unnecessary anxiety till they experience the reality of transitions, that might not be negative, over the life course. This is also set in the context of the societal environment where adults and family members (Jindal-Snape and Foggie, 2008) and popular media, such as story books, novels and movies also depict transitions as problematic with bullying from peers, stricter teachers, massive difference in academic level etc. However, in practice, educational transitions are in the main positive. For instance, transitions to higher education, especially for international students, is a result of hard work at school, opportunity to study in their chosen discipline, with international academics and being successful in getting prestigious and competitive scholarship (Jindal-Snape and Rienties, 2016). Therefore, it is important that the discourse is balanced and researchers investigate both, positive and negative, transition experiences (e.g., see Jindal-Snape and Cantali, 2019).

Another aspect worth considering is that the research designs that were used in the articles reviewed by Jindal-Snape et al. (2020), most had collected data just before and after the move to secondary school, capturing the immediate perceptions of children without considering their adaptation over time. This was also the case in healthcare literature as suggested by a literature review of 23 studies published between 1994 and 2004, with only one study using a longitudinal design and the rest focusing on a specific event in a short time frame (Kralik et al., 2006). From our four stage longitudinal study (final year of primary school to end of second year of secondary school), it was clear that different children adapted to change at different times and that their transitions were dynamic and multiple with different transition experiences at the same time (Jindal-Snape and Cantali, 2019).

Similarly, the way professionals and policy makers conceptualize transitions has a major impact on transitions practices. For instance, if transition is conceptualized as a one-off significant event such as starting secondary school, the transitions support is likely to be provided just before and/or after the move to secondary school. This can be highly problematic as we found in our four-stage study which collected data from students in the final year of primary school, and twice in the first year and once in the second year of secondary school (Jindal-Snape and Cantali, 2019). As mentioned earlier, we found that transitions were ongoing and dynamic, and the student participants reported that what they were excited about when in primary school continued to be good in secondary school. However, the concerns they had in primary school declined over time. More importantly, we found that a number of students experienced some problems at the start but adapted later; whereas for some students, problems emerged at the end of the first and at the start of the second year of secondary school with no transitions support available to them, apart from some students reporting that guidance staff were available to anyone experiencing any problems. This can be explained to some extent by the conceptualization of professionals in this primary-secondary school group. There was a lack of shared understanding about what transitions mean, with some conceptualizing them as a one-off event linked with major moves at particular

stages and time points, while others as an ongoing process of multiple changes, such as changes in the physical and social environment at school, with limited attention to adaptation over time (Jindal-Snape, 2018; Jindal-Snape and Cantali, 2019).

Defining successful transitions

Similar to the lack of clear definition of transitions, the term successful transitions is used without defining it, with the focus being on a dip in attainment and/or self-esteem.

In an educational context, there has been an emphasis on academic adjustment. A student was seen to have adjusted well if they had good attainment, engagement at school/university and regular attendance (e.g., Rice et al., n.d.). However, there has been some shift with a recognition that social and emotional adjustment at school/university are equally important, if not more so. These include connectedness and sense of belonging with the organization, and development of positive and mutually respectful relationships (Jindal-Snape and Rienties, 2016; Peters, 2010; Rice et al., n.d.).

Further, it is important to be mindful of the fact that successful transitions might mean something totally different, depending on whom we ask. For instance, from the perspective of the healthcare staff in Kirby et al. (2014) study on transitions to palliative care, successful transitions were seen to be an acceptance of medical futility on the part of the patient, pain and symptom management, and wellbeing of patient and family psychologically. The patients and families might have had a different perspective to this.

In Evangelou et al. (2008) study in an English primary-secondary school context, the professionals suggested that the indicators of successful transition was timely actioning of parental choices, children getting admission to their school of choice and reduction in placement appeals. However, children and parents reported that successful transitions meant making new friends, no concerns about settling in school, increase in engagement at the new school, adapting to the school routines, enhanced self-esteem and confidence, and curricular continuity.

In another study, for instance, at the time of the first interviews, two children reported being bullied after moving to the secondary school, which was having a very negative impact on their wellbeing (Jindal-Snape and Foggie, 2008). In the subsequent stage, one child reported that they had carried on but with similar issues whereas the other child and parent decided to leave that school to access an alternative curriculum where the child was happier. The question it raises is, which child was experiencing a successful transition? And who decides who is having a successful transition or not? Therefore, it is important that the person who is most affected by transitions is asked what successful transition for them might be.

Theoretical perspectives

The most prevalent theories in literature that have been used to explain any aspect of transitions have been Stage-Environment Fit Theory (Eccles and Midgley, 1989; Eccles et al., 1993), Bronfenbrenner's Ecological Systems Theory (1993), Life Course Theory (Elder, 1998), Self-determination Theory (Ryan and Deci, 2000), Resilience (Luthar, 2006; Newman and Blackburn, 2002 linked it with transitions; Rutter, 1987), Emotional Intelligence (Goleman, 1995) and Self-esteem (Mruk, 1999; Jindal-Snape and Miller, 2008 linked self-esteem and resilience with transitions). Usually these theories are referred to in the introduction of an article but there is little evidence of the theories feeding into the conceptualization, research design and/or being used to explain the results. Apart from the Life Course Theory, none of the other theories, explicitly refers to transitions. Despite this, none of the articles critiqued the theory/ies and, more importantly, they did not suggest any further development of these theories based on their research context, nor did they develop new transitions theory/ies.

Adapting theories borrowed from other contexts and disciplines

When considering the transitions of international students, initial research literature in the field focused on existing *acculturation* theories which were based on the transition experiences of migrants who were seen to be experiencing culture shock. Culture shock is based on the premise that when a person experiences a new culture, the difference between their own culture and that of the host country, will lead to the migrant experiencing stress and shock. Three groups of theories were used to explain acculturation: Stress and coping (Affect) theories which hypothesize that life changes are bound to be stressful and a migrant needs to develop coping strategies to help combat stress; Culture learning (Behavior) theories that highlighted that social interaction is important for migrants and that they need to learn culturally appropriate social skills to survive in the new culture and context; and Social identification (Cognitive) theories that hypothesize that any cross-cultural transition will lead to a change in cultural identity. This ABC of acculturation theoretical framework (Ward et al., 2001) offered a better understanding of the multi-dimensional processes involved in the acculturation of migrants.

However, all these theories seemed to consider the migrant and, in our context, the international student, as passive. This changed as Berry (2005) moved the thinking and suggested that migrants' attitudes toward acculturation and strategies used could lead to four outcomes, namely, integration (maintains original identity and accepts new culture), assimilation (rejects original and accepts host identity), separation (only identifies with original culture), and marginalization (rejection of both cultural identities) (Fig. 1). Choice in this context is important but integration is seen to be the healthiest acculturation strategy. He also discussed the

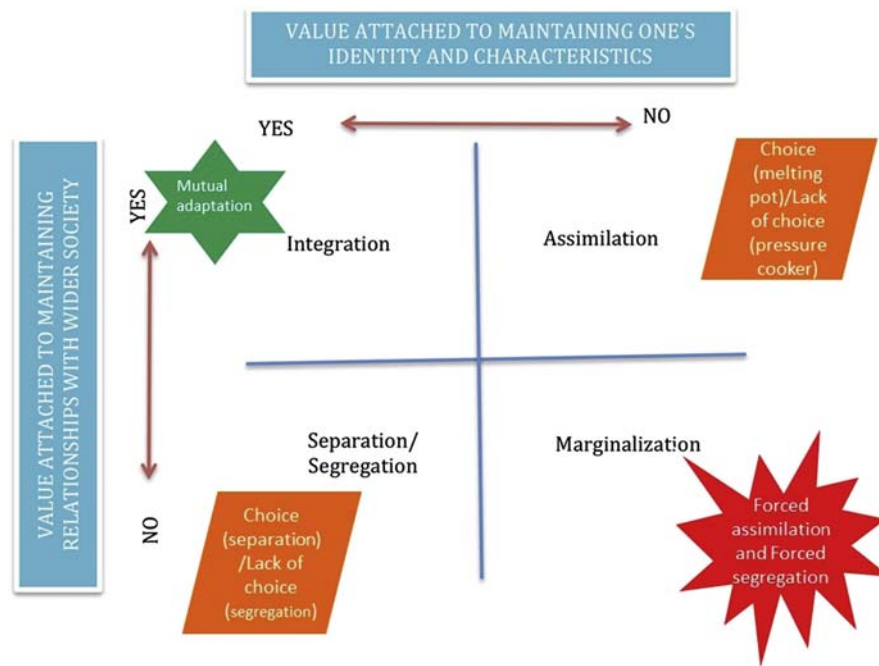


Fig. 1 Acculturation strategies. Based on [Berry \(1997\)](#).

mutual nature of acculturation, that is, it is bi-dimensional rather than uni-dimensional ([Berry, 2005, 2006; Smith and Khawaja, 2011](#)).

However, there is, potentially, an issue in applying theories, developed in the context of economic migrants, to international students' transitions. International students have a different motivation to study abroad; they are seen to be top level students, bringing with them large funding as fees, helping economic regeneration due to increased spending in the host nation, and are seen to enrich the host university's academic environment. They carefully choose the country and university they would like to move to and these decisions are made based on the best match with their aspirations and needs. Their main purpose when studying abroad is to integrate in the host environment.

Taking conceptualizations of migrants' transitions further, both the international student and the host university and community experience acculturation and there is mutual adaptation with strong emphasis on mutual connectedness. This moves the ABC theory further to a social, Ecological Systems Theory perspective ([Bronfenbrenner, 1979](#)), highlighting the role of the wider society in the acculturation of international students ([Bethel et al., 2016](#)). This change in the theoretical framework emphasizes that the international students are not passive and actively adapt to the differences in cultural and pedagogical approaches of the host nation. This was also apparent in an educational context; [Zhou et al. \(2008\)](#) found that UK higher education staff adapted their teaching style and assessment practices to accommodate international students' pedagogical needs. One could argue that these acculturation processes and attitudes are, in some ways, similar to pedagogical attitudes although they can be dictated, to a large extent, by assessment practices of the host nation and university. This has led to increasing calls for internationalization of curriculum, although few have raised voice about internationalization of pedagogical approaches.

However, it was important to consider how these theories could be applied to transitions of international students and whether they were fit for purpose. This led to the development of a theoretical framework called Social Network-Affective, Behavior and Cognition (SN-ABC) ([Rienties and Jindal-Snape, 2016a,b](#); see [Fig. 2](#)), pulling together ABCs of acculturation, Social Network theories and Social Capital Theory ([Putnam, 2001; Rienties et al., 2015](#)) by undertaking research in international classrooms.

Unlike previous theories used to understand international students' transitions, SN-ABC is based on the belief that students do not passively experience "culture shock" or "pedagogical shock", and that they proactively choose to behave in particular ways for adaptation to the cultural and pedagogical environment in the host country; admittedly sometimes these decisions might not be consciously made by the student ([Rienties and Jindal-Snape, 2016a,b; Rienties et al., 2015](#)). Further, the environment, whether human, physical or social, in which these interactions and transitions take place, plays a substantial role in the outcomes, as there will be an impact of the university and/or national policy, educational practices and assessments, and the local community on all students. Therefore, although the student characteristics and motivation to interact and adapt (see [Fig. 1](#)) might act as mediating factors, simple modifications of the environment could have an impact on the international students and the host context, such as hosting of students by community members for a weekend, assessments that involve small group work that require host and international students to work together, as well as community organizations and universities providing support to their family if they travel with them. This is an important example of understanding the limitations of using previous theories and the need to

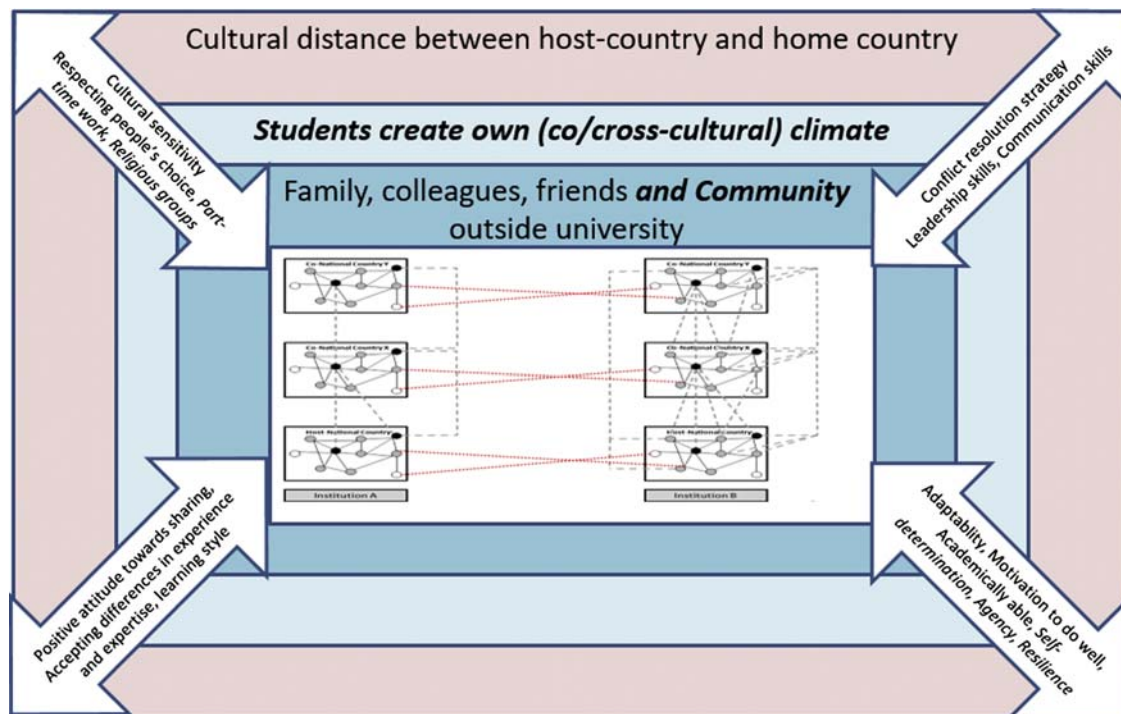


Fig. 2 SN-ABC theory. From Rienties and Jindal-Snape (2016a,b).

adapt them for one's particular context. Further, this shows how a new theory, that is fit-for-purpose for transitions research, policy and practice, can be developed.

Developing research informed transitions theories

Educational and life transitions (ELT) theory

Based on research with international doctoral students, an Educational and Life Transitions (ELT) theory was proposed (Jindal-Snape and Ingram, 2013) as it was found that not only did they experience educational transitions due to, for example, change in academic level, academic protocols, assessment styles but they also experienced other life transitions due to, for example, moving to a new country with different social and cultural norms, leaving or bringing family members with them and financial aspects. And that these transitions were dynamic and happened at the same time, with the possibility of one transition experience being positive and another being negative. Further, these different experiences interacted with each other. If the positive transition experiences were stronger than the negative experiences, they could act as a buffer; for example, if the academic transitions are going well, they could mitigate the negative peer relationship transitions (see Fig. 3). However, if the overall negative transitions experiences are overpowering, the positive transition experiences might be overlooked; for example, an international student struggling financially might find it so overwhelming that any success in their studies might not lead to any positive experiences.

It is also important to remember that the multiple transitions are dynamic and the student might have different experiences at different times. Further, the transition experiences might also be influenced by what is most important to the student at *particular* time points and this will fluctuate over the period of their university studies. This fluctuation was also observed in other transitions research. For example, primary-secondary transitions research found that at the start of secondary school, the focus of the students and families were relationship transitions which changed to educational transitions later (Galton, 2010).

However, even with the ELT theory, which emphasizes multiple transitions, the need to develop it further arose. This development was based on transitions research with international students and students in other contexts, research which highlighted the impact of the students' transitions on the people significant to them, e.g., their family. This led to the development of the Multiple and Multi-dimensional Transitions Theory.

Multiple and multi-dimensional transitions (MMT) theory

The Multiple and Multi-dimensional Transitions (MMT) theory proposes that each individual experiences multiple transitions and that their transitions can trigger transitions for significant others. Conversely, significant others will be experiencing their own transitions and these transitions will trigger and/or have an impact on that individual's transitions. These mutual interactions of transitions between individual and significant others lead to an almost ongoing ripple effect across the ecosystem, highlighting the multi-dimensional nature of the theory. Further, these transitions are dynamic and are not taking place in a vacuum; they are in an environment that is constantly changing. This can be explained in the COVID-19 context with several environmental changes

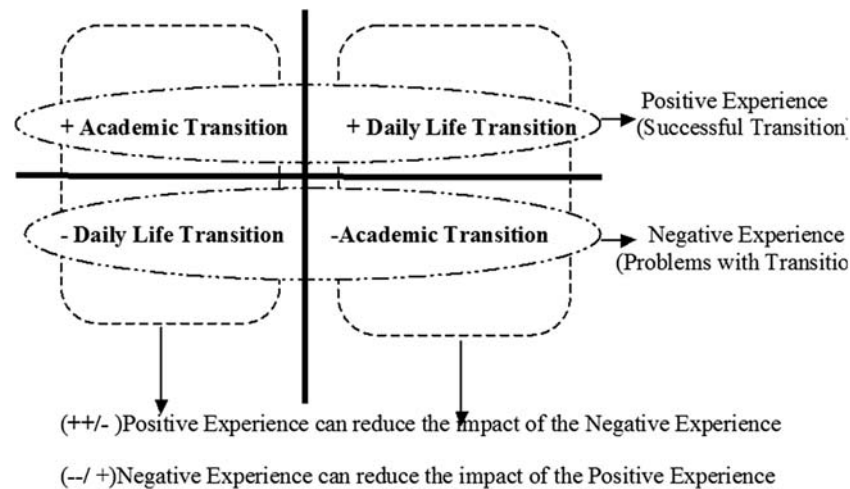


Fig. 3 Educational and life transitions model (Jindal-Snape and Ingram, 2013) Note: This is not a static model and individuals' experiences will fluctuate over the period of their study and stay. The overall experience can shift depending on the strength of the positive or negative transition experience. The authors have permission to reproduce Fig. 3 under Creative Commons License.

taking place which, in turn, led to school and university closures. As can be seen in Fig. 4, the school teacher was dealing with a rapidly changing situation with transitions in their work and home contexts (Stones et al., 2020). Their students' transitions were also triggering further transitions in terms of adapting educational provision and pastoral support that suited the students' emerging and ongoing transition needs. This created transition support needs not only for their students but also for them. Further, their transition support needs in the context of significant others from their family were also emerging over time.

Fig. 5 shows the multiple and multi-dimensional transitions of a university lecturer and her students, as well as her family members, for example, the impact of home schooling on her professional life (Glazzard et al., 2020b). It also highlights that transition is an ongoing process with adaptation happening over time (Jindal-Snape, 2016). For instance, this parent might experience transitions personally and professionally in the future, due to their child's developmental stage, for example when childcare responsibilities change and the child prefers to spend more time with peers and starts asserting the need for independence from parents. These transitions due to their child's transitions will likely happen alongside other transitions in the lecturer's personal, social and professional life. Further, the parent's transitions will trigger transitions for the child. Therefore, the child and parent will experience transitions as a result of one another's transitions and these transitions will interact with the transitions of others in their ecosystem such as other children at school, their parents, and school professionals.

To understand MMT better, it can be visualized using Rubik's cube as a metaphor (Fig. 6; Jindal-Snape, 2016; Jindal-Snape et al., 2017). If each individual and their significant others is seen as one color, we can visualize six individuals and their significant others as well as their mutual interactions. If anything changes for anyone in that single color, that change can trigger changes across the Rubik's cube not only for that one color but for *all* colors and sides. Those changes then lead to transitions for all those in the connected ecosystems.

This will be explained further through examples from research that has used MMT as the theoretical framework.

Multiple and multi-dimensional transitions of young adults and significant others

In a study conducted with young adults with life limiting conditions, MMT was used as the framework to design a longitudinal study in which data were collected from young adults ($n = 12$) at three time-points and significant others that they nominated at two time points (Jindal-Snape et al., 2019). The significant others included parents, grandparents, siblings, physician, family doctor and social care professionals ($n = 23$, however not all participated at both time points). None of the young adults nominated school or university teachers but they made reference to them in their interviews. All participants were asked questions about their own and others' transitions. The sample size was small but rich data were collected. The findings suggested that young adults experienced multiple transitions such as health transitions, educational transitions, social and relationship transitions, and developmental transitions (including moving out, independent living, death and dying). Further, these transitions also interacted over time. Their transitions triggered transitions for significant others, such as one grandparent took early retirement to become a fulltime carer, one sibling decided to place their educational transition, and another sibling their social transition, on hold. Significant others were also experiencing unrelated multiple transitions which also interacted with the young adults' transitions.

Multiple and multi-dimensional transitions of university students who identify as LGBTQ+

In another study that investigated transitions of students to, and through, university, MMT theory was used to provide a theoretical framework (Glazzard et al., 2020a). Data were collected from five students through yearly interviews across three years of study (63,875 transcribed words), longitudinal audio diaries (23 audio diaries) and photo elicitation (22 photos). The students reported

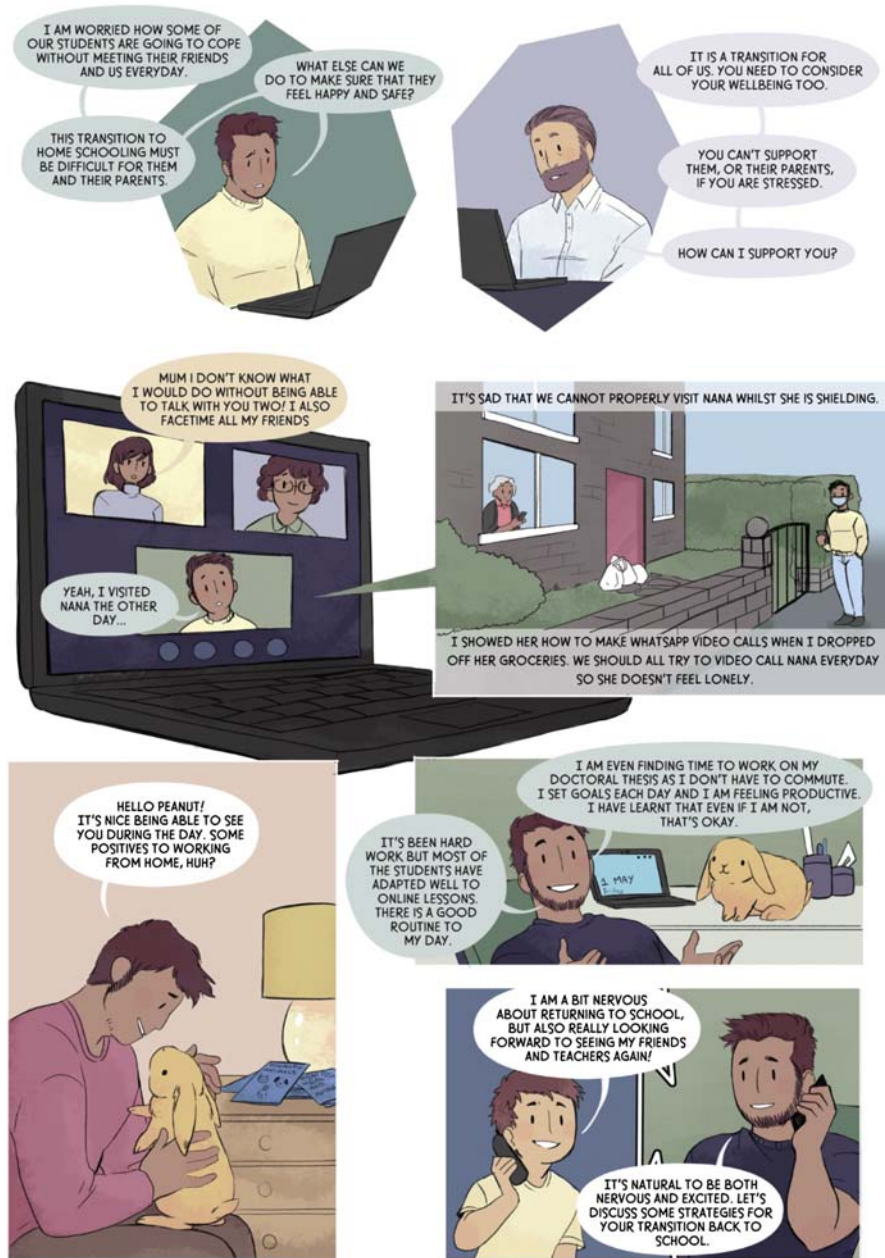


Fig. 4 School teacher and significant others' multiple and multi-dimensional transitions due to COVID-19 (Stones et al., 2020).

that they experienced multiple transitions, including academic (demands of the academic study and different approaches to teaching-learning and assessment), social (losing and developing new friendships), psychological, geographical (moving from home to a different city), professional (for those on qualifying programs), and identity (student identity, exploring and establishing LGBTQ + identity) transitions. Transitions were not linear or sequential and were experienced in multiple domains, every day. The students also reported having both positive and negative transition experiences at the same time, in different domains.

Multiple and multi-dimensional transitions of higher stage medical trainees

This study focused on the transitions of higher stage medical trainees from four health boards in Scotland (Gordon et al., 2017; see Fig. 6). This study used the MMT theoretical framework and undertook a longitudinal qualitative study which focused on participants' multiple and multi-dimensional transitions. Participants were 20 doctors within 6 months of completing postgraduate training. All participated in an interview at the start of the study, 18 participated in the interviews at the end of the study and those 18 also completed longitudinal audio/written diaries over a nine-month period. Participants submitted between 2 and 30 diary entries, an average of 13 diary entries. This study suggests that transitions happen over time and they are non-linear. Participants

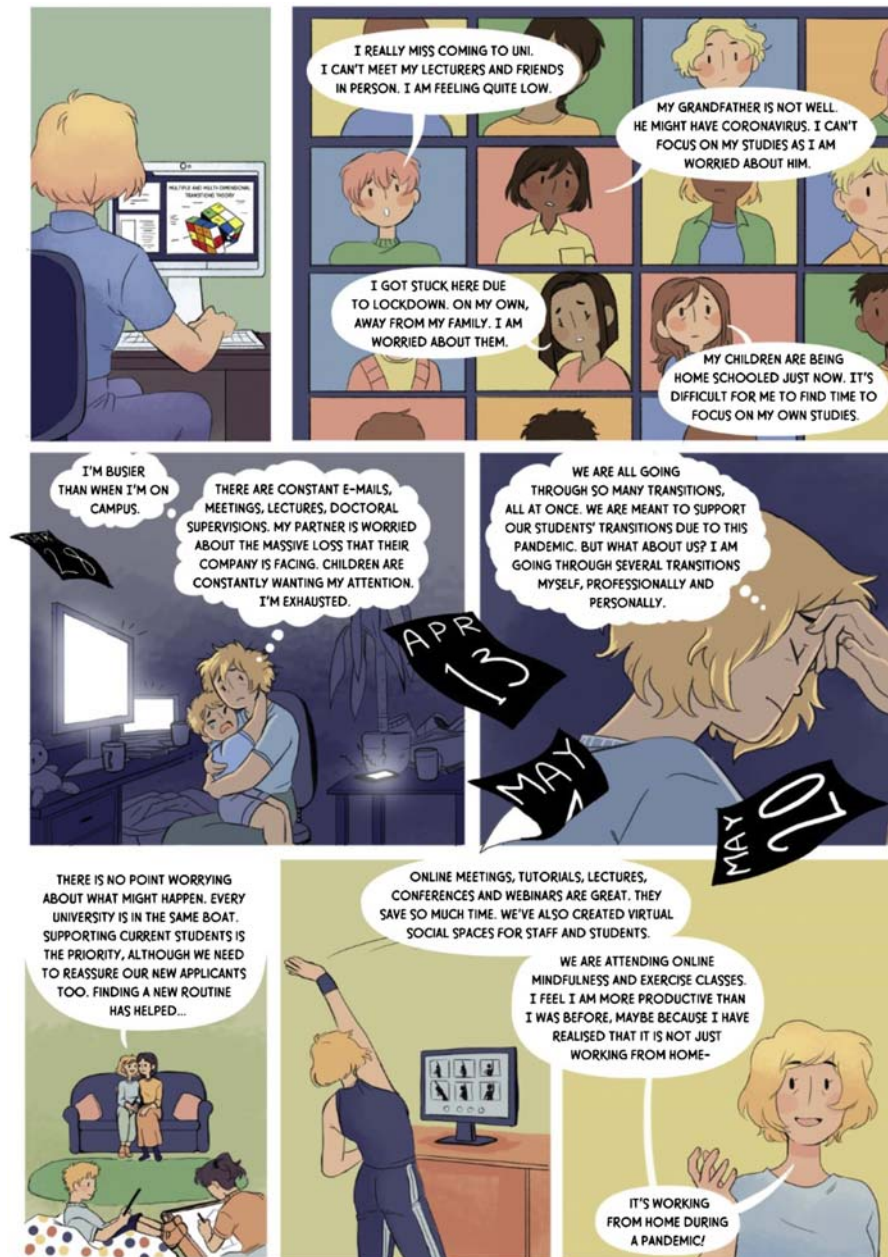


Fig. 5 University professor and significant others' multiple and multi-dimensional transitions due to COVID-19 (Glazzard et al., 2020b).

had both positive and negative multiple transition experiences, sometimes in the same or different domains of their life at the same time. These included work-related transitions (e.g., taking on new roles) and home related transitions (e.g., moving home due to change in organization and location). Transition experiences were complex and multi-dimensional, with their transitions interacting with or triggering transitions of significant others, such as a child spending more time in childcare setting due to changes in parental role and responsibilities. The participants were also affected by others' transitions, for example one indicated that they had to work flexibly to accommodate childcare due to their spouse's work transitions. Participants' work-related transitions also meant that they were able to spend less time with patients and had relationship transitions with existing colleagues when they took on a higher-level role in the same hospital, as well as those who moved hospitals, having to form new relationships.

Transition-to-trained-doctor (T3D) model

In keeping with the need to adapt and develop theories for one's own specific context, MMT was further expanded to include a Transition-To-Trained-Doctor (T3D) conceptual framework (see Fig. 7). The latter explicitly takes into consideration not only the multiple domains but also the multiple contexts by focusing on the contexts in Canada, Netherlands and the UK (Gordon et al., 2020). Using MMT as the theoretical framework, secondary data analysis was undertaken of 55 interview transcripts. Participants

MULTIPLE AND MULTI-DIMENSIONAL TRANSITIONS FROM TRAINEE TO TRAINED DOCTOR: A QUALITATIVE LONGITUDINAL STUDY

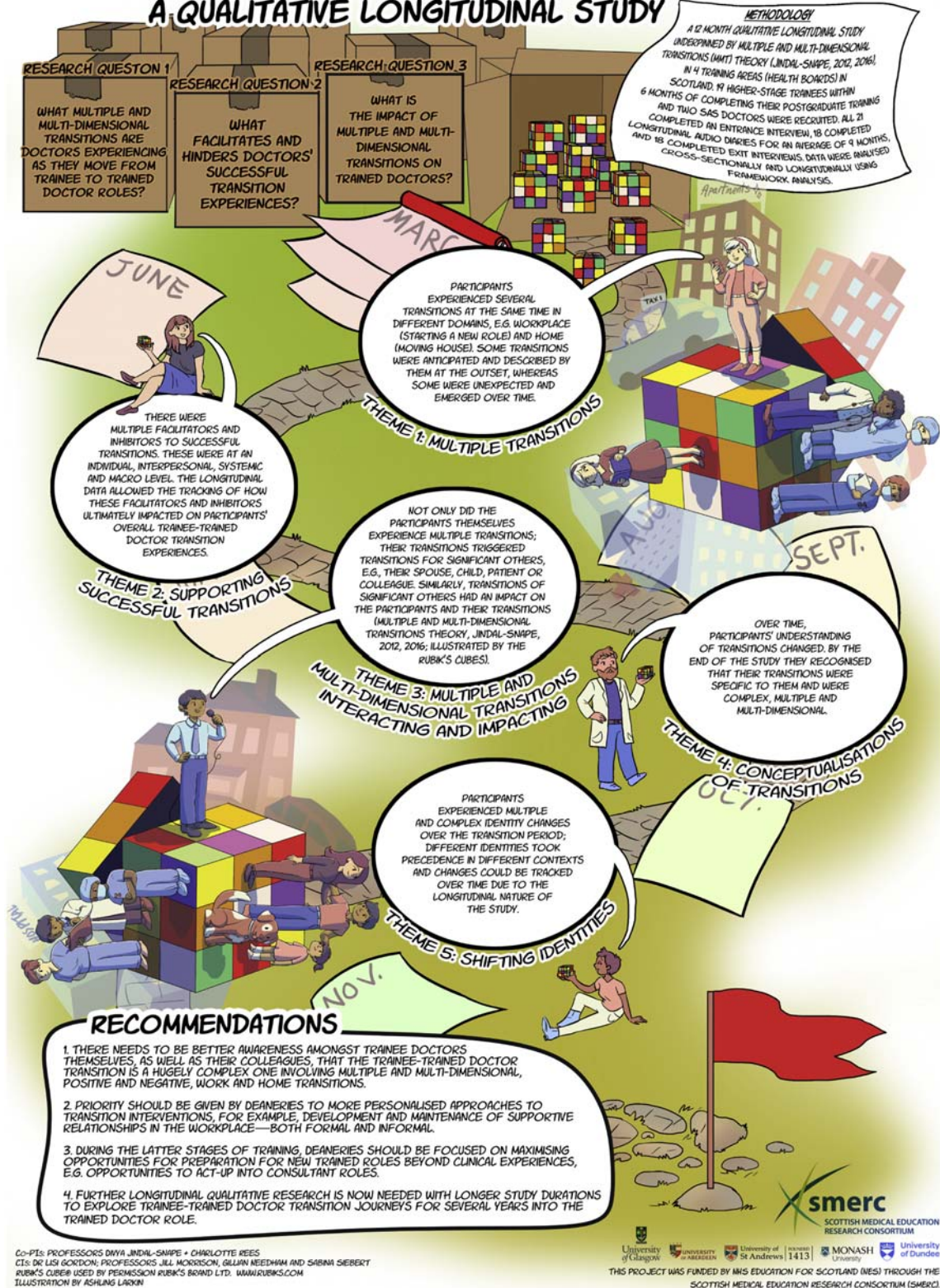


Fig. 6 Multiple and multi-dimensional transitions of higher stage medical trainees (Jindal-Snape et al., 2017).



Fig. 7 The T3D model. [Gordon et al. \(2020\)](#).

described the transition in four domains of their work and personal lives (physical, cultural, psychological, and social; [Jindal-Snape, 2016](#)) and four contexts, taking cognizance of [Bates and Ellaway's \(2016\)](#) multiple clinical context perspective (workplace, role, educational and home). As can be seen from [Fig. 7](#), the individual is situated in one or multiple contexts in line with their prominence for that individual at any given time, along with the overarching wider national context.

Within each context, all four transition domains exist, implying that an individual has multiple types of transitions happening in every context they are situated in every day. Further, the transition domains and contexts interact with each other. The multi-dimensional nature of transitions and interaction with the transitions of significant others is depicted through open circles. The T3D model takes cognizance of intentional fluidity and mobility.

The 12 pillars of successful multiple and multi-dimensional transitions

Our research suggests that certain dispositions and environmental factors can lead to successful educational and life transitions ([Glazzard et al., 2020a](#); [Gordon et al., 2017](#); [Jindal-Snape et al., 2020, 2021](#); [Jindal-Snape and Cantali, 2019](#); [Jindal-Snape et al.,](#)

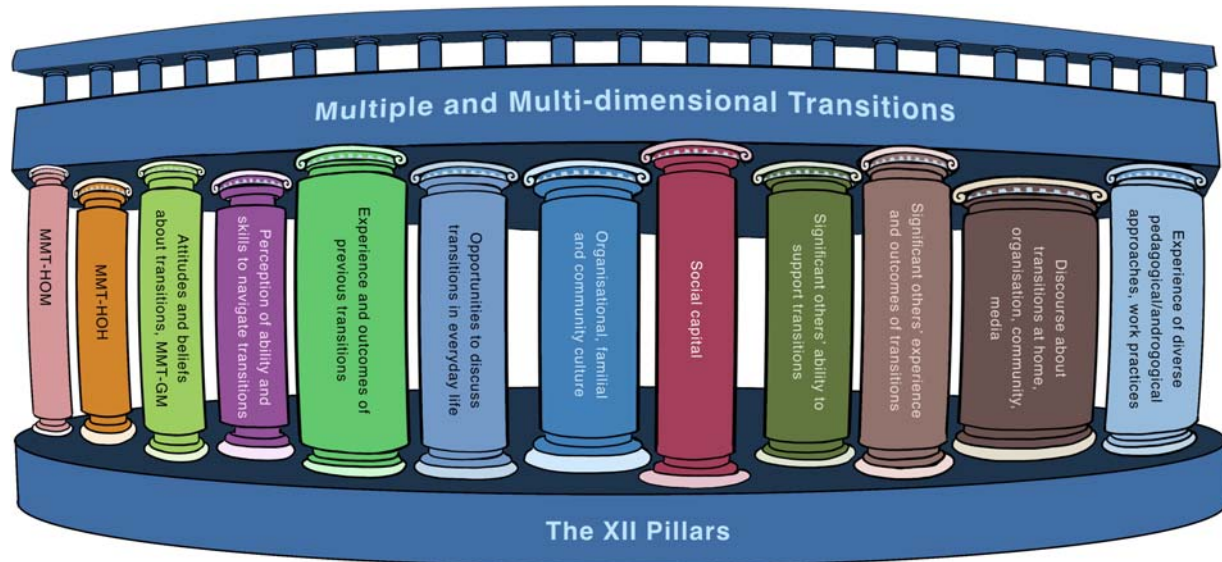


Fig. 8 The 12 pillars of successful multiple and multi-dimensional transitions. Divya Jindal-Snape & Clio Ding.

2019; Jindal-Snape and Miller, 2008). These can be summed up as 12 pillars of transitions, which interact and are interdependent but have been separated here for ease of understanding. The 12 pillars are related to dispositions, beliefs, support systems and environmental affordances including culture and discourse. It is worth noting that for successful transitions, not all pillars have to be present or have to be equal at the same time. Different pillars might be prioritised during different types of transitions. These pillars are useful for an individual across the life course and when navigating different types of transitions. It is also important to note that these are not within-person or fixed, rather they can be, and are, created and maintained over time (see Fig. 8).

The 12 pillars of successful MMT are:

1. Multiple and Multi-dimensional Transitions Habits of Mind (MMT-HOM)
2. Multiple and Multi-dimensional Transitions Habits of Heart (MMT-HOH)
3. Attitudes and beliefs about transitions (e.g., MMT Growth Mindset)
4. Perception of ability and skills to navigate transitions
5. Experience and outcomes of other multiple and concurrent transitions
6. Discourse about transitions at home, school, workplace, in the wider community, media
7. Social Capital (social connections/networks/relationships)
8. Significant others' ability to support transitions (e.g., family, professionals, peers)
9. Significant others' experiences and outcomes of their own transitions
10. Opportunities to discuss transitions in everyday life
11. Experience of diverse pedagogical/andragogical approaches/work practices
12. Organizational, familial and community culture

Habits of Mind have been defined as "thinking" dispositions to act intelligently when faced with a new situation or problem, and include a composite of diverse range of skills, attitudes and past experiences (Costa and Kallick, 2009). The skills and attitudes that can be developed to ensure MMT-HOM include providing opportunities to learn how to solve problems creatively, adaptability, open mindedness, curiosity, collaboration, resourcefulness, applying past learning to new situations, agency, persistence, and comfort with ambiguity and uncertainty (this is not an exhaustive list).

Habits of Heart have been defined as dispositions that can develop and enhance social-emotional intelligence. MMT-HOH focus on psychological adaptations and transitions related to, for example, change in relationships (and social capital), feelings about one's transition experiences, compassion for others, and most importantly self-compassion. Although separated here, MMT-HOM and MMT-HOH are habits that interact with one another. These are habits that can be developed and nurtured across the life course, whether in nursery/school/higher education/professional education and require diverse pedagogical/andragogical approaches/work practices.

Further, MMT-HOM and MMT-HOH will help an individual understand and access their social capital to navigate their transitions in new environments and with new people. On the other hand, an individual's social capital is also likely to influence their MMT-HOM and MMT-HOH. They will also have an impact on an individual's attitudes and beliefs about transitions (e.g., understanding of the dynamic nature of transitions and that transitions outcomes are also not static and they have the agency to change negative outcomes to positive over time, i.e., what I propose here as a Multiple and Multi-dimensional Transitions Growth Mindset (MMT-GM); see Dweck, 2017 for a discussion of Growth Mindset). THOH and THOM also have an impact on an individual's

perception of their ability and skills to successfully navigate transitions. For instance, an individual, when confronted by a new and unexpected situation, is more likely to have a positive attitude about transitions and would believe that they can manage them if they know how to solve a problem creatively and/or know how to adapt to change, and are self-compassionate when finding it difficult to solve the problem.

However, it is important to be mindful that the interactions between individuals and structures create dispositions over time and such interactions will have an impact on individual's attitudes toward transitions and perception of their own ability to navigate them successfully. These include unspoken rules and valuing of particular attributes in their environment, resulting in particular discourse about transitions in their environment, and other dimensions such as their experiences and outcomes of other multiple and concurrent transitions and significant others' ability to support the individual's transitions. The latter is affected by the significant others' experiences of their own transitions. Further, there are other pillars that interact with all of these, such as those related to opportunities to discuss about transitions in everyday life (e.g., opportunities to discuss excitement and concerns, strategies from those who have experienced those transitions), and organizational (whether of school or workplace), familial, and community culture.

These pillars are complex and impact on each other in multiple ways. They can be generated consciously and purposefully but require multiple approaches and partnership working between families, professionals, organizations, community, and media. Conceptualization of pillars in this way can highlight what resources an individual has to navigate transitions successfully and what can be done to increase them. However, it is important to note that it would be easier to influence some pillars than others.

It is a conceptual framework that can be used by teachers, for instance, to select pedagogical approaches that can develop young children's MMT-HOM and MMT-HOH. The teachers can also provide them with opportunities to discuss their beliefs about transitions, excitement/concerns, and strategies used by their peers, families and professionals when experiencing similar transitions. Further, teachers can ensure the discourse in the school about transitions is positive as well as involving parents and other stakeholders to support this discourse. However, it is important to realize that this discourse in the natural environment will be influenced by significant other's own transitions experiences, beliefs, attitudes and dispositions.

Conclusion

In this article, I have attempted to start a debate about the conceptualization and theorization of transitions. My argument is that until, and unless, we do this, we cannot advance transitions as a distinct area of study that can be pursued with rigor. This then has implications for the policy and practice that might be based on research in the area.

It is important, as is expected in any field, that authors explicitly present their conceptualization and operationalization of transitions. Ideally, they should also ask what transitions and successful transitions mean to their participants. It is also important that future research has a balanced discourse which focusses on both positive and negative aspects of transitions.

Further, it is crucial that future transitions research builds on previous transitions research, not just as a token mention in the introduction of an article but starting from the moment a researcher starts considering research questions and appropriate research design. It is important that future researchers are also mindful of the theoretical framework they are using and that it is threaded throughout their study. However, it is even more important that they critique the existing theories, for example, MMT, with regard to their particular context, to provide a more granular and robust framework that is fit-for-purpose.

In the context of policy and practice, it is important to understand the ongoing, dynamic nature of multiple transitions that necessitates support not only with one transition or at one particular time. Further, the transitions of significant others and their interplay with the students' transitions need to be understood. Professionals and families require support with their transitions if they can be expected to provide good transitions support to students.

I have also introduced some new concepts, which build on our research, conceptualization and theorization, namely MMT-HOM, MMT-HOH, MMT-GM and the XII Pillars of MMT. They provides insights into how, from a very young age, professionals can support the development of, at least some, habits and dispositions that will help them with navigation of transitions throughout their life. Further, it also can provide insights into where individuals might have less resources and how they can be supported to develop them.

References

- Andreas, J.B., Jackson, K.M., 2015. Adolescent alcohol use before and after the high school transition. *Alcohol Clin. Exp. Res.* 39 (6), 1034–1041.
- Arens, A.K., Yeung, A.S., Craven, R.G., Watermann, R., Hasselhorn, M., 2013. Does the timing of transition matter? Comparison of German students' self-perceptions before and after transition to secondary school. *Int. J. Educ. Res.* 57, 1–11.
- Bailey, S., Baines, E., 2012. The impact of risk and resiliency factors on the adjustment of children after the transition from primary to secondary school. *Educ. Child Psychol.* 29 (1), 47–63.
- Bates, J., Ellaway, R., 2016. Mapping the dark matter of context: a conceptual scoping review. *Med. Educ.* 50 (8), 807–816.
- Benner, A.D., Wang, Y., 2014. Shifting attendance trajectories from middle to high school: influences of school transitions and changing school contexts. *Dev. Psychol.* 50 (4), 1288–1301. <https://doi.org/10.1037/a0035366>.
- Berry, J.W., 1997. Immigration, Acculturation, and Adaptation. *Appl. Psychol.* 46 (1), 5–34. <https://doi.org/10.1111/j.1464-0597.1997.tb01087.x>.
- Berry, J.W., 2005. Acculturation: living successfully in two cultures. *Int. J. Intercult. Relat.* 29 (6), 697–712.

- Berry, J.W., 2006. Acculturative stress. In: Wong, P., Wong, L. (Eds.), *Handbook of Multicultural Perspectives on Stress and Coping*. Springer, New York, pp. 287–298.
- Bethel, A., Szabo, A., Ward, C., 2016. Parallel lives? Predicting and enhancing connectedness between international and domestic students. In: Jindal-Snape, D., Rienties, B. (Eds.), *Multi-dimensional Transitions of International Students to Higher Education*. Routledge, New York.
- Brewin, M., Statham, J., 2011. Supporting the transition from primary school to secondary school for children who are looked after. *Educ. Psychol. Pract.* 27 (4), 365–381. <https://doi.org/10.1080/02667363.2011.624301>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press, Cambridge, Massachusetts.
- Bronfenbrenner, U., 1993. The ecology of cognitive development: research models and fugitive findings. In: Wonziak, R., Fischer, K. (Eds.), *Development in Context: Acting and Thinking in Specific Environments*. Erlbaum, Hillsdale, NJ, pp. 3–44.
- Burchinal, M.R., Roberts, J.E., Zeisel, S.A., Rowley, S.J., 2008. Social risk and protective factors for African American children's academic achievement and adjustment during the transition to middle school. *Dev. Psychol.* 44 (1), 286–292.
- Costa, A.L., Kallick, B. (Eds.), 2009. *Habits of Mind Across the Curriculum: Practical and Creative Strategies for Teachers*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Draus, P., Roddy, J., Asabigi, K., 2014. Making sense of the transition from the detroit streets to drug treatment. *Qual. Health Res.* 25 (2), 228–240. <http://journals.sagepub.com/doi/pdf/10.1177/1049732314552454>.
- Dweck, C., 2017. The Journey to Children's Mindsets—and Beyond. *Child Dev. Perspect.* 11 (2), 139–144. <https://doi.org/10.1111/cdep.12225>.
- Eccles, J.S., Midgley, C., 1989. Stage-environment fit: developmentally appropriate classrooms for early adolescents. In: Ames, R., Ames, C. (Eds.), *Research on Motivation in Education*. Academic Press, New York, NY, pp. 139–181.
- Eccles, J.S., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C., et al., 1993. Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. *Am. Psychol.* 48, 90–101. <https://doi.org/10.1037/0003-066X.48.2.90>.
- Elder, G.H.J., 1998. The life course as developmental theory. *Child Dev.* 69 (1), 1–12. <https://doi.org/10.1111/j.1467-8624.1998.tb06128.x>.
- Evangelou, M., Taggart, B., Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., 2008. *Effective Pre-school, Primary and Secondary Education 3-14 Project (EPPSE 3-14) What Makes a Successful Transition From Primary to Secondary School?* DCSF Publications, Nottingham.
- Fortuna, R., 2014. The social and emotional functioning of students with an autistic spectrum disorder during the transition between primary and secondary schools. *Support Learn.* 29 (2), 177–191.
- Galton, M., 2010. Moving to secondary school: what do pupils in England say about the experience? In: Jindal-Snape, D. (Ed.), *Educational Transitions: Moving Stories From Around the World*. Routledge, New York, pp. 107–124.
- Glazzard, J., Jindal-Snape, D., Stones, S.O., 2020a. Transitions into and through higher education: the lived experiences of students who identify as LGBTQ+. *Front. Educ.* 5, 1–15. <https://doi.org/10.3389/educ.2020.00081>.
- Glazzard, J., Stones, S.O., Jindal-Snape, D., Murray, C., Laird, C., 2020b. University Lives in Transition. *Pandemic Tales, UniVerse*. <https://doi.org/10.20933/100001170>.
- Goleman, D., 1995. *Emotional Intelligence: Why It Can Matter More Than IQ for Character, Health and Lifelong Achievement*. Bantam Press, New York.
- Gordon, L., Jindal-Snape, D., Morrison, J., Muldoon, J., Needham, G., Siebert, S., Rees, C., 2017. Multiple and multi-dimensional transitions from trainee to trained doctor: a qualitative longitudinal study in the UK. *BMJ Open* 7 (11), e018583. <https://doi.org/10.1136/bmjopen-2017-018583>.
- Gordon, L., Teunissen, P.W., Jindal-Snape, D., Bates, J., Rees-Sidhu, C., Westerman, M., Sinha, R., van Dijk, A., 2020. An international study of trainee-trained transitions: introducing the transition-to-trained-doctor (T3D) model. *Med. Teach.* 42 (6), 679–688. <https://doi.org/10.1080/0142159X.2020.1733508>.
- Jindal-Snape, D., Cantali, D., 2019. A four-stage longitudinal study exploring pupils' experiences, preparation and support systems during primary-secondary school transitions. *Br. Educ. Res. J.* 45 (6), 1255–1278. <https://doi.org/10.1002/berj.3561>.
- Jindal-Snape, D., Foggie, J., 2008. A holistic approach to primary-secondary-transitions. *Improv. Sch.* 11, 5–18.
- Jindal-Snape, D., Ingram, R., 2013. Understanding and supporting triple transitions of International Doctoral Students: ELT and SuReCom models. *J. Perspect. Appl. Acad. Pract.* 1 (1), 17–24. <http://jpaap.napier.ac.uk/index.php/JPAAP/article/view/39>.
- Jindal-Snape, D., Miller, D.J., 2008. A challenge of living? Understanding the psychosocial processes of the child during primary-secondary transition through resilience and self-esteem theories. *Educ. Psychol. Rev.* 20, 217–236.
- Jindal-Snape, D., Rienties, B., 2016. Understanding multiple and multi-dimensional transitions of International Higher Education students: setting the scene. In: Jindal-Snape, D., Rienties, B. (Eds.), *Multi-dimensional Transitions of International Students to Higher Education*. Routledge, pp. 1–17. <https://www.routledge.com/products/9781138890909>.
- Jindal-Snape, D., Rees, C., Gordon, L., Morrison, J., Needham, G., Siebert, S., Larkin, A., 2017. *Multiple and Multi-dimensional Transitions From Trainee to Trained Doctor: A Qualitative Longitudinal Study (Infographic)*. University of Dundee.
- Jindal-Snape, D., Johnston, B., Pringle, J., Kelly, T., Scott, R., Gold, L., Dempsey, R., 2019. Multiple and multi-dimensional life transitions in the context of life-limiting health conditions: longitudinal study focusing on perspectives of young adults, families and professionals. *BMC Palliat. Care* 18, 1–12. <https://doi.org/10.1186/s12904-019-0414-9> [30].
- Jindal-Snape, D., Hannah, E.F.S., Cantali, D., Barlow, W., MacGillivray, S., 2020. Systematic literature review of primary-secondary transitions: international research. *Rev. Educ.* 8 (2), 526–566. <https://doi.org/10.1002/rev3.3197>.
- Jindal-Snape, D., Symonds, J., Hannah, E.F.S., Barlow, W., 2021. Conceptualizing primary to secondary school transitions: a systematic mapping review of worldviews, theories and frameworks. *Front. Educ.* 6, 540027. <https://doi.org/10.3389/educ.2021.540027>.
- Jindal-Snape, D., 2016. *A-Z of Transitions*. Palgrave Macmillan.
- Jindal-Snape, D., 2018. Transitions from early years to primary and primary to secondary schools in Scotland. In: Bryce, T., Humes, W., Gillies, D., Kennedy, A. (Eds.), *Scottish Education*, fifth ed. Edinburgh University Press.
- Keay, A., Lang, J., Frederickson, N., 2015. Comprehensive support for peer relationships at secondary transition. *Educ. Psychol. Pract.* 31 (3), 279–292. <https://doi.org/10.1080/02667363.2015.1052046>.
- Kirby, E., Broom, A., Good, P., 2014. The role and significance of nurses in managing transitions to palliative care: a qualitative study. *BMJ Open* 4, e006026. <https://doi.org/10.1136/bmjopen-2014-006026>.
- Knesting, K., Hokanson, C., Waldron, N., 2008. Settling in: facilitating the transition to an inclusive middle school for students with mild disabilities. *Int. J. Disabil. Dev. Educ.* 55 (3), 265–276. <https://doi.org/10.1080/10349120802268644>.
- Kralik, D., Visentin, K., van Loon, A., 2006. Transition: a literature review. *J. Adv. Nurs.* 55 (3), 320–329. <https://doi.org/10.1111/j.1365-2648.2006.03899.x>. PMID: 16866826.
- Langenkamp, A.G., 2010. Academic vulnerability and resilience during the transition to high school: the role of social relationships and district context. *Sociol. Educ.* 83 (1), 1–19. <https://doi.org/10.1177/0038040709356563>.
- Lofgran, B.B., Smith, L.K., Whiting, E.F., 2015. Science self-efficacy and school transitions: elementary school to middle school, middle school to high school. *Sch. Sci. Math.* 115 (7), 366–376. <https://doi.org/10.1111/ssm.12139>.
- Luthar, S.S., 2006. Resilience in development: a synthesis of research across five decades. In: Cicchetti, D., Cohen, D.J. (Eds.), *Developmental Psychopathology: Risk, Disorder, and Adaptation*. Wiley, New York, pp. 739–795.
- Mackenzie, E., McMaugh, A., O'Sullivan, K., 2012. Perceptions of primary to secondary school transitions: challenge or threat? *Issues Educ. Res.* 22 (3), 298–314.
- Madjar, N., Chohat, R., 2017. Will I succeed in middle school? A longitudinal analysis of self-efficacy in school transitions in relation to goal structures and engagement. *Educ. Psychol.* 37 (6), 680–694. <https://doi.org/10.1080/01443410.2016.1179265>.
- Maher, D., Maher, D., 2010. Supporting students' transition from primary school to high school using the internet as a communication tool. *Technol. Pedagog. Educ.* 19 (1), 17–30.

- Makin, H., Pellicano, E., 2017. The primary-to-secondary school transition for children on the autism spectrum: a multi-informant mixed-methods study. *Autism Dev. Lang. Impair.* 2, 1–18.
- McIntosh, K., Flannery, K.B., Sugai, G., Braun, D.H., Cochrane, K.L., 2008. Relationships between academics and problem behavior in the transition from middle school to high school. *J. Posit. Behav. Interv.* 10 (4), 243–255. <https://doi.org/10.1177/1098300708318961>.
- Mruk, C., 1999. *Self-Esteem: Research, Theory and Practice*. Free Association Books, London.
- Mudaly, V., Sukhdeo, S., 2015. Mathematics learning in the midst of school transition from primary to secondary school. *Int. J. Educ. Sci.* 11 (3), 244–252.
- Neal, S., Frederickson, N., 2016. ASD transition to mainstream secondary: a positive experience? *Educ. Psychol. Pract.* 32 (4), 355–373. <https://doi.org/10.1080/02667363.2016.1193478>.
- Neal, S., Rice, F., Ng-Knight, T., Riglin, L., Frederickson, N., 2016. Exploring the longitudinal association between interventions to support the transition to secondary school and child anxiety. *J. Adolesc.* 50, 31–43. <https://doi.org/10.1016/j.adolescence.2016.04.003>.
- Newman, T., Blackburn, S., 2002. *Transitions in the Lives of Children and Young People: Resilience Factors*. Scottish Executive Education Department, Edinburgh.
- Peters, R., Brooks, R., 2016. Parental perspectives on the transition to secondary school for students with Asperger syndrome and high-functioning autism: a pilot survey study. *Br. J. Spec. Educ.* 43 (1), 75–91.
- Peters, S., 2010. *Literature Review: Transition From Early Childhood Education to School*. Ministry of Education, New Zealand.
- Poorthuis, A.M.G., Thomaes, S., Aken, M.A.G., Denissen, J.J.A., de Castro, B.O., 2014. Dashed hopes, dashed selves? A sociometer perspective on self-esteem change across the transition to secondary school. *Soc. Dev.* 23 (4), 770–783.
- Putnam, R.D., 2001. *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster, New York.
- Rainer, P., Cropley, B., 2015. Bridging the gap—but mind you don't fall. Primary physical education teachers' perceptions of the transition process to secondary school. *Education* 43 (5), 445–461. <https://doi.org/10.1080/03004279.2013.819026>.
- Rice, F., Frederickson, N., Shelton, K., McManus, C., Riglin, L., Ng-Knight, T., (n.d.). The School Transition and Adjustment Research Study (STARS). Retrieved from <http://www.ucl.ac.uk/stars>.
- Rienties, B., Jindal-Snape, D., 2016a. A social network perspective on ABC of international and host-national students. In: Jindal-Snape, D., Rienties, B. (Eds.), *Multi-dimensional Transitions of International Students to Higher Education, New Perspectives on Learning and Instruction*. Routledge, New York, pp. 53–70.
- Rienties, B., Jindal-Snape, D., 2016b. Multiple and multi-dimensional transitions of international students to higher education: a way forward. In: Jindal-Snape, D., Rienties, B. (Eds.), *Multi-dimensional Transitions of International Students to Higher Education, New Perspectives on Learning and Instruction*. Routledge, New York, pp. 259–283.
- Rienties, B., Johan, N., Jindal-Snape, D., 2015. Bridge building potential in cross-cultural learning: a mixed method study. *Asia Pac. Educ. Rev.* 16 (1), 37–48. <https://doi.org/10.1007/s12564-014-9352-7>.
- Rutter, M., 1987. Psychosocial resilience and protective mechanisms. *Am. J. Orthopsychiatr.* 57, 316–331.
- Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55 (1), 68–78.
- Scanlon, G., Barnes-Holmes, Y., McEntegart, C., Desmond, D., Vahey, N., 2016. The experiences of pupils with SEN and their parents at the stage of pre-transition from primary to post-primary school. *Eur. J. Spec. Needs Educ.* 31 (1), 44–58.
- Serbin, L.A., Stack, D.M., Kingdon, D., 2013. Academic success across the transition from primary to secondary schooling among lower-income adolescents: understanding the effects of family resources and gender. *J. Youth Adolesc.* 42 (9), 1331–1347. <https://doi.org/10.1007/s10964-013-9987-4>.
- Smith, R.A., Khawaja, N.G., 2011. A review of the acculturation experiences of international students. *Int. J. Intercult. Relat.* 35 (6), 699–713. <https://doi.org/10.1016/j.jintrel.2011.08.004>.
- Stones, S., Glazzard, J., Jindal-Snape, D., Murray, C., Laird, C., 2020. School Transitions: Whose Transitions Are They Anyway? *Pandemic Tales, UniVerse*. <https://doi.org/10.20933/100001169>.
- Strnadova, I., Cumming, T.M., Danker, J., 2016. Transitions for students with intellectual disability and/or autism spectrum disorder: carer and teacher perspectives. *Aust. J. Spec. Educ.* 40 (2), 141–156.
- Tso, M., Strnadova, I., 2017. Students with autism transitioning from primary to secondary schools: parents' perspectives and experiences. *Int. J. Incl. Educ.* 21 (4), 389–403. <https://doi.org/10.1080/13603116.2016.1197324>.
- Vasquez-Salgado, Y., Chavira, G., 2014. The transition from middle school to high school as a developmental process among Latino youth. *Hisp. J. Behav. Sci.* 36 (1), 79–94. <https://doi.org/10.1177/0739986313513718>.
- Ward, C., Bochner, S., Furnham, A., 2001. *The Psychology of Culture Shock*, second ed. Routledge, New York.
- Weiss, C.C., Baker-Smith, E., 2010. Eighth-grade school form and resilience in the transition to high school: a comparison of middle schools and K-8 schools. *J. Res. Adolesc.* 20 (4), 825–839. <https://doi.org/10.1111/j.1532-7795.2010.00664.x>.
- West, P., Sweeting, H., Young, R., 2010. Transition matters: pupils' experiences of the primary-secondary school transition in the West of Scotland and consequences for well-being and attainment. *Res. Pap. Educ.* 25 (1), 21–50.
- Witherspoon, D., Ennett, S., 2011. Stability and change in rural youths' educational outcomes through the middle and high school years. *J. Youth Adolesc.* 40 (9), 1077–1090. <https://doi.org/10.1007/s10964-010-9614-6>.
- Zhou, Y., Jindal-Snape, D., Topping, K., Todman, J., 2008. Theoretical models of culture shock and adaptation in international students in higher education. *Stud. High Educ.* 33 (1), 63–75. <https://doi.org/10.1080/03075070701794833>.

Learning progressions: affordances, critiques, and pathways toward equity

Alicia C. Alonzo, Department of Teacher Education, Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	544
Variation in learning progressions	545
Learning progression components	546
Emphasis on ordering vs. pathway	546
Generalizability	547
Scope and grain size	547
Applications of learning progressions	547
Standards, curriculum, and instruction	547
Formative (classroom) assessment	548
Summative (classroom and large-scale) assessment	548
Teachers' knowledge/tool for professional development	548
Coherent educational systems	549
Equity critiques of learning progressions	549
Content of learning progression components	549
Standards, curricula, and instruction	550
Assessment	550
Ordering of learning progression components	551
Standards, curricula, and instruction	551
Assessment	551
Current and future research to explore and address equity concerns with and through learning progressions	551
Diversity of learning contexts	552
Providing information about learning context	552
Developing context-specific learning progressions	552
Student diversity	553
Exploring student diversity	553
Representing diversity	554
Responding to diversity	554
Conclusion	554
References	554

Introduction

With strong connections to past research (e.g., [Carpenter and Moser, 1984](#); [Driver, 1989](#)), contemporary efforts to describe and explain how students' learning unfolds over time have focused on *learning progressions*. Over the past two decades, work to develop and use learning progressions (and associated constructs) has spread from science and mathematics to varied subjects (see [Table 1](#)), conducted in countries around the world (see [Table 2](#)). Extant examples of learning progressions span preschool (e.g., [Allen, 2017](#); [Moreno et al., 2021](#)) through university (e.g., [Magana, 2017](#); [Sevian and Talanquer, 2014](#)) and beyond (e.g., in-service teachers; [Bargagliotti, 2021](#); [Furtak et al., 2012](#); [Schneider and Plasman, 2011](#)). In addition to numerous journal articles

Table 1 Examples of work on learning progressions in subjects other than math and science.

Subject	Example(s) of work on learning progressions
Art	Maras (2021)
Language	Bailey and Heritage (2014) , van Rijn et al. (2014)
Computer science	Rich et al. (2018)
Engineering	Magana (2017)
Geography	Larsen and Harrington (2018) , Maeng (2020) , Solem et al. (2014)
Humanities	Hughes et al. (2020)
History	Lee and Ashby (2000) , Zarmati (2021)
General	Thinking skills (e.g., Kamei and Pavlovic, 2021); collaboration (Scoular, 2021)

Table 2 Examples of work on learning progressions around the world.

Country	Example(s) of work on learning progressions
Australia	Callingham et al. (2021), Scoular (2021), Stephens et al. (2021), Zarmati (2021)
Canada	Liu and Tang (2004)
China	Bai (2020), Chen et al. (2017), Yao and Guo (2018)
England	Allen (2017), Johnson (2013), Russell and McGuigan (2019)
Finland	Wang et al. (2017)
Germany	Neumann et al. (2013)
Italy	Colantonio et al. (2018)
Korea	Jung et al. (2020), Maeng (2020), Paik et al. (2017)
Malaysia	Chin et al. (2021a,b)
New Zealand	New Zealand Ministry of Education (2010)
Singapore	Chang et al. (2020)
Spain	Bravo-Torija and Jimenez-Aleixandre (2018), Ivars et al. (2020), Moreno et al. (2021)
Taiwan	Chiu and Wu (2013), Wang et al. (2015)
Thailand	Ladachart et al. (2020), Pongprapan and Jituafula (2021)
United States	Confrey et al. (2014), Smith et al. (2006)
Wales	Hughes et al. (2020)

and reports, special journal issues (e.g., Alonzo, 2018; Bernholt and Sevan, 2018; Clements and Sarama, 2004; Duncan and Hmelo-Silver, 2009a; Mitchell, 2018; Oates and Seah, 2021; Talanquer, 2013), books (e.g., Alonzo and Gotwals, 2012; Furtak, 2017; Maloney et al., 2014; Siemon et al., 2019; Solem et al., 2014), and conferences (e.g., Learning Progressions in Science, 2009; Learning Progressions Footprint Conference, 2011) have been devoted to this work.

As with other constructs with widespread appeal (e.g., “inquiry”), a range of meanings has been attached to “learning progressions,” and the distinction between this and similar constructs is not always clear. In this article, I define a learning progression to be a framework that represents student learning of a competency through a series of components (e.g., levels, constituent “building blocks”) that describe the increasingly more difficult, complex, and/or sophisticated performances expected as students learn. This broad definition encompasses wide variation in frameworks that have been labeled “learning progressions”; it also includes *learning trajectories* (often in mathematics), for example Simon (1995) and Maloney et al. (2014). Although informed by disciplinary perspectives, learning progressions describe competencies from the point of view of the student, rather than disciplinary logic (e.g., Bailey and Heritage, 2014; Fonger et al., 2018). To capture the wide range of knowledge, practices, and other learning targets that learning progressions may represent, I draw on the notion of competencies as “complex cognitive abilities” and “affective and volitional dispositions” that underlie “real-world” performance (Blömeke et al., 2015, p. 6), such as engaging in the authentic work of a discipline (e.g., Wilson and Lehrer, 2021). *Teaching-learning sequences*, structured sequences of instructional activities with “well-documented teaching suggestions and expected student reactions” (Psillos and Kariotoglou, 2016, p. 12), and other frameworks for “teaching learning progressions” (Wickman, 2014) include attention to changes in students’ competencies. However, I exclude these frameworks from my definition as having a primary goal of describing teaching, rather than learning.

Learning progressions have been considered promising tools for supporting student learning by informing standards, curriculum, instruction, assessment, and teacher learning (e.g., Daro et al., 2011). Despite the promise of learning progressions to support various parts of the educational system, these tools have been critiqued for their depiction of students’ competencies and learning of those competencies (e.g., Hammer and Sikorski, 2015). Most crucially, concerns have been raised about learning progressions’ potential to perpetuate, and even exacerbate, educational inequities (e.g., Delgado and Morton, 2012; Myers, 2015).

In this article, I explore variations in and affordances of learning progressions, considering how different types of learning progressions are suited for different applications. I identify equity as a key challenge for learning progression research as the field moves from initial efforts to develop learning progressions to wider applications of these tools in classrooms and other educational spaces. Using examples of promising research, I call for additional research that is required if learning progressions are to have the envisioned impact on student learning.

Variation in learning progressions

There is significant variation in extant examples of learning progressions, including in the competencies they describe (see Table 3). However, learning progressions typically include some description of the resources for learning and/or nascent forms of the

Table 3 Examples of types of competencies described in learning progressions.

Type of competency	Example(s) of competency
Concept	• Mathematical equivalence (Fonger et al., 2018)
Practice	• Buoyancy (e.g., Gao et al., 2020; Paik et al., 2017; Yin et al., 2013) • Argumentation (e.g., Berland and McNeill, 2010; Bravo-Torija and Jiménez-Aleixandre, 2018; Osborne et al., 2016; van Rijn et al., 2014)
Ways of thinking or reasoning	• Modeling (Magana, 2017; Pierson et al., 2019; Schwarz et al., 2009) • Probabilistic thinking (Bai, 2020) • Statistical reasoning (Callingham et al., 2021) • Geocognition (Maeng, 2020)
Affective/volitional disposition	• Quantitative reasoning in science (Jin et al., 2019a; Mayes et al., 2014) • Epistemic agency (Taylor and Brand, 2021)
Combination	• Content and practices (e.g., Briggs and Furtak, 2019; Gotwals and Songer, 2013; Kaldaras et al., 2021; Wilson and Lehrer, 2021; Wyner and Doherty, 2017)

competency students are likely to have upon entering the progression (e.g., “lower anchor”), a description of the competency that is the goal for the progression (e.g., “upper anchor”), and portrayals of the intermediate competencies through which students might move as they progress from the lower to the upper anchor (e.g., Duschl et al., 2011; Jin et al., 2019b; Ketterlin-Geller et al., 2020). Traditionally, these descriptions are portrayed using one or more series of ordered levels (sometimes labeled *construct maps*, Wilson, 2009, or *progress variables*, e.g., Jin and Anderson, 2012); however, as discussed below (Section **Representing diversity**), other representations are being explored.

Over the years, others have characterized, and often classified, different types of learning progressions (e.g., Battista, 2011; Duschl et al., 2011; Lobato and Walters, 2017; Shavelson and Kurpius, 2012). Below, I highlight distinctions among learning progressions that are particularly important for their application (see **Applications of learning progressions** section) using four continua, instead of distinct categories.

Learning progression components

This continuum refers to variation in the nature of the components that make up a learning progression. On one end of this continuum, learning progressions are made up of smaller components (“building blocks”) that students must master or put together to reach the overall desired competency (e.g., Popham, 2007; Scoular, 2021). These learning progressions require attention to the structure of a competency, for example, what smaller components it comprises and which components are prerequisite to others. From this perspective, students accumulate the smaller building blocks to learn the overall desired competency.

On the other end of the continuum, learning progressions describe “successively more sophisticated” instantiations of a competency (e.g., Clements and Sarama, 2014; National Research Council, 2007, p. 219): a competency “in progress.” In this view, intermediate levels of a learning progression serve as stepping stones (Campbell et al., 2016; Wiser et al., 2012), which may include (or consist of) non-canonical ideas or practices important for developing more canonical versions of the competency (e.g., Duncan and Rivet, 2013). Learning progressions at this end of the continuum require attention to the way students engage with, and revise and/or refine, the competency as they learn.

Among the learning progressions in the middle of this continuum are those that combine the two perspectives, highlighting the ordering of components but also describing increasingly sophisticated instantiations of those components (e.g., Neumann et al., 2013). The knowledge integration approach (e.g., Lee and Liu, 2010) also falls in the middle of this continuum; in this perspective, increasing sophistication is the integration of increasing numbers of concepts (i.e., component building blocks).

Emphasis on ordering vs. pathway

Although most learning progressions include some motivation related to curriculum and instruction, that is ordering concepts so that children can learn most optimally (Allen, 2017), and encourage attention to “growth over status” (Briggs and Furtak, 2019, p. 146; Shepard, 2018), there is variation in the extent to which learning progressions are treated as pathways for learning. At one end of this continuum, researchers focus on the ordering of components, typically based on their (empirically determined) difficulty (e.g., Johnson, 2013; Kamei and Pavlovic, 2021; Liu, 2013). An underlying assumption of this approach is that the resulting ordered set of components will constitute a pathway for learning (e.g., Chiu and Wu, 2013); however, empirically derived ordering typically reflects the “status quo” instructional environment, which may not be optimal for student learning.

Researchers at the other end of the continuum focus on designing a pathway for learning, arguing that “the benchmarks of learning tend not to emerge unless someone carefully engineers and sustains the conditions that support them” (Lehrer and Schauble, 2012, p. 705). For example, design studies (e.g., Cardace et al., 2021; Clements et al., 2004; Taylor and Brand, 2021; Yoon et al., 2019) engineer instruction to support student learning of a competency and document, as a learning progression, the pathway that students take to learn that competency.

In between these two extremes, researchers may develop a learning progression by focusing on the difficulty of components and then work to refine the learning progression as a pathway for learning by using it to design instruction and documenting the student learning that results (e.g., [Castro-Faix et al., 2021](#)). Therefore, at different stages (e.g., initial development vs. refinement), work on a given learning progression may be situated at different points along the continuum.

Generalizability

In contrast to earlier work, such as Piaget's stages of development, "learning progressions are not developmentally inevitable but depend on instruction" ([Smith et al., 2006](#), pp. 5–6). At the same time, there may be (universal) developmental constraints ([Duncan and Gotwals, 2015](#)), and learning progressions differ in the extent to which specialized instruction is assumed necessary.

At one end of a generalizability continuum, many learning progressions are designed to be universally applicable. They are developed without much attention to context, under the assumption that "although some variability exists based on experience, culture, and instructional environments ... many students have similar patterns of development, learning, and thinking that can be generalized across the population" ([Ketterlin-Geller et al., 2019](#), p. 2).

At the other end of the continuum, acknowledging that teaching and learning are highly contextualized, some learning progressions describe learning under specific circumstances. Learning progressions of this type might apply to students engaging with a particular curriculum and/or students in a particular place (e.g., [Ladachart et al., 2020](#)). In these approaches, the learning progression may include not only the framework but also specific curriculum materials and/or other resources (e.g., [Fonger et al., 2018](#)). In the middle of this continuum is work that attends to context while trying to understand universal aspects of students' learning (e.g., [Taylor and Brand, 2021](#)).

Scope and grain size

As others have identified (e.g., [Gotwals, 2012](#)), learning progressions can be developed with varying scope (amount of time, whether expressed as ages or grade levels). For the most part, learning progressions with a broader scope have a larger grain size: the scope of individual components and/or the distance between them. Broader learning progressions generally contain less detailed accounts of, and larger shifts in, components, since grain size typically increases with increasing breadth to keep the number of levels manageable. For example, learning progressions associated with the *Next Generation Science Standards* ([NGSS Lead States, 2013](#)) describe students' competencies in four large grade bands: K–2, 3–5, 6–8, and 9–12.

At one end of the scope and grain size continuum, learning progressions cover many years (e.g., [Chang et al., 2020](#); [Hermann-Abell and DeBoer, 2018](#); [Kim et al., 2017](#); [Maras, 2021](#); [Mohan et al., 2009](#); [Paik et al., 2017](#); [Rich et al., 2018](#); [Smith et al., 2006](#)). At the other end of this continuum, learning progressions have a scope as narrow as a few weeks (e.g., [Alonzo and Steedle, 2009](#); [Yin et al., 2013](#)). In addition to learning progressions with scopes in between these two extremes, the middle of the continuum includes proposals to pair learning progressions with large and small grain sizes (e.g., [Fonger et al., 2018](#); [Gotwals, 2012](#)).

Applications of learning progressions

Since the early days of interest in "learning progressions," these frameworks have been promoted as tools to inform standards, curriculum, instruction, and assessment: both broad applications, such as achieving curricular coherence across multiple grades ([Smith et al., 2006](#); [Confrey et al., 2022](#)), and smaller-scale, classroom applications, such as support for teachers' formative assessment practices (e.g., [Black et al., 2011](#); [Clements et al., 2011](#); [Furtak, 2012](#); [Heritage, 2008](#); [Shepard, 2009](#); [Sztajn et al., 2012](#)). As described in the sub-sections below, many of these proposed applications are now coming to fruition. In this section, I provide a brief overview of some affordances of learning progressions, highlighting the type(s) of learning progressions well-suited to each application.

Standards, curriculum, and instruction

As others have noted, learning progressions are similar to [Bruner's \(2003\)](#) idea of a spiral curriculum (e.g., [Chang et al., 2020](#); [Huynh et al., 2015](#)), in which topics are revisited with increasing sophistication as students move through school. The ordering of components in learning progressions can support the development of standards and curricula so that students learn important building block ideas as they move toward desired competencies by the end of schooling (e.g., [Confrey and Maloney, 2014](#); [Foster and Wiser, 2012](#)). For example, development of new standards for mathematics ([National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010](#)) and science ([NGSS Lead States, 2013](#)) in the United States was informed by, and in terms of, broad learning progressions. Similarly, in Australia, learning progressions ([Australian Curriculum and Assessment and Reporting Authority, 2020a,b](#)) underpin the national curriculum.

Standards lay out a pathway for learning and may be informed by learning progressions that consider difficulty or complexity of components. To support coherence across grades and levels of schooling, learning progressions for this application have a large grain size. Standards are assumed to be broadly applicable (at least to students in the jurisdiction of the standards); therefore, learning progressions used to inform standards should be generalizable. In addition, because standards depict goals for student

learning, it is important that the components depicted are canonically accepted, rather than stepping stones that support development of canonical competencies but are themselves non-canonical (Foster and Wiser, 2012; see [Learning progression components](#) section).

In the classroom, a learning progression can influence instruction, whether through formal curriculum materials (such as those included with or as part of a learning progression) or teachers' own planning. The pathway laid out by a learning progression can serve as a guide when planning the specific learning experiences that will support movement within the learning progression. For example, Building Blocks (Clements and Sarama, 2007) and Carbon TIME (n.d.) are curricula designed to support movement through a learning progression. Both learning progressions that comprise building blocks and those that comprise increasingly more sophisticated or complex components can be used to inform instruction. To inform curriculum and instruction, learning progressions must have a small enough grain size to support the identification of specific instructional goals and strategies.

Formative (classroom) assessment

A highly touted affordance of learning progressions is their ability to support processes of formative assessment. A teacher could use a learning progression like a roadmap (e.g., Black et al., 2011): locating students' current competencies within the terrain of typical learning of the competency (e.g., Battista, 2011; Shepard, 2008) and, subsequently, using that location to identify appropriate support for students' movement from their current location toward mastery of the competency. This support could occur through targeted feedback (e.g., Furtak et al., 2018) or further instruction (e.g., Sztajn et al., 2012). By alerting teachers to potential student performances and, thus, influencing what ideas and practices are elicited and attended to, learning progressions might broaden assessment practices to include components characteristic of lower-performing students (Furtak, 2012; Sztajn et al., 2012; Wilson et al., 2013, 2015). To explore these affordances, researchers have developed and studied learning progression-aligned diagnostic assessment systems (e.g., Bennett, 2016; Confrey et al., 2017a; Confrey and Maloney, 2015; Daro et al., 2011; Wilson and Lehrer, 2021) and investigated how teachers make use of learning progressions to support formative assessment practices (e.g., Alonzo et al., 2022; Bailey and Heritage, 2014; Confrey et al., 2021; Covitt et al., 2018; Furtak et al., 2014).

For classroom use, learning progressions must capture "critical qualities of students' ways of thinking without being either so broad as to provide very little guidance for instruction and assessment, or so overwhelmingly fine-grained as to impede ready use" (Empson, 2011; Wilson and Lehrer, 2021, p. 2). The idea of learning progressions as pathways of learning is central to their use in formative assessment. Different types of learning progression components are useful for formative assessment, so long as the learning progression accurately describes the learning of students engaged in the formative assessment process. This could be accomplished either with a learning progression tailored to a particular environment or with a learning progression that can be generalized to that environment.

Summative (classroom and large-scale) assessment

In addition to informing formative assessment, learning progressions have the potential to inform summative assessment, at both classroom (e.g., Alonzo, 2019; Briggs et al., 2015) and large-scale (e.g., Alonzo et al., 2012) levels. Learning progressions describe a continuum of performance, so they can be used as the basis for criterion-referenced interpretations of students' performance and growth anywhere along the continuum (Briggs et al., 2015). Because students can be evaluated with respect to criteria laid out in the learning progression, rather than grade level expectations, assessments can be designed to capture a wide range of student performances, including those below grade level, and growth can be acknowledged, even when it does not result in fully meeting grade level expectations (e.g., Alonzo, 2019; Department of Education and Training, 2018; Waters, 2018).

Learning progressions used to inform summative assessment must be applicable to the students being assessed. At the classroom level, this means that the learning progression matches how learning is expected to unfold in that instructional context (i.e., specific to that context or generalizable across contexts); at the large-scale level, this means that learning progressions must be broadly generalizable (i.e., applicable to a wide range of learning contexts). In both cases, learning progressions should be at a grain size that matches the scope of the assessment: a classroom assessment covering a unit of instruction would have a smaller grain size than a national or international assessment covering a year or more of instruction. At the same time, reliability is crucial for summative assessment, and it is easier to characterize students' performance relative to a learning progression reliably when there are larger distinctions between components, such as levels, i.e., larger grain size. With smaller grain size, it becomes less likely that students respond at the same level across items or that graders consistently distinguish between responses at adjacent levels, since the difference between the levels is small.

Teachers' knowledge/tool for professional development

Teachers' understanding of a learning progression is key to its classroom use (e.g., Furtak et al., 2014; Heritage, 2011). At the same time, learning progressions can promote teachers' own learning: content knowledge (e.g., Wilson et al., 2013), pedagogical knowledge (e.g., Bennett, 2010); and pedagogical content knowledge (PCK; e.g., Gunkel et al., 2018; Sztajn et al., 2012), including content knowledge for teaching (Suh and Seshaiyer, 2015), awareness of student ideas and reasoning (e.g., Alonzo et al., 2022; Furtak, 2012; Ivars et al., 2020), and understanding of students' reasoning (e.g., Wilson et al., 2013) and learning (e.g., Bailey and Heritage, 2014). Researchers have also observed shifts in teachers' perspectives on teaching, learning, and/or academic performance (e.g.,

Alonzo et al., 2022; Furtak, 2012; Wilson et al., 2017). Teachers' learning may occur through their classroom use of learning progressions and/or their engagement with professional development (often provided to support learning progression use). The influence of learning progressions may be indirect, for example, when used to structure professional development experiences that lead to important shifts in practice without teachers' direct use of learning progressions (e.g., Alonzo et al., 2022; Furtak et al., 2015).

In addition, learning progressions may provide a conceptual framework that seeds teachers' own inquiries and knowledge generation. From this perspective, teachers do not learn only as they receive learning progressions from researchers but also as they make sense of these tools (e.g., Sztajn et al., 2012) in relation to their own practice. As models, LPs may, like models in science, be used to "notice and characterize phenomena that don't fit the model" and serve as a conceptual framework to ask and investigate questions (Alonzo and Elby, 2019, p. 6). For example, the typical pathway of student learning laid out in a learning progression may provide "a good starting point for deeper interrogation of learners' unique pathways" (Scoular, 2021, p. 118). Benefits of using learning progressions may even come not from teachers' direct use of these tools, but rather from the development of PCK and use of this knowledge to inform practice more generally (Alonzo and Elby, 2019; Parr, 2011). Teachers may benefit from learning progressions of all types. For example, they might explore vertical coherence through broader learning progressions (e.g., Birkhead et al., 2017) and explore details of students' competencies and learning through narrower learning progressions (Moreno et al., 2021).

Coherent educational systems

While learning progressions have potential for the individual applications described above, another central affordance is the ability to align multiple parts of the educational system to a learning progression to achieve greater coherence among standards, curriculum, instruction, assessment, and teacher professional development (e.g., Bennett, 2010; Black et al., 2011; Confrey and Maloney, 2014; Department of Education and Training, 2018; Jin et al., 2019b; National Research Council, 2006; Simeon, 2021). Such alignment among parts of the educational system is essential to visions of personalized instruction (e.g., Achieve, 2015; Department of Education and Training, 2018; Hughes et al., 2020), where students are engaged with learning opportunities matched to their current competencies, rather than age or grade level. In addition to the nationwide effort ongoing in Australia (e.g., Department of Education and Training, 2018), the diagnostic assessment systems described above (Section **Formative (classroom) assessment**) are examples of these ideas being put into practice.

Because of the variation in learning progressions required for different parts of the educational system, nested learning progressions will likely be needed, in which smaller learning progressions, which meet the requirements of classroom-level work, are embedded within larger learning progressions that meet the requirements of large-scale work (Gotwals, 2012). To achieve this vision, researchers (e.g., Confrey and Maloney, 2014; Lehrer et al., 2014) have begun to conceptualize learning progressions as boundary objects that facilitate work across the different communities of practice in an educational system.

Equity critiques of learning progressions

While learning progressions have generated interest because of their potential and, to some extent demonstrated ability, to support many parts of an educational system responsive to students' learning needs, they have also been subject to important critiques (e.g., Empson, 2011; Hammer and Sikorski, 2015; Pierson and Clark, 2019; Shavelson and Kurpius, 2012; Sikorski, 2019). As models of complex phenomena, learning progressions "do not just illuminate, they also eliminate information" (Alonzo and Elby, 2019; Corcoran et al., 2009; Empson, 2011; Lehrer and Schauble, 2015, p. 731).

Although there are many implications of omitting information about students' competencies and learning, I focus on equity implications in this article. In particular, learning progressions illuminate common patterns in students' competencies (e.g., Bailey and Heritage, 2014; Battista, 2011; Chiu and Wu, 2013; Gunkel et al., 2012); however, by reducing the complexity of the student competencies they represent, learning progressions inevitably leave out ways of knowing, reasoning, engaging with disciplinary practices, and learning (e.g., Duncan and Gotwals, 2015; Empson, 2011; Hammer and Sikorski, 2015; Lombard et al., 2018; Metz, 2009; Pierson et al., 2017; Salinas, 2009). In this section, I argue that such omissions have the potential to render some students less visible, less recognized, and less likely to benefit from learning progression-aligned instruction and that these impacts are especially likely for students from marginalized communities, who often are excluded from the research used to develop learning progressions (e.g., Delgado and Morton, 2012; Williams, 2018). Thus, rather than contributing to better learning outcomes for all students, as some have argued (e.g., Corcoran et al., 2009), learning progressions could create greater inequities.

In this section, I explore the consequences of reducing two types of complexity to create aggregate descriptions: variation in students' competencies and variation in students' learning. In particular, I focus on what happens when these aggregate descriptions omit alternatives, in terms of both the content and ordering of components of the learning progression.

Content of learning progression components

A learning progression foregrounds a set of building blocks or increasingly sophisticated ways of engaging with a competency and, necessarily, backgrounds others. By illuminating the components of a competency that precede full mastery, learning progressions provide insight into stepping stones in (and building blocks of) students' learning. Awareness of these components underlies key applications of learning progressions from informing large-scale standards, curriculum, and assessment to influencing teachers' knowledge and classroom assessment practices. Thus, students whose developing competency does not match the components

of a learning progression may experience exclusion at multiple levels of the educational system. Such exclusion is particularly problematic, given that learning progressions are designed to support students in using prior ideas, practices etc., as the foundation for learning new ones (e.g., Schramm et al., 2018) and because learning progressions may treat “students’ place-based experience, indigenous knowledge, and traditional ways of thinking” as noise (Ladachart et al., 2020, p. 995).

Standards, curricula, and instruction

Learning progression-informed standards, curricula, and instruction are designed to provide bridges from one component to the next in order to support students in developing new competencies on the basis of their prior ones. When the components of students’ developing competency do not match those in a learning progression, students’ opportunity to leverage their own resources may be curtailed. This exclusion may be particularly problematic toward the beginning of a learning progression, where the “lower anchor” describes the expected foundation for learning a competency (e.g., Duncan and Hmelo-Silver, 2009b).

A metaphor is useful here: learning progressions are “formations of stepping stones” across a river, where the goal is to help “students move from their current state of knowledge and skill to a learning goal on the other side of the river” (Achieve, 2015, p. 2). Here, and below, I use “stepping stones” to refer to both building blocks and less sophisticated versions of a competency. Extending the metaphor, the left side of Fig. 1 shows instruction providing a bridge from the bank of the river to the first stepping stone, as well as between stepping stones. However, a student down the river (right side of Fig. 1) has no bridge to get to the first stepping stone and must either jump from the bank into the river, hoping to hit a stepping stone, or travel down the riverbank to the starting point and cross the bridge from there. In both cases, the student is at a disadvantage: having to make a big leap to the first stepping stone and perhaps falling short (and into the river) or needing to travel a large distance on land before jumping to the first stepping stone (and potentially arriving after the rest of the class has moved on).

Assessment

Learning progressions have the potential to inform day-to-day student-teacher interactions by focusing teachers’ attention on significant features of students’ learning. However, there is a danger that learning progressions may narrow focus on those features, causing others to be missed (e.g., von Aufschnaiter and Alonzo, 2018). Continuing with the river metaphor, if teachers are only looking for students at the stepping stones represented in a learning progression, they may miss students like the one in Fig. 1 who is attempting to cross from a different starting point or, having ventured from the bank, is stuck in the middle of the river on a potentially productive stepping stone not included in the learning progression. If these students and their nascent competencies are not recognized, their ideas may not be highlighted in class discussions (e.g., Empson, 2011; Furtak, 2012; Sztajn et al., 2012) or taken up in other forms of instruction.

In addition, narrowed attention has the potential to result in misinterpretation of student achievement and learning needs. For example, despite their potential utility for large-scale assessment (e.g., Alonzo et al., 2012) and/or automated diagnostic/formative assessment (e.g., Chin et al., 2021b), ordered multiple-choice (OMC; Briggs et al., 2006) items may provide misdiagnoses of students whose developing competencies are not represented in the option choices and who are forced to choose an option that does not represent their engagement with the competency. When assessment results are then used to make instructional decisions, students whose competencies have been misdiagnosed are less likely to receive appropriate instruction.

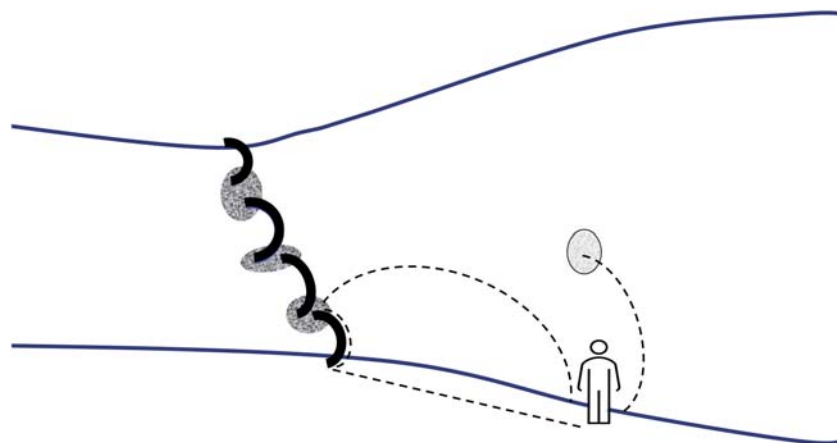


Fig. 1 River metaphor showing consequences of a mismatch between a student's starting position and that assumed in a learning progression
 Note: The pathway on the left is represented in the learning progression and supported by instruction (represented by the thick black lines). The student starts from a different point than that assumed in the learning progression and must take one of the three paths indicated by the thinner dashed lines.

Ordering of learning progression components

Layered on top of what components are included/excluded is the ordering of those components. Learning progressions “represent overarching aggregate arcs in the curriculum, which are rarely replicated in individual students’ context-dependent and non-linear learning trajectories” (Pierson et al., 2019, p. 834). Even if students’ learning can be described accurately by the components of a learning progression, if their pathways through those components are not represented, opportunities to learn may be diminished: both directly, through standards, curriculum, and instruction, and indirectly, through assessment.

Standards, curricula, and instruction

Laying out standards or curricula requires identifying one ordering (of content for students at different grade levels in standards, of lessons in a curriculum), based on the prediction of a pathway that will support as many students as possible to reach the upper anchor of the learning progression. However, learning unfolds through complex interactions with students’ prior competencies and resources for learning, interactions with teachers and peers, and other features of the learning environment (Eames et al., 2021). What is optimal for some students may be much less successful for others. For example, when learning progressions are created to portray cross-sectional differences in student achievement under status quo conditions, the resulting pathway is reflective of current instruction, which is often less successful for students from marginalized groups. There is also a danger that instruction will steer students along the pathway set out by the learning progression, at the expense of opportunities for students to “express, test, and revise their own ways of thinking” (Lesh and Yoon, 2004, p. 205; Sikorski and Hammer, 2017).

Assessment

Students’ achievements may be misidentified if their pathways of learning differ from that expected in the learning progression (e.g., Shavelson and Kurpius, 2012). Such interpretations, if used formatively, may result in students receiving inappropriate instruction. As a simple example, suppose that, based on prior learning and/or current circumstances, a student’s pathway should be “out of order” (e.g., A, C, B, D instead of A, B, C, D). Formatively, if the student’s responses to assessment tasks indicate mastery of component C, a teacher might assume that the student needs support to move from C to D. However, the student may struggle if that instruction builds on component B, which the student has not yet learned (see Fig. 2). Summatively, student responses that indicate difficulty with component B might lead to an assumption that the student has only mastered component A, not recognizing achievement with respect to component C.

Current and future research to explore and address equity concerns with and through learning progressions

Most researchers acknowledge that there are likely multiple pathways through the learning landscape for a particular competency (e.g., Battista, 2011; Smith et al., 2006). At the same time, many learning progressions are represented as single pathways with fixed components. In the previous section, I discussed what we might miss when (and the potential impacts of) trying to represent learning this way. In this section, I explore the possibilities when researchers respond to diversity: of students, their learning contexts, and the way they learn and demonstrate competencies. I highlight promising work, using it as a springboard to look forward to expanding these efforts.

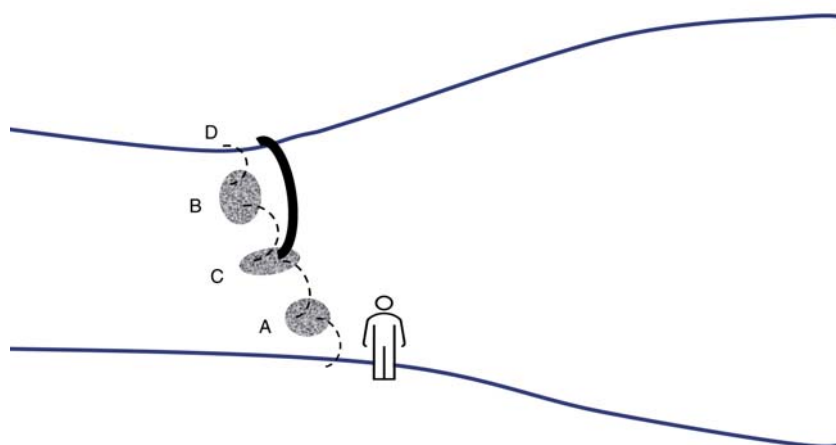


Fig. 2 River metaphor showing consequences of a mismatch between a student’s pathway for learning (ordering of components) and that assumed in a learning progression Note: The student’s expected pathway is represented by the thin dashed lines connecting stepping stones A, C, and B and the targeted competency (D). Instruction (indicated by the thick black lines) is expected to support the relatively small movement from C to D in the standard pathway, rather than the bigger jump necessitated by skipping over B in the student’s pathway.

Central to my call for future research is a need to understand the equity problem with learning progressions. The discussion above (Section **Equity critiques of learning progressions**) purposefully did not address the severity of divergence we might expect between a learning progression and students' competencies and learning. Across the work described in the sub-sections below, I ask the following overarching questions:

- When do patterns in students' competencies and learning represented in learning progressions seem widely applicable (generalizable across students and learning contexts)?
- When are students' competencies and learning not fully represented in learning progressions?
 - What divergences appear?
 - How significant are the divergences? For how many students?
 - Are there any patterns in the divergences (e.g., across types of learning progressions, subject areas, or groups of students)?

Better understanding of what is being eliminated when learning progressions illuminate common patterns is essential to moving forward to address potential equity concerns.

Diversity of learning contexts

As mentioned briefly above (Section **Generalizability**), one source of variation across learning progressions is the extent to which they are designed to be generalizable across contexts and populations. Increasingly, learning progression research attends to the environment in which learning is expected to occur.

Providing information about learning context

One way to begin exploring learning progressions' generalizability is to be more explicit about the contexts in which they were developed. For example, curricula and/or other instructional supports are an important form of contextualization of the learning environment. When learning progression development occurs in tandem with curriculum development and/or design work (e.g., [Fonger et al., 2018](#); [Lehrer et al., 2014](#)), descriptions of the curricular context usually accompany the learning progression. Specially designed curriculum is often needed for students to move through a learning progression (e.g., [Plummer and Maynard, 2014](#)). However, where status quo instruction is assumed, information is needed about what that entails, as "status quo" is context-dependent. A better understanding of the curricular and instructional context of a given learning progression is an important step toward being able to explore how much such contexts impact the learning progressions that are developed.

On a broad scale, there are important differences across countries in terms of standards, curriculum, and learning environments (e.g., [Kelly et al., 2020](#)), and research on learning progressions is being conducted in many countries. However, the resulting progressions are often presented as if they were universally applicable. In a rare example of cross-national work, [Yao et al. \(2017\)](#) refined a learning progression for energy developed in Germany ([Neumann et al., 2013](#)), using data from Chinese students; however, they did not treat discrepancies between the original learning progression and their new data as potentially revealing distinctive features of Chinese students' learning. Similarly, articles reporting data from multiple countries do not necessarily consider potential differences across the sample (e.g., [Lai et al., 2017](#); [Wilkins and Norton, 2018](#)).

However, some researchers have begun to acknowledge and investigate cross-national differences in student learning. For example, while not restricting their learning progression to the Chinese context in which it was developed, [Gao et al. \(2020\)](#) provided two explanations for why a Chinese learning progression for buoyancy might differ those developed in other contexts: the meaning of the word "buoyancy" in the Chinese language and the way buoyancy is presented in Chinese textbooks (in terms of force). Such targeted descriptions might serve as a step toward more explicitly cross-national work, for example, comparison between learning progressions for Chinese students and learning progressions for students who learn buoyancy in terms of a property (density) of the object. There are already a few examples of such cross-national work in the literature, such as [Jin et al.'s \(2009\)](#) and [Liu and Tang's \(2004\)](#) investigations of differences between students in China and in the United States and in Canada, respectively.

Like [Confrey \(2019\)](#), I see a need for additional cross-cultural work to investigate how differences in cultural context (e.g., country) impact different types of learning progressions. For some competencies, multiple learning progressions have been developed, sometimes in different countries (see [Table 2](#)). To what extent do differences reflect different ways of developing the learning progressions vs. differences in the cultural contexts in which they were developed? Researchers might investigate similarities and differences when a common learning progression is modified in response to evidence collected in different countries (or other types of contexts). At the country level, it would be particularly important to involve researchers from lower income countries, as learning progression research to date has been conducted in higher income countries: those classified as "high income" or "upper middle income" by [The World Bank \(2021\)](#).

Developing context-specific learning progressions

In the previous sub-section, I described aspects of the learning environment (curriculum/instruction and national context) that directly impact what happens in the classroom. However, learning is also influenced by the experiences and other resources students bring to the classroom, as well as by their communities. Here, I see a special role for localized learning progressions. For example, [Ladachart et al.'s \(2020\)](#) highly place-based learning progression describes students' explanations for haze pollution, recognizing the importance of haze pollution in northern Thailand and students' experiential knowledge of this phenomenon. Learning

progressions like this are purposefully not generalizable. However, deep attention to place and place-based learning can shed light on the ways that learning progressions may be influenced by place.

Some researchers have suggested that learning progressions are intended to be broad frameworks that teachers will adapt to their own contexts (e.g., [Lehrer and Schauble, 2015](#)). Yet research is needed to explore how teachers adapt learning progressions in this way. Although [Furtak and Heredia \(2014\)](#) explored how teachers in one school took up a learning progression developed by teachers in another school, they did not explore/report adaptations. Systematic exploration of teachers' adaptations in different contexts and with different students would also help to characterize the extent and nature of context-based variation in learning progressions. In addition to individual variations (discussed below in Section [Student diversity](#)), the students in a class impact the instructional environment and, thus, are an important consideration in instructional context. As those "on the ground," teachers may be best positioned to respond to this diversity by identifying ways that their students' learning differs from what is represented in a learning progression.

Student diversity

In learning progression research, it is typical to reduce empirical evidence such that differences between students are smoothed away (e.g., [Battista, 2011](#); [Pierson et al., 2019](#)) and patterns of competencies not reflective of the group as a whole are considered noise or "misfit." When collecting evidence to validate and/or refine a learning progression, researchers attend to variation in their data; however, this variation is typically used to make revisions to capture common ideas not already represented in the learning progression (e.g., [Castro-Faix et al., 2021](#)), rather to describe less common alternatives in an expanded version of the learning progression. Therefore, while such attention reflects an important step, I echo the call from [Hammer and Sikorski \(2015\)](#) for "approaches to research that can capture and consider idiosyncrasy, not as statistical noise to eliminate but as part of the phenomena of interest" (p. 429). This will require greater attention to individual students, particularly those from marginalized groups, as more than "misfit" to statistical models. In the sub-sections below, I describe and call for work to explore, represent, and respond to student diversity.

Exploring student diversity

Although curriculum development-based design work often produces a single learning progression, these approaches include detailed engagement with individual students' competencies over time. As such, they are well-positioned to contribute rich descriptions of how individual students engage with the curriculum and make progress with respect to the learning progression. As [Alonzo and Elby \(2019\)](#) suggest, learning progressions (as imperfect models) could be used as springboards for teachers' inquiry. Collectively, this work could yield important insights about alternative learning progression components and pathways.

Detailed qualitative work could be complemented by promising new quantitative approaches to explore diversity in students' learning. Cognitive Diagnostic Models have the potential to generate multiple pathways during learning progression development and validation. For example, [Gao et al. \(2021\)](#) used latent class analyses and the rule space model to identify multiple pathways (i.e., ordering of learning progression components) that students might take in learning a competency. In a more novel approach, [Chiu and Wu \(2013\)](#) have used a classification method from biology (*cladistics* or *phylogenetic classification*) to create a learning progression in the form of a *conceptual evolution tree* (CET). Instead of representing the evolution of organisms, the CET shows "evolutionary pathways for the development of mental models" (p. 375).

Additional research (using these and other approaches) is needed to attend to and document diversity in students' competencies and learning. This is important for characterizing the different ways that learning progressions might fail to capture students' competencies and learning, and for producing learning progressions that exclude fewer students.

Work with marginalized populations

While there are many reasons students' learning might differ from what is depicted in a learning progression, equity concerns demand that we pay particular attention to students from marginalized groups. Learning progressions intended to be "universal" are most likely to be applicable to students who are not marginalized in their particular contexts (e.g., ethnically/racially part of a dominant group, fluent in the language of instruction and assessment, having no disabilities, having medium to high socioeconomic status). Indeed, most learning progression research has been conducted with such students (e.g., [Delgado and Morton, 2012](#); [Williams, 2018](#); [Zahner and Wynn, 2021](#)). There are, however, exceptions, as described below.

Some researchers have purposefully developed learning progressions through work with marginalized students (e.g., [Taylor and Brand, 2021](#)); others have sought to include these students, along with others, in their studies (e.g., [Delgado and Morton, 2012](#); [Songer et al., 2009](#)). The latter type of work has allowed researchers to make comparisons between sub-groups, such as students from urban vs. rural schools or children with different ethnic/racial backgrounds, based on performance relative to a learning progression (e.g., [Yao et al., 2017](#)) and/or growth using learning progression-aligned curriculum and instruction (e.g., [Clements and Sarama, 2008](#); [Delgado and Morton, 2012](#)). However, there have been few efforts to investigate whether a learning progression differs between groups of students. Additional work is needed to explore not just how much (relative to a learning progression) students have learned, but how the components and pathway of a learning progression might differ for marginalized students.

Recent work has also explored learning progressions for students with particular learning needs. [Bailey and Heritage \(2014\)](#) developed a learning progression for emergent bilingual learners. The learning progression describes the way these students learn language, both for use in its own right and to support the learning of other content. Similarly, [Kamei and Pavlovic \(2021\)](#) and

Strickland (2018) report work on learning progressions describing foundational competencies required for students with intellectual disabilities, developmental disabilities, and/or autism spectrum disorder to “access the general curriculum” (Kamei and Pavlovic, 2021, p. 2). Further work is needed to explore other prerequisite competencies. Work is also needed to explore how these students draw on prerequisite competencies for further learning, as it is likely that the learning pathways for these students differ substantially from those developed from and for their mainstream counterparts.

Representing diversity

As researchers have grappled with the complexity of representing student learning, representations have been generated to express that students’ learning is not as simple as portrayed in a series of linear levels. Expanding from the portrayal of a single, linear pathway, researchers in the Teach to One project represented components of a learning progression as connected by a wide array of possible pathways (Achieve, 2015), and Confrey et al. (2017b) characterize student thinking like a climbing wall, where multiple handholds can be used to reach the top (rather than like a ladder, where there is a single sequence of rungs). Sevan and Talanquer (2014) use a potential energy surface to represent the conceptual landscape of a learning progression; they identify regions of relative stability (i.e., components) but do not arrange these in a sequence or pathway. I encourage both further work of this type and representations that might capture greater diversity in the components of a learning progression and not just flexibility in the pathway through those components.

Responding to diversity

In addition to documenting variation in students’ competencies and learning, it is important to document how this variation plays out in the classroom and other contexts (e.g., large-scale assessment) and, more broadly, how the use of learning progressions impacts marginalized students. For example, a teacher in Myers (2015) used a learning progression to justify deficit views of students. However, Suh et al. (2020, p. 229) found that learning progression-based lesson study “facilitated ... equitable access to instruction” and asset-based views of students’ thinking, and Myers et al. (2015) have developed a promising framework for how learning progression-based instruction could promote equity. Their framework could serve as a strong foundation for further research investigating how teachers use learning progressions in practice and how their use of these tools can be made more equitable. For example, Bernardo (2017) started to bring the learning progression construct into dialog with culturally responsive teaching. As another example, Graf et al. (2021) are developing a learning progression with two strands aligned with curriculum modules from the Algebra Project (Silva et al., 1990), which is designed to support students from historically marginalized communities. Further research could explore questions such as: What does it look like for a learning progression to be used in the context of culturally responsive pedagogy? Can learning progressions (especially those that acknowledge students’ diversity) be used to support teachers in noticing all students’ competencies, not just those represented in the learning progression? What support do teachers need to use learning progressions in equitable ways?

Conclusion

Learning progressions research has flourished over the past two decades, with a plethora of research to develop, validate, and study the use of learning progressions across grade levels, subject areas, and countries. These learning progressions represent diversity in terms of the type of competency they describe, the nature of their components, their emphasis on ordering or pathways for learning, their generalizability, and their grain size.

The good news is that learning progressions are beginning to fulfill their envisioned promise to inform standards, curriculum, instruction, formative and summative assessment, and teacher learning. The bad news is that there is a potential for these uses to exacerbate, rather than ameliorate, educational inequities. Work to address potential equity problems will take coordinated effort. Answers to the questions posed at the beginning of the **Current and future research to explore and address equity concerns with and through learning progressions** section will require researchers to work together to study variation in learning progressions across contexts and groups of students, attending to unique characteristics of teaching and learning in each case and to the nuanced ways that students’ learning may differ from that depicted in a “standard” learning progression. There is significant work being undertaken on learning progressions. This work could be even more powerful, and have less inequitable consequences, if researchers can find ways to coordinate and attend more closely to the contexts of their work.

References

- Achieve, 2015. The Role of Learning Progressions in Competency-Based Pathways. <https://www.achieve.org/files/Achieve-LearningProgressionsinCBP.pdf>.
- Allen, M., 2017. Early understandings of simple food chains: a learning progression for the preschool years. *Int. J. Sci. Educ.* 39 (11), 1485–1510. <https://doi.org/10.1080/09500693.2017.1336809>.
- Alonzo, A.C., Elby, A., 2019. Beyond empirical adequacy: learning progressions as models and their value for teachers. *Cognit. Instruct.* 37 (1), 1–37. <https://doi.org/10.1080/07370008.2018.1539735>.
- Alonzo, A.C., Gotwals, A.W. (Eds.), 2012. *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers. <https://doi.org/10.1007/978-94-6091-824-7>.
- Alonzo, A.C., Steedle, J.T., 2009. Developing and assessing a force and motion learning progression. *Sci. Educ.* 93 (3), 389–421. <https://doi.org/10.1002/sce.20303>.

- Alonzo, A.C., Neidorf, T., Anderson, C.W., 2012. Using learning progressions to inform large-scale assessment. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, pp. 211–240. https://doi.org/10.1007/978-94-6091-824-7_10.
- Alonzo, A.C., Wooten, M.W., Christensen, J.A., 2022. Learning progressions as a simplified model: examining teachers' reported uses to inform classroom assessment practices. *Sci. Educ.* 106 (4), 852–889. <https://doi.org/10.1002/sce.21713>.
- Alonzo, A.C., 2018. Formative assessment with science learning progressions. *Appl. Meas. Educ.* 31 (2).
- Alonzo, A.C., 2019. Defining trustworthiness for teachers' multiple uses of classroom assessment results. In: Brookhart, S.M., McMillan, J.H. (Eds.), *Classroom Assessment and Educational Measurement*. Routledge, pp. 120–145. <https://doi.org/10.4324/9780429507533-8>.
- Australian Curriculum, Assessment and Reporting Authority, 2020a. National Literacy Learning Progression (Version 3.0). Australian Curriculum, Assessment and Reporting Authority. <https://www.ofai.edu.au/media/01nixkio/national-literacy-progressions-v3-for-publication.pdf>.
- Australian Curriculum, Assessment and Reporting Authority, 2020b. National numeracy learning progression (Version 3.0). Australian Curriculum, Assessment and Reporting Authority. <https://www.ofai.edu.au/media/iwbecoj/national-numeracy-progression-v3.pdf>.
- Bai, S., 2020. Developing a learning progression for probability based on the GDINA model in China. *Front. Psychol.* 11 (569852). <https://doi.org/10.3389/fpsyg.2020.569852>.
- Bailey, A.L., Heritage, M., 2014. The role of language learning progressions in improved instruction and assessment of English language learners. *Tesol Q.* 48 (3), 480–506. <https://doi.org/10.1002/tesq.176>.
- Bargagliotti, A.E., 2021. Developing learning trajectories for teacher learning. In: Jenlink, P.M. (Ed.), *Learning to Teach: Curricular and Pedagogical Considerations for Teacher Preparation*. Rowman & Littlefield, pp. 11–30.
- Battista, M.T., 2011. Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *Math. Enthus.* 8 (3), 507–570. <https://doi.org/10.54870/1551-3440.1228>.
- Bennett, R.E., 2010. Cognitively based assessment of, for, and as learning (CBAL): a preliminary theory of action for summative and formative assessment. *Measurement* 8 (2–3), 70–91. <https://doi.org/10.1080/15366367.2010.508686>.
- Bennett, R.E., 2016. From cognitive-domain theory to assessment practice. *Educ. Psychol.* 51 (1), 82–107. <https://doi.org/10.1080/00461520.2016.1141683>.
- Berland, L.K., McNeill, K.L., 2010. A learning progression for scientific argumentation: understanding student work and designing supportive instructional contexts. *Sci. Educ.* 94 (5), 765–793. <https://doi.org/10.1002/sce.20402>.
- Bernardo, C., 2017. Integration of Culturally Relevant Pedagogy Into the Science Learning Progression Framework (Publication No. 10275105) [Doctoral Dissertation], Columbia University [ProQuest Dissertations and Theses Global].
- Bernholt, S., Sevan, H., 2018. Learning progressions and teaching sequences. *Chem. Educ. Res. Pract.* 19 (4).
- Birkhead, S., Suh, J., Gerasimova, D., Seshaiyer, P., 2017. Improving knowledge of algebraic learning progressions through professional learning in collaborative vertical teams. In: Galindo, E., Newton, J. (Eds.), *Proceedings of the 39th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Hoosier Association of Mathematics Teacher Educators, pp. 415–422.
- Black, P., Wilson, M., Yao, S.-Y., 2011. Road maps for learning: a guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research and Perspectives* 9 (2–3), 71–123. <https://doi.org/10.1080/15366367.2011.591654>.
- Blömeke, S., Gustafsson, J.-E., Shavelson, R.J., 2015. Beyond dichotomies: competence viewed as a continuum. *Z. Psychol.* 223 (1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>.
- Bravo-Torija, B., Jiménez-Aleixandre, M.-P., 2018. Developing an initial learning progression for the use of evidence in decision-making contexts. *Int. J. Sci. Math. Educ.* 16 (4), 619–638. <https://doi.org/10.1007/s10763-017-9803-9>.
- Briggs, D.C., Furtak, E.M., 2019. Learning progressions and embedded assessment. In: Brookhart, S.M., McMillan, J.H. (Eds.), *Classroom Assessment and Educational Measurement*. Routledge, pp. 146–169. <https://doi.org/10.4324/9780429507533-9>.
- Briggs, D.C., Alonzo, A.C., Schwab, C., Wilson, M., 2006. Diagnostic assessment with ordered multiple-choice items. *Educ. Assess.* 11 (1), 33–63. https://doi.org/10.1207/s15326977ea1101_2.
- Briggs, D.C., Diaz-Bilello, E., Peck, F., Alzen, J., Chattergoon, R., Johnson, R., 2015. Using a Learning Progression Framework to Assess and Evaluate Student Growth. Center for Assessment Design Research and Evaluation, University of Colorado, Boulder. <https://files.eric.ed.gov/fulltext/ED561889.pdf>.
- Bruner, J.S., 2003. *The Process of Education: A Landmark in Educational Theory*. Harvard University Press.
- Callingham, R., Watson, J., Oates, G., 2021. Learning progressions and the Australian curriculum: the case of statistics and probability. *Aust. J. Educ.* 65 (3), 329–342. <https://doi.org/10.1177/00049441211036521>.
- Campbell, T., Schwarz, C., Windschitl, M., 2016. What we call misconceptions may be necessary stepping-stones toward making sense of the world. *Sci. Scope* 53 (7), 28–33.
- Carbon TIME (n.d.). Carbon TIME: Unit Synopses. https://carbontime.bsccs.org/sites/default/files/educator_resources/unit_synopses.pdf.
- Cardace, A., Wilson, M., Metz, K.E., 2021. Designing a learning progression about micro-evolution to inform instruction and assessment in elementary science. *Educ. Sci.* 11 (609). <https://doi.org/10.3390/educsci11100609>.
- Carpenter, T.P., Moser, J.M., 1984. The acquisition of addition and subtraction concepts in grades one through three. *J. Res. Math. Educ.* 15 (3), 179–202. <https://doi.org/10.2307/748348>.
- Castro-Faix, M., Duncan, R.G., Choi, J., 2021. Data-driven refinements of a genetics learning progression. *J. Res. Sci. Teach.* 58 (1), 3–39. <https://doi.org/10.1002/tea.21631>.
- Chang, C.H., Tan, G.C.I., Tan, J., Kwek, C.H., 2020. Developing a Learning Progression for Climate Change in Geography Education [Technical Report, Office of Education Research]. National Institute of Education, Singapore.
- Chen, F., Yan, Y., Xin, T., 2017. Developing a learning progression for number sense based on the rule space model in China. *Educ. Psychol.* 37 (2), 128–144. <https://doi.org/10.1080/01443410.2016.1239817>.
- Chin, H., Chew, C.M., Lim, H.L., 2021a. Development and validation of a cognitive diagnostic assessment with ordered multiple-choice items for “multiplication of time”. *J. Comput. Educ.* 8 (2), 289–316. <https://doi.org/10.1007/s40692-020-00180-7>.
- Chin, H., Chew, C.M., Lim, H.L., Thien, L.M., 2021b. Development and validation of a cognitive diagnostic assessment with ordered multiple-choice items for addition of time. *Int. J. Sci. Math. Educ.* <https://doi.org/10.1007/s10763-021-10170-5>.
- Chiu, M.-H., Wu, W.-L., 2013. A novel approach for investigating students' learning progression for the concept of phase transitions. *Educ. Quim.* 24 (4), 373–380. [https://doi.org/10.1016/S0187-893X\(13\)72490-X](https://doi.org/10.1016/S0187-893X(13)72490-X).
- Clements, D.H., Sarama, J., 2004. Learning trajectories in mathematics education. *Math. Think. Learn.* 6 (2).
- Clements, D.H., Sarama, J., 2007. *Building Blocks Curriculum, Grade preK*. SRA/McGraw Hill.
- Clements, D.H., Sarama, J., 2008. Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *Am. Educ. Res. J.* 45 (2), 443–494. <https://doi.org/10.3102/0002831207312908>.
- Clements, D.H., Sarama, J., 2014. Learning trajectories: foundations for effective, research-based education. In: Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing, pp. 1–30.
- Clements, D.H., Wilson, D.H., Sarama, J., 2004. Young children's composition of geometric figures: a learning trajectory. *Math. Think. Learn.* 6 (2), 163–184. https://doi.org/10.1207/s15327833mtl0602_5.
- Clements, D.H., Sarama, J., Spitler, M.E., Lange, A.A., Wolfe, C.B., 2011. Mathematics learned by young children in an intervention based on learning trajectories: a large-scale cluster randomized trial. *J. Res. Math. Educ.* 42 (2), 127–166. <https://doi.org/10.5951/jresmetheduc.42.2.0127>.
- Colantonio, A., Galano, S., Leccia, S., Paddu, E., Testa, I., 2018. Design and development of a learning progression about stellar structure and evolution. *Phys. Rev.* 14 (010143). <https://doi.org/10.1103/PhysRevPhysEducRes.14.010143>.

- Confrey, J., Maloney, A.P., 2014. Linking standards and learning trajectories. In: Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing, pp. 125–160.
- Confrey, J., Maloney, A., 2015. A design research study of a curriculum and diagnostic assessment system for a learning trajectory on equipartitioning. *ZDM Math. Educ.* 47 (6), 919–932. <https://doi.org/10.1007/s11858-015-0699-y>.
- Confrey, J., Maloney, A.P., Nguyen, K.H., Rupp, A.A., 2014. Equipartitioning: a foundation for rational number reasoning: elucidation of a learning trajectory. In: Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing, pp. 61–96.
- Confrey, J., Gianopoulos, G., McGowan, W., Shah, M., Belcher, M., 2017a. Scaffolding learner-centered curricular coherence using learning maps and diagnostic assessments designed around mathematics learning trajectories. *ZDM Math. Educ.* 49 (5), 717–734. <https://doi.org/10.1007/s11858-017-0869-1>.
- Confrey, J., Maloney, A., Gianopoulos, G., 2017b. Untangling the “messy middle” in learning trajectories. *Measurement: Interdisciplinary Research and Perspectives* 15 (3–4), 168–171. <https://doi.org/10.1080/15366367.2017.1404370>.
- Confrey, J., Shah, M., Belcher, M., 2021. Using learning trajectory-based ratio curriculum and diagnostic assessments for promoting learner-centred instruction. In: Barkatsas, T., McLaughlin, P. (Eds.), *Authentic Assessment and Evaluation Approaches and Practices in a Digital Era: A Kaleidoscope of Perspectives*. Brill, pp. 187–216. https://doi.org/10.1163/9789004501577_009.
- Confrey, J., Shah, M., Maloney, A., 2022. Learning trajectories for vertical coherence. *Math. Teach.: Learn. Teach.* 115 (2), 90–103. <https://doi.org/10.5951/MTLT.2021.0012.PK-12>.
- Confrey, J., 2019. A Synthesis of Research on Learning Trajectories/Progressions in Mathematics. Organisation for Economic Co-operation and Development.
- Corcoran, T., Mosher, F.A., Rogat, A., 2009. Learning Progressions in Science: An Evidence-Based Approach to Reform (CPRE Research Report #RR-63). Consortium for Policy Research in Education, Philadelphia, PA, U.S. <https://doi.org/10.12698/cpre.2009.r63>.
- Covitt, B., Gunckel, K.L., Syswerda, S., Caplan, B., 2018. Teachers' use of learning progression-based formative assessment in water instruction. *Appl. Meas. Educ.* 31 (2), 128–142. <https://doi.org/10.1080/08957347.2017.1408627>.
- Daro, P., Mosher, F.A., Corcoran, T., 2011. Learning Trajectories in Mathematics: A Foundation for Standards, Curriculum, Assessment, and Instruction (CPRE Research Report #RR-68). Consortium for Policy Research in Education. <https://doi.org/10.12698/cpre.2011.r68>.
- Delgado, C., Morton, K., 2012. Learning progressions, learning trajectories, and equity. In: van Aalst, J., Thompson K., J., Jacobson, M.J., Reimann, P. (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*. International Society of the Learning Sciences, pp. 204–211. Volume 1, Full Papers. <https://repository.isls.org/handle/1/2205>.
- Department of Education and Training, 2018. Through Growth to Achievement: Report of the Review to Achieve Educational Excellence in Australian Schools. Department of Education and Training.
- Driver, R., 1989. Students' conceptions and the learning of science. *Int. J. Sci. Educ.* 11 (5), 481–490. <https://doi.org/10.1080/0950069890110501>.
- Duncan, R.G., Gotwals, A.W., 2015. A tale of two progressions: on the benefits of careful comparisons. *Sci. Educ.* 99 (3), 410–416. <https://doi.org/10.1002/sce.21167>.
- Duncan, R.G., Hmelo-Silver, C.E., 2009a. Learning progressions. *J. Res. Sci. Teach.* 46 (6).
- Duncan, R.G., Hmelo-Silver, C.E., 2009b. Learning progressions: aligning curriculum, instruction, and assessment. *J. Res. Sci. Teach.* 46 (6), 606–609. <https://doi.org/10.1002/tea.20316>.
- Duncan, R.G., Rivet, A.E., 2013. Science learning progressions. *Science* 339 (6118), 396–397. <https://doi.org/10.1126/science.1228692>.
- Duschl, R.A., Maeng, S., Sezen, A., 2011. Learning progressions and teaching sequences: a review and analysis. *Stud. Sci. Educ.* 47 (2), 123–182. <https://doi.org/10.1080/03057267.2011.604476>.
- Eames, C.L., Graf, E.A., van Rijn, P.W., Budzban, G., Voepel, T., 2021. The finite-to-finite strand of a learning progression for the concept of function: a research synthesis and cognitive analysis. *J. Math. Behav.* 62, 100864. <https://doi.org/10.1016/j.jmathb.2021.100864>.
- Empson, S.B., 2011. On the idea of learning trajectories: promises and pitfalls. *Math. Entus.* 8 (3), 571–598. <https://doi.org/10.54870/1551-3440.1229>.
- Fonger, N.L., Stephens, A., Blanton, M., Isler, I., Knuth, E., Murphy Gardner, A., 2018. Developing a learning progression for curriculum, instruction, and student learning: an example from mathematics education. *Cognit. Instruct.* 36 (1), 30–55. <https://doi.org/10.1080/07370008.2017.1392965>.
- Foster, J., Wiser, M., 2012. The potential of learning progression research to inform the design of state science standards. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: In Science: Current Challenges and Future Directions*. Sense Publishers, pp. 435–460. <https://doi.org/10.1007/978-94-6091-824-7>.
- Furtak, E.M., Heredia, S.C., 2014. Exploring the influence of learning progressions in two teacher communities. *J. Res. Sci. Teach.* 51 (8), 982–1020. <https://doi.org/10.1002/tea.21156>.
- Furtak, E.M., Thompson, J., Braaten, M., Windschitl, M., 2012. Learning progressions to support ambitious teaching practices. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, pp. 405–433. https://doi.org/10.1007/978-94-6091-824-7_17.
- Furtak, E.M., Morrison, D., Kroog, H., 2014. Investigating the link between learning progressions and classroom assessment. *Sci. Educ.* 98 (4), 640–673. <https://doi.org/10.1002/sce.21122>.
- Furtak, E.M., Kiemer, K., Swanson, R., de León, V., Circi, R.K., 2015. Learning Progressions, Formative Assessment, and Professional Development: Results of a Longitudinal Study. Paper Presented at the Annual International Conference of the National Association for Research in Science Teaching, Chicago, IL.
- Furtak, E.M., Bakeman, R., Buell, J.Y., 2018. Developing knowledge-in-action with a learning progression: sequential analysis of teachers' questions and responses to student ideas. *Teach. Teach. Educ.* 76, 267–282. <https://doi.org/10.1016/j.tate.2018.06.001>.
- Furtak, E.M., 2012. Linking a learning progression for natural selection to teachers' enactment of formative assessment. *J. Res. Sci. Teach.* 49 (9), 1181–1210. <https://doi.org/10.1002/tea.v49.9>.
- Furtak, E.M., 2017. Supporting Teachers' Formative Assessment Practice With Learning Progressions. Routledge. <https://doi.org/10.4324/9781315562636>.
- Gao, Y., Zhai, X., Andersson, B., Zeng, P.F., Xin, T., 2020. Developing a learning progression of buoyancy to model conceptual change: a latent class and rule space model analysis. *Res. Sci. Educ.* 50 (4), 1369–1388. <https://doi.org/10.1007/s11165-018-9736-5>.
- Gao, Y., Zhai, X., Cui, Y., Xin, T., Bulut, O., 2021. Re-validating a learning progression of buoyancy for middle school students: a longitudinal study. *Res. Sci. Educ.* <https://doi.org/10.1007/s11165-021-10021-x> [Advanced online publication].
- Gotwals, A.W., Songer, N.B., 2013. Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *J. Res. Sci. Educ.* 50 (5), 597–626. <https://doi.org/10.1002/tea.21083>.
- Gotwals, A.W., 2012. Learning progressions for multiple purposes: challenges in using learning progressions. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, pp. 461–472. https://doi.org/10.1007/978-94-6091-824-7_19.
- Graf, E.A., van Rijn, P.W., Eames, C.L., 2021. A cycle for validating a learning progression illustrated with an example from the concept of function. *J. Math. Behav.* 62 (100836). <https://doi.org/10.1016/j.jmathb.2020.100836>.
- Gunckel, K.L., Covitt, B.A., Salinas, I., 2018. Learning progressions as tools for supporting teacher content knowledge and pedagogical content knowledge about water in environmental systems. *J. Res. Sci. Teach.* 55 (9), 1339–1361. <https://doi.org/10.1002/tea.21454>.
- Gunckel, K.L., Covitt, B.A., Salinas, I., Anderson, C.W., 2012. A learning progression for water in socio-ecological systems. *J. Res. Sci. Teach.* 49 (7), 843–868. <https://doi.org/10.1002/tea.21024>.
- Hammer, D., Sikorski, T.-R., 2015. Implications of complexity for research on learning progressions. *Sci. Educ.* 99 (3), 424–431. <https://doi.org/10.1002/sce.21165>.
- Heritage, M., 2008. Learning Progressions: Supporting Instruction and Formative Assessment. Council of Chief State School Officers.
- Heritage, M., 2011. Commentary on road maps for learning: a guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research and Perspectives* 9 (2–3), 149–151. <https://doi.org/10.1080/15366367.2011.599647>.

- Hermann-Abell, C.F., DeBoer, G.E., 2018. Investigating a learning progression for energy ideas from upper elementary school through high school. *J. Res. Sci. Teach.* 55 (1), 68–93. <https://doi.org/10.1002/tea.21411>.
- Hughes, S., Makara, K., Stacey, D., 2020. Learning progression in the humanities: identifying tensions in articulating progression in humanities in Wales. *Curric. J.* 31 (2), 276–289. <https://doi.org/10.1002/curj.28>.
- Huynh, N.T., Solem, M., Bednarz, S.W., 2015. A road map for learning progressions research in geography. *J. Geogr.* 114 (2), 69–79. <https://doi.org/10.1080/00221341.2014.935799>.
- Ivars, P., Fernández, C., Linares, S., 2020. A learning trajectory as a scaffold for pre-service teachers' noticing of students' mathematical understanding. *Int. J. Sci. Math. Educ.* 18 (3), 529–548. <https://doi.org/10.1007/s10763-019-09973-4>.
- Jin, H., Anderson, C.W., 2012. A learning progression for energy in socio-ecological systems. *J. Res. Sci. Teach.* 49 (9), 1149–1180. <https://doi.org/10.1002/tea.21051>.
- Jin, H., Choi, J., Anderson, C.W., 2009. Development and validation of assessments for a learning progression on carbon cycling in socio-ecological systems. In: Gotwals, A.W., Alonzo, A.C. (Eds.), *Conference Proceedings of the Learning Progressions in Science Conference*. In: <http://www.education.msu.edu/projects/leaps/proceedings/Jin.pdf>.
- Jin, H., Delgado, C., Bauer, M.I., Wylie, E.C., Cisterna, D., Lort, K.F., 2019a. A hypothetical learning progression for quantifying phenomena in science. *Sci. Educ.* 28 (9–10), 1181–1208. <https://doi.org/10.1007/s11191-019-00076-8>.
- Jin, H., Mikeska, J.N., Hokayem, H., Mavronikolas, E., 2019b. Toward coherence in curriculum, instruction, and assessment: a review of learning progression literature. *Sci. Educ.* 103 (5), 1206–1234. <https://doi.org/10.1002/sce.21525>.
- Johnson, P., 2013. A learning progression towards understanding chemical change. *Educ. Quím.* 24 (4), 365–372. [https://doi.org/10.1016/S0187-893X\(13\)72489-3](https://doi.org/10.1016/S0187-893X(13)72489-3).
- Jung, J., Chang, J., Park, P., 2020. An analysis of different grade levels of elementary school students' reasoning about the changes of state of water within a learning progression. *Asia Pac. Sci. Educ.* 6 (2), 548–563. <https://doi.org/10.1163/23641177-bja10004>.
- Kaldaras, L., Akaze, H., Krajcik, J., 2021. Developing and validating Next Generation Science Standards-aligned learning progression to track three-dimensional learning of electrical interactions in high school physical science. *J. Res. Sci. Teach.* 58 (4), 589–618. <https://doi.org/10.1002/tea.21672>.
- Kamei, T., Pavlovic, M., 2021. Investigating differential item functioning to validate a thinking skills learning progression for students with intellectual disability and autism spectrum disorder. *Int. J. Educ. Res.* 106 (101726). <https://doi.org/10.1016/j.ijer.2020.101726>.
- Kelly, D.L., Centurino, V.A.S., Martin, M.O., Mullis, I.V.S. (Eds.), 2020. *TIMSS 2019 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*. TIMSS & PIRLS International Study Center.
- Ketterlin-Geller, L.R., Shivraj, P., Basaraba, D., Yovanoff, P., 2019. Considerations for using mathematical learning progressions to design diagnostic assessments. *Measurement: Interdisciplinary Research and Perspectives* 17 (1), 1–22. <https://doi.org/10.1080/15366367.2018.1479087>.
- Ketterlin-Geller, L.R., Zannou, Y., Sparks, A., Perry, L., 2020. Empirical recovery of learning progressions through the lens of educators. *J. Math. Behav.* 60 (100805), 1–14. <https://doi.org/10.1016/j.jmathb.2020.100805>.
- Kim, E.M., Haberstroh, J., Peters, S., Howell, H., Oláh, L.N., 2017. A Learning Progression for Geometrical Measurement in One, Two, and Three Dimensions (ETS RR-17-55). ETS Research Report Series. <https://doi.org/10.1002/ets2.12189>.
- Ladachart, L., Poothawee, M., Ladachart, L., 2020. Toward a place-based learning progression for haze pollution in the northern region of Thailand. *Cult. Stud. Sci. Educ.* 15 (4), 991–1017. <https://doi.org/10.1007/s11422-020-09981-9>.
- Lai, E.R., Kobrin, J.L., DiCerbo, K.E., Holland, L.R., 2017. Tracing the assessment triangle with learning progression-aligned assessments in mathematics. *Measurement: Interdisciplinary Research and Perspectives* 15 (3–4), 143–162. <https://doi.org/10.1080/15366367.2017.1388113>.
- Larsen, T.B., Harrington Jr., J.A., 2018. Developing a learning progression for place. *J. Geogr.* 117 (3), 100–118. <https://doi.org/10.1080/00221341.2017.1337212>.
- Lee, P.J., Ashby, R., 2000. Progression in historical understanding among students ages 7–14. In: Stearns, P.N., Seixas, P., Wineburg, S. (Eds.), *Knowing, Teaching, and Learning History*. New York University Press, pp. 192–222.
- Lee, H.-S., Liu, O.L., 2010. Assessing learning progression of energy concepts across middle grades: the knowledge integration perspective. *Sci. Educ.* 94 (4), 665–688. <https://doi.org/10.1002/sce.20382>.
- Lehrer, R., Schauble, L., 2012. Seeding evolutionary thinking by engaging children in modeling its foundations. *Sci. Educ.* 96 (4), 701–724. <https://doi.org/10.1002/sce.20475>.
- Lehrer, R., Schauble, L., 2015. Learning progressions: the whole world is NOT a stage. *Sci. Educ.* 99 (3), 432–437. <https://doi.org/10.1002/sce.21168>.
- Lehrer, R., Kim, M., Ayers, E., Wilson, M., 2014. Toward establishing a learning progression to support the development of statistical reasoning. In: Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing, pp. 31–60.
- Lesh, R., Yoon, C., 2004. Evolving communities of mind—in which development involves several interacting and simultaneously developing strands. *Math. Think. Learn.* 6 (2), 205–226. https://doi.org/10.1207/s15327833mtl0602_7.
- Liu, X., Tang, L., 2004. The progression of students' conceptions of energy: a cross-grade, cross-cultural study. *Can. J. Math. Sci. Technol. Educ.* 4 (1), 43–57. <https://doi.org/10.1080/14926150409556596>.
- Liu, X., 2013. Difficulties of items related to energy and matter: implications for learning progression in high school chemistry. *Educ. Quím.* 24 (4), 416–422. [https://doi.org/10.1016/S0187-893X\(13\)72495-9](https://doi.org/10.1016/S0187-893X(13)72495-9).
- Lobato, J., Walters, C.D., 2017. A taxonomy of approaches to learning trajectories and progressions. In: Cai, J. (Ed.), *Compendium for Research in Mathematics Education*. National Council of Teachers of Mathematics, U.S., pp. 74–101.
- Lombard, F., Merminod, M., Widmer, V., Schneider, D.K., 2018. A method to reveal fine-grained and diverse conceptual progressions during learning. *J. Biol. Educ.* 52 (1), 101–112. <https://doi.org/10.1080/00219266.2017.1405534>.
- Maeng, S., 2020. Investigating a learning progression for reasoning practices of geocognition using GeoMapApp-based assessment. *Asia Pac. Sci. Educ.* 6 (2), 584–614. <https://doi.org/10.1163/23641177-bja10009>.
- Magana, A.J., 2017. Modeling and simulation in engineering education: a learning progression. *J. Prof. Issues Eng. Educ. Pract.* 143 (4), 04017008. [https://doi.org/10.1061/\(ASCE\)E1943-5541.0000338](https://doi.org/10.1061/(ASCE)E1943-5541.0000338).
- Maloney, A.P., Confrey, J., Nguyen, K.H. (Eds.), 2014. *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing.
- Maras, K., 2021. This time without “feeling”: children's intuitive theories of art as a logical basis for learning progression in visual arts. In: *Research Conference 2021: Excellent Progress for Every Student: Proceedings and Program*. Australian Council for Educational Research, pp. 42–46. https://doi.org/10.37517/978-1-74286-638-3_3.
- Mayes, R.L., Forrester, J.H., Christus, J.S., Peterson, F.I., Bonilla, R., Yestness, N., 2014. Quantitative reasoning in environmental science: a learning progression. *Int. J. Sci. Educ.* 36 (4), 635–658. <https://doi.org/10.1080/09500693.2013.819534>.
- Metz, K., 2009. Rethinking what is “developmentally appropriate” from a learning progression perspective: the power and the challenge. *Rev. Sci. Math. ICT Educ.* 3 (1), 5–22. <https://doi.org/10.26220/rev.119>.
- Mitchell, J.T., 2018. Place and learning progressions. *J. Geogr.* 117 (3).
- Mohan, L., Chen, J., Anderson, C.W., 2009. Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *J. Res. Sci. Teach.* 46 (6), 675–698. <https://doi.org/10.1002/tea.20314>.
- Moreno, M., Sánchez-Matamoros, G., Callejo, M.L., Pérez-Tyteca, P., Linares, S., 2021. How prospective kindergarten teachers develop their noticing skills: the instrumentation of a learning trajectory. *ZDM Math. Educ.* 53 (1), 57–72. <https://doi.org/10.1007/s11858-021-01234-5>.
- Myers, M., Sztajn, P., Wilson, P.H., Edgington, C., 2015. From implicit to explicit: articulating equitable learning trajectories based instruction. *J. Urban Math. Educ.* 8 (2), 11–22.
- Myers, M., 2015. The unintended consequences of a learning trajectories approach. In: Bartell, T.G., Bieda, K.N., Putnam, R.T., Bradfield, K., Dominguez, H. (Eds.), *Proceedings of the 37th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Michigan State University, pp. 756–763.

- National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010. Common Core State Standards for Mathematics. National Governors Association Center for Best Practices & Council of Chief State School Officers. http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf.
- National Research Council, 2006. Systems for State Science Assessment. The National Academies Press. <https://doi.org/10.17226/11312>.
- National Research Council, 2007. Taking Science to School: Learning and Teaching Science in Grades K–8. The National Academies Press. <https://doi.org/10.17226/11625>.
- Neumann, K., Viering, T., Boone, W.J., Fischer, H.E., 2013. Towards a learning progression of energy. *J. Res. Sci. Teach.* 50 (2), 162–188. <https://doi.org/10.1002/tea.21061>.
- New Zealand Ministry of Education, 2010. The Literacy Learning Progressions: Meeting the Reading and Writing Demands of the Curriculum. Learning Media Limited. <https://literacyprogressions.tki.org.nz/>.
- NGSS Lead States, 2013. Next Generation Science Standards: For States, By States. The National Academies Press, Washington, DC. <https://doi.org/10.17226/18290>.
- Oates, G., Seah, R., 2021. Learning progressions in mathematics and science. *Aust. J. Educ.* 65 (3).
- Osborne, J.F., Henderson, B., MacPherson, A., Szu, E., Wild, A., Yao, S.-Y., 2016. The development and validation of a learning progression for argumentation in science. *J. Res. Sci. Teach.* 53 (6), 821–846. <https://doi.org/10.1002/tea.21316>.
- Paik, S.-H., Song, G., Kim, S., Ha, M., 2017. Developing a four-level learning progression and assessment for the concept of buoyancy. *Eurasia J. Math. Sci. Technol. Educ.* 13 (8), 4965–4986. <https://doi.org/10.12973/eurasia.2017.00976a>.
- Parr, J., 2011. Repertoires to scaffold teacher learning and practice in assessment of writing. *Assess. Writ.* 16 (1), 32–48. <https://doi.org/10.1016/j.asw.2010.11.002>.
- Pierson, A.E., Clark, D.E., 2019. Sedimentation of modeling practices: dimensions of co-operative action at a classroom scale. *Sci. Educ.* 28 (8), 897–925. <https://doi.org/10.1007/s11191-019-00050-4>.
- Pierson, A., Clark, D.B., Sherard, M.K., 2017. Learning progressions in context: tensions and insights from a semester-long middle school modeling curriculum. *Sci. Educ.* 101 (6), 1061–1088. <https://doi.org/10.1002/sce.21314>.
- Pierson, A., Clark, D., Kelly, G., 2019. Scientific practices, epistemic aims, and learning progressions: tensions in prioritizing content, epistemic practices, and social dimensions of learning. *Sci. Educ.* 28 (8), 833–841. <https://doi.org/10.1007/s11191-019-00070-0>.
- Plummer, J.D., Maynard, L., 2014. Building a learning progression for celestial motion: an exploration of students' reasoning about the seasons. *J. Res. Sci. Teach.* 51 (7), 902–929. <https://doi.org/10.1002/tea.21151>.
- Pongprapan, P., Jituaflua, A., 2021. Developing and assessing learning progression for botanical literacy using Rasch analysis. *Sci. Educ. Int.* 32 (2), 125–130. <https://doi.org/10.33828/sei.v32.i2.5>.
- Popham, J.W., 2007. The lowdown on learning progressions. *Educ. Leader.* 64 (7), 83–84.
- Psillos, D., Kariotoglou, P., 2016. Theoretical issues related to designing and developing teaching-learning sequences. In: Psillos, D., Kariotoglou, P. (Eds.), *Iterative Design of Teaching-Learning Sequences*. Springer, pp. 11–34. https://doi.org/10.1007/978-94-007-7808-5_2.
- Rich, K.M., Binkowski, T.A., Strickland, C., Franklin, D., 2018. Decomposition: a K-8 computational thinking learning trajectory. In: *Proceedings of the 2018 ACM Conference on International Computing*. Association for Computing Machinery.
- Russell, T., McGuigan, L., 2019. Developmental progression in learning about evolution in the 5–14 age range in England. In: Harms, U., Reiss, M.J. (Eds.), *Evolution Education Reconsidered*. Springer, pp. 59–80.
- Salinas, I., 2009. Learning progression in science education: two approaches for development. In: Gotwals, A.W., Alonzo, A.C. (Eds.), *Conference Proceedings of the Learning Progressions in Science Conference*. In: <http://www.education.msu.edu/projects/leaps/proceedings/Salinas.pdf>.
- Schneider, R.M., Plasman, K., 2011. Science teacher learning progressions: a review of science teachers' pedagogical content knowledge development. *Rev. Educ. Res.* 81 (4), 530–565. <https://doi.org/10.3102/0034654311423382>.
- Schramm, J.W., Jin, H., Keeling, E.G., Johnson, M., Shin, H.J., 2018. Improved student reasoning about carbon-transforming processes through inquiry-based learning activities derived from an empirically validated learning progression. *Res. Sci. Educ.* 48 (5), 887–911. <https://doi.org/10.1007/s11165-016-9584-0>.
- Schwarz, C.V., Reiser, B.J., Davis, E.A., Kenyon, L., Acher, A., Fortus, D., Shwartz, Y., Hug, B., Krajcik, J., 2009. Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *J. Res. Sci. Teach.* 46 (6), 632–654. <https://doi.org/10.1002/tea.20311>.
- Scular, C., 2021. Identifying and monitoring progress in collaboration skills. In: *Research Conference 2021: Excellent Progress for Every Student: Proceedings and Program*. Australian Council for Educational Research, pp. 112–120. https://doi.org/10.37517/978-1-74286-638-3_15.
- Sevian, H., Talanquer, V., 2014. Rethinking chemistry: a learning progression on chemical thinking. *Chem. Educ. Res. Pract.* 15 (1), 10–23. <https://doi.org/10.1039/C3RP00111C>.
- Shavelson, R.J., Kurius, A., 2012. Reflections on learning progressions. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, pp. 13–26. https://doi.org/10.1007/978-94-6091-824-7_2.
- Shepard, L.A., 2008. Formative assessment: caveat emptor. In: Dwyer, C.A. (Ed.), *The Future of Assessment: Shaping Teaching and Learning*. Lawrence Erlbaum Associates, pp. 279–303.
- Shepard, L.A., 2009. Commentary: evaluating the validity of formative and interim assessment. *Educ. Meas.* 28 (3), 32–37. <https://doi.org/10.1111/j.1745-3992.2009.00152.x>.
- Shepard, L.A., 2018. Learning progressions as tools for assessment and learning. *Appl. Meas. Educ.* 31 (2), 165–174. <https://doi.org/10.1080/08957347.2017.1408628>.
- Siemon, D., Barkatsas, T., Seah, R. (Eds.), 2019. *Researching and Using Progressions (Trajectories) in Mathematics Education*. Brill.
- Sikorski, T.-R., Hammer, D., 2017. Looking for coherence in science curriculum. *Sci. Educ.* 101 (6), 929–943. <https://doi.org/10.1002/sce.21299>.
- Sikorski, T.-R., 2019. Context-dependent “upper anchors” for learning progressions. *Sci. Educ.* 28 (8), 957–981. <https://doi.org/10.1007/s11191-019-00074-w>.
- Silva, C.M., Moses, R.P., Rivers, J., Johnson, P., 1990. The Algebra Project: making middle school mathematics count. *J. Negro Educ.* 59 (3), 375–391. <https://doi.org/10.2307/2295571>.
- Simeon, D., 2021. Learning progressions/trajectories in mathematics: supporting reform at scale. *Aust. J. Educ.* 65 (3), 227–247. <https://doi.org/10.1177/00049441211045745>.
- Simon, M.A., 1995. Reconstructing mathematics pedagogy from a constructivist perspective. *J. Res. Math. Educ.* 26 (2), 114–145. <https://doi.org/10.2307/749205>.
- Smith, C.L., Wiser, M., Anderson, C.W., Krajcik, J., 2006. Implications of research on children's learning for standards and assessment: a proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspectives* 4 (1–2), 1–98. <https://doi.org/10.1080/15366367.2006.9678570>.
- Solem, M., Huynh, N.T., Boeh, R. (Eds.), 2014. *Learning Progressions for Maps, Geospatial Technology and Spatial Thinking: A Research Handbook*. Association of American Geographers.
- Songer, N.B., Kelcey, B., Gotwals, A.W., 2009. How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *J. Res. Sci. Teach.* 46 (6), 610–631. <https://doi.org/10.1002/tea.20313>.
- Stephens, M., Day, L., Home, M., 2021. An empirically based practical learning progression for generalisation, an essential element of algebraic reasoning. *Aust. J. Educ.* 65 (3), 313–328. <https://doi.org/10.1177/00049441211044798>.
- Strickland, J., 2018. *The Design and Use of a Learning Progression of Emergent Numeracy to Support the Instruction of Students With Intellectual and Developmental Disabilities*. [Doctoral dissertation]. The University of Melbourne. <http://hdl.handle.net/11343/219848>.
- Suh, J., Seshaiyer, P., 2015. Examining teachers' understanding of the mathematical learning progression through vertical articulation during Lesson Study. *J. Math. Teach. Educ.* 18 (3), 207–229. <https://doi.org/10.1007/s10857-014-9282-7>.
- Suh, J.M., Birkhead, S., Frank, T., Baker, C., Galanti, T., Seshaiyer, P., 2020. Developing an asset-based view of students' mathematical competencies through learning trajectory-based lesson study. *Math. Teach. Educ.* 9 (3), 229–245. <https://doi.org/10.5951/MTE.2020.0033>.
- Sztajn, P., Confrey, J., Wilson, P.H., Edgington, C., 2012. Learning trajectory based instruction: toward a theory of teaching. *Educ. Res.* 41 (5), 147–156. <https://doi.org/10.3102/0013189X12442801>.
- Talanquer, V., 2013. Progresiones de aprendizaje: promesa y potencial. *Educ. Quím.* 24 (4).

- Taylor, L., Brand, B., 2021. Enhancing place-based learning progression through epistemic agency: a response to toward a hypothetical place-based learning progression for haze pollution in the northern region of Thailand. *Cult. Stud. Sci. Educ.* 16 (2), 429–446. <https://doi.org/10.1007/s11422-021-10022-2>.
- The World Bank, 2021. The World by Income and Region. World Bank. <https://datatopics.worldbank.org>.
- van Rijn, P.W., Graf, E.A., Deane, P., 2014. Empirical recovery of argumentation learning progressions in scenario-based assessments of English language arts. *Psicol. Educ.* 20 (2), 109–115.
- von Aufschnaiter, C., Alonzo, A.C., 2018. Foundations of formative assessment: introducing a learning progression to guide pre-service physics teachers' video-based interpretation of student thinking. *Appl. Meas. Educ.* 31 (2), 113–127. <https://doi.org/10.1080/08957347.2017.1408629>.
- Wang, C.C., Ho, H.C., Cheng, Y.Y., 2015. Building a learning progression for scientific imagination: a measurement approach. *Think. Skills Creativ.* 17, 1–14. <https://doi.org/10.1016/j.tsc.2015.02.001>.
- Wang, C.-C., Niemi, H., Cheng, C.-L., Cheng, Y.-Y., 2017. Validation of learning progression in scientific imagination using data from Taiwanese and Finnish elementary school students. *Think. Skills Creativ.* 24, 73–85. <https://doi.org/10.1016/j.tsc.2017.02.014>.
- Waters, C., 2018. Learning Progressions in ACER's Work. Centre for Assessment, Reform and Innovation, Australian Council for Education Research (ACER), Melbourne, Australia. <https://research.acer.edu.au>.
- Wickman, P.-O., 2014. Teaching learning progressions. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research in Science Education*, 2. Routledge, pp. 145–163. <https://doi.org/10.4324/9780203097267>.
- Wilkins, J.L.M., Norton, A., 2018. Learning progression toward a measurement concept of fractions. *Int. J. STEM Educ.* 5 (27), 1–11. <https://doi.org/10.1186/s40594-018-0119-2>.
- Williams, R., 2018. The Impact of the Use of Science Notebooks in Conjunction With a Learning Progression-Based Science Unit in an Urban Middle School [Doctoral Dissertation]. University of Missouri, St. Louis. <https://irl.umsl.edu/dissertation/783>.
- Wilson, M., Lehrer, R., 2021. Improving learning: using a learning progression to coordinate instruction and assessment. *Front. Educ.* 6 (654212). <https://doi.org/10.3389/educ.2021.654212>.
- Wilson, P.H., Mojica, G.F., Confrey, J., 2013. Learning trajectories in teacher education: supporting teachers' understandings of students' mathematical thinking. *J. Math. Behav.* 32 (2), 103–121. <https://doi.org/10.1016/j.jmathb.2012.12.003>.
- Wilson, P.H., Sztajn, P., Edgington, C., Myers, M., 2015. Teachers' uses of a learning trajectory in student-centered instructional practices. *J. Teach. Educ.* 66 (3), 227–244. <https://doi.org/10.1177/0022487115574104>.
- Wilson, P.H., Sztajn, P., Edgington, C., Webb, J., Myers, M., 2017. Changes in teachers' discourse about students in a professional development on learning trajectories. *Am. Educ. Res. J.* 54 (3), 568–604. <https://doi.org/10.3102/0002831217693801>.
- Wilson, M., 2009. Measuring progressions: assessment structures underlying a learning progression. *J. Res. Sci. Teach.* 46 (6), 716–730. <https://doi.org/10.1002/tea.20318>.
- Wiser, M., Smith, C.L., Doubler, S., 2012. Learning progressions as tools for curriculum development. In: Alonzo, A.C., Gotwals, A.W. (Eds.), *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, pp. 359–403. https://doi.org/10.1007/978-94-6091-824-7_16.
- Wyner, Y., Doherty, J.H., 2017. Developing a learning progression for three-dimensional learning of the patterns of evolution. *J. Res. Sci. Teach.* 101 (5), 787–817. <https://doi.org/10.1002/sce.21289>.
- Yao, J.-X., Guo, Y.-Y., 2018. Validity evidence for a learning progression of scientific explanation. *J. Res. Sci. Teach.* 55 (2), 299–317. <https://doi.org/10.1002/tea.21420>.
- Yao, J.-X., Guo, Y.-Y., Neumann, K., 2017. Refining a learning progression of energy. *Int. J. Sci. Educ.* 39 (17), 2361–2381. <https://doi.org/10.1080/09500693.2017.1381356>.
- Yin, Y., Tomita, M.K., Shavelson, R.J., 2013. Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *Int. J. Sci. Educ.* 36 (4), 531–552. <https://doi.org/10.1080/09500693.2013.787556>.
- Yoon, S.A., Goh, S.-E., Yang, Z., 2019. Toward a learning progression of complex systems understanding. *Complexity* 16 (1), 1–19. <https://doi.org/10.29173/cmplt29340>.
- Zahner, W., Wynn, L., 2021. Rethinking learning trajectories in light of student linguistic diversity. *Math. Think. Learn.* 16 (1), 1–19. <https://doi.org/10.1080/10986065.2021.1931650>.
- Zarmati, L., 2021. How might we identify and measure learning progression in history?. In: *Research Conference 2021: Excellent Progress for Every Student: Proceedings and Program*. Australian Council for Educational Research, pp. 126–131. https://doi.org/10.37517/978-1-74286-638-3_19.

Optimizing learning: an overview

Stuart McNaughton, The University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	560
Description: limited optimization is typical	561
Explanation: the limits to optimizing	561
Fidelity to designs	561
Adaptation on the ground	562
Resourcing for sustainable change	562
More to know in instructional science	562
What is needed for successful optimization	563
Clarity about what counts as success	563
A science of optimization	564
Mitigating risks to optimizing	564
Differentiated optimization	564
Focusing on higher order and interdependent skills	565
Recognizing the significance of identity, language and culture	566
Conclusion: is optimization a chimera?	566
References	566

Introduction

Developmental and educational psychology are optimistic sciences. That optimism was captured in an early vision for the aim of developmental science which was to describe, explain and optimize intraindividual change and of interindividual differences in that change (Baltes et al., 1977). The assumption was that if we could get the descriptions right, and had robust explanations, then we would be in the position to optimize. The focus for Baltes et al. (1977) was development, but the same logic has been present for learning. However, it turns out that the sequence from description to optimization is not as linear as suggested, and the third part, which is the focus of this chapter on optimizing learning, is nowhere near as inevitable or easily achievable as implied.

The reason for having a more cautious optimism comes from knowing much more about the phenomena of learning than we did in the 1970s. We know much more about the proximal and distal conditions that affect learning, we know much more about the embedded systems in which a learner learns, and we know much more about the complexity of trying to change conditions and systems to optimize.

The complexity is captured in another programmatic statement which rejects the idea that we can discover “what works” in some universal and unqualified sense. In recognition of the variability, contextual adaptations and inconsistency in our attempts to optimize, especially when the focus is wide impact, the proposal is that we need to know what works for whom, under what conditions, and at scale (Bryk et al., 2015).

Baltes et al. (1977) were concerned with optimizing development across socialization settings. In this chapter, the focus is optimizing learning specifically in educational contexts. The concept of optimization is elaborated through the following discussion, but a preliminary definition is suggested here. Optimization means enabling what is valued educationally to be learned effectively, considered against five very closely related criteria. These are the robustness of learning; its transferability; the fluency and accuracy of that learning; its generalizability to abstracted principles and concepts; and the enhancement of further development through what is learned (McNaughton, 2018). The last criterion requires a horizon which marks having successfully optimized. Knowing what counts as success is especially significant for learners whose progress is lower than expected, or who have traditionally not been well served by schooling.

The present discussion of optimization adopts a systems framework for thinking about optimization. Much is known about specific variables which can impact on the individual learning, and especially when conditions are controlled in experimental analyses. For example, the properties of effective feedback, or the role of deliberate practice and opportunities to learn, have been described and explained in detail (Ericsson, 2006; Hattie and Timperley, 2007). But the more pressing questions, the ones we have yet to have definitive answers for, are more systemic and are more in keeping with Bryk et al.’s (2015) proposal.

There are two questions in particular which this chapter addresses. They are how to optimize at scale, and how to mitigate the risks. The former focuses on challenges in instructional designs to optimize learning at scale, and the latter on the seemingly intractable problems of optimization for all, conceived here as risks. To aid clarity, the discussion follows the Baltes et al. (1977) logic: describing the issues to do with optimizing, explaining them, and then considering what it would take to optimize learning.

Description: limited optimization is typical

Our attempts at optimization are hardly optimal: the current picture is one of variability in outcomes when optimizing at scale and limited sustainability over time. Optimal conditions and outcomes prove hard to replicate beyond initial experimental demonstrations (Kim, 2019).

The reasons for this description of the current state are illustrated in recent evaluation studies of Reading Recovery, an effective early intervention program accelerating literacy learning for those not making appropriate progress. Despite finding an overall strong effect size on a range of literacy outcome measures (a median effect size of 0.7), the school to school variation was substantial, varying from negative effect sizes, meaning intervention children in that site actually got worse than their control group peers, to effect sizes greater than +2.0, a positive impact almost unheard of in social science research interventions (May et al., 2013, 2016).

The variation was attributable to aspects of teaching termed “instructional strength,” reflecting the individual Reading Recovery teacher’s expertise with the instructional procedures. But it was also attributable at the school level to the degree to which the Reading Recovery teacher was well integrated into the school community and received endorsement by leaders. Further, at a district level, the constraints and enablers were about administrative support, in the forms of political recognition and funding, and this support needed to be ongoing. Finally, a constraint on implementation at a district level was turnover in the administrators. A follow up study replicated the primary finding of overall large effects on immediate outcomes, but also found very limited sustainability of achievement gains (May et al., 2022). This adds to the picture of variability, given that studies in other jurisdictions have detected long term gains (Appleton-Dyer et al. (2019)).

This pattern of variability in programs when implemented at scale is not limited to Reading Recovery or literacy; it is universal. It is found across many domains (Hill and Erickson, 2019; Kim, 2019) and across ages, from pre-K programs focused on self-regulation (Nesbitt and Farran, 2021) through to learning social and emotional skills and anti-bullying programs in secondary schools (Yeager et al. (2015)). More than this, variability is inherent in “business as usual” too, the everyday attempts to optimize learning; at every level of education; from granular to macro units of analysis (McNaughton, 2011); within classrooms and between classrooms; within schools and between schools; within and between clusters of school and districts; and within and between countries.

Sustaining the changes needed to optimize learning also has proven to be difficult as May et al. (2022) found, and not only at the individual student level. Schools that successfully implement reforms and improve achievement find it difficult to sustain them in the face of competing priorities, changing demands, changing student demographics and teacher and administrator turnover (Datonow, 2005).

The issue of sustainability of the conditions for optimizing learning when the unit of analysis is the school or a wider grouping of schools is a temporal one. The changes are hard to maintain over time for those involved. However, there is a second sense of sustainability when the unit of analysis is the individual’s learning. At this level it is a question of ongoing learning, the developmental significance of what is learned and what needs to be learned over time. In this sense what is optimized at one point is a condition for ongoing learning. The risk is that what is learned may not be optimal or as developmentally progressive as it could be for priming further learning and longer-term outcomes. This issue is readily apparent in core subjects such as literacy and numeracy learning.

Dual pathway models of early literacy learning identify two developing strands (Paris, 2005): one set of “constrained” knowledge and skills (e.g., phonological knowledge and alphabet knowledge), and a second set of “unconstrained” knowledge and skills (e.g., lexical knowledge and textual knowledge). They have different trajectories of development. The constrained set is learned faster and generally knowledge acquisition is completed early. The development of unconstrained skills is more gradual and open ended. Development of text reading for meaning is dependent on both sets being learned. But ongoing development of the unconstrained set is dependent on the how effectively the constrained set is learned, judged against the criteria noted earlier. In this sense, the constrained set are causal, they are building blocks which are necessary, but not sufficient, for development to proceed. The point for optimization is that optimizing which only focuses on the constrained skills, meets the necessary conditions for four of the criteria (robustness, transferability, fluency and accuracy, and in a principled way) but may not be sufficient for ongoing development. Indeed, a primary focus in the early years on the constrained skills risks opportunity costs for the learner and slowing development, by reducing opportunities to learn the unconstrained set (McNaughton, 2018). Meeting the final criterion of sustainability when considered as learning being developmentally progressive, is dependent on other learning being in place concurrently and consequentially.

Explanation: the limits to optimizing

Following the Baltes et al. (1977) sequence, this description of optimization as often variable at scale and over time requires explaining there are four major explanations for why we are not good at optimizing learning.

Fidelity to designs

The usual explanation for programmatic interventions being variable, which is also applicable to everyday teaching, focuses on the degree of specification or prescription in, and adherence to, the program design, curricula, syllabi, or lesson plans. This is treatment

fidelity or integrity. Fidelity to the “treatment” design is an established problem in the implementation of educational programs, and one of the major sources of influences on treatment effects (Hill and Erickson, 2019; Schochet et al., 2014; Coalition4evidence, 2013). Programs are hard to put in place reliably if the criterion is being consistent with the original program design. Durlak and DuPre (2008) reviewed 500 health and education programs for children and adolescents in studies from 1976 to 2006. There were generally low levels of compliance to the program design. They identified a long list of 23 contextual factors that influenced the fidelity of implementation.

The local context creates special circumstances for putting a program in place (Schochet et al., 2014). The 23 contextual features identified by Durlak and DuPre (2008) are illustrative of the sources identified for the Reading Recovery example. They include leadership, teacher skills and knowledge, resources, and the presence of other programs. When taking programs to scale, or when they are embedded as business as usual, the sources also include differences between the original demonstrations under carefully controlled and well-resourced conditions, and the conditions faced in wider implementation and in everyday settings when the original conditions such as researcher input, funding and participant buy-in are not present.

Adaptation on the ground

It matters whether the evidence-based instructional processes are followed with integrity. But there is a second explanation for the variability in optimizing. Teachers in the local context adapt. Even in the most prescriptive programs which attempt to control variation by close specification, it is present (Bryk et al., 2015; Hill and Erickson, 2019). The variability on the ground is not just through a failure to follow the design.

To adapt and innovate is in the nature of complex human activity. Teachers use their tools and deploy instruction in ways that reflect aspects of their expert knowledge, skills, values and goals. Some variability comes about because educators in local systems will, to a greater or lesser extent, modify and adapt the instructional designs and practices specified by their curricula and programs. Being an expert within a community of experts, with one’s own motivations, values and levels of expertise, mean practices are dynamic and sensitive; they change through responses to new conditions and knowledge.

Hill and Erickson, 2019 point out that studies of adaptation of interventions are rare. Some of the adaptations will be detrimental and undermine effectiveness and compromise optimization. Some will be progressive and contribute to optimization. Some of that adaptation may reflect the need to contextualize the generally effective instructional designs to the local circumstances. Some variability will reflect effective innovation. Each of these underline the need to know what works for whom under known conditions at scale, leading to much better nuanced specifications for the instructional design. Contemporary descriptions of effective teaching use the model of adaptive expertise to explain how optimization in the moment-by-moment implementation of curricula and pedagogy depends on the adaptations that skillful professionals make (Darling-Hammond and Bransford, 2005).

The conundrum that we face in educational systems when optimizing at scale is where to set the balance (McNaughton, 2021). It is between enough prescription to guarantee consistency in application of what is known to be effective across contexts, but “loose” enough in the specifications to enable the moment-by-moment adaptations and the progressive forms of innovation. Solving that balance problem is crucially dependant on enabling conditions such as those noted in the Reading Recovery example: the training of the teacher, and supporting conditions within and across sites.

Resourcing for sustainable change

The role of enabling conditions for optimizing leads to the third explanation—the resourcing of teaching, such as physical or social resources, strong leadership and professional development for teaching and the educational resources available to the schools and the communities that schools serve.

An example of realizing how optimization depends on enabling conditions (and not just the curriculum and pedagogical specifications) comes from the KEEP research with schools in the Kamehameha district in Hawai’i in the 1970s. The intervention focused on building higher achievement levels in reading than the typical low levels for indigenous Hawaiian students (Tharp et al., 2000). It was very successful and pronounced gains in reading achievement occurred, both in the original schools and in scaling up to multicultural public schools over an 18 year period.

But in 1997 it was formally terminated, and the researchers identified two issues. One was too rapid expansion, with consequent diminution of resources and a shift away from the original instructional design. The second was the intervention focus, which was based on culturally responsive reading instruction, but designed and implemented only in the early grades. It was not a whole school approach, addressing the issues of sustainability and developmental significance. Like other interventions based on culturally responsive pedagogy, it matters what the school leadership provides and it matters how the pedagogy is applied across year levels and subject domains. More than this, it matters that the resources to support these continue to be available.

More to know in instructional science

The fourth explanation is a truism: we always will need to know more. One example of this can be found in the idea of “instructional risk” (McNaughton, 2018). This proposes that all instruction carries risks to being effective. Our knowledge of how to optimize learning currently is limited because of not having the best evidence for how to mitigate the risks. We are still unraveling how the risks emerge and how to mitigate them.

Instructional risks come about in a number of ways. One is through limited specification of the components needed for learning. An example is how inquiry and discovery approaches in science and mathematics are problematic without teacher mediation. Guided inquiry or discovery is superior to inquiry alone and more direct didactic instruction (Alfieri et al., 2011; Furtak et al., 2012). Another example occurs across a range of applications of strategy instruction where the focus of instruction shifts to the strategies themselves and formulaic teaching of these, which has been called the risk of “proceduralisation” (Wilkinson and Son, 2010). Yet another example is how to mitigate the default to classroom discourse comprised of limited question and answer sequences, and reliably increase extended discourse which triggers more epistemic talk from the learner (Lefstein et al., 2021).

There is more to know about how to identify and incorporate cultural practices in optimizing the learning for culturally diverse children. For example, the qualities of complex dialogic exchanges in classrooms can be conceived in ways that reflect cultural meanings for students from Pacific nations, leading to more effective professional development to change classroom practices (Oldehaver, 2018). Oldehaver’s framing of dialogic turns using cultural meanings follows the instructional design of “cultural modeling” in English Language Arts (Lee, 2007). Effective teaching needs to be able to activate the learner’s existing knowledge and skill embedded in cultural practices and make explicit how to transfer from their cultural repertoires to using similar practices on unknown and canonical texts in school curricula. Among the questions which challenge systems to be more consistent at scale is what do all teachers need to know and how are they best prepared to engage with learners in the ways Lee and Oldehaver and others have described.

What is needed for successful optimization

Resolving the above issues requires greater clarity about what counts as successfully optimizing. It also requires a research and development capability based on the science of implementation. This is needed to meet the Bryk et al. (2015) challenge to nuance our understandings for different contexts, conditions, groups of students and across time.

Clarity about what counts as success

Optimizing learning has three properties, the outcomes are stochastic, the criteria are dynamic, and the criteria are values based. Each of these properties is present in how interventions at scale can view success for low progress, low achieving groups traditionally not well served by schools. In this illustrative case, the interplay between rate of learning and level of achievement of what has been learned is key to what schools can do to address the disparities associated with poverty and marginalization.

Given differential rates of progress emerging in the early years on school related measures of learning such as school readiness, optimization through schooling can be defined as *accelerated learning* such that students; (a) make more than the expected rate of rates of progress, and (b) as a group come to approximate expected distributions of achievement. For groups traditionally not well served by schooling, the goal is to accelerate sufficiently to match the distribution of all students in the defined “universe” of students. These are usually the population on which standardized assessments are based, with the probability of being in the top, middle and lower bands of that distribution being the same as the appropriate population at scale.

Values are central to such a definition of success (Kyriakides and Creemers, 2011). The matched distribution criterion is based on a value of equitable outcomes, as distinct from equitable input or opportunities. Because curricula are templates for what a society values in its citizens, the success criterion also is by definition values based. Whether the curriculum values themselves are equitable is an additional question, and part of the conditions for enabling optimization to occur.

In this example, the phenomena of success are probabilistic in the sense that a distribution is expected. But the outcomes are probabilistic in a second sense that both the processes and outcomes are variable, and as will be argued below, there should be some, tolerable levels of variability. Finally, the criterion is changeable as new curricula and new measures such as discussed below for critical literacy come to be adopted.

Optimizing in business as usual and within planned interventions has to solve a classic balance problem related to success criteria and the variability inherent in attempts to optimize. It is getting the balance right between prescription to ensure consistency through fidelity to effective instructional design, and enabling local capability to innovate.

In another form it is also the degrees of freedom needed to adapt what is generic to the particular needs of the individual learner. This is the challenge of being an “adaptive expert” (Darling-Hammond and Bransford, 2005). Being both expert and adaptive requires something akin to jazz extemporization, or what theories of social structures also call “deep improvisation” (Tilly, 1998). The capability to extemporize, to be adaptive, arises from high levels of local knowledge and moderate levels of scripting.

Solving the variability conundrum requires a balance between prescription and opportunity to innovate. The three properties of being stochastic, dynamic and values based suggest that the way to address the variability inherent in our attempts to optimize is to set tolerance levels for variability. Rather than having a fixed binary view of having optimized, interventions or more broadly, educational sites, could set shared targets for tolerance levels for variability. The desired net effect of setting these should be to increase consistency without reducing innovation.

A science of optimization

The implementation science position taken by Bryk et al. (2015) with which this chapter started, sets a direction for optimization research. Considering what works for whom under what conditions and at scale, and is sustainable over time, directs attention to understanding interventions at scale as a scientific problem in itself. This is also the case for optimization, solving the variability issues requires better science.

Implementation science and the wider agenda for design-based research (Snow, 2015) provides a methodology for optimization research. Using this methodology, variability can be a source of evidence for local innovation and the solving of educational challenges. Capitalizing on progressive adaptations adds power to the design of interventions. Even though teachers' knowledge of the innovative practices may be incipient and "pre scientific" (Vygotsky, 1978), progressive innovations are potentially able to be disciplined by research partners through technical language, testing and by being made explicit. The core premise here is that difficulties in problems of scaling and sustainability can be better solved with elements of local solutions in the design. This is because local solutions are likely to be more relevant and acceptable to others in that context, and also more contextually valid in that they are reflective of conditions in the local context not appreciated by outsiders.

This requires a different view of treatment integrity as not just prescription with adherence to the program design. We should see it as building high levels of capability to act flexibly and adaptively as "adaptive experts" based on knowing and acting with the core components of the program design, within a community of experts. This capability is dependent on a moderate level of prescription rather than a high level. Low prescription can provide opportunity for capability building but increases the likelihood of inequitable variability below tolerance limits. High prescription is antithetical to longer term capability building. It is also dependent on the embedded discourse and pedagogically productive talk that occurs in the community of practice to which teachers belong (Lefstein et al., 2021).

Clearly, building the capability to optimize is not just a function of setting a "moderate" level of prescription in the instructional design, be it from an intervention program or in business as usual. It is dependent on the preparation and consequent proficiency levels of teachers. It is dependent on ongoing professional development. It is dependent on the conditions and contexts identified in the Reading Recovery evaluation and the summary of context effects on implementation fidelity identified by Durlak and Dupree (2008).

Mitigating risks to optimizing

Trying to optimize comes with risks and potential untoward side effects. But these can be mitigated. Some of these have been noted already, but three closely related risks to successful optimizing are particularly problematic: inequitable resourcing and outcomes, narrowing the educational benefit, and the risk of cultural blindness. The principles for mitigating these risks include differentiating optimization, setting a wide focus for optimizing, and preserving identity language and culture.

Differentiated optimization

Unfortunately, the default condition in instruction, and at the level of interventions and business as usual in schools is rich get richer/poor get poorer effects, or "Matthew Effects" (McNaughton, 2011; Stanovich, 1986). The risk is some may benefit more from optimizing than others. The phenomena take two forms. A "match" form occurs where those with more relevant and more articulated knowledge and skills can make greater sense or use of conditions that are universally provided. In the beginning stages of literacy learning at school, those children already with advanced literacy skills and knowledge can make greater progress from the instruction provided and thereby gain faster access to more and more complex activities for learning, such as more complex texts for reading which expose them to more complex text features including vocabulary. And learning these increases skills levels and access further.

There is also a systemic form where levers such as funding or the provision of high quality courses and highly qualified teachers are systematically different for schools serving different communities. Both phenomena are ubiquitous over a wide range of structural features and activities in education (Ceci and Papierno, 2005), both in universal provision as well as targeted interventions.

Mitigating the risk is a matter of how well differentiated and fit-for-purpose interventions and everyday instruction are. The need for differential targeting in everyday provisions is illustrated by the "summer slide" patterns, where progress in reading, writing and maths during the summer break in schooling tends to be slower for students from communities with less access to the social, cultural and financial capital that make a difference to school based learning. These are often children from low SES and from marginalized communities. Universal conditions during school time are not sufficient to offset the out-of-school conditions and early differences in school-related learning tend to persist or grow across time at school. Several optimization responses are possible ranging from increasing the "dosage" of effective instruction for learners at risk of Matthew effects, through to changing the outside conditions by providing resources for school-related learning such as sending books home (Kim and Quinn, 2013).

Focusing on higher order and interdependent skills

A more general challenge to being successful facing many school systems currently, is whether we optimize learning across all the domains we value or have come to value—not only the usual core subjects of reading writing and mathematics, but also the higher order, so-called, 21st century skills. These are commonly divided into three categories or clusters. Two of these clusters, which have been called non-cognitive skills (Kautz et al., 2014), consist of interpersonal (e.g., perspective taking, empathy and prosocial skills) and intrapersonal (e.g., self-regulation and persistence) skills. Despite the categorization both clearly include cognitive elements. For example, regulation of attentional, emotional and behavioral actions and guiding or making reflective decisions about learning rely on strategies of prospection and metacognition deployed in recursive situation–attention–appraisal–response sequences (Duckworth et al., 2014). The third cluster focuses specifically on the cognitive areas of reasoning and analysis (e.g., critical thinking, critical literacy, and critical reasoning/argumentation).

The answer to the question of whether we attempt to optimize learning across valued domains as much as we should is—“not enough.” Take for example, critical literacy, which is a focus in many educational systems and is now a core part of international assessments (www.oecd.org/pisa/test). Critical literacy (“media literacy,” “epistemic vigilance”) skills are increasingly recognized as important for resilience and more generally as core citizenship skills. These do not just sum to a predisposition to questioning things or the status quo, but include cognitive skills of “criticality” (e.g., critical evaluation, reasoned judgments, identifying accuracy and credibility of information).

Access to information online, online communities and social media markedly increase the potential for mis-, dis- and mal-information, manipulation and prejudice. Risks to our valued outcomes for schooling come from how opinions can be manipulated with the rapid and extensive access to information and with the use of smart social-media platforms. Self-perpetuating and reinforcing systems of knowledge are created whereby the various degraded information sources, inaccuracies or untruths are taken as truths through the repetition and support within the network (Supovitz et al. (2017). The potential consequence is the uncritical adoption of positions on political and other issues, and uninformed resistance to alternative views.

Despite the curriculum focus and the increased need, students’ ability to reason effectively, navigate sources of information confidently, and engage critically with the systems of knowledge in content areas, are not taught well, across schooling and even at college level (Fraillon et al., 2018; McGrew et al., 2017).

Optimization will need to guarantee transfer from classroom to everyday use with these problematic information sources. Currently, the consensus is that a “mixed” approach involving explicit instruction in thinking critically embedded in major subjects, coupled with some dedicated course work focused on the critical literacy (media literacy) skills, such as in civics or philosophy courses, likely work best for transfer (Horn and Veermans, 2019). In addition to embedded and deliberate forms of instruction, communities of practice at a classroom and school level, which feature caring teacher and student relationships, clear norms, expectations and practices, create the effective conditions needed for the acquisition of these skills (Abrami et al., 2015). As with the opening example of Reading Recovery, conditions at a whole school level, as well as specific instructional moves need to be in place for optimizing.

These different domains of learning valued by schooling systems have interdependencies (Resnick et al., 2015). The reciprocal relationships between academic achievement and well-being for example are clear. Interpersonal skills (e.g., sociability) and intrapersonal skills (e.g., self-regulation) predict long-term educational and post-schooling success (Kautz et al., 2014). More specifically, levels of self-control and self-regulation are consistently associated with achievement outcomes (Duckworth et al., 2014; Zimmerman, 2008) and these skills and their relationships with cognitive development are identifiable early. For example, relationships are found between self-regulation, empathy and prosocial behavior, with emergent literacy and numeracy at the transition to school (OECD, 2020).

A research agenda for optimization includes designing for these interdependencies as a means of adding value to outcomes, but also because building the interdependencies is necessary for citizenship skills and resilience. The need and opportunities to optimize the learning of critical literacy and social and emotional skills has increased with the rapid expansion and extension of digital environments creating ubiquitous access and use by learners.

Digital worlds add urgency because social media increases threats to social and emotional development (Courage et al., 2015; Reid Chassiakos et al., 2016). Users of social media tend to have reduced capacity to judge the appropriateness of their own behavior and new vehicles for bullying are created (Aiken, 2016; Foody Samara Carlbring, 2015; Gardella et al., 2017).

An example of building on interdependencies is the instructional design for argumentation or collaborative reasoning. The designs require critical thinking skills as well as perspective taking and aspects of cognitive and emotional empathy (Kuhn, 2005; Rapanta et al., 2013). In well-designed interventions with classroom communities, skills are found to be generalizable across curriculum areas, from oral to written modes and from peer to peer (e.g., Reznitskaya et al., 2001; Kuhn, 2005; Rapanta et al., 2013).

However, as with critical literacy, low levels of argumentation and weak skills generally have been reported across grade levels and disciplines (Litma and Greenleaf, 2017; Kuhn, 2005), largely attributed to the absence of explicit instructional designs and deliberate socialization at class and school levels. The issue once again is how to take to scale the designs known to be effective. In addition to the specific instructional events, it means overarching enablers such as curriculum designations and widely socialized values such as those present in Finland’s anti-fake news campaign and revision to their critical thinking curriculum (Horn and Veermans, 2019). It also means a widely available and equitable research and development capability to partner with schools to co-design classroom, school and district level conditions.

Recognizing the significance of identity, language and culture

One form of the differential impact that educational systems have is through their treatment of culturally diverse students. Mainstream systems, even those identified as a “high quality” system, can have near optimal outcomes for many but which undermine or exclude culturally based skills and practices of some. This is despite evidence about the significance of identity, language and culture in equitable educational success.

The evidence from specific frameworks to optimize learning for culturally diverse students such as cultural modeling (Lee, 2007), and more generally, culturally responsive pedagogy is that learning is more robust by creating instructional bridges for generalizable learning and through recursive impacts on self-efficacy and well-being for robust learning. The relationships are detectable from large-scale studies showing that within and between countries the learner’s sense of belonging and being respected predict achievement in reading literacy at 15 years (OECD, 2017).

Optimizing for culturally diverse students from marginalized communities can be seen as a special case of personalization of learning (Cazden, 2001). Optimizing for this at scale poses daunting challenges for jurisdictions to solve. Reiterating the implications made above, it includes patterns of recruitment and training in initial teacher education, as well as interventions which are inclusive and additive, rather than exclusive and subtractive.

Conclusion: is optimization a chimera?

These descriptions of the limitations of our attempts at optimizing might suggest that we should reject the idea as unattainable. Certainly, the idea that a state of having optimized can be derived from a process of describing and then explaining was naïve, given the features of processes and outcomes being probabilistic, the criteria for what counts as success being changeable, and our intentions being framed by values.

However, despite the caveats, there are clear directions current evidence indicates we can take. Two lessons in particular can be drawn from the explanations of optimizing. These are the need to be very clear about what counts as optimizing, and leveraging off the science of implementation to inform optimization: in different contexts and under different conditions, for different groups of students, and in ways that are sustainable.

References

- Abrami, P.C., Bernard, R.M., Borokhovski, E., Waddington, D.I., Wadwe, C.A., Persson, T., 2015. Strategies for teaching students to think critically: a meta-analysis. *Rev. Educ. Res.* 85 (2), 275–314.
- Aiken, M., 2016. *The Cyber Effect*. John Murray Publishers.
- Alfieri, L., Brooks, P.J., Aldrich, N.J., Tenenbaum, H.R., 2011. Does discovery-based instruction enhance learning? *J. Educ. Psychol.* 103 (1), 1–18.
- Appleton-Dyer, S., Boswell, A., Reynolds, J., 2019. *Reading Recovery Evaluation: Report for the Ministry of Education*. Reading-Recovery-Evaluation-Report-13-May-2020.pdf. educationcounts.govt.nz.
- Baltes, P.B., Reese, H.W., Nesselroade, J.R., 1977. *Life-Span Developmental Psychology: Introduction to Research Methods*. Brooks/Cole, Monterey.
- Bryk, A.S., Gomez, L.M., Grunow, A., LeMahieu, P.G., 2015. *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard Education Press, Cambridge, MA.
- Cazden, C., 2001. *Classroom Discourse: The Language of Teaching and Learning*, second ed. Heinemann, Portsmouth, NH.
- Ceci, S.J., Paperno, P.B., 2005. The rhetoric and reality of gap closing: when the “have nots” gain but the “haves” gain even more. *Am. Psychol.* 60 (2), 149, 16.
- Coalition4evidence, 2013. <http://coalition4evidence.org/wp-content/uploads/2013/06/IES-Commissioned-RCTs-positive-vs-weak-or-null-findings-7-2013.pdf>.
- Courage, M.L., Bakhtiar, A., Fitzpatrick, C., Kenny, S., Brandeau, K., 2015. Growing up multitasking: the costs and benefits for cognitive development. *Dev. Rev.* 35, 5–14. <https://doi.org/10.1016/j.dr.2014.12.002>.
- Darling-Hammond, L., Bransford, J., 2005. *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, San Francisco, CA.
- Datnow, A., 2005. The sustainability of comprehensive school reform models in changing district and state contexts. *Educ. Adm. Q.* 41 (1), 121–153. <https://doi.org/10.1177/0013161X04269578>.
- Duckworth, A.L., Gendler, T.S., Gross, J.J., 2014. Self-control in school-age children. *Educ. Psychol.* 49 (3), 199–217. <https://doi.org/10.1080/00461520.2014.926225>.
- Durlak, J.A., DuPre, E.P., 2008. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41 (3–4), 327–350.
- Ericsson, A.K., 2006. The influence of experience and deliberate practice on the development of superior expert performance. In: *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, Cambridge, UK, pp. 683–703.
- Foody, M., Samara, M., Carlbirg, P., 2015. A review of cyberbullying and suggestions for online psychological therapy. *Internet Interv.* 2 (3), 235–242.
- Frailon, J., Ainley, J., Shulz, W., Friedman, T., Duckworth, D., 2018. *Preparing for Life in a Digital World: IEA International Computer and Information Literacy Study 2018 International Report*.
- Furtak, E.M., Seidel, T., Iverson, H., Briggs, D.C., 2012. Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev. Educ. Res.* 82 (3), 300–329.
- Gardella, J.H., Fisher, B.W., Teurbe-Tolon, A.R., 2017. A systematic review and metaanalysis of cyber-victimization and educational outcomes for adolescents. *Rev. Educ. Res.* 87 (2), 283–308. <https://doi.org/10.3102/00346554316689136>.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Hill, H.C., Erickson, A., 2019. Using implementation fidelity to aid in interpreting program impacts: a brief review. *Educ. Res.* 48 (9), 590–598.
- Horn, S., Veermans, K., 2019. Critical thinking efficacy and transfer skills defend against “fake news” at an international school in Finland. *J. Res. Int. Educ.* 18 (1), 23–41.
- Kautz, T., Heckman, J.J., Diris, R., Bas ter Weel, B., Borghans, L., 2014. *Fostering and Measuring Skills. Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success*. NBER Working Paper Series. Working Paper 20749. <http://www.nber.org/papers/w20749>.
- Kim, J.S., 2019. Making every study count: learning from replication failure to improve intervention research. *Educ. Res.* 48 (9), 599–607.

- Kim, J.S., Quinn, D.M., 2013. The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: a meta-analysis of classroom and home interventions. *Rev. Educ. Res.* 83 (3), 386–431.
- Kuhn, D., 2005. *Education for Thinking*. Harvard University Press.
- Kyriakides, L., Creemers, B.P.M., 2011. Can schools achieve both quality and equity? Investigating the two dimensions of educational effectiveness. *J. Educ. Stud. Placed Risk* 16 (4), 237–254. <https://doi.org/10.1080/10824669.2011.610269>.
- Lee, C., 2007. *Culture, Literacy, and Learning: Taking Bloom in the Midst of the Whirlwind*. Teachers College Press, New York, NY.
- Lefstein, A., Vedder-Weiss, D., Segal, A., 2021. Relocating research on teacher learning: toward pedagogically productive talk. *Educ. Res.* 19 (5), 360–368.
- Litma, C., Greenleaf, C., 2017. Argumentation tasks in secondary English language arts, history, and science: variations in instructional focus and inquiry space. *Read. Res. Q.* 53 (1), 107–126.
- May, H., Gray, A.M., Gillespie, J., Sirinides, P.M., Sam, C., Goldsworthy, H., Armijo, M., Tognatta, N., 2013. Evaluation of the i3 Scale-Up of Reading Recovery. Year One Report, 2011–2012. Consortium for Policy Research in Education, Philadelphia.
- May, H., Sirinides, P.M., Gray, A.M., Goldsworthy, H., 2016. Reading Recovery: An Evaluation of the Four-Year i3 Scale-Up. Consortium for Policy Research in Education, Philadelphia, PA.
- May, H., Blakeney, A., Shrestha, P., Mazal, M., Kennedy, N., 2022. Long-Term Impacts of Reading Recovery Through Third and Fourth Grade: A Regression Discontinuity Study From 2011–12 Through 2016–17. Paper Presented at the Annual Conference of the American Education Research Association (AERA) in San Diego, CA April 23, 2022.
- McGrew, S., Ortega, T., Breakstone, J., Wineburg, S., 2017. The Challenge that's bigger than fake news. *American Educator* Fall 2017 Civic Reason. Soc. Reason. Environ. 4–9.
- McNaughton, S., 2011. *Designing Better Schools for Culturally and Linguistically Diverse Children: A Science of Performance Model for Research*. Routledge, London.
- McNaughton, S., 2018. *Instructional Risk in Education*. Routledge, New York, NY.
- McNaughton, S., 2021. The conundrum of research-practice partnerships face with system variability. *Stud. Educ. Eval.* 70, 1–7. <https://doi.org/10.1016/j.stueduc.2021.101048>.
- Nesbit, K.T., Farran, D.C., 2021. Effects of prekindergarten curricula: tools of the mind as a case study. *Monogr. Soc. Res. Child Dev.* 86 (No. 1), Serial No. 340.
- OECD, 2017. PISA 2015 Results (Volume III): Students' Well-Being. OECD Publishing. <https://doi.org/10.1787/9789264273856-en>.
- OECD, 2020. Early Learning and Child Well-Being: A Study of Five-Year-Olds in England, Estonia, and the United States. OECD Publishing. <https://doi.org/10.1787/39904071-en>.
- Oldehaver, J.L., 2018. Developing a "culturally validated" dialogic indicator tool: a reconceptualised analytic framework using talanoa to code classroom talk. *Waikato J. Educ.* 23 (1), 25–41.
- Paris, S.G., 2005. Reinterpreting the development of reading skills. *Read. Res. Q.* 40 (2), 184–202.
- Rapanta, C., Garcia-Mila, M., Gilbert, S., 2013. What is meant by argumentative competence? An integrative review of methods of analysis and assessment in education. *Rev. Educ. Res.* 83 (4), 483–520.
- Reid Chassiakos, Y.L., Radesky, J., Christakis, D., AAP council on communications and media, et al., 2016. Children and adolescents and digital media. *Pediatrics* 138 (5), e20162593.
- Resnick, L.B., Asterhan, C.S.C., Clarke, S.N. (Eds.), 2015. *Socializing Intelligence Through Academic Talk and Dialogue*. AERA, Washington, DC.
- Reznitskaya, A., Anderson, R.C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., Kim, S.Y., 2001. Influence of oral discussion on written argument. *Discourse Process.* 32 (2–3), 155–175.
- Schochet, P.Z., Puma, M., Deke, J., 2014. Understanding Variation in Treatment Effects in Education Impact Evaluations: An Overview of Quantitative Methods. NCEE 2014-4017. National Center for Education Evaluation and Regional Assistance.
- Snow, C.E., 2015. 2014 Wallace foundation distinguished lecture: rigor and realism: doing educational science in the real world. *Educ. Res.* 44 (9), 460–466. <https://doi.org/10.3102/0013189X15619166>.
- Stanovich, K.E., 1986. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21 (4), 360–407.
- Supovitz, J., Daly, A.J., del Fresno, M., Kolouch, C., 2017. #Commoncore Project. Retrieved from. <http://www.hashtagcommoncore.com>.
- Tharp, R.G., Estrada, P., Dalton, S.S., Yamauchi, L.A., 2000. *Teaching Transformed: Achieving Excellence, Fairness, Inclusion and Harmony*. Westview Press, Boulder, CO.
- Tilly, C., 1998. *Durable Inequality*. University of California Press, Los Angeles, CA.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Wilkinson, I.A.G., Son, E.H., 2010. A dialogic turn in research on learning and teaching to comprehend. In: Kamil, M.L., Pearson, P.D., Moje, E., Afflerbach, P. (Eds.), *Handbook of Reading Research*, vol. 4. Routledge, New York, NY, pp. 359–387.
- Yeager, D.S., Fong, C.A., Lee, H.Y., Espelage, D.L., 2015. Declines in efficacy of anti-bullying programs among older adolescents: theory and a three-level meta-analysis. *J. Appl. Dev. Psychol.* 37, 36–51. <https://doi.org/10.1016/j.appdev.2014.11.005>.
- Zimmerman, B.J., 2008. Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *Am. Educ. Res. J.* 45 (1), 166–183. <https://doi.org/10.3102/0002831207312909>.

Apprenticeship learning

Julia Eberle, Faculty of Philosophy and Education, Ruhr-University Bochum, Bochum, Germany

© 2023 Elsevier Ltd. All rights reserved.

Characteristics of apprenticeship learning	568
Roots and developments of apprenticeship learning	569
Historical roots and development	569
Roots and development in educational research	570
Apprenticeship learning from the learning perspective	570
Apprenticeship learning in communities of practice	570
Developmental perspective on apprenticeship learning	571
Apprenticeship learning from the teaching perspective	571
Apprenticeship learning in vocational training	571
Apprenticeship learning in academic fields: the cognitive apprenticeship framework	572
Learning content: several types of knowledge	572
Methods in the learning process	572
Sequencing learning activities	573
Social context of learning	573
Conclusion	573
References	573

Apprenticeship learning is a central macro-approach to how learning and teaching can take place, standing in contrast to schooling as the second central macro-approach (Eberle, 2018). In the apprenticeship learning approach, learning takes place in the real-life context in which the content to be learned is of immediate relevance. Learning is understood as developing expertise in practices relevant in the learner's current life and comes with engaging in these practices. Furthermore, learning takes place in a fully contextualized way, within the real-life context in which experienced practitioners carry out their activities, and learning goes alongside the regular flow of practices. In addition, apprenticeship learning is situated in a meaningful social context of practitioners, the so-called community of practice. Social relations in the community of practice usually can be characterized either as person-centered, defined by a close relationship between a guiding master and one or several apprentices, or as decentered, when the learner is more loosely related to the members of the community of practice (Nielsen, 2010). Apprenticeship learning is mostly applied to observable practices but can also be applied to cognitive skills.

This chapter first addresses the central characteristics of apprenticeship learning and contrasts them to the schooling approach as the second major macro-approach. Next, the roots and the development of the apprenticeship learning approach are described, on the one hand from a historical perspective, on the other hand from an educational research perspective. Then, two strands of research on apprenticeship learning are presented that take the learner perspective into account and provide insights into the mechanisms behind apprenticeship learning: apprenticeship learning in communities of practice, and a developmental perspective based on apprenticeship. In subsequent sections, the teaching/educational organization perspective is addressed. The common goal is to bring the apprenticeship learning and the schooling approaches together, looking first at the formal organizational perspective in vocational training, and next at an instructional design-approach that allows for combining both approaches in the classroom.

Characteristics of apprenticeship learning

Apprenticeship learning is a complex concept that comes in many different forms. A generally accepted definition does not exist as most concise definitions tend to focus on only some central characteristics of the concept but neglect others based on the purpose and context in which the definition is used. Starting with the relatively vague definition that apprenticeship learning is one of two major macro-approaches for the organization of teaching and learning, the characteristics of apprenticeship learning can best be understood by contrasting it to schooling as the second major macro-approach.

Apprenticeship learning is based on a "participation metaphor" of learning (Sfard, 1998). That is, learning is understood holistically as becoming a person who actively participates in a community of practice, knowledgeably performs relevant practices, and acts in accordance with the norms and values of this community. This idea of learning stands in contrast to the schooling approach that focuses primarily on the acquisition of explicit knowledge and the increase in a set of predefined skills. This core difference gives reason for further differences in how learning and teaching take place in the two approaches. These differences can be found in nine core aspects: the epistemic/functional context, the relevance of learning, the sequencing of learning tasks, the social context, the role of the expert, the responsibility for learning, the relationship between expert and learner, the learning content, and the central form of learning (Collins, 2015; Eberle, 2018).

Regarding the *epistemic/functional context*, apprenticeship learning takes place in a real situation when a real problem occurs or a real task is performed. This situation occurs naturally and learning is rather a side-effect, in contrast to schooling where artificial situations are created and specifically designed to foster learning. Apprenticeship learning is embedded in the same context in which the apprentices will still be when they have become experts; that is, learning and application are not separated in time and space but learning takes place during application. For example, a baker apprentice usually learns in the bakery, surrounded by experienced bakers and participates from the start in preparing bread that will be sold to customers. Consequently, what is learned in an apprenticeship context is of immediate *relevance* to performing a task or solving a problem. In the baker scenario, for example, the baker apprentice learns how to prepare a certain type of bread and how to treat the dough when the bread is needed. And when the dough does not rise as it is supposed to do, solutions for this problem needs to be found on the spot. The schooling approach, in contrast, aims at preparing learners for situations in a distant future.

This example also shows the difference between the two approaches regarding the *sequencing of learning tasks*. In the apprenticeship learning approach, learning opportunities emerge from the context and as part of the practices, either when incidents occur that provide a new situation in which learning is required or when the apprentice initiates specific learning situations to gain better understanding during regular practices. In general, learning tasks are primarily created from the bottom up and the central goal or outcome is very clear while learning about details follows the sequence of required activities toward the achievement of a goal. Apprentices learn by participating in the different tasks according to what is possible at their current state of expertise. For example, the baker apprentice may, in the beginning, do very simple tasks such as providing ingredients or cleaning the bakery, and only later do more complex activities such as preparing the dough and forming bread. In the schooling approach, in contrast, learning is structured top-down following a specifically designed learning sequence based on the structure of knowledge (e.g., going from one topic to the next, providing information chronologically) while possibilities for applying this knowledge later on are usually secondary.

In addition, the *social context* in apprenticeship learning is a real social context that existed before an apprentice entered the picture, and becoming a respected member of this social community is part of the apprenticeship learning experience. In opposition to this, the social context in schooling is created artificially with the best learning outcomes in mind, usually bringing together an expert and often a homogenous group of learners.

In both approaches, an expert has an important part, but the *role of the expert* differs. Following an apprenticeship learning approach, the expert is an experienced master in the practices and acts primarily as a coach during naturally occurring learning situations. The master's central task is, however, to perform the practices while the coaching role of the otherwise autonomously performing and self-regulated learning apprentice is secondary. In the schooling approach, the expert is usually a teacher who structures the learning experiences and designs the learning process.

Consequently, the *responsibility for the learning process* is very different for the two approaches. The teacher, on the one hand, is the central person who pre-structures and initiates learning processes. The master, on the other hand has a more reactive role. A master may notice learning affordances of the context and use this for an ad-hoc lesson, (e.g., when the bread is about to burn, the master baker may use the situation to explain to their apprentices the importance for setting a timer and that a certain smell indicates that the bread is ready). In many cases, the apprentices will be the ones to request teaching from the master, when they seek more information or guidance, and the master can react to these requests.

The *relationship between expert and learner* in the apprenticeship learning approach is, accordingly, relatively close and the master knows the individual strengths and weaknesses of the apprentices well. Apprenticeship learning also relies on many one-on-one interactions and the master is usually responsible for a very small group of apprentices. These aspects give the master room for individually tailored feedback and guidance that is in the zone of proximal development for each apprentice. Unlike masters, teachers can educate larger numbers of students at the same time but serving individual learners' needs is a challenge.

The *learning content*, for which apprenticeship learning is best suited, are observable practices. Schooling, in contrast, has its strength in academic skills that are "in the head" of people and largely unobservable. However, unobservable academic skills and practices can also be the content of apprenticeship learning but requires some specific methods (see cognitive apprenticeship below). The *central form of learning* is, therefore, observational learning and active doing in the apprenticeship approach, unlike the schooling approach, in which instruction and practice are the most dominant forms of learning.

Roots and developments of apprenticeship learning

Historical roots and development

Before industrialization, apprenticeship learning was the most common form of occupational learning. The term refers to a form of training agreement between a master and an apprentice regarding a craft, art, or trade and the earliest historical references of apprenticeship learning date back to the Babylonian age ([The Editors of Encyclopedia Britannica, 2020](#)). However, apprenticeship agreements often went much further than mere occupational training and were often described as part of or similar to an adoption in which the apprentice becomes, at least temporarily, a part of the master's family (e.g., [The Code of Hammurabi](#)). The reason for this is that many occupations require more than just the knowledge and skills to perform specific occupational activities but come with a specific lifestyle and values and are embedded in the larger society as well as in specific communities associated with the occupation. For example, being a baker requires not only knowledge about recipes and baking skills but also requires that the baker starts to prepare the dough very early in the morning, so the bread is ready when customers come to buy it for breakfast.

In medieval Europe, apprenticeship learning was organized in craft guilds. And although this learning was mostly applied in observable hands-on crafts and the acquisition of physical skills, predominantly cognitive occupations such as scientists or clerics also followed an apprenticeship approach ([The Editors of Encyclopedia Britannica, 2020](#)) of which elements are still visible today.

Alongside the apprenticeship approach, the schooling approach also has a long tradition but was used to provide training only for children of a small elite and only in cognitive domains. With the changed organization of work that came with the industrialization, and even more with the transition into a knowledge society, the schooling approach replaced the apprenticeship approach as the dominant way of training in many areas ([Nielsen, 2010](#)). However, as both approaches have distinct advantages that are missing when relying on only one of them, we increasingly see attempts to bring them together by bringing elements of apprenticeship learning into school or by combining apprenticeship learning and schooling in occupational learning.

Roots and development in educational research

In educational research, apprenticeship learning gained broader attention around the late 1980s. Researchers from multiple backgrounds started to look at aspects closely related to apprenticeship learning from different perspectives ([Nielsen, 2010](#)). Anthropologists contributed ethnographic studies on situated learning and learning as a social practice. Educational philosophers, cognitive scientists, and computer scientists explored situated cognition—how knowledge, and practices are represented in the human mind and how they can be simulated by technology—drawing conclusions for learning. Learning scientists, finally, explored ways to successfully implement elements of apprenticeship learning into classroom settings, drawing from Vygotsky's work on socio-cultural learning and Bandura's social learning theory.

Furthermore, in the upcoming branch of artificial intelligence, the term apprenticeship learning is used for describing a certain type of machine learning. Specifically, in this context, apprenticeship learning means that machines are trained by observing an expert model performing a task ([Abbeel and Ng, 2004](#)). This approach is especially helpful for complex learning tasks for which the desired outcome cannot easily be put into an algorithm.

Based on the different roots of research on apprenticeship learning, we can find different strands and perspectives on the topic. First, there is a research strand examining the learning process of apprenticeship learning. This research is predominantly conducted using ethnographic methodology. One part within this research strand investigates “classical” apprenticeship learning in communities of practice, looking mostly at newcomers in different contexts and analyzing the details of how they become recognized, knowledgeable, and skilled members. The other part of this research strand applies the concept of apprenticeship learning to child development, seeing the child as an apprentice in their surrounding world who increasingly learns to participate in everyday life while growing up. The second major research strand examines apprenticeship learning from a teaching perspective and how it can best be brought together with schooling. A sub-strand of this approach takes a macro-perspective on vocational education systems. Another sub-strand takes an instructional design perspective and explores ways of bringing elements of apprenticeship learning into classroom settings and applying the approach to mainly cognitive learning situations. This is called the cognitive apprenticeship approach.

Apprenticeship learning from the learning perspective

Apprenticeship learning in communities of practice

Apprenticeship learning is an inherently social approach to structure teaching and learning and the community of practice is the social and epistemic context in which it takes place. A community of practice is usually not externally designed but emerges on its own. The term “community of practice” has become widely known. It has even received attention beyond academia and has become popular in organizational contexts. This popularity has led to a still evolving definition of the term and lack of coherence in its use ([Lindkvist, 2005](#)). Nevertheless, several characteristics distinguish communities of practice from other social groups ([Barab and Duffy, 2000](#); [Barab et al., 2003](#)). The reason for forming the community is a common practice and/or a mutual enterprise shared by several people. To be successful in this enterprise, community members need each other; that is, mutual interdependence exists. Accordingly, the community creates opportunities for its members to interact with each other and allows all members, especially newcomers, to participate to some extent in the practices of the community. The diverse perspective that the members bring together, is valued as a core element of the community of practice, including minority perspectives. However, the community members have a certain amount of shared knowledge, values, and beliefs that holds the community together. These stem from a shared history over a significant part of the membership. The relationships among community members are considered as meaningful with a long-term perspective. Members of the community see value in it beyond the individual relationships but tend to care for the existence of the community in the long-term future, and, consequently, develop mechanisms to bring in new members, who at some point in time can take over the community.

The idea of communities of practice, and the apprenticeship learning process that takes place in them, dates back to [Lave and Wenger's \(1991\)](#) book in which they provide a variety of examples of apprenticeship learning situations, such as midwives in Yucatan, butchers in German butcher shops, and Alcoholics Anonymous. The authors define the apprenticeship learning process as legitimate peripheral participation in communities of practice. This emphasizes that, to increasingly become a member of the community, learners need legitimate access to the community, its practices and knowledge, and the interaction opportunities of the members. When legitimate access is granted, learners need opportunities to participate peripherally in the beginning and,

increasingly, opportunities for participation in more central aspects of the community's practices. Peripheral participation allows newcomers to actively participate already from the beginning and contribute to the community but also provide opportunities to listen and observe what more experienced members do and to get acquainted with the practices and values. With increasing understanding, knowledge, and skills, the former newcomer can take over more complex tasks. The "curriculum" of this learning starts with understanding the central result, such as delivering a healthy baby for midwives in Yucatan, and follows the necessary steps to achieve this goal with an increasingly active role in the process. The apprentices often work together with a master and other community members of varying experience levels, who can be observed and guide the learning process. How easy or hard becoming a members of the community of practice is depends on participation support structures used by the community, which enable legitimate peripheral participation (Eberle et al., 2014; Nistor, 2016).

The legitimate peripheral participation process during apprenticeship situations has been observed in various occupational contexts, such as police officers, nurses, psychiatrists and teachers (e.g., Lave and Wenger, 1991; Lambson, 2010). However, legitimate peripheral participation also takes place in contexts unrelated to occupational learning, such as for parents in home schooling communities (Safran, 2010) or in witch circles (Merriam et al., 2003).

Developmental perspective on apprenticeship learning

The concept of apprenticeship learning can also be applied to child development by seeing children as apprentices in their culture or as "universal" legitimate peripheral participants, who learn to use language, develop an understanding of the world, and ways to think, through interpersonal mechanisms (Rogoff, 1991; Rogoff et al., 2003). Building on Vygotsky's (1978) work on the zone of proximal development, children learn when being supported by more knowledgeable others through guided participation (Rogoff et al., 1993). Interaction with adults and peers can be seen as the driving factor for development and increasing participation in the world (Rogoff, 1991).

Ethnographic research in this field found that caregivers in highly industrialized cultures, in which the schooling approach is the major form of education, tend to adopt a more school-like approach early on in their interaction with toddlers. By contrast, caregivers in less industrialized cultures with primarily apprenticeship approaches to education show a more apprenticeship-oriented behavior toward toddlers. Each seeks to prepare children for their further education (Rogoff et al., 1993).

Similar to apprenticeship in occupational learning, the developmental apprenticeship approach sees the responsibility for initiating learning with the children instead of with the caregivers. Children peripherally participate in the everyday activities of their caregivers, such as performing household chores, and this context provides opportunities for learning. Children can observe what their caregivers do and caregivers guide children's observations and attention to the general context, often using few verbal instructions. Children can initiate further interaction around occurring activities or everyday objects they find interesting. Caregivers then react to children's interests and specifically show them how to do what they are interested in, and help them participate in the practices. This guided participation allows children to participate peripherally until they are able to do the activity without support. Caregivers with a predominantly schooling-oriented approach, in contrast, design specific learning situations outside of their own everyday activities. They also feel responsible for initiating learning situations instead of reacting to the children.

Apprenticeship learning from the teaching perspective

As the learning perspective shows, apprenticeship learning has several benefits that the schooling approach lacks. A central advantage of the apprenticeship learning approach is that instruction usually takes place when learners face an immediate need for new knowledge and are, therefore, willing to listen, observe, and learn, whereas the benefits for learning in the schooling approach are often more distal and abstract, making the maintenance of learners' motivation a challenge for teachers. Another advantage is that the learned knowledge can be immediately applied to a concrete problem when it is most helpful. The emphasis on the social context and the wholistic perspective on learning inherent in the apprenticeship learning approach is an urgently needed addition to the schooling approach. Schooling traditionally has focused on individual knowledge acquisition in a distinct domain and neglected social and affective aspects, as well as the importance of identity development. Two research strands explore how apprenticeship learning and schooling can best be brought together, the first taking a macro-institutional perspective on vocational education systems, the second taking a micro-instructional-design approach on how to implement aspects of apprenticeship learning in the classroom.

Apprenticeship learning in vocational training

In the context of vocational training, apprenticeship learning is one of the most common forms of work-based learning. It is often not clearly distinguished from related forms such as traineeship, on-the-job training, or internships. All of them are characterized as forms of learning that take place in the real work environment and are often combined with schooling that takes place in a classroom separate from the work environment (European Center for the Development of Vocational Training, 2021).

The European Center for the Development of Vocational Training (2021) defines apprenticeship learning as paid work experiences of at least 6 months that occur as a mandatory part of the curriculum in formal vocational training. Formal apprenticeship learning as part of vocational learning is very common, not only in Europe but worldwide, while the structure of apprenticeship

learning—how exactly it is combined with schooling, and how much time is spent in each of the settings—differs significantly between countries. However, a general finding across European countries is that combining an apprenticeship phase with schooling has a positive effect on employability ([European Center for the Development of Vocational Training, 2021](#)).

Apprenticeship learning in academic fields: the cognitive apprenticeship framework

To apply the beneficial aspects of the apprenticeship learning approach in highly academic settings, Collins developed the cognitive apprenticeship framework ([Collins 2015](#); [Collins & Kapur 2014](#)) developed the *cognitive apprenticeship framework*. As academic activities are hard to observe due to their largely cognitive nature, the cognitive apprenticeship approach describes aspects that need to be considered to make cognitive processes of experts accessible and to ensure that novices engage in meaningful cognitive activities and receive support when needed. The framework was originally targeted to be applied in K-12 schools and is based on research in such settings, but it has also been widely investigated in higher education settings ([Minshew et al., 2021](#)). Bringing together insights from a multitude of previous studies, the cognitive apprenticeship framework addresses four dimensions for successfully implementing the apprenticeship approach into academic settings: the content dimension (the types of knowledge that need to be focused on), the method dimension (how cognitive processes can be initiated and fostered), the sequencing dimension (in which order certain methods and cognitive processes should be implemented), and the social context (what social aspects need to be fostered).

Learning content: several types of knowledge

The goal of the cognitive apprenticeship approach is to promote the development of expertise and to enable learners to become active members of a community of academic practice. To achieve this goal, learners need several types of knowledge that go beyond the declarative and procedural knowledge of a given domain as is predominantly taught under the schooling approach ([Collins and Kapur, 2014](#)). Nevertheless, specific domain knowledge is also a central component in the cognitive apprenticeship approach. In the context of mathematics education, this may include understanding of specific terminology and mathematical signs, but also how to perform mathematical procedures. Besides domain knowledge, strategic knowledge is highly important as it allows learners to flexibly use their domain knowledge and to adapt to specific situations. Strategic knowledge, in the cognitive apprenticeship approach, refers to three types of knowledge ([Collins and Kapur, 2014](#)): heuristic strategies, which are effective ways to approach and solve typical problems; metacognitive strategies, which are ways to plan, execute, and reflect on learning and problem solving processes and assess the quality of the problem solving process and its outcome; and learning strategies, which enable learners to acquire new domain knowledge that seems necessary to solve a problem.

Methods in the learning process

According to the cognitive apprenticeship approach, it is necessary to (1) support learners in understanding and performing relevant practices on specific exemplary tasks, (2) help them generalize beyond the exemplary task, and (3) provide opportunities for them to proceed autonomously and in a self-regulated way beyond teacher provided tasks ([Collins, 2015](#)). The cognitive apprenticeship framework defines six specific methods that serve these goals ([Collins and Kapur, 2014](#)): cognitive modeling, coaching, and scaffolding serving the first purpose; articulation and reflection serving the second purpose; and exploration serving the third purpose.

Cognitive modeling aims at making invisible thinking processes of experts visible to the learners. To do so, teachers or experienced practitioners explicate what they think while solving a problem, such as a mathematical equation or interpretation of research data. This way, the order of thought processes and the way to approach a problem can be “observed.” Cognitive modeling can be implemented into learning processes in several forms. One prominent form is a live or video-recorded modeling example, in which a person demonstrates how to perform a problem solving process; another form is vicarious learning, in which another learner is observed while being taught to solve a problem by a knowledgeable person ([van Gog et al., 2019](#)). Worked examples are a third way of cognitive modeling: a classical worked example consists of an exemplary problem statement, a step-by-step description of the procedure to solve the exemplary problem, and the solution, whereas a heuristic worked example focuses on the underlying reasoning why certain choices are made during the problem solving process ([Mulder et al., 2014](#)). Combinations of these forms of cognitive modeling can also be used. Often not only correct ways to solve a problem are shown but also failed or erroneous attempts to approach the problem ([van Gog et al., 2019](#)).

Participating in relevant practices and doing on their own or in groups what experts do is a core aspect of (cognitive) apprenticeship. However, coming from the Vygotskian perspective that learning takes place in the zone of proximal development, until the learners become proficient in these activities and while they are still struggling, they need support when engaging in the practices. The cognitive apprenticeship framework describes coaching and scaffolding as the two core methods for supporting learners during their attempts to perform the practices. *Coaching* is a form of individual and personal support, for which the teacher needs to monitor students’ learning processes and intervene in the right moment, when a learner is stuck ([Collins and Kapur, 2014](#); [Collins, 2015](#)). However, monitoring cognitive processes in a typical classroom setting with a large group of learners as well as providing the right kind of support exactly when it is needed is a huge challenge. Researchers are, therefore, exploring ways to use technology to support teachers in these challenges. For example, teacher dashboards are a form of technological support which provides various information to teachers to help them see where individuals and groups of learners are in the learning process and whether help is needed ([van Leeuwen et al., 2019](#)). When help is needed, *scaffolding* comes into play. Scaffolding is a metaphor for all support measures that can be provided to students. An effective scaffold is contingent, that is, specifically tailored to the learner’s current need and, consequently, highly dependent on the teacher’s monitoring of the learning process;

it fades when learner abilities increase to avoid dependence on the scaffold; and while the scaffold fades, the responsibility for performing the task to be learned is slowly transferred to the student (van de Pol et al., 2010). Specifically helpful in this regard is feedback that makes learners aware when they are performing well and when problems show. Useful feedback may come from teachers or peers (e.g., Bolzer et al., 2015), but also from the learning environment. Intelligent tutoring systems are a digital learning environment that provides coaching and scaffolding in the form of a personal learning environment. Based on a model of the knowledge to be learned and a continuously updated model of the learner's current state in the learning process, intelligent tutoring systems adapt their support to the individual learner's needs and provide the right information and feedback exactly when it is needed, often using artificial intelligence (Mousavinasab et al., 2021). Intelligent tutoring systems have been successfully applied in many fields, including medical education, computer science, and language learning (Mousavinasab et al., 2021). Such systems are not meant to replace human teachers but allow them to provide optimal coaching and scaffolding in a clearly defined area of a domain when a large number of learners face a limited number of teachers, as it simulates a personal and adaptive master-apprentice-learning situation (Collins, 2015).

To foster deep understanding of underlying principles and avoid a situation where learning stays limited to specific situations and exemplary tasks or cases, the cognitive apprenticeship approach emphasizes *articulation* of and *reflection* about thought and reasoning processes. Prompting is an effective way to foster articulation and reflection, for example by triggering self-explanation of meta-cognitive processes (e.g., Bannert et al., 2015; Rau et al., 2015; Heitzmann et al., 2015). Having learners create representations to visualize complex content or arguments is another form of effective support (e.g., Ainsworth et al., 2011; Bell, 1997).

Finally, *exploration* fosters learners' autonomy in solving problems beyond what was done in the classroom and encourages them to find opportunities for learning and problem-solving on their own (Collins and Kapur, 2014; Collins, 2015).

Sequencing learning activities

In line with the curriculum of apprenticeship learning in communities of practice, the sequencing of learning activities is different from the schooling approach. However, as in a classroom setting, learning opportunities do not naturally emerge but need to be designed. Three rules can help to come as close as possible to the classical apprenticeship learning approach (Collins et al., 1988; Collins and Kapur, 2014; Collins, 2015). First, the bigger picture and main goal for learning needs to be clear to the learners before going into details of specific skills or tasks. Learners need a conceptual model of the overall task and what is important to complete the overall tasks, so they find it easier to understand why the details are important and can better monitor their learning processes accordingly. Second, to simulate peripheral participation, learners need to start with simple tasks and complexity increases in line with learners' increasing skills and knowledge. When the cognitive apprenticeship situation is well designed in this regard, learners often perceive feelings of success and rarely experience frustration. Third, to counteract the possible disadvantage of apprenticeship learning, that learners' do not generalize principles they learned but only apply them to the context in which the principles were learned, teachers need to foster cognitive flexibility (Spiro et al., 1991). To broaden but also deepen learners' knowledge and skills, the complexity of tasks and problems, as well as their diversity, needs to increase with learners' abilities.

Social context of learning

As the social context in which intermental activities take place plays a central role for apprenticeship learning, designing for meaningful social interactions and relationships is also crucial for the cognitive apprenticeship approach. The learning context in general requires a real-world or highly authentic learning task that comes with the inherent need for learners to acquire new knowledge and skills to solve a relevant problem. On the social level, this learning context needs to come with an inherent need for exchange and collaboration with other learners so a sense of community arises, as well as motivation and a sense of ownership for solving the problem (Collins, 2015; Collins and Kapur, 2014). This may be achieved by thoughtful implementation of collaborative learning or a learning community approach.

Conclusion

Apprenticeship learning has a long tradition and although it has, in many areas, been replaced by a schooling approach, it still exists in other fields, formally as well as informally. Looking at formal education, we see the benefits for bringing apprenticeship learning and schooling together and research provides evidence for the benefit of doing so. In vocational training, current trends show that combining phases of both approaches can be very beneficial to train persons that are well set for working in their job. On the school level, we see benefits of merging both approaches and this is also a strong field of ongoing research.

References

- Abbeel, P., Ng, A.Y., 2004. Apprenticeship learning via inverse reinforcement learning. In: Design Automation Conference: Proceedings 2004. Association for Computing Machinery; IEEE. <https://doi.org/10.1145/1015330.1015430>.
- Ainsworth, S., Prain, V., Tytler, R., 2011. Drawing to learn in science. *Science* 333 (6046), 1096–1097.
- Bannert, M., Sonnenberg, C., Mengelkamp, C., Pieger, E., 2015. Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Comput. Hum. Behav.* 52, 293–306. <https://doi.org/10.1016/j.chb.2015.05.038>.

- Barab, S.A., Makinster, J.G., Sheckler, R., 2003. Designing system dualities: characterizing a web-supported professional development community. *Inf. Soc.* 19 (3), 237–256. <https://doi.org/10.1080/01972240309466>.
- Barab, S.A., Duffy, T.M., 2000. From practice fields to communities of practice. In: Jonassen, D.H., Land, S.M. (Eds.), *Theoretical Foundations of Learning Environments*. Lawrence Erlbaum Associates, pp. 25–55.
- Bell, P., 1997. Using argumentation representation to make thinking visible for individuals and groups. In: Hall, R., Miyake, N., Enyedy, N. (Eds.), *Proceedings of the 2nd International Conference on Computer Support for Collaborative Learning*. International Society of the Learning Sciences, pp. 10–19.
- Bolzer, M., Stribos, J.W., Fischer, F., 2015. Inferring mindful cognitive-processing of peer-feedback via eye-tracking: role of feedback-characteristics, fixation-durations and transitions. *J. Comput. Assist. Learn.* 31 (5), 422–434. <https://doi.org/10.1111/jcal.12091>.
- Collins, A., 2015. Cognitive Apprenticeship: NAPLeS Webinar. ISLS. <http://isls-naples.psy.lmu.de/intro/all-webinars/collins/index.html>.
- Collins, A., Kapur, M., 2014. Cognitive apprenticeship. In: Sawyer, R.K. (Ed.), *Cambridge Handbooks in Psychology*. The Cambridge Handbook of the Learning Sciences, second ed. Cambridge University Press, pp. 109–127.
- Collins, A., Brown, J.S., Newman, S.E., 1988. Cognitive apprenticeship: teaching the craft of reading, writing and mathematics. *Thinking: The Journal of Philosophy for Children* 8 (1), 2–10.
- Eberle, J., 2018. Apprenticeship learning. In: Fischer, F., Hmelo-Silver, C., Goldman, S.R., Reiman, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge.
- Eberle, J., Stegmann, K., Fischer, F., 2014. Legitimate peripheral participation in communities of practice: participation support structures for newcomers in faculty student councils. *J. Learn. Sci.* 23 (2), 216–244. <https://doi.org/10.1080/10508406.2014.883978>.
- European Center for the Development of Vocational Training, 2021. The Role of Work-Based Learning in VET and Tertiary Education. Publications Office. <https://doi.org/10.2801/799490>.
- Heitzmann, N., Fischer, F., Kühne-Eversmann, L., Fischer, M.R., 2015. Enhancing diagnostic competence with self-explanation prompts and adaptable feedback. *Med. Educ.* 49 (10), 993–1003. <https://doi.org/10.1111/medu.12778>.
- Lambson, D., 2010. Novice teachers learning through participation in a teacher study group. *Teach. Teach. Educ.* 26 (8), 1660–1668. <https://doi.org/10.1016/j.tate.2010.06.017>.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Lindkvist, L., 2005. Knowledge communities and knowledge collectivities: a typology of knowledge work in groups. *J. Manag. Rev.* 42 (6), 1189–1210.
- Merriam, S.B., Courtenay, B., Baumgartner, L., 2003. On becoming a witch: learning in a marginalized community of practice. *Adult Educ. Q.* 53 (3), 170–188.
- Minshew, L.M., Olsen, A.A., McLaughlin, J.E., 2021. Cognitive apprenticeship in STEM graduate education: a qualitative review of the literature. *AERA Open* 7 (4). <https://doi.org/10.1177/23328584211052044>, 233285842110520.
- Mousavinasab, E., Zarifsanaley, N., Kalhori, S.R.N., Rakhshan, M., Keikha, L., Ghazi Saeedi, M., 2021. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interact. Learn. Environ.* 29 (1), 142–163. <https://doi.org/10.1080/10494820.2018.1558257>.
- Mulder, Y.G., Lazonder, A.W., de Jong, T., 2014. Using heuristic worked examples to promote inquiry-based learning. *Learn. Instr.* 29, 56–64. <https://doi.org/10.1016/j.learninstruc.2013.08.001>.
- Nielsen, K., 2010. Apprenticeship approach to learning. In: Peterson, P., Baker, E., McGraw, B. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, pp. 469–475. <https://doi.org/10.1016/B978-0-08-044894-7.00465-6>.
- Nistor, N., 2016. Quantitative analysis of newcomer integration in MMORPG communities. In: Li, Y., Chang, M., Kravcik, M., Popescu, E., Huang, R., Kinshuk, Chen, N.-S. (Eds.), *Lecture Notes in Educational Technology. State-of-the-Art and Future Directions of Smart Learning*. Springer, Singapore, pp. 131–136. https://doi.org/10.1007/978-981-287-868-7_15.
- Rau, M.A., Alevan, V., Rummel, N., 2015. Successful learning with multiple graphical representations and self-explanation prompts. *J. Educ. Psychol.* 107 (1), 30–46. <https://doi.org/10.1037/a0037211>.
- Rogoff, B., 1991. *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press.
- Rogoff, B., Mistry, J., Göncü, A., Mosier, C., 1993. Guided Participation in Cultural Activity by Toddlers and Caregivers, vol. 236. *Monographs of the Society for Research in Child Development*.
- Rogoff, B., Paradise, R., Arauz, R.M., Correa-Chavez, M., Angellillo, C., 2003. Firsthand learning through intent participation. *Annu. Rev. Psychol.* 54, 175–203. <https://doi.org/10.1146/annurev.psych.54.101601.145118>.
- Safran, L., 2010. Legitimate peripheral participation and home education. *Teach. Teach. Educ.* 26 (1), 107–112. <https://doi.org/10.1016/j.tate.2009.06.002>.
- Stard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27 (2), 4–13. <https://doi.org/10.3102/0013189X027002004>.
- Spiro, R.J., Feltovich, P.J., Jacobson, M.J., Coulson, R.L., 1991. Cognitive flexibility, constructivism, and hypertext: random access instruction for knowledge acquisition in ill-structured domains. *Educ. Technol.* 31 (5), 24–33.
- The Code of Hammurabi. <https://avalon.law.yale.edu/ancient/hamframe.asp>.
- The Editors of Encyclopedia Britannica, 2020. Apprenticeship. *Encyclopaedia Britannica*. <https://www.britannica.com/topic/apprenticeship>.
- van de Pol, J., Volman, M., Beishuizen, J., 2010. Scaffolding in teacher–student interaction: a decade of research. *Educ. Psychol. Rev.* 22 (3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>.
- van Gog, T., Rummel, N., Renkl, A., 2019. Learning how to solve problems by studying examples. In: Dunlosky, J., Rawson, K.A. (Eds.), *The Cambridge Handbook of Cognition and Education*, vol. 8. Cambridge University Press, pp. 183–208. <https://doi.org/10.1017/9781108235631.009>.
- van Leeuwen, A., Rummel, N., van Gog, T., 2019. What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *Int. J. Comput. Support. Collab. Learn.* 14 (3), 261–289. <https://doi.org/10.1007/s11412-019-09299-x>.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Argumentation and critical thinking

Chrysi Rapanta^a and Kalypso Iordanou^b, ^a Universidade Nova de Lisboa, Lisbon, Portugal; and ^b University of Central Lancashire, Larnaka, Cyprus

© 2023 Elsevier Ltd. All rights reserved.

Critical thinking as a starting point	575
Intentionality	576
Sensitivity to context	576
Clarity and precision	576
Boundedness to principles	576
Argumentation as an alternative construct to critical thinking	576
Reasonableness	577
Relevance	578
Sufficiency	578
Acceptability	578
Reasoning abilities, skills and dispositions implied in argumentative discourse	579
Generative abilities	579
Inferential abilities	580
Appraisal abilities	580
Evaluation of information	580
Evaluation of arguments	580
Evaluation of context	580
Metastrategic skills	580
Epistemic and epistemological awareness	581
Reasoning/arguing disposition	581
Evidence from argument-based pedagogical methods designed to foster students' critical thinking skills and dispositions	582
Transfer from social to individual: the case of writing	582
Transfer to other topics/contexts	582
From skills to dispositions	582
Critical thinking as an endpoint	583
Values	583
Reflexivity	583
Media and information literacy	584
Conclusion	584
References	585

Critical thinking as a starting point

Critical thinking (CT) is a widely accepted educational goal and ideal. At the same time, it is a highly ill-defined construct, with scholars not yet in agreement regarding what exactly it refers to, and most importantly how it can be fostered and assessed.

Regarding its definition, a main division among scholars concerns whether CT refers to a constructive process of problem solving and decision making, or solely to an appraisal of intellectual products. Scholars defending the former (e.g., [Bailin, 1987](#); [Ennis, 2001](#); [Glassner and Schwarz, 2007](#); [Lipman, 1987](#); [Paul, 1987](#); [Perkins, 1985](#)) also tend to claim that critical thinking and creative thinking overlap, because generating any type of idea, solution, or decision is itself a creative process. Scholars defending the latter (e.g., [Ennis, 1993](#); [Scriven, 1976](#); [Sternberg, 1986](#); [Watson and Glaser, 1964](#)) tend to separate critical thinking from any demonstration of creativity. A second division within the CT scholarly community regards whether CT skills are subject-dependent or subject-independent. Although some scholars argue in favor of subject specificity in the development of CT skills (see for example [Glaser, 1984](#); [McPeck, 1981](#)), the majority of scholars argue in favor of a subject-general nature of CT (e.g., [Ennis, 1987](#); [Higgins and Baumfield, 1998](#); [Paul, 1990](#); [Perkins and Salomon, 1989](#)), justifying the transfer of CT skills to several contexts. However, and this leads to the third main division, CT is composed not only of skills but also of attitudes or dispositions as well as knowledge. If subject-matter knowledge is considered a component of CT, the question arises as to whether this subject-dependency extends to CT skills, or attitudes, or both ([Ennis, 1989](#)). The relation between these CT components remains an issue under exploration and an object of academic divide ([Hitchcock, 2020](#)).

Despite the ill-defined and variable nature of the term critical thinking, authors agree on (at least) its following characteristics:

Intentionality

Ennis (1996) defines CT as “reasonable reflective thinking focused on deciding what to believe or to do” (p. 166). Similarly, Halpern (1998) perceives CT as a way of thinking that can be applied to everything as long as it includes a “deliberate use of skills and strategies that increase the probability of a desirable outcome” (p. 449). This conscious and deliberate exercise of higher-order reflection is what makes thinking, applied to any context, critical. Intentional thinking is any thinking that is aimed at a specific purpose, and as such it implies a learner’s reflection on how (much) and whether the purpose has been achieved. According to Dewey (1966), in intentional thinking “natural active tendencies shall be fully enlisted in doing something, while seeing to it that the doing requires observation, the acquisition of information, and the use of a constructive imagination” (p. 137). In other words, intentional thinking is active thinking and learning (Freire, 1993, called it *authentic learning*.) Intentional thinking, and subsequent learning, is the core and outcome of reflexivity and criticality: the more intentional one becomes in one’s thinking, the more reflective and critical the thinking becomes, and vice versa (Bennett et al., 2016).

Sensitivity to context

As Lipman (1987) argued, critical thinking is self-corrective thinking, but for this to be possible, sensitivity to context and boundedness to principles are necessary conditions. The situatedness of CT in a concrete context is indisputable: We cannot ask someone to think critically in general, but we can ask her to, for example, critically analyze a text, or engage in a critical discussion of a topic (Moore, 2013). This is one aspect of context in relation to CT—i.e., the fact that it is contextualized in a situated practice. A second main aspect is the fact that it is discursive, and therefore it is both dialogical and dialectical, and therefore it needs to take several contexts into consideration. The fact that CT is discursive (i.e., expressed in intra- or inter-personal discourse) is rooted in its relation to reflection (see Dewey’s definition of CT above). CT needs some kind of information as an object of processing and reflecting on in order to formulate a judgment about something. This implies that critical thinkers engage in some type of situated, discursive reflection in order to validate the information being assessed or communicated to/from them. If they do not do that, their only other recourse, as Mezirow (1997) affirms, is to turn to an authority or tradition to do this validation for them. In this sense, CT is dialogical, as engaging in discourse means engaging in “a dialog devoted to assessing reasons presented in support of competing interpretations, by critically examining evidence, arguments, and alternative points of view” (Mezirow, 1997, p. 6). And it is also dialectical, as any argument made as part of a critical thinking process needs to confront existing or potential counterarguments (Paul, 1981).

Clarity and precision

According to Halpern (2014), “learning the skills of clear thinking can help everyone recognize propaganda and thus not fall prey to it, analyze unstated assumptions in arguments, realize when there is deliberate deception, consider the credibility of an information source, and think a problem or a decision through in the best way possible” (p. 11). However, such clarity and precision must always leave a space for flexibility in order to be critical. As Halpern (2014) again affirms, “a substantial increase in the likelihood of a desirable outcome is the best that critical thinking can promise” (p. 9). With this awareness acquired, a critical thinker sets goals and does the best to achieve these goals (there is much literature affirming that critical thinking is also persisting thinking; see for example Dewey, 1910; Faccione, 1990; Halpern, 2014). Yet if the desired outcomes do not materialize, the critical thinker is eager to modify the plan of action, as evidenced in problem solving involving critical thinking attitudes and skills (Paul and Elder, 2006). Being clear and precise while at the same time flexible is a quality of critical thinkers.

Boundedness to principles

Thinking well in any area requires knowledge, but it is its relation to principles of good reasoning that thinking becomes critical thinking (Siegel, 1980). Those principles, however universal or subject-related, are the ones that make thinking sound, credible, and accountable. Examples of these include principles of interpretation and criticism, as well as relations between principles of interpretation and standards of assessment, and between arguers and their audience (Govier, 2018). Groarke and Tindale (2004) also remind us that the key to success is not simply knowing the principles and criteria of good reasoning but, above all, employing them in practice. This can be done in two main ways: either individually, through critical thinking or problem-solving exercise, or publicly, through argumentation practices that require, and therefore imply, the application of such skills and standards of good reasoning. This idea that CT is individual reflective, rational and goal-oriented reasoning, while argumentation is the social practice of CT is further explained in **Argumentation as an alternative construct to critical thinking** section onwards.

Argumentation as an alternative construct to critical thinking

Argumentation can be defined as the use of language to support an uncertain or potentially uncertain claim, typically contending with a difference between two or more parties (Rapanta and Macagno, 2019). This difference includes the incompatibility of points

of view concerning an expressed position; however, it is not limited to it. An argument can be advanced to address different types of doubts: “an unsolved problem, an unproven hypothesis, or even a situation where both parties are blocked from further actions they are trying to carry out” (Walton, 1990, p. 411). According to Groarke and Tindale (2004), a good arguer, just like a critical thinker, “must be able to recognize and judge: what counts as a good reason for a claim; when claims are relevant to an argument, and when they are not; what conclusions reasonably follow from different kinds of evidence; the difference between sufficient and insufficient evidence; and the expectations that attend different contexts and different audiences” (p. XII). Critical thinking is about reasoning in a critical manner, and this is what argumentation is all about. What is then the difference between the two? The difference lies in the distinction between argumentation and reasoning.

Reasoning is defined by the process of making inferences, or inferring, with argument being one of the various purposes reasoning can have (Angeles, 1981). According to Woods et al. (2004), “arguments in the broad sense are interpersonal, social phenomena—a kind of public event. Inferences, in contrast, are essentially private affairs. They involve the adjusting of an individual’s mental states, or the transition from one belief state to another. Thus, they are something the individual handles solo” (p. 6). Said differently, reasoning is a tool for arguing, manifested in discourse *with* others (as in the case of academic writing) or *together with* others (as in the case of dialogical argumentation).

One of the main limitations of the construct of CT is the absence of consensus as to how to assess it and foster it. Although scholars tend to broadly agree that CT is reasonable, reflective, intentional thinking involving knowledge, skills, and dispositions, there is no consensus as to when and how CT is manifested. According to Siegel (1985), the important question is not, “Is there a generalized skill of critical thinking?”, but rather, “How does critical thinking manifest itself?” (p. 75). Kuhn (1999) similarly asks “What does it mean to be accomplished in thinking, and how does this accomplishment manifest itself?” (p. 16).

Argumentation as a discursive social practice is the “battlefield” in which CT’s main characteristics are not only manifested, but also developed. Many would claim this occurs through engagement in dialog with others (e.g., Burbules and Berk, 1999; Kuhn, 2019; Rapanta, 2019a; Walton, 1989). Here we focus on how argumentative reasoning—that is, reasoning to argue—invokes explicit, assessable standards for critical thinking to manifest in discourse and further develop. Such standards overlap with argument assessment criteria of an informal logic nature, such as: reasonableness, relevance, sufficiency, and acceptability.

Reasonableness

Reasonableness is the umbrella criterion that has the same function for informal logic arguments as the concept of validity has for formal logic—that is, providing a means for distinguishing between sound and unsound arguments (Zarefsky, 1981). Reasonableness, also called rationality, refers to the degree that the premises chosen as the most relevant to support one’s claims and arguments lead to “reasonable” or “rational” conclusions, or are reasonable or rational themselves. A reasonable or rational judgment is non-arbitrary, impartial, and objective (Siegel, 1980).

Among CT scholars, “reasonable” has been used to describe either a position (Ennis, 2011) or the criteria implied to support a position (Faccione, 1990). In terms of the former, it is not clear what a reasonable position refers to. Regarding the latter, Ennis (2011) adds that the reasonableness of criteria refers not only to their quality but also to their comprehensiveness: the critical thinker must “employ a reasonable critical thinking checklist” (p. 17). However, within the CT literature, reasonableness is not well-defined.

How can argumentation theory help us define the reasonableness/rationality of an argument? Several scholars (Chin and Osborne, 2010; Mayweg-Paus et al., 2016; Nussbaum and Edwards, 2011; Song and Ferretti, 2013) claim that this is possible through posing the right critical questions. Walton (Walton, 1996; Walton et al., 2008) made the connection between critical questions and argument context more explicit by identifying critical questions for various types of everyday arguments, according to whether they express deductive (e.g., causal, consequential), inductive (e.g., reasoning from example, analogical reasoning) or abductive (e.g., reasoning from sign, from best explanation) reasoning (Macagno, 2015). The so-called argumentation schemes are “stereotypical patterns of inference, combining semantic-ontological relations with types of reasoning and logical axioms and representing the abstract structure of the most common types of natural arguments” (Macagno, 2015, p. 192). Walton et al. (2008) provide a list of critical questions accompanying each scheme, according to the topical, material and dialectical relations expressed between the argument premises and its conclusion, as well as between the argument and other possible arguments (Walton, 1998). For the argument from expert opinion, for instance, the following questions apply (see also Rapanta and Walton, 2016):

- (1) Expertise Question (Q1): How credible is E as an expert source?
- (2) Field Question (Q2): Is E an expert in the field F that A is in?
- (3) Opinion Question (Q3): What did E assert that implies A?
- (4) Trustworthiness Question (Q4): Is E personally reliable as a source?
- (5) Consistency Question (Q5): Is A consistent with what other experts assert?
- (6) Backup Evidence Question (Q6): Is E’s assertion based on evidence?

Different critical questions apply to different types of arguments. Reasonableness relates to the CT aspect of intentionality, because the degree to which the arguer searches to anticipate and reply to the type of critical questions listed above is largely defined by how (much) she recognizes the (dia)logical purpose of the argument. In other words, the capacity of an arguer to anticipate potential counterarguments corresponding to the type of critical questions each argument may elicit is a sign of reasonableness and

a condition for the construction of sound arguments, as opposed to fallacious ones (for the relation between critical questions, argument premises and fallacies see [Walton, 2010](#)).

If reasonableness is the “umbrella” criterion for assessing informal arguments, or for creating valid ones, relevance, sufficiency and acceptability express how this informal soundness or cogency can be operationalized in the field of informal (everyday, ordinary) logic.

Relevance

Argument relevance does not merely imply that contributions are on the topic and coherent with the contents discussed. This is only one aspect of relevance, which, in argumentative dialog, corresponds to the so-called *structural* relevance—how the argument components logically inter-relate one with another ([Macagno, 2008, 2016](#)). The other aspect of relevance refers to being consistent with the *pragmatic* function—that is, the purpose of the dialog, and the different forms it may take throughout interaction ([Macagno, 2016, 2019](#)). In argument skill development terms, the former type of relevant participation is contained in argument strategic and metacognitive skills, whereas the latter is expressed through the idea of metastrategic and epistemological awareness (see [Reasoning abilities, skills and dispositions implied in argumentative discourse](#) section). Being sensitive to the criterion of relevance is equivalent to being “dialectically astute”; this means being “alert to the need for enough evidence of the right kinds, and to the possibilities of counter-arguments and conflicting evidence” ([Blair and Johnson, 1987](#), p. 51). Therefore, argument relevance largely fulfills the CT characteristic of sensitivity to context (see [Critical thinking as a starting point](#) section).

Sufficiency

If probative relevance is assessed by the criterion that “the premises of an argument ought to be relevant to the conclusion”, argument sufficiency refers to the requirement that “the premises ought to provide sufficient support for the conclusion” ([Blair and Johnson, 1987](#), p. 54). Similarly to relevance, sufficiency also has a dialectical component: to decide whether or not an arguer sufficiently defends a claim and its premises, we must imagine the dialogical context within which those claims and premises will be put forward. Critical questions accompanying argumentation schemes are useful in this task. For example, when deciding whether an argument from expert opinion is sufficiently supported or not, questions assessing the sources of this expertise can be asked (e.g., How many years of practice does Doctor X have in the field?). Or, when assessing the validity of an argument from generalization, the following sufficiency question can be asked: How big shall a sample be in order for a research study to be valid? Overall, it can be said that sufficiency is an aspect of relevance applying to the relation between premises and conclusion, referring to how explicitly (i.e., clearly and precisely) this relation is expressed. If an argument is sufficient in terms of support provided for the conclusion, it follows that it is also clear and precise. In other words, the more one reveals the different premises that support a claim, the more content and context is added to the presuppositions of an argument (see, for example, the statement “Jack is overweight” used by [Paul and Elder \(2013\)](#), to show how a claim can be clear but not sufficiently precise, because of the lack of explicit premises). Therefore, argumentative reasoning connects to the CT criterion of clarity and precision (see [Critical thinking as a starting point](#) section) through the informal logic criterion of argument sufficiency.

Acceptability

Acceptability of an argument refers to whether its premises should be accepted by an audience as they are, without further support, or not. Even in arguments that appear to be deductively valid (i.e., when the premises are true, then the conclusion is also true), it is often the case that the premises are debatable or open to discussion. In those cases, the burden of proof lies with the arguer, meaning that the arguer needs to provide further support for these premises to show they are themselves convincing. Similarly to relevance and sufficiency, acceptability is something that can only be assessed in the context of dialog, due to its strong relation to the dialogical concept of burden of proof ([Blair, 2012](#)). Moreover, because when an argument’s premises are relevant to the conclusion and support it sufficiently, they are also acceptable, we can also say that acceptability corresponds to the CT component of boundedness to principles (see [Critical thinking as a starting point](#) section).

In summary, in this section we showed that argumentation can be considered an alternative construct to CT, because it binds to the informal logic criteria of reasonableness, relevance, sufficiency and acceptability, which strongly relate to the main CT components described in [Critical thinking as a starting point](#) section. Moreover, we argued that these criteria are assessable or make sense in a dialogic context: given that argumentation is dialectical in nature ([Blair and Johnson, 1987](#)), an interlocutor is at least implicitly present or taken into consideration for argumentation to take place. The situatedness and functionality of these argument assessment criteria in the context of dialogic practice is highly important as they become criteria of both generation and appraisal, combining therefore the critical and creative aspects of argumentative reasoning (see also [Glassner and Schwarz, 2007](#), for a similar perspective).

In the next section we further explain how argumentative discourse reveals (i.e., brings to the surface) all the critical thinking abilities, skills, and dispositions otherwise difficult to access and assess.

Reasoning abilities, skills and dispositions implied in argumentative discourse

Ennis (1993) remarked on the difficulty of several critical thinking assessment tests to truly capture the CT skills and dispositions they attempt to assess, particularly, when the skills and dispositions are expected to be transferred and applied to different contexts. As an alternative, a “strong sense” approach to CT is necessary, one that does not rely on the assumption that CT is “a battery of atomic technical skills” (Paul, 1981, p. 3), accessible through an assessment test. Argumentation provides the opportunity for individuals to engage in a practice that includes the following: construct an argument to support a claim, identify and make use of relevant evidence, generate/anticipate counterarguments, and respond to counterarguments with rebuttals (Kuhn, 1991). In oral discourse, this series of activities translates into a persuasive dialog, where the goal of the participants is to adopt a position, to defend it as well as possible, and to seek to persuade the interlocutor of its merits over an alternative. In written discourse, this translates into persuasive or argumentative writing, where the goal of the writer is to consider various, often contrary, perspectives and to provide an integrated (two-sided) argument as result of the weighing of positive and negative evidence of each position. The inherent goal of persuasion in the practice of argumentation as manifested in externalized discourse (both oral and written) implies the presence of key reasoning abilities, skills and dispositions of a similar nature to those associated with CT.

Generative abilities

More and more in educational policy documents around the world, students are expected to “generate and evaluate scientific evidence and explanations” and “participate productively in scientific practices and discourses” (Duschl et al., 2008, p. 2). Science here is conceived broadly, as a type of reasoning based on inquiry and argumentation (Kuhn, 1992, 1993). The generation and appraisal of evidence and explanations as well as the productive participation in argumentative discourse are all generative/constructive abilities entailed in the practice of argumentation (as we saw in **Critical thinking as a starting point** section, this is not always so in the case of CT).

Generating arguments and participating productively in argumentative discourse imply the successful (reasonable, relevant, sufficient, and acceptable) construction of argument claims and premises, as well as theory-evidence coordination. Premises (reasons offered in support of a claim) can be of various types, also according to the type and content domain of the claim. For example, a claim about a social issue (e.g., abortion) can be of the type “in favor/against” whereas a claim about a scientific phenomenon needs to contain a hypothesis (e.g., “If I drop an object, it will fall”). Among the types of premises used to support a claim, Toulmin (1958) identified two, namely a warrant, or the inferential link between claim and data, and a backing, namely the evidence used to further support this inference. The warrant replies to the “Why do you say so?” aspect of argumentative reasoning, while the backing to the “How do you know this is so?” (Kuhn, 1999, 2001). The theory-evidence coordination is a necessary aspect of generating an argument not only because of the selection of the right premises that will function as a support of it but also because of the necessary consideration of limitations (restrictions) to one’s own reasoning, structurally conceived by Toulmin (1958) with the argument element called “rebuttal.” Fig. 1 shows an example of the main argument structure of both a social and a scientific argument, with their premises (warrant and backing) and a potential rebuttal for each.

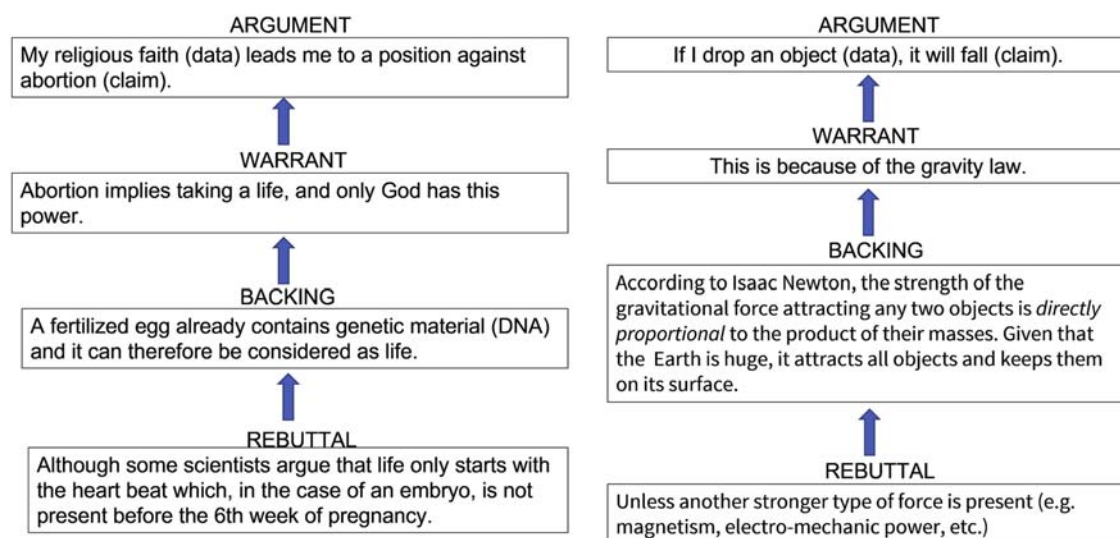


Fig. 1 An example of a social (left) and a scientific argument (right).

Inferential abilities

According to [Henderson and Horgan \(2009\)](#), “inferential processes, broadly understood, are simply those cognitive processes in which beliefs are formed or maintained on the basis of the information possessed” (p. 301). These inferential processes are usually considered equivalent to reasoning, as reasoning describes “the making or granting of assumptions called premises (starting points) and the process of moving toward conclusions” ([Walton, 1990](#), p. 403). When it comes to the relation between inference/reasoning and argumentation, [Walton \(1990\)](#) claims that reasoning occurs *within* discourse or argument, whereas other scholars attest that reasoning is discourse or argument (see [Walton, 1990](#), for the distinction). In either case, inferential reasoning is a requirement for argumentative discourse, and at the same time appears during it. A view of argumentation as a dynamic and dialectical type of reasoning occurring in certain contexts of dialog is thus justified under this approach.

Appraisal abilities

Appraisal, or evaluation, abilities can be regarded as part of generative abilities, but can also be considered independently, when assessing arguments produced by others. From an argumentation perspective, the following appraisal activities are part of the reasoning processes necessary for the practice of argumentation to take place:

Evaluation of information

A complementary process to argumentation is that of inquiry. Inquiry abilities are about “recognizing that there is something to find out—that new evidence can be distinguished from and potentially bears on existing understanding—and identifying what this something is within the context of any specific inquiry activity” ([Kuhn and Pease, 2008](#), p. 513). Evaluating available information and its potential use as evidence to support or rebut one’s arguments is part of the ability to evaluate information in relation to the context of the activity in which such information is used. [Dewey \(1910\)](#) was the first to claim that thinking is uncritical when it lacks reflective inquiry. For him, reflective, critical, and good thinking are synonyms, although he preferred using the term reflective thinking to refer to “active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends” (p. 6). This implies that a certain degree of judgment suspension is necessary, including the “search for new materials to corroborate or to refute the first suggestions that occur” (p. 13). However, such judgment cannot be ongoing; otherwise it cannot direct inquiry and examination (p. 108). Argumentation provides a context for inquiry to be directed toward the goal of constructing adequate, coherent, and consistent explanations based on information at hand ([Sampson et al., 2011](#)).

Evaluation of arguments

All the “good reasoning” criteria described in **Argumentation as an alternative construct to critical thinking** section are dialectical in nature, meaning that they all have a dialogical component that defines their satisfaction or not, even when arguments are constructed individually (i.e., not socially). Furthermore, in order for individual arguments to be reasonable (i.e., fallacy-free), the arguer must consider any objections a reasonable critic of these arguments might raise ([Blair and Johnson, 1987](#)). This is possible through the argumentation skill of antilogos, which is the ability to critically evaluate whether specific information can support different claims ([Billig, 1987](#); [Glassner and Schwarz, 2007](#)), and subsequently coordinate theory (claims) with evidence (information) ([Kuhn, 1999](#)). This skill is manifested in the production of counterarguments to one’s own arguments, and of responding to those through oral rebuttals or through written integrated (two-sided) arguments.

Evaluation of context

Evaluating when and whether to engage in critical thinking is another ability involving argumentation, evidenced by several argument-based instructional programs around issues that are worthy of investing mental energy in order to choose the best among several explanations. As [Halpern \(1998](#), p. 452) puts it, “an extended session of generating alternatives and calculating probabilities is a reasonable response to a diagnosis of cancer; it is not worth the effort when the decision involves the selection of an ice-cream flavor”. But what makes an issue one worthy of argumentative inquiry and contemplation? The following are some of the characteristics of an issue meriting argument: (a) the issue is related to a problem or phenomenon under consideration; (b) it is questionable or open to discussion; and (c) it is addressable through use of information as supporting evidence ([Rapanta, 2019b](#)). Recently, [Kuhn \(2018a\)](#) provided a list of authentic real-world issues that serve well as a context for argumentation. These issues relate to topics of gradually expanding scope, from personal to community to national and global; their common dimension is that they are addressable in terms of two or more contrasting alternatives with evidence pertinent to both sides.

Metastrategic skills

A series of empirical studies show that argument skills are multifaceted, involving both strategic skills, such as the ability to construct counterarguments and use evidence, but also metastrategic ones. Studies examining the development of argument skills, using the microgenetic method, revealed that a metastrategic understanding of the norms of argumentation was developing and supported development of argument skill ([Kuhn et al., 2008, 2013](#); [Iordanou and Constantinou, 2015](#); [Shi, 2020](#)). Participants who engaged in extensive practice in dialogic argumentation, in addition to the gains at the strategic level—use of counterargument strategies

aiming to weaken other's position—showed significant gains in meta-level communication *about* the discourse (Kuhn et al., 2008). The meta-level communication became more reciprocal over time and took a directive character, showing an attempt to direct the argumentation process (Kuhn et al., 2013). Those findings suggest that meta-level awareness and understanding of argumentation is developing through engagement in dialogic argumentation and supports the development of argumentation strategies, at the performance level. The co-existence of gains at the strategic and metastrategic level is consistent with other findings in cognitive development literature suggesting a bootstrapping relation between strategic and metastrategic levels (Kuhn and Pearsall, 1998), that is, between applying argumentation strategies at the performance level and the meta-level awareness of the process of argumentation.

Epistemic and epistemological awareness

The meta-level component of argumentation involves besides metastrategic knowledge, epistemic awareness. Epistemic awareness involves appreciation that counterargument and evidence use are critical components of argumentation (Kuhn et al., 2013). The growing attention to others' position, evident in individuals' inquiry skills—looking for information about other's position—and the use of counterarguments weakening others' position, while engaging in dialogic argumentation (Iordanou and Kuhn, 2020) is evidence of an implicit developing epistemic awareness that other's position might have some merit and is worthy of our attention. The increasing use of counterarguments and evidence may also reveal an epistemic appreciation of the function of counterargument and evidence as a means for persuading others, but also reveal another prominent epistemic achievement, namely, an appreciation of the subjective nature of human knowledge and the acknowledgment that there might be alternative interpretations of a particular phenomenon which are worthy of our attention.

Epistemic awareness also involves one's epistemic standards, that is, the standards used to evaluate knowledge claims (Chinn et al., 2020). Epistemic standards change developmentally, moving from egocentric to more objective ones. For example, three-year-olds show preference for egocentric standards, such as familiarity with the informant, while older children employ more objective standards, showing preference toward informers who have been reliable in their past performance (Corriveau and Harris, 2009). Underlying developmental changes in epistemic awareness are also changes in epistemic understanding regarding evidence and its coordination with theory, such as understanding that evidence is different from a claim and that different information leads to different beliefs (see Iordanou, 2016a).

Epistemic standards involve issues such as judging the credibility of the source of a piece of information and trustworthiness—Who can I trust to get evidence? Involving further questions such as “Who is an expert?” but also “Which media source can I trust?” and are closely related with one's science media literacy (Allchin, 2021).

Reasoning/arguing disposition

Engagement in argumentation, involving both construction and critical evaluation of arguments, depends not only on whether individuals have the knowledge and skills to do so—including both strategic and metastrategic knowledge—but also on whether individuals have the disposition to engage in this task. The disposition to engage in argumentation is affected by individuals' values and epistemic thinking (Kuhn, 2001). Epistemic thinking, besides involving one's epistemic commitments regarding which sources are trustworthy and what kinds of evidence are reliable, involve one's appreciation of which standards and methodologies are most appropriate for deciding what is true (Chinn et al., 2020). Valuing argumentation as a reliable and appropriate methodology for achieving truth and as the most powerful means of influencing others is a prerequisite for engaging in fruitful argumentation.

According to the developmental model of epistemic cognition (Kuhn et al., 2000), epistemic understanding develops, progressing from the absolutist level—where knowledge is conceived as an objective, external entity, which is knowable with certainty—to the multiplist level—where knowledge is conceived as a subjective construct of the human mind—to the evaluativist level—where a balance is achieved between the objective and subjective components of knowledge and knowledge is viewed as the product of evaluation based on argument and evidence (Kuhn et al., 2000). Empirical findings showed that there is a relation between mature epistemic understanding—the evaluativist level—and engagement in both argument evaluation (Iordanou et al., 2019) and argument construction (Iordanou et al., 2020; Mason and Scirica, 2006).

In addition to epistemic beliefs, egocentric biases and affect influence reasoning disposition and the quality of reasoning. Individuals tend to show my-side bias in reasoning when dealing with an issue that they hold strong beliefs about, that is, a favoritism toward egocentric arguments, which are in line with one's own views or beliefs (Baron, 1995). Depending on the source and/or whether they read something that is in agreement with their beliefs or not, individuals with strong beliefs on an issue show avoidance or engage in different kind of reasoning (Iordanou et al., 2020). Finally, high emotional investment affects one's reasoning disposition and leads individuals to biases in reasoning, such as rating opposing arguments as weaker, known as the disconfirmation bias (Edwards and Smith, 1996).

Finally, argumentativeness can be defined as the disposition “to attack the other arguer's case, reasons, position, and evidence” (Rapanta and Hample, 2015, p. 266). Argumentativeness is a motive for someone engaging or not in an argumentative dialog, with either negative (anti-social) or constructive consequences (Infante, 1988). The idea behind argumentativeness as a positive disposition is that “when individuals engage in argumentativeness, they attack the positions that others take or hold ... When individuals engage in verbal aggression, they attack the self-concept of the other” (Rancer, 1998, p. 152). Seen as a contrary disposition to being verbally aggressive, being argumentative can be related to critical thinking as it requires a degree of self-evaluation and self-reflection.

Evidence from argument-based pedagogical methods designed to foster students' critical thinking skills and dispositions

According to Kuhn (2019), rather than an individual ability or skill, "critical thinking is a dialogic practice people engage in and commit to, initially interactively and then in interiorized form with the other only implicit" (p. 146). This idea is rooted in socio-cultural theories of learning (Mead, 1934; Vygotsky, 1987a), according to which thinking is a social practice that is facilitated and developed through interaction with others. It is only after engaging with others intermentally, that practice and its gains are also transferred to the intramental level (Wertsch, 1985). This view of thinking in general, and of critical thinking in particular (see Larrain et al., 2019), underlies the vast quantity of empirical research on dialogical, argument-based pedagogical methods reporting gains subsequently transferred to individual tasks. This section is dedicated to review of some of these works by identifying some major ways in which skills and dispositions exercised or "put in service" in dialogical argumentation may transfer to individual practices manifesting critical thinking.

Transfer from social to individual: the case of writing

From a sociocultural learning theory perspective, it is expected that part of the knowledge negotiation process taking place socially, through dialog, is then transferred to individual writing products (Vygotsky, 1978). For example, Chen et al. (2016) showed that students' writing in science improved over time following engagement in peer-to-peer negotiations and whole-class presentations followed by peers' critique. They maintain that the most important reason for this improvement is the "advanced awareness of audience" (p. 313), which then become part of students' reasoning in their development of written arguments. Similarly, Sampson et al. (2013) showed how a writing-intensive instructional approach may enable adolescents to advance their writing skills through engaging in group writing processes as part of science laboratory courses. In a non-science context, results are similar. For example, Reznitskaya et al. (2001) showed how primary school students' engagement in small-group discussions on controversial issues about texts they had read significantly improved their subsequent persuasive essays.

Kuhn and colleagues in a series of studies (e.g., Hemberger et al., 2017; Kuhn et al., 2016a,b; Kuhn and Moore, 2015; see also Iordanou and Rapanta, 2021, for a review) have shown that dialogical argumentation is a path to argumentative writing and its development. They showed that young adolescents engaged in an extended dialog-based program of peer argumentation, "Argue with Me," not only improve their discourse but also their argumentative writing. For example, Kuhn and Crowell (2011) found the intervention group wrote more dual-perspective arguments than a comparison group at the end of Year 2, and more integrated-perspective and more arguments in general in Year 3.

Transfer to other topics/context

Due to the prevailing methods of assessing CT as a set of individual responses to a set of closed-ended content-specific questions, the issue of whether CT skills and dispositions are transferable to different domains remains unresolved (Ennis, 1989; Halpern, 1998). Arguments in favor of CT domain-dependence consist of the following: that thinking always needs something to be applied to (Paul, 1985); that different fields have different types and challenges of reasoning (Toulmin et al., 1979); and that skills applied to resolve one task or problem may not be necessary for a different task or problem (McPeck, 1981). Arguments in favor of CT domain-independence mainly consist of the following: that high order cognitive skills are not content- but context-bound (Perkins and Salomon, 1989), where context can bring them to the surface (Resnick, 1987); that the question is whether knowledge of the field is a necessary or sufficient requirement for CT to emerge (Ennis, 1989); and that dispositions, which are content-free, may influence the ways knowledge and skills are applied (Ennis, 1996).

Empirical studies of argumentation have addressed the issue of transferability of CT skills to other contexts. Iordanou (2013) showed that the gains in the electronic dialog mode, in particular increased levels of counterarguments and rebuttals, successfully transferred to the face-to-face dialog mode. Iordanou (2010) also showed that transfer of argumentative reasoning gains occurred in both directions from arguing in the science to social domain and vice versa. Several other studies (e.g., Hemberger et al., 2017; Iordanou and Constantinou, 2014; Kuhn et al., 2016a; Iordanou and Kuhn, 2020; Papathomas and Kuhn, 2017; Reznitskaya et al., 2012; Shi, 2019) confirm that transfer of reasoning gains to a new topic beyond the one students engage with in discourse is possible. Finally, recent studies also showed that transfer from "learning to argue" to "arguing to learn" contexts (Muller-Mirza and Perret-Clermont, 2009; Von Aufschnaiter et al., 2008) is possible. In particular, students who engaged in intensive dialogical argumentation practice on a topic either showed reasoning gains on another, curriculum-related topic (Rapanta, 2021), or even content knowledge learning gains (Iordanou et al., 2019b). In summary, because meaningful learning requires deep engagement with ideas, and the practice of argumentation provides the context for this (Jonassen and Kim, 2010), gains related to both the "deep engagement" process as well as "ideas/knowledge" learning are possible, confirming the two-fold metaphor of learning as participation and as acquisition (Sfard, 1998).

From skills to dispositions

Another type of transfer is that from manifested skills to dispositions. Immersion in the practice of dialogical argumentation may affect not only how much and how well students argue but also their stance regarding the value, purpose and norms of argument.

This can be considered a type of transfer from CT skills to CT dispositions. Studies, for example, reporting on improvements in argument evaluation tasks after students engage in dialogical argumentation, or after their recognition of argumentation norms, confirm this relationship (e.g., [Ryu and Sandoval, 2012](#); [Weinstock et al., 2004](#)). Learning the epistemological norms of dialogical argumentation through applying them in practice and transferring such learning to a dialog construction and evaluation task is another type of evidence of transfer from skills to disposition (e.g., [Kuhn et al., 2013](#)). Transfer is also confirmed when the tasks manifesting improvement or change in epistemological dispositions are related to concepts of knowledge and reasoning in general. Again, engagement in dialogical argumentation practice was shown to produce such transfer. For example, [Iordanou \(2016b\)](#) showed, in a pre/post-intervention assessment, that students' engagement in intensive dialogic practice had a positive effect on their understanding or reasoning and knowledge in general.

Transfer from skills to dispositions can also be implicit (i.e., without manifestation in a transfer task) as several argumentation skills "contain" certain dispositions. For example, [Kuhn \(1991, 1999\)](#) showed that advanced use of argument skills (generation of valid arguments, counterarguments, and rebuttals) implies the presence of advanced epistemologies pertaining to an evaluativist in contrast to an absolutist or a multiplist (see also [Reasoning/arguing disposition](#) section). What is more, the acquisition of such advanced epistemologies is itself developed through people's immersion in dialogical argumentation ([Iordanou and Kuhn, 2020](#)). According to [Anderson et al. \(2001\)](#), "the ability and disposition to take more than one perspective arises from participating in discussions with others who hold different perspectives" (p. 2). This multiperspectivism requires that at least two different or opposite perspectives are acceptable at the same time, and what makes one better than the other is the fact that it better fulfills the criteria of relevance and sufficiency discussed in [Argumentation as an alternative construct to critical thinking](#) section. Fulfillment of such criteria increases the plausibility of a perspective against an alternative ([Walton, 1989](#)). Considering that the degree of plausibility may change according to new information and its use as relevant support for one position or another is the basis of the evaluativist epistemology, developed through participation in persuasive argumentation, also referred to as inquiry dialog ([Reznitskaya and Gregory, 2013](#)).

Critical thinking as an endpoint

Argumentation, as discussed above, is a socio-discursive practice that functions at several levels (i.e., cognitive, metacognitive, epistemological) and those functions imply the application of CT abilities, skills, and dispositions. Therefore, argumentation provides a "visible practice" tool for identifying and assessing students' CT. At the same time, much empirical research shows that systematic engagement in dialogical argumentation helps develop reasoning skills and dispositions of a CT nature. Therefore, argumentation is also a vehicle for fostering and developing CT. However, for argumentation to be of a high quality, a degree of criticality is necessary. Such criticality is reflected in the values, reflexivity, and media and information literacy implicit in good argumentation.

Values

Values, both personal and social, can be used to support one's claims during argumentation. Values are considered to provide legitimate support to claims when they express "generalized conceptions of what are desirable ends or ways of behaving" ([Rieke et al., 1997](#), p. 10). Not all personal values are legitimate, as they can go against the common-sense "good"; therefore, we can talk about negative social values, such as social hatred. When such values are used as support for one's claims, a functional argument analysis does not always reveal the value-based unreasonableness of the arguments made. An analytical system combining functional analysis with content analysis (e.g., based on social representations) is necessary to this end (see [Rapanta and Trovão, 2020](#), for an example).

From an educational practice point of view, for argumentation to be critical, authenticity of the issues discussed as well as possible engagement in action as a result of the decisions taken are necessary criteria. Authentic issues are situated in real or realistic contexts ([Evagorou et al., 2011](#)), so that participants relate to them, and show willingness to engage at all levels of interaction (i.e., cognitive, metacognitive, epistemological). The enactment of the decisions taken as part of an interpersonal argumentation activity can help foster an ethical and activist stance ([Moore, 2013](#)) related to informed actions emerging from the discussions.

Reflexivity

On one hand, engagement in dialogical argumentation is necessary, as we showed above, for the social aspects of CT to be manifest and taken forward. On the other hand, individuals' engagement in argumentation is not enough for critical decision making and problem solving to take place. Finding a balance between the individual thinking process implied in CT and its social influence is key ([Rieke et al., 1997](#), p. 18). This balance can be achieved through the process and practice of reflexivity, which "occurs whenever a person interacting within a context elaborates the effects of their actions in an inter-subjective dynamic (...) gives new meaning to the processes (...) and recognizes the active role they assumed in guiding events" ([Freda and Esposito, 2017](#), p. 5). During argumentation, this can be achieved through the combination of the social reasoning activities with self-reflexive or peer-to-peer reflective activities as the one introduced as part of the "Argue with Me" method ([Kuhn et al., 2016b](#)): As pairs of peers argue against each

Let's think... Starting with our argument

One of our MAIN ARGUMENTS was:

Their COUNTERARGUMENT against our argument was:

Our COMEBACK was:



Fig. 2 Extract from the reflection tools used as part of the “Argue with Me” method. Courtesy of Prof. Deanna Kuhn.

other on a debatable topic of their interest, they also reflect on how their arguments are getting stronger or weaker during interaction by completing the diagram in Fig. 2. In general, graphical and diagrammatic representations have been shown to be effective in enhancing the self-reflexivity required for critical thinking and argumentation (Harrell and Wetzel, 2015).

Media and information literacy

Another challenge that adds to being critical in arguing relates to a competence complementary to argumentation, namely inquiry. According to Kuhn (2019), “inquiry and argument can be regarded as twin pillars of skilled thinking, replacing the more generic critical thinking label, if we regard inquiry as an input process of seeking and analyzing new information and argument as an output process of drawing and communicating justified conclusions” (p. 147). Inquiry skills are related to media and information literacy skills (Chu et al., 2011). Research needs to be done on how media, information, and argument literacy skills can be fostered together as a whole, with different but related tasks aiming at inquiry and argument.

Conclusion

The ideas presented in this article further substantiate the argument already formulated by Piaget (1928), and later articulated by Vygotsky (1987b), that reflective/critical thinking develops only in argument and discussion. It is through justification, counterargument and response (Leitão, 2000) that argumentation supports not only the construction of content knowledge (i.e., conceptual learning) but also and mainly the acquisition and development of argument knowledge, what is known as “learning to argue” (Muller-Mirza and Perret-Clermont, 2009). Argument knowledge entails a range of abilities, skills, and dispositions inherent to argumentative reasoning and discourse that almost fully cover the range of abilities, skills, and dispositions expected from a critical thinker. This justifies the extant use of argument-based instructional programs around the world explicitly aiming at the development of critical thinking skills through engagement in argumentation, and in particular, dialogic argumentation. The skill of *antilogos* (Billig, 1987), a necessary aspect of arguing, is also the basis of thinking critically. As argumentation theorist Walton (1989) puts it: “the common core of basic critical thinking skills underlying critical reasoning (...) is the key ability to look at both sides of an argument. The structure behind this ability is the concept of argument as dialogue” (p. 182).

The inherent relation of argumentation and critical thinking must become more explicit in educational research, so that the use of argumentative practices as a vehicle for critical thinking development can be further validated. One way to do so is through making explicit use of theoretical concepts and ideas from informal logic and argumentation theory and adapting them to pedagogical purposes. An example is Erduran et al. (2004) adaptation of Toulmin’s (1958) structural model of argument as a rubric for assessing classroom-based dialogical arguments. Another example is Kuhn’s (Kuhn et al., 2016a,b) inspiration from Walton (1989, 2014) to construct an argument-based curriculum focusing on increasing the plausibility of one’s own arguments while at the same time decreasing that of his/her components, first orally, as part of a social practice, and then individually, as part of a written essay. Another way is through making the connection between inquiry and argumentation more explicit in the various

instructional programs aiming at critical thinking development. Although argument-based inquiry and inquiry-based argumentation are already in use as instructional paradigms (see [Rapanta and Felton, 2019](#), for a review), especially in science, their broader use in contexts combining media, information, and argument literacy skills is still scant.

To conclude, argumentation and critical thinking (CT) are two sides of the same educational ideal: preparing reasonable and reflective democratic citizens, able to take decisions trusting their own analytical, appraisal, and argument construction skills. What is more, against the failure of CT assessment tests to provide the necessary link between what they define as good CT performance and what the necessary skills are to do so ([Kuhn, 2018b](#)), argumentation provides a way of both appraising and fostering CT abilities, skills, and dispositions as described in **Reasoning abilities, skills and dispositions implied in argumentative discourse** section. Together with inquiry skills, necessary for information search and selection, argumentation gets us closer to “empirically identifiable skills or behaviors than does the term critical thinking, while capturing much of what critical thinking is envisioned to encompass” ([Kuhn, 2019](#), p. 147). Although much work remains to be done in regard to assessing and fostering argument, adaptable to different contexts and domains (see [Iordanou and Rapanta, 2021](#)), argumentation as an instantiation of CT and as a pedagogical method and objective is a promising pathway to preparing 21st century students and citizens.

References

- Allchin, D., 2021. Who speaks for science? *Sci. & Educ.* <https://doi.org/10.1007/s11191-021-00257-4>.
- Anderson, R.C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., Tillmanns, M., Gilbert, L., 2001. The snowball phenomenon: spread of ways of talking and ways of thinking across groups of children. *Cognit. Instruct.* 19, 1–46.
- Angeles, P.A., 1981. *Dictionary of Philosophy*. Barnes & Noble, New York.
- Baillin, S., 1987. Critical and creative thinking. *Informal Log.* 9, 23–30.
- Baron, J., 1995. Myside bias in thinking about abortion. *Think. Reas.* 1, 221–235.
- Bennett, D., Power, A., Thomson, C., Mason, B., Bartleet, B.L., 2016. Reflection for learning, learning for reflection: developing indigenous competencies in higher education. *J. Univ. Teach. Learn. Pract.* 13, Art. 7.
- Billig, M., 1987. *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge University Press, Cambridge.
- Blair, J.A., Johnson, R.H., 1987. Argumentation as dialectical. *Argumentation* 1, 41–56.
- Blair, J.A., 2012. Relevance, acceptability and sufficiency today. In: Tindale, C. (Ed.), *Groundwork in the Theory of Argumentation*. Springer, Dordrecht, pp. 87–100.
- Burbules, N.C., Berk, R., 1999. Critical thinking and critical pedagogy: relations, differences, and limits. In: Popkewitz, T.S., Fendler, L. (Eds.), *Critical Theories in Education: Changing Terrains of Knowledge and Politics*. Routledge, New York, pp. 45–65.
- Chen, Y.C., Hand, B., Park, S., 2016. Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Sci. Educ.* 25, 277–320.
- Chin, C., Osborne, J., 2010. Supporting argumentation through students' questions: case studies in science classrooms. *J. Learn. Sci.* 19, 230–284.
- Chinn, C.A., Barzilai, S., Duncan, R.G., 2020. Disagreeing about how to know: the instructional value of explorations into knowing. *Educ. Psychol.* 55, 167–180.
- Chu, S.K.W., Tse, S.K., Chow, K., 2011. Using collaborative teaching and inquiry project-based learning to help primary school students develop information literacy and information skills. *Libr. Inf. Sci. Res.* 33, 132–143.
- Corriveau, K.H., Harris, P.L., 2009. Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Dev. Sci.* 12, 188–193.
- Dewey, J., 1910. *How We Think*. D.C. Heath, Boston.
- Dewey, J., 1966. *Democracy and Education: An Introduction to the Philosophy of Education*. The Free Press, New York.
- Duschl, R.A., Schweingruber, H.A., Shouse, A.W., 2008. *Taking Science to School: Learning and Teaching Science in Grades K-8*. National Academy Press, Washington, DC. Retrieved from: <http://www.nap.edu>.
- Edwards, K., Smith, E.E., 1996. A disconfirmation bias in the evaluation of arguments. *J. Pers. Soc. Psychol.* 71, 5–24.
- Ennis, R.H., 1987. Critical thinking and the curriculum. In: Heiman, M., Slomianko, J. (Eds.), *Thinking Skills Instruction: Concepts and Techniques*. National Education Association, Washington, D.C., pp. 40–48.
- Ennis, R.H., 1989. Critical thinking and subject specificity: clarification and needed research. *Educ. Res.* 18, 4–10.
- Ennis, R.H., 1993. Critical thinking assessment. *Theor. Pract.* 32, 179–186.
- Ennis, R.H., 1996. Critical thinking dispositions: their nature and assessability. *Informal Log.* 18, 165–182.
- Ennis, R.H., 2001. Argument appraisal strategy: a comprehensive approach. *Informal Log.* 21, 97–140.
- Ennis, R., 2011. Critical thinking: reflection and perspective. Part I. *Inq. Crit. Think. Across Discip.* 26, 4–18.
- Erduran, S., Simon, S., Osborne, J., 2004. Tapping into argumentation: developments in the application of Toulmin's Argument Pattern for studying science discourse. *Sci. Educ.* 88, 915–933.
- Evagorou, M., Sadler, T.D., Tal, T.R., 2011. Metatalk: assessment, audience, and authenticity for teaching SSI and argumentation. In: Sadler, T.D. (Ed.), *Socio-scientific Issues in the Classroom*. Springer, Dordrecht, pp. 161–166.
- Faccione, P., 1990. Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction. In: *The Delphi Report: Research Findings and Recommendations Prepared for the American Philosophical Association*. ERIC Doc, pp. ED315–423.
- Freda, M.F., Esposito, G., 2017. Promoting reflection and reflexivity through narrative devices. *Qual. Res. J.* 17, 2–19.
- Freire, P., 1993. In: Ramos, M.B. (Ed.), *Pedagogy of the Oppressed*. Continuum, New York.
- Glaser, R., 1984. Education and thinking: the role of knowledge. *Am. Psychol.* 39, 93–104.
- Glassner, A., Schwarz, B.B., 2007. What stands and develops between creative and critical thinking? *Argumentation? Think. Skills Creativ.* 2, 10–18.
- Govier, T., 2018. *Problems in Argument Analysis and Evaluation*, vol. 6. University of Windsor, Windsor.
- Groarke, L.A., Tindale, C.W., 2004. *Good Reasoning Matters! A Constructive Approach to Critical Thinking*, third ed. Oxford University Press, Don Mills.
- Halpern, D.F., 1998. Teaching critical thinking for transfer across domains: disposition, skills, structure training, and metacognitive monitoring. *Am. Psychol.* 53, 449–455.
- Halpern, D.F., 2014. *Thought and Knowledge. An Introduction to Critical Thinking*, fifth ed. Psychology Press, New York and London.
- Harrell, M., Wetzel, D., 2015. Using argument diagramming to teach critical thinking in a first-year writing course. In: Barnett, R., Davies, M. (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education*. Palgrave Macmillan, New York, pp. 213–232.
- Hemberger, L., Kuhn, D., Matos, F., Shi, Y., 2017. A dialogic path to evidence-based argumentative writing. *J. Learn. Sci.* 26, 575–607.
- Henderson, D., Horgan, T.E., 2009. Epistemic virtues and cognitive dispositions. In: Damschen, G., Schnepf, R., Stüber, K. (Eds.), *Debating Dispositions: Issues in Metaphysics, Epistemology and Philosophy of Mind*. Walter de Gruyter, Berlin, pp. 296–319.
- Higgins, S., Baumfield, V., 1998. A defence of teaching general thinking skills. *J. Philos. Educ.* 32, 391–398.

- Hitchcock, D., 2020. Critical thinking. In: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*, fall 2020 ed. Available at. <https://plato.stanford.edu/archives/fall2020/entries/critical-thinking/>
- Infante, D.A., 1988. *Arguing Constructively*. Waveland Press, Prospect Heights.
- Iordanou, K., Constantinou, C.P., 2014. Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. *Learn. Instr.* 34, 42–57.
- Iordanou, K., Constantinou, C.P., 2015. Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Sci. Educ.* 99, 282–311.
- Iordanou, K., Kuhn, D., 2020. Contemplating the opposition: does a personal touch matter? *Discourse Process.* 57, 343–359.
- Iordanou, K., Rapanta, C., 2021. "Argue with me": a method for developing argument skills. *Front. Psychol.* 12. <https://doi.org/10.3389/fpsyg.2021.631203>.
- Iordanou, K., Kuhn, D., Matos, F., Shi, Y., Hemberger, L., 2019a. Learning by arguing. *Learn. Instruct.* 63, 101–207.
- Iordanou, K., Muis, K.R., Kendeou, P., 2019b. Epistemic perspective and online epistemic processing of evidence: developmental and domain differences. *J. Exp. Educ.* 87, 531–551.
- Iordanou, K., Kendeou, P., Michalinos, Z., 2020. Examining my-side bias during and after reading controversial historical accounts. *Metacogn. Learn.* 15 (3).
- Iordanou, K., 2010. Developing argument skills across scientific and social domains. *J. Cognit. Dev.* 11, 293–327.
- Iordanou, K., 2013. Developing face-to-face argumentation skills: does arguing on the computer help? *J. Cognit. Dev.* 14, 292–320.
- Iordanou, K., 2016a. From theory of mind to epistemic cognition. A lifespan perspective. *Frontline Learn. Res.* 4, 106–119.
- Iordanou, K., 2016b. Developing epistemological understanding in scientific and social domains through argumentation. *Z. Padagog. Psychol.* 30, 109–120.
- Jonassen, D.H., Kim, B., 2010. Arguing to learn and learning to argue: design justifications and guidelines. *Educ. Technol. Res. Dev.* 58, 439–457.
- Kuhn, D., Crowell, A., 2011. Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychol. Sci.* 22, 545–552.
- Kuhn, D., Moore, W., 2015. Argument as core curriculum. *Learn. Res. Pract.* 1, 66–78.
- Kuhn, D., Pearsall, S., 1998. Relations between metastrategic knowledge and strategic performance. *Cognit. Dev.* 13, 227–247.
- Kuhn, D., Pease, M., 2008. What needs to develop in the development of inquiry skills? *Cognit. Instruct.* 26, 512–559.
- Kuhn, D., Cheney, R., Weinstock, M., 2000. The development of epistemological understanding. *Cognit. Dev.* 15, 309–328.
- Kuhn, D., Goh, W., Iordanou, K., Shaenfield, D., 2008. Arguing on the computer: a microgenetic study of developing argument skills in a computer-supported environment. *Child Dev.* 79, 1311–1329.
- Kuhn, D., Zillmer, N., Crowell, A., Zavala, J., 2013. Developing norms of argumentation: metacognitive, epistemological, and social dimensions of developing argumentative competence. *Cognit. Instruct.* 31, 456–496.
- Kuhn, D., Hemberger, L., Khait, V., 2016a. Tracing the development of argumentative writing in a discourse-rich context. *Writ. Commun.* 33, 92–121.
- Kuhn, D., Hemberger, L., Khait, V., 2016b. *Argue With Me: Developing Thinking and Writing Through Dialog*. Routledge, New York.
- Kuhn, D., 1991. *The Skills of Argument*. Cambridge University Press, New York.
- Kuhn, D., 1992. Thinking as argument. *Harv. Educ. Rev.* 62, 155–179.
- Kuhn, D., 1993. Science as argument: implications for teaching and learning scientific thinking. *Sci. Educ.* 77, 319–337.
- Kuhn, D., 1999. A developmental model of critical thinking. *Educ. Res.* 28, 16–46.
- Kuhn, D., 2001. How do people know? *Psychol. Sci.* 12, 1–8.
- Kuhn, D., 2018a. *Building Our Best Future: Thinking Critically about Ourselves and Our World*. Wessex, New York.
- Kuhn, D., 2018b. A role for reasoning in a dialogic approach to critical thinking. *Topoi* 37, 121–128.
- Kuhn, D., 2019. Critical thinking as discourse. *Hum. Dev.* 62, 146–164.
- Larrain, A., Freire, P., López, P., Grau, V., 2019. Counter-arguing during curriculum-supported peer interaction facilitates middle-school students' science content knowledge. *Cognit. Instruct.* 37, 453–482.
- Leitão, S., 2000. The potential of argument in knowledge building. *Hum. Dev.* 43, 332–360.
- Lipman, M., 1987. Critical thinking: what can it be? *Anal. Teach.* 8, 5–12.
- Macagno, F., 2008. Dialectical relevance and dialogical context in Walton's pragmatic theory. *Informal Logic* 28, 102–128.
- Macagno, F., 2015. A means-end classification of argumentation schemes. In: van Eemeren, F., Garssen, B. (Eds.), *Reflections on Theoretical Issues in Argumentation Theory*. Springer, Cham, pp. 183–201.
- Macagno, F., 2016. Argument relevance and structure. Assessing and developing students' uses of evidence. *Int. J. Educ. Res.* 79, 180–194.
- Macagno, F., 2019. Coding relevance. *Learn. Cult. Soc. Interact.* 1–15.
- Mason, L., Scirica, F., 2006. Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learn. Instruct.* 16, 492–509.
- Mayweg-Paus, E., Thiebach, M., Jucks, R., 2016. Let me critically question this!—insights from a training study on the role of questioning on argumentative discourse. *Int. J. Educ. Res.* 79, 195–210.
- McPeck, J., 1981. *Critical Thinking and Education*. St. Martin's, New York.
- Mead, G.H., 1934. *Mind, Self, and Society*. University of Chicago Press, Chicago.
- Mezirow, J., 1997. Transformative learning: theory to practice. *N. Dir. Adult Cont. Educ.* 74, 5–12.
- Moore, T., 2013. Critical thinking: seven definitions in search of a concept. *Stud. High Educ.* 38, 506–522.
- Muller-Mirza, N., Perret-Clermont, A.N., 2009. *Argumentation and Education: Theoretical Foundations and Practices*. Springer, Dordrecht.
- Nussbaum, E.M., Edwards, O.V., 2011. Critical questions and argument stratagems: a framework for enhancing and analyzing students' reasoning practices. *J. Learn. Sci.* 20, 443–488.
- Papathomas, L., Kuhn, D., 2017. Learning to argue via apprenticeship. *J. Exp. Child Psychol.* 159, 129–139.
- Paul, R., Elder, L., 2006. Critical thinking: the nature of critical and creative thought. *J. Dev. Educ.* 30, 34–35.
- Paul, R., Elder, L., 2013. Critical thinking: intellectual standards essential to reasoning well within every domain of human thought. Part two. *J. Dev. Educ.* 37, 32–36.
- Paul, R., 1981. Teaching critical thinking in the "strong" sense: a focus on self-deception, world views, and a dialectical mode of analysis. *Informal Log.* 4, 1–7.
- Paul, R.W., 1985. Bloom's taxonomy and critical thinking instruction. *Educ. Leader* 42, 36–39.
- Paul, R.W., 1987. Dialogical thinking: critical thought essential to the acquisition of rational knowledge and passions. In: Baron, J.B., Sternberg, R.J. (Eds.), *Series of Books in Psychology. Teaching Thinking Skills: Theory and Practice*. Freeman, New York, pp. 127–148.
- Paul, R.W., 1990. *Critical Thinking*. Sonoma State University, Sonoma.
- Perkins, D.N., Salomon, G., 1989. Are cognitive skills context-bound? *Educ. Res.* 18, 16–25.
- Perkins, D.N., 1985. Reasoning as imagination. *Interchange* 16, 14–26.
- Piaget, J., 1928. *Judgment and Reasoning in the Child*. Harcourt, Brace.
- Rancer, A.S., 1998. Argumentativeness. In: McCroskey, J.C., Daly, J.A., Martin, M.M., Beatty, M.J. (Eds.), *Communication and Personality: Trait Perspectives*. Hampton Press, Cresskill, pp. 149–170.
- Rapanta, C., Felton, M., 2019. Mixed methods research in inquiry-based instruction: an integrative review. *Int. J. Res. Method Educ.* 42, 288–304.
- Rapanta, C., Hampe, D., 2015. Orientations to interpersonal arguing in the United Arab Emirates, with comparisons to the United States, China, and India. *J. Intercult. Commun. Res.* 44, 263–287.
- Rapanta, C., Macagno, F., 2019. Pragmatics, education and argumentation: introduction to the special issue. *Learn. Cult. Soc. Interact.* <https://doi.org/10.1016/j.lcsi.2019.100371>.

- Rapanta, C., Trovão, S., 2020. Shall we receive more refugees or not? A comparative analysis and assessment of Portuguese adolescents' arguments, views, and concerns. *Pedagog. Cult. Soc.* 28, 581–603.
- Rapanta, C., Walton, D., 2016. Identifying paralogs in two ethnically different contexts at university level. *Infanc. Aprendiz.* 39, 119–149.
- Rapanta, C., 2019a. Argumentation as critically oriented pedagogical dialogue. *Informal Log.* 39, 1–31.
- Rapanta, C., 2019b. *Argumentation Strategies in the Classroom*. Vernon Press, Wilmington.
- Rapanta, C., 2021. Can teachers implement a student-centered dialogical argumentation method across the curriculum? *Teach. Teach. Educ.* 105.
- Resnick, L.B., 1987. *Education and Learning to Think*. National Academy Press, Washington, D.C.
- Reznitskaya, A., Gregory, M., 2013. Student thought and classroom language: examining the mechanisms of change in dialogic teaching. *Educ. Psychol.* 48, 114–133.
- Reznitskaya, A., Anderson, R.C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., Kim, S.Y., 2001. Influence of oral discussion on written argument. *Discourse Process.* 32, 155–175.
- Reznitskaya, A., Glina, M., Carolan, B., Michaud, O., Rogers, J., Sequeira, L., 2012. Examining transfer effects from dialogic discussions to new tasks and contexts. *Contemp. Educ. Psychol.* 37, 288–306.
- Rieke, R.D., Sillars, M.O., Peterson, T.R., 1997. *Argumentation and Critical Decision Making*. Longman, New York.
- Ryu, S., Sandoval, W.A., 2012. Improvements to elementary children's epistemic understanding from sustained argumentation. *Sci. Educ.* 96, 488–526.
- Sampson, V., Grooms, J., Walker, J.P., 2011. Argument-driven inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: an exploratory study. *Sci. Educ.* 95, 217–257.
- Sampson, V., Enderle, P., Grooms, J., Witte, S., 2013. Writing to learn by learning to write during the school science laboratory: helping middle and high school students develop argumentative writing skills as they learn core ideas. *Sci. Educ.* 97, 643–670.
- Scriven, M., 1976. *Reasoning*. McGraw-Hill, New York.
- Stard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27, 4–13.
- Shi, Y., 2019. Enhancing evidence-based argumentation in a mainland China middle school. *Contemp. Educ. Psychol.* 59. <https://doi.org/10.1016/j.cedpsych.2019.101809> [Online first].
- Shi, Y., 2020. Talk about evidence during argumentation. *Discourse Process.* 57, 770–792.
- Siegel, H., 1980. Critical thinking as an educational ideal. *Educ. Forum* 45, 7–23.
- Siegel, H., 1985. Educating reason: critical thinking, informal logic, and the philosophy of education. *Informal Log.* 7, 69–81.
- Song, Y., Ferretti, R.P., 2013. Teaching critical questions about argumentation through the revising process: effects of strategy instruction on college students' argumentative essays. *Read. Writ.* 26, 67–90.
- Sternberg, R.J., 1986. *Critical Thinking: Its Nature, Measurement, and Improvement*. National Institute of Education, Washington, DC.
- Toulmin, S., Reike, R.D., Janik, A., 1979. *An Introduction to Reasoning*. McMillan, New York.
- Toulmin, S., 1958. *The Uses of Argument*. Cambridge University Press, New York.
- Von Aufschnaiter, C., Erduran, S., Osborne, J., Simon, S., 2008. Arguing to learn and learning to argue: case studies of how students' argumentation relates to their scientific knowledge. *J. Res. Sci. Teach.* 45, 101–131.
- Vygotsky, L.S., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge.
- Vygotsky, L.S., 1987a. Lectures on psychology. In: Minick, N., Rieber, R.W., Carton, A.S. (Eds.), *The Collected Works of L.S. Vygotsky: Problems of General Psychology*, vol. 1. Plenum Press, New York, pp. 289–358.
- Vygotsky, L.S., 1987b. Thinking and speech. In: Minick, N., Rieber, R.W., Carton, A.S. (Eds.), *The Collected Works of L.S. Vygotsky: Problems of General Psychology*, vol. 1. Plenum Press, New York, pp. 39–285.
- Walton, D., Reed, C., Macagno, F., 2008. *Argumentation Schemes*. Cambridge University Press, New York.
- Walton, D.N., 1989. Dialogue theory for critical thinking. *Argumentation* 3, 169–184.
- Walton, D.N., 1990. What is reasoning? What is an argument? *J. Philos.* 87, 399–419.
- Walton, D.N., 1996. *Argumentation Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah.
- Walton, D.N., 1998. *The New Dialectic: Conversational Contexts of Argument*. University of Toronto Press, Toronto.
- Walton, D.N., 2010. Why fallacies appear to be better arguments than they are. *Informal Log.* 30, 159–184.
- Walton, D.N., 2014. *Dialogue Theory for Critical Argumentation*. John Benjamins, Amsterdam.
- Watson, G., Glaser, E.M., 1964. *Watson-Glaser Critical Thinking Appraisal Manual: Forms YM and ZM*. Harcourt, Brace, and World, New York.
- Weinstock, M., Neuman, Y., Tabak, I., 2004. Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemp. Educ. Psychol.* 29, 77–94.
- Wertsch, J.V., 1985. *Vygotsky and the Social Formation of Mind*. Harvard University Press, Cambridge.
- Woods, J., Irvine, A., Walton, D.N., 2004. *Argument: Critical Thinking, Logic, and the Fallacies*. Pearson/Prentice Hall, New York.
- Zarefsky, D., 1981. "Reasonableness" in public policy argument: fields as institutions. In: Ziegelmüller, G., Rhodes, J. (Eds.), *Dimensions of Argument: Proceedings of the Second Summer Conference on Argumentation*. Speech Communication Association, Annandale, pp. 88–100.

Computer supported collaborative learning

Sanna Järvelä^a, Päivi Häkkinen^b, and Piia Näykki^c, ^a University of Oulu, Department of Educational Sciences and Teacher Education, Oulu, Finland; ^b University of Jyväskylä, Finnish Institute for Educational Research, Jyväskylä, Finland; and ^c University of Jyväskylä, Faculty of Education and Psychology, Jyväskylä, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	588
Productive CSCL	588
Shared meaning making in social interactions	589
Designing and supporting CSCL	590
Future perspectives	590
References	591

Introduction

Research on *computer supported collaborative learning (CSCL)* covers how to design the technological contexts for productive collaboration and how people learn in the context of collaborative activity. This is to optimize high level collaborative activities for sharing and building knowledge. The sharing and construction of knowledge among the participants in groups supported by technology is at the center of communication and can become a shared resource of group members (Stahl and Hakkarainen, 2021). CSCL can be designed to facilitate high-level cognitive achievements in educational settings in both online and classroom learning environments. It can occur synchronously or asynchronously, and can be realized in dyads, small study groups or work teams, but also in broader social networks and communities. Currently, CSCL has been implemented as part of various teaching and learning models at different educational levels and in work-life teams. It has increasingly been acknowledged for empowering learners as agentic participants in communities of learners to become authors of their future (Rosé and Järvelä, 2020). *Collaborative learning (CL)* is a powerful way of enhancing individual learning and can also be effective in developing group working skills and practices. It has origins both in sociocognitive theory, in terms of the social construction of knowledge, and in sociocultural theory, extending the view from learners' thinking and interactions to community- and cultural-level knowledge practices. Knowledge construction in CSCL is built by students' interactions, by means of the negotiation of meaning and the co-creation of knowledge (Resnick, 1991). The social construction of knowledge is commonly made via technology-mediated collaborative efforts through dialogs and facilitated by differences in persons' perspectives. By definition, collaborative learning is a coordinated effort to solve a problem or perform a task together by constructing and maintaining a shared conception of that problem (Roschelle and Teasley, 1995).

Productive CSCL

Productive CSCL can be explained from *sociocognitive* and *sociocultural*, as well as *linguistic perspectives*. Foundational sociocognitive research on small group interaction (e.g., Mugny and Doise, 1978) has focused on individual cognition and learning and has shown, for example, that confrontations in terms of socio-cognitive conflict are beneficial for learning when the conflict confrontation leads to conflict regulation in an epistemic manner. Sociocultural theory, having roots in Vygotsky (1978) thinking, emphasizes learning taking place first intersubjectively and then intrasubjectively and as a process of constructing a shared meaning (Stahl, 2004). Learning is viewed as a process of enculturation and participation in a community of practice. In addition to group interactions and social practices, tool-mediated activities (e.g., through or around computers) also play a crucial role in cognitive development and learning. From the linguistic perspective, building and maintaining a common ground means that individuals negotiate about common goals and construct shared understanding, knowledge, beliefs, assumptions, and presuppositions (Clark and Schaefer, 1989). These are prerequisites for the coordination of collaborative activities and knowledge sharing. However, while not merely sharing a common ground triggers learning, the core is the cognitive effort put into building shared understanding (Schwartz, 1995).

Productive collaboration can be enabled and enriched by computers and digital technologies. Group participants learn when they engage in collaborative activities, which trigger particular kinds of learning mechanisms. Collaborative learning situations supported by computers can provide a natural setting for externalizing thinking by explaining, questioning, arguing and sharing regulation. Collaborative technologies offer tools for exploring and experimenting with alternative solutions to problems as well as for representing and modeling knowledge (Fischer and Mandl, 2005). Therefore, CSCL provides environments that mediate collaborative interaction and learning in diverse ways. The tools through which interactions between learners, groups and artifacts take place mediate both interactions and materially embodied interactions (Davidsen and Ryberg, 2017). This is to say, interactions in CSCL environments happen between minds, but are also mediated by a variety of resources such as technology, language, gestures, prior knowledge or previous interactions (Stahl et al., 2006).

Shared meaning making in social interactions

The process of grounding refers to coordination of both content and process in communication and has been used to explain short conversational events in CSCL (Clark and Schaefer, 1989). Later research has focused more on shared meaning making (Stahl, 2004). Meaning making refers to “the increased cognitive-interactional effort involved in the transition from learning to understand each other to learning to understand the meanings of the semiotic tools that constitute the mediators of interpersonal interaction” (Baker et al., 1999, p. 31). Collaborative learning takes place through processes of shared meaning making, and there is a dynamic relationship between shared meanings and individual interpretations. During this process, learners negotiate their individual views to reach a shared understanding, which is used as a resource for constructing further understanding (Stahl et al., 2006).

The positive effects of social interaction for individual learning have been demonstrated in many of the earlier studies focusing on causal links between the conditions and effects of collaboration. Later research has started to focus on particular processes, mechanisms and contexts that either support or constrain the co-construction of knowledge (Dillenbourg, 1999). The theoretical ideas behind knowledge co-construction follow Roschelle's (1992) notion of convergence; group members *construct shared meanings* by constantly monitoring the degree to which they understand each other's thinking, extend other people's ideas, acknowledge divergent interpretations and resolve inconsistencies between the ideas that have been proposed. In sum, prior research has shown that when learners externalize and explain their still-developing understanding to each other, they learn more effectively (Baker et al., 1999). This can be supported by shared workspaces with a variety of representation tools provided by computers that enable learners to externalize their ideas and to co-construct shared representations (Fischer and Mandl, 2005).

Decades of research has shown *communication and interaction patterns* that impact the quality of processes in sharing meaning (e.g., Baker, 2009). Key interaction patterns, such as task-focused participation, which impact the quality of information synthesis making among group members have been identified. Equity in participating interactions is the extent to which all members contribute to the discussion process. Inability to integrate important ideas from group members may lead to reduced potential for collaborative problem solving (Borge and White, 2016). By developing joint understanding, a group ensures ideas are understood among team members by requesting clarifications, rewording and rephrasing. Successful groups develop a joint understanding of ideas in order to establish common ground or mutual knowledge (Clark and Schaefer, 1989). Failure to develop a joint understanding hinders group task engagement. Building joint ideas is the extent to which a group elaborates interactions to ensure that contributions and ideas are not ignored or accepted without discussion. If this does not occur, loss of information may lead to poor decision-making outcomes (Barron, 2003).

Group level information processing covers several critical patterns in interactions that contribute to successful collaborative knowledge construction. Exploring alternative perspectives and explanations allows group members to present alternative ideas, opinions and claims. If the group agrees with the first possible opinion rather than exploring alternative solutions, this may lead to a lack of innovative thinking and to routine solutions. Studies show that the power of collaborative learning may fail if group members simply acknowledge or ignore other group members' contributions, rather than explore and extend them (e.g., Jeong and Chi, 2007). Then, a group simply transfers a set of existing individual ideas to other members rather than creating new knowledge that no group member had before (Webb et al., 1995).

Deep argumentation consists of high-quality claims, whereas lack of critical evaluation and analysis of information may lead to weak argumentation quality (Weinberger and Fischer, 2006). In general, collaborative argumentation allows students to consider and recognize divergent views, find discrepancies in understanding and negotiate meanings (Baker, 2009). For example, in CSCL, learners can argue collaboratively and co-constructively, deliberating and reasoning together in order to improve their understandings of the topic at hand, rather than each trying to “win” (Asterhan and Babichenko, 2015). Collaborative argumentation can also be regarded as a type of interaction process that may support group coordination. Computers can support the process of argumentation by making the interaction process more visible and tangible for the learners (Andriessen et al., 2003). Engaging in argumentative and constructive discussion consists of respecting social norms for other members' ideas and for equal participation (Blumenfeld et al., 1996), whereas disruptive discussion and unfavorable socioemotional interactions can lead to psychological unsafety that impacts group members' ability to learn from each other. These interaction patterns can impact how groups develop knowledge and the quality of the knowledge that is developed in CSCL. They affect the extent to which a group uses its members as cognitive, motivational and emotional resources; uses resources provided by technology; establishes productive common ground; and develops new knowledge or solutions that did not exist before collaboration.

Productive collaborative learning and working is a *multifaceted interaction process* that integrates cognitive, motivational and emotional components as the core of collaboration (Borge et al., 2018; Ludvigsen, 2016). In theory, collaborative learning requires group members to be aware of and to coordinate their cognitive, metacognitive, motivational and emotional resources and processes (Hadwin et al., 2018). In practice, starting points for successful CSCL are working in line with shared goals, sharing of knowledge and understanding, and joint monitoring of collaboration processes. Therefore, sharing has also been conceptualized as *metacognitive-level interactions*, and the *concept of regulation* in terms of self-regulation, co-regulation and socially shared regulation has been introduced to CSCL (Järvelä and Hadwin, 2013). Increasing evidence shows that it is not only an individual's but also a group's shared processes that matter, and for this reason, regulating collaborative learning is critical for success (Zheng et al., 2019).

Designing and supporting CSCL

Important questions that are being investigated in research on CSCL are: why do some groups perform better than others, how do learners engage in deep level collaborative interactions, why do they want to extend each other's understanding, and why do they sometimes fail in attempts to work as an effective group? As described earlier, research has explained these problems from different angles during the past decade, and efforts have been made to design and support productive CSCL. Effective collaborative learning does not occur automatically, but it can be designed. Technology affords learner opportunities to engage in a joint task, communicate, share resources, engage in productive collaborative learning processes, engage in co-construction, monitor and regulate collaborative learning, and find and build groups and communities (Jeong and Hmelo-Silver, 2016). All these can be used as guidelines for *designing forms and supports* in terms of (a) pedagogical designs and (b) technological tools for CSCL to overcome the interaction challenges and support critical CL processes. These can be enacted in advance to guide group activities—for example, during the groups' learning processes as a support for orienting to the productive collaboration or as a support for monitoring learning and interaction during the collaboration process.

Collaboration scripts are pedagogical designs of scaffolding that pre-structure socio-cognitive processes of CSCL. They define phases and tasks for students to perform, group composition, task distribution, roles of the group members, the mode and granularity of interaction, and timing and social levels of phases (Dillenbourg and Hong, 2008). Computer support in scripting can also mean that they are implemented in the interfaces of CSCL environments (e.g., prompts), aiming to facilitate cognitive, metacognitive and social activities of learners. CSCL scripts have been found to have some effect on domain learning but a significant effect on collaboration skills (Vogel et al., 2017). For example, prior studies have explored how scripting can be used as a support for learning through argumentation and how argumentation skills can be supported with scripting (Noroozi et al., 2012). Although there are studies showing that scripting CSCL enhances learning, some criticism has also been leveled at this practice: whether overscripting hinders interactions and learners' agency and has only a weak effect on transfer of skills, and thus subsequently undermines learners' motivation (Radkowsch et al., 2020). An alternative view is not to constrain activity during collaboration, but instead to emphasize the planning and monitoring of collaborative interactions before, during and after it. *Orchestration* refers to the management of teaching in the classroom context, particularly from the perspective of the coordination of activities such as individual work, small-group interaction and classroom discussions (Tchounikine, 2013). Teachers orchestrate, monitor and guide CSCL activities among peer learners. Other key components of orchestration cover materials or resources for collaboration and the technology environment with its variety of tools for working with knowledge and sharing it. The cornerstone of pedagogically designing CSCL environments is to find a delicate balance between free-form collaboration and the structuring, scaffolding, and orchestrating of collaborative activities.

Efforts have been made to implement technologies *supporting collaborative knowledge construction*. One of the first approaches was the development of knowledge building competences supported by the CSILE (Computer Supported Intentional learning Environment) and Knowledge Forum learning environments (Scardamalia and Bereiter, 1994). Cognitive group awareness tools, in turn, have been implemented to make group participants aware of each other's knowledge and opinions, for example, by subjective ratings or multiple choice knowledge tests (Sangin et al., 2011). The cognitive group awareness tools have shown positive effects on knowledge construction because comparisons of participating group members' knowledge, opinions, and understanding are available and trigger discussion and shared meaning making (Kirschner et al., 2015). Sociability tools have been developed for helping groups to increase sound social space, trust, and sense of community, and to overcome challenges in socioemotional interaction. For example, Kreijns et al. (2013) suggested that a sound social space is the key to allowing group members using CSCL to gain feelings of relatedness, group cohesiveness, trust, and respect for each other.

Because metacognition plays an important role in group learning, recent technologies have focused specifically on *supporting metacognition* in the context of CSCL (e.g., Bannert et al., 2015). The support is not only in individual metacognitive processes, but rather to raise groups' metacognitive awareness in terms of shared regulatory activities, such as planning, task performance, monitoring and reflection. Typically, the support has been provided in the form of scripts or prompts that are used to facilitate learners' awareness of their metacognitive processes at individual and group levels (Janssen and Bodemer, 2013). All share the common goal of designing supports that (1) increases learners' awareness of their own and others' learning process, (2) externalize students' and others' learning processes in a social plane and helps in sharing and interaction, and (3) prompts the acquisition and activation of regulatory processes (Järvelä et al., 2015).

Learning analytics innovations hold the promise of empowering teachers and learners by making certain aspects of learning interactions visible and by providing information that can prompt CSCL. Real time analytics tools and dashboards have been developed that can provide information about group participation and task progress, the current state of the CSCL script, inform teachers about student progress and support learners (Martinez-Maldonado, 2019). Learning analytics techniques could be a key to increasing learners' agency and supporting self- and socially shared regulation of learning.

Future perspectives

CSCL is well established in education and work-life and increasingly implemented in various interdisciplinary group contexts. Rapidly developing technologies offer opportunities for participation, new forms for interacting, and spaces for social construction of knowledge and social relations.

The traditional focus in CSCL has been on specialized applications designed to support collaboration, while more recent younger generation digital ecologies—for example, generic social media services and games (Martin et al., 2020)—are not designed with the critical processes of productive collaboration in mind but rather have a strong emphasis on social relations for their own sake. CSCL research could put more effort into investigating social networking and large-scale learning environments, not only collaborative learning environments and tools of own designs. Today's learning environments also increasingly integrate physical and virtual learning spaces, formal and informal learning contexts, and global and local settings. Understanding how learning occurs and how learning could be supported in these contexts provides new avenues for CSCL research. In the future, CSCL will be considered more and more within global contexts, contributing to educational change and ongoing learning opportunities.

While contexts and perspectives in CSCL have broadened, multidisciplinary efforts have emerged to understand collaborative learning processes, renew theoretical concepts, and find new and complementary methodological and analytical solutions for researching CSCL (Dindar et al., 2020). For example, physiological measures and eye tracking have helped capture core processes in collaboration, such as affective activation or joint visual attention (Törmänen et al., 2021). In addition, sociometric badges and social network analyses have opened up new ways to capture the traces of social interactions, levels of explanations for interactions and how collaborators are connected to each other through physical or conceptual closeness (Zhou et al., 2020).

Increasing methodological opportunities have produced rich datasets. CSCL generates digital learning traces, and in multidisciplinary collaboration, this type of data can be analyzed—for example, by experts in learning analytics and machine learning—to make predictions and build causal explanations of how certain processes that make collaborative learning productive occur in interactions. Even with innovative future methodological perspectives, the need for detailed qualitative analyses in CSCL research, such as video analyses, remains essential since they are able to explain the reactions, regularities and nuances of human collaborative learning and, thus, contribute to strong evidence-based CSCL designs.

References

- Andriessen, J., Baker, M.J., Suthers, D.D., 2003. Argumentation, computer support, and the educational context of confronting cognitions. In: Andriessen, J., Baker, M.J., Suthers, D. (Eds.), *Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning Environments*. Kluwer, Dordrecht, Netherlands, pp. 1–25.
- Asterhan, C.S.C., Babichenko, M., 2015. The social dimension of learning through argumentation: effects of human presence and discourse style. *J. Educ. Psychol.* 107, 740–755.
- Baker, M., Hansen, T., Joiner, R., Traum, D., 1999. The role of grounding in collaborative learning tasks. In: Dillenbourg, P. (Ed.), *Collaborative Learning: Cognitive and Computational Approaches*. Pergamon, Amsterdam, Netherlands, pp. 31–63.
- Baker, M., 2009. Argumentative interactions and the social construction of knowledge. In: Muller Mirza, N., Perret-Clermont, A.N. (Eds.), *Argumentation and Education*. Springer, Boston, MA, pp. 127–144.
- Bannert, M., Sonnenberg, C., Mengelkamp, C., Pieger, E., 2015. Short-and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Comput. Hum. Behav.* 52, 293–306.
- Barron, B., 2003. When smart groups fail. *J. Learn. Sci.* 12, 307–359.
- Blumenfeld, P., Marx, R., Soloway, E., Krajcik, J., 1996. Learning with peers: from small group co-operation to collaborative communities. *Educ. Res.* 25, 37–40.
- Borge, M., White, B.Y., 2016. Toward the development of socio-metacognitive expertise: an approach to 1014 developing collaborative competence. *Cognit. Instr.* 34, 323–360.
- Borge, M., Ong, Y.S., Rosé, C.P., 2018. Learning to monitor and regulate collective thinking processes. *Int. J. Comput. Support. Collab. Learn.* 13, 61–92.
- Clark, H.H., Schaefer, E.F., 1989. Contributing to discourse. *Cognit. Sci.* 13, 259–294.
- Davidson, J., Ryberg, T., 2017. "This is the size of one meter": children's bodily-material collaboration. *Int. J. Comput. Support. Collab. Learn.* 12, 65–90.
- Dillenbourg, P., Hong, F., 2008. The mechanics of macro scripts. *Int. J. Comput. Support. Collab. Learn.* 3, 5–23.
- Dillenbourg, P., 1999. What do you mean by collaborative learning? In: Dillenbourg, P. (Ed.), *Collaborative Learning: Cognitive and Computational Approaches*. Pergamon, Amsterdam, Netherlands, pp. 1–19.
- Dindar, M., Järvelä, S., Ahola, S., Huang, X., Zhao, G., 2020. Leaders and followers identified by emotional mimicry during collaborative learning: a facial emotions recognition study. *IEEE Trans. Affect. Comput.* 3045.
- Fischer, F., Mandl, H., 2005. Knowledge convergence in computer-supported collaborative learning: the role of external representation tools. *J. Learn. Sci.* 14, 405–441.
- Hadwin, A.F., Järvelä, S., Miller, M., 2018. Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In: Schunk, D.H., Greene, J.A. (Eds.), *Handbook of Self-Regulation of Learning and Performance*, second ed. Routledge, New York, NY, pp. 83–106.
- Janssen, J., Bodemer, D., 2013. Coordinated computer-supported collaborative learning: awareness and awareness tools. *Educ. Psychol.* 48, 40–55.
- Järvelä, S., Hadwin, A., 2013. New frontiers: regulating learning in CSCL. *Educ. Psychol.* 48, 25–39.
- Järvelä, S., Kirschner, P.A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., Koivuniemi, M., Järvenoja, H., 2015. Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educ. Technol. Res. Dev.* 63, 125–142.
- Jeong, H., Chi, M.T.H., 2007. Knowledge convergence and collaborative learning. *Instr. Sci.* 35, 287–315.
- Jeong, H., Hmelo-Silver, C.E., 2016. Seven affordances of computer-supported collaborative learning: how to support collaborative learning? How can technologies help? *Educ. Psychol.* 51, 247–265.
- Kirschner, P.A., Kreijns, K., Phielix, C., Fransen, J., 2015. Awareness of cognitive and social behaviour in a CSCL environment. *J. Comput. Assist. Learn.* 31, 59–77.
- Kreijns, K., Kirschner, P.A., Vermeulen, M., 2013. Social aspects of CSCL environments: a research framework. *Educ. Psychol.* 48, 229–242.
- Ludvigsen, S., 2016. CSCL: connecting the social, emotional and cognitive dimensions. *Int. J. Comput. Support. Collab. Learn.* 11, 115–121.
- Martin, F., Dennen, V., Bonk, C., 2020. A synthesis of systematic review research on emerging learning environments and technologies. *Educ. Technol. Res. Dev.* 68, 1613–1633.
- Martínez-Maldonado, R., 2019. A handheld classroom dashboard: teachers' perspectives on the use of real-time collaborative learning analytics. *Int. J. Comput. Support. Collab. Learn.* 14, 383–411.
- Mugny, G., Doise, W., 1978. Socio-cognitive conflict and structure of individual and collective performances. *Eur. J. Soc. Psychol.* 8, 181–192.
- Noroozi, O., Weinberger, A., Biemans, H.J.A., Mulder, M., Chizari, M., 2012. Argumentation-based computer supported collaborative learning (ABCSCCL): a synthesis of 15 years of research. *Educ. Res. Rev.* 7, 79–106.
- Radkowsky, A., Vogel, F., Fischer, F., 2020. Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *Int. J. Comput. Support. Collab. Learn.* 15, 5–47.
- Resnick, L.B., 1991. Shared cognition: thinking as social practice. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC, pp. 1–20.

- Roschelle, J., Teasley, S.D., 1995. The construction of shared knowledge in collaborative problem solving. In: O'Malley, C.E. (Ed.), *Computer-Supported Collaborative Learning*. Springer, Berlin, Germany, pp. 69–197.
- Roschelle, J., 1992. Learning by collaborating: convergent conceptual change. *J. Learn. Sci.* 2, 235–276.
- Rosé, C.P., Järvelä, S., 2020. Building community together: towards equitable CSCL practices and processes. *Int. J. Comput. Support. Collab. Learn.* 15, 249–255.
- Sangin, M., Molinari, G., Nüssli, M.-A., Dillenbourg, P., 2011. Facilitating peer knowledge modeling: effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Comput. Hum. Behav.* 27, 1059–1067.
- Scardamalia, M., Bereiter, C., 1994. Computer support for knowledge-building communities. *J. Learn. Sci.* 3, 265–283.
- Schwartz, D.L., 1995. The emergence of abstract representations in dyad problem solving. *J. Learn. Sci.* 4, 321–354.
- Stahl, G., Hakkarainen, K., 2021. Theories of CSCL. In: Cress, U., Rosé, C., Wise, A., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Springer, New York, NY, pp. 23–43.
- Stahl, G., Koschmann, T., Suthers, D., 2006. Computer-supported collaborative learning: an historical perspective. In: Sawyer, R.K. (Ed.), *Cambridge Handbook of the Learning Sciences*. Cambridge UP, Cambridge, UK, pp. 409–426.
- Stahl, G., 2004. Building collaborative knowing: elements of a social theory of CSCL. In: Strijbos, J.-W., Kirschner, P., Martens, R. (Eds.), *What We Know About CSCL: And Implementing It in Higher Education*. Kluwer, Boston, MA, pp. 53–86.
- Tchounikine, P., 2013. Clarifying design for orchestration: orchestration and orchestrable technology, scripting and conducting. *Comput. Educ.* 69, 500–503.
- Törmänen, T., Järvenoja, H., Mänty, K., 2021. All for one and one for all—how are students' affective states and group-level emotion regulation interconnected in collaborative learning? *Int. J. Educ. Res.* 109.
- Vogel, F., Wecker, C., Kollar, I., Fischer, F., 2017. Socio-cognitive scaffolding with computer-supported collaboration scripts: a meta-analysis. *Educ. Psychol. Rev.* 29, 477–511.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard UP, Cambridge, MA.
- Webb, N.M., Troper, J.D., Fall, R., 1995. Constructive activity and learning in collaborative small groups. *J. Educ. Psychol.* 87, 406–423.
- Weinberger, A., Fischer, F., 2006. A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Comput. Educ.* 46, 71–95.
- Zheng, J., Xing, W., Zhu, G., 2019. Examining sequential patterns of self- and socially shared regulation of STEM learning in a CSCL environment. *Comput. Educ.* 136, 34–48.
- Zhou, N., Kisselburgh, L., Chandrasegaran, S., Badam, K., Elmqvist, N., Ramani, K., 2020. Using social interaction trace data and context to predict collaboration quality and creative fluency in collaborative design learning environments. *Int. J. Hum. Comput. Stud.* 136.

Cooperative and collaborative learning

Noreen M. Webb and Eric Burnheimer, Department of Education, School of Education and Information Studies, University of California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Theoretical justification for the benefits for learning: the cognitive elaboration perspective	593
Example instructional programs	594
Process-outcome links: empirical evidence	594
Teacher interventions with collaborative groups to support productive interaction	595
When to intervene	595
How to intervene	595
Content versus process teacher interventions	595
Specific teacher discourse moves with small groups	596
Effects of teacher discourse in the whole class on small-group functioning	597
Conclusion	598
References	598

For decades, researchers, practitioners, and policy-makers have recognized the educational potential of students working with their peers in small groups to solve problems, complete tasks, and learn material. A large research base supports this enthusiasm. When compared with other instructional approaches such as whole-class or individualistic learning, peer-led group work has been shown to lead to higher student achievement, more positive attitudes toward learning and toward others, improved social skills, and increased perceptions of their own competence (e.g., [Slavin and Lake, 2008](#); [Kyndt et al., 2013](#)). The positive effects may be immediate, delayed, or both (e.g., [Howe et al., 2005](#)).

Peer-based learning approaches have been described in multiple ways, often in terms of cooperative and collaborative learning. Although the origins of cooperative and collaborative learning are different in many respects (for example, the former is often seen as more structured in terms of incentives and rewards, task design, role assignment, and group membership to foster positive interdependence among group members and individual accountability for one's own learning), the terms are often used interchangeably. In this article, we do not distinguish between them, but consider both under the umbrella of instructional approaches that provide students opportunities to learn from each other as they work together in face-to-face small groups.

The article "Peer learning in the classroom" in the third edition of the *International Encyclopedia of Education* ([Webb, 2010](#)) explored the mechanisms by which working with other students is thought to benefit student learning, and the many ways in which small-group work might be orchestrated to engender the learning benefits. This article focuses on research developments and directions that have appeared since then, particularly involving the links between student interaction during group work and learning outcomes, and the ways in which teacher practice can support productive group functioning.

Theoretical justification for the benefits for learning: the cognitive elaboration perspective

The cognitive elaboration perspective focuses on how overt communication processes during collaboration may trigger internal cognitive processes that are associated with learning. During group collaboration, students may explain their ideas to others in the context of conveying information about how to solve a problem or complete a task, helping others understand the material, justifying their suggestions in response to challenges, questions, or disagreements, or challenging others when perceiving conflicts or discrepancies. Students may engage with others' ideas in the context of constructing or elaborating solutions to problems or tasks, comparing perspectives, or addressing differences in opinions or problem-solving strategies.

Explaining ideas to others, whether in the context of making suggestions or responding to others' questions and challenges, encourages students to reflect on and monitor their own thinking. In the process of transforming what they know into communication that is relevant, coherent, and complete so that others can understand it, students may recognize their own misconceptions, or contradictions or incompleteness in their ideas ([Benware and Deci, 1984](#)). To resolve these issues, explainers may bring concepts or details together in ways they had not thought of previously, may see ways to make their arguments more compact, their ideas clearer or more elaborated, or their problem-solving strategies more efficient, or may seek out and use new information to build more complete or correct conceptions ([Bargh and Schul, 1980](#); [Coleman et al., 1997](#); [King, 1992](#); [Palincsar et al., 1993](#)).

Similarly, engaging with others' ideas can lead students to revisit or re-examine their own ideas, make connections between their own and others' ideas, generate new ideas that combine or extend those already proposed, and actively search for more information to resolve inconsistencies between students' perspectives ([Brown et al., 1992](#); [Roscoe and Chi, 2008](#); [Wittrock, 1990](#)). As a result, students may develop fuller conceptions of ideas, deeper understanding of material, and new perspectives.

Example instructional programs

Multiple instructional programs have been designed to promote productive communication in cooperative and collaborative groups. In many of them, teachers set ground rules for student interaction and provide instruction and practice in communication skills. To increase group reasoning (jointly analyzing problems, comparing possible explanations, making joint decisions), teachers in a set of studies conducted by Mercer and colleagues set ground rules for Exploratory Talk to help students engage critically and constructively with each other's ideas. Students are expected to share all relevant suggestions and information, provide reasons to justify assertions, opinions, and suggestions, ask others to give reasons for their suggestions, listen to others attentively, discussing alternatives before making decisions, and accept and respond to constructive challenges (Mercer, 1996, 2000; Mercer et al., 2004; Rojas-Drummond and Mercer, 2003).

Preparation in communications skills is also a central feature of the SPRinG (Social Pedagogic Research into Group work) program developed by Baines et al. (2008); see also Baines et al. (2007), Blatchford et al. (2006) to help teachers create inclusive and supportive classrooms. Teachers use a variety of activities in classrooms in order to develop students' skills in the following areas: taking turns speaking; engaging in active listening; asking and answering questions; making and asking for suggestions; expressing and requesting ideas and opinions; brainstorming suggestions, ideas and opinions; giving and asking for help; giving and asking for explanations; explaining and evaluating ideas; arguing and counter-arguing; using persuasive talk; and summarizing conversations. For example, students engage in debates and follow-up debriefing to help them understand what reasons are, the importance and benefits of giving reasons, and possible words to use when providing or requesting reasons.

In an approach to literature discussion called Collaborative Reasoning, students work in small groups to discuss a central question (or moral dilemma) raised by a story and argue for and against positions they take on the central question (Chinn et al., 2001). Students are taught how to support their positions with reasons, to challenge one another with counterarguments and evidence, and to weigh the pros and cons of their own and others' positions (Reznitskaya et al., 2009).

Training of students in communications skills is a prominent feature of many other studies as well. In Gillies and colleagues' research program (e.g., Gillies, 2003, 2004; Gillies et al., 2014), teachers train students to actively listen to each other, to provide constructive feedback for each other's suggestions and ideas, to answer each other's questions, to encourage all group members to contribute to the group task, to try to understand other group members' perspectives, to explore disagreements, and to monitor and evaluate the progress of the group. As another example, the teachers in Veenman et al.'s (2005) study trained students in explanation-giving and help-seeking skills (e.g., giving explanations rather than answers, asking clear and precise questions).

These instructional approaches have been shown to increase the depth of group discussion and, often, group performance and/or student achievement. Some examples of positive outcomes include a higher incidence of detailed explanations when students jointly construct new knowledge or conceptions and resolve conflicting points of view, improved argumentation skills, higher reasoning scores, and higher achievement in multiple subject matters, including social studies, mathematics, science, reading, and English than in programs without a strong focus on productive cooperative or collaborative group work.

Process-outcome links: empirical evidence

Considerable evidence of the links between student interaction during group work and learning outcomes (e.g., in mathematics, science, educational psychology) comes from research that collects observational data about students' collaboration. Some studies have found that composites of communication skills—such as giving help with explanations and challenging others' suggestions with explanations and justifications (Veenman et al., 2005) or explaining one's own ideas and engaging with others' ideas (Ing et al., 2015)—significantly predict achievement. Other research has shown that specific communication dimensions, such as explaining (or justifying) one's own ideas, extending others' ideas, supporting others' arguments, rebutting others' suggestions, co-constructing reasoning with others, resolving conflicts with others using justification, and referring to and building on others' ideas by clarifying, correcting, completing, or extending the ideas (transactive communication), also significantly correlate with learning outcomes (Chinn et al., 2000; Howe et al., 2007; Jurkowski and Hanze, 2015; Van Boxtel et al., 2000; Webb et al., 2014).

Additional evidence of the benefits of certain kinds of engagement with others' ideas comes from studies of argumentation. Certain types of argumentation have been shown to be more effective than others in promoting conceptual gains in science understanding. For example, Asterhan and Schwarz (2007, 2009) showed that dyads who argued for and against competing solutions or positions (including explanations of pros and cons of them, sometimes called dialectical argumentation) showed greater gains than dyads who discussed only one proposed solution or position (one-sided argumentation). An important feature of dialectical argumentation is the attempt by group members to persuade each other of the merits of their held positions.

The importance of collaborative groups considering multiple ideas finds support in the work of Howe and colleagues (e.g., Howe and Zachariou, 2019) showing that discussing contrasting ideas is associated with greater conceptual understanding of science concepts. In particular, they found that exchanging and negotiating different opinions about solutions to science problems and discussing divergent views about science concepts were both positively associated with growth from pretest to posttest. In contrast, the more frequently groups engaged in discussion that confirmed one particular solution or perspective, the less they progressed.

Another way of conceptualizing communication in collaborative groups is the rank-ordered classification of engagement activity proposed by Chi and Menekse (2015). Their ICAP (interactive-constructive-active-passive) framework distinguishes between

ordered categories of engagement: (1) interactive and constructive engagement (students build on each other's contributions), (2) constructive without interacting engagement (students build on their own ideas without considering or extending others' ideas), (3) active engagement (students restate or repeat what was said), and (4) passive engagement (students listen without saying much). Their empirical studies show level of engagement of the dialog pattern to be significantly correlated with learning gains (see also Chi and Wylie, 2014).

Finally, some research has shown in-the-moment advances in learning that occur while students work on a task together. Roscoe and Chi (2008), for example, described how students trying to explain to another student the structure and functioning of the human eye realized or admitted that they did not fully understand. These realizations prompted students to revisit and review material, and to rethink their ideas in order to provide more complete and correct explanations. Brown et al. (1992) illustrate how the group's challenge of an explainer's incomplete or incorrect ideas caused the explainer to re-examine her prior knowledge, to formulate and test predictions based on her incorrect mental model, and to revise her ideas based on the information that her peers provided in response to her initial predictions (in this case, her ideas about animals' use of camouflage as a defense mechanism). Webb et al. (2021) showed how students made advances in their mathematical thinking (e.g., moving from strategies that were incomplete or partially formed, ambiguous, or invalid to strategies that were fully correct, complete, and unambiguous, generating new mathematical ideas, making connections to new mathematical representations, generating new and substantially different strategies) through the process of explaining their own thinking and engaging with others' ideas. Students corrected or extended their mathematical thinking when issuing challenges or being challenged, while elaborating on or extending others' ideas, and when making connections between different ideas being presented.

Teacher interventions with collaborative groups to support productive interaction

When to intervene

Many cooperative learning researchers and theorists advise teachers to monitor small group progress and to intervene when groups fail to progress or appear to be functioning ineffectively, when no group member can answer the question, when students exhibit problems communicating with each other, when students dominate group work without allowing true dialog, when students are denied opportunities to participate or are not treated with respect, and when students and/or their ideas are marginalized (e.g., Cohen, 1994; Ding et al., 2007; Johnson and Johnson, 2008). Similarly, Tolmie et al. (2005) also suggested that teacher guidance may help counteract students' tendencies to control group work or to disagree in unproductive ways (e.g., failing to provide reasons for their opinions and ideas), as well as enable children to explore ideas more effectively.

Should teachers always intervene when groups are off task? A recent study suggests it may depend on when the off-task activity takes place and how long it lasts (Langer-Osuna et al., 2020). Brief off-task activity that takes place at the beginning of a group's collaboration may serve to help students ease into the task (for example, quick remarks about interesting features of the task may reorient group members toward each other and facilitate the start of the task). Off-task activity that occurs in the middle of group conversations may serve to help certain students gain access (or re-access) to the conversation (for example, a previously ignored student was able to secure entry into the group collaboration by engaging in brief, off-task use of materials that drew others' attention and interaction). Off-task activity that lasts longer than a minute or so, however, might signal that a group is avoiding the task, which may warrant teacher intervention.

Should teachers always intervene when some students appear to dominate the conversation at the expense of other students' participation? Esmonde (2009) showed that students who positioned themselves as experts (and were granted this authority by other group members) tended to dominate interactions and have the most opportunities to explain their thinking and obtain feedback from peers. Students positioned as novices, on the other hand, tended to be excluded and have fewer opportunities to express their ideas. However, Shah and Lewis (2019) found that participatory equity, in which opportunities to participate are fairly distributed among group members, is fluid, not static, such that levels of inequity fluctuate over the course of the collaboration. Even collaborations in which one student generally dominates the work may nonetheless exhibit equitable participation at times, and vice versa. Furthermore, different factors may drive inequity, including a push to finish the group's work as quickly as possible, group members' interpretation of their relational role as the person in charge, or group members' perceptions that others may lack relevant knowledge or competence. These results show that teachers need to be vigilant about monitoring group work, to consider both short-term and long-term participation patterns among group members, and to take into account the multiple reasons why inequitable participation might arise.

How to intervene

Content versus process teacher interventions

Some research has considered whether teachers should intervene about content or about process, and the results are not clear cut. Some evidence suggests that process interventions may be more productive than content interventions. For example, in a comparative study by Dekker and Elshout-Mohr (2004), students in classrooms in which teachers were instructed to provide only process help to groups (e.g., encouraging students to explain and justify their work) engaged in more extensive discussions and exhibited more equal participation among group members, and showed higher mathematics posttest scores, than students in classrooms in which teachers were instructed to provide only content help (e.g., giving hints about the mathematical content and strategies).

Other work suggests that content help may sometimes be effective, depending on how explicit or direct it is. Chiu (2004) found that more-explicit teacher help (e.g., explaining a concept, or giving a solution tactic) reduced groups' subsequent time-on-task and the depth of their discussions (whether students provided new ideas and explanations) compared to less-explicit teacher help (e.g., drawing attention to an aspect of the problem through asking questions instead of providing a solution strategy or telling groups how to proceed). Gillies (2004) found similar results: students in classrooms in which teachers provided direct instruction and explicit content help provided fewer detailed explanations than students in classrooms in which teachers provided less-direct content help (e.g., asking open and tentative questions to probe, clarify, and focus student thinking). In a further refinement, Meloth and Deering (1999) reported that teacher explicit content help (e.g., providing direct instruction about content) facilitated productive group discussion (in terms of students communicating their ideas and supporting their ideas with explanations and elaborations) when the teacher listened to the group's ideas (for example, to find out whether groups were focusing on irrelevant information) before providing specific instruction, but had no effect on group discussion otherwise. When the teacher's content help did not relate to a group's discussion, groups were able to disregard the teacher's suggestions without reducing the quality of their conversations.

More important than the focus of teacher interventions—subject matter content or group process—may be whether teachers try to ascertain student thinking and base their interaction on what they learn about students' thinking on the task (see Cohen's 1994 caution for teachers to carefully listen to group discussions, and to make hypotheses about the groups' difficulties before deciding on what questions to ask or suggestions to make). Interestingly, Chiu (2004) observed that attending to students' thinking may influence the nature of teacher interventions with groups. Specifically, teachers who considered and evaluated students' suggestions tended to provide less-explicit help and issued few directives; teachers who did not evaluate groups' ideas tended to give more-explicit help and issue many directives. Considering and evaluating strategies that groups were using made it possible for teachers to work with students' ideas and to help in less direct ways, such as asking students to formulate a plan for answering part of a problem that the group had not yet addressed, or asking students to reread a part of the problem that they had misinterpreted.

Specific teacher discourse moves with small groups

Teachers can use specific discourse moves when intervening with small groups to improve group functioning and students' learning outcomes. For example, in Collaborative Reasoning (described above; Chinn et al., 2001), to spur students to generate different reasons for their positions and participate in counterarguments, and to weigh the pros and cons of their own and others' positions when discussing moral dilemmas, the teacher asks students to clarify their ideas, prompts them to give reasons or evidence for their positions, and encourages them to respond to each other's claims. These teacher moves support rich student discussions which in turn increase students' argumentation skills (as measured by posttest reflective essays on new topics; Reznitskaya et al., 2009).

Additional teacher moves shown to be productive include asking students probing and clarifying questions ("Can you tell me a little more about what you're intending to do here?"), acknowledging and validating students' ideas ("I can see you've worked really hard to find out how these items are related. I wonder what you could do now to identify a key category they can all fit into?"), identifying discrepancies in students' work and clarifying options they may take ("I wonder how you can include ... when you've already mentioned ... ?"), and offering suggestions in a tentative fashion ("I wonder if you've considered doing it this way?"; Gillies, 2004, p. 260). Such moves encourage students to acknowledge each other's ideas, expand on others' suggestions, challenge others' ideas, and probe each other's opinions, which, in turn, lead to achievement gains (improved scores on posttests of reasoning and problem solving; e.g., Gillies, 2003, 2006; Gillies and Boyle, 2008).

Recent research distinguishing between initial teacher invitations for students to explain their thinking and engage with others' ideas and teacher follow up on those invitations shows that teachers' initial invitations may not be sufficient to produce in-depth student conversations and that follow up may be critical. Franke and colleagues (Franke et al., 2015; Webb et al., 2019) found that teachers' initial moves to invite student explaining such as pressing students to explain, clarify, and justify their problem-solving strategies, and initial moves to invite student engagement with others' ideas such as asking students to compare their strategies to those generated by others, explain problem-solving strategies that other students use, and ask questions of each other were not always effective. Students were more likely to produce complete explanations of their ideas when teachers continued to probe student thinking to prompt students to give further details about their problem-solving strategies beyond their initial explanation, and when teachers persisted in asking questions in order to push students to clarify the ambiguous aspects of their explanations. Similarly, students were more likely to go beyond superficial and perfunctory engagement with each other's ideas when teachers followed up with additional questions that drew attention to specific details in students' mathematical ideas and encouraged students to engage with those particular details ("Do you understand why she [divided each whole] into five pieces?").

The teacher discourse moves such as those described above can have delayed, as well as immediate, effects on collaborative group functioning. Jadallah et al. (2011) showed how teacher moves can propagate through students' discussions. In their study of collaborative reasoning about moral dilemmas and policy issues, when the teacher asked students to clarify their suggestions, not only did students generally do so, but there was a fairly high probability that other students would subsequently engage with the clarified idea of the first student. More specifically, when a student gave clarification in response to a teacher prompt to do so, there was a 41.9% probability that a second student would comment on the first student's ideas, and a 49.1% probability that a third student would also comment on the idea. After the teacher prompted a student to clarify an idea, long chains of conversational turns among students (e.g., 27 turns) were devoted to engaging with the initial clarified idea (including repeating details in the idea, elaborating on the idea, and challenging the idea). Other teacher moves found to have delayed positive effects on students' elaborations of their thinking (particularly the extent to which students explicitly conveyed connections between ideas) included

praising students' use of specific cognitive strategies (e.g., using examples in their explanations), and regulating students' turn-taking (e.g., by asking students who were interrupted by other students to continue talking; [Lin et al., 2015](#)).

The delayed effects of teacher moves can persist even when the teacher is no longer present with the group. For example, [Webb et al. \(2009\)](#) found that teacher probing of students' explanations to uncover details or further thinking about problem-solving strategies often corresponded with students giving more details about their strategies during their discussions (and often providing complete explanations) after the teacher intervened and had departed from the group than they had offered before the teacher intervened.

One reason for the sustained effects of teacher moves such as prompting for clarification of ideas may be that the teacher's move lends credence to the idea and bestows intellectual authority on the student offering the idea. A teacher's attention to a particular student's idea may position the student as a credible source of information, with the result that the student's ideas gain special traction, attention, and positive evaluation during the group's discussion ([Langer-Osuna, 2016](#)).

Effects of teacher discourse in the whole class on small-group functioning

Teachers' discourse during whole-class instruction can also impact how students interact with each other in collaborative groups. A consistent theme is that teachers pressing students to explain their thinking during whole-class discussions is important for the depth of students' discussion in collaborative groups. In [Kazemi and Stipek's \(2001\)](#) "high press" classrooms, mathematics teachers pressed students to explain and to justify their problem-solving strategies during whole-class discussions. For example, a high-press teacher may have pressed students to further explain each step in their problem-solving procedures, to give mathematical arguments for their choices of strategies, to explain the significance of specific details in their strategies, and to demonstrate their problem-solving procedure in different ways (e.g., verbal description, shading diagrams, adding fractions). Some of the mathematics teachers in [Webb et al.'s \(2008, 2009\)](#) studies carried out similar moves, such as pushing students to make explicit the steps in their mental processes both when students' answers and strategies were correct and when they were incorrect. "High press" teachers make analogous moves in other subject matter domains as well. For example, teachers of literature interpretation in [Smagorinsky and Fly's \(1993\)](#) study invited students to participate in interpreting text and prompted them to elaborate on their responses (e.g., by directly requesting elaboration, repeating portions of students' suggestions in the form of questions).

In contrast, teachers in "low press" classrooms tend not to ask students to explain their thinking during whole-class discussions. They might ask students to describe the steps they used to solve problems but not ask students to link their strategies to mathematical reasons or to explain why they chose particular procedures, and may accept students' correct answers and procedures without follow-up ([Kazemi and Stipek, 2001](#)). Or they might generally accept the first explanations students offer even if they are incomplete or incorrect ([Webb et al., 2008, 2009](#)), request only answers, or ask students to provide just the results of calculations that the teacher set up, without asking students to explain their reasoning ([Webb et al., 2006](#)). Or, in the case of literature interpretation, teachers may assume responsibility for interpreting the text without asking students to contribute ([Smagorinsky and Fly, 1993](#)).

The differences between teachers' discourse during their whole-class instruction in high-press and low-press classrooms have been shown to correspond to differences in student-led peer interaction. In collaborative group work in "high press" classrooms, students use mathematical arguments to explain why and how their solutions work, provide explanations that are correct and complete, strive to arrive at mutual understandings of solutions, and, in the case of literature interpretation, elaborate on their ideas for how to generalize from the text, and frequently ask each other to explain their reasoning. In contrast, in collaborative group work in "low press" classrooms, students summarize, but do not explain, their steps for solving problems, do not debate the mathematics involved in the problems, and rarely ask about their peers' thinking or strategies, and, in the case of literature interpretation, give brief summaries of the text without elaborating on their suggestions.

Additional insights into specific moves teachers make during whole-class discussions to encourage students to engage with others' ideas that may transfer to student behavior during collaborative work come from a recent case study by [Webb et al. \(2018, pp. 300–301\)](#). First, the teacher enacted a variety of ways of engaging with students' ideas, including asking students specific questions about details in their problem-solving strategies, restating or rephrasing students' suggestions, noting similarities and differences between different students' ideas, challenging students' suggestions, adding on to students' ideas, and drawing connections between students' ideas. Second, the teacher supported students to engage with others' ideas by explicitly calling for students to engage with others' ideas in the ways she had modeled, such as making connections between different students' ideas ("Can you see any connections between the red strategy and the blue strategy?"), and asking them to consider particular details in a student's strategy ("Are you sure about this part [of the student's strategy]? Can you talk about how it's going to work?"). In many cases, she gave them specific guidance in what to ask or what to say, for example, when revocing another student's idea ("Can you talk about what was the first step that these partners took together? [Say] 'First, these partners ...'"') or when challenging another student ("If you disagree [say] 'I'm disagreeing because ...'"). In their collaborative groups, students carried out all of the ways of engaging with each other's ideas that the teacher had enacted herself or had supported students to carry out during whole-class instruction, including using some of the same language the teacher had used or suggested (e.g., "Why do you think that?", "Where did you get the 5 from?" "So you're saying we could do another 4 times 12.").

Conclusion

Over the past several decades, research has deepened our understanding of the kinds of communication in cooperative and collaborative groups that are productive for student learning, and has heightened our awareness into how teachers can intervene with groups to support productive group functioning. A predominant theme from recent research results is that teachers pressing students to clarify and elaborate on their own and others' thinking can have positive influences on group conversations that are both immediate and delayed. Moreover, analogous teacher moves with students during whole-class discussions can carry over to interaction among students in collaborative groups.

An intriguing possibility is that productive in-the-moment teacher discourse with cooperative and collaborative groups may complement other ways in which teachers can prepare students to work collaboratively. As described in the article "Peer learning in the classroom" in the third edition of this encyclopedia, to support group functioning teachers can, for example, provide students with instruction in communications skills, structure group interaction by assigning cognitive roles or by giving students high-level questions to ask each other, and pose open-ended tasks that invite multiple perspectives or require multiple areas of expertise. Future research will reveal how teachers may be able to use intervention strategies in tandem with these other supports to produce fruitful interaction in cooperative and collaborative groups and, hence, positive student learning outcomes.

References

- Asterhan, C.S.C., Schwarz, B.B., 2007. The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *J. Educ. Psychol.* 99 (3), 626–639.
- Asterhan, C.S., Schwarz, B.B., 2009. Argumentation and explanation in conceptual change: indications from protocol analyses of peer-to-peer dialog. *Cognit. Sci.* 33, 374–400.
- Baines, E., Blatchford, P., Chowne, A., 2007. Improving the effectiveness of collaborative group work in primary schools: effects on science attainment. *Br. Educ. Res. J.* 33, 663–680.
- Baines, E., Blatchford, P., Kutnick, P., Chowne, A., Ota, C., Berdondini, L., 2008. *Promoting Effective Group Work in the Classroom*. Routledge, London.
- Bargh, J.A., Schul, Y., 1980. On the cognitive benefit of teaching. *J. Educ. Psychol.* 72, 593–604.
- Benware, C.A., Deci, E.L., 1984. Quality of learning with an active versus passive motivational set. *Am. Educ. Res. J.* 21, 755–765.
- Blatchford, P., Baines, E., Rubie-Davies, C., Bassett, P., Chowne, A., 2006. The effect of a new approach to group work on pupil–pupil and teacher–pupil interactions. *J. Educ. Psychol.* 98, 750–765.
- Brown, A.L., Campione, J.C., Webber, L.S., McGilly, K., 1992. Interactive learning environments: a new look at assessment and instruction. In: Gifford, B., O'Connor, M.C. (Eds.), *Changing Assessments: Alternative Views of Aptitude, Achievement, and Instruction*. Kluwer Academic Publishers, Boston, MA, pp. 121–211.
- Chi, M.T.H., Menekse, M., 2015. Dialogue patterns that promote learning. In: Resnick, L.B., Asterhan, C., Clarke, S.N. (Eds.), *Socializing Intelligence Through Academic Talk and Dialogue*. American Educational Research Association, Washington, DC, pp. 263–274.
- Chi, M.T.H., Wylie, R., 2014. The ICAP Framework: linking cognitive engagement to active learning outcomes. *Educ. Psychol.* 49, 219–243.
- Chinn, C.A., O'Donnell, A.M., Jinks, T.S., 2000. The structure of discourse in collaborative learning. *J. Exp. Educ.* 69, 77–97.
- Chinn, C.A., Anderson, R.C., Waggoner, M.A., 2001. Patterns of discourse in two kinds of literature discussion. *Read. Res. Q.* 36, 378–411.
- Chiu, M.M., 2004. Adapting teacher interventions to student needs during cooperative learning: how to improve student problem solving and time on-task. *Am. Educ. Res. J.* 41, 365–399.
- Cohen, E.G., 1994. *Restructuring the Classroom: Conditions for Productive Small Groups*.
- Coleman, E.B., Brown, A.L., Rivkin, I.D., 1997. The effect of instructional explanations on learning from scientific texts. *J. Learn. Sci.* 6 (4), 347–365.
- Dekker, R., Elshout-Mohr, M., 2004. Teacher interventions aimed at mathematical level raising during collaborative learning. *Educ. Stud. Math.* 56, 39–65.
- Ding, M., Li, X., Piccolo, D., Kulm, G., 2007. Teacher interventions in cooperative-learning mathematics classes. *J. Educ. Res.* 100, 162–175.
- Esmonde, I., 2009. Mathematics learning in groups: analyzing equity in two cooperative activity structures. *J. Learn. Sci.* 18, 247–284.
- Franke, M.L., Turrou, A.C., Webb, N.M., Ing, M., Wong, J., Shin, N., Fernandez, C.H., 2015. Student engagement with others' mathematical ideas: the role of teacher invitation and support moves. *Elem. Sch. J.* 116, 126–148.
- Gillies, R.M., Boyle, M., 2008. Teachers' discourse during cooperative learning and their perceptions of this pedagogical practice. *Teach. Teach. Educ.* 24, 1333–1348.
- Gillies, R.M., Nichols, K., Burgh, G., Haynes, M., 2014. Primary students' scientific reasoning and discourse during cooperative inquiry-based science activities. *Int. J. Educ. Res.* 63, 127–140.
- Gillies, R.M., 2003. Structuring cooperative group work in classrooms. *Int. J. Educ. Res.* 39 (1–2), 35–49.
- Gillies, R.M., 2004. The effects of communication training on teachers' and students' verbal behaviors during cooperative learning. *Int. J. Educ. Res.* 41, 257–279.
- Gillies, R.M., 2006. Ten Australian teachers' discourse and reported pedagogical practices during cooperative learning. *Elem. Sch. J.* 106, 429–451.
- Howe, C., Zachariou, A., 2019. Small-group collaboration and individual knowledge acquisition: the processes of growth during adolescence and early childhood. *Learn. Instruct.* 60, 263–274.
- Howe, C., McWilliam, D., Cross, G., 2005. Chance favors only the prepared mind: incubation and the delayed effects of peer collaboration. *Br. J. Psychol.* 96, 67–93.
- Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., Donaldson, C., 2007. Group work in elementary science: towards organizational principles for supporting pupil learning. *Learn. Instruct.* 17, 549–563.
- Ing, M., Webb, N.M., Franke, M.L., Turrou, A.C., Wong, J., Shin, N., Hernandez, C.H., 2015. Student participation in elementary mathematics classrooms: the missing link between teacher practices and student achievement? *Educ. Stud. Math.* 90, 341–356.
- Jadallah, M., Anderson, R., Nguyen-Jahiel, K., Miller, B.W., Kim, I.-H., Kuo, L.-J., Dong, T., Wu, X., 2011. Influence of a teacher's scaffolding moves during child-led small-group discussions. *Am. Educ. Res. J.* 48, 194–230.
- Johnson, D.W., Johnson, R.T., 2008. Social interdependence theory and cooperative learning: the teacher's role. In: Gillies, R.M., Ashman, A., Terwel, J. (Eds.), *The Teacher's Role in Implementing Cooperative Learning in the Classroom*. Springer, New York, NY.
- Jurkowski, S., Hanze, M., 2015. How to increase the benefits of cooperation: effects of training in transactive communication on cooperative learning. *Br. J. Educ. Psychol.* 85, 357–371.
- Kazemi, E., Stipek, D., 2001. Promoting conceptual thinking in four upper-elementary mathematics classrooms. *Elem. Sch. J.* 102, 59–80.
- King, A., 1992. Facilitating elaborative learning through guided student-generated questioning. *Educ. Psychol.* 27, 111–126.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., Fochy, F., 2013. A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educ. Res. Rev.* 10, 133–149.
- Langer-Osuna, J.M., Gargroetzi, E., Munson, J., Chavez, R., 2020. Exploring the role of off-task activity on students' collaborative dynamics. *J. Educ. Psychol.* 112, 514–532.

- Langer-Osuna, J.M., 2016. The social construction of authority among peers and its implications for collaborative mathematics problem solving. *Math. Think. Learn.* 18, 107–124.
- Lin, T.-J., Jadallah, M., Anderson, R.C., Baker, A.R., Nguyen-Jahiel, K., Kim, I.-H., Kuo, L.-J., Miller, B.W., Dong, T., Wu, X., 2015. Less is more: teachers' influence during peer collaboration. *J. Educ. Psychol.* 107, 609–629.
- Meloth, M.S., Deering, P.D., 1999. The role of the teacher in promoting cognitive processing during collaborative learning. In: O'Donnell, A.M., King, A. (Eds.), *Cognitive Perspectives on Peer Learning*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 235–256.
- Mercer, N., Dawes, L., Wegerif, R., Sams, C., 2004. Reasoning as a scientist: ways of helping children to use language to learn science. *Br. Educ. Res. J.* 30, 359–377.
- Mercer, N., 1996. The quality of talk in children's collaborative activity in the classroom. *Learn. Instruct.* 6, 359–377.
- Mercer, N., 2000. *Words and Minds: How We Use Language to Think Together*. Routledge, New York, NY.
- Palincsar, A.S., Anderson, C., David, Y.M., 1993. Pursuing scientific literacy in the middle grades through collaborative problem solving. *Elem. Sch. J.* 93, 643–658.
- Reznitskaya, A., Kuo, L.-J., Clark, A.-M., Miller, B., Jadallah, M., Anderson, R.C., Nguyen-Jahiel, K., 2009. Collaborative reasoning: a dialogue approach to group discussions. *Camb. J. Educ.* 39, 29–48.
- Rojas-Drummond, S., Mercer, N., 2003. Scaffolding the development of effective collaboration and learning. *Int. J. Educ. Res.* 39, 99–111.
- Roscoe, R.D., Chi, M.T.H., 2008. Tutor learning: the role of explaining and responding to questions. *Instr. Sci.* 36 (4), 321–350.
- Shah, N., Lewis, C.M., 2019. Amplifying and attenuating inequity in collaborative learning: toward an analytical framework. *Cognit. Instruct.* 37, 423–452.
- Slavin, R.E., Lake, C., 2008. Effective programs in elementary mathematics: a best-evidence synthesis. *Rev. Educ. Res.* 78, 427–515.
- Smagorinsky, P., Fly, P.K., 1993. The social environment of the classroom: a Vygotskian perspective on small group process. *Commun. Educ.* 42, 159–171.
- Tolmie, A., Thomson, J.A., Foot, H.C., Whelan, K., Morrison, S., McLaren, B., 2005. The effects of adult guidance and peer discussion on the development of children's representations: evidence from the training of pedestrian skills. *Br. J. Educ. Psychol.* 96, 181–2004.
- Van Boxtel, C., van der Linden, J., Kanselaar, G., 2000. Cooperative learning tasks and the elaboration of conceptual knowledge. *Learn. Instruct.* 10, 311–330.
- Veenman, S., Denessen, E., van den Akker, A., van der Rijt, J., 2005. Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. *Am. Educ. Res. J.* 42, 115–151.
- Webb, N.M., Nemer, K.M., Ing, M., 2006. Small-group reflections: parallels between teacher discourse and student behavior in peer-directed groups. *J. Learn. Sci.* 15 (1), 63–119.
- Webb, N.M., Franke, M.L., Ing, M., Chan, A., De, T., Freund, D., 2008. The role of teacher instructional practices in student collaboration. *Contemp. Educ. Psychol.* 33, 360–381.
- Webb, N.M., Franke, M.L., De, T., Chan, A.G., Freund, D., Shein, P., Melkonian, D.K., 2009. Teachers' instructional practices and small-group dialogue. *Camb. J. Educ.* 39, 49–70.
- Webb, N.M., Franke, M.L., Ing, M., Wong, J., Fernandez, C.H., Shin, N., Turrou, A.C., 2014. Engaging with others' mathematical ideas: interrelationships among student participation, teachers' instructional practices, and learning. *Int. J. Educ. Res.* 63, 79–93.
- Webb, N.M., Franke, M.L., Johnson, N.C., Turrou, A.C., Ing, M., 2018. Dude, Why did you start without me? Fostering engagement with others' mathematical ideas. In: Chinn, C., Manalo, E., Uesaka, Y. (Eds.), *Promoting Spontaneous Use of Learning and Reasoning Strategies: Theory, Research, and Practice*. Routledge, pp. 292–309.
- Webb, N.M., Franke, M.L., Ing, M., Turrou, A.C., Johnson, N.C., Zimmerman, J., 2019. Teacher practices that promote productive dialogue and learning in mathematics classrooms. *Int. J. Educ. Res.* 97, 176–186.
- Webb, N.M., Franke, M.L., Johnson, N.C., Ing, M., Zimmerman, J., 2021. Learning through explaining and engaging with others' mathematical ideas. *Math. Think. Learn.* <https://doi.org/10.1080/10986065.2021.1990744>.
- Webb, N.M., 2010. Peer learning in the classroom, pp. 162–169. In: Jarvela, S. (Ed.), *International Encyclopedia of Education*, third ed., vol. 6. Elsevier, pp. 636–642.
- Wittrock, M.C., 1990. Generative processes of comprehension. *Educ. Psychol.* 24, 345–376.

Dialogic education: tensions and dilemmas

Rupert Wegerif, Shengpeng Shi, Ana Rubio-Jimenez, Yun Long, Qian Liu, and Chih Ching Chang, Faculty of Education, University of Cambridge, Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Questions of definition	600
Epistemological dialogic education: constructing knowledge together	601
Ontological dialogic education: new selves and new worlds	601
The historical origins of dialogic education	601
Socrates' contrast between "living" and "dead" meaning	602
Buber's philosophy of dialogue	602
Bakhtin's dialogism	602
Freire's pedagogy of the oppressed	602
Some contemporary approaches to dialogic education	603
Nystrand's dialogically organized instruction	603
Mortimer and Scott's communicative approach	603
Alexander's dialogic teaching framework	604
Teaching the dialogue so far	604
Accountable talk	605
Thinking together	605
Learning communities	606
Tensions, dilemmas and future directions	606
Dialogic ideal versus transmission of the curriculum	606
Cultural imperialism	607
Dialogic education and Chinese culture	607
Dialogic teaching and special educational needs and disability (SEND)	607
Technology and dialogic education	608
Methodological issues for researching educational dialogues	608
Summary and conclusions	609
References	609

Questions of definition

"Dialogic" combines the Greek word "dia" meaning "through or across" with the Greek word "logic" meaning pertaining to words and speech but also, perhaps most relevantly here, pertaining to reason. From these roots the word "dialogic" can be seen as referring to "reason across difference". The term dialogic was first used to refer to novels in the form of a dialogue. The philosopher Michael Bakhtin extended this meaning to refer to novels, such as those of Dostoevsky, where multiple characters interact with different points of view which are not reduced to a single reality. For Bakhtin and the Bakhtin circle, dialogism was extended beyond literature to become a general theory of meaning. Meaning is not contained by words and utterances but is found only in the dialogue between them and their context. The meaning of what is said in a dialogue is a response to things that have been said before and an anticipation of things that will be said in response. One implication of this is that wherever there is meaning there is more than one voice in play. Another implication is that there are no final meanings, only an ongoing open dialogue. Dialogic is a contrast to monologic or the view that there is a single correct voice or final view of things.

Dialogic education uses dialogue between students and between teachers and students with the goal that students learn how to engage more effectively in dialogue both as a way of learning and as a way of thinking. We are using dialogic education here as an umbrella term for the various traditions that we think fit, at least in part, our definition, including dialogic teaching (Alexander, 2008), dialogic learning (Flecha, 2000), thinking together (Mercer and Littleton, 2007), accountable talk (Michaels et al., 2008), and dialogic pedagogy (Matusov, 2009).

Traditions of dialogic education have been divided according to whether they are more epistemological or more ontological (Matusov, 2009; Wegerif et al., 2019; Kim and Wilkinson, 2019). Epistemological dialogic education focuses on knowledge construction and ontological dialogical education focuses on changing self, society and experienced reality in the direction of becoming more dialogic or dialogical (where dialogic and dialogical are synonyms). In practice, this epistemological/ontological theory divide tends to correspond with a "talk only" versus "dialogue more generally" divide. Where the focus is only on teaching a way of talking then the theory focus tends to be on shared knowledge construction, whereas if the practice extends beyond talk to include multi-media and embodied position taking perhaps through music, dance or drama then, whether this is implicit or explicit, the education programmes are likely to aim at increasing dialogicality as a way of being in the world and not just as a way of knowing.

Epistemological dialogic education: constructing knowledge together

Dialogic or dialogism refers to a theory of how we make meaning and construct knowledge (Linell, 2009; Wells, 1999). It starts with the claim that all meaning for us arises first as an answer to questions that we ask explicitly or implicitly within dialogues. In this respect, the meaning of utterances and other signs can only be understood in context. A word someone uses in a dialogue does not necessarily have the same meaning as the same word in the dictionary; its meaning is given through its position and role in the dialogue as a response to what has been said as an attempt to influence what will be said (Linell, 2009; Rommetveit, 1992). The big epistemological implication of this is that to understand the meaning of any sign or utterance we have to look at the context, especially the historical context, and be aware that there is always more than one relevant voice or perspective in play. The big educational or pedagogical implication is that, since knowledge is always dialogic, we should not teach facts without context but need to teach students how to construct knowledge and make sense of knowledge by teaching them how to engage in dialogue and how to understand the historical and evolving nature of knowledge.

This epistemological focus can be found in most approaches to dialogic education, such as Alexander's (2008) dialogic teaching (2008), Lipman's (2003) community of enquiry, Littleton and Mercer's (2013) exploratory talk, and Michaels et al.'s (2008) accountable talk.

Ontological dialogic education: new selves and new worlds

The above epistemological conception of dialogic education is close to constructivism, or, more specifically, social constructivism (Bereiter, 2005; Paavola and Hakkarainen, 2009), insofar as both dialogic education and social constructivism understand meaning or knowledge to be constructed socially. However, dialogic education differs from constructivist approaches to education in its ontological understanding: it places less emphasis on the knowledge that is constructed than on the relationships and spaces of dialogue opened up by the differences between perspectives (Wegerif, 2007). One aim of a more ontological understanding of dialogic is the expansion of dialogue or dialogue as an end in itself within the curriculum.

Different strands of the ontological conception of dialogic education focus variously on understanding and transforming (a) the self, (b) social and political reality or (c) experienced reality as a whole.

- (a) One strand of ontological dialogic education focuses on developing the self as a more dialogic author (Matusov, 2009; Sidorkin, 1999). Through dialogic education, students are more open to embracing and learning through differences and better at engaging in dialogue, thinking, and holding multiple perspectives together in reflective tension (Wegerif, 2007, 2013).
- (b) Freire (2005) proposed dialogic teaching as a way to change the social and political reality in the direction of greater social justice. Dialogic education, in this model, is a form of critical education. Whereas monologic education or what Freire referred to as the banking model, deposits fixed truths in the heads of students and so does not change the social order, dialogic education invites students to question and to name the world afresh for themselves.
- (c) Another strand of the ontological conception of dialogic education emphasizes understanding and transforming experienced reality as a whole. It considers education as a journey of development from monologic to dialogic. This journey begins with the monologic illusion that selves and objects are separate entities within an external fixed reality and moves toward the dialogic realization that all identities are aspects of a universal dialogue—multiple, chiasmic and co-constructed—with which people can learn to engage more reflectively through education (Wegerif, 2007; Kennedy, 2014).

Epistemological dialogic education and ontological dialogic education are not incompatible. In practice, it is common for dialogic approaches to engage students in an ongoing process of shared enquiry as a way to gain knowledge (i.e., epistemological) with an additional focus on developing a dialogic orientation to expand dialogue as a valued end in itself (i.e., ontological). For examples of the combination of epistemological with ontological dialogic education see Nystrand et al. (1997), Flecha (2000), Lefstein and Snell (2013), and Phillipson and Wegerif (2017). Given that there are already many terms for collaborative learning and social constructivist pedagogy, it could be argued that the use of the term dialogic education should be reserved for educational approaches in which there is an ontological as well as an epistemological dimension in that dialogue is being promoted and expanded as an end in itself as well as a means to knowledge construction.

The historical origins of dialogic education

The importance of dialogue and dialogic reasoning to education has roots deeper and broader than the Western enlightenment. Evidence of both dialogic educational practices and recognition of the educational value of dialogue can be found in non-Western thinking much earlier than the enlightenment. In the *Analects*, for example, Confucius encourages his students to interpret, persuade, and reason through dialogue (Li and Wegerif, 2014). The notion of "forming circles of learning" where different views are discussed together can be found in the traditional Islamic approach to education (Makdisi, 1990, p. 210), along with traditional Jewish education (Schwarz and Baker, 2016) and Indian Buddhist intellectual traditions (Sen, 2005).

Socrates' contrast between "living" and "dead" meaning

Some writers replace the word dialogue with "talk" preferring, for example, to write about research on classroom talk rather than classroom dialogue (Mercer, 2010). However, not all talk is dialogic. Barnes, one of the earliest researchers of classroom talk, distinguishes between "presentational talk" which is not dialogic and "exploratory talk" which is (1976). Exploratory talk is distinguished by the fact that thinking is visible in the talk as Barnes puts it. This theme can be traced back to the ancient Greek philosopher Socrates, who taught his students orally, through questioning. The technology of alphabetic writing was spreading to Greece in Socrates' lifetime. As recorded by Plato, Socrates questions the cognitive value of writing and claims that real intelligence exists only in verbal dialogue, not in written texts. He argues that words spoken in dialogue are "living", capable of answering back and defending themselves, whereas written words are dried up and "dead" like seeds left in the sun or orphans with no-one to defend them (Plato, 360 BCE). Of course, later writers like Bakhtin managed to show that written words too could be living dialogic words but nonetheless the important contrast Socrates drew between living and dead words is the essence of the dialogic versus monologic distinction taken up by all dialogic theorists.

Buber's philosophy of dialogue

Writing in Germany in the 1930s before the second World War, Jewish philosopher Buber (1958) distinguishes between an objectifying orientation ("I-it") and an inter-subjective orientation ("I-thou"). The "I-it" orientation locates and defines things externally with the monologic assumption of a single true perspective. The "I-thou" orientation enters into a responsive dialogue in which the perspective of the other or interlocutor is taken seriously. Interestingly, Buber's concept of the other includes nature, trees and dogs, and God in the form of the "eternal Thou". Buber is not saying that these are to be thought of as separate responsive consciousnesses but that dialogue is best thought of as an orientation that opens a shared space into which we can enter rather than a relationship between separate entities exchanging words or other tokens. Buber (1958) writes of the space "in-between", or the "space of meeting", thus perhaps first introducing the theme of dialogic space, found in dialogic education today (e.g., Yang, 2016; Wegerif, 2008; Mercer et al., 2010).

Bakhtin's dialogism

In his theory of dialogism, Bakhtin (1981, 1986), an influential point of reference for contemporary dialogic education, develops Socrates' and Buber's accounts of the essential dialogic distinction between the space inside dialogue and the space outside; the living voice of the act of saying and the dead voice of that which has been said. At one point he conceptualizes this distinction in terms of the difference between the authoritative voice he associates with schoolteachers and the persuasive voices he associates with peers. The authoritative voice keeps people outside the dialogue in an isolated and static condition; they cannot enter the dialogue but must simply accept or reject it. Therefore, the dialogue terminates without generating any newly constructed knowledge or perspectives. In contrast, the persuasive voice embeds people in the dialogue, being both inside the dialogue for interanimation and outside the dialogue for reflection. Bakhtin (1981) commends the persuasive voice for its ability to awaken new and independent voices productively and creatively and make the dialogue move forward. Applied to education, this essential dialogic distinction implies that the authoritative transmission of knowledge may reduce the diversity of voices and fail to promote deeper understanding. Students can learn more effectively if they are called into and able to locate themselves both within and outside persuasive dialogues to engage with and adopt different points of view.

Bakhtin (1986) emphasizes the potentially infinite nature of meaning in dialogues. Within persuasive dialogues, there are always more perspectives, pulling in different directions, that can be inter-animated, thought with, and incarnated. This resonates with Kristeva's (1966) claim, building on Bakhtin, that there is always "intertextuality" between dialogues.

Freire's pedagogy of the oppressed

Freire (2005), a Brazilian educator, explicitly argues for the social significance of dialogic education in his book *Pedagogy of the Oppressed*. He criticizes conventional education as following a "banking" model and thus oppressing students through manipulation, because it treats knowledge as fixed facts to be inserted into the minds of students. In contrast, according to Freire, dialogic education is about empowering students to speak their own words so that they can name the world in their own ways. In doing so, both an expansion of consciousness and a transformation of social reality can occur. According to some critics, this view may itself be considered manipulative (Matusov, 2009), because the criteria determining what counts as social justice are pre-established rather than dialogically constructed. Regardless of its political implications, however, Freire's account of the empowerment of students implies that dialogic teaching and learning should be based on the mutual respect and collaboration of teachers and students and designed to promote students' voices and responses rather than rendering them merely knowledge receivers. This is one of the key values of dialogic education. In addition, Freire's focus on liberating and empowering students to participate equally and collectively in dialogues that shape a shared social reality reflects a key objective of dialogic education.

Some contemporary approaches to dialogic education

From the above theoretical avenues, we can see that dialogic teaching and learning take place through opening the dialogic space, in which different perspectives can inter-animate and inter-illuminate each other for conceptual understanding. This approach not only offers a way to teach “for intelligence” but also indicates a more explicit purpose: education “for dialogues” (Resnick and Schantz, 2015). Contemporary dialogic approaches to education are, either explicitly or implicitly, designed to open up and expand dialogic spaces in which (1) knowledge can be co-constructed; (2) dialogic orientations, characterized by an openness to the possibility of learning from others and from otherness, can be promoted; (3) and social norms that support productive and reflective dialogue can be mutually developed.

This section outlines several contemporary dialogic models, approaches and frameworks with different foci in different contexts and outlines research evidence as to the impact of dialogic education.

Nystrand's dialogically organized instruction

Informed by Bakhtin's dialogism (1981, 1986), Nystrand and his colleagues investigated hundreds of English lessons in American secondary education, from which they synthesize and extract signs and features of effective teaching in classrooms. Based on their analysis, they suggest that the potential of language for conceptual understanding depends largely on how utterances are treated within instruction; that is, how teachers organize instruction (Nystrand et al., 1997). Hence, they propose a dialogic form of instruction, characterized by respecting and giving space to students' voices, which they call “dialogically organized instruction” (Nystrand et al., 1997).

In Nystrand's model of dialogically organized instruction, three “dialogue moves” frequently used by teachers to organize dialogic instruction are emphasized: uptake, authentic questions, and high-level evaluation. Uptake refers to the process of understanding the voices of students, allowing the teacher to validate particular students' ideas by incorporating their responses into subsequent questions (Nystrand et al., 1997). This reflects the role of the teacher in inter-connecting the voices of students and building on their contributions. Different from questions with pre-determined answers, authentic questions are designed to invite students' perspectives and ideas into educational dialogues, promoting teacher-student interactions. Similarly, high-level evaluation is opposite to low-level evaluation in which teachers respond to and assess students' ideas with simple repetition or perfunctory praise. Instead, high-level evaluation requires teachers to respond to students' ideas with elaboration, commentary, or follow-up questions (Nystrand et al., 1997). These three inter-related “dialogue moves” illustrate how teaching can be done dialogically to fuel students' understanding in classroom dialogues.

Mortimer and Scott's communicative approach

Through analysis of observed teaching episodes in science classrooms, Mortimer and Scott (2003) derive an analytical framework to characterize teacher-student discourses in science education. Within this framework, classroom discourses have two dimensions: dialogic–authoritative and interactive–noninteractive. The former refers to the extent to which teachers take into account students' ideas in the flow of discourse, whereas the latter represents the degree of teacher-student interaction (Mortimer and Scott, 2003). Combining these two dimensions, four types of teacher-student discourses are defined, as shown in Table 1.

Based on their empirical investigations in science classrooms, Mortimer and Scott (2003) argue that efficient teaching entails alternation between the four types of teacher-student discourse. Dialogic discourses (i.e., NI/D and I/D) suit the goal of engaging students in exploration of their pre-existing ideas about and experiences of scientific phenomena, whereas authoritative discourses (i.e., NI/A and I/A) fit the objective of introducing and focusing on the “scientific story” (e.g., the scientific meaning of concepts, Mortimer and Scott, 2003, p. 18).

Table 1 Types of teacher–student talk.

	<i>Non-interactive (NI)</i>	<i>Interactive (I)</i>
Authoritative (A)	NI/A: Monologic lecture	I/A: Recitation
Dialogic (D)	NI/D: Dialogic lecture, presenting various points of view	I/D: Discussion

(1) The NI/A type represents mere lecturing based on a monologic orientation, in which the teacher and students do not interact.

(2) The NI/D type can be considered as lecturing based on a dialogic orientation, in which the teacher and students do not take turns, but multiple voices and perspectives on educational content are valued and taken into account.

(3) The I/A type can be understood as recitation; although interaction occurs, different perspectives, including the voices of students, are not taken into account and attention is focused on a monologic perspective.

(4) The I/D type is dialogic discussion, in which interactions occur and different perspectives are taken into account (Mortimer and Scott, 2003).

This communicative approach implies that teachers have a repertoire of types of teacher-student discourse that they strategically deploy to suit educational purposes. This distinguishes it from Nystrand et al.'s (1997) idea that teachers should always be dialogic. One way to explain this tension is that the goal of introducing a "scientific story" can appear to some extent monologic; it requires students to learn the current consensus as to the scientific meanings of concepts. Authoritative discourse is more efficient for this than dialogue. Mortimer and Scott (2003) suggest that effective classroom discourse should follow a rhythm. In such a rhythm, the teacher-student discourses initially are dialogic, when the teacher explores students' ideas about scientific concepts. Subsequently, the discourses then shift toward a more authoritative mode as the teacher works on aspects of the scientific concepts "through shaping, selecting and marking ideas. Finally, the teacher synthesizes key points and reviews the students' progress. In other words, dialogic discourse calls for subsequent authoritative discourses in which the teacher intervenes to clarify the scientific perspective (Scott et al., 2007).

Alexander's dialogic teaching framework

Alexander offers a comprehensive framework for explicating dialogic teaching and supporting teachers who intend to improve the quality and power of educational dialogues. This framework is largely derived, he claims, from his comparative educational study on classroom discourses in English, Russian, American, Indian, and French primary education, underpinned by Bakhtin's (1981) dialogism and Vygotsky's (1978) socio-cultural theory. In the fieldwork for that comparative analysis, Alexander observed and interviewed teachers, students, and other stakeholders in the above mentioned five educational systems to explore their educational practices and how these practices related to the culture, policies, and structures of the different systems.

Within the framework, Alexander (2008, p. 185) defines dialogic teaching as a general pedagogical approach that "harnesses the power of talk to engage (students), stimulate and extend their thinking, and advance their learning and understanding," which can be characterized by a set of core principles and indicators. The principles are:

- Collectiveness: teachers and children address learning tasks together, whether as a group or as a class
- Reciprocity: teachers and children listen to each other, share ideas and consider alternative viewpoints
- Supportiveness: children articulate their ideas freely, without fear of embarrassment about "wrong" answers, and they help each other to reach a common understanding
- Cumulativeness: teachers and students build on their own and each other's ideas and chain them into coherent lines of thinking and enquiry
- Purposefulness: teachers design and steer classroom talk with specific educational goals in view

Similar to the communicative approach, Alexander's framework is based on the notion of a repertoire, requiring teachers to have a repertoire of methods of organizing interactions and engaging in discourse and to be able to strategically adopt different types of classroom discourse to achieve the goals of the curriculum. The types of classroom discourse include both teaching talk (e.g., exposition, rote, lecture) and learning talk. The latter enables students to explore, interpret and argue. Although this "repertoire" approach includes some monologic types of talks, it privileges discussion and dialogue, which stimulate students' thinking and advances their learning. For Alexander, discussion pursues the goals of information-sharing and problem-solving, whereas dialogue, as directed inquiry and guided questioning, can achieve a mutual understanding and scaffolding to foster independence. Alexander regards discussion and dialogue as "the forms of talk which are most in line with prevailing thinking on student's learning" (2008, p. 103) and "have by far the greatest cognitive potential" (2017, p. 31). This emphasis makes it more aligned with dialogically organized instruction (Nystrand et al., 1997), and distinct from the communicative approach (Mortimer and Scott, 2003).

Teaching the dialogue so far

The above three approaches share an epistemological position focusing on the power of discourse for knowledge co-construction. More recently, another approach to dialogic teaching has emerged, focusing on the ontological stance of dialogue. Grounded largely in Bakhtinian dialogism and Oakeshott's (1960) notion of education inducting students into "the conversation of mankind," and inspired by Langer's (2016) emphasis on the mindful learning of teaching everything from a range of perspectives and as emerging from debate, Phillipson and Wegerif (2017) argue that dialogic teaching is not just about constructing knowledge but is also about "teaching the dialogue so far." This builds on what Linell (2009) refers to as "double dialogic", that voices are not just in dialogue with each other but also with the cultural and historical context. In an educational dialogue in a classroom there is often engagement with a subject discipline which itself is a form of dialogue, albeit, a long-term cultural dialogue. Education involves inducting students into participation, through dialogue, in the long-term dialogue of culture which, although it has many strands, is ultimately one open-ended dialogue of humanity. Oakeshott (1960) refers to education in this sense as giving students access to their "inheritance."

The role of the teacher is not to transmit fixed and final knowledge but to summarize and present "the dialogue so far", so that latecomers can catch up (Phillipson and Wegerif, 2017). Knowledge should be presented in the ways in which it has previously been constructed, as a story about the dialogue that led to the "knowledge", the questions asked and how and why they were asked. The

significance of “teaching everything as fallible and dialogic” is to leave students free to build upon and reflect upon what they have been taught and to challenge and develop it in the future. In this way, “teaching the dialogue so far” enables students to challenge and develop it through participation in an ongoing and open-ended shared inquiry.

Accountable talk

The accountable talk approach places thinking at the core of the curriculum and emphasizes the primacy of talk in promoting thinking. The notion of “accountable talk” was originally abstract, introduced by Resnick (1999), Resnick and Hall (1998), but it was subsequently developed into a more concrete pedagogical approach (Resnick et al., 2010, 2018a,b). Talk is understood to promote learning and thinking when it is accountable to accurate knowledge, rigorous reasoning, and the learning community. “Accountability to knowledge” refers to the content of students’ discussion. As far as possible, they should argue based on specific and accurate information or solid evidence. From this perspective, students may challenge the evidence used by others within argument (e.g., “are those statistics accurate?”). Accountability to rigorous thinking is ensured when talk follows a logical and coherent line of argument from premises to conclusions. In other words, students must not only talk about talk but also talk about reasoning (e.g., “what’s your reasoning for that explanation?”). Accountability to the learning community is about how students talk to and with each other. They are expected to listen to and respect each other and focus on their classmates’ contributions, so that they can understand, challenge, build upon, and refine each other’s ideas. This can lead to facilitative questions (e.g., “does anyone want to add on?”).

These expectations are explicitly imposed on students as a way to promote learning and thinking. To some extent, this approach is similar to dialogically organized instruction (Nystrand et al., 1997), as a variety of dialogue moves, such as revoicing, are key to accountable talk. However, there is more focus on building social norms of classroom talk by enabling students to learn and value accountability in their talk. In other words, it aims not only to enable teachers to become dialogic teachers, but also to enable students to become dialogic learners. In doing so, dialogic norms can be developed by them all, forming a classroom culture.

Research on accountable talk suggests that it leads to “robust, sometimes remarkable, academic achievements for working-class and middle-class students alike, and for students from a range of linguistic backgrounds” (Michaels et al., 2008, p. 286). This research also reveals certain limitations of accountable talk, especially the extent to which its success depends on the commitment and ability of particular teachers. However, this challenge is common to all dialogic education programmes.

Thinking together

Another dialogic approach, Thinking Together, is similar to accountable talk but developed independently in the UK (Dawes et al., 2004). This approach is influenced by Bakhtin’s idea of persuasive voices and Habermas’ notion of an ideal speech situation (1979, 1984), but it is based primarily on Vygotsky’s (1991) theory that the “higher mental functions” originate in forms of social interaction. Thinking Together aims to promote dialogues in the form of “exploratory talk”. This is talk in which “participants pool ideas, opinions and information, and think aloud together to create new meanings, knowledge and understanding” (Mercer et al., 2019b, p. 368).

The Thinking Together approach further proposes that students should engage in “exploratory talk”, allowing differences of opinion to be discussed and explored and leading to the development of an understanding (i.e., ground rules) for everyone engaged in the interaction. This notion of ground rules can be understood as dialogic norms that are co-created and agreed upon by the teacher and students. Such ground rules are applied when students talk and solve problems together. Each class can develop its own set of ground rules, but these are all variations of the ground rules suggested by the original developers of the approach:

- All relevant information is shared openly.
- Each group member should be actively encouraged to contribute to the discussion.
- Everyone should listen to others attentively.
- Each suggestion should be carefully considered.
- Group members are asked to provide reasons for ideas and opinions.
- Constructive challenges to ideas are accepted and a response is expected.
- Alternatives are discussed before a decision is taken.
- The group works together with the purpose of reaching agreement.
- The group, not the individual, takes responsibility for decisions made.

Dawes et al. (2004).

Based on the notion of “exploratory talk” and collaborating closely with teachers, the Thinking Together research team members developed a series of “Talk Lessons” for teaching these ground rules and applying them within the formal primary education (Dawes, 2010; Phillipson and Wegerif, 2017). These lessons encourage teachers to model dialogue and guide children to use discourse as a tool for both individual reasoning and collaborative problem solving. Experimental implementations and evaluations of the Thinking Together approach have been conducted as part of various research projects in countries such as Mexico and South Africa. These studies demonstrate that significant improvements can be achieved in curriculum-specific areas and in scores on Raven’s non-verbal reasoning tests, which is a test that correlates highly with IQ (Mercer and Dawes, 2010; Webb et al., 2016).

Learning communities

The above three approaches focus on promoting thinking among students by enabling them to engage in and expand educational dialogues. Another approach, the learning communities project, reaches beyond the school. As well as promoting educational achievement in the curriculum, this approach is directed toward the social transformation of broader communities based around schools through egalitarian dialogue among all community members, including teachers, students, families, and volunteers. Drawn largely from Freire's ontological stance on dialogues and Habermas's notion of intersubjectivity, [Flecha \(2000\)](#) argues that the implementation of dialogic learning is most influential if it takes place not just in school but in all of a student's important relationship settings, including the home and the workplace. Inspired by similar projects, such as the "Accelerated Schools" project at Stanford ([Levin, 1998](#)) and worker cooperatives and democratic models of work distribution, Flecha intends to develop learning communities capable of analyzing the current situation of their schools and creating a joint vision of the future of those schools.

Specifically, the learning communities project starts with an awareness raising stage, during which issues of social inequality and information literacy are discussed. All participants must engage in this awareness raising; then, they are given the opportunity to decide, jointly, to implement the transformation of the school, forming a learning community. This community is expanded as teaching staff agree to participate, along with the parent association and other stakeholders. A blueprint of action is then drawn up, including actions such as establishing teaching in what are termed "interactive groups". Extra resources are often needed by teachers to handle this new pedagogy. If so, volunteer members of the community are encouraged to come in and help. One example, given by Flecha, is that of a Roma grandmother who herself was not literate but was able to come into school and very effectively organized a learning group. Families are engaged in and offered help with literacy and the use of ICT in the schools. Where conflicts or concerns arise, these are dealt with through dialogues where all parties are equal participants. Through interactive groups and the presence of parents within the classroom, conflicts between groups tend to disappear from the classroom. Evaluations using mixed methods reveal that, according to families in the program, the approach has had a very positive impact on children's achievement and engagement in education ([Flecha and Soler, 2014](#); [Flecha, 2014](#)). The learning community project offers an example of the potential of dialogues to shape and bring into being a shared social vision.

Tensions, dilemmas and future directions

This section explores some of the tensions that exist within dialogic education as well as responding to some of the criticisms that have been leveled at it from the outside. Responses to these tensions and challenges suggest future directions for the development of dialogic education as a strand within the ongoing global dialogue about what we should teach our children and how we should teach it.

Dialogic ideal versus transmission of the curriculum

A true dialogue is open-ended. It is impossible to predict where it will end up. This sets the dialogic ideal in conflict with the idea of a pre-planned curriculum with set learning goals specifying knowledge and skills. It also raises a challenge for assessment. As [Teo \(2019\)](#) points out, the current modes of assessment in many schools still treat "student knowledge and skills as fixed, discrete and assessable through individualized and standardized pen and paper tests" (p. 176).

[Matusov \(2009\)](#) criticizes much so-called "dialogic education" as being essentially instrumental, applying dialogue to serving curriculum goals without ever really becoming truly dialogic. To adopt a genuine "dialogic pedagogy" instead of instrumental dialogic education, teachers should see themselves as the first learners in classrooms, Matusov claims. Highlighting the equal power dynamics, teachers should be ignorant in a sense that they are not afraid to "suspend the certainty of their own knowledge" and to explore it with students again ([Matusov, 2009](#), p. 6). In that sense, both teachers' and students' consciousness should be taken into equal account as they seek information from and with each other ([Skidmore, 2010](#)).

Responding to Matusov's challenge, [Wegerif \(2013\)](#) argues that where there is a curriculum goal this can be approached in a way that also teaches students how to be more dialogic. The goal of promoting more dialogue can be achieved even within a relatively fixed knowledge focused curriculum. In another response to Matusov, [Miyazaki \(in Matusov and Miyazaki, 2014\)](#) points out that every individual has a unique relation with the world that surrounds us, which he refers to as "the ternary relation" (p. 17). That means, when two people interact whether in a monologic or dialogic way, there is always a third relation with the world itself. Miyazaki is perhaps pointing to the significance of the double dialogic referred to by [Linell \(2009\)](#), whereby voices in a dialogue are not just in dialogue with each other but also with their context. As described earlier, [Phillipson and Wegerif \(2017\)](#) turn this double-dialogic idea into a key principle of dialogic education. As well as short term dialogues in classrooms between students and teachers there are also, they claim, long-term dialogues of culture in which conceptual knowledge is developed and transmitted in the form of answers to questions. These long-term dialogues of culture, in science for example, are truly open-ended. While one of the goals of short-term dialogues in classrooms is to develop a capacity for dialogue as an end in itself, another goal is to induct students into

participation in the long-term dialogues of culture. In other words, there are levels of dialogue, and transmitting what has happened in the dialogue so far so that students can join in and take it forward into the future is not the opposite of dialogic education but a key part of dialogic education.

Cultural imperialism

Many approaches (e.g., Thinking Together; accountable talk) involve explicitly promoting a specific form of talk, characterized by largely pre-determined “ground rules” or social norms, which may fail to recognize and value the culture of the students. As Lam-birth (2006, 2009) claims, setting up the ground rules or social norms might be a form of cultural imperialism, in which ways of talking determined among the powerful are imposed on the less powerful. This is also the implication of a critique of the undemocratic way in which teachers sometimes impose ground rules (Barak and Lefstein, 2021). In some cases, the values and norms of the middle classes might be imposed on the working classes. In other contexts, in China for example, it might be argued that western ideas, values, and norms are being imposed upon other cultures.

Littleton and Mercer (2013) respond to this issue with the claim that the exploratory talk promoted in the Thinking Together project is not a class-based or culturally based norm of talking, but a way of talking and thinking together that is educationally effective. In the Thinking Together approach, shared ground rules are developed together by students and teachers on the basis of their experience with group work and their intuitions as to what is likely to be most effective. These ground rules are then meant to be revisited at frequent intervals, at least once a term, to see if they still work or need revising. Agreeing with and building on these points from Littleton, Kershner et al. (2020) argue that what is fundamental are not the specific ground rules but a certain dialogic orientation toward the other. How best to implement this dialogic orientation (in and through exploratory talk) in terms of ground rules will vary depending on the cultural and historical context, but the inter-subjective orientation itself is claimed to be a universal feature of communication that is productive for learning together and has emerged in many times and places.

Dialogic education and Chinese culture

In practice in the USA and the UK dialogic education involves talk in which students initiate dialogues. Fung (2014) points out some tension between dialogic education and Chinese classroom culture which he claims is characterized by “self-effacement” (谦逊, to achieve humility and avoid conflict) and “Li” (ritual, “礼”, to respect others and be less willing to challenge a hierarchy or authority). Teachers, historically seen as a sage or at least as authoritative figures, need to be respected in the Chinese context. As a result, asymmetrical interactions between the teacher and students are likely to be formed and habitually accepted. Additionally, seeking for harmony in Chinese traditional culture may impede argumentative discussions, challenging and competing views (Li and Wegerif, 2014). When contextualizing the Thinking Together program in China, Yang (2016) found that the teacher’s underlying assumptions contradicted some key expectations of dialogue, but she later negotiated the social and cultural aspects of the school with various parties and developed awareness of using talk. The implementation of the program led to the emergence of different ground rules than in the UK, more appropriate for the Chinese classroom context, but, nonetheless, effective for learning (Yang, 2016). As Lefstein (2010) argues, there is no need to remove the existing and inherent tensions of dialogic pedagogy, as new infinite possibilities such as innovative teaching methods are likely to be generated out of the tension of difference and conflict.

Chinese scholars define dialogic education from different perspectives. These definitions can be categorized in terms of Wegerif’s (2019) epistemological level (dialogue for learning) and ontological level (dialogue as learning). Luan and Pan (2010) conclude that educator, learner and learning materials are the dialogic authorities in Chinese classrooms. Li and Wegerif (2014) argue that the Confucian tradition is not against the dialogic but emphasizes cultivating reflection as silent dialogue between multiple voices. Hu et al. (2017) support this, arguing that silence is irreplaceable in Chinese classrooms as it provides time and space for learners to have dialogue with learning materials, ancients, and the universe, and this kind of silent dialogue is essential for students’ understanding of some complex knowledge. A simple hybrid dialogic education practice that has emerged in the Chinese context, for example, is to spend time in silent reflection and individual writing about an issue before engaging in spoken dialogue about it. Liu (2001) describes the potential of dialogic education in the Chinese context and refers to empirical studies conducted by Chinese scholars suggesting that dialogic approaches in Chinese classrooms can have positive effects on students’ learning (Yang, 2016; Hu et al., 2017).

Dialogic teaching and special educational needs and disability (SEND)

International declarations urge countries to ensure an inclusive education to students of all levels (e.g., United Nations, 2006; UNESCO, 1994). However, despite the international declarations, there are barriers to inclusion of those with SEND which include the views of teachers and SEND peers who think that productive dialogue with some SEND students is not possible (Dar-row, 2009). On the other hand, the findings of a literature review carried out by Fernandez-Villardón et al. (2020) suggest that peer support can be used to tackle this barrier. Students without disabilities participating in interventions aimed at improving

SEND students' social skills increased their social support, their requests for repairs in response to unintelligible interventions of their peers with disabilities (Aalsvoort and Gossé, 2007; Weiner, 2005), and their social initiations toward children with SEND. The quality of these exchanges became longer and more reciprocal (Kohler et al., 2007). At the same time, students identified with SEND enhanced their social and communication skills, their reciprocal interactions and their unprompted initiations (e.g., Harper et al., 2008; Kalyva and Avramidis, 2005; Nelson et al., 2007). The number of distractions was also reduced (McCurdy and Cole, 2014).

Learning to use augmentative and alternative communication systems (AAC) may also help to tackle knowledge barriers and facilitate communication with SEND students. AAC is effective in increasing the functional communication skills of students with SEND and it decreases problem behaviors in some individuals with limited or no verbal communication (Hart and Banda, 2010). Students identified with profound intellectual multiple disabilities (PIMD) can also benefit from multimodal communication systems, interactive computed activities and cooperative educational activities. At the same time, their peers without PIMD can also learn these systems to increase reciprocal peer interaction and conduct cooperative tasks with their peers with PIMD (Nijs and Maes, 2014). Computer activities with social interaction can help some SEND students (Lau et al., 2005).

Despite fears that dialogic education will not work with SEND students, it has been found that, given the opportunity and support, people identified with SEND contribute to the dialogues that concern them in their childhood (e.g., Black-Hawkins et al., 2021; Snelgrove, 2005), youth (e.g., Rubio-Jimenez and Kershner, 2020, 2021) and adulthood (e.g., Dew et al., 2019). Despite negative attitudes associated with learners who are perceived as a major challenge for educators (Norwich, 2008), social interactions and communication are always possible and desirable (e.g., Goodwin, 2000; Nijs and Maes, 2014; Teachman et al., 2018).

Technology and dialogic education

Most literature on dialogic education tends to assume dialogue mediated by words in face-to-face or in-person interactions. Others argue that spoken language is just one kind of communications technology that can support dialogue, that it has limitations as well as strengths, and that new forms of dialogic education can be designed for the affordances of different technology (Wegerif and Major, 2019; Trausan-Matu et al., 2021). The use of digital technology has the potential to scale up dialogic education to include a wider range of voices (Wegerif, 2010). A growing body of research has shown the pedagogically enacted affordances of technology can enhance participants' engagement in dialogue, thereby fostering collective knowledge building and collaborative creativity (e.g., Hennessy, 2011; Pifarré, 2020). For example, digital artifacts jointly created by students and teachers on an Interactive Whiteboard open up a shared dialogic space where diverse ideas can be viewed, saved, revisited and modified, not only making different perspectives explicit to participants, but visualizing the evolvement of ideas and externalizing learning trajectories (Mercer et al., 2019a).

More than tools to support dialogue, digital technology has been described as having a voice that can relate to and resonate with other voices in dialogue. Barad's (2007) "new materialist" perspective has been recently brought to the fore and applied to highlight the role of materials, including digital technology, in dialogue (Cook et al., 2019; Hetherington and Wegerif, 2018). The separation of matter and meaning is rejected; instead, "they are always in a relationship of intra-action and inter-connection and affecting each other mutually" (Lenz-Taguchi, 2010, p. 30). This is what Barad (2007) calls "material-discursive". Barad's materialist approach applied to understand classroom dialogues in science when equipment is involved has been described as opening up a "material-dialogic" pedagogy (Hetherington and Wegerif, 2018).

Methodological issues for researching educational dialogues

Published studies on the impact of dialogic education in relation to curriculum learning gains, the development of skills, and the promotion of shared social norms have been positive (Howe et al., 2019; Alexander, 2018; Resnick et al., 2018a,b). Yet, these evaluations tend to use methodologies such as Randomized Control Trials (RCT) which some argue are not really compatible with the philosophical assumptions of dialogism because they seem to assume a monologic view of the world. Much methodology in the social sciences assumes a monologic epistemology, claiming that all apparent differences in perspectives and voices can be reduced to a single correct perspective, a master perspective. Flecha (2000) has responded to this issue by developing what is called a "communicative method" of evaluation that actively engages the voices of participants in the research process. Others argue for adopting mixed methods or design-based research as the methodology for researching educational dialogues, as it is underpinned by a pragmatist epistemology that assumes truths are provisional and are given through both experience and experimenting (Johnson and Onwuegbuzie, 2004, p. 18).

Recently, a new methodology, chiasm, has been developed which aligns better with Bakhtinian dialogism (Kershner et al., 2020). Chiasm suggests that investigating a reality composed of dialogues should combine insider perspectives (emic) with an outsider perspectives (etic) in a systematic way. The insider perspectives could be obtained through the inside-out approach (interviews for example), whereas the outsider perspective could be achieved through the outside-in approach (statistical measures of outcomes for example). The coherence and interlinking of outside-in and inside-out approaches are how the chiasm differs from multi-methods approaches (which can simply juxtapose qualitative and quantitative approaches to answer different

research questions). The chiasm brings outside-in and inside-out approaches together as a whole (answering each research question dialogically, from both the outside in and the inside out), both addressing the research questions with empirical evidence from outside-in and informing them in terms of lived experience from within. This methodology is intended to produce knowledge in the form of a living productive dialogue between inside and outside without reducing either side into the terms of the other.

Summary and conclusions

This chapter offers a brief history of some of the ideas behind dialogic education, outlines some major contemporary approaches to implementing dialogic education, and summarizes research on the impact of dialogic education. Dialogic education means opening up dialogic spaces in which different perspectives can inter-animate and inter-illuminate each other such that learning happens. But dialogic education is more than teaching and learning *through* dialogue; it is also education *for* dialogue, seeking to expand and deepen the dialogic space, enabling students to engage in dialogue more productively, reflectively and collaboratively. As a result of dialogic education, students do not only learn knowledge with deeper conceptual understanding, they also become better at communicating, thinking, and learning together with others. However, there are of course a number of challenges to dialogic education and tensions within dialogic education. We considered just a few of these, focusing on the difficulty of teaching a fixed curriculum through dialogue, the charge of cultural imperialism, especially in the context of China, and the challenge of teaching through dialogue with SEND students. The final two sections on technology and methodology suggest ways in which dialogic education and research on dialogic education might be developing for the future.

The first handbook of research on dialogic education (Mercer et al., 2019b) with over 50 articles from many countries reveals a flourishing global community. Inevitably, given the focus on dialogue, there are many strands to dialogic education and tensions between voices, but there is also the dynamic coherence and optimism of a community moving forward investigating together the meaning and potential of education.

References

- Aalsvoort, G., Gossé, G., 2007. Effects of video recorded interactions and counseling for teachers on their responses to preschoolers with intellectual impairments. *Intellect. Dev. Disabil.* 45 (2), 103–115. [https://doi.org/10.1352/1934-9556\(2007\)45\[103:E0VIAC\]2.0.CO;2](https://doi.org/10.1352/1934-9556(2007)45[103:E0VIAC]2.0.CO;2).
- Alexander, R., 2008. *Towards Dialogic Teaching: Rethinking Classroom Talk*, fourth ed. Dialogos, York.
- Alexander, R., 2017. *Towards Dialogic Teaching: Re-thinking Classroom Talk*, fifth ed. Dialogos, York.
- Alexander, R., 2018. Developing dialogic teaching: genesis, process, trial. *Res. Pap. Educ.* 33 (5), 561–598.
- Bakhtin, M., 1981. *The Dialogic Imagination*. University of Texas Press, Austin Texas.
- Bakhtin, M.M., 1986. *Speech Genres and Other Late Essays*. University of Texas Press, Austin.
- Barad, K., 2007. *Meeting the Universe Half-Way: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, Durham.
- Barak, M., Lefstein, A., 2021. Opening texts for discussion: developing dialogic reading stances. *Reading research quarterly. Read. Res. Q.* 57 (2), 449–468.
- Barnes, D., 1976. *From Communication to Curriculum*. Boynton/Cook Publishers, Harmondsworth.
- Bereiter, C., 2005. *Education and Mind in the Knowledge Age*. Routledge.
- Black-Hawkins, K., Maguire, L., Kershner, R., 2021. Developing inclusive classroom communities: what matters to children? *Education 3–13*, 1–15. <https://doi.org/10.1080/03004279.2021.1873398>.
- Buber, M., 1958. *I and Thou*. R. Gregory Smith, trans., second ed. T & T Clark, Edinburgh.
- Cook, V., Warwick, P., Vrikki, M., Major, L., Wegerif, R., 2019. Developing material-dialogic space in geography learning and teaching: combining a dialogic pedagogy with the use of a microblogging tool. *Think. Skills Creativ.* 31, 217–231.
- Darrow, A., 2009. Barriers to effective inclusion and strategies to overcome them. *Gen. Music Today* 22 (3), 29–31. <https://doi.org/10.1177/1048371309333145>.
- Dawes, L., 2010. *Creating a Speaking and Listening Classroom: Integrating Talk for Learning at Key Stage 2*. Routledge, London.
- Dawes, L., Mercer, N., Wegerif, R., 2004. *Thinking Together: A Programme of Activities for Developing Speaking, Listening and Thinking Skills*, second ed. Imaginative Minds Ltd, Birmingham.
- Dew, A., Collings, S., Dillon, I., Gentle, E., Dowse, L., 2019. "Living the life I want": a framework for planning engagement with people with intellectual disability and complex needs. *J. Appl. Res. Intellect. Disabil.* 32 (2), 401–412. <https://doi.org/10.1111/jar.12538>.
- Fernandez-Villardón, A., Alvarez, P., Ugalde, L., Tellado, I., 2020. Fostering the social development of children with special educational needs or disabilities (SEND) through dialogue and interaction: a literature review. *Soc. Sci.* 9 (11), 97. <https://doi.org/10.3390/socsci9060097>.
- Flecha, R., 2000. *Sharing Words. Theory and Practice of Dialogic Learning*. Rowman & Littlefield, Lanham, MD.
- Flecha, R., 2014. Using mixed methods from a communicative orientation: researching with grassroots Roma. *J. Mix. Methods Res.* 8 (3), 245–254.
- Flecha, R., Soler, M., 2014. Communicative methodology: successful actions and dialogic politics. *Curr. Sociol.* 62 (2), 232–242.
- Freire, P., 2005. *Pedagogy of the Oppressed*. The Continuum International Publishing Group, New York.
- Fung, D., 2014. The influence of ground rules on Chinese students' learning of critical thinking in group work: a cultural perspective. *Pedagog. Cult. Soc.* 22 (3), 337–368. <https://doi.org/10.1080/14681366.2014.899611>.
- Goodwin, C., 2000. Pointing and the collaborative construction of meaning in Aphasia. *Tex. Ling. Forum* 43, 67–76.
- Habermas, J., 1979. What is universal pragmatics? Thomas McCarthy, trans. In: Habermas, J. (Ed.), *Communication and the Evolution of Society*. Beacon Press, Boston, MA.
- Habermas, J., 1984. *The Theory of Communicative Action*, vol. 1. Polity Press, Cambridge.
- Harper, C., Symon, J., Frea, W., 2008. Recess is time-in: using peers to improve social skills of children with autism. *J. Autism Dev. Disord.* 38 (5), 815–826. <https://doi.org/10.1007/s10803-007-0449-2>.
- Hart, S., Banda, D., 2010. Picture exchange communication system with individuals with developmental disabilities: a meta-analysis of single subject studies. *Remedial Special Educ.* 31 (6), 476–488. <https://doi.org/10.1177/0741932509338354>.

- Hennessy, S., 2011. The role of digital artefacts on the interactive whiteboard in supporting classroom dialogue. *J. Comput. Assist. Learn.* 27 (6), 463–489.
- Hetherington, L., Wegerif, R., 2018. Developing a material-dialogic approach to pedagogy to guide science teacher education. *J. Educ. Teach.* 1–17.
- Howe, C., Hennessy, S., Mercer, N., Vrikki, M., Wheatley, L., 2019. Teacher–student dialogue during classroom teaching: does it really impact on student outcomes? *J. Learn. Sci.* 28 (4–5), 462–512.
- Hu, X., Zheng, T., Zeng, X., Tan, X., 2017. 对话学习模式对小学生概念转变有效性的实证研究 [an empirical research on effectiveness of conceptual change influenced by the dialogue learning model to primary school students]. *Educ. Res. (《教育研究》)* 38 (05), 130–136.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26. <https://doi.org/10.3102/0013189X033007014>.
- Kalyva, E., Avramidis, E., 2005. Improving communication between children with autism and their peers through the “Circle of Friends”: a small-scale intervention study. *J. Appl. Res. Intellect. Disabil.* 18 (3), 253–261. <https://doi.org/10.1111/j.1468-3148.2005.00232.x>.
- Kennedy, D., 2014. Neoteny, dialogic education and an emergent psychoculture: notes on theory and practice. *J. Philos. Educ.* 48 (1), 100–117.
- Kershner, R., Hennessy, S., Wegerif, R., Ahmed, A., 2020. *Research Methods for Educational Dialogue*. Bloomsbury, London.
- Kim, M.Y., Wilkinson, I.A., 2019. What is dialogic teaching? Constructing, deconstructing, and reconstructing a pedagogy of classroom talk. *Learn. Cult. Soc. Interact.* 21, 70–86.
- Kohler, F., Greteman, C., Raschke, D., Highnam, C., 2007. Using a buddy skills package to increase the social interactions between a preschooler with autism and her peers. *Top. Early Child. Spec. Educ.* 27 (3), 155–163. <https://doi.org/10.1177/02711214070270030601>.
- Kristeva, 1966. Word, dialogue, and novel. Trans. Thomas Gora et al. In: Roudiez, L.S. (Ed.), *Desire in Language: A Semiotic Approach to Literature and Art*. Columbia U. P, New York, pp. 64–91.
- Lambirth, A., 2006. Challenging the laws of talk: ground rules, social reproduction and the curriculum. *Curric. J.* 17 (1), 59–71.
- Lambirth, A., 2009. Ground rules for talk: the acceptable face of prescription. *Curric. J.* 20 (4), 423–435.
- Langer, E.J., 2016. *The Power of Mindful Learning*. Hachette, UK.
- Lau, C., Higgins, K., Gelfer, J., Hong, E., Miller, S., 2005. The effects of teacher facilitation on the social interactions of young children during computer activities. *Top. Early Child. Spec. Educ.* 25 (4), 208–217. <https://doi.org/10.1177/02711214050250040201>.
- Lefstein, A., 2010. More helpful as problem than solution: some implications of situating dialogue in classrooms. In: Littleton, K., Howe, C. (Eds.), *Educational Dialogues: Understanding and Promoting Productive Interaction*. Taylor and Francis.
- Lefstein, A., Snell, J., 2013. *Better than Best Practice: Developing Teaching and Learning Through Dialogue*. Routledge, London, U.K.
- Lenz-Taguchi, H., 2010. *Going Beyond the Theory/Practice Divide in Early Childhood Education: Introducing an Intra-active Pedagogy*. Routledge, Abingdon.
- Levin, H.M., 1998. Accelerated schools: a decade of evolution. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *International Handbook of Educational Change*. Springer, Netherlands, pp. 807–830.
- Li, L., Wegerif, R., 2014. What does it mean to teach thinking in China? Challenging and developing notions of “Confucian education”. *Think. Skills Creativ.* 11, 22–32.
- Linell, P., 2009. *Rethinking Language, Mind and World Dialogically: Interactional and Contextual Theories of Human Sense-Making*. Information Age, Charlotte, NC.
- Lipman, M., 2003. *Thinking in Education*. Cambridge University Press.
- Littleton, K., Mercer, N., 2013. *Interthinking: Putting Talk to Work*. Routledge.
- Liu, Q., 2001. 对话教学初论 [A preliminary discussion on dialogic teaching. *Educ. Res.* 2001 (11), 65–69.
- Luan, X., Pan, H., 2010. 对话教学研究20年:回顾与展望 [20 Years of dialogue teaching research: review and prospects]. *J. Tianjin Normal Univ. (Element. Educ. Ed.)* 11 (04), 50–53.
- Makdisi, G., 1990. *The Rise of Humanism in Classical Islam and the Christian West*. Edinburgh University Press.
- Matusov, E., 2009. *Journey Into Dialogic Pedagogy*. Nova Science Publishers, NY.
- Matusov, E., Miyazaki, K., 2014. Dialogue on dialogic pedagogy. *Dialogic Pedagogy* 2. <https://doi.org/10.5195/dpj.2014.121>.
- McCurdy, E.E., Cole, C.L., 2014. Use of a peer support intervention for promoting academic engagement of students with autism in general education settings. *J. Autism Dev. Disord.* 44 (4), 883–893.
- Mercer, N., 2010. *Words and Minds: How We Use Language to Think Together*. Routledge, London.
- Mercer, N., Davies, L., 2010. Making the Most of Talk: Dialogue in the Classroom. English, Drama, Media. NATE, London, pp. 19–28, 16, March.
- Mercer, N., Littleton, K., 2007. *Dialogue and the Development of Children's Thinking: A Sociocultural Approach*. Routledge.
- Mercer, N., Hennessy, S., Warwick, P., 2019a. Dialogue, thinking together and digital technology in the classroom: some educational implications of a continuing line of inquiry. In: Mercer, N. (Ed.), *Language and the Joint Creation of Knowledge: The Selected Works of Neil Mercer*. Routledge, pp. 367–394.
- Mercer, N., Wegerif, R., Major, L. (Eds.), 2019b. *The Routledge International Handbook of Research on Dialogic Education*. Routledge.
- Mercer, N., Warwick, P., Kershner, R., Staarman, J.K., 2010. Can the interactive whiteboard help to provide “dialogic space” for children’s collaborative activity? *Lang. Educ.* 24 (5), 367–384.
- Michaels, S., O'Connor, C., Resnick, L.B., 2008. Deliberative discourse idealized and realized: accountable talk in the classroom and in civic life. *Stud. Philos. Educ.* 27 (4), 283–297.
- Mortimer, E.F., Scott, P.H., 2003. *Meaning-Making in Secondary Science Classrooms*. Open University Press, Buckingham.
- Nelson, C., Nelson, A., McDonnell, A., Johnston, S., Crompton, A., 2007. Keys to play: a strategy to increase the social interactions of young children with autism and their typically developing peers. *Educ. Train. Dev. Disabil.* 42 (2), 165–181. Retrieved from: <https://www.jstor.org/stable/23879993>.
- Nijs, S., Maes, B., 2014. Social peer interactions in persons with profound intellectual and multiple disabilities: a literature review. *Educ. Train. Autism Dev. Disabil.* 49 (1), 153–165. <https://www.jstor.org/stable/23880662>.
- Norwich, B., 2008. Dilemmas of difference, inclusion and disability: international perspectives on placement. *Eur. J. Spec. Needs Educ.* 23 (4), 287–304. <https://doi.org/10.1080/08856250802387166>.
- Nystrand, M., Gamoran, A., Kachur, R., Prendergast, C., 1997. *Opening Dialogue: Understanding the Dynamics of Language and Learning in the English Classroom*. Teachers College Press, New York.
- Oakeshott, M., 1960. *The Voice of Poetry in the Conversation of Mankind: An Essay*. Bowes & Bowes, Cambridge.
- Paavola, S., Hakkarainen, K., 2009. From Meaning Making to Joint Construction of Knowledge Practices and Artefacts—A Trialogical Approach to CSCL.
- Phillipson, N., Wegerif, R., 2017. *Dialogic Education: Mastering Core Concepts Through Thinking Together*. Routledge, London, U.K.
- Pifarré, M., 2020. Designing a dialogic technology-enhanced pedagogy to support collaborative creativity. In: Mercer, N., Wegerif, R., Major, L. (Eds.), *Routledge International Handbook of Research on Dialogic Education*. Routledge, London, pp. 425–438.
- Plato, 360 BCE/2006. Phaedrus. B. Jowett, trans. Available online at: <http://ebooks.Adelade.edu.au/p/plato/p71phs/>. (Accessed 1 December 2011).
- Resnick, L.B., 1999. Making America smarter. *Educ. Week* 18 (40), 38–40.
- Resnick, L.B., Hall, M.W., 1998. Learning organizations for sustainable education reform. *Daedalus* 127 (4), 89–118. <https://doi.org/10.1007/s00404-014-3434-y>.
- Resnick, L.B., Schantz, F., 2015. Re-thinking intelligence: schools that build the mind. *Eur. J. Educ.* 50 (3), 340–349.
- Resnick, L.B., Asterhan, C., Clarke, S., 2018a. Accountable Talk: Instructional Dialogue that Builds the Mind. International Academy of Education (IAE) and International Bureau of Education (IBE).
- Resnick, L.B., Asterhan, C., Clarke, S., Schantz, F., 2018b. Next generation research in dialogic learning. In: Hall, G.E., Quinne, L.F., Gollnick, D.M. (Eds.), *Wiley Handbook of Teaching and Learning*. Wiley-Blackwell, pp. 323–338.

- Resnick, L.B., Michaels, S., O'Connor, C., 2010. How (well-structured) talk builds the mind. In: Sternberg, R.J., Preiss, D.D. (Eds.), *Innovations in Educational Psychology: Perspectives on Learning, Teaching and Human Development*. Springer, New York, NY, pp. 163–194.
- Rommetsveit, R., 1992. Outlines of a dialogically based social-cognitive approach to human cognition and communication. In: Wold, A. (Ed.), *The Dialogical Alternative: Towards a Theory of Language and Mind*. Scandinavian Press, Oslo, Sweden, pp. 19–45.
- Rubio-Jimenez, A., Kershner, R., 2020. Promoting the self-determination of Mexican young adults identified with intellectual disability: a sociocultural discourse analysis of their discussion about goal setting. *Soc. Sci.* 9 (11), 200. <https://doi.org/10.3390/socsci9110200>.
- Rubio-Jimenez, A., Kershner, R., 2021. Transition to independent living: signs of self-determination in the discussions of Mexican students with intellectual disability. *Br. J. Learn. Disabil.* 49 (3), 352–364. <https://doi.org/10.1111/bld.12398>.
- Schwarz, B.B., Baker, M.J., 2016. *Dialogue, Argumentation and Education: History, Theory and Practice*. Cambridge University Press.
- Scott, P.H., Mortimer, E.F., Aguiar, O.G., 2007. The tension between authoritative and dialogic discourse: a fundamental characteristic of meaning making interactions in high school science lessons. *Sci. Educ.* 90 (4), 605–631. <https://doi.org/10.1002/sce.20131>.
- Sen, A., 2005. *The Argumentative Indian: Writings on Indian History, Culture and Identity*. Macmillan.
- Sidorkin, A.M., 1999. *Beyond Discourse: Education, the Self, and Dialogue*. SUNY Press, New York.
- Skidmore, D., 2010. From pedagogical dialogue to dialogical pedagogy. *Lang. Educ.* 14 (4), 283–296. <https://doi.org/10.1080/09500780008666794>.
- Snelgrove, S., 2005. Bad, mad and sad: developing a methodology of inclusion and a pedagogy for researching students with intellectual disabilities. *International J. Incl. Educ.* 9 (3), 313–329. <https://doi.org/10.1080/13603110500082236>.
- Teachman, G., McDonough, P., Macarthur, C., Gibson, B., 2018. A critical dialogical methodology for conducting research with disabled youth who use augmentative and alternative communication. *Qual. Inq.* 24 (1), 35–44. <https://doi.org/10.1177/1077800417727763>.
- Teo, P., 2019. Teaching for the 21st century: a case for dialogic pedagogy. *Learn. Cult. Soc. Interact.* 21 (March), 170–178. <https://doi.org/10.1016/j.lcsi.2019.03.009>.
- Trausan-Matu, S., Wegerif, R., Major, L., 2021. Dialogism. In: Cress, U., Wise, C., Rosé, A.F., Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning*. Computer-Supported Collaborative Learning Series. Springer Nature, pp. 219–239.
- United Nations, 2006. *Convention on the Rights of Persons With Disabilities*. Retrieved from: <https://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>.
- United Nations Educational, Scientific and Cultural Organization, UNESCO, 1994. *The Salamanca Statement and Framework for Action on Special Needs Education*. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000098427>.
- Vygotsky, L.S., 1978. *Mind in Society*. Harvard University Press, Cambridge, MA.
- Vygotsky, L.S., 1991. In: Light, P., Sheldon, S., Woodhead, B. (Eds.), *The Genesis of Higher Mental Functions*. Routledge, London.
- Webb, P., Whitlow, J.W., Venter, D., 2016. From exploratory talk to abstract reasoning: a case for far transfer? *Educ. Psychol. Rev.* 1–17.
- Wegerif, R., 2007. *Dialogic Education and Technology: Expanding the Space of Learning*. Springer, New York, NY.
- Wegerif, R., 2008. Dialogic or dialectic? The significance of ontological assumptions in research on educational dialogue. *Br. Educ. Res. J.* 34 (3), 347–361. <https://doi.org/10.1080/01411920701532228>.
- Wegerif, R., 2013. *Dialogic Education for the Internet Age*. Routledge, London.
- Wegerif, R., 2019. Dialogic education. In: *Oxford Research Encyclopaedia of Education*.
- Wegerif, R., 2020. Orientations and ground rules: a framework for researching educational dialogue. In: Kershner, R., Hennessy, S., Wegerif, R., Ahmed, A. (Eds.), *Research Methods for Educational Dialogue*. Bloomsbury Publishing (Chapter 4).
- Wegerif, R., Major, L., 2019. Buber, educational technology, and the expansion of dialogic space. *AI Soc.* 34 (1), 109–119.
- Wegerif, R., Mercer, N., Major, L., 2019. Introduction to the Routledge international handbook of research on dialogic education. In: *The Routledge International Handbook of Research on Dialogic Education*. Routledge, pp. 1–8.
- Wegerif, R., 2010. *Mind-Expanding: Teaching for Thinking and Creativity*. Open University Press/McGraw Hill, Buckingham, UK.
- Weiner, J., 2005. Peer-mediated conversational repair in students with moderate and severe disabilities. *Res. Pract. Pers. Severe Disabil.* 30 (1), 26–37. <https://doi.org/10.2511/rpsd.30.1.26>.
- Wells, G., 1999. *Dialogic Inquiry: Toward a Sociocultural Practice and Theory of Education*. Cambridge University Press, Cambridge.
- Yang, Y., 2016. Lessons learnt from contextualising a UK teaching thinking program in a conventional Chinese classroom. *Think. Skills Creativ.* 19, 198–209.

Exemplars

Alexander Renkl, University of Freiburg, Freiburg, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	612
Observational learning	612
Analogical reasoning	614
Learning from worked examples	615
Problem solving followed by instruction	617
Relations between theoretical perspectives	618
Instructional implications	618
Multiple examples	619
Inter-example features	619
Intra-example features	619
Learners' example processing	619
Outlook	620
References	620

Introduction

Exemplars are cases that represent a more general concept, procedure, or principle. This term is typically used when the case is a special one, either in the sense that it should be followed or copied (e.g., a very well performed procedure to solve a problem) or in the sense that it is a (proto-) typical or representative example. Sometimes exemplars are also special as they show learners that they have to attend to a specific aspect, for example, when solving a probability problem (e.g., order relevant or not).

At times, teachers introduce exemplars as these cases are (also) important in themselves and not (just) because they illustrate something more general. In literature instruction, for example, an outstanding piece by William Shakespeare, by Marcel Proust, by Miguel de Cervantes, or by Johann Wolfgang von Goethe may be read and discussed in schools, within the respective language areas. In history, a specific event might be part of the curriculum as it has heavily influenced the societal development in a country or even a continent (e.g., the French Revolution at the end of the 18th century). It is more typical, however, that cases are presented to enable learners to understand something more general, for example, the *concept* of mimicry in biology (e.g., [Brown and Kane, 1988](#)), the solution *procedure* for certain mechanics problems ([Chi et al., 1989](#)), or a *principle* that is helpful when students try to find a mathematical proof (e.g., [Reiss et al., 2008](#)). Moreover, exemplars can also be employed to show how concepts (e.g., motion or force), solution procedures (e.g., for certain types of mechanics problems), and principles (e.g., Newton's laws) interrelate in a specific domain (here: mechanics).

In this article, it is not possible to cover all the numerous instructionally relevant theoretical perspectives in which exemplars play an important role. We focus on four perspectives (a) that are supported by a substantial body of sound empirical research, (b) that are highly influential, and (c) that focus on student learning; readers interested in teacher education might refer to the literature on analyzing video-based classroom cases for the acquisition of professional vision or, as it has been recently often called, noticing skills (e.g., [Blomberg et al., 2014](#); [Gaudin and Chaliès, 2015](#); [Van Es and Sherin, 2021](#)).

In the next sections, we refer to the following approaches: (a) observational learning; (b) analogical reasoning; (c) example-based learning; (d) problem solving followed by instruction.

Observational learning

This approach is frequently called observational learning ([Bandura, 1971, 1986](#)). However, other labels are used as well: (Cognitive) modeling or social-cognitive theory. In this section, we primarily use the term observational learning. This means that a person learns by observing a model that, for example, solves a problem.

In his seminal book on modeling and observational learning, [Bandura \(1971\)](#) stated: "This volume is principally concerned with learning by example" (p. 1). He made the strong assumption that most behavior is learned by observing other people, that is, from example cases. Learning takes place if the learners pay attention to relevant behavior (*attention*), if they encode and remember the model activity (*retention*), and if they are capable of reproducing this type of behavior (*reproduction*); furthermore, the learners have to be motivated to reproduce the behavior (*motivation*) (e.g., [Bandura, 1986](#)).

Especially in academic learning, *abstract modeling* is relevant. This term means that a model demonstrates a behavior that follows an abstract rule or principle (e.g., of how to formulate well-substantiated arguments), in contrast to behavior that can be directly mimicked (e.g., how to make a specific knot in a rope). In the case of abstract modeling, the learners construct rules that can guide their future behavior ([Bandura, 1986](#)).

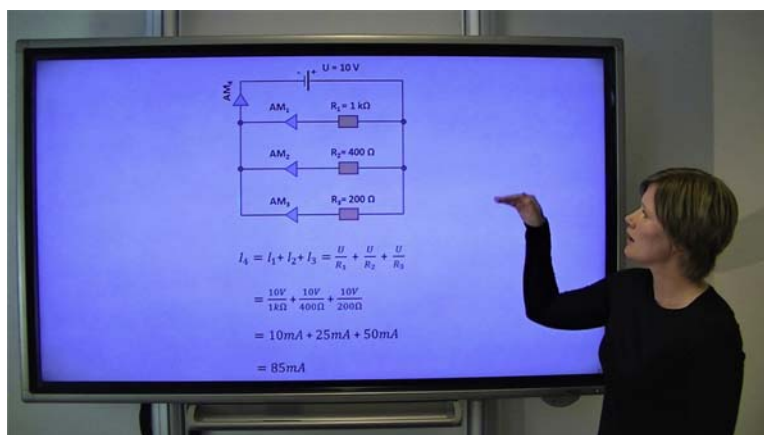


Fig. 1 A model showing how to solve a physics problem. Used by Hoogerheide et al. (2016), as shown in Van Gog et al. (2019).

In recent years, models in the internet (e.g., presented by YouTube or Khan Academy) have become an important source of learning. Fig. 1 shows a model that demonstrates how to solve an electrical circuits troubleshooting problem (Hoogerheide et al., 2016). The learners should learn abstract rules or principles to become able to solve similar new problems. Hence, this is a case of abstract modeling.

Bandura (1986) suggested providing multiple models (examples) to learners. Multiple examples make it not only easier to abstract the general rule or principle behind the models' behaviors; multiple models provide also a better chance for the learners to identify with at least one model and to learn effectively from it (e.g., Schunk et al., 1987).

Most research on observational learning in educational contexts did not try to prove the effectiveness of observational learning by comparing it to other forms of learning. Note that Bandura made the fundamental assumption that most behavior is learned by observing others. Instead, most studies analyzed the effects of different variants of observational learning, for example, learning from a smoothly performing mastery model or a coping model showing how to overcome impasses (e.g., Schunk and Hanson, 1985), learning by abstract modeling with(out) support to induce the underlying principles (e.g., Decker, 1980, 1984; Rummel et al., 2009), or learning from a model (dis-)similar to the observer (e.g., in terms of gender, age, or perceived competence; e.g., Braaksma et al., 2002); a similar model can lead to a higher self-efficacy of the learner and thereby better learning (e.g., Schunk et al., 1987).

Nevertheless, some research also tested the relative effectiveness of abstract modeling in comparison to other forms of learning. For example, it was found to be superior to foster students with problems in mathematics by cognitive modeling than by didactic instruction (e.g., Schunk, 1981); observational learning enhanced college students' writing skills more than working on practice problems (e.g., Zimmerman and Kitsantas, 2002); observing video models on productive cooperation was more effective for learning how to cooperate than learning by "scripted" doing (e.g., Rummel et al., 2009).

For developing self-regulation in academic domains such as writing, Schunk and Zimmerman (1997, 2007) have proposed a four-stage model with observational, emulative, self-controlled, and self-regulated stages. Over the development through these stages, regulation proceeds from social sources (observational, emulative) to self-regulation (self-controlled, self-regulated). At the observational stage, a learner profits from observing models that explain and demonstrate the execution of a skill, thereby learning the main features of a skill cognitively, that is, without necessarily performing the skill itself. At the emulative stage, a skill must be executed for further learning. The learner's skill execution approximates the model's performance. However, the learners do not fully copy the model but emulate the model's general pattern. As the learners need model cases for learning in the first two stages, these stages are considered as primarily social. At the self-controlled stage, the learner executes a skill independently when working on a problem, with a mental representation of the skill in mind. However, the skill is still structured after models. Learners have yet to develop a mental representation independent from the models. At the final self-regulated phase, learners are increasingly able to adapt a skill to changes in the problem-solving context. Note that the provision of example cases is pivotal in the first two stages (observational and emulative) to initiate skill acquisition. On this basis, self-regulated skill execution can developed in later stages.

Besides this social-cognitive model, there are also other instructionally relevant models that were influenced by the perspective of observational learning. For example, the well-known Self-Regulated Strategy Development (SRSD) approach by Harris and colleagues (e.g., Graham and Harris, 2018) was designed to foster students' writing strategies and skills; meanwhile SRSD has also been successfully applied in other subject areas such as mathematics or science (see McKeown et al., 2021). This instructional framework includes six stages, the third stage being "Model it." Interestingly, Dignath et al. (2008) found, in their meta-analysis of self-regulation training programs for primary school students, that interventions with a theoretical background in Bandura's perspective provide overall the best training effects (as compared to training interventions having their theoretical foundation in metacognitive or motivational theories).

In recent research, the concept of observational learning is also used to design and analyze forms of technology-enhanced learning. For example, a number of studies investigate how videos presenting models for academic learning (often called video modeling examples) could be optimized (Fig. 1). For example, research suggests that it is helpful when a model behavior is shown from a first-person perspective (i.e., video shows what the model sees) rather than a third-person perspective (i.e., video shows the model from an observer's point of view; Fiorella et al., 2017).¹ For other ways to improve such video models, the reader might refer to Mayer et al. (2020).

Another application in technology-enhanced learning is "Eye Movement Modeling Examples" (EMME), which show the eye-movements of an expert or a successful learner when working on a problem or when studying learning materials. For example, EMME can show the eye movements of an instructed model demonstrating effective processing strategies for multimedia learning materials with text and pictures. Such models can actually foster later learning from multimedia materials (e.g., Krebs et al., 2019; Mason et al., 2015).

In summary, a basic assumption of this perspective is that most behavior is learned from a specific form of exemplars, that is, models showing how to perform a task or solve a problem. Observational learning from models can also be successfully exploited in educational contexts, in which this form of learning is typically used for teaching strategies.

Analogical reasoning

Seminal research in this area was published in the 1980s as basic research on human reasoning (e.g., Gentner, 1983; Gick and Holyoak, 1980, 1983). In particular, it analyzed to what extent knowledge of exemplar cases can be exploited to solve analogous problems. Problems are analogous if they have the same structure with respect to their solution though the surface features (e.g., numbers, objects, or persons) may differ. For example, mathematical word problems are analogous when they can be solved by applying the same mathematical principle (e.g., addition rule in probability), even if the numbers, objects, or persons differ in the problem formulation. Another example is the Doppler effect which is usually known by laypersons as a shift in a siren's sound (i.e., a shift in wavelength) when an emergency vehicle is approaching us and then recedes again. By analogy, laypersons can also understand that similar changes in the wavelength of a star's light occur when it is approaching us or receding from the earth. More formally expressed, the superficial features such as numbers, objects, etc. can differ in analogous cases but the relations between these entities are the same (Gentner, 1983). To take up the two Doppler examples, they differ as to whether they refer to cars or stars or whether there are light waves or sound waves. However, the relation between the motion of the wave source (i.e., car or star) and the change in the wavelength corresponds in both cases. Cars and stars play the same role and light waves and sound waves play the same role; structurally, however, cars and stars as well as sound waves and light waves are exchangeable.

The process of analogical reasoning (and resulting learning) can be described by four stages (see Holyoak, 2012). Following is an explanation of what happens in the case of successful analogical reasoning. In stage 1, a student encounters a first example or, in some cases, also more analogous examples (e.g., probability word problem(s) to be solved by applying the addition rule).² This or these case(s) form the knowledge source for later solving analogous problems; for these reasons, the cases are often called *source* in this research area. In stage 2, a student encounters a new case (e.g., word problem to be solved by the addition rule), which is often called *target*. Memory traces of potentially relevant source cases are activated (e.g., word problems to be solved by the addition rule but maybe also word problems to be solved by the multiplication rule), and relevant cases (i.e., "addition rule cases") are selected. In stage 3, the new case (target) is mapped onto an analogous source case or even to several source cases. To take up again the Doppler examples, a student notices that, in the new case, the star plays the role of the car (mapping the star onto the car) and the light waves play the role of sound waves (mapping the light waves onto the sound waves). Most importantly, however, the identical relation between the wave source's motion and the shift in the wavelength is mapped one onto the other. In stage 4, mapping two or more examples to each other leads to the construction of a schema that represents relations and abstract role slots. In the ideal case, these abstract slots can be flexibly filled in by the superficial features of further new cases (e.g., Doppler echocardiogram in the medical domain).

Note that successful analogical reasoning leads not only to the solution of problems but also to substantial learning as schemata are constructed that can be applied to new cases. In addition, it was found that learners store both in their memory, the schema and the cases (Reeves and Weisberg, 1994; Ross, 1987, 1989.). Even if a schema has been constructed, the cases may play an important role in further problem solving and learning. It requires substantial expertise to "see" instantly that a schema applies when encountering a new case. A possibility for less expert learners is that a new case reminds them of an analogous case (e.g., Ross, 1987, 1989), and if the old case is mentally connected to an abstract schema, this schema can be applied to the new case. In a nutshell, example cases have several functions. They help to solve new problem cases, to construct abstract schemata, and they help to access these schemata by "reminders" (Ross, 1987, p. 629).

¹Of course, this can only hold true when the model's motoric behavior does not play an important role; often, such behavior cannot (fully) be seen from a first-person perspective.

²In basic analogical reasoning research as well as in educationally-oriented research, the rule or principle underlying the case solutions is sometimes already provided in stage 1. For the sake of simplicity, we do not consider this option here. This omission seems to be justified as it is the less typical option, and educationally it does not seem to be a preferable option (Alfieri et al., 2013).

Note, however, that the preceding description focused on successful cases of analogical reasoning. Yet, many students might not be able or willing to perform the required cognitive processes. Typical problems are as follows (see, e.g., [Gray and Holyoak, 2021](#); [Vendetti et al., 2015](#)): Students might have poor knowledge of the potential source cases (e.g., incomplete representations; stage 1). Another major problem is that many students do not spontaneously (i.e., without a hint) refer to a relevant source case (stage 2). Students might have also difficulties in mapping the source and target cases because they are “seduced” to map only superficial features (e.g., there may be children in two probability word problems but they play a different role so that they should not be mapped onto each other; stage 3). Students may also fail to infer the abstract structure behind analogous cases so that no (adequate) schema is constructed (stage 4). Overall, unsupported analogical reasoning is a demanding process imposing high cognitive load on the learners so that in particular learners with low prior knowledge frequently fail ([Gray and Holyoak, 2021](#)).

Against the background of these deficits, some reviews have provided guidelines on how to support students to make this learning method effective ([Alfieri et al., 2013](#); [Gray and Holyoak, 2021](#); [Vendetti et al., 2015](#)): For example, teachers may ensure that the students have a basic understanding of the source by explaining the source case and, thereby, highlighting its relevant features (stage 1). To help students select the right source(s) when trying to solve a new case, teachers might provide a hint to the source(s) or a source case and a case target can be presented side-by-side so that the source is visible when comparing the cases (stage 2). When students need to map cases, in particular with respect to structural correspondences, teacher can direct students’ attention to relations instead of to objects, persons, or numbers while comparing (stage 3). Finally, students are supported to form an abstract schema, if teachers provide explicitly the relational structure (i.e., principle) after comparing the examples (stage 4).

Overall, well-implemented learning from analogous cases is an effective method, as shown by the meta-analysis by [Alfieri et al. \(2013\)](#). This effectiveness was not only determined in controlled lab studies. For example, [Richland et al. \(2007\)](#) also found that US mathematics teachers used relational comparisons relatively often but they provided little help to their students such as keeping the source visible to learners during the comparison with the target or using gestures signaling important mappings. Hong Kong and Japanese teachers provided more support, which might explain some of the achievement differences between these countries in favor of the Asian region. The latter assumption is backed up by [Richland and McDonough \(2010\)](#) who found, in their classroom experiments, that providing visibly aligned source and target problems during initial learning, as well as using gestures to facilitate mappings, led to favorable learning outcomes.

In summary, this perspective emphasizes the role of exemplars stored in memory in three respects. First, stored source cases can help to solve new (target) problems; second, multiple cases form the basis for the induction of abstract schemata for later problem solving; third, when encountering a new problem, learners might be reminded of a previous case that helps accessing the relevant abstract problem solving schema. However, many learners need support in order to engage in successful analogical reasoning.

Learning from worked examples

Worked examples consist of a problem or task (e.g., forming a good counterargument to a claim) and its solution (e.g., a properly structured counter-argument). Worked examples with mathematical solution procedures usually also provide the solution steps leading to the final solution (e.g., mathematical word problem, equation(s), and the final solution). In some cases (see [Fig. 2](#)), worked examples are also “enriched” by elements designed to help learning, for example, explanations or color codes helping the students understand correspondences between a figure and an equation.

Seminal research on learning from worked examples (frequently also called example-based learning) was conducted in the 1980s by [Sweller and Cooper \(1985\)](#), as well as [Cooper and Sweller \(1987\)](#). They compared learning by solving problems with learning by studying worked examples. For example, students (9th grade) learned to solve algebraic equations ([Sweller and Cooper, 1985](#), Exp. 3). First, all students studied worked examples within the introduction to this topic. Following this, the students in the conventional problem-solving condition tried to solve eight problems covering four types of problems. The students in the worked-example condition encountered the same eight problems, but the solution steps of the first problem of each pair of structurally identical problems was shown (i.e., example-problem pairs). Students in the worked example condition, as compared to the conventional condition, made fewer errors in a final posttest with structurally identical problems and they needed less time for the acquisition phase well as for the posttest. [Cooper and Sweller \(1987\)](#) could also find positive effects of studying worked examples on the performance on later transfer (i.e., not structurally identical) problems. This advantage of studying worked examples—which was also found in many subsequent studies—is usually described as “worked example effect” (e.g., [Renkl, 2014](#); [Sweller et al., 2019](#); [Van Gog et al., 2019](#)).

To prevent a common misunderstanding, the use of one worked example between introducing a principle (e.g., addition rule in probability) and solving corresponding problems is not considered as example-based learning. Such a sequence was actually used in the conventional conditions (i.e., control conditions) in the seminal studies by [Sweller and Cooper \(1985\)](#), as well as [Cooper and Sweller \(1987\)](#). Example-based learning means that students study multiple examples before problem solving, either in the form of several example-problem pairs (as in [Sweller and Cooper, 1985](#)) or a sequence of examples only (see also [Van Harsel et al., 2019](#)).

The advantage of worked examples can be explained as follows ([Renkl, 2014](#)). If students have just started to learn, for example, about certain laws in physics (principles), they have developed just a rudimentary understanding of these laws. Hence, these students cannot rely on a real understanding of these laws when they work on corresponding problems. As consequence, the students frequently engage in superficial strategies to, nonetheless, find a solution (e.g., copying the solution of

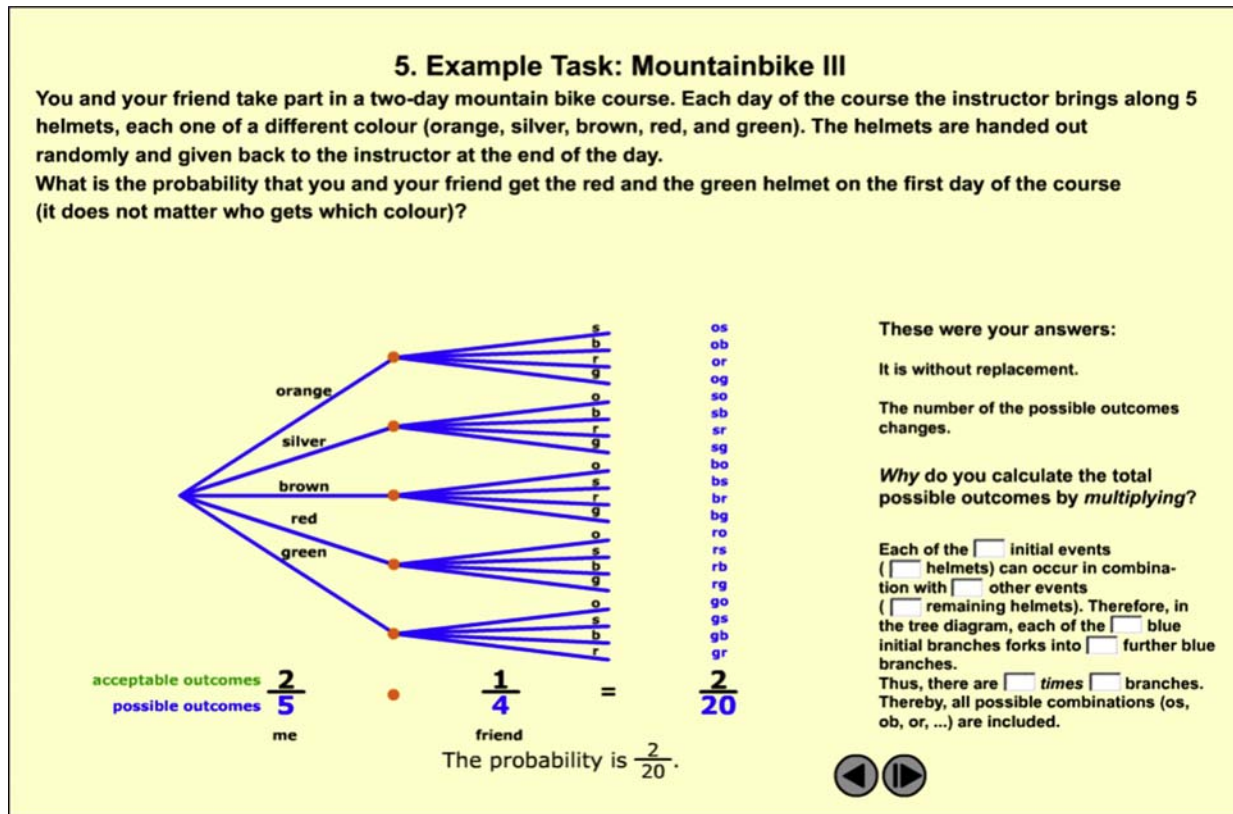


Fig. 2 A worked example from probability, with color coding and to be supplemented explanations. Taken from Berthold and Renkl (2009).

a former problem and adapting the numbers). Such superficial strategies, in many cases, may lead to the correct numerical solution but they barely foster the understanding of the physics laws and of how to apply them for solving problems. Thus, solving problems by relying on superficial strategies can be viewed as extraneous (not productive) cognitive load (Sweller et al., 2019). Hence, learners should first study worked examples to reach an understanding of underlying principles (e.g., physics laws, mathematical theorems) before they engage in solving problems. With a sufficient understanding of the principles, problems can then be solved in a principle-based way, which deepens understanding and may eventually also lead to automation (Renkl, 2014). Against this background, learning from worked examples is helpful for the initial acquisition of cognitive skills (e.g., learning to solve a certain type of physics problems) when the main goal is to gain an understanding of principles and their application. In later stages of skill acquisition, when the goals are fine-tuning a skill and, in part, to automate it, problem solving is superior (e.g., Kalyuga et al., 2001; Renkl, 2014).

Initially, the worked example effect was primarily shown in mathematics and in domains with mathematical solution procedures (e.g., in physics). Such learning domains are usually classified as well-structured. However, the worked example effect has also been found in many ill-structured domains where just heuristics ("rules of thumb") apply, instead of algorithmic solutions. Examples of such ill-structured domains in which worked examples effectively fostered learning are finding mathematical proofs (Reiss et al., 2008), cooperating effectively (Rummel et al., 2009), identifying styles of designers (Rourke and Sweller, 2009), effective concept mapping (Hilbert and Renkl, 2009), analyzing legal cases (Nievelstein et al., 2013), arguing properly (Hefter et al., 2014), and writing essays in English literature (Kyun et al., 2013); many more examples could be provided.

As it is the case with other learning methods (Pant, 2014; Renkl, 2015), learning from worked examples is not always helpful. Three important boundary conditions are as follows. First, as already mentioned when providing the theoretical explanation of the worked example effect, worked examples are for initial skill acquisition. Providing such examples to learners with good prior knowledge of the learning content is redundant and provides no advantage or even leads to less learning than solving problems (e.g., Kalyuga et al., 2001). Second, many learners study worked examples just superficially, for example, without referring to the underlying principle, which can lead to little learning (e.g., Hilbert and Renkl, 2009; Renkl, 1997). Third, studying worked examples induces less extraneous load than solving problems, which is a major reason for the worked example effect (see the theoretical explanation for the worked example effect). However, when extraneous load is increased by having worked examples with multiple representations (e.g., diagram and equation) that are difficult to map one to the other, the worked example effect disappears (e.g., Tarmizi and Sweller, 1988).

Appropriate instructional design can help to counteract these problems. First, to avoid the effect of redundant examples, the solution steps of the examples can be gradually faded out with increasing learner knowledge. After the initial study of fully worked examples, learners have to determine single solution steps and later they solve the whole problem on their own. Fading provides the learners with a smooth transition from studying worked steps to problem solving (Renkl and Atkinson, 2003). As the learners differ in their speed of acquiring knowledge, it is best to fade worked steps adaptively according to the individual increase in understanding, which can be accomplished by technology-enhanced learning environments (e.g., Salden et al., 2010). Second, to counteract superficial processing, the learners can be prompted to explain to themselves what principle(s) (e.g., physic laws) underlie an example's solution. Self-explaining fosters learning outcomes (self-explanation effect; e.g., Chi et al., 1989; Renkl and Eitel, 2019; Rittle-Johnson et al., 2017). Sometimes in example-based learning, several examples are presented simultaneously, and self-explanation prompts require the learners to compare these examples and their solution procedures (e.g., Schalk et al., 2020). Such an instructional procedure could also be regarded as an analogical-reasoning learning procedure; in this case, the boundaries between the two approaches becomes blurred. Third, to reduce extraneous load, the mapping between different representations can be eased for the learners (e.g., Renkl, 2022), for example, by color coding (i.e., using the same color for elements corresponding across different representation; Berthold and Renkl, 2009, Fig. 2) or by an integrated format (e.g., writing the equations parts to corresponding locations in the diagram; Tarmizi and Sweller, 1988). In the case of video-based worked examples, the presented information often quickly disappears and is replaced by other new information (transient information effect; Singh et al., 2012). Learners profit if the video stops after smaller meaningful event units (segmenting; Spanjers et al., 2011, 2012). For further instructional designs to optimize example-based learning, the reader might refer to Renkl (2014, 2017, 2022).

In summary, it is a very robust finding that learning from worked examples is an effective option to foster the initial acquisition of cognitive skills, when the main goal is to gain an understanding of principles and their application. This is true not only in domains with mathematical solutions but also in domains where just heuristics apply. However, to fully exploit the potential of example-based learning, appropriate instructional design has to be used.

Problem solving followed by instruction

The approach of problem solving followed by instruction—which means in this case direct forms of instruction with explanations of the to-be-learned contents—is typically abbreviated as PS-I (e.g., Loibl et al., 2017). Corresponding research usually tries to show that (under certain circumstances) PS-I is superior to the more traditional sequence of direct instruction followed by a phase of problem solving, in which students apply what has been explained before (I-PS). It is argued that it is too difficult for (many) students to get to an understanding of the principle(s) to be learned when beginning with direct instruction. The problem solving phase should prepare the students to profit optimally from the following direct instruction so that they can understand the principle(s) (e.g., Kapur, 2015). At first glance, this approach does not emphasize the role of exemplars during learning. However, when looking at how the two typical PS-I models, “productive failure” (e.g., Kapur, 2008, 2015) and “inventing” (e.g., Schwartz et al., 2011), arrange the PS phase, it is clear that specifically designed sets of example cases play a crucial role.

In productive failure (e.g., Kapur, 2008, 2015), students get a problem related to a to-be-learned concept (e.g., variance/standard deviation) that they cannot solve. For example, students (here: in triads) worked on the following problem without support or guidance (Kapur, 2012, 2015): A committee has to decide which one of three soccer strikers should be awarded with the prize for consistent performance. The students (9th grade) get a table with the number of the strikers' goals in the years from 1992 to 2011: Mike Arwen: 14, 9, 14, 10, 15, 11, 15, 11, 16, 12, 16, 12, 17, 13, 17, 13, 18, 14, 19, 14; Dave Backhand: 13, 9, 16, 14, 10, 11, 13, 14, 15, 19, 14, 12, 15, 14, 17, 13, 14, 18, 14, 15; Ivan Right: 13, 18, 15, 10, 16, 10, 17, 10, 12, 14, 19, 14, 18, 9, 10, 18, 11, 10, 18, 18. The committee wants to decide on the basis of a mathematical index for consistency. The specific instructions students received were as follows: “(1) Design as many different measures of consistency as you can. (2) Your measure of consistency should make use of all data points in the table. (3) Show all working and calculations on the papers provided” (Kapur, 2015, Appendix A). In the following phase, the teacher compared the different student-generated solutions and then introduced the canonical solution; the students could, thus, compare their wrong or suboptimal solutions to the canonical solution. Finally, the students worked on three practice problems, and the solutions were discussed in the class.

Note that the set of three cases was specifically designed (e.g., all three strikers had the same range and mean) and played a substantial role in the first PS phase. The cases had to be compared, like in analogical reasoning arrangements. Different from analogical reasoning, however, there is no general principle that is intrinsic to the (set of) cases (e.g., in other contexts, the cases could also be compared with respect to the average number of goals per season), and the students are not expected to infer a principle correctly. The prime intended function of this problem solving phase is that the students activate and differentiate their prior knowledge in relation to the target concept (here: variance/standard deviation).

The approach of inventing (e.g., Schwartz et al., 2011; Schwartz and Martin, 2004) is very similar. However, Loibl et al. (2017) identified two important differences. First, in inventing, the teacher might not take up the previous student solutions before providing instruction about the canonical solution (as it is common in productive failure). Second, “contrasting cases” are used. This means that there is a set of simpler cases, as compared to productive failure, and the cases are very similar to each other; they ideally differ just in one structurally relevant aspect. Thereby, the students' attention should be guided to this aspect. Loibl et al. (2017) illustrated how the already provided materials of Kapur (2012, 2015) might look in an invention study: “Contrasting cases #1: ● Player A: 9, 10, 10, 11 ● Player B: 5, 10, 10, 15/Contrasting cases #2: ● Player B: 5, 10, 10, 15 ● Player C: 5, 5, 15, 15” (p. 701; see also Loibl et al., 2020). Note

that this different design of the example sets makes also a difference with respect to whether the students can determine on their own to what extent their inventions work. By their simple and overlapping design, it is very easy to see that player A is more consistent than player B, and player B is more consistent than player C. Hence, it is not hard for the students to check whether their invented indices discern the cases in the appropriate way; they are less dependent on external feedback in this respect.

Loibl et al. (2017) concluded in their review that PS-I, as compared to I-PS, does not reliably foster the performance on problems analogous to the ones encountered during learning. However, there are positive effects on conceptual understanding and on transfer if the instructional procedure includes contrasting cases and/or taking up the students' solution in the instructional phase. If both are missing, no positive effects can be expected. In addition, their analyses suggest that the following set of mechanisms is responsible for the positive effects of an initial problem solving phase: (a) The students activate their prior knowledge; (b) they become aware of their knowledge gaps; (c) the students' attention is directed to structural features related to the underlying principle, which can, in particular, happen when working on contrasting cases or when student solutions are compared to the canonical solutions. This is why contrasting cases and/or taking up the students' solution in the instructional phase is important. In addition, the findings of Glogger-Frey et al. (2017) suggest that experience in inventing activities helps to make these activities productive. Note that the effects of PS-I do not seem to depend on students' success in the problem-solving phase; even failure prepares for further learning (e.g., Loibl and Rummel, 2014).

Loibl et al. (2017) as well as the discussion in the preceding paragraph focused on cognitive mechanisms. However, motivational mechanisms may also account for the positive effects of initial problem solving. Such an activity can lead to more epistemic curiosity about and situational interest in the learning content (e.g., Glogger-Frey et al., 2015). To date, however, there is no clear picture on the interplay between cognitive and motivational factors in PS-I learning.

In summary, PS-I requires the students to inspect specifically designed sets of example cases and to invent an index for comparing these cases. These invented indices, which are, in most cases, (partly) incorrect, are then related to the target concept (i.e., the correct index). Even without finding the correct index on their own (i.e., "failure"), students can profit from such problem solving activities as they prepare them to learn from subsequent direct instruction.

Relations between theoretical perspectives

There are cases in which different strands of research investigate very similar phenomena but hardly take notice of each other. Fortunately, this is not true in the case of the four perspectives just described. In many studies, researchers refer to more than one of these perspectives. This multiple referencing is usually not done to test competing hypotheses from the different perspectives; instead they are primarily exploited as complementary perspectives. It is beyond the scope of this article to review the findings of research relying on such multiple perspectives, but a "flavor" of such research should be provided by the following.

The research area that is most readily connected to the other areas is example-based learning. To illustrate these interconnections, we refer to selected studies. There are studies using worked examples in form of videos that show persons demonstrating how, for example, a mathematical problem is solved (e.g., on YouTube). Such worked examples directly call for taking factors into account that are usually considered in observational learning such as model-observer similarity or the models' perceived expertise (e.g., Hoojerheide et al., 2016, 2018). In other studies, there are several related principles that should be taught by example-based learning (probability problems with and without replacement; e.g., Berthold and Renkl, 2009). Then, sets of examples are typically provided that also help differentiate the different related problem types (e.g., with and without replacement). For this purpose, students are required to compare example cases. In such circumstances, it is sensible to also refer to analogical reasoning research for tried-and-tested support measures to support example comparison and, thereby, to optimize learning (e.g., presenting examples side-by-side; Schalk et al., 2020). Finally, there are studies testing whether not only the typical example-based arrangements (i.e., example-problem pairs or examples-only) lead to effective learning but also problem-example pairs, which share some similarities with the PS-I approach (e.g., Van Harsel et al., 2019).

Studies connecting the other perspectives are not so numerous but they can also be found. For example, Jacobson et al. (2020) combined ideas from productive failure and analogical reasoning to foster across-domain transfer with respect to applying complex systems principles. Hartmann et al. (2020) found that PS-I can be effective not only when learners "fail" themselves but also when they observe a model with failed problem-solving attempts prior to instruction. Finally, Wu and Weng (2013) found that models displaying analogical reasoning can help teach architecture design.

In summary, the four perspectives on exemplars discussed in this article can be seen as complementary. Against this background, largely shared instructional implications can be derived from this set of perspectives.

Instructional implications

It is beyond the scope of this article to discuss instructional design issues for learning from exemplars in detail (for this purpose see, e.g., Loibl et al., 2017; Gray and Holyoak, 2021; Renkl, 2014, 2022; Schunk and Zimmerman, 1997; Sinha and Kapur, 2021; Van Gog et al., 2019; Vendetti et al., 2015). However, we present selected instructional guidelines that relate to the theoretical perspectives discussed in this article. They refer to the use of the multiple examples, to how sets of examples might be composed (inter-example features), to how single examples should be designed (intra-example features; Atkinson et al., 2000), and to how the

learner's processing of the examples can be fostered. To facilitate the mapping between the previously discussed perspectives and the following practice recommendations, the following abbreviations will be used: OL (observational learning), AR (analogical reasoning), WE (worked examples), and PF-I (problem-solving before instruction).

Multiple examples

An obvious common recommendation of the four approaches is to use multiple example cases. In part the assumed advantages of multiple examples differ. However, this does not mean that the specific advantages emphasized in one perspective do not also apply to learning arrangements designed (primarily) from another perspective. More specific recommendations with respect to how to structure sets of multiple examples are provided in the subsequent sub-sections.

Inter-example features

Recommendable features of example sets (i.e., inter-example features) depend on the specific learning goal and the specific learning setting:

- (a) If there is a principle and its application that should be understood, it is recommendable to provide a set of analogous example cases (WE, AR). Two variants have to be differentiated in this context: the principle is or is not presented up-front. If the principle is presented up-front (typical of WE), subsequently studying multiple analogous examples (or example-problem pairs) can foster the learners' understanding of the principle and its application. Thereby, the learners get well prepared for learning by solving further related problems in a principle-oriented way. If the principle is not presented up-front (as typical of AR and obligatory in PS-I), the learners get the opportunity to (try to) infer the principle by "seeing" through the superficial features on the relational structure (i.e., principle; AR). In the case of productive failure arrangements, it is not expected that the learners will correctly induce the "full" principle but they are prepared for learning the principle by being sensitized for structural features (PS-I).
- (b) If the goal is to discern easily mixed-up problem categories (e.g., rule of three in mathematics for direct proportions or for inverse proportions), the learner should work on a set of examples that provide the opportunity to compare cases for the different categories (e.g., providing them side-by-side); thereby students learn to discern problems related to the different categories (WE).
- (c) If models are the cases to learn from (OL, WE), multiple examples have the advantage that different learners can find at least one model to identify with, which is a favorable condition for learning (OL).

Intra-example features

All perspectives emphasize that the learners' attention should be directed to structural aspects related to the underlying principles. Unnecessary processing demands and resulting extraneous working memory load, which can also lead to cognitive overload, should be avoided (WE, AR). Two important possibilities to avoid extraneous load are as follows.

- (a) As already mentioned in the section of example-based learning, it is essential in the case of worked examples with multiple representations (e.g., text, graphic, or equations) to ease mapping between these presentations, for example, by color coding or integrated format (e.g., Renkl, 2022; see Fig. 1). It is very likely that this finding generalizes also to the other types of learning arrangements (OL, AR, PS-I), as research on multimedia learning strongly suggests (see, e.g., Mayer, 2021).
- (b) In the case of video-based worked examples, models, or source analogs, which can be often found on the internet, the transience of information is a challenge for students that should not be underestimated (e.g., Singh et al., 2012). If the learners are briefly inattentive they might miss important information. Such inattentiveness can be due to mind wandering (e.g., Risko et al., 2012, 2013) or attention to salient, but not important information in the video (e.g., a case of unusual wording in the explanatory narration or the accent of a model). It might also be that a learner thinks a bit longer about an important to-be-learned aspect so that new information is missed. To mitigate this problem, segmenting can be used, that is, the videos stop after each meaningful event unit (e.g., Spanjers et al., 2011).

Learners' example processing

A well-known problem is that many learners do not spontaneously engage in self-explaining models or worked examples (OL, WE) or in comparing example cases (AR, PS-I). There are several possibilities to address such problems.

A tried-and-tested option is to use prompts that ask the learners to explain to the themselves the principles underlying a model (OL) or a worked example (WE). Prompts can also be used to enhance learners' processes of example comparison (AR, WE). Note, however, that such support may not have an added value in the case of PS-I where it is not crucial that the comparison processes lead to the induction a principle (see Loibl and Rummel, 2014). In this case, it seems more important that the student solutions and the canonical solution are compared in an instructional phase.

Another tried-and-tested option in the case of worked examples or examples to be compared is to use short training interventions (AR, WE). Such an intervention to foster the processing of worked examples was, for example, developed and tested by Renkl et al. (1998). This intervention (10–15 min) included the following components: (a) information on the importance of self-explanations, (b) modeling self-explanations with one worked example (here: from interest calculation), and (c) coached practice with another example (for a similar training intervention see Hodds et al., 2014). Gentner et al. (2003) has developed a successful short training intervention for learning by example comparison (AR).

Another option leading to enhanced learning is imagery. For example, an observed model behavior can be mentally repeated (e.g., Eaves et al., 2016). Similarly, it fosters learning when the students close the eyes after having studied a worked example and then imagine the solution procedure (e.g., Ginns et al., 2003).

Outlook

Overall, it is fair to say that the topic of this article is relatively well-researched. This is at least true with respect to the functions of exemplars for learning, to the learning processes, to the effects of corresponding learning arrangements, and to instructional design options for optimizing learning. Nevertheless, there are important issues for future research.

Although the cognitive aspects are well researched, there is relatively little research on motivational processes and factors as well as their interplay with cognitive factors. This is especially true for research on example-based learning, analogical reasoning, and PS-I. In observational learning, at least the motivational concept of self-efficacy has received substantial attention (e.g., Braaksma et al., 2018; Schunk, 1981; Schunk and Hanson, 1985). Future research on motivational issues and on the interplay of motivation and cognition could provide very helpful information for implementing effective forms of learning from exemplars in classrooms.

Up to date, there are some studies, in which forms of learning from exemplars were implemented in classrooms (e.g., Reiss et al., 2008). However, such studies are rare. This is more than pity, as there is an ample research base (at least on cognitive issues) that allowed for promising implementations. Whether the last statement is just a “pious wish” of a researcher in this area—the author of this article—or an “accurate expert judgment” should be investigated in future implementations and design-based studies.

References

- Alfieri, L., Nokes-Malach, T.J., Schunn, C.D., 2013. Learning through case comparisons: a meta-analytic review. *Educ. Psychol.* 48, 87–113.
- Atkinson, R.K., Derry, S.J., Renkl, A., Wortham, D.W., 2000. Learning from examples: instructional principles from the worked examples research. *Rev. Educ. Res.* 70, 181–214.
- Bandura, A., 1971. *Psychological Modeling: Conflicting Theories*. Aldine Atherton, Chicago, IL.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, NJ.
- Berthold, K., Renkl, A., 2009. Instructional aids to support a conceptual understanding of multiple representations. *J. Educ. Psychol.* 101, 70–87.
- Blomberg, G., Sherin, M.G., Renkl, A., Gloger, I., Seidel, T., 2014. Understanding video as a tool for teacher education: investigating instructional strategies to promote reflection. *Instr. Sci.* 42, 443–463.
- Braaksma, M.A.H., Rijlaarsdam, G., van den Bergh, H., 2002. Observational learning and the effects of model-observer similarity. *J. Educ. Psychol.* 94, 405–415.
- Braaksma, M., Rijlaarsdam, G.C.W., van den Bergh, H.H., 2018. Effects of hypertext writing and observational learning on content knowledge acquisition, self-efficacy, and text quality: two experimental studies exploring aptitude treatment interactions. *J. Writ. Res.* 9, 259–300.
- Brown, A.L., Kane, M.J., 1988. Preschool children can learn to transfer: learning to learn and learning from example. *Cognit. Psychol.* 20, 493–523.
- Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., Glaser, R., 1989. Self-explanations: how students study and use examples in learning to solve problems. *Cognit. Sci.* 13, 145–182.
- Cooper, G., Sweller, J., 1987. Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *J. Educ. Psychol.* 79, 347–362.
- Decker, P.J., 1980. Effects of symbolic coding and rehearsal in behavior-modeling training. *J. Appl. Psychol.* 65, 627–634.
- Decker, P.J., 1984. Effects of different symbolic coding stimuli in behavior modeling training. *Person. Psychol.* 37, 711–720.
- Dignath, C., Buettner, G., Langfeldt, H.P., 2008. How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programs. *Educ. Res. Rev.* 3, 101–129.
- Eaves, D.L., Riach, M., Holmes, P.S., Wright, D.J., 2016. Motor imagery during action observation: a brief review of evidence, theory, and future research opportunities. *Front. Neurosci.* 10, 514.
- Fiorella, L., van Gog, T., Hoogerheide, V., Mayer, R.E., 2017. It's all a matter of perspective: viewing first-person video modeling examples promotes learning of an assembly task. *J. Educ. Psychol.* 109, 653.
- Gaudin, C., Chaliès, S., 2015. Video viewing in teacher education and professional development: a literature review. *Educ. Res. Rev.* 16, 41–67.
- Gentner, D., Loewenstein, J., Thompson, L., 2003. Learning and transfer: a general role for analogical encoding. *J. Educ. Psychol.* 95, 393–408.
- Gentner, D., 1983. Structure-mapping: a theoretical framework for analogy. *Cognit. Sci.* 7, 155–170.
- Gick, M.L., Holyoak, K.J., 1980. Analogical problem solving. *Cognit. Psychol.* 12, 306–355.
- Gick, M.L., Holyoak, K.J., 1983. Schema induction and analogical transfer. *Cognit. Psychol.* 15, 1–38.
- Ginns, P., Chandler, P., Sweller, J., 2003. When imagining information is effective. *Contemp. Educ. Psychol.* 28, 229–251.
- Gloger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., Renkl, A., 2015. Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learn. Instruct.* 39, 72–87.
- Gloger-Frey, I., Gaus, K., Renkl, A., 2017. Learning from direct instruction: best prepared by several self-regulated or guided invention activities? *Learn. Instruct.* 51, 26–35.
- Graham, S., Harris, K., 2018. An examination of the design principles underlying a Self-Regulated Strategy Development study. *J. Writ. Res.* 10, 139–187.
- Gray, M.E., Holyoak, K.J., 2021. Teaching by analogy: from theory to practice. *Mind Brain Educ.* 15, 250–263.
- Hartmann, C., van Gog, T., Rummel, N., 2020. Do examples of failure effectively prepare students for learning from subsequent instruction? *Appl. Cognit. Psychol.* 34, 879–889.

- Heffter, M.H., Berthold, K., Renkl, A., Riess, W., Schmid, S., Fries, S., 2014. Effects of a training intervention to foster argumentation skills while processing conflicting scientific positions. *Instr. Sci.* 42, 929–947.
- Hilbert, T.S., Renkl, A., 2009. Learning how to use a computer-based concept-mapping tool: self-explaining examples helps. *Comput. Hum. Behav.* 25, 267–274.
- Hodds, M., Alcock, L., Inglis, M., 2014. Self-explanation training improves proof comprehension. *J. Res. Math. Educ.* 45, 62–101.
- Holyoak, K.J., 2012. Analogy and relational reasoning. In: Holyoak, K.J., Morrison, R.G. (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, pp. 234–259.
- Hoogerheide, V., van Wermeskerken, M., Loyens, S.M.M., van Gog, T., 2016. Learning from video modeling examples: content kept equal, adults are more effective models than peers. *Learn. Instruct.* 44, 22–30.
- Hoogerheide, V., van Wermeskerken, M., van Nassau, H., van Gog, T., 2018. Model-observer similarity and task-appropriateness in learning from video modeling examples: do model and student gender affect test performance, self-efficacy, and perceived competence? *Comput. Hum. Behav.* 89, 457–464.
- Jacobson, M.J., Goldwater, M., Markauskaite, L., Lai, P.K., Kapur, M., Roberts, G., Hilton, C., 2020. Schema abstraction with productive failure and analogical comparison: learning designs for far across domain transfer. *Learn. Instruct.* 65, 101222.
- Kalyuga, S., Chandler, P., Tuovinen, J., Sweller, J., 2001. When problem solving is superior to studying worked examples. *J. Educ. Psychol.* 93, 579–588.
- Kapur, M., 2008. Productive failure. *Cognit. Instruct.* 26, 379–424.
- Kapur, M., 2012. Productive failure in learning the concept of variance. *Instr. Sci.* 40, 651–672.
- Kapur, M., 2015. Learning from productive failure. *Learn. Res. Pract.* 1, 51–65.
- Krebs, M.-C., Schüler, A., Scheiter, K., 2019. Just follow my eyes: the influence of model-observer similarity on eye movement modeling examples. *Learn. Instruct.* 61, 126–137.
- Kyun, S., Kalyuga, S., Sweller, J., 2013. The effect of worked examples when learning to write essays in English literature. *J. Exp. Educ.* 81, 385–408.
- Loibl, K., Rummel, N., 2014. The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instr. Sci.* 42, 305–326.
- Loibl, K., Roll, I., Rummel, N., 2017. Towards a theory of when and how problem solving followed by instruction supports learning. *Educ. Psychol. Rev.* 29, 693–715.
- Loibl, K., Tillema, M., Rummel, N., van Gog, T., 2020. The effect of contrasting cases during problem solving prior to and after instruction. *Instr. Sci.* 48, 115–136.
- Mason, L., Pluchino, P., Tornatora, M.C., 2015. Eye-movement modeling of integrative reading of an illustrated text: effects on processing and learning. *Contemp. Educ. Psychol.* 41, 172–187.
- Mayer, R.E., Fiorella, L., Stull, A., 2020. Five ways to increase the effectiveness of instructional video. *Educ. Technol. Res. Dev.* 68, 837–852.
- Mayer, R.E., 2021. *Multimedia Learning*, third ed. Cambridge University Press, Cambridge, UK.
- McKeown, D., FitzPatrick, E., Ennis, R., Sanders, S., 2021. Self-Regulated Strategy Development: a framework for effective instruction across the content areas. *Learn. Disabil. Res. Pract.* <https://doi.org/10.1111/ldrp.12256>.
- Nievelstein, F., van Gog, T., van Dijk, G., Boshuizen, H.P., 2013. The worked example and expertise reversal effect in less structured tasks: learning to reason about legal cases. *Contemp. Educ. Psychol.* 38, 118–125.
- Pant, H.A., 2014. Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung [Synthesizing evidence for educational policy and educational decisions: meta-analyses in educational sciences]. *Z. für Erziehungswiss.* 17, 79–99.
- Reeves, L.M., Weisberg, R.W., 1994. The role of content and abstract information in analogical transfer. *Psychol. Bull.* 115, 381–400.
- Reiss, K., Heinze, A., Renkl, A., Groß, C., 2008. Reasoning and proof in geometry: effects of a learning environment based on heuristic worked-out examples. *ZDM* 40, 455–467.
- Renkl, A., Atkinson, R.K., 2003. Structuring the transition from example study to problem solving in cognitive skills acquisition: a cognitive load perspective. *Educ. Psychol.* 38, 15–22.
- Renkl, A., Eitel, A., 2019. Self-explaining: learning about principles and their application. In: Dunlosky, J., Rawson, K. (Eds.), *Cambridge Handbook of Cognition and Education*. Cambridge University Press, Cambridge, UK, pp. 528–549.
- Renkl, A., Stark, R., Gruber, H., Mandl, H., 1998. Learning from worked-out examples: the effects of example variability and elicited self-explanations. *Contemp. Educ. Psychol.* 23, 90–108.
- Renkl, A., 1997. Learning from worked-out examples: a study on individual differences. *Cognit. Sci.* 21, 1–29.
- Renkl, A., 2014. Towards an instructionally-oriented theory of example-based learning. *Cognit. Sci.* 38, 1–37.
- Renkl, A., 2015. Different roads lead to Rome: the case of principle-based cognitive skills. *Learning* 1, 79–90.
- Renkl, A., 2017. Learning from worked examples in mathematics: students relate procedures to principles. *ZDM* 49, 571–584.
- Renkl, A., 2022. The worked example principle in multimedia learning. In: Mayer, R.E., Fiorella, L. (Eds.), *The Cambridge Handbook on Multimedia Learning*, third revised ed. Cambridge University Press, Cambridge, UK.
- Richland, L.E., McDonough, I.M., 2010. Learning by analogy: discriminating between potential analogs. *Contemp. Educ. Psychol.* 35, 28–43.
- Richland, L.E., Zur, O., Holyoak, K.J., 2007. Cognitive supports for analogies in the mathematics classroom. *Science* 316 (5828), 1128–1129.
- Risko, E.F., Anderson, N., Sarwal, A., Engelhardt, M., Kingstone, A., 2012. Everyday attention: variation in mind wandering and memory in a lecture. *Appl. Cognit. Psychol.* 26, 234–242.
- Risko, E.F., Buchanan, D., Medimorec, S., Kingstone, A., 2013. Everyday attention: mind wandering and computer use during lectures. *Comput. Educ.* 68, 275–283.
- Rittle-Johnson, B., Loehr, A.M., Durkin, K., 2017. Promoting self-explanation to improve mathematics learning: a meta-analysis and instructional design principles. *ZDM* 49, 599–611.
- Ross, B.H., 1987. This is like that: the use of earlier problems and the separation of similarity effects. *J. Exp. Psychol. Learn. Mem. Cognit.* 13, 629–639.
- Ross, B.H., 1989. Distinguishing types of superficial similarities: different effects on the access and use of earlier problems. *J. Exp. Psychol. Learn. Mem. Cognit.* 15, 456–468.
- Rourke, A.J., Sweller, J., 2009. The worked-example effect using ill-defined problems: learning to recognize designers' styles. *Learn. Instruct.* 19, 185–199.
- Rummel, N., Spada, H., Hauser, S., 2009. Learning to collaborate while being scripted or by observing a model. *Int. J. Comput. Support. Collab. Learn.* 4, 69–92.
- Salden, R., Alevyn, V., Schwonke, R., Renkl, A., 2010. The expertise-reversal effect and worked examples in tutored problem solving. *Instr. Sci.* 38, 289–307.
- Schalk, L., Roelle, J., Saalbach, H., Berthold, K., Stern, E., Renkl, A., 2020. Providing worked examples for learning multiple principles. *Appl. Cognit. Psychol.* 34, 813–824.
- Schunk, D.H., Hanson, A.R., 1985. Peer models: influence on children's self-efficacy and achievement. *J. Educ. Psychol.* 77, 313–322.
- Schunk, D.H., Zimmerman, B.J., 1997. Social origins of self-regulatory competence. *Educ. Psychol.* 32, 195–208.
- Schunk, D.H., Zimmerman, B.J., 2007. Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Read. Writ. Q.* 23, 7–25.
- Schunk, D.H., Hanson, A.R., Cox, P.D., 1987. Peer model attributes and children's achievement behaviors. *J. Educ. Psychol.* 79, 54–61.
- Schunk, D.H., 1981. Modeling and attributional effects on children's achievement: a self-efficacy analysis. *J. Educ. Psychol.* 73, 93–105.
- Schwartz, D.L., Martin, T., 2004. Inventing to prepare for future learning: the hidden efficiency of encouraging original student production in statistics instruction. *Cognit. Instruct.* 22, 129–184.
- Schwartz, D.L., Chase, C.C., Oppezzo, M.A., Chin, D.B., 2011. Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *J. Educ. Psychol.* 103, 759–775.
- Singh, A.M., Marcus, N., Ayres, P., 2012. The transient information effect: investigating the impact of segmentation on spoken and written text. *Appl. Cognit. Psychol.* 26, 848–853.
- Sinha, T., Kapur, M., 2021. When problem solving followed by instruction works: evidence for productive failure. *Rev. Educ. Res.* 91, 761–798.
- Spanjers, I.A., Wouters, P., van Gog, T., van Merriënboer, J.J., 2011. An expertise reversal effect of segmentation in learning from animated worked-out examples. *Comput. Hum. Behav.* 27, 46–52.
- Spanjers, I.A.E., van Gog, T., van Merriënboer, J.J.G., 2012. Segmentation of worked examples: effects on cognitive load and learning. *Appl. Cognit. Psychol.* 26, 352–358.

- Sweller, J., Cooper, G.A., 1985. The use of worked examples as a substitute for problem solving in learning algebra. *Cognit. Instruct.* 2, 59–89.
- Sweller, J., van Merriënboer, J.J., Paas, F., 2019. Cognitive architecture and instructional design: 20 years later. *Educ. Psychol. Rev.* 31, 261–292.
- Tarmizi, R.A., Sweller, J., 1988. Guidance during mathematical problem solving. *J. Educ. Psychol.* 80, 424–436.
- Van Es, E.A., Sherin, M.G., 2021. Expanding on prior conceptualizations of teacher noticing. *ZDM* 53, 17–27.
- Van Gog, T., Rummel, N., Renkl, A., 2019. Learning how to solve problems by studying examples. In: Dunlosky, J., Rawson, K. (Eds.), *Cambridge Handbook on Cognition and Education*. Cambridge University Press, Cambridge, UK, pp. 183–208.
- Van Harsel, M., Hoogerheide, V., Verkoeijen, P., van Gog, T., 2019. Effects of different sequences of examples and problems on motivation and learning. *Contemp. Educ. Psychol.* 58, 260–275.
- Vendetti, M.S., Matlen, B.J., Richland, L.E., Bunge, S.A., 2015. Analogical reasoning in the classroom: insights from cognitive science. *Mind Brain Educ.* 9, 100–106.
- Wu, Y.W., Weng, K.H., 2013. Using an analogical thinking model as an instructional tool to improve student cognitive ability in architecture design learning process. *Int. J. Technol. Des. Educ.* 23, 1017–1035.
- Zimmerman, B.J., Kitsantas, A., 2002. Acquiring writing revision and self-regulatory skill through observation and emulation. *J. Educ. Psychol.* 94, 660–668.

Feedback for student learning in higher education

David Carless, Faculty of Education, University of Hong Kong, Pok Fu Lam, Hong Kong

© 2023 Elsevier Ltd. All rights reserved.

Introduction	623
Partnerships in teacher and student feedback literacy	624
Feedback seeking strategies	625
Peer feedback	625
Digitally-enabled feedback	626
Feedback in online learning environments	627
Conclusion	627
References	628
List of relevant webpages	629

Introduction

In order to improve, individuals need to generate, receive and act upon feedback of different forms. It is necessary to appreciate what has been done well, become more aware of weaknesses and develop action plans for ongoing improvement. Indeed, it is well-recognized that feedback is one of the most significant levers to enhance student achievement (Hattie and Timperley, 2007), yet at the same time feedback processes are complex, and often fail to produce a worthwhile return on the time and resources invested (Carless and Boud, 2018).

In higher education, there is a wide range of evidence that undergraduate students do not appear to be particularly satisfied with how feedback processes are generally managed (e.g., Buckley, 2020). National quality assurance surveys in the UK, Australia and elsewhere consistently suggest that feedback is one of the least satisfying aspects of the student experience. Admittedly, it is not clear how students are defining feedback in these kinds of surveys, and they may also be conflating it with assessment grades which often provoke emotional reactions.

This sense of dissatisfaction about feedback processes is reinforced by evidence that students do not find the feedback they receive from their teachers easily understandable, useful or actionable (Winstone et al., 2017). Teachers' frustrations with grading and marking are also apparent in that they often experience these types of work as being time-consuming and relatively unproductive, not least because of feedback doing double duty involving grade justification, commentary and quality assurance demands (Winstone and Carless, 2021). A related challenge is that students and teachers seem to have different views of what feedback is and what it should do (Adcroft, 2011; Carless, 2006). First-year undergraduates are influenced by the close relationships and guidance that they experienced in high school. This might reinforce naïve perceptions that feedback means being told what to do in order to obtain good grades. When students think of feedback in terms of telling and guiding, they are sometimes under-prepared for the autonomous learning culture of higher education. Teachers often envisage a more active student role in feedback processes than some learners are prepared to adopt (Adcroft, 2011).

A consequence is that how feedback is conceptualized and defined is both critical and disputed. A useful conventional view is to see feedback as mainly comprising information provided by an agent (e.g., teacher, peer or self) about aspects of performance or understanding (Hattie and Timperley, 2007). This view of feedback is roughly similar to dictionary definitions of feedback: "If you get feedback on your work or progress, someone tells you how well or badly you are doing, and how you could improve" (Collins COBUILD, 2018). Conceptualizing feedback as an information-sharing process has been referred to as the "old paradigm" of feedback (Winstone and Carless, 2019) in that it underplays recent developments in emphasizing the learner role in generating, making sense of, and using feedback of different forms. Accordingly, a more process-oriented social constructivist view emphasizes the student role in using feedback for enhancement purposes (e.g., Carless and Boud, 2018). This way of thinking has been referred to as a "new paradigm" of feedback (Winstone and Carless, 2019) placing the learner firmly at the center of feedback processes through strategies, such as peer feedback and self-assessment. New paradigm feedback practices are defined as approaches which emphasize students generating and acting upon feedback inputs of different forms.

Internal feedback is an important concept congruent with new paradigm feedback processes. Internal feedback denotes the insights that students generate when they compare their current knowledge and competence against some reference information, such as the work of their peers (Nicol, 2021). For Nicol, external information received by learners needs to be transformed into internal feedback if it is to influence ongoing development. Building on these strands of thinking, feedback is defined in this chapter as interactive processes through which learners generate performance-relevant information to enhance their work or learning strategies. Performance-relevant information encompasses how learners interpret their learning trajectory: it includes the performance of others, assessment outcomes, and responses of peers (van der Leeuw et al., 2018).

There are various learning theories underpinning feedback research and practice. Cognitivist approaches emphasize information-processing in that individuals receive feedback inputs, process them, and act to close gaps between current and desired

performance levels (Thurlings et al., 2013). Social constructivist feedback research takes the perspective that shared and individual interpretations are developed through dialog, sense-making, and through co-construction between participants (Price et al., 2011). Sociocultural approaches to feedback highlight the role of participation and meaning-making mediated through activity within social and cultural contexts embedded in the norms of different disciplines (Esterhazy and Damşa, 2019). Socio-material approaches suggest that feedback practices are entangled with social, material, spatial and temporal influences which may impede productive student involvement (Gravett, 2020).

In the last 15 years, there has been a significant increase in attention afforded to feedback in undergraduate education, and corresponding expansion of the research base. Two significant inter-related developments within the new paradigm of feedback are worth noting. The first focuses on feedback impacts in that the core of feedback processes lies in their impact on students, how learners improve their work or learning capacities. This line of thinking places particular emphasis on the closing of feedback loops, and evidence that students have used feedback information for enhancement (Boud and Molloy, 2013; Carless, 2019; Henderson et al., 2019). The second key development is in conceptualizing and researching student feedback literacy: the capacities to make sound academic judgments and work with feedback information so that it produces useful impacts (Carless and Boud, 2018). Unless students are developing a set of capabilities to engage productively in feedback processes, inputs may not lead to impact.

The aim of this chapter is to illustrate how new paradigm feedback practices can be implemented in undergraduate education in order to produce feedback impact and the development of student feedback literacy. Five inter-related sub-themes are discussed: partnerships in feedback literacy; feedback seeking strategies; peer feedback; digitally-enabled feedback; and feedback in online learning environments. For each of these sub-themes, key rationales, challenges and further possibilities are addressed.

Partnerships in teacher and student feedback literacy

Feedback literacy has been a focus of vigorous research interest in the last five years or so, driven by the appreciation that effective feedback processes rely on the capacities that students bring to the feedback partnership as well as the support provided by teachers. In their heavily cited paper, Carless and Boud (2018) defined student feedback literacy as the understandings, capacities and dispositions needed to make sense of information and use it to enhance work or learning strategies. Key frameworks suggested that student feedback literacy principally involves seeking, generating, engaging and using feedback; the development of capacities in making academic judgments; and acknowledging and working with emotions (Carless and Boud, 2018; Chong, 2021; Molloy et al., 2020). Emotional and relational support is a pre-requisite for opening oneself to critique, revealing vulnerability and engaging with hard-hitting commentary (Steen-Utheim and Wittek, 2017). Feedback literate individuals are able to maintain emotional equilibrium and avoid defensiveness when receiving critical feedback (Carless and Boud, 2018). Feedback interactions seem to work best within supportive relationships, although these kinds of close relationships are often hard to build and maintain within massified higher education.

How feedback literacy is manifested within different disciplines is an important consideration. The social norms and dynamics of different disciplinary learning activities entail various socioculturally-based affordances and constraints for the emergence of productive feedback exchanges (Esterhazy, 2018). A study of National Qualifications Frameworks and Subject Benchmark Statements, drawing on data from 6 countries, suggests that discipline-specific feedback literacies can be developed through authentic learning activities and assessment tasks (Winstone et al., 2020a). These resonate with the concept of authentic feedback, mirroring the feedback practices of the discipline, profession or workplace (Dawson et al., 2021). Building on the notion of signature pedagogies, the characteristic ways feedback processes are carried out in specific disciplines are represented by the concept of *signature feedback practices* (Carless et al., 2020; Quinlan and Pitt, 2021). Signature feedback practices are potentially valuable in integrating disciplinary ways of thinking with authentic and motivating feedback experiences (Carless et al., 2020).

The initial focus of feedback literacy research mainly focused on the student contribution in line with the orientation of new paradigm feedback practices. There are risks, however, that a focus on individual learner capacities neglects socio-material factors and assumes more agency than students possess within neoliberal mass higher education (Gravett, 2020; Nieminen et al., 2021). There is still, however, potential for feedback literacy to be developed through shared responsibilities between teachers and students, working in partnership toward common goals of enhancing student learning (Matthews et al., 2021; Nash and Winstone, 2017). Accordingly, teacher feedback literacy has begun to become a research focus because unless teachers facilitate opportunities for students to experience using feedback productively, student development of feedback literacy may be constrained. In an initial conceptualization of teacher feedback literacy, Carless and Winstone (2020) proposed a framework of three dimensions: a design dimension focusing on designing feedback processes for student uptake; a relational dimension representing the interpersonal side of feedback exchanges; and a pragmatic dimension addressing how teachers manage the compromises and workload implications inherent in disciplinary and institutional feedback practices. In an empirically-derived framework, Boud and Dawson (2021) added macro (program design and development), meso (course design and implementation) and micro (student assignments) levels of activity. Yet teacher feedback literacies may also involve learning to negotiate institutional structures, socio-material considerations, and power relations that influence or constrain new paradigm feedback practices (Tai et al., 2021).

To sum up this sub-theme, student and teacher feedback literacy represents the capacities to make the most of feedback opportunities. The co-ordinated interplay between teacher and student feedback literacy represents a promising means of tackling the challenges of reaping learning gains from feedback processes. Ongoing research might continue teasing out disciplinary differences

in feedback literacies and involve research and development projects focused on the mutual development of teacher and student feedback literacy. The publication in November 2021 of the first student feedback literacy scale (Zhan, 2021) also promises a new line of quantitative inquiry into feedback literacy.

Feedback seeking strategies

If students are to adopt a pro-active role in feedback processes congruent with new paradigm feedback practices, they need to elicit comments from suitable others and continue dialog with them as necessary (Carless and Boud, 2018). Eliciting feedback information on self-identified priority areas forms part of the characteristics and repertoire of feedback literate students. Feedback seeking behaviors are defined as learners intentionally eliciting information about their own work for the purposes of improvement. Feedback seeking is a key element of student self-assessment because it enables learners to calibrate and refine their own judgments (Yan and Brown, 2017; Yan and Carless, 2021). Feedback seeking has been somewhat under-explored in research on higher education pedagogy, although there are significant strands in relation to the workplace (e.g., Anseel et al., 2015) and in medicine (e.g., Bok et al., 2013).

The literature identifies two main forms of feedback seeking strategies: inquiry and monitoring (Leenknecht et al., 2019). Inquiry involves directly eliciting comments from others on progress or self-identified issues, either spontaneously or after being prompted by a teacher. These kind of feedback requests begin a dialog which potentially creates the conditions for effective feedback exchanges. The inquiry function is illustrated through students seeking comments on selected aspects of their work, thereby enabling teachers to tailor all or part of their input to students' preferences and perceived needs. Malecka et al. (2020) proposed that feedback requests should be embedded purposefully within the curriculum in cumulative ways so that eliciting becomes increasingly sophisticated with students adjusting requests in the light of ongoing progress. Teachers can model and encourage feedback seeking strategies because sometimes students are wary of eliciting feedback because they are wary of appearing too needy or revealing their limitations, and/or because they are worried about potential threats to their self-esteem if the response is critical. Supportive course cultures and trusting relationships create a climate to encourage feedback requests in which admitting doubts through eliciting information from others is a normalized aspect of pedagogy.

The second strand of feedback seeking strategies involves monitoring: drawing on information available in the environment, such as making comparisons with the previous performance of oneself or others, exemplars, assessment criteria, or consulting other resources deemed relevant. This form of feedback seeking stimulates internal feedback through interpreting cues or evidence. Monitoring is an aspect of feedback literacy in that it involves learners seeking cues from the environment, the assessment task criteria and exemplars (Joughin et al., 2021; Molloy et al., 2020). A further aspect of monitoring involves drawing on previous performance and related feedback to consolidate inferences for current and future development. The feedback literate student synthesizes previous comments from a wide range of sources to inform current performance.

To sum up this sub-theme, feedback seeking exemplifies new paradigm feedback practices in highlighting the active student role in feedback processes. Capacities to elicit feedback from others and continue interactions with them, where necessary, are part of feedback literacy. In terms of ongoing research, the knowledge base would benefit from more evidence of how students use feedback seeking strategies to enhance their work or approaches to learning.

Peer feedback

Peer feedback, also known as student peer review, involves learners evaluating and making judgments about the work of peers on a task that they have already attempted (Nicol et al., 2014). The main benefits from being exposed to the work of peers arise from the internal feedback generated when students make comparisons between their own production and that of others (Nicol, 2021). Through being exposed to a variety of work of different styles and standards, feedback literate students develop a sense of what quality work looks like. As they enhance their awareness of quality, and apply insights to their own work, students develop capacities in monitoring their own work in progress. These are enhanced when teachers design opportunities for digitally-mediated dialogs where students have opportunities to clarify peer feedback and co-construct opportunities to use it (Wood, 2022).

While well-implemented by some teachers, peer feedback is often carried out in sub-optimal ways and a number of challenges commonly arise (Winstone and Carless, 2019). Students sometimes worry that they or their classmates lack sufficient knowledge or competence to provide useful feedback. They may also find that some classmates do not engage seriously in composing peer feedback or that their efforts in offering constructive comments are not reciprocated. Students are also generally wary of forms of peer assessment which involve the award of grades because of concerns about unfairness and bias. By potentially disrupting some of the power dynamics of teacher-dominated feedback, peer feedback also faces socio-political challenges (Tai et al., 2021). Skillful designs of peer feedback sequences in the hands of feedback literate teachers potentially mitigate some of these challenges.

A significant, but frequently underplayed, aspect is the need for coaching in carrying out peer review as it cannot be assumed that students understand its principles and how to do it well. Coaching can include: explaining the rationale for peer feedback and its potential benefits; teachers modeling their experiences of academic peer review and strategies for managing the emotional impact of critical feedback; supporting students to generate and apply criteria; and advising them on how to carry out peer review more effectively (Carless and Boud, 2018). A key point is that the composing of peer feedback is often more beneficial than receiving

comments because it is a rich process of diagnosing problems and suggesting actionable solutions (Nicol et al., 2014). Whenever possible, multiple peer reviews should be arranged so that students can be exposed to work of different quality and receive more than one commentary on their work (e.g., Harland et al., 2017). This prepares the ground for students to make comparisons between their own work and that of others (Nicol, 2021).

To sum up, peer feedback is a complex lifelong learning capacity that needs sustained and cumulative practice over time. Designing effective peer feedback sequences and supporting students in involving themselves productively are elements of teacher feedback literacy. Composing peer feedback and learning from comparing one's own work with that of others are two of the most beneficial aspects of peer feedback processes.

Digitally-enabled feedback

Digitally-enabled feedback processes carry the promise of timeliness, convenience, and portability with the potential to streamline staff workloads and facilitate prompt feedback interactions with students even with large classes. Technology has considerable potential to enable promising feedback practices, especially when it is context, students' needs and pedagogical considerations which drive technology choices.

Building on the previous sub-section, technology is often deployed to facilitate student peer review, especially with large classes. PeerMark within the Turnitin suite of applications and Moodle Workshop for peer assessment enable the anonymizing of student submissions and automatic allocation of peer reviewers. Moodle Workshop features, such as scoring guides, were found to be particularly useful in facilitating formative peer feedback on draft essay assignments within a macro-economics class of 800 first year students (Mostert and Snowball, 2013). Challenges still arise, however, in terms of motivating students to engage in peer dialogs that go beyond perfunctory exchanges of low-level information.

Audio and video modalities enable the production of peer feedback. In a study of audio peer feedback within large online courses, it was found that providing and receiving audio feedback encouraged deep approaches to learning and a sense of commitment (Filius et al., 2019). Peer-to-peer audio or video feedback carry potential to promote student interaction, engagement, and the development of a learning community. Whereas teacher audio or video feedback risks perpetuating information transmission approaches (Mahoney et al., 2019), peer-generated variations potentially harness the active role of students in generating internal feedback and making judgments.

Digitally-enabled feedback of different forms can be implemented in various user-friendly ways. Online annotated exemplars can serve as proxies for teacher feedback by clarifying what good performance in a task looks like and acting as inputs which stimulate students to generate internal feedback and then enhance their own work (Carless, 2020). Collaborative exchanges involving peers and teachers facilitated through cloud applications enable learners to discuss feedback and use insights to revise work in progress (Wood, 2021). Electronic marking and feedback through tools, such as "Track Changes" or annotated PDF documents can sometimes be useful although they are mainly positioned as old paradigm transmission of information. When teacher inputs of this kind are designed to encourage self-regulation or invite student follow-up, they are moving in a new paradigm direction.

Recent developments in learning analytics make it possible to track a range of students' behaviors in digital learning environments, such as whether they have viewed a resource or opened a file of feedback comments, and the time spent on these activities. Information based on learning logs and digital traces can enable formative feedback for large cohorts to be delivered just in time for student action (Pardo et al., 2019). Research into the impacts of analytics-based feedback suggests that this kind of feedback can stimulate self-regulated learning by enabling students to adjust their learning strategies to optimize their performance (Lim et al., 2020). Learning analytics focuses on individual behaviors while tending to downplay other aspects, such as social and collaborative learning, so users of analytics-based data need to be aware that what can be traced through analytics is only part of a student learning trajectory (Fawns, 2019). When presented in a student-facing way, learning analytics does, however, offer potential in contributing insights into students' learning behaviors through providing information about facets of their own progress and in relation to other members of their cohort.

A promising example of digitally-enabled feedback is the Feedback Engagement and Tracking System (FEATS), a feedback e-portfolio which enables learners to synthesize feedback from various sources, and monitor their progress toward self-identified enhancement targets (Winstone, 2019). This e-portfolio involves a feedback review and synthesis tool where students collate multiple feedback inputs to identify focal areas for improvement; a bank of relevant resources such as videos, websites or articles to enable the development of self-identified target skills; and a student-generated personalized feedback implementation plan with tasks that need to be completed by certain dates. Through repeated review of feedback information and planning for uptake, students focus on how feedback supports their learning beyond an individual unit or task (Winstone and Carless, 2019). This signals the important and under-researched dimension of program-level approaches to feedback: how can students be supported to transfer insights across the courses that make up their entire undergraduate program?

To sum up this sub-section, digitally-enabled feedback is most productive when learners are involved actively in ways which support the development of their feedback literacy. Technology is harnessed most effectively when students' needs are at the forefront, and technology facilitates new paradigm feedback practices rather than just the transmission of information.

Feedback in online learning environments

Emergency online teaching during the COVID-19 pandemic has further placed the spotlight on online feedback possibilities. Many of the digitally-enabled possibilities discussed above are relevant but there are also some additional strategies or issues that merit consideration. Automated online quizzes, for example, are part of a flexible repertoire of online feedback methods and seem particularly popular in the hard sciences. These kinds of quiz seem to be most useful when hints rather than correct answers are provided to increase cognitive engagement (Förster et al., 2018). If students retake an online quiz on multiple occasions, reflect on previous performance, and/or access relevant resources, productive learning can accrue.

Social presence is important in reducing isolation of the online or distance learner and feedback interactions can play a role. In the previous sub-section, the potential of peer-to-peer audio or video feedback was highlighted and teachers can also provide commentary through video feedback to highlight key issues, and enhance social presence. Video feedback should probably be short and to the point in order to minimize cognitive load, and is more oriented toward new paradigm feedback practices when some kind of student response or follow-up is required. The qualitative synthesis by Mahoney et al. (2019) found, however, that few studies explicitly incorporated opportunities for dialog or response to teacher video feedback inputs, so the potential for social interaction was not fully exploited.

Screencasting tools enable teachers to provide audio and/or visual commentary on student work. A screencast is a digital recording that shows actions taking place on a teacher's computer screen. By capturing what is on teachers' monitors, screencasts can show as well as tell the teacher response to student work. A useful variation congruent with new paradigm feedback practices is for students to generate their own screencasts or respond to those of the teacher. In the context of a distance learning Spanish program, Fernández-Toro and Furnborough (2014) used screencasting to enable students to record their responses to feedback through thinking aloud for 5 min. This strategy enabled students to have time and space to reflect on feedback inputs and articulate their responses. It afforded dialogic interaction between students and teachers, and helped teachers see how students engaged with feedback. Screencasts can also be combined with cloud applications, such as Google Docs, to enable students to respond to and question inputs through digitally-mediated dialogs (Wood, 2021).

Learning management systems (LMSs) are probably not particularly well-designed to facilitate new paradigm feedback practices. Challenges for students include the timing and easy accessibility of feedback messages on the LMS, reinforcing difficulties of student engagement with these kinds of feedback and grading inputs (Winstone et al., 2020b). Low achieving students often did not access feedback when it was reported separately to the grade on the LMS, reinforcing a challenge for feedback practice that students most in need of feedback are least likely to engage with it (Mensink and King, 2020). With principled pedagogical adjustments, LMSs might become useful repositories for feedback information and students could be encouraged to show how they are using feedback from previous assignments to inform current submissions in analogous ways to FEATS above. Discussion forums on LMSs could, in principle, enable students to involve themselves in dialogs around course content or work in progress but in practice students often do not seem to find these asynchronous written discussions particularly motivating, user-friendly or engaging. If well-designed, discussion forums might generate useful feedback exchanges but they need to enable student agency and interaction so that they do not just end up as busy work that does not contribute to substantive course assessment outcomes.

Social media, such as Facebook or Twitter are increasingly being used for academic interaction and may be particularly useful in online learning environments. There is potential to capitalize on the ubiquity of social media by integrating feedback exchanges. A challenge, however, lies in bridging interpersonal and academic domains by using a predominantly social platform for learning purposes. An implication is that participating in these kinds of feedback exchanges might be a voluntary option rather than a compulsory course element. Under these circumstances, there is potential in social media for facilitating collaborative learning and feedback dialogs.

To sum up this sub-section, online feedback needs to be carefully planned and designed so that inputs go beyond than unidirectional transfer of information in old paradigm ways. Future research and development work might aim to investigate the intersection between new paradigm feedback practices and online feedback, and the teacher and student feedback literacies required for successful implementation.

Conclusion

Feedback is a complex social interaction which carries implications for well-being and self-esteem. Effective feedback processes rest on shared responsibilities between participants within cultural and disciplinary contexts. This perspective resonates with new paradigm feedback practices in highlighting active student roles in seeking, generating, processing and using feedback information. The potential value of digital feedback possibilities has been an area of focus for some time, but marrying the digital with purposeful feedback designs and the co-development of teacher and student feedback literacies remains a key aspect for further exploration.

References

- Adcroft, A., 2011. The mythology of feedback. *High. Educ. Res. Dev.* 30 (4), 405–419. <https://doi.org/10.1080/07294360.2010.526096>.
- Anseel, F., Beatty, A.S., Shen, W., Lievens, F., Sackett, P.R., 2015. How are we doing after 30 years? A meta-analytic review of the antecedents and outcomes of feedback-seeking behavior. *J. Manag.* 41 (1), 318–348. <https://doi.org/10.1177/0149206313484521>.
- Bok, H., Teunissen, P., Spruijt, A., Fokkema, J., van Beukelen, P., Jaarsma, D., van der Vleuten, C., 2013. Clarifying students' feedback-seeking behavior in clinical clerkships. *Med. Educ.* 47 (3), 282–291. <https://doi.org/10.1111/medu.12054>.
- Boud, D., Dawson, P., 2021. What feedback literate teachers do: an empirically-derived competency framework. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.1910928>.
- Boud, D., Molloy, E., 2013. Rethinking models of feedback for learning: the challenge of design. *Assess. Eval. High. Educ.* 38 (6), 698–712. <https://doi.org/10.1080/02602938.2012.691462>.
- Buckley, A., 2020. Crisis? What crisis? Interpreting student feedback on assessment. *Assess. Eval. High. Educ.* 46 (7), 1008–1019. <https://doi.org/10.1080/02602938.2020.1846015>.
- Carless, D., Boud, D., 2018. The development of student feedback literacy: enabling uptake of feedback. *Assess. Eval. High. Educ.* 43 (8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>.
- Carless, D., Winstone, N., 2020. Teacher feedback literacy and its interplay with student feedback literacy. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2020.1782372>.
- Carless, D., To, J., Kwan, C., Kwok, J., 2020. Disciplinary perspectives on feedback practices: towards signature feedback practices. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2020.1863355>.
- Carless, D., 2006. Differing perceptions in the feedback process. *Stud. High. Educ.* 31 (2), 219–233. <https://doi.org/10.1080/03075070600572132>.
- Carless, D., 2019. Feedback loops and the longer-term: towards feedback spirals. *Assess. Eval. High. Educ.* 44 (5), 705–714. <https://doi.org/10.1080/02602938.2018.1531108>.
- Carless, D., 2020. From teacher transmission of information to student feedback literacy: activating the learner role in feedback processes. *Act. Learn. High. Educ.* <https://doi.org/10.1177/1469787420945845>.
- Chong, S.W., 2021. Reconsidering student feedback literacy from an ecological perspective. *Assess. Eval. High. Educ.* 46 (1), 92–104. <https://doi.org/10.1080/02602938.2020.1730765>.
- Collins COBUILD Advanced Learner's English Dictionary, 2018. HarperCollins Publishers. Retrieved from: <https://www.collinsdictionary.com/dictionary/english/feedback>.
- Dawson, P., Carless, D., Lee, P.P.W., 2021. Authentic feedback: supporting learners to engage in disciplinary feedback practices. *Assess. Eval. High. Educ.* 46 (2), 286–296. <https://doi.org/10.1080/02602938.2020.1769022>.
- Esterhazy, R., Damsa, C., 2019. Unpacking the feedback process: an analysis of undergraduate students' interactional meaning-making of feedback comments. *Stud. High. Educ.* 44 (2), 260–274. <https://doi.org/10.1080/03075079.2017.1359249>.
- Esterhazy, R., 2018. What matters for productive feedback? Disciplinary practices and their relational dynamics. *Assess. Eval. High. Educ.* 43 (8), 1302–1314. <https://doi.org/10.1080/02602938.2018.1463353>.
- Fawns, T., 2019. Postdigital education in design and practice. *Postdigit. Sci. Educ.* 1, 132–145. <https://doi.org/10.1007/s42438-018-0021-8>.
- Fernández-Toro, M., Fumbarough, C., 2014. Feedback on feedback: eliciting learners' responses to written feedback through student-generated screencasts. *Educ. Media Int.* 51 (1), 35–48. <https://doi.org/10.1080/09523987.2014.889401>.
- Filius, R., de Kleijn, R., Uijl, S., Prins, F., van Rijen, H., Grobbee, D., 2019. Audio peer feedback to promote deep learning in online education. *J. Comput. Assist. Learn.* 35 (5), 607–619. <https://doi.org/10.1111/jcal.12363>.
- Förster, M., Weiser, C., Maur, A., 2018. How feedback provided by voluntary electronic quizzes affects learning outcomes of university students in large classes. *Comput. Educ.* 121, 100–114. <https://doi.org/10.1016/j.compedu.2018.02.012>.
- Gravett, K., 2020. Feedback literacies as sociomaterial practice. *Crit. Stud. Educ.* <https://doi.org/10.1080/17508487.2020.1747099>.
- Harland, T., Wald, N., Randhawa, H., 2017. Student peer review: enhancing formative feedback with a rebuttal. *Assess. Eval. High. Educ.* 42 (5), 801–811. <https://doi.org/10.1080/02602938.2016.1194368>.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77 (1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Henderson, M., Molloy, E., Ajjawi, R., Boud, D., 2019. Designing feedback for impact. In: Henderson, M., Ajjawi, R., Boud, D., Molloy, E. (Eds.), *The Impact of Feedback in Higher Education: Improving Assessment Outcomes for Learners*. Palgrave Macmillan, Basingstoke, pp. 267–285.
- Joughin, G., Boud, D., Dawson, P., Tai, J., 2021. What can higher education learn from feedback seeking behaviour in organisations? Implications for feedback literacy. *Assess. Eval. High. Educ.* 46 (1), 80–91. <https://doi.org/10.1080/02602938.2020.1733491>.
- Leenknecht, M., Hompus, P., Van der Schaaf, M., 2019. Feedback seeking behaviour in higher education: the association with students' goal orientation and deep learning approach. *Assess. Eval. High. Educ.* 44, 1069–1078. <https://doi.org/10.1080/02602938.2019.1571161>.
- Lim, L.-A., Gentili, S., Pardo, A., et al., 2020. What changes and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learn. Instr.* 72, 101202. <https://doi.org/10.1016/j.learninstruc.2019.04.003>.
- Mahoney, P., Macfarlane, S., Ajjawi, R., 2019. A qualitative synthesis of video feedback in higher education. *Teach. High. Educ.* 24 (2), 157–179. <https://doi.org/10.1080/13562517.2018.1471457>.
- Malecka, B., Boud, D., Carless, D., 2020. Eliciting, processing and responding to feedback: mechanisms for embedding feedback literacy within the curriculum. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2020.1754784>.
- Matthews, K., Tai, J., Enright, E., Carless, D., Rafferty, C., Winstone, N., 2021. Transgressing the boundaries of “students as partners” and “feedback” discourse communities to advance democratic education. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2021.1903854>.
- Mensink, P., King, K., 2020. Student access of online feedback is modified by the availability of assessment marks, gender and academic performance. *Br. J. Educ. Technol.* 51 (1), 10–22.
- Molloy, E., Boud, D., Henderson, M., 2020. Developing a learning-centred framework for feedback literacy. *Assess. Eval. High. Educ.* 45 (4), 527–540. <https://doi.org/10.1080/02602938.2019.1667955>.
- Mostert, M., Snowball, J., 2013. Where angels fear to tread: online peer assessment in a large first-year class. *Assess. Eval. High. Educ.* 38 (6), 674–686. <https://doi.org/10.1080/02602938.2012.683770>.
- Nash, R., Winstone, N., 2017. Responsibility-sharing in the giving and receiving of assessment feedback. *Front. Psychol.* 8, 1519. <https://doi.org/10.3389/fpsyg.2017.01519>.
- Nicol, D., Thomson, A., Breslin, C., 2014. Rethinking feedback practices in higher education: a peer review perspective. *Assess. Eval. High. Educ.* 39 (1), 102–122. <https://doi.org/10.1080/02602938.2013.795518>.
- Nicol, D., 2021. The power of internal feedback: exploiting natural comparison processes. *Assess. Eval. High. Educ.* 46 (5), 756–778. <https://doi.org/10.1080/02602938.2020.1823314>.
- Nieminen, J.H., Tai, J., Boud, D., Henderson, M., 2021. Student agency in feedback: beyond the individual. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.1887080>.
- Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., Miriahi, N., 2019. Using learning analytics to scale the provision of personalised feedback. *Br. J. Educ. Technol.* 50 (1), 128–138. <https://doi.org/10.1111/bjet.12592>.
- Price, M., Handley, K., Millar, J., 2011. Feedback: focusing attention on engagement. *Stud. High. Educ.* 36 (8), 879–896. <https://doi.org/10.1080/03075079.2010.483513>.

- Quinlan, K.M., Pitt, E., 2021. Towards signature assessment and feedback practices: a taxonomy of discipline-specific elements of assessment for learning. *Assess. Educ.* 28 (2), 191–207. <https://doi.org/10.1080/0969594X.2021.1930447>.
- Steen-Utheim, A., Wittek, A., 2017. Dialogic feedback and the potentialities for student learning. *Learn. Cult. Soc. Interact.* 15, 18–30. <https://doi.org/10.1016/j.lcsi.2017.06.002>.
- Tai, J., Bearman, M., Gravett, K., Molloy, E., 2021. Exploring the notion of teacher feedback literacies through the theory of practice architectures. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.1948967>.
- Thurlings, M., Vermeulen, M., Bastiaens, T., Stijnen, S., 2013. Understanding feedback: a learning theory perspective. *Educ. Res. Rev.* 9, 1–15. <https://doi.org/10.1016/j.edurev.2012.11.004>.
- van der Leeuw, R., Teunissen, P., van der Vleuten, C., 2018. Broadening the scope of feedback to promote its relevance to workplace learning. *Acad. Med.* 93 (4), 556–559. <https://doi.org/10.1097/acm.0000000000001962>.
- Winstone, N., Carless, D., 2019. *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. Routledge, London.
- Winstone, N., Carless, D., 2021. Who is feedback for? The influence of accountability and quality assurance agendas on the enactment of feedback processes. *Assess. Educ.* 28 (3), 261–278. <https://doi.org/10.1080/0969594X.2021.1926221>.
- Winstone, N., Nash, R., Rowntree, J., Parker, M., 2017. "It'd be useful, but I wouldn't use it": barriers to university students' feedback seeking and reciprocity. *Stud. High. Educ.* 42 (11), 2026–2041. <https://doi.org/10.1080/03075079.2015.1130032>.
- Winstone, N., Ballou, K., Carless, D., 2020a. Discipline-specific feedback literacies: a framework for curriculum design. *High. Educ.* <https://doi.org/10.1007/s10734-020-00632-0>.
- Winstone, N., Bourne, J., Medland, E., Niculescu, I., Rees, R., 2020b. "Check the grade, log out": students' engagement with feedback in learning management systems. *Assess. Eval. High. Educ.* 46 (4), 631–643. <https://doi.org/10.1080/02602938.2020.1787331>.
- Winstone, N., 2019. Facilitating students' use of feedback: capturing and tracking impact using digital tools. In: Henderson, M., Ajjawi, R., Boud, D., Molloy, E. (Eds.), *The Impact of Feedback in Higher Education: Improving Assessment Outcomes for Learners*. Palgrave Macmillan, Basingstoke, pp. 225–242.
- Wood, J., 2021. A dialogic technology-mediated model of feedback uptake and literacy. *Assess. Eval. High. Educ.* 46 (8), 1173–1190. <https://doi.org/10.1080/02602938.2020.1852174>.
- Wood, J., 2022. Making peer feedback work: the contribution of technology-mediated dialogic peer feedback to feedback uptake and literacy. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.1914544>.
- Yan, Z., Brown, G.T.L., 2017. A cyclical self-assessment process: towards a model of how students engage in self-assessment. *Assess. Eval. High. Educ.* 42 (8), 1247–1262. <https://doi.org/10.1080/02602938.2016.1260091>.
- Yan, Z., Carless, D., 2021. Self-assessment is about more than self: the enabling role of feedback literacy. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.2001431>.
- Zhan, Y., 2021. Developing and validating a student feedback literacy scale. *Assess. Eval. High. Educ.* <https://doi.org/10.1080/02602938.2021.2001430>.

List of relevant webpages

- Developing Engagement With Feedback Toolkit, <https://www.advance-he.ac.uk/knowledge-hub/developing-engagement-feedback-toolkit-deft>.
- Enhancing Feedback, University of Edinburgh, <https://www.enhancingfeedback.ed.ac.uk/>.
- FEATS, University of Surrey, <https://www.surrey.ac.uk/institute-of-education/learning-lab/feedback-engagement-tracking-surrey>.
- Feedback for Learning Project, Led by Monash University, <http://newmediaresearch.educ.monash.edu.au/feedback/>.
- Y1 Feedback Project, Led by Maynooth University, <http://y1feedback.ie/about/>.

Inquiry-based learning

Ard W. Lazonder, Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	630
An example of inquiry-based learning	631
Conditions for inquiry-based learning	631
Criticisms of inquiry-based learning	631
Guidance to facilitate inquiry-based learning	632
Differentiated guidance to optimize inquiry-based learning	633
Conclusion	634
References	635

Introduction

Ancient philosophers such as Aristotle, Socrates and Jean-Jacques Rousseau believed that students do not learn by being told. Knowledge, in their view, is constructed through active and curiosity-driven processes of observation and reasoning. These sentiments echo in the writings of American psychologist John Dewey (1900, 1938), who planted the seed for the instructional method currently known as inquiry-based learning. Dewey argued that teachers should create opportunities for students to discover for themselves, and guide them in the process of becoming active and independent learners. More recent influential scholars such as Jean Piaget and Jerome Bruner provided the first empirical support for Dewey's philosophical propositions by showing how and under which conditions people learn through self-directed explorations (Bruner et al., 1956; Inhelder and Piaget, 1958).

Inquiry-based learning has become a common instructional method in many countries, although its application varies depending on the educational level and definitions used. In science education, which is the main focus of this article, the method is defined as one in which "students investigate scientifically oriented questions, conduct experiments, formulate explanations based on evidence, evaluate their explanations in light of alternative explanations, and communicate and justify their proposed explanations" (National Research Council, 2000, p. 25). These activities are often combined in a so-called inquiry cycle that generally consists of five phases (Pedaste et al., 2015). In the orientation phase, students explore the topic of investigation, for example by reading introductory texts or making initial observations. During the conceptualization phase, students generate one or more research questions and formulate hypotheses, which they test through observation or experimentation during the investigation phase. The data that result from these investigations are interpreted during the conclusion phase where students decide on the validity of their hypotheses or develop new ideas. In the final discussion phase, students look back on their inquiry and communicate their findings.

Inquiry-based lessons do not necessarily include all of these activities. Teachers can, for example, decide to take over or skip activities that are too difficult or time consuming. As a result, inquiry-based lessons can take many forms and range from highly teacher directed to highly student centered (Furtak et al., 2012). Bell et al. (2005) proposed a four-level framework to evaluate instructional activities in terms of the amount of inquiry students perform themselves. At level 1, the lowest level of inquiry, students investigate a teacher-given question using a prescribed procedure to confirm a known outcome. At level 2, the activity is slightly more open in that the outcome is unknown to students in advance. Inquiry tasks at both these levels prevail in textbook-based curricula (Chinn and Malhotra, 2002) and are sometimes mockingly referred to as "cookbook science". At level 3, only the research question is given so students have to design and conduct their own investigations, whereas level 4 activities are completely open in the sense that research questions, methods and outcomes all have to be conceived by the students.

A well-designed curriculum gradually exposes students to more open forms of inquiry (Bell et al., 2005). This learning progression assumes that students become increasingly more adept in doing research, which is indeed a main goal of inquiry-based learning. However, not all students succeed in improving their inquiry skills and even the ones who do, tend to struggle during a self-directed investigation. So in addition to selecting a level of inquiry, teachers should assist students in order for inquiry-based learning to be effective, and recent research suggests that this guidance should be adapted to the needs of individual students. These issues of student readiness, guidance and differentiation form the backbone of this article. They are discussed in more detail in the sections below, with a main emphasis on the inquiry skills employed in elementary and secondary science classrooms. This is not to say that inquiry-based learning is exclusively meant for science education; the method has been successfully applied in other school subjects such as economics (Grol et al., 2016) and history (Reisman, 2012; Voet and De Wever, 2017), but these applications are beyond the scope of this article.

An example of inquiry-based learning

An example from an elementary school science project developed by the Science Education Hub Radboud University (Dekker and Van Baren-Nawrocka, 2017) illustrates inquiry-based learning in action. The sixth-graders in this example were first introduced to molecular collisions, which was the topic of inquiry, by observing how long it takes for colored ink to dissolve in hot and cold water. Following a teacher-led discussion of the findings, the children started thinking of a research question. One group decided to investigate how temperature affects the application of paint. They hypothesized that warm paint is more spreadable than cold paint because its molecules move faster. To test this idea, the children asked ten classmates to apply three samples of paint on a sheet of paper and indicate on a 3-point scale how well each one spread. The paint samples differed in temperature while all other materials were kept the same. When the ratings were displayed in a table, the children noticed a clear pattern in scores: paint from the fridge (7 °C) was perceived less spreadable than paint at room temperature which, in turn, was regarded less easy to spread than paint heated to 50 °C. These results confirmed children's hypothesis: warm paint is indeed more spreadable than cold paint. They presented this conclusion on a science fair poster, together with a brief description of their hypothesis, procedure, materials and results.

The children's teacher had a supportive role in every phase of the inquiry. She planned each activity and monitored progress to ensure that all groups would finish in time. In addition to this regulative support, the teacher helped pupils with the procedural parts of the inquiry, for instance by checking whether research questions were feasible and hypotheses testable, suggesting ways to set up a good experiment, explaining how data can be presented in a table, and so on. Although some of this guidance was available through worksheets, the teacher was on standby to offer ad hoc assistance where needed. She also offered content support at the beginning and end of each lesson to activate children's prior knowledge and discuss the learning experiences in light of molecular collisions. And on top of that, she took care of all practical arrangements. Her continuous and substantial guidance enabled children to successfully complete an open inquiry (i.e., "level 4" according to the framework of Bell et al., 2005) that would otherwise be beyond their reach.

Conditions for inquiry-based learning

Not every topic lends itself well to inquiry-based learning. Plain facts (e.g., there are 2.54 cm in an inch) and simple operations (e.g., $2 + 2 = \dots$) leave little room for scientific investigation and are probably best taught through didactic methods. But when the learning content is more ill-structured or open to multiple interpretations, inquiry-based learning can foster a deep understanding. A good topic for inquiry encourages students' curiosity and enables them to pose research questions and hypotheses to test their initial beliefs. Although it may sound paradoxical, an inquiry is generally more effective if students enter the classroom with some prior knowledge of the topic of investigation—regardless of its truthfulness—as this will help generate hypotheses and provide a focus for the inquiry (Lazonder et al., 2008; Lazonder et al., 2009). In other words, domain knowledge facilitates the use of inquiry skills.

Inquiry skills, like the ones exemplified in the previous section, generally start developing in Kindergarten. Four-year-olds display an initial proficiency in experimentation and evidence evaluation (Koerber and Osterhaus, 2019; Lapidow and Walker, 2020; Van der Graaf et al., 2018) whereas the more difficult skill of hypothesizing emerges around age 12 (Kuhn et al., 2009; Piekny and Maehler, 2013). Still, the development of all skills is rather slow and erratic, and differs considerably among same-age children (Lazonder et al., 2021; Schlatter et al., 2020). These developmental differences might explain why some high school students still face problems during a self-directed investigation (Krajcik et al., 1998; Mulder et al., 2010)—but it could also be that these students fail to apply their learned inquiry skills when left on their own account. So the bottom line is that students of all ages are ready for inquiry-based learning, but need help from their teachers to make their investigations productive for learning.

As teachers are the key element in inquiry-based classrooms, it is problematic that many of them find inquiry-based learning demanding and difficult to implement in their lessons (Dobber et al., 2017). Those capable of successfully applying the method possess certain attitudes and skills to promote student learning. Research suggests that these teachers have rich content knowledge and hold the epistemological belief that inquiry is a quest for meaning rather than a test to determine if something works or a set of pre-determined activities that lead to a known outcome. Successful teachers also use well-vetted strategies to guide students through the phases of the inquiry cycle, including ways to encourage their curiosity, promote productive classroom discourse and develop a culture of inquiry (Danielsson and Warwick, 2014; Dobber et al., 2017; Yoon et al., 2012).

Criticisms of inquiry-based learning

Although inquiry-based learning has enjoyed widespread support by educational policy makers and field experts, opinions in the scientific community are divided (see, for an overview, Zhang, 2016). Back in the 1960s, scholars such as Ausubel (1962) considered the method—which, at the time, was called "discovery learning"—to be developmentally appropriate for children in the concrete operational stage; older learners capable of formal operational reasoning should be taught by expository methods, which back then were deemed more effective and efficient. More recently, Klahr and Nigam (2004) fanned the flames by showing that even pupils in third and fourth grade learn more from interactive teacher demonstrations than through unguided discovery. Based on

a historical analysis of the effectiveness of such unguided discovery methods, Mayer (2004) argued against what he called “pure discovery learning” and in favor of more guided methods of instruction. A similar argument was raised by Kirscher et al. (2006), who based their objections on the limited processing capacity of the human brain. When students have to find things out on their own, the processing demands on working memory become so high that there is little cognitive capacity left for the actual learning of new information.

Hmelo-Silver et al. (2007) replied by demonstrating that Kirschner et al. (2006) confounded inquiry-based learning with unguided discovery learning. Hmelo-Silver and colleagues argued that in practice inquiry-based activities are often heavily guided—or scaffolded as they called it—in order to reduce cognitive load and facilitate student learning. This guidance can take many forms and does not necessarily have to precede the inquiry. Rather, effective guidance can just as well be, and often is, embedded in the activity itself. This reply set the research agenda for the next decade and inspired many scholars to start examining the effectiveness of different types of guidance during inquiry-based learning (e.g., Hagemans et al., 2013; Kuhn and Katz, 2009; Lazonder and Kamp, 2012; Morgan and Brooks, 2012).

Meanwhile, outcomes of international benchmarking studies such as the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) fueled the debate over the value of inquiry-based learning. For example, Kaya and Rice (2010) showed that the frequency of inquiry-based lessons was negatively associated with elementary schoolchildren’s scores on the science items of the TIMSS 2003 assessment. More recently, Jerrim et al. (2019) analyzed the PISA 2015 dataset and found that the frequency of inquiry-based learning was negatively associated with English teenagers’ PISA test scores (cf. Oliver et al., 2019). These findings might be interpreted as disqualifying inquiry-based learning as an effective way of teaching scientific knowledge. Such criticism should nevertheless be interpreted with caution because the TIMSS and PISA data merely convey how often students were exposed to inquiry-based learning. As it remains unknown which activities students engaged in and how much guidance they received, the comparability and accumulation of results across schools and countries remains problematic (Rönnebeck et al., 2016) and points to a need for more systematic comparisons of the effects of guidance.

Guidance to facilitate inquiry-based learning

The relative effectiveness of instructional methods has intrigued educational scientists for many years and inquiry-based learning has often been included in their comparative analyses. While individual studies contrasting two or more instructional methods often fail to provide rigorous evidence due to a lack of experimental control (see, for an exception, Eysink et al., 2009; Martella et al., 2020), meta-analyses synthesizing the results of these studies are generally considered a powerful means to establish true effects and estimate their magnitude. In one such research integration, Alfieri et al. (2011) concluded on the basis of 164 primary studies that unguided inquiry-based learning is less effective than direct instruction. The Cohen’s d effect size estimate indicated that posttest scores were 0.38 standard deviation higher for students who learned the same material through direct instruction, which represents a medium effect. This result clearly undergirds the argument presented by Mayer (2004) and Kirschner et al. (2006) that unguided discovery is ineffective. However, the second part of Alfieri et al.’s meta-analysis compared expository methods such as direct instruction with guided inquiry-based learning and found favorable results for the latter, again with a medium effect size ($d = 0.30$). In other words, inquiry-based learning is effective if adequately guided.

But what exactly does “adequate guidance” entail? One answer that immediately comes to mind is that students should receive sufficient guidance to be able to complete their inquiry and learn from the activity. A meta-analysis by Furtak et al. (2012) confirmed that the amount of guidance makes a difference. Participants of the 37 synthesized studies tended to learn more from a minimally guided inquiry than through expository instruction, but the magnitude of this effect was more than doubled in situations where teacher guidance was intensified. This evidence, although insightful, is of course one piece of the puzzle (cf. Vorholzer and Aufschneider, 2019). Questions remain regarding whether guidance should focus on the content or the procedural part of the inquiry, whether the specificity of the guidance matters, and whether the effectiveness of giving more or more specific guidance depends on the students’ age and experience.

These questions were addressed in a meta-analysis by Lazonder and Harmsen (2016), who synthesized 72 studies on inquiry-based learning in science and mathematics education. The authors based their categorization of guidance on a taxonomy originally proposed by De Jong and Lazonder (2014), which comprises six types of guidance that differ in terms of specificity or explicitness. As shown in Table 1, these types of guidance become increasingly more specific and were therefore assumed to suit the needs of learners with increasingly less inquiry skills. In line with Hmelo-Silver et al. (2007), all types of guidance can be provided by the classroom teacher, but (part of) it can also be offered by curriculum materials, educational software, teaching assistants, or classmates.

All studies included in Lazonder and Harmsen’s (2016) meta-analysis contrasted a group of students who received some kind of guidance during an inquiry task against a group of students who performed the same task without this guidance. Results showed that, overall, guidance empowered students’ inquiry skills during the task, enhanced the quality of the artifacts they created (e.g., lab reports, science fair posters), and increased learning outcomes as assessed by a posttest. Effect sizes were in the range of 0.50–0.71 and differed considerably across studies. Contrary to expectations, this variability was largely independent of the type of guidance. Although this finding might seem counterintuitive—it might be assumed that more specific guidance produces larger effects—the taxonomy in Table 1 presents a possible explanation: more specific guidance is intended for younger, less experienced students so the net effect could be the same. To test this possibility, Lazonder and Harmsen examined whether the effectiveness of the types of

Table 1 Taxonomy of guidance for inquiry-based learning.

<i>Type of guidance</i>	<i>Basic idea</i>	<i>Intended audience</i>
Process constraints	Restrict the comprehensiveness of the learning task	Learners who are able to perform and regulate the basic inquiry process, but still lack the experience to do so under more demanding circumstances
Status overviews	Make task progress or learning visible	Learners who are able to perform the basic inquiry process, but lack the skills to plan and keep track of their learning trajectory
Prompts	Remind to perform an action	Learners who are able to perform an action but may not do so on their own initiative
Heuristics	Remind to perform an action and suggest how to perform that action	Learners who do not know exactly when and how an action should be performed
Scaffolds	Explain or take over the more demanding parts of an action	Learners who do not have the proficiency to perform an action themselves or cannot perform the action from memory
Explanations	Specify exactly how to perform an action	Learners who are (largely) incognizant of the action and how it should be performed

Note. From: Lazonder, A.W., Harmsen, R., 2016. Meta-analysis of inquiry-based learning: effects of guidance. *Rev. Educ. Res.* 86, 681–718.

guidance listed in [Table 1](#) varied for different-aged students. This was not the case: children, teenagers and adolescents were equally responsive to any particular type of guidance. These findings contradict the assumed match between the specificity of guidance and students' prior knowledge and skills, and leaves unanswered the question why more specific guidance is not more productive for learning.

To conclude, there is a vast body of evidence demonstrating that guided inquiry-based learning is more effective than expository methods of instruction. Researchers have therefore started to examine the role of guidance in more detail but these attempts have thus far failed to provide a solid understanding of “what works when and why”. One possible reason is that the lion's share of these studies administered one type of guidance to groups of peers. As students of the same age, and even in the same classroom often differ considerably in terms of general cognitive abilities and readiness for inquiry-based learning, a paradigm shift toward more differentiated forms of guidance might be needed to understand how teachers can support students' inquiry in more effective ways.

Differentiated guidance to optimize inquiry-based learning

[Koerber et al. \(2015\)](#) examined what second-, third- and fourth-graders know about “doing research”. The results of this cross-sectional study implied that children gradually develop a more advanced understanding of scientific inquiry, but occasionally fall back on more naïve conceptions. This inference was generally confirmed in a three-year longitudinal assessment ([Lazonder et al., 2021](#)), which additionally revealed substantial individual variation in children's development of inquiry skills. Some low-performing children in this study hardly made any progress throughout the years, whereas others managed to improve their proficiency scores from 7 to 87% on a performance-based inquiry-skills test. This overall progress does not apply equally to all inquiry skills. Experimentation skills, for example, are generally learned rather quickly whereas the skills of hypothesizing and evaluating evidence require more than a single lesson series to improve ([Schlatter et al., 2020](#)).

This snapshot of the developmental psychology research serves to illustrate the sizable individual differences among students and across inquiry skills. Accommodating these differences is a promising way to make inquiry-based learning more effective, and requires abandoning the “one size fits all” approach in favor of more differentiated forms of guidance. Research suggests that these adaptations can be planned in advance based on students' proficiency in language and math as well as their performance in previous school inquiry projects. Language and math achievements are reliable sources of adaptation because of their consistent positive association with students' performance on inquiry tasks (e.g., [Koerber et al., 2015](#); [Lazonder et al., 2021](#); [Schlatter et al., 2020](#); [Van der Graaf et al., 2018](#)). Other cognitive and demographic characteristics such as intelligence, executive functioning, gender and socioeconomic status have a less determinate influence and are therefore a less reliable source.

The taxonomy presented in [Table 1](#) provides suggestions for how guidance can be differentiated to meet the needs of individual students. Teachers could, for example, decide to offer experimentation prompts to the more advanced students in their class (e.g., “remember to design a well-controlled experiment”), heuristics to the average achievers (e.g., “remember to design a well-controlled experiment, so only vary the variable you want to find out about”) and full-fledged explanations about experimental design procedures to students who are expected to struggle most. This example illustrates what [Tomlinson \(2001\)](#) referred to as “differentiation in process”. Differentiation can also target the learning content, for instance by having advanced students investigate more difficult

research questions; the products students create to demonstrate their learning; or the learning environment, which is often synonymous to grouping arrangements in collaborative classrooms.

Eysink et al. (2017) exemplified how these differentiation modes can be combined in an inquiry-based classroom. Pupils in this study collaborated in groups of four to investigate elementary school physics themes such as magnetism, buoyancy and sound. Each theme comprised four topics that differed in complexity, and pupils were assigned to a topic that matched their ability level (as determined by a pretest). The better-achieving children, in addition, received coarse-grained written directions to perform their part of the inquiry, whereas stepwise instructions were given to the lower-achieving children. Teacher assistance during the lessons was adapted accordingly, meaning that more specific guidance was given to lower-achieving pupils and vice versa. Toward this end, group composition changed periodically from heterogeneous ability groups (where students could help each other) to homogeneous ability groups (where the teacher could give ability-matched support).

This example indicates that differentiated guidance does not stand on its own as it should align with instructional decisions regarding the learning content, the products students create and the arrangement of the learning environment. This “balancing act” is one of the reasons why many teachers find it difficult to differentiate their lessons (e.g., Gaitas and Alves Martins, 2017) and often refrain from trying. Those who do succeed possess sufficient domain knowledge, have adopted a growth mindset and reflect on their own teaching (Parsons et al., 2018). These inherent teacher characteristics match the conditions for successful inquiry-based learning presented earlier and develop from experience and through formal training. The merits of differentiation training were demonstrated by Vogt and Rogalla (2009), who offered elementary and secondary science teachers a two-day seminar and nine follow-up coaching sessions to help apply learned differentiation strategies in class. The participating teachers were found to exhibit superior differentiation behavior compared to teachers from an untrained control group, and their students learned more from the differentiated science lessons than students in the control classes.

These findings can be summarized in a punny phrase: “teachers need guidance to differentiate their guidance”. This was also one of the conclusions from an unpublished systematic review of differentiation practices in language, math and science education (Lazonder, 2018). Across the 32 synthesized studies, continual support was found to be more effective than initial support (i.e., just an introductory training), which, in turn, was as ineffective as no teacher support at all. A lack of teacher support was probably the main reason why the integrated differentiation approach of Eysink et al. (2017) failed to enhance student learning. Studies that continued to coach and assist teachers during differentiation proved more successful (e.g., Mastropieri et al., 2006; McCrea Simpkins et al., 2009).

The difficulties teachers anticipate or encounter with differentiation can be circumvented by embedding guidance in learning materials or the learning environment. Such adaptive guidance is becoming increasingly popular in software applications for language and math learning, and is beginning to find its way into science education research. For example, Neubrand et al. (2016) adapted the prompts tenth-graders received during an online evolutionary biology lesson based on pretest scores. Contrary to expectations, students did not learn more when supported by ability-matched prompts. In a more successful attempt using the Go-Lab online learning environment, a software agent generated adaptive feedback on the hypotheses middle school students’ drafted while investigating electrical circuits. Students who consulted this feedback were much more likely to create syntactically correct hypotheses than those who ignored it (Kroeze et al., 2019). Still, studies investigating differentiation in science education are remarkably scarce. Researchers should therefore devote more attention to this important topic in order to unveil how science teachers can guide students’ inquiry in more effective ways.

Conclusion

The research reviewed in this article paints an optimistic portrait of inquiry-based learning. Although not without criticism, the method can certainly be more effective than expository forms of teaching and learning—if used thoughtfully. Teachers who successfully apply inquiry-based learning in their classrooms ensure that their students have some initial understanding of the topic of investigation and offer appropriate guidance to compensate for anticipated and observed skill deficiencies. This guidance should preferably be attuned to the needs and preferences of every single student, which admittedly puts high demands on the teacher. Distributing the guidance across curriculum materials, educational software and even fellow students could be a fruitful way to alleviate this burden.

In a broader perspective, inquiry-based learning belongs to a family of student-centered approaches that all share a focus on active learning, which is rooted in the belief that knowledge is constructed from experience. Related approaches relevant to the field of science education include design-based learning (e.g., De Jong et al., 2010), model-based learning (e.g., Mulder et al., 2015) and problem-based learning (e.g., Hmelo-Silver et al., 2018). All these approaches include some form of student research that generally follows the steps in the inquiry cycle, although minor variations exist due to the purpose of the activity: designing an artifact or solving a problem requires a slightly different research approach than investigating scientifically oriented questions. Conversely, models of inquiry-based learning also include elements of other instructional approaches discussed in this section of the encyclopedia, such as collaborative learning (students often conduct their inquiry in dyads or small groups) and argumentation (students evaluate evidence in order to draw conclusions).

Computer simulations and virtual labs are gaining momentum in inquiry-based science classrooms. These technological applications have some clear advantages over experiments performed with “real” materials in the school lab, such as shorter preparation time, increased safety and the production of clean datasets. Research has shown that both experimentation modes have a similar

positive effect on students' acquisition of conceptual knowledge and inquiry skills. Virtual and real experiments can also be combined, for instance by having students explore a topic through simulation-based experiments before entering the school lab. Alternating this sequence so that lab experiments precede simulation work is equally effective, thus offering teachers the possibility to combine virtual and real experiments based on practical considerations (Kapici et al., 2019; Lazonder and Ehrenhard, 2014; Zacharia, 2015).

Technological advances also open new opportunities for offering and differentiating student guidance. Rather than placing this responsibility entirely in the hands of the teacher, intelligent software agents could monitor students' ongoing activities and provide tailor-made support, either on demand or automatically on a just-in-time basis. Such blended learning environments are currently within reach and raise the question as to how guidance should be distributed across teachers and adaptive software agents. Pedaste et al. (2016) envisioned a scenario where the two act in symbiosis: software agents could handle the more straightforward needs for guidance while the teacher takes care of the more intricate issues. Although software agents may eventually be capable of fully taking over the role of the teacher, this option is still beyond reach—and one may wonder whether the complete absence of a human teacher is in fact desirable for social and motivational reasons.

Another topic worthy of consideration is the domain-general nature of inquiry skills. Most models of inquiry-based learning are loosely based on the "scientific method": a highly generalized portrayal of how knowledge is constructed by the scientific community rather than a description of what individual researchers actually do. As research practices vary across scientific disciplines (Dunbar, 2000), students should be familiarized with these domain-specific idiosyncrasies. This argument is central to the Next Generation Science Standards (NGSS Lead States, 2013) in the US, but other countries still seem somewhat hesitant to embrace this science-as-practices vision. From an instructional perspective, one may wonder whether a disciplinary approach to scientific inquiry is always desirable. Young learners, for example, might benefit more from repeated engagement with the same, generic model of inquiry; the non-unitary perspective becomes relevant once a solid basic understanding is reached. Questions like these are discussed at length in the volume edited by Fisher et al (2018).

Many other works dealing with or touching on inquiry-based learning have been published over the past decade, and their disciplinary focus demonstrates that inquiry-based learning remains a timely topic in science education (e.g., Trundle and Saçkes, 2015; Tsivitanidou et al., 2018). The uptake in other content areas lags behind, but perhaps the emerging emphasis on disciplinary differences in scientific inquiry could serve as a catalyst for a broader application. Schools aspiring to increase the frequency of student inquiry throughout their curriculum are recommended to develop a coherent rationale for when and how the domain-general and domain-specific aspects of inquiry will be addressed, and establish a dialog between teachers of different subjects to materialize this vision in lesson materials. Initiatives like these will enable educational researchers and practitioners to join forces in working toward increasing the scientific literacy of our future generations.

References

- Affieri, L., Brooks, P.J., Aldrich, N.J., Tenenbaum, H.R., 2011. Does discovery based instruction enhance learning? *J. Educ. Psychol.* 103, 1–18.
- Ausubel, D.P., 1962. Learning by discovery. *Educ. Leader* 20, 113–117.
- Bell, R.L., Smetana, L., Binns, I., 2005. Simplifying inquiry instruction. *Sci. Teach.* 72, 30–33.
- Bruner, J.S., Goodnow, J.J., Austin, G.A., 1956. *A Study of Thinking*. Wiley, New York.
- Chinn, C.A., Malhotra, B.A., 2002. Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Sci. Educ.* 86, 175–218.
- Danielsson, A.T., Warwick, P., 2014. "You have to give them some science facts": primary student teachers' early negotiations of teacher identities in the intersections of discourses about science teaching and about primary teaching. *Res. Sci. Educ.* 44, 289–305.
- De Jong, T., Lazonder, A.W., 2014. The guided discovery learning principle in multimedia learning. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, New York, pp. 371–390.
- De Jong, T., Van Joolingen, W.R., Gienza, A., et al., 2010. Learning by creating and exchanging objects: the SCY experience. *Br. J. Educ. Technol.* 41, 909–921.
- Dekker, S., Van Baren-Navrocka, J., 2017. Wetenschappelijke doorbraken de klas in: molecuulbotsingen, stress en taal der zintuigen [Scientific breakthroughs in the classroom: molecular collisions, stress and the language of the senses]. Wetenskapsknooppunt Radboud Universiteit, Nijmegen.
- Dewey, J., 1900. *The School and Society*. The University of Chicago Press, Chicago.
- Dewey, J., 1938. *Experience and Education*. Kappa Delta Pi, New York.
- Dobber, M., Zwart, R., Tanis, M., Van Oers, B., 2017. Literature review: the role of the teacher in inquiry-based education. *Educ. Res. Rev.* 22, 194–214.
- Dunbar, K., 2000. How scientists think in the real world: implications for science education. *J. Appl. Dev. Psychol.* 21, 49–58.
- Eysink, T.H.S., De Jong, T., Berthold, K., Kollöffel, B., Opfermann, M., Wouters, P., 2009. Learner performance in multimedia learning arrangements: an analysis across instructional approaches. *Am. Educ. Res. J.* 46, 1107–1149.
- Eysink, T.H.S., Hulsbeek, M., Gijlers, H., 2017. Supporting primary school teachers in differentiating in the regular classroom. *Teach. Teach. Educ.* 66, 107–116.
- Fischer, F., Chinn, C.A., Engelman, K., Osborne, J., 2018. In: *Scientific Reasoning and Argumentation: The Roles of Domain-Specific and Domain-General Knowledge*. Routledge, New York.
- Furtak, E.M., Seidel, T., Iverson, H., Briggs, D.C., 2012. Experimental and quasiexperimental studies of inquiry-based science teaching: a meta-analysis. *Rev. Educ. Res.* 82, 300–329.
- Gaitas, S., Alves Martins, M., 2017. Teacher perceived difficulty in implementing differentiated instructional strategies in primary school. *Int. J. Incl. Educ.* 21, 544–556.
- Grol, R., De Vries, B., Sent, E.-M., 2016. Effects of economic classroom experiments on economic knowledge and reasoning in secondary education. *Think. Skills Creativ.* 22, 129–141.
- Hagemans, M.G., Van der Meij, H., De Jong, T., 2013. The effects of a concept map-based support tool on simulation-based inquiry learning. *J. Educ. Psychol.* 105, 1–24.
- Hmelo-Silver, C.E., Duncan, R.G., Chinn, C.A., 2007. Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). *Educ. Psychol.* 42, 99–107.
- Hmelo-Silver, C.E., Kapur, M., Hamstra, M., 2018. Learning through problem solving. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 210–220.

- Inhelder, B., Piaget, J., 1958. *The Growth of Logical Thinking from Childhood to Adolescence*. Basic Books, New York.
- Jerrim, J., Oliver, M., Sims, S., 2019. The relationship between inquiry-based teaching and students' achievement: new evidence from a longitudinal PISA study in England. *Learn. Instruct.* 61, 35–44.
- Kapici, H.O., Akcay, H., De Jong, T., 2019. Using hands-on and virtual laboratories alone or together—which works better for acquiring knowledge and skills? *J. Sci. Educ. Technol.* 28, 231–250.
- Kaya, S., Rice, D.C., 2010. Multilevel effects of student and classroom factors on elementary science achievement in five countries. *Int. J. Sci. Educ.* 32, 1337–1363.
- Kirschner, P.A., Sweller, J., Clark, R.E., 2006. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86.
- Klahr, D., Nigam, M., 2004. The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychol. Sci.* 15, 661–667.
- Koerber, S., Osterhaus, C., 2019. Individual differences in early scientific thinking: assessment, cognitive influences, and their relevance for science learning. *J. Cognit. Dev.* 20, 510–533.
- Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., Sodian, B., 2015. The development of scientific thinking in elementary school: a comprehensive inventory. *Child Dev.* 86, 327–336.
- Krajcik, J., Blumenfeld, P.C., Marx, R.W., Bass, K.M., Fredricks, J., Soloway, E., 1998. Inquiry in project-based science classrooms: initial attempts by middle school students. *J. Learn. Sci.* 7, 313–350.
- Kroeze, K.A., Van den Berg, S.M., Lazonder, A.W., Veldkamp, B.P., De Jong, T., 2019. Automated feedback can improve hypothesis quality. *Front. Educ. Sect. Educ. Psychol.* 3, 116.
- Kuhn, D., Katz, J., 2009. Are self-explanations always beneficial? *J. Exp. Child Psychol.* 103, 386–394.
- Kuhn, D., Pease, M., Wirkala, C., 2009. Coordinating the effects of multiple variables: a skill fundamental to scientific thinking. *J. Exp. Child Psychol.* 103, 268–284.
- Lapidow, E., Walker, C.M., 2020. Informative experimentation in intuitive science: children select and learn from their own causal interventions. *Cognition* 201, 104315.
- Lazonder, A.W., Ehrenhard, S., 2014. Relative effectiveness of physical and virtual manipulatives for conceptual change in science: how falling objects fall. *J. Comput. Assist. Learn.* 30, 110–120.
- Lazonder, A.W., Harmsen, R., 2016. Meta-analysis of inquiry-based learning: effects of guidance. *Rev. Educ. Res.* 86, 681–718.
- Lazonder, A.W., Kamp, E., 2012. Bit by bit or all at once? Splitting up the inquiry task to promote children's scientific reasoning. *Learn. Instruct.* 22, 458–464.
- Lazonder, A.W., Wilhelm, P., Hagemans, M.G., 2008. The influence of domain knowledge on strategy use during simulation-based inquiry learning. *Learn. Instruct.* 18, 580–592.
- Lazonder, A.W., Wilhelm, P., Van Lieburg, E., 2009. Unraveling the influence of domain knowledge during simulation-based inquiry learning. *Instr. Sci.* 37, 437–451.
- Lazonder, A.W., Janssen, N., Gijlers, H., Walraven, A., 2021. Patterns of development in children's scientific reasoning: results from a three-year longitudinal study. *J. Cogn. Dev.* 22, 108–124.
- Lazonder, A.W., 2018. *De jonge onderzoeker [The Young Investigator]*. Inaugural address, Radboud University, Nijmegen, the Netherlands.
- Martella, A.M., Klahr, D., Li, W., 2020. The relative effectiveness of different active learning implementations in teaching elementary school students how to design simple experiments. *J. Educ. Psychol.* 112, 1582–1596.
- Mastropieri, M.A., Scruggs, T.E., Norland, J.J., et al., 2006. Differentiated curriculum enhancement in inclusive middle school science: effects on classroom and high-stakes tests. *J. Spec. Educ.* 40, 130–137.
- Mayer, R.E., 2004. Should there be a three-strike rule against pure discovery learning? A case for guided methods of instruction. *Am. Psychol.* 59, 14–19.
- McCrea Simpkins, P., Mastropieri, M.A., Scruggs, T.E., 2009. Differentiated curriculum enhancements in inclusive fifth-grade science classes. *Remedial Special Educ.* 30, 300–308.
- Morgan, K., Brooks, D.W., 2012. Investigating a method of scaffolding student designed experiments. *J. Sci. Educ. Technol.* 21, 513–522.
- Mulder, Y.G., Lazonder, A.W., De Jong, T., 2010. Finding out how they find it out: an empirical analysis of inquiry learners' need for support. *Int. J. Sci. Educ.* 32, 2033–2053.
- Mulder, Y.G., Lazonder, A.W., De Jong, T., 2015. Key characteristics of successful science learning: the promise of learning by modelling. *J. Sci. Educ. Technol.* 24, 168–177.
- National Research Council, 2000. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. National Academy Press, Washington.
- Neubrand, C., Borzikowsky, C., Harms, U., 2016. Adaptive prompts for learning evolution with worked examples: highlighting the students between the 'novices' and the 'experts' in a classroom. *Int. J. Environ. Sci. Educ.* 11, 6774–6795.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, by States*, vol. 1. The National Academies Press, Washington. The Standards.
- Oliver, M., McConney, A., Woods-McConney, A., 2019. The efficacy of inquiry-based instruction in science: a comparative analysis of six countries using PISA 2015. *Res. Sci. Educ.* 49.
- Parsons, S.A., Vaughn, M., Scales, R.Q., et al., 2018. Teachers' instructional adaptations: a research synthesis. *Rev. Educ. Res.* 88, 205–242.
- Pedaste, M., Mäeots, M., Siiman, L.A., et al., 2015. Phases of inquiry-based learning: definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61.
- Pedaste, M., Lazonder, A.W., Raes, A., et al., 2016. Grand challenge problem 3: empowering science teachers using technology-enhanced scaffolding to improve inquiry learning. In: Eberle, J., Lund, K., Tchounikine, P., Fischer, F. (Eds.), *Grand Challenge Problems in Technology Enhanced Learning II: MOOCs and Beyond*. Springer, Cham, pp. 17–20.
- Piekny, J., Maehler, C., 2013. Scientific reasoning in early and middle childhood: the development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *Br. J. Dev. Psychol.* 31, 153–179.
- Reisman, A., 2012. Reading like a historian: a document-based history curriculum intervention in urban high schools. *Cognit. Instruct.* 30, 86–112.
- Rönnebeck, S., Bernholt, S., Ropohl, M., 2016. Searching for a common ground: a literature review of empirical research on scientific inquiry activities. *Stud. Sci. Educ.* 52, 161–197.
- Schlatter, E., Molenaar, I., Lazonder, A.W., 2020. Individual differences in children's development of scientific reasoning through inquiry-based instruction: who needs additional guidance? *Front. Psychol. Sect. Dev. Psychol.* 11, 904.
- Tomlinson, C.A., 2001. *How to Differentiate Instruction in Mixed-Ability Classrooms*, second ed. Association for supervision and curriculum development, Alexandria.
- Trundle, K.C., Saçkes, M., 2015. *Research in Early Childhood Science Education*. Springer Science+Business Media, Dordrecht.
- Tsivitanidou, O.E., Gray, P., Rybska, E., Louca, L., Constantinou, C.P., 2018. *Professional Development for Inquiry-Based Science Teaching and Learning*. Springer, Cham.
- Van der Graaf, J., Segers, E., Verhoeven, L., 2018. Individual differences in the development of scientific thinking in kindergarten. *Learn. Instruct.* 56, 1–9.
- Voet, M., De Wever, B., 2017. Preparing pre-service history teachers for organizing inquiry-based learning: the effects of an introductory training program. *Teach. Teach. Educ.* 63, 206–217.
- Vogt, F., Rogalla, M., 2009. Developing adaptive teaching competency through coaching. *Teach. Teach. Educ.* 25, 1051–1060.
- Vorholzer, A., Von Aufschnaiter, C., 2019. Guidance in inquiry based instruction: an attempt to disentangle a manifold construct. *Int. J. Sci. Educ.* 41, 1562–1577.
- Yoon, H.-G., Joong, Y.J., Kim, M., 2012. The challenges of science inquiry teaching for pre-service teachers in elementary classrooms: difficulties on and under the scene. *Res. Sci. Educ.* 42, 589–608.
- Zacharia, Z.G., 2015. Examining whether touch sensory feedback is necessary for science learning through experimentation: a literature review of two different lines of research across K-16. *Educ. Res. Rev.* 16, 116–137.
- Zhang, L., 2016. Is inquiry-based science teaching worth the effort? *Sci. Educ.* 25, 897–915.

Intelligent tutoring systems and conversational agents

Arthur C. Graesser^a and Haiying Li^b, ^a University of Memphis, Memphis, TN, United States; and ^b Iowa State University, Ames, IA, United States

© 2023 Elsevier Ltd. All rights reserved.

Intelligent tutoring systems	638
ITS mechanisms	638
Performance of ITS in improving learning	639
Example ITS systems without conversational agents	639
Learning environments with conversational agents	639
Conversational patterns in human and computer tutors	640
Example systems with conversational agents	641
AutoTutor	641
Inq-ITS	644
Conversational agents in an informal learning setting	644
Challenges for the future	645
References	645

Computer-based training (CBT) systems are digital technologies that range from digital books that are read by learners at their own pace to adaptive systems that follow simple learning principles at a course-grain level. For example, in a typical adaptive CBT that applies mastery learning, the student (a) studies material presented in a lesson, (b) gets tested with a multiple-choice test or other objective test, (c) receives feedback on the test performance, (d) re-studies the material if the performance is below a mastery threshold, and (e) progresses to a new topic when done.

Intelligent tutoring systems (ITSs) are more advanced digital environments that help students master dozens-to-hundreds of specific knowledge and skill components by implementing intelligent computational procedures that adapt to students at a fine-grained level and that instantiate complex principles of learning (Graesser et al., 2018; VanLehn, 2006; Woolf, 2009). ITSs have traditionally worked with one student at a time in order to adapt to the idiosyncrasies of individual learners. However, recent ITS efforts have included team training that goes beyond a single individual to small groups (Sottolare et al., 2018a).

Conversational agents have been developed during the last 30 years to serve as substitutes for human tutors, teachers, peers, and students. Computer agents attempt to simulate text, speech, gestures, body movements, and/or facial expressions in ways similar to people, but may go beyond what humans do by implementing pedagogical mechanisms that humans cannot or do not provide (Graesser, 2020; Johnson and Lester, 2016). The agents are programmed to apply learning science principles systematically at varying degrees of complexity and grain size, with more reliability than humans. Inflexible conversational agents are aligned with CBT instruction when they present information in a rigid choreographed sequence and model ideal behavior, strategies, reflections, and social interactions. In contrast, intelligent agents are flexible because their behavior is adaptive to the learners' actions, verbal contributions, emotions, and/or other psychological states. For example, *AutoTutor* helps students learn by holding a dialogue in natural language (Graesser, 2016; Nye et al., 2014) or engaging in a three-way "trialogue" that has two agents (a student and a peer agent) interacting with the human learner (Graesser et al., 2017). *AutoTutor* conversations that analyze the meaning of students' verbal contributions are unique for each student.

This chapter covers advances in ITSs and conversational agents, with an emphasis on the intersection between these two technologies. Together these technologies have the promise of meeting the needs of education both in the past and into the future. The report on *How People Learn II* (HPL-II) of the National Academy of Sciences, Engineering, and Learning (NASEM, 2018) emphasized the uniqueness of individual learners in their background knowledge, culture, and contexts of learning. Therefore, educational activities need to shift from one-size-fits all to individualized instruction. The HPL-II report also emphasized the importance of digital technologies to provide broad and deep personalized instruction. The ITS environments provide this opportunity. Human instructors will of course be a major part of the future educational ecosystem as they have in the past. Most training over previous millennia has been one-on-one through an apprenticeship framework with conversations between the learner and a master, mentor, or tutor (Collins and Halverson, 2009). Print became popular after the printing press, particularly during the industrial revolution, because of its efficiency in dissemination and training. Nevertheless, one-on-one communication has continued to be mainstream. Conversational agents provide some facsimile of similarity to human interaction, and thereby fill this need when access to high quality human instruction is limited.

Intelligent tutoring systems

All ITSs adopt the principle of active student learning rather than resorting to the mere delivery of information through lectures, films, and books (Chi and Wylie, 2014; Lesgold, 2019). However, there is tremendous variation in the learning principles that are implemented in the different ITSs. One way of analyzing and classifying ITSs is to identify the “technological affordances” of the learning environments that were identified in the HPL-II report (NASEM, 2018). An affordance is a feature or property of the digital system that makes possible a particular way for the learner to use it. The following affordances nearly always occur in an ITS:

Interactivity. The system responds to actions of the student contingently, that is, it responds differently depending on what the student’s actions are.

Adaptivity. The system presents information and activities that are contingent on the knowledge, skills, abilities and other psychological characteristics of the student.

Feedback. The system gives feedback to the student on the quality of the student’s responses, including qualitative information that improves a correct response or corrects an error.

The following affordances frequently occur in ITSs but not always:

Choice. The system gives students options on what to learn in order to improve motivation and encourage self-regulation of learning.

Nonlinear access. The system allows the student to select or receive learning activities in an order that deviates from a rigid curriculum order.

Linked representations. The system provides quick connections between representations that emphasize different conceptual viewpoints and media, such as verbal descriptions, symbolic formulae, diagrams, tables, and simulations.

Open-ended learner input. The system allows the students to express themselves through natural language, drawing pictures, and other forms of open-ended communication.

One affordance that is comparatively rare in ITSs is *communication with other people*, such that the student communicates with one or more other people who are either peers or experts on the subject matter. However, some ITSs have an option that allows the student to have a chat interaction when they get stuck, such as www.tutor.com being used with a math tutor, with hundreds of human tutors available for chat interactions. Also, Hausmann et al. (2008) set up their ITS in physics with two students interacting with the ITS (without an agent) and showed benefits of the dyad over single student interactions. However, these two examples are exceptions rather than a frequent affordance of ITSs. Once again, however, there are many ITSs with conversational agents that serve as substitutes for communication with other people and thereby satisfy this affordance.

ITS mechanisms

The *Generalized Intelligent Framework for Tutoring* (GIFT; www.gifttutoring.org; Sottolare et al., 2018b) has articulated the frequent practices, standards (both pedagogical and technical), and computational mechanisms in ITS development. A framework is needed to scale up ITS for greater use in schools, industry, government agencies and the public. The GIFT initiative has a free publication series (in www.gifttutoring.org) with 20–30 experts contributing annually to particular thematic topics: Learner modeling, instructional strategies, authoring tools, domain modeling, assessment, team learning, self-improving systems, data visualization, and proficiency-based scenario design.

GIFT describes characteristics of the primary and frequent ITS components as well as more exotic components that are used infrequently. The *Learner Record* automatically stores the knowledge, skills, abilities (abbreviated as KSA’s) and other characteristics of each particular student that is collected before, during, and after the student interacts with the ITS. This is a rich history of interactions stored in log files, performance scores on tests, mastery of specific knowledge components, and often noncognitive characteristics (such as persistence, initiative, and particular emotions). The volume of data is far greater than what a human instructor could ever collect. A *Sensor* module senses and encodes multimodal input (if there is any) such as physiological states, geographical location, facial expressions, speech, and gestures. A *Pedagogical* module is the set of tutoring strategies and tactics that apply to particular types of knowledge, skills, and subject matter. A *Domain* module specifies the knowledge components, procedures, and other content associated with the subject matter being tutored. This content can range from texts and PowerPoint presentations to virtual reality and augmented reality environments.

One computational construct that is popular in the domain module of ITSs is a *knowledge component* (Koedinger et al., 2012). Intuitively, a knowledge component (KC) is a bite-size piece of information or skill that can be tracked and assessed in different contexts. For example, consider Newton’s second law of motion (net force = mass X acceleration). One could track how well a student’s performance reflects an understanding of this law in different contexts, such as answering a word problem, computing the speed of an object, or interacting with a physics simulation. The student might fail in one context but succeed in another. Newton’s second law is likely to involve a set of more specific KC’s, which researchers attempt to resolve through empirical research and debate. The important point is that such KC’s are fine grained and can be tracked over the student’s learning history. A *Q-matrix* is constructed for all of the tasks and problems in an ITS activity that specifies what KC’s are involved (there can be more than one for a problem) and what scores (0–1, continuous or discrete) to assign in the Learner Record store for each of the KC’s. New

problems and learning tasks can then be selected intelligently (through a *Recommender System*) that fills in missing KC's at the frontier (slightly beyond) what the student has mastered and gives opportunities to correct prior errors. It is important to acknowledge that the KC's are a major unit of learning in this approach. There are typically dozens to hundreds of KC's associated with a subject matter. This is a very different grain size than the traditional approach in education that assesses performance on a lesson or items on a test.

One limitation of the KC approach is that it assumes that there are correct answers that can be precisely specified. Some subject matters are more ill-defined, yet informative assessment and feedback can still be provided. For example, verbal answers to open-ended questions and essay grading have advanced considerably and are scored quite accurately, on par with human experts in many cases (Yan et al., 2020). Successful systems sometimes have no KC's. There is a rubric for assessing the quality of student contributions but no KC's per se.

A second computational construct that is popular in the Pedagogical module is the *production rule*. This consists of an IF <state> THEN <action> expression that specifies what the ITS should do under specific conditions (Anderson et al., 1995; Laird, 2012). The ITS watches over the landscape of current states in the Learner Record and Sensor modules. If a particular state exists or reaches some threshold of activation, a production rule is fired probabilistically. Contemporary rules need not be brittle, but rather are activated to some degree and probabilistically. They also are executed at any grain size, ranging from a very specific context (e.g., computing fractions with negative numbers) to an abstract generalization (e.g., suggestion to spending more time comprehending the lessons).

Performance of ITS in improving learning

Reviews and meta-analyses have shown that ITS technologies frequently improve learning compared to conditions in which students study text or experience classroom lectures and other activities, with some attempts to control for learning time. These meta-analyses normally report effect sizes (*d*) to convey differences between the ITS condition and a comparison condition in standard deviation units. The reported meta-analyses show positive effect sizes that vary from 0.05 (Dynarsky et al., 2007) to 1.08 (Dodds and Fletcher, 2004), but most hover between 0.40 and 0.80 (Kulik and Fletcher, 2015; Ma et al., 2014; VanLehn, 2011). A 0.60 performance is comparable to human tutoring which varies between 0.20 and 0.80 (Cohen et al., 1982; VanLehn, 2011), depending on the expertise of the tutor. Human tutors have not varied greatly from ITSs in terms of effect size in direct comparisons between ITSs and trained human tutors (Graesser, 2016; Olney et al., 2012; VanLehn, 2011).

Example ITS systems without conversational agents

ITSs have frequently been developed for mathematics and other well-formed subject matters in which there are objectively correct answers to questions. For example, *MATHia* (formerly *Cognitive Tutors*, Alevan et al., 2009; Anderson et al., 1995; Ritter et al., 2007), and *ALEKS* (Falmagne et al., 2013) cover basic mathematics, algebra, geometry, statistics, and more advanced quantitative skills. ITSs exist in other STEM areas, such as physics (Van Lehn et al., 2005) and information technology (Mitrovic et al., 2007). ITSs also cover topics that have a stronger verbal foundation as opposed to mathematics and precise analytical reasoning (Johnson and Lester, 2016). These are covered in the next section that focuses on ITSs with conversational agents.

Learning environments with conversational agents

Conversational agents may be rigidly scripted or may be highly interactive with conversations in natural language. Sometimes a scripted interaction is fine, as when students who have such low domain knowledge that they can only observe and not generate anything. Sometimes intelligent, adaptive, interactivity is crucial so that the student gets the right information at the right time.

There are many functions of conversational agents that should be considered in making decisions on the selection of a learning environment. A single agent conversing with a student can have different roles (mentor, tutor, peer, student to teach, adversary, help assistant, etc.), personalities, and persona (Baylor and Kim, 2005). An agent can be assertive and barge into conversations or wait to be called. The communication can range from simple chat without any persona image to 3D avatars in virtual environments (Johnson and Valente, 2009). Adding a second agent can dramatically increase the power of agents, as in the case of dialogues with a student interacting with a peer agent. Graesser et al. (2017) have identified various designs of agents in dialogues such as those below.

- (1) *Vicarious learning with student observer*. Two agents interact and model ideal social and verbal interactions, answers to questions, or reasoning. Craig et al. (2012) reported the benefits of students observing peer agents asking deep questions and receiving answers from tutor agents on physics and computer science domains.
- (2) *Vicarious learning with limited student participation*. The same as #1 except that the agents occasionally turn to the student and ask a prompt question, with a yes/no or single word answer. The question that invites an answer maintains human engagement and assesses student performance.
- (3) *Tutor agent interacting with student and peer agent*. There is a tutorial dialogue with the student, but the peer agent periodically contributes and receives feedback. When the student has low self-esteem, negative short feedback can be given to the peer agent

on bad answers (the agent takes the heat) whereas similar answers by the student receive neutral feedback. This triologue design has been applied in studies with AutoTutor for struggling adult learners who have limited ability to type verbal input (Graesser et al., 2019).

- (4) *Expert agent staging a competition between the student and a peer agent.* There is a competitive game (with score points) between the student and peer agent, with the competition guided by the expert agent. This would be appropriate for students who are motivated by competition. Jackson and McNamara (2013) used this triologue design in their instructional game iSTART-ME to improve reading comprehension on science topics.
- (5) *Student teaches/helps a student agent with facilitation from the tutor agent.* This is a “teachable agent” design (learning by teaching), with the tutor agent rescuing problematic interactions. More advanced students who can generate material might particularly benefit from this design. Biswas et al. (2016) used this design in *Betty’s Brain* on topics in biology.
- (6) *Student interacts with two peer agents that vary in proficiency.* The peer agents can vary in knowledge and skills. The Educational Testing Service is using this approach in conversation-based assessment (Jackson and Zapata-Rivera, 2015). The Program for International Student Assessment on collaborative problem solving (OECD, 2017) used this design to track whether the student corrects an incorrect contribution by a peer, correctly answers a peer question, and takes initiative in guiding the exchange.
- (7) *Student interacts with two agents expressing contradictions, arguments, or different views.* The discrepancies between agents stimulate cognitive disequilibrium, confusion, and potentially deeper learning. D’Mello et al. (2014) used this design to demonstrate that contradictions triggered cognitive disequilibrium in some students who experienced sufficient confusion that they resolved the contradiction and showed improvements in learning.

These examples of dialogues illustrate the potential of conversational agents in manipulating contexts and social situations in learning environments. There is an exciting future in research on conversational agents that go beyond two agents (i.e., including many agents), with each taking on different roles and functions.

Conversational patterns in human and computer tutors

Some agent-based ITSs have been influenced by systematic analyses of the discourse and pedagogical strategies of human tutors who vary from novices to experts (Chi et al., 2001; D’Mello et al., 2010; Graesser et al., 2009). Effective tutoring goes beyond lecturing and into the realm of guiding learners in a process of co-constructing solutions to problems, answering challenging questions, and designing projects. VanLehn (2006) has distinguished between the outer loop and inner loop in ITS structure. The outer loop is the set of major tasks for the tutor and student to work on whereas the *inner loop* consists of the steps that attempt to make progress within a major task. The inner loop of many human tutoring sessions and ITSs follows systematic patterns of interaction. One of these is the *5-step dialogue frame* (Graesser et al., 2009). Once a major task is selected, the 5-step tutoring frame is initiated with the following main steps: (1) the tutor or student present the main task, (2) the student generates an initial attempt to handle the task, (3) the tutor gives short feedback on the quality of the answer, (4) the tutor and student improve the quality of the answer through interaction, and (5) the tutor assesses whether the student understands the correct answer and follows up if necessary. Interestingly, classroom teaching normally (though not always!) involves the first three steps (Sinclair and Coulthart, 1975) but rarely steps 4 and 5.

An important question among ITS researchers is how to optimize step 4? AutoTutor (Graesser, 2016) and most human tutors (Graesser et al., 2009; D’Mello et al., 2010) apply an *expectation and misconception tailored dialogue*. The ITS typically has a list of *expectations* (anticipated good answers, steps in a procedure) and a list of anticipated *misconceptions* (errors or bugs) associated with each task. The ITS guides the student in generating the expectations through a number of dialogue moves: *pumps*, *hints*, and *prompts* for the student to fill in missing information. A pump is a move to get the student to provide more information, such as “What else?” or “Tell me more.” Hints and prompts are questions selected by the ITS to get the student to make decisions, generate content at a step, or articulate missing words, phrases, and propositions in the expectations. As the student generates contributions over many interactional turns and the tutor guides them or directly fills in missing content, the list of expectations is eventually covered and the task is completed. The ITS keeps track of how well the student does in covering expectations and not having misconceptions. The selection of the next task in the outer loop depends on the student’s profile in the prior history of performance on tasks in tutorial sessions.

Human tutors and ITSs give feedback to the student on their performance. There is both short feedback (positive, negative or neutral) and often qualitative explanations that justify correct versus incorrect answers. A typical ITS turn in a dialogue interaction has three components:

Tutor Turn → Short Feedback + Dialogue Advancer + Floor Shift

The short feedback addresses the quality of the student’s prior turn whereas the dialogue advancer gives either qualitative feedback or moves the tutoring agenda forward with hints, prompts, requests, or information to fill gaps. The third component, floor shift, provides a cue (e.g., question, gesture, facial expression) that it is the student’s turn in the exchange.

Most ITS attempt to accommodate a mixed-initiative dialogue by allowing students to ask questions, ask for help, and select tasks to work on. However, it is difficult or impossible for an ITS to anticipate all of the student questions, requests for help, or tasks to work on. This difficulty also applies to human tutoring to the extent that the tutor does not understand the student’s questions or

requests. When student initiatives are not understood or meaningfully taken up by the tutor, there is the risk of the student giving up trying to initiate questions and requests. This is one explanation of why the frequency of student questions and requests for help is low, even when students are encouraged to take the initiative (Graesser et al., 2009).

Example systems with conversational agents

Table 1 shows the highlights of 22 studies of ITS with conversational agents. All of these studies report empirical evidence on whether or not there are benefits from these agent-based ITSs over various comparison conditions. There are many different outcome measures in assessments of the systems' added value, such as subject matter learning gains, scores on psychometric tests of achievement, engagement with the system, self-reported impressions of the system, and other psychological attributes. It is beyond the scope of this chapter to summarize the results, many of which are subtle and involve aptitude-treatment-content interactions. Instead, we emphasize the broad scope of subject matters, learners, and pedagogical functions of these agent-based ITSs.

The main claim we wish to make from the systems in **Table 1** is that there is a broad scope of subject matters, learners, and pedagogical functions of the agents. The subject matters include mathematics, science, computer literacy, computer programming, electronics, human-computer interaction, and reading comprehension. There are also other systems not included in the table that address cultural practices. For example, the *Tactical Language and Culture Training System* has adults learn languages and cultural practices through virtual worlds with avatars (Johnson and Valente, 2009) and has been used by thousands of US soldiers deployed in Iraq. The learner age groups vary also, starting with children ages 3–6 in storybook reading, and encompassing K12 (elementary, middle, and high school), college, the military, and the workforce. The pedagogical functions run the gamut from constructing mental models of STEM topics to self-regulated learning and emotions. Given the diverse scope of these developed systems, we are convinced that agent-based ITSs could fill out the curriculum of schools, the government, the workforce, and extended education for citizens throughout the world. The combination of intelligent systems and human-like agents is at the very least a promising direction. At this point we will describe a few systems in more detail for readers who desire a more nuanced understanding of the nature of these agent-based ITSs.

AutoTutor

AutoTutor (Graesser et al., 2004; Graesser, 2016; Nye et al., 2014) was developed to simulate the discourse and pedagogical strategies of tutors and thereby help students learn by holding a conversation in natural language. AutoTutor goes beyond didactic lectures and into the realm of presenting problems to solve and difficult questions to answer that require a multi-turn conversation. AutoTutor implements the 5-step tutoring frame, the expectation and misconception-tailored dialogue mechanism, and the tutor turn composition that were described earlier. A list of *expectations* (anticipated good answers, steps in a procedure) and a list of anticipated *misconceptions* (bad answers, incorrect beliefs, errors, bugs) is associated with each task. As the students articulate their answers over multiple conversational turns, the contents of their contributions are compared with the expectations and misconceptions. The students' answers within each turn are typically short (one to two speech acts on average), vague, ungrammatical, and not semantically well-formed. AutoTutor compares the students' content to expectations and misconceptions through pattern matching processes with semantic evaluation mechanisms that incorporate research in computational linguistics. Examples of AutoTutor conversations can be found in on-line venues on the internet (sites.autotutor.org).

Listed below are tutor dialogue moves that frequently occur in the Expectation and Misconception Tailored dialogues in AutoTutor and in most human tutoring sessions.

Main question or problem. This is a challenging question or problem that the tutor is trying to help the student answer.

Short feedback. The feedback to a student's previous contribution is either *positive* ("yes," "correct," head nod), *negative* ("no," "almost," head shake, long pause, frown), or *neutral* ("uh huh," "okay").

Pumps. The tutor gives nondirective pumps ("Anything else?" "Tell me more.") to say more or to take some action.

Hints. The tutor gives hints to get the students to do the talking or take action, but directs the students along some conceptual path.

The hints vary from generic statements or questions ("What about X?", "Why?") to speech acts that nudge the student toward a particular answer.

Prompts. The tutor asks a leading question to get the student to articulate a particular word or phrase. Sometimes students say very little so these prompts are needed to get the student to say something specific.

Prompt completions. The tutor expresses the correct completion of a prompt.

Assertions. The tutor expresses a fact or state of affairs.

Summaries. The tutor gives a recap of the answer to the main question or solution to the problem.

Mini-lectures. The tutor expresses didactic content on a particular topic.

Corrections. The tutor corrects an error or misconception of the student.

Answers. The tutor answers a question asked by the student.

Table 1 Systems with conversational agents.

<i>Authors</i>	<i>System</i>	<i>Subject matters</i>	<i>Learners</i>	<i>Pedagogical function</i>
Adamson et al. (2014)	Bazaar: Academically Productive Talk (APT) agents	Biology, chemistry, engineering design	High school and undergraduate students	APT conversational agent supports on-line collaborative learning discussions
Azevedo et al. (2010)	MetaTutor	Science	High school and college students	Using agents to train students on developing metacognitive and self-regulated learning strategies while exploring information with hypermedia
Biswas et al. (2016)	Betty's brain	Science	Middle school students	Leveraging the learning-by-teaching paradigm by having the student teach a peer agent, with guidance from a coach agent
Crockett et al. (2017)	OSCAR-CITS (Conversational Intelligent Tutoring System)	Structured Query Language (SQL)	Undergraduate students	Predicting learning styles based on conversational features in an ITS
D'Mello and Graesser (2012b)	AutoTutor	Computer literacy	Undergraduate psychology students	Analyzing emotions (boredom, frustration, confusion, flow) and dynamic transitions between emotions during AutoTutor sessions
Forsyth et al. (2020)	Operation ARA	Research methodology	Undergraduate psychology students	Examining predictors of shallow versus deep learning in a game with multiple agents and natural language interaction
Geden et al. (2020)	Crystal Island	Microbiology	Middle school students	Predicting learning outcomes from actions in a virtual world with agents and natural language responses to in-game reflection prompts
Graesser et al. (2004)	AutoTutor	Newtonian qualitative physics and computer literacy	College students	AutoTutor helps students learn by holding a conversation in natural language through questions, hints, feedback, corrections and other tutor dialogue moves
Hampton et al. (2018)	PAL3 (Personal Assistant for Life-Long Learning)	Electronic circuits	Navy sailor trainees	Knowledge decay is minimized by a personal conversational agent that gives the trainee a daily electronics problem that maintains particular knowledge components
Hutt et al. (in press)	Guru	Biology	High school students	An ITS with an agent attempts to detect student mind wandering during learning and reorient the student's attention to the material and thereby improve learning and retention
Jackson and McNamara (2013)	iSTART (Interactive Strategy Training for Active Reading and Thinking)	Reading comprehension	High school students	Improving comprehension through self-explanation strategies with multiple iSTART agents, including one version being motivationally enhanced through gamification
Jackson et al. (2018)	CBA (Conversation-Based Assessment)	Volcanos	Middle school students	Improving assessment of science knowledge by agent trialogues that implement CBA compared with a constructed response approach
Meyer et al. (2011)	ITSS (Intelligent Tutoring System for the Structure Strategy)	Reading comprehension	5th grade students	Improving comprehension with an agent that individualizes reading instruction focusing on the rhetorical structure of texts

Table 1 Systems with conversational agents.—cont'd

<i>Authors</i>	<i>System</i>	<i>Subject matters</i>	<i>Learners</i>	<i>Pedagogical function</i>
Olney et al. (2012)	Guru	Biology	High school students	An ITS with an agent is compared with human tutors and no tutoring in some high school classes.
Snow et al. (2015)	iSTART (Interactive Strategy Training for Active Reading and Thinking)	Constructing self-explanations	College students	Exploring the importance of giving students choices (i.e., agency) in the game version of iSTART that trains students how to construct self-explanations during reading
Tärning et al. (2019)	The math game	Math	10–11 year old students	Testing the effects of the self-efficacy expressed by a teachable agent on students in a math game.
Taub et al. (2020)	Crystal Island	Microbiology	College students	Testing the impact of agency on learning, problem solving, and affect when learning with Crystal Island, a game with agents in a virtual world
Tegos et al. (2016)	MentorChat	Principles of human-computer interaction	Undergraduate computer science students	Testing the impact of an agent that mentors them on academically productive talk in dyads in collaborative activities
VanLehn et al. (2007)	Why2-Atlas, Why2-AutoTutor	Physics	College students	Comparing conversational interaction with agents and reading texts, with an eye to the difficulty of the material and zone of proximal development
Ward et al. (2013)	MyScience tutor	Science	Elementary school students	Agents help you children learn science through a virtual tutor in multimedia environments.
Wijekumar et al. (2017)	ITSS (Intelligent Tutoring System for the Structure Strategy)	Reading comprehension	5th grade students	Improving comprehension on both standardized tests and experimenter-designed measures with an agent-based intervention that focuses on the rhetorical structure of texts
Xu et al. (2021)	Google home mini	Storybook reading	Children aged 3–6	An agent guides conversation during story reading to improve comprehension, reading skills, and engagement

AutoTutor has shown learning gains of approximately 0.80 standard deviations compared with reading a textbook for an equivalent amount of time (Graesser, 2016). There are also different versions of AutoTutor that go beyond collaborative conversations on the subject content. One version of AutoTutor scaffolded students to take actions in a simulation world in addition to prompting verbal contributions of the students (Graesser et al., 2005). AutoTutor 3-D had the conversational tutor interact with students in a simulation on Newtonian physics problems. An example main question is: *When a car without headrests on the seats is struck from behind, the passengers often suffer neck injuries. Why do passengers suffer neck injuries?* The students are encouraged to manipulate values of variables (such as the mass, speed, and direction of vehicles) to see what happens when collisions occur in a 3-D world. The interactive simulation is a resource to help them answer the question verbally and to correct their answers when AutoTutor suggests they have misconceptions. Ideally, the students would spontaneously enact the Vary-One-Thing-at-a-Time (VOTAT) strategy in which values of one variable are varied while other variables are held constant (Greiff et al., 2013). VOTAT requires multiple runs of the simulation with systematic changes of the values and variables in order to construct a deep analysis of the system. Unfortunately, most students ran the simulation once and did not enact the VOTAT strategy. This is where an agent can step in and guide them to run multiple simulations and see the results of each simulation.

Another version of AutoTutor is sensitive to the learner's affect states, namely boredom, frustration, confusion, and flow (D'Mello and Graesser, 2012a). An empathetic AutoTutor helps learning for low knowledge students when they are frustrated or experiencing other forms of negative affect. ITS researchers have been sensitive to learner's affective states in addition to their states of knowledge and cold cognition (Arroyo et al., 2014; Baker et al., 2010; DeFalco et al., 2018).

Inq-ITS

Inq-ITS (Inquiry Intelligent Tutoring System; www.inqits.com) is a simulation-based intelligent tutoring and assessment system that supports teaching, learning, and assessment of scientific inquiry practices in the domain of life, earth, and physical science (Gobert et al., 2013). These scientific practices are aligned with the Next Generation Science Standards (NGSS Lead States, 2013) and a national framework for K-12 science education (NRC, 2012) with the goal of helping middle school students to develop key inquiry skills as scientists do. Students practice five primary scientific inquiry practices within Inq-ITS: asking a question/forming a hypothesis, collecting data to test it, interpreting data, warranting the claim with data, and communicating findings via written claim, evidence, and reasoning statements. The first four practices involve practicing science, whereas the last one involves writing about science. A pedagogical agent, Rex (a cartoon dinosaur), offers students real-time, personalized scaffolding that adapts to the quality of students' inquiry practices as they engage in virtual labs (Gobert et al., 2018; Li et al., 2018). The teacher dashboard Inq-Blotter, a mobile alert system, informs teachers of students' performance on inquiry practices and provides them instant alerts on which students need one-on-one support most and when whole-class instruction is necessary.

The implementation of personalized assistance within Inq-ITS and teacher's alerts within Inq-Blotter depends on the automated evaluation of students' science inquiry practices by leveraging techniques from educational data mining, knowledge engineering, and natural language processing. The computational procedures and algorithms are employed to automatically score students' actions and interactions within each practice with robust features that were developed through manually scoring log files and automated training assessment models. For example, these models can identify an independent variable (IV), identify a dependent variable (DV), and interpret/warrant the IV/DV and hypothesis/claim relationship. The automated scoring has been shown to robustly match human scores with high precision, ranging from 84% to 99% accuracy among diverse groups of students (Gobert et al., 2018).

Inq-ITS has also applied pattern match NLP techniques and if-then algorithms to automatically score the communicated findings involving written claim, evidence, and reasoning statements in the open-ended response format (Li et al., 2017a,b). Specifically, regular expressions are used to determine the presence or absence of key concepts or complex phrases within written responses. Regular expressions identify alternative forms of words and expressions including misspellings (e.g., /indic/capturing different forms of indicate, i.e., *indicate*, *indicates*, *indicated*, *indicating*, *indication*, *indicote*, and *inducution*) and syntactic precision of wording (e.g., *I increased the temperature so it increased the number of animals with green short furs* vs. *I increased the number of animals with green short fur so it increased the temperature*). The automated scoring is highly correlated with human scores with r equal to 0.94, 0.97, and 0.84 for written claim, evidence, and reasoning, respectively (Li et al., 2017a,b).

The automated assessment of students' practices allows for real-time adaptive scaffolds (Gobert et al., 2018; Li et al., 2018) in addition to building the teacher dashboard. The pedagogical agent (Rex) provides students different types of scaffolds based on their performance, consisting of orienting scaffolds (i.e., directing attention to a particular step), conceptual scaffolds (i.e., explaining an inquiry practice required for the step), procedural scaffolds (i.e., informing students of the procedure for the step), and instrumental scaffolds (i.e., guiding students how to move forward on a step). The adaptive scaffolds effectively improve students' inquiry practice competencies, with both near and far transfer of these practices over time and across inquiry topics with effect sizes ranging from 0.18 to 1.80 (Gobert et al., 2018; Li et al., 2018).

Conversational agents in an informal learning setting

The increasing emergence of digital devices, particularly mobile technology, facilitates an upsurge of informal learning that is not constrained to learning at schools. The museum is a primary location where students become motivated and engaged in science and technology. If a museum guide can enhance the interactive exchange with visitors rather than merely present information, the visitors will gain a deeper understanding of museum content (Swartout et al., 2013). Due to the short supply of human museum guides, the virtual human guide is a potential museum guide to improve the quality of interactions between visitors and the guide. Swartout and colleagues have investigated the efficacy of virtual museum guides in interacting with visitors, engaging visitors, and increasing their understanding of the museum content (Swartout et al., 2013) in the *InterFaces* project, a collaboration between the University of Southern California and the Museum of Science in Boston. *InterFaces* is an exhibit where virtual human guides, called Ada and Grace, are housed and designed to promote interest and engagement in STEM topics, targeting children between ages 7 and 14. Ada and Grace are twins designed to increase engagement using the following four elements: (1) an appealing appearance of a young adult female with an indeterminate racial background that the museum visitors selected with traits of being friendly, smart, and patient, (2) near photoreal appearance of the twins that are created using light stage technology and digital artists to appear and animate realistically, (3) generation of responses to questions posed by museum visitors to better maintain the pace of the conversation and visitor interest, and (4) natural language understanding through the acoustic speech and language models that are customized to individual staff members' voices to increase the speech recognition accuracy. Evaluation of natural language performance showed that 70% of the public's spoken utterances to the characters matched expected utterances by the computer system. Among the known utterances, over 80% were answered correctly, with failures typically attributed to poor speech recognition. Aside from the accuracy of answers to questions, the Ada and Grace twins were impressive in both entertaining and motivating children to learn about science.

Challenges for the future

In this chapter we have articulated the need for ITS, particularly those with conversational agents. We have identified many successful ITS learning environments that are backed by empirical support—they have had a positive impact on learning, engagement, and/or impressions on the learner. However, the picture is not uniformly rosy. There are limitations and challenges to face in the future for those who want to build, test, and adopt these digital learning environments. This section identifies the major obstacles that have prevented ITS from reaching many millions of learners throughout the globe. Perhaps these obstacles will be solved in the next generation.

One obstacle is that it has historically taken a large amount of time and money to build ITSs because of the complexity of the mechanisms and the limited number of competent experts to build them. The GIFT community (Sottilare et al., 2018a,b; gifttutoring.org) has been attempting to reduce the cost and time through modularity, standards, and better authoring tools. However, the cost is still measured in millions of dollars and the time is measured in years. Once a good system is built, it can be disseminated to millions of people, which dramatically reduces the cost per learner. However, this requires capital investments and marketing, which has not yet occurred for ITSs at the volume of other digital learning technologies, such as Massively Open Online Learning (MOOC) systems (Siemens et al., 2015).

A second obstacle is that it is difficult to develop authoring tools that can be productively used by individuals without expertise in computer science. The ideal person or R&D team would have expertise in the subject matter knowledge, intelligent tutoring systems, conversational agents, discourse processing, and pedagogy—a rare combination of skills to assemble.

A third obstacle is uncertainty as to how much human intervention and support is needed to implement ITSs (either with or without conversational agents) in real world contexts. Instructors are increasingly turning to *blended learning* environments in which the instructor attempts to weave in these sophisticated technologies like ITSs (Lesgold, 2019; Siemens et al., 2015). Unfortunately, we are at a point in history when the existing professional development (i.e., training of teachers) is not keeping pace with advances in technology and consequently not meeting the needs to have instructors integrate computer technologies in the curriculum and practice. This challenge of course exists for all digital learning technologies. However, ITSs are comparatively complex and some computational components are difficult to explain to instructors who want to know how they work.

A fourth obstacle is relevant to those ITSs with conversational agents. These systems have limited mixed-initiative dialogue and affordances for self-regulated learning. Each system covers a rather specific topic so they cannot provide support for all questions, topics, problems, or tasks that students initiate unless the ITS has already developed them ahead of time. Even for a specific topic, students ask a broad diversity of questions, with many unanticipated by the ITS. Consequently, the learning leans toward being instruction-centered or tutor-centered more than student-centered. Perhaps these limitations will be reduced in the future when there are thousands of ITSs developed on a wide range of topics and when data repositories are so large that the likelihood of a student asking a unique question about a topic is vanishingly small. That will hopefully evolve in the emerging era of Big Data.

References

- Adamson, D., Dyke, G., Jang, H., Rosé, C.P., 2014. Towards an agile approach to adapting dynamic collaboration support to student needs. *Int. J. Artif. Intell. Educ.* 24 (1), 92–124.
- Aleven, V., McLaren, B.M., Sewall, J., Koedinger, K., 2009. A new paradigm for intelligent tutoring systems: example-tracing tutors. *Int. J. Artif. Intell. Educ.* 19, 105–154.
- Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R., 1995. Cognitive tutors: lessons learned. *J. Learn. Sci.* 4, 167–207.
- Arroyo, I., Muldner, K., Burleson, W., Woolf, B.P., 2014. Adaptive interventions to address students' negative activating and deactivating emotions during learning activities. In: Sottilare, R., Graesser, A., Hu, X., Goldberg, B. (Eds.), *Design Recommendations for Intelligent Tutoring Systems: Instructional Management*, vol. 2. Army Research Laboratory (US), Orlando, FL, pp. 79–91.
- Azevedo, R., Moos, D., Johnson, A., Chauncey, A., 2010. Measuring cognitive and metacognitive regulatory processes used during hypermedia learning: issues and challenges. *Educ. Psychol.* 45, 210–223.
- Baker, R.S.J.D., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C., 2010. Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum. Comput. Stud.* 68 (4), 223–241.
- Baylor, A.L., Kim, Y., 2005. Simulating instructional roles through pedagogical agents. *Int. J. Artif. Intell. Educ.* 15, 95–115.
- Biswas, G., Segedy, J.R., Bunchongchit, K., 2016. From design to implementation to practice a learning by teaching system: Betty's Brain. *Int. J. Artif. Intell. Educ.* 26 (1), 350–364.
- Chi, M.T.H., Wylie, R., 2014. ICAP: a hypothesis of differentiated learning effectiveness for four modes of engagement activities. *Educ. Psychol.* 49, 1–25.
- Chi, M.T.H., Siler, S., Yamauchi, T., Jeong, H., Hausmann, R., 2001. Learning from human tutoring. *Cognit. Sci.* 25, 471–534.
- Cohen, P.A., Kulik, J.A., Kulik, C.C., 1982. Educational outcomes of tutoring: a meta-analysis of findings. *Am. Educ. Res. J.* 19, 237–248.
- Collins, A., Halverson, R., 2009. *Rethinking Education in the Age of Technology: The Digital Revolution and Schooling in America*. Teachers College Press, New York.
- Craig, S.D., Gholson, B., Brittingham, J.K., Williams, J.L., Shubeck, K., 2012. Promoting vicarious learning of physics using deep questions with explanations. *Comput. Educ.* 58, 1042–1048.
- Crockett, K., Latham, A., Whitton, N., 2017. On predicting learning styles in conversational intelligent tutoring systems using fuzzy decision trees. *Int. J. Hum. Comput. Stud.* 97, 98–115.
- DeFalco, J.A., Rowe, J.P., Paquette, L., Georgoulas-Sherry, V., Brawner, K., Mott, B.W., Baker, R.S., Lester, J.C., 2018. Detecting and addressing frustration in a serious game for military training. *Int. J. Artif. Intell. Educ.* 28 (2), 152–193.
- D'Mello, S.K., Graesser, A.C., 2012a. AutoTutor and affective AutoTutor: learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Trans. Interact. Intell. Syst.* 23, 1–38.
- D'Mello, S., Graesser, A., 2012b. Dynamics of affective states during complex learning. *Learn. Instruct.* 22 (2), 145–157.
- D'Mello, S., Olney, A.M., Person, N., 2010. Mining collaborative patterns in tutorial dialogues. *J. Educ. Data Min.* 2 (1), 1–37.
- D'Mello, S., Lehman, B., Pekrun, R., Graesser, A.C., 2014. Confusion can be beneficial for learning. *Learn. Instruct.* 29, 153–170.

- Dodds, P.V.W., Fletcher, J.D., 2004. Opportunities for new "smart" learning environments enabled by next generation web capabilities. *J. Educ. Multimed. Hypermedia* 13, 391–404.
- Dynarsky, M., Agodina, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., et al., 2007. Effectiveness of Reading and Mathematics Software Products: Findings From the First Student Cohort. Institute of Education Sciences, Washington, DC.
- Falmagne, J., Albert, D., Doble, C., Eppstein, D., Hu, X., 2013. *Knowledge Spaces: Applications in Education*. Springer, Berlin-Heidelberg.
- Forsyth, C.M., Graesser, A., Millis, K., 2020. Predicting learning in a multi-component serious game. *Technol. Knowl. Learn.* 25 (2), 251–277.
- Geden, M., Emerson, A., Carpenter, D., Rowe, J., Azevedo, R., Lester, J., 2020. Predictive student modeling in game-based learning environments with word embedding representations of reflection. *Int. J. Artif. Intell. Educ.* 1–23.
- Gobert, J.D., Sao Pedro, M., Raziuddin, J., Baker, R.S., 2013. From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *J. Learn. Sci.* 22 (4), 521–563.
- Gobert, J., Moussavi, R., Li, H., Pedro, M.S., Dickler, R., 2018. Real-time scaffolding of students' online data interpretation during inquiry with Inq-ITS using educational data mining. In: Auer, M.E., Azad, A.K.M., Edwards, A., de Jong, T. (Eds.), *Cyber-Physical Laboratories in Engineering and Science Education*. Springer, Cham, Switzerland, pp. 191–218.
- Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H.H., Ventura, M., Olney, A., Louwerse, M.M., 2004. AutoTutor: a tutor with dialogue in natural language. *Behav. Res. Methods Instrum. Comput.* 36 (2), 180–192.
- Graesser, A.C., Chipman, P., Haynes, B.C., Olney, A., 2005. AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Trans. Educ.* 48, 612–618.
- Graesser, A.C., D'Mello, S.K., Person, N., 2009. Meta-knowledge in tutoring. In: Hacker, D.J., Dunlosky, J., Graesser, A.C. (Eds.), *Metacognition in Educational Theory and Practice*. Erlbaum, Mahwah, NJ.
- Graesser, A.C., Forsyth, C., Lehman, B., 2017. Two heads are better than one: learning from agents in conversational dialogues. *Teach. Coll. Rec.* 119, 1–20.
- Graesser, A.C., Hu, X., Sottolare, R., 2018. Intelligent tutoring systems. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, pp. 246–255.
- Graesser, A.C., Greenberg, D., Olney, A.M., Lovett, M.W., 2019. Educational technologies that support reading comprehension for adults who have low literacy skills. In: Perin, D. (Ed.), *Wiley Adult Literacy Handbook*. Wiley, New York, pp. 471–493.
- Graesser, A.C., 2016. Conversations with AutoTutor help students learn. *Int. J. Artif. Intell. Educ.* 26, 124–132.
- Graesser, A.C., 2020. Learning science principles and technologies with agents that promote deep learning. In: Feldman, R.S. (Ed.), *Learning Science: Theory, Research, and Practice*. McGraw-Hill, New York, pp. 2–33.
- Greiff, S., Wüstenberg, S., Holt, D.V., Goldhammer, F., Funke, J., 2013. Computer-based assessment of complex problem solving: concept, implementation, and application. *Educ. Technol. Res. Dev.* 61, 407–421.
- Hampton, A.J., Nye, B.D., Pavlik, P.I., Swartout, W.R., Graesser, A.C., Gunderson, J., 2018. Mitigating knowledge decay from instruction with voluntary use of an adaptive learning system. In: *International Conference on Artificial Intelligence in Education*. Springer, Cham, pp. 119–133.
- Hausmann, R.G.M., van de Sande, B., VanLehn, K., 2008. Trialog: how peer collaboration helps remediate errors in an ITS. In: *Proceedings of the 21st International FLAIRS Conference*. AAAI Press, Menlo Park: CA, pp. 415–420.
- Hutt, S., Krasich, K., Brockmole, J., D'Mello, S.K., (in press). Design and evaluation of an attentional-aware intelligent tutoring system. In: *The ACM Conference on Human Factors in Computing Systems (CHI)* 2021.
- Jackson, G.T., McNamara, D.S., 2013. Motivation and performance in a game-based intelligent tutoring system. *J. Educ. Psychol.* 105, 1036–1049.
- Jackson, G.T., Zapata-Rivera, D., 2015. Conversation-based assessment. *R&D Connect.* 25, 1–8.
- Jackson, G.T., Castellano, K.E., Brockway, D., Lehman, B., 2018. Improving the measurement of cognitive skills through automated conversations. *J. Res. Technol. Educ.* 50 (3), 226–240.
- Johnson, W.L., Lester, J.C., 2016. Twenty years of face-to-face interaction with pedagogical agents. *Int. J. Artif. Intell. Educ.* 26, 25–36.
- Johnson, W.L., Valente, A., 2009. Tactical Language and culture training systems: using AI to teach foreign languages and cultures. *AI Mag.* 30, 72–83.
- Koedinger, K.R., Corbett, A.C., Perfetti, C., 2012. The Knowledge-Learning-Instruction (KLI) framework: bridging the science-practice chasm to enhance robust student learning. *Cognit. Sci.* 36 (5), 757–798.
- Kulik, J.A., Fletcher, J.D., 2015. Effectiveness of intelligent tutoring systems: a meta-analytic review. *Rev. Educ. Res.* 85, 171–204.
- Laird, J., 2012. *The SOAR Cognitive Architecture*. MIT Press, Cambridge, MA.
- Lesgold, A.M., 2019. *Learning for the Age of Artificial Intelligence: Eight Educational Competencies*. Routledge, New York.
- Li, H., Gobert, J., Dickler, R., 2017a. Automated assessment for scientific explanations in on-line science inquiry. In: Hu, X., Barnes, T., Hershkovitz, A., Paquette, L. (Eds.), *Proceedings of the 10th International Conference on Educational Data Mining*. EDM Society, Wuhan, China, pp. 214–219.
- Li, H., Gobert, J., Dickler, R., 2017b. Dusting off the messy middle: assessing students' inquiry skills through doing and writing. In: André, E., Baker, R., Hu, X., Rodrigo, M., du Boulay, B. (Eds.), *Artificial Intelligence in Education. AIED 2017, Lecture Notes in Computer Science*, vol. 10331. Springer, Cham, Switzerland, pp. 175–187.
- Li, H., Gobert, J., Graesser, A.C., Dickler, R., 2018. Advanced educational technology for science inquiry assessment. *Policy Insights Behav. Brain Sci.* 5 (2), 171–178.
- Ma, W., Adesope, O.O., Nesbit, J.C., Liu, Q., 2014. Intelligent tutoring systems and learning outcomes: a meta-analytic survey. *J. Educ. Psychol.* 106, 901–918.
- Meyer, B.J., Wijekumar, K.K., Lin, Y.C., 2011. Individualizing a web-based structure strategy intervention for fifth graders' comprehension of nonfiction. *J. Educ. Psychol.* 103 (1), 140.
- Mitrovic, A., Martin, B., Suraweera, P., 2007. Intelligent tutors for all: the constraint-based approach. *IEEE Intell. Syst.* 22, 38–45.
- National Academy of Sciences, Engineering, and Medicine, 2018. *How People Learn II: Learners, Contexts, and Cultures*. National Academies Press, Washington, D.C.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- Next Generation Science Standards Lead States, 2013. *Next Generation Science Standards: For States, By States*. National Academies Press, Washington, DC.
- Nye, B.D., Graesser, A.C., Hu, X., 2014. AutoTutor and family: a review of 17 years of natural language tutoring. *Int. J. Artif. Intell. Educ.* 24, 427–469.
- OECD, 2017. *PISA 2015 Results (Volume V): Collaborative Problem Solving*. OECD Publishing, Paris.
- Olney, A., D'Mello, S., Person, N., Cade, W., Hayes, P., Williams, C., Lehman, B., Graesser, A.C., 2012. Guru: a computer tutor that models expert human tutors. In: Cerri, S.A., Clancey, B. (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems (ITS 2012)*. Springer-Verlag, Berlin, pp. 256–261.
- Ritter, S., Anderson, J.R., Koedinger, K.R., Corbett, A., 2007. Cognitive Tutor: applied research in mathematics education. *Psychon. Bull. Rev.* 14, 249–255.
- Siemens, G., Gasevic, D., Dawson, S. (Eds.), 2015. *Preparing for the Digital University: A Review of the History and Current State of Distance, Blended, and Online Learning*. MOOC Research Initiative.
- Sinclair, J., Coulthart, M., 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford University Press, London.
- Snow, E.L., Allen, L.K., Jacovina, M.E., McNamara, D.S., 2015. Does agency matter? Exploring the impact of controlled behaviors within a game-based environment. *Comput. Educ.* 82, 378–392.
- Sottolare, R., Graesser, A.C., Hu, X., Sinatra, A. (Eds.), 2018a. *Design Recommendations for Intelligent Tutoring Systems: Team Tutoring*, vol. 6. U.S. Army Research Laboratory, Orlando, FL.
- Sottolare, R.A., Baker, R.S., Graesser, A.C., Lester, J.C., 2018b. Special issue on the Generalized Intelligent Framework for Tutoring (GIFT): creating a stable and flexible platform for innovation in AIED research. *Int. J. Artif. Intell. Educ.* 28, 139–151.
- Swartout, W., Artstein, R., Forbell, E., Foutz, S., Lane, H.C., Lange, B., Morie, J., Noren, D., Rizzo, S., Traum, D., 2013. Virtual humans for learning. *AI Mag.* 34 (4), 13–30.
- Tärning, B., Silvervarg, A., Gulz, A., Haake, M., 2019. Instructing a teachable agent with low or high self-efficacy—does similarity attract? *Int. J. Artif. Intell. Educ.* 29 (1), 89–121.

- Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., Lester, J., 2020. The agency effect: the impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Comput. Educ.* 147, 103781.
- Tegos, S., Demetriadis, S., Papadopoulos, P.M., Weinberger, A., 2016. Conversational agents for academically productive talk: a comparison of directed and undirected agent interventions. *Int. J. Comput. Support. Collab. Learn.* 11 (4), 417–440.
- VanLehn, K., Lynch, C., Schultz, K., Shapiro, J.A., Shelby, R., Taylor, L., et al., 2005. The Andes physics tutoring system: lessons learned. *Int. J. Artif. Intell. Educ.* 15 (3), 147–204.
- VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., Rose, C.P., 2007. When are tutorial dialogues more effective than reading? *Cognit. Sci.* 31, 3–62.
- VanLehn, K., 2006. The behavior of tutoring systems. *Int. J. Artif. Intell. Educ.* 16 (3), 227–265.
- VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. *Educ. Psychol.* 46, 197–221.
- Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., Weston, T., 2013. My science tutor: a conversational multimedia virtual tutor. *J. Educ. Psychol.* 105, 1115–1125.
- Wijekumar, K.K., Meyer, B.J., Lei, P., 2017. Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *J. Educ. Psychol.* 109 (6), 741.
- Woolf, B.P., 2009. *Building Intelligent Tutoring Systems*. Morgan Kaufman, Burlington, MA.
- Xu, Y., Wang, D., Collins, P., Lee, H., Warschauer, M., 2021. Same benefits, different communication patterns: comparing Children's reading with a conversational agent vs. a human partner. *Comput. Educ.* 161, 104059.
- Yan, D., Rupp, A., Foltz, P., 2020. In: *Handbook of Automated Scoring: Theory Into Practice*. CRC Press/Taylor and Francis, New York.

Mindfulness and learning

Darren L. Dunning^{a,b}, ^a Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom; and ^b University of Hertfordshire, Hatfield, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	648
How could mindfulness training improve learning?	648
What is the evidence that mindfulness training improves learning?	649
Method	649
Search strategy and inclusion criteria	649
Data extraction and synthesis	650
Analysis	650
Risk of bias	651
Results	651
Heterogeneity	651
Risk of bias	651
Moderator analysis	652
Discussion	652
Conclusion	653
Appendices	653
Appendix 1—Table showing effect sizes (Cohen's <i>d</i>) immediately after training and at follow-up. Positive effect sizes show superiority of MBI over controls	653
Appendix 2—Forest plots	654
Appendix 3—Funnel plots	655
References	656

Introduction

Mindfulness is the intentional moment-to-moment focus on the present, while treating inevitable intrusive thoughts and feelings with an attitude of curiosity and nonjudgement (Bishop et al., 2004). Mindfulness-based interventions (MBIs) that train this ability are an increasingly popular way of attempting to enhance the cognitive, emotional and behavioral outcomes of adults and young people, with studies showing improvements in a host of areas including anxiety and depression (Hofmann et al., 2010), cognitive skills (Chiesa et al., 2011) and disruptive behavior (Klingbeil et al., 2017). However, studies seldom target mindfulness' utility in improving learning. Learning, the process of gaining new skills, knowledge and understanding, is crucial for improving performance and sustaining achievements. The following review and meta-analysis elucidates the relationship between mindfulness and learning based on existing literature. Although learning is a continuous process that occurs throughout the lifespan, the following focuses on young people in education.

How could mindfulness training improve learning?

Mindfulness has been practiced for around 2500 years, though formal MBIs have been around for less than 50. Three of the most popular and enduring MBIs that were originally designed for adults, but have been adapted for use with children and adolescents are: Mindfulness-Based Stress Reduction (MBSR; Kabat-Zinn, 1982, 2003), Acceptance and Commitment Therapy (ACT; Hayes et al., 2009) and Mindfulness-Based Cognitive Therapy (MBCT; Segal et al., 2002). MBSR was the earliest example of a formalized MBI and was developed to help adults learn to manage illness, pain and stress. Later, ACT sought to help individuals overcome negative thoughts and feelings through acceptance, and MBCT focused on helping to prevent the recurrence of depression. So, if MBIs were typically designed to help us manage stress or prevent reoccurrence of depression, why could they also improve our ability to learn?

First, there is evidence to suggest that MBIs may improve executive functions (Frank et al., 2021; Grabovec et al., 2011; Lam and Seiden, 2020; Shapiro et al., 2006). Executive functions are a set of mental skills required for managing our thoughts and behaviors that are also important for learning and academic achievement (Duncan et al., 2007; St. Clair-Thompson and Gathercole, 2006). These include working memory, planning, organization, task initiation, flexible thinking and most relevant to this review—emotional and attentional control, inhibition, and self-regulation. Using MBI's to improve learning via enhancement of executive functions makes theoretical sense. The ability to focus attention is a cornerstone of both mindfulness practice and successful learning (Fisher, 2019; Steinmayr et al., 2010; Valentine and Sweet, 1999). Indeed, sustaining attention, in particular, appears to play an essential supporting role in enhancing other cognitive functions such as memory (Fortenbaugh et al., 2017) and is related to

academic success (Polderman et al., 2010; Posner and Rothbart, 2005). Similarly, MBIs have also been shown to improve self-regulation skills, such as self-monitoring (i.e., understanding and adjusting our behaviors) and inhibition (i.e., how we control our impulsive responses), which are again important for learning (Holzel et al., 2011; Panadero et al., 2012; Van der Sluis et al., 2004; Zimmerman and Shunk, 2011). If mindfulness practice requires attending to one's own present moment thoughts and feelings and treating intrusive thoughts nonjudgmentally without trying to engage with or change those thoughts, then, in theory, this requires sustained attention, self-monitoring, and inhibition.

Second and more indirect, learning during the school years coincides with a particularly vulnerable period for the onset of psychopathology; indeed, adolescents, in particular, are susceptible to mental illness with 50% of mental health problems having their onset before the age of 18 and 75% before the age of 24 (Kessler et al., 2005). Emotional turmoil that manifests as anxiety, depression or loneliness can make the school years a challenging time for young people and can negatively affect their learning experience. Because MBIs are useful in helping reduce depression and anxiety in young people (e.g., Dunning et al., 2019) and in preventing relapse of depression (e.g., Kuyken et al., 2015), it seems sensible to suggest that if a student is less concerned with mental health issues, they would more likely find it easier to learn.

What is the evidence that mindfulness training improves learning?

Given the importance of learning and the increased use of MBIs to improve outcomes in young people, the evidence for mindfulness improving learning is surprisingly sparse. A meta-analysis by Klingbeil and colleagues included five studies of varied design (i.e., randomized controlled trials (RCTs), quasi-randomized trials) that included a wide range of measures of academic achievement (e.g., school grades) and school functioning (e.g., classroom participation, student behavior). Although the meta-analysis showed MBIs improved academic achievement and school functioning (hedges' $g = 0.32$ (pre-post trials), $g = 0.39$ (controlled trials)) the authors acknowledged that the results contained publication bias and large standard errors, which were likely the result of the broadness of included outcomes and varied trial designs (Klingbeil et al., 2017).

Maynard et al. (2017) in a review and meta-analysis of school-based MBIs found five studies that included measures of academic outcomes (school grades and academic credits). They found small but non-significant effects on learning (hedges' $g = 0.27$, $p = 0.08$) and further concluded that the studies showed a high risk of bias. Indeed, the studies included in their review were a mixture of RCTs, quasi-experimental trials, single-group pre-post designs and single-subject designs.

Since the publication of these two reviews, there has been a considerable number of new studies examining the effects of MBIs on young people, including a twofold increase in the number of RCTs (Dunning et al., 2022b). This allows a more comprehensive analysis of the effects of MBIs on learning.

In the following, I present a meta-analysis of the benefits of MBIs for improving the learning outcomes of young people. A key feature of this meta-analysis is that the included studies are all RCTs, the gold standard in research methodology. Including only RCTs protects against both test-retest effects and the effects of maturational changes. Furthermore, as participants are randomized to group, it increases the likelihood that there are no systematic differences between groups that might be confounded with the treatment condition. The following also allows the separate analyses of MBIs compared against passive control groups (e.g., no contact, teaching as usual, wait-list) or active control groups (e.g., another intervention). Meta-regressions are used to investigate several variables of interest. First, as bias has been shown in other meta-analyses, we can ascertain if the quality of individual studies (risk of bias) affect results when analyzing only tightly controlled RCTs. Second, we can explore if the dosage of training is related to learning improvements—specifically does more mindfulness practice mean better results. Third, we can assess if MBIs are more beneficial for improving the learning of younger or older age groups. Additionally, we examine if publication bias is present and, by means of studies that included follow-up assessments, whether effects of MBIs on learning are sustained or relatively short-lived.

Method

Search strategy and inclusion criteria

Literature searches for published and unpublished articles from January 1990 to January 2022 were conducted, using the electronic databases Pubmed, PsycINFO, EMBASE and Google Scholar. The search terms and Boolean operators used were ((“mindful*”) AND (“child*” OR “school” OR “adolescen*” OR “youth”)). Because “learning” is a broad concept, this was not included in the search terms as it was feared studies would be missed if the search terms were too prescriptive. The reference lists of reviews and meta-analyses were also searched for suitable studies. After searches were collated and duplicates removed, the titles/abstracts of the remaining articles were reviewed. If the title/abstract suggested that the study included a learning outcome (e.g., reading, maths, grades, academic achievement) then the full-text article was downloaded and compared against the following inclusion criteria:

- (i) The effects of mindfulness were evaluated against a control condition and the participants were randomized to condition;
- (ii) The outcome measures used were related to learning outcomes and the paper provided quantitative data from which effect sizes could be calculated (means and standard deviations).

Ten studies met these criteria and were synthesized in the analyses.

Data extraction and synthesis

For each study the following variables were recorded: age of sample, total number of participants (i.e., number in MBI group plus number in control group), type of control condition(s), total hours of mindfulness training (i.e., sessions x duration in minutes, excluding home practice), type of mindfulness practice (e.g., MBSR, MCBT), learning outcome reported and follow-up period (if applicable). All of this information can be seen in [Table 1](#).

Between-group effect sizes (Cohen's *d*; [Cohen, 2013](#)) were calculated based on the mean pre-post change in the MBI group minus the mean pre-post change in the control group, divided by the pooled pre-test standard deviation ([Morris, 2008](#)). The pooled pre-test standard deviation was used so the intervention could not have influenced the standard deviation, based on the recommendation by [Carlson and Schmidt \(1999\)](#). A positive effect size indicates that the MBI group benefitted more than the comparison group. Effect sizes were interpreted on the following basis: $d = 0.20$, a small effect; $d = 0.50$, a moderate effect; and $d = 0.80$, a large effect ([Cohen, 2013](#)). If a study used multiple measures to assess learning, then a mean of the effect sizes was calculated on pre-to post-intervention effects. A full list of the learning outcomes used, with their effect sizes can be seen in [Appendix 1](#).

Analysis

All analyses were conducted using the Comprehensive Meta-Analysis program version 3.3.70 ([Borenstein et al., 2021](#)). Confidence intervals were calculated for effect sizes.

Due to variation in studies in features such as age of sample, learning outcome used, and type of MBI used, a random effects model was chosen for all analyses. All 10 studies were included in initial analyses. Additionally, separate analyses were conducted on the six studies that used an active control condition, the five studies that used a passive control condition (note: [Zelazo et al., \(2018\)](#) used both active and passive control groups), and the six studies that included a follow-up assessment.

Individual, random-effects meta-regressions assessed the impact of three moderator variables: (1) age of sample, to establish if younger or older participants benefitted most from MBIs; (2) total hours of MBI, to explore if the dose of mindfulness training affected learning outcomes; and (3) risk of bias, to ascertain if study quality influenced learning outcome. For the studies that included a follow-up assessment, an additional moderator variable was used: 4) time since follow-up (i.e., the time between post-assessment and follow up) to establish if any effects of MBIs strengthen or fade over time.

Table 1 Details of all studies included in the meta-analysis.

Authors	<i>n</i> Sample	<i>n</i> Mindfulness Controls	Age (mean)	Mindfulness training	Control group	Learning outcome	Duration of training (hrs)	Follow up period (mths)	Risk of bias
De Voy (2019) study 1	Typical 20	14	14.5	dot.be	Active—progressive muscle relaxation	Spelling, maths, reading	6	3	- + + - -
De Voy (2019) study 2	Typical 41	15 58	14.5	dot.be	Active—study skills Active—progressive muscle relaxation	comprehension Spelling, maths, reading	6	3	- + + - -
De Voy (2019) study 3	Typical 6	36 5	14.33	dot.be	Active—study skills Active—progressive muscle relaxation	comprehension Spelling, maths, reading	6	3	- + + - -
Flook et al. (2015)	Typical 29	6 37	4.7	Kindness curriculum	Active—study skills Passive—wait list	comprehension School grades	10.00	N/A	? ? + + -
Franco et al. (2010)	Typical 31	30	16.8	Meditation fluir	Passive—wait list	School grades	15.00	N/A	? + ? ? -
Lassander et al. (2020)	Typical 58	59	13.5	dot.be	Active—relaxation	Verbal fluency	6.75	3	- ? - - +
Lin and Mai (2018)	Typical 42	35	University students	Not stated	Passive—teaching as usual	Course assessment	26	N/A	? + + ? -
Moreno-Gómez and Cejudo (2019)	Typical 48	26	5.08	MindKinder	Passive—no contact	Verbal fluency, reading, writing	36	N/A	+ ? ? ? ?
Schonert-Reichl et al. (2015)	Typical 48	51	10.2	MindUP	Active—social responsibility program	Maths grades	9.00	N/A	+ ? + + +
Zelazo et al. (2018)	Typical 72	76 68	4.75	Mindfulness/ reflection	Active—literacy Passive—no contact	Literacy	10	1	? ? ? + ?

Note: for risk of bias: (1) random sequencing; (2) allocation concealment; (3) blinding of outcome assessment; (4) incomplete outcome; (5) selective reporting.

Heterogeneity, the amount of diversity in the characteristics of the outcome measures, was quantified using the Q statistic and I^2 estimates. For I^2 estimates, a value of 0% equates to no heterogeneity, 25% to low heterogeneity, 50% to moderate heterogeneity and 75% to high heterogeneity (Higgins et al., 2003).

Risk of bias

The Cochrane Collaboration's Risk of Bias Tool (Higgins et al., 2008) was used to assess study quality. This involves making evidence-based judgments about specific features of each study to establish if biases exist that could lead to an overestimation or underestimation of the true intervention effects. Risk of bias was graded in terms of the following categories: (1) random sequencing—if the method used to randomly allocate participants to group was appropriate; (2) allocation concealment—if allocation to condition was concealed from participants during enrollment; (3) blinding of outcome assessment—if assessors were unaware of study condition; (4) incomplete outcome data—if missing outcome data were appropriately reported and/or given appropriate statistical treatment; (5) selective reporting—the degree to which studies reported all pre-specified outcomes.

Each of the five risk of bias categories for each study was awarded one of three ratings: a “minus” indicated a low risk of bias; a “plus” indicated a high risk of bias, and a “question mark” was used if the risk of bias was unclear. To enable the use of risk of bias as a moderator of the effects of MBIs on learning outcomes, a discrete variable was calculated. For each of the five categories for each study, each “plus” given a value of 1, each “question mark” a value of 0 and each “minus” a value of -1 . Therefore, individual studies could have a risk of bias score of between -5 and 5 , with higher scores indicating a higher risk of bias.

To investigate publication bias, the extent to which the studies included were representative of the population of studies, a series of Begg's funnel plots and Egger's regressions were conducted for each analysis. Missing data were imputed using the Duval and Tweedie trim and fill method.

Results

Table 2 shows the results of the meta-analysis. Across all studies, the learning outcomes of participants receiving an MBI improved significantly more than those receiving the control conditions (across active and passive conditions), with a small effect size ($d = 0.26$). For sub-analyses of the six studies that compared MBIs against active control conditions and the five studies that compared MBIs against passive control conditions, results were non-significant—though MBIs versus active control groups approached significance ($d = 0.16$, $p = 0.06$).

Analysis of the six studies that included follow-up assessments showed that MBIs were more effective than control conditions (the analysis contained a mixture of active and passive control groups) in improving learning outcomes, with a small effect size ($d = 0.24$).

Forest plots of all analyses can be found in Appendix 2.

Heterogeneity

Heterogeneity was present for analysis of all studies and the studies that compared MBIs against passive control groups, with significant Q values of 26.60 and 23.75 respectively. I^2 estimates suggested heterogeneity for significant outcomes was moderate to high in size. Heterogeneity statistics can be seen in Table 2.

Risk of bias

There was a low risk of bias in 40% of the studies for random sequencing, 0% for allocation concealment, 10% for blinding, 40% for incomplete outcome data, and 60% for selective reporting. A high risk of bias existed in 20% of the studies for random sequencing, 50% for allocation concealment, 60% for blinding, 30% for incomplete outcome data and 20% for selective reporting. In all other cases the risk of bias was unclear. For the risk of bias of individual studies, see Table 1.

Table 2 Analysis of effects of MBIs on learning, when compared with active/passive control groups, and at follow-up assessment.

	k	Number of effect sizes	Total n	Intervention effects					Heterogeneity			Publication bias (Eggers)	
				Mean effect size (d)	SE	95% CI	Z	p	Q value	p	I^2 (%)	t	P
All studies	10	33	911	0.26	0.13	[0.01–0.50]	2.05	<0.05	26.60	<0.01	66	1.12	0.29
Active controls	6	21	565	0.16	0.09	[−0.01 to 0.33]	1.89	0.06	3.04	0.69	0	0.76	0.49
Passive controls	5	12	418	0.32	0.25	[−0.18 to 0.81]	1.26	0.21	23.75	<0.01	83	1.24	0.30
Follow up	6	22	608	0.24	0.12	[0.11–0.48]	2.06	<0.05	8.41	0.14	41	0.18	0.86

There was no evidence of publication bias in any analysis, suggesting that the studies included here are representative of the wider literature (see [Table 2](#)). Funnel plots for all analyses can be found in [Appendix 3](#).

Moderator analysis

Neither age nor hours of mindfulness moderated outcome in analysis of all 10 studies, MBIs versus active controls or MBIs versus passive controls (all $ps > 0.5$). However, for MBIs that included a follow-up assessment, age significantly moderated learning outcomes with younger participants experiencing larger learning gains following MBIs (compared to controls) than older participants. Likewise, hours of mindfulness moderated learning outcomes at follow-up, with a bigger dose of mindfulness equating to larger effects. Risk of bias moderated improvements in learning following MBI for studies that compared MBIs against passive control groups and for studies that included a follow-up. The pattern here was different with a higher risk of bias equating to a lower effect size for MBIs compared to passive controls and lower risk of bias equating to lower effect sizes for MBIs at follow up. Length of follow-up (i.e., time elapsed since post-assessment) did not moderate performance at follow up.

Discussion

This is the first meta-analysis composed entirely of RCTs that examines if MBIs improve learning. It included separate analyses of MBIs versus active and passive control groups, and analysis of studies with follow-up assessments. A previous meta-analysis showed that MBIs significantly improved learning outcomes ([Klingbeil et al., 2017](#)), whereas another showed non-significant findings ([Maynard et al., 2017](#)). However, both previous studies were small, with only five studies included in each, and demonstrated high bias, possibly due to the varied research designs included in the study samples. Although the meta-analysis reported here is also relatively small, it includes double the number of studies of previous meta-analyses. Additionally, as its focus was exclusively on RCTs that by design are methodologically more robust than non-RCTs, there should be a lesser chance of bias.

Results revealed that MBIs did not significantly improve learning outcomes when compared to either passive ($k = 6$) or active ($k = 5$) control groups. However, when all studies were considered (i.e., both active and passive control groups included in the same analysis), MBIs showed small beneficial effects on learning ($d = 0.26$). For studies that included a follow-up assessment ($k = 6$), there was also a significant small positive effect of MBIs compared to controls ($d = 0.24$). Taken together, this evidence, though not compelling, is promising and warrants further investigation.

Moderator analysis revealed that age and hours of mindfulness did not moderate performance in analysis of MBIs versus active controls, or MBIs versus passive controls, or all studies. However, risk of bias did moderate performance in analysis of MBIs versus passive controls, with a higher risk of bias associated with lower effect sizes. For moderator analysis of follow-up studies, a larger dose of mindfulness (i.e., more hours of practice) suggested better learning outcomes, age moderated results at follow-up with younger participants benefitting more from MBIs than older ones and risk of bias moderated learning outcomes with a higher risk of bias associated with larger effect sizes, suggesting that studies of lower methodological quality may have inflated the overall effect size.

There was no evidence of publication bias, signifying that the trials included here were not systematically different from unpublished studies. However, risk of bias analysis showed that bias was present in some form in all 10 studies included in the review, particularly with regard to allocation concealment (if allocation to condition was hidden from participants during enrollment) and blinding (if assessors were unaware of study condition).

Even though RCTs are the gold standard research design, there were still relatively few mindfulness studies suitable for inclusion in this review. Indeed, the positive effects for studies at follow-up is based on analysis of only six studies and includes three taken from a single unpublished doctoral dissertation. Furthermore, the longest follow-up period of included studies was three months after cessation of training. The prospects of the effects of MBIs enduring for longer should therefore be explored. Although the focus here was on young people in education, there was a wide age range of participants in the included studies that stretched from kindergarten (4 years of age) up to 1st-year university students (18+ years). This makes understanding exactly how mindfulness improves learning a challenge. If we accept that improvements in executive functions lead to downstream improvements in learning outcomes, then we also need to consider that as different executive functions emerge at different times (and rates) across childhood and adolescence (e.g., inhibitory control usually emerges in early childhood and predates skills such as cognitive flexibility), then MBIs might also work in very different ways at different stages of childhood and adolescence. Another issue worth considering is that most of the studies included in this review tested different MBI protocols rather than run replication studies with established MBIs. Indeed, of the studies included in this review, there was only one MBI that was used more than once (i.e., dot.be). This problem extends to the wider field (see [Dunning et al., 2019](#) for a discussion) and, as mindfulness is a broad construct and interventions using mindfulness programs differ in terms of content, implementation, fidelity and duration (including home mindfulness practice), it makes drawing meaningful conclusions about how MBIs work difficult. There was also heterogeneity present in several analyses that suggests a lack of similarity between studies, in this case perhaps with regards to the different types of MBIs used or the wide age range of included participants.

All of the studies included in the meta-analysis were with typical groups. Future studies should investigate if MBIs improve the learning of children and adolescents with a clinical diagnoses of learning difficulties or with disorders such as attention deficit/hyperactivity disorder (ADHD) that can impede learning progress. Indeed, previous non-RCT studies have shown benefits in

both individuals with learning difficulties (Beauchemin et al., 2008) and with ADHD (Van der Oord et al., 2012; Zylowska et al., 2008).

Conclusion

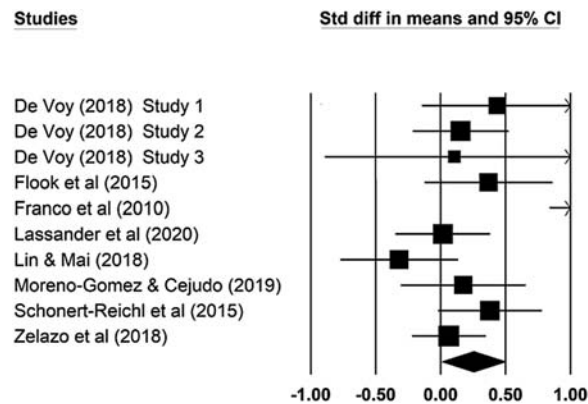
There is some evidence that MBIs might be useful in improving learning outcomes as well as a suggestion that effects of MBIs might be sustained and be particularly beneficial under certain circumstances (i.e., if the MBI contains more hours of training and includes younger participants). However, although there has been an influx of recent RCTs, some of which were fully powered, with large school-based representative samples and follow-ups that test the sustainability of effects (e.g., Kuyken et al., 2022; Volanen et al., 2020; Lassander et al., 2020; Dunning et al., 2022a), none of these studies included a measure of learning in their outcome variables, suggesting that improving learning is not a primary focus of current mindfulness research. Indeed, the existing evidence base for MBIs improving learning reviewed here is sparse (in both number and in size of studies) and shows evidence of bias. Although MBI offers some promise in improving learning and sustaining effects over time, all findings should be interpreted with caution.

Appendices

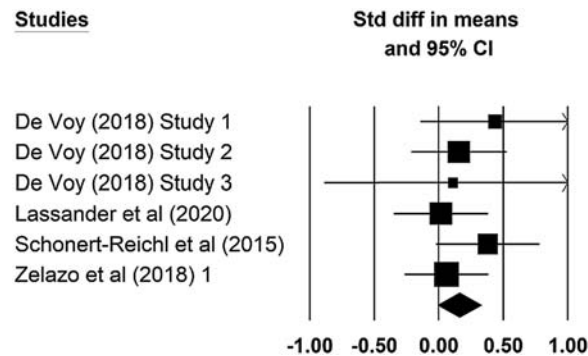
Appendix 1—Table showing effect sizes (Cohen's *d*) immediately after training and at follow-up. Positive effect sizes show superiority of MBI over controls

	<i>Assessment following training</i>	<i>Assessment at follow up</i>
De Voy (2019)1	Versus active study skills: spelling = 0.234, maths = 0.062, reading comprehension = 0.394	Versus active study skills: spelling = 0.328, maths = 0.199, reading comprehension = 0.322
	Versus active muscle relaxation: spelling = 0.136, maths = -0.058, reading comprehension = 0.475	Versus active muscle relaxation: spelling = 0.111, maths = -0.229, reading comprehension = 0.266
De Voy (2019)2	SS: spelling = -0.028, maths = 0.32, reading comprehension = -0.027	SS: spelling = 0.014, maths = 0.178, reading comprehension = -0.049
	Versus active muscle relaxation: spelling = 0.117, maths = 0.345, reading comprehension = 0.211	Versus active muscle relaxation: spelling = 0.074, maths = 0.273, reading comprehension = 0.28
De Voy (2019)3	SS: spelling = -0.028, maths = 0.055, reading comprehension = 0.174	SS: spelling = -0.182, maths = 0.319, reading comprehension = 0.189
	Versus active muscle relaxation: spelling = 0.251, maths = 0.386, reading comprehension = -0.198	Versus active muscle relaxation: spelling = 0.189, maths = 0.266, reading comprehension = -0.117
Flook et al. (2015)	School grades (learning = 0.556, health = 0.554, social-emotional = 0.964, cognition = -0.32, language = 0.092)	
Franco et al. (2010)	School grades = 1.4	
Lassander et al. (2020)	Verbal fluency = 0.017	Verbal fluency = -0.097
Lin and Mai (2018)	Course assessment = -0.32	
Moreno-Gómez and Cejudo (2019)	Verbal fluency = -0.024, reading = 0.03, writing = 0.046, learning problems = 0.647	Learning problems = 0.657
Schonert-Reichl et al. (2015)	Maths grades = 0.38	
Zelazo et al. (2018)	Versus active literacy: literacy = 0.06	Versus active literacy: literacy = 0.178
	Versus passive no contact: literacy = 0.067	Versus passive no contact: literacy = 0.696

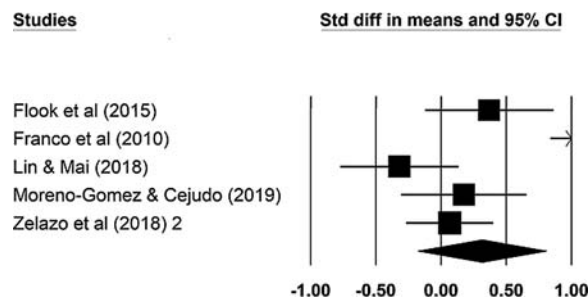
Appendix 2—Forest plots



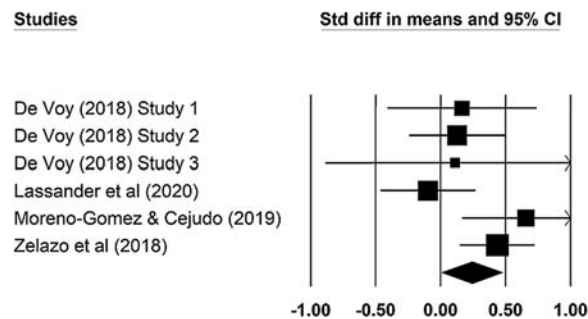
Forest plot showing all MBIs compared to controls in improving learning outcomes. Positive standard differences in means shows superiority of MBIs.



Forest plot showing MBIs compared to active controls in improving learning outcomes. Positive standard differences in means shows superiority of MBIs.

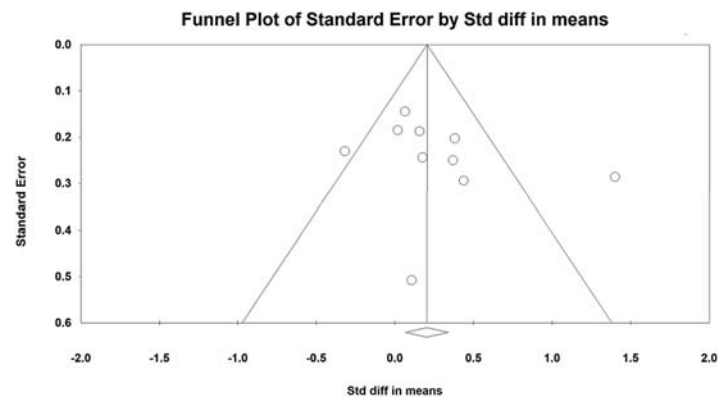


Forest plot showing MBIs compared to passive controls in improving learning outcomes. Positive standard differences in means shows superiority of MBIs.

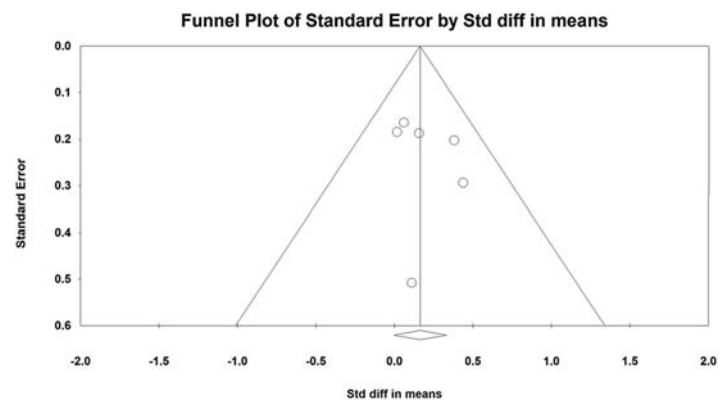


Forest plot showing MBIs compared to controls in improving learning outcomes at follow-up. Positive standard differences in means shows superiority of MBIs.

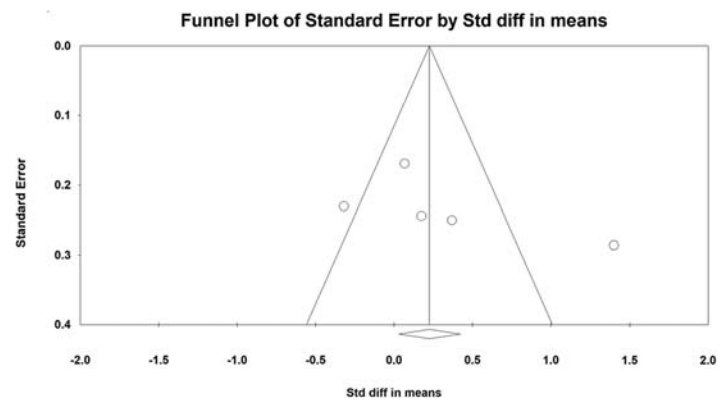
Appendix 3—Funnel plots



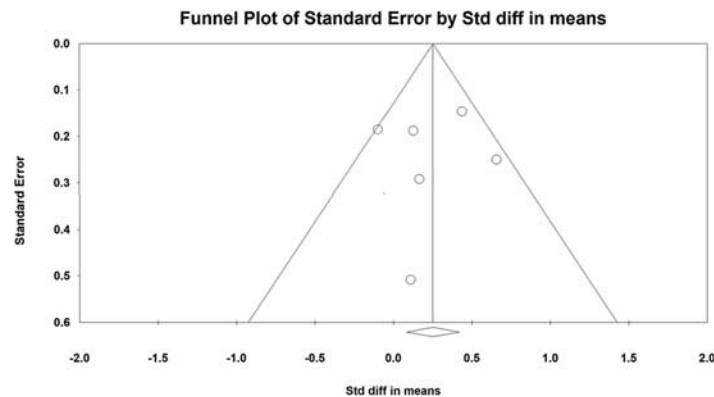
Funnel plot of publication bias for all MBI studies.



Funnel plot of publication bias for MBI studies versus active controls.



Funnel plot of publication bias for MBI studies versus passive controls.



Funnel plot of publication bias for MBI studies at follow-up assessment.

References

- Beauchemin, J., Hutchins, T.L., Patterson, F., 2008. Mindfulness meditation may lessen anxiety, promote social skills, and improve academic performance among adolescents with learning disabilities. *Compl. Health Pract. Rev.* 13 (1), 34–45.
- Bishop, S.R., Lau, M., Shapiro, S., Carlson, L., Anderson, N.D., Carmody, J., Segal, Z.V., Abbey, S., Speca, M., Velting, D., Devins, G., 2004. Mindfulness: a proposed operational definition. *Clin. Psychol. Sci. Pract.* 11 (3), 230.
- Borenstein, M., Hedges, L.V., Higgins, J.P., Rothstein, H.R., 2021. *Introduction to Meta-analysis*. John Wiley & Sons.
- Carlson, K.D., Schmidt, F.L., 1999. Impact of experimental design on effect size: findings from the research literature on training. *J. Appl. Psychol.* 84 (6), 851.
- Chiesa, A., Calati, R., Serretti, A., 2011. Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. *Clin. Psychol. Rev.* 31 (3), 449–464.
- Cohen, J., 2013. *Statistical Power Analysis for the Behavioral Sciences*. Routledge.
- De Voy, M., 2019. The Impact of Mindfulness Training on Schoolchildren's Academic Performance (Order No. 28357344). Available from ProQuest Dissertations & Theses A&I. (2475221279). Retrieved from: <https://ezp.lib.cam.ac.uk/login?url=https://www.proquest.com/dissertations-theses/impact-mindfulness-training-on-schoolchildrens/docview/2475221279/se-2?accountid=9851>.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., 2007. School readiness and later achievement. *Dev. Psychol.* 43 (6), 1428.
- Dunning, D.L., Griffiths, K., Kuyken, W., Crane, C., Foulkes, L., Parker, J., Dalgleish, T., 2019. Research Review: the effects of mindfulness-based interventions on cognition and mental health in children and adolescents—a meta-analysis of randomized controlled trials. *JCPP* 60 (3), 244–258.
- Dunning, D.L., Ahmed, S., Foulkes, L., et al., 2022a. The impact of mindfulness training in early adolescence on affective executive control, and on later mental health during-COVID-19: a randomised controlled trial. *Evid. Base Ment. Health*.
- Dunning, D.L., Tudor, K., Radley, L., et al., 2022b. Do mindfulness-based programmes improve the cognitive skills, behaviour and mental health of children and adolescents? An updated meta-analysis of randomised controlled trials. *Evid. Base Ment. Health*.
- Fisher, A.V., 2019. Selective sustained attention: a developmental foundation for cognition. *Curr. Opin. Psychol.* 29, 248–253.
- Flook, L., Goldberg, S.B., Pinger, L., Davidson, R.J., 2015. Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Dev. Psychol.* 51 (1), 44.
- Fortenbaugh, F.C., DeGutis, J., Esterman, M., 2017. Recent theoretical, neural, and clinical advances in sustained attention research. *Ann. N. Y. Acad. Sci.* 1396 (1), 70–91.
- Franco, C., Mañas, I., Cangas, A.J., Gallego, J., September 2010. The applications of mindfulness with students of secondary school: results on the academic performance, self-concept and anxiety. In: *World Summit on Knowledge Society*. Springer, Berlin, Heidelberg, pp. 83–97.
- Frank, J.L., Broderick, P.C., Oh, Y., Mitra, J., Kohler, K., Schussler, D.L., Geier, C., Roeser, R.W., Berrena, E., Mahfouz, J., Levitan, J., 2021. The effectiveness of a teacher-delivered mindfulness-based curriculum on adolescent social-emotional and executive functioning. *Mindfulness* 12 (5), 1234–1251.
- Grabovac, A.D., Lau, M.A., Willett, B.R., 2011. Mechanisms of mindfulness: a Buddhist psychological model. *Mindfulness* 2 (3), 154–166.
- Hayes, S.C., Strosahl, K.D., Wilson, K.G., 2009. *Acceptance and Commitment Therapy*. American Psychological Association, Washington, DC.
- Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. *BMJ* 327 (7414), 557–560.
- Higgins, J.P., Green, S., Scholten, R.J.P.M., 2008. Maintaining reviews: updates, amendments and feedback. In: *Cochrane Handbook for Systematic Reviews of Interventions*, p. 31.
- Hofmann, S.G., Sawyer, A.T., Witt, A.A., Oh, D., 2010. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. *J. Consult. Clin. Psychol.* 78 (2), 169.
- Hölzel, B.K., Lazar, S.W., Gard, T., Schuman-Olivier, Z., Vago, D.R., Ott, U., 2011. How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspect. Psychol. Sci.* 6 (6), 537–559.
- Kabat-Zinn, J., 1982. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. *Gen. Hosp. Psychiatr.* 4 (1), 33–47.
- Kabat-Zinn, J., 2003. Mindfulness-based stress reduction (MBSR). *Constr. Hum. Sci.* 8 (2), 73.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., Walters, E.E., 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey replication. *Arch. Gen. Psychiatr.* 62 (6), 593–602.
- Klingbeil, D.A., Renshaw, T.L., Willenbrink, J.B., Copek, R.A., Chan, K.T., Haddock, A., Yassine, J., Clifton, J., 2017. Mindfulness-based interventions with youth: a comprehensive meta-analysis of group-design studies. *J. Sch. Psychol.* 63, 77–103.
- Kuyken, W., Hayes, R., Barrett, B., Byng, R., Dalgleish, T., Kessler, D., Lewis, G., Watkins, E., Brejcha, C., Cardy, J., Causley, A., 2015. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. *Lancet* 386 (9988), 63–73.
- Kuyken, W., Ball, S., Crane, C., et al., 2022. Effectiveness and cost-effectiveness of a universal school-based mindfulness training programme compared with normal school provision in reducing risk of mental health problems and promoting well-being in adolescence: results of the MYRIAD cluster randomised controlled trial. *Evid. Base Ment. Health*.

- Lam, K., Seiden, D., 2020. Effects of a brief mindfulness curriculum on self-reported executive functioning and emotion regulation in Hong Kong adolescents. *Mindfulness* 11 (3), 627–642.
- Lassander, M., Hintsanen, M., Suominen, S., Mulla, S., Fagerlund, Å., Vahlberg, T., Volanen, S.M., 2020. The effects of school-based mindfulness intervention on executive functioning in a cluster randomized controlled trial. *Dev. Neuropsychol.* 45 (7–8), 469–484.
- Lin, J.W., Mai, L.J., 2018. Impact of mindfulness meditation intervention on academic performance. *Innovat. Educ. Teach. Int.* 55 (3), 366–375.
- Maynard, B.R., Solis, M.R., Miller, V.L., Brendel, K.E., 2017. Mindfulness-based interventions for improving cognition, academic achievement, behavior, and socioemotional functioning of primary and secondary school students. *Campbell Syst. Rev.* 13 (1), 1–144.
- Moreno-Gómez, A.J., Cejudo, J., 2019. Effectiveness of a mindfulness-based social-emotional learning program on psychosocial adjustment and neuropsychological maturity in kindergarten children. *Mindfulness* 10 (1), 111–121.
- Morris, S.B., 2008. Estimating effect sizes from pretest-posttest-control group designs. *Organ. Res. Methods* 11 (2), 364–386.
- Panadero, E., Tapia, J.A., Huertas, J.A., 2012. Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learn. Indiv. Differ.* 22 (6), 806–813.
- Polderman, T.J., Boomsma, D.I., Bartels, M., Verhulst, F.C., Huizink, A.C., 2010. A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatr. Scand.* 122 (4), 271–284.
- Posner, M.I., Rothbart, M.K., 2005. Influencing brain networks: implications for education. *Trends Cognit. Sci.* 9 (3), 99–103.
- Schonert-Reichl, K.A., Oberle, E., Lawlor, M.S., Abbott, D., Thomson, K., Oberlander, T.F., Diamond, A., 2015. Enhancing cognitive and social-emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a randomized controlled trial. *Dev. Psychol.* 51 (1), 52.
- Segal, Z.V., Teasdale, J.D., Williams, J.M., Gemar, M.C., 2002. The mindfulness-based cognitive therapy adherence scale: inter-rater reliability, adherence to protocol and treatment distinctiveness. *Clin. Psychol. Psychother.* 9 (2), 131–138.
- Shapiro, S.L., Carlson, L.E., Astin, J.A., Freedman, B., 2006. Mechanisms of mindfulness. *J. Clin. Psychol.* 62 (3), 373–386.
- St. Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59 (4), 745–759.
- Steinmayr, R., Ziegler, M., Träuble, B., 2010. Do intelligence and sustained attention interact in predicting academic achievement? *Learn. Indiv. Differ.* 20 (1), 14–18.
- Valentine, E.R., Sweet, P.L., 1999. Meditation and attention: a comparison of the effects of concentrative and mindfulness meditation on sustained attention. *Ment. Health Relig. Cult.* 2 (1), 59–70.
- Van der Oord, S., Bögels, S.M., Peijnenburg, D., 2012. The effectiveness of mindfulness training for children with ADHD and mindful parenting for their parents. *J. Child Fam. Stud.* 21 (1), 139–147.
- Van der Sluis, S., De Jong, P.F., Van der Leij, A., 2004. Inhibition and shifting in children with learning deficits in arithmetic and reading. *J. Exp. Child Psychol.* 87 (3), 239–266.
- Volanen, S., Lassander, M., Hankonen, N., et al., 2020. Healthy learning mind—Effectiveness of a mindfulness program on mental health compared to a relaxation program and teaching as usual in schools: a cluster-randomised controlled trial. *J. Affect. Disord.* (260), 660–669. <https://doi.org/10.1016/j.jad.2019.08.087>.
- Zelazo, P.D., Forston, J.L., Masten, A.S., Carlson, S.M., 2018. Mindfulness plus reflection training: effects on executive function in early childhood. *Front. Psychol.* 9, 208.
- Zimmerman, B.J., Schunk, D.H., 2011. *Handbook of Self-Regulation of Learning and Performance*. Routledge/Taylor & Francis Group.
- Zylowska, L., Ackerman, D.L., Yang, M.H., Futrell, J.L., Horton, N.L., Hale, T.S., Pataki, C., Smalley, S.L., 2008. Mindfulness meditation training in adults and adolescents with ADHD: a feasibility study. *J. Atten. Disord.* 11 (6), 737–746.

Learning through play-in-interaction

K. Aronsson, Stockholm University, Stockholm, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Beyond play dichotomies	658
Play, mediation and everyday disputes: Vygotsky versus Piaget	658
The role of play and games for learning	659
Play as participation	660
Play politics	660
Gameplay, animation and learning	661
Language play	662
Learning what?	662
References	662
Further reading	663

Beyond play dichotomies

Historically, play has often been described as a mode that constitutes a natural way of being, close to children, innocence, and nature. Play has been perceived as a contrast to school, where pupils have been seen as unduly tamed citizens, colonized by methods that are alien to the child itself. Thinkers like Rousseau and Huizinga, and acclaimed authors like Goethe and Tolstoy, have all foregrounded and celebrated the creative nature of play. Consequently, they have advocated pedagogies that offer leeway for the child's own free roaming, and playful explorations. For instance, on his estate in tsarist Russia, Tolstoy educated the children of poor peasants through Socratic dialogues, play, and creative writing.

Within the history of social sciences, play has been subordinated to instrumental goals. Following Sigmund Freud, Erik H. Erikson, Donald Winnicott and other psychodynamic thinkers, play has been seen as the therapeutic method par excellence in clinical work with children. For instance, play-world methods, which are widely used within child psychology and psychiatry, draw on the basic idea that children display and learn to handle envy, jealousy, and relational difficulties through the use of play. Thereby, play becomes both a diagnostic and therapeutic method. Similarly, play has been advocated as an important method for learning. In both cases, diagnostic/therapeutic or instructional goals are set as the privileged or superordinate goals of play. In much of such work, play has not been studied per se but subsumed under ulterior societal goals, such as diagnosis, therapy, or instruction. Moreover, it is situated within a set of implicit dichotomies that form a hierarchy, where play and exploration are generally cast as what is less "developed" or valuable and that has to be upgraded through connections to what is serious. Within such theorizing, play is often discussed in terms of a number of more or less fixed dichotomies, shown in [Table 1](#).

In this chapter, we explore the ways in which some of these binaries can be seen to dissolve when analyzing children's play in contexts of social interaction. Beyond discussions of fixed binaries, sociocultural theorizing (e.g., [Vygotsky, 1976, 2002](#)) casts play as a basic resource in social interactions in that it is presented as the dominant mode in the child's life. The human being should be seen as a playing creature, someone who has an inborn capacity and predisposition for play, what Huizinga has called *homo ludens*, and this goes for both children and adults. Within sociocultural theory, play thus has a constitutive role for development.

Play, mediation and everyday disputes: Vygotsky versus Piaget

Two of the most influential developmental theorists of the twentieth century, Jean Piaget and Lev Vygotsky, both discussed the role of exploration in children's learning, arguing that playful variations constitute core aspects of learning.

Table 1 Play – work dichotomies.

<i>Play</i>	<i>Work (school)</i>
Exploratory learning	Instruction
Individual	Group work
Fantasy play	Play with rules
Role play	Construction play
Simplicity	Complexity
Improvisation	Rules
Nature	Culture

Piaget (1945) claimed that learning always involves adaptation. From the infant's first attempts at world making, learning and cognition draw both on attempts to mimic others, on imitation, and on play and subjective meaning-making. In his terminology, adaptation thus involves both imitation (accommodation) and play (assimilation). For instance, infants may imitate words of their caregivers but endow them with idiosyncratic meaning. In their very first appropriations of words in their mother tongue, Swedish infants may, for instance, employ *pappa* (daddy) for *lampa* (lamp) as the two Swedish words are phonologically fairly similar, and both phenomena tend to be matters of interest within the young child's world making.

In much of his early work in the 1920s–1940s, Jean Piaget had a phenomenological focus, manifested in a series of ethnographic studies of children as philosophers. His phenomenological period led to extensive theorizing on play with objects, on the basis of observations of his three children's actions in nursery-room contexts (Piaget, 1945). He claimed that concept formation involved a constant balancing (equilibration) between the known and the unknown. The child therefore repeats, experiments with, and recombines what is not yet quite mastered. The learning of a first language is a clear example of the ways in which children learn, not through explicit instruction, but through observations, exploration, and experimenting, which can, for instance, be seen in children's hyper-corrections and other self-corrections (for a review of contemporary work, see Rogoff et al., 2003). Within Piagetian theorizing, such exploratory work characterizes all the semiotic modes of young children – language, art, imitation (including pantomime), and fantasy play (role play) – as well as children's play in construction games (Piaget, 1945). A basic notion is that play is a form of intelligence that is parallel to and equivalent to language, imitation, and art in the child's early development. However, in his theorizing, Piaget primarily cast fantasy play as an aspect of the younger children's early learning.

In contrast, Vygotsky (1976, 2002) located both play and game activities at the heart of his theorizing on children's activities and development. In fact, he theorized play (including games) as something that defines much of the child's imagination and combinatorial thinking. In contrast to Piaget, who did not discuss the role of fantasy play in later learning or modernists like Herbert Read (who have discussed children's art work as something that is arrested through schools and schooling), Vygotsky (2002) claimed that the child's imagination continues to develop all though the school years. This is partly an effect of the fact that mediation is an important aspect of learning in sociocultural theorizing. A given mode of mediation, for example, creative writing, may involve difficulties that the child does not encounter in oral story telling (Vygotsky, 2002). Each medium (e.g., speaking vs. writing, words vs. pictures, virtual play vs. play in "real life") involves its specific difficulties or challenges.

The role of play and games for learning

In Vygotskian theorizing, one important role of teachers and other masters is to build on the children's play activities by introducing topics that appeal to their imagination, and by choosing a type of mediation (drama, play, writing, art work or digital activities) that might appeal to a child in relation to the child's age and prior experiences or a specific area of learning. Today, it can be seen that children spontaneously engage in quite complex cognitive activities when engaging in, for instance, digital activities on the Internet. Play as such involves one type of mediation, what Vygotsky called a 'mental tool' in children's development. This also means that play may boost intellectual development all through life.

Drawing on his work on children's thinking around games with rules, Piaget (1932) presented a model of informal learning and social order, claiming that children of school age develop much of their moral thinking from social exchange with peers and, in particular, from intellectual disputes with peers about games with rules (e.g., marbles). Piaget claimed that peer relations are more symmetrical than adult–child dialogues. They therefore engage more wholeheartedly in critical disputes with peers, challenging the other person's thinking. Thereby, they develop perspective-talking skills in ways that they would not do with adult authorities. In contrast to their conflicts with adults, children who argue with peers tend to engage in more extended and more elaborate arguments. However, even though Piaget emphasized the role of peers for critical thinking, he did not document sibling or peer-play dialogues in everyday life, whereas he did document parent–child-play interactions (Piaget, 1945) and adult–child interviews about games with rules (Piaget, 1932).

In contrast, Vygotsky (1976) claimed that all human behavior and social life includes rules. Therefore, it is not meaningful to cast fantasy play as an activity that is necessarily less complex than game activities or construction play. In Russian, there is moreover but one word, "igra", for both play and games. Vygotsky (1976, 2002) explicitly challenged a major divide between role play and play with rules, as well as any major divide between exploration and instruction. Instruction does not arrest imagination and free exploration. In fact, imagination increases as the individual child appropriates more knowledge of the world.

Vygotsky theorized on the role of apprenticeship in children's learning, and on informal master–apprentice relations that form zones of proximal development, where children may surpass themselves, moving into a higher stage of mental processes, if assisted by a master. He claimed that "In play, a child is always above his average age, above his daily behavior; in play, it as though he were a head taller than himself" (Vygotsky, 1976, p. 552). Play thus constitutes a potential zone of proximal development, where children may assist each other in exploring the world. Moreover, play is about recollections (memory) and this means that children collaboratively advance their imaginative faculties, as they gradually appropriate more knowledge. This, of course, also means that collaborative play may profit from the accumulated prior experiences of all co-present peers.

Whereas Piaget celebrated verbal dispute and the developmental role of games with rules, Vygotsky valued all kinds of play, including fantasy play or role play. Along with Piaget's focus on peer disputes, Vygotsky's foregrounding of language and mediation, and his view that play is the child's dominant mode, have prepared the ground for a novel focus on peer dialogues in children's learning. In contemporary language socialization studies, and sociocultural theorizing, play and play communities have indeed

become central foci, and there is a growing awareness of the role of play for children's learning in peer group and family life contexts (Evaldsson, 2005; Goodwin, 1990, 2006, 2007).

Play as participation

Barbara Rogoff et al. (2003) reviewed the various ways in which children acquire new knowledge without formal instruction, but through keen observations and listening-in, what they called "intent participation." Such informal apprenticeship often takes place in different types of horizontal collaborations, for example, between siblings or peers or with adults if children are allowed to freely watch and listen-in. In contrast to the lesson-style conversational practices that often characterize traditional middle-class adults and their ways of socializing children, for example, at dinner time or in school contexts, such informal apprenticeships may involve keen observations, mimicking, and exploration, rather than instruction. These practices may involve various aspects of children's appropriations of cultural know-how: acquiring a first language, learning other tasks, and the appropriation of local norms for what is an accepted community member.

A substantial body of work on informal practices has been devoted to "participation" and its role in informal apprenticeships, and more specifically, to the ways in which people become fully competent members of specific communities of practice. Lave and Wenger (1991) argued that apprenticeship is linked to membership in communities of practice, where members advance from onlookers and peripheral participants to fully accepted members. They primarily highlighted adult communities of practice, such as alcoholics anonymous, where participants gradually move from peripheral to full participation. In the field of play, such emergent participation may of course also involve informal play communities, related to, for instance, sandbox play (Corsaro, 1979), pog-game activities (Sparman and Aronsson, 2003), computer gaming (Sjöblom and Aronsson 2012a, 2012b) or hopscotch (Goodwin, 2006), that is, various play and game activities where children and young people in informal groups learn from each other within informal communities of practice.

Vygotsky (2002) described children's play products as crystallized imagination. In line with his theorizing, recent research shows that such imagination is not located in a social vacuum. Imagination emerges within a given social order, and it is socially organized. The value that is placed on, for instance, play materials tends to be strictly regulated within local communities of practice, where peers play together and in various ways assess each other's play performances and play materials. Children who, for instance, play with pogs (Sparman and Aronsson, 2003) constantly assess each others' pogs and their trade values in ways that establish local hierarchies of what Bourdieu would call symbolic capital.

Within a sociocultural perspective, children's imagination and increasing skills in play and games are related to their communities of practice, be these related to family life and sibling groups or to peer groups, groups of schoolmates or fellow guild members in, say, World of Warcraft. Intent participation is linked to play communities and play hierarchies in these communities in that key emotions (e.g., excitement, affiliation, disaffiliation) are triggered by the play and games of everyday life.

Play politics

Play politics concerns who gets to play, and even more importantly, who does not get to play. Erving Goffman (1971, 1981) often deployed "play" as a metaphor for social life – for example in his discussions of ratified and non-ratified players and for his analyses of play and by-play – testifying to the basic role of play in social life. Play is both non-serious and serious and, in order to understand its full meaning and seriousness, we have to explore how it is situated in social life.

At first glance, children's play worlds may seem totally free and accessible. Yet one of the first matters to be learnt by young children is how to get access to play – to playmates, to peers' sustained attention, and to play materials. In his documentations of young toddlers' play practices, Corsaro (1979) documented how young children acquired various access rituals for gaining the attention of their (potential) playmates. Drawing on Goffman's (1971) notion of access rituals, he showed how such rituals are learnt through prior successful and unsuccessful attempts at achieving joint attention in play. Thereby, Corsaro also illuminated how even very young children engage in quite complex perspective taking when trying to secure access to ongoing play activities. In her work on mixed-age peers, Whalen (1995) has similarly documented how quite young preschoolers engage in complex tactics, calibrating their play directives in relation to peers' prior utterances, and how they delicately avoid alliances that might upset co-participants (and eventually lead to the arrest of play). It could be argued that children form what could be called play communities, where it is important for participants to be able to take part in local teasing, joking, and other play practices. Obviously, such communities are closely tied to participation, and to local rules for inclusion or exclusion in play (for recent work on school age children's play politics, see Karlsson, 2018).

On the basis of a broad review of prior work on play in different cultures and in contemporary society as well as in historical settings, Schwartzman (1978) has documented how play negotiations take up a great proportion of children's playtime throughout different phases in the child's life, and she has advocated increased attention to issues concerning access to play. These negotiations may concern who gets access to material resources as well as who gets to play what role or in what ways or with whom. Play negotiations – or what we would call play politics – thus concern both physical and virtual play resources. In his work on preschoolers' playful improvisations, Sawyer (1997) observed that young children spent more time than older children on planning, negotiating,

and talking about play. Moreover, he found that friendship had an effect that was similar to age and experience in that friends would go more seamlessly into play, needing less time for planning and coordination.

Research indicates that jokes and playful teasings are important conversational resources for securing participation in local play communities. Teasing, threats, and play politics have been documented in great detail by Marjorie Goodwin, who has analyzed social interaction as parts of her ethnographies of children's play practices in neighborhood and schoolyard areas (Goodwin, 1990, 2006). Over the years, she has analyzed ways in which children and adolescents spend much time, negotiating with each other about inclusion and exclusion and play rules, as well as about material resources. Moreover, she has moved beyond gender stereotypes, showing that both boys and girls engage in insults, gossip, teasing, and other play practices that, at times, border on what is aggressive or violent. In fact, verbal insults, at times, take the form of highly complex entire or partial repeats of prior utterances, where the participants build upon each other's contributions. Through creative repetitions of the prior speaker's insults, a child may thus build on peers' repertoires of invectives and insults, building new and more elaborate versions (Goodwin, 1990, 2006).

Such conversational work can also be seen in the play of preschoolers. Traditionally, Piagetian and other theories on play have generally positioned games and rule play as more complex than fantasy play or so-called imaginative play genres. However, contemporary research, based on video recordings of children's play worlds, reveals a picture where role play might be quite complex, requiring substantial account work, that is, social explanations or reason giving, and other types of mitigations on the part of the participants. On the basis of their analyses of the conversational work of young children in play dyads, Aronsson and Thorell (1999) have shown in detail some of the many ways in which preschool children engaged in quite elaborate perspective-taking in order to negotiate about and solve play conflict scenarios (for instance, involving which play figure – a parent doll or child doll – is to use the only available play vehicle (a bike) for moving in ongoing role play).

Gameplay, animation and learning

Björk-Willén and Aronsson (2014) have documented and analyzed how children "animate" play during gameplay in a preschool context in that they, for instance, talk back to game characters, protesting about their ways of treating the player or their co-players "Listen you duck don't laugh at her!" On other occasions, the children warned game characters "watch out, watch out!" Moreover they animated their conversations during gameplay in a number of other ways, for example, through outcries like "oah pow," "oops," along with "ouch," "WOW" and "oh no, oh NO."

Such blurted affective outcries – what Goffman (1978) has called response cries – thus offer conversational resources for animating play. It can be noted that, along with code-switches, they make it possible for someone to signal a change in footing, that is, in his or her overall orientation to events (Goffman, 1979). Work on children's gaming in the everydayness of their homes has shown that quite young children deploy "response cries" in their gameplay with peers as ways of marking affiliative as well as disaffiliative moves in the game (Aarsand and Aronsson, 2009).

Response cries, laughing and shouting also form parts of the social interaction in youths' gameplay when engaging in World of Warcraft and other computer games, that involve guilds and teamwork (Piirainen-Mars and Tainio, 2009; Sjöblom and Aronsson, 2012a, 2012b). In these game contexts, response cries and other improvised commentaries on the game are important for signaling what is newsworthy or noteworthy in an ongoing game. Such outcries signal affiliation or disaffiliation in ways that make sense to the co-players, but, at the same time, they also form ways of qualifying for participation, that is, for being a fully qualified member of a target team in that they show the player's mastery or lack of mastery in making relevant or irrelevant comments. Moreover, response cries can be seen as the players' resources for orienting to their audience, engaging in action aesthetics, making gameplay more interesting to watch for any co-present siblings or peers. Thereby, it might also be possible to secure the sustained interest and participation of any by-stander to the game.

When the players let out spill cries like "oops!" ("ojdå!" in Swedish) or other response cries, they can be seen to accentuate their shifting affiliations and disaffiliations with other players, but at the same time, they also enhance the dramatic aspects of gameplay. Moreover, they at times animate gameplay talk in that they address the enemy's avatars through aggressive verbalizations, animating the enemy through talking to him as a real person. Apprenticeship in gameplay is not only about playing skillfully, it is also about talking and about joking, rejoicing, making glee or pain cries like someone in the know, someone who is a legitimate participant in the game. In children's gameplay, it is not just the game software that animates gameplay, it is recurrently the players themselves who animate gameplay (Aarsand and Aronsson, 2009; Björk-Willén and Aronsson, 2014; Sjöblom and Aronsson, 2012a, 2012b).

Within the social context of an afterschool gameplay café (game hall) environment, Sjöblom and Aronsson (2012a, 2012b) have moreover shown that youths in their late teens recurrently get engaged in lengthy "negotiations" during the gaming or, soon afterward, about who was to blame for a poor strategy or who was ultimately accountable for that the team lost the game. Another aspect of informal learning in computer gaming is thus how the players engage in negotiations both about what has happened and hypothetical discussions about what might not have happened if a co-player (other member of the same team) had, for instance, not made specific disastrous moves. Such hypothetical talk at times turns into detailed discussions about in what order things "really" happened or why or how something started to go wrong, and who was to blame for the team's failure to win over the enemies in, for instance, World of Warcraft. In terms of reasoning, the players at times engaged in quite complex arguments and counterarguments, and these post-game discussions at times turned into quite sophisticated, somewhat academic disputes.

Language play

Imitation or repetition in classroom contexts has traditionally been seen as an impoverished way of acquiring new knowledge. But in their work on children's language learning and joking in a first-grade Swedish immersion classroom for migrant children, [Cekaite and Aronsson \(2004, 2005\)](#) documented how the children recurrently engaged in repetitions as a type of meta-pragmatic play. Such play took place as indirect playful comments on other persons' ways of saying things. For instance, children enacted performances of the ways in which persons of authority, like teachers, would say something. During informal work in the classroom, the children and their peers also made funny role enactments or teased each other for their "funny" ways of talking. The pupils would, for instance, say something where they spontaneously deployed a "teacher voice," talking in an overly authoritative or somewhat stylized type of teacher voice, and engaging in improvised role reversals. On other occasions, repetitions involved playful and creative recyclings where the children accentuated parts of words that invoked something tabooed. During informal spontaneous conversations in the classroom, peers made such role enactments for the mutual enjoyment of the classroom community. In order to keep conversations going with classmates, the children had to find ways of capturing and attaining the sustained attention of their peers. As noted by [Cathcart-Strong \(1986\)](#) in her discussion of second language (L2) learning in a preschool setting, one of the unwritten conversational maxims of peer group life was to "be entertaining!" In order to recruit play partners, you have to be someone who is fun to play with.

At times, children's enactments involve artful alliteration and other forms of poetics, as well as code switches and variations in pitch, volume, and voice quality. In his original paper on changes of footings, [Goffman \(1979\)](#) discussed code switches as one method for highlighting speaker's playful accentuations of what is newsworthy or remarkable, and code switches have been documented in children's play, for instance in multilingual school settings (e.g., [Cromdal and Aronsson, 2000](#)). Work on playful teasing practices in both gameplay and informal school contexts have shown how children and adolescents create complex tactics, where they build upon prior insults in subtle ways, turning the weapons of their attacker against them as it were ([Tholander and Aronsson, 2002](#)). On other occasions, repetitions merely involved creative partial repetitions of, for instance, teacher talk, accentuating parts of words that invoke something taboo.

In their analyses of a high-school classroom for beginner learners of Swedish as a second language, [Åhlund and Aronsson \(2015\)](#) documented how youths in their teens (16–18 year-old migrants) deployed code switches and stylizations as important resources for creating peer group alignments through joking and laughing together, playfully animating peer's or teacher's ways of speaking that could be seen as noteworthy or laughable. Such playfulness in the classroom was an important part of forming a community of practice in that it generally involved events where the pupils "laughed with" rather "laughed at" each other.

Artful ways of stylizing authorities or of speaking as a teacher (when, for instance, enacting, teacher roles) constitute important conversational resources in classroom performances and classroom struggles about legitimacy as a "clever" student within the local peer group context. Stylizations of teacher talk, or posh ways of speaking English, are deployed by school-age children in their ways of joking with each other and having a laugh, for instance, when deploying mock politeness or mock authoritative fashions of talking to each other. In different ways, these studies all illustrate the need to integrate play in models of language learning.

Learning what?

What do children learn in their play dialogues? Apparently, much of a first language (or a second language in immersion contexts) is learnt through playful imitation (repetitions), and increasingly more complex recyclings and variations. Such informal learning involves different types of renderings, of both what is said and how it is said, that is, matters of style that require so-called pragmatic awareness. Children are neither formally instructed in their first language, nor in how this language should be used in terms of everyday diplomacy or politics, that is, when and how to deploy mitigations or other so called politeness strategies. Yet they obviously acquire quite broad repertoires of styles that can be deployed to upgrade or downgrade the intensity (and offensiveness) of potentially sensitive messages. Play dialogues are important sites for displaying and learning such strategies. If we assume that there is substantial learning through intent participation, it is possible to explain the skills with which even preschool-age children engage in intricate play dialogues, adopting the perspective of others, and predicting the ways in which conversational strategies may upgrade or mitigate potential challenges to the hearer.

In much of the work discussed here, role play certainly appears to be as complex as play with rules in that children have to take into account both their co-participants' past moves, and their projected uptake. At the same time, all of this has to be calibrated in terms of local rules of comportment, and local diplomatic protocols for what might be said and in what ways. Yet, this largely takes place without any explicit instruction. It means that play theories should move away from any rigid dichotomies between role play, on the one hand, and play with rules, on the other. [Vygotsky \(1976\)](#) claimed that there is no major divide between work and play, and in children's lives, play is serious business. Contemporary studies show that play dialogues involve important relational work in that children display and acquire conversational tactics for handling conflicts. As discussed, this can be seen in children's access rituals, negotiations and play politics, as well as in their sustained participation in play activities.

References

- Aarsand, P., Aronsson, K., 2009. Response cries and other gaming moves. Toward an intersubjectivity of gaming. *J. Pragmat.* 41, 1557–1575.
 Åhlund, A., Aronsson, K., 2015. Stylizations and alignments in a L2 classroom: multiparty work in forming a community of practice. *Lang. Commun.* 43, 11–26.

- Aronsson, K., Thorell, M., 1999. Family politics in children's play directives. *J. Pragmat.* 31, 25–47.
- Björk-Willén, P., Aronsson, K., 2014. Preschoolers' animation of computer games. *Mind Cult. Activ.* 21 (4), 318–336.
- Cathcart-Strong, R., 1986. Input generation by young second language learners. *Tesol Q.* 20, 515–530.
- Cekaite, A., Aronsson, K., 2004. Repetition and joking in children's second language conversations. *Playful recyclings in an immersion classroom. Discourse Stud.* 6, 373–392.
- Cekaite, A., Aronsson, K., 2005. Language play, a collaborative resource in children's L2 learning. *Appl. Linguist.* 26 (2), 169–191.
- Corsaro, W., 1979. 'We're friends, right?' Children's use of access rituals in a nursery school. *Lang. Soc.* 8 (3), 315–336.
- Cromdal, J., Aronsson, K., 2000. Footing in bilingual play. *J. Sociolinguistics* 4, 435–457.
- Evaldsson, A.-C., 2005. Staging insults and mobilizing categorizations in a multiethnic peer group. *Discourse Soc.* 16, 763–786.
- Goffman, E., 1971. *Relations in Public*. Basic Books, New York, NY.
- Goffman, E., 1978. Response cries. *Language* 54 (4), 1787–1815.
- Goffman, E., 1979. Footings. *Semiotica* 25, 1–29.
- Goffman, E., 1981. *Forms of Talk*. Basil Blackwell, Oxford.
- Goodwin, M.H., 1990. *He-Said-She-Said. Talk as Social Organization Among Black Children*. Indiana University Press, Bloomington, IN.
- Goodwin, M.H., 2006. *The Hidden Life of Girls. Games of Stance, Status and Exclusion*. Blackwell, Oxford, UK.
- Goodwin, M.H., 2007. Occasioned knowledge exploration in family interaction. *Discourse Soc.* 18, 93–110.
- Karlsson, S., 2018. 'Do you know what we do when we want to play'? Children's hidden politics of resistance and struggle for play in a Swedish asylum centre. *Childhood* 25 (3), 311–324.
- Lave, J., Wenger, E., 1991. *Situated Learning. Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Piaget, J., 1932. *The Moral Judgment of the Child*. Routledge Kegan Paul, London.
- Piaget, J., 1945. La formation du symbole chez l'enfant. Imitation, jeu et rêve, image et représentation. ((1951); *Play, Dreams, and Imitation in Childhood*. Norton; Delachaux et Niesté, New York and Paris.
- Pirainen-Marsh, A., Tainio, L., 2009. Other-repetition as a resource for the activity of playing a video-game. *Mod. Lang. J.* 93, 153–169.
- Rogoff, B., Paradise, R., Arauz, R.M., Correa-Chá vez, M., Angelillo, C., 2003. Firsthand learning through intent participation. *Annu. Rev. Psychol.* 54, 175–203.
- Sawyer, R.K., 1997. *Pretend Play as Improvisation*. Erlbaum, Mahwah, NJ.
- Schwartzman, H., 1978. *Transformations. The Anthropology of Children's Play*. Plenum Press, New York, NY.
- Sjöblom, B., Aronsson, K., 2012a. Participant categorizations of gaming competence: Noob and imba as learner identities. In: Erstad, O., Sefton-Green, J. (Eds.), *Identity, Community and Learning Lives in the Digital Age*. Cambridge University Press, Cambridge, pp. 181–197.
- Sjöblom, B., Aronsson, K., 2012b. Disputes, stakes and game involvement: facing death in computer gaming. In: Danby, S., Theobald, M. (Eds.), *Disputes in Everyday Life: Social and Moral Orders of Children and Young People. Sociological Studies of Children and Youth*, 15. Emerald Books, Cambridge, pp. 377–405.
- Sparman, A., Aronsson, K., 2003. Pog game practices, learning and ideology. Local markets and identity work. In: Walford, G. (Ed.), *Investigating Educational Policy Through Ethnography*. Elsevier, Amsterdam, pp. 169–192.
- Tholander, M., Aronsson, K., 2002. Teasing as serious business. Collaborative staging and response work. *Text* 22, 559–595.
- Vygotsky, L., 1976. Play and its role in the mental development of the child. In: Bruner, J., Jolly, A., Sylva, K. (Eds.), *Play. Its Role in Development and Evolution*. Penguin Books, Harmondsworth, pp. 537–554.
- Vygotsky, L., 2002. Imagination and creativity in childhood. *J. Russ. East Eur. Psychol.* 42 (1), 7–97.
- Whalen, M.R., 1995. Working toward play. Complexity in children's fantasy activities. *Lang. Soc.* 24, 315–348.

Further reading

- Andersen, E., 1990. *Speaking With Style. The Sociolinguistic Skills of Children*. Routledge Kegan Paul, London.
- Aronsson, K., 2018. Daily practices and the time politics of family life. In: Hedegaard, M., Aronsson, K., Højholt, C., Ulvik, O.S. (Eds.), *Children, Childhood and Everyday Life. Information Age Publishing, Charlotte, NC*, pp. 37–53.
- Aronsson, K., 2020. Myths about bilingual learning in family life settings: Leopold's child language biographies and contemporary children's play practices. *Learn. Cult. Soc. Interact* 26, 1–9. <https://doi.org/10.1016/j.lcsi.2018.03.011>.
- Aronsson, K., Cekaite, A., 2011. Activity contracts and directives in everyday family politics. *Discourse Soc.* 22 (2), 137–154.
- Aronsson, K., Gottzén, L., 2011. Generational positions at family dinner: food morality and social order. *Lang. Soc.* 40 (4), 1–22.
- Aronsson, K., Hedegaard, M., Højholt, C., Skjær Ulvik, O., 2018. Rethinking development: situated studies of children's perspectives. In: Hedegaard, M., Aronsson, K., Højholt, C., Ulvik, O.S. (Eds.), *Children, Childhood and Everyday Life. Information Age Publishing, Charlotte, NC*, pp. 1–13.
- Aronsson, K., Morgenstern, A., 2021. "Bravo!": Co-constructing praise in French family life. *J. Pragmat.* 173, 1–14. <https://doi.org/10.1016/j.pragma.2020.12.002>.
- Aronsson, K., Thorell, M., 2002. Voice and collusion in adult-child talk. Toward an architecture of intersubjectivity. In: Blum-Kulka, S., Snow, C.E. (Eds.), *Talking to Adults*. Erlbaum, Mahwah, NJ, pp. 227–293.
- Garvey, C., Kramer, T., 1989. The language of social pretend play. *Dev. Rev.* 9 (4), 364–382.
- Gottzén, L., Aronsson, K., 2019. "A Darwinist at our dinner table?": discursive transitions between school and home. In: Hedegaard, M., Fleer, M. (Eds.), *Children's Transitions in Everyday Life and Institutions*. Bloomsbury.
- Huizinga, J., 2004. *Homo Ludens. Natur och Kultur, Stockholm* (original in German, 1938).
- James, A., 1996. Learning to be friends. Methodological lessons from participant observations among English schoolchildren. *Childhood* 3, 313–330.
- Maynard, D., 1985. How children start arguments. *Lang. Soc.* 14 (1), 1–30.
- Pauletto, F., Aronsson, K., Galeano, G., 2017. Endearment and address terms in family life: children's and parents' requests in Italian and Swedish dinnertime interaction. *J. Pragmat.* 109, 82–94.
- Rindstedt, C., Aronsson, K., 2012. Children's intent participation in a pediatric community of practice. *Mind Cult. Activ.* 19 (4), 325–341.
- Sheldon, A., 1996. You can be the baby brother, but you aren't born yet. Preschool girls' negotiations for power and access in pretend play. *Res. Lang. Soc. Interact.* 29 (1), 57–80.

Optimizing learning through retrieval practice and spacing

Shana K. Carpenter, Department of Psychology, Iowa State University, Ames, IA, United States

© 2023 Elsevier Ltd. All rights reserved.

Retrieval practice	664
Spacing	666
Retrieval practice and spacing combined	667
Students' use of retrieval and spacing	668
Potential reasons for students' study decisions	668
Toward successful implementation of retrieval practice and spacing	669
Summary and conclusions	670
Acknowledgment	671
References	671

Now more than ever there is a need for effective ways to learn information. Advances in technology, rapid increases in the amount and accessibility of information, and more flexible course delivery methods are bringing forth changes to the landscape of education that raise questions about the best ways to design and deliver instruction. Such questions will most certainly increase as we head into a future that will involve further challenges, uncertainties, and increasingly non-traditional approaches to education.

One of the most valuable things that research can offer to education, therefore, is reliable evidence-based practices that stand the test of time. When it comes to learning, a great deal of scientific effort has been devoted to understanding how the learning process works and how to improve it. In particular, over 100 years of cognitive science research has revealed specific strategies and techniques that substantially and reliably boost student learning, and that can be readily applied in any course material. This chapter reviews two of the most powerful techniques to emerge from this research: *retrieval practice* and *spacing*. The sections that follow include a description of these techniques, the research demonstrating their effectiveness, how often they are utilized by students, and practical recommendations for incorporating the techniques into education.

Retrieval practice

Retrieval practice involves practicing to retrieve information from memory. For example, after reading a section of a textbook chapter on the visual system, a student might lay the text aside and try to recall everything that she has just read. This could include anatomical information about the eye, the function of rods and cones, and visual pathways within the brain. Even if she does not remember everything perfectly, the act of trying to remember it and then checking her accuracy makes it more likely that she will successfully remember that information again in the future.

Fig. 1 shows the design of one recent study on retrieval practice. Rowley and McCrudden (2020) had middle school students learn information about overfishing by first reading a text passage and taking notes over the information, and then completing a follow-up activity in which they either re-wrote the information from their notes or tried to retrieve the information from the passage. On a later essay-type test over the information, the group that practiced retrieval scored significantly higher than the group that re-wrote their notes.

A large number of studies has shown similar benefits of retrieval practice. One of the earliest studies was carried out by Sones and Stroud (1940), who found that seventh grade students learned information about the history of paper significantly better by reading a text passage and then answering questions about it, compared to reading the passage and then rereading it over again. Of particular importance are studies carried out in authentic educational environments showing the real-world advantages of retrieval practice. When compared to non-retrieval-based activities such as re-reading or additional studying of the information, retrieval has produced significant benefits on elementary school students' learning of scientific information (Jaeger et al., 2015), new vocabulary (Goossens et al., 2014), and spelling skills (Jones et al., 2016), as well as on middle and high school students' learning of history (Carpenter et al., 2009; McDermott et al., 2014), social studies (Roediger et al., 2011), and science (Rowley and McCrudden, 2020). Retrieval practice has also been shown to benefit university students' learning in a variety of courses, including biology (Carpenter et al., 2016), psychology (Ebersbach et al., 2020), anatomy (Dobson et al., 2017), and mathematics (Hopkins et al., 2016), as well as medical residents' learning of information about medical conditions (Larsen et al., 2009).

The exact way that retrieval practice is done varies somewhat across studies. Fortunately, the technique is quite robust and can be flexibly adapted in a number of ways. The key element involves having some opportunity to retrieve information that is being learned. This could take the form of practice questions that students work on during class or in a homework assignment, "clicker" questions that are part of the class lesson, or online review questions that students complete outside of class. A variety of different types of questions (e.g., multiple-choice, short answer) are beneficial as long as they require students to retrieve the answers

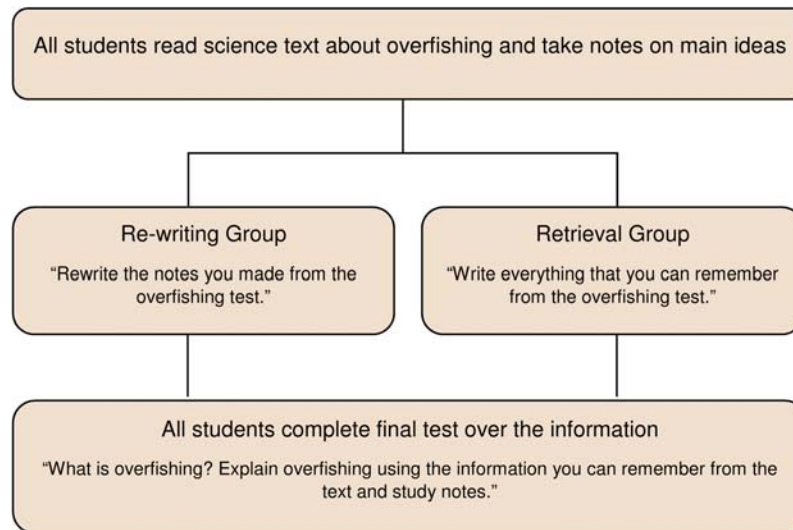


Fig. 1 Overview of the design from Rowley and McCrudden (2020).

(McDermott et al., 2014). Classroom-based studies have shown consistent advantages of retrieval practice in the form of low-stakes quizzes administered in class or online, with a variety of different question types (Sotola and Credé, 2021).

Many studies have shown benefits of retrieval practice on memory of the same information that was retrieved. For example, if students practice retrieving the English translations for Spanish vocabulary words (e.g., Gato - ?), this will boost their future memory for those same vocabulary words. The benefits of retrieval can also be flexible, however, such that they can apply to situations in which students might need to apply their knowledge to answering a question that they have not seen before. For example, Dobson et al. (2019) had undergraduate students learn texts on physiology topics through reading the texts multiple times, versus reading the texts and then trying to retrieve them. On a test one week later, students recalled more information from the texts that they had learned through retrieval compared to rereading. Retrieval practice also facilitated students' performance on transfer questions that required them to generalize knowledge they learned from the texts. For example, after practicing to retrieve a text containing information about how aerobic adenosine triphosphate (ATP) production is affected by oxygen pressure in the atmosphere, students had more success answering a question about how cardiorespiratory function would be affected by a hypoxic (high-altitude) environment.

Studies exploring the durability of retrieval practice have shown that these benefits are long-lasting. Practicing retrieval at one point in time boosts learning of that information even on a final test that is given several weeks later (Carpenter et al., 2008), or several months later (Carpenter et al., 2009). Thus, retrieval practice is an effective learning method that is flexible and long-lasting, and can be adapted to fit any course material.

Why is retrieval practice beneficial? Researchers have conducted many studies on the theoretical and neuroscientific mechanisms involved in retrieval (e.g., Carpenter and Yeung, 2017; Carpenter, 2011, 2009; Karpicke et al., 2014; Rickard and Pan, 2018; Pyc and Rawson, 2010; Wing et al., 2013). This research has shed light on the mental processes that occur during retrieval and how they contribute to learning. Perhaps of greatest practical importance, one reason that retrieval is beneficial is because it reveals the state of our knowledge. That is, it effectively tells us what we know and do not know. Only by trying—and sometimes failing—to retrieve something can we be aware of what we know (information that we can retrieve successfully) and what we do not know (information that we could not retrieve). Such awareness is critical to successful learning because it identifies the information that is not yet learned and needs to be studied more.

Accurate awareness of one's own knowledge does not occur naturally. Students have a strong tendency to overestimate what they know. For example, in some classroom studies, prior to taking an exam, students are asked to predict what they think their grade on the exam will be. These studies consistently show that students' predictions of their grades are higher than the actual grades that they earn on the exam (Foster et al., 2017; Hacker et al., 2000; Hartwig and Dunlosky, 2017; Miller and Geraci, 2011). The tendency to overestimate one's knowledge is extremely common and has been demonstrated in a multitude of studies, ranging from simple laboratory-based memory tasks (Carpenter and Olson, 2012; Finn and Metcalfe, 2007) to exams in real courses. This tendency extends beyond academic situations too, as surveys of the general public show that people commonly overestimate their own knowledge, skills, and abilities in a variety of domains (Dunning et al., 2004).

Overconfidence is so common, it is possible that it reflects a general "optimism bias" that people adopt without even being aware of it. In the context of a learning situation, overconfidence can be detrimental because a student may choose not to put forth effort into learning something if they believe they already know it. To cease studying prematurely then leads to the unfortunate reality check when students see their exam scores and realize they did not know the information as well as they thought they did.

Retrieval practice provides a simple solution to this all-too-common problem. By using retrieval as a strategy for learning information, students get this reality check at an early stage, well in advance of a high-stakes exam. This allows them to focus their study

efforts where they are needed most, on the information that they were unable to retrieve. Retrieval thus serves as a diagnostic check that helps students guide their learning in a way that maximizes their success. Indeed, research shows that retrieval practice improves students' awareness of their own learning, a key step in guiding study decisions in a more informed and effective way (Agarwal et al., 2008; Little and McDaniel, 2014; Tullis et al., 2013).

Spacing

Learning is not easy. Even if we use the right learning techniques, achieving full understanding of a concept requires us to devote sufficient time and effort to learning it. When we encounter a new concept for the first time, we do not instantly understand it 100%, but instead have to engage in further study, thought, and reflection about that concept.

Perhaps it goes without saying that repeated practice is necessary for learning. However, a question that is not so obvious to answer is how that repeated practice should be scheduled. When learning to apply mathematical formulas to story problems, for example, is it best to practice applying the same formula over and over again right away? Or, should the formula be practiced a little at a time on different occasions?

Given the same overall number of learning opportunities, research shows that those opportunities are more effective if they occur separated in time, rather than immediately back-to-back. This finding is known as the *spacing effect* (also sometimes referred to as *distributed practice*), due to the advantage of repeated learning opportunities that are "spaced out" or distributed in time.

The spacing effect was first reported by Ebbinghaus (1885/1913), who measured his own memory for nonsense syllables and found that repeated attempts to learn the syllables were more effective if they were carried out at different points in time, rather than within the same session. The spacing effect has since become one of the most widely-studied and consistent findings in the learning sciences, having been demonstrated in several hundred laboratory- and classroom-based studies.

In a typical study on the spacing effect, students have two or more opportunities to learn some information. If those opportunities are spaced, they occur at different points in time. If they are not spaced (often referred to as *massed*), the same learning opportunities occur back-to-back immediately. Even though the number of learning opportunities and overall time spent learning are the same in both cases, spaced scheduling leads to significantly better learning.

Fig. 2 shows the design of one recent study on the spacing effect. Emeny et al. (2021) had mathematics students complete 12 practice problems on permutations and Venn diagrams. Following a tutorial on these topics, students worked through four practice problems in each of three sessions one week apart (spaced), or they completed the problems all in the same session (massed). On a test four weeks later containing new problems over these same concepts, performance was significantly greater for students who practiced the problems using the spaced schedule compared to the massed schedule.

A number of additional studies conducted in authentic educational environments show real-world advantages of spacing. For example, spacing has produced significant learning enhancements in a variety of educational contexts, including elementary school children's learning of new vocabulary (Sobel et al., 2011), language skills (Ambridge et al., 2006; Seabrook et al., 2005), and mathematics (Schutte et al., 2015), middle school students' learning of biology (Randler et al., 2008) and mathematics (Nazari and Ebersbach, 2019), and high school students' learning of geometry (Yazdani and Zebrowski, 2006) and French vocabulary (Bloom and Shuell, 1981). Spacing has also been shown to benefit university students' learning of mathematics (Lyle et al., 2020), English language skills (Miles, 2014), and meteorology (Kapler et al., 2015), along with medical students' learning of nutrition (Raman et al., 2010), laparoscopic tasks (Spruit et al., 2015), and surgical skills (Moulton et al., 2006; van Dongen et al., 2011).

The exact way that spacing is carried out varies across studies. Spaced learning opportunities are sometimes scheduled on consecutive days, and other times with longer periods of time in-between such as days or weeks. In studies involving more than two learning opportunities, the time in-between those learning opportunities can also be evenly-spaced (e.g., two days in-between each learning opportunity) or non-evenly spaced (e.g., with the time in-between each learning opportunity increasing or decreasing). Like retrieval practice, spacing is a robust technique that can be flexibly adapted. The key is to arrange the spaced learning opportunities with enough time in-between to allow some forgetting of the information, but not with so much time in-between that the initial learning opportunity has been completely forgotten. Many of the classroom studies on spacing have demonstrated significant learning benefits by scheduling repeated learning opportunities on consecutive days or every few days.

Although many studies show that spacing benefits memory for the same information that was practiced, spacing also benefits the ability to apply the learned information to new situations. In one study (Vlach and Sandhofer, 2012), elementary school children learned about food chains. The children were taught general scientific principles (such as the tendency for larger animals to eat

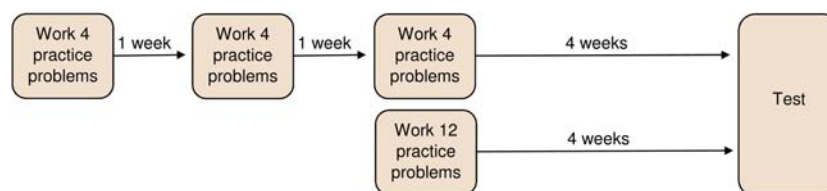


Fig. 2 Overview of the design from Emeny et al. (2021).

smaller animals) within a particular biome such as an ocean, desert, grassland, swamp, or arctic environment. Four lessons over this information were given during the same day (massed), spread across two days with two lessons per day (clumped), or spread across four days with one lesson per day (spaced). On a test one week later, children were asked to generalize these principles to a biome that they had not previously learned about. For example, if they learned about the grassland biome (and the tendency for a snake to eat a rat) but did not learn about the desert biome, a question on the one-week delayed test asked specifically about the tendency for a particular desert animal (e.g., the scorpion) to eat other desert animals (e.g., lizard, beetle, cactus, or fox). Children performed significantly better on this test if they had received the spaced lessons one week prior, compared to the clumped or massed lessons. Along similar lines, other studies have shown that spacing enhances students' application of other knowledge and skills, such as diagnosis of clinical disorders (Zulkiply et al., 2012) and proficiency at evaluating the credibility of information on websites (Foot-Seymour et al., 2019).

The benefits of spacing are long-lasting, having been demonstrated over time intervals of several weeks (Kapler et al., 2015; Sobel et al., 2011), several months (Carpenter et al., 2009), and in one study up to a year (Cepeda et al., 2008). Thus, like retrieval practice, spacing is an effective learning method that is flexible and long-lasting, and can be adapted to fit any course material.

Why is spacing beneficial? Many studies have looked into the theoretical and neuroscientific mechanisms involved in spacing. These studies have revealed important insights into the mental processes involved in repeated practice and how those processes change when the spacing between repeated practice opportunities is increased (for recent reviews, see Carpenter, 2017, 2021). An early proposal that is still quite relevant today is the idea that spacing promotes students' attention to what they are learning (Hintzman et al., 1973). When students learn something through a spaced schedule rather than a massed schedule, the amount of time spent in a given learning session is shorter. For example, a student who spends a total of 90 min studying anatomical terms could study those terms repeatedly in the same 90-min period (massed) or could study the same terms for 30 min at a time on three different days (spaced). Even when learners try hard to focus on a task, attention naturally declines over time, such that a longer learning session is less likely to receive full attention than is a shorter learning session. Thus, even with sustained effort, students are less likely to be able to devote their full attention to the single 90-min session than to the three different 30-min sessions.

It has also been proposed that spacing is beneficial because it engages retrieval practice (Benjamin and Tullis, 2010). When information is learned in a massed fashion, the same information repeats itself on such a short timeline that the previous occurrence of that information is still fresh in mind and does not need to be retrieved. When information is spaced, however, enough time goes by in-between repetitions such that the earlier occurrence of that information is not still in mind. In this case, a given piece of information (e.g., a foreign language vocabulary item) can remind students, and encourage retrieval, of the previous occurrence of that information. Indeed, studies have shown that spacing is more effective for learning when the spaced repetitions encourage retrieval of the earlier repetitions (Wahlheim et al., 2014).

Retrieval practice and spacing combined

Retrieval practice and spacing can be combined to produce better learning than either technique alone. Although spacing is significantly better than massing even when information is simply presented without requiring retrieval, spacing that involves retrieval practice is even more beneficial. For example, Dobson et al. (2017) had kinesiology students learn information about muscles by reading the information repeatedly, or by reading the information and then trying to retrieve it. Students either read the information multiple times across two or three consecutive days, or they alternated between reading the information and trying to retrieve it the same number of times across the same number of consecutive days. These spaced strategies were compared to a massed strategy in which students engaged in both reading and retrieval of the information all on the same day. On a test over the information given one week later, performance was highest for information learned through reading and retrieval spaced across three consecutive days.

One particularly effective method that combines retrieval and spacing is known as successive relearning. Here, students take some material that they are learning (e.g., foreign language vocabulary) and they practice retrieving it, with feedback, until they retrieve all of the information 100% correctly. Students then repeat the process again for a pre-determined number of times at particular time intervals. For example, Bahrck (1979) had students retrieve Spanish vocabulary words until they got each of them right. Students then engaged in five additional learning sessions in which they practiced retrieving the same words again until they got them all right. These "relearning" sessions either occurred back-to-back (massed), spaced apart by one day, or spaced apart by 30 days. Although it was harder to retrieve the Spanish words every 30 days—that is, students forgot more of the information in-between the learning sessions—the 30-day interval led to superior long-term learning on a final test given 30 days later. A follow-up study showed that the learning advantage for the group that practiced at 30-day intervals was still present even after several years (Bahrck and Phelps, 1987).

Successive relearning at spaced intervals has proven to be quite effective for enhancing students' learning of course information. Rawson et al. (2013) had undergraduate psychology students learn the definitions for concepts (e.g., classical conditioning) by practicing to retrieve them, with feedback, until they retrieved all of the definitions 100% correctly three times. Two days later students practiced retrieving again until they could retrieve each definition correctly one time. Compared to learning concepts through additional study opportunities at the same time intervals, successive relearning resulted in significantly higher exam scores. This was the case even though the exam questions covered the same concepts but were not identical to the practice questions that students encountered during successive relearning. Furthermore, the advantage of successive relearning was still present on a follow-up test over the concepts that was given more than three weeks later.

Students' use of retrieval and spacing

Retrieval practice and spacing are highly effective techniques for boosting learning. We have seen that they are also easy to implement and can be adapted to any learning material. These techniques therefore have great potential to enhance students' academic achievement. However, research indicates that these techniques are underutilized in education.

Some of this research involves surveys in which students are asked about the methods they use to learn information in their courses. About 70% of undergraduate students report using retrieval-based methods such as practice questions (Geller et al., 2018; Hartwig and Dunlosky, 2012). However, when asked why they use these methods, the majority of students report that they use practice questions and other retrieval-based methods to determine how well they have learned the material. This certainly makes sense, as trying to retrieve something is a direct way to see if that information has been learned. Importantly however, this research indicates that students are largely unaware of the direct benefits of retrieval on learning. Indeed, a much smaller portion of students in these studies (under 30%) reported using retrieval-based methods because they believe these methods help them learn. Thus, students view retrieval primarily as a means of checking their learning, rather than as a means of directly enhancing their learning.

Surveys show that students underutilize spacing in their studying. In the study by Geller et al. (2018), only 17% of students reported that they space out their studying over multiple days or weeks. In a study by Corral et al. (2020), only 8% of students reported that they do their studying throughout the semester as they learn the information. These and other surveys show that instead of spacing their studying, students tend to do most of their studying just a handful of days before the exam (Blasiman et al., 2017).

Other research involves collecting observational data on students' choices to engage in retrieval and spacing. In some studies, questions were made available in the form of practice quizzes prior to mandatory exams. Students were given the option to use the practice quizzes if they wanted to, but the quizzes were optional and not incentivized. Because practice quizzes represent a form of retrieval practice, the number of students who choose to complete practice quizzes can give an estimate of how often students engage with retrieval practice. In one study involving a practice quiz in multiple courses, the completion rate of the quiz varied by course, such that the quiz was completed by about 40%–50% of students in psychology courses, about 60% of students in a genetics course, about 66% of students in an ecology course, and 91%–92% of students in medical and zoology courses, respectively (Carillo-de-la-Peña et al., 2009). In other studies, completion rates of an optional practice quiz were about 46% of students in dentistry courses (Olson and McDonald, 2004), and about 75% of students in pre-med courses (Velan et al., 2008).

In other studies, practice quizzes were made available, but in smaller doses such that short quizzes over different parts of the course material were provided at regular time points throughout the academic term. For example, Johnson (2006) provided two optional quizzes over each of 14 textbook chapters in an educational psychology course. In two other studies, a short review quiz was made available to students after each individual class meeting in introductory biology (Carpenter et al., 2017) and introductory psychology (Corral et al., 2020). Students were made aware of these quizzes by the instructors of the courses and were encouraged to use them as a study aid. Even so, the completion rate of the quizzes was quite low. The percentage of students who completed all of the available quizzes was fewer than 10% in Johnson's study, about 30% in the study by Carpenter et al. and fewer than 4% in the study by Corral et al.

What about students' use of spacing? One way to examine this is to look at when students choose to complete practice quizzes. If practice quizzes are made available online after each topic is covered in class, do students complete those quizzes in a manner that reflects spacing (for example, completing each quiz a few days after each topic is covered)? Or do they tend to delay completion of all quizzes until shortly before the exam? Although few studies have tracked the timing of students' completion of optional online practice quizzes, the consistent pattern from these studies is that students tend to delay completion of the practice quizzes until shortly before the date of a mandatory exam or other type of assessment (Corral et al., 2020; see also Lui et al., 2019). Thus, the date of the mandatory assessment, rather than spacing, appears to be a primary factor influencing the decision of when to complete practice quizzes. These studies show that students have a tendency to "bunch" all of the quizzes into a brief time window shortly before they will be tested over the information, which is consistent with an approach based more on cramming than spacing.

Potential reasons for students' study decisions

Why do students avoid retrieval and spacing? Research suggests that these tendencies could arise from students' beliefs about the effectiveness of these techniques. More specifically, retrieval and spacing require effort. It is harder to try to retrieve something from memory than to simply read it again, and it is harder to retrieve something that was seen two days ago compared to something that was just seen a moment ago. The fact that retrieval and spacing involve effort has been confirmed through studies in which students are asked to use these techniques and then report how much mental effort was involved. In these studies, students rate retrieval as more effortful than study (Foster et al., 2018; Hanham et al., 2017; van Gog et al., 2011), and they also rate spacing as more effortful than massing (Kirk-Johnson et al., 2019).

Despite the benefits of retrieval and spacing, there is a common tendency for students to perceive these strategies as counter-effective for learning. For example, Yeo and Fazio (2019) taught participants the solution steps to statistical problems by having them study an example problem with solution steps and then try to retrieve the problem and solution, versus simply reading

the problem and solution steps that were already provided. Afterward, participants rated how likely they would be to successfully recall this information on a later test. Even though participants recalled significantly more of the information on a one-week delayed test if they had learned through retrieval compared to reading (i.e., the usual advantage of retrieval practice), participants' predictions of their own performance were the opposite. That is, participants believed they had learned the information significantly better after reading it compared to trying to retrieve it. Similarly, Foster et al. (2018) found that students who were learning probability problem solutions believed that reading an example problem with solution steps was more effective for learning, compared to practicing to solve the problem by retrieving the solution steps.

Along similar lines, studies of the spacing effect have involved opportunities to learn information through spacing versus massing, and then rating the effectiveness of these methods for learning. In one study, participants were taught to identify the characteristics of different butterfly species (Birnbaum et al., 2013). Participants saw several examples of butterfly species, either presented in a way that the examples representing the same species occurred back-to-back (massed), or with other examples from different species in-between (spaced). On a later test requiring participants to classify new examples from the same species, they performed better if they had learned the species through spacing compared to massing. However, when asked which method they believed helped them learn better, the majority of participants (67%) believed that they learned better through massing. Similar results were observed by Kornell and Bjork (2008), who found that students' ability to classify the painting styles of different artists benefited from spacing compared to massing, but 78% of students believed that massing was as effective or better than spacing. In the study by Emeny et al. (2021), spaced practice led to significant enhancements in mathematics performance, yet students erroneously expressed higher confidence in massed practice.

This research reveals important insights into students' perspectives about retrieval and spacing. Specifically, students find both of these techniques to be effortful and not particularly effective for helping them learn. Recent research has shown that these two perspectives could be related, such that students interpret the effort they experience during learning as a sign that the learning has failed. For example, Kirk-Johnson et al. (2019) conducted a series of studies in which participants learned information (e.g., categorizing different bird families, and reading scientific text passages) through massing versus spacing (Experiment 1), and through retrieval versus reading (Experiment 2). Participants had the chance to experience both strategies (massing and spacing in Experiment 1; retrieval and reading in Experiment 2), and then rated the mental effort involved and the perceived effectiveness of the strategy for helping them learn. Afterward, they were given a new set of material to learn (a new set of bird families in Experiment 1, and a different scientific passage in Experiment 2), and chose which strategy they wanted to use to learn that material. Participants rated retrieval and spacing as more effortful, and less effective for learning, than reading and massing, respectively. Of particular interest, these judgments were linked with participants' decisions of which strategy to use to learn the new information. The majority of participants chose to use massing (68% of students) and reading (69% of students) to learn the new material. Furthermore, the effort involved in using spacing and retrieval was negatively associated with the perceived effectiveness of each technique—that is, more effort invested coincided with perceptions that the technique was less effective—and this in turn predicted participants' decisions about which technique to use. In other words, when learning was effortful it was perceived as less effective, and this influenced students' decisions to use what they believed was the more effective (less effortful) strategy.

Thus, even though retrieval and spacing are quite effective for learning, these techniques do not generally *feel* effective, possibly because of the mental effort they involve. Effort plays an interesting, and sometimes counterintuitive, role in learning. A long tradition of research on “desirable difficulties” shows that the most effective ways to learn are often difficult and require effort (Bjork et al., 2013). Often however, when given the choice students avoid effortful learning situations, opting instead for passive learning experiences that feel smooth and fluent (such as sitting through a lecture or reading over their textbook), but do not actually enhance learning (Carpenter et al., 2020). Such a tendency could stem from a misunderstanding of the role of effort in learning and brings to light the important need to educate students on how to think about and invest effort in order to make the best use of their study time.

Toward successful implementation of retrieval practice and spacing

As demonstrated, students are not naturally inclined to use retrieval and spacing on their own, in part, due to the perceived effort involved and the belief that these techniques are not effective for learning. Further contributing to the difficulty in utilizing these techniques is the fact that they require preparation. Utilizing retrieval practice requires that students have practice questions with good quality feedback. Although some materials may naturally lend themselves to retrieval more readily (e.g., foreign language vocabulary), many others require questions to be generated. Especially when learning material for the first time, students may encounter difficulty or uncertainty in generating their own questions. Along similar lines, spacing requires planning a schedule of when to study, which may be challenging for students if they are uncertain about how best to schedule the spacing of study sessions. Thus, getting students to use retrieval and spacing likely requires instructors to provide the tools to facilitate it.

Fortunately, research has revealed a number of ways that instructors can do this. First, instructors can provide practice questions for students over the content they are learning. Research shows that such questions are effective for enhancing learning whether they are in the form of in-class paper-based questions (Carpenter et al., 2009; Carpenter et al., 2016), in-class “clicker” questions (Mayer et al., 2009; Shapiro and Gordon, 2012), or online practice questions that students complete outside of class (Carpenter et al., 2017; Corral et al., 2020).

Second, instructors can encourage spacing by arranging practice questions and assignments in such a way that previously-learned information is re-visited. Many times instructional materials are set up in a way that does not encourage spacing. Instead, course content is usually “blocked” such that one concept is focused on for a period of time, and then the next concept is introduced and focused on, then the next concept, and so on, without returning to previous concepts. In one recent study, Rohrer et al. (2021) examined middle school mathematics textbooks and found that the ordering of practice problems did not make use of spacing. Instead, practice problems over a given concept (e.g., the Pythagorean Theorem) tended to follow immediately after the chapter on that concept, and practice problems over that concept tended to be grouped together such that a number of problems in a row addressed the same concept. Although such an approach does not make use of spacing, research fortunately shows that spacing can be encouraged by including practice problems from previously-learned concepts into current lessons. For example, Rohrer et al. (2015) observed significant improvements in middle school students’ mathematics achievement by taking some of the practice problems from a given lesson and including them in subsequent assignments over the academic term. Along similar lines, Lyle et al. (2020) found that undergraduate students’ achievement in precalculus was significantly enhanced by completing weekly quizzes that contained practice problems from both current and previous lessons, compared to practice problems from just the current lesson. Implementing spacing in this way is quite straightforward, as it simply involves modifying the schedule of practice problems and could be applied to any academic subject.

Third, successive relearning could be utilized to give students clear and concrete guidance on how to use retrieval and spacing. This method might be particularly useful if students have opportunities to use retrieval and spacing but are not entirely sure how to go about it. By giving students a specific approach to use (e.g., continue practicing retrieval until you get all of the information right), along with a specific schedule (e.g., repeat the process every couple of days), students may find it easier to implement these techniques. Indeed, research shows that providing students with specific guidance about successive relearning increases their tendencies to use it (Dunlosky and Rawson, 2015). The method could be implemented via a low-tech approach in which students make flash-cards and carry out the relearning sessions on scheduled days. Or instructors could make use of available online course management systems and set up the relearning sessions in the form of recurring online quizzes that students complete at scheduled times.

Fourth, making students aware of the benefits of retrieval and spacing could help build their knowledge of the effectiveness of these techniques, and might increase the likelihood that students will use them. In the study by Carpenter et al. (2017) involving biology students’ completion of optional online practice quizzes, after the first two exams students were shown the average exam scores of students who completed the practice questions, which were significantly higher than the exam scores of students who did not complete the practice questions. This coincided with a significant increase in the number of students who completed the practice questions before the next exam. Hui et al. (2021) also observed that giving students individualized feedback about the benefits of practice questions on their own learning led to an increase in students’ decisions to use practice questions.

Fifth, technology can be leveraged to implement retrieval practice and spacing. In one recent study (Cleary et al., 2021), practice questions were delivered to students via smartwatches. Students wore the watches which periodically displayed practice questions over material that they were trying to learn. On a later test, the practice questions led to significantly better memory for the material compared to statements of the same information that were delivered for reading only. This benefit occurred even when students were outside of a supervised class environment and engaged in other things. This study suggests that mobile devices may be an effective tool for delivering practice questions, and the scheduling of such practice questions could be arranged to involve spacing as well. The benefits of retrieval and spacing are due to the processes that these techniques engage, so any mobile device could be utilized to engage these benefits. Indeed, a recent review of the research on educational technology reported that mobile devices can be quite effective for enhancing learning, as long as they engage effective learning principles like retrieval practice and spacing (Yeung et al., 2021).

Finally, instructors can encourage spacing by using cumulative exams. By including exam questions from previously-learned topics, cumulative exams involve spacing of information on the exams themselves, and would also encourage spacing by incentivizing students to continue studying material from previous topics. Consistent with these ideas, cumulative exams have been shown to enhance learning more than non-cumulative exams (Beagley and Capaldi, 2016; Lawrence, 2013).

Summary and conclusions

Retrieval practice and spacing are effective learning techniques with a long history of research supporting their positive effects in educational contexts. Despite the widespread benefits of retrieval and spacing, students are reluctant to make use of these techniques for learning course information. One hindrance to adopting these techniques could be the effort that they involve, and the tendency for students to misinterpret this effort as a sign of ineffective learning. A step forward in encouraging the use of retrieval and spacing, therefore, could be to make students aware of the positive role that effort can play in learning, and to experience for themselves how retrieval and spacing can benefit their own learning.

Given that both techniques require some setup and planning, teachers can facilitate students’ use of retrieval and spacing by providing them with the necessary tools in the form of practice questions, review activities that re-visit previously-learned concepts, opportunities for successive relearning, feedback on their performance, and cumulative exams. Education will most certainly face future challenges and changes that can be hard to predict. By incorporating solid and reliable practices for enhancing learning that are backed by strong research evidence, students and educators can be armed with the tools they need to meet future challenges in a way that promotes durable and meaningful learning.

Acknowledgment

This material is based upon work supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition, Collaborative Grant 220020483.

References

- Agarwal, P.K., Karpicke, J.D., Kang, S.H.K., Roediger III, H.L., McDermott, K.B., 2008. Examining the testing effect with open- and closed-book tests. *Appl. Cognit. Psychol.* 22, 861–876.
- Ambridge, B., Theakston, A.L., Lieven, E.V.M., Tomasello, M., 2006. The distributed learning effect for children's acquisition of an abstract syntactic construction. *Cognit. Dev.* 21, 174–193.
- Bahrick, H.P., 1979. Maintenance of knowledge: questions about memory we forgot to ask. *J. Exp. Psychol. Gen.* 108, 296–308.
- Bahrick, H.P., Phelps, E., 1987. Retention of Spanish vocabulary over eight years. *J. Exp. Psychol. Learn. Mem. Cogni.* 13, 344–349.
- Beagley, J.E., Capaldi, M., 2016. The effect of cumulative tests on the final exam. *Primus* 26, 878–888.
- Benjamin, A.S., Tullis, J., 2010. What makes distributed practice effective? *Cognit. Psychol.* 61, 228–247.
- Birnbaum, M.S., Kornell, N., Bjork, E.L., Bjork, R.A., 2013. Why interleaving enhances inductive learning: the roles of discrimination and retrieval. *Mem. Cognit.* 41, 392–402.
- Bjork, R.A., Dunlosky, J., Kornell, N., 2013. Self-regulated learning: beliefs, techniques, and illusions. *Annu. Rev. Psychol.* 64, 417–444.
- Blasiman, R.N., Dunlosky, J., Rawson, K.A., 2017. To what, how much, and when of study strategies: comparing intended versus actual study behavior. *Memory* 25, 784–792.
- Bloom, K.C., Shuell, T.J., 1981. Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *J. Educ. Res.* 74, 245–248.
- Carillo-de-la-Peña, M.T., Baillès, E., Caseras, X., Martínez, A., Ortet, G., Pérez, J., 2009. Formative assessment and academic achievement in pre-graduate students of health sciences. *Adv. Health Sci. Educ.* 14, 61–67.
- Carpenter, S.K., 2009. Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J. Exp. Psychol. Learn. Mem. Cogni.* 35, 1563–1569.
- Carpenter, S.K., 2011. Semantic information activated during retrieval contributes to later retention: support for the mediator effectiveness hypothesis of the testing effect. *J. Exp. Psychol. Learn. Mem. Cogni.* 37, 1547–1552.
- Carpenter, S.K., 2017. Spacing effects on learning and memory. In: Byrne, J.H. (Ed.), *Cognitive Psychology of Memory, Learning and Memory: A Comprehensive Reference*, second ed., vol. 2. Academic Press, Oxford, pp. 465–485.
- Carpenter, S.K., Olson, K.M., 2012. Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *J. Exp. Psychol. Learn. Mem. Cogni.* 38, 92–101.
- Carpenter, S.K., Yeung, K.L., 2017. The role of mediator strength in learning from retrieval. *J. Mem. Lang.* 92, 128–141.
- Carpenter, S.K., 2021. Distributed practice/spacing effect. In: Zhang, L. (Ed.), *Oxford Research Encyclopedia of Education*. Oxford University Press.
- Carpenter, S.K., Lund, T.J.S., Coffman, C.R., Armstrong, P.I., Lamm, M.H., Reason, R.D., 2016. A classroom study on the relationship between student achievement and retrieval-enhanced learning. *Educ. Psychol. Rev.* 28, 353–375.
- Carpenter, S.K., Pashler, H., Cepeda, N.J., 2009. Using tests to enhance 8th grade students' retention of U. S. history facts. *Appl. Cognit. Psychol.* 23, 760–771.
- Carpenter, S.K., Pashler, H., Wixted, J.T., Vul, E., 2008. The effects of tests on learning and forgetting. *Mem. Cognit.* 36, 438–448.
- Carpenter, S.K., Rahman, S., Lund, T.J.S., Armstrong, P.I., Lamm, M.H., Reason, R.D., Coffman, C.R., 2017. Students' use of optional online reviews and its relationship to summative assessment outcomes in introductory biology. *CBE-Life Sci. Educ.* 16, ar23.
- Carpenter, S.K., Witherby, A.E., Tauber, S.K., 2020. On students' (mis)judgments of learning and teaching effectiveness: where we stand and how to move forward. *J. Appl. Res. Mem. Cogni.* 9, 181–185.
- Cepeda, N.J., Vul, E., Rohrer, D., Wixted, J.T., Pashler, H., 2008. Spacing effects in learning: a temporal ridgeline of optimal retention. *Psychol. Sci.* 19, 1095–1102.
- Cleary, A.M., McNeely-White, K.L., Hausman, H., Dawson, J., Kuhn, S., Osborn, R.M., Huebert, A.M., Rhodes, M.G., 2021. Wearable technology for automatizing science-based study strategies: reinforcing learning through intermittent smartwatch prompting. *J. Appl. Res. Mem. Cogni.* 10, 444–457.
- Corral, D., Carpenter, S.K., Perkins, K., Gentile, D.A., 2020. Assessing students' use of optional online lecture reviews. *Appl. Cognit. Psychol.* 34, 318–329.
- Dobson, J.L., Linderholm, T., Stroud, L., 2019. Retrieval practice and judgments of learning enhance transfer of physiology information. *Adv. Health Sci. Educ.* 24, 525–537.
- Dobson, J.L., Perez, J., Linderholm, T., 2017. Distributed retrieval practice promotes superior recall of anatomy information. *Anat. Sci. Educ.* 10, 339–347.
- Dunlosky, J., Rawson, K.A., 2015. Practice tests, spaced practice, and successive relearning: tips for classroom use and for guiding students' learning. *Scholar. Teach. Learn. Psychol.* 1, 72–78.
- Dunning, D., Heath, C., Suls, J.M., 2004. Flawed self-assessment: implications for health, education, and the workplace. *Psychol. Sci. Publ. Interest* 5, 69–106.
- Ebbinghaus, H., 1885/1913. *Memory*. H. A. Ruger & C. E. Bussenius, Transl. Teachers College, Columbia University, New York.
- Ebersbach, M., Feierabend, M., Nazari, K.B.B., 2020. Comparing the effects of generating questions, testing, and restudying on students' long-term recall in university learning. *Appl. Cognit. Psychol.* 34, 724–736.
- Emeny, W.G., Hartwig, M.K., Rohrer, D., 2021. Spaced mathematics practice improves test scores and reduces overconfidence. *Appl. Cognit. Psychol.* 35, 1082–1089.
- Finn, B., Metcalfe, J., 2007. The role of memory for past test in the underconfidence with practice effect. *J. Exp. Psychol. Learn. Mem. Cognit.* 33, 238–244.
- Foot-Seymour, V., Foot, J., Wiseheart, M., 2019. Judging credibility: can spaced lessons help students think more critically online? *Appl. Cognit. Psychol.* 33, 1032–1043.
- Foster, N.L., Rawson, K.A., Dunlosky, J., 2018. Self-regulated learning of principle-based concepts: do students prefer worked examples, faded examples, or problem solving? *Learn. InStruct.* 55, 124–138.
- Foster, N.L., Was, C.A., Dunlosky, J., Isaacson, R.M., 2017. Even after thirteen class exams, students are still overconfident: the role of memory for past exam performance in student predictions. *Metacogni. Learn.* 12, 1–19.
- Geller, J., Toftness, A.R., Armstrong, P.I., Carpenter, S.K., Manz, C.L., Coffman, C.R., Lamm, M.H., 2018. Study strategies and beliefs about learning as a function of academic achievement and achievement goals. *Memory* 26, 683–690.
- Goossens, N.A.M.C., Camp, G., Verkoijen, P.P.J.L., Tabbers, H.K., 2014. The effect of retrieval practice in primary school vocabulary learning. *Appl. Cognit. Psychol.* 28, 135–142.
- Hacker, D.J., Bol, L., Horgan, D.D., Rakow, E.A., 2000. Test prediction and performance in a classroom context. *J. Educ. Psychol.* 92, 160–170.
- Hanham, J., Leahy, W., Sweller, J., 2017. Cognitive load theory, element interactivity, and the testing and reverse testing effects. *Appl. Cognit. Psychol.* 31, 265–280.
- Hartwig, M.K., Dunlosky, J., 2012. Study strategies of college students: are self-testing and scheduling related to achievement? *Psychon. Bull. Rev.* 19, 126–134.
- Hartwig, M.K., Dunlosky, J., 2017. Category learning judgments in the classroom: can students judge how well they know course topics? *Contemp. Educ. Psychol.* 49, 80–90.
- Hintzman, D.L., Block, R.A., Summers, J.J., 1973. Modality tags and memory for repetitions: locus of the spacing effect. *J. Verb. Behav.* 12, 229–238.
- Hopkins, R.F., Lyle, K.B., Hieb, J.L., Ralston, P.A.S., 2016. Spaced retrieval practice increases college students' short- and long-term retention of mathematics knowledge. *Educ. Psychol. Rev.* 28, 853–873.
- Hui, L., de Bruin, A.B.H., Donkers, J., van Merriënboer, J.J.G., 2021. Does individual performance feedback increase the use of retrieval practice? *Educ. Psychol. Rev.* 33, 1835–1857.
- Jaeger, A., Eisenkraemer, R.E., Stein, L.M., 2015. Test-enhanced learning in third-grade children. *Educ. Psychol.* 35, 513–521.

- Johnson, G., 2006. Optional online quizzes: college student use and relationship to achievement. *Can. J. Learn. Technol.* 32.
- Jones, A.C., Wardlow, L., Pan, S.C., Zepeda, C., Heyman, G.D., Dunlosky, J., Rickard, T.C., 2016. Beyond the rainbow: retrieval practice leads to better spelling than does rainbow writing. *Educ. Psychol. Rev.* 28, 385–400.
- Kapler, I.V., Weston, T., Wiseheart, M., 2015. Spacing in a simulated undergraduate classroom: long-term benefits for factual and higher-level learning. *Learn. Instr.* 36, 38–45.
- Karpicke, J.D., Lehman, M., Aue, W.R., 2014. Retrieval-based learning: an episodic context account. *Psychol. Learn. Motiv.* 61, 237–284.
- Kirk-Johnson, A., Galla, B.M., Fraundorf, S.H., 2019. Perceiving effort as poor learning: the misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognit. Psychol.* 115, 1–31.
- Kornell, N., Bjork, R.A., 2008. Learning concepts and categories: is spacing the “enemy of induction?”. *Psychol. Sci.* 19, 585–592.
- Larsen, D.P., Butler, A.C., Roediger III, H.L., 2009. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med. Educ.* 43, 1174–1181.
- Lawrence, N.K., 2013. Cumulative exams in the introductory psychology course. *Teach. Psychol.* 40, 15–19.
- Little, J.L., McDaniel, M.A., 2014. Metamemory monitoring and control following retrieval practice for text. *Mem. Cognit.* 43, 85–98.
- Lui, A.K.-F., Poon, M.H.M., Wong, R.M.H., 2019. Automated generators of examples and problems for studying computer algorithms. *Interact. Technol. Smart Educ.* 16, 204–218.
- Lyle, K.B., Bego, C.R., Hopkins, R.F., Hieb, J.L., Ralston, P.A.S., 2020. How the amount and spacing of retrieval practice affect the short- and long-term retention of mathematics knowledge. *Educ. Psychol. Rev.* 32, 277–295.
- Mayer, R.E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bulger, M., Campbell, J., Knight, A., Zhang, H., 2009. Clickers in college classrooms: fostering learning with questioning methods in large lecture classes. *Contemp. Educ. Psychol.* 34, 51–57.
- McDermott, K.B., Agarwal, P.K., D’Antonio, L., Roediger III, H.L., McDaniel, M.A., 2014. Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *J. Exp. Psychol. Appl.* 20, 3–21.
- Miles, S.W., 2014. Spaced vs. massed distribution instruction for L2 grammar learning. *System* 42, 412–428.
- Miller, T.M., Geraci, L., 2011. Unskilled but aware: reinterpreting overconfidence in low-performing students. *J. Exp. Psychol. Learn. Mem. Cognit.* 37, 502–506.
- Moulton, C.-A.E., Dubrowski, A., MacRae, H., Graham, B., Grober, E., Reznick, R., 2006. Teaching surgical skills: what kind of practice makes perfect? *Ann. Surg.* 244, 400–409.
- Nazari, K.B., Ebersbach, M., 2019. Distributing mathematical practice of third and seventh graders: applicability of the spacing effect in the classroom. *Appl. Cognit. Psychol.* 33, 288–298.
- Olson, B.L., McDonald, J.L., 2004. Influence of online formative assessment upon student learning in biomedical science courses. *J. Dent. Educ.* 68, 656–659.
- Pyc, M.A., Rawson, K.A., 2010. Why testing improves memory: mediator effectiveness hypothesis. *Science* 330, 335.
- Raman, M., McLaughlin, K., Violato, C., Rostom, A., Allard, J.P., Coderre, S., 2010. Teaching in small portions dispersed over time enhances long-term knowledge retention. *Med. Teach.* 32, 250–255.
- Randler, C., Kranich, K., Eisele, M., 2008. Block schedule versus traditional biology teaching—an educational experiment using the water lily. *Instr. Sci.* 36, 17–25.
- Rawson, K.A., Dunlosky, J., Sciarrelli, S.M., 2013. The power of successive relearning: improving performance on course exams and long-term retention. *Educ. Psychol. Rev.* 25, 523–548.
- Rickard, T.C., Pan, S.C., 2018. A dual memory theory of the testing effect. *Psychon. Bull. Rev.* 25, 847–869.
- Roediger III, H.L., Agarwal, P.K., McDaniel, M.A., McDermott, K.B., 2011. Test-enhanced learning in the classroom: long-term improvements from quizzing. *J. Exp. Psychol. Appl.* 17, 382–395.
- Rohrer, D., Dedrick, R.F., Hartwig, M.K., 2021. The scarcity of interleaved practice in mathematics textbooks. *Educ. Psychol. Rev.* 32, 873–883.
- Rohrer, D., Dedrick, R.F., Stershic, S., 2015. Interleaved practice improves mathematics learning. *J. Educ. Psychol.* 107, 900–908.
- Rowley, T., McCrudden, M.T., 2020. Retrieval practice and retention of course content in a middle school science classroom. *Appl. Cognit. Psychol.* 34, 1510–1515.
- Schutte, G.M., Duhon, G.J., Solomon, B.G., Poncy, B.C., Moore, K., Story, B., 2015. A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *J. Sch. Psychol.* 53, 149–159.
- Seabrook, R., Brown, G.D.A., Solity, J.E., 2005. Distributed and massed practice: from laboratory to classroom. *Appl. Cognit. Psychol.* 19, 107–122.
- Shapiro, A.M., Gordon, L.T., 2012. A controlled study of clicker-assisted memory enhancements in college classrooms. *Appl. Cognit. Psychol.* 26, 635–643.
- Sobel, H.S., Cepeda, N.J., Kapler, I.V., 2011. Spacing effects in real-world classroom vocabulary learning. *Appl. Cognit. Psychol.* 25, 763–767.
- Sones, A.M., Stroud, J.B., 1940. Review, with special reference to temporal position. *J. Educ. Psychol.* 31, 665–676.
- Sotola, L.K., Credé, M., 2021. Regarding class quizzes: a meta-analytic synthesis of studies on the relationship between frequent low-stakes testing and class performance. *Educ. Psychol. Rev.* 33, 407–426.
- Spruit, E.N., Band, G.P.H., Hamming, J.F., 2015. Increasing efficiency of surgical training: effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surg. Endosc.* 29, 2235–2243.
- Tullis, J.G., Finley, J.R., Benjamin, A.S., 2013. Metacognition of the testing effect: guiding learners to predict the benefits of retrieval. *Mem. Cognit.* 41, 429–442.
- van Dongen, K.W., Mitra, P.J., Schijven, M.P., Broeders, I.A.M.J., 2011. Distributed versus massed training: efficiency of training psychomotor skills. *Surg. Tech. Dev.* 1 (e17), 40–42.
- van Gog, T., Kester, L., Paas, F., 2011. Effects of worked examples, example-problem, and problem-example pairs on novices’ learning. *Contemp. Educ. Psychol.* 36, 212–218.
- Velan, G.M., Jones, P., McNeil, H.P., Kumar, R.K., 2008. Integrated online formative assessments in the biomedical sciences for medical students: benefits for learning. *BMC Med. Educ.* 8, 52.
- Vlach, H.A., Sandhofer, C.M., 2012. Distributing learning over time: the spacing effect in children’s acquisition and generalization of science concepts. *Child Dev.* 83, 1137–1144.
- Wahlheim, C.N., Maddox, G.B., Jacoby, L.L., 2014. The role of reminding in the effects of spaced repetitions on cued recall: sufficient but not necessary. *J. Exp. Psychol. Learn. Mem. Cognit.* 40, 94–105.
- Wing, E.A., Marsh, E.J., Cabeza, R., 2013. Neural correlates of retrieval-based memory enhancement: an fMRI study of the testing effect. *Neuropsychologia* 51, 2360–2370.
- Yazdani, M.A., Zebrowski Jr., E., 2006. Spaced reinforcement: an effective approach to enhance the achievement in plane geometry. *J. Math. Sci. Math. Educ.* 1, 37–43.
- Yeo, D.J., Fazio, L.K., 2019. The optimal learning strategy depends on learning goals and processes: retrieval practice versus worked examples. *J. Educ. Psychol.* 111, 73–90.
- Yeung, K.L., Carpenter, S.K., Corral, D., 2021. A comprehensive review of educational technology on objective learning outcomes in academic contexts. *Educ. Psychol. Rev.* 33, 1583–1630.
- Zulkiply, N., McLean, J., Burt, J.S., Bath, D., 2012. Spacing and induction: application to exemplars presented as auditory and visual text. *Learn. Instr.* 22, 215–221.

Service-learning and experiential education

Christian Winterbottom^a, Vickie E. Lake^b, and Ithel Jones^c, ^a University of North Florida, Jacksonville, FL, United States;

^b University of Oklahoma, Norman, OK, United States; and ^c Florida State University, Tallahassee, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Settlement houses	673
Black social settlement movement	674
Black educators	674
W.E.B. DuBois	674
George Edmond Haynes	675
Institutional service-learning	675
Conceptualizing service-learning	676
Theorizing service-learning	676
Reconceptualizing service-learning for the 21st century	677
A critical service-learning	678
Moving forward	678
Conclusion and implications	679
References	679

The purpose of service-learning has always been to impact the thinking, attitudes, and development of students, or as [Stoecker \(2016, p. 5\)](#) states, to “liberate students.” Service-learning has also been accepted and promoted as a powerful teaching and learning approach. Researchers such as [Wade \(2000\)](#), [Billig \(2000\)](#), and [Lake and Jones \(2008, 2012\)](#) have highlighted the potential contribution of service-learning in enhancing the social studies curriculum and developing civic responsibility in students. Service-learning achieves these outcomes by actively involving students in the process of learning; service-learning is experiential learning. As such, many service-learning practitioners and researchers have attributed the beginning of service-learning to the work of Dewey; however, doing so privileges higher education–centric models and federal service-learning programs ([Stoecker, 2016](#)). Additionally, starting with Dewey places emphasis on the learning half of the phrase and Stoecker claims that the real emphasis should be placed on the service half. Also focusing on the service half, in [2010](#), Jacoby and Mutascio stated that higher education was being asked to renew its historic commitment to service-learning by accepting a leadership role in addressing society’s increasing problems and meeting growing human needs. However, if one actually examines higher education’s historic commitment to service-learning, it takes the analyst to the period of time before the establishment of settlement houses in the 1880s ([Stoecker, 2016](#)) and the “unrecognized roots of service-learning embedded in African American social thought and action” ([Stevens, 2003, p. 25](#)).

Beginning with settlement houses, this article provides a historical perspective of service-learning rooted in the African American community. The authors then examine the theoretical underpinnings of service-learning and how it has become a pedagogy used throughout many institutions in the United States. To conclude the article, the authors draw on the work of critical service-learning scholars as they look at the next steps in service-learning and experiential education.

Settlement houses

Toynbee Hall opened in East London on Christmas Eve in 1884 and is considered the first settlement house. However, before it opened, a social movement had been taking place across Europe that “knowledge and understanding are the springs of social action and that the surest way to increase knowledge and promote understanding is to practice the art of living together as neighbors and friends” ([Ellis, 1948, p. 170](#)). Paving the way were Christian Socialists like Edward Denison, Arnold Toynbee, and Samuel Barnett, but this social movement lacked organization and the documentation of those in the movement was particularly scarce. Therefore, much is attributed to Samuel and Henrietta Barnett who opened Toynbee Hall in honor of their friend and fellow social reformer, Arnold Toynbee, who died the previous year at 30 ([Ellis, 1948](#)).

Toynbee was an economic historian at Balliol College in Oxford whose life interest was helping Britain’s growing number of working-class poor by developing trade unions, establishing housing and libraries, and offering them free classes. He also encouraged his students and other social reformers to teach free classes in their own communities to working-class audiences. Barnett’s vision of Toynbee Hall was to have university students (men) from Oxford and Cambridge develop solutions to poverty by witnessing it firsthand by living in the hall, volunteering in the community, and sharing their knowledge ([Ellis, 1948](#)). Toynbee Hall is considered the first university affiliated settlement house.

After visiting Toynbee Hall, Jane Addams and Ellen Gates Starr rented an abandoned mansion in Chicago built by Charles Hull and opened Hull House, the first social settlement in North America ([Hull House Museum, n.d.](#)). Hull House offered childcare for working mothers, kindergarten, help with employment and interviewing, English and citizenship classes, as well as classes in

theater, music, and art. Although Hull House was not affiliated with a university, it helped launch the College Settlements Association in 1894 that linked university affiliated settlement houses across major cities in the US (Stoecker, 2016). The mission of these settlement houses varied from charity, to building the power of the dispossessed, to defusing class struggles while maintaining class inequality (Mieras, 2008).

According to Stevens (2003), at the same time that settlement houses were expanding in the US, “a practical application of experiential learning and education devoted to race and community development, and ideas later associated with service-learning, might be observed in most “Negro” institutions like the church, black colleges, and other civic-minded organizations” (p. 26). Focusing on the years between 1890 and 1930, Stevens explored how the activity of African American women, the agenda of the black women’s social club movement, and African American educators should be seen as philosophical and pedagogical precursors to service-learning.

Black social settlement movement

At the forefront of community service, social welfare systems, and educational programs intended for racial uplift and enrichment, one would find African American women and educators participating in service to their communities (Stevens, 2003). The value of service began in the family, was reinforced in the church, and implemented either formally or informally at most black colleges. Like their white counterparts, black women were social reformers and social change agents who believed in educating, instructing, and caring for the young and old in their communities. Anna Julia Cooper, an educator and social reformer, helped organize the black social settlement movement by forming relationships between black college men and women and their communities to problem solve, address public issues, and provide education on citizenship and democracy. However, the social settlement movement was only successful when the college students learned to speak the language of the immigrant groups and the working poor and demonstrated respect for their social lives.

In the late 19th and early 20th century, Ida B. Wells-Barnett was a prominent journalist, activist, and researcher who was instrumental in the formation of the NAACP (Norwood, 2017; Steptoe, n.d.). After three of her friends and fellow businessmen were lynched in 1908, Wells-Barnett became an anti-lynching advocate and created the Negro Fellowship League, which became a social settlement agency in 1910 focused on criminal justice issues. The League offered educational programs for black migrants and ex-prisoners from Joliet State Prison who had no help reentering society and who were often mistreated by a biased system (Stevens, 2003). The League was a way for the local black community to become involved in education and social action surrounding issues that impacted their daily lives.

During the black social settlement movement, organized social clubs for black women were also forming (Stevens, 2003). In 1896, at a conference held for black women’s social clubs, The National Association of Colored Women’s Clubs was created. Josephine Ruffin’s statement became the mission of this newly formed organization.

The reasons why we should confer are so apparent ... We need to talk over not only those things which are of vital importance to us as women, but also the things that are of special interest to us as colored women, the training of our children, openings for our boys and girls, how they can be prepared for occupations and occupations may be found or opened to them, what we especially can do in the moral education of the race with which we are identified, our mental elevations and physical development, the home training it is necessary to prepare them to meet [the] peculiar [special; difficult] conditions in which they find themselves, how to make the most of our own ... opportunities, these are some of our questions to be discussed.

National Association of Colored Women’s Clubs (n.d., p. 5).

These clubs combined social outreach with education to address issues in their communities. Clubs such as the Locust Street Social Settlement in Hampton, Virginia focused on social services and education for wayward girls. The Working Girls Home Association in Cleveland provided education to black girls and women on how to become proactive in their community and civic life, as well as giving girls a place to live when they arrived in the city. Cooper, Wells-Barnett, and members of the black social settlement movement and black women’s social clubs were able to identify social issues of their day, mobilize community members, and connect learning and action to better their community, thereby implementing the philosophy of service-learning, although not as a formal educational approach.

Black educators

W.E.B. DuBois

Male black educators were as involved as the black women in connecting social action and education to combat impediments of racial progress (Stevens, 2003). A student at Fisk University, DuBois was exposed to experiential education and the ideals of service that had a transformative impact on his life. After graduation he lived and taught in rural Tennessee where he learned to speak the language of the people he wanted to teach and the value of non-traditional education. He used the skills and lessons learned while teaching in Tennessee and applied them to the study of the Philadelphia Negro. Following the Hull House, or settlement house model, DuBois lived in and studied Philadelphia’s Seventh Ward for two years. Through this research he once again experienced

the power of experiential learning, but he also realized that change must take place *in* and *with* the community. While not termed service-learning at the time, DuBois' work shares the principles of service-learning pedagogy.

George Edmond Haynes

George Edmond Haynes, a professor at Fisk University, created an experiential education course that is considered a precursor to the modern service-learning courses. Speaking at the *National Conference of Charities and Corrections* in 1911, Haynes outlined how "the problem of social uplift is so great that, in addition to expert social workers, all Negro ministers, doctors, lawyers, teachers and others should have the benefit of instruction in scientific methods and the new social point of view" (p. 385). He outlined a three-point plan implemented at Fisk, as well as other Negro and White Universities.

First, all Negro students were to be instructed in the issues, conditions, and needs of the people in their communities. Students were required to take courses in the following areas: industrial history, economic theory, sociology, economic and labor problems, social work methods, and History of the Negro in America and the Negro Problem. They also attended expert lectures on the "religious problems among Negroes in cities; delinquents and probation problems; special problems of Negro women in cities; of Negro children; and on Principles of Relief" (Haynes, 1911, p. 385). In their final semester, students spent 10 h per week working with their community on the conditions facing Negro people in Nashville. Building knowledge and training of college students was a huge priority for Haynes. He stated that "if there is any one fact well known, it is that the Negro's situation cannot be helped by inefficient and inexperienced enthusiasts" (p. 387).

Second, promising graduates of Fisk University and other colleges were to be provided further study in New York by the Committee on Urban Conditions Among Negroes (Haynes, 1911). The purpose of this committee was to foster cooperation among existing agencies working among Negroes in urban settings and to discover where additional work was needed, to provide resources to the identified areas in order to remedy the situation, and to further train Negro social workers.

Third, agencies in New York and other major urban cities were to be standardized (would all be similar in the way that the agencies would be established), coordinated, and run by trained workers (Haynes, 1911). Haynes discussed how agencies would work together and share information to improve Negro communities, specifically focusing on housing and neighborhoods, creating employment agencies, thrift agencies and cooperative business enterprises, recreation facilities, religious institutions, and exchanging training methods and practices that work.

Barnett, Addams and Starr, Cooper, Wells-Barnett, DuBois, and Haynes should be considered pioneers in service-learning. They sought solutions to issues, mostly racial issues, in their communities and connected education and service to address them. Steven's (2003) stated that "there are relevant lessons that seem to connect service-learning with the action-oriented expression of African American social thought around the turn of the 20th century" (p. 32). The African American social movement and service-learning share many of the same themes: creating knowledge of institutions and issues, investment in the community, and an understanding and appreciation of the circumstances faced by community members. Stevens argued that, had these scholars and their work been included in mainstream higher education, the knowledge base surrounding service-learning would be more expansive and advanced. Instead, many service-learning scholars and practitioners look to John Dewey for the inspiration and philosophical foundation of service-learning (Stoecker, 2016).

Institutional service-learning

In reference to Dewey, Stoecker (2016) noted "It is odd that institutionalized service learning would claim as its hero someone who did not practice service learning and focused his educational theories on children" (p. 13). In his critique of institutionalized service-learning, Stoecker posited that using Dewey as its foundation meant that higher education never had to change, nor was it responsible for taking action to address social issues. Stoecker argued that traditional service learning benefits students and also reinforces power structures that result in detriment to community partners with limited power and resources. Dewey did not practice experiential learning or engage his college students in service even though he collaborated with Jane Addams by lecturing at Hull House.

Most of service-learning history leaves out the Highlander Folk School and its important role in the civil rights movement (Stoecker, 2016). Already discussed is the absence of African American scholars and the black social movement, but also absent is the work by Native American students who occupied Alcatraz Island in 1969. The 60s and 70s witnessed campus service and activism at its peak. Professors and students focused on issues of poverty, racism, sexism, and other -isms and exclusions in society. "And all of it is completely missing from the official histories of institutionalized service learning" (Stoecker, 2016, p. 15). The *National Service Learning Clearinghouse* (2013) historical timeline, does not mention any of these events.

Scholarship Reconsidered, a 1990 publication by Ernest Boyer set the course for institutionalized service-learning and was "one of the first signs of the institutional colonization of service learning's potential" (Stoecker, 2016, p. 18). The National and Community Service Trust Act was signed into law by President Bush in 1990 and defined institutionalized service-learning.

SERVICE-LEARNING—The term "service-learning" means a method—(A) under which students or participants learn and develop through active participation in thoughtfully organized service that (i) is conducted in and meets the needs of a community (ii) is coordinated with an elementary school, secondary school, institution of higher education, or community service program, and with the community; and (iii) helps foster civic

responsibility; and (B) that (i) is integrated into and enhances the academic curriculum of the students, or the educational components of the community service program in which the participants are enrolled; and (ii) provides structured time for the students or participants to reflect on the service experience.

Stoecker argued that *change* and *action*, which were part of the original principles of community based organizations, were changed to *education* and *service* and that there was no empowerment philosophy that undergirded its foundation. In essence, higher education institutions control how service is defined and how students engage in service. Since 1990, critics of institutionalized service-learning have questioned if there is any learning or service happening and if it is happening, is it any good (Eyler and Giles, 1999). If higher education returned to its historical commitment to service-learning, as Jacoby and Mutascio claimed in 2010, it would mean addressing social issues that are prominent today: Black Lives Matter and racial injustice, voting rights, climate change, healthcare, refugee crisis, income gap, gun violence, hunger and food insecurity, and equality (9 Biggest Social Justice Issues of 2020, 2020). As much as we would like to think that institutions of higher education would embrace engaging in politicized civic engagement, it is well documented that they will not approve such activities for their professors or students (Bickford and Reynolds, 2002; Bringle et al., 2006; Kajner et al., 2013).

Conceptualizing service-learning

Service-learning has a long history in education, particularly at the high-school and higher education levels and, more recently in teacher education programs. It is an approach that emphasizes the importance of grounding learning in experience within schools. Proponents of service learning maintain that children's learning emerges from curricula of authentic and meaningful experiences, purposeful activities, and opportunities for social interaction. More recently, there is increasing evidence that integrating service into the school curriculum contributes to students' motivation and academic achievement (e.g., Dewsbury-White, 1993; Hamilton and Zeldin, 1987; Scales and Roehlkepartain, 2004; Scales et al., 2006). Consequently, it is recognized as an effective pedagogical approach that combines authentic community service with academic learning experiences (Billig, 2004). While service-learning is considered an effective approach across all grade levels, we believe that as a pedagogy, it can even be used in the early grades.

Service-learning is considered a "natural fit" for the early grades (i.e., PK-3rd grade) because of its emphasis on first-hand experiences and active engagement. This is because early childhood educators have long recognized the importance of hands-on learning for young children. Indeed, from the early beginnings of the kindergarten movement, components of a service-learning pedagogy have been the staples of most approaches to early childhood education. Froebel's (1899) "gifts and occupations", together with "gardening and nature study", were used to develop the child according to his nature. Montessori (1966) then recognized that children learn important concepts through their own activity, and used practical life exercise or real world experiences to teach her students. Similarly, Kilpatrick, Dewey, and others were proponents of experiential learning through the use of projects and community-oriented themes. Thus, Dewey's (1918) proposition that kindergarten should be "*more natural, more direct, more real representations of current life*" (p. 117) lends support to service-learning pedagogy with young children.

Typically, service-learning is defined as the integration of community service projects with academic skills and content. It is, however, a broad approach that is applied in different ways across different contexts. As such, service-learning represents a wide array of activities such as projects that entail raising money for various worthy causes (e.g., fundraisers, canned food drives) or advocacy activities such as writing letters to public officials about a community issue or concern. Other types of service-learning include direct service whereby students engage with the recipients of the service activity, and indirect service in which the service is offered to the community as a whole but without any direct interaction with the recipients.

Early childhood teachers integrate such service activities into the curriculum not only in areas such as social studies, but also reading, writing, science, and mathematics. It has been recommended (Lake and Jones, 2012) that in the early grades preference should be given to direct service because it allows children to experience the result of their service efforts. Underlying this recommendation are theories of children's early learning, as well as the theoretical bases of a pedagogy of service-learning.

Theorizing service-learning

The theoretical underpinnings of service-learning was first considered almost three decades ago (e.g., Giles and Eyler, 1994; Saltmarsh, 1996). Initially, attempts to develop a theoretical rationale for a service-learning approach drew on Dewey's (1916) pragmatic philosophy. Dewey's philosophical pragmatism challenged the dualistic relationship between knowledge and experience and argued that reality can and should be changed through reason and action. For Dewey, therefore, knowledge and learning should be connected to experience. Proponents of service-learning claimed that service-learning is a form of experiential learning that relies on reflection. Then, in efforts to legitimize service-learning, they reasoned that the paradigm of pragmatism can be employed as a philosophical underpinning for the approach (Maddux and Donnett, 2015). According to Dewey (1938), "*all genuine education comes about through experience*" (p. 35). In other words, education is a social process whereby learning and knowledge emerge from positive experiences.

Service-learning practitioners and researchers (e.g., [Giles and Eyler, 1994](#); [Maddux and Donnett, 2015](#)) maintain that two of Dewey's primary works, *How We Think* (1933) and *Experience and Education* (1938) offer a plausible theoretical basis for the approach. This is because they address the three fundamental service-learning elements of *experience*, *inquiry*, and *reflection*. For Dewey, learning takes place in the context of experience and interaction, whereby learning is derived through a process of reflective thinking.

Beyond Dewey, service-learning advocates also draw on David Kolb's (1994) work on experiential learning theory. [Kolb \(1984\)](#) posited service-learning in terms of a four-stage cyclical process involving concrete experiences, reflection, abstract conceptualization, and active experimentation. Thus, when they are engaged in service-learning, students first participate in a service experience. Then, through reflection they integrate abstract concepts learned in the classroom and consequently, actively understand their experience. For Kolb (2015) and proponents of experiential education, experience is a fundamental and critical part of the learning process. It is hardly surprising, therefore, that this model is often cited as a plausible theoretical rationale for service-learning as pedagogy. Yet, Kolb's theory has been criticized by educators and researchers (e.g., [Miettinen, 2000](#); [Seaman et al., 2017](#)). The main criticism of Kolb's experiential learning theory is that it lacks empirical support ([Coffield et al., 2004](#)). Furthermore, it has been claimed that it is too eclectic ([Miettinen, 2000](#)), that dimensions of the model do not connect to each other cohesively ([Jarvis, 2012](#)), and that it can be somewhat of a barrier to successful facilitation of experiential learning and subsequent understanding ([Seaman et al., 2017](#)). Thus, a revision of Kolb's model has been proposed ([Morris, 2019](#)), as well as other alternative models (e.g., [Bergsteiner and Avery, 2014](#); [Miettinen, 2000](#)). To date, however, Kolb's model remains the most influential as it relates to experiential learning and service-learning approach ([Seaman et al., 2017](#)).

The work of Paulo [Freire \(1970\)](#) has also been invoked by proponents of service-learning as a convincing theoretical rationale. According to the Brazilian educator's critical pedagogy, education should not only nurture the intellect but also promote emancipatory change. Freire's defining work on critical pedagogy in his *Pedagogy of the Oppressed* (1970) has had a significant impact on education and pedagogy worldwide. According to Freire, teaching cannot be separated from critical analysis of how society works. Thus, teaching should challenge learners to critically engage with the world, which in turn can lead to social change. Like Freire, proponents of service-learning believe that education should be rooted in personal experience leading to a process that enhances community, builds social capital, and action for social change.

Several aspects of [Freire's \(1970\)](#) work contribute to a plausible theoretical rationale for service-learning pedagogy. First, praxis, or action that is informed, was a central concern for Freire. Second, Freire emphasized the importance of situating educational activity in the lived experience of participants. Finally, dialog was significant for Freire because it involves individuals working together and, in turn, individuals reflecting and acting on reality. For service-learning proponents, education should engage participants in problem solving, reflection, social action, and transformation. Processes that, like Freire, view education as a constructive, experiential process.

The perspectives of Dewey, Kolb, and Freire offer theoretical support for service-learning. For some service-learning advocates (e.g., [Erikson and Santmire, 2001](#)), however, psychological theory offers a more plausible theoretical basis. For [Erikson and Santmire \(2001\)](#), Contact Theory (CT) offers a useful theoretical rationale for service-learning. First developed by Gordon [Allport \(1954\)](#), CT suggests that contact between two groups of individuals can promote tolerance and acceptance ([Pettigrew and Tropp, 2006](#)), at least under certain conditions. For example, friendship between members of different groups is a form of contact that can minimize inter-group conflict ([Dovidio et al., 2003](#)), which in turn yields individual benefits in the form of "self-expansion." Viewed this way, CT can be interpreted as a motivational process that encourages individuals to grow and integrate new elements into their lives ([Aron et al., 1998](#)). As such, CT seems to primarily rely on emotional processes as opposed cognitive.

Classic cognitive and sociocultural theories of [Bruner \(1968\)](#), [Gardner \(1987\)](#), [Vygotsky \(1978\)](#), and others also lend support to a service-learning approach. Vygotsky, in particular appeals to proponents of service-learning in early childhood contexts ([Lake and Jones, 2012](#)), presumably because of its emphasis on situating thinking and speaking within the context of activity ([Wertsch et al., 1995](#)). That is, children's cognitive development is embedded in experience (i.e., activity), social relationships, and social interaction ([Lave and Wenger, 1991](#); [Rogoff, 1990](#)). Vygotsky's theory suggests that children can achieve much more when they are engaged in collective activities. Such activities are evident in a service-learning context during which the language that is exchanged within peer and adult-child interaction facilitates both the joint activity and subsequent internalization. Through this co-constructive process, the teacher and child interact to create the zone of proximal development (ZPD). In doing so, they each bring different levels of understanding to any given task and, through a process of interaction, compromise, and shared experiences, reach a new level of understanding ([Goldstein, 1999](#)).

Reconceptualizing service-learning for the 21st century

As the previous sections have suggested, service-learning has historical connections to community and education that began many hundreds of years ago. Starting in the settlement homes, and then through community-school collaborations espoused by Dewey, service-learning has significantly impacted education. This section will examine how we perceive the current adaptations of service-learning and its potential as a pedagogy moving forward out of a pandemic.

A critical service-learning

It has been argued that traditional service-learning has historically placed the emphasis on student learning and outcomes related to student development of new ideas and perspectives in relation to meeting a community need. Service-learning for social justice, however, provided a new means to move beyond just “volunteering” and was “designed to enhance practice through exploration of issues of oppression, individual voice, [and] empowerment” (Maybach, 1996, p. 224).

Scholars such as Artz (2001) have long argued that a serious limitation to traditional service-learning pedagogy is the *service-learning as charity* model often employed in higher education. He posits that this model provides students with opportunities to identify community needs, serves as a way to address the needs, yet often does not provide an outlet for reflection and consideration of the structural forces that perpetuate the injustices. It also fails to provide opportunities to meaningfully reflect *with* the community in which they serve as to why the injustice occurs and how their own (the student’s) individual biases may serve as limitations to social justice.

The foundation for Artz’s (2001) argument was that, “charity, by itself, as a basis for service-learning, will not likely lead to a sustained social critique or action” (p. 240). He noted, “charity frequently denies the possibility of social change by implying that the poor or oppressed are less competent and less able than those who have more social, cultural, and economic capital” (p. 240). Artz also stressed the need to move beyond a service-learning pedagogy that fails to challenge structural forces, and allows for a discourse that promotes mutually inclusive dialog between service-learning students and the *oppressed* in order to promote social action, not charity. Mitchell (2008) similarly argued that the goal of critical service-learning, as opposed to traditional service-learning, is to deconstruct and dismantle power structures that perpetuate the need for service and serve to sustain social inequalities. It is these perspectives that we believe will move service-learning into a new *era*. The Covid-19 pandemic, in particular, has helped highlight the inequalities that currently exist, with homeless shelters filling to capacity and food banks becoming depleted quicker than can be refilled. Service-learning scholars should start to stress the importance of change and how we can work collaboratively with community to enable these changes.

An important point emphasized by Artz (2001), and one that we concur with, is that critical service-learning must be approached with care at the institutional level, as the very nature of critical service-learning hinges on a direct focus on the “relationships between power and culture, communication and ideology” (p. 242), which are often materialized as tensions between the academy and the community. Mitchell (2008) also acknowledged that a critical approach to service-learning may be difficult to implement within institutional and societal structures that are traditionally resistant to social change, yet she predicts that the “promise of the approach” to question the status-quo and activate change will invariably lead to critical service-learning as the “next direction” of the field (p. 62). Mitchell (2008) noted that to successfully engage students in meaningful critical service-learning pedagogy, there must be a redistribution of power among all stakeholders, authentic, reciprocal relationships between students and community, and the precinct of a social change orientation. In order for this shift to occur in tangible ways, recognizing that the academy is not the only holder of knowledge and valuing the expertise and knowledge of community stakeholders and partners as co-educators is necessary.

In critical service-learning pedagogy, it is suggested that the stakeholders are the students, organizations/institutions, and resident community members, including their individual and collective interests and histories. By focusing equitably and inclusively on stakeholders’ agency in the critical service-learning process there is opportunity to evaluate the reciprocity of each stakeholder involved, providing valuable insight into perceptions and meaning that serve to empower. Bringle et al. (2009) identified an Exploitative, Transactional, or Transformational (ETT) Framework of service focused on relationship outcomes ranging from exploitive to transactional to transformational. Much of the experiential learning found within the institution are transactional at best and exploitive at worst.

Moving forward

To develop a community engagement ethic in our students, we need to move beyond these traditional modes of partnership and embrace a more transformational experience for all stakeholders. To do this, it is necessary to include community partners as co-educators, to value community partners as holders of knowledge, and to assess outcomes from an academic *and* community perspective. Furco (2000) indicated that, “Service-learning programs are distinguished from other approaches to experiential education by their intention to equally benefit the provider and the recipient of the service as well as to ensure equal focus on both the service being provided and the learning that is occurring” (p. 12). Equality demands reciprocity and reciprocity requires proximity. As critical service-learning educators, it is our responsibility to promote a climate of *service with* as opposed to *service for*, a proximal relationship between students and the communities with which we work, to dismantle deficit-based structural hierarchies often prevalent in schools and communities, in order to foster inclusivity and equity through an asset-based lens. This asset-based lens becomes more focused when critical reflection is added.

Reflection can be accomplished independently or as a collective endeavor. Yet, however done, reflection demands consideration of one’s internal state (beliefs, feelings, assumptions) and external circumstances (actions, relationships, power dynamics, obstacles). Reflection also “demands a self-honesty and humility that will hold its own against affront from any quarter” (Hall, 2019). As a required component of high-quality critical service-learning, reflection plays a necessary and integral part in helping students connect their own identity to the world around them, think critically about their role in the world, and analyze how their own perceptions and beliefs shape the lens through which they *experience* that world. Kolb’s (1984) model of experiential learning

(see figure below) denotes that reflection is necessary for learners to move from experience to conceptualization and from conceptualization to action.

Although educators recognize the benefits of reflection, it is not often *done well* in the classroom. Too often, students are asked to provide a surface-level summary of their experiences and to make obvious connections to the content, leading to little to no critical reflection, adjustment of perceptions, and/or innovative approaches to problem-solving (Nemeth and Winterbottom, 2015). To address these inconsistencies between what we, as educators, know is best-practice and what we actually see in the classroom, there needs to be a better understanding of what critical reflection is, why it matters to students' learning and community impact, and how to facilitate critical reflection that reaches desired academic, social/emotional, and civic outcomes.

Conclusion and implications

In this article, we have taken a step back and examined the roots of service-learning, the transition via Dewey into the larger community, and the theoretical underpinnings that exist within service-learning in the educational institutions we serve. We believe that as service-learning scholars continue in their work, students are best served by faculty and staff working with them who avoid an already agreed upon definition of social justice in the classroom, which challenges students to consider views of the world that differ from their own. Service-learning seems undeniably well positioned to raise important questions about social conditions, to confront students with diverse ideas about these conditions, and to engage them in considering and weighing evidence about conditions they encounter through their community engaged work. Now more than ever, with the implications of a pandemic that has seen hundreds of thousands lose their lives, while millions more become infected, all partners should consider the service-learning they are embarking on. Critical thinking's emphasis on "examining assumptions, exploring various viewpoints, analyzing positions, engaging in self-reflection, and developing [one's] own perspectives" (Goodman, 2011, p. 40) is ideally suited for students' in the 21st Century, specifically, through an authentic processing of their service-learning experiences in light of existing evidence and competing perspectives.

References

- Allport, G.W., 1954. *The Nature of Prejudice*. Addison-Wesley, Cambridge, MA.
- Aron, A., Norman, C.C., Aron, E.N., 1998. The self-expansion model and motivation. *Represent. Res. Soc. Psychol.* 22, 1–13.
- Artz, L., 2001. Critical ethnography for communication studies: dialog and social justice in service learning. *South. Commun. J.* 66 (3), 239–250.
- Bergsteiner, H., Avery, G.C., 2014. The twin-cycle experiential learning model: reconceptualizing Kolb's theory. *Stud. Cont. Educ.* 36, 257–274.
- Bickford, D., Reynolds, N., 2002. Activism and service-learning: reframing volunteerism as acts of dissent. *Pedagogy* 2 (2), 229–252.
- Billig, S.H., 2000. Research on K-12 school-based service-learning—the evidence builds. *Phi Delta Kappan* 81 (9), 658–664.
- Billig, S.H., 2004. Heads, hearts, hands: the research on K-12 service-learning. In: *Growing to Greatness 2004: The State of Service-Learning Project*. National Youth Leadership Council, St. Paul, MN, pp. 12–25.
- Boyer, E.L., 1990. *Scholarship Reconsidered: Priorities of the Professoriate*. Carnegie Foundation for the Advancement of Teaching, Princeton, N.J.
- Bringle, R.G., Hatcher, J.A., McIntosh, R.E., 2006. Analyzing Morton's typology of service paradigms and integrity. *Mich. J. Community Serv. Learn.* 13 (1), 5–15.
- Bringle, R.G., Clayton, P.H., Price, M.F., 2009. Partnerships in service learning and civic engagement. *Partnerships* 1 (1), 1–20.
- Bruner, J., 1968. *Toward a Theory of Instruction*. W. W. Norton, New York.
- Coffield, F., Moseley, D., Hall, E., Ecclestone, K., 2004. *Learning Styles and Pedagogy in Post-16 Learning. A Systematic and Critical Review*. Learning and Skills Research Centre, London.
- Dewey, J., 1916/1966. *Democracy and Education*. The Free Press, New York.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York.
- Dewsbury-White, K.E., 1993. *The Relationship of Service Learning Project Models to the Subject Matter Achievement of Middle School Students*. Doctoral dissertation. Michigan State University.
- Dovidio, J.F., Gaertner, S.L., Kawakami, K., 2003. Intergroup contact: the past, present, and the future. *Group Relat. Intergr. Relat.* 6, 5–21.
- Ellis, L.F., 1948. Toynbee Hall and the university settlements. *J. Roy. Soc. Arts* 96, 167–178. <https://www.jstor.org/stable/41363547>.
- Erikson, J.A., Santmire, T., 2001. Psychological bases of effective service learning. In: Andersn, J.B., Swick, K.J., Joost, J. (Eds.), *Service Learning in Teacher Education: Enhancing the Growth of New Teachers, Their Students, and Communities*. AACTE, New York, pp. 19–38.
- Eyler, J., Giles Jr., D.E., 1999. *Where's the Learning in Service-Learning*. CA Jossey-Bass, San Francisco.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York.
- Furco, A., 2000. *Service-Learning: A Balanced Approach to Experiential Education*. In: *Introduction to Service-Learning Toolkit*. Campus Compact, Washington, DC.
- Gardner, H., 1987. *The Mind's New Science*. Basic Books, New York.
- Giles Jr., D.E., Eyler, J., 1994. The theoretical roots of service-learning in John Dewey: toward a theory of service-learning. *Mich. J. Community Serv. Learn.* 1 (1), 77–85.
- Goldstein, L., 1999. The relational zone: the role of caring relationships in the co-construction of mind. *Am. Educ. Res. J.* 36, 647–673.
- Goodman, D.J., 2011. *Promoting Diversity and Social Justice: Educating People From Privileged Groups*, second ed. Taylor & Francis, New York, NY.
- Hall, A., 2019. *REAL Reflection Reflective Learning Communities: A Handbook for Facilitating Critical Reflection in the Classroom and Beyond*. Virginia Commonwealth University, Richmond, VA.
- Hamilton, S.F., Zeldin, R.S., 1987. Learning civics in the community. *Curric. Inq.* 17 (4), 407–420.
- Haynes, G.E., 1911. Cooperation with colleges in securing and training Negro social workers for urban conditions. *Proc. Natl. Conf. Charities and Correct.* 38, 384–387.
- Hull House Museum, (n.d.). About Jane Adams and Hull-House Settlement, <https://www.hullhousemuseum.org/about-jane-addams>.
- Jacoby, B., Mutascio, P. (Eds.), 2010. *Looking in, Reaching Out: A Reflective Guide for Community Service-Learning Professionals*. Campus Compact, Boston, MA.
- Jarvis, P., 2012. *Adult Learning in the Social Context*. Routledge, London/New York.
- Kajner, T., Chovanec, D., Underwood, M., Mian, A., 2013. Critical community service learning: combining critical classroom pedagogy with activist community placements. *Mich. J. Community Serv. Learn.* 19 (2), 36–48.

- Kolb, D., 1984. *Experiential Learning: Experience as the Source of Learning and Development*. NJ, Englewood Cliffs.
- Lake, V.E., Jones, I., 2008. Service-learning in early childhood teacher education: using service to put meaning back into learning. *Teach. Teach. Educ.* 24 (8), 2146–2156.
- Lake, V.E., Jones, I., 2012. *Service Learning in the PreK-3 Classroom: The What, Why, and How-To Guide for Every Teacher*. Free Spirit Publishing, MN.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, New York.
- Maddux, H.C., Donnett, D., 2015. John Dewey's pragmatism: implications for reflection in service-learning. *Mich. J. Community Serv. Learn.* 21 (2), 64–73.
- Maybach, C.W., 1996. Investigating urban community needs: service learning from a social justice perspective. *Educ. Urban Soc.* 28 (2), 224–236.
- Mieras, E., 2008. Latter-day knights. In: Knupher, A., Woyshner, C. (Eds.), *The Educational Work of Women's Organizations, 1890–1960*. Palgrave Macmillan, New York. https://doi.org/10.1057/9780230610125_6.
- Miettinen, R., 2000. The concept of experiential learning and John Dewey's theory of reflective thought and action. *Int. J. Lifelong Educ.* 19, 54–72.
- Mitchell, T.D., 2008. Traditional vs. critical service-learning: engaging the literature to differentiate two models. *Mich. J. Community Serv. Learn.* 14 (2), 50–65.
- Montessori, M., 1966. *The Secrets of Childhood*. Fides Publishers, New York.
- Morris, T.M., 2019. Experiential learning—a systematic review and revision of Kolb's model. *Interact. Learn. Environ.* 28, 1064–1077.
- National and Community Service Trust Act of 1990, http://www.nationalservice.gov/pdf/cncs_statute.pdf.
- National Association of Colored Women's Clubs, (n.d.). Our History, <https://www.nacwc.com/history>.
- National Service Learning Clearinghouse, 2013. Historical Timeline. http://www.servicelearning.org/what_is_service-learning/history.
- Nemeth, E.A., Winterbottom, C., 2015. Communities of practice: youth and social justice service-learning. In: Tinkler, A., Tinkler, B., Strait, J.R., Jagla, V.M. (Eds.), *Service-Learning to Advance Social Justice in a Time of Radical Inequality*. Information Age Publishing, Charlotte, NC, pp. 297–317.
- Norwood, A., 2017. Ida B. Wells-Barnett. National Women's History Museum. www.womenshistory.org/education-resources/biographies/ida-wells-barnett.
- Pettigrew, T.F., Tropp, L., 2006. A meta-analytic test of intergroup contact theory. *J. Pers. Soc. Psychol.* 90, 751–783.
- Rogoff, B., 1990. *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press, Oxford.
- Saltmarsh, J., 1996. Education for critical citizenship: John Dewey's contribution to the pedagogy of community service learning. *Mich. J. Community Serv. Learn.* 3, 12–21.
- Scales, P.C., Roehlkepartain, E.C., 2004. *Community Service and Service-Learning in U.S. Public Schools, 2004: Findings From a National Survey*. National Youth Leadership Council, St. Paul, MN, pp. 1–48.
- Scales, P.C., Roehlkepartain, E.C., Neal, M., Kielsmeir, J.C., Benson, P.L., 2006. Reducing academic achievement gaps: the role of community service and service learning. *J. Exp. Educ.* 29 (1), 38–60.
- Seaman, J., Brown, M., Quay, J., 2017. The evolution of experiential learning theory: tracing lines of research in the JEE. *J. Exp. Educ.* 40, NP1–NP21.
- Steptoe, T., (n.d.). Ida B. Wells-Barnett, Journalist, Civil Rights Activist and a Founding Member of the NAACP, Social Welfare History Project, <http://socialwelfare.library.vcu.edu/eras/civil-war-reconstruction/wells-barnett-ida-b/>.
- Stevens, C.S., 2003. Unrecognized roots of service-learning in African American social thought and action, 1890–1930. *Mich. J. Community Serv. Learn.* 19 (2), 25–34.
- Stoecker, R., 2016. *Liberating Service Learning and the Rest of Higher Education Civic Engagement*. Temple University Press, Philadelphia, PA.
- Vygotsky, L.S., 1978. *Mind in Society*. Harvard University Press, Cambridge, MA.
- Wade, R.C., 2000. From a distance: service-learning and social justice. In: O'Grady, C.R. (Ed.), *Integrating Service Learning and Multicultural Education in Colleges and Universities*. Erlbaum, Mahwah, NJ, pp. 93–111.
- Wertsch, J.M., Hagstrom, F., Kikas, E., 1995. Voices of thinking and speaking. In: Martin, L.M., Nelson, K., Tobach, E. (Eds.), *Sociocultural Psychology: Theory and Practice of Doing and Knowing. Learning in Doing: Social, Cognitive, and Computational Perspectives*. Cambridge University Press, New York, pp. 276–290.

Simulations and games for learning

Tasos Hovardas^a, Judith Ter Vrugte^b, Zacharias C. Zacharia^a, and Ton de Jong^{b, a} University of Cyprus, Nicosia, Cyprus; and
^b University of Twente, Enschede, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	681
A typology of simulations and games	681
Hybrid forms of simulations and games designed for learning	684
Learning with simulations, games, and their hybrid forms	685
Implications and recommendations for future research	686
References	687

Introduction

The use of computer simulations and games for educational purposes has increased dramatically during the last decade, following a corresponding increase in the technological capacity to support strategies in pedagogical design, instruction, and learning. This increase has been accompanied by several studies, reviews and meta-analyses of the various learning effects of simulations and games. This article aims to (a) present the major characteristics of simulations and games designed and used for learning, as well as, their hybrid forms, (b) discuss gamification in the transition from simulations to games, and (c) highlight recent findings and challenges for the use of simulations and games in education, especially with regard to inquiry-based learning (IBL) in science education. Even though, games have been traditionally associated with learning through game-based learning (GBL) theory (Li and Tsai, 2013; Qian and Clark, 2016; Wouters and van Oostendorp, 2013), we examine, through the extant literature, whether a transition from GBL to IBL is possible, since science education has been strongly connected to IBL. The ultimate goal is to examine whether the best of these two digital worlds could be combined and offered together, through a hybrid form of a game and a simulation, within the same computer supported inquiry-based learning environment. In this regard, the unique affordances of games and simulations are merged for enhancing learning. In the next section, we present a typology of simulations and games in an attempt to identify the different types of simulations and games, as well as their hybrid types, that currently exist in the extant literature. Then, we report on the relationship that these types of simulations, games, and hybrids have with science learning. Finally, we report on the implications that our review has for science learning and point to unresolved issues for future research.

A typology of simulations and games

Computer simulations and games share some core features but also diverge in several characteristics. Both operate on the basis of models, which are sets of rules that determine how their constituent variables are interconnected, and which refer to the systems that simulations and games represent. These underlying models regulate how simulations and games will respond to user input. A design concern, which has been much more expressed for games than simulations, however, is to weigh the complexity of the represented system against usability—the less complex the system, the easier it is for users to interact purposefully with the game (Lukosch et al., 2018; Wardaszko, 2018). In other words, the amount of story (i.e., background information) for a game should be as little as possible to allow game playing as soon as possible. In the case of simulations, the amount of story needed is usually greater for establishing meaningful interaction. The “richer” (i.e., the amount of information and the diversity of the relations a user can extract) or the less transparent (i.e., how “obvious” are the variables, objects and relations involved to users) a simulation is, the more story will be needed to run the simulation purposefully (Swaak et al., 1998; Zacharia and Olympiou, 2011).

There are several other major differences between simulations and games. First, simulations lack a narrative or scenario, which is a distinctive feature of games (Wilson et al., 2009). Such narratives interrelate three essential game elements: (a) The goal to be accomplished (which is different for simulations); (b) the gameplay with a start and end point that unfolds according to certain rules or constraints (whereas simulations are governed by information than rules); and (c) feedback for monitoring user progress in the form of levels, scores, or leaderboards (Parker and Becker, 2013). Second, games and simulations differ in that games demand decision-making by the user to determine the course of user experience (Lamb et al., 2018; Michael and Chen, 2006). For example, in games the users could proceed in a non-linear way, whereas in simulations users are required to make decisions in a more linear fashion. Third, simulations and games differ in terms of dynamics (i.e., user interaction with the game or simulation) and emotions

(i.e., the affective states that are formed after interaction with the game or simulation). Although both simulations and games use experience to convey meaning, games additionally include design elements for eliciting user emotions to stimulate engagement (Marsh, 2011).

Another major difference has to do with the fact that simulations are strictly tied to reality, which results in imitating how the physical or material world works. The representations used for presenting the objects, processes, systems and phenomena involved are drawn from real life. Games, on the other hand, do not usually have strong ties with reality. In fact, many games are based on fictional scenarios, which involve imaginary objects, processes, systems and phenomena.

In order to illustrate the, sometimes, subtle differences between simulations and games, four examples are provided below. Each example is based on an underlying model of projectile motion. The first figure represents an educational simulation, the second an educational gamified simulation, the third a simulation game, and the fourth a serious game.

Fig. 1–4 employ a model of projectile motion (i.e., 2D motion). The educational simulation in Fig. 1 allows students to manipulate variables to study this model. The PhET gamified educational simulation (Fig. 2) does the same, while adding context (i.e., a narrative scenario that puts a playful context to the model, including a specified goal). It lets you shoot a car (the projectile) out of a canon to see where it lands. While you are challenged to set the parameters so that you hit the target, you manage to study, as in Fig. 1, 2D motion. Fig. 3 depicts “Frog Jump,” a simulation game, where levels and challenge are introduced (e.g., the spacing of the lily pads goes from constant to variable, when you miss a pad the frog falls in the water and has to start over) and feedback is provided (e.g., when missing a jump, the formula to optimize the next attempt is prompted). Again, implicitly you are enacting physics laws of 2D motion. Fig. 4 shows the serious game “Angry Birds” by Rovio. In this game you shoot birds from a catapult to reach the levels’ objectives, which is to destroy all the green pigs that are hiding in shaky constructions. Aside from a playful narrative, game elements include, among others, challenge and levels (e.g., the number of birds is limited, the constructions create obstacles to overcome, pigs come in different sizes and might be harder to destroy, birds have unique projectile attributes), points, and leaderboards. Just like the former cases, the game is physics based and built around a model of projectile motion. Similar to the previous cases, users get to understand through the game, in an implicit manner, the variables (e.g., initial velocity of the projectile, the angle that the projectile is thrown) that affect the motion and trajectory of a projectile. However, this last example was not designed to train users to understand projectile motion beyond what they need to play the game better.

Figs. 1–4 illustrate how both simulations and games are basically a functional representation of an underlying model that shares the same principles. In these examples, the principles are those of projectile motion. When characteristic elements are adapted or added (e.g., narrative, goal, esthetics), the representation of the model can shift toward a playful simulation or a game. In the case of the four examples presented above, we have the educational simulation of Fig. 1 and the serious game of Fig. 4 being the two ends of the simulation-game continuum (if one excludes simulations designed for scientific research and games developed for entertainment as shown in Table 1), and between these two ends, the two hybrids (i.e., an educational gamified simulation, a simulation game) of Figs. 2 and 3. Hybrid forms of simulations and games could emerge by either gamifying simulations or by subtracting game elements from games. In the next section, we present hybrids of simulations and games as identified through the relevant literature. It should be noted that there is no consensus in the literature in terms of the terminology used. We present the most prevalent ones among academic scholars.

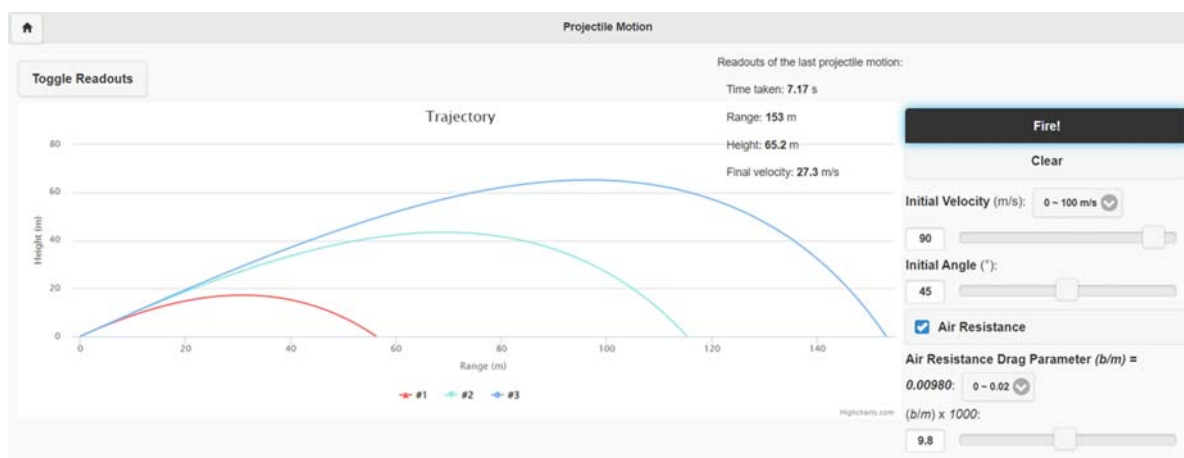


Fig. 1 An example of an educational simulation representation on projectile motion. Reproduced with permission from Michael Fowler. Projectile Motion Virtual Lab. Retrieved December 2020 from https://galileoandeinstein.phys.virginia.edu/more_stuff/Applets/Projectile/projectile.html.

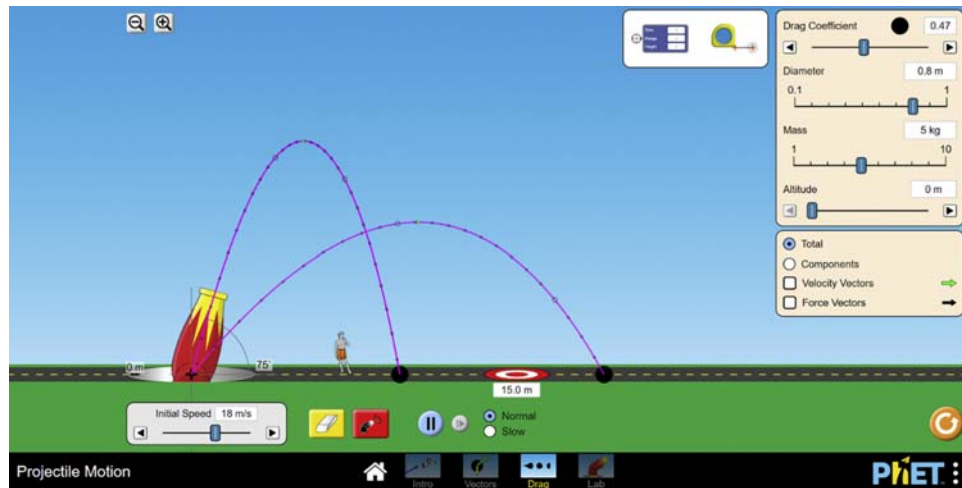


Fig. 2 An example of an educational gamified simulation representation on projectile motion (a PhET simulation). Reproduced with permission from Kathy Perkins, Phet Interactive Simulations. Projectile Motion Virtual Lab. Retrieved December 2020 from https://phet.colorado.edu/sims/html/projectile-motion/latest/projectile-motion_en.html. Screenshot by authors.



Fig. 3 An example of a simulation game representation on projectile motion. Reproduced with permission from Lookang Lawrence Wee, Iwant2study. Frog Jump Game. Retrieved December 2020 from https://iwant2study.org/lookangejss/02_newtonianmechanics_2kinematics/ejss_model_projectile_frog15/projectile_frog15_Simulation.xhtml. Screenshot by authors.



Fig. 4 An example of a serious game representation on projectile motion (made on the Angry Birds game). Reproduced with permission from Minna Eloranta, Rovio Entertainment Corporation. Angry Birds 2. Retrieved December 2020 from the android app Angry Birds 2. Screenshot by authors.

Table 1 Comparison between different types of simulations and games across several design parameters.

	<i>Simulations designed for scientific research</i>	<i>Educational simulations (including virtual labs)</i>	<i>Modeling tools with simulation capabilities</i>	<i>Gamified educational simulations</i>	<i>Simulation games</i>	<i>Serious games</i>	<i>Games developed for entertainment</i>
Play or learning as the main design objective	Neither play nor learning	Learning	Learning	Learning prevalent but balanced against play	Learning prevalent but balanced against play	Learning prevalent but balanced against play	Play prevalent; learning may occur as a side-effect of use
Narrative and challenge	Absent	Absent	Absent	Either absent (structural gamification) or present (content gamification)	Present	Present	Present
Modeling options	Research-driven, when present	Absent for simulations with provided experiment setup; present for virtual labs that allow building own experimental setups	Present	Absent	Absent	Absent	Absent
Data extraction and processing taken up by the user	Should allow for research and development purposes	Necessary; should allow for educational purposes	Necessary; should allow for educational purposes	Necessary; should allow for educational purposes	Not foreseen within game flow	Not foreseen within game flow	Not foreseen within game flow
MDE framework	Not incorporated	Not incorporated	Not incorporated	Integrating aspects of MDE framework	Based on the MDE framework	Based on the MDE framework	Based on the MDE framework
Opportunities for inquiry-based learning	Irrelevant	Considerable	Considerable	Considerable	Existing	Existing	Irrelevant
Iterations determined by	Variable selection	Variable selection for simulations with provided experiment setup; model building and refinement for virtual labs that allow building own experiment setups	Model building and refinement	Variable selection	Simulation constraints	User actions in the gameplay	User actions in the gameplay
Course of user experience	Pre-defined or partly determined by user actions	Pre-defined for simulations with provided experiment setup; partly determined by user actions for virtual labs that allow building own experiment setups	Partly determined by user actions	Mostly pre-defined	Mostly pre-defined	Partly determined by user actions	Partly determined by user actions

Note: MDE: Mechanics, Dynamics, and Emotions Framework (Robson et al., 2015).

Hybrid forms of simulations and games designed for learning

Although simulations were first employed for research purposes and games for entertainment, there were versions of both simulations and games designed and used for educational purposes. Apart from educational simulations and serious games, which have been the archetypical versions of simulations and games, respectively, to serve educational purposes, there are also hybrid, in-between forms, which attempt to employ and marry the unique affordances of both simulations and games (e.g., modeling tools with simulation capabilities; gamified educational simulations; simulation games). In this section we present core characteristics of educational simulations, serious games, and hybrid, in-between forms (see Table 1).

Educational simulations allow users to change one or more values of variables and then investigate the effect of this intervention on values of other variables built in the system (de Jong, 2011). They may even allow the building of different scenarios or setups. For instance, in the case of science education, users could use virtual labs that allow building different experiment setups, before starting to investigate relationships between variables (Zacharia, 2007). Virtual laboratories are a specific form of educational digital simulations, which can be combined with software scaffolds and other applications to provide a holistic learning experience (de Jong et al., 2021). Although educational simulations allow users to explore the system freely, the underlying model cannot be changed by users, and user experience is largely pre-defined. In the case of virtual labs, however, the possibilities may increase. For instance, there are virtual labs that allow the building of different experiment setups, which allow the user to switch from one underlying model to the other, or even create their own model which, of course, follows the already pre-determined algorithms. In addition, there are modeling tools with simulation capabilities. These tools provide users with modeling options for model building and refinement (see, for instance, GoModel; <https://www.golabz.eu/app/gomodel>). In this case, the course of user experience is partly determined by user actions, as in the case of the virtual labs. This latter characteristic is also found in serious games (see below).

Educational simulations may be enriched with game design aspects (i.e., game-based mechanics, esthetics, and game thinking), an instructional design approach which aims to facilitate user engagement and which is termed “gamification” (see Darejeh and Salim, 2016; Hamari, 2017). Gamification may employ points, badges, and leaderboards (Brom et al., 2019). These elements do not affect content and are characterized as *structural* gamification. *Content* gamification employs game-elements like narrative and challenge, which are expected to not only foster user motivation but also to enhance metacognition (Anderson and Barnett, 2013; Hou, 2015). Gamified educational simulations may display characteristics of structural and content gamification separately or in a way that complements each other.

Simulation games are full-fledged games engaging students in problem-solving tasks organized around a core challenge or mission (Hong et al., 2020). Simulation games incorporate in their game mechanics and dynamics a full content gamification which unifies both the user’s role and decision making (user input), as well as the processing of user input to deliver the system output. Since most simulation games are rigid-rule games (Gust and Klabbers, 2005), the gameplay is largely pre-determined by simulation constraints. Serious games, instead, evolve around more open-ended tasks (for a discussion of simulation games and serious games see Kurbjuhn, 2012). Serious games often have a real-life operational task as their topic (e.g., how to act in emergency situations; e.g., Steinrück et al., 2020). The adjective “serious” is meant to denote that learning is prevalent in serious games, although it is balanced against play (Prensky, 2001; Michael and Chen, 2006).

Learning with simulations, games, and their hybrid forms

Simulations, including virtual laboratories, are considered most useful for promoting inquiry-based science learning (Hovardas et al., 2018). IBL is currently among the most advocated educational strategies for teaching science across K-16 (Pedaste et al., 2015). Its main goal is to have students simulate the methods and practices that scientists follow in order to learn (for details on IBL, see Pedaste et al., 2015). Using simulations in an IBL context requires students to assume the ownership of the learning process by formulating their own questions or hypotheses and to collect data to address them (de Jong, 2019). Simulations, including virtual laboratories, allow students to identify and select variables to perform their investigations and enable them to manipulate the values of independent variables and examine the effect of this manipulation on the values of dependent variables (de Jong et al., 2013).

A series of recent studies, focusing on educational simulations and applications hosted in the Go-Lab ecosystem (www.golabz.eu), showed promising results. For instance, the inclusion of virtual laboratories and applications for preparing experimental designs and constructing graphs in digital lessons was found to enhance the corresponding student skills (e.g., Efstathiou et al., 2018; Xenofontos et al., 2020). Given that IBL is a complex and demanding procedure, software scaffolds are usually employed to support students (Zacharia et al., 2015). Of course, scaffolds can be adapted to learner progress (de Jong, 2019) by fading out when students no longer need the support provided. Some web-based applications with authoring tools offer teaching analytics functionalities to let teachers evaluate the guidance planned for their students in their IBL educational designs (Sergis et al., 2019). These developments increase the capabilities of teachers to enact formative assessment in web-based learning environments and better support their students’ educational paths. However, the need for support in IBL also raises concerns as to why guidance is needed for IBL to deliver its expected positive outcomes (e.g., Aditomo and Klieme, 2020).

Despite the complexity of enacting IBL, even when it is computer supported, researchers and educators are strongly advocating its enactment for science learning purposes (e.g., Alfieri et al., 2011; Furtak et al., 2012; Minner et al., 2010). For instance, the use of simulations in an IBL context was found to be more or equally conducive to student science learning, as opposed to using physical counterparts/manipulatives for experimentation purposes, across K-16 (e.g., Zacharia et al., 2012; Zacharia and de Jong, 2014; Zacharia and Michael, 2016). As a result, IBL is considered by now as an integral part of an individual’s science learning. The question that is raised at this point is whether games or hybrids of simulations and games could also be situated in an IBL context and enhance student science learning.

The use of games and game elements (gamification) for learning (GBL) have also been found to have a positive effect on students' learning outcomes. GBL has been recently documented to deliver a range of desirable learning outcomes, including cognitive, behavioral, and affective effects (Clark et al., 2016), which paralleled previous findings (e.g., Connolly et al., 2012). Further, GBL has been found to promote the development of 21st learning skills (Qian and Clark, 2016). The literature on GBL singles out flow (Chang et al., 2017), competition (Chen et al., 2020) and narrative (Naul and Liu, 2020) as major strengths of games, which may promote cognitive (e.g., achievement and retention) and non-cognitive goals (e.g., enjoyment).

Well established psychological theories provide explanations for how effective learning mechanisms are triggered by games and game elements (Sailer and Homner, 2020; van Roy and Zaman, 2017). The engaging effect of games is commonly explained using Self-Determination Theory (SDT) (van Roy and Zaman, 2017). SDT explains that engagement is the result of an internal drive to fulfill a set of basic psychological needs: competence, relatedness, and autonomy. To foster this kind of engagement in games, designers employ multiple "mechanics" (systems of elements and rules) that appeal to these needs. For instance, game elements such as badges appeal to the need for competence, leaderboards or collaborations appeal to the need for relatedness, and choices (when to play or where to go next) appeal to the need for autonomy.

In addition the success of the game depends on an opaque interaction between the mechanics, the behavior of the player (dynamics), and the emotions that these components evoke (according to the Mechanics, Dynamics and Emotions (MDE) framework, see Robson et al., 2015). Mechanics concern the model that shows how the game works, in terms of its elements, rules, goals, and progression. Dynamics occur through the interaction of the mechanics and the user. However, it is hard to predict how this interaction will evolve due to other unforeseen variables that might become involved by bending or breaking, for example, the mechanics. Emotions concern the affective state that the user is at in regards to the mechanics, dynamics, and content of a game.

The engaging effect of GBL has been confirmed in science education as well, with game flow being positively correlated with reflective processes (Hou, 2015). However, similar to IBL, GBL benefits from support for students during and after the gameplay, like scaffolds and reflection tasks targeting specific knowledge and skills (Iten and Petko, 2016), to safeguard learning objectives and facilitate transfer (Ter Vrugte and de Jong, 2017). In this regard, possible trade-offs between game design aspects and learning outcomes should be considered (Imlig-Iten and Petko, 2018; Landers and Landers, 2015; Ter Vrugte and de Jong, 2017). Designers should take into account that instructional add-ons could distract users from the flow of simulation games (Erb, 2015) or serious games (Baker and Delacruz, 2016). In addition, challenges posed to students and instructional support need to align with their current level of skills so that flow and engagement are maintained (Qian and Clark, 2016; Yen and Lin, 2020). Depending on the level of skills, specific support might have a unique or even disruptive effect (see Ter Vrugte et al., 2015, for an example in GBL). A similar concern is that the net effect of game elements like goals, rewards, and feedback needs to be carefully examined, as far as their overall effect on cognitive load is concerned (Brom et al., 2019).

Although IBL has been strongly advocated for science education (Li and Tsai, 2013), GBL has not, partly because it has not been advocated for and linked with IBL. Some initial attempts have shown considerable potential (e.g., Taub et al., 2018) but more research is needed to substantiate this synergy. The integration of IBL and GBL is expected to accommodate a balanced contribution of both play and learning in student experience (National Research Council, 2011). Such an integration should also incorporate an appropriate level of guidance provided to students in the form of instructional support (e.g., Ter Vrugte and de Jong, 2012, 2017; Wouters and van Oostendorp, 2013). In our view, one way to link IBL to GBL is through the identification of overlaps between the processes involved for their enactments. For instance, both IBL and GBL involve data collection and processing within each of their iterations. Of course, as user immersion in simulation games and serious games increases, as compared to gamified educational simulations, data selection and processing for IBL would require a reallocation of time-on-task toward inquiry activities, which might compromise user immersion and flow. This is an issue to be addressed by future research.

Implications and recommendations for future research

The salience of educational simulations and serious games has risen sharply during the COVID-19 pandemic, which, in 2020, impacted more than 1.5 billion students globally. Simulations and games could be employed to lead science learning and inquiry enactments under the restrictions imposed during this crisis. However, more research is needed to identify how simulations, games or their hybrids could be used to enhance students' online science learning experiences. It is quite clear from the literature that the more we move toward the game end of the simulations-games continuum, the less we now in terms of enacting games and their hybrids in an IBL context. This is a result of researchers and educators' failure to examine the possibilities of using games or hybrids in an inquiry context. We believe that this is feasible for hybrids, since they have a significant overlap with simulations, which have been successfully used in IBL context. In the case of games, it becomes harder to understand whether games could be grounded in IBL, due to the absence of research evidence. Thus, this is a vital area for future research.

Although educational simulations have been well situated in an IBL context, we are still missing pieces of the puzzle concerning the integration of GBL with IBL. For instance, it would be insightful to investigate if simulations, games, and their hybrids could be employed in science education and if GBL could be combined with IBL. Would it be as easy to do so in all cases, or would it be easier through the use of the hybrids? Needless to say, research should be employed to answer these questions and to shed light on possible trade-offs that GBL and IBL could suffer. For instance, immersion and flow (sought for GBL) could be jeopardized (in an integration of GBL with IBL), while trial-and-error user behavior (the usual heuristic in GBL) would compete with the identification of variable relationships through systematic experimentation (as practiced in IBL).

The idea of combining simulations and games is to bring the best of both these worlds into students' science learning experiences. Both of them offer unique affordances that have been found to positively influence science learning. The issue now is how to combine the two digital worlds in ways that optimize students' science learning. This is an agenda for future research.

References

- Aditomo, A., Klieme, E., 2020. Forms of inquiry-based science instruction and their relations with learning outcomes: evidence from high and low performing education systems. *Int. J. Sci. Educ.* 42, 504–525. <https://doi.org/10.1080/09500693.2020.1716093>.
- Alfieri, L., Brooks, P.J., Aldrich, N.J., Tenenbaum, H.R., 2011. Does discovery-based instruction enhance learning? *J. Educ. Psychol.* 103 (1), 1–18. <https://doi.org/10.1037/a0021017>.
- Anderson, J.L., Barnett, M., 2013. Learning physics with digital game simulations in middle school science. *J. Sci. Educ. Technol.* 22, 914–926. <https://doi.org/10.1007/s10956-013-9438-8>.
- Baker, E.L., Delacruz, G.C., 2016. A framework to create effective learning games and simulations. In: O'Neil, H.F., Baker, E.L., Perez, R.S. (Eds.), *Using Games and Simulations for Teaching and Assessment*. Routledge, New York, pp. 3–20. <https://doi.org/10.4324/9781315817767>.
- Brom, C., Stárková, T., Bromová, E., Děchtěrenko, F., 2019. Gamifying a simulation: do a game goal, choice, points, and praise enhance learning? *J. Educ. Comput. Res.* 57, 1575–1613. <https://doi.org/10.1177/0735633118797330>.
- Chang, C.C., Liang, C., Chou, P.N., Lin, G.Y., 2017. Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Comput. Hum. Behav.* 71, 218–227. <https://doi.org/10.1016/j.chb.2017.01.031>.
- Chen, C.H., Shih, C.C., Law, V., 2020. The effects of competition in digital game-based learning (DGBL): a meta-analysis. *Educ. Technol. Res. Dev.* 68, 1855–1873. <https://doi.org/10.1007/s11423-020-09794-1>.
- Clark, D.B., Tanner-Smith, E.E., Killingsworth, S.S., 2016. Digital games, design, and learning: a systematic review and meta-analysis. *Rev. Educ. Res.* 86, 79–122. <https://doi.org/10.3102/0034654315582065>.
- Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., Boyle, J.M., 2012. A systematic literature review of empirical evidence on computer games and serious games. *Comput. Educ.* 59, 661–686. <https://doi.org/10.1016/j.compedu.2012.03.004>.
- Darejeh, A., Salim, S.S., 2016. Gamification solutions to enhance software user engagement—a systematic review. *Int. J. Hum. Comput. Interact.* 32, 613–642. <https://doi.org/10.1080/10447318.2016.1183330>.
- de Jong, T., Linn, M., Zacharia, Z.C., 2013. Physical and virtual laboratories in science and engineering education. *Science* 340, 305–308. <https://doi.org/10.1126/science.1230579>.
- de Jong, T., Gillet, D., Rodríguez-Triana, M.J., et al., 2021. Understanding teacher design practices for digital inquiry-based science learning: the case of Go-Lab. *Educ. Technol. Res. Dev.* 69, 417–444. <https://doi.org/10.1007/s11423-020-09904-z>.
- de Jong, T., 2011. Instruction based on computer simulation. In: Mayer, R.E., Alexander, P.A. (Eds.), *Handbook of Research on Learning and Instruction*. Routledge, New York, NY, pp. 446–466.
- de Jong, T., 2019. Moving towards engaged learning in STEM domains; there is no simple answer, but clearly a road ahead. *J. Comput. Assist. Learn.* 35, 153–167. <https://doi.org/10.1111/jcal.12337>.
- Efstathiou, C., Hovardas, T., Xenofontos, N., Zacharia, Z., de Jong, T., Anjewierden, A., van Riesen, S.A.N., 2018. Providing guidance in virtual lab experimentation: the case of an experiment design tool. *Educ. Technol. Res. Dev.* 66, 767–791. <https://doi.org/10.1007/s11423-018-9576-z>.
- Erb, U., 2015. Possibilities and limitations of transferring an educational simulation game to a digital platform. *Simulat. Gaming* 46, 817–837. <https://doi.org/10.1177/1046878115621980>.
- Furtak, E.M., Seidel, T., Iverson, H., Briggs, D.C., 2012. Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev. Educ. Res.* 82 (3), 300–329. <https://doi.org/10.3102/0034654312457206>.
- Gust, M., Klabbers, J.H.G., 2005. Unterschiede zwischen offenen und geschlossenen Planspielen. In: Blötz, U. (Ed.), *Planspiele in der beruflichen Bildung: Auswahl, Konzepte, Lernarrangements, Erfahrungen—Aktueller Planspielkatalog*. W. Bertelsmann, Bielefeld, pp. 58–67.
- Hamari, J., 2017. Do badges increase user activity? A field experiment on the effects of gamification. *Comput. Hum. Behav.* 71, 469–478. <https://doi.org/10.1016/j.chb.2015.03.036>.
- Hong, J.-C., Hwang, M.-Y., Liu, M.-C., Tsai, C.R., Tai, K.-H., 2020. Metacognition in covariation reasoning relevant to performance achievement mediated by experiential values in a simulation game. *Educ. Technol. Res. Dev.* 68, 929–948. <https://doi.org/10.1007/s11423-020-09904-z>.
- Hou, H.-T., 2015. Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: a video-based process exploration. *Comput. Hum. Behav.* 48, 424–435. <https://doi.org/10.1016/j.chb.2015.02.010>.
- Hovardas, T., Pedaste, M., Zacharia, Z., de Jong, T., 2018. Model-based science inquiry in computer-supported learning environments: the case of Go-Lab. In: Azad, A.K.M., Auer, M., Edwards, A., de Jong, T. (Eds.), *Cyber-Physical Laboratories in Engineering and Science Education*. Springer, Dordrecht, pp. 241–268. https://doi.org/10.1007/978-3-319-76935-6_10.
- Imlig-Iten, N., Petko, D., 2018. Comparing serious games and educational simulations: effects on enjoyment, deep thinking, interest and cognitive learning gains. *Simulat. Gaming* 49, 401–422. <https://doi.org/10.1177/1046878118779088>.
- Iten, N., Petko, D., 2016. Learning with serious games: is fun playing the game a predictor of learning success? *Br. J. Educ. Technol.* 47, 151–163. <https://doi.org/10.1111/bjet.12226>.
- Kurbjuhn, B., 2012. Serious and simulation games—a definition approach. In: Abramowicz, W., Domingue, J., Węcel, K. (Eds.), *Business Information Systems Workshops*. Springer, Heidelberg, pp. 177–185. <https://doi.org/10.1007/978-3-642-34228-8>.
- Lamb, L.R., Annetta, L., Firestone, J., Etopio, E., 2018. A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Comput. Hum. Behav.* 80, 158–167. <https://doi.org/10.1016/j.chb.2017.10.040>.
- Landers, R.N., Landers, A.K., 2015. An empirical test of the theory of gamified learning: the effect of leaderboards on time-on-task and academic performance. *Simulat. Gaming* 45, 769–785. <https://doi.org/10.1177/1046878114563662>.
- Li, M.C., Tsai, C.C., 2013. Game-based learning in science education: a review of relevant research. *J. Sci. Educ. Technol.* 22, 877–898. <https://doi.org/10.1007/s10956-013-9436-x>.
- Lukosch, H.K., Bekebrede, G., Kurapati, S., Lukosch, S.G., 2018. A scientific foundation of simulation games for the analysis and design of complex systems. *Simulat. Gaming* 49, 279–314. <https://doi.org/10.1177/1046878118768858>.
- Marsh, T., 2011. Serious games continuum: between games for purpose and experiential environments for purpose. *Entertain. Comput.* 2, 61–68. <https://doi.org/10.1016/j.entcom.2010.12.004>.
- Michael, D.R., Chen, S.L., 2006. *Serious Games: Games That Educate, Train, and Inform*. Thomson Course Technology PTR, Massachusetts, USA.
- Minner, D.D., Levy, A.J., Century, J., 2010. Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984–2002. *J. Res. Sci. Teach.* 47 (4), 474–496. <https://doi.org/10.1002/tea.20347>.
- National Research Council, 2011. *Learning Science Through Computer Games and Simulations*. The National Academies Press, Washington DC.

- Naul, E., Liu, M., 2020. Why story matters: a review of narrative in serious games. *J. Educ. Comput. Res.* 58, 687–707. <https://doi.org/10.1177/0735633119859904>.
- Parker, J.R., Becker, K., 2013. The simulation-game controversy: what is a ludic simulation? *Int. J. Gaming Comput. Mediated Simulat.* 5, 1–12. <https://doi.org/10.4018/jgcms.2013010101>.
- Pedaste, M., Mäeots, M., Siiman, L.A., de Jong, T., van Riesen, S.A.N., Kamp, E.T., Manoli, C.C., Zacharia, Z.C., Tsourlidaki, E., 2015. Phases of inquiry-based learning: definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61. <https://doi.org/10.1016/j.edurev.2015.02.003>.
- Prensky, M., 2001. Digital natives, digital immigrants. *Horizon* 5, 1–6.
- Qian, M., Clark, K.R., 2016. Game-based Learning and 21st century skills: a review of recent research. *Comput. Hum. Behav.* 63, 50–58. <https://doi.org/10.1016/j.chb.2016.05.023>.
- Robson, K., Plangger, K., Kietzmann, J.H., McCarthy, I., Pitt, L., 2015. Is it all a game? Understanding the principles of gamification. *Bus. Horiz.* 58, 411–420. <https://doi.org/10.1016/j.bushor.2015.03.006>.
- Sailer, M., Homner, L., 2020. The gamification of learning: a meta-analysis. *Educ. Psychol. Rev.* 32, 77–112. <https://doi.org/10.1007/s10648-019-09498-w>.
- Sergis, S., Sampson, D.G., Rodríguez-Triana, M.J., Gillet, D., Pelliccione, L., de Jong, T., 2019. Using educational data from teaching and learning to inform teachers' reflective educational design in inquiry-based STEM education. *Comput. Hum. Behav.* 92, 724–738. <https://doi.org/10.1016/j.chb.2017.12.014>.
- Steinrück, J., Veldkamp, B.P.M., de Jong, T., 2020. Information literacy skills assessment in digital crisis management training. *Front. Educ.* 5. <https://doi.org/10.3389/educ.2020.00140>.
- Swaak, J., Van Joolingen, W.R., De Jong, T., 1998. Supporting simulation-based learning; the effects of model progression and assignments on definitional and intuitive knowledge. *Learn. Instr.* 8 (3), 235–252. [https://doi.org/10.1016/S0959-4752\(98\)00018-8](https://doi.org/10.1016/S0959-4752(98)00018-8).
- Taub, M., Azevedo, R., Bradbury, A.E., Millar, G.C., Lester, J., 2018. Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learn. Instr.* 54, 93–103. <https://doi.org/10.1016/j.learninstruc.2017.08.005>.
- Ter Vrugte, J., de Jong, T., 2012. How to adapt games for learning: the potential role of instructional support. In: de Wannemacker, S., Vandercruysse, S., Clarebout, G. (Eds.), *Serious Games: The Challenge*. Springer, pp. 1–5. https://doi.org/10.1007/978-3-642-33814-4_1.
- Ter Vrugte, J., de Jong, T., 2017. Self-explanations in game-based Learning: from tacit to transferable knowledge. In: Wouters, P., van Oostendorp, H. (Eds.), *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*. Advances in Game-Based Learning. Springer, Cham, pp. 141–159. https://doi.org/10.1007/978-3-319-39298-1_8.
- Ter Vrugte, J., de Jong, T., Wouters, P., Vandercruysse, S., Elen, J., van Oostendorp, H., 2015. When a game supports prevocational matheducation but integrated reflection does not. *J. Comput. Assist. Learn.* 31, 462–480. <https://doi.org/10.1111/jcal.12104>.
- van Roy, R., Zaman, B., 2017. Why gamification fails in education and how to make it successful: introducing nine gamification heuristics based on self-determination theory. In: Ma, M., Oikonomou, A. (Eds.), *Serious Games and Edutainment Applications*. Springer, Cham, pp. 485–509. https://doi.org/10.1007/978-3-319-51645-5_22.
- Wardaszko, M., 2018. Interdisciplinary approach to complexity in simulation game design and implementation. *Simulat. Gaming* 49, 263–278. <https://doi.org/10.1177/1046878118777809>.
- Wilson, K.A., Bedwell, W.L., Lazzara, E.H., Salas, E., Burke, C.S., Estock, J.L., Orvis, K.L., Conkey, C., 2009. Relationships between game attributes and learning outcomes: review and research proposals. *Simulat. Gaming* 40, 217–266. <https://doi.org/10.1177/1046878108321866>.
- Wouters, P., van Oostendorp, H., 2013. A meta-analysis review of the role of instructional support in game-based learning. *Comput. Educ.* 60, 412–425. <https://doi.org/10.1016/j.compedu.2012.07.018>.
- Xenofontos, N., Hovardas, T., Zacharia, Z.C., De Jong, T., 2020. Inquiry-based learning and retrospective action: problematizing student work in a computer-supported learning environment. *J. Comput. Assist. Learn.* 36, 12–28. <https://doi.org/10.1111/jcal.12384>.
- Yen, W.-C., Lin, H.-H., 2020. Investigating the effect of flow experience on learning performance and entrepreneurial self-efficacy in a business simulation systems context. *Interact. Learn. Environ.* <https://doi.org/10.1080/10494820.2020.1734624>.
- Zacharia, Z.C., de Jong, T., 2014. The effects on students' conceptual understanding of electric circuits of introducing virtual manipulatives within a physical manipulatives-oriented curriculum. *Cognit. Instr.* 32 (2), 101–158. <https://doi.org/10.1080/07370008.2014.887083>.
- Zacharia, Z.C., Michael, M., 2016. Using physical and virtual manipulatives to improve primary school students' understanding of concepts of electric circuits. In: Riopel, M., Smyrniou, Z. (Eds.), *New Developments in Science and Technology Education, Innovations in Science Education and Technology*, vol. 23. Springer, Cham. https://doi.org/10.1007/978-3-319-22933-1_12.
- Zacharia, Z.C., Olympiou, G., 2011. Physical versus virtual manipulative experimentation in physics learning. *Learn. Instr.* 21 (3), 317–331. <https://doi.org/10.1016/j.learninstruc.2010.03.001>.
- Zacharia, Z.C., Loizou, E., Papaevripidou, M., 2012. Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Child. Res. Q.* 27, 447–457. <https://doi.org/10.1016/j.ecresq.2012.02.004>.
- Zacharia, Z.C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S.A., et al., 2015. Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. *Educ. Technol. Res. Dev.* 63, 257–302. <https://doi.org/10.1007/s11423-015-9370-0>.
- Zacharia, Z.C., 2007. Comparing and combining real and virtual experimentation: an effort to enhance students' conceptual understanding of electric circuits. *J. Comput. Assist. Learn.* 23 (2), 120–132. <https://doi.org/10.1111/j.1365-2729.2006.00215.x>.

The efficacy of learning strategies from the ICAP perspective

Micheline T.H. Chi, Nicole S. Boucher, and Jesse Ha, Arizona State University, Tempe, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	689
The ICAP theory	690
Passive	690
Active	691
Constructive	691
Interactive	692
Criteria for discriminating between modes	692
The ICAP hypothesis	692
Using the ICAP theory to evaluate cognitive strategies	692
<i>Passive</i> mode strategies	694
<i>Active</i> mode strategies	694
<i>Constructive</i> mode strategies	694
Compliance	695
More or less complicated processing	695
Invisible thinking processes	695
Generating vs. retrieval	695
<i>Interactive</i> mode strategies	695
Evidence in support of ICAP hypothesis	696
Conclusion: the value of ICAP	697
References	697

Introduction

Using and applying strategies to learn new information in a content domain have been investigated in the literature for decades. The terms “using or applying strategies” refer to an intentional conscious action undertaken by learners. Such a definition refers to at least two kinds of strategies, the *cognitive* strategies that are used for learning new information, and the *metacognitive* strategies used for regulating the use of cognitive strategies. Cognitive strategies include examples such as rehearsing and elaborating that can be used to learn new information; whereas metacognitive strategies include planning which strategy to use and when to apply it, as well as monitoring and evaluating one’s understanding. Regulation requires processes such as self-assessment, deciding which strategy to use based on the goal of learning in that instance, and so forth. In short, strategies for learning are distinct from strategies for regulation of learning. The focus of this paper is on cognitive strategies for learning (specific strategies referring to what learners actually do will always be italicized).

Cognitive strategies generally refer to efforts undertaken by learners or students to increase their learning of a given content domain of new information, such as learning expository information about Civil war history or learning procedural knowledge on how to solve geometry problems. The term “learning” has been used broadly in many laboratory studies to mean memory retention of words, facts, and procedures, instead of learning with deep understanding of difficult concepts and processes. Learning with deeper understanding is the goal of classroom instruction, especially in higher education and STEM (science, technology, engineering and math) disciplines. Thus, the goals of classroom instruction are often different from many laboratory studies. This discrepancy between what is measured in many laboratory studies (memory retention) and the goal of instruction (deeper learning), especially in college STEM classrooms, will be noted throughout this chapter.

Many cognitive strategies have been studied for decades, and many meta-analyses have been undertaken (Donker et al., 2014; Fiorella and Mayer, 2016; Hattie, 1996). On the whole, there is no question that using cognitive strategies enhances learning, especially during initial learning of new information, which is the focus of this chapter. But the crucial questions are: Are some strategies better for learning with deeper understanding than other strategies? If they do have differential effectiveness, how can we tell which strategies are more or less effective?

There are many ways to evaluate or predict the potential effectiveness of a learning strategy. One way is to analyze its underlying cognitive processes. For example, the category of “elaborating” strategies, with examples such as summarizing and paraphrasing, involves the processes of “building internal connections between the items to be learned and already existing knowledge,” according to Donker et al. (2014, p. 3). However, it is not easy to know that these two strategies map to the underlying processes of “building internal connections between the items to be learned and already existing knowledge,” since cognitive thinking processes are not directly observable, especially in a classroom context. Other considerations and methods used for evaluating the utility of a strategy include the ease with which students can be trained to adopt the strategy (Fiorella and Mayer, 2016), the ease of implementation in

classrooms (Dunlosky et al., 2013), the age group for whom they are suitable (Brod, 2020), and so forth. The most optimal evaluation method is to seek empirical evidence for a strategy's effectiveness by comparing them side-by-side, as in the case of many laboratory studies.

In this chapter, although we prioritize the importance of having empirical evidence for determining the utility of a strategy and appeal to the possible underlying cognitive processes, we point out a major flaw in the literature for determining effectiveness by contrasting it with an alternative strategy or control condition. This flaw will be pointed out below. We propose, instead, to analyze the utility of a strategy from the perspective of a new theory of how students engage to learn, called ICAP.

Using ICAP, we illustrate how the effectiveness of a variety of strategies can be evaluated without relying on empirical findings. These strategies are chosen using several informal criteria such as: (a) strategies that do not require technology support nor advanced preparation by the teachers (as those may not be easily implemented in a classroom), (b) evidence-based effective strategies (i.e., those that have been investigated in laboratory studies), (c) strategies that are commonly used in classrooms, and (d) strategies that students often utilize (a criterion also used by Dunlosky et al., 2013). We arrived at 50 examples of strategies.

The ICAP theory

ICAP is a theory of cognitive engagement and it provides a framework that outlines different levels of student engagement and learning. ICAP focuses on classifying students' engagement with instruction rather than classifying instructors' instructional objectives or assessment items as commonly discussed in the literature (Bloom and Krathwohl, 1956; Dick and Carey, 2005; Krathwohl, 2002). Moreover, ICAP has a unique approach that uses student behaviors and student-produced outputs for assessing students' engagement level.

ICAP distinguishes among four modes of cognitive engagement: *Interactive*, *Constructive*, *Active*, and *Passive*. The taxonomy of the four modes is discussed in detail along with extensive supporting evidence from the literature in the work of Chi (2009), Chi and Wylie (2014), Fonseca and Chi (2011), and Chi et al. (2018). Four criteria, presented in Table 1 (left column), are provided to facilitate a more accurate discrimination of learning strategies from each mode, because some criteria are more salient for discriminating between one mode from another for some strategies, whereas other criteria are better at discriminating the modes of other strategies. The four criteria are: (1) physically overt behaviors, (2) visible outputs produced, (3) additional focusing or re-focusing on selected parts of the instructional materials, and (4) outputs contain new information. Caveats of each mode are also mentioned.

These four modes are briefly described next, along with their plausible underlying cognitive or thinking processes relevant to learning. We propose that learning involves these six elementary, knowledge-change thinking processes (Chi, 2021): receiving and storing new information, activating prior knowledge, connecting new information with prior knowledge, changing some parts of prior knowledge, inferring from new information and/or prior knowledge, and reflecting on new information or prior knowledge.

Passive

The *Passive* mode of cognitive engagement refers to learners' overt behavior when they are oriented toward instruction (i.e., commonly thought of as paying attention) without producing any outputs. For example, listening to a lecture without taking notes, reading a textbook without highlighting, or watching a video without taking notes, would be considered activities in the *Passive* mode of cognitive engagement. Because learners are not producing any outputs, they are not focusing on a selection of the

Table 1 Four criteria to discriminate among the ICAP modes.

Criteria	ICAP mode			
	Passive	Active	Constructive	Interactive
Physically overt behaviors (include talking and writing)	<i>Oriented toward instruction</i> (or paying attention)	<i>Manipulating</i> some aspects of the instructional materials	<i>Generating</i> based on some aspects of the instructional materials	<i>Reciprocally co-generating</i> from each other's contributions, which were based on some aspects of instructional materials
Visible outputs produced	No	Yes (often)	Yes (often)	Yes (often)
Additional focusing or re-focusing on selected parts	No	Yes	Yes	Yes
Outputs contain new information	No	No	Yes, produced by an individual	Yes, often produced by partners jointly

instructional materials involved, nor any new information produced. So, *Passive* mode has only a physical behavior component or criterion feature (see [Table 1](#), Column 2). When students are not paying attention, then they are not even engaged with instruction, such as when they chat with a friend, doze off, or talk on the phone.

When students are engaging in *Passive* activities, such as watching a video, ICAP assumes that students' cognitive processes are likely to only be receiving information and presumably at best storing it in their memory probably in isolation, meaning it is not connected to some other relevant prior knowledge. A caveat about the *Passive* mode is that, when students are not displaying any overt behaviors other than attentively listening or watching, they could be covertly engaged in deeply processing information, such as inferring, reflecting, or connecting new information with prior knowledge. However, ICAP makes the probabilistic assumption that when students are being *Passive*—not displaying any overt learning behaviors while learning—they are more likely to be engaged only in the thinking processes of receiving and storing.

Active

The *Active* mode of cognitive engagement occurs when students display the overt behavior of physically manipulating instructional materials, such as highlighting or underlining specific sections of a text ([Katayama et al., 2005](#)), pausing and rewinding sections of a video ([Chi et al., 2008](#)), copying verbatim portions of the problem solution steps ([VanLehn et al., 2007](#)), or choosing a justification from a menu ([Conati and VanLehn, 2000](#)). The overt outputs are the highlighted or underlined sentences, the replayed portions of the video, and so forth.

Manipulation often requires students to select portions of the materials, such as selecting a specific sentence or passage to underline or selecting which parts of an instructor's presentation to copy. Selection can often be made based on syntactic cues in the text, such as signaling an important principle, putting important information inside a box, and so forth. The consequence of selecting in manipulation is that it allows students to re-focus on that selected information, thereby paying more attention to it.

Although students do produce overt outputs, the outputs do not contain any new information beyond what was already presented in the instructional materials resulting from the physical behavior of manipulating that the *Active* mode affords. For example, the underlined sentences were already in the text, and the justification chosen from a menu was already provided.

The underlying knowledge-change processes associated with the *Active* mode can be surmised to be activating prior knowledge relevant to the selected and focused parts of the instructional materials, facilitating the encoding of new information to be connected with the activated parts of prior knowledge.

In short, the four criteria of *Active* engagement are: some physical manipulation of some selected aspects of the instructional materials, resulting in visible outputs that can lead to re-focusing, without producing any new information when compared with the given instructional materials. These four criteria are listed in [Table 1](#), under the *Active* column.

Constructive

The *Constructive* mode of cognitive engagement refers to overt behaviors in which learners generate or produce an output that goes beyond what is already provided in the instructional materials, such as a concept map, a diagram, or an explanation. Re-focusing is involved in that the output is generated from some selected aspect of the instructional materials. Finally, the output must contain additional information not included in the available course material or instructional content.

For example, suppose a student is reading about a historical account of an event in a history textbook ([Wineburg and Martin, 2004](#)). If a student starts to create a timeline in his or her notes that sequences the dated events mentioned in the course content (such as the year and actions taken by different historical figures discussed), the student would have produced something that was not present in the instructional materials, but all the information was already provided. This would be an example closer to a manipulative activity, because the student physically created a timeline, in much the same way as students plotting a graph from a provided data set. However, if the student, in addition, enhanced the timeline with some cause-and-effect relations because the sequencing of events allows her to see plausible antecedent causes, then this timeline activity is more generative, and she is thereby engaging in a *Constructive* mode. In contrast, if the textbook provided a diagrammed timeline of events and the student copied the timeline into their notes without making any changes, the student would only be engaging in an *Active* attending activity. Thus, producing an external output per se is not necessarily engaging in a generative activity (i.e., engaging in the *Constructive* mode), unless new information is present in the visible output.

The key knowledge-change process involved in the *Constructive* generating mode is inferring. That is, to generate knowledge that has not been presented, students need to focus on an aspect of the instructional materials and generate new knowledge from either prior knowledge, prior knowledge connected with new information, or new information per se. For example, the *Constructive* activity of self-explaining engages the inferring process ([Chi, in press](#)).

A caveat for the *Constructive* mode is that, even though it requires students to be generative beyond the course content or instructional materials, the *Constructive* mode does not require students to create novel or profound information. Rather, outputs produced in the *Constructive* generating mode only define new information in relation to the course materials. For example, a learner could generate a new idea by inferring a connection to an everyday example while reading a course textbook. The generated new idea is novel to the learner herself, but may not be novel to the content domain. For example, after reading about the septum in the heart in the context of a passage about the circulatory system, a student might wonder about whether the septum is solid or porous. If the student eventually concludes that the septum is solid, this information is new to the learner and not stated in the text passage, but

not novel to the topic of the human heart. Another caveat is that the new information the learner generated is most likely new information the learner induced, and not retrieved from her prior knowledge. See the *Constructive* column in [Table 1](#).

Interactive

We define the *Interactive* mode as overt collaborative dialoging among two peers or more. Although interactions can occur in other ways, dialoging or conversing is the most prevalent. The dialogs themselves can be considered overt outputs, such as explanations and feedback given to each other. In order to engage in the *Interactive mode*, each partner must also engage in the *Constructive* generating mode, bringing new information into the conversation; additionally, each partner must build on the other partner's contributions. Because both partners are *Constructive*, and moreover, because they each can build on the other's partner's contributions (i.e., being reciprocally co-generative), it is plausible that the outputs they jointly produce may contain ideas that neither partner could have produced alone, thus resulting in innovations. See [Table 1](#), the *Interactive* column.

One caveat about the *Interactive* mode is that the dialogs are more likely to be co-generative if a significant amount of turn-taking occurs. This is because more frequent turn-taking allows each partner to incorporate her partner's understanding of the domain into her own thinking and make more frequent adjustments to her own mental model ([Chi et al., 2017](#)). With few turns, essentially, the two peers are each giving mini-lectures, instead of building off of each other's contributions. Such reciprocally co-generative mutual exchanges of ideas can be manifested in several types of standard and conventional dialog patterns, such as defending or arguing a position in a debate ([Schwarz et al., 2000](#)), asking and answering each other's questions ([Webb, 1989](#)), and elaborating on each other's contributions, such as clarifying, building upon, correcting, and so forth ([Hogan et al., 1999](#)).

Criteria for discriminating between modes

Some criteria in [Table 1](#) are more salient than others for identifying the mode of an activity or strategy, or for discriminating among some of the modes. The *Passive* mode is the only one without visible outputs, so the availability of visible outputs is the most salient criterion. The critical criterion for the *Active* attending mode is whether the produced outputs contain some information that has already been presented in instruction, thereby allowing students to re-focus on this selected information. The critical criterion discriminating between an *Active* attending and a *Constructive* generating activity is whether the outputs from the activity contain information beyond what was provided in the instructional materials. This could be a minute inference or it could be a larger unit of knowledge. Finally, the critical criterion discriminating *Constructive* from *Interactive* is whether the new information co-generated in the *Interactive* mode could only be produced by both peers in a reciprocally co-generated way. This determination may be difficult to make, so one may have to rely on other behavioral features, to be explained in another section below.

The ICAP hypothesis

ICAP assumes a hierarchical relationship among the modes based on the behavioral definition of each mode. That is, each mode subsumes the other, such that reciprocally co-generating (the *Interactive* mode) requires students to also be generative (*Constructive*), and generating also requires students to be re-focusing and attending to information (*Active*). In addition, one could postulate the same hierarchical relationship among the four ICAP modes based on the underlying knowledge-change processes. In short, the ICAP hypothesis proposes that the *Interactive* mode can result in the highest level of learning, followed by the *Constructive* mode, followed by *Active*, followed by *Passive* as the lowest level of learning ([Chi, 2009](#); [Chi and Wylie, 2014](#)). Moreover, because only the *Constructive* and *Interactive* modes require inferring and going beyond the course materials, there is a bigger gap in learning between the two "higher" (*Constructive* and *Interactive*) modes compared to the two "lower" (*Active* and *Passive*) modes. We could assume that engaging in the two higher modes promotes deeper learning whereas engaging in the two lower modes promotes shallower learning. Thus, both the behaviors and the plausible underlying cognitive processes implicate the hierarchical relationships of the ICAP hypothesis.

There are two important caveats to note about the ICAP hypothesis. The first is that the deeper learning achievable in the *Constructive* and *Interactive* modes can only be detected if deeper assessment questions are designed to test students' understanding, such as questions that ask students to explain the cause of a historical event, when the cause was not stated, or to reason about the consequence of a historical event. Shallower question, those that typically ask students to recall some information, are perfectly adequate for detecting learning from the *Passive* and *Active* modes. The second caveat about the ICAP hypothesis is that it does not make predictions about learning outcomes for engagement within the same mode, yet there can be large differences in the actual processing of engagement within a single mode. For example, asking students to self-explain while reading a text is a *Constructive* activity, and so is the task of designing an experiment. However, self-explaining takes less processing efforts than designing an experiment, if we consider the amount of time as an indicator of effort, even though both are within the *Constructive* mode. Thus, ICAP's predictions of students' learning levels are reflective of learning levels between ICAP modes, and not within ICAP modes.

Using the ICAP theory to evaluate cognitive strategies

In this section, we use the operational definitions provided for each ICAP mode to determine which ICAP mode a learning strategy or activity likely belongs, as the mode it belongs to predicts its effectiveness for learning. In [Table 2](#), we identify 50 strategies that

Table 2 50 Strategies investigated in the literature.*Passive strategies (5)*

Salient criterion: any outputs produced?

1. Reading and re-reading (Rawson and Kintsch, 2005)
2. Observing, watching, and looking (Cocco et al., 2021)
3. Listening (Feng and Webb, 2020)
4. Tactile learning (Savaiano et al., 2016)
5. Turn-taking reading (since, as a whole, each individual student speaks aloud only for a short portion) (Duffy, 1983)

Active strategies (20)

Salient Criterion: any outputs produced that are present in the original instruction?

1. Clickers and polling (Arthurs and Kreager, 2017)
2. Copying (VanLehn et al., 2007)
3. Enacting: enacting how fast a ball moves or the slope of a line; mimicking; or show the slope of a line with arms (Fiorella and Mayer, 2015)
4. Gallery walk (Chin et al., 2015; Namaziandost et al., 2018).
5. Highlighting and underlining (a selected passage or equation), clip-and-paste (Ponce and Mayer, 2014; Bauer and Koedinger, 2007; Katayama et al., 2005)
6. Identifying a flaw (Friesem, 2019)
7. Listing (generating a list) from memory (Epstein, 1969; Roberts, 1972)
8. Mapping or aligning two sets of given information: this includes matching, choosing an option (Yin et al., 2005)
9. Measuring and other lab activities (Yaron et al., 2010; Tobin, 1990; Aladejana and Aderibigbe, 2007)
10. Ranking (Lombardi et al., 2021)
11. Reading a selected passage or sentence out loud [reading an entire passage would be passive], recite a poem (Pergams et al., 2018)
12. Recall (Endres et al., 2017)
13. Rehearsing (Oberauer, 2019)
14. Review (Cho and MacArthur, 2011)
15. Total physical response (e.g., having students stand or sit for answering true false, students walk to different corners of the room to represent different answer choices) (Asher, 1966, 1969)
16. Guided notes with fill-in-the-blank (Kotsopoulos et al., 2021; Konrad et al., 2009; Austin et al., 2002; Neef et al., 2006; Ponce and Mayer, 2014).
17. Summarizing by deleting irrelevant information (Reilly et al., 2017; Low and Over, 1989)
18. Following step by step instructions (Liefoghe et al., 2018)
19. Modeling, plotting or graphing (Balwaria and Bavishi, 2014; Doer and Zangor, 2000)
20. Testing and re-testing, interleaved testing (Rawson and Dunlosky, 2012; Roediger and Butler, 2011)

Constructive strategies (20).

Salient Criterion: any outputs produced containing information beyond the instructional materials?

1. Analogizing (Gentner et al., 2003; Chinn and Malhotra, 2002; Gentner and Hoyos, 2017; Hoyos and Gentner, 2017)
2. Comparing and contrasting: to derive similarities and differences (Schwartz et al., 2011; Hoyos and Gentner, 2017)
3. Critiquing or criticizing (Okada and Simon, 1997)
4. Justifying (Asterhan and Schwarz, 2016; Lin and Lehman, 1999; May, 2015)
5. Organizing (categorizing, putting in table format, putting events in a time line, etc.) (Donker et al., 2014)
6. Concept-mapping (Chang et al., 2002; Novak, 1990a,b)
7. Drawing, graphing, imagining, visualization (diagrams or other representations, or images, not already presented) (Fiorella and Zhang, 2018; Martin and Schwartz, 2014; Schwamborn et al., 2010)
8. Explaining to another (Roscoe and Chi, 2008; Biswas et al., 2005; Lombrozo, 2012; Coleman et al., 1997; Fiorella and Mayer, 2014; Legare and Lombrozo, 2014)
9. Integrating two or more sets of information or ideas to arrive at a third set; integrating two single or multi-media sources (Britt and Sommer, 2004; Rouet et al., 1996; Butcher, 2006; Bodemer et al., 2004; Mayer, 2005)
10. Posing questions to others, to self, or to authors (Beck et al., 1997; Morris and Chi, 2020; Mestre, 2002)
11. Predicting the outcome of a demonstration (Brod et al., 2018; Schauble et al., 1995)
12. Self-explaining and elaborating (Chi et al., 1989, 1994, Chi, in press)
13. Taking notes or paraphrasing in one's own words; writing an interpretive summary (Trafton and Trickett, 2001; Bretzing and Kulhavy, 1979); writing an interpretive review (Cranton, 2001; Doctorow et al., 1978).
14. Connecting theories or course concepts with everyday examples (Kothari et al., 2019; Rathburn, 2015)
15. Designing experiments in inquiry learning (Ariyawong and Phongsatha, 2017; Gao, 2015; Lombardi et al., 2021).
16. Hypothesizing explanations for a set of data (Hulsizer et al., 2018)
17. Modeling: aligning data with theories (Lombardi et al., 2021); deriving equations from data (VanLehn, 2013).
18. Reflecting and re-assessing one's own understanding to look for gap, coming up with the "muddiest" point (Akhtar and Saeed, 2020).
19. Solving challenging problems with real world scenarios, such as in project-based learning (Fortus et al., 2004)
20. Debugging and troubleshooting (Li et al., 2019; McCauley et al., 2008)

(Continued)

Table 2 50 Strategies investigated in the literature.—cont'd*Interactive strategies (5)*

Salient Criteria: both partners must be reciprocally co-generative and the outputs produced cannot be produced by a single partner (These two criteria are stringent and difficult to assess).

1. Asking and answering each other's questions (Rivers et al., 2017; Webb, 1989); or reciprocal questioning (King, 1990a,b); asking each other for justifications and explanations (Okada and Simon, 1997)
2. Debating with a partner, offering claims and justifications (Schwarz et al., 2000; Okada and Simon, 1997; Darby, 2007; Roy and Macchiette, 2005)
3. Improv theater (Rossing and Hoffmann-Longtin, 2016; Berk and Trieber, 2009; Hoffmann-Longtin et al., 2018); jazz ensemble (Larsson and Georgii-Hemming, 2019; Sawyer, 2008)
4. Peer tutoring (Beasley, 1997; Mastropieri et al., 2003)
5. Reciprocal learning or teaching (Whalon and Hanline, 2008; Palinscar and Brown, 1984)

have been investigated in the literature. For each mode (except the *Interactive* mode), we use its most salient criteria, based on the nature of the outputs (a combination of the second and fourth criterion listed in [Table 1](#)), to determine the ICAP mode of a strategy. These most salient criteria are listed under each mode in [Table 2](#).

Passive mode strategies

The most salient criterion for the *Passive* mode is whether students are producing any outputs while they are paying attention (we assume that students in all the modes are paying attention). We ignore situations in which students are not even paying attention, as that means students are not even engaged. Using these criteria, [Table 2](#) lists five learning strategies that can be determined to be *Passive*. That is, in silent reading, listening, watching and so forth, in which students are merely oriented toward instruction and not doing anything else nor producing any outputs.

Active mode strategies

For a strategy or activity to qualify as engaging students in the *Active* attending mode, students must be producing some physical manipulations, either with or without outputs, but the outputs are part of the original instructional materials. In [Tables 2](#), 20 learning strategies are listed that likely engage students in the *Active* attending mode. These 20 strategies seem obviously *Active* because students are actively manipulating instructional material, often resulting in re-focusing on a selected portion of the instructional information. From this list, total physical response, such as raising hands or stand up/down (*Active* Strategy #15), do not produce any outputs per se, but in responding, students are re-focusing on the questions asked. Clicking on clickers to choose an answer (*Active* Strategy #1), is also physically manipulative, but it also produces an output, the output is an option among the choices provided. For a third example, when students are asked to align two sets of given information via matching (*Active* strategy #8), they are overtly selecting information from one set and linking it to information in a second set by drawing a line or writing out the alignment while not producing any new information or ideas in the process. Similarly, when students are asked to fill out guided notes with fill-in-the-blanks (which typically require students to recall key terms to fill in the blanks), the key terms have already been presented during the lecture or within the instructional materials.

A caveat about these *Active* strategies is that, although they tend to elicit *manipulative* responses, students, at times, may go beyond the task by generating additional information. For example, while summarizing, a student may decide to paraphrase a portion of the paragraph for better clarity and flow, which would turn the *Active* task of summarizing into a generative one. As another example, students may improvise and fill in the blank with their own answers instead of following the prescribed lecture. But overall, if students perform these 20 strategies as designed and expected, then they are likely engaging in the *Active* mode.

Constructive mode strategies

As stipulated in our salient criterion in [Table 2](#), in the *Constructive* mode, the operational criterion is that students are typically generating external information not presented in the instructional materials, or that goes beyond the instructional materials. In [Table 2](#), 20 cognitive strategies are listed that likely engage students in the *Constructive* generating mode because some new knowledge beyond instruction is produced. For example, in analogizing, a new analogy or relations between analogies are produced; in comparing-and-contrasting, new similarities and differences are produced; in justifying, new reasons or justifications are produced, and in organizing, a new organization structure is produced.

A similar logic goes for assuming that strategies such as concept-mapping, drawing diagrams or graphs, explaining-to-another, integrating two sets of information, justifying, posing questions, predicting, self-explaining/elaborating, taking notes are all strategies that are clearly generative in the sense that students are creating some visible outputs or products containing information not already presented in instruction.

There are several caveats that are more relevant for strategies in the *Constructive* mode, although they can apply for the other modes as well. These caveats are addressed below.

Compliance

Just because an activity or strategy expects students to be generative does not guarantee that all students will comply by generating new information. The caveat is that students can instead produce information that does not go beyond instruction, thereby they are responding in the *Active* attending mode. For example, if students are drawing diagrams similar to ones already presented in the text, then they have essentially just copied the diagram and did not generate new information and, thus, are not responding in the *Constructive* generating mode. That is, students can respond in a higher mode (such as responding generatively to an *Active* elicitation, as pointed out in the caveat above for *Active* strategies) or respond in a lower mode (such as manipulatively to a *Constructive* elicitation), as the copy diagram example illustrates. The non-correspondence between the mode of students' responses and the mode elicited, can be determined by analyzing students' responses. However, by-and-large, instructors can only hope that the majority of students do respond in the mode a learning strategy elicits, as it is too difficult or too time-consuming to monitor compliance, either in real time (in a classroom context) or in the context of homework assignments. The important point is that, even without 100% compliance, strategies designed to elicit a generative response will enhance learning more so than strategies designed to elicit a manipulative response.

More or less complicated processing

Although the *Constructive* strategies listed in Table 2 all satisfy the criteria of generating outputs beyond instruction, they do differ substantially, depending on the amount of materials to which a strategy is applied or the amount of time needed to process the information (such as coming up with a solution). The amount of materials can vary within a strategy as well as between strategies. For example, take the strategy of integrating two or more sets of information (#9). The two sets of information can be two sentences, two paragraphs, or two pages. This difference will obviously show that integrating two pages might take more time and is more complicated than integrating two sentences. "More complicated" can be measured in a number of ways, such as the amount of instructional materials a strategy applies to or the amount of time it takes to complete. Thus, "more complicated," obviously means that they take longer to do (sometimes multiple days), thus requiring multiple processing steps, and typically encompass deeper understanding of a larger body of knowledge, in order to be able to carry it out. These strategies, include: coming up with the "muddiest" point of the class instruction (#18), hypothesizing for a set of data (#16), modeling or aligning data with theories (#17), writing an interpretive review (#13), and a variety of other complicated tasks, such as designing investigations and experiments (#15), often involved in project-based (#19) and other forms of robust inquiry learning (#15). A complicated engaging task, as described above, often requires multiple external steps, and numerous repetitive internal thinking processes, in contrast to engaging in simpler *Constructive* strategy on a smaller unit of knowledge, such as self-explaining one sentence, which requires a simple inferring process. The point of this caveat is to explain that ICAP cannot make accurate predictions of learning levels within a single mode, as the processes could vary considerably because of the variations in the time and the number of instructional materials involved. Instead, ICAP is only intended to make learning comparisons among modes.

Invisible thinking processes

In carrying out a *Constructive* strategy, it is possible for students not to display any outputs. For example, a learner can imagine or visualize a diagram without actually drawing it. However, for many of the strategies we described above, some externalized outputs are required, but this requirement does not preclude students from not complying and carrying out generative thinking processes covertly.

Generating vs. retrieval

Another caveat is that for some strategies, we cannot determine whether students' outputs actually generated new information or merely retrieved prior knowledge. Again, the liberal assumption is that for the majority of time, students in the *Constructive* mode are generating new knowledge rather than retrieving prior knowledge, as they are engaged in the context of learning new information.

Interactive mode strategies

To collaborate in the *Interactive* mode means that the interactions must be co-generative, meaning that each partner is being generative beyond the information provided in the instructional materials. In addition, to be reciprocally co-generative means that each partner must build on some selected aspect of the other partner's contributions. We can surmise that the five types of interactions mentioned in Table 2 for the *Interactive* mode are likely to meet these requirements. For example, in debating, each partner is generative in supplying a rationale or justification for defending his/her own point; moreover, each partner often builds on the other partner's points by refuting it. However, in order to be certain, one must analyze the process and content of the dialog patterns.

The process and content of a dialog pattern refers to how and what each partner contributes in a dialog turn. If we categorize the contribution of each partner in one turn of any hypothetical dialog, we can come up with 15 types of dialog patterns. For example, partner 1 can give an *Active* comment such as "I agree," "That's what I believe," and partner 2 can be *Constructive*, with comments

such as “What do you think might happen if we did this?” or “I suspect this might happen if ...”. This exchange would be an example of an *Active-Constructive* (A-C) dialog pattern. An A-A dialog pattern is one in which neither partner made any substantive comments that went beyond either what the instructional materials already provided or beyond what the other partner has contributed, but instead, merely re-described either what was already presented in instruction or what one’s partner said. A C-C dialog pattern involves both partners each *generating* beyond the information provided, but neither of the *Constructive* comments were generated upon the contribution of the partner. Using the four ICAP modes to describe each turn of a dialog gives rise to 15 type of dialog patterns (see Chi and Menekse, 2015). Thus, using ICAP modes to categorize each partner’s contribution shows clearly that only one out of 15 dialog patterns are truly co-generative. This suggests that it is difficult to obtain true *Interactive* collaboration in the ICAP sense of reciprocally co-generative.

Reciprocally co-generative means that each partner is selecting and building upon by inferring some aspect of the partner’s contributions, which may already have been inferred information beyond the information provided in instruction. Because each partner can come up with or infer some ideas that the other partner did not think of, these kinds of ideas are innovative, in the sense that these ideas could not have been generated by either of the partners alone. However, it would be extremely difficult to determine whether an output from collaboration was made jointly or whether one of the partners could have generated it alone.

It is therefore extremely difficult to ascertain, in practice, whether collaborators are interacting in the reciprocally co-generative way, without fine-grained analyses of their dialogs to see how an innovative output is produced. However, there are two options to assess in practice (such as in a classroom) whether collaborators are truly *Interactive*. First, one can rely on gross behaviors, such as sitting and facing each other, gazing at partner, talking to partner, taking turns, pointing and looking at where a partner points, and so forth. Second, instead of identifying dialog patterns by the contribution of each individual turn, we can identify three types of overarching non-co-generative, dialog patterns that we will refer to as: *cooperate*, *dominate*, and *parallel*. In order to understand these overarching collaboration patterns, we need to further introduce the notion of interaction, meaning whether one partner is having conversational (or other types of) exchanges with another partner. For example, the parallel collaboration pattern refers to two of our 15 dialog patterns, such as C-C or A-A. That is, both partners are working either *Constructively* or *Actively*, but they are not interacting in the sense of co-generating or building on the other’s contributions. In the other instance, they can each be *Active* by manipulating aspects of the instructional materials, such as when both students manipulate the lab equipment, but each student measures a different aspect of the lab equipment. A second overarching pattern is that the partners do interact, but one partner dominates the other partner. This usually means, in the literature, that one partner is doing all the work and all the talking, and basically takes control of the task. If the dominant partner is *Constructive* and the other partner merely supports and agrees with the dominant partner, the dialog pattern from an ICAP perspective is either C-A or C-P. The third overarching pattern is referred to as cooperation, which means that the two partners agree to divide a problem or a task up so each partner works on a separate component of the task or problem. In cooperation, there are no interactions between the partners, aside from the initial interactions to negotiate and agree upon what subcomponents to assign to each partner.

In short, although the *Interactive* mode has the greatest potential for promoting deep learning, it is difficult to achieve and ascertain that the students are collaborating *Interactively* in the reciprocally co-generative sense. However, one can assess in practice how *Interactive* partners are by their gross collaboration behaviors and by the three gross types of interacting patterns. In addition, one can anticipate more reciprocal co-generation by designing collaborative tasks that require generative outputs (to prevent parallel manipulation), that are not decomposable (to prevent cooperation), and that are sufficiently difficult that students cannot solve it alone (to prevent parallel generation or one partner dominating).

Evidence in support of ICAP hypothesis

Now that we have illustrated how the mode of a strategy can be determined based on the criteria listed in Table 2, we can verify whether ICAP’s predictions are true by studies that contrasted strategies from different ICAP modes. Because the majority of studies in the literature compare two strategies (that is, comparing a treatment condition with a control condition), we can make many pair-wise comparisons that are consistent with the ICAP hierarchy, such as comparing the *Interactive* mode with the *Constructive* (I vs. C), or comparing the *Constructive* with the *Passive* (C vs. P), with predictions that *Interactive* would be better or more effective for learning than C ($I > C$), and *Constructive* more beneficial for learning than *Passive* ($C > P$), and so forth. Such comparisons across all combinations of pair-wise comparisons have been reported in Chi (2009) and in Chi and Wylie (2014). However, in Chi and Wylie, we focused only on three strategies: self-explaining, concept-mapping, and note-taking. Here we will illustrate a few other examples by selecting studies that we have not described in prior papers, focusing on more difficult tasks at the college level, whether on learning or problem solving.

First, performing in the *Interactive* mode should be better than in a *Constructive* mode. In a study by Okada and Simon (1997), students were asked to solve a molecular genetic discovery task, so this was a *Constructive* activity because they had to derive a new solution that was not shown previously. Pairs of students succeeded better on solving the task (and focused more on alternative hypotheses and justifying their thought processes) than individual students working alone, suggesting that $I > C$, assuming that the individuals were trying to be generative.

Second, the *Interactive* mode should also be better than the *Active* mode for learning. In a reciprocal teaching context, one group of students practiced reciprocal question asking about a text (a generative activity), in addition to other *Constructive* activities, and the other group was locating information from a text (a manipulative activity). The reciprocal teaching group performed better in reading comprehension than the group that only practiced locating information in a text ($I > A$, Palinscar and Brown, 1984).

Third, undertaking a strategy in the *Constructive* mode should be better than a strategy in an *Active* mode. In a study by Britt and Sommer (2004), students were given two texts to read. The first group of students read two texts with instructions to focus on comparing and contrasting the texts and then answered some short-answer questions after each text, clearly a *Constructive* generating activity. A second control group just read the two texts, clearly a *Passive* attending activity. The results showed that the first group was better able to integrate the two texts than the second group ($C > P$, Britt and Sommer, 2004).

Fourth, a strategy in the *Active* mode should be clearly better than a strategy in the *Passive* mode. Many studies make such a contrast. For example, in a study by Kotsopoulos et al. (2021), in a calculus classroom, students in one group used guided notes to fill in the blanks for implications, connections, solutions, while the instructor was annotating or writing on a screen during class, so essentially students could copy the instructor's work, making this a manipulative strategy (*Active* mode). This group was compared to a group from a prior year, in which the instructor only lectured without providing any guided notes for students to fill out (*Passive* mode). The Guided-note group learned to solve calculus math problems better than the group without any guided notes ($A > P$).

Although the studies we cite typically compared two ICAP modes, there are a few studies in the literature that compared three conditions (Bretzing and Kulhavy, 1979; Gobert and Clement, 1999; Coleman et al., 1997) and one that compared all four ICAP modes (Menekse et al., 2013). The results of all four studies support the ICAP hierarchical predictions.

Conclusion: the value of ICAP

Although many strategies have been studied in the laboratory and all have been hailed as effective learning strategies, ICAP shows that these studied strategies are distinctly different in terms of their ICAP mode and thereby in terms of their effectiveness for learning. The *Active* strategies are good at enhancing recall and retention of relevant information, whereas the *Constructive* strategies are good at promoting deeper learning, with the *Interactive* strategies having the potential of creating innovations. The reason that so many studies in the literature have found so many strategies to be effective (Table 2 lists a sample of 50 strategies), is because their effectiveness was compared to a weaker control, meaning a control condition that utilizes an activity that is either in the *Passive* mode or a lower mode than the mode of the treatment condition. The four studies cited in the preceding section, as well as all the others cited in this chapter and earlier papers, all achieved a beneficial outcome of the treatment strategy because the treatment strategy was compared to another strategy of a lower ICAP mode. In order to claim that a strategy is absolutely effective, a study needs to compare it with another strategy at the same ICAP mode, as well as other strategies at the higher ICAP modes.

To illustrate this point with a popular conceptual phenomenon, it has often been shown that when children manipulate objects, such as toys, to show what the characters in a story event do (the treatment group), their recall of the story passage is better than children who only read the story without access to manipulating toys (the control group, Glenberg et al. 2004). This shows that manipulation is a beneficial *Active* strategy, assuming children do not further elaborate the stories with additional information. The control group clearly participated in a *Passive* only activity by just reading, so it is not surprising, based on ICAP's prediction, that enacting is better than reading only. With ICAP, we can now ask, are there much better *Constructive* strategies that can enhance reading, compared to enacting? Setting aside a consideration of the age-appropriateness of a strategy, one needs to ask whether examples such as posing questions (a *Constructive* strategy), either posing them to self (as in self-explaining, Chi, in press, or in reflection), posing them to others (Mestre, 2002), or posing them to the authors (Beck et al., 1997), would be a more effective learning strategy for learning than enacting? The same question can be asked about many other effective strategies. The point is that we can no longer claim in an absolute basis that a strategy is effective, without bracketing and comparing it with what other modes of strategies it is better than.

So, a theoretical value of the ICAP theory is that it can evaluate the effectiveness of a target strategy based on the mode it belongs, without needing to carry out an empirical study. It also implies caution in designing studies that compares a treatment with a *Passive* mode control condition. Being able to evaluate the relative effectiveness of strategies based on their mode also has the advantage of allowing instructors to know whether one strategy or activity can substitute for another. Finally, ICAP has practical value for informing instructors on how to design activities in various ICAP modes.

References

- Akhtar, M., Saeed, M., 2020. Effect of Frayer model, choral response and muddiest-point on students' academic achievement. *J. Educ. Sci. Res.* 7 (2), 71–86.
- Aladejana, F., Aderibigbe, O., 2007. Science laboratory environment and academic performance. *J. Sci. Educ. Technol.* 16 (6), 500–506.
- Ariyawong, O., Phongsatha, T., 2017. The student's self-designed learning activities influencing their learning achievement: a case study of an English tutorial center in Krabi. *ABAC ODI J. Vision. Action. Outcome* 4 (2), 28.
- Arthurs, L.A., Kreager, B.Z., 2017. An integrative review of in-class activities that enable active learning in college science classroom settings. *Int. J. Sci. Educ.* 39 (15), 2073–2091.
- Asher, J.J., 1966. The learning strategy of the total physical response: a review. *Mod. Lang. J.* 50 (2), 79–84.
- Asher, J.J., 1969. The total physical response approach to second language learning. *Mod. Lang. J.* 53 (1), 3–17.
- Asterhan, C.S., Schwarz, B.B., 2016. Argumentation for learning: well-trodden paths and unexplored territories. *Educ. Psychol.* 51 (2), 164–187.
- Austin, J.L., Lee, M.G., Thibeault, M.D., Carr, J.E., Bailey, J.S., 2002. Effects of guided notes on university students' responding and recall of information. *J. Behav. Educ.* 11 (4), 243–254.
- Balwaria, R.R., Bavishi, S., 2014. A study of graphical representation of data in social science of standard IX. *Online Int. Interdiscipl. Res. J.* 4 (3), 397–405.

- Bauer, A., Koedinger, K.R., 2007. Selection-based note-taking applications. In: Rossan, M.B., Gilmore, D. (Eds.), CHI 07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, NY, pp. 981–990.
- Beasley, C., 1997. Students as teachers: the benefits of peer tutoring. *Learn. Teach.* 21–30.
- Beck, I.L., McKeown, M.G., Hamilton, R.L., Kucan, L., 1997. Questioning the Author: An Approach to Enhancing Student Engagement With Text. International Reading Association, Newark, DE.
- Berk, R.A., Trieber, R.H., 2009. Whose classroom is it, anyway? Improvisation as a teaching tool. *J. Excel. Coll. Teach.* 20 (3), 29–60.
- Biswas, G., Leelawong, K., Schwartz, D., Vye, N., The Teachable Agents Group at Vanderbilt, 2005. Learning by teaching: a new agent paradigm for educational software. *Appl. Artif. Intell.* 19 (3–4), 363–392.
- Bloom, B.S., Krathwohl, D., 1956. Taxonomy of Educational Objectives: The Classification of Educational Goals, by a Committee of College and University Examiners. Handbook I: Cognitive Domain. Longman, Green, New York, NY.
- Bodemer, D., Ploetzner, R., Feuerlein, I., Spada, H., 2004. The active integration of information during learning with dynamic and interactive visualisations. *Learn. Instr.* 14 (3), 325–341.
- Bretzing, B.H., Kulhavy, R.W., 1979. Notetaking and depth of processing. *Contemp. Educ. Psychol.* 4, 145–153. [https://doi.org/10.1016/0361-476X\(79\)90069-9](https://doi.org/10.1016/0361-476X(79)90069-9).
- Britt, M.A., Sommer, J., 2004. Facilitating textual integration with macro-structure focusing task. *Read. Psychol.* 25 (4), 313–339.
- Brod, G., Hasselhorn, M., Bunge, S.A., 2018. When generating a prediction boosts learning: the element of surprise. *Learn. Instr.* 55, 22–31. <https://doi.org/10.1016/j.learninstruc.2018.01.013>.
- Brod, G., 2020. Generative learning: which strategies for what age? *Educ. Psychol. Rev.* 1–24.
- Butcher, K.R., 2006. Learning from text with diagrams: promoting mental model development and inference generation. *J. Educ. Psychol.* 98, 182–197.
- Chang, K.-E., Sung, Y.-T., Chen, I.-D., 2002. The effect of concept mapping to enhance text comprehension and summarization. *J. Exp. Educ.* 71, 5–23.
- Chi, M.T.H., Menekse, M., 2015. Dialogue patterns that promote learning. In: Resnick, L.B., Asterhan, C., Clark, S.N. (Eds.), *Socializing Intelligence Through Academic Talk and Dialogue*. AERA, Washington, DC, pp. 263–274.
- Chi, M.T.H., Wylie, R., 2014. ICAP: a hypothesis of differentiated learning effectiveness for four modes of engagement activities. *Educ. Psychol.* 1–29.
- Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P., Glaser, R., 1989. Self-explanations: how students study and use examples in learning to solve problems. *Cognit. Sci.* 13, 145–182.
- Chi, M.T.H., Deeeuw, N., Chiu, M.H., Lavacher, C., 1994. Eliciting self-explanations improves understanding. *Cognit. Sci.* 18 (3), 439–477.
- Chi, M.T.H., Roy, M., Hausmann, R.G.M., 2008. Observing tutorial dialogues collaboratively: insights about human tutoring effectiveness from vicarious learning. *Cognit. Sci.* 32, 301–341.
- Chi, M.T.H., Kang, S., Yaghmourian, D.L., 2017. Why students learn more from dialogue-than monologue-videos: analyses of peer interactions. *J. Learn. Sci.* 26 (1), 10–50.
- Chi, M.T.H., Adams, J., Bogusch, E.B., Bruchok, C., Kang, S., Lancaster, M., et al., 2018. Translating a theory about modes of cognitive engagement to practice: teachers implementing ICAP. *Cognit. Sci.* 42 (6), 1777–1832. <https://doi.org/10.1111/cogs.12626>.
- Chi, M.T.H., 2009. Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Top. Cognit. Sci.* 1, 73–105.
- Chi, M.T.Y., 2021. Translating a theory of active learning: an attempt to close the research-practice gap in education. *Top. Cognit. Sci.* 1–23.
- Chi, M.T.H., (in press). The self-explanation principle in multimedia learning. In: Mayer, R.E., Fiorella, L. (Eds.), *The Cambridge Handbook of Multimedia Learning*. third ed. Cambridge University Press.
- Chin, C.K., Khor, K.H., Teh, T.K., 2015. Is gallery walk an effective teaching and learning strategy for biology?. In: *Biology Education and Research in a Changing Planet*. Springer, Singapore, pp. 55–59.
- Chinn, C.A., Malhotra, B.A., 2002. Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Sci. Educ.* 2 (86), 175–218.
- Cho, K., MacArthur, C., 2011. Learning by reviewing. *J. Educ. Psychol.* 103 (1), 73–84.
- Cocco, V.M., Bisagno, E., Di Bernardo, G.A., Cadamuro, A., Riboldi, S.D., Crapolicchio, E., Trifiletti, E., Stathi, S., Vezzali, L., 2021. Comparing story reading and video watching as two distinct forms of vicarious contact: an experimental intervention among elementary school children. *Br. J. Soc. Psychol.* 60 (1), 74–94.
- Coleman, E.B., Brown, A.L., Rivkin, I.D., 1997. The effect of instructional explanations on learning from scientific texts. *J. Learn. Sci.* 6 (4), 347–365.
- Conati, C., VanLehn, K., 2000. Toward computer-based support of meta-cognitive skills: a computational framework to coach self-explanation. *Int. J. Artif. Intell. Educ.* 11, 389–415.
- Cranton, P., 2001. Interpretive and critical evaluation. *N. Dir. Teach. Learn.* 2001 (88), 11–18.
- Darby, M., 2007. Debate: a teaching-learning strategy for developing competence in communication and critical thinking. *J. Dent. Hyg.* 18 (4).
- Dick, W., Carey, S., 2005. *The Systematic Design of Instruction*, sixth ed. Allyn and Bacon, New York City, NY.
- Doctorow, M., Wittrock, M.C., Marks, C., 1978. Generative processes in reading comprehension. *J. Educ. Psychol.* 70 (2), 109.
- Doer, H.M., Zangor, R., 2000. Creating meaning for and with graphic calculator. *Educ. Stud. Math.* 41, 143–163.
- Donker, A.S., de Boer, H., Kostons, D., Dignath van Ewijk, C.C., van der Werf, M.P.C., 2014. Effectiveness of learning strategy instruction on academic performance: a meta-analysis. *Educ. Res. Rev.* 11, 1–26.
- Duffy, G.G., 1983. First in series-factor related to reading performance from turn taking to sense making broadening the concept of reading teacher effectiveness. *J. Educ. Res.* 76 (3), 134–139.
- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.T., Willingham, D.T., 2013. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Publ. Interest* 14 (1), 4–58.
- Endres, T., Carpenter, S., Martin, A., Renkl, A., 2017. Enhancing learning by retrieval: enriching free recall with elaborative prompting. *Learn. Instr.* 49, 13–20.
- Epstein, W., 1969. Recall of word lists following learning of sentences and of anomalous and random strings. *J. Verb. Learn. Verb. Behav.* 8 (1), 20–25.
- Feng, Y., Webb, S., 2020. Learning vocabulary through reading, listening, and viewing: which mode of input is most effective? *Stud. Sec. Lang. Acquis.* 42 (3), 499–523.
- Fiorella, L., Mayer, R.E., 2014. Role of expectations and explanations in learning by teaching. *Contemp. Educ. Psychol.* 39 (2), 75–85.
- Fiorella, L., Mayer, R.E., 2015. *Learning as a Generative Activity: Eight Learning Strategies That Promote Understanding*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107707085>.
- Fiorella, L., Mayer, R.E., 2016. Eight ways to promote generative learning. *Educ. Psychol. Rev.* 28 (4), 717–741.
- Fiorella, L., Zhang, Q., 2018. Drawing boundary conditions for learning by drawing. *Educ. Psychol. Rev.* 30 (3), 1115–1137.
- Fonseca, B., Chi, M.T.H., 2011. The self-explanation effect: a constructive learning activity. In: Mayer, R., Alexander, P. (Eds.), *The Handbook of Research on Learning and Instruction*. Routledge Press, New York, pp. 270–321.
- Fortus, D., Dershimer, R.C., Krajcik, J., Marx, R.W., Mamlok-Naaman, R., 2004. Design-based science and student learning. *J. Res. Sci. Teach.* 41 (10), 1081–1110.
- Friesem, Y., 2019. Teaching truth, lies, and accuracy in the digital age: media literacy as project-based learning. *J. Mass Commun. Educ.* 74 (2). Retrieved from. <https://doi.org/10.1177/1077695819829962>.
- Gao, R., 2015. Incorporating students' self-designed, research-based analytical chemistry projects into the instrumentation curriculum. *J. Chem. Educ.* 92 (3), 444–449.
- Gentner, D., Hoyos, C., 2017. Analogy and abstraction. *Top. Cognit. Sci.* 9 (3), 672–693.
- Gentner, D., Loewenstein, J., Thompson, L., 2003. Learning and transfer: a general role for analogical encoding. *J. Educ. Psychol.* 95 (2), 393–408.
- Glenberg, A.M., Gutierrez, T., Levin, J.R., Japuntich, S., Kaschak, M.P., 2004. Activity and imagined activity can enhance young children's reading comprehension. *J. Educ. Psychol.* 96 (3), 424.

- Gobert, J.D., Clement, J.J., 1999. Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *J. Res. Sci. Teach.* 36, 39–53.
- Hattie, J.A., 1996. *Visible Learning: A Synthesis of 800+ Meta-Analysis on Achievement*. Routledge, Abingdon.
- Hoffmann-Longtin, K., Rossing, J.P., Weinstein, E., 2018. Twelve tips for using applied improvisation in medical education. *Med. Teach.* 40 (4), 351–356.
- Hogan, K., Nastasi, B.K., Pressley, M., 1999. Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognit. Instr.* 17, 379–432.
- Hoyos, C., Gentner, D., 2017. Generating explanations via analogical comparison. *Psychon. Bull. Rev.* 24 (5), 1364–1374.
- Hulsizer, H., Huq, A., Wei, W., 2018. Early introduction of hypothesis testing in introductory statistics: a pilot study. *Online J. New Horiz. Educ.* 8 (3), 10–19.
- Katayama, A.D., Shambaugh, R.N., Doctor, T., 2005. Promoting knowledge transfer with electronic note taking. *Teach. Psychol.* 32, 129–131.
- King, A., 1990a. Enhancing peer interaction and learning in the classroom through reciprocal questioning. *Am. Educ. Res. J.* 27 (4), 664–687.
- King, A., 1990b. Reciprocal peer-questioning a strategy for teaching students how to learn from lectures. *Clear. House* 64 (2), 131–135.
- Konrad, M., Joseph, L.M., Eveleigh, E., 2009. A meta-analytic review of guided notes. *Educ. Treat. Child.* 421–444.
- Kothari, D., Hall, A.O., Castañeda, C.A., McNeil, A.J., 2019. Connecting organic chemistry concepts with real-world contexts by creating infographics. *J. Chem. Educ.* 96 (11), 2524–2527.
- Kotsopoulos, D., Weatherby, C., Woolford, D.G., 2021. Using guided notes to support learning in first-year calculus. *Int. J. Math. Educ. Sci. Technol.* 1–16. <https://doi.org/10.1080/0020739X.2021.1910742>.
- Krathwohl, D., 2002. A revision of Bloom's taxonomy: an overview. *Theor. Pract.* 41 (4), 212–218.
- Larsson, C., Georgii-Hemming, E., 2019. Improvisation in general music education—a literature review. *Br. J. Music Educ.* 36 (1), 49–67.
- Legare, C.H., Lombrozo, T., 2014. Selective effects of explanation on learning during early childhood. *J. Exp. Child Psychol.* 126, 198–212.
- Li, C., Chan, E., Denny, P., Luxton-Reilly, A., Tempero, E., 2019. Towards a framework for teaching debugging. In: Simon, Luxton-Reilly, A. (Eds.), *Proceedings of the Twenty-First Australasian Computing Education Conference*. Association for Computing Machinery, New York, NY, pp. 79–86.
- Liefooghe, B., Braem, S., Meiran, N., 2018. The implications and applications of learning via instructions. *Acta Psychol.* 184, 1–3.
- Lin, X., Lehman, J.D., 1999. Supporting learning of variable control in a computer-based biology environment: effects of prompting college students to reflect on their own thinking. *J. Res. Sci. Teach.* 36 (7), 837–858.
- Lombardi, D., Shipley, T.F., Astronomy Team, Biology Team, Chemistry Team, Engineering Team, Geography Team, Geoscience Team, Physics Team, 2021. The curious construct of active learning. *Psychol. Sci. Publ. Interest* 22 (1), 8–43.
- Lombrozo, T., 2012. Explanation and abductive inference. In: Holyoak, K.J., Morrison, R.G. (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, pp. 260–276.
- Low, R., Over, R., 1989. Detection of missing and irrelevant information within algebraic story problems. *Br. J. Educ. Psychol.* 59 (3), 296–305.
- Martin, L., Schwartz, D.L., 2014. A pragmatic perspective on visual representation and creative thinking. *Vis. Stud.* 29, 80–93.
- Mastropieri, M.A., Scruggs, T.E., Spencer, V., Fontana, J., 2003. Promoting success in high school world history: peer tutoring versus guided notes. *Learn. Disabil. Res. Pract.* 18, 52–65.
- May, B., 2015. *An Introduction to the Justification Principle and Its Associated Benefits and Challenges Within the Mathematics Classroom*. Master of Teaching Research Project. University of Toronto, Canada.
- Mayer, R.E., 2005. *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.
- McCauley, R., Fitzgerald, S., Lewandowski, G., Murphy, L., Simon, B., Thomas, L., Zander, C., 2008. Debugging: a review of the literature from an educational perspective. *Comput. Sci. Educ.* 18 (2), 67–92.
- Menekse, M., Stump, G., Krause, S., Chi, M.T.H., 2013. Differentiated overt learning activities for effective instruction in engineering classrooms. *J. Eng. Educ.* 102, 346–374.
- Mestre, J.P., 2002. Probing adults' conceptual understanding and transfer of learning via problem posing. *J. Appl. Dev. Psychol.* 23, 9–50.
- Morris, J., Chi, M.T.H., 2020. Improving teacher questioning in science using ICAP theory. *J. Educ. Res.* <https://doi.org/10.1080/00220671.2019.1709401>.
- Namaziandost, E., Esfahani, F.R., Nasri, M., Mirshekaran, R., 2018. The effect of gallery walk technique on pre-intermediate EFL learners' speaking skill. *Lang. Teach. Res. Q.* 8, 1–15.
- Neef, N.A., McCord, B.E., Ferreri, S.J., 2006. Effects of guided notes versus completed notes during lectures on college students' quiz performance. *J. Appl. Behav. Anal.* 39 (1), 123–130.
- Novak, J.D., 1990a. Concept mapping: a useful tool for science education. *J. Res. Sci. Teach.* 27 (10), 937–949.
- Novak, J.D., 1990b. Concept maps and Vee diagrams: two metacognitive tools to facilitate meaningful learning. *Instr. Sci.* 19 (1), 29–52.
- Oberauer, K., 2019. Is rehearsing an effective maintenance strategy for working memory? *Trends Cognit. Sci.* 23 (9), 798–809.
- Okada, T., Simon, H.A., 1997. Collaborative discovery in a scientific domain. *Cognit. Sci.* 21, 109–146.
- Palinscar, A.S., Brown, A.L., 1984. Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognit. Instr.* 1 (2), 117–175.
- Pergams, O.R.W., Jake-Mathews, C.E., Mohanty, L.M., 2018. A combined read-aloud think-aloud strategy improves student learning experiences in college-level biology courses. *J. Coll. Sci. Teach.* 47 (5), 10–15.
- Ponce, H.R., Mayer, R.E., 2014. Qualitatively different cognitive processing during online reading primed by different study activities. *Comput. Hum. Behav.* 30, 121–130. <https://doi.org/10.1016/j.chb.2013.07.054>.
- Rathburn, M.K., 2015. Building connections through contextualized learning in an undergraduate course on scientific and mathematical literacy. *Int. J. Scholarsh. Teach. Learn.* 9 (1).
- Rawson, K.A., Dunlosky, J., 2012. When is practice testing most effective for improving the durability and efficiency of student learning? *Educ. Psychol. Rev.* 24 (3), 419–435. <https://doi.org/10.1007/s10648-012-9203-1>.
- Rawson, K.A., Kintsch, W., 2005. Rereading effects depend on time of test. *J. Educ. Psychol.* 97 (1), 70.
- Reilly, T.R., Haynes, J.M., McCarley, N.G., Williams, J.L., Whetzel, T.A., Bailey, L., Sharpe, J.L., Phillips, J., 2017. Using irrelevant information to aid the identification process. *N. Am. J. Psychol.* 19 (3), 731–744.
- Rivers, J., Smith, A., Higgins, D., Mills, R., Maier, A.G., Howitt, S.M., 2017. Asking and answering questions: partners, peer learning, and participation. *Int. J. Stud. Partn.* 1 (1).
- Roberts, W.A., 1972. Free recall of word lists varying in length and rate of presentation: a test of total-time hypotheses. *J. Exp. Psychol.* 92 (3), 365–372.
- Roediger, H.L., Butler, A.C., 2011. The critical role of retrieval practice in long-term retention. *Trends Cognit. Sci.* 15 (1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>.
- Roscoe, R.D., Chi, M.T., 2008. Tutor learning: the role of explaining and responding to questions. *Instr. Sci.* 36 (4), 321–350.
- Rossing, J.P., Hoffmann-Longtin, K., 2016. Improv(ing) the academy: applied improvisation as a strategy of educational development. *To Improv. Acad.* 35 (2), 303–325.
- Rouet, J.F., Britt, M.A., Mason, R.A., Perfetti, C.A., 1996. Using multiple sources of evidence to reason about history. *J. Educ. Psychol.* 88 (3), 478.
- Roy, A., Macchiette, B., 2005. Debating the issues: a tool for augmenting critical thinking skills of marketing students. *J. Market. Educ.* 27 (3), 264–276.
- Savaiano, M.E., Compton, D.L., Hatton, D.D., Lloyd, B.P., 2016. Vocabulary word instruction for students who read braille. *Except. Child.* 82 (3), 337–353.
- Sawyer, R.K., 2008. Learning music from collaboration. *Int. J. Educ. Res.* 47, 50–59.
- Schauble, L., Glaser, R., Duschl, R.A., Schulze, S., John, J., 1995. Students' understanding of the objectives and procedures of experimentation in the science classroom. *J. Learn. Sci.* 4 (2), 131–166.
- Schwaborn, A., Mayer, R.E., Thillmann, H., Leopold, C., Leutner, D., 2010. Drawing as a generative activity and drawing as a prognostic activity. *J. Educ. Psychol.* 102 (4), 872–879. <https://doi.org/10.1037/a0019640>.

- Schwartz, D.L., Chase, C.C., Oppezzo, M.A., Chin, D.B., 2011. Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *J. Educ. Psychol.* 103 (4), 759.
- Schwarz, B.B., Neuman, Y., Biezuner, S., 2000. Two wrongs may make a right ... if they argue together! *Cognit. Instr.* 18, 461–494.
- Tobin, K., 1990. Research on science laboratory activities: in pursuit of better questions and answers to improve learning. *Sch. Sci. Math.* 90 (5), 403–418.
- Trafton, J.G., Trickett, S.B., 2001. Note-taking for self-explanation and problem solving. *Hum. Comput. Interact.* 16, 1–38.
- VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., Rosé, C.P., 2007. When are tutorial dialogues more effective than reading? *Cognit. Sci.* 31 (1), 3–62.
- VanLehn, K., 2013. Model construction as a learning activity: a design space and review. *Interact. Learn. Environ.* 21 (4), 371–413.
- Webb, N.M., 1989. Peer interaction, problem solving, and cognition: multidisciplinary perspectives. *Int. J. Educ. Res.* 13, 1–119.
- Whalon, K., Hanline, M.F., 2008. Effects of a reciprocal questioning intervention on the question generation and responding of children with autism spectrum disorder. *Educ. Train. Dev. Disabil.* 367–387.
- Wineburg, S., Martin, D., 2004. Reading and rewriting history. *Educ. Leader.* 62, 42–45.
- Yaron, D., Karabinos, M., Lange, D., Greeno, J.G., Leinhardt, G., 2010. The ChemCollective—virtual labs for introductory chemistry courses. *Science* 328 (5978), 584–585.
- Yin, Y., Vanides, J., Ruiz-Primo, M.A., Ayala, C.C., Shavelson, R.J., 2005. Comparison of two concept-mapping techniques: implications for scoring, interpretation, and use. *J. Res. Sci. Teach.* 42 (2), 166–184.

Student-centered instruction: inquiry-, problem-, project-, and case-based learning

Sofie M.M. Loyens^{a,b}, Lisette Wijnia^c, and Remy M.J. P. Rikers^a, ^aUniversity College Roosevelt, Utrecht University, Utrecht, the Netherlands; ^bDepartment of Education, Psychology, and Child Studies (DPECS), Erasmus University Rotterdam, Rotterdam, the Netherlands; and ^cOpen University of the Netherlands, Heerlen, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	701
Inquiry-based learning (IBL)	702
Design of the learning question in IBL	703
The level of student/teacher centeredness in IBL	703
Sequence of information acquisition	703
Problem-based learning (PBL)	703
Design of the learning task in PBL	704
Building on prior knowledge	704
Eliciting discussion	704
Stimulating self-directed learning	705
Encouraging knowledge integration and transfer	705
Relevance for future profession	705
The level of student/teacher centeredness in PBL	705
Sequence of information acquisition in PBL	706
Project-based learning (PjBL)	706
Design of the learning task in PjBL	706
The level of student/teacher centeredness in PjBL	707
Sequence of information acquisition in PjBL	707
Case-based learning (CBL)	707
Design of the learning task in CBL	707
The level of student/teacher centeredness in CBL	708
Sequence of information acquisition in CBL	708
Acquaintances of the student-centered family: team-based and challenge-based learning	709
Team-based learning	709
Challenge-based learning (ChBL)	709
Directions for future research	709
References	710

Introduction

Student-centered instruction or learning pedagogies promote teaching practices that stem from a constructivist view of learning. Although constructivism is an ill-defined concept, this view on learning starts from the idea that the learner is an active agent in the learning process, constructing one's own meaning while relating new information to prior knowledge. In other words, there is an emphasis on the construction of knowledge. Often, this knowledge construction process happens in a setting that involves peer-to-peer discussion and social interaction, although constructivist researchers acknowledge the importance of collaboration to a varying degree (Dunbar and Yadav, 2022; Loyens and Rikers, 2017). In addition, goal setting, plan making, and monitoring one's learning process (i.e., self-directed learning), are important foci of constructivist learning views (Loyens et al., 2008). Students need to "own" their learning processes and they can do this when they determine and accomplish learning goals. This process happens with the support of teachers as it is a common misconception that students will perform independently without external guidance. Instead, students should be scaffolded through the process from owning the learning question, problem or project, researching it, and sharing its outcomes. By varying the locus of perceived control, students maintain personal responsibility for learning (Hannafin et al., 2014; Lee and Hannafin, 2016). Finally, constructivist learning approaches usually use problems, cases, projects, and questions that are ill-structured (i.e., without a clear path to the solution or without a solution at all) and "authentic" (i.e., resemble real-life situations) in order to promote the application and transfer of knowledge (Loyens et al., 2021). A description of somebody who hurt herself and feels a throbbing pain around the wound a few days later can stimulate students to discuss topics such as the natural healing process of a wound, infection, and characteristics of inflammation (Dolmans et al., 1997). These single processes and features become more meaningful in the context of a story and it is believed that students construct personal meaning through working on these problems, questions, and projects (Hannafin et al., 2014).

A constructivist view of learning provides a common ground for student-centered instruction, as the principles of knowledge construction (i.e., connecting to students' prior knowledge), collaborative learning, autonomy, and working with realistic learning assignments (i.e., problems, projects, questions, cases) can be found in various forms of student-centered instruction (Loyens et al., 2021). In all formats, teachers create room for students to formulate and explore their answers to the posed question instead of providing students with a question and answer. This approach has crystallized into several teaching strategies that place an emphasis on the learner as an active agent in his or her learning process instead of the learner as a passive receiver of information. These instructional formats have been labeled 'student-centered' and employ classroom practices such as observations, generating questions, discovering gaps in one's knowledge base, and studying resources to try to overcome these gaps (Loyens and Rikers, 2017).

However, differences exist among several student-centered instructional methods. Barrows (1986) identified three variables that can vary in different student-centered approaches: the design and format of the problem, project, question or case; the degree to which learning is teacher centered or learning centered; and the sequence in which problems or tasks are offered and information is acquired. These varying factors have led to a variety of approaches, often lacking conceptual clarity, which has posed problems for practitioners and researchers in higher education (Gallagher and Savage, 2020). After all, effectiveness studies become difficult without a clear consensus on definitions and shared as well as different characteristics, as the risk for comparing apples and oranges becomes substantial. Given additional, more methodological issues associated with effectiveness research (e.g., lack of control group and lack of information on fidelity to the instructional approach; Granger et al., 2012), providing conceptual clarity seems to be a first step in the desired direction.

The present chapter aims to shed light on different student-centered approaches commonly found in the research literature. We successively discuss inquiry-based learning, problem-based learning, project-based learning, and case-based learning. Each of the instructional approaches is discussed in the light of the factors that Barrows (1986) put forward: 1) the design of the problem, project, case, question, 2) the level of student/teacher centeredness, and 3) the sequence in which information is acquired. In addition, we briefly describe two relatively newer methods that have emerged, team-based learning and challenge-based learning. These methods combine elements of specific student-centered formats and general beneficial learning strategies.

Inquiry-based learning (IBL)

Instruction based on inquiry emphasizes classroom practices that involve active participation of the learner as well as a prominent role for questions and issues generated by the learner. Generally stated, inquiry is based on a scientific investigation through classroom practices such as posing questions and it is aimed at knowledge acquisition and development (e.g., Blanchard et al., 2008). Because of this broad definition, IBL is often seen as an umbrella term for several student-centered classroom practices such as observations, generating questions, discussion, and self-study in the context of meaningful tasks (Loyens and Rikers, 2017; Tawfik et al., 2020). However, other researchers have described inquiry-based learning as a process related to scientific inquiry. In that case, IBL entails iterative cycles of hypothesis generation, experimentation (i.e., designing experiments suitable for generating or testing hypotheses), and evidence evaluation (i.e., critical assessment of the experimentation results) which lead to acceptance, rejection or revision of hypotheses (e.g., Lazonder and Harmsen, 2016). Indeed, Pedaste et al. (2015) define IBL as "a process of discovering new causal relations, with the learner formulating hypotheses and testing them by conducting experiments and/or making observations" (p. 47). Within this definition, inquiry is closely linked to reasoning and critical thinking and the empirical circle of scientific research, also referred as the *Scientific Inquiry Cycle* (SIC, Klahr and Dunbar, 1988; Lazonder and Harmsen, 2016). The SIC has been put forward as the core element in scientific reasoning (Schiefer et al., 2019). It is therefore not surprising that IBL is popular in science education (Loyens and Rikers, 2017). In a similar vein, IBL is often viewed as an approach to solving problems and involves the application of several problem-solving skills. Table 1 provides an overview of the different characteristics of IBL.

Table 1 IBL analyzed according to Barrows' (1986) factors.

Factor	Inquiry-based learning
1. Design of learning task	(Research) question E.g. "is sustainable agriculture good for the environment?"
2. Teacher versus student centered	Teacher = facilitator, providing guidance, having expert knowledge
3. Sequence of information acquisition	1) hypothesis generation 2) experimentation: Designing experiments suitable for generating or testing hypotheses 3) evidence evaluation: Critical assessment of the experimentation results, which lead to acceptance, rejection, or revision of hypotheses 4) scientific inquiry circle (SIC) can start again

Design of the learning question in IBL

In IBL, pupils formulate their own research question and design and conduct an investigation to answer the question (Windschitl, 2003). It can therefore be seen as a bottom-up approach (Van Uum et al., 2021). Pedaste et al. (2015) called the generation of research questions the “conceptualization” phase, which comes right after the initial “orientation” phase. While the orientation phase is focused on getting the learner started with a topic for investigation (i.e., exploring or observing a phenomenon to trigger interest and reading some theory in order to know the scientifically oriented questions related to the particular phenomenon), the process of generating research questions based on the stated problem constitutes the core of the conceptualization phase. The assignment of the learning task in IBL consists of activities such as ‘ask’, ‘question’, ‘ask questions’, ‘developing a question’, ‘initial inquiry question’, ‘generating a scientific question’, ‘set up of inquiry question’, ‘raising and revising questions’, and ‘decide my inquiry question’ (Pedaste et al., 2015, p. 52).

The level of student/teacher centeredness in IBL

The fact that pupils formulate their own research question, and design and conduct an investigation to answer the question, does not imply that teachers only observe their pupils. On the contrary, they have an important role in facilitating inquiry-based learning and scaffolding the knowledge and skills of their pupils (Van Uum et al., 2016, 2017). Indeed, Grant and Hill (2006) identified several factors influencing a student-centered learning environment such as IBL. One of these factors is the recognition and acceptance of new roles and responsibilities. This implies the acknowledgment that the teacher’s role is de-centered and hence becomes a facilitator of learning and a partner in learning with students. In addition, these authors mention that student-centered learning environments also demand flexibility from teachers to deal with potential ambiguity in students’ learning processes in IBL. Grant and Hill (2006) cited examples in this respect such as “The teacher is not the sole source of knowledge.” “Many lines of inquiry may be underway by students”, and “Student progress may be more difficult to discern” (p. 18). These examples clearly require a different role of the teacher. A recent study by Murphy et al. (2021) showed that some teachers appeared to see IBL as student-centered independent discovery, which created tension between loss of teacher control and the impact on student learning. The results of this study demonstrated that IBL was perceived as a challenge to the authority of the teacher. However, on the other hand, it was also acknowledged that the teacher’s role in managing dialogue in the classroom is paramount (Murphy et al., 2021). Indeed, the role of the guidance students receive during the IBL process has been directly linked to the effectiveness of IBL with “a reasonable degree of guidance being more effective than low guidance” (Lazonder and Harmsen, 2016, p. 683).

Sequence of information acquisition

As can be seen in Table 1, IBL follows a specific sequence of steps in the learning process. Generally speaking, the process starts with a question, formulation of hypothesis, experimentation, and evidence evaluation based on whether the hypothesis is rejected or accepted. Some researchers have added steps to this sequence of information acquisition such as the orientation phase discussed above (Pedaste et al., 2015). Some examples of IBL tasks were explained by Aditomo et al. (2013). In one of the IBL activities they described in the field of history, students had to decide on an individual research topic in local history, from which they developed research questions or aims, with assistance from the teacher to keep the research topic manageable and to find relevant resources. Based on both primary and secondary literature resources, they wrote a short history on their research project. An example in the domain of science (e.g., biomedicine) was that students were given three sets of DNA sequences and needed to use the computational tools and databases they had learnt about to discover the respective functions of the proteins. They were not told which tools to use, but the teacher did provide scaffolding if needed. Finally, an example put forward in the field of management was that students had to interview a family member, a flat-mate, a friend, or a colleague or boss using a particular theory of motivation (i.e., goal-setting, expectancy or equity theories), and analyze their current levels of motivation. Using the same theory, students had to identify areas for improving their motivation (Aditomo et al., 2013). In all examples, the starting point was a (research) question such as “By which theory could one’s current level of motivation best be explained and how could it be improved?”, followed by hypothesis generation, experimentation, and evidence evaluation.

Problem-based learning (PBL)

The central feature of PBL is that the learning process starts with a complex, ill-structured problem that describes one or more observable phenomena or events (Schmidt, 1983). As is true for all formats of student-centered instruction, a plethora of definitions and formats of PBL exist. Nevertheless, Barrows (1996) defined six core characteristics of PBL. The first characteristic is that learning needs to be student-centered. Second, learning has to occur in small student groups under the guidance of a teacher, a so-called tutor. The third characteristic refers to the tutor as a facilitator or guide instead of a provider of knowledge. Fourth, so-called “authentic” (i.e., relating to real life) problems are primarily encountered in the learning sequence before any preparation or study has occurred. Fifth, the problems are used as a trigger for students’ prior knowledge before any other curriculum input, which leads to the discovery of knowledge gaps. Finally, these knowledge gaps are overcome through self-directed learning, which requires sufficient time for self-study (Schmidt et al., 2009).

Besides the core characteristics, shared characteristics can also be found in terms of the process, which always starts with a problem that is discussed in a group meeting. Issues that need further clarification and study are collaboratively agreed on and formulated in the form of questions (i.e., learning issues). Next, students individually try to find answers to these learning issues during a period of self-directed, individual study that usually lasts a couple of days. Finally, students return to their groups and discuss the answers found (Wijnia et al., 2019), critically evaluating all students' findings and weighing the evidence found for their answers to the learning issues (Loyens et al., 2022). Table 2 provides an overview of the different characteristics of PBL.

Design of the learning task in PBL

As mentioned earlier, the problem is the first input that PBL students encounter. On some occasions (e.g., in medical education), these problems originate from professional practice; in other cases, they tackle problems or events typical for a particular domain of study (Barrows, 1996; Norman and Schmidt, 1992; Schmidt et al., 2007). Since these problems need to be understood in terms of their underlying theoretical explanations, they need to have several characteristics that ensure sufficient scaffolding for the brainstorm, the formulation of learning issues, and students' self-study activities. Several researchers have proposed central features for problem construction in PBL (Dolmans et al., 1997; Otting and Zwaal, 2006). All taxonomies can be narrowed down to five "rules" for effective problems. PBL problems need 1) to build on prior knowledge, 2) elicit discussion, 3) stimulate SDL, 4) encourage knowledge integration and transfer, and finally, 5) they need to be relevant for students' future profession (Loyens et al., 2012).

Building on prior knowledge

Research in cognitive psychology has demonstrated that prior knowledge influences the quantity and quality of new knowledge (e.g., Anderson, 1990). Students need to be acquainted with at least part of the knowledge necessary to solve the problem. If not, the brainstorm is impossible. Problems that are too difficult will lead to frustration and a decrease in motivation. Problems that are too easy will be perceived as boring and insufficiently challenging and might be more effective for individual learning than for collaborative learning (Kirschner et al., 2009). Therefore, problem complexity needs to be tuned to students' prior knowledge (Otting and Zwaal, 2006). An implication of this problem characteristic is that a PBL curriculum needs to have a balanced structure. A specific problem about diabetes, for example, can only be presented to medical students after they have gained knowledge about the anatomy and physiology of the human body (Dolmans et al., 1997).

Eliciting discussion

A problem not only needs to build on students' prior knowledge, it also has to make sure it can retrieve this knowledge and subsequently elicit discussion. The retrieval of prior knowledge is crucial since new information can then be related to it. A problem can

Table 2 PBL analyzed according to Barrows' (1986) factors.

Factor	Problem-based learning
1. Design of learning task	Real-life problem E.g., "Jan and Erik, identical twins, lost their parents at the age of 7 from a car accident. Jan was fostered in the (3 children) family of a contract-laborer. Erik was fostered by the (also 3 children) family of a scientist. At the age of 6, both Jan and Erik were psychologically evaluated. Jan's IQ was measured as 111, that of Erik was 108. At the age of 12, both boys were evaluated again. Jan's IQ was now 93, and that of Erik 123. Explain the difference in the results" (Schmidt and Bouhuijs, 1981, p. 6).
2. Teacher versus student centered	Teacher = facilitator and stimulator of group discussion, providing guidance, ideally possessing subject matter expertise, cognitive congruence, and social congruence (Schmidt and Moust, 1995)
3. Sequence of information acquisition	1. Pre-discussion of a complex problem (i.e., <i>brainstorm</i>) in a group meeting. This pre-discussion of the problem ends with a collaborative formulation of learning issues (e.g., what are possible treatments for dyslexia?) 2. Individual self-directed learning activities (i.e., selecting and studying literature resources) 3. Reporting phase: Sharing and critically evaluating the literature findings in the second group meeting (Wijnia et al., 2019)

elicit discussion when it contains certain cues such as opposite views, so that students can generate arguments for and against each view and discuss which view is best. Most suitable in this respect are so-called *ill-structured* or *ill-defined problems*, because these problems can have multiple solutions and can be solved in multiple ways (Otting and Zwaal, 2006). Furthermore, ill-structured problems better represent problems that one can encounter in daily life and are hence more realistic (Loyens et al., 2012).

Stimulating self-directed learning

Problems need to lead to the formulation of learning issues that guide students' self-study activities (e.g., Hmelo-Silver, 2004), since PBL students determine themselves what they find relevant for their learning based on the learning issues (Barrows, 1996). The self-study activities to answer the learning issues are assumed to prepare students for autonomous problem solving later in life (Dolmans et al., 1997). An ill-structured problem can stimulate the formulation of learning issues and further self-directed learning because of the multiple solutions as described earlier. This can generate discussion in the tutorial group but can also cause cognitive conflict within the student (Dolmans et al., 2001; Loyens et al., 2015).

Encouraging knowledge integration and transfer

Problems need to encourage students to integrate their acquired knowledge so that they can apply this information in a subsequent, new situation. In other words, a problem needs to establish transfer of knowledge. To accomplish this, information needs to be presented in a broad context so students can better understand the purpose of the single problem elements. A description of somebody who hurt herself and feels a throbbing pain around the wound a few days later can stimulate students to discuss topics such as the natural healing process of a wound, infection, and characteristics of inflammation (Dolmans et al., 1997). These single processes and features become meaningful in the context of a story. Knowledge acquired in a meaningful context becomes more easily accessible, promoting transfer (Otting and Zwaal, 2006; Van der Vleuten et al., 2000).

Relevance for future profession

Finally, PBL problems are, preferably, relevant for students' future profession. Such problems are considered to be more motivating and lively for students (Loyens et al., 2008; Otting and Zwaal, 2006). Furthermore, they are thought to narrow the bridge between the learning situation and praxis, since they become more similar (Van der Vleuten et al., 2000). Presenting medical students with a patient's file is a good example in this respect.

The problem example presented in Table 2 about IQ and the nature-nurture debate can be presented to first-year PBL psychology students after they have studied the topic of intelligence. In this way it can build on students' prior knowledge since students typically first learn about theories of intelligence as well as the development of intelligence before tackling the nature-nurture debate. The problem deals with an important topic within psychology, intelligence and the nature-nurture debate, and is hence relevant for some of the students' future profession as a school, clinical, or educational psychologist. It contains different, relevant elements about DNA (i.e., the fact that Erik and Jan are identical twins), the families in which Jan and Erik were brought up, their IQ scores at the age of six and at the age of 12 etc., that are important for understanding the problem and that need to be integrated. This leads to a discussion about the difference in IQ scores between the age of six and 12 and most importantly, how this can be explained. Multiple options are possible and students need to unravel these options during self-study based on the formulated learning issues (e.g., "How does IQ develop?" "What is the role of genetics and environment in (the development of) IQ?").

The level of student/teacher centeredness in PBL

As mentioned, the starting point of the learning process in PBL is the problem. Problems are designed by instructors, which might be considered more teacher-centered compared to, for example, project-based learning (PjBL, see below) where students can often choose their own project to work on. However, Barrows (1996) identified "the teacher as facilitator or guide" as a core characteristic of PBL with the teacher's role being significantly different compared to that in lecture-based curricula. The student-centeredness also lies in the fact that, in the PBL process, students determine themselves what they find relevant for their learning based on the learning issues they collaboratively formulate (Barrows, 1996). The role of the tutor in PBL is to facilitate and stimulate group discussion, make sure the problem content is considered in-depth, and evaluate the group members' contributions to addressing the problems (Barrows, 1985; Loyens et al., 2012; Schmidt et al., 2007). Hence, whenever needed, the tutor is expected to ask open-ended, metacognitive questions such as "Explain in your own words what this article says about the learning issue?" when students are primarily summarizing instead of discussing or "How does this article differ from the other?" or "What do you think?" By asking such questions and catalyzing group progress, the tutor helps support knowledge building. Ideally, tutors' interventions diminish over time as students become more knowledgeable with the PBL process on the one hand and become more self-directed on the other hand, although empirical evidence on this issue is scarce and not decisive (Hmelo-Silver and Barrows, 2008).

In terms of tutor characteristics, Schmidt and Moust (1995) put forward the idea that effective tutors need cognitive congruence, social congruence, and expertise. Cognitive and social congruence emphasize the personal qualities of the tutor. *Cognitive congruence* refers to the ability to "frame his or her contributions in a language that is adapted to the level of the students" (Schmidt and Moust, 1995, p.709), whereas *social congruence* implies the tutor's willingness to be involved with the students' life and learning. A tutor also needs a suitable knowledge base for the topic under study, also referred to as *subject-matter expertise* (Schmidt and Moust, 1995, 2000). A review of research on the role of tutors in PBL supported the claim that both congruence and expertise are needed (Dolmans et al., 2002) as does a study in which good tutors were described by students as those who know when and how to intervene

without dominating (Maudsley et al., 2008). Another study on students' perceptions of the tutor showed that students saw effective tutoring as being equivalent to stimulating active learning. Distinctions between poor and excellent tutors came down to three dimensions in the eyes of students: degree of direction (i.e., not too directive, but not too passive either), content knowledge, and adequate evaluations of both content and process (Dolmans et al., 2006).

Sequence of information acquisition in PBL

Wijnia et al. (2019) highlighted that PBL as a process consists of three separate stages: an initial discussion phase, a self-study phase, and a reporting stage. Firstly, students are given a meaningful problem that describes an observable phenomenon or event. During the first stage, prior knowledge of the problem is important, as students come up with theories to explain the problem. As mentioned, students brainstorm about the problem before they receive other curriculum input and, therefore, they need to rely on their prior knowledge. Together, students try to construct understanding of the problem and discuss possible explanations or solutions. Because their knowledge is limited and insufficient, learning issues (formulated as questions) to guide students' research and study are formulated for further self-study. All of this is undertaken in a class discussion of five to 12 students.

During the second stage, students consult learning resources to gain knowledge about the problem and to answer the learning issues. These resources can be selected by the students, the tutor, or a combination of both (Wijnia et al., 2019). Next to selecting relevant literature about the topic to be studied, students need to plan their study activities to optimally prepare themselves for the next tutorial group meeting, assess whether their self-study activities were sufficient to fully understand the subject matter introduced in the problem, etc. (Loyens et al., 2008).

In the final stage, students reconvene under the guidance of their tutor to share and evaluate their findings critically and elaborate on their newly acquired knowledge. This knowledge is then applied to the problem in order to identify solutions or explanations (Loyens et al., 2022; Loyens and Rikers, 2017; Wijnia et al., 2019).

In general, these stages can be identified in the PBL process, although the specific steps that are to be followed can differ (for an overview of the different formats of the PBL process, see Wijnia et al., 2019).

Project-based learning (PjBL)

The central characteristic of project-based learning or project-centered learning (PjBL) is the project, through which students learn central concepts and principles of a discipline (Blumenfeld et al., 1991). During collaborative and problem-solving activities, students acquire new domain-specific knowledge and thinking skills useful in solving real-world problems, with a facilitating role for instructors (Chiu, 2020). Projects last for a considerable period of time (Helle et al., 2006).

Also, in PjBL, different implementations can be found, but most formats start with a driving question or problem and seem to incorporate the following features (Krajcik, 2015; Thomas, 2000): (a) the projects for which the students seek solutions or clarifications are relevant to learners' lives; (b) PjBL involves planning and performing investigations to answer questions; (c) students collaborate with other students, teachers, and members of society; and (d) PjBL is centered around producing artifacts. In addition, often technology tools are used when appropriate (Krajcik, 2015). A distinction should be made between regular projects used in education and in PjBL. The Buck Institute for Education makes this distinction:

We find it helpful to distinguish a "dessert project" – a short, intellectually-light project served up after the teacher covers the content of a unit in the usual way – from a "main course" project, in which the project is the unit. In Project Based Learning, the project is the vehicle for teaching the important knowledge and skills students need to learn. The project contains and frames curriculum and instruction

Buck Institute for Education (n.d.)

Through working on a project in PBL, skills such as critical thinking, problem solving, collaboration, and various forms of communication are believed to be fostered (Buck Institute for Education, n.d.). In sum, students are involved in a series of collaborative inquiry activities that help them acquire new, domain-specific knowledge and thinking skills useful in solving real-world problems (Saad and Zainudin, 2022). The PjBL end products are believed to reflect a learner's knowledge in information search, knowledge related to the project's topic, and metacognitive knowledge (Grant and Branch, 2005). Table 3 provides an overview of the different characteristics of PjBL.

Design of the learning task in PjBL

A project can be a problem to solve, a phenomenon to investigate, a model to design, or a decision to make, and usually starts with a driving question. An example from elementary school is: "How can we build a tiny house, taking into consideration budget, location, and personal preferences?" (Buck Institute for Education, n.d.).

Specific end products need to be reached, and these are clearly defined, while the processes to get to the end products can vary. The end products (e.g., a website, presentation, report) serve as the basis for discussion, feedback, and revision (Blumenfeld et al., 1991; Helle et al., 2006; Nurbekova et al., 2020; Tal et al., 2006).

Table 3 PjBL analyzed according to Barrows' (1986) factors.

Factor	Project-based learning
1. Design of learning task	Project that starts with a driving question E.g., "how can we build a tiny house, taking into consideration budget, location, and personal preferences?" (Buck Institute for education, n.d.)
2. Teacher versus student centered	Teacher = facilitator, helps with framing and structuring the projects, monitors the development of the end product, and assesses what students have learned
3. Sequence of information acquisition	The process starts with a driving question. The process to reach the end product (e.g., website, microbot, tiny house) may vary, but stages can often be distinguished (i.e., project organization, project implementation, and project evaluation; Nurbekova et al., 2020)

The level of student/teacher centeredness in PjBL

The projects are student-driven, which means that students have a significant degree of control of the project they will work on and what they will do in the project. This is in contrast to PBL in which the problems are designed by the instructor. The role of the instructor consists of facilitating the project. That is, the instructor helps with framing and structuring the projects, monitors the development of the end product, and assesses what students have learned (Chiu, 2020; David, 2008; Helle et al., 2006). An instructor can also choose to include experts and clients. For example, in the tiny house example for elementary school, part of the project consisted of interviewing clients and making an inventory of their wishes for the tiny house. In these interviews, learners needed to obtain crucial information from their "clients" such as the budget available and minimum square feet for their tiny house. The expert in this project was a structural engineer who demonstrated and explained a blueprint of a tiny house.

In sum, the projects in PBL are student-centered and the teacher takes on a facilitating role. In this facilitation, the teacher can bring in external persons and experts to fulfill a specific role in the project (Buck Institute for Education, n.d.).

Sequence of information acquisition in PjBL

As the project is key in PjBL, specific end products need to be reached. While the end products are clearly defined (e.g., develop a microrobot, design a specific plan, build a tiny house for a client), the processes to reach the end products can vary. In some projects, the stages of project organization, implementation, and evaluation can be distinguished (Nurbekova et al., 2020). For project organization, activities at the conceptual stage take place (e.g., formation of the project team, project selection based on the driving question, problem definition, objective of the project) as well as team building with function distribution (e.g., defining project tasks and exploring resource capabilities). During project implementation, the actual development of the product takes place (e.g., designing and programming a microrobot) as well as a first testing phase (e.g., testing and debugging the software and hardware of a microrobot). During the subsequent project evaluation phase, the project is presented (e.g., the project presentation is prepared and the "public defense" of the project takes place) and finally, the work performed is analyzed (Are the goals achieved, what went well, what went less well and why? etc.; Nurbekova et al., 2020).

Case-based learning (CBL)

In line with the previously discussed student-centered formats, there is also no consensus as to the definition of case-based learning (CBL), also referred to as case study teaching and case method learning. Nevertheless, it is considered a form of student-centered instruction and fits on the continuum between structured and guided learning (Thistlethwaite et al., 2012).

The goal of CBL is to prepare students for practice through the use of authentic cases and it centers instruction around those cases (Tawfik et al., 2020). It links theory to practice, through the application of knowledge to the cases, using inquiry-based learning methods (Thistlethwaite et al., 2012). CBL is used in many business schools, law curricula, and in medical education (Tawfik et al., 2020; Thistlethwaite et al., 2012). In contrast to other student-centered strategies, CBL focuses more on an instructor-driven discussion of the case with the student (Williams, 1992). Table 4 provides an overview of the different characteristics of CBL.

Design of the learning task in CBL

CBL is a form of collaborative learning where learners are presented with a case. For example, in a teacher education course, students can be presented with a case on classroom management in which the teacher wants to change the arrangement of children's desks because two boys are constantly talking and interrupting the class instruction. However, one of the two boys seems to have few

Table 4 CBL analyzed according to Barrows' (1986) factors.

Factor	Case-based learning
1. Design of learning task	The learning task starts with a case that students prepare. E.g., case on classroom management for students to prepare. "The teacher wants to change the arrangement of children's desks because two boys are constantly talking and interrupting the class instruction"
2. Teacher versus student centered	Teacher = facilitator, but has a more prominent role in the discussion of the cases. Whereas in PBL, the facilitator can remain silent when the group discussion is going well, in CBL the group discussion is led by the teacher.
3. Sequence of information acquisition	1. Students prepare the case before they come to class, using prior knowledge but also literature resources. For example, for cases in law school, students will search legal dictionaries, encyclopedias, textbooks, statutes, and treatises to help them understand the case (Williams, 1992). 2. The teacher poses a number of questions to the group of students related to the case. For example, in law education, the teacher may question the decision made by the court described in the case. "The discussion session serves as the venue for students to collaboratively apply learned principles and data analyses and evaluate the usefulness of various strategies to achieve optimal resolutions of the case" (Bowe et al., 2009, p. 835, p. 835) 3. Optional: The teacher presents a series of related hypothetical cases that represent variations or perturbations of the original case.

friends and is experiencing problems adjusting to school, and the boy next to him seems to be the only one he is close with (Choi and Lee, 2009, p. 104). Cases are similar to problems in PBL. However, in Barrows' terms (1986), the sequence of the problem and gathered information is different. In PBL, the problem is the starting point. In CBL, students need to prepare in advance for the group session and can ask questions during the session, when the case is discussed under the guidance of a facilitator (Srinivasan et al., 2007; Williams, 2005). Therefore, CBL is sometimes considered as a special form of PBL. Both PBL and CBL share a focus on acquiring new knowledge (Aditomo et al., 2013). The overlap between PBL and CBL can also be seen in the characteristics of the learning task, which share significant overlap: cases need to be "authentic" and typical (i.e., based on real-life events or patients or situations to enhance the relevance of the subject matter by focusing on the actual performance of professionals.), stimulate interest, tell a story, promote decision making, and be aligned with defined learning outcomes. Cases can be real (e.g., real patients in health professional education), simulated, virtual, or text-based (Thistlethwaite et al., 2012). Williams (1992) added in this respect that the cases used in CBL reflect real cases, but these cases often contain less complexity and do not contain irrelevant details.

The level of student/teacher centeredness in CBL

CBL focuses more on an instructor-driven discussion of the case with the student. Also, students need to prepare the case at home and, whereas learning issues/questions are collaboratively decided upon in PBL, the learning issues are determined by the teacher in CBL (Thistlethwaite et al., 2012). The teacher also has a more prominent role in the group discussion. As noted in Table 4, whereas in PBL the teacher (i.e., tutor) can remain relatively silent during the group discussion when the students are doing a good job and little 'steering' is necessary, the situation is different in CBL, where the teacher leads the discussion of the case.

Even though the teacher plays a key role in facilitating the discussion in CBL, this method can still be considered student-centered. As Bowe et al. (2009) stated: "The Case Method Teaching format encourages students to articulate their perspectives and propose problem-solving strategies. It also provides the faculty valuable insight into the students' understanding of key concepts and their ability to effectively apply them" (p. 835).

Some researchers have argued that CBL is more structured than PBL, but that is to a large extent dependent on the method used for the learning process, as PBL formats can also employ structured methods to guide the group discussion (Wijnia et al., 2019).

Sequence of information acquisition in CBL

Typically, CBL takes place in face-to-face groups but may also take place online. Students review cases prior to sessions in order to be able to discuss them in groups. By preparing the case independently at home, students can identify the key issues of the case and

study additional resources (see [Table 4](#) for examples). The preparation at home is crucial to ensure a thorough and in-depth discussion of the case in the classroom. After preparing the case at home, students return to their classes where the teacher starts a facilitated discussion with trigger questions about the case ([Bowe et al., 2009](#)). As an optional third step, the teacher can present related hypothetical cases that represent variations of the original case, which allow students to apply their newly learned knowledge to a new and different case. Practically, a teacher can introduce a case in a lecture and let students prepare the case independently. The lecture is then followed by a group session in which the case is analyzed in depth. Typically, students have little work after the session in which the case is discussed ([Bowe et al., 2009](#)).

Acquaintances of the student-centered family: team-based and challenge-based learning

As mentioned, the lack of clear definitions makes it difficult to distinguish the different student-centered formats. Concepts are often confused (e.g., problem- and project-based learning and problem- and case-based learning), or incorrectly explained (e.g., calling a method PBL when the problem is not the starting point of the learning process). What also adds to the difficulty in reaching conceptual clarity is the fact that new methods have emerged throughout the years, sometimes combining elements of specific student-centered formats and general beneficial learning strategies.

Team-based learning

One such example is Team-based learning (TBL), which could be considered a PBL variant since students also work on problems in teams. It is similar to CBL in that students have to prepare before coming to class. One could argue that TBL overcomes some of the problems put forward with respect to PBL. For example, an often-heard criticism of PBL concerns the facilities (small rooms are needed to host the tutorial group discussions of 10–12 students each) and associated costs (all groups need a tutor to facilitate the PBL discussion, requiring a large number of tutors; [Loyens et al., 2012](#)). In TBL, one teacher facilitates various small teams, sometimes 20 or even more ([Dolmans et al., 2015](#)). Students gather in teams and they are all present in the same large room in TBL. TBL also makes use of the beneficial learning strategy of testing. Practice testing one's knowledge has been shown to be a very effective learning strategy ([Dunlosky et al., 2013](#)). Before group sessions, students have to complete a mandatory reading assignment and their individual knowledge of the materials is assessed in a multiple-choice test. In the teams, students discuss their individual answers to the exact same test items with each other to reach team consensus and receive immediate feedback on the team answers ([Haidet et al., 2012](#); [Dolmans et al., 2015](#)). This test provides the teacher with valuable information about difficult concepts for students, so that extra attention can be paid to the clarification of those concepts. Subsequently, students work in the same teams to resolve professionally relevant, challenging, and complex problems. After teams have worked on the problem, they present their solutions in plenary to the other teams. Solutions to problems are exchanged and challenged through discussion, facilitated by the teacher ([Dolmans et al., 2015](#)). Often experts are invited to attend these plenary sessions. In terms of the level of student/teacher centeredness, one could say that TBL is more structured regarding the teams, the pre-activity testing, and the formation of the problems ([Davis, 2017](#)). In addition, there is a strong focus of TBL on collaboration and the acquisition of teamwork skills ([Dolmans et al., 2015](#)).

Challenge-based learning (ChBL)

ChBL is another approach that is identified as a student-centered format with links to the above-mentioned approaches. In ChBL, “students collaborate with academia, industry and extra-academic actors to solve real-world challenges through creative and authentic experience” ([Gallager and Savage, 2020](#), p. 2–3). Hence, stakeholders from multiple settings are used in ChBL to support students. This aspect is different from other formats, where teachers usually act as facilitators (see [Tables 1–4](#)), with the exception of PjBL, where extra-academic actors can also be invited to the classroom. [Gallager and Savage \(2000\)](#) reviewed 100 articles to discover the key principles of ChBL. They found that ChBL is usually focused on themes of global importance such as sustainability and other real-world challenges. The solution of, or decision on, the challenge often requires input from multiple disciplines, creativity and innovation, and the use of technology. Depending on the methodology followed, the challenges in ChBL can be either defined by the students or by the teachers. Unlike PjBL, the focus is not just on the final product but also on the process, which is viewed as being as important ([Gallager and Savage, 2000](#)).

Directions for future research

This chapter aimed to achieve conceptual clarity on several distinct student-centered methods. Although significant overlap can be observed in IBL, PBL, PjBL, and CBL, they all have their unique features. These unique features make it paramount to include definitions as well as details on the procedures and methodology. Both are of key importance in research articles. Only by doing so, will it become possible for researchers to fully understand student-centered formats and conduct reviews and meta-analyses. In a recent review on critical thinking in PBL and PjBL, the most prevalent reason for exclusion of research articles was a lack of a clear definition and/or description of the methodology ([Loyens et al., 2022](#)).

Although this chapter dealt with the conceptualizations of student-centered instruction, an inevitable follow-up question is how effective these methods are in fostering learning. The majority of studies deal with experiences of teachers and students with these methods, while there is a paucity of controlled effect studies (Granger et al., 2012; Loyens et al., 2012). Effect studies should not serve the sole purpose of answering questions such as “Which method is better than the other?”; nonetheless, the effects of instructional formats can only be established with a comparable control group. Whereas the effects of student-centered formats on knowledge acquisition have been well documented (e.g., Chen and Yang, 2019; Lazonder and Harmsen, 2016; Walker and Leary, 2009), the effects on skill acquisition have been less frequently studied. Only with well-conducted and well-documented research, can we combine the best of all worlds and advance the use of student-centered approaches in classrooms.

References

- Aditomo, A., Goodyear, P., Blui, A., Ellis, R.A., 2013. Inquiry-based learning in higher education: principal forms, educational objectives, and disciplinary variations. *Stud. High Educ.* 38, 1239–1258. <https://doi.org/10.1080/03075079.2011.616584>.
- Anderson, J.R., 1990. *Cognitive Psychology and its Implications*. Freeman, New York.
- Barrows, H.S., 1985. *How to Design a Problem-Based Curriculum for Preclinical Years*. Springer, New York.
- Barrows, H.S., 1986. A taxonomy of problem-based learning methods. *Med. Educ.* 20, 481–486.
- Barrows, H.S., 1996. Problem-based learning in medicine and beyond: a brief overview. In: Gijssels, W.H. (Ed.), *New Directions for Teaching and Learning*, vol. 68. Jossey-Bass, San Francisco, pp. 3–11.
- Blanchard, M.R., Souterland, S.A., Granger, E.M., 2008. No silver bullet for inquiry: making sense of teacher change following an inquiry-based research experience for teachers. *Sci. Teach Educ.* 93, 322–360.
- Blumenfeld, P.C., Soloway, E., Marx, R.W., Krajcik, J.S., Guzdial, M., Palincsar, A., 1991. Motivating project-based learning: sustaining the doing, supporting the learning. *Educ. Psychol.* 26 (3–4), 369–398.
- Bowe, C.M., Voss, J., Aretz, H.T., 2009. Case method teaching: an effective approach to integrate the basic and clinical sciences in the preclinical medical curriculum. *Med. Teach.* 31 (9), 834–841. <https://doi.org/10.1080/01421590902922904>.
- Buck Institute for Education, (n.d.). What Is Project-Based Learning? Retrieved May 20, 2022, from <https://www.pblworks.org/what-is-pbl>
- Chen, C.-H., Yang, Y.-C., 2019. Revisiting the effects of project-based learning on students' academic achievement: a meta-analysis investigating moderators. *Educ. Res. Rev.* 26, 71–81. <https://doi.org/10.1016/j.edurev.2018.11.001>.
- Chiu, C.-F., 2020. Facilitating K-12 teachers in creating apps by visual programming and project-based learning. *Int. J. Eng. Technol. Learn.* 15 (1), 103–118. <https://www.learnlib.org/p/217066/>.
- Choi, I., Lee, K., 2009. Designing and implementing a case-based learning environment for enhancing ill-structured problem solving: classroom management problems for prospective teachers. *Educ. Technol. Res. Dev.* 57, 99–129.
- David, J.L., 2008. Project-based learning. *Educ. Leader: J. Depart. Supervision Curric. Develop.* 65 (5), 80–82.
- Davis, D.A., 2017. Getting started with team-based learning. *Interdiscip. J. Prob. Based Learn.* 11 (1). <https://doi.org/10.7771/1541-5015.1701>.
- Dolmans, D.H.J.M., Gijssels, W.H., Moust, J.H.C., de Grave, W.S., Wolfhagen, I.H.A.P., Van der Vleuten, C.P.M., 2002. Trends in research on the tutor in problem-based learning: conclusions and implications for educational practice. *Med. Teach.* 24, 173–180.
- Dolmans, D.H.J.M., Janssen-Noordman, A., Wolfhagen, I.H.A.P., 2006. Can students differentiate between PBL tutors with different tutoring deficiencies? *Med. Teach.* 28, E156–E161.
- Dolmans, D.H.J.M., Michaelsen, L., van Merriënboer, J., van der Vleuten, C.P.M., 2015. Should we choose between problem-based learning and team-based learning? No, combine the best of both worlds. *Med. Teach.* 37 (4), 354–359. <https://doi.org/10.3109/0142159X.2014.948828>.
- Dolmans, D.H.J.M., Snellen-Balendong, H., Wolfhagen, I.H.A.P., Van der Vleuten, C.P.M., 1997. Seven principles of effective case design for a problem-based curriculum. *Med. Teach.* 19, 185–189.
- Dolmans, D.H.J.M., Wolfhagen, I.H.A.P., Van der Vleuten, C.P.M., Wijnen, W.H., 2001. Solving problems with group work in problem-based learning: hold on to the philosophy. *Med. Educ.* 35, 884–889.
- Dunbar, K., Yadav, A., 2022. Shifting to student-centered learning: influences of teaching a summer service learning program. *Teach. Teach. Educ.* 110, 103578. <https://doi.org/10.1016/j.tate.2021.103578>.
- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., Willingham, D.T., 2013. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest: A Journal of the American Psychological Society.* 14 (1). <https://doi.org/10.1177/1529100612453266>.
- Gallagher, S.E., Savage, T., 2020. Challenge-based learning in higher education: an exploratory literature review. *Teach. High. Educ.* 1–23. <https://doi.org/10.1080/13562517.2020.1863354>.
- Granger, E.M., Bevis, T.H., Saka, Y., Southerland, S.A., Sampson, V., Tate, R., 2012. The efficacy of student-centered instruction in supporting science learning. *Science* 338, 105–108.
- Grant, M.M., Branch, R.M., 2005. Project-based learning in a middle school: tracing abilities through the artifacts of learning. *J. Res. Technol. Educ.* 38 (1), 65–98. <https://doi.org/10.1080/15391523.2005.10782450>.
- Grant, M.M., Hill, J.R., 2006. Weighing the rewards with the risks? Implementing student-centered pedagogy within high-stakes testing. In: Lambert, R., McCarthy, C. (Eds.), *Understanding Teacher Stress in the Age of Accountability*. Information Age, Greenwich, CT, pp. 19–42.
- Haidet, P., Levine, R.E., Parmelee, D.X., Crow, S., Kennedy, F., Kelly, A., Perkowski, L., Michaelsen, L., Richards, B.F., 2012. Guidelines for reporting team-based learning activities in the medical and health sciences education literature. *Acad. Med.* 87 (3), 292–299. <https://doi.org/10.1097/ACM.0b013e318244759e>.
- Hannafin, M.J., Hill, J.R., Land, S.M., Lee, E., 2014. Student-centered, open learning environments: research, theory, and practice. In: Spector, J., Merrill, M., Elen, J., Bishop, M. (Eds.), *Handbook of Research on Educational Communications and Technology*. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-3185-5_51.
- Helle, L., Tynjälä, P., Olkinuora, E., 2006. Project-based learning in secondary education – theory, practice and rubber sling slots. *High Educ.* 51, 287–314. <https://doi.org/10.1007/s10734-004-6386-5>.
- Hmelo-Silver, C.E., 2004. Problem-based learning: what and how do students learn? *Educ. Psychol. Rev.* 16, 235–266. <https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>.
- Hmelo-Silver, C.E., Barrows, H.S., 2008. Facilitating collaborative knowledge building. *Cognit. Instruct.* 26, 48–94.
- Kirschner, F., Paas, F., Kirschner, P.A., 2009. A cognitive load approach to collaborative learning: united brains for complex tasks. *Educ. Psychol. Rev.* 21, 31–42.
- Klahr, D., Dunbar, K., 1988. Dual space search during scientific reasoning. *Cognit. Sci.* 12, 1–48. https://doi.org/10.1207/s15516709cog1201_1.
- Krajcik, J., 2015. Project-based science. *Sci. Teach.* 82 (1), 25–27. http://www.nsta.org/store/product_detail.aspx?id=10.2505/4/tst15_082_01_25.
- Lazonder, A.W., Harmsen, R., 2016. Meta-analysis of inquiry-based learning: effects of guidance. *Rev. Educ. Res.* 86 (3), 681–718. <https://doi.org/10.3102/0034654315627366>.
- Lee, E., Hannafin, M.J., 2016. A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it. *Educ. Technol. Res. Dev.* 64, 707–734.

- Loyens, S.M.M., Broughton, S., Mikkers, J., van Gog, T., 2015. Problem-based learning as a facilitator of conceptual change. *Learn. Instruct.* 38, 34–42. <https://doi.org/10.1016/j.learninstruc.2015.03.002>.
- Loyens, S.M.M., Kirschner, P., Paas, F., 2012. Problem-based learning. In: Graham, S., Bus, A., Major, S., Swanson, L. (Eds.), *APA Educational Psychology Handbook: Vol. 3. Application to Learning and Teaching*. American Psychological Association, Washington, DC, pp. 403–425.
- Loyens, S.M.M., Magda, J., Rikers, R.M.J.P., 2008. Self-directed learning in problem-based learning and its relationships with self-regulated learning. *Educ. Psychol. Rev.* 20, 411–427.
- Loyens, S.M.M., Rikers, R.M.J.P., 2017. Instruction based on inquiry. In: Mayer, R.E., Alexander, P.A. (Eds.), *Handbook of Research on Learning and Instruction*, second ed. Routledge, pp. 405–431.
- Loyens, S.M.M., van Meerten, J.E., Schaap, L., Wijnia, L., 2022. Situating Higher-Order, Critical, and Critical-Analytic Thinking in Problem- and Project-Based Learning Environments: A Systematic Review. Manuscript submitted for publication.
- Loyens, S.M.M., Wijnia, L., van der Sluijs – Duker, I., Rikers, R.M.J.P., 2021. Problem-based learning. In: Zhang, L. (Ed.), *The Oxford Encyclopedia of Educational Psychology*. Oxford University Press.
- Maudsley, G., Williams, E.M., Taylor, D.C., 2008. Problem-based learning at the receiving end: a 'mixed methods' study of junior medical students' perspectives. *Adv. Health Sci. Educ.* 13 (4), 435–451. <https://doi.org/10.1007/s10459-006-9056-9>.
- Murphy, C., Abu-Tineh, A., Calder, N., Mansour, N., 2021. Teachers and students' views prior to introducing inquiry-based learning in Qatari science and mathematics classrooms. *Teach. Educ.* 104, 103367. <https://doi.org/10.1016/j.tate.2021.103367>.
- Norman, G.R., Schmidt, H.G., 1992. The psychological basis of problem-based learning: a review of the evidence. *Acad. Med.* 67, 557–565.
- Nurbekova, Z., Tolganbaiuly, T., Nurbekov, B., Sagimbayeva, A., Kazhiakparova, Z., 2020. Project-based learning technology: an example in programming microcontrollers. *Int. J. Emerg. Technol. Learn. (IJET)* 15 (11), 218–227. <https://doi.org/10.3991/ijet.v15i11.13267>.
- Otting, H., Zwaal, R., 2006. Critical task characteristics in problem-based learning. *Ind. High. Educ.* 10, 347–357.
- Pedaste, M., Mäeots, M., Siiman, L.A., De Jong, T., Van Riesen, S.A., Kamp, E.T., Tsoouridaki, E., 2015. Phases of inquiry-based learning: definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61.
- Saad, A., Zainudin, S., 2022. A review of project-based learning (PBL) and computational thinking (CT) in teaching and learning. *Learn. Motiv.* 78. <https://doi.org/10.1016/j.lmot.2022.101802>. Article 101802.
- Schiefer, J., Golle, J., Tibus, M., Oschatz, K., 2019. Scientific reasoning in elementary school children: assessment of the inquiry cycle. *J. Adv. Acad.* 30 (2), 144–177. <https://doi.org/10.1177/1932202X18825152>.
- Schmidt, H.G., 1983. Problem-based learning: rationale and description. *Med. Educ.* 17, 11–16.
- Schmidt, H., Bouhuijs, P., 1981. *Vijfentwintig Problemen*. Maastricht University, Maastricht, Netherlands.
- Schmidt, H.G., Cohen-Schotanus, J., Van der Molen, H.T., Splinter, T.A.W., Bulte, J., Holdrinet, R., Van Rossum, H.J.M., 2009. Learning more by being taught less: a "time-for-self-study" theory explaining curricular effects on graduation rate and study duration. *High Educ.* 60 (3), 287–300. <https://doi.org/10.1007/s10734-009-9300-3>.
- Schmidt, H.G., Loyens, S.M.M., van Gog, T., Paas, F., 2007. Problem-based learning is compatible with human cognitive architecture: commentary on Kirschner, Sweller, and Clark (2006). *Educ. Psychol.* 42, 91–97. <https://doi.org/10.1080/00461520701263350>.
- Schmidt, H.G., Moust, J.H.C., 1995. What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula. *Acad. Med.* 70, 708–714.
- Schmidt, H.G., Moust, J.H.C., 2000. Factors affecting small-group learning: a review of the research. In: Evensen, D.H., Hmelo, C.E. (Eds.), *Problem-based Learning: A Research Perspective on Learning Interactions*. Erlbaum, Mahwah, NJ, pp. 19–52.
- Srinivasan, M., Wilkes, M., Stevenson, F., Nguyen, T., Slavin, S., 2007. Comparing problem-based learning with case-based learning: effects of a major curricular shift at two institutions. *Acad. Med.* 82, 74–82.
- Tal, T., Krajcik, J.S., Blumenfeld, P.C., 2006. Urban schools' teachers enacting project-based science. *J. Res. Sci. Teach.* 43 (7), 722–745. <https://doi.org/10.1002/tea.20102>.
- Tawfik, A.A., Hung, W., Giabbanelli, P.J., 2020. Comparing how different inquiry-based approaches impact learning outcomes. *Interdiscip. J. Prob. Based Learn.* 14 (1). <https://doi.org/10.14434/ijpbl.v14i1.28624>.
- Thistlethwaite, J.E., Davies, D., Ekeocha, S., Kidd, J.M., MacDougall, C., Matthews, P., Purkis, J., Clay, D., 2012. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. *Med. Teach.* 34 (6), e421–e444. <https://doi.org/10.3109/0142159X.2012.680939>.
- Thomas, J.W., 2000. A Review of Research on Project-Based Learning. <http://www.bie.org/files/researchreviewPBL.pdf>.
- van Uum, M.S.J., Verhoeff, R.P., Peeters, M., 2016. Inquiry-based science education: towards a pedagogical framework for primary school teachers. *Int. J. Sci. Educ.* 38 (3), 450–469. <https://doi.org/10.1080/09500693.2016.1147660>.
- van Uum, M.S.J., Verhoeff, R.P., Peeters, M., 2017. Inquiry-based science education: scaffolding pupils' self-directed learning in open inquiry. *Int. J. Sci. Educ.* 39 (18), 2461–2481. <https://doi.org/10.1080/09500693.2017.1388940>.
- van Uum, M.S.J., Peeters, M., Verhoeff, R.P., 2021. Professionalising primary school teachers in guiding inquiry-based learning. *Res. Sci. Educ.* 51, 81–108. <https://doi.org/10.1007/s11165-019-9818-z>.
- Van der Vleuten, C.P.M., Dolmans, D.H.J.M., Scherpbier, A.J.J.A., 2000. The need for evidence in education. *Med. Teach.* 22, 246–250.
- Walker, A., Leary, H., 2009. A problem based learning meta-analysis: differences across problem types, implementation types, disciplines, and assessment levels. *Interdiscip. J. Prob. Based Learn.* 3 (1). <https://doi.org/10.7771/1541-5015.1061>.
- Wijnia, L., Loyens, S.M.M., Rikers, R.M.J.P., 2019. The problem-based learning process: an overview of different models. In: Moallem, M., Hung, W., Dabbagh, N. (Eds.), *The Wiley Handbook of Problem-Based Learning*. John Wiley & Sons, USA, pp. 273–295. <https://doi.org/10.1002/9781119173243.ch12>.
- Williams, S.M., 1992. Putting case-based instruction into context: examples from legal and medical education. *J. Learn. Sci.* 2 (4), 367–427.
- Williams, B., 2005. Case based learning – a review of the literature: is there scope for this educational paradigm in prehospital education? *Emerg. Med. J.* 22, 577–581.
- Windschitl, M., 2003. Inquiry projects in science teacher education: what can investigative experiences reveal about teacher thinking and eventual classroom practice? *Sci. Educ.* 87, 112–143. <https://doi.org/10.1002/sce.10044>.

Universal Design for Learning in its 3rd decade: a focus on equity, inclusion, and design

Kavita Rao^a, Jenna W. Gravel^b, David H. Rose^b, and T. Nicole Tucker-Smith^c, ^a University of Hawai'i at Mānoa, College of Education, Honolulu, HI, United States; ^b CAST, Wakefield, MA, United States; and ^c Lessoncast, Baltimore, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	712
What is UDL?	713
UDL as an orientation	713
Embracing learner variability	713
Reframing of barriers to learning and relocating the “problem” to be fixed	714
Reconsidering beliefs about the capabilities of all learners	714
UDL rising to equity: more fully developing the UDL framework	715
UDL as an instructional design framework	715
The UDL Design Cycle: inclusive instructional design	715
Flipped, blended and online: UDL for all formats	717
UDL Reporting Criteria	718
Conclusion	718
References	719

Introduction

In 2010, two of the authors of this current chapter wrote an article on Universal Design for Learning (UDL) for the third edition of the *International encyclopedia of education* (Rose and Gravel, 2010). In that chapter, we introduced UDL, a framework for teaching and learning that guides the design of accessible, inclusive, and challenging learning environments (Rose and Meyer, 2002). At the time, the UDL field was just in its first decade, and we framed UDL as “still in the very early stages of development” (p. 124). We described that while UDL was gaining traction at the policy level, “most educators do not really know what UDL means” (p. 119). Thus, the goal for that 2010 chapter was “to clarify what is meant by UDL and especially to provide the evidentiary basis for its principles and a review of its guidelines” (p. 119).

Now more than 11 years later, we have the opportunity to contribute once again to the *International encyclopedia of education*, and the field of UDL has dramatically changed. UDL has now entered into its third decade and has gained remarkable momentum within individual classrooms, school systems, and international educational initiatives. Schools and institutions of higher education throughout the U.S. and abroad are embarking on UDL initiatives in order to design inclusive learning environments that reduce barriers and support learners to thrive. For example, Bartholomew Consolidated School Corporation (BCSC) in Columbus, Indiana, USA, implemented a district-wide UDL initiative more than a decade ago, and UDL has become an overarching framework for schools across the district (BCSC, 2021). Further, institutions of higher education such as Boston College, Goodwin College, and the California State University System have developed UDL initiatives to promote the application of UDL to course design (UDL on Campus, 2021). There has also been remarkable growth in professional learning around UDL. In the past year alone, CAST’s professional learning team has collaborated with more than 4000 educators across 44 states and 25 countries. Finally, UDL has taken hold at the national policy level. In the US, for example, UDL has been defined and endorsed in the Every Student Succeeds Act (ESSA) of 2016 and incorporated into the Strengthening Career and Technical Education for the 21st Century Act of 2018 (CAST, 2021). Further, UDL has been emphasized in two recent versions of the National Education Technology Plan, a policy document that articulates the vision of technology in US schools (CAST, 2021).

UDL has been shaped by various stakeholders in the educational community, including practitioners, researchers, and developers of educational materials such as digital curriculum and instructional technology tools. Established in 2011, the UDL Implementation and Research Network (UDL-IRN), a grassroots organization, was instrumental in bringing together stakeholders and fostering ongoing dialog, cross-pollination of perspectives, and expertise around UDL (UDL-IRN, 2022). In the past decade, members of the UDL-IRN have collaborated to develop various tools and resources and disseminate information including white papers, databases of UDL studies, and tools for designing and reporting on UDL (see udl-irn.org/). Further, more researchers are exploring UDL through lenses such as equity, instructional design, learning technologies, and teacher education, with major educational journals such as the *Harvard Educational Review* dedicating a forum (Alim et al., 2017) and *Remedial and Special Education* dedicating a special issue (King-Sears, 2020) to the topic. Finally, UDL has gained a marked presence in social media and a weekly #UDLchat on Twitter which attracts a dedicated number of followers. Thus, UDL today represents a framework that has matured through the integration of ideas, perspectives, applications, and research by the educational community itself.

In this chapter, we trace the maturation and development of the UDL framework in the past decade. UDL today has come to represent both a set of values around inclusive education and an instructional design framework that can be applied to the design and development of curriculum and instruction. At a macro level, UDL is an orientation toward teaching and learning—a framework that has the potential to prompt practitioners to uncover and wrestle with their preconceived attitudes and beliefs. At a practical level, UDL is a way to consider educational practices and develop flexible and engaging curricula. In this chapter, we address these two aspects of UDL and also share a vision for the future.

What is UDL?

Universal Design for Learning (UDL) offers a framework for teaching and learning that guides the design of accessible and challenging learning environments (CAST, 2021a). The UDL framework was first developed at CAST (formerly known as Center for Applied Special Technology), a not-for-profit research and development organization formed in 1984 by a group of about 20 clinicians (mostly neuropsychologists) and educators (mostly teachers and evaluators of children with disabilities) who worked at the Medical and Educational Evaluation Center within a children's hospital near Boston, Massachusetts. The original focus of that group was on identifying how new and emerging technologies might be helpful for the children who were diagnosed as having disabilities that interfered with their learning in traditional schools. Eventually that work led to a significantly different mission: identifying and expanding the ways that new learning technologies could reduce the barriers and impediments that many children faced in traditional schools. Eventually CAST left the hospital setting—where the focus was on “fixing” children who were considered “deficient” in some way—and launched as a separate organization whose focus was on “fixing” or transforming the curriculum to support and embrace the natural variability of learners.

Eventually CAST's approach attracted a variety of educators, designers, technology developers, researchers, and learning scientists. Together that group began to formalize an approach to educational re-design that had much in common with the Universal Design movement in architecture, and began calling the approach Universal Design for Learning. Given the focus on learning rather than physical structures, CAST developed a foundational framework and guidelines that were based in the learning sciences, particularly in cognitive and affective neuroscience. That work had roots in classical developmental psychology and neurology (e.g., Piaget, 1952; Vygotsky, 1978; Wood et al., 1976; Luria, 1976), but also in contemporary cognitive and affective neuroscience. Referencing Luria's original formulation and many others who followed, the UDL framework is based on the foundation of the three main divisions of the brain: the affective network, responsible for emotional functioning; the recognition network, responsible for perceptual and pattern recognition functioning; and the strategic network, responsible for motor and executive functioning (Meyer et al., 2014). These three brain networks provide the foundation for the three principles of UDL: provide *multiple means of engagement* to motivate learners; provide *multiple means of representation* to ensure that learners can access information; and provide *multiple means of action and expression* to support learners to act on and express what they know (Meyer et al., 2014).

CAST's *UDL Guidelines*, which are available at udlguidelines.cast.org (CAST, 2018b), are a tool to support stakeholders in applying the UDL principles to practice. They too are based on the basic structure of the brain, in this case, the “hierarchical” structure of each of the three brain systems, from the primary sensory, motor, and limbic cortex up through the “tertiary” prefrontal cortex where affective, executive, and cognitive functions merge. More informally, many describe the three resulting rows in the guidelines as the following: The “access” row offers guidelines to ensure that learners have access to the learning goal by recruiting interest and by offering options for perception and physical action; the “build” row offers guidelines to develop effort and persistence, language and symbols, and expression and communication; and the “internalize” row represents the guidelines to develop self-regulation, comprehension, and executive function (CAST, 2021b). The guidelines culminate to the ultimate goal of supporting the development of “expert learners” who “are, each in their own way, resourceful and knowledgeable; strategic and goal-directed; purposeful and motivated” (Rose and Gravel, 2013, pp. 82–84).

UDL as an orientation

Over the past three decades, UDL has developed as a framework for reflecting on orientations and commitments toward teaching and learning. Many stakeholders within the UDL community—including preK–12 educators, faculty members in higher education, curriculum developers, researchers, parents, and others who apply UDL to the design learning environments—have leveraged UDL as a tool to begin to explore the attitudes and beliefs they bring to their practice. Thus, UDL has matured as a framework that can potentially prompt practitioners to uncover and consider their preconceived values, supporting practitioners to wrestle with their orientations toward learner variability, their framing of barriers to learning, and their beliefs about the capabilities of all learners.

Embracing learner variability

UDL encourages educators to examine their beliefs with regard to individual differences among learners. The framework is rooted in the notion of learner variability—the idea that learners are highly variable in the ways that they learn. Meyer et al. (2014) emphasize that “individual brains are remarkably distinctive in their anatomy, chemistry, and physiology” and that “like each person's

fingerprints, every brain is unique” (p. 49). This notion stands in stark contrast to the tendency to design for the “mythical average learner” that is deeply entrenched across the education system in the United States—and in other countries as well.

The UDL framework’s rejection of the “average learner” is closely aligned to the social model of disability, a model that positions normalcy as a social construct and upholds disability as a result of the interaction between the individual and the environment. In contrast to the medical model that interprets disability as a deficit that must be fixed, the social model frames disability as a natural variation of the human experience (Baglieri and Lavani, 2019; Shakespeare, 2010). Similarly, the UDL framework positions disability—and individual differences among learners more broadly—as a natural and inevitable form of human variation. Thus, UDL has the promise to prompt educators to examine their beliefs with regard to the meaning of “average” and “normal” and to embrace the natural variability of learners.

Reframing of barriers to learning and relocating the “problem” to be fixed

Identifying and reducing barriers to learning is a central tenet of UDL. Oftentimes, educators are socialized into practices and deficit orientations that locate barriers within learners themselves. For example, an educator may believe that a learner is unable to understand the plot of a story because the themes are too complex, or a learner is unable to complete a long-term project given their struggle to stay organized. The UDL framework supports educators to relocate the “problem” from within the individual learner to within the design of the learning opportunity, emphasizing that it is the curriculum rather than the learner that needs to be “fixed” (Meyer and Rose, 2005). UDL prompts educators to ask: What barriers stand in the way of the learning goal? How might I apply UDL to reduce these barriers and increase access to the goal? For example, if the goal is understanding plot, the UDL framework supports educators to consider options for reducing barriers to the text, including text-to-speech, audiobook versions, and graphic organizers to uncover plot structure. If the goal is to complete a long-term assignment, UDL prompts educators to consider options for supporting organization and executive function such as developing short-term goals, creating checklists, and engaging in consistent check-ins with peers.

This relocation of the “problem” can be a powerful way to support educators to begin to reframe their understandings and beliefs with regard to barriers in their classrooms. UDL encourages educators to approach challenges through a mindset that shifts away from identifying barriers within individual learners and instead focuses on ways to reduce barriers in the learning environment through proactive design. (See [Box 1](#) for examples).

Reconsidering beliefs about the capabilities of all learners

The UDL framework has also developed to be a potentially powerful tool that educators can use to examine their expectations and reimagine the capabilities of all learners. Far too often, educators are socialized into entrenched cultures of low expectations (Baglieri et al., 2011; Causton-Theoharis et al., 2011; Gartner and Lipsky, 1987; Hehir, 2002). UDL provides a lens for educators to examine and confront limiting beliefs.

As noted above, UDL is focused on the goal of supporting learners to develop as “expert learners” (Rose and Gravel, 2013). Instead of focusing on basic facts and rote memorization, UDL emphasizes the importance of designing learning goals that engage learners in meaningful ways of thinking and doing. This emphasis on the goal prompts educators to interrogate the nature of their learning goals and to explore the values and beliefs that undergird them. Is the goal rigorous? Is the goal authentic to learners? Do educators believe that all learners can engage in these rigorous and authentic goals? UDL encourages educators to begin examining these challenging questions and to reconsider the kinds of goals that they set for their learners—as well as the kinds of goals they support learners to set for themselves.

Box 1 Teacher perspectives of a UDL mindset

In 2018, CAST partnered with a New England state to launch a multi-year, state-wide professional learning opportunity designed to support educators to apply UDL to their practice. Now in its fourth year, CAST is collaborating with more than 70 schools and 600 educators from across the state. Over the four years, evidence of the UDL framework supporting educators to shift their understanding of barriers from within individual learners to within the design of the environment has emerged. The following reflections from participating educators are representative of this shift:

- “I am looking for barriers, not learning deficits.”
- “Changing our own mindsets to engage our students and identify barriers in the classroom rather than always focusing on the students was eye opening.”
- “Understanding UDL allows me to consider all my students, what they need to be successful, and what barriers are in the classroom and lessons that prevent learning from occurring.”
- “I’m making more conscientious decisions and efforts to implement strategies and options to eliminate barriers so all can access learning and share what they know.”

UDL rising to equity: more fully developing the UDL framework

Although UDL has prompted various stakeholders to begin wrestling with their orientations toward learner variability, their framing of barriers to learning, and their beliefs about the capabilities of all learners, it is important to note that exposure to UDL alone does not prompt these kinds of reflections. As with so many professional learning journeys, leveraging UDL as a tool for reflecting on one's orientations toward teaching and learning takes sustained leadership and commitment over time (Fullan, 2007; Garet et al., 2001; Knapp, 2003). Practitioners who have made these shifts in their orientations have been a part of learning communities that prioritize time to grapple with the tensions and new ways of thinking that UDL can surface.

Further, the *UDL Guidelines* are intended to be a living, dynamic tool that can be developed over time based on new research and feedback from stakeholders (CAST, 2021b). CAST intentionally called the first version of the guidelines "Version 1.0" to emphasize the beginning of an iterative process. Thus, the *UDL Guidelines*—and the UDL framework more broadly—continue to evolve to better support stakeholders to wrestle with the attitudes and beliefs that they bring to the design of learning opportunities and to push even further in terms of how UDL can be more fully developed as a lever for equity. Over the past several years, there has been a growing call from stakeholders to explore how UDL intersects with concepts such as power, privilege, oppression, and identity. For example, researchers and practitioners are advocating for UDL to make connections to Culturally Sustaining Pedagogy (CSP) (Waitoller and King Thorius, 2016), antiracism (Fitzgerald, 2020), and Disability Studies and Critical Race Theory (DisCrit) (Fornauf and Mascio, 2021). This push from stakeholders is propelling the continued iteration of UDL as a framework to support the design of more equitable learning environments.

CAST's *UDL Rising to Equity* is one example of an initiative that aims to respond to this growing call from the UDL field. *UDL Rising to Equity* is a transparent, inclusive, community-driven process to re-envision CAST's *UDL Guidelines* through an equity lens. Specifically, the process focuses on updating the guidelines to "address systemic barriers that result in inequitable learning opportunities and outcomes" (CAST, 2020). CAST is facilitating a collaborative, iterative process that includes a diversity of stakeholder experiences—including perspectives that may not have been considered in previous versions of the guidelines. *UDL Rising to Equity* focuses on fully developing the *UDL Guidelines* as a lever for designing learning environments where learners are honored, welcome, and valued—and as a lever for encouraging stakeholders to examine their own orientations and commitments toward equity.

UDL as an instructional design framework

In the past three decades, educators have discussed, debated, and developed ways to apply the guidelines to practice. The *UDL Guidelines* provide a basis and framework for inclusive instructional design. The 3 principles, 9 guidelines and 31 checkpoints that make up the UDL framework were derived from research (CAST, 2018a) and, when they were originally published, CAST provided detailed definitions of the guidelines and checkpoints as well as examples of how to apply the guidelines (Gravel and Rose, 2009; Rose and Gravel, 2010). However, in keeping with a foundational value of flexible use, CAST did not prescribe a particular way to implement the *UDL Guidelines*. Users of the *UDL Guidelines* were free to apply the principles in ways that were relevant to their students, settings, and needs. The inherent flexibility of the UDL framework is one of its greatest strengths, but also poses some challenges for UDL implementation efforts.

As a flexible framework that can be applied in many permutations and combinations to create inclusive educational environments, UDL is used in various ways and for varied purposes. In the past three decades, educators have applied UDL to curricula, courses, digital tools, and instructional environments. School districts have adopted UDL system-wide and states have recommended the integration of UDL. Researchers have applied UDL to a variety of practices and settings, using varied combinations of the guidelines and checkpoints (Rao et al., 2014). The literature on UDL includes applications of UDL in K12 and higher educational environments. Researchers have examined implementation of UDL in relation to designing inclusive environments (e.g., Kumar and Wideman, 2014; Lock et al., 2017; Scott and Temple, 2017), supporting student success in content areas (e.g., Gravel, 2018; King-Sears and Johnson, 2020; McMahon et al., 2016; Root et al., 2020), applications within digital environments (e.g., Daley et al., 2020; Dalton et al., 2011; Hall et al., 2015; Kennedy et al., 2014; Marino et al., 2014; Rappolt-Schlichtmann et al., 2013), and teacher education (e.g., Cunningham et al., 2017; Evmenova, 2018; Lanterman and Applequist, 2018; Wu, 2012). Researchers have also examined academic outcomes and socio-emotional outcomes (e.g., engagement, behavioral) in relation to the application of UDL (Ok et al., 2017).

The UDL Design Cycle: inclusive instructional design

In *Universal Design for Learning: Theory & Practice*, Meyer et al. (2014) present the core concepts that underlie the framework. First, they highlight the notion of learner variability, that every individual has different needs, abilities, backgrounds, and preferences. As described earlier in this chapter, they emphasize that learner variability is the norm in classrooms. Instead of designing for a mythical average learner, a focus of UDL is to design for all by integrating flexible options. UDL-based design also places a focus on reducing barriers in the learning environment and addressing learner variability through proactive and intentional development of flexible, supportive, and engaging learning environments (Nelson, 2021).

Educators can use UDL as part of an inclusive instructional design process. By considering UDL guidelines during the design process, it is possible to broaden the diversity of learners for whom instruction will work from the outset. UDL-based design starts with a consideration of instructional goals and learner variability. By first considering the knowledge needed to master instructional goals, educators can identify where potential barriers lie. Extending from that analysis, they can integrate flexible options and supports to address learner variability, including the varied needs, preferences, backgrounds and abilities of learners.

CAST originally defined how UDL can be applied to instructional goals, assessments, methods, and materials (Meyer et al., 2014). For educators interested in undertaking an inclusive instructional design process, Rao and Meo (2016) provided additional guidance and a systematic approach for designing with UDL during the lesson planning process. This approach, called the *UDL Design Cycle*, lays out a step-by-step approach for applying UDL to instructional goals, assessments, methods, and materials (see Fig. 1).

Similar to other instructional design processes, the *UDL Design Cycle* uses a backward design process starting with a learner analysis and linking instructional design decisions to the needs of learners. The *UDL Design Cycle* adds a uniquely inclusive focus to other models by emphasizing learner variability from the start. Teachers and instructional designers begin with a consideration of variability and analyze instructional goals in relation to potential barriers in the curriculum and learning environment. As they design instruction, decisions are made in each step to intentionally and proactively reduce and eliminate barriers and to provide supports and scaffolds that address learner variability. Table 1 depicts key considerations and questions to ask during the UDL based instructional design process.

The *UDL Design Cycle* provides a systematic way to consider socioemotional aspects of learning, in addition to cognitive aspects of learning. For example, when developing assessments (Step 3 of the cycle), educators can consider how to provide scaffolds that reduce threats for students by giving students opportunities to practice the format of assessments, providing low stake assessments (e.g., shorter weekly assessments instead of one high stakes test), and providing mastery-oriented feedback. When developing instructional methods (Step 4 of the cycle), the UDL framework provides guidance on reducing barriers in the socio-emotional realm as well by integrating supports that can help students persist and engage with learning. This includes strategies to increase relevance and authenticity, support meaningful interactions and collaborations, and promote self-regulation. Ultimately, UDL-based design focuses on creating inclusive learning environments that take into consideration the contexts for learning along with the content.

The last two steps of the *UDL Design Cycle* (Steps 5 and 6) address implementation of the lesson and reflection on whether the UDL-based elements addressed learner variability as intended. Designing with UDL can be an iterative process, building on the elements that worked well and adding flexible options and supports over time. For this reason, it is helpful to reflect on UDL-based lessons, retain the elements that supported learners, and include new strategies as needed to further reduce barriers.

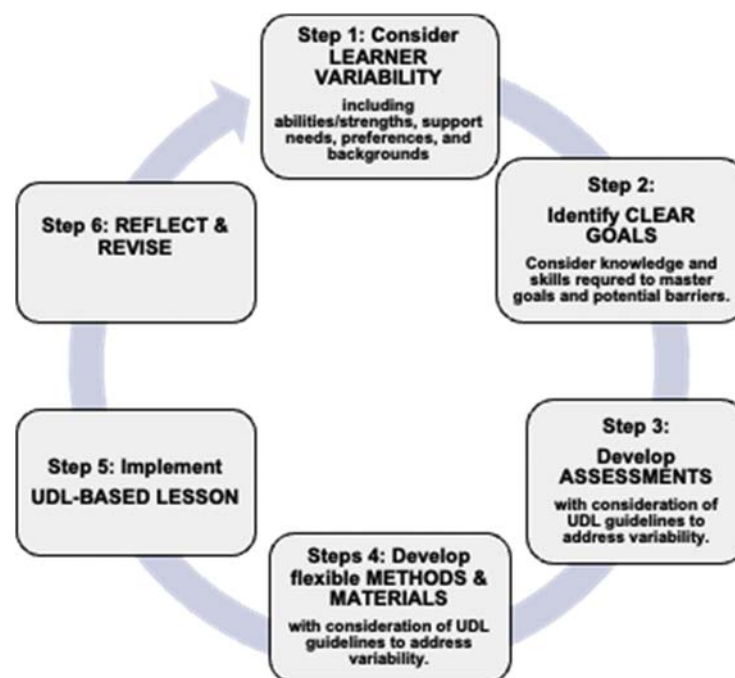


Fig. 1 UDL Design Cycle. Adapted from Rao and Meo (2016).

Table 1 Considerations and questions for the instructional design phase.

<i>UDL Design Cycle</i>	<i>Considerations and questions</i>
1. Consider learner variability	What is learner variability like in my class? Consider the following attributes of your learners: • Abilities/strengths • Backgrounds/experiences • Preferences/interest • Support needs
2. Identify clear goals	State 1–2 goals in clear and simple form. What knowledge and skills are required to meet these goals?
3. Develop assessments	What are potential barriers to reaching these goals? How can I use formative assessments that provide information on student mastery toward goals? How can I develop summative assessments that have construct relevance (measure knowledge rather than the format)?
4. Develop methods and materials	Taking UDL guidelines into consideration, plan strategies that address learner variability: • What strategies can I integrate to reduce barriers? • What scaffolds can I integrate for learners to use as needed? • How can I provide flexible options and choices in relation to lesson goals? • What materials/resources can I use to enhance flexibility, choice, and support options?

Flipped, blended and online: UDL for all formats

UDL-based design can be used to support inclusive practices in all instructional formats. Whether educators are designing for traditional classroom instruction or for online formats (e.g., blended, hybrid, or fully online), they can use the *UDL Design Cycle* to reduce barriers by integrating supports and scaffolds. When designing for online and blended learning, UDL can be especially useful as a way to make decisions about the most effective ways to use varied tools and formats (Rao, 2021). In online formats, UDL-based design can provide a systematic way to consider how to blend methods, for example, when and how to use asynchronous and synchronous opportunities to ensure that students have necessary supports for learning. For example, if a teacher is teaching online and uses both asynchronous and synchronous formats (e.g., online meetings using a system like Zoom or Google Meet), they can make instructional decisions on how to reduce barriers during asynchronous instruction (e.g., providing directions using screen-casts, using multimodal formats to deliver content) and how to use synchronous meetings to ensure that learners are engaged and motivated to learn.

The COVID-19 pandemic in 2020 resulted in a rapid shift to online learning worldwide. This highlighted the need for educators to understand how to proactively design learning experiences to support students in online learning environments. Educators worldwide had to quickly develop the skills to teach online, navigating learning management systems (e.g., Google Classroom, Blackboard), web-conferencing software (e.g., Google Meet, Zoom), and various digital tools available for teaching and learning (e.g., Padlet, Flipgrid, PearDeck, Quizlet, EdPuzzle). Using the *UDL Design Cycle*, teachers can make instructional decisions about how to use these types of digital tools meaningfully to support student learning in online environments.

Online learning brings with it new barriers that may not be as prominent in face-to-face learning environments. For example, in the online learning environment, students may work more independently without a teacher present as they learn. This results in fewer opportunities to clarify questions immediately. Teachers are not able to immediately observe and assist students just in time. To reduce these types of barriers, teachers can use UDL to make decisions about how and when to integrate digital tools and instructional strategies to provide supports that benefit students in online and hybrid learning environments. For example, using the features of digital text and e-books, teachers can support independent reading activities. Features such as text-to-speech, dual highlighting, and comprehension supports built into digital text environments are powerful ways to provide multiple means of representation. To give students multiple ways to express what they know, teachers can incorporate various multimodal tools that allow students to have a choice of how to demonstrate their knowledge (using video, audio, graphics, and text). These are just a few examples of the ways that digital tools can support diverse learners in the online learning environment (Kennedy and Boyle, 2021; Long et al., 2021; Rao et al., 2021; Speight and Kucharczyk, 2021).

Several studies have explored the application of UDL to online learning in higher education environments (i.e., Coy et al., 2014; Evmenova, 2018; Hollingshead, 2018; Lock et al., 2017; Tobin, 2014). In the post-secondary environment, UDL has been applied at a course level to provide students with options to persist and engage in the online-hybrid learning format. UDL has been used to support first-generation college students, students with disabilities, and students who are culturally and linguistically diverse. Using the *UDL Design Cycle*, instructors can design courses by breaking up a semester-long course into several modules. For each module, the instructor can identify key objectives and make decisions about the instructional strategies, digital tools, and assessment formats to use to provide options for students (Rao, 2021).

UDL Reporting Criteria

Because UDL can be applied for various purposes in ways that are flexible and localized to each setting, it is necessary for educators and researchers to clearly describe their applications of UDL in order to present models and allow others to replicate best practice with UDL. To provide a structure for describing UDL implementation, the Research Committee of the Universal Design for Learning Implementation and Research Network (UDL-IRN) developed and published the *UDL Reporting Criteria* (UDL RC) in 2019. Using the UDL RC, educators can describe the UDL-based aspects of lessons, programs, and/or environments. The UDL RC provide a checklist, organized into three essential areas of UDL based design:

1. Learner Variability and Environment
2. Proactive and Intentional Design
3. Implementation and Outcomes

The UDL RC were designed for use by many stakeholders, including practitioners, researchers, instructional designers, and curriculum developers, to describe their UDL-based lessons/curricula, environments and interventions. Researchers can use the RC as they design interventions, ensuring that core components are clearly articulated and able to be replicated and used by others. Teacher educators and academic coaches can also use the UDL RC as a tool for professional development on UDL. Educators can use the UDL RC to document their UDL-based planning processes and lesson implementation for review and discussion with peers and teacher educators. The UDL RC provide a structured way to describe how UDL was applied during the instructional design process and during implementation of a lesson or intervention, and what the related outcomes were, making it easier for others to understand how UDL was applied and to replicate or adapt models for their own settings.

Conclusion

Over the course of 20 years, UDL has grown from its origins within a small group of people interested in changing conceptions of ability/disability to a widely-accepted framework for inclusive education that is used throughout the world. At its core, UDL exhorts educators to consider their values around equity and inclusion and to infuse those values into the way they design educational environments. Most importantly, UDL encourages educators to think deeply about learning, diversity, and the goals of education. Not surprisingly, that means that the UDL guidelines themselves depend upon continual reflection and revision in order to adequately reflect the learning that emerges from their application. The wide application of UDL has revealed both strengths and weaknesses of the existing guidelines.

As mentioned in this chapter, CAST has spearheaded efforts to review, revise, and update the guidelines in the light of what we have learned so far. For example, CAST's *UDL Rising to Equity* has been one way to ensure that future versions of the UDL guidelines are guided by a much more representative group than previous iterations. The process integrates a transparent, community-driven structure to re-envision the next version of the guidelines through an equity lens and to emphasize the guidelines as a living, dynamic tool, that depends on iterative feedback from a diversity of stakeholders. It is this kind of inclusive collaboration that will sustain the growth of UDL as a lever for designing equitable learning environments where learners can thrive. We conclude by sharing our predictions and visions for the future of UDL.

First, although the UDL guidelines were timely and innovative at their origin, ongoing efforts will ensure that the guidelines remain a tool that is reflective of current research. These continual updates include three major areas: (a) advances in the affective and cognitive neurosciences with special focus on the bases for individual variability in ability and disability; (b) advances in evidence-based pedagogies that have proven effective for students who have been marginalized or excluded altogether (including research on UDL practices); and (c) advances in learning technologies, especially those adaptive learning technologies that can not only teach but also learn (e.g., can learn about individual students and their optimal pathways for success).

Second, UDL was originally developed within the limited context of the USA. Although many innovative individuals have applied UDL throughout the world, the guidelines provide for them a foundation that is too narrow and rigid, particularly in their research basis and ultimately their recommendations. The rich experience derived from applying UDL internationally provides a broader and stronger foundation for UDL, making it more suitable for international application but also more flexible and inclusive to meet the growing diversity that characterizes the USA. There is much to be learned from international partners who apply UDL in various contexts; their experiences support the continued evolution of the guidelines, adding a global perspective on its application to promote inclusion and equity.

Third, although the UDL guidelines address a wide range of potential barriers to learning, they do not address all of the barriers and they do not address the barriers that are most problematic for some learners. Pointedly, for the most part, the UDL guidelines address barriers based on ability but not identity. For some students, the barriers they face have little to do with their actual abilities and much more to do with other people's perceptions (biases) about their abilities (and motivation, etc.) based on their skin color, their sexual orientation or gender, their age, their ethnic origin, etc. These implicit biases, and the barriers they (often inadvertently) erect, are much harder to address than explicit barriers. Addressing them will require thoughtful collaborations with other individuals and organizations for whom barriers of identity have already been a primary focus. Because the two kinds of barriers inevitably intersect, such collaborations will ensure that the UDL guidelines are more carefully articulated to be effective for everyone.

Fourth, the UDL guidelines were focused primarily on the barriers that arise within an individual classroom or school curriculum. But many barriers are systemic, institutional, even broadly cultural. Classrooms and curricula are inevitably embedded within such larger entities which impose various implicit and explicit biases and barriers. Since most institutions are, like classrooms, learning organizations, the UDL guidelines should apply to the barriers they impose.

Continued iteration and development are critical to the future success of UDL. We look forward to this third decade as a time of ongoing learning from stakeholders—from both within the current UDL community as well as from stakeholders whose perspectives and experiences have not yet been considered.

References

- Alim, H.S., Baglieri, S., Ladson-Billings, G., Paris, D., Rose, D.H., Valente, J.M., 2017. Responding to “cross-pollinating culturally sustaining pedagogy and universal design for learning: toward an inclusive pedagogy that accounts for dis/ability”. *Harv. Educ. Rev.* 87 (1), 4–25.
- Baglieri, S., Lalvani, P., 2019. *Undoing Ableism: Teaching About Disability in K-12 Classrooms*. Routledge.
- Baglieri, S., Bejorian, L., Broderick, A., Connor, D., Valle, J., 2011. [Re]claiming “inclusive education” toward cohesion in educational reform: disability studies unravels the myth of the normal child. *Teach. Coll. Rec.* 113 (10), 2122–2154.
- Bartholomew Consolidated School Corporation (BCSC), 2021. Universal Design for Learning. Retrieved from. <https://www.bcscschools.org/udl>.
- CAST, 2018a. Research Evidence. Retrieved from. <https://udlguidelines.cast.org/more/research-evidence>.
- CAST, 2018b. Universal Design for Learning Guidelines Version 2.2. Retrieved from. <http://udlguidelines.cast.org>.
- CAST, 2020. CAST Announces a Community-Driven Process to Update UDL Guidelines. Retrieved from. <https://www.cast.org/news/2020/community-driven-process-update-udl-guidelines>.
- CAST, 2021. UDL in Public Policy. Retrieved from. <https://www.cast.org/impact/udl-public-policy>.
- CAST, 2021a. About Universal Design for Learning. Retrieved from. <https://www.cast.org/impact/universal-design-for-learning-udl>.
- CAST, 2021b. The UDL Guidelines. Retrieved from. <https://udlguidelines.cast.org/>.
- Causton-Theoharis, J., Theoharis, G., Orsati, F., Cosier, M., 2011. Does self-contained special education deliver on its promises? A critical inquiry into research and practice. *J. Spec. Educ. Leader* 24 (2), 61–78.
- Coy, K., Marino, M.T., Seriani, B., 2014. Using universal design for learning in synchronous online instruction. *J. Spec. Educ. Technol.* 29 (1), 63–74. <https://doi.org/10.1177/016264341402900105>.
- Cunningham, M.P., Huchting, K.K., Fogarty, D., Graf, V., 2017. Providing access for students with moderate disabilities: an evaluation of a professional development program at a Catholic Elementary School. *J. Catholic Educ.* 21 (1). <https://doi.org/10.15365/joce.2101072017>.
- Daley, S.G., Xu, Y., Proctor, C.P., Rappolt-Schlichtmann, G., Goldowsky, B., 2020. Behavioral engagement among adolescents with reading difficulties: the role of active involvement in a universally designed digital literacy platform. *Read. Writ. Q.* 36 (3), 278–295. <https://doi.org/10.1080/10573569.2019.1635545>.
- Dalton, B., Proctor, C.P., Uccelli, P., Mo, E., Snow, C.E., 2011. Designing for diversity: the role of reading strategies and interactive vocabulary in a digital reading environment for fifth-grade monolingual English and bilingual students. *J. Literacy Res.* 43 (1), 68–100. <https://doi.org/10.1177/1086296X103978732>.
- Evmenova, A.S., 2018. Preparing teachers to use Universal Design for Learning to support diverse learners. *J. Online Learn. Res.* 4, 147–171.
- Fornauf, B.S., Mascio, B., 2021. Extending DisCrit: a case of universal design for learning and equity in a rural teacher residency. *Race Ethn. Educ.* 1–16. <https://doi.org/10.1080/13613324.2021.1918409>.
- Fritzsche, A., 2020. *Antiracism and Universal Design for Learning: Building Expressways to Success*. CAST Professional Publishing.
- Fullan, M., 2007. Change the terms for teacher learning. *J. Staff Dev.* 28 (3), 35–36. <http://michaelfullan.ca/wp-content/uploads/2016/06/13396074650.pdf>.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., Yoon, K.S., 2001. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38 (4), 915–945. <https://doi.org/10.3102/00028312038004915>.
- Garner, A., Lipsky, D.K., 1987. Beyond special education: toward a quality system for all students. *Harv. Educ. Rev.* 57 (4), 367–396. <https://doi.org/10.17763/haer.57.4.kj517305m7761218>.
- Gravel, J.W., Rose, D.H., 2009. Getting from here to there: UDL, global positioning systems, and lessons for improving education. In: Gordon, D.T., Gravel, J.W., Schifter, L.A. (Eds.), *A Policy Reader in Universal Design for Learning*. Harvard Education Press, pp. 5–18.
- Gravel, J.W., 2018. “Going deep”: leveraging Universal Design for Learning to engage all learners in rich disciplinary thinking in ELA. *Teach. Coll. Rec.* 120 (3), 1–40.
- Hall, T.E., Cohen, N., Vue, G., Ganley, P., 2015. Addressing learning disabilities with UDL and technology: strategic reader. *Learn. Disabil. Q.* 38 (2), 72–83. <https://doi.org/10.1177/0731948714544375>.
- Hehir, T., 2002. Eliminating ableism in education. *Harv. Educ. Rev.* 72 (1), 1–32. <https://doi.org/10.17763/haer.72.1.03866528702g2105>.
- Hollingshead, A., 2018. Designing engaging online environment: universal Design for Learning principles. In: Milheim, K.L. (Ed.), *Cultivating Diverse Online Classrooms Through Effective Instructional Design*. IGI Global.
- Kennedy, M.J., Boyle, J.R., 2021. That really escalated quickly—online learning moves into the mainstream: introduction to the special issue. *J. Spec. Educ. Technol.* <https://doi.org/10.1177/01626434211006052>.
- Kennedy, M.J., Thomas, C.N., Meyer, P., Alves, K.D., Lloyd, J.W., 2014. Using evidence-based multimedia to improve vocabulary performance of adolescents with LD: a UDL approach. *Learn. Disabil. Q.* 32 (2), 71–86. <https://doi.org/10.1177/0731948713507262>.
- King-Sears, M.E., Johnson, T.M., 2020. Universal design for learning chemistry instruction for students with and without learning disabilities. *Remedial Spec. Educ.* 41 (4), 207–218. <https://doi.org/10.1177/0741932519862608>.
- King-Sears, M.E., 2020. Introduction to special series on universal design for learning. *Remedial Spec. Educ.* 41 (4), 191–193.
- Knapp, M.S., 2003. Professional development as a policy pathway. *Rev. Res. Educ.* 27, 109–157. <https://doi.org/10.3102/0091732X027001109>.
- Kumar, K.L., Wideman, M., 2014. Accessible by design: applying UDL principles in a first year undergraduate course. *Can. J. High Educ.* 44 (1), 125–147. <https://www.learntechlib.org/p/157510>.
- Lanterman, C.S., Applequist, K., 2018. Pre-service teachers’ beliefs: impact of training in Universal Design for Learning. *Except. Educ. Int.* 28 (3), 102–121.
- Lock, J., Altowairiki, N.F., Hill, S.L., Johnson, C., 2017. A journey through the development of online environments: putting UDL theory into practice. In: *Handbook of Research on Innovative Pedagogies and Technologies for Online Learning in Higher Education*, pp. 218–235.
- Long, H.M., Bouck, E.C., Jakubow, L.N., 2021. Explicit instruction in mathematics: considerations for virtual learning. *J. Spec. Educ. Technol.* 36 (2), 67–76. <https://doi.org/10.1177/0162643421994099>.
- Luria, A.R., 1976. *The Working Brain: An Introduction to Neuropsychology*. Penguin Books.
- Marino, M.T., Gotch, C.M., Israel, M., Vasquez, E., Basham, J.D., Becht, K., 2014. UDL in the middle school science classroom: can video games and alternative text heighten engagement and learning for students with learning disabilities? *Learn. Disabil. Q.* 37 (2), 87–99. <https://doi.org/10.1177/0731948713503963>.

- McMahon, D.D., Cihak, D.F., Wright, R.E., Bell, S.M., 2016. Augmented reality for teaching science vocabulary to postsecondary education students with intellectual disabilities and autism. *J. Res. Technol. Educ.* 48 (1), 38–56. <https://doi.org/10.1080/15391523.2015.1103149>.
- Meyer, A., Rose, D.H., 2005. The future is in the margins: the role of technology and disability in educational reform. In: Rose, D.H., Meyer, A., Hitchcock, C. (Eds.), *The Universally Designed Classroom: Accessible Curriculum and Digital Technologies*. Harvard Education Press, pp. 13–35.
- Meyer, A., Rose, D.H., Gordon, D.T., 2014. *Universal Design for Learning: Theory and Practice*. CAST Professional Publishing.
- Nelson, L.L., 2021. *Design and Deliver: Planning and Teaching Using Universal Design for Learning*, second ed. Brookes Publishing.
- Ok, M.W., Rao, K., Bryant, B.R., McDougall, D., 2017. UDL in the preK-12 classroom: a systematic review of research. *Exceptionality* 25 (2), 116–138. <https://doi.org/10.1080/09362835.2016.1196450>.
- Piaget, J., 1952. *The Origins of Intelligence in Children*. W.W. Norton & Company, New York, NY.
- Rao, K., Meo, G.J., 2016. Using universal design for learning to design standards-based lessons. *Sage Open* 6 (4), 1–12. <https://doi.org/10.1177/2158244016680688>.
- Rao, K., Ok, M.W., Bryant, B.R., 2014. A review of research on universal design educational models. *Remedial Spec. Educ.* 35 (3), 153–166. <https://doi.org/10.1177/0741932513518980>.
- Rao, K., Torres, C., Smith, S.J., 2021. Digital tools and UDL-based instructional strategies to support students with disabilities online. *J. Spec. Educ. Technol.* <https://doi.org/10.1177/0162643421998327>.
- Rao, K., 2021. Inclusive instructional design: applying UDL to online learning. *J. Appl. Instr. Des.* 10 (1). <https://doi.org/10.51869/jaid2021101>.
- Rappolt-Schlichtmann, G., Daley, S.G., Lim, S., Lapinski, S., Robinson, K.H., Johnson, M., 2013. Universal design for learning and elementary school science: exploring the efficacy, use, and perceptions of a web-based science notebook. *J. Educ. Psychol.* <https://doi.org/10.1037/a0033217>.
- Root, J.R., Cox, S.K., Saunders, A., Gilley, D., 2020. Applying the Universal Design for Learning Framework to mathematics instruction for learners with extensive support needs. *Remedial Spec. Educ.* 41 (4), 194–206. <https://doi.org/10.1177/0741932519887235>.
- Rose, D.H., Gravel, J.W., 2010. Universal design for learning. In: Peterson, P., Baker, E., McGraw, B. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, pp. 119–124.
- Rose, D.H., Gravel, J.W., 2013. Using digital media to design student-centered curricula. In: Wolfe, R.E., Steinberg, A., Hoffmann, N. (Eds.), *Anytime, Anywhere: Student-Centered Learning for Students and Teachers*. Harvard Education Press, pp. 77–101.
- Rose, D.H., and Meyer, A., 2002. *Teaching every student in the digital age: Universal Design for Learning*. ASCD.
- Scott, L., Temple, P., 2017. A conceptual framework for building UDL in a special education distance education course. *J. Educ. Online* 14 (1).
- Shakespeare, T., 2010. The social model of disability. In: Davis, L. (Ed.), *The Disability Studies Reader*. Routledge, pp. 266–273.
- Speight, R., Kucharczyk, S., 2021. Leveraging positive behavior supports to improve engagement in virtual settings. *J. Spec. Educ. Technol.* 36 (2), 90–96. <https://doi.org/10.1177/0162643421992704>.
- Tobin, T.J., 2014. Increase online student retention with universal design for learning. *Q. Rev. Dist. Educ.* 15 (3), 13–24.
- UDL-IRN, 2022. Universal Design for Learning Implementation and Research Network. <https://udl-irn.org>.
- UDL on Campus, 2021. Post-secondary Institutions With UDL Initiatives. Retrieved from. <http://udloncampus.cast.org/home>.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Waitoller, F.R., King Thorius, K.A., 2016. Cross-pollinating culturally sustaining pedagogy and universal design for learning: toward an inclusive pedagogy that accounts for disability. *Harv. Educ. Rev.* 86 (3), 366–389. <https://doi.org/10.17763/1943-5045-86.3.366>.
- Wood, D.J., Bruner, J.S., Ross, G., 1976. The role of tutoring in problem solving. *J. Child Psychiatr. Psychol.* 17 (2), 89–100.
- Wu, X., 2012. Promoting interface and knowledge sharing: a joint project between general and special education preservice teachers. *Inq. Educ.* 3 (2). <https://digitalcommons.nli.edu/ie/vol3/iss2/4>.

Well-being and learning

Tina Hascher, University of Bern, Institute of Educational Science, Department of Research in School and Instruction, Bern, Switzerland

© 2023 Elsevier Ltd. All rights reserved.

This is an update of T. Hascher, Wellbeing, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 732–738, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00633-3>.

Understanding well-being	721
Well-being in psychology and education	721
Predictors and outcomes of well-being	722
Well-being and learning in school	723
Understanding student well-being	723
What does empirical evidence tell about student well-being?	724
Predictors of student well-being	724
Understanding teacher well-being	725
Predictors and outcomes of teacher well-being	726
Supporting student and teacher well-being in school	726
Future perspective on well-being and learning	727
References	727

Since it came into discussion in the early 1970s, the idea of well-being has been strongly promoted as an individual and societal value as well as a basic psychological need (Biswas-Diener et al., 2005; Diener and Lucas, 2000; OECD, 2017; Ryan and Deci, 2001; Strack et al., 1991). Well-being as an indicator of quality of life is important for human health, development, and flourishing. The basic objectives of well-being research are to identify the core elements, predictors, and outcomes of well-being and to test programs that aim at improving well-being. Well-being emerged as a research topic of multi- and transdisciplinary interest as it is expected to influence various dimensions of individual and societal life. Specifically, positive psychology has dedicated itself to research in well-being (Linley et al., 2006). As regards well-being and learning, major contributions are acknowledged from disciplines that focus on schools, learning and instruction such as educational sciences and educational psychology. Understanding learners' needs for both academic achievement and well-being is crucial for 21st century education (Seligman et al., 2009).

Understanding well-being

The variety of disciplines that investigate well-being has led to different research approaches and definitions of well-being (e.g., Diener, 1984; Hascher and Waber, 2021; Larsen and Diener, 1987; Veenhoven, 1991). Across this variety, four general lines of the conceptualization of well-being can be identified: (1) Well-being is defined as a specific emotional quality of feeling well. (2) Well-being is used as a supra-term for various positive emotions such as enjoyment, pride, satisfaction etc. (3) Well-being is used as a supra-term for the absence of diverse negative emotions, psychological and physiological variables such as anxiety, depression, exhaustion, or burnout. (4) Well-being is defined as a multi-dimensional concept combining specific positive and negative dimensions (e.g., satisfaction and positive emotions combined with an absence of emotional exhaustion).

Well-being in psychology and education

As a psychological term, well-being usually refers to a subject's own feelings and cognitive evaluations and is defined as a multi-dimensional construct (Ryff and Keyes, 1995). A set of dimensions covering emotions and evaluations are considered to form well-being as can be seen in the PERMA-model of Seligman (2002), which consists of positive emotion, engagement, relationships, meaning, and accomplishment. For example, if individuals experience happiness or pride and evaluate their actual situation positively, they are expected to feel well; on the contrary, if individuals experience sadness or anger and evaluate the situation negatively, they experience a low level of well-being. Feelings and evaluations depend on a person's previous experiences, expectations, goals, and values. Thus, reasons for positive or negative emotions and evaluations may differ across individuals. Well-being also consists of both mental and physical components; that is, a person's thoughts as well as body sensations have to be considered. Defined this way, well-being differs significantly from a single feeling of satisfaction or from a singular emotional experience like enjoyment.

Along with the multi-dimensional character of well-being, different forms of well-being are discussed. According to the World Health Organization's definition of health, well-being can be specified as a psychological, physical, or social form (e.g., WHO, 2000). Ryan and Deci (2001), for example, suggested a distinction between hedonic well-being that is predominantly related to

positive affective experience and need fulfilment and eudemonic well-being that is based on the experience and pursuit of a meaningful life. Also, forms of domain-specific well-being can be differentiated, such as workplace well-being that is defined as consisting of organizational well-being and subjective well-being (Burns and Machin, 2013).

Due to the co-existence of positive and negative events in life, well-being can be described by the interaction of positive and negative aspects: The more positive life is in general or the more positive specific relevant life domains are experienced, the better the well-being; negative feelings and negative evaluations of important circumstances or events in one's life result in lower levels of well-being. Thus, to understand and explain a person's well-being, positive and negative dimensions must be taken into account simultaneously, and well-being can be understood as a "positive imbalance" that represents a dominance of positive emotions and evaluations over negative emotions and evaluations (Hascher et al., 2021). The more pronounced this difference between the positive and negative dimensions, the higher the well-being.

Well-being can develop in a short- or long-term perspective (e.g., Kim-Prieto et al., 2005). In the short-term, actual well-being is situationally determined and an expression of a predominance of current positive emotions and cognitions while experiencing a situation. This might include, for example, students feeling well while working in pairs on a problem-solving task during mathematics class or teachers feeling well while collaborating with a colleague in preparing a set of lessons. In the long-term, habitual well-being is sustainable and an indicator of a more general, subjective evaluation of life situations and circumstances—for example, feeling well in school overall apart from single situations that are negatively experienced. This means, repeated positive experiences such as feeling enjoyment, mastering academic challenges, or feeling competent in school combined with a low number of negative occurrences such as feeling stressed, losing control, or feeling under pressure will lead to well-being in school. Short-term and long-term well-being are related to each other as habitual well-being develops through frequent experiences of short-term well-being, and short-term well-being is determined by the frame of habitual well-being. However, habitual well-being cannot be seen as a simple or accurate aggregation of actual well-being. This is because specific experiences of actual well-being can contribute differently to habitual well-being, in that actual well-being in situations of special importance may have a stronger impact on well-being than situations of lower relevance.

For the assessment of well-being, an array of quantitative and qualitative instruments exists, ranging from affect-balance-scales to well-being questionnaires to event sampling methods and happiness interviews (Andrews and Robinson, 1991; Diener, 1984; Diener and Biswas-Diener, 2000; Hascher and Waber, 2021; Lucas et al., 1996). Of specific importance for the accurate assessment of well-being is a multi-faceted approach. Various dimensions of student well-being, for example, the emotional and cognitive as well as the positive and negative, need to be considered. Single items are less valid for measuring the complexity of well-being, and multi-method assessments which explicitly integrate different perspectives on well-being are recommended.

Predictors and outcomes of well-being

Well-being is sensitive to many factors. The more general well-being is defined and measured the more sources can influence it (Diener, 1984). Although researchers have attempted to specify predictors for the different components of well-being, research so far has only limited been successful in uncovering the most potent variables. However, social relationships turn out to play a crucial role for the development of well-being. For example, social support can enhance an individual's functioning and serve as a buffer against adverse experiences (e.g., Chu et al., 2010). Based on Bronfenbrenner's paradigm (1979) that the individual is embedded within and interacts with different ecological systems and the notion that this interplay determines well-being, three main categories of predictors have been identified:

- (1) Well-being is influenced by living conditions such as an individual's financial or health status that are of relevance for a general evaluation of life and contribute directly to well-being. Other contextual factors such as an organizational or school culture might be intertwined with specific resources (e.g., positive social interactions), which influence the fulfilment of goals and aspirations and, as a result, well-being (e.g., Diener et al., 1995).
- (2) Personal characteristics can predict well-being. In addition to gender, the role of introversion versus extraversion for well-being has been identified (e.g., Vittersø and Nilsen, 2002). Extraverted individuals seem to experience more positive situations in life and, thus, feel better than more introverted persons (bottom-up approach). Also, extraverted individuals might interpret life more positively, react in a more positive way to life events and circumstances, and feel better than more introverted individuals (top-down approach).
- (3) Subjective and temporal differences have been shown to affect how an event is evaluated and interpreted. A person's evaluation might be based on various points of reference (e.g., Stein et al., 1997), such as social comparisons, individual goals and expectancies, processes of adaptation to new life-circumstances, coping with challenging situations, control beliefs or causal attributions.

As with the array of predictors, the outcomes of well-being are recognized to be manifold, with strong support for the idea that well-being can contribute to successful adjustment, positive development, and a meaningful life, whereas a lack of well-being may be harmful for the individual (e.g., Wood and Josephs, 2010). Regarding the relevance of well-being for learning, the idea that well-being supports learning and development across the lifespan has been broadly accepted (Hascher, 2010). According to the Broaden-and-Built Theory (Fredrickson, 2001), well-being is expected to enable positive experiences that, like positive emotions, support cognitive functioning and learning. Although there might be varying effects due to contextual factors, studies have confirmed that well-being affects various aspects of learning and achievement (e.g., Abele et al., 2005). Positive mood might support

more creative, holistic, flexible thinking, and might be associated with better performance. Empirical findings also suggest that positive learning emotions are aligned with factors that influence learning and achievement, such as academic-ability concept, academic interest and engagement, time-management strategies, and meta-cognition (e.g., Efklides and Petkaki, 2005; Reschly et al., 2008). These positive associations are also mirrored in the concept of *flourishing* that more generally represents high levels of well-being (Seligman, 2011). Regarding educational settings such as schools, research reveals that not only negative but also positive emotions deserve closer attention because of their beneficial effects on learning and achievement as well as reciprocal effects between emotions and students' and teachers' learning and achievement (Frenzel, 2014; Pekrun and Linnenbrink-Garcia, 2012).

Well-being and learning in school

During recent years, student and teacher well-being have gained increasing attention. There are several reasons for this. First and foremost, well-being has been recognized as an important aspect of successful education and individual growth, as learning is based on an interplay of cognitive, motivational, and emotional factors. Second, schools have increasingly developed as caring communities for students and teachers that aim to support individual growth as well as social skill development by incorporating well-being as a goal of school improvement. Third, well-being has been acknowledged as a protective source in facing and coping with societal challenges that possibly increase psychological and emotional problems (Green et al., 2012). With regard to the multi-dimensional nature of well-being, well-being in school can be described by several components that equally can be applied to student and teacher well-being (see Fig. 1).

For students and teachers alike, well-being in school may be viewed as a subjective, emotional, and cognitive evaluation of school reality (Van Petegem et al., 2006). Well-being in school is seen as an imbalance resulting from positive emotions and cognitions related to the school context that outperform negatively experienced and evaluated aspects of the school context.

Understanding student well-being

Student well-being has been recognized as a pedagogical aim in educational theory and practice. However, there has been some controversy around the issue. Sceptics have voiced doubt about the solely positive function of well-being and asked if student well-being could be harmful (Boekaerts, 1993). They stress the likelihood of neglecting the attainment of academic goals for the benefit of a positive atmosphere in school. This scepticism is unjustified, as there are several strong reasons for maintaining and fostering well-being in school (Hascher and Hagenauer, 2020). Apart from the idea that well-being in school serves as a value in itself, it has been found that student well-being is an indicator of successful psychosocial adjustment to school (Buyse et al., 2009). Also, there is empirical evidence that student well-being positively affects students' learning processes and learning outcomes (see Hagenauer and Hascher, 2014; Hascher, 2011; Morinaj and Hascher, 2019). Although student well-being might not directly enhance an individual's achievement, it is intertwined with various factors that contribute to effective learning and is an important criterion for a positive school climate which is necessary for students' learning—particularly in a highly structured, achievement-oriented, non-optional learning context such as a school (UNICEF Office of Research, 2017). Student well-being can also serve as a resource for coping with failure as well as social and academic problems. Students who feel well in school may also transfer their positive attitudes toward learning and achievement in school to learning and achievement in general. Furthermore, student well-being is a crucial indicator of children's and adolescents' health status (Inchley et al., 2016) and mirrors school quality and effectiveness (Bradley et al., 2018).

Among studies of student well-being, various views as to how to investigate student well-being can be found. One important difference is how precise and detailed student well-being is defined and how strongly it is related to the school setting. An overview of the research reveals five different approaches to student well-being (see also Hascher, 2008):

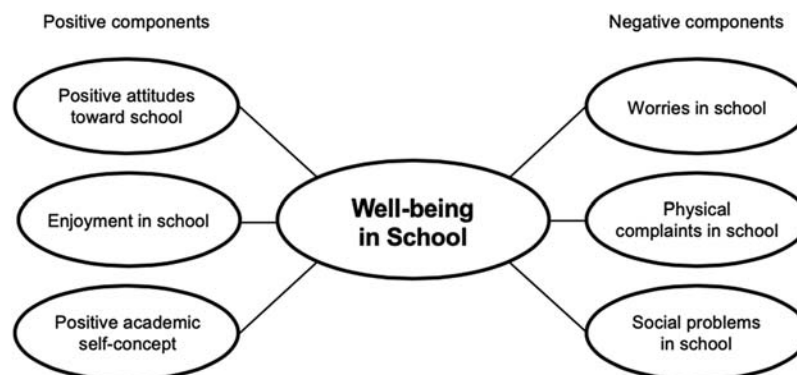


Fig. 1 Six dimensions of well-being in school.

- (1) Students are asked about aspects of their general mood states, mostly negative states like depression (e.g., Benjet and Hernández-Guzmán, 2002; Undheim and Sund, 2005). Although student well-being is not directly addressed and school context might be neglected, the results can help to elucidate how students generally feel.
- (2) Students are asked to report their well-being by using a general well-being scale for children or adolescents (e.g., Jin and Moon, 2006; Huang and Chien, 2013; Konu et al., 2002). Both dimensions of well-being, positive (e.g., enjoyment) and negative (e.g., worries), might be addressed, and the results can contribute to an understanding of well-being in young individuals.
- (3) Well-being is operationalized by a selective combination of various emotional, social, and cognitive school-related variables. These include, for example, test anxiety, social integration into the classroom, and satisfaction in school (Huebner et al., 2001). Results from these studies provide insight into school-specific psychological phenomena that are related to student well-being.
- (4) Student well-being is defined and investigated as a specific quality in the school setting. This research explicitly addresses student well-being, school well-being, or well-being in school (e.g., Engels et al., 2002; Hascher, 2003; Hascher and Hagenauer, 2020; Morinaj and Hascher, 2021) and is grounded in a theory of psychological well-being. With its explicit focus on the school environment, this research can be seen as a domain specific approach to student well-being. The strengths of these studies lie in their contribution to a deeper understanding of student well-being, its correlates and predictors.
- (5) In terms of the WHO definition of health, student well-being is discussed under the broader perspective of health and health promotion (e.g., Konu and Linthonen, 2005, 2006). Related to school as an important context for children and adolescents, these studies contribute to an understanding of student well-being as a crucial component of general physical and mental health.

What does empirical evidence tell about student well-being?

In asking the question about well-being in school, an interesting contradiction can be found: Repeatedly, it has been shown that school can turn into an aversive context and central source of stress for children and adolescents (e.g., OECD, 2017; Undheim and Sund, 2005). Accordingly, when asked about daily experiences in school, students report a dominance of negative situations and emotions (Hascher, 2007, 2008). However, when asked more generally about well-being in school, students in many countries indicate average, or even high, scores (OECD, 2017). One possible explanation for this discrepancy is that the results depend on the way students are asked about their school. To report about everyday events in school may lead to a focus on negative aspects because students feel a stronger need to communicate them. Conversely, a questionnaire on student well-being that stresses negative as well as positive components is better able to take both sides of school into consideration simultaneously. Another explanation is given by the theoretical assumption that student well-being is more than simply a compilation of daily events. A more holistic view on school leads to different, more positive, results than a focus on singular experiences.

Independent of the method of measurement, low to moderate inter-individual differences in student well-being have been empirically confirmed (see Hascher and Hagenauer, 2020; Konu and Linthonen, 2006; Morinaj and Hascher, 2021). High achieving students feel better in school than low achievers. Girls report higher scores on well-being but are also more likely to describe higher scores on negative components like physical complaints, in comparison to boys. Younger students feel better in school than students in higher grades; however, toward the end of compulsory education, an increase of student well-being can be found. It has to be noted, that these differences in student well-being cannot be attributed exclusively to individual characteristics, as learning conditions in terms of the school setting, instructional quality, and teaching subjects are likewise relevant (see Opdenakker and Van Damme, 2000; Van Petegem et al., 2006). Schools have been shown to differ in their level of student well-being, indicating the relevance of school culture variables. Also, students report better well-being scores in schools with a higher academic ranking in comparison to schools with a lower academic ranking. Higher levels of student well-being are reported in lessons with a predominant student orientation in comparison to teacher-centred lessons. Student well-being differs across subjects showing that highly selective subjects are more often associated with lower well-being than subjects with lower achievement pressure. With regard to the sub-dimensions of student well-being, interesting differences can be found (e.g., see Hascher and Hagenauer, 2020; Van Petegem et al., 2007). Students report low degrees of physical complaints and equally low degrees of social problems. However, worries in school are frequent, and students often report a lack of enjoyment in school. Students' intrinsic motivation correlates positively with student well-being, whereas test anxiety correlates negatively with student well-being.

Predictors of student well-being

According to research on general well-being, there might be an array of sources that contribute to student well-being (see Fig. 2). More systematically, predictors of student well-being can be divided into three main groups:

- (1) Student well-being is influenced by *contextual* conditions, primarily by the school and learning environment. Characteristics of school structures as well as school-specific variation in terms of action plans in school, school culture, pedagogical orientation, infrastructure, and facilities are expected to have a distinct impact on student well-being. Situation-specific classroom factors such as the quality of instruction, facilities, participation in the classroom or the fulfillment of basic needs are of comparable importance. Teachers, as well as classmates, play an essential role in student well-being (Hascher, 2003; Van Petegem et al., 2007). Teachers' pedagogical and social competencies and a supportive teaching style, along with high quality teacher-student relationship, have been shown to be predictive of student well-being (Hagenauer et al., 2015), whereas pressure to achieve, lack of instructional quality along with social isolation and a problematic teacher-student-relationship can be harmful for student well-being (Baumeister and Leary, 1995; Hall-Lande et al., 2007; OECD, 2017). Similarly, the social support of peers, a cohesive

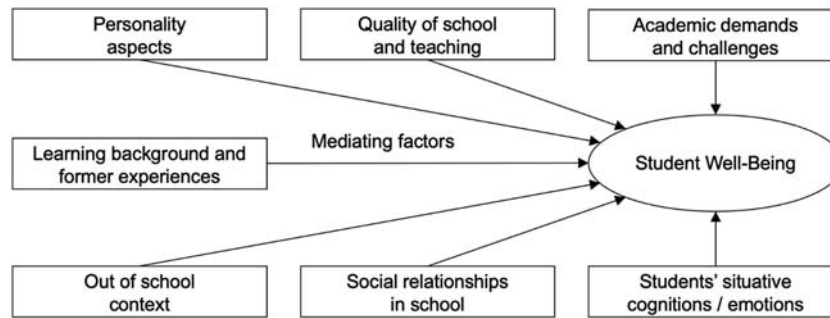


Fig. 2 Predictors of student well-being.

classroom, and positive interactions during breaks foster student well-being by promoting positive attitudes toward school and student academic self-concept and by preventing social problems in the classroom. Social discrimination in the classroom or rejection by peers, however, reduces student well-being by increasing social problems in school. Additionally, students' living conditions, family structure, parental academic orientation, and school orientation of peers can influence student well-being (Martin, 2005).

- (2) Student well-being can be predicted to some degree by personal characteristics. In regards to gender, it has been repeatedly found that girls express higher scores of positive as well as negative dimensions of student well-being. This may be because they experience student well-being more intensively or because they are more open to reporting their student well-being (Brebner, 2003). There is also evidence that predictors of student well-being show gender-specific relevance (Bergman and Scott, 2001; Liu et al., 2016; Løhre et al., 2014). In addition to gender, school-related factors such as engagement and effort, self-efficacy and control beliefs, goal orientation or tolerance to frustration might influence student well-being, as they can contribute to students' ability to manage academic challenges. Factors that support success in school also help to foster student well-being.
- (3) Although some context and personality variables might directly influence student well-being, most effects may result from person-environment interaction. The interaction is represented by subjective differences in how school-related conditions and events are evaluated and interpreted. For instance, differences in student evaluations can result from social comparisons, processes of adaptation to the school situation, personal interests, coping with challenging situations, causal attributions, etc. These variables affect how school is perceived and how an individual's situation is related to the context, and they are powerful for students' self-evaluations. These variables can have a direct impact on student well-being as well as serve as mediators between school conditions, personal characteristics, and student well-being. Accordingly, the alignment of individual goals, values, and expectancies with the school situation seems to be crucial for the development of student well-being—the better the alignment, the higher student well-being.

Understanding teacher well-being

Teacher well-being in school as an indicator of the quality of school life has also been considered as a crucial factor for teachers' functioning and mental health. Teacher well-being is acknowledged as a major driver of effective teaching (Duckworth et al., 2009) and a preventive resource against emotional exhaustion, burnout and job dismissals (McCallum et al., 2017; Renshaw et al., 2015; Skaalvik and Skaalvik, 2018). Among the plethora of empirical studies on teacher well-being (for an overview, see Hascher and Waber, 2021), various aspects of psychological functioning such as job satisfaction, intrinsic motivation, flow, and emotions at work or dysfunctional factors such as exhaustion, burnout or stress at work are studied as indicators of teacher well-being.

Among the various approaches, job satisfaction has often been considered a main indicator or core element of teacher well-being, although it is frequently defined as a cognitive job attitude, neglecting the emotional dimension of well-being (e.g., Veldman et al., 2013). It has been found that teachers experiencing high satisfaction differ from less satisfied teachers in terms of subjective health and thoughts about leaving the job. Low satisfaction seems to go along with stress and unsuccessful coping strategies. However, the level of job satisfaction has been found to vary between different occupational areas: For example, teachers might be highly satisfied with their core professional tasks (e.g., work with children, instruction, and teaching) but dissatisfied with organizational factors (e.g., classroom size, number of administrative tasks) and professional support. Also, significant differences in terms of school type have been reported, indicating high levels of job satisfaction for nursery and primary school teachers in comparison to secondary teachers. It must also be noted that relevant predictors of job satisfaction (e.g., interaction with students) can simultaneously serve as causes of stress. As a result, a more differentiated approach to satisfaction and teacher well-being in terms of the quality of experience and the evaluated subject is recommended. For example, different forms of satisfaction can be identified (e.g., Inauen et al., 2015).

Interestingly, although it has been repeatedly mentioned that teacher well-being is a risk, research has shown that the majority of teachers are satisfied with their job and feel well in school (Hascher and Waber, 2021). At the same time, however, increasing

evidence has revealed that teachers experience high levels of stress. These results point to the necessity to carefully define teacher well-being and factors that might influence it. There is a need for a more precise conceptualization of teacher well-being. Based on the original idea of well-being as a multi-dimensional construct, teacher well-being could be defined as a positive imbalance of work-related dimensions. Teacher well-being could be seen as a clearly defined set of positive dimensions that are more pronounced than a clear set of negative dimensions. This definition could also help differentiate teacher well-being from a sole absence of health issues, stress, burden, or burnout symptoms.

Predictors and outcomes of teacher well-being

According to research on general well-being, various predictors for teacher well-being can be assumed. Generally, they can be divided into individual and contextual factors and their interplay. Regarding contextual factors, the workplace is the main source of teacher well-being. Diverse conditions of working in schools and teachers' subjective perceptions need to be taken into account. These are, for instance, high autonomy, high responsibility for students' learning outcomes, support of colleagues, a continuous process of educational reforms, high achievement-pressure due to school evaluation and assessment, challenging task of handling disruptive student behavior. Furthermore, a differentiation of intrinsic factors, like contributions to children's education and development, and extrinsic factors, like salary and social status, are thought to be helpful in understanding the factors that facilitate or hamper teacher well-being.

Resources at work are of considerable importance for teacher well-being (Salanova et al., 2006). Two forms of resources and their interaction can be differentiated: personal resources (e.g., self-efficacy beliefs) and organizational resources (e.g., autonomy and social support at work). Research on personal resources confirms that high self-efficacy beliefs can foster job satisfaction and teacher well-being (e.g., Capone and Petrillo, 2018). Similarly, organizational resources have been shown to influence teacher well-being (e.g., Ilgan et al., 2015), although the relationship between these constructs may be bidirectional. For example, organizational resources may influence teacher well-being, which, in turn, may affect future personal and organizational development. Among the array of contextual factors, social relationships have been repeatedly confirmed as a major source of teacher well-being (Hascher and Waber, 2021). Teaching as a social profession that is based on interactions with students, colleagues, and parents seems to be specifically vulnerable to social factors with social interactions at the heart of teacher well-being and crucial in fostering teacher well-being. Positive teacher-student-relationships nourish teachers' needs for social relatedness (Collie et al., 2016; Klassen et al., 2012; Spilt et al., 2011) and they mirror a successful fulfilment of teachers' role as educators that, in turn, has been shown to promote teacher well-being.

What are the possible effects of teacher well-being with regard to learning? Although empirical evidence is scarce and only few longitudinal studies have been conducted, three main effects have been identified. First, teacher well-being influences teacher learning and supports teacher professional development directly or indirectly. Teaching is known as a very challenging profession that requires continuous professional development. As a protective source against teacher stress, teacher negative affect and burnout (Renshaw et al., 2015; Vazi et al., 2013), teacher well-being can contribute to individual motivation for in-service training and a supportive learning environment alike. With its association with teaching professional values, job satisfaction, and social support, and the experiences of meaningful work (Brunzell et al., 2018; Tang et al., 2018), teacher well-being can foster the development of individual resources and strengths that also might positively extend to the classroom and school climate. Second, teacher well-being has been found to be associated with quality teaching (Turner and Thielking, 2019). Teachers who feel well in school are more likely to engage in instruction, provide individual learning support for students, and support school improvement action plans. It has also been found that students' perceptions of interpersonal teacher behavior that predict student well-being are related to teacher well-being (Van Petegem et al., 2007). Third, teacher well-being seems to influence teachers' professional development such that a dominance of negative emotions can impede their willingness to implement educational reforms (e.g., Borko, 2002; Clement, 2014).

Supporting student and teacher well-being in school

Along with growing demands on schools, such as achievement pressure through high stakes testing or educating for 21st century skills and reports about increasing student and teacher mental health issues, the importance of student and teacher well-being has been amplified. Accordingly, there is broad agreement that schools need to cultivate a social climate that is characterized by physical and emotional safety, physical and mental health, as well as learning support and enjoyment (Hascher, 2003; Roffey, 2012; Shoshani and Slone, 2013). Also, the fulfilment of students' and teachers' psychological needs alike is seen as a valuable contribution to a school in building up a supportive learning environment. Moreover, empirical evidence suggests that student well-being and teacher well-being interact (Coleman, 2009; McCallum and Price, 2010).

Although some schools may intuitively succeed in fostering well-being in school, one can expect a high probability of failure if a school does not deliberately care about student and teacher well-being. Support for those schools might be given through intervention research findings that examines the effectiveness of specific programs on well-being in school. Intervention research related to well-being in school has been dedicated to the individual, to groups, or the whole school. Intervention programs that aim at fostering well-being in school have been based on different approaches such as mental health (e.g., Dawood, 2013; Simovska and McNamara, 2015), positive psychology (e.g., Green et al., 2012; Sin and Lyubomirsky, 2009), self-determination theory (Cheon et al., 2014), school improvement initiatives regarding school climate (Bradley et al., 2018) or socio-emotional learning (Durlak et al., 2011). Common to all these programs is the shared idea of the need for a school environment that capitalizes on

the strengths of the individual, the positive power of the group and community, and the efficacy of an educational setting that is dedicated to offering the best possible opportunities for students' and teachers' learning and development.

Across the different approaches to supporting well-being in school, most interventions have been found to hold considerable potential for establishing a caring and empowering learning climate (e.g., Denny et al., 2011). Although effect sizes have usually been shown to be low to moderate, converging evidence indicates that well-being in school can be maintained and improved. Ideally, well-being interventions have been targeted to teachers as well as students, have been integrated into the curriculum instead of providing an additional or extra-curricular program, and have been closely related to students' and teachers' roles in school (e.g., Shankland and Rosset, 2017). Following a whole-school-approach has considerable potential for sustainable school improvement (Quinlan and Hone, 2020; Shoshani and Steinmetz, 2014). These results are encouraging, as they support the notion that well-being in school can be significantly fostered. The effective implementation of strategies that equally boost student and teacher well-being and support the alignment of academic/professional learning and well-being can be regarded as of utmost importance for future research and practice.

Future perspective on well-being and learning

Around the globe, educational systems and schools, in particular, have recognized the importance of teacher and student well-being. This is based on the finding that well-being can contribute to learning and educational success as much as other factors such as motivation, cognition, or social interaction. Given the complexity and the multi-dimensional character of the construct of well-being, a broad range of conceptualizations and models exists. Equally, a variety of factors that influence well-being has been identified. This complexity might be enriching, but the existing ambiguity in definitions and interrelations of constructs can also blur our understanding of well-being. Consequently, this might hamper discussion about well-being in school and its relationship with learning. Thus, it would be helpful to identify some core dimensions of well-being and to define more precisely the specific quality of well-being such as a positive imbalance where positive emotions and cognition outperform negative emotions and cognition.

In the future, well-being and learning need further careful attention in research and practice. It is important to elucidate the interplay between well-being and learning. Reciprocal effects that show the role of well-being for learning and the importance of learning for the development of well-being need to be investigated. Specifically, research is needed that identifies how well-being in school and learning can best be fostered. An answer to this question might be based on a sound partnership of researchers and school practitioners who collaboratively develop and evaluate programs to promote well-being in school such as self-regulation trainings, resilience trainings, social competence trainings or health education. As well-being is related not only to the individual but also classrooms and school teams, a whole school approach that takes various levels into account might be helpful.

References

- Abele, A., Silvia, P.J., Zöller-Utz, I., 2005. Flexible effects of positive affect on self-focused attention. *Cognit. Emot.* 19, 623–631.
- Andrews, F.M., Robinson, J.P., 1991. Measures of subjective well-being. In: Robinson, J.P., Shaver, P.R., Wrightsman, L.S. (Eds.), *Measures of Personality and Social Psychological Attitudes*. Academic Press, San Diego, pp. 61–114.
- Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117 (3), 497–529.
- Benjet, C., Hernández-Guzmán, L., 2002. A short-term longitudinal study of pubertal change, gender, and psychological well-being of Mexican early adolescents. *J. Youth Adolesc.* 31, 429–442.
- Bergman, M.M., Scott, J., 2001. Young adolescents' wellbeing and health risk behaviors: gender and socio-economic differences. *J. Adolesc.* 24, 183–197.
- Biswas-Diener, R., Vittersø, J., Diener, E., 2005. Most people are pretty happy, but there is a cultural variation: the Inughuit, the Amish, and the Massai. *J. Happiness Stud.* 6, 205–226.
- Boekaerts, M., 1993. Being concerned with well-being and with learning. *Educ. Psychol.* 28, 149–167.
- Borko, H., 2002. Professional development: a key to Kentucky's educational reform effort. *Teach. Teach. Educ.* 18, 969–987.
- Bradley, C., Cordaro, D.T., Zhi, F., et al., 2018. Supporting improvements in classroom climate for students and teachers with the four pillars of wellbeing curriculum. *Transl. Issues Psychol. Sci.* 4 (3), 245–264.
- Brebner, J., 2003. Gender and emotions. *Pers. Individ. Differ.* 34, 387–394.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press, Cambridge, MA.
- Brunzell, T., Stokes, H., Waters, L., 2018. Why do you work with struggling students? Teacher perceptions of meaningful work in trauma-impacted classrooms. *Aust. J. Teach. Educ.* 43 (2), 116–142.
- Burns, R.A., Machin, M.A., 2013. Employee and workplace well-being: a multi-level analysis of teacher personality and organizational climate in Norwegian teachers from rural, urban and city schools. *Scand. J. Educ. Res.* 57 (3), 309–324.
- Buyse, E., Verschueren, K., Verachtert, P., Van Damme, J., 2009. Predicting school adjustment in early elementary school: impact of teacher-child relationship quality and relational classroom climate. *Elem. Sch. J.* 110 (2), 119–141. <https://doi.org/10.1086/605768>.
- Capone, V., Petrillo, G., 2018. Mental health in teachers: relationships with job satisfaction, efficacy beliefs, burnout and depression. *Curr. Psychol.* 39 (5), 1757–1766.
- Cheon, S.H., Reeve, J., Yu, T.H., Jang, H.R., 2014. The teacher benefits from giving autonomy support during physical education instruction. *J. Sport Exerc. Psychol.* 36 (4), 331–346.
- Chu, P.S., Saucier, D.A., Hafner, E., 2010. Meta-analysis of the relationship between social support and well-being in children and adolescents. *J. Soc. Clin. Psychol.* 29, 624–645.
- Clement, J., 2014. Managing mandated educational change. *Sch. Leader. Manag.* 34 (1), 39–51.
- Coleman, J., 2009. Wellbeing in schools: empirical measure, or politician's dream? *Oxf. Rev. Educ.* 35 (3), 281–292.
- Collie, R.J., Shapka, J.D., Perry, N.E., Martin, A.J., 2016. Teachers' psychological functioning in the workplace: exploring the roles of contextual beliefs, need satisfaction, and personal characteristics. *J. Educ. Psychol.* 108 (6), 788–799.
- Dawood, R., 2013. Positive psychology in school-based psychological intervention: a study of the evidence-base. *Procedia Soc. Behav. Sci.* 113, 44–53.

- Denny, S.J., Robinson, E.M., Utter, J., Fleming, T.M., Grant, S., Milfont, T.L., et al., 2011. Do schools influence student risk-taking behaviors and emotional health symptoms? *J. Adolesc. Health* 48, 259–267.
- Diener, E., Biswas-Diener, R., 2000. *New Directions in Subjective Well-Being Research: The Cutting Edge*. University of Illinois, Department of Psychology, Champaign, IL.
- Diener, E., Lucas, R.E., 2000. Subjective emotional well-being. In: Lewis, M., Haviland, J.M. (Eds.), *Handbook of Emotions*. Guilford Press, New York, pp. 325–337.
- Diener, E., Diener, M., Diener, C., 1995. Factors predicting the subjective well-being of nations. *J. Pers. Soc. Psychol.* 69, 851–864.
- Diener, E., 1984. Subjective well-being. *Psychol. Bull.* 95, 542–575.
- Duckworth, A.L., Quinn, P.D., Seligman, M.E.P., 2009. Positive predictors of teacher effectiveness. *J. Posit. Psychol.* 4, 540–547.
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., 2011. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev.* 82 (1), 405–432.
- Efkides, A., Petkaki, C., 2005. Effects of mood on students' metacognitive experiences. *Learn. Instr.* 15 (5), 415–431.
- Engels, N., Aelterman, A., Van Petegem, K., Schepens, A., 2002. Factors which influence the well-being of pupils in Flemish secondary schools. *Educ. Stud.* 30, 127–143.
- Fredrickson, B.L., 2001. The role of positive emotions in positive psychology. *Am. Psychol.* 56 (3), 218–226.
- Frenzel, A.C., 2014. Teacher emotions. In: Linnenbrink-Garcia, E.A., Pekrun, R. (Eds.), *International Handbook of Emotions in Education*. Routledge, New York, pp. 494–519.
- Green, L.S., Oades, L.G., Robinson, P.L., 2012. Positive education programs: integration coaching and positive psychology in schools. In: van Nieuwerburgh, C. (Ed.), *Coaching in Education: Getting Better Results for Students, Educators, and Parents*. Karnac Books, London, pp. 115–132.
- Hagenauer, G., Hascher, T., 2014. Early adolescents' enjoyment experienced in learning situations at school and its relation to student achievement. *J. Educ. Train. Stud.* 2, 20–30.
- Hagenauer, G., Hascher, T., Volet, S., 2015. Teacher emotions in the classroom: associations with students' engagement, classroom discipline and the interpersonal teacher-student relationship. *Eur. J. Psychol. Educ.* 30 (4), 385–403.
- Hall-Lande, J., Eisenberg, M.E., Christenson, S., Neumark-Sztainer, D., 2007. Social isolation, psychological health, and protective factors in adolescence. *Adolescence* 42 (166), 265–286.
- Hascher, T., Hagenauer, G., 2020. Swiss adolescents' well-being in school. *Swiss J. Educ. Res.* 42 (2), 367–390.
- Hascher, T., Waber, J., 2021. Teacher well-being: a systematic review of the research literature from the year 2000 to 2019. *Educ. Res. Rev.* 34 (8).
- Hascher, T., Beltman, S., Mansfield, C., 2021. Teacher wellbeing and resilience: towards an integrative model. *Educ. Res.* <https://doi.org/10.1080/00131881.2021.1980416>.
- Hascher, T., 2003. Well-being in school—why students need social support. In: Mayring, P., von Rhöneck, C. (Eds.), *Learning Emotions—The Influence of Affective Factors on Classroom Learning*. Lang, Bern, pp. 127–142.
- Hascher, T., 2007. Exploring students' well-being by taking a variety of looks into the classroom. *Hellenic J. Psychol.* 4, 331–349.
- Hascher, T., 2008. Quantitative and qualitative research approaches to assess student wellbeing. *Int. J. Educ. Res.* 47, 84–96.
- Hascher, T., 2010. Learning and Emotion: perspectives for theory and research. *Eur. Educ. Res. J.* 9, 13–28.
- Hascher, T., 2011. Wellbeing. In: Järvelä, S. (Ed.), *Social and Emotional Aspects of Learning*. Elsevier, Oxford, pp. 99–105.
- Huang, F.-M., Chien, Y.-N., 2013. Working, schooling, and psychological well-being: evidence from longitudinal data for Taiwanese youth. In: Yi, C.-C. (Ed.), *The Psychological Well-Being of East Asian Youth*. Springer, Dordrecht, pp. 185–207.
- Huebner, E.S., Ash, C., Laughlin, J.E., 2001. Life experiences, locus of control, and school satisfaction in adolescence. *Soc. Indic. Res.* 55 (2), 167–183.
- Ilgan, A., Özü-Cengiz, O., Ata, A., Akram, M., 2015. The relationship between teachers' psychological well-being and their quality of school work life. *J. Happiness Well-Being* 3 (2), 159–181.
- Inauen, A., Jenny, G., Bauer, G.F., 2015. Discriminating five forms of job satisfaction: investigating their relevance for occupational health research and practice. *Psychology* 6, 138–150.
- Inchley, J., Currie, D., Young, T., Samdal, O., Torsheim, T., Augustson, L., Mathison, F., Aleman-Diaz, A., Molcho, M., Weber, M., Barnekow, V., 2016. Growing up unequal: gender and socioeconomic differences in young people's health and well-being. In: *Health Behaviour in School-Aged Children (HBSC) Study: International Report From the 2013/2014 Survey*. WHO Regional Office for Europe, Copenhagen.
- Jin, S.-U., Moon, S.M., 2006. A study of well-being and school satisfaction among academically talented students attending a science high school in Korea. *Gift. Child. Q.* 50, 169–184.
- Kim-Prieto, C., Diener, E., Tamir, M., Scollon, C., Diener, M., 2005. Integrating the diverse definitions of happiness: a time-sequential framework of subjective well-being. *J. Happiness Stud.* 6, 261–300.
- Klassen, R.M., Perry, N.E., Frenzel, A.C., 2012. Teachers' relatedness with students: an underemphasized component of teachers' basic psychological needs. *J. Educ. Psychol.* 104 (1), 150–165.
- Konu, A., Lintonen, T.P., 2005. Theory-based survey analysis of well-being in secondary schools in Finland. *Health Promot. Int. Adv. Access.* <https://doi.org/10.1093/heapro/dai028>.
- Konu, A., Lintonen, T.P., 2006. School well-being in grades 4–12. *Health Educ. Res.* 21, 633–642.
- Konu, A.I., Lintonen, T.P., Autio, V.J., 2002. Evaluation of well-being in schools—a multilevel analysis of general subjective well-being. *Sch. Effect. Sch. Improv.* 13, 187–200.
- Larsen, R.J., Diener, E., 1987. Affect intensity as an individual difference characteristic: a review. *J. Res. Pers.* 21, 1–39.
- Linley, P.A., Joseph, S., Harrington, S., Wood, A.M., 2006. Positive psychology: past, present, and (possible future). *J. Posit. Psychol.* 1, 3–16.
- Liu, W., Mei, J., Tian, L., Huebner, S.E., 2016. Age and gender differences in the relation between school-related social support and subjective well-being in school among students. *Soc. Indic. Res.* 125, 1065–1083.
- Lohre, A., Moksnes, U.K., Lillefjell, M., 2014. Gender differences in predictors of school wellbeing? *Health Educ. J.* 73 (1), 90–100.
- Lucas, R.E., Diener, E., Suh, E., 1996. Discriminant validity of well-being measures. *J. Pers. Soc. Psychol.* 71, 616–628.
- Martin, P., 2005. *Making People Happy. The Nature of Happiness and Its Origins in Childhood*. Fourth Estate, Harper Collins, London, UK.
- McCallum, F., Price, D., 2010. Well teachers, well students. *J. Stud. Wellbeing* 4 (1), 19–34.
- McCallum, F., Price, D., Graham, A., Morrison, A., 2017. Teacher Wellbeing: A Review of the Literature. The Association of Independent Schools of New South Wales, Canberra. <https://apo.org.au/sites/default/files/resource-files/2017/10/apo-nid201816-1133141.pdf>.
- Morinaj, J., Hascher, T., 2019. The longitudinal relation between student well-being and academic achievement. In: Castelli, L., Marcionetti, J., Plata, A., Ambrosetti, A. (Eds.), *Well-Being in Education Systems*. Bern, Hogrefe, pp. 273–277.
- Morinaj, J., Hascher, T., 2021. Well-being of primary and secondary school students in Switzerland: a longitudinal perspective. In: McLellan, R., Faucher, C., Simovska, V. (Eds.), *Wellbeing and Schooling: Cross Cultural and Cross Disciplinary Perspectives*. Springer, New York, XX-XX.
- OECD, 2017. *PISA 2015 Results Volume III: Students' Well-Being*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264273856-en>.
- Opdenakker, M.-C., Van Damme, J., 2000. Effects of schools, teaching staff and classes on achievement and well-being in secondary education: similarities and differences between school outcomes. *Sch. Effect. Sch. Improv.* 11, 165–196.
- Pekrun, R., Linnenbrink-Garcia, L., 2012. Academic emotions and student engagement. In: Christenson, S.L., Reschly, A.L., Wylie, C. (Eds.), *Handbook of Research on Student Engagement*. Springer, New York, pp. 259–282.
- Quinlan, D.M., Hone, L.C., 2020. *An Educators' Guide to Whole-School Wellbeing*. Routledge, London.
- Renshaw, T.L., Long, A.C.J., Cook, C.R., 2015. Assessing teachers' positive psychological functioning at work: development and validation of the teacher subjective wellbeing questionnaire. *Sch. Psychol. Q.* 30 (2), 289–306.
- Reschly, A.L., Huebner, E.S., Appleton, J.J., Antaramian, S., 2008. Engagement as flourishing: the contribution of positive emotions and coping to adolescents' engagement at school and with learning. *Psychol. Sch.* 45 (5), 419–431.

- Roffey, S., 2012. Pupil wellbeing—teacher wellbeing: two sides of the same coin? *Educ. Child Psychol.* 29 (4), 8–17.
- Ryan, R.M., Deci, E.L., 2001. On happiness and human potentials: a review of research on hedonic and eudemonic well-being. *Annu. Rev. Psychol.* 52, 141–166.
- Ryff, C.D., Keyes, C.L.M., 1995. The structure of subjective well-being revisited. *J. Pers. Soc. Psychol.* 69, 719–727.
- Salanova, M., Bakker, A.B., Llorens, S., 2006. Flow at work: evidence for an upward spiral of personal and organizational resources. *J. Happiness Stud.* 7, 1–22.
- Seligman, M.E.P., Ernst, R., Gillham, J., Reivich, K., Linkins, M., 2009. Positive education: positive psychology and classroom interventions. *Oxf. Rev. Educ.* 35, 293–311.
- Seligman, M.E.P., 2002. *Authentic Happiness: Using the New Positive Psychology to Realize Your Potential for Lasting Fulfillment*. Nicholas Brealey, London.
- Seligman, M.E.P., 2011. *Flourish: A New Understanding of Happiness and Well-Being—And How to Achieve Them*. Nicholas Brealey, London.
- Shankland, R., Rosset, E., 2017. Review of brief school-based positive psychological interventions: a taster for teachers and educators. *Educ. Psychol. Rev.* 29, 363–392.
- Shoshani, A., Slone, M., 2013. Middle school transition from the strengths perspective: young adolescents' character strengths, subjective well-being, and school adjustment. *J. Happiness Stud.* 14, 1163–1181.
- Shoshani, A., Steinmetz, S., 2014. Positive psychology at school: a school-based intervention to promote adolescents' mental health and well-being. *J. Happiness Stud.* 15, 1289–1311.
- Simovska, V., McNamara, P. (Eds.), 2015. *Schools for Health and Sustainability: Theory, Research and Practice*. Springer, New York.
- Sin, N.L., Lyubomirsky, S., 2009. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. *J. Clin. Psychol.* 65, 467–487.
- Skaalvik, E.M., Skaalvik, S., 2018. Job demands and job resources as predictors of teacher motivation and well-being. *Soc. Psychol. Educ.* 21, 1251–1275.
- Spilt, J.L., Koomen, H.M.Y., Thijs, J.T., 2011. Teacher wellbeing: the importance of teacher–student relationships. *Educ. Psychol. Rev.* 23, 457–477.
- Stein, N.L., Folkman, S., Trabasso, T., Richards, T.A., 1997. Appraisal and goal processes as predictors of psychological well-being in bereaved caregivers. *J. Pers. Soc. Psychol.* 72, 872–884.
- Strack, F., Argyle, M., Schwarz, N. (Eds.), 1991. *Subjective Well-Being*. Pergamon Press, Oxford.
- Tang, Y., He, W., Liu, L., Li, Q., 2018. Beyond the paycheck: Chinese rural teacher well-being and the impact of professional learning and local community engagement. *Teach. Teach.* 24 (7), 825–839.
- Turner, K., Thielking, M., 2019. Teacher wellbeing: its effects on teaching practice and student learning. *Issues Educ. Res.* 29 (3), 938–960.
- Undheim, A.M., Sund, A., 2005. School factors and the emergence of depressive symptoms among young Norwegian adolescents. *Eur. Child Adolesc. Psychiatr.* 14, 446–453.
- UNICEF Office of Research, 2017. *Building the Future: Children and the Sustainable Development Goals in Rich Countries* (Innocenti Report Card 14). Innocenti, UNICEF Office of Research, Florence.
- Van Petegem, K., Aelterman, A., Rosseel, Y., Creemers, B., 2006. Student perception as moderator for student well-being. *Soc. Indic. Res.* <https://doi.org/10.1007/s11205-006-9055-5>.
- Van Petegem, K., Aelterman, A., Van Keer, H., Rosseel, Y., 2007. The influence of student characteristics and interpersonal teacher behavior in the classroom on student's well-being. *Soc. Indic. Res.* <https://doi.org/10.1007/s11205-007-9093-7>.
- Vazi, M.L.M., Ruiters, R.A.C., van den Borne, B., Martin, G., Dumont, K., Reddy, P.S., 2013. The relationship between wellbeing indicators and teacher psychological stress in Eastern Cape public schools in South Africa. *SA J. Ind. Psychol.* 39, 1–10.
- Veenhoven, R., 1991. Questions on happiness: classical topics, modern answers, blind spots. In: Strack, F., Argyle, I., Schwarz, N. (Eds.), *Subjective Well-Being*. Pergamon Press, Oxford, pp. 7–26.
- Veldman, I., Van Tartwijk, J., Brekelmans, M., Wubbels, T., 2013. Job satisfaction and teacher–student relationships across the teaching career: four case studies. *Teach. Teach. Educ.* 32, 55–65.
- Vittersø, J., Nilsen, F., 2002. The conceptual and relational structure of subjective well-being, neuroticism, and extraversion: once again, neuroticism is the important predictor of happiness. *Soc. Indic. Res.* 57, 89–118.
- WHO, 2000. *Health and health behavior among young people*. In: A WHO-Cross-National-Survey. World Health Organization, Copenhagen.
- Wood, A.M., Josephs, S., 2010. The absence of positive psychological (eudemonic) well-being as a risk factor for depression: a ten year cohort study. *J. Affect. Disord.* 122, 213217.

FOURTH EDITION

International Encyclopedia of Education

Approaches to Curriculum
and its Politics

VOLUME
7

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Berit Karseth
Ninni Wahlström



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 7

Approaches to Curriculum and its Politics

VOLUME 7 EDITORS

Berit Karseth and Ninni Wahlström



ELSEVIER

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

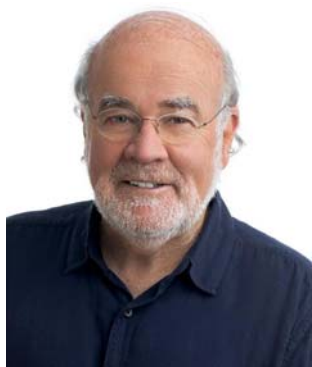
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati Mamoe Ngati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



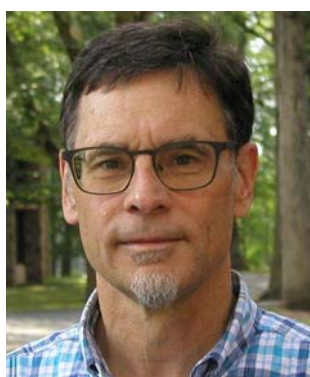
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 7: Approaches to the Curriculum and Its Politics

Berit Karseth and Ninni Wahlström

© 2023 Elsevier Ltd. All rights reserved.

Introduction

In this, we introduce the themes and contributions that form the basis of the volume “Approaches to curriculum and its politics” within the 4th edition of the *International Encyclopedia of Education*. A common theme that binds these chapters together is unraveling the importance of the curriculum for different aspects and perspectives on education in general and in schools and preschools, in particular. Furthermore, two areas recur across the chapters: the curriculum as part of a framework for the governance of schools and the curriculum as an arena for competing perspectives on knowledge.

The definition and meaning of the term “curriculum” have been the cause of lengthy discussions. According to the *Oxford English Dictionary*, the term curriculum means “a regular course of study or training, as at a school or university”. However, as early as the beginning of the 20th century, John Dewey (among others) considered such a definition too narrow, although he never tried to formulate a new definition. Instead, he highlighted the gap between the child and the curriculum in a text that employed this exact title (Dewey, 1976). Highlighting young children’s interests and activities with respect to the formal course of study was an effective way of illustrating the need for some kind of transformative pedagogical activity that would enable a connection between the two parties (Jackson, 1992). Dewey posed two questions that underpinned all curriculum making: “What do we want to do, and why?” and “How can we do it?” However, these questions cannot be answered through the text of the curriculum alone, even if they constitute an important basis. Rather, they should be understood from a broader perspective that includes the actual institutional and organizational frameworks that shape the practice of the school (Westbury, 2008).

As Apple (2008) noted, the choices and emphases made in the field of curricula also raise ethical and political issues. The curriculum is not only a “course of study,” as it is also a site of conflict about what kind of knowledge content should be selected and included in the curriculum and how teaching and learning should be assessed. Moreover, how curricula are selected, taught, and evaluated cannot be understood as isolated phenomena in the field of education. Instead, these activities need to be understood as relational to the society in which they exist and to prevailing power relations (Apple, 2008). Based on such a perspective, Pinar’s understanding of curriculum as “an extraordinarily complicated conversation” in which we “choose what to remember about the past, what to believe about the present, what to hope for, and fear about the future” offers a more multidimensional meaning of the concept of curriculum than simply “a course” to be followed (Pinar, 2008). Hence, the *understanding* of a curriculum as a contextualized and experienced phenomenon is an important aspect of curriculum studies (Jung and Pinar, 2016). Drawing on Basil Bernstein (1975), the curriculum can be viewed as being included in a broader message system of education, distinguishing between the curriculum as representing valid knowledge, pedagogy as the valid transmission of knowledge, and evaluation as the valid realization of knowledge (Wyse et al., 2016). In this way, the selection of knowledge, the teaching of the subject content, and the assessment of student knowledge can be understood as constituting core interests in the field of curriculum research.

A critical curriculum theory approach comprises two questions: what is included in the curriculum and what is excluded. For this purpose, the curriculum definition that guides this section is what the child and the student “have an opportunity to learn” through education, which emphasizes the continuum between the curriculum and teaching (Cherryholmes, 1988, p. 133). This definition of what the students have an opportunity to learn acknowledges the curriculum as a guideline for planning and evaluation. This definition also includes what students learn from the school structure and organization, from their peers, and from what is excluded from the discourse of the teaching content. In the following, this broad understanding of critical curriculum theory research guided the selection of themes that constitute the basis of this section. The chapters in this section are largely inspired by studies that are characterized by a critical reading of different approaches to curricula and their relationships with politics and practice.

The chapters in this volume are situated in different national contexts and regions. Despite globalization, the distinctiveness of national history and culture can play an important role in how curricula are structured and designed (Deng et al., 2022; Tröhler, 2014), enacted in classrooms, and recontextualized in pedagogical practices (Ensor, 2015; Schulte, 2018). On the one hand, acknowledgment of the situatedness of the curriculum (as well as curriculum research) can be perceived as an obstacle to comparative research. On the other hand, these differences and particularities invite us to learn about other systems, helping us understand our own curriculum history and system of reasoning.

The chapters in the first section examine aspects of curriculum governance and global education policy. Curriculum reforms are viewed from both historical and contemporary approaches, and they reflect specific purposes and conditions that change over time. As a modern invention, various actors set the agenda and formulate purposes with regard to given circumstances. The historical and theoretical roots of the dilemma between curricular centralization and teacher autonomy are discussed in this section by drawing attention to the complexity of curriculum governance, which is linked to every aspect of the education system. Moreover, it relates to curriculum models, visions of what is relevant to learn, and implications for the role and agency of both teachers and students. In addition, the processes of decentralization and centralization are discussed. Furthermore, the chapters reveal the impact of current transnational curriculum models and examine the discourse pertaining to the position of 21st-century skills in the curriculum. This represents a manifestation of the concerns of diverse stakeholders in society about the current relevance of the formal education curriculum in these times of rapid changes in society. In one of the contributions, an argument is made that the logic of current transnational curriculum models is based on learning outcomes that under stipulate curriculum content. This places new burdens

on all teachers, especially those in developing countries. From a different empirical point of departure, another chapter demonstrates how global curriculum models are recontextualized according to traditions and dynamics that exist within an educational system. This analysis highlights the complex, ever-changing dynamics that underlie contemporary curricular policies. The section also presents critical frameworks for analyzing curriculum reforms and offers different theoretical lenses to unpack curriculum reforms beyond global trends.

The second theme is termed “curriculum theory,” which includes contemporary trends of two topical curriculum discussions: understanding the curriculum as part of the governance of education and arguments for different kinds of knowledge approaches that should underpin curriculum. These trends can be described as displacements from new public management to more pronounced multilevel governance (often in terms of networks) and from an unproblematic approach to knowledge to a discussion of knowledge that focuses on the relationships between humans and the “real” world. The topic of realism is followed up in a chapter that focuses on a discussion of the philosophy of critical realism. Furthermore, the notion of “critical” in curriculum theory is discussed in terms of “new materialism,” which goes beyond divisions into subjects and objects. It is argued in this discussion that critical theory brings into view an agential notion to resist the seemingly “normal” in educational practices.

From a historical perspective, the development of curriculum studies as an educational subdiscipline is reconstructed, and three models of curriculum theory on three different levels of theoretical abstractions are outlined. From another historical angle, an analysis of world systems suggests the need for a reorganization of the social sciences that breaks with the development of the social sciences as knowledge disciplines in the historical trajectory of the capitalist world system. Finally, it is argued that the theories of new institutionalism, especially its latest model of discursive institutionalism, constitute a vital theoretical starting point for analyzing curriculum change in a globalized world.

In the next section of curriculum research, the focus is on curriculum making at the local school level, which entails deliberative decision-making at the planning stage and in the work within the classroom. It is argued that curriculum making is dependent upon teacher agency, which is strengthened by meaningful curriculum making as a social practice. The words we choose to describe the actions (or positions) that teachers take in relation to the curriculum matter because they reveal how we consider teachers as professionals and what we perceive (or expect) their role to be with respect to curriculum policy directives. Deliberative decision-making requires the use of theory to illuminate and interpret specific issues and problems and involves fundamental questions related to what teachers do and think. Hence, curriculum making constitutes an important part of teacher education.

Innovative pedagogy and curricula should be understood as accommodating divergent interpretations and actions that are negotiated and exercised as part of the pedagogical processes of interpreting the curriculum and transforming it into teaching at the school level. Moreover, teachers, children, and students should be involved in curriculum making, as this can also be understood as a learning activity. Accordingly, the ontologies, theories, and ideas that enable children to participate in local curriculum making are presented and discussed. In many early childhood curricula, children’s interests are purported to be central. Thus, this subsection includes definitions of the term “children’s interests” and ways in which these can inform approaches to interest-related curricula. At the other end of the age scale, curriculum making in higher education is explored from the perspective of activity theory, emphasizing academics as actors. Shifting the focus to a macro level, curriculum development can also be understood as a specific part of the broader process of curriculum making. Currently, it is argued that it is often more appropriate to speak about curriculum revisions or reviews rather than curriculum reforms.

Under the title “Curriculum design and curriculum innovations,” the subsequent set of chapters highlights new and emerging approaches. There are different approaches to curriculum design (such as instrumental and deliberative), and it is argued that an approach that emphasizes the important role of research in the process is preferable. Common themes in many of the chapters are an emphasis on the learner and questions about how to develop a learning environment that enhances student involvement in decision-making processes. In particular, investment in bottom-up initiatives is necessary to successfully implement a curriculum design that invites cooperation when designing a curriculum. Moreover, the distinction between the cocreation of the curriculum and cocreation *in* the curriculum is described. Curriculum changes are also affected by how school buildings are designed and used. School architecture is not only defined by the architect’s pen; it is created through policies and pedagogies, as well as teachers’ and students’ everyday use of space.

Within the theme of “Curriculum, school subjects, and Didaktik,” one of the chapters provides an overview of the theoretical and practical domains constituted by European Continental Pedagogy, emphasizing key aspects of its history and theory. Here, Didaktik is discussed in terms of the relationships illustrated through the “didactic triangle,” which combines the educator, the *educand* (the one being educated), and the world (the selection of content). In another chapter, Didaktik is contextualized as tightly connected to the curriculum and Bildung in a European continental (German-speaking) tradition. These three concepts are presented in their origins from a German discourse on education and in relation to an Anglo-American understanding of curricula. It is also argued that the three concepts have changed in certain ways due to competence-based curricula and standardized testing. It is apparent that different orientations of the view of knowledge have consequences in terms of schools’ selection of teaching content and the organization of teaching. Generic competencies, as highlighted in the OECD project “The Future of Education and Skills: Education 2030,” can contribute to transforming education policy in national arenas. The transnational perspective of the OECD and international knowledge tests results in a discussion of how the notion of competence varies across contexts and over time. In another text, the concept of competence is explored, including its role in a differentiated curriculum, the connection between competence and assessment, the relationship between competence and standards, and how this relates to the increasing importance of accountability in education systems. Within a closely related area to the discourse on competencies, the question of how an interdisciplinary approach can be understood and applied in different countries is raised, in addition to why interdisciplinarity should be promoted in basic education.

In school subjects, tensions often arise between different perspectives when selecting teaching content. The curriculum of the school subject history is discussed in relation to human activity throughout history, which significantly influences the climate and ecosystem on the earth. However, turning the subject of history in the direction of a deepened understanding of the development of climate change could challenge the current understanding of the body of knowledge that constitutes the school subject of history. In addition, new emphases for social studies teaching are outlined, where students examine inequitable power relations as a way of deepening their understanding of society. Another field of tension is explored in the literacy curriculum. Here, one trend is toward expanding and pluralizing the scholarly understanding of literacy, while another is toward curriculum and policy focused on measurable units and standardization. The subject of science is discussed in terms of always being a representation of scientific content rather than being a scientific content in itself. This is due to pedagogical considerations that inevitably involve certain selections and simplifications of the teaching content. Another field of tension that is actualized is between public education and supplementary curriculum out of schooling provided by educational business industries (referred to as shadow education).

In the section of “Lived curriculum,” the intersection between curriculum implementation and making is explored and problematized. While the curriculum generally comprises implementing governing regulations, curriculum making is about teachers educating learners in a local context. In the intersection between these arenas, issues pertaining to the enactment of teacher knowledge in context, competing stories of schools, and alternate versions of teacher communities emerge. The boundaries between school and life outside school are also challenged in terms of “living schools,” which reflect the creativity inherited from the life process itself. The significance of teachers and their curriculum making is clearly noticeable when on the front line, such as when facing structural inequalities in educational systems or challenges from the impacts of COVID-19. The process of individual and collective meaning making constitutes the core of curriculum enactment, which is a process that can adopt many different forms and adapt to different local situations. For example, the concept of identity options foregrounds the need to value learners’ cultural, linguistic, and epistemic resources in curriculum making. This has also promoted the consideration of how curricula are lived or experienced by learners and how this can affect their sense of self in the world.

The last section in this volume is entitled “Curriculum, democracy, and inclusion,” which consists of five chapters that share the common concern of education in the context of democracy. The global education policy of the United Nations’ Education for Sustainable Development is examined, and it is suggested that the common good should be included in a curriculum theory that aims to critically explore the curriculum processes involved.

It is important for the well-being of society that the younger generation is provided with opportunities to understand the possibilities that digital technology brings to sociopolitical participation as well as the threats it can pose to upholding the core values of social justice and democracy. Ideas around democracy and younger children have focused on children’s rights rather than children’s participation in democratic processes. Accordingly, one of the chapters contains a discussion on the curricular potential for democracy in early childhood, which would allow them to develop their own “voice” and actively participate in democratic processes.

The curriculum sends powerful messages to students about whose voices and perspectives are valued by society at large. Hence, as a reflection of dominant groups, the curriculum functions to marginalize LGBTQ+ students unless educators design an inclusive curriculum based on equity, democratic governance, and social justice.

The curriculum and inclusive education relate to children with diverse abilities and preferences having access to the curriculum and represent two tightly wound initiatives. No system can claim to serve the public unless it is accessible to all its members. The Universal Design for Learning framework is one approach for ensuring that education can be accessible to all participants in an educational setting. Although Universal Design for Learning is a promising traveling concept in many ways, it does not mean education for all, as some forms of disabilities continue to be excluded from their national curriculum.

References

- Apple, M.W., 2008. Curriculum planning. Content, form, and the politics of accountability. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, pp. 24–44.
- Bernstein, B., 1975. *Class, codes and control: towards a theory of educational transmissions*, vol. 3. Routledge and Kegan Paul.
- Cherryholmes, C.H., 1988. *Power and Criticism. Poststructural Investigations in Education*. Teachers College Press.
- Deng, L., Wu, S., Chen, Y., Wang, Y., Peng, Z., 2022. A comparative study of twenty-first century competencies in high school mother tongue curriculum standards in China, the United States and Finland. *Lang. Cult. Curric.* <https://doi.org/10.1080/07908318.2022.2072506>.
- Dewey, J., 1976. The child and the curriculum. In: Boydston, J.A. (Ed.), *John Dewey, The Middle Works, 1899–1924*, vol. 2. Southern Illinois University Press, pp. 271–291. Original work published in 1902.
- Ensor, P., 2015. Regulative discourse, ritual and the recontextualising of education policy into practice. *Learn. Cult. Soc. Interact.* 6, 67–76.
- Jackson, P.W., 1992. Conceptions of curriculum and curriculum specialists. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, pp. 3–40.
- Jung, J.-H., Pinar, W.F., 2016. Conceptions of curriculum. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *Curriculum, Pedagogy and Assessment*. Sage, pp. 29–46.
- Pinar, W.F., 2008. Curriculum theory since 1950. Crisis, reconceptualization, internationalization. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, pp. 491–513.
- Schulte, B., 2018. Envisioned and enacted practices: educational policies and the “politics of use” in schools. *J. Curric. Stud.* 50 (5), 624–637. <https://doi.org/10.1080/00220272.2018.1502812>.
- Tröhler, D., 2014. International curriculum research: why and how? In: Pinar, W.F. (Ed.), *International Handbook of Curriculum Research*, second ed. Routledge, pp. 60–66.
- Westbury, I., 2008. Curriculum in practice. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, pp. 1–4.
- Wyse, D., Hayward, L., Pandya, J., 2016. Introduction. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *Curriculum, Pedagogy and Assessment*. Sage, pp. 1–26.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 7

Daniel Alvunger

Linnaeus University, Växjö, Sweden

Johanna Annala

Tampere University, Tampere, Finland

Liesbeth Baartman

HU University of Applied Sciences Utrecht, Research Centre for Learning and Innovation, Research group Vocational Education, Utrecht, the Netherlands

Jessie S Barrot

National University, Manila, Philippines

Maria Birbili

School of Early Childhood Education, Aristotle University of Thessaloniki, Thessaloniki, Greece

Noah Borrero

School of Education, Urban Education and Social Justice, University of San Francisco, San Francisco, CA, United States

Catherine Bovill

Institute for Academic Development, University of Edinburgh, Edinburgh, United Kingdom

Steven Camicia

School of Teacher Education and Leadership, Utah State University, Logan, UT, United States

Liz Chesworth

The University of Sheffield, South Yorkshire, UK

Carla Ruthes Coelho

Western University, School of Education, John George Althouse, London, ON, Canada

Cheryl J Craig

Department of Teaching, Learning, and Culture, Texas A&M University, College Station, TX, United States

Elly de Bruijn

HU University of Applied Sciences Utrecht, Research Centre for Learning and Innovation, Research group Vocational Education, Utrecht, the Netherlands

Majella Dempsey

Department of Education, Maynooth University, Maynooth, Co. Kildare, Ireland

Dominik Dvořák

Institute for Research and Development of Education, Faculty of Education, Charles University, Prague, Czechia

Norm Friesen

Boise State University, College of Education, Boise, ID, United States

Judith Glaesser

Tübingen School of Education and Methods Center, University of Tübingen, Tübingen, Germany

Tom G Griffiths

Oslo Metropolitan University, Oslo, Norway

Ming Fang He

Department of Curriculum, Foundations, and Reading, College of Education, Georgia Southern University, Statesboro, GA, United States

Helen Hedges

The University of Auckland, Auckland, New Zealand

Rachel M Heydon

Western University, School of Education, John George Althouse, London, ON, Canada

Ursula Hoadley

University of Cape Town, Cape Town, South Africa

Bernadette Hörmann

University of Oslo, Oslo, Norway

Patrick Howard

Cape Breton University, School of Education and Health, Sydney, NS, Canada

Junzi Huang

Peking University, Beijing, China

David Hyatt

School of Education, University of Sheffield, Sheffield, United Kingdom

Christopher Johnstone

*University of Minnesota Minneapolis, Minneapolis,
MN, United States*

Jung-Hoon Jung

*Pusan National University, Pusan Metropolitan City,
South Korea*

Berit Karseth

*University of Oslo, Department of Education, Oslo,
Norway*

Karsten Kenklies

*University of Strathclyde, School of Education, Glasgow,
United Kingdom*

Young Chun Kim

*Chinju National University of Education, Jinju-si, South
Korea*

Maike Koopman

*HU University of Applied Sciences Utrecht, Research
Centre for Learning and Innovation, Research group
Vocational Education, Utrecht, the Netherlands*

Ole Andreas Kvamme

*Department of Teacher Education and School Research,
University of Oslo, Oslo, Norway*

Nancy Law

*University of Hong Kong, Faculty of Education,
University of Hong Kong, Hong Kong*

David Leat

*Research Center for Learning and Teaching, School of
Education, Communication and Language Science,
Newcastle University, Newcastle, United Kingdom*

Wing On Lee

*Institute for Adult Learning, Singapore University of
Social Sciences, Singapore*

Yves Lenoir

University of Sherbrooke, Sherbrooke, QC, Canada

Elizabeth Macedo

Rio de Janeiro State University, Rio de Janeiro, Brazil

Margaret Macintyre Latta

*Okanagan School of Education, Faculty of Education,
University of British Columbia, Vancouver, BC, Canada*

Cécile Mathou

*Institute of Social Sciences, University of Lausanne,
Lausanne, Switzerland*

Claire Jane McLachlan

*Institute of Education, Arts and Community, Federation
University Australia, Ballarat, VIC, Australia*

Johan Muller

University of Cape Town, Cape Town, South Africa

Jaqueline Naidoo

*University of KwaZulu-Natal, School of Education,
Teacher Development Studies, Scottsville, South Africa*

Hayley Niad

*University of Minnesota Minneapolis, Minneapolis,
MN, United States*

Nienke M Nieveen

*University of Twente, Department of Teacher
Professional Development ELAN Enschede, the
Netherlands*

Kenneth Nordgren

Karlstad University, Karlstad, Sweden

Andreas Nordin

*Linnaeus University, Department of Pedagogy and
Learning, Växjö, Sweden*

Natalie N Pareja Roblin

*University of Amsterdam, Department of Child
Development and Education, Amsterdam, the
Netherlands*

Stavroula Philippou

University of Cyprus, Nicosia, Cyprus

Louise G Phillips

Southern Cross University, Gold Coast, QLD, Australia

Thomas S Popkewitz

*University of Wisconsin-Madison, Madison, WI, United
States*

Mark Priestley

*Faculty of Social Sciences, University of Stirling, Stirling,
United Kingdom*

Axel Rivas

Universidad de San Andrés, Argentina

Lisa Rosén Rasmussen

*Danish School of Education, Aarhus University,
Denmark*

Belén Sanchez

Universidad de San Andrés, Argentina

Kirsten Sivesind

*Department of Education, University of Oslo, Oslo,
Norway*

Tiina Soini

Tampere University, Tampere, Finland

Marcin Starnawski

University of Lower Silesia, Wrocław, Poland

Daniel Sundberg

*Faculty of Social Science, Linnaeus University, Växjö,
Sweden*

Keith S Taber

*Faculty of Education, University of Cambridge,
Cambridge, United Kingdom*

Daniel Tröhler

*University of Vienna, Department of Education, Vienna,
Austria*

Jan JH van den Akker

*Curriculum Research & Consultancy Bergen, the
Netherlands*

Joke M Voogt

*University of Amsterdam, Department of Child
Development and Education, Amsterdam, the
Netherlands*

Ninni Wahlström

*Linnaeus University, Department of Education and
Teachers' Practice, Växjö, Sweden*

Csilla Weninger

*Nanyang Technological University, National Institute of
Education, Singapore*

Leesa Wheelahan

*Ontario Institute for Studies in Education, University of
Toronto, Toronto, ON, Canada*

Alison Whelan

*Research Center for Learning and Teaching, School of
Education, Communication and Language Science,
Newcastle University, Newcastle, United Kingdom*

Christine Winter

*School of Education, University of Sheffield, Sheffield,
United Kingdom*

Mary C Woolley

*Faculty of Arts, Humanities and Education, Canterbury
Christ Church University, Canterbury, United Kingdom*

Weili Zhao

*Jing Hengyi School of Education, Hangzhou Normal
University, China; and Department of Curriculum and
Instruction, The Chinese University of Hong Kong, Hong
Kong SAR, China*

Ilya Zitter

*HU University of Applied Sciences Utrecht, Research
Centre for Learning and Innovation, Research group
Vocational Education, Utrecht, the Netherlands*

This page intentionally left blank

CONTENTS OF VOLUME 7

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 7: Approaches to the Curriculum and Its Politics</i>	<i>xix</i>
<i>Contributors to Volume 7</i>	<i>xxiii</i>

CURRICULUM GOVERNANCE AND GLOBAL EDUCATIONAL POLICY

Changing forms of curriculum and governance in developing countries <i>Johan Muller and Ursula Hoadley</i>	1
Contemporary curriculum policies: globalization, recontextualization and hybridization of pedagogic models <i>Cécile Mathou</i>	8
Curriculum governance: the dilemma of the centralized curriculum <i>Axel Rivas and Belén Sanchez</i>	19
The rise and relevance of curriculum reform <i>Kirsten Sivesind</i>	28
Critical curriculum analysis framework (CCAF) <i>Christine Winter and David Hyatt</i>	35
Curriculum and 21st century skills <i>Joke M Voogt and Natalie N Pareja Roblin</i>	49

CURRICULUM AND GLOBALIZATION

Transnational curriculum inquiry <i>Elizabeth Macedo</i>	56
Re-turning to the cultural, historical, and philosophical foundations of international curriculum <i>Weili Zhao</i>	65

CURRICULUM THEORY

Contemporary trends in curriculum research <i>Berit Karseth and Ninni Wahlström</i>	74
Realism and theorizing curriculum—the role of knowledge <i>Leesa Wheelahan</i>	85
World-systems analysis and curriculum theory <i>Marcin Starnawski and Tom G Griffiths</i>	96
Rethinking critical theory/study of education: contributions of the “posts/new materialism” <i>Thomas S Popkewitz and Junzi Huang</i>	105
Curriculum theory and education history <i>Daniel Tröhler</i>	117
Curriculum theory and new institutionalism <i>Andreas Nordin</i>	126

CURRICULUM MAKING

Innovative pedagogies in relation to curriculum <i>David Leat and Alison Whelan</i>	132
Teachers mediating of curriculum making processes <i>Maria Birbili</i>	142
Curriculum development <i>Dominik Dvořák</i>	149
Curriculum and teacher education: the pre-service teacher as curriculum maker <i>Majella Dempsey</i>	155
Curriculum making and higher education: an activity-theoretical perspective <i>Johanna Annala</i>	163
Children’s interests and curriculum making in early childhood education <i>Liz Chesworth and Helen Hedges</i>	173
Children’s participation in local curriculum-making <i>Louise G Phillips</i>	181
Curriculum making and teacher agency <i>Mark Priestley, Daniel Alvunger, Stavroula Philippou, and Tiina Soini</i>	188

CURRICULUM DESIGN AND CURRICULUM INNOVATIONS

Curriculum design <i>Nienke M Nieveen, Jan JH van den Akker, and Joke M Voogt</i>	198
Learning environments in vocational education: how to elicit learning for and in practice <i>Maaïke Koopman, Ilya Zitter, Liesbeth Baartman, and Elly de Bruijn</i>	206
Language education <i>Jessie S Barrot</i>	216

Designing school(s): school architecture and curriculum <i>Lisa Rosén Rasmussen</i>	224
Students and staff co-creating curriculum in higher education <i>Catherine Bovill</i>	235

CURRICULUM, SCHOOL SUBJECT AND DIDAKTIK

Continental pedagogy & curriculum <i>Norm Friesen and Karsten Kenklies</i>	245
Curriculum and knowledge content <i>Daniel Sundberg</i>	256
Interdisciplinarity in basic education: interpretations, relevance and principles <i>Yves Lenoir</i>	267
Curriculum and competences <i>Judith Glaesser</i>	281
Shadow curriculum as a new form of curriculum: its evolution, definition, and implications for curriculum studies <i>Young Chun Kim and Jung-Hoon Jung</i>	288
History curriculum in the Anthropocene; how should we tell the story? <i>Kenneth Nordgren</i>	296
Curriculum and literacy <i>Csilla Weninger</i>	308
Curriculum and science <i>Keith S Taber</i>	314
Curriculum and social studies education <i>Steven Camicia</i>	327
Curriculum, Didaktik, and Bildung <i>Bernadette Hörmann</i>	333

LIVED CURRICULUM

Challenges to lived curriculum making: 1990–2020 <i>Cheryl J Craig</i>	339
Living schools—reframing education for future possibility <i>Patrick Howard</i>	350
Teachers living the curriculum: culturally relevant pedagogy and action research <i>Noah Borrero and Jaqueline Naidoo</i>	357
Teachers' lived experience: voices of history teachers on the English National Curriculum <i>Mary C Woolley</i>	363
Curriculum's multiplicities as participatory modes of interaction <i>Margaret Macintyre Latta</i>	371

Curriculum and identity options	381
<i>Rachel M Heydon and Carla Ruthes Coelho</i>	
Cross-cultural narrative inquiry—cross-cultural lives	392
<i>Ming Fang He</i>	

CURRICULUM, DEMOCRACY AND INCLUSION

Curriculum and the United Nations' sustainable development goals	406
<i>Ole Andreas Kvamme</i>	
Curriculum and digital citizenship	414
<i>Nancy Law and Wing On Lee</i>	
Curriculum and democracy: building understandings of democracy through early childhood practice	424
<i>Claire Jane McLachlan</i>	
Curriculum and LGBTQ+ inclusion	434
<i>Steven Camicia</i>	
Curriculum and inclusive education: universal design for learning as a "traveling" phenomenon	440
<i>Christopher Johnstone and Hayley Niad</i>	

Changing forms of curriculum and governance in developing countries

Johan Muller and Ursula Hoadley, University of Cape Town, Cape Town, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Recent curriculum reform in the developing world	1
The emergence of transnational curriculum reform discourse	2
The social logic of curriculum models: competence and performance	3
South Africa's curriculum 2005: a cautionary tale	5
Conclusion	5
References	6

Introduction

Discussions of curriculum, and proposals for curriculum change, usually swing between two poles: between a 'traditional' model based on school subjects and their content, driven predominantly by the teacher, and a 'progressive' model which favors weaker subject boundaries, thematic organization, looser specification of the content to be learnt, promoting initiative and autonomy on the part of learners. There are a number of alternative terms for these two polar models, but we will use these until we introduce a more analytical classification in the discussion below.

These are of course ideal types, and curricula have historically varied widely in the way these have been implemented, with hybrid sub-types readily discernible, though usually with one or other predominating. Disillusionment with the one seems to trigger an enthusiasm for the other. This has recently been given renewed impetus by international tests like TIMSS and PIRLS which have inevitably led to invidious comparisons between 'winning' and 'losing' schooling systems. The finger of blame is most often pointed at the curriculum model, which has led in turn to political and lay enthusiasm for supposedly 'winning' models, like the 'Finnish model' for example. In practice, curriculum systems vary widely. One key reason for this lies with the embeddedness of educational practices which often persist in their old tried and tested manner despite the overlay of a new system. Newer prescripts then become grafted onto, absorbed by, or re-interpreted in the light of established practices. This gives rise to hybrids, which at best work fairly well or, at worst, fall victim to incompatible design logics which undermine each other. The resulting failure then gives rise to the next cyclical swing.

The principal difference between curriculum reform in developed as against developing countries used to be more distinctive than it is now. In the past, the Anglo-American developed countries together with those of Europe and Asia had reasonably distinct national curriculum/pedagogical traditions – the early specializing, streamed system in England as against the late specializing quasi-system in the USA, with autonomy for the different states to set their own criteria; the *didaktik/bildung* wellspring of the Germanic and Nordic countries; the Confucian system in China and the East Asian countries to mention only a few distinct traditions. Also, European countries like France and Germany have long experience with national curricula, compared to countries like the USA where federal control of the curriculum is still strong. More latterly, the global ecumenism of TIMSS and the agencies pursuing a globalizing agenda – like the IEA, UNESCO, OECD – have constructed a marketplace for national comparisons which, despite claims of a decades-long convergence of curricula (McEneaney and Meyer, 2000), has given this convergence accelerated impetus. Convergences are now everywhere discernible, at least at the symbolic and policy level. Nevertheless, the lineaments of the national traditions of the developed countries are still readily evident (Tröhler, 2016).

Many developing countries, particularly those with a colonial past, have schooling systems which bear a distinct imprint of their colonial pasts. To take one example, South Africa's schooling system owes something to its English colonial roots and something to the Germanic roots brought to South Africa by the Dutch colonists from whence sprang apartheid ideology and 'black' education. The once-black schooling system, only partly transformed, still today educates about 80% of the school-going population in what van den Berg and Louw (2006) call 'schools for the poor'. The 80% resemble the schools in many developing countries.

Recent curriculum reform in the developing world¹

A common set of attributes is to be found in the curriculum and pedagogy of many developing countries: the lessons are run by the teacher who does most of the talking; the learners are passive unless asked to respond; the response is often in the form of collective chanting; the learners copy material into their books, and are drilled on the material (Schweisfurth, 2011; see also Carnoy et al., 2012). The pace of the lesson is usually slow, often repetitive, so that curriculum coverage is limited. When learners proceed to

¹This section and ones that follow draw on Hoadley (2018); Muller and Hoadley (2019).

the next grade, they are rarely well prepared to encounter the more complex material, and knowledge gaps are frequent, posing a hurdle to progression especially in conceptually-rich subjects like science and mathematics.

Two other features of the pedagogy are worth noting. First, learners engage curriculum content orally far more than they do textually, in either reading or written form. Books, or other forms of text, are not a regular feature of many developing country classrooms. Secondly, teachers engage the whole class far more often than they do individual learners. Since there is little explicit marking out of individual performances, performance level is a collective affair, which invariably means regression to a fairly low common denominator. Small wonder then that empirical studies have found a persistent pattern of low performance outcomes (Glewwe and Muralidharan, 2015; Pritchett and Beatty, 2012). It is this pattern that curriculum reform has sought to reverse.

Perhaps because this pattern must have seemed to observers to be a version of traditionalism, the majority of proposals for reform were a variant of progressive curriculum and pedagogy, an emphasis on relevance to both the worlds of the learners and the world of work, and a thematically driven curriculum with weakened boundaries between subjects. This model was carried to the developing world in the 1980s first by the international lenders as part of structural adjustment packages, and later with more zeal by international aid organizations with the UNESCO 'Education for All' World Declaration a high-water mark for globalizing progressivism (UNESCO, 1990).

By the early 2000s the pendulum was swinging again, and donor and international aid agencies were transferring allegiance to a model of generic skills and competencies said to be essential for the rapidly changing world of work in the 21st century. Superficially akin to progressivism – they both promoted the initiative of the learner, interdisciplinarity and relevance, for example – these features mask a more fundamental shift in the social logic of the curriculum model which marks a relative departure from progressivism. We discuss this further below.

Progressivism of one form or another has, until recently at least, remained popular in developing countries. The longevity of the popularity of progressivism has been variably assigned, but at least one feature has been the elision of education and politics: political progressivism and educational progressivism have been assumed to be two sides of the same coin, such that to be politically progressive meant to be educationally progressive. To question the one was to question the other. Small wonder that post-colonial elites have found it easy to persuade their followers that this was the way to go.

The impact of the reforms, however, has been hard to judge. The effects on student outcomes have been disappointing, and traditional practices have persisted to a greater or lesser extent. In almost all cases, evaluations have noted the observance of the outward forms of progressivism but no substantive uptake of the deeper social logic of the model. What is evident in this recurrent pattern is the recruitment of the forms of progressive pedagogy, like group work for example, to the collectivizing practices of a particular form of traditional pedagogy. This is ascribed by some to traditional cultural practices within which the teachers remain embedded (Schweisfurth, 2011).

Interesting reports from some developed countries also signal how curriculum reforms graft imperfectly onto long established traditional practices. In Singapore, for example, the pedagogy has remained relatively immune to progressive influence despite government embrace of progressive policy prescriptions from the OECD. What makes Singapore interesting is the fact that although the government has made concerted efforts to promote features of progressivism, through the introduction of programmes like 'Innovation and Enterprise' in 2004, and 'Teach Less, Learn More' in 2005, the schools have retained 'a strong focus on direct instruction and traditional pedagogic practices' (Deng and Gopinathan, 2016, p. 460). They also continue to perform near the top of the ladder on the OECD's PISA test. Interesting too is the case of late-industrializing Finland, where one principal cause of their pre-eminence in PISA outcomes has been ascribed to 'a hierarchical and traditional schooling climate' (Sahlgren, 2015, p. iii), and their more recent fall in performance to their belated shift to progressive learner centered classroom methods.

Our focus in this chapter is, however, on curriculum reform in developing contexts, where there have been few unambiguously successful cases of progressive curriculum reform. A form of traditional practice continues to hold sway, one that is more or less effective. Indeed, where there have been demonstrable achievement effects in the developing world, as in Cuba (Carnoy and Marshall, 2005) and Ecuador, and to a lesser degree in Peru and Chile (Schneider et al., 2017), these have coincided with curriculum reforms that look very like a stringent version of traditional pedagogy – with 'a strong focus on direct instruction and traditional pedagogic practices' (Carnoy and Marshall, 2005, p. 460). We are left unclear as to whether the failure of progressive reforms is because of failures of implementation, low dosage, sedimented cultural practices, or some other factor. What has not so far been examined is the social logic of progressivism and its difference from that of traditionalism. We examine this below.

The emergence of transnational curriculum reform discourse

In the shift from the Millennium Development Goals (for the period 2000 to 2015) to the 'Sustainable Development Goals' (for 2015 to 2030; see UN, 2015), the shift toward generic skills was sustained but with an additional education reform discourse around 'learning outcomes' (Verger et al., 2018). In 2018 the World Bank declared a 'global learning crisis' (World Bank, 2018). Despite impressive achievements in access to education across nations, this had not translated into increased levels of learning. Wide-spread testing using the Early Grade Reading Assessment had shown that across developing countries learners were leaving school barely able to read and write in the language of instruction despite completing six years of primary education (World Bank, 2018). Gone were ubiquitous references to learner-centered pedagogies.

What followed was a 'new learning agenda', and an emphasis on learning and learning outcomes evident in academic literature (Verger et al., 2018) and also in key reports of multilateral and donor organizations and NGOs (UNESCO, 2017/2018). The World Bank attributed one of the central reasons for the learning crisis to education systems in the developing world having little evidence around who was learning and who was not. Consequently, in the period after 2015, there was a shift to transnational measurement of basic numeracy and literacy and the establishment of the Global Alliance to Monitor Learning (GAML) to develop methodological tools and standards to 'ensure global comparability among indicators related to learning outcomes' (UNESCO, 2016).

Criticism of the new learning agenda focused on the lack of an explicit theory of learning (Tikly, 2017), 'distanced from discussion of pedagogy as debate and focused mainly on questions relating to the outcomes of learning – which outcomes we should care about and how can they be measured' (Barrett et al., 2015, p. 231). In October 2020, the GAML published a set of 'proficiency frameworks' for reading and mathematics, with 144 pages of domains, constructs, subconstructs and global proficiency descriptors for mathematics Grades 1 to 9 (GAML, 2020). This drive for ever greater specificity is the fate of outcomes-specified curricula that refrain from extensive content specification. As Knight (2001, p. 373) remarked, 'Clarification leads to complication which is why lists of outcomes grow like mold and become unwieldy'. This too was the consequence of South Africa's experimentation with a radical form of outcomes-based curriculum reform, as we briefly show below.

One response to the focus on learning outcomes has been an assertion of the need to go beyond narrow skills statements of learning to incorporate '21st century skills'. Frameworks for these sets of skills are proliferating (Global Partnership for Education, 2020). The notion of 'skill' in these documents is a relatively common-sensical one, with none of the nuance to be found in more conceptual analyses of skill like that of Winch (2012). Some argue that these skills can't be prised away from the transmission of subject content knowledge: 'While it is hard to argue against the importance of creativity, communication, collaboration and critical thinking, one has to create, communicate, collaborate and think critically about *something*' (Thornton, 2020, p. 2). Nevertheless, considerable effort is being expended in trying to work out *how* they are taught, with such efforts often framed in the abstract (Taylor et al., 2020).

The OECD has recently produced their own iteration of '21st century skills' in their 2030 Learning Framework (OECD, 2018; 2019a) and 'Learning Compass' (OECD, 2019b) which includes a range of competences that 'include knowledge, skills, attitudes and values that allow a person to act in coherent and responsible ways that change the future for the better' (OECD, 2019b, p. 2). Despite actually having been around for at least the last three decades in various iterations of core, generic, key, transferable or 21st century skills, (Tight, 2021), little progress has been made in figuring out how these are to feature in curriculum frameworks.

The OECD pulls in two directions, asserting a commitment to subject disciplinary knowledge – 'meaningful sequential structure of topics that reflect the logic of the academic discipline/disciplines on which they draw' (OECD, 2020a, p. 26) – but also to generic skills, focusing on skills to prepare for a changing workplace, but also skills that are intended to foster social cohesion and an ecological orientation to the world. The framework thus appears to want it both or all ways, pointing compass-like in all directions but with no clear true north. The metaphor of the compass itself is telling, aiming to 'emphasize the need for students to learn to navigate by themselves through unfamiliar contexts, and find their direction in a meaningful and responsible way, instead of simply receiving fixed instructions or directions from their teachers' (OECD, 2019b, p. 2). This spells out clearly the ideal learner as an autonomous life course entrepreneur. The neoliberal overtones are hard to ignore.

The social logic of curriculum models: competence and performance

The name 'traditional' is misleading in several senses. When related to curriculum, the term takes its first meaning, at least in Europe where schooling first began to massify, in the sixteenth century with Peter Ramus in Paris who was the first to codify a curriculum. Up until then, instruction was an oral art, and each teacher decided what to do as he (invariably he) proceeded. There were no set subjects or topics, and pedagogical authority vested in the oral presentation of the teacher, deriving in the last instance from Aristotle, hence also known as 'classical'. Ramus did not only produce the first written curriculum, but he also ordered the curriculum topics, both horizontally – designing topic clusters into proto-disciplines he called 'arts'; and hierarchically, from particulars to universals. In this way was born the modern 'traditional' curriculum with its ordering into subjects which were internally structured by curriculum sequencing and progression principles (Ong, 1961; Hamilton, 1990; Triche and McKnight, 2004). This new pedagogical object could also for the first time be printed by Gutenberg's printing press, and the result went through 150 editions in 45 years. For the first time the curriculum could be disseminated for consideration at a distance, and knowledge progression could be topographically displayed. This allowed the widely distributed expertise of European scholars to be harnessed and focused on refining an increasingly common project. This new pedagogical procedure came to be associated not with anything progressive but with method and seen as a tool for producing stultifying conformity.

German romanticism formed the philosophical basis for a revolt against this radical externalization of knowledge, and the basis for proto-progressivist alternatives to what was by now the mainstream in schooling. These provided the basis for nineteenth century alternatives such as Froebel, Hebart, Pestalozzi, Steiner and Montessori. Even Dewey's system derived from Hegel. The main point about these alternative systems was that, with the partial exception of Dewey, they all remained peripheral to the mainstream traditional curriculum. What they all had in common was a greater or lesser reaction to traditionalism's rationalism, an emphasis on the inner life of the learner, and the desire to release the inner nature of the learner through minimal direction and imposition.

Even at the time, not everyone welcomed the romantic promise. Antonio Gramsci said of the progressive Gentile reforms in 1929, 'The new education created a kind of church that paralyzed pedagogical research. It produced curious aberrations like "spontaneity," which supposed that the child's brain is like a ball of string that the teacher should help unwind. In reality, each generation educates and forms each new generation. Education opposes the elemental biological instincts of nature' (quoted at <http://www.nationalreview.com/article/445038/educational-reformer-hirsch-promotes-knowledge-against-its-enemies>).

These are, in brief, the antecedents to 'traditionalism' and 'progressivism'. But that tells us very little about the particular purchase progressivism has exercised on mainstream educational thinking since the 1960s, when progressivism seemed to experience a second life. Why and how did progressivism emerge from the alternative shadows and make a bid for the mainstream in the late 1960s? Bernstein (2000) traces this to a marked convergence in the social and psychological sciences around the concept of 'competence'. A 'competence' is an inner or in-built potential for some or other specialized repertoire. It is something that must be activated and actualized, not something that is acquired from outside. Bernstein goes on to anatomize the 'social logic' of the concept of 'competence':

1. One has a 'competence' – say, a competence for math – by definition, just as one has for one's mother tongue. If one is not competent, it is not something that one lacks. Rather, the 'competence' has not as yet been activated. Hence, as Bernstein (2000, p. 43) says, given time and the right conditions, 'competence' prefigures a 'universal democracy of acquisition'.
2. The learner is active and creative. External imposition of instructional content can hinder rather than enlighten the learner.
3. This means in turn that the learner self-regulates; learning should be emancipatory, not regulatory.
4. This leads naturally to a suspicion of pedagogical hierarchy, that between learners and between the learners and the teacher. The best teacher is thus a facilitator and midwife to the emergence of competence.
5. The focus is on the present, not the past (the canonical curriculum and its content) nor on an anticipated future where the reality of work and its specialized cognitive requirements lies in wait; the active emancipatory present is the target.

'Competence' is thus the emblem for the rejection of the externalized knowledge necessary for a common curriculum; 'acquisition', like 'transmission', points to educational activities of the outside and thus transgresses against the creative inner, the ideal wellspring of knowledge in the progressive classroom (see also Wahlström, 2016). Bernstein traces the new mainstream allure of this idea in education to various social movements of the 1960s – the rise of militant student protests, rock 'n roll, the emergence of the 'youth' as a sociological and political subject, and anti-authoritarianisms of various kinds. We might add, the emergence of popular democracy movements. It is small wonder that a younger cohort of educators and teachers would be drawn to the promise of emancipation, democracy and equality that progressivism embodied. In practice, many progressive schools do not have quite the disdain for specialized knowledge that this depiction might suggest, and some have a tradition of producing high levels of performance outcomes. This has not been the case for schools in developing countries, as we have discussed above.

Above all, 'competence' presaged a genuine rupture with the staid and traditional model of schooling that had reproduced inequality so remorselessly, because of a special property. Competence did not so much embrace a route out of power relations; instead, it sought to circumvent them. As Bernstein put it, 'competences are beyond the reach and restraints of power relations and their differential unequal positionings' (Bernstein, 2000, p. 147). No wonder it was, and is, seen as radical. It is a self-legitimizing notion, a 'means of bestowing ideological purity' (Bernstein, 2000, p. 148). To put it simply, the competence-progressive promise was to transcend the cruel strictures of class-determined educational destiny.

If the competence model locates its principle of dynamism as internal to the learner, the principle of dynamism of the alternative model, which Bernstein calls a performance model, is located externally, either in the intended curriculum and its design or in the market. A performance-based curriculum is driven by the amount of content to be covered or tasks to be mastered, by the pace at which they are to be covered, and by the sequence in which they are set. What is evaluated in the competence model is the competence, or potential, of the learner as activated; in the performance model, it is the learner's performance on, or mastery of, the content or tasks prescribed by the intended curriculum.

Two sub-forms of the performance model are relevant for present purposes. They are, first, a specialized sub-form developed during the slow growth of elite into mass schooling. This sub-form is sometimes aligned with elite schooling as a basis for its rejection. Undoubtedly, performances form a hierarchical rank which delivers 'winners' and 'losers' (Sandel, 2020). We will refer to this sub-form as the 'autonomous performance form'. The second is a sub-form we shall call the 'generic performance form', where the external dynamism comes not so much from the knowledge base itself as from the presumed requirements of the marketplace, either the one currently in place, increasingly the ones presumed to be in force in some or other future. The content to be mastered is embodied in, and indexed by, generic skills such as 'working in a team', 'communication skills', sometimes referred to collectively as 'core', 'key' 'critical', 'transversal' (EU 2007) or '21st century skills' (Trilling and Fadel, 2009; McPhail and Rata, 2018). Qualifications frameworks and outcomes-based curriculum systems are systems that stipulate the curriculum in such generic skills-based terms, which if left unqualified, lead to an under-specified curriculum.

The generic performance sub-form is a confusing one because, as Bernstein says, it shares fundamental features in common with the competence model, in that it stipulates general attributes which all learners should, and the implication is could, attain. It shares some of the democratic aura of competence, and is consequently at times seen as progressive, as a component part of a competence model. Wahlström (2016), in an analysis of the impact of recent EU prescripts of 'transversal skills' on the Swedish school curriculum, using Bernstein's competence and performance template, sees it in these terms.

Our argument is rather that future oriented generic skills are external to learners, not part of their internal creative spark. As Allais (2014) points out, the rationale for all generic skills is to be of instrumental service to the present or future market, whether that 'market' is couched in labor market terms or, as is increasingly the case, in terms of the future life course of the individual learner. This makes generic skills integral to a neoliberal rather than a progressive emancipatory social project which would regard any attempt to bend the inner creative potential of learners to instrumental ends as a deformation of creativity. We view the generic performance sub-form as a close fit with '21st century skills', albeit often couched in competence-style rhetoric.

What this means in terms of Bernstein's conceptualization is that the 'third wave' of transnational curriculum policy marks a shift in tension, from that between competence and performance, to one between sub-forms of the performance model. While this might superficially seem to be a less pernicious tension, it masks a central feature of the generic sub-form, namely, that it is inimical not only to disciplinarily based knowledges but is more particularly inimical to specifying curriculum content in anything other than learning outcome terms. We will illustrate this by means of a brief excursus into South African curriculum history, only because South Africa encountered a radical form of a generic curriculum early on, in the 1990s. The point of the argument then is to illustrate what this 'third wave' generic performance form might mean for developing country learners if it becomes the next export to these countries.

South Africa's curriculum 2005: a cautionary tale

South Africa launched its first post-apartheid curriculum with a radical skills agenda, comprising an outcomes-based framework coupled with a commitment to learner-centred and constructivist pedagogy in the classroom. This national curriculum, Curriculum 2005, started first in the junior grades, but was to be progressively rolled out to all grades by 2005, hence the name (Department of Education, 1997). At its heart was something called radical outcomes-based education. Although couched in discourses of transformation and democracy, the intended national curriculum was an extreme version of an outcomes-stipulated curriculum, a radical generic performance curriculum somewhat ahead of the global vogue spelt out above.

Like all outcomes-based systems, it looked upon traditional school subject boundaries as impediments to be transcended, so in their place were eight transdisciplinary learning areas, to be translated into learning programmes by means of a set of design features, none of which stipulated the content to be learnt: there were twelve critical outcomes, which would now be called '21st century skills'; 66 specific outcomes which despite the name were generic across grades; 237 assessment criteria; range statements specifying level and depth (but not content); performance indicators; and Expected Levels of Performance. Teachers, in addition to supplying all their own content, were expected to report on all these indicators for every class (Hoadley, 2018).

Problems arose almost from the start, which the designers responded to by adding generic design features, 'stipulating finer and finer levels of outcomes' (DOE, 2000, p. 30) in line with Knight's remark above. When that did not help, the Minister of Education commissioned a review. Despite phrasing the result in diplomatic terms, the verdict was damning, finding the curriculum 'technically over-designed, yet remains under-specified' (DOE 2000, p. 41). The Review further found incoherence amongst the design features, with no conceptual sequence and progression path stipulated. This was particularly disabling for knowledge domains dependent on selection and sequencing, like Mathematics. This left under-prepared teachers and learners without a curricular road map (Jansen, 1998).

The resultant chaos was gradually revised in three iterations. This entailed a gradual shift to a knowledge-based, or autonomous performance curriculum model, leading to the Curriculum and Assessment Policy Statement, or CAPS (DBE, 2011). CAPS entailed detailed content specification and strong controls over the sequence and pace of curriculum coverage, viewed in some quarters as a neoconservative shift back toward the traditional curriculum to which it bore some structural similarities (Shalem, 2018). Under the CAPS there has been gradual improvement in student achievement outcomes, although off a low base.

At the end of 2020, pressure from the global generic skills movement is again felt in the Department of Basic Education with calls to re-integrate '21st century skills' into the curriculum. The thrust this time is a neoliberal one, unlike the progressive intentions and pedagogy of Curriculum 2005, focused predominantly on the changing nature of work now and in the future. In addition, and particularly under the conditions of the COVID-19 pandemic in 2020, a discourse of 'self-directed learning' would appear to stress again the virtues of the 'autonomous learner' and personal responsibility for success, or as the case so often is in developing contexts, failure.

Conclusion

This chapter has looked at the age-old tension between traditional and progressive curriculum models, between what Bernstein has called performance and competence curricular forms, and has discussed how a progressive learner-centered form of the competency was exported to less-developed countries. We have argued that while this tension can be said to have framed curriculum debates in the past, a new tension has more recently come to the fore. This is the tension between the autonomous and generic sub-forms of the performance model. The generic sub-form has adopted aspects of the competence model, at least rhetorically – the preference for an integrated rather than a subject-based curriculum, and a focus on the self-directed individual. But the differences are more significant; the competence model might prefer not to over-prescribe content, but outcomes-specification design blocks it. Where the first takes a moral position against content, allowing it to creep in by the back door, the latter's design precludes it. Finer and finer levels

of outcome specification do not mitigate this design problem. Both forms take the individual as their principal actor, but where the competence individual progressively releases and specializes their inner creativity and innovativeness, the generic individual accumulates outcomes in Sysiphean fashion (Muller, 1998), each of which has a limited specialization span. The former curriculum faces toward a world of individual innovation, the latter toward the market.

The final question is why generic curricula have proved so attractive. The proponents will say that the generic curriculum is the only way to prepare learners for a world of change, now but especially in the future as the rate of change speeds up. A more mundane but nonetheless pressing reason might be that it becomes progressively harder to make sensible selection of content, to cater for both breadth and depth. This may be because specialization has advanced to such an extent that new knowledge keeps crowding into the curriculum, or it may be because of the proliferation of new fields, like media and data literacy, bidding for inclusion in a curricula vehicle that has limited possibilities to expand. Either way, curriculum overload has become a serious problem globally (OECD, 2020b). It is no longer enough to say that the curriculum should concentrate on 'big ideas'. It may seem that the solution is, by means of outcomes-stipulation, to sidestep the problem of content proliferation altogether.

As we have argued, however, outcomes-stipulation comes with consequences that makes life difficult for teachers everywhere, but especially in classrooms in less-developed countries, where poorly prepared teachers rely on an explicit roadmap of the kind that the generic performance curriculum model is unable to provide. If this kind of model is exported to less-developed countries, then their chances of narrowing the performance gap with the more-developed countries will be vanishingly small.

References

- Allais, S., 2014. *Selling Out Education: National Qualifications Frameworks and the Neglect of Knowledge*. Sense Publishers, Rotterdam.
- Barrett, A., Sayed, Y., Schweisfurth, M., Tickly, L., 2015. Learning, pedagogy and the post-2015 education and development agenda. *Int. J. Educ. Dev.* 40, 231–236.
- Bernstein, B., 2000. *Pedagogy, Symbolic Control and Identity*. Rowman and Littlefield Publishers, London.
- Carnoy, M., Chisholm, L., Chilisa, B., 2012. *The Low Achievement Trap: Comparing Schools in Botswana and South Africa*. HSRC Press, Pretoria.
- Carnoy, M., Marshall, J., 2005. Cuba's academic performance in comparative perspective. *Comp. Educ. Rev.* 49 (2), 230–261.
- Deng, Z., Gopinathan, S., 2016. PISA and high performing education systems: explaining Singapore's educational success. *Comp. Educ.* 52 (4), 449–472.
- Department of Basic Education, 1997. *Curriculum 2005: Lifelong Learning for the Twenty-First Century*. Department of Basic Education, Pretoria.
- Department of Basic Education, 2000. *A South African Curriculum for the Twenty First Century: Report of the Review Committee on Curriculum 2005*. Department of Basic Education, Pretoria.
- Department of Basic Education, 2011. *Curriculum and Assessment Policy Statement (CAPS)*. Department of Basic Education, Pretoria.
- European Union, 2007. *Schools for the Twenty First Century*, Commission Staff Working Paper, SEC (2007), 1009. Commission of the European Communities, Brussels.
- GAML, 2020. *Global Proficiency Framework for Mathematics: Grades 1 to 9*. UNESCO, Paris.
- Global Partnership for Education, 2020. *21st Century Skills: What Potential Role for the Global Partnership for Education?* Retrieved from: <https://www.globalpartnership.org/sites/default/files/document/file/2020-01-GPE-21-century-skills-report.pdf>.
- Glewwe, P., Muralidharan, K., 2015. *Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications*. University of Oxford, Research on Improving Systems of Education (RISE), Oxford.
- Hamilton, D., 1990. From curriculum to 'bildung' (some preliminary considerations). In: Paper presented to the International Standing Conference for the History of Education, Prague, 23–26 August 1990.
- Hoadley, U., 2018. *Pedagogy in Poverty: Lessons of Twenty Years of Curriculum Reform in South Africa*. Routledge, London.
- Jansen, J., 1998. Curriculum reform in South Africa: a critical analysis of outcomes-based education. *Camb. J. Educ.* 28 (3), 321–331.
- Knight, P., 2001. Complexity and curriculum: a process approach to curriculum-making. *Teach. High. Educ.* 6 (3), 369–381.
- McEneaney, E., Meyer, J., 2000. The content of the curriculum: an institutionalist perspective. In: Hallinan, M. (Ed.), *Handbook of the Sociology of Education*. Springer, Boston, pp. 189–211.
- McPhail, G., Rata, E., 2018. A theoretical model of curriculum design: 'Powerful knowledge' and '21st century learning'. In: Barrett, M., Hoadley, U., Morgan, J. (Eds.), *Knowledge, Curriculum and Equity: Social Realist perspectives*. Routledge, London, pp. 63–79.
- Muller, J., 1998. The well-tempered learner: self-regulation, pedagogical models and teacher education policy. *Comp. Educ.* 34 (2), 177–193.
- Muller, J., Hoadley, U., 2019. Curriculum reform and learner performance: an obstinate paradox in the quest for equality. In: Spaull, N., Jansen, J. (Eds.), *South African Schooling: The Enigma of Inequality – A Study of the Present Situation and Future Possibilities*. Springer, Dordrecht, pp. 80–101.
- OECD, 2018. *The Future of Education and Skills Education 2030: The Future We Want*. Retrieved from: [http://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](http://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf).
- OECD, 2019a. *Future of Education and Skills 2030. Conceptual Learning Framework: Knowledge for 2030*. OECD, Paris.
- OECD, 2019b. *Future of Education and Skills 2030. Conceptual Learning Framework: Learning Compass 2030*. OECD, Paris.
- OECD, 2020a. *Curriculum (Re)design: A Series of Thematic Reports from the OECD Education 2030 Project. Overview*. OECD, Paris.
- OECD, 2020b. *Curriculum Overload: A Way Forward*. OECD Publishing, Paris.
- Ong, W., 1961. Ramist classroom procedure and the nature of reality. *SEL Stud. Engl. Lit.* 1500–1900 1 (1), 31–47.
- Pritchett, L., Beatty, A., 2012. *The Negative Consequences of Overambitious Curricula in Developing Countries*. Centre for Global Development, Washington. CGD Working Paper 293.
- Sahlgren, G., 2015. *Real Finnish Lessons: The Story of an Education Superpower*. Centre for Policy Studies, London.
- Sandel, M., 2020. *The Tyranny of Merit: What's Become of the Common Good?* Farrar, Strauss and Giroux, New York.
- Schneider, B., Estarellas, P., Bruns, B., 2017. *The Politics of Transforming Education in Ecuador: Confrontation and Continuity, 2006–2017*. Unpublished paper. World Bank, Washington.
- Schweisfurth, M., 2011. Learner-centered education in developing country contexts: from solution to problem? *Int. J. Educ. Dev.* 31, 425–432.
- Shalem, Y., 2018. Scripted lesson plans: what is visible and invisible in visible pedagogy? In: Barrett, M., Hoadley, U., Morgan, J. (Eds.), *Knowledge, Curriculum and Equity: Social Realist perspectives*. Routledge, London, pp. 183–199.
- Taylor, R., Fadel, C., Kim, H., Care, E., 2020. *Competencies for the 21st Century: Jurisdictional Progress*. Washington D.C: Brookings Institute and Centre for Curriculum Redesign. Retrieved from: https://www.brookings.edu/research/competencies-for-the-21st-century-jurisdictional-progress/?utm_campaign=Center%20for%20Universal%20Education&utm_medium=email&utm_content=100263582&utm_source=hs_email.
- Thornton, S., 2020. (Re)asserting a knowledge-building agenda in school mathematics. *Math. Educ. Res. J.* <https://doi.org/10.1007/s13394-020-00322-1>.

- Tight, M., 2021. Twenty-first century skills: meaning, usage and value. *Eur. J. High Educ.* 11 (2), 160–174. <https://doi.org/10.1080/21568235.2020.1835517>.
- Tikly, L., 2017. The future of Education for All as a global regime of educational governance. *Comp. Educ. Rev.* 61 (1), 22–57.
- Triche, S., McKnight, D., 2004. The quest for method: the legacy of Peter Ramus. *Hist. Educ.* 31 (1), 39–54.
- Trilling, B., Fadel, C., 2009. *21st Century Skills: Learning for Life in Our Times*. Wiley Publishers, New Jersey.
- Tröhler, D., 2016. Curriculum history or the educational construction of Europe in the long nineteenth century. *Eur. Educ. Res. J.* 15 (3), 279–297.
- UNESCO, 1990. World declaration on education for all: meeting basic learning needs. In: *World Conference on Education for All*, Jomtien, March 1990. UNESCO, Paris.
- United Nations, 2015. Sustainable Development Goals. <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>.
- UNESCO, 2016. Education for People and Planet: Creating Sustainable Futures for All. EFA Global Monitoring Report. UNESCO, Paris.
- UNESCO, 2017/2018. Accountability for All: Meeting Our Commitments. EFA Global Monitoring Report. UNESCO, Paris.
- Van den Berg, S., Louw, M., 2006. Lessons learnt from SACMEQII: South African student performance in regional context. In: *Paper presented to the Investment Choices for Education in Africa Conference*, Johannesburg, 19–21 September, 2006.
- Verger, A., Novelli, M., Altinyelken, H.K. (Eds.), 2018. *Global Policy and International Development. New Agendas, Issues and Policies*. Bloomsbury, London.
- Wahlström, N., 2016. A third wave of European education policy: transnational and national conceptions of knowledge in Swedish curricula. *Eur. Educ. Res. J.* 15 (3), 298–313.
- Winch, C., 2012. *Dimensions of Expertise: A Conceptual Exploration of Vocational Knowledge*. Continuum, London.
- World Bank, 2018. *World Development Report 2018: Learning to Realize Education's Promise*. World Bank, Washington, DC. <https://doi.org/10.1596/978-1-4648-1096-1>.

Contemporary curriculum policies: globalization, recontextualization and hybridization of pedagogic models

Cécile Mathou, Institute of Social Sciences, University of Lausanne, Lausanne, Switzerland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	8
Bernstein's pedagogic models: a heuristic framework to compare curricular transformations	10
Curriculum policies in France and Quebec: a comparison	11
Competence-based curricular reforms in two contrasting education systems	11
The Pedagogic Renewal in Quebec	12
The Common Core in France	13
The recontextualization of common trends into two national contexts	14
Conclusion	16
References	16

Introduction

Since the late 1990s, important reforms aimed at transforming curricula in compulsory education have been implemented in many industrialized countries. Such developments have taken place in a context where comparisons of education systems and students' competences have taken center stage through the work of the Organization for Economic Co-operation and Development (OECD). Thus, in the 1990s the OECD has defined key competences, that are independent from national curricula:

In its PISA and DeSeCo projects, the OECD promotes a view that what matters is not what is learned about the world, but what competencies and orientations (such as motivation to be a "life-long learner") are being produced in learners, as if these are essentially culture-neutral and context-free. The assessment programs themselves are not tied to what schools have taught in particular countries, but to what problem-solving abilities and "competencies" are agreed to be the desired outcomes of mathematics or literacy or science.

Yates and Grumet (2011, pp. 5–6).

This has increased concerns about the requirements of the "knowledge society," shifting the attention of policy makers to the curriculum (Robertson, 2005; Yates and Young, 2010). Contemporary curriculum reforms share important common features. A perceived change in the skills, knowledge and dispositions needed for success in modern societies and a globalized economy has contributed to a shift toward the use of "competency or skills-based curriculum" (Whitty, 2010). According to Young (2011), assumptions of the knowledge society has led governments to focus on four priorities: (i) a "generic" approach which plays down specific knowledge contents in favor of general and "content-free" competences and skills, (ii) individual learner choice rather than curriculum prescription playing a major role in how the sequencing of learning is structured, (iii) the increasing importance of informal or "non-school" learning and experience, and (iv) the shift from syllabus to learning outcomes. In contrast, "traditional" curricula centered on subject-specific contents are being dismissed as less valuable in the knowledge economy (Mangez, 2008; Robertson, 2005).

As a result, some scholars argue that we can observe the diffusion of a "global curriculum" (Meyer, 2007). Looking at worldwide patterns in compulsory education curricula, it can indeed be argued that much of the formal content of these curricula have become standardized. According to world culture theory, educational systems have converged toward the same "world standards" with regard to the structure, organization, and content of education. Meyer and colleagues (Kamens et al., 1996; Meyer et al., 1992) have shown that even curricula within highly dissimilar countries have converged in the last few decades, especially at primary and lower secondary education. More recently, the analysis of curriculum guidelines from about a hundred countries collected by UNESCO points to a shift from content-oriented to learning-oriented curricula, while individuals are seen as self-regulated in their approach to knowledge (Rosenmund, 2007).

World culture theorists argue that a single global model of schooling has spread around the world, however the idea that globalization has led to policy convergence is one that continues to be much debated among education policy researchers.

More recent analysis invites us to use the world "global" cautiously. Based on a literature review of available research on competence-based education reforms, Anderson-Levitt and Gardinier (2021) found that such reforms seem to have strongly influenced European countries, peaking in the early 2000s. They have been embraced by some African countries, and have affected some Australian states and some Canadian provinces, but only touched the United States lightly, and only anecdotally affected Latin America and Asia.

Secondly, whereas world culture theorists have identified common patterns, researchers looking closely at specific countries or regions tend to find that curricular transformations are more complex and less unidirectional. In certain cases, curricular policies have moved “back” to older forms, for example the standards movement in the US that work at “countercurrent to reforms in which students drive the curriculum” (Anderson-Levitt, 2003, p. 12). Other countries have implemented competency-based reforms but later stepped back from some or all of the reform, such as Sweden, England, Japan, Poland and the Flemish community of Belgium, shifting back toward more discipline-focused curricula in a pendulum swing movement (Anderson-Levitt and Gardinier, 2021).

In some national contexts, contrasting policy responses to globalization have been identified, illustrating the tensions arising between the desire to experiment new forms of learning and teaching, and the pressure to perform on standardized tests (Lingard and McGregor, 2014). Some scholars highlight the tension between the logic underlying contemporary curricular policies, for example between the rise of so-called constructivist models, the call for active learning and creativity and, more recently, a focus on fundamental and “basic” knowledge (notably language and mathematics) and performance measured by standardized assessments (Lingard and McGregor, 2014; Rochex et al., 2011). These performative pressures are linked to the implementation of accountability policies and “post-bureaucratic” modes of regulation (Maroy, 2009), i.e. involving control based on results and no longer on compliance with the law and rules. As highlighted by Anderson-Levitt and Gardinier (2021), much more investigation is needed to establish the link between competency-based approaches to New Public Management and standardized assessments.

Finally, while a certain convergence seems to be effective in terms of national policy statements (Rosenmund, 2007; Voogt and Roblin, 2012), other research focusing on national case studies shows that curricular models do not spread identically, in particular because of translation phenomena (Anderson-Levitt, 2003) favored by the semantic vagueness of the concepts at the heart of these reforms, such as the notion of competence (Ropé and Tanguy, 1994). This polysemous notion has led to hybrid forms of discourse and policy implementation. Recent empirical work in comparative education has therefore focused on the ways in which policies change as they travel across different political, social, and geographic contexts, and how they are variously and uniquely taken up in national policy settings, using a variety of concepts from indigenization, vernacularization, recontextualization, or creolization:

Empirical research in the global education policy field suggests that national institutions and domestic politics are key to understanding, on the one hand, the uneven level of diffusion and penetration of global education policy ideas in different territories, and, on the other, the re-contextualization and (on occasion) the drastic transformation of such global ideas within local institutions, networks of rules, and local practices.

Mundy et al. (2016, p. 9).

Despite renewed interest in “curricular” issues (Lingard, 2010; Priestley and Biesta, 2013) the analysis of these complex dynamics is still underdeveloped, as curricular issues (knowledge, pedagogy, in other words, the internal aspects of “schooling”) have long been neglected by comparative education in favor of structures and organization of systems (Kazamias, 2009). Few empirical, in-depth comparative research have been carried out on contemporary curricular policies, while macro-level, cross-national comparisons only allow to identify broad, common trends. Large-scale comparative analyses are often based on international documentation and look at the curriculum from their most external and superficial dimension (Forquin, 2008). Nor do these studies make it possible to take account of the temporal depth of curricula and the changes that have taken place in them. Finally, the peer-reviewed academic literature is relatively slim compared to a much more extensive gray literature on the topic (Anderson-Levitt and Gardinier, 2021).

The findings presented in this chapter aimed to contribute to this growing but under-developed field through a synchronic and diachronic comparison of curricular transformations in two francophone education system: France and Quebec.

In many francophone education systems, the “curricular reconfigurations” in progress since the end of the 1990s are underpinned by a “competence-based approach” (Hasni and Lebeaume, 2010; Legendre, 2008). France and Quebec provide two examples of such competence-based reforms. In these two education systems, the study programs for compulsory schooling underwent a major overhaul in the mid-2000s. These curricular transformations have given rise to heated controversies, both professional and scientific. Within the scientific community, interpretations of the nature and meaning of these transformations differ widely, from the profound restructuring of learning through the advent of a new “paradigm,” to the instrumentation of teaching contents to satisfy the needs of the “knowledge economy.” These interpretations are not always based on empirical research, and above all, they are based on oppositions that create artificial dichotomies that limit our understanding of contemporary educational policies and their implications. Using Alexander’s terms, the debates revolve around an “oppositional discourse polarized around dichotomous pedagogies” (2006, p. 733)—e.g. progressive vs. traditional, child-centered vs. subject centered.

The rest of the chapter is divided into three sections. First, we present the main conceptual framework that informed the study. The sociology of the curriculum, and particularly the concepts developed by Bernstein (1975, 1990) provide heuristic tools to analyze curricular changes comparatively. The following and central section focuses on the comparison of the two policies and shows how common orientations are recontextualized according to logics and dynamics specific to each educational system, resulting in both countries in hybrid forms or “mode mixing” (Muller, 1998) combining elements of contrasting pedagogic models. Finally, the conclusion returns to the partial convergences and variations observed between the two cases, highlights the ever-changing dynamics underlying contemporary curricular policies and what it means for the further development of comparative curriculum research.

Bernstein's pedagogic models: a heuristic framework to compare curricular transformations

The analysis that led to the findings presented in this chapter drew extensively on the conceptual framework Bernstein developed in his latest work in relation to educational change in the 1980s and 1990s. His concepts provide important tools for understanding the complex “dynamics” of educational change (Sadovnik, 2006) and accounting for the hybrid forms that contemporary curricula can take, mixing different models to respond to conflicting demands (Muller, 1998).

The “language of description” he developed aimed precisely at facilitating empirical research in a variety of contexts, allowing researchers to map empirically curricular changes across three key dimensions, which he referred to as “message systems”:

Formal educational knowledge can be considered to be realized through three message systems: curriculum, pedagogy and evaluation. Curriculum defines what counts as valid knowledge, pedagogy defines what counts as a valid transmission of knowledge, and evaluation defines what counts as a valid realization of this knowledge on the part of the taught.

Bernstein (1971, p. 41).

This ability to embrace curriculum, pedagogy and assessment all at once contributes to the strength of this conceptualization (Broadfoot and Pollard, 2006). In particular, the concepts of classification, framing and evaluation provided a common language to characterize curricular transformations in both education systems. These are three constitutive dimensions for the study of pedagogic discourses and practices.

Classification refers to the organization of knowledge or how different contents co-exists within the curriculum. A strong classification refers to a highly differentiated curriculum, a high degree of separation between categories (between disciplines, and within a discipline, between different contents); a weak classification refers to an integrated curriculum in which boundaries are blurred, permeable, with less specialized discourses. A strong classification can also mean that the boundaries between school and everyday knowledge are made visible. There is a strong classification between inside and outside the institution, and a hierarchy of knowledge exists between “common sense” knowledge and school knowledge. This is referred to as “collection” type curricula as opposed to the “integrated” type in which contents are subordinated to an overall conception and become part of a whole (Bernstein, 1971).

Framing refers to the transmission of knowledge and the distribution of control over pedagogic communication or “who controls what.” Strong framing is where the transmitter has explicit control over the communication; weak framing gives the learner more apparent control over the communication. Framing refers, on the one hand, to the control that is exercised over the selection of pedagogic communication, its sequencing (what comes first, what comes next) and pacing (or rate of expected acquisition). On the other hand, framing refers to the “regulatory discourse” which governs the forms that hierarchical relations take in the pedagogic relationship and expectations in terms of conduct, character, manner. These will vary according to a more or less strong framing—learners are expected to be conscientious, attentive, careful, hard-working, and at the opposite pole, to be creative, interactive, and to make his or her own mark.

Evaluation is the third key dimension of the pedagogic relationship. Evaluation criteria may be explicit or implicit. When the criteria for producing the expected text are explicit, the assessments are such that the learner knows on what he or she is being assessed in a particular assessment. It is not so much a question of producing an original and personal text as of reproducing specialized knowledge. Conversely, when the assessment is such that only the teacher knows what he or she is trying to assess, the criteria are implicit or diffuse.

These concepts are central to understanding the specificity of what Bernstein calls the pedagogic discourse, that is, the specialized form of communication encompassing curriculum, pedagogy and evaluation: what is taught to whom, when, where, how it is transmitted and how it is evaluated or deemed as acquired (Bernstein, 1990).

On the basis of these three dimensions, Bernstein (1975) identified two generic types of pedagogic discourse and practice. To describe them briefly, the visible model is characterized by a strong classification of contents, a strong framing of pedagogic relations, and explicit evaluation criteria. Visible pedagogy always places the emphasis on the learner's performance, and the extent to which it is meeting the expected criteria. An “objective” grid exists to assess learners, a grading system allows to compare learners and schools. In terms of sequencing, the rules are explicit and mark very clearly stages (what should be acquired when, in which order). This pedagogy is often based on behaviorist approaches to learning. In the case of an invisible pedagogy, classifications weaken and tend toward an integrated type of school knowledge code. Sequencing is loose, implicit, and only the teacher knows the specific objectives he or she is pursuing, objectives that are underpinned by a theory of the child's development. Invisible pedagogy is less concerned with ranking learners and comparing productions, because differences between learners' achievements reveal the uniqueness of each individual—thus they cannot be used to compare and stratify learners. Their focus is not upon a “gradable” performance but upon procedures internal to the acquirer (cognitive, linguistic, affective, motivational).

Bernstein (1990) later enriched his concepts (and the main, dichotomous opposition visible/invisible) in particular to capture the contemporary curricular reforms taking place in the UK. He further distinguished between two modalities of visible pedagogy. One is called “autonomous,” the other “dependent,” primarily oriented toward external demands (especially relevance to the labor market):

Here the opposition is between a pedagogic practice dependent upon the market place for its orientation and legitimation [...] and a pedagogic practice independent of the market place, claiming for itself an orientation and legitimation derived from the assumed autonomy of knowledge (1990, p. 63).

Table 1 Invisible pedagogy and visible pedagogy (autonomous and dependent modalities).

<i>Dimensions of pedagogic discourse</i>	<i>Invisible pedagogy</i>	<i>Visible pedagogy—dependent (or market oriented)</i>	<i>Visible pedagogy—autonomous</i>
Classification	Weak boundaries between disciplinary subjects and between school and non-school knowledge.		Strong—usually discipline-based curricular structures.
Framing	Weak Learner conceived as self-regulating, creative and active in the construction of knowledge and meaning. Personalized or individualized control over learners.	Strong Learners have relatively little control over selection, sequence and pace. Hierarchies in teaching-learning relationships are more explicit.	
Evaluation	Evaluation criteria and rules for producing the legitimate texts are implicit and diffuse. A theory of child development underlies the evaluation made by the teacher (interpreting the manifestations of inner competence).	Evaluation criteria and rules for achieving a legitimate text are relatively explicit. Emphasis is on a specific output expected from the learner. Learners' texts (performances) are objectified by grades and stratified.	

Adapted from Bernstein (1990).

The “autonomous” visible pedagogy (which, from a diachronic point of view, precedes the dependent model) aims to construct autonomous subjects, emancipated from their family and local belonging, thanks to a teaching body largely independent from the political power and local communities. It promotes a school culture that preserves a high degree of autonomy from economic and social pressures, emphasizing knowledge as a factor of intellectual enrichment: “The autonomous visible pedagogy justifies itself by the intrinsic worthwhileness and value of the knowledge it relays and by the discipline its acquisition requires” (Bernstein, 1990, p. 87).

The second form of visible pedagogy has developed in a context centered on the efficiency of education and relevance to market demands. Since the late 1980s, the fundamental transformation in the field of educational policy lies in the institutionalization of models that value above all performance and the objective of integrating individuals into the market, in a context where work and life can no longer be based on stable predictions of the future. In that sense, it is externally regulated. This pedagogic model is characterized by instrumentalism (focusing on school rankings, skills as an investment to increase future educational and professional opportunities). The learning acquired at school becomes resources for inter-individual competition, linked to the imperative of employability (van Zanten, 2009). According to Bernstein, this dependent, instrumental model incorporates some of the criticisms that could have been expressed with regard to the autonomous model: elitism, knowledge that refers only to itself, indifference to the consequences in terms of stratification, obsolete curricula that do not arouse the interest of children from working-class backgrounds. This dependent model thus tends to translate into “mixed” pedagogic regimes drawing on features of invisible pedagogy: for example, greater permeability between school and life. In this sense, this new pedagogic discourse incorporates features of apparently oppositional discourses (Table 1).

Curriculum policies in France and Quebec: a comparison

Competence-based curricular reforms in two contrasting education systems

The two reforms under study resulted from large-scale national debates which insisted on the need to adapt the curricula to the needs of the 21st century but also to ensure the success of all pupils and address the problem of early school leaving (Box 1).

If the French and Quebec education systems share common characteristics, these are also two contrasting systems in terms of organization, weight of the different institutions, school culture, values and curricular traditions. The French education system is characterized by a degree of administrative centralization that remains strong, while in Quebec regional education bodies have a relatively high degree of autonomy. Values of universalism (Holmes and McLean, 1989) remain central to the French system, while at the core of the Quebec system is the project of building a specific education project for Quebec in an Anglophone Canadian context (Rocher, 2004). North American influences such as pragmatism have also been present in Quebec since the 1960s (Lenoir, 2005). Indeed, since the Quiet Revolution, Quebec has distanced itself from the French-style education model (typical of the classical colleges, where the elites of French Canada had been educated until then) in the name of democratizing education. The “Parent Report” (1963–1965), which led to the most important educational reform in Quebec, included a strong criticism of the obsolete and elitist model of the classical colleges, with its bookish pedagogy and emphasis on scholarship. The “renewed” humanism, drawing on humanistic psychology and pragmatism, valued so-called active pedagogy, the taking into account of children’s particularities, and the recognition of the pluralistic character of society (Lenoir, 2005).

Box 1 Research methodology

This chapter draws on a doctoral research project carried out between 2012 and 2017 which aimed at analyzing contemporary curricular transformations in lower secondary education in France and Quebec (Canada).

The multi-level comparative research focused on compulsory education, excluding vocational education and higher education which have their own distinct history of competency-based reforms. The *collège* in France and first cycle of secondary school in Quebec correspond to the last years of compulsory education, where all pupils follow a common pathway (before moving into different tracks, e.g. general upper secondary education or vocational education and training for instance).

The study aimed to compare how similar “competence-based reforms” in lower secondary education—partly drawing on the same ideas, pedagogic models and experts—were recontextualized at the “national” and meso level. In Canada, education comes under provincial jurisdiction. Each province has the power to establish its own education system and to make all decisions regarding schools, teachers and curriculum. We therefore considered Québec’s education system as a unit of analysis comparable to the French one. The second strand of the research (which is not presented in this chapter) examined the practices of pedagogic advisors (Quebec) and inspectors (France) based on fieldwork carried out in three mid-level education authorities.

The research focused on what the Ministry legitimizes as “official pedagogic discourse,” constituted by the legal and regulatory texts governing the functioning of schools, official syllabus, teaching guidelines and other support documents. The Quebec corpus comprised 18 texts and the French corpus 24 texts. They were constituted by the following texts:

- Legal and regulatory texts. These include the study programs which constitute the main mediating tools between the level of decision-making and work in schools (Mangez, 2008). We focused on two disciplines, French as a language of instruction and mathematics, because of their weight (hours of teaching) and relative importance in the curricula, but also in the evaluations (local, national, international) of pupils’ achievements.
- Administrative documents (service notes, circulars, ministerial instructions) that specify how legal and regulatory texts are to be implemented.
- Supporting or accompanying documents (various guides and pedagogical tools developed by the ministry and its departments) which do not have a prescriptive value but can nevertheless be central in teachers’ work.

Our approach and empirical material differed entirely from the methodology adopted in large-N comparative analyses. For example, Kamens et al. (1996) used reports on the curricular outlines of secondary education programs worldwide found in compendia produced by international and regional conferences. This is the formal curriculum in its most “external” and “superficially institutional” aspect (Forquin, 2008, p. 104), with data processing being limited to identifying the presence or absence of different disciplines and their respective place. On the contrary, the choice of our data aimed to reveal the plurality of forms that changes take in different national spaces, as well as the hybridization processes between different logics and models that underpin national education policies (Maroy, 2009).

To analyze the corpus, we developed an analytical grid based on the Bernstein’s key concepts which allowed us to construct a common theoretical denominator beyond national contexts. The thematic analysis aimed at characterizing the transformations of the official pedagogic discourse along its key dimensions. The analysis of these data was carried out to reveal diachronic variations (2000–2012) in each system and synchronic variations (between the two education systems). For the full presentation of the methodology, see Mathou (2018).

The Pedagogic Renewal in Quebec

The decision to undertake a reform of Quebec’s education system stemmed largely from the Estates General on Education (EGE) held in 1996. In 1995, the left-wing Parti Québécois (PQ) government established the Commission for the Estates General on Education (EGE) to review the entire education situation in Quebec. The overhaul of the curriculum was intended to allow for the democratization of educational success.

The Ministry of education then began elaborating new study programs for primary and secondary education, based on a new approach to curriculum-making. Until the late 1990s, curriculum development was entrusted to groups of experts in their discipline, with little cooperation between these groups (Carpentier, 2010). Moving away from this compartmentalized approach, more than four hundred teachers, principals, counselors, education professionals and academics actively participated in the design of the new programs (Lessard et al., 2007).

The new Quebec Education Program (QEP) for lower secondary education was published in 2003 (MEQ, 2003) and was gradually introduced from 2005, when the reform was renamed “Pedagogic Renewal.” The QEP is organized around the development of competences (subject-related and cross-curricular competences) which reflect the goals to be attained at the end of compulsory education. Disciplines are grouped into five “learning fields” (languages, mathematics, science and technology, social sciences and the arts and social development) that are supposed to highlight the interdependence of these fields of knowledge. In addition, the QEP puts a very strong emphasis on two key elements. The first is the definition of “general education fields” such as health and well-being, media literacy, environmental awareness and consumer rights. The second is the definition of cross-curricular competences (defined as the intellectual, methodological, personal, social, and communication skills) common to all learning fields. Finally, the QEP is structured around 2-year learning cycles, allowing the progression of each individual pupil at his/her own rhythm.

Following the publication of the QEP the Ministry published its policy on the assessment of learning in the spirit of the Pedagogic Renewal. Assessment was defined as an integral part of the learning process, taking into account a whole range of observations collected in different situations. In line with this vision, the Basic School Regulation (Quebec, 2005) loosely defined the form of the communication to be handed in to parents. The pupils’ results could be expressed in the form of letters using scales of competence levels to reflect pupil’s progress and development. Numerical grades and pupils’ comparisons were strongly criticized.

The orientations defined by the Parti Québécois government were not called into question by the arrival in power of the Liberal Party (PLQ) in 2003. However, from 2007 onwards major changes took shape, in a context characterized by a combination of pressures. The discontent of parents—denouncing incomprehensible jargon and unclear assessment methods—converged with strong teacher union pressure. Around that time, the Liberal Party embarked on a major legislative project, the adoption of the law on results-based management in education, which introduced new rules on accountability and performance evaluation in schools (Maroy et al., 2014).

From 2007 to 2010, the Ministry of Education modified the Basic school regulation several times. The number of report cards were defined on an annual basis, rather than cycles, their content was standardized. Marks obtained for each subject and the group average had to be included in the report card, allowing comparison between pupils within a group. Standardized ministerial assessments were also introduced in one key subject (French) at the end of lower secondary school. Each school was required to set performance targets in advance for pupils' results in these assessments. An addendum to the program, called the "Progression of learning," was published in 2010/2011 specifying the contents to be covered each year, for each subject.

Finally, a "reading intervention" toolkit (MELS, 2011) was published by the Ministry to identify "effective" practices, putting at the forefront practices based on explicit, teacher-centered teaching such as direct instruction. Despite these changes, the central curriculum document, the QEP itself, remained unchanged.

The Common Core in France

In France the curriculum reform followed the national debate on the future of education, chaired by C. Thélot, which formalized the idea of a common core in its final report (Thélot, 2004). The ensuing Guidance and Planning Law of April 23, 2005 called for "every pupil to be guaranteed the means needed to acquire a common core consisting of a set of knowledge and skills that it is indispensable to master in order to successfully complete school education, continue training, build one's personal and professional future and play a successful part in society." Based on the recommendations from a newly created body, the high Council for Education (HCE), the Decree of July 11, 2006 (MEN, 2006) defined the seven broad competences (or "pillars") that all pupils must progressively acquire throughout their compulsory schooling (from 6 to 16).

Subject-specific programs were subsequently revised in order to take into account these competences. However, the groups in charge of the subject-specific program revision worked independently, and more than half of the members belonged to the inspection bodies, contradicting the stated supra-disciplinary logic of the Common Core (Clément, 2013).

The new programs were published in 2008. New assessment tools were also elaborated in line with the "competence-based" approach outlined in the Common Core. The personal competence booklet was meant to be used to progressively "validate" the seven competences of the Common Core throughout compulsory education, allowing teachers to take stock of the pupil's achievements (MEN, 2010).

In the following years, numerous documents with no legal value were published. A first set of circulars defined various "education for" (e.g. health, sustainable development, career choice ...) which would then have to be taken into account in the validation of the competences of the Common Core. A second set of circulars introduced various forms of "individual support" provided to pupils during specific slots dedicated to "differentiated" learning. Between 2005 and 2011, three types of measures were introduced: the personalized programs for educational success or PPRE (2005), educational support (2008), and personalized support (2011).

Following the 2012 presidential elections, the newly elected socialist government passed a new Guidance and Planning Law (July 2013), which announced the review of the Common Core and the creation of a new independent body, the High Council for Programs (CSP). From 2013 to 2015, the CSP drafted a charter to be used as a guide for drafting the new Common Core and new study programs. Working groups corresponding to the three cycles of compulsory schooling were set up, each group being interdisciplinary and including representatives of the various actors working in schools, as well as academics specializing in pupil learning and the areas of knowledge included in the curricula.

Unlike the 2006 Common Core, in which five of the seven competences referred to clearly recognizable disciplinary pillars, the five new "domains" of the 2015 Common Core referred to educational areas or domains and led to a much greater blurring of disciplinary boundaries (MEN, 2015b) (Table 2).

Table 2 The Common Core in 2006 and 2015.

<i>Seven competences or pillars of the Common Core of knowledge and competences (2006)</i>	<i>Five areas of the Common Core of knowledge, competences and culture (2015)</i>
(1) Mastering the French language	Area 1: Languages for thinking and communicating
(2) Speaking a modern foreign language	Area 2: Methods and tools for learning
(3) Basic knowledge in mathematics and scientific and technological culture	Area 3: The formation of the person and the citizen
(4) Mastering common information and communication technologies	Area 4: Natural and technical systems
(5) Humanist culture	Area 5: Representations of the world and human activity
(6) Social and civic skills	
(7) Autonomy and initiative.	

This movement toward greater integration of content was, however, accompanied by a cautious approach toward the risks that may arise from a weak classification of knowledge, a potential source of confusion for pupils. Particular attention was also paid to what was expected of the pupil, the Common Core promoting the explicit teaching of school language, lexicon, understanding of instructions, using school books and note-taking (MEN, 2015a).

Multi-year cycles and “progressivity benchmarks” (indicating the notions to be introduced at the beginning or end of the cycle) replaced annual sequencing. Texts published in 2014–2015 also emphasized less hierarchical relationships and, above all, the recognition of the child’s fulfillment and personality development, taking into account the pupil as an adolescent in the process of physical and psychological evolution (MEN, 2015a).

The recontextualization of common trends into two national contexts

Common trends underpin the curricular transformations observed in the two education systems during the period 2000–2015. The two reform movements originating in the late 1990s/early 2000s, both mark a shift toward the weakening of classifications between types of knowledge and toward the relaxation of the framing of pedagogic relations.

In the two corpus analyzed, we identified a rise in cross-curricular themes that weaken disciplinary boundaries, also the boundaries between the inside and outside of school (Whitty et al., 1994). They are indicators of a strong integration of knowledge. This is accompanied by a shift in the foundation of the legitimacy of knowledge, increasingly based on the relevance with regard to life outside of school, leading to varying degrees of de-differentiation (Young, 2006) between forms of knowledge. These changes are closely linked to changes in teacher-pupil relationships and expectations of pupils, such as the increased personalization of pedagogic relations.

However, there are important variations in the recontextualization of these common orientations, particularly among the logics presiding over the weakening of classifications and framing. In Quebec, this weakening is reflected in a strong permeability between school knowledge and experiential knowledge. Everyday life issues become the stakes of knowledge while subject-specific knowledge becomes a resource among others to deal with these issues. For instance, in the section of the QEP dedicated to mathematics, the theme “Environment and consumption” is tackled through pupils’ daily lives and habits (such as calculating taxes, rebates). The example of “learning situation” provided by the program presents mathematics as a resource among others for dealing with a recycling problem, overshadowing scientific conceptualization (Hasni, 2010).

The permeability between types of knowledge is much less strong in France. While greater legitimacy is granted to the contextualization of school knowledge, everyday, common-sense knowledge is treated with great caution. Cross-cutting themes such as environmental education can give rise to disciplinary crossovers, but each discipline treats the theme in a different way. Thus, the so-called “convergence” themes (MEN, 2008) are somewhat absorbed in the disciplinary code (Whitty et al., 1994): they are distributed across a variety of separate subjects, without significantly impacting the boundaries between these subjects. We can therefore observe two contrasting approaches to the introduction of cross-curricular themes related to “life skills”: in Quebec the subjects are supposed to be taught through the themes, while in France the themes are taught through the subject. Although from 2012 onwards, the official pedagogic discourse grants increased legitimacy to the contextualization of school knowledge, it nevertheless remains cautious about the blurring of boundaries.

Similarly, while the framing of the teacher/pupil relationship is relaxed on both sides, the differentiated approach to teaching is much more institutionalized in Quebec. The affective, personal dimension is at the core of pedagogic relations: teachers must take into account pupils’ fields of interests, learning styles and tastes, as well as their family, personal and social contexts. Differentiation, a leitmotiv of the key texts adopted in the context of the Pedagogic Renewal, is supposed to be the ordinary modality of the pedagogic relationship in the classroom. Teachers are expected to put in place the conditions conducive to “the expression of pupils’ concerns, tastes, feelings, emotions” (MEQ, 2003, p. 23), leading to a form of social control which “involves personality at a deeper level” (Forquin, 2008, p. 33).

By contrast, “differentiation” in the French official discourse is essentially limited to taking into account pupils’ previous knowledge, and remains focused on remediation to bring pupils to the necessary prerequisites (Houssaye, 2012). Measures introducing forms of individualization tend to weaken the framing of teacher-pupil relations, but within well-defined spaces and time frames, and they focus first and foremost in the provision of support to complete homework and revise lessons. These are typical interventions aiming at “giving more education” (Power and Whitty, 2008), thus increasing the frequency and intensity of educational activities without weakening the boundaries between home and school and between teachers and pupils (strong classification and strong framing). The prevailing form of “individualization” in France therefore remains centered on school performance and aimed above all at low achieving pupils.

Secondly, the comparative analysis highlights the complex hybridization of pedagogic models at work in French and Quebec curricular policies. In both cases, the transformations that took place over more one decade entailed the layering of multiple texts resulting in mixed pedagogic models (Muller, 1998). In relation to the three models outlined by Bernstein (1990)—an invisible pedagogy, a visible pedagogy with autonomous and market-oriented modalities—the diachronic analysis shows that these three models and modalities are mixed in varying degrees and according to different arrangements, depending on the period under consideration (Table 3).

In Quebec, the central texts of the Pedagogic Renewal (2000–2007) promoted a highly integrated code of knowledge and an invisible pedagogy. However, toward the end of the decade, the official discourse evolved toward a more hybrid form in which invisible and dependent visible pedagogies coexist. The official discourse has preserved in part the rhetoric of pupils’ personal

Table 3 Main pedagogic orientations underlying the official discourse according to the periods considered.

	2000	2005	2007	2012
Québec	Period 1 2000–2007 Invisible pedagogy strongly asserted through very homogeneous texts		Period 2007–2015 Incremental shift toward dependent, market-oriented visible pedagogy through the layering of new texts to the QEP	
France		Period 2005–2012 Centrality of autonomous visible pedagogy despite the marginal opening toward invisible pedagogic practices		Period 2012–2015 Invisible pedagogy takes center stage in official discourse although adopted in a “cautious” mode

development, but has moved toward a more interindividual competition model, centered on achieving performance measured by standardized tests. Meanwhile, knowledge pertaining to everyday life has retained a central place, and characteristic features of invisible pedagogy have been maintained in sub-segments of the system. Thus, teachers may continue to use competence scales and report cards centered on learning outcomes with pupils with learning difficulties, a common situation according to Bernstein (1990), where the practices that tend most toward invisible pedagogy are reserved for certain sub-groups (disadvantaged socio-economic categories, pupils defined by a “deficit”) or certain “non-verbal” activities (Bernstein, 1975, p. 110). Most notably, the QEP has remained untouched.

In France, the period 2005–2012 was characterized by the marginal shift toward the invisible model. The integration of knowledge was contradicted by the decoupling between the logic of the Common Core and that of subject-specific programs. Moreover, invisible pedagogy was reserved to certain categories of pupils (low achievers), and limited to specific times “outside” the regular classroom. Invisible pedagogy was incorporated in a circumscribed way, within a system where the autonomous visible model predominated. Although the changes adopted in 2012 marked a strong movement toward the invisible model, the official discourse remained cautious with regards to the weakening of boundaries, as a form of protection against the potentially differentiating effects that penalize pupils from disadvantaged backgrounds. This cautious and critical approach to the blurring of boundaries and de-differentiation of knowledge draws on the work of researchers from the ESCOL research group (see e.g. Bautier and Rayou, 2013), some of whom participated in the working groups responsible for developing the programs. Thus, the 2012–2015 discourse can be seen as an attempt to combine “therapeutic” orientations that value the expression and development of the “self,” more permeability between school and non-school knowledge, more inter-disciplinarity, with the preservation of explicit rules of recognition and realization.

Finally, three main interpretative hypotheses can be put forward to account for the variations observed in the recontextualization of common curricular trends. Firstly, historical and cultural factors partly account for the variations in the intensity and modality of the transformations observed influence the degree of legitimacy of invisible pedagogy in each education system. In Quebec, these orientations have been institutionalized for a longer period of time. The reform resulting from the EGE was in part a continuation of the orientations adopted since the 1960s, and initiated by the Parent Report. In France, the recognition of differences (through the individualization of relations with pupils) is still problematic in a system historically built to promote a “universalism” that requires ensuring a uniform distribution of knowledge to all, throughout the country, and through examinations that sanction fair competition (Duru-Bellat and van Zanten, 2012). Thus individualism is valued more than group work, teaching tends to be decontextualized from the daily experience of pupils, the affective dimension is clearly separated from the cognitive dimension (Osborn et al., 2003).

Secondly, factors relating to the processes presiding the production of official texts play a role in the modalities of curricular change. In both cases, the very marked movements toward the weakening of classification have been accompanied by attempts to break with the traditional modes of curriculum production, driven by disciplinary logics (subject-specific professional associations and general inspectorates). New curriculum making arrangements were put in place in Quebec (early 2000s) and France (2012), characterized by the role of a single public body that coordinates the drafting process of all the curriculum components (study programs, teaching guidelines, assessment tools); interdisciplinary working groups organized by cycle (not by year of study) to ensure vertical and horizontal alignment; the involvement of actors at the forefront of the implementation of the curriculum (teachers, pedagogical advisors) but also academic experts specialized in learning processes. Interestingly, our research, although it did not explore this aspect in detail, points toward processes of policy borrowing. For instance, before embarking on the new reform in 2012, the newly created CSP carried out an international comparison exercise to inform its reflection on the new programs and the new common core, based on the experience of a selection of countries including Quebec.

Thirdly, more contingent factors, linked to the strategies of actors and to political and electoral contexts, contributed to processes of juxtaposition and hybridization. Thus, the 2007 turning point in Quebec can be explained by a particular electoral and parliamentary context, combined with pressure from unions and parents. The introduction of results-based management, a major

legislative project in 2008, was thus explicitly linked to the “realignment” of the curriculum (standardized report cards, changes to the programs to prioritize core subjects or “basic” knowledge) in order to increase school performance. As Bernstein and others (see Broadfoot and Pollard, 2006; Halpin et al., 2004; Sadvnik, 2006) have argued, invisible pedagogies are less amenable to public scrutiny and accountability, as their outputs are more difficult to evaluate according to objective and standardized frameworks. In such a model, the differences between learners’ achievements cannot be used as a basis for comparison and stratification, as these differences are supposed to reveal the uniqueness of each individual. Instead, accountability is facilitated by the standardization of the sequencing and pace of learning, aligned with a uniform temporality, and the measurement of performance objectified by numerical assessment. Grades and exams make it possible to construct rankings in order to compare pupils, schools, etc. (Muller, 1998).

In France, the extreme politicization of education policies resulted in the proliferation of texts adopted by successive ministers (Meuret, 2013). The succession of ministers from the right-wing majority in power (F. Fillon, G. De Robien, X. Darcos, L. Chatel) between the adoption of the 2005 Law and the change of government in 2012 produced a mix of heterogeneous texts lacking a clear direction in the changes undertaken. However, unlike in Quebec (where the change in the party in power did not lead to scrapping entirely the Pedagogic Renewal), the change of political majority in 2012 re-launched the whole process of curriculum development leading to a new Common Core and new programs.

Conclusion

The objective of this chapter was to understand how common curricular trends are translated into national curricula. The study presented analyzed contemporary curricular transformations in two national contexts, France and Canada (Quebec), and over a 15 year period.

Our findings illustrate two distinct national declinations of the “global curriculum.” In both contexts, the official pedagogic discourse in the 2000s led to a movement toward the weakening of classification and framing and the invisibilization of evaluation criteria. But far from resulting in a standardization of national curricula, these common orientations were recontextualized according to logics and dynamics specific to each educational system. Long-term historical factors (including curricular traditions), existing institutional arrangements (such as curriculum-making processes), and factors related to the political context constituted powerful filters that maintained strong national specificities in terms of curricular evolutions and policy trajectories.

The analysis also revealed the complexity of contemporary curricular developments, producing composite, mixed pedagogic models, far from clear-cut dichotomies. The composite models correspond neither to what critical perspectives denounce as a device aimed at satisfying market needs nor to what critics of competence-based reforms describe as pedagogism or child-centrism. The analysis thus makes it possible to highlight the complex, ever-changing dynamics underlying contemporary curricular policies.

In France notably, following the presidential elections of 2017, the official pedagogic discourse seems to have taken a new direction, once again. The changes adopted by the new Minister of Education seem to undermine the affirmation of the invisible model. This could be the beginning of a new pendulum swing, shifting back to the autonomous visible pedagogy or toward a mixed regime embedding more instrumental, dependent modalities.

Drawing on the sociology of the curriculum means indeed apprehending the curriculum as the product of an ongoing process of elaboration and institutionalization in which skills and resources are mobilized, interests are confronted, value conflicts and power issues are at stake (Forquin, 2008).

Finally, these results invite us to examine with caution the idea of a necessarily convergent movement between so-called “competency-based” reforms, New Public Management (NPM) and neo-liberal policies. While the rhetoric around “competency-based” seems to converge, there are fundamental differences between the pedagogic discourses and practices in which the notion is recontextualized. An in-depth analysis of the logics that preside over the weakening or strengthening of classifications, framing and evaluation is therefore essential to avoid reducing pedagogic discourse to a relay of ideological messages that are external to it (Bernstein, 1990). This in turn argues in favor more attention dedicated to the curriculum in an era of globalized education policies, and for the further development of comparative curriculum research attentive to on-going recontextualization processes between and within transnational and national policies that influence national school systems (Wahlström et al., 2018).

References

- Alexander, R., 2006. Dichotomous pedagogies and the promise of cross-cultural comparison. In: Lauder, H., Brown, P., Dillabough, J.A., Halsey, A.H. (Eds.), *Education, Globalisation and Social Change*. Oxford University Press, Oxford, pp. 722–733.
- Anderson-Levitt, K., Gardinier, M.P., 2021. Introduction contextualising global flows of competency-based education: polysemy, hybridity and silences. *Comp. Educ.* 57.
- Anderson-Levitt, K., 2003. A world culture of schooling. In: Anderson-Levitt, K. (Ed.), *Local Meanings, Global Schooling: Anthropology and World Culture Theory*. Palgrave Macmillan, New York, pp. 1–26.
- Bautier, É., Rayou, P., 2013. *Les inégalités d'apprentissage. Programmes, Pratiques et malentendus scolaires*, second ed. PUF, Paris.
- Bernstein, B., 1971. On the classification and framing of educational knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control*. Collier-Macmillan, London, pp. 47–69.
- Bernstein, B., 1975. Class and pedagogies: visible and invisible. *Educ. Stud.* 1, 23–41.
- Bernstein, B., 1990. *Class, Codes and Control Volume IV. The Structuring of Pedagogic Discourse*, vol. IV. Routledge, London.

- Broadfoot, P., Pollard, A., 2006. The changing discourse of assessment policy: the case of English primary education. In: Lauder, H., Brown, P., Dillabough, J.A., Halsey, A.H. (Eds.), *Education, Globalisation and Social Change*. Oxford University Press, Oxford, pp. 760–765.
- Carpentier, A., 2010. Étude de la mise en œuvre de la réforme du curriculum québécois du primaire, de 1997 à 2003. These de doctorat. Université de Montréal. Retrieved from: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/4125/Carpentier_Anylene_2010_these.pdf?sequence=4.
- Clément, P., 2013. Réformer les programmes pour changer l'école ? Une sociologie historique du champ du pouvoir scolaire. Thèse de doctorat. Université de Picardie Jules Verne. Retrieved from: <http://www.theses.fr/s27205>.
- Duru-Bellat, M., Van Zanten, A., 2012. *Sociologie de l'école*, fourth ed. Armand Colin, Paris.
- Forquin, J.-C., 2008. *Sociologie du curriculum*. Presses Universitaires de Rennes, Rennes.
- Halpin, D., Dickson, M., Power, S., Whitty, G., Gewirtz, S., 2004. Curriculum innovation within an evaluative state: issues of risk and regulation. *Curric. J.* 15 (3), 197–206.
- Hasni, A., Lebeaume, J., 2010. Introduction. L'enseignement scientifique et technologique: nouvelles orientations curriculaires, nouveaux défis. In: Hasni, A., Lebeaume, J. (Eds.), *Enjeux contemporains de l'éducation scientifique et technologique*. Presses de l'université d'Ottawa, Ottawa, pp. 1–16.
- Hasni, A., 2010. L'éducation à l'environnement et l'interdisciplinarité: de la contextualisation des savoirs à la scolarisation du contexte? In: Hasni, A., Lebeaume, J. (Eds.), *Enjeux contemporains de l'éducation scientifique et technologique*. Presses de l'Université d'Ottawa, Ottawa, pp. 179–222.
- Holmes, B., McLean, M., 1989. *The Curriculum: A Comparative Perspective*. Unwin Hyman, London.
- Houssaye, J., 2012. La gestion pédagogique des différences entre les élèves: variations françaises. *Carrefours de l'éducation* 34 (2), 227–245.
- Kamens, D.H., Meyer, J.W., Benavot, A., 1996. Worldwide patterns in academic secondary education curricula. *Comp. Educ. Rev.* 40 (2), 116–138.
- Kazamias, A., 2009. On educational knowledge—a neglected theme in comparative education. In: Cowen, R., Kazamias, A. (Eds.), *International Handbook of Comparative Education*. Springer Netherlands, pp. 803–812.
- Legendre, M.F., 2008. La notion de compétence au cœur des réformes curriculaires: effet de mode ou moteur de changements en profondeur? In: Audigier, F., Tutiaux-Guillon, N. (Eds.), *Compétences et contenus: les curriculums en questions*. De Boeck Supérieur, Bruxelles, pp. 27–50.
- Lenoir, Y., 2005. Le « rapport Parent », point de départ de l'ancrage de l'école québécoise dans la logique anglophone nord-américaine. *Can. J. Educ./Revue canadienne de l'éducation* 28 (4), 638–668.
- Lessard, C., Henrripin, M., Larochelle, M., Cournoyer, É., Carpentier, A., 2007. Inventaire des politiques d'éducation au Québec portant sur l'enseignement primaire et secondaire de 1990 à 2006. Analyse et impacts des politiques scolaires canadiennes sur le personnel scolaire. Chaire de recherche du Canada sur les métiers de l'éducation, Montréal.
- Lingard, B., McGregor, G., 2014. Two contrasting Australian Curriculum responses to globalisation: what students should learn or become. *Curric. J.* 25 (1), 90–110.
- Lingard, B., 2010. Towards a sociology of pedagogies. In: Apple, M.W., Ball, S.J., Gandin, L.A. (Eds.), *The Routledge International Handbook of the Sociology of Education*. Routledge, New York, pp. 167–178.
- Mangez, É., 2008. *Réformer les contenus d'enseignement: une sociologie du curriculum*. Presses Universitaires de France, Paris.
- Maroy, C., Mathou, C., Vaillancourt, S., Voisin, A., 2014. La construction de la politique de "Gestion axée sur les résultats" au Québec: récits d'action publique et trajectoire de la politique. *Éduc. Comp.* 12, 45–69.
- Maroy, C., 2009. Convergences and hybridization of educational policies around "post-bureaucratic" models of regulation. *Compare* 39 (1), 71–84.
- Mathou, C., 2018. Transformations et recontextualisations du discours pédagogique : une comparaison des politiques curriculaires en France et au Québec (2000–2015). Thèse de doctorat. Université de Montréal. Retrieved from: <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/21213>.
- Ministère de l'éducation, du loisir et du sport (MELS), 2011. *Référentiel d'intervention en lecture pour les élèves de 10 à 15 ans*. Gouvernement du Québec, Québec.
- Meuret, D., 2013. Pour une école qui aime le monde: les leçons d'une comparaison France-Québec (1960–2012). Presses Universitaires de Rennes, Rennes.
- Meyer, J.W., Kamens, D.H., Benavot, A., 1992. *School Knowledge for the Masses: World Models and National Primary Curricular Categories in the Twentieth Century*. Falmer Press, Washington, DC.
- Meyer, J.W., 2007. World models, national curricula, and the centrality of the individual. In: Benavot, A., Braslavsky, C. (Eds.), *School Knowledge in Comparative and Historical Perspective*. Springer, Dordrecht, pp. 259–271.
- Ministère de l'éducation du Québec (MEQ), 2003. *Programme de formation de l'école québécoise—Enseignement secondaire, premier cycle*. Gouvernement du Québec, Québec.
- Ministère de l'Éducation nationale (MEN), 2006. *Le socle commun de connaissances et de compétences*. Décret du 11 juillet 2006. Tout ce qu'il est indispensable de maîtriser à la fin de la scolarité obligatoire. Direction générale de l'enseignement scolaire, Paris.
- Ministère de l'Éducation nationale (MEN), 2008. *Programmes du collège. Programmes de l'enseignement de mathématiques*. Arrêté du 8-7-2008, J.O du 5-8-2008. Bulletin officiel spécial du 28 août 2008.
- Ministère de l'Éducation nationale (MEN), 2010. Circulaire n° 2010-087 du 18-6-2010. Mise en œuvre du livret personnel de compétences. Bulletin officiel no 27 du 8 juillet 2010.
- Ministère de l'Éducation nationale (MEN), 2015a. Arrêté du 9-11-2015. J.O. du 24-11-2015. Programme d'enseignement du cycle des approfondissements (cycle 4). Bulletin officiel spécial n°11 du 26 novembre 2015.
- Ministère de l'Éducation nationale (MEN), 2015b. Décret n° 2015-372 du 31-3-2015. J.O. du 2-4-2015. Socle commun de connaissances, de compétences et de culture. Bulletin officiel n° 17 du 23 avril 2015.
- Muller, J., 1998. The well-tempered learner: self-regulation, pedagogical models and teacher education policy. *Comp. Educ.* 34 (2), 177–193.
- Mundy, K., Green, A., Lingard, B., Verger, A., 2016. Introduction: the globalization of education policy: key approaches and debates. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley-Blackwell, Chichester, UK, pp. 1–20.
- Osborn, M., Broadfoot, P., McNess, E., Planel, C., Ravn, B., Triggs, P., 2003. *A World of Difference? Comparing Learners Across Europe*. Open University Press, Maidenhead.
- Power, S., Whitty, G., 2008. A Bernsteinian Analysis of Compensatory Education. Paper Presented at the 5th Basil Bernstein Symposium, Cardiff University, July 9th–12th.
- Priestley, M., Biesta, G., 2013. *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. Bloomsbury Publishing, London.
- Québec, 2005. *Règlement modifiant le Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire*. Décret 488-2005, 25 mai 2005. Éditeur officiel du Québec, Québec.
- Robertson, S.L., 2005. Re-imagining and rescripting the future of education: global knowledge economy discourses and the challenge to education systems. *Comp. Educ.* 41 (2), 151–170.
- Rocher, G., 2004. Un bilan du Rapport Parent: vers la démocratisation. *Bull. Hist. Polit.* 12 (2), 117–128.
- Rochex, J.Y., Francia, G., Greger, D., Le Breton, J., 2011. Les adaptations et transformations curriculaires et pédagogiques et leurs agents dans les politiques d'éducation prioritaire. In: Demeuse, M., Frandji, D., Greger, D., Rochex, J.Y. (Eds.), *Les politiques d'éducation prioritaire en Europe. Tome II: Quel devenir pour l'égalité scolaire?* ENS Éditions, Lyon.
- Ropé, F., Tanguy, L. (Eds.), 1994. *Savoirs et Compétences. De l'usage de ces notions dans l'école et l'entreprise*. L'Harmattan, Paris.
- Rosenmund, M., 2007. The current discourse on curriculum change: a comparative analysis of national reports on education. In: Benavot, A., Braslavsky, C. (Eds.), *School Knowledge in Comparative and Historical Perspective*. Springer, Dordrecht, pp. 173–194.
- Sadovnik, A., 2006. Toward a sociology of educational change: an application of Bernstein to the US "No Child Left Behind" Act. In: Moore, R., Beck, J., Arnot, M., Daniels, H. (Eds.), *Knowledge, Power and Educational Reform. Applying the Sociology of Basil Bernstein*. Routledge, New York.
- Thélot, C., 2004. *Pour la réussite de tous les élèves: rapport de la Commission du débat national sur l'avenir de l'Ecole*. La Documentation française, Paris.
- van Zanten, A., 2009. *Choisir son école. Stratégies familiales et médiations locales*. Presses Universitaires de France, Paris.
- Voogt, J., Roblin, N.P., 2012. A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J. Curric. Stud.* 44 (3), 299–321.

- Wahlström, N., Alvunger, D., Wermke, W., 2018. Living in an era of comparisons: comparative research on policy, curriculum and teaching. *J. Curric. Stud.* 50 (5), 587–594. <https://doi.org/10.1080/00220272.2018.1502814>.
- Whitty, G., Rowe, G., Aggleton, P., 1994. Discourse in cross-curricular contexts: limits to empowerment. *Int. Stud. Sociol. Educ.* 4 (1), 25–42.
- Whitty, G., 2010. Revisiting School Knowledge: some sociological perspectives on new school curricula. *Eur. J. Educ.* 45 (1), 28–45.
- Yates, L., Grumet, M., 2011. *World Yearbook of Education 2011: Curriculum in Today's World: Configuring Knowledge, Identities, Work and Politics*. Routledge, New York.
- Yates, L., Young, M., 2010. Globalisation, knowledge and the curriculum. *Eur. J. Educ.* 45 (1), 4–10.
- Young, M., 2006. Curriculum studies and the problem of knowledge; updating the enlightenment. In: Lauder, H., Brown, P., Dillabough, J.A., Halsey, A.H. (Eds.), *Education, Globalization and Social Change*. Oxford University Press, Oxford, pp. 734–741.
- Young, M., 2011. Curriculum policies for a knowledge society? In: Yates, L., Grumet, M. (Eds.), *World Yearbook of Education 2011. Curriculum in Today's World: Configuring Knowledge, Identities, Work and Politics*. Routledge, New York, pp. 125–138.

Curriculum governance: the dilemma of the centralized curriculum

Axel Rivas and Belén Sanchez, Universidad de San Andrés, Argentina

© 2023 Elsevier Ltd. All rights reserved.

Introduction	19
A historical approach to centralized curriculum	20
Theoretical distinctions on curriculum governance	21
Does centralization work?	22
Making sense of centralization	23
Conclusions	25
References	26

Introduction

The presentation of polarized debates impoverishes discussions and decision making in education. Simplification is attractive; it prevents us from thinking, debating, and listening. In the field of curriculum this happens with different issues that have been pushed into a scheme of opposites. One of these is centralized curriculum governance vs. teacher autonomy and diversity. In this article we explore this debate and seek to escape its simplification through a key question: Should curriculum governance be centralized to achieve better learning outcomes?

In recent decades we have witnessed the constant emergence of the temptation of curriculum centralization. There are two major parallel histories of this governance shift. The first is the argument that there should be greater curricular centralization in educational systems with lower educational quality. The second is the story of the United States and other Anglo-Saxon countries with a tradition of strong curricular decentralization, where centralization is proposed as a decisive reform path.

One of the most widely read and cited documents on international education policy is the “McKinsey Report” (Mourshed et al., 2010). There, the argument for curriculum centralization in the countries with the lowest educational results is presented at its full potential.

The McKinsey Report invites policymakers around the world to follow a linear and straightforward instructions manual: they should locate their education system in the league table of learning performance and act depending on whether they are among the “poor,” “regular,” “good” or “very good” performing systems. For systems with poor performance, the report, based on international cases of systemic improvement, recommends curricular centralization policies such as scripted teaching materials, coaching on curriculum and instructional time, and provision of textbooks. Instead, for countries with higher results, the report recommends “decentralizing pedagogical rights to schools and teachers” (Mourshed et al., 2010, p. 28).

The McKinsey report has been criticized for its lack of comparative rigor. In generalizing educational policy formulas, it ignores the contexts and the enormous cultural variations of educational systems (Alexander, 2010; Morris, 2015). However, its argument is shared by numerous authors and approaches linked to the need to centralize curricular governance in education systems with a low level of teaching skills.

The argument for curriculum centralization in fragile contexts has also led to top-down solutions. In some cases, it has been proposed that the distribution of free textbooks is a way to establish a practical curricular floor throughout the system. In other cases, it has been proposed to advance more rapidly in curricular centralization through standardized assessments. In the face of low effectiveness of curricular standards, high-stakes standardized assessments can boost incentives to teach what is highlighted by the tests. This is one way of rapid and systemic reform.

Another way of curriculum centralization, which can complement the previous ones, is the development of scripted lessons. An extreme case of this variant is Bridge International Academies (BIA), the world’s largest network of private schools, which operates in several African countries. BIA schools work based on a pre-programmed curriculum delivered to teachers through electronic devices. These contain completely scripted lesson guides that offer word-by-word and action-by-action steps for every lesson teachers are meant to teach. This model, which BIA describes as an “academy-in-a-box,” operates under the assumption that scripted guides are necessary in developing countries where teachers, allegedly and according to BIA, cannot read nor solve primary-level mathematical problems.

In other countries with historically limited state capacities, the private market also influences curricular governance. For example, in Brazil there are “systems of teaching” (*sistemas do ensino*) that are responsible for providing schools with a comprehensive set of assistance: textbooks, teacher training, digital platforms, management systems and professional advice. The schools that hire these systems seek to simplify their task and give more prescriptive lessons and teaching routes to their teachers (Chizzotti and Ponce, 2012).

In these models, curricular centralization seems to be the transitory solution until the capacity of teachers to work with greater autonomy and creativity is achieved. The centralization shift argues that developing countries need a common floor in the face of

fragile contexts. Even in the case of BIA, the curriculum is designed outside the countries, through a private company that seems to follow a logic of efficiency in curriculum governance that the State cannot achieve.

Another version of curriculum centralization comes from the specific context of the United States. Some authors have analyzed the case of the United States as a world exception in its extremely decentralized model of curricular government (Judge et al., 1994). An interpretative approach on this exceptionality says that the US is epitomized by two traditions—"democratic localism" and "individualism"—that oppose state centralization (Timar et al., 1998).

In the 1990s, this diagnosis became a reform laboratory for curricular centralization. Several states began processes of designing a centralized curriculum. The No Child Left Behind (NCLB) Act reform went further, seeking to consolidate curricular foundations for the entire country (Deng, 2010).

In some states the ideas of curricular centralization reached the extreme of tightly regulating teaching work, through the introduction of third-party (mostly for-profit) scripted curriculum programs, understood as ways to homogenize teaching in ways matching "what works," allegedly defined by programs with scientific evidence on their effectiveness (Maheady et al., 2016). Florida offers an example. The Miami-Dade Public School District, one of the largest in the United States, introduced prescribed programs which align perfectly to the requirements of yearly standardized testing. It also issued pacing guides that set the standards needed to be covered during each grade level, with the purpose of ensuring curricular continuity across the district. The state department of education has promoted scripted programs and pacing guides, issuing explicit instructions on how teachers should use the materials they receive, recommending actions such as "verbatim reading" of bold phrases during class.

These two parallel narratives in favor of curriculum centralization converge in many arguments and differ in others. Is the centralization of curriculum governance the appropriate way to improve educational systems? Does centralization improve the equity of the system, providing a common core for all, regardless of the social context? Is it a way to solve the problem of poor teaching quality in developing countries? Is it the way to improve overly decentralized systems?

These questions open the dilemma of curricular centralization. In this article we analyze the historical genesis of this debate in the origins of the educational system. Then we argue that it is necessary to escape the simplification of polarized arguments. Curriculum governance is like a set of Russian dolls or Chinese boxes. What appears to be a simple decision involves other debates. Curricular governance encompasses all spheres of organization and meaning of an educational system. We propose a map to integrate these debates and look for ways to overcome the dilemma.

A historical approach to centralized curriculum

The historical emergence of the concept "curriculum" is associated with the moment of initial development of educational systems in the stage of the Wars of Religion. David Hamilton shows how the concept was born at the intersection between Ramus' educational doctrines and Calvinist religious doctrines in Northern Europe. The term "curriculum" substituted earlier Latin terms such as *ratio* and *disciplina*. The adoption of "curriculum" by Calvinists is closer to the connotations of sequence and linearity and, above all, the Ramist conceptions of order and method (Hamilton, 1990, p. 28).

During the 17th and 18th centuries the first school systems were born close to the religious campaigns. In this context the first models of centralized curriculum emerged. The Ratio Studiorum was the curricular text of the Jesuits, which played a key role in France and other European Catholic countries (Durkheim, 1977). The Ratio was an extremely detailed curriculum, which defined exactly what should be taught at each moment of the school progression. The Ratio allowed different schools to teach the same thing at the same time.

This model of government cannot be dissociated from the organization of the modern school. Together with the Ratio, the Jesuits organized the classrooms with a gradual model of courses organized by age. With the Jesuits individual exercises grew under the control of the teacher, along with grades and exams. The graduated and simultaneous system became an organized machine that allows massive teaching and certification processes.

The pedagogue most recognized for his contribution to the creation of modern educational systems is Jon Amos Comenius. The obsession of the Protestant pastor Comenius was to design a system that could teach "everything to everyone." In his *Didactica Magna*, he raised the idea of a centralized model of government based on the homogeneity, gradualness and sequence of the curriculum. Even Comenius contributed to the creation of what can be considered the first illustrated textbook, the *Orbis Pictus*, which was very popular in 17th century Europe (Comenius, 2019).

As Hamilton indicates, "17th-century theories about schooling contributed to 17th-century theories of political absolutism. The top-down efficiency of the state machine would be assured via the bottom-up inculcation, by means of schooling, social discipline and intellectual deference" (Hamilton, 1989, p. 33).

The origin of the curriculum must be found in this religious-based context of educational systems. The closest concept to the curriculum is that of the Holy Scriptures, the texts that must be reproduced to install beliefs shared by an entire population. It is not by chance that at the birth of schools the Bible was the central learning text, which could be reproduced and distributed on a large scale thanks to the invention of the printing press in the 16th century (Hunter, 1994).

This introductory look at the history of the curriculum shows us the importance of a comprehensive theory. The role of the curriculum cannot be understood without analyzing the functioning of educational systems. In most countries, the consolidation of a centralized curricular governance model was an inherent part of the advancement of state education systems. The French model is perhaps the most recognized in this regard in the 19th century. With a centralizing top-down perspective, the French tradition

installed a homogeneous curricular management model that governed the educational system from the state administration. An example of this systemic power of simultaneous control indicates that in 1896 the French Minister of Public Instruction, taking out his pocket watch, affirmed that at that time, all the 5th-grade students in France were reading the sixth book of *The Aeneid* (taken from Ozouf, 1970).

The roots of educational systems are not homogeneous. In Germany, the concept of *Didaktik* has its own genealogy that allocates a space of autonomy to teachers at the local level. The process of curricular transformation from content (*Inhalte* in German) to teaching (*Gehalte*) was essentially left to the teachers themselves. *Didaktik* is the academic knowledge that includes the specific instructional design as well as the process of performance assessment and creates a space of teaching authority that is distinctive from in the Anglo-Saxon tradition focused on the curriculum (Hopman, 2015).

Theoretical distinctions on curriculum governance

The debate over the centralization of curricular government has many faces. A classic text in the Anglo-Saxon tradition defended the importance of a centralized curriculum. Smith and O'Day (1990) argue for a systemic reform that entails a proactive role by the central units of educational governance (particularly the states), based on the diagnosis that a fragmented policy system with multiple, often incoherent and short-term goals have led to mediocrity in resources and conservatism in instructional practice in the United States. Under the premise that states can overcome this fragmentation through coordinating long-range instructional guidance instruments and that they are in the best position to promote equity, the authors propose the creation of state-led coherent systems of instructional guidance that combine and coordinate three key functions of curriculum governance: curriculum frameworks, pre- and in-service teacher training, and assessment. They suggest this should be accompanied by a revision of the responsibilities attached to each level of governance, and combined with professional discretion by schools, that should be left to decide on staffing, in-service strategies, pedagogies, with support from district and state agencies (Smith and O'Day, 1990).

Faced with the arguments of curricular centralization, Clune (1993) raised a position contrary to that of Smith & O'Day. He argues that a standard curriculum could not address the diversity of needs, that high-stakes testing could lead to the phenomenon of teaching to the test, and that the fragmented, multi-layered governance structure of the United States is not compatible with a system such as that proposed by Smith and O'Day (1990). He claims that it is possible to achieve curricular coherence without centralization, and proposes a system based on decentralized guidance, which he considers more feasible and powerful. According to Clune, instead of defining a standard curriculum, governments should encourage schools to undergo curriculum upgrading processes by joining approved curriculum development networks that provide guidance to generate school-level coherence in terms of content, pedagogies, teacher training and instructional materials. The role of the state and other centralized structures should be that of providing incentives for schools to upgrade their curriculum, developing indicators to track school progress through inspections, and remove obstacles that could come from regulation.

This first presentation of the centralization-decentralization debate of the curricular government is based on its systemic advantages and disadvantages. But this is only one side of the debate. If we open the black box of curricular governance we find another box that takes us to the place that the circulation of knowledge occupies in the educational system. In this sense, Basil Bernstein defines two types of pedagogical discourse: horizontal and vertical. This distinction is essential to open the dilemma of curricular centralization.

The vertical discourse has strong distributive rules regulating access, regulating transmission and evaluation. The vertical discourse is strongly associated with the scientific architecture of knowledge. There are certain closed rules that define what is valid and then distribute it with clear boundaries of organization, rhythm, and validation. The rules are specialized and hierarchical: "vertical discourse consists not of culturally specialized segments, but of specialized symbolic structures of explicit knowledge" (Bernstein, 1999, p. 161).

Instead, "a horizontal discourse entails a set of strategies which are local, segmentally organized, context specific and dependent, for maximizing encounters with persons and habitats" (Bernstein, 1999, p. 160). Horizontal discourse opens the doors of particular views of cultural realities.

The shift from vertical to horizontal curriculum has evident sociological and political consequences. As Bernstein mentions, horizontal discourse could be a crucial resource for "pedagogic populism in the name of empowering or unsilencing voices to combat the elitism and alleged authoritarianism of vertical discourse" (Bernstein, 1999, p. 169). This process marks a shift from equality of vertical reproduction of the curriculum to the recognition of diversity in horizontal discourse.

The debate over the type of organization of knowledge also opens the box of the disciplines. Not everything that is taught in school has the same properties. Within the vertical discourse of academic knowledge production, disciplines belonging to the physical and the social sciences have different knowledge structures. A hierarchical structure is most commonly present in the physical sciences, where general propositions integrate knowledge at lower levels; and a horizontal one is mostly present in social sciences, where a set of specialized languages coexist and have developed segmentally (Bernstein, 2000 cited in Pountney and McPhail, 2017). Some studies have even shown that the structure of the curriculum varies widely by teaching discipline (Schmidt et al., 2001).

This axis of the curricular debate takes on new meanings in the light of recent curricular reforms, promoting a competency-based curriculum aimed at teaching skills for the 21st century (Center for Curriculum Redesign, 2015). Is it possible to integrate the

teaching of skills or competences in contexts of vertical pedagogical discourses? Isn't it necessary to open experiences in the use of knowledge in order to give them local meaning and distribute skills not only anchored in encyclopedic knowledge?

David Perkins indicates that skills require knowledge, but also feelings and motives. Skills are inclinations formed to act in the world, or "big know-how" abilities that connect thinking with understanding and contexts (Perkins, 2014). To teach them is vital to rethink traditional learning organizations and create networks of curriculum design and horizontal questioning.

The organization of the pedagogical discourse opens another black box: the role of teachers in the educational system. In vertical curriculum models, the role of teachers may seem passive. This was exacerbated by traditional centralized views for the organization of educational systems. In the models of curriculum by objectives (Bloom, 1956) there is a taxonomic definition of the path that teaching must take through the world of knowledge: the work of teachers cannot be diverted. This is the tradition of the curriculum as a model to be applied.

The classic Tyler (1969) and Taba (1962) developments present this vision of the systematic, goal-based curriculum. In these models, evaluation plays a central role in keeping the system together and ensuring that objectives are met. According to Taba, the curriculum is, in essence, "a blueprint for learning." This vision of the curriculum as a model has close links with the vertical pedagogical discourse.

In contrast, the horizontal curriculum favors the emergence of diverse and situated educational experiences. This is a conception of the curriculum as a process (Kelly, 2009), which integrates teachers as part of the design of pedagogical paths. Curriculum cannot be isolated from instruction. The pioneering works of Schwab (1959) challenged the passive role of teachers in vertical models of the curriculum. Instead, Schwab proposes to discuss the difficult or "impossible" role of the teacher who must leave the theoretical space to enter the uncertain, problematic, changing, and practical field of real teaching.

Other authors will expand this problematizing vision of the curriculum, where the teacher is a central actor. In 1975 Lawrence Stenhouse defined curriculum as "an attempt to communicate the essential principles and features of an educational proposal in such a form that it is open to critical scrutiny and capable of effective translation into practice" (1975). This definition has a clear correlate in the educational theory of Dewey (1997), who assigned a central place to teachers as creators of unique learning experiences, rather than serial repeaters of content external to them and to the lives of students.

This is how the role of teachers leads us to the next hidden box within the dilemma of curricular centralization: the learning experience of students. The great paradox of curricular centralization is that the greater the verticality of the pedagogical discourse, the greater the scientific rigor is achieved at the cost of losing local significance. Students tend to experience knowledge as something distant and alien. It is a model that favors homogeneous and simultaneous systemic regulation at the expense of the intrinsic motivation of students. In contrast, horizontal pedagogical discourse models tend to take more advantage of contexts and favor individual and group experiences of localized learning (Cox, 2018).

These boxes of the curriculum centralization dilemma show the complexities at stake. Curriculum governance is linked to every aspect of education systems'. It relates to models of curriculum, visions on what is relevant to learn, sociological and cultural implications of the construction and distribution of the pedagogical discourse of knowledge, the role and agency of teachers and the motivation and contexts of students.

Does centralization work?

The boxes opened in the previous section have multiple philosophical, political and pedagogical dimensions. The theoretical conceptions of curricular government are linked to the entire functioning of educational systems. But these questions also have an empirical dimension: does curriculum centralization really work?

The aforementioned McKinsey report (Mourshed et al., 2010), based on PISA data, is a key pillar of the literature in favor of curriculum centralization in developing countries. But there are many other empirical endeavors in this same direction, based on data from international assessments and several other comparative research methods. In the study "Searching for the Golden Model of Education," Bodovski et al. (2017) analyze which characteristics of education systems have influences on the educational results of countries in mathematics. The authors indicate that "a higher level of standardization of educational systems was associated with higher average math achievement" (Bodovski et al., 2017, p. 722). The study found that having a standardized textbook for math was particularly associated with higher average math achievement in a country.

An example of the studies systematized by the authors shows that having mandated curricular guidelines provides a sufficiently balanced curriculum to every student irrespective of the school they attend (English and Steffy, 2001); it also ensures that no inappropriate repetition of the taught material happens as children progress through school (Vulliamy and Webb, 1995). In some countries, changes toward a more centralized curriculum increased the results in international assessments (Walberg et al., 2000).

A more recent study of 19 educational projects across 13 countries in the Global South showed that programs that introduced teacher's guides impacted significantly on learning outcomes, with other factors like teacher professional development and coaching strategies determining the size of these effects. When looking at the intensity of scripting, they found a moderately negative relationship with outcomes: the provision of fully scripted lessons at the beginning and the transition to less structured proposals as the year went on tended to improve outcomes more than word-by-word scripting strategies (Piper et al., 2018b).

As regards textbooks, they certainly affect teaching and learning in several ways. They determine taught contents and ways of teaching (Stern and Roseman, 2004, cited in Behnke, 2018). Studies that have looked at their impact on student achievement show mixed results (Piper et al., 2018a; Polikoff, 2015). Studies conducted in the 1980s and 90s found positive effects but relying

on methods that did not allow to isolate the causal effect of textbooks (Piper et al., 2018a). The most recent ones emphasize the fact that the book itself is not enough to improve performance; but other factors interact with textbook availability to produce. Other authors stress the importance of textbook design when looking at their impact: Glewwe et al. (2009) point out that books may be aiming too high or in a language that children are only starting to learn (Piper et al., 2018a).

Some studies found an effective role of curriculum-based external exit examinations, as systems with this type of high-stakes assessments outperform countries or states without standard exit exams on math and science (Bishop and Woessman, 2004; Carnoy and Loeb, 2002; Fuchs and Woessmann, 2007; Schutz et al., 2007). There are also comparative studies that show the positive impact of centralized examinations on reducing socioeconomic inequalities and tracking (Bol et al., 2014).

However, other studies emphasize that the differences on models of curricular governance cannot be understood without addressing the history and culture of the countries. One of the largest comparative studies of the effects of curricular organization models on student learning, based on data from the TIMSS assessment, indicates this: “variations (on curriculum structures) were consistent with the cultural hypothesis that suggested that curriculum, learning, and their relationship were country specific and that this also likely reflected deeper cultural differences” (Schmidt et al., 2001, p. 282).

A profound study on classroom practices in five countries also found deep culturally embedded differences in curricular governance traditions (Alexander, 2001). A comparative study on classroom practices and educational policies from Cuba, Chile and Brazil found that one of the main differences that explain Cuba’s better performance in international assessments was related to more curricular coherence. Teacher education prepared teachers for the real curriculum in Cuba, while in Brazil and Chile they were far away from classroom realities. However, these differences were related to the profound differences in the state social cohesion of the socialist regime of Cuba and were impossible to separate from this cultural heritage (Carnoy, 2007).

Other studies have discussed the findings of some comparative research that found favorable results in centralized education systems. Westbury and Hsu (1996) discuss the thesis of Stevenson and Baker (1991) on the positive effect that the centralization of the curricular government would have in less variation in the curricular coverage of teachers. The authors point out that countries are not the central variable to explain variations in curricular coverage, but rather the way in which courses are framed. When asked about how curriculum is structured and how are course regulated, the authors find “that the categories of the school and its curriculum reflect and are embedded within larger social understandings defining what schooling or a curriculum mean” (Westbury and Hsu, 1996, p. 136). It is culture and symbolic definitions that are shared within a society at large that explain differences in the governance of curriculum and the enactment process on classrooms and teachers’ practices.

However, it is not only culture that shapes the curriculum voyage to the classrooms. The most sophisticated studies have managed to escape the simplified definitions of curricular centralization vs. decentralization. For example, the classic study “Why schools matter” shows that countries achieve better results when they structure the curriculum with greater focus and rigor: “curricular materials in high-performing nations focus on fewer topics, but also communicate the expectation that those topics will be taught in a deeper, more profound way” (Schmidt et al., 2001, p. 303).

Another study found that the policy tools had little effect on curriculum implementation or on teachers’ perceptions of alignment. “Instead, implementation was related to the availability of multiple curriculum materials in teachers’ schools and to differences between the state’s and teachers’ construal of what standards could be taught with the materials” (Penuel et al., 2009, p. 657). The enactment of the curriculum was much more complex than what centralized policies expected.

Another research drew data from 37 countries that participated in TIMSS to test two different hypotheses related to curriculum governance: location vs. credibility. Schmidt and Prawat (2006) explore whether the location of final authority for curriculum governance (at the national, the state or the local level, for example) matters as a factor for curriculum coherence. The study found that as long as there is credibility around an institutional center with authority over system goals, its location is not relevant when analyzing the overlap in topic coverage between different policy instruments and the consistency among teachers in terms of curriculum coverage.

These studies show that centralization is too much of a broader concept to grasp any significant meaning by itself. Instead, we should look at the many complexities that shape the enactment process of curriculum governance. This is a concrete way of escaping the temptation of the polarization of the dilemma “centralization vs. teacher autonomy.” To move forward this “gray arena,” we conclude this article with some of the key lessons learned on how to rethink the centralization of curriculum governance.

Making sense of centralization

The analysis so far allowed us to open the hidden boxes of the debate on curricular centralization. We have seen that the curriculum cannot be isolated from the rest of the educational system: a systemic perspective is necessary (Schmidt et al., 2001, p. 358). Instead of assuming a fixed position on the dilemma, other authors have proposed a combination. As Deng puts it, “a successful combination model needs to strike a balance between top-down and bottom-up approaches. It needs to acknowledge, on the one hand, the key role of classroom teachers as curricular change agents and, on the other hand, the need for institutional and programmatic curriculum planning in guiding, supporting, and enabling curricular change at the classroom level” (Deng, 2010, p. 388).

Dimensions involved in the design of curriculum policy pile up over each other as overlapping layers of complexity making up a sophisticated enterprise. Unpacking them will provide a way to overcome the centralization/autonomy dichotomy.

A fine-grained look at government action should pay particular attention to the broad spectrum of ways in which curricular messages are (intended to be) delivered to teachers. A way of escaping the simplification of the centralization dilemma is to

shed light on the technical sophistication implied in curriculum governance. Instructional components or instruments for instructional guidance (Cohen and Spillane, 1992) are of several types, each of which operates through relatively stable and predictable mechanisms to influence what and how is taught.

Curriculum frameworks or standards, the documents specifying expectations for what should be taught, express the intended curriculum and usually set the terms of reference for the whole enterprise, which means other components of the instructional complex are based on the directions written in curricular frameworks (Cohen and Spillane, 1992; Remillard and Heck, 2014).

Another relevant instrument for guidance are instructional materials as resources specially designed to support or complement teaching, work as mediators between the intended and implemented curriculum, and thus exert a translational role by offering teachers interpretations of instructional goals through concrete suggestions of instructional activities and classroom sequences (Schmidt et al., 2001).

Teacher training, both pre-and in-service, can work toward guidance as well. Pre-service training can be more or less centrally regulated, through mechanisms as diverse as direct inspection, curriculum frameworks, or certification standards. Its institutional design and other rules surrounding it need to be understood in relation to rules concerning teacher recruitment and licensing policies, which themselves open up a further layer of complexity as regards this instructional component (Cohen and Spillane, 1992). Workshops, courses or programs of different extension and with different kinds of consequences for the teaching career make up in-service training. It shares with initial teacher training its focus in capacity-building, although the extension of training actions in CPD make it a shorter-term strategy with different implications.

Standardized assessments can be curriculum-based, and thus intend to produce evidence on the attained curriculum; or independent of curricula (Eckstein and Noah, 1992). Most importantly, assessments can comprise high or low-stakes for students and teachers. In some systems, exam performance determines student promotion or further education options and thus tend to powerfully orient teaching efforts, even up to the extent of generating unwanted effects such as reducing instruction to the tested curriculum and educational experiences to test item drills. In other cases, exams serve diagnostic purposes and entail no major consequences for school actors, and probably less guidance potential.

These policies have grown over the past 2 decades. Test-based accountability focuses on students' learning outcomes and involves the generation of data through large-scale standardized evaluation instruments. The rise of this new regulatory governance in education is rooted in a new ideology that promotes the modernization and rationalization of education systems through new public management, economic pressures for educational reforms and a global landscape of testing industries (Verger and Parcerisa, 2018).

Another instrument is school inspection or oversight of instruction, which can also steer teaching at the school level. Models vary from more advice-oriented to more punitive ones, and also in terms of the attention paid to curricular content coverage. All of these variables have effects in the extent to and the ways in which inspection can guide instruction.

The curricular environment these instruments make up is crossed by more or less explicit divisions of authority and capacity among a broad range of institutional actors (Elmore and Sykes, 1992). In every setting, around each of the instructional guidance instruments a map of stakeholders can be traced, which adds to the complexity of the matter. The State (at its different levels and through agencies with different degrees of autonomy), the market, subject matter specialists, universities, teacher unions, parents' associations and religious entities all have a stake in curriculum decisions where their spheres of influence are involved. The possible combinations of participating actors, roles they can assume, rules for participation and degree of involvement are hard to fit into the simple centralization/decentralization divide. For example, universities that set admission standards or train teachers affect curriculum; and testing companies are usually interested in supplying assessment instruments that match State-established curricular guidelines (Elmore and Sykes, 1992).

Textbook policies can help depict the sophistication of the matter. The State can either produce the books, regulate their production or use by schools (for example, by setting a list of limited approved options or by making some textbooks compulsory), purchase them from private providers (mediated or not by bidding processes where, for example, curriculum alignment rules can be introduced) or not participate in the textbook provision and distribution circuits. These options impact on the role of publishing houses. Also, subject matter specialists (often university-based) may be part of State-led commissions for textbook assessment, or be contracted by publishers to produce textbook content. Teachers' choices of instructional materials are influenced by this kind of institutional dynamics, but not determined by them, since most teachers have access to an increasingly wide array of materials in digital formats. This complex and unorganized architecture builds up around each of the instructional components and makes it difficult to talk in terms of the simple centralization/autonomy dichotomy.

The broad range of actors involved in curriculum governance can represent a challenge to the alignment between different instructional components (Porter, 2002), which is considered desirable under the premise that poor coherence among these elements configure conflicting sets of demands on teachers and thus compromises education policy implementation (Fuhrman, 1993). Well-aligned systems, instead, provide common expectations to teachers through a coherent "curricular environment" of instruments intended to steer teaching (Valverde, 2009). This view is critical of policies based on a "hopeful pedagogy," where no deliberate efforts to achieve a coherent scheme are made (Atuhurra and Kaffenberger, 2020).

Alignment considerations add an additional layer of complexity to the design of curriculum governance. However, it can lead to the temptation of equating high levels of alignment with policies undermining teacher autonomy via centralized curriculum development. A closer look at policy options that can promote alignment shows this is far from being the only way to coherence.

Schmidt and Prawat (2006) consider alignment as a relevant criterion for curriculum development, but not a sufficient one. They stress the importance of working with a broader conceptualization of curricular coherence, which Schmidt et al. (2005) elaborate

further. This introduces a new layer of considerations pertaining to the internal architecture of curricular standards. Based on Bruner's definition of understanding, they define content standards to be coherent if they are "articulated over time as a sequence of topics and performances consistent with the logical and, if appropriate, hierarchical nature of the disciplinary content from which the subject matter derives" (2005, p. 528).

Another dimension of the internal structure of curriculum standards is brought up by Pritchett and Beaty (2015), who find out that many countries show small learning gains in children after several years of schooling (a situation to which they refer through the idea of "flat learning profiles"). They stress the importance of attending curricular pace as a relevant criterion: "if the official school curriculum covers too much, goes too fast and is too hard compared to the initial skill of the students and the ability of the schools to teach this can produce disastrous results. An overambitious curriculum causes more and more students get left behind early and stay behind forever" (2015, p. 280).

This perspective shows curriculum governance as a highly sophisticated policy realm made up of overlapping layers of technical complexity. The choice and design of instruments for instructional guidance, the network of institutional (state and non-state) actors involved in their design and deployment, the extent to which instruments are aligned with each other and the internal architecture of curriculum guidelines make up just a few of the multiple dimensions entailed in the technical design of curriculum governance.

But the question is not limited to the technical dimension. There is a political dimension that allows us to strike the final blow to the centralization vs. autonomy dichotomy. The extent to which governments build direct bridges with schools through listening and feedback mechanisms and the ways in which they open the game to plural participation in defining curricular priorities are part of this dimension. The State can provide detailed guidelines on how to teach and still give freedom to teachers in their choice, trying to introduce other policy instruments that build teachers' capacities for the judicious selection of these materials.

There are multiple alternatives and it is not possible to reduce them to such a binary dichotomy as centralization-autonomy. The analysis shows us that these are not separate paths but ingredients that are combined in different layers of curricular governance; in proportions that are not fixed but which, in the best of cases, are adjusted to the needs of the cultural and historical roots of every educational system and to the capacities and sense-making processes by teachers.

Besides, the dichotomy becomes weaker as governance itself shifts toward network-based structures made up of state and non-state actors and based on disparate sites, where the centralization-decentralization divides no longer captures the complex architecture of regulation. As Ball puts it, there is a shift toward a more polycentric state: "a controlled decontrol, the use of contracts, targets and performance monitoring to 'steer' from a distance, rather than the use of traditional bureaucracies and administrative systems to deliver or micro-manage policy systems" (Ball, 2008, p. 41).

A democratic and iterative position in terms of curricular governance need not be centralized or decentralized. It can escape top-down imposition and control through capacity-building, plural dialogs, evidence-informed decision making, teacher-led curriculum development commissions and so many other non-dichotomous alternatives.

Conclusions

Vertical curriculum models have gained increasing weight in recent curriculum debates in many countries. There is a temptation to centrally regulate the curriculum through different policy instruments: standards, textbooks, centralized tests, or scripted lessons. This curriculum governance shift clashes with professional autonomy and cultural diversity.

Franklin Bobbitt's classic text already posed this dilemma in 1918: Should a curriculum be envisaged as the blueprint (map, journey, destination) that underpins (and predates) the day-by-day activities of schooling? Or should a curriculum be envisaged primarily as the "consciously directed training experiences" themselves? (Bobbitt, 1918, p. 43).

In this text we open multiple questions to escape the simplifications of the dilemma in three ways. In the first place, looking for the complexity of the curricular government. There are different policy instruments, diverse actors and technical layers of curricular design. An overemphasis on centralization may be associated with perspectives more focused on accountability as external pressure for improvement. It is possible to overcome this perspective with the analysis of the sophistication of the field of curricular government, taking the lessons of internal organization of the curricular models that show to have better results in terms of student learning (Cohen and Spillane, 1992; Elmore and Sykes, 1992; Schmidt et al., 2001; Valverde, 2009).

Second, it is important to assume the complexity of the curriculum enactment process. A simplified view of centralization can fall into the trap of thinking that the rules are applied by sheer force. This ignores the weight of cultures and traditions that may be more oriented toward the fulfillment of state mandates or toward the diversity of interpreting practices in schools. As Remillard and Heck indicate: the education system is a map of "places within the system that curricular elements are translated and transformed" (Remillard and Heck, 2014, p. 705). Everywhere in the diversity of situations of learning there are disputes, interpretations and transformations of the curricular mandates.

Third, the centralization dilemma requires recognition of the identity and agency of teachers. There is a repeated lesson in curriculum policy studies: programmatic curriculum planning affects classroom curriculum planning only if teachers understand and employ the adopted framework and materials (Snyder et al., 1992). A democratic curriculum is one that consults teachers in their work context to adapt the curricular text to reality and to give a voice to the actors of the educational system.

These three escapes from the dilemma allow us to take the lessons of the history of centralized curriculum models and rethink them in the 21st century. We live in a time of wide cultural diversity, constant change and an uncertain future. The curriculum

cannot be something fixed or excessively external to teacher's realities or alien to the culture of a country. Neither can we ignore the importance of accountability, alignment and systemic coherence, as different studies teach us. Finding the balances, mixes, and blends is a central challenge of curriculum governance today.

References

- Alexander, R., 2001. *Culture and Pedagogy. International Comparisons in Primary Education*. Blackwell, London.
- Alexander, R., 2010. "World class schools"—noble aspiration or globalized hokum? *Compare* 40, 801–817.
- Atuhurra, J., Kaffenberger, M., 2020. System (In) Coherence: Quantifying the Alignment of Primary Education Curriculum Standards, Examinations, and Instruction in Two East African Countries (No. 20/057). RISE Working Paper Series. RISE Program.
- Ball, S., 2008. *The Education Debate*. The Policy Press-University of Bristol, Bristol. <https://doi.org/10.1038/022312a0>.
- Behnke, Y., 2018. Textbook effects and efficacy. In: Fuchs, E., Bock, A. (Eds.), *The Palgrave Handbook of Textbook Studies*. Palgrave Macmillan, New York.
- Bernstein, B., 1999. Vertical and horizontal discourse: an essay. *Br. J. Sociol. Educ.* 20, 45–65. <https://doi.org/10.4324/9780203964248>.
- Bernstein, B., 2000. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Rowman & Littlefield, New York.
- Bishop, J.H., Woessman, L., 2004. Institutional effects in a simple model of educational production. *Educ. Econ.* 12, 17–38.
- Bloom, B., 1956. *Taxonomy of Educational Objectives*. Longmans, Ann Arbor.
- Bobbitt, J., 1918. *The Curriculum*. Houghton Mifflin, New York.
- Bodovski, K., Byun, S.-Y., Chykina, V., Chung, H.J., 2017. Searching for the golden model of education: cross-National Analysis of Math Achievement. *Compare* 47, 722–741.
- Bol, T., Witschge, J., Van de Werfhorst, H.G., Dronkers, J., 2014. Curricular tracking and central examinations: counterbalancing the impact of social background on student achievement in 36 countries. *Soc. Forces* 92 (4), 1545–1572. <https://doi.org/10.1093/sf/sou003>.
- Carnoy, M., Loeb, S., 2002. Does external accountability affect student outcomes? A cross-state analysis. *Educ. Eval. Policy Anal.* 24, 305–331.
- Carnoy, M., 2007. *Cuba's Academic Advantage: Why Students in Cuba Do Better in School*. Stanford University Press, Stanford.
- Center for Curriculum Redesign, 2015. *Four-Dimensional Education: The Competencies Learners Need to Succeed*. Center for Curriculum Redesign.
- Chizzotti, A., Ponce, B.J., 2012. O currículo e os sistemas de ensino no Brasil. *Curriculo sem Front.* 12, 25–36.
- Clune, W.H., 1993. The best path to systemic educational policy: standard/centralized or differentiated/decentralized? *Educ. Eval. Policy Anal.* 15, 233–254. <https://doi.org/10.3102/01623737015003233>.
- Cohen, D.K., Spillane, J.P., 1992. Chapter 1: policy and practice: the relations between governance and instruction. *Rev. Res. Educ.* 18, 3–49. <https://doi.org/10.3102/0091732X018001003>.
- Comenius, J.A., 2019. *Orbis Pictus*. Creative Media Partners.
- Cox, C., 2018. Curriculum: categorías de análisis, tendencias, gobernanza. In: Arratia, A., Oсандón, L. (Eds.), *Políticas Para El Desarrollo Del Currículum. Reflexiones y Propuestas*. Ministerio de Educación, UNESCO, Santiago, Santiago de Chile.
- Deng, Z., 2010. Curriculum planning and systems change. In: McGaw, B., Baker, E., Peterson, P. (Eds.), *International Encyclopedia of Education*. Elsevier, Oxford, pp. 384–389.
- Dewey, J., 1997. *Experience and Education*. Touchstone, New York.
- Durkheim, E., 1977. *The Evolution of Educational Thought: Lectures on the Formation and Development of Secondary Education in France*. Routledge & Kegan Paul, London.
- Eckstein, M., Noah, H., 1992. *Examinations: Comparative and International Studies*. Oxford University Press, Oxford.
- Elmore, R., Sykes, G., 1992. Curriculum policy. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 185–215.
- English, F., Steffy, B., 2001. *Deep Curriculum Alignment: Creating a Level Playing Field for All Children on High-Stakes Tests of Educational Accountability*. Rowman & Littlefield Education, Lanham.
- Fuchs, T., Woessmann, L., 2007. What accounts for international differences in student performance? A re-examination using PISA data. *Empir. Econ.* 32, 433–464.
- Fuhrman, S., 1993. *Designing Coherent Education Policy: Improving the System*. Jossey-Bass, Washington D.C.
- Glewwe, P., Kremer, M., Moulin, S., 2009. Many children left behind? Textbooks and test scores in Kenya. *Am. Econ. J. Appl. Econ.* 1 (1), 112–135. <https://doi.org/10.1257/app.1.1.112>.
- Hamilton, D., 1989. *Towards a Theory of Schooling*. Routledge, London.
- Hamilton, D., 1990. *Curriculum History*. Deakin University Press.
- Hopmann, S., 2015. "Didaktik meets curriculum" revisited: historical encounters, systematic experience, empirical limits. *Nordic J. Stud. Educ. Policy* 1, 27007. <https://doi.org/10.3402/nstep.v1.27007>.
- Hunter, I., 1994. *Rethinking the School: Subjectivity, Bureaucracy, Criticism*. Routledge.
- Judge, H., Lemosse, M., Paine, L., Sedlak, M., 1994. *The University and the Teachers: France, the United States, England*. Traingle Books, Oxford.
- Kelly, A., 2009. *The Curriculum: Theory and Practice*. SAGE Publications, Thousand Oaks.
- Maheady, L., Rafferty, L., Patti, A., Budin, S., 2016. Leveraging change: influencing the implementation of evidence-based practice to improve outcomes for students with disabilities. *Learn. Disabil. A Contemp. J.* 14, 109–120.
- Morris, P., 2015. Comparative education, PISA, politics and educational reform: a cautionary note. *Compare* 45, 470–474.
- Mourshed, M., Chijioke, C., Barber, M., 2010. *How the World's Most Improved School Systems Keep Getting Better*. McKinsey & Company, Washington, DC.
- Ozouf, M., 1970. *L'École, L'Église et La République (1870–1914)*. Gamallard, Paris.
- Penuel, W., Fishman, B.J., Gallagher, L.P., Korbak, C., Lopez-Prado, B., 2009. Is alignment enough? Investigating the effects of state policies and professional development on science curriculum implementation. *Sci. Educ.* 93, 656–677. <https://doi.org/10.1002/sce.20321>.
- Perkins, D., 2014. *Future Wise: Educating Our Children for a Changing World*. Jossey-Bass, San Francisco.
- Piper, B., Simmons Zuilkowski, S., Dubeck, M., Jekemei, E., King, S.J., 2018a. Identifying the essential ingredients to literacy and numeracy improvement: teacher professional development and coaching, student textbooks, and structured teachers' guides. *World Dev.* 106, 324–336. <https://doi.org/10.1016/j.worlddev.2018.01.018>.
- Piper, B., Sitabkhan, Y., Mejia, J., Betts, K., International, R.T.I., 2018b. *Effectiveness of Teachers' Guides in the Global South: Scripting, Learning Outcomes, and Classroom Utilization*. Occasional Paper. RTI Press Publication OP-0053-1805. RTI Int.
- Polikoff, M.S., 2015. How well aligned are textbooks to the common core standards in mathematics? *Am. Educ. Res. J.* 52, 1185–1211. <https://doi.org/10.3102/0002831215584435>.
- Porter, C., 2002. *Measuring the content of instruction: uses in research and practice*. In: 2002 Presidential Address. University of Wisconsin, Madison.
- Pountney, R., McPhail, G., 2017. Researching the interdisciplinary curriculum: the need for "translation devices". *Br. Educ. Res. J.* 43, 1068–1082. <https://doi.org/10.1002/berj.3299>.
- Pritchett, L., Beatty, A., 2015. Slow down, you're going too fast: matching curricula to student skill levels. *Int. J. Educ. Dev.* 40, 276–288. <https://doi.org/10.1016/j.jiedudev.2014.11.013>.
- Remillard, J., Heck, D.J., 2014. Conceptualizing the curriculum enactment process in mathematics education. *ZDM* 46, 705–718. <https://doi.org/10.1007/s11858-014-0600-4>.

- Schmidt, W.H., Prawat, R.S., 2006. Curriculum coherence and national control of education: issue or non-issue? *J. Curric. Stud.* 38, 641–658. <https://doi.org/10.1080/00220270600682804>.
- Schmidt, W.H., McKnight, C.C., Houang, R., Wang, H., Wiley, D., Cogan, L., Wolfe, R., 2001. *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. Jossey-Bass, San Francisco.
- Schmidt, W.H., Wang, H.C., McKnight, C.C., 2005. Curriculum coherence: an examination of US mathematics and science content standards from an international perspective. *J. Curric. Stud.* 37, 525–559. <https://doi.org/10.1080/0022027042000294682>.
- Schutz, G., West, M., Woessmann, L., 2007. *School Accountability, Autonomy, Choice, and the Equity of Student Achievement: International Evidence From PISA 2003* (No. 14). Education Working Paper, Paris.
- Schwab, J., 1959. The "impossible" role of the teacher in progressive education. *Sch. Rev.* 67, 139–159.
- Smith, M.S., O'Day, J., 1990. Systemic school reform. *J. Educ. Pol.* <https://doi.org/10.1080/02680939008549074>.
- Snyder, J., Bolin, F., Zumwalt, K., 1992. Curriculum implementation. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 402–435.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development*. Pearson Education.
- Stern, L., Roseman, J., 2004. Can middle-school science textbooks help students learn important ideas? Findings from project 2061's curriculum evaluation study: life science. *J. Res. Sci. Teach.* 41, 538–568.
- Stevenson, D., Baker, D., 1991. State control of the curriculum and classroom instruction. *Sociol. Educ.* 64, 1–10.
- Taba, H., 1962. *Curriculum Development. Theory and Practice*. Harcourt, Brace and World, New York.
- Timar, T., Kirp, D., Kirst, M., 1998. *Tionalizing Mathematics and Science Reform: Do State Education Agencies Matter?* WestEd, San Francisco.
- Tyler, R.W., 1969. *Basic Principles of Curriculum and Instruction*. The University of Chicago Press, Chicago, London.
- Valverde, G., 2009. Estándares y evaluación. In: Schwartzmann, S., Cox, C. (Eds.), *Políticas Educativas y Cohesión Social En América Latina*. CIEPLAN, Santiago de Chile.
- Verger, A., Parcerisa, L., 2018. Test-based accountability and the rise of regulatory governance in education: a review of global drivers. In: Wilkins, A., Olmedo, A. (Eds.), *Education Governance and Social Theory: Interdisciplinary Approaches to Research*. Bloomsbury, London.
- Vulliamy, G., Webb, R., 1995. The implementation of the National Curriculum in Small Primary Schools. *Educ. Rev.* 47, 25–41.
- Walberg, H., Paik, S., Komukai, A., Freeman, K., 2000. Decentralization: an International Perspective. *Educ. Horizons* 1.
- Westbury, I., Hsu, C., 1996. Structures of curriculum governance and classroom mathematics. *Educ. Eval. Policy Anal.* 18, 123–139.

The rise and relevance of curriculum reform

Kirsten Sivesind, Department of Education, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

The rise of curriculum reform in Europe (1500–1800s)	28
How curriculum reform changed scientific meaning at the turn of the 20th century	30
Revolutionary approaches and the reinvention of curriculum in the 21st century	31
References	33

Curriculum as an institutional arrangement engages networks of actors, including public ministries, agencies, textbook publishers, universities and colleges, unions, school practitioners, and public interest groups (Sivesind and Westbury, 2016). In this context, both public and private school authorities formulate curricula to coordinate their education systems. Formal curricula are traditionally presented in the form of books or booklets and conveyed via web pages or electronic platforms. In any form, curricula can mediate general principles and guidelines, overall goals and objectives, and content that is, what to teach and learn over a period (Karseth and Sivesind, 2010). They can also present evaluation standards and assessment criteria that help decide when students improve their learning and complete a study course (Scholl, 2012; Westbury, 2007).

When school authorities reform their programs and study courses, academics develop knowledge, relevant in teacher education. Here, students learn about past, present, and future orientations to reform work and how to make sense of curriculum in everyday school practice. Academics are also involved in curriculum implementation within and across policy realms and levels (Priestly et al., 2021; Wahlström and Sundberg, 2017). Through a branch of activities and by various mandates and purposes, educational researchers deploy knowledge in various forms, which, over time, has created a field known as curriculum studies (Hopmann, 2015).

Currently, experts in the field of media and technology develop perspectives on education that challenge this study field. For example, Collins and Halverson (2018, p. 67) argue that technology-based seeds of new systems revolutionize education by innovations and developments that allow for a new type of flexibility in pedagogy and learning that outpaces both curriculum and schooling as traditional knowledge fields. They contend that the old curriculum tradition corresponds with a “grammar of schooling” that does not resonate with rationales, needs, and expectations of how to educate children and adolescents within the 21st century. By considering the 20th century industrialization of North America as the breaking shift toward a modern school system, they ignore the fact that various modes of educational reasoning modernized public education in Europe long before Americans invented their mass education system.

This chapter aims to elucidate the various modes of reasoning that have legitimized curriculum reform and public schooling across centuries. It sheds light on how curriculum reform has developed into an institutional arrangement of administrations and schools in both Western Europe and North America. The chapter provides historical insights into relevant concepts and perspectives to understand various rationales for curriculum reform, both about the way reform gained interest by school authorities in Europe before the 20th century and how curriculum turned into the vocabulary of reformers in the United States and elsewhere over the two last centuries.

I begin with a historical review of the first European efforts to develop comprehensive school reforms legitimized by a philosophy of *Bildung* (Sivesind, 2021, 2022; Sivesind and Luimes, 2017). Interestingly, the idea of general schooling appeared in academic literature in the 16th century before state administrations became the site for organizing reforms that implied curriculum renewals. In the second part, I present how reform became a key feature of North American education and how historical and empirical conceptions of curriculum reform emerged in the 20th century. Reform efforts in the United States during the early and mid-20th century emphasized the significance of disciplinary knowledge and thereby ideational streams of thought that were invented at European universities during the 16th and 17th century.

In the third part, I discuss whether the modes of reasoning that have evolved through history are relevant to the 21st century. As mentioned, some researchers today claim that curriculum studies are out of date due to new technology and media that revolutionize education. I argue, however, that without acknowledging the democratic value of curriculum reform as an institutional arrangement, it is almost impossible to understand how public schooling can serve the democratization of society. Thus, I conclude that both for continuing to develop democratic societies and to handle various types of social, cultural, political, and educational problems, research on curriculum reform is highly relevant. This also means considering the intended and unintended impacts of technological innovations on public schooling, which is a topic that further research needs to examine more closely.

The rise of curriculum reform in Europe (1500–1800s)

Comparative history studies examine how ideas and institutions of relevance for public schooling have evolved over centuries in concert with societal developments. They illuminate how curriculum embraces ideas of public schooling and explore how schooling

develops in concrete and diverse forms (Hamilton, 1989; Hopmann, 1999, 2007; Künzli, 1998). Broadly seen, researchers in this field consider historical changes in curriculum practices as movements underpinned by political, public, and academic ideas and traditions (Kliebard, 1992; Schubert, 1986; Tanner and Tanner, 1990). Also, researchers review structural and purposeful changes in teaching practices as an important dimension of curriculum and schooling.

Curriculum implies a systematic and structured educational practice. As an institution, it lasts over a period different from a cursus where students do not necessarily comply in accordance with a planned study course (Hamilton, 1989). Researchers, such as Reid (1999, p. 97), consider curriculum as an institution that is characterized by both a beginning, a process, and an end. The beginning can be the enrolling procedures of including students in a program, the process may well reflect teaching and learning activities in the classroom, and the end can be the final exam or another confirmation that signalizes that the study course is completed.

As already noted, curriculum reform as a modern idea draws on classical modes of reasoning that include conceptions of a public order (Hopmann, 1999; Weniger, 1999/1952). A semantical analysis of reform helps unravel the meaning of this concept. At first, intellectual scholars used the term revolution to capture a particular order of the world, as the concept reflected a natural order as much as a societal one. According to Hansson (1991, p. 7), the term revolution originated from medieval astronomy, referring to the regular return of planets along their orbits in space. Against this backdrop, reform means to reproduce the original form. Hansson (1991) emphasizes that the Christian Reformation leading to Protestantism at the beginning of the 16th century received its name from the concept of reform. In this context, reform was meant to reinstate a form of God-given order. Although actors were considered as individuals with a free will, they were not in a position to change the order of the Genesis and could not change the structural basis for what to preach and teach. The catechetical method of, for example, the Lutheran contract school that resulted from the Protestant Reformation was founded on this conception of a structural order (Künzli, 1998; Sivesind, 2021, 2022).

At the same time, a reconstitution of the church during the Protestant reformation aimed at accomplishing a practical order that substantiated rhetorical principles of creating trust, which in the church related to the interaction between the priest and his congregation. The theological doctrine allowed for a worldly relationship as well that were human by being informed by reason. Thus, rationales beyond a predetermined order evolved along with the development of Protestantism that influenced national constitutions as well as public schooling in, for example, Germany, and the Nordic countries in Europe (Tröhler, 2012). It is in this new context that reformists developed ideas of reform and *Bildung* that enlightened a theory of teaching called *Didaktik* that conceptualized schooling beyond a theological world view (Künzli, 1998).

The reform history in Europe serves as a brilliant example of how curriculum and reform first entered the vocabulary of academic scholars in education. Because of the new possibilities in publishing text materials arose in the 16th century, scholars created textbooks that shaped teaching and learning in substantial ways. These materials, as well as the organization of teaching activities, created modern conceptions of curriculum content. For subjects other than religion, university professors outlined scientific disciplines of significance for educating students (Nordkvelle, 2011). In European states, the invention of print media resulted in new knowledge that also influenced elementary schooling during the 19th century. New books for teaching were written along with the growing number of schoolhouses and school buildings (Sivesind, 2008).

One of the classical texts from this period is Ratke's (1664/1967) book that inspired educational thinking and reasoning, relevant for curriculum and teaching. As an enlightenment scholar, Ratke wrote about a system of grades that structured a public school system along with the Latin notion of *Currere* (eng. to run). Also, the philosopher Comenius (1659/2007) wrote his *Didactica Magna* in this period, which drew on encyclopedic reasoning, including ideas about educational governance besides moral reasoning of significance in teaching practices (Norlin, 2020). Thereafter, in the 18th century in Europe, noble people, including popes and kings, began organizing elementary teaching for children. Their highly elitist projects transformed over time and turned into public schooling under the motto "education for everybody". This expression was inspired by Comenius' work and his ideas of *Allgemeinbildung*, which aimed at cultivating the entire population in a country (Klafki, 2000).

In German states, it became essentially important to establish and regulate compulsory schooling along with the invention of state administrations. In the Nordic region, curriculum-making processes were first considered part of law revisions that were legitimized by religious reforms that resulted in new regulations and ordinances of the state as well as the school. These law revisions first gained their legitimacy through the approval of the monarchs who authorized their rationales on a religious manifesto. More generally, reforms became decisive for building and authorizing public systems or sectors that differed from larger, more dramatic changes in society associated with revolutions and wars. Moreover, during the century, the adoption of national constitutions in various parts of Europe legitimized powerful steps toward establishing compulsory schooling by law (Tröhler, 2016).

During the 19th century, school owners used texts and other materials as powerful resources to differentiate curriculum reform into various subject areas. In addition to textbooks, collections of materials, such as stuffed animals, were used to bring the material to life for students. This form of secularization legitimized encyclopedic knowledge as part of school reforms for both elementary and secondary schooling. It was no longer just the law and religious texts that served as a canon, but texts about nature and history in general. This idea served to implement Friedrich Schleiermacher's (1768–1834) hermeneutic doctrine of what he considered an "art". This form of hermeneutics—that is, an approach of text interpretation that reflected upon the objective basis of sequencing and selecting content beyond what is legally and religiously predefined—became essentially important for legitimizing curriculum reforms in Europe.

Along with the development of this art, an evolving public has acquired the right to legitimize changes. As a result, regulations did not arise out of a natural right, nor were they constituted by laws alone; rather, they arose “out of right, established by relations”, to use Lorenz von Stein’s formulations from 1852 (Koselleck, 1985, p. 63). By a dialectic mode of reasoning, reform makers emphasized encyclopedic ideas that shaped the content of curriculum reforms and an increasingly number of subjects. The explosion of content that emerged during the 19th century demonstrates why formalized curricula with canonical knowledge became so important. A core aim was to rebuild relations between the state and the people. Throughout the next centuries, this arrangement became institutionalized by European states, which were dependent on political, professional, public legitimacy, and support to establish themselves as democratic institutions (Sivesind, 2008). In England, private institutions also gained legitimacy with a public right to provide education for the masses.

How curriculum reform changed scientific meaning at the turn of the 20th century

Despite the combination of administrative and academic initiatives to create a “school for all” that drew on scientific reasoning and aimed at developing a societal-relevant school system, researchers in the US argued for rationalizing the curriculum by reform during the 20th century. In the beginning, cultural and social differences between those born into noble families, those who belonged to the bourgeoisie, and those who were poor without titles and occupations marked societies as highly segregated. Moreover, conflicts and insecurity resulting from wars in continental Europe brought about immigration in the United States, which again resulted in a highly cultural-heterogenic society in the country. At that point in history, industrialization and new beliefs in technological progress created hope for a better future as well as ideas about systems that could make education more efficient.

Of key significance for developing systematic thinking about curriculum was as Kliebard (1988a, p. 45) interpreted: the “systematic thought about education generally” that implied general consideration of historical conditions. He points to the rising population of high school students that doubled over a forty-year period, every decade after 1890, as a core reason for curriculum developments. Moreover, transformations in American society resulted in a wish for an Americanization of immigrants that turned into a new vision of schooling. Professors in education who participated in the National Education Association (NEA), which became an important meeting point for national discussions in education, also influenced curriculum development, and in particular by the initiation of the Committee of Ten.

The commission was mandated to deal with the issue of heterogeneity in high schools and the question of how this school program should prepare students for enrollment in colleges. Under the direction of Charles W. Elliot, the commission argued for a highly disciplined curriculum based on the conception that everyone could learn academic knowledge independent of their cultural and social background. Therefore, no separate tracks were needed for college admission. Another prominent scholar in this period was G. Stanley Hall, who was critical of the conception that every student should be taught school subjects in the same way regardless of their destination. Thereby, he argued for “curriculum differentiation” (Kliebard, 1988a, p. 49). Also, he did not accept Elliot’s argument that preparation for college is a preparation of life, and in the end, it was Hall’s perspective that won out; that is, the goal of public schooling for adolescents was to prepare students for practical and social responsibilities rather than to accomplish intellectual attainment.

Together with a steep growth of knowledge that outpaced traditional subject scholars, John Franklin Bobbitt, in agreement with Werrett W. Charters and David Snedden developed a set of basic principles of a “scientific curriculum”. This movement implemented ideas of scientific management that Fredrick Winslow Taylor promoted in the field of organizational studies. In this era, curriculum goals were perceived as a natural phenomenon that could be discovered by observing skillful human beings and their ways of making elementary operations into an efficient series of movements. Kliebard 1988a summarized this expert notion of curriculum reform:

Curriculum development became an effort to standardize the means by which predetermined specific outcomes might be achieved. Success was judged by the extent to which the actual outcomes matched the predictions.

Kliebard (1988a, p. 51).

However, with increased urbanization, known reformists, such as John Dewey (1858–1992), also started to write about ways to deal with the problems of men (Dewey, 1946/2014). Although he failed in setting the agenda for curriculum research in his own country, as Bobbitt and Charter did (Waldow, 2015), Dewey inspired reform work far beyond the borders of the United States through his many projects and publications. A core theme related to the societal conditions that served as a frame for Dewey’s theoretical deliberations (Lundgren, 1989). A possible reason for the variation in his outreach is the fact that Dewey developed his pedagogy in a context where reforms and plans had a different authority compared with what was the case in European states, where school content was selected by professionals in collaboration with state officials. Nonetheless, for the American case, both societal and scientific reasoning and disciplined knowledge become essential dimensions for selecting and organizing curriculum knowledge. Moreover, curriculum reform did not merely refer to state public schooling but to a highly differentiated practice of various reform initiatives.

American curriculum projects that developed during the early 20th century and onward selected curriculum content through scientific rationales, an approach that the federal government intensively urged after the Sputnik shock in 1957. Then, the Soviet Union accomplished its mission and scientific advancement of sending the first satellite, Sputnik-1, into the space with a dog, Laika. Scientific reasoning was already a decisive basis for legitimizing school content in American curriculum theory at that moment. Dewey, for example, also included assessment by scientific rationales as a key to students' learning, which differed from German scholarship, which emphasized practical judgment by teachers in schools (English, 2013).

The Industrial Revolution and an increasing population in the cities were, above all, the core explanations why compulsory education developed into a mass schooling system in the United States at the beginning of the 20th century. Curriculum reform reflected societal changes that affected teaching and learning, as well as the way educators communicated about reforming programmes and courses. Social aspirations of various kinds and assessment criteria for selecting students for various programmes and study courses served as important parameters for organizing public schooling in America (Labaree, 2012). Through assessments and examination, geography, age, ability, former academic achievements as well as needs, motivations, and interests served as criteria for differentiating students in various grades and classes. These criteria were important for thinking and writing about curriculum reform as well. Thus, curriculum as a concept reflected the architecture of the education system, known as the "grammar of schooling" (Tyack and Cuban, 1995; Westbury, 1998).

Internationally, different pedagogies referring to various types of curriculum reform developed as parallel modes of thought during the 20th century. This phenomenon can be explained by the fact that there were no clear linkages between curriculum reform as an institutional arrangement across national borders and how schools responded to new regulations. Moreover, curriculum reform was also considered an expression of various cultural ideologies. This conception resulted in views of contemporary reforms in the North American tradition as a conscious move between poles of value-oriented ideologies, conceptualizing changes as "pendulum swings" and as "cycles that ebb and flow" (Cuban, 1990, p. 7).

Short-lived reform cycles have been the subject of many reports and analyses in the field of curriculum history and reform. Kliebard (1988b) suggests that one of main reasons for the steady change of reform orientations, or the ebb and flow of "curriculum fashions", is the actual absence of purpose of curricular decisions that characterizes many reform efforts. This again means that structural reforms of merely legislative or legal order do not necessarily correspond with the core purposes of curriculum ideologies reflected by the philosophy that legitimize curriculum reforms from a cultural point of view.

Moreover, Cuban (1990) uses "waves" as a metaphor for explaining how teaching and learning in schools are targeted by curriculum reform, however, not necessarily leading to a transformation of curriculum practices. He writes about school districts that initiate large-scale reforms on a national basis, but like a wave, they generated only surface change. Cuban (1990, p. 9) also observes that administrators turn "religiously to school-based solutions for national problems" while teachers continue teaching their subjects to meet expectations within the public domain. Therefore, reforms are frequently reintroduced to restructure education and reorder the inner lives of schools, however, without necessarily accomplishing their missions and goals by changing classroom practices.

Revolutionary approaches and the reinvention of curriculum in the 21st century

By investigating discourses on curriculum reform, one discovers how scientific and technological reasoning frame modes of reasoning to close the gap between reform aspirations and practices. Today, international nongovernmental organizations and expert networks serve as sites for authorizing curriculum programs in which assessment standards and competence goals serve as a common core. Many efforts to create international programs with implications for curriculum reform across nations have resulted in several definitions of competencies and skills for the 21st century. Numerous international organizations use competency and skills for the 21st century as a scientific-normative expression of what human beings should achieve through their life-course to accomplish work-life demands and societal transformations beyond what teachers can teach in a normal classroom.

The scientific effort to improve learning and instruction developed in a Cold War context and resulted in large reform programmes in the United States and elsewhere, motivated by the Sputnik shock (Tröhler and Lenz, 2015). However, since then, expressions of scientific-legitimized curriculum reform efforts have reflected paradoxical and conflicting interests, as acknowledged by UNESCO in the early 1970s. Faure et al. (1972–73) described the educational traditions and legacies of the past that put focus on both literacies and humanism, grounded within a scientific rationale for developing education within a learning society. The following quote demonstrates how the commission, consisting of ministers and experts, characterized the mastery of learning as a perplexing purpose of education:

For these reasons, the commission considered that it was essential for science and technology to become fundamental, ever-present elements in any educational enterprises for them to become part of all educational activities designed for children, young people and adults, so as to help the individual to control not only natural and productive forces, but social forces too, and in so doing to acquire mastery over himself, his choices and actions; and, finally, for them to help man to develop a scientific frame of mind in order to promote the sciences without becoming enslaved by them.

Faure et al. (1973/1972).

To argue for the development of “a scientific frame of mind in order to promote the sciences without becoming enslaved by them” (Faure et al., 1973/1972) seems to be one of the dilemmas that policymakers, professionals, and even populations could face concerning this policy agenda. This concern is entangled with international large-scale assessments that draw on particular research methodologies and scientific reasoning to legitimize curriculum reforms in the 21st century (Sivesind, 2013).

A range of programs launched by international organizations have emerged along reform efforts, such as the programs of comparative studies from the 1960s through the 1980s organized by the International Association for the Evaluation of Educational Achievement (IEA) and international assessments by the Organization for Economic Co-operation and Development (OECD), which entered this comparative field during the 1970s (Sivesind, 2019). Today, several research reviews define what experts mean by “competence” and “21st century skills” across areas and contexts (Binkley et al., 2012; Dede, 2010; Kereluik et al., 2013; Tahirsylaj and Sundberg, 2020; Voogt and Roblin, 2012). National and international organizations present “competence” and “skills” as if they were the most relevant to master in a “lifelong” perspective, a term that suggests that each person has an imperative responsibility to see him or herself as an active lifelong learner and to act accordingly. Moreover, these initiatives have various forms of support in different parts of the world (Nordin and Sundberg, 2016), depending on conditional recourses (Hopmann, 2003), political decision-making capacities (Tahirsylaj and Sundberg, 2020) as well as shared historical memories of regional and local reform initiatives (Sivesind and Karseth, 2019).

Manifestly, there has been an intensified focus on competence and learning within large projects that prospect the future school beyond 21st century skills, cognitive, and social skills that are essential to success (Hilt et al., 2019). The same goes for innovation and integration brain development—an optimal time to help young people shape their intelligence, identity, and personality. As Restad and Mølstad (2020) have shown, the focus has recently changed from a general competence-based orientation of school reforms with emphasis on literacy within and across school subjects to a more psychological orientation of the inner life of students, which draws attention to deep learning and the performance of social and emotional skills.

There is also a shift from competence built into aims and objectives for teaching and learning to a more specific focus on attitudes associated with certain values (e.g., OECD, 2020). More specifically, “deeper learning” is the process of preparing and empowering students to master a set of outcomes through learning processes that are characterized by a set of quality features (Bellanca and Boss, 2015; Mehta and Fine, 2019). The topic is not described as entirely new by educational scholars, but considered as much needed by think tanks and other stakeholders because of political and societal problems, such as the achievement gap that is not diminished by traditional school systems (All4Ed, 2022). Thus, it is interesting to investigate the way this and similar concepts and strategies are used by policymakers and experts to determine if and how curriculum reform efforts in contemporary systems fit with old styles of reasoning.

Even though current discourses are characterized by a constructive discourse of creating and recreating new curriculum reforms for the sake of democratizing society, the “grammar of schooling” as a system that resists changes serves as a strong narrative that researchers and experts criticize. Currently, revolution seems to be a scattered concept among those who promote paradigmatic changes in education and would like to reinvent and renew curricula along with novel technological innovations. Rather than focusing on curriculum reform and affiliated pedagogies, authors like Collins and Halverson (2018, p. 67) argue that technology-based seeds of new systems revolutionize education and thereby outpace both schooling and education research. The authors claim that technological revolutions alter not only the lives of children and adolescents but also every aspect of work that demands a paradigm shift in education.

Collins and Halverson (2018, p. 49) refer to modernization processes of the education system in the United States in the first half of the 19th century and Horace Manns’ education visionary for creating a social cohesive system through state involvement in education. They also reference the rationales of elitist education of the Greek gymnasium, the medieval university, and the English grammar schools as well as the invention of the printing press, the reformation, the American and the Industrial Revolutions as core explanations for the rise of universal schooling in the United States. Despite being aware of this historical trajectory, the “grammar of schooling” is considered more or less fixed without a capacity to integrate new technology as such.

A possible reason why the authors look upon the grammar as more or less fixed without a potentiality to deliberately change education in accordance with contemporary demands is the empirical fact that a system, where the state serves a passive role as a school provider, has largely been replicated as a curriculum model in large parts of the American education system. The first signs of this model date back to the mid-17th century, when Calvinists adopted formal laws for general education (p. 50). Horace Mann serves as an example of a movement in New England, where the state became a provider of general education. However, along with progressivist reformers, Collins and Halverson (2018) consider the Industrial Revolution as the prime driver that transformed “the universal schooling movement from an evangelical movement to a practical necessity” (p. 53). Thus, the Industrial Revolution is considered the key to understanding how society became modernized in the United States in the 20th century and how a “grammar of schooling” established procedures and patterns that the authors consider as not congruent with current demands.

However, in their critique, Collins and Halverson (2018) overlook the variation of traditions and modes of reasoning that legitimized schooling as a democratic project beyond the evangelic movement. Thereby, they base their view on a highly biased narrative of curriculum reform and schooling. For example, in Norway and Denmark, which belongs to the Nordic part of Europe, nation states developed education systems in which parents were considered the prime responsible actor for taking care of children’s upbringing. The contract-school model that arose during the Protestant Reformation legitimized this idea (Sivesind, 2021). In this system, monarchies developed their education systems not merely by preparing “people for their proper place in the social order” nor for preparing “students to choose public over private interests”, as claimed by Republican leaders in the United States, referenced by Collins and Halverson (2018, p. 53). Rather, the education system was developed as a pragmatic project to ensure

children's education in their own rights, which was legitimated by progressive teachers and intellectual reformists in the early 20th century (Sivesind, 2008). This is particularly true for monarchies that democratized their systems before industrialization changed their economies into a new societal order (Sejersted, 2011).

Conclusively, due to various modes of reasoning that legitimized public schooling during the 20th century and because of different forms of administrative control across countries and continents, there are no such thing as one "grammar of schooling", neither in Western Europe nor in the United States. Although more post-modern approaches to reforms have made the world into a more complex space, one may still consider curriculum reform as a place-based phenomenon with a highly particular history and therefore of relevance for developing public schooling in the 21st century. For this reason, researchers argue for re-inventing schools, the curriculum, and reforms within a global perspective (Bangs et al., 2010; Priestley and Biesta, 2013; Takayama, 2010).

Westbury and Milburn (2006) argue that "core issues policy-makers have faced have been typically set in a time and place, and framed within the discourses of their immediate worlds". Thus, the life-world of both policymakers, teachers, and all other parties are all important for how curriculum reform evolves and transforms (Goodson, 2019). I argue that digitization processes and the active involvement of edtech companies in education challenge traditions as well as curriculum reform daily (Williamson, 2017). Still, it is also of high relevance to rethink the novelty of technology in education:

Given the breathless enthusiasm with which new learning technologies are introduced, one of the most useful dispositions in evaluating edtech is to regularly ask the question, "What's really new here?"

Reich (2020, p. 21).

This question may lead to a renewed interest in the ways technologies and media interact with curriculum reform to handle social, cultural, political, and educational problems in contemporary school systems.

References

- All4Ed, 2022. Equity. Justice. Education. <https://all4ed.org/>.
- Bangs, J., MacBeath, J., Galton, M., 2010. Reinventing Schools, Reforming Teaching: From Political Visions to Classroom Reality. Routledge.
- Bellanca, J.A., Boss, S., 2015. Deeper Learning: Beyond 21st Century Skills. Solution Tree Press.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., Rumble, M., 2012. Defining twenty-first century skills. In: Griffin, P., et al. (Eds.), Assessment and Teaching of 21st Century Skills. Springer, pp. 17–66.
- Collins, A., Halverson, R., 2018. Rethinking Education in the Age of Technology: The Digital Revolution and Schooling in America. Teachers College Press.
- Comenius, J.A., 1659/2007. Orbis Sensualium Pictus. English Translation. <http://www.uned.es/manesvirtual/Historia/Comenius/OPictus/OPictusAA.htm>.
- Cuban, L., 1990. Reforming again, again, and again. Educ. Res. 19 (1), 3–13.
- Dede, C., 2010. Comparing frameworks for 21st century skills. In: Bellanca, J., Brandt, R. (Eds.), 21st Century Skills, vol. 20. Solution Tree Press, pp. 51–76.
- Dewey, J., 1946/2014. Problems of Men. Open Road Media.
- English, A.R., 2013. Discontinuity in Learning: Dewey, Herbart and Education as Transformation. Cambridge University Press.
- Faure, E., Herrera, F., Kaddoura, A.-R., Lopes, H., Petrovsky, A.V., Rahnama, M., Ward, F.C., 1973/1972. Learning to Be: The World of Education Today and Tomorrow. UNESCO.
- Goodson, I., 2019. From mystification to markets: the evolution of curriculum history and life history. In: McCulloch, G., Goodson, I., González-Bailón, S. (Eds.), Transnational Perspectives on Curriculum History. Routledge, pp. 12–26.
- Hamilton, D., 1989. Towards a Theory of Schooling. The Falmer Press.
- Hansson, S.O., 1991. Idéer om reformism. Texter av Thomas Humphrey Marshall, Adam Przeworski, Gøsta Esping-Andersen, Jon Elster. Ideas about Reformism. Texts by Thomas Humphrey Marshall, Adam Przeworski, Gøsta Esping-Andersen, Jon Elster. Tidens Förlag.
- Hilt, L.T., Riese, H., Søreide, G.E., 2019. Narrow identity resources for future students: the 21st century skills movement encounters the Norwegian education policy context. J. Curric. Stud. 51 (3), 384–402.
- Hopmann, S.T., 1999. The curriculum as a standard of public education. Stud. Philos. Educ. 18, 89–105.
- Hopmann, S.T., 2003. On the evaluation of curriculum reforms. J. Curric. Stud. 35 (4), 459–478.
- Hopmann, S.T., 2007. Restrained teaching: the common core of Didaktik. Eur. Educ. Res. J. 6 (2), 109–124.
- Hopmann, S.T., 2015. "Didaktik meets curriculum" revisited: historical encounters, systematic experience and empirical limits. Nordic J. Stud. Educ. Policy 1.
- Karseth, B., Sivesind, K., 2010. Conceptualizing curriculum knowledge—within and beyond the national context. Eur. J. Educ. (45), 103–120.
- Kereluik, K., Mishra, P., Fahnoe, C., Terry, L., 2013. What knowledge is of most worth: teacher knowledge for 21st century learning. J. Digital Learn. Teach. Educ. 29 (4), 127–140.
- Klafki, W., 2000. Didaktik analysis as the core of preparation of instruction. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), Teaching as a Reflective Practice. The German Didaktik Tradition. Lawrence Erlbaum Associates, pp. 139–159.
- Kliebard, H.M., 1988a. The drive for curriculum change in the United States, 1890–1920. Part I—the ideological roots of curriculum as a field of specialisation. In: Gress, J.R., Prurpel, D.E. (Eds.), Curriculum. An Introduction to the Field. McCutchan Publishing Corporation, pp. 44–58. Reprint from Journal of Curriculum Studies 11:3 (1979): 191–202.
- Kliebard, H.M., 1988b. Fads, fashions and rituals: the instability of curriculum change. In: Tanner, L.N. (Ed.), Critical Issues in Curriculum. The University of Chicago Press, pp. 16–34.
- Kliebard, H.M., 1992. Forging the American Curriculum. Essays in Curriculum History and Theory. Routledge.
- Koselleck, R., 1985. Futures Past. On the Semantics of Historical Time. Massachusetts Institute of Technology.
- Künzli, R., 1998. The common frame and the places of Didaktik. In: Gundem, B.B., Hopmann, S. (Eds.), Didaktik and/or Curriculum. Peter Lang, pp. 29–45.
- Labaree, D.F., 2012. Someone Has to Fail: The Zero-Sum Game of Public Schooling. Harvard University Press.
- Lundgren, U.P., 1989. Pedagogik och psykologi. Om John Dewey's filosofi [Pedagogy and Psychology. About John Dewey's philosophy. In: Thuen, H., Vaage, S. (Eds.), Oppdragelse til det moderne. Emile Durkheim, George Herbert Mead, John Dewey, Bourdieu. Universitetsforlaget, pp. 121–146.
- Mehta, J., Fine, S., 2019. In Search of Deeper Learning: The Quest to Remake the American High School. Harvard University Press.
- Nordin, A., Sundberg, D., 2016. Travelling concepts in national-curriculum policy making: the example of competencies. Eur. Educ. Res. J. 5 (3/4).

- Nordkvelle, Y.T., 2011. Petrus Ramus. Libertas philosophandi og universitetspedagogisk martyrium [Petrus Ramus. Libertas philosophandi and university pedagogical martyrdom]. In: Skodvin, A., Flyum, K.H., Knudsen, G., Simonsen, E. (Eds.), *Forelesningens kunst*. UNIPUB, pp. 73–90.
- Norlin, B., 2020. Comenius, moral and pious education, and the why, when and how of school discipline. *Hist. Educ.* 49 (3), 287–312. <https://doi.org/10.1080/0046760X.2020.1739759>.
- OECD, 2020. PISA 2018 Results (Volume V). Effective Policies, Successful Schools. Retrieved from. <https://www.oecd.org/pisa/publications/pisa-2018-results-volume-v-ca768d40-en.htm>.
- Priestley, M., Biesta, G., 2013. *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. A&C Black.
- Priestley, M., Philippou, S., Alvunger, D., Soini, T., 2021. Curriculum making: a conceptual framing. In: *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*, pp. 1–28.
- Ratke, W., 1664/1967. *Kleine pädagogische Schriften*. Julius Klinkhardt.
- Reich, J., 2020. *Failure to Disrupt*. Harvard University Press.
- Reid, W.A., 1999. *Curriculum as Institution and Practice. Essays in the Deliberative Tradition*. Lawrence Erlbaum Associates.
- Restad, F., Mølsted, C.E., 2020. Social and emotional skills in curriculum reform: a red line for measurability? *J. Curric. Stud.* 1–14.
- Scholl, D., 2012. Are the traditional curricula dispensable? A feature pattern to compare different types of curriculum and a critical view of educational standards and essential curricula in Germany. *Eur. Educ. Res. J.* 11 (3), 328–341.
- Schubert, W.H., 1986. *Curriculum. Perspective, Paradigm, and Possibility*. Macmillan Publishing Company.
- Sejersted, F., 2011. In: Adams, M.B., Daly, R. (Eds.), *The Age of Social democracy: Norway and Sweden in the Twentieth Century*. Princeton University Press.
- Sivesind, K., Karseth, B., 2019. An officially endorsed national curriculum: institutional boundaries and ideational concerns. *Curric. Perspect.* 39 (2), 193–197.
- Sivesind, K., Luimes, M., 2017. Didaktik and curriculum studies: a European perspective. In: Chi-Kin Lee, J., Kennedy, K.J. (Eds.), *Theorizing Teaching and Learning in Asia and Europe. A Conversation Between Chinese Curriculum and European Didactics*. Routledge, pp. 217–231.
- Sivesind, K., Westbury, I., 2016. State-based curriculum-making, part I. *J. Curric. Stud.* 48 (6), 744–756. <https://doi.org/10.1080/00220272.2016.1186737>.
- Sivesind, K., 2008. *Reformulating Reform. Curriculum History Revisited*. The Faculty of Education Research, University of Oslo.
- Sivesind, K., 2013. Education in the spirit of archimedes: pertaining the buoyancy of PISA. In: Nordin, A., Sundberg, D. (Eds.), *Transnational Policy-Flows in European Education. Conceptualizing and Governing knowledge*. Symposium.
- Sivesind, K., 2019. Nordic reference societies in school reforms in Norway: an examination of Finland and the use of international large-scale assessments. In: Waldow, F., Steiner-Khamisi, G. (Eds.), *Understanding PISA's Attractiveness: Critical Analyses in Comparative Policy Studies*. Bloomsbury, pp. 89–107.
- Sivesind, K., 2021. Historical trajectories of the contract-school model in Norway. In: Zhao, W., Tröhler, D. (Eds.), *Euro-Asian Encounters on 21st-Century Competency-Based Curriculum Reforms*. Springer.
- Sivesind, K., 2022. Public schooling in a “post-factual” society: a call for historically nuanced discussions. In: Krogh, E., Qvortrup, A., Graf, S.T. (Eds.), *Bildung, knowledge, and Global Challenges in Education Didaktik and Curriculum in the Anthropocene Era*. Routledge.
- Tahirsylaj, A., Sundberg, D., 2020. The unfinished business of defining competences for 21st century curricula—a systematic research review. *Curric. Perspect.* 1–15.
- Takayama, K., 2010. Politics of externalization in reflexive times: reinventing Japanese education reform discourses through “Finnish PISA Success”. *Comp. Educ. Rev.* 54 (1), 51–75.
- Tanner, D., Tanner, L., 1990. *History of the School Curriculum*. Macmillan Publishing Company.
- Tröhler, D., Lenz, T., 2015. Between the national and the global. In: Tröhler, D., Lenz, T. (Eds.), *Trajectories in the Development of Modern School Systems: Between the National and the Global*. Routledge, pp. 4–9.
- Tröhler, D., 2012. *Languages of Education: Protestant Legacies, National Identities, and Global Aspirations*. Routledge.
- Tröhler, D., 2016. Curriculum history or the educational construction of Europe in the long nineteenth century. *Eur. Educ. Res. J.* 5 (3).
- Tyack, D.B., Cuban, L., 1995. *Tinkering Towards Utopia. A Century of Public School Reform*. Harvard University Press.
- Voogt, J., Roblin, N.P., 2012. A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J. Curric. Stud.* 44 (3), 299–321.
- Wahlström, N., Sundberg, D., 2017. Transnational curriculum standards, curriculum reforms and classroom practices—an introduction. In: Wahlström, N., Sundberg, D. (Eds.), *Transnational Curriculum Standards and Classroom Practices*. Routledge, pp. 1–14.
- Waldow, F., 2015. From Taylor to Tyler to no child left behind: legitimating educational standards. *Prospects* 45 (1), 49–62.
- Weniger, E., 1999/1952. Didaktik as a theory of education. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice. The German Didaktik Tradition*. Lawrence Erlbaum Associates, Publishers, pp. 111–125. Original Publication *Didaktik als Bildungslehre, Teil 1: Die Theorie der Bildungsinhalte und des Lehrplans*. Weinheim: Beltz.
- Westbury, I., Milburn, G., 2006. Introduction: rethinking schooling—twenty-five years of the *Journal of Curriculum Studies*. In: *Rethinking Schooling*. Routledge, pp. 13–18.
- Westbury, I., 1998. Didaktik and curriculum studies. In: Gundem, B.B., Hopmann, S. (Eds.), *Didaktik and/or Curriculum*. Peter Lang, pp. 47–78.
- Westbury, I., 2007. Making curricula: why states make curricula, and how. In: Connelly, F.M. (Ed.), *The Sage Handbook of Curriculum and Instruction*. Sage, pp. 45–65.
- Williamson, B., 2017. *Big Data in Education. The Digital Future of Learning, Policy and Practice*. Sage.

Critical curriculum analysis framework (CCAF)

Christine Winter and David Hyatt, School of Education, University of Sheffield, Sheffield, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	35
Curriculum context	36
Curriculum from a critical discourse studies (CDS) perspective	37
The critical curriculum analysis framework (CCAF)	38
Applying the CCAF to a curriculum policy text ensemble: case study—geography GCSE in England	39
Introduction to the case study	39
Curriculum context	39
Curriculum policy deconstruction	40
Curriculum knowledge	41
Conclusion	44
Appendix 1	45
1 Curriculum context	45
1.1 Curriculum ensemble	45
1.2 Curriculum drivers/warrants	45
2 Curriculum policy deconstruction	46
2.1 Policy Archeology	46
2.2 Textual deconstruction	46
3 Curriculum <i>per se</i> structure	46
3.1 Curriculum construct-ors/-ion	46
3.2 Curriculum structure	46
4 Curriculum knowledge	47
4.1 Curriculum knowledge epistemology/ontology	47
References	47

Introduction

Curriculum is both a key concept in the study of education and a concept that is deeply contested. To some, the term refers to the seemingly objective knowledge content and skill-set of a program of educational study. In contrast, Apple understands curriculum as a profoundly political phenomenon: “... always part of a selective tradition, someone’s selection, some group’s vision of legitimate knowledge” (1996, p. 22), while Todd describes curriculum as “the symbolic raw material” of education (2001, p. 446). In attempts to pin down and control the slippery concept and practice of curriculum, governments globally have introduced prescriptive curriculum policies. Reform of curriculum policy does not only affect the curriculum message system,¹ it also impacts on the message systems of pedagogy and assessment which interweave symbiotically with curriculum. This is important because it is through the three message systems in schools that dominant social groups construct, select, classify and relay what is constituted as socially legitimate knowledge, values and attitudes (Bernstein, 1971). So, for example, the selection of knowledge in a curriculum specification influences its pedagogic enactment, its assessment and the values entrenched within that knowledge.

Such is the significance of curriculum policy reform in driving change in pedagogy and assessment in recent years, that Gerrard and Farrell describe curriculum policy as “a cornerstone of educational governance” (2013, p. 4). Taking this a step further, Lingard et al. (2013) identify “curriculum as systemic policy” (p. 549). In other words, curriculum policy-effects extend into classrooms through the enactment of curriculum knowledge, pedagogical practices, student-teacher relationships, school spaces, time and cultures, assessment processes, as well as through the very identities of participants. Since curriculum forms such an important site of political power in education, analyzing curriculum becomes a key concern. Hence, the central focus of this article is the language through which curriculum is constructed and the discourses it evokes.

¹Bernstein describes the three message systems of educational knowledge as curriculum, pedagogy and evaluation: “Curriculum defines what counts as valid knowledge, pedagogy defines what counts as a valid transmission of knowledge, and evaluation defines what counts as a valid realization of this knowledge ...” (1971, p. 47).

In the past, curriculum research and academic texts were located in the curriculum studies tradition, with the aim of supporting curriculum planning and development (Tyler, 1949; Schools Council, 1970; Goodlad, 1979; Kelly, 2009). While these curriculum planning/development models and perspectives can be adapted for analytical purposes, curriculum analysis did not form their *raison d'être*. More recently, education policy analysis frameworks have been adapted for curriculum analysis, such as the policy cycle model offered by Bowe et al. (1992) and Rizvi and Lingard's key policy analysis questions (2010, pp. 54–56). In addition, curriculum researchers have developed particular approaches to analyze curriculum in specific international contexts (Chun-Lok and Wing-Yan, 2009 Hong Kong; Roofe, 2020 Jamaica; Singh et al., 2013 Australia; Priestley and Sinnema, 2014 Scotland and New Zealand; Winter, 2018 England; Macedo, 2011 Brazil; Gleeson et al., 2020 Australia and Ireland; Fitz and Nikolaidis, 2020 US; Ghebru and Lloyd, 2020 Ethiopia; Nomura, 2017 Japan). In spite of these theoretical and empirical tools available for curriculum analysis, a gap exists in terms of a critical systematic text-analytic framework of value to students, scholars and practitioners across a range of educational phases, from early years to schools, through to adult/Further/Higher Education, and across different national contexts. The framework offered in this article aims to fill this gap.

The contemporary demand for such a framework arises primarily from the need to identify and challenge the seemingly innocent purveyance of neo-liberal education policies globally through curriculum reform and its impact on social justice (Lingard et al., 2013, p. 541). Ball (2008) describes how supranational organizations like the OECD, World Bank and EU drive global policy reform to embed new processes of educational governance through accountability measures to promote the economization of education and human capital formation. In the advancement of curriculum reform, common policy technologies of governance, consisting of the market, new public management and performance are operationalized (Ball, 2008, p. 42). However, shaped by national "vernacularization" (Lingard, 2006), these policy technologies find expression in particular ways in different countries. National curriculum reform policies manifest in the form of curriculum texts and discourses. For the purposes of this study, these consist of curriculum policy texts, curriculum specifications, textbooks, test/examination questions, worksheets, teaching artifacts, online teaching material, but always interpellating and/or co-occurring with other texts (intertextuality) and the discourses contained within (interdiscursivity). A critical analytical framework like the one offered here allows education professionals to deconstruct these texts and discourses to unpack and unmask how language and images expose the very power relations, values, politics and ethics at work in curriculum construction, meaning and enactment at global, national and local scales.

The context described above demonstrates the need for a clear, flexible, accessible and critical curriculum audit and deconstruction framework to support Higher Education (HE) students and staff and school/college teachers engaging in curriculum analysis/critique. Educators require a rigorous approach to investigate curriculum processes and motivations via the language, power, discourse, ideology and epistemology embedded in texts. While offering a generic and comprehensive audit tool which can be applied across a range of curriculum texts in any geographical location, the framework is not a technical, one-size-fits-all mechanism to be applied in a linear, step-by-step fashion. Instead, it requires the curriculum analyst to research the historical, political, and cultural context of the curriculum under investigation, and to make decisions about the focus of the analysis, depending on their own curriculum context and on their own reflexive stance in terms of their positionality.

Following this introduction, the article discusses insights from the literature (divided into a focus on curriculum contexts and on a critical discourse studies (CDS-oriented) approach) before presenting the framework itself. Then the framework is discussed, through an illustrative case study curriculum in the secondary school in England. We end the article with our final thoughts.

Within the article, we aim to foster cognizance by education researchers and practitioners of the ways in which potential construals in curriculum texts and enactments can obscure circulations and articulations of power. Through this, we aim to reveal the ways in which these articulations can impact upon teachers and learners. Our framework focuses on the investigation and elucidation of the means by which textual practices should be seen to be historically, ideologically and politically located social practices. The critical curriculum analysis framework (CCAF) offers an analytical structure and process for interrogating curriculum discourses. In doing so, we aim to offer those engaged in curriculum analysis the means to deconstruct the language, and its ideological essence as construed through discourse, of curriculum avowals.

Curriculum context

Before introducing the approach to discourse analysis adopted in this article and the critical curriculum analysis framework itself, it is important to situate curriculum, albeit briefly, in the global setting. One starting point is Sahlberg's Global Educational Reform Movement (GERM) (2011, 2015) which illustrates the impact of globalization and neo-liberalization on curriculum. Sahlberg describes how supranational organizations, corporations and venture philanthropists invest finances, business concepts and practices into global education systems (2015, p. 143), to propel them toward an economic, commercial mode of operation. Along with international standardized student tests like PISA as a key driver of curriculum reform, the impact of GERM around the world includes externally prescribed, mandatory curricula; prioritization of core subjects of literacy, numeracy and science; outcome-based assessment, standardization of performance through benchmarks; and test-based accountability, expressed in

school performance tables and driven by competition between schools for student recruitment (2015, p. 145). Such technicization of curriculum has involved the re-fashioning of traditional academic curricula into competency/skill-based curricula which aim to prepare students for the 21st century global knowledge economy (Savage, 2017, p. 158). Unsurprisingly, national policy players mediate global curriculum reform tendencies in different ways in different national spaces as curriculum policies are “borrowed” and “travel” between countries. The resulting policies are influenced by political, economic and historical conditions particular to each nation. Taking the cue from Appadurai, Lingard describes this process as “vernacularization” (2000). But shifts to competency-based curricula have met with a backlash around the world in the form of the claim that it forces the “downgrading” of disciplinary knowledge at the expense of generic learning skills, with the effect of damaging students’ life chances (Rata, 2012; Young, 2007).

This claim has given rise to the curriculum debate over “powerful” versus “everyday” knowledge. Proponents of “powerful knowledge” posit that disciplinary or subject knowledge is “powerful” because it is conceptual, systematic, specialized and takes students beyond their personal experience. As such, they claim, it is distinctive from everyday knowledge and should be an entitlement of all learners because it offers opportunities for educational attainment and social mobility by providing a more reliable and objective foundation for understanding the world than knowledge arising from students’ own experiences (Young, 2014). On the other hand, opponents argue that powerful knowledge emphasizes the cognitive realm and neglects social justice issues (Zipin et al., 2015). Beck points to tensions around inducting “disadvantaged” students into an esoteric specialized epistemology founded on a minority conception of “high culture” (2013) and in a similar vein, Catling and Martin (2011) state that powerful knowledge neglects learners’ “ethno-knowledges” as a learning base for pluralist epistemologies. These discussions are enriched via Rudolf et al.’s thesis that powerful knowledge erases the power relations of colonialism (2018). Opportunities to consider the pros and cons of powerful knowledge arise later in the article in the case study application of the CCAF. Before introducing the case study, an introduction to the critical discourse studies (CDS) approach to curriculum is needed.

Curriculum from a critical discourse studies (CDS) perspective

As Flowerdew and Richardson (2017, p. 1) note:

CDS is an interdisciplinary approach to language in use which aims to advance our understanding of how discourse figures in social processes, social structures and social change.

Building on the field of critical discourse analysis (CDA) (Fairclough, 2003), CDS-oriented discourse analytic tools, such as our own, allow for more systematic attention to be paid to the macro and micro linguistic features of the textual articulations of curriculum. Following Fairclough, Hyatt writes:

... relationships of texts, practices and ideologies to social structures, processes and interrelationships should be seen as modalities of power and can be more systematically investigated (2013, p. 44).

CDS draws its linguistic foundations from Halliday’s (1985) systemic functional linguistics, which emphasizes the centrality of viewing language within the contexts in which it is produced and utilized. In doing so it offers affordances for the interrogation of the ways in which language and power interrelate, and the manner in which language is implicated in other social relations, processes and the social actors involved therein.

For us, the value of using a CDS-informed approach to the analysis of curriculum lies in its potential for conceptualizing the realization of curriculum as a site of struggle. The processes of construction and representation of curriculum are discursive, and so our approach seeks avenues by which such constructions can be deconstructed and then potentially reconstructed in more socially equitable forms. Following Gee (2000, p. 197), we view discourses as:

... socially and culturally formed, but historically changing, ways of talking and writing about, as well as acting with and toward, people and things. These ways are circulated and sustained within various texts, artefacts, images, social practices, and institutions, as well as in moment-to-moment social interactions. In turn, they cause certain perspectives and states of affairs to come to seem or be taken as normal or natural and others to seem or be taken as deviant or marginal.

CDS’s orientation to interdisciplinarity allows the analysis of curriculum to be approached in dialogue with other disciplinary lenses and frames. Its semiotic focus on construction and representation through language and other modes allows for the unclinking of such representations of social structures and events which themselves are mediated “through social practices—forms of social

activity, which constitute and, in turn, are constituted by, fields, institutions and organizations in a dialectical relationship” (Hyatt, 2013, p. 45). Such a nexus of social practices finds its articulation through discourses, and a CDS-aligned approach offers the curriculum analyst a systematic framework to uncover how curriculum articulations represent particular worldviews. As such it can orient those on whom curriculum policy impacts to specific understandings of relationships, identities and institutions—and to unmask the relations of power that mould and characterize these discursive constructions and representations through their co-option of language and other semiotic modes.

We contend therefore that our critically-oriented curriculum analysis framework can elucidate an engagement with curriculum texts, where such texts are viewed as the product of recurring contestations “between contenders of competing objectives, where language—or more specifically discourse—is used tactically” (Fulcher, 1989, p. 7). We next introduce the CCAF before putting our claim to work in the case of the Geography curriculum in England.

The critical curriculum analysis framework (CCAF)

For the purpose of this study, we understand curriculum as multi-faceted and dynamic: as text and/or process and/or discourse. It occupies spaces—of production, influence, recontextualization (from the policy field to the educational institution) and of enactment. Regardless of form or spaces of occupation, curriculum policy and curriculum *per se*² in our minds are always and already steeped in values and relations of power, and are subject to influences both beyond the national realm and in the processes of classroom enactment. The diagram below shows the constituent parts of the framework.

Before investigating the framework in more detail, it is important to explain its provenance. It has emerged from an amalgamation of three core approaches to education policy and curriculum analysis, namely:

- Rizvi and Lingard’s, 2010 key questions for policy analysis (pp. 54–56)
- Hyatt’s policy analysis framework using critical discourse analysis (2013) and
- Winter’s deconstructive approach to curriculum *per se* and curriculum policy analysis (2006, 2011, 2014, 2018).

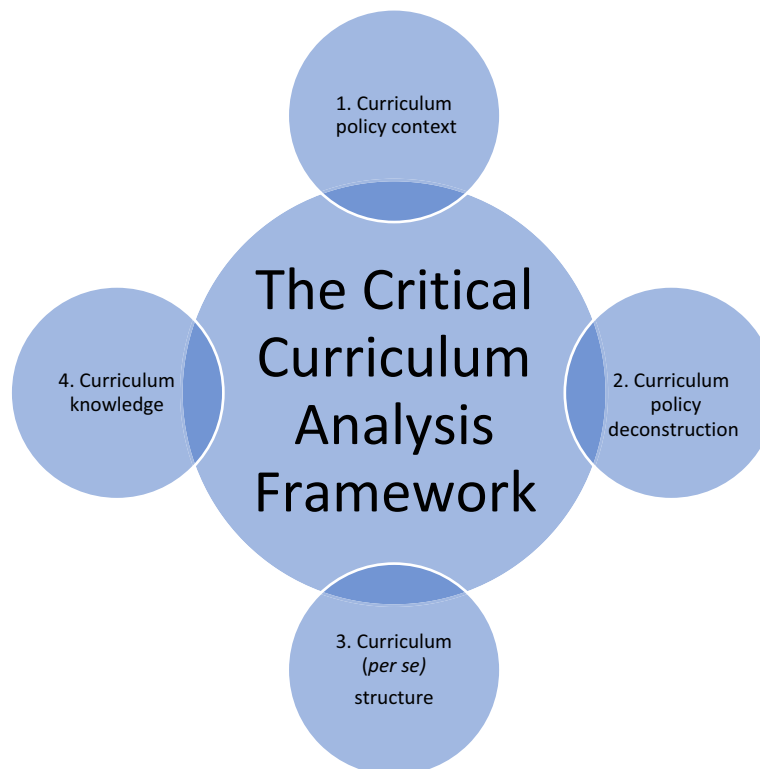


Fig. 1 The critical curriculum analysis framework.

²By the term “curriculum *per se*”, we refer to the materiality of the curriculum, for example, schemes of work, lesson plans, textbooks, teaching and assessment materials. These exist in textual, multi-modal or digital form. The process of translating curriculum policy into “curriculum *per se*” was described by Bernstein as “re-contextualization” (1990). Knowledge re-contextualization occurs through the pedagogic device, an analytical model to translate disciplinary knowledge into school knowledge (Singh, 2002, p. 572).

The first framework (Rizvi and Lingard, 2010) highlights policy origin, policy and textual issues and implementation, and policy outcome issues. These facets can all be applied to curriculum policy and curriculum *per se* analysis. The second (Hyatt, 2005, 2013) takes a semiotic route to investigate how authors of texts represent and thereby constitute the social world discursively, and the ideological power relations at work within and across these discourses. The final approach of Winter's work draws on a Derridean perspective aiming to disrupt and transgress the naturalized messages of curriculum policy and curriculum *per se* to open a space for something more socially just. The full framework can be found in Appendix 1, pp. 45–47. We acknowledge that the framework is incomplete in at least one significant aspect—that of curriculum enactment, in other words, the “diverse and complex ways in which sets of education policies are ‘made sense of’, mediated and struggled over, and sometimes ignored ... in schools”. Curriculum enactment involves “creative processes of interpretation and recontextualization” (Ball et al., 2012, p. 3), and we hope to extend the framework to encompass enactment, in collaboration with practitioners in the future.

In the section to follow, in order to demonstrate the potential of the framework, we use it to analyze a case study curriculum, the contemporary Geography General Certificate in Secondary Education (GCSE) in England.³ This analysis is necessarily limited in scope owing to the word restriction of the article, but its focus also demonstrates that the responsibility for selection of appropriate sections relies on the purpose and positionality of the analyst. For example, in the analysis to follow, a selection of sections of the CCAF was made for the analysis because of word restriction. In Fig. 1 (p. 38), the Curriculum policy deconstruction section, 2.1 Policy Archeology and 3.0 Curriculum *per se* structure were considered to be of lesser importance than other sections for the article and so were excluded from the analysis.

Applying the CCAF to a curriculum policy text ensemble: case study—geography GCSE in England

Introduction to the case study

The GCSE curriculum is usually taught to students aged 14–16 in English schools, during the educational phase known as Key Stage 4. While Geography (and History) are optional subjects for students at this stage, there is pressure on teachers to increase student recruitment to at least one of these subjects in order to comply with a performance indicator known as the E-Bacc or English Baccalaureate, which serves as one measure of quality of annual performance of schools. GCSE subject knowledge content is prescribed by the Department for Education (DfE), and passed on to the regulatory body, Ofqual which sets out the regulatory conditions and requirements for examination boards (or awarding bodies). These examination boards construct the curriculum specifications which must comply with Ofqual's conditions and requirements. In turn, schools are required to follow examination board specifications in order to prepare their students to engage with the curriculum and the associated examination. In the section below, we show how a selection of parts of the CCAF can be used to analyze a curriculum policy and curriculum *per se* text.

Curriculum context (p. 45)

Curriculum ensemble

For the purpose of this exercise, one curriculum policy text and two texts from one GCSE textbook have been selected as the “ensemble” for analysis. The policy text is *Educational Excellence Everywhere* White Paper (DfE, 2016) (Text 1) and the textbook is *GCSE Geography AQA*, written by Ross et al. (2016) (Texts 2 & 3). These three texts have been selected because their discursive analysis neatly illustrates the politics driving the configuration of curriculum policy knowledge and its manifestation in teaching materials. A wider range of texts could have been selected, including an earlier White Paper (DfE, 2010); the DfE subject knowledge content (DfE, 2014a), Ofqual Conditions and Requirements (2015); the examination board specification, examination question papers; mark schemes and examiners' reports, but word limit precludes these.

In order to address the curriculum context section of the framework below (1.2 *Curriculum drivers/warrants*, CCAF, p. 45), some research beyond the immediate documents under examination is necessary as will become evident.

Text 1 Educational Excellence Everywhere White Paper (DfE, 2016) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/508447/Educational_Excellence_Everywhere.pdf.

Curriculum drivers/warrants

Origin of this curriculum

The contemporary Geography GCSE curriculum has multiple origins, including the establishment of the subject 2200 years ago (Bonnett, 2008, p. 4) and its historical role in the establishment of the British Empire (Livingstone, 1992; Winter, 2022). The

³The secondary school Geography curriculum was chosen as the focus because of the research interests of one of the authors, a former Geography teacher who researches in this area.

contemporary configuration of school Geography in England links back to earlier curricula, for example, the first Geography National Curriculum in England (Department of Education and Science, 1991) which was described as a resurgence of “back to basics” neo-conservative knowledge by Ball (1994). Since then, global economic success has become a priority for Governments and curriculum reform identified as a means of achieving it:

What really matters is how we’re doing compared with our international competitors. That is what will define our economic growth and our country’s future. The truth is, at the moment we are standing still whilst others race past.

DfE (2010, para 11).

While prioritizing school assessment demands introduced by international performance tests like PISA, the British Government, since 2010, has borrowed curriculum knowledge policy from the cultural literacy movement in the US, led by E.D. Hirsch, to re-institute traditional academic knowledge into the contemporary Geography curriculum at GCSE level (Winter, 2014).

Why was this curriculum adopted (drivers) & why was it adopted at that time?

As a backlash to the 1997–2010 Labor Government’s focus on a “learning-to-learn” or a skills-based approach to pedagogy and an integrated approach to subject knowledge, the Coalition Government argued for the return to a traditional academic approach to the curriculum, this time, supported by Young’s “powerful knowledge” discourse (2008).

Who were the “players” (groups, interests, individuals) involved in establishing the contemporary geography GCSE curriculum agenda?

A wide range of policy-makers and influencers have participated in the reconstitution of Geography in schools over the last 30 years. These include Government Ministers, such as the Secretary of State for Education, Michael Gove (2010–2014), DfE, Ofqual, Geography teacher education and Geography researchers, officials of examination boards and subject associations, freelance consultants, textbook writers and publishers.

How were the policy drivers (warrants, justifications) disseminated to teachers?

A range of mechanisms was used to disseminate the policy drivers to teachers, including White Papers, news media, Government speeches, teacher union newsletters and subject associations, as well as academic articles and social media. Some of these authors promoted the curriculum reform policy, some critiqued it.

Text 1 Educational Excellence Everywhere White Paper (DfE, 2016, Ch 6, paras 6.1–6.24, pp. 88–93) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/508447/Educational_Excellence_Everywhere.pdf.

Curriculum policy deconstruction (p. 46)

Textual deconstruction

Section 2 of the CCAF, Curriculum Policy Deconstruction (2.2), will be used to guide this part of the analysis of the 2016 DfE White Paper (See **Appendix 1**, p. 46). The first question (2.2.1) in the CCAF is “How is this curriculum legitimated?”, in other words, what is its warrant, or the justification for its existence? Appeal to rationality is used in the policy document in the form of reference to outcomes, since the policy promotes certain outcomes with which it would be hard to disagree: “High expectations and a world-leading curriculum for all” (DfE, 2016, Ch 6, pp. 88–93), stating:

The best possible education for adult life in 21st century Britain is one that equips children and young people with the knowledge, skills, values, character traits and experiences that will help them to navigate a rapidly changing world with confidence (6.1).

Other accountability warrants are made throughout the policy document, for example, “world leading assessments and qualifications”, making these “as challenging as those in the highest performing countries around the world” (6.15), valorizing global competition, international benchmarks and performance tables.

Responses to the CCAF question (2.2.2, p. 46) “In what ways does the curriculum draw on discourses from outside education?” are found in the 2016 curriculum policy text (Interdiscursivity). The discourse of colonial celebration is promoted as the aspiration against a background of falling educational standards. This appears in the form of the good future story vs. the bad past story, or Legitimation via Mythopoesis. Evidence is found in the reference to knowledge about “the best which has been thought and said” (Arnold, 1869, p. 5), from a quotation by Matthew Arnold, a 19th century poet, cultural commentator and school inspector, who supported the cultural canon as the preferred guide to social development. The text also refers to a British-centered curriculum which

“balance(es) three Shakespeare plays and the study of a broad sweep of British history” (6.6), as well as “the development of the character traits and fundamental British Values that will help children succeed ... ” (6.1, e).

The second example of interdiscursivity occurs in the interpenetration of discourses of science and education in the policy document. Cognitive Science is “new” (“has shed fresh light”), is “ambitious” and “maximise(s) pupil understanding and minimize(s) confusion, giving teachers professional autonomy over how to teach” (6.6). Evocative evaluation is evaluation which on the surface appears ideationally neutral but which evokes a positive or negative reading (Martin, 2000). The text resonates with such evoked evaluation, with many references to “academic rigor”, “robustness”, and “a new gold standard”. The word “rigor” occurs six times. Likewise, discourses of modernity and entrepreneurialism occur. Both are presented as self-evidently “good”, for example, “a 21st century education that equips them [students] with the knowledge and character necessary for success in modern Britain”, (Inset box, p. 88), and “These new qualifications will command the confidence of students, employers, and further and higher education institutions” (6.22).

Text 1 Educational Excellence Everywhere White Paper (DfE, 2016, Ch 6, paras 6.1–6.24, pp. 88–93).

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/508447/Educational_Excellence_Everywhere.pdf.

Curriculum knowledge

The third section of the analysis focuses on the questions in Part 4 of the CCAF, *Curriculum knowledge* (Appendix 1, p. 47). The recurrence of the word “knowledge” (21 times) in the curriculum policy indicates that it forms a key underpinning concept (CCAF 4.1.2), and the knowledge promoted here is described as “a core body of knowledge” which is “set out” in a certain “format”, stating it is presented in a particular pre-determined way. References to: “the best that has been thought and written”; “the sort of knowledge-based, ambitious, academically rigorous education which every child should experience” and to the English Baccalaureate⁴ (EBacc) signal this “core body of knowledge” to be traditional academic knowledge. This statement is confirmed by the repetition, as noted above, of adjectives which serve as accountability warrants in the promotion of the curriculum, such as “robust”, “rigorous”, “ambitious”, “academically demanding”, “challenging”. All these characteristics allude to the notion of “powerful knowledge”, discussed in the context section of the article (p. 37). Participation in the competitive global education community is referenced: “national assessments and qualifications ... that ... are as challenging as those in the highest performing countries around the world” (6.15, p. 91). Two other key concepts (CCAF 4.1.2, p. 47) appear several times in this extract. The first is the equal opportunities discourse, evidenced in the statement:

We want every child, wherever they live and whatever their background or needs, to receive a 21st century education that equips them with the knowledge and character necessary for success in modern Britain.

Inset box, p. 88.

The second is the conceptualization of learning as “knowledge acquisition”, knowledge accumulation (“new knowledge is acquired through situating it within existing knowledge” and “the more a child knows the more they are capable of learning” (6.4, p. 89)). In summary, five key discourses have been identified in the curriculum policy extract: the knowledge-based curriculum, a core body of knowledge, the promotion of “quality” knowledge of a certain kind, equal opportunities/entitlement and a transmission/acquisition conception of learning. All, one, or a selection of these discourses can be analyzed to a deeper level in line with the four points of problematization contained within the CCAF 4.1.4: [i] underpinning assumptions, [ii] standpoint; [iii] responses and [iv] implications.

Text 2: Ross et al. (2016, p. 203) “Colonialism” Curriculum knowledge.

In this section, we present a critical analysis of the second text using the CCAF. The following paragraph about colonialism was published in a Geography GCSE textbook (Ross et al., 2016), currently used in schools under the heading of “Historical causes” [of uneven global development]:

... By the end of the 19th century much of Africa and parts of South America and Asia had been divided up between the European superpowers. Countries such as the UK, Germany, Spain and France had powerful empires and colonies. Since 1950 former European colonies have gained independence. In many cases this has been a difficult process, resulting in civil wars and political struggles for power. Money has been spent on armaments and some governments have been corrupt. This political instability has held back development.

Ross et al. (2016, p. 203).

⁴The EBacc is not an award or a qualification, but an accountability indicator which contributes to the assessment of performance by schools in terms of their ranking in performance tables, making it a high-stakes indicator. When first introduced, it consisted of a percentage of students achieving high GCSE grades in five traditional academic subjects. This later increased to eight subjects in 2013.

We focus on the CCAF questions found in 4.1, p. 47. The key discourse is colonialism (4.1.2) and the kind of knowledge (4.1.1) represented here fits the category of out-dated, inaccurate, evasive disciplinary knowledge, since it is telling “half a story” as Olusoga states (cited in Akbar, 2016). The history of colonialism is not only over-simplified in the paragraph (4.1.3), but is grossly and offensively mis-represented, raising the question “why?” Stuart Hall reminds us of:

... the profound historical forgetfulness—what I want to call the loss of historical memory, a kind of historical amnesia, a decisive mental repression—which has overtaken the British people about race and Empire since the 1950s. Paradoxically, it seems to me, the native, homegrown variety of racism begins with this attempt to wipe out and efface every trace of the colonial and imperial past.

Hall (1978, p. 26).

The assumption underpinning the paragraph is that following independence, former European colonies were unable to govern themselves and became violent, corrupt and politically unstable. In other words, the knowledge informing these assumptions is imperialist, colonial knowledge—driven by an assumption that colonial rule was a “good thing”. This represents the “single story” standpoint of colonialism (Adiche, 2009), told from the perspective of those holding colonial power, since the voices of the less powerful, the colonized, are silenced. The role of Europeans in colonial genocide, violence, dispossession, slavery and exploitation is missing, erased. Questions of who benefits and who loses out are relevant here, as the paragraph indicates that the reputation of the European powers remains intact, while calling into disrepute the moral character of those living in former colonial countries. This reductive binary of “coloniser = ‘good’/colonized = ‘bad’” indicates further deployment of evoked evaluation. This echoes through the curriculum in the form of the depiction of “High Income Countries” (HICs) as superior to “Low Income Countries” (LICs), highlighting the “West is Best” discourse. At the same time an implicit racialization is at work in the paragraph as well as in the curriculum overall, but never explicitly stated, that people racialized as white predominate in HICs and people racialized as black in LICs.

In relation to opportunities for inquiry into and critique of this paragraph by students, and gaining an understanding of student/teacher responses to it (4.1.4, [iii], p. 47), we need to examine this and other curriculum texts deployed in the classroom more closely and comprehensively, and conduct classroom studies to investigate the reactions and emotions that these discourses may generate among students and teachers. But, at the same time, it is important to acknowledge the authority and power of the textbook, and the over-crowded curriculum in a high-stakes assessment regime such as GCSE, which, combined, constrain the amount of inquiry and critique teachers can facilitate in English schools (Winter, under review). The fourth section, (4.1.4 [iv]) of the CCAF poses questions about the effect of such discourses about colonialism in terms of educational equality, student subjectivity and citizenship. Again, recourse to further research is needed, but the following questions may be considered: how do students and teachers racialized as black and those with heritages in former colonial countries feel? How do students and teachers racialized as white feel?

In the third and final text discussed in this article, the CCAF is applied to a textbook image associated with a popular concept found in school geography, the Demographic Transition Model (DTM). The analysis here follows Teo and Zhu (2018) approach which aims “to better understand and interpret how meaning is constructed and construed through the co-deployment of linguistic and visual semiotic resources”. This concept of the DTM is not currently a compulsory part of Department for Education Subject Content for Geography (2014) but does form an important part of the examination board specification (AQA, 2016), which teachers are expected to follow.

Text 3: Textbook image—multimodal analysis.

In geography, the Demographic Transition Model (DTM), “shows changes over time in the population of a country” according to the textbook under examination (Ross et al., 2016, p. 198). The DTM shows how, as a country moves in an assumedly linear path from “traditional” to “developed” state, its population characteristics change in stages in terms of birth and death rates as it “progresses” from stage to stage. According to the graph on the preceding page to the image below, countries in stage one of the model have high birth rates and high death rates (“traditional rain-forest tribes” in Indonesia, Brazil and Malaysia) and those in stage 5, the final stage, have very low birth and death rates (examples given are Italy, Germany, eastern European countries and Japan).

The following image is taken from the textbook, p. 199. It shows “Countries at different stages of development”. Questions in the CCAF, section 4, Curriculum Knowledge (p. 47) are drawn on in the analysis of this image.

Countries at different stages of development

Stage 1: Traditional rainforest tribes

In parts of Indonesia, Brazil and Malaysia, small groups of people live separately with little contact with the outside world. They have high birth and death rates (photo **B**).



B Rainforest tribe

Stage 2: Afghanistan

Afghanistan is one of the poorest and least developed countries in the world. Its birth rate is 38 per 1000 and its death rate is 14 per 1000. About 80 per cent of the population are farmers who often need children to support them in the fields and tending livestock.

Stage 3: Nigeria

Nigeria is a newly-emerging economy (NEE) experiencing economic growth. The death rate is much lower than birth rate (see page 220). The country's population is growing rapidly.



C Germany's ageing population

Stage 5: Germany

Germany is a well-developed country experiencing population decline as death rate exceeds birth rate. The birth rate is 9.5 per 1000, one of the lowest in the world. Women often have careers and few children. With an ageing population, Germany's death rate (11.2 per 1000) will continue to rise (photo **C**).

Stage 4: USA

The USA is one of the most developed countries in the world. Good-quality health care means death rates are low (8 per 1000). Women tend to have small families, choosing to study and follow careers. Therefore, birth rate is low (12 per 1000). Population growth is due mainly to immigration.

ACTIVITIES

- 1 Draw a sketch to show the DTM. Indicate on your sketch the links with economic development.
- 2 How does a falling birth rate reflect increased economic development?
- 3 How can an increasing death rate reflect high levels of development?
- 4 Consider the possible impacts for Germany of being in Stage 5 of the DTM. How might it have a negative effect on development?

Stretch yourself

Investigate the population and developmental characteristics of Afghanistan. How far do you agree that it is in Stage 2 with a low level of development?

Practice question

Explain how the demographic transition model is linked to economic development. (6 marks)

The kind of knowledge promoted in this image is traditional academic knowledge (4.1.1). The DTM is a long-established model used in the geography education community to teach about population change and global development. According to the textbook, the model is based on “changes that took place in Western countries such as the UK” (Ross et al., 2016, p. 198). While this is true in general terms, the model was conceived and developed in the USA (Szreter, 1993), where it was used in the weaponisation of population control in global anti-communist policy. In terms of simplification (4.1.3), three points stand out: the model focuses on birth and death rates only, ignoring migration as a factor of population change; the model’s complex political provenance is erased, and it alludes to the Malthusian idea of “disease, famine and war” as maintaining a steady population level in traditional societies, without critique. The DTM provides a Western population change process as a bench-mark which is used, unproblematically, to compare, categorize and “other” non-Western countries.

The perspectives of those without power, in other words, of those people living in non-Western countries, are not evident in the textbook discourse of the DTM (4.1.4 [ii]). The model promotes an unashamedly Western viewpoint. The binary at work here is the contrast between the “traditional” and the “developed”, with the authority of the assumedly universal application of the model left unquestioned. A close examination of the images reveals the construction of racialized difference and differential value in terms of level of “civilization” ascribed to social groups. This point is relevant to the CCAF question of promotion of equality (4.1.4 [iv]). The “stage 1” image (“Traditional rainforest tribes”) shows a family group, wearing traditional clothing which exposes bare skin, sitting in a relaxed way, on the floor outdoors in the rain forest, implying “primitive” surroundings. In contrast, the German senior citizen sits, fully clothed, upright on an IKEA-type chair, indoors. A degree of exoticization of the “rainforest tribe” is evident when contrasted with the image of the white European norm of the German woman, and the textual reference to the “tribes” isolation from “the outside world” reinforces this negative idea of the “tribes” being “left behind”. Indeed, this point recurs in the graph on p. 198, which indicates that the UK was at stage 1 (the stage of the “traditional tribes” in 2016) in the DTM pre-1760. The text attached to each image suggests the negativity of high natural population increase in “tribal” societies compared with the positivity of high natural population decrease in “civilized” societies.

In relation to the CCAF sections 4.1.4 [iii] “responses” and [iv] “implications” of this image, (p. 47), no opportunities are provided in the textbook for student critical reflection on the DTM or the image under scrutiny, and the nature of affective responses from students and teachers can only be imagined. The DTM discourse and image do not conjure a sense of equality between different social groups, but instead indicate the construction of difference, similar to that described by Willinski (1999), in his account of how dominant educational discourses serve to “divide the world”. The implicit message carried by the texts about relative superiority and inferiority and racialization may be engaged with by students and teachers or discarded as “business as usual”, or may create alienation and stigmatization for those racialized as black or having family heritage in Indonesia, Brazil and Malaysia. A third possibility is that students and teachers challenge the othering produced by the textbook—a response which represents an action of critical and responsible citizenship.

Conclusion

As argued earlier, curriculum plays a crucial and central role in education. This is because it not only involves a certain selection of knowledge, but also because the language of curriculum carries political power to constitute certain meanings and thereby influence the reader in certain ways. This is the case in both curriculum policy and in curriculum texts of all kinds. Curriculum is a political phenomenon. It exerts power by both propagating and obstructing, discursively, the configuration and circulation of “truths” (Foucault, 1972, p. 49). Since curriculum is such a powerful political phenomenon, critical analysis is necessary to excavate its political drivers, and to understand its ideological goals. This is particularly salient in an era of increasing control through educational neo-liberalization and securitization around the world. Several examples can be found in England. First, the National Curriculum reform (DfE, 2014b) increased the volume of prescribed knowledge content in the English school curriculum, stifling the creativity and professional autonomy of teachers as critical intellectuals, subject specialists and pedagogues, positioning them effectively as educational technicians. Second, recent policies require teachers to teach a “Fundamental British Values” curriculum, the origin of which lies in a counter-terrorism strategy (HM Government, 2011) which coerces teachers to operate as state security agents (Winter et al., 2021). Third, Relationships, Sex and Health curriculum policy (DfE, 2020) states that schools should not use materials produced by organizations that take “extreme political stances”, including “a publicly stated desire to abolish or overthrow democracy, capitalism ...”. In effect, this amounts to state censorship of sections of British school History, Sociology and Politics curricula (Busby, 2020). Fourth, the powerful knowledge movement, by canonizing core knowledge that represents “... respect for the best that has been thought and written” (Gove, 2012),⁵ restricts student and teacher accessibility to the full range of potential knowledge, authorizes certain prescribed configurations of knowledge as superior to others, while obscuring the provenance of that knowledge deemed to be powerful.

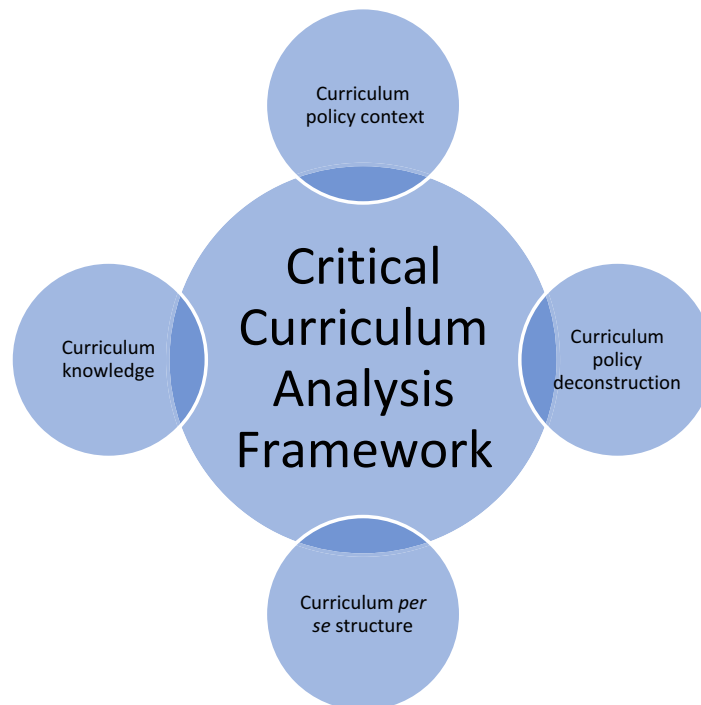
Each of these curriculum policies above constitute meanings in certain ways for students and teachers. At the same time, the underpinning high-stakes accountability regime holds curriculum compliance firmly in place, allowing little room for challenge and change. Teacher autonomy has been eroded under technologies of performance which Ball et al. describe as “deliverology”

⁵This quote is adapted from Matthew Arnold (1869) book *Culture and Anarchy*. The actual quote is “... the best which has been thought and said in the world” (Preface, p. 5).

aimed to turn students and teachers into “economically useful citizens” (2012, p. 139). The problem, however, extends beyond students, teachers and the school curriculum. According to Carr (1998), a dialectical relationship exists between the kind of curriculum we create and the kind of society in which we wish to live. Society creates the curriculum and the curriculum prepares students for life as citizens in society. If the goal of the education system is a pluralistic democratic society, then a pluralistic democratic curriculum is needed: democracy and curriculum “stand in a reciprocal relationship such that each provides the foundations on which the other is erected” (p. 337). The creation of a pluralistic democracy requires a curriculum in which reflective enquiry, critical intelligence, collective deliberation, collaborative social relations and the habit of resolving difficult practical, moral, political and social problems is exercised (See Carr, p. 336). An intelligent pluralistic democracy cannot be born out of a curriculum driven by coercion, competition, nationalism and political censorship. The Critical Curriculum Analysis Framework offers one approach by which students and educators might unravel and transgress the complex political discourses of curriculum policy and curriculum *per se* to construct a curriculum fit for an intelligent pluralistic democracy. Just as curriculum is a complex, dynamic and evolving phenomenon, so is this article.

Christine Winter acknowledges funding provided by The British Academy/Journal of Moral Education Trust and thanks are extended to Stephanie Borg, Gareth Bramley, Louise Kay, Georgina Spencer and Daniel Whittall for their constructive comments on an earlier draft.

Appendix 1



The critical curriculum analysis framework CCAF.

1 Curriculum context

1.1 Curriculum ensemble

1.1.1 What documents/texts make up the curriculum ensemble? Why were these documents/texts selected?

1.2 Curriculum drivers/warrants

1.2.1 Where did this curriculum originate, including consideration of any relevant global factors/institutions? Does it have links to earlier curricula?

1.2.2 Why was this curriculum adopted (drivers) & why was it adopted at that time?

1.2.3 Who were the “players” (institutions, groups, interests, individuals) involved in establishing the curriculum agenda?

1.2.4 How were the policy drivers (warrants, justifications) disseminated to teachers? Were the drivers hidden or explicit?

2 Curriculum policy deconstruction

2.1 Policy Archeology

- 2.1.1 What discourses frame the curriculum policy text? Are these globalized discourses?
- 2.1.2 What evidence is there here of interdiscursivity?
- 2.1.3 To which “problem” is the curriculum policy constructed as a solution?
- 2.1.4 How is the curriculum policy problem conceptualized? Will the curriculum policy as constructed “solve” the problem to which it is a response?
- 2.1.5 What alternative curriculum policy problem constructions have been rejected/neglected/not articulated, or actively considered & rejected or ignored?
- 2.1.6 How have the curriculum policy players constructed its context and/or history?
- 2.1.7 What complementary curriculum policies (i.e., policy levers, the mechanics of policy implementation, e.g., inspection, funding, metrics) were put in place to support the achievement of curriculum goals?
- 2.1.8 Does the curriculum policy contain vagueness and ambiguity? Is there recourse to “personal responsibility” in liberal values mode, with the purpose of avoiding official accountability?

2.2 Textual deconstruction

- 2.2.1 How is the curriculum legitimated? By an appeal to authority (legislation, etc. “we must do this”); by an appeal to rationality (“it’s the only sensible thing to do”); by an appeal to morality (“it’s the right thing to do”) or through narratives of potential success/failure if the policy is or isn’t enacted.
- 2.2.2 In what ways does the curriculum draw on discourses from outside education (economics, science, religion, etc.)? In what ways does the curriculum draw on other texts (citation, referencing, quoting, statistics, paraphrasing, etc.)? What are the impacts/implications of these discursive/textual borrowings?
- 2.2.3 How does the language used construct assertions as positive or negative? How overt/explicit or covert/implicit are these constructions?
- 2.2.4 What presupposition exist in the curriculum and how are they realized via the language?
- 2.2.5 What lexical and grammatical features of the curriculum construe certain readings of the text in certain ways—this may be applied to each of the categories above as they are all realized in language.
- 2.2.6 What do any images/tables/figures included in the curriculum imply/evoke in terms of the information imparted?

3 Curriculum *per se* structure

Teaching texts: curriculum specifications, textbook pages, test/examination questions, worksheets, teaching artifacts, online teaching materials, etc.

3.1 Curriculum construct-ors/-ion

- 3.1.1 Is the curriculum constructed from top-down (by curriculum experts, Governments) or bottom-up (by students, teachers, communities) or both or by another approach?
- 3.1.2 How was the curriculum introduced: quickly, with adequate teacher consultation, training and resources; in negotiation with teachers, students and communities?
- 3.1.3 Have research and curriculum precedents been used in support/justification/production of the curriculum? How credible is this evidence?
- 3.1.4 What is the relationship between the content of the official curriculum requirements and the teaching texts? What curriculum features stand out as different in the official curriculum text and teaching texts (“translation/recontextualization”)? What are the implications of these differences?

3.2 Curriculum structure

- 3.2.1 Is a technical (objectives/outcomes led) curriculum structure promoted?
- 3.2.2 What is the relationship between the curriculum structure (including knowledge), pedagogy and assessment?
- 3.2.3 How will the outcomes be measured/assessed/evaluated? What does this imply—exams might imply this is to be learned by rote and not engaged with critically?
- 3.2.4 Is the curriculum targeted toward a specific group of learners, e.g., “giftedness”/“learning disability”/“less able” or is it inclusive? What are the arguments supporting the targeting of curriculum to specific groups of learners?
- 3.2.5 Is the curriculum part of a high-stakes assessment regime?

4 Curriculum knowledge

4.1 Curriculum knowledge epistemology/ontology

- 4.1.1 What kind of knowledge is promoted in the curriculum—disciplinary (traditional academic subject knowledge; knowledge “what?”) or competency (skill-based knowledge; knowledge “how?”), vocational knowledge or a mix or something different?
- 4.1.2 What key concepts/ideas/discourses recur through the curriculum?
- 4.1.3 Are key concepts/ideas/discourses over-simplified? What elements of the concept/idea/discourse are over-simplified? Why?
- 4.1.4 For a selection of key concepts/ideas/discourses:
- [i] Underpinning assumptions:
 - What assumptions underpin the key concepts/ideas/discourses? (politics, ethics, values)
 - What knowledge informs them?—out-dated theory? Business theory? Public consciousness? Recent rigorous research evidence? What are the academic arguments and evidence in favor and against?
 - [ii] Standpoint:
 - Is the “single story” represented? Or is the concept/idea/discourse viewed from different perspectives ie those with power; those without power?
 - Whose viewpoint is being promoted?
 - What is missing from the key concepts/ideas/discourses? (historico-political background (e.g., colonialism); a global dimension; power relations; counter-arguments and narratives)?
 - Are binaries evident in the curriculum? What purpose do they serve? Do they operate masquerade/veneer to imply value-freeness/a-politicalness?
 - Are universal truths promoted? What purpose do they serve?
 - What curriculum knowledge is missing and why?
 - Who/what benefits and who loses out when the concept/idea/discourse is applied?
 - [iii] Responses:
 - Are teachers and students encouraged to be critically reflective about their own stances and lived experiences relating to the concepts/ideas/discourses?
 - Are concepts/ideas/discourses open to inquiry by students and teachers?
 - Are concepts/ideas/discourses presented in such a way to engender a feel-good response from dominant groups who engage with them? (What does the discourse make students and teachers feel?)
 - [iv] Implications:
 - Do the concepts/ideas/discourses promote equality?
 - What are the implications of the concepts/ideas/discourses for constituting student & teacher subjectivity?
 - What kind of citizen do the concepts/ideas/discourses promote?
 - Are there opportunities for students to create their own curriculum knowledge?

Andreotti, V., Stein, S., Sutherland, A., Pashby, K., Susler, R., Amsler, S., 2018. Gesturing Decolonial Futures Collective Policy and Practice: A Development Education Review Mobilizing Different Conversations About Global Justice in Education: Toward Alternative Futures in Uncertain Times. pp. 9–41.

Pashby, K., Sund, L., 2020. Teaching for Sustainable Development Through Ethical Global Issues Pedagogy: A Resource for Secondary Teachers. Manchester Metropolitan University. 27-03-2021.

References

- Adiche, C.N., 2009. The danger of a single story. In: TED Talk. https://www.ted.com/talks/chimamanda_ngozi_adichie_the_danger_of_a_single_story?language=en.
- Akbar, A., 2016. Interview: David Olusoga: “there’s a dark side to British history, and we saw a flash of it this summer”. Guardian. <https://www.theguardian.com/books/2016/nov/04/david-olusoga-interview-black-history>.
- Apple, M.W., 1996. Cultural Politics and Education. Open University Press, Bucks.
- Arnold, M., 1869. Culture and Anarchy. Smith, Elder and Co, London.
- AQA, 2016. GCSE Geography (8035) Specification. Version 1.0. Manchester UK.
- Ball, S.J., Maguire, M., Braun, A., 2012. How Schools Do Policy: Policy Enactments in Secondary Schools. Routledge, London.
- Ball, S.J., 1994. Education Reform: A Critical and Post-structural Approach. Open University Press, Bucks.
- Ball, S.J., 2008. The Education Debate. The Polity Press, Bristol.
- Beck, J., 2013. Powerful knowledge, esoteric knowledge, curriculum knowledge. Camb. J. Educ. 43 (2), 177–193.
- Bernstein, B., 1971. On the classification and framing of educational knowledge. In: Young, M.F.D. (Ed.), Knowledge and Control. Collier-Macmillan, London, pp. 47–69.
- Bernstein, B., 1990. The Structuring of Pedagogic Discourse. Routledge, London.
- Bonnett, A., 2008. What is Geography? Sage, London.
- Bowe, R., Ball, S., Gold, A., 1992. Reforming Education and Changing Schools. Routledge, London.
- Busby, M., 2020. Schools in England Told Not to Use Material From Anti-capitalist Groups. <https://www.theguardian.com/education/2020/sep/27/uk-schools-told-not-to-use-anti-capitalist-material-in-teaching-Guardian-Education-27-Sept>. (Accessed 14 March 2021).
- Carr, W., 1998. The curriculum in and for a democratic society. Curric. Stud. 6 (3), 323–340.
- Catling, S., Martin, F., 2011. Contesting powerful knowledge; the primary geography curriculum as an articulation between academic and children’s (ethno-)geographies. Curric. J. 22 (3), 317–335.

- Chun-Lok, F., Wing-Yan, Y., 2009. The politics of re-introducing liberal studies into Hong Kong secondary schools. *Educ. Res. Pol. Pract.* 9, 17–40.
- Department for Education, 2010. The Importance of Teaching Schools White Paper. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf. (Accessed 12 February 2021).
- Department for Education, 2014a. Geography GCSE Subject Content Crown Copyright. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/301253/GCSE_geography.pdf. (Accessed 12 February 2021).
- Department for Education, 2014b. Statutory Guidance: National Curriculum in England: Framework for Key Stages 1–4. <https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4>. (Accessed 18 March 2021).
- Department for Education, 2016. Educational Excellence Everywhere White Paper. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/508447/Educational_Excellence_Everywhere.pdf. (Accessed 12 February 2021).
- Department for Education, 2020. Guidance: Plan Your Relationships, Sex and Health Curriculum. Information to Help School Leaders Plan, Develop and Implement the New Statutory Curriculum. <https://www.gov.uk/guidance/plan-your-relationships-sex-and-health-curriculum>. (Accessed 14 March 2021).
- Department of Education and Science, 1991. Geography in the National Curriculum. HMSO, London.
- Fairclough, N., 2003. Analyzing Discourse and Text: Textual Analysis for Social Research. Routledge, London.
- Fitz, J.A., Nikolaidis, A.C., 2020. A democratic critique of scripted curriculum. *J. Curric. Stud.* 52 (2), 195–213.
- Flowerdew, J., Richardson, J., 2017. Introduction. In: *The Routledge Handbook of Critical Discourse Studies*, pp. 1–10.
- Foucault, M., 1972. *The Archaeology of Knowledge*. Vintage, New York.
- Fulcher, G., 1989. *Disabling Policies? A Comparative Approach to Education Policy and Disability*. Falmer Press, London.
- Gee, J.P., 2000. Discourses and sociocultural studies in reading. In: *Handbook of Reading Research*, vol. 3, pp. 195–207.
- Gerrard, J., Farrell, L., 2013. “Peopling” curriculum policy production: researching educational governance through institutional ethnography and Bourdieuan field analysis. *J. Educ. Pol.* 28 (1), 1–20.
- Ghebru, B., Llyod, M., 2020. From civic to citizenship education: towards a stronger citizenship orientation I the Ethiopian curriculum. *Int. J. Educ. Dev.* 72, 102143.
- Gleeson, J., Klenowski, V., Looney, A., 2020. Curriculum change in Australia and Ireland: a comparative study of recent reforms. *J. Curric. Stud.* 52 (4), 478–497.
- Goodlad, J.I., 1979. *Curriculum Inquiry: The Study of Curriculum Practice*. McGraw-Hill, New York.
- Gove, M., 2012. Letter from Michael Gove to Tim Oates Chair of the Expert Panel 110612. <https://www.bera.ac.uk/bera-in-the-news/background-to-michael-goves-response-to-the-report-of-the-expert-panel-for-the-national-curriculum-review-in-england>. (Accessed 14 December 2021).
- Hall, S., 1978. Racism and reaction. In: *Five Views of Multi-racial Britain*. Commission for Racial Equality and BBC, London, pp. 23–35.
- Halliday, M.A.K., 1985. *An Introduction to Functional Grammar*. Edward Arnold, London.
- HM Government, 2011. Prevent Strategy. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/97976/prevent-strategy-review.pdf. (Accessed 1 March 2021).
- Hyatt, D., 2005. A critical literacy frame for UK secondary education contexts. *Engl. Educ.* 39 (1), 43–59.
- Hyatt, D., 2013. The critical higher education policy discourse analysis framework. In: Huisman, J., Tight, M. (Eds.), *Theory and Method in Higher Education Research*. Routledge, London, pp. 41–59.
- Kelly, V., 2009. *The Curriculum: Theory and Practice*, sixth ed. Sage, London.
- Lingard, B., Martino, W., Rezaei-Rashi, G., 2013. Testing regimes, accountability and education policy: commensurate global and national developments. *J. Educ. Pol.* 28, 539–556.
- Lingard, B., 2000. It is and it isn't: vernacular globalization, educational policy and restructuring. In: Burbules, N.C., Torres, C.A. (Eds.), *Globalization and Education: Critical Perspectives*. Routledge, London.
- Lingard, B., 2006. Globalization, the research imagination and deparochializing the study of education. *Glob. Soc. Educ.* 3 (2), 287–302.
- Livingstone, D.N., 1992. *The geographical tradition*. Blackwell, Oxford.
- Macedo, E., 2011. Curricula policies in Brazil: the citizenship discourse. In: Yates, L., Grunet, M. (Eds.), *The World Yearbook of Education: Curriculum in Today's World: Configuring Knowledge, Identities, Work and Politics*. Routledge, New York, pp. 44–57.
- Martin, J.R., 2000. Beyond exchange: APPRAISAL systems in English. In: Hunston, S., Thompson, G. (Eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford University Press, Oxford.
- Nomura, K., 2017. Project-based learning in post-World War II Japanese school curriculum: an analysis via curriculum orientations. *Curric. J.* 28 (4), 626–641.
- Ofqual, 2015. GCSE Subject Level Conditions and Requirements for Geography 15/5629. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/410512/gcse-subject-level-conditions-and-requirements-for-geography.pdf. (Accessed 12 February 2021).
- Priestley, M., Sinnema, C., 2014. Downgraded curriculum? An analysis of knowledge in new curricula in Scotland and New Zealand. *Curric. J.* 25 (1), 50–75.
- Rata, E., 2012. *The Politics of Knowledge in Education*. Routledge, London.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Roofe, C., 2020. Curriculum as place: the social relevance of the secondary school curriculum in Jamaica. *Curric. J.* 32 (2), 165–181.
- Ross, S., Rowles, N., Holmes, D., Digby, B. (Eds.), 2016. *GCSE Geography AQA*. Oxford University Press, Oxford.
- Rudolf, S., Sriprakash, A., Gerrard, J., 2018. Knowledge and racial violence: the shine and the shadow of “powerful knowledge”. *Ethics Educ.* 13 (1), 22–38.
- Sahlberg, P., 2011. *Finnish Lessons*. Columbia University: Teachers College Press.
- Sahlberg, P., 2015. *Finnish Lessons 2.0*. Columbia University: Teachers College Press.
- Savage, G., 2017. Neoliberalism, education and curriculum. In: Savage, G. (Ed.), *Powers of Curriculum*, pp. 143–165.
- Schools Council/Nuffield, 1970. *The Humanities Project: An Introduction*. Heinemann, London.
- Singh, P., Thomas, S., Harris, J., 2013. Recontextualizing policy discourses: a Bernsteinian perspective on policy interpretation, translation, enactment. *J. Educ. Pol.* 28 (4), 465–480.
- Singh, P., 2002. Pedagogizing knowledge: Bernstein's theory of the pedagogic device. *Br. J. Sociol. Educ.* 23 (4), 571–582.
- Szreter, S., 1993. The idea of demographic transition and the study of fertility change: a critical and intellectual history. *Popul. Dev. Rev.* 19 (4), 659–701.
- Teo, P., Zhu, J., 2018. A multimodal analysis of affect and attitude education in China's English textbooks. *Multimodal Commun.* 7 (1).
- Todd, S., 2001. “Bringing more than I contain”: ethics, curriculum and the pedagogical demand for altered egos. *J. Curric. Stud.* 33, 431–450.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, Chicago.
- Willinski, J., 1999. *Learning to Divide the World: Education at Empire's End*. University of Minnesota Press, Minnesota.
- Winter, C., Heath-Kelly, C., Kaleem, A., Mills, C., 2021. A moral education? British values, colour-blindness, and preventing terrorism. *Crit. Soc. Policy*.
- Winter, C., 2022. The White Geography GCSE Curriculum in England: A Curriculum of Deceit.
- Winter, C., 2006. Doing justice to geography in the secondary school: deconstruction, invention and the National Curriculum. *Br. J. Educ. Stud.* 54 (2), 212–229.
- Winter, C., 2011. Curriculum knowledge and justice: content, competence and concept. Whiteness and Education Available online: <https://www.tandfonline.com/doi/full/10.1080/23793406.2022.2132179>.
- Winter, C., 2014. Curriculum knowledge, justice, relations: the schools' white paper (2010) in England. *J. Philos. Educ.* 48 (2), 276–292.
- Winter, C., 2018. Disrupting colonial discourses in the geography curriculum during the introduction of British values policy in schools. *J. Curric. Stud.* 50 (4), 456–475.
- Young, M., 2007. *Bringing Knowledge Back in: From Social Constructivism to Realism in the Sociology of the Curriculum*. Routledge, London.
- Young, M.D.F., 2014. *Knowledge and the Future School*. Bloomsbury, London.
- Zipin, L., Fataar, A., Brennan, M., 2015. Can social realism do social justice? Debating the warrants for curriculum knowledge selection. *Educ. Soc. Change* 19 (2), 9–36.

Curriculum and 21st century skills

Joke M. Voogt and Natalie N. Pareja Roblin, University of Amsterdam, Department of Child Development and Education, Amsterdam, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

21st century skill in the curriculum	49
Conceptualizing 21st century skills	50
Position in the curriculum	50
Macro level: 21st century skills in curriculum frameworks and standards	51
21st century skills in school curricula	52
Implementation in the curriculum: prominent issues	52
Assessment	52
Capacity building	53
Time lag	53
References	53
Further reading	55
Relevant websites	55

21st century skill in the curriculum

Due to technological developments, societies all around the world are rapidly changing with major implications for the ways we live, work and learn. Several scholars acknowledge the impact of these changes on the nature and types of jobs (e.g., [Levy and Murnane, 2006](#); [Van Laar et al., 2019](#)). Others emphasize the increasing importance of ideas and knowledge as the new commodity in a technology-driven society (e.g., [Anderson, 2008](#)). Further, the need for a normative re-orientation on human values such as privacy and equity is called for (e.g., [Duff, 2008](#); [Baeva et al., 2021](#)), because of technological developments such as machine learning and the use of algorithms ([Webb et al., 2020](#)). These developments require a reconsideration of the competences people need to be able to actively contribute to a technology-driven society (cf. [Ananiadou and Claro, 2009](#); [Gordon et al., 2009](#)). Therefore, a key curriculum issue in many countries nowadays is what the implications of these changes are for what needs to be taught and learned in schools.

The call for 21st century skills in the curriculum is a manifestation of the concerns diverse stakeholders in society have about the current relevance of the formal education curriculum in times of rapid changes in society. Curriculum can be defined as design for learning ([Van den Akker and Nieveen, 2021](#)) and includes decisions about objectives, content and organization of what is worth teaching and learning ([Walker, 2003](#)). Taking decisions about what is worth teaching and learning in the 21st century is not simple. It is always the outcome of a social-political process, in which views on the purposes of the curriculum in society are discussed among stakeholders and turned into concrete goals and content for teaching and learning. Part of this process is to find consensus about three key functions of education in society: qualification, socialization and personal development (or subjectivation in terms of Biesta) ([Biesta, 2013](#)). Qualification refers to the knowledge, skills and understanding we need to equip students with in order to prepare them for contributing to society in their later life. Socialization is the preparation of students to internalize a society's culture, including its norms and values. Finally, through the curriculum, education aims to contribute to the personal development of each individual learner.

The curriculum as design for learning ([Van den Akker and Nieveen, 2021](#)) becomes tangible in the form of curriculum artifacts (materials, programs), which are prepared for use at different educational levels. The most well known levels are the macro level (the curriculum of a jurisdiction, usually a nation, a state or province), the meso level (the curriculum of a school, a school board or a district) and the micro level (the classroom curriculum). Because of globalization and an increased need for lifelong learning two levels have been added recently: the supra level (curriculum frameworks proposed by supranational organizations such as UNESCO, the OECD and the European Union) and the nano level to be able to pay attention to individual learners (the curriculum to personalize learning) ([Van den Akker, 2003](#)). It is important to realize that these five curriculum levels do not form a hierarchy, but interact with each other. Curriculum materials describe the intentions of what needs to be taught and learned in formal curriculum documents—the intended curriculum. But the curriculum also becomes manifest in two other representations: the implemented curriculum shows how the curriculum is implemented in the practice of teaching and learning in schools and classrooms. Finally, the attained curriculum reflects what students actually experience and learn ([Van den Akker, 2003](#)).

Conceptualizing 21st century skills

In many countries the discussion about what needs to be taught and learned in a technology-driven society started outside the educational arena. In the first decade of the 21st century supranational organizations and public-private partnerships fueled the discourse. The DeSeCo project (the Definition and Selection of Competencies) initiated by the Organization of Economic Cooperation and Development (OECD) was among the first projects that formulated lifelong learning competencies that were considered relevant in a technology-driven society (Rychen and Salganik, 2001). Similarly, the European Union formulated key competences (Gordon et al., 2009). Hard- and software companies were particularly active in public-private partnerships that aimed to formulate what 21st century skills imply for school curricula and how these skills can be implemented in the practice of teaching and learning. Examples of such partnerships were the Partnership for 21st century skills, which was mostly active in the US (Partnership for 21st Century Skills, 2009) and the Assessment and Teaching of 21st Century Skills (ACT21S), which operated at an international level (Binkley et al., 2012). More recently the OECD launched “The future of Education and Skills—Education 2030” project, which aims to update the lifelong learning competencies formulated in the DeSeCo project. This project places particular emphasis on the importance of individual and collective wellbeing as an important aspect of reconsidering the curriculum with a view to society in the future (OECD, 2018). These various initiatives resulted in concrete definitions and frameworks, which describe and justify the need for generic goals that are important for 21st century curriculum.

While in the US the term “21st century skills” is mostly used for addressing the goals that need attention in the 21st century curriculum, the term competency (e.g., lifelong learning competency and key competency) became more popular in Europe. Despite these different terms, there is consensus that the general features of these learning goals: (1) are generic and transversal, which means that they are relevant across many domains; (2) not only refer to skills, but also include knowledge and attitudes; and (3) are context-specific as they prepare learners to cope with complex problems with wicked tendencies (e.g., Gordon et al., 2009; Veltman et al., 2019). Klieme et al. (2008) use the term disposition instead of skill or competency to refer to the learning goals relevant in the 21st century. They define these as “context-specific cognitive dispositions that are acquired by learning and needed to successfully cope with certain situations or tasks in specific domains” (Klieme et al., 2008, p. 9). Their definition refers not only to cognitive components, but includes also behavioral, motivational, social and ethical, aspects. Because the term 21st century skills is internationally used, in this contribution we adhere to this term.

Despite consensus in the overall features of 21st century skills, the different frameworks use different language and different concepts to describe these skills. Voogt and Pareja Roblin (2012) analyzed eight frameworks (including the ones described above) leading to a set of skills that were considered core in almost all frameworks. These are: cooperation, communication, social and cultural skills, problem solving, critical thinking, digital literacy, creativity and self-regulation. From this list it is clear that most 21st century skills are by no means new. For instance, already in 1910 Dewey (1910) mentioned critical thinking as an important skill that education should pay attention to. And although digital literacy may seem new, it has been discussed as educational goal in many countries already in the 1990s (Plomp et al., 1996). While many of these skills are not new, the call for explicit attention to 21st century skills in the curriculum was, in part, a reaction to the narrow emphasis on basic skills and high stakes testing in many countries (Au, 2011; Bryderup et al., 2009), which left little room for the development of the generic skills that are needed to become flexible lifelong learners in a society that is rapidly changing.

It must be noted that the call for 21st century skills in formal education is also disputed. Hilt et al. (2019) argue that the call for 21st century skills takes a neo-liberalistic stance on education. Biesta (2013) further argues that an uncritical acceptance of 21st century skills in the curriculum represents a narrow view on the purposes of education. He warns that a focus on 21st century skills seems to emphasize the qualification (i.e., to prepare students for living and working in a technology-driven society) and the socialization functions of education (i.e., taken the technology-driven society for granted without a critical stance) at the cost of personal development. Probably as a response to the latter critique, in recent developments values and student well-being have become much more at the forefront of the discussion around the question of what needs to be learned in order to prepare students for a future society (OECD, 2018).

Position in the curriculum

Notwithstanding agreement on what can be understood by 21st century skills and why they are important, the position of these skills in the formal curriculum and their realization in curriculum practice is often the subject of debate. One of the reasons is that 21st century skills cannot easily be separated from each other and are often overlapping, which makes it difficult to formulate unambiguous learning goals and to assess them. Another reason is the fear that these skills become an extra add-on to the curriculum, which may result in curriculum overload. We briefly illustrate these points using two skills: critical thinking and digital literacy.

Critical thinking can be approached from three perspectives: the philosophical, the psychological and the critical pedagogical perspective. Central in the philosophical perspective is the formation of a reasonable and just judgment, whereas the psychological perspective emphasizes higher order thinking skills, and the critical pedagogical perspective mainly pays attention to effects of argumentation and reasoning in the social context (Volman and Ten Dam, 2015). To be able to formulate concrete learning goals, which can be taught and assessed it is important to provide a clear definition in which these three perspectives are incorporated. For this reason, and after a broad consultation with many experts, the American Psychological Association (APA) published a definition in

which critical thinking is seen as goal-oriented and the ability to give an independent judgment. They formulated six cognitive skills (interpretation, analysis, evaluation, inference, explanation and self-regulation) and nineteen dispositions (e.g., clarity, orderliness, reasonableness, persistence, trust, self-confidence, flexibility, open-mindedness, willingness to reconsider) that are required to develop critical thinking (Falcione, 1990). This broad definition of critical thinking clearly demonstrates that it is difficult to isolate critical thinking from other 21st century skills, such as for instance problem solving, creativity, social and cultural skills and communication. The definition of the APA suggests that critical thinking can be understood as a set of cognitive skills and dispositions that can be applied in various disciplines (Ennis, 2018). Although the definition was meant to give guidance to teaching and learning, it remains a challenge to incorporate this broad understanding of critical thinking in curriculum design. A prominent discussion about generic skills, such as critical thinking, is to what extent they are domain specific. Scholars such as Ennis (2018) argue that critical thinking as a generic skill enables transfer to various contexts, although research shows that this is not self-evident (Lazonder and Egberink, 2014). Ennis (2018) argues that some substantive knowledge is necessary but not crucial in teaching critical thinking skills. Others argue that domain knowledge is essential for being able to apply critical thinking skills (and comparable skills such as problem solving and creativity) at an in-depth level (McPeck, 1981). In this view, critical thinking skills should be given a domain-specific interpretation, as is the case, for example, with historical reasoning (Stoel et al., 2017), the teaching of science inquiry in primary school science (Kruit et al., 2018), the thinking of the architect (Hamel, 1990), or the clinical reasoning by doctors (Norman, 2005).

Digital literacy is not a simple technical ability but a complex skill based on the integration of (meta)cognitive and technological skills (Aesaert et al., 2014) and linked to motor, emotional and social skills. It is more than just operating digital technology. The aim is to be able to understand the impact of digital technology on life and work and to be able to use digital technology for various purposes that are relevant in society and for personal use, such as: problem solving, creating information, providing information/ accessing information, communication, cooperation and safe and responsible use of the internet. Current digital literacy frameworks and curricula mainly emphasize problem solving, information acquisition and processing, critical thinking and creative and innovative use of digital technology (Siddiq et al., 2016). A well-known framework is the European DigComp reference framework (Ferrari, 2013), which structures digital technology skills in five categories: information, communication, creation, safety and problem solving. These categories show overlap with other 21st century skills. Many countries worldwide have formulated policies for integrating digital literacy in the curriculum for compulsory education (Eickelmann, 2018; UNESCO, 2016) and have started the implementation process. In few countries digital literacy is a separate subject in the curriculum, such as England. However, in many countries the goals for digital literacy are integrated in traditional school subjects (e.g., Canada, BC, Germany NRW, Belgium, Australia) also because countries try to avoid extra add-ons to the curriculum, as this may result in feelings of curriculum overload of teachers. However, integration of digital literacy in other subjects requires teachers who are qualified to teach digital literacy and see the relevance for their subject (Voogt and Brummelhuis, 2014).

These opposite views, as illustrated by critical thinking skills and digital literacy, fuel the discourse on where to situate 21st century skills in the curriculum. Should these skills become integrated across subject domains, do they require a separate subject in the curriculum, or do they need to be taught in cross-curricular projects? The two examples show that two different types of considerations are at stake. At the one hand the answer to this question is guided by the discussion on the nature of the skills (i.e., whether they generic or domain-specific). On the other hand the answer to this question is also driven by concerns about curriculum expansion. Although it is generally acknowledged that the curriculum needs to respond to developments in society, expanding the curriculum with new content also implies less time for traditional content.

Macro level: 21st century skills in curriculum frameworks and standards

Countries all around the world have been responding to calls from society to renew curriculum frameworks and standards in primary and secondary education. The attention for 21st century skills has a prominent place in these calls. Yet, countries are at different stages in the curriculum renewal process and they differ significantly in the way they handle it (Voogt and Nieveen, 2017). While some countries are in the middle of the curriculum renewal process itself (e.g., The Netherlands, Wales), other countries have formulated frameworks, set new standards and are implementing the new frameworks and standards in schools (e.g., Singapore, New Zealand, Canada BC, Scotland). Some countries (e.g., England, Germany NRW) have renewed parts of their curriculum by, for instance, adding subjects or adapting standards for specific subjects (e.g., digital literacy), while other countries have chosen to review and renew the complete primary and secondary school curriculum. There are countries that periodically renew the curriculum, such as Finland and Norway, who do this every ten years (Pietarinen et al., 2017; Siveskind and Westbury, 2016). However, in many countries (e.g., The Netherlands, Ireland) curriculum renewal is an ad hoc process (e.g., FitzPatrick and O'Shea, 2013; Kuiper et al., 2013), often fueled by a shared feeling of urgency (Barber et al., 2010). The different approaches toward national or state curriculum reform affect the socio-political process of adoption of 21st century skills in the curriculum.

The curriculum renewal process usually starts with creating awareness, engaging stakeholders, identifying the needs for a new curriculum and a political decision to formally start the renewal process. For instance, the process in New Zealand started with an analysis of the social, economic and technological changes taking place in society, a discussion on the changing meaning of knowledge due to technological developments, and the observation that the learning sciences offer increasing understanding on how people learn. This analysis led to the argument that schools need to contribute to the development of competencies that enable students to contribute to the wicked problems of the 21st century (Sinnema, 2011). These discussions at the country level

are often fed by the supranational initiatives as described earlier. It is not always the governments who take the initiative to start the curriculum renewal process. In England the Royal Society started the discussion about the need for Computing in the curriculum (The Royal Society, 2012). In the Netherlands advisory councils to the government initiated the process of curriculum renewal with a discussion about the position of technology in society and the potential for learning (e.g., [Wetenschappelijke Raad voor het Regeringsbeleid, 2013](#)). Usually many different stakeholders are involved in the awareness process, from within the educational arena (e.g., teachers and teacher unions, school boards), the community (e.g., parents' associations) and the private sector (Law, 2014; Pietarinen et al., 2017). Technology is increasingly used to inform stakeholders about the content and process (e.g., Website; twitter), and to promote stakeholder engagement (e.g., discussion forums) (Nieveen et al., 2014). For a system-wide impact on school and classroom practices, it is important to have a shared vision about the needs for and affordances of 21st century skills in the curriculum as a starting point (Van den Akker and Nieveen, 2021). Such a shared vision provides a rationale for specific decisions that need to be made during the design and development of curriculum frameworks and standards for a 21st century curriculum.

21st century skills in school curricula

Although 21st century skills are often not yet systematically integrated in national and school curricula, schools and teachers are increasingly paying attention to these skills. Voogt and Pareja Roblin (2012) distinguished three ways of integrating 21st century skills in the school curriculum: as new subjects or new content in existing subject domains, as cross-curricular competences acquired in specific projects or through curriculum transformation. All three approaches are found in studies of schools that put efforts in the implementation of (a subset of) the 21st century skills in curriculum practice. Tan et al. (2017a) studied Singaporean schools that integrated some of the 21st century skills. They found examples of schools that implemented a subset of the skills in single subjects, such as math, science, arts and physical education. Also schools were found that implemented 21st century skills across the curriculum. For example a school that infused the arts throughout the school curriculum, and a school that emphasized globalization and practiced 21st century skills in an exchange with schools in the US. The 21st Century skills were also an important component of the educational concept of innovative school's in the Netherlands (Volman et al., 2020). This study investigated how self-regulation, collaboration and creativity were embedded in the curriculum. In schools that had just started to implement 21st century skills in their curriculum, the skills were seen as an add-on to the traditional curriculum. However, schools where 21st century skills were part of the school's mission for quite some time a learner-centered approach was adopted and project-based and inquiry learning was emphasized, indicating a cross curricular approach. This study showed that teaching 21st century skills in the curriculum positively correlated with students' perception that they mastered them. In an international case study of innovative pedagogical practices using technology from 28 countries, Voogt and Pelgrum (2005) found an emphasis on fostering 21st century skills through the use of technology. They observed three foci in the way technology was used: single subject, thematic and school-wide. In all three foci 21st century skills were addressed, but in a different manner. In all three foci students and teachers were positive about the innovative practices, but evidence on the impact of the 21st century skills on student outcomes were lacking. In the Single Subject Curricular Focus technology was used to teach traditional content at a deeper level, emphasizing problem solving and critical thinking skills. The innovative technology-based practices in the Thematic Curricular Focus explicitly aimed at developing 21st century skills, such as problem solving, digital literacy and collaboration. This was accomplished through cross curricular project-based or thematic learning activities that were organized in addition to a subject based curriculum. In the School-wide Curricular Focus the innovative technology-based practices were embedded throughout the school curriculum and were part of the school's mission. These schools had the view that students are responsible for their own learning, and emphasized self-regulation skills while also paying attention to the other 21st century skills. However, although (innovative) schools are actively integrating 21st century skills in their curriculum, the implementation in mainstream education is often challenging. In the next section we discuss some prominent issues that affect the implementation of 21st century skills.

Implementation in the curriculum: prominent issues

Assessment

Assessment requirements often steer the way the curriculum is being implemented. Several studies showed that lack of alignment between the intention to teach 21st century skills and assessment practices is a threat to the implementation of 21st century skills (e.g., Au, 2011; Sinnema, 2011; Tan et al., 2017b). To overcome this problem assessment practices need to be developed that allow the assessment of 21st century skills in an authentic context. Webb and Gibson (2015) argue that the affordances of technology makes it easier to assess 21st century skills, because with technology it becomes possible to embed the assessment of 21st century skills in authentic performance tasks. An example is the OECD PISA study of the assessment of collaborative problem solving skills (OECD, 2017). Building on the work of Hesse et al. (2015), who developed an assessment framework for assessing collaborative problem solving skills, the OECD PISA study developed a computerized environment in which students had to interact with virtual agents to solve problems. This approach has proven to be a valid and reliable way of assessing such skills (Stadler et al., 2020). Another example is the International Computer and Information Literacy Study (ICILS). In this study aspects of digital literacy are assessed in a computerized environment. Students have to complete both closed tasks (answering a question. Clicking a link) and authentic open tasks (e.g., creating an information product) (Fraillon et al., 2020).

Capacity building

The implementation of a curriculum that incorporates 21st century skills usually requires attention and time for capacity building at all levels of the school system (governance, leadership, teachers) (Desha et al., 2009). Leadership continuity is essential in the implementation of complex change in schools, such as the implementation of 21st century skills in teaching and learning. Leaders need to trust their teachers and provide a supportive school culture for the change (UNESCO, 2010). Teachers need to be qualified to teach 21st century skills and understand what is expected from them in order to prevent misunderstandings. For example, in New Zealand teachers had difficulty in seeing the relation between the new curriculum that contained standards about 21st century skills and subject-related standards. They needed support to think in-depth about the connections between the two parts of the standards and to prevent that they perceived 21st century skills as an add-on to the curriculum (Hipkins, 2016).

At the school and teacher level guiding materials can help to exemplify how the curriculum can be realized in school and classroom practice (McAra et al., 2013). Several studies show the importance of a middle layer, local or private agencies which act as a broker between the requirements set forward in national standards and the possibilities and constraints of the local context. Such agencies can help leadership and teachers in developing local interpretations of the curriculum (Toh et al., 2016; Pietarinen et al., 2019).

Time lag

While transformations in society are changing fast, the process of curriculum development is usually slow. It might take 10–15 years before we can see effects of attention for 21st century skills in society, as this is the period it usually takes from the start of the curriculum development process until students equipped with 21st century skills leave the school system (Desha et al., 2009). This implies that we need to prepare students now for a future that is unknown (Hilt et al., 2019). It is too simple to state that because 21st Century skills are generic skills we may expect that these skills are not only relevant today but also in the future. Changes in society often imply a need to review the rationale for teaching the skills as well as relevant content for teaching the skills. Moreover, skills such as digital literacy need constant adaptation because of technological developments. The need to update the lifelong learning competencies formulated by the OECD in 2005 and in 2018 is just an illustration of the uncertainties that are inherent to the development of a curriculum for the future.

References

- Aesaert, K., Van Nijlen, D., Vanderlinde, R., Van Braak, J., 2014. Direct measures of digital information processing and communication skills in primary education: using item response theory for the development and validation of an ICT competence scale. *Comput. Educ.* 76, 168–181.
- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries. Organization for Economic Cooperation and Development. EDUWorking paper no. 41. OECD, Paris.
- Anderson, R.E., 2008. Implications of the information and knowledge society for education. In: Voogt, J., Knezek, G. (Eds.), *International Handbook of Information Technology in Primary and Secondary Education*. Springer, New York, pp. 5–22.
- Au, W., 2011. Teaching under the new Taylorism: high-stakes testing and the standardization of the 21st century curriculum. *J. Curric. Stud.* 43 (1), 25–45.
- Baeva, L.V., Khrapov, S.A., Azhmukhamedov, I.M., 2021. "Smart technologies" in education: development opportunities and threats. In: Popkova, E.G., Sergi, B.S. (Eds.), "Smart Technologies" for Society, State and Economy. ISC 2020. Lecture Notes in Networks and Systems, vol. 155. Springer, Cham.
- Barber, M., Chijioke, C., Mourshed, M., 2010. How the World's Most Improved School Systems Keep Getting Better. McKinsey & Company, New York.
- Biesta, G.J.J., 2013. Responsive or responsible? Education for the global networked society. *Pol. Futures Educ. Internet* 11, 734–745.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Rumble, M., 2012. Defining 21st century skills. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, New York, pp. 17–66.
- Bryderup, I., Larson, A., Trentel, M.Q., 2009. ICT-use, educational policy and changes in pedagogical paradigms in compulsory education in Denmark: from a lifelong learning paradigm to a traditional paradigm? *Educ. Inf. Technol.* 14, 365–379.
- Desha, C.J., Hargroves, K., Smith, M.H., 2009. Addressing the time lag dilemma in curriculum renewal towards engineering education for sustainable development. *Int. J. Sustain. High Educ.* 10 (2), 184–199.
- Dewey, J., 1910. *How We Think*. D.C. Heath & Co, Lexington, MA.
- Duff, A., 2008. The normative crisis of the information society. *Cyberpsychology* 2 (1).
- Eickelmann, B., 2018. Cross-national policies in information and communication technology in primary and secondary schools: an international perspective. In: Voogt, G., Knezek, G., Christensen, R., Lai, K.W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer International Handbooks of Education, Cham, Switzerland, pp. 1227–1238.
- Ennis, R.H., 2018. Critical thinking across the curriculum: a vision. *Topoi* 37, 165–184.
- Falcone, P.A., 1990. Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations. American Philosophical Association, Newark, DE (ERIC Document Reproduction Service No. ED315423).
- Ferrari, A., 2013. DigComp: A Framework for Developing and Understanding Digital Competence in Europe. European Commission, Luxembourg.
- FitzPatrick, S., O'Shea, M., 2013. Moving up the line: schools at the hub of policy development in Ireland. In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom across Europe*. SLO, Enschede, the Netherlands, pp. 119–137.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Duckworth, D., 2020. Preparing for Life in a Digital World. IEA International Computer and Information Literacy Study 2018 International Report. Springer, Cham.
- Gordon, J., Halsz, G., Krawczyk, M., Leney, T., Michel, A., Pepper, D., Putkiewicz, E., Wisniewski, J., 2009. Key Competences in Europe. Opening Doors for Lifelong Learners across the School Curriculum and Teacher Education. Center for Social and Economic Research on Behalf of CASE Network, Warsaw.
- Hamel, R., 1990. Over het denken van de architect [About the thinking of an architect]. AHA Books, Amsterdam.
- Hesse, F., Care, E., Buder, J., Sassenberg, K., Griffin, P., 2015. A framework for teachable collaborative problem solving skills. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, New York, pp. 37–56.

- Hilt, L.T., Riese, H., Sørdeide, G.E., 2019. Narrow identity resources for future students: the 21st century skills movement encounters the Norwegian education policy context. *J. Curric. Stud.* 51, 384–402.
- Hipkins, R., 2016. New Zealand's journey to build a competency-based curriculum. In: *Proceedings of Recent Challenges to the School Curriculum and Necessary Responses*. Korean Society of Curriculum Studies, Seoul, pp. 183–201.
- Klieme, E., Hartig, J., Rauch, D., 2008. The concept of competence in educational contexts. In: Hartig, J., Klieme, E., Leutner, D. (Eds.), *Assessment of Competencies in Educational Contexts*. Hogrefe & Huber Publishers, Göttingen, pp. 3–22.
- Kruit, P.M., Oostdam, R.J., Van den Berg, E., Schuitema, J.A., 2018. Effects of explicit instruction on the acquisition of students' science inquiry skills in grades 5 and 6 of primary education. *Int. J. Sci. Educ.* 40, 421–441.
- Kuiper, W., Nieveen, N., Berkvens, J., 2013. Curriculum regulation and freedom in The Netherlands—a puzzling paradox. In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom Across Europe*. SLO, Enschede, pp. 139–162.
- Law, W.W., 2014. Understanding China's curriculum reform for the 21st century. *J. Curric. Stud.* 46, 332–360.
- Lazonder, A.W., Egberink, A., 2014. Children's acquisition and use of the control-of-variables strategy: effects of explicit and implicit instructional guidance. *Instr. Sci.* 42, 291–301.
- Levy, F., Murman, R.J., 2006. *How Computerized Work and Globalization Shape Human Skill Demands*. MIT, Department of Urban Studies and Planning & Harvard University, Graduate School of Education, Boston/Cambridge, MA.
- McAra, K., Broadley, E., McLauchlan, J., 2013. Curriculum for excellence in Scotland—local flexibility or national exemplification? In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom Across Europe*. SLO, Enschede, pp. 211–228.
- McPeck, J., 1981. *Critical Thinking and Education*. Oxford University Press, Toronto, Canada.
- Nieveen, N., Fisser, P., Muller, A., Voogt, J., 2014. *Create the Debate: Towards a Set of Design Principles for National Curriculum Debates and Dialogues with Technology Support*. SLO, Enschede.
- Norman, G., 2005. Research in clinical reasoning: past history and current trends. *Med. Educ.* 39, 418–427.
- OECD, 2017. *PISA 2015 Results (Volume V): Collaborative Problem Solving*. OECD, Paris.
- OECD, 2018. *The Future of Education and Skills: The Future We Want*. OECD, Paris.
- Partnership For 21st Century Skills, 2009. *Framework for 21st Century Learning*. Partnership For 21st Century Skills, Hilliard, OH.
- Pietarinen, J., Pyhäntö, K., Soini, T., 2017. Large-scale curriculum reform in Finland—exploring the interrelation between implementation strategy, the function of the reform, and curriculum coherence. *Curric. J.* 28 (1), 22–40.
- Pietarinen, J., Pyhäntö, K., Soini, T., 2019. Shared sense-making in curriculum reform: orchestrating the local curriculum work. *Scand. J. Educ.* 63, 491–505.
- Plomp, T., Anderson, R.E., Kontogiannopoulou-Polydorides, G. (Eds.), 1996. *Cross-National Policies and Practices on Computers in Education*. Kluwer Academic Publishers, Dordrecht.
- Rychen, D.S., Salganik, L.H. (Eds.), 2001. *Defining and Selecting Key Competencies*. Hogrefe & Huber, Göttingen.
- Siddiq, F., Hatlevik, O.E., Olsen, R.V., Throndsen, I., Scherer, R., 2016. Taking a future perspective by learning from the past—a systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educ. Res. Rev.* 19, 58–84.
- Sinnema, C., 2011. *Final Evaluation Report on the Implementation of the New Zealand Curriculum 2008–2009*. Ministry of Education, Wellington, New Zealand.
- Siveskind, K., Westbury, I., 2016. State-based curriculum work and curriculum-making: Norway's læreplanverket 1997. *J. Curric. Stud.* 48, 766–782.
- Stadler, M., Herborn, K., Mustafi, M., Greiff, S., 2020. The assessment of collaborative problem solving in PISA 2015: an investigation of the validity of the PISA 2015 CPS tasks. *Comput. Educ.* 157.
- Stoel, G.L., Van Drie, J.P., Van Boxtel, C.A.M., 2017. The effects of explicit teaching of strategies, second-order concepts, and epistemological underpinnings on students' ability to reason causally in history. *J. Educ. Psychol.* 109, 321–337.
- Tan, J.P.L., Choo, S.S., Kang, T., Liem, G.A.D., 2017a. Educating for twenty-first century competencies and future-ready learners: research perspectives from Singapore. *Asia Pac. J. Educ.* 37, 425–436.
- Tan, L.S., Koh, E., Lee, S.S., Ponnusamy, L.D., Tan, K.C.K., 2017b. The complexities in fostering critical thinking through school-based curriculum innovation: research evidence from Singapore. *Asia Pac. J. Educ.* 37, 517–534.
- The Royal Society, 2012. *Shut Down or Restart? The Way Forward for Computing in UK Schools*. The Royal Society, London.
- Toh, Y., Hung, W.L.D., Chua, P.M.H., He, S., Jamaludin, A., 2016. Pedagogical reforms within a centralized-decentralized system: a Singapore's perspective to diffuse 21st century learning innovations. *Int. J. Educ. Manag.* 30, 1247–1267.
- UNESCO, 2010. *Engineering: Issues, Challenges and Opportunities for Development*. UNESCO, Paris.
- UNESCO, 2016. *A Policy Review: Building Digital Citizenship in Asia-Pacific Through Safe, Effective and Responsible Use of ICT*. UNESCO Bangkok, Bangkok.
- Van den Akker, J., Nieveen, N., 2021. Combining curriculum and teacher development through design research. In: Philippakos, Z., Pellegrino, A., Howell, E. (Eds.), *Design based Research in Education: Theory and Applications*. Guilford Press, New York.
- Van den Akker, J., 2003. Curriculum perspectives: an introduction. In: van den Akker, J., Kuiper, W., Hameyer, U. (Eds.), *Curriculum Landscapes and Trends*. Kluwer, Dordrecht, pp. 1–10.
- van Laar, E., van Deursen, A.J.A.M., van Dijk, J.A., de Haan, J., 2019. Twenty-first century digital skills for the creative industries workforce: perspectives from industry experts. *Clin. Hemorheol. Microcirc.* 24 (1).
- Veltman, M.E., van Keulen, J., Voogt, J.M., 2019. Design principles for addressing wicked problems through boundary crossing in higher professional education. *J. Educ. Work* 32 (2), 135–155.
- Volman, M., Ten Dam, G., 2015. Critical thinking for educated citizenship. In: Davies, M., Barnett, R. (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education*. Palgrave Macmillan, New York, pp. 593–603.
- Volman, M., Karssen, M., Emmelot, Y., Heemskerck, I., 2020. The focus of schools on twenty-first-century competencies and students' experience of these competencies. *Curric. J.* 31 (6).
- Voogt, J., Nieveen, N., 2017. Conceptualizing Time Lag Dilemma in Curriculum Change—An Exploration of the Literature. Report Prepared Under the Auspices of the OECD Future of Education and Skills 2030 Project. University of Amsterdam, Amsterdam.
- Voogt, J., Pareja Roblin, N., 2012. Teaching and learning in the 21st century. A comparative analysis of international frameworks. *J. Curric. Stud.* 44, 299–321.
- Voogt, J., Brummelhuis, A., 2014. Information literacy in the Netherlands Information literacy in the Netherlands: rise, fall and revival. In: Tatnall, A., Davey, B. (Eds.), *Reflections of the history of computers in education*. Springer, New York, pp. 83–92.
- Voogt, J., Pelgrum, H., 2005. ICT and curriculum change. *Hum. Technol.* 1, 157–175.
- Walker, D., 2003. *Fundamentals of Curriculum: Passion and Professionalism*. Lawrence Erlbaum Associates, Mahwah.
- Webb, M.E., Gibson, D., 2015. Technology enhanced assessment in complex collaborative settings. *Educ. Inf. Technol.* 20, 675–695.
- Webb, M.E., Fluck, A., Magenheimer, J., Malyn-Smith, J., Waters, J., Dechènes, M., Zagami, J., 2020. *Machine Learning for Human Learners: Opportunities, Issues, Tensions and Threats*. Educational Technology Research and Development.
- Wetenschappelijke Raad voor het Regeringsbeleid, 2013. *Naar Een Lerende Economie. Investeren in Het verdienvermogen van Nederland (Towards a Learning Economy. Investing in the Earning Capacity of the Netherlands)*. WRR, Den Haag.

Further reading

- Dede, C., 2010. Comparing frameworks for 21st century skills. In: Bellanca, J., Brandt, R. (Eds.), *21st Century Skills*. Solution Tree Press, Bloomington, pp. 51–76.
- Erstad, O., Voogt, J., 2018. The twenty-first century curriculum: issues and challenges. In: Voogt, G., Knezek, G., Christensen, R., Lai, K.W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer International Handbooks of Education, Cham, Switzerland, pp. 19–36. Springer.
- Fraillon, J., 2018. International large-scale computer-based studies on information technology literacy in education. In: Voogt, G., Knezek, G., Christensen, R., Lai, K.-W. (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer International Handbooks of Education, Cham, Switzerland, pp. 1161–1179. Springer.
- Griffin, P., McGaw, B., Care, E., (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, Springer.
- Kereluik, K., Mishra, P., Fahnoe, C., Terry, L., 2013. What knowledge is of most worth: teacher knowledge for 21st century learning. *J. Digit. Learn. Teach. Educ.* 29 (4), 127–140.
- Voogt, J., Erstad, O., Dede, C., Mishra, P., 2013. Challenges to learning and schooling in the digital networked world of the 21st century. *J. Comput. Assist. Learn.* 29 (5), 403–413.
- Williamson, B., 2013. *The Future of the Curriculum. School Knowledge in the Digital Age*. The MIT Press, Cambridge, MA.

Relevant websites

- Assessment and Teaching of 21st Century Skills (ACT21S), <http://www.atc21s.org/>.
- OECD The Future of Education and Skills—Education 2030, <http://www.oecd.org/education/2030/>.
- International Computer and Information Literacy Study (ICILS), <https://www.iea.nl/studies/iea/icils>.
- Key competences in Europe, <https://www.econstor.eu/bitstream/10419/87621/1/613705459.pdf>.
- OECD DeSeCo Project: EDU Working Paper (41), [http://www.ois.oecd.org/ois/2009doc.nsf/linkto/edu-wkp\(2009\)20](http://www.ois.oecd.org/ois/2009doc.nsf/linkto/edu-wkp(2009)20).
- Partnership for 21st Century Skills, <https://www.battelleforkids.org/networks/p21>.

Transnational curriculum inquiry

Elizabeth Macedo, Rio de Janeiro State University, Rio de Janeiro, Brazil

© 2023 Elsevier Ltd. All rights reserved.

Globalization and transnational curriculum	57
Internationalization of curriculum studies	59
Acknowledgment	62
References	62

It can be said that postcolonial thought is in many respects a globalized way of thinking, even if initially it does not use that term. In the first place, it shows that there is little disjunction between the history of the nation and that of the empire (...). Postcolonial thought demonstrates that colonialism itself was a global experience which contributed to the universalization of representations, techniques and institutions (in the case of the nation state, even of merchandise of the modern kind). It shows that this process of universalization, far from being a one-way street, was basically a paradox, fraught with all sorts of ambiguities.

Mbembe (2006, p. 5).

By opening this text with Mbembe's epigraph, I denounce some decisions that inhabit it. Besides offering me the option for a post-colonial and post-structural reading—possibly due to the geographic place where I was born, I live and work—this epigraph marks a willingness to get closer to the discussion of the curriculum field from the perspective of policy. However, this does not mean taking state curricular policies—at the interface between national and global demands—as the focus of this text, but understanding that the constitution of the field itself is an antagonistic struggle for hegemony (Laclau, 2011; Lopes and Macedo, 2014) as well as a performative action of normativity (Butler, 2015). There is also, in such a disposition, an ethical desire to question power operations that sustain “the relationship between the recognizable and the unrecognizable” (Butler, 2015, p. 5), thus implying a need to seek to understand the field production policies “from within [contingency and global incommensurability] and against it” (Bhabha, 1998, p. 21).

Terms such as globalization and transnational do not refer to new phenomena. Although such an interpretation is not consensual, as Mbembe (2006) and other post-colonial authors point out (Bhabha, 1998), these terms can even be associated to colonial experiences. Especially when dealing with this statement from the South American context, where colonialism refers to the navigation of the 16th century, there is possibly some temporal imprecision—which does not disturb post-colonial thinking in view of its distrust in relation to Modern temporal linearity. Without the pretension of origin, I take the second half of the last century as a time when globalization, defined by Mbembe (2006) as “the central phenomenon of our epoch” (p. 7), became an object of great value to sociological analyzes (Giddens, 1999; Castells, 2004; Bauman, 1998, 2005). I specifically refer to the post-war moment, when the integration of markets and also the search for major transnational agreements—for example, the Universal Declaration of Human Rights—were intensified. With the end of the Cold War and the large-scale growth of Internet use, the discussion of the global seems to have reached its peak in the late 1980s, to the point where Rosenberg (2005) calls it the *Zeitgeist* of the 1990s. The reiteration of the concept produced its displacement and terms such as transnational—also widely used—or globalization resurfaced in the arena. At times, they aimed to mark relevant theoretical distinctions in relation to globalization, at other times, to minimize the emphasis on the economic that this discussion had assumed.

I here do not aim to discuss such phenomena in their sociological and economic nuances, but only to address their intersections with the curriculum field over the past few decades. Such intention finds its first challenge in the proliferation of meanings for curriculum, recognized by authors with different theoretical views (Pinar, 2004; Young, 2013; Hopmann, 2015). When putting my post-structural option into operation, I deviate myself from an epistemological definition, and I assume that the field, at a given moment, is only the effect of the power relations that constitute it as a performative discourse (Butler, 2015). In this understanding, curriculum studies could not even be defined by an object that they seek to address—as in the question of authors as different as Pinar (2004) and Young (2013) about “what curricular theory is”. On the contrary, their answers [and even the question] sanction what is legitimate as a curriculum in the political game that constitutes the field as something heterogeneous and contingent. Without disdaining the failure with which all normativity lives (Butler, 2015), I operate here with the fiction that, named and summoned, goes by the name—which intends to give it identity—of curriculum studies. It is important to highlight that fictions matter because they have effects, including the sanctioning of what can be thought and said at a given moment about, in this case, the curriculum.

With this understanding of the field, the methodological path I have taken here has enabled me to seek meanings shared by researchers in “places” established in the political games that constitute it. I have done so in a non-exhaustive and less rigorous way, not because of sloppiness, but because I only aim, theoretically, to question some meanings that circulated and continue to circulate in these “places”. I thus have chosen to wander, like Poe (1993) when diving into the night of the city behind the man in the crowd, or follow Benjamin (2006) suggestion: to really know a city, it is necessary to get lost in it. So, I combed through, without quantitative pretensions, the two most cited journals explicitly dedicated to the curriculum theme—*Journal of Curriculum Studies* (JCS); and *Curriculum Inquiry* (CI)—to read texts that, in the last twenty years, dealt with phenomena such as globalization and trans/internationalization. The very decision to wander around the Global North reflects—and reproduces—the strength of the

researchers based there in the struggles for hegemonizing meanings for the internationalized field. I acknowledge the existence of extensive work on curriculum studies in different countries in Latin America, Africa and Asia, but I must confess that the little I know of its content came from what gained visibility in these (global?) “places” that I have studied via my sojourns through publications such as these two major curriculum studies journals.

Although both journals present themselves as international and as thus supporting the very idea that an academic field requires a certain sharing of traditions, it was curious that I felt myself in different places when exploring each one. Obviously, there is a circularity of meanings between the two, but there is also some untranslatability that cannot go unnoticed when discussing the inter/transnationalization of curriculum studies. Each journal has a different set of authors and, in many cases, the themes brought up for discussion or the emphases they assume differ as well. Most notably, possibly, the fact that the *JCS* favors a more sociological approach, with a greater presence of the analysis of the State’s role in global phenomena, while in *CI* non-national identities gain space.

Like [Calvino \(1990\)](#) traveling narrator, I here have invented two narratives that are even possible as means to escape tourist maps. My narratives are not about the most common themes in one or the other journal, nor will I bring only or predominantly what I read during my journal travels. I purposefully mixed these texts with other readings that constitute me as a researcher in the field to produce reflections on these narratives that I explore throughout the two sections into which this paper is divided.

The first of them has a dialog with Sociology, Economics and Political Science and takes concepts, such as nation-state, (neo) liberalism and globalization as important analytical tools. Thus, curriculum theory is considered in a close relationship with state curriculum policies and the narrative focuses on the impacts of global flows on curricula. I assume that globalization has broadened the reach of (neo)liberal normativity and I scrutinize the effects of that normativity on national and local curricular policies, especially due to the historical importance of curricula in the construction of a nation’s (sometimes unitary) metaphor. At the same time, I am concerned with the reduction that this movement has been producing in the sense of curriculum, what [Miller \(2014\)](#) called return “with a vengeance” (p. 18) of the “exclusively managerial and technical-rational orientations” (p. 18). However, I try to react to think of normativity in a saturating way, which, faithful to post-structural lessons, highlights its iterative or performative character.

In the second narrative that I compose, which was only implied through the two journals, I start from the movements that, in the last twenty years, have broadened the connection of researchers of the field at the international level, assuming that there is an effective desire to deal with difference and alterity. This is undeniable in the many initiatives coordinated by William Pinar and other researchers that culminated in the creation of the International Association for the Advancement of Curriculum Studies (IAACS) and its journal, *Transnational Curriculum Inquiry (TCI)*.¹ Thus, like the Association, the two versions of the *International Handbook of Curriculum Research* ([Pinar, 2003, 2014a](#)) and the project, “Curriculum Studies: Intellectual Histories, Present Circumstances” ([Pinar, 2010, 2011, 2014b, 2015](#)), aimed to expand the debate in the field beyond North America and Europe.

When I lost myself in both journals, I found the same desire in the movement for greater internationalization of editorial committees, but I also came across a strong concentration of authors and references in the Global North. I have no doubt that “as editors, [our colleagues] are committed to publishing articles that reflect diverse approaches to curriculum studies and a wide range of theoretical frameworks” ([Gaztambide-Fernández and Thiessen, 2012](#), p. 1). The homogeneity of text styles and research formats narrated in them, however, indicate that perhaps, as noted by the authors themselves, “the internationalization Project (...) runs the risk of perpetuating dynamics of colonization that have historically marked academic and educational enterprises” (p. 9). Thus, in a movement somewhat opposite to that of my first narrative, in the second, I seek to think about the challenges of producing an internationalization of the field that is not intended to be colonial, but that has to deal with asymmetric flows of power or with “traces of imperialism and colonialism (...) discernible [even in cosmopolitan projects]” ([Pinar, 2010](#), p. 2).

Globalization and transnational curriculum

[the] interest in policy (and in policy/practice relations) is itself a form of thinking about curriculum that is locating specific, arising from countries that have a history of formal state policy frameworks for curriculum, as distinct, for example, from the history of curriculum in the USA where strands of theorizing curriculum flourish (reconceptualists theorizing in particular) that have very little origin in or concern with “policy”.

[Yates \(2016, p. 367\).](#)

I start from Yates’ assertion that interest in policy is a way of thinking about curriculum; however, I disagree with the possibility of attributing geographic—and historical location—to that theorization. For the traveling narrator who toggles between *JCS* and *CI*, this was undoubtedly a tempting link. However, if it is true that the North American field, which is known, in great part, as one influenced by the reconceptualization and its proliferations that [Malewski \(2010\)](#) calls post-reconceptualization,² operates with other concerns, one could argue that the standards & testings binomial is the focus of much of the country’s curricular discussion.

¹In constituting this scene, I think it is important to highlight that I have been working within the Association since the mid-2000s, having participated in the projects carried out by it.

²I assume a broad reading of the movement—contesting the hegemony of the standards & testings model in Curriculum Studies—enclosing phenomenological and critical studies, as well as a whole set of post-structural approaches ([Malewski, 2010](#)).

This is explicit in different works, among which are the historical retrospective of the field carried out by Pinar et al. (1995) and the book, *Teaching by Numbers*, in which Taubman (2009) analyzes recent North American curriculum policies. Hopmann (2015) will even maintain that the North American curriculum movement at the turn of the 20th century, focused on “learning results [and] (...) toward the psychology of learning (...) was then exported during the 1960s into the continental European context (...) as a ‘better’ form of education planning” (p. 17). To the author, this has been repeated recently, for example, with the expansion of the Organization for Economic Co-operation and Development (OECD) governance over European curricula. Indeed, the organization itself (OECD, 2005) admits that the definition of competencies in its 2003 project is depository of the idea of generative skills or standards that comprised an American document from years before.

Still, in line with Yates’ (2016) argument, the scene I construct for my first reflection came to mind when browsing through the JCS texts and is, in some way, located in Europe. This is because it goes beyond the links between curriculum studies and state policies, referring to the impact of global flows on the non-uniform metaphor of national cohesion.³ I do not want to produce an origin for such a metaphor—in the European Enlightenment—nor to deny its spread across the Global South, for example as attested by the narratives gathered in the *International Handbook of Curriculum Research* (Pinar, 2014a). Even in the USA, where the idea of nation has its own nuances (Scott, 2007), there were several moments when curricula were called upon to play a relevant role in nation building (Kliebard, 1992; Pinar et al., 1995; Goodson, 2002; Franklin, 1986).

I begin my scene at some point in the second half of the last century when curriculum policies, sometimes with the adjective “national” included, began to make reference to transnational or global citizenship. It may be appropriate to consider that this movement took place in very different ways. On the one hand, the post-war world had begun to question totalizing discourses. The nation metaphor, described by Bauman (2005) as one of the most powerful mechanisms for creating the illusion of belonging via birth, was one of them. Identity, reason, subject—categories belonging to the same theoretical spectrum—were read with suspicion. While Cha-terjee (2000) denounced the nation’s bonds with authoritarianism and its disrespect for ethnicities and cultural differences, Bhabha argued that, as a fragile and unstable representation, the nation was unable to guarantee the promised unity. On the other hand, national states were also weakened by the creation of economic blocs, as well as by the expansion of neoliberal governance, which Brown described as “the peculiar form of reason that configures all aspects of existence in economic terms” (Brown, 2015, p. 17). Despite a certain antagonism between these two movements, that synchrony may have produced an atmosphere more conducive to the weakening of the nation state and national identities. It is not possible, however, to measure the cross-impact—in the sense of reinforcement, but also of opposition—between global governance and the disruptive potential of post-colonial criticism.

In the field of curriculum studies, this scenario raised proposals in defense of cosmopolitan, multicultural or global curricula, with biases ranging from liberal to critical, as well as with approaches focused on the environment, peace and citizenship (Heilman, 2010). In most of these proposals, the former links between curriculum and national identity are weakened, although universal curriculum visions are not necessarily called into question. Although I stumbled upon some of them on my journeys through the journals,⁴ especially through *CI*, my scene will not emphasize them. Instead, I center my attention on the effects of the rise of neoliberal normativity on national states, which shifted the focus “from a political to an economic semantic order” (Brown, 2015, p. 41), and on curricula national policies produced in the last twenty years. I am interested in highlighting the action of multi-lateral organizations such as the World Bank (BM), United Nations Educational, Scientific and Cultural Organization (UNESCO), OECD and Inter-American Development Bank (IDB), as well as of the networks that are replacing and strengthening them, in the governance of national policies. As Ball (in Avelar, 2016) points out,

in many countries, the framework, limits and possibilities of policy thinking are actually constrained, at least to some extent, by the language, assumptions and concepts which the World Bank and the OECD are articulating, advocating, promulgating, writing about, speaking about in conferences, and talking about with education ministries during visits and conferences. (p. 10)

In agreement with Ball, Sellar and Lingard (2013) argue that the expansion of global governance of curriculum policies intensified when the OECD mobilized its expertise in collecting and comparing statistical data for the creation of Program for International Student Assessment—PISA. Elsewhere, Lingard (2016) claims that, since then, the organization has disseminated, through well-structured propaganda, “a type of Weberian technical rationality as a demonstration of modernity of any nation” (Lingard, 2016, p. 618). To this end, it has been established soft governance exercised through the creation of infra-structural conditions, as well as through a “well-established capacity to shape the view of key actors in education across local, national and global scales” (Sellar and Lingard, 2013, p. 722).

With the expansion of the power of the OECD in the definition of global policies, agencies such as UNESCO, with a more cultural profile, lost space. Some of their concepts, such as competence, were appropriated receiving a more economic and instrumental guise. Since the 2000s, the OECD has redefined this concept, establishing what Mausethagen (2013) calls conceptual governance. The definition still used today, with huge impact on national curriculum policies, was created in mid-2000s: “a competency is more than just knowledge and skills, it involves the ability to meet complex demands, by drawing on and mobilizing psychosocial resources (including skills and attitudes) in a particular context” (OECD, 2005). Anderson-Levitt (2017) defends that this

³Papers about the pervasive effects of globalization on former curriculum policies in *CI* focus mostly on subjectivity or identity, as can be seen in Casella (2005) and Beltrán (2018). In *JCS*, although national identities play a big role, the focus on identity is not totally absent (LaSpina, 2003).

⁴Just to mention a few: Warriner (2017), Rizvi and Beech (2017), Hull and Stornaiuolo (2014), Pashby (2015), Jickling and Wals (2008), Saito (2010).

definition, broader than the previous ones and presented in less clearly economic terms, responded to the worsening of the international financial crisis, when the rhetoric of employability that competence would guarantee has eroded.

For Lingard (2016), among others, OECD governance via PISA has “legitimized a new educational agenda in favor of further school reforms” (Lingard, 2016, p. 622). Less than having its results used for decisions on national educational policies, PISA creates a “shared political enterprise” (Addey, 2016, p. 690) at the international level. The evaluation is used “for externalization purposes” (Lingard, 2016, p. 621), providing national governments with a rhetoric of modernity and external legitimation for their actions (Anderson-Levitt, 2017). By this means, the OECD governance constitutes “epistemic communities of policy analysts, bureaucrats and politicians within the Organization and in member countries” (Sellar and Lingard, 2013, p. 712).

Despite recognizing the importance of multilateral organizations, especially OECD, in formulating policy, Ball (Avelar, 2016) has stressed the need for attention to new social actors in the global education landscape. The author suggests the expansion of global governance networks to encompass networks exercised by “various philanthropic organizations and foundations” (p. 10), as well as by “edu-businesses and business foundations (...), social enterprises, charity global organizations, [and] (...) well established NGOs” (p. 10). Such decentralized “political communities” (Ball, 2012, p. 5) are organized around common problems and with a view to their solutions, producing “new forms of sociability” (Ball, 2012, p. 9). As with OECD governance, while antagonizing national states, global networks do not dispense with them as a sphere that will guarantee legitimacy for the solutions they promote. Expanding such governance, they create “new narratives about what counts as good policy” (Ball, 2012, p. 6), which put in place a new public management capable of producing solutions, but which mainly worked in defining the problems and its relevance. Ball and collaborators (Ball and Junemann, 2012; Avelar and Ball, 2019; Ball, 2016) have been producing countless mappings that have given visibility to an intricate set of relationships in those new forms of management that, after all, promote the same economic logic envisioned in OECD governance for the past two decades.

The movement for global curriculum policies generates theoretical demands for Curriculum Studies. On the one hand, technologies for curriculum design focused on the definition of standards and their testing are developed. Differently from what use to happen decades ago, most of these networks are nowadays creating their own epistemic communities in order to develop research in their own centers outside the Universities or, in some cases, in closed privately financed nuclei within the Universities themselves. If “developing curriculum” (Pinar et al., 1995) has strongly moved to the global networks, a community of researchers traditionally linked to the Universities, on the other hand, have been kept mobilized by the task of “understanding” it (Pinar et al., 1995). Thus, these new ways of producing national curriculum policies, as well as their content, gained prominence as a major research theme within curriculum studies.

In my wanderings through the journals, I indeed perceived that many articles dedicated themselves to the analysis of the impact of global governance on the field and on national policies. As Yates (2016) suggested, communities where the fiction of the national state was more strongly established were more concerned about that, although the meanings of “national” varied greatly. In JCS, for example, a whole issue number dedicated to the “phenomena such as globalization and transnational policy (...) explored from a national/local perspective” (Wahlström et al., 2018, p. 587) was released in 2008. One year later, another issue explored the links between citizenship and curriculum, focusing on the “the changes and challenges facing citizenship education policy and curricula by way of case studies from Europe and beyond” (Philipou et al., 2009, 291). The articles together show that even within Europe, there are major differences⁵: “supra-national citizenship and its relationship to national citizenship and other groups [are conceptualized] in differing ways” (p. 291). In an even more diverse perspective, Zhao (2020) explores the impact on China of ancient cultural values of “globalizing West-centric discourses and practices” (p. 109).

In exploring the national and local effects of global governance, these studies are foregrounding the performative character of neoliberal normativity. As Brown (2015) points out, such normativity exists only in action and, therefore, is not able to saturate everything around it. While focusing on the normative strength of neoliberal governance, these papers make explicit its “inconstant, differentiated, unsystematic, impure character” (p. 20). In some other papers, the proliferation of differences in the normalized environment is even more prominent. Anderson-Levitt (2018) has argued that, despite the undeniable convergence global policies have been producing, it is fallacious to believe in a stronger unified global curriculum than that produced by the Enlightenment. In a paper analyzing the potentiality of comparative ethnographic and case studies, the author argues that they “reveal a richer picture of national patterns (...) and how the organization and tone of classroom activities vary from country to country” (p. 357). This is the case, for example, in the study by Wahlström (2018) in Swedish schools, in which the author “explores how education policy (...) is both enabled and constrained by transnational policy flows and national policy built up by social, cultural, historical traditions are enacted through curriculum at the classroom level” (p. 655). Among others, those papers point out to the always-possible popping-up of difference, no matter how asphyxiating normativity seems to be.

Internationalization of curriculum studies

... consciousness of the world arises from the actualization of what was already possible in me, achieved through my encounter with the lives of others, my responsibility towards the lives of others and towards seemingly distant worlds and, above all, towards people with whom I have apparently no connection: the intruders.

Mbembe (2006, p. 5).

⁵Ireland (Keating, 2009); Germany and Spain (Engel and Orloff, 2009); Cyprus (Philipou, 2009); Slovakia and Estonia (Michaels and Stevick, 2009); UK (Marshall, 2009).

As I explained, in this second narrative I make a movement somewhat opposite to the previous one, in which I highlight the political-associative and academic articulation of constituting an internationalized field in the curriculum. Inverse, because, as highlighted, among others, by Pinar (2010, 2014a), Carson (2009) and Gough (2000, 2004), this articulation responded to the acceleration of global rationality based on standards & testing: “internationalization can provide scholars with critical and intellectual distance (...) from those standardizing processes of globalization” (Pinar, 2010, p. 1). Also inverse because such an answer pointed to an explicit and conscious commitment to difference, as stated in the IAACS constitution, in which the association is described as being “established to support a worldwide—but not uniform—field of curriculum studies (...) [and] to provide support for scholarly conversations within and across national and regional borders about the content, context, and process of education, the organizational and intellectual center of which is the curriculum”. This commitment has echoed, in different ways, since the creation of the Association: in debate by CI (Carson, 2009; Gaztambide-Fernández and Thiessen, 2002), in papers in the JCS (Gough, 2000; Kim, 2010) and in the TCI (Gough, 2004; Miller, 2006), among others:

... the internationalization incurs an ongoing ethical engagement with alterity, resulting in the reconstruction of one's own self-understanding and social relation.

Pinar (2014a, p. 525).

... the more significant challenge for the internationalization of curriculum studies is one of understanding, of finding our place and forging new directions in these global circumstances that face education.

Carson (2009, p. 157).

... the “internationalization” of curriculum studies might then be understood not so much in terms of translating local representations of curriculum into a universalized discourse but, rather, as a process of creating transnational spaces in which local knowledge traditions in curriculum inquiry can be performed together.

Gough (2000, p. 339).

... in considering possibilities of transnational and transcultural collective exchange within the contexts of a worldwide curriculum studies field inflected by flows and mobilities, such “unknowingness” [at the core of what we know] might enable those of us attempting to create a worldwide field to avoid a scenario wherein the non-Western “other” is positioned to desire recognition only through assimilation or nativism, for example

Miller (2006, p. 36).

The difficulties that the internationalization of an academic field presents to those who launch themselves with a commitment and real desire to do so without controlling the proliferation of difference—as expressed in Miller's warning—are spotted by most of the scholars involved in the process. Gaztambide-Fernández and Thiessen (2002), for example, foreground that “curriculum scholars committed to internationalization (...) must remain skeptical of the impetus and the practice of ‘internationalization’ without a critical stance toward colonialism” (p. 10). While Pinar (2010), after remembering that colonialism is not absent from projects committed to alterity, argues that it is necessary to “continuing caution in (...) analytic efforts, an anxiety that any analysis risked neocolonial appropriation” (p. 2) in order to “engage in international conversation cognizant of world history and present injustices but not fated to reenact them” (p. 2).

The difficulties of the internationalization experience require that it is “further theorized and enacted by [in Miller's suggestion] considering differing points of view on transnational flows and mobilities of people, ideas, and communication and their effects on our abilities to engage in collective exchanges” (Miller, 2006, p. 34). Here, in a way, I want to launch myself into this task, in line with what was proposed by Miller,⁶ suggesting that internationalization, thought beyond the framework of national communities, has the potential to trouble the universal normative voice. It would not be a question of rearticulating national fields, but of perceiving them as the “resistance of the singular as a stain on the universal” (Hidleston, 2010, p. 296). In this perspective, the inter/transnational would appear in its complexity and could only be conceptualized as a retroactive effect of its own process.

To explore this argument, I expanded my wanderings through the *International Handbooks of Curriculum Research* (Pinar, 2003, 2014a), as well as through the 5 vol of the project, “Curriculum Studies: Intellectual Histories, Present Circumstances” (Pinar, 2010, 2011, 2014b, 2015). Even though both are thought from the distinctive perspective of the national—as “the nation (...) remains the imaginary and material site in which much of humanity experiences daily civic life” (Pinar, 2010, p. 2)—the sense of national operates in different ways in each of them. The handbooks bring together narratives from the curriculum field in 28 countries in the first edition (Pinar, 2003) and 34 in the second (Pinar, 2014a), which together make up the volumes as international handbooks. They are, therefore, somewhat marked by the national perspective that I want to avoid. The books in the curriculum studies project are collections of texts by authors from the same country, in which they “provide an intellectual history, an

⁶The author proposes that the post-foundational autobiographical inquiry can be the way “to activate, situate and/or interrogate conceptions, memories, fantasies, idealizations, or normalizations” (p. 45).

account of the present circumstances, and an overview of curriculum studies in [the specific country]. They also bring in their own research agendas" (Hawkins & Jacob as cited in Pinar, 2011, p. xi). I bring this project as a powerful example for thinking the national "as a stain on the universal", as suggested by Hiddleston.

It is not that, within the national narratives of the handbooks, it is not possible to perceive a strong transnational flow of authors and theorizations. On the contrary, they show that the field has always been internationalized, at least for those outside the great centers of production of universalisms. In many narratives, especially produced in the Global South, references to North American curriculum studies—Tylerian rationality, Schwab's diagnosis of the field, Reconceptualization, among other milestones in the history of the field in the USA—are constant. Other national histories also show a more regionalized influence potential, being shared, as a reference, by researchers from neighboring countries. In general, external interference in the constitution of national fields is tensioned by the authors of the texts, in readings with different levels of criticality, which does not prevent the sharing of these speeches from sanctioning meanings for the curriculum. Although in my travels through *TCI* and *CI*, I have come across few texts outside the Europe-North America axis, it was also possible to see the influence of certain national narratives on the naturalization of what is meant by curriculum.

In the case of handbooks, due to their nature of reference, the performative action of certain speeches on national fields is more evident, as general overviews of the field are presented in different countries. Although the stories are local, the narrated object is sanctioned as a curriculum by a universal normativity that is alien to them.⁷ With this analysis, I do not want to take away from the exercise-of-making-known-to-other-local-stories the potential, which it certainly has, of displacing the normative meanings of curriculum. I do not deny the reiterative or performative potential of normativity, but I suggest the need for a more obviously anti-colonial attitude. I am concerned that the distinctive experience of the other is put in place of that on which a non-distinctive curriculum theory is going to be produced. This opens, in my view, a colonial vision that is difficult to overcome, inscribed in what Bhabha (1998) calls the logic of diversity.

The theoretical-practical task proposed by the post-colonial landscape to overcome such logic is expressed in Butler's defense (2015) of "a critical approach to the norms of recognition, [which] can... begin to dismantle those more vicious forms of logic that uphold forms of racism and anthropocentrism" (p. 37). This is not an easy task for any field. It demands negotiating [or living] with alterity, risking the very recognition that is established according to the rules that are being challenged. Perhaps that is why Miller (2006) appeal, "trust me" (p. 44), when she says something that should be obvious: "there is no collective 'we' in the field of curriculum studies" (p. 44). After all, the "collective we" sanctions an acknowledgment, which, if displaced, can dismiss and/or put anyone in the position of the other. Even with this risk, the critical approach proposed by Butler (2015) needs to be taken because it is the responsible [and impossible] response that must be given to the other (Derrida, 2003).

I thus start from Miller's warning—that there is alterity in the [North-American] curriculum field and that it is not always perceived—to defend that the internationalization experience can help in the critical task proposed by Butler, in producing "the radical affirmation of the density of proximity, of displacement, even of dislocation" (Mbembe, 2006, p. 5). The innumerable examples of homogenization that globalization has been providing, some of which I explored in the previous section, perhaps disallowed such a statement, although it was also explicit there that there is no total saturation. At the same time, if there is no "collective we" in the North-American field, this displacement that Mbembe talks about does not depend on blurring national borders. Even if this is true, it is also important to note that the most prevalent "collective we" in Western Modernity is, possibly, national identities. By erasing them, post-colonial thinking, transnational in nature, advanced in demands out of respect for difference. For this reason, I advocate the power of the shock produced by trans/international contact also in the field of Curriculum Studies.

I use the 5 vol⁸ of the "Curriculum Studies" project as an expression of a powerful experience of internationalization of curriculum studies. With slightly different structures, all volumes include, in addition to the texts of scholars from the countries in focus and a brief presentation by their authors⁹, synoptic articles of the stages of the project that involved more intense exchanges among researchers from different traditions—national, disciplinary, theoretical, among others. I take these exchanges as the core of the project and as moments of productive—and iterative—shocks among traditions:

... the exchanges between panel members and participating (...) scholars occurred via the Internet over a (...) [certain] period. (...) [The] two rounds of questions and replies (...) [they were studied by Pinar] for their discursive movements, including the concepts those movements conveyed. (...) Participants [were invited] to critique [the] commentary, the first time in draft form (...) and the second time publicly in the epilogue, a "final word" articulated by the participants themselves.

Pinar (2011,¹⁰ p. 4–5).

Exchanges between panel members and participants—as well as Pinar's commentaries on the fields in the countries in focus, which are read in conjunction with the epilogues—offer glimpses of the difficulties and potentials of the internationalization of curriculum studies. In the articles, as in the texts of the handbooks, there are strong international influences in the constitution

⁷Interestingly, this apparent agreement on what the curriculum "is" divides space with analyzes that indicate strong hybridity in the field, which, to some, even leads to a certain lack of characterization.

⁸In a way, my experiences as one of those who participated as authors in the volume on Brazil as well as a member of the panel of experts in the South Africa stage of the project are also present in my presumptions throughout this article.

⁹Produced by Pinar from interviews/questionnaires focusing on the authors' biographies and their perceptions of the field.

¹⁰Text taken from the volume about Brazil, describing the process that occurred throughout the project.

of national fields, with abundant reference to the international literature on curriculum, as well as on related areas. In one of the volumes, Pinar (2011) goes so far as to claim that scholars “grapple with local legacies and present situations through concepts imported from Europe as well as from North America and elsewhere in South America” (p. 1). The exchanges, however, indicate that the colonial normativity that produced the sharing of universalized meanings of curriculum was being dislocated—with greater or lesser intensity—in its use. In some passages, there are even heated debates, since different statements were not always acceptable to those who saw themselves as owners of the displaced stories. In one disagreement here, in another there, some panel members let out feelings of possession of the universalized knowledge produced in the Global North or that there would be more correct or updated appropriations of the canon. As Pinar (2014b) points out, in the volume on China, “difference may always exceed our capacities to understand it, but ethical engagement with alterity requires not only the suspension of one’s assumptions—to the extent possible; it is a subjectively demanding undertaking—but active listening to whatever others (...) write and say” (n.p.).

At the end of each volume, Pinar produces syntheses (which in one of the volumes he classifies as synergistic), in which he intends to identify “concepts around which ‘discursive movements’ were organized”, highlighting that they circulate and “craft an unmistakable sense of shared difference” (Pinar, 2011, p. 198). Despite the attention to the difference, they seek to produce a “distinctiveness” (p. 198) of the fields in the countries, in a suture that is, in its own way, colonial, regardless of who produced it. Again, in the epilogues, however, the difference breaks out in recognition of another self—“what my local vision saw in those texts was a path of differentiation, and what internationalization presented to me was a displacement of approximation” (Macedo, 2011, p. 225). Or even, as a disagreement to be seen by the foreign and synoptic gaze: “I question the coherence of enunciation with my experimentations with images and words (...) I propose to work with the concept of fabulation” (Amorim, 2011, p. 219).

There are many other moments of disturbances created by exchanges among the scholars working in the project, some made explicit, others certainly silent. I do not intend to romanticize the experiences of internationalization that I brought as example—exchanges like these happen all the time, not dependent on international projects. My argument has been that spaces like these help the movement toward a wider recognition of the “others”. To sustain such an argument, it is necessary to make explicit that I do not assume internationalization and its institutionalization in IAACS, for example, as inclusive projects from the Global North. Rather, I posit these as necessary responses to the challenges and demands of a specific historical time. They are radical echoes of shifts in totalizing notions and have emerged in every arena as unexpected effects of the global norm, as local movements exploding around the world can exemplify. Thus, they are not magical places, but rather are being “regulated by norms of recognition that are themselves hierarchical and exclusionary” (Butler, 2015, p. 38). At the same time, with Butler, I assume that while “avowing and showing certain forms of interdependency, [they] stand a chance of transforming the field of appearance itself” (p. 43). Changings in the field of appearance are crucial to trouble the marginalizations produced by exclusionary experiences of recognition. And they are being made by those persons who, suddenly, appear through the always-there breaches. In that sense, from a USA perspective, it can make sense to bring together internationalization and reconceptualization (Pinar, 2014a), as the desire to open up exclusionary norms also animated the latter.

In this sense, I am treating “the experience of internationalization”—like so many others in which we are urged to live in difference—as an ethical imperative. It is as if the strangeness, to which the contact with otherness gives place, would put everyone in an eccentric position from where the fantasies of identity could be more radically displaced. It is not that the other who insists on appearing, displacing the identity metaphors—“the collective we”—will not disappear in a new recognition. It is precisely in this aporetic ambivalence between refusing and wanting—after all, we want to be “we”—that the ethical obligation to respond to the interpellation of the other lies (Derrida, 2003). When the unrecognizable other—the foreigner, in the sense of Derrida (2003)—insinuates himself/herself, we are placed in front of such an obligation.

It is common sense that globalization has been expanding regulation, producing a normativity from which, at times, it seems there is no way out—an interpretation that I, myself, enunciated in the previous section. Perhaps it could be potent to think of it as a kind of backlash, a response to the proliferation of others who do not fit in the fields of the appearance of colonial modernity. Perhaps, as proposed by Bhabha (2013), this globalization in which we are living is only “a transition space captured between a past that refuses to die and a future that will not wait to be born, a place deeply marked by the ambivalences of barbaric transmission” (p. 60). If so, the collectives that populate the field of curriculum studies are likely to be more and more transformed in order to expand the “we”.

Acknowledgment

This paper is related to research financed by CNPq (Brazil) and FAPERJ (State of Rio de Janeiro, Brazil).

References

- Addey, C., 2016. O PISA para o desenvolvimento e o sacrifício de dados com relevância política. *Educ. Soc.* 37 (136), 685–706. <https://doi.org/10.1590/ES0101-73302016166001>.
- Amorim, A.C., 2011. Last words..., at the beginning of Difference. In: Pinar, W. (Ed.), *Curriculum Studies in Brazil: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York, pp. 217–220.
- Anderson-Levitt, K.M., 2017. Global flows of competence-based approaches in primary and secondary education. *Res. Pap. Educ. Knowl.* 16 (1), 47–72.
- Anderson-Levitt, K.M., 2018. Comparing ethnographies of teaching when comparison seems impossible. *J. Curric. Stud.* 50 (5), 608–623. <https://doi.org/10.1080/00220272.2018.1502809>.

- Avelar, M., Ball, S.J., 2019. Mapping new philanthropy and the heterarchical state: the mobilization for the National Learning Standards in Brazil. *Int. J. Educ. Dev.* 64, 65–73. <https://doi.org/10.1016/j.jedudev.2017.09.007>.
- Avelar, M., 2016. Interview with Stephen J. Ball: analyzing his contribution to education policy research. *Arch. Anal. Educ. Polit.* 24 (24), 1–17. <https://doi.org/10.14507/epaa.24.2368>.
- Ball, S.J., Junemann, C., 2012. *Networks, New Governance and Education*. The Policy Press, Bristol.
- Ball, S.J., 2012. *Global Education Inc.: New Policy Networks and the Neoliberal Imaginary*. Routledge, New York.
- Ball, S.J., 2016. Following policy: networks, network ethnography and education policy mobilities. *J. Educ. Pol.* 31 (5), 549–566. <https://doi.org/10.1080/02680939.2015.1122232>.
- Bauman, Z., 1998. *Globalization: The Human Consequences*. Polity, Cambridge.
- Bauman, Z., 2005. *Identidade*. Jorge Zahar Editor, Rio de Janeiro.
- Beltrán, A.C.D., 2018. The nowhere of global curriculum. *Curric. Inq.* 48 (3), 273–292. <https://doi.org/10.1080/03626784.2018.1474712>.
- Benjamin, W., 2006. *Passagens*. UFMG, Belo Horizonte.
- Bhabha, H., 1998. *The Location of Culture*. Routledge, London.
- Bhabha, H., 2013. *Nuevas minorías, nuevos derechos: Notas sobre cosmopolitismos vernáculos*. Sigo Veituno Editores, Buenos Aires.
- Brown, W., 2015. *Undoing the Demos: Neoliberalism's Stealth Revolution*. Zone Books, New York.
- Butler, J., 2015. *Notes Towards a Performative Theory of Assembly*. Harvard University Press, Cambridge.
- Calvino, I., 1990. *As Cidades Invisíveis*. Companhia das Letras, São Paulo.
- Carson, T.R., 2009. Internationalizing curriculum: globalization and the worldliness of curriculum studies. *Curric. Inq.* 39 (1), 145–158. <https://doi.org/10.1111/j.1467-873X.2008.01442.x>.
- Casella, R., 2005. The uses of rules for strategies of containment and self-regulation in a global economy. *Curric. Inq.* 35 (2), 183–211. <https://doi.org/10.1111/j.1467-873X.2005.00324.x>.
- Castells, M., 2004. *The Network Society: A Cross-Cultural Perspective*. Edward Elgar, London.
- Chatterjee, P., 2000. *Comunidade imaginada por quem?* In: Balakrishnan, G. (Ed.), *Um mapa da questão nacional*. Contraponto, Rio de Janeiro, pp. 227–238.
- Derrida, J., 2003. Anne Dufourmantelle convida Jacques Derrida a falar da hospitalidade. *Escuta*, São Paulo.
- Engel, L.C., Orloff, D.H., 2009. From the local to the supranational: curriculum reform and the production of the ideal citizen in two federal systems, Germany and Spain. *J. Curric. Stud.* 41 (2), 179–198. <https://doi.org/10.1080/00220270802541147>.
- Franklin, B.M., 1986. *Building the American Community*. The Falmer Press, London & Philadelphia.
- Gaztambide-Fernández, R., Thiessen, D., 2012. Fomenting flows and internationalizing curriculum studies. *Curric. Inq.* 42 (1), 1–11. <https://doi.org/10.1111/j.1467-873X.2011.00578.x>.
- Giddens, A., 1999. *Runaway World*. Profile Books, London.
- Goodson, I., 2002. *The Making of Curriculum: Collected Essays*. The Falmer Press, Washington.
- Gough, N., 2000. Locating curriculum studies in the global village. *J. Curric. Stud.* 32 (2), 329–342. <https://doi.org/10.1080/002202700182790>.
- Gough, N., 2004. A vision for transnational curriculum inquiry. *TCI* 1 (1), 1–11.
- Hawkins, J.N., Jacob, W.J., 2011. Series editors' introduction. In: Pinar, W. (Ed.), *Curriculum Studies in Brazil: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York, pp. xi–xii.
- Heilman, E.E., 2010. Global education. In: Kriedel, C. (Ed.), *Encyclopedia of Curriculum Studies*. Sage, Los Angeles, pp. 408–412.
- Hidleston, J., 2010. *Poststructuralism and Postcoloniality: The Anxiety of Theory*. Liverpool University Press, Liverpool.
- Hopmann, S., 2015. "Didaktik meets Curriculum" revisited: historical encounters, systematic experience, empirical limits. *Nordic J. Stud. Educ. Policy* 2015 (1), 14–21. <https://doi.org/10.3402/nstep.v1.27007>.
- Hull, G.A., Stornaiuolo, A., 2014. Cosmopolitan literacies, social networks, and "proper distance": striving to understand in a global world. *Curric. Inq.* 44 (1), 15–44. <https://doi.org/10.1111/curi.12035>.
- Jickling, B., Wals, A.E.J., 2008. Globalization and environmental education: looking beyond sustainable development. *J. Curric. Stud.* 40 (1), 1–21. <https://doi.org/10.1080/00220270701684667>.
- Keating, A., 2009. Nationalizing the post-national: reframing European citizenship for the civics curriculum in Ireland. *J. Curric. Stud.* 41 (2), 159–178. <https://doi.org/10.1080/00220270802467475>.
- Kim, Y.C., 2010. Transnational curriculum studies: reconceptualization discourse in South Korea. *Curric. Inq.* 40 (4), 531–554. <https://doi.org/10.1111/j.1467-873X.2010.00500.x>.
- Kliebard, H.M., 1992. *Forging the American Curriculum*. Routledge, New York.
- Laclau, E., 2011. *Emancipação e diferença*. EdUERJ, Rio de Janeiro.
- LaSpina, J.A., 2003. Designing diversity: globalization, textbooks, and the story of nations. *J. Curric. Stud.* 35 (6), 667–696. <https://doi.org/10.1080/0022027032000135058>.
- Lingard, B., 2016. PISA: fundamentações para participar e acolhimento político. *Educ. Soc.* 37 (136), 609–627. <https://doi.org/10.1590/ES0101-73302016166670>.
- Lopes, A.C., Macedo, E., 2014. The curriculum field in Brazil since the 1990s. In: Pinar, W. (Ed.), *International Handbook of Curriculum Research*, second ed. Routledge, New York, pp. 86–100.
- Macedo, E., 2011. Simply words. In: Pinar, W. (Ed.), *Curriculum Studies in Brazil: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York, pp. 221–226.
- Malewski, E., 2010. *Curriculum Studies Handbook: The Next Moment*. Routledge, New York.
- Marshall, H., 2009. Educating the European citizen in the global age: engaging with the post-national and identifying a research agenda. *J. Curric. Stud.* 41 (2), 247–267. <https://doi.org/10.1080/00220270802642002>.
- Mausethagen, S., 2013. Governance through concepts: the OECD and the construction of competence in Norwegian education policy. *Berkeley Rev. Educ.* 4 (1), 161–181.
- Mbembe, A., 2006. What is postcolonial thinking? An interview with Achille Mbembe. *Esprit* 12, 117–133.
- Michaels, D.L., Stevick, E.D., 2009. Europeanization in the "other" Europe: writing the nation into "Europe" education in Slovakia and Estonia. *J. Curric. Stud.* 41 (2), 225–245. <https://doi.org/10.1080/00220270802515919>.
- Miller, J.L., 2006. Curriculum studies and transnational flows and mobilities: feminist autobiographical perspectives. *TCI* 3 (2), 31–50.
- Miller, J.L., 2014. Curriculum theorizing in the throes of audit culture. *Curric. Teach. Dialog.* 16 (1&2), 13–30.
- Organization for Economic Cooperation and Development (OECD), 2005. *The Definition and Selection of KCs: Executive Summary*. OECD, Paris.
- Pashby, K., 2015. Conflations, possibilities, and foreclosures: global citizenship education in a multicultural context. *Curric. Inq.* 45 (4), 345–366. <https://doi.org/10.1080/03626784.2015.1064304>.
- Philippou, S., Keating, A., Orloff, D.H., 2009. Citizenship education curricula: comparing the multiple meanings of supra-national citizenship in Europe and beyond. *J. Curric. Stud.* 41 (2), 291–299. <https://doi.org/10.1080/00220270802475015>.
- Philippou, S., 2009. What makes Cyprus European? Curricular responses of Greek-Cypriot civic education to "Europe". *J. Curric. Stud.* 41 (2), 199–223. <https://doi.org/10.1080/00220270802558406>.
- Pinar, W., et al., 1995. *Understanding Curriculum*. Peter Lang, New York.
- Pinar, W., 2003. *International Handbook of Curriculum Research*. Lawrence Erlbaum Associates, New Jersey.
- Pinar, W., 2004. *What is Curriculum Theory?* Lawrence Erlbaum, Mahwah, NJ.

- Pinar, W., 2010. *Curriculum Studies in South Africa: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York.
- Pinar, W., 2011. *Curriculum Studies in Brazil: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York.
- Pinar, W., 2014a. *International Handbook of Curriculum Research*, second ed. Routledge, New York.
- Pinar, W., 2014b. *Curriculum Studies in China: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York.
- Pinar, W., 2015. *Curriculum Studies in India: Intellectual Histories, Present Circumstances*. Palgrave Macmillan, New York.
- Poe, E.A., 1993. *O homem da multidão*. Paraula, Porto Alegre.
- Rizvi, F., Beech, J., 2017. Global mobilities and the possibilities of a cosmopolitan curriculum. *Curric. Inq.* 47 (1), 125–134. <https://doi.org/10.1080/03626784.2016.1254500>.
- Rosemberg, J., 2005. Globalization theory: a post mortem. *Int. Polit.* 42, 2–74.
- Saito, H., 2010. Actor-network theory of cosmopolitan education. *J. Curric. Stud.* 42 (3), 333–351. <https://doi.org/10.1080/00220270903494261>.
- Scott, J.W., 2007. *The Politics of the Veil*. Princeton University Press, New Jersey.
- Sellar, S., Lingard, B., 2013. The OECD and global governance in education. *J. Educ. Pol.* 28 (5), 710–725. <https://doi.org/10.1080/02680939.2013.779791>.
- Taubman, P., 2009. *Teaching by Numbers: Deconstructing the Discourse of Standards and Accountability in Education*. Routledge, New York.
- Wahlström, N., Alvunger, D., Wermke, W., 2018. Living in an era of comparisons: comparative research on policy, curriculum and teaching. *J. Curric. Stud.* 50 (5), 587–594. <https://doi.org/10.1080/00220272.2018.1502814>.
- Wahlström, N., 2018. When transnational curriculum policy reaches classrooms—teaching as directed exploration. *J. Curric. Stud.* 50 (5), 654–668. <https://doi.org/10.1080/00220272.2018.1502811>.
- Warriner, D., 2017. Theorizing the spatial dimensions and pedagogical implications of transnationalism. *Curric. Inq.* 47 (1), 50–61. <https://doi.org/10.1080/03626784.2016.1254501>.
- Yates, L., 2016. Europe, transnational curriculum movements and comparative curriculum theorizing. *Eur. Educ. Res. J.* 15 (3), 366–373.
- Young, M.F., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45 (2), 101–118. <https://doi.org/10.1080/00220272.2013.764505>.
- Zhao, W., 2020. Problematizing “epistemicide” in transnational curriculum knowledge production: China’s suyang curriculum reform as an example. *Curric. Inq.* 50 (2), 105–125. <https://doi.org/10.1080/03626784.2020.1736521>.

Re-turning to the cultural, historical, and philosophical foundations of international curriculum[☆]

Weili Zhao^{a,b}, ^aJing Hengyi School of Education, Hangzhou Normal University, China; and ^bDepartment of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong SAR, China

© 2023 Elsevier Ltd. All rights reserved.

Prelude	65
Curriculum studies and globalization & my entry points	65
“Language/discourse” as a (de)colonial lens to disentangle the cultural-historical-philosophical foundations of international curriculum studies	67
Re-turning to the cultural-historical-philosophical foundations of China’s curriculum knowledge	69
Re-turning to the cultural-historical-philosophical foundations of international curriculum: some on-going efforts	70
“Returning to the historical-cultural-philosophical foundations” to better cut into the present schooling practices in China and beyond	71
Epilogue	72
Acknowledgment	72
References	72

Prelude

When I was invited to write a chapter for this volume, *Approaches to Curriculum and its Politics*, to be included in Elsevier’s fourth edition of the *International Encyclopedia of Education*, a temporary title was even suggested for my to-do chapter, i.e., Curriculum and globalization: Cultural, historical, and philosophical foundations of international curriculum. Reading that title, I was quite moved by the editors’ considerateness as these terms are what appeared on my personal website as “expedient” descriptors for my research interests. “Expedient” in the sense that there is indeed no single accepted sub-field of curriculum studies that is called “cultural, historical, and philosophical foundations of international curriculum”. Rather, I list them as boundary-setting descriptors of three broad domains, albeit hardly separable from each other and better put as “cultural-historical-philosophical”, to describe roughly what I have been exploring in international curriculum studies. I have now added “re-turning to” to the suggested title to foreground the fact that these three dimensions are not newly created ones in curriculum studies, but rather have been often marginalized or overwritten over the past few decades when curriculum and educational studies have always accented the, not less important though, political, social, and organizational entanglements along a neoliberal commodification logic. This being said, I put together this chapter as an overview of how I have come to delve into these entangled dimensions of international curriculum studies, say, what kind of curriculum scholarship and whose scholarship have guided me in my own learning and research trajectory. Henceforth, this chapter is in a sense an autobiographical summary of my recent research, far from a comprehensive review of the important works in international curriculum studies.

Curriculum studies and globalization & my entry points

Curriculum studies has always been a contested field since its very initial stage. According to Pinar (2008), the field has witnessed a four-stage evolution since early 1900s: (a) the Tylerian instrumentalization of curriculum development, (b) the reconceptualization movement of curriculum as a running course (*currere*) advocated by Pinar and his colleagues; (c) the multi-directional expansion of curriculum studies along the lines of narrative studies and phenomenological, historical, cultural, and philosophical foundations research; and (d) the post-2000 internationalization of curriculum studies. Pinar has undoubtedly made unparalleled contributions to internationalizing curriculum studies beyond the US boundaries, editing by himself or with colleagues a series of international or nation-based books on curriculum studies (say, on United States, Mexico, Brazil, South Africa, India, and China). Other similar efforts include, among others, international curriculum books edited by Wang and Eppert (2007), Baker (2009), Roofe and Bezzina (2018), Ashwani (2019), Hébert et al. (2019), and Ryan et al. (2020).

In the 50th anniversary issue of *Journal of Curriculum Studies*, Deng (2018) argues that Pinar’s and his colleagues’ reconceptualization, and the ensuing multi-dimensional expansion, movement of curriculum studies are in many ways a response to Joseph Schwab’s critique of the 1960s curriculum field as “moribund” and his call for inventing “new principles” and “new methods” in the field. These two movements, as a multi-discursive academic effort, have gradually and successfully replaced traditional curriculum studies focused upon curriculum development along a Tyler Rationale with a paradigm shift, namely, re-understanding

[☆]This project is partially supported by the startup grant from the Hangzhou Normal University, China.

curriculum from, among others, an array of historical, political, racial, (auto)biographical, aesthetical, theological, phenomenological, institutional and international, postmodern, and post-structural lens. This paradigm shift expands curriculum to an all-inclusive learning experience entangled with knotty issues of politics, gender, subjectivity, and stretches curriculum theorizing/studies to a broader scope intricated with new cultural, historical, philosophical, theoretical, and methodological perspectives. This multi-discursive academic effort was pushed further since the 21st century as a “postreconceptualizing” movement that intersects with posthumanism, queer theory, critical postqualitative studies and so forth, finally turning the curriculum field to encompass literally any topic related to politics, culture, society, environment, technology and human existence. To [Deng \(2018\)](#), the curriculum field has somewhat been transformed into a kind of “postmodern cultural studies” that is all-inclusive but neglects its core subject – practice and the inner work of schooling as a complex institution as its basic concern.

While Pinar’s and Deng’s comprehensive literature review retrospectively clarifies the apparently mixed field of curriculum studies in its dynamic, multidimensional, and rhizomatic growth and expansion, I would like to story-tell how my own research interests have grown along a transdisciplinary, transnational, and transcultural trajectory and ended up situating itself as a “re-turning to” the historical-cultural-philosophical foundations in international curriculum studies. Not particularly (or merely) inspired by the reconceptualization movement spearheaded by Pinar and other colleagues or the post-reconceptualization expansion in the Western scholarship, my research interests in curriculum studies were initially grounded in a different intellectual background of sociolinguistics and critical discourse studies along van Dijk’s and Norman Fairclough’s scholarship. I learned critical discourse studies through my former colleagues at Zhejiang University, [Shi-xu \(2005, 2016\)](#) and [Zongjie Wu \(2011, 2014\)](#), who have respectively worked very closely with van Dijk and Fairclough and endeavored to explicate a cultural and de-colonial discourse analysis paradigm that could help to explicate Chinese discourses as they are, not bogged down to the modern Western discourses and frameworks. In other words, my initial intellectual training was intimately intertwined with the issue of globalization (specifically the globalization of Western discourses) and (de)colonization, unpacking how globalizing discourses marginalize the Chinese cultural forms of knowledge (epistemes) in the guise of modernization and globalization.

Retrospectively speaking, my initial intellectual training provided two significant signposts or placeholders in developing my research interests in curriculum studies: a discourse/language lens and a transnational/transcultural framework. The discourse/language lens has been largely neglected in curriculum research except a few scholars like [Dwayne Huebner \(1966/1999\)](#) and [Daniel Tröhler \(2013\)](#). However, it smoothed my intellectual transition or expansion from a linguistic discourse analysis to a Foucauldian discourse critique of the issues of representation and signification, which in turn worked as a detour strategy and alerted me to the Chinese language as not a non-representational or a-representational language to the English phonetics but rather a disparate language system and being. Put differently, Foucauldian critique to modern Western language as a representational system not only moved me further to understand Heidegger’s ontological interpretation of language but also a historical mode of inquiry. [Heidegger \(1966/1996\)](#) vociferates, “thinking itself can be transformed only by a thinking which has the same origin and calling”. Inspired, I turned ‘discourse’ back to ‘language’ as a critical lens in my curriculum research and came to explore the Chinese language as an ontological being, an epistemic expression and trace of Chinese cultural mode of thinking or thesis, even though my former mentors commonly argue that modern Chinese language terms have been totally Westernized such that it is no longer a visible medium for traditional educational thinking and practice (see, [Wu, 2014](#)). One example is that I have discerned a Chinese “wind-education” discourse/language that expresses Confucius’ educational vision which I will unpack in detail in the next section.

This language lens has been continuously complicated by my exploration of transnational and global knowledge transfer and (re)production in relation to the issue of (de)colonialism and (de)coloniality, building off [Chakrabarty’s \(2000\)](#) well-known invocation: Western terms and frameworks are indispensable yet inadequate in explicating non-Western sensibilities. Pushing it further, I ask how it is possible to first discern non-Western sensibilities at this globalized modern world (as we researchers are already habituated to a modern Western and globalized representational epistemology) and then to portray them as they are and in the English wor(l)d, not through the English presentational framework. [Joao Parakseva’s \(2016, 2018\)](#) work on curriculum epistemicide helped me better link my own research interests with the (de)colonial scholarship that tackles the issues of modernity, coloniality, knowledge, being, and power. Specifically, as a form of colonialism, epistemicide “is constitutive of, rather than derivative from, modernity” ([Andreotti, 2011](#), p. 383), and “coloniality”, as the memory or legacy of colonialism, defines culture, labor, intersubjective relation, and knowledge production well beyond the limits of colonialism and long after the end of a colonial administration ([Maldonado-Torres, 2007](#)). This cohort of de-colonial scholars still ask the political critical question, say, *what/whose knowledge gets taught at schools as legitimate knowledge and what/whose not* (see, [Apple, 2000](#)), but they crossover the schooling boundary to the nation-state level. Furthermore, they alert us to another equally significant and yet often neglected question in mainstream Western scholarship, namely, *what other forms of knowledge exist(ed) in the world outside the Western epistemological harbor*. This question, as they rightly argue, is highly crucial toward reaching a de-colonial “cognitive justice” worldwide (see, [Sousa Santos, 2007](#); [Paraskeva, 2016, 2018](#)).

Intriguingly, these scholars also advocate language as a de-colonial gesture in such endeavor, perfectly abridging my two signposts of language and globalization in curriculum research. This resonates well with my own research on, through, and with language and discourse and sharpens my argument that language is both a colonial and de-colonial tool and expression. By that it means language can become a window for us to historically decolonize the Western modernity coloniality in transnational knowledge transfer and (re)production. It is not an investigation of the language system, but rather, to turn language as a window to glimpse into the cultural-historical-philosophical foundations of international curriculum studies. Such

a linguistic-cultural-historical-philosophical lens, as I show through my own research projects as below, revisions the otherwise rhizomatic yet ahistorical “ideal” linking and hybridization of culturally distinct onto-epistemologies, worldviews, languages, spatiotemporalities, and styles of reasoning. Instead, it explicates Other forms of knowledge outside of the bracket of the English wor(l)d, providing not an “alternative thinking of curriculum”, but rather, an “alternative thinking of alternatives” (see, [Sousa Santos, 2007](#)).

This being said, I detail out below my “re-turning to” trajectory in exploring the historical-cultural-philosophical foundations in curriculum research through a discourse/language and transnational framework. While taking up an autobiographical review of my own research projects, it also speaks volume to a broader curriculum scholarship that contextualizes, conditions, as well as transforms my own running course of learning and unlearning.

“Language/discourse” as a (de)colonial lens to disentangle the cultural-historical-philosophical foundations of international curriculum studies

“Language” and “discourse” are two popular yet contested notions. Linguistics and sociolinguistics usually view language as a self-justified linguistic system and discourses as specific language terms in use. Western philosophy and literary studies closely link language with the grounding issues of metaphysics, signification, and representation. With Michel Foucault and other post-modern or post-structural thinkers, language and discourse are no longer treated merely as a given linguistic system or term impregnated with fixed semantic meanings, but become a theoretical construct, an embodiment of historical-cultural styles of reasoning, which orders our thinking about social-political reality and truth. Unpacking language/discourse this way is to render visible the broader landscape of thought and practice. For example, Tröhler’s *Languages of Education* (2013) provides an original narrative on how the rooted religious languages of Protestantism and republicanism, in confrontation or negotiation, have historically, ideologically, and politically structured the schooling in Germany, US, and Swiss. Focused on contextualized religious and political languages (*langues*) rather than arguments (*paroles*), Tröhler’s book conducts a transnational archeology of educational arguments in a historical, analytical, and comparative way, adding new insights to Whig and Marxist interpretations. Similarly, [Horlacher \(2016\)](#) historicizes the German notion of *Bildung* as an educational language/discourse and explores how this notion gets deployed in other linguistic and cultural contexts, say, the Scandinavian countries and the English-speaking world, to impart varied educational aspirations and styles of reasoning. [Blumenfeld-Jones \(2016\)](#) pinpoints that the Protestant discourse/language derived from the Hellenic and subsequently Christian thinking has structured and ordered the Western world, marginalizing and excluding a Judaic rabbinic interpretive paradigm. All of them foregrounds language as an epistemic trace of historical and cultural style of reasoning, i.e., a language-episteme conjuncture.

Back in the 1960s, Dwayne Huebner, “the most important scholar in the field of curriculum studies” ([Pinar, 1999](#)), already alerted us to this language-episteme conjuncture when he vociferated that curriculum research was in search of a new language. To him, curriculum scholars back then were often overly dependent on scientific thought patterns (say, Tyler Rationale), which unfortunately blinded them to many other significant and important intellectual, say, religious and spiritual, traditions of both East and the West. Drawing upon Heidegger’s language thinking, Huebner envisioned five sets of languages and value systems, *technical, political, scientific, esthetic, and ethical*, that comes with a rich, wholistic, and meaningful curriculum. However, a curriculum delimited to the Tylerian managerial and instrumental language would only accent the *technical* value system, with ends or objectives carefully and accurately specified in behavior terms and activities designed to fulfill the ends. While such curriculum also attends to the *political* and *scientific* valuing, they often completely ignored the *esthetic* and *ethical* valuing, failing to account for the symbolic meaning of education as well as the very human-human encountering in educational engagements. Huebner’s critique remains true with today’s educational and curriculum practices that are still largely enthralled by a predominant learning logic, an instrumental trope, and managerial governance. Again, language represents an epistemic style of reasoning (see, [Zhao, 2021](#)).

[Michel Foucault \(1973\)](#) understands “episteme” as some historical “epistemological unconsciousness” that grounds and configures knowledge formation and its discourses within an historical era. Simply put, what language terms we choose to use and how we use them are closely entwined with our epistemic, conscious or unconscious, styles of reasoning. In this light, Huebner advises us to examine how people talk about or describe educational activities, to pay attention to the language they use, and to explicate the values or epistemic styles of reasoning intended through the used language. Only after such scrutiny can we produce a systemic rationality to maximize each of the five languages-values toward developing a wholistic curriculum and more meaningful educational activities. For example, drawing upon religious and spiritual wisdom and languages like “spirit”, “spiritual”, and “transcendental”, [Huebner \(1985/1999\)](#) intends to encompass the riches of education beyond the parameters of the Tyler Rationale. Furthermore, he (1966/1999) artfully rephrases the word “responsibility” into “response-ability” as a sought-after educational agenda. Namely, education is to cultivate students’ “response-in-the-world” (p. 112), a responsive sensitivity as well as sensibility with which students are awakened to, and become responsive to, the unveiling of the unconditioned, the new, and the unique in the real world.

Placed in the field of trans/inter-national curriculum knowledge (re)production, Tröhler’s and Huebner’s curriculum language research has implications for the issue of epistemic colonialism that has been constitutive of the modernization process of non-Western nation states and of modernity as coloniality (see, [Zhao, in press](#)). Specifically, the globalization of the Tyler Rationale language and other Eurocentric discourses not merely reduces curriculum and educational thinking into

an instrumental and managerial enclosure. More importantly, it marginalizes and overwrites other historical-cultural forms of knowledge or epistemes, both within and outside of the West, causing an effect of “curriculum epistemicide” (Paraskeva, 2016). It is to be noted here that linguistic epistemicide doesn’t merely mean a replacement of discourses per se. Rather, it is a killing of a mode of reasoning and a form of knowledge-making that come with different cultural discourse and language. In alignment with [wa Thiong’s \(1986\)](#) claims that language is the most important vehicle in subjugating the spiritual mindset of the colonized, [Paraskeva \(2016\)](#) foregrounds “linguistic genocide” as the “very core of the colonial and the neocolonial project” (p. 202). That is, “the production and reproduction of hegemonic forms of knowledge are precisely the institutionalizations of a linguistic or cultural epistemicide” (p. 241). Furthermore, these scholars argue that language can be flipped aside as a de-colonial tool in exposing and suspending the “darker side” of the Western modernity-coloniality of knowledge, power, and being toward global “cognitive justice”, a de-colonial gesture that hopes to bring to the foreground “a silenced and different genealogy of thought” ([Paraskeva, 2016](#), p. 80).

Over the past 10 years and with a Heideggerian-Foucauldian language-discourse perspective, I have been problematizing both the modernization of the Chinese language and the popularization of West-centric discourses in China as an imprint of modernity-coloniality (see, [Zhao, 2019a, 2019b, 2020a, 2021, in press](#)). However, compared with Paraskeva’s and others’ arguments, I have reiterated that things are a bit different with China and the Chinese language: Even though China was partly colonized from mid-19th to early 20th century, the modern Chinese language was not colonized or replaced by English and Chinese people still speak Chinese language. Nevertheless, the meaning making rationale behind the Chinese language is indeed Westernized-modernized into a conceptual signifier-signified mode of signification (see, [Zhao, 2019a, 2019b](#)). That is, the Chinese words are not dead, but the way they are read has been transformed to the extent that the original cultural-historical senses nurtured within each Chinese character has been overwritten by its modernized Western terms. Such modernization of the Chinese language happens as an effect of modernity coloniality since most Chinese academics and policy makers since the 20th century have critiqued the Chinese language as being illogical and welcomed a replacement by the modern Western terms (see, [Zhao, 2019b](#)). Put differently, the Chinese intellectuals and institutions have subscribed to the murder of their own cognitive matrix ([Paraskeva, 2016](#)). The fact that Chinese policy makers and academics are still feverishly borrowing Western curriculum policies and practices re-affirms the long-time effect of modernity-coloniality on the colonized ([Maldonado-Torres, 2007](#)).

Nevertheless, I have also argued that, with its rising nationalism and fast economic development in the past few decades, China is striving to recover its cultural soft power to maintain its Chinese characteristics on the global platform (see, [Zhao, in press](#)). Toward that end, more and more traditional-cultural discourses are being re-invoked and re-generated to counter the Western power and foreground its Chinese features. Within such a context, it is opportune to say China currently speaks a language of heteroglossia as old Chinese monographs and terms are continuously re-invoked, re-combined, and re-interpellated as glosses or counter-glosses of Western discourses. Such a mixture of discourses makes the expression of epistemicide more entangled and the unpacking of epistemicide more complicated. Epistemicide expresses itself less in unpacking what forms of discourses are circulating, than in how the varied forms of discourses are ordered, juxtaposed, interpreted, and manipulated by the Chinese academicians and policy makers. In other words, it is through unpacking the academicians’ and policy makers’ way of understanding and reasoning the varied discourses in assemblage that we can glimpse into the operativity and expression of epistemicide. Epistemicide links itself not only to the circulating discourses, but how the modern Chinese language and schooling discourses, whether Western or traditional Chinese, are *interpreted* semantically, culturally, and cross-culturally, and more importantly, in the *English* language to the international audience (see, [Zhao, in press](#)).

For example, I have examined language and discourse as both a colonial and de-colonial gesture, a very site of epistemological dilemma ([Zhao, 2020a, 2021, in press](#)) that confronts China’s ongoing competency-based curriculum reform. China released its *Core Suyang Definitions for Chinese Student Development* in 2016, invoking the traditional-cultural notion of *suyang* as both a gloss and counter-gloss of the English terms of competency, skills, and literacies. Claiming to be more than a replica of the OECD’s and USA’s frameworks, China’s *suyang* definitions also foregrounds the Confucian tradition of *learning, body-cultivation (xiushen), and governing* in nurturing holistic student development. However, I have argued that Chinese policy makers’ and academics’ efforts of making a Chinese characteristic competency-based curriculum are thwarted by their unconscious subjection to the modernity-coloniality, and the reinvocation of cultural *suyang* discourse is nothing but a linguistic trope ([Zhao, 2020a](#)).

There, I examined how the *suyang* curriculum discourses are *interpreted* semantically, culturally, and cross-culturally in relation to the Western discourses of competency-skill-literacy and found Chinese intellectuals, still largely subjugated to the modernity-coloniality, have subscribed to the murder of their own cognitive matrix ([Paraskeva, 2018](#)) in two expressions. First, Chinese scholars commonly presume a *suyang* = competency ordering and interpret *suyang* merely as a semantic equivalent of competency. The *suyang* = competency ordering well betrays scholars’ subscription to the modernity-coloniality episteme of conceptual thinking, or what [Foucault \(1973\)](#) puts as a representational language and a signifier-signified mode of signification, reducing language to a trap of philology which assumes the a priori existence of grammatical arrangements in a language in which meaning is deduced from the grammatical structure. Second, even though *suyang* foregrounds cultural elements of “desirable character-traits” and “emotions-attitudes-values” (on top of the Western/modern “key competencies-skills-literacies”), they are largely treated as psychological traits, ordered in no way different from competencies-skills-literacies and failing to explicate the cultural-traditional episteme of “body-thinking” ([Wu K. 1997](#)) that goes with a Confucian tradition of *learning, body-cultivation (xiushen), and governing* in nurturing holistic student development.

Inspired by Foucault’s and Heidegger’s thinking on language/discourse and representation, I have bracketed *suyang* as a modern concept and rewritten it into a grammatical form of “su + yang”, probed into the etymological sayings of *su* and *yang*, and explicated

a Chinese (w)holistic body-thinking (Wu, K. 1997) episteme that could be nurtured within the cultural notion of *suyang* (see, Zhao, 2020a, 2021, in press). The Chinese notion of “body” registers not the physical or biological body of the modernity episteme but a “psychosomatic self” (Ames, 1993, p. 165), a (w)holistic mind-body-heart ensemble. Thus, rather than “bodily thinking” or “thinking has a bodily dimension”, the (w)holistic “body-thinking” enacts thinking ontologically such that the body is thinking and thinking embodies itself in bodily engagements. This (w)holistic body-thinking epistemically responds to the Confucian “body-cultivation” (*xiushen*) in that it is by playing/living varied contextualized roles that a Confucian person learns to make his/her personhood in relation to others, not as a foundational and autonomous individual like “the learner”. This (w)holistic body-thinking episteme opportunely re-treats the Confucian virtues or character traits such as “human compassion, righteousness, ritual propriety, wisdom and trustworthiness” no longer as being reducible to a modern conceptualization of ideas. Cultivating such virtues becomes a mode of being to be lived and practiced, say, through parents’ bringing up the children and children’ repaying their parents by showing filial piety and reverence when their parents get old (Zhang X., 2015). In other words, I turned language into a de-colonial tool, rendering visible a silenced and different genealogy of thought (Paraskeva, 2018), which provides an alternative paradigm and direction for reconfiguring China’s competency-based curriculum beyond the grip of Western/modern discourses/languages and frameworks.

Re-turning to the cultural-historical-philosophical foundations of China’s curriculum knowledge

As shown above, examining language as epistemic traces and unpacking translingual practices of translation help to draw our attention to what is otherwise marginalized or suppressed Other cultural-historical forms of knowledge. Yet, as I have elsewhere argued (see, Zhao, 2019b), explicating such Other cultural-historical epistemes as they are and not in the representation of the modern present is hardly possible without a historical mode of inquiry, because, as Heidegger (1966/1996) rightly argues, “thinking itself can be transformed only by a thinking which has the same origin and calling”. To Heidegger, a historical inquiry is distinct from the usual history of ideas mode of inquiry, namely, historical reporting of the various opinions, viewpoints, and propositions, about something through the centuries. The history-of-ideas mode of inquiry already presupposes the static existence of something and only asks about the formula and the definitions of the essence of something, whereas a historical mode of inquiry asks about the happening and the movements that something has taken historically, of which the formula and the definition are only the residuum and sediment. Furthermore, Heidegger (1967) argues that the historical happenings and movements of something are not necessarily the already established and past actual movements, but could also be in a form of *acquiescence* (my italics) no longer going on in the present yet still constituting as a basic form of the presence of the history. This is because what is merely actually past does not exhaust what has been. The purpose of this historical mode of inquiry, Heidegger further argues, is not to merely acknowledge how something was before, but to pose for decision how essentially it still is today. Put differently, the detour into the past is aimed to access the present.

To Foucault, such historical mode of inquiry is archeological or History of the Present which he grants a privilege over all the other disciplines in his discussion about today’s human sciences. He clarifies that discourses are linked in a historical fashion such that the present is an effect of conditions of historical possibilities. Henceforth, in studying an ensemble of theoretical discourses concerning language, economy and living beings in *The Order of Things* (1973), Foucault didn’t try to establish the a priori possibilities or impossibilities of such knowledge but do a historian’s work by showing the simultaneous functioning of these discourses and the transformations that accounted for their visible changes. In other words, archaeology concerns knowledge constitutive in its condition of possibility rather than a history of ideas or sciences. It is an inquiry that, by going back upstream in the history of discursive formations, knowledge, and practices, seeks to discover on what basis knowledge and theory became possible (Agamben, 2009). This archaeological-historical mode of inquiry thus involves an epistemological shift from viewing time, history, and tradition as being some static, irreversible and irreducible essences (data) to re-configuring them as happening, events, movement, and moments of arising possibly external to any timeline. The aim of disrupting the fixed and canonized tradition and history is to set back into motion the original inner happenings of something by tracing its way back to its self-laid grounds to found itself anew out of them. This new founding and finding are expected to shed new light on and render accessible the otherwise enclosed present educational reforms and discourses (see, Zhao, 2019b).

The field of curriculum studies lacks a historical perspective in general (Phillips, 2014) even though it is highly needed to reclaim some “historical-philosophical-cultural motifs” (Kazamias, 2001, p. 37) as the very root of education to re-humanize the currently predominant neoliberal order (Baker, 2009; Popkewitz, 2008, 2013). Since Herbert Kliebard’s classical historicization of the American curriculum from 1893 to 1958, historical unpacking of educational thinking and practices in the Foucauldian sense of the history of present has emerged, if not flourished, as a new field. For example, Bernadette Baker’s (2009) edited book, entitled *New Curriculum History*, indexes the challenges in locating the irreducible “heritage” of transnational curriculum studies which meanwhile does not centralize curriculum, education, knowledge, and power as boundary fields or reproduce the second-order normativity embedded in the a priori definitions of these terms. Thomas Popkewitz et al.’s edited book (2017) problematizes the “reason” of schooling by historicizing the conditions of possibilities of curriculum, pedagogy, teacher education, and the sociology of knowledge in varied cultural (con)texts. Zazanian and Nordgren (2019) edited a special issue of the Journal of Curriculum Studies that revisits historical consciousness as a theoretical construct and explores its broad understanding for contemporary education and educational research.

I did my doctoral studies with Thomas Popkewitz and Bernadette Baker at University of Wisconsin-Madison. Inspired by their historical unpacking of transnational curriculum knowledge, my doctoral research was to discover on what basis knowledge and styles of reasoning about China's education and curriculum become possible, and within what space of order they are constructed. It explicates the historical-cultural a priori which makes it possible for teaching, learning, and the teacher-student difference in past and present China to be said and enacted as they were and are. In other words, my doctoral research didn't track the historical vicissitudes of China's education and curriculum reform; rather, it got deeper into the epistemological field, examining the positivity of knowledge as well as the historical-cultural principles as conditions of possibility for today's educational thinking and practice in China on a cross-cultural landscape (see, Zhao, 2019b).

For my doctoral research, the turning-point signpost is my serendipitous discerning of the prevalent yet silent "wind-education" discourses, literally put, "school/teaching/learning wind" in contemporary China's schooling. The Heideggerian-Foucauldian historical-archaeological mode of inquiry enabled me to ask not what "school wind" means but how it is historically possible for wind and education to intersect in Chinese culture. Such a discerning set me on to a journey of dancing with the "wind", or rather, following the dancing "wind" to be open to the (con)textures and (over)tones of the Chinese language itself and ancient texts. For example, rather than "air in movement" as defined by the Oxford dictionary, the Chinese character of "wind" etymologically says "wind blows, and insects get germinated and transformed within 8 days" (Shuowenjianzi), disrupting the modern-western conceptual style of reasoning and signifier-signified mode of signification. In other words, the Chinese "wind" has an inherent sense of "transforming" or as Kuriyama rightly puts it, "the fascination of the Chinese winds lay in their power to *transform*" (my italics) (1994, p. 23). It is this inherent "transforming" sense that links Chinese "wind" back to Confucian teaching and learning narratives, say, "wind (is) teaching, wind moves and teaching transforms" (Han Preface to *Book of Odes*, around 200AD) and "gentleman's virtue is (like) wind and petty man's virtue is (like) grass; when wind blows across it, the grass must bend" (*Analec*s, around 500BCE). Put succinctly, the imagination of Chinese "wind" instead looms large in and across domains of education, politics, poetry, and Chinese medical body (see, Zhao, 2019b).

Following further the dancing "wind" finally exposed me to Confucius' own envisioning of teaching and learning as expressed in his commentary on one Yijing hexagram called *guan* (observation) with a hexagram-image of wind blowing over the earth (see, Zhao, 2019a, 2019b, 2019c). This hexagram, itself having nothing to do with teaching and learning, zooms on one ritual scenario where the Son of Heaven washed his hands but not yet offered the sacrifice, with a pious outlook and in front of his subjects. Confucius added a couple of comments, leveraged upon the "transforming" sense of Chinese "wind", and envisioned it into an ideal, exemplary and contingent, teaching-learning movement that happens like wind blowing over the earth. Indeed, I was accidentally exposed to the Yijing *guan*-observation hexagram, which I argued as an originary, not original, re-source of the entire Confucian teaching and learning, and the notion "wind" can be re-viewed as the signature language of Chinese education (Zhao, 2019a, 2019b). Neither point has so far been discerned, let alone scrutinized in scholarship on Confucius and Confucian educational culture in China and beyond. Here, "originary" needs to be understood in the sense that there is always something unsaid that could be further tapped, that what was said within the Chinese notion "wind" is always in being in advance of it, and yet is never expressly thought in its own right and as the Originary.

In other words, Heideggerian and Foucauldian historical-archaeological mode of inquiry enabled me to disrupt and suspend the modern Westernized representational thinking, setting back into the motion the original inner happenings of the presently silenced "wind-education" rationale by tracing its way back to its historical-cultural-philosophical self-laid grounds to found itself anew out of them. The historical-cultural-philosophical returning is not for a nostalgia purpose, but to better cut into the present, exposing deeper the coloniality of knowledge, being, and power of modernity on current Chinese people, say, my examination into China's competency-based *suyang* curriculum reform as detailed above. Such historical-cultural-philosophical returning provides not merely an alternative thinking of China's curriculum, but rather, an alternative thinking of alternatives. My "wind" research along a detour-and-access (detouring into the Western scholarship to access the Chinese self, and detouring into the Chinese past to better access its present) cartography of knowledge, language, and education blurs the geographical and epistemic boundaries of East, West, past, present, human subjects, and the researched objects. Instead, I am thrust into a new pluri-verse where to engage in varied forms of knowledge. Within such pluri-verse, language is no longer assumed to exist as a legitimized and enclosed system; meaning is no longer decided by the grammatical arrangements of statements; education is no longer about a neoliberal seeking of outcome and profit; and doing research is no longer merely to apply a pre-given framework to certain contexts. My dancing with the "wind" is more of an itinerant encountering, a de-centering and re-centering of various signposts, a processing, something that can only be recounted backward rather than anticipated ahead of time (see, Zhao, in press).

Re-turning to the cultural-historical-philosophical foundations of international curriculum: some on-going efforts

Currently, I am pushing further the cultural-historical-philosophical foundations of China's curriculum studies in deeper connectivity with the Western scholarship, as well as editing books on furthering transnational dialogues on the cultural-historical-philosophical foundations of international curriculum studies.

For the former, I am writing my second monograph, *Edusemiotics, (New) Materialism, and (Body) Governance: Flashpoints in 21st Century Education in China and Beyond* for the book series, *Flashpoints: Education in the Age of Interconnection*, edited by Baker, Saari, and Prasad for Routledge. This project is to revision the current rhizomatic "ideal" and de-historical linking of culturally distinct

onto-epistemologies, worldviews, languages, spatiotemporalities, and styles of reasoning in social science and educational research. Such rootless “ideal” connectivity has engendered trans-philosophical flashpoints and ethical difficulties for researchers, practitioners, and policymakers in education and beyond. With China as its locus of analysis, my ongoing book project unpacks the technologies of interconnection, broadly conceived, in (re)producing onto-epistemic contestations and clashes on the issues of meaning, being, reality, language, knowledge, body, and governance, and in educational settings. Using edusemiotics, new materialism, and Foucault’s governance theory as three expressions of the latest cultural-philosophical flashpoints, it hopes to provide a healthier intellectual interconnection, historicizing the dynamic collisions, confrontations, and merging of trans-philosophical flashpoints as events, illustrating the distinct worldviews, cosmologies, languages, spatiotemporalities, and rituals not as essentialized “ideas”, but in their fluidities (see, [Zhao, forthcoming](#)).

I have also expanded such transnational “epistemic” connectivity issue in two other projects – one book project for Routledge (see, [Zhao et al., 2022](#)) and one journal special issue for *Discourse* (see, [Zhao and Popkewitz, 2022](#)). Both projects bring into international curriculum and educational studies a relational “epistemic” network to re-examine knowledge translation within and among nation-states on a global landscape both historically and at the present. It explores the trans-national and trans-cultural negotiation, translation, and confrontation of varied historical-cultural-philosophical *epistemes* in their material implications within the (curriculum) modernization and globalization movement. Specifically, it explicates the epistemological groundings of varying forms of knowledge in crisscrossing as well as the historically specific conditions for the happening of “epistemological colonialism”, i.e., the hidden or explicit marginalization, eclipse, or overriding of one cultural form of knowledge by another on a global landscape, whether within or cross nation-states, and in connection with economic development and power struggle.

Other than that, I recently finished a book project with Daniel Trohler on *Euro-Asian Encounters on 21st-Century Competency-Based Curriculum Reforms: Cultural Views on Globalization and Localization* (see, [Zhao and Trohler, 2021](#)). Provoked by the on-going competency-based curriculum reforms on a global scale, this edited book examines where such global framework as OECD’s core competency definitions is rooted in and how they are borrowed, resisted, and/or recontextualized in some European nation-states with a Protestant educational-cultural heritage and some Asian nation-states with a Confucian educational-cultural heritage. Put differently, this book foregrounds the roles that varied mechanisms, say, history, culture, religious attitudes, ideology, and state governance, play in nation-states’ recontextualization of global curriculum policies and practices beyond a simplistic and dualistic globalism/power and nationalism/resistance dynamic. In so doing, it provides a global context to better understand individual nation-state’s on-going curriculum reforms and school practices. The selection of the two educational-cultural systems and wisdom of thought, the Protestant and the Confucian, also provides a unique springboard for conducting international curriculum studies beyond the usual confinement of geopolitical nation-state construct.

“Returning to the historical-cultural-philosophical foundations” to better cut into the present schooling practices in China and beyond

In the 50th anniversary issue of *Journal of Curriculum Studies*, [Deng \(2018\)](#) critiques that the curriculum field has somewhat been transformed into a kind of “postmodern cultural studies” that encompass an array of topics but neglects its core subject – practice and the inner work of schooling as a complex institution as its basic concern. While it is an opportune critique, I have been trying to translate my historical-historical-philosophical unpacking of curriculum knowledge into classroom teaching and learning, as an historical inquiry is to better cut into the present. Specifically, I explore the possible challenges, new openings, as well as intervention strategies, of embodying and translating the historical-cultural-philosophical thinking into classroom practices to enhance teaching and learning efficacy at the nexus, and as the (dis)assemblage, of tradition and modernity, East and West.

For example, China’s Ministry of Education launched a “classroom revolution” in 2017 that intends to transform current teacher-centered pedagogies and rationales with brand-new philosophies and technologies and hopes to cultivate students’ higher-order thinking skills. However, my observations of China’s frontline teaching and learning over the years pinpointed an epistemological flashpoint that enthralls teachers in their pedagogical reform. Namely, teachers commonly believe that their commonly adopted teacher-centered didactic (old Confucian) pedagogies are not conducive to students’ thinking cultivation and yet they have no other choice but to continue such practice due to testing pressure and institutional constraints. Nevertheless, my cultural-historical-philosophical unpacking of China’s (Confucian) curriculum knowledge helps disrupt such naturalized (mis)belief and some stereotypical (mis)understandings of Confucian teaching and learning, demonstrating instead an epistemological compatibility between Confucian didactic (which I term “do-after-me”) pedagogy and students’ thinking cultivation.

Using this epistemological re-understanding as a new entry point for classroom research, I did a case study with two math classrooms through an ethnomethodology approach and as a teaching intervention project on two layers (see, [Zhao, 2020b](#)). First, I explicate the presumed epistemological clash that the teachers have between the claimed Confucian pedagogies and critical thinking. Second, I mobilize an unrecognized yet educative ‘repair moment’ in a Chinese math classroom, generating a possible re-conjoining between teachers’ self-despized yet habitually implemented ‘Confucian do-after-me pedagogy’ and critical mathematical reasoning. Through observations and interviews, I illustrate how this ‘repair moment’ can become pedagogically significant, overturning teachers’ naturalized (mis)belief, redeeming them from the moral-practical dilemma, and cultivating their critical pedagogical consciousness. In so doing, I showcase how to employ a philosophical-plus-empirical paradigm to do classroom research in China and beyond.

Epilogue

The field of (international) curriculum studies, as Pinar (2008) and Deng (2018) review, has indeed grown into a rhizomatic landscape. While practice and the inner work of schooling as a complex institution are still one of the basic concerns of curriculum research, there are many ways to approach them and often times roundabouts and new starting points are needed to try to get out of the grip of the modernity-coloniality trap. As Paraskeva (2016) and others rightly aver, curriculum epistemicide has reached a “quasi-irretrievable point” (p. 3) and continued to intersect with the daily praxis of schools in and beyond the West. The happening of epistemicide, however, is not only an effect of Western power hegemony, but also anchored in a fabricated eugenic claim that Western epistemological perspective is “unique and the only cognitive possible” (p. 3). For example, OCED’s and USA’s core competencies-skills discourses have become dominant ones global wise, suppressing and overwriting alternative local discourses and forms of knowledge in developing countries. The fact that many countries in the world are still jumping on the competency-based curriculum reform wagon, hoping to go global with the more advanced West, shows that non-Western countries are still naturally taking the Western super-discourses of “competencies” and “skills” as epistemologically better, reproducing Western hegemonic curriculum knowledge (above local knowledge forms) in non-Western countries before they know it.

To suspend the “darker side” of the Western modernity-coloniality of knowledge, power, and being, we need to work toward a global “cognitive justice”, that is “sentient of the wor(l)ds behind and beyond the Western epistemological platform, wor(l)ds that are non-monolithic” (Paraskeva, 2016, p. 86). Anticipating a cognitive justice worldwide calls for a new pluri-verse of varied epistemes and cultural forms of knowledge. It pushes the theorists into a pluri-directional path, allowing us not only to start anew from the beginning of the West, but also from a new beginning point, i.e., the non-West, hoping to render visible “a silenced and different genealogy of thought” (p. 80). My research trajectory of re-turning to the cultural-historical-philosophical foundations of China’s curriculum knowledge with new implications for international curriculum research, as explicated above, is one such decolonial effort. Just as Friedrich Hölderlin says: For as you began, so you will remain.

Acknowledgment

The project was partially supported by the startup grant (No. 4045C50221204110) from Hangzhou Normal University, China.

References

- Andreotti, V., 2011. (Towards) decoloniality and diversity in global citizenship education. *Glob. Soc. Educ.* 9 (3–4), 381–397.
- Agamben, G., 2009. *The Signature of All Things: On Method*. Zone Books, New York.
- Ames, R.T., 1993. The meaning of body in classical Chinese thought. In: Kasulis, T.P., Ames, R., Dissanayake, W. (Eds.), *Self as Body in Asian Theory and Practice*. State University of New York Press, Albany, pp. 39–54.
- Ashwani, K., 2019. *Curriculum in International Contexts: Understanding Colonial, Ideological, and Neoliberal Influences*. Palgrave Macmillan.
- Apple, M., 2000. *Official Knowledge: Democratic Education in a Conservative Age*. Routledge, New York.
- Baker, B. (Ed.), 2009. *New Curriculum History*. Sense Publisher.
- Blumenfeld-Jones, D., 2016. The violence of words, words of violence: Keeping the uncomfortable at bay, a Jewish perspective. *J. Curriculum Theor. (JCT)* 31 (1), 4–12.
- Chakrabarty, D., 2000. *Provincializing Europe: Postcolonial Thought and Historical Difference*. Princeton University Press, Princeton, NJ.
- Deng, Z., 2018. Contemporary curriculum theorizing: crisis and resolution. *J. Curric. Stud.* 50 (6), 691–710.
- Foucault, M., 1973. *The Order of Things: An Archaeology of the Human Sciences*. Vintage Books, New York (Original work published 1966).
- Heidegger, 1966/1996. Spiegel interview with Martin Heidegger. *Philos. Today* 20, 268–284.
- Heidegger, M., 1967. *What Is a Thing?* (W. B. Barton, Jr and V. Deutsch Trans. with an analysis by E. T. Gendlin.) Henry Regnery Company, Chicago.
- Hébert, C., Ng-A-Fook, N., Ibrahim, A., Smith, B., 2019. *Internationalizing Curriculum Studies: Histories, Environments, and Critiques*. Palgrave Macmillan.
- Horlacher, R., 2016. *The Educated Subject and the German Concept of Bildung: A Comparative Cultural History*. Routledge, New York.
- Huebner, D., 1966/1999. Curricular language and classroom meanings. In: Hills, V. (Ed.), *The Lure of the Transcendent: Collected Essays by Dwayne E. Huebner*. Lawrence Erlbaum Associates, Abingdon-on-Thames, pp. 101–117.
- Huebner, D., 1985/1999. Spirituality and knowing. In: Hills, V. (Ed.), *The Lure of the Transcendent: Collected Essays by Dwayne E. Huebner*. Lawrence Erlbaum Associates, Abingdon-on-Thames, pp. 340–352.
- Kazamias, A.M., 2001. Re-inventing the historical in comparative education: reflections on a protean episteme by a contemporary player. *Comp. Educ.* 37 (4), 439–449.
- Kuriyama, S., 1994. The imagination of winds and the development of the Chinese conception of the body. In: Zito, A., Barlow, T.E. (Eds.), *Body, Subject, and Power in China*. Chicago University Press, Chicago, pp. 23–41.
- Maldonado-Torres, N., 2007. On the coloniality of being. *Cult. Stud.* 21 (2–3), 240–270.
- Paraskeva, J.M., 2016. *Curriculum Epistemicide: Towards an Itinerant Curriculum Theory*. Routledge.
- Paraskeva, J.M. (Ed.), 2018. *Toward a Just Curriculum Theory: The Epistemicide*. Routledge.
- Phillips, D., 2014. ‘Comparatography’, history and policy quotation: some reflections. *Comp. Educ.* 50 (1), 73–83.
- Pinar, W.F., 2008. Curriculum theory since 1950: crisis, re-conceptualization, internationalization. In: Connelly, F.M., He, M.F., Phillion, J. (Eds.), *The Sage Handbook of Curriculum and Instruction*. Sage, Thousand Oaks, CA, pp. 491–513.
- Pinar, W.F., 1999. Introduction. In: Hills, V. (Ed.), *The Lure of the Transcendent: Collected Essays by Dwayne E. Huebner*. Lawrence Erlbaum Associates, Abingdon-on-Thames, pp. xv–xxv.
- Popkewitz, T.S., 2008. *Cosmopolitanism and the Age of School Reform: Science, Education, and Making Society by Making the Child*. Routledge.
- Popkewitz, T. (Ed.), 2013. *Rethinking the History of Education: Transnational Perspectives on its Questions, Methods, and Knowledge*. Palgrave MacMillan, New York.
- Popkewitz, T., Diaz, J., Kirchgasser, C. (Eds.), 2017. *A Political Sociology of Educational Knowledge*. Routledge, New York.
- Roofe, C., Bezzina, C., 2018. *Intercultural Studies of Curriculum: Theory, Policy and Practice*. Palgrave Macmillan.
- Ryan, A., Tocci, C., Moon, S., 2020. *Curriculum Foundations Reader*. Palgrave Macmillan.

- Sousa Santos, B., 2007. Beyond abyssal thinking: from global lines to ecologies of knowledges. *Review XXX* (1), 45–89.
- Shi-xu, 2005. *A Cultural Approach to Discourse*. Palgrave Macmillan, London.
- Shi-xu, 2016. Cultural discourse studies through the *Journal of Multicultural Discourses*: 10 years on. *J. Multicult. Discourses* 11 (1), 1–8.
- Tröhler, D., 2013. *Languages of Education: Protestant Legacies, National Identities, and Global Aspirations*. Routledge, New York.
- Wa Thiong, N., 1986. *Decolonising the Mind: The Politics of Language in African Literature*. East African Educational Publishers, Nairobi/London.
- Wang, H., Eppert, C., 2007. *Cross-cultural Studies in Curriculum: Eastern Thought, Educational Insights*. Routledge.
- Wu, K.M., 1997. *On Chinese Body Thinking: A Cultural Hermeneutic*. Brill Academic Publishing, Leiden and New York.
- Wu, Z.J., 2011. Interpretation, autonomy, and transformation: Chinese pedagogic discourse in a cross-cultural perspective. *J. Curric. Stud.* 43 (5), 569–590.
- Wu, Z.J., 2014. 'Speak in the place of the sages': Rethinking the sources of pedagogic meanings. *J. Curric. Stud.* 46 (3), 320–331.
- Zanazanian, P., Nordgren, K., 2019. Introduction. *J. Curric. Stud.* 51 (6), 771–778.
- Zhao, W., 2019a. Re-invigorating the being of language in international education: unpacking Confucius' 'wind-pedagogy' in Yijing as an exemplar. *Discourse: Studies in the Cultural Politics of Education* 40 (4), 474–486.
- Zhao, W., 2019b. *China's Education, Curriculum Knowledge and Cultural Inscriptions: Dancing with the Wind*. Routledge, New York.
- Zhao, W., 2019c. "Observation" as China's civic education pedagogy: a historical perspective and a dialogue with Michel Foucault. *Discourse: Studies in the Cultural Politics of Education* 40 (6), 789–802.
- Zhao, W., 2020a. Problematizing "epistemicide" in transnational curriculum knowledge production: China's suyang curriculum reform as an example. *Curric. Inq.* 50 (2), 105–125.
- Zhao, W., 2020b. Epistemological flashpoint in China's classroom reform: (how) can A confucian 'do-after-me' pedagogy cultivate critical thinking? *J. Curric. Stud.* 52 (1), 101–117.
- Zhao, W., 2021. Toward a de-colonial language gesture in transnational curriculum studies. In: Green, B., Roberts, P., Brennan, M. (Eds.), *Curriculum Challenges and Opportunities in a Changing World: Transnational Perspectives in Curriculum Inquiry*. Palgrave Macmillan, pp. 87–101.
- Zhao, W., in press. Re-turning to the "cultural" foundations of China's curriculum reform: ICT and confucian "wind" pedagogy. In: Paraskeva, J. (Ed.), *Itinerant Curriculum Theory A Declaration of Epistemological Liberation*. Palgrave, New York.
- Zhao, W., forthcoming. Edusemiotics, (New) Materialism, and (Body) Governance: Flashpoints in 21st-Century Education in China and Beyond. To Be Included in Bernadette Baker, Antti Saari, Amit Prasad's Book Series, *Flashpoints: Education in the Age of Interconnection*. Routledge.
- Zhao, W., Trohler, D. (Eds.), 2021. *Euro-Asian encounters on 21st-century competency-based curriculum reforms: cultural views on globalization and localization*. Springer.
- Zhao, W., Popkewitz, T., & Tero, A. (Eds.), 2022. *Epistemic colonialism and the transfer of curriculum knowledge across borders: applying a historical lens to contest unilateral logics*, Included in William Pinar's Series, *Studies in Curriculum Theory*, Routledge, New York <https://doi.org/10.4324/9780429323027>.
- Zhao, W., Popkewitz, T.S., 2022. Introduction: critiquing the onto-epistemic coloniality of modernity in/beyond education. *Discourse: Studies in the Cultural Politics of Education* 43 (3). <https://doi.org/10.1080/01596306.2022.2041327>.
- Zhang, X.L., 2015. *Shangshu-Yaodian Jieshuo: Yi shi yu xiao weiyan de zhengzhi* ([Interpreting Shangshu-Yaodian: Govern-mentality with Time and Filial Piety]). Sanlian Chubanshe.

Contemporary trends in curriculum research

Berit Karseth^a and Ninni Wahlström^b, ^a University of Oslo, Department of Education, Oslo, Norway; and ^b Linnaeus University, Department of Education and Teachers' Practice, Växjö, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Increasing global interest in the policy of national school systems	74
Modes of curriculum governance	75
Centralization, decentralization, and recentralization	76
From NPM to network governance	77
Hard and soft governance in an era of new technology	78
Multilevel governance and curriculum policy	78
Curriculum, knowledge, and the view of realism	79
Social realism as a response to postmodernism	80
Experience as a concept of knowledge: transactional realism	80
Bildung—an environment of culture	81
The three knowledge concepts and curriculum	81
A concluding reflection/discussion	82
References	82

The term “curriculum” has multiple definitions. It is a contested term that reflects the historical, cultural, and epistemological characteristics of curricula, as well as institutional and political conditions. The most basic curriculum question is often formulated in terms of “what should count as knowledge” (Deng and Luke, 2008). Thus, the curriculum research area comprises substantial aspects, such as the purpose of education, the teaching content, and the knowledge expectations of the learners, as well as different approaches to address the relationships between society, knowledge, pedagogy, learning, and assessment. Researchers of curriculum theory tend to be interested in two or more of the three arenas—the societal arena, the programmatic arena, and the classroom arena—as well as the relationships between these different arenas. Thus, curriculum research includes processes of decision-making on different levels of the school system by different actors in accordance with different national traditions and regulations. During the 2000s, educational policies at both the international and national policy levels captured the interest of curriculum researchers. The curriculum research literature reminds us that there are different ways to understand and approach the relationship between the learner, the school, the school district/municipality, the state, and society. Curriculum researchers are also interested in how curriculum texts explicitly or implicitly influence power relations between teachers and students, teachers and the school as an institution, and the school and external stakeholders, such as the government, professional associations, and employee organizations.

The seminal book *Understanding Curriculum* (1995) by William Pinar and associates represents an important contribution to the contemporary history of research and theorizing in the American curriculum field. In particular, the book addresses the reconceptualization of the curriculum field to capture the curriculum as different texts representing various discourses. The book is still highly relevant almost 30 years after it was published. Curriculum can be approached in different ways: as a political, racial, and gendered text, as well as phenomenological text emphasizing the lived curriculum.

In this chapter, we explore the field of research through the lens of two curriculum theory questions: *How can we understand the organization of the school as a societal institution?* And *How can we understand the organization of knowledge in school?* These focus areas have been debated and have changed over the past two decades in Europe and other parts of the world. The first part of the chapter highlights the increasing interest in educational and curriculum policy at the international level. Even if curriculum policy is basically a national or regional responsibility, there is an interest from intergovernmental organizations (IGOs) to influence policy formation. The second section explores different modes of governing and governing trends for curricula over time. The third part of the chapter highlights the current debate on what sort of knowledge best underpins curricula and the role of realism in different knowledge concepts. The fourth and final sections present the main points in a concluding discussion.

Increasing global interest in the policy of national school systems

During the latter half of the 20th century, the term *globalization* emerged, characterized by fundamental changes in our perceptions of distance in time and space. Globalization concerns the intertwining of social relations and social activities “at a distance” with those in local contexts. In this dialectic between the global and the local, both the individual and society are interrelated to a global milieu (Giddens, 1991). The development of globalization, crossing national borders of economy and communication, has simultaneously led to the internationalization of educational discourse. It has also resulted in a policy-directed international education comparison based on an interest in policy problems and solutions. Thus, international knowledge comparisons, such as the

Program for International Student Assessment (PISA) implemented by the Organization for Economic Co-operation and Development (OECD), are concerned with a policy agenda, which is often presented as “self-evident” and with an aim to inform others engaged in similar policy definitions and policymaking (Alexander, 2001). The global trend has evoked an increasing interest in comparative research in areas such as pedagogy, curricula, and policy transformations at the intersection of the global and the national/local levels.

The phenomenon of globalization, based on an ideology of modernity, can be critically examined from at least three different perspectives (Anderson-Levitt, 2008). First, world-system theory admits the material and economic power in globalization but claims it has been imposed by the most dominant nations on other countries. Consequently, if there are certain international convergences in curricula across regions and nations, this is because it is in the interest of powerful nations and international organizations. Second, world-culture theorists, who base their theories on neo-institutionalism, argue that ideologies and ideas, rather than material causes, attract international collaborations between states, organizations, and individuals. Rather, it is a cultural notion of what is thought of as modern and global competition among states that motivates governments to borrow ideas internationally. Third, a critical perspective on modernity and globalization is represented by postcolonial theory. From a post-colonial perspective, ideas matter, but they matter in relation to power. There is asymmetry in the value of different ideas related to the different origins of ideas. For a policy idea to be perceived as important and influential, it needs to be disseminated through powerful and rich actors, both nations and international organizations. By contrast, those who are assumed to receive the ideas are considered to be subordinated as independent actors rather than active players based on their own knowledge and experiences. As Rizvi and Lingard (2010) noted, the pressures for educational change have emerged not only within national systems but also from IGOs, and this has been the case particularly for nations in the Global South. According to Anderson-Levitt (2008), official curricula are becoming more uniform around the world, especially regarding elementary subjects. However, when it comes to how curricula are structured, locally interpreted, and transformed into actual teaching, there are broad variations.

The process of transnational generalizing of curricula and pedagogy across borders has often proceeded as part of development programs, without much criticism. However, policies do not always travel well. The policy solutions are part of “policy selections,” characterized by the inclusion of some educational research and the exclusion of others. As a result, many education systems now encounter policies of standardization of knowledge achievements in curricula where equity is rationalized to mean sameness. According to Luke (2013), this policy movement represents an ideology of marketization and standardization supported by multinational educational enterprises and other international actors. Instead, Luke (2013) argued that policy borrowing needs to be considered from an understanding that effective education policy is often highly context-dependent and therefore cannot move freely across cultural, social, and historical traditions. In addition, the trends toward a global curriculum solution focus mainly on measurable goals, while leaving other goals, such as the democratic goals of civics, civility, language, and culture, outside the curriculum. However, these global processes do not affect all educational systems in the same way.

Rizvi and Lingard (2010) noted that policy studies and curriculum studies are usually considered separate arenas without any direct connection to each other, either theoretically or practically. However, through globalization, the gap between policy and curriculum research has been narrowed, and schooling, curriculum, pedagogy, and evaluation are currently presented in terms of solutions in a discourse of global challenges. Thus, education policy reform is now intimately linked to the international reframing of how to understand curriculum, pedagogy, and evaluation. An important feature of a globalized education policy is the notion of accountability and a major concern for testing. Increasingly, national governments and national educational actors have begun to govern the school system by using standardized tests, both nationally and internationally. Thus, it can be argued that testing has become a central element in policy regimes and that policymakers are trying to influence local school practices through various tests with implications for both curriculum and pedagogy (Rizvi and Lingard, 2010).

Modes of curriculum governance

In the modern history of curriculum development as a governing tool, countries have followed different patterns and routes. European countries such as Germany, France, and the Scandinavian countries have a tradition of state-based curriculum making (Westbury et al., 2016). In countries like the United States, Australia, and the United Kingdom, curriculum development was traditionally in the hands of local authorities and schools, and the idea of a curriculum controlled by the state was highly contested. In the last 30 years, developments have indicated that curriculum has become a national concern and a part of the jurisdiction of state governments in most of these countries (Biesta and Priestly, 2013, pp. 230–231).

Despite the convergence toward recognizing the national level in curriculum making, historical trajectories and institutional settings show that the emergence of a national curriculum has been legitimated in different ways. As shown by Hopmann (2003), two different approaches to curriculum control and evaluation were observed by the end of the 20th century. Continental European and Nordic countries followed a process-centered approach characterized by the enhancement of the independence of the teaching profession and where *Didaktik* represents the common core of professionalism. The Anglo-Saxon tradition followed a product-centered approach where efficiency becomes the core of the profession and “the best teacher is the one who gets his students to ‘stand up and deliver’ at the right time—whatever the cause for their ability to do just that might be” (p. 472). While the curriculum control within a process-centered approach is input-based with plans and frameworks as the main instruments, the product-centered approach is output-based with learning outcomes and standards as the main instruments.

These differences reflect how countries have dealt with processes of centralization and decentralization in the policy of curriculum making. This issue has been a concern throughout modern schooling. The concept of governance, although slippery, is relevant for understanding the policy and politics of curriculum making. Governance is “the interactive processes through which society and the economy are steered toward collectively negotiated objectives” (Ansell and Torfing, 2016, p. 4). This definition underlines the importance of interaction with regard to the exchange of ideas, resources, and competences that are required.

Centralization, decentralization, and recentralization

Contemporary research on curriculum governance has paid great attention to the processes of centralization and decentralization. In the 1970s and 1980s, many Western countries promoted decentralized modes of governance. One line of argument called for replacing a more central administrative model of governance with a model that gave the profession and the school a more influential role (Beare and Sturmann, 1991). Another approach to decentralization called for a participatory model that involved the profession and the educational administration while inviting a wider range of participants, such as parents, local communities, and social movements, to play roles in the decision-making.

Curriculum reform initiatives in the 1980s and 1990s in the Nordic countries followed trajectories of decentralization that moved some of the power over the decision-making processes to the local level and the schools (Lavonen, 2017; Lundgren, 1990; Mølstad, 2015; Alvunger and Wahlström, 2021; Wermke and Prøitz, 2021). However, the story also shows differences in the trajectories of Nordic countries. The curriculum reform in Norway in the 1990s represented a recentralization in which the regulation of the content of the curriculum became more detailed than in the former curriculum. This change reflected the idea of a strong state combining central management with social democratic confidence in the state as a guarantor for the fair distribution of welfare services (Telhaug et al., 2006). In Sweden, the strategies chosen in the 1990s included the marketization of the school system with the introduction of school vouchers (Wahlström, 2014; Wermke and Prøitz, 2021). According to Wahlström (2014), the 1991 school reform of decentralization in Sweden led to a denationalized conception of education characterized by the state’s withdrawal from the national policy arena in favor of transnational policymaking and a displacement from a more corporative to a more state-centered approach. The state’s presence was strengthened in terms of governing through accountability and results.

Weiler (1990) described three commonly advocated rationales for greater decentralization:

1. Decentralization as the redistribution of power.
2. Decentralization as a means of enhancing the efficiency of educational governance.
3. Decentralization as a means of enhancing a culture of learning.

Weiler (1990) argued that these rationales encounter substantial difficulties despite their claims. The notion of decentralization as a redistribution of power may be counterproductive to effective governing of education and attempts to secure equality and a unified educational system for all. The second notion of decentralization as a means of enhancing the efficiency of educational governance, represents a critique of public administration and follows ideas from new public management (NPM) such as a belief that decentralized system can utilize available resources more efficiently. Weiler (1990) also questioned this rationale, as it appears to depend on certain premises such as possibilities for different local stakeholders to have an influence on the educational system. The third notion of decentralization as a means of enhancing a culture of learning means that the local context and cultural arrangements are considered when determining how to regulate education. However, Weiler (1990) argued that the claim for learning may be less oriented to the particularities of the local culture; rather, global terms that include more universalities and traveling ideas characterize the discourses of learning.

Despite these criticisms of the three rationales, an overt commitment to decentralization in educational governance has important political utility. Weiler (1990) noted that “it becomes an important vehicle for both the management of social conflict and for what I have called ‘compensatory legitimization’” (p. 51). Weiler (1990) continued as follows:

The frequent wavering between centralized and decentralized modes of policy behavior – or, to be more exact, between decentralization rhetoric and centralization behavior – may very well have its cause in this difficult task of walking the fine line between the conflicting imperatives of control and legitimacy (p. 55).

Weiler’s (1990) contribution seems relevant to the ongoing debate on decentralization. Walking the fine line means balancing political and professional power as well as central and local power. Furthermore, the distinction between rhetoric and behavior is important when studying modes of governance.

Based on data across 33 countries (mostly OECD members and European countries) from the PISA surveys in 2000 and 2015, Jeong and Luschei (2018) concluded that, despite the narrative about decentralization, school governance appears to be moving toward centralization, most dramatically in the area of curriculum and instruction. Although the authors pointed to some important limitations, such as the fact that the measure of teacher responsibility was based on questionnaires administered to school principals and that perceptions of autonomy may not be a perfect proxy for actual autonomy, they asserted that centralized governance has resulted in teachers losing their power and responsibility in educational classroom decisions. Conversely, principals are

maintaining or increasing their responsibilities (p. 299). This point is interesting because it calls for more power-sharing arrangements between governments and principals. Jeong and Luschei (2018) asked whether this trend is a type of “neo-centralization.”

In Nordic research and elsewhere, scholars have highlighted municipalization as a central government strategy (Verger et al., 2016; Wahlström, 2020). While this approach can be viewed as a strategy of decentralization at the local level, it may also mean less power for teachers and a restriction of professional autonomy.

The arguments above, which have been used to advocate for greater decentralization, can, as previously mentioned, be recognized in the arguments put forward within the discourse of NPM.

From NPM to network governance

The NPM policy approach that came onto the scene in the 1980s altered the governance of educational institutions by downsizing government regulations, decentralizing the governing of schools to local levels, introducing results-based performance indicators, initiating a process of commodification of the public sector, and implementing privatization and consumer choice (Theisens et al., 2016; Verger and Curran, 2014). Following NPM policies, the state strengthens its role as a regulator, evaluator, and distributor of incentives to providers (Verger and Curran, 2014). While some of the policy measures connote neoliberal ideas, others, such as curriculum content and school-based curriculum development, do not.

Several researchers have pinpointed the recent transformation of network governance that extends NPM policies. They described these changes as a re-regulation of society through the ways in which these networks of actors make use of new governance tools (Helgøy et al., 2007). Theisens et al. (2016) emphasized that the state is still central in the political context, with public sector organizations operating as knowledge providers. Reform policy today emphasizes state-centered solutions, while acknowledging the mix of governance mechanisms and institutional complexity that characterizes the public sector and has generated renewed interest in collaborative governance (Karseth and Sivesind, 2022). As a result, policymaking involves multiple actors, including international actors, non-state actors, interest groups, and professional and governmental actors, meaning that governments are under constant pressure to interact and negotiate with and mediate among a plethora of policy actors (Steiner-Khamisi et al., 2022). The arguments put forward in policymaking call for broader involvement and more deliberative processes.

Although the picture of policy governance emphasizes network, assemblages, and temporality, state bureaucracies are still operative institutions that govern reform processes in the Nordic countries, but they are “less driven by politicians, more planned, and less contested by the unions, and there is more public involvement” (Greve et al., 2016, p. 201). This condition is certainly true when it comes to national curriculum making, as evidenced by the recent curriculum renewal in Norway. The incremental reform that came into effect in 2020 continued with components introduced by the fundamental reform in 2006, such as competence aims and basic skills. In addition, the 2020 reform (LK20) introduced descriptions of core elements in different school subjects, three interdisciplinary topics, and formative assessment. The new curriculum framework also entails a new core curriculum (Steiner-Khamisi et al., 2020; Sundby and Karseth, 2021).

The educational agency, the Norwegian Directorate of Education and Training (DET), orchestrated the process, and the concept of co-creation became a signifier and a mobilizer in handling the complexity of curriculum making (Karseth et al., 2020). The work was organized by establishing a partnership group headed by the Minister of Education and consisting of representatives in leader positions from different teacher and school leader associations/unions, the Norwegian Association of Local and Regional Authorities, the School Student Union of Norway, the Sami Parliament, and the Higher Education Council. Representatives from these organizations constituted a “working group” according to the DET. Some groups were appointed to produce drafts of the core elements of the school subject, while others drafted the school subject curricula. The resulting work was sent out for public responses—and several actors were invited to pose their arguments and suggestions at an early stage—in addition to formal public hearings.

We argue that the process follows the idea of collaborative governance. From a critical perspective, we question whether a co-creation process hides how power unfolds. First, the power of the state is made invisible because of an impression of participation, although the premises for the work have already been established. Second, what appears to be a democratic and open process is, in fact, rather closed and opaque with regard to final decision-making. The capacity to join different arenas to give feedback and suggestions is also uneven. The third point is the types of policy instruments offered by the DET (progress plans, lists of approved verbs, templates for completion, requirements for digitization) that define the space for not only how to write and structure the curriculum but also how to interpret it.

These issues indicate that DET acts as a main vehicle for policy implementation but that it also plays a role in setting the policy agenda by being a core actor in the governance network of cooperation. As such, the division of labor between the Ministry of Education and its agency is rather blurred (Szescilo, 2020). Furthermore, DET also plays an important role in connecting national and international policy agendas by participating in networks orchestrated through the OECD. As previously indicated, education policy reform is closely connected to the international reframing of how to understand curriculum, pedagogy, and evaluation. Through the project known as *The Future of Education and Skills 2030: OECD Learning Compass 2030* (OECD, 2018), the OECD has taken a new step in governing education globally, with a call for redesigning curricula to meet the challenges of the 21st century. To design and implement a future-oriented curriculum, the different national agencies are important players to legitimate and act as knowledge brokers.

Ministerial agencies, such as the DET, are defined as organizations that are formally part of public administration while enjoying considerable autonomy. In the field of curriculum, the DET plays a significant role in not only implementing a curriculum reform but also orchestrating the process with regard to collecting evidence, knowledge, and information and acting as a mediator and interpreter between international reform initiatives and the agenda of national curriculum policy. Moreover, in the case of Norway, the DET is also instrumental in producing digital resources and infrastructure that can be used at the school level. These instruments are not part of legal curriculum regulation and can be described as soft governance tools.

Hard and soft governance in an era of new technology

In the educational sector, researchers have identified a line to walk between governing through hard laws and soft laws. They have recognized a shift from “formal legal instruments to more flexible forms of steering and involving decentral organizations and actors in implementation” (Theisens et al., 2016, p. 468). This trend implies the introduction of more subtle “soft” forms of control linked to the increased use of evaluation and quality control (Hudson, 2007, p. 278).

The use of soft governing tools has been central in an attempt to standardize higher education in Europe. Approaches to regulation are through a political convergence that includes implicit discursive mechanisms, such as language used, knowledge making, and meaning making in general (Le Bianic and Svensson, 2008, p. 573). Despite the lack of hard governance in the form of legally binding laws, the European Union (EU) and the Bologna Process represent powerful driving forces for the processes of national decision-making through coordination procedures (e.g., meetings and recommendations), benchmarking, and monitoring activities (e.g., reviews, reports, and scorecards), and guidelines (e.g., procedures, templates, and checklists). These are typical examples of soft laws providing clear advice to national governments and higher education institutions that normally have been left to academics (Karseth and Solbrekke, 2016). Such advice might be impossible to ignore if one wants to operate under the Bologna Process. The EU policy rhetoric opens the door for agents to develop guidelines and best practice examples that are difficult to ignore at national or institutional levels (Veiga and Amaral, 2012).

Likewise, state authorities’ use of soft digital technology is not only to support learning in schools but also to serve governance purposes in terms of what can be called e-governance (Sivesind and Skedsmo, 2020). School leaders and teachers are increasingly expected to use digital software to collect and distribute digital data to confront challenging issues in curriculum, assessment, and school development. In this sense, digital platforms and software have become necessary to comply with governmental expectations of performance measurement. Digital technologies are now providers of evidence and important actors in identifying what best practice is and what it should be (Lunde and Gunnulfsen, 2020, p. 197). Recent research on digital education governance has leaned on actor-network theory, in which digital curriculum instruments are explored as actors (Lunde and Ottesen, 2021).

Multilevel governance and curriculum policy

As previously mentioned, curriculum governance is complex and takes place at different levels and sites. Researchers from different educational fields of study have drawn attention to understanding intergovernmental relations and the interplay of different levels in the context beyond the national or federal state, as well as the changing dynamic of the state–society relationship (Bache et al., 2016).

From a new institutional theory approach, Meyer (2007) argued that curriculum reform worldwide follows common ideas. Based on an analysis of curriculum research in several countries, he stated that modern world society builds on the idea of “an expansive conception of the individual human person, being a member of a human society as a whole rather than principally as the citizen of a nation-state” (p. 264). Although curriculum researchers have identified significant variations in the implementation and management of reforms, these reforms are not legitimated by references to national structures (Meyer, 2007, p. 265). Hence, Meyer (2007) questioned the role of states in creating global heterogeneity through national reforms.

Researchers have criticized the arguments of convergence and have noted that similar education reforms are translated, recontextualized, and modified in national and local contexts as they spread globally. Nation-states understand and handle global reform manuscripts on curriculum policy differently and adjust their policies to meet global requests and demands (Maroy and Pons, 2019a,b; Nordin and Sundberg, 2016; Steiner-Khamisi, 2009; Wahlström, 2016). Therefore, curriculum work is embedded in a context guided by global policy initiatives, as well as national mindsets and institutional procedures and practices.

Maroy and Pons (2019a,b) offered a theoretical framework for analyzing education policy, particularly accountability policy, to capture the dynamic and interaction between a global policy agenda and deeper-rooted national institutional or political processes. According to their framework, the trajectory of education policy is the product of empirically intertwined dynamics of path dependence, bricolage, and interpretation. Furthermore,

The trajectory leading to vernacular forms of globalization (rather than a convergence of policies deriving from irresistible global discourse or forces) is also shaped by the processes of policy implementation at the intermediate level (regulatory authorities falling between the central authorities and local schools) or at the local level (p. 53).

Finally, the orientations and local meaning of the policy processes depend upon the instrumentation of this policy and how local actors use or do not use different devices and tools.

The empirical findings from Maroy and Pons' (2019b) study support another variant of NPM than the neoliberal one, and the authors argue that the two cases they analyzed (France and Quebec) follow a "neo-statist" trajectory that implies strengthened state power and administrative regulation.

Maroy and Pons (2019a,b) analyzed accountability policy and not curriculum. However, their framework also offers theoretical and conceptual lenses to capture the multiplicity of curriculum governance. Mathou (2022) analyzed the curriculum in the same two cases and concluded that there is not necessarily a convergent movement between competency-based curriculum reforms, NPM, and neo-liberal policies. She highlighted the ever-changing dynamics and complexity underlying curricular policies. Future research needs to examine the processes of recontextualization between and within transnational and national policies and to address how school subjects are defined and laid out in the formal curriculum as well as in the classroom.

Even if curricula reforms differ between countries and different countries respond differently to international debates and influences, the question of curriculum content and its foundational perspectives can hardly be avoided for those interested in curriculum research. Thus, a topical debate is what kind of knowledge the curriculum should be based on. The OECD project *The Future of Education and Skills: Education 2030* (OECD, 2018, p. 3) claims that the organization is "committed to helping every learner develop as a whole person, fulfill his or her potential and help shape a shared future built on the well-being of individuals, communities, and the planet." The OECD suggests the concept of *competences* as the best way to prepare students for a rapidly changing world. The concept of competence is a broad concept that includes knowing, acting, and valuing. However, the concept of competence is based mainly on policies for contemporary society rather than on a deepened interest in theories of knowledge or knowledge traditions. To bring a well-grounded knowledge debate into the educational sphere, there is now an increasing need to explore different knowledge traditions and their basic assumptions about how individuals can develop knowledge about the world. In the next section, three different traditions of knowledge are discussed, based on their different basic assumptions and their potential implications for school knowledge formulated in the curriculum.

Curriculum, knowledge, and the view of realism

In the second section of the chapter, the central interest is the current discussion on the concept of knowledge and its implications for a knowledge-based curriculum. At the center of the debate is humans' perception of reality; that is, how humans can gain reliable knowledge about the world.

With the publications of *Knowledge and Control* (1971) and *Bringing Knowledge Back In* (2008a), the British sociologist of knowledge, Michael Young, has played a prominent role in debates on knowledge and curriculum. In the former text, Young (1971) noted that the questions of how knowledge is selected, organized, and assessed in educational institutions were largely neglected by sociologists at the time. Young (1971) insisted that education is not a product; it cannot be compared to goods manufactured in a factory. Instead, education is about "selection and organization from the available knowledge at a particular time which involves conscious and unconscious choices" (Young, 1971, p. 24). From such a starting point, studies on methods of assessment, selection, and organization of knowledge, as well as the principles underlying them, become central. Moreover, according to Young (1971), sociologists have overlooked the process of transmission of knowledge. The space where the transmission of knowledge from the social culture of society to the new generation is taking place is the curriculum. Thus, the curriculum became a central interest for studies of stratification of knowledge and social control.

In this "new sociology" in the 1970s, in which Michael Young and Basil Bernstein in England and Pierre Bourdieu in France were significant actors, students' educational failure was less related to genetic theories on intelligence and more to socially controlled cultural transmission. The "new" studies in sociology focused directly on the content of education and on the school's internal organization of the curriculum content to different groups of students. The new "interpretive paradigm" emerged in opposition to an earlier "positivistic" and "functional" paradigm. A main critique was that, until then, curriculum researchers had taken the problems formulated by the educators for given rather than formulating their own problems for research on curriculum and pedagogy. While interpretive sociology dominated educational sociology in Great Britain during the 1970s, conflict theory emerged as the dominating critical research approach in educational sociology in the US (Karabel and Halsey, 1976).

A focus in both approaches was on education and differential power. In its early phase, education was seen simply as reflecting the dominating relations of economic and social forces in society. The critical approach concerned the ways in which national school systems and curricula reproduced prevailing power relations. Over the years, a more complex set of contradictory relationships has been recognized in critical curriculum studies, influenced by other traditions such as cultural studies, and feminist and postcolonial theories (Apple, 2003). Drawing on Max Weber, the starting point in critical studies in education was the insight that dominant groups in society have the capacity to impose their own educational and cultural ideals on schools. These ideals become reflected in both what is taught in school and how it is taught. Thus, what counts as knowledge in school is ultimately a political question, even when there seems to be a social consensus (Karabel and Halsey, 1976).

Against this background, Young's book *Bringing Knowledge Back In* attracted a great deal of attention when it was published in 2008. Young was concerned by what it really means that knowledge is "social." In the book, he takes a critical stance to the post-modern tradition, which "appear to provide powerful support for the cultural demands of subordinate groups" but "by arguing that knowledge is inseparable from how it is constructed, they cannot avoid the conclusion that all knowledge, whether based on

professional expertise, research, or the experience of particular groups, is of equal value" (Young, 2008a, p. 22). About 40 years after the book *Knowledge and Control* was published, with its warnings of overlooking the connections between school knowledge represented in curriculum and the social structure of society by taking the cultural content represented in education for granted, Young (2008) now asked himself what he meant more exactly when claiming that knowledge is social.

Young (2008a) is critical of two dominant currents of thought in education. First, there is the longstanding idea termed neo-conservatism, which emphasizes a certain content in the curriculum representing the respective canons of the subjects. Although neo-conservatists' focus of canonical texts, neo-conservatists are not interested in epistemological issues in the first place but are rather driven by the idea that the traditional discipline required in learning subject knowledge also promotes traditional values in society. Second, neo-conservatism has been challenged by a more recent broad discourse on education that Young (2008a) terms "technical-instrumentalism." This conception of education takes its starting point in the need of the industry for skilled labor and the need of the nation to be economically competitive. Within this view of curriculum, education and knowledge are not an end in themselves but rather a means to achieve other goals. Most critical, however, was Young (2008a) of postmodernism and the "linguistic turn," which he believes has led to relativism between different forms of knowledge.

Social realism as a response to postmodernism

Young (2008a) criticized postmodernists for claiming that academic curriculum essentially relies on arbitrary and general assumptions about knowledge and culture and that the curriculum thereby contributes to the perpetuation of social inequalities. When arguing that "knowledge is inseparable from how it is constructed" postmodernists end up with the conclusion that all knowledge "whether based on professional expertise, research, or the experience of particular groups, is of equal value" (Young, 2008a, p. 22). The problem, Young (2008a) argued, is that since postmodernists lack a theory of knowledge, they cannot do much more than disclose power relations in the policy of curricula without offering any real explanations. He considered postmodernists to have a reductionist way of understanding knowledge by treating knowledge as a standpoint rather than a fact. According to Young (2008a), postmodernists highlighted silenced and excluded voices and equated the experiences of these groups with already-established knowledge. From this relativism among different sources of knowledge, Young (2008a) claimed that several problems related to the sociology of knowledge and knowledge as a base for curriculum occurred. One problem is the upholding of the intrinsic autonomy of knowledge. A second problem is that if knowledge is reduced to being the standpoint of any social group, there is no basis for the selection of knowledge in the curriculum. A third problem is that reducing knowledge to experience means that there is no possibility of generalizing knowledge beyond different social situations. The core of Young's critique was directed at the concept of experience, because if experiences of different social groups should represent actual knowledge, there is no longer any sociology of knowledge but only a sociology of knowers, leaving school curriculum without any logic of what should be included and what for.

To solve the problem of holding onto the social character of knowledge without falling into relativism, Young (2008a) suggested an approach to knowledge, which he termed *social realism*. A main argument in social realism is that a social theory can constitute the basis for claims of truth and objectivity, provided that the processes and codes through which knowledge is produced are strictly defined. A second main argument is that some knowledge is objective because it transcends the social and historical conditions of its production. A third basic argument is that social realism acknowledges the precedence of internal cognitive conditions for the production and acquisition of knowledge over external social processes in society. The social realist approach to knowledge and curriculum is social in that it recognizes the role of human agency in the production of knowledge. It is called "realism" because social realists claim that due to the specific processes that are valid for the production of knowledge, knowledge is independent of context. This context independence is what characterizes the difference between knowledge and common sense. Since knowledge is socially and historically constructed, it is always temporary and fallible.

These basic assumptions forming the concept of social realism as an approach to knowledge and curriculum have subsequently been followed up in a series of articles and books (e.g., Young, 2008b, 2010, 2013; Young and Lambert, 2014; Young and Muller, 2010). It has also led to a broad debate within curriculum theory on the concept of knowledge and its relationship to curriculum and school subjects. Some of the more prominent conceptions of knowledge that have become salient in this debate are the concept of transactional realism developed by John Dewey within the American philosophy of pragmatism, and the concept of *Bildung* based on a German historical and philosophical tradition from the beginning of the 19th century.

Experience as a concept of knowledge: transactional realism

Instead of considering the concept of "experience" as something strictly subjective and distinguished from the domain of knowledge, as Young (2008a) does in social realism, Dewey puts experience at the center of his intersubjective and transactional understanding of knowledge. Dewey's conception of philosophy is often thought of as instrumental, where knowledge is sought for, not for its own sake, but for the sake of action. Sleeper (2001) pointed out that this instrumentalism leads to an understanding of action itself as instrumental. Action becomes a means of social change. Sleeper (2001) denotes Dewey's pragmatism "to be a radical form of realism—a transactional realism" (p. 3). Both knowing and thinking represent actions that involve independent reality. "Even reflection is a means of changing, or reconstructing, the world" (Sleeper, 2001, p. 3). Central to this transactional conceptualization of human action and knowledge is the idea that thinking as a form of action is an important force in social change.

Dewey (1916/2008) took his starting point in the individual's interactions with the environment. The environment is the specific part of the physical and social surroundings with which humans actively interact; thus, the environment includes selectivity. This transactional approach between humans and their environment, based on human experiences in terms of knowledge about the environment, transcends the dualism between subject and object. The individual (the subject) is always already from the beginning part of a world full of objects, which means that he/she is always already in contact with reality. The question of attaining knowledge is thus not about bridging a knowledge gap that separates humans from objects but about distinguishing and examining the various objects and phenomena with which humans interact in their environment. Thus, a main concept in Dewey's theory of knowledge is the concept of experience. Experience is, in Dewey's meaning of the term, "the way in which living organisms are connected with reality" (Biesta and Burbules, 2003, p. 28). Experience is not a process going on inside the subject; instead, it is an interactive process which "changes in some degree the objective conditions under which experiences are had" (Dewey, 1938/1991, p. 22). Thus, an experience is both internal and objective. Experience is transactional because of its interactive character and reciprocal impact. Through the reciprocal impact between humans and their surroundings, experience is transactional; experience incorporates the world in human knowledge (Levine, 2019). Getting knowledge comprises the stages of (a) defining the problem, (b) inquiring into relationships of different possible solutions and consequences, and (c) reflecting on what might be the most reasonable conclusion to handle the problem (Dewey, 1916/2008).

As noted above, experience should be understood as a transactional and communicative approach to specific parts of the environment with which individuals interact. Experience is instrumental in developing human knowledge and, thus, being able to handle both social and physical problems in society. However, it is not instrumental in the sense of finding the best practical solution. Instead, both "practical" and "instrumental" are used in a broader meaning than are usually attributed to the terms. Within pragmatism, theories and intellectual knowing are viewed as parts of practical and instrumental processes aimed at gaining new or deeper meaning of an object or a phenomenon. Already achieved knowledge contributes to the quality of successful experiences.

Bildung—an environment of culture

The idea of Bildung has its roots in the neo-humanist tradition flourishing in Germany in the period between 1770 and 1830. In English, Bildung can be understood in terms of "culture" and "cultivation." Wilhelm von Humboldt, a German philosopher and educationist, held a pioneering role in the political and intellectual history of *Bildung*. A fundamental idea of von Humboldt was that humans need to be both contemplative and active. The noun *Bildung* is derived from the verb *bilden*, "to form." For their activities, people need the world to live their lives and gain knowledge, but they also need the world for their inner cultivation and development. While the world is influenced and understood through human thought and action, it is also the source of the formation of the inner moral and intellectual fulfillment of the individual. Embedded in the theory of Bildung is the dialectic between human activities in the external world and inner reflections on these activities within the individual. This dialectic is important to uphold for the individual not to lose oneself in a fragmented world. Thus, in the case of Bildung, subject knowledge in a single subject is insufficient. It is only when the connections between different knowledge fields transform and extend humanity in our individual minds that Bildung occurs. The individual's freedom is a prerequisite for social life and social contacts because it is only in a state of freedom that people can claim their personal autonomy that arises from their formative inner cultivation (Lüth, 2000).

Inherited in the tradition of Bildung is the insight that it is not the academic disciplines, but rather life as a whole, and the individual's part in that whole, that forms the basis of Bildung. The understanding of life in this context is culture, drawing on the philosophy of Wilhelm Dilthey and the foundations of the human sciences. The German concept of (general) Didaktik, built on a tradition of Bildung, represents the task of identifying the content contributing to Bildung through an analysis of the whole culture of life. In classical general Didaktik, academic disciplines are neither the source of content nor the criteria for achievement (Künzli, 2000). With its roots in Bildung, Didaktik is characterized by elements of self-determination, independence, freedom, and responsibility on the one hand, and humaneness, world, objectivity, and the general on the other.

Thus, Didaktik is about the individual in the world, with the intellectual and moral commitments that it entails. The general educational question becomes "What wider or general sense or reality does this content exemplify and open up to the learner?" (Klafki, 1995, p. 22). The content is the mediator of reality, rather than a gap to overcome, since the individual is always from the beginning embedded in a social and historical world of culture. The focus in both Bildung and Didaktik is not on what could be thought of as reality but on what is understood as human culture or human life. A central idea in the German Didaktik is that the presentation of content in itself is insufficient for education to happen. Learning, or Bildung, only takes place in the encounter between specific content that engages a specific individual. "Then, Bildung is what remains beyond this situated engagement" (Hopmann, 2007, p. 115).

The three knowledge concepts and curriculum

Taking Young's (2013) call for knowledge-based curricula seriously, it becomes crucial what concept of knowledge constitutes the foundation for curriculum. Drawing on three historical curriculum orientations (Deng and Luke, 2008), the social realist approach is in line with an orientation of academic rationalism. In this tradition, the importance of the transmission of disciplinary knowledge is emphasized, since academic disciplines are viewed as authoritative sources from which curriculum content is derived. The curriculum orientation of social reconstructionism corresponds with the concept of knowledge as transactional realism. In social

reconstructionism, social contexts, social issues, and futures are central, with the purpose of helping students reconstruct their own analyses and standpoints. This curriculum orientation comprehends school subjects as sources for meaningful knowledge if the knowledge also constitutes the basis for reflective thinking about what this means for individuals and society. *Bildung*, finally, is closer to the curriculum tradition of humanism. Within humanism, the needs and interests of the individual provide the central starting point for selecting the teaching content. The knowledge approach is experiential, aiming at developing the individual's self-realization and understanding of the world through a mutual development of both culture and the individual.

A concluding reflection/discussion

Two questions derived from curriculum theory have structured this chapter: How can we understand the organization of the school as a societal institution, and how can we understand the organization of knowledge in school? A main argument for a globalized impact on educational policy is that the international interest in education policy, even for compulsory schooling, has been growing over the last two decades, with the OECD, the World Bank and the United Nations Educational, Scientific and Cultural Organization (UNESCO) as prominent international actors. The global dimension is reinforced by the fact that different organizations advocate a similar policy message and refer to each other in their argumentation. In this way, international policy has at the same time approached issues that concern the school's governance, the content and construction of the curriculum as well as the view of knowledge underpinning national and regional curricula. The national reforms can be justified with the help of similar vocabulary, but the very design of the reform is determined primarily by the individual country's historical and social traditions regarding education and school system.

In the 1970s and 1980s, several school systems with strong state-centered control over the national school system turned to decentralization as a way to respond to the demands for pluralism in the supply of school, to limit the development of costs, and to increase local influence over the school. This trend of decentralization seems to have succeeded by a movement of recentralization at the beginning of the 2000s as the result of a policy of local accountability and a focus on outcome-based curricula. A major impact on the changing logic of governing the school sector can be attributed to the success of the launch of the NPM. Within this logic, the state governs the school by means of "soft" technologies as evaluations of knowledge results and allocation of economic resources, rather than using "hard" governance technologies in terms of laws and regulations. However, the recontextualization of international policy to national education policy and the recontextualization from national educational policy to curriculum reform are only the first steps toward actual changes in teaching activities. A third step, the process of transformation from curriculum content to actual teaching in classrooms, remains to be explored to a greater extent than has been the case to date.

A more internal debate in the field of curriculum research has emerged in recent years regarding the view on knowledge and the relationships between knowledge and the real world in terms of realism. This new interest in realism linked to the view of knowledge can be viewed as a reaction to recent forms of knowledge that claim that knowledge only represents a linguistic construction of reality without really incorporating the world outside humans in the understanding of knowledge. In this chapter, three concepts of knowledge have been explored in relation to the function of forming the basis for school curriculum: social realism based on the sociology of knowledge, transactional realism developed within the philosophy of American pragmatism, and the German philosophy of *Bildung*, with its roots in the early 19th century. All three concepts relate to the outer world, albeit in different ways. Both *Bildung* and transactional realism understand knowledge as a relation between the human and the world. In transactional realism, society and common affairs are placed at the center, while culture represents the common interest in *Bildung*. Common to both the pragmatist terms of growth and *Bildung* is the human being acting in the world, who learns and develops from their increasingly in-depth experiences. Both concepts emphasize the importance of teaching content that arouses students' interest and commitment. For social realists, the question of "what" to learn is more related to academic disciplines, and in this way, the content is more predictable. Instead, within social realism, the question of "how" becomes central—how should students get access to this content?

Finally, although the two questions addressed in this chapter point to different aspects and theoretical lenses of curriculum research, the questions need to be connected, as they are mutually dependent and deal with the phenomenon called curriculum.

References

- Alexander, R., 2001. *Culture and Pedagogy. International Comparisons in Primary Education*. Blackwell Publishers, Malden, MA.
- Alvunger, D., Wahlström, N., 2021. Understanding transnational curriculum policies and curriculum making in local municipal arenas: the case of Sweden. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald Publishing Limited, Bingley, UK, pp. 223–245.
- Anderson-Levitt, K.M., 2008. Globalization and curriculum. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, Thousand Oaks, CA, pp. 349–368.
- Ansell, C., Torfing, J., 2016. Introduction: theories of governance. In: Ansell, C., Torfing, J. (Eds.), *Handbook on Theories of Governance*. Edward Elgar Publishing, Cheltenham, pp. 1–17.
- Apple, M.W., 2003. The state and the politics of knowledge. In: Apple, M.W., Aasen, P. (Eds.), *The State and the Politics of Knowledge*. Routledge Falmer, London, pp. 1–24.
- Bache, I., Bartle, I., Flinders, M., 2016. Multi-level governance. In: Ansell, C., Torfing, J. (Eds.), *Handbook on Theories of Governance*. Edward Elgar Publishing, Cheltenham, pp. 486–498.
- Beare, H., Sturmman, A., 1991. Centralization and decentralization in curriculum and evaluation: looking both ways at once. In: Walker, W., Farquar, R., Hughes, M. (Eds.), *Advancing Education. School Leadership in Action*. Routledge, London, pp. 16–28.

- Biesta, G.J.J., Burbules, N.C., 2003. *Pragmatism and Educational Research*. Rowman & Littlefield, Lanham.
- Biesta, G., Priestly, M., 2013. A curriculum for the twenty-first century? In: Priestley, M., Biesta, G. (Eds.), *Reinventing the Curriculum*. New Trends in Curriculum Policy and Practice. Bloomsbury, London, pp. 229–236.
- Deng, Z., Luke, A., 2008. Subject matter: defining and theorizing school subjects. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage Publications, Los Angeles, pp. 66–87.
- Dewey, J., 1991. Experience and education. In: Boydston, J.A. (Ed.), *John Dewey. The Later Works, 1925–1953*, vol. 13. Southern Illinois University Press, Carbondale, pp. 2–62 (Original work published 1938).
- Dewey, J., 2008. Democracy and education. In: Boydston, J.A. (Ed.), *John Dewey. The Middle Works, 1899–1924*, vol. 9. Southern Illinois University Press, Carbondale, pp. 3–370 (Original work published 1916).
- Giddens, A., 1991. *Modernity and Self-Identity. Self and Society in the Late Modern Age*. Polity, Cambridge.
- Greve, C., Lægred, P., Rykkja, L.H., 2016. The Nordic model revisited: active reformers and high performing public administrations. In: Greve, C., Lægred, P., Rykkja, L.H. (Eds.), *Nordic Administrative Reforms*. Palgrave Macmillan, London, pp. 189–212.
- Helgøy, I., Homme, A., Gewirtz, S., 2007. Local autonomy or state control? Exploring the effects of new forms of regulation in education. *Eur. Educ. Res. J.* 6, 198–202.
- Hopmann, S., 2003. On the evaluation of curriculum reforms. *J. Curric. Stud.* 35, 459–478.
- Hopmann, S., 2007. Restrained teaching: the common core of Didaktik. *Eur. Educ. Res. J.* 6, 109–124.
- Hudson, C., 2007. Governing the governance of education: the state strikes back? *Eur. Educ. Res. J.* 6, 266–282.
- Jeong, D.W., Luschei, T.F., 2018. Are teachers losing control of the classroom? Global changes in school governance and teacher responsibilities, 2000–2015. *Int. J. Educ. Dev.* 62, 289–301.
- Karabel, J., Halsey, A.H., 1976. The new sociology of education. *Theor. Soc.* 3, 529–552.
- Karseth, B., Solbrekke, T.D., 2016. Curriculum trends in European higher education: the pursuit of the Humboldtian university ideas. In: Slaughter, S., Taylor, B.J. (Eds.), *Higher Education, Stratification, and Workforce Development*. Springer, Cham, pp. 215–233.
- Karseth, B., Sivesind, K., 2022. Conclusion: toward a renewed understanding of evidence-based policy in education. In: Karseth, B., Sivesind, K., Steiner-Khamsi, G. (Eds.), *Evidence and Expertise in Nordic Education Policy*. Palgrave Macmillan, London, pp. 409–429.
- Karseth, B., Kvamme, O.A., Ottesen, E., 2020. Fagformylensens læreplanverk. Politiske intensjoner, arbeidsprosesser og innhold [The Curriculum renewal: Political intentions, work processes and content.] Report 1 from EVA2020 [Evaluation of the subject renewal: Intentions, processes and practices]. Det utdanningsvitenskapelige fakultetet [Faculty of Educational Sciences]. University of Oslo, Oslo.
- Klafki, W., 1995. Didaktik analysis as the core or preparation of instruction. *J. Curric. Stud.* 27, 13–30.
- Künzli, R., 2000. German Didaktik: models of re-presentation, of intercourse, and of experience. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice. The German Didaktik Tradition*. Routledge, London, pp. 41–61.
- Lavonen, J., 2017. Governance decentralisation in education: Finnish innovation in education. *Revista de Educación a Distancia (RED)* (Issue 53).
- Le Bianic, T., Svensson, L.G., 2008. European regulation of professional education. A study of documents focusing on architects and psychologists in the EU. *Eur. Soc.* 10, 567–595.
- Levine, S., 2019. *Pragmatism, Objectivity, and Experience*. Cambridge University Press, Cambridge.
- Luke, A., 2013. Generalizing across borders. Policy and the limits of educational science. In: Luke, A., Woods, A., Weir, K. (Eds.), *Curriculum, Syllabus Design and Equity. A Primer and Model*. Routledge, New York, pp. 144–161.
- Lunde, I.M., Gunnulfsen, A.-E., 2020. Governance through digital formations—the case of ‘What Works’ in a Norwegian education context. In: Krejsler, J.B., Moos, L. (Eds.), *What Works in Nordic School Policies*. Springer, Cham, pp. 195–212.
- Lunde, I.M., Ottesen, E., 2021. Digital technologies in policy assemblages in Ireland and Norway: a visual network analysis. *Eur. Educ. Res. J.* 20, 193–211.
- Lundgren, U.P., 1990. Educational policymaking, decentralisation and evaluation. In: Granheim, M., Kogan, M., Lundgren, U.P. (Eds.), *Evaluation as Policymaking. Introducing Evaluation Into a National Decentralised Educational System*. Jessica Kingsley Publishers, London, pp. 23–41.
- Lüth, C., 2000. On Wilhelm von Humboldt’s theory of Bildung. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice. The German Didaktik Tradition*. Routledge, London, pp. 63–84.
- Maroy, C., Pons, X., 2019a. Theoretical framework. In: Maroy, C., Pons, E. (Eds.), *Accountability Policies in Education: A Comparative and Multilevel Analysis in France and Quebec*. Springer, Cham, pp. 53–94.
- Maroy, C., Pons, X., 2019b. Conclusion. In: Maroy, C., Pons, X. (Eds.), *Accountability Policies in Education: A Comparative and Multilevel Analysis in France and Quebec*. Springer, Cham, pp. 215–229.
- Mathou, C., 2022. Contemporary curriculum policies: globalization, recontextualization and hybridization of pedagogic models. In: Tierney, R., Rizvi, F., Ercikan, K., Smith, G. (Eds.), *International Encyclopedia of Education*, vol. 4. Elsevier (in press).
- Meyer, J.W., 2007. World models, national curricula, and the centrality of the individual. In: Benavot, A., Braslavsky, C. (Eds.), *School Knowledge in Comparative and Historical Perspectives. Changing Curricula in Primary and Secondary Education*. Springer, Cham, pp. 259–271.
- Mølstad, C.E., 2015. State-based curriculum-making: approaches to local curriculum work in Norway and Finland. *J. Curric. Stud.* 47, 441–461.
- Nordin, A., Sundberg, D., 2016. Travelling concepts in national curriculum policy-making: the example of competencies. *Eur. Educ. Res. J.* 15, 314–328. <https://doi.org/10.1177/1474904116641697>.
- OECD, 2018. *The Future of Education and Skills. Education 2030*. The Organisation for Economic Co-operation and Development.
- Pinar, W., Reynolds, W.M., Slattery, P., Taubman, P.M., 1995. *Understanding Curriculum. An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. Peter Lang, New York, NY.
- Rizvi, F., Lingard, B., 2010. *Globalizing education policy*. Routledge.
- Sivesind, K., Skedsmo, G., 2020. Norway: educational governance, gap-management strategies, and reorganizational processes of the state authorities in Norway. In: Årlestig, H., Johansson, O. (Eds.), *Educational Authorities and the Schools*. Springer, Cham, pp. 75–92.
- Sleeper, R.W., 2001. *The Necessity of Pragmatism. John Dewey’s Conception of Philosophy*. University of Illinois Press, Urbana, IL (First published 1986).
- Steiner-Khamsi, G., 2009. Knowledge-based regulation and the politics of international comparison. *Nord. Pedagog.* 29, 61–71.
- Steiner-Khamsi, G., Karseth, B., Baek, C., 2020. From science to politics: commissioned reports and their political translation into white papers. *J. Educ. Policy* 35 (1), 119–144. <https://doi.org/10.1080/02680939.2019.1656289>.
- Steiner-Khamsi, G., Baek, C., Karseth, B., Nordin, A., 2022. How much is policy advice changed and lost in political translation. In: Karseth, B., Sivesind, K., Steiner-Khamsi, G. (Eds.), *Evidence and Expertise in Nordic Education Policy*. Palgrave Macmillan, London, pp. 281–320.
- Sundby, A.H., Karseth, B., 2021. ‘The knowledge question’ in the Norwegian curriculum. *Curric. J.* <https://doi.org/10.1002/curj.139>.
- Szescilo, D., 2020. Agencification revisited: trends in consolidation of central government administration in Europe. *Int. Rev. Adm. Sci.* <https://doi.org/10.1177/0020852320976791>.
- Teisens, H., Hooge, E., Walsander, S., 2016. Steering dynamics in complex education systems. An agenda for empirical research. *Eur. J. Educ.* 51, 463–477.
- Telhaug, A.O., Mediås, O.A., Aasen, P., 2006. The Nordic model in education: education as part of the political system in the last 50 years. *Scand. J. Educ. Res.* 50, 245–283.
- Veiga, A., Amaral, A., 2012. Soft law and the implementation problems of the Bologna process. *Educ. Sociedade Cult.* 36, 121–140.
- Verger, A., Curran, M., 2014. New public management as a global education policy: its adoption and re-contextualization in a Southern European setting. *Crit. Stud. Educ.* 55, 253–271.

- Verger, A., Fontevila, C., Zancajo, A., 2016. *The Privatization of Education: A Political Economy of Global Education Reform*. Teachers College Press, New York, NY.
- Wahlström, N., 2016. A third wave of European education policy: transnational and national conceptions of knowledge in Swedish curricula. *Eur. Educ. Res. J.* 15, 298–313. <https://doi.org/10.1177/1474904116643329>.
- Wahlström, N., 2020. Tracing implications of transnational policy in curriculum events. *Curric. J.* 31, 587–604.
- Wahlström, N., 2014. *Utbildningens villkor II – en denationaliserad utbildningskonception* [The conditions for education II – a denationalized conception of education.] *Utbildning & Demokrati – tidskrift för didaktik och utbildningspolitik Educ. Democr.* 23, 77–94.
- Weiler, H.N., 1990. Decentralisation in educational governance: an exercise in contradiction? In: Granheim, M., Kogan, M., Lundgren, U.P. (Eds.), *Evaluation as Policymaking. Introducing Evaluation Into a National Decentralised Educational System*. Jessica Kingsley Publishers, London, pp. 42–65.
- Wemke, W., Prøitz, T.S., 2021. Integration, fragmentation, and complexity: governing the teaching profession and the Nordic model. In: Larsen, J.E., Schulte, B., Thue, R.F.W. (Eds.), *Schoolteachers and the Nordic Model. Comparative and Historical Perspectives*. Routledge, London, pp. 216–228.
- Westbury, I., Aspfors, J., Fries, A.-V., Hansén, S.-E., Ohlaver, F., Rosenmund, M., Sivesind, K., 2016. Organizing curriculum change: an introduction. *J. Curric. Stud.* 48, 729–743.
- Young, M.F.D., 1971. An approach to the study of curricula as socially organized knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control. New Directions for the Sociology of Education*. Collier Macmillan, London, pp. 19–46.
- Young, M.F.D., 2008a. Bringing Knowledge Back In. *From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, Abingdon, Oxon.
- Young, M., 2008b. From constructivism to realism in the sociology of the curriculum. *Rev. Res. Educ.* 32, 1–28.
- Young, M., 2010. Alternative educational futures for a knowledge society. *Eur. Educ. Res. J.* 9, 1–12.
- Young, M., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45, 101–118. <https://doi.org/10.1080/00220272.2013.764505>.
- Young, M., Lambert, D., 2014. *Knowledge and the Future School. Curriculum and Social Justice*. Bloomsbury, New York.
- Young, M., Muller, J., 2010. Three educational scenarios for the future: lessons from the sociology of knowledge. *Eur. J. Educ.* 45, 11–27.

Realism and theorizing curriculum—the role of knowledge

Leesa Wheelahan, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	85
Critical realism	85
Depth ontology: domains of the real, actual and empirical	86
Similarities and differences between the natural and social worlds	87
Intransitive domain (the natural and social world) and transitive domain (bodies of knowledge)	87
Knowledge of the powerful and powerful knowledge	88
Implications of critical realism for curriculum	88
Archer's morphogenetic social realism	89
The social realist school in the sociology of education	90
Social justice and social and epistemic access to knowledge	91
Horizontal and vertical discourses	91
Transactional realism	92
Conclusion	93
References	93

Introduction

Critical realism and social realism have emerged as important approaches within the social sciences and consequently in theorizing the nature of curriculum. Curriculum is inevitably shaped by what we understand the world to be, the way we understand the similarities and differences between the natural and social worlds, how we understand human agency, the nature of knowledge, and how we access knowledge. Critical realism emerged in the 1970s and social realism somewhat later in response to positivism and empiricism on the one hand, and on the other, different relativist approaches now covered under the broad heading of post-modernism (Bhaskar, 2008 [1975]; Bhaskar, 2011a). Critical realism and social realism are often used interchangeably, but it is important to distinguish the different ways in which they are used. Critical realism is a philosophy that has implications for Margaret Archer's social realism (also called the morphogenetic approach) and the social realist school in the sociology of education, but both are elaborated theoretical frameworks that draw on, though are not reducible to, critical realism. Moreover, Archer's social realism has affinities with, but is not the same as, the social realist school within the sociology of education. And, both these approaches differ from transactional realism which is inspired by Dewey's pragmatism. It is important, as Basil Bernstein (1999) explained, to understand the different "languages" within sociology and other social sciences to gain access to the different conceptual systems of meaning within each.

This article first explores the philosophy of critical realism and the ways in which it "under-labors" the social sciences (Bhaskar, 1998b), and in particular, the way in which it can be used to theorize curriculum, including the goals of curriculum. The second section explores the social realism of prominent sociologist Margaret Archer (1995). The third section discusses the social realist school within the sociology of education which is primarily concerned about the place of theoretical knowledge in curriculum (Young, 2008; Maton and Moore, 2010; Moore, 2013; Muller, 2000; Wheelahan, 2010). The final section discusses "transactional realism" (Biesta, 2014; Carlgren, 2020). This article provides an introduction into these different realist approaches, while it isn't able to enter into deep exploration about the debates between and within the different approaches. As with any substantive philosophy and social science, there is robust debate between and within these approaches and this is one way that knowledge develops and grows. The aim of this article is to provide a starting point for further investigation into these different realist approaches.

Critical realism

All social theories (such as social realism) are premised on a metaphysical premise about the nature of the world, society, and human agency. That is, our theory of ontology (what the world is like) shapes our theory of epistemology (the nature of, and how we gain access to, knowledge about the world) (Sayer, 2000). Sayer (2000, p. 78) explains that "Critical realism is a philosophy but its main focus is on ontology, not epistemology". There are three central propositions within critical realism: first, the real world exists independently of our knowledge of it; second, our knowledge of the real world is fallible; and third, there are grounds for judging some accounts as better than others (Bhaskar, 2008 [1975]). Critical realism is a materialist philosophy because it argues that our knowledge of the world arises from our practices in it. We find that things work or don't work, and we modify our understandings as a consequence. As we will see, things aren't quite so simple, but the notion of knowledge arising from practice is an important cornerstone.

Depth ontology: domains of the real, actual and empirical

Bhaskar (2011a, p. 146) explains that the world is “structured, differentiated and changing”. Planets, mountains, global warming, and Covid-19 exist, and so do societies, capitalism, nations, class, racism, misogyny, schools, and conspiracy theories. They exist regardless of whether we believe they exist. Critical realists argue that the natural and social worlds are complex and stratified. For example, Collier (1998b, p. 263) explains that:

Everything is governed by the laws of physics; since a few hundred million years ago, some of those but not all have been governed by the laws of biology; for a few hundred years, some but not all of those have been governed by the laws of capitalist economics, and so on.

This depth ontology allows critical realists to examine the interplay between different “levels” and objects (for example, the laws of physics, capitalist economics, factory production and strikes) (Collier, 1998b) to understand the emergent outcomes that ensue. Sayer (2000, p. 12) explains that critical realism is an emergent philosophy in positing that:

the world is characterized by emergence, that is situations in which the conjunction of two or more features or aspects gives rise to new phenomena, which have properties which are irreducible to those of their constituents, even though the latter are necessary for their existence.

This is in contrast to philosophies such as empiricism which is based on methodological individualism, in which the world consists of independent objects. The total is understood by adding up all the individual components and these components are not causally related. For example, Hume (1921 [1777], p. 27) who was one of the founders of British empiricism in the 1700s, in explaining how billiard balls move around the table, argued that “Motion in the second Billiard-ball is a quite distinct event from motion in the first; nor is there anything in the one to suggest the smallest hint of the other”. All we can know is that the billiard balls hit each other (something happened) and observe how often they hit each other. Similarly, while somewhat simplified, a “traditional” understanding of science is limited to observing and counting the constant conjunction of events—that something happened and how many times it happened.

Bhaskar developed critical realism in part in opposition to this reductionist account of science. He identifies three levels of reality in the natural and social worlds: the first is the domain of the real, which consists of structures and generative mechanisms. The second is the domain of the actual in which events take place. The third is the domain of the empirical in which experiences occur. This hierarchy and the relationship between each domain is presented in Table 1. Most of the time generative mechanisms are not observable, and we can only discern their effects in interacting with other generative mechanisms; for example, we can’t see gravity or class, but we can see their effects. Mechanisms interact with each other in open systems and this shapes the way they act, or sometimes, stops them from acting or changes the way that they act. For example, a virus will be stopped by a vaccine, or in the absence of a vaccine, its spread will be accelerated or slowed by the actions of communities and individuals (practicing social distancing or not, wearing masks or not). And, its effects will be shaped by relations of poverty and inequality. This example shows the difficulty of sharply distinguishing between the natural and social worlds, as does the problem of global warming (Benton and Craib, 2001).

The domain of the real shapes the domain of actual, in which events actually happen. Outcomes cannot be fully predicted in advance because events are always co-determined by multiple stratified structures and causal mechanisms (Collier, 1998b), as in the example of the spread of a virus. The extent to which events can be observed is an empirical one; for example, the development of microscopes allowed scientists to see viruses. The domain of the empirical is where we directly perceive events. For this to occur, events must have been generated by the interplay of mechanisms and structures. Empiricist philosophies focus on experiences and counting how many times they happen. According to Bhaskar (1998a), a focus on the “empirical world” to the exclusion of the actual and real reduces the ontological question of what exists to the epistemological question of what we can know or experience, and he refers to this as the “epistemic fallacy”. This is one reason why an over-reliance on large-scale data as the foundation of education research can lead to problematic conclusions about “what works” in practice. This depth ontology also allows us to understand absences, or why things do not happen, whereas counting the constant conjunction of events misses this.

Table 1 The domains of the real, actual and empirical.

	<i>Domain of real</i>	<i>Domain of actual</i>	<i>Domain of empirical</i>
Mechanisms	✓		
Events	✓	✓	
Experiences	✓	✓	✓

Source: Bhaskar (2008 [1975], p. 2).

According to emergent philosophies such as critical realism, the total is always more than the sum of its parts. A society is more than the sum of the total number of people as it has its own properties, structures and mechanisms that affect people, and society in turn is affected by what people do. Bhaskar (1998c) explains that when people get up in the morning to go to work, their intention is not to reproduce capitalist economic relations, yet that is the outcome of their actions because they live and work in a capitalist society. Societies are structured by economies and by relations of power and domination, such as class, race and gender. Archer (2020, p. 141) explains that social structure, culture and human agency “are *different kinds of emergent entities* as is shown by the differences in their properties and powers, *despite* the fact that they are crucial for each other’s formation, continuation and development”. She illustrates this point about different properties by explaining that an education system may be centralized, but an individual cannot be centralized, while an individual can be reflexive, while an education system cannot.

Similarities and differences between the natural and social worlds

Bhaskar (1998b) argues that both the natural and social worlds can be studied scientifically, but that there are important differences between the two. Reality in both the natural and social worlds is stratified, complex and emergent, but the nature of the structures and objects in each differ even though they interact in emergent ways, as in the examples of a virus or global warming. First, social structures in the social world depend on human activity, unlike structures and mechanisms in the natural world (although this notion is challenged by global warming, at least to some extent). Consequently, social structures are only relatively enduring compared to the more enduring structures in the physical world. Social structures tend to change slowly as human activity changes and adapts to its environment, but they may be punctuated by revolutions or crises. An example of change is the change from feudalism to capitalism, and within capitalism, from the capitalist welfare state to the neoliberal market society (Harvey, 2007). One of Archer’s contributions is to theorize this process, and this is discussed in the next section.

The second difference is that the social world is, at least to some extent, concept dependent—it depends in part on the reasons and beliefs agents have, which help to shape their actions (for example, to wear a mask—or not—in the middle of a pandemic). An important part of understanding society is understanding meanings and beliefs, even if these meanings and beliefs are fallible (for example, that the Covid-19 disease is a conspiracy meant to destroy democracy). Sayer (2000, p. 17) explains that meaning can’t be counted, it can only be understood and this demonstrates the necessity for hermeneutical or interpretative research methods in the social sciences. While we may be able to count the number of people who believe that Covid-19 is a conspiracy and it may be important to do so, we need to understand why they think that. However, while beliefs may be a part of the social reality that we seek to explore, they do not explain everything about that social reality. It is also necessary to explore social structures, social institutions, and the way in which humankind engages with the natural, material and social worlds to reproduce and sustain itself.

Third, the social sciences have the capacity to react back on society and to help create change, for example, in providing insights into why different groups of young people drop out of school. Social science can offer alternative explanations to “common sense understandings” and this may help change social practices. Collier (1998a, p. 446) contrasts this to the natural sciences in explaining: “There can be no equivalent of this in the natural sciences. Black holes may be unpleasant things to contemplate, but that is no criticism of them. They exist—or don’t—and there’s an end of it”.

Bhaskar (2011b, p. 143) argues that the implications of critical realism for the aims of science (and social science) are not to seek control and prediction, but explanation, understanding and enlightenment. Control and prediction are possible in closed systems, but not in open systems, and the social world is by definition an open system because it comprises human agents. It may be possible to create a closed system to identify and “see” the actions of causal mechanisms in some sciences such as physics, biology or chemistry, where the possibility of controlling all variables exist. This is not possible in all natural sciences and it is impossible in the social sciences. This is because society is peopled by agents who act on their beliefs and reasons, and it is thus impossible to completely control all variables. Bhaskar (2011b, p. 144) explains that: “It follows from this that decisive test situations are in principle impossible; so that criteria for theory-choice and theory-development must be exclusively explanatory and non-predictive”. However, in contrast to relativist arguments that reduce theories to standpoints held by different actors so that theory is no more than an expression of power, Bhaskar argues that one needs to test theories empirically, which includes qualitative and quantitative research. (In critical realism, quantitative research in open systems does not require constant conjunctions of events or hypotheses testing.) Archer (1995, p. 23) explains that if we find the social world to be different from our conceptions as a consequence of our research, then we should modify our conceptions. There is an iterative relationship between our knowledge of the world and the way that the world is.

Intransitive domain (the natural and social world) and transitive domain (bodies of knowledge)

However, this relationship is imperfect, and our theories are not a mirror of the world (Scott, 2018). There is not a direct “correspondence” between our theories and the objects they describe. If this were the case, then knowledge would never change because it would always truthfully reflect the way the world is. Critical realists distinguish between the intransitive domain which refers to the “real” (the domains of the real, actual and empirical) in the natural and social worlds, and the transitive domain which refers to our knowledge of the real, or in other words, our knowledge of the natural and social worlds. It is this distinction that allows critical realists to distinguish between a realist ontology (the real world exists), fallible knowledge or epistemology (we might be wrong), and judgmental rationality (some theories work better than others, because we’ve checked) (Collier, 1994). According to this

approach, our theoretical knowledge reflects our best efforts so far in understanding the natural and social worlds, but that our knowledge will develop as we gain better understandings through our research and practice in the world.

Objects and structures on the one hand, and our knowledge of objects and structures on the other, are not the same and each has different properties. Our knowledge of the world is always concept-dependent, and these concepts will in part shape what we see and how we understand what we see. This is why our knowledge will always be fallible and provisional, because our experience in the world is always shaped by our theories, but it is not determined by theories (Sayer, 1992). We “see” the world through our conceptual lenses. For example, two teachers, each with contrasting understandings about class and disadvantage, may view the same group of students differently and respond to them differently.

The conditions that are needed to build theoretical knowledge are not the same as the referents (objects) of that knowledge. While black holes and societies exist, their existence doesn’t rely on what we know about them. Societies may change as a result of theoretical knowledge, but this process is usually slow and protracted, and only occasionally punctuated by revolutions and crises. Bhaskar (1998b, p. 34) explains that the raw materials used in constructing new theories in science “include established results and half-forgotten ideas, the stock of available paradigms and models, methods and techniques of enquiry ...” Once codified, theoretical knowledge has properties of its own that can’t be reduced to the conditions of its production or the standpoint of those who produced it. This is for two reasons. The first is that this knowledge is about something, and not just the perceptions or standpoint of those who developed that knowledge. The second reason is, as Chalmers (1999) explains, because theories have properties that are independent of their knowledge producers, they may have implications about which the original knowledge producers are unaware. People build new knowledge using (accepting, extending, modifying, transforming, rejecting) knowledge that went before. Bhaskar (1998b, p. 34) explains that “the scientific innovator comes to appear in retrospect as a kind of cognitive bricoleur”.

This distinction between the intransitive and transitive dimension allows us to understand that each may change independently of the other. For example, the objects of knowledge may undergo a fundamental change, while theoretical knowledge hasn’t caught up; or, theoretical knowledge may develop and change toward a better understanding of its object. This is one reason why we can never definitively say we have arrived at the “truth”, even though we may be getting closer to it.

Knowledge of the powerful and powerful knowledge

Knowledge depends on social communities for its production. Creating knowledge requires “institutional structures and cultures to support develop, codify, curate and transmit knowledge” (Wheelahan, 2014, p. 128). The academic disciplines provide the main framework for developing knowledge, although they are not the only way such knowledge is created. The disciplines have rules for creating, codifying and testing knowledge which are fallible, but the best tools we have to date. Universities are the main institutions which have provided the academic disciplines with a home, and in the Western world both universities and the disciplines have changed since the founding of the University of Bologna in 1088 (see Hayhoe, 2005, 2019 for a discussion of the historical development of scholarly institutions in other civilizations). This elaboration of institutions of knowledge production in turn reflects the growth and complexity in society, and requirements for an increasingly complex and interdependent division of labor in society (Bernstein, 2000).

Because knowledge is produced socially, it bears the marks of those who have produced it. The academic disciplines in universities have historically been dominated by social elites, and the production of knowledge bears these marks, including racism, sexism, homophobia, colonialism, class and other relations of domination and subordination. And historically, those who suffer the deleterious effects of these relations of domination have had to fight for access to universities and other sites of knowledge production, just as they have had to do in every other social institution in society. These struggles have resulted in the establishment of new disciplines such as women’s studies, Black studies, and cultural studies in the latter part of the 20th Century, and the challenging of these relations of domination and subordination within existing disciplines such as medicine, education, engineering, history, economics and philosophy. These challenges have only been partially successful and will always only be partially successful because universities reflect, even if they sometimes challenge and help to change (at least a little bit), society’s relations of power and social structures.

Even though theoretical knowledge will always reflect relations of power, this doesn’t mean that such knowledge is reducible to power. This is because such knowledge is *about* something *other* than relations of power. Theoretical knowledge is powerful because it provides at least some epistemic access to the natural and social worlds, even if this access is imperfect, fallible and shaped in part by relations of power. Michael Young (2013), a leading scholar within the social realist school in the sociology of education, distinguishes between powerful knowledge and knowledge of the powerful. Powerful knowledge is powerful because it provides (at least some) epistemic access to its objects in the natural and social worlds. On the other hand, knowledge of the powerful is shaped and distinguished by unequal social access to knowledge production, acquisition and use. The social elites seek to ensure their privileged access to theoretical knowledge precisely because of the way it can be used in society and the economy—because of its power in understanding the natural and social worlds.

Implications of critical realism for curriculum

There are several implications of a critical realist analysis for curriculum. The first is that knowledge and knowing are not the same. Theoretical knowledge is codified knowledge embedded in relational systems of meaning. Each academic discipline is a body of

abstract knowledge about its object. No academic discipline can represent the world in its entirety, and can only ever have as its referent a small slice of “the real” (Collier, 1997). The task of curriculum is to help students enter different system of meaning and make sense of them. Students are apprentices at different stages of their journey who need to learn their theoretical tools of the trade. They need to make sense of these tools, engage with them skillfully and critically and perhaps contribute to their evolution, but the main point is to help students learn to use these tools to enter and move around within systems of meaning, and to apply knowledge in contextually specific and appropriate ways.

The second implication is that seeking truth about an objective natural and social world should be a normative goal of curriculum even if our knowledge is always partial and revisable, and marked by the social conditions under which it was produced. The objectivity of knowledge is an important principle for curriculum, but as discussed earlier, this is not based on a “correspondence theory” of truth where knowledge directly reflects its objects. Collier (2003) explains that the objectivity of knowledge can be understood in three ways: first, that something exists independently of our knowledge and/or awareness of it. The earth has always been round, even when most of the world’s population thought it was flat. The second meaning of objectivity concerns the nature of human judgments. Knowledge is objective (but still fallible) if it is caused by its object, and not, as Collier (2003, p. 135) explains “by some feature of its subject other than that subject’s openness to the effects of the object”. That is, such knowledge, even if it is imperfect, is about something external and not just a reflection of our internal state. Collier explains that the third meaning of objectivity is objectivity as a human attitude. We seek knowledge (the second meaning of objectivity) about an independent reality (the first meaning of objectivity). So, as a consequence of this third meaning of objectivity, we need to develop the means and tools to seek such knowledge, develop criteria that we use to judge knowledge claims, and be open to the possibility that we are wrong. Otherwise, as Collier (2003, p. 137) explains: “the alternative is to be shut up in one’s own subjectivity”. These three dimensions of objectivity demonstrate why and how knowledge may be fallible and revisable, but still objective.

The third implication for curriculum is that students need to understand how knowledge is produced in different fields and to gain access to the criteria needed to judge knowledge claims. This goes far beyond teaching students propositional “know that” knowledge as truth. Winch (2013, p. 130) explains that propositional knowledge is “embedded within a conceptual structure which is itself embodied in further related propositions”. Students need access to this relational conceptual structure within bodies of knowledge, to their inferential relationships, and to the procedures used to make these claims and to evaluate them. This is not an argument for “the canon as truth”. This is an argument for inducting students into a critical stance within bodies of knowledge where they engage with the criteria used to judge existing and new knowledge. This treats knowledge as objective knowledge yet socially produced and fallible, but about which judgments can be made. The aim is to give students the epistemic tools they need to gain access to the social and natural worlds, and to participate in debates about what our society should be like.

Critical realism provides a foundation for thinking about the nature of curriculum (Priestley, 2011; Scott, 2000, 2014), but it doesn’t provide all the answers. It provides a philosophical foundation, and as such cannot do all the work that is needed to theorize the nature of curriculum. Critical realism under-labours the social sciences by providing a realist framework for complementary social theories that explore society. Bhaskar (1998b, p. 5) explains that philosophy can tell us that the world is stratified, but “it cannot tell us *what* structures the world contains or *how* they differ”. Archer (1995, p. 22) explains that ontology “acts as both gate-keeper and bouncer for methodology”. For example, theories premised on methodological individualism are ruled out, while those based on an emergentist and realist philosophy are in play (see Archer, 2020). This leaves considerable scope for debate within emergentist realist social theories such as Archer’s and the social realist school within the sociology of education.

Archer’s morphogenetic social realism

Margaret Archer’s (1995, 2000a) morphogenetic realist social theory builds on critical realism and it has become very influential in sociology. Her approach theorizes the relationship between the individual and the social, with both premised on a materialist and emergentist ontology. She uses “analytical dualism” to distinguish between agency and social structure, arguing that each has different emergent properties, and that neither is reducible to the other. For example, a society can have a young or old demographic structure, be more or less unequal, and have different types of political systems and economies. In contrast, individuals have capacities for cognition, perception and consciousness, and reflection.

Archer insists that in analytically distinguishing between agency and structure she avoids under-socialized accounts of the individual based on methodological individualism in which society is the sum total of individuals scurrying about their business. Such approaches are guilty of “upwards conflation”. She also avoids “downwards conflation” in which personhood is the gift of society with the individual who is discursively constructed in society which, she argues, results in the over-socialized individual (for example, that an individual is nothing more than a product of their culture or religion). And, she avoids the problem of “central conflation” which she attributes to Giddens (1984) in his theory of structuration. Giddens defines social structure as rules and resources that individuals draw upon in acting and in the process they “instantiate” social structure. Archer (2010b) critiques this approach for collapsing the distinction between structure and agency, for collapsing the temporal dimension between actions of agents and social structure, and for excessive voluntarism (all that is needed for social change to occur is for individuals to change what they do).

Archer (1995) explains that while society emerges from social relations between agents (and their interactions with the natural and material worlds), it is not possible to add up all individuals in society to explain what a society is. A society would not exist without agency and is thus activity dependent, but it has emergent properties which mean that it is not reducible to agency and is able to act back on agency. For example, a society with high unemployment is also likely to be highly unequal and to deny opportunities to young people which may well shape their life chances. And relations of patriarchy and racism are also real and shape the life chances of individuals. Gorski (2017, p. 34) explains that the consequence of this depth ontology is that “human persons and societies may have certain latent potentials whose development has been suppressed and whose absence manifests in ill-being”.

The temporal dimension is very important in Archer’s approach. She explains that structure predates human activity; that is, social structures are the outcomes of past agential activity and they provide the parameters that individuals act within in the present. She argues that social structures condition human agency but do not determine what individuals will do precisely because they are agential beings. The outcome of social interactions by agents within existing social structures determines whether no change occurs (morphostasis), or whether social change occurs (morphogenesis). This is why her approach is called the morphogenetic approach.

Her analysis of the nature of the individual is a key component of her approach, and she insists on the ontological status and irreducibility of the self (Archer, 2000a). Archer (1995) draws on the seminal work of Marcel Mauss ([1938] 1985) (French scholar and Emile Durkheim’s nephew) to argue that the sense of self is universal, while the concept of the self is historically and socially situated. This helps us to see that the conception of the self of Western individualism is a socially, temporally and historically specific construction, while not denying that there are selves. That is, there is an enduring “self” who has continuity even though this self may change in engaging with the natural, practical and social worlds, and over time in engaging with others in socially conditioned circumstances (which includes relations of class, gender and race). This self has a past, present and a future and acts reflexively in responding to the situation in which they find themselves, which may result in the emergence of new identities and subjectivities (Clegg, 2006).

Archer (2003) refers to the internal reflexivity that all individuals engage in as the “internal conversation”. We reflect, consider, make decisions, change our minds and act in engaging in the social, practical and natural worlds. This resonates with Bhaskar’s (1998b) insistence that our knowledge of the world arises from our practice in the world. Archer provides insight into the iterative relationship between our inner conversation and our practice in the world. She argues that the self is defined by things that matter to us, including who and how we choose to love, our priorities and where we choose to “invest” ourselves, even this results in detrimental outcomes for ourselves. Collier (1999) argues that our emotions are our internal commentary on the things that matter to us, and consequently emotions are not irrational, although they are fallible if our understanding of the situation in which we find ourselves is mistaken. Our “inner conversation” mediates our relationship with the world and is not purely self-referential (Archer, 2000a, 2003). While agreeing with Archer about the irreducibility of the self, Archer’s account of the inner conversation is arguably over-intellectualized in discounting the ways in which our experiences acquired over a lifetime are inscribed in our “habitus” and help to orient our understandings of the normal, natural and taken-for-granted (see the debate between Sayer, 2010, and Archer, 2010a).

Archer’s work has been used in theorizing about pedagogy (particularly teaching and learning in post-secondary education) (see Case, 2015, 2017; Clegg, 2005, 2006; Ashwin, 2012), but there has been less work that considers the implications of her work for curriculum (but see Priestley, 2011; Maton, 2014; Wheelahan, 2007). Regardless of any reservations, Archer’s is an expansive notion of human agency which stands in stark contrast to the homo economicus (Devas, 1883, p. 27) of human capital theory (Archer, 2000b). This is the notion of the human agent which underpins curriculum in which the purpose of education is reduced to preparing people for work. This is why Archer’s approach has implications for theorizing curriculum and not just pedagogy because the purposes of a curriculum are more expansive based on a more expansive notion of the agent. Her notion of investment encompasses all the things that matter to us, and not the instrumental self-maximizing economically “rational” self whose main goal in life is realization of their market selves.

The social realist school in the sociology of education

The main focus of the social realist school within the sociology of education is on theorizing the role of theoretical knowledge in curriculum (Young, 2008; Muller, 2000; Moore, 2004; Maton and Moore, 2010; Maton, 2014; Wheelahan, 2010). All those within the social realist school acknowledge their debt to the British sociologist Basil Bernstein who was an English sociologist of education from the 1960s until the end of the 20th century (see Muller and Gamble, 2010). Bernstein (2000) in turn, acknowledges his debt to Durkheim who was a founder of the discipline of sociology in the 19th century, as do all within the social realist school. However, not all those within the social realist school draw on critical realism, or draw on critical realism as markedly as does this article even though all are committed to a realist ontology (Young and Muller, 2007).

As part of his oeuvre, Bernstein theorizes the nature of, and structures of, theoretical knowledge and everyday knowledge, and in doing so he addresses a key shortcoming in critical realist theorizing about the nature of knowledge. If the transitive domain is relatively independent of the intransitive domain and has its own properties, then we should seek to theorize those properties. There hasn’t been a great deal of critical realist work within the sociology of education that takes on this task. Knowledge is mostly taken as a given (for exceptions, see Scott, 2010; Wheelahan, 2010; Maton, 2014). From Bernstein (and before him, Durkheim), social realists insist on the distinction between theoretical knowledge on the one hand, and everyday knowledge on the other, arguing that each has different properties and that this has implications for structuring curriculum.

Bernstein's approach is fundamentally Durkheimian because of his analysis of the nature of knowledge. Bernstein shares with Durkheim the view that all societies distinguish between esoteric knowledge on the one hand and mundane knowledge on the other (Bernstein, 2000; Durkheim, 1967). Esoteric knowledge is theoretical and conceptual knowledge, while mundane knowledge is "every day" common sense knowledge (Bernstein, 2000, p. 157). Bernstein argues that the distinction between the two is universal, while the content of each is culturally and historically specific. It is the fact of this distinction, and the social relations underpinning it, that shapes the social division of labor and knowledge production, rather than the relationship between knowledge and the object that knowledge is about.

This is an idealist ontology that makes primary the ideas that humans use to make sense of their world and make their way in the world (Johnson et al., 1984). Or as Bernstein (2000, p. 82) says, "Word before the World". In contrast, critical realism is a materialist ontology because its premise is that knowledge arises from our practice in the world. Or, world before word. However, Bernstein and Durkheim are realist because they regard the social structures that arise from the ideational division of labor as real structures that are not reducible to the individuals that comprise a society (Bernstein, 2000; Durkheim, [1897] 1951). Durkheimian approaches are thus idealist and realist—the real is the ideal, and because it is real, knowledge has objective qualities and is not relativist (although still fallible). However, arguably, as well as being realist, Durkheim's approach is also emergentist (Sawyer, 2002). Durkheim ([1897] 1951, p. xxxvi) argues, for example, that "It is not realized that there can be no sociology unless societies exist, and that societies cannot exist if there are only individuals". Moore (2006) argues that even though Bernstein eschewed labels, that Bernstein's approach is nonetheless best understood as a form of critical realism because of his realist and emergentist emphasis, and because of his insistence that theory describe something external to itself; theory needs to describe and explain practices, processes, structures and outcomes and their inner logics.

Social justice and social and epistemic access to knowledge

Bernstein is fundamentally concerned with social justice and distributive justice. He argues that access to theoretical knowledge provides the grounds for democracy. Theoretical knowledge is the means through which society conducts its conversation about itself, and debates what it should be like in the future. It provides the knowledge that society needs to consider alternative futures, such as ameliorating global warming or responding to the refugee crisis. It is fallible and marked by social elites, but it is society's best efforts so far. He argues that individuals have the "right to the means of critical understanding and to new possibilities", the right to be "included socially, intellectually, culturally and personally", and, "the right to participate" (Bernstein, 2000, p. xx). Access to knowledge is a key precondition for participation. Young (2010, p. 6) explains that consequently the purpose of education is to provide students with access to the knowledge that "takes them *beyond their experience* and which they would be unlikely to have access to at home, at work or in the community". This is a precondition for all students to be able to participate in debates about what society should be like and in debates and controversies within their occupation. Bernstein shows that there can be no social access to theoretical knowledge without epistemic access, and his life's work was to demonstrate how working class students were systematically denied this access through the structuring of knowledge in the curriculum.

Bernstein's main focus was on theorizing the nature of curriculum and how the "pedagogic device" mediates access to theoretical knowledge. He explains how distributive, recontextualizing and evaluative rules mediate access to pedagogic codes (how knowledge is classified into different domains and how it is framed, paced and sequenced), and how the construction and structure of these codes either obscures or facilitates students' access to theoretical knowledge (Muller and Gamble, 2010; Singh, 2002). While this is so, the pedagogic codes used in school that draw on abstract and conceptual knowledge (and arguably, in college and university) are more congruent with those used in middle class homes, with the latter providing "an effective second site of acquisition" (Arnot and Reay, 2004, p. 204). Middle class students are thus doubly privileged as the curriculum they engage in at school provides them with privileged access to theoretical knowledge, while the curriculum for students from working class backgrounds is more likely "to stress operations, local skills rather than the exploration of principles and general skills, and the pacing is likely to be weakened" (Bernstein, 2003, p. 77).

Horizontal and vertical discourses

Bernstein (1999, p. 157) explains that it wasn't until relatively late in his career that he turned his attention from a "theory of the construction of pedagogic discourse" to theorizing the nature of theoretical and everyday knowledge, their structures and "the internal principles of their construction and social base". This was to become a fundamental and enduring part of his work. He describes everyday knowledge as "horizontal discourse" and theoretical knowledge as "vertical discourse", with each having different rules for its generation and acquisition. Horizontal discourse is "common-sense", everyday knowledge which is "likely to be oral, local, context dependent and specific, tacit, multi-layered, and contradictory across but not within contexts" (Bernstein, 1999, p. 159). Vertical discourse, in contrast, consists of "specialized symbolic structures of explicit knowledge" (p. 161) that "takes the form of a coherent, explicit, and systematically principled structure" (Bernstein, 1999, p. 159).

Within vertical discourses, Bernstein distinguishes between the sciences which are hierarchically organized (for example, physics), and the social sciences and humanities which "take the form of a series of specialized languages". An example of a discipline with multiple languages is sociology, which comprises competing paradigms which include, among others, conflict theory and functionalism. Bernstein calls the pure academic disciplines "singularity" because they face inwards to a single knowledge

structure, while he calls applied disciplines “regions”, because they sit at the interface between the pure disciplines and the occupational field of practice, and they integrate and recontextualize knowledge from the “pure” disciplines for the field of practice. The pure and applied academic disciplines have authoritative speakers, principles and rules for generating new knowledge, and criteria for judging knowledge claims. Students need access to the principles for constructing the discipline or applied discipline (its key concerns and how they are distinguished from other disciplines and fields) and knowledge building, and for evaluating knowledge claims. They also need to be inducted into debates and controversies within the disciplines. Hierarchically structured disciplines (such as physics) must ensure that students can understand what comes first before they can understand what comes second, and the social sciences and humanities must ensure students understand the different languages, speakers and authoritative texts (Muller, 2009).

As with Bernstein’s work (Muller and Gamble, 2010), there are debates about, on the one hand, the extent to which social realism is indeed seeking to induct students into the conservative “authoritative cannon” with rigid and untouchable boundaries, and in which the voice of authority provides the grounds for belief, thus legitimating the cultural domination of elites. And, on the other, whether it is a progressive argument designed to support students’ social and epistemic access to knowledge that is emergent, fallible, and revisable based on a critical stance in which students are provided with access to the criteria for judging knowledge claims (Young and Muller, 2010, 2013; Muller and Young, 2019; Morgan et al., 2019).

This is why Young’s (2009) explanation of the distinction between, but also the relationship between, powerful knowledge and knowledge of the powerful is important (Muller and Young, 2019; Beck, 2018). Each has a different kind of power; powerful knowledge provides social and epistemic access to knowledge needed to understand and navigate one’s way in the world. This is potentially available to all, but access to it is restricted by relations of power and domination, in which social elites dominate access to knowledge by dominating access to elite schooling and elite curriculum, and then to elite universities and elite professions. This is knowledge of the powerful. Social elites also have a relative monopoly over the production of such knowledge, and knowledge produced by social elites will reflect these conditions of its production. This is another reason why democratic access to knowledge is so important. However, it is not just democratic access to knowledge that is important, it is also democratic access to knowledge production that is important.

Transactional realism

Like Bernstein, the famous, influential and long-lived United States philosopher and psychologist John Dewey (1859–1952) was concerned with democracy and social justice (see Dewey, 1997 [1916]). He was a founder of pragmatism within both disciplines (Hildebrand, 2018). Hildebrand (2018, p. 1) argues that Dewey was “arguably the most prominent American intellectual for the first half of the twentieth century”. And, arguably, his book *Democracy and Education* is still of fundamental importance in thinking about the aims of education (Dewey, 1997 [1916]). Dewey’s work provides the basis for “transactional realism” and this has been elaborated in the philosophy and sociology of education by scholars such as Biesta (2014, 2020), Carlgren (2020), and Wahlström (2020a,b) (see also Garrison, 1994).

Dewey’s pragmatism emphasizes processes of knowledge creation and meaning-making in which education is a “continuous reconstruction of experience” (Dewey, 1997 [1916], p. 80). The emphasis is experiential, which means that the contextual and situated takes priority. However, Dewey is clearly not a relativist. His insistence on the scientific method of inquiry demonstrates his belief that the external world is not purely a social construction brought about by acts of signification (Dewey, 1997 [1910], p. 27). Biesta (2020, p. 96) explains that “we live ... in a real world, not a fantasy” and we “encounter this reality when our initiatives meet resistance” in the material and social world. Biesta (2014) explains that Dewey’s theory of education is communication-centered, which emphasizes the connection—or the transaction—between the student and the curriculum in the case of education, or between experience of “living organisms” and the environment more broadly. It is realist because not just any connection is possible, as a consequence of butting up against the world. Biesta also problematizes the “learnification” of education, in which the emphasis moves away from education to learning (Biesta, 2009, 2014). Learnification results in “the translation of everything there is to say about education in terms of learning and learners” (Biesta, 2009, p. 38). This ignores the fact that students “learn something, that they learn it for a reason, and that they learn it from someone” (Biesta, 2020, p. 91).

However, arguably, Dewey (1997 [1916], p. 339) collapses the distinction between knowledge and knowing by focusing on the contextual in arguing that: “knowledge is a perception of those connections of an object which determine its applicability in a given situation.” Moreover, the action oriented nature of knowing so that knowing is “in action” lends itself to a “what works” or a pragmatist criterion for evaluating knowledge. Biesta (2014, p. 39) cites Dewey in saying that “The question of truth is not as to whether Being or Non-Being, Reality or mere Appearance is experienced, but as to the *worth* of a certain concretely experienced thing”. In this way, Dewey arguably ends up at subjectivism or more charitably at an intersubjective world (see Biesta, 2014, p. 43 for a discussion of this point). Biesta (2014, p. 45) explains that “the most important question to ask from the point of view of pragmatism is not ‘What is true?’ but ‘What is the problem?’” From a critical realist perspective, Sayer (2000, p. 42) explains that “The question is not only what works, but what is it about the world which makes it work”. A pragmatist “what works” from the perspective of a particular problem doesn’t necessarily provide insights into the structures and mechanisms of the real, and consequently, knowledge derived in this way may “work” but still be wrong in important respects (Chalmers, 1999).

Dewey is not relativist about the world, but he is arguably relativist about knowledge. Moreover, his realism about the world is very attenuated. Biesta (2014, p. 40) explains that:

Dewey's approach also has implications for how we understand the objects of knowledge. Whereas in the dualistic approach the objects of knowledge are seen as "things" that exist in a world "out there" and are there for us to discover and depict, Dewey's transactional view sees the objects of knowledge as the *outcomes* of processes of inquiry.

This is a quite distinct approach to critical realism which distinguishes between the intransitive and transitive domains, with each having different conditions necessary for their existence. Experience, while a precondition for knowledge, is not sufficient if that knowledge is to be codified, elaborated, and institutionalized in an elaborated system of meaning.

It is important to distinguish between curriculum and the production of knowledge in the academic disciplines. It is possible to agree with Carlgren (2020, p. 332), a proponent of transactional realism, who argues that "the aim of education is to produce knowledgeable people" and that this is a precondition for democracy (Wahlström, 2020a,b). However, the production of knowledge and the curriculum have different aims, dynamics and conditions for their realization. The structures of knowledge do have implications for curriculum depending on whether they are hierarchically structured or structured as horizontal languages (Muller, 2009). Students must engage experientially with the objects of these knowledge structures, as well as engaging with the relational system of meaning about those objects of knowledge. The challenge for a curriculum is to work out how that should be done for different groups of students at different stages of development who are studying different sorts of programs for different purposes at different types of institutions (schools, colleges and universities). Finally, one can agree with Russell (1945, p. 819) who explains that while he greatly admires Dewey and that he wishes that he could "agree completely" with him, but that

... to my regret I am compelled to dissent from his most distinctive philosophical doctrine, namely the substitution of "inquiry" for "truth" as the fundamental concept of logic and theory of knowledge.

Conclusion

Realist approaches have implications for theorizing the aims and purposes of curriculum, and also for what should be included in curriculum. The purpose of curriculum is to produce "knowledgeable people" (Carlgren, 2020) about the world, a real world which is stratified, complex and emergent (Bhaskar, 1998b). This is based on an expansive view of the human actor, their motivation and concerns and the complexity of their worlds, and not the homo economicus "man" of human capital theory (Archer, 1995). We want students to be able to participate in debates and controversies in society about what our world should be like, and to participate in and shape debates and conversations about their occupation (Bernstein, 2000). Theoretical knowledge is powerful knowledge, and students require epistemic access to this knowledge, as a precondition for social access. We need to distinguish between powerful knowledge and knowledge of the powerful if we are to provide students with access to the former, while challenging relations of power intrinsic to the latter (Young, 2013). The structures of knowledge as either hierarchical or multiple languages has important implications for curriculum and for curriculum coherence (Muller, 2009). The aim is to ensure students have access to the way knowledge is built, and to the criteria for judging knowledge claims, which is a critical position that rejects the canon as authority and truth. The curriculum needs to be based on the normative principle of seeking truth about a real world, while recognizing that our knowledge of that world is fallible, but about which we can still make judgments (Collier, 2003).

References

- Archer, M.S., 1995. *Realist Social Theory: The Morphogenetic Approach*. Cambridge University Press, Cambridge.
- Archer, M.S., 2000a. *Being Human: The Problem of Agency*. Cambridge University Press, Cambridge.
- Archer, M.S., 2000b. Homo economicus, Homo sociologicus and Homo sentiens. In: Archer, M., Tritter, J.Q. (Eds.), *Rational Choice Theory Resisting Colonization*. Routledge, London.
- Archer, M.S., 2003. *Structure, Agency and the Internal Conversation*. Cambridge University Press, Cambridge.
- Archer, M.S., 2010a. Can reflexivity and *habitus* work in tandem? In: Archer, M. (Ed.), *Conversations About Reflexivity*. Routledge, London.
- Archer, M.S., 2010b. Morphogenesis versus structuration: on combining structure and action. *Br. J. Sociol.* 61, 225–252.
- Archer, M.S., 2020. The morphogenetic approach; critical realism's explanatory framework approach. In: Róna, P., Zsolnai, L. (Eds.), *Agency and Causal Explanation in Economics*. Springer International Publishing, Cham.
- Arnot, M., Reay, D., 2004. The framing of pedagogic encounters: regulating the social order in classroom learning. In: Muller, J., Davies, B., Morais, A. (Eds.), *Reading Bernstein, Researching Bernstein*. RoutledgeFalmer, London.
- Ashwin, P., 2012. *Analysing Teaching-Learning Interactions in Higher Education: Accounting for Structure and Agency*. Bloomsbury Publishing Plc, London, United Kingdom.
- Beck, J., 2018. "Beyond the present and the particular": similarities and differences between Michael Young's and Charles Bailey's arguments for the public provision of liberating forms of education for all. In: Guile, D., Lambert, D., Reiss, M.J. (Eds.), *Sociology, Curriculum Studies and Professional Knowledge: New Perspectives on the Work of Michael Young*. Routledge, London.
- Benton, T., Craib, I., 2001. *Philosophy of Social Science: The Philosophical Foundations of Social Thought*. Palgrave, Hampshire.
- Bernstein, B., 1999. Vertical and horizontal discourse: an essay. *Br. J. Sociol. Educ.* 20, 157–173.

- Bernstein, B., 2000. *Pedagogy, Symbolic Control and Identity*. Rowman & Littlefield Publishers, Oxford.
- Bernstein, B., 2003. *The Structuring of Pedagogic Discourse*. In: *Class, Codes and Control*, vol. IV. Routledge, New York.
- Bhaskar, R., 1998a. The logic of scientific discovery. In: Archer, M., Bhaskar, R., Collier, A., Lawson, T., Norrie, A. (Eds.), *Critical Realism: Essential Readings*. Routledge, London.
- Bhaskar, R., 1998b. The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences. Routledge, London.
- Bhaskar, R., 1998c. Societies. In: Archer, M., Bhaskar, R., Collier, A., Lawson, T., Norrie, A. (Eds.), *Critical Realism: Essential Readings*. Routledge, London.
- Bhaskar, R., 2008 [1975]. *A Realist Theory of Science*, third ed. Routledge, London.
- Bhaskar, R., 2011a. *Philosophy and the Idea of Freedom*. Routledge, London.
- Bhaskar, R., 2011b. *Reclaiming Reality*. Routledge, London.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21, 33–46.
- Biesta, G., 2014. Pragmatizing the curriculum: bringing knowledge back into the curriculum conversation, but via pragmatism. *Curric. J.* 25, 29–49.
- Biesta, G., 2020. Risking ourselves in education: qualification, socialization, and subjectification revisited. *Educ. Theor.* 70, 89–104.
- Carlgren, I., 2020. Powerful knowns and powerful knowings. *J. Curric. Stud.* 52, 323–336.
- Case, J.M., 2015. *A Social Realist Perspective on Student Learning in Higher Education: The Morphogenesis of Agency*. Higher Education Research & Development.
- Case, J., 2017. *Higher Education and Social Justice: Engaging the Normative with the Analytical*. Centre for Global Higher Education working paper series. Centre for Global Higher Education, Institute of Education, University College London.
- Chalmers, A.F., 1999. *What is This Thing Called Science?* Open University Press, Buckingham.
- Clegg, S., 2005. Theorizing the mundane: the significance of agency. *Int. Stud. Sociol. Educ.* 15, 149–164.
- Clegg, S., 2006. The problem of agency in feminism: a critical realist approach. *Gend. Educ.* 18, 309–324.
- Collier, A., 1994. *Critical Realism: An Introduction to Roy Bhaskar's Philosophy*. Verso, London.
- Collier, A., 1997. Unhewn demonstrations. *Radic. Philos.* 81, 22–26.
- Collier, A., 1998a. Emancipation and explanation. In: Archer, M., Bhaskar, R., Collier, A., Lawson, T., Norrie, A. (Eds.), *Critical Realism: Essential Readings*. Routledge, London.
- Collier, A., 1998b. Stratified explanation and Marx's conception of history. In: Archer, M., Bhaskar, R., Collier, A., Lawson, T., Norrie, A. (Eds.), *Critical Realism: Essential Readings*. Routledge, London.
- Collier, A., 1999. *Being and Worth*. Routledge, London.
- Collier, A., 2003. *In Defence of Objectivity and Other Essays*. Routledge, New York.
- Devas, C.S., 1883. *Groundwork of Economics*. Longmans, Green, & Co, London.
- Dewey, J., 1997 [1910]. *How We Think*. Dover Publications, Mineola, New York.
- Dewey, J., 1997 [1916]. *Democracy and Education: An Introduction to the Philosophy of Education*. The Free Press, New York.
- Durkheim, E., 1951 [1897]. *Suicide*. Routledge, London.
- Durkheim, É., 1967. *The Elementary Forms of Religious Life*. The Free Press, New York.
- Garrison, J., 1994. Realism, Deweyan pragmatism, and educational research. *Educ. Res.* 23, 5–14.
- Giddens, A., 1984. *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press, Berkeley.
- Gorski, P., 2017. Human flourishing and human morphogenesis: a critical realist interpretation and critique. In: Archer, M.S. (Ed.), *Morphogenesis and Human Flourishing*. Springer International Publishing, Switzerland.
- Harvey, D., 2007. *A Brief History of Neoliberalism*. Oxford University Press, Oxford.
- Hayhoe, R., 2005. Universities, women, and the dialogue among civilizations. In: Jones, G.A., Mccarney, P.L., Skolnik, M.L. (Eds.), *Creating Knowledge, Strengthening Nations: The Changing Role of Higher Education*. University of Toronto Press, Toronto.
- Hayhoe, R., 2019. The gift of Indian higher learning traditions to the global research university. *Asia Pac. J. Educ.* 39, 177–189.
- Hildebrand, D., 2018. John Dewey. In: Zalta, E.N. (Ed.), *Stanford Encyclopedia of Philosophy*, winter 2018 ed.
- Hume, D., 1921 [1777]. *An Enquiry Concerning Human Understanding and Selections from a Treatise of Human Nature*. The Open Court Publishing Company, Chicago.
- Johnson, T., Dandeker, C., Ashworth, C., 1984. *The Structure of Social Theory*. Macmillan, Hampshire.
- Maton, K., Moore, R. (Eds.), 2010. *Social Realism Knowledge and the Sociology of Education: Coalitions of the Mind*. Continuum, London.
- Maton, K., 2014. *Knowledge and Knowers*. Routledge, London.
- Mauss, M., 1985. A category of the human mind: the notion of person; the notion of self. In: Carrithers, M., Collins, S., Lukes, S. (Eds.), *The Category of the Person*. Anthropology, Philosophy, History. Cambridge University Press, Cambridge.
- Moore, R., 2004. *Education and Society: Issues and Explanations in the Sociology of Education*. Polity Press, Cambridge.
- Moore, R., 2006. Knowledge structures and intellectual fields: Basil Bernstein and the sociology of knowledge. In: Moore, R., Arnot, M., Beck, J., Daniels, H. (Eds.), *Knowledge, Power and Educational Reform: Applying the Sociology of Basil Bernstein*. Routledge, London.
- Moore, R., 2013. Social realism and the problem of the problem of knowledge in the sociology of education. *Br. J. Sociol. Educ.* 34, 333–353.
- Morgan, J., Hordern, J., Hoadley, U., 2019. On the politics and ambition of the “turn”: unpacking the relations between future 1 and future 3. *Curric. J.* 30, 105–124.
- Muller, J., Gamble, J., 2010. Curriculum and structuralist sociology: the theory of codes and knowledge structures. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, Oxford.
- Muller, J., Young, M., 2019. Knowledge, power and powerful knowledge re-visited. *Curric. J.* 30 (4), 1–19.
- Muller, J., 2000. *Reclaiming Knowledge. Social Theory, Curriculum and Education Policy*. RoutledgeFalmer, London.
- Muller, J., 2009. Forms of knowledge and curriculum coherence. *J. Educ. Work* 22, 205–226.
- Priestley, M., 2011. Whatever happened to curriculum theory? Critical realism and curriculum change. *Pedagog. Cult. Soc.* 19, 221–237.
- Russell, B., 1945. *History of Western Philosophy*. Simon and Schuster, New York.
- Sawyer, R.K., 2002. Durkheim's dilemma: toward a sociology of emergence. *Socio. Theor.* 20, 227–247.
- Sayer, A., 1992. *Method in Social Science: A Realist Approach*. Routledge, London.
- Sayer, A., 2000. *Realism and Social Science*. Sage, London.
- Sayer, A., 2010. Reflexivity and the habitus. In: Archer, M. (Ed.), *Conversations About Reflexivity*. Routledge, London.
- Scott, D., 2000. *Realism and Educational Research: New Perspectives and Possibilities*. RoutledgeFalmer, London.
- Scott, D., 2010. *Education, Epistemology and Critical Realism*. Routledge, London.
- Scott, D., 2014. Knowledge and the curriculum. *Curric. J.* 25, 14–28.
- Scott, D., 2018. Powerful knowledge and the formal curriculum. In: Guile, D., Lambert, D., Reiss, M.J. (Eds.), *Sociology, Curriculum Studies and Professional Knowledge: New Perspectives on the Work of Michael Young*. Routledge, London.
- Singh, P., 2002. Pedagogizing knowledge: Bernstein's theory of the pedagogic device. *Br. J. Sociol. Educ.* 23, 571–582.
- Wahlström, N., 2020a. Democracy and curriculum—the task still before us. *Eur. Educ. Res. J.* 19, 351–363.
- Wahlström, N., 2020b. Tracing implications of transnational policy in curriculum events. *Curric. J.* 31 (1).
- Wheelahan, L., 2007. Blending activity theory and critical realism to theorize the relationship between the individual and society and the implications for pedagogy. *Stud. Educ. Adults* 39, 183–196.
- Wheelahan, L., 2010. *Why Knowledge Matters in Curriculum: A Social Realist Argument*. Routledge, London.
- Wheelahan, L., 2014. Babies and bathwater: revaluing the role of the academy in knowledge. In: Barnett, R., Gibbs, P. (Eds.), *Thinking About Higher Education*. Springer, New York.

- Winch, C., 2013. Curriculum design and epistemic ascent. *J. Philos. Educ.* 47, 128–146.
- Young, M., Muller, J., 2007. Truth and truthfulness in the sociology of educational knowledge. *Theor. Res. Educ.* 5, 173–201.
- Young, M., Muller, J., 2010. Three educational scenarios for the future: lessons from the sociology of knowledge. *Eur. J. Educ.* 45, 11–27.
- Young, M., Muller, J., 2013. On the powers of powerful knowledge. *Rev. Educ.* 1, 229–250.
- Young, M., 2008. *Bringing Knowledge Back In: From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, London.
- Young, M., 2009. Education, globalization and the “voice of knowledge”. *J. Educ. Work* 22, 193–204.
- Young, M., 2010. Alternative educational futures for a knowledge society. *Eur. Educ. Res. J.* 9, 1–12.
- Young, M., 2013. Powerful knowledge: an analytically useful concept or just a “sexy sounding term”? A response to John Beck’s “Powerful knowledge, esoteric knowledge, curriculum knowledge”. *Camb. J. Educ.* 43, 195–198.

World-systems analysis and curriculum theory

Marcin Starnawski^a and Tom G. Griffiths^b, ^a University of Lower Silesia, Wrocław, Poland; and ^b Oslo Metropolitan University, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	96
Structures of knowledge	96
(Unidisciplinary) historical social science	98
Historical social science and school curricula	99
Global education as historical social science	100
Beyond school curriculum: the role of antisystemic movements	102
References	103

Introduction

World-systems analysis (WSA) is an ongoing intellectual movement that [Wallerstein \(2001\)](#) described as “a critique of existing dominant views in the various social sciences, and primary of developmentalism and modernization theory ...” (p. 266). The WSA project has been wide-ranging in its scope and ambition, promoting academic work with explicit commitments to the construction of a world-system that is more equal, just, and democratic than our current, historical, capitalist world-economy. As a knowledge movement in the social sciences, WSA elaborated an approach of “historical social science” that challenged institutional boundaries between disciplines and the established structures of knowledge within them. Historical social science brought together the nomothetic pursuit of universal and generalizable truth and particular, idiographic, conceptions of the good, in an effort to “construct understandings of reality that, albeit inherently approximate and certainly not deterministic, will be useful heuristically in focusing us on the historical options we have in the present in which we all live” ([Wallerstein, 2004a](#), p. 3).

This statement from [Wallerstein's \(2004a\)](#) introduction to *The uncertainties of knowledge*, expresses an important feature of the WSA project, with implications for the question of curriculum theorizing and contents. The project explicitly invokes ideas of the “usefulness” of knowledge generated by world-systems analysis, as criteria for assessing its worth. Usefulness is marked in terms of what Wallerstein identifies as three key tasks of WSA: the intellectual task of understanding social reality; the moral task of imagining alternatives; and the political task of taking action to transform our reality. In WSA terms this involves developing our understanding of the origins and historical trajectory of our existing world-system, as the primary unit of analysis of social reality, locating analyses of present reality within these terms, and engaging in the “utopistics” project of the

sober, rational, and realistic evaluation of human social systems, the constraints on what they can be, and the zones open to human creativity. Not the face of the perfect (and inevitable) future, but the face of an alternative, credibly better, and historically possible (but far from certain) future.

[Wallerstein \(1998, p. 2\).](#)

As articulated by [Wallerstein \(2012\)](#), if world-systems analysis “ends up as being merely one more theoretical position within the social sciences, it will have failed in what it had hoped to accomplish” (p. 521). What did it hope to accomplish? According to [Wallerstein \(2012\)](#), the ambition was unambiguously focused on a “rejection of the entire framework of mainstream social science,” leading to its “unidisciplinary reorganization” (p. 520). The reorganization rejected the divisions of knowledge across economics, politics, and socio-cultural areas of analysis, insisting that all three are “intrinsically linked” and “define each other” (p. 518). Crucially, given some critiques of the Marxist and so-called economic determinist character of WSA, Wallerstein stressed that none of these intrinsically linked areas is “primary,” hence the emphasis on an unidisciplinarity.

Structures of knowledge

WSA operates with a conceptual framework involving the notion of historical social system (including a typology of systems) and a systematic view of the capitalist world-economy with its structural (core–semiperiphery–periphery), temporal (cyclical and secular) and political (hegemony, ideologies, antisystemic movements) dynamics. WSA also encourages epistemological reflection on the social sciences. This reflection was undertaken by Wallerstein and the Gulbenkian Commission on the Restructuring of the Social Sciences, and the Fernand Braudel Center's working groups, such as the Structures of Knowledge Research Working Group, the Categories of Social Knowledge RWG and RWG on Methodology across the Disciplines ([Report, nd](#); [Wallerstein, 1996](#)).

Wallerstein's critique of the historical shaping of social science(s), and the "two cultures" rift within and between the social sciences and humanities was a constant theme in his writings. A collection of essays published between 1982 and 1991 (Wallerstein, 2001) was presented as a call to overcome "the limits of 19th-century paradigms" and set out the task to "unthink" the prevailing assumptions. As Wallerstein (2001, p. 1) explained:

I have entitled this book, "unthinking social science" and not "rethinking social science." It is quite normal for scholars and scientists to rethink issues. When important new evidence undermines old theories and predictions do not hold, we are pressed to rethink our premises. In that sense, much of nineteenth-century social science, in the form of specific hypotheses, is constantly being rethought. But, in addition to rethinking, which is "normal," I believe we need to "unthink" nineteenth-century social science, because many of its presumptions—which in my view are misleading and constructive—still have far too strong a hold on our mentalities. These presumptions, once considered liberating of the spirit, serve today as the central intellectual barrier to useful analysis.

Wallerstein suggested six themes, where this task of "unthinking" social science knowledge can lead to its fruitful reconstruction. The first theme is the critical understanding of the invention of the social sciences as a feature of the historical dynamics of the modern world-system. This centers around the question of how and why the social sciences emerged and became institutionalized in the 19th century as separate disciplines to study the past and the present; the state, the market, and the society; the "West" and the "rest." These "cleavages" were associated with the disciplines of history, political science, economics, sociology, social/cultural anthropology, and oriental studies, that eventually split into social sciences "proper" and humanities. The social sciences were also caught up in the epistemological division described by Wallerstein (2001) as "the false nomothetic-idiographic antinomy" (p. 2), whereby the natural sciences applied a nomothetic epistemology to produce universal truths, while philosophy and the humanities involved idiographic projects focused on particular descriptions social phenomena. A consequence was the reinforcement of the divide between work in pursuit of the truth and that in pursuit of the good.

The second theme focuses on the notion of *development*, presented as "the key, and most questionable, concept of 19th-century social science" (Wallerstein, 2001, p. 2). Wallerstein assumes the need for the critical understanding of the concept's centrality not only in the later part of the 20th century, where it served to capture disparities between core and periphery in the capitalist world-economy, but as "an avatar of the concept of an 'industrial revolution' ... and ... the axis not only of most historiography but of all the varieties of nomothetic analysis" (Wallerstein, 2001, p. 2). As such, the concept has been only partially correct and rather than an accurate heuristic device it served as an ideological signpost for major political camps (from pro-capitalist conservatives to liberals to socialists and communists). The concept led to "false expectations," like earlier theories of social evolution, and to overlooking the actual unequal interconnectedness of core and peripheral zones of the world-economy. Importantly for WSA, by failing to account for the relational aspect of the world-system, development-oriented scholarship privileged the nation-state as a unit of comparative analysis, whereas WSA points to the *world-system* as the more adequate level to explore global economic (axial division of labor) and political (hegemony) distribution of power. In other words, WSA postulates that rather than tracing development of a nation-state society we should examine the expansion of and transformations within the capitalist world-economy, processes that result in underdevelopment as much as growth, and that the relation-ality of core-periphery dynamics is an analytical key to this.

The third theme is a call to problematize the often neglected notions of time and space. Wallerstein (2001) suggests that "one of the most remarkable achievements of the epistemology that has dominated social science has been to eliminate TimeSpace from the analysis," (pp. 2–3) even though geographic and temporal considerations remained prominent in social research. Rather than fixed, stable, and eternally universal, TimeSpace in human history should be seen as a social creation, and furthermore this concept can help distinguish between different levels of historical analysis. Here WSA draws on Braudel's (1982) conceptualization, which points to different time-spans and to objects of measurement associated with them (Wallerstein, 2001, p. 136). Braudel proposed three subcategories of these time-spans: history of events (*l'histoire événementielle*), cyclical history (*l'histoire conjoncturelle*) and structural history (*l'histoire structurelle*). It is mostly on the first level—that of events, episodes, and directly observable short-term processes—that a vast part of traditional historical studies operated. Wallerstein (2001, pp. 139–144) argues that each of these categories of time, or time-span, is associated with certain space thus forming different TimeSpace realities: geopolitical timespace associated with the history of events, or episodic time; ideological timespace associated with cycles (the Cold War "East-West" or later "North-South" divisions can serve as examples here); and world-system, that is, capitalist world-economy, timespace associated with structural or long-term history (*longue durée*). Wallerstein also introduces the concept of TimeSpace of human choice, based on Tillich's (1948) distinction between "quantitative" and "qualitative" time, that is, between *chronos* and *kairos*, or "formal time" and the "right time," and this transformational TimeSpace is perhaps best signified by terms such as "crisis" or "transition" (Wallerstein, 2001, p. 146). The notion of transformational TimeSpace opens up the question of systemic change. Historical social systems persist over long periods as structural TimeSpace, yet the "tension between cyclical rhythms and the secular trends" (Wallerstein, 2001) is where mid-term and long-term changes occur, and they can be understood deterministically rather than otherwise. However, when the system faces its structural crisis, the time of transformation makes human beings face moral choice concerning the future. *Kairos*

is the TimeSpace of human choice. It is the rare moment when free will is possible. ... Human beings therefore, faced with *kairos*, faced with what I shall term transformational TimeSpace, cannot avoid moral choice. It is thrust upon them, at the "right" time and space, at a moment of qualitative time and space whose length and breadth are of uncertain and unpredictable locus. But despite all these uncontrollable elements, we can be certain that there is a *kairos*, a TimeSpace at which transformation does occur, a TimeSpace in which we all exercise our free will for good or ill. And when it comes, we choose our new order.

Wallerstein (2001, p. 147).

It is this aspect of human choice at the time of a systemic crisis, or "bifurcation," that pertains to Wallerstein's concept of *utopistics*, that is, a reflection on possibilities and directions of action for change.

Wallerstein's reading of Braudel also underpins the fourth major theme in his effort of "unthinking" social science. Braudel advocated interdisciplinary cooperation in defining common themes and directions of inquiry. He saw the plurality of social time and long-term perspective as having a particular value for such efforts. Moreover, Braudel's approach to economic history provides an important angle for WSA, namely his distinction between capitalism and the market. The distinction rests on a long-term view that profit-oriented exchange, i.e., markets, pre-dated the capitalist world-economy. Wallerstein (2001, p. 207) echoes Braudel's claims that contrary to dominant views of liberals and most Marxists, capitalism is effectively anti-market. The reading of Braudel also inspires a critical scrutiny of the common-sense view that positions the market against the state. Wallerstein (2001, pp. 210–211) argues that while capitalists may oppose the state in its role of *regulator* (such as in price controls, or taxation), historically they favored the state as a *guarantor* of monopoly. This line of "unthinking" suggests that the market and capitalism are in contradiction, and that capitalism rewards only those able to sustain their capacity to endlessly accumulate capital leaving behind smaller ownership, the labor and, under monopoly, eventually also customers. The "pure" market in the capitalist world-economy is more of an ideology, whereas the state is an integral and necessary element of this economy (Wallerstein, 2004b).

Some of these issues connect with the fifth theme that is the reading of Karl Marx. Wallerstein (2001, p. 3) argues that as a 19th century thinker, Marx "sought to rise above the limitations of his time." What needs reconsideration is how Marx's ideas were adapted by the historical Marxism (or Marxisms) of modern political parties and movements, predominantly those associated with the Second and Third Internationals (i.e., social democratic, socialist and communist). Wallerstein (2001, p. 151) suggests that "Marx would have written the *Communist Manifesto* differently in 1948 than in 1848, and *Capital* differently in 1959 than in 1859. We must do the same." He highlights three assumptions of the orthodox Marxism, which need revisiting or unthinking: the centrality of the industrial proletariat as a force in capitalist processes and as a major transformative actor; the priority of developed countries in progress toward a socialist revolution; and the priority of industrial capital over merchant capital in understanding capitalist expansion (Wallerstein, 2001, pp. 152–153). Additionally, he argues that "Marx was unequivocally wrong" on the presumption that "capitalism represented progress over what existed before, and it will ineluctably be followed by the dawning of the classless society" (Wallerstein, 2001, p. 167). Therefore, a new analysis must also take into account other forms of labor, including informal and unwaged, the political and economic dynamics of peripheral zones, and the fact that "industrial vs. merchant" is a false dichotomy when the totality of accumulation is considered. Some of Marx's concepts that inspired WSA are: class struggle, polarization, ideology and alienation (Wallerstein, 1995).

(Unidisciplinary) historical social science

The sixth theme concentrates on WSA itself as a way to move beyond the older "paradigms" and the location of social science disciplines within them, toward a more open and *unidisciplinary* field of knowledge: *historical social science*. Wallerstein (2000a) described historical social science as

a singular task ... based on the epistemological assumption that all useful descriptions of social reality are necessarily simultaneously "historical" (that is, they take into account not only the specificity of the situation but the continual and endless changes in the structures under study as well as in their enviroing structures) and "social scientific" (that is, they search for structural explanations of the *longue durée*, the explanations for which, however, are not and cannot be eternal) (p. 34).

WSA is not a theory per se but an analytical framework, a perspective combining: (a) long-term macro-historical inquiry; (b) conceptual tools developed by and across different social science disciplines; (c) epistemological and moral reflection on how social realities are studied, what the implications are of the existing methodologies, and how human choices can be informed by social research. Inspired by post-WWII non-orthodox Marxist debates, non-traditional historiography of the French *Annales* school, and Latin American dependency theory, since the late 1970s WSA has become a label for diverse scholarship oriented toward understanding *historical social systems* rather isolated societies (nation-states). WSA situates the processes characterized as globalization in a long-term historical perspective, in two ways. Given basic qualities of the capitalist world-economy (i.e., the modern world-system), we can see that the spatial expansion combined with structural subjugation of new zones was in fact all about the "globalization" of the accumulation of capital. That today we can talk about truly global world-system is a result this very process, beginning in European conquests of Americas, Africa and Asia, and intensifying over the past 150 years or so through

capital mobility in search of profitability and control over labor, using “fixes” that are not only spatial but also technological/organizational and product-oriented (Harvey, 2007; Silver, 2003). Drawing on WSA we can also move beyond thinking of globalization as a linear and inevitable process and rather see the tendencies which indicate the world-system’s structural and historical crisis that demands choices and actions. Traditional disciplinary divisions and epistemological premises are a major limitation and a challenge to be overcome, rather than adequate tools to understand the capitalist world.

Historical social science requires that we unthink a number of common assumptions and consider their alternatives (Wallerstein, 2001, pp. 238–255). Firstly, the disciplinary divisions that were shaped in the 19th and 20th centuries via the formation of academic departments, journals, specialized learned societies etc., enforce intellectual and institutional constraints but do not reflect how humans act. Overlapping occurs between, for example, sociology and socio-cultural anthropology, political science and history, psychology, and education and in all other directions. As an academic discipline, education itself is a good example of multi-disciplinary influences. Moreover, a growing number of fields (rather than disciplines) establish themselves as interdisciplinary, cross-disciplinary or transdisciplinary, and more and more journals and conference calls invite contribution from across a wide range of disciplines. Overlooking of disciplinary boundaries is increasingly seen as a necessary step for specialists researching an issue, and cross-disciplinary theoretical inspirations are also frequent. WSA is an attempt to facilitate such an integrative approach.

Secondly, the epistemological split between the social sciences and humanities, which is captured by the *nomothetic* (oriented at generalizations and theoretical formulations of regularities, or rules governing the social world) and *idiographic* (in-depth description of unique sequences of events and patterns) ways of analysis, is also problematic. This can manifest itself as a clash between social scientists “proper” and historians. However

World-systems analysis offers the heuristic value of the *via media* between transhistorical generalizations and particularistic narrations. ... It argues that the optimal method is to pursue analysis within systemic frameworks, long enough in time and large enough in space to contain governing “logics” which “determine” the largest part of sequential reality, while simultaneously recognizing and taking into account that these systemic frameworks have beginnings and ends and are therefore not to be conceived of as “eternal” phenomena. This implies, then, that at every instant we look both for the framework (the “cyclical rhythms” of the system), which we describe conceptually, and for the patterns of internal transformation (the “secular trends” of the system) that will eventually bring about the demise of the system, which we describe sequentially. This implies that the task is singular. There is neither historian nor social scientist, but only a historical social scientist who analyses the general laws of particular systems and the particular sequences through which these systems have gone.

Wallerstein (2001, p. 244).

Overcoming the nomothetic–idiographic distinction, along with a call for crossing, indeed abolishing, disciplinary boundaries, is a move toward *unidisciplinarity*. This is an organizational as much as theoretical and methodological task. WSA assumes the need for “the wholesale reorganization of the social science sector of our universities and our libraries” (Wallerstein, 2001, p. 272) and, in effect, a transformation of curriculum systems and theorizing about curricula.

WSA also raises the question whether dominant forms of social and historical inquiry have been able to understand the current “moment of transition” of the modern world-system and. If, indeed, these forms of human science crystallized “at the moment of fullest unchallenged triumph of the logic of our present historical system,” they may be conceptually incongruent with the historical challenges faced nowadays. Therefore, WSA proposes a historical social science “that contributes to the transformation of the world by illuminating the choices without appealing to the crutch of a belief in the inevitable triumph of good” (Wallerstein, 2001, p. 256).

A field called *utopistics* is meant as an inquiry into the historical choices in the “moment of transition” or the system’s “bifurcation.” It is described as

the attempt to clarify the real historical alternatives that are before us when a historical system enters into its crisis phase, and to assess at the moment of extreme fluctuations the pluses and minuses of alternative strategies. ... We can clarify the network of forces at work, elaborate possible vectors (and therefore loci of possible conscious interference), and thereby illuminate the real historical choices that are before us. This is not a matter of speculation but of serious research.

Wallerstein (2001, p. 270).

Historical social science and school curricula

How does the WSA ambition for the “rejection of the entire framework of mainstream social science” and its reorganization as a singular and holistic historical social science, relate to questions and debates for curriculum thinking and theorizing within formal education settings? Here we will consider WSA’s potential application to curriculum theory and practice in terms of the rigid disciplinary structure of school curricula globally. The WSA project of unidisciplinary, historical social science can be seen in terms of established theorizing about and initiatives for more integrated and democratic approaches to curricular contents and structures.

At the level of formal curriculum subjects, research has identified a degree of convergence between national systems of education across the globe (Meyer et al., 2017). Acknowledging the inevitable differences between common school subjects found in different countries, and so the limits of claimed curriculum convergence (Anderson-Levitt, 2003), some familiar patterns in the organization

and teaching of knowledge in schools remains. [Anderson-Levitt \(2003\)](#) identified this as the “language arts, mathematics, social sciences, natural sciences, esthetic education, and physical education” (p. 6), but the point is that the dominant model of organization persists: age-based and school subject based classes, with prescribed core curricular contents for school subjects.

The separation of knowledge across school subjects is a curricular problem, reflecting the historical construction of the knowledge disciplines and the organizational dynamics of models of mass school education. In response we find well established lines of argument, and accompanying initiatives, within formal education settings at the levels of policy and practice, to better integrate knowledge from across different subjects, or subject disciplines. [Beane's \(1995\)](#) argument for an integrated curriculum, for example, made the case for shifting the starting point for curriculum organization and contents away from separate knowledge disciplines, asserting that “the sources of curriculum ought to be problems, issues, and concerns posed by life itself” (p. 616). A multi-disciplinary or cross-disciplinary approach follows, in which knowledge contents are selected from across “regular” or established subject disciplines and brought to bear on “problems, issues, and concerns” that students are engaged with. There is a clear influence of Paulo Freire's work here in terms of student involvement in problem posing, connected to their lived experience, as a starting point for setting educational activity including curriculum contents ([Darder, 2018](#); [Darder et al., 2015](#); [Freire, 2014](#)).

An integrated curriculum, from this perspective, “is rooted in a view of learning as the continuous integration of new knowledge and experience so as to deepen and broaden our understanding of ourselves and our world ... It concerns the active construction of meanings rather than the passive assimilation of others' meanings” ([Beane, 1995](#), p. 622). [Apple and Beane's \(1999\)](#) work on a democratic curriculum extends this approach, emphasizing students' knowledge construction that includes the inclusion of non-dominant voices and perspectives, and grounding classroom activity and associated curriculum choices in meaningful problems, events, or issues as experienced by students. The democratic impetus similarly promotes critical approaches to knowledge and knowledge claims or truths about the world, acknowledging the selective, non-democratic, nature of official knowledge in curriculum structures. In this sense, an integrated and democratic curriculum seeks, at least in part, to address the curriculum's role in the reproduction of social inequalities through schooling, and formal schooling's failure to achieve what [Dale \(2007\)](#) described as its “redemptive project” of “severing the links between social origins and destinations” (p. 299).

WSA and its unidisciplinary, historical social science knowledge project, clearly connects directly with efforts to promote inter-disciplinarity within schools. WSA's critique of the disciplinary separation of knowledge, and the associated dominance of (European) scientific universalism, presented the latter as “an essential element in the functioning and legitimation of the political, economic and social structures of the system” ([Wallerstein, 2006](#), p. 59). The disciplinary structures work against the efforts of integrated approaches seeking to ground student learning and knowledge in the urgent social concerns and problems confronting students. [Wallerstein \(2006\)](#) argued that the operation of a capitalist world-economy in pursuit of the endless accumulation of capital, “requires a cultural-intellectual scaffolding to make it work smoothly” (p. 54). Three key elements of this “cultural-intellectual scaffolding” were identified as:

a paradoxical combination of universalistic norms and racist-sexist practices; a geoculture dominated by centrist liberalism; and the seldom noticed, but quite crucial, structures of knowledge based on an epistemological division between the so-called two cultures (p. 54).

Unidisciplinary, historical social science would consciously address each of these elements, tied to a utopistic (not utopian) project of learning and structuring of knowledge that contributes to the political project of transforming our existing world-system (see [Beyond school curriculum: the role of antisystemic movements](#) section below).

Global education as historical social science

Operationalizing an unidisciplinary conception of knowledge for application to school curricula is a challenge, just as efforts to institute integrated or interdisciplinary approaches to knowledge across school subjects confront institutional and organizational barriers. [Aguirre Rojas \(2000\)](#) argued that world-systems analysis has the potential to move beyond the “limited projects to defend and promote a ‘multi,’ ‘pluri,’ ‘trans,’ or ‘inter’ disciplinarity within which, however, the core itself is left untouched regarding the division of social knowledge into ‘disciplines.’ This is the foundation that should really be challenged and radically demolished” (p. 751). Can we envision a school curriculum or university programs that shift away from traditional divisions of social knowledge such as society, politics, and economics; past and present; studies of us and others? And even within subject-based curricula, can we imagine history taught as a way to analyze long term processes and world-scale structures not merely for the sake of studying the closed past but to understand the historicity of our present; or geography classes incorporating socio-economic themes in a similar mode, grounded in WSA; or cultural subjects informed partly by such a framework when discussing intercultural circulation of ideas, marginalized status of some and hegemonic of others?

[Wallerstein \(1997\)](#) is quite explicit about the project of unthinking dominant Eurocentric assumptions, and about avoiding the trap of relying on dominant paradigms to make the critique of European universalism, arguing for example that:

The idea that science is over here and socio-political decisions are over there is the core concept that sustains Eurocentrism, since the only universalist propositions that have been acceptable are those which are Eurocentric. Any argument that reinforces this separation of the two cultures thus sustains Eurocentrism (p. 106).

He affirms this point, asserting that “if social science is to make any progress in the 21st century, it must overcome the Eurocentric heritage which has distorted its analyses and its capacity to deal with the problems of the contemporary world” (p. 94). The broad convergence in school subject disciplines alluded to earlier does not preclude a wide range of particular configurations, including various combinations of subjects into, for example, “key learning areas,” which include multiple more traditional “subjects” being brought together to explore thematic learning areas. For example, schools in New South Wales, Australia, include a “key learning area” of “human society and its environment” which includes subjects of Aboriginal studies, Commerce, Geography, and History (<https://education.nsw.gov.au/teaching-and-learning/curriculum/key-learning-areas/hsie>). Another example could be Poland’s recent introduction of the “History and Present” subject that, while manipulated by the current government and motivated by its conservative nationalism (Starnawski et al., 2021), could provide a promising space for exploring history and its impact beyond dominant narratives of the nation-state and beyond a story of the nation’s glorious events. These are just two of countless examples from across the world, our point being that WSA clearly has a lot to offer to such initiatives, inspired by a historical social science approach.

Curricular reforms like these include “cross curricular themes” that all teachers and subjects are expected or required to attend to, such as attention to Sustainable Development via SDG 4.7 (United Nations, 2015). These reforms are being elaborated within existing systems, linking content selections to central questions of concern to students, societies, and indeed the world. Drake and Reid (2020), for example, note how current attention to 21st century skills development and Problem Based Learning (PBL) supports “an interdisciplinary approach because its underlying philosophy is student-centered, inquiry based, and oriented to big picture concepts and skills applied to real world issues” (p. 6).

One current and promising development in this sort of direction, with ambitions for dissemination on a global scale, is the *Global History of Humanity* initiative, sponsored by the European Union and the SDGs. The initiative involves the production and dissemination of a three-volume *global history of humanity* from 70,000 BCE (Before the Common Era) to the present (Bernardi and Vanhaute, 2021a,b,c). The project is being tested in twelve European countries, with an explicit link to the international SDG agenda, as expressed by the authors:

This textbook is intended as one of the instruments of a broader programme aimed at combining the methods and principles of Global Citizenship Education with the contents of the UN Sustainable Development Goals (SDGs) in school curricula. This textbook and the Get up and goals! project focus especially on quality education for all (target 4.7) and on gender and international inequalities, climate change and global migration. The resulting contents and educational approach aim at overcoming ethnocentrism and enhance students’ ability to place specific issues in a broader and more global dimension.

Bernardi and Vanhaute (2021a, p. 6).

The *Global History of Humanity* initiative offers a current example of curriculum materials, with global ambitions, designed in line with the principles of an unidisciplinary historical social science approach, and with an explicit goal of students “learning from the past to change the present” (Bernardi and Vanhaute, 2021a, p. 16). As expressed by the authors, the goal is for students to “to analyze, criticize, compare, test and broaden our views on the history of humans and the planet they inhabit. This knowledge is also fundamental for making the planet a sustainable, equal, and safe place for all humans, nature and creatures living on it.” (p. 16).

The textbooks take a broad chronological approach, seeking to cover the last 70,000 years through a “firm non-Eurocentric perspective,” and one involving “different time and space scales” (Bernardi and Vanhaute, 2021a, p. 11). The three chronological chapters in each of the volumes are organized not as a linear sequence of events, though simple timelines are included for “readers to remember some significant transformations the planet and humanity underwent, such as climate change and global processes ... [and to highlight] major political and social formations that had global ambitions, like empires and colonial powers.” (Bernardi and Vanhaute, 2021a, pp. 12–13). Rather, each chapter develops a narrative organized into four themes to “look at the different transformations that have occurred in the world” (p. 12). The four themes are: (1) humans change nature, (2) humans on the move, (3) social organization and inequality, and (4) worldviews.

Work like this explicitly and organizationally applies a deeply integrated, historical social science approach to the study of human and non-human history together. In volume 2 for example, Bernardi and Vanhaute (2021b) provide an example of the trajectory of the capitalist world economy in keeping with WSA’s concept of TimeSpace as they describe how “By the 16th century, Western Europe had created new economic and political formations. A radically different relationship between humans and nature was forged; this relationship still exists today and is the root of both the world’s opulence and the present-day problems” (p. 87). Capitalism is defined as a historical social system with

a mode of production that is geared toward “endless” commodification, surplus accumulation and self-expansion. It is very pliable since it has changed the way it functions over time; and it has used different ways of extracting surpluses from nature and labor in different places simultaneously. It is a globalising mode of production that creates inequalities to reproduce itself: spatial inequalities, like the one between core regions and peripheries; social inequalities, including sexual, racial, and inequality between classes. Capitalism reproduces itself to create value from formerly uncommodified nature and human lives. Frontiers of extraction of nature, land and labor are continuously remade and moved to nourish this expansive and unequal mode of production.

Bernardi and Vanhaute (2021b, p. 87).

As with educational approaches inspired by critical pedagogy and working through antisystemic movements, WSA brings a critical angle to education for sustainable development, known also as global education. As pointed by Connell (2020), dominant patterns of critical knowledge on global inequalities or cultural patterns have been typically produced in the “Global North.” Similarly, Buchowski (2004) discussed the “hierarchies of knowledge” in West-dominated anthropology focused on post-socialist Central and Eastern Europe. While such arguments have limitations in that they may divert attention from *what* (an issue) to the potentially moralizing question of *who* (an author and his or her geographic location), they have merit in signaling the problem of the adequacy of concepts and analytical perspectives. Discussing the case of Poland as a former “Second World” and contemporary semi-peripheral country, Starnawski (2015) and Kuleta-Hulboj (2020) attempt to unpack the well-established notions of “North” and “South.” WSA serves as an essential tool in this task as it helps to account for historical dependencies of the given country as a peripheral zone of Europe and the capitalist world-economy, and its contemporary modernization aspirations and contradictions as a semi-peripheral zone within the European Union. The case of Poland is shaped by West-inspired cosmopolitan patterns of imitative modernization on the one hand and nationalist politics on the other, being both a sending and receiving country in terms of migration, advocating environmental policies and continuing to rely on extractivism (coal mining), belonging to the “rich club” of the OECD and having GDP per capita significantly lower than that of capitalist core zones. As such, Poland, along with many other semiperipheries, is not only an example that resists the simple North-South dichotomy prevalent in global education programs, but one that requires a lens that modifies common notions of global interdependencies. The case of CEE countries may also problematize the understanding of peace education, given the existing tensions concerning international security systems, nationalisms, imperialist rivalries and the quest for peace and stability in the post-communist world.

Introducing themes and perspectives inspired by WSA in curricula within national schooling systems requires efforts in educational policy and practice at all levels of scale. In recent decades, initiatives promoting ideas pertaining to global education both within schooling and via non-formal education, have come from educational projects of NGOs and media linked to social movements. WSA’s conceptual framework positions such movements as essentially antisystemic actors striving to change parameters of the capitalist world-economy though until recently mostly limited to concerns of particular social groups (e.g., industrial workers or suppressed nationalities) and mostly bound territorially (i.e., within particular polities’ legislation and electoral systems). Both past and present, and likely in the future, the antisystemic movements should be seen as more than merely partisan or non-partisan actors within nation-states’ political systems and social structures. They also operate as crucial producers of ideas on how to understand and how to overcome contradictions, inequalities, and destructive tendencies of the capitalist world-economy.

Beyond school curriculum: the role of antisystemic movements

Interrelations between antisystemic movements and legitimate knowledge within social and human sciences are central for WSA’s both historical-epistemological reflection and moral concerns. Following Wallerstein (2014) we see antisystemic movements as those emancipatory social movements that: (1) emerged in the “long 19th century” (Hobsbawm, 1973) and unfolded into what was later called the “old Left,” as a response to socio-economic contradictions of capitalist expansion (labor movements and political parties attached to them, most notably social-democratic, socialist and communist ones, and some radical peasant parties too) and a response to imperial, nationalist and elitist political suppression of certain peoples and ethnicities (national liberation, autonomist and abolitionist movements) and women (suffrage movements); (2) challenged the old Left and the post-1945 world order (US hegemony and Soviet alternative) in the form of anti-militarism, the rekindling of anti-colonial, anti-racist and other ethnic struggles, extending activism into areas of gender (second-wave feminism) and sexuality (LGBT+ movements), maintaining universal human rights as a key concept in opposing dictatorships and calling for a turn to environmentalism; and (3) challenged the capitalist world-economy’s transformations in the late 20th and early 21st centuries, commonly referred to as “globalization” (Wallerstein, 2000b, 2005), claimed global solidarity and justice, and sought to reintegrate the “old Left’s” agenda (labor, public services) with some of the New Left’s concerns (identity, certain degree of anti-statism etc.) (Graeber, 2002; Klein, 2000, 2002; Sen and Waterman, 2009; Wallerstein, 2002), yet beyond the latter’s neoliberal diversions (Fraser, 2009).

Antisystemic movements have contributed to the crystallization of dominant modes of knowledge since the late 19th century. Pro-capitalist liberals and conservatives shared a focus on the idea (though not particular politics) of “development” within the nation-state, shared also with radicals, that is, national-liberation movements and workers’ movements. The centrist *geoculture* linked the three ideological strands together, emphasizing—to varying degrees—the triad of civil and political rights, social policies, and nationhood (most notably shaped via the education system and the army, and later via mass culture, e.g., sports). At the same time, the liberal centrism was an ideological foundation for thematic, epistemological, and disciplinary distinctions in the social and human sciences. Since the emergence of the Cold War order after World War-II and the break-up of the colonial empires, research in area studies was pursued with growing cross-disciplinarity. Critical ideas, such as Latin American dependency theory or the perspective of “development of underdevelopment,” produced by the movements and scholars challenging the imperialist and neocolonial dominations, further prompted the shift in focus of social research.

The challenge against the “old Left” was most fully expressed by a series of political and cultural revolts that took place worldwide around the year 1968. These revolts, however distant geographically and specific politically (from Paris to Chicago to Prague and Warsaw to Mexico, Tunisia and Japan) targeted the inability of the old Left to effect systemic solutions to capitalism, which would be egalitarian and democratic at the same time. They certainly targeted the US international hegemony as well but they also criticized the Soviet-style socialism (Wallerstein, 1989, 2004b). And they inspired new paths of scholarly inquiry in social

and human sciences, leading eventually to the establishment of a number of new fields such as ethnic and race studies, women's and feminist studies, sexuality and queer studies, human rights and peace studies, environmentalism etc. WSA developed as an interplay of a search for a critical universalist perspective in studying historical capitalism (and other world-systems) with an understanding of particularity of knowledge(s) as emphasized by the abovementioned new "studies." The shifts in the patterns of knowledge were eventually institutionalized in new departments, journals and learned societies (or new sections of the latter). These transformations should be seen as effected by groups of intellectuals rooted in and inspired by new antisystemic movements, that continue to be essential actors in seeking, formulating, and spreading new modes of knowledge and posing new problems as political and ethical tasks for the generations facing current uncertainties of the future course of the world-system. The movements thus acted more and more purposefully as educational agents.

Various critical pedagogy or critical education traditions have focused on its potential or promise to contribute to action that transforms the world (e.g., [Cho, 2013](#); [Darder et al., 2015](#); [Freire, 2014](#)). Critical pedagogy looks to critical educators within formal education settings to facilitate this sort of action. [Schultz et al. \(2010\)](#), for example, examine the Social Action Curriculum Project (SACP) initiative in a college classroom, involving student-driven "problem-posing" for "real world problem solving" (p. 370). That project, as one of many others, was grounded in a "justice-oriented" conception of education for democracy, and in curricular reforms that were emergent, relevant, and connected to students' "real-life" problems, while still meeting the formal "standards for every subject area" (p. 378). More broadly, critical pedagogy looks to educational work within and by social and popular movements, that connects to such projects for transformative and creative change. [Stetsenko \(2015\)](#) for example argues for a "Transformative Activist Stance" in social research in which the pursuit of the truth about the world simultaneously "creates the world rather than mirrors it" (p. 113); and in which the act of knowing is about "creating the world and getting to know its truth in the very act of bringing about transformative and creative change" (p. 115). Research included in [Darder et al.'s \(2015\)](#) volume includes international examples of social movements as educational sites, such as [Hickling-Hudson's \(2016\)](#) work on Freirean seminar groups in Grenada and theater arts pedagogy in Jamaica, highlighting for example participants increased "ability to challenge and change entrenched public injustices" (p. 259).

[Freire's \(2014\) *Pedagogy of the Oppressed*](#), first published in 1970, continues to be a foundation for critical educators' work across formal and informal educational settings, and within popular/social movements (for a comprehensive account of the scope and depth of Freire's work see [Darder, 2015, 2018](#)). [Freire \(2014\)](#) makes a crucial contribution to work in popular or social movements through his clear and consistent focus on human liberation, on women and men's "ontological and historical vocation of becoming more fully human" (p. 66), and so the pedagogy of the oppressed as "the pedagogy of people engaged in the fight for their own liberation" (p. 53). The text is grounded in the capacity for cultural and educational work that facilitates liberation as praxis—critical reflection on the world that becomes action to transform the world, avoiding both "verbalism" that stops short of action, and misguided "activism" without serious critical reflection. [Freire's \(2014\)](#) explicit purpose is to work toward people's "adherence to liberation" (p. 168), not through the imposition of an alternative worldview, but through dialogical teaching that involves the "the organized, systematized, and developed 're-presentation' to individuals of the things about which they want to know more" (p. 93), effectively re-presenting reality as problems. The liberatory purposes of Freire's approach, fixed firmly on critical reflection and action to transform reality, connect clearly with WSA's focus on systemic transformation. To cite just one example, [Wallerstein \(1998\)](#) highlighted the need for intellectual work that involved both the critical understanding of social reality and the creative imagining of utopistic alternatives expressed as an "alternative, credibly better, and historically possible (but far from certain) future" (p. 2), followed by political action guided by the project of constructing such alternatives.

[Wallerstein \(2014\)](#) pointed to a paradox about the antisystemic movements. While they challenge the institutions of the capitalist world-economy, at the same time such movements are themselves a product of it, which means that hitherto for the most part massive efforts lead to changes *within* but not *of* the world-system. Wallerstein argues that at the point of bifurcation shaped by the structural crisis of capitalism and its ideologies, and characterized by growing uncertainty about the future of the planet and humanity, there may be more hope for various political efforts to have a lasting, that is, transformative impact on the world-systemic level. The ultimate direction and the results of such efforts cannot be predicted, and the capitalist world-economy may transform into a more exploitative and more authoritarian world-system. Still, the bulk of contemporary antisystemic movements work on ideas and actions aimed at achieving a world-system that would be relatively more egalitarian and more democratic than the world we know. While mainstream social sciences study social movements and political behavior, Wallerstein's notion of *utopistics* meant a more far-reaching task for politically and morally informed scholarship, that is, a systematic reflection on historical choices that individuals and populations are facing in our times.

References

- Aguirre Rojas, C.A., 2000. Rethinking current social sciences: the case of historical discourses in the history of modernity. *J. World Syst. Res.* VI, 750–766.
- Anderson-Levitt, K. (Ed.), 2003. *Local Meanings, Global Schooling: Anthropology and World Culture Theory*. Palgrave Macmillan, New York.
- Apple, M.W., Beane, J.A., 1999. *Democratic Schools: Lessons From the Chalk Face*. Open University Press, Buckingham.
- Beane, J.A., 1995. Curriculum integration and the disciplines of knowledge. *Phi Delta Kappan* 616–622.
- Bernardi, C., Vanhaute, E., 2021a. *A Global History of Humanity Volume 1: Expanding Frontiers: From Foragers to Empires (From 70,000 BCE to 1,000 CE)*. CISP (International Committee for the Development of Peoples), Rome.
- Bernardi, C., Vanhaute, E., 2021b. *A Global History of Humanity Volume 2: Connecting Frontiers: From Empires to a Capitalist World Economy (From 1,000 CE to 1,870 CE)*. CISP (International Committee for the Development of Peoples), Rome.

- Bernardi, C., Vanhaute, E., 2021c. A Global History of Humanity Volume 3: Intensifying Frontiers: From a World of Nations to Global Capitalism (From 1870 CE to 21st Century). CISP (International Committee for the Development of Peoples), Rome.
- Braudel, F., 1982. *On History*. The University of Chicago Press, Chicago.
- Buchowski, M., 2004. Hierarchies of knowledge in central-eastern European anthropology. *Anthropol. East Eur. Rev.* 22, 5–14.
- Cho, S., 2013. *Critical Pedagogy and Social Change: Critical Analysis on the Language of Possibility*. Routledge, New York.
- Connell, R., 2020. *Southern Theory: The Global Dynamics of Knowledge in Social Science*. Routledge, New York.
- Dale, R., 2007. *Educating*. In: Matthewman, S., West-Newman, C.L., Curtis, B. (Eds.), *Being Sociological*. Palgrave Macmillan, London.
- Darder, A., Mayo, P., Paraskeva, J. (Eds.), 2015. *International Critical Pedagogy Reader*. Routledge, New York.
- Darder, A., 2015. *Freire and Education*. Routledge.
- Darder, A., 2018. *The Student Guide to Freire's Pedagogy of the Oppressed*. Bloomsbury Academic, London.
- Drake, S.M., Reid, J.L., 2020. 21st century competencies in light of the history of integrated curriculum. *Front. Educ.* 5.
- Fraser, N., 2009. Feminism, capitalism and the cunning of history. *N. Left Rev.* 97–117. *New Left Review*.
- Freire, P., 2014. *Pedagogy of the Oppressed*. Bloomsbury Academic & Professional, New York.
- Graeber, D., 2002. The new anarchists. *N. Left Rev.* 13, 61–73.
- Harvey, D., 2007. *The Limits of Capital*. Verso, London.
- Hickling-Hudson, A., 2016. Striving for a better world: lessons from Freire in Grenada, Jamaica and Australia. In: Darder, A., Mayo, P., Paraskeva, J., Darder, A., Mayo, P., Paraskeva, J.O. (Eds.), *International Critical Pedagogy Reader/[Electronic Resource]*. Routledge, New York, London, [England. New York, [New York]; London, [England].
- Hobsbawm, E., 1973. *The Age of Revolution 1789–1848*. Abacus, Great Britain.
- Klein, N., 2000. *No Logo: Taking Aim at the Brand Bullies*. Flamingo, London.
- Klein, N., 2002. *Fences and Windows: Dispatches From the Front Lines of the Globalization Debate*. Flamingo, London.
- Kuleta-Hulboj, M., 2020. Critical and postcolonial perspectives on global education: the case of Poland. *J. Soc. Sci. Educ.* 19, 8–22.
- Meyer, J.W., Kamens, D., Benavot, A., Meyer, J.W., 2017. *Introduction. School Knowledge for the Masses: World Models and National Primary Curricular Categories in the Twentieth Century*. Routledge Ltd.
- Report nd. *Report on an Intellectual Project: The Fernand Braudel Center, 1976–1991*.
- Schultz, B.D., Baricovich, J.E., Mcsurley, J.L., 2010. Beyond these tired walls: social action curriculum induction as public pedagogy. In: Sandlin, J.A., Schultz, B.D., Burdick, J. (Eds.), *Handbook of Public Pedagogy: Education and Learning Beyond Schooling*. Routledge.
- Sen, J., Waterman, P. (Eds.), 2009. *World Social Forum: Challenging Empires*. Black Rose Books, Montreal, New York and London.
- Silver, B.J., 2003. *Forces of Labor: Workers' Movements and Globalization since 1870*. Cambridge University Press, Cambridge.
- Sarnawski, M., Gawlicz, K., Duda, D., 2021. Children's educational rights in Poland: policy, school realities and ideological tensions. In: Gillett-Swan, J., Thelander, N. (Eds.), *Children's Rights From International Educational Perspectives. Transdisciplinary Perspectives in Educational Research*. Springer, Cham.
- Sarnawski, M., 2015. Edukacja globalna—perspektywa Polski jako europejskiego kraju półperyferyjnego [Global education—the perspective of Poland as a European semi-peripheral country]. In: Kuleta-Hulboj, M., Gontarska, M. (Eds.), *Edukacja Globalna. Polskie konteksty i inspiracje [Global Education: Polish Contexts and Inspirations]*. Wydawnictwo Naukowe Dolnośląskiej Szkoły Wyższej, Wrocław.
- Stetsenko, A., 2015. Theory for and as social practice of realizing the future: implications from a transformative activist stance. In: Martin, J., Sugarman, J., Slaney, K.L. (Eds.), *The Wiley Handbook of Theoretical and Philosophical Psychology: Methods, Approaches, and New Directions for Social Sciences*. Wiley Blackwell, Canada.
- Tillich, P., 1948. *The Protestant Era*. The University of Chicago Press, Chicago.
- United Nations, 2015. *Transforming our world: the 2030 agenda for sustainable development*. eSocialSciences. <https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>.
- Wallerstein, I., 1989. 1968, revolution in the world-system. *Theor. Soc.* 18, 431–448.
- Wallerstein, I., 1995. Marxism after the collapse of the communisms. In: Wallerstein, I. (Ed.), *After Liberalism*. The New Press, New York.
- Wallerstein, I., 1996. *Open the Social Sciences: Report of the Gulbenkian Commission on the Restructuring of the Social Sciences*. Stanford University Press, Stanford, California.
- Wallerstein, I., 1997. Eurocentrism and its avatars: the dilemmas of social science. *N. Left Rev.* 1, 93–107.
- Wallerstein, I., 1998. *Utopistics: or, Historical Choices of the Twenty-First Century*. The New Press, New York.
- Wallerstein, I., 2000a. From sociology to historical social science: prospects and obstacles. *Br. J. Sociol.* 51, 25–35.
- Wallerstein, I., 2000b. Globalization or the age of transition?: a long-term view of the trajectory of the world-system. *Int. Sociol.* 15, 249–265.
- Wallerstein, I., 2001. *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Temple University Press, Philadelphia.
- Wallerstein, I., 2002. New revolts against the system. *N. Left Rev.* 29–39.
- Wallerstein, I., 2004a. *The Uncertainties of Knowledge*. Temple University Press, Philadelphia.
- Wallerstein, I., 2004b. *World-Systems Analysis: An Introduction*. Duke University Press, London.
- Wallerstein, I., 2005. After developmentalism and globalization, what? *Soc. Forces* 83, 1263–1278.
- Wallerstein, I., 2006. *European Universalism: The Rhetoric of Power*. The New Press, New York.
- Wallerstein, I., 2012. Conclusion: world-systems analysis as a knowledge movement. In: Babones, S.J., Chase-Dunn, C. (Eds.), *Routledge Handbook of World-Systems Analysis*. Routledge, London.
- Wallerstein, I., 2014. Antisystemic movements, yesterday and today. *J. World Syst. Res.* 20, 158–172.

Rethinking critical theory/study of education: contributions of the “posts/new materialism”

Thomas S. Popkewitz^a and Junzi Huang^b, ^a University of Wisconsin-Madison, Madison, WI, United States; and ^b Peking University, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

The political: fabrications in the making of kinds of people	106
The art of governing: reason and fabricating kinds of people	106
Fabricating kinds of people: the exteriorities of interiorities	107
The politics and the political of schooling	108
Affect: attachments of desires and potentialities	109
Temporal ordering and the desires of differences	110
Locating the homeless and location-less reason	111
Critical studies	113
References	114

We focus on the notion of critical studies as concerned with power as a productive force governing exclusions and abjections in its drives for inclusion. Our direct attention is to the infrastructures of pedagogy and research as the affective and cognitive structuring of body and soul. The infrastructures configure and distribute categorical differences about kinds of people through a comparative reason that creates spaces of action. The political lies in the distributions of differences. The effects of power are elided in representational logics of the critical studies in education; at least in the structural and phenomenological traditions populated in Anglo-American traditions. Our strategy is to make visible and thus disrupt the productive rules and standards of the procedures of othering. The representational logics in social and education research

are practices or operating mechanisms which do not explain power, since they presuppose its relations and are content to “fix” them, as part of a function that is not productive but reproductive. There is no State, only state control, and the same holds for all other cases.

(Deleuze, 1986/1988, p. 75.)

The arguments enter into a non-representational logic that eschews the philosophical distinctions of nominalism and realism that have driven critical studies in education. A non-representational approach takes the onto-epistemic forms of the representations of people as having a materiality. The materiality occurs as the representations of people and events “act” as transcendental objectifications of the “Real” and “World”, calculative spaces of action that exclude and abject in the borders of its inclusionary thrusts (Popkewitz, 2020).

The chapter is organized in the following manner. The first section considers the connections of the affective and cognitive structuring of schooling as a material “actor” in the governing of modernity. Our main concern is the comparative reason that fabricates kinds of people through distributing and governing *differences*. This logic and principles, we argue, are elided in the structural and phenomenological traditions of critical studies populated in Anglo-American contexts. The next three sections focus on the affective and cognitive infrastructure in the reason of schooling to examine how the exteriority of representations—abstractions of people, time, and space—constitute architectures and topographies of the subject, the body, and the soul. The affect inscribed in the infrastructure of knowledge and critique embodies desires and potentialities that act as aspirations and fears of the dangers of difference and particular populations.¹ The infrastructure, we argue, invents and stabilizes the comparative reasoning of difference as agential spaces of actions that cross reformist and Hegelian logics in critical theories of education.

The chapter draws on a spectrum of studies related to post-colonial/decolonial and feminist studies and the new materialism as they relate to the field of curriculum studied in education. We engage the broad band of studies but not to find a homogeneity. Our task is to intellectually “play” with the nuances, distinctions and dispositions of these studies to explore the political of knowledge at the interstices of the modern education.² It directs attention to “what is accepted as authority through a critique of the conditions of what is known, what must be done, what may be hoped” (Foucault, 1984, p. 38). Critical theory, as argued here, brings into view an agential notion integral to the Enlightenment attitude toward reason as *critique*.³ The critique we engage is to resist the political of educational practices; their spaces of normalization, differentiation and abjection in their concreteness that are never merely of schooling.

¹For general discussions of affect, see e.g., Deleuze, 1986/1988; Anderson, 2010, 2014; Brinkema, 2014; in education, Demikos et al., 2020; Lesko and Van den Berg, 2019.

²Our use of *modern* and *modernity* are not to suggest a particular epoch or period, but as to think of the particular interstices of historical lines that change over time in epistemes that “act” affectively and cognitively as fold and folding to calculate and produce comparatively kinds of people (see, e.g., Popkewitz, 2008).

³For this discussion, see, e.g., Blumenberg, 1966/1983; Foucault, 1984; also Popkewitz, 2008, 2020.

The political: fabrications in the making of kinds of people

The notion of critique that we engage considers the affective and cognitive structuring of schooling as a material “actor” in the governing of modernity. The art of governing is examined as an infrastructure, the historical and material (re)configuring of differences as a comparative reason fabricating kinds of people.⁴ At the interstices of the Enlightenment, the welfare state, and coloniality that emerged in the long 19th century and continue today are different settlements in governing the distributions of differences. The governing practices perform as the political through its representation logics and onto-epistemic principles. This logic and principles, we argue, are elided in the critical studies populated in Anglo-American traditions.

The art of governing: reason and fabricating kinds of people

The formation of modern nations recognizes the importance of governing through the production of particular kinds of people whose participation was necessary for government. The politics of the state entails new relations and practices. Collective belonging joins with notions of the family, civil participation and the self-governing of individuality as a productive apparatus of power that is different from the prior practices of a sovereign, the Church or the Emperor.

The notion of the welfare state expresses the new relations. The state entailed a reciprocity that operated in care of people to connect public projects with subjectification, acts of inscription of moral order in the economy of the family and schooling that include the invention of childhood that traverse the late 19th up to today (see, e.g., [Hultqvist and Dahlberg, 2001](#); [Zsuzsa, 2014](#)). The spatiality and sociality were intimately, although indirectly, bound to modes of living, and interiority as the degrees of individual freedom.

This governing was not merely the art of governing of the modern welfare state but the economy of subjectification and coloniality. It entails the emergence of an apparatus that connects cognitive and affective structures to calculate representations of people and events. The objectifications of people entail identifiable forms that generated the possibilities for codification and administration. The creative child, for example, is a particular way of thinking, differentiating, and ordering who the child is and should be; a comparative way of thinking about the child who is not “creative” that appears in the 19th century as a concept of science (See, e.g., [Martins, 2017, 2018](#)). The representations of kinds of people, however, have certain specificities that are not merely reproductive and isomorphic, sharply delineated lines but entangled in multiple and mobile lines that are rearranged and reinscribed in settlements of different cultural and social spaces.

Science was a practice that embodied a calculative knowledge. That knowledge was given as the apotheosis of reason as wisdom itself by the turn of the 20th century. Science was the sacred knowledge that connected and generated desires about the potentialities with theories and programs about human development. While with debates and dissents, the European Enlightenments sensibilities about human reason and science were brought into the formation of the republics as salvific themes about collective belonging and the redemptive themes that gave optimism to the potentialities of the future ([Becker, 1932](#); [Bercovitch, 1978](#)). The optimism and potentialities, we will argue, were double gestures in which the anterior topography of the differences were inscribed in the interiority of the body and soul.

The economy of subjectification and coloniality was embodied internationally in the New Educational Fellowship and American Progressive Education in the early decades of the 20th century. An idealized and analytic notion of scientific method as a mode of thinking and logic of classifying the objects of the real world was assembled in the pedagogical practices and the modern school curriculum. Embodied in the notion of scientific method was a logic that instantiated cultural theses about the responsibilities, modes of living, and dispositions of kinds of people ([Rudolph, 2019](#); [Kirchgasler, 2017](#)). The notion of the adolescence, for example, embodies cultural and philosophical ideals of the “adult” that the adolescent is not, yet to become and, at the same time, thinking of “youth” as a period of “natural” transition where the deviances and moral disorder of urban life can be controlled.⁵

The science of observing and governing bodily as preemptive logics of psychology operated as the affective economy. The social sciences, psychology and anthropology folded into pedagogy as an affective and cognitive structuring of indigeneity. Guatemalan schools at the turn of the 20th century produced sciences and a visual culture that differentiated “the Indian” as the exotic other whose body and soul were objects of abjection and civilizing ([López, 2013](#); [Lopez and Coello, 2020](#); also see [McLeod and Paisley, 2016](#)). The normalizing and differentiating capacities of people inform today’s education through ordering how to observe, classify and govern the hyperactivity, attention-deficiency, as well as boredom, depression and lack of motivation among American youths. The care and governing of bodily capacities are often built upon a racialized perspective, for example, of a young, American black teenager as “moody” ([Lesko, 2001](#)).

The reason of sciences is not merely the structuring of knowledge. It has a materiality as the infrastructures in pedagogy are attached to the body and soul. The continual objects of change are the interiority or “the soul” and the body. Modern governing no longer addresses the body and soul as binary categories, instead, the body is turned into a non-corporeal reality and the soul is made into domains of knowledge in the infrastructures of pedagogy and psychology ([Foucault, 1978](#)). [G. Stanley Hall \(1904/1928\)](#), one of the founders of the child study movement at the turn of the 20th century, argued that the psychology of adolescence was to work on “the soul” of the child. Psychology was concerned with the inner moral qualities and characteristics of people that

⁴Our use of infrastructure has homologies to [Kuhn’s, \(1970\)](#) notion of paradigm and [Hacking’s, \(2004\)](#) styles of reasoning. Linking infrastructure with reason in this chapter is to give greater attention to the materiality of epistemology and ontology as a historical practice related to the body and soul.

⁵This relation of giving a nature to things and moral order in science is explored in [Daston \(2019\)](#).

are today expressed as changing people’s “behaviors”, “dispositions” and “mindsets,” that later as not merely about ideations of the mind but as modes of living and being. The non-corporeal “soul”, argues Foucault, is given historically as the power/truth in mastering the body as the bodiless reality. The soul inhabits the body and brings the subject of the child to existence, and “the soul is the prison of the body” (Foucault, 1978).

Fabricating kinds of people: the exteriorities of interiorities

We focus on the practices of pedagogy and its sciences as folds and folding structuring people and differences through its onto-epistemic representations. We initially approach schooling as the fabrication of human kinds. Fabrication is to engage two seemingly different and oppositional meanings in talking about education; that is, the determinant categories that represent the subjects and objects of schooling are simultaneously fictions and a materiality manufacturing differences of kinds of people.

But fabrication is more. It provides a way to explore how the exteriorities of the forms of knowledge work to produce the interiority of the body and the soul. The psychologies of the child as a learner, youth and the adolescent, for example, are desires about the kinds of people that schools are to produce. The notion of the youth and the adolescent child were fictions at the turn of the 20th century, for example, to think about, classify, and plan for correcting the moral disorders associated with the new urban populations that were coming into mass schooling of Western Europe and North America.

The fictions become spaces of action in the manufacturing of kinds of people. The classifications become determinant categories about people that are taken as real. The classification of teachers, social worker, family and community become categories whose objectifications act on as modes of living for the adolescents’ transition to adulthood and make possible the potentialities, for example, of the creative, or the scientifically able child. The classifications connect individualities and communities as a particular intimate closure in a field of futurity. The adolescents’ experience, memory and passages of becoming are enclosed and interned in the anticipated potentialities. They function in everyday life through theories, programs, and stories. The classifications are no longer only to think about people; rather, the categories work their ways into what teachers, children, and family do and the degrees of freedom associated with the actions.

What appears as single objects for reflection are formed through multiplicities; a complex assemblage of ideas, concepts, technologies and principles that produce the objects as “actionable”. These multiplicities of historical lines and foldings of affect and the cognitive structuring are given as settlements of human potentialities and *subjectification*. Notions of childhood and youth, for example, entail folds of political theories of government, abstractions of *the* people, probability theories of populations, and psychological ideas of childhood in the long 19th century to manufacture the citizen as new kinds of person.⁶ The cognitive structuring embody folds of feelings and emotions inscribing particular cultural and normative settlements in the distinctions and ordering practices about the child and childhood in school programs (see, e.g., Friedrich and Colmenares, 2017; Zhao, 2018).

The fictions work into everyday life and have a materiality for ordering of experiences and the degrees of freedom about who people are and should be. The fabrications are spaces of action. The psychology works into the schools as a place where children go in the making and revisioning of the interiority of people. Pedagogy, curriculum and the sciences of learning are the sacred knowledge to effect that governing. The concerns with learning science, mathematics and civic education, for example, occur at the spaces of social psychologies and of the child that work affectively and cognitively as anticipatory to order and “develop” the modes of living of the kinds of people that children are and should be. These kinds of people are often named today as the child who will be an active, participatory, lifelong learner and the mathematically able child. But the names are more than names as they have a materiality in the designing of practices that anticipate modes of living.

The fabrications work as an exterior to enunciate pedagogical designs of an imagined interiority of the able, satisfied, self-motivated and engaged child (Yolcu and Popkewitz, 2018; Andrade-Molina and Valero, 2015). Assembled as an alchemy, the spaces of production in science, mathematics, and the arts are reterritorized in the knowledge practices of schooling as analytics of the subject content of teaching. (Popkewitz, 2018). Phantasmagrams of modes of life are produced through the temporal and ontological distinctions that order the objects of effective teaching, efficient children’s learning, and imagined families and communities. The curriculum is a fold with psychologies about the potentialities of who the child is to become (theories of learning). The body and the soul becomes objects of pedagogical observations activated as learning, attitudes, motivation, and “civic” virtues.

The issues at hand are not the particular naming of the child as representational objects, but how the determinant categories are assembled and connected in particular historical grids that govern the possibilities of existence—making of kinds of people. These possibilities are performed on the body and the interiority or the soul. Today that soul is spoken of as changing teachers’ habits of mind sedimented into the organizational language of benchmarks and empirical evidence. These fictions embody “the double gestures” that we explore later. They affectively generate the hope of progress and fears of populations that prevent that progress (Lesko and Talburt, 2011; Ó, 2003, 2005; McLeod and Paisley, 2016; Kirchgasser, 2019, 2020).

The double quality of fabrication as fiction and manufacturing is given reference to what Hacking (2004) has called dynamic nominalism and Barad (2007) agential realism. The notions of fabrication as a fiction and manufacturing of people correlates

⁶While there is a particular philosophical literature that links the contemporary political idea of the citizen to the Ancient Greek cities, that evolutionary and conceptual history homogenizes and loses site of the historical distinctions of time, space and cosmologies through which life was ordered and its modes of life (see discussion of this issue in Popkewitz, 2008).

to a range of non-representational theories that cross feminist theories that make gender and sexuality as a cultural practice that intersects with the body (Butler, 1993; Judovity, 2001), theories of racialization (Bullock, 2018, 2019; Koza, 2021; Lopez, 2017; Hall, 1986; Gilroy, 1993; Weheliye, 2008; Dixon-Román, 2016; Wynter, 1996, 2001; Pinar, 2004; Britzman, 1991), and post-colonial and de-colonial studies (Spivak, 1999; hooks, 2014; Wynter, 2003; Paraskeva, 2016; Lopez et al., 2019). The corpus of work entails questioning the onto-epistemic principles what is known and how that knowing occurs.

The politics and the political of schooling

The different historical enunciations in the making of the subject are elided in representational logic of the critical studies in education; at least in the Anglo-American traditions. These studies follow undulating pathways given visibility in the middle of the 20th.⁷ That literature, represented in different ways, in the early 1970s educational sociology about “the hidden curriculum.” Expressed in Philip Jackson’s *Life in the Classroom* (1968) and Michael F. D. Young’s *Knowledge and Control* (1971) the concern was with the reproduction of social class structures that exclude. The focus on social class explored, for example, how accountability and reform programs harmed particular social and racial groups produced, inequality, and denied minority community right, voice and representation. Today class has either disappeared or intersects with the objectifications representing race, gender and ethnicity as determinant categories to account for how dominant group are represented in the school to reproduce inequality. Populated with concepts like ideology and hegemony in educational anthropology and sociology of education as well as curriculum studies, differences are expressed today within a universalized, globalized framework of neoliberalism viewed as embodying practices of domination and reproduction.

The politics of representation in these critical studies have been important to political movements to challenge the effects of inequality produced and reproduced in modern institutions. The exclusion produced in welfare and educational institutions through denying or limiting access and opportunities have provided new avenues for the struggles of equal rights. The phantasma of the harmonious world occupied by Northern European and North American families living in white picket fence houses in childrens books are no longer adequate as providing the futurity of education. Questions of diversity, differences and multiculturalism have become mainstreamed in the study of the practices of schooling.

Referenced primarily with Frankfurt critical theory and Gramscian notion of hegemony, Anglo-American critical research has particular limits and paradoxes in enacting its social commitments. There is a duality of the utopic that co-exists with a realism of zero-sum world. The realism embodies and defines representations of different social groups as transcendental signifiers to explain the origins of inequities and correcting the social wrongs. The categories carry the utopic dream of a world that, at least, theoretically, the subaltern or indigenous speaks through the right guidance of Reason as the originary forces of power.

The aporia of the realism manifests in where difference is made thinkable, malleable and subjected to styles of critique. Diversity, a concept valorized in biology, is difference measured, universalized to a presumed, utopic vision of progress and future. The comparative reasoning in the theory of difference manifests what Wells (1938) refers to as a strict parallelism in which categorical and conceptual differences—habits, cultures, races, genders, traditions, etc.—are welcomed but ultimately should be mediated and directed by sameness. The critical pulses are homologous to the grand inclusion in the long nineteenth century when the social problems—the workers, sinners, criminals, the poor, etc., were invented as determinant categories of deviance displaced outside of the individual or religious enterprises of philanthropy and included and incorporated by the state apparatus. The new images of the marginality are objects of knowledge, hospitality and desires of critique institutionalized into the system of self-other relations anticipating a utopic futurity—a symptom of a cruel optimism of the hope of the good life that produces its Others as outside of its liveable spaces (see, e.g., Berlant, 2011).

The zero-sum world represented in the critical research is that of the absolutes of the sovereign and disciplinary power. The world is divided into binary categories of populations that differentiate those who dominate and rule from those who are ruled and repressed. The production of differences embodied in the representations function as the originary points of the real and change. The realism of the objectifications work as a sum-zero game world to define freedom in absolute languages of empowerment, voice, and inclusion, among others. The acts of freedom and empowerment are historically given as outside of the relational assemblies and connections in which they are produced.

The major differences between earlier research and its more recent elaborations of critical theories of education are the changes in subject matter, sometimes with the languages that have resemblances of the non-representational theories but maintain their representational infrastructure (see, e.g., Braidotti, 2013). The conceptual and categorical differences are inscribed in a representational infrastructure that operates with the question of “how” to re-configure and redistribute the relations of the other who becomes natural but not completely true yet (Anderson, 2016). The promotion and articulation of “diversity”, for example, manifests to subjugate *difference* to the coordinates of the unified normativity of the subject, problematizing and “correcting” the disdainful misrepresentation and stereotyping of the minoritized.

The colonial story of the traveling critical theory rarely explores the principles of representation and the realism of the subject as the system of reason that overtly forms its problem. The narratives are of a global history of governing the social and psychic life of

⁷We are casting a general statement about particular principles that seem to cross a wide range of educational literature. We can also differentiate the Anglo-American literature in education to that of Latin American and Japanese Marxist traditions; different from critical literatures of Latin American Liberation and the Catholic Worker Movements, the latter in the work of Pablo Friere but which was given its neomarxist theoretical grounding as it traveled by his interlocutions in North America.

the colonized, of seizing the body and soul of the indigenous and incorporating the reproducibility of a historically particular fold in Western conscience, knowledge and emancipation. The procedures of truth and sublimation traveled with the medical missions and ethnographic science alongside the imperial expansions to Asia, Africa and Latin America in the early twentieth century, and further institutionalized and dispersed into various pedagogical languages and infrastructure to perpetuate the self-other relations. A “non-polemic” space is produced where the apparition of the sensible builds up a hierarchical humanism that traverses the subject to render the marginal in the inscriptions of their own condition (Rancière, 2004).

The style of critique in the studies of the politics of representation misrecognizes the politics of schooling for the political. Politics, the political theorist [Jacque Rancière \(2006\)](#) argues, is the policing of the boundaries and hierarchies of the sensible (and the sensibilities of that sensible) as the object of change. The epistemic structuring and ontologies that maintain the contemporaneous frameworks are re-language but not undone with the politics of representation. Left unscrutinized is a critical analysis of the onto-epistemic principles through which the objects of reflection and action are produced. The normative positiveness of change inscribed in research ignites the sense of the possibility for transformation that creates its own impossibilities. The affective desires for change by centering on “the Child” is assembled as a particular futurity that abjects queer bodies from the boundaries of spaces of the inscribed normalcy ([Edelman, 2004](#)).

Our attention is to a notion of critique in this chapter resonances with Rancière’s differentiating of politics from the political. *The later is the historically enunciated partitions of the sensible and, we can add, the sensibilities* that encloses and interns the partitions of the sensible as the real and materialities for action. [Rancière’s \(2006\)](#) notion of the political has a homology to Deleuze’s theory of difference and repetition, [Foucault’s \(1979\)](#) notion of governmentality and knowledge/critique as cutting concerned with the principles generated in governing the production of desire and potentialities.

Affect: attachments of desires and potentialities

The infrastructure, we have argued, entails affective and cognitive apparatus as folds and folding in the making of kinds of people. This focus is a “decentering” of the subject that Foucault spoke about to examine the inside of the exterior that constitute the representational objects as the phenomena of schooling. [Deleuze’s \(1986\)](#) reading of Foucault viewed this radical critique to “continually submits interiority” (p.53) to the outside that is not fixed. The folds are animated by uneven movements of historical lines “that together make up an inside: they are not something other than the outside, but precisely the inside of the outside.” (p. 54) This strategy brings the interiority (of thought, as the unthought; as finitude, as the ground of subjectivity) to the surface, made exterior to constitute a new topography of the subject, the body, and knowledge in the process ([Brinkema, 2014](#), pp. 22–23).

At this point, we explore the exteriority in the mapping the typography of the objects of schooling as embodying an affective structure. The affective is not as a subsidiary quality of knowledge but entangled in what makes legible the expertise of science, for example, as it settles in policy, social debates, pedagogical practices and research. The focus on affect directs attention away from emotion, feeling, tears, awe, sensation, and aesthetics as a function of moral sentiments, psychological categories of the interior states of individual, or as an identity that functions as entry into the knowledge of a subject ([Brinkema, 2014](#)).

The focus on affective attachment to infrastructures of social life indicates how an exteriority of knowledge holds for and directs aspirations and fears that constitute its inside and interiority. The anticipatory promise of the calculations representing of youth as a category of a kind of person, for example, embodies national and personal aspiration and fears of dangers of diverse publics ([Lesko and Niccolini, 2017](#)). Inscribed as the spatial configuration and distributed in the determinant category of youth is the fabrication of a kind of person, produced in the infrastructure that enables and designs an exteriority to govern the interiority by seizing, dispersing and governing attention, perception, energy and “mood” ([Kowalczyk, 2014, 2016](#)).

Affective structuring works as desires of what should be done to shape action and to manage living. Acting in the rationalities about life and people are the inscriptions of desires about the potentialities of social life and humanness anticipated ([Deleuze, 1995; Deleuze and Parnet, 1977, 1987; Agamben, 1999](#)). Folded in the infrastructure, desire is considered a substantial force of life yet it is directed to what is to be changed and what should be. Inscribed in schooling are desires, express as, for example, the potentialities of kinds of people such as creative child, the expert teacher, and the nurturing mother. The desires are not psychological interiority but the becoming of a historical reality, a predisposition that regulates actions through routine anticipations, habitual responses, patterns of recognition as unconscious and non-conscious characteristics.

The production of desires are elided through different historical lines given as the “non-polemic” space that tells of the truth of schooling. Prominent in this elision are the calculative practices that represent the achievements of students through statistical indicators and indexes of national conditions of education. The OECD’s International Student Assessments, for example that compares approximate 100 nations education is taken as providing for the competencies and well-being of nations and societies. The statistical knowledge and numbers seem as having a mechanical objectivity. *The charts and graphs, however, are imbued with representations that move and mobilize* ([Popkewitz, 2020; Popkewitz, 2022a; Anderson, 2010](#)). The numbers and statistics are embedded in a visual culture that is not merely descriptive of educational conditions ([Dussel, 2013](#)). The data and facts are generative of emotions and feelings, inviting and tasteful as an something that is working toward the fulfillment of national aspirations and in processes of subjectification as actualizing students ability to be economically productive and emotionally prepared for the demands of the uncertain future society that they will encounter.

The affective quality of knowledge is a fold in the calculated domains of numbers visualized in tables, charts, and graphic images. The indexes of educational attainments act effectively to engender an optimism that change can prevent catastrophes in which nations fall behind in the mists of the transformation as well as bring about psychological “well-being”. The statistical calculations visualized in the lists and ranking of PISA generates indexes to locate the progress or failure of nations to grasp the utopic world that PISA assessment assumes and anticipate. The titles of reports, for example, express the measures as significant in locating what nations have to do to stay at the top and what nations need to do to sustain growth and development guided through PISA’s expertise.

The desire of potentialities are formed as philosophical ideals embodied in the classifications and distinctions through which probabilities theories of the statistics are applied. In the international assessments, desires are embodied in the distinctions about children’s academic success assembled with notions of literacy, well-being, motivation, and participation. These qualities of the child are connected as the potentialities of the economic and social prosperity through the pathways identified in the OECD models of change. Generated through the infrastructure are conceptions of the modes of living and dispositions about future needs that are not empirical or evidential, in the sense of merely descriptive of what “is”.

The affective interface between collective belonging and the body and soul is embodied in the cascades of data. The knowledge and skills measured are spoken of as representations of the *literacies, competences and benchmarks* that children need for global competence of the anticipated future (Popkewitz, 2022a). Literacy is affective, a plurality that is “constantly infusing embodied practices, resonating with discourses, coalescing around images, becoming part of institutions, animating a comparative reason in catalyzing political communities, and being known and intervened in, among much else (Anderson, 2014, p.6)”. “Literacy”, at least European contexts, reserved for learning how to read the bible is transposed affectively to generate the necessity of the successful participation in modern societies. Its affect is of a generic, general world, inhabited with a cosmopolitanism that students possess for competence in navigating a global world not yet here but filled with uncertainty. The assessment’s affect of literacy activates desires of potentialities of the reasonable and rational person preparing for the indeterminacies through determined rational knowledge and skills.

Temporal ordering and the desires of differences

The reasonableness given to kinds of people and events of schooling never occur in solitude, but collectively and historically formed within a temporal ordering. Time becomes a social attribute in the identities of differences and changes of human life that crosses reformist and Hegelian logics. Time becomes a social “fact” ordered as processes that compare human development, evolution, and growth. The regularization, irreversibility of time made possible new forms of memory and forgetting that were tied to external sites as museums as a representation of human qualities and characteristics as marks of difference. The past, present and future are placed into “the arrow of time” that moves within the pedagogical practices as well as revisioned in the Hegelian notion of dialectics (for latter, see, e.g., Brinkema, 2014).

The significance of time is its assemblage as a fold in the comparative differentiating of people in hierarchies and continua of values.⁸ In the nineteenth century, for example, there occurred the rise of a wholesale awareness of change, the future, and history, with a Faustian desire for the knowledge of becoming, rather than being. The Faustian becoming brings the relation of the past and the present as giving direction to the future that science anticipates. Research rearticulates and (re)visions John Stuart Mills’ expression: “the idea of comparing one’s own age with former ages, or with our notion of those which are yet to come, had occurred to philosophers; but it never before was itself the dominant idea of any age” (Eksteins, 1985, p. 3).

The movements of a regular, irreversible time focus human attention and perceptions around processes that stabilize the representations of identities by which differences are defined. Linked with the irreversible movement from past, present to future brings into view principles of change, differences and their administration through calculating everyday experiences, such as found in studies of households and labor. The past inhabits the conceptions and classifications that give intelligibility to the objectifications of the child’s development and growth for example, in configured spaces that generate potentialities. The intersection of past, present and future is captured in the notions of youth and adolescent as transitional stages of life, represented in the social and psychological sciences.

Time is an element of subjectification—that is, how the exteriority of forms and knowledge order the interiority for feeling and giving meaning to experience and the trajectories of one’s life. Cultural changes in the 19th century, for example, made it possible to think of one’s life as governed with a planned trajectory of biography in “the arrow of time.” This planning in the arrow of time was articulated in the idea of people having careers, a word that previously described racetracks for horses (Bledstein, 1976). The idea of a career, however, was not only about an individual planning a life but was intertwined with the emergence of the expertise of professional theories for codifying and standardizing who people were and could be as a course of life.

The recognition of people to be included and as different, paradoxically, establishes continua of value between *normalcy and pathology* that haunt the social sciences and psychologies related to education (Popkewitz, 2008). The assemblages of sociality and individuality in the sciences bound critical pulses with the grand inclusion in the long nineteenth century. New distinctions of kinds of people were given as the origin of the social problems, reclassified in the ideas of the workers, criminals, the poor and racialized groups. New forms of philanthropy, the emergence of welfare institutions and settlement houses, social science, and the modern, common school wove into the art of government.

⁸Carl L. Becker, *The Heavenly City of the Eighteenth-Century Philosophers* (New Haven: Yale University Press, 1932).

Installed in this governing was the notion of social and individual improvement as utopic (Hetherington, 2001). The new sociologies and psychologies associated with American Progressive pedagogies and its sciences,⁹ for example, gave focus to the undoing the moral disorder of urban life and its populations (Popkewitz, 2020). With the introduction of “normal schools” in the middle of the 19th century, the professionalization of the teacher concerned the “awakening” the moral character, later replaced on training in communication skills and the standard techniques of teaching (Lee, 2020).

The cultural space of childhood becomes a temporal ordering in the making of the interiority of the “child,” (see, e.g., Steedman, 1995). Today formed as fictions about modes of living, concepts as adolescence, the lifelong learner, and the disadvantaged child instantiate the spatial unity of the whole from which differences are registered through processes of “learning” and notions of well-being, engagement and “grit”. The processes stabilize the unity as representations from which a comparative reason differentiates and divides the qualities, and capabilities of the child. The distinctions and differentiations embody an unspoken but present normativity about the interiority of the child who lacks motivation and engagement and of the “fragile” (or “chaotic”) family.

The representations of the child are stabilized, transcendental signifiers that locate the origins of change. Time as processes become the object of change. The objectifications are represented as the identities of the populations. Change becomes the processes of remediation, rescue and redemption for entrance to the closed space of society and its “humanity.”¹⁰ The stabilizing of space and the privileging of time produce systems of exclusion and abjection. The desires of potentialities generate double gestures that distribute differences about the kinds of people in the enclosed “other” spaces of a temporal order; the “urban” populations of the turn of the 20th century and the 21st century as urban populations who “lack” self-esteem, engagement, grit, and motivation-people and families abjected as being consigned to unlivable spaces (Popkewitz, 1988/2017; Popkewitz et al., 2017; Kirchgasser, 2016, 2018, 2019).

The representational reasoning in the infrastructures of schooling and its sciences conceptually and physically banished populations as not being of *the average*, a concept of space ordered through time. The objectifications that represent and identify differences in a temporal order cast out kinds of people who are made as strange, “foreign,” racialized, and as endangering the future. The principles of comparison reside in the distinctions and differentiations around who the child is or is supposed to be as the norms that produce the unity, but also around its “others,” who do not fit into these spaces of normalcy - the immigrant, the member of an ethnic and radicalized populations, and the child today who falls into the achievement “gap” and is left behind. The representations revision the structuralism and realism of the rulers and ruled into the binary resignified as the learner and the disadvantaged. The production of comparative reason entails, as Casid (2015) argues, “the colonial machinery of dominance” (p. 122).

Locating the homeless and location-less reason

The different dimensions of affective and cognitive structuring that we have explored are entangled with a particular kind of consciousness of modernity. That consciousness embodies particular kinds of abstractions that “act” as representations or determinate categories about kinds of people and differences in social and personal life.¹¹ The abstractions of society, individuality, and populations are such examples that emerge in the long 19th century. They produce phenomena that no longer have a direct relation to the conditions of human activities but arrive as “homeless” and location-less concepts and calculation of life outside of bodies. The homelessness and location-less logic makes childhood as a particular “space” in which time is developmental, turning the child into a prophetic, governable surface and interface. The child becomes an object of its own history contained within notions of nature that brings to bear its contextualization as the movement of stages of development and transitions. The child is objectified through coordinates promoted through affects of potentiality of the good life and flows of experience and memory into universal, flatten spatial relations in a regularized time. The relations of childhood seem to have no history or social location to authorize it other than coordinates of its interior conceptualizations.

The abstractions are universal, globalized concepts without any home or historical location, what Deleuze and Guattari (1980/1987) express as the deterritorializing and reterritorizing. The abstractions of childhood and the citizen, for example, remove the immediacy of experiences by deterritorializing their coordinate in space and time and reterritorialize experiences as phenomena that can be calculable and as seemingly objective knowledge actionable in many different spaces. Whether in China, Spain, Ethiopia or the United States, the abstractions of child appear as universal and real distinctions in teacher education, psychologies and the school curriculum. Generated as seemingly global principles about the potentialities of the child, the abstractions “act” as determinant categories of people that hold promises, although in different ways, of collective progress, individual participation, and the competence for future success. The images, interfaces, and concepts created appear as transcendental planes about life, learning and the self that seem to have no cultural and social boundaries.

The distancing practices of these abstractions enable reflective practices that narrate a human history, collective belonging, and the self in a flow of events that have little to do with face-to-face relations. But the “explanations” create coordinates of those relations that have a materiality. To return to the earlier notion of literacy, it is an abstraction of a school subject historically formed in

⁹The use of Progressive not to delineate a sharply defined temporality given to the past; its configurations embody multiple folds and folding of historical lines that produce objects of desire and change produced, hence the phrase “the long 19th century.”

¹⁰For counter-narratives, (see, for example: Ariès, 1961/1962; Dekker, 2010; Finkelstein, 1989).

¹¹We are focusing on a particular quality of homelessness as it relates to universalizing kinds of people as determinate categories from which differences are produced. As a broader historical mode of reasoning, this quality of distancing, abstracting and ordering relations is embodied in this argument but with different historical and relational approach.

the designing of the religious soul of the child through pedagogical practices that in the late 19th century was a normalizing practice related to national belonging that simultaneously differentiated the abled body from those disabled and abjected (Green, 2003, 2017; Cormack and Green, 2009; Hunter, 1988).

We can now explore more specifically the materiality of the epistemic-ontological principles of schooling. The homeless, location-less phenomena produce relevancies in which dimensions of time and space provide principles of the real and the access to the experience of everyday life. Abstractions about childhood and learners, for example, are objectification activated by teachers as what is real in classrooms and the object of the knowledge and skills required for their professional expertise. The abstractions about teachers' communication skills in the models of teacher professionalism materialize as a universal and global knowledge for forming ways of acting and assessing immediate experience. This new knowledge is a reclassification, ordering and re-spatializing of the objects in which concepts of childhood and psychologies of development represent principles of the real phenomena actionable as learning.

It is not that humans did not use abstractions previously to reason about, for example, theology. What makes this quality of abstractions historically interesting and related to the political of modernity is that the categories and organization of phenomena embody a *homelessness* and *location-less logic* in which the classification and ordering the events of the present are taken as the origin of reflection and desires about human potentialities. The potentialities are about what should be and a comparative reason of abjecting the qualities about what does not “fit” in its normalized spaces. No longer life is organized through the time of gods and theology, *the abstractions function as concepts* that fix interpretive framing for how everyday life is lived, whether in Buenos Aires, New York or Shanghai.

This moving “away” and distancing produce the spatial representations and identities as a particular exterior designed to project and activate an interiority as the pedagogical objectifications. The classifying of the child as motivated, problem solving, self-actualizing and engaged are designed to order the interiority of the child—the soul. The concepts, however, appear as analytical distinct, folding into an ontology of the child as an exteriority of the objects of teaching and childhood. *The distinctions about learning form as the outside/surface/exterior is where design and installation of pedagogy happens.* The design of the surface projects its mastery over communication processes and classroom activities to an imagined, invented interior of the soul that is no longer named as such but as “the learner”, “mindsets”, and learning.

The concepts become determinant categories that deterritorialize the immediacy of experiences into abstractions and deterritorializes experience into concepts as the natural properties of the child (childhood), the school “context” and the pedagogical content. With the mind as the intermediary, the concepts of the child, learning and the curriculum serve to generate principles shaping and fashioning what matters as the spaces of action in the classroom. While *the concepts are given as representations of the objects of the world, they are intricately bound with the conditions that give intelligibility to the knowledge organized to change social life and people.*

The homelessness, however, works materially to produce “homes” of belonging and identities through the pedagogical practices of the school. Importance about these homes in the homeless reason for critical theory is the inscriptions of a comparative reason. The homeless, paradoxically, inscribes norms and rules about the borders of the normativity that excludes and adjoins particular kinds of people; qualities outside of the rules and standards of reason and bandished as in unassimilable spaces to which subjects are bound. This double qualities of potentialities radically exclude through the comparative reason that differentiates, distinguishes and divides capabilities and capacities of people. The abstractions about childhood and learners, for example, locate oneself in a broader world not accessible in immediate experience yet provide principles to experience “motivation”, “creativity” and the mundane practices that racialize and differentiate the body as lacking motivation.

This paradox of double gestures that exclude and abject in efforts to include are historical lines in the interstices of contemporary schooling, and the epistemologies of equality/inequality. Change today speaks about social and educational reforms as changing the “mindsets” and “habits of the mind,” providing “the funds of knowledge” and making the globally competent citizen, the professional teacher, and “the lifelong learner.” These kinds of people inscribed in contemporary teacher effectiveness and models of school improvement are anticipatory, a modeling of institutional life directed as normalizing practices to enable personal happiness and the betterment of society. The notion of change are salvific and redemptive, desires about the potentialities of the literate child that stands in opposition to the child left behind.

The increasing significance of the comparative reasoning of statistics and numbers embodies this homeless-ness and location less logic. The statistical measures and algorithms are fabrications that reclassify and reterritorialize everyday life and traditions through technologies of distancing and representations. Statistical measures perform seemingly without any historicity. Numbers standardize and calculate events and people as centers of calculation or observation decks for seeing, thinking and acting on the phenomena of education as a location-less logic to order “what matters” in schools.

Yet what is taken as the real and the origin of reflection and action through data points create new phenomena for thinking and acting about people and the future potentialities (Popkewitz, 2018). Pedagogy is the reterritorialization of life that creates new phenomena in a temporal order directed to the future. This desired future is expressed, for example, through notions of action, problem-solving and development. The classifications become focal points for what is real and comparative for the precautionary and preventive actions governing social life.

The distancing and location-less quality of modernity is not necessarily bad but continually dangerous. Modern science, as Daston (2011) explores, requires this homeless and location-less logic of abstractions, a process that involves learning to “see” and record events and things through abstractions organized in systematic notes that occurs between the 1600s and 1800s. In the sciences and practices of pedagogy, the homeless-ness and location-less logics are continually inscribed in a comparative reason.

The linking of the child with family and community generated principles as boundaries of inclusion and exclusion in an inscription of the moral order. The moral order (and disorder) affectively overlap with collective notions of “civic virtues” and “the common good” that generate its Others outside its boundaries of normalcy.

The question of a critical study of education being raised in this discussion is bound to the epistemological principles of difference embodied in contemporary reasoning. Our focus has been on the representations that act as identities located in a continua of values generated in the affective and cognitive structuring-the onto-epistemic principles- of schooling. The continua of values proceeds from some notion of sameness, embodied in the distinctions that differentiate in a manner in which the representations normalize and divide the qualities of people. The divisions are embedded in categories and ordering principles that represent identities in the curriculum, institutions, and social and psychological theories of schooling.

Critical studies

Our task joined critical theory’s complicated and ongoing relationship with the Enlightenment. We reference the notion of coloniality in our discussion to recognize the entanglements and multiplicities of forms of knowledge, affect and being in the enunciations of schooling as the disavowing and absorbing of differences (see [Mignolo, 2013](#); [Sobe, 2013](#); [Sobe and Kowalczyk, 2018](#); [Zhao et al., 2022](#)). The comparative reasoning of differences and specifically systems of differentiation presume utopic vision of progress and the future mediated and woven into a sameness as the art of government. The comparative reason has a materiality, articulated as “diversity” in policy, research and school programs that unify and coordinate differences within the normativity of the subject.

This recognition of the political and power embodied in the comparative reason of school has different coordinates, relations, and value to certain peoples/societies while disenfranchizing others as the infrastructures of schooling and its practical sciences settle in different historical spaces. Phrases like Eurocentric elide the varied and multiple settlements and the activations of desires and expectations about experience and futures. Social science traveled to China in the late 19th century, for example, as an agency of Enlightenment. Produced were lens of visualization and grids of intelligibility that colonized notions of reason, freedom and equality to elide any pre-existing thoughts and practices among the Chinese ([Huang, 2022](#)). This particular historical settlement of universalized reason and modernity re-territorializes China into a field of Truth that privileges a particular historical notion of humanism as an agential freedom of “becoming” over *life*.

The potentialities generate double gestures about the kinds of people excluded and abjected the “urban” populations of the turn of the 20th century and again the 21st century as urban populations who “lack” self-esteem, engagement, grit, and motivation-people and families abjected as being outside, as being consigned to unlivable spaces ([Popkewitz, 1988/2017](#); [Popkewitz et al., 2017](#); [Kirchgasler, 2016, 2018, 2019](#)).

The representational reasoning in the infrastructures of schooling and its sciences conceptually and physically banished populations as not being of *the average*. The objectifications that represent and identify differences cast out kinds of people who are made as strange, “foreign,” racialized, and as endangering the future. The principles of comparison reside in the distinctions and differentiations around who is to be, but its “others,” who does not fit into these spaces of normalcy.

If we return to our earlier discussion, today’s representations of critical theories have expressed and privileged the intersectionality of race, gender, ethnicity in Anglo-American literatures but within the onto-epistemic framing that conserves its specializations.¹⁰ Stabilized are the globality of space and regularized time that are distributed unevenly in the global order through a spectrum of state critiques and postcolonial literatures. Haunted by the shadow of Enlightenment related to Cartesian and representational logics, educational reforms are globalized, and often articulated and transcoded within salvific and redemptive logics and desires of the emerging nation-states and local intellectual production.

While haunted with shadows of the European Enlightenment, our discussion recognizes both its singularity and also the multiplicities of connections, assemblies, and disconnections in manifold historical spaces (see, e.g., [Zhao, 2018](#); [Popkewitz et al., 2014](#); [Popkewitz, 2022b](#); [Wu, 2013](#)). Important to the Enlightenment ideas of progress was its critique, the continual asking of how we have arrived at the present and its limits as a method of the possibilities of who we are, who we should be, and change. Critique is to cut into the causalities of the present to make fragile what seems inevitable, cutting into what seems self-evident to give historical contingencies to objectifications and the ontological determinism. To return to Foucault, criticism is a matter of flushing out what is thought and trying to change it: to show that things are not as self-evident as one believed, to see that what is accepted as self-evident will no longer be accepted as such.

Critique of *difference* is to hold open possibilities outside of contemporary frameworks that inscribes differences within the unity of oneness. The historicizing focuses on the contingencies of what would otherwise be taken as the unquestionable objects of thought, action, and change.

As [Deleuze \(1986/1988\)](#) and Foucault argued in the 1980s with an indebtedness to the attitude of the Enlightenment, if not contemporary dogma, critique is the diagnosis and the dermatology of apparatus which power installs on the surface of our being. It is to crack things open yet does not stop where the previous grids of intelligibility are ‘destroyed’ or replaced with a new one. It is not to mythologize power but to make visible the internments, enclosures and forces being produced so as to make possible alternatives.

The strategies of critique we believe are entangled in the spaces of a range of critical studies associated with scholarship in the “posts”, queer and de-colonial literatures. The critical impulses are to resist the spaces of normalization, differentiation and abjection in its concreteness that are never merely “of schooling.”

The engagement in critique is an optimism about the possibilities of change. That optimism about change lies in making unstable what is taken as natural, “bound up more with circumstances than necessities, more arbitrary than self-evident, more a matter of complex, but temporary, historical circumstances than with inevitable anthropological constants” (Foucault, 1988, 156). It embodied an optimistic suspicion of the present. To borrow from Stoler (2016),

Critique here is not about “fault finding” and judgment but about restoring the forms that occlusion takes and the questions that its effects may lead us to ask. Thus, the effort is to understand that occlusion is an ongoing, malleable process, sometimes in a form already congealed and seemingly over as it acts on the present, making of us unwittingly compliant observers, nearly always belated in identifying just how it works. (p. 14).

The testing the limits of the present does not give up notions of freedom and agency. Agency, as explored here, lies in challenging the habitual ways of working and thinking in constituting society and people. This challenging of the habitual ways is “grasping for freedom and liberty” by historically exploring the “practical test of the limits that we may go beyond, and thus work carried out by ourselves upon ourselves as free beings” (Foucault, 1984, p. 47). Agency can be thought of as testing the limits of the present, freeing ourselves from the particular dogma of the present through a resistance to what seems inevitable and necessary.

Reflecting on critique as an agential power is to examine how life is divided, measured and produced as onto-epistemic forms of universal reason and progress. The freedom of agency is not the “intrasection” of agential power by projects of making kinds of people and making the body and “soul.” Critique is to open up collective spaces of action that are not directed to phantasmagrams of the prophetic future but to affirm immanent planes of life and potential art of resistance.

References

- Agamben, G., 1999. In: Heller-Roazen, D. (Ed.), *Potentialities: Collected Essays in Philosophy*. Stanford University Press, Stanford. Trans.
- Anderson, B., 2010. Pre-emption, precaution, preparedness: anticipatory action and future geographies. *Prog. Hum. Geogr.* 34 (6), 777–798.
- Anderson, B., 2014. *Encountering Affect. Capacities, Apparatuses, Conditions*. Ashgate Publishing Company, Surrey, England.
- Anderson, B., 2016. *Taking-place: Non-representational theories and geography*. Routledge.
- Andrade-Molina, M., Valero, P., 2015. The sightless eyes of reason: scientific objectivism and school geometry. In: Paper Presented at the Ninth Congress of European Research in Mathematics Education (CERME), Prague, Czech Republic.
- Ariès, P., 1961/1962. *Centuries of Childhood: A Social History of Family Life*. R. Baldick, Trans.. Vintage Books, New York.
- Barad, K., 2007. Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press, Durham.
- Becker, C., 1932. *The Heavenly City of the Eighteenth-Century Philosophers*. Yale University Press, New Haven.
- Bercovitch, S., 1978. *The American Jeremiad*. University of Wisconsin Press, Madison.
- Berlant, L., 2011. *Cruel Optimism*. Duke University Press, Durham and London.
- Bledstein, B., 1976. *The Culture of Professionalism, the Middle Class, and the Development of Higher Education in America*. Norton & Co, New York.
- Blumenberg, H., 1966, 1983. *The Legitimacy of the Modern Age*. Translated by R. Wallace. MIT Press, Cambridge.
- Braidotti, R., 2013. *The Posthuman*. Polity Press, Cambridge, UK.
- Brinkema, E., 2014. *The Forms of the Affects*. Duke University Press, Durham.
- Britzman, D., 1991. *Practice Makes Practice: A Critical Study of Learning to Teach*. State University of New York Press, Albany.
- Bullock, E.C., 2018. Intersectional analysis in critical mathematics education research: a response to figure hiding. *Rev. Res. Educ.* 42 (1), 122–145.
- Bullock, E.C., 2019. Mathematics curriculum reform as racial remediation: a historical counterstory. In: Davis, J., Jett, C.C. (Eds.), *Critical Race Theory in Mathematics Education*. Routledge, New York, pp. 75–97.
- Butler, J., 1993. *Bodies that Matter: On the Discourse Limits of “sex”*. Routledge, New York.
- Casid, J., 2015. *Scenes of Projection. Recasting the Enlightenment Subject*. University of Minnesota Press, Minneapolis.
- Cormack, P., Green, B., 2009. Re-reading the historical record: curriculum history and the linguistic turn. In: Baker, B. (Ed.), *New Curriculum History*. Sense Publishers, Rotterdam, pp. 223–236.
- Daston, L., 2011. The empire of observation, 1600–1800. In: Daston, L., Lunbeck, E. (Eds.), *Histories of Scientific Observation*. The University of Chicago Press, Chicago, pp. 81–114.
- Daston, L., 2019. *Against Nature*. The MIT Press, Cambridge.
- Dekker, J.J., 2010. *Educational Ambitions in History: Childhood and Education in an Expanding Educational Space from the Seventeenth to the Twentieth Century*. Peter Lang, Frankfurt.
- Deleuze, G., 1995. *Difference and Repetition*. P. Patton, Trans.. Athlone Press of Columbia University, New York.
- Deleuze, G., 1986/1988. *Foucault. S. Hands*, Trans.. University of Minnesota Press, Minneapolis, MN.
- Deleuze, G., Guattari, F., 1980/1987. *A Thousand Plateaus: Capitalism and Schizophrenia*. B. Massumi, Trans.. University of Minnesota Press, Minneapolis.
- Deleuze, G., Parnet, C., 1977/1987. *Dialogues*. H. Tomlinson & B. Habberjam, Trans.. Columbia University Press, New York.
- Dermikos, B., Lesko, N., McCall, S., Niccolini, A. (Eds.), 2020. *The Affective Turn in Education. Theory, Research, and Pedagogies*. Routledge, New York.
- Dixon-Román, E., 2016. Algo-Ritmo: more-than-human performative acts and the racializing assemblages of algorithmic architectures. *Cultur. Stud. Critic. Methodol.* 165, 482–490.
- Dussel, I., 2013. The visual turn in the history of education. In: Popkewitz, T.S. (Ed.), *Rethinking the History of Education: Transnational Perspectives on its Questions, Methods, and Knowledge*. Palgrave, New York, pp. 29–50.
- Eidman, L., 2004. *No Future*. Duke University Press.
- Eksteins, M., 1985. Anthropology and degeneration. In: Chamberlin, J.E., Gilman, S.L. (Eds.), *Degeneration: The Dark Side of Progress*. Columbia University Press, New York, pp. 24–48.
- Friedrich, D., Colmenares, E. (Eds.), 2017. *Resonances of El Chavo del 8 in Latin American childhood schooling*. Bloomsbury, London.

- Finkelstein, B., 1989. *Governing the Young: Teacher Behavior in Popular Primary Schools in Nineteenth-Century United States*. Falmer, New York.
- Foucault, M., 1978. *Discipline and Punish: The Birth of the Prison*. Vintage Books, New York, p. 30.
- Foucault, M., 1979. Governmentality. *Ideol. Consciou.* 6, 5–22.
- Foucault, M., 1984. What is the enlightenment? Was ist Aufklärung? In: Rabinow, P. (Ed.), *The Foucault Reader*. Pantheon Books, New York, pp. 32–51.
- Foucault, M., 1988. Practicing criticism. A. S. a. others, Trans.. In: Kritzman, L. (Ed.), *Politics, Philosophy, Culture: Interviews and Other Writings, 1977–1984*. Routledge, New York, pp. 153–156.
- Gilroy, P., 1993. *The Black Atlantic: Modernity and Double Consciousness*. Harvard University, Cambridge, MA.
- Green, B., 2017. Curriculum, politics and the postmodern; or, beyond the knowledge question in curriculum inquiry. *Currículo sem Fronteiras* 17 (3), 501–514. Published in Portuguese under the title *Curriculo, Política E A Pósmodernidade: além da questão do conhecimento na pesquisa em currículo*. Retrieved from. <http://www.curriculosemfronteiras.org/artigos.htm>.
- Green, B., 2003. (Un)changing English: past, present, future? In: Doecke, B., Homer, D., Nixon, H. (Eds.), *English Teachers at Work: Narrative, Counter Narratives and Arguments*. Wakefield Press, Adelaide, pp. 135–148.
- Hacking, I., 2004. *Historical Ontology*. Harvard University Press, Cambridge, MA.
- Hall, G.S., 1904/1928. *Adolescence: Its Psychology and its Relation to Physiology, Anthropology, Sociology, Sex, Crime, Religion, and Education*, vol. 1. Appleton and Co, New York.
- Hall, S., 1986. The problem of ideology-Marxism without guarantees. *J. Commun. Inq.* 10 (2), 28–43.
- Hetherington, K., 2001. Moderns as ancients: time, space and the discourse of improvement. In: May, J., Thrift, N. (Eds.), *Timespace, Geographies of Temporality*. Routledge, London, pp. 49–72.
- Hooks, B., 2014. *Teaching to Transgress*. Routledge, New York.
- Huang, J., 2022. *Making Underdog: Colonial Desires and Affect in Educational Science and Reformatory Discourse*. Doctoral Dissertation. University of Wisconsin-Madison.
- Hultqvist, K., Dahlberg, G. (Eds.), 2001. *Governing the Child in the New Millennium*. Psychology Press, New York.
- Hunter, I., 1988. *Culture and Government: The Emergence of Literary Education*. Macmillan, Hampshire, England.
- Jackson, P., 1968. *Life in the Classroom*. Rinehart & Winston, New York.
- Judovity, D., 2001. *The Culture of the Body. Genealogies of Modernity*. University of Michigan, Ann Arbor.
- Kirchgasler, C., 2016. The limits of “knowledge for all”: historicizing transnational school reforms in Kenya. *Knowl. Cult.* 4 (2), 73–94.
- Kirchgasler, C., 2018. True grit? making a scientific object and pedagogical tool. *Am. Educ. Res. J.* 55 (4), 693–720.
- Kirchgasler, C., 2019. Haunted data: the colonial residues of transnational school reforms in Kenya. In: Salajan, F., Jules, T. (Eds.), *The Educational Intelligent Economy: Big Data, Artificial Intelligence, Machine Learning, and the Internet of Things in Education*. Emerald, Bingley, pp. 215–232.
- Kirchgasler, K.L., 2017. Scientific Americans: reading intercultural school policy through Italian cosmopolitanism. In: Popkewitz, T.S., Diaz, J., Kirchgasler, C. (Eds.), *A Political Sociology of Educational Knowledge: Studies of Exclusions and Difference*. Routledge, New York, pp. 89–104.
- Kirchgasler, K.L., 2020. Dangers of “making diversity visible”: historicizing metrics of science achievement in U.S. educational policy. In: Fan, G., Popkewitz, T. (Eds.), *The International Handbook of Education Policy Studies: Volume II*. Springer, Singapore, pp. 331–349.
- Kowalczyk, J., 2014. Nationalizing interculturalism: reading intercultural school policy through Italian cosmopolitanism. In: Pereyra, M.A., Franklin, B.M. (Eds.), *Systems of Reason and the Politics of Schooling: School Reform and Sciences of Education in the Tradition of Thomas S. Popkewitz*. Routledge, New York, pp. 277–295.
- Kowalczyk, J., 2016. Homo pontem: teachers as bridges to a multicultural Italy. *Paedagogica Historica. Int. J. History Educ.* 52 (5), 438–452.
- Koza, J.E., 2021. *Destined: Carl Seashore's World of Music, Education, and Eugenics*. University of Michigan, Ann Arbor.
- Kuhn, T., 1970. *The Structure of Scientific Revolutions*, second ed. University of Chicago Press, Chicago.
- Lee, S.Y., 2020. Seeing the difference: anticipatory reasoning of observation and its double gesture in teacher education. *Curric. Inq.* 50 (5), 378–399.
- Lesko, N., 2001. *Act Your Age: A Cultural Construction of Adolescence*. Routledge, New York.
- Lesko, N., Niccolini, A.D., 2017. Progressive: historicizing affect in education. In: Popkewitz, T.S., Diaz, J., Kirchgasler, C. (Eds.), *A Political Sociology of Educational Knowledge: Studies of Exclusions and Difference*. Routledge, New York, pp. 70–86.
- Lesko, N., Talburt, S. (Eds.), 2011. *Youth Studies: Keywords and Movement*. Routledge, New York.
- Lesko, N., van den Berg, G., 2019. New curriculum histories. In: Fitzgerald, T. (Ed.), *Handbook of Historical Studies in Education*. Springer International Handbooks of Education. https://doi.org/10.1007/978-981-10-0942-6_6-1.
- López, L.L., 2013. The problem: historicizing the Guatemalan projection and protection of the “Indian”. In: Popkewitz, T.S. (Ed.), *The “reason” of Schooling: Historicizing Curriculum Studies, Pedagogy, and Teacher Education*. Routledge, New York, pp. 229–246.
- López, L.L., 2017. *The Making of Indigeneity, Curriculum History, and the Liits of Diversity*. Routledge, New York.
- López, L.L., Coello, G., 2020. In: López, L.(L.)L., Coello, G. (Eds.), *Indigenous Futures and Learnings Taking Place*. Routledge, New York.
- López López, L., de Wildt, L., Moodie, N., 2019. ‘I don’t think you’re going to have any aborigines in your world’: minecraft ing terra nullius. *Br. J. Sociol. Educ.* 40 (8), 1037–1054.
- Mignolo, W., 2013. On comparison: who is comparing what and why? In: Felski, R., Friedman, S.S. (Eds.), *Comparison*. Johns Hopkins University Press, Baltimore, MD, pp. 99–199.
- Martins, C.S., 2017. From scribbles to details: the invention of stages of development in drawing and the government of the child. In: Popkewitz, T.S., Diaz, J., Kirchgasler, C. (Eds.), *A Political Sociology of Educational Knowledge: Studies of Exclusions and Difference*. Routledge, New York, pp. 105–118.
- Martins, C.S., 2018. The alchemies of the arts in education: problematizing some of the ingredients of the recipe. In: Jörissen, B., Unterberg, L., Klepacki, L., Engel, J., Flasche, V., Klepacki, T. (Eds.), *Spectra of Transformation*. Waxmann, Munster and New York, pp. 51–67.
- McLeod, J., Paisley, F., 2016. The modernisation of colonialism and the educability of the ‘native’: transpacific knowledge networks and education in the interwar years. *Hist. Educ. Q.* 56 (3), 473–502.
- Ó, J.R.d., 2003. The disciplinary terrains of soul and self-government in the first map of the educational sciences. In: Smeyers, P., Depaepe, M. (Eds.), *Beyond Empiricism: On Criteria for Educational Research (Studia Paedagogica 34)*. Leuven University Press, Belgium, pp. 105–116.
- Ó, J.R.d., 2005. Government of the soul and genesis of the modern educational discourse (1879–1911). *Paedagog. Hist.* 41 (1&2), 243–258.
- Paraskeva, J., 2016. *Curriculum Epistemicides: Towards an Itinerant Theory*. Routledge, New York.
- Pinar, W., 2004. *What Is Curriculum Theory?* Peter Lang, New York.
- Popkewitz, T.S., 2020. *The Impracticality of Practical Research: A History of Contemporary Sciences of Change that Conserve*. University of Michigan Press, Ann Arbor.
- Popkewitz, T.S., 1988/2017. *Teacher Education and Teaching as Struggling for the Soul: A Critical Ethnography*. Routledge, New York.
- Popkewitz, T.S., 2018. What is “really” taught as the content of school subjects? Teaching school subjects as an alchemy. *High Sch. J.* 101 (2), 77–89.
- Popkewitz, T.S., 2008. *Cosmopolitanism and the Age of School Reform: Science, Education, and Making Society by Making the Child*. Routledge, New York.
- Popkewitz, T.S., 2022a. International assessments as the comparative desires and the distributions of differences: infrastructures and coloniality. In: *Discourse: Studies in the Cultural Politics of Education*.
- Popkewitz, T.S., 2022b. Making the scientific self: a location-less logic with locations. In: Zhao, W., Popkewitz, T.S., Autio, T. (Eds.), *Epistemic Translation of Curriculum Knowledge on a Global Landscape: A Historical and International Dialogue*. Routledge, New York, pp. 21–44.
- Popkewitz, T.S., Diaz, J., Kirchgasler, C. (Eds.), 2017. *The Reason of Schooling and Educational Research: Studies of Exclusions and Difference*. Routledge, New York.
- Popkewitz, T.S., Khurshid, A., Zhao, W., 2014. Comparative studies and the reasons of reason: historicizing differences and “seeing” reforms in multiple modernities. In: Vega, L. (Ed.), *Empires, Post-coloniality, and Interculturality: New Challenges for Comparative Education*. Sense, Rotterdam, pp. 21–43.
- Rancière, J., 2006. *Hatred of Democracy*. Translated by S. Corcoran. Verso, London.
- Rancière, J., 2004. *The Politics of Aesthetics: The Distribution of the Sensible*. translated with an introduction by Gabriel Rockhill. Continuum, London; New York.

- Rudolph, J., 2019. *How We Teach Science: What's Changed, and Why it Matters*. Harvard University Press.
- Sobe, N., 2013. Entanglement and transnationalism in the history of American education. In: Popkewitz, T.S. (Ed.), *Rethinking the History of Education: An Intercontinental Perspective on the Questions, Methods, and Knowledge of Schools*. Palgrave MacMillan, New York, pp. 93–108.
- Sobe, N., Kowalczyk, J., 2018. Context, entanglement and assemblage as matters of concern in comparative education research. In: Seddon, T., Ozga, J., Sobe, N.W. (Eds.), *World Yearbook of Education Time-Space and Mobility*. Routledge, New York, pp. 197–204.
- Spivak, G.C., 1999. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Harvard University Press, Cambridge, MA.
- Steedman, C., 1995. *Strange Dislocations: Childhood and the Idea of Human Interiority, 1780-1930*. Harvard University Press, Cambridge, MA.
- Stoler, L.A., 2016. *Duress. Imperial Durabilities in Our Times*. Duke University Press, Durham.
- Weheliye, A.G., 2008. After man. *Am. Lit. Hist.* 20 (1/2), 321–336.
- Wells, H.G., 1938. *World Brain*. Methuen & Co, London.
- Wu, Z., 2013. Chinese mode of historical thinking and its transformation in pedagogical discourse. In: Popkewitz, T.S. (Ed.), *Rethinking the History of Education: Transnational Perspectives on its Questions, Methods, and Knowledge*. Palgrave, New York, pp. 51–74.
- Wynter, S., 1996. Is “development” a purely empirical concept or also teleological? A perspective from “we-the-underdeveloped”. In: Yansane, A.Y. (Ed.), *The Prospects for Recovery and Sustainable Development in Africa*. Greenwood Press, Westport, pp. 299–316.
- Wynter, S., 2001. Towards the sociogenic principle: fanon, the puzzle of conscious experience, of identity and what its like to be “black”. In: Durán-Cogan, M., Gómez-Moriana, A. (Eds.), *National Identity and Sociopolitical Change: Latin America between Marginalization and Integration*. routledge, New York & London, pp. 30–66.
- Wynter, S., 2003. Unsettling the coloniality of being/power/truth/freedom: towards the human, after man, its overrepresentation—an argument. *CR N. Centen. Rev.* 3 (3), 257–337.
- Yolcu, A., Popkewitz, T.S., 2018. Making the Able Body: School Mathematics as a Cultural Practice. *ZDM Mathematics Education* (Advance online publication). <https://doi.org/10.1007/s11858-018-1003-8>.
- Young, M., 1971. *Knowledge and Control; New Directions for the Sociology of Education*. Collier-Macmillan, London.
- Zhao, W., 2018. *China's Education, Curriculum Knowledge, and Cultural Inscriptions: Dancing with the Wind*. Routledge, New York.
- Zhao, W., Popkewitz, T.S., Autio, T. (Eds.), 2022. *Epistemic Colonialism and the Transfer of Curriculum Knowledge across Borders: Applying a Historical Lens to Contest Unilateral Logics*. Routledge, New York.
- Zsuzsa, M., 2014. The cultural politics of “childhood” and “nation”: space, mobility, and a global world. *Glob. Stud. Chidhood* 4 (3), 137–142. www.words.co.uk/GSCH.

Curriculum theory and education history

Daniel Tröhler, University of Vienna, Department of Education, Vienna, Austria

© 2023 Elsevier Ltd. All rights reserved.

Curriculum studies: the cultural configuration of a sub-discipline of education research	118
Denominational roots of the cultural configuration of “curriculum” and “Didaktik”	119
National curricula and national epistemologies	121
National discourses, international research and curriculum theory	122
References	124

In the academic world, “theory” refers to a system of research-based statements that serve to explain parts of reality and its underlying regularities. These explanations often make timeless claims, not least because the research process is characterized by “reliable” and “valid” methods that were developed to describe the actual factuality of this reality, i.e., to generate “objective” knowledge that is strictly separated from and unaffected by the inevitably contingent researching individual (“subjectivity”). Initially developed between the seventeenth and nineteenth centuries for those sciences aiming at discovering natural laws, this model of research began to oscillate around and after 1900 more and more into the exploration of human practices. Armed with an arsenal of this kind of objective knowledge, interventions in social and thus educational realities became more legitimate (“evidence-based”), and ultimately these interventions upstaged negotiations of different views and opinions. These developments paved the way to what can be called “expertocracy” as an alternative to (too much) democracy.

Under these assumptions, theoretical progress has always been interpreted either as a correction of older attempts at explaining parts of reality—a phenomenon to be witnessed often in philosophy or history of education—or as a supplement to them, leading to a more precise or more in-depth explanation—as we find more often in psychologically based research in education. The first case is about the falseness of previous research; the second case is about (as a silent rule) confirmation and differentiation. In both cases, however, new research is performed to formulate theories describing parts of social reality in a way that “adequately” or “objectively” defines a possible necessity and strategies of intervention. In the broad field of all social sciences, there have been few disciplines that have not only quickly placed themselves at the service of these interventionist motives but have also sprung directly from interventionist aspirations as research in education in general and curricular research in particular have done.

In the different histories of science, of knowledge, of theory and even of education as an academic discipline, however, there is no theoretical agreement among the approaches defining the research objects, the research methods and the theoretical insights resulting from the respective research activities in the different research institutions across the world. Rather, we are faced with quite different epistemological configurations of theories which, although they do indeed move internationally, morph in the process (Cowen, 2009, 2021), and they can usually hardly discard the normative premises from which they originated. Accordingly, in order to understand the meaningfulness of their statements, insights or theories, these normative premises must be sought, excavated and understood, an approach for which the concept of archeology is often used (Foucault, 1969; Pocock, 1987, p. 23). It leads to the discovery of different—I call it that way—cultural milieus that are composed of two fundamental and often intertwined normative systems of thought related to the making sense of the world: religion or denomination on the one hand and the nation on the other. Both of these systems of thought are often understood as discourse, discursive formation or *langue* (Billig, 1995, p. 8; Calhoun, 2007, p. 3; Özkirimli, 2010, pp. 206–217, Tröhler, 2011).

Following the basic assumptions of the linguistic and cultural turns in the social sciences, the appropriate attitude or answer to this plurality is obviously not the search for a more comprehensive meta-theory, but it is the understanding of how religious-national or denomi-national cultures each generate different perceptions of what a problem is that is in need of being addressed and accordingly formulated with different (research) questions, which usually entail different research methods, and which, not surprisingly, lead to different kinds of theories. This cultural milieu-sensitive understanding of plurality in both everyday and scientific production of knowledge and theory as different ways of making sense of the world cannot but be historically and comparatively informed. Yet, only few committed researchers and activists in the field of education in general and in curriculum studies in particular are committed to wider historical reconstructions, either because they share an understanding of historical research as dealing merely with a past that no longer has any relevance to today’s challenges or because they want to prevent the usually hidden value limits of their own perception of these challenges from being historically reconstructed. Just as theory traditionally claims to be independent of time, one wants to be “objective” as an actor in the definition of the need and the formulation of theory. These claims are made as if Ludwik Fleck’s ideas of “thought styles,” which are supported (and defended) by “thought collectives,” had not been formulated over 80 years ago (Fleck, 1935/1979), explaining how particular “styles of reasonings” (to refer to Ian Hacking, 1992) arise and are cultivated and handed down in—usually—national associations, which in turn act as powerful gatekeepers in academic journals, academic book-series and thus in academic careers (Tröhler, 2020a).

In contrast to these claims, it seems evident that theories are inevitably historically bound expressions of particular ways of thinking about and making sense of the world, and as a rule they fight for recognition and supremacy. Some, then, become dominant, others regressive, and even if in the historical process they have had an impact on each other, they are not or are

only partly traceable to each other. This is also true, and particularly so, for curriculum theory, which has emerged as a distinct genre or sub-discipline of education above all in the Anglo-Saxon-Protestant milieu and has had, and still has, difficulties in establishing itself as a globally recognized research field, especially on the European continent. This is so for good reasons that can be recognized historically by discovering that the object of research, “curriculum,” is a very specific, and not universally shared, perception of what school should be (doing) in the context of a given society. “Curriculum” is not simply a word that could be translated arbitrarily, but a *parole*, whose meaning it takes from a particular *langue* or normative thought system in the making sense of the world. This being said, “curriculum” is, for instance, something different from the German *Lehrplan*, even though most dictionaries would translate *Lehrplan* into curriculum, and vice versa (the literal translation of “curriculum” is the “plan of teaching” or the “plan of what should be taught”) (Horlacher, 2018). This plurality, already even in the object configuration, makes research incredibly diverse and comparative international cooperation for the purpose of clarifying theory building rather challenging.

The following article attempts to analyze this complexity of curriculum and to reflect on what this means for curriculum theory in four steps. First, the history of the development of curriculum studies as an educational sub-discipline is reconstructed in its specific cultural milieu characterized by particular ideological preferences which became heavily contested in the last third of the twentieth century (1). In a second step, an excursus into the early seventeenth century demonstrates that “curriculum” is associated with a specific denominational religious interpretation of the world and that another religious system contained an idea of education and learning of its own, which was then called *Didaktik*, which differed substantially from “curriculum,” and with which *Lehrplan* was later associated (2). Building on this, the third step shows how these different models became nationalized during the long nineteenth century and were thus understood as an effective means of educating loyal, national citizens in the respective contexts and that this nationalization affected not only schools but also science, both of which are shaped by national epistemologies (3). In a final step, there is a reflection on how research can deal with this institutional and epistemological plurality, and, on this basis, three different theoretical models are formulated, the most comprehensive version of which is tightly connected to a comparative educational history (4).

Curriculum studies: the cultural configuration of a sub-discipline of education research

In the beginning, when research began to deal more systematically with “curriculum,” there was no such concept as “curriculum theory.” One spoke of research activities by using “curriculum inquiry” or “curriculum studies.” These terms date back to the early 1900s, when the United States was faced with the challenge of a massive mass immigration of people who came from countries that were culturally very different from the more or less established national and political culture of the United States (or at least of its cultural and political stakeholders). Since, in the United States, the content of teaching in schools was defined by the individual school communities and locally acting superintendents—and not centrally or at least regionally, as in most parts of Europe—there was or seemed to be a considerable need for academic support to adjust the curricula to the perceived challenges of the local schools.

A first overview of these various studies and inquiries about curriculum reform was published by the Canadian-born Werrett Wallace Charters in his 1923 book *Curriculum construction* (Charters, 1923). None other than one of the main representatives of the so-called social-efficiency-education movement, John Franklin Bobbitt, arguably the first “star” of curriculum construction in the United States, reviewed Charters’ study under the title *The technique of curriculum-making* (Bobbitt, 1923). Thereby, Bobbitt assumed that the professional field in education, which had hitherto been satisfied with rather vague assumptions about the objectives of school education, “has been growing dissatisfied with prescientific methods of guessing at the objectives of education” and has therefore “undertaken the task of scientifically determining its objective” (p. 787). Bobbitt identified Charters’ book as a kind of meta-analysis, by which he honored the author—inexplicitly—as the founder of academic curriculum study, for the book “summarize[s] the results of the many investigations of the past decade and present[s] in systematic form the theory and the technique of curriculum investigation” (p. 787). Hence, the “book is obviously written for scientific investigators in the field of education and for students of the technique to be employed” and precisely “not written ... for the purpose of guiding the practical superintendent in his immediate labors of formulating a curriculum in its totality for next year’s work” (p. 788).

The translation from or implementation of theory into practice, that is, of scientifically identified technologies into the practice of local school authorities, had become part of Bobbitt’s lifetime commitment in the context of an industrializing economy and its alleged social needs (Bobbitt, 1918, 1924), and by far the most famous publication on this subject, *Basic principles of curriculum and instruction*, was published by Bobbitt’s like-minded fellow, Ralph W. Tyler (1949). The “science-based” proposals formulated therein presupposed both the psychological scientism in education, which had been cultivated above all at the University of Chicago and at the Teachers College in New York, and the American characteristic of adapting the curricula locally and from year to year. In a rather intriguing way, Tyler’s *Basic principles*, which aimed at giving advice to local policy makers acting in a unique national characteristic of the radical local principle in school organization, was then praised as a universally valid rational method of problem-solving in education.

Hence, what characterized curriculum studies or curriculum theory in the first 50 or 60 years of its sub-disciplinary existence in education was a very specific concept of cooperation between scientific expertise and school policy on the one hand and a school political reality that was extraordinarily, strongly localist on the other. It represented a political, psychologically based epistemology that was strongly influenced by conservative Protestant, often Presbyterian, circles and aimed at helping a political school reality that was greatly shaped by Congregationalist Protestantism, which was placing communal autonomy at the center. Tyler, who had in fact

once wanted to become a Congregational missionary to Rhodesia (today's Zimbabwe) and who had always professed himself as belonging to the conservative circles in the Congregationalist Church (Tjerandsen and Chall, 1987), made, then, an excellent career in countless committees as a scientific advisor or expert of school policy, which, especially in the post-Sputnik era of the United States, expanded globally almost in a missionary style via the expertocratic culture of the OECD (Tröhler, 2014).

However, in the course of the 1960s, strong intellectual and political-agitative opposition arose against this concentrated charge of system-stabilizing conservative WASP culture (not only) in education, which was later subsumed under the term "reconceptualist movement." In 1977, B. J. Benham and Billy Askins published a literary review with 50 annotated studies identified as part of this *Reconceptualist movement* (Benham and Askins, 1977). The notion of "movement," here, might perhaps be misleading, as one of the authors noted, as it lacked an "organized structure." In addition, he stated that "it is a fluid coalition of individuals who share a similar conceptual frame-work and value system, and who object to the mechanistic production metaphor that predominates in American schooling at this time" (p. 3). Some of the proponents of this movement were neo-Marxists, some democratic socialists, and, a little bit later on, some of them were Foucauldians or they were a blend of these ideologies. One of the first major exponents of this movement, William F. Pinar, remembered later how he and his combatants rejected the traditional role of curriculum and curriculum research in the United States (Pinar, 2004, p. 209):

... the official or institutional curriculum (especially the secondary school curriculum) was settled, more or less, during the early 1960s; it would be directed toward and articulated with post-secondary destinations, primary among them the university and the workplace. It was—and remains—a gendered and racialized rite of (academic) passage. By the early 1960s, what remained of the progressive dream—education for democratization which meant schooling for psycho-social as well as intellectual development—was past.

In this movement, among others, the term "hidden curriculum" became a pejorative battle term. First, it was used in the early 1960s to describe social learning processes in the classroom that were not defined in the formal curriculum:

There is a hidden curriculum in every school which teaches, in addition to reading, writing and arithmetic, how to get along with one's fellow-man and how to think about and evaluate oneself and others ... each student is pressed to conform to the kinds of attitudes, beliefs, mores, and behavior which surround him

Pinderhughes (1964/1966, p. 2196), see also Jackson (1968), Overly (1970).

During the 1960s, this term became one of the centerpieces of the "reconceptualist movement" as it served to expose the capitalist-conservative, and therefore *eo ipso*, unjust system which was exemplified in questions of race, religion, women, migrants and workers (e.g., Apple, 1971; Snyder, 1971). As a research "product," theory should not focus on providing efficient advice to the local stakeholders but reveal the disciplining and oppressive mechanism of the U.S.-American curricula and provide advice and inspiration for more individual or collective creativity, freedom and justice. Out of this movement, the *Journal of Curriculum Theorizing* emerged in 1979, and the first issue included lead articles written by, among others, Michael W. Apple (*On analyzing hegemony*), Janet L. Miller (*Curriculum theory: a recent history*) and William F. Pinar (*What is the Reconceptualization?*).

The activists of this movement have triggered an impressive, but also very diverse, intellectual development in the field of curriculum studies, which has found numerous international collaborators and supporters, especially in those countries that traditionally had close ties with the United States. Notwithstanding, this movement has not fundamentally changed the field of curriculum research or theory to "their favor," but they have broadened it in a sustainable way. Today, curriculum studies—after all, a prestigious division of AERA since 1964—represents an extremely heterogeneous field of research whose original motives, namely research-based policy advice, are still prominently advocated today, even in an explicit polemical rejection of the "reconceptualist movement" (Deng, 2018). However, this internationality disguises the fact that it is nevertheless a relatively one-sided phenomenon, not with regard to what has sometimes been called "developing countries," but at least in particular with regard to Europe and more precisely to Central and Northern Europe. It is not that researchers in these areas do not care about school or the content and skills to be taught in the classroom, but they do so based on a very different understanding of schooling and the organization of teaching and learning. These differences emerged from competing religious cultures that, although in various ways, emphasized the role of education with regard to salvation.

Denominational roots of the cultural configuration of "curriculum" and "Didaktik"

In the 1990s, an international, or at least transatlantic, collaboration discussed the nature of the Anglo-American curriculum tradition as compared to the German *Didaktik* tradition (Gundem et al., 1991; Gundem and Hopmann, 1998; Hopmann and Riquarts, 1995; Doyle, 2017; Sivesind and Luimes, 2017); interestingly, and not coincidentally, as will be shown, the notion of the *Lehrplan* was not included in this comparison. The initiators of this collaboration spoke of "differing attitudes toward curriculum planning and implementation," as the Anglo-American curriculum tradition would build on "method-centered teacher education" and act in local environments, and the German and Nordic *Didaktik* tradition and its teacher education would build on content and operate in a "centralized education system with state control over curricula and syllabi" (Gundem and Hopmann, 1998, p. 1). What made the

conversation particularly difficult, the initiators held, was the fact that the central concepts and terms of the German and Nordic tradition lacked “counterparts in English,” as for instance *Didaktik* or *Bildung* (p. 2)—but the same would of course also apply in the other direction since “curriculum” is not simply “*Lehrplan*,” even if dictionaries would suggest so.

Looking at this phenomenon or challenge of translation from the perspective of the linguistic turn, one can take this phenomenon to a deeper ideological level by, following the suggestion of Ferdinand de Saussure or John G. A. Pocock, understanding these terms as *paroles*, each of which gets its own meaning from a particular *langue* (Tröhler and Horlacher, 2019). It then turns out that the difference between curriculum and *Didaktik* lies on a very fundamental religious-ideological level which is characterized by two different Protestant traditions: the Swiss-Reformed on the one hand and the German-Evangelical on the other. They express a fundamental cultural difference in the construction of curriculum and the perception of subjectivity, which Tero Autio already indicated in 2006 (Autio, 2006). These differences become apparent in one of the first conflicts that arose in the context of schooling and learning, namely, in the example of Johann Amos Comenius and his didactic conception, i.e., in the beginning of the seventeenth century. In this time, there was a generally intense discussion of learning and knowledge that developed between Bacon’s *The advancement of learning* (Bacon, 1605) and Comenius’ *Didactica Magna* [Great didactic], the most famous part of his *Didactica Opera Omnia* (1657). Obviously, most of these discussions took place during one of the greatest tragedies in European history, the Thirty Years’ War (1616–1648), and it included, foremost, Protestant scholars from England, Northern Germany and Moravia (today’s Czech Republic).

While in England an early empiricism fostered an idea of learning, research, new knowledge and progress also for the benefit of the people, an attitude which was reflected among other things in the foundation of the Royal Society (1660) (Tröhler, 2020b), German thinking was caught up in religious idealism in so far as it was assumed that all relevant knowledge had already been given biblically and that it was only a matter of teaching it in an appropriate way. It was the intersection of the emphasis on method, as promoted by Bacon (1605), and on absolute biblical authority that lead the German Lutheran Wolfgang Ratke to the creation of the notion of “*didactica*” after 1610 (Knecht, 1984, pp. 102–104). Didactics was therefore a matter of bringing the spirit of the Lutheran Reformation and the authority of the Bible into schools. Accordingly, Ratke promoted German lessons at school on the basis of Luther’s translation of the Bible as the prime school subject. Apparently, this idea of *Didaktik* was what the Lutheran audience in Germany had wanted, for the concept of didactics spread at lightning speed and had become so popular (and correspondingly ambiguous) that even Comenius could not escape it (Knecht, 1984, pp. 107–108).

Comenius, who was a member of the Protestant Bohemian-Moravian Brethren Unity—actually a third alternative to the Swiss and German Reformation, but with much less success in its institutionalization, had been on the run since 1620 as a result of the Thirty Years’ War. He began to write, from 1626 on, an educational treatise that was only completed years later and published in 1657 in Amsterdam: the *Didactica Opera Omnia* (Comenius, 1657). In 1638, he sent a draft version of it to Joachim Hübner in England, who belonged to the *Hartlieb-circle*, named after the Protestant activist Samuel Hartlieb and to which, for instance, the Scottish Presbyterian John Dury belonged. This *circle*—a precursor of the Royal Society—was often called the “invisible college” of pansophic Protestant activists with strong educational aspirations among whom knowledge was created, collected and exchanged, and even designs for calculating machines, double writing machines, seed machines and siege machines were distributed. They helped persecuted Protestants on the Continent, and they tried, in the end unsuccessfully, to convince Comenius to move to England permanently. Hübner’s response to Comenius’ request for feedback of the draft version of *Didactica Opera Omnia* (Comenius, 1657) reveals the fundamental difference in the understanding of the basic question of schooling, teaching and learning, for Hübner criticized Comenius’ major ideas as belonging to the “didactical sect” by which the rules of teaching were elaborated completely free from any content to be taught, and that, therefore, was in fact senseless (Hübner, 1639/1898, p. 81). This reproach is essential and deserves a closer look.

In the published version of *Didactica Opera Omnia* in 1657, the content of teaching was defined in the most famous part, the *Didactica Magna*. It was, simply, everything, as the subtitle of the *Didactica Magna, Universale Omnes Omnia* (The whole art of teaching all things to all men), indicates. In principle, this included everything that one understands as the object of “curriculum studies”: the learning content (all things), the addressee (all men) and the type of teaching (whole). The defined learning content was divided into the three areas of science, good morals and piety. Comenius’ religious conception of man is optimistic, and, therefore, the mediation between content and addressee—that is, teaching—should be, as Comenius said, “quick, pleasant and thorough” (*compendiose, jucunde, solide*). Built on these principles, Christianity would contain “less darkness, confusion and dispute, but more light, order, peace and quiet” (Comenius, 1657, pp. 15–16). Why, then, had the empiricist Hübner been complaining that Comenius had elaborated rules of teaching free from content?

It was Comenius’ deepest conviction that—anthropologically speaking—man, by his very nature and in spite of his original sinfulness, has within himself all the capacities for good and that by means of didactics he brings out this goodness by educating himself in scholarship, virtue and religion (*Eruditionis, Virtutis, Religionis*) (1657, p. 31). Since man is understood as a microcosm of the (divine) universe, didactics does not mean to learn new things (content) that come from “outside,” but to “only evolve” (*solum evolvi*) what is already given in him by his divine descent (1657, p. 32). Comenius’ understanding of didactics thus implies the process of consciousness-making of divinely predetermined things in a universal cosmos. This conceptualization differs fundamentally from an understanding of learning that the science-oriented Reformed Protestants in England pursued. Admittedly, none of these English exponents of empiricism doubted the divinity of nature, not even Isaac Newton, but at most, they took the order of this nature as the starting point of an inductive empiricism without understanding man as a micro-cosmologic mirror of a divine universe.

While the German Lutherans subsequently thought of the idea of progress—popular everywhere in the course of the seventeenth century—strictly inwardly in the human souls as a development process toward the image of God and his creation, the English Calvinists thought of progress scientifically and thus technically and ultimately socially (Tröhler, 2017). In the Calvinist milieu, with its mission to shape God's kingdom on earth, progress had to be approached in an organized and orderly manner, and, in the field of education, it was the term “curriculum” that was used precisely for this purpose, first, already in 1582, at the Calvinist University of Leyden and then in 1633 at the University of Glasgow (Hamilton, 1980, pp. 12–15):

From this perspective there is a homologous relationship between curriculum and discipline; curriculum was to Calvinist educational practice as discipline was to Calvinist social practice (p. 14).

During these decades of the late sixteenth and seventeenth centuries, a reservoir of ideas was created that was implemented in the establishment of the school system and updated exactly when it was believed that the school would no longer be able to properly fulfill the tasks set in it. Hence, better organization of learning and discipline were conceivable as organizational principles when, at the beginning of the twentieth century, the United States faced the question of how to deal with migration and the social consequences of capitalism. This orientation toward discipline or disciplining subsequently favored, especially against the background of crisis perception, an increasing sciento-methodological orientation that should lead to a certain standardization of the localist school reality in the United States, thus promising the empowerment of future citizens building a godly future. Conversely, in the German tradition of reading and receiving Comenius' work, teaching and learning were less about a strict adherence to curricularly defined content and more about the—according to Comenius, free, pleasurable and progressive—spiritual appropriation of the sacred world. After the aesthetic turn of Lutheran thought in the eighteenth century, this appropriation was called *Bildung*, the harmonious inward perfection of the soul. This tradition was never about knowledge per se, but about godly wisdom that transforms the individual soul of the human being as a human being, while the other, Anglo-American tradition is about curricularly defined knowledge, the transmission of which is removed from chance so that the citizen was formed to transform the world. “Curriculum” represents an idea of method, learning and controlled empowerment, whereas *Didaktik* indicates the appropriation and aesthetic of *Bildung*. This is exactly what perhaps the most prominent European representative of the curriculum-didactics debate in the 1990s meant when he wrote that his didactic concept was influenced by German *Geisteswissenschaftliche Pädagogik* (Klafki, 1998, pp. 307–309, 312–314), a deeply Lutheran discourse of education (Tröhler, 2003) that never aimed at knowledge empowering an “atomistic individual”—this had been the standard accusation of German Lutheran scholars against the political culture of the United States since Max Weber (Tröhler, 2011, pp. 42–49)—but that aimed at the empowering of the “connection of the individual to the general” (Klafki, 2000, p. 93) with the aim of “fundamental reasonableness, to concrete universality” (p. 92).

Against this background it becomes understandable why the term *Lehrplan* has hardly appeared in the mentioned transatlantic collaboration discussing the nature of the Anglo-American curriculum tradition as compared to the German *Didaktik* tradition. Facing an envisaged universality and *Bildung*, the aesthetic transformation of the soul, questions of curriculum as *Lehrplan*, of deciding what knowledge and skills should be taught to future citizens, have hardly played a role. *Lehrplan* or *Lehr-Ordnung*, did not even exist before the second half of the eighteenth century; the first use of the term *Lehr-Plan* is interpreted as the commission of Jesus to his disciples (Hanauische Berichte, 1751). Accordingly, research on *Lehrplan*, that is *Lehrplanforschung*, played, in the context of the Lutheran countries in Europe, at best an extremely marginal role.

National curricula and national epistemologies

While today the terminologies have (almost) lost their religious obviousness, the Protestant reflex to entrust not only all kinds of social problems but also the imaginaries of the self, the world and the future to education has (almost) become a global matter of course (Tröhler, 2016a). Discursively, though not obviously terminologically, educational and curricular thinking are religiously shaped. However,—to add theoretical complexity—since the long nineteenth century, this thinking has been configured and sacralized more and more in the context of what has been called nation-building, i.e., in the context of the nation-states, which have been both sacralized and educationalized and have thus acted as the answer and countermovement to what is called (institutional) secularization.

Since the beginning of the early modern age, when we find the first more or less coordinated attempts at educating the masses of the inhabitants of the different political entities, the business of school education has been more than simply teaching the skills of reading, writing and, at best, arithmetic, but it has been about making children become a certain kind of person, i.e., good Christians, depending on how “good Christian” was defined in each case. In the Age of Enlightenment, which can be dated as beginning in the middle of the seventeenth century, there were first reflections on and reforms of modern academic elite education in the English Commonwealth. Then, around 1700, there were broader educational reforms for social outsiders, mostly in cities of Protestant Europe (children of poor or delinquent parents, orphans), and later for the urban citizenry and then deliberately of political leaders. All these educational conceptions had different ideas about the form, content, mediation and aims of education, and from this huge repertoire of educational conceptions developed during this Age of Enlightenment, the emerging nation-states of the long nineteenth century, triggered in the wake of the French Revolution, were able to be selective in order to turn their inhabitants into loyal citizens (Tröhler, 2016b, 2020b).

Although the relationship between nation-building and education is considered important for research on nationalism, it remains completely under-researched, just as the topic is largely ignored by educational research in general and curriculum research in particular (Tröhler, 2020a). Yet, modern constitutional states have depended on the loyalty of their citizens, which is promoted by citizens identifying themselves with the cultural thesis called “nation,” *their* nation, whose characteristics and dignity they see promoted and defended by the state. The state has the privileged means of promoting this national identity, be it through its own transportation systems, television channels, stamps, currencies, work-free holidays, national teams at the Olympics and world championships, and the school and its curriculum or its *Lehrplan* in connection to *Didaktik*. In the school, we detect subtle practices of “doing nation” and the development of “national literacies” that allow citizens to recognize national symbols and flaggings as theirs as opposed to others (Tröhler, 2020c).

As at the beginning of the early modern era, school education has been about making a certain type of person into a morally loyal national citizen who is imagined quite differently in various places: No one would claim—then or now—that a French citizen (*citoyen*) is basically the same as a German citizen (*Bürger*), or a U.S.-American citizen as a Canadian citizen, and not even the Scandinavians, who claim to be part of a shared Nordic culture, have the same imaginaries when it comes to their national citizenship and identity (Hjerm, 1998; Kellas, 2004, pp. 126–137). Nationalism, as a national sentiment or identity, has to be learned somehow, as no one has been born a citizen. It is not with a single school subject that this is done, however, for civic and political education are not the same as educating the future citizens (Horlacher, 2020), but with a cross-curricular aim that not only concerns the more obvious subjects of history or geography but also includes, for instance, mathematics. As a rule, children learn national mathematics (Boser, 2020), and, since the foundation of the republic, a U.S.-American child does not simply learn the English language (how to spell, write, read) but becomes a U.S.-American by virtue of spelling, writing and reading (Bynack, 1984). Thus, even orthography could equally become a means of nationalization, as in the case of the Netherlands (Schoemaker and Rutten, 2019), the same as the struggle over typefaces being used at school (Boser and Hofmann, 2019). In these fabrics of educating the future loyal citizens, the same historical events become very different stories depending on which nation tells its future citizens these events (Gotling, 2020; Peckovic, 2017).

This adds serious complexity to curriculum research aiming at theory. One could perhaps still deal with this incredible plurality in the fabrication of loyal nationalities if one had a more or less uniformity in the perception of issues or if one had “only” to deal with curriculum and *Didaktik* as two (and there are probably more) fundamentally different ways of conceiving of the essence of schooling. But the epistemologies themselves not only have a dual structure corresponding to the various Protestantisms, they also have national connotations and attempt in complex ways to unfold hegemonic effects in an imperial manner (Fox, 2020). Of course, one can ignore epistemological differences by making one’s own epistemology absolute, and then it is easy to level out these national differences in curriculum and curriculum research and to explore the whole world according to the paradigm of the Anglo-American understanding of curriculum, coming to the conclusion of the advent of one “world culture,” a “world society” and a “world polity” (Meyer, 1980; Boli and Thomas, 1997) that create and result from a “world curriculum” of the “global village,” thereby indicating the “relative unimportance of the national, so far as mass curricular outlines go” (Meyer, 1992, p. 6). However, there is actually very little evidence for this, and certainly not in the contemporary age of coarse nationalism demonstrated by the very many (elected) political leaders on all continents of the earth. Rather, it seems that these theories are more results of a self-fulfilling prophecy (Tröhler, 2009) as they explain less a reality than they actually produce imaginaries that act as depicting reality (Carney et al., 2012, p. 368).

Against this background, it has been suggested to distinguish between “internationalization” understood as an “ethical engagement with difference” and “globalization” understood as “neo-imperialism enforced through economic and educational standardization” (Pinar, 2015, p. 36). With these terms, we are obviously dealing with a fundamental, epistemological challenge that has massive consequences for school policy and curriculum reform. It points to the core of the challenges facing curriculum research and curriculum theory.

National discourses, international research and curriculum theory

So far, two different theoretical models of the curriculum have been discussed and presented which are at different levels of theoretical abstraction. Before presenting a theoretical approach at a third, even higher level of abstraction, these two models will be briefly recapitulated. The first model of theory, at the lowest theoretical abstraction level, aims at an educational technology that is meant to be first developed and then applied, copied and pasted in the same cultural environments according to the doctrine of “what works,” even if international scholars have critically discussed *Why “what works” doesn’t work* (Smeyers and Depaepe, 2008; Biesta, 2007). Even though it is difficult to define this kind of knowledge as an academic “theory,” it must nevertheless be accorded a high degree of attractiveness, firstly, because it can be copied and is therefore inexpensive, and secondly, because it at least promises to implement research, which, to varying degrees, is always promising. Especially in the American idea(1) of social efficiency, this model has a certain plausibility, precisely because it is so closely aligned with the normative premises that had initially given rise to the idea of “curriculum” as a specific form of organizing schooling. This approach or attitude to research is very akin to a policy that does not want to “waste” time in negotiating too much but wants to be all the more efficient, as can be seen from an interview with the chancellor of the New York City Department of Education, Joel Klein, on March 9, 2008. In an interview with the New York Times, Klein told the journalist:

What you need to create ... is a school district that people from other cities can come to and say: “This works. All we’ve got to do is replicate this”
How many billionaires (2008).

It is evident, however, as the “reconceptualist movement” in the United States has criticized—and this critique indicates the second model of theory at a middle theoretical abstraction level of curriculum theory—that this kind of research and policy understanding of the first level does not seek to change the dominant notions of social order and good citizenship. On the contrary, it does not seek to question them, but rather to reinforce them. When the U.S.-American society of the 1960s became for many intellectuals an expression of falsehood, injustice, racism, sexism, oppression and exploitation that had to be fought, some of these indignant intellectuals turned their attention to the curriculum, developing this second model at a middle theoretical abstraction as they reflected the ideological bias of both curriculum and traditional curriculum research which represented the first level of abstraction. However, the discussion of this reflective approach representing the second level of abstraction remained largely within the normative realm that has characterized “curriculum” as a cultural institution and organization of creating future citizens. Hence, “curriculum” was not so much targeted by critics as a particular organizational form of institutionalized settings of teaching and learning at school, but rather for its normativity that preserves the social and political system. Certainly, there were demands that compulsory education should be softened, that the teacher-pupil relationship should be less hierarchical, that the school should not be selective and should place more emphasis on creativity, peacefulness and justice. A differently oriented curriculum should, in other words, by virtue of the normative power inherent in the organizational form of “curriculum,” create a different, more just society, one which is Marxist, socialist, democratic or freed from all unnecessary (or even from all) constraints. Today, after almost 50 years since the “reconceptualization movement,” the concerns are perhaps not less radical, but the forum of discussion has become less dogged. Marxist motives still play a role in curriculum criticism (e.g., Apple, 2001), as does the French post-structuralist inspired criticism of the idea(l) of the empowered self, which is tightly related to the traditional understanding of curriculum (e.g., Popkewitz, 2008). Still others have reformulated their commitment to plural democracy and seen in this context the “self-reflexive individual” willing to communicate with others as the goal of education and curriculum (Pinar, 2004, p. 224).

The sometimes astonishingly refreshing critical radicality of these theories of the second level can even be surpassed once again if they were to become less normative and more descriptive. This includes the epistemological emancipation from both the curriculum and the *Didaktik* traditions, which evidently want to—as all ideologies or discourses do—perpetuate themselves. This indicates a third level of abstraction in theory formation, which inevitably becomes historical, comparative and international, and it needs to be contextualized in a broader context that starts to see education in general as a medium of salvation. At least a first essential step has been undertaken by Pinar in his *International handbook of curriculum research* (Pinar, 2014a), in which curriculum research in 34 countries are described (in 39 reports), some of them written in light of the “reconceptualist movement.” In his own introduction, Pinar attempts to de-culturalize the connotations associated with “curriculum” in order to give the developments and traditions in the different countries the right to present themselves from their own ways of thinking (Pinar, 2014b). “Curriculum” thus necessarily becomes a “floating signifier,” a formal concept that is materialized everywhere, similar to the concept of the citizen. The close relationship between “curriculum” and “citizen” points to the logics of the nation-states in which previously developed educational aspirations and ideas of salvation were organized and instrumentalized by state authorities to strengthen the countries in several respects (Tröhler, 2019). Hence, the concrete organization of schools, their formal architecture consisting of levels and tracks and transition regimes as well as their curricula, textbooks and kinds of teacher training are thus a national configuration of educational concepts that are much older than compulsory schooling itself.

A curriculum theory of this third level of abstraction—transcending the second level of theory formation (the “reconceptualist movement”) that, itself, criticizes the service providers of the first level of abstraction in theory formation—therefore takes into account the powerful motives behind the development of modern compulsory education and the entire educational system, which are both religious in their origins and national in their configurations. They are manifestations of what is labeled as educationalization, not only of social problems and challenges (Smeijers and Depaepe, 2008) but of the world and of the modern self (Tröhler, 2016a). In this context, it is no coincidence that the modern self, which has become a major educational project, is worked on from birth to the grave, as first prominently described by Jean Calvin. In his major work, *Institutes of the Christian religion*, Calvin (1536) states in the very first two sentences of the first article that all wisdom, deserving the name of wisdom, consists of two parts, namely the knowledge of God and the knowledge of ourselves, and that they are interwoven in such a way that it is impossible to decide which knowledge results from the other (Calvin, 1536/1823, p. 1). The self, in this Reformation view, has a duty to be conscious of the Christian doctrine, and a large number of strategies of enforcing self-reflection, such as diaries (Blaak, 2009), have subsequently become media of Protestant self-assurance. Reading and writing thus became prominent cultural techniques with the aim of educating the religious person, of forming the religious self. They were often combined with singing, and only later did arithmetic come into play because this religious self had to be able to calculate in the transformation to an industrial society (Cohen, 1999). As a national self, history, geography or even gymnastics had to be learned and incorporated (Tröhler et al., 2017).

Curriculum theory of this third level of abstraction historically traces religious motives and how they were educationalized and then brought into the organizational form of the school, where they shaped what can be called “curriculum.” The global dissemination of these educationalized models of salvation, which grew out of originally Protestant concerns and ambitions, and how they changed over space and time and became nationalized in the long nineteenth century have certainly not yet been sufficiently researched in an internationally comparative way that respects the different idiosyncrasies in the performativity of the different school systems. There is a need for archaeological reconstructions of how these concerns and ambitions have created the idiosyncratic curricular structures and characteristics that were different in various places and that created different research traditions, which therefore developed theories of the first level of abstraction (service-providers for efficiency) that became criticized by theories of the second level of abstraction (the “reconceptualist movement,” which blamed the theories of the first level of abstraction for being conservatively bourgeois). This reconstruction itself represents the third level of abstraction and is, in fact, as far as we can see

today, the key to understanding internationally institutionalized educational practices. Engaging in this kind of curriculum theory certainly does not imply any aims at advising curriculum-makers but pursues, in principle, more epistemological than practical goals. It can be labeled “practical” in that sense as it provides a broad, historically-informed international-comparative repertoire of ideas that may indeed be intellectually helpful for those engaged with the second level of abstraction who are struggling for ethically “better” standards in curricula as instruments in the making of loyal national citizens. Needless to say, however, this kind of theory does not meet the ideal of theory as objective in the classical sense, but it at least appropriately faces the plurality of curricula. In any case, one cannot expect more from a theory devoted to the practices of human action.

References

- Apple, M.W., 1971. The hidden curriculum and the nature of conflict. *Interchange* 2, 27–40.
- Apple, M.W., 2001. *Educating the “Right” Way: Markets, Standards, God, and Inequality*. Routledge Falmer, New York.
- Autio, T., 2006. *Subjectivity, Curriculum, and Society. Between and Beyond the German Didaktik and Anglo-American Curriculum Studies*. Lawrence Erlbaum Publishers, Mahwah.
- Bacon, F., 1605. *Of the Proficiency and Advancement of Learning, Divine and Humane*. Henrie Tommes, London.
- Benham, B.J., Askins, B., 1977. *Curriculum Theory in the 1970's: The Reconceptualist Movement*, 50 Annotated Sources. Texas Tech University, College of Education, Lubock.
- Biesta, G., 2007. Why “what works” won't work: evidence-based practice and the democratic deficit in educational research. *Educ. Theor.* 57 (1), 1–22. <https://doi.org/10.1111/j.1741-5446.2006.00241.x>.
- Billig, M., 1995. *Banal Nationalism*. SAGE, London.
- Blaak, J., 2009. *Literacy in Everyday Life: Reading and Writing in Early Modern Dutch Diaries*. Brill, Leiden.
- Bobbitt, J.F., 1918. *The Curriculum: A Summary of the Development Concerning the Theory of the Curriculum*. Houghton Mifflin, Boston.
- Bobbitt, J.F., 1923. The technique of curriculum-making [review of Charters, W.W., Curriculum construction]. *Elem. Sch. J.* 23 (10), 787–788.
- Bobbitt, J.F., 1924. *How to Make a Curriculum*. Houghton Mifflin, Boston.
- Boli, J., Thomas, G.M., 1997. World culture in the world polity: a century of international non-governmental organization. *Am. Socio. Rev.* 62 (2), 171–190.
- Boser, L., Hofmann, M., 2019. Fraktur or Antiqua in primary schools? The struggle for a unified typeface in German-speaking Switzerland between the mid-nineteenth and early twentieth centuries. *Paedagog. Hist.* 55 (6), 792–811. <https://doi.org/10.1080/00309230.2019.1653945>.
- Boser, L., 2020. Nations and numbers: elementary mathematics education as a nationalizing tool. *Croat. J. Educ.* 22 (Sp.Ed.No.2), 47–63 (in press).
- Bynack, V.P., 1984. Noah Webster's linguistic thought and the idea of an American national culture. *J. Hist. Ideas* 45 (1), 99–114. <https://doi.org/10.2307/2709333>.
- Calhoun, C., 2007. *Nations Matter: Culture, History, and the Cosmopolitan Dream*. Routledge, New York.
- Calvin, J., 1532/1823. *Institutionen der christlichen Religion*. Büchler'sche Verlagsbuchhandlung, Elberfeld.
- Carney, S., Rappleye, J., Silova, I., 2012. Between faith and science: world culture theory and comparative education. *Comp. Educ. Rev.* 56 (3), 366–393.
- Charters, W.W., 1923. *Curriculum Construction*. Macmillan, New York.
- Cohen, P.C., 1999. *A Calculating People: The Spread of Numeracy in Early America*. Routledge, New York.
- Comenius, J.A., 1657. *Didactica Opera Omnia*. Amsterdam: [s.n.].
- Cowen, R., 2009. The transfer, translation and transformation of educational processes: and their shape-shifting? *Comp. Educ.* 45 (3), 315–327. <https://doi.org/10.1080/03050060903184916>.
- Cowen, R., 2021. Recent developments in comparative education: myths, muddles, and marvels. *Bildungsgeschichte* 11 (1), 11–22.
- Deng, Z., 2018. Contemporary curriculum theorizing: crisis and resolution. *J. Curric. Stud.* 50 (6), 691–710.
- Doyle, W., 2017. The *Didaktik*/curriculum dialogue: what did we learn?. In: Uijens, M., Yilmaki, R.M. (Eds.), *Bridging Educational Leadership, Curriculum Theory and Didaktik*, Educational Governance Research, vol. 5. Springer International Publishing, Cham, pp. 219–227. https://doi.org/10.1007/978-3-319-58650-2_5.
- Fleck, L., 1935/1979. In: Bradley, F., Trenn, T.J. (Eds.), *Genesis and Development of a Scientific Fact*. University of Chicago Press, Chicago.
- Foucault, M., 1969. *L'archéologie du savoir*. Gallimard, Paris.
- Fox, S., 2020. The entanglement of nation and epistemology. glances into the backyard of academic dignity. *Croat. J. Educ.* 22 (2), 193–210.
- Gotling, N., 2020. National textbook narratives and historiography: presenting a same that is never the same. *Croat. J. Educ.* 22 (Sp.Ed.No.2), 65–82.
- Gundem, B.B., Hopmann, S. (Eds.), 1998. *Didaktik and/or Curriculum. An International Dialogue*. Peter Lang, New York.
- Gundem, B.B., Engelsen, B.U., Karseth, B. (Eds.), 1991. *Curriculum Work and Curriculum Content Theory and Practice Contemporary and Historical Perspectives: Conference Papers*. University of Oslo, Institute for Educational Research 10–12 October 1990. University of Oslo, Oslo.
- Hacking, I., 1992. “Style” for historians and philosophers. *Stud. Hist. Philos. Sci.* 23 (1), 1–20.
- Hamilton, D., 1980. Notes on the Origins of the Educational Terms Class and Curriculum. Paper Presented at Annual Convention of the AERA. Retrieved from ERIC, 17 August 2020.
- Hanausche Berichte von Religions-Sachen, Zwölfter Theil, 1751.
- Hjerm, M., 1998. Reconstructing “positive” nationalism: evidence from Norway and Sweden. *Socio. Res.* <https://doi.org/10.5153/sro.163>.
- Hopmann, S., Riquarts, K. (Eds.), *Didaktik und/oder Curriculum. Grundprobleme einer international vergleichenden Didaktik*, pp. 9–36. Weinheim, Beltz.
- Horlacher, R., 2018. The same but different: the German Lehrplan and curriculum. *J. Curric. Stud.* 50 (1), 1–16. <https://doi.org/10.1080/00220272.2017.1307458>.
- Horlacher, R., 2020. Doing nation by civic education. The case of Switzerland. *Croat. J. Educ.* 22 (2), 83–99.
- How many billionaires does it take to fix a school system? N. Y. Times, 2008. <http://www.nytimes.com/2008/03/09/>.
- Hübner, J., 1639/1898. Letter to Comenius. In: Kvačala, J. (Ed.), *Korrespondence Jana Amosa Komenského*, vol. I. Nákladem České Adad, Praz, pp. 73–82.
- Jackson, P.W., 1968. *Life in Classrooms*. Holt, Rinehart and Winston, New York.
- Kellas, J.G., 2004. *Nationalist Politics in Europe: The Constitutional and Electoral Dimensions*. Macmillan, New York.
- Klafki, W., 1998. Characteristics of critical-constructive Didaktik. In: Gundem, B.B., Hopmann, S. (Eds.), *Didaktik and/or Curriculum, An International Dialogue*. Peter Lang, New York, pp. 307–330.
- Klafki, W., 2000. The significance of the classical theories of Bildung for a contemporary concept of Allgemeinbildung. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice: The German Didaktik Tradition*. Erlbaum, Mahwah, pp. 85–107.
- Knecht, I., 1984. Zur Geschichte des Begriffs Didaktik. *Arch. Begr.* 28, 100–122.
- Meyer, J.W., 1980. The world polity and the authority of the nation-state. In: Bergeson, A. (Ed.), *Studies of the Modern World-System*. Academic Press, New York, pp. 109–137.
- Meyer, J.W., 1992. Introduction. In: Meyer, J.W., Kamens, D.H., Benavot, A. (Eds.), *School Knowledge for the Masses: World Models and National Primary Curricular Categories in the Twentieth Century*. Falmer Press, Washington, DC, pp. 1–16.
- Overly, N.V. (Ed.), 1970. *The Unstudied Curriculum: Its Impact on Children*. Association for Supervision and Curriculum Development, Washington, DC.
- Özkirimli, U., 2010. *Theories of Nationalism: A Critical Introduction*, second rev. and extended ed. Palgrave Macmillan, Basingstoke.
- Peckovic, A., 2017. Nation divided: building identities through Bosnian history. *Rev. Soc. Stud.* 4 (1), 23–40.
- Pinar, W.F., 2004. *What is Curriculum Theory?* Lawrence Erlbaum Publishers, Mahwah.

- Pinar, W.F., 2014a. International Handbook of Curriculum Studies, second ed. Routledge, New York.
- Pinar, W.F., 2014b. Introduction. In: Pinar, W.F. (Ed.), International Handbook of Curriculum Studies, second ed. Routledge, New York, pp. 1–16.
- Pinar, W.F., 2015. Internationalization. In: Pinar, W.F. (Ed.), Educational Experience as Lived: Knowledge, History, Alterity. Routledge, New York, pp. 36–46. The selected works of William F. Pinar.
- Pinderhughes, C.A., 1964/1966. Effects of ethnic group formation upon educational processes, personality formation, and mental health. In: Elementary and Secondary Education Act of 1966, Hearings before the Subcommittee on Education of the Committee on Labor and Public Welfare. Printed for the Use of the Committee on Labor and Public Welfare, Washington, DC, pp. 2195–2202.
- Pocock, J.G.A., 1987. The concept of a language and the *métier d'historien*: some considerations on practice. In: Pagden, A. (Ed.), The Languages of Political Theory in Early-Modern Europe. Cambridge University Press, Cambridge, pp. 19–38.
- Popkewitz, T.S., 2008. Cosmopolitanism, the citizen and processes of abjection: the double gestures of pedagogy. In: Peters, M.A., Britton, A., Blee, H. (Eds.), Global Citizenship Education: Philosophy, Theory and Pedagogy. Sense Publishers, Rotterdam, pp. 133–152.
- Schoemaker, B., Rutten, G., 2019. One nation, one spelling, one school: writing education and the nationalization of orthography in The Netherlands (1750–1850). *Paedagog. Hist.* 55 (6), 754–771. <https://doi.org/10.1080/00309230.2019.1622573>.
- Sivesind, K., Luimes, M., 2017. Didaktik and curriculum studies: a European perspective. In: Chi-Kin Lee, J., Kennedy, K.J. (Eds.), Theorizing Teaching and Learning in Asia and Europe: A Conversation between Chinese Curriculum and European Didactics. Routledge, New York, pp. 217–231. <https://doi.org/10.4324/9781315751900-13>.
- Smeijers, P., Depaepe, M., 2008. Educational Research: Why "What Works" Doesn't Work. Springer, Dordrecht.
- Snyder, B.R., 1971. The Hidden Curriculum. Knopf, New York.
- Tjerandsen, C., Chall, M., 1987. Educating America: how Ralph W. Tyler taught America to teach. In: Tjerandsen, C., Chall, M., Tyler, R.W. (Eds.), Education: Curriculum Development and Evaluation. Oral History Transcript. University of California Press, Berkeley.
- Tröhler, D., Horlacher, R., 2019. Histories of ideas and ideas in context. In: Fitzgerald, T. (Ed.), Handbook of Historical Studies in Education. Debates, Tensions, and Directions. Springer, Cham. https://doi.org/10.1007/978-981-10-0942-6_2-1.
- Tröhler, D., Westberg, J., Berg, A. (Eds.), 2017. Physical education and the embodiment of the nation, *Nordic J. Educ. Hist.* 4 (2).
- Tröhler, D., 2003. The discourse of German *Geisteswissenschaftliche Pädagogik*—a contextual reconstruction. *Paedagog. Hist.* 34 (6), 759–778.
- Tröhler, D., 2009. Globalizing globalization: the neo-institutional concept of a world culture. In: Popkewitz, T.S., Rizvi, F. (Eds.), Globalization and the Study of Education [2009 Yearbook of the National Society for the Study of Education], vol. 108, pp. 29–48.
- Tröhler, D., 2011. Languages of Education. Protestant Legacies, National Identities, and Global Aspirations. Routledge, New York.
- Tröhler, D., 2014. Change management in the governance of schooling: the rise of experts, planners, and statistics in the early OECD. *Teach. Coll. Rec.* 116 (9), 13–26.
- Tröhler, D., 2016a. Educationalization of social problems and the educationalization of the modern world. In: Peters, M. (Ed.), Encyclopedia of Educational Philosophy and Theory. Springer, Singapore. https://doi.org/10.1007/978-981-287-532-7_8-1.
- Tröhler, D., 2016b. Curriculum history or the educational construction of Europe in the long nineteenth century. *Eur. Educ. Res. J.* 15 (3), 279–297. <https://doi.org/10.1177/1474904116645111>.
- Tröhler, D., 2017. Progressivism. In: Noblit, G.W. (Ed.), Oxford Research Encyclopedia of Education. Oxford University Press, Oxford. <https://doi.org/10.1093/acrefore/9780190264093.013.111>.
- Tröhler, D., 2019. Curriculum history. In: Rury, J.L., Tamura, E.H. (Eds.), The Oxford Handbook of the History of Education. Oxford University Press, New York, pp. 523–538.
- Tröhler, D., 2020a. Knowledge, medias and communication. In: Harford, J., O'Donoghue, T. (Eds.), A Cultural History of Education in the Modern Age. Bloomsbury, London, pp. 35–56.
- Tröhler, D., 2020b. Learning, progress, and the taming of change: the educational aspirations of the age of enlightenment. In: Tröhler, D. (Ed.), A Cultural History of Education in the Age of Enlightenment. Bloomsbury, London, pp. 1–23.
- Tröhler, D., 2020c. National literacies, or modern education and the art of fabricating national minds. *J. Curric. Stud.* 52 (5), 620–635. <https://doi.org/10.1080/00220272.2020.1786727>.
- Tyler, R.W., 1949. Basic Principles of Curriculum and Instruction. The University of Chicago Press, Chicago.

Curriculum theory and new institutionalism

Andreas Nordin, Linnaeus University, Department of Pedagogy and Learning, Växjö, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	126
New institutionalism	126
Discursive institutionalism	127
Curriculum as a discursive practice	128
Curriculum as an increasingly complicated discourse	129
The contribution of discursive institutionalism to curriculum theory	129
A language to connect the analytical levels of curriculum theory	129
Analytical concepts to understand curriculum change (and continuity)	130
Curriculum actors as agents of change	130
Concluding remarks	130
References	131

Introduction

One of the major concerns in the field of curriculum theory (CT) today has to do with the theoretical understanding of curriculum change. Who are the actors and forces driving curriculum change? I would argue that in order to understand these issues in a globalized world, we have to have theories and theoretical concepts that enable us to move beyond limited modernist approaches focused on linear processes of implementation. In this article, I will build an analytical argument for the use of new institutionalism, and especially its latest branch, discursive institutionalism (DI), as developed by Vivien Schmidt (2008, 2011), as a point of departure for such a dynamic analysis of curriculum change, which has hitherto been used very sparingly within the field of CT (cf. Young, 2015). In building my argument, I take my point of departure from Pinar et al.'s (1995) distinction between curriculum as development and curriculum as understanding. However, instead of treating them as historically exclusive, I see them as complementary. Curriculum as development refers to the different ways in which CT contributes to the development of curriculum texts and syllabuses and produces research of relevance to such developmental processes. Curriculum as understanding, in turn, refers to CT—at the same time—being able to make use of a variety of theoretical approaches in critically examining those very developmental processes. Such a critical approach gives the field of CT relative independence with respect to political, economic, or any other interests that are external to education (cf. Uljens and Ylimäki, 2017). Both development and understanding are thus legitimate and necessary approaches to CT. However, when looking at recent developments, there has been a strong move toward curriculum researchers providing useful knowledge to solve problems relating to policy and/or practice in line with curriculum as development. Following Biesta et al. (2019), I would argue that this development has become problematic as it seems to happen at the expense of educational research making use of its critical potential. Against this background, the contribution of new institutionalism to CT, as discussed in this article, should be seen first and foremost as a contribution to strengthen a theoretical understanding of curriculum, although I would argue that critical analysis also contributes to curriculum development.

New institutionalism

New institutionalism emanated during the 1990s out of the work that John W. Meyer and colleagues at Stanford University undertook in the late 1970s and 1980s. In their analyses, educational institutions were treated as loosely coupled systems in which societal legitimacy replaced technical efficiency as the core organizing principle. The basic assumption was that educational institutions did not seem to conform to general principles of organizational theory. Although loosely coupled, educational institutions, at the same time, seemed to be unexpectedly stable over time. This anomaly to organizational theory was explained by their capability of conforming to norms and values and, in doing so, maintaining the trust and legitimacy of the public at large. However, although acknowledging the role of values, norms, and shared beliefs as constitutive for educational institutions, their analyses still took place within a framework where “[E]ducation was seen as being fully controlled by government and the professions and thus beyond the grip of the market forces” (Meyer and Rowan, 2006, p. 3). During the 1990s, such an analytical approach became increasingly problematic as advanced capitalist countries entered into a period of neoliberalism characterized by flexible labor markets, minimized welfare states, and free-market solutions to economic problems, thus reducing the importance of geographical barriers. These were fundamental societal changes that called for new analytical concepts and approaches in order to understand how nations would adopt new ideas and pursue different policies during the process of reorganizing their political-economic activity. The development of new institutionalism can therefore, in many ways, be seen as a theoretical and analytical response to the societal and institutional

changes that were taking place in the wake of the emerging neoliberal paradigm (Campbell and Pedersen, 2001). Eventually, this new institutionalism developed into different traditions, such as historical institutionalism (HI), rational choice institutionalism (RI), sociological institutionalism (SI), and DI. In general terms, HI directed its analytical focus onto historical regularities and the ways in which historical paths continued to shape and guide future action. RI placed its analytical focus on the incentive structures within which rational people make decisions, and SI focused on the cultural norms framing human action. Taken together, these strands of new institutionalism first and foremost contributed to an understanding of the different ways in which institutional rules operate as constraints to human action. They provided explanatory power for issues of institutional continuity and the way in which humans conform to historical and structural rules, rather than for issues concerning institutional change. Following Schmidt (2008), the fourth tradition, referred to as DI, could be seen as a reaction to this lack of analytical interest in why human institutions change. She suggested that in order to understand institutional change in a globalized world, we would have to turn to communicative concepts, such as discourse and ideas, and adopt a more dialectical approach to human institutions as simultaneously “given (the context within which agents think, speak, and act) and as contingent (as the results of agents’ thoughts, words, and actions)” (Schmidt, 2008, p. 314). Understanding institutions as internal to actors thus adds a reflexive element to the institutional analysis, as “people constantly generate representations of what they do as part of what they do” (Chouliarakis and Fairclough, 1999, p. 25f). DI thus places greater emphasis on agency than the other branches of new institutionalism do, which makes it especially suitable for the analysis of curriculum change.

Discursive institutionalism

As is often the case with different branches of theoretical traditions, there are no Chinese walls to separate them from each other; instead, they overlap in many ways, which is also the case with the different branches of new institutionalism. Therefore, DI should be seen as complementary rather than as exclusive to the other traditions, distinguishing itself by paying greater analytical attention to the role of discourse and ideas. In taking on a communicative approach, DI enables a more dynamic analysis of institutional change. While the other traditions adopt a rather one-directional analytical approach “by positioning institutions as in stable equilibria” (Schmidt, 2011, p. 107), DI takes on a more multidirectional approach following a communicative rationale.

Discursive institutionalism by contrast, takes a more dynamic view of change (and continuity) by concentrating on the substantive ideas developed and conveyed by “sentient” agents in discursive interactions that inform their policy-oriented actions which in turn serve to alter (or maintain) institutions. Schmidt (2011, p. 107).

In acknowledging the possibility of agency as a force that not only maintains but also alters institutions, DI adds explanatory power to the analysis of institutional change. In the following, I will discuss ideas and discourse as central DI concepts enabling a more dynamic and multidirectional approach to institutional analysis.

DI distinguishes between three kinds of ideas: policy ideas, programmatic ideas, and philosophical ideas. Policy ideas are the most rapid form of ideas whereby these new ideas turn up in times of an event or crisis and open up a window of opportunity (Kingdon, 1984). However, as pointed out by Schmidt (2011), this should not be misinterpreted as if it is only when a window of opportunity opens up, triggered by an external event or crisis, that new ideas drive policy change. Policy ideas themselves can also contribute internally to institutional change by building up reform pressure that might open up such a window of opportunity. Such a multidirectional approach suggests that processes of institutional change are more complex than what can be captured in simple formulas of stimuli and response. It also shows that there are no watertight barriers between the communication taking place inside and outside of human institutions; instead, they are intertwined, influencing each other in both directions. In turn, rapidly changing policy ideas are often constitutive elements of more general sets of programmatic ideas. These programmatic ideas change more slowly as they are built into the life and structure of institutions where they function as frames of reference. Although some frames of reference prevail as dominant over time, it can also be the case that different frames of reference at any given time struggle for dominance (cf. Young, 1971). Therefore, frames of reference are here understood not so much as homogeneous sets of ideational elements, but rather as a result of discursive struggles. Finally, philosophical ideas refer to a deeper and longer-lasting form of ideas. These are confined largely to the political and public spheres, operating as broad concepts tied to moral principles and values. Such ideas are manifested in ideologies, public philosophies, and world views, and thus they change very slowly.

Turning to the concept of discourse, DI distinguishes between two types of discursive interaction: coordinative discourse and communicative discourse. The two types of discursive interaction relate to what could be understood as different spheres of communication. Coordinative discourse refers to the policy sphere, involving actors such as policymakers, politicians, government officials, experts, lobbyists, union leaders, and so on, operating in transnational, national, and/or local policy arenas. The discursive interaction taking place in this sphere of communication strives toward cognitive justification among policy actors in different kinds of discourse coalitions. Communicative discourse is directed toward the public sphere and involves actors such as the media, interest groups, social movements, and the everyday talk of people in digital fora as well as in real life. Here, the cognitive ideas deliberated over by policy elites during coordinative discourse are communicated to the public in an attempt to gain normative justification. In the context of DI, a focus on discursive interaction directs attention to the importance of the actors involved, as discourses never

float freely; they are acted out by thinking and speaking agents who communicate their ideas through deliberation and contestation, leading to collective action (Schmidt, 2011). This is important as it makes a difference who speaks to whom about what, where it is done, and for what purpose.

Taking on such a differentiated approach to ideas in understanding discursive interaction as a multidirectional process thus facilitates a more dynamic analysis of institutional change than the other branches of new institutionalism do. The different forms of ideas lend explanatory power to a more profound analysis of both institutional change and continuity. In turn, the division of discursive interaction into coordinative and communicative discourses enables a balanced approach to the analysis of institutional change, taking both structure and agency into account.

Curriculum as a discursive practice

From having had quite a functionalistic approach to the curriculum as something to be developed in an almost technical manner, during the 1970s, CT started to take a greater interest in the curriculum as a societal phenomenon (Pinar et al., 1995). Here, Michael Young's edited book *Knowledge and Control*, published in 1971, was an important contribution to the establishment of a new interpretative paradigm (Torres and Antikainen, 2003). The analytical focus was now directed toward the selection and organization of official knowledge and issues of power and control. CT adopted a more societal approach, trying to understand the role of schools and educational systems in the cultural reproduction of societies. The field of CT no longer regarded "the problems of curriculum and teaching as 'technical' problems, that is, problems of 'how to.' The contemporary field regards the problems of curriculum and teaching as 'why' problems" (Pinar et al., 1995, p. 8). Asking why questions then opened up a more complex understanding of the curriculum as a social practice involving many different actors acting out of different interests and with different resources at their disposal. At the same time, the critical question of why implied that things could be different. The task for CT was therefore not only to provide a better understanding of the current curriculum, but also to discuss what alternatives might be at hand.

Conceptualizing the curriculum as an interactive practice among different stakeholders and interests also led to the adoption of an interactionist approach to the central question of CT: What counts as knowledge? The idea of representational knowledge had to give way to philosophical and sociological approaches emphasizing the social construction of answers to questions of what counted as knowledge. Central here was the idea that all knowledge is related to social interests. In addition, theorists such as Richard Rorty, Jürgen Habermas, John Austin, and John Searle, among others, contributed widely to the understanding of such processes of social construction as fundamentally communicative or discursive in character. One could say that they understood the relationship between language and the world as internal, where language was no longer seen as just a representation of the world, but also as part of it and its continuous reconstruction. As pointed out in this tradition, we do things when we speak to each other, conceptualizing language as a social practice among other social practices (Chouliaraki and Fairclough, 1999). Such an approach directs attention to curriculum as a discursive practice. That what is referred to as the official knowledge of a national or local curriculum is the result of discursive interaction among different stakeholders representing different social interests. And that the official curriculum continues to communicate and challenge as it is being recontextualized into different social settings, while other simultaneously deliberate and plan for new curriculum reforms. Answers to questions of what counts as knowledge are thus never static, but continuously constructed and reconstructed in discursive interaction among a diversity of stakeholders within and between different geographical spaces.

Drawing on the work of Habermas, Fritzell (1996) suggests that formal education as a discursive practice at its most fundamental level can be understood following two rationales. One is a communicative rationale that refers to educational actors' intersubjective construction of meaning, and the other is a system rationale that has to do with issues of governability and strategic decision-making. Although touched upon very briefly here, this division helps us to identify formal education as a teleological practice guided by different rationalities and interests, where some can be said to be of more interest to education than others. Here, Pinar (2019, p. 108) points out that:

[T]he educational point of the public school curriculum is understanding, understanding the relations among academic knowledge, the state of society, processes of self-formation, and the character of the historical moment in which we live, in which others have lived, and in which our descendants may live.

From a curricular point of view, this discursive construction of understanding should not be mistaken for conforming to a pre-determined end, but rather it should be seen as an invitation to engage in curriculum as a complicated conversation between diverse actors (and interests) (Pinar, 2019). In making use of such a discursive conception of the curriculum, the role of CT can never only be about providing easy answers so as to develop policy and practice along a system rationale. Instead, by following a communicative rationale, it also has to include a critical component aiming at provoking educational actors at all levels to critically reflect upon themselves, their practices, and the world (cf. Pinar, 2019). As such, CT can be described as a critical science of change, as "critique is central to the issues of freedom and change" (Popkewitz, 2020, p. 222). Thus, this is not merely a critique for its own sake, but a productive one, provoking people to engage in discursive interaction on curricular issues.

Curriculum as an increasingly complicated discourse

The complicated curriculum discourse has traditionally been conceived as a practice taking place within national borders among a variety of educational actors, such as teachers, students, local school authorities, politicians, policymakers, unions, media, academics, and the public. Through the communicative interaction among this multitude of actors, collective national narratives are developed, expressing national values, shared myths, and symbols, placing people in a collective whole, where the authoritative curriculum-as-document is understood as an interim resolution to such a continuous, ongoing national discourse (Westbury, 2008). Irrespective of the construction of such narratives, the curriculum is, by definition, a selective practice where educational actors struggle for interpretive precedence. It is not only a question of what should count as knowledge, but is also always a question of “Whose knowledge should be taught, by whom, and in what ways?” (Apple, 2008, p. 40). Furthermore, this is not only a question that is discussed by a multitude of actors, as these actors also operate at different levels of curriculum making (Deng and Luke, 2008). Deng and Luke distinguish between three curriculum levels: the institutional/societal level, the programmatic level, and the classroom level. The institutional/societal level refers to the overall educational ideas of a specific society and their public legitimation. The programmatic level, in turn, has to do with the way(s) in which values and content knowledge are organized in curriculum texts and syllabuses. Finally, the classroom level refers to students’ and teachers’ experience of the curriculum as it is acted out in the classroom. Such a multilevel approach contributes considerably to the understanding of the curriculum as a complicated discourse connecting different actors at different levels of curriculum making. However, due to the emergence of new transnational policy spaces influenced by powerful organizations, such as the Association for the Evaluation of Educational Achievements (IEA), the Organization for Economic Co-operation and Development (OECD), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the World Bank, and so on, the nation state has become destabilized as the natural site and scale of educational politics (Lingard and Ozga, 2007). The curriculum discourse has thus become even more complicated, as it is now increasingly intertwined with conversations taking place in the transnational arena, thus challenging the field of CT to make use of theoretical perspectives and concepts by widening its analytical gaze to also include a transnational dimension (Nordin and Sundberg, 2018; Wahlström and Sundberg, 2017).

The argument made in this article is that new institutionalism could contribute to such a multilayered and multidirectional understanding of the curriculum by acknowledging the transnational dimension. Especially DI offers a theoretical language capable of discussing both institutional continuity and change in a complex global context. As pointed out by Uljens and Ylimäki (2015, p. 39), “Discourse-oriented curriculum theory provides a language for talking about the human interaction and interpersonal dimensions of any level, from classroom to transnational.” In the following section, I will bring CT and DI together to further explicate my argument for the relevance of DI to CT in order to better understand this multilayered discursive interaction in a global context.

The contribution of discursive institutionalism to curriculum theory

By understanding the curriculum as a discursive practice taking place in a global context, DI has a lot to offer to CT. In this section, I will point to some of its central contributions. First, I will say something about DI offering a language through which to connect the different analytical levels of CT from the local to the transnational. I will then continue by saying something about the ways in which DI lends explanatory power to issues of curriculum change and continuity, and finally, I will discuss the critical potential of agency.

A language to connect the analytical levels of curriculum theory

The emergence of new transnational policy spaces that revolve around powerful actors, such as the IEA, the OECD, UNESCO, the World Bank, and others, has been revolutionary to education, both in terms of governance and to the way in which education has been incorporated as a natural part of a neoliberal policy discourse. In the last few decades, these transnational policy fields have increasingly come to center around the production and use of numerical data gathered through the use of international large-scale assessments (ILSAs), such as the OECD’s Program for International Student Assessment (PISA), and the IEA’s Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), and many more. Although they are not part of the formal governing structure for national educational systems, these transnational spaces operate as systemless systems (Lawn, 2013) exercising soft power; that is, as an indirect form of governing exercised through competition visualized in league tables and ranking lists, building up pressure for states and regions to act. These policy spaces are by no means value-neutral; they harbor ideas about what norms, values, and knowledge are of most worth. Since they have been built up around the production and use of educational statistics, objective and decontextualized knowledge have become imperative here. The knowledge produced in the educational system is only seen as useful if proven objective. Although this is not talked about as an educational ideal, that is what it has become, as educational systems have started to adjust to the political quest for objective knowledge at all curriculum levels (cf. Nordin, 2018). Among world-culture theorists coming from the tradition of new institutionalism, these processes have been discussed in terms of an increased world-culture of schooling, as national school systems conform to discourses dominating the transnational policy space (cf. Meyer and Ramirez, 2012). In so doing, world-culture theory has provided important insights into the ways in which national school systems conform to external pressure. However, at the same time, the approach has received criticism for not

paying enough attention to the influence of national and local contexts in processes of translation and transformation as global ideas are recontextualized within national borders (cf. [Nordin and Sundberg, 2014](#)). However, I would argue that adopting a discursive approach provides a theoretical language enabling CT to overcome such contrapositions connecting the different levels and facilitating a better understanding of processes of change as ideas are reinterpreted and translated when traveling within and/or between different curriculum levels. Furthermore, what new institutionalism has shown is that such processes can no longer be understood solely as processes taking place within national contexts. Therefore, adding a transnational dimension to the analysis of curriculum change seems vital in order to be able to critically examine the reconfiguration of national educational systems today (cf. [Autio, 2009](#); [Uljen, 2018](#)). Here, new institutionalism offers a fruitful approach to the different ways in which transnational policy spaces and their powerful actors today operate as part of national educational institutions, governing at a distance through soft power.

Analytical concepts to understand curriculum change (and continuity)

DI contributes to a nuanced understanding of curriculum change and continuity as a multidirectional process in adopting a stratified approach to the concept of ideas. First, in conceptualizing ideas as the substantive content of discourse, DI acknowledges that educational discussions on knowledge always mean talking about something, which, irrespective of interpretation, cannot be reduced simply to discourse. Questions of how are always related to questions of what. However, having said that, it is also true that ideas never float freely; they are carried by thinking and speaking actors operating out of different interests and with different resources at their disposal. Answers to the question of what counts as official knowledge are therefore never given in advance; instead, curriculum actors are involved in a continuous struggle for discursive dominance. The lesson to be learned from new institutionalism is that issues of what counts as knowledge in a society, and in schools, are discursive in character. Using a stratified concept of ideas enables curriculum analyses of such discourses by acknowledging aspects of both time and space. By using such a stratified approach to ideas, DI contributes to a context-sensitive analysis of curriculum change and continuity by showing that when decontextualized transnational discourses are recontextualized within geographical borders, they are confronted with a variety of institutional ideas, of which some can easily be changed, while others are more resilient, depending on the contextual conditions and to what extent they are part of more deeply rooted philosophical ideas. Such a stratified approach to the discursive interaction of ideas thus facilitates a multidirectional analysis of what happens when ideas travel within and/or between different curriculum levels, and why some ideas change in the complicated discourse of the curriculum while others do not.

Curriculum actors as agents of change

DI criticizes the other traditions of new institutionalism for not paying enough attention to agency. Although, from slightly different positions, they all treat human institutions as relatively stable rule-following structures external to human actors. As such, institutional rules serve mainly as constraints for human thinking and action. In adopting a dialectical approach, DI instead sees human institutions as “internal to the actors serving both as structures that constrain actors and as constructs created and changed by those actors” ([Schmidt, 2008](#), p. 14). Institutions are thus understood as reflexive discursive practices, placing greater emphasis on agency, as people, through their discursive actions, are acknowledged as capable of not just maintaining but also of altering the very institutions of which they themselves are part. From such a point of departure, educational actors at all levels are recognized as agents of change in the complicated curriculum discourse. Of course, some actors do become more influential than others due to differences regarding resources, but from a theoretical point of view, it is important that all actors who are involved are recognized as possible agents of change, since windows of opportunity not only open up as a response to external stimuli, but because the actors themselves can build up pressure to open up such windows, making change possible. A focus on agency thus carries a critical component into the analyses by acknowledging that things could be different. As pointed out by [Popkewitz \(2020, p. 222\)](#), “[C]ritique is an optimism about the possibilities of change.” Such a critical approach to CT means directing why questions to processes of institutional rule-following, examining the norms, values, and ideas guiding such processes, and what actors and interests are involved. The critical potential of CT then lies in its relative independence from any interest external to education, be it political, economic, or other, enabling curriculum researchers to pose critical questions about what is of most worth to education. In making use of its critical potential, curriculum research can then contribute to opening up discourses that are seemingly inescapable and closed in order to facilitate a broader understanding of curriculum change in a globalized context. By insistently asking inconvenient questions of what might be of most worth to education, CT then contributes to curriculum development not just through a provision of solutions to practical problems, but also through provocation, challenging educational actors to make use of their critical potential.

Concluding remarks

In this article, I have argued for the usefulness of DI to CT and to the study of curriculum change. Adopting a DI approach enables the dynamic and multidirectional analyses of the curriculum as a complicated discursive practice. DI also provides a theoretical language for connecting different curriculum levels and for analyzing processes of change as ideas travel within and/or between

them. In understanding institutions as internal to humans, DI positions educational actors at all curriculum levels as agents of change. As such, educational actors always hold the potential of both maintaining and altering the very institutions of which they themselves are part.

To finish, I would like to stress that in a time when the functionalist rationale has once again come to dominate curriculum research in many parts of the world, the need for a critically oriented CT has become more important than ever. Without such research, CT runs the risk of getting out of balance if its analytical interest is narrowed down to a matter of finding solutions to practical problems formulated by educational policy and/or practice. In order to maintain a theoretical balance between curriculum as development and curriculum as understanding, CT also has to conduct critical research, asking inconvenient questions, and provoking actors at all curriculum levels to think differently about the educational institutions of which they are part. What I have tried to do in this article is to argue for the usefulness of DI in such critical and reflexive analyses of curriculum change (and continuity) in today's global educational context.

References

- Apple, M., 2008. Curriculum planning: content, form, and the politics of accountability. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. SAGE, Los Angeles, pp. 25–44.
- Autio, T., 2009. Globalization, curriculum and new belongingness of subjectivity: beyond the nexus between psychology (curriculum) and the nation-state (Didaktik). In: Ropo, E., Autio, T. (Eds.), *International Conversations on Curriculum Studies*. Sense, Dordrecht, pp. 1–20.
- Biesta, G., Filippakou, O., Wainwright, E., Aldridge, D., 2019. Why educational research should not just solve problems, but should cause them as well. *Br. Educ. Res. J.* 45 (1), 1–4. <https://doi.org/10.1002/berj.3509>.
- Campbell, J.L., Pedersen, O.K., 2001. The rise of neoliberalism and institutional analysis. In: Campbell, J.L., Pedersen, O.K. (Eds.), *The Rise of Neoliberalism and Institutional Analysis*. Princeton University Press, Princeton, pp. 1–23.
- Chouliarakis, L., Fairclough, N., 1999. *Discourse in Late Modernity: Rethinking Critical Discourse Analysis*. Edinburgh University Press, Edinburgh.
- Deng, Z., Luke, A., 2008. Subject matter: defining and theorizing school subjects. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, Los Angeles, pp. 66–87.
- Fritzell, C., 1996. Pedagogical split vision. *Educ. Theor.* 46 (2), 203–216.
- Kingdon, J.W., 1984. *Agendas, Alternatives and Public Policies*. Little, Brown and Company, New York.
- Lawn, M., 2013. The internalization of education data: exhibitions, tests, standards and associations. In: Lawn, M. (Ed.), *The Rise of Data in Education Systems*. Symposium Books, Oxford, pp. 11–25.
- Lingard, B., Ozga, J., 2007. Introduction: reading education policy and politics. In: Lingard, B., Ozga, J. (Eds.), *The RoutledgeFalmer Reader in Education Policy and Politics*. Routledge, London, pp. 1–8.
- Meyer, J.W., Ramirez, F.O., 2012. The world institutionalization of education. In: Schriewer, J. (Ed.), *Discourse Formation in Comparative Education*. Peter Lang, Berlin, pp. 111–132.
- Meyer, H.D., Rowan, B., 2006. Institutional analysis and the study of education. In: Meyer, H.D., Rowan, B. (Eds.), *The New Institutionalism in Education*. State University of New York Press, Albany, pp. 1–13.
- Nordin, A., Sundberg, D., 2014. The making and governing of knowledge in the educational field. In: Nordin, A., Sundberg, D. (Eds.), *Transnational Policy-Flows in European Education: The Making and Governing of Knowledge in the Education Policy Field*. Symposium Books, Oxford, pp. 9–20.
- Nordin, A., Sundberg, D., 2018. Exploring curriculum change using discursive institutionalism: a conceptual framework. *J. Curric. Stud.* 50 (6), 820–835. <https://doi.org/10.1080/00220272.2018.1482961>.
- Nordin, A., 2018. The idea of distance in data-driven curriculum policy making: a productive critique. *TCI* 15 (2), 26–37.
- Pinar, W.F., Reynolds, W.M., Slattery, P., Taubman, P.M. (Eds.), 1995. *Understanding Curriculum*. Peter Lang, New York.
- Pinar, W.F., 2019. *What is Curriculum Theory?* Routledge, New York.
- Popkewitz, T.S., 2020. *The Impracticability of Practical Research: A History of Contemporary Sciences of Change that Conserve*. University of Michigan Press, Ann Arbor.
- Schmidt, V.A., 2008. Discursive institutionalism: the explanatory power of ideas and discourse. *Annu. Rev. Polit. Sci.* 11, 303–326.
- Schmidt, V.A., 2011. Speaking of change: why discourse is key to the dynamics of policy transformation. *Crit. Pol. Stud.* 5 (2), 106–126.
- Torres, C.A., Antikainen, A., 2003. *The International Handbook on the Sociology of Education: An International Assessment of New Research and Theory*. Rowman & Littlefield, Lanham.
- Uljens, M., Ylimaki, R., 2015. Towards a discursive and non-affirmative framework for curriculum studies, Didaktik and educational leadership. *Nordic J. Stud. Educ. Policy* 2015 (3). <https://doi.org/10.3402/nstep.v1.30177>.
- Uljens, M., Ylimaki, R., 2017. Non-affirmative theory of education as a foundation for curriculum studies, didaktik and educational leadership. In: Uljens, M., Ylimaki, R.M. (Eds.), *Bridging Educational Leadership, Curriculum Theory and Didaktik: Non-affirmative Theory of Education*. Springer, Cham, Switzerland, pp. 3–150.
- Uljens, M., 2018. A contribution to re-theorizing curriculum research. *TCI* 15 (2). <https://ojs.library.ubc.ca/index.php/tci/index>.
- Wahlström, N., Sundberg, D., 2017. Discursive institutionalism towards a framework for analyzing the relation between policy and curriculum. *J. Educ. Pol.* 33 (1), 163–183.
- Westbury, I., 2008. Making curricula: why do states make curricula, and how? In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. SAGE, Thousand Oaks, pp. 45–65.
- Young, M. (Ed.), 1971. *Knowledge and Control: New Directions for the Sociology of Education*. Collier-McMillan, New York.
- Young, M., 2015. Curriculum theory and the question of knowledge: a response to the six papers. *J. Curric. Stud.* 47 (6), 820–837. <https://doi.org/10.1080/00220272.2015.1101493>.

Innovative pedagogies in relation to curriculum

David Leat and Alison Whelan, Research Center for Learning and Teaching, School of Education, Communication and Language Science, Newcastle University, Newcastle, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	132
Convergent and divergent pedagogy and curriculum	132
Productive pedagogies	134
What might divergent pedagogy look like?	134
Philosophy for children	135
Singapore's Integrated Programme	135
Thinking through geography	135
Changing the learner	136
Rolling the stone uphill	137
Barriers for individual teachers	137
School level barriers	138
Discussion	138
Conclusion	139
References	140
Further reading	141

Introduction

The OECD (Cerna, 2014) has reported that in the early part of the current century there had been large increases in pedagogic practices across all countries in such areas as relating lessons to everyday life, high order cognitive skills, data and text interpretation. Paradoxically, however, there is a long tradition of concluding that schools are very resistant to change. The educational literature is full of accounts of and commentaries on pedagogic and curriculum innovation, large and small, reflecting a lack of change:

It can be said that by and large [schools] have not changed fundamentally over some decades.

Tangerud and Wallin (1986, p. 45, Sweden).

The most remarkable feature of the educational system is its capacity for continuity and stability in the face of efforts at change [We are confounded by] the inability of innovations to transform the schools.

House (1979, pp. 9 ± 10, USA).

When we speak of an innovative school, we mean one that tries one new thing after another without making any of them work.

Runkel (1984, p. 178, USA).

These evaluations might come from thirty plus years ago, but in the late 1990s conclusions had not changed substantially (Cuban, 1998; Spillane, 1999), nor in the new century (Elmore, 2004, Supovitz, 2008). Hallgarten et al. (2015) reach the same conclusion and attribute the atrophy to the constraint of high stakes testing systems. The paradox arises substantially because there is confusion about what constitutes innovation in pedagogy and curriculum. There is no shortage of mandated reform and there are frequent changes in the content of curricula, but the effects are often negligible.

Convergent and divergent pedagogy and curriculum

The main tenet of this chapter is that innovation in pedagogy is best understood through the degree to which the teacher pursues convergent or divergent outcomes. This distinction is well articulated through the work of Torrance and Pryor (1998) in relation to assessment, one of the three message systems of schooling identified by Bernstein (1971), the other two being curriculum and pedagogy. However, it is also reflected in the work of a range of other authors, including Sfard (1998) and Lingard (2005). Torrance & Pryor developed their framework through the observation of primary school classrooms in England focusing on formative

assessment practice. They summarized the two extremes that they encountered in [Table 1](#), while acknowledging that some practice occupied an intermediate position.

A convergent pedagogy proceeds from the intention that the learner will learn or master something that is pre-determined and thus not open to negotiation. This aligns most closely with curriculum as content model ([Kelly, 2009](#)) where the driving purpose is for students to learn and reproduce subject content, which is generally taught in subject specific lessons. The teacher is positioned in authority, assessing the degree to which the specific content has been mastered. There is likely to be frequent teacher talk episodes, many tasks will relate to comprehension, understanding and application of subject content (facts, concepts and ideas) and assessment is geared to checking performance or retention of subject content. It is highly significant that interaction is characterized by the I-R-F sequence ([Sinclair and Coulthard, 1975](#)). In this pattern, the teacher initiates (I), usually with a closed question, a student responds (R) and the teacher provides feedback (F) based on the degree to which the student has matched the expected, “correct” answer. [Sarason \(1990\)](#) summarized a large number of US studies and reported that teachers asked 45–150 questions per 30 min, but students asked less than two in return! The contrasting, divergent assessment and pedagogical style has significantly different characteristics starting from the intention which is to stimulate the development of the young person and assessment aimed at determining what has been learned, whatever that may be. Torrance & Pryor made the point that this mode reflected different understandings of learning and how to use assessment to support it. In this pedagogical mode tasks are more open, but it is in the resultant interaction that more critical differences occur as the teacher is provocative or exploratory in feedback, generating debate and discussion and inching toward reflection and metacognitive insight. Furthermore, the curriculum is to some extent malleable and can be shaped by the students’ response, and thus those students are not subservient to the curriculum. The I-R-F pattern is disrupted and there is space for the student to initiate on the basis of curiosity, which is too often neglected in curriculum formulations. In Torrance & Pryor’s work convergent pedagogy was the norm and evidence generally suggests that it is a universal pattern.

The significance of a divergent curriculum is indicated by the concept of *subjectification* in education ([Biesta, 2012](#)). Subjectification sits alongside qualification and socialization as one of the purposes of education, with qualification representing the process of accrediting the mastering of knowledge, skills and dispositions, and socialization representing the process of becoming a citizen through the adoption of cultures, traditions and acceptable behaviors. Subjectification, however, is the development of the unique individual human, growing into their talents, such as the arts or sports or humor or, indeed, empathy. Biesta attacks the lack of attention given to the purposes of education, and this concern is echoed by many others ([Alexander, 2010](#); [Au, 2012](#)). Furthermore, Biesta draws attention to the centrality of interaction in education as teachers seek to influence their students, but he problematizes this process in that interaction is not simple teaching action and a resultant learning “effect”, it is a “hermeneutic” process, because students interpret the teaching action, whether it is through teacher words or tasks set. There is no perfect match between education “input” and “output” as is sometimes assumed in the widespread use of objective-led teaching geared to high stakes testing. For Biesta education assumes that teachers are aiming to foster independence in adult life, fulfilling their human potential in leading “good” lives, and he pinpoints (p. 585) unpredictability in interaction and interpretation:

If teaching is going to have any impact on students, it is because of the fact that students interpret and make sense of what they are being taught, not because the teaching would simply flow into their minds and bodies. Yet the art of interpretation is a fundamentally open art which lodges an element of unpredictability at the very heart of education.

The threats brought by uncertainty underpin much of the debate and struggle in the domain of pedagogic and curriculum innovation. Increasingly education systems and their manifestation in curriculum crave predictable and controllable outcomes ([Ranson, 2003](#)). This is partly because political systems want schools and teachers to be accountable for the task and monies devolved to them ([Rizvi and Lingard, 2006](#)). More widely such accountability reflects an increasing managerialism in public life, in which numbers govern the conduct of public services through a series of targets and technologies to gather data to monitor those targets. The emergence of this dominant thinking in education reinforces convergent pedagogies, with set objectives and an educational mindset and system that is geared to sticking close to those objectives and banishing uncertainty. Many in public office experience

Table 1 Some characteristics of convergent and divergent assessment based on [Torrance and Pryor \(1998\)](#).

<i>Convergent assessment</i>	<i>Divergent assessment</i>
Assessment which starts from the aim to discover if the learner knows, understands or can do a predetermined thing. Some implications:	Assessment which starts from the aim to discover what the learner knows, understands or can do. Some implications:
a. Precise planning by the teacher focused on objectives;	a. Flexible or complex planning which incorporates alternatives;
b. Closed or pseudo-open teacher questioning;	b. Primarily open tasks with “helping” questions;
c. Feedback focused on performance and the successful completion of the task;	c. Feedback as part of discussion prompting reflection on the task and its context with a view to constructing understanding;
d. An interaction based on an Initiation-response-feedback (IRF) sequence;	d. On-going dialog between and among learners and teachers;
e. Assessment controlled mainly by the teacher.	e. Assessment as a collaboration between and among teachers and students.

this as an oppressive form of surveillance (Ball, 2003) which constrains their professional agency and creates a dissonance between their values and actions, which are often out of kilter (Fuller, 2019). Divergent pedagogy does not readily fit in such regimes.

There are interesting links between convergent and divergent pedagogies and Basil Bernstein's concept of framing which refers to "the degree of control teacher and pupil possess over the selection, organization, pacing and timing of the knowledge transmitted and received in the pedagogical relationship" (1996). Strong framing is found in classrooms where teachers maintain control over decisions made concerning the selection of content, the activities applied to them and the pace of lessons. That control is evident in the classroom, through fact that the teacher asks all the questions of consequence, with scarcely any space for student curiosity and questions. Teachers generally plan lessons to set objectives rather to stimulate debate and thought. Weak framing, however, offers some space for the student to determine subject matter and questions to be pursued and would be associated with inquiry-based approaches. Weak is not a pejorative term and it signals a more equal power relationship between teacher and student. The teacher has to be more contingent, responding to the impulses, questions and motivation of the students. Furthermore, it does not mean that subjects have no place, rather that at times they serve the curriculum and the curiosity of students, with teachers taking responsibility that there is coherent subject progression with no glaring gaps.

Productive pedagogies

A further valuable reference point in articulating the nature of pedagogic innovation, is provided by the Queensland School Reform Longitudinal Study (QSRLS) (Lingard, 2007). This research project was focused on analyzing and theorizing school pedagogy and tried to square a difficult circle of promoting positive identity construction and the learning of critical subject knowledge as a partial antidote to disadvantage. To put it another way they were trying to calibrate powerful theoretical constructs with powerful classroom practice. This is against a backdrop of increasing government intervention directly or indirectly affecting pedagogy, with the effect of leaching out attention to difference and contemporary global citizenship. Lingard (p. 50) quotes Bourdieu (1990) as follows:

If all pupils were given the technology of intellectual enquiry, and if in general they were given rational ways of working (such as the art of choosing between compulsory tasks and of spreading them over time), then an important way of reducing inequalities based on cultural inheritance would have been achieved.

The QSRLS finessed the argument that pedagogic and curriculum innovation are critical in achieving social justice outcomes. Consequently, drawing upon work on authentic instruction (Newmann et al., 1995) and wider literature QSRLS developed four dimensions of pedagogies of difference, with sub-categories, for use in classroom observation (Table 2).

What might divergent pedagogy look like?

Three examples, at contrasting scales, are given in this section of pedagogic innovation – Philosophy for (or with) Children (P4C), adopted across numerous jurisdictions, the Integrated Programme (IP) in Singapore, a nation state, and an informal and local network of geography teachers (Thinking Through Geography) in England collaborating around developing metacognitive practice. Whilst the aims and structure of these contrasting

Table 2 Dimensions and sub-categories of productive pedagogies.

<i>Dimensions</i>	<i>Exemplar sub-categories</i>
Intellectual quality	Higher-order thinking; Depth of knowledge; Substantive conversation; Meta-language
Connectedness	Connectedness to the world beyond the classroom Knowledge integration
Supportiveness	Social support Academic engagement Student self-regulation
Working with and valuing difference:	Cultural knowledge Active citizenship Group identities in learning communities Representation

innovations do reflect divergent pedagogy generally, it is important to note that in some contexts they may be distorted to more convergent formats and it is important to remember that high stakes, intrusive testing regimes can thin out pedagogies.

Philosophy for children

P4C is a program originally developed in the US by philosopher Matthew Lipman (Lipman et al., 1980) who was disconcerted by the evidence about the poor cognitive abilities of students leaving high school and entering college. P4C in some form is now used in over 60 countries. Teachers provide a stimulus for school students to excite their curiosity. The students are asked to voice questions prompted by the story or stimulus (which might be a photograph, recording or artifact). Through democratic processes a question is chosen for the class to pursue using protocols to promote listening and the principles of dialog. Haynes and Murris (2011) argue that there is an epistemological shift in the role of the teacher and a fundamental reframing of relationship between teacher and student. P4C does not induce an overnight transformation, as for both teachers and students there can be considerable period of adapting to new divergent norms. The Education Endowment Fund in England has funded a quasi-experimental study into the effects of P4C on maths and reading attainment in primary schools with pupils aged 8–10 years old, involving nearly 50 schools (Gorard et al., 2015). The results indicate a 2 month gain in English and mathematics test scores. Qualitative evidence from both teachers and pupils indicated that P4C had had a positive impact on pupils' confidence to speak, listening skills and self-esteem and thus they were more likely to initiate through their own questions. This evidence along with a number of other studies (Trickey and Topping, 2004) confirms the potency of P4C in breaking open the restraints of the routine talk pattern of I-R-F, in which teachers control talk, ask nearly all the questions in pursuit of right answers converging on the planned curriculum outcomes.

Singapore's Integrated Programme

While P4C is international in its status, the Integrated Programme (IP) is country specific, originating in Singapore and targeted at the top 10% of academic achievers. According to Koh et al. (2014) IP had radical intentions, with explicit mention of diverse experience. They also note that the IP represents a shift in the education system toward a student-centric phase (from 2011), having passed successively through survival-driven, efficiency-driven and ability-driven eras. The student-centric system pays greater attention to social-constructivist understandings of learning in which individual meaning making and creativity are valued. Introduced in 2004, the IP was a government initiative for higher attainers to experience an enriched curriculum surpassing a subject based approach to encourage critical, creative and reflective thinking as well as "leadership, teamwork and communication skills" (Ministry of Education, 2002). While the majority of the school population take "O" (ordinary) Level examinations at age 16, the IP students bypass this certification stage and only take the traditional "A" (advanced) Level qualification at age 18, thus freeing them from the restraints of more convergent "O" Level courses. The school based curriculum planners have had to work hard to adjust school structures to accommodate the IP modules alongside the traditional "O" Levels introducing foundation modules (Level 1), interest-driven modules reflecting staff specialties (Level 2) and modules for developing independent/inquiry learning (Level 3). It is highly relevant that the curriculum planners focus strongly on key concepts in structuring learning rather than more traditional content organization. There is a clear expectation of reshaping classroom practices.

Some teachers are teaching both the traditional "O" level focused course and the IP modules and they are clear about the difference in pedagogy, as reported in Koh et al. (p. 856):

The OP is very simple. You just teach, exam, any time of the day, they just go, no problem. The IP one, you must moderate a bit and put in more activity, so you make things a bit more fluid, a bit more interactive for them. (Teacher D)

Because after a while I realized, OK if I use this method, the IP students get a bit bored, but the OP are OK. If I use another method, the OP students don't get it and the IP students find it exciting. So after a while ... I started to vary the way I teach both classes. For IP, I really try and do more facilitation. That means I give them work, they do it, they present and then as a class, we critic and [it is] more self-directed. Whereas for the OP... I monitor them a bit more closely. So if I notice that they do not understand the particular concept, that's when I have to explain it." (Teacher E)

The IP is perhaps not an outright divergent pedagogy but these comments subtly demarcate the different teaching constructs that are being applied by teachers between the two student cohorts, with interactivity, facilitation, critique and responsibility being offered to the IP students. It is pertinent to reflect that it is only the brightest students that are being offered such opportunities with the sense that they are an elite who need a more expansive education for the long-term health of the Singaporean economy.

Thinking through geography

The divergent practice described by Leat and Lin (2003) focuses on an informal network of Geography teachers developing meta-cognitive outcomes for students and illustrates how teachers can create spaces (see below) for more innovative practice within the confines of a more convergent system. The development of metacognition is a critical element of divergent practice and this

emphasis relies on reflective discussion of thinking. It is interesting to note that the lessons were not part of a program or intervention. Through video analysis and student interviews Leat & Lin identified some key aspects of divergent practice. These included:

- Teachers giving care and attention to individuals and groups to the effect that they feel that their ideas are respected and a focus of interest of the teacher;
- Teachers encouraging students to ask questions;
- A strong press from the teachers for students to expand, explain or justify their answers, thus moving away from the IRF pattern as the teachers do not provide immediate feedback but, as in the Torrance and Pryor study, they provoke and stimulate;
- Perhaps more tellingly the teachers promoting and managing discussion

Overall, there was an emphasis on shared thinking, open discussion and making connections, with an explicit message to students that they are not just learning subject matter but that they are also developing their critical faculties and more complex identities. Furthermore, they are not being treated as a collection of individuals but as a learning collective. The students' interviews provide some insights into their unfolding conceptualizations of learning and learning environments stimulated by divergent dimensions to learning:

Female pupil: She was making us look at other peoples' work to understand, to see what they were writing about.

Female pupil: It's better to have someone else's opinion.

Teacher: Why? I don't think that you are wrong. I want to know why?

Female pupil: Because you can put your ideas and their ideas together and make it like better.

Teacher: Why is it important to talk about your ideas? If you and I sat down now and talked about our opinions or some of our ideas, why is that better than you deciding on your own?

Female pupil: You can find out better ideas than you had and make them better ... And make them more exciting and more serious.

There is no sense here of getting students to zero in on the right answer as quickly as possible, but instead that discussion is a place to share one's thoughts, listen to others, compare ideas, build understanding and be critical. Taken together and in their different ways, these three examples demonstrate an evident concern for developing outcomes that go far beyond the mastering and retention of subject content for examination purposes.

Changing the learner

Divergent pedagogy, as indicated by [Torrance and Pryor \(1998\)](#), has many dimensions that are different from convergent pedagogy, related to teacher constructs and classroom processes but there is every reason to believe that the effect on learners is different. This can be understood in a number of ways that are complementary to one another, as the psychological impact may differ from student to student. However, the central argument is that in divergent pedagogy the learner is interacting with their environment in more complex ways than being taught to master and remember tenets of subject disciplines. A first consideration is that divergent pedagogy is likely to be more "active", not necessarily in terms of making things or concrete actions, but active in the sense that the student initiates and is invited to think and respond. Indeed, in some divergent contexts such as project based, or inquiry-based learning there is explicit scope for action. Dewey's Transactional Theory of Knowing ([Biesta, 2014](#)) suggests that action is critical because it represents the learner acting in their environment, opening up the possibilities of what effect that action has. This is significant because the environment, whether human or natural, will respond to that action, creating some form of feedback, through which one can learn about the results of one's actions. This might be as varied as building something which does or does not fulfill one's goals or engaging in a discussion about an ethical issue that needs tackling. The experience of the action and the feedback constitutes learning from experience. Convergent pedagogy, because it is focused on a reproducing a right response or answer, cannot offer this mode of opportunity to learn. The importance of metacognition comes to the fore in action-oriented divergent pedagogies in that possible actions can be rehearsed in advance and their possible efficacy and consequences evaluated. [Dewey \(1922, p. 132\)](#) terms this intelligent action which he defines as "dramatic rehearsal (in imagination) of various competing possible lines of action". In the classroom environments described by Leat & Lin students are strongly encouraged to rehearse possible lines of action, including thought. Additionally, metacognitive reflection, after the event, permits not only recognition of missteps but also how the experience relates to wider patterns and concepts in subject disciplines.

For some students the experience of divergent learning is transformative ([Mezirow, 1991](#)). The learning experience changes the way the individual sees the world, they have new perspectives on their world and their place in it. It is, to a degree, a whole person change, involving knowledge, values and attitudes, feelings and behavior ([Illeris, 2014](#)). This change can be sudden or incremental ([Cranton, 2002](#)) but involves the individual coming to understand that their view of the world is skewed or limited, and as their perspective changes their assumptions and meaning making processes change. Divergent learning experiences stimulated by pedagogic innovation can have such effects of this magnitude on young people, particularly those who had limited life chances. These two students suggest such an effect after exposure to inquiry-based teaching in their school ([Leat and Reid, 2012, p. 203](#)):

it has made me think about what I can do and want to do when I leave, it has given me more confidence.

I didn't know learning was like this – it's really good because I know what I can do more – to get better at things

In these extracts identity development is evident. In the C21st identity has taken on new significance as the traditional social anchors of religion, class, work, binary gender roles and family have all lost their weight. We have more opportunity and freedoms to author ourselves through work, dress, networks, affiliation, activity and of course social media but it is a bewildering choice with an expectation that we can be anything we choose and that we will be a "somebody".

As Illeris observes (2014, p. 579):

For the generations which grew up during the 1980s and later, the task or duty of creating their own identities became more and more central. Who am I? Who do I want to be? How can I fulfil my dreams? For some, usually those with the more advantaged social backgrounds, possibilities might be great and never-ending. But for others the proliferation of choice became a strain, a continuous demonstration of their insufficient individual capacity to make things function and capture the success and happiness which should be the outcomes of all the choices and possibilities.

Divergent pedagogy, in which experience is planned for and valued, positions the student as an active agent in an environment in which they will have an effect and interact with others. At the more active spectrum of divergent pedagogy is inquiry and project-based learning, where, to use Bernstein's term, framing is weak and students can shape the curriculum in important ways. Students can take on meaningful roles in projects, they are not passive and they can be offered the chance to operate in more diverse social and physical environments beyond classrooms. They can go places, meet people and do and make things. The German psychologist Ziehe (2009) refers to this process as "search movements" where ideas, clothes, views and social roles might be tested out for fit. As competence is recognized by others, including teachers and peers, identity is refined and confirmed especially where there is a sense of collaboration or community (Lave and Wenger, 1991). As a consequence, participation through action is pivotal to identity development, but this stands in contrast to so many secondary schools where students play a needlessly passive role and where responsibility is largely confined to handing completed work to teachers.

It is especially important therefore that during formative years when they are soon to be faced by a precarious and challenging future, that young people should have the opportunity to visit a wide range of social environments and to meet as many people as possible. In addition, they should be challenged to work with these people and their peers and in these contexts, because these divergent experiences will provide them with a rich array of raw material with which to construct robust identities. It is also such processes that begin to establish particular competences or habits, such as creativity, criticality and empathy. Competences are more than skills, they are part of who we are and this comes from substantive experiences and challenges, not superficial classroom exercises which demand such a narrow slice of ourselves.

Rolling the stone uphill

If innovation in pedagogy can produce the beneficial outcomes associated with divergent pedagogies, the question arises as to the stubborn resistance to change as outlined at the start of this chapter. Pedagogic innovation can produce a reaction in schools that has similarities to the immune system in a human body responding to infection. The immune system produces antibodies which fight and (usually) overwhelm the infection. A further metaphor, rolling the stone uphill, captures the relentlessly recurring nature of failure of reform is found in Greek mythology. Sisyphus, the first king of Ephyra was punished by Zeus for his deceitfulness and arrogance, by being condemned to pushing a large boulder uphill, only for it to roll down again as he neared the top, with these events repeated eternally. There was no closure and the boulder never reached the summit. The failure in innovative pedagogy is systemic and by way of explanation one can point to some of the factors which operate at the individual teacher level but also some affecting schools as institutions, and indeed national and global discourse.

Barriers for individual teachers

Individual teachers are conditioned by a variety of socializing forces to teach in particular ways (Leat, 1999). Teachers have been students themselves, taught for thousands of hours which provides an apprenticeship of observation (Lortie, 1975) which creates a mental model of teaching. Once in training and induction there is a certain accepted model of teacherly activity relayed by teachers in the work environment and it is hard to step away from those shaping forces. This can be further reinforced by students themselves who have learned a particular role, usually fairly passive and cemented in the IRF questioning pattern. Many of these questions will be closed and pseudo-open questions in which the teachers stays firmly and deliberately in control. Teachers are generally taught to plan using objectives which demands that they predict what students will learn which becomes the bedrock of what and how they will teach and assess. Teachers are not taught or expected to teach in a manner where learning unfolds and students are given responsibility for some of the choices about manner, pace and direction of learning. A teacher experimenting with thinking skills approaches reported in Leat (1999, p. 391) describes the challenge of changing student expectations and roles:

Pupils were slow to get the hang of what was expected of them, and as a result the lesson quickly deteriorated into chaos The main lesson I have learned from my experiences ... is that they are not easy for pupils to do. The activities, although often apparently straightforward, usually require a certain way of thinking and depth of thought which pupils are simply not accustomed to. It is therefore necessary to 'train' children to learn in a different way, to teach them a variety of interchangeable skills, and to be more adaptable in their learning.

Another obstacle to be overcome is giving up classroom routines. Much of the process of learning to teach involves learning to make classrooms predictable as they are potentially chaotic and unruly in that a teacher is attempting to manage the learning of 25 students who have different needs and interests. It takes a big investment from teachers to establish and embed norms and routines for students, such that if the teacher gives particular instructions students will respond in predictable ways. [Brown and McIntyre \(1993\)](#) developed the concept of Normal Desirable States (NDS), which represents the criteria through which a teacher can judge whether a lesson, or part thereof, is proceeding along the expected lines toward the expected conclusions. If a lesson departs from the expected they will intervene to try to get it back to the NDS, which might be seen as a state of equilibrium. Curriculum innovation, not least in the shape of divergent outcomes, upsets the NDS and teachers and students have to learn a new normal.

School level barriers

Commonly schools do not make it easy for teachers to innovate toward more divergent pedagogies. [Payton and Williamson \(2008\)](#) summarizing the challenges faced by teachers trying to introduce *Enquiring Minds* (an enquiry model) in England detailed a range of very practical obstacles such as timetabling, resources and the dominance of subject teaching:

In none of the schools was there a concerted effort to ensure that the skills and aptitudes of enquiry-based learning, information management and so on were embedded across the routines of the school. While certain teachers did report that they had imported aspects of these initiatives into their subject teaching, these were individual examples and not endorsed systematically across any one participating institution.

In Spain, [Pozuelos et al. \(2010\)](#) in a study of the implementation of an enquiry curriculum project *Exploring Our World* (age 6–12) drew attention to very similar practical problems, but also to teachers' transactional conceptions of teaching. One of the interviewees provides a very trenchant analysis (p. 138):

Enquiry-based learning runs up against the whole fixed and immovable school machinery.

If curriculum innovation requires innovative schools then long-term prospects of change may be limited. [Giles and Hargreaves \(2006\)](#) identify three factors which undermine sustainability of such schools. The first is the additional funds that they may receive in start-up or conversion and their ability to attract key staff, which are rarely continued ad infinitum. The second factor is an alleged life span of innovative schools of creativity and experimentation, overreaching and entropy before they expire or lapse into a survival mode ([Fink, 2000](#)). Thirdly more innovative schools can be overtaken by national policies such as increasing centralized state control. Giles & Hargreaves note a tendency in some jurisdictions for large scale reform with a strong core of standardized testing and accountability to dislodge local innovation. There is a strong conservative default mode in the grammar of schooling, underpinned by governments aligning themselves to the global reform movement.

This chapter does not offer a particularly uplifting perspective, as pedagogy and curriculum remain fossilized in a traditional subject and teacher dominated pattern. Despite the recurring efforts of teachers, researchers and activists to reform there is a Groundhog Day quality to those efforts (Groundhog Day is a 1993 film in which a TV reporter becomes trapped in a time loop, forced to repeat the unsatisfactory events of the same day, February 2nd, over and over). This is no historical accident, although history plays a part. Rather it reflects the decisions that are made by those with political power, who either wish to maintain the status quo or who balk at the challenge of the scale of system change. Furthermore, that conservatism has been bolstered by the new managerial technology of testing and accountability, which believes in the power of markets to improve education, without ever seriously questioning what education might be for.

Discussion

It is instructive to see innovation as a form of social space, following sociological traditions. [Bourdieu \(1996\)](#) was a prime social space theorist through the development of field theory focusing on power for dominance. [Liu \(2021\)](#) argues that social spaces have "natural histories" and exhibit relationships with other social spaces and through this lens of relationships between spaces it is possible to see why pedagogic and curriculum innovation get squeezed. Liu proposes a typology of social spaces: kindred, symbiotic and oppositional. Kindred spaces are essentially similar to each other, while symbiotic spaces are dependent on one another and benefit from exchanges, whereas oppositional spaces are contrasting and potentially hostile to each other. The concept of

oppositional spaces is highly pertinent to innovation in the shadow of mainstream curriculum and pedagogy. For example, [Sannino \(2008\)](#) draws attention to the dominant activity in the context of an Italian school and how it suppresses a promising innovation supported by students from a local university. [Pozuelos et al. \(2010\)](#) and [Stanley et al. \(2012\)](#) provide further evidence that the dominant, convergent pedagogy and curriculum present a threat to innovation, in that the demands it places on educational infrastructure makes precious little allowance for innovation and its different demands. In addition, innovation can be seen as too risky in high stakes environments ([Greany and Waterhouse, 2016](#)). Where there is oppositional pressure there are a number of tactics that can ameliorate the threat such as hybridization of innovation so that it more closely resembles the normative ([Leat et al., 2014](#)), and temporal/spatial solutions whereby the innovation is given some preference at particular times of the year or with particular students (usually low attaining/low motivation).

Liu refers to two important types of actor in these areas of friction between different spaces—guardians and brokers. Guardians, in particular, would have an explicit awareness of the threats to innovation coming from the dominant, convergent school activity and can protect spaces. This is not to argue that convergent teaching is detrimental but that it is only part of a pedagogical repertoire. Guardians can also protect innovation spaces by articulating the benefits of a divergent approach and tracking both its quality of implementation and its impact on students, both areas that are ripe for professional learning focus, through approaches such as lesson study ([Vermunt et al., 2019](#)). It can also be achieved by putting protective borders around spaces where innovation can mature with a view to a longer-term curriculum development. It can be seen as a form of R&D. To feed the innovation process it is advisable to promote the agency of teachers to undertake curriculum development, creating the ecological conditions that support this process ([Priestley et al., 2015](#)).

The significance of the relationship between spaces, one divergent and other more convergent and test related, is signaled by the critique of the social realists ([Young and Muller, 2010](#)), who advocate the place of subjects in the curriculum as a form of cultural capital, without which disadvantaged students are locked out of opportunities for social mobility. While there are well grounded critiques of the approach ([Wrigley, 2017](#)) and a degree of parody in the depiction of more skills-based teaching, there is a point to the argument as exemplified in the QSRLS research, where much teaching even in well-regarded schools lacked intellectual rigor. Young and Muller have characterized three future scenarios for curriculum, Futures 1, 2 and 3, which are respectively traditional subject based curriculum maintaining power and privilege for elites, a progressive (but misguided) challenge to powerful elites with an emphasis on skills, and their preferred but somewhat ill-defined option (Future 3) a symbiosis between Future 1 and 2. They argue strongly for the place of subjects in that they are not arbitrary and represent important knowledge about the world created by communities of scholars over time, not least in mathematics and science, which have hierarchies of core concepts which need to be mastered sequentially.

Part of the value of the concept of Future 3 is the explicit attention given to boundaries. In this “Future” it is the variable relation between the two other “Future”s that is the condition for the creation and acquisition of new knowledge as the respective spaces of Future 1 and 2 become symbiotic spaces in which various forms of boundary crossing and exchange become productive. According to [Akkerman and Bakker \(2011, p. 2\)](#) “A boundary can be seen as a sociocultural difference leading to discontinuity in action or interaction” and the process of surmounting the barriers across spaces is increasingly seen as an important process as the territory is not familiar and there is a vital learning process involved in adapting. For example, [Phelan et al. \(1991\)](#) provide an account of the variable success of adolescents in moving between the world of their peers and that of their families. Thus learning to participate across a spectrum of sites is a measure of success and this applies both to young people ([Akkerman and van Eijck, 2013](#)) and to teachers, if they are to achieve fusion between deep subject knowledge and projects which demand that knowledge and put it to work.

Conclusion

A broadening of pedagogies and curriculum available to students is more likely to come from an acceptance that a range of educational outcomes, such as good subject knowledge, creativity, inclusion and mental and physical well-being, require a spectrum of pedagogies. Such a spectrum may require, among others, for older students something akin to good lecturing from teachers who are masters of explanation ([Wragg and Brown, 1993](#)), creative and performance opportunities and projects which draw on deep subject knowledge and address environmental and social issues in local and global contexts. This needs to be built around a meta-language of learning, shared by teachers and students and promoted by policy. For too long curriculum policy has been lost in an epistemological fog ([Leat et al., 2012](#)), inclined to load its resources into one pedagogical basket believing it to be the “best” or most effective, blind to the fact that the development of the whole person or human capability will need diverse opportunities and experience.

Divergent pedagogy is more likely to be embedded if teachers, schools and curriculum turn outwards and draw together partnerships to support curriculum in its widest sense. There are interesting examples in the US, such as the Community Funds of Knowledge model ([Gonzalez et al., 2005](#)) and the Ohio school improvement model ([Anderson-Butcher et al., 2008](#)). A very good example of what can be achieved comes from Australia ([Tytler et al., 2011](#)) in the Australian School Innovation in Science, Technology and Mathematics (ASISTM) initiative, involving partnerships between schools and community and industry organizations in developing local curriculum projects. These are aimed to overcome the failure of traditional science teaching to reflect the interdisciplinary nature of much science at the frontiers of social, ethical and personal application to real world issues. Projects were co-planned between teachers and associate partners from universities, public agencies or companies, but the best projects were reported to incorporate the perspectives of the partners. Projects were assembled, tested and refined with a changing cast of actors and

with a real purpose, both for improving student experience and providing some service to the community or partners. The authors describe some teachers as wanting to burst the artificial bubble inhabited by teachers and students doing school science and open the school to the community and vice-versa. There is this description of linked projects in one area (p. 25):

BioTech Units at Serendip Sanctuary: The BioTech project in this Geelong satellite area was devised in response to a felt need to strengthen the community engagement of youth in the local schools. The cluster of schools agreed to develop a four-year plan with sustainability as the theme whilst allowing diversity amongst the schools in their curriculum development within this agreed theme. Meg, the coordinator, managed a complex web of initiatives involving visiting scientists and other community members, environmental projects, excursion programs and project-based learning initiatives.

In the description above, Meg, the co-ordinator is acting as a broker across boundaries of different contexts. Subject knowledge is still critical, but the authors comment that ASISTM challenges “the manner and context in which it is developed”. This does imply that there is development needed in teachers’ in pedagogical content knowledge (Shulman, 1987) with a renewed focus on key concepts, if they are to cope with the unpredictability and contingency in pedagogical and curriculum innovation which allows more divergent outcomes.

Such an account places an onus on boundaries and boundary crossing (Akkerman and Bakker, 2011). Teachers need to be able to work with professionals from other settings and accommodate their voices, ideas and perspectives if they are to provide suitable contexts for divergent pedagogy and curriculum. A boundary represents a fault line between one practice or community and another and a divergent curriculum requires teachers who can negotiate those fault lines and draw together people, resources and ideas into a curriculum melting pot. In doing so they are also modeling a vital outcome for students, as increasingly the ability to cross boundaries, thereby displaying some chameleon qualities, is seen as a measure of a successful education for young people (Akkerman and van Eijck, 2013).

References

- Akkerman, S., Bakker, A., 2011. A conceptual review of literature on boundary crossing and boundary objects: new horizons in educational theory? *Rev. Educ. Res.* 81, 132–169.
- Akkerman, S., van Eijck, M., 2013. Re-theorising the student dialogically across and between boundaries of multiple communities. *Br. Educ. Res. J.* 39 (1), 60–72.
- Alexander, R.J. (Ed.), 2010. *Children, Their World, Their Education: Final Report and Recommendations of the Cambridge Primary Review*. Routledge, London.
- Anderson-Butcher, D., Lawson, H.A., Bean, J., Flaspohler, P., Boone, B., Kwiatkowski, A., 2008. Community collaboration to improve schools: introducing a new model from Ohio. *Child. Sch.* 30, 161–172.
- Au, W., 2012. What curriculum could be: Utopian dreams amidst a Dystopian reality. *Kappa Delta Pi Rec.* 48 (2), 55–58.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Bernstein, B., 1971. On the classification and framing of educational knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control*. Collier-Macmillan, London, pp. 47–69.
- Bernstein, B., 1996. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Taylor and Francis, London.
- Biesta, G., 2012. Philosophy of education for the public good: five challenges and an agenda. *Educ. Philos. Theor.* 44 (6), 581–593.
- Biesta, G., 2014. Pragmatising the curriculum: Bringing knowledge back into the curriculum conversation, but via pragmatism. *Curriculum Journal*, 25 (1), pp. 29–49.
- Bourdieu, P., 1990. Principles for reflecting on the curriculum. *Curric. J.* 1 (3), 307–314.
- Bourdieu, P., 1996. *The Rules of Art: Genesis and Structure of the Literary Field* (S. Emanuel, Trans.). Stanford University Press, Stanford.
- Brown, S., McIntyre, D., 1993. *Making Sense of Teaching*. Open University Press, Buckingham.
- Cerna, L., 2014. Innovation, Governance and Reform in Education. CERl (Centre for Educational Research and Innovation). Conference Background Paper, November 3–5, 2014. www.oecd.org/edu/ceri/CERI%20Conference%20Background%20Paper_formatted.pdf. (Accessed 14 June 2021).
- Cranton, P., 2002. Teaching for transformation. *N. Dir. Adult Cont. Educ.* 93, 63–71.
- Cuban, L., 1998. How schools change reforms: redefining reform success and failure. *Teach. Coll. Rec.* 99 (3), 453–477.
- Dewey, J., 1922. Human nature and conduct. In: Boydston, J.A. (Ed.), *The Middle Works (1899–1924)*, vol. 14. Southern Illinois University Press, Carbondale and Edwardsville.
- Dweck, C., 1999. *Self-Theories: Their Role in Motivation, Personality and Development*. Psychology Press.
- Elmore, R.F., 2004. *School Reform From the Inside Out: Policy, Practice, Performance*. Harvard Education Press, Cambridge, MA.
- Fink, D., 2000. *Good Schools/Real Schools: Why School Reform Doesn't Last*. Teachers College Press, New York.
- Fuller, K., 2019. “That would be my red line”: an analysis of headteachers’ resistance of neoliberal education reforms. *Educ. Rev.* 71 (1), 31–50.
- Giles, C., Hargreaves, A., 2006. The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educ. Adm. Q.* 42 (1), 124–156.
- Gonzalez, N., Moll, L.C., Amanti, C. (Eds.), 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Gorard, S., Siddiqui, N., See, B.H., 2015. *Philosophy for Children, Evaluation Report and Executive Summary*. Education Endowment Fund. Downloaded from: EEF_Project_Report_PhilosophyForChildren.pdf. ioe.ac.uk. (Accessed June 2022).
- Greany, T., Waterhouse, J., 2016. Rebels against the system: leadership agency and curriculum innovation in the context of school autonomy and accountability in England. *Int. J. Educ. Manag.* 30 (7), 1188–1206.
- Hallgarten, J., Hannon, V., Beresford, T., 2015. *Creative Public Leadership: How School System Leaders Can Create the Conditions for System-Wide Innovation*. WISE, Dubai.
- Haynes, J., Murris, K., 2011. The provocation of an epistemological shift in teacher education through philosophy with children. *J. Philos. Educ.* 45 (2), 285–303.
- House, E., 1979. Technology versus craft: a ten year perspective on innovation. *J. Curric. Stud.* 11, 11–15.
- Illeris, K., 2014. Transformative learning re-defined: as changes in elements of the identity. *Int. J. Lifelong Educ.* 33 (5), 573–586.
- Kelly, A.V., 2009. *The Curriculum; Theory and Practice*, sixth ed. Sage, London.
- Koh, E., Ponnusamy, L.D., Tan, S.T., Lee, S.-S., Ramos, M.E., 2014. A Singapore case study of curriculum innovation in the twenty-first century: demands, tensions and deliberations. *Asia-Pac. Educ. Res.* 23, 851–860.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Leat, D., 1999. Rolling the stone uphill: teacher development and the implementation of thinking skills programmes. *Oxf. Rev. Educ.* 25, 387–403.

- Leat, D., Lin, M., 2003. Developing a pedagogy of metacognition and transfer: some signposts for the generation and use of knowledge and the creation of research partnerships. *Br. Educ. Res. J.* 29 (3), 383–415.
- Leat, D., Reid, A., 2012. Exploring the role of student researchers in the process of curriculum development. *Curric. J.* 23 (2), 189–205.
- Leat, D., Thomas, U., Reid, A., 2012. The epistemological fog in realising learning to learn in European curriculum reform. *Eur. Educ. Res. J.* 11 (3), 400–413.
- Leat, D., Thomas, U., Reid, A., 2014. Reframing relationships between teachers, students and curriculum – the phenomenon of “hybridisation” in IBL. In: Blessinger, P., Carfora, J. (Eds.), *Inquiry Based Learning for Faculty and Institutional Development*. Emerald, Bingley.
- Lingard, B., 2005. Socially just pedagogies in changing times. *Int. Stud. Sociol. Educ.* 15 (2), 165–186.
- Lingard, B., 2007. Pedagogies of indifference. *Int. J. Incl. Educ.* 11 (3), 245–266.
- Lipman, M., Sharp, A., Oscanyan, E., 1980. *Philosophy in the Classroom*, second ed. Temple University Press, Philadelphia.
- Liu, S., 2021. Between social spaces. *Eur. J. Soc. Theor.* 24 (1), 123–139.
- Lortie, D., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago.
- Mezirow, J., 1991. *Transformative Dimensions of Adult Learning*. Jossey-Bass, San Francisco, CA.
- Ministry of Education, 2002. Performance Indicators, Sustained Achievement Awards, Best Practice Awards and National Education Awards for Secondary Schools. http://www.moe.gov.sg/media/press/2002/pr15082002_print.htm. (Accessed 12 July 2021).
- Newmann, F., Secada, W., Wehlage, G., 1995. *A Guide to Authentic Instruction and Assessment: Vision, Standards and Scoring*. ASCD, Alexandria, VA.
- Payton, S., Williamson, B., 2008. *Enquiring Minds—Innovative Approaches to Curriculum Reform*. Future Lab, Bristol.
- Phelan, P., Davidson, A.L., Cao, H.T., 1991. Students' multiple worlds: negotiating the boundaries of family, peer, and school cultures. *Anthropol. Educ. Q.* 22 (3), 224–250.
- Pozuelos, F., Gonzalez, G., Canal de Leon, P., 2010. Enquiry-based teaching: teachers' conceptions, impediments and support. *Teach. Educ.* 21 (2), 131–142.
- Priestley, M., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Academic, London.
- Ranson, S., 2003. Public accountability in the age of neo-liberal governance. *J. Educ. Pol.* 18 (5), 459–480.
- Rizvi, F., Lingard, B., 2006. Globalization and the changing nature of the OECD's educational work. In: Lauder, H., Brown, P., Dillabough, J., Halsey, H. (Eds.), *Education, Globalization & Social Change*. Oxford University Press, Oxford, UK, pp. 247–260.
- Runkel, P., 1984. Maintaining diversity in schools. In: Hopkins, D., Wideen, M. (Eds.), *Alternative Perspectives on School Improvement*. Falmer Press, Lewes.
- Sannino, A., 2008. Sustaining a non-dominant activity in school. *J. Educ. Change* 9, 329–338.
- Sarason, S.B., 1990. *The Predictable Failure of Educational Reform*. Jossey-Bass, Oxford, UK.
- Sfard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27 (2), 4–13.
- Shulman, L.S., 1987. Knowledge and teaching: foundations of the new reforms. *Harv. Educ. Rev.* 57 (1), 1–22.
- Sinclair, J., Coulthard, M., 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford University Press, Oxford.
- Spillane, J., 1999. External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *J. Curric. Stud.* 31 (2), 143–175.
- Stanley, G., Jones, M., Murphy, J., 2012. Implementing the opening minds curriculum in a secondary school in England: an alternative to the one-size-fits-all national curriculum? *Curric. J.* 23 (3), 265–282.
- Supovitz, J.A., 2008. Implementation as iterative refraction. In: Supovitz, J.A., Weinbaum, E.H. (Eds.), *The Implementation Gap: Understanding Reform in High Schools*. Teachers College Press, New York, pp. 151–172.
- Tangerud, H., Wallin, E., 1986. Values and contextual factors in school improvement. *J. Curric. Stud.* 18, 45–61.
- Torrance, H., Pryor, J., 1998. *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Open University Press, Buckingham.
- Trickey, S., Topping, K., 2004. Philosophy for children: a systematic review. *Res. Pap. Educ.* 19 (3), 365–380.
- Tytler, R., Symington, D., Smith, C., 2011. A curriculum innovation framework for science, technology and mathematics. *Educ. Res. Sci. Educ.* 41, 19–38.
- Vermunt, J., Vrikk, M., van Halem, N., Warwick, P., Mercer, N., 2019. The impact of Lesson Study professional development on the quality of teacher learning. *Teach. Teach. Educ.* 81, 61–73.
- Wragg, E., Brown, G., 1993. *Explaining*. Routledge, London.
- Wrigley, T., 2017. Canonical knowledge and common culture: in search of curricular justice. *Eur. J. Curric. Stud.* 4 (1), 536–555.
- Young, M., Muller, J., 2010. Three educational scenarios for the future: lessons from the sociology of knowledge. *Eur. J. Educ.* 45 (1), 11–27.
- Ziehe, T., 2009. “Normal learning problems” in youth: in the context of underlying cultural convictions. In: Illeris, K. (Ed.), *Contemporary Theories of Learning*. Routledge, London, pp. 184–189.

Further reading

- Au, W., 2008. Devising inequality: a Bernsteinian analysis of high-stakes testing and social reproduction in education. *Br. J. Sociol. Educ.* 29 (6), 639–651.
- Clayton, R., 2017. *Building Innovation Ecosystems in Education to Reinvent School A Study of Innovation & System Change in the USA*. Fellowship Report for the Winston Churchill Memorial Trust.
- Elmore, R.F., 2016. “Getting to scale...” it seemed like a good idea at the time. *J. Educ. Change* 17, 529–537.
- Facer, K., Buchczyk, M., 2019. Understanding learning cities as discursive, material and affective infrastructures. *Oxf. Rev. Educ.* 45 (2), 168–187.
- McKernan, J., 2008. *Curriculum and Imagination: Process Theory, Pedagogy and Action Research*. Routledge, Abingdon.
- Nussbaum, M., 2011. *Creating Capabilities: The Human Development Approach*. The Belknap Press of Harvard University Press, Cambridge, Massachusetts.
- Priestley, M., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Academic, London.
- Rata, E., 2012. The politics of knowledge in education. *Br. Educ. Res. J.* 38 (1), 103–124.
- Rata, E., 2019. Knowledge-rich teaching: a model of curriculum design coherence. *Br. Educ. Res. J.* 45 (4), 681–697.
- Young, M., 2014. What is a curriculum and what can it do? *Curric. J.* 25 (1), 7–13.

Teachers mediating of curriculum making processes

Maria Birbili, School of Early Childhood Education, Aristotle University of Thessaloniki, Thessaloniki, Greece

© 2023 Elsevier Ltd. All rights reserved.

Introduction	142
From curriculum implementation to curriculum making	142
Curriculum making revisited	143
Compliance, resistance, and creative mediation	143
How do teachers mediate the curriculum?	144
Teachers as mediators of curriculum: the role of teacher education and professional development	146
Concluding remarks	146
References	146
Further reading	148

Introduction

Teachers' relationship with curriculum is central to the implementation of educational policy and changing classroom practice therefore the words we choose to describe the actions, or the positions teachers take in the curriculum development process matter (Gerrard and Farrell, 2014). Do they execute the curriculum? Do they deliver it? Do they resist it? Do they implement it, or do they enact it? The words we choose matter because they reveal how we construct teachers as professionals and what we perceive, or expect, their role to be concerning curriculum policy directives (Gerrard and Farrell, 2014). They also have an impact on how we support teachers when new curricula, ideas, and practices are introduced.

Curriculum research has long shown that teachers do not passively implement externally determined innovations in curriculum and pedagogy as policymakers intended—they translate, interpret, and mediate them while in some cases they ignore, resist, or dismiss them (Hoyle and Wallace, 2005: 158; Priestley and Philippou, 2018). However, recent discourses and trends influencing global education policy, like for example that of investing schools and teachers with authority for curricular and instructional decisions (OECD, 2020; Priestley et al., 2016), have led to a renewed emphasis on understanding what it means, beyond the rhetorical level, to move the teacher to the center of curriculum decisions at the school level and how the intersection of curriculum and teachers everyday pedagogical and curricular decisions works in practice (Gerrard and Farrell, 2014: 639; Priestley et al., 2016: 188). Within this context, the concept of teacher mediation is assuming a critical role in understanding how teachers make sense of curriculum and negotiate both their positions and the positions of their students within it.

This article reviews the international literature on teacher mediation of curriculum making. We begin by considering how the teacher-curriculum relationship has changed and developed to recognize that teachers actively negotiate their positions in relation to policy directives. We briefly consider key concepts such as "curriculum making" and "teachers as curriculum makers" (these are discussed in greater depth elsewhere in the encyclopedia) before we take a closer look at teacher mediation of curriculum. We then synthesize the literature to provide an account of teacher mediation of curriculum making as a complex and situated agentic experience and practice. The review features studies from different national educational and cultural contexts to illustrate the connections among variables and how global trends create continuities across different policy contexts. We end the article with thoughts on how teacher education and professional development can support both individual and collective teacher sensemaking of curriculum.

From curriculum implementation to curriculum making

For many years the teacher-curriculum encounter (Ben Peretz, 1990) was discussed and researched from perspectives that conceived teachers and curriculum as two separate entities. The dominant image in the curriculum reform literature was that of a passive teacher who receives and implements curricula designed by experts and meant to be followed as "manuals" or "cookbooks" (Craig, 2020: 13). This top-down approach to curriculum aligned with a view of knowledge as "a set of accurate representations of an independent reality" (Rosiek and Clandinin, 2016: 296), a theory that separated the conceptualization and design of curricula from the process of their implementation in classrooms (Giroux, 2011) and models of teaching, embedded, as Parker (1987: 7) puts it, "in the particular value that teachers ought to be competent implementers of others' plans" (see for example behavioristic and cognitive models). As expected, this technical and linear conceptualization of the curriculum development process influenced teacher education and professional development which for several years focused on how to make teachers' behaviors and thinking "more conducive" to implementing externally mandated change (Giroux, 2011; Parker, 1987: 6).

While the phrase "curriculum making" is more than a century old (Catling, 2015: 134), the concept has gained more attention with Clandinin and Connelly's article in the 1992 Handbook of Research on Curriculum (Rosiek and Clandinin, 2016: 299). Based

on their exploration of how teachers find meaning in their teaching experiences, Clandinin and Connelly argued that classroom teachers do not merely implement curriculum as dictated by policymakers but actively experience and make curriculum from their personal and professional knowledge, and they do so every day alongside their students (Ciuffetelli Parker et al., 2011: 4; Craig, 2011: 21). Stated differently, their research indicated that curriculum is made “through the intertwining of the teacher’s life course of action with the students’ life courses of action” (Ciuffetelli Parker et al., 2011: 10) and not through a top-down implementation of policy and prescriptive teacher-proof curricula (Taylor, 2013).

Teacher as curriculum maker—as envisioned by Clandinin and Connelly—is an image that “acknowledges the teacher as a holder, user, and producer of knowledge, a self-directed individual,” an active agent, who negotiates curriculum in relationship with students “unhampered by, though not oblivious to, others’ mandates and desires” (Craig, 2010: 867). Thus, teachers as curriculum makers, in contrast to teachers as curriculum implementers, “not only ... act on their own initiative, [but they are] expected to do so” (Dewey, 1904 in Craig, 2020: 13). That makes curriculum making a deliberative process of interpretation, decision-making, and responsibility and positions both teacher and students as active agents in the negotiation of the curriculum (Deng, 2017; Doyle, 1992: 69; Restad, 2021). As teachers interpret and transform the mandated curriculum into learning experiences, they take into consideration not only their personal and professional knowledge and their students but also the subject matter and the milieu (Schwab, 1983).

Clandinin and Connelly’s work on how teachers’ and students’ lives interact with one another, and curriculum also gave prominence to the term “curriculum of lives”—“curriculum centered in the experiences of many individuals living in relation” (Ciuffetelli Parker et al., 2011: 10). The term captured not only the importance of lived experience but also the relationality of curriculum-making processes. Finally, their work also highlighted the value of studying teachers’ curriculum experiences from a narrative perspective as it allows the unfolding of experience, “as lived,” in the contexts it takes place (Caine et al., 2013). For Connelly and Clandinin (1990), life’s narratives are the context for meaning making of school experiences.

Curriculum making revisited

While the term *curriculum making* is now widely accepted, over the years it has been ascribed different meanings by different authors, many of whom fail to take into account its complexity as a practice (Priestley and Philippou, 2018). As curriculum research continues to explore the “on the ground” (Apple, 2013) realities of schooling and curricula, the development of more holistic and dynamic understandings of curriculum making is important if we want, as Priestley and Philippou (2018: 156–157) argue, to develop “practices that are more sensitive to broader pedagogical purposes and educational principles.”

Approaching school curriculum as “a complex system involving teachers, students, curricular content, social settings, and all manner of impinging matters ranging from the local to the international” (Connelly, 2013: ix), Priestley and Philippou (2018) and Priestley et al. (2021) define curriculum making as a social practice that occurs in a complex web of enactment, as teachers interpret, mediate, translate, negotiate and enact curriculum directives alongside their students and other social actors. The authors use the metaphor of the spider’s web to draw attention to the complexity, the uniqueness and the unpredictability of curriculum making. As a result, it produces unique social practices (Priestley and Philippou, 2018: 154).

A key concept in Priestley et al.’s (2021) approach to curriculum-making as a social practice is that it occurs in multiple “sites of activity.” In contrast with linear and hierarchical conceptions of curriculum making as happening in levels, the idea of “sites of activity,” suggests that curriculum making is made (and re-made) in different contexts, each one with its own kind of engagement and actions in curriculum making (Alvunger et al., 2021: 274). These sites interact and intersect in often unpredictable and context-specific ways (Priestley and Philippou, 2018: 154).

Priestley et al. (2021) identify five sites of activity that are connected in multiple, and non-linear ways, allowing actors to move between sites and influence curriculum making in different ways and multiple directions. These are the supra, the macro, the meso, the micro and the nano “sites of activities,” where activities range from transnational curriculum discourse generation from actors such as the OECD, the World Bank and the European Union (at the Supra layer), to curriculum making in classrooms and other learning spaces as teachers and students interact (at the nano layer).

This systemic way of thinking about curriculum making allows us to understand not only how activity happening in one part of the system influences other parts of it but also that curriculum is always in a state of becoming (Hewitt, 2006: 39; Lingard, 2021: 31): it is constructed and re-constructed as different actors and groups in different sites of activities interpret the curriculum using different discourses to make sense of it (Woods and Wenham, 1995), accommodate its requirements through a variety of mediations (Osborn et al., 2000: 15) and respond to local conditions. A systemic mode of thinking also draws our attention to the fact that as systems change, individuals and groups working within them, also change (Alvunger et al., 2021: 273). That makes Alvunger et al. (2021: 273) argue that “curriculum making is not just a matter of creating the curriculum as product, but a powerful process of learning for those involved as well.”

Compliance, resistance, and creative mediation

The understanding that teachers do not mechanically implement curricula that someone else has written for them and that “any curriculum is a set of meanings yet to be situated and made sensible” by different actors (Olson, 1980 in Parker, 1987) led to a critical question: “How do teachers come between students, curriculum, and milieu, and how does that coming-between, that agency or mediation, shape practice?” (Parker, 1987: 7).

This question has guided a series of studies that took a critical approach to arguments that positioned teachers either as submissive technicians or as contesters of policy and aimed to understand teachers' responses to policy initiatives and directives "as a series of mediations" (Osborn et al., 2000: 12). They did so from the perspective that individuals do not establish a direct relationship with the social-material world (Lantolf and Thorne, 2007). A focus on this direction aimed, according to Parker (1987) to provide

a more accurate and less constraining portrayals of teachers and schooling -more accurate because they do not exclude teachers' tacit and practical knowledge of schooling, and less constraining because they do not presuppose teachers as objects in a curriculum production system (p. 2).

Osborn et al. (2000: 231, 234) conceptualized teacher mediation of curriculum as "the product of successive interpretations and reinterpretations" which are embedded in a dynamic network of personal identity, values, beliefs, understandings, interactions, and structural constraints. In their study of teachers' and pupils' experiences of the 1988 Education Reform Act in England, they identified varied strategic teacher responses to curriculum, which ranged from compliance to imposed changes, through incorporation of imposed changes into existing modes of working, creative mediation, submission to the imposed changes without any fundamental change in values and professional ideology (retreatism) and resistance to the imposed changes. As the same researchers argue, these responses should not be seen "as static or unshifting typologies of teachers":

Rather, they represent role configurations or strategies which individual teachers might adopt and move through at different times and in different ways. Teacher strategies may have varying features at different moments and in response to differing situations. Thus, teachers may sometimes adopt the position of creative mediators and sometimes that of retreatists or resisters (p. 67).

In other words, teachers' mediation of policy and curriculum is not confined to one response or another but moves among strategies (Giroux, 1985; Parker, 1987). The decisions teachers make are more complex than a choice between compliance and resistance as they are based on and are influenced by a variety of variables: personal factors and the social and structural conditions of teachers' work (Coburn, 2005; Maniates, 2010).

Osborn et al.'s (2000) term "creative mediation" has probably been the most influential in the related literature and has been used to capture teachers' creative responses to or creative ways of working within externally imposed guidelines. As Nias (2000: xiv) points out, creative mediators—teachers "who feel able to take active control of educational changes and to respond to them in a creative way" despite difficult and/or uncertain conditions (Osborn et al., 2000: 81)—can be defined as curriculum makers rather than simply policy implementers. Osborn et al. (2000) have also identified four types of teacher response that fall under creative mediation. These are protective mediation, innovative mediation, collaborative mediation, and conspiratorial mediation and all of them indicate that some teachers seek creative ways of working within the mandated curriculum to develop practices in accordance with their values and working situations.

As evident in many studies from different national and educational contexts, curriculum interpretation and mediation by teachers is a continuous process that takes place even under the most prescriptive policies and teacher-proof curricula (Priestley et al., 2021: 6). Teachers function as selective agents (Ben Peretz, 1990: 114) and quickly find ways to mediate the curriculum for their contexts and negotiate their position within it based on their experiences, skepticisms and critiques (Ball et al., 2012: 15; Wallace and Priestley, 2011: 360). From this perspective, they are and will always be curriculum agents and curriculum makers (Alvunger et al., 2021: 284).

How do teachers mediate the curriculum?

When we position the teacher as a mediator and a maker of curriculum—and therefore a person with agency—it becomes clear that teachers' biographies and identities will always be part of what is happening in the classroom (Rosiek and Clandinin, 2016). While for years, several scholars and researchers have been critical of the over-emphasis on the role of the teacher in curriculum-making and the neglect of other important factors, today is widely acknowledged that the individual is still a critical variable in successful curriculum-making, as contextual influences are mediated through teachers' personal resources (Hizli Alkan and Priestley, 2019: 15; Rosiek and Clandinin, 2016; Spillane, 1999).

As mentioned earlier teacher mediation is about how teachers make sense of the curriculum (and their work) and how their understandings and interpretations shape practice. The realization that teachers interpret, mediate and adapt policy, rather than just mechanically implement it, drew researchers' interest in the process of teacher sensemaking and the factors that influence it (Ganon-Shilon and Schechter, 2017; März and Kelchtermans, 2013). A sensemaking perspective allows researchers to understand not only how and why teachers make certain decisions but also why the same policy messages generate such varied interpretations and responses among curriculum agents (Seashore Louis and Robinson, 2012; Wallace and Priestley, 2011: 378).

Sense-making is widely defined as an (inter) active and dynamic process of interpretation and constructing a reality that draws on the sense-maker's pre-existing knowledge, practices, experiences, beliefs, attitudes, emotions, and values (Ganon-Shilon and

Schechter, 2017; Ketelaar et al., 2012; März and Kelchtermans, 2013). As Ketelaar et al. (2012: 274) put it, “sensemaking is more than simply interpreting a message. It is an active cognitive and emotional process in which a person attempts to fit the new information into existing knowledge and beliefs.” For example, teachers’ sense of being powerful or powerless in reform efforts may affect how they respond to them (Wallace and Priestley, 2011).

In the context of curricular decisions and practice, sensemaking is the result of a complex and ongoing interaction between policy content, the relevant prior understandings and experiences of teachers, and the social and structural conditions in which sense-making occurs (Coburn, 2005; Craig, 2020: 125; Wallace and Priestley, 2011: 361). In other words, teachers come to understand policy ideas and curricula by “placing new information into pre-existing cognitive frameworks,” also called “worldviews” or “working knowledge” or “cognitive frames” or “cognitive maps” (Coburn, 2005: 478). As teachers’ personal resources interact with external factors, they mediate, in turn, teachers’ curricular decisions and practices. So, to borrow Spillane’s (1999: 169) argument, by virtue of who they are, teachers can construct very different worlds for learning in their classrooms.

Another influential concept in the literature studying teacher mediation of curriculum making processes is personal practical knowledge (Clandinin, 1985), a term which was coined to indicate that “curriculum is lived and experienced” and to acknowledge teachers “as both knowing and knowledgeable” (Craig and Ross, 2008: 291; Ross and Chan, 2016: 4). Personal practical knowledge is gained through personal and professional experiences and differs from teacher knowledge that is gained from expert sources such as professional documents and support (Clandinin and Connelly, 1996; Ross and Chan, 2016). It is unique to the circumstances and contexts of each teacher as it grows out of their experience and as teachers construct knowledge through their interactions with students, colleagues, parents, and others within and beyond their classroom and school settings (Ross and Chan, 2016: 2). Connelly and Clandinin (1988) understand personal practical knowledge as experiential, emotional, embodied, relational, contextual, and as shaped and reshaped in present situations by past experiences and future plans and actions (Rosiek and Clandinin, 2016: 299; Ross and Chan, 2016: 5). As Ross and Chan note (2016: 5), approaching teacher knowledge from this point of view “enables scholars to consider more carefully the dialectic between each teachers’ personal knowledge and his/her social contexts as knowledge shaped and lived out.” As a result of this “dialectic,” teachers’ personal practical knowledge develops throughout their professional lives (Swart et al., 2018). For Rosiek and Clandinin (2016: 301) personal practical knowledge constitutes “the most important resource in understanding the experienced curricula in classrooms.”

Studying teachers “as biographies with intertwining personal and professional lives” (Kontovourki et al., 2018: 259), research from different social and cultural contexts provides valuable insight into the combination of individual factors that interact to influence teachers’ sense-making and eventually curriculum-making. These include teachers’ histories of where they had taught previously and their memories of past teaching experience (Rosiek and Clandinin, 2016; Willis et al., 2019); previous career experience (Goodson, 2004); teachers’ own experiences as students (from school through university) and the resulting images about what schools should look and feel like (Ross and Chan, 2016); and their socialization as teachers from initial teacher education (Ball, 2014). For example, Kontovourki et al. (2018) in Cyprus, identified a strong interplay between a teacher’s positive previous work experience as “an exemplar classroom teacher” and her identity as a professional who can make curriculum making decisions.

In making sense of curriculum, teachers also draw upon their feelings about their job (Hizli Alkan and Priestley, 2018); the kind of relationship they have with their students and how they conceptualize student needs (Datnow and Castellano, 2000); and their representations of themselves as teachers (Kelchtermans, 2009). Other individual factors that influence how teachers interpret and engage with curriculum are the way they conceptualize curriculum and curriculum making (Hizli Alkan and Priestley, 2018); their dominant mode of reflexivity (Hizli Alkan and Priestley, 2018, 2019); the nature of their connection to policy directives (Kirk and MacDonald, 2001); their sense of familiarity with the changes introduced (Hizli Alkan and Priestley, 2018); their professional knowledge and language (Biesta et al., 2015); teacher subject knowledge and expertise and their subject-based professional identity (Mitchell, 2016). For instance, the less secure teachers feel in their subject knowledge the less capable and willing they feel to engage in curriculum making (Pountney, 2021). Or, as Biesta et al. (2015) found in a study in Scotland, teacher professional language, expressed, for example, in terms such as “bright” or “gifted” children, may guide teacher mediation of curriculum in particular ways.

While teacher individual cognitive processes, biography, and knowledge are important curriculum resources (Rosiek and Clandinin, 2016), sense-making is embedded and situated in social interactions and contexts so from this perspective, teacher sensemaking about curriculum is not solely an individual matter (Coburn, 2005: 478). It is also influenced by the collective sense-making that takes place in formal networks and informal interactions and alliances among teachers (for example the conversations teachers have with their colleagues about policy messages) (Coburn, 2001: 145). In this context, factors such as school culture (März and Kelchtermans, 2013), patterns of social or professional interaction with colleagues (Birbili and Myrovali, 2019; Hizli Alkan, 2020), or conflicting interpretations of policy and curriculum among teachers and school leaders can play a critical role in teacher mediation of curriculum making (Gouëdard et al., 2020).

The increasing policy interest in teachers’ professional communities and interactions (OECD, 2020), as important social sites for meaning making, has led to a growing body of research on the nature, the extent and the dynamics of teachers’ formal and informal connections and networks with the purpose of understanding the role they play or can play in helping teachers to manage the complexity of curriculum making (Coburn, 2001; Hizli Alkan, 2020; Hizli Alkan and Priestley, 2019). For example, how connections or interactions may mediate structural pressures (Coburn, 2001).

Studies also indicate that teacher mediation of curriculum making is integrally bound with the structural, the cultural and the material reality teachers operate in (Alvunger et al., 2021; Hizli Alkan and Priestley, 2018; Maguire et al., 2015; März and

Kelchtermans, 2013). Possible enablers and/or constraints that mediate curriculum-making practices include the size of institutions and subject departments (März and Kelchtermans, 2013); the quantity and the quality of physical resources (Erss et al., 2016); the available support and guidance on curriculum making from authorities, bodies and other actors (Kontovourki et al., 2018); the available curriculum material (Willis et al., 2019); the styles of school management and leadership (Vähäsantanen, 2015); and the “grammar of schooling” (e.g., timetabling, school protocols) (Tyack and Cuban, 1995; Tan, 2016). So, for example, the head teacher’s leadership style and skills may strengthen teachers’ experience in curriculum making and thus work as an enabler (Hizli Alkan and Priestley, 2019: 9). On the other side, a management culture with strong control over teachers’ work and the curriculum may narrow teachers’ agency, as was the case in a Finnish study (Vähäsantanen, 2015). Beyond school settings, the current accountability and examinations systems, the political agendas and even the catchment area of a school (Hizli Alkan and Priestley, 2019: 9), as well as cultural conditions such as parental expectations, school priorities and values and professional learning expectations may also influence how teachers mediate curriculum-making (Willis et al., 2019: 237).

In sum, teachers’ mediation of the official curriculum is a process influenced by a range of different variables at play in any given moment and varies from context to context as factors combine to create conditions of possibility and constraint (Priestley et al., 2012). The connection between these variables and teacher mediation is not a simple causal one—the embedded nature of the various supports and hindrances makes this a complex relationship (Buehl and Beck, 2015; Imants and Van der Wal, 2020; Skott, 2015).

Teachers as mediators of curriculum: the role of teacher education and professional development

In the emerging trend of educational reforms that “ask for teacher agency” and encourage teachers to make personal sense of reforms and enact them at the local level, mediation of curriculum making is perceived as a critical part of teachers’ role in the classroom (Darling-Hammond et al., 2017; Rajala and Kumpulainen, 2017: 312). For example, the idea of actively shaping the curriculum to include all students, their needs and interests, is assumed to be an element of meaningful education (Biesta et al., 2015). This perception has important implications for teacher education and ongoing professional development. Prospective and experienced teachers need the kind of support that would allow them to develop as curriculum thinkers: enriched understandings of the concept of teacher mediation as a process that involves identity, subjectivity, interpretation, and even contestation (Kohler, 2015: 132); the reflexivity to understand their role as mediators, themselves as a resource for mediating and how their own “identities, historicities and subjectivities” (Kohler, 2015: 196) may enhance or inhibit mediation and finally, tools that can help them manage both the mediations that comprise their professional role (Smith, 2017) and other mediators who can make teachers “do things in certain ways” (Alvunger et al., 2017: 3).

In this direction, frameworks and models of professional learning such as professional learning communities, communities of practice, participatory action research and critical collaborative professional inquiry, have become popular, as professional, and social sites which encourage teachers to see themselves as “curriculum agents or inquirers in practice” (Parker, 1987: 7) and engage in “discomforting” dialogs and organized opportunities of collective sense-making (Ganon-Shilon and Schechter, 2017; Kelchtermans, 2009; Molla and Nolan, 2020; Philpott and Oates, 2017; Priestley and Drew, 2017, 2019). Their effectiveness lies, among other things, in honoring the primacy of teachers’ own experiences (Craig, 2011) and acknowledging that individuals will always interpret, mediate, and adapt what is being suggested.

Concluding remarks

As mentioned earlier, by virtue of who they are teachers mediate curriculum making in unique ways (Spillane, 1999). It was also argued that teachers’ beliefs and the actions that result from them are influenced by the social, the structural, the cultural and the material environments in which teachers work. These observations may falsely lead someone to approach teacher mediation of curriculum as a passive reflection of teacher personalities and the environments in which they act (Edwards et al., 2017) rather than the embodied, agentic and dynamic process it is. Osborn et al.’s (2000) “creative mediation,” Priestley et al.’s (2021) definition of curriculum making and even Connelly and Clandinin’s (1988) “curriculum of lives” foreground exactly that: teacher mediation is an agentic, interactional, and adaptive response (Harris, 2013) to mandated curriculum, students and the contexts within teachers and students operate. It is also a process that requires serious commitment to meaningful learning, professional responsibility to pedagogy and accountability to students (Grimmett and Chinnery, 2009). To achieve this kind of agency, teachers need policy contexts and professional environments that create conditions of possibility.

References

- Alvunger, D., Sundberg, D., Wahlström, N., 2017. Teachers matter—but how? *J. Curric. Stud.* 49 (1), 1–6.
- Alvunger, D., Soini, T., Philippou, S., Priestley, M., 2021. Conclusions: patterns and trends in curriculum making in Europe. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald Publishing, Bingley, England, pp. 273–293.
- Apple, M.W., 2013. The gendered realities of managerialism in education. *Educ. Pol.* 28 (6), 916–924.
- Ball, S.J., Maguire, M.M., Braun, A., 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. Routledge, London.

- Ball, S.J., 2014. *The Micro-politics of the School: Towards a Theory of School Organization*. Routledge, London.
- Ben-Perez, M., 1990. *The Teacher-Curriculum Encounter: Freeing Teachers From the Tyranny of Texts*. State University of New York, Cop, Albany, NY.
- Biesta, G., Priestley, M., Robinson, S., 2015. The role of beliefs in teacher agency. *Teach. Teach.* 21 (6), 624–640.
- Birbili, M., Myrovali, A., 2019. Early childhood teachers' relationship with the official curriculum: the mediating role of professional and policy contexts. *Educ. Inq.* 11 (2), 110–125.
- Buehl, M.M., Beck, J.S., 2015. The relationship between teachers' beliefs and teachers' practices. In: Fives, H., Gill, M.G. (Eds.), *International Handbook of Research on Teachers' Beliefs*. Routledge, New York, NY, pp. 13–30.
- Caine, V., Estefan, A., Clandinin, D.J., 2013. A return to methodological commitment: reflections on narrative inquiry. *Scand. J. Educ. Res.* 57 (6), 574–586.
- Catling, S., 2015. Teachers' perspectives on curriculum making in primary geography in England. In: Catling, S. (Ed.), *Research and Debate in Primary Geography*. Routledge, London, pp. 133–159.
- Ciuffetelli Parker, D., Pushor, D., Kitchen, J., 2011. Narrative inquiry, curriculum making, and teacher education. In: Kitchen, J., Ciuffetelli Parker, D., Pushor, D. (Eds.), *Narrative Inquiries Into Curriculum Making in Teacher Education*. Emerald Publishing, Bingley, England, pp. 3–18.
- Clandinin, D.J., Connelly, F.M., 1996. Teachers' professional knowledge landscapes: teacher stories—stories of schools—stories of schools. *Educ. Res.* 25 (3), 24–30.
- Clandinin, D.J., 1985. Personal practical knowledge: a study of teachers' classroom images. *Curric. Inq.* 15 (4), 361–385.
- Coburn, C.E., 2001. Collective sensemaking about reading: how teachers mediate reading policy in their professional communities. *Educ. Eval. Pol. Anal.* 23 (2), 145–170.
- Coburn, C.E., 2005. Shaping teacher sensemaking: school leaders and the enactment of reading policy. *Educ. Pol.* 19 (3), 476–509.
- Connelly, F.M., Clandinin, D.J., 1988. *Teachers as Curriculum Planners: Narratives of Experience*. Teachers College Press, New York.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14.
- Connelly, F.M., 2013. Foreword. In: Deng, Z., Gopinathan, S., Lee, C.K.E. (Eds.), *Globalization and the Singapore Curriculum: From Policy to Classroom*. Springer, Singapore, pp. vii–xii.
- Craig, C.J., Ross, V., 2008. Cultivating the image of teachers as curriculum makers. In: Connelly, F.M., Fang He, M., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, Los Angeles, CA, pp. 282–305.
- Craig, C.J., 2011. Narrative inquiry in teaching and teacher education. In: Kitchen, J., Ciuffetelli Parker, D., Pushor, D. (Eds.), *Narrative Inquiries Into Curriculum Making in Teacher Education*. Emerald Publishing, Bingley, England, pp. 19–42.
- Craig, C.J., 2010. Teachers as curriculum makers. In: Kridel, C. (Ed.), *Encyclopedia of Curriculum Studies*. Sage, New York, NY, pp. 867–869.
- Craig, C.J., 2020. *Curriculum Making, Reciprocal Learning, and the Best-Loved Self*. Palgrave Macmillan, Cham, Switzerland.
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A.L., Hammerness, K., Low, E.-L., Zeichner, K., 2017. *Empowered Educators: How High-Performing Systems Shape Teaching Quality Around the World*. Jossey-Bass, San Francisco, CA.
- Datnow, A., Castellano, M., 2000. Teachers' responses to success for all: how beliefs, experiences, and adaptations shape implementation. *Am. Educ. Res. J.* 37 (3), 775–799.
- Deng, Z., 2017. Rethinking curriculum and teaching. In: Noblit, G.W. (Ed.), *Oxford Research Encyclopedia of Education*. Oxford University Press, New York, NY, pp. 1–25.
- Dewey, J., 1904. The relation of theory to practice in education. In: McMurray, C.A. (Ed.), *The third NSSE yearbook*. University of Chicago Press, Chicago, IL, pp. 27–28.
- Doyle, W., 1992. Constructing curriculum in the classroom. In: Oser, F.K., Dick, A., Patry, J. (Eds.), *Effective and Responsible Teaching: The New Syntheses*. Jossey-Bass, San Francisco, CA, pp. 66–79.
- Edwards, A., Montecinos, C., Cádiz, J., Jorratt, P., Manriquez, L., Rojas, C., 2017. Working relationally on complex problems: building capacity for joint agency in new forms of work. In: Goller, M., Paloniemi, S. (Eds.), *Agency at Work: An Agentic Perspective on Professional Learning and Development*. Springer, Cham, Switzerland, pp. 229–247.
- Ers, M., Kalmus, V., Autio, T.H., 2016. "Walking a fine line": teachers' perception of curricular autonomy in Estonia, Finland and Germany. *J. Curric. Stud.* 48 (5), 589–609.
- Ganon-Shilon, S., Schechter, C., 2017. Making sense of school leaders' sense-making. *Educ. Manag. Adm. Leader* 45 (4), 682–698.
- Gerrard, J., Farrell, L., 2014. Remaking the professional teacher: authority and curriculum reform. *J. Curric. Stud.* 46 (5), 634–655.
- Giroux, H.A., 1985. Intellectual labor and pedagogical work: rethinking the role of teacher as intellectual. *Phenomenol. Pedagog.* 3 (1), 20–32.
- Giroux, H.A., 2011. Thinking about schools. In: Blair Hilly, E. (Ed.), *Thinking About Schools: A Foundations of Education Reader*. Routledge, London, pp. 121–128.
- Goodson, I.F., 2004. *The Making of Curriculum: Collected Essays*, second ed. Falmer Press, Washington, D.C.
- Gouëdard, P., Pont, B., Hyttinen, S., Huang, P., 2020. *Curriculum Reform: A Literature Review to Support Effective Implementation*. OECD. Working Paper No. 239.
- Grimmett, P.R., Chinnery, A., 2009. Bridging policy and professional pedagogy in teaching and teacher education: buffering learning by educating teachers as curriculum makers. *Curric. Inq.* 39 (1), 125–143.
- Harris, M., 2013. Emergent indigenous identities: rejecting the need for purity. In: Harris, M., Nakata, M., Carlson, B. (Eds.), *Politics of Identity: Emerging Indigeneity*. UTS Press, Broadway, Australia, pp. 10–25.
- Hewitt, T.W., 2006. *Understanding and Shaping Curriculum: What We Teach and Why?* Sage, Thousand Oaks, CA.
- Hizli Alkan, S., Priestley, M., 2018. Exploring Teacher Mediation in Curriculum Making: Scotland and Wales. Paper Presented to the European Conference for Educational Research, Bolzano, 4–7 September.
- Hizli Alkan, S., Priestley, M., 2019. Teacher mediation of curriculum making: the role of reflexivity. *J. Curric. Stud.* 51 (5), 737–754.
- Hizli Alkan, S., 2020. Curriculum making as relational practice: a qualitative ego-network approach. *Curric. J.* <https://doi.org/10.1002/curj.98>.
- Hoyle, E., Wallace, M., 2005. *Education Leadership: Ambiguity, Professionals and Managerialism*. Sage, London.
- Imants, J., Van der Wal, M.M., 2020. A model of teacher agency in professional development and school reform. *J. Curric. Stud.* 52 (1), 1–14.
- Kelchtermans, G., 2009. Who I am in how I teach is the message: self-understanding, vulnerability and reflection. *Teach. Teach. Theor. Pract.* 15 (2), 257–272.
- Ketelaar, E., Beijard, D., Boshuizen, H.P.A., Den Brok, P.J., 2012. Teachers' positioning towards an educational innovation in the light of ownership, sense-making and agency. *Teach. Teach. Educ.* 28 (2), 273–282.
- Kirk, D., MacDonald, D., 2001. Teacher voice and ownership of curriculum change. *J. Curric. Stud.* 33 (5), 551–567.
- Kohler, M., 2015. *Teachers as Mediators in the Foreign Language Classroom*. Multilingual Matters, Bristol, Buffalo.
- Kontovourki, S., Philippou, S., Theodorou, E., 2018. Curriculum making as professionalism-in-context: the cases of two elementary school teachers amidst curriculum change in Cyprus. *Curric. J.* 29 (2), 257–276.
- Lantolf, J., Thorne, S.L., 2007. Sociocultural theory and second language learning. In: van Patten, B., Williams, J. (Eds.), *Theories in Second Language Acquisition*. Lawrence Erlbaum, Mahwah, NJ, pp. 201–224.
- Lingard, B., 2021. National curriculum making as more or less expressions of and responses to globalization. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald Publishing, Bingley, England, pp. 29–52.
- Maguire, M., Braun, A., Ball, S., 2015. "Where you stand depends on where you sit": the social construction of policy enactments in the (English) secondary school. *Discourse* 36 (4), 485–499.
- Maniates, H., 2010. *When Highly Qualified Teachers Use Prescriptive Curriculum: Tensions Between Fidelity and Adaptation to Local Context*. Unpublished doctoral dissertation. University of California, Berkeley.
- März, V., Kelchtermans, G., 2013. Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teach. Teach. Educ.* 29 (1), 13–24.
- Mitchell, D., 2016. Geography teachers and curriculum making in "changing times". *Int. Res. Geogr. Environ. Educ.* 25 (2), 121–133.
- Molla, T., Nolan, A., 2020. Teacher agency and professional practice. *Teach. Teach. Theor. Pract.* <https://doi.org/10.1080/13540602.2020.174019>.

- Nias, J., 2000. Preface. In: Osborn, M.J., McNess, E.M., Broadfoot, P.M., Pollard, A.J., Triggs, P.A. (Eds.), *What Teachers Do: Changing Policy and Practice in Primary Education*, xvi ed. Continuum, London, pp. xi–xiv.
- OECD, 2020. School Autonomy. <https://gpseducation.oecd.org/revieweducationpolicies/#?node=41701&filter=all>.
- Osborn, M.J., McNess, E.M., Broadfoot, P.M., Pollard, A.J., Triggs, P.A., 2000. *What Teachers Do: Changing Policy and Practice in Primary Education*, xvi ed. Continuum, London.
- Olson, J.K., 1980. Teacher constructs and curriculum change. *J. Curric. Stud.* 12, 1–11.
- Parker, W.C., 1987. Teachers' mediation in social studies. *Theor. Res. Soc. Educ.* 15 (1), 1–22.
- Philpott, C., Oates, C., 2017. Teacher agency and professional learning communities; what can learning rounds in Scotland teach us? *Prof. Dev. Educ.* 43 (3), 318–333.
- Pountney, R., 2021. Curriculum Makers and Thinkers: The Case for Curriculum Studies in Teacher Education. https://blogs.shu.ac.uk/kierng/2021/02/?doing_wp_cron=1621143317.1150150299072265625000.
- Priestley, M., Drew, V., 2017. Teacher sense-making in school-based curriculum development through critical collaborative professional enquiry. In: Peters, M., Cowie, B., Menter, I. (Eds.), *A Companion to Research in Teacher Education*. Springer, Singapore, pp. 769–784.
- Priestley, M., Drew, V., 2019. Professional enquiry: an ecological approach to developing teacher agency. In: Godfrey, D., Brown, C. (Eds.), *An Eco-System for Research-Engaged Schools*. Routledge, London, pp. 154–169.
- Priestley, M., Philippou, S., 2018. Curriculum making as social practice: complex webs of enactment. *Curric. J.* 29 (2), 151–158.
- Priestley, M., Edwards, R., Priestley, A., Miller, K., 2012. Teacher agency in curriculum making: agents of change and spaces for manoeuvre. *Curric. Inq.* 42 (2), 191–214.
- Priestley, M., Biesta, G., Philippou, S., Robinson, S., 2016. The teacher and the curriculum: exploring teacher agency. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. Sage, London, pp. 187–201.
- Priestley, M., Philippou, S., Alvunger, D., Soini, T., 2021. Curriculum making: a conceptual framing. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald Publishing, Bingley, England, pp. 1–27.
- Rajala, A., Kumpulainen, K., 2017. Researching teachers' agentic orientations to educational change in Finnish schools. In: Goller, M., Paloniemi, S. (Eds.), *Agency at Work: An Agentic Perspective on Professional Learning and Development*. Springer, New York, NY, pp. 311–329.
- Restad, F., 2021. Exploring the problems and potential of curriculum-making for social learning: implications for policy and practice. *Curric. J.* 1–16.
- Rosiek, J., Clandinin, D.J., 2016. Curriculum and teacher development. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. Sage, London, pp. 293–308.
- Ross, V., Chan, E., 2016. Personal practical knowledge of teacher educators. In: Loughran, J., Hamilton, M.L. (Eds.), *International Handbook of Teacher Education*. Springer, New York, NY, pp. 3–33.
- Schwab, J.J., 1983. The practical 4: something for curriculum professors to do. *Curric. Inq.* 13 (3), 239–265.
- Seashore Louis, K., Robinson, V.M., 2012. External mandates and instructional leadership: school leaders as mediating agents. *J. Educ. Adm.* 50 (5), 629–665.
- Skott, J., 2015. The promises, problems, and prospects of research on teachers' beliefs. In: Fives, H., Gill, M.G. (Eds.), *International Handbook of Research on Teachers' Beliefs*. Routledge, New York, NY, pp. 13–30.
- Smith, R., 2017. Three aspects of epistemological agency: the socio-personal construction of work-learning. In: Goller, M., Paloniemi, S. (Eds.), *Agency at Work: An Agentic Perspective on Professional Learning and Development*. Springer, New York, NY, pp. 67–84.
- Spillane, J.P., 1999. External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *J. Curric. Stud.* 31 (2), 143–175.
- Swart, F., de Graaff, R., Onstenk, J., Knezic, D., 2018. Teacher educators' personal practical knowledge of language. *Teach. Teach.* 24 (2), 166–182.
- Tan, C., 2016. Teacher agency and school-based curriculum in China's non-elite schools. *J. Educ. Change* 17 (3), 287–302.
- Taylor, M.W., 2013. Replacing the "teacher-proof" curriculum with the "curriculum-proof" teacher: toward more effective interactions with mathematics textbooks. *J. Curric. Stud.* 45 (3), 295–321.
- Tyack, D., Cuban, L., 1995. *Tinkering toward Utopia: a century of public school reform*. Harvard University Press, Cambridge, MA.
- Vähäsantanen, K., 2015. Professional agency in the stream of change: understanding educational change and teachers' professional identities. *Teach. Teach. Educ.* 47, 1–12.
- Wallace, C.S., Priestley, M., 2011. Teacher beliefs and the mediation of curriculum innovation in Scotland: a socio-cultural perspective on professional development and change. *J. Curric. Stud.* 43 (3), 357–381.
- Willis, J., McGraw, K., Graham, L., 2019. Conditions that mediate teacher agency during assessment reform. *Engl. Teach. Pract. Critiq.* 18 (2), 233–248.
- Woods, P., Wenham, P., 1995. Politics and pedagogy: a case study in appropriation. *J. Educ. Pol.* 10 (2), 119–141.

Further reading

- Buchanan, R., 2015. Teacher identity and agency in an era of accountability. *Teach. Teach. Theor. Pract.* 21 (6), 700–719.
- Huber, J., Murphy, M., Clandinin, D.J. (Eds.), 2011. *Places of Curriculum Making: Narrative Inquiries Into Children's Lives in Motion*. Adv. Res. Teach., vol. 14. Emerald Publishing, Bingley, pp. 1–10.
- Pyhältö, K., Pietarinen, J., Soini, T., 2018. Dynamic and shared sense-making in large-scale curriculum reform in school districts. *Curric. J.* 29 (2), 181–200.
- Sheikh, I., Bagley, C., 2018. Towards a policy social psychology: teacher engagement with policy enactment and the core concept of affective disruption. *Br. Educ. Res. J.* 44 (1), 43–60.
- Shkedi, A., 2009. From curriculum guide to classroom practice: teachers' narratives of curriculum application. *J. Curric. Stud.* 41 (6), 833–854.
- Stickney, J.A., 2012. Judging teachers: Foucault, governance and agency during education reforms. *Educ. Philos. Theor.* 44 (6), 649–662.
- Stith, I., Roth, W.M., 2010. Teaching as mediation: the cogenerative dialogue and ethical understandings. *Teach. Teach. Educ.* 26, 363–370.

Curriculum development

Dominik Dvořák, Institute for Research and Development of Education, Faculty of Education, Charles University, Prague, Czechia

© 2023 Elsevier Ltd. All rights reserved.

Terminology	149
Historical overview	149
Curriculum revisions today	151
Curriculum development infrastructure and capacity building	153
Conclusion	153
Acknowledgment	153
References	154
Relevant websites	154

Curriculum development may presuppose a narrow conception of the curriculum as a document or plan, and/or an outdated mechanistic idea of its production. Among the community of curriculum scholars, curriculum “making” has emerged in recent years as a more general and comprehensive term. Despite that, curriculum development seems to be a concept that is still firmly established in the discourse of many educators, among them curriculum developers in public and private organizations. Marsh (2004) noted that the relationship of scholars to curriculum development varies and can be categorized into three groups: prescriptive education-alists have produced models for curriculum development, descriptive researchers study real development practices, and finally, crit-ical theorists emphasize problematic aspects of current approaches to curriculum making.

Terminology

As understood today, the curriculum is a complex social practice that takes place across a number of different but connected places, layers, sites, or even arenas (Priestley and Philippou, 2018). In this text, curriculum development is conceived as a specific part of this broad curriculum-making, more or less confined to a particular level or locus, its desired outcome being some kind of formal curricular document or other resource. The following terms tend to be with a similar or the same meaning: curriculum planning, construction (reconstruction), or curriculum design, with a few more related concepts—primarily “review”—discussed below. The various uses of the term “curriculum design” (or redesign) render the terminology even more complicated. While one might distin-guish between curriculum development as a process and curriculum design as a structure or template of the resulting artifact, some highly competent users also understand by curriculum design a process/practice (e.g., OECD, 2020; International Bureau of Educa-tion, 2021). In such a case, in the design phase the overall architecture/template of the curricular document could be agreed, while the development phase means filling in the template (“design”) with actual subject matter or instructional activities. Such a two-step approach may be almost necessary in vocational education and training, where the challenge is the number of parallel curricula that need to be developed for each qualification.

This text mainly addresses curriculum development at the national (system) level, resulting—if everything goes as expected—in some kind of national or state curricular framework, standards, etc. Thus, we concentrate on macro processes, although sometimes a further distinction is made between the macro-level formulation of curricular policy at the government/nation-state level and the actual development of a curricular framework through the specialized infrastructure at the meso level. The creation of a curricular project in an individual school, often referred to as school-based curriculum development (SBCD), is considered here to be a different process, partly because it is usually more or less constrained by the outcomes of national/state framework development. The previous statement does not imply, however, that curriculum developers at the macro/meso level are unbound. Exactly the opposite, it is undeniable that international actors (supra levels) are increasingly influencing the curriculum development of indi-vidual nations or other jurisdictions, although the assessments of the reasons for these influences and their effects differ somewhat. One of the proofs and simultaneously tools of this influence is the fact that international organizations are key producers and disseminators of guidance on curriculum development and analyses of it (e.g. OECD, 2020).

Historical overview

The English-language literature has usually located the beginning of modern curriculum development into the first quarter of the 20th century. At that time, however, in many countries of Continental and Northern Europe, more or less detailed national or central curricula had been in use for decades. Such documents were construed in parallel with the formation of modern nation-states and were intended to serve them, and it is thus understandable that national governments usually appointed small

committees or commissions that actually wrote the central curricula or syllabi. Such an approach may seem to be undemocratic by today's standards. However, it is possible that—despite flashy and costly public consultations—the current real curriculum development behind the scenes is often still “carried out by a small selection of individuals associated with the Ministry/Department of Education” and individuals with power continue to invite “individuals with educational ideologies (...) attractive to them” to revise the national frameworks (Broom, 2020).

The development of a traditional curricular document (in German: *der Lehrplan*, plan of teaching) has been referred to as *Lehrplanarbeit* (curriculum work) or in similar terms. According to the theory of *Didaktik* (Hopmann, 2015), the basic criterion for the inclusion of a specific subject or content into the plan was to be its value for education (in the sense of “*Bildung*”). In Czech, the term “*warp*” (*osnova*) was used to describe such a document, which is somewhat static in comparison with the English “curriculum”, but promises to capture the essence of the field and at the same time assumes further processing (adding more material) to construct instruction.

Early sorting or tracking of students into different types of schools was common, and the ideas of what was valuable for different groups of students also differed. Thus, the curriculum development for different tracks was guided by different principles. The developers and their contemporaries were well aware, too, that the real process of curriculum development was also a kind of curriculum war (*der Streit/Kampf um den Lehrplan*) of educational powers, i.e., of government, churches, business, society, science and the arts, and tradition. Another form of this battle took place between the subject communities for their discipline's relative prominence and the time allocated to it within the *Lehrplan* (Hericks et al., 2008).

New 20th-century approaches to curriculum development, originating mainly in the United States of America, could be understood as an attempt at a new, more rational approach superior both to the unreflective transmission of a canon of disciplinary knowledge (which was no longer feasible as a result of the rapid development of science) and to partisan efforts to promote the agendas of specific pressure groups. New curriculum construction methods were expected to produce curricula corresponding to the needs of citizens of an industrial democratic society. (And it is precisely the very moment when the revision of the school curriculum was performed from progressive and pragmatic positions, which some authors consider to be the twilight of the great humanistic curricular tradition.) Beginning from John F. Bobbitt in the 1910s, these efforts undeniably culminated in the works of Ralph W. Tyler. According to his famous model (often called the Tyler rationale), curriculum development should proceed linearly by addressing four consecutive issues: (1) What educational purposes should the school seek to attain? (2) How can learning experiences be selected which are likely to be useful in attaining these objectives? (3) How can learning experiences be organized for effective instruction? (4) How can the effectiveness of learning experiences be evaluated (Tyler, 1949)? It can be said that, despite all the later criticism, this “ends-means” or product model of curriculum development still remains attractive to many educationalists and technical-professional knowledge of efficient development procedures is deemed to be crucial for curriculum designers.

It should be recalled here, however, that the Anglo-Saxon countries had a culture of strongly decentralized curriculum development, which was very different from continental Europe. Authors such as Bobbitt and Tyler addressed their recommendations directly to teachers and administrators of individual schools or districts. Their relatively simple technical models still remain very popular in areas such as the development of educational materials or curricula designed for a limited number of users, for example within a single university department. As the procedures for the development of frameworks or standards for entire national systems are being studied today, we understand the necessary knowledge of the developers as less normative, with a greater emphasis on understanding the implementation of the curriculum. The consideration of implementation leads to the fact that today's curriculum theorizing often relies much more on ideas about the role of the teacher and the mechanisms of change in the school as an organization.

Tyler and his successors were interested in curriculum development strategies that produced the necessary curriculum as efficiently as possible. In the family of product-oriented models, goals are set at the beginning and activities are selected according to them. But for a long time there have been different approaches to teaching that did not consider it necessary or appropriate always to start from pre-determined goals and objectives. The proponents of such process-oriented models stressed the value of activities or partial contents from which the teacher can find a way to more general goals. Such models of the planning of learning allow, among other things, greater openness of instruction toward the experiences, interests, voices and needs of students or communities. Education becomes more of an encounter between the teacher and the learner.

Tyler himself later pointed to the importance of three sources of the curriculum, which are the learner, subject matter, and contemporary life. That is how substantive expertise (related to specific disciplines/subjects and to learning processes) and socio-political factors found their way back into his model. This was partly in response to two new trends that influenced curriculum development in the second half of the twentieth century—the cognitive revolution and the reconceptualization of curriculum studies.

In the 1960s, the Cold War and the progress of technology brought a growing demand for new curricula that would facilitate school modernization. New approaches to the selection and sequencing of subject matter were promoted, some based on revolutionary ideas about the child's cognitive development and about the specific characteristics of human learning. The thoughts of Jerome Bruner and Lev Vygotskij had a great influence in this decade. The early encounter of all students with abstract concepts or advanced scientific topics can be an example of the principles used in Cold War curriculum development. In some cases, the resulting documents differed considerably from the previous tradition and their implementation in schools faced serious obstacles and public opposition. The perceived failure of the “curricular reform movement” that was just described had consequences that continue to this day.

Since the late 1960s, however, approaches have been emerging in which the experience gained through the creation of the curriculum and deliberation on it is equally important as, or more so than, the artifact that is produced. This experience of curriculum developers is often self-reflective in its nature. A significant number of educationalists (often called reconceptualists) no longer consider the actual production of the curriculum to be their task at all and instead focus on its analysis and critique, increasingly from postmodern or post-colonial positions, once again emphasizing the political and cultural factors inherent in every curriculum. This turn and growing gap between the language of the people actually responsible for curriculum development on one hand and academics on the other hand could be seen as one reason why in recent decades, theorists have played a lesser role in curriculum development, with other stakeholders such as teachers being given a greater say (Broom, 2016).

The critical attitude of the reconceptualists was further strengthened as curricular policy trends changed in the last decades of the 20th century. Centrally established national curricular frameworks began to appear in countries with a previous tradition of decentralized curriculum making. The two different approaches, Continental and Anglo-Saxon, thus became very close: "Under the perspective of development, curriculum and Lehrplan are exchangeable terms, both focusing on policy processes and state-issued guidelines" (Horlacher, 2018). From an Anglo-American point of view, recent decades of more centralized curriculum development have brought about the decline of teachers' curricular expertise. However, from the point of view of other countries, such as those in Central Europe, teachers have received a highly unusual degree of curricular autonomy.

In the United States, in 1989 the pioneering, albeit controversial, Curriculum and Evaluation Standards for School Mathematics represented an important milestone, being initiated by a professional association of teachers outside official government bodies. This provoked a wave of similar resources for other subjects created by various groups and institutions. Conflict between the global tendency toward comparability of student performance and opposition to a centralized curriculum has been resolved in recent years by presenting the development of a national curriculum as a "bottom-up" project growing from the grassroots or middle level. Since 2010, Common Core State Standards (CCSS) have been developed in this way in the United States as an overarching document. A similar development can be observed in Germany, where after World War II curriculum development was the responsibility of the individual Länder (member states of the Federation), but since 2000, standards have been rising at the national (federal) level (Klieme et al., 2004). In contrast, some traditionally centralized European school systems have introduced two- or three-level curriculum development, which is gradually specified at the national, regional, and school levels. Clearly, this is a permanent search for a balance between central regulation and local freedom and flexibility in curriculum development (Kuiper and Berkvens, 2013).

Curriculum revisions today

The processes described in the previous paragraphs have led to a situation where most nations now have framework curricula or standards, either official or de facto. As a result, macro curriculum development today rarely starts de novo, and in reality we usually encounter curriculum change that has to deal with existing traditions, documents, textbooks, and teaching practices. That is why we now usually come across curriculum revisions, redesigns, renewals, or enhancements. Even curriculum reviews do not limit themselves to analytical/evaluative activities (e.g., the "National Curriculum review for England" announced in 2011 was intended to produce draft Programmes of Study for all subjects). However, the aim of the review need not always be changing the status quo. Arguably, some curricular resources are created with the aim of codifying previous practice, rather than modifying it (Hodge et al., 2019). As noted before, curriculum development could be seen as just one phase in these review/revision processes comprising, e.g., analysis/research, development, implementation, and evaluation, leading to a new cycle again (Curriculum review cycle). However, these processes have been little studied by research and are also undertheorized (van den Akker, 2018; Priestley et al., 2021).

It can be said that these changes have similar goals in all countries—namely to adapt the aims and content of school education to advances in technology and society, or to eliminate existing inequalities in educational outcomes. Sometimes, curriculum revision can only be seen as a periodic housekeeping activity involving minor adjustments, but it is often part of a more ambitious education reform. The OECD (2015) points out to a trade-off between how bold the overall goals of the revision are and how precisely the goals are defined. Some curriculum developers focus on technical issues of the measurability of objectives, often stressing basic skills (and sometimes do achieve improvements in learning outcomes). The second approach, broadly corresponding to the family of processual models defined above, is represented by ambitious reforms setting unconventional goals, but not specific in defining them and often also not successful in achieving them. Today, many education systems are expected to instill key competences/capabilities/21st-century skills. Curriculum development can thus become a genuine search for experiences best suited to the building of capabilities, but also a bureaucratic game aimed at showing the compliance of national agencies with international trends. Multiple goals lead to voluminous curricular documents that tend to be ignored by teachers.

Today's curriculum development—in contrast to traditional models—usually not only builds on the older curricular documents of the given school system (diachronic aspect), but also and especially takes into account the concurrent curricular frameworks of other nations (synchronic comparison). Curricular frameworks subject to revision are often compared against those of nations that

perform well in international large-scale assessments as PISA or TIMSS. Such an exercise may be called international benchmarking (Donnelly, 2005; Oates, 2011, 2017). An example of an approach that simultaneously builds on the continental tradition of Didaktik and contemporary cognitive theory is the method of didactic reconstruction (Duit et al., 2012), combining the clarification of subject-specific knowledge, the analysis of students' understanding, and also the study of teachers' ideas. Another contemporary model that responds to the "back to knowledge" movement was proposed by Rata (2021).

Gradually, both inside and outside the professional community, the awareness has been growing that the actual creation of a document is only one step, and by far not the most complicated one. Thus, during the actual production of the curricular text, the chances of successful implementation or enactment could be increased by an iterative approach to design that includes repeated piloting of draft frameworks in schools. Such an approach is sometimes referred to as design research or design-based research or DBR (McKenney et al., 2006). Another of its key features is the close relation to some theoretical framing that not only guides the development of the curriculum, but is simultaneously elaborated.

Since not only content standards, but also performance standards, have begun to play the role of a curricular document in some countries, psychometric methods, specifically performance standard-setting techniques, have also become one of the curriculum development tools. Last but not least, as national curricular frameworks are no longer (just) glossy brochures and become online platforms supporting planning and teaching in schools, curriculum development also includes designing and coding for the technical aspects of the platform and user interfaces.

Despite similar goals and the above common characteristics, emerging comparative studies show that national approaches to achieving these goals differ significantly (Priestley et al., 2021). The differences include who initiates curriculum (re-)development and who is represented in development committees/teams, the timeframe and periodicity of revisions, the use of research and other data/evidence, the degree of continuity or discontinuity with the curricular tradition, etc. (van den Akker, 2018). In some cases, the development of new curricula for different levels of education takes place independently (sequentially), but some countries have developed new curricula for the entire initial general education. The approach to the subject curricula may also vary, as, for example, the digital skills curriculum is likely to require more frequent changes. Also in the field of technical and vocational education and training, the necessary revisions tend to be more frequent, the curriculum needs to be more flexible, and the involvement of employers and/or unions in its quasi-permanent revision is crucial.

Returning to similarities, development/revision is often hijacked by politicians or interest groups, and the attention of the wider audience gets focused on some specific aspect, such as the way the national past is presented in the curriculum. Van den Akker (2018) notes that the time needed to develop/revise a curricular framework varies greatly, from about two to eight years, the average being a framework of around six years. If the implementation and some preparatory phases are included, we reach the figure of about nine years. Clearly, the time scale of curriculum development exceeds the usual length of the election cycle, and politicians may be tempted to speed up the process so that positive effects can be reported within a single election period.

Moreover, curriculum change usually takes place simultaneously with the enactment of other educational reforms or policies that influence or even inhibit curricular processes. The larger and more complex the system, the greater the need to devote a shared understanding to building a new curriculum, both within and across levels (Pietarinen et al., 2017). To increase the legitimacy and ownership of the reform, some form of consultation with relevant actors is now an integral part of all stages of development. Various social networks or dedicated internet environments are sophisticatedly used for these consultations, but the real level of involvement of key groups such as teachers or parents seems to remain limited. It is usually easy to get a massive input on social networks quickly, but its representativeness is questionable and distilling a clear vision could be almost impossible. With cultural wars raging and the growing fragmentation of both the general public and professional communities, the future development of new frameworks could turn out to be mission impossible.

The Netherlands may serve as a case of a country where long-term and systematic reflection on curriculum development has been carried out (Thijs and van den Akker, 2009; Handelzalts et al., 2019) and both academic researchers and curriculum development around the world often draw on the work of Dutch authors. At the same time, even such a commitment to capacity building for design processes does not guarantee a smooth course of concrete curriculum changes (as the example of the most recent revision shows). The case of the Netherlands shows that even stabilized democracies have problems (re-) developing curricular frameworks and the whole process can take a longer time in comparison to the original estimates (Nieveen and Kuiper, 2021).

As noted above, in the 19th century, curriculum development was both a facilitator of the emergence of the nation-state and a reaction to it; in the 21st century the new "national" curricula are in a similar relationship to globalization (Lingard, 2018). Curriculum development can also be understood as a case of the diffusion of policies (Steiner-Khamsi, 2012). Globalization produces new relationships between the various levels at which curriculum development takes place, but they are most appropriately described not as "nice" logical models, but post-bureaucratic structures that are polycentric, transnational, and with a great diversity of actors involving the private and public spheres (the "reform industry"). England could be an example, with the transnational commercial Cambridge Assessment project originating from academia having been linked to the recent re-development of the national curriculum.

Above and beyond the educational policy realm, major social and political transitions frame some curricular revisions, e.g., in post-colonial settings. The transition to democracy and a market economy posed a great challenge for the development curricula in the post-Communist countries (Janík et al., 2020). Unfortunately, it seems that before this process could be completed, new revisions were launched in several countries aimed at suppressing liberal values in favor of national and conservative ones. Curriculum designers are also facing specific challenging situations in bicultural and multicultural settings, (post-) conflict environments and fragile states.

Curriculum trends are constantly evolving, but the capacity to take over and implement them in some countries is limited or borrowed innovations are entering educational practice at a time when they are already being revised or abandoned in the countries of their origin. Arguably, not all nation-states are able to produce quality original curricula (Ramirez and Meyer, 2002), but in some subjects (such as history or often the national language) they have no other choice.

Curriculum development infrastructure and capacity building

Recently, considerable attention has been paid to the importance of building the capacity needed for the active involvement of teachers in curriculum development (e.g., Pieters et al., 2019). Somewhat paradoxically, we know far less about how to build adequate capacity or agency at the macro and meso levels and what support structures there should be. In many European countries, for example, there are permanent agencies whose key role is to develop or revise national curricular frameworks (e.g., in the Nordic countries). Elsewhere, on the other hand, ad hoc teams, sometimes under the auspices of a prominent figure in academia or politics, are entrusted with this task by governments. In the United States, non-governmental actors—professional associations—play a key role in setting subject standards, while—according to the official narrative—the state governors prompted the creation of literacy and numeracy standards. However, it is possible that the real curriculum development processes that take place behind these seemingly different structures and arrangements are quite similar (see above).

In any case, three types of expertise defined above should be represented in the development team/agency: substantive knowledge (related to specific disciplines/subjects); understanding of the broader socio-political contexts of schooling and the effects of education, and the technical-professional knowledge and skills necessary for leading the processes of actual curriculum development (Kuiper et al., 2004).

The political nature of curriculum development has long been recognized. However, it is clear that bureaucrats are also entering the design processes in government agencies. Unlike the role of teachers and politicians, little is known about how design processes are affected by the (potentially self-interest-serving) behavior of civil servants. Another related question somewhat neglected by the research so far is the financial side of curriculum development and, more generally, the role of material flows (a welcome exemption being Dempsey et al., 2021). For example, curriculum development in Central and Eastern European countries has in several cases been influenced by the intention to draw on European funds (external funding-driven reforms).

Conclusion

Despite all the criticism, normative models of curriculum development continue to emerge. They can be useful especially where the development takes place relatively close to the learners and involves a small, relatively homogeneous group of actors. Very similar principles are often rediscovered under different names, and many approaches can be related either to a product (ends-means) model or a process (encounter) model of curriculum development.

Models look so elegant compared to the reality. We can read in them the modernist belief that such a new, fully rational restart of education is possible. The real world of curriculum development at a system level seems to be far more chaotic. Therefore, when creating a model of curriculum development, the educationalist is in danger of believing his/her own construct. Longstreet and Shane (quoted by Marsh, 2004) argue that the curriculum is not some elaborate structure, it is a historical coincidence. It was not developed to achieve any clear set of goals, but rather gradually emerged as the complexity of education increased. We would need more empirical research to better understand these processes.

At the macro level, we now encounter review, revision and renewal processes that try to adapt the historically sedimented curriculum to current requirements, making it also more socially just and inclusive (Young, 2008; Young and Muller, 2010). Supra-level transnational actors such as the OECD, the World Bank, UNESCO, or the EU tend to be directly or indirectly involved, e.g., by producing background papers (e.g., OECD, 2015). Explicit international benchmarking or less obvious borrowing processes become a usual part of such exercises (Baek et al., 2018). Much attention is paid to the transparent communication of all steps, an emphasis on broad participation, and the strengthening of ownership, and the planning of curriculum implementation and teacher support are necessary from the beginning of development activities. However, such a complex enterprise carries with it the risk that the traditional curriculum development issues of the selection of educational content and its organization will be somewhat neglected.

Acknowledgment

This work was supported by The Czech Science Foundation, grant number 15-25137S.

References

- Baek, C., Hörmann, B., Karseth, B., Pizmony-Levy, O., Sivesind, K., Steiner-Khamsi, G., 2018. Policy learning in Norwegian school reform: a social network analysis of the 2020 incremental reform. *Nord. J. Stud. Educ. Policy* 4, 24–37. <https://doi.org/10.1080/20020317.2017.1412747>.
- Broom, C.A., 2016. Power, politics, democracy and reform: a historical review of curriculum reform, academia and government in British Columbia, Canada, 1920 to 2000. *J. Curric. Stud.* 48, 711–727. <https://doi.org/10.1080/00220272.2015.1069402>.
- Broom, C.A., 2020. Procedural democracy: perceptions of the latest curriculum revision in British Columbia, Canada. *Citizsh. Soc. Econ. Educ.* 19, 51–68.
- Dempsey, M., Doyle, A., Looney, A., 2021. The craft of curriculum making in lower secondary education in Ireland. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald, Bingley, pp. 199–222. <https://doi.org/10.1108/978-1-83867-735-020211010>.
- Donnelly, K., 2005. Benchmarking Australian Primary School Curricula. Department of Education, Science and Training, Canberra.
- Duit, R., Gropengießer, H., Kattmann, U., Komorek, M., Parchmann, I., 2012. The model of educational reconstruction—a framework for improving teaching and learning science. In: Jorde, D., Dillon, J. (Eds.), *Science Education Research and Practice in Europe*. SensePublishers, Rotterdam. https://doi.org/10.1007/978-94-6091-900-8_2.
- Handelzalts, A., Nieveen, N., van den Akker, J., 2019. Teacher design teams for school-wide curriculum development: reflections on an early study. In: Pieters, J., Voogt, J., Pareja Roblin, N. (Eds.), *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham. https://doi.org/10.1007/978-3-030-20062-6_4.
- Hericks, U., Kunze, I., Meyer, M.A., 2008. Forschung zu Didaktik und Curriculum. In: Helsper, W., Böhme, J. (Eds.), *Handbuch der Schulforschung*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91095-6_30.
- Hodge, E., Salloum, S., Benko, S., 2019. The changing ecology of the curriculum marketplace in the era of the Common Core State Standards. *J. Educ. Change* 20, 425–446.
- Hopmann, S., 2015. "Didaktik meets Curriculum" revisited: historical encounters, systematic experience, empirical limits. *Nord. J. Stud. Educ. Policy* 2015, 27007. <https://doi.org/10.3402/nstep.v1.27007>.
- Horlacher, R., 2018. The same but different: the German Lehrplan and curriculum. *J. Curric. Stud.* 50, 1–16. <https://doi.org/10.1080/00220272.2017.1307458>.
- International Bureau of Education, 2021. Glossary of Curriculum Terminology. <http://www.ibe.unesco.org/en/glossary-curriculum-terminology>.
- Janík, T., Porubský, Š., Chrást, M., Kuzák, K., 2020. Curriculum Changes in the Visegrad Four: Three Decades After the Fall of Communism. Waxmann, Münster.
- Kuiper, W., Berkvens, J. (Eds.), 2013. *Balancing Curriculum Regulation and Freedom Across Europe*. SLO, Enschede.
- Kuiper, W., Nieveen, N., Visscher-Voerman, I., 2004. Curriculum development from a technical-professional perspective. In: van den Akker, J., Kuiper, W., Hameyer, U. (Eds.), *Curriculum Landscapes and Trends*. Springer, Dordrecht, pp. 177–198. https://doi.org/10.1007/978-94-017-1205-7_11.
- Klieme, E., et al., 2004. *The Development of National Educational Standards*. Federal Ministry of Education and Research, Berlin.
- Lingard, B., 2018. The Australian curriculum: a critical interrogation of why, what and where to? *Curric. Perspect.* 38, 55–65. <https://doi.org/10.1007/s41297-017-0033-7>.
- Marsh, C.J., 2004. *Key Concepts for Understanding Curriculum*. RoutledgeFalmer, London.
- McKenney, S., Nieveen, N.M., van den Akker, J., 2006. Design research from a curriculum perspective. In: van den Akker, J., Gravemeijer, K., McKenney, S., Nieveen, N. (Eds.), *Educational Design Research*. Routledge, London, pp. 67–90.
- Nieveen, N., Kuiper, W., 2021. Integral curriculum review in The Netherlands: in need of dovetail joints. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald, Bingley, pp. 125–150. <https://doi.org/10.1108/978-1-83867-735-020211007>.
- Oates, T., 2017. *A Cambridge Approach to Improving Education. Using International Insights to Manage Complexity*. Cambridge Assessment, Cambridge.
- Oates, T., 2011. Could do better: using international comparisons to refine the National Curriculum in England. *Curric. J.* 22, 121–150. <https://doi.org/10.1080/09585176.2011.578908>.
- OECD, 2015. *Improving Schools in Scotland: An OECD Perspective*. OECD. On-line. <http://www.oecd.org/edu/school/improving-schools-in-scotland.htm>.
- OECD, 2020. *Curriculum (Re)design. A Series of Thematic Reports From the OECD Education 2030 Project*. On-line. <https://www.oecd.org/education/2030-project/contact/brochure-thematic-reports-on-curriculum-redesign.pdf>.
- Pietarinen, J., Pyhäntö, K., Soini, T., 2017. Shared sense-making in curriculum reform: orchestrating the local curriculum work. *Scand. J. Educ. Res.* 63, 491–505. <https://doi.org/10.1080/00313831.2017.1402367>.
- Pieters, J., Voogt, J., Pareja Roblin, N. (Eds.), 2019. *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham. <https://doi.org/10.1007/978-3-030-20062-6>.
- Priestley, M., Philippou, S., 2018. Curriculum-making as social practice: complex webs of enactment. *Curric. J.* 29, 151–158.
- Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), 2021. *Curriculum Making in Europe: Policy and Practice Within and Across Diverse Contexts*. Emerald, Bingley.
- Ramirez, F.O., Meyer, J.W., 2002. *National Curricula: World Models and National Historical Legacies*. Stanford University, Palo Alto.
- Rata, E., 2021. The curriculum design coherence model in the knowledge-rich school project. *Rev. Educ.* 9, 448–495. <https://doi.org/10.1002/rev3.3254>.
- Steiner-Khamsi, G., 2012. Understanding policy borrowing and lending: building comparative policy studies. In: Steiner-Khamsi, G., Waldow, F. (Eds.), *Policy Borrowing and Lending*. World Yearbook of Education 2012. Routledge, London and New York, pp. 3–17.
- Thijs, A., van den Akker, J., 2009. *Curriculum in Development*. SLO, Enschede.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction*. University of Chicago, Chicago.
- van den Akker, J., 2018. *How National Curriculum Frameworks Are Developed: A Comparative Analysis*. A Paper Presented at European Conference of Educational Research, Bolzano.
- Young, M., 2008. From constructivism to realism in the sociology of the curriculum. *Rev. Res. Educ.* 32, 1–28.
- Young, M., Muller, J., 2010. Three educational scenarios for the future: lessons from the sociology of knowledge. *Eur. J. Educ.* 45, 11–27.

Relevant websites

- Australian Curriculum, Assessment and Reporting Authority (ACARA). <https://www.acara.edu.au/>.
- Cidree—Consortium of Institutions for Development and Research in Education in Europe. <https://www.cidree.org/>.
- IBE/UNESCO. <http://inprogressreflections.ibe-unesco.org/prototype-of-a-national-curriculum-framework/>.
- Finnish National Agency for Education. <https://www.oph.fi/en>.
- OECD. <https://www.oecd.org/education/2030-project/curriculum-analysis/>.
- Skolverket. The Swedish National Agency for Education. <https://www.skolverket.se/andra-sprak-other-languages/english-engelska>.
- The Institute for Educational Quality Improvement (IQE). <https://www.iqb.hu-berlin.de/>.
- The National Council for Curriculum and Assessment (NCCA)—Ireland. <https://ncca.ie>.
- The Netherlands Institute for Curriculum Development (SLO). <https://www.slo.nl/international/>, <https://curriculum.nu>.
- The Norwegian Directorate for Education and Training. <https://www.udir.no/en-english/>.

Curriculum and teacher education: the pre-service teacher as curriculum maker

Majella Dempsey, Department of Education, Maynooth University, Maynooth, Co. Kildare, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	155
Curriculum making and educational change	156
Curriculum making and teacher education	158
Space for curriculum making in initial teacher education (ITE)	159
Conclusion: key concepts in curriculum making, agency and curriculum literacy for PSTs	160
References	161

Introduction

Curriculum studies is the nexus where educational thinking and practice meet, by this I mean it links thinking across traditional humanities, social science, philosophy and psychology. Curriculum studies is concerned with what we define as knowledge or content, with discussions around whether knowledge can be defined as being powerful knowledge or knowledge of the powerful in a push away from knowledge relativism (Young, 2008, 2010; Wheelahan, 2010; Rata, 2012, 2019; Wahlström, 2016). Deng (2022) calls for us to think about the concept of a knowledge-rich curriculum. Curriculum studies as a field has gone through a crisis of identity from the assertion by Schwab in 1970's that the field was moribund, to talk of the field being in crisis (Wheelahan, 2010) to a call from Priestley and Philippou (2019) to put curriculum at the heart of practice. Deng in his critique of developments in the discipline of curriculum studies contends that there is an evolution of practice away from classroom concerns to a wider cultural studies perspective "that encompasses literally any topic pertaining to politics, culture, society, environment, technology and human existence" (2018, p. 692). Taking all this into consideration clearly the role of curriculum studies in teacher education is complex. Teacher education strives to make new what is familiar about schooling, what Popkewitz (2014, p. 3) calls "denaturalizing" what seems natural. Pinar (2004, p. 25) calls for teacher education that takes "the generative roles of creativity and individuality in teaching" serious to push against a de-professionalization of the practice. To do this Pre Service Teachers (PSTs) need to develop curriculum thinking (Deng, 2020), they need to develop the capacity to be curriculum makers (Priestley et al., 2021; Deng, 2020) and be afforded the agency (Priestley et al., 2015) to negotiate spaces within sites for curriculum making (Priestley et al., 2021).

The role of the teacher in curriculum has begun to change with the emergence worldwide of new forms of curriculum development and policy, which explicitly reposition the teacher as an agent of curriculum change and active maker of the curriculum in local settings (see for example Priestley and Philippou, 2018; Priestley and Biesta, 2013). Curriculum implementation requires more from the teacher and it demands more from schools than what was historically the case (Dempsey et al., 2021). We see examples of these new forms of curriculum in Ireland in the Junior Cycle Reforms (NCCA, 2015), Curriculum for Excellence in Scotland (Scottish Executive, 2004) and more recently the curriculum reforms in Wales. What is notable about these reforms is how they position the teachers as curriculum interpreters, implementers, makers and enactors. This is a new role for some teachers who traditionally worked with a system that relied on input regulation with tightly prescribed curriculum documents defining what to teach and, in some cases, how to teach content (Nieveen and Kuiper, 2012). While we have moved on from the concept of a teacher proof curriculum, Priestley and Xenofontos (2020) talk about the need for schools to be sites where curriculum questions are posed and addressed constructively, and I would add in some cases where curriculum policy needs to be opposed or advocated for. This role of teachers as curriculum makers is not altogether new as arguably teachers make curriculum in their daily pedagogical decisions, as Bruner (1999) reminds us that pedagogy is never innocent, it carries its own messages. Clarification of the concept of curriculum has the potential to deepen our understanding of/insights into pedagogical processes. Such enrichment has a special relevance for teacher education.

Clearly this role for teacher as curriculum maker needs attention in teacher education courses. Curriculum making requires PSTs to relate theory and practice in their university learning and in their practicum experience, the boundaries between both sites are blurred. Priestley and Philippou (2018) contend that curriculum making occurs

Across multiple sites, in interaction and intersection with one another, in often unpredictable context-specific ways, producing unique social practices, in constant and complex interplay, wherein power flows in non-linear ways, thus blurring boundaries between these multiple sites. (p. 154).

For the purposes of this Chapter, I will use Priestley's (2019) definition of school curriculum as "the multi-layered social practices, including infrastructure, pedagogy and assessment, through which education is structured, enacted and evaluated" (p. 8). This definition in tandem with the definition of curriculum making above recognizes how curriculum is a social

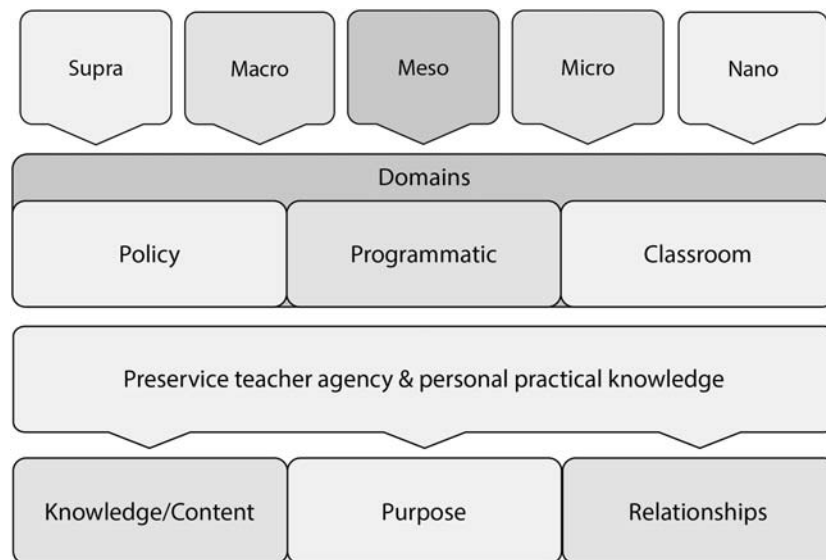


Fig. 1 Pre service teacher as curriculum maker – complex webs of enactment.

practice, between teachers, between a teacher and their students, between school leaders and teachers, parents and students, policy makers and teachers and so on, there are multiple actors involved in making curriculum. It also acknowledges how curriculum is made in multiple sites, for example, at the policy table, the school, classroom level and individual level, among others (Priestley et al., 2021). This definition also recognizes how curriculum is political, and involves power and occurs across different layers (Priestley et al., 2021). For curriculum change to happen there needs to be alignment between the content or knowledge, the teaching, the learning and the assessment practices (Dempsey et al., 2021). If any one of these are not aligned curriculum making is stifled. Looney (2001) contends that curriculum, assessment, pedagogy and organization are treated in isolation. It is the intersection of these message systems at the school level that is important, and this is often not analyzed. These message systems are central to curriculum making. In addition, it is difficult to address teachers' role in curriculum making without looking at the concept of agency and the ability to exert agency in the sites for curriculum making (Priestley et al., 2015).

In this Chapter, I will outline the complexity of curriculum making for PSTs as they navigate sites where it happens (Priestley et al., 2021), within the policy, programmatic and classroom domains (Deng, 2018, 2020; Doyle 1992), and make judgments on what content/knowledge to teach (Young, 2008; Rata, 2012; Biesta, 2015; Deng, 2020), and for what reason (Biesta, 2015; Deng, 2020) while they seek to bring the curriculum alive in fruitful experiences for the students they teach (for example, Shulman and Shulman, 2004; Lambert and Biddulph, 2014; Schulte, 2018). Fig. 1 gives an overview of how the article links the thinking on curriculum making within the Initial Teacher Education (ITE) space.

Curriculum making and educational change

Across the world there are multiple actors involved in processes of curriculum reform, often with conflicting ideas and priorities. These actors can represent cultural, political or epistemological stances, or a combination of all three (Slattery, 2013). Apple writing in 2006 from the perspective of the USA, reminded us that education is a site of crucial struggles over authority and identity and over what it means to be educated. Issues such as an audit culture and privatization are even more pertinent today considering new forms of curriculum development where there is a push for skills linked to the economy, personalization of learning, a move to output regulation, datafication and a marketization of education (Lingard, 2011; Lingard, 2013; Biesta, 2015; Winter, 2017; Tahirshylaj and Wahlström, 2019). These big issues are not easy to address, they are complex. Curriculum has a major role to play in addressing them. At the core of each new wave of educational change is the curriculum. Curriculum development, implementation, evaluation and curriculum policy development and analysis are central to educational change (Connelly and Xu, 2010). Without changing the curriculum there can and will only be structural change. We have seen this in Ireland with successive efforts to change the experience at lower secondary for students (Gleeson, 2012). Until the present reform there was little success (Smyth and McCoy, 2011). This was due in part to a lack of alignment between teaching, learning, assessment and curriculum content (Dempsey et al., 2021). Without considering the variety of ecological factors in schools, policy change is doomed to fail or at best be partially implemented (Priestley et al., 2012). In 2018 Apple still had concerns for the field of education and in particular the decisions around curriculum policies and practices. He wrote.

With neoliberal, neoconservative, authoritarian populist, and new managerial forces increasingly occupying the space of real policies and practices, we have little voice in the public debates over the realities of schooling and the decisions of curriculum policies and practices. The field of education deserves more.
(Apple, 2018, p. 689).

Deng (2020) building on the work of Doyle (1992) describes three levels of curriculum making, the policy, the programmatic and the classroom level and in doing so he highlights how the teacher can exert power and have voice in education. Curriculum making within the policy level is concerned about what schooling should be with respect to society and culture. It is about the big questions such as what is education for? (see Biesta, 2015), in this domain curriculum making is viewed as “something with social, economic and political meaning and significance” and contributing to the well-being of society (Deng, 2011, p. 547). In the programmatic context curriculum making is concerned with school subjects and program content. This is what some term the official or written curriculum (see van den Akker, 2003). Curriculum making at the classroom level is how teachers transform the official or programmatic into “educative experiences for the students” (Deng, 2020, p. 31), where the programmatic curriculum “invites critical testing rather than acceptance” (Stenhouse, 1975, p. 142). This central role for teachers in curriculum making in unlocking what “*curriculum potential* a material has depends on how the teachers uses the material” (Deng, 2011, p. 538 emphasis in original). It is important to note that the choices teachers and PSTs make is grounded in knowledge that is personal (Clandinin and Connelly, 1988; Haraldstad and Kristiansen, 2020) and entails professional judgment (Goodson, 1997; Biesta, 2015).

Priestley et al. (2021, p. 12) posit that curriculum making occurs within and between layers or sites of social activity and is defined and determined by the nature of the curriculum making activity rather than the site at which it occurs. Table 1 provides a typology for the sites where curriculum making occurs, this work by Priestley et al. (2021) builds on earlier work by Thijs et al. (2009) and seeks to move from a linear concept of layers to sites. The influence of transnational curricular discourse at the supra site is more evident since Lisbon in 2000 where education was linked with social policy, labor market and overall economic policy and with the introduction of Open Method Coordination^a (European Council, 2000). This has led to policy borrowing and peer learning, but also to ideation convergence and an obsession, some would say, with global data (Raffe, 2011; Dale and Robertson, 2012; Lingard, 2011). Curriculum is not immune to this supra influence, however, it is not as prevalent as one might first think. Even with all these international data (such as PISA and TALIS) available for comparisons, there is little evidence of it having direct impact on educational policies, curricula and the way teaching and learning are organized in schools according to some commentators (Drent et al., 2013). Some influence can be seen in new forms of curriculum development where key competences or versions of them can be found in most curricula (Gordon et al., 2009; Halász and Michel, 2011; Biesta and Priestley, 2013; Dempsey, 2016).

The question arises as to where in this typology does teacher education exist? Is it between the meso and the micro sites? The policy governing the organization and structure of teacher education programmes rests in the macro space and is influenced by the supra space. Conway and Murphy (2013) contend that the interrelated influences of the European higher education space, education legislation and professional self-regulation policies from local regulatory bodies all exert influence on teacher education. The three interrelated domains of curriculum making described by Doyle (1992) and expanded on by Deng (2018, 2020): the policy curriculum, the programmatic curriculum, and the classroom curriculum are not unlike these sites above. Deng (2010, 2017) contends that these domains are not only inter-related but also inter-dependent, Priestley et al. (2021) talk about the how curriculum making at these sites is complex and multidimensional with porous borders between sites. A key challenge is for PSTs to locate themselves within this curriculum making space and to appreciate how the layers/domains or sites are inter-dependent and influence their professional practice.

Table 1 Sites of curriculum making.

<i>Site of activity</i>	<i>Examples of activity</i>	<i>Examples of actors</i>
Supra	Transnational curricular discourse generation, policy borrowing and lending; policy learning	OECD; World Bank; EU
Macro	Development of curriculum policy frameworks;	National governments; curriculum agencies; teaching regulatory bodies
Meso	Production of guidance; leadership of and support for curriculum making; production of resources	National governments; curriculum agencies; textbook publishers; professional development providers
Micro	School-level curriculum making; program design; lesson-planning	Principals; senior leaders; middle management; subject teams; teachers; PSTs
Nano	Curriculum making in classrooms and other learning spaces; pedagogical decisions	Teachers; PSTs; students

Source: Adapted from Priestley et al. (2021), p. 13.

^aThe Open Method of Coordination (OMC) is an EU policy-making process, or regulatory instrument, formally initiated by the Lisbon European Council in 2000. The OMC does not result in EU legislation, but is a method of soft governance which aims to spread best practice and achieve convergence toward EU goals in those policy areas which fall under the partial or full competence of Member States.

PSTs are in university being inducted into a profession, forming their teacher identities, developing their pedagogical knowledge, developing professional judgment while negotiating what it means to be a curriculum maker. Priestley et al. (2021) describe this complexity of how one needs to negotiate this social activity and;

Operate in education systems as curriculum is made and remade in different settings; as sites of social activity with changing social actors, who are moving between sites, are being present in more than one, or actually becoming a site themselves depending on the social activity they are engaging in (p. 12).

PSTs are straddling the university and the school site, working with social actors in both sites. It is interesting that PSTs or teacher education do not feature in the representations of layers or sites to date. Priestley et al. (2021, p. 14) see this typology as “an analytical tool for understanding how curriculum making occurs and emerges in different contexts” and for this reason is very useful when thinking about the teacher education space. The work of teacher educators with PSTs in the university setting around curriculum could be construed as micro curriculum making, this is where the PSTs either work individually or in teams to plan topics for teaching when in school. It is both linked back to the macro site through subject specification, curriculum planning materials and supporting documentation and the nano site of the classroom where the teaching will occur.

Curriculum making and teacher education

The term curriculum means different things to different people depending very much on their context and the site they are working in. Deng (2017, p. 1) contends that “the matter of defining curriculum is in a state of disarray.” In his 2017 paper Deng provides multiple definitions of curriculum and teaching and in doing so not only highlights the complexity of the area but also offers a promising way to revitalize curricular and pedagogical theory as applied to practice in schools and classrooms in context. In recent years scholars including Deng (2017) and Priestley (2019) have provided definitions that aim at dealing with this complexity. Building on the definition for curriculum provided by Priestley (2019, p. 8) as “the multi-layered social practices, including infrastructure” it is important to unpack what the concept of multi-layered means in the context of teacher education. Deng (2020) provides a succinct description of how curriculum thinking across curriculum policy, curriculum planning and classroom teaching is needed for curriculum making by teachers.

Curriculum planning at the institutional level entails an interpretative, deliberative process within a particular social context, with specific groups of learners in mind. Through this process, various types of knowledge are selected and organised into content in the light of their potentials for the cultivation of human powers. In the classroom a teacher interprets and enacts the content of a school subject – embedded in the institutional curriculum – to create “fruitful meetings” between students and content that give rise to the cultivation of human powers (p. 91).

Curriculum making in the context of this chapter is seen as a social practice, where each actor draws on “personal practice knowledge” (Connelly and Clandinin, 1988, p. 25), linked to the concept of reconstructing the past and the projective intentions for the future within the cultural, structural and material of the present (Priestley et al., 2015; Wallace and Priestley, 2017). Cornbleth (1990) contends that curriculum is contextually shaped. Understanding curriculum making and agency in this way helps us to acknowledge the many dilemmas experienced by the PSTs as they negotiate and use this personal practice knowledge in developing curriculum within the dominant social, cultural, political and institutional narratives structuring educational contexts (Cornbleth, 1990; Goodson, 2003; Priestley et al., 2015). In the Irish context we are working within an education system that has a history in the post-primary sector, at least, of predominantly top-down linear curriculum development, with a high stake externally assessed examination leading to little space for curriculum making at the micro and nano domains (Banks and Smyth, 2015). However, the process of curriculum development in recent years has changed with more emphasis on research, consultation and working with schools and curriculum specifications that encourage local interpretation (Dempsey et al., 2021).

Loughran (2006, 2010) reminds us that students of teaching have a natural desire to learn about teaching, however, he asserts that learning to teach about teaching is confusing, full of uncertainty and complexity. Learning how to be a curriculum maker amid this complexity is no easy task but an essential one. Biesta (2015) reminds us that “the point of education is that students learn something, that they learn it for a reason, and that they learn it from someone” (p. 76). Deng provides an excellent synopsis of this vital role the teacher plays in curriculum making.

That teaching is an ethical and intellectual undertaking that is vital for social reproduction and innovation, human development and the flourishing of students – and for which content is an essential resource. And teachers, being (as they are) at the heart of such an undertaking, are curriculum makers (or theorists) who must grapple with the intellectual and moral questions of what content should be taught, why it should be taught and how it should be taught within a particular classroom context (2020, p. 65).

This central role for teachers is also echoed by Biesta (2015) when he describes teachers as educational professionals that should have scope for judgment and discretion in how, why and what they teach. Schwab’s (1978) four commonplaces recognizes the

teacher as central to curricular decision making further supporting the central role the teacher has in bridging the gap between the curriculum content and the experienced curriculum of the student.

Current teacher education reforms emphasize teachers learning more disciplinary knowledge so that they can become more competent (Conway and Murphy, 2013). Some see this as a result of the neoliberal push for benchmarking, standards and numerical ranking (Lingard, 2011; Popkewitz, 2014; Sellar and Lingard, 2014). The thinking is that if teachers know more, they will be better teachers. There is some merit to this kind of thinking, see, for example, the large body of research on Pedagogical Content Knowledge (PCK), a concept first mooted by Shulman (1986) and later expanded to include Mathematical Knowledge for Teaching (MKT) the mathematical knowledge required for teaching mathematics (Ball et al., 2008) and Technological Pedagogical Content Knowledge (TPCK) the knowledge needed to integrate technology into teaching (Mishra and Koehler, 2006) among others. PCK and associated concepts have been the subject of intensive investigation (Depaepe et al., 2013). Seidel and Stürmer (2014) state that one of the main aims of teacher education is to help teachers acquire integrated knowledge of theory and practice. However, Deng (2020, p. 60) contends that PCK needs to be expanded to include “curriculum thinking, alongside specialist curriculum knowledge” for teachers to identify and interpret the content and unlock its potential. The important concept here is how teachers link their knowledge of their discipline, with their knowledge of pedagogy with their understanding of the individuals they are teaching in the context of their classroom at a particular time while acknowledging the influences from within and outside the classroom. This is what curriculum making entails in each moment in teaching, see Fig. 1.

Phelps and Lee (2003) contend that teachers tend to revert to teach in a way which they themselves were taught. For example, Conway and Murphy (2013) found that PSTs working with schools in reform-oriented mathematical contexts are presented with several dilemmas. These are the mismatch between their experience as a secondary school student studying mathematics, their experience as a university student studying mathematics and education, and their experience as a pre service teacher on placement in a school. The impact of these three types of experience on how PSTs construct, deconstruct and reconstruct their professional identities is well documented (see for example Flores and Day, 2006; also, Osgood, 2010 with early childhood educators). Part of this development of professional identity is how PSTs approach curriculum, with what Olson refers to as “unique narratives of experience” (2000, p. 173). Gholami et al. (2020) refer to an “internal dialog” that PSTs established between their personal and professional selves while observing real conditions of students in schools, this dialog contributes to their construction of professional identities. This dialog, discourse and language is crucial to the process of constructing and construing the world and self (Popkewitz, 2014) and is imperative to PSTs developing curriculum making capacities. Teacher education can underplay and misrepresent the process of curriculum making and the complexity of teaching as a social process (Priestley and Philippou, 2018). As stated earlier the complexity of developing this specialist curriculum knowledge (Deng, 2020) and curriculum thinking for curriculum making (Deng, 2022) in teachers’ education is influenced by personal practical knowledge and personal agency (Priestley et al., 2015). Priestley and colleagues describe agency as;

An emergent phenomenon- as something that is achieved by individuals, through the interplay of personal capacities and resources, affordances and constraints of the environment by means of which individual act (p. 19. emphasis in original).

If we apply this thinking on agency to PSTs, we can see how their life and histories are intertwined in their personal practical knowledge found “in the person’s past experience, in the person’s present mind and body, and in the person’s future plans and actions” (Connelly and Clandinin, 1988, p. 25). Clandinin and Connelly (1992) drew of Dewey’s (1938) work on experience and Schwab’s (1978) four commonplaces, that is, teacher, learner, subject matter and milieu and conclude that curriculum might be viewed as “an account of teachers’ and students’ lives together in schools and classrooms” (p. 392). They see curriculum making as something that is dynamic, experienced in relation to others with as esthetic and moral component. This view is also echoed in Deng’s (2020) concept of cultivation-via knowledge and in Priestley et al. (2015) conception of agency. For Deng (2020), there are three propositions for cultivation-via-knowledge platform, one is a vision of education, two an attendant theory of knowledge and three, a “theory of content which serves to inform curriculum planning and classroom teaching” (p. 91). However, Gleeson and O’Flaherty (2014) caution that there exists a theory/practice dichotomy for PSTs as they grapple with the own beliefs and that of schools, teachers and the external environment. Also some argue that teachers and PSTs need to develop “the complex skill of understanding the double-speak of policy documents” (Schulte, 2018, p. 634).

Space for curriculum making in initial teacher education (ITE)

From the earlier discussion in this chapter we see that curriculum making and agency is contingent on personal capacities but also dependent on cultural and structural working conditions (Priestley et al., 2015). For PSTs on their Initial Teacher Education (ITE) course the space for curriculum making is crucial as they navigate the cultural and structural conditions of the university and then also to the school where they have their practicum. Nieveen and Kuiper (2012) in their cross-case analysis of curriculum regulation provide a good overview of how the pendulum swings from tight to loose regulation in curriculum. They describe what this regulation means.

“Curriculum regulation” reflects a government’s intention to prescribe the high-fidelity implementation of directives at the input level (goals and contents, in terms of “goals to attain”) and at the output level (modes of assessments and examinations). Those prescriptions imply that the room for site-specific curricular choices is restricted. On the other hand, “curriculum deregulation” reflects a government’s intention to refrain from prescription and control at the input and output level by stimulating school-based decision-making “through soft tools and systems” (p. 359).

At the heart of curriculum deregulation is a culture of trust in schools and teachers having the professional freedom to make site-specific interpretations of curriculum specifications, in essence a desire to provide space for curriculum making (Priestley et al., 2021) and professional judgment (Biesta, 2015). In Scotland researchers found that pressures of attainment and other accountability mechanisms acted as a barrier for teachers to exert professional agency (Wallace and Priestley, 2017). Clearly curriculum policy can afford teachers and PSTs greater flexibility, autonomy and space to be curriculum makers, but must do so with an acknowledgment of local ecologies and the capacity and agency for curriculum making.

For PSTs to move away from a transactional classroom environment to one where they are curriculum makers they need to link their practice to theory and in cases to many theories. For example, if a PST wants to incorporate dialogical teaching into their mathematics class, they might draw on theorists such as Dewey (1938), Alexander (2017), and Ball (2000) including others. In planning for a class that incorporates dialogical teaching they are consulting philosophy of education, sociology of education and learning theories and making decisions around knowledge selection (Young, 2010; Lambert, 2018), select tasks for teaching and assessment (Dempsey and O'Shea, 2020). Schulte describes how this classroom enactment further interacts with various ideas and practices as follows:

They engage with school policies, school and teacher organization, and various forms of teacher training (both pre-service and in-service); the kind of teaching guidelines and materials at disposal (such as various national, regional, and local curricula and textbooks); those again are partially an outcome of academic research and discourse, as well as of various forms of policy-making among political bodies 2018, p. 626.

This classroom practice in schools, in the words of Deng (2018) requires our PSTs to use theories and models in “an eclectic, critical and innovative manner” rather than rely on one theory or doctrine. It calls for greater conceptual clarity between curriculum and pedagogy (Lambert and Biddulph, 2015). This use of theory and curriculum thinking (Deng, 2020) and teacher judgment is central to what it means to be a professional (Biesta, 2015; Deng 2020).

Conclusion: key concepts in curriculum making, agency and curriculum literacy for PSTs

Curriculum making entails three inter-related processes. The first is selection of content to teach. As stated earlier this can be highly prescribed in subject specifications or leave space of professional judgment. Regardless, the PST must make a judgment on what knowledge/content they will teach. This may include powerful knowledge (Young, 2010), skills or competences and attitudes or dispositions. Lambert reminds us that the curriculum is about knowledge selection, but teachers need to work with this knowledge and ensure that students are meaningfully engaged with it (2018, p. 14). The PSTs then makes decisions on how to teach the knowledge/content including decisions on the type of tasks to use (Dempsey and O'Shea, 2020), the kind of pedagogical experiences they will set up for their students and how they will assess the learning. At the heart of these decisions is relationships, the relationship between the PST and their discipline (subject), their students, colleagues and community. The pedagogical decisions are not only part of what happens in the classroom but “they further interact with various ideas and practices” from other sites of curriculum making (Schulte, 2018, p. 626; see also Priestley et al., 2021). Linking to the purpose of education they must teach this content for a reason. What is critical is the human interaction not just a knowledge transaction exercise.

Curriculum making entails deliberative decision-making at the planning stage and in the work within the classroom. These decisions are linked to the purpose of education and to the needs of the students within a specific context of a school or classroom it is both practical and projective (Priestley et al., 2015). Deliberative decision-making requires the use of theory to illuminate and interpret specific issues and problems (Deng, 2018) and involves fundamental questions of teachers thinking and teachers doing (Connelly and Clandinin, 1988). PSTs need to be able to reflect on practice within the context of their schools, their classes and their students. In doing this they bring with them their past and present educational experiences and have a focus on the future, the way they want to develop as teachers, how they wish their practice to be embodied in their professional career. What can be challenging for PSTs is how they grapple with the disconnect between their personal experience of education and the theory and practice they are being taught in their university pre service teacher education course. Equally the mismatch between their university course material and the observed practice in schools can cause dissonance. The school can be a “world of conflicting ideas and actions” (Teo, 2012, p. 672; see also Gleeson and O'Flaherty, 2014).

For PSTs to engage with curriculum making in their ITE courses, they must see themselves as having agency within the sites for curriculum making (Priestley et al., 2021), they must be provided with opportunities to develop curriculum making capacities. One such capacity is relating the theory to practice “in an eclectic, creative and innovative manner, for the purposes of constructing theory that matters in practice and in the world of schooling” (Deng, 2018 p. 707). Another is making judgments about appropriate content, pedagogy, organization of the classroom and so on (Lambert and Biddulph, 2014; Biesta, 2015; Deng, 2017, 2020). Finally PSTs need to become curriculum thinkers where they think about the what and why of teaching and make decisions, judgments on the educational potential of content (Deng, 2021; see also Biesta, 2015). Building on the work of Biesta on the three purposes of education that of qualification, subjectification and socialization, PSTs need to keep these “in an educationally meaningful balance” (2015, p. 78). This is a complex task that requires professional judgment. PSTs need to know “the learner” and become aware of the “educational milieu” is a major challenge. Teaching is a social practice and us working in teacher education need to recognize that something like a teaching experience during ITE is rooted in one context but that the chances are high that the PSTs will progress to

work in very different contexts. Thus, being able to listen to and to know the learner and, indeed, the educational milieu (school culture and so on) is a vital, and difficult skill to develop. At best we strive to “denaturalize” the taken for granted and to open up the space to engage with these curriculum questions (Popkewitz, 2014).

Priestley and Xenofontos (2020) tell us that we need to focus less on excellent teachers and more on excellent teaching, and in doing this we avoid unnecessary dichotomies and see different curriculum practices as interrelated. Fig. 1 strives to bring some of the complexity of curriculum making and thinking together. One of the purposes of education is “to enfold within each present moment the past, the present, and the future so that our lives will be illuminated with deep understanding” (Slattery, 2013, p. 90). It is the enfolding described here that makes ITE and curriculum making such a complex but rewarding endeavor.

References

- Alexander, R.J., 2017. *Towards Dialogic Teaching: Rethinking Classroom Talk*, fifth ed. Dialogos, York.
- Apple, M., 2006. *Educating the “Right” Way: Markets, Standards, God, and Inequality*, second ed. Routledge, New York.
- Apple, M., 2018. Critical curriculum studies and the concrete problems of curriculum policy and practice. *J. Curric. Stud.* 50 (6), 685–690.
- Ball, D.L., Thames, M.H., Phelps, G., 2008. Content knowledge for teaching: what makes it special? *J. Teach. Educ.* 59 (5), 389–407.
- Ball, D.L., 2000. Bridging practices: intertwining content and pedagogy in teaching and learning to teach. *J. Teach. Educ.* 51 (3), 241–247.
- Banks, J., Smyth, E., 2015. ‘Your whole life depends on it’: academic stress and high-stakes testing in Ireland. *J. Youth Stud.* 18 (5), 598–616.
- Biesta, G., Priestley, M., 2013. Capacities and the curriculum. In: Priestley, M., Biesta, G. (Eds.), *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. Bloomsbury Academic, London, pp. 35–49.
- Biesta, G.J.J., 2015. What is education for? On good education, teacher judgement, and educational professionalism. *Eur. J. Educ.* 50 (1), 75–87.
- Bruner, J., 1999. Folk pedagogies in leach. In: Moon, B. (Ed.), *Learners and Pedagogy*. Paul Chapman, London.
- Clandinin, J., Connelly, M.F., 1988. Studying teachers’ knowledge of classrooms: collaborative research, ethics, and the negotiated narrative. *J. Educ. Thought* 22 (2A), 296–282.
- Clandinin, D.J., Connelly, F.M., 1992. The teacher as curriculum maker. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum: A Project of the American Educational Research Association*. Macmillan, New York, pp. 363–401.
- Connelly, F.M., Clandinin, D.J., 1988. *Teachers as Curriculum Planners: Narratives of Experience*. Teachers College Press, New York, NY.
- Connelly, F.M., Xu, S., 2010. An overview of research in curriculum inquiry. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, Oxford, UK, pp. 324–334.
- Conway, P., Murphy, R., 2013. A rising tide meets a perfect storm: new accountabilities in teaching and teacher education in Ireland. *Ir. Educ. Stud.* 32 (1), 11–36. <https://doi.org/10.1080/03323315.2013.773227>.
- Combleth, C., 1990. *Curriculum in Context*. Falmer Press, Basingstoke.
- Dale, R., Robertson, S.L., 2012. Towards a critical grammar of education policy movements. In: *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. Routledge, London and New York, pp. 21–40.
- Dempsey, M., O’Shea, A., 2020. The role of task classification and design in curriculum making for preservice teachers of mathematics. *Curric. J.* 31 (3), 436–453.
- Dempsey, M., Doyle, A., Looney, A., 2021. The craft of curriculum making in lower secondary education in Ireland. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy, Practice Within and Across Diverse Contexts*. Emerald publishing, London.
- Dempsey, M., 2016. *Exploring the Impact of a Key Skills Approach to Teaching and Learning in Secondary Education*. Thesis. Trinity College Dublin, Dublin.
- Deng, Z., 2010. Curriculum planning and systems change. In: McGaw, B., Baker, E., Peterson, P. (Eds.), *International Encyclopaedia of Education*, third ed. Elsevier, Oxford, UK, pp. 384–389.
- Deng, Z., 2011. Revisiting curriculum potential. *Curric. Inq.* 41, 538–559.
- Deng, Z., 2017. Rethinking curriculum and teaching. In: Noblit, G.W. (Ed.), *Oxford Research Encyclopaedia of Education*. Oxford University Press, New York, NY, pp. 1–25.
- Deng, Z., 2018. Contemporary curriculum theorizing: crisis and resolution. *J. Curric. Stud.* 50 (6), 691–710. <https://doi.org/10.1080/00220272.2018.1537376>.
- Deng, Z., 2020. *Knowledge, Content, Curriculum and Didaktik: Beyond Social Realism*. Routledge, London.
- Deng, Z., 2021. Constructing ‘powerful’ curriculum theory. *J. Curric. Stud.* 53 (2), 179–196. <https://doi.org/10.1080/00220272.2021.1887361>.
- Deng, Z., 2022. Powerful knowledge, educational potential and knowledge-rich curriculum: pushing the boundaries. *J. Curric. Stud.* <https://doi.org/10.1080/00220272.2022.2089538>.
- Depaepe, F., Verschaffel, L., Kelchtermans, G., 2013. Pedagogical content knowledge: a systematic review of the way in which the concept has pervaded mathematics educational research. *Teach. Teach. Educ.* 34, 12–25.
- Dewey, J., 1938. *Experience and Education*. Macmillan, New York.
- Doyle, W., 1992. Constructing curriculum in the classroom. In: Oser, F.K., Dick, A., Patry, J. (Eds.), *Effective and Responsible Teaching: The New Syntheses*. Jossey-Bass, San Francisco, CA, pp. 66–79.
- Drent, M., Meelissen, M.R.M., van der Kleij, F.M., 2013. The contribution of TIMSS to the link between school and classroom factors and student achievement. *J. Curric. Stud.* 45 (2), 198–224.
- European Council, 2000. Lisbon European Council 23 and 24 March 2000. Presidency Conclusions. Council of European Union, Lisbon.
- Flores, M.A., Day, C., 2006. Contexts which shape and reshape new teachers’ identities: a multi-perspective study. *Teach. Teach. Educ.* 22, 219–232.
- Gholami, K., Faraji, S., Meijer, P.C., Tirri, K., 2020. Construction and deconstruction of student teachers’ professional identity: a narrative study. *Teach. Teach. Educ.* <https://doi.org/10.1016/j.tate.2020.103142>.
- Gleeson, J., O’Flaherty, J., 2014. Teachers’ perceptions of moral/ethics education in catholic secondary schools in Australia and Ireland. In: Paper Presented at the Annual Meeting of the American Educational Research Association (Philadelphia, PA).
- Gleeson, J., 2012. The professional knowledge base and practice of Irish post-primary teachers: what is the research evidence telling us? *Ir. Educ. Stud.* 31 (1), 1–17. <https://doi.org/10.1080/03323315.2011.617945>.
- Goodson, I.F., 1997. ‘Trendy theory’ and teacher professionalism. *Camb. J. Educ.* 27 (1), 7–22. <https://doi.org/10.1080/0305764970270102>.
- Goodson, I.F., 2003. *Professional Knowledge, Professional Lives: Studies in Education and Change*. Open University Press, Philadelphia.
- Gordon, J., Halasz, G., Krawczyk, M., Leney, T., Michel, A., Pepper, D., Putkiewicz, E., Wisniewski, J., 2009. CASE Network Report: Key Competences in Europe: Opening Doors for Lifelong Learners Across the School Curriculum and Teacher Education. Case-Center for Social and Economic Research, Warsaw.
- Halász, G., Michel, A., 2011. Key Competences in Europe: interpretation, policy formulation and implementation. *Eur. J. Educ.* 46 (3), 289–306.
- Haraldstad, A., Kristiansen, A., 2020. Building bridges – between the pre-service teachers’ school experiences and the teaching of an educational content. A narrative approach. *J. Curric. Stud.* 52 (5), 608–619. <https://doi.org/10.1080/00220272.2019.1658229>.
- Lambert, D., Biddulph, M., 2014. The dialogic space offered by curriculum making in the process of learning to teach, and the creation of a progressive knowledge led curriculum. *Asia Pac. J. Teach. Educ.* 43 (3), 210–224.

- Lambert, D., Biddulph, M., 2015. The dialogic space offered by curriculum-making in the process of learning to teach, and the creation of a progressive knowledge-led curriculum, Asia-Pac. J. Teach. Educ. 43 (3), 210–224. <https://doi.org/10.1080/1359866X.2014.934197>.
- Lambert, D., 2018. Editorial: teaching as a research-engaged profession: uncovering a blind spot and revealing new possibilities. Lond. Rev. Educ. 16 (3), 357–370. <https://doi.org/10.18546/LRE.16.3.01>.
- Lingard, B., 2011. Policy as numbers: ac/counting for educational research. Aust. Educ. Res. 38, 355–382. <https://doi.org/10.1007/s13384-011-0041-9>.
- Lingard, B., 2013. It is and it isn't: vernacular globalization, educational policy, and restructuring. Routledge, London.
- Looney, A., 2001. Curriculum as policy: some implications of contemporary policy studies for the analysis of curriculum policy, with particular reference to post primary curriculum policy in the Republic of Ireland. Curric. J. 12 (2), 149–162.
- Loughran, J., 2006. Developing a Pedagogy of Teacher Education: Understanding About Teaching and Learning About Teaching. Routledge, London.
- Loughran, J., 2010. Seeking knowledge from teaching teachers: moving beyond stories. Stud. Teach. Educ. 6 (3), 221–226.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a framework for teacher knowledge. Teach. Coll. Rec. 108 (6), 1017–1054.
- National Council for Curriculum and Assessment (NCCA), 2011. Framework for Junior Cycle. NCCA, Dublin.
- Nieveen, N., Kuiper, W., 2012. Balancing curriculum freedom and regulation in The Netherlands. Eur. Educ. Res. J. 11 (3), 357–369. <https://doi.org/10.2304/eej.2012.11.3.357>.
- Osgood, J., 2010. Reconstructing professionalism in ECEC: the case for the 'critically reflective emotional professional'. Early Years 30 (2), 119–133.
- Phelps, A., Lee, C., 2003. The power of practice: what students learn from how we teach. J. Chem. Educ. 80 (7), 829–832.
- Popkewitz, T.S. (Ed.), 2014. The Reason of Schooling: Historicizing Curriculum Studies, Pedagogy, and Teacher Education. Routledge, New York and London.
- Priestley, M., Biesta, G. (Eds.), 2013. Reinventing the Curriculum: New Trends in Curriculum Policy and Practice. Bloomsbury, London.
- Priestley, M., Philippou, S., 2018. Editorial: curriculum making as social practice: complex webs of enactment. Curric. J. 29, 151–158.
- Priestley, M., Philippou, S., 2019. Curriculum is – or should be – at the heart of educational practice. Curric. J. 30 (1), 1–7. <https://doi.org/10.1080/09585176.2019.1598611>.
- Priestley, M., Xenofontos, C., 2020. Curriculum making: key concepts and practices. In: Biddulph, J., Flutter, J. (Eds.), Inspiring Primary Curriculum Design. Taylor and Francis Group, London, pp. 10–23.
- Priestley, M., Edwards, R., Priestley, A., 2012. Teacher agency in curriculum making: agents of change and spaces for manoeuvre. Curric. Inq. 42 (2), 191–214.
- Priestley, M., Biesta, G., Robinson, S., 2015. Teacher Agency: An Ecological Approach. Bloomsbury, London.
- Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), 2021. Curriculum Making in Europe: Policy, Practice Within and Across Diverse Contexts. Emerald publishing, London.
- Priestley, M., 2019. Curriculum: concepts and approaches. Imp. J. Chart. Coll. Teach. Impact, 6: Broad and Balanced Curriculum. Retrieved from <https://impact.chartered.college/article/curriculum-concepts-approaches/>.
- Raffe, D., 2011. Policy Borrowing or Policy Learning? How (Not) to Improve Education Systems. CES Briefing, Edinburgh.
- Rata, E., 2012. The politics of knowledge in education. Br. Educ. Res. J. 38 (1), 103–124.
- Rata, E., 2019. Knowledge-rich teaching: a model of curriculum design coherence. Br. Educ. Res. J. 45 (4), 681–697.
- Schulte, B., 2018. Envisioned and enacted practices: educational policies and the 'politics of use' in schools. J. Curric. Stud. 50 (5), 624–637.
- Schwab, J.J., 1978. The practical: a language for curriculum. In: Westbury, I., Wilkof, N.J. (Eds.), Science, Curriculum, and Liberal Education: Selected Essays. University of Chicago Press, Chicago, IL, pp. 287–321.
- Scottish Executive, 2004. A Curriculum for Excellence. The Curriculum Review Group. Scottish Executive, Edinburgh.
- Seidel, T., Stürmer, K., 2014. Modeling and measuring the structure of professional vision in pre-service teachers. Am. Educ. Res. J. 51 (4), 739–771.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. British J. Educ. Res. 40 (6), 917–936. <https://doi-org.dcu.idm.oclc.org/10.1002/berj.3120>.
- Shulman, L., 1986. Those who understand: knowledge growth in teaching. Educ. Res. 15 (1), 4–14.
- Shulman, L.S., Shulman, J.H., 2004. How and what teachers learn: a shifting perspective. J. Curric. Stud. 36 (2), 257–271. <https://doi.org/10.1080/0022027032000148298>.
- Slattery, P., 2013. Curriculum Development in the Postmodern Era, third ed. Routledge, New York, NY.
- Smyth, E., McCoy, S., 2011. Improving Second-Level Education: Using Evidence for Policy Development. Renewal Series Paper 5. ESRI Web Published, Dublin.
- Stenhouse, L., 1975. An Introduction to Curriculum Research and Development. Heineman, London.
- Tahirsylaj, A., Wahlström, N., 2019. Role of transnational and national education policies in realisation of critical thinking: the cases of Sweden and Kosovo. Curric. J. 30 (4), 484–503. <https://doi.org/10.1080/09585176.2019.1615523>.
- Teo, T.W., 2012. Building Potemkin schools: science curriculum reform in a STEM school. J. Curric. Stud. 44 (5), 659–678.
- Thijs, A., van den Akker, J. (Eds.), 2009. Curriculum in Development. SLO–Netherlands Institute for Curriculum Development, Enschede. Retrieved from. <http://www.slo.nl/downloads/2009/curriculum-in-development.pdf>.
- van den Akker, J., 2003. Curriculum perspectives: an introduction. In: van den Akker, J., Kuiper, W., Hameyer, U. (Eds.), Curriculum landscapes and trends. Kluwer Academic Publishers, Dordrecht, pp. 1–10.
- Wahlström, N., 2016. A third wave of European education policy: transnational and national conceptions of knowledge in Swedish curricula. Eur. Educ. Res. J. 15 (3), 298–313.
- Wallace, C., Priestley, M., 2017. Secondary science teachers as curriculum makers: mapping and designing Scotland's new curriculum for excellence. J. Res. Sci. Teach. 54 (3), 324–349.
- Winter, C., 2017. Curriculum policy reform in an era of technical accountability: 'fixing' curriculum, teachers and students in English schools. J. Curric. Stud. 49 (1), 55–74. <https://doi.org/10.1080/00220272.2016.1205138>.
- Wheelahan, L., 2010. Why Knowledge Matters in Curriculum: A Social Realist Argument. Routledge, London.
- Young, M., 2008. Bringing Knowledge Back in. Routledge, London.
- Young, M., 2010. The future of education in a knowledge society: the radical case for a subject- based curriculum. J. Pac. Cir. Consort. Educ. 22 (1), 21–32.

Curriculum making and higher education: an activity-theoretical perspective

Johanna Annala, Tampere University, Tampere, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	163
Curriculum as an object of activity	164
Tools in curriculum making	166
Rules in curriculum making	167
Division of labor in curriculum making	168
Community in curriculum making	169
Conclusion	170
Acknowledgments	171
References	171

Introduction

In the present article, curriculum as an object of activity is the framework through which curriculum making is explored from the perspective of academics. Curriculum researchers have repeatedly agreed that the concept of curriculum has various interpretations in higher education research and practice (e.g., [Annala et al., 2016](#); [Coate, 2009](#); [Fraser and Bosanquet, 2006](#); [Knight, 2001](#); [Krause, 2020](#)). If the understanding of curriculum is not shared, actors in one specific curriculum change context may see the object of the activity from contradictory perspectives, which may result in contradictory actions and outcomes ([Annala and Mäkinen, 2016](#)). Depending on the scale of the curriculum making and the role of the actors, the consequences may be different.

Curriculum making is understood here as an intentional, dynamic and contradictory process that is always situated in a certain cultural, historical, societal and political environment. Regarding the contemporary environment in curriculum making in higher education, [Krause \(2020\)](#) has identified three “vectors of change” that are shaping higher education curricula: national and international political, economic and labor market forces; universalization; and the influence of technology. The role of national policy and regulatory frameworks varies; however, today in OECD countries especially, student employability, quality standards and institutional accountability have a noteworthy impact on shaping curricula ([Krause, 2020](#)). For example, the harmonization of European higher education degrees through the Bologna process ([European Commission/EACEA/Eurydice, 2018](#)) was executed through higher education policy, but in the end, it manifested in different ways in curriculum making in the local discipline-specific work contexts of academics ([Karseth, 2006](#); [Wihlborg and Teelken, 2014](#)).

The second vector of change is universalization, which refers to the widening access to university education ([Krause, 2020](#)). The diversity of the student body, with different knowledge, cultures and experiences, shapes the curriculum ([Krause, 2020](#)). For example, in South Africa, structural injustices and the equity of access and outcomes are topical in curriculum transformation, as the deficit view has predominated instead of recognition ([Luckett and Shay, 2020](#)). The question of knowledge is also significant here. Originally, [Gibbons et al. \(1994\)](#) defined mode 1 and mode 2 knowledge, mode 1 referring to more traditional forms of knowledge building and research in universities, which is hierarchical and specialized, whereas mode 2 knowledge is associated with industry, innovation and government, emphasizing knowledge development in interdisciplinary collaboration and the applicability and usefulness of knowledge ([Yates et al., 2017: 19](#); [Young, 2013](#)). The widened access attracts students, for example those with a nonacademic family background and who may be more engaged with working-life-oriented and integrated curricular knowledge (mode 2). Yet all the students in higher education should have access to powerful forms of knowledge (mode 1), because access to abstract theoretical knowledge is a precondition to joining the important discussions in society and to democracy ([Shay, 2013](#); [Wheelahan, 2010](#)).

All fields draw on some mix of conceptual knowledge and know-how of various kinds, but some disciplinary fields are more vocational and some more “truth-seeking,” as the higher education institutions also are. Nonetheless, in curriculum trends, the balance has moved from mode 1 knowledge toward mode 2 knowledge ([Lindén et al., 2017](#); [Yates et al., 2017: 20](#)). This is visible in the third vector of change, the advancement in technology and the simultaneous emphasis on industry-aligned learning outcomes ([Krause, 2020](#)). This has emerged as the unbundling of formal curriculum and the rise of microcredentials, digital badges and nano-units that shape the curriculum toward a “buffet table” service without curriculum coherence ([Cliff et al., 2020](#); [Krause, 2020](#)). Thus, the second and third “vectors of change” are linked to the first one, the political, economic and labor market forces around higher education curriculum.

These trends are very different from the traditional tasks and cultural norms of higher education curriculum, described as aiming to provide opportunities for purposeful, intended experiences of complex learning with indeterminate and non-linear nature that are “not easily reducible to precise statements predicting what the outcomes will be” ([Knight, 2001](#)). Wang also states that higher education should not be a pre-given way on a map in which the student is prepared to pursue professional competence and employability; instead, the goal of higher education is to educate a student to be “a cartographer to create his or her new life” (2015: 1558) and to invite students to “explore strangeness, thereby promoting their curiosity and assisting them in accommodating a future creatively” (2015: 1557).

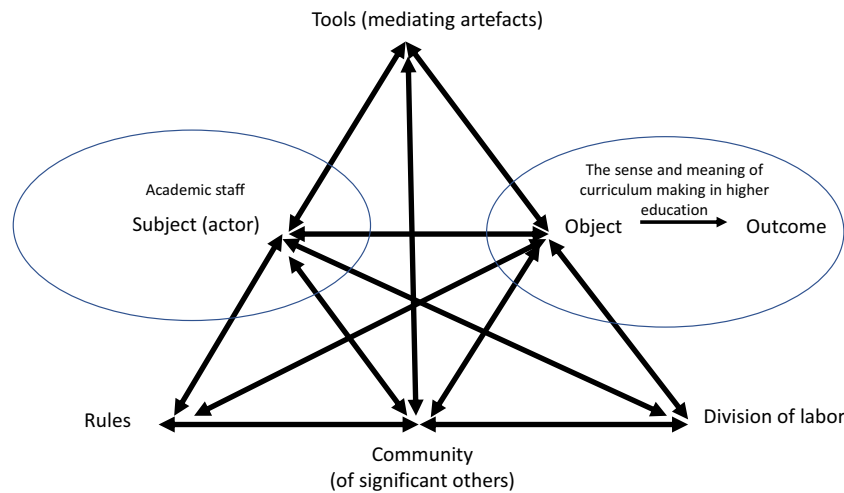


Fig. 1 Analyzing curriculum making as an activity system. Source: Engeström, 1987: 78.

Being aware of this contemporary environment of higher education curriculum, curriculum making is approached here as a multi-voiced and multilayered activity system (Engeström, 2000; Foot, 2014). The aim is to analyze curriculum making in higher education with the help of concepts from activity theory (Engeström, 1987). Activity theory is used as a broad framework by which it is possible to pay attention to micro- and macrolevel elements in curriculum making. Cultural-historical activity theory has different emphases and approaches as well as different generations. This paper is built on the concepts of the activity system modeled by Yrjö Engeström (1987: 78) as illustrated in Fig. 1. The activity system has six components, of which the first three are more visible: a subject (actor), an object and tools. Tools are employed by the actors to work toward the object of action and to achieve the desired outcome (Foot, 2014). The three other components—rules, community and division of labor—are less visible social mediators of activity (Engeström, 1999).

Inner contradictions between the six components create constant transformation and reconstruction and are a source of development and change (Engeström, 1999; Foot, 2014). Correspondingly, curriculum making is often depicted using terms like “development,” “reform” and “change” in the strategic discourse (Krause, 2020), but terms like “struggle” and “bargain” appear when academics describe their experience (Annala et al., 2020).

Activity in the system requires subjects (actors), referring to “the individual or subgroup whose position and point of view are chosen as the perspective of the analysis” (Engeström and Sannino, 2020: 5). Traditional curriculum making is often led by experts working inside the university system, but participatory approaches have broadened the co-ownership toward different stakeholders, such as educational developers, managers, students, representatives of the world of work and policy-makers (Alexander and Hjortso, 2019). In the following, the perspective of academics as subjects (actors) in the system is emphasized. During the last ten years, in our research group, we have completed several qualitative studies on curriculum from the perspective of academics. These studies are used here to illustrate curriculum making as an activity system.

Curriculum as an object of activity

The concept of “object of activity” is regarded to have a key role in the activity system (Leont’ev, 1978). Engeström and Kerosuo (2007: 337) state that “the object embodies the meaning, the motive and the purpose of a collective activity system.” According to Foot (2014), the object is comprised of three facets: a thing-to-be-acted-upon (e.g., curriculum), an objectified motive (e.g., developing or reforming the curriculum) and a desired outcome (e.g., superior quality of teaching and learning). Objects of activity have been described as a “horizon of possibility” (Blackler, 2009: 29) and as a “raw material or problem space at which the activity is directed” (Engeström and Sannino, 2020: 5).

When thinking of curriculum making, the desired outcome and “horizon of possibility” varies among academics. Based on interviews (N = 45) of teachers at a research university and university of applied sciences, representing a range of different disciplinary fields, we identified nine different objectives for curriculum development (Mäkinen and Annala, 2010a, Fig. 2). Our analytical framework relied on a schema developed by Barnett and Coate (2005), where three curricular objectives concern students’ knowing, acting and being. Knowing refers to the core knowledge of the subject or disciplinary field in question, acting refers to the kind of actions and practices, such as “acting like a dentist” or “thinking like a sociologist,” which the student is expected to acquire. Being has to do with the student’s identity and personality, “who I am” and “what kind of relationship I have with the knowledge and knowing” (Barnett and Coate, 2005.) In our analysis, we also used Bernstein’s (1996) concepts of introjection and projection, described as starting points of university education. Introjection refers to the construction of curriculum on the basis of internal disciplinary interests (e.g., the subject taught), whereas curriculum development on the basis of external demands (e.g., competence

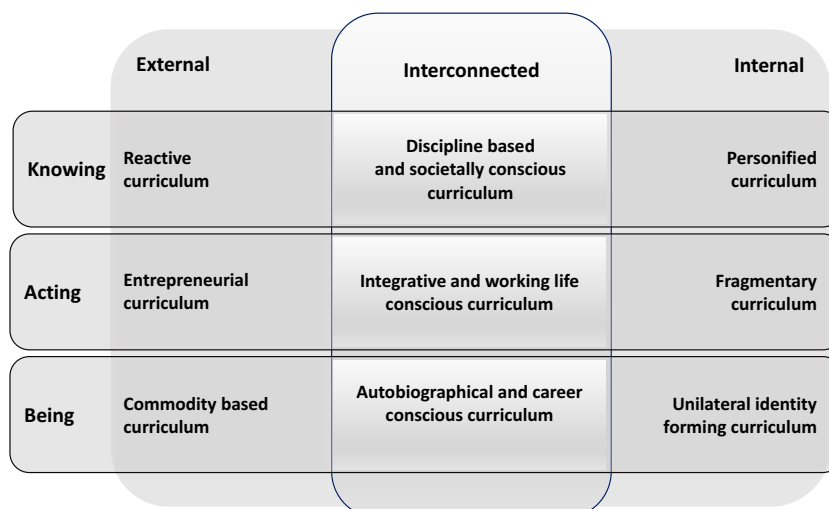


Fig. 2 Comprehensive curriculum framework Source: Mäkinen and Annala, 2010a: 12.

demand in working life) is called projection (Bernstein, 1996). We found external and internal objectives for curriculum development but also interconnected views between them.

At one extreme, the aim of curriculum emphasizes the external demands; it is reaction to the contemporary knowledge requirements of the society and aims to produce competences relevant to employment and to make a student “a fit product” for the market. At the other extreme, curriculum is made for the university and teachers themselves, to fit and preserve their areas of interest (cf. Shay, 2011), with a fragmentary and subject-specific approach without connection to the world outside. If one teacher wants to make a curriculum so that it retains teaching positions (requirement to have “my course” in the curriculum) and another aims to produce knowledge relevant to working life (giving the authority to outside stakeholders), the contradictions emerge (Mäkinen and Annala, 2010a.) These oppositions reflect, on the one hand, neoconservative traditionalism, emphasizing the given body of knowledge that is transmitted through educational institutions, and on the other hand, technical-instrumentalistic curriculum policy, where curriculum is seen as a way to direct the students toward the needs of the economy (Young, 2008). The oppositions also make visible the layers of higher education history influencing the present-day approaches (cf. Engeström and Sannino, 2020).

If contradictions in the activity system are regarded as a source of development and learning, the interconnected view could also be described as such. Remaining true to the core of a given discipline, but being conscious of what the world is outside the university, was taken as an opportunity to make a difference to the world outside and not just to follow the market voices (Annala and Mäkinen, 2016; Mäkinen and Annala, 2010a). In that case, mode 1 and mode 2 knowledge were combined (cf. Gibbons et al., 1994). Even though this could be a happy end to a story, it is important to note that only some of the informants shared the interconnected view and did not necessarily collaborate, as they were from different faculties. The object of activity remains various in the actual curriculum making process if it is not consciously shared and made meaningful inside the local community (Annala et al., 2021).

As the above-discussed study was a local case study representing teachers, the practitioners, in curriculum making, I will next look at the curriculum researchers’ perspective on what the object of activity in curriculum making is, as it offers a more global view of the object of activity. A few years ago, we did a systematic review of how curriculum is approached in research articles in higher education journals (Annala et al., 2016). Following curriculum theoretical ideas, we used as our reference point the traditional distinction of understanding curriculum as syllabus, product, process and praxis (Grundy, 1987; Kelly, 2009/1977). When we looked more closely, the 62 analyzed articles represented fundamental differences in their orientation toward two core themes: knowledge and ownership. This analysis brought to the fore the power relations and the nature of the outcomes, as if students were empty spaces to be filled or as if they had ownership in the process. What we found was that curriculum was conceptualized as (1) controlling of contents, (2) producing competences, (3) negotiating of potentials and (4) empowerment. If curriculum is seen from the perspective of control (e.g., mapping the knowledge and skill gaps defined by experts from industry) or from the perspective of empowerment (e.g., creating culturally sensitive curriculum or paying attention to students’ access to powerful forms of knowledge), the object of the activity is very different.

Sannino and Engeström (2018) state that the object of activity is difficult to define for the participants, as it cannot be reduced to conscious goals and actions. An important distinction here is the difference between action and activity. As the actions are directed toward short-term goals, at a given moment, in each action, the object of activity evolves historically and culturally, giving space to interpretation, personal sense making and societal transformation (Engeström, 1999; Sannino and Engeström, 2018). The activity approach is similar to the “curriculum as a process” approach (Fraser and Bosanquet, 2006; Knight, 2001; Stenhouse, 1975) and to “curriculum as complicated conversation” (Pinar, 2004), where the complexity of curriculum making is present. The object of a collective activity includes multiple alternative goals and actions that are turned into outcomes. The object is not stable, because

after the intermediate goal is reached, it escapes like the horizon does, and new goals and actions need to be negotiated and reconstructed. However, in curriculum development projects, often led by people without educational backgrounds (Lindén et al., 2017) or without training as leaders (Knight, 2001), the short-term goals, actions and technological solutions may be emphasized, leading toward “curriculum as product” or “control over contents” approaches (cf. Fraser and Bosanquet, 2006; Grundy, 1987). These become more visible when elaborating the tools, instruments and mediating artifacts in the curriculum making.

Tools in curriculum making

In the activity system, the tools can be either material or conceptual artifacts. In every activity system, actors use already existing tools but also create new tools with which to “engage, enact and pursue the object of their activity” (Foot, 2014: 336). Tools employed in an activity system reveal something of the actor’s understanding of the object of activity, but they also reflect the cultural-historical context.

Foot (2014) characterizes tools of different kinds with the help of Wartofsky’s (1979) three-tiered hierarchy of tools. The primary tools are quite simple, for example, a protocol for what curriculum should include. Ralph Tyler’s (1949) principles in curriculum design have had a major influence on curriculum making from primary to tertiary-level education (Flinders and Thornton, 2004), and these could be characterized as primary tools. The principles include defining the “goals,” “contents,” “methods” and “assessment,” which can be found as boxes to be filled in, in electronic curriculum systems around the world, and which direct curriculum thinking toward the “curriculum as product” approach (Grundy, 1987).

The secondary tools are representations of primary tools and mediate how actors employ them. In curriculum making, they may be, for example, models or explanatory frameworks. In higher education, the widely used model of constructive alignment by John Biggs and Catherine Tang (2011) could be characterized as such. Even though this model reflects Tylerian elements (Lindén et al., 2017), the emphasis here is in the shift from a teacher- and teaching-focused approach to a student- and learning-centered approach (Bovill and Woolmer, 2019). For academics writing curriculum texts, this means that instead of a list of contents to be taught, teachers should express what the intended learning outcomes are and what a student is expected to learn and achieve during the course. The language forms a tool, a repertoire with which to talk about the object of the activity, and it constructs the reality we are living. The role of language is different in the social sciences and natural sciences, and the conceptual world in curriculum making is different for academics, professional support staff, managers and politicians. In our study, one professor characterized this in the following way:

Some people talked about how we should only use a lot of verbs, you know, they kind of talked down what the whole thing really should be about, but I still can’t quite figure it out myself. Either it is such a big linguistic or mental twist that it just takes time to understand it, or then again, maybe I was just too exhausted to perceive anything.

Source: Annala and Mäkinen, 2016: 13.

Tertiary tools are abstractions, for example epistemological paradigms, that influence the overall direction of the activity system. They may be connected to ideologies and identities and are needed to understand the activity system (Engeström, 1990: 174). This was visible in the previous citation, expressing a need for an explanation and motivation behind the practical actions. Following this need, Fig. 2 was used as a tertiary tool in the university in question, to put on the table the different ideologies and aims and to identify and share what the primary object of activity in their curriculum making was.

Curriculum theorists are most often interested in the tertiary tools, that is, what the ideologies and driving forces are behind curriculum initiatives. For example, the roots of Tyler’s (1949) and Biggs and Tang’s (2011) approaches can be traced to Bobbit (1972/1918), who applied Taylor’s (2003/1911) scientific management in factory production to educational management and planning. At that time, one hundred years ago, curriculum was seen as a way to impact people through education and to respond to the challenges of society. This view is surprisingly similar to the views today in economics-led ideologies and, for example, in the OECD and EU initiatives to boost the economy through higher education and competences relevant to industry (Lindén et al., 2017).

As the more traditional curriculum models rely on educational research, the recent curriculum instruments, such as microcredentials or curriculum mapping, do not have similar roots; instead, “pedagogical jargon” is seen as a barrier (Rawle et al., 2017). These instruments reflect technological and commercial interests (Willis et al., 2016) and are most often used in the natural sciences, especially in engineering and medicine (Rawle et al., 2017). Examples of micro-modules can be found, for example, in the ECIU university network, an EU-funded initiative that “creates a groundbreaking and innovative educational model on a European scale” (ECUI, 2020). Even though marketed as improving educational quality, instead of supporting complex learning, the emphasis is on control and standardization (Wang, 2015) and on observable skills (Carey and Stefaniak, 2018). For example, the technical tools around microcredentials, also known as digital badges, are depicted as a student-centered, working-life-relevant, authentic and democratic way of offering university studies as a new ecosystem (Willis et al., 2016). These tools are described as being in “the intersection of education and motivation, business and data, and badge culture and technology,” and they “stand at the forefront of current credentialing systems” (Willis et al., 2016: 31). As such, they feature universalization, advancement in technology and an emphasis on industry-aligned learning outcomes (cf. Krause, 2020).

Tools of this kind are presented as though they were separate from the other components in the activity system. If academics tune up only their own courses with a focus on the controllable process, the consequences in the system might be unexpected. Besides producing a lack of curriculum coherence and a patchy growth of student expertise, it also changes the role of the university institution. Digital badges are characterized as a new currency of educational power, as they challenge higher educational “institutions privileging exclusivity of education, credentials, and evidence of learning” (Willis et al., 2016: 24). Educational technology companies are involved as stakeholders. This suggests that marketized curriculum space affects not only the curriculum and nature of learning in higher education but also aims to shake the credibility of the university institution (see also Barnett, 2000; Cliff et al., 2020; Krause, 2020).

Rules in curriculum making

According to Sannino and Engeström (2018: 45), “rules refer to the explicit and implicit regulations, norms, conventions and standards that constrain actions within the activity system.” Rules in curriculum making have their formal basis in regulations and standards that originated from educational policy, although higher education has more autonomy compared to other educational levels. As already described in discussing Krause’s (2020) three “vectors of change,” the economic and labor market forces have a growing impact on higher education curriculum from the outside, pushed by the policy requirement to strengthen the university–society relationship (see also Karseth and Solbrekke, 2016; Yates et al., 2017). Also, Priestley (2011) has identified three features in curriculum policy: the growing importance of national qualification frameworks, the shift (or return) to the emphasis on learning outcomes and a move from a subject and disciplinary emphasis toward a more generic curriculum. Yet cultures and histories in the higher education institutions shape the norms and conventions, resulting in the ways in which the rules are interpreted and enacted.

When subjects perceive tools as requirements or demands by those with power, instead of being useful for a subject to engage in the object of activity, tools may be employed as rules (Engeström, 1990: 90; Foot, 2014). As an example, curriculum mapping represents a tool by which to think of how to ensure that certain topics, skills and contents are included in curricula. Curriculum mapping focuses on student learning and encourages a communicative culture among teachers, as they need to collaborate (Rawle et al., 2017; Wang, 2015). In curriculum mapping, special attention is given to student employability and generic/transversal skills. Even though curriculum researchers have underlined the fact that generic skills cannot be isolated from knowledge (e.g. Young, 2011), it is increasingly required to map those skills separately in curriculum making. Online curriculum maps implemented as matrices or forms to be filled out may constrain the curriculum offerings to keep them on track and under control, primarily to serve the work-readiness of students (Wang, 2015). The instrumental way of using tools as rules does not necessarily engage academics who see the object of the activity more broadly. This is visible in the following professor’s comment:

We were given such instructions or restrictions that did not improve the work here. Rather, we were required to struggle against those to prevent or alleviate them so that we could have a territory where we could work reasonably.

Source: Annala et al., 2021.

Another example of rules in curriculum making is how the intended learning outcomes should be expressed in a written text. In Finland, there were guidelines suggesting that the verb “understand” was not allowed, as it was too ambiguous to be evaluated as a learning outcome (see also Biggs and Tang, 2011; Cedefop, 2017). However, some academics disagreed with this notion, as “understanding” was strongly related to complex learning in higher education, for example in the social sciences. This was also our motive for studying the “competence” language in EU policy guidelines in the translations from English to Finnish and studying how academics themselves understood the different concepts in their contexts. What we found was that there is a linguistic inconsistency between different concepts and translations (Mäkinen and Annala, 2010b). Because of this, it was important to see how academics as actors had their own negotiations of meaning: what does a certain concept mean in our context and in our language, and how do we interpret those in our curriculum practice, in our higher education?

We summarized our findings with two paradigms: linear competence and dynamic competency (Mäkinen and Annala, 2010b, 2012). The linear competence paradigm focuses on standardized learning objectives and their assessment. It is manifested in the mainstream guidelines, tools and rules for curriculum making. Dynamic competency relates to student potential and unforeseen learning processes. It is specific to university education, if we see the university as a site of the highest knowledge. In curriculum making, both paradigms are possible and needed, with different emphases in different disciplinary areas and in different phases of studies. Following this, the actors in curriculum making can define some but not all of the specific learning outcomes in detail if we expect our graduates to be future makers, to use their potential and find solutions to the problems the actors themselves (i.e., professors) have not yet been able to solve. Future graduates should know more than the professors of today, which requires a move toward curriculum as empowerment (cf. Annala et al., 2016). Based on our work, our notion is that if competence-based language and models are reflected along with the broader goals and ideas of the university, and if they negotiate a shared meaning between the members of academic community, the tools and rules in curriculum making start to make sense for academics.

The rules reveal how the object of activity is interpreted. People, faculties or institutions follow the rules differently, as curriculum making offers space for interpretations and sensemaking (Annala and Mäkinen, 2016; Sannino and Engeström, 2018). The previous examples concern practical and visible rules in curriculum making, but there are also hidden norms and conventions concerning curriculum making as a part of academic work, including the power relations in the academic community. The rules include some background assumptions about the right and the wrong way of acting, and it is important to ask who wrote the rules and why they should be accepted to govern academics' way of acting (cf. Taylor, 2009). The relations between the subject and community are mediated through the rules that regulate subjects' actions toward an object and relations with other participants (Foot, 2014; Engeström, 1987). This led us to explore the division of labor in the system and academics' roles there.

Division of labor in curriculum making

With regard to the activity system, the question of the division of labor in making curriculum was analyzed earlier as the internal relationships in higher education between academics, professional staff, and leaders, such as the unit chairs responsible for curriculum (Dracup et al., 2020; Cliff et al., 2020). The division of labor was also examined to include the external stakeholders, such as decision makers outside the unit, those in the private sector, NGOs and government (Alexander and Hjortso, 2019). The division of labor concerns the horizontal division of tasks, understood as what is being done by whom toward the object, and the vertical division of power and status, including power, positions, access to resources and rewards (Sannino and Engeström, 2018: 45). Next, I will look at the horizontal division of labor, in the light of the case studies of Kate O'Connor (2020) in the Australian context, and then at the vertical division of labor from the Finnish perspective (Annala et al., 2020). These studies do not reflect the division of labor as a purely internal or external question but contribute to exploring the inner contradictions in the activity system and its relation to the object of activity.

Kate O'Connor's (2020) study concerns the relationship between constructivism, knowledge and curriculum, but it simultaneously points out the role of academics in curriculum making with regard to online modules. In those cases, the role of the academics was to produce subject content and knowledge, but after that, to leave the teaching to external tutors or casual teachers or to students studying collaboratively by themselves online. The ownership and implementation of curriculum moved from academics to online experts and casual teachers, appearing as a cost-effective way to implement curriculum. The core concern was whether knowledge about teaching practice could be isolated from the knowledge about the content taught. The online modules were put forward as student-centered and constructivist ways of supporting active learning, but the implementation positioned curriculum knowledge as fixed and stable rather than as complex and negotiated with different cohorts of students (O'Connor, 2020). The role of academics with a deep understanding of the topic was not acknowledged; instead, the lecturers were blamed for overloading their subjects with too complex an approach. The leaders asked the lecturers to reduce the content, despite the nature of the subject, to be fixed and stable long term, in order not to waste time and in order to fit the new online formula.

This is an example of the consequences of unbundling the curriculum making from the teaching activities (Cliff et al., 2020; Tucker and Neely, 2010). Ryan et al., (2017) see that a strong division of labor in academia is a risk to university teaching quality and reputation while being the university's main source of income. The formal pressures on core academic staff to increase their research performance and ignore teaching threatens the "idea of a university" as a "social institution dedicated to both the creation and dissemination of knowledge in the public interest" (Ryan et al., 2017: 69). Simultaneously, as some stakeholders emphasize the cost-effectiveness of curriculum, others aim to keep its profile distinct from other educational institutions, emphasizing that teaching and research are closely intertwined in curriculum making (Annala and Mäkinen, 2011; Visser-Winjeveen et al., 2010). In our study, one professor described this as follows:

Vocational, professional and scientific goals are contradictory. It should not be denied. We should defend the status of scientific approach and research. But most of all, we should think how we could take advantage of these tensions pedagogically.

Source: Annala and Mäkinen, 2011: 13.

The vertical division of labor with respect to academics' power, positions and access to resources and rewards became visible in our study on agency and structure in curriculum making processes (Annala et al., 2020). We applied Bourdieu's field theory and looked at curriculum making as a specific field of struggle. When looking at the division of labor at the department level, the diversity of internal participants in curriculum making, including students, academic developers, casual teachers and doctoral researchers, was not always seen as a resource but was sometimes seen as a threat to the professoriate (Annala et al., 2020). This was especially the case when curriculum and teaching were seen as separate from research. Instead, the enhancement of the research-teaching nexus and shared conceptual tools had an important role in engaging academics in curriculum making and in regarding it as a meaningful part of academic work. This implies that curriculum making would benefit from being framed more holistically, not as instrumental and short-term action and not as administration and teaching-related activity but as teaching and research-related activity.

Yet academics' narratives of curriculum making demonstrated power relations when there appeared features of a game with winners, losers and survivors along with stable narratives featuring development-orientation, autonomy and opposition (Annala et al., 2020). Curriculum making had hidden rules and raised questions of who had agency and power and how to

have them in curriculum making. In a university-wide curriculum change, the status of different disciplines played a role. For some, the strategic guidelines and structural frameworks created an activity system with rules and tools that improved their position in the field, whereas for some other disciplinary fields, the consequences were the opposite. Ylijoki and Ursin (2013) also state that specific areas in the natural sciences, technology and medicine are often so valued by the university and society that they can shape the rules in academic practice. The question is whether these disciplinary fields place more value on tools of a certain type in curriculum making and how this affects curriculum making more broadly within different communities, institutions and universities (see also Barnett, 2000).

A vertical division of labor, emerging as broader power structures, seems rather influential in curriculum making in academia. Power relations can promote a collective capacity to innovate and create new practices, but they may also restrict collaboration and decision-making in the activity system (Haapasaari, 2020). Following the premises of activity theory, tensions and contradictions here could be transformed toward innovative ideas and solutions, through “incremental developmental steps, radical changes in tools, processes and practices, or system level changes that transform the entire course of action” (Haapasaari, 2020: 4). These all require collaboration in the academic communities as well as analysis of their curriculum making practices.

Community in curriculum making

A community of significant others consists of individuals and subgroups who share with the subject an interest in and involvement with the same object (Sannino and Engeström, 2018: 45). When thinking of the activity system, relations between the subject and the community are mediated through the rules that regulate subjects’ actions toward an object, their relations with other participants and the division of labor (Engeström, 1987). As such, features of the community are strongly linked to the earlier sections of this paper. When looking at the activity system from the perspective of academics, the way curriculum making is implemented in the community is important.

Professional staff and academic developers are often the mediators of the rules between the policy guidelines, management and academics. Handal et al. (2014) explored academic developers’ experiences between internal and external pressures in Norway. They analyzed the actions and arguments of academic developers in curriculum making when implementing the National Qualification Framework (NQF), part of the Bologna process. The academic developers struggled between compliance and resistance, as there were tensions between decisions and guidelines coming from outside, from the Ministry and from the complex nature of higher education where they acted. They had better knowledge that the directions for formulating learning outcomes did not fit all the disciplinary fields similarly, but the Ministry had power as a decision-maker superior to the universities. Academic developers had professional accountability to induct academics into practices and procedures defined by politicians who did not have better knowledge of the higher education curriculum, but simultaneously, the developers had the professional responsibility to facilitate more academically and theoretically driven pedagogical practices (Handal et al., 2014). Similarly to the notions of O’Connor (2020), Handal et al. (2014) state that the core concern is that knowledge about teaching practice cannot be isolated from the knowledge about the content taught.

Handal et al.’s (2014) study revealed that not only curriculum but academics as well are considered as objects of activity, to be inducted into certain ideas and practices. This signals that from the perspective of policy, academics are either reluctant to join curriculum making processes or they may approach them in an unfavorable way, with tools or rules that do not fit the dominant agenda. This represents “resource power,” described by Haapasaari (2020) as one layer of power in the activity system, that is, power over others and the control of resources. Another layer of power described by Haapasaari (2020) is “process power”, which refers to organizational decision-making processes, and a third layer is “meaning power”, which manifests in dominant perceptions in an organization. Following all these layers, close attention to all the elements in the activity system is needed. What is the community of significant others who are involved in curriculum making? Do the academics, casual teachers and professional developers have ownership in the curriculum making and implementation? What is the role of external stakeholders? And how and where are the final decisions made? Academic pedagogical leadership in the curriculum process in communities was a recurrent topic that emerged in our studies, exemplified in the following:

There should’ve been someone kind of coordinating the process [–] there we had this fine theoretical knowledge and a vision, a concept, the idea of what we were trying to do. So that one doesn’t just go right ahead to the implementation, and start talking about courses one wants to get rid of, or how the basic courses should be changed, instead first we should have had this phase with absolutely no comments allowed regarding any of the courses. We should have talked about what this or that means, what are the premises we can commit ourselves to.

Source: Annala and Mäkinen, 2016: 11

This notion brings to the fore the question of whether academics themselves are subjects in curriculum making and whether they have genuine access to the negotiation of meanings of the object of activity in the academic community—do they have meaning power (cf. Haapasaari, 2020)? Results from our study indicate that in many cases, the curriculum change deeply touched the academics’ expertise areas, pedagogical philosophies and research fields, thereby resulting in commitment to the curriculum making process (Annala et al., 2021; see also Priestley et al., 2012). The academics’ engagement in curriculum making was also connected to the changes in structure; for example, academics’ experience was that university-wide curriculum change favored more hierarchical

and hard qualities than dialogical, collegial and pedagogical. The hierarchical qualities are an example of Louvel's (2013) characterization of curriculum making as an "engineering type" of activity where a plan is followed to meet the expected optimal solution.

As an alternative, Louvel (2013) characterizes curriculum making as a creative "bricolage type" of activity. She refers to how individuals and groups create something new by interpreting the rules creatively, using the existing rules as a resource but stretching them to fit their reality. It is clear that academics do not simply cope as passive receptacles of suggested changes in curriculum making; instead, they adopt and reject, negotiate and contest (Brøgger, 2014). Similarly, we found that the institutional, community and individual layers enabled and impeded implementation of academics' agency in curriculum making (Annala et al., 2021). The direct work community played an especially significant role in engagement in curriculum making, despite how large the change or reform was. The role of community in the activity system of curriculum making could be characterized as follows: "When people are provided with opportunities, forums and tools for participating, they take agentive actions and initiate innovative solutions to collaboratively develop joint activity" (Haapasaari, 2020). If there is room for collaboration and a capacity to work with others, it is also possible to expand the object of activity.

Conclusion

In this article, the aim was to analyze curriculum making in higher education, with the help of concepts from the activity theory, with specific focus on curriculum as an object of activity. The emerged contradictions, often dichotomic, reveal that not only do universities, disciplinary fields, policy-makers, students and other stakeholders see the object of activity differently, but academics also do not have a shared view. The tensions between different disciplinary fields are present, especially when the rules that are aligned with the histories and cultures of some fields are pushed by the education policy to other fields, faculties and institutions of different types. The three vectors of change described by Krause (2020) appeared clearly in the analysis of curriculum making.

Taking fully into account the various perspectives, the contradictions and dynamics would require applying third-generation activity theory (Engeström, 2001), where two or more activity systems are brought together through a common object. Doing so would contribute to an analysis of how different actors are positioned in the activity systems and how they see the tools, rules and other components. Yet there is a risk that we may focus exclusively on the subjects and actors (Engeström et al., 2003), while the key question remains: is there a common object of activity, and how it is understood? Thus, it is necessary to put on the table the object of activity, what this is all about. My key argument is that a necessary glue in curriculum making is to focus on the object of activity (cf. Engeström et al., 2003). This is also the message from activity theoretical thinking: It is important to take objects of activity seriously (Blackler, 2009; Engeström et al., 2003).

It is also important to pay attention to tools, as they may simultaneously have enabling and constraining elements (Kuutti, 1996), that is, how they are used or could be used by academics and whether they fit all the disciplinary fields similarly. What is problematic here is that when curriculum scholars have focused on tertiary tools, i.e. ideological critiques and philosophical questions, they have offered little if any support for practical tools in curriculum making (see also Priestley, 2011; Young, 2013). Then again, academics, managers and politicians who are unfamiliar with educational sciences easily prioritize questions at hand in curriculum making without a need of or interest in historical, political or societal analysis (Lindén et al., 2017; Young, 2013). The primary and secondary tools are brought to the fore as they make the curriculum process concrete, effective and manageable. The cultural-historical background of curriculum initiatives, especially if they cross national and disciplinary borders, needs transparency. For this reason, I argue that tertiary tools are strongly needed in curriculum making, as they contribute to understanding the rules, motives and ideologies behind the curriculum protocols and models, and they help academics and other stakeholders evaluate the possible outcomes and consequences of the chosen actions in curriculum making, not only locally but globally.

One of the key contradictions in the activity system is the desired outcome and the use value and exchange value of the object (Engeström, 1999). The superior quality of teaching and learning was suggested as one possible outcome in curriculum making. Yet the analysis showed that unintended outcomes and consequences may appear for academics, university education and university institution if some stakeholders dominate in curriculum making. Following this, a fundamental question is what is of value in the higher education curriculum and is it possible to find interconnected views that bridge the different interests. Paul Ashwin (2020) summarizes this issue as follows:

The educational purpose of a university education is not to prepare someone for their role in the future workforce. Rather, the educational purpose of a higher education is to bring students into a transformational relationship to knowledge that changes their sense of who they are and what they can do in the world. This undoubtedly prepares them to make a contribution to society, including through their contribution to the labour market but this is a by-product of the central educational purpose of a higher education rather than its driving force.

Source: Ashwin, 2020: 3

Those who understand curriculum making more broadly are needed in the academic communities and as decision makers in curriculum changes; they may be educational developers or academics themselves if they are given access to be Subjects who define the rules and tools that make sense for academic purposes. When object and motive are reconceptualized, a radically wider horizon opens up (Engeström et al., 2003). This could thus enable the academic community to see curriculum making possibilities as a process to empower and to make changes to society, not just to follow what is there but to see the horizon of possibility.

Acknowledgments

I would like to express my deep gratitude to my fellow researchers, Professor Marita Mäkinen and university lecturer Jyri Lindén, for partnership in research on higher education curriculum during the last decade. I am also grateful to Professor Annalisa Sannino for her comments on this article.

References

- Alexander, I.K., Hjortso, C.N., 2019. Sources of complexity in participatory curriculum development: an activity system and stakeholder analysis approach to the analyses of tensions and contradictions. *High Educ.* 77 (2), 301–322. <https://doi.org/10.1007/s10734-018-0274-x>.
- Annala, J., Lindén, J., Mäkinen, M., Henriksson, J., 2021. Understanding academic agency in curriculum change in higher education. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2021.1881772>.
- Annala, J., Lindén, J., Mäkinen, M., 2016. Curriculum in higher education research. In: Case, J., Huisman, J. (Eds.), *Researching Higher Education. International Perspectives on Theory, Policy and Practice*. Society for Research into Higher Education & Routledge, London, pp. 171–189. Postprint available at: <http://urn.fi/URN:NBN:fi:uta-201709112416>.
- Annala, J., Mäkinen, M., 2016. Communities of practice in higher education: contradictory narratives of a university-wide curriculum reform. *Stud. High Educ.* 42 (11), 1941–1957. <https://doi.org/10.1080/03075079.2015.1125877>.
- Annala, J., Mäkinen, M., 2011. Research-teaching nexus in higher education in curriculum design. *TCI* 8 (1), 3–25. Available online at: <http://ojs.library.ubc.ca/index.php/tci/article/view/2441>.
- Annala, J., Mäkinen, M., Lindén, J., Henriksson, J., 2020. Change and stability in academic agency in higher education curriculum reform. *J. Curric. Stud.* <https://doi.org/10.1080/00220272.2020.1836261>.
- Ashwin, P., 2020. *Transforming University Education – a Manifesto*. Bloomsbury, London.
- Barnett, R., 2000. University knowledge in an age of supercomplexity. *High Educ.* 40, 409–422. <https://doi.org/10.1023/A:1004159513741>.
- Barnett, R., Coate, K., 2005. *Engaging the Curriculum in Higher Education*. McGraw-Hill Education, Berkshire, GBR.
- Bernstein, B., 1996. *Pedagogy, Symbolic Control and Identity. Theory, Research, Critique*. Taylor & Francis, London.
- Biggs, J., Tang, C., 2011. *Teaching for Quality Learning at University: What the Student Does*, fourth ed. Society for Research into Higher Education & Open University Press, Maidenhead.
- Blackler, F., 2009. Cultural-historical activity theory and organization studies. In: Sannino, A., Daniels, H., Gutiérrez, K.D. (Eds.), *Learning and Expanding with Activity Theory*. Cambridge University Press, New York, pp. 19–39.
- Bobbit, J., 1972/1918. *The Curriculum*. Arno Press, New York.
- Bovill, C., Woolmer, C., 2019. How conceptualisations of curriculum in higher education influence student-staff co-creation in and of the curriculum. *High Educ.* 78 (3), 407–422. <https://doi.org/10.1007/s10734-018-0349-8>.
- Brøgger, K., 2014. The ghosts of higher education reform: on the organisational processes surrounding policy borrowing. *Glob. Soc. Educ.* 12 (4), 520–541. <https://doi.org/10.1080/14767724.2014.901905>.
- Carey, K.L., Stefaniak, J.E., 2018. An exploration of the utility of digital badging in higher education settings. *Educ. Technol. Res. Dev.* 66, 1211–1229. <https://doi.org/10.1007/s11423-018-9602-1>.
- Cedefop, 2017. *Defining, Writing and Applying Learning Outcomes: A European Handbook*. Publications Office, Luxembourg. <https://doi.org/10.2801/566770>.
- Cliff, A., Walji, S., Mogliacci, R.J., Morris, N., Ivancheva, M., 2020. Unbundling and higher education curriculum: a cultural-historical activity theory view of process. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2019.1711050>.
- Coate, K., 2009. Curriculum. In: Tight, M., Mok, K.H., Huisman, J., Morpew, C.C. (Eds.), *The Routledge International Handbook of Higher Education*. Routledge, New York, pp. 77–90.
- Dracup, M., Austin, J.E., King, T.J., 2020. Applying cultural-historical activity theory to understand the development of inclusive curriculum practices in higher education. *Int. J. Incl. Educ.* 24 (8), 882–900. <https://doi.org/10.1080/13603116.2018.1492638>.
- ECUI, 2020. European consortium of Innovative Universities. <https://www.ecui.org/>.
- European Commission/EACEA/Eurydice, 2018. *The European Higher Education Area in 2018: Bologna Process Implementation Report*. Publications Office of the European Union, Luxembourg. In: https://eacea.ec.europa.eu/national-policies/eurydice/content/european-higher-education-area-2018-bologna-process-implementation-report_en.
- Engeström, Y., 1987. *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Orienta-Konsultit, Helsinki.
- Engeström, Y., 1990. *Learning, Working and Imagining: Twelve Studies in Activity Theory*. Orienta-Konsultit, Helsinki.
- Engeström, Y., 1999. Expansive visibilization of work: an activity-theoretical perspective. *Comput. Support. Coop. Work* 8, 63–93.
- Engeström, Y., 2000. Activity theory as a framework for analyzing and redesigning work. *Ergonomics* 43 (7), 960–974. <https://doi.org/10.1080/001401300409143>.
- Engeström, Y., 2001. Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14 (1), 133–156. <https://doi.org/10.1080/13639080020028747>.
- Engeström, Y., Engeström, R., Kerosuo, H., 2003. The discursive construction of collaborative care. *Appl. Ling.* 24 (3), 286–315. <https://doi.org/10.1093/applin/24.3.286>.
- Engeström, Y., Kerosuo, H., 2007. From workplace learning to inter-organizational learning and back: the contribution of activity theory. *J. Workplace Learn.* 19 (6), 336–342. <https://doi.org/10.1108/13665620710777084>.
- Engeström, Y., Sannino, A., 2020. From mediated actions to heterogenous coalitions: four generations of activity-theoretical studies of work and learning. *Mind Cult. Activ.* 28 (1), 4–23. <https://doi.org/10.1080/10749039.2020.1806328>.
- Flinders, D.J., Thornton, S.J. (Eds.), 2004. *The Curriculum Studies Reader*, second ed. Routledge Falmer, New York.
- Foot, K.A., 2014. Cultural-historical activity theory: exploring a theory to inform practice and research. *J. Hum. Behav. Soc. Environ.* 24 (3), 329–347. <https://doi.org/10.1080/10911359.2013.831011>.
- Fraser, S.P., Bosanquet, A.M., 2006. The curriculum? That's just a unit outline, isn't it? *Stud. High Educ.* 31 (3), 269–284. <https://doi.org/10.1080/03075070600680521>.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M., 1994. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Sage, London.
- Grundy, S., 1987. *Curriculum: Product or Praxis*. Falmer, Lewes.
- Haapasaaari, A., 2020. *The Hunters of Lost Parcels: An Activity-Theoretical Study of the Emergence and Sustainability of Workers' Transformative Agency*. Doctoral dissertation. Helsinki Studies in Education, p. 86. <http://urn.fi/URN:ISBN:978-951-51-6386-8>.
- Handal, G., Hofgaard Lycke, K., Mårtensson, K., Roxå, T., Skodvin, A., Solbrekke, T.D., 2014. The role of academic developers in transforming Bologna regulations to a national and institutional context. *Int. J. Acad. Dev.* 19 (1), 12–25. <https://doi.org/10.1080/1360144X.2013.849254>.
- Karseth, B., 2006. Curriculum restructuring in higher education after the Bologna process: a new pedagogic regime? *Rev. Española Educ. Comp.* 12, 255–284.
- Karseth, B., Solbrekke, T.D., 2016. Curriculum trends in European higher education: the pursuit of the Humboldtian university ideas. In: Slaughter, S., Taylor, B.J. (Eds.), *Higher Education, Stratification, and Workforce Development, Higher Education Dynamics*, vol. 45, pp. 215–233. https://doi.org/10.1007/978-3-319-21512-9_11.
- Kelly, A.V., 2009/1977. *The Curriculum: Theory and Practice*, sixth ed. Sage, London.

- Knight, P.T., 2001. Complexity and curriculum: a process approach to curriculum-making. *Teach. High. Educ.* 6 (3), 369–381. <https://doi.org/10.1080/13562510120061223>.
- Krause, K.-L.D., 2020. Vectors of change in higher education curricula. *J. Curric. Stud.* <https://doi.org/10.1080/00220272.2020.1764627>.
- Kuutti, K., 1996. Activity theory as potential framework for human-computer interaction research. In: Nardi, B.A. (Ed.), *Context and Consciousness*. MIT Press, Cambridge, MA, pp. 103–116.
- Leont'ev, A.N., 1978. *Activity, Consciousness, and Personality*. Prentice-Hall, Englewood Cliffs, NJ.
- Lindén, J., Annala, J., Coate, K., 2017. The role of curriculum theory in contemporary higher education research and practice. In: Tight, M., Huisman, J. (Eds.), *Theory and Method in Higher Education Research*, vol. 3. Emerald, pp. 137–154. Postprint available. <http://urn.fi/URN:NBN:fi:tuni-201902181250>.
- Louvel, S., 2013. Understanding change in higher education as bricolage: how academics engage in curriculum change. *High Educ.* 66 (6), 669–691. <https://doi.org/10.1007/s10734-013-9628-6>.
- Luckett, K., Shay, S., 2020. Reframing the curriculum: a transformative approach. *Crit. Stud. Educ.* 61 (1), 50–65. <https://doi.org/10.1080/17508487.2017.1356341>.
- Mäkinen, M., Annala, J., 2010a. Meanings behind curriculum development in higher education. *PRIME* 4 (2), 9–24. Postprint available at: <http://urn.fi/urn:nbn:uta-3-1004>.
- Mäkinen, M., Annala, J., 2010b. Osaamisperustaisen opetussuunnitelman monet merkitykset korkeakoulutuksessa (Diverse meanings of competency-based curriculum in higher education). *Kasvatus & Aika* 4 (4), 41–61. http://www.kasvatus-jaika.fi/dokumentit/artikkeli_makinen_&_annala_0712100353.pdf.
- Mäkinen, M., Annala, J., 2012. Osaamisperustaisen opetussuunnitelman kahdet kasvat. [The two ways of understanding the competency-based curriculum in higher education]. In: Mäkinen, M., Annala, J., Korhonen, V., Vehviläinen, S., Norrgrann, A.M., Kalli, P., Svärd, P. (Eds.), *Osallistava korkeakoulutus [Participatory Higher Education]*. Tampere University Press, Tampere, pp. 127–151. <http://urn.fi/URN:ISBN:978-951-44-8947-1>.
- O'Connor, K., 2020. Constructivism, curriculum and the knowledge question: tensions and challenges for higher education. *Stud. High Educ.* <https://doi.org/10.1080/03075079.2020.1750585>.
- Pinar, W., 2004. What is Curriculum Theory? Lawrence Erlbaum, Mahwah.
- Priestley, M., 2011. Whatever happened to curriculum theory? Critical theory and curriculum change. *Pedagog. Cult. Soc.* 19 (2), 221–237. <https://doi.org/10.1080/14681366.2011.582258>.
- Priestley, M., Edwards, R., Priestley, A., Miller, K., 2012. Teacher agency in curriculum making: agents of change and spaces for manoeuvre. *Curric. Inq.* 42 (2), 191–214. <https://doi.org/10.1111/j.1467-873X.2012.00588.x>.
- Rawle, F., Bowen, T., Murch, B., Hong, R., 2017. Curriculum mapping across the disciplines: differences, approaches, and strategies. *Empowering Learners, Effecting Change CELT* 10. <https://doi.org/10.22329/celt.v10i0.4765>.
- Ryan, S., Connell, J., Burgess, J., 2017. Casual academics: a new public management paradox. *Lab. Ind.: J. Soc. Eco. Relat. Work* 27 (1), 56–72. <https://doi.org/10.1080/10301763.2017.1317707>.
- Sannino, A., Engeström, Y., 2018. Cultural-historical activity theory: founding insights and new challenges. *Cul. Histo. Psychol.* 14 (3), 43–56. <https://doi.org/10.17759/chp.2018140304>.
- Shay, S., 2011. Curriculum formation: a case study from history. *Stud. High Educ.* 36 (3), 315–329. <https://doi.org/10.1080/03075071003706681>.
- Shay, S., 2013. Conceptualizing curriculum differentiation in higher education: a sociology of knowledge point of view. *Br. J. Sociol. Educ.* 34 (4), 563–582. <https://doi.org/10.1080/01425692.2012.722285>.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development*. Heinemann, London.
- Taylor, F.W., 2003/1911. *Scientific Management. The Early Sociology of Management and Organizations*, vol. 1. Routledge, NY.
- Taylor, J.R., 2009. The communicative construction of community: authority and organizing. In: Sannino, A., Daniels, H., Gutiérrez, K.D. (Eds.), *Learning and Expanding with Activity Theory*. Cambridge University Press, New York, pp. 228–239.
- Tucker, J., Neely, P., 2010. Unbundling faculty roles in online distance education programs. *Int. Rev. Res. Open Dist. Learn.* 11 (2), 20–32. <https://doi.org/10.19173/irrodl.v11i2.798>.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction*. The University of Chicago Press, Chicago.
- Visser-Wijnveen, G.J., Van Driel, J.H., Van der Rijst, R.M., Verloop, N., Visser, A., 2010. The ideal research-teaching nexus in the eyes of academics: building *profiles*. *High Educ. Res. Dev.* 29 (2), 195–210. <https://doi.org/10.1080/07294360903532016>.
- Wang, C.L., 2015. Mapping or tracing? Rethinking curriculum mapping in higher education. *Stud. High Educ.* 40 (9), 1550–1559. <https://doi.org/10.1080/03075079.2014.899343>.
- Wartofsky, M., 1979. *Models: representation, and the scientific understanding*. D. Reidel Publishers, Dordrecht, Netherlands.
- Wheelehan, L., 2010. *Why Knowledge Matters in Curriculum. A Social Realist Argument*. Routledge, London.
- Wihlborg, M., Teelken, C., 2014. Striving for uniformity, hoping for innovation and diversification: a critical review concerning the Bologna process – providing an overview and reflecting on the criticism. *Pol. Futures Educ. Internet* 12 (8), 1084–1100. <https://doi.org/10.2304/pfie.2014.12.8.1084>.
- Willis III, J.E., Flintoff, K., McGraw, B., 2016. A philosophy of open digital badges. In: Ifenthaler, D., Bellin-Mularski, N., Mah, D. (Eds.), *Foundation of Digital Badges and Micro-credentials. Demonstrating and Recognizing Knowledge and Competencies*. Springer, pp. 23–38.
- Yates, L., Woelert, P., Millar, V., O'Connor, K., 2017. *Knowledge at the Crossroads. Physics and History in the Changes World of Schools and Universities*. Springer, Singapore.
- Ylijoki, O.-H., Ursin, J., 2013. The construction of academic identity in the changes of Finnish higher education. *Stud. High Educ.* 38 (8), 1135–1149. <https://doi.org/10.1080/03075079.2013.833036>.
- Young, M., 2008. *Bringing Knowledge Back in: From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, London.
- Young, M., 2011. Curriculum policies for a knowledge society. *Curriculum in Today's World: Configuring Knowledge, Identities, Work and Politics*. In: Yates, L., Grumet, M. (Eds.), *World Yearbook of Education 2011*. Routledge, London, pp. 125–138.
- Young, M., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45, 101–118. <https://doi.org/10.1080/00220272.2013.764505>.

Children's interests and curriculum making in early childhood education

Liz Chesworth^a and Helen Hedges^b, ^aThe University of Sheffield, South Yorkshire, UK; and ^bThe University of Auckland, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	173
Section one: critiquing an entrenched coalition	173
Section two: theoretical framings of children's interests	174
Section 3: children's interests as inquiry and identity development	176
Section 4: "intellectually honest" curriculum making	177
Conclusion	179
References	179

Introduction

No curriculum is value-free. Decisions relating to what, and how, young children should learn are open to multiple influences and interpretations, and have long been the subject of international debate. This is because curriculum decision-making is implicated in the political, cultural and economic contexts within which it is situated. It is shaped by particular perspectives about childhood, learning, what counts as knowledge and valued ways of coming to know and understand. Ultimately, then, curriculum acts as a mirror that reflects political and societal values regarding the purpose of early childhood education (ECE) and ways young children are viewed in society.

With these points in mind, our planning and development of this chapter was informed by the following questions: Who are the curriculum makers in ECE? How does curriculum making reflect views/images of children? How does curriculum making reflect common social and political perspectives/views that a main purpose of ECE is school readiness? What, and whose, knowledge, underpins curriculum making? And in what ways might curriculum produce strong inquirers or compliant learners? Drawing upon our research, we show the potential of approaches to ECE in which children's interests are the focus of curriculum making decisions.

Children's interests have long been recognized as contributing to curriculum making in early childhood, but suffered from narrow interpretations. By the end of the first section of this chapter we have problematized two entrenched contributions to decisions about foci of interests-based curriculum making in early childhood and offered an alternate proposal. Our proposal is built from views of children as capable and motivated learners whose diverse life experiences can drive curriculum and therefore position children as leading curriculum making in ECE.

In section two we present three theoretical framings for children's interests: developmental psychology, sociocultural theory and post-humanism. We argue that each framing illuminates interests for curriculum making differently. In section three we develop our proposal about interests and capabilities to make a case for a dynamic approach to curriculum making that is driven by deep analysis and authentic interpretations of children's interests and play choices. We draw on our research to introduce and exemplify concepts and approaches that emphasize the ways that inquiry and identity are central to children's interests.

The final section attempts to resolve dilemmas the first section raised about curriculum making focused on narrow ideas about outcomes and school readiness. We address these concerns through the alternate ideas about curriculum making based on interests and selected ideas of Bruner (1960, 1996). We show how interests-based outcomes inspire children's learning and promote this as an alternative form of "readiness" in which children make sense of the world through curriculum making that has contextual meaning and relevance. Throughout the chapter we offer examples of children from our research projects that exemplify the curriculum making potential of children's interests.

Section one: critiquing an entrenched coalition

International research and debate about curriculum in ECE highlight two key drivers that exert a powerful influence upon curriculum making in many contemporary contexts (see, for example Wood and Hedges, 2016). First, developmental psychology and associated child development theories have had an enduring influence upon the construction of many national curriculum frameworks for ECE. These frameworks are informed by Piagetian notions of development that establish linear expectations and progressive sequences relating to developmental milestones by which to understand and assess children's learning and achievements. Second, the political appeal of investing in the early years for future economic prosperity has led to including approaches that prioritize academic outcomes. Such approaches foreground ECE as a site for school readiness in order to generate a productive workforce for the future.

Taken together, the coalition of child development theory and political ideology has created dominant framings of curriculum in which educational success and progression are measured in terms of decontextualized academic achievement, often with a focus on early literacy and mathematics. Driven by this ideology, the Organization for Economic Cooperation and Development's (OECD) controversial cross-national assessment of five-year olds, the International Early Learning and Child Well-being Study (IELS), is indicative of the global preoccupation with instrumental measures of children's progress. The priorities that drive the IELS have been subject to widespread criticism by early childhood scholars and practitioners (see, for example, Moss and Urban, 2020) and are likewise enacted in many national policies that construct reductive approaches to young children's progress and school readiness. Wood's (2020) critical discourse analysis of the Early Years Foundation Stage (EYFS) (DfE, 2017) curriculum in England reveals how the EYFS exemplifies a global preoccupation with linear notions of progress and school readiness.

This framing of curriculum is informed by a standardized conceptualization of knowledge in which predetermined learning outcomes are assumed to have universal relevance for all children (Chesworth, 2019). Assumptions of progress and readiness have played a powerful role in shaping dominant framings of ECE curriculum and have generated a simplified construction of knowledge that erases the social, cultural and historical contexts within which young children live and learn. The assessment of children in relation to universally imposed norms constructs notions of the "deficit" child, family and community by which children who do not align with normative learning outcomes are positioned as somehow lacking or problematic (Burman, 2008). Hence, curriculum making in ECE has often been viewed as a teacher-directed social and academic intervention, designed to redress assumed social and economic disparities, and enhance children's readiness for the formality of school-based instruction with an eye on future economic productivity.

Thus, an outcomes-driven agenda exerts a powerful influence upon what knowledge and ways of knowing are deemed to have value and validity in curriculum making. In this standardized enactment of curriculum, young children are ascribed a passive role in which they are positioned as recipients of knowledge. Such an approach runs counter to the overarching tenet of this chapter: we argue that curriculum making should start from the premise that all children, no matter their life circumstances, are eager, competent and capable learners. Children's lives are situated within specific cultural contexts. Normative outcomes mask the complexity and diversity of children's learning and development. "Treating European American middle-class people and their communities as the "norm" is misleading and undermines understanding of both the generalities and the specifics in human development world-wide" (Rogoff et al., 2018, p. 4), a point we exemplify shortly in refuting deficit accusations and instead highlighting children's capabilities as learners.

A credit-based approach to children's learning and curriculum provision requires that teachers pay attention to diverse knowledge and ways of knowing that are mediated by children's active participation in the everyday activities of their families and communities. Rogoff, Correa-Chávez and Dexter (2015) frame these everyday activities as "cultural repertoires" which shift the focus of development from an individual trajectory toward recognition that learning takes place in the context of social interactions and relationships. From the many experiences children participate in, some connect with and/or spark their curiosity more than others. These become the kinds of sustained interests this chapter champions. This perspective recognizes that young children and their families have the potential to play an active role in curriculum making. It places value upon contextually situated knowledge and diverse ways of knowing. In this way, the goal of curriculum making shifts from decontextualized measures of academic performance toward enabling children to move toward learning grounded in deeper understanding from and in meaningful contexts.

Furthermore, standardized curriculum frameworks can lead to a technicist framing of interests in which teachers' interpretations are informed by predetermined learning outcomes rather than the everyday lived experiences of children and families (Chesworth, 2016, 2019). For example, a child may be interested in a book about going fishing with a grandparent, and while reading it be eager to share their knowledge and experience with a teacher, who instead diverts the child's attention to the literacy features of the book aware of the accountability measures s/he will be subject to. As such, we suggest that curriculum making informed by children's interests requires a stronger conceptual framing that addresses a middle ground between child-initiated play and teacher-led learning, foregrounds children's perspectives in order to avoid narrow, tokenistic or instrumental connections between interests and curriculum, and requires a rethinking of what constitutes outcomes and school readiness.

In this section we have problematized approaches to curriculum making that are driven by linear interpretations of child development, standardized learning outcomes and deficit notions of diversity. We have presented a case for curriculum making that recognizes children as competent and capable learners, and that is responsive to children's interests derived from diverse experiences in families and communities. In the section that follows we consider how a focus upon children's interests can inform curriculum making from three theoretical perspectives.

Section two: theoretical framings of children's interests

Play has been established, both historically and presently, as a strong foundation for children's development and learning (see, for example, Brooker et al., 2014). The rhetoric of building a curriculum upon children's play choices and interests is also well established within ECE curriculum documents and practice, if somewhat unproblematized (Wood, 2014) or researched until recently. Children are motivated to learn through connecting new experiences to what they know and are interested in. Interests, therefore, afford possibilities for meaningful curriculum making that has relevance for children's diverse lived experiences. Interests offer a potential means for "children and teachers to spontaneously inquire into and explore intuitive ideas, and the content relating

to these, in highly participative and interpretive ways" (Hedges, 2015, p. 303) rather than curriculum making being prescribed by outcomes to which teachers are held accountable.

However, notions of what constitutes "interest" are not universal. These notions are shaped by different theoretical and ideological perspectives about children, childhood, learning and pedagogy that we explain next.

Justification for the longstanding interpretations of children's interests within many curricular frameworks as emphasizing upon children's individual engagement with materials or activities within ECE play environments have derived from developmental and educational psychology. Related research offered children choices of toys or play equipment in experimental studies and deemed the activity selected as the interest. However, an activity-based framing of interest first implies a surface or low-level interpretation that has a tendency to underestimate the depth and complexity of children's thinking, reasoning and motivations to learn that are represented in that choice of activity. Second, this framing does not take account of cultural repertoires, whereby many settings continue to rely on Western ideals of play equipment and neglect considerations of equity and diversity in the choices made available to children. Third, adult roles were viewed as "hands-off", with safety and supervision overriding foci rather than offering input into children's play and subsequently, knowledge development of interest to children that fostered ongoing curiosity, knowledge building and thinking.

From a sociocultural perspective, curriculum-making can be informed by participatory approaches in which young children's play-based learning is understood to be mediated by interaction in social and cultural contexts (Rogoff, 2003). This perspective calls for children's interests to be interpreted in relation to their participation in family and community activities. Viewed as such, children's interests can enable teachers to notice, understand and respond to the cultural repertoires (Rogoff et al., 2015) that inform children's construction of identity and dispositions to learn in diverse sociocultural contexts. In this way, building a curriculum in response to children's interests offers potential for children to apply and transform knowledge, leading to deeper understanding in meaningful contexts.

Supporting this alternative position, we have contributed to a recent and significant body of international research that highlights the benefits of using children's interests developed in their families and cultures as a basis for curriculum-making (Chesworth, 2016; Hedges et al., 2011). In this chapter we next focus upon studies from this body of work that have applied and extended the concept of funds of knowledge (González et al., 2005) to illuminate approaches to curriculum that value, connect with and build upon the interests that children amass from their participation in home and community activities. Informed by sociocultural theory, the notion of funds of knowledge highlights the richness of experiences arising from children's participation in household and community activities. Moll et al. (1992) have proposed that all families, whatever their economic, social and cultural circumstances, have "ample cultural and cognitive resources" (p. 134) arising from funds of knowledge associated with everyday household functioning, development, relationships and wellbeing. The principles that inform a funds of knowledge approach to understanding children's interests have much in common with Bruner's beliefs regarding the connections between culture and curriculum that we discuss later.

From a sociocultural perspective, children's interests are motivated by a desire to participate in, and make meaning of, the cultural lives of their families and communities. Funds of knowledge-based interests arise from children's active participation in household practices, including domestic tasks, holidays and community experiences as well as through the influence of family members' occupations and leisure activities (Hedges et al., 2011). Funds of knowledge-based interests can also be influenced by less tangible aspects of home and community activity including family and cultural values, beliefs and languages (Hedges et al., 2011). For example, a child might develop a deep interest in growth, change and caring for plants as a result of funds of knowledge amassed from helping a parent or grandparent in the garden. This could generate appreciation for sustainable gardening practices and respect for wildlife and ecology. Likewise, children might develop interests relating to their participation in household activities relating to food preparation, baking or contributing to the care of younger siblings. Funds of knowledge-based interests could also develop from participating in religious activities or from observing a parent's engagement in hobbies or sporting activities. Locating interests within the multiple sources that constitute funds of knowledge can consequently inform curriculum making that responds to and sustains the diversity of practices and interactions that constitute young children's everyday lives (Chesworth, 2014).

Contemporary studies have built upon González et al.'s (2005) work to acknowledge a broader spectrum of funds of knowledge that extend beyond children's participation in family activities. Such studies highlight the ways in which funds of knowledge-based interests arise from children's participation in multiple contexts, including their interactions with popular culture (Chesworth, 2016; Hedges 2011; Poole, 2017), and with teachers and peers in ECE settings (Hedges et al., 2011).

Funds of knowledge that are selected as meaningful, that is, are of interest, have recently been termed funds of identity (Esteban-Guitart, 2016; Esteban-Guitart and Moll, 2014) and extended into early childhood education (Hedges, 2021). We see the concept of funds of identity as being particularly helpful in illuminating how children's interests arising from lived experiences contribute to their sense of themselves as members of social and cultural communities. Viewed through a funds of knowledge and/or funds of identity lens, children's interests afford potentially meaningful opportunities for curriculum making that values, connects with, and extends children's diverse knowledge and ways of knowing.

A recent theoretical framing we see has potential with regard to children's interests is post-humanism. Many Indigenous Peoples have long understood the world as an entanglement of human and non-human elements (Pacini-Ketchabaw and Taylor, 2015) and recent years have seen interest in bringing these ideas to Western educational theory. For example, New Materialism and the broader ontology of posthumanism within which it resides, locate learning within vibrant ecologies in which humans and nonhumans are always intra-acting, always in relation with one another (Pacini-Ketchabaw et al., 2017).

A post-human framing can inform further work to be undertaken to understand how interests may emerge from children's engagement with aspects of their everyday worlds that do not sit within a classification of social or cultural experiences. This position affords possibilities to adopt a broader lens to explore children's interests (Chesworth, 2019) by extending interpretations beyond social and cultural factors in order to explore a wider range of "causal factors for understanding what is going on in an educational encounter" (Murris, 2016, p. 91). For example, a post-human perspective can enable teachers to move toward deeper interpretations of children's interests that arise from their relations with places, animals, insects and other non-human things to which children are often deeply attached.

We clarify that while a post-human lens offers potential for new insights into interests, such a reading does not reject the importance of children's participation in sociocultural practices; rather, it blurs the boundaries between social, cultural and nonhuman factors which interweave within and beyond children's everyday experiences. Locating young children's interests within their affective and practical relations with human and more than human aspects of their everyday lives recognizes that learning is dynamic, responsive and situated. Viewed as such, children's interests are constructed through connections, and are open to new connections that lead to new knowledge and understanding. A post-human understanding of interests complements sociocultural positions that foreground children "being part of the world" (Haraway, 2020) and therefore central to curriculum making. As such, post-humanism affords possibilities to build upon, not replace, participatory perspectives of interest (Chesworth, 2019). However, there is to date no empirical work that identifies and evaluates the ways in which this position can inform interests-based approaches to curriculum-making in ECE.

Of these framings, our research on children's interests has been situated in sociocultural theory. We have identified play, funds of knowledge, inquiry and identity as significant in interpretations of interests that can inform curriculum making. We have offered examples of play that draws on funds of knowledge in this section. Such curriculum is authentic to children because it respects their lives and experiences. It builds personal, academic and societal knowledge in the context of children's interests-motivated participation in play-based activities that have deep representational possibilities and meaning for children. We now go on to explore ideas from our research of children's interests framed as inquiry and identity development, and illustrate their power for curriculum making.

Section 3: children's interests as inquiry and identity development

This section introduces a dynamic approach to curriculum making that is informed by interpretations of children's interests as centered on meaningful inquiry and identity development. In contrast to constructions of curriculum that position children as passive recipients of knowledge, a dynamic curriculum recognizes that children and teachers play an active role in curriculum making together in both planned and spontaneous ways. Such an approach is driven by respect for children's capabilities; informed by sociocultural theory, we locate these capabilities within children's relationships, local contexts and cultures. These understandings of children and culture lay the foundations for curriculum making that is respectful of, and responsive to, children's diverse lives and experiences. Such an approach requires teachers to know children and their families well enough to identify and respond to the funds of knowledge-based interests and "real questions" (Hedges and Cooper, 2016) that children are investigating through their play and inquiries. This highlights the central role of relational pedagogy: respectful dialog and responsive relationships with children, parents and caregivers in order to inform in-depth interpretations of, and responses to, the individual and collective interests that children enact in early childhood settings.

Children's play choices in early childhood settings represent and mediate their interests and intentions (Göncü, 1987). As such, self-initiated play-based learning affords possibilities for children to explore and extend personal and shared interests with peers and teachers. For example, Chesworth's (2014) research indicated that children's play choices were driven by a motivation to connect with, and deepen their understanding of, the people, practices and places that constituted their everyday worlds. As such, play has potential to be a cultural broker (Walker and Nokon, 2007) in which funds of knowledge-based interests connect ECE settings, homes and community contexts, and enable children to making meaning of the different activities that constitute their sociocultural worlds.

Our research exemplifies the possibilities that play affords for children to draw upon familiar adult scripts (Göncü, 1987) that have interest to them, and to imagine themselves into identities that are complex and creative (Holzman, 2010). For example, Ellie and Lily, both five-years old, co-constructed a play narrative in which they prepared for an imaginary party. The children enacted mutual funds of knowledge of baking and birthday celebrations and also drew upon their shared interest in Disney princess movies to introduce a fantasy element to their play. In this episode of play, Ellie and Lily co-ordinated their activity through a complex blending of meanings, in which funds of knowledge arising from diverse sources were interwoven to imagine new scenarios and identities (Chesworth, 2016).

Noticing the funds of knowledge-based interests that children are exploring during play offers important possibilities for teachers to respond to and extend learning in ways that have meaning and relevance for children's learning and identities. This approach offers a framework for a dynamic model of curriculum making that is attuned to the complex meanings and motivations that children bring to their play (Hedges, 2021; Hedges and Cooper, 2016). Furthermore, this approach offers a potential lens to notice the ways in which funds of knowledge-based interests connect with children's identity construction and to value the ways in which individuals create their identities through experiences that have particular interest and meaning for them (Esteban-Guitart and Moll, 2014; Hedges, 2021).

For example, in the case of Ellie and Lily, teachers could draw upon their knowledge of the children to bring greater depth to their interpretations of interest in ways that progress beyond noticing the sources of interest (baking, birthday parties and Disney films) to identifying the “real questions” that the children were exploring in their play. Such questions might, for example, revolve around children’s inquiries into present and future identities, such as: How do my experiences enable me to connect with others? Who am I and what can I do? Who else can I be? In another example, Hunter had immense skills in playing the drums. Analyzing this interest as activity-based only would have missed the deep cultural meaning music played in Hunter’s life. This interest represented his funds of knowledge, and his interest in developing his identity as a member of his family, cultures and communities. In turn, the curriculum experiences teachers developed for him led to him developing an identity as a confident learner and communicator that would stand him in good stead for schooling in ways a focus on academic readiness would not have achieved (Cooper and Hedges, 2014).

Dynamic modes of curriculum making, therefore, afford opportunities to affirm children’s identity construction and to extend learning through open ended, flexible responses that are attuned to children’s authentic questions and intentions. This means that teachers’ in-the-moment pedagogical interactions with children, as well as subsequent planning for future learning experiences, need to be synchronous with children’s interests, motivations and inquiries. In this way knowledge and skills can be brought forth/contribute to children’s learning through the sensitive, relational and conceptual input of teachers. This approach to curriculum making recognizes children’s curiosity and motivation to explore their interests in new contexts. Children actively seek new information that stimulates further inquiry and investigation when the source and context of the curriculum is their interests.

Foregrounding inquiry and identity is indicative of a holistic approach to curriculum making. An inquiry-driven approach to curriculum-making is informed by respect for children’s motivations and capabilities to move toward deeper understandings and apply their learning in new contexts. This mode of curriculum foregrounds the co-construction of meaningful knowledge through a dynamic and collaborative process involving children and adults. In this way, children’s interests and playful explorations are valued while teachers draw upon their deep knowledge of children to mediate new opportunities for children’s knowledge development and critical thinking (Hedges and Cooper, 2018). Moreover, as Peters et al. (2018) argue, a curriculum that foregrounds mutual exploration, challenge and curiosity leads to more complex understandings for children and teachers alike. Children seeing teachers as enthusiastic models of learning further ignites their inquiry and knowledge building.

In summary, a holistic view of curriculum recognizes the interweaving of cognitive and affective aspects of learning and offers an alternative to curricula frameworks that emphasize academic outcomes over affective and emotional domains of learning. This holistic view constitutes a significant shift from the deficit models of curriculum delivery that were outlined and problematized earlier in the chapter. Deeper interpretations of children’s interests as inquiry and identity construction offer rich possibilities for teachers to interact with children in relational ways that create broad, balanced and playful learning experiences that are responsive to children’s diverse experiences and ideas (Hedges and Cooper, 2018). These interpretations also have the potential to incorporate the knowledge valued by society represented in more recognizable ways as academic outcomes, an idea we explore next in our final section connecting culture with learning and knowledge building.

Section 4: “intellectually honest” curriculum making

The conceptualizations of children’s interests presented in this chapter offer ways to interpret and respond to children’s interests with authenticity. This is critical to curriculum making that is cognizant of children’s interests and highlights the vital roles teachers play in co-constructing learning experiences that have meaning and relevance for children.

Teachers notice the activities that children choose, valuable in themselves for play and learning experiences, but when aligned with deeper interpretations of interests as representing funds of knowledge, inquiry and identity development can recognize, respond to, extend and revisit them differently than as mere activities. Extending and revisiting enables children’s knowledge to grow in topics/subjects that are meaningful to them, that they are motivated to learn. However, all play-based activities available represent the explicit and implicit cultural values and priorities of the context they are located in, hence connections between culture and learning are always central to educational possibilities.

The relationship between culture and learning was a career-long focus of Jerome Bruner’s. In his 1960 book on *The process of education* Bruner made a provocative statement that aligns with the central idea in this chapter that children are capable and competent learners.

We begin with the hypothesis that any subject can be taught effectively in some intellectually honest form to any child at any stage of development. It is a bold hypothesis and an essential one in thinking about the nature of the curriculum. No evidence exists to contradict it; considerable evidence is being amassed that supports it (p. 33).

This statement is also powerful for considerations of the conceptual possibilities inherent in children’s interests, possibilities that may have the potential to address considerations of school readiness in more child-focused ways. There are naturally considerations of caveats about the degree to which a subject can be taught related to children’s maturity and life experiences. Perhaps this is partly what Bruner meant by an “intellectually honest form”?

Nevertheless, we have all met a child who, for example, knew more about dinosaurs or sharks than we will ever know. Campbell had a deep interest in all sea creatures, an interest that had been fostered by family fishing expeditions as a source of funds of knowledge. He brought this interest to kindergarten through his choice of playing with plastic animals. A teacher who listened intently to Campbell realized the focus of his interest and initially added to the resources provided (more plastic creatures and books about sea creatures). Campbell's interest sparked further curiosity and a thirst for knowledge. For example, he and a teacher used a tape measure in the playground to measure and demonstrate the length of a whale shark. In these ways early literacy and numeracy prized by many curricular outcomes were incorporated in Campbell's learning, and his knowledge of the "subject" of sea creatures was both intellectually honest and developed over time in increasingly complex and additive ways (Hedges, 2004).

There are then also understandings of the associated processes of children's learning and thinking and curriculum making that Bruner was very clear about, well ahead of his time. His points relate closely to our arguments about children's interests. Bruner outlined the importance of the roles of curiosity, inquiry, perception and intuition in children's learning, and ways that research to that date – over 60 years ago – acknowledged children's capabilities. These we argue form the ways that a subject could be addressed through inquiry in an "intellectually honest form".

Bruner argued that interest in a subject – perhaps a topic of deep interest as exemplified by Campbell's interest in sea creatures – was the motivation for learning. Aligning with Vygotsky's (1986) statement that "direct teaching of concepts is impossible and fruitless" (p. 150) Bruner reasoned that didactic teaching did not enable the "excitement about discovery" (p. 20) nor develop the dispositions of curiosity and inquiry necessary for learning. For Bruner, learning and knowledge building required access to information alongside ongoing opportunities to apply and transform knowledge in new situations while simultaneously evaluating the value of that information. He implied that where children had great motivation and deep interest that related subject information could be taught through participatory and mediated teaching and learning.

For example, when Liz was teaching she noticed that four-year old Faizah regularly brought brushes and water from the painting area and used them to wash the nursery cars and trikes. Subsequent conversations with Faizah and her parents revealed that Faizah's father had recently set up a car valeting business following a long period of unemployment. Careful listening to the family enabled Liz and her colleagues to understand that Faizah's interest was closely connected to her relationship with her father and to the family's collective pride in his new business venture. In response, they introduced additional sponges, cloths and brushes so that Faizah could explore, extend and share her interest with other children. Faizah's interest led to the children and their teachers creating a car wash in the nursery garden that generated further opportunities for learning. For example, Liz encouraged children to predict, observe and talk about the effect of different cleaning materials and solutions (for example, soapy, hot and cold water). On another occasion Faizah brought in a valeting menu from her dad's business and engaged with a teacher to produce an itemized list that presented the washing prices for each of the nursery vehicles. These examples are indicative of Bruner's argument regarding the potential for knowledge building in contexts that have meaning and relevance for children's lives. Bruner's position resonates with Vygotsky's (1987) seminal work in which he argued that the relational connection between concrete experiences and academic concepts underpins and drives the process of learning. Faizah's interest offered a meaningful context in which everyday experiences and academic concepts could connect and lead to deeper mathematical and scientific understanding.

With these critical points, and the associated examples of Campbell and Faizah in mind, we offer Bruner's view of curriculum making: the notion of a spiral curriculum. He stated that:

If one respects the ways of thought of the growing child, if one is courteous enough to translate material into his logical forms and challenging enough to tempt him to advance, it is possible to introduce him at an early age to the ideas and styles that in later life make an educated man. We might ask, as a criterion for any subject taught in primary school, whether, when fully developed, it is worth an adult's knowing, and whether having known it as a child makes a person a better adult. If the answer to both questions is negative or ambiguous, the material is cluttering the curriculum. If the hypothesis with which this section was introduced is true - that any subject can be taught to any child in some honest form - then it should follow that a curriculum ought to be built around the great issues, principles, and values that a society deems worthy of the continual concern of its members.

(Bruner, 1960, p. 52, p. 52).

The idea of great issues, principles and values as the basis for curriculum has resonance with the notion of fundamental inquiry questions as the ideas that underpin children's interests and that drive their learning. Perhaps the example of Campbell shows ways teachers could introduce authentically issues of sustainability and conservation as representing issues, principles and values vital to the world that need ongoing problem-solving and innovative responses. The fundamental inquiry questions might be: How can I understand the world I live in? What do intelligent, responsible, and caring adults do? If his schooling experiences continued to develop his interests, inquiry, knowledge and skills, perhaps Campbell has grown up to study marine engineering and conservation. If so, there would be little doubt Campbell would be the kind of "productive, economic citizen" some societies' curricular documents promote, as well as being a citizen and contributor to concerns about world sustainability issues.

These ideas represent a dynamic and evolving curriculum. Therefore, they also align with Bruner's (1996) later proposal in his book *The culture of education* that curriculum "is like an animated conversation on a topic that can never be fully defined ... animated ... because it is always lively if it is honest, but also because one uses animation ... props, pictures, texts, films, and even "demonstrations" (p. 116). This description of an ongoing conversation between teachers and children that is supported by planned and spontaneous resources, equipment and experiences has a good alignment with interests and play-based ECE provision in the way articulated in this chapter. From this perspective, culture is the context for understanding meanings and interpretations of both

interests and content, and also the means to innovate and go beyond what is known. This view is important to instil/enable in young children as our conceptualization of interests shows that they are inquirers into their worlds, and might be keen to/encouraged/prompted through ongoing interests that involve inquiry and identity development to innovate and problem-solve on a larger scale later in their lives.

The spiral notion refers to curriculum making as beginning with selected principles and information aligned with children's intuitive and other early interests-based understandings that is returned to in a recursive, iterative and increasingly complex manner many times throughout education. The example of Faizah offers a snapshot from this dynamic spiral of curriculum making in which the teacher used the nursery car wash as a meaningful context for building upon children's understanding of scientific and mathematical concepts. Over time, these concepts could be revisited so as to enable Faizah's interests to lead to deeper learning in a widening range of contexts. Bruner was clear that no knowledge is beyond children's understanding and that the teacher's role was to locate the ideas of interest and grow the child's learning through careful and deepening provision of experiences and scaffolding of topics/knowledge/subjects. Understanding, rather than the performance of decontextualized academic achievement, was Bruner's focus.

Clearly this spiral curriculum, learning and education can begin in ECE where interests are valued as a source of curriculum that can build knowledge. This is an idea Bruner championed long ago to redress concerns about the didactic nature of teaching, a concern that has crept into ECE over the years through a decontextualized focus on school readiness as being achieved through teacher-led approaches. As a longstanding scholar of education, Bruner's ideas about "readiness" in 1960 were these of a child-focused approach to curriculum making that kept returning to, and deepening, interests and knowledge. It was not about the imposition of content to reduce so-called academic deficits and socioeconomic disparities, content that would lead to later continued context-free mastery of learning.

The principles of a spiral curriculum in summary form then are content, coherence, extension/expansion, and progression that align with children's ways of knowing and learning. Each principle has implications for the role and responses of a teacher. Selecting *content* involves deep interpretation of interests. *Coherence* involves the revisiting of related ideas and experiences over time and in gradually more *progressive* and *expansive* ways. Notwithstanding this principle, we note that children's learning can take retrogressive steps – the eventual and intended direction of a spiral is upwards, but that may take time and repeated opportunities.

Conclusion

Curriculum decision-making is implicated in political, cultural and economic contexts. In the context of ECE children's interests are commonly included as a source of curriculum in policy documents but have been overridden by concerns about school readiness that lead to decontextualized societal interests as an influence on curriculum. We have offered a conceptualization of curriculum making built on a view of children as competent and capable learners who bring a range of interests to their educational experiences from their participation in families, communities and cultures. We presented three theoretical framings for children's interests, and then argued for sociocultural interpretations as leading to more authentic and dynamic curriculum in ECE. Throughout we offered examples of children's interests and their incorporation into curriculum.

We then offered selected intellectual contributions from Jerome Bruner to show that a more dynamic, relational and responsive approach to curriculum making in ECE can still lead to the kinds of outcomes desired by a society. Consideration of knowledge as "intellectually honest" and curriculum as a "spiral" enables knowledge and curriculum to be led by children and fostered by teachers in ways that meet shared personal, community, political and societal objectives. We suggest that synthesizing funds of knowledge-based interests and Bruner's spiral curriculum can inform new thinking about children's "readiness" in which the focus shifts from current preoccupations with standardized and decontextualized approaches to learning and progress in ECE. Instead, readiness is reframed as a relational process of curriculum making and pedagogical interactions that foster children's motivation to learn in contexts that have meaning and relevance to their everyday experiences and relationships.

References

- Brooker, L., Blaise, M., Edwards, S. (Eds.), 2014. The SAGE International Handbook of Play and Learning in Early Childhood. Sage, London, England.
- Bruner, J., 1960. The Process of Education. Harvard University Press, Cambridge, Mass.
- Bruner, J., 1996. The Culture of Education. Harvard University Press, Cambridge, Mass.
- Burman, E., 2008. Developments: Child, Image, Nation. Brunner Routledge, London.
- Chesworth, L., 2014. Multiple Perspectives of Play in an English Reception Class: Exploring Insider Meanings. PhD thesis. Leeds Beckett University.
- Chesworth, L., 2016. A funds of knowledge approach to examining play interests: listening to children's and parents' perspectives. *Int. J. Early Years Educ.* 24 (3), 294–308.
- Chesworth, L., 2019. Theorising young children's interests: making connections and in-the-moment happenings. *Learn. Cult. Soc. Interact.* 23, 100263. <https://doi.org/10.1016/j.lcsi.2018.11.010>.
- Cooper, M., Hedges, H., 2014. Beyond participation: what we learned from hunter about collaboration with Pasifika children and families. *Contemp. Issues Early Child.* 15 (2), 165–175. <https://doi.org/10.2304/ciec.2014.15.2.165>.
- DfE (Department for Education), 2017. Statutory framework for the Early Years Foundation Stage: setting the standards for learning, development and care for children from birth to five. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596629/EYFS_STATUTORY_FRAMEWORK_2017.pdf. (Accessed 22 February 2021).
- Esteban-Guitart, M., 2016. Funds of Identity: Connecting Meaningful Learning Experiences in and Out of School. Cambridge University Press, New York, NY.

- Esteban-Guitart, M., Moll, L.C., 2014. Funds of identity: a new concept based on the Funds of Knowledge approach. *Cult. Psychol.* 20 (1), 31–48. <https://doi.org/10.1177/1354067X13515934>.
- Göncü, A., 1987. Toward an interactional model in developmental changes in social pretend play. In: Katz, L., Steiner, K. (Eds.), *Current Topics in Early Childhood Education*, vol. 7, pp. 126–149.
- González, N., Moll, L.C., Amanti, C., 2005. Introduction: theorizing practices. In: González, N., Moll, L.C., Amanti, C. (Eds.), *Funds of Knowledge. Theorizing Practices in Households, Communities and Classrooms*. Routledge, London, pp. 1–24.
- Haraway, D., 2020. "Education Beyond the Human: a conversation with Donna Haraway" [Interviewed by Iveta silova in preparation for CIES 2020 "Education Beyond the Human: toward Symptoiesis"]. *Comp. Int. Educ. Soc.* [Online]. Available: <https://vimeo.com/453292469>.
- Hedges, H., 2021. The place of interests, agency and imagination in funds of identity theory. *Mind Cult. Activ.* 28 (2), 111–124. <https://doi.org/10.1080/10749039.2020.1833931>.
- Hedges, H., 2015. Sophia's funds of knowledge: theoretical and pedagogical insights, possibilities and dilemmas. *Int. J. Early Years Educ.* 23 (1), 83–96.
- Hedges, H., 2011. Rethinking Sponge Bob and Ninja Turtles: popular culture as funds of knowledge for curriculum co-construction. *Australas. J. Early Child.* 36 (1), 25–29.
- Hedges, H., 2004. A whale of an interest in sea creatures: the learning potential of excursions. *Early Child. Res. Pract.* 6 (1). <http://www.ecrp.uiuc.edu/v6n1/hedges.html>.
- Hedges, H., Cooper, M., 2018. Relational play-based pedagogy: Theorising a core practice in early childhood education. *Teach. Teach. Theor. Pract.* 24 (4), 369–383.
- Hedges, H., Cooper, M., 2016. Inquiring minds: theorizing children's interests. *J. Curric. Stud.* 48 (3), 303–322. <https://doi.org/10.1080/00220272.2015.1109711>.
- Hedges, H., Cullen, J., Jordan, B., 2011. Early years curriculum: funds of knowledge as a conceptual framework for children's interests. *J. Curric. Stud.* 43 (2), 185–205. <https://doi.org/10.1080/00220272.2010.511275>.
- Holzman, L., 2010. Without creating ZPDs there is no creativity. In: Connery, M.C., John-Steiner, V., Marjanovic-Shane, A. (Eds.), *Vygotsky and Creativity: A Cultural-Historical Approach to Play, Meaning Making and the Arts*. Peter Lang, Oxford, pp. 27–40.
- Moll, L., Amanti, C., Neff, D., González, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141.
- Moss, P., Urban, M., 2020. The organisation for economic co-operation and development's international early learning and child well-being study: the scores are in!, *Contemp. Issues Early Child* 18 (2), 250–258 <https://doi.org/10.1177%2F1463949120929466>.
- Morris, K., 2016. *The Posthuman Child: Educational Transformation Through Philosophy With Picturebooks*. Routledge, London.
- Pacini-Ketchabaw, V., Kind, S., Kocher, L., 2017. *Encounters With Materials in Early Childhood Education*. Routledge, Abingdon.
- Pacini-Ketchabaw, V., Taylor, A., 2015. *Unsettling the Colonial Places and Spaces of Early Childhood Education*. Routledge, Abingdon.
- Peters, S., Davis, K., McKenzie, R., 2018. Children's "working theories" as curriculum outcomes. In: Trevarthen, C., Delafield-Butt, J., Dunlop, A.-W. (Eds.), *The Child's Curriculum*. Oxford University Press, Oxford, pp. 296–310.
- Poole, A., 2017. Funds of knowledge 2.0: towards digital funds of identity. *Learn. Cult. Soc. Interact.* 13, 50–59.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press, Oxford.
- Rogoff, B., Dahl, A., Callanan, M., 2018. The importance of understanding children's lived experience. *Dev. Rev.* 50 (a), 5–15. <https://doi.org/10.1016/j.dr.2018.05.006>.
- Rogoff, B., Moore, L.C., Correa-Chávez, M., Dexter, A.L., 2015. Children develop cultural repertoires through engaging in everyday routines and practices. In: Grusec, J.E., Hastings, P.D. (Eds.), *Handbook of Socialization: Theory and Research*. The Guilford Press, pp. 472–498.
- Vygotsky, L.S., 1987. The development of scientific concepts in childhood. N. Minick, Trans.. In: Rieber, R.W., Carton, A.S. (Eds.), *The Collected Works of L. S. Vygotsky: Problems of General Psychology*. Plenum Press, New York, pp. 167–242.
- Vygotsky, L.S., 1986. *Thought and Language*. MIT Press, Cambridge, MA.
- Walker, D., Nokon, H., 2007. Boundary-crossing competence: theoretical considerations and educational design. *Mind Cult. Activ.* 14 (3), 178–195.
- Wood, E., 2014. Free choice and free play in early childhood education: troubling the discourse. *Int. J. Early Years Educ.* 22 (1), 4–18.
- Wood, E., 2020. Learning, development and the early childhood curriculum: a critical discourse analysis of the Early Years Foundation Stage in England. *J. Early Child. Res.* 18 (3), 321–336.
- Wood, E., Hedges, H., 2016. Curriculum in early childhood education: critical questions about content, coherence, and control. *Curric. J.* 27 (3), 387–405.

Children's participation in local curriculum-making

Louise G. Phillips, Southern Cross University, Gold Coast, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Origins of curriculum-making	181
Curriculum in contemporary times	181
Child as social actor	182
Children as project protagonists and collaborators	182
Children as curriculum informants	182
Children's participation in curriculum-making authorized in national curricula	183
Children as citizens in curriculum-making	184
Children in rhizomatic curriculum-making	184
Children's curriculum-making with place, home and community	185
Decentering the child in posthuman curriculum-making	185
Conclusion	185
References	186

Children's participation in curriculum-making as the co-construction of learning content is the focus of this article. As children scarcely (if at all) have a voice in national curriculum development, the examples discussed are about curriculum-making at the local level, that is, with a class or school. Beginning with Indigenous origins of curriculum-making, influential thinking on children's participation in local curriculum-making is described somewhat chronologically, noting key influences such as the conception of mass schooling, UN Convention of the Rights of the Child, Reggio Emilia preschool approach to curriculum, emergent curriculum, mosaic approach, children as citizens, rhizomatic thinking, place-based curricula, and posthumanism.

Origins of curriculum-making

Entangled relations with all entities have informed curricula making for millennia, without demarcating children, curricula and education. By looking to the world's longest living continuous culture, traditional Aboriginal Australian culture (Behrendt, 2016), in which curriculum-making is and has been a shared inclusive relational and organic activity with all entities (plants, skies, waterways, land, climate and animals) they co-exist with (Martin, 2016; Proud et al., 2017) for more than 50,000 years (Devlin, 2016). Children learn with the environment and community, alongside family members sourcing tucker (fish, meat, fruit, vegetables and bush honey) and telling stories (Marika, 2000) to come to know lore and kinship for living peacefully and sustainably with other living beings (Pascoe, 2014). Across in the lands that are now called New Mexico, Tewa science educator Gregory Cajete (1994) describes Pueblo ways-of-knowing-in-being (believed to be more than 1500 years (Duwe, 2020)) as an ecology of relationships porously extending to space and time to holistically be of the world rather than in. And in Aotearoa (now known as New Zealand), Māori, one of the youngest cultures in the world (estimated to be 700–800 years (Anderson, 2016)) enact curriculum-making in which children are accorded mana (respect) and freedom to participate within everyday activities of their whānau (family) and hapū (tribe or clan), accompanied by their Elders imparting myths and legends of the cosmos (Te Rangi Hiroa, 1950). Māori curriculum-making unfolds through responsibly knowing with nature with wise others (Elders).

In western ontologies, the conception of mass schooling from the industrial age led to the regulation of curricula through state authorities. The liveliness of curriculum-making that Bobbitt (1921) described as human activity "in the light of actual human needs" (p. 607) was perhaps more widespread across the globe through master-apprenticeship model prior to the establishment of mass schooling.

Curriculum in contemporary times

In formal schooling, the people involved are mostly students and teachers, with some contribution from families and community members. Across recent decades, curricula has shifted from input regulated to output regulated (Leat et al., 2013), that is, from specification of content to specification of skills, via measurement of performance data (e.g., see Wilkins, 2011). Such cultures of performativity (e.g., Keddie et al., 2011) erode teacher and student autonomy. Curriculum in these neoliberal agendas is firmly under the control of government education policy (Biddulph, 2011). In neoliberal societies, curriculum is nationalized, and standardized to meet global performance targets. The very people school curricula are designed for scarcely have any say or are silenced (Brooker and MacDonald, 1999).

Children's participatory contributions in curriculum-making has been minimally permitted and explored in western education, as it is widely understood and accepted that teachers have the responsibility of curriculum-making (Clandinin and Connelly, 1992) as regulated by state education authorities (Levin, 2007). Formal schooling has in the most been constructed as what Freire (1970) referred to as a banking model in which teachers narrate subject content and students passively listen as containers to be "filled." He proposed instead that such a hierarchy be countered with partnerships between students and teachers and the world through merging of roles (student-teacher and teachers-students) in problem-based curriculum-making. Curriculum studies scholarship further theorizes and advocates for greater agency in curriculum-making, through seeing curriculum as co-created inquiries (Edwards et al., 1998), a lived experience (Dewey, 1938/1998; Aoki, 2005), as an account of teachers and children's lives together in schools and classrooms (Clandinin and Connelly, 1992), as a space of emergence (Biesta, 2004; Jones and Nimmo, 1994), as a rhizomatic practice (Gough, 2007; Olsson, 2009; Chan, 2011), and as a process of living in and through the world with more-than-human entities (Ross and Mannion, 2012). These ideas of curriculum-making, open up voice and agency for children.

Child as social actor

The creation and endorsement of the United Nations Convention of the Rights of the Child (1989) has provided a strong case for children's participation in matters that affect them, in particular Article 12—right to express their views on all matters affecting them and for their views to be taken seriously. Article 13 (the child shall have the right to freedom of expression) furthers a case for children's contributions, and Article 30 (to enjoy his or her own culture, to profess and practice his or her own religion, or to use his or her own language) sanctions each child's diverse participation and contribution. Following the endorsement of the UNCRC in 1989, sociological interest in and attention to children and childhood has grown and formed burgeoning scholarship in sociology of childhood which proposes a view of "children as competent social actors" (James and Prout, 1995, p. 78) "who shape their identities, create and communicate valid views about the social world and have a right to participate in it" (MacNaughton et al., 2007, p. 460). In essence, children are understood as agential, inviting understandings of children as contributors and co-constructors of curriculum-making.

Children as project protagonists and collaborators

The preschools of the Italian province Reggio Emilia have attracted global attention and forged a movement in early childhood education known as the Reggio Emilia approach, for a shift away from a planned curriculum to projects that are co-created between children, teachers and families through a pedagogy of listening and a pedagogy of relationships (Edwards et al., 1998). Malaguzzi (1998) (founder of the philosophy that informs the Reggio Emilia preschools) and Rinaldi (2006) (world leading pedagogist on Reggio Emilia preschools) both explain that the term curriculum (and its associated terms) are not suited for representing the understanding of learning or knowledge in their practice. Instead they propose the word project to describe knowledge and learning creation at Reggio Emilia preschools. Children are understood as protagonists and collaborators in the curation of knowledge-building projects, with teachers as partners, who led curation of children's ideas, interests and activities. Though many romantically idealize the practices of the Reggio Emilia preschools, it is important to realistically recognize the limitations of what children can actually decide on. Preschools are institutions governed by local and national rules and policies, and "even in a democratically functioning preschool, the child structurally has a specifically subordinated position" (Hočevar et al., 2013, p. 485).

Children as curriculum informants

One interpretation of the Reggio Emilia preschool practices was coined emergent curriculum by Betty Jones (Jones and Nimmo, 1994) and employed in early childhood education in other countries (e.g., Australia and the US). In an emergent curriculum, "children's interests, worries, desires, understandings, and activities of learning interest expressed by both children and educators are generated as the beginning points for curriculum" relying on "developmentally appropriate" standards and "well-developed observation skills of early childhood teachers" (Wien and Stacey, 2000, p. 1) to plot in a webbed pattern as documentation of the program/pathways of learning (Nimmo & Jones). Teachers actively look for and follow the interests of the children (Birmer, 2015), documenting a curriculum as it unfolds. The use of the word curriculum is not synonymous with the intentions on the Reggio Emilia approach, as noted earlier. Emergent curriculum is one particular idea of looking to children's interests, ideas and activities as starting points for curriculum-making. Children are positioned as informants but not necessarily active decision-makers of an emergent curriculum.

A suite of research tools known as the mosaic approach developed by Clark and Moss (2001), Clark (2017) in the UK for gathering children's perspectives in the Listening to Young children's study have also been applied as mechanisms for children's input into curriculum-making. The intent of the mosaic approach as a multi-method, polyvocal approach is to "enable young children to create a 'living picture' of their lives" through a range of modes of expression (role play, photos, tours, map making) that are accessible to young children beyond reliance on the written or spoken word (Clark and Moss, 2001, p. 12). Ideas from the Reggio Emilia approach and the New Zealand early childhood curriculum Te Whāriki of listening to children and engaging with them as active

competent participants were influential in the development of the mosaic approach (Clark, 2017). The mosaic approach has come to be known as a framework that enables “teachers to become the researchers of their own class in a flexible, adaptable and fun way that respects and celebrates children’s rights and strengths” (Rouvali and Riga, 2019, p. 999) offering a conduit between adults and children to discuss and negotiate meanings (Clark and Moss, 2005). Children’s perspectives of learning in the world are worked with in curriculum-making on the grounds that this is necessary for children to identify with the educational program, to attend to Brooker’s (2002) provocation that “unless adults are alert to children’s own ways of seeing and understanding and representing the world to themselves, it is unlikely that the child will ever manage to identify with the school’s and teacher’s ways of seeing” (p. 171).

Emergent curriculum and mosaic approach are informed by social-constructivism, which foregrounds collaboration in learning and a strengths-based approach to curriculum-making. Young children are positioned as providers of knowledge and co-constructors of meaning, with teachers as facilitators or opportunity providers rather than definitive managers (Chan, 2011).

Children’s participation in curriculum-making authorized in national curricula

Early childhood curricula in some nations such as New Zealand, Australia and Canada have authorized children’s participation in curriculum-making.

In Aotearoa, New Zealand, the Te Whāriki curriculum was launched in 1996 (updated in 2017) after 5 years of comprehensive community and sector development. A 5-year development duration is a rarity in government-initiated policy, though as one of the curriculum writers, May (2012), explains “it takes time to develop and implement a curriculum that is accepted, inclusive, meaningful and makes a difference for children.” In Māori language, Te Whāriki translates to “a woven mat for all to stand on,” offering a metaphor for teachers, families and children weaving curriculum patterns shaped by different cultural perspectives, ages, philosophies for a child centered curriculum that grows from children’s learning dispositions. Te Whāriki was a landmark curriculum as it was the first bicultural curriculum in Aotearoa, New Zealand. The curriculum foregrounds Maori knowledge and rights and children’s, as Tilly Reedy (Ngati Porou) (1995), a Māori partner on the development of Te Whāriki asserted: “Our rights are recognized and so are the rights of everyone else ... Te Whāriki recognizes my right to choose, and your right to choose too” (p. 13). Children’s rights scholar, Anne Smith (2007), described Te Whāriki, as taonga (treasure) encapsulating aspirations for children based on children’s rights. The curriculum has attracted global attention, with UK esteemed early childhood provision scholar, Moss (2007) observing that through Te Whāriki, New Zealand was “leading a wave of early childhood innovation” (p. 27).

The first national Australian early childhood curricula, the Early Years Learning Framework for Australia (Australian Government Department of Education, Employment and Workplace Relations for the Council of Australian Governments, 2009) drew inspiration from Te Whāriki. The framework explicitly attends to the UNCRC broadly and Article 12 explicitly with recognition of children’s right to be active participants in all matters affecting their lives. When the framework was being developed, the first version welcomed children, educators and families as active, valued contributors to learning, with children’s perspectives and agency in decisions respected and responded to (Millei and Sumson, 2011). In the final version of the framework, children are described as active participants and decision makers, though all mentions of curriculum decision-making are firmly led by educators, such as “all children and families are respected and actively encouraged to collaborate with educators about curriculum decisions in order to ensure that learning experiences are meaningful” (p. 12) aligning more with notions of children as curriculum informants. A survey of Australian early childhood teachers in child care, kindergarten and pre-primary settings implementing the Early Years Learning Framework saw that authentic comprehensive engagement of young children in decision-making was limited, with many seeing making choices, one-off agendas as decision-making and did not mention including children in decisions that relate directly to curriculum structures (Hudson, 2012).

Another example of a national early childhood curriculum which welcomes children’s participation in curriculum-making is the British Columbia Early Learning Framework first released in 2008 and revised in 2019 (Government of British Columbia, 2019). This framework also responded to the UNCRC and sociology of childhood and is framed on a “image of the child as capable and full of potential” (p. 3). Children are also recognized as decision-makers, in fact it is the only mention of decision in the curriculum document. Curriculum is defined as emergent and responsive from the union of children, adults, ideas, and materials, constructing knowledge “in ways that are local, inclusive, ethical, and democratic” (p. 11). The British Columbia Early Learning Framework also implements the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP) to explicitly decolonize curriculum by honoring First Nation communities and knowledges and contributing to lasting reconciliation and self-determination for Métis children, families and Elders.

These progressive early childhood curriculum documents counter global emphasis on high-stakes assessment dominating schooling practice, policy and decision-making, though the uptake of children’s participation in curriculum-making is in the most found in small pockets, driven by individual teachers who advocate for children’s participation. The lack of resources, knowledge and relevant training on how to listen to children’s voices and how to include them in the decision-making process have been noted as significant hindrances (Rudduck and McIntyre, 2007), along with the continuously increasing curriculum workload, reducing teacher ability to dedicate time to listening to children and children’s participation (Rouvali and Riga, 2019).

Children as citizens in curriculum-making

Democracy as the participatory practice of citizens (Loenen, 1997) in education is not new, it has been discussed and theorized since its inception in the *polis* of Athens. An individualistic conception of democracy in education is based on the thinking of Enlightenment philosophers such as Kant (1784/1992), which translated into curriculum-making emphasizes freedom of expression and choice for individuals. A social conception of democracy in education, based on the influential text *Democracy and education* (Dewey, 1916), emphasizes group cooperation played out through both children and adults considering the direction of actions in reference to others. A political conception of democracy in education, informed by Arendt's (1958/1998) theory of action views the first step as taking initiative (Biesta, 2007). In curriculum-making, it is through actions, that people bring something new into the world that are responded to by others. It is in the interplay of initiated actions and supportive responses to those initiated actions that democracy is practiced.

The positioning of teacher authority reduces student agency and enablement of full democracy in curriculum-making. As mentioned earlier Freire (1970) opposed the banking model of education. Instead he proposed that the relationship between teacher and student be transformed to be equitable, forming a collective in which the democratically minded educator consciously critiques ethics in participation. To Freire (1998), democracy in education is a respectful practice, where teacher and students collaborate and involves teachers respecting the autonomy, identity, and knowledge of students orchestrated through cultivating a balance between freedom and authority.

To authentically embrace children's rights and citizenship in curriculum-making it is not just about making choices, but rather citizenship practices of offering ideas, discussing, deliberating and reaching consensus with groups. To reach a collective decision requires the civic art of deliberation, that is, deliberative in the sense of deliberative democracy, striving for agreement about what to do (collective-will formation) (Samuelsson, 2016). Samuelsson recognizes that attention to deliberative democratic skill building in education has in the most been minimal. Deliberative democracy in curriculum-making involves all members of a learning community having the opportunity to state claims, give reasons, listen to and reflect on others' suggestions and ideas, and work toward locating curriculum un-folding in collaboration with others. Tomas Englund (2006, 2015) offered a roadmap for deliberative democracy in curriculum-making working toward deliberative communication between students without teacher facilitation. Collective will formation processes are at the core of deliberative curriculum-making, that is, reaching joint agreement about what to do. Equitable participation of all parties (students, teachers, community members) in curriculum-making involves decision-making through collective-will formation. Deliberative curriculum-making has been found in empirical studies (e.g., see Andersson, 2012) to increase knowledge, articulation of points of view, civic capacity and participation. Children and young people are more likely to participate in the curriculum if they are active constructors of the curriculum (Enright and O'Sullivan, 2010) and their views have been included (Harris and Manatakis, 2013).

Children in rhizomatic curriculum-making

Application of Deleuze and Guattari's (1987) rhizomatic philosophy as non-hierarchical, heterogeneous, multiplicitous, and acentered to curriculum-making offers a means for organic equitable unfolding curriculum-making. Through such an ontology, preconceptions of children/students and teachers are abandoned with all participants in the learning community engaging in a dynamic ongoing negotiated process of transformational learning (Olsson, 2009). Rhizomatic curriculum-making is a shift away from essentialist curricula based on lack or deficit framing ("you don't know this or can't do this, so I need to teach you"). Curriculum-making approached from a rhizomatic onto-epistemology is multiplicitous, adventitious, nonlinear and nonhierarchical assemblages to other things. Children's desires, ideas, questions are listened to and valued, producing new and multiple thinking, action and creation. Children are not homogenously constructed, rather everyone, regardless of age, is understood and engaged with in a constant state of "becoming." The common becoming of human life is embraced. The British Columbia Early Learning Framework (Government of British Columbia, 2019) recognizes learning as rhizomatic "moving in unexpected and surprising directions as children are in relationships with people, place, ideas, and materials" (p. 25). Rhizomatic curriculum making is therefore "impossible to predict, plan, supervise or evaluate according to predefined standards" (Olsson, 2009, p. 117). Through rhizomatic curriculum-making, all involved form rhizomes with the world.

What spurs rhizomatic curriculum-making is desire, as the state of the unconscious forces and "the principal and primal force in everything, the immanent source of all production" (Chan, 2011, p. 117). Desire fuels curriculum-making produced through rhizomatic mapping. Chan suggests "we must consider the ways in which children's desires can be related to curriculum making. Doing this would require us to think, question, and critically analyze our ways of knowing and to open ourselves to otherness, complexity, and multiplicity" (p. 118). The binary of teacher and student is dissolved, and rather the collective is understood as an education community (children, families, educators, community members) "working together in an ongoing process of constructing and reconstructing the problem" (Chan, 2011, p. 119). Decision-making participation of all parties regardless of age across all domains is welcomed, as opposed to behaviorist model curriculum-making in which teachers emphasize the importance of getting children to make the *right* decisions to follow social and community rules (Hudson, 2012).

Children's curriculum-making with place, home and community

Attention to children and young people's curriculum-making beyond classrooms is growing. Huber et al. (2011) reconceptualize children's curriculum-making as inclusive of family and community members in home and community places, decentering curriculum-making as school-centric. In essence they propose curriculum of lives featuring children, teacher, family and community member identities. Ross and Mannion (2012) applying Tim Ingold's (2000) notion of dwelling and see curriculum-making as a process of living in and through the world with more-than-human entities, beginning "with a concern with the material context of learning and/or the lived experience of participants" (p. 304), so that curriculum is a lived story.

Many across the globe identify with home and community-based curriculum-making and place-based curriculum-making. Through growing movements of un-schooling (learning through life at home and in the community) and place-based education, that may be enacted through forest schools, place-responsive pedagogies, place-conscious pedagogies and environmental education centers. Forest schools see curriculum-making unfolding between child and nature explorations. Place responsive pedagogies create curricula through teachers collaborating with students and others, as historically embodied subjects, who explicitly seek to create new place-based practices and place-based relations (Ross and Mannion, 2012). Place-conscious pedagogies produce curricula through educators with an intimate knowledge of the ecology, history and pedagogical possibilities of the place, engaging students in sharing, questioning, and inquiring (Renshaw and Tooth, 2018). In essence these approaches see curriculum-making unfolding through children's interactions with places, home and community. Children's agency is foregrounded, though not necessarily through explicit collective decision-making.

Decentering the child in posthuman curriculum-making

A posthuman approach to curriculum-making seeks to decenter the child, and more broadly the human centric focus. Fikile Nxumalo proposes an uptake of "both critical posthumanisms and Indigenous relational ontologies that suggest we cannot continue with the universalized, individual human developing child as the center of what we do" (in Nxumalo et al., 2020, p. 199). Rather posthuman curriculum-making invites embracing learning with and from other entities, be they living or inanimate, following their rhythms, instead of all non-human matter being there to serve humans in the education project. Quantum physicist and feminist philosopher Karen Barad (2000) prompts ethical considerations in the decentering of the human in curriculum-making by asking: "what kind of (curriculum and) pedagogy help students (and teachers) to learn about practicing responsible science?" (p. 239) that does not continue colonizing, dispossessing and manifesting ecocide. Barad (2007) offers a theory of agential realism, which applied to curriculum-making, would see curriculum and all other participants emerge in intra-actions. Curriculum and participants are not predefined, but rather they become defined in actions of "learning how to intra-act responsibly within the world" (Barad, 2000, p. 237).

In precarious times of climate crises and global pandemics, sustenance for planet Earth is prime. Kroeger and Myers (2019) propose that curriculum enhances children's understanding and centrality with Earth, so rather than placing the child at the center of the curriculum (such as developmentally appropriate curriculum, Reggio Emilia approach, emergent curriculum, mosaic approach), it is the Earth that is at the center. Drawing from critical, feminist, Indigenous, post-colonial, post-foundational and post-human ideas, they propose that curriculum-making is a more-than-human activity of reciprocal relations between Earth, Elders and children. And to attend to the global crises, the United Nations asserts that it is necessary that children and youth are actively included in international communities "to achieve peace, security, justice, climate resilience, and sustainable development for all" (Clark et al., 2020, pp. 616–617).

Conclusion

Batched approaches to teacher-led curriculum are not working as demonstrated through the most frequent unproductive behavior evidenced by teachers being disengaged behavior (Sullivan et al., 2014). Children need to feel connected to the curriculum to be motivated to learn, otherwise they switch off from teacher directed learning and partake in disruptive practice (Rudduck and Flutter, 2000). Children feel connected to the curriculum if they contribute to curriculum-making so that the curriculum aligns with their interests, needs, strengths, capabilities and pace (Subban, 2016). They then are more likely to experience school education as relevant, valuable preparation for life. Education prior to formal schooling often has greater freedoms to collaborate with children and families to collectively make curricula, as in the examples provided by Quintero's (2015) book on child-led participatory curriculum. Compulsory schooling years are regulated by state written standardized curricula, restricting school freedoms to hear from the very people education is designed for through children's participation in curriculum-making. Independent and community governed schools are more likely to welcome children's participation in curriculum-making. The COVID-19 global pandemic has interrupted and shaken education in monumental ways that require reconceptualizing education and curriculum-making so that the voice of children is heard to ensure curricula is relevant and customized to their precarious and precious lives.

References

- Andersson, K., 2012. *Deliberativ undervisning. En empirisk studie* [Deliberative teaching: An empirical study]. Göteborg Studies in Politics 128. Göteborg.
- Anderson, A., 2016. *The First Migration. Māori Origins 3000BC–AD1450*. Bridget Williams Books.
- Aoki, T.T., 2005. Curriculum in a new key. In: Pinar, W., Irwin, R. (Eds.), *Curriculum in a New Key: The Collected Works of Ted T. Aoki*. Erlbaum.
- Arendt, H., 1958/1998. *The Human Condition*, second ed. The University of Chicago Press.
- Australian Government Department of Education Employment and Workplace Relations, 2009. *Belonging, Being and Becoming: The Early Years Learning Framework for Australia*.
- Australian Government Department of Education, Employment and Workplace Relations for the Council of Australian Governments. <https://education.gov.au/early-years-learning-framework>.
- Barad, K., 2000. Reconceiving scientific literacy as agential literacy. In: Reed, R., Traweek, S. (Eds.), *Doing Science + Culture*. Routledge, pp. 221–258.
- Barad, K., 2007. *Meeting the universe halfway: quantum physics and the entanglement of matter and meaning*. Duke University Press.
- Behrendt, L., 2016. Indigenous Australians know we're the oldest living culture—it's in our dreamtime. *Guardian*. <https://www.theguardian.com/commentisfree/2016/sep/22/indigenous-australians-know-were-the-oldest-living-culture-its-in-our-dreamtime>.
- Biddulph, M., 2011. Articulating student voice and facilitating curriculum agency. *Curric. J.* 22 (3), 381–399.
- Biermeier, M., 2015. Inspired by Reggio Emilia: emergent curriculum in relationship-driven learning environments. *Young Child.* 70 (5), 72–79. <https://www.naeyc.org/resources/pubs/yc/nov2015/emergent-curriculum>.
- Biesta, G., 2004. Mind the gap! Communication and the educational relation. In: Bingham, C., Sidorkin, A.M. (Eds.), *No Education Without Relation*. Peter Lang, pp. 11–22.
- Biesta, G., 2007. Education and the democratic person: towards a political conception of democratic education [Electronic Version]. *Teach. Coll. Rec.* 109, 740–769.
- Bobbitt, F.W., 1921. A significant tendency in curriculum-making. *Elem. Sch. J.* 21, 607–615.
- Brooker, R., Macdonald, D., 1999. Did we hear you? Issues of student voice in a curriculum innovation. *J. Curric. Stud.* 31 (1), 83–97.
- Brooker, L., 2002. *Starting School—Young Children Learning Cultures*. Open University Press, Buckingham.
- Cajete, G., 1994. *Look to the Mountain: An Ecology of Indigenous Education*. Kivaki Press, Durango, CO.
- Chan, K.H., 2011. Rethinking children's participation in curriculum making: a rhizomatic movement. *Int. Crit. Childh. Pol. Stud.* 4 (1), 107–122.
- Clandinin, D.J., Connelly, F.M., 1992. Teacher as curriculum maker. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 363–401.
- Clark, A., Moss, P., 2001. *Listening to Children: The Mosaic Approach*. National Children's Bureau, London.
- Clark, A., Moss, P., 2005. *Spaces to Play: More Listening to Young Children Using the Mosaic Approach*. National Children's Bureau, London.
- Clark, H., Coll-Seck, A.W., Banerjee, A., Peterson, S., Dalglish, S.L., Ameratunga, S., et al., 2020. A Future for the World's Children? A WHO–UNICEF–Lancet Commission. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(19\)32540-1/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(19)32540-1/fulltext).
- Clark, A., 2017. *Listening to Young Children, Expanded Third Edition: A Guide to Understanding and Using the Mosaic Approach*. ProQuest Ebook Central.
- Deleuze, G., Guattari, F., 1987. *A Thousand Plateaus: Capitalism and Schizophrenia* (B. Massumi, Trans.). University of Minnesota Press.
- Devlin, H., 2016. Indigenous Australians most ancient civilization on Earth, DNA study confirms. *Guardian*. <https://www.theguardian.com/australia-news/2016/sep/21/indigenous-australians-most-ancient-civilisation-on-earth-dna-study-confirms>.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. The Macmillan Company, New York.
- Dewey, J., 1938/1998. *Experience and Education: The 60th Anniversary Edition*. Kappa Delta, Indianapolis, Indiana.
- Duwe, S., 2020. *Tewa Worlds: An Archaeological History of Being and Becoming in the Pueblo Southwest*. The University of Arizona Press.
- Edwards, C., Gandini, L., Forman, G. (Eds.), 1998. *The Hundred Languages of Children: The Reggio Emilia Approach—Advanced Reflections*, second ed. Ablex Pub. Corp, Greenwich, Conn, pp. 179–198.
- Englund, T., 2006. Deliberative communication: a pragmatist proposal. *J. Curric. Stud.* 38 (5), 503–520.
- Englund, T., 2015. Toward a deliberative curriculum? *Nordic J. Stud. Educ. Pol.* 1, 26558. <https://doi.org/10.3402/nstep.v1.26558>.
- Enright, E., O'Sullivan, M., 2010. "Can I do it in my Pyjamas?" Negotiating a physical education curriculum with teenage girls. *Eur. Phys. Educ. Rev.* 16 (3), 203–222. <https://doi.org/10.1177/1356336X10382967>.
- Freire, P., 1970. *Pedagogy of the Oppressed* (M. Bergman Ramos, Trans.). Seabury Press.
- Freire, P., 1998. *Pedagogy of Freedom: Ethics, Democracy and Civic Courage* (P. Clarke, Trans.). Rowman & Littlefield.
- Gough, N., 2007. Changing Planes: rhizomatic play in transnational curriculum inquiry. *Stud. Philos. Educ.* 26 (3), 279–294.
- Government of British Columbia, 2019. *British Columbia Early Learning Framework*. <https://www2.gov.bc.ca/gov/content/education-training/early-learning/teach/early-learning-framework>.
- Harris, P., Manatakis, H., 2013. *Children's Voices: A Principled Framework for Children and Young People's Participation as Valued Citizens and Learners*. University of South Australia in partnership with the South Australian Department for Education and Child Development, Adelaide.
- Hiroa, T.R., 1950. *The Coming of the Maori*. Maori Purposes Fund Board. Whitcomb and Tombs, Wellington, NZ.
- Hočevar, A., Sebart, M.K., Stefanc, D., 2013. Curriculum planning and the concept of participation in the Reggio Emilia pedagogical approach. *Eur. Early Child. Educ. Res. J.* 21 (4), 476–488. <https://doi.org/10.1080/1350293X.2013.845437>.
- Huber, J., Murphy, M.S., Clandinin, D.J., 2011. *Places of Curriculum Making: Narrative Inquiries Into Children's Lives in Motion*. Emerald Publishing Ltd.
- Hudson, K., 2012. Practitioners' views on involving young children in decision making: challenges for the children's rights agenda. *Aust. J. Early Childh.* 37 (2), 4–9.
- Ingold, T., 2000. *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. Routledge.
- James, A., Prout, A., 1995. Hierarchy, boundary and agency: toward a theoretical perspective in childhood. *Sociol. Stud. Child.* 7, 77–101.
- Jones, E., Nimmo, J., 1994. *Emergent Curriculum*. NAEYC.
- Kant, I., 1784/1992. An answer to the question: what is enlightenment? In: Waugh, P. (Ed.), *Postmodernism: A Reader*. Edward Arnold, pp. 89–95.
- Keddie, A., Mills, M., Pendergast, D., 2011. Fabricating an identity in neo-liberal times: performing schooling as "number one". *Oxf. Rev. Educ.* 37, 75–92.
- Kroeger, J., Myers, C.Y., 2019. Introduction: why nurture nature and the environment with young children? In: Kroeger, J., Myers, C.Y., Morgan, K. (Eds.), *Nurturing Nature and the Environment With Young Children: Children, Elders, Earth*. Taylor & Francis Group, pp. 1–8.
- Leat, D., Livingston, K., Priestley, M., 2013. Curriculum deregulation in England and Scotland—different directions of travel? In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom Across Europe*. CIDREE Yearbook. SLO Netherlands Institute for Curriculum Development, pp. 229–248. <http://www.slo.nl/organisatie/recentepublicaties/cidree2013/>.
- Levin, B., 2007. Curriculum policy and the politics of what should be learned in schools. In: Connelly, F.M., He, M.F., Phillion, J.A. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage Publications, pp. 7–24.
- Loenen, N., 1997. *Citizenship and Democracy*. Dundurn Press.
- MacNaughton, G., Hughes, P., Smith, K., 2007. Young children's rights and public policy: practices and possibilities for citizenship in the early years. *Child. Soc.* 21 (6), 458–469.
- Malaguzzi, L., 1998. History, ideas, and basic philosophy: an interview with Lella Gandini. In: Edwards, C., Gandini, G., Forman, L. (Eds.), *The Hundred Languages of Children: The Reggio Emilia Approach to Early Childhood Education*. Ablex Publishing, pp. 49–97.
- Marika, R., 2000. *Milthun ilthun Latju Wana Romgu Yolnu*: Valuing Yolnu knowledge in the education system. *TESOL in Context* 10 (2), 45–52.
- Martin, K., 2016. Aboriginal worldview, knowledge frameworks and processes. In: Martin, K. (Ed.), *Voices Visions: Aboriginal Early Childhood Education in Australia*. Pademelon Press, pp. 3–8.

- May, H., 2012. Te Whāriki: An Early Childhood Curriculum for Inclusion in Aotearoa-New Zealand. *Heimatkunde*. <https://heimatkunde.boell.de/de/2012/08/01/te-whariki-early-childhood-curriculum-inclusion-aotearoa-new-zealand>.
- Millei, Z., Sumsion, J., 2011. The "work" of community in belonging, being and becoming: the Early Years Learning Framework for Australia. *Contemp. Issues Early Child.* 12 (1), 71–85.
- Ministry of Education, New Zealand, 2017. Te Whāriki: He whāriki mātauranga mō ngā mokopuna o Aotearoa Early Childhood Curriculum. <https://www.education.govt.nz/assets/Documents/Early-Childhood/ELS-Te-Whariki-Early-Childhood-Curriculum-ENG-Web.pdf>.
- Moss, P., May 2–7, 2007. Leading the wave. New Zealand in an international context. Travelling pathways to the future - Ngā huarahi arataki. Early childhood education symposium proceedings 2–3. Ministry of Education, Wellington, pp. 27–36.
- Nxumalo, F., Diaz-Diaz, C., Semenec, P., 2020. Interview with Fikile Nxumalo. In: *Posthumanist and New Materialist Methodologies: Research After the Child*. Springer, Singapore, pp. 195–203.
- Olsson, L.M., 2009. Movement and Experimentation in Young Children's Learning: Deleuze and Guattari in Early Childhood Education. Routledge.
- Pascoe, B., 2014. *Dark Emu Black Seeds: Agriculture or Accident*. Magabala Books, Broome, WA.
- Proud, D., Lynch, S., á Beckett, C., Pike, D., 2017. "Muck-about": aboriginal conceptions of play and early learning. In: Lynch, S., á Beckett, C., Pike, D. (Eds.), *Multidisciplinary Perspectives on Play From Birth and Beyond*. Springer Nature, pp. 79–91.
- Quintero, E.P., 2015. Storying Learning in Early Childhood: When Children Lead Participatory Curriculum Design, Implementation, and Assessment. Peter Lang, New York, NY.
- Reedy, T., 1995. Knowledge and power set me free. In: *Proceedings of the Sixth Early Childhood Convention*, vol. 1. Convention Committee, Auckland, pp. 13–32.
- Renshaw, P., Tooth, R. (Eds.), 2018. *Diverse Pedagogies of Place*. Routledge.
- Rinaldi, C., 2006. In *Dialogue With Reggio Emilia: Listening, Researching and Learning*. Routledge.
- Ross, H., Mannion, G., 2012. Curriculum making as the enactment of dwelling in places. *Stud. Philos. Educ.* 31 (3), 303–313. <https://doi.org/10.1007/s11217-012-9295-6>.
- Rouvali, A., Riga, V., 2019. Redefining the importance of children's voices in personal social emotional development curriculum using the Mosaic Approach. *Educ.* 3–13 47 (8), 998–1013. <https://doi.org/10.1080/03004279.2018.1553990>.
- Rudduck, J., Flutter, J., 2000. Pupil participation and pupil perspective: "carving a new order of experience". *Camb. J. Educ.* 30 (1), 75–89. <https://doi.org/10.1080/03057640050005780>.
- Rudduck, J., McIntyre, D., 2007. *Improving Learning through Consulting Pupils*. Routledge.
- Samuëlsson, M., 2016. Education for deliberative democracy. A typology of classroom discussions. *Democr. Educ.* 24 (1), 1–9.
- Smith, A.B., 2007. Children and young people's participation rights in education. *Int. J. Child Right.* 15 (1), 147–164.
- Subban, P., 2016. Australian students are becoming increasingly disengaged at school—here's why. *Conversation*. <https://theconversation.com/australian-students-are-becoming-increasingly-disengaged-at-school-heres-why-51570>.
- Sullivan, A.M., Johnson, B., Owens, L., Conway, R., 2014. Punish them or engage them? Teachers' views of unproductive student behaviors in the classroom. *Aust. J. Teach. Educ.* 39 (6). <https://doi.org/10.14221/ajte.2014v39n6.6>.
- Wien, C.A., Stacey, S., 2000. Spring conference on emergent curriculum. In: *Certification Council of Early Childhood Educators of Nova Scotia*.
- Wilkins, C., 2011. Professionalism and the post-performative teacher: new teachers reflect on autonomy and accountability in the English school system. *Prof. Dev. Educ.* 37, 389–409.

Curriculum making and teacher agency

Mark Priestley^a, Daniel Alvunger^b, Stavroula Philippou^c, and Tiina Soini^d, ^a Faculty of Social Sciences, University of Stirling, Stirling, United Kingdom; ^b Linnaeus University, Växjö, Sweden; ^c University of Cyprus, Nicosia, Cyprus; and ^d Tampere University, Tampere, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	188
Curriculum making	189
Teacher agency	191
Teacher agency and curriculum making	192
Teacher agency for curriculum making	193
Curriculum making for teacher agency	194
Concluding thoughts	195
References	196

Introduction

While education has traditionally been utilized for instrumental social and political reasons by governments worldwide, the 21st century has heralded an intensification of such tendencies, especially visible in profound changes to the framings of school curriculum policy. These changes, characterized as the “new curriculum” (Priestley and Biesta, 2013), have been transnational in scope, driven by significant discourses produced and disseminated via major institutional actors on the world stage, such as the OECD, the European Union and UNESCO (e.g., Zapp, 2019; Lingard, 2021). For example, the OECD’s *Education 2030* initiative has a strong focus on curricular purposes, relevance and currency in a fast-changing world, and the forms of knowledge and competence required in that world (OECD, 2018). The discourses at the heart of *Education 2030* both drive and reflect emerging policy trajectories worldwide, as its publications have emerged from discussions involving key education stakeholders from many countries, including policymakers. Key policy themes associated with new curricula include common approaches to curricular regulation, focus, and school autonomy.

In relation to curricular regulation, there have emerged changed methodologies for the mobilization of policy to frame practice in schools. This is most evident in the shift from input regulation to output regulation (Nieveen and Kuiper, 2012; Leat et al., 2013). The former relates to specification of content to be taught; the latter concerns the measurement of outputs, for example using attainment data and performance indicators to evaluate the effectiveness of schools and teachers. This changed emphasis is strongly associated with changes to the technical form of the curriculum (Luke et al., 2012). National curricula now tend to be presented as highly technical curriculum frameworks, articulated through learning outcomes structured into linear and hierarchical levels, which can be utilized to evaluate the performance of individual students, as well as schools and teachers. These policy approaches have been heavily critiqued for contributing to a performative culture in schools and eroding teacher autonomy (e.g., Kelly, 2008; Luke et al., 2012; Priestley et al., 2015). Changes in the form of curricular regulation are linked to changes to the focus of the new curricula. A shift from the specification of substantive knowledge/content to genericism has been well-documented and subject to criticism by many writers (e.g., Young, 2007; Yates and Collins, 2010; Young and Muller, 2010). This trend has been accompanied by a strong emphasis on skills/competencies. Sinnema and Aitken (2013) point to transnational commonalities in curricula. They highlight common goals for curriculum policy, including the use of curriculum as a lever for improvement and to promote equity. They identify common new emphases in policy: competencies, values, pedagogy, student agency, partnerships and reduced prescription. This final emphasis is an important – and contradictory – aspect of new curricula. On the one hand, there has been strong rhetoric about reducing prescription, enhancing school autonomy, empowering schools and teachers, and promoting local choice in building curricula in schools (e.g., OECD, 2018; ICEA, 2018); on the other hand, and as noted above, the recourse to output regulation – including the use of curricular frameworks to monitor performance – has the potential to diminish teachers’ professional agency, and to narrow the scope for local curriculum making as a creative and innovative pedagogical process.

In the above-described context, two concepts assume great importance. These are curriculum making and teacher agency. Curriculum making – the idea that the curriculum is constructed in different educational contexts such as policymaking arenas, national and local agencies, schools and classrooms – has a long history, for instance being coined by Bobbitt (1921) over a hundred years ago, albeit with a very different meaning to its later theorization, which is also the focus of this chapter. Bobbitt used the term in a way closer to the technical manner described above and which prevails in curriculum policy today, with a strong focus on a “scientific method” in education. Teacher agency is a relatively new concept, which has become prominent in the past decade in different forms (e.g., Pyhältö et al., 2014; Priestley et al., 2015). The first concept is important insofar as the changing discursive and policy landscape is opening up new possibilities for curriculum practice; this requires new understandings of how curriculum, as a set of social practices, is made in different settings. The latter idea is a vital accompaniment to the first, as strongly developed professional

agency is required to maximize possibilities for constructive curriculum making in those different settings. In this chapter, we will first provide an overview of the concept of curriculum making, including some recent conceptual developments that have some resonance for the changing curricular landscape described above. We then explore the concept of teacher agency, offering an ecological understanding of agency as something that is achieved in specific situations. This is a temporal/relational view of agency, which is highly relevant to the creation of propitious conditions for curriculum making. Finally, we conclude the chapter by examining the complex interplay of the two concepts; both the implications of teacher agency for curriculum making, and vice versa (as constructive engagement by teachers in curriculum making across different sites is a means for the development of teacher agency).

Curriculum making

We first explore the concept of curriculum making. The discussion that follows relates to different types of curriculum making across different settings within education systems; while the primary focus of the chapter is on curriculum making by teachers, and their professional agency, it is important to locate these issues within a wider systemic understanding of curriculum making. We preface this section with three key ideas, before exploring the concept of curriculum making. The first is that the school curriculum can be defined as “the multi-layered social practices, including infrastructure, pedagogy and assessment, through which education is structured, enacted and evaluated” (Priestley, 2019, p. 8). We wish to draw attention to three linked dimensions of this definition. The first is the idea of curriculum as social practice; in other words, curriculum is something that is done, or made, by people working with one another in particular social settings of schooling. The second dimension is that the curriculum is made in different ways within and across various layers or sites of education systems, for example schools and national agencies, and is not confined to classrooms. The third dimension concerns the different practices which comprise curriculum, such as selection of knowledge/content, pedagogical approaches and the organization of teaching.

The second idea is that curriculum making is systemic. According to Connolly, curriculum:

... is a complex system involving teachers, students, curricular content, social settings, and all manner of impinging matters ranging from the local to the international. It is a system that needs to be understood systemically. The question is not which of the various factors explain high achievement, the current crime-solving model at work in the literature, but, rather, how it all works together.

Connolly (2013, p. ix).

This implies that changes to one part of the system will inevitably exert effects (often unintended) on other parts of the system.

The third idea is that practitioners mediate curriculum policy, even within the most prescriptive and “teacher-proof” policy framings. This process is subject to many variables, for example beliefs (Priestley et al., 2015), school culture (Cuban, 1998), the nature of their networks (Hizli Alkan and Priestley, 2019), external demands such as accountability, the resources available to teachers (Luke et al., 2012), and notions of teacher professionalism (Kontovourki et al., 2018). Mediation can be constructive, focused on educational purposes, local conditions and the needs of students; or it can be more subversive and instrumental (e.g., Bowe et al., 1992; Osborn et al., 1997; Supovitz, 2008; Braun et al., 2011). This sort of active curriculum making is both inevitable and a desirable activity for professionals. As Doyle (1992, p. 69) suggested, it is important that curriculum making is “a deliberative process of interpretation, judgment and responsibility”.

The above ideas – curriculum as social practice, curriculum as a system, and the ongoing mediation of curriculum by practitioners – pose a serious challenge to prevailing linear product/implementation/delivery conceptions of curriculum that hold sway in many education systems. In particular, they point to the importance of thinking holistically about how policy is framed, and enactment supported. We will come back to this issue in the later sections of the chapter, but first we will explore some ideas about how curriculum operates as a system.

There is a rich vein of literature which has sought to theorize the ways in which curriculum is made systemically. Goodlad’s (1979) seminal work suggests three “levels” of curriculum – the societal, institutional, instructional (he later added a fourth level: the personal). Goodlad’s model was criticized as linear and as hierarchical, partly due to the term “levels”, and despite his emphasis on the transactional (negotiation) and two-way flows of communication, and the positioning of the levels as institutional sites, rather than as forms of activity. Later writers have sought to rework this thinking to address these critiques, but essentially keeping the conceptual framing of levels. Doyle (1992) posited two levels: institutional and classroom levels that represent “distinct but interrelated domains of curriculum discourse, and thus, different types of curriculum processes” (p. 69). Doyle’s institutional level comprises two aspects: 1] the abstract or ideal, a conversation connecting schooling and society; and 2] the analytical or formal, which produces practical curriculum programmes. In his view, however, this is not a linear or hierarchical process; Doyle saw teachers as active “curriculum makers who guide students through the texts, shape the interpretations that are allowed on the floor, and, importantly, define the tasks that students are to accomplish with respect to these texts” (p. 76). More recent writing by Deng (2012) posits three levels: [1] societal (ideal or abstract); [2] programmatic (technical or official); [3] classroom (enacted). Deng similarly does not see this as a linear process; in his view, translation across levels means that a school subject, for example, is not the same as the discipline[s] it derives from, and “classroom curriculum making entails transforming the programmatic curriculum (embodied in curriculum materials) into ‘educative’ experiences for students” (Deng et al., 2013, p. 7).

An alternative approach to defining curriculum making levels has been developed by [Thijs and van den Akker \(2009\)](#). This is a more fine-grained typology, which in a sense challenges curriculum-as-intended/prescribed versus curriculum-as-enacted/experienced binaries, utilizing the following terminology: supra (international); macro (national/system); meso (school/institute); micro (classroom/teacher); and nano (pupil/individual). However, the primary focus of this approach is on curriculum products (e.g., frameworks, programmes of work, etc.) rather than on processes. It can be critiqued on the grounds that it reifies the institutional settings; it does not allow, for example, easy theorization of how actors situated within a particular level (e.g., school) might contribute to curriculum making in other levels (e.g., teachers as curriculum policymakers). Nor, in common with earlier and less fine-grained theorizations explored above, does it easily take account of curricular components such as professional support infrastructure, which sit between schools and policy making, and which are key components in supporting effective curriculum making in schools.

In our recent work on curriculum making ([Priestley et al., 2021b](#)), we have sought to elaborate these approaches, utilizing the Thijs and van den Akker typology (supra, macro, etc.), while adapting it to reflect the insight that curriculum making, as social practice, occurs within different sites of activity across the system. We offer the following caveats. First, the typology should not be read as a hierarchy of levels or layers. Instead, it is a heuristic to analyze curriculum making from the inside out or bottom up, as well as top down, illustrating two way flows of influence, information and activity between the various layers. Second, we do not position sites as institutional levels or even institutional sites of formal or prescribed activity. Instead, these are types of curriculum practice that occur across education systems, as curriculum is made and remade in different contexts. These sites engage various social actors, often moving between sites, and/or present in more than one site, or actors adopting multiple identities, for example teachers as teacher educators (e.g., subject-area counsellors in Cyprus; see [Kontovourki et al., 2021](#)), or teachers as members of curriculum policy committees (e.g., [Theodorou et al., 2017](#)). This heuristic is illustrated in the diagram in [Fig. 1](#), below, showing the sites of activity (e.g., meso), examples of actors (e.g., teachers), and types of activity (e.g., lesson planning).

Curriculum making comprises many types of social practice in distinct contexts within an education system; in Westbury's (2008, p. 49) words, "loosely coupled settings for curriculum decision making are in fact contexts associated with very different activities". These practices produce various forms of curriculum in diverse settings, for example prescribed, described, enacted, and received, and occur for multifarious purposes, for example discourse generation, policy production, program design, pedagogy.

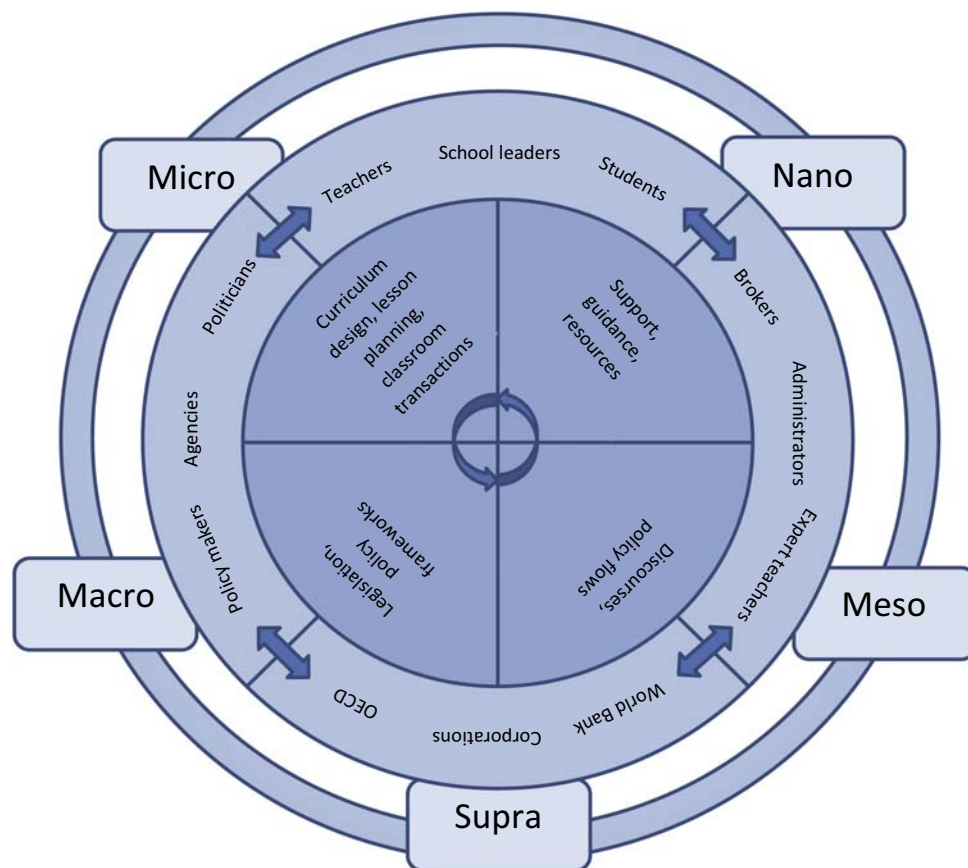


Fig. 1 Sites, actors and activities. Source: [Alvunger et al. \(2021\)](#).

There might be two-way or even multiple-way relationships between or cutting across different sites of activity. For example, policy making may be shaped by both supra discourses and by diverse local imperatives; curriculum making in schools is likely to be shaped by a variety of mediating conditions, only one of which may be a top-down policy prescription.

The focus on sites of activity rather than institutional settings, means that the heuristic is not a normative framework stating how curriculum making should occur; instead, it is an analytical tool for understanding how curriculum making occurs and emerges in different contexts – and for understanding both the considerable variation from country to country and the historicized complexity from context to context within countries. This is, as we shall demonstrate in later sections of the chapter, important and perhaps necessary when we theorize teachers as social actors and consider how teacher agency might be enhanced.

Teacher agency

We next examine the concept of teacher agency. This is an important concept in relation to curriculum making, for the reasons already outlined: the inevitability (and indeed desirability) of teachers mediating curriculum policy; the need for teachers, in the words of Doyle (1992, p. 77), to be “responsible curriculum theorists”; and recent contradictory trends in curriculum policy to position teachers as agents of change and active curriculum makers, while simultaneously undermining the conditions within/through which they can achieve agency. The latter point thus needs to be understood against a backdrop of complexity in teachers’ professional lives, particularly tensions within and between policies, which create dilemmas, encourage teachers to play safe (e.g., instrumental selection of content, teaching to the test, etc.) and effectively reduce teacher agency (e.g., see Priestley et al., 2015; Smith, 2019). The ecological understanding of agency (Priestley et al., 2015), which we expound in the following paragraphs, offers a practical approach to understanding this problem.

This approach to understanding agency approach differs from traditional sociological accounts of agency in significant ways. First, the ecological understanding posits agency as an emergent phenomenon, rather than as a variable in social action, as it is often characterized in the longstanding structure/agency debate.

[T]his concept of agency highlights that actors always act *by means of* their environment rather than simply in their environment [so that] the achievement of agency will always result from the *interplay of individual efforts, available resources and contextual and structural factors as they come together in particular and, in a sense, always unique situations.*

Biesta and Tedder (2007, p. 137); emph. added.

Agency is, therefore, not something that a person possesses; it is something that people do or, more precisely, something they achieve (Biesta and Tedder, 2007). In this understanding, agency is thus a “quality” of the engagement of actors with temporal-relational contexts-for-action, rather than an innate quality of the actors themselves.

This ecological understanding of agency draws heavily on the social theory of Emirbayer and Mische (1998), which suggests that agency is always shaped by the dynamic interplay of three temporal dimensions – influences from the past, orientations toward the future and engagement with the present – and which takes into consideration “how this interplay varies within different structural contexts of action” (Emirbayer and Mische, 1998, p. 963). Emirbayer and Mische refer to these three dimensions as the iterative, the projective and the practical-evaluative. All three dimensions play a role in social action, but the degree to which they contribute varies, thus agency may be oriented more or less, and variably over time and across different contexts, toward past, present or future influences. Emirbayer and Mische refer to these variable influences as a “chordal triad of agency within which all three dimensions resonate as separate but not always harmonious tones” (Emirbayer and Mische, 1998, p. 972; emphasis in original). Thus, they define agency as “the temporally constructed engagement by actors of different structural environments – the temporal-relational contexts of action – which, through the interplay of habit, imagination, and judgement, both reproduces and transforms those structures in interactive response to the problems posed by changing historical situations” (Emirbayer and Mische, 1998, p. 970; emph. in original). Agency therefore manifests as a “temporally embedded process of social engagement, informed by the past (in its habitual aspect), oriented toward the future (as a capacity to imagine alternative possibilities) and ‘acted out’ in the present (as a capacity to contextualize past habits and future projects with the contingencies of the moment)” (Emirbayer and Mische, 1998, p. 963).

Priestley et al. (2015) have applied this ecological understanding of agency to the professional work of teachers (see Fig. 2 below).

This diagram illustrates the key dimensions of teacher agency, enabling an analytical separation of the key elements of each dimension in the chordal triad. In respect of the iterative dimension, the model draws attention to the influence of the more general life histories of teachers and their more specific professional histories (which include both their own education as a teacher and the accumulated experience of being a teacher). This draws attention to the fact, as we earlier stated, that while agency is not merely personal capacity, such capacity – for instance beliefs and values, professional knowledge and confidence – is an important component of teacher agency. This in turn draws attention to the importance of good teacher education and rich professional experience to foster future agency.

The projective dimension of teacher agency relates to teachers’ abilities to form professional aspirations – both long term and short term. Such aspirations may be positive in nature, relating to the development and welfare of students (Lasky, 2005), and lead to agency that is protective of students’ interests. Future oriented agency may support policy intentions, or it may act in resistance to

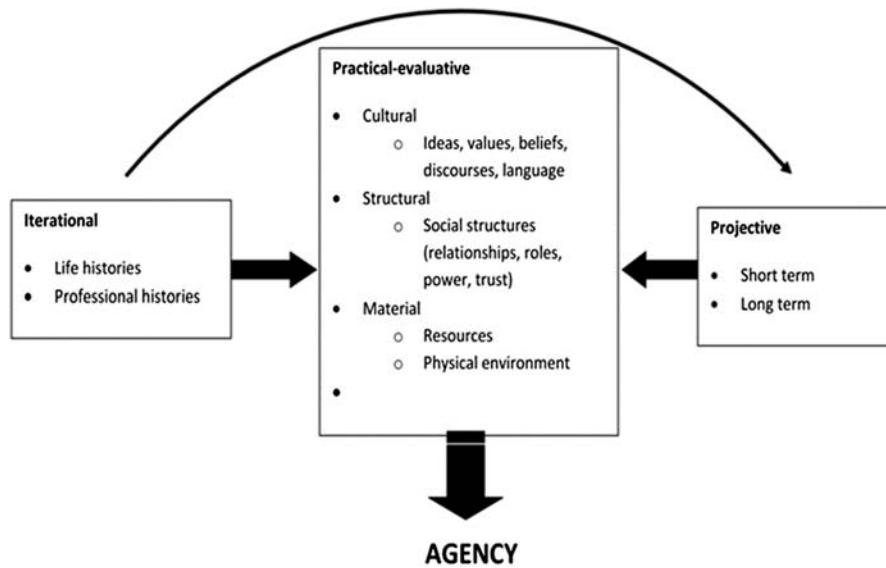


Fig. 2 An ecological understanding of teacher agency. Source: [Priestley et al. \(2015\)](#).

them ([Osborn et al., 1997](#)). Conversely, aspirations may be more narrowly instrumental, for example underpinning selection of content to support teaching to the test ([Smith, 2019](#)), or geared to maintaining a “normal desirable state” in the classroom ([Brown and McIntyre, 1992](#)).

The iterational and projective dimensions of agency are anchored in past experience, and to a large extent involve personal capacity. The practical-evaluative dimension, conversely, is more directly related to the present context – that is, the day-to-day working environment within which teachers work – by means of which agency is achieved. Teaching is a profession rife with ambiguity and professional dilemmas (e.g., [Helsby, 1999](#)), context dependent and contestable in terms of its aims (e.g., [Kontovourki et al., 2018](#)). Teachers make daily decisions that are difficult, often conflicting with their aspirations, and often with insufficient time to reflect and to engage in professional dialogue with colleagues. The practical-evaluative dimension powerfully shapes agency. This is practical, both offering affordances for agency (for example, making available resources) and inhibiting it (for example, imposing barriers to action through the exercise of coercive power). It is also evaluative, for instance being shaped by judgments relating to risk and cost/benefit. The practical-evaluative dimension of agency involves various types of affordances and constraints: cultural (e.g., access to cognitive resources, ideas); structural (e.g., relational resources available via teacher networks); and material (e.g., the physical layout of schools and classrooms).

Here, we have set out a brief summary of some key aspects of an ecological understanding of teacher agency. For a more detailed account of this conceptualization of teacher agency, readers should refer to [Priestley et al. \(2015\)](#). We now proceed to the concluding section of this chapter, where we further explore the dynamic relationship between the twin concepts of teacher agency and curriculum making.

Teacher agency and curriculum making

Teacher agency and curriculum making are mutually constitutive concepts – a chicken and egg situation, to draw upon a well-used metaphor. This echoes to some extent [Stenhouse's \(1975\)](#) observation that there can be no curriculum development without teacher development. Two facets are of crucial importance in understanding how teacher agency is essential for constructive curriculum making by teachers: personal capacity, in terms of professional knowledge, beliefs, dispositions and confidence, as well as the ability to form expansive aspirations for practice; and the affordances for agency that are available in the professional contexts within which, and by means of which, teachers work. Conversely, curriculum making as social practice – especially where actors move across sites – affords important experiences and resources, originating from different settings/environments in an interplay between actors and different agendas, for fostering and enhancing teacher agency. A key question is, thus, “how can systemic thinking and deliberate action across multiple sites of activity enhance teacher agency?”. We contend that “a better understanding of the processes of curriculum making, which reveals such complex underlying and regulating dynamics, would facilitate and enable social actors’ agency toward curriculum change” ([Alvunger et al., 2021](#), pp. 273–4). In this section of the chapter, we first illustrate how teacher agency can enhance curriculum making. We then examine how nuanced understandings of curriculum making might provide the insights for further enabling teacher agency. We conclude with some reflections about how policymakers, teacher educators and school leaders might act differently to develop the professional agency of teachers and facilitate more constructive and meaningful curriculum making in schools.

An important question relates to why teacher agency is important for curriculum making? We do not have the space to examine this question in depth, however, a wealth of research literature over the past 30 years has explored the consequences of teacher-proof policy that seeks to reduce the teacher's role to that of a technician, simply delivering prescribed policy. These include "playing the game" (Gleeson and Gunter, 2001); a game that can take the form of fabrication of the school's image through careful impression management and discourses of excellence (Keddie et al., 2011), the concealing of "dirty laundry" (Cowie et al., 2007), and more serious corruption and cheating (Ball, 2003; Sahlberg, 2010). Recent literature has highlighted the dangers associated with curriculum making that is too tightly focused on meeting external demands, for example instrumental selection of knowledge/content (Smith, 2019) and curriculum narrowing (Shapira and Priestley, 2020). Moreover, as we illustrated previously, teacher mediation of policy is an inevitability, and needs to be enabled constructively; hence there is also need to foster teacher agency.

Teacher agency for curriculum making

Curriculum making is inherently a social practice. It "involves an interplay between different actors, contested spaces and power relations, framed by particularities and contextual factors of social practices, shaping unique settings, and producing multifarious meanings of what curriculum is and for whom it is made" (Alvunger et al., 2021, p. 274). There is considerable resonance here with the ecological understanding of agency, which is also contingent on conditions, "such as the room for maneuver and conceptual resources of policy, as well as the beliefs, values and professional knowledge of the participating stakeholders" (Alvunger et al., 2021). But what does this mean in practical terms? We offer here some examples, structured around the three temporal dimensions of the chordal triad.

The iterational dimension of agency concerns "*the selective reactivation by actors of past patterns of thought and action, routinely incorporated in practical activity, thereby giving stability and order to social universes and helping to sustain identities, interactions, and institutions over time*" (Emirbayer and Mische, 1998, p. 971; *emph. in original*). A key word here is "selective". Agency is not about following habit – going with the flow – but is, for example, about recognizing, appropriating and refashioning past patterns of behaviors and experience in order to maneuver among repertoires in dealing with present dilemmas. The richness of prior experiences (both professional and personal) appears to be a key factor here. Iterational aspects that contribute to teacher agency include personal capacity (skills and knowledge), beliefs (professional and personal) and values. An important consideration is how teachers might become more agentic in the future. This points to the value of good professional education, for instance approaches based around collaborative professional enquiry, which aim to stimulate deep reflection on practice in relation to educational purposes, and which offer interruption to existing ways of thinking enquiry through, for example, engagement with academic reading (e.g., Priestley and Drew, 2019). Furthermore, professional education constitutes only a small part of the formation of teachers' professional experience. Day to day experience in schools – dialogue with colleagues, exposure to school culture, and other professional engagement is also significant, and it is likely that some contexts, such as highly performative cultures that serve to deprofessionalize teachers, will have long term consequences on future teacher development and professional agency (Priestley et al., 2015). Curriculum making requires rich experience; teachers are only able to maneuver between repertoires if they possess knowledge and reflexive understanding of the different possibilities, and research (e.g., Swann and Brown, 1997; Priestley and Minty, 2013) suggests that a more usual scenario in the face of mandated curriculum reform is to fit existing practice to the requirements of new policy, tweaking where necessary and engaging in strategic compliance in relation to terminology and paperwork. Capacity building, including via formal teacher education, would thus seem to be an essential prerequisite to successful curriculum making, as illustrated, for example in Finland's approach to collective sense-making (Soini et al., 2021).

A key implication here, is that teachers able to draw upon a rich reservoir of experience might develop the capacity to envisage more expansive alternative visions about their future practice, than might be the case with less experienced colleagues. Emirbayer and Mische recognize, that agency is "motivated," including possible intentions to bring about a future that is different from the present and the past. This is encapsulated in the projective element of agency, which encompasses "*the imaginative generation by actors of possible future trajectories of action, in which received structures of thought and action may be creatively reconfigured in relation to actors' hopes, fears, and desires for the future*" (Emirbayer and Mische, 1998, p. 971., *emph. in original*). Research in Scotland (Priestley et al., 2015) suggests that many teachers are effective in achieving short- and medium-term goals, for example getting through the lesson smoothly, or focusing on the attainment levels of their students. However, this research suggests a fairly narrow engagement with consideration of the longer term "effects" of education, and moreover a limited imaginary about what constitutes educational practice. Again, capacity-building through teacher education, particularly with a focus on purposes of education and linked to alternative practices, would seem to be an important prerequisite for curriculum making. Our key argument here is that, in contrast to common-place discourses about teacher-proof curricula, "to ensure a high-quality curriculum, we should educate curriculum-proof teachers, who are skilled and knowledgeable, and able to consider and reflect, and to modify the curriculum to meet their own ends and the ends of their students, regardless of the quality of the curriculum-as-product per se" (Alvunger et al., 2021, p. 285).

Although agency is influenced by orientations to the past and the future, it is exercised in the present. This relates to the practical-evaluative dimension, which entails "*the capacity of actors to make practical and normative judgements among alternative possible trajectories of action, in response to the emerging demands, dilemmas, and ambiguities of presently evolving situations*" (Emirbayer and Mische, 1998, p. 971, *emph. in original*). As we indicated previously, judgements are both practical, shaped by the affordances and constraints of particular contexts for action, and evaluative, for example influenced by perceptions of risk. Contexts for action

comprise cultural, structural and material elements. This insight has significant implications for curriculum making. Highly performative cultures, and/or hierarchically structured education systems can impede or restrict the scope of curriculum making; through constraining possibilities for action and rendering certain courses of action risky in terms of, for example, a school's reputation and/or the career/material interests of the teachers; and by rewarding certain instrumental behaviors such as teaching to the test or curriculum narrowing (Priestley et al., 2015). Conversely, contexts may provide the necessary resources for curriculum making, for example material resources (e.g., textbooks, computers), cultural resources (e.g., access to new ideas) and structural/relational resources (e.g., teacher support networks).

Curriculum making for teacher agency

As we have emphasized, agency is not a quality of the individual, but a property that is achieved through the interplay of individual capacity and contextual conditions. In the case of teacher professional agency in respect of curriculum making, we argue that it is constructed, developed and achieved through engagement in professional activities associated with the processes of curriculum making. We contend that participative and collaborative processes, leading to collective knowledge creation, are especially powerful in fostering teacher agency. This requires holistic system thinking, so that curriculum making across educational systems is joined-up, so to speak, and actively enhances teacher agency across different sites of activity.

Macro policy formation should thus be undertaken with teacher agency in mind. We offer several observations in relation to how this might be addressed. First, there is the issue of regulation. This raises questions about the extent to which governments and national agencies should specify content and pedagogy, and the degree to which local actors such as teachers should have autonomy in these matters. As we have illustrated in this chapter, too much regulation can stifle innovation and lead to instrumental curriculum making in schools. Conversely, too little regulation can deprive practitioners of the resources needed to be effective curriculum makers. Both approaches diminish teachers' professional agency. Furthermore, the wrong type of regulation – the technical form of the curriculum (Luke et al., 2012) – can impact negatively upon teacher agency, with the potential, if linked to accountability systems and performance indicators, to drive performative cultures in schools. Second, there is the question of who is involved in macro curriculum making. There is emerging evidence in the current Welsh curricular reforms (e.g., Priestley et al., 2019) that involvement by teachers in the formation of national policy has had a significant impact upon their curriculum making capacity, and their subsequent agency in leading curriculum reform in their own schools and districts. Third, there is the significant role that national policy makers can play in the development of meso infrastructure to support curriculum making.

Evidence from the European cases featuring in our edited collection *Curriculum Making in Europe* (Priestley et al., 2021a) suggests that meso curriculum making is highly significant for both supporting curriculum making in schools and fostering teacher agency.

[T]he meso site of activity emerges as a space “in between”, where ideas and experiences of nano and micro curriculum making meet supra and macro curriculum making. The meso site of activity is, in a way, comparable with a “melting-pot”, where policy, support and guidance on curriculum making are (although with varying success) provided, coordinated and communicated.

Alvunger et al. (2021, p. 282)

Meso curriculum making is a highly complex and multi-dimensional process, which needs to account for the porous and fluid borders between sites of activity. From our European cases, it is clear that it can take many forms, and is undertaken by many different actors, organizations and individuals, including government agencies, private brokers and companies, and bottom-up practitioner initiatives. It is clear that there is increasing interest, in many European systems at least, in developing meso curriculum support, although we also note the risks involved when a putative devolution of state control, accompanied by rhetorics of school autonomy and empowerment, continues to heavily focus heavily on systems for evaluation and audit, along with an emphasis on professional development for teachers stressing instrumentality and performance, rather than engagement with curricular questions. Well-established meso structures provide teachers with the opportunities to work beyond their immediate professional settings, to work collaboratively and develop networks with colleagues from other settings, and to enhance their relational resources (Priestley et al., 2015). This moving of teachers across sites can, in turn, enhance teacher agency, improving confidence, broadening aspirations, changing professional dispositions, and providing the tools – curricular literacy – to enable teachers to act in and respond to environments restrained by accountability and performativity culture (e.g., Priestley and Drew, 2019).

Finally, we briefly reflect on how curriculum making in schools (micro and nano sites) might be utilized to enhance teacher agency. Issues are cultural and structural. Systematic approaches to develop sense-making are important, to provide a vision, to facilitate engagement with new ideas, and especially to develop relational working. In this latter respect, Priestley et al. (2015) documented two high schools with differing approaches to relational working – one predominantly hierarchical, and the other with well-developed horizontal relational structures. In the former case, the authors reported a lack of teacher agency, characterized by low confidence and general confusion about curricular aims and practices; in the latter case, they witnessed more developed teacher agency in respect of the curriculum reforms being undertaken. These effects were evident in both micro and nano curriculum making in the schools. Teacher agency is especially important for the nano curriculum making site, where teachers and students negotiate the often conflicting demands created by the intersection of official goals, teacher lesson plans, student biographies and the contingencies of daily classroom life. Such transactions are invariably uncertain and

messy, emergent and will often diverge considerably from the intentions set out in national policy. We suggest that school leadership plays a significant role here, in creating the curriculum making spaces and practices that foster teacher agency, including establishing relational structures and the time to utilize them, providing resources (e.g., computers, textbooks), instigating the development of a vision for curriculum making, and affording permission and protection for innovation (both necessary to mitigate risk). Conversely, bottom-up influences can significantly enhance teacher agency through the provision of a rich vein of resources for micro and nano curriculum making, for example as captured by [Leat and Thomas \(2019\)](#) in their studies of community curriculum making.

Concluding thoughts

We conclude by reiterating our argument that teacher agency and curriculum making are mutually constitutive concepts. Teacher agency is necessary for the practices that comprise meaningful curriculum making in schools; if we are to see curriculum making as enactment across multiple sites, this requires pre-existing teacher agency, viz. teachers who are able to maximize the resources afforded by those contexts. Conversely, curriculum making across multiple sites has the potential to develop teacher agency, through the very act of curriculum making in different contexts, particularly in respect of the new experiences gained and the cross-fertilization of ideas that occurs when teachers interact in this way. We suggest the following implications of viewing the concepts in this way.

The first concerns participatory curriculum making. As we point out in our book, *Curriculum Making in Europe*, drawing upon insights gleaned from nine countries cases, “curriculum making strategies that allow actors to experience themselves as trusted and capable participants in curriculum making and make sense of it together with others are the most effective ones – ‘effective’ meaning here that people relate to the aims of the curriculum they co-construct and feel ownership, and through that are willing to adapt and develop not only curriculum, but also the educational system and settings within which they work” ([Alvunger et al., 2021](#), pp. 288–9). The second implication relates to accountability based on trust. We have illustrated how high-levels of top-down accountability and mistrust in practitioners have the potential to drive performativity, narrow curriculum making, lead to system-gaming and undermine professional responsibility. We would argue that there needs to be more trust in teachers, as highly qualified graduate-level professionals, accompanied by the resources for those professionals to work with that trust. Conversely, trust requires capacity-building, and we would argue that curriculum making needs to be supported by quality inputs (resources, support and professional education), rather than being dominated by a focus of measurement of outputs. Third, we point to the importance of middle ground and mobility. Meso activity is clearly indicated in all of the European case studies as being a significant and essential component of curriculum making in schools, even when conspicuous by its absence, particularly as a conduit to enable engagement with policy ([Priestley et al., 2021a](#)). We argue that a systemic and holistic approach to curriculum making should include strategies to enable actors to move between sites (which promotes learning and sense-making), and that these sites should include the in-between space provided by meso sites of activity. Finally, we make the case for a more balanced approach to curriculum regulation. This involves ensuring that regulation is neither too directive nor too unspecific. The former restricts through a spiral of specification ([Wolf, 1995](#)); the latter denies practitioners the cultural resources required for teacher agency. Moreover, there needs to be care, to ensure that policy frameworks do not set up impossible dilemmas for teachers, for example a tension between curricular aims and accountability practices. All of the above require systemic thinking by the architects of curriculum policy, and we would argue that this means appropriately involving the local experts (e.g., teachers) in macro curriculum making, both to design policy frameworks, but also to establish the meso infrastructure needed for effective enactment. We are not advocating here a bottom-up approach where responsibility is handed over mainly to teachers; as [Nieveen and Kuiper \(2021\)](#) have illustrated in respect of the Netherlands, this can lead to a narrowing of expertise. Instead, we advocate a whole system approach, which involves different types of expertise in symbiosis, and which serves in turn to enhance the agency of participants and the collective capacity of the system to engage in curriculum making. We acknowledge here that simply adding teachers to curriculum working groups does not guarantee that their voices will be articulated or even heard; research has demonstrated that issues relating to pre-existing hierarchies, expectations about roles, [lack of] trust and teacher confidence and capacity can limit their involvement in such groups (e.g., [Kirk and MacDonald, 2001](#); [Priestley et al., 2019](#); [Bradfield and Exley, 2020](#); [Theodorou et al., 2017](#)), and clearly such issues need to be acknowledged and addressed when groups are composed and facilitated to work.

We conclude by returning to the point at which the chapter started, viz. the changing landscape of policy – new curriculum – that has once more positioned teachers at the heart of the curriculum making process. While we welcome these trajectories, we also warn against the inherent dangers. These include a tendency to over-estimate the capacity of teachers to undertake full-scale curriculum making, particularly at times of economic austerity, and temptations to cut the resources required for such curriculum making. We especially point to risks of a “responsibilization” discourse in policy. While there has been, for some years, a strong emphasis in rhetoric about teacher importance, their agency, that teachers matter, etcetera, in *supra* (e.g., [OECD, 2005](#); [McKinsey and Co., 2007](#)) and national sites (e.g., [ICEA, 2018](#)), such rhetoric has easily slipped into scapegoating those same teachers, for example when PISA and TIMSS country results are published. Against this trend, we emphasize most strongly that it is teachers (as social actors) and enabling curriculum making contexts (as social practices) that matter. We theorize that teacher agency and curriculum making, as symbiotic concepts, enable each other; this highlights how teacher education, policy making, professional development and other related sites of social activity, should be informed by this reciprocal understanding.

References

- Alvunger, D., Soini, T., Philippou, S., Priestley, M., 2021. Conclusions: patterns and trends in curriculum making in Europe. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 273–293.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18, 215–228. <https://doi.org/10.1080/0268093022000043065>.
- Biesta, G.J.J., Tedder, M., 2007. Agency and learning in the lifecourse: towards an ecological perspective. *Stud. Educ. Adults* 39, 132–149. <https://doi.org/10.1080/02660830.2007.11661545>.
- Bobbitt, F.W., 1921. A significant tendency in curriculum-making. *Elem. Sch. J.* 21, 607–615.
- Bowe, R., Ball, S., Gold, A., 1992. *Reforming Education and Changing Schools: Case Studies in Policy Sociology*. Routledge, London.
- Bradfield, K.Z., Exley, B., 2020. Teachers' accounts of their curriculum use: external contextual influences during times of curriculum reform. *Curric. J.* 31, 757–774. <https://doi.org/10.1002/curj.56>.
- Braun, A., Ball, S.J., Maguire, M., Hoskins, K., 2011. Taking context seriously: towards explaining policy enactments in the secondary school. *Discourse* 32, 585–596. <https://doi.org/10.1080/01596306.2011.601555>.
- Brown, S., McIntyre, D., 1992. *Making Sense of Teaching*. Open University Press, Buckingham.
- Connelly, F.M., 2013. Foreword. In: Deng, Z., Gopinathan, S., Lee, C.K.E. (Eds.), *Globalization and the Singapore Curriculum: From Policy to Classroom*. Springer, Singapore, pp. vii–xii.
- Cowie, M., Taylor, D., Croxford, L., 2007. "Tough, intelligent accountability" in Scottish secondary schools and the role of standard tables and charts (stacs): a critical appraisal. *Scot. Educ. Rev.* 39, 29–50.
- Cuban, L., 1998. How schools change reforms: redefining reform success and failure. *Teach. Coll. Rec.* 99, 453–477.
- Deng, Z., Gopinathan, S., Lee, C.K.E., 2013. Introduction. In: Deng, Z., Gopinathan, S., Lee, C.K.E. (Eds.), *Globalization and the Singapore Curriculum: From Policy to Classroom*. Springer, Singapore, pp. 3–14.
- Deng, Z., 2012. School subjects and academic disciplines: the differences. In: Luke, A., Woods, A., Weir, K. (Eds.), *Curriculum, Syllabus Design and Equity: A Primer and Model*. Routledge, New York, NY, pp. 40–53.
- Doyle, W., 1992. Constructing curriculum in the classroom. In: Oser, F.K., Dick, A., Patry, J. (Eds.), *Effective and Responsible Teaching: The New Syntheses*. Jossey-Bass, San Francisco, pp. 66–79.
- Emirbayer, M., Mische, A., 1998. What is agency? *Am. J. Sociol.* 103, 962–1023.
- Gleeson, D., Gunter, H., 2001. The performing school and the modernisation of teachers. In: Gleeson, D., Husbands, C. (Eds.), *The Performing School: Managing, Teaching and Learning in a Performance Culture*. RoutledgeFalmer, London, pp. 139–158.
- Goodlad, J., 1979. The scope of the curriculum field. In: Goodlad, J., et al. (Eds.), *Curriculum Inquiry: The Study of Curriculum Practice*. McGraw-Hill, New York, pp. 17–42.
- Helsby, G., 1999. *Changing Teachers' Work*. Open University Press, Buckingham.
- Hizli Alkan, S., Priestley, M., 2019. Teacher mediation of curriculum making: the role of reflexivity. *J. Curric. Stud.* 51, 737–754. <https://doi.org/10.1080/00220272.2019.1637943>.
- ICEA, 2018. *International Council of Education Advisers; Report 2016–18*. Scottish Government, Edinburgh. Retrieved from: <https://www.gov.scot/publications/international-council-education-advisers-report-2016-18/>. (Accessed 4 January 2021).
- Keddie, A., Mills, M., Pendergast, D., 2011. Fabricating and identity in neo-liberal times: performing schooling as "number one". *Oxf. Rev. Educ.* 37, 75–92. <https://doi.org/10.1080/03054985.2010.538528>.
- Kelly, A.V., 2008. *The Curriculum: Theory and Practice*, fifth ed. Sage, London.
- Kirk, D., MacDonald, D., 2001. Teacher voice and ownership of curriculum change. *J. Curric. Stud.* 33, 551–568. <https://doi.org/10.1080/00220270010016874>.
- Kontovourki, S., Philippou, S., Theodorou, E., 2018. Curriculum making as professionalism-in-context: the cases of two elementary school teachers amidst curriculum change in Cyprus. *Curric. J.* 29, 257–276. <https://doi.org/10.1080/09585176.2018.1447308>.
- Kontovourki, S., Theodorou, E., Philippou, S., 2021. "Sites" of curriculum making in Cyprus: tracing the emergence and transformation of expert teacher-subjects. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 53–76.
- Lasky, S., 2005. A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teach. Teach. Educ.* 21, 899–916. <https://doi.org/10.1016/j.tate.2005.06.003>.
- Leat, D., Thomas, U., 2019. Community curriculum making: mixing the local with the national curriculum. In: *Impact, 6: Broad and Balanced Curriculum*. Retrieved from: [https://impact.chartered.college/article/community-curriculum-making-mixing-local-national-curriculum/\(19/01/21\)](https://impact.chartered.college/article/community-curriculum-making-mixing-local-national-curriculum/(19/01/21)).
- Leat, D., Livingston, K., Priestley, M., 2013. Curriculum deregulation in England and Scotland: different directions of travel? In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom across Europe*, CIDREE Yearbook 2013. SLO, Enschede, The Netherlands, pp. 229–248.
- Lingard, R., 2021. National curriculum making as more or less expressions of and responses to globalization. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 29–52.
- Luke, A., Woods, A., Weir, K., 2012. Curriculum design, equity and the technical form of the curriculum. In: Luke, A., Woods, A., Weir, K. (Eds.), *Curriculum, Syllabus Design and Equity: A Primer and Model*. Routledge, New York, NY, pp. 1–5.
- McKinsey & Co, 2007. *McKinsey Report: How the World's Best Performing School Systems Come Out on Top*. Retrieved from: [https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top/\(19/01/21\)](https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top/(19/01/21)).
- Nieveen, N., Kuiper, W., 2012. Balancing curriculum freedom and regulation in The Netherlands. *Eur. Educ. Res. J.* 11, 357–368. <https://doi.org/10.2304/eej.2012.11.3.357>.
- Nieveen, N., Kuiper, W., 2021. Integral curriculum review in The Netherlands: in need of dovetail joints. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 125–149.
- OECD, 2005. *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. OECD, Paris.
- OECD, 2018. *Future of Education and Skills: Education 2030*. OECD, Paris.
- Osborn, M., Croll, P., Broadfoot, P., Pollard, A., McNess, E., Triggs, P., 1997. Policy into practice and practice into policy: creative mediation in the primary classroom. In: Helsby, G., McCulloch, G. (Eds.), *Teachers and the National Curriculum*. Cassell, London, pp. 52–65.
- Priestley, M., Biesta, G.J.J. (Eds.), 2013. *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. Bloomsbury Academic, London.
- Priestley, M., Drew, V., 2019. Professional Enquiry: an ecological approach to developing teacher agency. In: Godfrey, D., Brown, C. (Eds.), *An Eco-System for Research-Engaged Schools*. Routledge, London, pp. 154–169.
- Priestley, M., Minty, S., 2013. Curriculum for Excellence: "a brilliant idea, but..." *Scot. Educ. Rev.* 45, 39–52.
- Priestley, M., Biesta, G.J.J., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury Academic, London.
- Priestley, M., Crick, T., Hizli Alkan, S., 2019. Co-construction of a national curriculum: the role of teachers as curriculum policy makers in Wales. In: Paper Presented at the ECER Conference, 5 September 2019, 3–6 September 2019, Hamburg, Germany.
- Priestley, M., Alvunger, D., Philippou, S., Soini, T., 2021a. *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, London.
- Priestley, M., Philippou, S., Alvunger, D., Soini, T., 2021b. Curriculum Making: a conceptual framing. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 1–27.
- Priestley, M., 2019. Curriculum: concepts and approaches. In: *Impact, 6: Broad and Balanced Curriculum*. Retrieved from: [https://impact.chartered.college/article/curriculum-concepts-approaches/\(30/06/20\)](https://impact.chartered.college/article/curriculum-concepts-approaches/(30/06/20)).

- Pyhäntö, K., Pietarinen, J., Soini, T., 2014. Comprehensive school teachers' professional agency in large-scale educational change. *J. Educ. Change* 15, 303–325. <https://doi.org/10.1007/s10833-013-9215-8>.
- Sahlberg, P., 2010. Rethinking accountability in a knowledge society. *J. Educ. Change* 11, 45–61. <https://doi.org/10.1007/s10833-008-9098-2>.
- Shapira, M., Priestley, M., 2020. Do schools matter? An exploration of the determinants of lower secondary school subject choices under the Scottish Curriculum for Excellence. *Rev. Educ.* 8, 191–238. <https://doi.org/10.1002/rev3.3180>.
- Sinnema, C., Aitken, G., 2013. Trends in international curriculum development. In: Priestley, M., Biesta, G.J.J. (Eds.), *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. Bloomsbury, London, pp. 141–164.
- Smith, J., 2019. Curriculum coherence and teachers' decision-making in Scottish high school history syllabi. *Curric. J.* 30, 441–463. <https://doi.org/10.1080/09585176.2019.1647861>.
- Soini, T., Pyhäntö, K., Pietarinen, J., 2021. Shared sense-making as key for large scale curriculum reform in Finland. In: Priestley, M., Alvunger, D., Philippou, S., Soini, T. (Eds.), *Curriculum Making in Europe: Policy and Practice within and across Diverse Contexts*. Emerald, Bingley, pp. 247–271.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development*. Heinemann, London.
- Supovitz, J.A., 2008. Implementation as iterative refraction. In: Supovitz, J.A., Weinbaum, E.H. (Eds.), *The Implementation Gap: Understanding Reform in High Schools*. Teachers College Press, New York, pp. 151–172.
- Swann, J., Brown, S., 1997. The implementation of a national curriculum and teachers' classroom thinking. *Res. Pap. Educ.* 12, 91–114. <https://doi.org/10.1080/0267152970120106>.
- Theodorou, E., Philippou, S., Kontovourki, S., 2017. Caught between worlds of expertise: primary teachers amidst official curriculum development processes in Cyprus. *Curric. Inq.* 47, 217–240. <https://doi.org/10.1080/03626784.2017.1283591>.
- Thijs, A., van den Akker, 2009. *Curriculum in Development*. SLO, Enschede, the Netherlands.
- Wolf, A., 1995. *Competence-Based Assessment (Assessing Assessment)*. Open University Press, Buckingham.
- Yates, L., Collins, C., 2010. The absence of knowledge in Australian curriculum reforms. *Eur. J. Educ.* 45, 89–101. <https://doi.org/10.1111/j.1465-3435.2009.01417.x>.
- Young, M., Muller, J., 2010. Three educational scenarios for the future: lessons from the sociology of knowledge. *Eur. J. Educ.* 45 (1), 11–27. <https://doi.org/10.1111/j.1465-3435.2009.01413.x>.
- Young, M., 2007. *Bringing Knowledge Back in: From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, London.
- Zapp, M., 2019. Empowerment for individual agency: an analysis of international organizations' curriculum recommendations. *Glob. Soc. Educ.* 17, 231–260. <https://doi.org/10.1080/14767724.2019.1577717>.

Curriculum design

Nienke M. Nieveen^a, Jan J.H. van den Akker^b, and Joke M. Voogt^c, ^a University of Twente, Department of Teacher Professional Development ELAN Enschede, the Netherlands; ^b Curriculum Research & Consultancy Bergen, the Netherlands; and ^c University of Amsterdam, Department of Child Development and Education Amsterdam, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Conceptualizing curriculum design	198
Approaches to curriculum design	200
Curriculum design as process	200
The instrumental approach	201
The communicative approach	201
The artistic approach	201
The pragmatic approach	201
Curricular design research approach	202
Actors in curriculum design	202
Curriculum policy and actors	202
Role of networks	203
References	204
Further Reading	205
Relevant Websites	205

Conceptualizing curriculum design

Defining related concepts as curriculum, curriculum design and curriculum development is an elusive activity. The concept of curriculum has been defined in literature in many ways, as exemplified during recent decades in three major curriculum handbooks (Connelly, 2008; Jackson, 1992; Wyse et al. 2016). Those usually long definitions often heavily depend on the specific contexts, roles and (ideological) perspectives of the authors (Jackson, 1992). We prefer a shorter, more open, neutral and flexible definition in order to recognize the complexity of the term (without premature closure to more narrow interpretations) and to facilitate wider use (Kridel, 2010). As curricula for educational settings are not given by nature but represent artifacts that are deliberately created by humans (Simon, 1996), it is obvious that they represent a typical *design* challenge, with many choices to be made. Simon (1996) defined design as “devising a course of action”, which also captures the essence of curriculum work. If one considers “learning” as the central action and overall aim in education, then the definition of curriculum as a *design for learning* suits well as a simple, open-ended, universal definition which allows for multiple specifications to various contexts and perspectives. The term design is also very appropriate as it refers to both “product” and “process” dimensions of curriculum (cf. Priestley and Biesta, 2013): curriculum artifacts (programmes and materials) as well as the series of acts that lead toward such artifacts and the way they are enacted in practice. The interpretation of curriculum as process underlines that curriculum is not a static document but may better be characterized as ongoing work-in-progress.

Traditionally, curriculum is primarily associated with the aims, content and organization of learning (Walker, 1990), but various authors (Klein, 1991; Van den Akker, 2003) have expanded this list of components to include: vision or rationale, goals and objectives, contents, materials and resources, learning activities, teaching strategies, assessment, grouping, time, and space. Especially when trying to redesign a curriculum and making it work in practice, it is wise to pay attention to the coherence of those components. The (normative) vision on the overarching, broader aims of learning and teaching serves as a central link, connecting all other curriculum components. A metaphor to illustrate this viewpoint is a spider web (Van den Akker, 2003; see Fig. 1), including guiding design questions for the many curriculum components. The curricular spider’s web points to both the flexibility and the vulnerability of curriculum, as every chain is as strong as its weakest element while all components are interrelated and interconnected.

Besides a visual representation of the challenging components, the curricular spider’s web can serve as an analytic tool to explore and clarify the discrepancies between the existing and desired curriculum as well as a design tool that assists developers in prioritizing next steps in the process of getting to a coherent curriculum.

In particular when “grand”, ambitious curriculum (re)design is at stake, one needs an underlying (or overarching) rationale (or vision), articulating the values and principles that offer a sense of direction and may help in choosing a position and striking a proper balance between the various broader aims of education: (i) creating a knowledge base for further learning and studies, (ii) preparing for the world of work, (iii) equipping for living in society, and (iv) stimulating personal development. (For some inspiration, including examples, for discourses about the curriculum rationale, see the OECD website on project “Education, 2030”). Obviously, priorities will vary over different stages of schooling and also depend on the dominant values in the specific educational context.

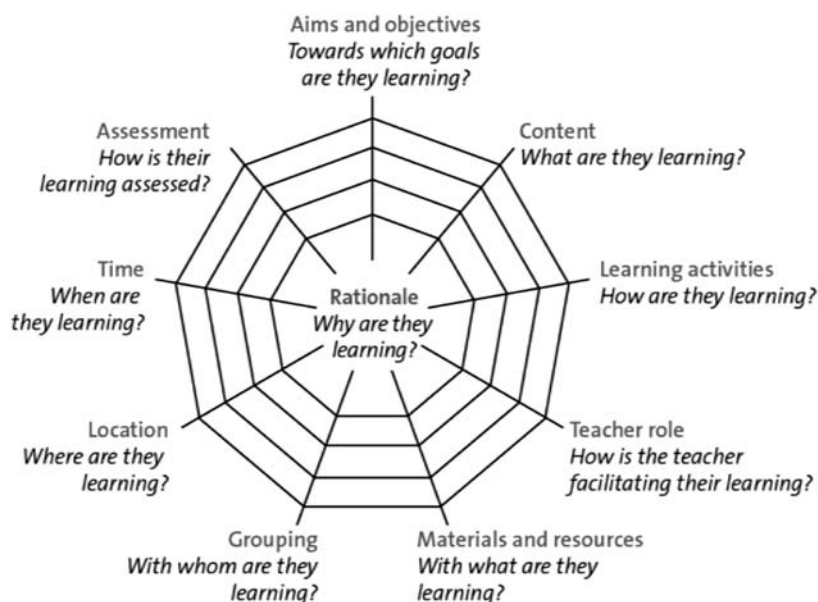


Fig. 1 The curricular spider's web.

If one considers curriculum design as a process, it may be seen as the core activity within a broader, usually iterative and lengthy process of curriculum development, that includes activities such as: analysis of needs and context (reviewing the current situation as well as forward-looking explorations of desirable options); formulating principles for the to-be-designed curriculum; drafting initial versions (prototypes) of (parts of) the new curriculum; seeking informative feedback about such drafts; constructing and trying-out more concrete curriculum parts, seeking a balance between ideals and practicality; exploring implementation chances; synthesizing parts into a coherent whole and fitting these into the broader picture. The curriculum development process usually involves many more (cyclic) interim steps and activities, also related to socio-political issues of decision-making. Eventually, often also challenges of scalability and sustainability become more prominent. Thus, curriculum development can be portrayed as an overall, iterative process of analyzing, designing and constructing, evaluating and implementing curriculum change. Curriculum design (in reality often redesign) cannot be easily isolated as a distinct phase – it is better to interpret it as an embedded activity within such a trajectory. We would argue that curriculum design regards the less routinized, intellectually most demanding tasks of imagining and creating new destinations and pathways of learning and teaching. As such it represents the heart and driving force of any ambitious curriculum development. In practice and over time, curriculum developers have followed various approaches and strategies to structure their design and development work. See section two for some elaboration on that.

However, before describing various approaches of curriculum design and development, it is clarifying to distinguish various curriculum levels or layers. For starters, it does make a difference whether the curriculum is meant for a specific school site or for a generic context. One may even elaborate such differences by identifying six possible levels (Van den Akker, 2003; Thijs and Van den Akker, 2009), including some examples of respective curricular “products”:

- Supra (inter-/transnational); e.g., the International Baccalaureate programmes, or the Common European Framework of Reference of Languages, or the curricular discussion documents or frameworks of international organizations as UNESCO or OECD.
- Macro (national, regional); e.g., Scotland’s “Curriculum for Excellence” or comparable documents of many other countries.
- Meso (school, institution); e.g., the school-wide curriculum of any specific educational organization
- Micro (classroom, group); e.g., lesson plans, curriculum materials, or textbooks for learning and teaching of specific parts of the curriculum.
- Nano (individual); e.g., the personal plan of studies of an individual student.

Thus, one should keep in mind that “the” curriculum does not exist. It is always necessary to specify the levels of curriculum one is referring to, in order to support clarity in communication. And, obviously, the level(s) of curriculum at stake will have implications for the most suitable development approach. Moreover, it is important to realize that curriculum change at scale will eventually have to relate to all those levels, thus implying much interaction between the various levels.

Another conceptually helpful distinction for understanding curriculum (design and development) is to differentiate between various curriculum representations (Goodlad and associates, 1979; Van den Akker, 2003):

- Visionary: The ideas, ideals and intentions for learning (and teaching), giving directions to or underpinning choices in the curriculum.
- Written: How the intentions are elaborated and specified in documents/materials.

- Perceived: How the intended curriculum is interpreted, in particular by teachers.
- Operational: How the curriculum is enacted in classroom practice.
- Experiential: How the curriculum is experienced, in particular by students.
- Learned: The learner outcomes of the enacted curriculum.

Often, there are tremendous gaps between the aspirations of the intended curriculum (visionary and written), the realities of the implemented curriculum (perceived and operational), and the results of the attained curriculum (experiential and learned). Hence, one could characterize the process of curriculum development as a comprehensive, systematic effort to transform ideas about desirable aims and contents for learning into documents and materials to stimulate instructional practices that foster learning activities and experiences leading toward intended learning outcomes. This systematic endeavor has usually many substantive, technical-professional and practical aspects. However, curriculum development processes also often carry many socio-political and contextual issues, with many uncertainties in making choices by many actors. Curriculum choices inevitably relate to (often conflicting) values and interests of various stakeholders and interest groups. Hence, curriculum development is sometimes even characterized as a war zone or battlefield (e.g., [Carnocahn, 1993](#)). Although tensions will always remain in the process of designing curricula, it does help if participants in curriculum decision making do have a clear understanding of the different curriculum perspectives in levels, representations, components, values and interests. Context always matters, and notably in curriculum approaches. It is important to always keep in mind that one size does not fit all. Moreover, while grand visions may have an inspiring and mobilizing effect, “big bang” approaches to curriculum change should be avoided. Instead, it makes sense to start with relatively small steps in designing and piloting new parts of an envisaged curriculum, departing from an analysis of the current situation in school practice and aiming at the *zone of proximal development* of those involved. Essentially, at this point, the idea of curricular design research becomes relevant as it may help to address uncertainties on various quality issues (which will be elaborated later on) and to improve quality.

Approaches to curriculum design

Curriculum design as process

As introduced in the previous section curriculum design is a core activity within the overall process of curriculum development. Curriculum development usually consists of various activities related to phases in the process: analysis, design, construction, evaluation and implementation (see for an overview [Gustafson and Branch, 2002](#)). While these phases may suggest that the process of curriculum development is linear, in reality the process is cyclic and iterative. In particular activities addressing evaluation and implementation of initial versions (prototypes of (parts of) the new curriculum) need to start already early in the process to prevent the development of curriculum products that may be valid from a scientific point of view but are not practical in the daily practice of teaching and learning.

Grounded in theoretical views about curriculum and curriculum design typically three approaches to curriculum design are described in the literature ([Marsh and Willis, 2007](#)): the instrumental approach ([Tyler, 1949](#)), the communicative approach ([Walker, 1971](#)) and the artistic approach ([Eisner, 1979](#)). More recently two other approaches are distinguished. The design of educational software applications used in software engineering resulted in the pragmatic approach ([Visscher-Voerman and Gustafson, 2004](#)). The last approach – curricular design research – emerged from the field of design research ([McKenney and Reeves, 2018](#); [Van den Akker and Nieveen, 2021](#)). **Table 1** presents an overview of the five approaches. After that we will discuss each approach.

Table 1 Curriculum design approaches: Overview.

	<i>Instrumental approach</i>	<i>Communicative approach</i>	<i>Artistic approach</i>	<i>Pragmatic approach</i>	<i>Curriculum design research approach</i>
<i>Characteristics of the approach</i>	Rational process	Intensive deliberation (especially at the start)	Creative reflection	Frequent evaluation of prototypes, user-centeredness	Interwoven design and research activities
<i>Design context</i>	Generic	Generic/specific	Specific	Generic	Generic/specific
<i>Focus of the process</i>	Coherent design	External consistency	Coherent design	Coherent design/external consistency	Coherent design/external consistency
<i>Teacher roles</i>	Implementer	Co-designer	Designer	Implementer/(Co-)designer	Implementer/(Co-)designer
<i>Conception of quality</i>	Design meets pre-specified requirements	Design meets requirements derived from broad consensus among stakeholders	Design meets the vision of the designer	Design meets requirements of the users	Design is relevant, justifiable, consistent, practical, effective, sustainable and scalable

Inspired by [Thijs and Van den Akker \(2009\)](#).

The instrumental approach

Tyler's seminal work "Basic principles on curriculum and instruction" (1949) is a classical example of the instrumental approach to curriculum design. Tyler articulated four questions to guide the curriculum designer: 1. In order to identify concrete learning objectives, what goals are pursued by this curriculum? 2. With which learning experiences can students achieve these objectives? 3. How to organize the learning experiences? 4. How to determine the effectiveness of the learning experiences? Answers to these questions can be found in several places such as the subject domains and the learning sciences (Sawyer, 2015). Curriculum development based on the instrumental approach puts an emphasis on analysis and design. Evaluation and implementation are – when they occur – conducted at the end of the process for summative purposes. The instrumental approach is rooted in prescriptive theories in searching for heuristics that guide a design process leading to curricula of the highest quality (Marsh and Willis, 2007). Such an approach strives after a design that is coherent, see the spider web introduced above (Van den Akker, 2003).

The main elements of the instrumental approach are often used in large-scale curriculum development processes with a generic focus, but increasingly also in school-based curriculum development processes. While in the instrumental approach teachers are above all seen as implementers of a high quality curriculum, increasingly teacher designers use elements of this approach when developing curriculum (McKenney et al., 2015; Pieters et al., 2019).

The communicative approach

While the instrumental approach strives after a coherent design, the communicative approach emphasizes the importance of shared understanding and creating ownership among different stakeholders for the process and the product of curriculum development, Kessels and Plomp (1999) refer to this as external consistency. The communicative approach is rooted in descriptive theories, which study what people actually do and how design decisions are taken. Walker's (1971) deliberative approach is a typical example. In this view, curriculum design is above all a process of negotiation among key stakeholders such as policy makers, teachers and scientists, who each bring their own knowledge, experiences, beliefs and values to the curriculum design process. That is why the deliberative approach starts with an open platform of ideas to discuss with major stakeholders the intended curriculum. An open discussion contributes to consensus about the rationale for the design and a shared understanding of the task. This phase is followed by deliberations about possible alternatives and their consequences leading to a decision about the most promising alternative. Although these two phases are distinguished in reality there is not a strict boundary. The emphasis of the approach is on building consensus about the curriculum at the start of the process, and on needs and context information during deliberations of alternatives. Little attention is paid to evaluation and implementation of the design. The communicative approach is used in large-scale generic curriculum design as well as in school-based curriculum design. Teachers are seen as major stakeholders in the design process. In large-scale generic design processes, they are frequently invited as co-designer for parts of the process (Turcotte et al. 2010), while in school-based curriculum design, teachers, often collaborating in teams, design for their own school or subject domain (Pieters et al., 2019).

The artistic approach

According to Eisner (1979), the process of curriculum design is similar to the process of creating art. Eisner sees curriculum design as a process of sense making based on the expertise and experience of the connoisseur. As such the artistic approach is rooted in the theory of situated cognition (Greeno, 2011). It is a social practice that takes place in a specific context and culture. Curriculum design is a process of construction and reconstruction of teachers, who uses their "wisdom of practice" (Shulman, 1986) to fully understand and enact the curriculum. It is often an intuitive and implicit process, and not based on an explicit course of action. Because means and ends are interdependent, the process of curriculum design is not linear but open ended. However, like Tyler, Eisner agrees that the curriculum elements, as visualized in the spider web, need to be addressed in relation to each other for a coherent design. The artistic approach is applied in curriculum design processes at the level of classrooms and schools. The teacher as designer of curriculum needs some curriculum design expertise to be able to arrive at a more systematic approach to realize a coherent design (Huizinga et al. 2014; McKenney et al., 2015).

The pragmatic approach

Due to the increasing use of educational software in education the pragmatic approach, derived from software engineering, is gaining ground in curriculum design (e.g., Nieveen, 1999; Visscher-Voerman and Gustafson, 2004). The pragmatic approach is highly iterative, because of the formative evaluation of successive prototypes. Successive prototypes provide concrete pictures of the intended curriculum already from an early stage of the design process. Formative evaluation leads to revision suggestions and adaptation of the evolving ideas that underpin the product. The user-centeredness of this approach leads to more intensive discussions with the end-users about the requirements of the product, provides better opportunities to negotiate and justify design ideas, increases chances for user-commitment and ownership of the final deliverable and moreover prevents curriculum developers from a tendency to "design for themselves". Each cycle represents the evolution of intentions of the final deliverable. End users, often

teachers and their students (and not their representatives or school leaders), are actively involved in the formative evaluation of the prototypes (Tripp and Bichelmeyer, 1990). The pragmatic approach leads to a final product that works in practice and is considered practical by its users (Visscher-Voerman and Gustafson, 2004).

Curricular design research approach

Curricular design research (CDR) can be conceived as a special form of educational design research (McKenney and Reeves, 2018; Van den Akker and Nieveen, 2021). Characteristic for this type of research is the close interaction between design, its iterative and cyclic nature, and targeted research activities to inform the design process. CDR builds on ideas of the instrumental, the communicative and the pragmatic approach: a systematic process that leads to coherent high quality design (instrumental approach); attention for shared understanding of stakeholders (communicative approach) and (rapid) prototyping with formative evaluation as core activity in the design process (pragmatic approach). More prominent than in the other approaches is the important role of research in all phases of the process with the aim to contribute to valid, practical, effective and sustainable curricular artifacts and practices. Moreover, CDR produces knowledge about curriculum design and development in the form of design principles that are ecologically valid and may guide future curriculum design in similar contexts.

CDR has been successfully applied in large-scale curriculum development and in school-specific curriculum development processes. In the latter case teachers often collaborate in teams. The active involvement of participants (e.g. teachers, school leaders) in the process contributes to their professional development.

To ensure that the curriculum that results from CDR is of high quality, the following criteria are formulated (see Van den Akker and Nieveen, 2021):

- **Relevance:** Is there a clear and shared rationale for the new curriculum?
- **Justification:** Is the curriculum supported by a credible knowledge base?
- **Consistency:** Are the various curriculum components (see spider's web) well aligned?
- **Practicality:** Is the curriculum feasible in practice (especially for teachers and learners)?
- **Effectiveness:** Does the curriculum result in the intended learner outcomes?
- **Sustainability:** Is the curriculum change successful in traveling over time?

And in case of large-scale development processes also.

- **Scalability:** Can the curriculum operate across various contexts?

Apart from an extensive preliminary research at the start and the implementation and scaling at the end, the development process entails a cyclical process of analysis, design/construction, and evaluation as a phase in the middle. Working on the relevance and justification are important criteria in the preliminary research phase. Here, stakeholder consultation, the analysis of the current practice and a review of research are important activities. During the cyclical analysis, design and evaluation phase in the middle, the focus is mainly on consistency and practicality. Typical activities during this phase are the formative evaluation of prototypes through walk-throughs, micro-evaluations and try-outs. During the implementation and scaling phase research activities focus on determining effectiveness, sustainability and scalability. Effectiveness can be found through empirical testing. The sustainability and scalability of the design can be determined through case studies and surveys.

Actors in curriculum design

Curriculum policy and actors

Starting from the idea that curricular documents are to be found at various levels in the educational system (supra, macro, meso, micro and nano; see earlier in this contribution) and that large-scale curriculum change will have to relate to decisions at all those levels, it is central to note that curriculum development is the responsibility of many different players in the educational field (illustrated with a system web, Fig. 2), and many more actors are usually involved in the development process itself. The nature of their involvement and responsibilities depend on the curriculum policy of a country.

In countries with a centralized curriculum policy, the curriculum at the macro level is typically highly specified and prescriptive from the idea that the curriculum change in the classroom should closely correspond to/align with the prescribed intentions. Here the curriculum policy starts from a "fidelity" approach (Fullan and Pomfret, 1977) that directs teachers to select and use particular lesson units in specified ways. This approach aims at uniformity, requiring that all schools teach the same topics in similar ways, in which the room for site-specific curricular choices is restricted.

In countries with a decentralized curriculum policy, the curriculum at the macro level leaves ample room for teachers and students to make local curricular decisions, based on for instance what happens in the classroom or in the context of the school. This means that next to their teaching duties, teachers will be involved in the design of a school-based curriculum. At the heart of this "mutual adaptation" approach is trusting schools and teachers to make site-specific interpretations of curriculum guidelines (Hopkins, 2001). Adjustments and insights of teachers and other stakeholders are also used to feed changes in the curriculum at the macro level in order to improve, for instance, its relevance and practicality.

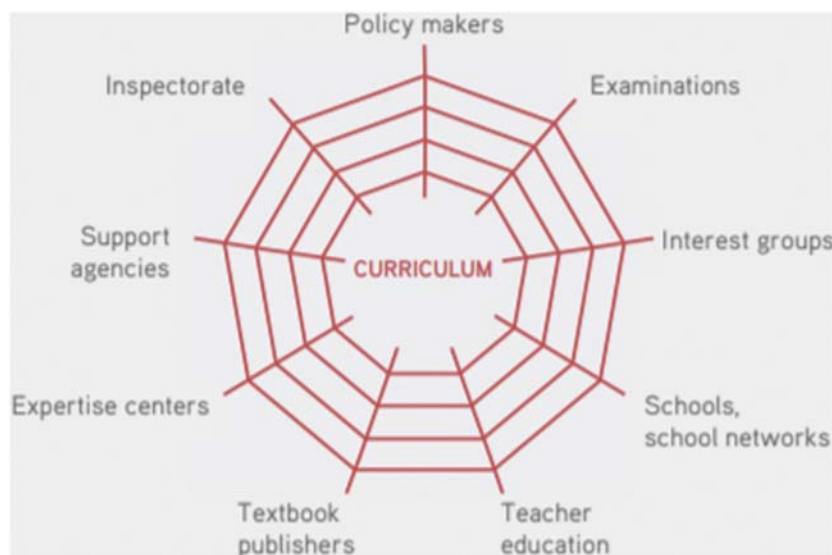


Fig. 2 Curriculum system web (Nieveen, van den Akker and Resink, 2010).

As was already stated, the nature of the involvement and responsibilities of the many actors in the educational field depend on the curriculum policy of a country. In countries with centralized policies, the actors in the system web have a role in directing and standardizing the curriculum implementation in the classroom. This happens usually at great cost to professionalism and motivation of teachers (cf. Hargreaves and Shirley, 2009). In countries with a decentralized approach, the role of the actors in the system web is much more diffuse. In these countries the needs and wishes of the schools, teachers and learners are steering their work. Where schools and teachers are in charge of the actual implementation of the curriculum in practice, all actions taken by others in the system web (policy makers, school boards, support agencies, test developers, textbook developers, pre- and in-service education, inspectorate, and so on) can to be considered as implementation encouragement (Nieveen et al., 2014).

Role of networks

Scaling of curriculum change that has been initiated at the macro level is a serious challenge for many countries (Fullan, 2008) and related to that countries show a difficulty in striking a good balance between both (decentralized and centralized) positions. Both prescriptive and more flexible models have their advantages and disadvantages (Fullan, 2008; Hargreaves, 2003). In short, the dilemma is that prescriptive models obtain better short-term results but usually lack sustainable change, while flexible models usually last longer but often lack focus (Fullan, 2008). Flexible models tend to last longer because at their heart is trust in schools and teachers leading the curriculum renewal (Hopkins, 2001). The latter is generally perceived as a prerequisite for securing sustainable change (Hargreaves and Shirley, 2009). At the same time, prescriptive models also need deliberately measured doses of pressure and support to put change into practice.

Concerning curriculum change, all dimensions of scale as elaborated by Coburn (2003) are at stake: spread over more people and organizations; shift in ownership from original source to other people and contexts; depth in the practice of teaching and learning in schools; sustainability over longer periods of time. To address these challenges, more mixed collaboration approaches between teachers, school leaders, teacher educators, textbook developers, assessment developers, researchers and the like, e.g. in networked improvement communities (Bryk et al., 2015), seem to be promising for common professional capacity building as well as gradually coming to scale. These kind of networks fit best with a mutual adaptation philosophy with sufficient space for and active encouragement of context-specific choices for local communities, parents, schools, teachers, and students. A clear example of such a localization approach (within the boundaries of the national framework) can be found in Finland (Halinen and Holappa, 2013), where local communities are expected and supported to make, translate, and add their own curriculum choices, together with the schools in that particular community. This also requires that teachers, being the key players in curriculum change, have not only an appropriate degree of professional autonomy in the pedagogical and instructional domain, but also have a sufficient degree of curricular capacity to recognize and utilize the potential to cooperate in designing their common school-based curriculum. Such curricular-capacity building of teachers individually and in teams, while also including school leaders, should be an important aim of professional development strategies (Huizinga et al., 2014).

In order to bridge the gap between the work of individual teachers (within their own subjects and classrooms) and school-wide curriculum renewal aspirations, teacher design teams are to be of help (cf. Handelzalts et al., 2019; Lieberman and Miller, 2004; McLaughlin and Talbert, 2001). These small networks of teachers from adjacent subjects center their work on students, share their responsibilities, work together on their common curriculum, make meaningful connections between topics that are usually addressed in different subject areas, reflect on and learn from their experiences (at the individual, group and school level). Although

all teachers need to be involved, one way or another, in some parts of the process of developing the new curriculum and feel that they have collective responsibility. Skilbeck (1998) points out that as talents and interests vary, and as there are different kinds of development tasks to perform, matching procedures are inevitable, meaning that not all teachers need to be actively involved in all curriculum development tasks.

Ultimate success, with deep curriculum change in classrooms, where teachers use new materials, engage in new teaching practices and incorporate new beliefs, will require lots of investments in bottom-up initiatives (Fullan, 2016). Teacher design teams constitute a promising vehicle for this kind of school-based curriculum development, potentially addressing all four of Coburn's dimensions. This implies that these teams have sufficient facilities (in particular development time) and access to adaptable exemplifications (e.g., educative lesson materials; digital portals). Such scenarios and examples can be the result of the earlier mentioned work of heterogeneous networked improvement communities.

Moreover, schools and teacher teams are in need of a clear curriculum at the meso level in order to bridge the curricular aims of the school and the instructional decisions at the classroom level. As long as this curriculum at the school level remains a "black box", it serves a situation where "anything goes," which in the end merely leads to incoherent situations for teachers and learners (Nieveen et al., 2010). Schools need curricular leaders who feel and act responsibly with the curriculum and who encourage others to strive for decisions that fit the curriculum choices at the other levels. Again, professional development of teachers as cooperative designers requires time facilities, a cooperative culture and energizing distributed leadership.

Of course, the differences between primary and secondary schools need to be taken in account. In secondary schools, the role of subject matter teachers is more dominant, asking for other collaboration patterns than in primary schools, where usually more generic teachers are responsible for a wide range of domains. However, regardless the specific education sector, the strong point of combining teacher design teams, the role of curriculum leaders in the school and networked improvement communities is that they expand the basic repertoire of individual teachers from designing lessons for their own classes to more collaborative, curriculum-oriented and research-supported patterns which is beneficial for the professional development of all involved and the quality of the curriculum change at hand.

References

- Bryk, A., Gomez, L., Grunow, A., LeMahieu, P., 2015. *Learning to Improve*. Harvard Education Press, Cambridge, MA.
- Carnocahn, W., 1993. *The Battleground of the Curriculum*. Stanford University Press, Stanford, CA.
- Coburn, C., 2003. Rethinking scale: moving beyond numbers to deep and lasting change. *Educ. Res.* 32 (6), 3–12.
- Connelly, F. (ed.), (2008). *The Sage Handbook of Curriculum and Instruction*. Los Angeles: Sage.
- Eisner, E.W., 1979. *The Educational Imagination. On the Design and Evaluation of School Programs*. Macmillan, New York.
- Fullan, M., Pomfret, A., 1977. Research on curriculum and instruction implementation. *Rev. Educ. Res.* 47 (1), 335–397.
- Fullan, M., 2008. Curriculum implementation and sustainability. In: Connelly, F. (Ed.), *The Sage Handbook of Curriculum and Instruction*. Sage, Los Angeles, pp. 113–122.
- Fullan, M., 2016. *The New Meaning of Educational Change*, fifth ed. Teachers College Press, New York.
- Goodlad, J., associates, 1979. *Curriculum Inquiry: The Study of Curriculum Practice*. McGraw-Hill, New York.
- Greeno, J.G., 2011. A situative perspective on cognition and learning in interaction. In: Koschmann, T. (Ed.), *Theories of Learning and Studies of Instructional Practice*, vol. 1. Springer, New York, pp. 41–71.
- Gustafson, K., Branch, R., 2002. *Survey of Instructional Development Models*, fourth ed. Clearinghouse of Instructional Technology, Syracuse University, New York.
- Hälinen, I., Holappa, A.-S., 2013. Curricular balance based on dialogue, cooperation and trust – the case of Finland. In: Kuiper, W., Berkvens, J. (Eds.), *Balancing Curriculum Regulation and Freedom across Europe*. SLO (Netherlands Institute for Curriculum Development), Enschede, pp. 136–153 [CIDREE Yearbook 2013].
- Handelzalts, A., Nieveen, N.M., Van den Akker, J., 2019. Teacher design teams for school-wide curriculum development: reflections on an early study. In: Pieters, J., Voogt, J., Roblin, N.P. (Eds.), *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham, Switzerland, pp. 55–82.
- Hargreaves, A., 2003. *Teaching in the Knowledge Society: Education in the Age of Insecurity*. Teachers College Press, New York.
- Hargreaves, A., Shirley, D., 2009. *The Fourth Way: The Inspiring Future for Educational Change*. Corwin, Thousand Oaks, CA.
- Hopkins, D., 2001. *School Improvement for Real*. Routledge Falmer, London.
- Huizinga, T., Handelzalts, A., Nieveen, N., Voogt, J., 2014. Teacher involvement in curriculum design: need for support to enhance teachers' design expertise. *J. Curric. Stud.* 46, 1–25.
- Jackson, P., 1992. Conceptions of curriculum and curriculum specialists. In: Jackson, P. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 3–40.
- Kessels, J., Plomp, T., 1999. A systematic and relational approach to obtaining curriculum consistency in corporate education. *J. Curric. Stud.* 31, 679–709.
- Klein, F., 1991. *The Politics of Curriculum Decision Making*. SUNY Press, Albany, NY.
- Kridel, C., 2010. *Encyclopedia of Curriculum Studies*. Sage, Los Angeles, CA.
- Lieberman, A., Miller, J., 2004. *Teacher Leadership*. Jossey-Bass, San Francisco.
- Marsh, C.J., Willis, G., 2007. *Curriculum: Alternative Approaches, Ongoing Issues*, fourth ed. Pearson Education, Upper Saddle River, NJ.
- McLaughlin, M.W., Talbert, J.E., 2001. *Professional Communities and the Work of High School Teaching*. The University of Chicago Press, Chicago.
- McKenney, S.E., Kali, Y., Markuskaukaite, L., Voogt, J., 2015. Teacher design knowledge for technology enhanced learning: an ecological framework for investigating assets and needs. *Instr. Sci.* 43, 181–202.
- McKenney, S., Reeves, T., 2018. *Conducting Educational Design Research*, second ed. Routledge, London.
- Nieveen, N., 1999. Prototyping to reach product quality. In: van den Akker, J., Branch, R., Gustafson, K., Nieveen, N., Plomp, T. (Eds.), *Design Approaches and Tools in Education and Training*. Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 125–136.
- Nieveen, N., Sluismans, L., Van den Akker, J., 2014. Encouraging curriculum change in The Netherlands: the next episode. In: Nyhamn, F., Hopfenbeck, T.N. (Eds.), *From Political Decisions to Change in the Classroom*. The Norwegian Directorate for Education and Training, Oslo, Norway, pp. 162–183.
- Nieveen, N., Van den Akker, J., Resink, F., 2010. Framing and supporting school-based curriculum development in The Netherlands. In: Law, E., Nieveen, N. (Eds.), *Schools as Curriculum Agencies: Asian and European Perspectives on School-based Curriculum Development*. Sense Publishers, Rotterdam, The Netherlands, pp. 273–283.
- Pieters, J., Voogt, J., Pareja Roblin, N. (Eds.), 2019. *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham, Switzerland.
- Priestley, M., Biesta, G., 2013. *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*. Bloomsbury Academic, London.
- Sawyer, R.K., 2015. In: *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, Cambridge.

- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14.
- Simon, H., 1996. *The Sciences of the Artificial*, third ed. MIT Press, Cambridge, MA.
- Skilbeck, M., 1998. School based curriculum development. In: Lieberman, A., Fullan, M., Hopkins, D.W. (Eds.), *Hargreaves, International Handbook of Educational Change*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 121–144.
- Thijs, A., Van den Akker, J. (Eds.), 2009. *Curriculum in Development*. SLO, Enschede, The Netherlands.
- Tripp, S.D., Bichelmeyer, B., 1990. Rapid prototyping: an alternative instructional design strategy. *Educ. Technol. Res. Dev.* 38, 31–44.
- Turcotte, S., Laferrière, T., Hamel, C., Breuleux, A., 2010. Multilevel innovation in remote networked schools. *Syst. Pract. Action Res.* 23, 285–299.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, Chicago.
- Van den Akker, J., 2003. Curriculum perspectives: an introduction. In: Van den Akker, J., Kuiper, W., Hameyer, U. (Eds.), *Curriculum Landscapes and Trends*. Kluwer Academic Publishers, Dordrecht, pp. 1–10.
- Van den Akker, J., Nieveen, N., 2021. Combining curriculum and teacher development through design research. In: Philippakos, Z., Pellegrino, A., Howell, E. (Eds.), *Design based Research in Education: Theory and Applications*. Guilford Press, New York.
- Visscher-Voerman, J.I.A., Gustafson, K., 2004. Paradigms in the theory and practice of education and training design. *Educ. Technol. Res. Dev.* 52 (2), 69–89.
- Walker, D.F., 1971. A study of deliberation in three curriculum projects. *Curric. Theor. Netw.* 7, 118–134.
- Walker, D., 1990. *Fundamentals of Curriculum*. Harcourt, San Diego (Brace, Jovanovich Publishers).
- Wyse, D., Hayward, L., Pandya, J. (Eds.), 2016. *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. Sage Publications, London.

Further Reading

- Marsh, C.J., 2009. *Key Concepts for Understanding Curriculum*. Routledge, London.
- Van den Akker, J., Gravemeijer, K., McKenney, S., Nieveen, N., 2006. *Educational Design Research*. Routledge, London.
- Voogt, J., Pieters, J., Handelzalts, A., 2016. Teacher collaboration in curriculum design teams: effects, mechanisms and conditions. *Educ. Res. Eval.* 22, 121–140.
- Voogt, J., Pieters, J., Pareja Roblin, N., 2019. Collaborative curriculum design in teacher teams: Foundations. In: Pieters, J., Voogt, J., Pareja Roblin, N. (Eds.), *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham, Switzerland, pp. 5–18.

Relevant Websites

- Consortium of Institutions for Development and Research in Education in Europe: <http://www.cidree.org/>.
- OECD Future of Education and Skills 2030: <https://www.oecd.org/education/2030-project/>.
- Educational Design Research, Part A: An Introduction Enschede: SLO: <https://www.slo.nl/zoeken/@4315/educational-design/>.
- Educational Design Research, Part B: Illustrative cases. <https://www.slo.nl/zoeken/@14467/educational-design-research-part/>.

Learning environments in vocational education: how to elicit learning for and in practice

Maaïke Koopman, Ilya Zitter, Liesbeth Baartman, and Elly de Bruijn, HU University of Applied Sciences Utrecht, Research Centre for Learning and Innovation, Research group Vocational Education, Utrecht, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	206
Review process	208
Results	209
Learning how to act in vocational practice	209
Learning processes and activities	209
Designable elements affording learning how to act in vocational practice	209
Developing vocational knowledge	210
Learning processes and activities	210
Designable elements affording developing vocational knowledge	210
Developing a vocational identity	211
Learning processes and activities	211
Designable elements affording developing a vocational identity	212
Conclusion and discussion	212
References	214

Introduction

The focus of this chapter is on learning for vocations, including both learning *for* practice and learning *in* vocational practice, and learning environments that afford such learning. Learning *for* vocational practice pertains to learning in the context of vocational education.¹ Learning *in* vocational practice concerns learning in the context of continuing education, further professional development, and lifelong learning. Learning for vocations, whether *in* or *for* vocational practice, involves the individual process of developing a vocation in terms of competencies of and being a professional. It entails preparing for participating in vocational practice and/or remaining qualified to contribute in work in a rapidly changing society. Learning for both initially becoming a professional, as well as learning for further professionalization demands from learners that they participate in communities of practice related to work, that they cooperate with other professionals from both comparable and dissimilar backgrounds, and that they continue to develop both themselves and the professional community they work in. Learning for vocations occurs within the context of formal education and in the context of lifelong learning. In line with Billett (2011), we use the term vocational education for all learning for vocations, whether low skilled or high skilled (such as physiotherapy or accounting), primarily aimed to prepare and qualify for work. Vocational education serves many purposes, mainly related to vocational education being a means for individuals to get or remain qualified for occupations throughout their (working) life (Billett, 2011). A large variety of learners is involved in learning for vocations, ranging from adolescents and young people to adults in different stages of their lives and careers. And while some learning for vocations is considered prestigious (such as learning to become a medical doctor), other trajectories are set up for young people to overcome disadvantages on the labor market (Behle, 2017), to help learners to get qualified despite prior poor school performance (Chan, 2013), or to handle unemployment and shortages of skilled labor (Ferm et al., 2018). While such learning environments are arranged in educational programs governed by different regulations, norms, and enrollment procedures and are situated in various social-cultural settings and conditions, there are also similarities and insights with respect to the learning processes and activities that can be instructive for investigating and eliciting learning in these differing settings.

The question of how (future) professionals learn and how this learning can be supported is a pressing and central issue for educators, whether in school or at the workplace (De Bruijn, 2012). It is related to the complex relation between (the individual learners' perception of) learning and the learning environments eliciting this learning. Evidence shows that learning is embedded in and results from social practice. Social practice serves as learning environment and can also be intentional (i.e. designed) (Entwistle and Petersen, 2004; Tynjälä et al., 2021). The assumption is that the way (future) professionals learn is—at least in part—afforded by the learning environments they are part of and the interactions with significant others while active in these learning environments (Koenen et al., 2015; Bouw et al., 2021a).

We believe that learning of individual learners is intertwined in social practice. Hence, in this chapter we use the body of knowledge in which learning *for* and *in* vocational practice is conceptualized from a socio-cultural perspective for our analyses (Hökkä

¹(Upper) Secondary Vocational Education (ISCED/EQF 3-4) and Higher Education (ISCED/EQF 5-7) are within the scope of this chapter.

et al., 2020). From this socio-cultural view, learning is perceived as increasing participation in and contribution to a professional community, which may emanate in developing occupational roles and behaviors, vocational knowledge, and a vocational identity (see e.g. Lave and Wenger, 1991). Learning is situated: it is viewed as growing engagement in and belonging to a vocation, i.e. a professional community of practice (Chan, 2015). Knowledge and actions are developed simultaneously in such Communities of Practice (CoP) (Hodkinson et al., 2008). In CoPs, learning may result in changes pertaining to the learners themselves (e.g. new insights, newly developed skills or actions) or changes to the CoP in question, also affecting the more experienced professionals in it (Engeström, 2001; Smith, 2012). The concept of boundary crossing plays an important role: boundaries are “socio-cultural differences leading to discontinuity in action or interaction” (Akkerman and Bakker, 2011, p. 133). Such differences between domains or communities and experiences of discontinuity may elicit learning when part of a learning environment.

In a recent systematic review of literature on learning for and in vocational practice (Koopman et al., in preparation), we focused on the way in which (future) professionals’ learning processes and accompanying learning activities are conceptualized and empirically studied. In this review, we made a distinction between which *learning processes* are defined and which *learning activities* are described (see Table 1). Learning processes are defined as configurations of several learning activities. Such processes take time and lead to change (e.g. in knowledge, actions, or attitudes). Learning activities are defined as actions (doing or thinking) related to task performance (e.g. cooking, thinking aloud, having a conversation) that take place at a particular point in time.

In this chapter, we look at learning processes and learning activities in the reviewed literature in relation to the learning environments which may elicit such learning processes and activities as described in this literature.

Taking a broad view on learning environments, these are considered as “both the socio-cultural setting and the physical/digital setting in which learners perform their tasks (Carvalho and Goodyear, 2018; Zitter and Hoeve, 2012)” (Bouw et al., 2021a, p. 489). Learning can be enhanced by the interplay between interactions and affordances. Interaction involves all interactions between (future) professionals and “significant others,” such as teachers, supervisors, peers, experienced professionals from the workplace, clients, etc. (Ceelen et al., 2019). Affordances refer to the perceived elements of environments that may elicit learning processes and activities, for example, tasks, roles, (professional) tools and objects, digital environments, physical locations and spaces at school and/or at the workplace, and timeframes (e.g. deadlines, sequences, and schedules) (Bouw et al., 2021a). Both interactions and affordances may arise from intentionally shaping the so-called *designable elements* of an environment and as such aim to enhance learning. We refer to these environments as learning environments. It should be noted that the learning processes and activities themselves cannot be designed, but the epistemic, physical, and social elements (i.e. the designable elements) of the environment in which this learning unfolds can be intentionally designed (Carvalho and Goodyear, 2018). That is, in learning environments interaction and learning processes and activities are emergent and the designable elements are there to elicit these, which consequently may lead to the intended (learning) outcomes. In this way learning environments can elicit or support certain learning processes or activities to be carried out by learners (see Fig. 1).

The following designable elements can be distinguished (Bouw et al., 2021a,b):

- *Epistemic elements*: tasks and assignments that represent ways of knowing—including skills and attitudes—that are seen as worthwhile in the professional domain(s) in question. Moreover, these elements are related to how the ways of knowing can best be presented and structured to elicit learning. Specifically when learning for and in vocational practice, this involves determining which work practices are suited for (the intended) learning and what kind of vocational tasks learners are supposed to engage in while active in these work practices;
- *Spatial elements*: the locations and spaces (both physical and digital) in which learning takes place. Locations/spaces are at school, at different workplaces, or at third places (neither school, nor work) which are relevant for the professional domain (for example, a specific area or neighborhood);
- *Instrumental elements*: the instruments, tools, objects, and artifacts needed for the (intended) learning processes and activities. Some are more oriented toward schools, while others are necessary to perform the vocational tasks in question (see epistemic elements). Specifically when learning for and in vocational practice, artifacts that facilitate the interaction between school and work, i.e. boundary objects, are likely. Smartly chosen professional artifacts can function as boundary objects;
- *Social elements*: the actors involved in a learning environment, the roles these actors fulfill, how they are grouped and how tasks are divided between them. When learning for and in vocational practice, actors from both school and work may be involved, fulfilling both education roles (e.g. learner, coach, or assessor) and roles related to the relevant professional domain (e.g. junior or senior colleague and managerial roles). One person can fulfill several roles at the same time, e.g. a students can be “learner,” junior colleague in relation to a seasoned professional, and senior colleague in relation to an even more junior peer. Roles can also be fulfilled by different persons concerned, e.g. a teacher can also explicitly fulfill the role of senior colleague.

Table 1 Distinction between learning processes and learning activities.

	<i>Learning processes</i>	<i>Learning activities</i>
Definition	Configuration of several learning activities Leads to changes in knowledge, actions or attitudes	Actions (doing or thinking) related to task performance e.g. thinking aloud, having a conversation
Timeframe	Longer time span	Particular point in time

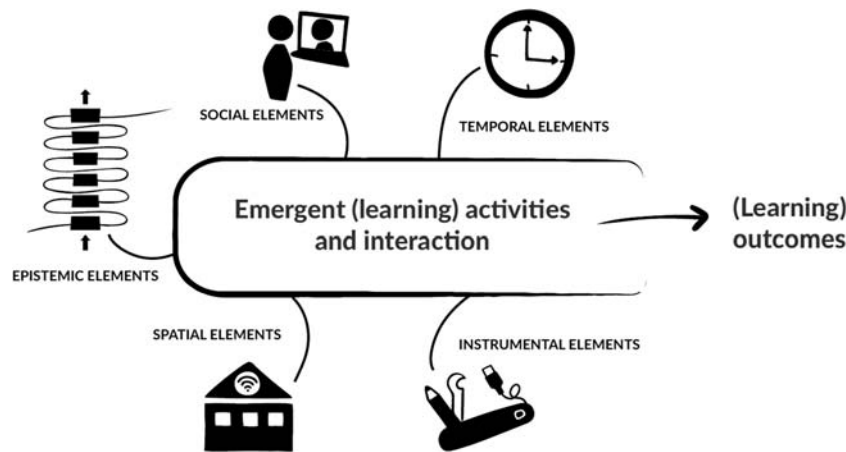


Fig. 1 Designable elements of learning environments.

- *Temporal elements*: these elements are included to clarify the importance of affordances related to time. Designable temporal elements are, for example, work pace (e.g. available amount of time-on-task, deadlines, time pressure), and overall intensity and timespan. When learning for and in vocational practice, opportunities to interrupt the work process can be designed, in order to slow down, accelerate, or pause for educational purposes. Temporal elements also include sequencing of tasks as described under epistemic elements and the position of a task in overall learning trajectories.

When consistently designed for eliciting specific intended learning processes and activities, the epistemic, spatial, instrumental, social, and temporal elements, may afford this desired learning.

The main question we answer in this chapter is: *Which designable elements can be identified in the social contexts where learning for and in vocational practice occurs?* To answer this question, we make use of the results of the systematic literature review, searching for the designable elements mentioned in the included studies. As such, the chapter intends to contribute to the knowledge base on the relation between (intended) learning processes and activities and designing for learning in vocational education. It may also provide concrete directions for educational designers and educators on how to support (future) professionals' learning for and in vocational practice.

Review process

We conducted a systematic review of literature on learning processes, and learning activities related to learning *for* and *in* vocational practice (Koopman et al., in preparation). We used an interpretative approach to analyze the included studies in order to broaden our understanding of these learning processes (Dixon-Woods et al., 2004). We conducted the search in ERIC, PsychInfo, and Web of Science (period 2009–2020). Four sets of search terms were used simultaneously to search the title, keywords, and abstract of publications:

- Set 1: learn* OR develop* OR becom* AND
- Set 2: knowledge* OR skill* OR attitude* OR identit* OR competenc* OR member* OR activit* OR practice OR expert* AND
- Set 3: vocational OR occupational AND
- Set 4: process*

This resulted in 944 unique hits. From these hits, we made a selection based on reading the title, keywords, and abstract using the following inclusion criteria for an article to be selected for further scrutiny: (a) the article had to describe learning processes (i.e. there is a learner; the research design and method leads to data on learning processes; the results and/or conclusions are about learning processes), (b) the article had to include empirical research (i.e. data have been used; data describe learning sufficiently concrete), and (c) the article had to focus on learning for vocations in the context of initial *vocational education* or *non-formal education* (instead of working communities without the intention to learn). 35 primary studies were included, 30 of which described student learning from a socio-cultural perspective; the five other studies described learning from cognitivist or constructivist perspectives. These 30 studies described learning for and in vocational practice in Australia, Brunei, Canada, Denmark, the Netherlands, New Zealand, Norway, South Africa, Sweden, the United Kingdom, and the United States. Most studies reported on learning for and in vocational practice organized via secondary vocational education or higher education ($n = 23$); the other seven studies reported on learning in continuing education (learning at work).

In this chapter, we limited our analyses to the dominant view of socio-cultural learning theory. Because of this focus, we were able to zoom in on the social practice in which learning for vocations is embedded and identify designable elements that might

afford such learning. For our analyses, we used the frame of designable elements of learning environments as described by Bouw et al. (2021a,b), i.e. epistemic, spatial, instrumental, social, and temporal elements as explained above.

Results

Learning for and in vocational practice involves participating in vocational practice, development of the knowledge that is relevant for practising a vocation, and development of one's identity as a (future) professional. In the 30 articles included, these three specific foci of learning could be identified; formulated more precisely (a) learning how to act in vocational practice, (b) developing vocational knowledge, and (c) developing a vocational identity. We used these three foci of learning to guide our analyses of which designable elements were mentioned in the included studies and how these elements were described to afford learning. In some of the articles one or two of these foci were recognizable. And, in case of two foci, all kinds of combinations existed, for example a focus on both acting in vocational practice and identity development or a focus on both acting in vocational practice and knowledge development. In other articles all three foci seemed present. From a socio-cultural view on learning these three foci are related and can be seen as inseparable (Colley et al., 2003). As meaning-making and sense-making are embedded in social practice (Bijlsma et al., 2016), acting in vocational practice will be reflected in both knowledge development and identity formation. Together these three foci can be considered to form the core of what it means to become a professional in a certain field (see e.g. Ord and Nuttall, 2016; Stets and Burke, 2000; Wallin et al., 2019). For analytical purposes, in the next sections the learning processes and activities are described in more detail per focus, zooming in on designable elements of learning environments we found in the reviewed literature that can afford learning.

Learning how to act in vocational practice

Learning processes and activities

Participation is typical for the way in which learning how to act in vocational practice is perceived (Bijlsma et al., 2016). The transition from peripheral to full participation in a CoP or activity system is key. Students' situated *learning processes* can be characterized by socialization, experiential learning, and increasing commitment and independence (Poortman et al., 2011; Pridham et al., 2012; Reegård, 2015). In some studies, learning processes related to boundary crossing, characterized by reflection and transformation, are added (Akkerman and Bakker, 2012). *Learning activities* of students can be characterized by doing (all kinds of work-related tasks), exerting agency of their own learning (e.g. taking initiative, asking for demonstrations or feedback, having discussions), thinking (in order to support development of acting), and contributing to building a CoP (Ferm et al., 2018; Chan, 2015; Jonasson, 2015; Pridham et al., 2012). Learning activities are typically combinations of actions and interactions, with students negotiating about their engagement in activity (Smith, 2012). For example, during workplace learning future industrial workers learn through doing everyday work in industry, paying attention to what happens in the workplace, and having others demonstrate tasks (Ferm et al., 2018). They are invited to participate in activities by their co-workers and hence become engaged in those activities, resulting in them learning the informal rules and ways of working of a specific workplace.

Designable elements affording learning how to act in vocational practice

Epistemic elements

Tasks that may elicit learning how to act in vocational practice are designed with a focus on acting and doing the work relevant to a profession. In many of the included studies, learning while doing occupational tasks is somehow part of the overall learning trajectory that results in formal qualification. Learning by working on actual occupational tasks can afford socialization and experiential learning (Chan, 2013; Lie et al., 2016). The participation of students in a CoP can be encouraged via a combination of increasingly difficult tasks and fading support (Chan, 2015). Sometimes, doing work tasks is combined with supporting assignments or projects (Knightbridge, 2014). In such cases, a more explicit link is designed between school-based and work-based learning. Experiences from the work tasks can be discussed, which—in some cases—seems to support meaning-making and understanding (Ferm et al., 2018; Johansson et al., 2019). In other cases, theory-oriented courses are provided on work-oriented content, which can result in students' experiencing boundaries or discrepancies between school and work. These experienced boundaries can also be discussed and can thus instigate learning (Akkerman and Bakker, 2012; Bakker and Akkerman, 2014).

Spatial elements

The physical environment in which (future) professionals learn how to act is often a real work setting, such as a shop, factory, bakery, or fitness facility (e.g. Chan, 2013; Ferm et al., 2018; Poortman et al., 2011). In case of learning *for* vocational practice in formal vocational education, this often takes place in the form of apprenticeships, internships, or placements in (future) occupational workplaces. In some studies, learning how to act as a professional is attempted to be afforded in a school-based setting, which can be characterized by in-school work settings similar to vocational practice (for example, a school-based mini enterprise or an in-school kitchen with restaurant) (Heusdens et al., 2019). In case of learning *in* vocational practice, it generally involves learning on the job.

Instrumental elements

Regarding learning how to act in vocational practice, a variety of tools, objects, and artifacts is described. These instrumental elements are generally related to and typical for the vocational tasks the (future) professional has to become experienced in. Examples of such instruments are lesson plans provided by the university that support the development of student teachers toward becoming a teacher (Goh, 2015); ingredients and equipment in a bakery or kitchen that enable learning on the job by bakers-in-training (Chan, 2013); and lab materials, machines, and protocols in laboratory education elicit acting (and thinking) similar to vocational practice (Bakker and Akkerman, 2014). Sometimes, such instrumental elements may serve as boundary objects that enable learning how to act. In case of learning how to be a security officer, for example, a uniform may serve as a transformational artifact eliciting discussions about the contradiction between (personal) threats and protection (of others) inherent to the work (Johansson et al., 2019). The discussion of the purpose and meaning of wearing uniforms, may support security officers-in-training in prioritizing their future actions and in becoming familiar with the norms and standards of the profession.

Social elements

Regarding social elements that encourage learning how to act in vocational practice, the importance of interaction between learners themselves and between learners and supervisors, mentors, or (future) colleagues is quite clear. (Future) professionals form communities of practice or communities of learners, do group assignments, have discussions, consider each other as experts or sources of knowledge, and are thus able to help each other (Pridham et al., 2012). Settings in which several learners meet, facilitate conversations in which learners make their implicit experiences more explicit, which—in turn—supports their thinking about their actions (Johansson et al., 2019). Learners are also supported by co-workers (who sometimes have a formal role as supervisor or mentor). In their role of junior colleague, learners tend to ask these more experienced professionals for demonstrations of vocational actions which they can observe (Pridham et al., 2012). In some studies, co-workers are described as being brokers who represent a certain professional domain and who make students aware of the discrepancies between professional communities they are already familiar with (e.g. a former job) and new professional communities (Andersson and Köpsén, 2019). A particular role is described for educators in school or at the workplace: they stimulate acting as professional by inviting to participate, encouraging, providing feedback, and acknowledging (growing) competence (Emad and Roth, 2016; Johansson et al., 2019). In some cases, learners are described to mirror this guidance by exerting agency over their own learning, taking up responsibility, and doing increasingly difficult tasks (Goh, 2015; Reegård, 2015).

Temporal elements

Affordances for learning how to act in vocational practice in the sequence of an overall learning trajectory are often strictly timeboxed, with separate timeboxes related to learning at work and separate timeboxes related to learning at school. In other words, learning how to act in vocational practice is clearly divided in timeframe and planning between settings at school or at the workplace. In the included literature, learning how to act is often described from the perspective of learning at work (e.g. Bisholt, 2012; Chan, 2013; Smith, 2012). Some other studies report slightly more fluid presence of the temporal elements. In these studies, learning at work is interspersed with “come back to school time” often dedicated to reflection on work experiences (Akkerman and Bakker, 2012; Johansson et al., 2019; Ferm et al., 2018). Learning at school and learning at work can even occur more or less simultaneously, in hybrid settings, but such settings are hardly described in the reviewed literature.

Developing vocational knowledge

Learning processes and activities

Development of vocational knowledge—like learning how to act in vocational practice—is closely related to participation (Chan, 2013; Heusdens et al., 2019). Important *learning processes* specifically related to developing vocational knowledge are meaning-making and sense-making. Meaning-making pertains acquiring meaning through interactions with other (prospective) practitioners of the vocation, for example via interpreting collective knowledge and concretizing personal knowledge (e.g. becoming acquainted with collective meanings, reflection on concepts, values and beliefs, or “vocationalizing” concepts) (Bijlsma et al., 2016). Sense-making is related to making knowledge personal and, through that, developing new collective knowledge (Heusdens et al., 2019; Johansson et al., 2019). It involves perspective taking, contextualizing, interpreting and integrating collective knowledge, and personalizing vocational knowledge. *Learning activities* related to developing vocational knowledge can be characterized by the link between thinking and acting as well. Granting meaning to concepts is typical, for example, and occurs by means of using and refining concepts in or for particular work tasks (see e.g. Bijlsma et al., 2016).

Designable elements affording developing vocational knowledge

Epistemic elements

Like learning how to act in vocational practice, developing vocational knowledge in many cases involves or is closely related to learning while doing occupational tasks. From this participation in work, development of know-how or vocational knowledge can emerge (Chan, 2013; Heusdens et al., 2016, 2019). Epistemic elements that may afford vocational knowledge development include consecutive tasks that seem to be created or provided in such a manner that they become increasingly difficult for learners. Moreover, intentionally designing difficult tasks seems to play a role in developing vocational knowledge (Jonasson, 2015). In the

study of Jonasson (2015), for example, learners in culinary education are confronted with deliberately complicated tasks that demand improvisation and are prone to mistakes. These tasks are a starting point for interaction among learners or between learners and educators, resulting in (future) professionals extending their repertoire and their knowledge about cooking, improving the techniques and procedures they use, and strengthening their ability to combine work tasks under time pressure.

Spatial elements

Though developing vocational knowledge is likely to take place during workplace learning, affordances for learning processes and activities related to knowledge development often seem to be school-based (Heusdens et al., 2019; Jonasson, 2015; Pridham et al., 2012). These spatial elements can sometimes be characterized by simulation of a real-life work setting (e.g. a catering company within a school).

Instrumental elements

An example of developing vocational knowledge while acting at a real workplace can be found in the work of Johansson et al. (2019), who describe shoemakers-in-training developing embodied knowledge during operating machines, while working with leather and feeling how the leather behaves. Several tools, objects, and artifacts play a role in eliciting development of vocational knowledge. Because of the close relation between acting and knowing, these instrumental elements are overlapping with the elements described in **Designable elements affording learning how to act in vocational practice** section. Additionally, the studies in which learning for the development of vocational knowledge is described, instrumental elements such as books and literature are mentioned explicitly (e.g. Holmgren et al., 2017).

Social elements

Regarding social elements that encourage the development of vocational knowledge, a combination of observing (peers, co-workers), interacting, posing questions, and having discussions seems to support learning (Bakker and Akkerman, 2014; Chan, 2013; Holmgren et al., 2017; Johansson et al., 2019). Together, and in close relation to vocational actions, these strategies seem to help (future) professionals in sharing and co-constructing knowledge via, for example, reasoning, articulation and elaboration of concepts (Bakker and Akkerman, 2014; Heusdens et al., 2019; Holmgren et al., 2017). For example, students that learn how to cook, develop knowledge by connecting concepts (e.g. the concept of Julienne cut) to their vocational activities (e.g. cutting a hot pepper) or by both interpreting and articulating what they do (e.g. skinning a tomato using first boiling and then iced water; Heusdens et al., 2019). Hence, designable elements which afford making knowledge explicit are an important way to give shape to the social elements of learning environments in order to make knowledge development more likely. Educators, in the role of senior colleague, may also fulfill a prominent role in this; they can support learners in developing vocational knowledge by helping them in conceptualizing and concretizing knowledge (Heusdens et al., 2019). Conceptualizing involves understanding what a concept means in a specific situation in relation to the meaning of other concepts. Concretizing involves understanding the use of concepts in a situation and interpreting that situation to be able to transform and use concepts appropriately. Both are complex processes, which can be supported by elaboration and questions from a teacher in the role of senior colleague. A teacher in the role of expert can also be seen as a source of knowledge and may give direct support in knowledge development (for example related to articulating knowledge and helping students reason), but also more indirect support (for example in the form of the subdivision of tasks and encouraging learners).

Temporal elements

At the level of distinct learning processes and activities, in the aforementioned simulated settings are explicitly designed to make pausing or interrupting ongoing work processes for instructional purposes possible (Bouw et al., 2021a), which may support vocational knowledge development. At the level of an overall learning trajectory, the sequence learners are involved in related to developing vocational knowledge seems to be characterized by alternation between school-based learning, workplace learning in the form of apprenticeships, and/or shorter workplace visits (e.g. Andersson and Köpsén, 2019; Chan, 2013; Johansson et al., 2019). As stated earlier, the actual development of vocational knowledge, or at least making this knowledge explicit, seems to occur relatively frequently within the school setting.

Developing a vocational identity

Learning processes and activities

Learning in terms of the development of a vocational identity, is often described in terms of “becoming” (becoming a baker, shoemaker, teacher, occupational therapist, etc.) (Chan, 2013, 2019). This involves *learning processes* that can be characterized by going through the phases of belonging, becoming, and being (Chan, 2013). This process starts with the feeling of fitting in a certain professional world, followed by becoming competent in the professional role that is demanded, and finally complete identification with the profession (to sometimes even seeing the profession as a calling) (Chan, 2019). In other words, it involves a process of positioning oneself toward the profession that is learned and transformation of an identity via participation in the professional community. In some publications, a process of identity work, characterized by integrating or combining identity positions and thus transforming identity, is described (Akkerman and Bakker, 2012). *Learning activities* related to the development of a vocational identity include creating opportunities to learn (by taking initiatives) and contributing to building a CoP (with a focus on sharing

motives and stories, discussing occupational roles, rules, norms, and values) (Ferm et al., 2018; Johansson et al., 2019; Pridham et al., 2012). In some studies learning activities are described that involve negotiations, particularly about the way the learner is engaged in a professional community and personal versus occupational norms (Smith, 2012). For example, based on his prior experiences a fire fighter negotiates regarding the way he personally likes to prepare and test his breathing apparatus compared to official protocol, enabling him to work in a more safe and quick way. As such he negotiates the way he is engaged in practice and contributes to that practice (Smith, 2012). Adaptation to a new or changed occupational role can often instigate such activities.

Designable elements affording developing a vocational identity

Epistemic elements

The main practice that seems to elicit vocational identity learning is participating in the profession. Examples we found are:

- Participating in an action research community by occupational therapists that elicits learning about what it entails to become more scholarly (Ennals et al., 2016);
- Participation in Saturday clinics for an underserved population that elicits future health professionals' thinking about representing their profession and working in teams with professionals from other disciplines (Lie et al., 2016);
- Teaching internships that elicit thinking about how student teachers see themselves as (future) teachers (Schoeman and Mabunda, 2012).

Other, more pre-structured and supportive tasks or assignments related to participation in work are hardly mentioned. The work itself seems to afford learners thinking about the type of professional they want to become and the way they identify themselves with the profession in question. Participating in work also seems to elicit negotiations about (future) occupational roles and the accompanying norms and values of a profession.

Spatial elements

Following from the previous paragraph, the workplace can thus be considered the most important setting that initiates thinking and learning of (future) professional about their vocational identity. Workplace learning, for example via apprenticeships, can be a way for students to cross a boundary (Akkerman and Bakker, 2012), which elicits their thinking about the kind of professional they want to be (Schoeman and Mabunda, 2012).

Instrumental elements

Several tools, objects, and artifacts play a role and they are closely related to the vocational practice in question (e.g. a bakery setting including ingredients and machines or blog posts written by occupational therapists doing research; see e.g. Ennals et al., 2016; Chan, 2013). Typical for vocational identity development is that these instruments are actively used by the (future) professionals or sometimes even created by them. Creating an artifact, e.g. writing a blog post, can encourage learners to reflect about who they are and who they want to be (Howell, 2009).

Social elements

Development of a vocational identity is supported by participating in a community of other learners. Other learners are seen as fulfilling the role of representative of the profession, as (future) colleague, in such communities. In the communities, they share experiences and reflect collectively on what it means to work in a certain practice and on how to align with the norms and values of that practice (Ennals et al., 2016; Herrera et al., 2015; Howell, 2009). Besides peers, a broad range of significant others, such as educators in school or at the workplace, co-workers, and even the learner's family can support identity development. Educators frequently fulfill the role of role model or socializing agent (Ferm et al., 2018; Schoeman and Mabunda, 2012) and in that role they are being observed by learners or provide feedback. Clients or, in case of student teachers, pupils can also play a role as significant other. They may serve as a mirror for (future) professionals and make them realize how they have to adapt their occupational role to their needs (Smith, 2012).

Temporal elements

In the included studies, little direct information can be found regarding temporal elements that afford developing a vocational identity. What we do see, is that vocational identity development seems to occur at several points in time during relatively long learning trajectories (e.g. when switching to a new profession, when starting a new internship at a new workplace, when participating in series of site visits). This supports a common assumption that identity development is an ongoing process during a person's career (Beijaard et al., 2004). The transition to a new real-life work setting or to a new occupational role is an important starting point for renewed thinking about oneself as a professional (Smith, 2012).

Conclusion and discussion

In the included studies, a variety of designable elements which may afford learning *for* and *in* vocational practice can be found. The studies focused on learning processes and activities which could be related to three foci: learning how to act in vocational practice, developing vocational knowledge, and developing a vocational identity. In many studies, these foci were described in close relation

to each other, which can be explained from the socio-cultural view that was present in the included articles (Colley et al., 2003). From this view and in line with Leont'ev (1978), thinking and acting are often seen as inseparable and embedded in social practice. For example, a learning process like vocationalizing concepts will probably most directly contribute to developing vocational knowledge, but acting in vocational practice is needed in order to be able to vocationalize concepts and the process of vocationalizing concepts might, in turn, improve learners' ability to act as a professional.

Looking at the designable elements of learning environments, some overlap between which elements might enhance which learning can be found as well. Regarding *epistemic elements*, learning while working on occupational tasks, for example during apprenticeships, has a very prominent role. Other, more classical tasks or assignments are sometimes found to support developing vocational knowledge, but such tasks play a much smaller role in learning how to act in vocational practice and almost no role in developing a vocational identity.

Regarding *spatial elements*, a combination of locations/spaces at school, at work, and sometimes integration of those two via incorporation seem to be typical for eliciting learning for and in vocational practice. Of course, this is not a new insight, but our review shows which learning processes and activities are supported by which of those settings. Regarding learning how to act in vocational practice and developing a vocational identity, learning can be linked most clearly to work-based settings. With respect to developing vocational knowledge, we still see a more prominent role for school-based learning. At the same time, simulated work settings (e.g. school-based restaurant or mini-enterprise) also afford developing vocational knowledge.

Instrumental elements, like tools, objects, and artifacts play an important role in learning for and in vocational practice. They enable (future) professionals to do the work at the workplace or in simulation work settings in school and, as such, support learners in their learning how to act in vocational practice and their vocational knowledge development. Tools, objects, and artifacts are actively used and sometimes also created by learners. In relation to development of a vocational identity, they may serve as boundary objects or markers of a profession and of the norms, values, habits, and procedures related to it.

The importance of *social elements* in a learning environment becomes clear from the results of our analyses. Interaction between (future) professionals and significant others fulfilling different roles is key for learning for and in vocational practice: without interaction, the other designable elements will lose most of their strengths. A broad range of possible significant others becomes visible, including peers, educators in school or at the workplace, co-workers, clients, and family. The learners and significant others fulfill a variety of roles. Interaction between learners or peers in the role of colleague-in-training is visible in many studies. They support each other's learning by observing one another, discussing experiences, asking questions, providing or receiving feedback, thus functioning as a community of learners. Educators play an important role in the interaction regarding development of vocational knowledge. In the role of senior colleague or expert, they support learners in articulating concepts, vocationalizing concepts, or conceptualizing and concretizing concepts. Teachers are also role models in encouraging students to take initiative and responsibility, and as such their learning how to act as professionals.

The *temporal elements* which may afford learning for and in vocational practice can be discussed at two levels: at the level of distinct learning processes and activities or at the level of an overall learning trajectory. At the level of distinct processes and activities, we can conclude that the studied articles do not provide much insights. While studies on learning environments in vocational education sometimes mention intentionally designing opportunities to pause or interrupt ongoing work for instructional purposes (Bouw et al., 2021a), the reviewed literature on learning for vocational practice did not include evidence on what such measures may afford in terms of learning processes and activities. At the level of overall learning trajectories, we can conclude that timeframes and planning can be either more strict or more fluid. Strict temporal elements include distinct timeboxes for learning at school or learning at work. More fluid temporal elements include alternating learning in different settings, interspersing learning at school with shorter or longer periods of learning at work. Time becomes even more fluid when learning at school and at work occur more or less simultaneous. In some studies we saw the benefits of more fluid, simultaneous planning, for example when work experiences are immediately discussed, thus affording extension of a vocational repertoire, developing vocational knowledge, and contributing to the development of a vocational identity.

In sum, the main differences between the three foci that we found can be described as follows. In the included literature, learning how to act in vocational practice seems to be frequently supported in or by means of work-based settings or in-school work settings similar to vocational practice. Instrumental elements and social elements (co-workers as significant others) play a crucial role in this respect. Development of vocational knowledge, conversely, seems to be supported in school-based settings, possibly because of the presence of an explicit teaching role (social elements) and opportunities to pause in order to discuss the use of concepts (temporal elements). The third focus, i.e. the development of a vocational identity, seems to be enhanced in learning environments related to or resembling every day work, which might elicit (future) professionals' thinking about the kind of professional they are or want to be. Such environments often consist of a configuration of all designable elements, but are not always explicitly designed for the purpose of identity learning. Intervention meetings, focused on looking back at learning experiences in those every day work settings might strengthen vocational identity development (e.g. Akkerman and Bakker, 2012).

By focusing on designable elements that were described in studies about learning for and in vocational practice, we were able to provide insight into the elements of learning environments that can be influenced by educational designers or educators when they create these learning environments. When creating learning environments, a question that remains is whether and how affordances are perceived by learners and whether the learning processes and activities actually emerge as intended. It is known that learners, and also significant others, tend to interpret a learning environment and its affordances from their own frame of reference. Personal interpretation, combined with personal preferences they might have for certain types of learning and their current vocational repertoire, determine the actual learning processes and activities that unfold. For this chapter, we started our analyses from these actual

learning processes and activities and reasoned back to what might have afforded such learning. While the frames of reference and repertoires of the learners and significant others are of course still a factor, we think this approach provides relevant insights about how to design for learning.

References

- Akkerman, S.F., Bakker, A., 2011. Boundary crossing and boundary objects. *Rev. Educ. Res.* 81, 132–169.
- Akkerman, S.F., Bakker, A., 2012. Crossing boundaries between school and work during apprenticeships. *Vocat. Learn.* 5 (2), 153–173.
- Andersson, P., Köpsén, S., 2019. VET teachers between school and working life: boundary processes enabling continuing professional development. *J. Educ. Work* 32 (6–7), 537–551.
- Bakker, A., Akkerman, S.F., 2014. A boundary-crossing approach to support students' integration of statistical and work-related knowledge. *Educ. Stud. Math.* 86 (2), 223–237.
- Behle, H., 2017. Developing vocational competences during secondary school? *Eur. J. Train. Dev.* 41 (1), 39–49.
- Beijaard, D., Meijer, P.C., Verloop, N., 2004. Reconsidering research on teachers' professional identity. *Teach. Teach. Educ.* 20 (2), 107–128.
- Bijlsma, N., Schaap, H., De Bruijn, E., 2016. Students' meaning-making and sense-making of vocational knowledge in Dutch senior secondary vocational education. *J. Vocat. Educ. Train.* 68 (3), 1–17.
- Billett, S., 2011. *Vocational Education*. Springer, Dordrecht.
- Bisholt, B.K., 2012. The professional socialization of recently graduated nurses—experiences of an introduction program. *Nurse Educ. Today* 32 (3), 278–282.
- Bouw, E., Zitter, I., De Bruijn, E., 2021a. Designable elements of integrative learning environments at the boundary of school and work: a multiple case study. *Learn. Environ. Res.* 24 (3), 487–517.
- Bouw, E., Zitter, I., De Bruijn, E., 2021b. Exploring co-construction of learning environments at the boundary of school and work through the lens of vocational practice. *Vocat. Learn.* 14, 559–588.
- Carvalho, L., Goodyear, P., 2018. Design, learning networks and service innovation. *Des. Stud.* 55, 1–27.
- Ceelen, L., Khaled, A., De Bruijn, 2019. Begeleiden van studenten op de werkplek [Guiding students at the workplace]. *Onderwijs en Gezondheidszorg* 43 (5). <https://onderwijsengezondheidszorg.nl/jaargangen/2019/5-sep/begeleiden-van-studenten-op-de-werkplek.html>.
- Chan, S., 2013. Learning through apprenticeship: belonging to a workplace, becoming and being. *Vocat. Learn.* 6 (3), 367–383.
- Chan, S., 2015. Apprentices' learning of occupationally informed practical judgment. *Vocat. Learn.* 8 (3), 335–351.
- Chan, S., 2019. From job to calling: vocational identity and the role of apprenticeship. *Vocat. Learn.* 12 (3), 387–403.
- Colley, H., James, D., Diment, K., Tedder, M., 2003. Learning as becoming in vocational education and training: class, gender and the role of vocational habitus. *J. Vocat. Educ. Train.* 55 (4), 471–498.
- De Bruijn, E., 2012. Teaching in innovative vocational education in the Netherlands. *Teach. Teach. Theor. Pract.* 18 (6), 637–653.
- Dixon-Woods, M., Agarwal, S., Young, B., Jones, D., Sutton, A., 2004. *Integrative Approaches to Qualitative and Quantitative Evidence*. Health Development Agency, London, p. 181.
- Emad, G.R., Roth, W.M., 2016. Quasi-communities: rethinking learning in formal adult and vocational education. *Instr. Sci.* 44 (6), 583–600.
- Engeström, Y., 2001. Expansive learning at work: toward an activity theoretical reconceptualization. *J. Educ. Work* 14 (1), 133–156.
- Ennals, P., Fortune, T., Williams, A., D'Cruz, K., 2016. Shifting occupational identity: doing, being, becoming and belonging in the academy. *High. Educ. Res. Dev.* 35 (3), 433–446.
- Entwistle, N.J., Peterson, E.R., 2004. Conceptions of learning and knowledge in higher education: relationships with study behaviour and influences of learning environments. *Int. J. Educ. Res.* 41 (6), 407–428.
- Ferm, L., Persson Thunqvist, D., Svensson, L., Gustavsson, M., 2018. Students' strategies for learning identities as industrial workers in a Swedish upper secondary school VET programme. *J. Vocat. Educ. Train.* 70 (1), 66–84.
- Goh, A.Y.S., 2015. An individual learning journey: learning as becoming a vocational teacher. *Int. J. Lifelong Educ.* 34 (6), 680–695.
- Herrera, H., Brown, D., Portlock, J., 2015. Foundation degree learning: an educational journey of personal development. *J. Furth. High. Educ.* 39 (6), 839–861.
- Heusdens, W.T., Bakker, A., Baartman, L.K.J., De Bruijn, E., 2016. Contextualising vocational knowledge: a theoretical framework and illustrations from culinary education. *Vocat. Learn.* 9, 151–165.
- Heusdens, W., Baartman, L., De Bruijn, E., 2019. Know your onions: an exploration of how students develop vocational knowledge during professional performance. *Scand. J. Educ. Res.* 63 (6), 839–852.
- Hodkinson, P., Biesta, G., James, D., 2008. Understanding learning culturally: overcoming the dualism between social and individual views of learning. *Vocat. Learn.* 1 (1), 27–47.
- Hökkä, P., Vähäsantanen, K., Paloniemi, S., 2020. Emotions in learning at work: a literature review. *Vocat. Learn.* 13 (1), 1–25.
- Holmgren, R., Haake, U., Söderström, T., 2017. Firefighter learning at a distance—a longitudinal study. *J. Comput. Assist. Learn.* 33 (5), 500–512.
- Howell, D., 2009. Occupational therapy students in the process of interprofessional collaborative learning: a grounded theory study. *J. Interprof. Care* 23 (1), 67–80.
- Johansson, M.W., Wärvik, G.B., Choy, S., 2019. Vocationalising specialized concepts: appropriating meanings through feedback. *Vocat. Learn.* 12 (2), 197–215.
- Jonasson, C., 2015. Interactional processes of handling errors in vocational school: students attending to changes in vocational practices. *Vocat. Learn.* 8 (1), 75–93.
- Knightbridge, L., 2014. Experiential learning on an alternative practice education placement: student reflections on entry-level competency, personal growth, and future practice. *Br. J. Occup. Ther.* 77 (9), 438–446.
- Koenen, A.K., Dochy, F., Berghmans, I., 2015. A phenomenographic analysis of the implementation of competence-based education in higher education. *Teach. Teach. Educ.* 50, 1–12.
- Koopman, M., De Bruijn, E., Baartman, L.K.J., in preparation. *Learning Processes in Vocational Education and Vocational Practice: Conceptualizations, Operationalizations, and Empirical Evidence*.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Leont'ev, A.N., 1978. *Activity, Consciousness, and Personality*. Prentice-Hall, Hoboken.
- Lie, D.A., Forest, C.P., Walsh, A., Banzali, Y., Loheny, K., 2016. What and how do students learn in an interprofessional student-run clinic? An educational framework for team-based care. *Med. Educ. Online* 21 (1), 31900.
- Ord, K., Nuttall, J., 2016. Bodies of knowledge: the concept of embodiment as an alternative to theory/practice debates in the preparation of teachers. *Teach. Teach. Educ.* 60, 355–362.
- Poortman, C.L., Illeris, K., Nieuwenhuis, L., 2011. Apprenticeship: from learning theory to practice. *J. Vocat. Educ. Train.* 63 (3), 267–287.
- Pridham, B., O'Mallon, S., Prain, V., 2012. Insights into vocational learning from an applied learning perspective. *Vocat. Learn.* 5 (2), 77–97.
- Reegård, K., 2015. Sales assistants in the making: learning through responsibility. *Vocat. Learn.* 8 (2), 117–133.
- Schoeman, S., Mabunda, P.L., 2012. Teaching practice and the personal and socio-professional development of prospective teachers. *S. Afr. J. Educ.* 32 (3), 240–254.
- Smith, R., 2012. Clarifying the subject centred approach to vocational learning theory: negotiated participation. *Stud. Cont. Educ.* 34 (2), 159–174.
- Stets, J.E., Burke, P.J., 2000. Identity theory and social identity theory. *Soc. Psychol. Q.* 63 (3), 224–237.

- Tynjälä, P., Beusaert, S., Zitter, I., Kyndt, E., 2021. Connectivity between education and work: theoretical models and insights. In: Kyndt, E., Beusaert, S., Zitter, I. (Eds.), *Developing Connectivity Between Education and Work*. Routledge, London, pp. 3–14.
- Wallin, A., Nokelainen, P., Mikkonen, S., 2019. How experienced professionals develop their expertise in work-based higher education: a literature review. *High. Educ.* 77 (2), 359–378.
- Zitter, I., Hoeve, A., 2012. *Hybride leeromgevingen: Het verweven van leer-en werkprocessen* [Hybrid learning environments: The interweaving of learning and work processes]. ECBO, 's-Hertogenbosch.

Language education

Jessie S. Barrot, National University, Manila, Philippines

© 2023 Elsevier Ltd. All rights reserved.

Background	216
English curriculum in the Philippines	217
Language curriculum 5.0	217
Forms of language curriculum	217
Balancing outcome and process	217
Technology integration	218
Sociocultural aspect of LC5	218
Self-management and life skills	219
Outcomes-based pedagogy	219
LC5 components	219
Implementing an evidence-based language curriculum 5.0	220
Implementation challenges	221
Conclusion	222
References	222

Background

The recent advancement in information and communication technology (ICT) and the growing interdependence of the world's economies and cultures are constantly reshaping the education landscape. Recently, the World Economic Forum has released a report on the future of jobs in the fourth industrial revolution (4IR). 4IR builds on the third revolution (i.e., the digital revolution) and represents a fundamental change in our society and world economies. As an era of cyber-physical systems, the internet of things, and network, 4IR promotes a real-time world where information is instantly provided, and almost everything can be delivered directly to your door. It is the convergence of digital, physical, and biological innovations, such as artificial intelligence, robotics, augmented reality, 3D printing, and genome editing. This revolution has resulted in profound social, economic, political, and cultural transformations on a global scale. Some jobs will disappear, while others grow. Some skills will be displaced, while others thrive. [Table 1](#) shows the top ten trending and declining skills by 2022 ([World Economic Forum, 2018](#), p. 12).

However, 4IR has been heavily criticized for being technology-centric in which smart technologies, automation, and connectivity are overly emphasized. This developing scenario has eclipsed the service of humanity and instigates a lack of confidence of humans in new technologies that take humans out of the industry. As a response to this emerging concern, the fifth industrial revolution (5IR) came about, which is the synergy between human intelligence and technology to further increase efficiency ([Nahavandi, 2019](#)). It is a period of development where humanity is brought back to industry along with the advancement in technology. To put it simply, 5IR is the humanistic version of 4IR with the ultimate purpose of contributing to the quality of human life and not displacing humans.

These developments fundamentally compel the field of education to adapt (i.e., Education 5.0) and produce 21st century multiliterate learners who are capable of contributing to and engaging in this knowledge-based innovation-driven society. In response, educational agencies in many countries have revisited and beefed up their respective curricula to address societal and economic demands. However, these curriculum reforms should not only focus on producing employable and ICT literate students but also socially critical and productive citizens ([Jackson, 2014](#)). The same effort has transcended to the field of language education, in which several language curriculum reviews have been initiated. However, several studies (e.g., [Barrot, 2019a](#); [San Isidro, 2018](#); [Zhu and Shu, 2017](#)) have reported some challenges in developing and implementing state-of-the-art language curricula. Thus, this paper seeks to address some of those by offering insights into how a language curriculum innovation may be adopted and implemented. In other words, this paper offers a new curriculum design known as Language Curriculum 5.0 (LC5). LC5 is an innovative-producing curriculum that focuses on developing future-ready and competent communicators within the 5IR milieu. Specifically, the first section briefly explains the context of language curriculum innovation. The second section discusses the characteristics and components of the LC5, while the third section talks about key challenges in implementing this curriculum innovation. The paper concludes with suggestions on how can an evidence-based language curriculum innovation be systematically implemented. Since language curriculum is evolutionary rather than revolutionary, this paper does not advocate that time-tested practices be abandoned. Instead, it seeks to add value to what has already been existing. It is also worthy to note that this article is written within a mainly Philippine curriculum research tradition to highlight different perspectives and contexts.

Table 1 Trending and declining skills by 2022.

<i>Trending skills</i>	<i>Declining skills</i>
Analytical thinking and innovation	Manual dexterity, endurance, and precision memory
Active learning and learning strategies	Verbal, auditory, and spatial abilities
Creativity, originality, and initiative	Management of financial and material resources
Technology design and programming	Technology installation and maintenance
Critical thinking and analysis	Reading, writing, math, and active listening
Complex problem solving	Management of personnel
Leadership and social influence	Quality control and safety awareness
Emotional intelligence	Coordination and time management
Reasoning, problem solving, and ideation	Visual, auditory, and speech abilities
Systems analysis and evaluation	Technology use, monitoring, and control

English curriculum in the Philippines

As a response to students' poor performance in the National Achievement Test (NAT) and standardized English proficiency tests, the Philippine government introduced the Language Arts and Multiliteracies Curriculum (LAMC) in 2013. This curriculum reform sought to produce multiliterate and communicatively competent students to thrive in this innovation-driven society (Department of Education, 2016). Unlike its predecessor, LAMC is decongested by adding two more years of English courses (i.e., *Reading and Writing*, *Oral Communication*, and *English for Academic and Professional Purposes*) in the senior high school. LAMC also incorporated some pedagogical changes. First, it adopted the Mother-tongue-based Multilingual Education (MTBMLE), which uses the students' mother tongue as a medium of instruction from grades 1 to 3. Second, it shifted from NAT to end-of-grade 6, 10, and 12 assessments. Finally, it reduced the time allotment for English subjects both in elementary and junior high school levels. Despite these efforts, recent studies (e.g., Barrot, 2019a) have revealed that LAMC fell short in aligning it with 21st century learning context and established language learning and teaching principles. The findings also revealed that LAMC lacks specificity and coherence, which may contribute to widening the gap between the intended and implemented curriculum as well as what learners need to learn and what teachers want to teach. Thus, Barrot (2019a) highlighted the need to adopt an evidence-based and future-proof English curriculum.

Language curriculum 5.0

Forms of language curriculum

A language curriculum is a document usually prepared by a leading or government education agency that guides the teachers, teacher-trainers, textbook developers, language testers, and other key stakeholders (Celce-Murcia and Olshtain, 2000). It is an educational document that undergoes constant evolution as the field of language education expands. A language curriculum can be content-based, process-based, and product-based. A content-based curriculum focuses on at least three different contents, namely linguistic, thematic or situational, and subject-matter. A process-based curriculum, on the one hand, focuses on language learning and the teaching process. It may either be a needs-based, task-based, or learner-based curriculum. And then, there is the product-based curriculum that puts emphasis on the outcome of the course.

Given the current language educational landscape and advances in technology, language educators are urged to consider taking a paradigm shift in curriculum development. A state-of-the-art language curriculum should be able to see the future and be able to match the target skills of the curriculum to the demands of future jobs and society. As recently reported in the *World Economic Forum* (2018), 65% of children who entered schools today are likely to land in jobs that are not yet existing today. This scenario compels all curricula to be "ahead of its time" and shift toward Curriculum 5.0, which can be adapted in different subject areas. In the case of language education, it is LC5. Unlike the existing curriculum that is primarily product-based, this language curriculum innovation is a layered type of curriculum that systematically fuses the features of the different forms of curriculum (Barrot, 2015a; Brown, 1995). This means that while the curriculum targets the achievement of outcomes, it equally emphasizes the process of achieving such outcomes. Meanwhile, content is heavily fused into the curriculum as a means to achieving the outcomes and performing processes.

Balancing outcome and process

Equally important as the output is the process. LC5 highlights the significance of teaching processes in a language classroom. This requires curriculum developers to cover learning strategies as one of the focus areas in LC5, a feature underrepresented in the current English curriculum, particularly in the higher grade levels. Teachers can also be urged to demonstrate or model for the students the processes in creating spoken, written, or multimodal outputs (*Curriculum Planning and Development Division*, 2020) through the teacher's guide in LC5. The value of emphasizing the process is it makes learners realize that a language

is a tool for higher purposes and not an end in itself. The process orientation is associated with the task-based, experiential, and humanistic approaches in which the focus is on the learning process rather than the learning products alone and formative rather than summative assessment. An example of a language curriculum that integrates process orientation is the Singapore English Syllabus 2020.

Technology integration

The third important characteristic of a language curriculum in a 5IR environment is technology integration. ICT has significantly changed the landscape of language education across contexts. It even led to the emergence of a new subfield in language education known as technology-enhanced language learning (TELL) which refers to learning a language in any context with, through, and around technologies. Given this, it is inevitable that ICT be systematically integrated into language teaching and curriculum, not as a method, but as a powerful tool in enriching content and creating a diverse learning environment (Kenning, 2007; Wilkinson, 2016). Unlike the LAMC that is silent about ICT integration (Albuero et al., 2019; Barrot, 2019a), LC5 comprehensively and explicitly discusses how ICT can be used as a “tool” for teaching, learning, and assessment and how it can be used as a “tutor” for developing learners’ ICT skills. However, doing so may raise few issues, such as the lack of technical skills of teachers and students and insufficiency of ICT resources.

Closely related to ICT integration is the development of learners’ multiliteracies. Multiliteracy refers to learners’ ability to collect and/or retrieve information, organize and manage information, evaluate the relevance, quality, and usefulness of information, and generate accurate information through the use of available resources (Educational Testing Service, 2007). With the introduction of multiliteracy, a new form of pedagogical construct emerged, coined as multiliteracy pedagogy. It tasks LC5 to provide learners with opportunities to access, evaluate, search, sort, gather, and read data from various multimedia and multimodal sources. The multiliteracy skills that need to be developed among learners are basic literacy (numeracy and language proficiency), scientific literacy (knowledge of scientific concepts and processes), economic literacy (knowledge of economy), technological literacy (knowledge and skills in using technology), visual literacy (ability to process moving and non-moving images), information literacy (ability to process and evaluate information from various media), and multicultural literacy (ability to understand and relate with one’s culture and of others).

Sociocultural aspect of LC5

Aside from focusing on how language is learned and should be treated, a state-of-the-art language curriculum puts equal emphasis on how language functions as a medium for building interpersonal relations, social transactions, and intercultural competence. While there are attempts to fuse the sociocultural aspect of language learning into the LAMC, such integration is trivial as it stops at the realm of intention. Therefore, LC5’s goal is not just to help learners acquire language skills but also to help them perform social functions, actively participate in the community where they belong, and contribute to self- and social transformation. This idea is based on the principles of transformative learning and critical pedagogy that takes a functional-interactional view of language. Transformative learning goes beyond the skills, process, and content knowledge acquisition; it is a learning theory that allows learners to question how and what they know about life experiences such as culture, religion, politics, and personalities (Mezirow, 2000). In the same vein, critical pedagogy refers to a set of education principles that relate classroom experience to social transformation and wider social context to promote justice and equality (Akbari, 2008; Crookes, 2012). To translate these principles into tangible concepts, LC5 may need to specify themes, texts, and tasks that are not just prepared to impart knowledge but believed to be socially and culturally constructed to empower people. LC5 should also be able to help learners reflect on their own culture and life experience, critically view the society through dialog with others and transform society through active social actions. Community-based service learning or participatory learning communities may be explored for this purpose as it allows learners to have the power and control, through their active participation, to make their lives and world better (Schneider, 2019). Through these strategies, the curriculum links the classroom with students’ realities beyond their school.

Globalization and localization in language education are not two opposing views. Rather, they are complementary views that make language curriculum global without compromising local identity. Reconciling these two views has led to the new construct known as glocalization which is a mixture of global and local knowledge and practices. To simply put it, glocalization takes on the mantra, *think globally and act locally*. Having a glocalized language curriculum allows learners to understand the language and behavior of other cultures and communicate with people from other cultures without losing their local identity (Byram, 1997; St Clair and Phipps, 2008). There are two ways that glocalization can be adopted for LC5 development. First, LC5 should present English as an international commodity (also known as world Englishes) rather than a property of English-speaking countries. It has to recognize its own variety of English appropriate for its local context. Second, LC5 needs to consider intertextuality when selecting texts. This means that suggested texts in LC5 should be carefully selected in a way that there is a balance and link between local and foreign texts (i.e., spoken, written, or multimodal). Therefore, learners should be exposed to various texts of local and international concerns, allowing them to bring a local issue to a global context.

Culture and curriculum are two interdependent constructs (Byrnes, 2010; Lavrenteva and Orland-Barak, 2015). Any language curriculum cannot exist without considering the culture where it is situated. LC5 developers are expected to consider the setting (interactional and physical), behavioral environment (kinetic and nonverbal), language, and extracontextual factors (sociopolitical,

historical, and environmental). Although curricular borrowing (i.e., adopting the language curriculum of other countries) may look practical, it usually fails locally as it contains elements that do not fit with local culture and norms (Carless and Harfitt, 2013). Take for example the use of strong task-based language teaching in some Asian countries (Butler, 2011; Cao, 2018). LC5 has to be personalized as well by considering the personal interest, motivations, and learning style of the target learners prior to its development. In many cases, curriculum developers use their common-sense intuition based on experience and practical knowledge during the development process. Although it is practical and has some significance, decisions solely based on intuition may have some serious setbacks as it is prone to overestimation or underestimation (Barrot, 2019b). As such, curriculum developers are urged to survey learners to get hold of data needed for a more personalized language curriculum. Nonetheless, LC5 needs to be flexible as well to cater to the learning needs and styles of other learners (Nel et al., 2016). A flexible curriculum is designed to be accessible for diverse learners regardless of their background and creates many opportunities for learners to attain the ILOs. It also allows teachers to explore various possibilities in executing a lesson, thereby empowering them and developing their creativity as facilitators of learning. Adopting a flexible curriculum means prioritizing quality over quantity. Certain provisions in LC5 need to provide opportunities for schools and teachers to individualize and contextualize the curriculum to their respective learners.

Self-management and life skills

Another humanistic feature of LC5 is its emphasis on learners' development of their self-management and life skills. Regrettably, these skills seem to be neglected in many curricula, as in the case of the Philippine K-12 curriculum. This set of skills comprises, but is not limited to, adaptability, managing complexity, self-direction, self-monitoring, curiosity, resource utilization, decision making, risk-taking, emotional control, and sound reasoning. Young and adult learners alike are prone and vulnerable to high-risk situations and may eventually succumb to these challenges. But when equipped with essential life skills coupled with psychosocial support from their environment, learners can deal with these challenges effectively (The Partnership for 21st Century Learning, 2015). The question is how they can be incorporated into LC5. One approach is to embed these skills into the affective ILOs in the learning guide. In such a case, life and self-management skills may be selected based on the set of texts or themes assigned in each lesson, quarter, or unit of LC5. Additionally, these skills may be reinforced in the suggested activities and assessment sections of the curriculum.

One way to reinforce self-management and life skills is through problem-based language learning, where learners are asked to perform a task similar to real-world problems that they may encounter beyond the four corners of the classroom. It specifically taps learners' creative, innovative, and analytical skills while allowing them to experience complex problem-solving situations (Chen and Kuo, 2019; Moallem et al., 2019). To incorporate this into the curriculum, well-defined intended learning outcomes (ILOs) and contexts that target the development of learners' problem-solving skills may be explored.

Outcomes-based pedagogy

In the case of some curriculum reforms (e.g., LAMC), although the target learning competencies or outcomes are comprehensive, problems on constructive alignment and specificity have been observed. In addition, these learning competencies or outcomes tend to be overly cognitive and academic, making it difficult for learners to bring their learning beyond the four walls of the classroom. On this note, LC5 systematically incorporates the principles of outcomes-based pedagogy to produce work-ready graduates who possess skills and abilities that match the demands of industry and community. Central to outcomes-based pedagogy are the ILOs and the ways on how the curriculum and teaching-learning processes revolve around achieving these ILOs (Biggs, 1996; Spady, 1994).

To achieve this, Biggs (1996) proposed the idea of constructive alignment, which emphasizes the consonance among the teaching-learning activities, assessment methods, and ILOs. To apply this in LC5 development, four steps must be taken: (1) clearly define the ILOs; (2) specify the teaching-learning activities that will scaffold the achievement of the ILOs; (3) assess students' performances and check how they match the ILOs; and (4) rate the performance. Note that multicomponential assessment plays a key role in an outcomes-based curriculum; thus, it should be clearly specified in LC5. By multicomponential, we mean using various forms of assessment (formative and summative, traditional and alternative, self-, peer and teacher assessment, formal and informal, assessment of learning, assessment for learning and assessment as learning) to get the full picture of learners' progress. During assessment activities, emphasis should be given to the process, formative assessment, and authenticity to mimic real-world contexts. However, all assessment activities should be within the social realities of the learners. For instance, a grade 3 pupil should not be given a bank transaction task. Although it is authentic, the task is beyond the social reality of a grade 3 pupil.

LC5 components

Given the features above, Table 2 presents the suggested components of LC5. The list is not exhaustive, but it will provide a good idea of how LC5 looks like.

Table 2 LC5 components.

<i>Components</i>	<i>Description</i>
Context of language use	This section explains how language is viewed and used in a given context and the local and global factors that shape it. It concludes with the desired outcomes for language learners of the curriculum.
Language curriculum framework	This section presents the schematic framework and key features of LC5. It also explains the general and specific aims of the language curriculum and how its different elements contribute to achieving them. These elements include the theories of language and language learning (e.g., sociocognitive theory), the language teaching and learning principles (e.g., social interaction and collaboration, macroskills integration, microskills integration, form-meaning connection, contextualization, intercultural competence, ICT integration, spiral progression, process orientation, differentiation, reflective learning, and 21st century learning) and the learning processes (e.g., problem-based, project-based, and task-based learning) that underpin the curriculum (Barrot, 2014, 2015b).
Areas of language learning	This section maps out the various language learning ILOs at multiple levels and shows their constructive alignment. The different ILO levels include core learning area ILOs (i.e., overarching ILO across all grade levels) per macroskill or cluster of macroskills; key stage ILOs (i.e., overarching ILO for primary and for secondary levels), grade-level ILOs (i.e., ILOs specific to one grade level); and focus area ILOs (i.e., a set of grade-level ILOs in each focus area, such as listening, reading, speaking, writing, representing, viewing, grammar, and vocabulary). It is expected that content, skills, strategies, and attitude/values are sufficiently addressed in the crafting of ILOs at multiple levels.
Language assessment guide	Since multicomponential assessment plays a key role in LC5, a comprehensive language assessment guide merits serious consideration. This section details the language assessment framework and activities that align with the goals of LC5 and its different ILOs. It also provides specific guidelines on how assessment can be implemented in different language learning contexts.
Language teaching guide	This section details the role of teachers in language classrooms and offers some approaches, methods, techniques, and strategies in implementing the lesson. It also offers tips on ICT integration and additional resources for language teaching. Finally, this section provides sample lessons that reflect LC5 principles and ILOs.

Implementing an evidence-based language curriculum 5.0

To better understand and implement curriculum reforms, it requires to be viewed from multiple perspectives. Unfortunately, teachers and other key stakeholders have very limited involvement in the work of the panel and the development process in many countries, such as the Philippines. As a way to move forward, the government education agencies may take a more transactional approach that allows end-users (i.e., teachers, school leaders, and students) and other key stakeholders to be actively involved in the development process from designing the curriculum to its implementation. One way to do this is to intensify public hearings and meetings that invite end-users to make comments on the drafts before the curriculum is finalized.

To do this, the following steps are recommended. As shown in Fig. 1, the first step is to create a technical panel that will bring together language curriculum experts (i.e., practitioners and researchers), teachers, students, school leaders, government education agencies, professional organizations, industry, community leaders, and parents. This multisectoral panel is tasked to provide multiple perspectives in curriculum development and implementation. Second, it is crucial to survey the students on their learning preferences, learning style, interest, and motivations. The collected information can be used in selecting tasks, texts, and topics, crafting learning competencies, and choosing appropriate resources, among others. Third, the crafting of the preliminary curriculum may be initiated once all necessary information is collected and analyzed. A team of expert curriculum developers may take the lead during the drafting stage. Fourth, the LC5 draft may be subjected to experts' face and content validation to further refine it prior to piloting. During content validation, experts evaluate the alignment, specificity, and coherence of the LC5. Face validation, on the one hand, focuses on the clarity, accuracy, and aesthetics of the curriculum. After the experts' validation phase, the draft curriculum may be piloted in strategic areas that represent various demographics of learners for at least one year. Consider the diversity of the selected areas during piloting so that the curriculum would be able to capture as many learning contexts as possible. Diversity relates to location (urban, suburban, and rural areas), performance (high performing and low performing schools based on achievement tests), resources (rich and low resources), socioeconomic status (class A, B, C, D, or E), and language proficiency level (high, mid, and low proficient) among others.

The Evaluation and Monitoring Team (EMT) plays a crucial role in the success of the curriculum implementation. It will take the lead in assessing the impact of LC5 on teachers, school leaders, and learners. The EMT may conduct classroom observation that focuses on the implementation challenges and impact of resources on instructional delivery and learning. The team may also interview teachers and school leaders to assess their level of understanding of LC5 and to seek suggestions from them in addressing the encountered challenges. For learners, the EMT may take the lead in assessing their performance before and after the implementation to see the viability of the curriculum in terms of learning. Learners may also be engaged in a focus group discussion, allowing them to voice out their learning experience. In the first place, they are the ultimate consumers of curriculum innovation. All the information collected during the monitoring and evaluation phase may be used to further refine the curriculum and its implementation strategies. Another round of piloting is preferred but not necessary. Once the curriculum and the implementing guidelines are finalized, the curriculum developers and the technical panel may begin developing the teaching guide, sample lesson plans, and

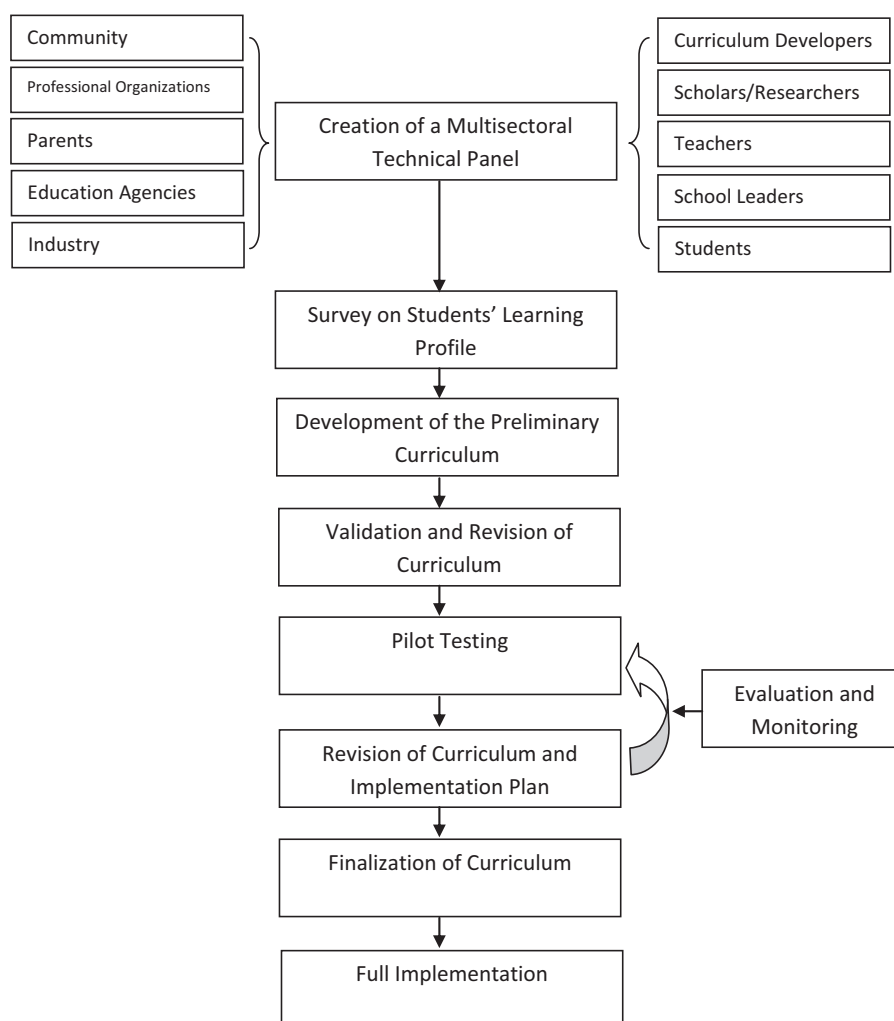


Fig. 1 Schematic representation of implementation framework.

assessment guide to ensure that sufficient information is given to the end-users (i.e., teacher, school leaders, and textbook developers). The assessment guide should be comprehensive enough to cover all the ILOs in LC5. When all things are set, full implementation of the new curriculum may commence. Since monitoring and evaluation is as crucial as the implementation, the central governing body or the technical panel may institutionalize an evaluation and monitoring department to ensure the effective implementation of LC5.

Implementation challenges

Once LC5 is developed, one may think that the job is done, but it is not. Development is just one part of the pie; curriculum implementation is another matter. Curriculum implementation is a complex process, which needs thorough planning and coordination with various stakeholders. However, in some cases, it is not given the level of attention it deserves. And so many problems arise, but I will only focus on few major ones based on [Vidovich's \(2002\)](#) policy cycle framework that discusses three contexts: the contexts of influences, context of practice, and context of policy text production. The context of influence refers to the struggles among stakeholders over the construction of policy discourses. The context of policy text production refers to the articulation of the curriculum despite its inconsistencies. Finally, the context of practice relates to how various stakeholders (e.g., teachers and school leaders) interpret the curriculum.

The first major challenge confronted when implementing curriculum innovation is the gap between curriculum intentions and how they are translated and learned at the classroom level. One factor that contributes to this widening gap is the centralized and bureaucratic process in its development and implementation. Common to the language development and implementation process is the top-down approach in which policies and procedures emanate from a central governing body. The problem with this approach is that teachers lack ownership and participation in the reform. And if ever they attempt to implement it, such an effort

is superficially adopted. Aggravating the situation is the potential clash between what the curriculum says and what the teacher believes and practices. In such a case, it is unlikely that teachers take curriculum innovation seriously (Carless, 2013).

The second major implementation challenge arises from the curriculum itself. Developing a pedagogically sound, specific, and coherent curriculum is not an easy task because it requires the government to bring together the right people for the job. So if the development process fails to properly execute coordination or fails to choose the right people, such failures may lead to inconsistencies and content issues. The common problems seen in a not so well-thought off language curriculum are its lack of alignment, specificity, and coherence (Barrot, 2019a; Howard and Millar, 2008). It has been observed that some language curricula fell short in integrating the established principles of language pedagogy and lack solid pedagogical frameworks that would underpin the curriculum. Consequently, the language curriculum becomes a hodgepodge of many elements without focus. Details in the elements and procedure of any language curriculum are important to its successful implementation. However, not all sufficiently explain their intentions, principles, and features. Take for example a language curriculum that lacks details on ICT integration as a tool and as a tutor, lacks domains of language skills (e.g., viewing and representing), and lacks an assessment guide. In terms of coherence, some learning outcomes and topics are introduced without sufficient scaffolding. Meaning, there is no systematic progression of learning outcomes and standards. Within one lesson, constructive alignment is sometimes breached. For instance, a week's lesson requires learners to produce a narrative, but the target linguistic form under the grammar domain is *prepositions* instead of *past tense*. In some instances, the target grammatical item is *prepositions for giving directions*, but the target text under the writing domain is a *feature article*.

The last implementation challenge is how the teachers and other stakeholders interpret the curriculum. As earlier mentioned, lack of substantial participation from teachers and school leaders and the curriculum incoherence, vagueness, and misalignment contribute to its misinterpretation and misapplication in the classroom. The other factors that contribute to the misapplication of the curriculum are uncommitted school leadership, lax monitoring and evaluation, minimal involvement of parents, limited school resources, poor school environment, teachers' administrative load, illiteracy levels of learners, limited training, mismatch between national standardized assessment thrust of the curriculum, and insufficient time (may be linked to overcrowded curriculum) that forces teachers to prioritize quantity over quality.

Conclusion

This article has shown the recent social and economic developments and how they impacted education, particularly language education. As a way to address the demands of the 5IR environment, LC5 is proposed. This state-of-art curriculum aims to develop future-ready and competent communicators who can thrive in an innovative-driven environment. The key features that set LC5 apart from the conventional are as follows: problem-based, process-oriented, technology-based, multiliteracy focused, socially and community-oriented, life skills and self-management oriented, outcomes-based, glocalized, and contextualized and personalized yet flexible. Nonetheless, possible challenges for its implementation were presented, and strategies to address them were offered.

While this paper offers insights into designing and implementing a state-of-the-art language curriculum, will and commitment of key people remain to be the key ingredient. No amount of curriculum innovation will succeed without the concerted effort of its key stakeholders, namely the policymakers, the school leaders, the teachers, and the students. As a final note, the proposed language curriculum innovation is suggestive rather than prescriptive. It is merely the beginning of the new era of language curricula. And as long as the society evolves, so as the curriculum.

References

- Akbari, R., 2008. Transforming lives: introducing critical pedagogy into ELT classrooms. *ELT J.* 62 (3), 276–283.
- Alburo, R.M., Flores, E.A., Yee, J., Barrot, J.S., 2019. Examining the secondary English curriculum in the Philippines: a step to move forward. *Curric. Teach.* 34 (1), 5–21.
- Barrot, J.S., 2014. A macro perspective on key issues in English as Second Language (ESL) pedagogy in the postmethod era: confronting challenges through sociocognitive-transformative approach. *Asia Pac. Educ. Res.* 23 (3), 435–449.
- Barrot, J.S., 2015a. A sociocognitive-transformative approach to teaching writing. *Indones. J. Appl. Linguist.* 4 (2), 113–122.
- Barrot, J.S., 2015b. A sociocognitive-transformative instructional materials design model for second language (L2) pedagogy in the Asia Pacific: development and validation. *Asia Pac. Educ. Res.* 24 (2), 283–297.
- Barrot, J.S., 2019a. English curriculum reform in the Philippines: issues and challenges from a 21st century learning perspective. *J. Lang. Ident. Educ.* 18 (3), 145–160.
- Barrot, J.S., 2019b. Examining the task complexity in ELT coursebooks. *Asia Pac. Educ. Res.* 28 (6), 469–481.
- Biggs, J., 1996. Enhancing teaching through constructive alignment. *High Educ.* 32 (3), 347–364.
- Brown, J.D., 1995. *The Elements of Language Curriculum. A Systematic Approach to Program Development.* Heinle & Heinle, Boston, MA.
- Butler, Y.G., 2011. The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annu. Rev. Appl. Ling.* 31, 36–57.
- Byram, M., 1997. *Teaching and Assessing Intercultural Communicative Competence.* Multilingual Matters, Clevedon.
- Byrnes, H., 2010. Revisiting the role of culture in the foreign language curriculum. *Mod. Lang. J.* 94 (2), 315–317.
- Cao, P.T.H., 2018. Task-based language teaching: affordances and challenges in TBLT implementation at the Vietnamese Tertiary Level. *J. Asia TEFL* 15 (2), 510–515.
- Carless, D., Harfitt, G., 2013. Innovation in secondary education: a case of curriculum reform in Hong Kong. In: Hyland, K., Wong, L.L.C. (Eds.), *Innovation and Change in English Language Education.* Routledge, Abingdon, pp. 172–185.
- Carless, D., 2013. Innovation in language teaching and learning. In: Chapelle, C. (Ed.), *The Encyclopedia of Applied Linguistics.* Wiley-Blackwell, Oxford, pp. 1–4.
- Celce-Murcia, M., Olshtain, E., 2000. *Discourse and Context in Language Teaching: A Guide for Language Teachers.* Cambridge University Press, Cambridge.

- Chen, C.M., Kuo, C.H., 2019. An optimized group formation scheme to promote collaborative problem-based learning. *Comput. Educ.* 133, 94–115.
- Crookes, G., 2012. Critical pedagogy in language teaching. In: Ortega, L. (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell, Oxford.
- Curriculum Planning & Development Division, 2020. English Language Syllabus: Primary & Secondary (Express/Normal [Academic]). Ministry of Education, Singapore.
- Department of Education, 2016. K to 12 Curriculum Guide for English. Available from: http://www.deped.gov.ph/sites/default/files/page/2016/English%20CG_0.pdf. (Accessed 6 June 2016).
- Educational Testing Service, 2007. Digital Transformation: A Framework for ICT Literacy, a Report of the International ICT Literacy Panel. Educational Testing Service, New Jersey.
- Howard, J., Millar, S., 2008. Ongoing challenges with language curriculum innovations in Asia: a South Korean case study. *Pac. Asian Educ. J.* 20 (2), 59–75.
- Jackson, L., 2014. *Muslims and Islam in US Education: Reconsidering Multiculturalism*. Routledge, New York.
- Kenning, M.M., 2007. ICT and language learning. In: Kenning, M.M. (Ed.), *ICT and Language Learning: From Print to the Mobile Phone*. Palgrave Macmillan, London, pp. 135–170.
- Lavrenteva, E., Orland-Barak, L., 2015. The treatment of culture in the foreign language curriculum: an analysis of National Curriculum Documents. *J. Curric. Stud.* 47 (5), 653–684.
- Mezirow, J., 2000. *Learning as Transformation: Critical Perspectives on a Theory in Progress*. Jossey-Bass, San Francisco.
- Moallem, M., Hung, W., Dabbagh, N. (Eds.), 2019. *The Wiley Handbook of Problem-Based Learning*. Wiley Blackwell, London.
- Nahavandi, S., 2019. Industry 5.0—a human-centric solution. *Sustainability* 11 (16), 4371.
- Nel, M., Nel, N., Hugo, A., 2016. Inclusive education: an introduction. In: Nel, M., Nel, N., Hugo, A. (Eds.), *Learner Support in a Diverse Classroom: A Guide for Foundation, Intermediate and Senior Phase Teachers of Language and Mathematics*, second ed. Van Schaik, Pretoria, pp. 3–33.
- San Isidro, X., 2018. Innovations and challenges in CLIL implementation in Europe. *Theor. Pract.* 57 (3), 185–195.
- Schneider, J., 2019. Teaching in context: integrating community-based service learning into TESOL education. *TESOL J.* 10 (1), 1–15.
- Spady, W.G., 1994. *Outcome-Based Education: Critical Issues and Answers*. American Association of School Administrators, Virginia, USA.
- St Clair, R., Phipps, A., 2008. Intercultural literacies. *Lang. Intercult. Commun.* 8 (2), 69–71.
- The Partnership for 21st Century Learning, 2015. P21 Framework Definition. Available from: http://www.p21.org/storage/documents/docs/P21_Framework_Definitions_New_Logo_2015.pdf. (Accessed 15 March 2020).
- Vidovich, L., 2002. *Expanding the Toolbox for Policy Analysis: Some Conceptual and Practical Approaches*. Comparative Education Policy Research Unit. Occasional Paper Series, 2. University of Hong Kong, Hong Kong.
- Wilkinson, M., 2016. Language learning with ICT. In: Renandya, W., Widodo, H. (Eds.), *English Language Teaching Today*. Springer, Cham, pp. 257–276.
- World Economic Forum, 2018. *The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution*. World Economic Forum, Geneva, Switzerland.
- Zhu, Y., Shu, D., 2017. Implementing foreign language curriculum innovation in a Chinese secondary school: an ethnographic study on teacher cognition and classroom practices. *System* 66, 100–112.

Designing school(s): school architecture and curriculum

Lisa Rosén Rasmussen, Danish School of Education, Aarhus University, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction	224
Materializing curriculum and educational discourses	224
Classroom curricula: forms and variations	226
The relationship between school architecture and educational practice	228
Material transfers of curricular designs	230
Building curriculum and future models of schooling	232
References	233

Introduction

In all its visible and invisible forms, school architecture has undergone an extensive transformation. In the middle of the 19th century, the most common type of school in Denmark was the village school, later known as “the thatched school” because of the frequently used roofing material. Such schools would typically be comprised of a small building with a schoolroom at one end and the teacher’s residence at the other. Inside the schoolroom, the children sat at long tables with fixed benches, looking up at the teacher’s desk. A blackboard that could be written on with chalk hung on the wall or stood on the floor next to the teacher’s desk. It was also common for city schools to house both the schoolroom and the teacher’s residence. If there were a lot of children, there might be two separate teaching spaces (Gjerløff et al., 2014). Schools in Denmark today are not only much bigger in size, but also have a wide range of designs and may be built using bricks, steel or glass, with white, bright green or wooden facades. Some stick to traditional classrooms dotted along long corridors. Others contain so-called innovative learning spaces, with rooms in various sizes and shapes. The interior might comprise a simple arrangement of chairs and tables in more or less orderly rows, but teaching spaces often include clustered arrangements, sofas and broad stairs meant for sitting as part of the core furnishings.

The history of schooling is continuously layered within the compound landscape of school buildings with different historical origins, often lying side by side or built into each other as expansions or modernizations of the original schoolhouses. At the same time, school architecture and the physical landscape of schooling exhibit explicit attempts at change when new material designs are invented and hold inertia in the sense of often long-lasting buildings (see Figs. 1 and 2). Buildings from the same time, but with highly dissimilar architectural forms may even be shaping the same local scenery, revealing the coexistence of different educational convictions, as in Mark Dudek’s example of how the conflict between traditional and radical approaches to education was reflected in the compound school architecture at the beginning of the 20th century (Dudek, 2002).

School architecture is about school buildings. It is walls, doors, windows and roofs. It is the building materials, colors, shapes and dimensions of teaching areas and classrooms. School architecture may also be understood as concerning the use and organization of space and about bodies, affects and cultural and sensory perceptions. Moreover, it may be seen as unavoidable that approaches to education, processes of learning and matters of more formal curricula inscribe or entangle themselves into this multi-dimensional and processual understanding of school architecture. This chapter aims to examine both the immediately decipherable and the more subtle links between school architecture and curriculum by viewing school architecture as not only a matter of the architect’s pen, but also as created through policies and pedagogies, as well as through teachers’ and students’ everyday use of space (e.g., Mulcahy and Morrison, 2017; Juelskjær and Rosén Rasmussen, 2019; Grosvenor and Rasmussen, 2018a; Boys, 2011). The paper primarily applies a European perspective, with particular focus on Danish and other Scandinavian examples. Even though the text seeks to explore a diverse range of empirical examples situated differently in space and time, the empirical outset will inevitably shape the analysis and its conclusions, not least when it comes to the question of the presence or absence of an explicitly formulated curriculum.

Materializing curriculum and educational discourses

The relationship between school architecture, the layout of school space and curriculum has a long history. In this history, school architecture and school interior designs of schools have sometimes been built with a direct reference to curriculum, but have more often been circuitously connected through practices and ideas of pedagogy and education. The variations in the presence and role of an explicitly formulated curriculum have been manifold and shifting across countries and through time (Tröhler, 2016). Therefore, exploration of the relationship between school architecture and curriculum entails ongoing discussion of the national and local sense-making regarding a curriculum—or sometimes its absence—vis-à-vis educational ideas and strategies. Nevertheless, the focus on the relationship between school architecture and the particular purpose of schooling at a given time and in a given place has been



Fig. 1 Two Danish schools, part of the same landscape of schooling in Copenhagen (2021) Top (**Fig. 1**): Guldberg Skole inaugurated 1887. Designed by architect Hans J. Holm. Renovations and expansions have been made several times. Photographer: Henning Sørensen. Source: Hovedstaden.dk.



Fig. 2 Two Danish schools, part of the same landscape of schooling in Copenhagen (2021) Bottom (**Fig. 2**): Copenhagen International School Nordhavn inaugurated 2017. Designed by C.F. Møller Architect. Photographer: Adam Mørk.

a recurrent theme for educationalists, architects and politicians. In 1849, the American educationalist Henry Barnard wrote in his famous work *“School Architecture. Or, Contributions to the Improvement of School-Houses in the United States”*:

The style of the exterior should exhibit good, architectural proportion, and be calculated to inspire children and the community generally with respect for the object to which it is devoted. (...) Every school-house should be a temple, consecrated in prayer to the physical, intellectual, and moral culture of every child in the community, and be associated in every heart with the earliest and strongest impressions of truth, justice, patriotism, and religion.
Barnard (1848: 40).

Barnard writes about the school building in terms embedded in the discourses of his contemporaries, spun into a religious set of norms and significant to the development of schooling in the late 19th century. When evaluating if specific schools were “conducted as to promote habits of neatness and order, and cultivate good manners and refined feelings” (Barnard, 1848), Barnard reflected complementary discourses of hygiene and discipline that were significant in the development of not only schooling but also of school architecture in the late 19th century in the US but also in Europe (Dudek, 2002). Meanwhile, Barnard creates an explicit link between the purpose and the designed materiality of schooling (Willis, 2017). Likewise when expressing his concern for the state of his contemporaneous schoolhouses—their lack of ventilation, privies and, more generally, suitable facilities for schooling—he directly accentuates the importance of a proper schoolhouse for “the bodies, the minds and the hearts of the young and rising generations” (Barnard, 1848: 37).

The research literature on educational architecture has studied how shifting architectural, political and pedagogical ideas have helped shape the design of school buildings (Burke and Grosvenor, 2008; Darian-Smith and Willis, 2017). Analyzing school architecture based on its material surface, radical alterations in the architectural styles of school buildings stand out and may be patterned according to time and geographical space (Willis, 2017). As such, school architecture may be read as a materialization of educational—and architectural—discourses of schooling and thus of different historically situated ideas of schooling, its purpose and form (see Fig. 3).

Classroom curricula: forms and variations

School architecture may be read as a materialization of the understanding of—and disputes over—questions of curriculum, often in a broader sense of knowledge and pedagogy concerning what is considered important for the child and future citizen to learn and, not least, how this is best accomplished. One of the key elements in school architecture over the years has been the classroom. Looking at the classroom, its variations and transformations, its appraisals and critiques, enables an exploration of the interweaving of curriculum and school architecture.

John and Evelyn Dewey’s critique of contemporary schooling of the Western world in the early 20th century includes a description of the classroom—or schoolroom as they call it—depicting “rows of children sitting still at desks and listening to the teacher, speaking only when they are spoken to” (Dewey and Dewey, 1915). Thus, it is a school characterized by a certain disciplinary



Fig. 3 In 2016 the renovation of Engstrandskolen in the Danish city of Hvidovre included the mounting of gymnastics rings and climbing walls in the school corridor. This was in line with the latest (2014) school reform’s decisions on schools having to integrate at least 45 min of exercise into the ordinary school day for all students. Source: Unisport Scandinavia ApS.



Fig. 4 Classroom with large windows, desks in rows and light coming from the left to avoid shadows on the paper when writing with the right hand. The picture is taken in 1931 at Ingerslev Boulevard School in the city of Aarhus, Denmark. The school was inaugurated 1901. Source: Den Gamle By Photo Collection. Photographer: Thomas Pedersen.

instructional mode that is evinced by the physical ordering of tables and children. The description of the classroom also echoes how Danish classrooms at that time reflected and reinforced the dominant educational ideas. With the increasing number of children living in cities in the latter part of the 19th century, more schools were gradually being built. These schools were often public schools; they were larger, with several classrooms and the need for more teachers, who no longer lived on school premises (Gjerløff et al., 2014). For the most part, the schools were spread across several floors, with square classrooms symmetrically located around landings or long corridors. With the focus at the time on the importance of hygiene, sunlight and fresh air, classrooms contained large windows and the students' two-person desks were often available in three sizes to match the child's height (see Fig. 4). The ideal was that the child sat with a straight back with elbows in to the body. The light should preferably come in from the left as students were expected to write with their right hand. The teacher had an elevated desk facing the students (Coninck-Smith, 2011). Thus, teachers, with their gaze, were able to maintain order and induce students' self-discipline. Both were essential to the organized class teaching promoted by the Danish educator Ernst Kaper that gained ground in Danish schools in the early 20th century (Kaper, 1903).

The above image (Fig. 4) represents what is perhaps considered the archetypal classroom: symmetrically opposed square spaces; students seated in pairs at rows of tables, backs straight and facing the teacher's elevated desk. Nevertheless, as we shall see in the following, the form and function of the classroom gradually transformed during the 20th century through its entanglement with shifts in educational policy and the emergence of new educational rationales. Ernst Kaper's philosophy was itself a critique of the earlier interrogative classroom praxis, where the teacher selected and then questioned individual students (Knudsen, 2015). Kaper recommended that the teacher asked the whole class a question and then paused. The students able to answer should be allowed to raise their hand, after which the teacher could select a student to answer the question (Kaper, 1903). Thus, encouraging all students to be involved at all times through organized class teaching, Kaper's ideas could be seen as representing a subtle transformation of earlier uses of the classroom.

The rearrangements of the classroom were tied to the rethinking of teaching strategies and changing view of the child. In 1957, the Danish Ministry of Education's Building Research Committee published a report in collaboration with the National Building Research Institute. The aim of this report was "to explore how the classroom could best be utilized, and how a change in the room's organization and layout could support the new forms of teaching that had emerged with emphasis on both group teaching and individual teaching" (Undervisningsministeriet, 1957). The report gave examples of how grouped table arrangements, rather than tables arranged in rows, were conducive to more individualized and differentiated teaching. The tables could also be set up in a circle or horseshoe shape. In general, it was encouraged to change the physical arrangement of the classroom to suit the teaching about to take place. Of course, this required non-fixed tables and chairs instead of the old-fashioned desks, where the bench and table were attached. Furthermore, the new ways of using the classroom meant that larger rooms were preferable. While the classroom proposed in the report was still a room with four walls, the message was that teaching should look beyond these walls, using corridors and other spaces during lessons. The proposals in the report resonated strongly with the pedagogical experiments that had taken place within the framework of progressive pedagogy as early as the 1920s. In the Danish "Vanløse experiments" carried out 1924–28, there had been a focus on the individual child's developmental opportunities and creativity, the fixed sloping desks had been replaced by tables and chairs and students were allowed to move around freely in the classroom (Christensen, 2008; Gjerløff

et al., 2014). With the Ministry's report, these ideas were tied to the material arrangement of the classroom and written into the broader and more general guidelines and regulations for the school.

Open-plan schools were a more radical break with the classroom model in both architectural and educational terms. The first open-plan schools were developed in the US and in Great Britain in the 1950s and 1960s, before spreading to a number of countries in the following decades (Bennett and Hyland, 1979). Open-plan schools were built with large open spaces and have been described as "the crystallization of progressive ideals in both architecture and education" in the sense of being concerned with the child's individuality and curiosity and striving for social progress by extending teaching beyond traditional knowledge disciplines (Logan, 2017). With flexibility as the key principle in the organization of space, these schools had fewer walls and doors, movable wall dividers, mobile blackboards and easy-to-move furniture. The furnishings were usually less institutional and homogeneous and spaces were decorated with exhibitions of work produced by the students and in the classroom (Burke, 2019; Darian-Smith and Willis, 2017). Thus, the architecture of open-plan schools reflected a shift in the material conceptualization of schools, further evolving a so-called open-plan pedagogy took form in its multiple variations (Franklin, 2012). In general, open-plan school buildings did not come with a ready-made curriculum, but with an openness to new ideas and a call for teachers to experiment with curriculum design. For countries like Great Britain, where there was no national curriculum at the time, and Denmark, where bottom-up processes and teachers' freedom to choose teaching methods were highly valued, schools and teachers had genuine autonomy and space for curricular creativity (Krogh-Jespersen and Nørgaard, 2011). Nevertheless, despite significant variations in the layout and curriculum of open-plan schools, they were generally accompanied by calls for more democratic and socially transformative schooling and by ambitions of establishing less hierarchical relations between teachers and students, more participatory learning and a student-led curriculum (McLeod and Rosén Rasmussen, 2021; Silberman, 1973).

Despite the challenge represented by new approaches such as open-plan schools, the classroom remained a key element in Danish schools, and globally, throughout the 20th century. However, there was a re-emergence of a radical critique of the conventional classroom in the new millennium, where it was argued that there was a need for innovative learning spaces (Dumont and Istance, 2010; Benade, 2017). In opposition to the standard classroom, innovative learning environments have been characterized as multimodal, technology infused and flexible learning spaces, engendering pedagogies for students to become self-directed learners capable of navigating the complexities of a technology-mediated and knowledge based society (Byers et al., 2018). Whereas local curriculum development was embedded in the pedagogy of the open-plan school, it has been argued that innovative learning environments can help reimagine the school and support students' development of the skills and competences anticipated as necessary for 21st century societies (Cardellino and Woolner, 2020; Dumont and Istance, 2010; Sigurðardóttir and Hjartarson, 2011; Carvalho and Yeoman, 2018).

Critical examination of the classroom as a key feature of school architecture and a core element in teaching practices shows variations in how it has been conceived. Moreover, it exemplifies the shifting but ever-present relation between, on one hand, physical buildings and layout, and on the other hand, pedagogical ideas and practices—a relation within which architecture and curriculum are mutually entangled.

The relationship between school architecture and educational practice

The topic of school architecture and curriculum brings up another central question already raised in the text above and further developed in the examples from Barnard's writing, the shifting design and conceptualization of the classroom and the open-plan movement: What is the relationship between school architecture and educational practice? The relationship between the school's material properties, its educational purpose and educational practice has been a matter of ongoing discussion among educational practitioners and in the field of educational research. Empirical understandings of the entanglements of spaces in new school buildings and teachers and students' occupation of these spaces have ranged from viewing architecture as an expression of the (intended) practice of schooling to seeing it as supporting, shaping or even determining pedagogy and possible teaching strategies.

There has long been a belief in the impact of the physical design of school space on educational practice (see Fig. 5) in empirical approaches to school architecture and school space among politicians, educational thinkers, architects and other practitioners (Cooper, 1981; Wood, 2020). In a historical perspective, this belief was already present between the lines in Barnard's writings on the ability of physical surroundings to *promote* certain habits, manners and feelings. In the open-plan schools, the physical environment was to a large extent addressed as supporting the chosen educational activities and strategies (Rosén Rasmussen, 2021). In the sense of balancing between providing structure and encouraging exploration, the school environment has been designated "the third teacher" (Strong-Wilson and Ellis, 2007; Burke and Grosvenor, 2008). This understanding of space as an active agent, rooted in the philosophy of the Italian educational project Reggio Emilia, can be found in contemporary understandings of educational architecture (Pezzetti, 2020).

Belief in the transformative force of school buildings has been crucial in political approaches to the matter. Both UNESCO and OECD, as large international institutions, have had programs for developing the quality of school buildings worldwide that also strived for educational change and improvement (Grosvenor and Rasmussen, 2018b; Charteris et al., 2017; Townshend, 1998). The aim of the British school-building program Building Schools for the Future, which ran from 2005 to 2020, was framed in terms of educational and social change. While based on national guidelines, schools were asked to tailor their curriculum to the needs of their local community and the school's particular educational aims, before the material spaces of the new school buildings were designed (Krafft, 2012). Around the same time but on a smaller scale, the Danish municipality of Gentofte ran a project to renovate school buildings that had a similar goal of facilitating a transformation of the form and content of education (Leiringer and Cardellino, 2011). Some have argued that a utopian ideal of a bottom-up transformation of schooling and the curriculum lay at the



Fig. 5 A rendering showing how architects have interpreted, translated, and developed political and educational visions for a new school in Sundby, Denmark, to be built in 2022. The school is by the architects named a landscape school and is said to have a potential of drawing history and nature inside as well as encouraging spontaneous and integrated movement. Source: Henning Larsen Architects.

core of the design and practice of open-plan schools in the 1960s and 1970s (McLeod and Rosén Rasmussen, 2021). Meanwhile, the innovative learning spaces of the present day are designed as a contrast to the traditional classroom-based school (see Fig. 5) and are expected to support or even enhance strategies suitable for teaching the curriculum of the future (Deed and Lesko, 2015; Carvalho and Yeoman, 2018).

The academic analysis of school architecture and its relationship to curriculum and broader educational matters has been molded by the thinking of Michel Foucault and his notions of panopticon and governmentality (Piro, 2008; Simmons, 2010; Peim, 2005; James, 1991). Along these lines, the role of governance in educational architecture has been highlighted and school architecture has been interpreted as imprints of policies and pedagogies believed to have the potential to govern and shape the school and its inhabitants (Burke, 2005; Grosvenor and Rasmussen, 2018b; Wood, 2020). The conceptualization of architecture as means of governance grasps the perception of school buildings as supportive of curricular aims. It also pinpoints the belief in the buildings' ability to promote educational ideas and strategies recognizable in Barnard's writings and extrapolated in later visions of the physical environment as a key to broader transformations of schooling, the curriculum and the school system as a whole.

Within academia, there has been a move toward an understanding of space as social and relational, acting as a critique of the idea of space as a container (Massey, 2005). This echoes and reintroduces Lefebvre's focus on the social production of space. In these studies, the power of school architecture is further complicated as users or inhabitants of school buildings, and in a broader sense educational practices, are not merely seen as governed by the physical architecture of these buildings, but also as affecting and shaping it in both material and immaterial ways. Such understandings of space as social and processual further complicate the relationship between architecture and curriculum.

A socio-material approach developed with an outset in cultural geography and actor-network theory has reintroduced and reinforced a focus on the material aspects of learning environments by stressing the agency of the material, but still maintaining a social and relational understanding of space and even introducing it to analyses of the affective dimension of learning environments (Mulcahy, 2015; Mulcahy et al., 2015; Mulcahy and Morrison, 2017; McGregor, 2003, 2004). In these studies, there is particular focus on educational change and how this may be induced by a new learning environment. Nevertheless, the socio-material approach leads to an understanding of space, pedagogy and change as mutually constitutive. Recent influence from new materialist theories has shifted theoretical understandings of school architecture even further from the idea of space as a container and thus of school buildings as an external support for or promoter of an easily decipherable curriculum. In the new materialist approach to school architecture, school designs, pedagogical ideas and practices—and one might add curriculum in a narrow sense—are intrinsically linked and explained as coming into being and obtaining their specific meanings *through* each other. The focus is on process and entanglements, weaving both material and immaterial meaning-making of school architecture and design, teaching practices, pedagogical understandings and curricular provisions closely together over time, across space and in situ. Studies applying this approach are occupied with the detailed processes of becoming, for instance exploring the protracted construction and continuous adjustment of a school building and of everyday teaching practices. At the same time, these studies all pay attention to how this process of becoming is tied up in, draws on and feeds into broader local, national and transnational discursive processes, which include but also go far beyond political decision-making (Juelskjær, 2014; Coninck-Smith, 2017; Thyssen, 2018; Carvalho and Yeoman, 2018).

The theoretical conceptualization of space and materiality as something that is shaped by, shapes or is entangled in educational discourses and practices grasps to varying degrees the empirical and practical understandings of and approaches to school buildings

and school spaces among architects, school leaders and teachers. Moreover, the development of theoretical understandings of the physical (school) space and its implications for comprehending the relationship between school architecture and curricula has influenced the design of and the political and practical approach to school architecture and school space in general. Thus, acknowledging how theoretical and empirical understandings of space interweave is important coming to terms with changes in school design and, not least, the expectations to and the conduct of school architecture and school space.

Material transfers of curricular designs

The history of the relationship between material school design and curriculum has a strong international and transnational dimension. Because of the physicality of its matter, it is easy to trace and identify how architecture and material designs travel across national borders. Studies of school architecture in an international or transnational perspective most often single out how pedagogical ideas and architectural models travel together (Coninck-Smith, 2010a; Thyssen, 2018; Lundahl, 2016). Regardless of the subsequent processes of translation and implementation, the spread of material school designs often seems to have functioned as a spearhead for the introduction of new educational ideas and practices (see Figs. 6 and 7).

The entangled travel of school architecture and curriculum has assumed various forms over time. World's fairs have been put forward as an example of an international meeting point for the materiality and curriculum of schooling where particular school designs and new teaching methods have been promoted as coherent models. In the 1860s and 1870s, various educational artifacts, such as scientific instruments, charts, desks and tables, were displayed at the world's fairs. In 1871, the Swedish state funded the exhibition of a full-scale and fully equipped schoolhouse at the World's Fair in London. This was a prototype for the ideal schoolhouse of the future rather than an example of an actual Swedish schoolhouse, but nevertheless, it was an expression of the Swedish education system, including the use of object teaching and more specific national pedagogical ideas such as Sloyd (Lundahl, 2016; Lundahl and Lawn, 2015). The exhibition of the Montessori demonstration classroom at the Panama-Pacific International Exposition in San Francisco, California in 1915 provides a similar example. This time, the classroom was populated by children and visitors could watch classroom activities through glass walls. Based on Maria Montessori's ideas, the classroom was equipped with child-sized furniture and the focus was on what has been called a pedagogy of attention. The Montessori demonstration classroom has been described as important in spreading Montessori's pedagogical ideas in the US (Sobe, 2018).

Through support programs and numerous recommendations for and evaluations of school buildings around the world, UNESCO and OECD have both been influential in the transnational movement and entanglement of school design and school curriculum since the 1950s and 1960s. For example, the OECD Program on Educational Building (PEB) began as a forum for discussion and analysis of educational buildings in the early 1970s and laid the foundations for a network of responsible officials in the national ministries of education, led by the OECD. Recently, the OECD has actively promoted the transition to innovative learning environments as a way to align school architecture with the curricular framework for 21st century learning (OECD, 2017).

Not surprisingly, architects have played a profoundly important role in the travel of physical school design. The British architect and member of the London School Board 1871–1889 Edward Robert Robson's famous book *School Architecture* (1874) includes long chapters on his visits to the US, Switzerland, Germany, France, Belgium and Holland. Robson had been traveling "in search of the best schools" and he does not only engage with different architectural design features and building materials. Through a discussion of different teaching methods and their corresponding physical layouts that he encountered in America and continental Europe, he developed design recommendations for the new schools to be built in London (Robson, 1972 (1874)). As an example, the book includes advice regarding the size, layout and not least lighting of classrooms. This advice reflects new approaches to instruction, such as the benefits of strengthening the possibility of surveillance by increasing "the teacher's angle of vision (...) far beyond that of 45 degrees" (Robson, 1972 (1874)). The way in which a small Danish architectural firm became involved in developing modular school buildings in the 1950s offers another example of the great importance of architects' study tours and exchange of ideas for the diffusion and materialization of new educational ideas throughout the 20th century. Contact with the British architect couple Mary Crowley and David Medd had a particular influence on the small company's state-financed design of new learning spaces for the independent and self-reliant student (Coninck-Smith, 2010a,b).

As already mentioned, transnational connections and transfers played an important role in the spread and development of open-plan schooling. In the 1960s and 1970s, architectural designs of US and British open-plan schools traveled across national borders through the work of architects and influential educationalists (Bennet and Hylland, 1979). These architectural models inspired architects and educational policymakers, but they were also often modified in response to national and local policies and educational systems (e.g., Franklin, 2012). The pedagogical ideas tied to open-plan schooling traveled in tandem with the designs themselves, but most often the use and development of open-plan schooling also involved the travel of books and educationalists through study trips and transnational networks involving teachers and school leaders as well as school developers and politicians (McLeod and Rosén Rasmussen, 2021). Thus, the travel of physical design models is not only connected to and brings with it simplified aspects of pedagogy and curriculum; it involves a range of compound ideas, processes and material matters situated within, but also transcending, local and national borders. These ideas, processes and matters entangle and potentially transform the physical design, as well as the associated pedagogical understandings, perhaps also entangling with different regulations and curricular ideas.



Fig. 6 Hellerup stairs is known from Hellerup Skole in Denmark built in the early 2000s. Architects: Arkitema. Photograph by Kontraframe.



Fig. 7 Architectural elements and curricular ideas travel together across borders, entangle and transform in new settings: The Hellerup stairs has become a core element in the construction of hub schools in Victoria, Australia, in the late 2010s. The photo is from South Melbourne Primary School. Architects: Hayball. Photographer: Dianna Snape.

Building curriculum and future models of schooling

Significant changes in curriculum, pedagogy and technology have been seen to change the order of the classroom; however, such change has not always been considered successful (Lawn and Grosvenor, 2021). On the contrary, a vast number of studies have pointed out how such intended change is often not realized in the physical learning environment built to incite and enhance curricular transformations. These studies have demonstrated how the absence of change—often accompanied by explicit frustration among school staff—entails “teaching as usual,” despite a new physical environment—sometimes even leading to refurbishments and rebuilds more in line with earlier teaching practices and educational beliefs (Bennett and Hyland, 1979; Cardellino and Woolner, 2020; Gislason, 2015). In some studies, the failure to bring about change is explained by unforeseen aspects of the physical architecture. One example of this is noise problems in some of the open-plan schools in the 1960s and 1970s, with similar problems afflicting some of the open innovative learning spaces built since the turn of the millennium (Shield et al., 2010; Cardellino and Woolner, 2020; Mulcahy and Morrison, 2017). In other cases, teachers’ motivation or their lack of training in and for the new environment have been given as explanations for their reversion to more familiar teaching strategies (Niemi, 2021; Brogden, 2007; Higgins et al., 2005). Recently, interrupted and unclear leadership has been suggested as an obstacle preventing the pedagogical transformation expected to result from profound alterations to school architecture (Daniels, 2019). A wave of so-called post-occupational studies have dived further into the question of the links between physical design and education, exploring what happens when buildings that are expected to inspire and enhance new pedagogical approaches are put into operation (Alterator and Deed, 2018). The answers are manifold but they all engage with the complexity of the issue; for instance, Neil Gislason’s model of how organizational matters, student dynamics and educational culture all help determine the use of a school’s (new) physical design (Gislason, 2018). With the lack of unequivocal answers, more or less all such studies seem to support the argument that there is no causal link between changes to physical learning spaces and pedagogical change (Mulcahy et al., 2015). Despite the missing or unintended effects of new school architecture, the ambition of transforming schooling and learning through educational technology and innovative learning environments has intensified in recent decades (Cardellino and Woolner, 2020; OECD, 2013, 2017; Dovey and Fisher, 2014; Woolner et al., 2018). This is accompanied by the development of strategies for aligning the educational intentions of physical changes with actual teaching practices through collaborative processes (Fig. 8) (Carvalho and Yeoman, 2018; Bøjer, 2019).

As a potential tool of change, the coming together of school architecture and curriculum has generally been future-oriented, sometimes even turning its back on the past. With the persistence of the utopian element of school architecture as a jumping-off point, Keri Facer has addressed the ethical implications of building schools for an idealized and politicized future (Facer, 2018). Facer recognizes how a school building is a “jetty into the future” and how this both concerns the educational vision involved in the design process and the fact that the building will probably survive “into a time that is not now.” Moreover, the future is seen as at the heart of educational discourses among politicians, academics and practitioners in the sense that education is seen as a way to shape, prepare or defend students for or against a projected future. However, Facer also emphasizes the temporality of education itself—how educational processes continuously generate new ideas about the future and the important role played



Fig. 8 School children engaged in collaborative workshops developing ideas and material elements for their (future) school spaces. Workshops held by Rune Fjord Studio. Photographer: Bodil Hovaldt Bøjer.

by dialog between the past and the present. With this in mind, Facer suggests designing schools as laboratories that allow teachers and students to play with the future.

Regardless of the complex relationship between physical design and curriculum in its broadest sense, the importance of the physical learning environment on educational processes is underlined both empirically and theoretically. The current paper's reflections on shifting empirical and theoretical understandings of the relationship between school architecture and curriculum may be helpful in maintaining a more holistic and caring approach to designing school(s), literally and figuratively.

References

- Alterator, S., Deed, C., 2018. *School Space and Its Occupation: Conceptualising and Evaluating Innovative Learning Environments*. Brill Sense, Boston.
- Barnard, H., 1848. *School Architecture: Or, Contributions to the Improvement of School-Houses in the United States*. A. S. Barnes & Company.
- Benade, L., 2017. Is the classroom obsolete in the twenty-first century? *Educ. Philos. Theor.* 49, 796–807.
- Bennett, N., Hyland, T., 1979. Open plan—open education? *Br. Educ. Res. J.* 5, 159–166.
- Bøjer, B.H., 2019. *Unlocking Learning Spaces: An Examination of the Interplay Between the Design of Learning Spaces and Pedagogical Practices*. Ph.D. Institute of Visual DesignThe Royal Danish Academy of Fine Arts, Schools of Architecture, Design and Conservation.
- Boys, J., 2011. *Towards Creative Learning Spaces: Re-thinking the Architecture of Post-compulsory Education*. Routledge, London.
- Brogden, M., 2007. Plowden and primary school buildings: a story of innovation without change. *Forum For Promot. 3-19 Compr. Educ.* 49, 55.
- Burke, C., Grosvenor, I., 2008. *School. Reaktion*, London.
- Burke, C., 2005. Containing the school child: architectures and pedagogies. *Paedagog. Hist.* 41, 489.
- Burke, C., 2019. Designing for “touch”, “reach” and “movement” in postwar (1946–1972) English primary and infant school environments. *Senses Soc.* 14, 207–220.
- Byers, T., Imms, W., Hartnell-Young, E., 2018. Evaluating teacher and student spatial transition from a traditional classroom to an innovative learning environment. *Stud. Educ. Eval.* 58, 156–166.
- Cardellino, P., Woolner, P., 2020. Designing for transformation—a case study of open learning spaces and educational change. *Pedagog. Cult. Soc.* 28, 383–402.
- Carvalho, L., Yeoman, P., 2018. Framing learning entanglement in innovative learning spaces: connecting theory, design and practice. *Br. Educ. Res. J.* 44, 1120–1137.
- Charteris, J., Smardon, D., Nelson, E., 2017. Innovative learning environments and new materialism: a conjunctural analysis of pedagogic spaces. *Educ. Philos. Theor.* 49, 808–821.
- Christensen, G., 2008. Individ og disciplinering: det pædagogiske subjekts historie. *Samfundslitteratur*, Frederiksberg.
- Coninck-Smith, N.D., 2010a. Danish and British architects at work: a micro-study of architectural encounters after the Second World War. *Hist. Educ.* 39, 713–730.
- Coninck-Smith, N.D., 2010b. Experts at work: a micro-study of architects and school buildings in Denmark, 1940–1970. In: Lundqvist, A., Petersen, K. (Eds.), *In Experts We Trust: Knowledge, Politics and Bureaucracy in Nordic Welfare States*. 1. Oplag. University Press of Southern Denmark, Odense.
- Coninck-Smith, N.D., 2011. Barndom og arkitektur: rum til danske børn igennem 300 år. *Klim, Århus*.
- Coninck-Smith, N.D., 2017. Making schools and thinking through materialities: Denmark 1890–1960. In: Darian-Smith, K., Willis, J. (Eds.), *Designing Schools: Space, Place and Pedagogy*. Routledge, London.
- Cooper, I., 1981. The Politics of Education and Architectural Design: the instructive example of British primary education. *Br. Educ. Res. J.* 7, 125–136.
- Daniels, H., 2019. *School Design Matters: How School Environments Relate to the Practice and Experience of Teaching and Learning*. Routledge, Taylor & Francis Group, London.
- Darian-Smith, K., Willis, J., 2017. *Designing Schools: Space, Place and Pedagogy*. Routledge.
- Deed, C., Lesko, T., 2015. “Unwalling” the classroom: teacher reaction and adaptation. *Learn. Environ. Res.* 18, 217–231.
- Dewey, J., Dewey, E., 1915. *Schools of To-Morrow*. J. M. Dent, London.
- Dovey, K., Fisher, K., 2014. Designing for adaptation: the school as socio-spatial assemblage. *J. Architect.* 19, 43–63.
- Dudek, M., 2002. *Architecture of Schools: The New Learning Environments*. Architectural, Oxford.
- Dumont, H., Istance, D., 2010. Analysing and designing learning environments for the 21st century. In: Dumont, H., Istance, D. (Eds.), *The Nature of Learning. Using Research to Inspire Practice. Organisation for Economic Cooperation and Development*.
- Facer, K., 2018. Governing education through the future. In: Grosvenor, I., Rosén Rasmussen, L. (Eds.), *Making Education: Material School Design and Educational Governance*. Springer International Publishing, Switzerland.
- Franklin, G., 2012. “Built-in variety”: David and Mary Medd and the child-centred primary school, 1944–80. *Architect. Hist.* 55, 321–367.
- Gislason, N., 2015. The open plan high school: educational motivations and challenges. In: Woolner, P. (Ed.), *School Design Together*. Routledge, Oxfordshire, England.
- Gislason, N., 2018. The whole school: planning and evaluating innovative middle and secondary schools. In: Alterator, S., Deed, C. (Eds.), *School Space and Its Occupation: Conceptualising and Evaluating Innovative Learning Environments*. Brill Sense, Boston.
- Gjerløff, A.K., Jacobsen, A.F., Nørgaard, E., Ydesen, C., 2014. *Da skolen blev sin egen: Dansk skolehistorie 1920–1970*. Aarhus Universitetsforlag.
- Grosvenor, I., Rasmussen, L.R. (Eds.), 2018a. *Making Education. Material School Design and Educational Governance*. Springer International Publishing AG, Switzerland.
- Grosvenor, I., Rasmussen, L.R., 2018b. Making education: governance by design. In: Grosvenor, I., Rasmussen, L.R. (Eds.), *Making Education. Material School Design and Educational Governance*. Springer International Publishing AG, Switzerland.
- Higgins, S., Hall, E., Wall, K., Woolner, P., Mccaughey, C., 2005. *The Impact of School Environments: A Literature Review*. Design Council, London.
- James, R., 1991. Observing and normalizing: Foucault, discipline, and inequality in schooling: big brother is watching you. *J. Educ. Thought* 25, 104–119.
- Juelskjær, M., Rosén Rasmussen, L., 2019. Klasserummets arkitektur. In: Johansen, M.B. (Ed.), *10 døre til klasserummet*. Akademisk Forlag, København.
- Juelskjær, M., 2014. Changing the organization: architecture and stories as material-discursive practices of producing “schools for the future”. *Tamara J. Crit. Organ. Inq.* 12, 25.
- Kaper, E., 1903. *Den daglige Undervisnings Form : Kritik og Forslag: 3 didaktiske Afhandlinger ud fra Praxis*. København.
- Knudsen, H., 2015. De blinde pletter i synlig læring: Hvordan evidens blev svaret på det hele og glemte de andre spørgsmål, man også kan stille. In: Bjerg, H., Vaaben, N. (Eds.), *At lede efter læring*. Samfundslitteratur, Frederiksberg.
- Kraftl, P., 2012. Utopian promise or burdensome responsibility? A critical analysis of the UK government's building schools for the future policy. *Antipode* 44, 847–870.
- Krogh-Jespersen, K., Nørgaard, E., 2011. Udviklingsarbejde og skoleudvikling. In: Kristensen, H.J., Fibæk Laursen, P. (Eds.), *Gyldendals pædagogikhåndbog: otte tilgange til pædagogik*. 1. udgave. Gyldendal, København.
- Lawn, M., Grosvenor, I., 2021. Finding traces, researching silences, analysing materiality: notes from the United Kingdom. *Educació i història: Revista d'història de l'educació* 38, 33–54.
- Leiringer, R., Cardellino, P., 2011. Schools for the twenty-first century: school design and educational transformation. *Br. Educ. Res. J.* 37, 915–934.
- Logan, C., 2017. Open shut them. In: Darian-Smith, K., Willis, J. (Eds.), *Designing Schools: Space, Place and Pedagogy*. Routledge, Taylor & Francis Group, London, New York, NY.
- Lundahl, C., Lawn, M., 2015. The Swedish schoolhouse: a case study in transnational influences in education at the 1870s world fairs. *Paedagog. Hist.* 51, 319–334.
- Lundahl, C., 2016. Swedish education exhibitions and aesthetic governing at World's fairs in the late nineteenth century. *Nordic J. Educ. Hist.* 3, 3–30.
- Massey, D., 2005. *For Space*. SAGE, London.

- McGregor, J., 2003. Making Spaces: teacher workplace topologies. *Pedagog. Cult. Soc.* 11, 353–377.
- McGregor, J., 2004. Spatiality and the place of the material in schools. *Pedagog. Cult. Soc.* 12, 347–372.
- McLeod, J., Rosén Rasmussen, L., 2021. Open-plan schooling and everyday utopias: Australia and Denmark in the 1970s. *Oxf. Rev. Educ.* 47, 659–680.
- Mulcahy, D., Morrison, C., 2017. Re/assembling 'innovative' learning environments: affective practice and its politics. *Educ. Philos. Theor.* 49, 749–758.
- Mulcahy, D., Cleveland, B., Aberton, H., 2015. Learning spaces and pedagogic change: envisioned, enacted and experienced. *Pedagog. Cult. Soc.* 23, 575–595.
- Mulcahy, D., 2015. Re/assembling spaces of learning in Victorian government schools: policy enactments, pedagogic encounters and micropolitics. *Discourse* 36, 500–514.
- Niemi, K., 2021. "The best guess for the future?" Teachers' adaptation to open and flexible learning environments in Finland. *Educ. Inq.* 12, 282–300.
- OECD, 2013. *Innovative Learning Environments*.
- OECD, 2017. *The OECD Handbook for Innovative Learning Environments*. OECD, Paris.
- Peim, N., 2005. Towards a social ecology of the modern school: reflections on histories of the governmental environment of schooling. *Paedagog. Hist.* 41, 627–639.
- Pezzetti, L.A., 2020. Space-places and third teacher: the issue of architectural space in the age of knowledge cities and schools 3.0. In: Della Torre, S., Bocciarelli, M., Daglio, L., Neri, R. (Eds.), *Buildings for Education: A Multidisciplinary Overview of the Design of School Buildings*. Springer International Publishing, Cham.
- Piro, J.M., 2008. Foucault and the architecture of surveillance: creating regimes of power in schools, shrines, and society. *Educ. Stud.* 44, 30–46.
- Robson, E.R., 1972 [1874]. *School Architecture*. With an Introduction by Malcolm Seaborne. Leicester University Press, New York. Humanities Press.
- Rosén Rasmussen, L., 2021. Building Pedagogies. A historical study of teachers' spatial work in new school architecture. *Educ. Inq.* 12, 225–248.
- Shield, B., Greenland, E., Dockrell, J., 2010. Noise in open plan classrooms in primary schools: a review. *Noise Health* 12, 225–234.
- Sigurðardóttir, A.K., Hjartarson, T., 2011. School buildings for the 21st century. Some features of new school buildings in Iceland. *CEPS J.* 1, 25–43.
- Silberman, C.E., 1973. *The Open Classroom Reader*. Vintage Books, New York.
- Simmons, L., 2010. The Docile body in school space. In: Monahan, T., Torres, R., Hirschfield, P., Simmons, L., Steeves, V., Apple, M., Bracy, N.L., Casella, R., Gilliom, J., Hope, A. (Eds.), *Schools Under Surveillance: Cultures of Control in Public Education*. Rutgers University Press, New Brunswick, NJ.
- Sobe, N., 2018. Boredom and classroom design: the affective economies of school engagement. In: Grosvenor, I., Rosén Rasmussen, L. (Eds.), *Making Education : Material School Design and Educational Governance*. Springer International Publishers, Switzerland.
- Strong-Wilson, T., Ellis, J., 2007. Children and place: Reggio Emilia's environment as third teacher. *Theory Into Pract.* 46, 40–47.
- Thyssen, G., 2018. Boundlessly entangled: non-/human performances of education for health through open-air schools. *Paedagog. Hist.* 54, 659–676.
- Townshend, J., 1998. *Under One Roof the Integration of Schools and Community Services in OECD Countries*. OECD Publishing, Paris.
- Tröhler, D., 2016. Curriculum history or the educational construction of Europe in the long nineteenth century. *Eur. Educ. Res. J.* 15, 279–297.
- Undervisningsministeriet, 1957. *Klasserummet: Funktion og udformning. Nyt skolebyggeri 5*. Teknisk Forlag, København.
- Willis, J., 2017. Architecture and the school in the twentieth century. In: Darian-Smith, K., Willis, J. (Eds.), *Designing Schools. Space, Place and Pedagogy*. Routledge.
- Wood, A., 2020. Built policy: school-building and architecture as policy instrument. *J. Educ. Pol.* 35, 465–484.
- Woolner, P., Thomas, U., Tiplady, L., 2018. Structural change from physical foundations: the role of the environment in enacting school change. *J. Educ. Change* 19, 223–242.

Students and staff co-creating curriculum in higher education

Catherine Bovill, Institute for Academic Development, University of Edinburgh, Edinburgh, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	235
What is co-created curriculum in higher education?	235
What is the higher education curriculum?	235
What is co-creation?	236
What is co-created curriculum in higher education?	236
Theoretical frameworks and models	237
Examples of practice	239
Co-creation <i>of</i> the curriculum	239
Course design teams at Elon University, North Carolina, USA	239
Co-creating a science skills curriculum at the University of Glasgow, Scotland, UK	239
Co-creating a chemistry course at Uppsala University, Sweden	239
Co-creating an interdisciplinary program, University of Technology, Sydney, Australia	239
Co-creation <i>in</i> the curriculum	240
Shared PowerPoint slides, University College Northern Denmark (UCN), Aalborg, Denmark	240
Co-creating essay titles, University of Vienna, Austria	240
Co-creating courses within a Masters in Education, Universiti Utara Malaysia, Sintok, Malaysia	240
Co-creating a business course, University of St Andrews, Scotland, UK	240
Benefits	240
Challenges	241
Can we afford not to co-create the curriculum?	242
References	242

Introduction

Traditionally, the curriculum is perceived to be the domain of teaching staff (Bourner, 2004); something designed by staff and “done to” students. Curriculum has often been considered synonymous with a syllabus or canon of knowledge, captured in a document, which in some countries is considered a legally binding statement of what an institution guarantees a student will be taught. Therefore, students and staff co-creating the curriculum in higher education is a reconceptualization of the way many people think about curriculum—what it is, who does it, and how it is done. Co-creating the curriculum is not a new idea, but it has attracted growing attention over the last decade (Bovill, 2020a; Cook-Sather et al., 2014; Lubicz-Nawrocka, 2020; Mercer-Mapstone et al., 2017). Within recent literature, co-created curriculum is considered to be a dynamic negotiation of learning and teaching between staff and students (Bovill and Woolmer, 2019; Fraser and Bosanquet, 2006).

In this chapter, I will explore the nature of higher education curriculum and definitions of co-creation, as well as outlining the differences between co-creation *of* the curriculum and co-creation *in* the curriculum. I will present a range of theoretical frameworks and models used to inform thinking about co-created curriculum. In trying to understand co-created curriculum, I present some examples from around the world, which help to illustrate the diversity of what is possible. I will also explore some of the benefits and challenges of co-creation. Finally I ask the question of whether we should be offering all students in higher education the opportunity to co-create the curriculum, in the face of the increasing evidence of the benefits from co-creation processes and outcomes. Before proceeding, I make just one note on the text: I use the term staff in this article to refer to both academic staff and professional services staff working in higher education who have any form of responsibility for teaching, learning, assessment and curriculum design.

What is co-created curriculum in higher education?

In order to explore the nature of co-created curriculum in the higher education setting, we need to ask first what is the higher education curriculum? We also need to ask what is co-creation? And finally, what is co-created curriculum in higher education?

What is the higher education curriculum?

The curriculum is frequently considered synonymous with the syllabus. The syllabus is usually a list of topics and concepts which students should be taught, and students are expected to know by the end of a course or program. Where curriculum is thought of in

a similar way to the syllabus, this is a knowledge-centric, often static version of curriculum that relates to the idea of students learning a disciplinary “canon.” This perspective fails to recognize that the curriculum for many scholars and practitioners is a more dynamic process, which cannot exist as a document on a shelf, or without students as part of that process (Barnett and Coate, 2005; Fraser and Bosanquet, 2006).

The curriculum framework from Barnett and Coate (2005) based on empirical research in the UK, highlights the importance of balancing students’ development of “knowing,” “acting” and “being” if students are to be prepared for a complex future world. Typically, higher education curricula emphasize content knowledge or “knowing.” “Acting” refers to putting into practice approaches that are common in a discipline but often tacit to disciplinary experts, emphasizing the skills needed within a particular discipline. And finally, “being” refers to the development of the student as a person, encompassing the idea of how a student takes a stance in relation to the world—“being-in-the-world,” self-realization and capability. Barnett and Coate (2005) highlighted that different disciplines place different emphasis on knowing, acting and being, but all of these elements are essential in any curriculum. These three building blocks of curriculum align well with work by Entwistle (2003) and McCune and Hounsell (2005), who described in their research, students learning “ways of thinking and practicing” in the subject. For example, students need to learn not only biology knowledge, but importantly, also how to practice as a biologist and scientist. Barnett and Coate (2005) envisage the curriculum as a form of engagement, which involves deliberately designing in opportunities for students to engage in knowing, acting and being.

Australian researchers Fraser and Bosanquet (2006) carried out important research that examined how academic staff define curriculum. Participants in their study defined curriculum in four ways: “A: the structure and content of a unit (subject); B: the structure and content of a programme of study; C: the students’ experience of learning; D: a dynamic and interactive process of teaching and learning” (Fraser and Bosanquet, 2006, p. 272). Categories A and B are similar to one another but differentiated by whether the focus is on a course or program, in other words, by the scale of the curriculum. Category C expands the idea of the curriculum to be more than just structure and content, to include students’ perspectives of learning—the teacher provides a framework within which students begin to negotiate the curriculum. Category D sees the curriculum as involving an integral role for students as collaborators in designing curriculum and co-constructing knowledge alongside academic staff. This final definition provides less framing by the teacher than Category C. Both categories C and D move “... away from curriculum as a product provided for students, to a process that enables student learning” (p. 274).

There are other higher education curriculum models and frameworks such as the constructive alignment model from Biggs (1996), which emphasizes an assessment-led, outcomes-focused model of course and program design. Also there is the taxonomy of significant learning from Fink (2003, 2007) and five approaches to curriculum design from Toohey (1999). However, the frameworks from Barnett and Coate (2005) and from Fraser and Bosanquet (2006) are particularly helpful in explaining why many colleagues have come to recognize the role of students in curriculum processes. We also need to acknowledge the key role of curriculum scholars from schools education, particularly those who have promoted the idea of the curriculum as a process (Beane, 1997; Boomer, 1978; Boomer et al., 1992; Eisner, 1979; Joseph, 2011; Stenhouse, 1975). In conceiving of the curriculum as a process, these authors were also seeing the curriculum as an opportunity for students to be involved in decision making about their education. Arguments for process-based definitions of curriculum are often informed by a recognition that the Latin root of curriculum translates as “a race” and its derivative, “currere” as running the race—suggesting a process (Bron et al., 2022; Pinar, 1975). Indeed Boomer “playfully uses the verb ‘curriculuming’ for this, to imply action and process” (Bron et al., 2022, p 3).

What is co-creation?

Co-creation is “the values-based implementation of staff and students working together in an ongoing, reciprocal, creative, and mutually beneficial process to negotiate and share decision-making regarding aspects of higher education curricula” (Lubicz-Nawrocka, 2020, pp. i–ii). This definition captures a sense that co-creation both differs from, and also shares characteristics with, some similar terms which are often used inter-changeably. It can be helpful to think about the terms student engagement, active learning, co-creation and partnership as broadly “an overlapping continuum of student participation where the level of meaningful participation, shared decision-making, shared responsibility and equality increases as you move from student engagement to partnership” (Bovill, 2020a, p. 31). Finally, Iversen and Stavnskaer Pedersen (2017) remind us that both the “co” and “creation” elements signal to us that this is a social process, in which something new appears from that process.

What is co-created curriculum in higher education?

The idea of co-created curriculum is not new. In the early 1900s, Dewey argued for more democratic school classrooms (Dewey, 1916). In the 1960s, Rogers built upon Dewey’s work by proposing alternative approaches to existing formal schooling. He argued students need the “freedom to learn,” emphasizing the learner as a contributor to the learning process, including students co-creating curriculum (Rogers and Freiberg, 1983). In the 1980s, Giroux continued to build upon these ideas in his foundational work advocating for critical pedagogy (Giroux, 1983). Critical pedagogy promoted the idea that students and staff should negotiate learning and teaching in order to co-create new forms of knowledge from their experiences and in order to challenge existing views of the world—ultimately with the goal of students re-creating a new society (Darder et al., 2003; Giroux, 1983). Boomer and colleagues in Australia drew heavily upon the critical pedagogy literature in their work focused on negotiating the curriculum (Boomer et al., 1992). According to Boomer et al. (1992) negotiation of the curriculum relies upon challenging existing

understandings of power relationships in learning and teaching. This legacy from schools education has been influential in higher education, with increasing attention being paid to co-created curriculum in higher education over the last decade in particular.

Co-created curriculum practices have spread across a range of disciplines in higher education, with some scholars claiming the benefits offer alternatives to the neoliberal and marketized university, while others challenge these claims (Gravett et al., 2020; Holen et al., 2020; Waddington, 2016; Wijaya Mulya, 2019). But other common motivations for co-creation often refer to the wish to enhance students' experiences, to democratize higher education, or to offer a more socially just approach to teaching and learning (Lubicz-Nawrocka, 2016; Owusu-Agyeman and Fourie-Malherbe, 2019). Co-created curriculum can be found in many different disciplines, and while not a comprehensive list, examples can be found in: language teaching (Barrineau et al., 2019; Breen and Littlejohn, 2000); teacher education (Bergmark and Westman, 2016; Mihans et al., 2008; Van Werven et al., 2017); geography (Moore Cherry et al., 2016; Hill et al., 2016); ecology (Huxham et al., 2015); environmental justice and environmental history (Barrineau et al., 2019; Fischer, 2005; Grudens-Schuck, 2003; Wilkinson and Scandrett, 2003); gender studies (Kenny, 2019); mathematics and statistics (Delpish et al., 2010; Duah and Croft, 2012); science (Woolmer et al., 2016); engineering (Bovill, 2020a); chemistry (Bengtson et al., 2017); psychology (Bovill, 2020a); business studies, innovation and entrepreneurship (Cook-Sather et al., 2014); medicine (Barrineau et al., 2019; Harden and Fawcner, 2019); nursing (Fawcett and Watson, 2019); film studies (Bell et al., 2019); and classics (Cook-Sather et al., 2014). In addition, there are resources which bring together co-creation examples, research and practice across a diverse range of subjects (Barrineau et al., 2019; Bovill, 2020a,b; Cook-Sather et al., 2014; Mercer-Mapstone and Abbot, 2020).

Bovill and Woolmer (2019) argue that in higher education, it can be helpful to distinguish between co-creation *of* the curriculum and co-creation *in* the curriculum. Co-creation *of* the curriculum involves co-design of an institutional curriculum or a program or course, usually before the curriculum, program or course takes place, and often involving only a selected group of students. In contrast, co-creation *in* the curriculum involves co-design of learning, teaching and assessment within a curriculum, program or course as it takes place, and often involving all students in the class or cohort (Bovill, 2020b; Bovill and Woolmer, 2019). There has been little debate about the connections between these two forms of co-created curriculum, and indeed, many examples of co-creation tend to be one or the other form of co-creation. It appears uncommon for students to be involved in co-creation *of* a curriculum and then also to get an opportunity to make further decisions about this curriculum as a student in the class, taking part in co-creation *in* that same curriculum. There are some examples, (see for example Mihans et al., 2008), but for students to be involved in both forms of co-created curriculum requires co-creation *of* the curriculum to involve future students who will be studying a course or program. In addition, most examples of co-creation *of* the curriculum involve a selected group of students and therefore, even where there has been an opportunity to co-create the curriculum before a course or program commences, the likelihood is that most students in the class will not have had a chance to co-create the curriculum before arriving in class.

Co-creation *of* the curriculum might be expected to focus on decisions about different things than co-creation *in* the curriculum, but in reality, there are examples of co-creation in the curriculum which include decisions about, for example, learning outcomes, method of assessment, and weighting of assessments. However, rules about what is expected to be pre-set in any curriculum before a course starts will differ from country to country and from institution to institution. Building on one of Fraser and Bosanquet's (2006) definitions of the curriculum, as a dynamic negotiation between teachers and students, I argue that it is important we view the curriculum not as something set in stone, but as Breen and Littlejohn (2000, p. 277) argue, "... teachers may need to come to see their own plans for classroom work as simply proposals ... which learners have the right to reformulate, elaborate upon or even reject".

Theoretical frameworks and models

A number of researchers have argued that much of the student engagement, partnership and co-creation literature tends to be atheoretical (Gravett et al., 2020; Macfarlane and Tomlinson, 2017). Yet others have argued a range of theoretical frames are used in partnership work and that there is space for creativity and theories emerging from practice (Matthews et al., 2019). It would be fair to say that some of the partnership literature is small scale scholarship providing examples of partnership in action and tends not to connect to theory explicitly. However to claim that the partnership and co-creation literature tends to be atheoretical may be over-stating this point. In their work Matthews et al. (2019) identified that partnership and co-creation work had previously drawn on the following concepts: threshold concepts, participative reality, communities of practice, student engagement, deliberative democracy, and ecology of participation literature. They also argued that partnership and co-creation work had drawn on the following constructs: identity, power, and gender. In Bovill (2019a) I argue that critical pedagogy has underpinned much of my work, but critical pedagogy also informs much co-created curriculum either explicitly or implicitly. Authors of critical pedagogy literature recognize that all education is political and has a role to play in developing critical citizens who not only contribute to, but also define society. Critical pedagogy promotes: greater freedom for students to make choices and decisions about their own learning; more dialogic student-teacher relationships; and a significant shift in perspective to view the learner as a competent contributor to education (Bovill, 2019a; Giroux, 1983). Using the lens of critical pedagogy enables us to see the potential of co-created approaches to contribute to developing negotiation, citizenship and leadership skills, to transform thinking and practice, and ultimately to change society (Bron et al., 2022). Clearly not all authors and practitioners of co-creation and partnership literature are motivated or influenced by critical pedagogy, but some of the characteristics of critical pedagogy can be found in a great

deal of the discourse. In addition, there are many other frameworks and models which are referenced in relation to co-created curriculum, some of which are described briefly here.

Arnstein's Ladder of Citizen Participation (Arnstein, 1969) has informed a number of adapted models, and it is explicitly translated for higher education co-created curriculum in the Ladder of student participation in curriculum design (Bovill and Bulley, 2011). The Bovill and Bulley (2011) model includes different ladder rungs illustrating different possible levels of participation by students in curriculum design. Arnstein's original work was in the field of town planning, and she expressed a deep frustration with planning approaches where decisions were made democratically only in name—essentially empty claims of participation. Her ladder helped to highlight the lack of power held by citizens in many planning decisions, consistent with the importance of power highlighted by Matthews et al. (2019).

The participation matrix is often used in the business and international development sectors to consider stakeholders' roles in a project—the possibilities include: inform, consult, participate, partnership, or control. Each stage of a project can then be considered in terms of the role each stakeholder adopts. For curriculum co-creation, use of the matrix helps to highlight that different students and teachers can be involved in different ways at different stages of the work (Bovill, 2017). The participation matrix helps to push an existing concept of stakeholder consultation to a deeper level of analysis by highlighting the range of stakeholders involved and the complexity of who should be involved when and in what ways. This tool can enable people to see past simplistic notions of someone being involved or not involved, toward more nuanced negotiation of the different roles people can adopt at different stages of co-creation work. It also moves past business notions of focusing the analysis of stakeholders on whether they are important or influential.

The decision-modes framework from Heron (1992) draws from a humanist perspective to highlight the early decisions made by teachers in learning and teaching collaborations. Heron highlights that in any teaching situation, staff make choices whether to “direct,” “negotiate” or “delegate” authority (Heron, 1992). Direction involves staff making decisions for students, negotiation involves staff making decisions with students, and delegation is where staff step back and students make decisions. Heron (1992) argues that whichever “decision-mode” is used is ultimately decided by staff. He also highlights that these decisions are taken on multiple levels in relation to any teaching. This work is helpful in highlighting the different levels at which teachers and students can co-create, as well as likely patterns of who will be making decisions at each level.

The co-creation of learning and teaching typology (Bovill, 2019b) draws together many previous ideas and models to pose a range of questions to those involved in co-creation work to help them frame their work and be more explicit about the kind of co-creation they are doing. These questions enable distinctions to be made between different forms of co-creation, for example: whole class and project-based co-creation; the focus of co-creation work; which students are involved; and the nature of student involvement. Used at an institutional level, it can enable colleagues to see trends in the kinds of co-creation in practice. It acts as an accessible planning tool, enabling staff and students to be clearer about the kind of work they are planning, implementing or evaluating.

The student roles model (Bovill et al., 2016), highlights four different but overlapping roles students commonly adopt when co-creating learning and teaching: “(1) consultant, sharing and discussing valuable perspectives on learning and teaching; (2) co researcher, collaborating meaningfully on teaching and learning research or subject-based research with staff; (3) pedagogical co-designer, sharing responsibility for designing learning, teaching and assessment; and (4) representative, student voices contributing to decisions in a range of university settings” p. 197–198. The authors highlight differences between the roles in terms of whether they tend to be student or staff-led, and whether students are typically elected, selected or all students have the opportunity to adopt these roles in co-creation. The model particularly highlights differences between representation and other co-creation roles, and echoes the discussion of identity in Matthews et al. (2019).

The students as partners in learning and teaching in higher education—an overview model (Healey et al., 2014) is a commonly referenced model within partnership literature. The model comprises a Venn diagram of overlapping circles: Learning, teaching and assessment; curriculum design and pedagogic consultancy, scholarship of teaching and learning; and subject-based research and enquiry. At the heart of the model is “partnership learning communities.” The model was used as the basis for a systematic review of partnership literature (Mercer-Mapstone et al., 2017) and has been used widely as a reference point in discussions of partnership. It has not been used as widely to inform discussion of co-created curriculum. Perhaps the separation of learning, teaching and assessment from curriculum within the model creates some tensions for conceptualizations of co-created curriculum. If we adopt a process conceptualization of the curriculum, as more than a list of topics to be taught and learned, curriculum will then include learning, teaching and assessment rather than being treated separately from the curriculum, as in this model.

The Value in Use model from Ranjan and Read (2016) highlights the three key concepts of personalization, experiences, and relationships in measuring the value of products within a marketing context. Dollinger and Lodge (2020) apply this model as a lens through which to explore their research findings and to judge the value of partnership to the student experience. They argue there might be benefits for universities to adopt the marketing concept of value to benchmark and assess the student experience.

The student partnership framework (Olsen, 2007) is divided into four quadrants representing different ways of understanding student partnership from an institutional perspective: apprentice; democratic participant; customer or consumer; or follower of political agendas. The roles are influenced by two axes: one which is focused on internal or external pressure toward student participation and the other focused on actors having either shared or conflicting norms and objectives (Holen et al., 2020). These different partnership roles are informed by other frameworks—apprentice, coming from the Humboldt vision of the university; the democratic participant, from a vision of a representative democracy; the customer or consumer, based on a service enterprise vision; and the follower of political agendas, based on viewing universities as places that can influence politics.

The theoretical model for students as change agents from [Dunne and Zandstra \(2011\)](#) is also a model based on quadrants, which focuses on four roles students adopt as change agents: evaluators of their HE experience (the student voice); participants in decision-making processes; partners, co-creators and experts; and agents for change. The two axes divide the quadrants by whether there is emphasis on the university or the student as a driver, and between whether there is emphasis on the student voice or student action. This model connects well with the construct of power noted as important by [Matthews et al. \(2019\)](#), in highlighting the changes in power necessary to enable students to gain agency to make change in their university. The original concept of change agents comes from business models of change.

Although there are overlaps between ideas in some of the models, each individual model or framework offers a lens through which we can conceptualize co-created curriculum (and wider partnership and student engagement initiatives). Each framing has the potential to support greater understanding of co-created curriculum. This is not a comprehensive list of relevant models and frameworks, but simply offers an introduction to some of the useful frameworks available.

Examples of practice

Having explored different ways in which co-created curriculum is conceptualized, it is worth taking a closer look at some of the many examples of co-created curriculum from around the world. I have classified the small selection of examples included here as either examples of co-creation *of* the curriculum or co-creation *in* the curriculum. Note that the examples of co-creation *of* the curriculum more commonly involve selecting a small group of students to be involved. In contrast, examples of co-creation *in* the curriculum commonly involve the whole cohort, class or group of students studying the course. Further examples can be found in [Barrineau et al. \(2019\)](#), [Bovill \(2020a\)](#), [Cook-Sather et al. \(2014\)](#), [Werder and Otis \(2010\)](#).

Co-creation *of* the curriculum

Course design teams at Elon University, North Carolina, USA

Mihans and colleagues describe bringing together a course design team in which students and staff worked together to re-design an existing education course ([Mihans et al., 2008](#)). The 10 person course design team was deliberately made up of more students than staff in an attempt to change the power dynamics typical in curriculum design discussions. The team comprised seven undergraduates, two faculty who taught the course; and one faculty developer. The students included some students who had previously studied the course and some students who were going to be studying the course. The team met about 12 times over a period of two to three months to carry out the course design work. Time was considered a key element of success in this work, with the regular meetings providing opportunities to achieve the course design aims as well as develop strong relationships, trust and shared respect.

Co-creating a science skills curriculum at the University of Glasgow, Scotland, UK

Woolmer and colleagues co-created a new inter-disciplinary science skills course for students from Physics and Astronomy, Geographical and Earth Sciences, and Chemistry ([Woolmer et al., 2016](#)). A staff member and a student from each of the subject areas worked together over the two phases of the project. First, the group surveyed students and staff from all the discipline areas to ask their views about the most important science skills students should learn in their discipline and which of these were currently taught well at the University. Several focus groups also explored these issues in more depth. In the second part of the project, students and staff co-created teaching materials for a new science skills course focused on the key skills identified in the first part of the project, and a plan for a teaching approach for the course based on research into science skills teaching. Students created high quality lecture notes, team projects, laboratory exercises, example sheets and exam questions. This new course ran so successfully in its first iteration that the course credits were doubled and the course expanded further in the following year.

Co-creating a chemistry course at Uppsala University, Sweden

[Bengtson et al. \(2017\)](#) co-created a first year undergraduate chemistry course. The co-creation project was inspired by a downturn in recruitment for the course due to senior students describing the course to junior students as difficult and of limited use. In the first stage of the co-creation process, they interviewed students to ensure their views were heard about the course. The second stage of the process involved creating a development team of students and teachers numbering eight people in total, working together to re-design the ways the course was taught. This involved for example, ensuring active student participation was a key part of every lecture, but also making changes to the course reading list and creating a table of formulas. "A difficult part of the process was to find a balance between defining the project and giving students enough power and freedom" ([Bengtson et al., 2017: 6](#)), an observation shared by many people co-creating curriculum. The changes made to the chemistry course led to much higher levels of student satisfaction with the course.

Co-creating an interdisciplinary program, University of Technology, Sydney, Australia

[Baumber et al. \(2020\)](#) co-created the curriculum for the interdisciplinary final year of a creative intelligence and innovation program. The course program is focused on addressing "... complex real-world challenges through collaboration and mutual learning across disciplines and with a variety of industry, government and community partners." p. 4. The work took place in three stages, the first of which was to hold participatory design sessions with staff and students who were going to be studying the

program. The second stage involved ensuring that all the different perspectives and experiences in the group were fed into the design process. This stage also involved ensuring these processes were documented in order for each stage to inform and enhance following stages. The third stage comprised trying out alternative co-creation approaches with a second cohort of students. The entire process was cyclical and iterative, emphasizing reflection on, and learning from, each stage of the design, and feeding this into the next stages.

Co-creation in the curriculum

Shared PowerPoint slides, University College Northern Denmark (UCN), Aalborg, Denmark

Louise Esko Refshøj and Steffen Holme Helledie at University College Northern Denmark teach a class focused on innovation. They were committed to students being a meaningful part of creating the course, as they recognized that students have ideas about what innovation is before they take a course on innovation. In order to tap into students' existing expertise, they created a shared PowerPoint document and gave all the students editing rights to the document. They asked each student to create a couple of slides that would represent their perspectives about what innovation is. The students created a large set of slides, and as the course progressed, in each class there would be time for students to present their slides to the rest of the class. It was just one element of the course design which helped to communicate to students that they were important contributors and collaborators in learning and teaching (Bovill, 2020a).

Co-creating essay titles, University of Vienna, Austria

Professor Peter Kruschwitz at the University of Vienna teaches Classics, and for many years now, has invited his students to co-create their own essay titles. He provides students with six to eight keywords, and then asks his students to create their essay title based on these keywords. This simultaneously keeps them focused on the topic, but gives them the freedom to choose which elements of the topic they are most interested in and to construct their essay title accordingly. He talks with students about their titles and sometimes suggests changes, but usually on the basis of scope—where essays will be difficult to write because the title is too narrow or too broad. Over the years he has evaluated this work and demonstrated increased grade point averages for students. He puts this down to increased student interest and motivation in their assignments compared with assignments set by the staff (Bovill, 2020a; Cook-Sather et al., 2014).

Co-creating courses within a Masters in Education, Universiti Utara Malaysia, Sintok, Malaysia

Kaur et al. (2019) describe their work to co-create four different courses within a Masters in Education. Each course involved all the students on the course (the 4 courses involved 7, 15, 16 and 28 students) working with the course instructors in partnership for one semester to plan, design and also teach the course. Students worked in small groups and were enabled to select elements of the curriculum that they were most interested in, and then they were provided with the relevant learning outcomes and assessment guidelines. Students and instructors discussed their preferred ways of learning so that this could inform the work they would do to co-create the courses. The students and instructors met throughout the semester to design and enhance the teaching plan. Each course ended up having between two and four sessions which were co-created by students and staff. The staff and students worked alongside each other to teach these co-created sessions. Importantly, Kaur et al. (2019) explore in their research whether some students in an East Asian setting might find co-creation an alien concept, and while some students were a little reluctant (similar to some students in all contexts), ultimately, many students adapted well to co-creating the curriculum.

Co-creating a business course, University of St Andrews, Scotland, UK

Professor Ignacio Canales co-created an Entrepreneurship and Business Planning course with 50 3rd and 4th year students when based at the University of St Andrews. As one of the main assignments for the course, the students are required to work together on group projects to create a business idea. Once they have been assigned to groups and have agreed what they are focusing on, the whole class is invited to share ideas about what they need to learn in order to be able to successfully achieve the group project. All the students and Professor Canales then negotiate sharing out responsibility for reading about, and creating engaging teaching materials to teach the rest of the class about these key ideas (Bovill, 2020a; Cook-Sather et al., 2014).

Benefits

One likely reason for the increased interest in students and staff co-creating curriculum over recent years, is the growing evidence demonstrating the benefits of co-creation. Many of the benefits are connected with student-staff partnership in learning and teaching, but there are many overlaps between the findings from studies specifically focused on co-created curriculum too. Partnership in learning and teaching has been demonstrated to lead to: enhanced engagement and motivation for learning and teaching; increased meta-cognitive awareness of learning and teaching and a more developed sense of identity; and enhanced classroom practices (Cook-Sather et al., 2014). Interestingly, many of these benefits are shared by students and staff, even if experienced in slightly different ways. Mercer-Mapstone et al. (2017) in their systematic review of "students as partners" literature also draw attention to beneficial enhancements for students to their sense of belonging, confidence, positive relationships, and trust.

Bovill (2020b) synthesizes additional benefits specifically demonstrated to result from co-creation of learning and teaching with a whole class of students, which include: enhanced academic performance; greater academic aspirations; enhanced skills development; opening up the learning process to become more transparent; a shift in the focus from grades to learning; increased ownership, autonomy and self-regulation; students feeling valued. Other findings from Bovill (2020b) highlight the role co-created curriculum can play in: providing an opportunity to practise working democratically; enabling students to experience a more equal relationship with a teacher; and the curriculum becoming more socially relevant—the latter findings backed up by other research demonstrating co-created curriculum is considered to contribute to social justice and civic engagement (Bergmark and Westman, 2016; Lubicz-Nawrocka, 2016; Owusu-Agyeman and Fourie-Malherbe, 2019). Finally, co-created curriculum has also been described as an authentic and enjoyable process (Lubicz-Nawrocka, 2019; Temple Clothier and Matheson, 2019).

These are all compelling outcomes, but also important is to try to capture the sense of personal transformation and significance from co-creation experiences. In Lubicz-Nawrocka's (2020) research, she found that every student in her study who was exposed to co-created curriculum (through a range of different co-creation initiatives across Scotland), reported that the co-created course that they experienced at university was the best course they had studied. We should also acknowledge that personal testimonies can be powerful and offer a different, but equally important form of evidence. One student in Lubicz-Nawrocka (2019) comments on her experiences of co-creating curriculum,

to me it is about being conscious, not only about what you are doing, but what you are capable of doing and definitely to challenge staff and other students. ... It is about going to university and not being passive and waiting for people to tell you what to do. It's getting into an environment where you are able to really get your creative side out and trying to see what you can do ... and be confident in what you are doing (p. 207).

Another student in the same study comments,

I was actually considering dropping out throughout last year so having this course to look forward to was the main reason why I stayed really. (Lubicz-Nawrocka (2019, p. 205).

Guy Finkill, in Barrineau et al. (2019), argues that

to experience any other form of education now would feel like nothing less than a fundamental step into the past ... (p. 58).

From a staff perspective:

It was liberating ... we moved from teaching ... that just didn't work to... [teaching] ... that ... was put together in ways that I never even imagined were possible ... so ... it's really transformed how I think about teaching and how I teach

Bovill (2014, p. 18).

Alison Cook-Sather, interviewed at Uppsala University in 2015, also highlights the value of co-creation from a staff perspective

I value your perspectives on what learning is like in my classroom, I can't know that, I can only know what it's like to teach in this classroom but only you know what it is like to learn in this classroom. So you have a perspective that I don't have and that I would benefit from hearing so that I can make sure that the learning is the best experience that it can be for you

Cited in Barrineau et al. (2019, p. 174).

Challenges

We are enculturated to believe that learning, teaching and curriculum design should be done in certain ways in higher education. Therefore, co-creating curriculum requires many students and staff to adapt their approaches to learning and teaching as well as the ways they relate to one another. In other words, for many people, co-creation requires a change of mind-set. It is perhaps unsurprising, therefore, that we sometimes encounter resistance from student and staff colleagues to this suggestion of change (Bovill et al., 2016). Yet, there are many examples of successful co-creation practice, processes and outcomes, which can be used to help overcome this resistance. Other common challenges include the perception that the constraints that exist to changing the way we "do" curriculum are more rigid than they actually are. As an example, professional body requirements in some programs may often set some quite distinct graduate outcomes that students are expected to achieve, yet there is commonly a good deal of flexibility as to how to achieve these outcomes (Bovill et al., 2016; Hutchings et al., 2011).

Bovill et al. (2016) argue that institutional systems and structures often act as a barrier to co-creating curriculum. For example, many institutions have long processes of course and program approval, and large class sizes are common. Larger classes can make co-creation more challenging, although there are some useful examples of co-creation in larger classes (Bovill, 2020a; Cook-Sather et al., 2014). The sustainability of co-creation initiatives has also been highlighted as a potential challenge (Cook-Sather et al., 2014), and another widely discussed challenge has focused on how to ensure co-created curriculum initiatives are inclusive (Dwyer, 2018; Marquis et al., 2018; Mercer-Mapstone et al., 2021).

Colleagues who are less used to teaching and learning in an online environment may struggle to adapt the concept of curriculum co-creation to an online environment. Yet, this is not a concept that “belongs” to in-person, face-to-face settings. This is an approach, a set of values and a mindset, which is not mode- or setting-dependent. The more we share examples and ideas, the more feasible co-creation may seem to those who feel less confident in adopting this approach. Co-created curriculum requires authenticity and sincerity to the offer to co-create, it also requires shared responsibility, respect and reciprocity (Cook-Sather et al., 2014).

Can we afford not to co-create the curriculum?

Many researchers and practitioners around the world are beginning to recognize the powerful possibilities of meaningful student involvement in curriculum design in higher education. Students are key stakeholders in their own learning. Involving students’ multiple perspectives in curriculum design enhances the relevance of curriculum to students as well as their experiences of the curriculum. As we recognize the benefits of co-creation, I have noted that there are increasing calls for co-created curriculum and partnership opportunities to be inclusive. One step in this direction is to co-create the curriculum with a whole class of students (Bovill, 2020b), while not assuming that inviting a whole class or group to co-create means that the offer is automatically inclusive. Particular attention needs to be paid to ensuring underserved students have genuinely equal opportunities to participate. Another area of work related to co-creation, which is gaining increased attention, is the key role played by positive relationships as a powerful driver of successful outcomes for students (Bovill, 2020a; Chickering and Gamson, 1987; Felten and Lambert, 2020; Mayhew et al., 2016). Indeed, positive student-staff relationships are not only a foundational requirement for co-creation, but co-creation leads to enhanced relationships between students and staff, and between students and other students (Bovill, 2020a).

Freeman et al. (2014) in their significant paper outlining the evidence in support of active learning from science, technology, engineering and mathematics, suggest that if active learning were a treatment in a drug trial, the control group in the trial would be suspended, because the evidence is so strong that active learning is beneficial. There is currently less quantitative evidence for co-creation and partnership, due in part perhaps to the nuanced nature of the values and processes involved. Yet as we reflect on the growing evidence base for the benefits of co-created curriculum, including Lubicz-Nawrocka’s (2020) findings that the course which students co-created was the best one they experienced at university, I am left wondering whether we may soon reach a point where we need to ask ourselves if we can justify not providing opportunities for every university student to experience both co-creation *of* and *in* the curriculum.

There are certainly areas of our practice which may not lend themselves perfectly to co-creation. And as with any major shift in the way we think about curriculum, changes will not happen overnight. However, we have witnessed significant change over the last decade as a very small group of enthusiastic staff and student colleagues has grown to become a much larger group; perhaps even a movement of people, committed to changing the way we create curricula with students rather than for students.

References

- Arnstein, S.R., 1969. A ladder of citizen participation. *J. Am. Inst. Plan.* 35, 216–224.
- Barnett, R., Coate, K., 2005. *Engaging the Curriculum in Higher Education*. Open University Press, Maidenhead.
- Barrineau, S., Engström, A., Schnaas, U., 2019. *An Active Student Participation Companion*. Uppsala University, Uppsala. Available at: www.diva-portal.org/smash/get/diva2:1286438/FULLTEXT02.pdf.
- Baumber, A., Kligyte, G., Van der Bijl-Brouwer, M., Pratt, S., 2020. Learning together: a transdisciplinary approach to student–staff partnerships in higher education. *High Educ. Res. Dev.* 39, 395–410.
- Beane, J., 1997. *Curriculum Integration: Designing the Core of Democratic Education*. Teachers College Press, New York.
- Bell, A., Potter, S., Morris, L.-A., Strbac, M., Grundy, A., Zaki Yawary, M., 2019. Evaluating the process and product of a student-staff partnership for curriculum redesign in film studies. *Innovat. Educ. Teach. Int.* 56, 740–750.
- Bengtson, C., Ahlqvist, M., Ekeröth, W., Nilsen-Moe, A., Proos Vedin, N., Rodiouchkina, K., Ye, S., Lindberg, M., 2017. Working as partners: course development by a student–teacher team. *Int. J. Scholarsh. Teach. Learn.* 11, 6.
- Bergmark, U., Westman, S., 2016. Co-creating curriculum in higher education: promoting democratic values and a multidimensional view on learning. *Int. J. Acad. Dev.* 21, 28–40. <https://doi.org/10.1080/1360144X.2015.1120734>.
- Biggs, J., 1996. Enhancing teaching through constructive alignment. *High Educ.* 32, 347–364.
- Boomer, G., Lester, N., Onore, C., Cook, J. (Eds.), 1992. *Negotiating the Curriculum: Educating for the 21st Century*. Falmer Press, London.
- Boomer, G., 1978. *Negotiating the curriculum*. *Engl. Aust.* 44, 16–29.
- Bourner, T., 2004. The broadening of the higher education curriculum, 1970–2002: an ipsative enquiry. *High Educ. Rev.* 36, 39–52.
- Bovill, C., Bulley, C.J., 2011. A model of active student participation in curriculum design: exploring desirability and possibility. In: Rust, C. (Ed.), *Improving Student Learning* (18)—Global Theories and Local Practices: Institutional, Disciplinary and Cultural Variations. The Oxford Centre for Staff and Educational Development, Oxford, pp. 176–188.
- Bovill, C., Woolmer, C., 2019. How conceptualizations of curriculum in higher education influence student-staff co-creation *in* and *of* the curriculum. *High Educ.* 78, 407–422.

- Bovill, C., Cook-Sather, A., Felten, P., Millard, L., Moore Cherry, N., 2016. Addressing potential challenges in co-creating learning and teaching: overcoming resistance, navigating institutional norms and ensuring inclusivity in student-staff partnerships. *High Educ.* 71, 195–208.
- Bovill, C., 2014. An investigation of co-created curricula within higher education in the UK, Ireland and the USA. *Innovat. Educ. Teach. Int.* 51, 15–25.
- Bovill, C., 2017. A framework to explore roles within student–staff partnerships in higher education: which students are partners, when and in what ways? *Int. J. Stud. Partn.* 1, 1–5.
- Bovill, C., 2019a. Student–staff partnerships in learning and teaching: an overview of current practice and discourse. *J. Geogr. High Educ.* 43, 385–398.
- Bovill, C., 2019b. A co-creation of learning and teaching typology: what kind of co-creation are you planning or doing? *Int. J. Stud. Partn.* 3, 91–98.
- Bovill, C., 2020a. Co-creating Learning and Teaching: Towards Relational Pedagogy in Higher Education. Critical Publishing, Northwich.
- Bovill, C., 2020b. Co-creation in learning and teaching: the case for a whole-class approach in higher education. *High Educ.* 79, 1023–1037.
- Breen, M.P., Littlejohn, A., 2000. The practicalities of negotiation. In: Breen, M.P., Littlejohn, A. (Eds.), *Classroom Decision-Making: Negotiation and Process Syllabuses in Practice*. Cambridge University Press, Cambridge.
- Bron, J., Bovill, C., Veugelers, W., 2022. Students experiencing and developing democratic citizenship through curriculum negotiation: the relevance of Garth Boomer's approach. *Curric. Perspect.* 36, 15–27.
- Chickering, A.W., Gamson, Z.F., 1987. Seven principles for good practice in undergraduate education. *AAHE Bull.* 3, 3–7.
- Cook-Sather, A., Bovill, C., Felten, P., 2014. *Engaging Students as Partners in Learning and Teaching: A Guide for Faculty*. Jossey Bass, San Francisco.
- Darder, A., Baltodano, M., Torres, R.D., 2003. Critical pedagogy: an introduction. In: Darder, A., Baltodano, M., Torres, R.D. (Eds.), *The Critical Pedagogy Reader*. RoutledgeFalmer, New York, pp. 1–21.
- Delpish, A., Holmes, A., Knight-McKenna, M., Mihans, R., Darby, A., King, K., Felten, P., 2010. Equalizing voices: student-faculty partnership in course design. In: Werder, C., Otis, M. (Eds.), *Engaging Student Voices in the Study of Teaching and Learning*. Sterling, Stylus, pp. 96–113.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. The Macmillan Company, New York.
- Dollinger, M., Lodge, J., 2020. Understanding value in the student experience through student–staff partnerships. *High Educ. Res. Dev.* 39, 940–952.
- Duah, F., Croft, T., 2012. Students as partners in mathematics design. In: Waller, D. (Ed.), *Centre for Excellence in Teaching and Learning of Mathematics, Statistics and Operational Research (CETL-MSOR) Conference. The Maths, Stats and OR Network*, pp. 49–55. In: <http://www.sigma-network.ac.uk/wp-content/uploads/2016/11/CETL-MSOR-Proceedings2011.pdf>.
- Dunne, E., Zandstra, R., 2011. *Students as Change Agents. New Ways of Engaging With Learning and Teaching in Higher Education*. Higher Education Academy/University of Exeter, Bristol, ESCalate.
- Dwyer, A., 2018. Toward the formation of genuine partnership spaces. *Int. J. Stud. Partn.* 2, 11–15.
- Eisner, E.W., 1979. *The Educational Imagination on the Design and Evaluation of School Programs*. MacMillan Publishers, New York, NY.
- Entwistle, N., 2003. Concepts and conceptual frameworks underpinning the ETL project. In: *Enhancing Teaching-Learning Environments in Undergraduate Courses Project*. University of Edinburgh, Edinburgh. Occasional Reports, 3.
- Fawcett, T., Watson, J., 2019. Co-creating Care Home Nursing Education in the Curriculum. University of Edinburgh Teaching Matters blog, 21st March. <https://www.teaching-matters-blog.ed.ac.uk/>.
- Felten, P., Lambert, L., 2020. *Relationship-Rich Education: How Human Connections Drive Success in College*. John Hopkins University Press, Baltimore.
- Fink, L.D., 2003. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. Jossey Bass, San Francisco.
- Fink, L.D., 2007. The power of course design to increase student engagement and learning. *Assoc. Am. Coll. Univ. Peer Rev.* 9, 13–17.
- Fischer, M.C.B., 2005. The methodology of “systemization” and its relevance to the academy. In: Crowther, J., Galloway, V., Martin, I. (Eds.), *Popular Education: Engaging the Academy, International Perspectives*. Leicester: Niace.
- Fraser, S., Bosanquet, A., 2006. The curriculum? That's just a unit outline, isn't it? *Stud. High Educ.* 31, 269–284.
- Freeman, S., Eddy, S.I., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P., 2014. Active learning increases student performance in science, engineering and mathematics. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8410–8415.
- Giroux, H.A., 1983. *Theory and Resistance in Education. A Pedagogy for the Opposition*. Heinemann, London.
- Gravett, K., Kinchin, I.M., Winstone, N.E., 2020. “More than customers”: conceptions of students as partners held by students, staff, and institutional leaders. *Stud. High Educ.* 45, 2574–2587.
- Grudens-Schuck, N., 2003. No beginners: teaching participation at the graduate level. *PLA Notes* 48, 11–14.
- Harden, J., Fawcett, S., 2019. A Student Partnership Project to Enhance Curriculum Development in Medical Education. University of Edinburgh Teaching Matters blog, 24th January. <https://www.teaching-matters-blog.ed.ac.uk/>.
- Healey, M., Flint, A., Harrington, K., 2014. *Students as Partners in Learning and Teaching in Higher Education*. Higher Education Academy, York.
- Heron, J., 1992. The politics of facilitation: balancing facilitator authority and learning autonomy. In: Mulligan, J., Griffin, C. (Eds.), *Empowerment Through Experiential Learning: Explorations of Good Practice*. Kogan Page, London, pp. 66–75.
- Hill, J., Thomas, G., Diaz, A., Simm, D., 2016. Borderland spaces for learning partnership: opportunities, benefits and challenges. *J. Geogr. High Educ.* 40, 375–393.
- Holen, R., Ashwin, P., Maassen, P., Stensaker, B., 2020. Student partnership: exploring the dynamics in and between different conceptualizations. *Stud. High Educ.* <https://doi.org/10.1080/03075079.2020.1770717>.
- Hutchings, P., Huber, M.T., Ciccone, A., 2011. *The Scholarship of Teaching and Learning Reconsidered: Institutional Integration and Impact*. Jossey-Bass/Carnegie Foundation for the Advancement of Teaching, San Francisco. Jossey-Bass.
- Huxham, M., Hunter, M., McIntyre, A., Shilland, R., McArthur, J., 2015. Student and teacher co-navigation of a course: following the natural lines of academic enquiry. *Teach. High. Educ.* 20, 530–541.
- Iversen, A.-M., Stavnskaer Pedersen, A., 2017. Co-creating knowledge: students and teachers together in a field of emergence. In: Chemi, T., Krogh, L. (Eds.), *Co-creation in Higher Education: Students and Educators Preparing Creatively and Collaboratively to the Challenge of the Future*. Sense, Rotterdam, pp. 15–30.
- Joseph, P.B. (Ed.), 2011. *Cultures of Curriculum*, second ed. Routledge, New York & London.
- Kaur, A., Awang-Hashim, R., Kaur, M., 2019. Students' experiences of co-creating classroom instruction with faculty—a case study in eastern context. *Teach. High. Educ.* 24, 461–477.
- Kenny, M., 2019. Student-Staff Co-creation of a Course: Understanding Gender in the Contemporary World. University of Edinburgh Teaching Matters blog, 10th January. <https://www.teaching-matters-blog.ed.ac.uk/>.
- Lubicz-Nawrocka, T., 2016. Co-creation of the Curriculum and Social Justice: Changing the Nature of Student–Teacher Relationships in Higher Education. Paper Presentation, Higher Education Close up Conference: Locating Social Justice in Close-Up Research in Higher Education, Lancaster, UK. <http://www.lancaster.ac.uk/fass/events/hecu8/abstracts/lubicz-nawrocka.htm>.
- Lubicz-Nawrocka, T., 2019. Creativity and collaboration: an exploration of inclusion, empathy, and resilience in co-creation of the curriculum. *Stud. Engagem. High. Educ. J.* 2, 199–213.
- Lubicz-Nawrocka, T., 2020. *An Exploration of How Curriculum Co-creation Advances Student and Staff Aims for Scottish Higher Education*. Unpublished PhD Thesis. University of Edinburgh, Edinburgh. <https://doi.org/10.7488/era/496>.
- Macfarlane, B., Tomlinson, M., 2017. Critiques of student engagement. *High Educ. Pol.* 30, 6–21.
- Marquis, E., Jayaratnam, A., Mishra, A., Rybkina, K., 2018. “I feel like some students are better connected”: students' perspectives on applying for extracurricular partnership opportunities. *Int. J. Stud. Partn.* 2, 64–81.

- Matthews, K.E., Cook-Sather, A., Acai, A., Dvorakova, S.L., Felten, P., Marquis, E., Mercer-Mapstone, L., 2019. Toward theories of partnership praxis: an analysis of interpretive framing in literature on students as partners in teaching and learning. *High Educ. Res. Dev.* 38, 280–293.
- Mayhew, M.J., Rockenbach, A.N., Bowman, N.A., Seifert, T.A., Wolniak, G.C., Pascarella, E.T., Terenzini, P.T., 2016. *How College Affects Students. Volume 3 21st Century Evidence that Higher Education Works.* Jossey Bass, San Francisco.
- McCune, V., Hounsell, D., 2005. The development of students' ways of thinking and practicing in three final year biology courses. *High Educ.* 49, 255–289.
- Mercer-Mapstone, L., Abbot, S., 2020. In: *The Power of Partnerships: Students, Staff, and Faculty Revolutionizing Higher Education.* Elon University Center for Engaged Learning, Elon, N.C.
- Mercer-Mapstone, L., Dvorakova, S.L., Matthews, K.E., Abbot, S., Cheng, B., Felten, P., Knorr, K., Marquis, E., Shammass, R., Swaim, K., 2017. A systematic literature review of students as partners in higher education. *Int. J. Stud. Partn.* 1, 1–23.
- Mercer-Mapstone, L., Islam, M., Reid, T., 2021. Are we just engaging “the usual suspects?” challenges in and practical strategies for supporting equity and diversity in student–staff partnership initiatives. *Teach. High. Educ.* 26, 227–245.
- Mihans, R., Long, D., Felten, P., 2008. Power and expertise: student–faculty collaboration in course design and the scholarship of teaching and learning. *Int. J. Scholarsh. Teach. Learn.* 2, 1–9.
- Moore-Cherry, N., Healey, R., Nicholson, D.T., Andrews, W., 2016. Inclusive partnership: enhancing student engagement in geography. *J. Geogr. High Educ.* 40, 84–103.
- Olsen, J.P., 2007. The institutional dynamics of the European University. In: Maassen, P., Olsen, J.P. (Eds.), *University Dynamics and European Integration.* Springer, Dordrecht, pp. 25–54.
- Owusu-Agyeman, Y., Fourie-Malherbe, M., 2019. Students as partners in the promotion of civic engagement in higher education. *Stud. High Educ.* <https://doi.org/10.1080/03075079.2019.1666263>.
- Pinar, W., 1975. *Currere: towards reconceptualization.* In: Pinar, W. (Ed.), *Curriculum Theorizing: The Reconceptualists.* McCutchan Publishing Corporation, Berkeley, pp. 396–414.
- Ranjan, K.R., Read, S., 2016. Value co-creation: concept and measurement. *J. Acad. Market. Sci.* 44, 290–315.
- Rogers, C., Freiberg, H.J., 1983. *Freedom to Learn*, third ed. Macmillan Publishing, New York.
- Stenhouse, L., 1975. *An Introduction to Curriculum Research and Development.* Heinemann, London, UK.
- Temple Clothier, A., Matheson, D., 2019. Using co-creation as a pedagogic method for the professional development of students undertaking a BA (Hons) in Education Studies. *J. Furth. High. Educ.* 43, 826–838.
- Toohey, S., 1999. *Designing Courses for Higher Education.* Society for Research into Higher Education/Open University Press, Maidenhead.
- Van Werven, I., Haagsma, M., Ruckstuhl, I., Gelissen, L., Van der Stouwe, A., Elting, P., 2017. Internationalization of teacher education curricula through student involvement. In: Coelen, R., Van der Hoek, K.-W., Blom, H. (Eds.), *Valorization of Internationalization: About Internationalization of Higher Education.* Stenden University of Applied Sciences, Stenden, pp. 71–85.
- Waddington, K., 2016. The compassion gap in UK universities. *Int. Pract. Dev. J.* 6, 1–9. Article 10. <https://www.fons.org/library/journal-ipdj-home>.
- Werder, C., Otis, M. (Eds.), 2010. *Engaging Student Voices in the Study of Teaching and Learning.* Stylus, Sterling.
- Wijaya Mulya, T., 2019. Contesting the neoliberalization of higher education through student–faculty partnership. *Int. J. Acad. Dev.* 24, 86–90.
- Wilkinson, M., Scandrett, E., 2003. A popular education approach to tackling environmental injustice and widening participation. *Concept* 13, 11–16.
- Woolmer, C., Sneddon, P., Curry, G., Hill, B., Fehertavi, S., Longbone, C., Wallace, K., 2016. Student staff partnership to create an interdisciplinary science skills course in a research intensive university. *Int. J. Acad. Dev.* 21, 16–27.

Continental pedagogy & curriculum

Norm Friesen^a and Karsten Kenkies^b, ^a Boise State University, College of Education, Boise, ID, United States; and ^b University of Strathclyde, School of Education, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Pedagogy: structures of relation	246
Educator-educand relation: the pedagogical relation	248
Establishing the relation	248
The personal character of the relation	249
The educator-world relation	249
Changing the relation between educand-world: didactics	250
Emphasizing the educator-world relation: representation	250
Emphasizing the educator-educand relation: communication	250
Emphasizing the educand-world relation: experience	251
The educand-world relation—<i>Bildung</i> and the educational telos	252
Pedagogy as a “science”: theory and art	252
Ways of knowing: continental pedagogy and curriculum studies	253
References	254

Glossary

Art *techne*, a practical knowing, doing, and reflection that involves craft and technique and that is variable and context dependent

Bildung The process of the forming of humans, as well as the determination of the goal and purpose of human existence (Benner and Brüggem, 2004, p. 180)

Didactic transposition (or reduction) The specific arrangement of the pedagogical situation through the educator (often involving simplifications and/or exemplifications) to allow for learning and transformation to happen

Didactics The study (and doctrine) of learning and teaching in general, the professional science of the teacher

Dignity of practice A value and integrity intrinsic to practice that gives it a certain independence from and precedence over theory

Educand The one receiving education as a type of influence, regardless of the context

Educator The one influencing the educand, with an intention to improve them and their relation to the world

Negative education Education or influence of the educand, conceived as a protection of the educand from other influence seen as undesirable

Pedagogy The practice of influencing someone (including oneself) with the intention of improving them, including also both informal and more theoretical reflection on this practice

Pedagogics A rigorous attempt at a theory of education

Representation Deliberately presenting aspects of ways of life and the world to the educand (Mollenhauer, 2014, pp. 34–55)

Science Any rigorous academic pursuit aiming at theorizing any and all realms of the world (e.g. nature and the natural sciences; culture and the humanities or human sciences)

Self-activity Activity on the educand's own initiative, for the educand's own purposes. May be prompted or elicited by an educator

Continental Pedagogy refers to a way of thinking about educational practice that coalesced in Europe some 250 years ago and that continues to be elaborated today. Developed from the Greek παιδαγωγία (*paidagōgīā*; *paidos* = child; *agōgos* = leader) and rendered as *la pédagogie* in French, *pedagogía* in Spanish and *Pädagogik* in German, Continental Pedagogy is based on understandings that are radically different from those taken for granted in education in the English-speaking world.

Continental Pedagogy¹ sees education as being primarily concerned neither with school and curricula, nor with the pedagogy in the English sense of the techniques and politics of teaching (Friesen, 2019). Instead, it begins with the insight that education and

¹The phrase “Continental Pedagogy” represents something of a neologism and a compromise; notably, it does *not* simply refer to the conjunction of continental *philosophy* and pedagogy. Although it has occasionally been used to denote the pedagogical tradition and way of thinking outlined here (e.g., Ponte and Rönnerman, 2009; Kudláčová and Rajskeý, 2019), we recognize that it designates a heterogeneous and heteronomous field or set of fields in Europe. Continental Pedagogy is like continental philosophy in that it is much more expressive of Anglophone awareness than of any European self-descriptions.

pedagogy are both about the influence that is exercised by one person or group (i.e., the educator) on another (i.e., the educand), and that this influence exists in special relation to the world. The resulting configuration of self, other and world forms what has been called the “pedagogical triangle” (Fig. 1; e.g., see: [Briançon, 2019](#)). Within this configuration, pedagogy is particularly concerned with the personal experience of the educand with the educator (i.e., the pedagogical relation; [Friesen, 2017](#)) and with the educand’s relationship with the world (*Bildung, formación, le formation*).

Defining Continental Pedagogy through the question of influence and the tripartite configuration of self, other and world gives it a distinctive emphasis on basic relations—ones that can be found in many spheres of human activity. The triangular paradigm of this pedagogy maps to a wide variety of life experiences and human interactions—not at all limited to modern or Western cultures—and it offers its own modes of knowing, reflecting and acting.

The relationship between educator and educand has traditionally been framed intergenerationally as the relation of *adult* and *child*. In this context the central point of pedagogy has been formulated in various ways: As “what ... the older generation ... want[s] with the younger” ([Schleiermacher, 1826/2022](#), p. 46), as “the reaction of society to the fact of [human] development” ([Bernfeld, 1925/1973](#), p. 90) or as “our attitude toward the fact of natality” ([Arendt, 1958/1968](#), p. 196). Correspondingly, Continental Pedagogy does not regard education as something to be studied primarily through disciplines like sociology or psychology. Instead, it sees pedagogy as designating a distinctive perspective, as constituting an autonomous discipline with its own, specifically pedagogical questions, terminology, and ways of knowing.

Continental Pedagogy also represents a tradition of thought in which theoretical and reflective texts—whether ancient or modern, fictional, philosophical, or theological—are seen as providing a general framework for its evolving questions and changing responses to them. Continental Pedagogy has reflected the influence of the Enlightenment as well as of Romantic, conservative and liberal cultural moments. Particularly in resonance with its Romantic inheritance, Continental Pedagogy today constitutes a site of resistance to neoliberal influences that continue to dominate educational policy and practice (e.g., see: [Horlacher, 2016](#), pp. 118–129; see also [Meirieu, 1996](#)).

Curriculum Studies—whether conceived by Franklin Bobbitt, Ralph Tyler or reconceptualized by [Pinar \(2019\)](#)—clearly differs from Continental Pedagogy in its focus and emphases. While Continental Pedagogy starts with a relation between educator and educand situated in the world, Curriculum Studies begins with the *experience* of the *student* in specifically *academic* settings. It focuses on how this academic experience is formed or manipulated—for better or worse—through specifically *institutional* instrumentalities. Nonetheless, some contemporary curriculum theorists have referred to Continental Pedagogy (as defined here) as the “tradition of *Bildung/Didaktik*” and characterized it as yet another type of “curriculum theory” (e.g., [Autio, 2014](#); [Pinar, 2014](#), p. 2). As this article demonstrates, however, the tradition of Continental Pedagogy not only claims a broader conceptual and practical terrain than Curriculum Studies, it also predates it by centuries. Finally, Continental Pedagogy’s personalized non-psychological and -managerial account of education also presents an alternative to instrumentalization long decried in reconceptualist Curriculum Studies.

This article provides an overview of the theoretical and practical domain constituted by Continental Pedagogy, emphasizing key aspects of its history and theory. It concludes by suggesting ways of reconsidering its relationship to Curriculum Studies and by emphasizing how this pedagogy offers ways of knowing and acting that are themselves irreducibly educative or *pedagogical*.

Pedagogy: structures of relation

As one of the first to provide a formal, theoretical account of pedagogy as a field, F.D.E. Schleiermacher (1786–1834) identified the temporally situated relation between the older and younger generations as the basis not just for all educative activity, but for pedagogical theory itself:

A significant part of the activity of the older generation extends toward the younger, and it is less complete or perfect, the less aware the older generation is about what it is doing and why it is doing it. Therefore, there has to be a theory that is based on the relation of the older generation and the younger This relationship between the older and the younger, and the obligations of the one to the other, form the basis on which we will build everything that lies in the scope of our theory (1826/2022, p. 46).

The starting point for Schleiermacher’s theory is the relation of the older and younger generations and their time-bound situation in the world. Within this context, it is the “response” of the older generation that serves as the basis for subsequent pedagogical action and reflection. These and other elements of Schleiermacher’s definition are later reaffirmed and reformulated by Hannah Arendt (1906–1975) an admirer of Schleiermacher’s: “What *concerns* us all is the relation between grown-ups and children in general or, putting it in even more general and exact terms, our *attitude* toward the fact of *natality*” (1954; p. 196; emphases added). In simple and traditional terms, then, pedagogy is a disposition, a way of knowing and thinking begins with the older generation asking itself about its relationship with and its response (or *responsibility*) to the younger. Arendt makes it clear that this concern entails more than just activity and obligation in some general sense; she explicitly positions educators in relation to those being educated as “representatives” of the world, as introducing the educand or child to the world.

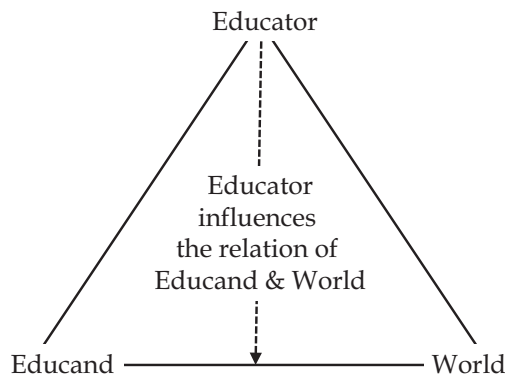


Fig. 1 The pedagogical triangle, showing the relationships between the educator, the one being educated, and the world. Each relationship brings with it different qualities, roles and purposes. The significance of the educator's influence on educand-world relation (indicated by the dashed vertical line) is discussed below (Illustration by authors).

Insofar as the child is not yet acquainted with the world, he must be gradually introduced to it ... the educators here stand in relation to the young as representatives of a world ... it is as though he [the educator] were a representative of all adult inhabitants, pointing out the details and saying to the child: This is our world (1954, p. 189).

As discussed below, the representational and mediating acts of pointing and showing have multiple meanings for pedagogical theory and practice. More important, though, are the three components that constitute this particularly pedagogical situation and their interrelationship. Educand, educator and world together constitute the fundamental structure of pedagogy, the pedagogical triangle (Fig. 1). Each relationship within this triangle—the educator's relation to the educand, the educator's relation to the world, and the child's or educand's relation to the same—constitute the sides of this triangle. As the descriptions of Arendt and Schleiermacher indicate, each relationship in this triadic configuration has a particular character: The *educator's* relationship to the *one being educated*, as Schleiermacher puts it, is a matter of "activity," "awareness" and obligation;" for Arendt, it implies a role for the educator that is particularly "representative." This special relationship is known as the "pedagogical relation"—one characterized by the intention of the educator specifically to influence the educand. More specifically, its orientation and purpose is to change the educand or child's own relation to the world—for example, to have them relate to it with greater familiarity, involvement and investment. In this context, the relation of the *educator* to the *world* takes on a character that it generally does *not* possess outside of pedagogical situations. It is *not* one, for example, where the educator works either to satisfy their own purposes or to address the demands that the world may place upon *them*. It is not a relation of expertise, but rather one in which the educator engages with *the world in terms of its significance for the educand*. For example, as Arendt points out, it is one in which the educator "introduces the world" in various ways, for example, by "pointing out the details" in relation to the child. It is also one in which the educator helps the educand, in part by showing them what is like to be a person who is simultaneously free and responsible. Arendt has the educator say to the one being educated: "this is our world." Through the educand's *Bildung*, their rich and variegated engagement with the world, these words may gradually change in significance: From articulating a kind of claim or challenge to more a statement of fact.

The three relations that constitute the sides of the pedagogical triangle can also be characterized more concretely and pragmatically: The relation between educator and the one being educated encompasses the parent-child and teacher-student relations, and unavoidably entails a personal dimension. The "world" in this sense can refer to the material and natural world, the familial, social and political one, and even the world constituted by one's own self or psyche. The world in this sense has been understood as (human) culture, and in more formal settings, as (educational) "content" (Klafki, 1958/1995)—with the educand's relation to it understood in terms of "knowledge," *savoir* (Houssaye, 2014). *Educand* and *world*, meanwhile are linked through activities like study, play and exercises, through personal interest, life circumstances and through knowing itself. Jean Houssaye, whose work has been widely used in French teacher education, indicates how particular types of educational processes tend to be associated with the specific relationships or "axes" in the pedagogical triangle: "teaching ... favors the teacher-knowledge axis," "learning favors the student-knowledge [or -world] axis," and the teacher-student axis is where Houssaye locates "formation," the French correlative for *Bildung* (2014, p. 12). These relations and the key process of change they can imply provide the outline for much of the remainder of the article and can be summarized as follows:

- **Educator-educand:** a personal pedagogical relation, formative for the educand, allowing the most direct form of educator influence.
- **Educator-world:** a relation focused not on expertise, but on understanding how the world is relevant to the educand, with the aim of improving the educand's relation to it.

- **Change in the Educand-World Relation:** insofar as the educator works to change the educand-world relation, the educator can be said to have their own pedagogical relation to this relation, as indicated by the dotted vertical line in Fig. 1. Educating means to engage with the Educand-World relation in order to change it.
- **Educand-world:** the relation that is the focus of the educator's intention and didactic efforts. Educand and world, meanwhile, are joined through formative experience.

Educator-educand relation: the pedagogical relation

Establishing the relation

Given that it is marked by purpose and intention, the educator's relationship to the one being educated tends to arise deliberately, often through an inaugural moment or explicit decision. Plato's *Republic* (375 BCE/1997), a text that has long served the West as a paradigm for pedagogy, illustrates this through the Allegory of the Cave. The prisoners shackled in the depths of the cave must first be freed from their chains, and then very deliberately "compelled" and "dragged ... by force, up the rough, steep path ... into the sunlight" (515 e). Socrates asks: "wouldn't he be pained and irritated at being treated that way ... [and] unable to see a single one of the things now said to be true?" (516 a). Being thus influenced, gaining insight into the truth, having one's relation to the world changed, can on this account verge on the traumatic. It highlights that the relation of educator and educand requires concerted care and effort on the part of the one educating.

The pedagogical triangle, it should be noted, is here realized as a relation *not* between adult and child—in clear distinction to Schleiermacher's and Arendt's characterizations. It instead arises between adults, between those who are prisoners of their own ignorance and others who undertake to liberate them, with their relation to the world characterized by the morally and epistemologically charged relation between darkness and light, bondage and freedom. Those acting as educators very intentionally change or improve this relation, moving it from imprisoning darkness to enlightened liberation.

Almost 2000 years later, the relationship between educator and educand was recast in the context of what can be regarded as another educational "primal scene." Taken from J.A. Comenius' 1658 *Orbis Pictus* or; *the Visible World*, one of the first picture books for children, it shows an old man pointing and speaking to a child, saying "come boy, learn to be wise" (Fig. 2). The old man can be seen as inaugurating the relation through his gesture as a specifically *pedagogical* relation—one that brings with it the experience of a particularly pedagogical space and time: Space is marked by proximity of adult and child ("come, boy") and time, at least implied by a concern with the *future* ("learn to be wise"). The rest of Comenius' book can be seen as a vast elaboration on this space and time, and of the representative role of the educator as outlined by Arendt. For it goes on to show all aspects of the world and history—as they were then known—not only to one child, but in multiple languages and to a vast range of readers.

Despite the considerable and wide-ranging effort that can be involved, the entry into the space and time of pedagogy is not always explicitly invoked or announced by the educator. In the Romantic novel of education, *Emile; or on Education*, J.J. Rousseau (1712–1778) describes what he calls a "negative education." This is one in which any deliberate or explicit pedagogical plans or intentions are carefully hidden from the educand, the young Emile. Rousseau argues that

the first education ought to be purely negative. It consists not at all in teaching virtue or truth but in securing the heart from vice and the mind from error. If you could do nothing and let nothing be done, if you could bring your pupil healthy and robust to the age of twelve ... at your first lessons the eyes of his understanding would open up to reason (1762/1979, p. 93).



Fig. 2 The pedagogical "primal scene"—with educator, educand and world in indexical interrelation—as shown on the first page of Jan Amos Komenský's (Comenius; 1592–1670) *Orbis Pictus* (Public Domain).

In this initial negative education, the pedagogue works largely if not exclusively not to influence the child directly, but to shield the him or her from damaging external influences. Philippe Meirieu observes that Rousseau's vision has "become the starting point of all pedagogies that want to place the child's activity at the heart of their approach" (p. 83). It entails a pedagogical relation between tutor and child that lasts multiple years (as it would between parent and child), but one—in distinction to Comenius' pedagogical "scene"—in which any direct pedagogical influence is itself hidden from view.

The personal character of the relation

Herman Nohl (1879–1960), whose account is used here as an example, coined the phrase "pedagogical relation" early in the 20th century. He describes it as arising not through specialization or subject-matter expertise, but in terms of a basic attitude or stance (*Grundeinstellung*) taken up by the educator:

This basic stance ... is decisively characterized by the fact that its perspective is unconditionally that of the child or student ... This means that its task is not to draw the child towards ... specific, predetermined, objective goals [of] the state, the church, law, the economy Instead, it sees its goal in the subject and his/her physical and *personal* realization or unfolding. That this child here comes to his life's purpose, that is [the] ... autonomous and inalienable task [of this pedagogy] (1926/2022, p. 65; emphasis added).

Following Rousseau, Nohl here depicts the role of the educator as a largely "negative" one of protection and preservation—from influences of the state, law and economy. Nohl's point is to allow for the uninhibited flourishing of the nascent personality of the child—the development of the "life purposes" that dwell within the child and that have yet to clearly emerge. As a result, the educator fulfills not just a professional but also a personal role: "The basis of education," as Nohl later remarks, "is the passionate relation between a mature person and one who is becoming specifically for the sake of the latter, so that he comes to his life and form" (1933/2022, p. 79).

The pedagogical relation is in this sense an *asymmetrical* one: The educator is there for the educand or child in a way that the child is simply *not* for the educator. This personal "relation of the educator to the child," Nohl continues, "is always doubly determined: by the love for the child as he is and by the love for his educational goal, the child's ideal" (p. 80). This means that educators must balance their engagement with the child between the wants of the present and expectations for the future. The fact that the pedagogical present is in tension with the future in this way further distinguishes the "time" of pedagogy from everyday temporality. As Schleiermacher points out (see, 1826/2022, pp. 50–56), our being with children is neither marked by sustained indulgence in the present nor by the continuous "sacrifice" of present enjoyment for the future (as might happen at work). Nohl explains that the "peculiar opposition and entwinement of two directions of work [oriented the present and future] constitutes the pedagogical stance and gives the educator a singular distance to his subject as well as to his child or student. In its most refined expression," Nohl concludes, "this distance is called pedagogical tact" (1933/2022, p. 81). This tactful disposition and the distance it implies is illustrative of the space of pedagogy, as understood in this case—one in which the educator gives the child room to become themselves, but remains close enough to help.

In both Comenius, Rousseau and Nohl (and also in many more recent conceptions), the relation between educator and educand is understood as a one-to-one relationship. The influencing constitutive of pedagogy, however, is in reality often directed at multiple individuals. Many Continental theorists have thus understood educator, educand and the educational process as situated in a special type of community. For example, the Swiss educational reformer, Johann Heinrich Pestalozzi (1746–1827), expressed in 1799 his desire to transform an orphanage he was setting up "into the spirit of a single large household"—a domestic community in which Pestalozzi himself be the *pater familias* (1799/2022, p. 21). In this role, Pestalozzi sought to cultivate the minute but practical interdependencies of everyday domestic life—ones still indispensable to the development of trust and wellbeing of children today. "In this way," Pestalozzi explains, "the children may be reminded of their own feelings and supplied with a moral and communal understanding of their lives and interrelationships" (1799/2022, p. 21).

In the pedagogical relation, finally, the educator educates through their own self and via the dimensions of the relation itself. Here, the educator *may* serve as a role model, even a representative of all adults or humanity. But what is more important is that they still remain a particular person, with their individual love for the educand, and their own ways of relating. Only by being with a *particular* other can the educand be sensitized to "their own feelings" and develop an "understanding of their lives and interrelationships" in ways that are practical, moral, communal and personal. In this sense, as Houssaye points out, pedagogical relations can themselves be profoundly formative, an intrinsic part of the process of *Bildung*.

The educator-world relation

As mentioned, in relating pedagogically to the world, the educator is not necessarily concerned with knowing or improving the world for *its* sake, but rather, for the sake of the one being educated. The educator's relation to the world is thus *not* one marked by expertise, by a focus on some specific aspect of the world to the exclusion of the others. Just as the educator is not necessarily a subject-matter expert, the goal of pedagogy is not fulfilled via the development of new experts in science, math or any other

subject. Instead, its goal is the development of “well-formed laypersons” (*gebildete Laie*), well-rounded individuals ready for participation in democratic society (Klafki, 1958/1995, pp. 17–18).

Educators engage with the world for the sake of the educand, working to influence the educand’s relation to the world by mediating it in various ways for them. Initially, parents or guardians are generally the best educators of their children. This is not because of what parents may know about the world, but because they have the most developed relationship with the child. They are in the best position to know the relevance of things in the world to the child (e.g., bed-time, a favorite toy) and how to change and improve their child’s relation to them. Like the one being educated, educators have direct access to the corner of the pedagogical triangle designated as “world,” allowing them to develop some sense about how the child or educand sees it while simultaneously showing positive ways of relating to it.

The educator must continuously sense the way of trying to shape the educand’s relation to the world (through the educator-educand or the educator-world relations) that would make a desired improvement most probable. This is what defines the educator’s professional knowledge and ultimately what represents their ability to educate. Those who are most effective as pedagogues are able to change the relation of the educand with the world—whether it is the physical or human world or the world of the educand’s own mind and emotions.

Changing the relation between educand-world: didactics

The techniques that the educator can use to change the world-educand relation are known collectively as “didactics,” *didáctica*, *la didactique*, *Didaktik*. These all trace back to the Greek root to “show” or “point,” with didactics as a pedagogical topic or specialization receiving one of its earliest treatments in *Didactica Magna* (1657) by Comenius. “Didactics” can be interpreted very broadly or narrowly: as referring to all educating and learning, to practices of showing, pointing and clarifying in general, to matters of formal instructional content and technique, and to these matters specifically as they apply to one educational subject area (Gundem, 2000; Plaisance and Vergnaud, 1993). Following Künzli (2010) it is possible to understand the different sides of the pedagogical triangle as corresponding to different but mutually interpenetrating didactical emphases or approaches (i.e., “pedagogies” in the English sense): (1) The *educator-world* side corresponds to a “representational” approach to didactics; (2) the *pedagogical relation* to a communicative one; and (3) the *educand-world relation* to one that is experiential.

Emphasizing the educator-world relation: representation

The relation that the educator has to specific entities of the world serves as both a model and entryway for the educand to develop similar relations. Emphasizing the educator’s direct relation to the world, to actions, ideals, understandings and knowledge—in short, to the world as a whole—highlights their power to select, simplify, adapt and otherwise prepare aspects of the world for the one being educated. In the classroom, selection, simplification and adaptation is an important part of the expertise that constitutes teaching. Very often, though, it is not the world itself that is thus prepared, but *representations* of selected objects and processes within it—in the form of pictures, games, exercises, books and toys. These are often effectively “pre-adapted” and “pre-simplified” to be more easily integrated in the educand-world relation—shaped precisely to improve this relation in some particular way: building blocks for children’s hands and coordination, or books for their interests and reading level. “We long ago accepted that the realm of schooling—and of childhood in general—“consists of a huge montage of images and representations which are not ‘the things themselves’ but instead ‘point out’ things and phenomena,” as Klaus Mollenhauer puts it (1983/2014, pp. 34–35).

The process of adapting and simplifying the world—or representations thereof—has been referred to as the *transposición didáctica*, *la transposition didactique* (Chevallard, 1997), and the *didaktische Reduktion* (Lehner, 2020). Focusing on knowledge, Chevallard asks: “Why is a *transposition didactique* needed?” “Because the didactic function of knowledge is different than [its] scholarly function,” he answers, “because there are two regimes of knowledge [didactic and expert] that are interrelated but not superimposable” (as quoted in Makassikis, 2017, p. 99). Again, an educator is someone *not* necessarily in possession of subject-matter expertise, but of an ability to transpose and thus bring aspects of the world more capably and sensitively into relation with the educand. The transposition from expert knowledge to didactic knowledge can occur in a range of ways: through the use of examples that cover a broad range of cases (e.g., a bean sprout for all plant life), by focusing on part of a larger whole (e.g., a passage from a book), and by making knowledge appealing (e.g., through pictures, activities, mnemonic rhymes, etc.). This process extends to the person of the educator themselves—in Arendt’s sense of fulfilling a broadly “representative” role—in modeling the relation to the world that the educator wants to achieve for the child.

Emphasizing the educator-educand relation: communication

If the relation between educator and educand is emphasized over the other two sides of the pedagogical triangle, then the didactic approach becomes broadly communicative. Here, the student’s relation to others and to the world is to be improved precisely via collaborative negotiations with others. It may involve, for example, Socratic dialog between teacher and student, or collaboration within student groups. Whether educational groups are understood as “communities of practice” or as Pestalozzi’s domestic family, these members are seen as having the potential to reciprocally guide and educate one another.

Emphasizing the educand-world relation: experience

Emphasizing the educand's relation to the world entails an experiential approach, one that encourages direct engagement with and relation to the phenomena being studied. Such "educative experiences" as Dewey noted, are doubly important: Experiential education occurs precisely *through* the educand-world relation while pedagogy itself simultaneously seeks to improve this relation: "Hence the central problem of an education based upon experience," Dewey observes, is to foster "the kind of present experiences that live fruitfully and creatively in subsequent experiences" (1938, pp. 27–28).

A didactic reduction or transposition is implied in communicative and experiential approaches as well as in "representational" didactics. In these cases, this reduction may mean that communication with others is directed in such a way as to be positive and cooperative rather than conflicted, or that experiences are deliberately cultivated that reward certain types of action, knowledge and awareness, and discourage others. Although Rousseau characterizes Emile's education as "negative" in nature, he places both communicative and experiential approaches at the center of his novel of education. And as Meirieu points out, Rousseau's "cunning" in this regard consists not only "in organizing his pedagogy around the 'interest of the child:' but by making sure, thanks to skillfully calculated situations, that 'what interests [Emile]' and 'what is in his interest' converge" (Meirieu, 1996, pp. 56–57). For example, while with his master, Emile loses his way in a forest, becomes hungry and begins to cry. Comforting Emile, his master reminds him of a recent lesson in astronomy and points out the direction of the shadows in the forest. Through this didactic transposition or reduction of the complexity of his environment, Emile is able to deduce the direction to the town and to find his way, crying: "Let's have lunch! ... Let's run fast! Astronomy is good for something!" (1762/1979, p. 181).

This incident also highlights a second important way that the educator can influence the educand-world relation, namely through the *redirection* of the educand's will and desire. In this case the master redirects Emile's hunger and despair into an effort to deduce the direction of the town—so that they might ultimately enjoy a warm meal (Fig. 3). In fact, Schleiermacher describes all educational efforts, all efforts at changing the educand's relation to the world, as either representing a kind of "support" or "counteraction" of the educand's actions and inclinations: One supports the efforts of the person being educated to try new things (in the world), to become generous, cooperative or knowledgeable, all while counteracting certain reactions, dependencies or habits seen as less desirable. In all of this, the educator works to elicit and encourage the educand's spontaneous "self-activity," activity on the educand's own initiative and with its own objectives.



Fig. 3 Rousseau shows how a difficult situation can be didactically transposed and Emile's desire pedagogically redirected (Public Domain).

Despite the techniques and approaches available to change the educand-world relation, those being educated are fundamentally free in relation to any and all attempts to influence them. All that education can do, as a result, is to continue to transpose aspects of the world and redirect the educand's desire—shaping a context through words and objects—to make such change more, rather than less, likely.

The educand-world relation—*Bildung* and the educational telos

Given that education seeks to improve the educand's relationship to the world, pedagogy must be able to tell what is good or better in this relationship and what is not. This has traditionally been done by positing an ideal state for this relation and by seeing it as the final ideal or *telos* of the educand's development, indeed all of education itself. Throughout the centuries, one image or another of an all-embracing ideal person has offered a unifying framework for this telos, from the Greek *areté* (i.e., virtue) (Jaeger, 1939–46), *kalokagathia* (from *kalós* & *agathós*, i.e., “beautiful” and “good”) (Marrou, 1956) and *enkýklios paideia* (i.e., general education), through to the Roman *humanitas* (i.e., the human as moral and cultivated being) or Christian Godlikeness. The term *Bildung* itself first emerged from the idea of humans developing in the image (*Bild*) of God, with such ideals gradually being secularized through notions of “reason” and *civilitéé*. Since the Renaissance, this ideal became increasingly more open and less defined, just as the old *scientia* (i.e., knowledge one could possess), became the new “science” (a procedure which someone undertakes regardless of what is ultimately to be known; Gadamer, 1972). As the idea of a specific, external educational human *telos* has gradually faded away—a process seen as intrinsic to modernity itself (Blumenberg, 1966/1983)—the acquisition of a canon, of “true” knowledge of the world and life, has been replaced by various modes of relating to the world, namely ones that are seen as reasonable, critical, autonomous and authentic.

For example, in his classic definition of *Bildung*, Wilhelm von Humboldt (2000) describes the process as one that is intrinsically open—as nothing less than “the ultimate task of our existence:”

This [task] can be fulfilled only by the linking of the self to the *world* to achieve the most general, most animated, and most *unrestrained interplay* [emphasis added]. This alone is the yardstick by which each branch of human knowledge can be judged. For in each case, the only true path is that on which the eye is able to follow steady progress toward the final goal (1794/2000: 58–59).

In addition to the emphasis it places on the educand-world relation (i.e., “unrestrained interplay”), Humboldt's definition of *Bildung* is also noteworthy for its continued relevance in Northern and Central European pedagogy today: “Globalization and neo-liberalism,” as one historian observes, have recently only lent increasing “international support to *Bildung* as a concept based on higher principles and moral values that are not easily reduced to simple measures” (Horlacher, 2016, p. 118). *Bildung* itself has come to represent “an unquantifiable excess value” (p. 3) that signals both the dignity of ongoing collective and individual struggle and advancement as well as the potential transformation of the self through the other (Koller, 2012).

Pedagogy as a “science”: theory and art

The idea that pedagogy is not just a collection of concerns, arrangements and values, but that it constitutes a comprehensive theoretical account of a distinct “art” or practice, emerged in Northern Germany in the late 18th century. This occurred through the development of a distinctly pedagogical or “educational science”—an *Erziehungswissenschaft*—with *Wissenschaft* referring to any domain of rigorous research. This research work began with the anthologizing of texts (e.g., of Comenius, Locke, Rousseau) by a group of Enlightenment educators known as the “philanthropists” (i.e., lovers of humanity) and with Ernst Christian Trapp's 1780 *Attempt at a Pedagogics* (pedagogics being a rigorous attempt at a pedagogical theory). A few decades later, F.D.E. Schleiermacher developed his own *Outline of the Art of Education* on the basis, as he was quoted above as saying, of the relation of the older and younger generations. Schleiermacher begins these lectures, however, by asking how the “entirely general statements” of theory could apply to the particularities of pedagogical practice that had for years—millennia, really—preceded it:

Originally, it was parents who educated, and as is commonly acknowledged, they did not do so following a “theory.” And it is true for every domain that can be called an art in the narrow sense of the word that practice always precedes theory. Therefore, it would be incorrect to say that practice gains its character and specificity only through theory. The dignity of practice exists independently from theory. Theory only makes practice more conscious. Schleiermacher (1826/2022, p. 44).

The dignity, integrity and value of practice arises not just because it came first, but because it is always addressing the unique configurations of educator, educand and world—moments that must by definition escape “entirely general” theoretical statements. Practice in this sense designates an “art,” a *τέχνη* or *techné*, a practical knowing and doing that is variable and context dependent. Theory, correspondingly, is not based on empirical and psychological findings that would dictate optimal arrangements or

“best practices” for all situations. “Theory,” as Schleiermacher has already noted, “makes practice more conscious.” It poses questions about the all-but unavoidable dilemmas of education—questions of the “relationship between the older and younger” generations, of moments of present and future, of tactful distance and supportive proximity in the experience of the child (Kenklies, 2012).

These questions and dilemmas are not posed to be answered once and for all or to tell educators exactly what to do; they do not indicate which course of action would be right or wrong, better or best. Instead, they are meant to engage educators in similar questioning, for example, regarding habitual solutions, fashionable prescriptions, actual actions and possible alternatives. Such dilemmas and reflections—whether analyzed in academic discourse or in terms of immediate educational engagement—are seen as constitutive of pedagogy as a “science.” Instead of giving answers, the point of such theory is to “attune” practice, as Schleiermacher’s contemporary in pedagogical theory, J.F. Herbart (1776–1841) maintained. Like Schleiermacher, Herbart also began one of his most famous lectures in pedagogy by reflecting on the opposition of theory and practice. Herbart saw this as an opposition that could be reconciled, specifically through the addition of pedagogical tact as a mediating element:

But for every theorist who puts his theory into practice in particular cases ... a link intermediate between theory and practice involuntarily inserts itself. By this I mean a certain *tact*, a quick judgment and decision that is not habitual and eternally uniform Based on continuous practice, there inevitably develops in humans a mode of action that is dependent on feeling but that only remotely relies on certainty of belief (1802/2022, p. 32).

Like Nohl after him, pedagogical tact for Herbart is a kind of enacted disposition that provides the educator with a certain distance and also proximity in relation to the educand. Again, as reflected in the work of later educationists—from Nohl to Arendt—the basis of such sensitive engagement in pedagogical practice is a matter of one’s stance, frame of mind or attitude. And improving one’s engagement, as Herbart makes clear, is first of all a matter of reflectively preparing and cultivating this stance or disposition precisely through theory or “science.”

by reflection, reasoning, inquiry—in short, by science—the educator must prepare not his future action in individual cases as much as *himself, his disposition*. He must prepare his *head as well as his heart* to correctly receive, perceive, feel, and judge the phenomena awaiting him and the situation in which he will be placed. If he has, in anticipation, indulged in extensive plans, the practical circumstances will mock him (emphases added 1802/2022, p. 33).

The topic of pedagogical tact has recently experienced a Renaissance both in the German and French languages, with Herbart’s (and others’) characterizations being updated for contemporary theory and practice. For example, Burghardt and Zirfas (2019) open their recent book-length examination of *Pedagogical Tact* by characterizing this tact as marking “a central perspective in the field of tension between determinacy and indeterminacy in education, *Bildung* and instruction” (p. 9). It is a tension, they add, “that simply cannot be converted into systematic ‘determinacy’ or empirical ‘univocality’” (p. 9).

Ways of knowing: continental pedagogy and curriculum studies

In defining pedagogy as both a theory and a practice, and in understanding the two as connected through a reflective and responsive disposition, Herbart, Schleiermacher and others also give effective definition to pedagogy as a *way of knowing*. On the one hand, it represents an art, a practical way of knowing—an attuned awareness or disposition—that may not even reach to the level of theoretical articulation; on the other hand, it is a way of reflecting, a particular perspective on life, even a “theory.” Together, pedagogical theory and practice constitute a way of knowing that combines emotion, feeling and commitment together with a more reflective, awareness—a “*besonnenes Bewusstsein*,” a considerate consciousness, as Schleiermacher puts it.

Indeed, the configuration of the pedagogical triangle itself can be said to represent a “theory” of pedagogy. But this is not one that stands or falls on the basis of statistical generalizability or empirical evidence (e.g., how many times this configuration can be traced in a given classroom). It instead represents a model that allows us to identify education “out there” in the world, both in a theoretical and practical sense. Processes, persons, objects and arrangements need no longer be labeled “educational” simply because they happen to be found in places known as “school,” “college” or the “classroom.” Instead, this theory provides educational thought and research with an intrinsically human set of phenomena that constitute a clearly-bounded set or an object for their investigations. It is a curious but telling fact that most English-language reference works that claim to cover “education” overall (including this one) conspicuously fail to provide either a definition of education, or—as suggested earlier—explicit definitions of “pedagogy” as well (Friesen, 2019, 2021). Not only does Continental Pedagogy define such terms and their interrelation, it also identifies a timeline of historical texts in which their meaning has been refined and reaffirmed. As they have been briefly reviewed above, these sources show that pedagogy is not a set of prescriptions for the classroom, but rather, a set of relations that can arise in the widest variety of contexts, both real and imaginary—from Plato’s cave through Rousseau’s natural tutelage to Humboldt’s *Bildung*. Such examples illustrate that pedagogy not only represents a way of generating questions and identifying dilemmas that are specifically pedagogical and educational, but also that it offers a specifically *pedagogical* way of understanding not just education, but the world in which it

arises. As Gert Biesta has said, it offers “a distinctively *educational* perspective on education” (2015, p. 15)—one that goes well beyond educational institutions and embraces the world of human relation, influence and intention as a whole. It holds out the promise—as a recent French translation of Herbart’s pedagogical works puts it—that pedagogical thought is able to speak to “the overall contexts of pedagogy,” contexts which have long been neglected “in the framework of particular specializations or disciplines” (2007, cover).

Significantly, Herbart himself wrote about the dangers of “the science of education ... being conquered and ruled from afar,” precisely by other sciences or disciplines. Herbart argued that “it would be better if the science of education remained as true as possible to its own *intrinsic conceptions* and cultivated more an *independent mode of thought*” (1806/1908, p. 83; emphases added). In addition to illustrating how education can indeed cultivate this independence, this article has also attempted to introduce a number of the “conceptions” that Herbart saw as “intrinsic”—words like *educand* (*Zögling, le élève, alumno*), *Bildung*, pedagogical tact, and *la didactique*. We have further extended this vocabulary with ideas and practices like representation, self-activity, *l’éducation négative*, and the didactic transposition or reduction. Meanwhile, in the case of English language education, we can say that the colonization of education by other disciplines that Herbart warned of is now complete: The dominance of the social and psychological sciences, together with elements of philosophy and critical theory, is now so total in education as to have rendered it simply a “field” of phenomena to be studied only from *other* disciplinary perspectives (see: Bellmann, 2014, p. 277).

Curriculum Studies—whether contemporary or traditional—can also be said to accept today’s understanding of education as a “field” to be addressed by managerial, critical and other types of disciplines and theories. In an age of managerialism and standardized testing, its aforementioned concern with institutional forms has given contemporary Curriculum Studies a focus on the “catastrophic,” even “fascistic” results of corporate, governmental and administrative instrumentalities (Pinar, 2019, p. 124). In this context, contemporary curriculum theory has advocated for possibilities of renewal, for a new conception of curriculum that would function “as a bridge between the social and subjective, connecting as it reconstructs both” (p. 124). The result is that curriculum itself would be “enacted as [a] complicated conversation,” something that unfolds among adults—and is “intensely historical, political, racial, gendered, phenomenological, autobiographical, esthetic, theological, and international” in its nature (Pinar et al., 1995, p. 124, p. 847).

In this light, the contributions that Continental Pedagogy could make in mutual dialog with Curriculum Studies are manifold: First, it provides a non-managerial, non-psychological but nonetheless *practical* account of the pedagogical. As a part of a personal relation situated in the world and aimed at the educand becoming “who they are,” Continental Pedagogy conceptualizes both pedagogical action and reflection in ways that are irreducible to measures of efficiency or outcomes. Unlike many diagrams of educational systems and processes that are ubiquitous today, the pedagogical triangle does not map out any educational inputs our outputs, or any mechanisms for optimization or value generation. This triangle—and the theory it implies—instead portrays pedagogy as an ongoing engagement that is above all ethical and relational. And even if this type of relation and action may be subject to conditions characterized by some as catastrophic or fascistic, Schleiermacher and Arendt’s analyses suggest that there would still inhere within them a dignity and humanity that exceeds political categorization.

References

- Arendt, H., 1958/1968. The crisis in education. In: Arendt, H. (Ed.), *Between Past and Future*. Penguin, London, pp. 173–196.
- Autio, T., 2014. The internationalization of curriculum research. In: Pinar, W. (Ed.), *International Handbook of Curriculum Research*. Routledge, London and New York, pp. 17–31.
- Bellmann, J., 2014. Educational theory, nature of. In: Phillips, D.C. (Ed.), *Encyclopedia of Educational Theory and Philosophy*. Sage, Thousand Oaks, CA.
- Benner, D., Brüggem, F., 2004. *Geschichte der Pädagogik*. Reclam, Stuttgart.
- Bernfeld, S., 1925/1973. *Sisyphus or the Limits of Education*. University of California Press, Berkeley.
- Biesta, G., 2015. On the two cultures of educational research. *Eur. Educ. Res. J.* 14 (1), 11–22. <https://doi.org/10.1177/1474904114565162>.
- Blumenberg, H., 1983/1985. *The Legitimacy of the Modern Age*. MIT Press, Cambridge, MA.
- Briançon, M., 2019. *The Meaning of Otherness in Education*. Wiley, San Francisco.
- Burghardt, D., Zirfas, J., 2019. *Pädagogischer Takt*. Beltz/Juventa, Weinheim & Basel.
- Chevillard, Y., 1997. *La transposición didáctica: Del saber sabio al saber enseñado*. AIQUE, Buenos Aires.
- Comenius, J.A., 1657. *Didactica Magna*. Amsterdam.
- Comenius, J.A., 1658/1777. *Orbis Pictus or the Visible World*. S. Leacroft, London.
- Dewey, J., 1938. *Experience and Education*. Macmillan, New York.
- Friesen, N., 2017. The pedagogical relation past and present: experience, subjectivity and failure. *J. Curric. Stud.* 49 (6), 743–756. <https://doi.org/10.1080/00220272.2017.1320427>.
- Friesen, N., 2019. Educational research in America today. *Erziehungswissenschaft* 30 (59), 77–83. <https://doi.org/10.3224/ezw.v30i2.09>.
- Friesen, N., 2021. The necessity of translation in education: theory and practice. In: Jörnitz, S., Willmers, A. (Eds.), *Transatlantic Discourse in Education Research*. Buderich, Opladen.
- Gadamer, H.-G., 1972. The science of the life-world. In: Tymieniecka, A.-T. (Ed.), *Analectica Husserliana, The Later Husserl and the Idea of Phenomenology*. Springer, Dordrecht.
- Gundem, B.B., 2000. Understanding European didactics. In: Moon, B., Brown, S., Ben-Peretz, M. (Eds.), *Routledge International Companion to Education*. Routledge, London & New York, pp. 235–262.
- Herbart, J.F., 1802/2022. Introductory lecture to students in pedagogy. In: Friesen, N. (Ed.), *Tact & the Pedagogical Relation: An Introduction*. pp. 25–37, Peter Lang, New York.
- Herbart, J.F., 1806/1908. *Science of Education*. D.C. Heath & Co, Boston.
- Herbart, J.F., 2007. *Tact, autorité, expérience et sympathie en pédagogie*. Economica, Paris.
- Horlacher, R., 2016. *The Educated Subject and the German Concept of Bildung*. Routledge, New York & London.
- Houssaye, J., 2014. *Le triangle pédagogique, Les différentes facettes de la pédagogie*. ESF, Paris.
- Jaeger, W., 1939–46. *Paideia. The Ideals of Greek Culture*. Oxford University Press, Oxford.
- Kenkles, K., 2012. Educational theory as topological rhetoric. *Stud. Philos. Educ.* 31, 265–273.

- Klafki, W., 1958/1995. Didactic analysis as the core of preparation of instruction. *J. Curric. Stud.* 27 (1), 13–30.
- Koller, H.C., 2012. *Bildung anders denken*. Kohlhammer.
- Kudláčová, B., Rajský, A., 2019. *Education and Pädagogik*. Peter Lang, Bern.
- Künzli, R., 2010. German Didaktik: models of re-presentation, of intercourse, and of experience. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as Reflective Practice: The German Didaktik Tradition*. Routledge, London & New York, pp. 41–54.
- Lehner, M., 2020. *Didaktische Reduktion*, second ed. Haupt.
- Makassikis, M., 2017. *Apprendre l'orthographe française quant on est étudiant allophone*. Publibook, Paris.
- Marrou, H.I., 1956. *A History of Education in Antiquity*. Mentor Books, New York.
- Meirieu, P., 1996. *Frankenstein pedagogue*. ESF, Paris.
- Mollenhauer, K., 2014. *Forgotten Connections. On Culture and Upbringing*. Routledge, London & New York.
- Nohl, H., 1926/2022. Thoughts on the educational practice of the individual with special reference to the findings of Freud and Adler. In: Friesen, N. (Ed.), *Tact & the Pedagogical Relation: An Introduction*. pp. 61–73, Peter Lang, New York.
- Nohl, H., 1933/2022. The pedagogical relation and the formative community. In: Friesen, N. (Ed.), *Tact & the Pedagogical Relation: An Introduction*. pp. 75–84, Peter Lang, New York.
- Pestalozzi, J.H., 1799/2022. Letter to a friend on his work at Stans. In: Friesen, N. (Ed.), *Tact & the Pedagogical Relation: An Introduction*. pp. 15–23, Peter Lang, New York.
- Pinar, W., Reynolds, W.F., Slattery, P., Taubman, P.M., 1995. *Understanding Curriculum*. Peter Lang, New York.
- Pinar, W., 2014. Introduction. In: Pinar, W. (Ed.), *International Handbook of Curriculum Research*. Routledge, London & New York, pp. 17–31.
- Pinar, W., 2019. *What is Curriculum Theory?*, third ed. Routledge, London & New York.
- Plaisance, E., Vergnaud, G., 1993. *Les Sciences de l'éducation*. Editions La Découverte. Peter Lang, Berne.
- Plato, 375 BCE/1997. *Plato: Complete Works*. Hackett, Indianapolis.
- Ponte, P., Rönnerman, K., 2009. Pedagogy as human science, Bildung and action research: Swedish and Dutch reflections. *Educ. Action Res.* 17 (1), 155–167.
- Rousseau, J.J., 1762/1979. *Emile; or on Education*. W.W. Norton, New York.
- Schleiermacher, D.F.E., 1826/2022. Outlines of the Art of Education—Introductory Lecture—Selections. In: Friesen, N. (Ed.), *Tact & the Pedagogical Relation: An Introduction*. pp. 39–57, Peter Lang, New York.
- von Humboldt, W., 2000. Theory of Bildung. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as Reflective Practice: The German Didaktik Tradition*. Routledge, London & New York, pp. 41–54.

Curriculum and knowledge content

Daniel Sundberg, Faculty of Social Science, Linnaeus University, Växjö, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	256
Framing the issue of knowledge traditions and curriculum structure	256
Premises for knowledge divisions in curricula	257
The institutional curriculum: three basic knowledge concepts	258
Four curriculum ideologies and different views of knowledge in modern schooling	258
View of knowledge in a scientifically rational curriculum ideology	259
View of knowledge in learner-centered curriculum ideology	259
View of knowledge in social efficiency as curriculum ideology	259
View of knowledge in social reconstruction as curriculum ideology	259
The programmatic curriculum	260
Curriculum standardization	260
Knowledge and curriculum change: the case of Sweden	261
Framing the question of knowledge in the curriculum	262
Competing views of knowledge in the Swedish curriculum discourses	262
The textual curriculum: formulating the official view of knowledge in Sweden	263
Curriculum theorizing on the knowledge question	263
Conclusions	264
References	265

Introduction

It has been argued for a long time that today's information- and knowledge-intensive society, together with changes in globalization and values, is changing the conditions for education and schooling. New knowledge flows where information is available to everyone, which means that the accepted answers to what is valuable to know and learn in school are no longer self-evident. At the same time, most countries have a long tradition of nationally and culturally institutionalized curricula that formulate and structure the schools' knowledge content and forms. How the established school subject structure is influenced by international discourses about lifelong learning or 21st-century transversal competencies is an open question (Anderson-Levitt and Gardinier, 2021). This chapter addresses the question of knowledge in curricula and introduces theories, concepts, and research on how different orientations of the view of knowledge and their respective foundations in different traditions, such as scientific rationalism, social efficiency, and learning-centered and social reconstruction, have consequences for the schools' knowledge content and the organization of teaching in schools.

Three dimensions are crucial in dealing with these issues—the institutional, the programmatic, and the textual curriculum—which accordingly make up the outline of the text. First, after framing the general topic, the history of the curriculum structure (i.e., the institutional curriculum) is addressed. Usually, the so-called cultural grammar of schooling provides background assumptions that are taken for granted, such as the structure of the school subjects. However, knowledge traditions do change. Special attention is paid to research on how the view of knowledge can be related to different competing curriculum ideologies and displacements over time. Second, the programmatic curriculum is addressed, which includes the formal regulation of national curriculum making. With reference to the Swedish case of curriculum change, the features of the curriculum as a program are outlined. As many countries today apply different forms of knowledge in their curricula, for example, knowledge and competencies, the case study of Sweden highlights the international trend on curriculum standardization and convergence present in many countries worldwide. Third, the textual curriculum, how transversal key competencies are related to the curriculum as text, is tackled in the final part of this chapter and triggers induces the concluding discussion on theorizing the knowledge question of contemporary schooling.

Framing the issue of knowledge traditions and curriculum structure

Let us start with the most fundamental questions about curricula. What should be counted as knowledge? What selection of knowledge, values, and so on, should be made in compulsory curricula and syllabi? This is not a technical issue but one that involves epistemic and ideological, as well as cultural and political dimensions. When we discuss knowledge in curricula, a number of culturally, socially, and politically established premises need to be considered. It is neither a given that we can talk about *one* legacy of knowledge to be passed on from one generation to the next, nor is the institutionalized subject structure an unchanging entity.

Currently, we need to take a more open and pluralistic approach to different perspectives and contexts for the knowledge, as well as for the students' varied backgrounds. Interestingly, an opposite trend drives toward curriculum standardization in increasing comparability across nations and educational systems.

All societies are based on socialization and educational processes, where schools play an important role, but learning is also shaped by our own activities and close social interactions. Civic education enables participation in democracy and social life. However, there is not one but several different ideals and traditions around such socialization and educational processes. *Selective tradition* is a concept used in curriculum theorizing to show how schools knowledge content (e.g., in the form of a subject's predominantly taken-for-granted assumptions about what constitutes knowledge), how learning takes place, and what kind of teaching should be conducted, are based on certain traded perceptions (Englund, 2005). In both theory and practice, the curricula contain a normative dimension regarding knowledge selection, classification, and framing, based on human knowledge archives (Sundberg, 2021). Various forms of cultural canons, collective memories, social imaginaries, and political narratives form a horizon for all work with and decisions on curricula and syllabi in transnational, national, and local arenas.

In line with a linguistic turn in social theory and curriculum analysis, as demonstrated in the following text, several hidden premises have been placed in the spotlight. This applies to questions about why (the legitimacy of the content), what (the content and its organization), how the content is communicated in the school's teaching, when new content should be introduced (progression), and who is mandated to decide on these priorities. On one hand, the aims, goals, content, and forms of education and schools exist in the context of a constant struggle among different perceptions of what knowledge is, how teaching should be organized, what should be counted as valid knowledge, and which learning should be rewarded, historically configured by dominant curriculum ideologies (Englund et al., 2012). On the other hand, there is always a hidden dimension to culturally established norms that affect the design and content of curricula, which the sociological turn of education has clarified since the 1970s through the concept of the *hidden curriculum* (Jackson, 1968).

The cultural reproduction explanation model that emerged with the *new sociology of education* in the 1970s (i.e., the official school knowledge reproduces the dominant class power patterns in society) has been challenged and succinctly replaced by more dynamic models explaining the structure and knowledge content of schools. Along with the linguistic and discursive turn in social theory, the curriculum concerns not only the dominant *power over ideas* (i.e., the imposition of ideas and ideological reproductive gatekeeping) but also the *power in ideas* and the *power through ideas* (i.e., the ideational capacities, individual and collective, to influence normative and cognitive beliefs in communicative discourses) (Carstensen and Schmidt, 2016). Endogenous and exogenous drivers of change, continuities, and the discontinuous discursive controversies reconfigure the curriculum in the nexus of scientific, public and political, administrative, and professional arenas. However, the latent or background ideas that structure the forms and knowledge content of schools are highly formative.

Such an often-hidden dimension of schools' substantive organization concerns the division of subject domains into theoretical and practical subjects that can be traced back to Aristotle. It is a historically established hierarchical division that contributes with a kind of cognitive and institutional architecture (Deng and Luke, 2008). Tyack and Cuban call this phenomenon the cultural grammar of schooling for the knowledge content that remains strong in universities as well as schools in the Western world (1995). However, it is not a given, and in an information- and knowledge-intensive society, the ecologies and cycles of knowledge change, as we will see (later on in this chapter) reflected in the case of Swedish curriculum policymaking.

Premises for knowledge divisions in curricula

To understand the knowledge dimension of curricula, we need to, on one hand, address the discursive dimension and, on the other hand, acknowledge that the structure of knowledge is partly about deep intellectual epistemological traditions that can be traced to ancient Greeks, from Socrates (469–399 BC) and Plato (429–348 BC) to Aristotle (384–322 BC) and Alexander the Great (356–323 BC), to mention some key figures. Often, since then, construed by philosophers as an epistemological problem, the structure of knowledge invokes multiple ways of knowing, such as experimentation, empiricism, revelation, authority, rationality, intuition, meditation, and embodied knowing. In the classical curriculum, trivium and quadrivium comprised the basic architecture of the knowledge content for schooling. The word *trivium* comes from Latin and means the place where three roads meet. Originally, these consisted of grammar, rhetoric, and dialectics (Hamilton and Zufiaurre, 2013). In higher education, these were later complemented with four domains of knowledge (*quadrivium*): arithmetic (mathematics), geometry, music, and astronomy.

The modern curriculum conformed to tradition by retaining school subjects as the building blocks of the curriculum—although, to be sure, the modern subjects were not necessarily the same subjects or in the same form as those in the classical curriculum (Lundgren, 1984). Curriculum scholars have helpfully clarified such transformations. In the 1970s, Hirst (1974) extended R.S. Peters' (1919–2011) work in the analytic philosophy of education to address the issue of how the structures differed in various domains of knowledge (the disciplines). Hirst (1974) identified four characteristics of structures, each of which differed across knowledge domains. The core concepts varied across disciplines, as did the logical relations among them and the tests against experience, and the different forms of knowledge have developed unique techniques and skills. Based on these differences in their structures, Hirst (1974) identified seven forms of knowledge or disciplines: (1) mathematics, (2) physical sciences, (3) human sciences, (4) history, (5) religion, (6) literature, and (7) the fine arts.

Indeed, the realist stance that Hirst (1974) represents (i.e., school subjects mirror or correspond to basic objective realities) has been challenged. Joseph Schwab (1909–1988) called for an alternative position in structuring knowledge in the curriculum using

Deweyan lenses and by drawing on Aristotle's tripartite treatment of inquiry (theoretical, practical, and productive). He called for pragmatic discarding of moribund stable theoretical inquiry that focused on overgeneralized problems and empirical methods that sought lawlike principles of education. Instead, he advocated replacing it with the fluid, practical inquiry that would look for situational insights to interact with and remediate specific problems by engaging in productive eclectic arts of matching and tailoring theories to situational needs. Thereby, curriculum scholars should create a precedent to anticipate and generate alternative practices in order to enhance decision and action as social realities change and as we face new social challenges that confront education (Schwab, 1969, 1978).

The institutional curriculum: three basic knowledge concepts

There are profound traditions regarding the basic structures for the organization of schools' knowledge content; it is not a simple matter of either objective and enduring subject structures or problem-based and situational inquiry. In Deng and Luke's (2008) more recent thorough review of subject classifications in education and schools, they have identified three basic knowledge concepts (ideal types) that have structured schools' subject content worldwide and historically over time. (1) A disciplinary conception of knowledge constructs it in terms of canonical academic knowledge that embraces different traditions but revolves around on facts, concepts, and ideas that have been formulated and tested in scientific discourses over time. Here, pieces of reality are arranged and rationalized in hierarchies, categories, and taxonomies that systematically arrange different domains of knowledge. (2) A practical knowledge concept is action oriented toward certain practitioners and focuses on the use of knowledge around practical and socio-cultural problems. It can be about a set of procedures for successfully managing activities, as well as about considering different courses of action (metacognitive strategies) and embodied learning (compare, e.g., Aristotle's concept of *phronesis* with Pierre Bourdieu's concept of the logic of practice). (3) An experience-based knowledge concept focuses on social and cognitive elements in understanding and making phenomena in the everyday environment understandable and manageable. Here, the focus is on knowledge in everyday life and situations between the actor and his or her environment and between the subject and the object. From a pragmatic perspective (as in John Dewey's work), knowledge entails an ongoing construction of meaning that social actors create based on bio-social environments where both linguistic and semiotic interactions, as well as material artifacts and resources, are used for common dealings (Dewey, 1899/2009).

The different knowledge traditions have—to varying degrees—delivered premises for curricula. Most traditional definitions of knowledge for curricula tend to emphasize the former, cognitive, and fact-oriented subject or disciplinary conception. It is an uncommon procedure—although recommendable—that the question of knowledge and curriculum content is made part of public discourse. In connection with the work on the curriculum reforms in the 1990s in Sweden, a discussion was held on precisely the different dimensions of the concept of knowledge. The four forms of knowledge (facts, understanding, skills, and judgment) became ways to create space for different conceptions of knowledge, rather than just the epistemological and disciplinary one (Carlgren, 2009). However, the former, theoretical disciplinary conception has historically been dominant as a premise in the public understanding of and organizing principles for the construction of curricula.

Nonetheless, current changes in the curriculum landscape challenge such clear-cut knowledge divisions. In the discussions among curriculum researchers, the distinctions between cognitive and non-cognitive, theoretical and practical, or formal and informal education comprise the manipulation of those symbolic resources (words, numbers, pictures, etc.), which point to external demands for skills in changing labor markets, challenged democracies, cultural pluralism, and so on. Competency-based education and skill-based knowledge emphasize socio-cognitive and procedural knowledge (rather than cognition and declarative knowledge), along with the active inquiry of the learning in defining the learning objects. A purely skill-based competency curriculum that underpins an international comparative system of testing, such as the Program for International Student Assessment (PISA) (Organization for Economic Co-operation and Development [OECD], 2019), would mistakenly reject this distinction and its designation as an epistemic activity. Superficially, the PISA literacy concept attempts to take the critique of the separation of formal and informal education into account by stressing the importance of basic competencies to meet challenges for learning in real-world situations. However, there are different ways in which cross-disciplinary competencies are added on, are integrated, or transform subject structures of national curriculum frameworks (Nordin and Sundberg, 2021).

Distinguishing between knowledge of how to do something (procedural, practical, and experiential) and knowledge of something (declarative and disciplinary) is important, but both are, in essence, knowledge-making activities. For example, dispositional knowledge is one concept that transcends such a division and refers to relatively stable habits of mind and body, sensitivities to the occasion, and participation repertoires. The three types of knowledge thus have different forms in their original states and as a result, various pedagogical structures and expressive or performative modes. They can only be assessed functionally in relation to their different internal relations, that is, there should be diverse ways of assessing or evaluating them. However, as argued by Bellmann (2010), the PISA literacy concept as functional for the real-life situations outside schools borrows from the language of progressivism; still, there are potential contradictions in aligning generic skills with disciplinary school knowledge.

Four curriculum ideologies and different views of knowledge in modern schooling

To understand how curriculum content and knowledge are changing, a realistic account of the various ideological forces at play is necessary (Sundberg, 2021), although this dimension usually goes under the radar of policymakers. In international curriculum

research, four different dominant historical curriculum orientations are identified, linked to the emergence of modern schooling (see, e.g., Kelly, 2009; Schiro, 2013). They represent various systems of ideas, culturally established assumptions and ideas, which in diverse ways, form the basis for curriculum constructions and knowledge conceptualizations. Let us look at the view of knowledge within these.

View of knowledge in a scientifically rational curriculum ideology

This orientation focuses on and emphasizes the transfer of disciplinary subject knowledge in order to develop the learner's intellectual capacities and reproduce a cultural tradition based on such basic knowledge and common frames of reference. Academic disciplines here constitute authoritative resources through which the school subjects also gain their legitimacy. Essentially, it is a curriculum of knowledge, techniques, and ways of knowing that is passed on. Many curricula have strongly relied on a traditional (essentialistic) subject structure, what has been called *the cognitive architecture of modern curricula*. A content-oriented curriculum usually provides concrete instructions for the content that students must encounter in the teaching, in terms of subject matter and work and teaching methods. The emphasis is on the substance to be conveyed. In some countries, canon lists are also made, indicating cultural frames of reference in different areas. An underlying assumption is that the content conveyed by the teaching is also what the students learn (i.e., the impact premise). In the international curriculum discourse of the 1990s, this disciplinary notion of school knowledge was challenged as more learner-centered ideas of active participating became influential, with the so-called first wave of competence-based curricula (Sundberg, 2021).

View of knowledge in learner-centered curriculum ideology

A curriculum orientation toward a liberal humanistic direction is driven by the focus on the individual's learning in what should constitute central school knowledge content. The school should foster personal development, self-realization, innovation, and creativity. Disciplinary knowledge is relevant as long as it contributes to the development process and self-realization of the individual. A curriculum can consequently contain a wide repertoire of knowledge, experiences, and resources that can have a liberating effect. Historically, pedagogical progressivism (and reform pedagogy) has been a movement against this type of curriculum construction. In recent international curriculum development, there are also examples of how this ideology has driven competency-oriented curricula (Englund, 1996). In recent decades, many countries have focused on identifying and systematizing competency goals for students (Anderson-Levitt and Gardinier, 2021). Different kinds of generic abilities are formulated, which, among other things, are about learning to master different practices and problems by distinguishing, assessing, resolving real-life issues, and so on. They are usually designed according to goal and result management principles, something that content- and process-oriented curricula do not have to be and have not been. In many countries, including Sweden, this curriculum ideology was highly influential in the restructuring of curricula and syllabi in the 1990s. At the same, the pressures for education efficiency and measurable outcomes framed the implementation and enactment of the competency-based curriculum in practice.

View of knowledge in social efficiency as curriculum ideology

This curriculum orientation stresses the importance of equipping prospective citizens with the necessary skills, knowledge, and cultural capital for economic and societal productivity and growth. The questions about curricula are related in this way to questions about strengthening human capital. The benefit, that is, the functional value of the knowledge, becomes central to the strengthening. Subject knowledge is linked in this way to the various skills required in working life. In a post-industrial society, particularly those generic competencies that can be used in different spheres (social and economic) have formally and informally come to the fore.

The result-oriented curriculum, which is an expression of this curriculum ideology, is usually a combination of a content-based, a process-oriented, or a competence-focused curriculum. However, the result-oriented curriculum is designed so that it also states and specifies what the student should have learned, relating to different specific minimum levels based on the so-called performance standards (Kelly, 2009). By linking content requirement levels (referred to as alignment in the research literature) with assessment and grading criteria, results in the curriculum and in the student's learning are focused on in a distinctive way (Meyer and Benavot, 2013). In many countries, curriculum reform gradually became more result-oriented as the pressures on the competitiveness of educational systems in the new knowledge economy pushed for efficiency, according to international performance indicators.

View of knowledge in social reconstruction as curriculum ideology

This conception emphasizes the sociocultural and sociopolitical context rather than the individual in curriculum decisions. Education and schooling are important tools in a societal reform toward increased equality, justice, and equity. In accordance with this orientation, a curriculum also needs to contain a study of societal contexts, issues, and possible futures to enable the individual to develop one's own positions and future paths. Subject knowledge then becomes a set of more diverse activities that contribute to meaningful learning experiences for one's own agency. Different value dimensions become central for the curricula constructors and developers to consider. Process-oriented curricula are focused more on students' learning experiences and learning processes and less on predetermined content. They emphasize characteristics of the development of students' learning and thus also cover skills and attitudes. The curriculum is a means of trying to control the various processes involved in education and teaching. Content and

methods are only aids to achieve desirable development processes at the individual, the group, or the societal level. Thus, the focus is on societal challenges, working methods, and cognitive tools to broaden students' experiences and insights, rather than on subjects or a given body of knowledge. Despite its historically strong influences in the curriculum reforms in the Nordic countries, this curriculum ideology has clearly become marginalized over the last decades.

The programmatic curriculum

Regarding the history of curricula, there is a need to acknowledge the differences between the North Continental and the Anglo-Saxon curriculum traditions, as they have largely structured how curricula have been constructed (Gundem and Hopmann, 1998). The Continental tradition has a history dating back to the 17th century and Johann Amos Comenius. He was behind what has become the prototype for modern curricula, *Didactica Magna* (1657/1999). Some distinctive features of this entire North Continental curriculum and *Didaktik* tradition are that they have been philosophically oriented and the theorizing have built on conceptual models, with a focus on schools' social mission and knowledge content. The Anglo-Saxon curriculum tradition does not have the same long history. For example, in the USA and England, curriculum issues have to a greater extent been practical and empirical. Taken together, however, both traditions have also been based on some common starting points. Hopmann (1999) stresses that the underlying premises are based on the following premises:

1. The required knowledge, attitudes, and abilities are known and can be specified in curricula.
2. What is specified can also be taught.
3. What can be learned through teaching (i.e., its actual outcome) also corresponds to the desired outcome.

Hopmann describes them, respectively, as the *planning*, *learning*, and *impact* premises for curricula (1999). However, none of these can be secured in a school's history as empirically substantiated. They are assumptions that must be made but whose uncertainties are always in play. They represent a kind of hidden action theory that governs the work with curricula. They are general and deeply institutionalized but at the same time, raise questions about the power to define the school's knowledge content and about political, expert-based, and professional control in curriculum-making processes at various levels and arenas. On one hand, the democratic premise says that there should ultimately be democratic decision making behind the curriculum, based on scientific expertise. On the other hand, a *living curriculum* presupposes that professionals perceive ownership of the taught curriculum.

However, the space for professional curriculum making is restricted and framed as workable organizational frameworks for the delivery of schooling. The programmatic curriculum consists of the regulation of formalized curriculum making and national political governance of the goals and content of the educational system (Westbury, 2008) as part of the political jurisdiction and as an instrument for steering the school system. The curriculum as part of a steering program does not only consist of specific document goals, content specifications, assessment criteria, and so on, but it also includes the interrelated programs of in-service education, textbooks, guidelines, the assessment system, and so on. The programmatic curriculum include "routine educational and administrative activity of many ministries and boards of education as they support and ratify ongoing revisions and changes in subject areas and courses of study" (Sivesind and Westbury, 2016, p. 744).

The programmification of the curriculum thus results in specific configurations of content selection, as well as emphasizes and prioritizes what, how, and when content should be taught. Subjects are building blocks for the routinization of the inner work of schools. The administration of the school system and the management in the schools are organized around the subject structures. When comparative scholars refer to an ongoing curriculum convergence globally, it concerns these textual and programmatic dimensions of the curriculum. In many educational systems across the world, a kind of curriculum programmification is observed as loosely coupled systems are replaced by an effort to align goals and content according to performance standards and tests. In some countries, mainly Anglo-Saxon, this results in a highly prescriptive, detailed, and scripted teacher-proof curriculum, with guidelines for teachers to follow in order to climb in the rankings of international achievements tests such as PISA.

Curriculum standardization

Although the curriculum is a contested issue at the very center of education and schools today, a slow global convergence is discernible, which mostly concerns the programmatic and technical aspects of the curriculum. Meyer et al. (1992) refer to the modern curriculum as a tamed and professionalized curriculum. Global standardization of school subjects can be identified over the last 100 years. This does not merely involve curriculum rhetoric and terminological categories. Comparative analysis reveals increasing convergence, homogeneity, and standardization in curriculum systems, policy programs, and assessment systems (Meyer and Rowan, 2006). While a general common-sense understanding and a general definition of education standards are the expected levels of achievement for all children of a particular age, it is vital to place them in the historical context of curriculum making in order to understand the emergence of such standards in education. Kliebard (2004) specifically refers to scientific curriculum making and the rise of social efficiency as an educational ideal in the early decades of the 20th century. The allure of having a clear and coherent system for teaching, learning, and assessment is closely related to the scientific outlook. "Standards are presented as a rational and scientific way of making and controlling a curriculum" (Waldow, 2015, p. 59).

Standards are usually represented as technical objects in contemporary educational discourses (i.e., fixed, stable, and universal). However, a closer critical investigation indicates the contested nature and ambiguities inherent in standardization processes. First, conceptualizing standards as epistemic objects means that they are defined by a struggle among different power alliances (i.e., standards are carriers of implicit normativity). Second, presenting them in a uniform way across contexts underestimates the dynamics of standardization by different actors and institutions in translating them into norms and actions (Brunsson and Jacobsson, 2000). Broadly speaking, regarding education, *content standards* include prescribed goals and curriculum knowledge (input), and *performance standards* define levels of student achievement (outcomes). In this chapter, we further explore the gradual shift in curriculum discourses, from content to performance standards, that has occurred from the late 1960s to the present day. In general terms, standardization can be defined as a process of building uniformities in time and space through agreed rules (Bowker and Star, 1999).

Examining modern educational systems today shows that standards proliferate and are particularly relevant when thinking about regulating curricula beyond national boundaries (Djelic and Sahlin-Andersson, 2006). In a broad sense, standards can be defined as constituting a particular type of rules—usually voluntary rules that are explicitly formulated, pertaining to a wide range of actors (individuals or organizations)—that are recommended to be adopted in order to allow comparisons across contexts. Indeed, a variety of organizations are involved in developing, sustaining, and implementing curriculum standards (content and performance), including, but not limited to, standard makers, adopters, monitoring and certification agencies, and the larger public. Standards and standardization are often addressed as part of the broader discussion on organizational regulation. Regulation involves creating and propagating more or less explicit rules and thus fosters the formation of social order. Thus, studying standards allows us to consider both the *demand side* of order (i.e., how organizations and individuals are affected by organizing efforts) and the *supply side* (i.e., how organizing elements are produced).

Despite their pervasiveness and significance in modern life, standards have received comparatively little serious attention from curriculum researchers. It is only over the last few years when researchers started to systematically explore the social dynamics of standardization in education—the dynamics involved in standard development, standard setting, standard compliance, and standard enforcement. An example is the case of key competencies and their associated performance indicators promoted by the OECD (2019). While standards are often associated with rigidity and formalism, a closer study reveals standardization as a highly dynamic process involving multiple actors (individuals and organizations) and often fostering creativity and innovation on the side of both standard setters and followers. Curriculum scholars have lately directed attention to how actors (scientists, experts, bureaucrats, etc.) deploy techniques to mobilize other actors and to govern at a distance, beyond local spaces, and via calculative devices and translation processes into new patterns of actions (Bowker and Star, 1999; Brunsson and Jacobsson, 2000).

One main topic among researchers and policymakers in the field of curriculum and instruction concerns such processes of standardization and its benefits and risks. Curriculum coherence and curriculum alignment are central concepts in addressing not only the external but also the internal factors involved in curriculum standardization. The notion of curriculum coherence highlights the links between curriculum aims and content, teaching practices, and assessment. Some findings have suggested that a high level of curriculum coherence is associated with high-performing systems (e.g., Oates, 2011). Curriculum alignment refers to the process of ensuring consistency between the intended curriculum, as stated in its aims and objectives, and different learning and assessment activities in classrooms. Nonetheless, many studies have shown that curriculum content (whether we refer here to its essence, its legitimacy, or its genealogy) is contested and thus requires choices to be made among different formulations, conceptions, and arrangements (Englund, 1996; Sundberg and Wahlström, 2012). In turn, this has implications for the types of pedagogy that can be employed and the types of evaluative procedures that should be adopted. The Swedish case of the recontextualization of its curriculum and the enactment by teachers and schools shows that this is far from a univocal linear process of implementation.

Knowledge and curriculum change: the case of Sweden

Sweden is a good example of a country that is in the middle of two fundamentally different ways of governing with curricula, in the field of tension that Hopmann (2003, p. 460) has described as “content and process control” and “outcome and product control.” Sweden has had strong cultural and educational links to the North Continental European tradition, especially the German Didaktik tradition. This is characterized by content and process control, a licensing model for teachers, professional self-control, and self-evaluation (Sivesind et al., 2012). The political leadership has trusted that the professionals will handle and carry out their assignments internally. However, a new type of management philosophy of more product control has quickly gained ground in recent decades. It is characterized by a focus on outcomes and efficiency and is largely based on an Anglo-Saxon curriculum tradition (Gundem and Hopmann, 1998).

On a general level, the curriculum as a system of ideas (i.e., the institutional curriculum) has also been pushed into the background under the international pressures of curriculum standardization, along with an Anglo-Saxon framing, leaving the knowledge heritage unexamined. Instead, it is the curriculum as a program that has come to the fore in contemporary educational policy in Sweden, as well as in many other countries. It is about building a goal management system that is aligned and operationally

efficient. A look inside the work with curricula makes it clear that there are far from rational and well-planned processes behind new curricula, without them also being governed by different pragmatic and contextual conditions. In an important research study from within the work with the curricula in the early 1990s, the scholar Ingrid Carlgren described the curriculum work as a form of practical problem solving rather than a rational deliberation on governing principles and ideologies:

The curriculum proposals (that were considered within the 1993 curriculum commission) did not come about as compromises in relation to substantive and ideological discussions, but from an interaction between the factors framing the process, incidental events, and practical considerations. Part of understanding curriculum change is, therefore, to understand what problems there are to solve (1995, p. 412).

In research on curriculum work at the system level, it is now well established that there is a categorical simplification of the complex task involved in constructing a curriculum. An administrative simplification reduces different perspectives and discourses about a school's mission, subject content, and forms. Haft and Hopmann (1990, p. 25) have introduced the term "compartmentalization" for such a simplification. It is then usually a matter of adjusting within an existing system rather than changing the whole curriculum with a new one, which could threaten its legitimacy and stability. The work of organizing knowledge for curricula thus includes several different activities and processes at different levels. At a more societal level, there are societal and cultural ideologies and symbols, politics and organized interest groups, various organizational and political structures and processes, as well as local understandings and practices. Research shows that given the complexity of curriculum issues and ideological value loads, they are often reduced to technical and practical problem solving based on presumed premises (Westbury, 2008).

Framing the question of knowledge in the curriculum

The curriculum's view of knowledge has in many respects been strongly embedded in cultural traditions (i.e., it has been implicit). In connection with the radical Swedish school reforms of the 1990s, the concept of knowledge became more explicit, discussed, and re-thought. The directives for the Curriculum Committee stated that its work should be formulated on the basis of a view of knowledge that "shall rest on current research on knowledge and learning," that is, not on traditions and common sense (Curriculum Committee, 1994, p. 7; see also Carlgren, 1995). The idea was that knowledge and the schools' subject content could no longer be taken for granted but that the knowledge issues had to be dealt with by the schools' actors, not least the teachers themselves. Instead of a centrally prescribed view of knowledge, it was thought and required that at the local level, there would be conversations about different types of knowledge and how they would relate to one another and to each school's goals. The report *School for Education* (Curriculum Committee, 1994) emphasizes that the pedagogical professionals should be assigned the responsibility for developing a reflexive view on knowledge, learning, and teaching on which the activities are based, that is, stressing the continuity of the historically established licensing model and division of labor between politicians and education professionals. It is "the interplay between the teachers' theoretical ideas and the practical design of the activities that should be the basis for an increasingly reflected practice" (Curriculum Committee, 1994, p. 62).

According to the Curriculum Committee (1994), the previously implicit normative notions of how the schools' knowledge content and forms should be structured would be replaced by explicit knowledge discussions for which the schools themselves were responsible based on the various contextual conditions. The goals in the curricula would not govern directly but indirectly through teachers' interpretations and reflections. At the same time as the so-called participatory goal management opened up for extensive professional decision making on curriculum issues, the management also became driven by goals and results. The previous rule management that prescribed the schools' subject content and organization would be replaced by an outcome control. The state formulated an assignment, and the schools were expected to deliver the expected results. The licensing model of professional curriculum making was thereby accompanied with a new form of contract to deliver according to the prescribed performance indicators.

Competing views of knowledge in the Swedish curriculum discourses

The Swedish (and Nordic) curricula and knowledge traditions constitute an interesting case internationally. The Swedish school has been part of the model of the modern welfare state since the 1950–1960s. It has been founded on a political premise of redistribution. Within the framework of a common school, the school must function to compensate for different life opportunities. Concepts such as a school for all, equality, and diversity set certain standards for teaching and learning. They are also anchored in philosophical and moral ideas about the purpose and meaning of school and teaching, ideologically related to social reconstructionism. Altogether, since the 1960s, they have provided certain frames of reference or background understandings among various curriculum stakeholders when new curricula have to be produced (Wahlström and Sundberg, 2017).

Curricula and syllabi are always based on certain notions of knowledge and learning, something that can be formulated in the different types of curriculum ideologies. The question of what view of knowledge lies behind a curriculum construction has always been important at the same time as it has received varying degrees of attention and debate. It is not only about what is to be counted as important knowledge but also about the nature and significance of the knowledge for teaching and learning. When we discuss the view of knowledge in curricula, both explicit formulations in the governing documents (curricula, syllabi, commentary materials,

advice and guidelines, grading criteria, etc.) and the implicit, background ideas and assumptions should be included. For example, a scientifically rational curriculum ideology was a basic feature of the welfare model and the systemic reforms in the 1950s and the 1960s in Sweden. Gradually, it became an institutionalized background assumption in curriculum discourses until the 1990s.

However, throughout the 20th century, a traditional organization, in the sense of the teacher-centered communication of schools' subject content based on a scientifically rational principle and an essentialist curriculum ideology, had been challenged from various quarters. Not least, Ellen Key (1900/1996) contributed in Sweden with the *Century of the Child* to direct criticism at superficial factual knowledge transfer and the one-sided division of subjects and placed the learning student's entire development in the center instead. A traditionalist and essentialist versus a progressivist and learner-centered view of knowledge has taken opposite sides in an educational policy field of tensions. The dichotomies between teacher- and learner-centered pedagogies and between teacher control and pupil autonomy have impeded attempts to understand and address complex educational problems and have thwarted productive discussions among policymakers, researchers, professionals, and the public. The tensions were accentuated when competency-based curricula became a central theme of the curriculum discourse of the 1990s in Sweden, which was partly backed up by historical-background assumptions of a state or administrative progressivism, along with a learner-centered curriculum ideology.

The textual curriculum: formulating the official view of knowledge in Sweden

Instead of prescribing a certain view of knowledge from national educational agencies, the concept of knowledge began to be problematized and discussed in the national curriculum discourse of the 1990s. Not least, knowledge acquisition in terms of learning participation became a starting point for the discussions (instead of learning as memorizing). In line with the critique of Piaget and individual constructivism, the reintroduction of Vygotsky and the Russian cultural-historical theory formation came to form an important foundation (Adolfsson, 2013). Here was a critical point toward traditional knowledge assumptions, technical rationality, as well as dualisms that characterized the view of knowledge in curricula: thought-action, cognitive-social, empirical-normative, practical-theoretical, and so on. A traditional view of knowledge as fixed knowledge to be transferred from teachers to students in a given order and pace was challenged from several different directions, such as pragmatism, sociocultural theory, and phenomenographic research, with the context of knowledge and different values emphasized instead (Carlgren et al., 2009). Carlgren (2009, p. 20) summarizes the view of knowledge in the new curriculum in three categories, as follows:

The basis of knowledge

- The knowledge is constructed by people in certain contexts for certain purposes (i.e., it is intentional, contextual, and functional in nature).
- The knowledge is carried by humans but is also found in cultural artifacts (texts, tools, and languages).

The form of knowledge

- The knowledge is expressed in the form of facts, understanding, skills, and familiarity, which correspond to knowledge, such as information, practice, interpretation, and judgment.

The content of knowledge

- School subjects are not formulated primarily in terms of material properties but as subject-specific qualities (expressed in the aim of striving for and the basis for assessment).

Such a view of knowledge as formulated here means that clear boundaries between theoretical and practical subjects have become blurred. All subjects are assumed to contain both theoretical and practical forms of knowledge. The hierarchy from theoretical to practical and down to less valuable subjects is also dissolved, as is the hierarchy spanning the four forms of knowledge. Instead, they are perceived as forming different but necessary aspects of knowledge. Formulations such as knowledge and skills were removed from national curricula and syllabi, as the concept of knowledge also contains skills. The view of knowledge expressed in the curricula from 1994 onward is also invoked to apply in the follow-up from 2011 and the present curriculum reform of 2022 (Official Swedish Government Report, 2007; Swedish National Agency for Education [NAE], 2022, 2011). However, the questions that arise are to what extent this view of knowledge is the dominant one in today's curriculum discourse, and what impact it has had on the implementation of the curriculum reform. Empirical research has shown that a traditional subject-centered view of knowledge, in combination with the instrumental values of knowledge, has become dominant (Wahlström and Sundberg, 2017), which has also been reinforced by the pressures on the competitiveness of educational systems pushed for efficiency, according to international performance indicators. The case of Sweden concerning the knowledge question in the curriculum makes it clear that institutionalized background ideas and subject structures provide a grammar of schooling that is difficult to change.

Curriculum theorizing on the knowledge question

Internationally, there is an intensive discussion on the question of knowledge in the field of curriculum theory (Wyse et al., 2015). Young (2008), in particular, has taken up a new discussion on curriculum knowledge over the last decade. In a general analysis of how the knowledge economy is about the conditions for education and schools, he believes that the schools' knowledge content has

been subordinated to various external objectives regarding productivity, growth, employability, and so on. Qualification and employability in emerging labor markets have created a management mindset around curricula centered on various performance indicators. His critique is that standardization of performance criteria and generic skills denies students access to what the disciplines can provide in the form of powerful knowledge (access to powerful knowledge). This social realist position thus criticizes the curriculum ideology of social efficiency from within the scientifically rational curriculum ideology by stressing disciplinary knowledge as an end in itself (intrinsic value), not as a means for economic and societal productivity (extrinsic value).

Scott (2014) discusses the question of knowledge in curricula based on what kind of knowledge is sought, the central principles of how it should be formed, and how it can be legitimized. He believes that these are necessary—although vast and difficult—issues to work with around the curricula. Here, consensus and hidden premises are major and difficult obstacles. He advocates a visualization and an identification of different curriculum ideologies that work so that hidden premises become the subject of discussion and deliberation. In what he calls essentialism (foundationalism), instrumentalism, and pragmatism, there are critical aspects that need to be considered to be able to construct solid and sustainable curricula. An essentialist (scientifically rational) conception can become too internally preoccupied, the instrumental conception can become too focused on external goals, and a pragmatic conception can be problematic due to its preoccupation with social goals.

Recent curriculum research clearly indicates that contemporary curriculum changes must be analyzed on the basis of the global processes driven by powerful transnational actors and national and local contexts with their historical, cultural, and political conditions in order to avoid simplified and/or overestimated analyses (Nordin and Sundberg, 2016; Sundberg and Wahlström, 2012). In recent decades, transnational policy actors have become important premise suppliers for curriculum work, as education and economics have become increasingly intertwined activities in the knowledge-intensive society. Young (2008) refers to a technical–instrumental orientation to the curriculum where knowledge itself is not valued. Instead, it is the knowledge as an instrument for other external goals that has been a non-questionable point of departure in education policymaking.

The Swedish case of curriculum change illustrates the influence of transnational education policy movements, the significance of previous Swedish curriculum reforms, and the reforms' impact on teachers in terms of their way of organizing teaching and assessing students' knowledge development. Closer analyses show a fundamental shift in the curriculum design that can be attributed to a transnational influence in terms of standard-based curricula characterized by a clear link across goals, content, and results. It represents a kind of denationalized and instrumental educational thinking created by two international influences—a technical–instrumental curriculum discourse focusing on results, and a scientifically rational curriculum discourse on given subject knowledge. Not least, the standardized knowledge requirements focus on clear (summative) assessments of the degree of achievement, which also affects the choices of teaching content and of teaching methods, as teachers' work with assessment and grading. Teachers' professional agency and curriculum work have become more focused on the application of the curriculum (the curriculum as a matter of facts for transmission) rather than on further developing content and forms together with students (the curriculum as a common matter of concern (Wahlström, 2022)).

Conclusions

The questions on knowledge and the curriculum are enduring concerns. In this chapter, key concepts, distinctions, and developments that frame the international discourse have been addressed. Three fundamental dimensions of the curriculum have been illustrated—the institutional, the programmatic, and the textual curriculum concerning the question of knowledge. Given the present situation with international comparisons and standardization in curriculum policy, these provide critical observanda. There is much debate today between subject-matter specialists in curriculum and instruction and curriculum generalists. While subject specialists hold that knowledge is structured differently in each discipline or field, generalists think that foundational considerations must address knowledge from myriad realms of human endeavor and not epistemology alone. Thus, it may be correct to conclude that the structures of knowledge conversation remain unsettled, enabling educators to keep alive basic questions of what is worth knowing in fast-changing knowledge economies.

The knowledge questions about the curriculum designs (i.e., as policy programs) are highly topical in most countries. A strong transnational policy trend in the curriculum area that contributes to this is the so-called key competencies of the future or the capabilities of the 21st century. They emphasize competence, skills, and learning outcomes in favor of knowledge content (see, e.g., Nordin and Sundberg, 2016). Powerful political actors, such as the OECD and the European Commission (EC), have come to advocate a competency-based curriculum and soft governance, together with increasing policy coordination, which have a significant influence on national curriculum discourses and curriculum models.

Whether the selection, organization, and assessment of knowledge or the institutionalized subject structure will change remains an open question. The dominant and traditional premises for organizing the schools' knowledge content are based on an additive logic. When different societal changes call for changes or when new subject areas are added within the various subjects' knowledge development (i.e., within research and the subject disciplines), new content is usually added to the existing subject structure. A historical review clearly shows significant stability regarding the subject structure in Swedish curricula (Sundberg, 2021). Since 1842 (when compulsory schooling was introduced), subjects such as mathematics, the mother tongue, history, geography, religion, music, biology, and drawing have made up the cognitive architecture. Since the 1960s, social studies, foreign languages (English, German, and French), gymnastics, physics, chemistry, and technology have made their way into the Swedish curricula. The only new subject since 1980 is the mother tongue for immigrants. This clearly indicates that a scientifically rational conception of knowledge

(essentialism) is strongly institutionalized within the educational system and constitutes a guiding premise in curriculum work (Sundberg, 2021).

However, in many countries worldwide, curriculum reforms are where subject renewal and integrative principles of content organization are tested. Given the rapid development of knowledge and new societal challenges, it has been argued that new selection principles and opportunities to integrate new content with subjects in new subject constellations will likely become increasingly important to avoid ending up in an overloaded curriculum preparing for tests and where new content is constantly added without existing content disappearing. Sustainable development is an example of such an area that is given an interdisciplinary character and more projected space in the curriculum. In this shift, a learning-oriented curriculum ideology has, above all, been a driving force. Social reconstructionism as a curriculum ideology addresses the future challenges and values at stake in setting the direction for future education.

References

- Adolfsson, C.-H., 2013. The Question of Knowledge: A Curriculum Theoretical Study of the Swedish Upper Secondary School's Reforms Between the 1960s and the 2010s. Linnaeus University, Diss. Växjö.
- Anderson-Levitt, K., Gardinier, M.P., 2021. Introduction contextualising global flows of competency-based education: polysemy, hybridity and silences. *Comp. Educ.* 57 (1), 1–18. <https://doi.org/10.1080/03050068.2020.1852719>.
- Bellmann, J., 2010. Deweyan pragmatism as the philosophy of PISA? On the language of progressivism and its adaptability. In: Tröhler, D., Schlag, T., Osterwalder, F. (Eds.), *Pragmatism and Modernities*. Sense Publishers, Amsterdam, pp. 191–207. <https://doi.org/10.1163/9789460913457>.
- Bowker, G.C., Star, S.L., 1999. *Sorting Things Out: Classification and Its Consequences*. MIT Press, Cambridge, MA.
- Brunsson, N., Jacobsson, B., 2000. *A World of Standards*. Oxford University Press, Oxford.
- Carlgrén, I., 1995. National curriculum as social compromise or discursive politics? Some reflections on a curriculum-making process. *J. Curric. Stud.* 27 (4), 411–430.
- Carlgrén, I., 2009. The view of knowledge in the curriculum construction of the 1990s. In: Carlgrén, I., Forsberg, E., Lindberg, V. (Eds.), *Perspectives on the Swedish School's Knowledge Discussion*. Stockholms universitetsförlag, Stockholm, pp. 16–38.
- Carlgrén, I., Forsberg, E., Lindberg, V., 2009. *Perspective on the Swedish School's Knowledge Discussion*. Stockholms universitetsförlag, Stockholm.
- Carstensen, M.B., Schmidt, V.A., 2016. Power through, over and in ideas: conceptualizing ideational power in discursive institutionalism. *J. Eur. Publ. Pol.* 23 (3), 318–337. <https://doi.org/10.1080/13501763.2015.1115534>.
- Curriculum Committee, 1994. *Education and Knowledge: Reprints From the Curriculum Committee's Report School for Education (SOU 1992: 94)*. Statens skolverk, Stockholm.
- Deng, Z., Luke, A., 2008. Subject matter. Defining and theorizing school subjects. In: Connelly, F.M. (Ed.), *The SAGE Handbook of Curriculum and Instruction*. SAGE, Thousand Oaks, CA, pp. 66–87.
- Dewey, J., 2009. *The School and Society and The Child and the Curriculum*. BN Publishing, United States.
- Djelic, M.-L., Sahlin-Andersson, K. (Eds.), 2006. *Transnational Governance—Institutional Dynamics of Regulation*. Cambridge University Press, Cambridge, UK.
- [Utbildningspolitiskt systemskifte?]. In: Englund, T. (Ed.), 1996. *Education Policy System Change?*. HLS, Stockholm.
- Englund, T., 2005. *The Curriculum and the Political Dimension of School Knowledge [Läroplanens och skolkunskapens politiska dimension]*. Daidalos, Göteborg.
- [Vad räknas som kunskap? Läroplansteoretiska utblickar och inblickar i lärarutbildning och skola]. In: Englund, T., Forsberg, E., Sundberg, D. (Eds.), 2012. *What Counts as Knowledge? Curriculum Theoretical Prospects and Insights Into Teacher Education and School*, first ed. Liber, Stockholm.
- Gundem, B.B., Hopmann, S. (Eds.), 1998. *Didaktik and/or Curriculum: An International Dialogue*. P. Lang, New York.
- Haft, H., Hopmann, S. (Eds.), 1990. *Case Studies in Curriculum Administration History*. Palmer, London.
- Hamilton, D., Zufiaurre, B., 2013. *Blackboards and Bootstraps: Revisioning Education and Schooling*. Sense, Rotterdam.
- Hirst, P.H., 1974. *Knowledge and the Curriculum: A Collection of Philosophical Papers*. Routledge, London.
- Hopmann, S., 1999. The curriculum as a standard of public education. *Stud. Philos. Educ.* 18 (1), 89–105.
- Hopmann, S., 2003. On the evaluation of curriculum reforms. *J. Curric. Stud.* 35 (4), 459–478.
- Jackson, P.W., 1968. *Life in Classrooms*. Holt, Rinehart & Winston, New York.
- Kelly, V.A., 2009. *The Curriculum: Theory and Practice*, sixth ed. Sage, Los Angeles, CA.
- Key, E., 1900/1996. *The Century of the Child [Barnets århundrade]*, new ed. Informationsförl, Stockholm.
- Kliebard, H.M., 2004. *The Struggle for the American Curriculum: 1893–1958*. RoutledgeFalmer, London.
- Lundgren, U.P., 1984. *Between Hope and Happening: Text and Context in Curriculum*. Deakin University, Victoria.
- Meyer, H.-D., Benavot, A. (Eds.), 2013. *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Symposium Books, Oxford.
- Meyer, H.-D., Rowan, B. (Eds.), 2006. *The New Institutionalism in Education*. State University of New York Press, Albany, NY.
- Meyer, J.W., Ramirez, F.O., Soysal, Y.N., 1992. World expansion of mass education, 1870–1980. *Sociol. Educ.* 65 (2), 128–149.
- Nordin, A., Sundberg, D., 2016. Traveling concepts in national curriculum policy-making: the example of competencies. *Eur. Educ. Res. J.* 15 (3), 314–328.
- Nordin, A., Sundberg, D., 2021. Transnational competence frameworks and national curriculum-making: the case of Sweden. *Comp. Educ.* 57 (1), 19–34. <https://doi.org/10.1080/03050068.2020.1845065>.
- Oates, T., 2011. Could do better: using international comparisons to refine the National Curriculum in England. *Curric. J.* 22 (2), 121–150.
- Official Swedish Government Report, 2007. *Tydliga Mål Och kunskapskrav I Grundskolan. Förslag till Nytt Mål- Och Uppföljningssystem [Clear goals and knowledge requirements in compulsory school education: proposal for a new system of goals and monitoring]*. Official Swedish Government Reports, Stockholm.
- Organisation for Economic Co-operation and Development (OECD), 2019. *PISA 2018 Assessment and Analytical Framework*. PISA, OECD Publishing, Paris. <https://doi.org/10.1787/b25efab8-e>.
- Schiro, M., 2013. *Curriculum Theory: Conflicting Visions and Enduring Concerns*, second ed. Sage, Thousand Oaks, CA.
- Schwab, J.J., 1969. The practical: a language for curriculum. *Sch. Rev.* 78 (1), 1–23. <http://www.jstor.org/stable/1084049>.
- Schwab, J.J., 1978. *Science, Curriculum, and Liberal Education: Selected Essays*. University of Chicago Press, Chicago.
- Scott, D., 2014. Knowledge and the curriculum. *Curric. J.* 25 (1), 14–28.
- Sivesind, K., Westbury, I., 2016. State-based curriculum work and curriculum-making: Norway's Læreplanverket 1997. *J. Curric. Stud.* 48 (6), 766–782. <https://doi.org/10.1080/00220272.2016.1186741>.
- Sivesind, K., Van Den Akker, J., Rosenmund, M., 2012. The European curriculum: restructuring and renewal. *Eur. Educ. Res. J.* 11 (3), 320–327. <https://doi.org/10.2304/eej.2012.11.3.320>.
- Sundberg, D., 2021. *Swedish Curricula—Curriculum Theory for Education Professionals [Svenska läroplaner – Läroplansteori för de pedagogiska professionerna]*. Studentlitteratur, Lund.
- Sundberg, D., Wahlström, N., 2012. Standards-based curricula in a denationalized conception of education—the case of Sweden. *Eur. Educ. Res. J.* 11 (3), 342–356.

- Swedish National Agency for Education (NAE), 2011. Curriculum for the Compulsory School, Preschool Class and the Leisure-Time Centre: Lgr 11. Swedish National Agency for Education, Stockholm. <https://www.skolverket.se/publikationsserier/styrdokument/2018/curriculum-for-the-compulsory-school-preschool-class-and-school-age-educare-revised-2018>.
- Swedish National Agency for Education (NAE), 2022. Curriculum for Compulsory School [Läroplan för grundskolan]. Fritze Förlag, Stockholm. <https://www.skolverket.se/undervisning/grundskolan/aktuella-forandringar-pa-grundskoleniva/andrade-laroplaner-och-kursplaner-hosten-2022>.
- Tyack, D., Cuban, L., 1995. *Tinkering Toward Utopia—A Century of Public School Reform*. Harvard University Press, Cambridge, MA.
- Wahlström, N. (Ed.), 2022. *Equity, Teaching Practice and the Curriculum: Exploring Differences in Access to Knowledge*. Routledge, Oxon. <https://doi.org/10.4324/9781003218067>.
- Wahlström, N., Sundberg, D., 2017. *Transnational Curriculum Standards and Classroom Practices—The New Meaning of Teaching*. Routledge, London.
- Waldow, F., 2015. From Taylor to Tyler to no child left behind: legitimating educational standards. *Prospects* 45, 49–62. <https://doi.org/10.1007/s11125-014-9334-x>.
- Westbury, I., 2008. Why do states make curricula, and how? In: Connelly, F., Michael, H., Ming, F., Phillion, J. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage, Los Angeles, CA, pp. 45–65.
- Wyse, D., Hayward, L., Pandya, J.Z. (Eds.), 2015. *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, first ed. Sage, Los Angeles.
- Young, M., 2008. *Bringing Knowledge Back In: From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, London.

Interdisciplinarity in basic education: interpretations, relevance and principles

Yves Lenoir, University of Sherbrooke, Sherbrooke, QC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	268
Some interdisciplinary conceptions in education	268
Three different logics	268
Interdisciplinary conceptions and practices in Quebec education	270
Some arguments to implement interdisciplinarity in school education	272
Fundamental principles to guide the implementation of the interdisciplinary approach in education	274
Conclusion	277
Acknowledgment	277
References	278

Glossary

Circumdisciplinarity (from Latin *circum*, “around”, accusative adverbial of *circus*, “circle”) takes into account not only academic and disciplinary knowledge but also others knowledges (experiential knowledge, common-sense knowledge, knowledge of otherness, etc.) that practitioners (specialists, teachers, learners, etc.) incorporate into their practices (see Lenoir, 2000; Lenoir et al., 2006)

French Republican tradition « the French Republican tradition [...] posits state guaranteed non-domination as a definition of liberty, and is built on the foundation of individuals’ ability to go beyond individual interests and adherences in order to work together in constructing an egalitarian political society » (Lenoir et al., 2015, p. 44)

Integration refers to three complementary meanings related to interdisciplinarity. First, it refers to the interrelated learning processes implemented by students (integrating processes); second, it refers to cognitive outcomes achieved by students (integrated knowledge), and third, to an interdisciplinary approach implemented by the curriculum or by the teacher (integrative approach)

Intradisciplinarity refers to the interrelations established in an area of a disciplines based on its internal logic. For instance, interdisciplinary links can be made between geometry, arithmetic and algebra, as these are all components of mathematics. In education, it is also possible, for example, to conceive the establishment of links between history, geography and economy as humanities

Interdisciplinarity in its broader (and common) sense, is often used as a generic term to designate any type of link that can be made between disciplines. Using the term polydisciplinarity would surely be more useful. Strictly speaking, interdisciplinarity refers to real interactions between two or more disciplines concerning their concepts, methodological procedures, techniques, etc. It is therefore incompatible with any kind of cumulative perspective, as it imposes authentic interactions

Mediation two-tier regulatory system: (1) as a process by which a learning subject establishes a cognitive relationship to the object of knowledge (intrinsic mediation); (2) as a process of didactical-pedagogical intervention of the teacher on cognitive mediation carried out by the pupil (extrinsic mediation)

Monodisciplinarity (or unidisciplinarity) refers to the involvement of one single discipline

Multidisciplinarity simply refers to the involvement of two or more disciplines, without specifying the presence or absence of links between them, or the types of links established

Pluridisciplinarity refers to the basic juxtaposition of two or more disciplines where disciplinary specialists work side by side in an additive manner, and without integrating effect. It also implies the absence of any kind of direct link between these disciplines, with no integrative effect. In the academic context, pluridisciplinarity is often characterized by a thematic approach in which a question is addressed separately by different academic disciplines, the only common thread being in fact the theme in question. It refers to pseudo-interdisciplinarity

Transdisciplinarity like many other concepts, is loaded with many senses: (1) it can be understood, first of all, in the sense of cross-disciplinarity (transversality) between two or more disciplines (across) (for example, Barbier, 1997); (2) in the sense of exceeding a discipline, leading to scientific unity based on a group of unifying principles, concepts, methods and goals acting on a metascientific level (beyond) (for example, Nicolescu, 1996); (3) as understood for example by D’Hainaut (1986), in the sense of focusing on behaviors (below), or, as Hamel and Orival (2017), in the sense of a philosophical knowledge to guide practical intervention; (4) it is also used instead of the term “interdisciplinarity”, thereby undermining its own relevance; (5) in the sense of cross-disciplinary mobilization in the context of a project, as understood by Fourez et al. (2002). Morin (1990), who also repudiates previous conceptions of transdisciplinarity, suggests that it should be considered as an approach that allows « to distinguish, separate, oppose, therefore disjoin relatively these scientific fields, but which can make them communicate without operating the reduction » (p. 127), to be able to « account for the multidimensional characteristics of

any reality studied (Morin, 1990, p. 309); (6) in the sense of circumdisciplinarity, taking into account at the same time disciplinary knowledge and also others knowledges (for example, Klein et al., 2001; Lawrence, 2019)

Introduction

Over the past few decades, some of the teaching curricula in the Western world from primary school up through university have advocated the use of interdisciplinary approaches to education. This is shown, for example, by the CROSSCUT project (Cross-Curricular Teaching) funded by the European Commission since 2016 (Reverdy, 2019), by the book edited by Darbellay et al. (2019) which includes the implementation of school interdisciplinarity in different European countries, or by the 2010 special issue of the journal *Issues in Integrative Studies* (Lenoir and Klein, 2010) as regards primary and secondary education (designated as “basic education” in this article) across nine Anglophone, Francophone and Hispanic societies. It is now clear that teaching curricula have undergone substantial change over the past decades. For example, some have grouped together school subjects into subject domains, as in Quebec programs. Even more importantly, they have opened up to questions directly connected to social preoccupations, i.e., to peace, health, civic, environmental, and other types of education; this is what is illustrated by the “educations to” in Quebec, the “convergence themes” in France, the “transversal themes” in the programs of the French-speaking community of Belgium or the “transversal capacities” in French-speaking Switzerland. In addition, today’s curricula stress a strong contextualization of teaching–learning situations, to root them in experienced realities. These pragmatic and/or cultural necessities (social practices of reference) consequently call also for transdisciplinarity that cannot be treated here.

In this article, we will examine interdisciplinarity only in connection with these primary and secondary levels of instruction. The interested reader can consult the glossary or Klein (1990, 1996, 2002) and Lenoir and Hasni (2010) for the taxonomy of interdisciplinarity to which we will be referring in this text as needed. We wish, firstly, to highlight at once the existence of three different fundamental conceptions of interdisciplinarity in education, which, in our view, should be strongly complementary and, then, to present results of many studies we have conducted for 40 years on teaching practices in Quebec to expose that the concept of interdisciplinarity has multiple significations (four overall) that create semantic confusion, to say the least, and ultimately practices that are not so interdisciplinary. Secondly, we suggest arguments which can justify this approach. In other words, which aims can be pursued to apply school interdisciplinarity, as opposed to monodisciplinarity? Or why implement interdisciplinarity in basic education? Thirdly, we suggest six foundational principles by which to found, describe and advise the implementation of interdisciplinarity in education. We will conclude by a definition of school interdisciplinarity.

Some interdisciplinary conceptions in education

Interdisciplinarity has been a preoccupation of school systems since the end of the 19th century in the United States and has been the subject of numerous debates since then (Boix-Mansilla and Lenoir, 2010; Klein, 1990; Kliebard, 1992a). Only in the 1960s and 1970s, however, with the development of economic and cultural globalization, the internationalization of exchanges, and various social and economic pressures, did it significantly emerge in the European world (Ansart, 1990; Apostel et al., 1972), and later still in the Spanish-speaking world (Fazenda, 1994, 1995; Miñana Blasco, 2001). This has been documented in particular by the articles published in the special issue edited by Lenoir and Klein (2010), which respectively provide a historical portrait of the emergence of interdisciplinarity in basic education in North America (United States, Ontario and Quebec), South America (Colombia), Europe (Spain, France and Switzerland), and Australia. While these articles may show that the conditions and circumstances under which interdisciplinarity emerged are different in each country, they also show that « the emergence of the concept can essentially be traced to two sources: government decisions and economic pressures in line with the requirements of a global market economy » (p. xi). They also suggest the existence of « a close link today between the competency-based approach and the recourse to the interdisciplinary perspective in curricula » (p. xi), a link which, moreover, is based on various epistemological foundations and/or neoliberal and/or utilitarian perspectives.

Three different logics

It is worth noting here that each society has approached the issue of interdisciplinarity in its school system in its own way, depending on its unique historical, political, cultural, social, educational, and other features. An anecdote eloquently attests to these different interpretations of educational realities. In 2002 a reform of initial teacher education was announced during a departmental assembly at the university where we work, in a context where the previous reform had not yet been completed and the new one was to be applied the following year. Several of our colleagues have the reaction: « What are we going to do? » This reaction left me a strong impression. During a trip to France the following month, we presented this issue to colleagues at the universities of Nantes and Toulouse. Their reaction was completely different. They asked, « Why the reform? », « How does it change the training content? » Two months later, we found ourselves at the Catholic Pontifical University of São Paulo where we have been invited by our colleague

Ivani Fazenda. Her reaction and the reaction of her students to whom we have presented the same issue were very different: « Would I feel good in this reform? » These were three very different reactions to the same situation, already indicating the existence of distinct interpretations.

By more extensively examining the Anglophone North American and Francophone European perspectives (and, later, a certain Latin American and Brazilian perspective), we have been able to identify a series of factors that might explain the existence of distinct socio-historical logics and their underlying rationales. These logics can be illustrated based on a second real-life anecdote. In 2000, at an international conference, we invited 32 researchers from various European countries (Belgium, France, Switzerland), North America (Canada, United States), and South America (Argentina, Brazil, Chile). The European participants extensively discussed the relevance of implementing interdisciplinarity, the meaning it should be given, and its impact on the hierarchy of scientific disciplines (the tree of science). Gordon Vars and Julie Klein from United States then took the floor and wrote on a board to explain, in a few very systematic points, how interdisciplinarity can be operationalized. As for the Brazilians, they put on music and projected slides

Although we must avoid over-generalization and caricature, these two concrete examples taken from direct experience, and supported by critical analysis of many publications originating from these three cultures, show just how differently the notion of interdisciplinarity can be interpreted. We have been able to highlight the existence of three distinct logics (Fig. 1) that result from sociohistorical and cultural factors (Lenoir, 1999, 2002, 2005; Lenoir and Hasni, 2004; Lenoir et al., 2015, 2001).

In Francophone Europe, and especially in France, interdisciplinarity is strongly marked by rational Cartesian thought and the influence of philosophers (including Voltaire) and the Encyclopedists in the Enlightenment period (Sachot, 2001, 2004). Hence, in this country, the search for meaning, rational logic, and consequently the epistemological perspective (Resweber, 1981) and the relationship to knowledge (Beillerot, 1989; Charlot, 1997) are central and omnipresent to interdisciplinary reflection. In the school context, the debate most often centers on protecting disciplines' specific identities, rather than forging interdisciplinary links (Durkheim, 1969/1938; Stichweh, 1991). Socialization in the school context, for its part, is above all conceived as the transmission

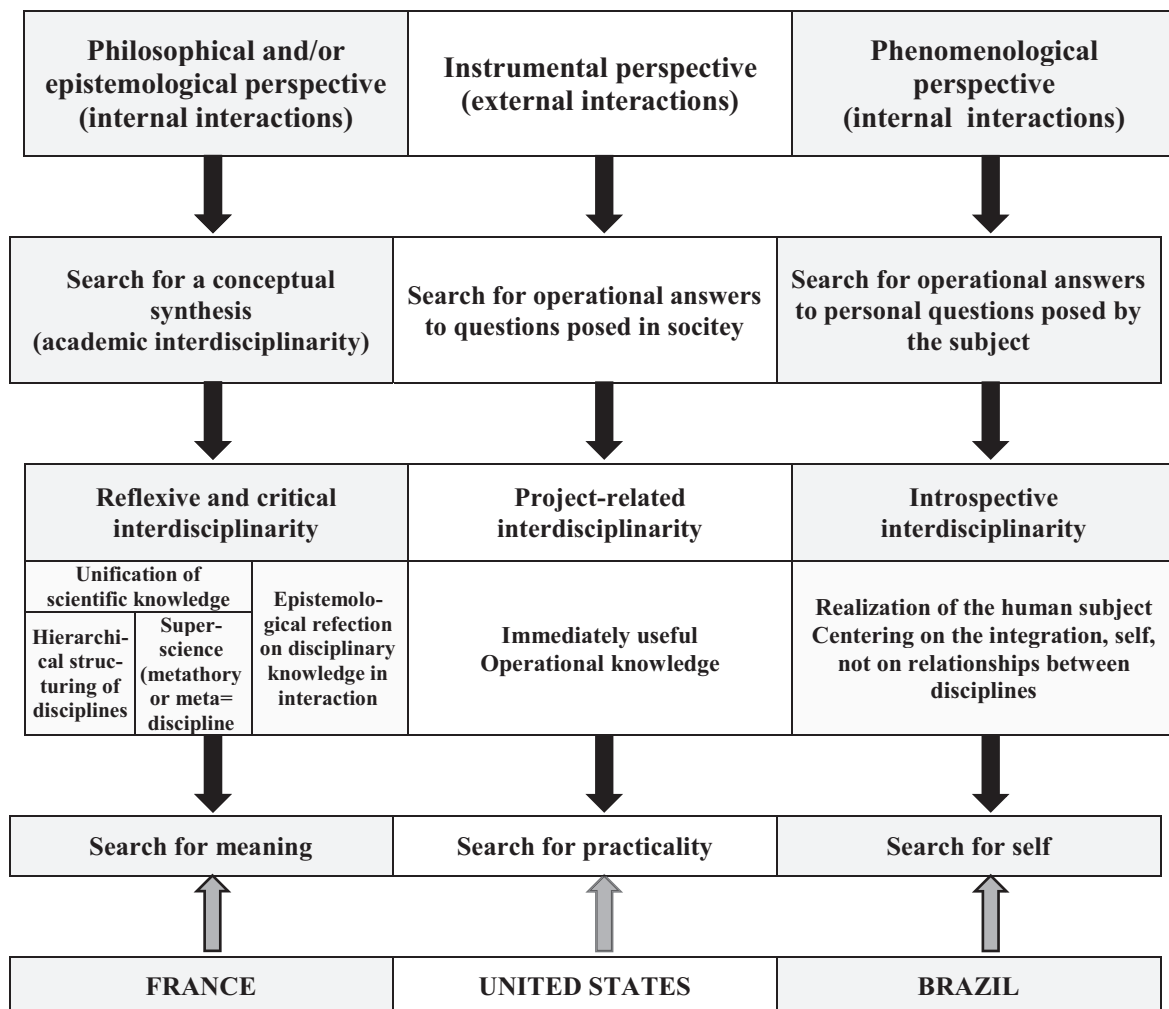


Fig. 1 Three distinct readings of interdisciplinarity.

of cultural heritage in direct connection with a universalist vision consistent with the French Republican model. Subject-specific didactics is therefore omnipresent since it intends to analyze the knowledge to be taught in order to make it accessible to students (for example: Fourez, 1994, 1998; Fourez et al., 2002; Lenoir et al., 2015; Reuter, 2019; Lenoir, 2000).

The North American English-speaking logic is based on other references. The break at the end of the 19th century with the educational model inherited from the United Kingdom led to the promotion of pragmatism (Tanner and Tanner, 1990; Rudolph, 1977). This change has highlighted the importance of know-how (how to do), a functional approach. In the United States, which is dominated by an instrumental (pragmatic and organizational) educational perspective (Klein and Newell, 1996; Kliebard, 1992a,b; Lenoir, 2002; Newell, 2001; Pinar, 1998), school interdisciplinarity is thought of as a search for operational answers to questions that arise within society. This focus on solving societal problems can be characterized as a “project” interdisciplinarity (Klein, 1990; Klein and Newell, 1996; Lenoir et al., 2015) in which the knowledge involved is immediately useful and operational. In schools, the importance of know-how is clearly illustrated by publications that give operational methods for applying interdisciplinarity in class (e.g., Clarke and Agne, 1997; Fogarty, 1991, 1993; Frazee and Rudnitski, 1995; Jacobs, 1989; Pollard and Tann, 2002; Tann, 1988; Vars, 1993). As noted in Lenoir et al. (2015), « In the United States, the focus on the functionality of learning can be traced back to the pragmatism established at the end of the 19th century. This pragmatism should be associated with the preoccupation, dating to the same period, with placing the individual student at the heart of learning (the “subject pole”). Indeed, the need to socially integrate masses of immigrants necessitated a focus on pedagogical practices (to facilitate learning processes) and socialization processes (to foster civic skills). These two fundamental traits of the North American curricular conception have underpinned the professionalizing orientation of its education system. In the United States, the priority appears to be the adherence to the “virtues” and values of American society (personal skills) as well as know-how geared toward problem solving » (p. 66). Hence, the French epistemological search for meaning stands in contrast to American functionality, even if this does not preclude the existence of other interdisciplinary currents in the school context.

In addition to these two distinct approaches to interdisciplinarity—one primarily involving knowledge and the other know-how and certain personal skills—it is important to acknowledge a third, which is strongly expressed in a certain Brazilian conception of interdisciplinarity. The perspective adopted in this case is clearly phenomenological and involves a symbolic interactionism. The emphasis is on the actor as a medium for interdisciplinarity, and therefore on the subjectivity of individuals who are integrated into everyday life and on their methodological intersubjectivity. The main concern is the process of continually constructing the individual in a situation, a process involving not only self-knowledge, but also knowledge of others (Fazenda, 1994, 1998). In the Brazilian logic, which is profoundly subjective (and intersubjective), school interdisciplinarity is mainly focused neither on a reflexive function with regard to knowledge, nor on a pragmatic function. The predominant aim is rather to pursue the realization of the human subject by stressing the integrating “I”, not the relationships between disciplines, and even less functional and pragmatic action. From this affective perspective, interdisciplinarity above all has to do with the internal interactions of the subject, as he or she searches for answers to personal questions. The idea might be described as an introspective interdisciplinarity that is founded on a phenomenological search for the self. We will come back to these three perspectives at the end of the following section in order to underscore the need for complementary relationships between these perspectives in school interdisciplinarity.

Finally, in order to make the problem of interdisciplinarity even more complex, it would be essential to specify the type of interdisciplinarity we are talking about and which we intend to use. In fact, interdisciplinarity is part of the field of practical, scientific, vocational and academic operationalization. Each of these perspectives has specific purposes, reference systems, study objects, and operational modalities (Lenoir, 2019). Here, let us say again, we are only talking about school interdisciplinarity.

Interdisciplinary conceptions and practices in Quebec education

It is not enough to use the word “interdisciplinarity” for a given practice to accurately correspond to the concept. The many studies we have conducted for almost 40 years on teaching practices in Quebec have allowed me to identify Quebec teachers’ representations of interdisciplinarity at school as well as its implementation in their teaching practices (Lenoir and Hasni, 2010). The results of these researches in schools have brought to light that interdisciplinarity exhibits a polysemy sometimes bordering on cacophony (Fourez et al., 2002). This polysemy is of little help in circumscribing the term’s meaning(s) in terms of its scientific basis. The word is clearly fraught with incoherence that has led to errors and obscurities. The concept of interdisciplinarity is given multiple significations that create semantic confusion, to say the least, and ultimately practices that are not so interdisciplinary. Analysis over the past 30 years has identified four groups of conceptions and teaching practices in the school context—pseudo-interdisciplinary, hegemonic, eclectic and holistic—which show different degrees of proximity to the scientific foundations of school interdisciplinarity (Fig. 2).

Placed on a cartesian axis (x and y), these four dominant approaches make up the extreme poles of two intersecting continua. The x axis (holism-eclecticism) concerns the degree of fusion or dispersion of school disciplines, while the y axis (hegemony-pseudo-interdisciplinarity) concerns the intensity of relations between disciplines, ranging from domination to the absence of real links. The two-sided arrows ac and bd illustrate the fact that interdisciplinary practices can be associated with one approach or lie somewhere between two approaches. The circle at the center of the diagram represents the interdisciplinary perspective we will present further on.

The results of various researches show that a theme-based pseudo-interdisciplinary is especially often preferred by first-cycle primary teachers. This tendency is mainly caused by their significant preoccupation with awakening student interest, in which case the relational and psycho-affective dimensions (together with the organizational dimension) take up a preponderant place

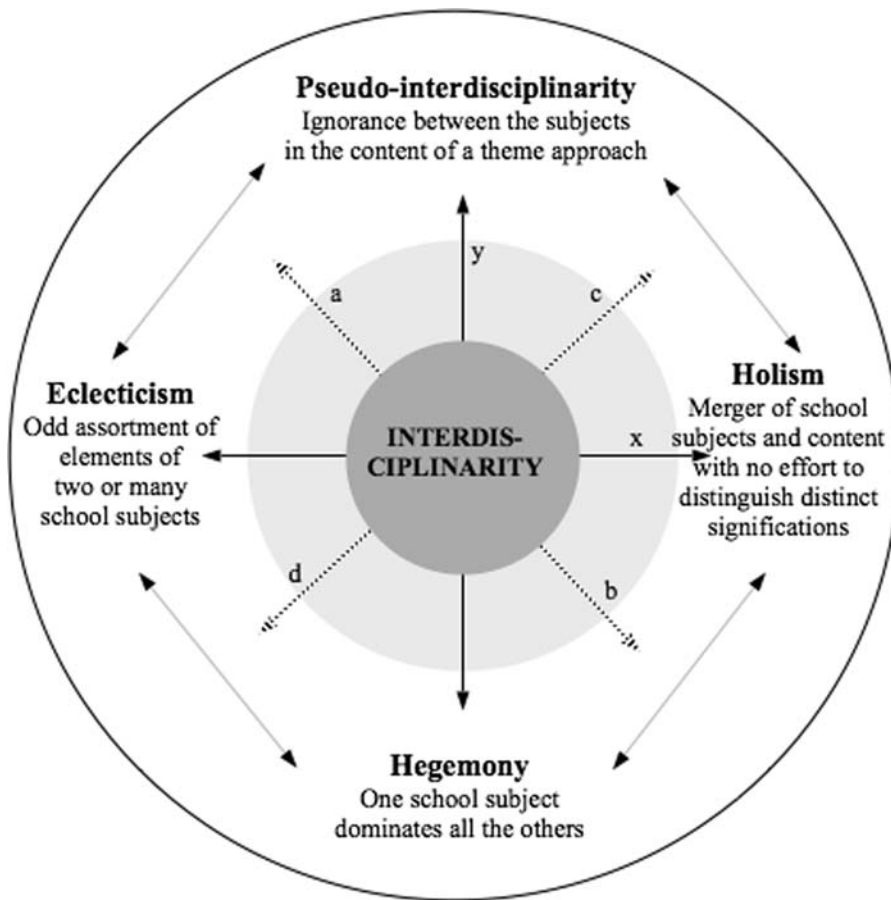


Fig. 2 The poles of interdisciplinary practice.

in their interventions with students, to the detriment of cognitive dimensions. On the other side of the spectrum, the hegemonic approach, in which certain disciplines are in fact no more than a pretext for teaching other disciplines, may be explained by the priority certain teachers give to teaching the language of instruction. The eclectic approach, which is profoundly destructuring, since, to quote [Jacobs \(1989\)](#), it perceives learning content as a “pot-pourri” from which one can pick at random, is also commonly observed in teaching practices. The holistic approach, which is based on the negation of all disciplinary specificities in the name of a natural approach, is frequently observed among teachers adhering to pedagogical conceptions that promote an open and active pedagogy centered on student interest, thereby sidelining cognitive objectives. These four approaches are in some cases used by teachers whose primary aim is to meet curricular requirements from a strictly administrative point of view. When this happens, these approaches serve mostly as justifications for the absence or near-absence of the teaching of certain school disciplines that are officially mandatory and planned for in the program, but in fact considered secondary from a societal point of view. This situation especially applies to the teaching of the arts, natural sciences, and humanities. These different trends have been presented in [Lenoir and Klein \(2010\)](#). For example, pseudo-interdisciplinarity, if not eclecticism, can be found in Australia ([Long et al., 2010](#)) and eclecticism, along with the hegemonic perspective and holism, can be found in Colombia ([Rodriguez and Miñana Blasco, 2010](#)). Likewise, the hegemonic perspective can be noted in teaching practices in Ontario ([Clausen and Drake, 2010](#)). It has also been identified in an article by Baillat and Niclot on France ([2010](#)) and an article by [Segovia et al. \(2010\)](#).

The discourse on interdisciplinarity therefore conceals practices most often marked by the primacy of certain socially valued disciplines and a dilution of the socialized knowledge of “secondary” disciplines. It can be hypothesized that teachers’ use of “interdisciplinary” practices is based on two key preoccupations: first, gaining time, and second, boosting student interest and motivation via thematic approaches or projects. The cognitive contributions ultimately receive little consideration. This state of affairs connects with the “practical logic” of [Bourdieu \(1980\)](#). Practices are “practical” in that they are « capable of producing intrinsically coherent practices » (p. 145) in a given situation, and in that they involve the principle of an economy of logic. In other words, they are simple and useful because they are based solely on a few generative principles that are strongly interrelated and impose the structuring of (and structuring within) action. In short, this logic—which we associate with the world of common sense—can be easily mastered, managed, and put into practice, thus requiring little time and energy.

Among many other results, we have highlighted the strong hierarchy of school disciplines. The reasons are not primarily school-related, but social, political, and ideological-economic. The importance of the disciplines does not, however, depend *a priori* on

government discourse, which is also influenced by pressure from lobbies and social groups. Their importance depends on the value that society and institutions actually assign to them. In this vein, British sociology research (especially by Bernstein (1971a,b) and Young (1971), who contrast a compartmentalized curriculum with an integrated one) underscores the intrinsic hierarchical character of a compartmentalized curriculum and the very powerful process of social selection and control that it implies.

In contrast to these different and more or less skewed interpretations, we consider that school interdisciplinarity postulates the establishment of mutual dependence, free of predominance or neglect, between school subjects. We will come back to this point. Such an epistemological, social, and political posture leads to a different reading of curricular components, one based on a search for the specific identity of each discipline (its place and function on the cognitive and social levels) as well as the complementarity of contents that is needed in order to apprehend, communicate, and relate to natural, human, and social reality. This posture also calls for a critical position toward the hierarchy of school disciplines and the resulting function of social selection. To return to the perspective developed by Bernstein, the “serial” (additive or cumulative) conception of curriculum implies opacity, rigidity, and hierarchy, while the “integrated” (we might say “integrative”) conception promoted by interdisciplinarity involves permanent inter-activeness, a sharing of power between teachers and between teachers and students, and a necessary transparency about the aims and goals that are being pursued.

Some arguments to implement interdisciplinarity in school education

As most interdisciplinary authors point out, current social and economic transformations would be at the origin of curricular changes with an interdisciplinary, if not transdisciplinary, tendency. Traditional disciplinary and often compartmentalized models would no longer meet the needs of students and societies, of complex social and environmental issues that require comprehensive, integrated analysis and understanding. It is important to recognize the complexity of the natural, human, and social reality in which we live, as well as the fragmentation of the knowledge transmitted in schools. This perspective has been elaborately presented and described by Apostel et al. (1972), Klein (2002) and Torres Santomé (1998), among others.

Thus, interdisciplinarity would become an element of response to the complexity of the actual world. This perspective, which fights against fragmentation, was presented at length and explained among other things by Morin (1977, 1990). While simplification has led to great scientific discoveries in the natural sciences, « nature is a polysystemic whole » recalls Morin (1977, p. 100), the study of human beings, their relationships and social practices absolutely require the use to the complexity where the interrelationships between the constituent elements are grasped as organized systems. So, we are simply pointing out this central argument, which is certainly a very important one, to justify the use of interdisciplinarity. But in the context of youth education, we would like to point out other arguments that directly bear on teaching–learning processes. These arguments are based on the fact that school learning takes place in a specific context, namely the physically closed space of the classroom, that is, a system of interactions among pupils and between them and the teacher. So, we would thus consciously like to examine aspects that are less often considered when dealing with school interdisciplinarity in general.

It is important to recall, first of all, that students 6 to 16–17 years old do not generally study out of a love for the school disciplines, but instead, one might say, out of a “love” for the teacher. More specifically, although we could not develop here the entire set of arguments that originally stemmed from dialectic Hegelian thinking (Hegel, 1941/1807, 1969/1805) and that has been developed by so many researchers, including Kojève (1947), Lacan (1966), Lordon (2010, 2011/2006, 2013) and Honneth (2013)—and well before them by Spinoza (1849/1667), who described it as the essence of human nature and the driver of action (the *conatus*)—the student strives to capture the desire to know that he/she decodes in the eyes, attitudes, and discourse of the teacher. The student’s behavior in class is then dictated not by the need to know, but originally by social, intellectual, and affective relationships that he/she develops with the teacher with the intent of appropriating the qualities decoded in the teacher and obtaining recognition of value as a human being (Lenoir, 2017/2014). A love of knowledge may develop only subsequently. The teacher’s role is therefore crucial. An interdisciplinary approach in this context assumes an important function of facilitation; however, it will only do so insofar as it takes into account the following.

First, interdisciplinarity requires teaching–learning situations that stimulate research practice—for example, “a-didactic situations” (Brousseau, 1998) and “problem situations” from a constructivist perspective—and that can make sense for students on the epistemological (the relationship to knowledge), ontological (the relationship to the subject) and social (the relationship to life in society) levels. The situation must be contextualized, and such a contextualization must necessarily be multi-dimensional, given that, to be understood the situation requires an interdisciplinary approach.

Yet for many students, distinguishing between everyday life and school life is problematic. Scholars such as Bernard Rey of the Université Libre de Bruxelles (private communication) have observed misunderstandings that arise when students are confronted with a proposed situation. Let us take a real-life example: when Professor Rey, who had been invited to a classroom, asked students how to take the train to go home, the route to take, the distance, the duration of the journey (which required the use of mathematical and geographical knowledge), a student immediately answered him that he shouldn’t worry, that his father was not working that day and would be pleased to drive him to the station on time. Joël Bisault (2011), of the École normale supérieure de Cachan in Paris, has likewise underlined problems of interpretation in school assignments among primary students by showing that some adopt a “domestic point of view.” This misunderstanding has also been noted in other studies of classroom practices in Quebec. For example, a study of four teachers’ instruction on plant parts (biology) shows that hypothetical situations that are strongly anchored in everyday life (an aunt’s organization of a food shopping basket into fruits and vegetables) can be an obstacle to posing

scientific problems. In this case, the students' interpretations of the situation led them to focus their verbal exchanges on their everyday knowledge of fruits (such as their sweetness and juiciness) or on men and women's culinary skills (use of fruits and vegetables) (Hasni, 2011).

There is a real trap that students can become engaged in exclusively relational, affective, domestic, and other readings of the situation if the contextualization remains superficial and incomplete. This problem persists in secondary school. Suffice it to say that if contextualization is not adequately developed, there is a risk that students will differently interpret the instructions and tasks given by their teachers, because they will not decode the school task within the situation. The issue is with the student's task, as opposed to the child's or adolescent's task, that is expected of them, because they may perceive only the visible (external) facet of the task to complete. They might also be unable to assign meaning to the situation or the task itself (the internal context) and lose interest and motivation in the work that is expected; in short, they "disconnect." Indeed, this is one of the factors leading to a lack of interest in school and cases of dropout. In short, the interdisciplinary perspective in school fosters students' search for meaning, in terms of the knowledge (epistemological perspective), the subject (ontological perspective), and the engagement in life activities (social perspective) that are involved.

Second, because of the need to use varied disciplinary knowledge, interdisciplinarity calls for the use of what Boix-Mansilla (2017) refers to as a pragmatic constructionist epistemology and to what Martinand (1986) called "basic social practices" and Vygotsky (1987, 2014) "everyday concepts". Pastré (2011) works on vocational education in France highlighted the conceptual structure of a situation and the centrality of effective knowledge in practice, as well as the importance of these social practices and more particularly of practical concepts needed to organize the action. That is why we suggest the use of the concept of circumdisciplinarity or transdisciplinarity if the latter is properly designed.

The fundamental *raison d'être* of such a pluralist epistemological concept is to promote and support students' construction of meaning. Precisely because students struggle to distinguish between everyday life and school life, between common sense learning and school-related learning, we believe it is necessary to use the springboard of teaching-learning processes that involve situations anchored in reality. This allows teachers to leverage students' life experiences, as well as the pragmatic notions and concepts that they use in their everyday lives and that are indispensable for organizing action. Vygotsky (2014) had already established a link between everyday concepts and scientific conceptualization processes. Real-life situations allow students to invest themselves based on their life experiences, so that they can inter-subjectively debate their different conceptions, and then, with the teacher's help, progress toward the use of scientific approaches and toward the production of abstract and general concepts.

We postulate here, without being able to develop it, that knowledge is not in the head of individuals and that it is not their property, but that it is a dialogical, intersubjective process that establishes itself in class. Here we join the conceptions of social constructionism of Gergen (1994, 1995). However, if, in everyday life, pupils communicate, build knowledge, solve problems, experiment, these different processes are part of the common sense. But school must be the place where they move from these processes acquired in everyday life to scientific approaches that involve the critical and thoughtful objectification of actions systematically and explicitly laid: communicating, conceptualizing, problem solving, experimenting, etc. To take only this last approach, it is necessary to distinguish between expressing an opinion and carrying out an experiment. That is why test and error and empirical groping are starting points to be overcome from an integrative perspective (Lenoir, 2017/2014).

Even if, in everyday life, students communicate, construct knowledge, solve problems, and carry out experiments, these different processes tend to come down to common sense. School should be a place where they make the transition from these processes acquired in everyday life to scientific approaches that involve a critical and well-considered objectivization of actions that are carried out systematically and explicitly: communicating, conceptualizing, problem-solving, realizing, experimenting, etc. To address only this last process, a distinction must be made between giving an opinion and conducting an experiment. This is why trial and error, and empirical efforts at understanding, are a point of departure that must be moved beyond from an integrative perspective. Student action is therefore a point of departure, but also a stumbling block. Education must indeed focus on the need to apprehend a subject who is "capable", who can say "I can" or "I cannot," and not only epistemological subjects who "know," who master knowledge and who can say "I know" or "I do not know." Interdisciplinarity carries with this conception of an inseparable bond between cognition and functional action, for "I can" if "I know". In short, the interdisciplinary perspective precisely requires the interrelated use of a number of scientific processes.

Allow us to give a real-life example that illustrates the need to anchor teaching-learning processes in the realities of students. Invited to a classroom by a teacher in order to do a science activity, we asked the students, based on the program objectives at the time, to design, draw, describe, and explain the production and functioning of an object that could make noise as it moved (thereby using science, technology, drawing, and French). The students worked in teams and one of them helped out most of these teams. However, before entering the classroom, the teacher had told us that the student who was helping all the others was the most difficult, the least cooperative; that was why he was seated at the back of the class. When we congratulated the student, who was one of the most pleasant and helpful during the entire activity, for his valuable help and the quality of guidance he gave to other students, we also asked how he had developed these skills. The student answered that his father was an engineer and that he liked to imitate him. This may have been a happy coincidence, yet interdisciplinarity clearly offers the possibility of giving meaning to learning activities.

A third contribution of interdisciplinarity in basic education has to do with the need for complementarity and overlap in the scientific processes to which we have just referred. Rather than naively or unconsciously—or neglectfully—employing only problem solving, interdisciplinarity can (among other things via a problem- or project-based approach) support the development of integrating learning processes. It is able to do so through the use of several approaches required by the necessary complementarity

of diverse disciplinary outlooks, namely conceptualization, problem-solving, communicational, experimental, esthetic and other processes. It also stresses the fact that problematization is central to a teaching/learning process, and that scientific approaches are cognitive processes that must in themselves be an object of learning. It is impossible, for instance, to produce an experimental protocol without having first conceptualized the experiment's components; and the results, for their part, require a communicational approach. As a result, interdisciplinarity in the school, in its training function, must promote the use of an integration of scientific processes.

Fourth, and this point connects with the question of metacognition and the self-regulation of learning, interdisciplinarity is only genuinely meaningful if metacognitive processes are introduced and promoted by the teacher, and if they are applied by students. What has been observed in everyday life is that teachers teach knowledge, sometimes by making links between different school disciplines, but they forget or neglect the metacognitive processes that are indispensable to ensure effective learning. While they are often led to ask students whether they enjoy an activity, and hence call upon affective dimensions, they forget the cognitive and procedural dimensions. Indeed, only rarely does one hear a teacher asking the students, « What have we just learned? », « How did we arrive at this knowledge? », « What pathway, what steps have we taken? », « What difficulties did we come up against? », « How did we resolve these difficulties? », etc. Students' initial representations need to be collected and held on to so that they can be referred to and compared during a culminating synthesis' activity, in order to take stock of the cognitive ground that has been covered.

The following example clearly illustrates this absence of a metacognitive process. In a class during last year of primary school, a teacher gave students an aerial photograph of their village. Several students immediately said that the picture was of Quebec. We will not describe all the activities held by the teacher, including a field trip to the village with direct observation, graphic representations, the production of texts, etc. Following these activities, the same students said, « We knew it all along, it's our village! », thereby forgetting their initial representations, which had not been collected. Moreover, the teacher had not given the students an opportunity to apply the metacognitive processes that would have allowed them to integrate both the procedural processes by which they had constructed the new knowledge and the knowledge itself. As a result, there was no guarantee that this knowledge and especially the approaches implemented have been acquired and that the capacity to reinvest it in the future had been developed.

Finally, fifth, and not least, interdisciplinarity needs the establishment of a reciprocal dependence, without any predominance or ignorance, between school disciplines. Such a posture, at the same time epistemological, social and political, leads to another reading of the components of a curriculum based at the same time on the search for the specificity of each discipline (its place and its function on the cognitive and social planes) and the necessary complementarity of its contents to apprehend and communicate the natural, human and social reality, and to enter into a relationship with it. It also elicits a critical position on the hierarchy of school disciplines and the resulting social selection function (Bernstein, 1971b; Gosselin et al., 2005; Lenoir et al., 2000; Pate et al., 1997; Wineburg and Grossman, 2000).

In the few points we have just presented to justify the use of interdisciplinarity in basic education and to discuss a few of its aims, the reader will note that we have called upon the mind, the hand, and the heart, that is to say, the cognitive, pragmatic, and ontological dimensions.

Fundamental principles to guide the implementation of the interdisciplinary approach in education

While several characteristics to be respected are already in the previous section, we would like to stress here six principles which we believe are fundamental to ensuring a proper use of interdisciplinarity in school education.

The first principle is that there can be no interdisciplinarity without disciplinarity, that is, without formalized cognitive contents, and without their related instrumental and procedural methods, which we refer to in the educational jargon as "learning processes," i.e., methodological processes that make it possible to construct the knowledge prescribed by the curriculum, to experiment it, to resolve it, to communicate it. Although this might be construed as stating the obvious, we would like to recall, following Klein (1996), Newell (1990, 1998), Petrie (1992) and Wineburg and Grossman (2000), among others, that reflection on interdisciplinarity is only meaningful in a disciplinary context and that it presupposes at least two reference disciplines as well as the presence of reciprocal action. The term "interdisciplinarity" itself makes this relationship as a requirement: "inter-disciplinarity" subsumes "disciplinarity"! Numerous definitions of interdisciplinarity that have been researched attest to this position by admitting, at the very least, some kind of relationship between various core components of at least two disciplines, i.e., subjects, contents, processes, techniques, etc. We could talk about the "disciplinary complexity principle". In other words, interdisciplinarity must be designed to facilitate cognitive learning that is at the heart of school function (Lenoir et al., 2016).

The second principle is that interdisciplinarity conceives all school disciplines as important and complementary in constructing and understanding reality. As Haynes (2002) points out, « No disciplinary viewpoint is inherently or universally true or superior to others » (p. xv). School interdisciplinarity must be understood from a relational standpoint, since, to achieve learning, it strives to establish exchanges, connections, and complementary links between school disciplines. This is why it is important to move beyond the views, at first glance antithetical, between theoretical, methodological, instrumental, and critical interdisciplinaries (Gosselin et al., 2005; Klein and Newell, 1996) in the school context. Indeed, each discipline in the curriculum exhibits these features. The disciplines should be used in complementary fashion, particularly to avoid extreme practices such as Platonic pure ideas or a mere technicist application. Hermerén (1985) further notes the complementarity that is needed, from a knowledge integration perspective, between knowing "that," associated with school interdisciplinarity, and knowing "how," associated with instrumental

interdisciplinarity, i.e., between the knowledge he describes as propositional (declarative knowledge) and operational knowledge (procedural knowledge). The “principles of equality, necessity and complementarity” between different disciplinary contents is indispensable; moreover, it stands in contrast to a distinction between primary disciplines, seen as more important, and secondary disciplines, seen as inferior. Instead, it suggests the importance of taking into account the specific identity and functions of each mobilized discipline and its contributions.

This is why interdisciplinarity must be clearly distinguished from multidisciplinary and, above all, from pluridisciplinarity. It is not based on a cumulative perspective, just as a house is more than just a pile of stones, to paraphrase a metaphor from Poincaré (1917). Indeed, interdisciplinarity is incompatible with the tendency to believe that mere physical proximity between people—in this case teachers of different subjects—would be enough to make an educational activity interdisciplinary. Nor is interdisciplinarity compatible with an “additive” vision of disciplines and content. Moreover, it is insufficient to acknowledge the multidimensionality of a real-life situation and the possibility of interpreting it in different ways to bring about interdisciplinarity. The fact that students may be taking a course in one or more other disciplines does not make their education interdisciplinary; at first glance, in any case, we are dealing with multidisciplinary, if not eclectic education. Nor does a field trip to a farm, and the subsequent use of this theme to produce activities in French, mathematics, science, arts, etc., make the approach interdisciplinary. In this case we are dealing with pseudo-interdisciplinarity that is in fact pluridisciplinary in nature. Gusdorf (1967) refers to this frequent “magical mentality” toward interdisciplinarity as an erroneous view that consists in « bringing together specialists from different areas believing that such an encounter would be sufficient to establish a common ground and common language between these individuals who, in fact, have nothing in common » (p. 1089). In short, interdisciplinary practice requires a real complementary interaction between the disciplinary contents where all the school disciplines possess an equal importance, due to their required complementarity and their necessary collaboration, according to the intentions pursued in the teaching–learning process. Again, it is here both the “principles of equality, complementarity and necessity”.

The third principle, which clearly illustrates the complexity of the concept of interdisciplinarity, is that there are several epistemological conceptions with regard to the function of interdisciplinarity, which is illustrated in Table 1. The relational perspective, which refers to the “principle of relationality”, is the academic and vocational perspective that characterizes school interdisciplinarity. If the ampliative perspective is especially characteristic of scientific interdisciplinarity, the restructuring approach is essentially linked to the transdisciplinary perspective or to the challenge of science (Lenoir and Sauvé, 1998a,b).

The fourth principle, and perhaps the most important in our view, is that interdisciplinarity in education belongs to the means, rather than the ends. The aim of interdisciplinarity is the integration of learning processes and the integration of the resulting knowledge. The purpose of using interdisciplinarity is to foster the mobilization of knowledge and processes with a view to realizing and successfully carrying out an action. In other words, it is to promote and facilitate students’ integration of integrating processes and integrated knowledge, as well as their mobilization and application of these elements in real-life situations (Beane, 1997; Henry, 1958; Hopkins, 1937). It therefore requires the establishment of integrative approaches by the teacher, and not the imposing of an integrated curriculum where the integrating process itself has already been established from outside, from above, by the designers of the curriculum, textbooks or activities. This is the “principle of integrative finality”. As Klein (1996) rightly points out, « Interdisciplinarity work gets done by moving across the vertical plane of depth and the horizontal plane of breadth. Breadth connotes a comprehensive approach based in multiple variables and perspectives. Depth connotes competence in pertinent disciplinary, professional, and interdisciplinary approaches. Synthesis connotes creation of an interdisciplinary outcome through a series of integrative actions. Synthesis does not derive from simply mastering a body of knowledge, applying a formula, or moving in linear fashion from point A to point B. [...] It requires active triangulation of depth, breadth, and synthesis » (p. 212).

Interdisciplinarity consequently requires the adoption of a socio-constructivist epistemological posture, given that it requires that learning subjects obtain recognition and assert themselves as knowledge producers, and requires that this conceptualization process be part and parcel of the social context. Students’ experience of reality is then no longer a representation of the world “as it is” but rather, as Stengers (1993) aptly states, “a risky invention” (p. 62). This is why it is indispensable for the teacher to introduce the conditions deemed favorable to support and guide students’ implementation of an interdisciplinary approach to the cognitive questions that are addressed. At the same time, however, it must be kept in mind that integration is above all a cognitive process that must be performed by the students and not the teacher or the curriculum.

At the school level, this implies the implementation of an interstructuration model, based on dialog, exchange, where the teacher acts as a mediator on the interaction students establish with the learning contents (Lenoir, 2017/2014). This also requires, among

Table 1 Three epistemological conceptions of the function of interdisciplinarity.

<i>Epistemological options</i>	<i>Characteristics</i>
1. Relational approach (education)	Establish links (complementarities, convergences, connections, etc.), “gateways” (<i>bridge-building</i>)
2. Ampliative approach (research)	Bridging the gap between two existing sciences (research) (emergence of new scientific disciplines) (<i>no man’s land</i>)
3. Restructuring approach (epistemological criticism) that sometimes becomes radical	To question the nature of knowledge itself and to promote the birth of a new concept and organization of scientific knowledge Substitute another structure for the disciplinary structure (for example deconstructionist criticism) (<i>tabula rasa</i>)

other things, a strong interaction between curricular, didactic and pedagogical plans. Indeed, an interdisciplinary approach requires an integrative articulation between the curriculum and classroom practice (Fig. 3).

The transition from the curricular plan to the pedagogical plan or, to put it another way, to the practice of teaching in the classroom, requires the use of the mediation of the didactics of the school disciplines. The pedagogical time of the teacher's updating in class requires to understand the aims pursued by each school discipline, their specificities and the mutual contributions. But it is also important, with this understanding in mind, the design of teaching-learning activities uses didactic dimensions that ensure that learning objects are arranged in a way that makes them accessible to students.

As a fifth principle we call the "principle of integrated logics", and as we have already reported, we must take into account three perspectives to approach teaching and learning: meaning, functionality, and affectivity. These three perspectives are complementary and should be interwoven when using interdisciplinary approaches in teaching practices. Otherwise put, interdisciplinarity requires the mind, the hand, and the heart, meaning a strong interrelationship between the cognitive (or epistemological), practical (or pragmatic), and human (ontological and social) dimensions, which, in our view, are fundamental components of human beings and should be at the core of any educational process.

Finally, the sixth and final principle we would like to suggest is that it is no longer possible today to truly master several disciplines. The era of Giovanni Pico della Mirandola, who in the 15th century set out to conquer all existing knowledge, is long gone. What is needed instead is to pool disciplinary competencies. Consistent with Bernstein (1971a,b), an interdisciplinary approach based on integrated intervention by a teaching team will tend toward an integrating curriculum, characterized by a system of open contents and an "ample education", as opposed to a compartmentalized curriculum based on a hierarchy of disciplines and a rigid and opaque system of closed-off contents. This "principle of collaboration" requires organizational management which is often the real weakness. An interdisciplinary team must not only receive institutional recognition and real support from its superiors, but also a leader who is respected, rigorous and demanding of the team members, who is also able to perform several

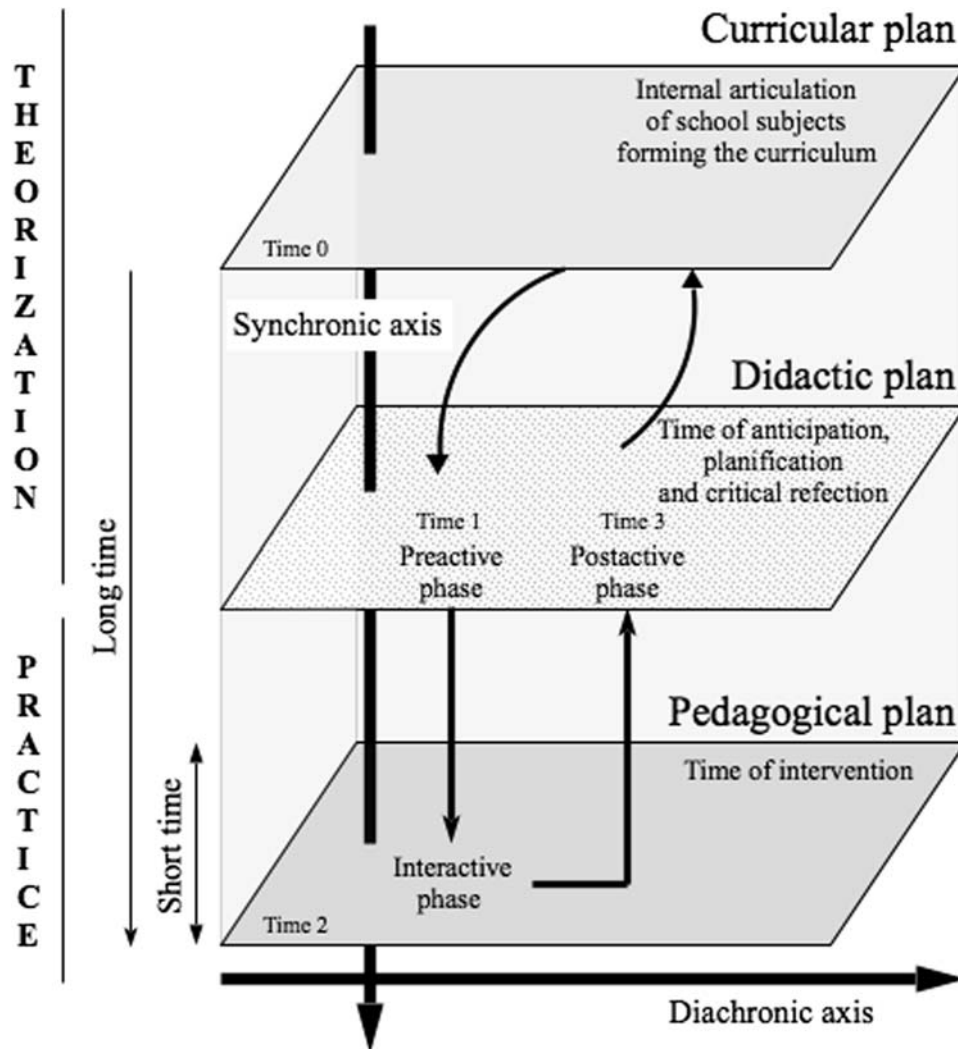


Fig. 3 Relations between curricula, didactic and pedagogical plans.

functions: *ringmaster, bridge specialist, gatekeeper, boundary agent, ombudsman, polymath, dynamo, metascientist, specialist/generalist, strong entrepreneur* (Klein, 1990; Klein and Porter, 1990). This leader must also have the necessary competencies to manage team teaching and pedagogical projects, and to regulate actions and interactions (regular fixed, individual and group meetings). Numerous authors (Crawford Burns, 2002; Davis, 1995; Fourez, 1994; 1998; Hasni et al., 2008; Klein and Newell, 1996; Wentworth and Davis, 2002; Wineburg and Grossman, 2000) have shed light on the need for team teaching, which must be undertaken in a spirit of genuine collaboration. In interdisciplinary education, team teaching is required because solving a problem, addressing a question, or carrying out a project all require complementary processes, cognitive contents from different disciplines, etc. Crawford Burns (2002) stresses that « interdisciplinary teams should consider the needs and interests of students as well as the essential content and processes in each discipline and their application to authentic tasks » (p. 68).

Conclusion

In light of the various clarifications we have suggested regarding interdisciplinarity, we would like to propose the following definition: school interdisciplinarity is the action of putting two or more school disciplines into relation on the curricular, cognitive, and operational levels, leading to the establishment of complementarity or cooperative links, interpenetration, or reciprocal action from various standpoints (the aims, objects of study, concepts and notions, learning approaches, technical abilities, etc.). These interactions are aimed at promoting human beings' integration of processes and knowledge (Lenoir and Sauvé, 1998b). In teaching-learning practices, the teacher's role is to establish the conditions judged to be the best, the most appropriate, to promote and support student learning processes. To draw on interdisciplinarity at school is to introduce conditions that are normally favorable to human beings' implementation of integrating processes by mobilizing various and interrelated disciplinary angles. For, indeed, it is not the teacher who must do the integrating, but the students. They are the actors who will do the integrating, insofar as the two following basic conditions are met: first, examining the principal elements of school disciplines and their structure (conceptual knowledge and disciplinary processes) with a view to identifying interdisciplinary perspectives; and, second, implementing teaching-learning situations that require the construction of conceptual knowledge from two or more school disciplines and that require the complementary use of different learning methods.

As a result, rather than tending toward a merging of disciplines, an eclectic addition of knowledge elements, a thematic pseudo-interdisciplinarity, or a hegemony of one school discipline over the others, interdisciplinary practices involve relationships that foster mutual dependence between disciplines, free of any dominance or neglect. They are based on a consideration of the educational aims pursued, acknowledging the wealth of complementary links and of real and essential interrelationships between cognitive contents and methods, which are necessary in order to construct, express and interact with human, social, and natural reality. In a school context, interdisciplinary perspective therefore requires an open relationship between cognitive and practical dimensions, but this relationship must also introduce affective dimension. This brings us back to the idea of the mind, the hand and the heart, insofar as reason (reflected-on knowledge) mediates between the heart (empirical human dimension) and the hand (operational dimension).

Owing to its specific features, the school interdisciplinarity requires that the knowledge stemming from different school subjects be explicitly addressed and treated to enable learning in connection with various learning situations (problems, projects, compelling scientific or current societal question, etc.). In short, it is important to take into account the purposes and the specific identities of school disciplines in any interdisciplinary classroom approach. This is why school interdisciplinarity can under no circumstances be an end in itself: « on an educational level, one can no longer interpret interdisciplinarity without referring to the concept of integration » (Lenoir and Klein, 2010, p. xviii). This is because its aim is to enable students to develop integrative cognitive processes and the cognitive integration of the knowledge that is gained. We thus consider that the notion of integration must include a dynamic and constructive outlook on learning, one that is meaningful and necessitates the use of cognitive elements from different school disciplines, as well as complementary and interrelated inquiry processes across disciplines. Moreover, it requires the realization of a set of principles that ensure it will pursue objectives that are complex, but accessible to students thanks to the mediation of teachers.

It must be kept in mind that the foremost mission of basic (primary and secondary) education is to educate individuals who are capable of living and finding fulfillment in society, while being able to understand the world in which they live and being able to act on it in a thought-out and critical way. As a result, even if they must adapt to the realities and requirements of our times, primary and secondary school curricula should not fall into an exacerbated individualistic utilitarianism in their attempt to provide young generations with basic human and social education that is rich, diversified, and open. In sum, we must conceive the education of young people from an emancipatory perspective and it is our belief that the inter- and transdisciplinary approaches can significantly contribute to reaching this goal.

Acknowledgment

Our sincere thanks go to the Social Sciences and Humanities Research Council of Canada (SSHRC), which has funded our researches for 30 years, and to all colleagues who have collaborated with us and to all our former Master's and Ph.D. students who have been actively involved in our research activities and subsequent publications.

References

- Ansart, P., 1990. Pluralisation des savoirs et formation scientifique. In: Racette, G., Forest, L. (Eds.), *Pluralité des enseignements en sciences humaines à l'université*. Éditions Noir sur blanc, Montreal, pp. 21–29.
- Apostel, L., Berger, G., Briggs, A., Michaud, G. (Eds.), 1972. *L'interdisciplinarité. Problèmes d'enseignement et de recherche dans les universités*. OCDE/Centre pour la recherche et l'innovation dans l'enseignement, Paris.
- Baillat, G., Niclot, D., 2010. In search of interdisciplinarity in schools in France: from curriculum to practice. *Issues Integr. Stud.* 28, 170–207.
- Barbier, R., 1997. *L'approche transversale. L'écoute sensible en sciences humaines*. Anthropos, Paris.
- Beane, J.A., 1997. *Curriculum Integration. Designing the Core of Democratic Education*. Teachers College, New York, NY.
- Beillerot, J., 1989. Le rapport au savoir: une notion en formation. In: Beillerot, J., Bouillet, A., Blanchard-Laville, C., Mosconi, N. (Eds.), *Savoir et rapport au savoir. Élaborations théoriques et cliniques*. Éditions universitaires, Paris, pp. 165–202.
- Bernstein, B., 1971a. *Class, Codes and Control*. Routledge & Kegan Paul, London.
- Bernstein, B., 1971b. On the classification and framing of educational knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control. New Directions for the Sociology of Education*. Collier-Macmillan, London, pp. 47–69.
- Bisault, J., 2011. Contribution à l'élaboration curriculaire d'une éducation scientifique à l'école primaire. Modélisation des moments scolaires à visée scientifique. Mémoire pour l'habilitation à diriger des recherches, École normale supérieure de Cachan, Paris.
- Boix-Mansilla, V., Lenoir, Y., 2010. Interdisciplinarity in United States schools: past, present, and future. *Issues Integr. Stud.* 28, 1–28.
- Boix-Mansilla, V., 2017. Learning to synthesize: the development of interdisciplinary understanding. In: Frodeman, R., Klein, J.T., Pacheco, R.C.S. (Eds.), *The Oxford Handbook of Interdisciplinarity*. Oxford University Press, Oxford, pp. 261–279.
- Bourdieu, P., 1980. *Le Sens Pratique*. Éditions de Minuit, Paris.
- Brousseau, G., 1998. *Théorie des situations didactiques. La Pensée sauvage*, Grenoble.
- Charlot, B., 1997. *Du rapport au savoir. Éléments pour une théorie*. Economica-Anthropos, Paris.
- Clarke, J.H., Agne, R.M. (Eds.), 1997. *Interdisciplinary High School Teaching: Strategies for Integrated Learning*. Allyn and Bacon, Boston, MA.
- Clausen, K., Drake, S., 2010. Interdisciplinary practices in Ontario: past, present, and future. *Issues Integr. Stud.* 28, 69–108.
- Crawford Burns, R., 2002. Interdisciplinary teamed instruction. In: Klein, J.T. (Ed.), *Interdisciplinarity Education in K-12 and College. A Foundation for K-16 Dialogue*. College Board Publications, New York, NY, pp. 47–69.
- Darbellay, F., Louviot, M., Moody, Z. (Eds.), 2019. *L'interdisciplinarité à l'école. Succès, résistance, diversité*. Éditions Alphil-Presses universitaires suisses, Neuchâtel.
- Davis, J.R., 1995. *Interdisciplinary Courses and Team Teaching. New Arrangements for Learning*. American Council on Education/The Oryx Press, Phoenix, AZ.
- D'Hainaut, L., 1986. L'interdisciplinarité dans l'enseignement général (Étude de Louis D'Hainaut à la suite d'un colloque international sur l'interdisciplinarité dans l'enseignement général organisé à la Maison de l'Unesco du 1er au 5 juillet 1985). Unesco, Division des sciences de l'éducation, des contenus et des méthodes, Paris.
- Durkheim, M., 1969/1938. *L'évolution pédagogique en France*, first ed. Presses universitaires de France, Paris.
- Fazenda, I.C.A., 1994. *Interdisciplinariedade: História, teoria e pesquisa*. Papirus Editora, Campinas, SP.
- Fazenda, I.C.A., 1995. Critical-historical review of interdisciplinary studies in Brazil. *Assoc. Integr. Stud. Newsl.* 17 (1), 1–9.
- Fazenda, I.C.A., 1998. La formation des enseignants pour l'interdisciplinarité: synthèse de recherches effectuées au Brésil. *Rev. Sci. Éduc.* XXIV (1), 95–114.
- Fogarty, R., 1991. *How to Integrate the Curricula*. IRI/Skylight Publishing, Palatine, IL.
- Fogarty, R., 1993. *Integrating the Curricula*. IRI/Skylight Publishing, Palatine, IL.
- Fourez, G., Maingain, A., Dufour, B., 2002. *Approches didactiques de l'interdisciplinarité*. De Boeck Université, Bruxelles.
- Fourez, G., 1994. Alphabétisation scientifique et technique. Essai sur les finalités de l'enseignement des sciences. De Boeck Université, Bruxelles.
- Fourez, G., 1998. Se représenter et mettre en œuvre l'interdisciplinarité à l'école. *Rev. Sci. Éduc.* XXIV (1), 31–50.
- Frazee, B., Rudnitski, R.A., 1995. *Integrated Teaching Methods: Theory, Classrooms Applications, Field-Based Connections*. Delmar, Albany, NY.
- Gergen, K.J., 1994. *Realities and Relationships*. Harvard University Press, Cambridge, MA.
- Gergen, K.J., 1995. Social construction and the educational process. In: Steffe, L.P., Gale, J. (Eds.), *Constructivism in Education*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 17–39.
- Gosselin, M., Lenoir, Y., Hassani, N., 2005. La structuration par domaines du nouveau curriculum de l'enseignement primaire: une analyse critique du modèle retenu. In: Lenoir, Y., Larose, F., Lessard, C. (Eds.), *Le curriculum de l'enseignement primaire: regards critiques sur ses fondements et ses lignes directrices*. Éditions du CRP, Sherbrooke, pp. 169–200.
- Gusdorf, G., 1967. *Les sciences de l'homme sont-elles des sciences humaines?* Université de Strasbourg, Faculté des lettres, Strasbourg.
- Hamel, J., Orival, T., 2017. Qu'est-ce exactement que la transdisciplinarité? In: Caron, M.-A., Turcotte, M.-F. (Eds.), *La transdisciplinarité et l'opérationnalisation des connaissances scientifiques*. Éditions JFD, Montréal, pp. 23–40.
- Hasni, A., Lenoir, Y., Larose, F., Samson, G., Bousadra, F., Dos Santos, C., 2008. Enseignement des sciences et technologies et interdisciplinarité: point de vue d'enseignants du secondaire au Québec. In: Hasni, A., Lebeaume, J. (Eds.), *Interdisciplinarité et enseignement scientifique et technologique*. Éditions du CRP/INRP, Sherbrooke-Lyon, pp. 75–110.
- Hasni, A., 2011. Problématiser, contextualiser et conceptualiser en sciences: point de vue d'enseignants du primaire sur leur pratique de classe. In: Hasni, A., Baillat, G. (Eds.), *Pratiques d'enseignement des sciences et technologies: regards sur la mise en œuvre des réformes curriculaires et sur le développement des compétences professionnelles des enseignants*. Éditions et presses universitaires de Reims, Reims, pp. 105–140.
- Haynes, C., 2002. Laying a foundation for interdisciplinary teaching. In: Haynes, C. (Ed.), *Innovations in Interdisciplinary Teaching*. The Oryx Press/American Council on Education, Westport, CT, pp. xi–xxii.
- Hegel, G.W.F., 1941/1807. In: Hippolyte, J. (Ed.), *La phénoménologie de l'esprit* (2 t.), first ed. Aubier-Montaigne, Paris.
- Hegel, G.W.F., 1969/1805. *La première philosophie de l'esprit* (Iéna, 1803–1804), first ed. Presses universitaires de France, Paris.
- Henry, N.B. (Ed.), 1958. *The Integration of Educational Experiences: The Fifty-Seventy Yearbook of the National Society for the Study of Education*. The University of Chicago Press, Chicago, IL.
- Hermerén, G., 1985. Interdisciplinarity revisited—promises and problems. In: Levin, L., Lind, I. (Eds.), *Interdisciplinarity Revisited: Re-assessing the Concept in the Light of Institutional Experience*. Oecd/Ceri, Swedish National Board of Universities and Colleges, Linköping University, Stockholm, pp. 15–25.
- Honneth, A., 2013. Du désir à la reconnaissance. La fondation hégélienne de la conscience de soi. In: Honneth, A. (Ed.), *Ce que social veut dire. I. Le déchirement du social*. Gallimard, Paris, pp. 84–107.
- Hopkins, L.T., 1937. *Integration: Its Meaning and Application*. Appleton-Century, New York, NY.
- Jacobs, H.H. (Ed.), 1989. *Interdisciplinary Curriculum: Design and Implementation*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Klein, J.T., Newell, W.H., 1996. Advancing interdisciplinary studies. In: Gaff, J.G., Ratcliff, J.L., Associates (Eds.), *Handbook of the Undergraduate Curriculum. A Comprehensive Guide to Purposes, Structures, Practices, and Change*. Jossey-Bass, San Francisco, CA, pp. 393–415.
- Klein, J.T., Porter, A.L., 1990. Preconditions for interdisciplinary research. In: Birnbaum-More, P.H., Rossini, F.A., Baldwin, D.R. (Eds.), *International Research Management. Studies in Interdisciplinary Methods from Business, Government, and Academia*. Oxford University Press, Oxford, pp. 11–19.
- Klein, J.T., Grossenbacher-Mansuy, W., Häberli, R., Bill, A., Scholz, R.W., Welti, M., 2001. *Transdisciplinarity: Joint Problem Solving Among Science, Technology, and Society*. Birkhäuser Verlag, Basel.

- Klein, J.T., 1990. *Interdisciplinarity: History, Theory, and Practice*. Wayne State University Press, Detroit, MI.
- Klein, J.T., 1996. *Crossing Boundaries Knowledge, Disciplinarity, and Interdisciplinarity*. University Press of Virginia, Charlottesville, VI.
- Klein, J.T., 2002. Introduction: interdisciplinarity today: why? what? and how? In: Klein, J.T. (Ed.), *Interdisciplinarity Education in K-12 and College. A Foundation for K-16 Dialogue*. College Board Publications, New York, NY, pp. 1–17.
- Kliebard, H.M., 1992a. *Forging the American Curriculum. Essays in Curriculum History and Theory*. Routledge, New York, NY, London.
- Kliebard, H.M., 1992b. Constructing a history of the American curriculum. In: Jackson, P.H. (Ed.), *Handbook of Research on Curriculum. A Project of the American Educational Research Association*. Macmillan, New York, NY, pp. 157–184.
- Kojève, A., 1947. *Introduction à la lecture de Hegel. Leçons sur la phénoménologie de l'esprit professées de 1933 à 1939 à l'École des Hautes-Études réunies et publiées par Raymond Queneau*. Gallimard, Paris.
- Lacan, J., 1966. *Écrits I*. Seuil, Paris.
- Lawrence, R.J., 2019. Convergences: défis disciplinaires et réponses transdisciplinaires. In: Darbellay, F., Louviot, M., Moody, Z. (Eds.), *L'interdisciplinarité à l'école. Succès, résistance, diversité*. Éditions Alphil-Presses universitaires suisses, Neuchâtel, pp. 49–67.
- Lenoir, Y., Hasni, A., 2004. La interdisciplinaridad: por un matrimonio abierto de la razón, de la mano y del corazón. *Ibero Am. J. Educ.* 35, 167–185.
- Lenoir, Y., Hasni, A., 2010. Interdisciplinarity in Quebec schools: 40 years of problematic implementation. *Issues Integr. Stud.* 28, 238–294.
- Lenoir, Y., Klein, J.T. (Eds.), 2010. *Interdisciplinary in schools: a comparative view of national perspectives*. *Issues Integr. Stud.* 28, 1–331.
- Lenoir, Y., Sauvé, L., 1998a. De l'interdisciplinarité scolaire à l'interdisciplinarité dans la formation à l'enseignement: un état de la question. 1—Nécessité de l'interdisciplinarité et rappel historique. *Rev. Fr. Pédagog.* 124, 121–153.
- Lenoir, Y., Sauvé, L., 1998b. De l'interdisciplinarité scolaire à l'interdisciplinarité dans la formation à l'enseignement: un état de la question. 2—interdisciplinarité scolaire et formation interdisciplinaire à l'enseignement. *Rev. Fr. Pédagog.* 125, 109–146.
- Lenoir, Y., Larose, F., Grenon, V., Hasni, A., 2000. La stratification des matières scolaires chez les enseignants du primaire au Québec: évolution ou stabilité des représentations depuis 1981? *Rev. Sci. Éduc.* XXVI (3), 483–514.
- Lenoir, Y., Rey, B., Fazenda, I. (Eds.), 2001. *Les fondements de l'interdisciplinarité dans la formation à l'enseignement*. Éditions du CRP, Sherbrooke.
- Lenoir, Y., Larose, F., Dirand, J.-M., 2006. Formation professionnelle et interdisciplinarité: quelle place pour les savoirs disciplinaires? In: Fraysse, B. (Ed.), *Professionalization des élèves ingénieurs*. Éditions L'Harmattan, Paris, pp. 13–35.
- Lenoir, Y., Hasni, A., Froelich, A., 2015. Curricular and didactic conceptions of interdisciplinarity in the field of education: a socio-historical perspective. *Issues Interdiscip. Stud.* 33, 39–93.
- Lenoir, Y., Adigüzel, O., Lenoir, A., Libâneo, J.C., Tupin, F. (Eds.), 2016. *Les finalités éducatives scolaires. Une étude critique des approches théoriques, philosophiques et idéologiques. T. 1: Fondements, notions et enjeux socioéducatifs*. Éditions Cursus universitaire, Saint-Lambert.
- Lenoir, Y., 1999. Interdisciplinarité. In: Houssaye, J. (Ed.), *Questions Pédagogiques. Encyclopédie Historique*. Hachette, Paris, pp. 291–314.
- Lenoir, Y., 2000. Formation à l'enseignement et interdisciplinarité: un mythe ou une exigence? Dépasser l'interdisciplinarité et penser circumdisciplinarité. *Eur. J. Teach. Educ.* 23 (3), 289–298.
- Lenoir, Y., 2002. Les réformes actuelles de la formation à l'enseignement en France et aux États-Unis: éléments de mise en perspective socio-historique à partir du concept d'éducation. *Swiss J. Educ. Sci.* 24 (1), 91–126.
- Lenoir, Y., 2005. Três interpretações da perspectiva interdisciplinar em educação em função de três tradições culturais distintas. *E-Curric. Rev.* 1 (1), 1–24.
- Lenoir, Y., 2017/2014. *Les médiations au cœur des pratiques d'enseignement-apprentissage: une approche dialectique. Des fondements à leur actualisation en classe. Éléments pour une théorie de l'intervention éducative (second revised and expanded edition)*, first ed. Éditions Cursus universitaire, Saint-Lambert.
- Lenoir, Y., 2019. De l'interdisciplinarité scientifique à l'interdisciplinarité scolaire. In: Darbellay, F., Louviot, M., Moody, Z. (Eds.), *L'interdisciplinarité à l'école. Succès, résistance, diversité*. Éditions Alphil-Presses universitaires suisses, Neuchâtel, pp. 7–48.
- Long, J., Moran, W., Harris, J., 2010. Following the yellow brick road: interdisciplinary practices in the land of Oz. *Issues Integr. Stud.* 28, 28–68.
- Lordon, F., 2010. *Capitalisme, désir et servitude. Marx et Spinoza*. La Fabrique, Paris.
- Lordon, F., 2011/2006. *L'intérêt Souverain. Essai D'anthropologie Économique Spinoziste*, first ed. La Découverte, Paris.
- Lordon, F., 2013. *La société des affects. Pour un structuralisme des passions*. Seuil, Paris.
- Martinand, J.-L., 1986. *Connaître et transformer la matière*. Peter Lang, Berne.
- Miñana Blasco, C. (Ed.), 2001. *Interdisciplinarity and curriculum. Construcción de proyectos escuela-universidad*. Universidad Nacional de Colombia, Bogotá.
- Morin, E., 1977. *La méthode. T. 1: La nature de la nature*. Seuil, Paris.
- Morin, E., 1990. *Science Avec Conscience*, second ed. Seuil, Paris.
- Newell, W.H., 1990. Interdisciplinary curriculum development. *Issues Integr. Stud.* 8, 69–86.
- Newell, W.H., 1998. Interdisciplinary curriculum development. In: Newell, W.H. (Ed.), *Interdisciplinarity. Essays from the Literature*. The College Board, New York, NY, pp. 51–65.
- Newell, W.H., 2001. A theory of interdisciplinary studies. *Issues Integr. Stud.* 19, 1–26.
- Nicolescu, B., 1996. *La transdisciplinarité*. Éditions du Rocher, Monaco.
- Pastré, P., 2011. *La didactique professionnelle. Approche anthropologique du développement chez les adultes*. Presses universitaires de France, Paris.
- Pate, E.P., Homestead, E.R., McGinnis, K.L., 1997. *Making integrated curriculum work. Teachers, students, and the quest for coherent curriculum*. Teachers College, New York, NY.
- Petrie, H.G., 1992. Interdisciplinarity education: are we face with insurmountable opportunities? *Rev. Res. Educ.* 18, 299–333.
- Pinar, W.F., 1998. *Dreamt into existence by others: notes on school reform in the US*. In: Tardif, M., Lessard, C., Gauthier, C. (Eds.), *Formation des maîtres et contextes sociaux. Perspectives internationales*. Presses universitaires de France, Paris, pp. 201–229.
- Poincaré, H., 1917. *La science et l'hypothèse*. Ernest Flammarion, Paris.
- Pollard, A., Tann, S., 2002. *Reflective Teaching in the Primary School. A Handbook for the Classroom*. Cassell Education, London.
- Resweber, J.-P., 1981. *La Méthode Interdisciplinaire*. Presses universitaires de France, Paris.
- Reuter, Y., 2019. Didactiques, disciplines et interdisciplinarité. In: Darbellay, F., Louviot, M., Moody, Z. (Eds.), *L'interdisciplinarité à l'école. Succès, résistance, diversité*. Éditions Alphil-Presses universitaires suisses, Neuchâtel, pp. 133–149.
- Reverdy, C., 2019. Interdisciplinarité et projets collectifs: vers une nouvelle professionnalité enseignante? In: Darbellay, F., Louviot, M., Moody, Z. (Eds.), *L'interdisciplinarité à l'école. Succès, résistance, diversité*. Éditions Alphil-Presses universitaires suisses, Neuchâtel, pp. 197–214.
- Rodriguez, J.G., Miñana Blasco, C., 2010. Interdisciplinary and research on local issues in schools: policies and experiences from Colombia. *Issues Integr. Stud.* 28, 109–137.
- Rudolph, F., 1977. *Curriculum. A history of the American undergraduate course of study since 1636*. Jossey-Bass, San Francisco, CA.
- Sachot, M., 2001. L'interdisciplinarité entre disciplines et curriculum: retour sur un impensé en matière de formation. In: Lenoir, Y., Rey, B., Fazenda, I. (Eds.), *Les fondements de l'interdisciplinarité dans la formation à l'enseignement*. Éditions du CRP, Sherbrooke, pp. 51–66.
- Sachot, M., 2004. Disciplines du maître, disciplines de l'élève: contre une « disciplinarisation » du primaire. In: Sachot, M., Lenoir, Y. (Eds.), *Les enseignants du primaire entre disciplinarité et interdisciplinarité: quelle formation didactique?* Les Presses de l'Université Laval, Québec, pp. 19–33.
- Segovia, I., Lupiáñez, J.L., Molina, M., González, F., Miñán, A., Real, I., 2010. The conception and role of interdisciplinarity in the Spanish education system. *Issues Integr. Stud.* 28, 208–237.
- Spinoza, B., 1849/1667. In: Saisset (Translator), *L'éthique. Ille partie: De l'origine et la nature des passions*, first ed., pp. 81–139 Available at: <http://www.spinozaetnous.org>.
- Stengers, I., 1993. *L'invention des sciences modernes*. La Découverte, Paris.
- Stichweh, R., 1991. In: Blaise, F. (Translator), *Études sur la genèse du système scientifique moderne*. Presses universitaires de Lille, Lille.

- Tann, S. (Ed.), 1988. *Developing Topic Work in the Primary School*. The Falmer Press, London, Philadelphia, PA.
- Tanner, D., Tanner, L., 1990. *History of the School Curriculum*. Macmillan/Collier Macmillan, New York, NY, London.
- Torres Santomé, J., 1998. *Globalización e interdisciplinariedad: el currículo integrado*. Ediciones Morata, Madrid.
- Vars, G.F., 1993. *Interdisciplinary Teaching: Why & How*. National Middle School Association, Columbus, OH.
- Vygotsky, L.S., 1987. The collected works of L. S. Vygotsky. In: Minick (Translator), N., Rieber, R.W., Carton, A.S. (Eds.), *Problems of General Psychology Including the Volume Thinking and Speech*, vol. 1. Plenum Press, New York, NY.
- Vygotsky, L.S., 2014. In: Sève (Translator), F., Brossard, M., Sève, L. (Eds.), *Histoire du développement des fonctions psychiques supérieures*. La Dispute, Paris.
- Wentworth, J., Davis, J.R., 2002. Enhancing interdisciplinarity through team teaching. In: Haynes, C. (Ed.), *Innovations in Interdisciplinary Teaching*. The Oryx Press/American Council on Education, Westport, CT, pp. 16–37.
- Wineburg, S., Grossman, P., 2000. *Interdisciplinary Curriculum. Challenges to Implementation*. Teachers College Press, New York, NY.
- Young, M.F.D., 1971. An approach to the study of curricula as socially organized knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control*. New Directions for the Sociology of Education. Collier Macmillan, London, pp. 19–46.

Curriculum and competences

Judith Glaesser, Tübingen School of Education and Methods Center, University of Tübingen, Tübingen, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	281
The changing role of competences in the curriculum	281
Competences in a differentiated curriculum and in mass education	282
Assessment	283
Competence and standards	283
Standards as an instrument in accountability and monitoring performance	284
Competences internationally: the PISA study	285
Conclusion	286
Acknowledgment	287
References	287
Further reading	287

Introduction

Most people have an intuitive sense of what competence means and whom they consider to be acting competently, and they would like professionals or tradespeople to be competent in their areas of work. Given that (school) education is increasingly seen to be the means of preparing young people for the workplace, school curricula are expected to include the provision of the required competences. In this chapter, I give an overview of how the notion of competence has varied across different contexts and how its importance within curricula has changed over time. In doing so, I explore the role of a differentiated curriculum and the related topic of mass education, the connection between competence and assessment within the curriculum, the way in which the notion of competence is related to standards and how this in turn relates to the rising importance of accountability and monitoring in the education system, and I discuss the role of PISA as an example of the international prominence of competence. In the concluding section, I draw these strands together and I evaluate the place of competence in modern curricula.

The changing role of competences in the curriculum

There are numerous definitions of competence (and they can vary by context; see Carr, 1993). I do not intend to discuss their relative merits, but I would like to draw on Messick's (1984) description since it is interesting in the context of the curriculum:

Educational achievement essentially refers to what an individual knows and can do in a specified subject area as a consequence of instruction. Educational achievement should be distinguished from the closely related construct of competence, which refers to what an individual knows and can do in a subject area however that knowledge and skill is acquired, whether through instruction or experience or whatever. Like achievement, the term "competence" as used here refers to a continuous variable reflecting various degrees of proficiency rather than to its other common usage as a particular standard of adequate or sufficient performance. (p. 217)

Several aspects of this are worth noting. To begin with, competence here clearly is not reliant on teaching; instead, it is defined, in contradistinction to educational achievement, as a skill "however ... acquired". This does not preclude attaining competence as a result of schooling, but schooling is not necessary. This conceptualization of competence is to be found also, for example, in the UK's vocational education framework NVQ (National Vocational Qualification) where, in principle, certificates are to be conferred on the basis of demonstration of competence rather than on the basis of how much time had been spent in an educational institution (Wolf, 2000) (though problems arise in the implementation of this principle, but this is another matter). Another issue to highlight in Messick's definition is the distinction of competence (and educational achievement) from knowledge: competence relies on knowledge and goes beyond it since being able to act on one's knowledge is also required for a person to be deemed competent. A final point is the mention of standards which will be discussed below since, unlike in Messick's usage, competences actually are central to establishing standards in many contexts today.

Competence has not traditionally been a central goal of school teaching. Indeed, Hirst (1969) suggested that "[t]he objectives of education are surely certain developments of the pupil which are achieved in learning, and I suggest that these are all basically connected with the development of a rational mind." (p. 147) However, competence is not a new concept within formal education. Traditionally, it has been more commonly associated with vocational education and teacher education (Carr, 1993; Whitty and Willmott, 1991), having been used within teacher education as early as the 1920s by Ralph Tyler, as Norris (1991) points out.

Nowadays, it is one of the central features in general education curricula of many countries: schools are expected to produce competent individuals who are able to function in the workplace. Thus, the role of the education system is seen as to improve employability and to increase economic growth (for example, Wahlström and Sundberg, 2018a; Young, 2013), with a curriculum centered on competences being a means to foster such an improvement. Indeed, Wahlström and Sundberg (2018a) explicitly link these two phenomena: (1) the notion of employability as the aim of education and (2) the emergence of competence in the curriculum across many countries (p. 9). However, the nature of the link between education and economic growth is not as clear-cut as it is sometimes made out to be. It is clearly the case that in modern society, highly educated individuals on average tend to earn more than less highly educated ones, but it does not follow that expanding the time spent in education and/or changing the nature of education will result in economic growth and more prosperity for all (Wolf, 2002, 2004), nor that, for an individual, a high level of education is a guarantee of an attractive job (Bukodi and Goldthorpe, 2019).

In the next section, I explore further the notion of education as preparation for the workplace, discussing how this is linked to a differentiated curriculum as opposed to mass education for all, and what role competences play in this.

Competences in a differentiated curriculum and in mass education

As noted in the previous section, Hirst's view is that the overarching objective of education is the development of a rational mind in each child (Hirst, 1969), and that this applies regardless of the child's ability. Content and teaching methods may have to be adapted in various ways to make it possible to achieve this objective, but he evidently does not support the view that less able children should be taught mere skills (as opposed to knowledge and understanding) in order to function in a workplace where they are expected to carry out routine tasks. However, such a view has been expressed at least implicitly in various curricula. For example, Cooper (1985) analyses the history of the mathematics curriculum in England, discussing how in this subject (as in others), the curriculum has been differentiated according to school type and the associated type of career children of different abilities (and, frequently, different social backgrounds) can be expected to have. In such a curriculum, abstract knowledge and thinking is the preserve of the most able, and the others are to be taught skills which will enable them to function in simple jobs and in their daily lives. He quotes handbooks on Modern Schools (the most basic type of school in the old English tripartite system) which stress arithmetic at the expense of mathematics. Teachers he interviewed argued similarly that the role of the school for less able children was limited to providing them with the basic skills they could expect to need once they left school. The most highly academically selective grammar schools, by contrast, offered a mathematics curriculum which encouraged learning for its own sake as well as preparation for university study.

We can see that it is not a new idea for the school's role to be to prepare young people for the workplace. Historically, this link was made explicit only for the least educated groups who had to learn basic skills at school so that they would be able to cope with fairly low-level demands. But at the other end of the scale, while learning for its own sake was stressed, the most selective schools also prepared their pupils for their future careers, even if this was not their stated aim. The difference was in the type of career they were expected to pursue. Indeed, "learning for its own sake" does not necessarily have to mean that such learning has no purpose at all other than private enjoyment. Dore (1976, p. 8) describes why one of the reasons learning and knowledge may be appreciated is because they can enable the learner to do a better job. He contrasts "learning to *do* a job" with "learning to *get* a job" (emphasis in the original). The goal of learning in the former case is to acquire the knowledge and/or skills necessary to carry out a task in a satisfactory way; in the latter, learning is undertaken in order to gain a certificate which will then enable the bearer to obtain a job or access to a course of more advanced education (which can in turn lead to yet another certificate). In this latter situation, the learner may or may not be particularly interested in the actual content to be learned. Understood in this way, "learning for its own sake" is not just a way of describing the goal of learning; instead, it embodies a certain ideal way of living one's life (see also his discussion of the history of education and qualifications in England in his chapter 1).

Throughout the second half of the 20th century and into the 21st, alongside the expectation that education should prepare young people for the labor market, the role of the education system was seen by some parts of society as being a means of providing children of all backgrounds with equal opportunities. This is captured in the British Labour Party's slogan from the early 1960s which was "grammar schools for all", indicating that the type of education associated with advantageous subsequent careers available to former grammar school pupils should be made accessible for everyone. However, Cooper's analysis shows that this was not seen as feasible by many teachers, who assumed that the tripartite system reflected the fact that there really is a wide range of abilities and that not all children are equally well suited to studying the type of abstract content associated with the grammar school curriculum. By that time, England had moved to a largely comprehensive system, but the expectation that not all teaching content was equally suitable for children of different abilities remained in place nevertheless. This was also reflected in the way teaching was organized: many—if not most—comprehensive schools continue to employ some form of ability grouping at least for some subjects. The teaching content is differentiated within these groups (termed bands, streams or sets) according to the students' (assumed or measured) ability. More recently, Young (2013, p. 112) made a similar point noting that many teachers expect that only a minority of children would be able to benefit from a curriculum based on what he terms "powerful knowledge", i.e., the sort that makes access to privileged positions in society possible. Thus, differences still exist as to what type of content is taught to which children even in a notionally comprehensive system.

However, the expectation that it is the role of schools to provide a well qualified workforce is as strong as ever. According to Young (2013), the expansion of education seen in many countries from the end of the Second World War onwards is a result of

pressure from “global capitalism” (p. 106), and the face of education has changed accordingly. Thus, while it may be the case that children in developed countries generally have access to schooling for an extended period of time, this has not had the liberating effect proponents of increased access to education might have hoped it would have, because, in Young’s analysis, the demands arising from the economy and the pressure to acquire employability have all but replaced learning for its own sake and its “emancipatory role”. “Learning for its own sake” is “dismissed as elitist” when the goal of education is employability (p. 106, see also Rosenmund, 2007; Wolf, 2002).

Thus, while “knowledge for its own sake” has in fact always constituted a form of preparation for a child’s future career, the idea of acquiring such knowledge seems to have gone out of fashion altogether anyway, to be replaced by the proclaimed need for different forms of competence which will enable young people to participate in the modern labor market. Given the rapidly changing demands placed on modern employees, some policy-makers argue (as reported, e.g., by Rosenmund, 2007) that knowledge becomes obsolete too quickly and, therefore, schools should aim to offer more general competences instead which enable individuals to identify quickly any information that is needed to carry out a task, to locate this information and to apply it to the relevant context. This is a long way from Hirst’s (1969) notion of the development of a rational mind as the universal goal of schooling—he assumed that the acquisition of different types of knowledge was suited to this goal—though there are some parallels: in principle, a rational mind can be applied to any problem that is open to being solved rationally. Individuals with the capability to think rationally can transfer their knowledge to new contexts. But it seems that this is not currently pursued as a central goal of education (if it ever was). At the same time, as with older forms of a differentiated curriculum where “knowledge for its own sake” was taught to some and basic practical skills to others, young people today are likely to acquire different forms and levels of competence in the education system, as Cooper’s and Young’s analyses have shown. And, also as with older forms of a differentiated curriculum, they can expect to have different types of career. In that sense, employability is also differentiated: most young people who leave school are employable, but not necessarily in the same types of jobs.

Assessment

Rosenmund (2007, see previous section) analyzed the shift from a curriculum centered on knowledge to one focusing more on competences. Given that assessment is a central part of teaching and learning, and given that any assessment has to be suited to the content to be assessed, it may be necessary to develop new forms of assessment to match the new type of curricular content.

As we shall see, in the same way that the concept of competence moved from vocational education and teacher education into general education, so have ideas arising out of assessment of vocational education spread into general education. Specifically, the notion of standards, while, again, not a new concept in educational assessment, has gained more prominence in general education. It is discussed in the section “**Competence and standards**”. The section “**Standards as an instrument in accountability and monitoring performance**” focuses on how standards have been taken further: they are used not only to assess individual learning but also as an instrument in monitoring schools and educational systems.

Before I go on to discuss standards and monitoring, a more general point: It may be problematic to transfer concepts from one educational sphere to another. Young (2008) cautions against the simple import of principles which may be useful in vocational education into general education. To use learning outcomes as an indicator of a person’s competence, in his view, is more sensible in a vocational context where “in many cases it is important to know that someone is capable of competently carrying out specific tasks” (p. 123). But in general education, Young argues, goals cannot be expressed as “precisely specified outcomes”. Similarly, Pring (2004) discusses potential dangers arising from a reliance on skills only as the aim of education: these dangers partly stem from the fact that it is quite hard to define and measure skills appropriately, but also that many activities need knowledge and understanding as well as skills.

Competence and standards

Taken seriously, the principle to teach competences rather than, say, knowledge, does not just replace one form of educational content with another. Instead, competences are conceptualized in relation to some clearly defined task to be performed. Performance can then be assessed, though in the strictest sense, it cannot be graded since the underlying competence cannot be partially present—the individual either is or is not (yet) competent in relation to the task in hand. However, such a strict conceptualization may not be appropriate for most types of competence, and in the general education context, it is more common to refer to competence as a continuous variable which can be measured on a scale, as Messick (1984) has done (see section “**The changing role of competences in the curriculum**”). Nowadays, a combination of the two approaches may frequently be found: a widely applicable conceptualization is a threshold model, where grading on a scale is possible and appropriate, and where a minimum level on the scale has to be reached for someone to be considered competent.

In assessing competence, assessment criteria are derived directly from the task. Thus, competence-based assessment is a form of criterion-referenced assessment. That is, the assessment does not take the form of a comparison with the performance of other learners, as would be the case with norm-referenced assessment (Wolf, 2000). Instead, candidates are judged in relation to a standard, i.e., the threshold described above. Standards, then, are a key element in the assessment of the acquisition of competences. In certain areas, minimum standards are used to ensure that individuals act sufficiently competently in a given situation. An example are driving tests where safety considerations dictate that any driver should be competent enough so as not to endanger themselves

and other road users, though even in such a context, there may be a political component to standard-setting if there is a reason why not too few or not too many candidates should pass a competence test, thus introducing an element of norm-referencing. In the example of driving tests, it may not be politically viable if too many candidates fail their tests, making it hard for them to, for example, commute to work in areas with little public transport. Nevertheless, minimum standards can be very useful in certain contexts, though [Wolf \(2000\)](#) notes that there are dangers associated with them as well. One of them is that a minimum standard as defined in the curriculum may too easily be taken for the maximum standard when the curriculum is implemented. Thus, if a minimum standard is defined for a school subject, teachers may focus on ensuring that as many pupils as possible reach this minimum standard, something which in itself is desirable, but which may come at the expense of the learning of pupils who are able in principle to learn more than required by the minimum standard but who may not receive adequate support to do so.

There is clearly a close relationship between competences and standards. Given the provenance of the concept of competences in the vocational education context, this is unsurprising: assessment of competences in this context involved setting standards against which performance could be judged in order to determine whether the candidate could be deemed competent. Equally clearly, they are not the same thing, though this sometimes can appear to be the case when curriculum documents use the two terms as if they are interchangeable ([Glaesser, 2019](#)). In addition, the term standard is not always employed to mean a threshold separating competent from not-competent individuals. Instead, competence is measured on a continuous scale and several thresholds may be imposed after the fact to give several levels of competence. Such thresholds are not standards in the original sense of the term, since they are derived from the data rather than from theoretical conceptualizations of the underlying construct of competence.

Standards as an instrument in accountability and monitoring performance

The problem of setting minimum standards described in the previous section is exacerbated when not just the candidate's, but the teacher's or the school's performance is monitored using the minimum standard as a benchmark. For instance, the proportion of pupils achieving A* to C (i.e., top grades) in at least five subjects at GCSE (General Certificate of Secondary Education, a qualification taken by 16-year-olds in England and Wales) is a common measure of a school's performance, with the result that many teachers focus on pupils whose level of performance is around the C/D boundary, aiming to have as many as possible achieve at least a C in the actual exam. The result can be that pupils securely above (and well below) this borderline receive much less attention. Teaching to the test is a related and well-known problem: if the achievement of some standard is used to monitor teachers' or schools' performance, then the achievement of that standard becomes the overriding goal rather than the goal of imparting the associated knowledge or competence ([Pring, 2013](#)). The two goals may overlap, but not necessarily fully. This has become known as Goodhart's law: "As Charles Goodhart pointed out some years back, any measure used for control purposes becomes unreliable; so if qualifications are used for monitoring and funding, it is overwhelmingly likely that they will, in the process, become less reliable as sources of information on their holders' skills" ([Wolf, 2004](#), p. 329) (see also [Goldstein, 2020](#)). An example of such a situation is provided by [Schoenfeld \(2014\)](#) who describes how, as a result of the American No Child Left Behind Act (NCLB), high stakes tests were introduced on the basis of which students', teachers' and principals' fates might be decided, so that teachers focused on teaching whatever was demanded by the test. Referring to Hugh Burkhardt's phrase "What You Test Is What You Get" (see also Burkhardt's own chapter in the same volume, [Burkhardt, 2014](#)) Schoenfeld concludes that, in the case of high stakes tests, "testing drives the curriculum" (p. 60) rather than the other way round.

Thus, while it is desirable that the content of the curriculum determines what should be assessed in tests, the effect—arguably less desirably—also runs in the other direction: the way in which tests are designed and used can have a great impact on the curriculum as enacted, i.e., on the teaching itself (e.g., [Wolf, 2000](#), p. 86). For the reasons already discussed, this is problematic, since it is impossible to assess everything which may be regarded as a desirable outcome. Instead, tests are designed to be cost-effective and efficient with regard to time, covering content which is easy to administer and to score. But then teaching to the test may prevent pupils from acquiring other kinds of knowledge and competences which are also aims of the curriculum.

I will illustrate this point by drawing on another example, the case of foreign language learning: there is a wide range of goals associated with foreign language teaching and learning in schools. An obvious one is the acquisition of communicative competence, that is, the ability to communicate with other speakers of the foreign language. It is possible to define and operationalize this, something the Council of Europe has done in its Common European framework of reference for languages (CEFR, [Council of Europe/Conseil de l'Europe, 2001](#)) (for a discussion of the CEFR and its use of the concept of competence see [Glaesser, 2019](#)). The standards arising out of this framework can then in principle be transformed into suitable instruments of assessment. Incidentally, this is a good example of a continuous competence scale with (potentially) meaningful thresholds: there is a point below which there is no communicative competence—the person may know a few words of the language but is still unable to communicate—but above this point there are degrees of competence. For example, an individual may be able to communicate effectively but only in certain situations, or they can express only a limited range of views etc. In the CEFR, there are indeed several such thresholds, resulting in the levels A1 to C2, with clearly described competences an individual can be expected to have given performance at a particular level.

However, foreign language teaching and learning has far wider goals than the acquisition of communicative competence. Other aims are to gain an insight into a different culture including its literature and, possibly, an enhanced understanding of the human condition (e.g., [Paesani, 2018](#); [Paran, 2008](#)). Such things are harder to measure through standardized tests (e.g., [Paran, 2010](#)), and, accordingly, they cannot form part of the accountability framework. This potentially results in a neglect of such topics in the foreign language curriculum.

Accountability and monitoring may partly be a result of the move toward mass education discussed above: if an increased number of young people spend an increased amount of time in the education system, the cost of funding their education rises accordingly. Both politicians and the general public would like to be assured that the money is spent profitably. To this end, they aim to measure schools' performance in order to ensure that monetary input results in output in the form of competences, but, as discussed in this section, measurement of educational output is a complex undertaking with plenty of pitfalls. Such a monitoring approach to ensuring the quality of the education provided can be described as an output-oriented approach, where educational progress and success are measured against standards which are used as instruments of quality control. An input-oriented approach, by contrast, is one where the government aims to achieve high quality of the education system by focusing on teacher training, and by issuing curriculum guidelines and teaching materials. The actual results of teaching, that is, exam performance, dropout rates, and the like, are not the focus of attention for governments using an input-oriented approach. The two are not mutually exclusive, instruments from both forms of quality control are used by most governments, though the focus in recent years has shifted, as discussed in this section, towards more output control (e.g., Ertl, 2006; Glaesser, 2019).

Competences internationally: the PISA study

The Organisation for Economic Co-operation and Development (OECD) regularly conducts surveys of young people's proficiency in the subjects of reading, mathematics and science towards the end of their period of compulsory schooling. The OECD refers to these subjects as "domains of competence": "the survey measures students' proficiency in three foundational domains of competence—reading, mathematics and science" (OECD, 2019) and, indeed, the title of that publication, "What students know and can do" is identical to Messick's definition of both competence and educational achievement quoted in the section "**The changing role of competences in the curriculum**", with educational achievement being, according to Messick, the "consequence of instruction" and competence referring to "what an individual knows and can do ... however that knowledge and skill is acquired". Thus, in this definition, competences are not linked to any particular curriculum, and indeed, PISA does not set out to evaluate performance with respect to curricular content (Hopmann, 2008). Instead, it aims to assess how well the young people perform in relation to competences which they are expected to acquire in order to meet future demands in the workplace (OECD, 2019, p. 63). PISA is thus clearly concerned with competences, but since its focus is not on curricular content, a discussion of PISA may seem out of place in this chapter. However, it has an outsized influence on national curricula and policy-making, as I discuss in this section, and it does set out to compare the performance of different countries' education systems. It is therefore highly relevant to our discussion.

One of the ways in which PISA affects national debates on education and policy-making is how it assesses education systems' performance: its assessment is based solely on external indicators, and it relies on the concept of accountability as discussed in the previous section. Thus, it constitutes a purely output-oriented approach as a way of monitoring and improving performance; its claims concerning its ability to improve education systems are made on that basis alone, not on input-oriented measures. Wahlström (2018) concludes that, in the international context, PISA is a "key player" (p. 720) in the move towards monitoring of educational systems via an international comparison of students' performance. However, output-oriented policies preceded PISA, so it cannot be considered to have invented them, though it has contributed to the spread of the concept and its increased use (Lingard and Lewis, 2016). In that sense, it is both a symptom and a cause of such measures. With regard to its usefulness for reforming and improving educational systems, it is worth noting that PISA's sole focus on performance measures means that it is actually not particularly well suited to examining educational processes in any detail. It does not provide any evidence of what happens in the classroom and therefore no conclusions can be drawn regarding the success of particular teaching methods, and, while it is possible to break down PISA results by countries, schools systems, social background of individual students, gender etc., any differences in the results between the different groups and/or countries do not, in themselves, explain why such difference are found. But this lack of causal explanations does not prevent commentators from assuming that PISA results can tell us which features of particular education systems are responsible for good or bad performance, or—an example of a much debated topic—which education systems are better at ensuring social equality (however defined) (Wahlström, 2018). In fact, PISA results have been used to support opposing arguments, for example arguments both in favor of and against comprehensive schooling (Hopmann, 2008). Similarly, policy makers occasionally use PISA results to support a reform they were planning to carry out anyway, and, given its high public profile, PISA serves as a convenient reason, whether or not its results actually support the proposed reform (Lingard and Lewis, 2016; Tillmann et al., 2008).

In the preceding paragraph, I discussed PISA's influence on national educational systems through the accompanying drive towards monitoring and accountability (see also section "**Standards as an instrument in accountability and monitoring performance**"). Another way in which it exerts some influence is on curricular content directly. PISA describes itself as follows: "PISA is a triennial survey of 15-year-old students around the world that assesses the extent to which they have acquired key knowledge and skills essential for full participation in social and economic life." (OECD, 2019, p. 26) National curricula have taken on board this notion that schools are supposed to prepare young people for their future "participation in social and economic life" and therefore it is important to acquire competences needed later on through schooling. Germany, for example, introduced changes to its curriculum following the so-called PISA shock, i.e., the fact that the country's performance in the first round of PISA in 2000 was not as good as had been hoped and expected. These changes, which were made alongside changes to oversight and the focus of educational research, placed greater emphasis on competences than previously (Ertl, 2006). This is not a result of any teaching to the test,

which would be hard to do in the case of PISA since there is no published practice material. In addition, neither teachers nor individual pupils are held directly to account on the basis of their PISA results, so, for them, there is no direct incentive to achieve highly on PISA. Instead, the focus on competences arises directly out of changes to the curriculum (Glaesser, 2019; Küster, 2016; Prenzel et al., 2015; Stanat, 2018). Similar changes have taken place in Austria (Hopmann, 2008) and in Australia (via the NAPLAN scheme, see Lingard and Lewis, 2016, p. 396), for example. Without linking this directly to PISA (though see Hopmann, 2008, p. 435), Wahlström and Sundberg (2018b) describe how the curricula in Nordic countries have also been changing in recent years to place greater emphasis on the competences associated with employability. They do note, however, that in recent years, there has been “curriculum convergence” globally (p. 4) and that “another area of convergence, and perhaps a stronger one today, concerns the use and development of accountability in output measures and standardized assessments in international knowledge tests like the Programme for International Student Assessment (PISA).”

In summary, PISA’s focus on the type of competences expected to improve participation in social and economic life has influenced national curricula both through changes to curricular content directly and through the greater use of accountability and output-based policies in many national educational systems.

Conclusion

As noted above, Messick (1984) implied that competences may be acquired anywhere, inside as well as outside the classroom. However, the concept has become central to school curricula. Writing as early as 1991, Norris (1991) commented: “Everybody is talking about competence. [...] Words like ‘competence’ and ‘standards’ are good words, modern words; everybody is for standards and everyone is against incompetence.” (p. 331) The discussion in this chapter has shown that this focus on competences has resulted in various developments, some of which may be more welcome than others.

The decision on whether to plan school curricula which include competences to be taught cannot be based on empirical evidence alone: “But, before considering the nature of evidence, it is worth saying that public policy decisions are, and should be, influenced by considerations beyond the research evidence, such as priorities, feasibility, acceptability, ethics and so on. All of these will involve judgements, typically subjective ones” (Goldstein, 2020, p. 38). In the case of planning curricula, an example of choices which have to be made on the basis of such considerations is whether content which may be regarded as “knowledge for its own sake” is so valuable that it should be taught, regardless of whether or not it can be expected to have any immediate benefits apart from, hopefully, enjoyment of the content. More generally, decisions have to be made concerning what the aims of any curriculum should be, whether economic aims such as employability trump everything else or whether different kinds of aims are considered important, too. Budgetary constraints may also play a part in designing curricula. In addition, certain questions are very hard if not impossible to answer empirically. Assuming that employability is a key goal, an empirical test of whether a curriculum is suitable for achieving this goal would have to take account of what demands will be made by the labor market around twenty years later. There are ways of modeling this, but there will still be great uncertainty. It is also quite difficult to determine, empirically, what suitable teaching methods and content look like with regard to achieving a particular goal. Even if it was clear what effective teaching looks like—something that is actually very hard to determine, though some helpful frameworks exist (e.g., Coe et al., 2014)—it would not necessarily be straightforward to implement it. A curriculum not only has to be designed but also enacted (e.g., Burkhardt, 2014). Thus, a focus on both content and implementation is necessary. The wider political context should also be considered: what reasons might politicians, policy makers and various interest groups have for wanting to include competences in the curriculum, or indeed for not wanting to include them? What unintended consequences might there be, for example with regard to assessment and standards (see section “Standards as an instrument in accountability and monitoring performance”)? Some of the questions raised in this paragraph, rather than being amenable to empirical analysis, call for theoretical or philosophical analysis.

There are some reasons to argue that caution may be required in incorporating competences into curricula. Some potential pitfalls have been discussed in this chapter. On differentiation, I would like to stress that there is likely to be no panacea to the problem of unequal opportunities and outcomes in the education system and the labor market. Talent and privilege greatly affect career prospects, and it seems unlikely that competence-based curricula can change this. With regard to competence-based assessment and standard-based monitoring of individual pupils, schools and education systems, unintended consequences have already become apparent, and they have been well documented by numerous authors (e.g., Ball, 2008; Hopmann, 2008; Lingard and Lewis, 2016; Münch, 2012; Pring, 2013; Ravitch, 2010).

A final comment on the link between a competence-based curriculum and the employability it is expected to produce: given that the link between educational and economic growth is much less clear-cut than we are sometimes led to believe (Wolf, 2002, 2004) (see also Bukodi and Goldthorpe, 2019), more caution would be in order concerning the expectations placed on the education system with regard to its contribution to employability. Obviously, schools play an enormously important part in preparing young people for the work place, but it is less clear exactly how they do this, and what particular types of content are most likely to achieve this aim (not least because employability can mean different things—see the point regarding “differentiated employability” at the end of section “Competences in a differentiated curriculum and in mass education” above). Therefore, more theoretical analyses and more empirical research are in order if we wish to obtain a clearer picture of what the curriculum can contribute to the various aims that have been entrusted to it, while at the same time more humility might be required in the claims made with regard to what it can achieve.

Acknowledgment

My thanks go to Barry Cooper for his very helpful comments on this chapter.

References

- Ball, S.J., 2008. *The Education Debate*. Policy Press, Bristol.
- Bukodi, E., Goldthorpe, J., 2019. *Social Mobility and Education in Britain. Research, Politics and Policy*. Cambridge University Press, Cambridge.
- Burkhardt, H., 2014. Curriculum design and systemic change. In: Li, Y., Lappan, G. (Eds.), *Mathematics Curriculum in School Education*. Springer, Dordrecht.
- Carr, D., 1993. Questions of competence. *Br. J. Educ. Stud.* 41, 253–271.
- Coe, R., Aloisi, C., Higgins, S., Major, L.E., 2014. What Makes Great Teaching? Review of the Underpinning Research. The Sutton Trust.
- Cooper, B., 1985. Secondary school mathematics since 1950: reconstructing differentiation. In: Goodson, I. (Ed.), *Social Histories of Secondary Curriculum: Subjects for Study*. Falmer, London.
- Council of Europe/Conseil de l'Europe, 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press, Cambridge.
- Dore, R., 1976. *The Diploma Disease. Education, Qualification and Development*. Allen & Unwin, London.
- Ertl, H., 2006. Educational standards and the changing discourse on education: the reception and consequences of the PISA study in Germany. *Oxf. Rev. Educ.* 32, 619–634.
- Glaesser, J., 2019. Competence in educational theory and practice: a critical discussion. *Oxf. Rev. Educ.* 45, 70–85.
- Goldstein, H., 2020. The role of evidence in public policy making. *Significance* 17, 38–40.
- Hirst, P.H., 1969. The logic of the curriculum. *J. Curric. Stud.* 1, 142–158.
- Hopmann, S.T., 2008. No child, no school, no state left behind: schooling in the age of accountability. *J. Curric. Stud.* 40, 417–456.
- Küster, L., 2016. Kompetenzorientierung im Kontext des Lernens und Lehrens von Sprachen. In: Bausch, K.-R., Burwitz-Melzer, E., Krumm, H.-J., Mehlhorn, G., Riemer, C. (Eds.), *Handbuch Fremdsprachenunterricht*, sixth ed. Francke, Tübingen.
- Lingard, B., Lewis, S., 2016. Globalization of the Anglo-American approach to top-down, test-based educational accountability. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, London, New York.
- Messick, S., 1984. The psychology of educational measurement. *J. Educ. Meas.* 21, 215–237.
- Münch, R., 2012. Mit dem PISA-Express in die globale Wissensgesellschaft. In: Pfadenhauer, M., Kunz, A.M. (Eds.), *Kompetenzen in der Kompetenzerfassung. Ansätze und Auswirkungen der Vermessung von Bildung*. Beltz Juventa, Weinheim.
- Norris, N., 1991. The trouble with competence. *Camb. J. Educ.* 21, 11.
- OECD, 2019. *PISA 2018 Results. What Students Know and Can Do*. Pisa, Oecd Publishing, Paris.
- Paesani, K., 2018. Researching literacies and textual thinking in collegiate foreign language programs: reflections and recommendations. *Foreign Lang. Ann.* 51, 129–139.
- Paran, A., 2008. The role of literature in instructed foreign language learning and teaching: an evidence-based survey. *Lang. Teach.* 41, 465–496.
- Paran, A., 2010. Between scylla and charybdis: the dilemmas of testing language and literature. In: Sercu, L., Paran, A. (Eds.), *Testing the Untestable in Language Education*. Multilingual Matters, Bristol.
- Prenzel, M., Blum, W., Klieme, E., 2015. The impact of PISA on mathematics teaching and learning in Germany. In: Stacey, K., Turner, R. (Eds.), *Assessing Mathematical Literacy: The PISA Experience*. Springer, Cham.
- Pring, R., 2004. The skills revolution. *Oxf. Rev. Educ.* 30, 105–116.
- Pring, R., 2013. *The Life and Death of Secondary Education for All*. Routledge, London.
- Ravitch, D., 2010. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, New York.
- Rosenmund, M., 2007. The current discourse on curriculum change: a comparative analysis of national reports on education. In: Benavot, A. (Ed.), *School Knowledge in Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education*. Springer Netherlands, Dordrecht.
- Schoenfeld, A.H., 2014. Reflections on curricular change. In: Li, Y., Lappan, G. (Eds.), *Mathematics Curriculum in School Education*. Springer, Dordrecht.
- Stanat, P., 2018. 15 Jahre kompetenzorientierte Bildungsstandards—eine kritische Reflexion. In: Jungkamp, B., John-Ohnesorg, M. (Eds.), *Können ohne Wissen? Bildungsstandards und Kompetenzorientierung in der Praxis*. Friedrich-Ebert-Stiftung, Berlin. <http://library.fes.de/pdf-files/studienfoerderung/14457.pdf>.
- Tillmann, K.-J., Dederig, K., Kneuper, D., Kuhlmann, C., Nessel, I., 2008. PISA als bildungspolitisches Ereignis. Fallstudien in Vier Bundesländern. VS Verlag für Sozialwissenschaften, Wiesbaden.
- Transnational curriculum standards and classroom practices. In: Wahlström, N., Sundberg, D. (Eds.), 2018a. *The New Meaning of Teaching*. Routledge, Abingdon.
- Wahlström, N., Sundberg, D., 2018b. Transnational curriculum standards, curriculum reforms and classroom practices—an introduction. In: Wahlström, N., Sundberg, D. (Eds.), *Transnational Curriculum Standards and Classroom Practices. The New Meaning of Teaching*. Routledge, Abingdon.
- Wahlström, N., 2018. Where is “the political” in curriculum research? *J. Curric. Stud.* 50, 711–723.
- Whitty, G., Willmott, E., 1991. Competence-based teacher education: approaches and issues. *Camb. J. Educ.* 21, 10.
- Wolf, A., 2000. *Competence-Based Assessment*. Open University Press, Buckingham.
- Wolf, A., 2002. Does Education Matter? Myths about Education and Economic Growth. Penguin, London.
- Wolf, A., 2004. Education and economic performance: simplistic theories and their policy consequences. *Oxf. Rev. Econ. Pol.* 20, 315–333.
- Young, M.F.D., 2008. Contrasting approaches to qualifications and their role in educational reform. In: Young, M.F.D. (Ed.), *Bringing Knowledge Back In. From Social Constructivism to Social Realism in the Sociology of Education*. Routledge, London.
- Young, M.F.D., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45, 101–118.

Further reading

- Dale, R., 1999. Specifying globalization effects on national policy: a focus on the mechanisms. *J. Educ. Pol.* 14, 1–17.
- Mcculloch, G., 2016. History of the curriculum. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. SAGE.

Shadow curriculum as a new form of curriculum: its evolution, definition, and implications for curriculum studies

Young Chun Kim^a and Jung-Hoon Jung^b, ^aChinju National University of Education, Jinju-si, South Korea; and ^bPusan National University, Pusan Metropolitan City, South Korea

© 2023 Elsevier Ltd. All rights reserved.

Need for a new definition: shadow curriculum	288
Evolution of the concept	289
Defining shadow curriculum	290
The characteristics of shadow education	290
Shadow curriculum: definition and characteristics	291
Focusing on the academic success of students in schooling	291
Fulfilling the academic needs and abilities of individual students	291
Providing differentiated content based on the amount of tuition	291
Increasing the competitive nature of education	292
Shadow curriculum for future curriculum studies	292
Conclusion	293
References	293

Curriculum is a complicated concept and is defined in different ways by scholars who are working in different countries and regions, in different academic fields, and who have different research purposes (Aoki, 1980/2005; Jung and Pinar, 2015). As a result, the concept is elastic and multifaceted; it has been described as involving a “cacophony of voices” (Pinar et al., 1995). As in other academic fields, discourse has advanced the field of curriculum studies, leading to the inclusion of new concepts such as hidden curriculum, null curriculum, etc. This chapter presents another concept, shadow curriculum, which has emerged from the international phenomenon of private supplementary tutoring, also known as shadow education. The following discussion explores the aspects of shadow education that have led to the need for shadow curriculum. It also reviews the history of cognate concepts that are related to shadow education, presents a broad definition of the phenomenon, and concludes with areas for future research within the field of curriculum studies.

Need for a new definition: shadow curriculum

Why do we need to develop a new curricular concept? The following discussion explores four important and interrelated elements related to the worldwide phenomenon of shadow education and how these have led to the need for a new concept in the field of curriculum studies.

First, many students now learn in shadow education spaces for both remedial and accelerated learning. Historically, shadow education has been prominent in some East Asian countries such as South Korea, Japan, Hong Kong, Taiwan, and Singapore (Bray, 2007[1999], 2017; Dawson, 2010; Entrich, 2017; Kim, 2016; Mori and Baker, 2010). It has now expanded to become a worldwide phenomenon. Since the beginning of the 21st century, it has emerged in many countries including Canada (Aurini and Davies, 2004), China (Zhang and Bray, 2017), Egypt (Hartmann, 2013), England (Jerrim, 2017), Australia (Watson, 2008), and many others in South Asia, Southern and Western Europe, North Africa, and South America (see Bray and Kwo, 2014; Kim and Jung, 2019a,b). Specifically, Jerrim (2017) reported that approximately one-third of students aged 11–16 years in England and Wales engaged in shadow education. Notably, he found that half of the students in London accessed private tutoring, depending on grades, regions, and social-economic status of families. Jerrim also found that 11th graders in England spend on average 9.5 h per week in private tutoring, and that this is significantly less than students in 12 other European countries. Aurini and Davies (2004) found that in the Canadian context, shadow education was school-like in terms of organizational structure, curriculum, and participation rates. They found that as of 2002, 24% of Ontario students were taking tutoring classes, and that the tutoring business had increased by 200–500% in major Canadian cities during the previous 30 years. Sweden, along with other Nordic countries, has experienced rapid growth in the market for private supplementary tutoring (Bouakaz, 2012; Forsberg et al., 2020; Hallsén and Karlsson, 2019).

Second, research has demonstrated that shadow education serves various functions to ensure academic success, including supporting student learning in schools, improving academic performance, and obtaining admission to a prestigious schools or universities. While the findings have yielded some inconsistent results, they generally suggest that shadow education significantly improves academic achievement of specific groups of students (see Dang and Rogers, 2008; Choi and Park, 2016), and that the effects are generally positive (Buchmann et al., 2010; Mischo and Haag, 2002; Davis et al., 2011; Brand and Xie, 2010; Kim and

Jung, 2019a,b; OECD, 2012; Carr and Wang, 2015; Entrich, 2014; Askew et al., 2010). More specifically, research has demonstrated that shadow education can help students improve their scores on university examinations such as the SAT and KSAT (Byun and Park, 2012; Appelrouth et al., 2017). Hu et al. (2015) reported that supplementary math tutoring had significant and positive effects on math scores among students in Hong Kong, Japan, South Korea, and Shanghai. Byun and Park (2012) found that East Asian students in the United States were more likely to take SAT preparation courses provided by shadow education services and that these improved their test scores. Appelrouth et al. (2017) found that individual tutoring was more effective than group tutoring for students in the United States. In the South Korean context, Hyung Jai Choi (2008) reported a positive relationship between private tutoring and college admission. Jung-Ho Yang (2012) also investigated the relationship between shadow education participation and the likelihood of university entrance; an analysis involving multinomial logistic regression revealed a strong relationship between private tutoring and admission to a four-year university. Kim (2016) collected in-depth narratives of how Korean students learn in shadow education to increase and/or maintain school grades and eventually obtain college admission. Although some scholars have reported that the effects of shadow education differ depending on the characteristics of students (Hof, 2014; Zhang, 2013), research appears to indicate that students in most countries can benefit from shadow education.

Third, with the expansion of shadow education, students who are simultaneously in and out of schooling are constructing a new culture of learning (Williams, 2013; Jung, 2018; Kim et al., 2018). This transboundary culture of learning is emerging as students crisscross the spatial, temporal, institutional, and curricular boundaries between public schooling and shadow education (Kim and Jung, 2019). Within this new culture, students are active agents who decide what, when, where they learn, and from whom. This has been weakening the status and authority of school curriculum; one scholar argued that the new culture of learning is a form of attack on institutionalized schooling and a commitment to self-determined learning that “reanimates the countercultural deschooling agenda” (Williams, 2013, p. 42). Shadow education provides more options for learning, allowing students and parents to be important decision-making agents in terms of curriculum. Students have traditionally been expected to learn the materials they are presented with in school, and this is still the case in many countries, but the emergence and expansion of shadow education is dramatically changing the ecology of education, eroding the authority of school curriculum and making students key decision-makers in their own learning.

Fourth, although private tutoring currently coexists with public education, it historically pre-dates it. Worldwide, people learned from private tutors long before public education systems emerged. The sophists in Ancient Greece and Rome could be described as freelance or personal teachers of philosophy who were “very strictly private tutors with no official standing whatsoever” (Winter, 1997, p. 25). Gordon (1988) traced the history of tutoring from Ancient Greece to the United States. Tutors have been called sophists, paidagogós, tutela, thetores, domi, fori, pedagogue, pedisequis, comes, custos, rector, governor, and governess. In Medieval Europe they were sometimes called escalate or theological. Gordon argued that in the 20th Century, American public education became almost exclusively a school experience, with tutorial philosophies and methods forgotten or quietly absorbed by tax-supported public schools, but that the American progressive education movement continued to embrace the concepts of individual differences and child-centered education first proposed by philosopher-tutors like Erasmus, Elyot, Vives, Thomas More, Locke, Fenelon, Rousseau, and others. Gordon also highlighted the roles played by tutors in the ‘visiting teacher movement’ during the ‘domestic education’ period in the 1920s in the United States and in many educational capacities throughout modern Great Britain; tens of thousands of teachers in England supplemented their income through private tutoring, but in 1985, when Gordon was conducting his research, this activity remained highly secretive for tutors, children, and parents. In English, shadow education is usually called private supplementary tutoring; it is referred to as private tuition and coaching in Bangladesh, India, and Pakistan; juku in Japan; hakwons and sagyoyuk in South Korea; and buxiban in Taiwan. In Ireland, shadow education might be called fosterage; in Italy, courtesy education; and in Germany, wohnstübenerziehung, erziehungskunstler, or Nachhilfe (Bray, 2007(1999); Kim and Jung, 2019a,b).

Overall, the worldwide phenomenon of shadow education cannot be ignored by curriculum studies scholars. The next section explores the new concept of shadow curriculum, which curriculum studies scholars can apply in future research.

Evolution of the concept

This section focuses on academic discourse related to K-12 educational contexts, where shadow education is conceived as private tutoring that supplements public education. For example, the term ‘shadow education system’ has been used to describe worker education contexts, but this is beyond the scope of this discussion (Wirtz, 1979; Craig and Evers, 1982).

In the 1970s and 1980s, scholars, educators, and policymakers were urged to pay attention to the phenomenon of shadow education. In his 1976 book *The Diploma Disease*, Ronald Dore critiqued the failing public schooling system for being geared toward “mere qualification-earning” for the modernized job markets in England, Japan, Sri Lanka, and Kenya. He noted that supplementary tutoring was an important element of this, helping students pass examinations to earn qualifications. In 1976, the Japanese government conducted a survey to investigate shadow education (Mori and Baker, 2010). This inspired more work including Rohlen’s books *The Juku Phenomenon: An Exploratory Essay* (1980) and *Japan’s High Schools* (1983), in which Rohlen referred to shadow education as the pursuit of educational advantage among middle-class families. In the South Korean context, Jung-Han Kim (1985) wrote on the intense home-visit form of private tutoring (gwaoe) used to prepare for university entrance examinations.

In the 1990s, scholars made efforts to develop the terms “shadow education” and “private supplementary tutoring” as academic concepts. de Silva et al. (1991) used the term “private supplementary tutoring” in their book *Extra-school Instruction, Social Equity and*

Educational Quality (Sri Lanka). Marimuthu et al. (1991) were the first to use the shadow metaphor in their Malaysian study. The next year, Stevenson and Baker (1992) defined shadow education as “a set of educational activities that occur outside formal schooling and are designed to enhance the student’s formal school career” (p. 1639). Bray popularized the term “shadow education” in his 1999 book *The Shadow Education System: Private Tutoring and Its Implications for Planners*.

Since 2000, more scholars have conceived of shadow education as an international phenomenon. Bray 2007(1999) book was the first to incorporate data from many countries and provided a comprehensive picture of shadow education in the international context. It inspired cross-national comparisons in terms of the intensity of shadow education, educational inequality, and government policies. Baker et al. (2001) extended the discourse in their book *Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement*; they argued that shadow education is prevalent worldwide and described the strategies and characteristics used in different countries and regions. Since then, many researchers have framed shadow education as a common international phenomenon (Ireson, 2004; Silova et al., 2006; Bray and Lykins, 2012; Manzon and Areepattamannil, 2014; Entrich, 2017).

Until recently, the concept of shadow education has mainly been defined and used in academic fields including comparative education (Mori and Baker, 2010; Ventura and Jang, 2010), education and policy (Bray and Kwo, 2014; de Castro and de Guzman, 2014; Ireson, 2004; Park et al., 2016), sociology of education (Byun, 2010; Yamamoto and Brinton, 2010), education and economics (Entrich, 2014; Zhang, 2013), and lifelong education (Ozaki, 2015). Few scholars have tried to define shadow education from the perspective of curriculum studies. In *Shadow Education: Evolution Flaws and Further Development of the Term*, Malik (2017) pointed out the ambiguity of the concept of shadow education and identified subcategories including shadow teaching, shadow curriculum, and prerecorded academic aids. These subcategories are helpful for revealing various possible, and understudied, research topics, but we argue that Malik’s use of the term “curriculum” is limiting. Malik defined curriculum as “textual” and conceived it as a static entity of materials. In reality, a curriculum is neither static nor singular (Pinar et al., 1995). The next section presents a new concept for inclusion in the field of curriculum studies: shadow curriculum.

Defining shadow curriculum

The following discussion draws from existing terminologies, definitions, and characteristics of shadow education, beginning with the commonly accepted characterization of shadow education (Bray, 2007(1999)). Next, we draw from our extensive literature review and our own fieldwork to present four main characteristics of shadow education. Finally, based on these findings, we present a definition of shadow curriculum.

The characteristics of shadow education

Shadow education can be generally defined as “private supplementary tutoring in academic subjects beyond the hours of mainstream formal schooling” (Bray, 2006, p. 515; also see e.g., Marimuthu et al., 1991; Stevenson and Baker, 1992; Silova et al., 2006). It has five general forms: private tutoring institute, home-visit private tutoring, Internet-based tutoring, subscribed learning program, and academic after-school program (Kim and Jung, 2019a,b).

According to Bray (2007[1999]), shadow education involves three main characteristics: supplementation, privateness, and academic subjects. “Supplementation” refers to both remediation and acceleration for student learning in mainstream schooling. In other words, shadow education can help low-achieving students catch up and help high-achieving students advance faster and more (Carr and Wang, 2015; Cho, 2015; Entrich, 2014; Kim and Jung, 2019a,b).

“Privateness” refers to funding source: parents/families pay tuition to shadow education providers such as educational personnel and companies (Stevenson and Baker, 1992; Bray, 2007[1999]; Malik, 2017; Kim, 2016). Importantly, this characteristic excludes any form of schooling, both public and private—including home-schooling funded and regulated by governments. It also excludes individual or group self-directed study at home, as this does not involve an instructor. It also excludes various available learning sources such as online videos uploaded by the Kan Academy or individual creators as they do not specifically target registered students—however, these may be components of shadow education when they are intentionally assigned to specific students by shadow education instructors. Many scholars have focused on this characteristic of shadow education because it is generally seen as exacerbating educational inequality (Melot, 2007; Klemm and Klemm, 2010; Majumdar, 2014).

“Academic subjects” refers to the explicit objective and contents of shadow education. Shadow education is in general intended to ensure academic success by helping students improve their scores on school examinations and obtaining admission to a university of choice (Lee, 2007; Yamamoto and Brinton, 2010). For this reason, the most popular subjects in shadow education are language arts, mathematics, English (and other foreign languages), and science (see Bray and Lykins, 2012; Kim, 2016; Kim and Jung, 2019a,b).

These three characteristics are important elements of shadow education, but Bray’s analysis was incomplete. We argue that shadow education involves a fourth characteristic: individualness. “Individualness” refers to students seeking learning opportunities that best meet their individual needs, educational purposes, and level of learning (Bray, 2007[1999]; Ozaki, 2015; Kim, 2016; Entrich, 2017; Malik, 2017; Kim and Jung, 2019a,b). In most countries, mainstream education is based on the values of “education for all” and “education for equality” (Mori and Baker, 2010; Bray, 2017). As a result, its instruction, programs, and evaluations are usually age-based to make education equal and fair. In contrast, shadow education is delivered based on each

student's intellectual abilities, needs, interests, and learning speed (Ozaki, 2015; Kim, 2016; Entrich, 2017; Kim and Jung, 2019a,b)—and shadow education businesses only profit when they meet the needs of individual customers (Bray, 2007 [1999]; Kim 2016; Malik, 2017).

Shadow curriculum: definition and characteristics

Building on the previous discussion, we can broadly define shadow curriculum as: supplementary curriculum out of schooling provided by educational business industries for academic success of individual students in formal education.

Shadow curriculum has four distinctive characteristics that differentiate it, at least in part, from curriculum used in mainstream education. The following sections explore these characteristics, based on our extensive literature review and our own field research in South Korea.

Focusing on the academic success of students in schooling

The first characteristic of shadow curriculum is that its primary focus is to help students achieve academic success. This is the main reason that students access shadow education (Mori and Baker, 2010; Stevenson and Baker, 1992; Silova et al., 2006; Bray, 2007(1999); Byun and Park, 2012; Byun, 2010; Kim, 2016; Zhang and Bray, 2017; Entrich, 2018). In most countries, school curricula are designed to promote holistic development in students by balancing intellectual, affective, and psychomotor development. In contrast, shadow curriculum is oriented toward academic success, and especially improving scores on university admissions examinations (Kim, 2016; Ozaki, 2015; Stevenson and Baker, 1992). In this respect, it is developed to be goal-specific and task-oriented for a specific test or improve a grade in a key course (Aurini and Davies, 2004; Harnisch, 1994; Kim, 2016; Kim and Kim, 2012; 2015; Stevenson and Baker, 1992; Yamamoto and Brinton, 2010). Many studies have focused on this element of shadow education in Japan (Stevenson and Baker, 1992), where intense engagement in *juku* has been linked with university admission and overall future success in life (Ozaki, 2015; Harnisch, 1994). Similar findings have been reported in South Korea, where shadow curriculum is highly oriented toward guiding students to prepare for school examinations and the College Scholastic Ability Test (Kim, 2016; Kim and Jung, 2019a,b). Aurini and Davies (2004) found that shadow education in Canada also tends to be goal-specific and task-oriented, usually to pass an upcoming test or improve a grade in a key course. Another example is Finland, which is known for its excellent education system. Kosunen (2018) recently explored the emerging processes of hidden privatization in education and found that tutoring in preparatory courses for the competitive university entrance examinations in Finland could be interpreted as a form of private education as it “functions alongside the public education system, the aim being to enhance the position of clients by facilitating access to higher education ... while collecting tuition fees” (p. 67).

Fulfilling the academic needs and abilities of individual students

Shadow curriculum provides individualized, personalized, and differentiated learning specific to each student (Medway, 1995). This kind of education was common before the emergence of public schooling, based on tutorial philosophies and methods (Gordon, 1988). Now, most mainstream schools teach students the same content as a group, at the same time, using the same methods. Some schools have made efforts at individualized learning, but students tend to be more motivated by private tutoring (Silova, 2010). This is particularly true in South Korea, where many students rely on *hakwon* education to achieve their goals because public schooling does not meet their academic needs (Kim, 2016)—and shadow curriculum actively and aggressively addresses consumer need (Kim, 2016; Kim and Kim, 2012; 2015; Mawer, 2015). In some countries such as Japan, Australia, and South Korea, the public education system provides vouchers to students whose needs are not met in schooling so that they get extra academic support from shadow curriculum providers (Dawson, 2010; Kim and Jung, 2019a,b). Generally, students and their parents seek shadow education because public education systems are less able to provide individualized learning (Bray, 2007(1999); Bray and Kwo, 2014; Kim, 2016; Ozaki, 2015; Park et al., 2016).

Shadow curriculum providers respond to this need (Bray, 2007(1999); also see Bray and Kwo, 2014; Kim, 2016; Kim and Jung, 2019a,b; Entrich, 2017). They may use techniques such as one-to-one tutoring (Bray, 2009). In South Korea, many *hakwons* use ability grouping, differentiated instruction, tailored curricular programs, and counseling informed by portfolios, learning diaries, university exam preparation schedules, and accumulative evaluations (Kim, 2016). In Japan, *hingaku*, *hoshu juku*, and *juku* that specialize in *kobetsu shido*, a type of *juku* that is focused on individual instruction, provide highly individualized instruction for both academic study and more general guidance (Mawer, 2015). Internet-based private tutoring also provides highly personalized content with many choices in terms of subject areas, levels of courses, materials to be used, and teaching methods (Kim and Jung, 2019a,b; Kim, 2016; Park et al., 2016; Ventura and Jang, 2010). In contrast to the curriculum provided in public schools, where students are assigned teachers and learning materials, students engaging with shadow curriculum can manage their own learning by selecting specific instructors, course content, learning speed, as well as where and when to study.

Providing differentiated content based on the amount of tuition

The third characteristic of shadow curriculum is tuition, which is generally not a part of mainstream education. Families pay to access shadow curriculum, and in general, higher tuition will ensure a smaller teacher-student ratio, more frequent sessions, more

qualified instructors, and better curricular materials (Kim and Jung, 2019a,b). Wealthy families are more able to live near public schools with better quality education, and may even move to certain areas (e.g., Gangnam in Seoul) to better their children's education (Kim, 2016; Park et al., 2015). However, parental investment in shadow curriculum tends to function more powerfully—and much more quickly. One-to-one tutoring, which is generally considered the most personalized, is usually the most expensive (Bray and Kwo, 2014; Kim, 2016). Internet-based tutoring is the most reasonably priced, but it becomes more expensive if a student needs more courses or individual services.

Shadow curriculum can be considered a form of positional good because it is highly subject to a family's socio-economic status (Halliday, 2016). Students from wealthy families are more likely to access—and benefit from—shadow curriculum than their less resource-rich counterparts (Bray and Kwok, 2003; Byun and Kim, 2008; Dang, 2007; Kim, 2016; Rowley et al., 2020). Empirical research has demonstrated that shadow education can reproduce and/or strengthen educational inequality (Park et al., 2015; Sun and Braeye, 2012; Lee et al., 2010), and many scholars have criticized it for this reason (Dawson, 2010; Halliday, 2016). For example, researchers have documented the positive relation between familial investment and entrance rates to prestigious schools in Korea (Kim, 2015; Kim and Park, 2010).

However, Rowley et al. (2020) analyzed the OECD Program for International Student Assessment (PISA) results of students from wealthy countries and concluded that scores are not necessarily affected by socio-economic status. Additionally, some researchers have found that shadow curriculum can actually equalize educational opportunity by providing extra support to disadvantaged children (Aurini et al., 2013; Bray, 2009; Entrich, 2017). Overall, the issue of whether shadow curriculum contributes to educational inequality is far more complex than it first seems.

Increasing the competitive nature of education

The fourth main characteristic of shadow curriculum is that it perpetuates the idea of education as competition (Jung, 2014, 2016). Students work to obtain the best grades and test results, and shadow education is contributing to intense competition (Halliday, 2016; Kim, 2003; Kim and Kim, 2012, 2015; Lee, 2007; Mawer, 2015). High-achieving students are much more likely to engage in shadow education because they want to maintain their superior position in terms of academic achievement (Toyama-Bialke, 1997; Bray, 2007[1999]; Kim, 2016; Tseng, 1998). Under this kind of intense competition, students tend to consider other students as rivals to defeat, and to be preoccupied with test scores and school grades rather than focusing on the intrinsic value of learning (Kim, 2016; Park et al., 2015; Zhang, 2013). Halliday (2016) argued that this can lead to a kind of “educational arms race” at the expense of educational development, noting that students become vulnerable when competition comes before the intrinsic value of learning. Taubman (2010) came to similar conclusions in his discussion about the discourse of standards and accountability in education, arguing that teaching and learning for numbers extracted from students, teachers, and schools makes students vulnerable to centers of control while providing the illusion of objective accountability and meritocracy. Overall, although competition is an increasingly important element of education in many countries, it should not be the primary or only component of student learning culture. As far as the role of shadow curriculum is concerned, it would be hard for shadow curriculum free from the accusation that many scholars criticized.

Shadow curriculum for future curriculum studies

This section presents areas for future research in the field of curriculum studies. Certain elements of shadow education remain largely understudied (Zhang and Xie, 2016), especially within the field of curriculum studies (Kim and Jung, 2019a,b). The advantages of shadow curriculum are largely unknown and require rigorous investigation. Most discourse related to shadow education has focused on negative aspects of its practices (Bray, 2007[1999]), especially the increased workload for students (Yamamoto and Brinton, 2010) and the financial burden for families (Park et al., 2016, 2015) that may also perpetuate educational inequalities (Bray and Kwo, 2014; Burch, 2009; Stevenson and Baker, 1992). More recently, some scholars have begun to focus on the advantages of shadow education. For example, Hajar (2018) found that in England, private tutoring not only provides tangible benefits such as academic achievement but also intangible benefits among students by boosting their self-esteem, confidence, and interest in learning. Other scholars have reported that shadow education provides similar benefits to student wellbeing (Jung and Jung, 2019; Ozaki, 2015). Still, the advantages of shadow curriculum remain largely unknown, possibly because of the power dynamic between mainstream education and shadow education. More research is needed to explore issues such as these. Shadow curriculum involves innovative practices. Although it “shadows” mainstream curriculum, it is not tightly controlled by governments, and its reach continues to expand; for example, advances in technology have led to rapid expansion of Internet-based private tutoring.

The power dynamic between schooling and shadow education is an important issue for further research. Shadow education has been framed as subordinate or inferior to public education (Bray, 1999; de Silva et al., 1991; Marimuthu et al., 1991) or as an “exotic” cultural practice in East Asian nations (Mori and Baker, 2010). This kind of framing is related to the long-standing prioritization of public schooling within the field of curriculum studies. This stance was rationalized based on the almost exclusive emphasis on the “backwash effects” of shadow education that have led to educational policies intended to control or suppress shadow education (Bray and Kwo, 2014; Doherty and Dooley, 2018; Kim and Jung, 2019a,b). However, this issue is far more complex than it first appears. Shadow curriculum offers what public schools do not, including remedial instruction (Roesgaard,

2006). In this way, shadow curriculum can actually reduce inequality in educational opportunities (Kim, 2016; Mori, 2012; Entrich, 2014; Kim and Jung, 2019a,b; Tsuneyoshi, 2001; Dawson, 2010). Shadow curriculum can also play an important role in preserving ethnicity, cultures, and languages (Sun and Braeye, 2012; Zhou and Kim, 2006; Kim and Jung, 2019a,b). More research is needed to explore why students are drawn to private tutors more than public school teachers (Paramita, 2014), why and how the phenomenon of “inverted roles” has emerged between schools and shadow education institutes (Yang and Kim, 2010), and why and how shadow education in some countries has become school-like (Aurini and Davies, 2004).

Another issue for future research is how shadow curriculum affects school curriculum, and how school teachers teach. Some research suggests that shadow education can lead to corruption among teachers; a study in Vietnam revealed some teachers provide tutoring classes for students for whom they are already responsible, and many exclusively disclose examination questions to the tutored students (Vu et al., 2011). Similar findings have been reported in Cambodia (EIC et al., 2006) and Central Asia and the Caucasus (Silova, 2010). On the other hand, student participation in shadow curriculum also affects school teachers. Many teachers are facing more challenges as more students access shadow education. Teachers in South Korea report that many students do not pay attention in class because they have already learned the targeted contents in shadow education (Kim and Jung, 2019a,b). In Japan, up to 30% of junior high-school students reportedly sleep through class (Hattori et al., 2010), often due to their attending juku (Ozaki, 2015). Some countries have introduced policies regulating shadow education (Bray & Kwo, 2014). More research is needed to explore how shadow curriculum influences school curriculum and instruction, and this should include different contexts and involve diverse research foci.

Another area for further research is transboundary curriculum. Some scholars have argued that technological innovations in education have led to a countercultural “deschooling” agenda that is against institutionalized schooling and assembly-line learning, and instead promotes self-determined learning through informal networks and community bonds (Williams, 2013). Student learning has become increasingly flexible, and the line between schooling and shadow curriculum is becoming increasingly blurry. Schools and school teachers are being decentered as the sole educational authorities as student learning is happening “in and beyond school walls” (Park et al., 2016), “physical and cyber learning spaces” (Ventura and Jang, 2010; Williams, 2013), and “official and nonofficial curricular programs and materials” (Kim and Jung, 2019a,b). In this respect, curriculum scholars need to explore both mainstream curriculum and shadow curriculum, specifically how students learn in complex and blurry new ways. It is vital that we understand how shadow curriculum fits into the whole picture of student achievement. One way to do this is through analysis of international comparison test results, e.g., PISA and TIMSS scores. One report (OECD, 2012) has already recognized the importance of shadow curriculum for students’ achievement, but it provides limited data. More comprehensive data collection and analysis will help reveal the quality of student learning in different learning contexts and to what extent their learning in shadow curriculum affects their achievement. PISA or TIMSS could aid this kind of research by develop new evaluation designs to assess how, where, what, and with whom students learn.

Conclusion

Educational scholars can no longer ignore shadow education. It has become an important component of the ecology of education worldwide, and should no longer being considered an intruder or “invasive species” (Bray and Kobakhidze, 2015). Instead of being framed as having only negative effects on students, families, and mainstream education, the phenomenon deserves scholarly attention and acceptance. More research on shadow curriculum will greatly broaden our understanding of the phenomenon, and break new theoretical ground in the field of curriculum studies. The power dynamic between public education and shadow education is complex, and should not be limited to the dominant antagonistic stance toward shadow education (Takeda, 2005; Ozaki, 2015; Kim and Jung, 2019). We hope that as an initial academic effort in curriculum studies, our new concept of shadow curriculum opens intellectual space for more nuanced and complicated conversations.

References

- Aoki, T.T., 2005 [1985/1991][J]. Signs of vitality in curriculum scholarship. In: Pinar, William F., Irwin, R.L. (Eds.), *Curriculum in a New Key*. Lawrence Erlbaum, Mahwah, NJ, pp. 229–233.
- Appelrouth, J.I., Zabrocky, K.M., Moore, D., 2017. Preparing students for college admissions tests. *Assess Educ. Princ. Pol. Pract.* 24 (1), 78–95.
- Askew, M., Hodgen, J., Hossain, S., Bretscher, N., 2010. *Values and Variables: Mathematics Education in High-Performing Countries*. London: Nuffield Foundation.
- Aurini, J., Davies, S., 2004. The transformation of private tutoring: education in a franchise form. *Can. J. Sociol.* 29 (3), 419–438.
- Aurini, J., Davies, S., Dierkes, J., 2013. Out of the Shadows: The Global Intensification of Supplementary Education. Emerald, Bingley.
- Baker, D.P., Akiba, M., LeTendre, G.K., Wiseman, A.W., 2001. Worldwide shadow education: outside-school learning, institutional quality of schooling, and cross-national mathematics achievement. *Educ. Eval. Pol. Anal.* 23 (1), 1–17. <https://doi.org/10.3102/01623737023001001>.
- Bouakaz, L., 2012. Att behålla mitt och lära mig något nytt: föräldraengagemang i mångkulturella miljöer- Föräldraengagemang i mångkulturella miljöer. Ett diskussionsunderlag framtaget för Kommission för ett socialt hållbart Malmö [“To Keep My Own and Learn Something new” – Parent Involvement in Multicultural Contexts. A Discussion Piece by the Commission for a Socially Sustainable Malmö]. Malmö kommun, Malmö.
- Brand, J.E., Xie, Y., 2010. Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *Am. Sociol. Rev.* 75 (2), 273–302.
- Bray, M., Kobakhidze, M.N., 2015. Evolving ecosystems in education: the nature and implications of private supplementary tutoring in Hong Kong. *Prospects* 45 (4), 465–481.

- Bray, M., Kwo, O., 2014. Regulating Private Tutoring for Public Good: Policy Options for Supplementary Education in Asia. UNESCO, Bangkok.
- Bray, M., Kwok, P., 2003. Demand for private supplementary tutoring: conceptual considerations, and socio-economic patterns in Hong Kong. *Econ. Educ. Rev.* 22, 611–620.
- Bray, M., Lykins, C., 2012. Shadow Education Private Supplementary Tutoring and its Implications for Policy Makers in Asia. Asian Development Bank, Thailand.
- Bray, M., 2006. Private supplementary tutoring: comparative perspectives on patterns and implications. *Compare* 36 (4), 515–530.
- Bray, M., 2007[1999]. The Shadow Education System: Private Tutoring and its Implications for Planners. UNESCO International Institute for Educational Planning (IIEP), Paris.
- Bray, M., 2009. Confronting the Shadow Education System. UNESCO, Paris.
- Bray, M., 2017. Schooling and its supplements: changing global patterns and implications for comparative education. *Comp. Educ. Rev.* 61 (3), 469–491.
- Buchmann, C., Condon, D.J., Roscigno, V.J., 2010. Shadow education, American style: test preparation, the SAT and college enrollment. *Soc. Forces* 89, 435–462.
- Burch, P., 2009. Hidden Markets: The New Educational Privatization. Routledge, New York, NY.
- Byun, S.Y., Kim, K., 2008. Parental involvement and student achievement in South Korea: focusing on differential effects by family background. *Korean J. Sociol. Educ.* 18 (1), 39–66 in Korean.
- Byun, S.Y., Park, H., 2012. The academic success of East Asian American youth: the role of shadow education. *Sociol. Educ.* 85 (1), 40–60.
- Byun, S., 2010. Does policy matter in shadow education spending? Revisiting the effects of the high school equalization policy in South Korea. *Asia Pac. Educ. Rev.* 11, 83–96. <https://doi.org/10.1007/s12564-009-9061-9>.
- Carr, D., Wang, L.C., 2015. The Effect of After-School Classes on Private Tuition, Mental Health and Academic Outcomes: Evidence from Korea. Available at: <http://journals.sagepub.com/doi/abs/10.1177/0038038516677219?journalCode=soca>.
- Cho, J.S., 2015. Where are Korean Bright Students at School Studying? Qualitative Case Study on Curriculum of an Einstein Hakwon for Smart Students. Master's Dissertation. Jinju National University of Education. in Korean.
- Choi, Y., Park, H., 2016. Shadow education and educational inequality in South Korea: examining effect heterogeneity of shadow education on middle school seniors' achievement test scores. *Res. Soc. Stratif. Mobil.* 44, 22–32.
- Choi, H.-J., 2008. The effects of private supplementary tutoring for the university entrance. *J. Int. Econ.* 14 (1), 73–110 in Korean.
- Craig, R.A., Evers, C.J., 1982. Employers and educators: the shadow education system. In: Gold, G.G. (Ed.), *New Directions for Experiential Learning: Business and Higher Education—Towards New Alliances*. Jossey-Bass, San Francisco, CA, pp. 29–46.
- Dang, H.A., Rogers, F.H., 2008. The growing phenomenon of private tutoring: does it deepen human capital, widen inequalities, or waste resources? *World Bank Res. Obs.* 23 (2), 161–200.
- Dang, H.A., 2007. The determinants and impact of private tutoring classes in Vietnam. *Econ. Educ. Rev.* 26 (6), 683–698.
- Davis, G.A., Rimm, S.B., Siegle, D., 2011. *Education of the Gifted and Talented*, sixth ed. Pearson, Boston, MA.
- Dawson, W., 2010. Private tutoring and mass schooling in East Asia: reflections of inequality in Japan, South Korea, and Cambodia. *Asia Pac. Educ. Rev.* 11 (1), 14–24.
- de Castro, B.V., de Guzman, A.B., 2014. A structural equation model of the factors affecting Filipino university students' shadow education satisfaction and behavioral intentions. *Asia Pac. J. Educ.* 34 (4), 417–435.
- de Silva, W.A., Gunawardena, C., Jayaweera, S., Perera, L., Rupasinghe, S., Wijetunge, S., 1991. Extra-school Instruction, Social Equity and Educational Quality (Sri Lanka). The International Development Research Centre, Singapore.
- Doherty, C., Dooley, K., 2018. Responsibilising parents: the nudge towards shadow tutoring. *Br. J. Sociol. Educ.* 39 (4), 551–566.
- Economic Institute of Cambodia (EIC), Deline, S., Chamroeun, H., Sethy, Y., 2006. Local Public Services: Performance and Unofficial Fees. Phnom Penh.
- Enrich, S.R., 2014. German and Japanese education in the shadow—do out-of-school lessons really contribute to class reproduction? *IAFOR J. Educ.* 2 (2), 17–53.
- Enrich, S.R., 2017. Shadow Education and Social Inequalities in Japan: Evolving Patterns and Conceptual Implications. Springer, Dordrecht.
- Enrich, S.R., 2018. Worldwide Shadow Education and Social Inequality. When do shadow systems become a threat to equality of educational opportunities.
- Forsberg, E., Hallsén, S., Karlsson, M., Bowden, H.M., Mikhaylova, T., Svahn, J., 2020. Lärhjälp as shadow education in Sweden: The logic of equality in “a school for all”. *ECNU Review of Education*, 2096531120966334.
- Gordon, E.E., 1988. Centuries of Tutoring: A Perspective on Childhood Education. Doctoral Dissertation. Loyola University, Chicago.
- Hajar, A., 2018. Exploring Year 6 pupils' perceptions of private tutoring: evidence from three mainstream schools in England. *Oxf. Rev. Educ.* 44 (4), 514–531.
- Halliday, D., 2016. Private education, positional goods, and the arms race problem. *Polit. Philos. Econ.* 15 (2), 150–169. <https://doi.org/10.1177/1470594X15603717>.
- Hallsén, S., Karlsson, M., 2019. Teacher or friend? Consumer narratives on private supplementary tutoring in Sweden as policy enactment. *J. Educ. Pol.* 34 (5), 631–646.
- Harnisch, D.L., 1994. Supplemental education in Japan: juku schooling and its implication. *J. Curric. Stud.* 26 (3), 323–334.
- Hartmann, S., 2013. Education “home delivery” in Egypt: private tutoring and social stratification. In: Bray, M., Mazawi, A.E., Sultana, R.G. (Eds.), *Private Tutoring across the Mediterranean: Power Dynamics and Implications for Learning and Equity*. Sense, Rotterdam, pp. 57–75.
- Hattori, S., Nonoue, K., Tada, T., 2010. The relationship between nodding off during classes and academic performances, complaints of subjective symptoms, daily time management of junior high school students. *Jpn J. School Health* 52.
- Hof, S., 2014. Does private tutoring work? The effectiveness of private tutoring: a nonparametric bounds analysis. *Educ. Econ.* 22 (4), 347–366.
- Hu, Y., Fan, W., Ding, W., 2015. Does shadow education aggravate inequality of educational outcomes. *Peking. Univ. Educ. Rev.* 13 (3), 29–46.
- Ireson, J., 2004. Private tutoring: how prevalent and effective is it? *Lond. Rev. Educ.* 2 (2), 109–122. <https://doi.org/10.1080/1474846042000229458>.
- Jerrim, J., 2017. Extra Time: Private Tuition and Out-of-School Study, New International Evidence. The Sutton Trust, London.
- Jung, S.W., Jung, J.-H., 2019. Qualitative study about Korean students' accelerated learning in shadow education. *J. Learner Center. Curric. Instruct.* 19 (17), 663–691.
- Jung, J.H., Pinar, W.F., 2015. Conceptions of curriculum. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. SAGE, Thousand Oaks, pp. 29–46.
- Jung, J.-H., 2014. Hakbeolism: a historical and curriculum consideration of Korean test-focused education. *TCI* 11 (2), 48–66.
- Jung, J.-H., 2016. *The Concept of Care in Curriculum Studies*. Routledge, New York, NY.
- Jung, J.-H., 2018. Decolonizing educational/curriculum studies in East Asia: problematizing shadow education in South Korea. *Inter Asia Cult. Stud.* 19 (2), 269–281.
- Kim, H., Park, G., 2010. The impact of special purpose high schools on the private tutoring expenditure by the middle school students in Korea. *J Econ Educ* 19 (4), 75–104.
- Kim, Y.C., Jung, J.-H., 2019a. Conceptualizing shadow curriculum: definition, features and the changing landscapes of learning cultures. *J. Curric. Stud.* 51 (2), 141–161. <https://doi.org/10.1080/00220272.2019.1568583>.
- Kim, Y.C., Jung, J.-H., 2019b. *Shadow Education as Worldwide Curriculum Studies*. Palgrave Macmillan, New York.
- Kim, Y.C., Kim, P.S., 2012. Hakwon Does not Die; Fathers Die: Secrets of Korean Education that Obama did not Know. Academy Press, Paju, South Korea in Korean.
- Kim, Y.C., Kim, P.S., 2015. The Best Hakwons Top 7. Academy Press, Paju, South Korea in Korean.
- Kim, Y.C., Gough, N., Jung, J.-H., 2018. Shadow education as an emerging focus in worldwide curriculum studies. *Curric. Matters* 14, 8–30.
- Kim, J.H., 1985. An Analytic Study on the Change and Problems of the System of Entrance Examination. Masters' Thesis. Ewha University.
- Kim, M., 2003. Private institute education: competition and anxiety of the South Korean middle class. *Korean J. Soc. Educ.* 13 (3), 67–87.
- Kim, K.R., 2015. Korean Familial Expenditure, 15 Times Higher than Finland. Hankyeoye. Retrieved from: http://www.hani.co.kr/arti/economy/economy_general/709689.html.
- Kim, Y.C., 2016. *Shadow Education, Curriculum, and Culture of Schooling in South Korea*. Palgrave Macmillan, New York, NY.
- Klemm, K., Klemm, A., 2010. Ausgaben für Nachhilfe: Teurer und unfairer Ausgleich für fehlende individuelle Förderung. Bertelsmann Stiftung, Gütersloh.
- Kosunen, S., 2018. Access to higher education in Finland: emerging processes of hidden privatization. *Nordic J. Stud. Educ. Policy* 4 (2), 67–77.

- Lee, C.J., Lee, H., Jang, H.-M., 2010. The history of policy responses to shadow education in South Korea: implications for the next cycle of policy responses. *Asia Pac. Educ. Rev.* 11 (1), 97–108.
- Lee, J., 2007. Two worlds of private tutoring: the prevalence and causes of after-school mathematics tutoring in Korea and the United States. *Teach. Coll. Rec.* 109 (5), 1207–1234.
- Majumdar, M., 2014. *The Shadow School System and New Class Divisions in India*. Max Weber Stiftung, London.
- Malik, M.A., 2017. Shadow education: evolution, flaws and further development of the term. *Soc. Sci. Educ. Res. Rev.* 4 (1), 6–29.
- Manzon, M., Areepattamannil, S., 2014. Shadow education: mapping the global discourse. *Asia Pac. J. Educ.* 34 (4), 389–402.
- Marimuthu, T., Singh, J.S., Ahmad, K., Lim, H.K., Mukherjee, H., Oman, S., et al., 1991. Extra-school Instruction, Social Equity and Educational Quality. International Development Research Centre, Singapore.
- Mawer, K., 2015. Casting new light on shadow education: snapshots of juku variety. *Contemp. Jpn.* 27 (2), 131–148.
- Medway, E.J., 1995. Tutoring. In: Anderson, L.W. (Ed.), *International Encyclopedia of Teaching and Teacher Education*. Pergamon, Cambridge, pp. 271–274.
- Melot, L., 2007. *Le Marché du Soutien Scolaire [The Market for Schooling Support]*. Precepta, Paris. Summary Available on. <http://www.xerfi.fr/etudes/7SME04.pdf>. (Accessed 15 May 2011).
- Mischo, C., Haag, L., 2002. Expansion and effectiveness of private tutoring. *Eur. J. Psychol. Educ.* 17 (3), 263.
- Mori, I., 2012. *The effect of supplementary tutoring on students' mathematics achievement: A comparative study of Japan and the United States* (Doctoral dissertation). The Pennsylvania State University.
- Mori, I., Baker, D., 2010. The origin of universal shadow education: what the supplemental education phenomenon tells us about the postmodern institution of education. *Asia Pac. Educ. Rev.* 11 (1), 36–48. <https://doi.org/10.1007/s12564-009-9057-5>.
- Organization for Economic Co-operation and Development (OECD), 2012. PISA 2009 Technical Report. Retrieved September 2, 2017 from. <https://www.oecd.org/pisa/pisaproducts/50036771.pdf>.
- Ozaki, M., 2015. *A Juku Childhood: Children's Experiences in Juku Attendance and its Relation to Their Well-Being in Japan* (PhD dissertation). University of Bath, Claverton Down, UK.
- Paramita, S., 2014. "We follow the private tutors not the teachers": an ethnographic insight into educational practices among the students of an Indian city. *Int. J. Res. Soc. Sci.* 4 (4), 819–840.
- Park, S., Lim, H., Choi, H., 2015. "Gangnam Mom": a qualitative study on the information behaviors of Korean helicopter mothers. In: *iConference 2015 Proceedings*. Retrieved from. https://www.ideals.illinois.edu/bitstream/handle/2142/73636/124_ready.pdf?sequence=2&isAllowed=y.
- Park, H., Buchmann, C., Choi, J., Merry, J.J., 2016. Learning beyond the school walls: trends and implications. *Annu. Rev. Sociol.* 42 (1), 231–252. <https://doi.org/10.1146/annurev-soc-081715-074341>.
- Pinar, W.F., Reynolds, W.M., Slattery, P., Taubman, P.M., 1995. *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. Peter Lang, New York, NY.
- Rohlen, T.P., 1980. The juku phenomenon: an exploratory essay. *J. Jpn. Stud.* 6 (2), 207–242.
- Rohlen, T.P., 1983. *Japan's High Schools* (No. 21). Univ of California Press.
- Roesgaard, M.H., 2006. Japanese education and the cram school business: Functions, challenges and perspectives of the juku. Nordic Institute of Asian Studies Press, Copenhagen.
- Rowley, K.J., Edmunds, C.C., Dufur, M.J., Jarvis, J.A., Silveira, F., 2020. Contextualizing the achievement gap: assessing educational achievement, inequality, and disadvantage in high-income countries. *Comp. Educ.* 1–25.
- Silova, I., Bray, M., Zabulonis, A., Budiene, V., 2006. Education in a Hidden Marketplace: Monitoring of Private Tutoring. Education Support Program (ESP) of the Open Society Institute, Budapest.
- Silova, I., 2010. Private tutoring in eastern Europe and central Asia: policy choices and implications. *Compare* 40 (3), 327–344.
- Stevenson, D.L., Baker, D.P., 1992. Shadow education and allocation in formal schooling: transition to university in Japan. *Am. J. Sociol.* 97 (6), 1639–1657.
- Sun, M., Braeye, S., 2012. Comparing supplementary ethnic schools and the academic achievement of Chinese immigrant students in Quebec and Flanders. *Diversité Urbaine* 12 (1), 105–124.
- Takeda, T., 2005. *Juku and School: Possibility in Coexistence and Mutual Prosperity*. Juku Journal. Look Data publishing, Tokyo.
- Taubman, P.M., 2010. *Teaching by Numbers: Deconstructing the Discourse of Standards and Accountability in Education*. Routledge.
- Toyama-Bialke, C., 1997. *Jugendliche sozialisation und familiär Einflüsse in Deutschland und Japan: eine empirische Untersuchung zu jugendlicher Alltagsstruktur und elterlichen Erziehungspraktiken* (PhD Dissertation). University of Hamburg, Hamburg.
- Tseng, L.J., 1998. *Private Supplementary Tutoring at the Senior Secondary Level in Taiwan and Hong Kong*. MEd Dissertation. The University of Hong Kong.
- Tsuneyoshi, R., 2001. *The Japanese Model of Schooling: Comparisons with the United States*. Routledge Falmer, New York.
- Ventura, A., Jang, S., 2010. Private tutoring through the internet: globalization and offshoring. *Asia Pac. Educ. Rev.* 11 (1), 59–68.
- Vu, X.N.H., Ngo, M.T., Ta, M.T., Nguyen, M.T., 2011. Forms and Effects of Corruption on the Education Sector in Vietnam. Towards Transparency (TT) and Transparency International (TI), Hanoi.
- Watson, L., 2008. Private expectations and public schooling: The growth of private tutoring in Australia. Conference of the Australian Association for Research in Education (AARE).
- Williamson, B., 2013. The Future of the Curriculum: School Knowledge in the Digital Age. The MIT Press, p. 152.
- Winter, B.W., 1997. *Philo and Paul Among the Sophists*, vol. 96. CUP Archive.
- Wirtz, W., 1979. *Tuition-aid Revisited: Tapping the Untapped Resource*. National Manpower Institute.
- Yamamoto, Y., Brinton, M.C., 2010. Cultural capital in East Asian educational systems: the case of Japan. *Sociol. Educ.* 83 (1), 67–83.
- Yang, I., Kim, B.C., 2010. A qualitative case study on the school life of middle school students who go to private educational institute. *Korea Educ. Rev.* 16 (3), 117–153 in Korean.
- Yang, J.-H., 2012. Effects of private tutoring on college entrance. *J. Econ. Fin. Educ.* 21 (2), 27–52 in Korean.
- Zhang, W., Bray, M., 2017. Micro-neoliberalism in China: public-private interactions at the confluence of mainstream and shadow education. *J. Educ. Pol.* 32, 63–81.
- Zhang, Y., Xie, Y., 2016. Family background, private tutoring, and children's educational performance in contemporary China. *Chinese Sociol. Rev.* 48 (1), 64–82.
- Zhang, Y., 2013. Does private tutoring improve students' National College entrance exam performance? A case study from Jinan, China. *Econ. Educ. Rev.* 32, 1–28 c.
- Zhou, M., Kim, S., 2006. Community forces, social capital, and educational achievement: the case of supplementary education in the Chinese and Korean immigrant communities. *Harv. Educ. Rev.* 76 (1), 1–29.

History curriculum in the Anthropocene; how should we tell the story?

Kenneth Nordgren, Karlstad University, Karlstad, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	296
The curricular challenge and the reciprocal device	297
Storylines for the Anthropocene	299
The long duration	299
The short story	300
A new historical culture	300
Curricular strategies	301
Master narratives' rule and role	302
Possible alternatives	304
References	305

Introduction

In 1956, Swedish author and Nobel laureate in literature Harry Martinson (1998) published a book-length science fiction poem about a spaceship, *Aniara*. The story recounts that humanity has made the Earth uninhabitable through pollution and nuclear war. Today, the poem might be read as an anticipation of the Anthropocene's predicament—not only the story's plotline but also its depiction of how humans gain a god-like power to cross the boundaries of the planet, but lost control over the mechanisms and are now endangering the life-supporting systems of the planet (Rockström et al., 2009). The eponymous spaceship, the *Aniara*, begins to shuttle refugees to a new colony on Mars, but an accident veers the spaceship off course, destroying its steering ability. Slowly, the passengers start to realize that they are trapped on a relentless path into infinite space. This insight comes in phases, and the passengers react based on their different positions and experiences. At first, everyday routines continue, along with the trust that the ship's engineers will restore the steering technology. Meanwhile, Miman, a kind of interactive computer, offers solace through nostalgic images of the lost Earth. Soon, they enter a more nihilistic phase, and many passengers start living as if each day were their last. Hedonists clash with groups harboring fundamentalist ideals about penance, who believe that strict rituals will appease a power greater than themselves. Finally, the passengers succumb to apathy.

According to Mills (2000), dystopias are a sort of a sociological imagination of the fears of one's own time. As such, *Aniara* can be read as an allegory for the atomic bomb anxieties during the Cold War period and, for contemporary readers, the fright of an ecological collapse. Beyond these present threats, the poem also expresses a more subtle ambivalence toward the past and the future. As Koselleck (2004) notes, expectations of a crisis seem inherent to historical consciousness. In premodern times, the apocalypse was the final crisis, and it was unavoidable. Totally submissive to the will of God or gods, humans could do little to intervene. In modernity, crises have specific causes that bring possibilities for change: crises should be overcome, and they trigger evolution. While the brutal catastrophes of the twentieth century undermined the expectations that carried the European Enlightenment toward a future that was open to refinement (Hartog, 2015), the general notion that progress is achieved by overcoming the restraints of nature prevailed (Chakrabarty, 2018). Hence, *Aniara* is dystopic since it envisions the loss of not only an open horizon of expectation but also a relatable past. When the *Aniara's* flight from the ruined Earth fails, past deeds catch up with the lost refugees, but everything is already lost to infinity.

This chapter aims to discuss history curricula as an educational project in the Anthropocene epoch. Due to human influence, widespread and rapid changes have affected the atmosphere, ocean, cryosphere, and biosphere, fundamentally reshaping the conditions for life on Earth (Arias et al., 2021). These events are not dystopic fiction but, rather, they are the actual events that have situated us in an unprecedented predicament. As the effects of crossing planetary boundaries are irreversible, this seems to negate the notion of an open future—which, for Koselleck (2004), constitutes the core of modernity. However, it also seems to negate history as a relevant source of orientation, as narratives that separate human experiences from their environment (nature, climate, animals), is becoming obsolete. The present challenge is to find workable steering mechanisms to mitigate global warming and avoid a sixth mass species extinction. However, while investment in new technologies is necessary, it is not enough. According to The Intergovernmental Panel on Climate Change (IPCC), major lifestyle changes in how we live, travel, and consume are also required (Arneeth et al., 2019). Perhaps paradoxically, the mere thought of entering a new type of normality seems to trigger societal uses of history. The search for a new historical self-understanding is basically about ideals of *Bildung* or formation and should be of concern to anyone who takes an interest in the contemporary relevance of history education.

In outlining the scope of this chapter in greater detail, I must also remark on what not to expect. Combining education and history with the Anthropocene seems to raise expectations that might not be fulfilled. First, this *educational approach* does not

concern how best to teach or learn about climate change or sustainability. Rather, the task concerns the intended curriculum and strategies for incorporating new perspectives into the whole of historical learning. Second, *history teaching* in this context does not refer to specific environmental inquiries but to the impossible yet necessary task of teaching the bigger picture. As Shemilt (2000) observes, “The logic and methodological apparatus of historical enquiry can be applied to fragments and episodes in the past, but not to the past as a whole” (p. 85). The debate on climate change reveals that all claims are not equally valid and that, therefore, how the past is interpreted matters (Dunlap and McCright, 2015; Hunt, 2018). Drawing on the philosophy of history (Rüsen, 2017), humans’ historical consciousness affects how we orient in the world—that is, we cannot help but interpret the past in relation to what we expect from the future while we make sense of what is going on. Now, the Anthropocene is an unprecedented predicament that has caught curriculum-makers (including teachers) in a new situation where they might need to rethink what is relevant for study and what becomes significant when explaining historical processes, as well as who and what should be included or excluded as historical actors (Bladh, 2020; Nordgren, 2021; Van Straaten et al., 2016).

Third, taking the *Anthropocene* concept into the field of history causes epistemic confusion—not only because it is a geological epoch (which is still under investigation)—but also because it comes with a perspective that, by and large, has been a blind spot for the humanities’ generally anthropocentric self-understanding (Chakrabarty, 2009; Simon, 2019). The pacing and naming of geological epoch shifts are seldom of public concern. This shift does not take place in early evolution, or as slow-moving interglacial periods—nor does it originate in an alien meteorite collision, like the one that ended dinosaurs’ rule; rather, it results from recent human activity. The very term *Anthropocene* implies a matter of concern beyond the stratigraphic societies, that decides on geological successions. This shift is not traced from the slow-moving evolution of plankton—nor does it originate in an alien meteorite collision, like the crisis that ended dinosaurs’ rule; rather, it results from human activity (Crutzen, 2002). The *Anthropocene* is not an uncontested concept since many scholars have rejected a definition that starts with a generalized *Anthropos/humanity* and argued for a name that more specifically attributes blame. Among the alternative suggestions are the *Capitalocene*, pinpointing capitalism (Moore, 2017), and the *Plantationocene*, tracing the crisis’s causes to colonial agriculture (Haraway, 2015). This chapter will not engage much with this debate but will accept the *Anthropocene* as a historical periodization. The concept may be useful since it can be open to simultaneous perspectives: it can encapsulate a range of endangered planetary systems (not just climate change). Geologically, it does define a specific cause and time for transition—but without reducing the matter’s complexity to teleological closures in specific historical processes (Thomas, 2019). Moreover, it does indicate generalized human capabilities—but without denying that actions are linked to specific actors and regimes of powers. It situates humans at the center of the narrative, but it does so from a planetary perspective that decenters humans as entangled in—and dependent on—larger ecological systems. Naming an epoch is a discursive and homogenizing project, but not to the extent that the label will restrict what can be storied. As a product of periodization, the *Anthropocene* does not provide answers; rather, it offers an outlook from which to raise questions.

From here, we can reconnect to the curricular perspective with the help of Latour (2014), who formulated the simple but precise question: “How do we tell such a story?” (p. 3). He poses this question about the *Anthropocene* to the academic field of knowledge production, displaying the insecurity surrounding the consequences of producing new knowledge. This chapter suggests that Latour’s question is equally relevant to the field of educational recontextualization. History education research, so far, has produced surprisingly little on this matter (McGregor et al., 2021; Metzger and Harris, 2018). Historicity is intimately associated with the formation of modernity (Hunt, 2014), and curricular attempts to answer Latour’s question will eventually test whether history, as a specialized gaze, can also provide powerful knowledge for orientation in the age of the *Anthropocene*. To paraphrase Young (2003, p. 102): if, as curriculum theorists, we cannot tell what is important to know in the age of the *Anthropocene*, the epoch’s story will likely not be told, or it will be filtered by administrators and politicians’ pragmatic and ideological decisions.

The curricular challenge and the reciprocal device

Chakrabarty (2009) has proposed that the distinctions between *natural* and *cultural* history might collapse, arguing that from now on, planetary conditions are what will give history meaning. In 1916, when Dewey (1980) reflected upon history and geography as school subjects, he ascribed both of them with particular features, but rightly studied—that is, according to Dewey, jointly—they could offer students a holistic method with which to master the world. As complementary resources, they would reveal how human civilization is associated with its natural surroundings:

While geography emphasizes the physical side and history the social, these are only emphases in a common topic, namely, the associated life of men. For this associated life, with its experiments, its ways and means, its achievements and failures, does not go on in the sky nor yet in a vacuum. It takes place on the earth. This setting of nature does not bear to social activities the relation that the scenery of a theatrical performance bears to a dramatic representation; it enters into the very make-up of the social happenings that form history. Nature is the medium of social occurrences. It furnishes original stimuli; it supplies obstacles and resources. Civilization is the progressive mastery of its varied energies (p. 247).

Dewey’s argument for combining nature and history was not to critique modernity or Western civilization. For Dewey, the learning object was the trajectory of humans’ progressive mastery, for which nature was an obstacle or a resource. Still, he noticed that nature was lacking as a dimension in the historical explanation of culture. Today, policy-makers and curriculum-makers are

again more inclined to promote holistic or cross-disciplinary approaches, arguing that subject boundaries are obsolete or hindering change (e.g., Gough, 2021; Wals and Corcoran, 2006). However, this assumption that disciplines fragmentize a coherent whole obscures that knowledge beyond common sense derives from specialized ways of exploring the world (Deng, 2020; Young and Muller, 2016). Undoubtedly, an understanding of the Anthropocene presupposes interdisciplinary collaboration, in both science and education. Such collaboration presupposes, rather than replaces, epistemic boundaries (Nordgren, 2021). However, if history as subject-specific education fails to respond and does not provide useful tools with which to explore global challenges, it will solidify into dogma.

Education requires ideas about what it is to know something and how this knowledge relates to students' existing understandings. It requires the means to structure this content, making it accessible and conceivable for students at each stage of their progression. Young and Muller (2016) stress that education's fundamental task is to offer tools with which to explore the world and think beyond immediate circumstances. These tools are not randomly collected but rooted in specialized knowledge, which in turn is dependent on disciplinary communities. However, as Deng (2020) notes, education is also a normative project of *Bildung* or formation that involves students in the process of becoming a person among other persons and a member of specific communities. Making history education relevant in a changing world requires both a rooted idea of its specialized gaze and a notion of the times' apparent epochal challenges (Klafki, 1995).

The first questions about learning history are: Why bother? Why study history when previous experiences seem to expire so quickly? What can history possibly offer more than nostalgia when human activity had brought us to an unprecedented situation? One possible answer is that crises tend to trigger the human impulse to historicize predicaments. While climate change calls for immediate action, the willingness to change arguably requires a renewed historical self-understanding. Another possible answer is that crises are temporary, but learning to live with their consequences is a long-term engagement. This is a reason to reflect upon the direction of educational formation or *Bildung* in this new epoch. The human capacity for history seems inherent to our neurological disposition to remember and reflect, as well as our social urge to communicate through symbols. As Woolf (2019) points out in his book on global historiography, the human capacity to remember and communicate makes historical cultures inescapable. At the heart of any culture, one can discern the communication of collective memories—and how some of these discourses tend to dominate others since they have been selected and institutionalized in narratives, traditions, symbols, calendars, education, et cetera (Assmann, 1995; Rüsen, 2007). In his bestselling book *Sapiens: A brief history of humankind*, Harari (2014) collects examples of how politics, faith, law, and currency that signify complex societies are fundamentally based on expansive narratives. Again, how and by whom the Anthropocene is storied matters.

Obviously, disciplines and school subjects are inherently historical constructions and, thus, changeable. From this changeability, it does not follow that all knowledge claims are equally arbitrary or fortuitous. Everyday knowledge is situationally grounded, and it differs from specialized knowledge, which emanates from methodologies, categories, and epistemic communities that evaluate their findings. This logic of specialization creates borders that determine what is accepted as valid and relevant knowledge and what is not. Hence, disciplinary borders both hinder and drive advancement and communication. Their language can exclude, since access requires a special inauguration (Foucault, 2020), yet it is also inclusive since its concepts and methods allow for precision and transparency (Young and Muller, 2016). Specialized knowledge regimes develop a grammar that provides learners with tools to think beyond a given context, which Young and Muller (2016) call *powerful knowledge*. Thus, history, when disciplinarily framed, risks excluding relevant information (e.g., from geography) while also including an apparatus to produce knowledge that other perspectives cannot.

Keeping history teaching meaningful in a changing world requires a constant hermeneutical process that opens the school subject to a reciprocal exchange between what is to be explained and what that explanation requires (Fig. 1). That is: What is considered significant to explain, and what epistemic considerations does that explanation require? The discipline or subject explores the world through the tools of its grammar of specialization. This grammar can effectively enable a better understanding of the Anthropocene as a complex yet comprehensible phenomenon by providing perspectives on, for example, continuity and change in humans' relationships with nature. However, a problem arises in that the subject's epistemic traditions tend to disregard environmental relationships as being historically significant, and consequently, the disciplinary grammar might not be adjusted or fully equipped (Chakrabarty, 2009; Simon, 2019). Since the Anthropocene is a geological conceptualization and a predicament with

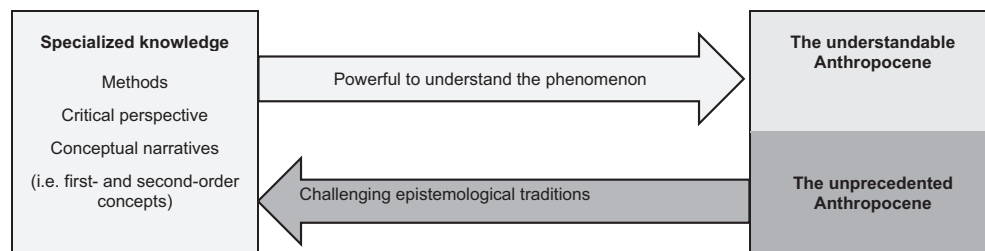


Fig. 1 A reciprocal device to consider the Anthropocene The epistemology of school history is based on a grammar of specialized knowledge. This grammar can be powerful when exploring and understanding the Anthropocene. However, the ambition to explore the historical dimensions of the Anthropocene will challenge epistemological traditions within the subject.

historical, political, social, and ethical consequences, its complexity challenges the borders of historical knowledge. Hence, the Anthropocene is not just another phenomenon to be explored and learned; rather, it requires explanations that make the historical landscape look different.

Nonetheless, the Anthropocene is not mysterious. Historians cannot hope to give a holistic explanation of it but, rather, to offer a relevant grammar of specialized knowledge with which to explore and add historical dimensions. Translating the Anthropocene into new narratives will conflict with and transform dominant ideas about historical significance, and previously marginalized events—such as climate change, volcanic eruptions, and the relocation of microbes, plants, and animals—will become more central. Such a shift in curriculum might open a floodgate to additional events and perspectives, which returns to the question of how to tell this story.

Storylines for the Anthropocene

When people reflect on their own time, they tend to sense an unprecedented period of transformative upheaval. Occasionally, the adjective *unprecedented* proves accurate. The Earth is now leaving behind the relatively stable climate of the Holocene period, which started after the most recent glaciation, about 12,000 years ago. Currently, the International Commission on Stratigraphy (ICS) is tracing the origins of this transition. The core of geological time division is the dramatic evolutions of life on Earth from its emergence, photosynthesis, and the Cambrian surge of complex life forms (Lewis and Maslin, 2018). Periodization is set on an official scale, based on methods and procedures for adding and changing time deviations. This system derives from the physical observation that any given stratum of rocks is older than the strata above it and younger than the strata below it. By comparing rocks and fossils throughout the world, geologists have divided and subdivided the Earth's 4.5-billion-year history into eons, eras, periods, epochs, and ages. These shifts mainly correspond to fossilized organisms' origins or extinction—which are, thus, traceable in sediment. The resultant stratigraphic question is: What physical reference point in the strata can define the beginning of this new epoch?

The candidates for such a *golden spike* are: the Neolithic revolution, with its consequences of domesticating plants and animals; the initiation of globalization through the Columbian Exchange in 1610; James Watt's improvement of the steam engine in 1788, which sped up the Industrial Revolution with the use of fossil fuel; the plastic revolution, which started in the 1920s; and the nuclear fallout after the Trinity test on July 16, 1945. In one sense, the Anthropocene does not differ from other geological types of periodization. Whether it will be accepted as an epoch depends on what can be scientifically established from sediment. In another sense, this periodization is novel. An unusual consequence of a geological label is that it inevitably signals that humanity's past and future choices will affect the future conditions for life on Earth. Since measures can be taken to reduce these effects, they are highly controversial. Lewis and Maslin (2018, p. 73) remark, "Geology plus humans equals politics." In fact, the concept's impact has already begun—regardless of any ICS decision, the Anthropocene has escaped seminar rooms and laboratories. Hence, as geological periodization has entered the realms of historical time, and the quest for a golden spike has centered sediments in contemporary political attention, it has also started to reshape the historical understanding of human cultures as not solely a human affair.

This scientific, conceptual, and political intricacy sets the scene for curriculum development as a reciprocal device. The geological concept brings new dimensions to historical thinking that can practically orient the discipline to discover climate, weather, natural forces, biodiversity, and environmental entanglement as fundamental dimensions in historical processes (Chakrabarty, 2018; Robin, 2013). However, geological timescales are based on different logics and methods and are not easily incorporated by historians' less paradigmatic and more contextual uses of periodization. Historical inquiries do not start or stop with golden spikes in rock strata and ice cores. A new periodization will trigger historians to trace origins and patterns from previous times.

Incorporating the Anthropocene into the historical gaze will rupture both established analytical models of inquiry and historical syntheses and overviews. For instance, notions of historical significance seem to be pushed in at least three different directions—all of which are slightly outside the comfort zone of history teaching. First, to perceive how natural and climatic continuity and change interweave with human cultures, a *longue durée* (long duration) of geological changes and early human history has been advanced. This perspective implies a timescale more familiar to archaeologists and geologists than historians (Ljungqvist et al., 2021; Robin and Steffen, 2007). Second, to grasp the Anthropocene's direct causes and effects, the field of study is compressed into the highly politicized timeframe of contemporary history (McNeill and Engelke, 2016). The third direction concerns historical consciousness; as the present reshapes the past, it also reshapes expectations for the future (Chakrabarty, 2018).

The long duration

To historicize nature, and to put climate change and the loss of biodiversity into perspective, a long storyline with a broad outlook from natural history or "Big History" is needed. Several periods of cold, drought, humidity, and storms have, arguably, influenced the whole of human history—from the exodus out of Africa to the droughts that doomed Akkad and Mayan culture, the Medieval warmth, and the cold period that preceded the French Revolution (Headrick, 2020). Climate change affects geographical areas differently since the respective underlying connections are complex, and because the current circumstances are unprecedented, these effects are difficult to predict. For example, from 7000 to 3000 BCE, the average temperature in the northern hemisphere was about 2–3 °C (Celsius) higher than the current average. Based on big data analysis, Zhang et al. (2011) show that cold periods correlated

with population dips from the 800s to the 1800s CE. War, conflicts, and migration are often closely related to climate change. Hence, fluctuations in temperature are not unprecedented, nor is species' extinction. What is unprecedented about the Anthropocene period is the causes, pace, and situation, with a human population exceeding 7 billion. The current rapid warming cannot be explained by the Earth's systems or solar activity (Arneth et al., 2019; Pachauri et al., 2014). To grasp these trends, points of reference are needed that only deep history can provide.

The short story

To understand the seriousness of the Anthropocene's not "just" being about how any human society has affected its surrounding ecosystem, humans' current effects on the planetary system must be understood. Whether its starting point is set at the Columbian Exchange or from James Watt's improved steam engine, history picked up speed after WWII. Since 1945, during a timespan no longer than it takes for a person to grow up and grow old, the world population has nearly tripled, while the animal wildlife population has decreased by more than half. This trend points toward a sixth mass extinction, including 75% of species going extinct in the next few centuries (Ceballos et al., 2015). The annual production of plastic has risen from 1 million tons to about 380 million tons, and microplastic particles are now virtually ubiquitous in our waterways and oceans. Nitrogen-synthesized fertilizer production increased from 4 million tons to more than 85 million tons, while chemical cycles are turning seas and rivers into dead zones. The number of cars has increased from 40 million to 850 million, and carbon dioxide has increased exponentially (McNeill and Engelke, 2016). Climate change is causing glaciers to melt, sea levels to rise, and severe weather events—such as floods, droughts, and hurricanes.

The short story of the Anthropocene is a controversial tale of the uneven consumption and destruction of natural resources. It is about economic globalization and the spread of wealth and education. The distribution of responsibility matters. Africa accounts for less than 3% of global emissions of greenhouse gases, but it experiences water stress due to climate change that could threaten the well-being of up to 600 million people (Malm and Hornborg, 2014). Extreme poverty has decreased over the last 40 years, but this trend is now reversing, partly due to the climate crisis (Arneth et al., 2019; Pachauri et al., 2014). The accumulation of wealth and the reduction of poverty have, so far, come with higher energy consumption. A burning question is how to get millions of people out of energy poverty without significantly increasing greenhouse gas emissions (González-Eguino, 2015). To reduce humans' impact on the Earth's systems, global and local solutions must be developed to even out injustices and eradicate poverty. Historical knowledge might offer powerful perspectives when exploring such a wicked problem and a perspective from which to discuss climate justice (Incropera, 2016).

A new historical culture

While awareness of the endangered planetary system is slowly sinking into our political and historical consciousness, the conclusion seems paradoxical: humans have become a force of nature and, thus, more powerful than ever—but this power has redefined nature from an infinite well of resources to a complex system of interconnected processes that are endangered. The growing insights that previous actions have mortgaged the future demand new narratives about historical origins, possible effects, and options (Chakrabarty, 2018, 2009). This perspective suggests a need for new historical orientations. One aspect of these necessary changes is the physical threat to memory culture, anthropology, and archeology. Rising sea levels are consuming coastal sites, and acidification and variations in precipitation are eroding buildings and statues (Fatoric and Seekamp, 2017). Climate change and deforestation are destroying ways of life for Indigenous peoples around the world (Chakravarty et al., 2012). Furthermore, rising soil temperatures are degrading the quality of archaeological evidence (Brunswig, 2014). Hence, traces of the past are literally eroding. Another aspect is the cultural consequences that can derive from both cognitive and emotional detachments from how we have learned to make sense of the past and future. Possibilities for change appear less as opportunities and more as necessities. Although cultural trends are unpredictable, threats also trigger investments in future technology, artificial intelligence, and genetic engineering, and crises might nurture new utopian visions.

Contradictory but connected trends seem to frame a new emerging historical culture. Thus, in Latour's (2014) question, the problem that must be storied is not simply about responding to the present situation or learning about the causes and consequences of the Great Acceleration. The question is an intellectual challenge to contemporary historicity, as well as critical theory and post-humanism. The obvious collapse might be the separation of culture and nature. Still, re-storying history must also find ways to encompass humans as both exceptional and non-exceptional, exploring both universal and particular causes and consequences. As Rüsen (2012) reminds us, the sweeping and widespread criticism of modernity and enlightenment is rather one-sided. While the division between nature and culture has been dominant, ideas of ecocriticism and other alternative narratives have always been parts of these multilayered traditions (Hinchman and Hinchman, 2001). Today, the fast-growing youth movement Fridays for Future combines an ecocentric vision and demands for climate justice with a rather technocratic reference to science and scientists (Zulianello and Ceccobelli, 2020).

Taking on the Anthropocene in history teaching is also about finding pragmatic strategies to tell this story. Possibly, the stories that come to be told are what shape the historical culture.

Curricular strategies

History as a school subject must be relatable to a range of curriculum goals (Chapman, 2021). Therefore, developing a history curriculum in the time of the Anthropocene raises a spectrum of questions: Does history have a specialized knowledge that can offer relevant perspectives on what it is to live and act in this new age? Can history provide useful answers about what brought us to this juncture? Can history offer relevant scenarios as inputs on what should be done from here? Answering such questions gives rise to new considerations; for instance, the postcolonial and Marxist critique of construing the human species as a unified historical actor evokes questions about agency and power structures since the consumption of the Earth's resources is not equally distributed (Malm and Hornborg, 2014; McEwan, 2021). In the field of posthumanism, the humanities' preoccupation with human culture is critiqued for decentering the larger ecosystem (Tyler, 2021; Wolfe, 2010), evoking questions of how to historize interrelations with non-humans. Eco-critical perspectives often direct their critique toward disciplinary borders and "the ranking of knowledge systems that places Western science at the top of an epistemological ladder" (Rose and Robin, 2004), which also evokes questions about different epistemic worldviews (Graeber and Wengrow, 2021).

Latour's (2014) question of how to story the Anthropocene is, indeed, an entry point into this challenge facing teachers and curriculum-developers. Teaching is practical, which shifts Latour's questions from rhetoric to practice: How do we, as teachers, recontextualize the long duration with the shorter story and in relation to the conflicting influences of a historical culture in crisis? A useful curriculum requires decisions and selections. One inspiration can be Seixas (2000), who has framed three possible teaching strategies to approach contested or controversial historical content. Using an example from Canadian colonialism, he discusses how teachers can address such historical problems. These strategies are outlined as generic epistemological positions, and as such, they offer a relevant starting point from which to reflect on integrating the Anthropocene into history curricula.

Seixas calls the first strategy "enhancing collective memory," which seeks to provide students with the one most accurate story, based on current knowledge. If different interpretations are available, one interpretation likely has more credibility. This is the best story and the one that should be taught. This strategy's advantages are that it gives teachers a cohesive storyline to tell and students an authorized point of view. Shared narrative creates a common experience and, perhaps, even a sense of a moral trajectory. The second strategy is to engage students in exploring the past, and drawing their own conclusions based on evidence. Rather than receiving a digested narrative, students learn to be critical but also to understand the limits of interpretation. This strategy is both educative and democratic, and as such, it fits well with an approach that Seixas calls "disciplinary." The third strategy accepts different opinions and focuses on the causes of the historical controversy, asking: Who are the advocates, and what are their interests? Seixas categorized this strategy as a "postmodern strategy" since it is not concerned with what happened in the past but, rather, how the past is narrated and used by different interest groups. This approach also invites students into inquiries—not about the past but about the use of historical narratives in contemporary culture.

Seixas is, to some extent, open to drawing from all three strategies. However, he advocates for the second, the disciplinary strategy. Seixas critiques the best story strategy since it cannot adequately present the complexity of a historical problem. Instead of enhancing critical historical thinking, it supplies a one-sided and, therefore, seductive collective memory. The postmodern strategy offers the advantage of drawing attention to narratives' seduction and pervasiveness. However, it does not engage with the past or actual conflict—only with its aftermath. Seixas warns that this approach can lead students to relativistic assumptions.

Including climate change, the environment, and biodiversity into curricula seems to thwart any ambition to tell one "best story." The Anthropocene adds more events, accents, motives, and inquiries to an already crowded agenda; new perspectives on time, agency, and causes seem antithetical to a cohesive collective memory. The postmodern strategy seems to be more useful in deconstructing anthropocentric narratives. However, per Seixas's argument, the postmodern approach would still fail to understand how the Anthropocene is based on real events. The disciplinary strategy would engage students in exploring the Anthropocene as a range of historical problems. As Gibson (2021) has noted, historical thinking concepts are well suited to helping teachers' inquiries of the Anthropocene and its significance, beginnings, causes, consequences, and ethical considerations. For instance, was the French Revolution caused by bad weather (Fagan, 2000), and did British industrialization bring the American bison to the brink of extinction (Taylor, 2011)? Hence, the "best narratives" strategy seems the most obviously dismissible. One aspect on which both the disciplinary and postmodern strategies agree is implementing multi-perspective curricula.

However, the choice of teaching strategy might not be so straightforward. Neither the disciplinary nor the postmodern strategy actually provides a basis for selection; everything that can be inquired about is a potential interest, and the means of deconstruction do not provide clues about what to construct. Nonetheless, teaching requires criteria (implicit or explicit) to select what is significant enough to be taught. Hence, the demarcation lines that separate the disciplinary from the postmodern, and the best-story approaches seem too definite, tending to obscure how all 'inquiries, deconstructions and collective memories' are embedded in the meaning-making processes that precede them. The curricular problem is not whether inquiries about the Anthropocene are teachable; rather, the problem is how to establish conventions in which the Anthropocene, as an epoch and a perspective, become inevitable and integrated parts of course activities. Consider a parallel with intercultural education: while nationalistic narratives play a less predominant role in many national curricula, educational metanarratives are still structured by narrative scripts that center nations and decenter migration and cultural encounters as core topics of exploration (see Carretero et al., 2012; Nordgren, 2017). Following this logic, the crux of renewing the curriculum is to find strategies that integrate the Anthropocene as a historically significant problem.

Historical significance is a prominent theme in the educational research literature. The process of selecting content is neither random nor arbitrary. Notions of what is significant can be linked to events of special importance, profundity, quantity, and durability (Lévesque, 2008). However, as Seixas (1997, p. 22) underlines, this process of selection does not emanate from the past itself but, rather, from the “interpretative frames and values of those who study it—ourselves.” Thus, we curriculum-developers engage with a past filtered through mental frameworks shaped by personal and educational experiences (Lévesque, 2008). Empirical studies on significance have often relied on questionnaires and interviews that collect what students, teachers, or textbooks identify as most important. According to Barton (2005), such methods are “well-suited to this field of inquiry because issues of significance are inherently about ranking—they involve deciding which topics are more important than others” (p. 12).

However, efforts to integrate the Anthropocene (and intercultural learning) into teaching reveal that significance is a gatekeeper that goes beyond individual interpretative frames and curriculum-developers’ personal and educational experiences. To analyze its logic requires, more than a ranking list, a critical methodology with which to read between the lines in order to transcend individual preferences and observe general patterns. When we make these connections, the content and inquiry tasks that structure our courses can reveal the structure of a master narrative; when we consider the range of events and processes that we find hard to integrate not just as missing parts, but as anomalies, we expose the paradigmatic framing of the historical regime (Hunt, 2014). Seixas’s repertoire of teaching strategies operates within the same overarching master narrative. The three identified strategies are helpful in discussing epistemic stances but insufficient as means to rethink what should be framed as historically significant. Hence, we must ask whether the Anthropocene indicates a new master narrative in history teaching.

Master narratives’ rule and role

Historians and history educators are generally skeptical about narratives with transcendent or universal ambitions—and for good reason. History has frequently been used by states and political groups to legitimize ideologies and homogenize worldviews, whether nationalistic, Marxist, or liberal. Agreeing with Sirkka Ahonen (2017) is easy when she dismisses such master narratives as “unsustainable as truths and socially exclusive as identity builders” (p. 58). Samuel Wineburg (2013) warns that critiques of Western-biased curricula should not try to replace one master narrative with another. For Evans (2011), history educators are obligated to debunk the myths of master narratives in order to promote multi-perspectivity. Hence, the ideal history curriculum seems to be free from master narratives, but is a curriculum with purely neutral ordering principles possible? Can historical synthesis be learned without overarching meaning transcending from the sum of many multi-perspective voices (cf. Ankersmit, 2001)?

The scholarly literature tends to present the notion of master narratives with rather persuasive definitions, attaching specific value-laden attributes to the concept. Such specific contextualization is often suggested as constitutive features. According to Moyers (1990), a master narrative is an ideological script imposed by an authority on everyone else. For Woodson (2017), it is a dominant mythology that mutes, erases, and neutralizes the features of racial struggle. According to Carretero and Bermudez (2012), it is a simplification and exclusion that are essential for nations and nationals. When Carretero and Van Alphen (2014) empirically examined how master narratives affect students’ learning, they defined the concept as the opposite of disciplinary thinking.

However, rethinking a curriculum seems to require a further contemplation of how narratives structure historical consciousness. First, I suggest that we avoid equating the organization of the story with the actual plot that gives it a specific meaning (see White, 2014). Nationalistic history persists not only because of its ideological content but also the efficacy that a nation, as an overarching idea, can be used to organize the historical who, what, where, and when. A master narrative is not the same as the myth that it is narrating but, rather, the logic by which it assembles different narrative elements into a shared meaning. For example, an account becomes a master narrative not by reproducing stories about national greatness but by presenting the nation as the obvious protagonist. As Halverson et al. (2011) suggest, (master) narratives are not just stories centered on specific and situated events but, rather, systems of stories that form culturally recognizable trajectories of overcoming conflicts. Likewise, I suggest that we avoid associating master narratives with only some cultures, that nation-states produce homogenizing narratives while advocates for cosmopolitanism do not, that Western modernity makes universalistic claims while postmodern and postcolonial claims are only local, and that only authoritarian regimes suppress alternative narratives while liberal democracies are open to multi-perspectives. The argument against such an oversimplified association seeks not to deny differences between historical, ideological, and epistemological positions but to acknowledge that the distance between them occurs as tensions, not dichotomies (cf. Mignolo, 2002). Klein (1995) presents an interesting example of culturally transcendent meanings’ critical persistence: Lyotard criticized historical narratives for reducing enslaved people’s local suffering to grand stories, while Clifford Geertz feared that Lyotard’s construct would reduce enslaved people to a locally unified suffering object. The paradox of writing history in a postmodern culture is well expressed by Klein (1995, p. 298): “We are living a golden age of global narratives in which universal history is not simply possible, but unavoidable.”

In anthropology and narrative psychology, stories are often viewed as constitutive for people’s identity formation (Bruner, 1990; Polkinghorne, 1988). Humans have the cognitive disposition—or burden, according to Nietzsche (2008)—needed to connect past experiences and utterances into coherent narratives that give or even impose meaning. Since stories are shared through language and communicative exchanges, they are what Geertz (2008) describes as the fabric that creates culture. This role implies that we live in a web of narrative significance spun from countless stories. To distinguish functions and systems within this narrative network, McLean and Syed (2015) suggest a threefold distinction: *master narratives*, *personal narratives*, and *alternative narratives*. A master narrative is not a specific story but, rather, dominant scripts that contain collective storylines, ranging from a group’s history to

notions of what inhabiting a particular social category means. These scripts can be identified in cultural products and discourses—for example, in curricula and textbooks (Hammack, 2010). Personal and alternative narratives are influenced by, but not reducible to, these dominant storylines. Tensions and contradictions occur between different scripts, and as McLean and Syed (2015) emphasize, master narratives are always present alongside narratives that will counter them:

As individuals construct a personal narrative they negotiate with and internalize these master narratives—they are the material they have to work with to understand how to live a good life. For many individuals whose lives fit in with societal structures, these master narratives are functional and unproblematic. Others, however, may need to construct or adopt an alternative narrative, which at minimum differs from, and at maximum resists, a master narrative (p. 320).

This sectioning of master, personal, and alternative narratives can help curriculum-makers explore the Anthropocene's integration into history education by identifying interrelated narrative structures that transcend the teaching structures Seixas discusses. Master and alternative narratives are not imposed as "the best story" but as storylines that strongly influence their audiences' orientation toward what seems central or peripheral, historically significant, and socially valuable. For example, these storylines can support ontologies, such as the hand of God, the invisible hand of the market, or history's expression in class struggle. They can collect diverse but influential ideological, scientific, and esthetic traditions—such as the Enlightenment or postmodernity—and they frame emerging narratives, such as the Capitalocene or the Anthropocene.

However, the relationship between alternative and master narratives is complex. McLean and Syed (2015) characterize this relationship as ironic: the challenge that the alternatives pose might influence and even change the master narrative. Most often, the alternative narrative paradoxically reinforces what it opposes since the act of expressing alternatives gives dominant scripts recognition and opportunities to adapt. For example, migration studies have—despite researchers' intentions—been used to mirror nationalist images of normality (Wimmer and Schiller, 2003), and endeavors to strengthen teaching in the United States about the civil rights movement might paradoxically reinforce notions that "the race-problem" is in the past and has been overcome (Woodson, 2017). When Seixas (2000) introduced his three alternative strategies, he referred to a specific event concerning a treaty between the Nisga'a people and the government of British Columbia, Canada. Now, this story describes a (real) event that makes sense in any contemporary history class. The topic might be provocative to a nationalistic agenda, but in a multicultural society, such cases are to be expected. As such, it is an alternative narrative that does not undermine the dominant scripts of methodological nationalism, just as single inquiries on climate change will not rupture the anthropocentric embeddedness of history.

The basic assumption is that master narratives cannot be escaped. Some ordering scripts, patterns of interpretation, understandings of significance, and values will dominate over others; however, they will never do so with total hegemony. If an alternative discourse becomes influential, it might structure a new master narrative. When exploring the possibilities for a history curriculum in the Anthropocene, slightly modifying McLean and Syed's (2015) model by adding one more position could be useful; as discussed by Somers (1992), cultural and institutional formations larger than individuals are also part of this game. Collectives, networks, and institutions are both recipients and producers of narratives. Academic disciplines and curriculum-makers produce conceptual narratives that interact with—but are not reducible to—sociocultural or national master narratives. Further, digital media have become accessible to a diverse range of interest groups who have added numerous public narratives that can both undermine and support master narratives (Levinger, 2018).

This modified model (Fig. 2) might help discern relevant positions in the curricular context when introducing the Anthropocene to history teaching. Since national and regional curricula reproduce master narratives to which ecological perspectives, migrations, and cultural encounters are anomalous, they seldom surface in history classrooms.

However, such alternative narratives can rupture stability because they are produced by public actors who can influence the historical culture. Curriculum-makers (such as disciplines and other epistemic communities) can, from their institutionalized positions, revise paradigmatic discourses. These conceptualized alternatives form the narratives that social movements use to pressure the political and economic systems. Networks with other interests can also use social media to produce public narratives that undermine or discredit scientific knowledge. Hence, when we curriculum-makers regard education and disciplines as part of public narratives, we can discern how they can simultaneously be fundamental to both the dominant master narrative and critical alternatives. Adolescents form personal narratives in the turmoil of competing narratives. If history education cannot constructively contribute to students' ability to inquire and deconstruct different narratives, and also to grasp a bigger picture, the subject might become irrelevant. The growing discrepancy between experiences of an endangered planet and a historical curriculum detached from these experiences might eventually marginalize history as powerful knowledge.

Telling the best story according to evidence might lay a good foundation for understanding the big picture of how humans spread across the world and fundamentally changed the face of the Earth—first locally, then globally, and now on a planetary level. Meanwhile, a disciplinary approach might inquire about the reasons behind events, such as the Lisbon earthquake of 1755, or the Lyton Creek wildfire in Canada in 2021. Conversely, the postmodern strategy can usefully deconstruct the oil and gas industries' lobbying compared to how the tobacco industry tried to dismiss the harms of smoking. However, none of these strategies have a grammar with which to re-evaluate what to study. Therefore, some isolated environmental events will be tolerated and thought to fit within the current master narrative. Then again, alternative narratives without ordering principles will not be embedded into

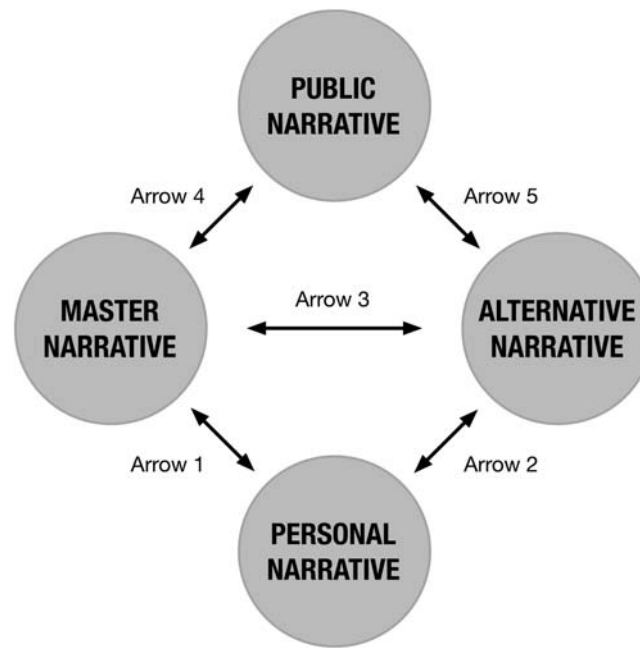


Fig. 2 The modified master narrative model. McLean and Syed's (2015) master narrative model describes personal narratives as derived from a balance between master narratives and alternative narratives. Drawing on Somers (1992), a relation to the public narrative is added. Public narratives are attached to cultural and institutional formations larger than the single individual. The bidirectional arrows indicate that individuals and public groups (e.g., disciplines and curriculum-makers), through their narratives, both internalize and create or sustain master narratives (arrows 1 & 4) and alternatives (arrows 2 & 5), and that alternative narratives are related to one another (arrow 3).

the synthesized line of development that lays the foundation for content selection and explanatory perspectives. These shortcomings of the three teaching strategies again raise the question of whether the Anthropocene indicates a need for a new master narrative in history teaching.

Possible alternatives

Sometimes, art and fiction dispense heuristic insights with which to explore how anticipating the future affects historical understanding. The condensed narrative of Harry Martinson's doomsday poem encompasses both the diversity among the passengers and their shared fate inside the *Aniara*. The spaceship, lost without steering capability, might also indicate another and an even more disturbing reflection: despite all scientific evidence revealing the need for immediate action, national and global authorities and enterprises are still reluctant to act. We who live in the part of the world that consumes most of its resources still have not incorporated this evidence into the narratives by which we orient our lives.

Historical knowledge is horizontally structured, allowing different and even competing theoretical and ontological outlooks within its disciplinary gaze. Nevertheless, historical consciousness seems to be affected by deep narrative patterns with paradigmatic qualities. These patterns form a web of coherence and significance that could be termed a *genetic historical regime*. School history can sometimes appear as one damned thing after another, but this appearance is an illusion. Selecting and pacing events to study is not a random process. It follows an established trajectory that tends to overstate geopolitical entities and marginalize migration, cultural encounters, nature, and climate as historical factors. In 1958, Braudel (1972) advocated for a long-term geographical perspective as a way to illuminate the environmental effect on human social and economic life, as well as on human mentality. History, he argued, is also the "long centuries of climates, of vegetations, of animal populations, of types of crop, of slowly constructed equilibria" whose elements cannot be transformed "without the risk of endangering everything" (p. 172). Today, we know the equilibria are raptured and that everything is endangered, but we have not found the key to writing this story.

Reforming the history curriculum must be about what signifies a sound historical culture in the Anthropocene. Certainly, we cannot just request or order a new and ready-made master narrative and, as curricula are ultimately politically framed, we should be very careful what we wish for. Based on specialized knowledge, the climate threat and human activities' impacts on the Earth's future have been detected. Teachers and educators must take responsibility as curriculum-developers and use a range of educational strategies to develop relevant and concrete inquiries, deconstructions, and individual expert narratives—but we must also dare to explore forms and principles for new synthesizing perspectives.

Several world-history projects have explored alternative narrative structures, including McNeill and McNeill's (2003) exploration of globalized networks. Frankopan (2015) has used the Silk Road as a way to unravel global development, and Big History is a multidisciplinary project to read history through big data (Christian, 2011; Bain, 2019; Galdi and Armitage, 2014). However, the environmental history field is, according to Hornborg (2010), reluctant to theoretically relate environmental change to political, economic, and cultural power structures. In teaching history, the future is most often a silent but salient point of reference. It often looms at the end of textbooks as fragmentized reports of contemporary threats and conflicts. Without a reciprocal device, the Anthropocene predicament will be added as events in the final chapters of the textbooks', while the subject outdates itself. By challenging and exploring narrative structures, we might develop ways to synthesize historical outlooks. This exploration might lie beyond the comfort zone of multiperspectivity, but it is not an ambition to tell a singular story.

If master narratives are understood not as hegemonic narratives but as scripts that help and, thus, also direct our ability to create coherence and meaning, we realize that the question is incorrectly formulated. The question is not *if* but *which*. Shemilt (2000) wrote about the need for a polythetic framework that can paint a big picture—a multidimensional, complex image “that, first gives perspective to the present by prompting us to take the long view and to look beyond what is happening to what might be going on, and second, allow us to fit present phenomena within a narrative and a polythetic framework” (p. 100). However, a framework is not a kaleidoscope of random patterns. The unavoidable submission to master narratives is both limiting and enabling: without the possibility to synthesize generalizations, our observations would only be individual fragments, but if the notions of historical significance are rigidly scripted in curricula and traditions, our outlook will be limited. Narratives are antennae through which we sense the world, so when the world changes, we must tell new stories to orient ourselves.

References

- Anonon, S., 2017. The lure of grand narratives: a dilemma for history teachers. In: Åström Elmersjö, H., Clark, A., Vinterek, M. (Eds.), *International Perspectives on Teaching Rival Histories*. Springer, New York, pp. 41–62.
- Ankersmit, F.R., 2001. *Historical Representation*. FR Ankersmit. Stanford University Press, Stanford.
- Arias, N., Bellouin, E., Coppola, E., et al., 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group 14 I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Technical Summary. Intergovernmental Panel on Climate Change (IPCC), Geneva.
- Arnell, H., Barbosa, T., Benton, T., et al., 2019. IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems. Technical report. Intergovernmental Panel on Climate Change (IPCC), Geneva.
- Assmann, J., 1995. Collective memory and cultural identity. *New Ger. Critiq.* 65, 125–133.
- Bain, R.B., 2019. Using big history to meet challenges in teaching and learning in the United States. In: *The Routledge Companion to Big History*. Routledge, London and New York, p. 372.
- Barton, K.C., 2005. “Best not to forget them”: secondary students’ judgments of historical significance in Northern Ireland. *Theory Res. Soc. Educ.* 33, 9–44.
- Bladh, G., 2020. GeoCapabilities, Didactical analysis and curriculum thinking—furthering the dialogue between Didaktik and curriculum. *Int. Res. Geogr. Environ. Educ.* 29, 206–220.
- Braudel, F., 1972. History and the social sciences. *Econ. Soc. Early Mod. Eur.* 18, 725–753.
- Bruner, J., 1990. Culture and human development: a new look. *Hum. Dev.* 33, 344–355.
- Brunswig, R.H., 2014. Risks and benefits of global warming and the loss of mountain glaciers and ice patches to archeological, paleoclimate, and paleoecology resources. *Ecol. Quest.* 20, 99–108.
- Carretero, M., Bermudez, A., 2012. Constructing histories. In: *Oxford Handbook of Culture and Psychology*. Oxford University Press, New York, pp. 625–646.
- Carretero, M., Van Alphen, F., 2014. Do master narratives change among high school students? A characterization of how national history is represented. *Cogn. Instr.* 32, 290–312.
- Carretero, M., Asensio, M., Moneo, M.R. (Eds.), 2012. *History Education and the Construction of National Identities*. Information Age, Charlotte, NC.
- Ceballos, G., Ehrlich, P.R., Barnosky, A.D., et al., 2015. Accelerated modern human-induced species losses: entering the sixth mass extinction. *Sci. Adv.* 1, e1400253.
- Chakrabarty, D., 2009. The climate of history: four theses. *Crit. Inq.* 35, 197–222.
- Chakrabarty, D., 2018. Anthropocene time. *Hist. Theory* 57, 5–32.
- Chakravarty, S., Ghosh, S., Suresh, C., Dey, A., Shukla, G., 2012. Deforestation: causes, effects and control strategies. In: *Global Perspectives on Sustainable Forest Management*. InTech, London, UK.
- Chapman, A. (Ed.), 2021. *Knowing History in Schools: Powerful Knowledge and the Powers of Knowledge*. UCL Press, London.
- Christian, D., 2011. *Maps of Time: An Introduction to Big History*. University of California Press, Berkeley.
- Crutzen, P.J., 2002. The “Anthropocene”. *J. Phys.* IV 12, 1–5.
- Deng, Z., 2020. Liberal education, Bildung and theory of content. In: *Knowledge, Content, Curriculum and Didaktik*. Routledge, London and New York.
- Dewey, J., 1980. The Middle Works of John Dewey. In: *Essays on Philosophy and Education, 1916–1917*, vol. 10. SIU Press, Carbondale, pp. 1899–1924.
- Dunlap, R.E., McCright, A.M., 2015. Challenging climate change. In: Dunlap, R., Brulle, R. (Eds.), *Climate Change and Society: Sociological Perspectives*. Oxford University Press, New York, pp. 300–331.
- Evans, R., 2011. The wonderfulness of us (the Tory interpretation of history). *Lond. Rev. Books* 33, 9–12.
- Fagan, B., 2000. *The Little Ice Age: How Climate Made History 1300–1850*. Basic Books, New York.
- Fatoric, S., Seekamp, E., 2017. Are cultural heritage and resources threatened by climate change? A systematic literature review. *Clim. Change* 142, 227.
- Foucault, M., 2020. *Power/Knowledge*. Routledge, London and New York.
- Frankopan, P., 2015. *The Silk Roads: A New History of the World*. Bloomsbury Publishing, London.
- Geertz, C., 2008. *Thick Description: Toward an Interpretive Theory of Culture*. Routledge, London and New York.
- Gibson, L., 2021. Open peer review. Part of the problem or part of the solution: historical thinking and the Anthropocene. *Hist. Weekly* 8, 9.
- González-Eguino, M., 2015. Energy poverty: an overview. *Renew. Sustain. Energy Rev.* 47, 377–385.
- Gough, A., 2021. *Education in the Anthropocene*. Oxford University Press, New York.

- Graeber, D., Wengrow, D., 2021. *The Dawn of Everything: A New History of Humanity*. Farrar, Straus, and Giroux, New York.
- Guldi, J., Armitage, D., 2014. *The History Manifesto*. Cambridge University Press, Cambridge.
- Halverson, J.R., Goodall, H.L., Corman, S.R., 2011. *Master Narratives of Islamist Extremism*. Palgrave Macmillan US, New York.
- Hammack, P.L., 2010. Identity as burden or benefit? Youth, historical narrative, and the legacy of political conflict. *Hum. Dev.* 53, 173–201.
- Harari, Y.N., 2014. *Sapiens: A Brief History of Humankind*. Random House, New York.
- Haraway, D., 2015. Anthropocene, Capitalocene, Plantationocene, Chthulucene: making kin. *Environ. Humanit.* 6, 159–165.
- Hartog, F., 2015. *Regimes of History: Presentism and Experiences of Time*. Columbia University Press, New York.
- Headrick, D.R., 2020. *Humans Versus Nature: A Global Environmental History*. Oxford University Press, New York.
- Hinchman, L.P., Hinchman, S.K., 2001. Should environmentalists reject the enlightenment? *Rev. Polit.* 63, 663–692.
- Hornborg, A., 2010. Toward a truly global environmental history: a review article. *Review* 33, 295–323.
- Hunt, L., 2014. *Writing History in the Global Era*. WW Norton & Company, New York.
- Hunt, L., 2018. *History: Why It Matters*. John Wiley & Sons, Hoboken.
- Incropera, F.P., 2016. *Climate Change: A Wicked Problem: Complexity and Uncertainty at the Intersection of Science, Economics, Politics, and Human Behavior*. Cambridge University Press, Cambridge.
- Klafki, W., 1995. Didactic analysis as the core of preparation of instruction. *J. Curric. Stud.* 27, 13–30.
- Klein, K.L., 1995. In search of narrative mastery: postmodernism and the people without history. *Hist. Theory* 34, 275–298.
- Koselleck, R., 2004. *Futures Past: On the Semantics of Historical Time*. Columbia University Press, New York.
- Latour, B., 2014. Agency at the time of the Anthropocene. *New Lit. Hist.* 45, 1–18.
- Lévesque, S., 2008. Thinking Historically: Educating Students for the Twenty-First Century. University of Toronto Press, Toronto.
- Levinger, M., 2018. Master narratives of disinformation campaigns. *J. Int. Aff.* 71, 125–134.
- Lewis, S.L., Maslin, M.A., 2018. *The Human Planet*. Yale University Press, New Haven.
- Ljungqvist, F.C., Seim, A., Huhtamaa, H., 2021. Climate and society in European history. *WIREs Clim. Change* 12, e691.
- Malm, A., Hornborg, A., 2014. The geology of mankind? A critique of the Anthropocene narrative. *Anthr. Rev.* 1, 62–69.
- Martinson, H., 1998. *Aniara: An Epic Science Fiction Poem*. Story Line Press, Brownsville.
- McEwan, C., 2021. Decolonizing the Anthropocene. In: Chandler, D., Müller, F., Rothe, D. (Eds.), *International Relations in the Anthropocene: New Agendas, New Agencies and New Approaches*. Springer International Publishing, Cham, pp. 77–94.
- McGregor, H.E., Pind, J., Kam, S., 2021. A “wicked problem”: rethinking history education in the Anthropocene. *Rethink Hist.* 25, 1–25.
- McLean, K.C., Syed, M., 2015. Personal, master, and alternative narratives: an integrative framework for understanding identity development in context. *Hum. Dev.* 58, 318–349.
- McNeill, J.R., Engelke, P., 2016. *The Great Acceleration: An Environmental History of the Anthropocene Since 1945*. Harvard University Press, Cambridge.
- McNeill, J.R., McNeill, W.H., 2003. *The Human Web: A Bird's-Eye View of World History*. W.W. Norton & Company, New York.
- Metzger, S.A., Harris, L.M.A., 2018. *The Wiley International Handbook of History Teaching and Learning*. John Wiley & Sons, Hoboken.
- Mignolo, W., 2002. The enduring enchantment (or the epistemic privilege of modernity and where to go from here). *S. Atl. Q.* 101, 927–954.
- Mills, C.W., 2000. *The Sociological Imagination*. Oxford University Press, New York.
- Moore, J.W., 2017. The Capitalocene, part I: on the nature and origins of our ecological crisis. *J. Peasant Stud.* 44, 594–630.
- Moyers, B., 1990. Exposing the “Master Narrative”: Teaching Toni Morrison's the Bluest Eye. *Feminist Teacher*. <https://feministteacher.com/2010/04/13/exposing-the-master-narrative-teaching-tonimorrison-the-bluest-eye>.
- Nietzsche, F., 2008. *On the Use and Abuse of History for Life*. Dodo Press, Gloucester.
- Nordgren, K., 2017. Powerful knowledge, intercultural learning and history education. *J. Curric. Stud.* 49, 663–682.
- Nordgren, K., 2021. Powerful knowledge for what? History education and 45-degree discourse. In: Chapman, A. (Ed.), *Knowing History in Schools*. UCL Press, London, pp. 177–201.
- Pachauri, R.K., Allen, M.R., Barros, et al., 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change (IPCC), Geneva.
- Polkinghorne, D.E., 1988. *Narrative Knowing and the Human Sciences*. SUNY Press, New York.
- Robin, L., Steffen, W., 2007. History for the Anthropocene. *Hist. Compass* 5, 1694–1719.
- Robin, L., 2013. Histories for changing times: entering the Anthropocene? *Aust. Hist. Stud.* 44, 329–340.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., et al., 2009. Planetary Boundaries: exploring the safe operating space for humanity. *Ecol. Soc.* 14 (2).
- Rose, D.B., Robin, L., 2004. The ecological humanities in action: an invitation. *Aust. Humanit. Rev.* 31–32. <http://australianhumanitiesreview.org>.
- Rüsen, J., 2007. *Time and History: The Variety of Cultures*. Berghahn Books, New York.
- Rüsen, J., 2012. Forming historical consciousness—towards a humanistic history didactics. *Antiteses* 5, 519–536.
- Rüsen, J., 2017. *Evidence and Meaning: A Theory of Historical Studies*. Berghahn Books, New York.
- Seixas, P., 1997. Mapping the terrain of historical significance. *Soc. Educ.* 61, 22–27.
- Seixas, P., 2000. Schweigen! Die Kinder! or, does postmodern history have a place in the schools? In: Stearns, P.N., Seixas, P., Wineburg, S. (Eds.), *Knowing, Teaching, and Learning History: National and International Perspectives*. NYU Press, New York, pp. 19–37.
- Shemilt, D., 2000. The caliph's coin: the currency of narrative frameworks in history teaching. In: Stearns, P.N., Seixas, P., Wineburg, S. (Eds.), *Knowing, Teaching, and Learning History: National and International Perspectives*. NYU Press, New York, pp. 83–101.
- Simon, Z.B., 2019. *History in Times of Unprecedented Change: A Theory for the 21st Century*. Bloomsbury Publishing, London.
- Somers, M.R., 1992. Narrativity, narrative identity, and social action: rethinking English working-class formation. *Soc. Sci. Hist.* 16, 591–630.
- Taylor, M.S., 2011. Buffalo hunt: international trade and the virtual extinction of the North American bison. *Am. Econ. Rev.* 101, 3162–3195.
- Thomas, J.A., 2019. Why the “Anthropocene” is not “climate change” and why it matters. *Asia Glob.* 10, 2019. <https://www.asiaglobalonline.hku.hk>.
- Tyler, T., 2021. The exception and the norm: dimensions of Anthropocentrism. In: McHugh, S., McKay, R., Miller, J. (Eds.), *The Palgrave Handbook of Animals and Literature*. Springer International Publishing, Cham, pp. 19–36.
- Van Straaten, D., Wilschut, A., Oostdam, R., 2016. Making history relevant to students by connecting past, present and future: a framework for research. *J. Curric. Stud.* 48, 479–502.
- Wals, A., Corcoran, P.B., 2006. Sustainability as an outcome of transformative learning. In: Holmberg, J., Samuelsson, B.E. (Eds.), *Drivers and Barriers for Implementing Sustainable Development in Higher Education*. UNESCO, Paris, pp. 103–111.
- White, H., 2014. *Metahistory: The Historical Imagination in Nineteenth-Century Europe*. Johns Hopkins University Press, Baltimore.
- Wimmer, A., Schiller, N.G., 2003. Methodological nationalism, the social sciences, and the study of migration: an essay in historical epistemology. *Int. Migr. Rev.* 37 (3), 576–610.
- Wineburg, S., 2013. Undue certainty: where Howard Zinn's. *Am. Educ.* 36, 27–34.
- Wolfe, C., 2010. *What is Posthumanism?* University of Minnesota Press, Minneapolis.
- Woodson, A.N., 2017. “There ain't no white people here” master narratives of the civil rights movement in the stories of urban youth. *Urban Educ.* 52, 316–342.

- Woolf, D., 2019. *A Concise History of History: Global Historiography From Antiquity to the Present*. Cambridge University Press, Cambridge.
- Young, M., Muller, J., 2016. *Curriculum and the Specialization of Knowledge*. Routledge, London.
- Young, M., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45, 101–118.
- Zhang, D.D., Lee, H.F., Wang, C., et al., 2011. The causality analysis of climate change and large-scale human crisis. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17296–17301.
- Zulianello, M., Ceccobelli, D., 2020. Don't call it climate populism: on Greta Thunberg's technocratic ecocentrism. *Polit. Q.* 91, 623–631.

Curriculum and literacy

Csilla Weninger, Nanyang Technological University, National Institute of Education, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	308
Expanding literacy	308
New literacy studies	308
Multiliteracies	309
Global trends in literacy curriculum	310
Education as human capital	310
Literacy curricula in global times	311
Conclusion	312
References	312

Introduction

Literacy has occupied a central position in national curricula since the early days of compulsory schooling, although the curricular form literacy has taken has varied across time and different contexts. On the one hand, the ability to read and write has been argued by governments to be a necessary prerequisite for individuals to function in society; in fact, the UNESCO has since its founding in 1946 maintained that literacy should be seen as a fundamental human right (Wagner, 2011). On the other hand, literacy and its curricular role has also been a highly political and politicized matter; what to teach and how under the label of literacy continues to be used as a tool in the ideological struggle for power and recognition (Blackledge, 2001; Curdt-Christiansen and Weninger, 2015; Luke, 1988, 2004; McKay, 2011). Tensions around literacy have only been made more pronounced due to globalization and the technological developments of the last 20 years, given that these have profoundly altered communication practices and thus have direct implications for literacy curricula.

In this chapter, I discuss curriculum and literacy with reference to two parallel trends that I argue have been affecting literacy curricula across the globe. First, I describe how the field of literacy has seen major shifts in terms of what we understand literacy to be and the implications of this change for school literacy education. In general, this shift is characterized by an expansion and pluralization of the notion of literacy, both in terms of its theoretical treatment in academic scholarship as well as with regards to its practice in everyday communicative encounters. Second, there has been a trend toward the reconfiguration of education in terms of human capital development; in other words, education and schooling is seen as a direct and causal enabler of individuals' earning power (Gillies, 2011) and by extension also as a key driver of national economic progress. This strong economic impetus, coupled with the international educational competition it has facilitated, brought with it a need to standardize literacy both in curriculum and assessment. These two trends, one toward expansion and pluralization, the other toward measurable units and standardization, have led to tensions which ultimately must be managed by those in charge of enacting curricula: teachers and educators.

Expanding literacy

New literacy studies

The notion of literacy as the subject matter of curricula has undergone major shifts in the last several decades. The answer to the question, what does it mean to be literate, has been the focus of intense scholarly discussions whose general direction has been toward pluralization. An important catalyst in this process was theory and research in language and communication that surfaced in the 1970s. On the one hand, functional (as opposed to mentalist) theories of language drew attention to the social motivations for the emergence, organization and development of lexical-grammatical systems (e.g., Fillmore, 1976; Halliday, 1973) while sociolinguists shed light on the social origins of linguistic variation in speech (Labov, 1972). On the other hand, cross-cultural, ethnographic studies of community and home literacy carried out by anthropologists in the 1970s–1980s (Heath, 1983; Street, 1984; Hymes, 1974; Besnier, 1988) revealed fundamental insights about the deeply contextual nature of reading and writing. Shaped by these developments toward a social and cultural understanding of language and communication, literacy came to be recognized as a complex social practice; that is, reading and writing are occasioned by social-communicative needs arising within communities or social groups, shaped by the values, histories and social relations within the group. Forms and technologies of literacy are thus embedded within systems of meaning-making that are fundamentally social and cultural in nature.

This notion of literacy as social practice had significant implications for school literacy curriculum and instruction, particularly within the anglophone world. It directly challenged what New Literacy Studies scholars called autonomous literacy: the idea that reading and writing were discrete, individual cognitive skills detached from their sociocultural context or community of use. As

Street (1995), Gee (1990) argued, such a conceptualization undergirded school curricula and teaching, resulting in “pedagogized literacy” (Street, 1995) which did little to recognize the diverse socio-cultural literate experiences of learners and instead imposed and expected standardized (and supposedly neutral) forms of literacy that served the national interest. School literacy was therefore best described as an ideological practice (Street, 1995) which aimed primarily to foster modes of communication, learning and identities associated with social elites and which reproduced extant social hierarchies.

New Literacy studies scholarship was crucial in highlighting the ideological nature of literacy. In that sense, the pluralization and expansion of literacy was partly a move away from cognitive and individual conceptualization toward a socio-cultural and political one. In addition to the question of what counts as literacy, now the focus was also on whose literacy practices should serve as the model for school literacy. This question has been especially consequential for children whose home/community literacy practices have been traditionally excluded from the school-sanctioned repertoire since those children’s linguistic expertise and identities could not act as resources in the acquisition of literacies that are valued by society. Much subsequent research has therefore explored the home-school literacy connections, seeking ways to highlight children’s funds of knowledge (González et al., 2005) that should be drawn upon in school literacy instruction.

While the issue of class divides and their impact on curriculum, pedagogy and student learning have long been discussed in educational research (e.g., Bernstein, 1990; Erickson, 1987) the question of how out-of-school literacy should inform what happens in school has been formulated and studied from a range of perspectives. For instance, literacy scholars in the 1990s and 2000s were interested in how popular culture as a crucial element of children’s culture and literacy could be harnessed within the school classroom (e.g., Dyson, 1997; Lewis, 1998; Marsh and Millard, 2000; Shegar and Weninger, 2010). Similarly, awareness also grew of the difficulties faced by children from families with languages or language varieties which are vastly different from the standard expected at school. With the global increase in migration in the last two decades, classrooms are becoming more and more diverse linguistically. Thus even in countries where literacy instruction is in a single national or majority language, recognizing children’s multilingual repertoires as a resource has been an important argument for enabling them to gain competence in the language and literacy practices valued by schools (Gort and Sembiante, 2015; Hornberger and Link, 2012). Most recently, the issue of divide has been studied from the perspective of differences between technologically mediated forms of literacy youth engage in for leisure and the lack of inclusion of such literacies in school instruction (Alvermann, 2012).

These developments have had obvious implications for curriculum. The new conceptualization situated literacy within histories and traditions of textual meaning-making tied to social groups and communities. It also recognized values, emotions and identity as implicated in the socio-cultural practice of reading and writing. Further, literacy as ideological practice highlighted how social valuations attached to different forms and norms of reading and writing are intertwined with the relative power various groups have within the social hierarchy, so that learning to read and write in school becomes exclusionary or marginalizing toward those whose non-school textual practices are not reflected in pedagogized literacy. In light of this, curriculum should incorporate and account for students’ home and out-of-school literacy practices as much as possible (Alvermann and Hinchman, 2012). Researchers have also argued for the need to expand the school literary canon to incorporate a wider variety of texts, both in terms of genre (e.g., those from pop culture such as comic books, or fantasy novels) as well as cultural or ethno-linguistic origins of authors (e.g., postcolonial works or migrant literature).

Additionally, the school subject of literacy had to be broadened in order to successfully account for the fact that reading and writing are not simply about encoding and decoding texts; rather, literacy is equally about understanding how texts communicate values and identities through the ways in which they represent the world. Starting in the early 2000s, strong arguments have been made to expand subject literacy to encompass critical and media literacy (Alvermann and Hagood, 2000; Hart and Hicks, 2002) as well as new literacies of the digital era (Lankshear and Knobel, 2003, 2011). The expansion of the school canon through diverse popular cultural, digital and media texts necessitated the teaching and learning of new competencies in addition to what has been termed “traditional literacy”. Media and critical literacy as part of the literacy curriculum should foster the critical reading and evaluation of (new) media texts, couched within an examination of how social context shapes the production of texts (literacy as social practice) as well as providing opportunities for students to engage in text production to explore questions of identity, representation and power (Buckingham, 2003; Hobbs, 2007; Kellner and Share, 2005). These developments were precipitated to a great extent by the publication of a paper that has come to be one of the most significant milestones in literacy theory of the past 30 years.

Multiliteracies

Scholarly arguments for the relevance of popular culture for school literacy was in a way part of much broader developments that were taking place in the 1990s and which had direct consequences for literacy curriculum and instruction. New communication technologies such as mobile phones and email, were becoming widespread thanks to advancements in digital technology. The internet, text-heavy as it was in the 1990s, enabled a range of new multimodal genres and ways for people to connect with others on a scale and at a speed not previously seen, although technological adoption was extremely uneven in those early years. In parallel, conditions of work and life were changing due to a decline in mass manufacturing (in the global North) and the dominance of knowledge-based industries. These changes created new conditions of work and employment as well as new competencies from those seeking employment as knowledge workers, such as communications skills and collaboration (OECD, 2001). As Gee and Lankshear (1995) argued in a more critical vein, a whole new discourse around the new work order and its requisite skills emerged, which was then taken up by educational establishments as rationale to realign curricula in the service of the global capitalist economy.

While literacy scholars were broadly critical of global capitalism and its commodification of communication (e.g., Fairclough, 1999), there was agreement that shifting communication practices had to be accounted for in the theoretical conceptualization of literacy as well as in educational practice. This was precisely the premise of the New London Group (1996) publication titled *A pedagogy of multiliteracies: Designing social futures*. It presented a succinct overview of social and technological changes of the late 20th century and argued for the need to redefine the what as well as the how of literacy education. The group's proposal was first of all motivated by a vision of education where learning benefitted all students by enabling them to fully participate in social, civic and economic activities. To facilitate this access, literacy education had to reflect the diversity and pluralism that characterized literacy practices outside of school:

First, we want to extend the idea and scope of literacy pedagogy to account for the content of our culturally and linguistically diverse and increasingly globalized societies, for the multifarious cultures that interrelate and the plurality of texts that circulate. Second, we argue that literacy pedagogy now must account for the burgeoning variety of text forms associated with information and multimedia technologies.

New London Group (1996, p. 61).

The group ardently advocated the need to de-prioritize language, especially in its schooled, prescriptive and monolingual form, and think of literacy as competence over the myriad modes of representation (visual, spatial, gestural, etc.) available to people to communicate. Central to their arguments about the “what” of literacy was the notion of design. Literacy as social-semiotic practice, according to the New London Group, had to be reconceptualized to highlight its dynamic aspect and transformative potential. On the one hand, we must acknowledge that the tools at our disposal for meaning-making (such as the grammar of language, or conventions of visual communication) are to some extent given to us—what the group called available design. Yet semiotic activity is never simply about strict adherence to conventions; there is always the possibility for agentive deviation or creative deployment of available designs through acts of literacy. This process is termed Designing, again to highlight the dynamic character of reading, writing, speaking and listening, as enabling Redesign: transformed meanings (representations, identities) which themselves become new available designs.

Literacy curriculum as envisioned by the New London Group was then “a design for social futures” (p. 73), prompted by fundamental changes in the social and economic organization of the world, and the kind of social subject schools needed to educate. Instead of hierarchical authority relations, stable knowledge and learning-as-transmission that characterized institutional education in the 20th century, 21st century curriculum must envision learners as agentive, and learning as transformative. As Kress (2000) put it, “Design asks for production of the new rather than replication of the old. Thus putting ‘design’ at the center of the curriculum and of its purposes is to redefine the goal of education as the making of individual dispositions oriented toward innovation, creativity, transformation and change” (p. 141).

Global trends in literacy curriculum

Education as human capital

Literacy scholars, recognizing the fundamental shifts that were taking place in the late 20th century, had a plan for literacy education: to pluralize the notion of literacy and to shift school curricula to embrace the creative possibilities of the diversity that globalization and a new digital communicative landscape brought about. As many of them recognized around the turn of the century, the global capitalist economy rather than national states would shape the direction of education and thus school curricula (Gee and Lankshear, 1995; Kress, 2000).

But in fact, the largely economic reshaping of (public) education began significantly earlier, with the advent and increasing influence of Human Capital Theory (HCT), originally formulated in the 1960s (Becker, 1962; Schultz, 1960). HCT posits a causal link between an individual's level of education and their earning power so that according to the model, more educational credentials will lead to higher levels of income (Gillies, 2011). Moreover, this direct and causal relationship can be scaled up such that a well-educated populace will also translate into higher economic gains at the national level. This fundamentally economic argument for education has been extremely influential for education policy globally (Fitzsimons, 2015). Specifically, it has facilitated the marketization of education in multiple ways. First, once the individual became a form of mobile capital, schooling became a process of capital accumulation; the acquisition of skills, competencies and attributes that are thought to add and increase value. Second, framing education in economic terms gave it an instrumental, transactional valuation: one where cost-benefit analysis will determine which schools one will choose (for their child), which courses, subjects, majors one will enroll in, which languages one will learn. Other motivations for pursuing education (passion, interest, talent) become secondary. Third, entire industries have sprung up around aspects of human capital thought to be particularly valuable on the global market: commercial literacy packages, STEAM enrichment programs, English language schools, coding camps, often offered at very young ages so that kids can “get a head-start”. Fourth, competition, a central driving force of market capitalism, has also become a central characteristic of education both at the individual as well as the national level. League tables and global rankings now sort schools, universities and entire countries according to how well they score on certain standardized measures, giving us supposed objective assessment of the country's human capital potential or the best schools or universities to attend for maximum economic return.

HCT has precipitated the emergence of the educational market over the last few decades where students as customers come to acquire commodified educational goods (skills, competencies, dispositions) in institutions of learning that operate under a management model that is based on accountability to customers, “stakeholders” and the government, who can withdraw funds for under-performance (Ball, 2003; Hursh, 2005). Measurable outcomes and measurable progress are key components of this model because they constitute the basis for comparison across schools, districts and countries. In this regard, international tests of educational achievement such as the OECD’s Program for International Student Assessment (PISA), the International Association for the Evaluation of Educational Achievement’s TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading and Literacy Study) have been particularly influential not only in facilitating global comparisons but also in directly shaping national educational policy and the direction of education reform (Lewis and Lingard, 2015; 2016 Special Issue of the European Educational Research Journal).

The widespread acceptance of HCT as a blueprint for education policy making had a significant role in reshaping education as a global business enterprise for the creation of human capital to feed the knowledge economy. This process has been largely overseen and managed not by national governments but by an emerging global network of supranational and non-governmental organizations (Ball, 2012; Rizvi and Lingard, 2010). One important impact of these global trends on national policy and curriculum has been toward homogenizing educational purpose, content and quality, irrespective of context or national histories (Zhao, 2015, 2020). Twenty-first century education espoused by global policy-making is a decidedly “employee-oriented education” (Zhao, 2012) in its purpose, comprised of sets of skills and competencies thought to be universally applicable and needed by a 21st century workforce and measurable through standardized tests. What does such a view of education then mean in terms of literacy and curriculum?

Literacy curricula in global times

In the first part of this chapter, I argued that a crucial point for 21st century literacy education from a scholarly standpoint was the need to empower learners to design new social futures. The model of education described in the preceding paragraphs however is in stark contrast with this vision. Indeed, scholars have critically examined the impact of the global trends sketched above on literacy curriculum and instruction and the results are not always encouraging, pointing toward processes of decontextualization, standardization and fragmentation in many cases.

First, and in line with Kress’ (2000) call for a curriculum of communication instead of language or even literacy, communication has indeed been afforded a central place in curriculum debates. In fact, it is frequently cited as one of a number of “C”s, along with collaboration, creativity and critical thinking that make up crucial skills for the 21st century (Battelle for Kids, 2019). However, communication as a curricular component of 21st century education has generally been stripped of its complexity and reduced to the notion of communication skills (Shin, 2016; Weninger, 2019): a set of desirable competencies that are thought to be context-independent, such as how to give a successful presentation or how to construct arguments. Often, communication curricula amount to little more than mastery of the genres of work (Weninger and Kan, 2013), which while certainly valuable, falls far short of lofty goals such as innovation and creativity. Quite the contrary; when we reduce communication to a bundle of decontextualized skills, we are erasing the very contingency and contextual situatedness which makes communication an ambivalent process, always open to possibilities (Fairclough, 1999).

Another important aspect of curriculum development has been the adoption of standards-based frameworks for literacy. In the United States, the Common Core State Standards Initiative was jointly developed by the National Governors Association and the Council of Chief State School Officers and has been adopted by the majority of US states. The Standards, with the tagline Preparing America’s Students for Success, represent core knowledge and skills that all high school graduates in the United States need for college and career-readiness in the two main areas of math and literacy. Similar efforts in the UK and Australia have led, over the past 20 years, to the formulation of literacy benchmarks; essentially standards of achievement formulated as part of a national literacy curriculum. The Council of Europe’s Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001; and the 2020 Companion Volume update) was specifically aimed to standardize educational administration concerning the purpose, content, methods and professional development of teaching and learning languages across Europe. The framework identifies detailed descriptors for various levels of language proficiency in areas of receptive, productive, pragmatic and sociolinguistic competence and has been increasingly adopted in regions and countries outside of Europe, such as Asia and Latin America (Read, 2019).

Literacy curricula today mostly take the form of national and supranational frameworks of standards and competences. Scholars have critiqued individual frameworks on various grounds; but a broader critique of standards-based literacy curricula centers on the observation that they are inevitably accompanied by assessment regimes which have serious washback effects on teaching and learning (Comber, 2012; Hardy, 2015). For instance, despite the CEFR being driven by a “vision of the user/learner as a social agent, co-constructing meaning in interaction, and by the notions of mediation and plurilingual/pluricultural competences” (Council of Europe, 2020, p. 21), the framework is mostly known for its proficiency scales which are used as the basis for assessment and developing national tests (Read, 2019). Standards-based curricula thus must be seen in the broader transformations of language education as a global affair; what is viewed as important knowledge and skill is informed by the needs of global capitalism, gets vernacularized (Rizvi and Lingard, 2010; Weninger, 2017) as local curriculum for the creation of human capital, and is in turn monitored and measured through accountability systems such as standardized tests. At every step of the way, there is a host of educational products being created for an educational marketplace for anxious parents and lifelong learning in the gig-economy.

The implications of these trends for literacy as a curricular subject are quite clear and well-documented in multiple contexts. Despite scholarly arguments and research evidence for the increasing pluralization of literacy as outlined earlier in this chapter, it seems that we are reverting back to the idea of literacy as a measurable asset of individual learners. Given the interlocking of outcome and standards-based literacy curricula with standardized testing, it is no surprise that “as the content of the curriculum moves to match what the tests require, the structure of curricular content knowledge similarly shifts toward the fragmentation demanded by the tests” (Au, 2011, p. 30). Curriculum is always the outcome of a process of selection (Apple, 1979) based on ideological principles. And a literacy curriculum influenced heavily by the need to measure will give weight to those aspects of literacy which are amenable to measurement and have a history of being measured. This means a strong emphasis on (select aspects of) language—as printed, monolingual text—at the expense of multiliteracies.

Perhaps even more subtly, intended curricula are frequently based on progressive ideas about the need for multiliteracies and an emphasis on effective, creative communication. This is certainly what we see in the globally influential frameworks that often incorporate critical thinking, media and research skills as part of literacy. However, by emphasizing results rather than the means (as the Common Core State Standards for English Language Arts explicitly does, for example), we are essentially putting the cart before the horse. In other words, when standardized high-stakes examinations are the primary measures of learning, the enacted curriculum will likely be shaped by what is on the test rather than the philosophical ideas behind the curriculum.

Conclusion

New literacy scholars introduced the expression literacy as social practice in opposition to what they saw as the prevailing cognitive, individualistic view that undergirded school literacy curriculum and instruction. They critiqued this pedagogized literacy as an autonomous or neutral technology that was routinely utilized as an “independent variable” (Street, 1984), measurable on its own as well as how it correlated with or resulted in various societal outcomes. Yet new literacy scholars were by no means the first ones to bemoan the negative effects of psychometric approaches to literacy and the overstatement of literacy’s socially and morally redemptive powers (cf. Graff, 1979). As early as 1921, writers of the Newbolt Report, which was instrumental in identifying the place of English as curricular subject in Great Britain, stated:

Now, as then, there is the danger that a true instinct for humanism may be smothered by the demand for definite measurable results, especially the passing of examinations in a variety of subjects, and if those who are anxious to do justice to English find it so hard to carry out their desire, what is to be expected from those who will remain indifferent?

The Newbolt Report (Board of Education, 1921, p. 55f).

It seems that this 100-year-old concern of those “anxious to do justice” to literacy is unlikely to go away or be resolved anytime soon. The creative potential of literacy as a tool for young people to design social futures, as the New London Group implored, lies in its messy, situated, deeply contextual and pluralistic manifestations. School curricula and instruction that is driven by the need to foster literacy as homogeneous, decontextualized and measurable skills is not going to lead to the transformation that the notion of Design encapsulates, even if influenced by noble ideas about global communication. Such literacy curricula are fundamentally system-preserving, effecting the replication of the old and thus the reproduction of the status quo.

References

- Alvermann, D.E., Hagood, M.C., 2000. Critical media literacy: research, theory, and practice in “New Times”. *J. Educ. Res.* 93, 193–205.
- Alvermann, D.E., Hinchman, K.A., 2012. *Reconceptualizing the Literacies in Adolescents’ Lives: Bridging the Everyday/Academic Divide*. Routledge, New York.
- Alvermann, D.E. (Ed.), 2012. *Adolescents’ Online Literacies: Connecting Classrooms, Digital Media and Popular Culture*. Peter Lang, New York.
- Apple, M., 1979. *Ideology and curriculum*. Routledge & Kegan Paul, London.
- Au, W., 2011. Teaching under the new Taylorism: high-stakes testing and the standardization of the 21st century curriculum. *J. Curric. Stud.* 43 (1), 25–45.
- Ball, S.J., 2003. The teacher’s soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Ball, S.J., 2012. *Global Education Inc.: New Policy Networks and the Neo-liberal Imaginary*. Routledge, New York.
- Battelle for Kids, 2019. Framework for 21st Century Learning. <https://www.battelleforkids.org/networks/p21/frameworks-resources>.
- Becker, G., 1962. Investment in human capital: a theoretical analysis. *J. Polit. Econ.* 70 (5), 9–49.
- Bernstein, B., 1990. *The Structuring of Pedagogic Discourse. Volume IV of Class, Codes and Control*. Routledge, London/New York.
- Besnier, N., 1988. The linguistic relationship of spoken and written Nukulaelae registers. *Language* 64 (4), 707–736.
- Blackledge, A., 2001. Literacy, schooling and ideology in a multilingual state. *Curric. J.* 12 (3), 291–312.
- Board of Education, 1921. *The teaching of English in England (The Newbolt Report)*. His Majesty’s Stationery Office, London. Available from: <http://www.educationengland.org.uk/documents/newbolt/newbolt1921.html#04>.
- Buckingham, D., 2003. *Media Education: Literacy, Learning and Contemporary Culture*. Polity, Cambridge, UK.
- Comber, B., 2012. Mandated literacy assessment and the reorganization of teachers’ work: federal policy, local effects. *Crit. Stud. Educ.* 53 (2), 119–136.
- Common Core State Standards Initiative, (n.d.). Introduction: Key Design Consideration for English Language Arts. Retrieved from <http://www.corestandards.org/ELA-Literacy/introduction/key-design-consideration/>.
- Council of Europe, 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe Publishing, Strasbourg.

- Council of Europe, 2020. Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion Volume. Council of Europe Publishing, Strasbourg.
- Curdtt-Christiansen, X.L., Weninger, C., 2015. Introduction. In: Curdtt-Christiansen, X.L., Weninger, C. (Eds.), *Language, Ideology and Education: The Politics of Textbooks in Language Education*. Routledge, London/New York, pp. 1–8.
- Dyson, A.H., 1997. *Writing Superheroes: Contemporary Childhood, Popular Culture, and Classroom Literacy*. Teachers College Press, New York.
- Erickson, F., 1987. Transformation and school success: the politics and culture of school success. *Anthropol. Educ. Q.* 18, 335–356.
- Fairclough, N., 1999. Global capitalism and critical awareness of language. *Lang. Aware.* 8 (2), 71–83.
- Fillmore, C.J., 1976. Frame semantics and the nature of language. *Ann. N. Y. Acad. Sci.* 280 (1), 20–32.
- Gee, J.P., Lankshear, C., 1995. The new work order: critical language awareness and “fast capitalism” texts. *Discourse* 16 (1), 5–19.
- Fitzsimons, P., 2015. Human capital theory and education. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore, pp. 1050–1057.
- Gee, J.P., 1990. *Social Linguistics and Literacies: Ideology in Discourses*. Falmer, London.
- Gillies, D., 2011. State education as high-yield investment: human capital theory in European policy discourse. *J. Pedagog.* 2 (2), 224–245.
- González, N., Moll, L.C., Amanti, C. (Eds.), 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum, Mahwah, N.J.
- Gort, M., Sembianti, S.F., 2015. Navigating hybridized language learning spaces through translanguaging pedagogy: dual language preschool teachers’ languaging practices in support of emergent bilingual children’s performance of academic discourse. *Int. Multidiscip. Res. J.* 9 (1), 7–25.
- Graf, H.J., 1979. *The Literacy Myth: Literacy and Social Structure in the Nineteenth-Century City*. Academic Press, New York.
- Halliday, M.A.K., 1973. *Explorations in the Functions of Language*. Edward Arnold, London.
- Hardy, I., 2015. A logic of enumeration: the nature and effects of national literacy and numeracy testing in Australia. *J. Educ. Pol.* 30 (3), 335–362.
- Hart, A., Hicks, A., 2002. *Teaching Media in the English Curriculum*. Trentham Books, London.
- Heath, S.B., 1983. *Ways With Words*. Cambridge University Press, Cambridge.
- Hobbs, R., 2007. *Reading the Media: Media Literacy in High School English*. Teachers College Press, New York.
- Hornberger, N.H., Link, H., 2012. Translanguaging and transnational literacies in multilingual classrooms: a biliteracy lens. *Int. J. Biling. Educ. Biling.* 15 (3), 261–278.
- Hursh, D., 2005. Neo-liberalism, markets and accountability: transforming education and undermining democracy in the United States and England. *Pol. Futures Educ. Internet* 3 (1), 3–15.
- Hymes, D., 1974. *Foundations in Sociolinguistics: An Ethnographic Approach*. University of Pennsylvania Press, Philadelphia.
- Kellner, D., Share, J., 2005. Toward critical media literacy: core concepts, debates, organization and policy. *Discourse* 26 (3), 369–386.
- Kress, G., 2000. A curriculum for the future. *Camb. J. Educ.* 30 (1), 133–145.
- Labov, W., 1972. *Sociolinguistic Patterns*. University of Philadelphia Press, Philadelphia.
- Lankshear, C., Knobel, M., 2003. *New Literacies: Changing Knowledge and Classroom Learning*. Open University Press, Maidenhead.
- Lankshear, C., Knobel, M., 2011. *New Literacies: Everyday Practices and Social Learning*, third ed. Open University Press, Maidenhead.
- Lewis, S., Lingard, B., 2015. The multiple effects of international large-scale assessment on education policy and research. *Discourse* 36 (5), 621–637.
- Lewis, C., 1998. Rock n’ roll and horror stories. *J. Adolesc. Adult Lit.* 42, 116–120.
- Luke, A., 1988. *Literacy, Textbooks and Ideology: Postwar Literacy Instruction and the Mythology of Dick and Jayne*. Taylor & Francis, Bristol, PA.
- Luke, A., 2004. Literacy and educational fundamentalism: an interview. *Engl. Q.* 36 (4), 12–17.
- Marsh, J., Millard, E., 2000. *Literacy and Popular Culture: Using Children’s Culture in the Classroom*. Sage, London.
- McKay, G.R., 2011. Policy and indigenous languages in Australia. *Aust. Rev. Appl. Ling.* 34 (3), 297–319.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- OECD, 2001. Competencies for the knowledge economy. Available from: <https://www.oecd.org/innovation/research/1842070.pdf>.
- Read, J., 2019. The influence of the Common European Framework of Reference (CEFR) in the Asia-Pacific Region. *Learn. J.* 12 (1), 12–18.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Schultz, T., 1960. Capital formation by education. *J. Polit. Econ.* 68 (6), 571–583.
- Shegar, C., Weninger, C., 2010. Intertextuality in preschoolers’ engagement with popular culture: implications for literacy development. *Lang. Educ.* 24 (5), 1–17.
- Shin, H., 2016. Language “skills” and the neoliberal English education industry. *J. Multiling. Multicult. Dev.* 37 (5), 509–522.
- Street, B.V., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Street, B.V., 1995. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Longman, London/New York.
- Wagner, D.A., 2011. What happened to literacy? Historical and conceptual perspectives on literacy in UNESCO. *Int. J. Educ. Dev.* 31, 319–323.
- Weninger, C., Kan, K.H.-Y., 2013. (Critical) language awareness in business communication. *Engl. Specif. Purp.* 32 (2), 59–71.
- Weninger, C., 2017. The “vernacularization” of global education policy: media and digital literacy as 21st century skills in Singapore. *Asia Pac. J. Educ.* 37 (4), 500–516.
- Weninger, C., 2019. *From Language Skills to Literacy: Broadening the Scope of English Language Education Through Media Literacy*. Routledge, Abingdon.
- Zhao, Y., 2012. *World Class Learners: Educating Creative and Entrepreneurial Students*. Corwin, Thousand Oaks.
- Zhao, Y., 2015. Globalization in education. In: Wright, J.D. (Ed.), *International Encyclopedia of the Social & Behavioral Sciences*, second ed. Elsevier, Amsterdam, pp. 247–253.
- Zhao, Y., 2020. Two decades of havoc: a synthesis of criticism against PISA. *J. Educ. Change* 21, 245–266.

Curriculum and science

Keith S. Taber, Faculty of Education, University of Cambridge, Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	314
Three phases of constructing a science curriculum	315
Demarcating science	315
Selecting science to teach	316
The science curriculum and educational purpose	316
Four frames for understanding the purposes of the science curriculum	317
The economic driver for science curriculum design	317
The developmental driver for science curriculum design	317
The cultural driver for science curriculum design	318
The civic driver for science curriculum design	319
Organization of the science curriculum	320
Disciplinary identify	320
A common core to the sciences	320
The STEM question	321
Differentiating the curriculum	322
Curriculum progression	323
Developing curriculum models to represent science	323
Learning progressions	325
References	325

Glossary

- Bildung** The process of formation, as in the formation of the person's character through suitably educative experience
- Canonical concept** A conceptualization that has broad currency in a disciplinary field, and so provides a useful referent for representing as target knowledge in the curriculum
- Confirmation bias** The tendency to pay more attention to data that provides supporting evidence for a conjecture or belief
- Control of variables** The technique whereby the effect of changing one variable (referred to as the independent variable) upon another variable of interest (referred to as the dependent variable) is tested by keeping all other potentially conflating variables constant (i.e., controlled)
- Curricular inertia** Where the merits of a well-established component of the curriculum come to be taken for granted
- Curricular models** The versions of target knowledge that appear in curriculum specifications and which are generally simplifications of the disciplinary knowledge they represent
- Curriculum subject** An identified component of the school curriculum (e.g., science, physics), usually assigned designated slots in the timetable
- Didactic transposition** The process by which disciplinary knowledge is inevitably imperfectly represented through the educational processes of curriculum development and teaching
- Discipline** In part this refers to an area of knowledge, but also the associated assumptions about how knowledge in that discipline is judged and developed, and the methodological techniques and standards, as well as the norms concerning such matters as how findings are communicated and are evaluated within a community of practitioners
- Learning progressions** Research-informed models to inform curriculum design, providing a framework encompassing learning trajectories in a concept area that are judged as viable for learners working over a defined range of grades
- Project-based learning** Curricular activities defined for the learners in terms of a substantive project that groups of students work on over an extended period

Introduction

Science has become established globally as a core aspect of the school curriculum, as is evidenced by the perceived significance of such international comparisons the Trends in International Mathematics and Science Study (Mullis et al., 2019). This article will consider the nature of science in the school curriculum focusing in particular on the theme of what should be included under school

science curriculum in relation to (i) the purposes of science education for all school age learners, and questions of (ii) how science itself is understood, and (iii) how science can be organized within a teaching curriculum (including in relation to “STEM”—science, technology, engineering and mathematics—as a curriculum area).

Three phases of constructing a science curriculum

The construction of a curriculum around any subject area (to form a timetabled curriculum subject) involves the specification of subject matter to be included, which encompasses three sets of considerations (see Fig. 1). “Subject matter” may refer to knowledge (topics, concepts, etc.), but could also mean skills (e.g., accurate weighing with a chemical balance) or attitudes (e.g., to value objectivity in enquiry). School curriculum subjects may be constructed that draw upon knowledge from across formal disciplinary boundaries (e.g., “personal, social and health education,” citizenship education) which offers curriculum developers considerable freedom to define the subject. However, commonly the school curriculum is largely framed in terms of subjects that nominally reflect academic disciplines or groups of disciplines—such as the natural sciences.

Firstly, it is necessary to identify *the scope* of potential material that might be included as falling within the subject area. Here, we might ask: what counts as *science*? (Or, what counts as *biology*, or *astronomy*, and so forth.) Then, there is a question of *selection*. Given that there is likely to be more potential material than can reasonably be included in a school curriculum, decisions have to be made about what to include, and what to omit. That is, which science is it most valuable to teach? That is a judgment that can only be made in the light of a position (whether explicit or otherwise) on the purpose (or purposes) of science education—and may lead to different answers for different groups of learners. However, once a selection has been made from among the potential range of disciplinary material, there is also a further consideration as disciplinary knowledge is often not immediately accessible to school age learners. There is the issue, then, of how the disciplinary content will be *represented* in the curriculum.

Demarcating science

It is tempting to assume that the distinction between science and non-science may be taken for granted, as it may seem that what counts as science is widely recognized. The sociologist Basil Bernstein (1971/2003) suggested that curriculum areas differ in terms of how permeable their boundaries are. In the Anglophile tradition in which he worked, science, with its well-established disciplines and associated learned societies and professional bodies, might be characterized to be a relatively “strongly classified” subject area well “insulated” from other parts of the curriculum, and traditionally a “strongly framed” curriculum area where there has been limited flexibility in pedagogy. It generally seems clear what does, and what does not, fall within its scope—and often what counts as “teaching science.”

However, in some national contexts, reference to “science” in the curriculum might be assumed to include mathematics as a component part of science, whereas these are clearly understood as distinct school subjects elsewhere (see also Fig. 3). Content can be found included in science curricula in some national contexts that would more commonly be considered to fit elsewhere in the school curriculum, such as under geography (Taber and Vong, 2020). This reflects how even in Universities today disciplines such as psychology and geography may be considered to fall partially under the natural sciences, and partially outside.

Despite its ubiquity in modern culture, science in its contemporary sense is a relatively recent invention. The term “scientist” was only introduced in the 19th century. The word science reflects the Latin term “scientia” which refers to knowledge in general. In German today the term Wissenschaft has a much broader meaning than its *nominal* equivalent, “science,” in English. So, whereas, generally, English speakers see “science” as synonymous with the natural sciences (and therefore *not* including the social sciences), the natural sciences would usually be *specified* as Naturwissenschaft in German.

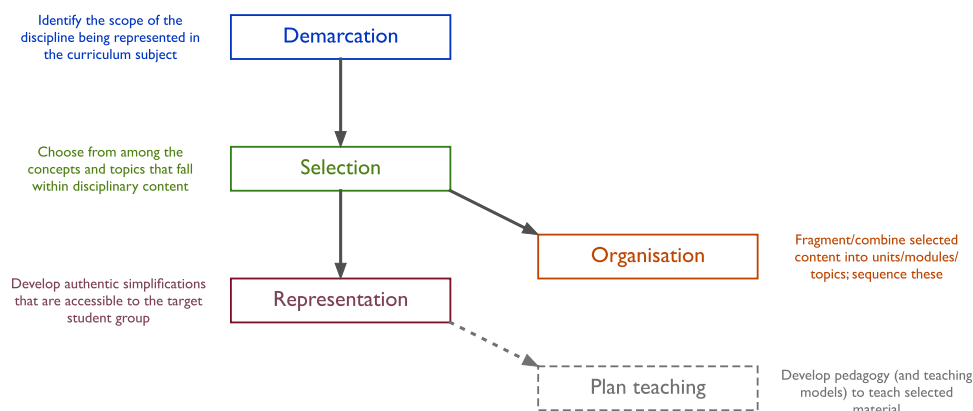


Fig. 1 Constructing the curriculum. From <https://science-education-research.com/teaching-science/the-science-curriculum/>.

The medieval universities of Europe tended to require a common core for all students (before they could progress to the advanced faculties such as law, theology and medicine): grammar, rhetoric, and logic; followed by arithmetic, astronomy, geometry, and music. Although we might recognize astronomy as a science in contemporary terms, it was then a subject largely concerned with practical issues such as determining the calendar, and the modern distinction between astronomy and astrology had not become so clear.

Modern science drew upon pre-existing traditions such as that of natural philosophy and alchemy, and indeed the notion of natural theology (in Christianity, but reflecting a view also widely adopted in Islam) where the study of the natural world was expected to complement and buttress the study of religious scriptures (Grumett, 2009). It is probably not possible to make a clear distinction between science and its historical precursors, nor indeed to offer an account of the precise bounds of contemporary science that would attract universal agreement.

Selecting science to teach

During the 19th century, science subjects such as physics, chemistry, and geology, the latter then having a much higher public profile than today (Bowker, 2005), began to appear in the school curriculum for those privileged to be offered more than an elementary education. During the twentieth century there were debates in various contexts about the amount and form of science education suitable for all or some groups of learners (based on ability judgments, age, or even gender). The sheer extent of knowledge in the natural sciences also grew enormously, making decisions about the selection of content more challenging. Material that has become established in a curriculum may acquire a kind of *curricular inertia*: teachers will know the topic well, effective pedagogy and learning resources exist, a bank of past assessment items considered to test understanding or skill level is available, and so on, and its presence as part of school science may become taken for granted.

Introducing “new” science (which may be seen as having currency in public discourse as well as greater technical, commercial or social relevance) may seem indicated. For example, in the 20th century relativity and quantum theory brought about fundamental shifts in physics, and in terms of our best current understanding of the universe. These topics would seem to have a strong claim to being introduced to at least more advanced school students in order to reflect and represent modern physics. Yet these are considered especially difficult topics due to their counter-intuitive nature and high mathematical content. It might in part be countered that research physics today is *generally* primarily conceptualized and communicated in highly abstract mathematical terms that are necessarily downplayed in school physics, and that the *descriptive*, qualitative nature of the treatment in school science of topics such as heating, electricity, fluids, and light reflects how scientific knowledge must be distorted in order to be accessible to school-level learners when it is *represented* in curriculum accounts.

Other examples of more recent science would be advances in reproductive biology (such as *in vitro* fertilization) and in developing models of climate change. These are less fundamental new areas of science (*cf.* relativity and quantum theory), but with great relevance to people’s lives.

However, introducing new topic areas requires either substituting for some existing curriculum content that will be omitted, or increasing the amount of material to be addressed by teachers and learners. Even when new science seems to *update* existing topics this may not be straightforward. Newtonian mechanics, though now considered displaced by relativity as a theory of the world, is still widely applied. It is said that the Apollo missions that took people to the moon and back managed with calculations based on Newtonian mechanics even though that is now strictly considered an incorrect theory.

Where new material is added to the curriculum without making omissions, the science curriculum becomes increasingly packed with content. Then the time typically available for developing any particular topic is reduced. This makes science more challenging for students who are introduced to an extensive carousel of different topics, each for relatively short periods before moving on (Cerini et al., 2003). This, in turn, is considered to encourage a focus on memorization and recall of key points rather than deep understanding of concepts and engaging higher level skills. This can also be demotivating for the highest achievers (sometimes labeled “gifted” learners) who find little to challenge their thinking and engage them in depth.

This issue of the science curriculum becoming extremely broad but shallow may be avoided if the curriculum is designed from a perspective that puts less emphasis on specific conceptual content and more on other foci (such as skills development, values and attitudes), which aligns with some views of the primary purpose of education.

The science curriculum and educational purpose

Considerations of curriculum choice should be related to the aims of education, as curriculum can be seen as the means of responding to such aims (see Fig. 2). When knowledge was largely viewed as something fixed and transferrable (consider the central place of Aristotle’s corpus in University education and the place of the classics—that is the Latin and Greek classics—in grammar school education over some centuries), education could be seen as the means by which knowledge was passed on, from teacher to pupils, some of whom would later become teachers and pass the same knowledge on themselves. Arguably, such a perspective remained influential into the start of the twentieth century—even in elite institutions (Perry, 1985). Today it is more common in education to reject the notion of knowledge as the kind of entity which can be transferred (or perhaps more accurately copied) from one mind to another, and to see knowledge as personally constructed, albeit usually socially mediated (Glaserfeld, 1989), and being contingent

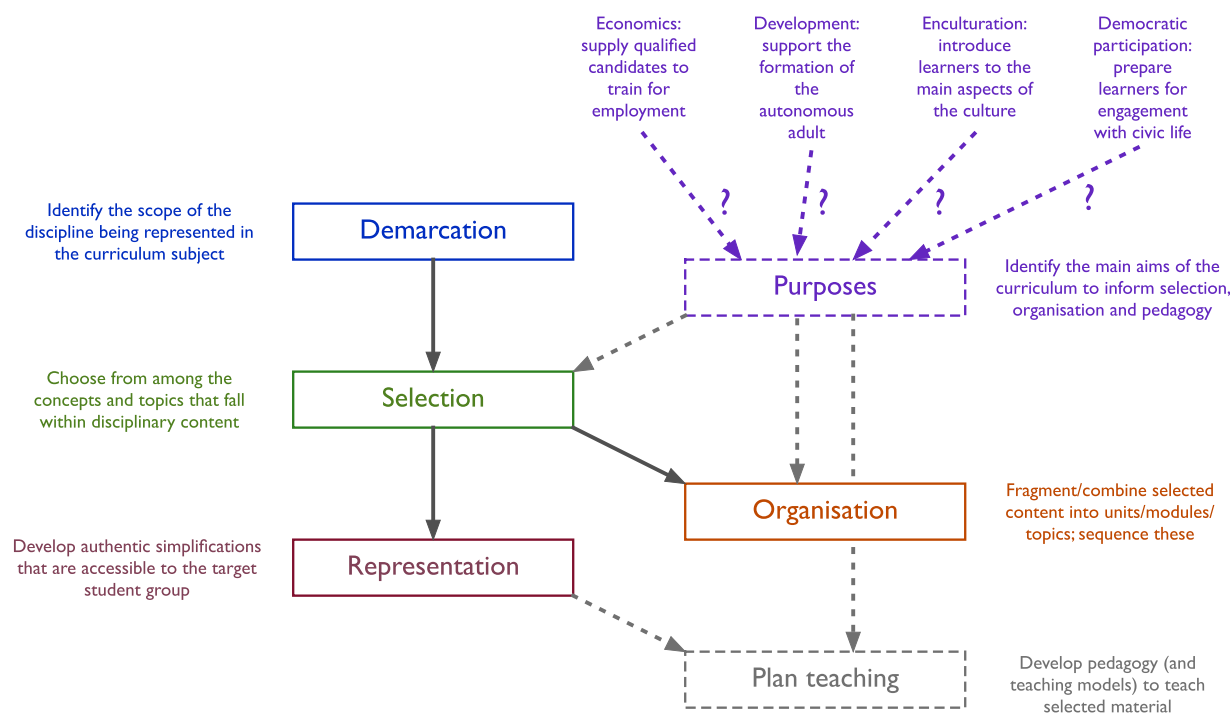


Fig. 2 Selection of curriculum content should reflect the perceived purposes of the educational program. From <https://science-education-research.com/teaching-science/aims-of-the-curriculum/>.

on the interpretive resources (prior knowledge, intuitions, experiences, sources for understanding similes and analogies, imagery, etc.) available to the learner, as well as the teacher's skill in "making the unfamiliar familiar."

A curriculum specification should then reflect the question "*what are we trying to achieve through this curriculum?*", which will inform not only the *selection* of material but also its *organization*—for example, organizing units according to disciplinary topics, or alternatively, through social issues or everyday contexts (Bennett et al., 2003)—as well as how teaching is organized. It is possible to suggest a number of complementary "drivers," such as the four suggested below, that might motivate different choices of (or priorities for) curriculum content and organization.

Four frames for understanding the purposes of the science curriculum

There are a number of perspectives informing the aims of schooling today. These can be considered complementary as all have merit, but where they indicate different kinds of curriculum specification they also compete for limited resources (for "space" in the curriculum).

The economic driver for science curriculum design

One perspective can be associated with the view that one important role of schools is to prepare young people to take up economically productive roles in society. Science education contributes to the so-called "STEM pipeline" (Cannady et al., 2014) that ultimately provides the scientists, engineers, school science teachers, technicians, medical professionals, technologists and so forth that society needs. Whether education at school level should have an *explicit* vocational flavor is a matter of contention, but schools are expected to provide sufficient candidates prepared to enter further education and training in the sciences.

While this argument has much force, not everyone wishes to enter a scientific or science-based career (and not all of those who seek to do so will be selected). It has been widely recognized for some decades now that designing a science curriculum *primarily* for the minority who will progress in this way is inappropriate (Fensham, 1985): that is, science education that is imposed on all through compulsory schooling should not be designed so that most either fail or are likely to disengage because they perceive it as aligned with vocational aspirations they do not share (Eilks and Hofstein, 2015).

The developmental driver for science curriculum design

Alternative perspectives consider that school science should offer clearer, direct relevance for *all* students (Millar and Osborne, 1998). One such viewpoint is that education is about supporting the formation of the learner to develop into a rounded, self-

confident and balanced person, and this can be supported by a liberal education that offers a wide range of experiences and challenges that can both help the individual develop all aspects of her personality as well as offering a taster of different domains to help her make informed career choices. Through this process a person develops as a unique individual and finds their path in life, something that can be seen as reflecting the German tradition of *Bildung* or formation (Hansen and Olson, 1996) to support self-cultivation.

This perspective might lead us to ask what science might offer to support such personal development. As one example, science is associated with rational thought and science can offer contexts for applying logic, but this will only occur if suitable *activities* are included in the curriculum. Teaching about scientific products, such as theories, does not in itself engage logical thought, unless that teaching encompasses the argumentation that led to the ideas being proposed and accepted (Newton et al., 1999). Such activities might include designing experiments to test hypotheses, drawing inferences by analyzing data, and deducing whether findings support the hypotheses being tested—that is, elements of scientific enquiry.

Widely respected educational theories from psychology suggest that cognitive development requires opportunities to engage in pertinent activities (Piaget, 1970/1972; Vygotsky, 1978). So, for example, research suggests that learners often struggle to apply a basic “control of variables” logic to make inferences from data sets (Kuhn and Dean, 2005). They may notice an *apparent* pattern where a change in a dependent variable seems to reflect a change in the independent variable without taking into account conflating factors due to other potentially relevant variables which are also changing. This perhaps reflects biases inherent to human perception and cognition such as confirmation bias. In the Wason selection task (a seemingly simple task involving four cards, where participants have to decide which cards need to be turned over to test a simple conjecture such as “if a card shows an even number on one face, then its opposite face is red”) most people, even from among samples of highly educated adults, initially fail to appreciate the information needed to logically complete the task (Jackson and Griggs, 1988). Normally developing human youngsters have the potential to understand and apply the principle of control of variables (something fundamental to understanding experimental methodology) but fully realizing this potential may require considerable guidance, support and opportunities.

Science offers myriad disciplinary contexts where experimental tests are designed based on how possible outcomes can logically support or cast doubt upon some conjectured relationship (e.g., the water weed will release more bubbles of gas if the illumination is brighter; the current flow will drop as more lamps are added to a circuit, the metal strip will react quicker in the more concentrated acid, etc.) Such activities need some material context, and in order to be accessible for learners this needs to be a context familiar to them and where they already have relevant domain knowledge—but in principle such contexts could be drawn from across a very wide range of science topics. So, this perspective offers considerable flexibility in selecting science topics to be included in the school curriculum.

It can also be claimed that science offers opportunities for relating to, and appreciating, nature and the environment—but again this depends on the activities undertaken. Perhaps comparing the anatomy of unpleasantly-smelling, pickled, dead specimens of a wide range of animal species *in vitro* (once a common school science activity) can support developing some knowledge, but it may not support the development of a strong appreciation of the living world as much as unobtrusively observing a more modest range of living specimens in their natural habitats.

Clearly, in this regard, there are questions about whether a curriculum subject, science here, offers anything that can be considered unique. Mathematics also potentially offers myriad opportunities for logical thought and argumentation; art offers opportunities for closely observing and so appreciating the natural world; and geography offers opportunities for experiencing the natural environment. Arguably, what is particularly inherent to science is the way it sets up a dialectic process where progress is made through iteration between theory and testing—that is, observations inform new conceptualizations (conjectures) that suggest testing through setting-up further observations which allow refinement of conceptualizations, and so forth—and accordingly enquiry has long been proposed as a core feature of authentic science education (Schwab, 1962).

The cultural driver for science curriculum design

Another perspective sees education as about enculturation: cultural induction into society. This view would suggest that the school curriculum should introduce young people to *all* the major areas of activity that are valued in that society. Science is an important part of culture today, so every young person should be introduced to this aspect of culture. This is not primarily about employability, but in terms of making sense of, and having access to, the (cultural) world in which they live.

It was argued by the English scientist/science administrator and novelist Snow (1959/1998) that there were two cultures in his society—a kind of arts versus science divide, such that it was considered possible to be highly educated and cultured, yet to shamelessly claim complete ignorance of basic scientific ideas. Such a claim seems much less viable in the 21st century, and from this perspective, science education is necessary for a person not to feel excluded from this key strand of their culture. This is not about enabling scientific practices, but a kind of literacy relating to public discourse (Roberts and Bybee, 2014).

This curriculum driver would lead to us asking *what science should all young people be introduced to if they are to appreciate science as a cultural phenomenon*. Perhaps having some notion of what the Higgs boson is, and why seeking it justified the expenditure of such vast amounts of public money might be considered a candidate today. Less esoteric perhaps, it might be asked whether a person should be considered educated in the 21st century if they are unable to engage in informed public discourse about such matters as climate change, loss of biodiversity, or sustainable development.

Similar points may be made about issues of equity and fairness in society and public discourse around such matters as gender, race, and health inequalities. These are matters that clearly invite input from ethics and social science education, but there is also

important scientific background that is needed for engaging in informed debate—and some might suggest that science has a *particular responsibility* here due to a past record of both biased and misused science (Gould, 1992; Harding, 1993).

An obvious example of science that might claim a place in the curriculum here is learning about genetics and indeed epigenetics. For example, race is a social category as from a biological perspective there are not distinct human races—we are all one species, and, if we choose to employ the term in a scientific sense, one race. Yet there are certainly substantial differences in mean frequencies of some important health-related genes in different populations. So, sickle cell anemia is more frequent in people from regions where malaria is endemic, and lactose intolerance is normal in people in large parts of the world (South America, Africa, The Middle East, Southern Asia).

When students are taught about evolution by natural selection they may misconstrue the mechanism as being Lamarckian, as Darwinian inheritance seems to be more counter-intuitive to most people than the (now rejected) inheritance of acquired characteristics, and science educators have developed teaching approaches to address such alternative conceptions (Geraedts and Boersma, 2006). However, it is now recognized that in populations subjected to extreme stresses (such as extended periods of famine or warfare or being enslaved) there may indeed be trans-generational epigenetic effects that continue to disadvantage those born *after* the sources of stress have been removed (Meloni, 2019). There may then be some science topics which though known to be challenging to many learners, deserve a place in the curriculum because of their value in understanding key social discourses that are considered important in the culture.

In a multi-cultural society this cultural driver raises the question of into who's culture learners are to be inducted. Science *seeks* to be an objective means of learning about the natural world that produces culturally-blind knowledge, and is an international enterprise that should in principle be the same regardless of where it is practiced. Yet such an assumption may be problematic in relation to minority indigenous cultures. Often such communities embrace a worldview that does not draw sharp distinctions between science, technology, art, myth and religion. So, such communities value what is often termed traditional ecological knowledge—of a more synthetic nature—which has supported sustainable engagement with the environment for many generations (Ingilis, 1993).

Not only can such knowledge, which may incorporate mythical stories offering allegorical “truth,” seem inconsistent with what is taught in the science classroom, but it may reflect somewhat different values (for example, regarding the acceptability of keeping laboratory animals in cages for ready observation). The key challenge here for the cultural driver is not so much that there may be inconsistencies between accounts or practices arising from these different systems of thought, but rather that inducting learners into science as a *discrete cultural activity* is in itself contrary to the cultural traditions of some groups—indeed the very label “traditional ecological” [sic] knowledge may be seen as an artificial way to demarcate and divide the knowledge of indigenous communities (Kim et al., 2017). This is then an ongoing challenge to developing an inclusive science curriculum.

The civic driver for science curriculum design

A slightly different, but related, perspective asks what science all young people will need to help them in their adult lives regardless of whether or not they are to seek science-based careers. Very few people will be substantially disadvantaged by not being able to use a graduated pipette or not knowing how a transformer works. However, virtually everyone will at some time be faced with decisions about such matters as civic engagement (political action in its broadest sense, e.g., environmentalism) and medical treatment.

For example, a person may be told that they have (or perhaps that their child has) some medical condition where there is an uncertain prognosis, and various treatment options, each with possible side effects. Information provided may be framed in terms of likely and less likely outcomes, with or without treatments. Making an informed decision here will likely depend on some basic scientific understanding. For some people knowledge of the kidneys will be critical in making such a personal life decision, but for others it may be knowledge of the liver, the central nervous system, the endocrine system, the skeletal system, etcetera. This raises the question of just how much material should be specified in a common curriculum when members of society will face unique individual challenges. Rather than indicating a broad survey of a wide range of science topics, this perspective would prioritize deep understanding of a limited set of basic *concepts* that underpin the sciences (perhaps such as conservation of energy; molecular theory; a basic perspective on function and form in biology, and an appreciation of different levels of organization—organ systems, tissues, cells, etc.), complemented with *skills* in confidently accessing and critically evaluating information from across the sciences.

Often, in reaching decisions concerning such real-life dilemmas, no amount of knowledge of the science is sufficient by itself to indicate a definitive best course of action. In the hypothetical case of a serious medical diagnosis, even given all relevant technical information, one might still have to balance how much one would value some likely extra time on earth if choosing some particular recommended treatment with serious side effects (such as chronic nausea, vertigo and drowsiness). In such a context, someone who believed that life was always precious (perhaps sacred) and should always be protected at any cost, and another person who considered the main imperative in life was to maximize the quality of experience to increase pleasure and avoid pain, might reasonably (and completely rationally) make quite different value-mediated decisions when faced with the same choice and the same level of scientific knowledge and understanding. Similarly, someone who had outlived all their close friends and relatives might rationally make a different choice from someone looking forward to the birth of their first grandchild.

As another example, whatever the scientific case for protecting a natural habitat, different people may quite rationally put a different value on this, and have different views about the merits of avoiding habitat destruction when this means a longer commute into town or paying higher taxes. The contribution of the science curriculum would be to prepare learners to engage in public discourse regarding decision-making.

No schooling can provide *answers* to questions such as these, but it can offer experiences to develop *skills* in responding to such socio-scientific issues (Zeidler, 2014) where judgment involves engaging with systems of personal values as well as technical knowledge. The selection of science topics here (beyond those concepts referred to above that are judged foundational for learning *across* the sciences) will be informed by their utility in setting up learning experiences which support students' engagement with socio-scientific decision-making. Choices of context may also have a strong local component—drawing on examples and issues that students will recognize as relevant in their own communities (Sadler, 2011). This might be about the proposed building of a dam; the storage of nuclear waste; the poaching of ivory; the banning of a herbicide, and so forth. Current affairs may offer time-sensitive examples: the COVID-19 pandemic will remain a relevant context for exploring social and ethical issues around the balancing of economic activity, individual freedoms, and disease control measures (informed by knowledge of viruses, vaccination, hygiene measures, individual differences in response to infection, and so forth) in many national settings for as long as this remains a focus of public discourse.

Organization of the science curriculum

These different perspectives on the purposes of teaching and learning science link not only to choices about what to include in the curriculum, but also how to organize science in the curriculum. School science has at various times and in various places been taught primarily: as a program for developing laboratory practice; as an enquiry-based activity to be motivated by and respond to the learners' own interests; as exploration of a sequence of everyday themes or contexts (such as transport, cooking, clothing, health); with a focus on processes (including such activities as classifying as well as laboratory skills); in terms of issues arising from science understood in a social and economic context (Hunt, 1988); as well as in terms of disciplinary topics.

Disciplinary identify

There is a choice of curriculum subject(s)—that is “science or sciences?” One tradition has seen discrete sciences (most commonly biology, chemistry, physics) offered as examination subjects that would be taught as discrete courses. This approach was often most associated with those students being prepared for, or expected to aspire to, higher education. In some secondary school contexts the separation might be quite substantive: different teaching departments with their own designated staff responsible for their specialist subjects. Commonly these sciences would be taught in parallel, but sometimes different science subjects are taught, or offered, at different grade levels.

This tradition has some strengths. Students are given a strong basis to find out if they enjoy and are successful at different science subjects they might wish to select for further study. Teachers can be subject specialists (Reid and Ali, 2020), and teacher preparation and development can have a strong subject-pedagogic focus. Progression within the school can be stronger for learners when a small team of teachers is associated with the same school subject across grades. Where a range of sciences are offered to students as electives they may choose which to study and which to “drop,” giving them some say in their individual program of study (cf. Article 12 of the United Nations Convention on the Rights of the Child).

One objection to this approach is that it is “academic” and designed for a minority of learners (cf. the economic driver discussed above), whereas a broader science curriculum is more suitable for most learners. The term “general science” is sometimes used to refer to a course considered suitable for a wide range of abilities, while other terms may be intended to suggest that academic rigor is being retained while disciplinary content is organized within a single course, whether seeking to identify specific topics from the disciplines (co-ordinated science) or not (integrated science).

Such choices may have significance beyond how science is presented in the timetable. For example, when the English National Curriculum was introduced (Statutory Instrument, 1989) “science” was introduced as a compulsory subject for all learners from ages 5–16 years of age. Although schools could still choose how to timetable the subject (e.g., as separate sciences), “science” was the official school subject, and this had consequences in areas such as staffing (where in many schools an imbalance developed over time, with teachers with biology backgrounds coming to be over-represented compared with those with physics and chemistry backgrounds) and teacher education. Providers of initial teacher education were expected to ensure all those preparing for school science teaching were able to teach across the science curriculum. This had an effect on recruitment in some areas as previously many physics and engineering graduates had prepared to teach physics and saw mathematics as their natural second teaching subject—but now they could only prepare to teach physics alongside chemistry and biology.

A key feature of science is its organization within disciplines which each have their own traditions and characteristics in terms of core commitments, and any authentic science education needs to offer some sense of this (not necessarily, but most easily, done when science is timetabled as separate science subjects). These disciplinary commitments relate to the most fundamental entities and principles, and methodological norms adopted (which become further differentiated within fields within a discipline and to some extent within research programmes within fields).

A common core to the sciences

However, there are also common fundamental commitments across science which reflect the wider nature of science—in particular in terms of values. Science (in all fields) develops by an iterative process whereby observations inform conjectures which motivate

tests and so further observations. Science seeks objectivity. Even though there is much scholarship showing how this can only be an ideal, science values knowledge that is independent of the observer—in principle replacing one trained researcher by another should not influence scientific findings. Science requires honest and full reporting which can support replication and checking of reported results. Again this should be seen an ideal, as in practice there are recognized problems with replicability (Collins, 1992). Scientific authority (again, in principle) relies upon what can be demonstrated and logically argued, not claims based on seniority. Scientists are expected to be self-critical and to always seek alternative explanations for their observations and results. All scientific ideas are open to re-evaluation in the light of new evidence or new suggestions for how to interpret the existing data. Scientists value simplicity, and (internal and external) coherence. The most valued theories are those that encompass the widest range of phenomena with the fewest different concepts and assumptions. Theoretical findings in one field which seem inconsistent with knowledge in other areas of science are considered problematic (they may be tolerated for the time being but contradiction is seen as something that will eventually need addressing).

Scientists are human beings, with their flaws and idiosyncrasies, so it is important for learners to appreciate that the ideals of scientific work are not always met in practice, but this set of characteristics (or something quite like them, commentators will disagree on details) may be considered the normative core of what is to be considered “science,” and arguably the key task of the science curriculum is to support learners in appreciating this nature of science—which is relevant to those seeking scientific careers as well as the majority who will only engage with science as consumers and citizens (Lederman and Lederman, 2014).

If the primary focus of school science is to teach about science as practices of a diverse professional community informed by common scientific values, then this could potentially be achieved by selecting any number of combinations of science topics. Therefore supporting learners in engaging with areas where science impinges on society and personal decision-making can be perfectly consistent with a focus on learning about the core values of science in general.

The flexibility offered by a core focus on scientific values also allows scope to offer some experience of the disciplinary nature of science by ensuring that key concept areas and topics can be included which are associated with different disciplines. Indeed having selected science content according to such a scheme (to teach about the nature of science; to prepare learners to engage with science as citizens; to offer an introduction to the disciplines of science) there is still scope to “package” this as simply science, or to organize the selected topics under distinct discipline-based headings.

The STEM question

This leads to the question of whether school science should itself be subsumed under a broader STEM monicker (Chesky and Wolfmeyer, 2015). A school science department teaching its curriculum subject(s) could be part of a STEM faculty as an administrative arrangement which need not impact on the curriculum presentation itself. However, a strong argument for going beyond this relates to the artificiality of the common school curriculum structure, where learners experience a constant rotation around timetabled slots reflecting different subjects treated as knowledge “silos.” School learning is experienced in discrete aliquots (of perhaps an hour, sometimes less) before moving to a completely different topic in a different subject where there is no attempt to acknowledge or link to what students were doing just before. This compartmentalizes learning and prioritizes “academic” knowledge mainly useful for passing examinations.

A different approach would be to organize the curriculum in terms of more authentic activities where learners engage with projects over extended periods (as is more likely in the world of work) and where problem-solving draws upon whatever areas of knowledge are relevant to the task in hand (Rennie et al., 2012). STEM offers an opportunity for setting up such challenges where students have to solve problems, and build solutions, and where they will learn science and apply mathematics in the context of a realistic challenge. A team of learners might be allocated a sequence of half-day sessions to work on a particular “STEM” challenge.

Such an approach can potentially address some of the criticisms of school learning as being too divorced from the world of work, and involving a great deal of learning for its own sake. There is certainly much to be said for project-based teaching that leads to outcomes that seem tangible to learners (solving a problem in the local community; or at least developing something to be presented to a wider audience) although that can also be undertaken within “science” or specific science subjects.

While it is commonly argued that much groundbreaking work in science, medicine and technology today does take place in inter-disciplinary teams, by definition these teams can only exist because they bring together complementary specialists trained within different disciplines, and this is therefore of itself a weak argument for reorganizing education on such lines. A potential problem of moving to STEM as the subject *within which* science will be taught is staffing. Would the course be taught by an inter-disciplinary team or would “STEM” teachers expected to work across the constituent areas be needed. In most national contexts the latter option would require major changes in teacher preparation and development (which would only be sensible if a strong commitment to a permanent shift in his direction was made). The former option sets challenges within schools—how does one timetable STEM accordingly? One possibility is to block timetable so that a number of school classes work with a team of teachers with different specialisms—but realistically this also requires substantial shared teaching spaces that encompass laboratory, workshop, greenhouse, seminar areas, etc. A less revolutionary possibility is that classes move between specialists as timetable slots within the “STEM” subject periods—but this loses much of the potential flexibility of the interdisciplinary approach. There are considerable practical challenges here that are only worth addressing if moving to genuine “STEM” teaching was considered to justify this.

Another argument might be that STEM as a school subject offers learners a chance to see how science and mathematics can be applied—which is important for many learners who may fail to see the relevance of learning the subjects for their own sake. Yet if

science and mathematics were only taught in this way then a potential gain in terms of relevance could be at a cost of a failure to appreciate the true natures of mathematics and science as cultural practices in themselves. The application of science is not the practice of science. If we want learners to appreciate science as a cultural activity, then they should study science as a school subject (or some sciences as school subjects). While science and mathematics are fundamentally different in some senses (mathematics manages with deductive logic, relates to ideal entities, and involves genuine proofs; science always involves inductive as well as deductive logic, always has an empirical referent, and only offers provisional knowledge open to re-examination), they share putting a high value on new knowledge for its own sake without *requiring* buttressing by real world application. In this way they are quite different in nature from engineering and technology. Fig. 3 caricatures the three main areas of STEM at school level, showing that they all require the creative use of imagination, but for different purposes—and so their imaginative products are evaluated in different ways: to be internally consistent; to match (interpretations of) empirical data from the natural world; and to attain a practical end.

That need not mean there is no place for integrated STEM work. There may be great value for learners who study separate school subjects of science/the sciences, mathematics, technology periodically experiencing collapse of the normal timetable to engage in substantive authentic cross-curricular projects that involve applying their learning. Students then experience the diversity of different curricular traditions with specialist support, but periodically also engage with motivating scenarios where their learning can be applied, consolidated, and developed, without the constraints of the usual discontinuous timetable and the artificiality of disciplinary boundaries.

Differentiating the curriculum

Decisions about what to include in, and how to organize, the science curriculum may be customized for different groups of learners. It has sometimes been the case that “general science” and more contextualized (theme and issue based) science courses have been seen as more suitable for lower achievers, and separate sciences arranged around traditional disciplinary topics more appropriate for the highest achieving. Indeed, sometimes science courses for less academic groups have been framed as applied science (sending an

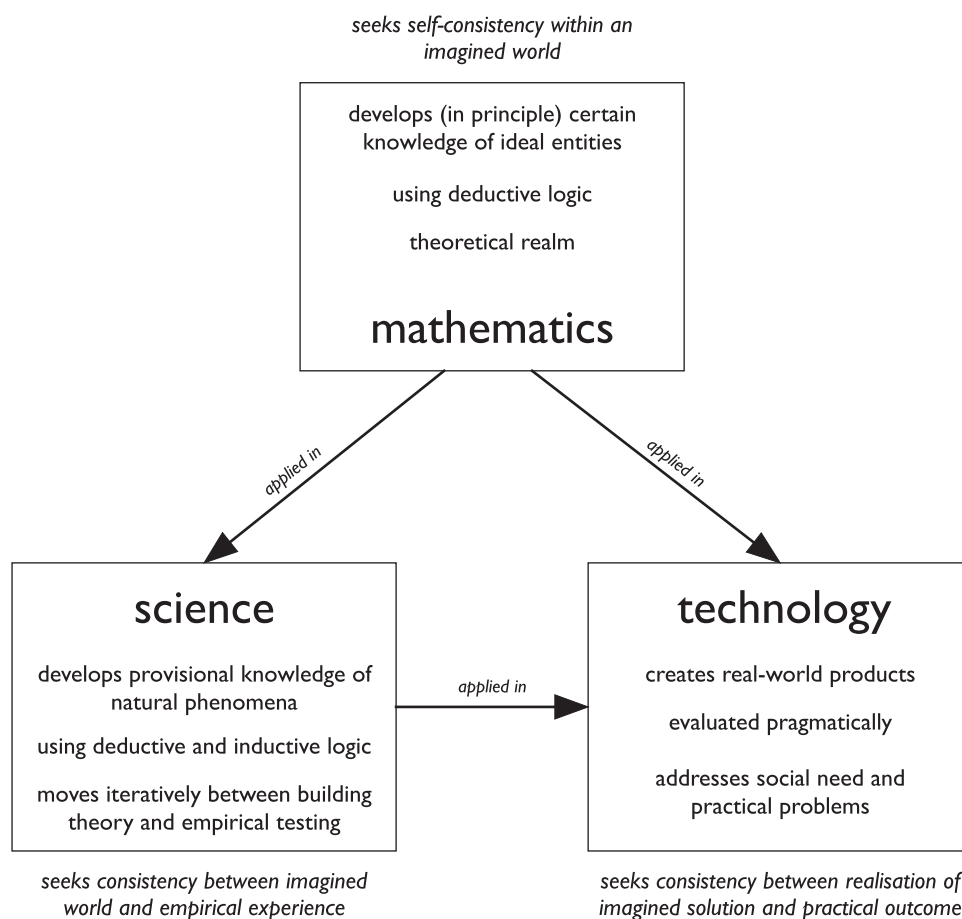


Fig. 3 Caricatures of science, mathematics and technology illustrating the fundamental differences in their natures. From <https://science-education-research.com/teaching-science/stem-education/>.

unwanted message about the relative status of engineering and technology). Yet, the considerations raised by the different drivers (see above) often apply across the ability range: indeed, learning to engage effectively in socio-scientific decision-making and appreciating the nature of science may be seen as requiring (and supporting the development of) higher levels of intellectual development (Taber, 2016).

Curriculum progression

There is a good case for modifying the nature of the science curriculum experience as students pass through their school career, reflecting how they mature and incrementally acquire experience of the natural world and on-going teaching. In early years education, for example, where teachers often lack formal qualifications and confidence in science, a focus on developing process skills (observing, recording, forming questions, etc.) in the context of nature study, and productive activities such as tending a garden, may lay important groundwork for later learning of formal concepts and theories (Taber, 2019a). The Japanese approach may offer a useful model here (Sumida, 2013). As an illustration of principle, Fig. 4 presents one model of how the science curriculum can be organized around different foci at different stages of schooling (very roughly as primary, middle and secondary phases) to better fit the interests and capabilities of learners of different ages.

Developing curriculum models to represent science

It is not usually the case that canonical knowledge can simply be selected and included in the curriculum “as is.” “School science” is not simply “science” transferred into a school context, but is rather a representation of it—something that may be more obvious in national contexts “such as German, Norwegian, Danish, and Japanese, [that] have different words to refer to science as an academic discipline and to science as a school subject” (Jenkins, 2007, p. 278).

For one thing, the referent for representation in the curriculum may need to be clarified. It may not be so easy for curriculum developers to unambiguously identify *canonical* knowledge in the sense of what scientists (as a community) mean by, for example, “an electronic orbital” or how scientists (collectively) understand “gravity.” In areas where research is currently active there may not be a single agreed “state of current knowledge.” In regards to areas of scientific knowledge that seem currently settled (and *strictly* all scientific knowledge is provisional in the sense of being open to be revisited), there will often be a historical sequence of models or theories. Some of these may now have uncertain status given that many models or accounts which have been strictly supplanted by more sophisticated (and so more difficult to understand and apply) models or accounts (i) are often retained by scientists for use in those contexts where they are considered to work well enough (e.g., Newton’s laws of motion); (ii) may have become partially incorporated into hybrid accounts included in textbooks (Justi and Gilbert, 2000).

As one example, it is not entirely clear if the canonical understanding of the acid concept should be considered as one particular conceptualization (the Lewis model) or a complex of several models (e.g., including the Brønsted-Lowry model) selected from pragmatically according to context. Specialists in a scientific area may not all entirely agree on such matters. In this sense the notion of canonical concepts may be considered as a useful fiction (Taber, 2019b). That is, it is helpful to talk of the placeholder of “the canonical account of a concept” as a referent when discussing curriculum design, without necessarily expecting to identify a version of the concept that attracts consensus among the scientific community. Such a position is not so alien to science itself which operates with such referents as non-existent “ideal” gases, and “laws” that are strictly never empirically followed (such as Raoult’s law), frictionless surfaces, point masses, and so forth.

Even when something can be identified as “the” scientific account of some topic adopted by those currently working in that area of science, it will likely be too complex and abstract for including in the school curriculum. Again, science itself is used to dealing with such situations by developing simplifying models that are useful for communicating ideas and inspiring testable hypotheses. Curriculum development also requires modeling of the science through the use of approximations and simplifications to provide “target knowledge” (i.e., that which it is hoped students will master) that can be engaged with, and understood, by learners at some specific level of cognitive development (Shayer and Adey, 1981). So, for example, if it is decided that the curriculum for a particular grade level will include learning about elasticity, acids and cell division, then this implies not the current scientific accounts *per se*, but specific *curricular models* of those accounts that will *represent* the science as it is to be taught at that level.

This is one area of ongoing scholarship in science education curriculum development, where curricular models are actively developed and proposed. The Institute of Physics has published an approach to teaching energy (an abstract, but foundational, concept where substantial learning difficulties have been documented) for 11–14 year old learners (Ireson, 2006). Specific teaching models have been developed to help move student thinking toward target knowledge, such as the “electronium” model of electric current intended for use with upper secondary (c. 18 year old) students (Buddle et al., 2002). This model is more sophisticated than the teaching models commonly used to introduce electrical circuit concepts to younger learners, which are based on analogies with familiar phenomena (Taber et al., 2015). This difference reflects ideas about progression in learning, and Bruner’s (1960) notion of the spiral curriculum. For example, target knowledge of the components of a plant cell at one grade level might be limited to a very small number of core elements each identified with a principal function, and then at higher grades further components, and more pluralistic accounts of their roles within the system, would be added to build up a more nuanced understanding.

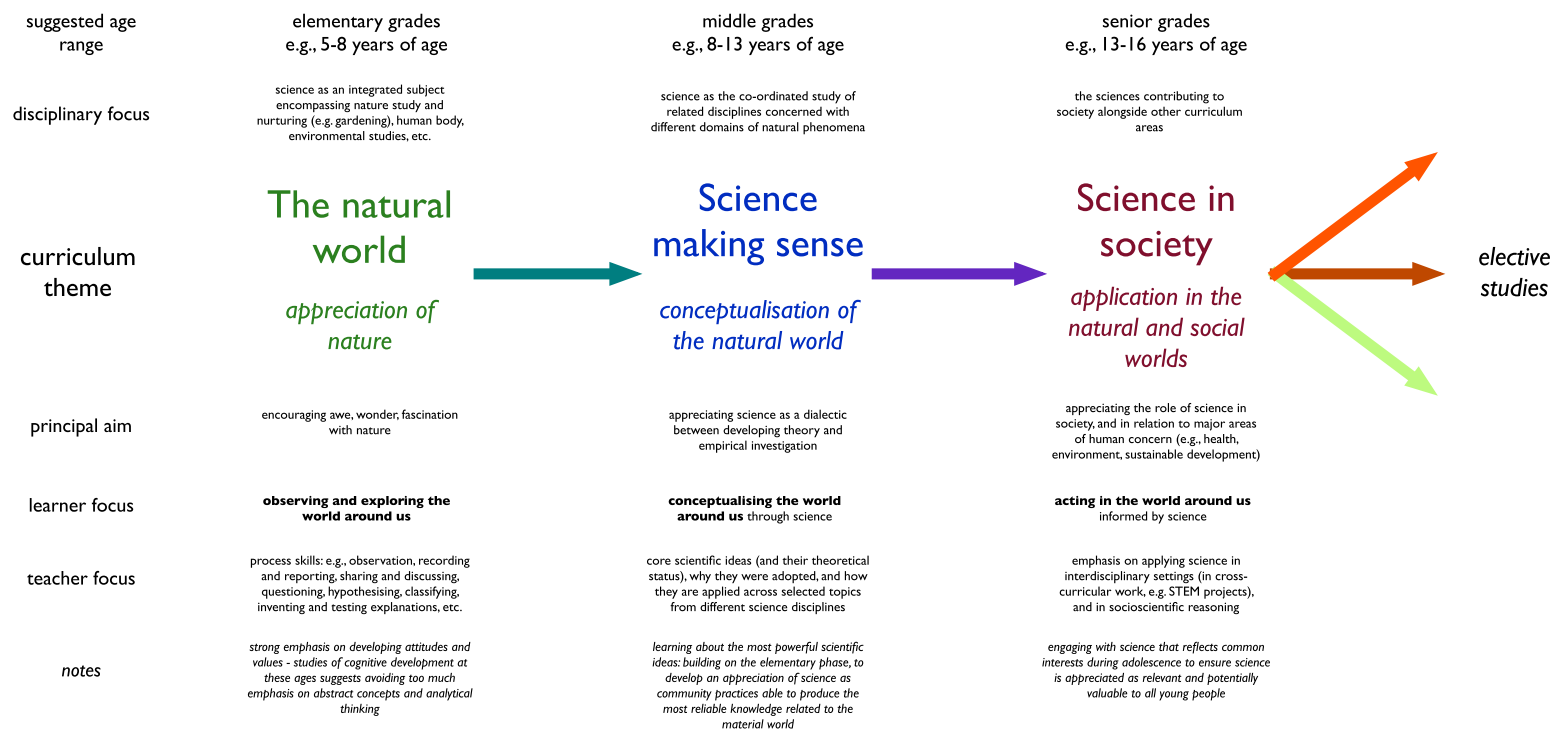


Fig. 4 An example of how the focus and nature of the science curriculum could shift across a learners' school careers reflecting their developing intellect and interests. From <https://science-education-research.com/teaching-science/the-science-curriculum/>, after Taber and Vong (2020), Fig. 2.

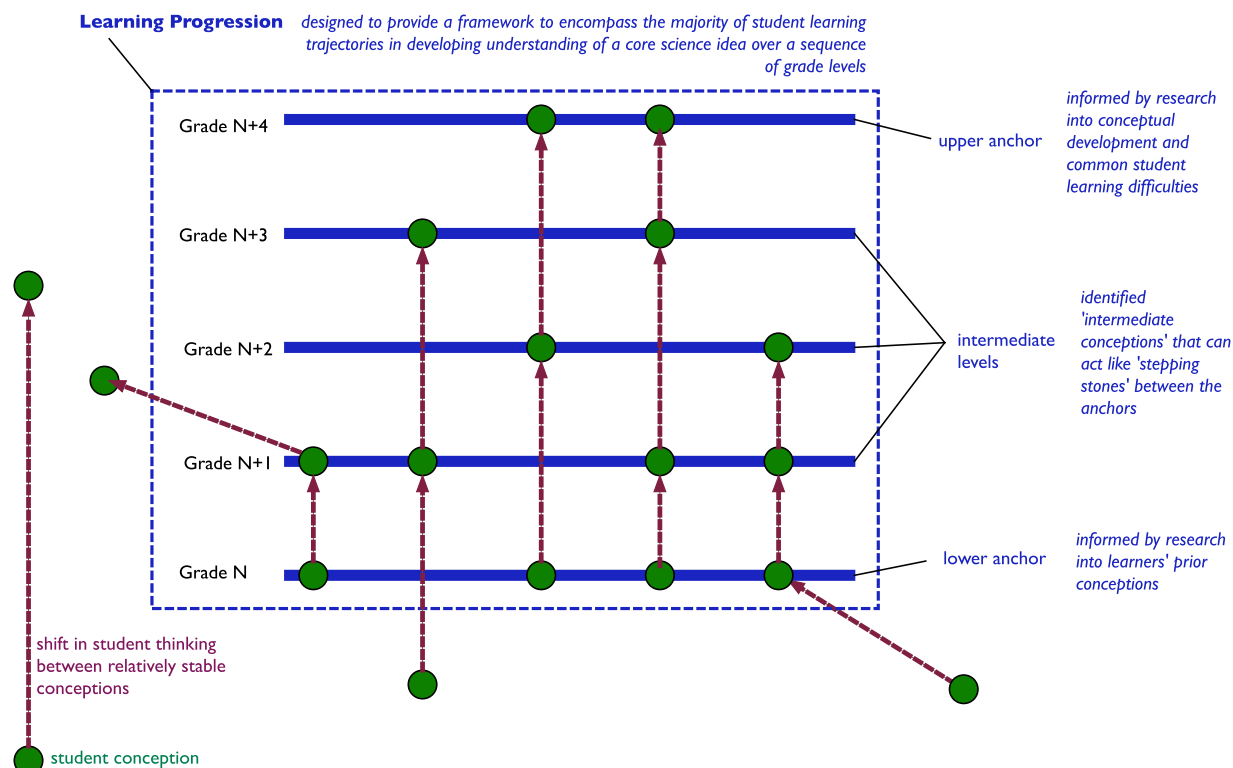


Fig. 5 Representation of a “learning progression”—a type of model to support curriculum development informed by research into learners’ ideas and studies of conceptual development. Figure from <https://science-education-research.com/learners-concepts-and-thinking/learning-progressions/> after Fig. 3, Taber (2021).

Learning progressions

This approach has been developed in recent years in an active research programme to develop tools referred to as “learning progressions” to guide the development of curriculum (Alonzo and Gotwals, 2012). While there is some ambiguity in the how the term is used (Taber, 2021), the approach seeks to understand optimal routes to building up more sophisticated target knowledge over successive school grades informed by the extensive literature on topic-specific learning difficulties and common alternative conceptions (see Fig. 5).

The representation of disciplinary knowledge in curriculum is inevitably distinct from what it is intended to represent. Curricular models deliberately simplify canonical content, and teaching models adopted in the classroom may further simplify target knowledge or represent it in more familiar, less abstract terms.

Even if some identified element of scientific content was considered suitable for including in the curriculum without deliberate simplification and modeling, that “element” would be excised from its wider context of historical development, scientific practice, and the embedding network of other “elements” of scientific knowledge against which it is understood and enacted by practicing scientists. Therefore, even when there is no deliberate attempt to modify the science to be taught, there is always a process of “didactic transposition” (Chevallard, 1989) which means that the knowledge *as presented in teaching* is inevitably a *representation* of the referent scientific content.

References

- Alonzo, A.C., Gotwals, A.W. (Eds.), 2012. *Learning Progressions in Science: Current Challenges and Future Directions*. Sense Publishers, Rotterdam.
- Bennett, J., Hogarth, S., Lubben, F., 2003. A Systematic Review of the Effects of Context-Based and Science-Technology-Society (STS) Approaches in the Teaching of Secondary Science: Review Conducted by the TTA-Supported Science Review Group. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.
- Bernstein, B., 1971/2003. On the classification and framing of educational knowledge. In: *Class, Codes and Control. Theoretical Studies Towards a Sociology of Language*, vol. 1. Routledge, Abingdon, Oxon, pp. 202–230.
- Bowker, G.C., 2005. *Memory Practices in the Sciences*. The MIT Press, Cambridge, Massachusetts.
- Bruner, J.S., 1960. *The Process of Education*. Vintage Books, New York.
- Buddle, M., Niedderer, H., Scott, P., Leach, J., 2002. “Electronium”: a quantum atomic teaching model. *Phys. Educ.* 37 (3), 197–203.
- Cannady, M.A., Greenwald, E., Harris, K.N., 2014. Problematizing the STEM pipeline metaphor: is the STEM pipeline metaphor serving our students and the STEM workforce? *Sci. Educ.* 98 (3), 443–460. <https://doi.org/10.1002/sce.21108>.
- Cerini, B., Murray, I., Reiss, M., 2003. *Student Review of the Science Curriculum: Major Findings*. Planet Science/Institute of Education/Science Museum, London.

- Chesky, N.Z., Wolfmeyer, M.R., 2015. *Philosophy of STEM Education: A Critical Investigation*. Palgrave Macmillan, New York.
- Chevallard, Y., 1989. On Didactic Transposition Theory: Some Introductory Notes Proceedings of the International Symposium on Selected Domains of Research and Development in Mathematics Education. Bratislava, pp. 51–62.
- Collins, H., 1992. *Changing Order: Replication and Induction in Scientific Practice*. University of Chicago Press.
- Eilks, I., Hofstein, A. (Eds.), 2015. *Relevant Chemistry Education: From Theory to Practice*. Sense Publishers, Rotterdam.
- Fensham, P.J., 1985. Science for all: a reflective essay. *J. Curric. Stud.* 17 (4), 415–435. <https://doi.org/10.1080/0022027850170407>.
- Geraedts, C.L., Boersma, K.T., 2006. Reinventing natural selection. *Int. J. Sci. Educ.* 28 (8), 843–870. <https://doi.org/10.1080/09500690500404722>.
- Glaserfeld, E.V., 1989. Cognition, construction of knowledge, and teaching. *Synthese* 80 (1), 121–140.
- Gould, S.J., 1992. *The Mismeasure of Man*. Penguin, London.
- Grumett, D., 2009. Natural theology after Darwin: contemplating the vortex. In: Northcott, M.S., Berry, R.J. (Eds.), *Theology After Darwin*. Paternoster, Milton Keynes, pp. 155–170.
- Hansen, K.-H., Olson, J., 1996. How teachers construe curriculum integration: the Science, Technology, Society (STS) movement as *Bildung*. *J. Curric. Stud.* 28 (6), 669–682. <https://doi.org/10.1080/0022027980280603>.
- Harding, S. (Ed.), 1993. *The “Racial” Economy of Science: Toward a Democratic Future*. Indiana University Press, Bloomington & Indianapolis.
- Hunt, A., 1988. SATIS approaches to STS. *Int. J. Sci. Educ.* 10 (4), 409–420. <https://doi.org/10.1080/0950069880100408>.
- Inglis, J.T., 1993. *Traditional Ecological Knowledge Concepts and Cases*. International Program on Traditional Ecological Knowledge International Development Research Centre, Ottawa.
- Ireson, G., 2006. *Energy: Supporting Physics Teaching*. Institute of Physics, pp. 11–14 [CD-ROM].
- Jackson, S.L., Griggs, R.A., 1988. Education and the selection task. *Bull. Psychonomic Soc.* 26 (4), 327–330.
- Jenkins, E.W., 2007. School science: a questionable construct? *J. Curric. Stud.* 39 (3), 265–282.
- Justi, R., Gilbert, J.K., 2000. History and philosophy of science through models: some challenges in the case of “the atom”. *Int. J. Sci. Educ.* 22 (9), 993–1009.
- Kim, E.-J.A., Asghar, A., Jordan, S., 2017. A critical review of Traditional Ecological Knowledge (TEK) in science education. *Can. J. Sci. Math. Technol. Educ.* 17 (4), 258–270. <https://doi.org/10.1080/14926156.2017.1380866>.
- Kuhn, D., Dean, D., 2005. Is developing scientific thinking all about learning to control variables? *Psychol. Sci.* 16 (11), 866–870. <https://doi.org/10.1111/j.1467-9280.2005.01628.x>.
- Lederman, N.G., Lederman, J.S., 2014. Research on teaching and learning of nature of science. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, vol. 2. Routledge, New York, pp. 600–620.
- Meloni, M., 2019. *Impressionable Biologies: From the Archaeology of Plasticity to the Sociology of Epigenetics*. Routledge, Abingdon, Oxon.
- Millar, R., Osborne, J., 1998. *Beyond 2000: Science Education for the Future*. King’s College, London.
- Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D.L., Fishbein, B., 2019. *Highlights. TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Centre, Boston.
- Newton, P., Driver, R., Osborne, J., 1999. The place of argumentation in the pedagogy of school science. *Int. J. Sci. Educ.* 21 (5), 553–576.
- Perry, W.G., 1985. Different worlds in the same classroom: students’ evolution in their vision of knowledge and their expectations of teachers. *Teach. Learn.* 1 (1), 1–17.
- Piaget, J., 1970/1972. In: Mays, W. (Ed.), *The Principles of Genetic Epistemology*. Routledge & Kegan Paul, London.
- Reid, N., Ali, A.A., 2020. *Making Sense of Learning: A Research Based Approach. Evidence to Guide Policy and Practice, With an Emphasis on Secondary Stages*. Springer Nature Switzerland AG, Cham, Switzerland.
- Rennie, L.J., Venville, G., Wallace, J., 2012. *Knowledge That Counts in a Global Community: Exploring the Contribution of Integrated Curriculum*. Routledge, Abingdon, Oxon.
- Roberts, D.A., Bybee, R.W., 2014. Scientific literacy, science literacy, and science education. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, vol. 2. Routledge, New York, pp. 545–558.
- Sadler, T.D. (Ed.), 2011. *Socio-scientific Issues in the Classroom: Teaching, Learning and Research*, vol. 39. Springer, Dordrecht.
- Schwab, J.J., 1962. The teaching of science as enquiry (The Inglis Lecture, 1961). In: Schwab, J.J., Brandwein, P.F. (Eds.), *The Teaching of Science*. Harvard University Press, Cambridge, Massachusetts.
- Shayer, M., Adey, P., 1981. *Towards a Science of Science Teaching: Cognitive Development and Curriculum Demand*. Heinemann Educational Books, Oxford.
- Snow, C.P., 1959/1998. The Rede Lecture, 1959: the two cultures. In: *The Two Cultures*. Cambridge University Press, Cambridge, pp. 1–51.
- Statutory Instrument, 1989. *The Education (National Curriculum) (Attainment Targets and Programs of Study in Science) Order 1989*. HMSO, England, London.
- Sumida, M., 2013. The Japanese view of nature and its implications for the teaching of science in the early childhood years. *Int. Perspect. Early Child. Educ. Care* 243.
- Taber, K.S., Vong, L.T.K., 2020. Lumping and splitting in curriculum design: curriculum integration versus disciplinary specialism. In: Bachmeier (Ed.), *Curriculum Perspectives and Development*. Nova Science Publishers, New York, pp. 1–66.
- Taber, K.S., Ruthven, K., Howe, C., Mercer, N., Riga, F., Hofmann, R., Luthman, S., 2015. Developing a research-informed teaching module for learning about electrical circuits at lower secondary school level: supporting personal learning about science and the nature of science. In: de Silva, E. (Ed.), *Cases on Research-Based Teaching Methods in Science Education*. IGI Global, Hershey, Pennsylvania, pp. 122–156.
- Taber, K.S., 2016. The nature of science and the teaching of gifted learners. In: Taber, K.S., Sumida, M. (Eds.), *International Perspectives on Science Education for the Gifted: Key Issues and Challenges*. Routledge, Abingdon, Oxon, pp. 94–105.
- Taber, K.S., 2019a. Exploring, imagining, sharing: early development and education in science. In: Whitebread, D., Grau, V., Kumpulainen, K., McClelland, M.M., Perry, N.E., Pino-Pasternak, D. (Eds.), *The SAGE Handbook of Developmental Psychology and Early Childhood Education*. Sage, London, pp. 348–364.
- Taber, K.S., 2019b. *The Nature of the Chemical Concept: Constructing Chemical Knowledge in Teaching and Learning*. Royal Society of Chemistry, Cambridge.
- Taber, K.S., 2021. Squaring the circle: circumnavigating an ontological tension between practical learning progression models and the complex, multi-faceted, and meandering nature of conceptual learning. In: Bowman, M.N. (Ed.), *Topics in Science Education*. Nova Science Publishers Inc, New York, pp. 1–100.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, Massachusetts.
- Zeidler, D.L., 2014. Socioscientific issues as a curriculum emphasis: theory, research, and practice. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, vol. 2. Routledge, New York, pp. 697–726.

Curriculum and social studies education

Steven Camicia, School of Teacher Education and Leadership, Utah State University, Logan, UT, United States

© 2023 Elsevier Ltd. All rights reserved.

Defining curriculum	327
Historical perspectives of curriculum and social studies	328
Contemporary perspectives on curriculum and social studies	329
Conclusion: future directions for social studies curriculum	330
References	331

The social studies curriculum sends powerful messages to students and communities about how they should understand themselves. Pinar (2004) wrote that “The school curriculum communicates what we choose to remember about our past, what we believe about the present, what we hope for the future” (p. 20). This is particularly true with social studies curriculum. When we think of the subject areas that are most associated with social studies, such as civics, history, geography, and economics, it becomes evident that these subjects communicate a worldview of the past, present, and future. As curriculum is implemented in schools, students can develop a better understanding of communities ranging in scale from the school to the globe. The National Council for the Social Studies (2022) in the United States explains that “the primary purpose of social studies is to help young people develop the ability to make informed and reasoned decisions for the public good as citizens of a culturally diverse, democratic society in an interdependent world” (emphasis in original, para. 2). Through a better understanding of the past, present, and future, the social studies curriculum emphasizes the need to increase student understanding of citizenship, democracy, multiculturalism, and the globe.

Because there are limitations on the time and resources devoted to social studies education in classrooms, decisions must be made about how to select from the vast universe of knowledge available. The results of these decisions communicate to students, educators, and communities what knowledge is of most value. Conversely, these decisions communicate what knowledge is of least value and not part of the social studies curriculum. Eisner (2002) described what is not included in the curriculum as the null curriculum. Historical and contemporary struggles over the social studies curriculum reform reside in what perspectives, individuals, and communities are portrayed. By including or excluding some individuals and groups from narratives, for example, the social studies curriculum can send powerful messages to students, educators, and communities about who is valued and who is not.

Due to a limited amount of instructional time in the school curriculum, the choices of the individuals, groups, and perspectives become central to the design and implementation of the social studies curriculum. Unfortunately, the social studies curriculum is often a reflection of inequitable power relations in society because dominant individuals and groups often use the social studies curriculum as a way to construct and maintain ideologies that benefit these dominant individuals and groups (Apple, 2018). When subjects such as civics and history reflect these ideologies, students are sent powerful messages about what individuals, groups, and perspectives are of value, as well as which ones are of little value and not worthy of inclusion (Au, 2012; Au et al., 2016; Camicia and Knowles, 2021).

This chapter examines curriculum and social studies education by first defining curriculum in general. There are many ways to describe curriculum, so it is important to provide a general description that can converge on a more specific discussion of social studies curriculum. As the introductory paragraph communicates, it is helpful to understand curriculum as representing a worldview or ideology. Next, I examine social studies education specifically. Because the universe of information involved with social studies subjects is vast, a thematic focus can help with this examination. I conclude the chapter with a discussion of current and future directions in social studies curriculum.

Defining curriculum

Perennial struggles to define the qualities and scope of “curriculum” have ebbed and flowed throughout the history of the United States (Kliebard, 2004). While it is beyond the scope of this chapter to develop a comprehensive historical account of curriculum and its reform, it is important to provide an operational description as a foundation for the rest of the chapter. This will provide a backdrop for understanding why social studies curriculum has been particularly contentious throughout its history. While many mistakenly attribute “instructional materials” or “instructional standards” to “curriculum,” these are only a partial understanding of the qualities and scope of the term “curriculum.”

The difficulty involved with defining curriculum is in the expansiveness of its facets. [Flinders and Thornton \(2013\)](#) capture the challenge posed by describing curriculum writing:

Even a simple definition of “curriculum” such as “what educating institutions should teach” compels answering a vast, possibly overwhelming, set of questions including: what subject matter should be selected and by what criteria, how should the subject matter be organized, to whom should it be taught and by what methods and with what materials, how is it most effectively learned, and how might we judge the success of the undertaking. No one is an expert on the entirety of this still incomplete set of questions every curriculum designer answers, whether directly or tacitly (p. ix).

There are at least three aspects of this description. First, it is apparent that there are multiple facets of curriculum design such as criteria of selection, which could involve the selection of aspects ranging from instructional materials to instructional methods. Second, there are individuals positioned differently in relationship to the curriculum. What are the relationships between students, and what are the relationships between students and educators? For example, does the curriculum reflect authoritarian or democratic relationships between the individuals who experience the curriculum ([Camicia and Si, 2021a](#); [Riddle and Apple, 2019](#)). Third, there are individuals and groups who make decisions about these different considerations. The three aspects above illustrate the complexity of considerations involved with curriculum. Because there are different individuals and groups that make decisions about what to include in the curriculum and what will be part of the null curriculum, the network of power relations in society at large is mirrored in the curriculum.

Curriculum is difficult to describe because it encompasses both explicit and implicit objectives and contexts. While daunting because of the scale and complexity involved with describing curriculum, [Beyer and Liston’s \(1996\)](#) provided a helpful list of questions that reflect the considerations about curriculum that I have included to this point:

1. What knowledge and forms of expression are most worthwhile?
2. What is the relationship between the knowledge embodied in the formal curricula of schools and those who are involved in enacting it?
3. What types of educational and social relationships are required or desirable in order to facilitate curricular experiences?
4. How do larger social, political, and institutional contexts affect the experiences students have with curriculum?
5. What are the implicit (and explicit) conceptions of democracy within the curriculum?
6. What are the implicit (and explicit) visions of students’ social, political, and economic futures, and how does curriculum prepare students for those futures? (pp. 218–219)

These questions serve as a heuristic for understanding and examining curriculum. Returning to [Pinar’s \(2004\)](#) definition of curriculum as a reflection of the intentions for understanding the past, present, and future, it is easy to see how Beyer and Liston’s questions reflect this. It is also easy to see how social studies curriculum might connect to these questions because social studies topics such as civics, geography, history, and economics connect past, present, and future concerns for our communities. I next examine social studies curriculum more specifically.

Historical perspectives of curriculum and social studies

There have been perennial struggles over how to frame social studies curriculum. Because the social studies curriculum has had a primary objective of developing informed citizens, the rationales for curricular reform have centered upon the types of communities we want to develop ([Camicia and Franklin, 2011](#); [Camicia and Knowles, 2021](#)). It is beyond the scope of this chapter to discuss a detailed history of social studies curriculum and its reform, but there are numerous accounts that go into more detail (e.g., [Evans, 2004](#); [Hertzberg, 1981a,b](#); [Lybarger, 1991](#); [Thornton, 2008](#); [Woyshner and Bohan, 2012](#)). The origins of social studies education is usually associated with the [National Education Association’s \(1916\)](#) report that placed citizenship education and history education as the primary focus of social studies curriculum in the United States ([Ross, 2006](#)). Other subject areas have been included over time, and there have been multiple reform movements related to progressive and critical orientations ([Kliebard, 2004](#)).

Within this history, as well as currently, one of the main themes is a struggle between an orientation of social studies toward a *social science* approach or a *social education* approach ([Thornton, 2005, 2008](#)). The social science approach emphasizes disciplinary boundaries as they mirror academic subject areas at the university level. While drawing upon other disciplines, the social science approach places a discipline such as geography centrally, as well as a course entitled “Geography.” The disciplinary approach focuses upon a conception of community where students gain an increased understanding of each discipline to better understand the world.

In contrast, the social education approach blurs disciplinary boundaries in favor of the needs of students and society ([Hertzberg, 1981a](#); [Thornton, 2008](#)). [Thornton \(2008\)](#) described the social education approach as “identifying the individual and social demands of associated living and then deciding what material from the social sciences (and collective material) is relevant to those demands” (p. 20). The social education approach focuses upon a conception of community where the goal is to examine topics and issues that are of importance to community members. In this approach, students play a central role in choosing these topics and

issues. As Thornton noted, these two approaches are not mutually exclusive, and there is a blurring of distinctions. These two approaches and the blending of these approaches in social studies curriculum communicate a stance on what knowledge and worldviews are privileged.

Historically, in the elementary grades, the expanding horizons (Wade, 2002) or expanding environments (Akenson, 1989) model has guided much of the scope and sequence of social studies curriculum in elementary grades across the United States. Wade (2002) described this model and its expanding scope for each grade as: “kindergarten-self, first grade-family, second grade-neighborhood, third grade-community, fourth grade-state and/or regions, fifth grade-United States, sixth grade-Canada and Mexico” (p. 115). As an organizing principle for social studies curriculum, this model emphasizes a worldview where children experience the world from one location and expand out from what they know to regions with which they aren’t familiar.

Unfortunately, the expanding horizons model in elementary grades social studies curriculum makes multiple problematic assumptions about the lived experiences of students in elementary grades. The assumption that children start with what they know in the location they grow up in is challenged by human migration, where students often experience multiple communities and places on the globe from a young age (Camicia and Saavedra, 2009; Saavedra and Camicia, 2010). A social studies curriculum that emphasizes the expanding horizon model risks excluding the experiences and perspectives of students in classrooms that have traveled and lived in different regions and parts of the globe. This is particularly problematic when we understand that one of the primary goals of social studies is to help students understand global interdependence. The exclusion of students’ perspectives that involve firsthand experiences of different locations on the globe harms these students because their knowledge is not valued, and this exclusion denies the rest of the students knowledge of global interdependence.

The expanding horizons model provides an illustration of how dominant ideologies and worldviews often guide social studies curriculum and reinforce social inequalities found in society at large. In addition, the expanding horizons model does not appear to be responsive to the realities of increasing media access and social media usage among elementary-level students. Although this model is ubiquitous in the United States, Wade (2002) noted that there are other alternatives, such as the “Toward the Common Good (TCG)” (p. 120) model, where the elementary level scope and sequence are based upon TGC themes. Wade proposed that these themes could be people, environment, exploration, conflict, democracy, human rights, and interdependence in kindergarten to sixth grade, respectively. Rather than an exclusionary force in social studies curriculum, the TGC model could serve as an inclusionary force by increasing the range of perspectives in the elementary grades.

The thematic approach is used in many of the current standards in social studies curriculum in the United States. These include the following themes, which are guidelines for each grade, K-12: Culture; time, continuity, and change; people, places, and environments; individual development and identity; individuals, groups, and institutions; power, authority, and governance; production, distribution, and consumption; science, technology, and society; global connections; and civic ideals and practices. These ten themes guide curriculum design and implementation at each grade level, and the themes are not mutually exclusive. Students examine one or more of these themes drawing upon the social sciences and other forms of information.

Contemporary perspectives on curriculum and social studies

Many of the historical perspectives of social studies curriculum continue today. In this section, I start by examining current descriptions of social studies curriculum, which are multifaceted and complex. As mentioned above, curriculum, in general, has many facets, but social studies curriculum multiplies these facets due to the number of disciplines included under the umbrella of social studies education. Next, I examine some contemporary perspectives that present opportunities for change and social justice. The current description of social studies reflects the expansiveness of social studies curriculum. According to the National Council for the Social Studies (NCSS) (2022) in the United States:

Social studies is the integrated study of the social sciences and humanities to promote civic competence. Within the school program, social studies provides coordinated, systematic study drawing upon such disciplines as anthropology, archaeology, economics, geography, history, law, philosophy, political science, psychology, religion, and sociology, as well as appropriate content from the humanities, mathematics, and natural sciences. The primary purpose of social studies is to help young people develop the ability to make informed and reasoned decisions for the public good as citizens of a culturally diverse, democratic society in an interdependent world.

(emphasis in original, para. 2).

This description illustrates the interaction of the two social studies curriculum approaches described in the previous section, the social science approach and the social education approach. While the social science approach is included, the National Council for the Social Studies (2022) description starts and finishes with an emphasis on civic competence. This indicates a social education approach because civic competence requires solving problems that are of importance to students and their communities. In other words, students as decision-makers are placed centrally within the learning objectives of social studies curriculum. In addition, the last sentence provides a view into current and future directions of social studies by describing students as citizens navigating “a culturally diverse, democratic society in an interdependent world” (emphasis in original, para. 2).

The foci of cultural diversity, democracy, and global interdependence in social studies curriculum as areas of emphasis create opportunities for students to gain nuanced understandings of different types and scales of community. Since community is a central aspect of citizenship, these areas of emphasis can provide guidance for curriculum development in the areas of multicultural education (e.g., Au et al., 2016; Banks, 1996, 2004; Banks et al., 2005; Castro, 2013; Ladson-Billings, 2004; Moll and González, 2004; Takaki, 2012) education for democracy (e.g., Camicia and Knowles, 2021; Dewey, 1966; Gay, 1997; Gibson, 2020; Hess and McAvoy, 2015; Lo, 2017; Ochoa-Becker, 2007; Pace, 2015; Parker, 2004; Riddle and Apple, 2019; Westheimer and Kahne, 2004), and global education (e.g., Abdi and Richardson, 2008; Andreotti, 2011; Banks, 2017; Becker, 2002; Gaudelli, 2013; Hanvey, 1976/1982; Merryfield and Subedi, 2001; Myers, 2006; Osler, 2015; Oxfam, 2006).

A read of the titles of the works cited above supports the idea that democratic citizenship, education for democracy, multicultural education, and global education are conceptually linked. They are linked through an emphasis on perspective consciousness and increasing perspectives through increasing inclusion. With this in mind, the emphasis upon identifying exclusions in social studies curriculum has been a significant theme of social studies research and curriculum development (e.g., An, 2016; Anderson and Metzger, 2011; Bickford, 2017; Brown and Brown, 2010; Busey and Walker, 2017; Camicia, 2016; Eargle, 2016; Hawkman and Shear, 2020; Journell, 2009; Shear et al., 2015).

Whether social studies curriculum follows a social science approach, social education approach, or a mix of the two, an emphasis on increasing perspectives in the curriculum addresses the democratic mission of schools to educate citizens within the contexts of multicultural and global communities. As the sections above have illustrated, social studies curriculum contains a vast universe of possibilities, approaches, and pedagogies. However, a focus on increasing perspectives is an increasing priority. When framed as a way to educate citizens to participate in democratic communities, an emphasis on increasing inclusion can be a way to increase the legitimacy of the social studies curriculum. Inclusion is closely connected with democratic legitimacy because democratic governance relies upon the inclusion of individuals, groups, and perspectives in this governance. In the next section, I conclude this chapter with considerations of inequitable power relations in society, a lens for inclusionary and exclusionary forces in curriculum.

Conclusion: future directions for social studies curriculum

As Beyer and Liston's (1996) questions prior to this section indicate, social studies curriculum is not designed and implemented in a vacuum. Instead, people are positioned differently in relation to the curriculum, and the curriculum is embedded in different socio-cultural and political contexts (Apple, 2018; Camicia, 2016). When the curriculum is portrayed as a neutral set of facts that are settled knowledge, students are denied an understanding of how knowledge is constructed in society and social science communities. Knowledge in such communities constitutes a political struggle between different understandings of the past, present, and future. By designing and representing social studies curriculum as neutral, curriculum designers present students with an apolitical view of the world. This is ironic when we consider that the stated objectives of the social studies curriculum involve civic education and political involvement in various communities.

An apolitical stance places higher value on authoritarian rather than democratic governance (Camicia and Si, 2021a). Apple (2018) pointed to the limited treatment of social movements in social studies curriculum as an example of how curriculum presents students with inaccurate perspectives on how society changes. Camicia and Si (2021b) examined social studies curriculum standards across the United States. They're analysis found the frequencies of social movement inclusion was extremely low in most states. By focusing on change from those in power rather than from individuals in society and their involvement in social movements, inequitable power relations are constructed and reinforced. Instead, the social studies curriculum needs to help students examine inequitable power relations in society and the ways that they can change their communities for the better with this increased understanding.

In conclusion, there are multiple directions that social studies curriculum can reform. Since "the primary purpose of social studies is to help young people develop the ability to make informed and reasoned decisions for the public good as citizens of a culturally diverse, democratic society in an interdependent world" (emphasis in original, para. 2), the social studies curriculum needs to emphasize education for democracy. Democracy in this sense is not reactionary but revolutionary (Camicia and Knowles, 2021). Democratic governance requires students to make informed choices drawing upon the knowledge from multiple disciplines.

The current national social studies standards emphasize an inquiry-based approach to social studies curriculum (National Council for the Social Studies, 2013; Swan et al., 2018). This is a good place to build upon since there is an emphasis upon multiple perspectives on topics and issues. This can serve as a point of departure by considering some of the references included above, which focus upon increasing student understanding of multiculturalism, democratic governance, and global interdependence. By valuing the knowledge that students bring to classrooms and increasing their awareness of multiple perspectives, students can build their understanding of how to solve problems in and out of their communities.

Although increasing the range of perspectives in social studies curriculum is a great start, increasing perspectives on topics and issues of shared concern is not enough. Perspectives in society are not valued equally, and to treat them as equal reinforces inequalities in society. Instead, social studies curriculum needs to locate different perspectives within networks of power relations. With the increasing movements toward authoritarianism and other forms of oppression, social studies curriculum needs to draw upon multiple sources of knowledge. This knowledge can emanate from students, communities, or social studies disciplines. But principles of democratic governance and understandings of inequitable power relations provide students with the tools necessary to place various perspectives within a social justice framework (Camicia and Knowles, 2021).

References

- Abdi, A.A., Richardson, G., 2008. Decolonizing Democratic Education.
- Akenson, J.E., 1989. The expanding environments and elementary education: a critical perspective. *Theor. Res. Soc. Educ.* 17 (1), 33–52.
- An, S., 2016. Asian Americans in American History: an AsianCrit perspective on Asian American inclusion in state U.S. history curriculum standards. *Theor. Res. Soc. Educ.* 44 (2), 244–276.
- Anderson, C.B., Metzger, S.A., 2011. Slavery, the civil war era, and African American representation in U.S. history: an analysis of four states' academic standards. *Theor. Res. Soc. Educ.* 39 (3), 393–415.
- Andreotti, V.D.O., 2011. *Actionable Postcolonial Theory in Education*. Palgrave Macmillan.
- Apple, M.W., 2018. *Ideology and Curriculum*, fourth ed. RoutledgeFalmer.
- Au, W., 2012. *Critical Curriculum Studies: Education, Consciousness and the Politics of Knowing*. Routledge.
- Au, W., Brown, A.L., Calderón, D., 2016. Reclaiming the Multicultural Roots of U.S. Curriculum: Communities of Color and Official Knowledge in Education. Teachers College Press.
- Banks, J.A., 1996. Multicultural Education, Transformative Knowledge, and Action: Historical and Contemporary Perspectives.
- Banks, J.A., 2004. Multicultural education: historical development, dimensions, and practice. In: Banks, J.A., Banks, C.A.M. (Eds.), *Handbook of Research on Multicultural Education*, second ed. Jossey-Bass, pp. 3–29.
- Banks, J.A. (Ed.), 2017. *Citizenship Education and Global Migration: Implications for Theory, Research, and Teaching*. American Educational Research Association.
- Banks, J.A., Banks, C.A.M., Cortés, C.E., Hahn, C.L., Merryfield, M.M., Moodley, K.A., Murphy-Shigematsu, S., Osler, A., Park, C., Parker, W.C., 2005. *Democracy and Diversity: Principles and Concepts for Educating Citizens in a Global Age*.
- Becker, J.M., 2002. Globalization and global education: ever the twain shall meet? *Int. Soc. Stud. Forum* 2 (1), 51–57.
- Beyer, L.E., Liston, D.P., 1996. *Curriculum in Conflict: Social Visions, Educational Agendas, and Progressive School Reform*. Teachers College Press.
- Bickford, J.H., 2017. The representations of LGBTQ themes and individuals in non-fiction young adult literature. *Soc. Stud. Res. Pract.* 12 (2), 182–198. <https://doi.org/10.1108/SSRP-05-2017-0021>.
- Brown, K.D., Brown, A.L., 2010. Silenced memories: an examination of the socio-cultural knowledge on race and racial violence in official school curriculum. *Equity Excell. Educ.* 43, 139–154. <https://doi.org/10.1080/10665681003719590>.
- Bussey, C.L., Walker, I., 2017. A dream and a bus: Black critical patriotism in elementary social studies standards. *Theor. Res. Soc. Educ.* 45 (4), 456–488. <https://doi.org/10.1080/00933104.2017.1320251>.
- Camicia, S.P., 2016. *Critical Democratic Education and LGBTQ-Inclusive Curriculum: Opportunities and Constraints*. Routledge.
- Camicia, S.P., Franklin, B., 2011. What type of global community and citizenship? Tangled discourses of neoliberalism and critical democracy in curriculum and its reform. *Glob. Soc. Educ.* 9 (3–4), 311–322.
- Camicia, S.P., Knowles, R.T., 2021. *Education for Democracy: A Renewed Approach to Civic Inquiries for Social Justice*. Information Age Publishing, Inc.
- Camicia, S.P., Saavedra, C., 2009. A new childhood social studies curriculum for a new generation of citizenship. *Int. J. Child. Right.* 17, 501–517.
- Camicia, S.P., Si, S., 2021a. Critical education for democracy under discourses of authoritarianism and nationalism. *Citizen. Teach. Learn.* 16 (3), 381–395. https://doi.org/10.1386/ctl_00072_1.
- Camicia, S.P., Si, S., 2021b. State Social Studies Standards and Social Movements: A Content Analysis. American Educational Studies Association, Portland, OR.
- Castro, A.J., 2013. What makes a citizen? Critical and multicultural citizenship and preservice teachers' understanding of citizenship skills. *Theor. Res. Soc. Educ.* 41 (2), 219–246.
- Dewey, J., 1966. *Democracy and Education: An Introduction to the Philosophy of Education*, first free press paperback ed. The Free Press; Collier-Macmillan.
- Eargle, J.C., 2016. The dominant narrative of slavery in South Carolina's history standards. *J. Soc. Stud. Res.* 40, 295–307.
- Eisner, E.W., 2002. *The Educational Imagination: On the Design and Evaluation of School Programs*, third ed. Prentice Hall.
- Evans, R.W., 2004. *The Social Studies Wars: What Should We Teach the Children*. Teachers College Press.
- Flinders, D.J., Thornton, S.J., 2013. Preface. In: Flinders, D.J., Thornton, S.J. (Eds.), *The Curriculum Studies Reader*. Routledge, pp. ix–x.
- Gaudelli, W., 2013. Critically theorizing the global. *Theor. Res. Soc. Educ.* 41 (4), 552–565.
- Gay, G., 1997. The relationship between multicultural and democratic education. *Soc. Stud.* 88 (1), 5–11.
- Gibson, M., 2020. From deliberation to counter-narration: toward a critical pedagogy for democratic citizenship. *Theor. Res. Soc. Educ.* 1–24. <https://doi.org/10.1080/00933104.2020.1747034>.
- Hanvey, R.G., 1976/1982. *An Attainable Global Perspective*.
- Hawkmán, A.M., Shear, S.B. (Eds.), 2020. *Marking the "Invisible": Articulating Whiteness in Social Studies Education*. Information Age Publishing.
- Hertzberg, H.W., 1981a. *Social Studies Reform, 1880–1980*. Social Science Education Consortium.
- Hertzberg, H.W., 1981b. Social studies reform: the lessons of history. In: Morrisett, I. (Ed.), *Social Studies in the 1980s*. Association for Supervision and Curriculum Development, pp. 2–14.
- Hess, D.E., McAvoy, P., 2015. *The Political Classroom: Evidence and Ethics in Democratic Education*. Routledge.
- Journell, W., 2009. An incomplete history: representation of American Indians in state social studies standards. *J. Am. Indian Educ.* 48 (2), 18–32.
- Kliebard, H.M., 2004. *The Struggle for the American Curriculum: 1893–1958*, third ed. RoutledgeFalmer.
- Ladson-Billings, G., 2004. New directions in multicultural education: complexities, boundaries, and critical race theory. In: *Handbook of Research on Multicultural Education*, pp. 50–65.
- Lo, J.C., 2017. Empowering young people through conflict and conciliation: attending to the political and agonism in democratic education. *Democr. Educ.* 25 (1). Article 2. <https://democracyeducationjournal.org/home/vol25/iss1/2>.
- Lybarger, M.B., 1991. The historiography of social studies: retrospect, circumspect, and prospect. In: Shaver, J.P. (Ed.), *Handbook of Research on Social Studies Teaching and Learning*. Macmillan Publishing Company.
- Merryfield, M.M., Subedi, B., 2001. Decolonizing the mind for world-centered global education. *Soc. Stud. Curric.* 277–290.
- Moll, L.C., González, N., 2004. Engaging life: a funds-or-knowledge approach to multicultural education. In: *Handbook of Research on Multicultural Education*, pp. 699–715.
- Myers, J.P., 2006. Rethinking the social studies curriculum in the context of globalization: education for global citizenship in the U.S. *Theor. Res. Soc. Educ.* 34 (3), 370–394.
- National Council for the Social Studies, 2013. *The College, Career, and Civic Life (C3) Framework for Social Studies Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History*. National Council for the Social Studies.
- National Council for the Social Studies, 2022. *What is Social Studies?* National Council for the Social Studies. <https://www.socialstudies.org/about>.
- National Education Association, 1916. *Social Studies in Secondary Education: A Six-Year Program Adapted to the 6-3-3 and the 8-4 Plans of Organization: Report of the Committee on Social Studies of the Commission on the Reorganization of Secondary Education*. U. S. Government Printing Office.
- Ochoa-Becker, A.S., 2007. *Democratic Education for Social Studies: An Issues-Centered Decision Making Curriculum*.
- Osler, A., 2015. Human rights education, postcolonial scholarship, and action for social justice. *Theor. Res. Soc. Educ.* 43 (2), 244–274.
- Oxfam, 2006. *Education for Global Citizenship: A Guide for Schools*. <http://www.oxfam.org.uk/coolplanet/teachers/globciti/downloads/gcguide06.pdf>.
- Pace, J.L., 2015. *The Charged Classroom: Predicaments and Possibilities for Democratic Education*. Routledge.
- Parker, W.C., 2004. Diversity, globalization, and democratic education: curriculum possibilities. In: *Diversity and Citizenship Education: Global perspectives*, pp. 433–458.
- Pinar, W.F., 2004. *What is Curriculum Theory?* Lawrence Erlbaum Associates, Inc. Publishers.

- Riddle, S., Apple, M.W. (Eds.), 2019. *Re-imagining Education for Democracy*. Routledge.
- Ross, W. (Ed.), 2006. *The Social Studies Curriculum: Purposes, Problems, and Possibilities*, third rev. ed. State University of New York Press.
- Saavedra, C., Camicia, S.P., 2010. Transnational childhoods: bodies that challenge boundaries. In: Cannella, G.S., Diaz Soto, L. (Eds.), *Diversity, Multiplicity and Childhoods*. Peter Lang Publishing, pp. 27–38.
- Shear, S.B., Knowles, R.T., Soden, G.J., Castro, A.J., 2015. Manifesting destiny: re/presenting of indigenous peoples in K-12 U.S. history standards. *Theor. Res. Soc. Educ.* 43 (1), 68–101.
- Swan, K., Lee, J., Grant, S.G., 2018. *Inquiry Design Model: Building Inquiries in Social Studies*. National Council for the Social Studies and C3 Teachers, United States.
- Takaki, R.T., 2012. In: Steffoff, R. (Ed.), *A Different Mirror for Young People: A History of Multicultural America*. Seven Stories Press.
- Thornton, S.J., 2005. *Teaching Social Studies That Matter: Curriculum for Active Learning*. Teachers College Press.
- Thornton, S.J., 2008. Continuity and change in social studies curriculum. In: Levstik, L.S., Tyson, C.A. (Eds.), *Handbook of Research in Social Studies Education*. Routledge, pp. 15–32.
- Wade, R., 2002. Beyond the expanding horizons: new curriculum directions for elementary social studies. *Elem. Sch. J.* 103 (2), 115–130.
- Westheimer, J., Kahne, J., 2004. What kind of citizen? The politics of educating for democracy. *Am. Educ. Res. J.* 41 (2), 437, 269.
- Woyshner, C., Bohan, C.H. (Eds.), 2012. *Histories of Social Studies and Race: 1865–2000*. Palgrave Macmillan.

Curriculum, Didaktik, and Bildung

Bernadette Hörmann, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	333
Curriculum	333
Didaktik	334
Bildung	335
Recent developments	336
Concluding remarks	337
References	337

Introduction

Curriculum, *Didaktik*, and *Bildung* represent three essential cornerstones of public schooling in general and teaching and learning in schools in particular. They are tightly related to each other and mutually dependent. Originating from a central European, mostly German discourse, they stand for a framework that has characterized schooling in Europe and beyond for almost two centuries, although their origins can be traced back to ancient philosophy and history. In a first step, this contribution presents the three phenomena by revealing their interrelatedness and some common grounds. In a second step, the article discusses how the recent trend of competence based curricula changes and affects the traditional concepts of *Didaktik* and *Bildung*. In a third step, current issues and future questions and tasks for research are presented.

Since the three terms curriculum, *Didaktik*, and *Bildung* refer mostly to theories originating from German, the translation of the terms into English is problematic. While “curriculum” is often used in research publications as a general term that also comprises the German concept of *Lehrplan*, the terms *Didaktik* and *Bildung* in fact do not have any equivalences in English. Each of the following sections briefly amplifies the challenges given with respect to the terms and their translation into English.

Curriculum

Curricula in terms of syllabi are a societal agreement that declare what knowledge a society considers relevant to be taught to its future generation in school and that should assure that societies preserve their identity and existence throughout generations (Hopmann and Künzli, 1994). In a German tradition, curricula (in German: *Lehrpläne* (plural), *Lehrplan* (singular)) deal with the what, when, and how of teaching and learning and specifically refer to a policy document that displays content descriptions for different subjects and grades in schools and in many cases a general description of values and principles that apply to the schools in the respective nation state. In contrast to this, the English term *curriculum* is more comprising and less specific by referring to practice in terms of what happens in a classroom, what is planned by the teacher, and what is expected to be the outcome of a class. The comparative research project “Didaktik meets Curriculum” has defined the term “Didaktik” as the comparative unit to the term “curriculum” (Gundem and Hopmann, 1998), which points to the different scope of curricula compared to the term *Lehrplan*. Even if the curriculum and *Lehrplan* are in fact no synonyms (Horlacher, 2018) and due to practical reasons, I will—in line with other research articles—use curriculum as a more comprising term that also includes syllabi in terms of *Lehrpläne*.

In the German speaking countries, curricula are steering tools in terms of laws that are not only provided to public schools, but that are developed as a subject of public deliberations. They gained significance within the emergence of the European nation states in the late 18th and 19th century. This process was accompanied by a strong desire and need to create capable and loyal citizens, not least by building an efficient school system controlled by the state and its administration body (Hopmann, 1991). An important challenge in curriculum making processes is to gain the acceptance for the new law and its content from practitioners, the public, politicians, or administrative bodies. In addition, the law has to cover a highly diversified landscape of schools and schooling with varying local characteristics and needs, different school types, and different age groups. A historical analysis of curriculum making processes in Prussia (Hopmann, 1991) revealed three basic features that reveal the close connection between a curriculum and the society in which it is enacted. First, the process of curriculum making and enactment has become highly differentiated. As a consequence of the formation of ever more differentiated structures in administration, curriculum work has become subject in a number of institutions, authorities, and administration bodies. These compartments provide a common ground for expertise in the multifaceted issues in curriculum making and enacting, while the people at the same time are able to hide between a corporate responsibility, where no one could be made responsible for the whole (Hopmann, 1991; Westbury, 2008). In so doing, the administration of a nation state ensures that the complex project of developing one national curriculum is processed on a formal level by organizing decision making processes. At the same time, compartmentalization contributes to the consolidation of a nation state’s administration body, because it helps to make reform and decision making processes more efficient. Second, curricula are by definition

instruments of “administrative supervision” (Westbury, 2008, p. 52), and therefore separated from instruction planning at the classroom level. This means that the teachers as executives of the curriculum gain freedom in operating and enacting the framework given in the curriculum and are free to choose between content matter and methods. When choosing content matter and making decision on methods, teachers draw on the curriculum in order to legitimize their choices and decisions. The feature of licensing therefore also allows for a division of responsibilities that diversifies the burden in case of failure. At the same time, it lays ground for a powerful discipline that will be presented in the next section: the discipline of *Didaktik*. The freedom of methods and its separation from the curriculum is thus a key concept for understanding how curriculum, *Didaktik* and *Bildung* are connected to each other. Third, curriculum making needs to take the demands and experience from various groups of people into consideration. By the segmentation of the levels of discourse in curriculum reform, different parts of the society that are involved in the reform matter, get an arena where they can be included into the process. The process of negotiating which content matter is supposed to be included in the curriculum becomes part of a public and partly transparent deliberation where practitioners and other people involved can bring forward arguments, take position, contribute aspects and perspectives to the discussion and give precise reasons for what they are demanding and claiming. This happens both in formally organized conferences and meetings between practitioners, public officials, and politicians and in informal ways by for example participating in media discourse. The three features of compartmentalization, licensing, and segmentation balance and steer the different interests in curricula as a governing tool. On the one hand, they assure not only that professional expertise is laying the ground for a curriculum, but also develop structures for how to organize this expertise (compartmentalization). The functioning of societies relies heavily on professional expertise as a bundle of knowledge and expertise that exaggerates the stock of knowledge that an individual person can have. The licensing feature as the core ground for the teachers’ professional practice allows teachers to justify their work toward the society on the basis of the curriculum. The segmentation of discourse assures that all people who have an interest in the future generation’s education can participate and share thoughts (Hopmann, 1991). Curricula from this perspective are not an instruction defined by one authority, but a matter that takes shape through societal deliberations and negotiations between different parts of the society.

The core element of a curriculum is of course the content matter that is listed and described in the document and that is supposed to be taught by the teachers. Content matter is Hopmann (2007) presented along three dimensions that also evolved one after another in the course of history. The first dimension, order, is about the structure and concepts of knowledge that is imparted to each subject and each content matter. Through curricula, content matter is on the one hand set into relation to a bigger picture, a superior classification of which the content is a part. On the other hand, content matter can be broken down into smaller aspects that are part of the content’s structure and that in fact characterize it. The second dimension deals with the sequence of subject matter, and the question how subject matter can be presented in a way that the students can follow and understand it. The sequence can reveal an order according to the difficulty of content matter, it can follow a chronological order, a thematic order, it can reach from the near to the far, and follow the different development levels of children. The third dimension, choice, is owed to the circumstance that a curriculum can never cover all knowledge that is available for humankind. It therefore can only present a choice of content matter, that either is exemplary, considered to be of high relevance, teachable, or which differs in other respects from what is decided to not be included in the curriculum. These three dimensions provide not only the basic framework for the formal document of curricula, but also of the teachers’ planning process, where they arrange content matter according to the needs of the students (Hopmann, 1991).

Historically regarded, the process of developing curricula was dedicated to the search for a universal, God given order of things, into which humankind should gain insight. However, in the course of time, the universal approach became increasingly reduced by following the above mentioned dimensions. After attempts to create universal pictures of knowledge, the sequence of content matter became important, followed by the necessity to choose specific content from the vast variety of knowledge of humankind (Hopmann, 2007). It was not before national states were founded that curricula have taken the shape we are familiar with today. The desire to develop a national identity and membership/loyalty among the nation states’ citizens, and especially among their future generation, has led the way for curricula that chose content matter shaped and in line with a national narrative (Tröhler, 2016). In contrast to the idea of the “*enkyklia paidea*” in ancient Greece, the aim was no longer to cover the whole available knowledge to humankind, but to define a nationally relevant choice of knowledge that can build the fundament of a society. The complex process of curriculum making reveals how close curricula are related and dependent on the society in which they are supposed to be implemented. Curricula in our times are developed within a specific society, and for a specific society, and therefore both are mutually dependent on each other.

Didaktik

Didaktik has again its origin in German theories and goes back to Ancient Greek terms. “*Didaskein*” means to teach and to show or perform something and has been developed to the “art of teaching.” It is the “pedagogy of teaching” (Oettingen, 2016) and provides, as a science of teaching, the knowledge teachers need for being able to teach (Oettingen, 2016). One can differentiate between subject matter didactics which deals with teaching of content matter of a specific subject and general didactics which is about general phenomena, premises and characteristics and of teaching.

As in case of the term *Lehrplan*, the term *Didaktik* is difficult to translate into English. While the English term “didactics” rather refers to methods of instruction and the science of teaching, the German term rather opens for the reflection on the question how content matter can be made meaningful to a group of students by a teacher. The important difference between these two

understandings lies in the concept of *Bildung* that is inherent to the German term *Didaktik* and that will be presented in the next section of the article.

Curricula in terms of *Lehrpläne* set essential conditions for teaching. They provide a legal and publicly accepted framework of content matter and give the teachers freedom in what content they choose and how they enact it in the classroom. This freedom allows teachers to develop their approaches in adaptation to their specific classroom at hand, with their students' specific needs and based on their professional experience and knowledge. The cornerstones under which this freedom develops become visible in a schematic representation of models of didactics, the didactic triangle (Künzli, 1998). It visualizes the relations between the general elements of teaching, i.e. the teacher, the student, and the content matter, which are indispensable conditions for institutionalized learning and define the core of didactics. Different *Didaktik* approaches emphasize specific relations within the triangle and put different amounts of weight to the role of each component. However, learning only unfolds within the relations between the three components teacher, student, and content. In search of the common core of *Didaktik*, Hopmann defines three distinctive corner stones of *Didaktik* that at the same time visualize its historical development in line with curricula: didactics as a matter of order, sequence, and choice (Hopmann, 2007). In terms of order, didactics deals with providing the students with a structured preparation of a certain content matter that helps the students with gaining insight into its inner structure and the higher order of which the content matter is a part. Simultaneously, any matter at hand features a specific logic that implies different steps in the teaching process. By drawing on their professional expertise and experience, teachers can thus prepare meaningful steps and tasks for the teaching of the same content. The matter of sequence relies mostly on the work of Johann Amos Comenius and his "Didactica Magna" where he elaborates not only on the structure of the content as a realm of teaching, but also on the student's relation and their presuppositions. Therefore, Comenius suggests to start teaching from the (student's) near and proceed to the far, from the simple to the complex, or from the more obvious and known to the more abstract and unknown. The third corner stone is choice, which refers to the necessity to choose between a variety of content depending on the appropriateness and the specific needs of the students at hand. The enormous wealth of knowledge that is available to humankind requires thoughtful choice of what is ought and worth being taught to tomorrow's generation. *Didaktik* unfolds here between a political framework in terms of the curriculum and the teachers' local practice: the curriculum serves as a basis for the legitimation of choosing (or not choosing) a specific content, while the actual enactment of the content underlies the teachers' professional expertise and experience. By finishing their education in teaching, teachers receive, in Hopmann's terms, a license for transforming and enacting content matter as provided by curricula according to best pedagogical practice (Hopmann, 1991). The three cornerstones describe the art of teaching as a way of restraining the students' exposure to an overwhelming flow of knowledge and content: teachers need to prepare content according to its structure, develop a sequence for the presentation of different steps, and choose specific content matter from the wealth of available knowledge. As argued by Herbart, the art of teaching mainly builds on the teachers' experience in preparing content matter according to the conditions and needs of their class at hand, and reasonate about it and legitimate it by drawing on the curriculum. The teachers' capability to realize pedagogical needs and necessities and to react in a decent way by making new decisions, is the actual art of teaching that is based on experience and ethics. By drawing on Herbart's concept of tact, Max van Manen describes the teachers' spontaneous pedagogical reactions as the foundation of teaching (Van Manen, 2016, 1986).

Bildung

The term *Bildung* is to be understood on the backdrop of a Jewish-Christ tradition of thinking. The "deo imagi" (picture of God) denotes the symbolical relation between God and the human being and points to the difference between the picture of the human being as it is, and the picture that she or he is supposed to become. This difference is expressed as a longing for self-improvement with the ideal picture of God in mind (Ehrenspeck, 2006). The word "picture" means *Bild* in German, which does not only mean the picture as such but the verb *bilden* also means the process of "forming." Also in its secularized meaning, the idea of *Bildung* refers to a mutual relation between the human being and the world, with a forming and changing effect for the human being, and, eventually, also for the world. In the context of schooling, the dialectic interplay between the learning human being and the content matter representing the world is extended with a social dimension. Not only the teachers as mediating persons participate in the learning, but also other students share the learning process and turn it into a common effort to develop a common understanding of the world. Learning is, from this perspective, not about accumulating knowledge in a material sense, but about making meaning of a specific content that represents a canon of knowledge to be learned. It presupposes a theoretical distinction between content matter and meaning which represents at the same time the intersection between curricula, *Didaktik*, and *Bildung*. While content matter is the knowledge defined in the curricula, the didactical question is how content can be made meaningful to the students. The meaning-making process is a unique, situated encounter of the student's previous, biographically determined stock of experience with the content matter under those didactical conditions that the teacher creates (Hörmann, 2011, 2015). *Bildung* therefore emerges as an individual meeting in which students activate "memories of prior personal experiences" (Menck, 2006, p. 116) and, in so doing, gain a "picture" in their imagination that represents the meaning of the content taught by the teacher. The fact, however, that learning in school is embedded in a social setting, makes the process of meaning making interactive. By communicating with their classmates and the teacher, students put forward different intuitive or immediate perceptions (Willbergh, 2008) and share a common, vivid presence in which they become participants of a common effort to understand, define, and categorize the world. The teacher navigates through this process and doses the extent of guidance according to the students' needs and becomes a participant in the common effort of understanding and exploring the issue at hand. By restraining teaching (Hopmann, 2007), teachers

create conditions under which both individual growth and social exploration and deliberation can become possible. The main purpose of this process is to develop educative substance or meaning which can be opened up by any content, while any given matter can represent many different meanings (Hopmann, 2007, p. 116). This relation allows the teacher to take into account the students' various backgrounds and different stocks of experience within the endeavor of making content matter meaningful to them. Content matter in a narrow sense and the academic disciplines in the form of school subjects in a wider sense denote/are of a rather exemplary character that serve as a framework for the students' exercises to understand concepts, structures, and categories of the world (Künzli, 1981). According to Klafki, one main purpose in lesson planning and teaching in general is to reflect on the meaning of the specific content for the presence and the future. This points to the school's mandate to prepare students for their future life as members of a democratic and highly complex community that requires citizens who are not only able to read and write, but who also can communicate with each other, share a common presence, and actively participate and engage in society (Klafki, 1967, 2007).

Especially by drawing on Humboldt's theory on *Bildung*, scholars and practitioners often promote *Bildung* as a transformation with the purpose of self-improvement by dealing with content matter from the classical arts and languages and focusing on esthetics. From a social historical perspective, *Bildung* has been a bourgeois ideology for promoting the interests of a privileged and aspiring social class in the 19th century (Horlacher, 2015). Despite of a variety of theories and practices around the idea of *Bildung* and different readings such as, for instance, by the *Frankfurter Schule* (Horlacher, 2015), modern schools have been strikingly characterized and formed by a bourgeois understanding and attitude toward learning and knowledge, not least because both policy makers and teachers often originate from these backgrounds. This leads to the situation that students from other backgrounds than the typically bourgeois educated family suffer from a mismatch between the school's and teachers' expectations and their biographically marked resources. Nevertheless, scholars in classical didactics emphasize a comprehensive understanding of *Bildung* which is not confined to specific sets of content, but regard any content matter as an exemplary representation of the world that can turned into meaningful learning experiences for all students (Klafki, 1967). Education is in this framework the nurturing of a natural development process that is inherent to each human being ("The human being carries the sufficient reason for its development in itself from the beginning of life until the point of its completion" origin: "*Der Mensch ist ein Wesen, welches den hinreichenden Grund seiner Entwicklung vom Anfang des Lebens an bis zum Punkt der Vollendung in sich trägt*" (Schleiermacher, 1826), own translation).

Different readings and theories on *Bildung* reveal a diverse picture of the phenomenon and turn it into a concept that is drawn on from many different angles and parties (Horlacher, 2015). In recent times, it is often used as a counter concept to education in the framework of competences and performance improvement. At the same time, proponents from standardized and competence oriented teaching proclaim the literacy concept to be a renewed version of the idea of *Bildung*. In the German language, *Bildung* can mean any kind of education and it is not confined to a specific concept. However, by drawing on a didactical framework, this article presented *Bildung* in relation to the didactic premises and thinking.

Recent developments

In the course of internationalization and globalization, many nation states have implemented competence-based curricula and oriented their education policies toward learning outcomes in terms of standardized testing. Curriculum making and education policy making have become both international and national projects which has led to a situation in which policies are oriented toward globally disseminated ideas, while these ideas simultaneously are reshaped and renegotiated on a national level (Sivesind and Wahlström, 2016; Steiner-Khamsi, 2012). Competence based curricula, for instance, have become a common reference point in education reforms in many countries all over the world, national governments, however, responded in fact differently to the common idea and reshaped and renegotiated it by drawing on their traditions and cultures (Karseth and Sivesind, 2010; Sivesind et al., 2016). For instance, in comparison to Norway, Finland's curricula balance both content oriented with a learning outcome oriented descriptions in the syllabi, whereas Norwegian curricula are based on a learning outcomes approach (Mølsted and Karseth, 2016). Anderson-Levitt also points to varying and inconsistent implementations of competence approaches in curricula in different countries with regard to the underlying theoretical premises or different foci on the students' individualized progress vs. integrative, project-based approaches (Anderson-Levitt, 2017).

The current trend of standardized testing and competence based teaching itself can be regarded as an adoption of elements from two long-standing traditions; the German *Didaktik* tradition on the one hand, and the Anglo-American curriculum tradition on the other hand. The test culture originating from the Anglo-American tradition combined with state based curricular formats have become a rapidly expanding movement, causing fundamental changes on the policy and the practical level (Hopmann, 2015). From a didactical perspective, competences connect content matter with a specific meaning and therefore suspend the distinction between the two aspects under which *Didaktik* unfolds (Hopmann, 2007). Within the competence paradigm, the focus of teaching lies on the students' competence development, whereas teaching in the tradition of *Bildung* rather deals with opening up content matter for the students in a way that it becomes meaningful to them. A vague concept like this is criticized because it sets mastery in a relative context and appears too subjective. Competence concepts however promise objective proofs for the successful mastery of knowledge by formulating cognitive goals and testing the students' performance with standardized tests (Köller, 2008). The licen-siation for teachers to enact content matter appropriate to their professional pedagogical standards is subject of change in a competence framework, because content matter becomes applied knowledge where the meaning is prescribed (Hopmann, 2007). Teaching under these terms means to "facilitate learning" (Biesta, 2013) by working with the students on their mastery of specific

applications of knowledge with a focus on each individual's progress (Masschelein and Simons, 2013). Since each individual's progress is to the fore, the meaning of the social dimension of learning and deliberating together in the classroom diminishes. This means that the focus of schooling lies on its qualifying function while its socializing and subjectivization functions become subordinated (Biesta, 2009). Bildung is redefined as the successful mastery of specific, applied knowledge represented by the concept of literacy that promises to provide students with the knowledge required in societies of the future (OECD, 2005).

Concluding remarks

This article has presented core concepts of education that originate from German discourse and at the same time have become subject of and involved in international renegotiations and processes of exchange. Curricula, *Didaktik*, and *Bildung* are cornerstones in public education in schools in modern nation states that have historically grown out of the need for institutions that secure the future generation's education and the continuance of the nation states. Recent policies are oriented toward international trends and global ideas and have led to profound changes of national curricula which are now based on competences. Besides from an international influence, an increased influence from the private market on educational practice lead to the situation that national states play a "denationalized role" in policy making where the role of the nation state tends to be transformed from an "enabling" to an "influencing-liberal" state (Sundberg and Wahlström, 2012; Wahlström, 2014). In the further course, the change in the conceptualization of curricula has empirical consequences for the enactment of teaching and the idea of the aim and purpose of schooling. These changes have been investigated extensively and critiques mostly refer to the narrow understanding of education in the competence and testing framework and the pressure that is put on students, teachers, and parents (for example Biesta, 2009; Hopmann, 2007, 2008; McGill-Franzen and Allington, 1993; Otto and Schrödter, 2010; Willbergh, 2015). An important area of future research however could be the way how old and new elements in curriculum and teaching are combined with each other and how accountability tools can be integrated into a genuine pedagogical understanding of schooling that balances the schools' functions and mandates with the purpose of opening the students future possibilities for an autonomous, good and succeeding life (Bachmann et al., 2021; Geppert, 2014).

References

- Anderson-Levitt, K., 2017. Global flows of competence-based approaches in primary and secondary education. *Cah. rech. éduc. savoirs* 16, 47–72.
- Bachmann, K., Hörmann, B., Sivesind, K., 2021. Bildung: alive and allowed? A critical study of work plan practices in Norwegian schools. *J. Curric. Stud.* 1–17. <https://doi.org/10.1080/00220272.2021.1936191>.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21, 33–46.
- Biesta, G., 2013. Receiving the gift of teaching: from "learning from" to "being taught by". *Stud. Philos. Educ.* 32, 449–461.
- Ehrenspeck, Y., 2006. Bildung. In: Krüger, H.-H., Grunert, C. (Eds.), *Wörterbuch Erziehungswissenschaft*, pp. 64–71. Basel.
- Geppert, C., 2014. Von "Value Added" zu "Freedom Added": Übertritts- und Verlaufsmuster niederösterreichischer SchülerInnen an der Bildungsübergangsschwelle zur Sekundarstufe I.
- Gundem, B.B., Hopmann, S.T. (Eds.), 1998. *Didaktik and/or Curriculum: An International Dialogue*. American University Studies, New York, 14, Education, Lang, P.
- Hopmann, S.T., Künzli, R., 1994. Topik der Lehrplanung: das Aarau Lehrplannormal. *Bildungsforschung Bildungspraxis* 16, 161–184.
- Hopmann, S.T., 1991. Retracing curriculum history: the multiple realities of curriculum making. In: *Curriculum Work and Curriculum Content. Theory and Practice. Contemporary and Historical Perspectives*, Rapport. Pedagogisk Forskningsinstitutt Universitetet i Oslo, Oslo, pp. 49–68.
- Hopmann, S.T., 2007. Restrained teaching: the common core of Didaktik. *Eur. Educ. Res. J.* 6, 109–124. <https://doi.org/10.2304/eej.2007.6.2.109>.
- Hopmann, S.T., 2008. No child, no school, no state left behind: schooling in the age of accountability. *Null* 40, 417–456. <https://doi.org/10.1080/00220270801989818>.
- Hopmann, S., 2015. "Didaktik meets Curriculum" revisited: historical encounters, systematic experience, empirical limits. *Null* 2015, 27007. <https://doi.org/10.3402/nstep.v1.27007>.
- Horlacher, R., 2015. *The Educated Subject and the German Concept of Bildung: A Comparative Cultural History*. Taylor & Francis.
- Horlacher, R., 2018. The same but different: the German Lehrplan and curriculum. *J. Curric. Stud.* 50, 1–16. <https://doi.org/10.1080/00220272.2017.1307458>.
- Hörmann, B., 2011. Capacities in diversified classrooms. In: Hillen, S.A., Sturm, T., Willbergh, I. (Eds.), *Challenges Facing Contemporary Didactics: Diversity of Students and the Role of New Media in Teaching and Learning*. Waxmann.
- Hörmann, B., 2015. *All Aboard the Good Ship Schooling? Marginalization in an Era of School Accountability*. University of Vienna.
- Karseth, B., Sivesind, K., 2010. Conceptualizing curriculum knowledge within and beyond the national context. *Eur. J. Educ.* 45, 103–120. <https://doi.org/10.1111/j.1465-3435.2009.01418.x>.
- Klafki, W., 1967. In: *Studien zur Bildungstheorie und Didaktik*, 8./9. Aufl. Beltz, Weinheim.
- Klafki, W., 2007. In: *Neue Studien zur Bildungstheorie und Didaktik: zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*, 6. neu ausgestattete Aufl. Beltz Bibliothek. Beltz, Weinheim [u.a.].
- Köller, O., 2008. Bildungsstandards in einem Gesamtsystem der Qualitätssicherung im allgemeinbildenden Schulsystem Deutschlands. In: Klieeme, E., Tippelt, R. (Eds.), *Qualitätssicherung Im Bildungswesen*, (Zeitschrift Für Pädagogik, Beiheft; 53). Beltz, Weinheim, pp. 59–75.
- Künzli, R., 1981. Das Schulfach als Denk- und Handlungsrahmen. *Bildungsforschung Bildungspraxis* 3, 25–31.
- Künzli, R., 1998. The common frame and the places of Didaktik. In: Gundem, B.B., Hopmann, S.T. (Eds.), *Didaktik and/or Curriculum. An International Dialogue*, Education. Peter Lang, New York, pp. 29–45.
- Masschelein, J., Simons, M., 2013. In: *Defence of the School: A Public Issue*. E-ducation, Culture & Society Publishers, Leuven.
- McGill-Franzen, A., Allington, R.L., 1993. Flunk'em or get them classified: the contamination of primary grade Accountability data. *Educ. Res.* 22, 19–22. <https://doi.org/10.2307/1177302>.
- Menck, P., 2006. *Unterricht - Was ist das? Eine Einführung in die Didaktik*. Books on demand, Norderstedt.
- Mølsted, C.E., Karseth, B., 2016. National curricula in Norway and Finland: the role of learning outcomes. *Eur. Educ. Res. J.* 15, 329–344. <https://doi.org/10.1177/1474904116639311>.

- OECD, 2005. The Definition and Selection of Key Competencies. France.
- Oettingen, A. von, 2016. Almen didaktik: mellem normativitet og evidens. Hans Reitzel, København.
- Otto, H.-U., Schrödter, M., 2010. "Kompetenzen" oder "Capabilities" als Grundbegriffe einer kritischen Bildungsforschung und Bildungspolitik? In: Krüger, H.-H., Rabe-Kleberg, U., Kramer, R.-T., Budde, J. (Eds.), *Bildungsungleichheit Revisited: Bildung Und Soziale Ungleichheit Vom Kindergarten Bis Zur Hochschule*. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 163–183. https://doi.org/10.1007/978-3-531-92201-0_9.
- Schleiermacher, F., 1826. *Texte zur Pädagogik : kommentierte Studienausgabe*. Suhrkamp-Taschenbuch Wissenschaft. Suhrkamp, Frankfurt am Main.
- Sivesind, K., Wahlström, N., 2016. Curriculum on the European policy agenda: global transitions and learning outcomes from transnational and national points of view. *Eur. Educ. Res. J.* 15, 271–278. <https://doi.org/10.1177/1474904116647060>.
- Sivesind, K., Afsar, A., Bachmann, K.E., 2016. Transnational policy transfer over three curriculum reforms in Finland: the constructions of conditional and purposive programs (1994–2016). *Eur. Educ. Res. J.* 15, 345–365. <https://doi.org/10.1177/1474904116648175>.
- Steiner-Khamsi, G., 2012. Understanding policy borrowing and lending: building comparative policy studies. In: *Policy Borrowing and Lending in Education*. Routledge, New York, pp. 3–17.
- Sundberg, D., Wahlström, N., 2012. Standards-based curricula in a denationalized conception of education: the case of Sweden. *Eur. Educ. Res. J.* 11, 342–356. <https://doi.org/10.2304/eej.2012.11.3.342>.
- Tröhler, D., 2016. Curriculum history or the educational construction of Europe in the long nineteenth century. *Eur. Educ. Res. J.* 15, 279–297. <https://doi.org/10.1177/1474904116645111>.
- Van Manen, M., 1986. *The Tone of Teaching*. Scholastic Book Services.
- Van Manen, M., 2016. *Pedagogical Tact: Knowing What to do When You Don't Know What to do*. Taylor & Francis.
- Wahlström, N., 2014. The changing role of the state in a denationalized educational policy context. In: Nordin, A., Sundberg, D. (Eds.), *Transnational Policy Flows in European Education: The Making and Governing of Knowledge in the Education Policy Field*, Oxford Studies in Comparative Education. Symposium Books, Oxford, pp. 159–182.
- Westbury, I., 2008. Making curricula: why do states make curricula, and how? In: Conelly, M.F., He, M.F., Phillion, J. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. SAGE Publications, Inc, Thousand Oaks, pp. 45–65.
- Willbergh, I., 2008. *Mimesis, didaktikk og digitale læremidler*. PhD-thesis. Trondheim.
- Willbergh, I., 2015. The problems of "competence" and alternatives from the Scandinavian perspective of Bildung. *J. Curric. Stud.* 47, 334–354. <https://doi.org/10.1080/00220272.2014.1002112>.

Challenges to lived curriculum making: 1990–2020

Cheryl J. Craig, Department of Teaching, Learning, and Culture, Texas A&M University, College Station, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Literature review	340
Boundaries	340
Curriculum	340
Teacher knowledge & knowledge-for-teaching	340
Images and metaphors of teaching	341
Images	341
Metaphors	341
Teacher community	342
Teacher identity	342
Narrative inquiry	342
Lived curriculum making stories	342
Stories of teachers/teacher stories: Benita Dalton 1990–1996	343
Stories of school/school stories: Henry Richards 1998–2004	343
Stories of teacher community/teacher community stories: Laura Curtis 2002–12	344
Story of evaluation/value-added evaluation story: Anna Dean (2006–2012)	345
Image of teaching as human enterprise/image of teaching as big data: Daryl Wilson (2013–2017)	346
Trends and issues on the boundaries where curriculum implementation and curriculum making merge	346
References	348

In the early 1900s, John Dewey detected a perplexing phenomenon among American teachers. Some, he noted, “[knew] how to teach, but they [were] not students of teaching” (Dewey, 1904, p. 215). They were not able to “... act on their own initiative, [even when they were] expected to do so.” “Tak[ing] on situation[s] for themselves” rather than following a “set method or scheme” seemed unlikely (Dewey, 1904, pp. 27–28). To Dewey,

a willingness [prevailed within] the teaching corps to accept without inquiry or criticism any method or device, which seem[ed] to promise good results... [and] teachers... [would] flock to those persons who give them clear-cut and definite instructions as to how to teach this or that (Dewey, 1904, p. 152).

Joseph Schwab, who became a professor the year Dewey's *Experience and Education* (Dewey, 1938) was released, similarly underscored the need for teachers who are able to make minded curriculum decisions. In Schwab's view, these teachers would be “agent[s] of education, not [simply] of subject matter” (Schwab, 1954/1978, p. 128)—with the delivery of subject matter being policymakers' main concern. For Schwab, the only plausible road to educational improvement comes through teachers who have the ability to reflect on their practices alone and together:

... only as the teacher uses the classroom as the occasion and the means to reflect upon education as a whole (ends as well as means), as the laboratory in which to translate reflections into actions and thus to test reflections, actions, and outcomes, against many criteria is he [sic] a good ... teacher (Schwab, 1959/1978, pp. 182–183).

To the observations of Dewey and Schwab, the insights of Philip Jackson will be added. Jackson was among the first to conduct research alongside flesh-and-blood teachers in schools. In *Life in classrooms*, he famously reported that teachers' work does not follow the straightforward trajectory of a bullet, which is what legislative bodies expect to happen. Instead, teachers' work zigs and zags, resembling the “flight of a butterfly” (Jackson, 1968, p. 67). Jackson's groundbreaking research, like that of Dewey and Schwab, draws attention to important conditions and phenomena that do not “scream out at you” in school milieus but “reveal through language [their] presence and meaning” (Eisner, 2007, p. 156). Written in the research tradition of Dewey, Schwab and Jackson, this chapter, building on the scholarship of Connelly and Clandinin (Connelly and Clandinin, 1988; Clandinin and Connelly, 1992, 1995), focuses on what has unfurled on the boundaries over the past 30 years as teachers have negotiated the mandated curriculum in tension with curriculum as it becomes lived. This work, which is culled from my research program, begins with an overarching literature review and a brief discussion of narrative inquiry. Serial interpretation, this chapter's primary research tool, is paid special attention. Five examples of bumping up places between teachers' implementing required curriculum and teachers'

animating required curriculum come next. This chapter ends with a discussion of the themes and trends evident in the exemplars and offers suggestions as to where the field of education might go from here, given the current state of lived curriculum, which alternately is called experienced curriculum (Connelly and Clandinin, 1988).

Literature review

Several important concepts underpin this synopsis of lived curriculum as it appeared in my research program. Concepts setting the backdrop for this three-decade review include (1) boundaries, (2) curriculum, (3) teacher knowledge/knowledge of teaching, (4) images and metaphors of teaching, (5) teacher community, and (6) teacher identity. Brief summaries of each appear next.

Boundaries

A boundary is an imaginary place where one field brushes against an “intellectual territory of another way of thinking” (Clandinin and Connelly, 2000, p. 21). The latter territory may suggest other ways of doing and being than those advocated by the initial field (Boyer, 1990). Not only do these “bumping up places” appear when research studies are enacted, they also happen in flesh-and-blood settings. This is because people’s lives—those who do research and those who are researched—“flow across the boundaries” (Clandinin and Connelly, 2000, p. 115). Hence, all research is “preceded and surrounded by other texts, some of which are evoked [and] some of [which are] not” (Lather, 1991, p. 9). Freema Elbaz-Luwisch (2005), for instance, used the terms, “the story before the story” (p. 358) and “coda” (p. 377), to capture the complexities of the contextual backdrop of her scholarship before and after her study. In this chapter, I explore “edges” in my work and “examine [more fully] the context in which ... [my] research was carried out” (Elbaz-Luwisch, 1997, p. 79).

Curriculum

Curriculum can be defined as a “course of study” (a document) or “a course of life” (Clandinin and Connelly, 1992, p. 393). When curriculum is not a manual (Westbury, 2000) or a cookbook (Oyoo, 2013), but something organic, rigorous and relatable, vital connections can be made to curriculum as a lived experience and the commonplaces of curriculum (Schwab, 1962; Craig and Ross, 2008). The curriculum commonplaces are near-unanimous considerations (Goodson, 2007) to which educators attend. Teacher, learner, subject matter and milieu are equally present in the making of curriculum. Combined, they bound “the curricular” (Connelly and Clandinin, 1988, p. 84). Curriculum making, in Schwab’s view, “res[ts] not on ideal or abstract representation, but on the real thing, on the concrete case, in all its completeness and with all its differences from other concrete cases ...” (Schwab, 1969, p. 11). Lived curriculum focuses on “practice and action, not [others’] theor[ies] and meaning[s]...” (Connelly and Xu, in press, p. 8). It places “people in front of ideas, experience in front of products and judgment in front of ... instructional scripts” (Hlebowitsh, 2012, p. 10). The real thing—the concrete case—is what Ted Aoki (1989/1990) termed the “full-life of curriculum” and what Jean Clandinin and Michael Connelly called the lived curriculum. Within particular classroom spaces, “teachers and students live out a curriculum ... An account of teachers’ and students’ lives over time is the curriculum, although intentionality, objectives [abstractions], and curriculum materials do play a part ...” (Clandinin and Connelly, 1992, p. 365). Hence, when Schwab declared that “a curriculum ought to be known by the persons it produces, as well as other signs and standards” (Schwab, 1983, p. 247), he meant that curriculum, viewed through a “curriculum of lives” (Clandinin and Connelly, 1992) lens, should be judged by how people’s lives are lived, in addition to how curriculum mandates have been satisfied. Distilled to its essence, curriculum should influence the kinds of human beings people become, given that democratic community is foundational to human flourishing as Dewey, Schwab, Jackson, Eisner, Connelly, Clandinin and others (i.e., Greene, 1995) have stressed.

Teacher knowledge & knowledge-for-teaching

Because teachers are the “fountainheads of the curriculum decision” (Schwab, 1983, p. 24), policymakers and researchers approach teachers’ primacy in the teaching-learning process in different ways. Other than teacher proofing (Tanner and Tanner, 1990) where teachers play no role at all, two knower and known relationships exist (Dewey and Bentley, 1949; Fenstermacher, 1994): (1) the knowledge-for-teaching approach, and (2) the teacher knowledge approach.

The knowledge-for-teaching perspective upholds the view that a body of knowledge exists that every teacher should know and do. For proponents of knowledge-for-teaching, such knowledge is determined by others and widely circulated “to build knowledge structures that are said to accumulate” (Eisner, 1997, p. 7). For them, educational improvement means mandating more or better knowledge for teachers to implement. In this view, teachers install others’ plans and the knowledge they deem most important (Clandinin and Connelly, 1992; Craig and Ross, 2008). In this scenario, teachers are “hand-servants” (Hansen, 2011) through whom others’ declarative knowledge passes. In these situations, curriculum is fixed like “a package ... delivered ... by bulk postal mailing” (Rosiek and Clandinin, 2016, p. 296).

Standing in juxtaposition to this view is the teacher knowledge stance, which understands teaching in teachers’ own terms. In this view, teacher knowledge is produced by teachers in the midst of, and as a response to, lived curriculum practices. At the heart of this version of teacher knowledge sits experience. This is because “knowledge arise[s] from experience and return[s] to experience for

validation" (Clandinin and Rosiek, 2007, p. 39). Even knowledge-for-teaching becomes sifted through the lens of experience in the teacher knowledge approach. There is no other way for policy or research imperatives to enter teachers' minds than through their experiences and no other way for them to enter teachers' bodies than through teachers' minds and embodied actions (Craig et al., 2018). Fenstermacher (1994), in his groundbreaking review of practical teacher knowledge programs, named three leading schools of thought: Russell and Munby (1991), following the Schön (1983) line; Connelly and Clandinin (Clandinin, 1986; Connelly and Clandinin, 1985), who with Elbaz-Luwisch (1983, 1991), built on Schwab's (1969) practical tradition; and Cochran-Smith and Lytle (1990, 1993) reflecting the teacher-as-researcher strand (Stenhouse, 1976). While somewhat different, all these research thrusts capture teachers personally and professionally coming to know their practices as minded human beings (Dewey, 1904, 1908/1981, 1938).

Following the Connelly-Clandinin line of research, my body of lived curriculum scholarship views teacher knowledge as "a narrative construct which references the totality of a teacher's personal practical knowledge gained from formal and informal educational experiences" (Xu and Connelly, 2009, p. 221). To Connelly and Clandinin (1988), personal practical knowledge is "... in the person's present mind and body, and in the person's future plans and actions ... It is seen and found in [teachers'] practices" (p. 25). Distilled to its essence, teachers' practices instantiate their personal practical knowledge, given that teachers are curriculum makers (Clandinin and Connelly, 1992; Craig and Ross, 2008), not merely implementers of others' knowledge injections.

Images and metaphors of teaching

Images

The widespread image of teachers-as-curriculum-implementers approaches teachers as agents of the state, charged with institutionalizing its edicts. In this view, shifts in teachers' practices happen because teachers are bound to the laws policymakers establish. In this scenario, the teacher is a "technician, consumer, receiver, transmitter, and implementer of other people's knowledge" (Cochran-Smith and Lytle, 1999, p. 16). This view of teachers being positioned in an educational conduit (Clandinin and Connelly, 1995; Craig, 2002), delivering curriculum like "passing on a bag of goods" (Hansen, 2011, p. 1), has spread internationally because of theory's long-standing estrangement from practice (Schwab, 1969). This hierarchical approach to change is highly problematic because, as emphasized earlier, teachers' practices embody their personal practical knowledge (Craig et al., 2018). This is far more complex and expansive than compliant teaching, which forestalls teachers' voices and agency (Rodgers, 2020b), returning us to Dewey's concerns about automaton teaching at the turn of the century.

Unfortunately, the dominant paradigm ignores an alternate image of teaching, one that recognizes teachers as active agents (Schwab, 1954/1978) and the mindedness Dewey (1908/1981) afforded them. In this view, teachers "*must* be involved in debate, deliberation, and decision about what and how to teach" (Schwab, 1983, p. 245, italics in original). The teacher-as-curriculum-maker construct underscores the preeminence of the teacher in organizing, planning, and orchestrating classroom interactions because only the teacher is positioned at the interstices of the curricular exchange and meets students in moment-to-moment ways. Thus, curriculum is what happens—what becomes animated—in the breath-to-breath instants when teaching and learning fuse. Distilled to the essence, experiential knowledge—practical knowledge—cannot be tested, packaged and delivered because human beings (teachers, students) are inseparable from the body of knowledge that exists (i.e., Dewey and Bentley, 1949; Fenstermacher, 1994) and essential to the curriculum making act (Schwab, 1983). Cut to the quick, the teacher is "the most responsive creator of curriculum" because he/she "negotiates the formal planned curriculum of governments and publishers within his/her practice, alongside the lives of learners" (Murphy and Pushor, 2010, p. 658). Consequently, in the teacher-as-curriculum-maker image, attention turns from written plans, authorized textbooks and government mandates (Clandinin and Connelly, 1992; Connelly and Clandinin, 1988), which typically privilege subject matter and destine teacher to dispensers of knowledge roles, to curriculum as it is lived within the context of people's lives (Downey and Clandinin, 2010). The latter image is an amalgam of Schwab's four curriculum commonplaces (teacher, learner, subject matter, and milieu) with the teacher serving as the intermediary. What becomes animated are the most important aspects of teacher knowledge, which are "ephemeral, passionate, shadowy and significant ... for the most part ... reflect [live of] teachers' lives" (Connelly and Clandinin, 2004, p. 42).

Metaphors

Like all other human beings, teachers employ metaphors to convey their pedagogical understandings to themselves and to carry across ("metepherein"—the original Greek word) their meaning to others. Using metaphors, they capture and communicate knowing of one phenomenon through the aperture of another phenomenon. In narrative research, lived metaphors that drive and reflect human experience are powerful tools for understanding teachers' knowledge and practices. Distinctions, however, exist between the kinds of metaphors that teachers use: (1) novel metaphors and stock metaphors, and (2) emergent and ascribed metaphors. Novel metaphors are imaginative and possess generative power (Schön, 1993). Such metaphors are compact, express the inexpressible and vividly illuminative. Conversely, stock metaphors risk over-use and "sedimentation" (Brown, 1976, p. 175). In short, stock metaphors are trivial and shape practice in non-illuminative ways. Emergent and ascribed represent two of metaphor's other distinctions. Emergent metaphors (Breault, 2006) are intuitively held and naturally surface in novel or stock ways in a teacher's language and unfolding practice. By way of contrast, ascribed metaphors (Breault, 2006) are novel or stock metaphors that researchers intentionally adopt to describe a phenomenon they have identified or teachers' perceived experiences of teaching (Nieto, 2003). Further to this, teachers can individually choose metaphors to live by (Lakoff and Johnson, 1980) as part of their professional growth (Wagenheim et al., 2009). Whether identified by a researcher or by a teacher, ascribed metaphors are choices

made from a vast range of alternatives. Once a metaphor is chosen, the teacher's practice is molded and shaped by the researcher and/or the teacher to fit within the boundaries of the metaphor. The metaphor then becomes the interpretative lens through which the teacher's pedagogical practice is bounded.

Teacher community

Many researchers have offered theories of teacher community (Grossman et al., 2001; Thomas et al. (1998); Wineberg and Grossman, 1998). Others such as Lieberman (1992) have examined the development of the scholarly community and networks (Lieberman, 2000) as reflective forums. Craig (1995b, 2007, Curtis et al., 2013; Craig et al., 2020a,b) and Craig and Olson (2002) further conceptualized the development of teachers' narrative authority in knowledge communities. For them, knowledge communities are "safe, storytelling place where educators narrate the rawness of their experiences, negotiate meaning for such experiences, and authorize their own and others interpretations of situation" (Craig and Olson, 2002, p. 116) as "legitimate tellings" (Mishler, 1986). Teachers' knowledge communities take shape around "commonplaces of experience" (Lane, 1988, p. 188) rather than bureaucratic and hierarchical relationships that determine who knows, what should be known, and what constitutes "good teaching" and "good schools." A further notion of teacher community, which frequently brushes up against knowledge communities, is the idea of professional learning communities (i.e., DuFour, 2003; DuFour and Eaker, 2009), a conceptualization introduced through the organizational literature. In this view of community, administrators preordain the communities in which teachers will interact and determine the topics they will discuss, which is antithetical to what is known about the knowledge communities teachers naturally form.

Teacher identity

Teachers' identities are their "stories to live by" (Connelly and Clandinin, 1999). Their identities are "composed and improvised as [they] live [their] lives embodying knowledge and engaging contexts" (Clandinin and Connelly, 1999, p. 4). Teachers' stories to live by affect "the making sense and reinterpretation of [their] ... values and experiences" (Flores and Day, 2006, p. 220). In addition, emotion is intimately involved in "how ... [their] selves come to be constituted" (Zembylas, 2005, p. 27). In short, teacher identity is the "product of shared story telling" (Sfard and Prusak, 2005, p. 10) and is manifested through teacher voice and agency (Rodgers, 2020b).

Narrative inquiry

Narrative inquiry is the "experiential study of experience" (Xu and Connelly, 2010, p. 354). In this research approach, narrative serves as both the phenomenon of inquiry and the method of investigation (Clandinin and Connelly, 2000; Connelly and Clandinin, 1990). Story is "the portal through which a person enters the world" (Connelly and Clandinin, 2006, p. 477) and a "portal to experience" (Xu and Connelly, 2010, p. 352). It is "not so much a structured answer to a question or a way for accounting for actions and events, as it is a gateway ... for narrative inquiry into meaning and significance" (Xu and Connelly, 2010, p. 356).

Relational in form, narrative inquiry has three commonplaces: temporality, sociality and place/situation. The term, commonplace, connects it to Schwab's (1969) "practical, a language for curriculum" whereas temporality, sociality and place/situation trace narrative inquiry's origins to Dewey's (1938) philosophy of experience. Thus, for a narrative inquiry to be field-based (Xu and Connelly, 2010), all three nodes must be present and interacting in the investigation.

Several field texts provide the narrative grist that morph into research texts. These include personal interviews, focus groups, and material objects of reflection. In addition, journal entries can be written and receive response. To this, participant observation notes from classroom observations, faculty meetings and department meetings, and lesson and unit plan documents can be added. Furthermore, teaching and meeting interactions can be videotaped and even digital narrative inquiries (Craig, 2013) can be prepared in advance of writing and publishing final research texts.

Broadening, burrowing and storying and restorying (Connelly and Clandinin, 1990) are the three main analytical tools that I used to gather the original narrative material having to do with lived curriculum. A fourth narrative inquiry tool, serial interpretation (Schwab, 1954/1978), will now be further elaborated. Serial interpretation is the tool I used to prepare this encyclopedia chapter. It allows one to identify "encompassing idea[s]" that traverse multiple studies. It involves "talking across" (Stone, 1988, p. 2) research projects—even countries—as I will do in this chapter. Serial interpretation adds heft and power to single case studies by bringing them together around a meta-topic, which is lived curriculum where this chapter is concerned.

Lived curriculum making stories

In this section, five stories representing five educators' experiences will be presented. Combined, the exemplars convey critically important lessons gleaned from my 30 years of conducting research in public schools in western Canada and the mid-southern

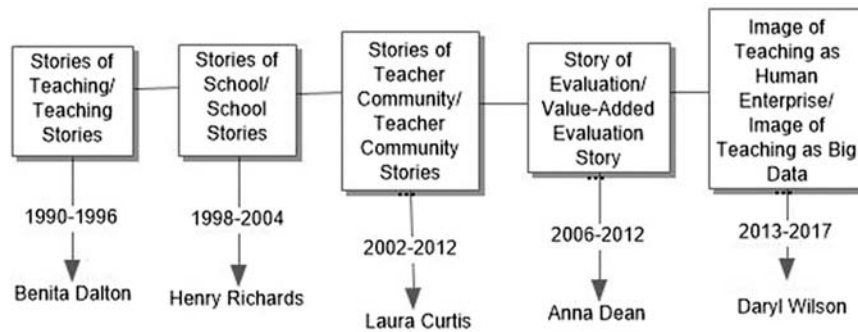


Fig. 1 Timeline of narrative accounts (1990–2020).

US. The first exemplar comes from Canada whereas the four remaining ones occurred in the US. I have included time frames for the specific stories so readers can situate their personal and professional experiences alongside the narrative accounts I share (Fig. 1).

Stories of teachers/teacher stories: Benita Dalton 1990–1996

My early research experiences in Canadian classrooms involved research participant, Benita Dalton, with whom I conducted both my doctoral and post-doctoral research studies. I followed Benita from her preservice teacher education program (Craig, 1995a, 1995b, 1995c), to her practicum experiences, and finally into her beginning teaching positions at three different schools in Western Canada. Benita learned how to team teach in her teacher preparation program and during her extended practicum experiences in the field (Craig, 1998) when she was placed on a six-member teaching team. However, when she entered her first school setting after graduation, her teaching partner was not enamored with co-teaching and physically and metaphorically shut the folding wall between their classrooms. This ended all possibilities of them team teaching from that moment onward. The situation taught me that team teaching partnerships must appeal to both partners. Both teachers—regardless of prevailing philosophies—must be committed to team instruction. At the most basic level, a shared agenda—a commonplace of experience (Lane, 1988)—must exist in order for teachers to overcome personal boundaries and join together in shared instruction.

At her second school, Benita experienced a different team teaching arrangement when she team-taught with two teacher peers. What the three teachers did this time around was to pretend to teach collaboratively without making too much of a time investment. In reality, they lived a cover story (Clandinin and Connelly, 1995; Olson and Craig, 2005). In actuality, they taught in collusion alongside one another, intentionally concealing the subterranean nature of their partnership (Craig, 1998). Benita and her co-partners pseudo-team taught because the school district and the school principal wanted them to co-instruct their multi-aged grades of students. However, the children's parents were adamantly opposed to the campus's administrative and pedagogical approach. The projection of a visible team teaching image became the way the teachers managed tensions they experienced at the boundaries between and among the school, its district and the community it served. This allowed them to persevere without being implicated in the brouhaha over who had the most power in setting the campus's agenda.

Finally, when Benita moved to her third school, she was placed with a new teaching partner and became part of a six-member, Grade 3 team where everyone—the principal, the parents, the teachers and the community—favored team teaching practices. Because Benita shared her developing personal practical knowledge (Clandinin, 1985) with me and allowed me to conduct research alongside her in her vastly different school settings, I learned—in an up-close way—that the contexts of teaching greatly shape what teachers like Benita can do and be, regardless of what they *know*. In short, through conducting research with Benita Dalton, I learned that team teaching is not simply an individual decision because it has to do with everyone else and everything else, most especially teachers' abilities to cross classroom boundaries, grade level boundaries and age level boundaries as well as parental boundaries, school district boundaries and philosophical and political boundaries. Reduced to the essence, teachers' knowing and doing is anchored in place. *Teachers are only able to do what their contexts allow them to do, in spite of their prior knowing and doing in other school milieus as this stories of teachers/teacher stories juxtaposition vividly shows through the experiences of one induction year teacher working in three consecutive school sites.*

Stories of school/school stories: Henry Richards 1998–2004

This second example revolves around the dragon in school backyards, a small story hard to forget (Davis, 2003). The school story, whose origin is folklore, is one told by Henry Richards, a former principal of Eagle High School. His campus, which expanded from 2200 to over 3000 students during the time I conducted research there, "changed almost overnight." (Craig, 2004, p. 1236). It initially was a nationally named Blue Ribbon School with mostly white students. It then became a sports powerhouse serving mostly Black and Latinx youth. Henry told the plotline of his changing school milieu to anyone who would listen: teachers, students, parents, researchers, other principals and members of a \$60 million dollar reform movement of which his campus was a part. The gist of Henry's tale went like this: "When you work in a school and there is a dragon in your school backyard, you

had better prepare for the dragon The dragon, of course, is the accountability system ...” (Craig, 2004, p. 1230). The accountability dragon of which Henry spoke had to do with the high stakes tests administered to students annually. These tests affected students’ life chances as well as the bonuses awarded teachers and principals. These exams became the measuring sticks through which the school district judged its principals, teachers and schools and through which the state, in turn, judged its school districts. Everything was about “measured experience”; very little, on the surface at least, was concerned about “lived experience” (Hlebowitsh, 2012, p. 2).

While digging deeply into Henry’s telling, a second dragon image appeared. Readers will recall that the first dragon was the fire-breathing accountability story of school, which demanded that student test scores meet or exceed the state mean (the dragon outside). However, the second dragon was the school story that Henry and the Eagle teachers called the relevancy dragon. The relevancy dragon was the curricular and teaching changes that the Eagle faculty were pursuing to improve students’ lived curriculum experiences (the dragon inside). Henry explained:

The first method [dragon] is “daily drip” but what we are doing is a much more honest and ethical way than simply teaching to the test ... It will take us a little longer—but we honor curriculum in our classes. We are dealing with the two accountability systems jointly as opposed to separately (Craig, 2004, pp. 1239–1240).

To give lived curriculum making the upper edge, Henry “put a circle of protection around [Eagle’s teachers] so [they] could work with students” (Craig, 2004, p. 1238). As one teacher explained,

Henry totally trusted in the professionalism of teachers. He had faith in us. He would never make us do anything perfunctory. He believed we knew best how to teach the children before us. He saw us as possessing very special knowledge of curriculum ...

(Craig, 2004, p. 1238).

Put differently, Henry Richards purposely created a context where Eagle’s teachers could live curriculum alongside their learners and where both would thrive.

However, when Eagle’s students wrote their end-of-year accountability tests, their scores did not show the dramatic increase that the school district expected. A head-on collision ensued where the two dragons—one of accountability, the other of relevancy—crossed paths. The mishap took the shape of a to-be-expected implementation dip, a phenomenon that “pacesetters and coercers ... have no empathy for ...” (Fullan, 1991, p. 40). Predictably, the school district had deaf ears when this possibility was raised. Henry Richards was called into the district office and told that “... a very prescriptive way is coming down the tubes ... The parameters have been well defined ... The reform programs are under review ... The translation of smart teachers into smart students has not taken place in the manner expected” (Craig, 2004, p. 1249). Shortly thereafter, Henry Richards “resigned” as Eagle’s principal and lead teacher. He moved up the interstate highway to work at another regional high school serving middle-and-upper-class students, a campus that greatly appreciated his insights and leadership acumen.

After Henry Richards departed, the Eagle High School students were administered another set of accountability tests the following year. In mathematics, writing and reading, their scores cumulatively increased by 25% across the board (12.6% in all tests taken), making Eagle High School a recognized campus in the state of Texas. Even more importantly, Eagle closed its achievement gaps between sub-populations (i.e., Black, Latinx), leading all other high schools in Henry’s former majority minority school district (Craig, 2004, p. 1247–1250). *Respecting teachers as curriculum makers and valuing their curriculum making alongside students made Eagle’s principal suspect in the eyes of those advancing the state accountability agenda. After the skirmish bearing serious employment consequences was over, Henry Richards maintained that how he managed the “dragons” at Eagle High School was “the honest, moral and authentic way to go”* (Craig, 2004, p. 1244). *He and the Eagle teachers concurred that “results” cannot be produced on command, especially in a high school the size of a small city.*

Stories of teacher community/teacher community stories: Laura Curtis 2002–12

The literacy teachers at T. P. Yaeger Middle School decided to adopt readers and writers workshop as their preferred teaching method. While all of them favored workshop teaching, they had mixed views concerning how their school’s wholesale adoption of the workshop would be rolled out. About half of the teachers wanted to work according to their own sensibilities; the other half favored the hiring a staff developer who would oversee their workshop implementation. Briefly, one group of teachers trusted their personal practical knowledge developed in their knowledge communities (Craig, 1995a, 2001, 2007; Craig et al., 2020a,b) whereas the other group wanted to privilege a staff developer’s knowledge over their own in professional learning communities. To make matters worse, their principal predetermined the topics around which each professional learning community would focus. As experienced teacher, Laura Curtis, explained:

Laura: The principal was in his Professional Learning Communities (PLC) groove and he wanted us to be a workshop school. Whatever we would be would be school-wide and PLC-related.

Cheryl: Was that every discussed with everybody?

Laura: It would not have been a problem, it was the way it was done by forcing people ... It was the forcing. He wanted us all as a PLC on the same page on the same day ... It has just been hard, very, very hard

(Craig, 2012a, p. 95).

Bent to the dictates of the staff developer, Laura metaphorically felt trapped like a “butterfly under a pin.” Her colleagues’ and her “slavish ... dependence on the specialist” (Schwab, 1954/1978, p. 128) meant that they no longer could make independent classroom decisions. In the past, the literacy teachers had been curriculum makers and “felt in charge of [their] teaching” (Craig, 2012a, p. 90). However, the forced curriculum change meant that everyone—even department chair Daryl Wilson—reported to the staff developer because the principal backed the staff developer and some of the literacy teachers but not the department chair and the remainder of the teachers. Unfortunately, the felt tensions came to a head on a professional development day led by the staff developer. In Laura Curtis’s words,

... it was confrontational ... everybody was confrontational. Our staff developer instigated it. It was a mean and ugly thing, something that should not be happening in a professional setting ... it just intimidated us so badly—the demeanor, the manner, the culture—what was said and how it was said ... I did not like being bullied like that

(Craig, 2012a, p. 97).

When the teachers returned to Yaeger after the weekend meeting, the teachers’ upset went up a notch. The principal stepped in, requiring the teachers to sign loyalty oaths. Laura explained how that felt: “We were under fire and we were muzzled. We had to sign affidavits. We were ordered” (Craig, 2012a, p. 97). As a result, four of 14 members of Yaeger’s literacy faculty left the campus (five member of the math faculty resigned as well) while others remained. Several of the latter took their alternate ways of living readers and writers workshop underground where they continued to tell and re-tell their counternarratives (Craig, 2020a) in the guise of cover stories (Clandinin and Connelly, 1995; Craig and Olson, 2002). This exemplar captures an epic clash at the boundaries between the two dominant teaching images and the versions of teacher community stemming from them: teaching-as-curriculum-implementer and teacher-as-curriculum-maker. *With Yaeger’s principal privileging the staff developer and the teacher-as-curriculum-implementer image, the curriculum making literacy teachers suffered a serious setback. Their identities and natural knowledge communities forcibly became imprinted with images of teaching (i.e., butterfly under a pin) and teacher community (butterflies under pins?) that were not of their making and choosing.*

Story of evaluation/value-added evaluation story: Anna Dean (2006–2012)

For the first five years of her career, Anna Dean participated in the workshop reform in the literacy department alongside Laura Curtis, an experienced teacher, and Daryl Wilson, their department chair. While metaphorically learning to teach in the “eye of a storm” (Craig 2013, p. 25), Anna was well aware of the tensions that existed on Yaeger’s bifurcated literacy faculty. She personally was of the opinion that “teachers should [not] have something forced down their throats ...” and that denying teachers’ ownership was “degrading” (Craig, 2013, p. 26). As a beginning teacher mentored by the literacy assistant principal, she “figure[ed] out that the ‘what’ and the ‘how’ are the teaching points [of] the [workshop method] ... and the connections between them are the ‘why.’” As Anna arrived at these important understandings, another experienced teacher called her “teacher-of-the-universe.” About the same time, Anna Dean announced that she had “nailed workshop.” A synergy had formed between the stories she and others (i.e., Laura Curtis, Daryl Wilson, other teacher) were telling about her teaching. After five years of induction into the workshop method, Anna achieved the highest score possible on her annual teacher evaluation. However, everything was about to be disturbed by a district superintendent change, a principal change, a literacy assistant principal change and only one of the school’s six assistant principals (the PE one) remaining on campus. Additionally, T. P. Yaeger admitted more minority-to-majority student transfers to the campus and the number of policies and procedures funneled into the campus intensified. The long and short of it was that Anna Dean who had earned her experienced teacher stripes over an extended period of time ended up being assessed by an entry-level assistant principal with no literacy background. That new administrator “victimized” Anna in what another teacher called the “value-added quota game” (Craig, 2019, pp. 20–21). The green-behind-the-ears assistant principal deducted 32 points from Anna’s previous evaluation, assigning her 3’s in categories where she previously had been awarded 5’s. Her newness to career may have made her susceptible to this kind of treatment. To Anna, her evaluation results were incommensurate with the quality of her curriculum making alongside Yaeger’s youth. As she quit the teaching profession, Anna reflected backward and said, “Cumulatively, it was too much for a veteran; much less a new teacher, but I had learned how to wade and swim backward in a sea that was slowly pulling me under. I decided to seek other alternatives for my career path ... The time had come to leave ...” (Craig, 2014, p. 107).

This close-up story illustrates how beginning teacher attrition played out for a well-mentored teacher-as-curriculum-maker. It shows how induction-year teachers risk being collateral damage when pernicious school/school district value-added teacher

evaluation formulas supported by state and national accountability policies become implemented. *Once again, the assertion that teachers can only be as good as their teaching milieus allow is reinforced—this time in the US milieu. This is especially the case in urban schools where the contexts of teaching are in constant flux due to changes in superintendents, principals and assistant principals and policies relating to student attendance and teacher evaluation, among other dictates raining down the conduit (Clandinin and Connelly, 1995; Craig, 2002).*

Image of teaching as human enterprise/image of teaching as big data: Daryl Wilson (2013–2017)

As foreshadowed, Daryl Wilson was the T. P. Yaeger's literacy department chair during the span of time when I conducted research on the campus (1997–2017). Over that sweep of time, he survived a multitude of tensions on the edges that Laura Curtis, Anna Dean and others who came before and after chronicled. Each year news form of standardization were introduced to T. P. Yaeger Middle School. For instance, after Anna Dean left, the school district mandated the use of standardized workbooks. This meant that a \$50 one-time-use workbook was purchased for every student in the school despite the literacy teachers using the workshop method, increasingly as curriculum makers. The upshot was that the teachers hid the workbooks in their classrooms. Meanwhile, another new Yaeger principal went along with their cover stories of compliant curriculum implementation. Not once did she unmask their creative rebelliousness. However, everything dramatically changed when the seventh principal in Daryl's career was assigned to Yaeger. That principal reviewed the pay-for-performance allocations received by teachers and department chairs at the school. She noted that some bonuses, while substantial, were less than previous years. Rather than attributing the evidence to Daryl teaching different students in different academic years and yielding different bonuses—or even considering that the overall school population had dramatically changed, the new principal removed Daryl from his department chair position. She also demoted another highly experienced teacher who was the grade level lead literacy teacher. When Daryl shared the sad news of what had happened to him and the other teacher, it seemed like age discrimination played a major role in the principal's theory of action. Concurrently, the administrator appeared to use big data without “[prerequisite] understandings and expertise to [make it] actionable” in defensible ways (Marsh, 2012, p. 3). Daryl Wilson was chagrined over what happened as his entire career had been devoted to advancing student learning and teacher learning through actively living the teacher as curriculum maker image. Although close to retirement, the decision to retire was not solely his own. As he left the profession to which he felt called, Daryl declared:

This is what we have come to ...

There is no other way to explain it ...

Data is [G]od ...

(Craig, 2020, p. 1).

In this experienced teacher attrition story, the privileging of decisions based on big data (even erroneously or simplistically interpreted data) superseded the value accorded teachers as curriculum makers and the knowledge generated in their emergent and ongoing communities of knowing. In the end, test data outweighed the primacy of teachers' and students' lived curriculum making, despite test data arising from instrumentalization of human experience (Muchmore, 2002), which effectively removes all of its humanity and robs it of its tears (Richards, 2020).

Trends and issues on the boundaries where curriculum implementation and curriculum making merge

This serial interpretation of what was learned from five different studies conducted over time involved lived curriculum experiences in three Grade K-6 schools in Calgary, Canada (Benita Dalton), a middle school in inner city Houston (Laura Curtis, Anna Dean, Daryl Wilson) and a high school in a Greater Houston school district (Henry Richards). As the five exemplars were unpacked, several overarching themes surfaced. One was that curriculum mandates are endlessly poured down the conduit to teachers, but teachers' ideas and solutions rarely travel upward to school district officials and state policymakers. All of the exemplars made this point in one way or another. A second theme was that those in charge (policy makers, school district officials) appeared to have only one theory of action in their playbooks: the use of power to dictate to teachers (and principals!) what they should know and do from a compliance—not a growth—mindset (Dweck, 2006). Third, the unpredictability of school and school district contexts is more-than-likely to topple best-laid plans. This was particularly the case in the four American exemplars. Henry Richards, Laura Curtis, Anna Dean and Daryl Wilson were all highly regarded in their schools; all vigorously supported lived curriculum making alongside students. However, the press for accountability played out in unimaginable ways for each of them. Even this researcher who cumulatively walked beside them had no inkling of the fates that would befall them nor the flagrant disregard paid curriculum making, which is widely recognized as the well-spring of lifelong learning (Craig, 2020a). There is little wonder that beginning (Anna Dean)

and experienced (Daryl Wilson) teacher attrition is skyrocketing, given that those with a call to teach (Hansen, 1995; Bullough and Pinnegar, 2009) are the ones who are first to abandon the profession (Yinon and Orland-Barak, 2017). Near the end of my site-based research with Daryl Wilson, his school district had in excess of 80% of its teachers with five or less years of experience and more than 50% of its principals with five or less years of experience (Craig, 2014). The fact that any middle to high performance Houston area teachers and principals would be experiencing difficulties of the proportions that Henry Richards, Laura Curtis, Anna Dean and Daryl Wilson faced is incomprehensible. It may very well be that those with thin career experiences are evaluating those with thick career experiences, which certainly seemed to be the case for Anna Dean and Daryl Wilson and possibly Henry Richards as well.

Two additional things need to be discussed: principals and surveillance. In the first exemplar drawn from my Alberta, Canada research, a principal is part of the educational hierarchy and the teachers are team teaching—pseudo-team-teaching, that is—because they know team teaching is part of their school's story. But the Alberta principal did not actively impose that agenda on them; they were hired from the outset to team teach at that district campus. Moving on to Henry Richards, he informed everyone (teachers included) about the metaphorical accountability and relevance dragons at Eagle High School and the possibility of a fracas if they sparred. Further to this, Henry favored the relevance dragon and refused to hobble the Eagle teachers by imposing the accountability dragon's agenda on them. The second-to-last T. P. Yaeger principal also turned "a blind eye" to 1500 workbooks purchased at a cost of \$7500 that were stacked up (actually "hidden in clear view" [Craig, 2020, p. 9]) in literacy classrooms. But there was one outstanding difference. In the Benita Dalton exemplar in Alberta, the principals and teachers belong to the same union and seek the same professional and classroom/school ideals. In addition, there is no financial gain attached to student performance on high stakes accountability tests in Calgary, Alberta. This is not true in Houston, Texas. Henry Richards made a moral decision to lend his support to curriculum making on his campus although he would have better served his financial gain by siding with district and state officials. Also, it is safe to say that the new T. P. Yaeger principal discussed earlier was aware of the glaring incongruences between the workshop approach to reading and writing—the campus's chosen mode of instruction—and the use of workbooks, the poster child of tradition. She too made a moral decision concerning how to proceed. However, other principals in the shared exemplars acted differently. For example, there was the early principal in Laura Curtis's and Daryl Wilson's stories who standardized the workshop approach for all teachers and students. That principal's move clipped the curriculum making wings of some teachers, dramatically defining the conflict in the school milieu for over a decade to follow. Also, Daryl's seventh principal demoted him and one other lead literacy teacher presumably because their high stakes accountability bonuses had dropped. That principal neither observed their teaching nor witnessed their interactions with other literacy department colleagues. She made her decision before the school year began. This suggests metrics may have been used in "dysfunctional and oppressive way[s]" (Mueller, 2018, p. xvii) by others in the background. Ravitch painted a possible scenario:

With sophisticated tools and capacity to do multivariate analysis, [there is] no need to enter classrooms ... Technical experts [can] evaluate teachers and schools without regard to ... actual lived experience ... what matters most [is] data

(Ravitch, 2010, p. 180).

Three other points need to be discussed. First, Diane Ravitch, mentioned above, was Assistant Secretary of Education during the administration of President George H. W. Bush. She confirmed that there is no other nation in the world except the US that evaluates its teachers on the rise and fall of students' test scores (Ravitch, 2013). Second, she recognized that countries such as Finland, Canada, Japan, and South Korea spend time and resources improving the skills of teachers, not selectively firing them in relation to student test scores (Ravitch, 2013). Third, rewarding teachers on their students' high stakes testing scores and rewarding principals on what teachers are able to accomplish with students as determined by test scores is destructive to human relationships, most especially those involving teaching, learning and supervision. Trusting relationships are nearly impossible to sustain when individuals discover that others are profiting on their labor. In addition, attention and resources can be withheld from some students and teachers in order to invest in others with a better chance of succeeding. Schwab called such orientations and enactments "bad eclectics" because they suffer from "poisoning ... at the source" (Schwab, 1959/1978, p. 171). Their root systems trace to incongruous theories and policies that limit and, in some cases, negate any positive impact they might have (Craig, 2012b).

This brings us to this chapter's conclusion. The takeaway point that the five exemplars drive home is that lived curriculum making is highly challenging and complex in the current educational environment due to the "messiness of school contexts" (Silin and Schwartz, 2003) that greatly constrains teachers' curriculum making at the juncture where it meets curriculum implementation/standardization/accountability. At the same time, it is equally apparent that positive change "depends on what teachers do and think—it is as simple and as complex as that," as Michael Fullan (2001, p. 155) phrased it. Therefore, it is imperative that this Gordian knot of teaching and teacher education (Craig et al., 2019) be untangled. Lastly, the Covid-19 pandemic stands a chance of paradoxically shifting the horizons in its own way and perhaps fueling the curriculum renaissance that Schwab (1969) foreshadowed. The timing seems opportune, given that citizens generally view teachers in a more positive light and parents are singing their praises. Further to this, the need for forward-moving educational change seems universally accepted as a foregone conclusion. Lastly, lived curriculum making accomplished through teachers' embodiments of the teacher-as curriculum-maker image (Craig et al., 2018) resonates with the American constitution whose foundation and principles are likely to be revisited on the brink of the 21st century.

References

- Aoki, T., 1989/1990. Beyond the Half-Life of Curriculum and Pedagogy. One World. Alberta Teachers' Association, Edmonton, Alberta.
- Boyer, E.L., 1990. Scholarship Reconsidered: Priorities of the Professoriate. Princeton University Press, Princeton, NJ.
- Breault, R., 2006. Finding the blue note: a metaphor for the practice of teaching. *J. Educ. Thought* 40 (2), 159–176.
- Brown, R.H., 1976. Social theory as metaphor: on the logic of discovery for the sciences of conduct. *Theor. Soc.* 3 (2), 1–29.
- Bullough Jr., R.V., Pinnegar, S., 2009. The happiness of teaching (as eudaimonia): disciplinary knowledge and the threat of performativity. *Teach. Teach. Theor. Pract.* 15 (2), 241–256.
- Clandinin, D.J., 1985. Personal practical knowledge: a study of teachers' classroom images. *Curric. Inq.* 15 (4), 361–385.
- Clandinin, D.J., 1986. Classroom Practice: Teacher Images in Action. The Falmer Press, Philadelphia, PA.
- Clandinin, D.J., Connelly, F.M., 1992. Teacher as curriculum maker. In: Jackson, P.W. (Ed.), *Handbook of Curriculum*. Macmillan, New York, pp. 363–461.
- Clandinin, D.J., Connelly, F.M., 1995. Teachers' Professional Knowledge Landscapes. Teachers College Press, New York.
- Clandinin, D.J., Connelly, F.M., 2000. Narrative Inquiry: Experience and Story in Qualitative Research. Jossey-Bass.
- Cochran-Smith, M., Lytle, S., 1990. Research on teaching and teacher research: the issues that divide. *Educ. Res.* 19 (2), 2–11.
- Cochran-Smith, M., Lytle, S., 1993. Inside/Outside: Teacher Research and Knowledge. Teachers College Press, New York.
- Cochran-Smith, M., Lytle, S., 1999. Relationships of knowledge and practice: teacher learning in communities. *Rev. Res. Educ.* 24 (1), 249–305.
- Connelly, F.M., Clandinin, D.J., 1985. Personal practical knowledge and the modes of knowing: relevance for teaching and learning. In: Eisner, E. (Ed.), *Learning and Teaching the Ways of Knowing*. University of Chicago Press, Chicago, IL, pp. 174–198.
- Connelly, F.M., Clandinin, D.J., 1988. Teachers as Curriculum Planners: Narratives of Experience. Teachers College Press, New York.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14.
- Connelly, F.M., Clandinin, D.J., 1999. Shaping a Professional Identity: Stories of Educational Practice. Teachers College Press, New York, NY.
- Connelly, F.M., Clandinin, D.J., 2004. Canadian teacher education in transition. In: Cheng, Y.C., Chow, K.W., Magdalena Mok, M.C. (Eds.), *Reform of Teacher Education in the Asia-Pacific in the New Millennium: Trends and Challenges*. Kluwer Academic Publisher, Dordrecht, The Netherlands, pp. 35–43.
- Connelly, F.M., Clandinin, D.J., 2006. Narrative inquiry. In: Green, J., Camilli, G., Elmore, P. (Eds.), *Handbook of Complementary Methods in Educational Research*. American Educational Research Association, Washington, DC, pp. 477–489.
- Clandinin, D.J., Rosiek, J., 2007. Mapping a landscape of narrative inquiry: borderland spaces and tensions. In: Clandinin, D.J. (Ed.), *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage Publications, Thousand Oaks, CA, pp. 35–75.
- Craig, C., 1995a. Dilemmas in crossing the boundaries on the professional knowledge landscape. In: Clandinin, D.J., Connelly, F.M. (Eds.), *Teachers' Professional Knowledge Landscapes*. Teachers College Press, New York.
- Craig, C., 1995b. Knowledge communities: a way of making sense of how beginning teachers come to know. *Curric. Inq.* 25 (2), 151–175.
- Craig, C., 1995c. Safe places in the professional knowledge landscapes. In: Clandinin, D.J., Connelly, F.M. (Eds.), *Teachers' Professional Knowledge Landscapes*. Teachers College Press, New York, NY, pp. 137–141.
- Craig, C., 1998. The influence of context on one teacher's interpretive knowledge of team teaching. *Teach. Teach. Educ.* 14 (4), 371–383.
- Craig, C., 2001. The relationships between and among teachers' narrative knowledge, communities of knowing, and school reform: a case of the monkey's paw. *Curric. Inq.* 31 (3), 303–331.
- Craig, C., 2002. The conduit: a meta-level analysis of lives lived and stories told. *Teach. Teach. Theor. Pract.* 8 (2), 197–221.
- Craig, C., 2004. The dragon in school backyards: the influence of mandated testing on school contexts and educators' narrative knowing. *Teach. Coll. Rec.* 106 (6), 1229–1257.
- Craig, C., 2007. Illuminating qualities of knowledge communities in a portfolio-making context. *Teach. Teach. Theor. Pract.* 13 (6), 617–636.
- Craig, C., 2012a. Butterfly under a pin: an emergent teacher image amid mandated curriculum reform. *J. Educ. Res.* 105 (2), 90–101.
- Craig, C., 2012b. Teacher education restructuring: learning environments and innovative pedagogies. In: Ben-Peretz, M., Kleeman, S., Richenberg, R., Shimoni, S. (Eds.), *Teacher Educators as Members of an Evolving Profession*. Rowman & Littlefield & MOFET Institute, Langham, MD, pp. 133–154.
- Craig, C., 2013. Opportunities and challenges in representing narrative inquiries digitally. *Teach. Coll. Rec.* 115 (4), 1–45.
- Craig, C., 2014. From stories of staying to stories of leaving: a US beginning teacher's experience. *J. Curric. Stud.* 46 (1), 81–115.
- Craig, C., 2019. From starting stories to staying stories to leaving stories: the career experiences of an English as a second Language teacher in the US. *Res. Pap. Educ.* 34 (3), 298–329.
- Craig, C., 2020. "Data is [G]od": the influence of cumulative policy reforms on teachers' knowledge in an urban middle school in the United States. *Teach. Teach. Educ.* 93, 103027.
- Craig, C., Turchi, L., McDonald, D.M. (Eds.), 2020a. Cross-disciplinary, Cross-Institutional Collaboration in Teacher Education: Cases of Learning while Leading. Palgrave Macmillan, New York, NY.
- Craig, C., 2020a. Curriculum Making, Reciprocal Learning and the Best-loved Self. Palgrave-Macmillan, New York, NY.
- Craig, C., Curtis, G.A., Kelley, M., Martindell, P.T., Pérez, M.M., 2020b. Knowledge Communities in Teacher Education: Sustaining Collaborative Work. Palgrave Macmillan, Cham, Switzerland.
- Craig, C., Evans, P., Li, J., Stokes, D., 2019. The Gordian knot of teacher induction: when context trumps teacher preparation and the desire to teach. In: McDonald, D. (Ed.), *Facing Challenges and Complexities in Retention of Novice Teachers*. Information Age Publishing, Charlotte, NC, pp. 185–204.
- Craig, C., Olson, M., 2002. The development of teachers' narrative authority in knowledge communities: a narrative approach to teacher learning. In: Lyons, N., LaBoskey, V. (Eds.), *Narrative Inquiry in Practice: Advancing the Knowledge of Teaching*. Teachers College Press, New York, NY, pp. 115–129.
- Craig, C., Ross, V., 2008. Cultivating teachers as curriculum makers. In: Connelly, F.M., He, M.F., Phillion, J. (Eds.), *Sage Handbook of Curriculum and Instruction*. Sage Publications, Thousand Oaks, CA.
- Craig, C., You, J., Zou, Y., Verma, R., Stokes, D., Evans, P., Curtis, G., 2018. The embodied nature of narrative knowledge: a cross-study analysis of embodied knowledge in teaching, learning, and life. *Teach. Teach. Educ.* 71, 329–340.
- Curtis, G., Reid, D., Kelley, M., Martindell, P.T., Craig, C.J., 2013. Braided lives: multiple ways of knowing, flowing in and out of knowledge communities. *Stud. Teach. Educ.* 9 (2), 175–186.
- Davis Jr., O.L., 2003. Foreword. In: Craig, C.J. (Ed.), *Narrative Inquiries of School Reform: Storied Lives, Storied Landscapes, Storied Metaphors*. Information Age, Greenwich, CT.
- Dewey, J., 1904. The relation of theory to practice in education. In: McMurray, C.A. (Ed.), *The Third NSSE Yearbook*. University of Chicago Press, Chicago, IL.
- Dewey, J., 1908/1981. The practical character of reality. In: McDermott, J. (Ed.), *The Philosophy of John Dewey*. University of Chicago Press, Chicago, pp. 207–222.
- Dewey, J., 1938. Experience and Education. Basic Books, New York, NY.
- Dewey, J., Bentley, A.F., 1949. Knowing and the Known. Beacon Press, Boston.
- Downey, C.A., Clandinin, D.J., 2010. Narrative inquiry as reflective practice: tensions and possibilities. In: Lyons, N. (Ed.), *Handbook of Reflection and Reflective Practice: Mapping a Way of Knowing for Professional Reflective Inquiry*. Springer, New York, p. XXVII (Chapter 19).
- DuFour, R., 2003. Building a professional learning community. *Sch. Adm.* 60 (5), 13–18.
- DuFour, R., Eaker, R., 2009. Professional Learning Communities at Work: Best Practices for Enhancing Student Achievement. Solution Tree Press.
- Dweck, C.S., 2006. Mindset: The New Psychology of Success. Random House Publishing Group, New York.

- Eisner, E.W., 1997. The promise and perils of alternative forms of data representation. *Educ. Res.* 26 (6), 4–10.
- Eisner, E., 2007. Finding phil. In: Hansen, D., Driscoll, M., Arcilla, R., Jackson, P. (Eds.), *A Life in Classrooms: Philip W. Jackson and the Practice of Education*. Teachers College Press, New York, NY, pp. 153–160.
- Elbaz-Luwisch, F., 1983. *Teacher Thinking: A Study of Practical Knowledge*. Routledge, London.
- Elbaz-Luwisch, F., 1991. Teachers as curriculum workers: teacher collegiality in Israel. *Curric. J.* 2 (3), 335–346.
- Elbaz-Luwisch, F., 1997. Narrative research: political issues and implications. *Teach. Teach. Educ.* 13 (1), 75–83.
- Elbaz-Luwisch, F., 2005. *Teachers' Voices: Storytelling & Possibilities*. Information Age, Greenwich, CT.
- Fenstermacher, G., 1994. The knower and the known: the nature of knowledge in research on teaching. *Rev. Res. Teach.* 20, 3–56.
- Flores, M.A., Day, C., 2006. Contexts which shape and reshape new teachers' identities: a multi-perspective study. *Teach. Teach. Educ.* 22 (2), 219–232.
- Fullan, M., 1991. *The New Meaning of Educational Change*. Teachers College Press, Toronto.
- Fullan, M., 2001. *Whole School Reform: Problems and Promises*. Ontario Institute for Studies in Education, University of Toronto. Paper commissioned by the Chicago Community Trust. <https://michaelfullan.ca/wp-content/uploads/2016/06/13396044810.pdf>.
- Greene, M., 1995. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*. Jossey-Bass, San , CA.
- Goodson, I., 2007. All the lonely people: the struggle for private meaning and public purpose in education. *Crit. Stud. Educ.* 48 (1), 131–148.
- Grossman, P., Wineberg, S., Woolworth, S., 2001. Toward a theory of teacher community. *Teach. Coll. Rec.* 103 (6), 942–1012.
- Hansen, D., 1995. *The Call to Teach*. Teachers College Press, New York, NY.
- Hansen, D., 2011. *The Teacher and the World: A Study of Cosmopolitanism as Education*. Routledge, New York, NY.
- Hlebowitsh, P., 2012. Being there: the ontological measure of teaching. *Curric. Teach. Dialog.* 14, 1–12.
- Jackson, P., 1968. *Life in Classrooms*. Holt Rinehart Winston, New York, NY.
- Lakoff, G., Johnson, M., 1980. The metaphorical structure of the human conceptual system. *Cognit. Sci.* 4, 195–208.
- Lane, B., 1988. *Landscapes of the Sacred: Geography and Narrative in American Spirituality*. Paulist Press, New York, N.
- Lather, P., 1991. *Getting Smart: Feminist Research and Pedagogy With/in the Postmodern*. Psychology Press.
- Lieberman, A., 1992. The meaning of scholarly activity and the building of community. *Educ. Res.* 21 (6), 5–12.
- Lieberman, A., 2000. Networks as learning communities: shaping the future of teacher development. *J. Teach. Educ.* 51, 221–227.
- Marsh, J., 2012. *Interventions Promoting Educators' Use of Data: Research Insights and Gaps*.
- Mishler, E.G., 1986. The analysis of interview-narratives. In: Sarbin, T.R. (Ed.), *Narrative Psychology: The Storied Nature of Human Conduct*. Praeger Publishers/Greenwood Publishing Group, pp. 233–255.
- Muchmore, J., 2002. Methods and ethics in a life history study of teacher thinking. *Qual. Rep.* 7 (4), 1–19.
- Mueller, J., 2018. *The Tyranny of Metrics*. Princeton University Press, Princeton, NJ.
- Murphy, M.S., Pushor, D., 2010. Teachers as curriculum planners. In: Kridel, C. (Ed.), *Encyclopedia of Curriculum Studies*, vol. 2. SAGE Publications, Inc, Thousand Oaks, CA, pp. 657–658.
- Nieto, S., 2003. *What Keeps Teachers Going?* Teachers College Press, New York, NY.
- Olson, M., Craig, C., 2005. Uncovering cover stories: tensions and entailments in the development of teacher knowledge. *Curric. Inq.* 35 (2), 161–182.
- Oyoo, S., 2013. Enhancing and sustaining teacher effectiveness as the “Trojan horse” in successful science education in Kenya. In: Craig, C., Meijer, P., Broeckmans, J. (Eds.), *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community*. Emerald Group Publishing Limited, Bingley, UK, pp. 457–477.
- Ravitch, D., 2010. *The Life and Death of the Great American School System*. Basic books, New York, NY.
- Ravitch, D., 2013. *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*. Vintage Books, New York, NY.
- Richards, S., 2020, September 29. Why Our Minds Can't Make Sense of the Enormous Covid-19 Death Toll. <https://www.nationalgeographic.com/science/2020/09/why-minds-brains-cannot-make-sense-coronavirus-enormous-death-toll/>.
- Rodgers, C., 2020b. *The Art of Reflective Teaching: Practicing Presence*. Teachers College Press, New York, NY.
- Rosiek, J., Clandinin, D.J., 2016. Curriculum and Teacher Development, vol. 1. *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, pp. 293–308.
- Russell, T.L., Munby, H., 1991. Reframing: the role of experience in developing teachers' professional knowledge. In: Schön, D.A. (Ed.), *The Reflective Turn*. Teachers College Press, New York, pp. 164–187.
- Schön, D., 1993. Generative metaphor: a perspective of problem-setting in social policy. In: Ortony, A. (Ed.), *Metaphor and Thought*. Cambridge University Press, New York, NY, pp. 137–163.
- Schwab, J., 1969. The practical: a language for curriculum. *Sch. Rev.* 78 (1), 1–23.
- Schwab, J., 1954/1978. Eros and education: a discussion of one aspect of discussion. In: Westbury, I., Wilkof, N. (Eds.), *Science, Curriculum and Liberal Education: Selected Essays*. University of Chicago Press, Chicago, IL.
- Schwab, J., 1959/1978. The 'impossible' role of the teacher in progressive education. In: Westbury, I., Wilkof, N. (Eds.), *Science, Curriculum and Liberal Education: Selected Essays*. University of Chicago Press, Chicago.
- Schwab, J.J., 1962. The teaching of science: the teaching of science as enquiry. In: Schwaband, J.J., Brandwein, O. (Eds.), *The Teaching of Science*. Harvard University Press, Cambridge, MA.
- Schwab, J., 1983. The practical 4: something for curriculum professors to do. *Curric. Inq.* 13 (3), 239–265.
- Schön, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York, NY.
- Sfard, A., Prusak, A., 2005. Telling identities: in search of an analytic tool for investigating learning as a culturally shaped activity. *Educ. Res.* 34 (4), 14–22.
- Silin, J.G., Schwartz, F., 2003. Staying close to the teacher. *Teach. Coll. Rec.* 105 (8), 1586–1605.
- Sterhouse, L., 1976. Case Study as a Basis for Research in a Theoretical Contemporary History of Action. University of East Anglia, UK (Unpublished paper).
- Stone, E., 1988. *Black Sheep and Kissing Cousins: How Our Family Stories Shape Us*. Transaction Publishers, New York, NY.
- Tanner, D., Tanner, L., 1990. *History of the School Curriculum*. Macmillan, New York, NY.
- Thomas, G., Wineburg, S., Grossman, P., Myhre, O., Woolworth, S., 1998. In the company of colleagues: an interim report on the development of a community of teacher learners. *Teach. Teach. Educ.* 14, 21–32.
- Wagenheim, G., Clark, R., Crispo, A., 2009. Metaphorical mirror: reflecting on our personal pursuits to discover and challenge our teaching practice assumptions. *Int. J. Teach. Learn. High. Educ.* 20 (3), 503–509.
- Westbury, I., 2000. Teaching as a reflective practice: what might Didaktik teach curriculum? In: Westbury, I. (Ed.), *Studies in Curriculum Theory. Teaching as Reflective Practice: The German Didaktik Tradition*. Erlbaum, Mahwa, NJ, pp. 15–39.
- Wineberg, S., Grossman, P., 1998. Creating a community of learners among high school teachers. *Phi Delta Kappan* 79 (5), 350.
- Xu, S., Connelly, F.M., 2009. Narrative inquiry for teacher education and development: focus on English as a second language in China. *Teach. Teach. Educ.* 25, 219–227.
- Xu, S., Connelly, M., 2010. Narrative inquiry for school-based research. *Narrat. Inq.* 20 (2), 349–370.
- Yinon, H., Orland-Barak, L., 2017. Career stories of Israeli teachers who left teaching: a salutogenic view of teacher attrition. *Teach. Teach. Theor. Pract.* 23 (8), 914–927.
- Zembylas, M., 2005. *Teaching with Emotion: A Post-modern Enactment*. Information Publishing, Charlotte, NC.

Living schools—reframing education for future possibility

Patrick Howard, Cape Breton University, School of Education and Health, Sydney, NS, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	350
Progressive education	351
Education for sustainable development	352
21st century teaching and learning	352
Living schools and the purpose of education	354
References	355

Introduction

It is increasingly apparent that our collective futures will be one marked by complexity, uncertainty, and vulnerability. The shocks and shifts to our societies from global threats related to pandemic, climate change, persistent inequities, political and religious extremism, and rising authoritarianism erode social cohesion and the robustness of public institutions to counter these pressures. As we enter the third decade of the 21st century, traditional models of economic growth and the underpinnings of a globalized economy are showing serious weaknesses in addressing future challenges. Models of development and the fundamental and long held purpose of education directly linked to the economic imperative are being re-examined.

For decades now, educators have been grappling with the challenge of determining how to transition outdated school systems that were developed for an industrial age to meet more contemporary learning needs. Recommendations for transforming education include developing 21st century competencies (C21, n.d.; Battelle for Kids, n.d.), integrated use of technology, increasing creativity, innovation, and entrepreneurship (Robinson and Aronica, 2015; Wagner, 2012; Zhao, 2012) through the use of New Pedagogies (Fullan and Langworthy, 2014). Economic prosperity and community development figure strongly as a rationale for these recommendations, though individual well-being is also an anticipated outcome (Kelly, 2016; Kelley and Kelley, 2013; Robinson, 2011; Wagner, 2012; Zhao, 2012). The national and international shifts to 21st century learning are due, in large part, to well-organized, well-funded education initiatives launched by non-governmental and not-for-profit organizations. Many of these initiatives were formed with the mandate to shift education priorities in new directions and disrupt education practices initiated over a hundred years ago (Ananiandou and Claro, 2009; Brooks and Holmes, 2014; C21, n.d.; Center for Educational Research and Innovation, n.d; Fullan and Langworthy, 2013, 2014; Battelle for Kids, n.d.).

However, quickly shifting global realities have rendered even these most recent efforts to reframe, reimagine, and transform education for the 21st century clearly wanting. Just a few years ago, proponents of 21st century teaching and learning called for “radical” (Fullan and Langworthy, 2014, p. 4) makeover of education for the perceived challenges and demands of the 21st century. Educators have turned to real-world design projects that require creative, critical problem solving, collaboration, communication, and integration of content knowledge, new learning, and skill development. Large multi-national tech companies and Silicon Valley billionaires have taken special interest in American K–12 education and are investing large sums of money to re-shape the classroom for 21st century learning goals (Singer, 2017a,b; Strauss, 2020).

While 21st century education initiatives are focused on competencies, dispositions, and work habits to prepare students for the future, in almost all cases an explicit discussion of the fundamental purposes and aims of education and what our common future will look like is avoided. Rather, in the discourse of 21st century skills development, the purpose, or aim of education, is implicit. Education is designed to meet functional ends—to develop creative, innovative, life-long learners, to solve the world’s problems, to answer the challenges of shifting workforce demands, and to develop good citizens. Few of these recent endeavors to transform K–12 education have incorporated what it means to educate for a world of increasing complexity, uncertainty and fragility. Instead they essentially aim to ensure that students are prepared for success in a fast-changing world—presumably contributing to the outdated and unsustainable economic patterns that rely on the over consumption of non-renewable resources and environmental degradation.

The purpose of this article is to situate any re-visioning and re-framing of education, classrooms and curriculum around the concepts of life and living. In this way we trace the concepts of a living curriculum in living schools (O’Brien and Howard, 2016, 2020) to its roots. We understand coalescing schooling around an ethos of life and human flourishing as a radical idea that has been cross-pollinated by rich and deep traditions of progressive education going back over a century. As we will see, reframing education and schooling for the future that awaits us is profound, far-reaching, and fundamental. In this article, I will demonstrate that the vision of a lived curriculum is closely related to initiatives characterized as 21st century teaching and learning but with an important difference. In a living school, the curriculum, the approaches to instruction, the physical environment, and the relationships between and among students, teachers, parents, and the community are guided by an ethos that is life-affirming. It reflects a way of being and becoming that roots education in what the philosopher McMurtry (1998) calls a “life-code of value” and situates the school in a living community of relations.

Living schools re-conceptualize education for our times and for future human flourishing. Such schools do not retreat from explicitly articulating a response to the question, “What is education for?” Historically, educational approaches, trends, and initiatives, including 21st century teaching and learning, have largely avoided engaging explicitly with *the* why of education. Commonly stated purposes of education most often align with raising test scores, helping students achieve by becoming better learners, and preparing students for emerging labor markets. These aims of schooling often deflect from deeper questions of the larger purposes of education and how we “navigate plural realities across communities in both the North and South” ([International Commission on the Futures of Education, 2020](#), p. 2). Without doubt larger purposes can be teased out of discussions of 21st century teaching and learning. The result of this effort invariably discovers that the why of education reflects a set of functional outcomes identified with an economic value system that in many ways commodifies the lives of students and “shapes the selves of learners in accordance with what are perceived to be current economic imperatives, rather than, say with what arises out of their sense of their own existence” ([Bonnett, 2009](#), p. 358).

Education cannot be simply framed in terms of economic growth; human flourishing and the accessibility of lives of purpose and dignity must be primary concerns. As we come to terms with the anthropocene as a geological era of human-caused change to the planet the key question before us is: what do we want to become? This is a question to be asked and answered through education ([International Commission on the Futures of Education, 2020](#), p. 1). Before we explore this more fully, let us look ahead by first looking back to discover the roots of a lived curriculum as a concept grounded in the traditions of progressive education. We will then look at education for sustainability and 21st century teaching and learning as important influences on the concept of a living school. Finally, we will examine what sets a living school with a lived curriculum apart from previous approaches by exploring how life values and sustainable well-being represent a profound and radical shift toward the kind of education required now and for the future.

Progressive education

Progressive ideas can be traced to the earliest writings about education. In the Western tradition, and particularly in the United States, pragmatist philosophers Charles Sanders Pierce, William James, and most notably John Dewey influenced progressive educators. [Dewey \(1938/1997\)](#) showed that complexity and change meant that education could not be reduced to formulaic preparation for life; it needed to be life itself, messiness and unpredictability included. Progressive educators looked for ways to realize Dewey’s vision through valuing diversity, building on the interests of students, organizing learning in larger more holistic units, connecting school to the surrounding community, and developing citizenship.

The central tenets of progressive education were formalized in 1919 at an organizational meeting of the newly created Progressive Education Association (PEA) in Washington, DC ([Little, 2013](#)). The founding principles were based on Dewey’s writing and research. Dewey ran an experimental school on the campus of the University of Chicago in the 1890s and into the turn of the 20th century. For Dewey, education connected to the vitality of life and living was the “great thing” ([Dewey, 2009](#); as cited in [Waks, 2013](#), p. 79). Dewey wrote,

We cannot overlook the importance for educational purposes the close and intimate acquaintance got with nature first hand with real things and materials, with the actual processes and their manipulation ... no number of object lessons can afford even a shadow of a substitute for acquaintance with plants and animals of the farm and garden acquired through actual living among them and caring for them.

[Dewey \(2009\)](#), as cited in [Waks \(2013, p. 79\)](#).

The reality of life in nature as the great thing is an enduring theme throughout Dewey’s writings. Progressive schools were community-based and connected to the life of the neighborhoods in which they were located. They promoted child-centered learning in which the unique self of the student was nurtured to meet their full potential. As we will see, the shadow of progressive education looms large in more recent initiatives, such as Education for Sustainable Development ([McKeown, 2002](#); [UNESCO, 2019a,b](#)). In *The School and Society* (1900/2009) Dewey included diagrams and designs based on his work at the Laboratory School in Chicago. The neighborhood, gardens and fields, businesses, and the university surrounded the school and connected the students to the life of the real world. Teaching and learning areas replaced traditional classrooms and the natural environment was connected to these learning areas. Gardens and fields supplied fresh food for the kitchen, another important learning area associated with learning for the occupations. More recent educational approaches, like the emergence of project-based learning (PBL) ([PBLWorks, 2018](#)), are also artifacts of progressive education re-purposed for today. The roots of PBL are found in *The Project Method* written in 1918 by William Heard Kilpatrick. John Dewey, a fellow faculty member at Columbia Teachers College at the time, introduced Kilpatrick to the method.

Innovative, progressive schools flourished in the US, Canada, and elsewhere until the 1950s then they waned and virtually disappeared. There was a resurgence in the 1960s and 1970s with the Open School concept, and indeed some progressive schools continued and are highly successful to this day. However, a conservative turn in the 1950s, rising concern of losing economic ground during the Cold War, and the emergence of a globalized economy combined with powerful ideological influences of neo-liberal and corporate elites largely sounded the death knell for progressive education. Transmissive teaching, traditional instruction, standardized curricula, measurable inputs and outputs in the form of learning outcomes, large-scale testing, and harsh

accountability regimes transformed Western education in the second half of the 20th century. The progressive purposes of education, the development of the child as a human-being-still-becoming were largely transplanted by commodifying the child's self in the service of the economic system.

While no one is advocating for a return to the schools of the 1930s, it is instructive to situate today's 21st century teaching and learning in the narrative of a longer educational story that leads to the concept of the Living School. Dewey and the other architects of progressive education were 19th and early 20th century people. The "why" of education for them was influenced by the historical and cultural realities of that time. The aim of education was to counter the destructive influences of the juggernaut of industrialism that was transforming society, displacing communities and households as the centers of production, and responsible for mass urbanization and global migration into cities. The factory model was powerful in shaping all aspects of life including schooling. The progressives countered these forces with a focus on individual development through experiential, real-world learning that strengthened community, cooperation, civic mindedness, and the democratic ideal.

Early in the 21st century we face different challenges for which our education systems also seem unsuited. Globalized systems, rising levels of inequity between rich and poor, and social, economic, technological, and ecological change are unprecedented, and the disruption is keenly felt at all levels. Again, we look to education for a response. The education initiatives of 21st Century Teaching and Learning (Griffin et al., 2011) and Education for Sustainable Development offer skills and competencies enhanced by emerging digital technologies that will prepare learners for new realities. The past preoccupation with raising test scores through standardized curricula and testing has been deemed a failure for developing the creativity, innovation, problem-solving, and critical thinking required to flourish in this new century (Ravitz et al., 2012). Each initiative signals a re-emergence of progressive ideals, but for very different purposes, and each has influenced the Living Schools concept. Let us begin by looking at Education for Sustainable Development.

Education for sustainable development

If there is one recent conceptual approach to organizing education to which the living school with a lived curriculum can be most directly linked it is Education for Sustainable Development (ESD). The overarching paradigm of sustainable development is based on greater equity between individuals and groups that is realized through values related to social justice, poverty reduction, and systems thinking that understands environmental, economic, and social interests as being inextricably intertwined (McKeown and Nolet, 2013). ESD was born out of the Earth Summit in 1992; Chapter 36 of the Agenda 21 (United Nations, 1993) document presented a vision of the world's education systems educating in ways that would lead to a more sustainable future. ESD, or Education for Sustainability (EfS) has emerged as a progressive approach to teaching and learning that is locally relevant, culturally appropriate, and addresses all realms of sustainability (environment, society, economy). In this regard, EfS is action-oriented and promotes similar competencies as 21st century learning (Council of Ministers of Education Canada, 2012; United Nations Economic Commission for Europe [UNECE], 2012); however, with the explicit vision "to help communities and countries meet their sustainability goals and attend to the well-being of the planet and all its living inhabitants" (McKeown and Nolet, 2013, p. 7). The Sustainable Development Goal 4 encompasses education for sustainability (UNESCO, 2019a,b) and situates education and knowledge as part of the living heritage of humanity in an interdependent world and a common global good (UNESCO, 2016).

ESD approaches are steeped in the tradition of progressive education. The child's experience is central and the importance of building on the experience of the student is recognized. Through dialog, experimentation, and engagement with real-world issues and projects, young people are able to clarify their own and others' worldviews and to recognize alternative frameworks and futures exist. The focus on experiential learning, individualized instruction, deep learning, real-world application, cooperative learning, and creating learning tasks with direct relevance to students' lives mirrors progressive ideas. However, ESD articulates unequivocally a larger purpose for education that Dewey and his contemporaries perhaps would not have been able to envision. "ESD helps develop the capacity for critical reflection and systems and futures thinking, as well as motivate actions that promote sustainable development" (UNECE, 2012, p. 7). ESD is aimed directly at allowing education to prepare the next generation with the skills, knowledge, attitudes, and values to address our most challenging issues. And yet it is in that language there exists what has become the most strident critique of ESD, and that is the term sustainable development. A full critique of EfS is beyond our scope here but suffice to say that the term development and the word sustainability have become problematic. It has been argued widely that these words have been co-opted by powerful corporate and economic interests to ensure status quo values, attitudes, and economic systems largely responsible for the ecological and societal challenges we face are, in effect, sustained. Jickling and Sterling (2017) ask, "Is sustainability the most useful term—the term that stands to create the most effective critical traction—to serve educational interests?" (p. 3). It is a fair question.

21st century teaching and learning

Global efforts are underway to re-conceptualize K–12 education for what is commonly known as 21st century teaching and learning. The international shift to 21st century learning is due, in large part, to numerous reports and white papers published in the last decade and well-organized, well-funded education initiatives launched by non-governmental and not-for-profit organizations. Many of these initiatives were formed in partnership with large corporations with the stated mandate to shift education

priorities to new education goals and disrupt deeply embedded traditional education practices (Ananiandou and Claro, 2009; Brooks and Holmes, 2014; C21, n.d; E21C, 2017; Fullan and Langworthy, 2013, 2014; P21, 2016).

Like ESD, 21st century teaching and learning approaches are progressive. To support the “radical” reformulation of education (Fullan and Langworthy, 2014, p. 4) for the challenges of the 21st century, educators have turned to real-world design projects that require creative, critical problem-solving, collaboration, communication, and integration of content knowledge, new learning, and skill development. The increase in popularity of the Maker Faire Movement (Kelly, 2016), a Do-It-Yourself (DIY) ethos (Knobel and Lankshear, 2010), robotic competitions, hack-a-thons, problem- and project-based learning (PBL), all enhanced by sophisticated information and communication technologies (ICTs), provide identifiable approaches for the application of the creative process. These learning opportunities offer accessible strategies and recursive steps to develop 21st century learning competencies (Stanford University School of Design, 2016).

21st century teaching and learning offers exciting, engaging learning for students and promotes strong relationships between teachers and young people as co-learners. The initiatives represent a moving away from traditional education approaches that are transmissive, controlling, authoritarian, and demanding of compliance in behavior and thought. The focus on creativity and imagination, collaboration, and critical thinking is hopeful, as is the renewed emphasis on the importance of student-teacher relationships. These are positive developments.

The radical change proposed by advocates of 21st century skills adheres to the dictionary definition of the word, a “thorough or extreme change to traditional forms” (Radical, 2013). But the word “radical” means something else. Words can lose their lived meaning through overuse and paying close attention can re-vitalize the language and connect us to the lived experience to which the word was originally tied. The word radical is charged with the political and the subversive. Calling for radical change in education is done for intentional effect. The word came from the Latin “radicalis” and meant “the direct source or sense” (Ayto, 1990, p. 428). It is a word rich in depth and nuance as it is related to roots and rooting. In medieval philosophy, the radical humor was inherent in all plants and animals, its presence being a necessary condition of life. It is this sense of the word radical is invoked regarding education reform; it is the word’s connection to the fundamental, primary, essential condition of life that may help us respond to the question of “radical change in education, for what purpose?” When the goals of education are aligned with an ethos of life and living, the fundamental question of what we are reforming education for can be posed as a counterpoint to the functional, future-oriented, technology-enhanced belief in education for its neo-liberal, economic, career, and problem-solving purposes.

Looking carefully at recent documents that espouse transforming education for the skills of 21st century learning provides a clearer sense of the demands placed on education. The international project New Pedagogies for Deep Learning: A Global Partnership (NPDL), led by Canadian education consultant and scholar Michael Fullan, is well funded and highly organized, and includes partners from multinational corporations like Intel, Microsoft, and Pearson Publishing (NPDL, 2019). NPDL is working with “clusters” of schools in each of seven countries. The goal is to build collective capacity within and across education systems and “to call to action, engage partners in the common purpose to mobilize deep learning, capture and cultivate new pedagogies that advance deep learning, and continuously measure and analyze what is working” (Fullan and Langworthy, 2013, pp. 19–20). The NPDL white paper *Toward a New End: New Pedagogies for Deep Learning* (Fullan and Langworthy, 2013) begins by describing “the crisis” in public education, namely student disengagement and teacher frustration. Fullan and Langworthy (2013) explain the crisis in education is “intrinsic to our societies’ transition to knowledge-based economies and global interdependency, enabled and accelerated by technology” (p. 1). The authors add, “Amid the challenges, we have begun *defining the what*: new goals for learning relevant to this new era. However, the immediate and expansive attempts to *innovate the how*; the process of education—are urgently required” (2013, emphasis in original, p. 1).

While the “what” and the “how” of education are the central concerns of the ambitious global initiative, the “why” is less well developed. Fullan and Langworthy (2013, 2014) attempt to address the “why,” the larger purpose of the global education reform initiative, by explaining that the NPDL partnership,

seeks to renew our goals for education and learning to include skills that prepare learners to be life-long, creative, connected problem solvers and to be happy, healthy individuals who can contribute to the common good in today’s global interdependent world. We need learning systems to encourage youth to develop their own visions about what it means to flourish and equip them with the skills to pursue these visions (2013, p. 2).

On the surface this seems to be perfectly reasonable. Who would not want children to be happy, healthy, flourishing people equipped to be successful in a complex and changing world? Yet, a claim like this becomes problematic when framed in the language of crisis and in an apparent value-neutral way that leaves young people to “develop their own visions” about what it means to flourish and “contribute to the common good.” In this way, the authors avoid the “why” of education for an instrumental, functional view tacitly related to “happiness, health, and flourishing” but the larger vision is never fully articulated.

The virtual absence of a purpose for learning, and a reluctance to articulate that to which the learning points, is problematic (Howard, 2018). A functional bias toward preparation for the knowledge economy, a fear of being left behind in a competitive global race for economic dominance is betrayed in Fullan and Langworthy’s language describing teachers as “human capital” who have “social capital (the relationships with students, parents and community)” and “decisional capital (negotiating the challenging learning experience)” (2013, p. 13). The language of “capital” has symbolic and rhetorical power. It speaks to the tacit

purpose of education reform for 21st century learning—to protect and strengthen the social and economic status quo firmly anchored in a market-based, competitive economic ideology. It is in this sense that the education transformation 21st century learning intends to represent is not the *radical rethinking* of education it positions itself to be.

Living schools and the purpose of education

The pioneers of progressive education working at the end of the 19th and early decades of the 20th century were responding to the social, political, and economic realities of their day. The larger purposes of education, as the progressives envisioned them, were shaped by strong democratic principles and the values of equity, diversity, the importance of individual freedom, agency, and experiential learning for human becoming. The education they worked hard to attain was at cross-purposes with the powerful market-driven values of schools as training grounds and labor pools for a burgeoning industrialized economy. The echoes of the factory model of schooling can still be heard today in the ringing of bells and buzzers to move students en masse; it can be seen in neat rows of individual desks, felt in the asphalt and chain-link of school yards, and understood in fragmented curricula cloaked in the language of fragmented learning outcomes and standardized test scores.

Today we face issues and challenges that would be unimaginable at the turn of the last century. One of most pressing existential questions for education that is generative and generational is how humans as a species will continue to live well, and sustainably on this planet. David Orr (1992) presents succinctly a purpose and vision for education in which the goal is not merely deeper learning and new pedagogies but learning *toward* life, making connections, and being explicit about what do we want to become. Orr (1992) writes of his vision for education:

First, it aims toward the establishment of a community of life that includes future generations, male and female, rich and poor, and the natural world. The essence of community is recognition, indeed celebration, of interdependence between all parts. Its indicators are the requisite sustainability, peace, harmony and justice and participation (p. 138).

Orr is advocating for a vision of education that is progressive but also encompassing, expansive, generous, affirming of all life, inclusive of the more-than-human, and reaches toward a place of deep transformation. The discourse of 21st century learning is one in which there is confidence and power, a techno-optimistic, and future-facing resolve that education, or more aptly, learning, specifically deep learning, will conserve a way of life founded on intensely held cultural beliefs about progress, growth, development, competition, and technology. Education has not been a panacea for the world's ills; as education levels rise so too do the myriad issues related to unsustainable consumption and inequities in the distribution of wealth. Students will only “develop their own visions about what it means to connect and flourish” (Fullan and Langworthy, 2013, p. 2) when educators, researchers, policy-makers, and all stakeholders radically deepen a sense of interconnectedness, earth-centredness, and humility that are both accepting and mindful of the paradoxical relation of the flux, uncertainty, and unpredictability of life.

Living schools with a lived curriculum are created on a shared vision for education that is life-forwarding. It is a vision of education that differentiates between what the philosopher McMurtry (1998, 2002) calls the life-code of value versus the money-code of value. He characterizes the dominant economic system that has fully been absorbed into culture, policy, government, education and our daily understanding of “reality,” as a system that promotes money and the market above all else. McMurtry (1998) describes the dominant market system as a value code “that underlies people’s normal decisions and actions [and] is one that affirms more money revenues as good and rejects less money revenues as bad” (p. 297). He calls this largely unquestioned principle of value and disvalue the *money-code of value*. It is represented simply as.

Money → Commodity for Sale → More Money.

Money is the beginning and the end of the sequence. In opposition to the money code of value, McMurtry (1998, 2002) proposes the life-code of value. The life-code of value is represented by the following sequence.

Life → Means of Life → More Life.

McMurtry defines life as “organic movement, sentience, and feeling, and thought” (1998, p. 298). Means of life is “whatever enables life to be preserved or to extend its vital range on these three planes of being alive” (p. 298) such as clean air, water, nutritious food, shelter, affective interaction, environmental space, and accessible learning conditions. Maintaining these conditions reproduces life; augmenting and deepening them increases life-value. Essentially, it is life-forwarding and life-affirming. The codes of value—money and life—are different in that in the money-code of value money is the beginning of the value sequence and money is the end of the sequence. “More money,” says McMurtry (1998), “and not more life is the regulating objective of thought and action” (p. 299). The more money that returns to the investor regardless of what happens to life, the better the investment.

Life values are beginning to re-assert themselves and it is these values that must guide discussions of 21st century education. McMurtry’s life-code of value offers a theoretical construct for an alternative value system with which most people are already familiar. It manifests itself in people’s lives everyday as the civil commons. The civil commons are what people come together to protect as distinct from the money-code of value logic. Universal health care, parks, green spaces, volunteer service groups, fire protection, clean water, food inspection, good government, a justice system, libraries, and public education are all examples of what we believe are life-based and life-protective. These are, as part of the civil commons, oriented to life values, not money

values. Commodifying any of these or attempts to turn them over to market-based, private, for-profit entities are met with great debate and resistance. These are all life-affirmative forms of capacity building and allow society to protect and enable open access to life goods or the means to life.

These are values built on human agency and enable access for all community members, not just a privileged few, to grow and express themselves. The community may be local, national, or global. In a life-code value, education is a life-good, both public and individual, and is not a commodity. Children and young people are not commodified learners, future knowledge workers, consumers, or human capital. Children are on the path of human “becoming.” They need pedagogical relationships with teachers entrusted to lead them in search of meaningful lives and the beauty of their full potential as flourishing individuals.

Living schools understand themselves to be in the service of life, and a life-code of value speaks to larger purposes that form the ethos of the school. In a living school, teaching and learning take on new meaning and vitality when allied with a life-affirming purpose. 21st century education is inherently experiential and working experientially offers a way to formulate alternative interpretations of reality, interpretations based on our felt contact with the world, that can challenge taken-for-granted historical and cultural beliefs largely shaped by the money-code of value. Education in a living school is interactive and dialogic. Building an experiential sense and involvement in the community of life develops awareness of how we are connected, interdependent, and intertwined, undoing an individualistic, competitive, isolating experience of the world. The experiential pedagogies of 21st century deep learning and progressive education, when aligned with a larger purpose as reflected in the life process, give authority to our experience. The emotional and physical well-being of both teachers and students can be supported through the open sharing of experiences that move toward actions that are *for* the life process, rather than against it—and that work to embrace all of the living world.

Living schools embody an ethos of bringing life to education and education to life. This means that the values of well-being for all (Hopkins, 2013; O’Brien, 2016) influence what is learned, how it is learned, where and when it is learned, and ultimately *why* it is learned. Living Schools demonstrate the inspiring and dynamic transformation that is possible when the foundational purpose of education is questioned, and the answer is that it ought to lead to well-being for all sustainably. This is both a radical reformulation of education as well as one that draws on longstanding views that education must be experiential (Dewey, 1997, p. 39), that respects holistic world views at the heart of many Indigenous cultures (Deer and Falkenberg, 2016), and that recognizes new pedagogies also include wisdom traditions associated with mindfulness and land-based learning. Living Schools bring a coherent perspective to the diverse and sometimes divisive efforts to reimagine education. These schools take us beyond a collection of social justice or environmental projects, which are laudable, but are not propelling a timely transformation of education.

The focus on well-being ensures that living schools support outdoor learning, social-emotional learning (Collaborative for Academic, Social, and Emotional Learning, 2012), positive education, and Health Promoting Schools where staff and students flourish (Joint Consortium for School Health, 2008). The presentation of a complete theoretical rationale for living schools and an explication and illustration of the attributes and practices indicative a living schools framework are beyond the scope of this article. However, this work can be found elsewhere as practical, professional research and resources (O’Brien and Howard, 2020; Howard and O’Brien, 2017; O’Brien and Howard, 2016).

Education in the service of life has, as its fundamental goal, the life-forwarding steps that emerge from making authentic contact with the world. This re-visioning and reframing of education takes advantage of the creativity inherent in the life process, the uncertainty, the new meanings, and the possibilities that arise with the unfolding of experience. Organizing schools with a lived curriculum based on life-forwarding principles grounded in the world - the social world, the human-built world, and the natural world—can foster contact and dialog *with* the world and, in so doing, schools become living schools that organize learning in ways to enhance *well-being for all*, sustainably. It is in this way that schooling re-conceptualizes education for our times and for future human flourishing.

References

- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries (OECD Education Working Papers, 41). Organization for Economic Co-operation and Development. <https://doi.org/10.1787/218525261154>.
- Ayto, J., 1990. Bloomsbury Dictionary of Word Origins. Bloomsbury.
- Battelle for Kids, (n.d.). P21 Partnership for 21st Century Learning: A Network of Battelle for Kids. <http://www.battelleforkids.org/networks/p21>.
- Bonnett, M., 2009. Education and selfhood: a phenomenological investigation. *J. Philos. Educ.* 43 (3), 357–370.
- Brooks, M., Holmes, B., 2014. Equinox Blueprint: Learning 2030. Waterloo Global Science Initiative. <http://www.wgsi.org/sites/wgsi-live.pi.local/files/Learning%202030%20Equinox%20Blueprint.pdf>.
- C21, (n.d.). Canadians for 21st Century Learning and Innovation. www.c21canada.org.
- Centre for Educational Research and Innovation, (n.d.). Centre for Educational Research and Education (CERI). Organization for Economic Co-operation and Development. <http://www.oecd.org/edu/ceri/>.
- Collaborative for Academic, Social, and Emotional Learning, 2012. Effective Social and Emotional Learning Programs: Preschool and Elementary School Edition. <http://casel.org/wp-content/uploads/2016/01/2013-casel-guide-1.pdf>.
- Council of Ministers of Education, Canada, 2012. Education for Sustainable Development in Canadian Faculties of Education. www.cmec.ca/publications/lists/publications/.../esd_dean_reporten.pdfCMEC.
- Deer, F., Falkenberg, T. (Eds.), 2016. Indigenous Perspectives on Education for Well-Being in Canada. Education for Sustainable Well-Being Press.
- Dewey, J., 1997. Education and Experience. Simon & Schuster (Original work published 1938).
- Dewey, J., 2009. The School and Society. Createspace Publishing (Original work published 1900).
- E21C, 2017. Education for the 21st Century. <http://www.e21c.co.uk/>.
- Fullan, M., Langworthy, M., 2013. Towards a New End: New Pedagogies for Deep Learning. Collaborative Impact. http://www.newpedagogies.nl/images/towards_a_new_end.pdf.

- Fullan, M., Langworthy, M., 2014. A Rich Seam: How New Pedagogies Find Deep Learning. http://www.michaelfullan.ca/wpcontent/uploads/2014/01/3897.Rich_Seam_web.pdf.
- Griffin, P., McGaw, B., Care, E., 2011. Assessment and Teaching of 21st Century Skills. <https://doi.org/10.1007/978-94-007-2324-5>.
- Hopkins, C., 2013. Educating for sustainability: an emerging purpose of education. *Kappa Delta Pi Rec.* 49, 122–125.
- Howard, P., O'Brien, C., 2017. Attributes and Practices of a Living School. *Sustainable Happiness*. <http://sustainablehappiness.ca/sh-extra/attributes-of-a-living-school/>.
- Howard, P., 2018. Twenty-first century learning as a radical re-thinking of education in the service of life. *Educ. Sci.* 8 (4), 189. <https://doi.org/10.3390/educsci8040189>.
- International Commission on the Futures of Education, 2020. Visioning and Framing the Futures of Education. Available. <https://en.unesco.org/futuresofeducation/sites/default/files/2020-03/Outcome%20Document%20First%20Meeting%20of%20International%20Commission%20on%20the%20Futures%20of%20Education.pdf>.
- Jickling, R., Sterling, S. (Eds.), 2017. Post-sustainability and Environmental Education: Remaking Education for the Future. Palgrave MacMillan.
- Joint Consortium for School Health, 2008. Investigating Positive Psychology Themes in School Health. <http://www.jcsh-cces.ca/index.php/resources>.
- Kelley, T., Kelley, D., 2013. *Creative Confidence: Unleashing the Creative Potential Within Us All*. Crown Business.
- Kelly, R., 2016. *Creative Development: Transforming Education Through Design Thinking, Innovation, and Invention*. Brush Education.
- Kilpatrick, W.H., 1918. *The Project Method: The Use of Purposeful Act in the Education Process*. Kessinger Publishing.
- Knobel, M., Lankshear, C. (Eds.), 2010. *DIY Media: Creating, Sharing and Learning with New Technologies*. Peter Lang.
- Little, T., 2013. 21st century learning and progressive education: an intersection. *Int. J. Prog. Educ.* 9 (1), 84–96.
- McKeown, R., Nolet, V., 2013. Schooling for sustainable development in Canada and the United States. In: McKeown, R., Nolet, V. (Eds.), *Schooling for Sustainable Development in Canada and the United States*. Springer, pp. 3–22.
- McKeown, R., 2002. Education for Sustainable Development Toolkit. <https://unesdoc.unesco.org/ark:/48223/pf0000152453>.
- McMurtry, J., 1998. *Unequal Freedoms: The Global Market as an Ethical System*. Garamond Press.
- McMurtry, J., 2002. *Value Wars: The Global Market Versus the Life Economy*. Pluto Press.
- NPDL, 2019. New Pedagogies for Deep Learning: A Global Partnership. Available. <https://www.npdl.global/>.
- O'Brien, C., Howard, P., 2016. The living school: the emergence of a transformative sustainability education paradigm. *J. Educ. Sustain. Dev.* 10 (1), 115–130.
- O'Brien, C., Howard, P., 2020. *Living Schools: Transforming Education. Education for Sustainable Well-being Press*, Winnipeg, MB. <https://www.eswb-press.org/publications.html>.
- O'Brien, C., 2016. *Education for Sustainable Happiness and Well-Being*. Routledge.
- Orr, D., 1992. *Ecological Literacy: Education and the Transition to a Postmodern World*. SUNY Press.
- P21, 2016. Partnership for 21st Century Learning. www.P21.org.
- PBL Works, 2018. Gold standard PBL: Project Based Teaching Practices. https://my.pblworks.org/resource/document/gold_standard_pbl_project_based_teaching_practices1.
- Radical, 2013. Oxford English Dictionary. Oxford University Press.
- Ravitz, J., Hixson, N., English, M., Mergendoller, J., 2012. Using Project Based Learning to Teach 21st Century Skills: Findings from a Statewide Initiative. Presented at Annual Meeting of the American Educational Research Association, Vancouver, 2012. https://www.academia.edu/1854322/Using_project_based_learning_to_teach_21st_century_skills_Findings_from_a_statewide_initiative.
- Robinson, K., Aronica, L., 2015. *Creative Schools*. Viking.
- Robinson, K., 2011. *Out of Our Minds: Learning to be Creative*. Capstone.
- Singer, N., 2017a. How Google took over the classroom. *N. Y. Times*. <https://www.nytimes.com/2017/05/13/technology/google-education-chromebooks-schools.html?action=click&contentCollection=Technology&module=RelatedCoverage®ion=Marginalia&pgtype=article>.
- Singer, N., 2017b. The Silicon Valley billionaires remaking America's schools. *N. Y. Times*. https://www.nytimes.com/2017/06/06/technology/tech-billionaires-education-zuckerberg-facebook-hastings.html?_r=0.
- Stanford University School of Design, 2016. K12 Lab Network Resource Guide. Retrieved March 5th 2017. <https://dschool.stanford.edu/resources/k12-lab-network-resource-guide>.
- Strauss, V., 2020. Cuomo questions why school buildings still exist—and says New York will work with Bill Gates to “reimagine school”. *Wash. Post*. Available. <https://www.washingtonpost.com/education/2020/05/06/cuomo-questions-why-school-buildings-still-exist-says-new-york-will-work-with-bill-gates-reimagine-education/>.
- United Nations Economic Commission for Europe, 2012. Learning for the Future: Competences in Education for Sustainable Development.
- United Nations Education Science Culture Organization, 2016. Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4: Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning Opportunities for All. <https://unesdoc.unesco.org/ark:/48223/pf0000245656>.
- United Nations Education Science Culture Organization, 2019a. Leading SDG 4: Education 2030. <https://unesdoc.unesco.org/ark:/48223/pf0000245656>.
- United Nations Education Science Culture Organization, 2019b. Education for Sustainable Development. <https://en.unesco.org/themes/education-sustainable-development>.
- United Nations, 1993. Agenda 21: Earth Summit—The United Nations Program of Action from Rio. <https://sustainabledevelopment.un.org/outcomedocuments/agenda21>.
- Wagner, T., 2012. *Creating Innovators: The Making of Young People Who Will Change the World*. Scribner.
- Waks, L., 2013. John Dewey and the challenge of progressive education. *Int. J. Prog. Educ.* 9 (1), 73–83.
- Zhao, Y., 2012. *World Class Learners: Educating Creative and Entrepreneurial Students*. Corwin.

Teachers living the curriculum: culturally relevant pedagogy and action research

Noah Borrero^a and Jaqueline Naidoo^b, ^a School of Education, Urban Education and Social Justice, University of San Francisco, San Francisco, CA, United States; and ^b University of KwaZulu-Natal, School of Education, Teacher Development Studies, Scottsville, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	357
Theoretical framework	358
Education for liberation	358
Culturally relevant pedagogy	358
Action research	359
Teachers living their curriculum	359
Notes to Mr. G.	359
Transforming understanding of electrical drawings using practical tasks	360
Discussion	361
Implications	362
References	362

Introduction

The world of education is changing before our very eyes. The impacts of the COVID-19 pandemic around the globe continue to create monumental shifts in the ways that education is delivered and envisioned. In our contexts—urban United States and rural South Africa—impacts of the COVID-19 pandemic are exacerbating structural inequities in our educational systems. That is, access to education is becoming increasingly challenging for communities that are the most vulnerable. To be clear, such inequities existed long before COVID-19. For example, both in South Africa the United States, teachers work in the contexts of shifting ethnic demographics of students. This significantly influences classroom teaching as teachers work alongside growing heterogeneous student populations with different ethnic, racial, social class, cultural, and linguistic backgrounds. In 2020, and with the onset of the COVID-19 pandemic, discrepancies associated with issues like access to remote learning, shelter in place, and safe learning contexts are under a new, intensified spotlight.

A central figure in these tumultuous educational times is the teacher. In many places, teachers are literally on the front lines—facing the challenges of a sea change in the ways that teaching and learning are created. In our work as teacher educators, we honor and acknowledge the tremendous work of teachers and place great value in the ways that their lived experiences shape the curriculum they deliver and impact the students they serve. In this article, it is our hope to portray ways in which K-12 and college teachers live their curriculum through critical pedagogy and action research.

Neither of these approaches—critical pedagogy and action research—are novel concepts in teacher education. Yet, we feel that when they are combined and infused within teacher training and support programs, they provide an important framework for teachers to generate meaningful curriculum, study their own pedagogy, and continually refine their practice in relevant ways. We believe that it is this type of praxis (Freire, 1970) that can foster the ingenuity, passion, and transformation that teachers bring to their work. We also feel that through highlighting the work of teachers in an article and a volume like this, we center their role in the changes before us in the field of education. Too often, teachers' voices are silenced in the narrative about educational reform, and too often teachers' experiences in classrooms are minimized to routines and strategies (Pour-Khorshid, 2020). We know—from our work as and with teachers for the past twenty-five years—that teaching is far more than a practice. It is a lifelong process rife with intellectual rigor and tenacity. By highlighting the ways in which teachers live their curriculum, we hope to honor the complexity of the work they do and their positive impact on young people and our society at large.

We hope to showcase the deep, reflective, and iterative work of teachers through presenting a conceptual framework that focuses on teaching for social justice (e.g., Nieto, 2002) through culturally relevant pedagogy (e.g., Ladson-Billings, 2016). In presenting this framework, we intend to explore some of the true purposes of education and the ways in which classroom teachers are the conduits for change in local communities (Freire, 1970). We then apply aspects of action research as a pedagogical and scholarly process that teachers engage in alongside their students to assess, reflect upon, and enhance their teaching. In the second half of the article, we present two case studies—from each of our local teaching contexts—to showcase examples of teachers enacting culturally relevant pedagogy and action research with young people. Our goals in describing and discussing these examples are to share the work of the inspiring teachers with whom we work in hopes of sparking dialogue, critique, and further innovation in our own and in other teachers' classrooms.

Theoretical framework

Education for liberation

As professionals, we take great pride in the roles we get to play in the educational system. We deeply believe in the transformative possibilities of teaching and learning, and we feel blessed to work alongside teachers and youth who will be the leaders of the future. We also hold strong critiques of the educational systems in our respective countries and communities because it is undeniable that schooling as we know it reflects and reproduces structural inequities in our society (Kumashiro, 2020; Rist, 1970). Formal education has been and continues to be a forceful tool used by those in power to promote assimilation and conformity (Emdin, 2016; Woodson, 2000). We do not feel that this critique dampens our faith in the value of education; instead, we believe that this type of critical social analysis is important in our systems as a way to envision and foster change.

We seek to build upon foundational theories and conceptualizations of education as a human right and a vehicle for social transformation within communities (e.g., Foucault, 1977; Giroux, 1992; Zinn, 2003). Our educational systems—and the very ideas of students, teachers, and classrooms—are part of a larger vision of meaningful learning, self-reflection, and action within communities for the purposes of change for and by those in the community. That is, we seek to follow Freire (2000) call for a form of education for liberation:

Education either functions as an instrument which is used to facilitate integration of the younger generation into the logic of the present system and bring about conformity or it becomes the practice of freedom, the means by which men and women deal critically and creatively with reality and discover how to participate in the transformation of their world (p. 34).

Heeding this call requires continual self-interrogation, a belief in the agency of teachers and students (including the co-construction and fluidity of these roles), and true commitment to a life-long learning process that centers critical consciousness and dialogue (Freire, 2000).

As teacher educators, we are continually learning from our students, who are continually learning from their students. We are part of a cycle of teaching and learning that must strive to stay culturally relevant amidst constantly evolving cultural contexts. As scholars, we feel that it is our responsibility to document and disseminate these learnings and teachings across academic settings and disciplines. We are keenly aware of many of the tensions and challenges inherent in this work. For example, as teachers of teachers, we do not fashion ourselves as experts. We do not proclaim to have answers for fixing the structural inequities in our schooling systems. Instead, in this article, we are attempting to present a framework and portray the work of classroom teachers as parts of a larger vision for meaningful changes in education and in society.

Culturally relevant pedagogy

As teacher educators, we are acutely aware of the importance of adopting and adapting culturally relevant pedagogies. We believe this develops our own and our students' knowledge, attitudes, and skills which contribute not only to academic and social empowerment but also to emotional and political empowerment. We also feel that embracing culturally relevant pedagogy allows us to teach effectively in classrooms that are culturally diverse. As teacher educators, we seek to bridge the cultural divide, and therefore integrate knowledge of our students' cultural backgrounds in our teaching. That is, we strive to adopt Ladson-Billings (1995) notion of a culturally relevant pedagogy:

... that not only addresses student achievement but also helps students to accept and affirm their cultural identity while developing critical perspectives that challenge inequities that schools (and other institutions) perpetuate (p. 469).

Like Gay (2000), our culturally relevant and responsive pedagogy takes into account "the cultural knowledge, prior experiences, frames of reference, and performance styles of ethnically diverse students to make learning more relevant to and effective" (p. 29). Our teaching is underpinned by Ladson-Billings (2014) three principles of culturally relevant pedagogy: academic achievement, socio-political consciousness and cultural competence. We agree with Ladson-Billings (2014) that culture is fluid and that we should take "previously developed ideas and synthesize them to create new and exciting forms" (p. 76). We believe that it is important for teachers to consider teacher knowledge and reflection when they design and implement culturally responsive lessons. In accord with Howard (2003), we argue that critical reflection about culture and race of teachers and students is essential in order to develop teaching strategies that are culturally relevant. Over the years, we have attempted to develop culturally relevant pedagogy and have tailored our teaching to the textual, social, cultural, and personal dimensions of our students' lives (Hefflin, 2002). We believe that culturally relevant pedagogy should have the following outcomes: empowerment, incorporation of culture in curriculum and teaching, and meaningful understanding of socio-political realities in society and how society works (Milner, 2011).

It is our hope that our teaching of a culturally relevant curriculum, will inspire and encourage our students to revise their classroom teaching to be more culturally relevant. We envision that a culturally relevant curriculum will serve as a vehicle to transform society to be more just and equitable. And in the context of today's rapidly-changing educational terrain, we feel that our own and

our students' abilities to strive for cultural relevance in teaching requires deep personal reflection, continual pursuit of enhanced critical consciousness, and active dialogue.

Action research

The last twenty years has seen exponential growth in interest in action research as more teacher educators and teachers engage in action research in various forms—including university credit-bearing courses and modules, professional development, and plans or projects to restructure schools and curriculum. As university teachers, we believe that action research forms an integral aspect of teacher education and is a key driver in educational change and professional development. According to Carr and Kemmis (1986), action research is:

... a form of self-reflective enquiry undertaken by participants in social situations in order to improve the rationality and justice of their own practices, their understanding of these practices, and the situations in which their practices are carried out (p. 162).

We are of the view that action research has the potential to bring about change in specific university, school and classroom contexts. That is, we are guided by Mills' (cited in Hagevik et al., 2012) notion of action research as:

any systematic inquiry conducted by teacher researchers, principals, school counsellors, or other stakeholders in teaching or learning to gather information about how their particular schools operate, how they teach and how well their students learn (p. 675).

As educational professionals, we believe that the principles and processes of action research provide us with information to improve facets of our day to day teaching practice. We heed the call of Kemmis (2006) and adopt a critical approach in action research. In adopting this critical approach to action research, we seek to foster in our students a critical edge aimed at transforming their social contexts and lived experience instead of accepting conventional ways of schooling and society. That is, we encourage our students, who are teachers themselves, to study their social situations and classroom contexts in order to improve their understanding of their practice as well as the quality of their teaching practice. We showcase how teachers engage with action research in their classrooms with their students and argue that this serves as a powerful teacher professional development activity to empower teachers and transform their practices.

Teachers in our case studies participated in the cyclical process of action research which comprised the steps of planning, acting, and observing, with reflection in all the steps (Macintyre, 2016). Our students were also acutely aware that action research refers to repetitive cycles of reflection comprising the following steps: planning a change, acting, and observing the process and consequences of the change, reflecting on these processes and consequences and then replanning, acting, and observing, reflecting and then repeating the process (Kemmis and McTaggart, cited in Koshy et al., 2011, p. 5). We believe that action research helps us and our students to find solutions to the challenges we experience in our teaching and brings about a positive change. That is, we support Koshy et al. (2011) assertion that action research involves problem solving, especially if the resolutions lead to improvements of practice. As professionals, our students conducted action research within the confines of their classrooms, to produce context-specific data about teaching and learning. It is our hope that through action research, both teacher educators and teachers are better able to adopt culturally relevant pedagogy and understand how it contributes to the lived curriculum in a globalized world.

Teachers living their curriculum

Below, we present two different examples of the work of teachers engaged in action research from our own cultural contexts. These cases are not intended to make generalizations about teachers, communities, or pedagogies but to provide examples of ways that teachers with whom we work develop and reflect upon their own practices. Additionally, we do not present these cases to portray a monolithic approach to culturally relevant pedagogy or action research. It is our hope that these examples can create dialogue and offer ideas for other teachers and educators.

Notes to Mr. G.

Mr. G. is a Fourth-Grade teacher in public elementary school in a large, urban school district in California. He is in second year of teaching and he lives and works in the community neighboring the school. In fact, it is the community where he grew up. Mr. G. identifies as Asian-American as do the vast majority of his students (nearly 80%). He also has students who identify as mixed-race, Latina/o, African American, and White. Mr. G. is passionate about building meaningful relationships with his students and cares deeply about fostering their unique interests and needs. Mr. G. reflects that much of this passion comes from his memories of his own schooling. He feels that there was so much pressure from his teachers and his parents to learn English quickly, become

a strong reader, get good grades, and excel on tests, that he never really thought about school as a place that he felt comfortable being himself.

To pursue this goal of relationship-building as a foundation of his teaching, Mr. G. wrote his Master's thesis on an action research project that he developed called "Notes to Mr. G." This project is an ongoing process for Mr. G.—he created it as a student teacher and continues to refine it during his first few years of full-time teaching. The premise of "Notes to Mr. G." is simple—it is an interactive journal that his students keep throughout the school year. At the beginning of the year, students design their own journal and Mr. G. tells them that their journal is theirs to keep and that it will not be graded; it is a tool to express themselves and work on their writing. Students design the covers of their journals to make them unique and they generally write 2–3 entries per week during the school day. On Fridays, Mr. G. takes the journals home with him, reads some of the students' entries, and writes back to them in the journal. The prompts for student writing range from free writing time (students get to choose to write about whatever they like) to fun/creative topics (e.g., "If you could have any animal as a pet, what would you choose, and why?") to school/classroom topics (e.g., "Do you think students should be able to bring phones to school? Why/Why not?") to current events (e.g., "What does the Black Lives Matter movement mean to you?"). Mr. G.'s written responses are not focused on technical aspects of students' writing, but rather are conversational—often asking students questions as follow up or sharing his own interests about a given topic.

As a process, "Notes to Mr. G." is an iterative, informal form of communication between student and teacher. As an action research project in which Mr. G. studies students' journals, his written responses, student learning, and his own learning, he is actively engaged in praxis through authentic dialogue—students come to learn about and trust Mr. G. just as he learns about and builds relationships with them. He then actively shapes his pedagogy and his curriculum through what he learns from the journals and his experiences with students in class. After studying one full year of "Notes to Mr. G.", he realized that students were opening up with him more and more. He was learning about parts of their identities that typically were not a part of classroom learning. For example, he learned about their passions for gaming, drawing, dancing, and wrestling through journal interactions. He also learned that they wanted to know more about him—they would often respond to his written questions with questions of their own. He learned that he was being more open with his student than any of his K-12 teachers were with him. He learned, even, about working through meaningful discussions with students (through the journals and through in-class conversations) about setting boundaries and respecting different people's privacy and individuality. As a representation of living and learning through the curriculum, this was an example of an unforeseen challenges that really pushed Mr. G. and his students to have difficult, uncomfortable conversations—ones that ultimately helped the class and teacher care more deeply for one another.

With the onset of COVID-19 and distance learning, Mr. G. has attempted to adapt "Notes to Mr. G." to a digital format, and while it has lost much of its personality through losing the written word, the core of the project remains—teacher and student connecting through writing in non-threatening, non-academic, non-evaluative ways. Clearly, the complexities of "Notes to Mr. G." as a teaching strategy and the study of it as an action research project extend far beyond what is written here, but the core concept is of value in considering ways in which teachers can live their curriculum.

Transforming understanding of electrical drawings using practical tasks

Mr. M. teaches electrical engineering at a college in an urban district in KwaZulu-Natal, South Africa. He has been teaching for seven years and he lives in the community neighboring the college. However, the majority of his students (nearly 85%) travel from outlying rural areas, using public transport, to the college. Mr. M. identifies as Asian while the vast majority of his students (nearly 85%) identify as African. Mr. M. is a dedicated teacher who genuinely cares about his students and is passionate about motivating his students to achieve good grades and find suitable employment. He finds teaching an enjoyable and rewarding experience. He is decidedly aware of the demands in the workplace for skilled and knowledgeable electrical engineers or electrical technicians to meet the growing shortage of such skills in society and globally. He feels that while his teachers and parents encouraged him to study and become an apprentice in the electrical industry, there was so much pressure on his students from parents and the community to complete their electrical engineering course quickly, achieve good grades in tests and examinations, and get a job and provide economic support to their families.

To pursue his goal of enhancing competence in interpretation of electrical drawings, Mr. M. conducted an action research project for his Bachelor of Education Honors degree called "Practical tasks to enhance interpretation and understanding of electrical drawings at a Technical Vocational Education and Training (TVET) college". His project aimed to transform his teaching strategies when teaching electrical drawings and improve students' competencies by incorporating more practical tasks in his teaching. At the beginning of the year, Mr. M. explained the rationale and purpose of the action research project to his students. Students were asked to keep a reflective journal in which they expressed the challenges they experienced in understanding and interpreting electrical drawings. The prompts for student writing on electrical drawings ranged from: "What did you enjoy about doing the practical task?", "What were the challenges you experienced?", "Do you believe that practical tasks should be done before, after or incorporated into the theory lesson?", "How did the practical task enhance your understanding of the drawing?", "To what extent are you confident to interpret electrical drawings at the workplace?", "Why do you think it is important to draw to learn or learn to draw?".

Mr. M. also observed students during practical tasks to note whether they enjoyed the practical tasks, developed requisite skills, and engaged actively. Focus group discussions were also conducted to discuss students' learning and challenges during the practical tasks. Mr. M. also shared his childhood experiences of drawings and the challenges he experienced with electrical drawings as well as some tips or hints about electrical drawings. This represented the informal process and authentic conversation between teacher and students.

The action research project afforded Mr. M. opportunities to reflect critically on his praxis, his learning as well as his students' learning—Mr. M. learned about using a new teaching strategy of incorporating practical tasks when teaching electrical drawings as students came to learn how to interpret and understand electrical drawings. Mr. M. purposefully adapted his pedagogy and the curriculum he taught based on what he learned from students' reflective journals, their dialogue and his observations and experiences. After completing his action research project, he realized that students communicated more freely during focus group discussions and shared that they enjoyed the practical tasks. He also learned that practical tasks enhanced students' understanding of electrical drawings. An important realization was that by emphasizing practical skills and tasks to enhance understanding of electrical drawings, he was able to bridge the theory and practical gap. Students shared that linking theory to practice enlightened them about electrical drawings and work opportunities. This showcased the living and learning through the curriculum of both the teacher and students.

The onset of the COVID-19 pandemic served as a catalyst to shift to online and distance teaching and learning. Mr. M. has attempted to adapt his teaching strategies to a digital design or flipped classroom. Although face-to-face teaching and learning has been relegated, online and distance teaching and learning has become the most prominent platform. Mr. M. attempts to retain the core principles of culturally relevant pedagogy and the lived curriculum through action research. It is clear that incorporating practical tasks improved understanding of electrical drawings, so Mr. M. has introduced the use of social media and Whatsapp Group Discussions. He shares videos and voice notes with his students and students share their concept maps, drawings and learning experiences as voice notes and messages on the Whatsapp Group. In this way, authentic dialogue between teacher and students continues and they can connect through non-threatening ways. Clearly, students are more open to engage through this online and distance platform as they have the requisite technological competencies. That is, action research and the shift to online and distance platforms affords both teachers and students opportunities to live their curriculum.

Discussion

It is important to re-state the idea that the cases above are not meant to portray answers, models, or best practices associated with culturally relevant pedagogy and action research. Rather, these two examples are presented in an attempt to showcase ways in which we have learned from our students, and the possibilities for action research to inform an ongoing, critical framework for teaching. Both teachers care deeply about education, their work as teachers, and their students. And, both teachers are enacting teaching strategies that have long been a part of classroom learning. The idea that both Mr. G. and Mr. M. are actively engaged in a cyclical process of planning, acting, observing, reflecting, and then replanning is what we find important in the context of action research. Both Mr. G. and Mr. M. acknowledged the cultural knowledge and prior experiences of their ethnically diverse students so that learning was more relevant and effective. It is also noteworthy that both teachers aligned their teaching not only with the principles of action research but also culturally relevant pedagogy.

"Notes to Mr. G." is a version of a classic teaching strategy—interactive journaling. The process of studying this journaling process—from the individualized design of journal covers, to the journal prompts, to student writing, to curricular connections—is an example of a teacher living the curriculum through studying a classroom project. The learning that Mr. G. was able to generate for his students and himself was exponentially expanded by his attention to the compounding, iterative influence of the journals. He was getting students to enjoy free writing while learning about them. He was also gauging his students' interests and writing abilities through his dialogue with them in the journals. And, most importantly for his objectives as a new teacher, he was building meaningful relationships with them through their written conversations. As such, Mr. G.'s action research was a tool for building community and collective learning in the classroom. While the journals may not have been a formal, summative assessment of his students' writing, "Notes to Mr. G." served a much higher purpose in the overall learning in the classroom. As a project it was an example of culturally relevant pedagogy in action. It was constantly evolving and providing opportunities for students and Mr. G. to voice opinions and perspectives about the classroom culture they were creating together. They were living the curriculum together.

Mr. M. critically reflected on his teaching praxis and transformed his teaching strategies to improve his students' understanding of electrical drawings. He revised and evolved his curriculum to include more practical tasks and activities—a standard teaching strategy—which students enjoyed and found very worthwhile. The processes Mr. M. engaged in—reflective journaling, observation of practical tasks and authentic dialogue in focus group discussions—showcased an example of a teacher living the curriculum through an action research project. Mr. M. was able to gauge the interests and practical competence of his students through observation and dialogue. This allowed Mr. M. to recognize and acknowledge the diverse personal and cultural knowledge of his students and the significance of their lived experiences. As such, the action research project allowed Mr. M. as well as his students to learn in culturally relevant ways and live the curriculum together. Mr. M. was able to get his students to enjoy electrical drawings while learning about them. And, most importantly for his aim to improve students' understanding of electrical drawings, he was constantly providing learning opportunities for his students to attain better results and equip them for the workplace. As such, Mr. M.'s action research served as a vehicle to enhance individual and collective learning in the classroom and maximize opportunities for student learning and academic success. And, through critical reflection, Mr. M. not only deepened his knowledge about his practices but also of his learning and that of his students. That is, through the action research project, Mr. M. showcased culturally relevant pedagogy, as he and his students constantly created a classroom learning culture and together lived the curriculum.

Implications

As stated earlier, the cases above are not presented to offer solutions or provide blueprints for culturally relevant pedagogy and action research. We don't believe that such solutions or blueprints exist. Rather, we portray the work of Mr. G. and Mr. M. to highlight the work of classroom teachers that we get to learn from and work alongside. Their work inspires us as teacher educators because they are enacting, reflecting upon, and continually refining their pedagogy in meaningful ways. The brief descriptions of their teaching above are incomplete and insufficient. The descriptions are our best attempts to provide a glimpse into the ways that practicing teachers are living their curriculum. Clearly, there are numerous limitations to these cases and we do not feel that the teaching and action research described above are generalizable or culturally relevant for all educators. What we do believe, is that teachers' experiences are generative, and the fluidity of their work with students is important in the context of continually creating meaningful curriculum. In this way, we feel that both Mr. G. and Mr. M. are engaged in processes that highlight the potential value of a teaching approach that combines culturally relevant pedagogy with action research.

More specifically, these case studies foreground action research as a pivotal tool to transform the lived experiences of teachers who explored their pedagogical practices in culturally diverse contexts. Furthermore, these case studies highlight that critical reflection is key to enhance teachers' understanding of culturally relevant pedagogical practices and the lived curriculum. We argue that the transformational and liberatory goals of teaching (e.g., Freire, 2000) and of action research offer resources for not only teachers but also teacher educators to engage with the challenge of social inequalities, the stark realities of which are highlighted by the global COVID-19 pandemic (e.g., Kumashiro, 2020). Action research created opportunities for teachers to critically reflect upon their pedagogy and broaden their understandings of culturally relevant pedagogy which considers not only the political but also the social dimensions that influence the lived curriculum, their pedagogy as well as their students' understanding and achievement.

As educators and scholars, the cultural contexts in which we develop our work are literally across the globe from one another (South Africa and the United States). Yet, we can see through the frameworks and the case studies presented in this article, that we have much to learn from the classroom teachers with whom we work. And, the similarities expressed through Mr. G.'s and Mr. M.'s action research projects show that there is important work to be done in this area of educational research. Namely, the connections between culturally relevant pedagogies and action research are potent when teachers apply them toward a vision of educational liberation. The truth is that many teachers are continually engaged in the process of planning, acting, observing, and reflecting—it is a core part of what we do with students. In this article, we argue that the purposeful study of this process—when combined with culturally relevant pedagogies—can enhance teacher and student learning for social transformation. In this way, we feel that it is important to continue to study ways in which teachers are living the curriculum.

References

- Carr, W., Kemmis, S., 1986. *Becoming Critical: Education, Knowledge and Action Research*. Falmer Press, Philadelphia.
- Emdin, C., 2016. *For White Folks Who Teach in the Hood ... And the Rest of Ya'll Too: Reality Pedagogy and Urban Education*. Beacon Press, Boston.
- Foucault, M., 1977. *Discipline and Punishment: The Birth of the Prison*. Pantheon Books, New York, NY.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York.
- Freire, P., 2000. *Pedagogy of the Oppressed*, 30th anniv. ed. Continuum, New York, p. 35.
- Gay, G., 2000. *Culturally Responsive Teaching*. Teachers College Press, New York.
- Giroux, H., 1992. Post-colonial ruptures and democratic possibilities: multiculturalism as anti-racist pedagogy. *Cult. Critiq.* 21, 5–40.
- Hagevik, R., Aydeniz, M., Rowel, C.G., 2012. Using action research in middle level teacher education to evaluate and deepen reflective practice. *Teach. Teach. Educ.* 28 (5), 675–684.
- Hefflin, B.R., 2002. Learning to develop Culturally Relevant Pedagogy: a lesson about crowded lives. *Urban Rev.* 34 (3), 231–250.
- Howard, T.C., 2003. Culturally relevant pedagogy: ingredients for critical teacher reflection. *Theory Pract* 42 (3), 195–202.
- Kemmis, S., 2006. Participatory action research and the public sphere. *Educ. Action Res.* 14 (4), 459–476.
- Koshy, E., Koshy, V., Waterman, H., 2011. *Action Research in Healthcare*. SAGE, Los Angeles, California.
- Kumashiro, K., 2020. COVID-19, racialized disease, and the yellow peril. *Insight Divers.* May 18, 2020. Retrieved July 24, 2020 from <https://www.insightintodiversity.com/covid-19-racialized-disease-and-the-yellow-peril/>.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harv. Educ. Rev.* 32 (3), 465–491.
- Ladson-Billings, G., 2016. “#Literate lives matter”: black reading, writing, speaking, and listening in the 21st century. *Lit. Res.* 65 (1), 141–151.
- Macintyre, C., 2016. *The Art of Action Research in the Classroom*. David Fulton Publishers, London.
- Milner, H.R., 2011. Culturally relevant pedagogy in a diverse urban classroom. *Urban Rev.* 43, 66–89.
- Nieto, S., 2002. *Language, Culture, and Teaching: Critical Perspectives for a New Century*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Pour-Khorshid, F., 2020. Teaching to heal, healing to teach: ethnic studies as a healing endeavor in and out of the classroom. In: Agarwal-Rangnath, R. (Ed.), *Planting the Seeds of Equity: Ethnic Studies and Social Justice in the K-2 Classroom*. Teachers College Press, New York, pp. 17–26.
- Rist, R., 1970. Student social class and teacher expectations: the self-fulfilling prophecy in ghetto education. *Harv. Educ. Rev.* 40 (3), 411–451.
- Woodson, C.G., 2000. *The Miseducation of the Negro. African-American Images*, Chicago.
- Zinn, H., 2003. *A People's History of the United States: 1492–2001*. Harper Collins, New York.

Teachers' lived experience: voices of history teachers on the English National Curriculum

Mary C. Woolley, Faculty of Arts, Humanities and Education, Canterbury Christ Church University, Canterbury, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

The history National Curriculum in England	363
History teachers' experience of curriculum since 2010	364
Contrasts from other countries	365
Methodology	366
English history teachers' experience of curriculum reform since 1985	366
Conclusions	369
References	370
Further reading	370
Relevant websites	370

Glossary

GCSE external examinations taken by 16 year olds in England after two years of teaching

Key Stage 3 part of the National Curriculum directed at 11–14 year olds

Ofsted the office for standards in education, inspecting services providing education and care

National Strategies programmes aimed to improve the quality of learning and teaching in schools in England. First introduced in 1998 they formed a key part of the government learning priorities

PISA Program for International Student Assessment (PISA), a worldwide study by OECD intended to evaluate educational systems by testing 15 year olds on maths, science and reading; Schools History Project set up in 1972 to undertake a radical re-think of the purpose and nature of school history

TIMMS part of a history of international comparative assessments in mathematics and science conducted by the International Association for the Evaluation of Educational Achievement

Traditionally, writing and research about curriculum reform, particularly in the case of a history National Curriculum, has tended to focus on policy documentation (Phillips, 1998; Guyver, 2012), which has led to some assumptions of direct translation between policy text and classroom practice. However, in recent years there has been a move to consider the approach teachers have taken to implementing or “enacting” curriculum reform (Ball et al., 2012). With government focus tending toward large-scale assessment data such as PISA and TIMMS, it is essential to explore the corollary and give voice to individual teachers’ experience of curriculum reform. Research with teachers can elicit personal experience of changes and continuities within the curriculum, revealing the complex network of agents teachers respond to when making changes to choice of substantive content or pedagogy. While there have been generic studies of class teachers (Priestly et al., 2015; Perryman, 2006), there is a need for subject-specificity in such research. Experiences of teaching “core” curriculum subjects such as English and maths have received some attention (Gibbons, 2017; Perryman et al., 2011), but there has been far less focus on the teachers of “foundation” curriculum subjects, for example, history. As Counsell (2000, p. 61) has argued, to “decide what history is to be taught, at school, regional, or national level, is to exercise phenomenal power.” Understanding the factors that underpin history teachers’ curricular decision making is therefore crucial if society is to reap the benefits of a relevant and rigorous history curriculum.

This article therefore sets my own in-depth research with history teachers in England within a context of other research, some more recent, some with larger data sets and some from comparative countries. First, a background to the National Curriculum in England is provided, specifically the changing details of the History National Curriculum. Second, temporally and geographically broader literature on history teachers and the history curriculum is explored to set a context. Third, a summary is given of the conclusions of my own research, exploring the lived experience of curriculum of thirteen veteran history teachers over a period of 25 years. The aim is to provide conclusions on what we can understand about history teachers’ lived experience of curriculum and how this knowledge needs to be shared in order to influence policy, practice and further research.

The history National Curriculum in England

The National Curriculum for England was first introduced in 1991 as a result of the 1988 Education Reform Act. It was initially underpinned by a focus on ten specified curriculum subjects, to be taught throughout the 5–16 age range. The National Curriculum

has been revised four times since 1991 (in 1994, 2000, 2008 and 2014), each time (until 2014) increasing flexibility and freedom, with history teachers in particular given less detailed prescription and more choice around substantive content. With the Labor landslide victory in 1997, education was declared a priority through the policy of "raising standards" (see Ball, 2008), and a series of National Strategies were introduced to achieve this aim. These initiatives, applicable across all curriculum subjects, gave an emphasis to "assessment for learning", learning objectives and student activities and came alongside significant funding to encourage dissemination and implementation. Concurrently, this period saw a gradual but significant increase in accountability measures. League tables were introduced in 1992, initially to compare success in the national GCSE examinations for 16 year olds. Ofsted, an inspection body for schools, was introduced the same year and moved over the next two decades from carrying out subject-specific inspections to a focus on progress data and school evaluation. This led to school leaders to needing evidence of progress for all children across the 11–16 age range, resulting in tight regimes of classroom observations, learning walks and data drops in many schools. It is against this backdrop of acute accountability pressures that teachers have instigated curriculum reform over the last 30 years in English schools. The particular experience of history teachers reveals the pressures on a subject where students could opt to discontinue the subject from the age of 14, concentrating on other curriculum subjects for the two years of GCSE preparation. This resulted in the 11–14 (Key Stage 3) curriculum in England becoming something of a marketplace, with subject teachers competing to recruit students to study past the age of 14 (Harris and Haydn, 2012). This in itself had a significant impact on the curricular choices of many history teachers. Besides this competition for students, the backdrop of accountability measures led to increasing pressure to start preparing students for the GCSE examination from a younger age.

In order to understand the position and experiences of history teachers across this period, it is necessary to give a brief history of the history curriculum in England. The original National Curriculum documentation for history, taught from 1991, was a merger between two competing approaches to the subject. There was notable influence from the "great tradition"; didactic teaching of a mainly British, political history to mainly passive pupils (Husbands et al., 2003) but the inclusion of evidential understanding demonstrated the influence of the Schools Council Project 13–16 (later Schools History Project [SHP]). It has been argued that in moving toward SHP, history teaching developed from a "technical act for conveying knowledge" to a "cultural act that taught students about warrant, about the nature of understanding and about their own role in making historical knowledge" (Stearns et al., 2000, p. 3). These quite different approaches to the nature and purpose of school history led to some debate during the construction of the National Curriculum. The initial version of the documentation was problematic due to "the overwhelmingly overloaded content" and "the difficulties of the assessment stages and targets" (Cannadine et al., 2011, p. 199). Eight detailed study units were prescribed, spanning over 2000 years of history. This content was to be assessed through three attainment targets covering knowledge, understanding, interpretations and sources. Theoretically, this put historical processes at the center of the history National Curriculum. However, the detail in prescribed content and assessment proved complex in practice. Foster, writing soon after about his experience as head of humanities during the introduction of the National Curriculum, suggested the proposals were "too prescriptive, detailed and unrealistic" (Foster, 1994, p. 376). In 1993 Lord Dearing was commissioned to simplify the demands of the National Curriculum across all subjects, including history.

As a result of this review, the History Program of Study for the National Curriculum was substantially slimmed down. Nine Level Descriptors were introduced to standardize assessment of progression across the age range from 5 to 14. This document also saw the move to five "Key Elements" which summarized the second-order concepts and processes underlying the teaching and learning of history, emphasizing a disciplinary framework. Further reforms to the Program of Study in 2000 and 2008 made few distinct changes, but it became less prescribed and more flexible. The 2008 document had a particular focus on "cross-curricular" approaches. Schools were encouraged to combine subjects, to develop thematic approaches and to organize their timetables in novel ways (Kitson et al., 2011).

By 2010 though, with New Labor losing power to a Conservative-led coalition government, school history was in a weak curriculum position. Ofsted argued that it was a very well-taught subject (OFSTED, 2011). However, the same report indicated that history was being marginalized in fourteen of the fifty-eight secondary schools visited between 2008 and 2010. A parallel report from the Historical Association led to similar findings, reporting increased teaching by non-specialists in the early years of secondary school, a reduction in the time allocated to the subject and restrictions on which pupils were allowed to take history beyond the age of 14 (Burn and Harris, 2011). The authors suggested, "Division between those encouraged and enabled to study history and those denied such expertise or deterred from continuing with the subject is based on socio-economic factors" (p. 1). It was at this point that exploring history teachers' lived experience of curriculum appeared vital, to understand the disconnect between the flexibility of official curriculum policy and the reality of teachers' experience in history classrooms.

History teachers' experience of curriculum since 2010

In 2010 a Conservative/Liberal Democrat coalition came to power in the United Kingdom. This led to reforms across the curriculum, with new specifications for external history examinations at 16 and 18 introduced in 2016 and 2015 and a new National Curriculum Program of Study introduced in 2013. While several new external examination specifications (for GCSE examined at 16 and A level examined at 18) and programmes of study (for 11–14 year olds) since 2010 would suggest great opportunity for curriculum reform, research has shown that accountability measures have tended to dominate history teachers' lived experience of curriculum in English schools. The Historical Association, the professional subject association for history teachers in England, has carried out annual online surveys of its members since 2009. Several reports and papers have used this data to draw conclusions about the current state of history education (Burn et al., 2018 see for example, Harris and Burn, 2016; Harris et al., 2019), but

conclusions can also be drawn concerning teachers' experience of the curriculum and particular curriculum reforms. For example, data collected through over 1100 survey responses between 2015 and 2017 showed that history teachers' decision making tended to be influenced by performative demands. Accountability measures dominated history teachers' thinking as they tried to implement curriculum reforms for external examinations at 18 and 16 (Harris and Graham, 2019). Teachers' attitudes, beliefs and values were explored alongside opportunities for teacher agency. Marketization of the curriculum and league table pressures for constant progression were found to influence what history teachers chose to do in non-examined phases of the curriculum, forcing teachers to choose between "educational outcomes and educational values" in the design of their Key Stage 3 curriculum. Teachers' agency when designing a curriculum was therefore found to be diminished.

There is however evidence of history teachers in England being passionate and informed about what should be within National Curriculum documentation and willing to fight for less prescription and more responsibility for design and choice at school level. Controversy surrounded the 2013 reform of the History National Curriculum where a draft version was published prescribing a long list of topics, events and individuals to be taught chronologically that all students would be expected to the study. Nearly 550 teachers responded to the Historical Association annual survey in 2013, two-thirds sharing written comments concerning this draft curriculum. The data provided a unique insight into history teachers' understanding of the curriculum and disciplinary history (Harris and Burn, 2016). Research into these comments found teachers communicating a variety of concerns, from the relevance of the proposed content to an over-emphasis on political history to the geographical spread (with an inappropriate balance between British history and the history of other countries and cultures). Over 95% of respondents found the amount of prescribed content inappropriate, showing a desire among English history teachers to take responsibility at school level for many decisions around substantive content. Following the negative reaction to the proposed curriculum, significant changes were made before the new version of the Program of Study was published (See <https://www.gov.uk/government/publications/national-curriculum-in-england-history-programmes-of-study>). However Harris and Burn (2016) were able to use the teachers' comments to gain an insight into their understanding of curriculum content. It was clear that these teachers saw history as being "more than a body of knowledge" (p. 537). Beyond that there was some desire from teachers for "relevant and interesting" material (p. 539) and an emphasis on a disciplinary approach. The authors concluded that "there remain powerful tensions for teachers caught between trying to develop young people's understanding of history, while seeking to serve the more generic aims of school education." (p. 539). This research and the reaction of history teachers to proposed changes to the history curriculum show the level to which teachers are engaged with the document, but also the extent to which they bring their views about the substantive and disciplinary history to their planning and enactment of the curriculum. This is far from a passive act. Not only is the document mediated, but engaged teachers are able to push back, through subject associations as lobbying bodies to reject extreme curriculum reforms. The authors noted the findings from the Historical Association survey formed part of the negative reaction to the proposals, resulting in significant changes being made to the final version of the National Curriculum.

Contrasts from other countries

While history teachers in England seem to struggle, at times, for curricular autonomy, it is interesting to contrast the experience of history teachers in other countries where there appears to be more teacher agency in curriculum planning. In Scotland, for example, teachers have an unusual degree of curriculum autonomy specifying no mandatory historical content in its social studies curriculum. A detailed piece of research with 21 Scottish history teachers explored how this curricular autonomy was used (Smith, 2019), finding that without access to theoretical debates about the nature of historical knowledge, teachers tended to fall back on instrumental justifications for content selection. This tended to lead to an extremely narrow and fragmented syllabus based upon pupil preference and teacher interests. Priestly et al. (2015) have suggested a need to increase teachers' agency through better access to research literature and greater collaborative working. Smith adds to this in demanding a professional conversation about the purpose of history.

Similar research has been carried out in New Zealand where the government also decided to move away from prescribing topics in a history curriculum. Ormond examined the complex nature of teachers' new role and responsibility in this context and found curriculum achievement objectives and national assessment placed significant constraint upon teachers' selection of historical knowledge (Ormond, 2017). She stressed that local curriculum making placed "a heavy burden of responsibility" upon teachers (p. 599). Further, Ormond (2019) found that the need for assessments to tightly fit the wording of national standards has limited flexibility around selection of historical content. Therefore in New Zealand, as in Scotland, high levels of teacher autonomy on paper through an open curriculum have not necessarily led to successful teacher agency and rigorously devised history curriculums at local level due to other accountability pressures in New Zealand and a perceived lack of subject-specific professional development in Scotland.

There is limited literature that explores the teachers' perspective on how history teaching has changed over time. Cuban's (2016) work on history teaching in America, however, has similarities to the research discussed here as it considers changes in history teaching over 50 years from the 1960s. Cuban chose a semi-autobiographical approach, detailing his own experience of teaching in two schools at the beginning of his career and then returning to those schools five decades later. He found considerable continuities in history teaching, estimating that only a quarter of schools were using what he termed the "New History" that prioritized the place of evidential work in the history classroom. These three examples show that changes in selection of content in the history classroom and approaches to how history is taught are complex and affected by many different factors beyond curriculum documentation, such as teacher agency, accountability measures and marketization within schools.

Methodology

This research sought to explore how teachers perceived, experienced and enacted curriculum and policy. Accessing teachers' experiences of curriculum is not a simple task as frequent changes to policy documents, examination specifications and local contexts impact the decisions teachers make concerning the way curriculum is enacted in the classroom. It was decided to explore such changes retrospectively by talking to history teachers who had been in the classroom for twenty-five years or more. The intention was to know more about changes in how the curriculum was enacted over time and the impact other agents had on the curriculum decisions made by teachers. Twenty in-depth, semi-structured interviews were carried out with thirteen experienced history teachers between 2011 and 2016, using an oral history approach to explore change over time.

As a "way of knowing," there are inevitable limitations to the use of oral history; potential problems associated with mis-memory, accuracy and selection. One problem in the study of memory is that any narrative is not necessarily linear, but sporadic and segregated, encroaching on history to differing degrees at different points in the past. Part of this problem arises from the issue of how to access and study memory, whether at an individual or collective level. In *History as an Art of Memory*, Hutton (1993) argued that memory is an internal activity of the living mind that can never be recovered. This poses a challenge for the historian who is limited to studying the ways in which memory manifests itself outside of the mind. Halbwachs (cited in Hutton, 1993) believed the historians' first task was to keep memory honest, "to remedy its distortions of the past by comparing its suspect claims to those based on documented historical evidence" (p. xxiv). By setting oral history accounts against a contextual backdrop gathered from other sources a justified, valid account of the past can emerge. As Cunningham (2006) put it, "documents and personal memory together can shed new light on important themes" (p. 16).

A distinction is also worth making here between life history and oral history. Although the teachers were asked to complete a timeline before the interview that included teacher education, changes in position etc., the open-ended style of questioning gave the teachers room to select the changes in history teaching that were important to them. Gardner (2003) has made a distinction between life history and oral history. While both may use life stories elicited through interview, arguing that oral history emerges when individual memories are probed "for more precise, detailed or verifiable information about the past" (p. 179). In choosing to explore a historical concept such as change and with an aim to enrich the existing historical record, it was decided to use a specifically oral history approach, seeking witness of events and changes that took place in the classroom, rather than a life history approach which might have focused more on the broader lives of participants themselves (Gardner, 2010). This research sought to provide and understand a story of change in history teaching across 25 years. It looked to the stories and experiences of individual history teachers for evidence of that change.

Interviews were transcribed, coded and set against secondary literature and contemporary policy documentation. The teachers involved in the study taught history in the secondary classroom from 1985 to 2011, with most teaching across the whole period. This led to a very particular group of teachers, only a small minority of whom had moved into senior leadership positions, but all of whom were passionate about their subject. Oral history interviews rarely provide a neat, complete account of chronological events. In this study, for example, teachers proved more familiar with the detail of school and classroom experience than with the minutiae of policy or curriculum documentation. The fragmentary nature of interview evidence, coupled with the potential problems of memory suggested a need to set a historical backdrop from alternative sources. This provided a chronological and contextual structure within which to frame the evidence of participants.

The findings provided evidence of a dedicated group of teachers who experienced and enacted sweeping curriculum and policy changes across the course of their careers. Their memories of the period add a rich layer of detail and an affective dimension to the existing literature in this area. Several research questions underpinned the study:

- How did history teachers perceive, experience and enact curriculum and policy changes 1985–2011?
- How did history teaching in English secondary schools change between 1985 and 2011?
- What agents acted on teachers to effect change over this period?

The findings shared in this article will focus on the teachers' lived experience of curriculum reform.

English history teachers' experience of curriculum reform since 1985

There was a dramatic shift in the nature and place of knowledge in the history classroom from the 1980s onwards. Various of the teachers described a traditional, rather didactic approach to teaching history in the early 1980s. Richard described a conversation with his mentor during teaching practice:

I said, look I'm sorry Mr D, what do you want me to do with your class? He said, "well" he said, "I've got up to page 37, page 38 looks good to me, does it to you?" And that was it. That was all I got.

Simon described similar textbook-based practice:

I can certainly remember when I first came into teaching, rows of very miserable-looking children, with a textbook and a pen, and, and an exercise book, with, well, in some cases dictation still going on, and I think that probably, that was probably the end of that era.

Many teachers followed a pragmatic, "dual practice" approach early in this period, including "new history" principles of source analysis and evaluation alongside more traditional approaches. Simon described his approach:

There'd be a balance between chalk and talk and question and answer and the things that go on all the time and the sort of sources-based, skills-based, evidence-based stuff that was quite popular at the time.

It was the introduction of the national, uniform GCSE examination from 1986 that meant such traditional approaches were no longer enough in the history classroom. With the introduction of this new qualification, replacing the more traditional O-level examination, it was no longer possible to teach history to 16 without teaching pupils to consider the provisional nature of historical knowledge. As Edward put it:

the source-based, evidence-based work was much more important than it was with the O-level, I mean the O-level that I taught was really about essay skills ...

Several of the teachers in this study acknowledged a necessary shift in their epistemology at this time. However, the research shows that such changes were neither linear nor necessarily embraced by all teachers. Twenty years after the National Curriculum was introduced, the specific disciplinary language of the History Program of Study was not at the forefront of these teachers' discourse about their practice. The National Curriculum can therefore be seen to have only partial responsibility for this shift in practice, with the shifting demands of external, national assessment in the shape of the GCSE examination acting as more of a catalyst for change within an era of accountability measures.

There was evidence of a shift in all the teachers in this study from "authentic professionals," concerned with the development of substantive historical knowledge in the 1980s to more "technical professionals" with a rather different discourse by the 2000s. According to Ball et al. (2012), where the "technical professional" is created by the policy artifacts of data collection, analysis and measurement (p. 5) the "authentic professional" is "required to bring judgment, originality and passion to bear upon the policy process" (p. 94). The intensified focus on examination results, particularly in the 2000s, led to an emphasis on technical knowledge or how to succeed in examinations and progress through National Curriculum "Levels" of assessment that in several cases displaced an earlier substantive focus. Teachers felt increasing pressure to teach to assessments. Alison compared the shift in her teaching from the early 1990s to twenty years later:

It used to be short questions, short answers and essays. And now it's enquiry-based, it is source-based. We gear our questions, even for Year 7 [5 years before the examination] very much along with lines of GCSE-type questions so that they get used to that formula The bigger influences were the exam boards, and what they required obviously because you know, you've got to get kids used in Year 7 to what's likely to come their way in Years 10, 11, 12 and 13.

Richard gave an honest account of the technical knowledge he needed to share with his pupils:

Certainly, at both A-level and GCSE I tell them exactly how to do it. My GCSE revision session tomorrow, it won't be on how to revise Nazi Germany, it will be on how to do question one. This is what they're asking for, this is how to answer it and make sure you do that. And at A-level even more so It's basically a game that you have to play and you've got to know the rules of the game.

For some teachers, there was a constant fear of Ofsted, placing constraints on the "authentic professional." There was a distinctive difference in the discourse teachers used to describe teaching in the 1980s and teaching in the 2000s. The assessment-focused language of the technical history teacher appears to have been particularly a product of the 21st century. Stories from these interviews about teaching in the 1980s were full of descriptions of substantive content or approaches to teaching such as resource boxes and trips. Laura and Alison were excited about the possibilities of local history; Patrick talked of the theme of women's history running through his schemes of work; William looked for opportunities to teach African history. Noticeably, there was virtually no mention of assessment outside the formal examinations of O-level and CSE. Descriptions of the 2000s, particularly the later part of the decade, however, were dominated by talk of assessment. This significant change in discourse suggests a change in priorities for the history teachers interviewed. For the majority of participants, the experience of teaching history had changed from a passionate exposition and sharing of the substantive content they valued to a vehicle for examination success. This was not a result of the National Curriculum or changes in the National Curriculum, but a direct result of a need for evidence of "progression" to meet the dual performative demands of league tables and Ofsted.

One key finding from this research was that different groups of teachers could be seen to respond to the same curriculum in different ways. While some teachers feared the level of detail in the initial Program of Study, a small minority minimized the impact of the National Curriculum and focused only on changes in assessment. A third group had their own pre-existing professional

agenda to realize and feared the National Curriculum would impede those personal goals. For each of these groups, the process of accommodation was different. Exploring detailed interviews with veteran history teachers enabled a move away from teachers being treated as a homogenous body in the research literature. One small subset of the teachers interviewed appeared to be able to resist or mediate some of the more performative policies around assessment. Context may have had a large role to play (Priestly et al., 2015), but each of these teachers spoke favorably about their early career experiences, sharing in particular detailed discussion and debate with colleagues around the subject of history and approaches to history. As Patrick described it:

I do remember we had quite profound philosophical debates between the deputy head, a real SHP man, who insisted it had to be pure, and myself and the head of department saying no, we've got to edit them [written sources], we've got to simplify them, we've got to make them accessible, so that was quite interesting.

It is quite possible that such opportunities to articulate and define their approach to the subject early in their careers had a long-term impact on their approach to teaching and the enactment of policy. This was a small, but convincing group of teachers. Smith (2017) has written of an "epistemological introspection" (p. 325) that seems to have provided these teachers with a rather more resilient, resistive approach to enacting policy. These teachers were noticeably more articulate about history as a discipline in comparison with many other teachers in the sample.

This research provided an insight into teacher agency, particularly how the teachers believed that agency had changed over time. One fundamental change in the experience of teachers interviewed was the reduction of autonomy that all of them experienced and chose to talk about. In 1985 teachers could choose the content that they wanted to teach across the time and space of history and select their own pedagogical and epistemological approach to teaching the subject. This autonomy was gradually eroded, first by the introduction of the GCSE examination and the National Curriculum, but to a far greater extent by the government policy of National Strategies and the surveillance and data pressures introduced alongside them in the decade following 2000. Not being able to choose how a lesson was taught, for how long the teacher could talk or whether to write lesson objectives on the board was seen by the teachers interviewed as a far greater intrusion into teacher autonomy than GCSE and National Curriculum interventions. Allan described his experience:

Perhaps it was just a new headteacher taking over and interpreting what Ofsted wanted in a more draconian way which led to half-termly observations for every teacher ... because they had a tick-list of what they were looking for it discouraged innovations.

The schedule Allan described would have seen all teachers in the school having formal lesson observations six times over the course of a year. This was much more a surveillance, performative culture than had previously existed in the school. When Allan was asked whether the "tick-list" he mentioned was subject-related, he responded:

They were generic, very much generic that teachers should not be at the front of the class doing whole class teaching for more than about 15% of the lesson time [So that, I imagine is quite a change from your previous approach?] Absolutely, absolutely and that's not really my style. I like to be at the front and leading discussions—breaking off and going into discussion groups for a few minutes and then coming back. [So did you take that on board, that advice, or did you mediate it in some way?] I tried to mediate it, but what I did do, I was observed by the head teacher and I exceeded the 15% quota and was marked down.

While "authentic" examples of autonomy, agency and mediation of policy were present in the evidence of teachers interviewed for this study, such examples were more apparent where teachers had strong disciplinary knowledge and professional confidence alongside working in departments or schools which offered pedagogical freedom. What happened in the history classroom over this period became increasingly prescribed in terms of content, pedagogy, assessment and the status of the subject on the curriculum. Pressure to constantly improve examination results led to a climate of fear in some history departments. While on paper the National Curriculum at this point, was open and free for teacher interpretation, in practice teaching approaches were limited by the generic demands of senior leadership teams, under the influence of the National Strategies and the "raising standards" agenda. Some history teachers and departments were seriously affected, feeling they were compromised in the way they taught the subject due to the demands of senior managers, exam boards and Ofsted. While some teachers saw the need to put into place every initiative and policy according to the letter, others felt more able to resist and mediate the way they enacted the policy, perhaps putting the disciplinary and learning needs of students before the need to conform to performative initiatives from senior leader.

This research approach allows some analysis of the impact of the National Curriculum and subsequent changes to Programmes of Study on the practice of history teachers. Evidence from these teachers suggests the main driving force behind changes in history teaching was the pressure for success in external assessment rather than the Programmes of Study of the National Curriculum. The vocabulary of GCSE mark schemes and National Curriculum Level Descriptors were far more prominent in the language used by teachers than the language of the "Key Concepts and Processes" or "Key Elements" that have made up the disciplinary focus of the National Curriculum. These teachers' experiences of the National Curriculum seem to focus on substantive content, rather than the

minutiae of changes in the various disciplinary "Key Elements" (see Counsell, 2011). This, therefore, suggests many teachers needed more than curriculum documentation in order to effectively teach the disciplinary structures explicit within the National Curriculum Program of Study.

Driving forces in changing approaches to history teaching can be seen in changing resources, particularly textbooks and technology. These were high on the priority list for teachers interviewed to talk about. However such tools do not seem to have promoted changed approaches, but rather emphasized existing approaches. Simon and Allan gave examples of particular textbooks supporting their development of overarching enquiry questions, which was critical in their development as history teachers, but this move did not happen in isolation. The support of teacher networks and conversations with other professionals played a significant role in professional development. Resources were rarely enough to provoke change unless they came as part of a package of professional development. All the teachers interviewed spoke with one voice about the decline in opportunities for subject-specific development in comparison to their early-career experiences. Allan, for example, described the professional development opportunities at the inception of the National Curriculum in the early 1990s:

one of the things we used to do was to run regular three-day courses. Often with between fifty and a hundred history teachers over 3 days that you could really immerse yourself in developing curriculum materials and working together. That's a big loss I think. That's a big loss.

From the late 1990s onwards, opportunities for subject-specific development seem to have become more limited, with generic development opportunities linked to National Strategies more widely available internally to schools. There were still opportunities for professional development through examination boards, but those teachers involved were adamant that opportunities for discussion had become limited at such events. Teachers welcomed the opportunity to work as mentors for initial teacher education courses. However, diminishing quotas for history teachers and the development of school-led teacher education programmes also limited the opportunity to work with universities and other subject-specific teachers in this way.

Conclusions

There are therefore benefits to this way of knowing or exploring teachers' lived experience of curriculum. The in-depth interviews with a small number of very experienced subject teachers provided an insight into significant change over a substantial period of time. The use of a semi-structured interview technique allowed teachers to select and talk about the curriculum opportunities and limitations that they felt passionate about, and this selection was illuminating in itself, with an obvious move from talking about substantive knowledge decisions early in their careers, to assessment dominating descriptions of more recent years. Through this in-depth approach, different groups of teachers could be identified, analyzed and followed across the three decades. Responses to National Curriculum measures from a central government could be compared to responses from senior leadership initiatives within schools and teachers were able to share their emotions and anecdotes, providing a rich and insightful level of detail to a well-known over-arching narrative. There were limitations to this methodological approach; whether such a small number of teachers could represent a national picture of experience; whether taking veteran teachers led to a certain slant on experiences of the National Curriculum that might not have been represented among younger teachers.

As recruitment and, particularly, retention of teachers remains a serious issue in England and beyond, it is crucial to find ways to listen to teachers' experiences. The teachers described above were uniformly passionate about the subject of history and the opportunities it could provide for their students. However, it is rare that they were given the opportunity to share that passion beyond the school walls. Several of the teachers in this study commented on what a cathartic and unusual experience the interview had been. They weren't used to being listened to or being offered a voice or for their experiences to be valued. If the problems of teacher retention in schools are to be addressed, then offering a voice to experienced, older teachers through research or otherwise is crucial. The nostalgia these teachers offered concerning local networks of history teachers may suggest one possible way forward; certainly, the valuing of continued subject-specific development for history teachers beyond of that offered by examination boards would be one recommendation for school senior management teams. The teachers within this research spoke powerfully about their early years in the profession and it is clear that new teachers need subject-specific nurturing at that stage of their career and beyond so that they are able to mediate any more generic government policies through a subject-specific and curricular lens. The opportunities teachers are given to articulate and define their understandings of the subject within subject-specific groupings in the earliest years of their career could prove fundamental to the way they are able to teach, lead and enact policy long into the future. This in turn could provide for a more engaged, contented and resilient long-term teacher workforce, that has an authentic approach to enacting curriculum policy.

Fundamentally, however, it is crucial to recognize and further explore the disconnect that often exists between the official, written curriculum and teachers' lived experience of curriculum in the classroom. If teachers are pulled in two directions by different government policies, in this case the detail of curriculum documentation and the accountability of the performative system, then this needs to be acknowledged. As this article is written, the history curriculum in England is in the media due to a perceived lack of coverage of Black British history in the classroom within the context of the Black Lives Matter movement (Leach et al., 2020). It is crucial that a diverse history is taught within our classrooms, and this is often the case as articles in the professional journal of the Historical Association, *Teaching History*, regularly show (Priggs, 2020; Mohamud and Whitburn, 2019; Lyndon, 2006). As this

research has shown, teachers often have the agency to go beyond the written curriculum and make choices based on their professional and disciplinary knowledge of the subject. The taught curriculum, teachers lived experience of curriculum, however, is so much more than the official written documentation of curriculum. In order to ensure that students can be taught the diverse histories that they deserve and need, the performative constraints acting on curriculum choices need to be acknowledged and then removed. If there are classrooms where teachers are not teaching Black British history, this is not the result of detail (or lack of) in a written curriculum, but the need for career-long subject-specific development, examination specifications that encourage innovation and professional conversations that challenge the status quo.

References

- Ball, S.J., Maguire, M., Braun, A., 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. Routledge, Abingdon.
- Ball, S.J., 2008. *The Education Debate*. The Policy Press, Bristol.
- Burn, K., Harris, R.J., 2011. Historical Association Survey of History in Schools in England. Historical Association.
- Burn, K., Harris, R.J., Sarazin, M., 2018. *Survey of History in Schools in England*. Historical Association.
- Cannadine, D., Keating, J., Sheldon, N., 2011. *The Right Kind of History: Teaching the Past in 20th-Century England*. PalgraveMacmillan.
- Counsell, C., 2000. Historical knowledge and skills: a distracting dichotomy. In: Arthur, J., Phillips, R. (Eds.), *Issues in History Teaching*. Routledge, London.
- Counsell, C., 2011. Disciplinary knowledge, the secondary history curriculum and history teachers' achievements. *Curric. J.* 22, 210–225.
- Cuban, L., 2016. *Teaching History Then and Now: A Story of Stability and Change in Schools*. Harvard Education Press, Cambridge, MA.
- Cunningham, P., 2006. Early years teachers and the influence of piaget: evidence from oral history. *Early Years Int. J. Res. Dev.* 26, 5–16.
- Foster, S., 1994. The painful lessons of introducing the National Curriculum in England. *Educ. Forum* 58, 374–377.
- Gardner, P., 2003. Oral history in education: teacher's memory and teachers' history. *Hist. Educ.* 32, 175–188.
- Gardner, P., 2010. *Hermeneutics, History and Memory*. Routledge, Abingdon.
- Gibbons, S., 2017. *English and Its Teachers: A History of Policy, Pedagogy and Practice*. Routledge, London.
- Guyver, R., 2012. The history working group and beyond: a case study in the UK's history quarrels. In: Taylor, T., Guyver, R. (Eds.), *History Wars and the Classroom*. Information Age Publishing, USA.
- Harris, R.J., Burn, K., 2016. English history teachers' views on what substantive content young people should be taught. *J. Curric. Stud.* 48, 518–546.
- Harris, R., Graham, S., 2019. Engaging with curriculum reform: insights from English history teachers' willingness to support curriculum change. *J. Curric. Stud.* 51, 43–61.
- Harris, R.J., Haydn, T., 2012. What happens to a subject in a "free market" curriculum? A study of secondary school history in the UK. *Res. Pap. Educ.* 27, 81–101.
- Harris, R.J., Courtney, L., UL-Abadin, Z., Burn, K., 2019. Student access to the curriculum in an age of performativity and accountability: an examination of policy enactment. *Res. Pap. Educ.* 35 (2), 228–248.
- Husbands, C., Kitson, A., Pendry, A., 2003. *Understanding History Teaching: Teaching and Learning About the Past in Secondary Schools*. Open University Press, Maidenhead.
- Hutton, P.H., 1993. *History as an Art of Memory*. University of Vermont Press, London, Vt.
- Kitson, A., Husbands, C., Steward, S., 2011. *Teaching History 11–18*. Oxford University Press, Maidenhead.
- Leach, A., Voce, A., Kirk, A., 2020. Black British History: The Row Over the School Curriculum in England. *The Guardian*.
- Lyndon, D., 2006. Integrating Black British history into the National Curriculum. *Teach. Hist.* 37–43.
- Mohamud, A., Whitburn, R., 2019. Anatomy of enquiry: deconstructing an approach to history curriculum planning. *Teach. Hist.* 28–39.
- OFSTED, 2011. *History for All: History in English Schools 2007–10*. Ofsted, London.
- Ormond, B., 2017. Curriculum decisions—the challenges of teacher autonomy over knowledge selection for history. *J. Curric. Stud.* 49, 599–619.
- Ormond, B.M., 2019. The impact of standards-based assessment on knowledge for history education in New Zealand. *Assess. Educ. Princ. Pol. Pract.* 26, 143–165.
- Perryman, J., Ball, S.J., Maguire, M., Braun, A., 2011. Life in the pressure cooker—school league tables and English and mathematics teachers' responses to accountability in a results-driven era. *Br. J. Educ. Stud.* 59, 179–195.
- Perryman, J., 2006. Panoptic performativity and school inspection regimes: disciplinary mechanisms and life under special measures. *J. Educ. Pol.* 21, 147–161.
- Phillips, R., 1998. *History Teaching, Nationhood and the State: A Study in Educational Politics*. Cassell, London.
- Priestly, M., Biesta, G., Robinson, S., 2015. *Teacher Agency: An Ecological Approach*. Bloomsbury, London.
- Priggs, C., 2020. No more "doing" diversity: how one department used year 8 input to reform curricular thinking about content choice. *Teach. Hist.* 10–19.
- Smith, J., 2017. Discursive dancing: traditionalism and social realism in the 2013 English history curriculum wars. *Br. J. Educ. Stud.* 1–23.
- Smith, J., 2019. Curriculum coherence and teachers' decision-making in Scottish high school history syllabi. *Curric. J.* 30, 441–463.
- Stearns, P.N., Seixas, P., Wineburg, S., 2000. *Knowing, Teaching and Learning History*. New York University Press, New York.

Further reading

- Biesta, G., Priestly, M., Robinson, S., 2017. Talking about education: exploring the significance of teachers' talk for teacher agency. *J. Curric. Stud.* 49, 38–54.
- Chapman, A., Burn, K., Kitson, A., 2018. What is school history for? British student-teachers' perspectives. *Arbor* 194. Article a443.
- Goodson, I.F., Hargreaves, A., 1996. *Teachers' Professional Lives*. Falmer Press, London.
- Phillips, R., 1999. *History Teaching, Nationhood and the State: A Study in Educational Politics*. Cassell, London.
- Woolley, M.C., 2020. *Teaching History in a Neoliberal Age: Policy, Agency and Teacher Voice*. Routledge, London.

Relevant websites

- <https://archives.history.ac.uk/history-in-education/about.html>.
- clioetcetera.com.
- www.history.org.uk.
- <https://larrycuban.wordpress.com/>.

Curriculum's multiplicities as participatory modes of interaction

Margaret Macintyre Latta, Okanagan School of Education, Faculty of Education, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction: slippery slopes of reconceptualized curriculum	371
Lived terms of curricular enactment's embrace of multiplicities	372
Meaning-making as elemental to being human	373
Meaning-making as generative, mindful of place and process	373
Meaning-making as in need of other(s)	374
Meaning-making as temporal/spatial negotiation	374
Meaning-making as imaginative, instilling embodied understandings	375
Meaning-making as reflexively informing self-understandings in relation to wider contexts	375
Curriculum's multiplicities as prompting humane becoming	376
Acknowledgment	378
References	378
Relevant websites	379

This article grapples with the needed participatory stance and associated lived understandings to embody curricular enactment as individual/collective meaning-making. In doing so, the article confronts what such enactment orients toward and why. A theorizing language of practice is found through the thinking of American Pragmatist, John Dewey, in conversation with contemporary curriculum theorists' reinterpretations and found kinships with Indigenous wisdom traditions. First, the lived terms of curricular enactment's multiplicities are foregrounded and then associated ways of being and habits of practice are explicated. Curriculum's multiplicities are illuminated through participatory modes of interaction, revealing stories of humane becoming as integral within traversing curricular terrain that attends to meanings-in-the-making. Worldwide, inequities have been foregrounded over this last year—made much more visible and tangible through local and global happenings' disregard for humane becoming. Capacities to foster and sustain diverse learner/learning spaces and invest in meaning-making that is genuinely inclusive, attentive to equity, are core professional skills for educators to embody. Students need to gain these habits and practices too. Communities need to, as well. The article positions readers to reflect on the seriousness of what has happened in the name of education and the lived consequences of ignoring curricular enactment's multiplicities manifesting through humane becoming. A shared commitment to the powers and possibilities of curricular enactments' multiplicities is generated.

Introduction: slippery slopes of reconceptualized curriculum

Meaning-making is elemental to being human. Human beings are meaning-makers—adapting, changing, and building understandings through knowledge-making discourses (e.g., Bruner, 1990; Dewey, 1934; Dissanayake, 2015; Duckworth, 2006; Greene, 1984; Scardamalia, 2017). Curricular situations that bring students, educators, community members together from across disciplines and interests to co-create learning experiences, invest in this elemental and formative nature of knowledge. Harkening back to the Latin root of *curriculum*, co-creating is at the heart of *currere*, reconceptualizing curriculum as a moving sense-making journey. Such characterization draws on the work of reconceptualists over the last five decades, and their re-envisioning efforts within the field of curriculum studies (Doll, 2017; Malewski, 2009; Pinar, 1975; Pinar and Grumet, 2015; Pinar et al., 1995). In 1978, Pinar, reflecting on the use of the term *reconceptualist* called it a “slippery” notion (p. 205); slippery, because the lived terms are interdependent with curricular enactment that embraces the contingencies of multiplicities encountered through knowledge-making discourses. It is such multiplicities that Pinar (2015) foregrounds in his notion of “complicated curricular conversations” engaging participants in ongoing “multi-referenced” conversations (p. 11). The complicated curricular conversations that ensue encourage the reconstruction of participants' lived worlds through “reanimation of the material they study” (Pinar, 2010, p. 5). But, such reanimation is limited and thwarted if preoccupied with predetermined evaluative educational measures and expectations. Pinar (2010) explains how it is persistent institutional neglect of the intellectual quality and character of the curriculum that holds lived consequences that orient curricular enactment very differently. For example:

- valuing singular perspectives and discounting varied histories, traditions, and narratives of experience, orients toward knowledge transmission;
- following set learning sequences and regulated learning paces, resulting in generic learning products that discount more fluid purposeful learning experiences through divergent learning processes and varied learning products, orients toward preconceived ways to learn particular content; and

- masking differences as catalysts in coming to know self and others, limiting, thwarting or losing opportunities to learn with and through others' experiences, orients toward prefigured ways of being and doing.

Thus, the slippery slopes of reconceptualized curriculum foreground the primacy of multiplicities as always present within all situations. But, concomitantly, needing to be valued as productive givens, catalytic within teaching/learning experiences of all kinds. Recently, [Pinar \(2022\)](#), reflecting on the reconceptualization of the field of curriculum studies, explicates how these complicated conversations traverse “across generations, reactivating the past to find the future, an ongoing dialogical encounter in service to conscientization, located in, animated by—as it supersedes—the historical moment and the individual lives of those submerged in it.” And, such an attentive gaze while traversing the curricular terrain entails locating and navigating these multiplicities as the needed matters or resources shaping all inquiry as it unfolds. It is this ongoing attention to process that continues to be persistently misunderstood within educational institutions and communities worldwide. When the traversing activities themselves are reduced solely to a predetermined approach, strategy, technique, or method, meaning-making is not occasioned nor the associated learning opportunities and powers. Rather, the activities can thwart and stultify multiplicities. Again, and again, such misunderstanding disregards what [Dewey \(1910\)](#) terms the presence and power of human beings' innate resources of curiosity inciting meaning-making suggestions and order-making, as habitual dispositions catalytic within learning of all kinds.

Lived terms of curricular enactment's embrace of multiplicities

Curricular enactment, unfolding as moving sense-making journeys, positions educators, students, and all involved as active participants, residing and contributing within such movement of thinking. Residing and contributing accordingly, entails dwelling at the conjuncture of multiplicities of meanings that educators, students, and others bring, encounter, and negotiate. It is the found nuances, intersectionality, and arising possibilities that such dwelling acknowledges ([Aoki, 2005](#); [Hansen, 2011](#); [Schubert, 1986](#)). Most importantly, these multiplicities are understood as deserving of everyone's attention. And, it is this capacity for greater awareness or sensitivity to the multiplicities at play that educators alongside their students must awaken and cultivate to concomitantly see and act within curricular enactment to further learners and learning. [Dewey \(1929\)](#) explains that traversing such moving terrain, marks a rhythmic movement “preserving and propagating a creative attitude” (p. 132).

It is through releasing of creativity as an attitude within meaning making of all kinds, that [Dewey \(1929\)](#) claims genuine inquiry is authorized, positioning all involved to enter as creators, makers of meaning (p. 132). Many contemporary theorists concur reiterating how entering into curriculum-making as an individual/collective co-creating undertaking insists on room to draw attention to the many relational complexities ([Aoki, 2005](#); [Biesta, 2013](#); [Bingham and Sidorkin, 2004](#); [Greene, 2000](#); [Macintyre Latta, 2013](#); [Noddings, 1984](#); [Seidel and Jardine, 2016](#); [Stengel, 2016](#); [Thayer-Bacon, 2004](#); [Wearing et al., 2020](#)). And, it is the deliberative engagement with these relationships that demarcates such sense-making terrain concerned with caring interactions, deliberations, debates, and discernments made within process. A vitality takes life that gives expression to the individual/collective movement that unfolds. [Dewey \(1934\)](#) describes the sense-making movement as being “carried out beyond ourselves to find ourselves” (p. 195). It seems such liberating movement is the heart, fueling and vivifying multiplicities vitality. Pursuing the lived terms of such vitality, [Aoki \(1991\)](#) confronts the static interpretations of many terms in education. He states, “Time-honored words, curriculum and pedagogy, with their frozen meanings need to be placed into turbulence, alert to the multiplicity of meanings and, as well, the legitimation and delegitimation of certain of these meanings” (p. ii). And, it is such unsettling of terms that he calls “inspired” (2005, pp. 357–365). Through multiplicity of meanings, inspired turbulence discloses telling learner/learning significances, instilling the needed liberating agency. It is this movement insisting that all involved situate themselves in the world again and again that forms the task of co-curricular-making, orienting toward greater individual and collective self-understandings. And, it is the vitality of such liberating terrain that participants alongside policies and practices, keep betraying and fearing within curricular enactment.

Grounded within co-curricular-making's concrete engagement, attention is given to the moving curricular terrain of embracing multiplicities. Associated ways of being and habits of practice emerge, giving expression to pedagogical makings which facilitate this movement. And, it is these participatory modes of being that have been, and continue to be, repeatedly circumvented. Curriculum severed from teaching and learning denies the needed modes of being and habits of practice that insist on educators, students, and communities traversing together the curricular terrain encountered, orienting toward individual/collective growth and well-being. Seeking greater physicality and materiality embodied within the very nature of humanity's being-ness, insists on personal investment within curricular enactment. It is only within the ongoing confronting and challenging of self-understandings that the contextual conditions and criteria to foster and nurture the lived curricular consequences emerge as productive for learners/learning and teachers/teaching. So, what are these needed modes of being and associated habits of practice? What do they orient toward and orient away from? What is at stake for liberating curriculum as a medium for sense-making, inspiring participatory modes and habits toward learning/living well in the world with others? Curriculum's multiplicities as openings offering ways of being and proceeding, prompt knowledge-building discourses. The associated co-curricular-making enters into a shared ethical commitment to openness with willingness to challenge values, assumptions and beliefs, alongside boldness to grapple with difficult and complex issues. Such attentiveness and investment in process, from within process itself, are integral. These knowledge-building discourses and associated meaning-making habits insist on what [Bal \(2002\)](#) terms, “traveling concepts”; operative processes informing how knowledge is acquired through our being-in-the-world, derived through the act of participation. As such, these concepts are not

fixed. They travel—between and across disciplines and the particulars of participants, contexts, and places. Varied sense-making paths emerge and as Bal explains they offer “... territory to be traveled in a spirit of adventure” (p. 23). It is the agency of these traveling concepts that Bal embraces for interdisciplinary work in the humanities. And, in seeking the lived terms of embodying the physicality and materiality of curricular multiplicities, I find similar traveling concepts to emerge, positioning all involved in attending to individual/collective co-curricular enactment from within the process itself—creating meaning—creating self (Macintyre Latta, 2013, pp. 14–21).

In what follows, these traveling concepts are identified and elucidated as curricular modes of being and doing. Grounded in Dewey's (1910, 1916, 1934, 1938) scholarship explicating knowledge and mind as verbs, giving flesh and blood to the ensuing “movement of thinking,” these traveling concepts are put in conversation with contemporary theorists, seeking a language for the associated participatory habits and interactive practices. Collectively, these theorists share a commitment to curricular conversations connecting self and world in an ever-enlarging conversation. Co-curricular enactment takes life, assuming continual adapting, changing, and building of individual/collective understandings. In doing so, the making of caring judgments en-route reveal again and again what it is to be human. The powers and possibilities of curriculum's multiplicities are illuminated through these participatory modes of interaction, revealing stories of humane becoming as integral within traversing such curricular terrain.

Meaning-making as elemental to being human

Dewey (1910) states that within the act of meaning-making “there is more than a desire to accumulate just information or heap up disconnected items In the feeling, however dim, that the facts which directly meet the senses are not the whole story, that there is more behind them and more to come from them, lies the germ of intellectual curiosity” (p. 32).

It is within physiological unease that Dewey (1910) claims curiosity manifests. Such unease may manifest as a question or wondering, through confronting an entirely foreign idea or perspective, or pursuing an unsettling feeling or stance that troubles what an individual knows. In doing so, curiosity's unease encounters ensuing multiplicities. Germinating sense-making terrain is incited. Such terrain assumes human beings interpret and make sense of their worlds through interactive engagement with these found multiplicities (Delaney and Mayer, 2021; Dewey, 1934; Dissanayake, 2015; Duckworth, 2006; Grumet, 1988; Johnson, 2007; Noddings, 1984). Such interactive engagement is personal—allowing for intersections with ideas, relations, and connections across multiple accounts; embodied—as it draws upon our lived bodily experiences that inform what we see, feel and respond to; and derived from narratives of experience—recalling prior narratives of experience and shaping narratives of experience ahead. Thus, individual/collective vulnerability is an ongoing accompaniment of germinating sense-making. Many contemporary education theorists articulate how entering into and living within such vulnerable terrain is necessary for genuine learning (Biesta, 2018). It is through the heightened visibility gained through individual/collective values, assumptions and beliefs made explicit, that prejudices are foregrounded, acknowledging the specifically situated and historically conditioned nature of all understandings. Gadamer (2000) explains “Whatever is being foregrounded must be foregrounded from something else, which in turn, must be foregrounded from it” (p. 305). Continually situating selves within such sense-making terrain gives life to the vitality integral to embracing the unfinished creative task of meaning-making as elemental to being human, fundamental to living itself. It is a vitality that incites curiosities that individuals may not know they ever had or could ever envision. Through embracing curricular enactment's multiplicities, interests germinate and foster ever-enlarging and deepening individual/collective understandings. “appealing to and fostering powers already active” within all participating (Dewey, 1910, p. 30).

Meaning-making as generative, mindful of place and process

Dewey (1916) states, “Never is method outside of material” (p. 65).

Dewey's insistence on the given particulars within every situation as the necessary starting place for all learning, puts mindful trust in the strengths, needs, and challenges of participants, context, and content, thus trust within unfolding contingencies. Collectively, these contingent multiplicities form what are particular and irreplaceable within each curricular situation (Chambers, 2008). Such investment is infused with emotional commitment in which learning is not simply an object or concept but permeated by a feeling or attitude that facilitates ongoing participation. Entering into relationship with learning in these ways insists on ongoing mindful attention (Aoki, 2005; Hansen, 2005; Ragoonaden, 2015; Rodgers, 2021; Van Manen, 2015). Dewey (1916) terms such mindfulness a confidence in process assuming a fitting organizing dynamic will emerge. The openings for seeing and responding within such an unfolding dynamic arise through building relations among students, teachers, context, and subject matter. It is these relationships that gather the materials informing curricular paths for intersections and deliberations. And, as these relationships emerge and develop between students and content, it is educators' and students' capacities to see, and concomitantly respond in ways that foster connections, through which curricular enactment comes into being. Derived through such an intimate understanding of situation, curriculum then is not applied or imposed as predetermined methods, techniques or strategies, but rather, unfolds as a mindful knowing in action that can never be fully anticipated. Learning can quickly become calcified without such intimacy.

Pinar (2009) conveys how conversing from within place, positions teachers and students to attend to the given and arising multiplicities as the generative materials of curricular conversations, building relations across self, others, and subject matter, offering productive ways to proceed. The generative movement of thinking that ensues is at the heart of what it means to educate. And, attention to this movement is the needed mindfulness for all involved. Thus, “mind” is understood as a verb assuming an inquiry stance

(Cochran-Smith and Lytle, 2009), seeking genuine concerted action from within process. The curricular task is to see what minds/bodies/spirits bring to every curricular situation. Aoki's (1992) pedagogical stance of tactfully following the pedagogic good in a caring situation provides guidance, making these multiplicities visible and then adapting, changing, and building meaning—moving minds/bodies/spirits. This stance is *watchful*—mindful of situation, relations, and action; and *thoughtful*—demanding a presence within the moment, taking in, receiving, and acting as situations call forth—oriented for learners'/learning's sake, deliberately seeking the well-being of students (p. 25).

Meaning-making as mindful curricular enactment is accessed through ongoing participatory contact and communication. Perceiving accordingly entails active searching alongside an intensely receptive responsiveness. Means and consequence or process and product simultaneously inform each other, rather than alternating or operating distinctively. Mindful curricular enactment generates a continual organizing and reorganizing experience, guided by the anticipation of the whole. For some time now, given concerns with compliance, uniformity, and standardization, teaching has been severed from notions of curriculum (Pinar, 2009). In such circumstances, setting pedagogical tone meant preparing students in advance to learn the curricular content. Such learning referred to covering/acquiring content and/or transferring knowledge. But, as pedagogy's roots suggest, to guide learners is to attend to pedagogical tone and conditions on a continual basis. Learning then entails creating the needed interactive context alongside continually seeking and building relations and connections across and among students, teachers, and subject matter. But, it is fear of the visibility and tangibility of relational multiplicities through curricular enactment that results in curricular enactment that separates form from matter, subject from object, and means from ends, and deliberately "obstructs", "deflects", and "prevents" (Dewey, 1934, p. 132) the needed vital interactions of Pinar's (2009) complicated conversations.

Meaning-making as in need of other(s)

Dewey (1934) states, "When the linkage of the self with its world is broken, then also the various ways in which the self-interacts with the world cease to have a unitary connection with one another" (p. 247).

It is lack of curricular room for multiple learning linkages with other(s) that devalues the roles of sense, feeling, desire, purpose, knowing, and volition within its enactment, instilling resistance, disregard, distrust, and fear as ways to exist according to Dewey (1934). Valuing these holistic participatory roles within sense-making of all kinds assumes that to learn about, with, and through other(s) is to create meaning, and concomitantly be created, always turning back on self through reciprocal interactions and modifications. Such reciprocal interactions and modifications, caringly connecting self and world, form curricular matter that very much matters. Thus, self-expression needs other(s). Dewey (1934) insists that it is the activity of the meaning-maker, each unique human being having a particular relationship to others, objects, events, and contexts, that converts an activity into an act of expression, holding "... together the various factors and phases of the self" (p. 252). Therefore, all understanding entails greater self-understandings.

Dewey (1934) maps out the reflexive interchange that exposes the makings of such expression beginning with impulsion, acknowledging the interdependency of self and surroundings, valuing resistance and obstacles, and enacted at the junctures of old and new encounters (pp. 58–60). Room to incite personal needs and interests manifests as participants are emotionally invested and guided by purpose and values. Embracing and following the creative flux, impulsion holds an agency all its own. Such agency values the active relations and intersections participants bring to situation. And, it is through the particular and varied narratives of experience that individuals confront and engage their understandings of self-in-the-world. Personal values, assumptions, and beliefs gain greater visibility to self and others. Seeing, hearing, and feeling one's way into thinking instills a belongingness that is interdependent with situation. It is belongingness that morphs through cognizance of selves within meaning making. And, such meaning-making reveals the importance of resistance and obstacles encountered to be productive. Feeling and interest, fear and hope, disappointment and elation must be concretely negotiated within meaning-making to value it as a self-confronting and self-challenging venture (Dewey, 1934, pp. 59–60). Living/learning accordingly takes life through the conjuncture of old and new understandings, through enlarging, deepening, and transforming sense-making through different situations and perspectives. Many contemporary learning theorists explicate how this reflexive interchange is not a solo venture but rather a venture that understands the centrality of the other(s) as being constitutive of the self (e.g., Bereiter and Scardamalia, 2016; Biesta and Stengel, 2016; Phelan, 2015; Pinar, 2019; Strong-Wilson, 2021; Rocha, 2020). They articulate how a separate sense of self is detached from the circumstances in which learning develops. They emphasize how a connected self is invested in the remaking of self. It is through the multiplicity of individual/collective connections that a movement of thinking is constructed and reconstructed in which all change in process with, from, and through other(s).

Meaning-making as temporal/spatial negotiation

Dewey (1934) states, "Space ... becomes something more than a void ... it becomes a comprehensive and enclosed scene within which are ordered the multiplicity of doings and undergoings in which man [sic] engages. Time ceases to be either the endless and uniform flow or the succession of instantaneous points ... it is an ordering of growth and maturations" (p. 23).

The multiplicities at play within attention to curricular agency and potentiality are more apt to be lost or undermined when learning space focuses on physical setting, excluding individuals' perceptions, emotions, and memories as present and infusing all spaces. And, further undermined, when time is understood as a commodity that controls and accounts for curricular planning and learning. Rather, Dewey (1934) "space-time" assumes a temporal movement of past-present-future within all sense-making.

Dewey terms this “a moving force” (p. 38). Attending to this movement of thinking from within the movement itself is depicted as “an ever-present process” (Dewey, 1938, p. 50). Delving into the nature of the present that positions teachers and students to seek the agency of “space-time” foregrounds what the immediate present gathers, what ought to shape the interactions taking place, and why it matters. Approaching curricular situations with this stance, understands that there is much already at play. Accessing these multiplicities, surfaces past understandings with implications for the future, while the immediate concerns and interests of the present ask all involved to seek worthwhile meanings for all involved. Time becomes a fluid organizational rhythm to be located within the present. Responsiveness to needs and challenges allows for multiple paces with varied pauses and rests, alongside a heightened sense of learning satisfaction and pleasure.

Active structuring of what is encountered on a continual basis becomes the curricular ground of “space-time.” Facilitating this ongoing structuring of meaning-making surfaces the role of growth within curricular situations as necessarily interdependent with “space-time.” Locating the needed sustenance for learners/learning within the present invests in growing individual/collective sense-making. It is such a turn towards the *contexts* for learning rather than the *causes* of learning that many contemporary theorists reiterate (e.g., Ayers et al., 2010; Biesta, 2007; Green and Reid, 2008; Hargreaves and Shirley, 2009; Palmer, 1998; Phelan et al., 2020; Wozniak and Fischer, 1993). In doing so, cautions are raised by discounting “space-time” including disconnected and numb learners and learning situations that history and contemporary times repeatedly reveal as being violent to individuals and violent to the world in-the-making (e.g., Battiste, 2013; Freire, 1970; Jardine et al., 2006; Kincheloe, 2008). It is the acceptance and surrender to “space-time” that must be heeded through understanding the lived costs of predetermined curricular control closing off opportunities afforded through attention to given multiplicities. Rather, embracing control as direction for learning derived or found through attending to the multiplicities shaping learning as a “moving force,” holds the potential the world keeps avoiding and fearing.

Meaning-making as imaginative, instilling embodied understandings

Dewey (1934) states, “Roominess, a chance to be, live, and move” (p. 209).

Dewey’s (1934) notion of roominess assumes participants are in-touch with context, seeking fitting ways to respond accordingly. The body’s role within sensemaking assumes that the human body is not separate from the mind, nor thinking from feeling, nor individual from social. Rather, such dualistic thinking is nonexistent within the needed reciprocal relationships of what Dewey (1934) terms, “mind-body” (p. 227). Curricular roominess allows for such reciprocal movement, actively making room for a multiplicity of suggestions accessed through mind-body multi-sensory mediums of deliberation, intuition, anticipation, natality, and enlarged realizations (Macintyre Latta, 2013, pp. 66–72). These mediums rely on embodied opportunities to make meanings that are interdependent with imagination as a means to image-forth possibilities. Deliberation images forth greater flexibility, learning through unexpected ways of being and responding. Intuition images forth greater trust within sensemaking processes. Anticipation images forth learning connections again and again, recursively enlarging individual/collective prior understandings. Natality images forth a plurality of tensions and differences as catalytic within all sensemaking. Enlarged realizations image forth continual learning momentum with conclusions understood as momentary syntheses that push forward into new learning terrain.

Thus, imagination serves as imaging-forth gateways into illuminating and appreciating multiple understandings (Baldachino, 2009; Greene, 2000). And, these understandings are further instilled through embodied mediums for seeking roominess within curricular enactment. But, it is the communicative, relational, and interactive multiplicities revealed through roominess within curricular enactment that many contemporary theorists document as being undermined and thwarted (e.g., Donald, 2009; Eisner, 2002; Irwin and de Cosson, 2004; Kumar, 2013; O’Loughlin, 2006; Sidorkin, 2002). Communication’s expressive powers can be hidden and suppressed through top-down curricular mandates and prescriptions repressing and controlling participation. Relational connections to subject matter are avoided when knowledge is understood as separate and disembodied, excluding interest, care, and pride in self and learning. Interactive practices are foreign if never or rarely engaged. The significances for learners and learning lacking lived embodiment of communicative, relational, and interactive multiplicities are lost through regulated curricular enactment divorced from the multiplicity of relational complexities intersecting through time, space, context, and participants.

Imaginary roominess for deliberation, intuition, anticipation, and natality open into multiple curricular paths and forms of engagement that compel participants’ “surrender and reflection” within the sense-making process itself (Dewey, 1934, p. 53). Such surrendering and reflecting opportunities are experienced as embodied impacts that hold the significances to be discovered through expressive communications, self-understandings, and found purpose occupied with forming learner/learning connections (O’Loughlin, 2006; Varela, 1997). It is the human body—living, being, moving—that needs to be entrusted with sense-making and empowered through curricular roominess, way-finding.

Meaning-making as reflexively informing self-understandings in relation to wider contexts

Dewey states, (1943), “the great thing for one as for the other is ... education which enables him [sic] to see within his [sic] daily work all there is in it of large and human significance” (p. 24).

Multiple reflexive opportunities to renew understandings of self through varied perspectives, invite concrete connections with others and embrace the agency of knowledge-building discourses. It is this agency that Dewey (1916) explains reframes individuality away from self-sufficiency and independence toward revising and transforming personal values, assumptions and beliefs, interdependent with openness to other(s) (p. 295). The needed attitude or stance taken must be one of ongoing inquiry rather than

mastery. The recursive opportunities availed for such self-renewal foster a dynamic rhythmic movement of thinking, positioning participants to locate selves within the given contexts, thus continually adapting, changing, building understandings of self in relation to the world. Contemporary theorists confront the need to recover a broader sense of world in which participants can break free from the “cage of subjectivity” (Smith, 2006, p. 80) confining self-understandings (e.g., Hansen, 2011; Kerdeman, 2003; Nussbaum, 2010; Pinar, 2011; Stitzlein, 2020; Waks, 2009). It is the opportunities seized to situate self within what is known and what is not known, said and unsaid, visible and invisible, that makes more explicit multiplicity’s renewing capacities and their lived consequences.

Knowledge-making discourses share a primary concern for attention to learner-learning connectedness mediated through the multiplicity of relationality. Seeking such connectedness invests in the particulars of self and situation first. Therefore, connectedness attends to histories of individuals and situations, positioning participants to confront these givens. Genuine concerted action is prompted as participants scrutinize connections and an individual/collective reorganizing movement of thinking manifests through divergent learning processes and products, divergent perspectives, traditions, and experiences. It is integral that participants respect and value such differences as catalysts for sense-making, placing more and more confidence in the reorganizing movement concerned with emergence and ongoing development. Spending more time and concrete practice living in such connected ways in classrooms is a pressing concern of many theorists. It is such concrete curricular enactment that they envision as enabling participants to live better in the world with others, seeking humanity’s common wealth and the needed modes of being to reside well together (Gaudelli and Hewitt, 2010; Granger, 2006; Hansen, 2011; Noddings, 2013; O’Loughlin, 2006; Pinar, 2009; Stitzlein, 2020).

The primacy of being faithful to all that unfolds through given multiplicities as shaping the task of curricular enactment, positions educators alongside their students residing at the nexus of curricular situations and interactions. Residing as such fosters heightened awareness of the choices that educators make and their consequences for learners/learning attending to the conditions and ongoing creation of genuine learning contexts. Such awareness demands that educators see what is at stake in given situations and further the movement of thought in self and others. Educators and students embrace the agency as integral to human flourishing (Cherkowski and Walker, 2018). Thus, curricular enactment embracing multiplicities is never scripted, nor formulaic, but rather, take multiple forms, interdependent with the particulars of situation. Such curricular enactment importantly emphasizes the catalytic role of educators committed to this pedagogical stance, alongside emphasizing the impacts for learners and learning, invested in growing strong student identities valuing interactions, deliberations, and debates. It is the difficulties, distances, and ideas comprising diverse multiplicities of interactions, deliberations, and debates that hold the significances for learning and learners. Curricular enactment so conceived is living inquiry, attentive to the givens of context, the particulars involved, and the relational complexities that ensue, generating seeing and acting to further learners/learning.

Curriculum's multiplicities as prompting humane becoming

Curricular multiplicities are brought to life through meaning-making’s elemental connectedness to being human, meaning-making’s embracing of the given within all situations, meaning-making’s need of others, meaning-making’s spatial/temporal agency, and meaning-making’s interdependence with imagination, instilling embodied understandings. Collectively, these interdependent and participatory modes of being and doing initiate the much-needed curricular terrain for educators alongside their students to see and act within the fecundity that transpires, bringing “... curriculum to life—to recapture the creative energy of all life, the esthetics that exist in being ...” (Doll, 2002, p. 34). The fecundity is characterized by Doll (2013) as “rich”, “related”, “recursive” and “rigorous” (pp. 215–222). And, it is the multiplicities traversed through the rich, related, recursive, and rigorous curricular terrain that bring all participants near to the moving particulars of situation and interaction. Such a lively spirit, seeking ways to proceed for learners/learning by attending to the sense-making terrain, holds tremendous potential to reframe education away from competition—good and bad students, teachers, schools, and communities; away from imposed top down controls—monolithic policies, methods, and practices; and, instead, invest within genuine learning experiences that foster strong student identities, growing strong school cultures, facilitated by strong educators and communities laboring to understand themselves and the work they inhabit. Enhanced educator and student empathy, openness, appreciation of differences of all kinds, and valuing of the process character of interactions, translate into greater self-understandings and learning confidence alongside developing capacities to live better in the world with others (e.g., Apple, 2013; Biesta, 2013; Groundwater-Smith and Mockler, 2009; Jardine, 2012; Macintyre Latta, 2018; Noddings, 2012; Stengel, 1991).

Troubling this account of rich, related, recursive, and rigorous curricular terrain is Doll’s (2002) attention to what he terms the curricular “ghost of control,” as a persistent haunting encounter that continues to thwart and betray the needed “livelier spirit of control” within curricular enactment across disciplines and interests of all kinds. This is not new thinking, but it is bold thinking that world-wide curricular enactment in schools and communities persists in avoiding and short-changing. The participatory modes identified through curricular enactment’s multiplicities reframe understandings of education and knowledge invested in such a livelier spirit of control. What education and communities need to embrace is just how lively, how revolutionary, the ensuing meaning-making modes can be, holding, the powers and possibilities of curricular enactment that is oriented toward learner/learning growth, greater self-understanding, enhanced well-being, and the opportunity to continually situate self in the world alongside others. Such meaning-making can be a vital medium for embodying and strengthening the roles of education within all institutions and communities, locally and globally (Tierney, 2018). It is world-wide human preoccupation with what Doll (2002) describes as “a method to bring control and order to life” that he argues haunts curricular visions (p. 34). Taking up Doll’s challenge to “ferry this spirit to the

other side," (p. 28) points to the primacy of investing in curricular enactment's multiplicities as disrupting long-held colonial learner/learning relationships (e.g., Archibald, 2008; Donald, 2012; Hare, 2016; Lowan-Trudeau, 2014; Ragoonaden and Mueller, 2017). The complex relational ground comprising the participatory modes of being and doing constitutes what Leggo (2007) conveys as "... a way of communing with others, with the whole of Creation; " a stance that sees the relational whole, "... calling out ... to other human beings on hopeful journeys of humane becoming" (p. 165). The phrase, hopeful journeys of humane becoming, sums up the impacts of curricular enactment's multiplicities in action. Not only does the phrase offer a wonderful way to respond to fundamental curricular questions concerning the nature of education and knowledge, but also, in emphasizing hope, it assumes potential in self, others, and situation; in drawing upon the notion of journey, it assumes movement; in grounding the venture in the uniqueness of individuals, it foregrounds humaneness; and in operationalizing the phrase through becoming, it assumes an ongoing mind/body/spirit/wholistic process, inviting participants to be and to become, again and again.

Collectively, curricular multiplicities offer ways to be and become through curricular journeys revealing how given situations and particulars shape the individual/collective movement of thinking incited and explicate how these journeys unfold as fundamentally contextual. Openness to the unanticipated is key alongside acknowledging that something is always already at play. Responsivity is assumed, inciting movement very much animated by the spirit of all that forms/informs each journey. It is after all the relational workings of self and other(s) encountered on each journey that explicates these features, suggesting a way to proceed and, concomitantly, bringing selves into being too. The connectedness to co-curricular-making manifests across disciplines and interests of all kinds, fostering the curricular conditions for ourselves and our students to continually locate and relocate themselves on varied curricular paths, creating meaning, creating self—heeding the call to journey with others in life's ventures, embracing learning as life-giving. The ongoing search, locating selves within these curricular journeys, is what it is to be alive. Furthermore, the research literature over the last two decades reveals that the world is only recently awakening to how such meaning-making finds kinships with long-held beliefs and modes of being embodied within Indigenous wisdom traditions (Arrows et al., 2010; McKinley and Tuhiwai Smith, 2019). World-wide, Indigenous connections to land, culture and the relational self, convey the need for such kinships (Haig-Brown, 2010; Kanu, 2011; Ng-A-Fook, 2014; Styres, 2017). And, it is within seeking such connections that the kinship of Indigenous commitments to interconnectedness, reciprocity, relationality, reverence and respect, emerge and offer the needed learning conditions, supports and participation toward decolonizing curricular efforts (Atleo, 2009, 2011; Archibald, 2008; Battiste and Henderson, 2009; Cajete, 2015; Dion, 2009; Donald, 2019; Carter et al., 2015; Cohen, 2001; Arrows et al., 2010; Restoule et al., 2013).

Current times offer no shortage of how multiplicities hold many teachings. The local/global intersections are blatantly apparent. For example, arising through escalating strife, given Russia's invasion of the Ukraine; the realities, intersections, and impacts of Truth and Reconciliation efforts; Black Lives Matter movement; and Orange Shirt Day confronting the violence of Residential Schools. This cross section of current concerns open the door to world-wide conversations on race, racialization, and racism, and reveal the genocidal legacies of ongoing costs and long-term learner/learning consequences; and the coming together of Indigenous, Black, and People of Color re concerns for equity, justice, and associated societal changes; with all foregrounding how systemic racism and monolithic thinking excludes and marginalizes, and have been and are, extremely harmful. In support of all human beings, movements such as SOGI 1 2 3 foreground how schools must be inclusive and safe for students of all sexual orientations and gender identities (SOGI), emphasizing the importance of gender and sexual orientation not limiting interests and opportunities, and where time and space to understand and express gender and sexualities are welcomed without discrimination. As the COVID-19 pandemic (2020) persists with unprecedented global impacts on societies, highlighting the disparities among individuals and nations, it importantly implicates how human connection, well-being, and health are intricately tied. Climate change, biodiversity, and the dire consequences of ecological imbalance are now much more commonly understood as shared urgent world concerns. This cross-section of global multiplicities is heightening worldly attention to human beings/non-humans/Mother Earth connectedness, foregrounding education's roles and responsibilities. Curricular multiplicities as participatory modes of interaction position all of us as complicit within situations together, world-wide. Such complicity entails investment within long-term human/non-human well-being in a shared world, productively reframing education locally, nationally, and internationally. Curricular enactment embracing its multiplicities, forms the much-needed practice ground that can no longer be ignored. It reframes education away from competition and imposed control and toward individual/collective growth and well-being, exposing the fertile ground of sense-makings' multiplicities and offering the sustenance the world needs to think and act together; collectively taking seriously ways to live well in the world with all our relations. Curricular multiplicities as participatory modes of interaction position all of us as complicit within situations together, world-wide (Noddings, 2012).



Social Sciences and Humanities
Research Council of Canada

Conseil de recherches en
sciences humaines du Canada

Canada

Acknowledgment

In part, the writing of this article was supported by research undertaken through a Partnership Grant from the Social Sciences and Humanities Research Council of Canada.

References

- Aoki, T., 1991. Foreword. *Inspiring Curriculum and Pedagogy: Talks to Teachers*. University of Alberta, Edmonton, pp. i–iii.
- Aoki, T., 1992. Layered voices of teaching. In: Pinar, W.F., Reynolds, W.M. (Eds.), *Understanding Curriculum as Phenomenological and Deconstructed Text*. Teachers College Press, New York, NY, pp. 17–27.
- Aoki, T., 2005. Inspiring the curriculum. In: Pinar, W.F., Irwin, R. (Eds.), *Curriculum in a New Key: The Collected Works of Ted T. Aoki*. Lawrence Earlbaum, Mahwah, NJ, pp. 357–365.
- Apple, M.W., 2013. *Can Education Change Society?* Routledge, New York.
- Archibald, J., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body and Spirit*. University of British Columbia Press, Vancouver.
- Arrows, F., Cajete, G., Lee, J., 2010. Critical Neurophilosophy and Indigenous Wisdom. *Sense Publishers*, Rotterdam, The Netherlands.
- Atleo, M.R., 2009. Understanding aboriginal learning ideology through storywork with elders. *Alberta J. Educ. Res.* 55 (4), 453–466.
- Atleo, E.R., 2011. *Principles of Tsawalk: An Indigenous Approach to Global Crisis*. University of British Columbia Press, Vancouver.
- Ayers, W., Kumashiro, K., Meiners, E., Quinn, T., Stovall, D., 2010. *Teaching Toward Democracy: Educators as Agents of Change*, first ed. Paradigm Publishers, Boulder, CO.
- Bal, M., 2002. *Traveling Concepts in the Humanities: A Rough Guide*. University of Toronto Press, Toronto.
- Baldacchino, J., 2009. *Educating Beyond Education: Self and the Imaginary in Maxine Greene's Philosophy*. Peter Lang, New York, NY.
- Battiste, M., Henderson, Y.J., 2009. Naturalizing indigenous knowledge in eurocentric education. *Can. J. Native Educ.* 32 (1), 5–18.
- Battiste, M., 2013. *Nourishing the Learning Spirit*. Purich, Saskatoon, SK.
- Bereiter, C., Scardamalia, M., 2016. "Good Moves" in knowledge-creating dialogue. *QWERTY* 11 (2), 12–26.
- Biesta, G., Stengel, B., 2016. Thinking philosophically about teaching. In: Gitomer, D., Bell, C. (Eds.), *Handbook of Research on Teaching*. AERA, Washington, DC, pp. 7–67.
- Biesta, G.J.J., 2007. Why "what works" won't work: evidence-based practice and the democratic deficit of educational research. *Educ. Theor.* 57 (1), 1–22.
- Biesta, G.J.J., 2013. *Beautiful Risk of Education*. Paradigm Publishers, Boulder, CO.
- Biesta, G.J.J., 2018. From the editor: the struggle for education. *Educ. Theor.* 67 (6), 653–656.
- Bingham, C., Sidorkin, A.M., 2004. *No Education Without Relation*. Peter Lang, NY, New York.
- Bruner, J., 1990. *Acts of Meaning*. Harvard University Press, Cambridge, MA.
- Cajete, G.A., 2015. *Indigenous Community: Rekindling the Teachings of the Seventh Fire*. Living Justice Press, St. Paul, MN.
- Carter, M.R., Prendergast, M., Belliveau, G. (Eds.), 2015. *Drama, Theatre and Performance Education in Canada: Classroom and Community Contexts*. Canadian Association for Teacher Education/Canadian Society for the Study of Education, Ottawa, ON.
- Chambers, C., 2008. Where are we? Finding common ground in a curriculum of place. *J. Can. Assoc. Curric. Stud.* 6 (2), 113–128.
- Cherkowski, S., Walker, K. (Eds.), 2018. *Perspectives on Flourishing*. Lexington, Lanham, MD.
- Cochran-Smith, M., Lytle, S., 2009. *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press, New York, NY.
- Cohen, W., 2001. The spider's web: creativity and survival in dynamic balance. *Can. J. Native Educ.* 25, 140–148.
- Delaney, M.K., Mayer, S.J., 2021. In: *In Search of Wonderful Ideas: Critical Exploration in Teacher Education*. Teachers College Press, New York, NY.
- Dewey, J., 1910. *How We Think*. Dover, New York, NY.
- Dewey, J., 1916. *Democracy and Education*. Free Press, New York, NY.
- Dewey, J., 1929. *The Quest for Certainty*. Putnam, New York, NY.
- Dewey, J., 1934. *Art as Experience*. Capricorn Books, New York, NY.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York, NY.
- Dewey, J., 1943. *The School and Society*. University of Chicago Press, Chicago, IL.
- Dion, S., 2009. *Braiding Histories: Learning From Aboriginal Peoples' Experiences and Perspectives*. University of British Columbia Press, Vancouver.
- Dissanayake, E., 2015. Aesthetic primitives: fundamental biological elements of a naturalistic aesthetics. *Aisthesis* 8 (1), 6–24.
- Doll, W., 2002. Ghosts and the curriculum. In: Doll Jr., W.E., Gough, N. (Eds.), *Curriculum Visions*. Peter Lang, New York, NY, pp. 23–70.
- Doll, W., 2013. The four R's—an alternative to the Tyler rationale. In: Flinders, D.J., Thornton, S.J. (Eds.), *Curriculum Studies Reader*, fourth ed. Routledge, New York, NY, pp. 215–222.
- Doll, M., 2017. *The Mythopoetics of Currere: Memories, Dreams, and Literary Texts as Teaching Avenues to Self-Study*. Routledge, New York, NY.
- Donald, D., 2009. Forts, curriculum, and indigenous métissage: imagining decolonization of aboriginal-Canadian relations in educational contexts. In: *First Nations Perspectives: The Journal of the Manitoba First Nations Education Resource Centre*, vol. 2, pp. 1–24 (1).
- Donald, D., 2012. Forts, curriculum, and ethical relationality. In: Ng-A-Fook, N., Rottman, J. (Eds.), *Reconsidering Canadian Curriculum Studies: Provoking Historical, Present, and Future Perspectives*. Palgrave Macmillan, New York, pp. 39–46.
- Donald, D., 2019. Homo economicus and forgetful curriculum: remembering other ways to be a human being? In: Tomlins-Jahnke, H., Styres, S., Lilley, S., Zinga, D. (Eds.), *Indigenous Education: New Directions in Theory and Practice*. University of Alberta Press, Edmonton, pp. 103–128.
- Duckworth, E., 2006. *The Having of Wonderful Ideas and Other Essays on Teaching and Learning*. Teachers College Press, New York, NY.
- Eisner, E., 2002. *The Arts and the Creation of Mind*. Yale University Press, London.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum.
- Gadamer, H.G., 2000. *Truth and Method*. Continuum, New York, NY. First published 1960.
- Gaudelli, W., Hewitt, R., 2010. The aesthetic potential of global issues curriculum. *J. Aesthetic. Educ.* 44 (2), 83–99.
- Granger, D., 2006. John Dewey, Robert Pirsig, and the Art of Living: Revisioning Aesthetic Education. Palgrave Macmillan, New York, NY.
- Green, B., Reid, J., 2008. Method(s) in our madness? Poststructuralism, pedagogy and teacher education. In: Phelan, A., Sumsion, J. (Eds.), *Critical Readings in Teacher Education: Provoking Absences*. Sense, Rotherdam, pp. 17–31.
- Greene, M., 1984. Excellence, meanings, and multiplicity. *Teach. Coll. Rec.* 86 (2), 283–297.
- Greene, M., 2000. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*. Jossey-Bass, San Francisco, CA.
- Groundwater-Smith, S., Mockler, N., 2009. *Teacher Professional Learning in an Age of Compliance*. Springer, Rotterdam, The Netherlands.
- Grumet, M., 1988. *Bitter Milk: Women and Teaching*. University of Massachusetts Press, Boston, MA.
- Haig-Brown, C., 2010. Indigenous thought, appropriation, and non-aboriginal people. *Can. J. Educ.* 33 (4), 925–950.
- Hansen, D.T., 2005. Creativity in teaching and building a meaningful life as a teacher. *J. Aesthetic. Educ.* 39 (2), 57–68.
- Hansen, D.T., 2011. *The Teacher and the World: A Study of Cosmopolitanism as Education*. Routledge, New York, NY.

- Hare, J., 2016. "All of our responsibility": instructor experiences with required indigenous education courses. *Can. J. Native Educ.* 38 (1), 101–120.
- Hargreaves, A., Shirley, D., 2009. The persistence of presentism. *Teach. Coll. Rec.* 111 (11), 2505–2534.
- Irwin, R., de Cosson, A., 2004. *A/R/Tography: Rendering Self Through Arts-Based Living Inquiry*. Pacific Educational Press, Vancouver, BC.
- Jardine, D., Friessen, S., Clifford, P., 2006. *Curriculum in Abundance*. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Jardine, D., 2012. *Pedagogy Left in Peace: On the Cultivation of Free Spaces in Teaching and Learning*. Continuum Books, New York.
- Johnson, M., 2007. *The Meaning of the Body: Aesthetics of Human Understanding*. University of Chicago Press, Chicago, IL.
- Kanu, Y., 2011. *Integrating Aboriginal Perspectives Into the School Curriculum: Purposes, Possibilities, and Challenges*. University of Toronto Press, Toronto, ON.
- Kerdeman, D., 2003. Pulled up short: challenging self-understanding as a focus of teaching and learning. *J. Philos. Educ.* 37 (2), 293–308.
- Kincheloe, J., 2008. *Knowledge and Critical Pedagogy*. Springer.
- Kumar, A., 2013. *Curriculum as Meditative Inquiry*. Palgrave Macmillan, New York.
- Leggo, C., 2007. Living poetically: pensées on literacy and health. *CCAHE J.* 6, 3–12.
- Lowan-Trudeau, G., 2014. Considering ecological métissage: to blend or not to blend? *J. Exp. Educ.* 37 (4), 351–366.
- Macintyre Latta, M., 2013. *Curricular Conversations: Play is the (Missing) Thing*. Routledge, New York, NY.
- Macintyre Latta, M., 2018. Curricular enactment matters reorienting Education. *J. Can. Assoc. Curric. Stud.* 16 (2), 34–51.
- Malewski, E., 2009. In: *Curriculum Studies Handbook: The Next Moment*. Routledge, New York.
- McKinley, E.A., Tuhiwai Smith, L., 2019. In: *Handbook of Indigenous Education*. Springer, Singapore.
- Ng-A-Fook, N., 2014. Provoking the very "idea" of Canadian curriculum studies as a counterpointed composition. *J. Can. Assoc. Curric. Stud.* 12 (1), 10.
- Noddings, N., 1984. *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press, Berkeley.
- Noddings, N., 2012. *Peace Education: How We Come to Love and Hate War*. Cambridge University Press, Cambridge.
- Noddings, N., 2013. *Education and Democracy in the 21st Century*. Teachers College Press, New York, NY.
- Nussbaum, M., 2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton University Press, Princeton, NJ.
- O'Loughlin, M., 2006. *Embodiment and Education: Exploring Creatural Existence*. Springer, Dordrecht.
- Palmer, P., 1998. *The Courage to Teach*. Jossey-Bass, San Francisco, CA.
- Phelan, A.M., Pinar, W.F., Ng-A-Fook, N., Kane, R. (Eds.), 2020. *Reconceptualizing Teacher Education: A Canadian Contribution to a Global Challenge*. University of Ottawa Press, Ottawa.
- Phelan, A., 2015. *Teacher Education and Curriculum Theorizing: Complicating Conjunctions*. Routledge, London.
- Pinar, W.F., Grumet, M.R., 2015. *Toward a Poor Curriculum*, third ed. Educator's International Press, Kingston, NY.
- Pinar, W.F., Reynolds, W.M., Slattery, P., Taubman, P.M., 1995. *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. Peter Lang, New York, NY.
- Pinar, W.F., 1975. In: *Curriculum Theorizing: The Reconceptualists*. McCutchan, Texas.
- Pinar, W.F., 1978. Notes on the curriculum field. *Educ. Res.* 7 (8), 5–12.
- Pinar, W.F., 2009. *The Worldliness of a Cosmopolitan Education: Passionate Lives in Public Education*. Routledge, New York, NY.
- Pinar, W.F., 2010. Notes on a blue guitar. *J. Educ. Controv.* 5 (1). Available at: <https://cedar.wvu.edu/jec/vol5/iss1/18>.
- Pinar, W.F., 2011. The Character of Curriculum Studies: Bildung, Currere, and the Recurring Question of the Subject. Palgrave, New York, NY.
- Pinar, W.F., 2015. *Educational Experience as Lived: Knowledge, History, Alterity*. Routledge, New York.
- Pinar, W.F., 2019. *What is Curriculum Theory?*, third ed. Routledge, New York.
- Pinar, W.F., 2022. *A Praxis of Presence in Curriculum Theory Advancing Currere Against Cultural Crises in Education*. Routledge, New York.
- Ragoonaden, K., Mueller, L., 2017. Culturally responsive pedagogy in higher education: indigenizing curriculum. *Can. J. High Educ.* 47 (2), 22–46.
- Ragoonaden, K. (Ed.), 2015. *Mindful Teaching and Learning: Developing a Pedagogy of Well-Being*. Lexington Books, Lanham, MD. An Imprint of Rowman & Littlefield.
- Restoule, J.P., Gruner, S., Metatawabin, E., 2013. Learning from place: a return to Traditional Mushkegowuk ways of knowing. *Can. J. Educ.* 36 (2), 68–86.
- Rocha, S.D., 2020. *The Syllabus as Curriculum*. Taylor and Francis, London.
- Rodgers, C., 2021. *The Art of Reflective Teaching: Practicing Presence*. Teachers College Press, New York.
- Scardamalia, M., 2017. Foreword: pedagogical innovations and new knowledge media. In: Hung, C.C., Divaharan, S. (Eds.), *Teaching and Learning Framework*. National Institute of Education, Singapore.
- Schubert, W.H., 1986. *Curriculum: Perspective, Paradigm, and Possibility*. Macmillan Publishing, New York, NY.
- Seidel, J., Jardine, D., 2016. *The Ecological Heart of Teaching: Radical Tales of Refuge and Renewal for Classrooms and Communities*. Peter Lang, New York, NY.
- Sidorkin, 2002. *Relational Pedagogy*. Peter Lang, New York, NY.
- Smith, D., 2006. *Trying to Teach in a Season of Great Untruth: Globalization, Empire and the Crises of Pedagogy*. Sense, Rotterdam.
- Stengel, B., 1991. *Just Education: The Right to Education in Context and Conversation*. Loyola University Press, Chicago, IL.
- Stengel, B., 2016. Educating homo oeconomicus? "The disadvantages of a commercial spirit" and the discourse of school failure. *Educ. Theor.* 66 (1), 245–261.
- Stitzlein, S.M., 2020. *Learning How to Hope: Reviving Democracy Through Our Schools and Civil Society*. Oxford University Press.
- Strong-Wilson, T., 2021. *Teachers' Ethical Self-Encounters With Counter-Stories in the Classroom: From Implicated to Concerned Subjects*. Routledge, New York, NY.
- Styres, S.D., 2017. *Pathways for Remembering and Recognizing Indigenous Thought in Education*. University of Toronto Press, Toronto, ON.
- Thayer-Bacon, B., 2004. Personal and social relations in education. In: Bingham, C., Sidorkin, A.M. (Eds.), *No Education Without Relation*. Peter Lang, New York, NY, pp. 165–179.
- Tierney, R., 2018. Toward a model of global meaning-making. *J. Lit. Res.* 50 (4), 397–422.
- Van Manen, M., 2015. *Researching Lived Experience: Human Science for an Action Sensitive Pedagogy*. Routledge, New York, NY.
- Varela, F., 1997. Patterns of life: intertwining identity and cognition. *Brain Cognit.* 34, 72–87.
- Waks, L.J., 2009. Inquiry, agency, and art: John Dewey's contribution to pragmatic cosmopolitanism. *Educ. Cult.* 25 (2), 115–125.
- Wearing, J., Ingersoll, M., DeLuca, C., Bolden, B., Ogden, H., Christou, T.M. (Eds.), 2020. *Key Concepts in Curriculum Studies: Perspectives on the Fundamentals*. Routledge, New York, NY.
- Wozniak, R.H., Fischer, K.W. (Eds.), 1993. *Development in Context: Acting and Thinking in Specific Environments*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.

Relevant websites

- The BIPOC Project: <https://www.thebipocproject.org>. A Black, Indigenous, and People of Color Movement building authentic and lasting solidarity among BIPOC in order to undo invisibility and anti-Blackness, dismantling white supremacy and advancing racial justice.
- Black Lives Matter: <https://blacklivesmatter.com>. Founded in 2013 in response to the acquittal of Trayvon Martin's murderer. Black Lives Matter Global Network Foundation, Inc. is a global organization in the US, UK, and Canada, whose mission is to eradicate white supremacy and build local power to intervene in violence inflicted on Black communities by the state and vigilantes.
- Convention on Biological Diversity: <https://www.cbd.int/convention/>. Building a Shared Future for All Life on Earth, United Nations Headquarters, New York, 21 May 2021.

Global Initiative for Justice, Truth and Reconciliation: <https://gijtr.org/about-us/>. Bringing advocates, activists, and practitioners from nine global organizations with expertise in a range of fields to assist communities in or emerging from conflict in creating just and peaceful futures.

NASA Global Climate Change: <https://climate.nasa.gov/news/645/climate-change-may-bring-big-ecosystem-changes/>. Documenting vital signs of the planet.

The National Centre for Truth and Reconciliation (NCTR): <https://nctr.ca/about/>. A place of learning and dialogue in Canada where the truths of Residential School Survivors, families and communities are honored and kept safe for future generations. The NCTR educates Canadians on the profound injustices inflicted on First Nations, Inuit and the Métis Nation by the forced removal of children to attend residential schools and the widespread abuse suffered in those schools.

Orange Shirt Day: <https://www.orangeshirtday.org>. A legacy of the St. Joseph Mission (SJM) Residential School (1891–1981) Commemoration Project and Reunion events that took place in Williams Lake, BC, Canada, in May 2013. The annual Orange Shirt Day on September 30th is an opportunity to create meaningful discussion about the effects of Residential Schools and the legacy they have left behind.

SOGI 1 2 3: <https://www.sogieducation.org>. Resources for educators and parents towards making schools inclusive and safe for students of all sexual orientations and gender identities (SOGI).

Worldometer: <https://www.worldometers.info/coronavirus/>. COVID-19 Coronavirus Pandemic reported cases and deaths affecting 220 countries and territories.

Curriculum and identity options

Rachel M. Heydon and Carla Ruthes Coelho, Western University, School of Education, John George Althouse, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Curriculum and identity options	381
Emergence of curriculum studies: identities on the horizon	382
Curriculum reconceptualization and an emphasis on identity	383
Feminist theorizing and identities	384
Phenomenological theorizing and identity	384
Identity and curriculum: the influence of socio-cultural theorizing	386
Socio-material approaches to identities and curriculum	386
Concluding thoughts	389
References	389

The conjunction of education and identity is a familiar trope; it is the core of the bildungsroman that characterize English curricula of the global North: education forms one's identity on the page and through the page. The curricula of this education can be formal and informal and are often connected to identities of survival and belonging. We are reminded, for instance, of Lucy Maude Montgomery's classic Canadian Anne novels (e.g., 1908), where protagonist Anne Shirley clutched to poetry to sustain her through the trauma of being an orphan, her adolescence in an adoptive family, and the pursuit of academic success, all to prove herself worthy of calling Green Gables home. More recently, there is Elena Ferrante's Neapolitan quartet (2012, 2013, 2014, 2015), an ocean and decades away from Anne's rural Prince Edward Island but similar in its treatment of education and identity. Set in the violence of a poor Neapolitan neighborhood in the 1960s, Ferrante's Elena Greco is schooled through the curriculum of that neighborhood, as well as the curricula of a hard-won institutional education, and at the hand of her brilliant, unschooled friend, Lila, all with effects on and implications for her identity.

There is, of course, more to say about these novels, but our point is not the novels themselves; instead, we begin an article about curriculum and identity options with a reminder of the bildungsroman to indicate the ubiquity of the concurrence between the concepts and to use it to argue that it is this very ubiquity that creates a danger of taken-for-grantedness about the potential link between curriculum and how and what identities may be. Tracing conceptualizations of identity becomes central to thinking with identities and what it might mean to have *identity options*. As such, we offer in this article opportunities for readers to see identity in/out of curriculum studies as a concept that has been diversely treated with each treatment having its own implications for curricular research and practice. We begin by offering a primer on conceptualizing curriculum and identity options. We then move through a history of curriculum and the ways in which each movement in the field has brought various ontological, ethical, and epistemological resources to the conversation of identity. We present a version of history that is tripartite in its construction: the emergence of the field to reconceptualism(s), including the rise of socio-cultural and poststructural theories, and into posthumanities. The construction is perhaps overly teleological yet provides a trajectory of theorizing that may be helpful to apprehending some resources of import for thinking with identities and curriculum.

Curriculum and identity options

Curriculum is a concept that has been diversely conceptualized. To deal with the multifaceted nature of curriculum conceptualizing, we call for qualifying what one means by curriculum, as we do in this article. Thus, we might speak of identity and *institutional curriculum*, (i.e., shared and abstract goals for education that respond to schools' socio-cultural and political contexts and demands (Doyle, 1992)), *programmatic curriculum* (i.e., written curriculum documents and materials for use in teaching (Doyle, 1992)), *classroom curriculum* (i.e., what actually happens in the classroom (Doyle, 1992)), *hidden curriculum* (i.e., values and messages that are implicitly learned through curriculum, but that are not explicitly taught (Apple, 2004)), *null curriculum* (i.e., that which is absent from being explicitly and implicitly taught, with its absence having implications for students' learning, knowledge, and values (Eisner, 1979)), *lived curriculum* (i.e., the phenomenological experience of a curriculum (Aoki, 1993)), and so on, marking specifically the concept and its origin(s) in the literature through our citational practices.

Much like curriculum, there is no single response in the literature to what identity is, what it does, or how it works. However, many theorists in diverse areas of scholarship have scrutinized the term. Gleason (1983) explained that its use is relatively recent in the social sciences. According to him, in the 1930s, when the first edition of the *Encyclopedia of Social Sciences* (Seligman and Johnson, 1930) was published, the term identity was not included in it, and, by 1968, only two entries contained the term. One conceptualized it by connecting it to psychosocial phenomenon (Erikson, 1968) and another to political affiliation (Buchanan, 1968). In 1983 when Gleason published the treatment of identity to date, he noted that though many areas of research had attempted to explain what identities are and how they are produced and expressed, there was no consensus on its definition or

conceptualization; a situation that our literature review in this article finds is still the case. There are numerous conceptions and uses of the term identity; conceptions that Gleason's research indicated were focused either on the individual and internal, psychological processes or more external, social processes of identity formation. He noted that Erikson (1963), for example, advanced a psychological approach to identity, connecting it to something internal and developmental, spanning across the lifecourse. An oppositional definition supported by sociologists is that identity is fluid, constantly changing according to context and social experiences (Gleason, 1983). As we will discuss below, identity was not on the radar in education when curriculum studies emerged; however, scholars in the field were concerned with the kinds of citizens institutional and programmatic curricula were trying to craft. The second movement of the field afforded the phenomenological conceptualization of *lived curriculum*, bringing identity to the fore. Shortly thereafter, socio-cultural theorizing birthed the notion of identity options.

Identity options refers to the ways in which all forms of curricula affect how learners can see themselves and their place in the world. This concept was extensively explored by Cummins (2000) in his work with culturally and linguistically diverse learners. Cummins (2000, 2001) argued that interactions between educators and learners are permeated with power relations with the potential to expand or constrain learners' academic engagement and affect their sense of self as competent learners. Cummins (2009) conceptualized power relations in learning spaces as collaborative or coercive. *Coercive relations* indicate control, hindering identity options for students; *collaborative relations* allow for power to be distributed among educators and learners, enabling recognition of learners' identities as valuable resources for curriculum making. One form of expanding learners' identity options is through *identity texts* (e.g., Cummins, 2006). Identity texts are said to promote equitable pedagogies that affirm diversity and learners' competence, expanding their identity options by allowing multiple modes of communication (e.g., visual, gestural), and welcoming diverse cultural knowledges and languages in students' academic productions. Cummins and Early (2011) explored the potential of identity texts in framing language learners in a positive light in multilingual educational settings. Their studies illustrate how learners benefit from pedagogies that enable them to express themselves in their own terms and through a multitude of resources.

Other researchers have also taken up the concept of identity options in investigations of multimodal literacy practices in intergenerational programs (Heydon, 2007), early literacy curriculum (Heydon et al., 2015), and transnational education (Zhang et al., 2020). Heydon and O'Neill's (2016) introduction of singing as a multimodal literacy practice into an intergenerational art program, for example, demonstrates that "expansive literacy options correlate with expansive identity options" (p. 136). The singing and art infused curriculum allowed elders and children to move across modes, experimenting with meaning-making practices that they might not have experienced without their peers' encouragement or in a narrower programmatic curriculum. The authors highlighted how "the bodily and vocal gestures that accompany singing articulate meanings and emotional states that when combined constitute a particular narrative" (p. 133). These narratives generated new forms of expression that allowed participants to define and share "aspects of who they are, who they were in the past, and who they might be in the future" (p. 133). Heydon and O'Neill's research illustrates how multimodal intergenerational curricula can offer opportunities for identity negotiation and expansion through relationship building and semiotic experimentation. Pedagogies that prioritize identity options aim at empowering students by creating a sense of belonging and expanding opportunities for them to express and pursue identities through myriad resources within and beyond canonical knowledge.

To understand how the concept of identity options sits within curriculum studies, its epistemological and ontological roots, commitments, and where its constraints might reside, we now turn to a discussion of identity and curriculum in the literature, moving from the early days of the curriculum studies field to contemporary theorizing.

Emergence of curriculum studies: identities on the horizon

The early history of curriculum as a field of study did not explicitly name identity. Identity was surely, however, part of the hidden curriculum. Much of the early field concentrated on the development of educational systems that would promote students' efficient and effective acquisition of skills and knowledge. Consider, for example, the twinned foci of scientific knowledge and economic growth from the United States which drove the shape of curriculum studies in the 1900s (Ingersoll et al., 2019). Bobbitt (1916, 1918) and Tyler (1949) are remarkable names of an era in which curriculum scholarship followed a course of "control, measurement" (Null, 2008, p. 479) and outcomes that were at the heart of many academic fields, though Dewey (1902) offered an early hint of the attention to identity and lived curriculum that was to come.

In the beginning of the 20th century, modern science advanced quickly, and education followed by adopting a discourse and logic that were compatible with the happenings of that context. Questions of curriculum were mainly those of programmatic curriculum and were answered "in the language of science" (Null, 2008, p. 479), with key concepts being objectivism and efficiency in relation to curriculum *development* and *transmission* (Doyle, 1992). Bobbitt (1918) envisioned that curriculum should be directed toward creating the kind of adults that could take their place within an industrial society (Null, 2008). His major work, *What the Schools Teach and Might Teach* (1916), outlines how observing adult life activities could help determine programmatic curriculum components that would build such citizens. In this orientation to curriculum, a technology of curriculum development and application could produce direct, predictable effects on people, including creating desired citizens (Null, 2008).

Also working in the vein of curricular efficiencies directed toward prescribed ends, is the oft-cited (e.g., Finder, 2004; Wraga, 2017) and analyzed (e.g., Kliebard, 1970, 1995) Tylerian rationale (Tyler, 1949). Tyler granted that a scientific process of teachers and school administrators responding objectively to a prescribed set of questions would, within his rationale, result in the

identification of “an effective educational program” (p. 125). The rationale was predicated on the idea that curricular experiences, once organized and planned, could generate the behavioral outcomes curriculum makers intended (Deng and Luke, 2008). With both Bobbit’s (1916, 1918) and Tyler’s (1949) models for curriculum development, we see curriculum conceptualized as something that could, with the right design, move smoothly from what was intended to what would be enacted. Curriculum was *done to* learners with the production of prescribed citizens or workers part of the hidden curriculum.

Contemporary to Tyler and Bobbit was, of course, Dewey. Although Dewey did not explicitly speak of identity, his focus on learners and their knowledge and experience opened the door for its later conceptualizations to come. Works such as *The Child and the Curriculum* (1902), were anathema to efficiency models of curriculum that separated knowledge from experience and the child from society. His work also renders problematic any idea of programmatic curriculum being directly translated into classroom curriculum; for instance, in talking of the “course of study”, Dewey expressed, “no such thing as imposition of truth from without ... All depends upon the activity which the mind itself undergoes in responding to what is presented from without” (p. 31). In statements like this, Dewey foreshadows what Schwab (1973) would later name the curricular *commonplaces* of educator and learner and shows some of the complexity of curriculum-making.

Dewey did not explicitly deal with identity as we would recognize the concept today; nonetheless, work like *The Child and the Curriculum* (1902) drew attention to learners and their experiences, demanding that learners be at the center of curriculum, ideas that were important for later identity theorizing in education. Of his foci, Null (2003) indicated that Dewey “admonished teachers not only to focus on subject-matter, but also to concern themselves with the needs, interests, experiences, and backgrounds of their students” (p. 66). Bringing learners into view in curriculum-making was expressed by Dewey and other *pragmatic progressives* as an important aspect of curriculum inquiry, alongside a focus on problem-solving and democratic participation (Wraga, 2019). According to Reese (2001), American progressivism also included child-centered ideas that challenged traditional schooling and rearing practices reflecting some of the shifts occurring in that period, such as changes in “family size ... gender roles ... and the softening of religious orthodoxy within Protestantism” (p. 3). Scholars such as Ingersoll et al. (2019), articulated one of the major differences between efficiency curricular models (e.g., Tylerian rationale) and Deweyan progressivism as a movement from viewing students as passive receptors of skills to understanding the child’s “experience of schooling” (p. 2).

Modernism and scientism placed emphasis on linearity, predictability, and reproducibility to understand human behavior and development. A critical understanding of identity had no place in this form of curriculum theorizing. Social movements and curriculum reconceptualization were about to change this.

Curriculum reconceptualization and an emphasis on identity

The reconceptualizing of curriculum was a turn toward the learner, their social positioning and relationship with/in curriculum writ large. It was a catalyst for explicit considerations of identity and a challenge to objectivism and dominant positivist paradigms of curriculum development (e.g., Pinar, 2008). The movement ignited multidisciplinary scholarship that considered curriculum in its interrelation with social concepts and identity markers like race, gender, and culture (Pinar, 2004). Robson and Sumara (2019) explained, for instance, that “the idea that our identities are at least in part the product of our socio-cultural environment grew from several streams of thought” (p. 146), and they pointed to grassroots movements led by feminist, black and LGBTQ+ activists in the 1970s, who pushed the boundaries of objectivism and universal truths by creating spaces in which previously minoritized experiences were debated and acknowledged.

Around the world, civil rights movements gained force while critical and then post-structural theories spread in/through curriculum studies and beyond. These new epistemic resources fueled a (re)consideration of seeing and doing education. Critical theory, for example, scrutinized how schooling and curriculum reproduced social inequalities, interrogating whose knowledges and interests were served by them. One of the most prominent names of critical pedagogy is Paulo Freire (1970), who advanced the idea that education should emancipate individuals instead of operating from a banking model in which the knowledge that education deposits into the learner is what one may withdraw—independent of who the learner is or their socio-cultural situation. Freire put into practice many of the tenets of critical theory in developing literacy projects where all curricula were based in learners’ own cultural and linguistic contexts. Literacy curriculum, in this way, was directed toward affirming learners’ experiences and identities while creating educative opportunities that raised their consciousness of how socio-political issues affected their lives. The question-to-come of identity options was also opened here with the focus on curriculum as related to power.

Integral to this movement was the questioning of curriculum and cultural and racial identities (Jackson, 1992). Robson and Sumara (2019), for instance, asserted that schools were spaces of competing values, where *hidden curriculum* operated through biases in the system, communicating that some knowledge was more desirable and desired by the cultural elites than others. Castenell and Pinar (1993) dug deeply into investigating *curriculum as racial text* as they sought to theorize representation, identity, and difference in education. In this work, they conceptualized identity as “a gendered, racialized, and historical construct” (p. 4), highlighting its situated socio-political nature. They also made the link between identities and curriculum explicit by considering curriculum a “significant form of representation” (p. 2), stating that “identity formation is constructed and expressed through representation, that is, the construction of ‘difference’, and negotiated in the public sphere” (p. 4). Further, Pinar (2008) expounded that multicultural education grew from an effort to acknowledge cultural diversity; however, he cited McCarthy’s (1993) critique alerting to the dangers of acritical notions of multiculturalism and exposing the complexities of race. He invited curriculum theorists and educators to recognize race and how it was part of curricula, textbooks, and pedagogies. He reprized his call with Castenell and Pinar (1993)

demanding that scholars engage with the complexities of race, rather than trying to find single solutions to racism or neglect how race was instantiated in curriculum and permeated experiences of learners and teachers.

Feminist theorizing and identities

The reconceptualization of curriculum was equally forwarded by critical feminist scholarship that laid the groundwork for future conversations of identity. This scholarship challenged education's sense of normalcy and (re)production of constraining gender and sexual roles. A key example is [Luke and Gore's \(1992\)](#) *Feminisms and Critical Pedagogy*, a collection that questioned progressive education and critical pedagogy for their lack of attention to the diverse ways language, power and knowledge operate depending upon one's social positioning and subjectivities. Bringing to the fore foundational feminist pedagogues of the era, each article confronted traditional understandings of critical pedagogy by exploring the localized experiences of women in academia. Challenging the generalizability of theoretical developments and pedagogies of masculine authorship, including a masculinist notion of emancipatory critical theory, the collection focused on the complexities and multiplicities of subjectivities, exposing how dominant educational and political discourses impact women's, children's, and racialized bodies. Their poststructural feminist analysis entwined research practices, theoretical thinking, and personal accounts, bringing identity to the forefront.

Collectively organizing, feminists of the day fashioned spaces for women to speak and be heard. The concepts of *voice* and its absence appear frequently in scholarship of the movement; for instance, [Lewis's \(1993\)](#) *Without a Word: Teaching Beyond Women's Silence* drew from her teaching and academic experiences to reveal "the lived contradiction between silence as oppression and silence as revolt" (p. 2). Lewis acknowledged that women's speech is transgressive and political. As such, her writings can also be understood as acts of defiance to dominant male voices in education. [hooks' \(1989\)](#) toiled with the idea of *voice* to fill up voids in educational theory in relation to racialized women's identities. Her collection of essays, *Talking Back: Thinking Feminist, Thinking Black*, tackled racism, feminism, pedagogy, and politics through an embodied perspective where her experiences as a black woman generated theory. [Kelly \(1997\)](#) introduced questions of affect and desire into conversations of identity in *Schooling Desire: Literacy, Cultural Politics, and Pedagogy*. Leaning on her schooling experiences in rural Canada, Kelly rethought the ways in which traditionally organized literacy curricula produce identities and regulate desire. She advocated for a consideration of how pleasure might enter curricula to challenge educational inequities, and her work encouraged the inclusion of feminist orientations to curriculum theorizing, arguing that they could open avenues for considering how "knowledge, power, and desire" (p. 1) operate through curriculum and affect identities.

The branches of the reconceptualist movement in curriculum, each in its own way, brought about new modes of thinking about education, questioning its purposes, who was responsible for making curricular choices, and the implications or effects of these choices ([Eisner, 1992](#)); a contiguous branch of this movement pursued a critique of traditional curriculum theory through phenomenology, and created additional resources for conceptualizing identity. Foremost among these is Ted [Aoki's \(1993\)](#) *lived curriculum*.

Phenomenological theorizing and identity

[Aoki \(1993\)](#) argued that the curriculum landscape of the day privileged *curriculum-as-plan*. Akin to the concept of the programmatic curriculum ([Doyle, 1992](#)), Aoki's curriculum-as-plan is a curriculum created by those who reside outside of the classroom (e.g., curriculum specialists, governments, etc.). This plan is what teachers (and students) must follow; it is not neutral because it is created by people and institutions and thus enmeshed with their philosophies/values of learning and knowledge. Prioritizing some knowledge and omitting the rest is sedimented in statements of intent, curricular goals, and objectives. Aoki critiqued the ways curriculum-as-plan divided knowledge into "epistemic categories" (p. 256), separating arts and sciences and creating a "curriculum topography" (p. 256) that represented Western Cartesian ways of organizing the world. Aoki attested that students' experiences often diverge from curriculum-as-plan. Whereas curriculum-as-plan is developed outside the classroom, using techno-scientific language to describe curriculum goals and intentions and "written for faceless people in a homogeneous realm" (p. 261), the lived curriculum is phenomenological, "embodied in the very stories and languages people speak and live" (p. 261).

[Aoki's \(1993\)](#) lived curriculum is "not the curriculum as laid out in a plan, but a plan more or less lived out" (p. 257). Lived curriculum refers to learners' and educators' curricular experiences or how they experience curricula. This curriculum, according to Aoki, is complex as there are likely as many lived curricula as learners' identities. Identities, for Aoki, are effects of curriculum in relation to learners, educators, in ongoing *becomings* these becomings are counter to a fixed, immutable conceptualization of identity. Identities and curriculum emerge in the *in betweenness* and tensions that exist in the curricular landscape of curriculum-as-plan and lived curriculum. This space produces possibilities for doing identities and curriculum in new ways.

According to [Aoki \(1993\)](#), to robustly comprehend curricula, one must recognize the tensions of curricular landscapes and how they create a multiplicity that "grows in the middle" (p. 259). Citing [Deleuze and Parnet \(1987\)](#); [Aoki \(1993\)](#) described that identities are not pre-existing but "ongoing effects of our becomings in difference" (p. 260). That is, identities emerge from relationships with others; they grow in the middle. Perhaps Aoki's separation of lived curricula and curriculum-as-plan reflects the very Cartesian separation that he rejects. However, this split of curricula promotes how lived curriculum and curriculum-as-plan are always enabling, producing, reproducing, or constraining these becomings. Aoki evoked postmodernist thinking, citing [Lyotard's \(1984\)](#) questioning of grand narratives of unity, rationality, and progress, and called for curriculum theory to do the same. Aoki suggested that curriculum should embrace the multiplicity of curriculum stories "legitimizing thoughtful everyday narratives" (p. 263) that unsettle curriculum, education structures, and the objectivity and separability (between categories of knowledge,

for example) in which they are founded. Furthermore, Aoki argued educators must move away from privileging curriculum-as-plan toward finding creative, open spaces in between curricular landscapes, that offer possibilities for learning to take different directions based on who learners and educators are and the contexts in which they live, giving “way to a more open landscape that offers possibilities by, in part, giving legitimacy to the wisdom held in lived stories of people who dwell within the landscape” (p. 267). Aoki called for knowledge and wisdom to be considered together through the recognition of educators’ competency as curriculum makers.

Lived curriculum was related to other crucial curricular theorizing of the movement that had implications for identity. In 1975, Pinar provided additional resources for shifting curriculum from techno-rationality by introducing the concept of *currere*. *Currere* is the Latin root of curriculum, meaning “the running (or lived experience) of the course” (Pinar, 2004, p. xiii). The idea of curriculum as a racecourse emphasized its continuous and dynamic trajectory, but enabled understanding curriculum theorizing, producing, and experiencing as a verb, focusing on the lived, embodied and negotiated experiences of learners. Pinar (2019) stated that curriculum is a “complicated conversation” (p. 50), in which “curriculum as plan” (p. 51) and curriculum as lived experience are often in relation to each other. Pinar explained that *currere* as a verb in its infinitive form, includes “actions, processes, and consequences” (p. 51), whereas perceiving curriculum as a noun eliminates its complexity, investing in the traditional idea of curriculum as a document that establishes what learners need to know to achieve pre-established goals. Corroborating Aoki (1993), Pinar (2019) said that “curriculum as plan” (p. 51) ignores the vibrant complexities of social, political, cultural aspects of life that are constantly in motion. For Pinar, this lived curriculum is embedded in a “subjective sphere” (p. 51) that is not present in institutional goals and efficiency orders. Rather, curriculum has consequences that can disturb any possible predictions of its outcomes (much as Dewey (1902) previously taught): “the consequences of academic study are ongoing, as they are, in lived terms, social and subjective as well as intellectual. Outcomes are, finally, immeasurable” (Pinar, 2019, p. 51). *Currere* also establishes a focus on identities in their relationship with all forms of curriculum. Pinar explained that *currere* concentrates on individual learning but recalls that learning always happens in relations with others since education and curriculum experiences are socially, politically, culturally, and economically situated. Pinar emphasized the power of curricular experiences and their impact on individuals by stating, “we are changed by what we study; such educational experience encourages subjective and social reconstruction” (p. 51).

An additional element of *currere* as it relates to conceptualizing identity is *difference*. Pinar (2019) extolled the value of difference calling for attention to the ways people become who they are. He stated that every person

has a different genetic and social makeup: different upbringings, families, and caretakers, and significant others like teachers. We are also different (and similar) in terms of race, class, culture, and gender, inflected as these are by place, time, and circumstance. While informed by these and by other homogenizing forces, each of us is, or can be, distinctive. Indeed, the method of *currere* invites us to cultivate that distinctiveness (p. 51).

Currere demands that learners and educators engage in ethical and reflexive practices that cultivate a “consciousness that remembers the past with an eye on the future while focused on the present” (p. 52). The method of *currere*, according to Pinar (2019) has four phases that would allow enacting “the concept of curriculum as self-formation through academic study and dialogical encounter” (p. 52): the “regressive phase” allows learners to return to the past and remember significant experiences; the “progressive moment” requires that learners imagine their future; the “analytic moment” calls learners to contemplate past and present experiences together and find what has been “obscured” or “neglected” in their lives, opening up new possibilities for subjectivities; and the “synthetic phase” allows the person “to act anew in the private and public worlds one inhabits” (p. 50). Curriculum as a verb moves beyond considering its goals, guidelines, and proposed outcomes. Instead, curriculum as lived experience means recognizing the stories, struggles, and identities that curriculum silences, praises, reinforces, and negotiates.

The concept of *currere* influenced other branches of curriculum studies, including the post-structural turn in the field. Kincheloe (1998), for example, proposed that *post-formal* thinking in relation to *currere* could produce new pedagogical commitments for teachers and students. He saw *currere* as meaning that “self and identity should be cultivated in relation to the learning process” (Kincheloe, 2013, p. 130) allowing an analysis of how identities and social positionings are constructed through “power discourses and political structures (p. 131). In his view, *currere* and post-formal thinking would enable

students to deconstruct their personal constructions of the purpose of their schooling; it induces men to expose the ways the privileged position of masculinity has shaped their self-images; it induces heterosexuals to refigure their identities vis-à-vis confrontation with the dynamics of heterosexism and sexual preference; it induces workers to reconsider the workplace and its social dynamics as profound influences on identity formation. The postmodern condition’s crisis of meaning and identity is also a crisis of thinking. As a culture, we have little idea how our identities are shaped by power relations and the impact of such a process on how we define intelligence. Post-formal thinking about thinking appropriates *currere*’s ideological disembedding, pushing cognition into the complexity of self-production. Transcending Piagetian formalism with its disembodied abstraction and its concern for disinterested procedure, post-formal thinking about thinking encourages a running meta-dialogue, a constant conversation with self (p. 132).

Kincheloe’s words here are indicative of how poststructural curriculum studies scholars worked alongside other attempts from the 70s, 80s and 90s to understand how identities were produced and reproduced.

Identity and curriculum: the influence of socio-cultural theorizing

Identity options emerged in the 2000s while the pace and intensity of thinking with curriculum and identities increased with new socio-cultural theorizing. Important for conceptualizing identity options and a lynchpin of the research in this vein is the concept of *funds of knowledge* (e.g., González et al., 2005). The concept was developed through an ethnographic project that invited teacher-researchers to learn about the household and community knowledges of Mexican, African American and American Indigenous children in the United States. The goal was to use this knowledge for responsive, *asset-oriented* (Heydon and Iannacci, 2008) curriculum-making. Visiting the neighborhoods where children lived and interviewing their families, the researchers studied “what households actually do and how they think about what they do” (González et al., 2005, p. 10). Funds of knowledge was based on the belief that “people are competent, they have knowledge, and their life experiences have given them that knowledge” (p. ix–x), a premise that was in direct opposition to a deficit-orientated curriculum that minoritized learners, their families, and communities (e.g., Shields et al., 2005). Buoyed by the theory of funds of knowledge, educators were poised to promote curricula that capitalized on what children knew and their community identities could be used for curriculum-making.

Growing out of funds of knowledge is the concept of *funds of identity* (e.g., Esteban-Guitart and Moll, 2014a). Combining cultural psychology with socio-cultural theories of learning, Esteban-Guitart and Moll (2014a) explained that “‘identity’ is not a thing, but a social construct vaguely referring to a vastly complex set of phenomena” (p. 32). They claimed that identity is part of people’s daily lives, something every person talks about and indexes to indicate who they are as they draw on categories such as gender, professional identity, and religious identity; identity is thus seen as an essential part of people’s self-expression and self-understanding. Identities too are lived, created in interactions between people, symbols, and artifacts in context. For Esteban-Guitart and Moll, identities are “social products, cultural devices, a kind of box of tools which can be used to define oneself” (p. 35). Rather than considering identity as a phenomenon that occurs exclusively in people’s minds, or a fixed trait, identity here is inextricably connected to culture, mediated, distributed among people, artifacts, contexts, and the like. People identify themselves always in relation to others, their known cultural worlds and social artifacts. Esteban-Guitart and Moll further explained that activities and artifacts such as songs and worship, become tools through which individuals make and express their identities: “Identity is in things as well as people In an important way, a person’s self can be viewed as a dynamic organization of various resources, socially, historically, and culturally created” (p. 37).

Funds of identity conceptualizes how people and communities come to build identities and express those identities based on what is available to them through their funds of knowledge and cultural capital (Esteban-Guitart and Moll, 2014b). Whereas language, religion, and geographic locations are part of funds of knowledge, when these elements become resources for self-identification, they become funds of identity. Funds of identity are unfixed; Esteban-Guitart and Moll (2014a) claimed that as people acquire life experiences, more funds of knowledge are available to them and, in turn, new resources for identifying themselves. Esteban-Guitart and Moll’s (2014b) theorizing of *lived experience* is central to identity. Lived experience, they expressed, is conditional since it emerges in the encounters between person and environment, previous experiences, and knowledges in contact with new possibilities that a social and cultural context offers over time. How a learner feels and understands the situations they live is their lived experience; it is an aggregation of internal and external factors. Esteban-Guitart and Moll (2014a) cataloged identity resources as follows: (a) “geographical funds of identity” (place, nationality); (b) “practical” (work or activity); (c) “cultural” (“national flags or social categories”); (d) “social” (“significant others such as relatives, friends”); (e) “institutional” (family, religion) (p. 38).

The premise of funds of knowledge in education works in parallel with funds of identity: by learning about and documenting students’ experiences, their funds of knowledge can be included in curricula and their funds of identity affirmed (Esteban-Guitart and Moll, 2014a). Esteban-Guitart and Moll (2014b) further suggested pedagogies for prioritizing expansive funds of identity. Like the multimodal learner and community-focused texts of identity texts (e.g., Cummins, 2006), Esteban-Guitart and Moll (2014a) called for pedagogies like the *significant circle strategy* in which the learner draws and/or writes in a circle what is important to them, such as family members, hobbies and objects. In all, Esteban-Guitart and Moll (2014b) encouraged educators to see learners’ funds of knowledge and identities as curricular resources and schools as spaces that reflect the community and the people who are part of it, bridging and valuing the different contexts that compose learners’ lives.

Funds of knowledge and identity are concepts that are like the aforementioned identity options. Both are founded on socio-cultural theory, prioritize the relationship between socio-cultural and semiotic contexts and the learner, and actively work against the minoritization of knowledges and languages. Moreover, the work of funds of knowledge and identity, like that of Cummins and Early’s (2011) identity texts and Heydon’s (2013) multimodal intergenerational curricula, considers how languages, modes, and out-of-school socio-cultural knowledge can be resources for school curricula and lead to the empowerment of people who might otherwise be constructed as pathological or deficient (e.g., Heydon and Iannacci, 2008).

Socio-material approaches to identities and curriculum

Most recently, socio-material theories and orientations have entered educational research, with major implications for conceptualizing curriculum and identity. One of these orientations is *posthumanities* which, according to Braidotti (2019), is an emergent field of study that grew in the contact zone between post-anthropocentrism and posthumanism, where the first offers a critique of a hierarchy of species in which humans are at the top and the second debunks the notion that the human is the “universal measure of all

things" (p. 32). The Anthropocene is what Haraway (2015) explained as the "immense irreversible destruction" (p. 160) of the planet. The era is also described by Braidotti (2019) as characterized by rapid technological developments and increased economic and social disparities. Education scholars engaged with ideas sympathetic to or directly announced as premised in posthumanities, come from a range of foci including literacy education (e.g., Hackett and Somerville, 2017; Kuby et al., 2019), Indigenous education (e.g., Jones and Hoskins, 2016; Tuck and McKenzie, 2015), and mathematics education (e.g., de Freitas, 2016). Each in its own way, counter to socio-cultural approaches just discussed, this scholarship de-centers learners as the focus of curriculum. Some have argued that this decentring is not a dismissal of learners; rather, moving the focus from the human provides opportunities to appreciate the constituents and movements of the time/space/matter in which learners are implicated (e.g., Kuby and Rucker, 2016), allowing for a questioning of how the entanglements that characterize education might better produce knowledge of how to live well together (Murrin, 2016). Indeed, the industrial products of traditional curriculum become profoundly upended through this movement. Finding the ability of any curricular system to predict and control knowledge and the creation of a particular kind of citizen a "fantasy" (Edwards, 2011, p. 39), posthumanities challenges the neoliberal projects of othering, in which the constituents of this planet (including the humans who are the students) are commodities for extraction and economic advancement (Tuhiwai Smith et al., 2018).

Posthumanities brings to the fore a new conversation where ethics is inseparable from ontology and epistemology (e.g., Barad, 2007). Barad (2007) described her approach to posthumanities as having to do with "responsibility and accountability for the entanglements 'we' help enact and what kinds of commitments 'we' are willing to take on, including commitments to 'ourselves' and who 'we' may become" (p. 382). In this sense, education scholars working with posthumanities understand that their practices cannot be dissociated from the world as they intra-act with all the human and non-human inhabitants that co-constitute it. Posthumanities ethics entail a project of breaking down ideas of universal experiences. In education, this can apply to the developmentalist idea of the universal child as well as social categories of identity that fall within simplistic binary oppositions such as boy/girl, black/white, high socio-economic standing/low socio-economic standing. Braidotti (2019) advocated instead for *situated knowledges*, which refers to the idea that political subjectivities and knowledges are always entangled with time and space, which makes each experience distinctive, ongoing, not universal. Thinking curriculum through this perspective attunes educators to the relationality of knowledge production and the local specificities that constitute materially and semiotically "the kind of knowing subjects currently constructed" (p. 34). This is one example of how posthumanities offer "a frame to understand the ongoing processes of becoming-subjects in our fast-changing times" (p. 34), without focusing exclusively on the human subject.

The focus of the identity conversation with the introduction of posthumanities shifts radically from human identity that comes from what is human-made, as in the case of socio-cultural approaches, to how subjectivities are co-produced through the kind of complex entanglements outlined above. Braidotti (2019) explained subjectivities as:

not restricted to bound individuals, but is rather a co-operative trans-species effort (Margulis and Sagan, 1995) that takes place transversally, in-between nature/technology; male/female; black/ white; local/global; present/past—in assemblages that flow across and displace the binaries. These in-between states defy the logic of the excluded middle and, although they allow an analytic function to the negative, they reject negativity and aim at the production of joyful or affirmative values and projects (Lloyd, 1996; Braidotti, 2011). Poststructuralism paved the way for this approach, but the posthuman turn materializes it and composes a new ontological framework of becoming-subjects (p. 33).

Developing ideas from Queer Theory, Braidotti (2019) rejected binary oppositions and fixed identities, forwarding the idea that being is a never ending, cooperative process of creation between humans and non-humans, and "within this worldview the posthuman subject emerges as multiple and inherently differentiated" (van der Zaag, 2016, p. 331). This insistence on difference is, again, not just ontological, but ethical.

Ethics has not, as we suggest earlier in this article, always been visible in discussions of curriculum. We saw in the early days of curriculum studies a major disavowal of identity—the person as a sentient entity with a history, a family, a story, was of no consequence to the curricular machine; the machine would make the right kind of citizen and implant the right kind of knowledge into this "faceless" (Aoki, 1993, p. 261) student. Subsequent waves of critically oriented theory into curriculum studies such as socio-cultural theory and feminist theory vehemently ascribed the face to the learner, emphasizing its relevance to what curriculum produced and who it favored. These faces were faces of asymmetrical power based in social and economic structures (e.g., Apple, 2008). Phenomenology of the day highlighted the face as also an internal face, produced through the experience of living curriculum; and the early poststructural movement in curriculum played a version of "a kind of intellectual version of the hokey-cokey" (Stronach and MacLure, 1997, p. 19) where identities could be multiple and unfixed, yet also related to critical readings of social structures and markers (e.g., consider Kincheloe and Pinar alongside each other). The ethico-onto-epistemological orientations of posthumanities scholarship create opportunities to see curriculum and identity in fresh and disruptive ways.

Erica Burman's (2021) *Developments: Child, Image, Nation* is an example of scholarship sympathetic to posthumanities that creates new resources for conceptualizing identities. Burman traced the connections between developmental psychology and economic development, claiming that ideas of progress from the global North are aligned with colonial, neoliberal agendas that decontextualize childhoods and aim to create manageable, universal subjects, much like in the early days of curriculum. Burman's focus is not identities, but her postdevelopmental theorizing is an important counterpoint to previous conceptualizations of identity. Her book, for example, considered subjectivities as they relate to developmental psychology's construction of the normal child, a construct which has influenced numerous curricula (e.g., Iannacci, 2019). In contrast, Burman (2021) eschewed notions of child

(and, in our extension, identity) that are autonomous, asking instead for “situated, nuanced ... narratives” (p. 32) of children and childhoods. A companion to a relational subject is the shaking of any claims to “authentic, unmediated exhibitions of inner subjectivities (whether others’ or their own)”; there are, Burman argued “limits of what one can claim to know, and the different knowledges that can be brought to bear in different contexts” (p. 57). These ideas have implications for thinking with any conceptualization of identity that is reliant on humans as singular, rationale beings who can fully know or design themselves or their identities.

There are times in her work when Burman (2021) addresses identity specifically and in so doing further ruptures links with old thought. One of these times was when Burman expressed the difficult roots of identity and the kind of identity categories that were suggested in some critical approaches to curriculum. To this point, Burman said,

the gender, class, and cultural relations—that privilege male over female, white over black, heterosexual over homosexual, middle over working class—are deeply implicated within the technological and colonial development of nineteenth-century Europe. Cultural analysts and philosophers have pointed out that one conceptual condition for all of this was the fixing of all these categories via the notion of “identity” (p. 159).

Burman was here again concerned with how identity, as a social marker, can be taken on as a psychological phenomenon and operate through a focus on the individual and a two-dimensional view between that individual and the social order.

The above is a complaint that resonates with Haraway’s (2016) advice that the Anthropocene (what she called the Chthulucene), is a time of deep relationships of all kinds which demands *making-with*, not *self-making*. Similarly, Ringrose and Renold (2016), writing about posthumanism and education, offered,

in Braidotti’s influential book *The Posthuman* (2013) she lays out her vision of posthuman feminism as that which “rebels” against identity politics and dominant categories of identity. Posthuman feminism goes beyond antihumanist critique (potestas) to act and move (potentia) (Ringrose and Renold, 2016, p. 26). Posthuman feminism incubates and injects lethal viruses to shake “the political economy of phallogocentrism and of anthropocentric humanism.” Braidotti specifically advocates that “posthuman feminists look for subversion not in counter-identity formations, but rather in pure dislocations of identities via the perversion of standardized patterns of sexualized, racialized and naturalized interaction” (Braidotti, 2013, p. 99). Evoking the famous Deleuzo-Guattarian phrase about bodily capacities, Braidotti (Braidotti, 2013) calls upon us to experiment with resistance and intensity in order to find out what “posthuman bodies can do” (p. 221).

But we ask what might it mean to dissolve the concept of identity and see what posthuman bodies can do in and through curriculum? And what is the ethico-onto-epistemology that accompanies this dissolution?

There are no firm answers to questions like these, and so any responses must *stay with the trouble* (Haraway, 2016), acknowledging, for example, the limitations, incommensurabilities, and aporia that accompany all related thinking. This staying with the trouble is itself an ethic. Witness Burman (2021) in her attempt to think with identity, the trouble with dualistic social categories, and the need to develop new lines of thought. Burman resourced the cyborg and cautioned against resourcing the cyborg:

There are clear difficulties in knowing how far to develop the cyborg metaphor. And at the outset there is the temptation to turn it into a subject, an identity, for this would recuperate the cyborg back into individualist psychology. (The fact that typically cyborgs are discussed only in the singular—“the” cyborg—bolsters this, as indeed do the many cultural representations that portray the cyborg as lonely, agonised, and longing to be, or to be like, a human.). These ideas become relevant here as a strategy to disturb psychological representations of subjectivity. For the cyborg resists the trappings of subjectivity and identification (Gordo López and Parker, 1999) and so exposes their limitations (p. 218).

Thinking with the cyborg is a troubled activity, but one that is productive in its trouble. It exposes some of the limitations of known approaches to identity (along with representation and subjectivity) and might be suggestive of paths forward. For instance, in thinking with the cyborg, Burman said, “the cyborg has no stable identity, and does not organise with others on the basis of similarity. Rather, its communities are temporary strategic alliances or affiliations organised around specific goals” (p. 217). In so saying, Burman evoked a suggestion offered by Ringrose and Renold (2016) to the difficulties of identity in Braidotti (2013).

The orientations of posthumanities allow for a sitting with the unease of incommensurabilities with ethics at the core. Ringrose and Renold (2016) suggested some of the constraints of Braidotti’s (2013) rejection of identity while simultaneously loving her. They also brought into view a contemporary version of the hokey-cokey, this one an avowedly critically oriented posthumanities. Ringrose and Renold (2016), for example, said:

Despite our longstanding love affair with Braidotti (see Renold and Ringrose, 2011 and forthcoming) we find her call for “pure dislocations of identity” utopic and problematic. Challenging simplistic identity politics around universal womanhood or girlhood has never meant throwing out or simply blowing up identity categories. Critically engaging with strategic essentialism in the context of everyday political activism has proven more useful (Spivak, 1990). Not satisfied with abstract treatises on what posthuman feminism should be, we want to use the important conceptual tools of posthumanism to think methodologically about the processes involved in feminist “becomings” through our empirical research on feminist groupings and politics in secondary schools (p. 221).

Burman's (2021) description of the cyborg's "strategic alliances" (p. 217) seem to connect to Ringrose and Renold's (2016) raising of *strategic essentialism* which in its classic critical, postcolonial sense referred to the "strategic use of essentialism" (Danius et al., 1993, p. 35) by the *subaltern* (Spivak, 1988) with Spivak emphasizing the strategic part of the term (and even wanting to abandon the term because of her readers' over-reliance on the essentialism part) (Danius et al., 1993). There is no resolution to the conundrum of identities as problematic, identities as politically useful, identities as stultifying, identities as a needed heuristic, and one of the affordances of thinking with identities in posthumanities is that they can be all these things at the same time and more.

Posthuman curriculum asks educators to think relationally about what composes curriculum and how curriculum produces networks of effects. Curriculum emerges from and generates relations between humans and non-humans (computer, classroom materials, learners, teachers, knowledges, places, time). Identities and learning are entangled in this web of relations and are not neutral. Curriculum choices become complex ethical commitments. Posthumanities in education requires a shift away from the view of humans as saviors or sources of power. Curriculum history has acknowledged the human as a "normative category" (Braidotti, 2019, p. 35), the point of reference of all other categories and divisions. Posthumanities complexifies the human category and makes it impossible to see humans as a point of reference for everything else. Curriculum must instead notice the interdependence and importance of all beings and becomings. Braidotti (2019) explained that posthumanities "offers a spectrum through which we can capture the complexity of ongoing processes of subject formation. In other words, it enables subtler and more complex analyses of powers and discourses" (pp. 35–36) promoting "complexity, embodied and embedded diversity and multiple becomings" (p. 37).

Concluding thoughts

In our examination of the treatment of identity theorizing across curriculum, we have found that the concept of identity options has numerous contributions to make to curriculum thought and practice. Identity options, as we have suggested, for instance, signals many things: the limits of technologies of curriculum that would be built without consideration for the learner; the need for epistemic and ontological diversity in curriculum; how curricula can colonize, minoritize and promote particular kinds of citizens; and the ethical requirement for curricula that promote expansive identity options for learners.

The concept of identity options is closely connected to lived curricula, with both forwarding awareness that learners experience and co-produce classroom curricula in potentially infinite ways, hence exploding the myth that programmatic curriculum can ever be enacted without alteration or be singularly apprehended.

Identity options, as a concept, owes an intellectual debt of gratitude to movements connected to the reconceptualization of curriculum, and has its lineages in related forms of critical, socio-cultural theorizing. Recent thinking oriented toward posthumanities has demanded that curriculum scholars rethink identity options and identity more generally to consider the constraints of human-centric thinking.

References

- Aoki, T.T., 1993. Legitimizing lived curriculum: towards a curricular landscape of multiplicity. *J. Curric. Superv.* 8 (3), 255–268.
- Apple, M., 2004. *Ideology and Curriculum*, third ed. RoutledgeFalmer, New York.
- Apple, M.W., 2008. Curriculum planning: content, form, and the politics of accountability. In: Connelly, F.M., He, M.F., Phillion, J. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage Publications, Inc, Thousand Oaks, pp. 25–44.
- Barad, K.M., 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, Durham.
- Bobbitt, F., 1916. *What the Schools Teach and Might Teach*. Survey Committee of Cleveland Foundation, Cleveland.
- Bobbitt, F., 1918. *The Curriculum*. Houghton Mifflin Company, Boston.
- Braidotti, R., 2011. *Nomadic Theory: The Portable Rosi Braidotti*. Columbia University Press, New York.
- Braidotti, R., 2013. *The Posthuman*. Polity Press, Cambridge.
- Braidotti, R., 2019. A theoretical framework for the critical posthumanities. *Theor. Cult. Soc.* 36 (6), 31–61. <https://doi.org/10.1177/0263276418771486>.
- Buchanan, W., 1968. Identification, political. In: Sills, D.L., Merton, R.K. (Eds.), *International Encyclopedia of the Social Sciences*. Macmillan, New York, pp. 57–61.
- Burman, E., 2021. *Developments: Child, Image, Nation*, second ed. Routledge, Abingdon, Oxon.
- Castenell, L.A., Pinar, W.F., 1993. Introduction. In: Castenell, L.A., Pinar, W.F. (Eds.), *Understanding Curriculum as Racial Text: Representations of Identity and Difference in Education*. State University of New York Press, Albany, pp. 1–30.
- Cummins, J., Early, M. (Eds.), 2011. *Identity Texts: The Collaborative Creation of Power in Multilingual Schools*. Trentham Books, Stoke-on-Trent.
- Cummins, J., 2000. *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Multilingual Matters, Clevedon.
- Cummins, J., 2001. *Negotiating Identities: Education for Empowerment in a Diverse Society*, second ed. California Association of Bilingual Education, Los Angeles.
- Cummins, J., 2006. Identity texts: the imaginative construction of self through multiliteracies pedagogy. In: Garcia, O., Skutnabb-Kangas, T., Torres-Guzman, M.E. (Eds.), *Imagining Multilingual Schools: Languages in Education and Globalization*. Multilingual Matters LTD, Toronto, pp. 51–68.
- Cummins, J., 2009. Transformative multiliteracies pedagogy: school-based strategies for closing the achievement gap. In: *Multiple Voices for Ethnically Diverse Exceptional Learners*, vol. 11, pp. 38–56 (2).
- Danius, S., Jonsson, S., Spivak, G.C., 1993. An interview with Gayatri Chakravorty Spivak. *Boundary 2* 20 (2), 24–50.
- de Freitas, E., 2016. Material encounters and media events: what kind of mathematics can a body do? *Educ. Stud. Math.* 91 (2), 185–202.
- Deleuze, G., Parnet, C., 1987. *Dialogues* (H. Tomlinson, B. Habberjam, Trans.). Columbia University Press, New York.
- Deng, Z., Luke, A., 2008. Subject matter: defining and theorizing school subjects. In: Connelly, F.M., He, M.F., Phillion, J. (Eds.), *The SAGE Handbook of Curriculum and Instruction*. Sage Publications Inc, Thousand Oaks, pp. 66–87.

- Dewey, J., 1902. *The Child and the Curriculum* (No. 5). University of Chicago Press, Chicago.
- Doyle, W., 1992. Curriculum and pedagogy. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 486–516.
- Edwards, R., 2011. Translating the prescribed into the enacted curriculum in college and school. *Educ. Philos. Theor.* 43 (S1), 38–54.
- Eisner, E.W., 1979. *The Educational Imagination: On the Design and Evaluation of School Programs*. Macmillan, New York.
- Eisner, E.W., 1992. Curriculum ideologies. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan Publishing Company, Toronto; New York, pp. 302–326.
- Erikson, E.H., 1963. *Childhood and Society*, second ed. W.W. Norton & Company, New York.
- Erikson, E.H., 1968. Identity, psychosocial. In: Sills, D.L., Merton, R.K. (Eds.), *International Encyclopedia of the Social Sciences*. Macmillan, New York, pp. 61–65.
- Esteban-Guitart, M., Moll, L., 2014a. Funds of identity: a new concept based on the funds of knowledge approach. *Cult. Psychol.* 20 (1), 31–48.
- Esteban-Guitart, M., Moll, L., 2014b. Lived experience, funds of identity and education. *Cult. Psychol.* 20 (1), 70–81.
- Ferrante, E., 2012. *My Brilliant Friend* (A. Goldstein, Trans.). Europa Editions, New York.
- Ferrante, E., 2013. *The Story of a New Name* (A. Goldstein, Trans.). Europa Editions, New York.
- Ferrante, E., 2014. *Those Who Leave and Those Who Stay* (A. Goldstein, Trans.). Europa Editions, New York.
- Ferrante, E., 2015. *The Story of the Lost Child* (A. Goldstein, Trans.). Europa Editions, New York.
- Finder, M., 2004. *Educating America: How Ralph W. Tyler Taught America to Teach*. Praeger, Westport.
- Freire, P., 1970. *Pedagogy of the Oppressed* (M. Bergman Ramos, Trans.). Herder and Herder, New York.
- Gleason, P., 1983. Identifying identity: a semantic history. *J. Am. Hist.* 69 (4), 910–931.
- González, N., Moll, L.C., Amanti, C. (Eds.), 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Inc, Mahwah.
- Gordo López, A.J., Parker, I. (Eds.), 1999. *Cyberpsychology*. Macmillan, Basingstoke.
- Hackett, S., Somerville, M., 2017. Posthuman literacies: young children moving in time, place and more-than-human worlds. *J. Early Child. Literacy* 17 (3), 374–391.
- Haraway, D., 2015. Anthropocene, Capitalocene, Plantationocene, Chthulucene: making kin. *Environ. Humanit.* 6 (1), 159–165.
- Haraway, D., 2016. *Staying With the Trouble. Making Kin in the Chthulucene*. Duke University Press, Durham.
- Heydon, R., Iannacci, L., 2008. *Early Childhood Curricula and the De-pathologizing of Childhood*. University of Toronto Press, Toronto.
- Heydon, R., O'Neill, S., 2016. *Why Multimodal Literacy Matters: (Re)conceptualizing Literacy and Wellbeing Through Singing-Infused Multimodal, Intergenerational Curricula*. Sense Publishers, Rotterdam.
- Heydon, R., Moffatt, L., Iannacci, L., 2015. "Every day he has a dream to tell": classroom literacy curriculum in a full-day kindergarten. *J. Curric. Stud.* 47 (2), 171–202.
- Heydon, R., 2007. Making meaning together: multi-modal literacy learning opportunities in an intergenerational art program. *J. Curric. Stud.* 39 (1), 35–62.
- Heydon, R., 2013. *Learning at the Ends of Life: Children, Elders and Literacies in Intergenerational Curricula*. University of Toronto Press, Toronto.
- hooks, b., 1989. *Talking Back: Thinking Feminist, Thinking Black*. South End Press, Boston.
- Iannacci, L., 2019. *Reconceptualizing Disability in Education*. Lexington Books, Lanham.
- Ingersoll, M., DeLuca, C., Bolden, B., et al., 2019. Introduction. In: Wearing, J., Ingersoll, M., DeLuca, C., et al. (Eds.), *Key Concepts in Curriculum Studies: Perspectives on the Fundamentals*, first ed. Routledge, New York, pp. 1–11.
- Jackson, P.W., 1992. Conceptions of curriculum and curriculum specialists. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum: A Project of the American Educational Research Association*. Macmillan, New York, pp. 3–40.
- Jones, A., Hoskins, T.K., 2016. A mark on paper: the matter of indigenous-settler history. In: Taylor, C.A., Hughes, C. (Eds.), *Posthuman Research Practices in Education*. Palgrave Macmillan, London, pp. 75–92.
- Kelly, U.A., 1997. *Schooling Desire: Literacy, Cultural Politics, and Pedagogy*. Routledge, New York.
- Kincheloe, J.L., 1998. Pinar's Currere and identity in hyperreality: grounding the postformal notion of intrapersonal intelligence. In: Pinar, W.F. (Ed.), *Curriculum: Toward New Identities*. Garland Publishing, Inc, New York, pp. 129–142.
- Kincheloe, J.L., 2013. Pinar's Currere and identity in hyperreality: grounding the postformal notion of intrapersonal intelligence. In: Pinar, W.F. (Ed.), *Curriculum: Toward New Identities*. Routledge, New York; Abingdon, Oxon, pp. 129–142.
- Kliebard, H.M., 1970. Reappraisal: the Tyler rationale. *Sch. Rev.* 78, 259–272.
- Kliebard, H.M., 1995. The Tyler rationale revisited. *J. Curric. Stud.* 27 (1), 81–88.
- Kuby, C.R., Rucker, T.G., 2016. *Go Be a Writer!: Expanding the Curricular Boundaries of Literacy Learning With Children*. Teachers College Press, New York.
- Kuby, C.R., Spector, K., Thiel, J.J. (Eds.), 2019. *Posthumanism and Literacy Education: Knowing/Becoming/Doing Literacies*. Routledge, New York.
- Lewis, M.G., 1993. *Without a Word: Teaching Beyond Women's Silence*. Routledge, New York.
- Lloyd, G., 1996. *Spinoza and the Ethics*. Routledge, London.
- Luke, C., Gore, J. (Eds.), 1992. *Feminisms and Critical Pedagogy*. Routledge, New York.
- Lyotard, J.F., 1984. *The Postmodern Condition: A Report on Knowledge*, vol. 10. University of Minnesota Press, Minneapolis.
- Margulis, L., Sagan, D., 1995. *What Is Life?* University of California Press, Berkeley.
- McCarthy, C., 1993. Multicultural approaches to racial inequality in the United States. In: Castenell, L., Pinar, W. (Eds.), *Understanding Curriculum as Racial Text: Representations of Identity and Difference in Education*. State University of New York Press, Albany, pp. 225–246.
- Montgomery, L.M., 1908. *Anne of Green Gables Series: Anne of Green Gables*. L.C. Page & Co, Boston.
- Murris, K., 2016. The posthuman child: educational transformation through philosophy with picturebooks. In: Olsson, L.M., Vandenbroeck, M. (Eds.), *Contesting Early Childhood Series*. Routledge, Abingdon, Oxon.
- Null, J.W., 2003. John Dewey's child and the curriculum 100 years later: lessons for today? *Am. Educ. Hist. J.* 30, 59–68.
- Null, J.W., 2008. Curriculum development in historical perspective. In: Connelly, F.M., He, M.F., Phillion, J. (Eds.), *The SAGE Handbook of Curriculum & Instruction*. Sage Publications Inc, Thousand Oaks, pp. 478–490.
- Pinar, W.F., 2004. *What is Curriculum Theory?* Routledge, London.
- Pinar, W.F., 2008. Understanding curriculum as racial text. In: Pinar, W.F., Reynolds, W.M., Slattery, P., Taubman, P.M. (Eds.), *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. Peter Lang, New York, pp. 315–358.
- Pinar, W.F., 2019. Currere. In: Wearing, J., Ingersoll, M., DeLuca, C., et al. (Eds.), *Key Concepts in Curriculum Studies: Perspectives on the Fundamentals*. Routledge, New York; Abingdon, Oxon, pp. 50–52.
- Reese, W.J., 2001. The origins of progressive education. *Hist. Educ. Q.* 41 (1), 1–24.
- Renold, E., Ringrose, J., 2011. Schizoid subjectivities?: re-theorising teen girls' sexual cultures in an era of 'sexualisation'. *J. Sociol.* 47 (4), 389–409.
- Ringrose, J., Renold, E., 2016. Cows, cabins and tweets: posthuman intra-active affect and feminist fire in secondary school. In: Taylor, C., Hughes, C. (Eds.), *Posthuman Research Practices in Education*. Palgrave Macmillan, London, pp. 220–241.
- Robson, C., Sumara, D., 2019. Normativity. In: Wearing, J., Ingersoll, M., DeLuca, C., et al. (Eds.), *Key Concepts in Curriculum Studies: Perspectives on the Fundamentals*. Routledge, New York, pp. 145–150.
- Schwab, J.J., 1973. The practical 3: translation into curriculum. *Sch. Rev.* 81, 501–522.
- Seligman, E.R.A., Johnson, A.S., 1930. *Encyclopedia of the Social Sciences*. Macmillan Publishers, New York.
- Shields, C.M., Bishop, R., Mazawi, A.E., 2005. *Pathologizing Practices: The Impact of Deficit Thinking on Education*, vol. 268. Peter Lang Publishing Incorporated.
- Spivak, G.C., 1990. *The Post-Colonial Critic: Interviews, Strategies, Dialogues*. Routledge, London.

- Spivak, G.C., 1988. Can the subaltern speak? In: Nelson, C., Grossberg, L.C. (Eds.), *Marxism and the Interpretation of Culture*. Macmillan Education, Basingstoke, pp. 271–313.
- Stronach, I., MacLure, M., 1997. *Educational Research Undone: The Postmodern Embrace*. Open University Press, Buckingham.
- Tuck, E., McKenzie, M., 2015. Relational validity and the “where” of inquiry: place and land in qualitative research. *Qual. Inq.* 21 (7), 633–638.
- Tuhiwai Smith, L., Tuck, E., Wayne Yang, K., 2018. Introduction. In: Tuhiwai Smith, L., Tuck, E., Wayne Yang, K. (Eds.), *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, New York; Abingdon, Oxon, pp. 1–23.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, Chicago.
- van der Zaag, A.C., 2016. On posthuman subjectivity. *J. Cult. Econ.* 9 (3), 330–336.
- Wraga, W.G., 2017. Understanding the Tyler rationale: basic principles of curriculum and instruction in historical context. *Space Time Educ.* 4 (2), 227–252.
- Wraga, W., 2019. The pragmatic progressives. *Am. Educ. Hist. J.* 46 (2), 111–129.
- Zhang, H., Heydon, R., Li, W., Malins, P., 2020. Literacies and identities in transnational education: a case study of literacy curricula in a Canadian Transnational Education Program in China. *Curric. J.* 31 (1), 132–156.

Cross-cultural narrative inquiry—cross-cultural lives

Ming Fang He, Department of Curriculum, Foundations, and Reading, College of Education, Georgia Southern University, Statesboro, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Origins of cross-cultural narrative inquiry	392
Creating a composite auto/biographical narrative method to study the bilingual/bicultural and multilingual/multicultural experience	393
Representing bilingual and bicultural lives and identities	393
Using metaphors to identify research themes and understand cross-cultural experiences	394
Tales in words and tales on canvas—telling stories through arts and poetry	394
Narrative and experience in multicultural education	394
Using life based literary narratives in multicultural teacher education	395
Research on the education of Asian American immigrant students in schools/families/communities	396
Activist practitioner inquiry into the life in schools/families/communities in the US South	397
Research on the education of disenfranchised individuals and groups in Hong Kong and Mainland China	398
East meets west in curriculum studies	398
Exile pedagogy/curriculum and teaching in-between	398
Diaspora curriculum theorizing	399
Cross-cultural narrative imagination in hard times	401
References	402

Origins of cross-cultural narrative inquiry

My experience of cross-cultural narrative inquiry was originated from my explorations during my master's program of studies at OISE/University of Toronto on immigrant education by understanding language, culture, learning, and teaching as a social process; by examining the political and social context of the education of language minority students; and by looking into the interconnectedness of language, culture, and identity. In my doctoral program, I studied narrative approaches to curriculum, teacher development, and research with F. Michael Connelly, where narrative, as phenomenon, as method, and as a form of representation is an approach about how people experience their lives, how they interact with others, how they shape and are shaped by the contexts in which they live and work. I explored a wide array of works related to narrative such as currere/feminist autobiography in the works of William Pinar and Madeleine Grumet in 1980s, 1990s, and 2000s; Janet Miller's work on teachers' auto/biography in 1980s, 1990s, and 2000s; bell hook's work in 1990s and 2000s; Richard Butt and Daniel Raymond's works on auto/biographical studies and collaborative autobiography, teachers thinking, and teacher's voice in 1980s and 1990s; Freema Elbaz's works on teacher thinking and personal practical knowledge in 1980s, 1990s, and 2000s; F. Michael Connelly and D. Jean Clandinin's works on teacher's personal practical knowledge, professional landscape, and narrative inquiry in 1980s, 1990s, 2000s; William Schubert's work on curriculum books, teacher practice, and teacher knowledge in 1980s, 1990s, 2000s; William Ayers' work on teacher practice, teacher education, teachers auto/biography, and social justice in 1980s, 1990s, 2000s; William Schubert and William Ayers' work on teacher lore/learning from one's own experience in 1992; Ivor F. Goodson and Andy Hargreave's work on teachers' professional lives in 1996; Marilyn Cochran-Smith & Susan L. Lytle's work on inside/out/teacher research and knowledge in 1993; Donald Schön's work on reflective turn/case studies in and on educational practice in 1991; Tom Russell & Hugh Munby's work on teachers as reflective practitioners and the role of experience in developing teachers' professional knowledge in 1990s, 1990s, and 2000s. However, the above-mentioned narrative approaches do not help capture the cross-cultural aspects of life.

I then turn to cross-cultural narrative inquiry which humanizes research that is philosophically compatible with cross-cultural lives and identities central to multiculturalism. I became involved in a Social Sciences and Humanities Research Council of Canada (SSHRC) funded long-term research program on teacher knowledge (1991–1998). I was a participant observer in a large urban adult ESL Center and conducted interviews with librarians, administrators, lead teachers, classroom teachers, and adult ESL learners while I was teaching Mandarin to Grade 6–12 public school children and English as a Second Language to immigrant adults and children, developed my understanding of the complexity of immigrant lives. I was also involved in other research with nursing educators and immigrant nurses, which helped further develop my research in cross-cultural inquiry of multicultural/multilingual education.

Cross-cultural narrative inquiry explores experience by bringing cross-cultural personal experience to bear on inquiry, seeing research as having autobiographical roots, as connected to, rather than disconnected from cross-cultural life; by thinking narratively, seeing cross-cultural experience as the starting point of inquiry, as fluid and changing rather than fixed, as contextualized rather than decontextualized; by being in the midst of cross-cultural lives, seeing research as long-term, passionate involvement in daily lives of participants, rather than short-term, in and out, detached observation; and by making meaning of cross-cultural experience in

relationship, rather than making meaning in isolation (Phillion and He, 2001, 2008; He and Phillion, 2008). Cross-cultural narrative inquiry, an approach that humanizes research, focuses on cross-cultural experience.

Creating a composite auto/biographical narrative method to study the bilingual/bicultural and multilingual/multicultural experience

My doctoral research was a study of cross-cultural identity development of three Chinese women teachers as they moved back and forth between Eastern and Western landscapes, cultures, and languages (He, 1998, 1999, 2003). Shiao, Wei, and Ming Fang, three participants in the study, were born during the Anti-Rightist Movement and the Great Leap Forward in the late 1950s, grew up during the Cultural Revolution (1966–1976), were exposed to foreign cultures and pursued undergraduate studies in Chinese universities (1978–1989) since the beginning of China's open door policy (1978–1989), pursued graduate studies in Canada (1988–1998), and moved to immigrant education (Shiao), cosmetic business (Wei), and North American academe (Ming Fang). As we moved between languages and cultures, the East and the West, the cultural, educational, and language strangeness we experienced was further complicated by rapid changes occurred in both landscapes. We did not feel at home in either place. Our foreignness and at-homeness in both landscapes were interchangeable, a phenomenon which characterizes our cross-cultural lives. I struggled to come to grips with new ways of thinking to bridge our lives in China, lives in Canada, lives in North American academy, and lives in-between. I started this life-based narrative inquiry, situated between non-fiction, fiction and academic discourses, with a hope to create possibilities for establishing a linkage between cross-cultural lives and identities, cross-cultural teacher education and curriculum studies in multicultural/multilingual contexts.

As my participants and I continued our inquiry of cross-cultural lives, we felt that our identities were threatened by telling our stories in public. Due to political reasons, I experimented with a composite auto/biographic narrative method to protect our identities and to capture the complexity of such identities (He, 1998, 1999, 2003). This method captured the essence of four bodies of literature in research in social sciences: auto/biography; fiction; narrative or closely related to narrative research in anthropology, ethnography, psychology, psychotherapy, psychiatry, theology, organizational theory, and social sciences in general; and narrative research in education. Other narrative works from people in different educational fields were also referred to such as teacher education; teaching and learning; writing; science education; curriculum studies; language teacher knowledge; educational psychology; Black feminist studies; multicultural teacher education, philosophy, and educational philosophy. Connelly and Clandinin (e.g., Clandinin and Connelly, 2000) pioneered narrative inquiry in the field of education and they have the longest sustained program of narrative research. Their work tremendously influenced my early work on cross-cultural narrative inquiry.

My participants and I worked through four procedures of my composite auto/biographical method. (1) We fictionalized our identities in the research text. My participants used pseudonyms of their choice. Their backgrounds were fictionalized according to their preference. We used the fictionalized first persona to tell our real stories. In my case, since my name and background were known to the reader, some of my stories were told through my participants with their permission. (2) We told and retold our stories. (3) We went back to our stories, relocated stories in a particular persona's voice, and developed a joint interpretation. The interpretations included agreement, disagreement, and puzzlement. (4) We went back to our stories and interpretations, and confirmed, or altered them. My participants and I, for example, tell stories of the Cultural Revolution (1966–1976) in a "stream of consciousness" narrative collage. This technique in writing with fragments of sentences ending in dots rather than complete sentences captures the shifting and evolving aspects of our displaced and disrupted lives. Three passionate and emotional voices tell stories within the same time span and across different spaces and settings as if watching a film with three pictured stories on the same screen. This extends a first person, singular account of narrative plot and temporality to a collective voice which represents stories located within a larger historical, socio-political, cultural, and linguistic context. Utilizing collage storytelling techniques with changing voices and fonts in the text helps capture a narrative unity out of fragments of stories by breaking the limits of time, space, and place. The cross-cultural narrative inquiry method moves beyond traditional narrative writing discourse and research boundaries and tells hidden and silenced narratives of suppressed and underrepresented groups to counter the preconceived metanarrative represented in the mainstream research literature that has often portrayed these groups as deficient and inferior. By telling counter stories, researchers begin to be aware of the importance of commitment to equity and social justice, and to recognize their obligation to link inquiry to social and educational change.

Representing bilingual and bicultural lives and identities

Along with my experiment with composite auto/biographical narrative method, I explore the bilingual and bicultural narrative representation of cross-cultural lives and identities (He, 1999, 2003). I derive the forms of narration from the storytelling styles of my participants instead of imposing my style upon them. For instance, when I tell the stories from my participants, I purposely use their words and expressions to reveal their cultural, linguistic, musical, or scientific backgrounds. Instead of polishing the interview and conversation transcripts, I deliberately allow orality (i.e., keeping in the realm of speech with its characteristic tones, pauses, and rhetorical transitions) to flow through while theorizing those stories in the left-hand column or in italics. As a bilingual and bicultural narrative researcher, I incorporate our Chinese voice into the text without losing the flow of English discourse. When I have difficulty in finding the English equivalent for specific Chinese words, phrases, slogans, and sentences, I mark them with

quotations and put the English translation and the Chinese in parentheses to satisfy both English and Chinese readers and to compensate for the loss of meaning in translation. However, in order to keep the flavor of both languages, I am caught in conflicts. Some of the expressions and metaphors convey different meaning depending on the language being used. I have to break those barriers to develop a new meaning based on both languages. If the conflict is irresolvable, an additional explanation is provided. It is crucial to develop a bilingual and bicultural methodology to research on bicultural lives and identities. Although my research has been conducted in Chinese and English only, I believe that it has implications for other bilingual and bicultural inquiries in the United States and other international contexts.

Using metaphors to identify research themes and understand cross-cultural experiences

As I collect stories from myself and my participants, I have found that my participants and I tend to narrate our experiences of acculturation and enculturation in China and Canada through metaphors such as journey metaphors, spatial metaphors, landscape metaphors, cultural metaphors including deceptional metaphor which conveys a philosophical truth through a fallacy as well as musical metaphors, and personal metaphors. As our stories of identity development and cultural transformation unfold, we realize that our modes of learning and ways of expressing ourselves are closely related to Chinese language and culture, which is strongly metaphorically oriented. The way metaphors developed implicitly from our cultural experiences helps explicitly portray the personal and professional identity development and knowledge transformation in our acculturation and enculturation processes in China and Canada. Thus I employ metaphor as both phenomenon (embodied in stories) and method (identifying stories) to explore the acculturation and enculturation processes of teachers and learners in China and Canada. I use a river metaphor to explore three phases in my life and the lives of my participants: our lives along the Yangtze River and the Yellow River before, during, and after the Cultural Revolution in China with multiple cultural movements; our lives in the cross-cultural flow in China and Canada; and our lives in the multicultural flow of the North American academy (He, 2002a, b, c). Special attention is paid to the relationship between our cross-cultural experience in China and Canada, the cultural impact of our move to Canada, and our experience of learning to become North Americans, a cultural process no less of an upheaval or displacement than the Cultural Revolution. I also use the river metaphor to create a new conception of multiculturalism perceiving different cultures as rivers running together to create sediment for a multicultural delta where lives of various cultural, ethnic, and linguistic groups merge with their own currents. This work is situated between nonfiction, fiction, and academic discourses. It is written in a way that is accessible to a large audience directly and indirectly involved in cross-cultural/bilingual/bicultural/multilingual/multicultural education, and transnational/diasporic studies. I see my work as fitting into an evolving, increasingly important, paradigm in the field of educational research on issues of language, culture, identity, and power in relation to bilingual/bicultural/multilingual/multicultural education and exile/in-between curriculum for dislocated, disenfranchised, and underrepresented individuals and groups.

Tales in words and tales on canvas—telling stories through arts and poetry

As I write about our stories in English, I encountered one of the major concerns over the generalizability of our stories of cross-cultural lives and identity development. Our narratives might strike the reader as being too individual, subjective, idiosyncratic, and anecdotal, and might be perceived to be noncomparable, non-generalizable, and non-scientific. I then discover a parallel way of storytelling through my collaboration with You Sha Liu, a Chinese woman artist. You Sha shares her cross-cultural life stories through painting, while my participants and I tell our stories in words. You Sha and I come together to make sense of our cross-cultural lives and to represent those stories through tales in words and tales on canvas (He, 1999, 2003). Between the lines of the text and colors on the canvases, our actions are transformed into imagination while imagination is converted into actions. The result of collaboration is a blend of narrative aesthetics: fiction ~ fact, dream ~ reality, theory ~ practice, retrospection ~ imagination, auto/biography ~ portraits. In this juxtaposition of narrative “onstage,” point of views have changed, characters shifted, sequences of events vanished, images clashed, and realms of narrative discourse collided. We also find that both in poetry and in the visual arts, the temporal, spatial, perceptual, social, or moral entities are not fixed but evolving with our inquiries of identities and lives in-between landscapes. I struggle to maintain the consistency of tenses in reflections upon past experiences. In Chinese there is much more freedom in tenses. We use variations of the third person all the time and sometimes we even omit tenses altogether. In English, however, there are specific rules for the proper use of verbs as they relate to different time periods. As a lifelong second language learner, I constantly struggle for the proper use of prepositions and articles in English. As I translate our stories and emotions into the text, poetry helps capture passions, emotions, and imaginations aroused by our complicated experiences without being inhibited by tenses, prepositions, and articles.

Narrative and experience in multicultural education

I also engaged in an edited book project, *Narrative and Experience in Multicultural Education* (Phillion et al., 2005). This chapter explores the potential that narrative and experiential approaches have for understanding issues of language, culture, identity,

and power related to multiculturalism in education. The research featured in the book reflects a critical way of understanding the nuances of multicultural life and the complexity of multicultural issues in education. The studies focus on the lives in schools, families, and communities. The particular quality of the work featured in this book that distinguishes it from other work in multicultural education is the emphasis on experience, understanding of experience, and the transformation of this understanding into social and educational change.

Using life based literary narratives in multicultural teacher education

I have been teaching in a predominantly White institution since 1998. As an educator of both preservice and in-service teacher education students, one of the major challenges is how to prepare White teachers in rural areas of the United States to meet the needs of increasingly diverse populations of students. Drawing upon the work of Martha Nussbaum (1997) on narrative imagination and Maxine Greene (1995) on literary imagination, I collaborate with JoAnn Phillion to engage our students in life based literary texts such as memoirs, autobiographies, and novels that focus on the daily experiences of diverse families, parents, students and teachers (Phillion and He, 2004). For Nussbaum (1997), narrative imagination means “the ability to think what it might be like to be in the shoes of a person different from oneself, to be an intelligent reader of that person’s story, and to understand the emotion and wishes and desires that someone so placed might have” (pp. 10–11). For Greene (1995), “[Literary] imagination is what, above all, makes empathy possible. It is what enables us to cross the empty spaces between ourselves and those we teachers have called other over the years” (p. 3).

In our teacher education courses, students first read Nussbaum’s (1997) *Cultivating Humanity: A Classical Defense of Reform in Liberal Education* and study the key concepts of narrative imagination, self-examination, and world citizenship. In the process of reading life based literary narratives and reflecting on their experiences, students develop narrative imagination (the ability to actively empathize with others), and engage in self-examination (the ability to critically examine ones’ traditions, beliefs and values), enabling them to begin to understand themselves as world citizens (the ability to see themselves as members of not only a local community, but also a world community) (Nussbaum, 1997). Our attempt is to engage our students fully in a quest “to cultivate their humanity ... to see themselves not simply as citizens of some local region, or group, but also, and above all, as human beings bound to other human beings by ties of recognition and concern” (Nussbaum, 1997, p. 10). In developing a sense of “world citizenship,” students learn to step out of their comfort zone, to challenge assumptions and question taken-for-granted truths, to experience differences, and to make connections between the worlds they live in and the worlds of others. They also read Greene (1995) *Releasing Imagination: Essays on Education, the Arts, and Social Change* and explore ideas of literary imagination, creating possibilities, and expanding community. The key concepts of Nussbaum and Greene provide the guiding philosophical framework for the courses.

We also draw upon a wide array of works on multicultural education such as culturally responsive/relevant/sustaining curriculum and pedagogy (Gay, 2000/2018; Ladson-Billings, 1994/2009, 2021; Paris and Alim, 2017), funds of knowledge (González et al., 2005), cultural modeling (Lee, 1996/2009), and third space (Gutiérrez, 2008; Gutiérrez et al., 1999). These works inspire teachers in schools to work with other teachers, administrators, parents, students, community workers, policy makers, and other educational workers to engage in critical reflections and dialogues; exemplify democracy for the common good; resist oppression, suppression, and repression; decolonize space and place; teach and learn toward freedom; “hope radically, imagine creatively, and act inspirationally” (Ball et al., 2017, p. 2); transgress epistemological, socio-political, cultural, and linguistic borders to become the light in troubling times; and invent a landscape of education for cultural and linguistic diversity where creative and worthwhile culturally sustaining (Paris and Alim, 2017) learning and teaching make the impossible possible in the increasingly multicultural/multiethnic/multilingual/multiracial world.

Then we read life based literary narratives such as Patrick Chamoiseau’s *School Days* (1994), Eva Hoffman’s *Lost in Translation* (1989), Toni Morrison’s *The Bluest Eye* (1970), and Maxine Hong Kingston’s *Woman Warrior* (1975). Our teachers are also encouraged to read autobiographical work such as Michele Foster’s (1997) *Black Teachers on Teaching* and Vivian Paley’s (1979) *White Teacher*; narrative inquiry research such as Chris Carger’s (1996) *Of Borders and Dreams: A Mexican American Experience of Urban Education* and Ming Fang He’s (2003) *A River Forever Flowing: Cross-Cultural Lives and Identities in the Multicultural Landscape*; and ethnographies such as Chritina Igoa’s (1995) *The Inner World of the Immigrant Child*, Lourdes Diaz Soto’s (1997) *Language, Culture, and Power: Bilingual Families and the Struggle for Quality Education*, and Guadalupe Valdés *Con Respeto: Bridging the Distances Between Culturally Diverse Families and Schools* (1996) and *Learning and Not Learning English: Latino Students in American Schools* (2001).

As our particular focus in multiculturalism is on language, culture and identity, we primarily use life based literary narratives written by immigrants of their experiences in a new land, and by minorities living in mainstream cultures (e.g., Anzaldúa, 1987; Blaise, 1993; Chamoiseau, 1994; Cisneros, 1994; Cowan, 1982; Crow Dog and Erdoes, 1990; Dorfman, 1998; Firmat, 1994; Hoffman, 1989; Kaplan, 1993; Kingston, 1975; Melanson, 1999; Rodríguez, 1982; Santiago, 1993). The purpose of choosing these narratives is to engage our students in the lives of others in dramatically different environments, and to feel how the characters struggle with new languages, cultures, and identities. As students are enveloped in these lives, we encourage them to reflect on their experiences, to recognize differences between their lives and the lives of the characters, to imagine what it would be like to live the life of the characters, and to envision what they would do if they had students with similar experiences as the characters.

In conjunction with reading life based literary narratives, autobiographies, narrative inquiries, and ethnographies which explore the experience of people different than themselves, we encourage our pre-service and in-service teachers to critically examine

themselves by composing autobiographical papers in which they reflect on their backgrounds and experiences, critically examine their values and beliefs to develop understanding of the ways in which their personal histories, cultures, and experiences affect who they were, how they become who they are, how they interact with others, and how they perceive the world. We encourage our teachers to compose narrative imagination papers where they reflect on a life based literary narrative and explore how it develops their ability to see how they are connected to others and to empathize with others; and a collaborative project in which students work together to extend their understanding of multicultural issues and to explore possibilities for changes in their classrooms, schools, and communities.

The focus is on the power of narrative to assist our students to increase knowledge of ways diverse students experience the world both inside and outside schools and classrooms, to make sense of lived worlds, to see themselves in the stories of others, to learn to listen to and hear the stories of others, to enter the realities of their lives often dramatically different than their own, not through abstract reasoning, but through developing their emotions and empathic understanding. We feel that in recognizing themselves as connected to all human beings, devoted to respect for difference, committed to equality and social justice, our students will see themselves as agents of much needed change in schools, societies, and other spaces and places.

Research on the education of Asian American immigrant students in schools/families/communities

The US Census 2020 reported six major race categories: White, Black or African American, American Indian or Alaskan Native, Asian, Native Hawaiian or other Pacific Islander, and other races. By 2020 of the total US population (331.9 million), 20.6 million people (6.2%) are identified as Asian, Native Hawaiian or other Pacific Islander. Of 44.9 million immigrants in the United States [13.7% of the total population (331.9 million)], 31% were Asian (14.1 million) (US Census Bureau, 2020). Asia is the second-largest region of birth for US immigrants, after the Americas, and since 2013 India and China have been the leading origin countries, displacing Mexico. Arrivals from Asia are projected to become the largest foreign-born group by 2055 surpassing Latinos as the largest group immigrating into the US (US Census Bureau, 2020). Asian immigrants have tripled over the past three decades representing the fastest-growing population of newcomers to the US, with higher levels of income and educational achievement than native-born Whites. Nevertheless, historically Asians have been the object of racist and nativist US legislation and policies relating to immigration, naturalization, civil and linguistic rights, and full access to education, including, *inter alia*, the Chinese Exclusion Act, Executive Order 9066, and *Lau v. Nichols* (Pew Research Center, 2022; Spring, 2016; Takaki, 1993). Asians are often *orientalized* (Saïd, 1979) as rote learning oriented and submissive learners, exotic others, or monolithic model minorities in times of peace and economic security, and vilified in times of war, economic adversity, and, now, zoonotic viral pandemics; and the cycle of perceived threat followed by xenophobic rhetoric and policy response continues with the rise in anti-Asian hate crimes in the wake of the COVID-19 pandemic and attempts to reverse the overrepresentation of Asian students at highly selective universities.

Further, the linguistic heritage, cultural knowledge, and experience Asian immigrants and their children bring to schools are often ignored or overlooked. The ethnic and cultural diversity within Asian and Pacific American groups is often obscured and ignored in mainstream scholarship. Asian and Pacific Americans are either excluded entirely from studies that focus on people of color or compared with European Americans and other minorities. One of the major reasons supporting the exclusion of Asian American and Asian Canadian students from the dominant discourse in education is “the stereotype that Asians do not have any problems (i.e., they are model minorities)” (Lee, 1996/2009, p. 5). The model minority myth masks struggles and difficulties Asian students face, perpetuates resentment and hostility from members of the majority and other minority groups, and bolsters crimes being committed against Asians. Many researchers call for attention to the needs of Asian students who are by and large neglected in educational policy making in terms of funding programs, services and educational resources. This relative absence from discourse in educational policy and practice, despite their growing presence in schools and societies (Park et al., 2001, 2003), are major reasons for developing more studies on Asian American immigrant students.

I have been engaged in a study in Savannah, Georgia for more than 15 years to explore how Chinese immigrant families and communities work with schools to facilitate their children’s academic success while maintaining their cultural and language heritage within diversities among minorities and other ethnic groups in the US South. This research project is part of a long-term program of research on language, culture, identity, and power in relation to bilingual/bicultural/biracial education, multilingual/multicultural/multiracial education, and cross-cultural curriculum studies and teacher education. It *builds on* my cross-cultural experiences in China, Hong Kong/China, Canada and the USA, and my ongoing cross-cultural educational research. It *extends* my previous cross-cultural narrative study of the identity development and cultural transformation of immigrant Chinese women teachers (He, 1998, 2003) *to an inquiry into* the impact of immigrant’s language, culture, and identity development on school-community relations and their bearing on immigrant children’s school success in an increasingly diversified society. This research links with significant documentary study, oral history research, memoir, ethnographic and narrative work done with immigrant families and children. This research illuminates silenced stories of Asian American families and children who are often neglected in educational policy making in terms of funding programs, services, and educational resources. Due to the fact that Asian immigrants are double-marginalized—they do not fit in the mainstream, dominant discourse nor do they fit in marginalized minority discourse, I inquire into Asian immigrants’ dilemmas, tensions, contradictions, and advantages associated with their lives in the East, in the West, and in-between.

Activist practitioner inquiry into the life in schools/families/communities in the US South

My cross-cultural narrative research and scholarship thrive upon my teaching and doctoral student supervision. I have established and sponsored a doctoral research seminar called *Works-In-Progress* where a group of faculty members and doctoral students gather regularly to invigorate discussions and conversations on current issues on dissertation work and educational research. An inquiry-oriented philosophy permeates all aspects of *Works-In-Progress*. I encourage doctoral students to perceive their dissertation work as a quest for possibilities for quality education in an increasingly expanding democratic society. I invite them to perceive this quest directly connected to their personal and professional concerns, research interests, and inquiries. I encourage them to bring in their personal/professional experience and practical knowledge to inform their research while moving beyond certain boundaries in their ways of thinking about, looking at, and writing about their research. I invite them to search for culturally relevant inquiries to locate research within complex social, cultural, historical, and linguistic contexts. I also invite them to create bridges between theory and practice, between worlds other than their own, to see themselves in the stories of others. I encourage students to be sensitive to the emotional aspects of language, cultural, and political concerns of their participants. I encourage them to understand, synthesize, analyze, and criticize others' theories of educational inquiries to develop their own through their dissertation work. I invite them to join the efforts of building a research community where diverse voices are heard, diverse ways of doing research appreciated, and diverse challenges initiated.

I invite them to keep asking themselves the questions: What kinds of educational research do we need to enhance democracy and social justice while advancing the knowledge that will make a difference in people's lives? What are the more pressing educational and social problems of our time we should address in our research? How do we position ourselves as educational researchers to help reinvigorate the discourse and the investment in the public good by offering research and scholarship that directly look at education and the public? How do people in schools, communities, and societies think and understand the ways educational research works? In which ways can we deeply engage ourselves in research that serves the public good but not merely our own interests? What role can and do education researchers play in challenging the reform efforts to live up to the standard of working in the public interest? *Works-In-Progress* generates some outstanding social justice-oriented dissertation work. Some of these social justice-oriented research have been and will continue to be published in academic outlets, and have been transformed into positive changes in institutions, schools, communities, and societies.

Building upon my teaching and doctoral student supervision, I have been developing a program of research with a particular focus on activist practitioner inquiry/research for social justice that explores life in schools, families, and communities in the US South (He and Ross, 2013; He et al., 2015). I co-edit a book series with Information Age Publishing, *Research for Social Justice*, which features educational research that connects the personal with the political, the theoretical with the practical, and inquiry with social change. The first book in this series, *Research for Social Justice: Personal ~ Passionate ~ Participatory Inquiry* (He and Phillion, 2008) features exemplary social justice work done by my advanced doctoral students and those of JoAnn Phillion. These inquiries demonstrate three distinct qualities. Each is personal, compelled by values and experiences researchers bring to the work. Each is passionate, grounded in a commitment to social justice concerns of people and places under consideration. Each is participatory, built on long term, heart-felt engagement, and shared efforts. The principal aspect of the inquiries featured in the book series that distinguishes it from others is that researchers are not detached observers, nor putatively objective recorders, but active participants in schools, families, and communities. Researchers have explicit research agendas that focus on equity, equality, and social justice. Rather than aiming solely at traditional educational research outcomes, positive social and educational change is the focal outcome of inquiry. The researchers are diverse and their inquiries are far ranging in terms of content, people and geographic locations studied. These studies reflect new and exciting ways of researching and representing experience of the disenfranchised, underrepresented, and invisible groups, and challenge stereotyped or deficit-oriented perspectives on these groups. This book series inspires pre-service and in-service teachers, educators, educational researchers, administrators, and educational policy makers to commit to the enactment of educational and social change that fosters equity, equality, and social justice.

I continue to create academic outlets to feature the cutting-edge work of diverse curriculum scholars, particularly those of my doctoral students who engage in social justice-oriented research and make positive social and educational changes in schools, communities, and societies. These social justice-oriented works focus on the power of counter narrative as a means to contest the official or meta narrative that often portrays disenfranchised individuals and groups as deficient and inferior. The counter narratives challenge traditional ways of engaging in and interpreting curriculum research and affirm the significance of curriculum inquiry as a form of liberatory or radical democratic practice. The counter narratives help tell silenced and neglected stories of experience of repressions, suppressions, and subjugations, stereotypes of Southern women, Blacks, and other disenfranchised individuals and groups; encourage examination of the force of slavery, racism, sexism, classism, religious repression and other forms of oppression and suppressions on the life and curriculum in schools, neighborhoods, and communities in the South; and promote a more equitable human condition that embodies cultural, linguistic, and ecological diversity and plurality of identities of individuals, groups, tribes, and societies that are conducive to the flourishing of creative capacities that invigorate intellectual, emotional, moral, and spiritual existence for all.

Research on the education of disenfranchised individuals and groups in Hong Kong and Mainland China

I have extended my research on language, culture, identity, and power in relation to bilingual/bicultural/biracial and multilingual/multicultural/multiracial education, and cross-cultural curriculum studies and teacher education in Canada and the United States to other international contexts such as Hong Kong and Mainland China where I have been enaging in two large scale research programs of study on minority and disenfranchised groups (He and Yu, 2017). Due to COVID-19, my research on the education of minority and disenfranchised individuals and groups in Hong Kong and Mainland China was delayed. This line of inquiry is exciting and challenging due to lack of concerns and actions on the education of these groups in history. I strive to work with multi-ethnic teacher-scholars, researchers, and practitioners to develop programs of research in which cross-cultural narrative inquiries are used to explore the life of schools, families, and communities to foster equity, equality, and social justice in Hong Kong, Mainland China, and other countries in Asia.

East meets west in curriculum studies

I am actively engaged in research in-between the Eastern and Western philosophy, curriculum, and pedagogy in education with a particular focus on Confucius, John Dewey, Tsunesaburo Makiguchi, Daisaku Ikeda, Weiming Tu, Martha Nussbaum, and Edward Saïd. I have been developing cross-cultural, bilingual/bicultural, and experiential approaches to curriculum, teaching, learning, and research in relation to bilingual/bicultural/biracial education and multilingual/multicultural/multiracial education, particularly the education of dislocated/exile/in-between, disenfranchised, and underrepresented individuals and groups. I particularly explore an East ~ West epistemological convergence of embodied democracy in education through cultural humanism illuminated in five main themes in the works of John Dewey (1859–1952), Confucius (551–479 BC), Ikeda Daisaku (b. 1928), and Tsunesaburo Makiguchi (1871–1944): *human-nature interconnection, self-cultivation, value creation, associated living, and joy of learning/happiness of living* (He, 2016). The cultural humanism, embedded in languages and cultures, flourishes with the ideal of embodied democracy that aims to educate for creative, harmonious, associative, joyful, and worthwhile living (e.g., Schubert, 2009a, b) for all in an increasingly diversified, complicated, and contested world.

The shared perspectives of Confucius, Ikeda, Makiguchi, and Dewey on cultural humanism, which transcend languages, cultures, and histories, create an East ~ West epistemological convergence of embodied democracy. In the midst of obstacles and challenges, I imagine cultivation of an East ~ West convergence of embodied democracy in education that fuses Confucianism, Makiguchianism, Ikedanism, Deweyanism, and philosophies from other Eastern and Western civilizations to counter globalization, standardization, and accountability; economic, racial, cultural-linguistic, social, gender and environmental injustice, acquisitiveness, and dislocation; and all forms of suppressions, repressions, and oppressions. Such a divergent East ~ West convergence of embodied democracy illuminated in cultural humanism in the works of Makiguchi, Ikeda, Confucius, and Dewey revitalizes human-nature interconnectedness, self-cultivation, value creation, associated living, and joy of learning/happiness of living, releases cross-cultural imagination, indigenizes Makiguchian, Ikedan, Confucius, and Deweyan educational ideals and philosophical traditions toward freedom and justice, cultivates humanity, and invigorates love, justice, and education for a better human condition for all (Schubert, 2009a).

Exile pedagogy/curriculum and teaching in-between

My cross-cultural narrative inquiry leads to my exploration of exile pedagogy, exile curriculum, and teaching in-between. I draw upon the works of Homi Bhabha on *location of culture* (1994) and *culture's in-between* (1996/2008); Gilles Deleuze and Félix Guattari (1972, 1980, 1991) on *rhizome, interbeing, intermezzo, multiplicity, and line of flight*; M. M. Bakhtin (1981) on *double-voiced, double accented, and double languaged hybrid*; Ted Aoki (2005a, b) on *teaching as indwelling between curriculum worlds*; Edward Saïd on *interstiti spaces and exile*; and Ming Fang He on *teaching, learning, and living in-between* (2003, 2006), *exile pedagogy—teaching in-between* (2010), *East ~ West epistemological convergence* (2016), *people in diaspora* (2018), and *diaspora curriculum theorizing* (2021).

Exile pedagogy and curriculum hold three qualities of in-betweenness: the epistemology of in-betweenness, the courageous aspect of teaching in-between, and the contested nature of teaching in-between. The epistemology of in-betweenness are illustrated in a wide array of literature such as philosophers in exile (e.g., Grathoff, 1989); women in exile (e.g., Afkhami, 1955/1983); writers in exile (e.g., Robinson, 1994); art of memory in exile (e.g., Píchová, 2002); exilic and diasporic filmmaking (e.g., Naficy, 2001); film, media, and the politics of place (e.g., Naficy, 1999); the making of exile cultures (e.g., Naficy, 1993); exiles and communities (e.g., Pagano, 1990); postmodern discourses of displacement (e.g., Kaplan, 1996); exiles, diasporas, and strangers in art (e.g., Mercer, 2008); reluctant exiles (e.g., Skelton, 1994); feminism, diasporas, and neoliberalisms (e.g., Grewal, 2005); dialectics of exile (e.g., McClennen, 2004); and contested landscapes: movement, exile, and place (Bender and Winer, 2001). There is more a sense of blurredness, intersectionality, or multiplicity and a sense of being in the midst in approaches to exile in arts, films, media, fictions, and poems. This discursive, multifaceted, complicated, sometimes contested nature of exile is in line with an evolving in-betweenness of exile pedagogy and curriculum.

The courageous aspect of teaching in-between is particularly influenced by radical democratic orientations illuminated through Freire's *pedagogy of the oppressed* (1970/1992), education for critical consciousness (1997), and *teachers as cultural workers* (1998);

Freire & Faundez's *learning to question/pedagogy of liberation* (1989); Palmer's (1998) *courage to teach*; public pedagogy of Sandlin et al. (2010); Schultz's *spectacular things happen along the way* (2008/2018), *listening to and learning from students* (2011), and *teaching in the cracks* (2017); Giroux *teachers as intellectuals* (1988) and *critical public pedagogy* (2004); Kincheloe's (2008) *critical pedagogy*; Asante's (2017) *revolutionary pedagogy: A primer for teachers of Black children*; King's (2015) *Dysconscious racism, Afrocentric praxis and education for human freedom*; McLaren's (2002) *critical pedagogy*; Grande's (2004/2015) *Red pedagogy*; Lather's (1991, 1998) *feminist pedagogy*; Kozol's *savage inequalities* (1991) and *letters to a young teacher* (2008); Crocco et al. (1999) *pedagogies of resistance*; Simon's (1991) *teaching against the grain*; Cochran-Smith's (1991, 2001, 2004) *learning to teach against the grain and teach for social change*; teaching for social justice of Ayers, Hunt, and Quinn (1998); *learning to teach for social justice* of Darling-Hammond, French and Garcia (2002); bell hooks' *teaching to transgress* (1994), *teaching community* (2003), and *teaching critical thinking* (2010), Said's *intellectuals as exiles* (1994), Ayer's (2004, 2006) *teaching toward freedom*, and *teaching the taboo* of Ayers and Ayers (2011).

The radical democratic quality of teaching creates an in-between space for educational workers to *exile voluntarily to teach in-between* (He, 2010). This aspect of exile in-between is illuminated in an oral tradition of Confucianism: A good teacher should be able to remove himself/herself from the crazy materialistic world, seek a balanced human condition in-between unbalanced and contested contradictions and complexities within nature and humanity, and develop a clear vision to *cultivate beauty, integrity, justice, and humanity* (also see Schubert, 2009a). Many educational workers who choose to teach in-between not only question [what] is worthwhile [to teach], for whom it is worthwhile [to teach], and how we make [teaching] worthwhile (Schubert, 2009b, p. 136), but also confront issues of equity, equality, social justice, societal change, and democratic human conditions through pedagogical theory and praxis.

The contested nature of teaching in-between draws substantially on the works of race, gender, class, and power of Black feminist thought/Black womanism (e.g., Baszile, 2015; Baszile et al., 2016; Davis, 1983; Guy-Sheftall, 1995; Collins, 1991/2000; hooks, 1999; James and Sharpley-Whiting, 2000/2001; McClaurin, 2001; Morrison, 1992, 2008; Phillips, 2006; Ross, 2015; Smith, 1983/2000; Walker, 1967/1983, 1997); Chicana feminist thought (e.g., Anzaldúa and Keating, 2002; Arredondo et al., 2003; García, 1997); critical race feminism (e.g., Wing, 2000, 2003); third world feminism (e.g., Mohanty, 2003/2005; Narayan, 1997); post/neocolonial feminism/ecofeminism (e.g., Anzaldúa, 1987, 1990; Mies and Shiva, 1993; Minh-ha, 1989; Mohanty, 2003/2005); Native American social and political thought (e.g., Grande, 2004/2015; Lomawaima, 1994; McCarty, 2002; Ng-A-Fook, 2007); indigenous ways of being, knowing, and doing (e.g., Archibald, 2008; Archibald et al., 2019; Tuhiwai Smith, 1999/2012; Tuhiwai Smith et al., 2018; Wilson, 2008; also Denzin et al., 2008); postmodern geography (e.g., Soja, 1989); the third space (Soja, 1996/2000; Wang, 2004, 2006, 2010); critical geography (e.g., Harvey, 2001); social and cultural geography (Del Casino and Vincent, 2009; Anderson et al., 2003); and spatial justice (Soja, 2010). While the intersectionality of race, gender, class, and power is always central to Black feminist thought/Black womanism, the intersectionality of repatriarchal historical analysis, spirituality, migration, displacement, slavery, racism, sexism, classism, imperialism, colonialism, heterosexism, ageism, ableism, anthropocentrism (i.e., human supremacism), speciesism, and other forms of oppression is illuminated in Chicana feminist thought, critical race feminism, third world feminism, and post/neocolonial feminism/ecofeminism. Indigenous ways of being, knowing, and doing connect critical theory with indigenous knowledge and socio-political contexts of indigenous education to develop transcendent theories of decolonization and advocate the liberty of indigenous language and cultural rights and intellectualism. Postmodern geographers, critical geographers, social and cultural geographers engage in this complex intersectionality within vibrant special dynamics of socio-political, cultural, and linguistic contradictions, complexities, and possibilities.

With three qualities distinguished from other forms of pedagogy and curriculum: the epistemology of in-betweenness, courageous aspect of teaching in-between, and the contested nature of teaching in-between, exile pedagogy and curriculum inspire us to teach to tensions, contradictions, and complexities in-between contested race, gender, class, and power with equity, equality, social justice, and human freedom as explicit goals. The power of exile pedagogy and curriculum lie in educational workers' strong advocacy on behalf of individuals, groups, families, tribes, communities, and societies that are often at controversy, underrepresented, misrepresented, or excluded in official narratives. Exile pedagogy and curriculum connect the personal with the political and the practical with the theoretical through passionate participation in and critical reflection on teaching, learning, inquiry, and life with an "epistemological curiosity" (Freire and Macedo, 1995, p. 382) and with passion particularly in hard times. Exile pedagogy and curriculum encourage us to teach courageously in-between contradictions; move beyond boundaries; transgress orthodoxies; question mandates and regulations; educate rather than profit (Nussbaum, 2010); and "build on long-term, heart-felt engagement and shared efforts driven by commitment to equity, equality, social justice, freedom, and human possibility" (He, 2010, p. 471). Exile pedagogy and curriculum create a culture of resistance and builds a participatory movement to promote a more balanced and equitable human condition through personal and political acts of teaching in an increasingly diversified and contested world.

Diaspora curriculum theorizing

Most recently, I venture into diaspora curriculum theorizing. Diaspora curriculum draws upon a wide array of theoretical traditions such as conceptions of diasporas (e.g., Bruneau, 2010; Safran, 1991; Tölölyan, 1991); the breadth, diversity, and complexity of diasporas (e.g., Bhabha, 1990, 1994; Clifford, 1994, 1997; Hall, 1999, 2017a,b, 2019); diaspora consciousness, diasporic space, and the in-betweenness of diasporas (e.g., Aoki, 2005a,b; Bakhtin, 1981; Bhabha, 1994, 1996/2008; Clifford, 1994, 1997; Deleuze and Guattari, 1972, 1980, 1991; He, 2003, 2006, 2010, 2016, 2018; Said, 1994, 2000, 2002); decolonizing diasporas (e.g., Dei,

2017; Dei and Lordan, 2016; Figueroa-Vásquez, 2020; Haig-Brown, 2009, 2012; Tuck and Yang, 2012); diasporic imaginaries and diasporic futurism (e.g., Antonsich, 2010; Axel, 2002; Mishra, 1996; Steger, 2008), Afrofuturism (e.g., Anderson and Jones, 2016; Dery, 1994; Gipson, 2019; Jackson and Moody-Freeman, 2011; Womack, 2013), and Indigenous futurism (e.g., Dillon, 2012; Lempert, 2014, 2018; also Alfred, 1999; Ortiz, 2009).

Diaspora curriculum draws upon a wide array of conceptions of diaspora (e.g., Bruneau, 2010; Safran, 1991; Tölölyan, 1991). Diaspora (in Greek, διασπορά—“a scattering [of seeds]”) refers to the movement of a population sharing common ethnic identity who are either forced to leave or voluntarily leave their indigenous or ancestral lands and become residents in areas often far removed from their former homes (He, 2010, 2018). In a broader sense, diaspora refers to situations when indigenous peoples, immigrants, and emigrants are forced to leave or voluntarily leave their tribes, native lands, territories, communities, or countries due to such reasons as imperialism, colonialism, political persecution, economic exploitation, trade or labor migrations. While people in diaspora might not maintain strong ties with their homelands or native lands, they lack full integration into the host lands. This mobile and unsettling existence of diasporas complicates and creates *multiple senses of belonging* (Bader and Schieren, 2019). Diaspora is dedicated to the multidisciplinary study of the history, culture, social structure, politics and economics of both the traditional diasporas—Armenian, Greek, and Jewish—and those transnational dispersions which in the past three decades have chosen to identify themselves as “diasporas.” These encompass groups ranging from the African-American to the Ukrainian-Canadian, from the Caribbean-British to the new East and South Asian diasporas.

While engendered by the mobile, unsettling, interdisciplinary, and transnational existence of diasporas, diaspora curriculum also draws upon the breadth, diversity, and complexity of diasporas. A group of anthropologists or cultural theorists (e.g., Bauböck and Faist, 2010; Bhabha, 1990, 1994, 1996/2008; Clifford, 1994, 1997; Hall, 1999, 2017a,b, 2019) broadened, diversified, and complicated the conceptions of diasporas. Drawing from Derrida’s (1982) notion of *différance* and Caribbean cultural diasporic aesthetics, Stuart Hall (1999) elaborated the unsettledness and unresolvableness of diasporas. For Hall (1999), the conception of diasporas cannot rest on “a binary conception of difference” (p. 7). Diaspora is “positional and relational, always on the slide along a spectrum without end or beginning” (p. 7). The reconfiguration of diasporas cannot be represented as “going back to where we were before ... there is always something else between” (p. 8). Diasporas “are dispersed forever from their homelands, to which they cannot literally return. Diasporas belong to “imagined communities to which there can be no return” (p. 172). They “must learn to inhabit more than one identity, dwell in more than one culture, and speak more than one language ... to speak in the unsettling place in-between languages means to constantly negotiate and translate across their differences” (p. 173). Such a diasporic experience invents “home from home” (Hall, 2017a,b, p. 172), a diasporic consciousness, a diasporic perspective, and a diasporic narrative to capture “the full complexities” and the displacements of “a collective predicament” (p. 171).

Paul Gilroy (1993) questioned “ethnic absolutism in cultural criticism” (p. 3) regarding the Black Atlantic diaspora. Gilroy perceived the Black Atlantic diaspora as “a counterculture of modernity” (pp. 1–40), a culture as “a non-traditional tradition, an irreducibly modern, exocentric, unstable, and asymmetrical cultural ensemble that cannot be apprehended through the Manichean logic of binary coding” (p. 198). This counterculture “is not specifically African, American, Caribbean, or British, but all of these at once, a Black Atlantic culture whose themes and techniques transcend ethnicity and nationality to produce something new” (Lott, 1993, cover). He also explores the Black Atlantic diaspora as it is manifested in black writing from the “double consciousness” of W. E. B. Du Bois to the “double vision” of Richard Wright to the compelling voice of Toni Morrison. By articulating the temporality, historicity, memory, and narrativity of the Black Atlantic countercultures, Gilroy (1993) reminds us “the long history of slavery, the legacy of scientific racism, and the complicity of rationality and terror in distinctively modern forms of domination” and “a long struggle for political and social emancipation, and critical visions of equality and difference” (Clifford, 1994, p. 317) that have been generated in the Black Atlantic diaspora. Drawing from Jacques Derrida’s notion of *différance* and cultural diasporic aesthetics, Homi Bhabha, James Clifford, and Stuart Hall broadened the conceptions of diasporas and complicates multiple senses of belonging for people in diaspora, which encapsulates the similar mobile, unsettling, interdisciplinary, and transnational nature of cross-cultural narrative inquiry. This diaspora curriculum theorizing is deeply embedded in my life experience as a Chinese woman scholar and a curriculum worker wondering amongst constantly changing and complicating Eastern and Western theoretical traditions, histories, languages, cultures, and geographical locations.

Diaspora curriculum is illuminated by the breadth, diversity, and complexity of diasporas through diaspora consciousness and diasporic space in the works of James Clifford (1994, 1997) on *diaspora consciousness and diasporic space*, Homi Bhabha on *location of culture* (1994) and *culture’s in-between* (1996/2008), Gilles Deleuze and Félix Guattari (1972, 1980, 1991) on *rhizome, interbeing, intermezzo, multiplicity, and line of flight*, M. M. Bakhtin (1981) on *double-voiced, double accented, and double language hybrid*, and Ted Aoki (2005a, b) on *teaching as indwelling between curriculum worlds* (He, 2021, 2022).

To theorize diasporas (Brazier and Mannur, 2003) begins with decolonizing diasporas. To decolonize diasporas, we need to “involve repatriation of Indigenous land and abolition of slavery” (Tuck, 2015, p. 444). Too often “diaspora theory ignores the presence of Indigenous peoples who were already in place when the first and the last diasporic peoples were forced or chose to come to a land” (Haig-Brown, 2009, p. 16). For Haig-Brown (2009), “to ignore the trauma of those people who have been displaced here to make room, first for the colonizers and then for those who came after under all sorts of the other conditions—from slavery to starvation to war to straightforward immigration for the promise [of] a new world—is to perpetuate ... epistemic violence” (p. 16). Haig-Brown (2009) called the diaspora theorists to “take seriously not only the land from which one comes, but the land and original people of the place where one arrives” (p. 16).

To decolonize diasporas, we need to “unsettle and dismantle the logics of colonial modernity/coloniality” and recognize the resistance of peoples of African descent as Afro-diasporic imaginaries—Afro-futurities—that “subvert and resist modernity and

coloniality” and offer new ways to approach questions of home, location, belonging, difference, and justice (Figueroa-Vásquez, 2020, p. 16). Decolonizing diasporas helps create a critical space to “examine how themes of intimacy, witnessing, dispossession, reparations, and futurities are remapped ... by tracing interlocking structures of oppression, including public and intimate forms of domination, sexual and structural violence, sociopolitical and racial exclusion, and the haunting remnants of colonial intervention” (Figueroa-Vásquez, 2020, pp. 1–2). Further, understanding of theories of decolonization and anti-colonialism is “a continuous process of challenging colonial and neo-colonial legacies, relations and power dynamics, taking up the subject of anti-colonial praxis and its specific implications—the larger questions of schooling and education in global and, particularly, diasporic contexts” (Dei and Lordan, 2016, p. vii). The *cartographies* (brah, 1996) of decolonizing diaspora can be also traced to a wide array of works on decolonization such as Karen R. Lawrence (1992) on *decolonizing tradition*, Linda Tuhiwai Smith (1999/2012) on *decolonizing methodologies*, Jo-ann Archibald (2008) on *Indigenous storywork: educating the heart, mind, body, and spirit*, Kuan-Hsing Chen (2010) on *Asia as method*, Angel Lin and Peter Martin (2005) on *decolonization and globalization*, Kagendo Mutua and Beth Blue Swadener (2004) on *decolonizing research in cross-cultural contexts*, Leigh Patel (2016) on *decolonizing educational research*, Bagele Chilisa (2012) on *Indigenous research methodologies*, Margaret Kovach (2009/2012) on *Indigenous methodologies*, and Chela Sandoval (2000) on *methodology of the oppressed*.

In addition to its breadth, diversity, and complexity, another illuminating aspect of diaspora curriculum is the evolving diasporic imaginaries (e.g., Axel, 2002). Instead of dwelling on the traditional notion of diaspora as a spatial separation between homeland and host country, the “diasporic imaginary” is a space of imagination which “account[s] for the creation of the diaspora ... through formations of temporality, affect, and corporeality” (Axel, 2002, p. 412). The “diasporic imaginary” reinterprets the imaginary as the creation of a shared diasporic space of dreams, imaginations, and visions to cultivate diasporic communities—“imagined communities” (Anderson, 1991). The “diasporic imaginary” (Mishra, 1991, 2007) is interconnected with “global imaginary” (Steger, 2008) as a consciousness of belonging to a global community of shared ideas and interests. Diasporic imaginaries represent a multiplicity of temporalities, localities, traditions, identities, and subjectivities engendered by shared diasporic experiences, multiple senses of belonging, and multiple processes of becoming.

Afrofuturism (Dery, 1994; Nelson, 2002) and Indigenous futurism help decolonize diasporic imaginaries to center shared diasporic experiences of dispossession, deprivation, oppression, exploitation, violence, exclusion, persecution, and invasion. Afrofuturism is a cultural aesthetic, philosophy of science and philosophy of history that explores the developing intersection of African diaspora culture with technology. Afrofuturism addresses themes and concerns of the African diaspora through technoculture and science fiction, encompassing a range of media and artists with a shared interest in envisioning Black futures that stem from Afrodiasporic experiences. An Afroturistic concept, *Astro-Blackness*, “represents the emergence of a black identity framework within global technocultural assemblages, migration, human reproduction, algorithms, digital networks, software platforms, bio-technical augmentation ... [where] racialized identities are increasingly materialized ... humans and techniques have co-evolved together” (Anderson and Jones, 2016). Afrofuturism is “an intersection of imagination, technology, the future and liberation” (Womack, 2013). Afrofuturism is the philosophy of science fiction, and history that traverses across African Diaspora cultures with technology to explore the African American experience, specifically slavery. Afrofuturism is “a way of imagining possible futures through a black cultural lens ... [It is] a way to encourage experimentation, reimagine identities, and active liberation” (LaFleur, 2011).

Indigenous Futurism (Dillon, 2012) is a movement consisting of art, literature, comics, games, and other forms of media which express Indigenous perspectives of the future, past, and present in the context of science fiction and related sub-genres. Such perspectives may reflect Indigenous ways of knowing, traditional stories, historical or contemporary politics, and cultural realities. Like Afrofuturism, Indigenous Futurism encapsulates multiple modes of art-making from literature to visual arts, fashion and music, processes of decolonization (Dillon, 2012). Using tools like slipstream, worldbuilding, science fiction and anthropological First Contact scenarios, Indigenous communities construct self-determined representations and alternative narratives about their identities and futures. Indigenous Futurists critique the exclusion of Indigenous peoples from the contemporary world and challenge notions of what constitutes advanced technology. Indigenous Futurists question the digital divide and recognize that Indigenous peoples have been purposefully excluded in history from accessing media technologies and exist outside of modernity. The widespread use of personal computers and the Internet following the Digital Revolution created conditions in which, to some extent, Indigenous peoples may participate in the creation of a network of self-representations.

Inspired by diasporic imaginaries, particularly Afrofuturism and Indigenous Futurism, diasporic futurism is a multicultural, multilingual, and multiracial aesthetic, philosophy of science and philosophy of history that explores the developing intersection of language, culture, identity, race, gender, class, power, and place in diaspora with technology. Diasporic futurism addresses themes and concerns of the multicultural, multilingual, and multiracial aesthetic diaspora through technoculture and science fiction, encompassing a range of media and artists with shared interests in envisioning alternative futures that stem from diasporic experiences. Such alternative futures shatter white supremacy, racism, xenophobia, misogyny, homophobia, transphobia, islamophobia and settler colonialism that perpetuating an emboldened hatred of the difference of minoritized communities.

Cross-cultural narrative imagination in hard times

In addition to its breadth, diversity, and complexity, another illuminating aspect of cross-cultural narrative inquiry is the evolving cross-cultural imaginaries, where we can keep our hopes and dreams alive in hard times when white supremacy, anti-Black/anti-Indigenous/anti-Asian racism, xenophobia, misogyny, homophobia, transphobia, Islamophobia, and settler colonialism are

perpetuated by hatred of differences (He et al., 2021). Such cross-cultural imaginaries invent cross-cultural futurities and cultivate radical possibilities (Anderson and Jones, 2016; Anyon, 2005; Womack, 2013) and radical imagination (Freire, 2007; Giroux, 2007; Greene, 1995; Jackson and Moody-Freeman, 2011; Kelley, 2002; Lake, 2013; Mills, 1959/2000; Olson and Worsham, 2007). Such cross-cultural futurities exhilarate cross-cultural consciousness that educates hope (Harvey, 2000), evokes “different histories and different futures” (Giroux, 2017, p. xiii), and cultivates radical love so that we “could find the strength, faith, and humility to establish solidarity and struggle together to transform the oppressive ideologies and practices” (Darder, 2017, p. 80) to build solidarity across differences to fight against myriad forms of oppression. Such cross-cultural futurities inspire “optimism over despair” (Chomsky, 2017), love over hatred, and possibilities over impossibilities (Ayers, 2016; hooks, 2000; Darder, 2017).

Such cross-cultural imaginaries inspire us to invent creative strategies to teach, research, and live toward freedom and social justice in troubling times. Such cross-cultural imaginaries continuously encourage us to cultivate “solidarity, integrity, and humility” (Stovall, 2016, p. 54) to counter authoritarian and dominant narratives and transgress orthodoxies and bureaucratic procedures (e.g., Archibald, 2008; Archibald et al., 2019; Chilisa, 2012; hooks, 1994; Love, 2019; Maparyan, 2012; Patel, 2016; Phillips, 2006; Sandoval, 2000; Tuck, 2015; Tuck and Yang, 2012; Tuhiwai Smith, 1999/2012; Tuhiwai Smith et al., 2018; Walker, 1997; Wilson, 2008). Such cross-cultural imaginaries enable us to build up courage to *voluntarily exile* (He, 2010) from *commodified* (Illich, 1970), *acquisitive* (Dewey, 1933; Schubert, 2009a), and *deskilling* societies (Apple, 1982, 1986, 2001) to *make the impossible possible* (Ayers, 2016), to keep the boundless human potential evolving (Nussbaum, 1997; Schubert, 2009b; Siddle-Walker, 1996), to create “more equitable and just public spheres within and outside of educational institutions” (Mohanty, 1989, p. 207), to build better human conditions in-between contradictions and complexities, and to passionately participate in the life of schools, families, communities, places, and spaces. Such cross-cultural futurities invigorate cross-cultural space for “imagined communities” (Bhabha, 1994, p. 5) where cross-cultural workers work with other educational workers such as researchers, educators, teachers, administrators, parents, students, community workers, and policy makers to heal the humanity and planet with “shared principles and visions” for “desirable collective futures” in “a world of increasing complexity, uncertainty, and fragility” (UNESCO, 2020, pp. 2–5) and to build diverse, resilient, healthy, peaceful, and “thriving sustainable futures” (UNESCO, 2022a,b, p. 5) for all.

References

- Afkhami, M., 1955/1983. *Women in Exile*. University of Virginia Press, Charlottesville, VA.
- Alfred, T., 1999. *Peace, Power, Righteousness: An Indigenous Manifesto*, second ed. Oxford University Press, North York, ON.
- Anderson, B., 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso, London.
- Anderson, K., Domosh, M., Pile, S., Thrift, N. (Eds.), 2003. *Handbook of Cultural Geography*. Sage, Thousand Oaks, CA.
- Anderson, R., Jones, C.E. (Eds.), 2016. *Afrofuturism 2.0: The Rise of Astro-Blackness*. Lexington Books, Lanham, MD.
- Antonsich, M., 2010. Searching for belonging: an analytical framework. *Geography Compass* 4 (6), 644–659.
- Anyon, J., 2005. *Radical Possibilities: Public Policy, Urban Education, and a New Social Movement*. Routledge, New York, NY.
- Anzaldúa, G., 1987. *Borderlands/La Frontera: The New Mestiza*. Spinsters/Aunt Lute, San Francisco, CA.
- Anzaldúa, G. (Ed.), 1990. *Making Face, Making Soul: Creative and Critical Perspectives by Women of Color*. Aunt Lute Foundation Books, San Francisco, CA.
- Anzaldúa, G., Keating, A. (Eds.), 2002. *This Bridge We Call Home: Radical Visions for Transformation*. Routledge, New York, NY.
- Aoki, T., 2005a. In the midst of slippery theme words: living as designers of Japanese Canadian curriculum (1992). In: Pinar, W.F., Irwin, R.L. (Eds.), *Curriculum in a New Key*. Lawrence Erlbaum, Mahwah, NJ, pp. 263–278.
- Aoki, T., 2005b. Teaching as in-dwelling between two worlds (1986/1991). In: Pinar, W.F., Irwin, R.L. (Eds.), *Curriculum in a New Key*. Lawrence Erlbaum, Mahwah, NJ, pp. 159–165.
- Apple, M., 1982. *Education and Power*. Routledge & Kegan Paul, New York, NY.
- Apple, M., 1986. *Teachers and Texts: A Political Economy of Class and Gender Relations in Education*. Routledge & Kegan Paul, New York, NY.
- Apple, M.W., 2001. *Education the “Right” Way*. Routledge; Falmer, New York, NY.
- Archibald, J., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. UBC Press, Vancouver.
- Archibald, J., Xie, Q.Q., Lee-Morgan, J.B.J., De Santolo, J. (Eds.), 2019. *Decolonizing Research: Indigenous Storywork as Methodology*. Zed Books, London.
- Arredondo, G.F., Hurtado, A., Klahn, N., Najera-Ramirez, O., Zavella, P. (Eds.), 2003. *Chicana Feminisms: A Critical Reader (Post-Contemporary Interventions)*. Duke University Press, Durham, NC.
- Asante, M.K., 2017. *Revolutionary Pedagogy: A Primer for Teachers of Black Children*. Universal Write, Brooklyn, NY.
- Axel, B.K., 2002. The diasporic imaginary. *Publ. Cult.* 14 (2), 411–428. <https://www.muse.jhu.edu/article/26278>.
- Ayers, W., 2016. *Demand the Impossible: A Radical Manifesto*. Haymarket Books, Chicago, IL.
- Ayers, W.C., 2004. *Teaching Toward Freedom: Moral Commitment and Ethical Action in the Classroom*. Beacon, Boston, MA.
- Bader, M., Schieren, 2019. *Diasporic imaginaries, multiple senses of belonging: call for proposal*. *Art Production and Art Theory in the Age of Global Migration*; German Center for Art History Paris, Paris.
- Bakhtin, M.M., 1981. Forms of time and of the chronotope in the novel: notes toward a historical poetics. In: Holquist, M. (Ed.), *The Dialogic Imagination: Four Essays by M.M. Bakhtin*. University of Texas Press, Austin, TX, pp. 84–258.
- Ball, D.L., O'Connor, C., Wilson, S., 2017, June 15. AERA 2018 Annual Meeting Call for Submissions. Available at: http://www.aera.net/Portals/38/docs/Annual_Meeting/2018%20Annual%20Meeting/AERA_AM2018_CallforPapers_1.1.pdf?ver=2017-06-01-194603-513×tamp=1496360770573.
- Baszile, D.T., 2015. Critical race/feminist currere. In: He, M.F., Schultz, B.D., Schubert, W.H. (Eds.), *The Sage Guide to Curriculum in Education*. Sage, Thousand Oaks, CA, pp. 119–126.
- Baszile, D.T., Edwards, K.T., Guillory, N.A., 2016. *Race, Gender, and Curriculum Theorizing: Working in Womanish Ways*. Lexington Books/An imprint of Rowman & Littlefield, Lanham, MD.
- Bauböck, R., Faist, T. (Eds.), 2010. *Diaspora and Transnationalism: Concepts, Theories and Methods*. Amsterdam University Press, Amsterdam.
- Bender, B., Winer, M. (Eds.), 2001. *Contested Landscapes: Movement, Exile and Place*. Berg, Oxford.
- Bhabha, H.K., 1990. Dissemination: time, narrative, and the margins of the modern nation. In: Bhabha, H.K. (Ed.), *Nation and Narration*. Routledge, London, pp. 291–322.
- Bhabha, H.K., 1994. *The Location of Culture*. Routledge, New York, NY.
- Bhabha, H.K., 1996/2008. *Cultures in-between*. In: Hall, S., du Gay, P. (Eds.), *Questions of Cultural Identity*. Sage, Thousand Oaks, CA, pp. 53–60.

- Blaise, C., 1993. *I Had a Father: A Post-modern Autobiography*. Addison-Wesley, Reading, MA.
- Brah, A., 1996. *Cartographies of Diaspora: Contesting Identities*. Routledge, New York, NY.
- Brazier, J.E., Mannur, A. (Eds.), 2003. *Theorizing Diaspora*. Blackwell, Malden, MA.
- Bruneau, M., 2010. Diasporas, transnational spaces and communities. In: Bauböck, R., Faist, T. (Eds.), *Diaspora and Transnationalism: Concepts, Theories and Methods*. Amsterdam University Press, Amsterdam, pp. 35–49.
- Chamoiseau, P., 1994. *School Days*. University of Nebraska Press, Lincoln, NE.
- Chen, K., 2010. *Asia as Method: Toward Deimperialization*. Duke University Press, Durham, NC.
- Chilisa, B., 2012. *Indigenous Research Methodologies*. Sage, Thousand Oaks, CA.
- Chomsky, N., 2017. *Optimism Over Despair: On Capitalism, Empire, and Social Change (Interviews by C. J. Polychroniou)*. Haymarket Books, Chicago, IL.
- Cisneros, S., 1994. *The House on Mango Street*. A.A. Knopf; Random House, New York, NY.
- Clandinin, D.J., Connelly, F.M., 2000. Narrative Inquiry: Experience and Story in Qualitative Research. Jossey-Bass, San Francisco, CA.
- Clifford, J., 1994. Diasporas. *Cult. Anthropol.* 9 (3), 302–338. Available at November 26, 2020. <http://www.jstor.org/stable/656365>.
- Clifford, J., 1997. *Routes: Travel and Translation in the Late Twentieth Century*. Harvard University Press, Cambridge, MA.
- Collins, P.H., 1991/2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge, New York, NY.
- Cowan, P., 1982. *An Orphan in History: Retrieving a Jewish Legacy*. Doubleday, Garden City, NY.
- Crow Dog, M., Erdoes, R., 1990. *Lakota Woman*. Harper Collins, New York, NY.
- Darder, A., 2017. *Reinventing Paulo Freire: A Pedagogy of Love*. Routledge, New York, NY.
- Davis, A.Y., 1983. *Women, Race, & Class*. Vintage Books, New York, NY.
- Dei, G.J.S., 2017. *Reframing Blackness and Black Solidarities Through Anti-Colonial and Decolonial Prisms*. Springer, Cham.
- Dei, G.J.S., Lordan, M. (Eds.), 2016. *Anti-Colonial Theory and Decolonial Praxis*. Peter Lang, New York, NY.
- Del Casino Jr., V.J., Vincent, J., 2009. *Social Geography*. Wiley-Blackwell, West Sussex.
- Deleuze, G., Guattari, F., 1972. *Anti-Oedipus: Capitalism and Schizophrenia*. Penguin Books, New York, NY.
- Deleuze, G., Guattari, F., 1980. *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press, Minneapolis, MN.
- Deleuze, G., Guattari, F., 1991. *What is Philosophy?* (Trans. Hugh Tomlinson & Graham Burchell) Columbia University Press, New York, NY.
- Denzin, N.K., Lincoln, Y.S., Tuhiwai Smith, L. (Eds.), 2008. *Handbook of Critical and Indigenous Methodologies*. Sage, Thousand Oaks, CA.
- Derrida, J., 1982. *Margins of Philosophy* (Trans. Alan Bass). University of Chicago Press, Chicago, IL.
- Dery, M., 1994. *Flame Wars: The Discourse of Cyberculture*. Duke University Press, Durham, NC.
- Dewey, J., 1933. Dewey outlines utopian schools. *N. Y. Times*. April 23, p. 7. (Also in Boydston, J. A. (Ed.), (1989), *Dewey: The later works, 1925–1953* (vol. 9; pp. 136–140). Carbondale, IL: Southern Illinois University Press.
- Dillon, G.L. (Ed.), 2012. *Walking the Clouds: An Anthology of Indigenous Science Fiction*. University of Arizona Press, Tucson, AZ.
- Dorfman, A., 1998. *Heading South and Looking North: A Bilingual Journey*. Penguin Books, New York, NY.
- Figuerola-Vásquez, Y.C., 2020. *Decolonizing Diasporas: Radical Mappings of Afro-Atlantic Literature*. Northwestern University Press, Evanston, IL.
- Firmat, G.P., 1994. *Life on the Hyphen: The Cuban-American Way*. University of Texas Press, Austin, TX.
- Freire, P., 2007. *Daring to Dream: Toward a Pedagogy of the Unfinished*. Organized and Presented by Ana Maria Araújo Freire. Forewords by Peter Park and Ana Lúcia Souza De Freitas. Paradigm, Boulder, CO.
- Freire, P.M., Macedo, D., 1995. A dialogue: culture, language, and race. *Harv. Educ. Rev.* 65 (3), 379–382.
- Gay, G., 2000/2018. *Culturally Responsive Teaching: Theory, Research, and Practice*. Teachers College Press, New York, NY.
- García, A.M., 1997. *Chicana Feminist Thought: The Basic Historical Writings*. Routledge, New York, NY.
- Gilroy, P., 1993. *The Black Atlantic: Modernity and Double Consciousness*. Harvard University Press, Cambridge, MA.
- Gipson, G., 2019. Creating and imagining black futures through afrofuturism. In: De Kosnik, A., Feldman, K.P. (Eds.), *#Identity: Hashtagging Race, Gender, Sexuality, and Nation*. University of Michigan Press, Ann Arbor, MI, pp. 84–103. Available at May 21, 2020. www.jstor.org/stable/j.ctvndv9md.9.
- Giroux, H., 2007. When the darkness comes and hope is subversive (foreword). In: Olson, G.A., Worsham, L. (Eds.), *The Politics of Possibility: Encountering the Radical Imagination*. Paradigm, Boulder, CO, pp. vii–xviii.
- Giroux, H., 2017, June 26. Thinking Dangerously: The Role of Higher Education in Authoritarian Times. Available at. <http://www.truth-out.org/opinion/item/41058-thinking-dangerously-the-role-of-higher-education-in-authoritarian-times>.
- González, N., Moll, L.C., Amanti, C., 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum, Mahwah, NJ.
- Grande, S., 2004/2015. *Red Pedagogy: Native American Social and Political Thought*. Roman & Littlefield, Lanham, MD.
- Grathoff, R., 1989. *Philosophers in Exile* (Trans. by J. C. Evans). Indiana University Press, Bloomington, IN.
- Greene, M., 1995. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*. Jossey-Bass, San Francisco, CA.
- Grewal, I., 2005. *Transnational America: Feminism, Diasporas, Neoliberalisms*. Duke University Press, Durham, NC.
- Gutiérrez, K., 2008. Developing a sociocritical literacy in the third space. *Read. Res. Q.* 43 (2), 148–164.
- Gutiérrez, K., Baquedano-Lopez, P., Tejeda, C., 1999. Rethinking diversity: hybridity and hybrid language practices in the third space. *Mind Cult. Activ.* 6, 286–303.
- Guy-Sheftall, B., 1995. *Words of Fire: An Anthology of African-American Feminist Thought*. The New Press, New York, NY.
- Haig-Brown, C., 2009. Decolonizing diaspora: whose traditional land are we on? *Cult. Pedagogical Inq.* 1 (1), 4–21.
- Haig-Brown, C., 2012. Decolonizing diaspora: whose traditional land are we on? In: Abdi, A.A. (Ed.), *Decolonizing Philosophies of Education*. Sense, Rotterdam, pp. 73–90.
- Hall, S., 1999. Thinking the diaspora: home-thoughts from abroad. *Small Axe* 6, 1–18.
- Hall, S., 2017a. *Fateful Triangle: Race, Ethnicity, and Nation*. Harvard University Press, Cambridge, MA.
- Hall, S., 2017b. *Familiar Stranger: A Life Between Two Islands*. Duke University Press, Durham, NC.
- Hall, S., 2019. Thinking the diaspora: home-thoughts from abroad. In: Hall, S., Morley, D. (Eds.), *Essential Essays*, vol. 2. Duke University Press, Durham, NC, pp. 206–226.
- Harvey, D., 2000. *Spaces of Hope*. University of California Press, Berkeley, CA.
- Harvey, D., 2001. *Spaces of Capital: Toward a Critical Geography*. Routledge, New York, NY.
- He, M.F., 1998. *Professional Knowledge Landscapes: Three Chinese Women Teachers Enculturation and Acculturation Processes in China and Canada*. Unpublished doctoral dissertation. University of Toronto, Canada.
- He, M.F., 1999. Featured article. A life-long inquiry forever flowing between China and Canada: crafting a composite auto/biographic narrative method to represent three Chinese women teachers cultural experiences. *J. Crit. Inq. Curric. Instr.* 1 (2), 5–29.
- He, M.F., 2002a. A narrative inquiry of cross-cultural lives: lives in China. Special issue on narrative multiculturalism. *J. Curric. Stud.* 34 (3), 301–321.
- He, M.F., 2002b. A narrative inquiry of cross-cultural lives: lives in Canada. Special issue on narrative multiculturalism. *J. Curric. Stud.* 34 (3), 323–342.
- He, M.F., 2002c. A narrative inquiry of cross-cultural lives: lives in North American Academe. *J. Curric. Stud.* 34 (5), 513–533.
- He, M.F., 2003. *A River Forever Flowing: Cross-Cultural Lives and Identities in the Multicultural Landscape*. Information Age, Greenwich, CT.
- He, M.F., 2006. In-between China and North America. In: Berry, T.R., Mizelle, N.D. (Eds.), *From Oppression to Grace: Women of Color and Their Dilemmas Within the Academy*. Stylus Publishing, Sterling, VA, pp. 68–76.
- He, M.F., 2010. Exile pedagogy: teaching in-between. In: Sandlin, J.A., Schultz, B.D., Burdick, J. (Eds.), *Handbook of Public Pedagogy: Education and Learning Beyond Schooling* (pp. 469–482). Routledge, New York, NY.

- He, M.F., 2016. Exploring an East~West epistemological convergence of embodied democracy in education through cultural humanism in Confucius~Makiguchi~Dewey. *J. Curric. Stud.* 48 (1), 36–57.
- He, M.F., 2018. Patriotism for people in diaspora is love of humanity. In: *The New Bank Street Occasional Paper Series—Am I Patriotic? Learning and Teaching the Complexities of Patriotism Here and Now*, vol. 40, pp. 80–83.
- He, M.F., 2022. Diaspora curriculum. with Associate Editors: Isabel Nuñez (Indiana University-Purdue University Fort Wayne), Patrick Roberts (Northern Illinois University), Sabrina Ross (Georgia Southern University), & Brian D. Schultz (Miami University, Oxford, OH). In: Schubert, W.H., He, M.F. (Eds.), *Oxford Encyclopedia of Curriculum Studies* (2086 Pages, 3 Volumes), Oxford Research Encyclopedias of Education. Oxford University Press, Oxford. G. W. Noblit (Ed.). <https://www.oxfordreference.com/view/10.1093/acref/9780190887988.001.0001/acref-9780190887988-e-1122>.
- He, M.F., Phillion, J. (Eds.), 2008. *Personal~Passionate~Participatory Inquiry into Social Justice in Education*. Information Age Publishing, Greenwich, CT.
- He, M.F., Yu, M., 2017. Education for cultural and linguistic diversity in the United States in hard times. *Curriculum Perspectives: The Journal of Australian Curriculum Studies Association* 37 (2), 205–210.
- He, M.F., Ross, S., 2013. Narrative of curriculum in the US South: lives in-between contested race, gender, class, and power. *Special Issue of Journal of Curriculum Theorizing* 28 (3), 196.
- He, M.F., Ross, S., Seay, K., 2015. Methodological dilemmas in social justice research. *J. Int. Crim. Justice* 1 (1), 70–106.
- He, M.F., Ngo, B., Bae-Dimitriadis, M., Matias, M.E., Sharma, S., Urrieta Jr., L., Rodriguez, S., 2021. Educating hope, radicalizing imagination, and politicizing possibility in hard times. *Educ. Stud.* 57 (3), 203–210.
- Hoffman, E., 1989. *Lost in Translation: A Life in a New Language*. Penguin Books, New York, NY.
- hooks, b., 1994. *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, New York, NY.
- hooks, b., 1999. *Aint I a Woman: Black Women and Feminism*. South End, Cambridge, MA.
- hooks, b., 2000. *All about Love*. William Morrow, New York, NY.
- hooks, b., 2003. *Teaching Community: A Pedagogy of Hope*. Routledge, New York, NY.
- hooks, b., 2010. *Teaching Critical Thinking: Practical Wisdom*. Routledge, New York, NY.
- Illich, I., 1970. *De-Schooling Society*. Harper and Row, New York, NY.
- Jackson, S., Moody-Freeman, J.E. (Eds.), 2011. *The Black Imagination: Science Fiction, Futurism and the Speculative*. Peter Lang, New York, NY.
- James, J., Sharpley-Whiting, T.D. (Eds.), 2000/2001. *The Black Feminist Reader*, third ed. Routledge, Malden, MA, New York, NY.
- Kaplan, A., 1993. *French Lessons: A Memoir*. University of Chicago Press, Chicago, IL.
- Kaplan, C., 1996. *Questions of Travel: Postmodern Discourses of Displacement*. Duke University Press, Durham, NC.
- Kelley, R.D.G., 2002. *Freedom Dream: The Black Radical Imagination*. Beacon, Boston, MA.
- Kincheloe, J.L., 2008. *Critical pedagogy*, 2nd ed. Peter Lang.
- King, J.E., 2015. *Dysconscious Racism, Afrocentric Praxis, and Education for Human Freedom: Through the Years I Keep on Toiling: The Selected Work of Joyce E. King*. Routledge, New York, NY.
- Kingston, M.H., 1975. *The Woman Warrior: Memoirs of Girlhood Among Ghosts*. Random House, New York, NY.
- Kovach, M., 2009/2012. *Indigenous Methodologies: Characteristics, Conversations, and Contexts*. University of Toronto Press, Toronto, ON.
- Ladson-Billings, G., 1994/2009. *The Dream Keepers: Successful Teachers of African American Children*. Jossey-Bass, San Francisco, CA.
- Ladson-Billings, G., 2021. *Culturally Relevant Pedagogy: Asking a Different Question*. Teachers College Press, New York, NY.
- LaFleur, I., 2011. Visual aesthetics of afrofuturism. Retrieved on June 30, 2022 from: <https://www.youtube.com/watch?v=x7bCaSzk9Zc>.
- Lake, R., 2013. *A Curriculum of Imagination in an Era of Standardization: An Imaginative Dialogue with Maxine Greene and Paulo Freire*. Information Age Publishing, Charlotte, NC.
- Lawrence, K.R., 1992. *Decolonizing Tradition: New Views of Twentieth-Century "British" Literary Canons*. University of Illinois Press, Urbana Champaign, IL.
- Lee, S.J., 1996/2009. *Unraveling the "Model Minority" Stereotype: Listening to Asian American Youth*. Teachers College Press, New York, NY.
- Lempert, W., 2014. Decolonizing encounters of the third kind: alternative futuring in native science fiction film. *Vis. Anthropol. Rev.* 30 (2), 164–176. [10.1111/var.12046](https://doi.org/10.1111/var.12046).
- Lempert, W., 2018. Indigenous media futures: an introduction. *Cult. Anthropol.* 33, 173–179.
- Lin, A., Martin, P.W., 2005. *Decolonisation, Globalisation: Language-in-Education Policy and Practice*. Multilingual Matters, Clevedon.
- Lomawaima, K.T., 1994. *They Called It Prairie Light: The Story of Chilocco Indian School*. University of Nebraska Press, Lincoln, NE.
- Lott, E., 1993. *Backcover blurb*. In: Gilroy (Ed.), P. (Ed.), *The Black Atlantic: Modernity and Double Consciousness*. Harvard University Press, Cambridge, MA.
- Love, B., 2019. *We Want to Do More Than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon, Boston, MA.
- Maparyan, L., 2012. *The Womanist Idea*. Routledge, New York, NY.
- McCart, T.L., 2002. *A Place to Be Navajo: Rough Rock and the Struggle for Self-Determination in Indigenous Schooling*. Lawrence Erlbaum, London.
- McClaurin, I. (Ed.), 2001. *Black Feminist Anthropology: Theory, Politics, Praxis, and Poetics*. Rutgers University Press, New Brunswick, NJ.
- McClennen, S.A., 2004. *The Dialectics of Exile: Nation, Time, Language, and Space in Hispanic Literatures*. Purdue University Press, West Lafayette.
- Melanson, Y., 1999. *Looking for Lost Bird: A Jewish Woman Discovers Her Navaho Roots*. Avon Books, New York, NY.
- Mercer, K. (Ed.), 2008. *Exiles, Diasporas, & Strangers*. Institute of International Visual Arts, Rivington Place and London.
- Mies, M., Shiva, S., 1993. *Ecofeminism*. Fernwood, Halifax.
- Mills, C.W., 1959/2000. *The Sociologic Imagination*. Oxford University Press, New York, NY.
- Minh-Ha, T.T., 1989. *Woman, Native, Other: Writing Postcoloniality and Feminism* (Midland Books). Indiana University Press, Bloomington, IN.
- Mishra, V., 1991. The diasporic imaginary: theorizing the Indian diaspora. *Textual Pract.* 10 (3), 421–447.
- Mishra, V., 1996. The diasporic imaginary: theorizing the Indian diaspora. *Textual Pract.* 10 (3), 421–447.
- Mishra, V., 2007. *The Literature of the Indian Diaspora: Theorizing the Diasporic Imaginary*. Routledge, New York, NY.
- Mohanty, C., 1989. On race and voice: challenges for liberal education in the 1990s. *Cult. Critiq.* 14, 179–208.
- Mohanty, C.T., 2003/2005. *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity*. Duke University Press, Durham, NC.
- Morrison, T., 1970. *The Bluest Eye*. A Lume Book, New York, NY.
- Morrison, T., 1992. *Playing in the Dark: Whiteness and the Literary Imagination*. Vintage, New York, NY.
- Morrison, T., 2008. In: Denard, C.C. (Ed.), *What Moves at the Margin: Selected Nonfiction*. University Press of Mississippi, Jackson, MS.
- Mutua, K., Swadener, B.B., 2004. *Decolonizing Research in Cross-Cultural Contexts: Critical Personal Narratives*. State University of New York Press, Albany, NY.
- Naficy, H., 1993. *The Making of Exile Cultures: Iranian Television in Los Angeles*. University of Minnesota Press, Minnesota, MN.
- Naficy, H. (Ed.), 1999. *Home, Exile, Homeland: Film, Media, and the Politics of Place*. Routledge, New York, NY.
- Naficy, H. (Ed.), 2001. *An Accented Cinema: Exilic and Diasporic Filmmaking*. Princeton University Press, Princeton, NJ.
- Narayan, U., 1997. *Dislocating Cultures: Identities, Traditions, and Third World Feminism*. Routledge, New York, NY.
- Nelson, A. (Ed.), 2002. *Afrofuturism*. Special Issue of *Social Text* 20 (2).
- Ng-A-Fook, N., 2007. *An Indigenous Curriculum of Place*. Peter Lang, New York, NY.
- Nussbaum, M.C., 1997. *Cultivating Humanity: A Classical Defense of Reform in Liberal Education*. Harvard University Press, Cambridge, MA.
- Nussbaum, M.C., 2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton University Press, Princeton, NJ.
- Olson, G.A., Worsham, L. (Eds.), 2007. *The Politics of Possibility: Encountering the Radical Imagination*. Paradigm, Boulder, CO.
- Ortiz, S.J., 2009. *A Poetic Legacy of Indigenous Continuance*. University of New Mexico Press, Albuquerque, NM.

- Pagano, J.A., 1990. *Exile and Communities: Teaching in the Patriarchal Wilderness*. State University of New York Press, Albany, NY.
- Paley, V.G., 1979. *White Teacher*. Harvard University Press, Cambridge, MA.
- Palmer, P.J., 1998. *The Courage to Each: Exploring the Inner Landscape of a Teachers Life*. Jossey-Bass, San Francisco, CA.
- Paris, D., Alim, H.S. (Eds.), 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College, New York, NY.
- Park, C.C., Goodwin, A.L., Lee, S.J. (Eds.), 2001. *Research on the Education of Asian and Pacific Americans*. Information Age, Greenwich, CN.
- Park, C.C., Goodwin, A.L., Lee, S.J. (Eds.), 2003. *Asian American Identities, Families, and Schooling*. Information Age, Greenwich, CN.
- Patel, L., 2016. *Decolonizing Educational Research: From Ownership to Answerability*. Routledge, New York, NY.
- Phillion, J., He, M.F., 2004. Using life based literary narratives in multicultural teacher education. *Multicult. Perspect.: An Official Journal of the National Association for Multicultural Education* 6 (3), 3–9.
- Phillion, J., He, M.F., 2001. Narrative inquiry in educational research. *Journal of Critical Inquiry Into Curriculum & Instruction* 3 (2), 14–20.
- Phillion, J., He, M.F., 2008. Multicultural and cross-cultural narrative inquiry in educational research. *Thresholds Educ.* 34 (1 &2), 2–12.
- Phillion, J., He, M.F., Connelly, F.M. (Eds.), 2005. *Narrative and Experience in Multicultural Education*. Sage, Thousand Oaks, CA.
- Phillips, L., 2006. *The Womanist Reader: The First Quarter Century of Womanist Thought*. Routledge, New York, NY.
- Pichová, H., 2002. *The Art of Memory in Exile*. Southern Illinois University Press, Carbondale, IL.
- Robinson, M. (Ed.), 1994. *Altogether Elsewhere: Writers on Exile*. A Harvest Book/Harcourt Brace & Company, San Diego, CA.
- Rodriguez, R., 1982. *Hunger for Memory: The Education of Richard Rodriguez: an Autobiography*. Bantam Books, New York, NY.
- Ross, S., 2015. The womanist/Black feminist milieu. In: He, M.F., Schultz, B.D., Schubert, W.H. (Eds.), *The Sage Guide to Curriculum in Education*. Sage, Thousand Oaks, CA, pp. 358–366.
- Safran, W., 1991. Diasporas in modern societies: myths of homeland and return. *Diaspora* 1 (1), 83–99.
- Said, E.W., 1979. *Orientalism*. Vintage Books, New York, NY.
- Said, E.W., 1994. *Representations of the Intellectual*. Vintage Books, New York, NY.
- Said, E.W., 2000. Invention, memory, and place. *Crit. Inq.* 26 (2), 175–192.
- Said, E.W., 2002. Reflections on exile and other essays. In: *Convergences: Inventories of the Present*, fourth ed. Harvard University Press, Cambridge, MA.
- Sandlin, J.A., Schultz, B.D., Burdick, J., 2010. *Handbook of Public Pedagogy: Education and Learning Beyond Schooling*. Routledge.
- Sandoval, C., 2000. *Methodology of the Oppressed*. University of Minnesota Press, Minneapolis, MN.
- Santiago, E., 1993. *When I Was Puerto Rican*. Vintage Books, New York, NY.
- Schubert, W.H., 2009a. Love, Justice, and Education: John Dewey and the Utopians. Information Age Publishing, Charlotte, NC.
- Schubert, W.H., 2009b. Whats worthwhile: from knowing and experiencing to being and becoming. *J. Curric. Pedagogy* 6 (1), 21–39.
- Siddle-Walker, E.V., 1996. *Their Highest Potential: An African-American School Community in the Segregated South*. University of North Carolina Press, Chapel Hill, NC.
- Skeldon, R. (Ed.), 1994. *Reluctant Exiles? Migration from Hong Kong and the New Overseas Chinese*. M. E. Sharpe, Armonk, NY.
- Smith, B. (Ed.), 1983/2000. *Home Girls: A Black Feminist Anthology*. Rutgers University Press, New Brunswick, NJ.
- Soja, E.W., 1989. *Postmodern Geographies: The Reassertion of Space in Critical Social Theory*. Verso, London.
- Soja, E.W., 1996/2000. *Thirdspace: Journeys to Los Angeles and Other Real-And-Imagined Places*. Blackwell, Malden, MA.
- Soja, E.W., 2010. *Seeking Special Justice*. University of Minnesota Press, Minneapolis, MN.
- Spring, J., 2016. *Deculturalization and the Struggle for Equality: A Brief History of the Education of Dominated Cultures in the United States*, eighth ed. Routledge, New York, NY.
- Steger, M., 2008. *The Rise of the Global Imaginary: Political Ideologies from the French Revolution to the Global War on Terror*. Oxford University Press, Oxford.
- Stovall, D.O., 2016. *Born Out of Struggle: Critical Race Theory, School Creation, and the Politics of Interruption*. State University of New York Press, Albany, NY.
- Takaki, R., 1993. *A Different Mirror: A History of Multicultural America*. Back Bay Books; Little, Brown and Company, New York, NY.
- Tölölyan, K., 1991. The nation state and its others: in lieu of a preface. *Diaspora* 1 (1), 3–7.
- Tuck, E., 2015. Indigenous land and decolonizing curriculum. In: He, M.F., Schultz, B.D., Schubert, W.H. (Eds.), *The Sage Guide to Curriculum in Education*. Sage, Thousand Oaks, CA, pp. 439–446.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- Tuhiwai Smith, L., 1999/2012. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books, London and New York, NY.
- Tuhiwai Smith, L., Tuck, E., Yang, K.W., 2018. *Indigenous and Decolonizing Studies in Education: Mapping the Long View*. Routledge, New York, NY.
- UNESCO, 2020. *Education in a Post-COVID World*. The United Nations Educational, Scientific and Cultural Organization, Paris.
- UNESCO, 2022a. *Culture in Times of COVID-19 Resilience, Recovery and Revival*. The United Nations Educational, Scientific and Cultural Organization & Department of Culture and Tourism, Paris, France.
- UNESCO, 2022b. *UNESCO Strategy on Education for Health and Well-Being*. The United Nations Educational, Scientific and Cultural Organization, Paris, France.
- Walker, A., 1967/1983. *In Search of Our Mothers Gardens*. Harcourt, Orlando, FL.
- Walker, A., 1997. *Anything We Love Can Be Saved: A Writers Activism*. A Ballantine Book, New York, NY.
- Wang, H.Y., 2004. *The Call from the Stranger on a Journey Home: Curriculum in a Third Space*. Peter Lang, New York, NY.
- Wang, H.Y., 2006. Speaking as an alien: is a curriculum in a third space possible? *J. Curriculum Theor.* 22 (1), 101–116.
- Wang, H.Y., 2010. Intimate revolt and third possibilities: co-creating a creative curriculum. In: Malewski, E. (Ed.), *Curriculum Studies Handbook: The Next Moment*. Routledge, New York, NY, pp. 374–386.
- Wilson, S., 2008. *Research Is Ceremony: Indigenous Research Methodologies*. Fernwood Publishing, Nova Scotia.
- Wing, A.K. (Ed.), 2000. *Global Critical Race Feminism: An international Reader (Critical America)*. New York University Press, New York, NY.
- Wing, A.K. (Ed.), 2003. *Critical Race Feminism: A Reader*, second ed. New York University Press, New York, NY.
- Womack, Y.L., 2013. *Afrofuturism: The World of Black Sci-Fi and Fantasy Culture*. Lawrence Hill Books, Chicago, IL.

Curriculum and the United Nations' sustainable development goals

Ole Andreas Kvamme, Department of Teacher Education and School Research, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	406
Agenda 2030 and the UN's sustainable development goals	406
Promises and insufficient action	407
Curriculum and the nation–state	407
The democratic impasse	407
Curriculum and the SDGs	408
Recontextualizations, hegemony, and democratic iterations	409
Deliberation and the ethical dimension involved	409
The SDGs as a common good?	410
Conclusion	411
References	411

Introduction

The United Nations (UN) initiatives on sustainable development have gained momentum in a world distinguished by climate crisis, ecological crisis, and global injustices. In the following, these initiatives are addressed with particular attention to the impact on curriculum. The scope is theoretical, identifying and discussing some vital issues for curriculum theory, that should be considered. I first give a short presentation of UN's Agenda 2030, including the 17 Sustainable Development Goals, followed by reflections on the key concepts of curriculum and democracy. A major contradiction within the history of the UN sustainable development initiatives is pointed out that consists of promises that persistently have been followed by the lack of sufficient action, also reflected in the so-called democratic impasse of sustainable development. Critical curriculum studies are called for, exploring both hegemonic imaginaries involved and the significance of context. A multi-level approach to curriculum processes is suggested and exemplified by an account that combines Basil Bernstein's conception of recontextualization with Seyla Benhabib's critical cosmopolitanism, conceiving the recontextualizations of universal norms as democratic iterations.

Finally, the conception of the common good is considered as a critical premise within a curriculum theory oriented toward studies of the UN's Agenda 2030 and the Sustainable Development Goals. Decisive is here the suggestion that the common good, in the SDGs addressed on a general level, also involves recontextualizations expressed in particular curriculum practices, on which hegemonic imaginaries continuously have an impact.

Agenda 2030 and the UN's sustainable development goals

On September 25, 2015, the United Nations General Assembly, through the Resolution 70/1 *Transforming our World: the 2030 Agenda for Sustainable Development* (United Nations, 2015), adopted the 17 Sustainable Development Goals (SDGs). The notion of SDGs, also referred to as the "Global Goals" (French and Kotzé, 2018), was first proposed at the 2012 UN Conference on Sustainable Development (the "Rio +20 Summit"), to replace the Millennium Development Goals (MDGs) which were launched in 2001 and expired in 2015. Through intergovernmental negotiations and open-ended dialogs with civil society including various groups, the draft featuring 17 goals and encompassing 169 targets and subsequent indicators was presented. Unlike the MDGs that were particularly oriented toward developing countries, the SDGs are universally applicable, including all nations globally.

The goals are framed by the concept of *sustainable development*, which, within United Nations, has a history going back to the 1980s and was famously introduced as a key concept by the World Commission on Environment and Development (appointed by United Nations General Assembly in 1983) in the report *Our Common Future* in 1987. The report states that sustainable development "meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987). With the concept of sustainable development, the United Nations brought together a long-term engagement strategy for social development and elimination of poverty with the acknowledgment of ecological crisis, including human-induced climate change. In a global call to action the commission states:

The issues we have raised in this report are inevitably of far reaching [sic] importance to the quality of life on earth—indeed to life itself. We have tried to show how human survival and well-being could depend on success in elevating sustainable development to a global ethic. In doing so, we have called for such major efforts as greater willingness and cooperation to combat international poverty, to maintain peace and enhance security world-wide, and to manage the global commons.

World Commission on Environment and Development (1987).

In 1992, *Our Common Future* was followed by the United Nations Conference on Environment and Development in Rio de Janeiro. From here emerged the UN conventions on biodiversity (United Nations, 1992a), climate change (United Nations, 1992b), desertification (United Nations, 1992c), the Rio Declaration (United Nations, 1992d), and Agenda 21 (United Nations, 1992e), a global action plan for the 21st century. With Agenda 2030 and the SDGs, adopted by UN General Assembly more than twenty years later (United Nations, 2015), sustainable development has now been granted an unprecedented status and role as the decisive strategy to guide future strategies and actions. While the 17 Sustainable Development Goals definitely constitute the backbone of Agenda 2030, other elements are included as well that concern financing, partnerships, technology facilitation, monitoring, reporting, and evaluation.

In Rio in 1992, education was already explicitly addressed in Chapter 36 of Agenda 21. Here, it is stated that education is “linked to virtually all areas in Agenda 21” (United Nations, 1992e). Chapter 36 even addresses public awareness and training and calls for the reorienting of education toward sustainable development. UNESCO was appointed as responsible for this educational engagement and monitored the UN Decade of Education for Sustainable Development (ESD) (2005–2014) as a major achievement. The significance and integration of education is restated in Agenda 2030 with quality education addressed as SDG 4, both conceived of as an aim in itself, and a means to promote other aims.

Promises and insufficient action

The UN resolution *Transforming our World: the 2030 Agenda for Sustainable Development* (United Nations, 2015) concludes that significant progress has been made in meeting several of the development challenges continuously addressed by the United Nations. Hundreds of millions of people have emerged from extreme poverty, global access to education has increased, and information and communication technology has been spread. However, the resolution makes equally clear that major threats to life are extended, and in vital areas, the situation has deteriorated distinguishably due to rising social and economic inequalities within and between countries, natural resource depletion and environmental degradation, climate change, putting the survival of societies and the biological support systems of the planet at risk.

More severe and disturbing when it comes to the history of this initiative, is the fact that in the early 1990s, social inequalities, climate change, and ecological crisis were addressed as pivotal, urgent concerns in United Nations conventions and declarations and subsequent calls for action. Almost 40 years later, the calls for action are even more urgent, particularly with regard to the ecological and climate crisis. The contradiction between promises and insufficient action appears to be a main flaw in the history of the policy of sustainable development, jeopardizing the legitimacy of the initiative itself. As far as education is a part of this story, the stated contradiction also has bearings on the global education policy of education for sustainable development.

Curriculum and the nation-state

Curriculum as a key concept in this entry is an elusive concept, partly due to the broadness of the curriculum field that encompasses student experiences on learning, teacher practices, written documents, and even ideological imaginaries with bearings on education. Moreover, within the context of an international encyclopedia, it is not without problems to refer to a concept emerging from Anglo-American traditions with an emphasis on curriculum-making in local contexts, reflecting school systems that differ considerably from structures and practices in other countries. Thoroughly documented and well explored are, for instance, similarities and differences between this curriculum tradition and the North European didactic tradition in which teachers have been licensed to interpret an open national curriculum while fulfilling the societal mandate of *Bildung* (“formation”; “edification”) (Hopmann and Riquarts, 2000; Künzli, 2000; Westbury, 2000).

“Curriculum is the heart of education,” Null (2011, p. 1) concludes, pointing at how curriculum expresses what should be taught, combining thought, action, and purpose. The emphasis on subject matter as a defining aspect of curriculum is summed up in the question coined by Spencer (1861): What knowledge is of most worth? Michael Apple, drawing attention to political and ideological aspects, rephrases the question as: Whose knowledge is of most worth? (Apple, 2004, p. xix). In the renewal of the German didactic tradition, in the 1980s, Wolfgang Klafki placed current, societal key challenges in the crucial content position of the didactic triangle (the other two apexes constituted by the students and the teacher), including concerns for environment, peace, and global justice (account in English, see Klafki, 1998).

Klafki exemplifies how global issues have been introduced to curriculum theory, challenging the impression of the general national framing of curriculum within the Westphalian order (see Kvamme, 2021). Since Klafki's works in the 1980s, these challenges have been elaborated within various forms of cosmopolitanism, including the critical cosmopolitanism of Benhabib (2006, 2011), considered in *Recontextualizations, hegemony, and democratic iterations* section below. The role of the nation-state also emerges as a decisive aspect when we now turn to democracy as another key concept in this entry.

The democratic impasse

Within the studies of governance, the contradiction stated above between universal claims adopted unanimously by the nations of the world and the lack of corresponding action have been discussed with regard to the Western, liberal democracies. William Lafferty

employs Carl Cohen's definition of democracy as "that system of community government in which, by and large, the members of a community participate, directly or indirectly, in the making of decisions which affect them all" (Lafferty, 2012, p. 299). Lafferty demonstrates how the notion of community in Western societies historically has been defined by geographical boundaries. Significant here is the development of nation-states, corresponding with a notion of citizenship based on individual membership from which representations of interest emerge. In contrast, sustainable development involves a notion of community based on ecological interdependence with an expanded notion of citizenship that incorporates the interests of future generations, the world's poor, and possibly the interests of other species (Lafferty, 2012, p. 305).

To Lafferty, the contradiction between the notions of community and citizenship demonstrates a main challenge involved in the United Nations' sustainable development agenda. Virtually all nations of the world have formally subscribed to the goal of sustainable development, and the goal is to be realized through democratic means, but the existing democratic structures privilege the interests of the demarcated jurisdictions of nation-states and do not accommodate the normative notions expressed in the goal. Lafferty's conclusion is expressed as "the impasse of dysfunctional democracy," cf. the subtitle of the paper referred to here (Lafferty, 2012).

The tensions between existing liberal democracies and implications of sustainable development are also addressed by Westall (2015). In addition to the restrictions constituted by defined political geographies and legally defined citizens, she points out how liberal democracies have given economic growth primacy while not taking environmental limits into account, often guided by short-termism that makes individual freedom a dominant ethic. In contrast, sustainable development implies considerations on long-term impacts on all affected people including intragenerational and intergenerational equity, an integration of economic, environmental and social considerations accentuating environmental limits to human activity, and an emphasis on shared goals and shared values (including future orientation and concern for nature).

Both Lafferty and Westall nevertheless accentuate democracy as the preferred model of governance facing the transformation to sustainable development with particular reference to participation and inclusion. This democratic framing is significant in the context of Agenda 2030. The general commitments in the crucial Goal 16 are to "promote just, peaceful, and inclusive societies", but democracy is not explicitly mentioned in the SDGs, according to Langford (2016, p. 173), ensured by China in the drafting process. Here, following Lafferty and Westall, I interpret Goal 16 within the context of a deliberative conception of democracy. As a consequence, Lafferty's formulation of the impasse of democracy should be taken into account.

Curriculum and the SDGs

As stated above, the United Nations' engagement with sustainable development has prompted the global education policy of Education for Sustainable Development (ESD) initiated and followed up on by UNESCO. Currently within Agenda 2030, this engagement is expressed in the *Framework for the implementation of ESD beyond 2019 "ESD for 2030" (2020–2030)* (UNESCO, 2019) and *Education for Sustainable Development A roadmap. ESD for 2030* (UNESCO, 2020). Both of these documents are quite general in their approach. The vision is "to build a more just and sustainable world through strengthening ESD and contributing to the achievement of all 17 Sustainable Development Goals" (UNESCO, 2020). The scope encompasses both formal and informal learning arenas. One of five priority action areas is to advance education policies, "including those that concern learning environments, curricula, teacher education as well as student assessment" (UNESCO, 2019, p. 26). Here, curriculum is quite narrowly conceived and clearly distinguished from education policy. A similar conception is reflected in the global indicator to target 4.7.1 referring to the extent "(i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment" (UNESCO, 2019, p. 51).

The framing accentuates a strategy in which UNESCO promotes curriculum development on a national level as a contribution to build a more just and sustainable world. The indicators are integrated in a monitoring process where nation states report back to UNESCO. As we see, curriculum is here positioned within the context of nation-states. The aim is definitely not to establish a global curriculum but to encourage nation-states to develop curricula that accommodate the global goals expressed in Agenda 2030.

This structure prompts more general considerations on the relationship between global education policy and curriculum. As stated above, UNESCO's strategy of education for sustainable development is an instance of global education policy as an aspect of institutional globalization (Langford, 2016). Rizvi and Lingard (2010) point out how global education policy is framed "by discourses which, while they often have their gestation in various global networks, are negotiated and rearticulated at regional, national, provincial and local levels" (p. 44). Consequently, it is decisive to recognize "the simultaneous plays of the local, national and global as spatial relations in the education policy cycle" (Rizvi and Lingard, 2010, p. 15; see also Kvamme, 2020a). The term "policy cycle", coined by Ball (1994), refers to the interactive set of relationships between various contexts. It rejects a one-way, linear account of relationships between the setting of policy agendas.

From this multilayered approach, the traditional divide between education policy studies and curriculum studies is not sufficient (Rizvi and Lingard, 2010, pp. 93–94). Rizvi and Lingard's suggestion is to include curriculum studies within the field of policy studies. As important as this might be, from the perspective of curriculum studies, one should as well consider the impact of education policy on curriculum.

Of particular relevance for the UNESCO follow-up on Agenda 2030 is the persistent discussion within curriculum studies of the relationship between theory and practice. In a classical take on this matter, Goodson (1994) points out the prescriptive tendencies in curriculum theories, whether the orientation is guided by scientific management and analysis or by the critique of such an

approach guided by meliorism or impossibilism. In either case, the approaches ignore “what is in pursuit of what *might be*” (Goodson, 1994, p. 29). In other words, Goodson calls for a curriculum theory that is sensitive to particular contexts.

With regard to curriculum studies involving the SDGs, a proposal here would first of all be to adopt a multilevel approach including the global, national, and local levels and second, ascertain that the mediations between education policy and curriculum making are accommodated in such a design. A third concern would be to avoid a one-sided emphasis on a top-down perspective and be aware of bottom-up processes as Rizvi and Lingard (2010) express in the concept of policy circle quoted above. Following Goodson (1994), the context of educational practices is particularly significant here.

Recontextualizations, hegemony, and democratic iterations

A multi-level perspective raises the problem of how movements and uptakes of values, imaginaries, ideas between the various levels are to be understood. Here various approaches have been established (see for instance Røvik's (2016) study of knowledge transfer as processes of translation). Addressing the democratic impasse considered above, the concept of recontextualization as formulated by Bernstein (2000) seems to offer a productive entrance. Bernstein draws attention to how discourses are *pedagogized* when recontextualized into educational contexts. A transformation is involved because “there is a space in which ideology can play” (Bernstein, 2000, p. 32). Within Bernstein's theory, recontextualizations concern how social discourse becomes part of legitimate educational knowledge (Loughland and Sriprakash, 2016). UN Agenda 2030 is fraught with normativity and values, and the research literature demonstrates that hegemonic structures seem to be involved in the recontextualization processes.

With regard to the ESD history, a few examples may suffice to demonstrate how the recontextualization processes involves ideology, here conceived of as the maintenance of current power relations (Ricoeur, 2008). Sauvé et al. (2005) note how the ESD agenda has prioritized economic development at the expense of environmental concerns. Furthermore, studies of the UN Decade of Education for Sustainable Development suggest that the ESD agenda has been increasingly adaptable to neoliberalism, reducing political issues to questions of individual preferences and behavior (Huckle and Wals, 2015). Others demonstrate how ESD has eased the socially conflicting nature of the ecological crisis (Gonzales-Gaudiano, 2016).

The possible impact of hegemonic structures on ESD processes has also been problematized. Stefan Bengtsson (2016) refers to what he claims to be a paradox in much of the scholarly literature criticizing ESD. ESD policy supposedly exercises a power of influence while reproducing hegemonic discourse. Still, Bengtsson claims, there seems to be a space for critical reflection and even deviance. With reference to the political theory of Laclau and Mouffe (1985), Bengtsson calls for a logic of contingency to replace the logic of necessity that he holds distinguishes scholars who assert that the ESD agenda serves hegemonic forces. However, as one of these scholars expressed in a rejoinder to Bengtsson, the claim that hegemonic structures are at work does not imply that there is no space for dissensus (Jickling, 2016). In any case, Bengtsson's reminder demonstrates the significance of the productive work exercised on all levels in the education policy cycle.

The moral philosopher and political theorist Seyla Benhabib offers with her critical cosmopolitanism an approach that accommodates both universal claims and the significance of context in processes of recontextualization. With the concept of democratic iterations, Benhabib (2006, 2011) notes how human rights are subject to reinterpretations when articulated in new contexts and make space for plurality, disagreements, and conflict. In one regard, Benhabib claims that such iterations will take place in any recontextualizing process, confirming the significance of context in empirical studies. However, democratic iterations has a normative dimension, as well, drawing attention to processes that open up or restrict the space for democratic iterations, a dimension expressed in the concept of jurisgenerativity, employed by Benhabib (2011) with reference to Robert Cover.

From her position within critical theory (Benhabib, 1986), Benhabib has developed a political theory of deliberate democracy (Benhabib, 1996) that challenges the boundaries of the nation-state. Central is her conception of enlarged thought (Benhabib, 1992), drawn from Arendt (1978), that includes everyone who may possibly be influenced by the action in question in the act of deliberation.

Benhabib's conceptions are consistent with the perspectives within critical pedagogy as accounted for above. In my own empirical studies, I have employed Benhabib's approach in a multi-level exploration of recontextualizations of environmental ethical values expressed within the UN agenda in Norwegian educational contexts and even in the school strikes for the climate (Kvamme, 2018, 2019, 2020a,b; see also Wahlström, 2009 for the employment of Benhabib within curriculum studies). I suggest that hegemonic processes are at work, including anthropocentrism, neoliberalism, and the boundaries of the nation-state, but also identify how the environmental ethical values formulated by the United Nations become resources for critique, empowerment, and transformative action. These studies demonstrate the decisive normative dimension of the United Nations' sustainable development agenda. This aspect is what I will address in the following section focused on relations between curriculum, deliberation, and ethics.

Deliberation and the ethical dimension involved

Deliberative perspectives in modern curriculum theory are strongly connected to the American curriculum scholar Joseph Schwab and his conceptions of the practical (Schwab, 1969). Decisive is the priority given to deliberative practices concerning practical problems, initiating further discussion by other scholars. William A. Reid's considerations on the relationships between deliberation and ethics are particularly interesting with regard to issues raised in this entry. Reid (1994) proposes that curriculum problems are

uncertain, practical, and moral problems. The practical aspect is mainly due to the fact that curriculum problems demand solutions in the form of action, whether as written curriculums or in the actual teaching in the classroom. The ethical aspect is made visible in the formulation of the major question posed as “What *ought* we to do?” leading to further questions of responsibility. The ethical responsibility involves all who are affected by the action being chosen. The uncertainty is pointing at the particularity of curriculum work. It always occurs within a unique context. How are curriculum problems to be dealt with, then? In his proposal, Reid refers to deliberation, specifically “the notion of inquiry” through which practical problems have to be resolved involving all who are affected, influencing and determining moral, political, and social actions (Reid, 1994, p. 18).

Reid's perspective is specific and practical, reflecting school systems in which curriculum work traditionally has been carried out locally. Although definitely not multi-layered in orientation, his perspectives nevertheless contribute to the discussion carried out here. First of all, Reid notes the significance of context, as emphasized above with reference to Benhabib. Second, Reid premises that curriculum practices as deliberations involve all who are affected, and finally, they emerge as an ethical responsibility.

Reid's contributions exemplify how curriculum theory until recently has been almost solely framed within geographical boundaries determining notions of community and participation. The message from cosmopolitan theory is that the local context is never just local (Beck, 2006; Mannion, 2015; Benhabib, 2006, 2011; Papastephanou, 2012). In numerous ways, the global has impact on the local and expands those affected by the chosen action. The intimate relationship between community and curriculum exposed in Reid's account is radically enhanced when the SDGs and Agenda 2030 are taken into account. Here, the question posed by Reid of what ought we to do emerges as a political issue and a policy task that, following the policy definition of Rizvi and Lingard (2010, p. 7), is concerned with “the authoritative allocation of values.” The ideal of deliberation is retrieved in the drafting process that produced UN Agenda 2030 and the 17 SDGs. Fox and Stoett (2016) contrast this process with the elite donor-driven process that led to the Millennium Development Goals and conclude that the 2030 Agenda process facilitated several arenas for citizen participation, and opened new paths toward the establishment of global democratic governance.

The participatory process is a main reason for ending up with as many as 17 goals (Langford, 2016) that constitute a thematic repertoire ranging from poverty, health, education, and inequality to water, energy, work, infrastructure (including climate, life on earth and underwater), peace, security, and good governance.

However, Fox and Stoett (2016) conclude that, despite their positive assessment, the drafting process remained far from an ideal participatory democracy. Others have been even more critical, demonstrating how hegemonic forces of anthropocentrism and neoliberalism are expressed in the document adopted by the United Nations General Assembly (Adelman, 2018; Kotzé, 2018).

This discussion confirms the normative and ethical dimension involved, echoing Reid's insistence on an ethical responsibility in curriculum practices conceived of as deliberation. Without the ethical dimension, the call for sustainable development is not justified (Kemp, 2011; Lafferty, 2012). Decisive is the concern for intragenerational and intergenerational justice for all and the protection of life on earth in all its diversity. Although the SDGs should probably not be considered a global curriculum, it is conspicuous how Agenda 2030, at its most general level, is a call to transform people, practices, and structures everywhere. This ambition transcends educational institutions and extends to informal and non-formal learning arenas alike and expresses an imaginary that is obviously pertinent when the impacts of the SDGs on national curricula are to be considered.

The SDGs as a common good?

In the previous section, I suggested that the drafting process of Agenda 2030 and the SDGs may be productively considered from the perspective of curriculum theory. The concern has not been to claim that this UN endeavor reflects a curriculum process in itself, but to demonstrate Agenda 2030's significance as more than a passive background for UNESCO's follow-up in the global education policy of Education for Sustainable Development. In this section I elaborate on this while exploring a position that connects the SDGs with *the common good*, as this notion emerges within the UNESCO context.

In the UNESCO report *Rethinking Education: Toward a global common good?* (UNESCO, 2015), a key contribution is the consideration of education and knowledge as a global common good. With reference to Deneulin and Townsend (2007), the report states that the common good is “constituted by goods that humans share intrinsically in common and that they communicate to each other, such as values, civic virtues and a sense of justice” (UNESCO, 2015, p. 77). The notion of the common good is different from the notion of the public good in which human well-being is framed by individualistic socio-economic theory. From the common good perspective, it is not just the good life of individuals that matters but also the goodness of the life that humans hold in common (UNESCO, 2015, p. 78). Moreover, the weight on community is expressed in the emphasis on participatory processes that stand out as a common good in themselves.

In the UNESCO report the notion of education as a common good is contrasted with tendencies in current education policy to conceive learning as an individual process of skill acquisition with little attention paid to questions of the purpose of education and the organization of learning opportunities as a collective social endeavor. Furthermore, learning has been essentially centered on the results of educational processes and neglected the processes of learning (UNESCO, 2015, pp. 80–81).

While this perspective has bearings on how SDG 4 on quality education is considered, the emphasis on both education and knowledge may be seen as an invitation to mediate between this particular goal and the sustainable development agenda as a whole. The UNESCO report actually suggests such a mediation itself:

Given the central concern for sustainable development in an increasingly interdependent world, education and knowledge should thus be considered global common goods. This means that the creation of knowledge, its control, acquisition, validation, and use, are common to all people as a *collective social endeavor*. The governance of education can no longer be separated from the governance of knowledge.

UNESCO (2015, p. 80).

A major aspect of Agenda 2030 and the SDGs is definitely the governance of knowledge, consequently opening up for connecting the SDGs to the common good. While a conception of the SDGs as a common good is an intriguing, but complicated matter that cannot be sufficiently unpacked in this entry (see also Lotz-Sisitka, 2017), the relationship between the SDGs and the notion of the common good should be considered by revisiting the democratic impasse accounted for in *The democratic impasse* section above, particularly addressing community and membership.

As Lafferty (2012, p. 306) pointed out, nation-states as communities have historically been seen as delimited spatial units organized for self-governance, while sustainable development involves a notion of community based on ecological interdependence incorporating the interests of future generations, the world's poor, and the interests of other species. Connecting the SDGs to the common good accentuates this expanded notion of community, and demonstrates how goods that are intrinsically shared in common across borders, are under threat. As long as interests of the nation-states are determining political priorities, the impasse will remain.

With regard to membership, sustainable development implies that citizenship cannot solely be conceived of within the nation states, but is global in scope. In *Resolution 70/1 Transforming our world*, a vision of solidarity is distinctly stated in the preamble, "As we embark on this collective journey, we pledge that no one will be left behind" (United Nations, 2015), and SDG 10 *Reduced inequalities* particularly address the social, economic and political inclusion of all (United Nations, 2015). When the SDGs in such ways are connected with the notion of the common good, the contextual and processual character of this notion should also be considered. The UNESCO report *Rethinking Education: Toward a global common good?* proposes: "What is meant by the common good can only be defined with regard to the diversity of contexts and conceptions of well-being and common life" (UNESCO, 2015, p. 78). This proposal opens for considering the recontextualizations of the SDGs in various regional, national, and local contexts as aspects of the common good. If so, the perspective is definitely also pertinent to curriculum practices as accentuated by Schwab and Reid.

An emphasis on community and participation offers a critical perspective for empirical curriculum studies on the SDGs and on ESD initiatives. A decisive ethical and political concern is whose perspectives "are privileged in the curriculum as well as in education institutions more generally" (Apple and Buras, 2006, p. 3). Such a critical perspective may be conceived of as an immanent critique "of social practices using principles that are reconstructed from those very practices, rather than using principles that have their validity established by moral arguments that do not refer to the practices to which the principles are applied" (Stahl, 2019). Identifications of contradictions and impasses involved, as addressed in this entry, are instances of such an immanent critique, considering values and concerns expressed in the UN sustainable development agenda itself.

Most probably, the democratic impasse as posed by Lafferty (2012), still distinguishes processes within Agenda 2030 and the SDGs in the decade we have entered, and the reproductions of hegemonic structures and practices still take place in the recontextualization processes, and should be subject to critique within curriculum studies. However, as Benhabib has emphasized, claims made within structures like the United Nations that express an institutional cosmopolitanism may have mobilizing and transforming functions (Benhabib, 2011). A prime example of such a bottom-up initiative is the school strikes for the climate (Kvamme, 2019). Thus, Bengtsson's (2016) appeal to a logic of contingency in curriculum studies on education for sustainable development should be considered.

Conclusion

This entry has addressed aspects of curriculum, democracy, and the UN's sustainable development goals. A concern has been to demonstrate how the various perspectives in many respects are intertwined, transcending the borders of the nation-states, and calling for a multi-layered approach. Still, a deliberative position on curriculum practices and democracy has been adopted with an emphasis on the significance of context. In the final part of the entry, a major suggestion has been to connect the SDGs—including both the drafting processes, the UN Resolution 70/1 (United Nations, 2015), and subsequent recontextualizations—with the notion of the common good. Such a connection may premise critical studies of the curriculum processes that are promoted in the global education policy of ESD. Finally, let us return to Reid and a pertinent reminder that demonstrates how the global common good explored in this entry resonates with a curriculum theory that emerges from the local and the practical. "We are never going to realize perfect curricula, but schools, teaching, assessments, and mandates will only serve the good of society insofar as they are guided by curricular ideals" (Reid, 2006, p. 2).

References

- Adelman, S., 2018. The sustainable development goals, anthropocentrism and neoliberalism. In: French, D., Kotzé, L.J. (Eds.), *Sustainable Development Goals: Law, Theory and Implementation*. Edward Elgar Publishing, Cheltenham, pp. 15–40.
- Apple, M.W., Buras, K.L., 2006. Introduction. In: Apple, M.W., Buras, K.L. (Eds.), *The Subaltern Speak: Curriculum, Power, and Educational Struggles*. Routledge, New York, pp. 1–39.

- Apple, M., 2004. *Ideology and Curriculum*, third ed. Routledge Falmer, New York.
- Arendt, H., 1978. *The Life of the Mind*. Harcourt Inc, San Diego.
- Ball, S., 1994. *Education Reform: A Critical and Poststructuralist Approach*. Open University Press, Buckingham, UK.
- Beck, U., 2006. *Cosmopolitan Vision*. Polity, Cambridge.
- Bengtsson, S.L., 2016. Hegemony and the politics of policy making for education for sustainable development: a case study of Vietnam. *J. Environ. Educ.* 47 (2), 77–90. <https://doi.org/10.1080/00958964.2015.1021291>.
- Benhabib, S., 1986. *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory*. Columbia University Press, New York.
- Benhabib, S., 1992. *Situating the Self: Gender, Community and Postmodernism in Contemporary Ethics*. Polity Press, Cambridge.
- Benhabib, S., 1996. Toward a deliberative model of democratic legitimacy. In: Benhabib, S. (Ed.), *Democracy and Difference. Contesting the Boundaries of the Political*. Princeton University Press, Princeton, pp. 67–94.
- Benhabib, S., 2006. *Another Cosmopolitanism*. Oxford University Press, Oxford.
- Benhabib, S., 2011. *Dignity in Adversity: Human Rights in Troubled Times*. Polity Press, Oxford.
- Bernstein, B., 2000. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Rowman & Littlefield Publishers, Inc, Lanham, NJ.
- Deneulin, S., Townsend, N., 2007. Public goods, global public goods and the common good. *Int. J. Soc. Econ.* 34 (1–2), 19–36.
- Fox, O., Stoett, P., 2016. Citizen participation in the UN sustainable development goals consultation process: toward global democratic governance? *Glob. Gov.* 22 (4), 555–573.
- French, D., Kotzé, L.J., 2018. Introduction. In: French, D., Kotzé, L.J. (Eds.), *Sustainable Development Goals: Law, Theory and Implementation*. Edward Elgar Publishing, Cheltenham, pp. 1–12.
- González-Gaudiano, E., 2016. ESD: power, politics, and policy: “Tragic optimism” from Latin America. *J. Environ. Educ.* 47 (2), 118–127. <https://doi.org/10.1080/00958964.2015.1072704>.
- Goodson, I.F., 1994. *Studying Curriculum*. Open University Press, Buckingham.
- Hopmann, S., Riquarts, K., 2000. Starting a dialogue: a beginning conversation between *Didaktik* and the curriculum traditions. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice. The German Didaktik Tradition*. Routledge, New York, pp. 3–11.
- Huckle, J., Wals, A.E.J., 2015. The UN decade of education for sustainable development: business as usual in the end. *Environ. Educ. Res.* 21 (3), 491–505.
- Jickling, B., 2016. Losing traction and the art of slip-sliding away: or, getting over education for sustainable development. *J. Environ. Educ.* 47 (2), 128–138.
- Kemp, P., 2011. *Citizen of the World: The Cosmopolitan Ideal for the 21st Century*. Humanity Books, Amherst, New York.
- Klafki, W., 1998. Characteristics of critical-constructive Didaktik. In: Gundem, B., Hopmann, S. (Eds.), *Didaktik and/or Curriculum. An International Dialogue*. Peter Lang, New York, pp. 307–328.
- Kotzé, L.J., 2018. The sustainable development goals: an existential critique alongside three new-millennial analytical paradigms. In: French, D., Kotzé, L.J. (Eds.), *Sustainable Development Goals: Law, Theory and Implementation*. Edward Elgar Publishing, Cheltenham, pp. 41–65.
- Künzli, R., 2000. German *Didaktik*: models of re-presentation, of intercourse, and of experience. In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice: The German Didaktik Tradition*. Routledge, New York, pp. 41–54.
- Kvamme, O., 2018. Blurring the image of the other? The recontextualization of environmental ethical values in Norwegian education policy documents. In: Ristiniemi, J., Skeie, G., Sporre, K. (Eds.), *Challenging Life: Existential Questions as a Resource for Education*. Waxmann Verlag, Münster, pp. 359–381.
- Kvamme, O., 2019. School strikes, environmental ethical values, and democracy. *Studier i Pædagogisk Filosofi* 8 (1), 6–27. <https://doi.org/10.7146/spf.v8i1.117967>.
- Kvamme, O., 2020a. Situating moral education in a globalized world: sustainability values and student experiences. In: Strand, T. (Ed.), *Rethinking Ethical-Political Education*. Springer, Dordrecht, pp. 45–65.
- Kvamme, O., 2020b. *Recontextualizing Environmental Ethical Values in a Globalized World: Studies in Moral Education*. Thesis submitted for the degree of Philosophiae Doctor. Faculty of Educational Sciences, University of Oslo, Oslo. <http://hdl.handle.net/10852/75162>.
- Kvamme, O., 2021. Rethinking Bildung in the Anthropocene: the Case of Wolfgang Klafki. *HTS Theol. Stud.* 77 (3). <https://doi.org/10.4102/hts.v77i3.6807>.
- Laclau, E., Mouffe, C., 1985. *Hegemony and socialist strategy. Towards a radical political politics*. Verso, London.
- Lafferty, W.M., 2012. Governance for sustainable development: the impasse of dysfunctional democracy. In: Meadowcroft, J., Langhelle, O., Ruud, A. (Eds.), *Governance, Democracy and Sustainable Development: Moving Beyond the Impasse*. Edward Elgar, Cheltenham, pp. 297–337.
- Langford, M., 2016. Lost in translation? The politics of the sustainable development goals. *Ethics Int. Aff.* 30 (2), 167–176.
- Lotz-Sisitka, H., 2017. Education and the common good. In: Jickling, B., Sterling, S. (Eds.), *Post-sustainability: Remaking Education*. Palgrave MacMillan, Basingstoke, pp. 63–78.
- Loughland, T., Sriprakash, A., 2016. Bernstein revisited: the recontextualization of equity in contemporary Australian School Education. *Br. J. Sociol. Educ.* 37 (2), 230–247. <https://doi.org/10.1080/01425692.2014.916604>.
- Mannion, G., 2015. Towards global pedagogies: some risks associated with education for global citizenship and how global pedagogies might avoid them. In: Friedman, J., Haverkate, V., Oomen, B., Park, E., Sklad, M. (Eds.), *Going Global in Higher Education: The Theory, Teaching and Measurement of Global Citizenship*. University College Roosevelt, Middleburg, pp. 19–34.
- Null, W., 2011. *Curriculum From Theory to Practice*. Rowman & Littlefield Publishers, Inc, Lanham.
- Papastephanou, M., 2012. *Thinking Differently About Cosmopolitanism: Theory, Eccentricity, and the Globalized World*. Paradigm Publishers, Boulder, CO.
- Reid, W.A., 1994. *Curriculum Planning as Deliberation*. Rapport Nr. 11, Pedagogisk Forskningsinstitutt. Universitetet i Oslo, Oslo.
- Reid, W.A., 2006. *The Pursuit of Curriculum: Schooling and the Public Interest*. Information Age Publishing, Greenwich.
- Ricoeur, P., 2008. Ideology and utopia. In: Ricoeur, P. (Ed.), *From Text to Action. Essays in Hermeneutics II*. Continuum, London, pp. 300–316.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, Abingdon.
- Røvik, K.A., 2016. Knowledge transfer as translation: review and elements of an instrumental theory. *Int. J. Manag. Rev.* 18 (3), 290–310.
- Sauvé, L., Brunelle, R., Beryman, T., 2005. Influence of the globalized and globalizing sustainable development framework on national policies related to environmental education. *Pol. Futures Educ. Internet* 3 (3), 271–283.
- Schwab, J.J., 1969. The practical: a language for curriculum. *Sch. Rev.* 78 (5), 1–24.
- Spencer, H., 1861. *Education: Intellectual, Moral, and Physical*. Williams and Norgate, London.
- Stahl, T., 2019. Immanent critique. In: Allen, A., Mendieta, E. (Eds.), *The Cambridge Habermas Lexicon*. Cambridge University Press. <https://doi.org/10.1017/9781316771303.051>.
- UNESCO, 2015. *Rethinking Education: Towards a Global Common Good?* <https://unesdoc.unesco.org/e-forum/RethinkingEducation.pdf>.
- UNESCO, 2019. *Education for Sustainable Development: Towards Achieving the SDGs (ESD for 2030). A Draft Framework for the Implementation of Education for Sustainable Development Beyond 2019*. <https://unesdoc.unesco.org/ark:/48223/pf0000370215.locale=en>.
- UNESCO, 2020. *Education for Sustainable Development: A Roadmap. ESD for 2030*. <https://unesdoc.unesco.org/ark:/48223/pf0000374802>.
- United Nations, 1992a. *The Convention on Biological Diversity*. <https://www.cbd.int/convention/text/default.shtml>.
- United Nations, 1992b. *United Nations Framework Convention on Climate Change*. <https://unfccc.int/resource/docs/convkp/conveng.pdf>.
- United Nations, 1992c. *United Nations Convention to Combat Desertification in Those Countries Experiencing Serious Drought and/or Desertification, Particularly in Africa*. https://www.unccd.int/sites/default/files/relevant-links/2017-01/UNCCD_Convention_ENG_0.pdf.
- United Nations, 1992d. *The Rio Declaration*. http://www.unesco.org/education/information/nfsunesco/pdf/RIO_E.PDF.
- United Nations, 1992e. *Agenda 21*. <https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf>.

- United Nations, 2015. Resolution 70/1. Transforming Our World: The 2030 Agenda for Sustainable Development. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf.
- Wahlström, N., 2009. Understanding the universal right to education as jurisgenerative politics and democratic iterations. *Eur. Educ. Res. J.* 8 (4), 520–533.
- Westall, A., 2015. The Relationship between Democracy and Sustainable Development. Briefing May 2015. Foundation for Democracy and Sustainable Development, London.
- Westbury, I., 2000. Teaching as a reflective practice: what might *Didaktik* teach curriculum? In: Westbury, I., Hopmann, S., Riquarts, K. (Eds.), *Teaching as a Reflective Practice: The German Didaktik Tradition*. Routledge, New York, pp. 15–39.
- World Commission on Environment and Development, 1987. Our Common Future. <https://sustainabledevelopment.un.org/content/documents/5987our-commonfuture.pdf>.

Curriculum and digital citizenship

Nancy Law^a and Wing On Lee^b, ^a University of Hong Kong, Faculty of Education, University of Hong Kong, Hong Kong; and
^b Institute for Adult Learning, Singapore University of Social Sciences, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	414
Digital citizenship in the school curriculum	415
Digital citizenship as digital competence	415
Digital citizenship development requires social emotional learning	415
Evolving conceptualizations of citizenship in context	416
Digital citizenship for civic, political and economic participation	417
Digital participation as citizenship rights	417
The rise of DIY citizens and borderless netizens	417
Internet activism and place-based social action	417
Tensions between empowering/empowered digital citizens and the (digital) citizenship curriculum	418
The future of digital citizenship curricula	419
Sociotechnical regime challenges in the context of educational transitions	419
Need for continuity between citizenship and digital citizenship education	420
Summary	421
Acknowledgment	421
References	421

Introduction

Rapid advances in digital technology and their pervasive adoption in all aspects of personal, social and workplace contexts have brought tremendous changes to the lives of citizens in all walks of life, and thus also the concept of citizenship. Digital citizenship as an area of study emerged around the turn of the millennium (Chen et al., 2021). Since then, there have been a lot of interest and development in this area of study from a diversity of disciplinary areas, including education, media studies, sociology, politics, and sociolinguistics. Public attention to digital citizenship as an integral part of the school curriculum started around 2007, coinciding with the publication of the first edition of *Digital Citizenship in Schools* by Ribble and Bailey (2007).

Conceptualizations of digital citizenship as an extension of the pre-Internet age construct of citizenship has been co-evolving in tandem with the rapid technological advances as well as the pervasive changes in personal, social, economic and professional practices in diverse contexts, be it local, national or global. This type of coupling and interaction, referred to as sociotechnical co-evolution (Geels, 2005), between technological innovations (such as in transport and energy industries) and a wide range of consequential changes in associated physical, technological, social, legal and cultural arenas have been observed in all cases where the technological innovation has brought about sustained social transitions.

Digital citizenship as part of the school curriculum has also been changing over time, though the scope of the change has been much narrower and slower compared to changes in the broader academic literature in this area. This article begins with an overview of digital citizenship and its variants in the school curriculum in different countries over time, which shows that much of the focus has been on digital competence, even though the influence of the Internet and social media on the social participation aspect of citizenship is starting to gain attention. Recent developments in digital citizenship curriculum conceptualizations have emerged from efforts of international organizations such as UNESCO, UNICEF, and OECD in connection with the concept of global citizenship (Chakroun and Keevy, 2018; Nascimbeni, 2019; OECD, 2019).

To gain further insight into the future development in the conceptualization of digital citizenship as well as policy and practice in digital citizenship curriculum, we further review the genesis in the conceptualization of citizenship over time since Aristotle. This is followed by a sketch of the more complex and nuanced conceptualizations of digital citizenship in the wider academic literature, as a consequence of the significant changes in how the Internet and social media have changed the landscape of sociopolitical action since the start of this millennium. The next section provides a brief overview of the citizenship curriculum within the broader educational and social context, highlighting its role in conserving the status quo and promoting the prevalent values and practices of the society. The final concluding section discusses the possible scenarios in the further development of the digital citizenship development based on a sociotechnical transformation perspective, highlighting the tension between the empowerment potential of digital technology for digital citizens and the conservative role of citizenship education in the curriculum.

Digital citizenship in the school curriculum

Citizenship has long been a part of the formal school curriculum in many countries, under different headings such as civics and citizenship education (e.g., Australia, [Moyle, 2014](#)), and often embedded within some school subjects such as social studies. Citizenship education generally focus on the rights and responsibilities of citizens within the specific socioeconomic, legal, and political context of the nation state within which the curriculum is situated. Hence, citizenship education has generally played the role of upholding (i.e., conserving) the status quo of the predominant culture and values of the nation state ([Subedi and Rhee, 2018](#)).

Digital citizenship has emerged as a distinguishable label in the formal school curriculum only in a few countries, e.g., Washington State in the US ([Tamayo, 2016](#)), Australia (e.g., the [NSW Department of Education, n.d.](#)), England ([South West Grid for Learning, n.d.](#)), New Zealand ([Netsafe, 2018](#)), and Singapore (referred to as cyber citizenship, [Singapore MOE, 2014](#)). An inspection of these curricula documents reveal that they share similar foci: ICT competence and related literacies, including the ability to address e-safety issues and for civic participation.

Digital citizenship as digital competence

The most influential definition for digital citizenship (DC in short) in the field of education is the one put forward by Ribble and colleagues through the three editions of a curriculum framework for digital citizenship ([Ribble and Bailey, 2007](#); [Ribble, 2011, 2015](#)) published by the International Society for Technology in Education (ISTE). This framework defines digital citizenship as “norms of appropriate, responsible behavior with regard to technology use” ([Ribble and Bailey, 2007](#), p. 10), and comprises nine dimensions grouped under three domains of a student’s everyday life: *digital access, digital learning, and digital communication* for learning and academic performance; *digital rights and responsibilities, digital etiquette, and digital security* as norms of behavior; and *digital health and wellness, digital law, and digital commerce* as necessary knowledge and skills for life beyond the school environment.

Ribble and colleagues’ competence-oriented DC framework has strongly influenced school curricula worldwide, as reflected by its citations as well as the commonality in content in curricula documents such as those discussed in the previous section. When DC is considered largely as digital competence, it is often considered as an integral part of a technology curriculum. For example, ISTE, a US-based professional society known for its influential technology standards for students, teachers and school leaders, has incorporated DC as one of the standards within the National Educational Technology Standards for Students (NETS.S, [ISTE, 2007, 2016](#)).

This competence-oriented conceptualization of DC is prevalent in more recent curricular frameworks, particularly those developed in the US, such as those by the DQ Institute¹ and Common Sense Media.² These newer frameworks tend to put more attention to the competence required by children and youth as their lives have become profoundly networked through the pervasive use of digital and social media. For example, Common Sense Media’s DC framework ([James et al., 2019](#)) has six topics: media balance and well-being; privacy and security; digital footprint and identity; relationships and communication; cyberbullying, digital drama, and hate speech; and news and media literacy.

While digital competence is widely acknowledged as important for the well-being of citizens, referencing such competence as digital citizenship is rare outside of US and Australia. The Council of Europe published a *Digital Citizenship Education Handbook* ([Richardson and Milovidov, 2019](#)), which is distinct from the Digital Competence Framework developed by the European Commission Joint Research Center ([Ferrari, 2013](#); [Vuorikari et al., 2016](#); [Carretero et al., 2017](#)). This will be further explored in a later section of this article.

Digital citizenship development requires social emotional learning

Similar to adults, children and youth are spending ever more time on digital devices and on the Internet. [James et al. \(2019\)](#) quotes statistics that show in the US, kids as young as age 5–8 spend on average almost 3 h per day on screen media, a third of which on mobile devices; and the time spent increases with age, with 45% of teens reporting being almost constantly online. Such extended exposure to the Internet and social media heightens the dangers due to cyber risks such as cyberbullying, identity thefts and data security breaches. While digital competence is needed to recognize and handle such risks, it is commonly recognized that social emotional learning is needed to cope with such challenges ([Datu et al., 2021](#)). There is a need for students to learn how to understand and manage emotions, to empathize and show respect for others, to establish positive relationships and make responsible decisions.

Social emotional learning is needed not simply to ensure e-safety, but to be able to live and contribute positively to a globalizing world. Socio-emotional competence is identified as one of the three core outcomes in the UNESCO Global Citizenship Framework ([Shin et al., 2019](#)) specified as “[having] a sense of belonging to a common humanity, sharing values and responsibilities, empathy, solidarity and respect for differences and diversity” (p. 75).

¹Official Website: <https://www.dqinstitute.org>.

²Official Website: <https://www.commonsense.org/education/digital-citizenship>.

Evolving conceptualizations of citizenship in context

The two most prominent perspectives on digital citizenship as encompassed in school curricula worldwide are introduced above, focusing on digital competence and e-safety. Both are brought about by the necessity of living in a world permeated by the use of digital technology and are not directly connected with the concepts and practices in citizenship education in the traditional curriculum. Given that digital technologies have brought deep changes in the landscape of sociopolitical engagement and participation worldwide, it is inevitable that digital citizenship in this broader sense need to have a place in the school curriculum, whichever subject discipline or curriculum aspect it may be relegated to. To provide a solid foundation for understanding changes in digital citizenship in the broader academic literature, we provide here a brief introduction to evolving conceptualizations of citizenship in historical contexts.

Citizenship mainly refers to the obligational relationship between an individual and the larger community that the individual belongs to. From a political perspective, this larger community refers to the nation-state, and the individual is therefore citizen of the nation-state (Zajda et al., 2009). This concept of citizenship can trace its origin to the city of Athens at the time of Aristotle, when only one-third of the population were qualified to be citizens, and the remaining two-thirds were slaves. Thus, Aristotle was able to propose the noble concept that all citizens have the responsibility to bear office, i.e., to contribute to governing the city (Akinboye, 2015). With a small proportion of citizens in a small city such as Athens, Aristotle's utopian citizenship was a democratic one, as all political decisions could technically be made by counting the votes cast by all citizens. This democratic model, being coined as "direct democracy," has continued in some states or cities in Europe, such as Switzerland, with major political decisions decided by referendum (Wagschal, 1997). It has to be noted that the citizenship concept further developed along with the elimination of slavery such that every individual is empowered to become a citizen of the state. Thus, the concept of citizenship grows alongside liberty, freedom, entitlement to social welfare, and most importantly equality and the right and duty of participation, especially active participation in civil affairs (Lee and Fouts, 2005).

The concept of citizen-state relationship was consolidated by the evolvement from the landlord system to the nation-state, and at that time, the concepts of nation and state were basically equivalent to each other, as for most of the cases at that time, one nation-one state. Thus, the dominant concept of citizenship was represented by national citizenship (Axtmann, 2004; Mandelbaum, 2013). It was only quite recently that the concept of supra-national citizenship emerged with the formation of the European Union. Many individuals want to identify themselves as European citizens, in addition to being citizens of a European country. This is especially apparent after the breakdown of the Communist bloc in Eastern Europe. Many Eastern European countries went through a democratization process and joined the European Union. Thus, individuals from these Eastern European countries particularly welcomed the identity of being a European citizen (Baubock, 2007).

The notion of national citizenship is challenged when the concept of citizenship is applied beyond the state to a larger entity such as Europe, giving rise to the concept of post-national citizenship (Tambini, 2001). Coupled with increasing international events and interactions in diverse, multiple domains, including economic, financial, cultural, and sports, this led to the emergence of global citizenship. Advanced transportation technologies have also contributed to the emerging prominence of global citizenship by enabling people to travel, which is one of the factors contributing to phenomenal mass migration, and the rise of cosmopolitan cities characterized by constant inflow and outflow of its population (Osler and Starkey, 2003). Some inhabitants of cosmopolitan cities are short-term residents, who may be laborers shouldering jobs not taken by the locals, or professionals residing as expatriates. They may not be citizens of the nation state, but they make significant contributions to the city and/or state where they reside. The concepts of resident and citizen thus become blurred if active participation is an important criterion in defining citizenship (Sadiq, 2008). The increasing uptake of the concept of global citizenship is also due to the increasing awareness that our behavior will not only affect those close to where we stay, but also the wellbeing of others in far corners of the world, and that issues of environmental protection and sustainable development can only be resolved through concerted global action (UNESCO, 2014).

Alongside with global citizenship emerged the concept of global competence, the attributes of which include openness to diversity and differences, ability to interact and communicate with people from different cultures and holding different ideas and values, with appreciation and respect. Global competence is considered important for generating new knowledge and novel solutions to unknown situations, and became a theme for comparative assessment of student competence in PISA 2018 administered by OECD (OECD, 2018).

Broadening of the concept of citizenship from nation-state, i.e., national, to post-national, supranational, and global levels indicates that citizenship rights and responsibilities are no longer confined to one's political affiliation, but to the communities that one relates to. Even at the national level, according to Anderson (1983), state is a rather abstract concept to most civilians, as unlike the Athenian times, it is impossible for people in large states with large populations to relate to each other on a daily basis. The concept of imagined communities is important for the extension of citizenship affiliation to broader communities beyond the nation-state toward post-national, supranational, and global communities.

Digitalization has become a fast growing and powerful platform for people to connect through multiple forms of interactions—at personal level, social and organizational levels. It has also infiltrated into the fabric of our everyday life in almost every sector: educational, economic, sociopolitical, cultural, etc. The digital world has become another space for global actions and interactions, giving rise to the concept of digital citizenship, which shares some of the characteristics of global citizenship. Indeed, the number of "inhabitants" on some of the social media platforms such as Facebook (the total monthly and daily active users are 2.6 billion and 1.73 billion respectively according to OBERLO (2020)) are even bigger than the size of the most populous country. For

many people, the digital communities they belong to are crucial to defining who they are, what they do, and their personal identity. Thus, it is meaningful to consider digital citizenship as an extension of citizenship in its full social, political, and cultural richness.

Digital citizenship for civic, political and economic participation

Grounded on the traditional association of citizenship with membership of geopolitical entities such as nation states, Marshall (1992) defined citizenship as “the right to share to the full in the social heritage and to live the life of a civilized being according to the standards prevailing in the society” (p. 74). Based on this definition, citizens are bestowed with rights in multiple domains: civil, political, and social. Cogan (2000) identifies five categories of citizenship attributes, which have been commonly accepted: sense of identity, enjoyment of specific rights, interest and involvement in public affairs, fulfillment of relevant obligations, and acceptance of basic social values. Digital technologies, in particular the Internet and social media, challenges a geographically bounded conceptualization of citizenship as people increasingly connect and interact with people and issues that may not even be locally relevant.

Digital participation as citizenship rights

Mossberger et al. (2007) put forward a conceptualization of digital citizenship as an extension of citizenship, defining digital citizens as “those who use the Internet regularly and effectively,” and digital citizenship as the “ability to participate in society online” (p. 1). This conceptualization focuses on three areas of digital participation: regular and effective use of prevailing forms of digital communication; participation as democratic citizens through the Internet; and the ability to take full advantage of job market opportunities via the Internet. Similar to Ribble and colleagues, Mossberger et al. (2007) also perceive digital citizenship as ability, but the focus is on equity. Their studies show that Internet use brings benefits to the individual as a conduit for social, political, and economic participation. They further provided empirical data that show digital participation ability is not simply a matter of competence but influenced strongly by pre-existing inequities such as socioeconomic disadvantages (lack of access to the Internet at home and in other settings) and racial inequalities.

While Mossberger et al. (2007) focus on digital participation includes civic and political participation, they did not take on a critical perspective when considering the citizen as a political subject. Political participation was considered within existing frameworks of democratic rights of participation of the nation state. However, just as historically civic and political participation of citizens is often not confined to prevalent legal and political frameworks, the same is true of digital sociopolitical participation. The prevalence of the Internet and social media has brought fundamental changes and challenges to civic and political participation as well as national and international politics and governance.

The rise of DIY citizens and borderless netizens

Law (2021) identified two major roles played by technology in the social creation of knowledge: augmenting human performance and connecting minds over time and space. The impact of these two features of the Internet and social media is evident in the evolution of digital participation since the turn of the new millennium. The accessibility of sophisticated technologies such as powerful search engines, global networks of communication and multimedia production software suites at low (or no additional) cost and low competence thresholds to the average person, especially those in developed economies has led to the phenomenon of “DIY citizenship.” The term “DIY citizens” was coined by Ratto and Boler (2014) to refer to individuals and groups who engage in civic/political actions to challenge and transform the status quo through online and offline means.

The concept of global citizenship (Falk, 1993) emerged before the popularization of the Internet, due to advances in transportation and communication technologies as well as increasing migration across countries and regions. Globalization and transnational cooperation created international/regional systems of rights, civil societies and public spheres that are not coterminous with the nation state, race or culture (Habermas, 1994). Dahrendorf (1994) put forward a framework of rights that transcends the nation state, comprising two dimensions: basic rights (human rights, civil rights and democratic political rights) and enabling rights (social rights such as social services and welfare to ensure that people have the capacity to exercise their basic rights).

The rise of social media has brought about more self-actualizing forms of citizenship. The Internet serves as the space for public engagement, and online network communities provide the sociotechnical infrastructure for propagating ideas, values, and norms as well as for supporting DIY organizational processes. Through various digital means, individuals are now able to choose or form their own online communities and hence their own “citizenship,” and to be more self-determined in their political perspectives and decisions.

Internet activism and place-based social action

Internet and social media bring new possibilities and challenges to democratic social participation. Neumayer and Rossi (2016) reviewed research literature on Internet/social media activism over the 15 years from 2000 to 2014, and identified changes over three dimensions: technology used, geographic location and nature of the activism. At first, the technologies used were referred to in general terms, such as Internet or cyberspace, but later references pinpointed the use of specific technology platforms, such

as “Facebook revolution” and “Twitter revolution.” Until 2008, most of the activism movements focused on US and Canada. From 2009, such activism emerged in places beyond North America, such as in Iran (where the first Twitter revolution took place), the Middle East (the Arab Spring), and China (the Occupy Central in 2014 in Hong Kong). There was also a distinct change in the nature of the activism that emerged, starting from a concern with long standing societal issues such as environmental, human rights, and racial justice to networked political activism that are specific to the local/national context such as the Arab Spring, Indignados and the Occupy movements.

In the early days of the Internet, cyberspace was conceptualized as distinct and separate from the world of physical interactions. *Life on the screen*, as Turkle’s (1995) research showed, provided an escape, an alternate reality for those who have the privilege of access to the Internet, which tended to be staff and students in colleges, universities and research lab in those days. Life on the screen was a separate existence from real life, and inhabitants in cyberspace had a separate virtual identity, often distinct from their identity known to family, friends, and co-workers.

When Turkle wrote her book *Alone Together* (Turkle, 2017), the scenario had totally changed. Talk about computers and the Internet was happening around kitchen tables and playrooms, and people, young and old, were inextricably connected to each other via different kinds of technology-enabled communication from text messaging to video calls to social media platforms. This kind of tremendous change taking place over a short period of 15 years was to a large part prompted by the popularity and accessibility of smartphones, which have become the most convenient channel of Internet access. This has also contributed much to connecting Internet activism to place-based social action.

Gerbaudo (2012) described a “choreography of assembly” (p. 5) orchestrated through skillful integrated use of media and technologies to connect the (cyber)space and place of social action. Social media tools can be used to spread awareness about a movement, broadcast/share information about meetings, rallies, and events, and attract further public attention through national/international news and mainstream media. The Twitter hashtag has become an important tool to connect disparate groups sharing a common cause to rally together. Smartphones with their location aware/location tracking feature serve as a perfect tool to orchestrate offline network activism. Sandoval-Almazan and Gil-Garcia (2014) described a 4-step model of political activism using digital technology: triggering event, media response, viral organization, and physical response. The first three steps take place primarily through the Internet, but the ultimate goal is to rally large numbers of place-based in-person participation in physical protest actions.

Through researching and interviewing key activists involved in three large scale Internet-based political movements that took place in 2011 (the Arab Spring in Egypt, the Spanish indignados, and Occupy Wall Street in New York), Gerbaudo (2012) concluded that social media use play a complementary role to existing forms of f2f interactions and on-street protest actions. Further, such social actions are not leaderless, as often claimed, but leadership is hidden through social media administrators and through the hierarchical nature of social networks in which a few people commands an inordinate amount of influence. However, the mode of communication with the leadership is changed. Organization becomes more fluid, following the communication channels, rather than the other way around as in traditional organizations, and participants communicate directly with the organizers through messages rather than orders.

Tensions between empowering/empowered digital citizens and the (digital) citizenship curriculum

In **Digital citizenship for civic, political and economic participation** section, we have highlighted how digital technology, particularly the Internet, social media, and the pervasive ease of access to smartphones have helped to facilitate grassroots political action of massive scales and the perils of disinformation. The empowerment effects of digital technology stand in contrast to the role played by citizenship education in promoting dutiful citizenship (Bennett, 2008). Citizenship education serves to support students’ assimilation into the mainstream culture and foster national belonging (Banks, 2007).

There are many ways to enact dutiful citizenship in the curriculum. Curriculum is a broad concept, and there are varied approaches. The first and most widely adopted approach is to consider curriculum as an organization of study, defined in terms of a set of subjects, a set of materials, a sequence of courses, a course of study, and/or a set of performance objectives, often referred to as syllabus, content outline, textbooks, etc. As an organization of study, curriculum can also be viewed as a process, referring broadly to school activities as a series of learning experiences—academic, athletic, emotional and social—in the school context (Oliva, 1992; Marsh, 1997; Posner, 1998). A second approach is to conceive curriculum as a learning orientation in response to the perceived overarching educational goals to be achieved. For example, Longstreet and Shane (1993) identify four types of curriculum orientations, namely society-oriented curriculum, student-oriented curriculum, knowledge-oriented curriculum, and eclectic curriculum. The latter looks at curriculum from the broadest possible perspective that involves various compromises related to the purpose and process of teaching and learning. A third approach focuses on classifying curricula based on how they are operationalized in schools. Esiner (1994) classifies school curricula into the formal (or official) curriculum, implemented curriculum, and hidden curriculum. The hidden curriculum is well known for its close relationship to citizenship education as it transmits implicit social norms and political expectations as a part of school life. MacDonald (1977) suggests that the curriculum plays a mediation role between the individual and the societal order and the political order, through a process of internalization developed from school experience.

By discussing curriculum in broader contextual terms, its linkage to citizenship education becomes more salient, especially in its social reconstructionist functions. It is particularly important to note that curriculum can be invisible and intangible, subsumed in

school life and in the society's social and political atmosphere that both teachers and students experience every day. Williams (1976) further argues that the school curriculum, including the hidden curriculum, is not ad hoc but intentionally selected, though often in a tacit manner, on the basis of a larger culture (Giroux, 1981). A whole possible area of past and present, certain meanings and practices can be chosen for emphasis, while certain other meanings and practices can be neglected and excluded. In this way, schools become agents of cultural incorporation, initiating the younger generation into selected cultural traditions. Giroux (1981) concedes that school is reproductive in the cultural sense, functioning to distribute and legitimize forms of knowledge, values, languages, and modes and styles of social etiquette that constitute the dominant culture and its interests. Freire (1974) further points out that education is a means of cultural action and schooling functions as a system of communicating a particular cultural message.

This broader perspective toward curriculum, referred to as the "institutional or policy curriculum" (Doyle, 1992a,b), describes the connection between schooling and society, embodying a conception of what schooling should be with respect to the wider society and culture. In situations when social and cultural contexts change rapidly, the institutional curriculum serves a mediating role between schooling as a socialization tool of the nation-state according to prevalent norms and responding to the changing expectations from stakeholders outside of schools, such as parents, employers, and socio-cultural agents. What we wish to highlight here is that this mediation role has two implications: (1) the institutional curriculum (including the citizenship curriculum, whether formally defined and implemented or as part of the hidden curriculum) will co-evolve over time as social norms and institutions change, and (2) there will be tensions experienced by school leaders and teachers in this process of mediation due to differences in expectations within the school community and with those held by the wider society. This tension is evident in examples of within campus conflicts and participation of students in internet-mediated political activism mentioned in **Digital citizenship for civic, political and economic participation** section. In the Anti-extradition Bill social movement that started in June 2019 in Hong Kong, we see children as young as 11-years-old participating in on-street protests and riot activities as well as occupation and vandalism of university campuses that were orchestrated through the Internet. Schools were also victims of the intensifying polarization in the community, with some facing intense criticisms from some sectors of the community due to the stance and participation of their teachers and students in the social movement. These tensions expose the conflictual role of citizenship curriculum: whether to function as enculturation or critical enculturation. The former is required for social solidarity, whereas the latter is often regarded as important for social progression as testified in the historical development of many countries.

The future of digital citizenship curricula

Throughout history, technological inventions from cavemen paintings to printing, telecommunication, to modern day digital technology has played instrumental roles in transforming human learning through augmenting human performance, providing more powerful ways of knowing, and connecting minds over time and space (Law, 2021). However, the adoption and impact of such technologies, and in particular whether these have really empowered learners in schools and classrooms have never been a simple matter of technology adoption.

Sociotechnical regime challenges in the context of educational transitions

Education is a sociotechnical system, and parallels can be drawn with other similar systems such as transportation and e-Government systems to help us consider the relationship between digital technology, digital citizenship, and the curriculum. Sociotechnical systems involve the integration and adoption of technology systems into organizational and social settings, and their successful deployment cannot be achieved by simply focusing on the technological perspective, without regard to the social and organizational aspects (Cherns, 1976). Geels (2005) analyzed how automobiles took over horse-drawn carriages as the dominant means of urban transport, which shows that the invention of the internal combustion engine around 1860 was only one step in that transition process. To transform the entire transportation system requires changes in the user practices, infrastructures, markets, maintenance networks, production systems, cultural/symbolic meanings, and public policies, etc. connected with the then predominant horse-drawn carriages transportation system to the corresponding ones for transportation systems centered around automobiles. These interconnected and interdependent systems associated with a sociotechnical system is referred to as sociotechnical regimes (Geels, 2005) e-Government is another example of a sociotechnical system. Abunadi and Alqahtani (2019) studied the sociotechnical factors influencing the acceptance by online users of e-transactions in e-Government systems in Saudi Arabia. They identified four key challenges, all belonging to the sociotechnical regime domain in user adoption of e-transactions: a lack of computer literacy and Internet access, an urban-rural digital divide in the availability of adequate technological infrastructures, a lack of public awareness of the available e-services, and inadequate legislation for e-transactions as an official means of conducting business. Based on their empirical findings, Abunadi & Alqahtani (Abunadi and Alqahtani, 2019) recommended changes to the technological, organizational (including policy), and social systems associated with e-transactions to address these challenges.

Integrating digital technology use into education systems is no less challenging and complex compared to the two sociotechnical system changes described above. Despite decades of strong policy directive in many countries to actively promote the use of information and communication technology (ICT) to foster the development of 21st century skills (Plomp et al., 2009), the overall findings have not been positive (e.g., Collins and Halverson, 2009). It is well documented in the technology-enhanced learning literature that deep changes need to take place in teachers' practices far beyond the mastery of technical skills (Albion and Tondeur,

2018). The vision of educating students to become lifelong learners capable of creative problem solving in collaboration with peers is a fundamental shift from the goals of universal education in the industrial era (Collins and Halverson, 2018). Collins and Halverson (2018) elaborated on the barriers to technology use in school and argue that many of the obstacles reflect deep incompatibilities between schools and the transformative visions of technology use in education. Using Geel's (2005) sociotechnical change analytical framework, Law and Liang (2019) show that even when pedagogically innovative technology adoption successfully emerges under conducive policy level support, for example, when digital technology becomes embedded into student-centered curriculum practices, the sustainability of the innovation may still be fragile. For emerging e-learning practices to become mainstream requires community-wide changes in the curriculum outcome goals embraced, as well as continuous adaptations in school routines, curriculum content and structure, forms and foci of assessment, technological and human resources infrastructure and priorities, as well as ways of working with other stakeholders such as parents and community partners, all of which are interconnected and interdependent elements of the sociotechnical regime within any education system.

The sociotechnical changes to education brought about by digital technologies is not only enhancing students' access to online learning, but also playing an increasingly significant role in students' socialization, which is a significant function of the education system, especially citizenship education. Properly recognizing the socialization role of digitalization will necessitate the integration of digital citizenship into the curriculum. In the previous section, we highlighted tensions between empowering/empowered digital citizens and the DC curriculum. Such tensions can be further examined within the broader context of sociotechnical change. On the one hand, for digital technology to realize its transformative educational potential requires sociotechnical regime level changes toward more student-centered pedagogy, more open, exploratory, and inquiry-oriented curriculum, and more democratic, non-hierarchical school culture that gives voice to all individuals. On the other hand, the (digital) citizenship curriculum is an integral part of the overall school curriculum, and generally serves an important role as a socialization tool of the nation-state to preserve the status quo according to prevalent norms. The prevalent focus on digital competence and socioemotional wellbeing in current digital citizenship curricular reflects an empowerment orientation. However, it is also evident that the empowerment is predominantly oriented toward the learners' personal and social wellbeing at an individual level, leaving the social political engagement aspects of citizenship education relatively untouched. There is little evidence that the DC curriculum has reflected the changing conceptualizations and debates around DC beyond technological competence. Such a situation is not surprising from a sociotechnical change perspective, as the development of a DC curriculum catering for social political participation in the digital era is a sociotechnical regime level change which is much more challenging. However, we also argue that this under-development is highly undesirable, and may cause serious ramifications to stability at the sociotechnical regime and landscape levels, as will be elaborated in the remainder of this section.

Need for continuity between citizenship and digital citizenship education

Choi (2016) identified four key categories of DC conceptualization through a systematic conceptual analysis of the literature: ethics, media and information literacy, participation and engagement, and critical resistance. The first two categories are the key foci in existing DC curricula as discussed in the **Digital citizenship as digital competence** section. The third category, participation and engagement (P&E), comprises two forms of engagement—macro and micro. The micro (or personalized) forms of DC engagement refers to interest-driven online activities, including games, popular culture, various forms of self-expression, and entertainment related activism as a form of civic engagement. This category of digital engagement is also increasingly addressed in some DC curricula that supports e-safety and social emotional learning as described in the **Digital citizenship development requires social emotional learning** section. Macro forms of DC engagement refers to ways of using the Internet as an extended sphere for political participation, which includes e-voting, online petitions for political/economic causes, and online electioneering. While macro forms of P&E addressed within the scope of traditional citizenship curricula remain conservative, these are not even included as a topic under existing DC curriculum at the school level.

Whereas P&E in Choi's (2016) DC conceptualization focuses on online participation that are legitimate under existing systems, the critical resistance category pursues more progressive and radical means and outcomes, through critique of existing power structures and/or political activism, as described in the **Internet activism and place-based social action** section on internet activism and place-based social action. Critical resistance is not a theme even under the traditional school curriculum. It can be argued that this is perfectly reasonable since an important role of education as an integral part of the sociotechnical regime is to preserve the existing sociopolitical system and values. On the other hand, there are many examples in the history of scientific and technological advancement, such as nuclear reactions, in which the operations and consequences of powerful inventions fail to be deliberately and systematically controlled. Regarding Internet technology and education, Glassman and Burbidge (2014) warn of the risk of "ceding control of the Internet's relationship with education to small, entrenched populations attempting to control rather than expand possibilities" (Glassman and Burbidge, 2014, p. 17). It is important for the wellbeing of our society that our younger generation have the opportunity to understand the possibilities that digital technology, and in particular digital social media, brings to socio-political participation, as well as the threats that these may pose to upholding the core values of social justice and democracy. In an in-depth action research on how a grassroots movement appropriates a variety of digital and social media tools for their science activism, Ghoshal and Bruckman (2019) showed that there are values embedded in the technology systems deployed, and that the core values of inclusivity, privacy/security, and social translucence are given different priorities by different technology tools. Proprietary owners of social media platforms and power users making decisions on the choice and set up of digital technologies

for P&E have advantaged access and influence that could jeopardize the very core values that participants in (place-based connected) digital social movements strive to uphold.

Place-based citizenship, global awareness and global citizenship, and digital citizenship are connected facets of one's dialectic interactions with the social, political, and digitally connected world. One notable effort in bridging these different facets of citizenship is the Digital Citizenship Education (DCE) framework developed by the Council of Europe (Richardson and Milovidov, 2019). The DCE framework comprises 10 domains categorized under three themes: (theme 1) being online—access and inclusion, learning and creativity, media and information literacy; (theme 2) well-being online—ethics and empathy, health and well-being, e-presence and communications; and (theme 3) rights online—active participation, rights and responsibilities, privacy and security, and consumer awareness. Within this framework, children as citizens have the right to both protection and participation, and the competences to be developed within the 10 domains are explicitly connected to active engagement and participation, underpinned by the core values of human dignity and human rights. The framework (Richardson and Milovidov, 2019) also puts forward a “five-pillar” model for its implementation. Four of the pillars are sociotechnical regime level elements: policies, strategies, infrastructure and resources, and evaluation, and the fifth pillar is multi-stakeholder engagement. This framework addresses the discontinuity between traditional citizenship education and the concerns prevalent in digital citizenship curricula and takes account of the need for regime level changes for its realization. Hopefully, there will be more efforts to extend DC curriculum along these two directions.

Summary

Digital citizenship in the school curriculum has been changing over time, but is still primarily confined to digital competence and e-safety/cyber wellness. Inclusion in the school curriculum regarding civic or political engagement and participation through using digital technology is still rare. This contrasts with the more complex and nuanced changes in the conceptualizations of digital citizenship in the wider academic literature to take account of the deep changes in the landscape of sociopolitical engagement and participation worldwide brought about by the use of digital technologies. Citizenship mainly refers to the obligational relationship between an individual and the larger community that the individual belongs to. For a long time, citizenship was synonymous with nation-state citizenship. With increasing mobility across countries and regions due to advances in global transportation systems and sociopolitical changes in many parts of the world, nation-state citizenship has been broadened to include post-national, supranational, and global citizenship. The digital world has become another space for global actions and interactions, and it is important to consider digital citizenship as an extension of citizenship engagement in holistic perspectives embracing its full social, political, and cultural richness. Social media platforms have complemented existing forms of face-to-face interactions and on-street protest action in a “choreography of assembly” (Gerbaudo, 2012, p. 5) to connect the (cyber)space and place of social action. The traditional focus of citizenship education on promoting dutiful citizenship is inadequate in addressing the tensions and challenges created by the empowerment effects of digital technology, as evidenced by the examples of within campus conflicts and participation of students in internet-mediated political activism. Digital citizenship education needs to take on the responsibility of ensuring that children and youth have the right to protection and participation by empowering students with appropriate competence, attitudes, and values for their digitally mediated participation to contribute to enhancing personal and social wellbeing.

Acknowledgment

Nancy Law acknowledges the support given by the Research Grants Council of the HKSAR Government, #T44-707/16/N, under the Theme-based Research Scheme for her contribution.

References

- Abunadi, I., Alqahtani, F.H., 2019. Sociotechnical factors in the endorsement of governmental e-transactions. *J. Global Inf. Manag.* 27 (3), 38–64.
- Akinboye, G., 2015. A (re)view of the notion of citizen and citizenship in the Athens of Aristotle. *Cross Cult. Commun.* 11 (5), 1–7.
- Albion, P.R., Tondeur, J., 2018. Section Introduction: Professional Learning and Development of Teachers. J. Voogt, G. Knezek, R. Christensen, K.W. Lai (Eds.), pp. 381–396.
- Anderson, B., 1983. *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. Verso, London.
- Axtmann, R., 2004. The state of the state: the model of the modern state and its contemporary transformation. *Int. Polit. Sci. Rev.* 25 (3), 259–279.
- Banks, J.A., 2007. *Diversity and Citizenship Education: Global Perspectives*. Jossey-Bass, Indianapolis, IN.
- Baubock, R., 2007. Why European citizenship? Normative approaches to Supranational Union. *Theor. Inq. Law* 8 (2), 453–460.
- Bennett, W.L., 2008. Changing citizenship in the digital age. In: Bennett, W.L. (Ed.), *Civic Life Online: Learning How Digital Media Can Engage Youth*. MIT Press, Cambridge.
- Carretero, S., Vuorikari, R., Punie, Y., 2017. *The Digital Competence Framework for Citizens*. Publications Office of the European Union. <https://doi.org/10.2760/38842>.
- Chakroun, B., Kevy, J., 2018. Digital Credentialing: Implications for the Recognition of Learning Across Borders. UNESCO, Paris. Retrieved from. <https://unesdoc.unesco.org/ark:/48223/pf0000264428?posInSet=16&queryId=e0c44f09-46ff-4fd3-b7de-12ec9c76afa6>.
- Chen, L.L., Mirpuri, S., Rao, N., Law, N., 2021. Conceptualizations and measurement of digital citizenship across disciplines. *Educ. Res. Rev.* <https://doi.org/10.1016/j.edurev.2021.100379>.
- Cherns, A., 1976. The principles of sociotechnical design. *Hum. Relat.* 29 (8), 783–792.
- Choi, M., 2016. A concept analysis of digital citizenship for democratic citizenship education in the internet age. *Theor. Res. Soc. Educ.* 44 (4), 565–607.

- Cogan, J.J., 2000. Citizenship education for the 21st century: setting the context. In: Cogan, J.J., Derricott, R. (Eds.), *Citizenship for the 21st Century: An International Perspective on Education*. Kogan Page Ltd, London, pp. 1–22.
- Collins, A., Halverson, R., 2009. *Rethinking Education in the Age of Technology: The Digital Revolution and the Schools*. Teachers College Press, New York.
- Collins, A., Halverson, R., 2018. *Rethinking Education in the Age of Technology: The Digital Revolution and Schooling in America*. Teachers College Press, New York.
- Dahrendorf, R., 1994. The changing quality of citizenship. In: van Steenberg, B. (Ed.), *The Condition of Citizenship*. SAGE Publications, London, pp. 10–19.
- Datu, J.A.D., Wong, G.S.P., Rubie-Davies, C., 2021. Can kindness promote media literacy skills, self-esteem, and social self-efficacy among selected female secondary school students? An intervention study. *Comput. Educ.* 161, 104062.
- Doyle, W., 1992a. Curriculum and pedagogy. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan, New York, pp. 486–516.
- Doyle, W., 1992b. Constructing curriculum in the classroom. In: Ser, F.K., Patry, J. (Eds.), *Effective and Responsible Teaching: The New Syntheses*. Jossey-Bass, San Francisco, pp. 66–79.
- Eisner, E., 1994. *The Educational Imagination*, third ed. Macmillan, New York.
- Falk, R., 1993. The making of global citizenship. In: Brecher, J., Childs, J.B., Cutler, J. (Eds.), *Global Visions: Beyond the New World Order*. South End Press, Boston, pp. 39–50.
- Ferrari, A., 2013. DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe. Publications Office of the European Union, Luxembourg.
- Freire, P., 1974. Education: domestication or liberation? In: Lister, I. (Ed.), *Deschooling*. Cambridge University Press, Cambridge, pp. 18–21.
- Geels, F.W., 2005. The dynamics of transitions in socio-technical systems: a multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860–1930). *Technol. Anal. Strat. Manag.* 17 (4), 445–476.
- Gerbaudo, P., 2012. *Tweets and the Streets: Social Media and Contemporary Activism*. Pluto Press, London.
- Ghoshal, S., Bruckman, A., 2019. The role of social computing technologies in grassroots movement building. *ACM Trans. Comput. Hum. Interact.* 26 (3), 1–36.
- Giroux, H., 1981. *Ideology, Culture, and the Process of Schooling*. Falmer Press, London.
- Glassman, M., Burbidge, J., 2014. The dialectical relationship between place and space in education: how the internet is changing our perceptions of teaching and learning. *Educ. Theor.* 64 (1), 15–32.
- Habermas, J., 1994. Citizenship and national identity. In: van Steenberg, B. (Ed.), *The Condition of Citizenship*. SAGE Publications, London, pp. 20–35.
- ISTE, 2007. *National Educational Technology Standards for Students*. ISTE (International Society for Technology Education), Eugene, Oregon.
- ISTE, 2016. *National Educational Technology Standards for Students*. Retrieved from Eugene, Oregon. <https://iste.org/standards>.
- James, C., Weinstein, E., Mendoza, K., 2019. *Teaching Digital Citizens in Today's World: Research and Insights Behind the Common-Sense K–12 Digital Citizenship Curriculum*. Common Sense Media.
- Law, N., Liang, L., 2019. SocioTechnical coevolution of an eLearning innovation network. *Br. J. Educ. Technol.* 50 (3), 1340–1353. Retrieved from: <https://onlinelibrary.wiley.com/doi/full/10.1111/bjet.12768>.
- Law, N., 2021. Design for learning in an age of rapidly evolving technology. In: Carroll, A., Cunnington, R., Nugent, A. (Eds.), *Learning Under the Lens: Applying Findings From the Science of Learning to the Classroom*. Routledge, New York, pp. 141–153.
- Lee, W.O., Fouts, J.T. (Eds.), 2005. *Education for Social Citizenship: Perceptions of Teachers in the USA, Australia, England, Russia and China*. Hong Kong University Press, Hong Kong.
- Longstreet, W.S., Shane, H.G., 1993. *Curriculum for a New Millennium*. Allyn and Bacon, Boston.
- MacDonald, M., 1977. *The Curriculum and Cultural Reproduction*. The Open University Press, London.
- Mandelbaum, M.M., 2013. One state-one nation: the naturalization of nation-state congruency in IR theory. *J. Int. Relat. Dev.* 16, 514–538.
- Marsh, C.J., 1997. Key Concepts for Understanding Curriculum, vol. 1. Falmer Press, London.
- Marshall, T.H., 1992. Citizenship and social class. In: Moore, R. (Ed.), *Citizenship and Social Class*. Pluto Press, London.
- Mossberger, K., Tolbert, C.J., McNeal, R.S., 2007. *Digital Citizenship: The Internet, Society, and Participation*. MIT Press, Cambridge, MA, p. 221.
- Moyle, K., 2014. Technologies, democracy and digital citizenship: examining Australian policy intersections and the implications for school leadership. *Educ. Sci.* 4 (1), 36–51.
- Nassacimbeni, F., 2019. *Digital Literacy for Children: Exploring Definitions and Frameworks*. Retrieved from: <https://www.unicef.org/globalinsight/media/1271/file/%20UNICEF-Global-Insight-digital-literacy-scoping-paper-2020.pdf>.
- Netsafe, 2018. *From Literacy to Fluency to Citizenship: Digital Citizenship in Education*. Netsafe, Wellington, NZ. Retrieved from: https://www.netsafe.org.nz/the-kit/wp-content/uploads/2018/07/From-literacy-to-fluency-to-citizenship_July-2018.pdf.
- Neumayer, C., Rossi, L., 2016. 15 years of protest and media technologies scholarship: a sociotechnical timeline. *Soc. Media Soc.* 2 (3), 2056305116662180.
- NSW Department of Education, (n.d.). *Digital Citizenship*. Retrieved from: <https://www.digitalcitizenship.nsw.edu.au/>.
- OBERLO, 2020. *Most Popular Social Media Platforms in 2020*. Retrieved from: <https://www.oberlo.com/statistics/most-popular-social-media-platforms>.
- OECD, 2018. *Preparing Our Youth for an Inclusive and Sustainable World: The OECD Global Competence Framework*. OECD, Paris. Retrieved from: <https://www.oecd.org/education/global-competency-for-an-inclusive-world.pdf>.
- OECD, 2019. *OECD Skills Outlook 2019: Thriving in a Digital World*. OECD, Paris. Retrieved from: https://books.google.com.sg/books?id=DRSXDWAAQBAJ&pg=PA166&lpg=PA166&dq=OECD+digital+citizenship+curriculum&source=bl&ots=oyWAwtf3D&sig=ACfU3U1yq7j0ibicrU3he2eMVPMS0jT_pg&hl=zh-TW&sa=X&ved=2ahUKEwiY-qSmruHuAhXCTX0KHRTVBXk4HhDoATAJegQIBxAC#v=onepage&q=OECD%20digital%20citizenship%20curriculum&f=false.
- Oliva, P.F., 1992. *Developing the Curriculum*, third ed. Harper Collins, New York.
- Osler, A., Starkey, H., 2003. Learning for cosmopolitan citizenship: theoretical debates and young people's experiences. *Educ. Rev.* 55 (3), 243–254. <https://doi.org/10.1080/0013191032000118901>.
- Plomp, T., Anderson, R.E., Law, N., Quale, A. (Eds.), 2009. *Cross-national Information and Communication Technology Policy and Practices in Education*, second ed. Information Age Publishing Inc, Greenwich, CT.
- Posner, G.J., 1998. Models of curriculum planning. In: Beyer, L.E., Apple, M.W. (Eds.), *The Curriculum: Problems, Politics and Possibilities*, second ed. State University of New York Press, Albany, pp. 79–100.
- Ratto, M., Boler, M., 2014. *DIY Citizenship: Critical Making and Social Media*. MIT Press.
- Ribble, M., Bailey, G., 2007. *Digital Citizenship in Schools*. International Society for Technology in Education, Eugene, OR.
- Ribble, M., 2011. *Digital Citizenship in Schools*, second ed. International Society for Technology in Education, Eugene, OR.
- Ribble, M., 2015. *Digital Citizenship in Schools: Nine Elements All Students Should Know*. International Society for Technology in Education.
- Richardson, J., Milovidov, E., 2019. *Digital Citizenship Education Handbook: Being Online, Well-Being Online, and Rights Online*. Council of Europe.
- Sadiq, K., 2008. *Paper Citizens: How Illegal Immigrants Acquire Citizenship in Developing Countries*. Oxford Scholarship Online. <https://doi.org/10.1093/acprof:oso/9780195371222.003.0004>. Retrieved from: <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780195371222.001.0001/acprof-9780195371222-chapter-4>.
- Sandoval-Almazan, R., Gil-Garcia, J.R., 2014. Towards cyberactivism 2.0? Understanding the use of social media and other information technologies for political activism and social movements. *Govern. Inf. Q.* 31 (3), 365–378.
- Shin, T.S., Hwang, H., Park, J., Teng, J., Dang, T., 2019. *Digital Kids Asia-Pacific: Insights Into Children's Digital Citizenship*. UNESCO Asia and Pacific Regional Bureau for Education, Bangkok.
- Singapore Ministry of Education, 2014. *2014 Syllabus Cyber Wellness Secondary*. Author, Singapore.
- South West Grid for Learning, (n.d.). *Digital Literacy & Citizenship*. Retrieved from: <https://digital-literacy.org.uk/>.

- Subedi, B., Rhee, J.-E., 2018. Transformative spirituality and citizenship. In: Davies, I., et al. (Eds.), *The Palgrave Handbook of Global Citizenship and Education*. Palgrave Macmillan, London, pp. 377–392.
- Tamayo, P.D., 2016. Digital Citizenship Recommendations. Report to the Legislature. Washington State Office of Superintendent of Public Instruction, Olympia, WA. Retrieved 13 Feb 2021 from: <https://www.k12.wa.us/sites/default/files/public/edtech/digitalcitmedialit/pubdocs/2016-12-digitalcitizenship-legislativereport.pdf>.
- Tambini, D., 2001. Post-national citizenship. *Ethn. Racial Stud.* 24 (2), 195–217. <https://doi.org/10.1080/01419870020023418>.
- Turkle, S., 1995. *Life on the Screen: Identity in the Age of the Internet*. Simon and Schuster, New York.
- Turkle, S., 2017. *Alone Together: Why We Expect More From Technology and Less From Each Other*. Hachette, UK.
- UNESCO, 2014. *Global Citizenship Education: Preparing Learners for the Challenges of the 21st Century*. UNESCO, Paris.
- Vuorikari, R., Punie, Y., Gomez, S.C., Van Den Brande, G., 2016. DigComp 2.0: The Digital Competence Framework for Citizens. Update Phase 1: The Conceptual Reference Model. Retrieved from: <https://publications.jrc.ec.europa.eu/repository/handle/JRC101254>.
- Wagschal, U., 1997. Direct democracy and public policymaking. *J. Publ. Pol.* 17 (2), 223–245.
- Williams, R., 1976. Base and superstructure in Marxist cultural theory. In: Dale, R., et al. (Eds.), *Schooling and Capitalism: A Sociological Reader*. Routledge & Kegan Paul, London.
- Zajda, J., Daun, H., Saha, L.J., 2009. *Nation-Building, Identity and Citizenship Education*. Springer Publishers, Dordrecht.

Curriculum and democracy: building understandings of democracy through early childhood practice

Claire Jane McLachlan, Institute of Education, Arts and Community, Federation University Australia, Ballarat, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Definitions of curriculum and democracy	424
Democracy	424
Definitions of curriculum in relation to democracy in early childhood education	425
Ideologies and curriculum	427
Children's rights: part of democratic practice?	429
Summary and conclusions	431
References	431

We conceive of democracy as embedded in participation and deliberation, with a strong emphasis on the participatory rights in the CRC (United Nations Convention of the Rights of the Child, 1989). These rights include a child's right to expression, the child's right to an opinion given due consideration by adults, access to appropriate information and to freedom of association; all in accordance with the age and maturity of the child. In our research projects, children are recognized as agents in society, not as immature objects to be excluded from deliberative and other political processes. The opportunity to be respected members of the public discourse provides young citizens with real life experiences that enable the development of language and reasoning skills, and especially skills in perspective taking and critical thinking that are fundamental to the actors in an equitable and sustainable democratic society

Carlson et al. (2002, p. 39).

This quote, focused on democracy and children in Costa Rica nearly twenty years ago, is a good place to start the discussion of democracy, curriculum and childhood for young children. It is clear that although the United Nations Convention on the Rights of the Child (UNCROC) (United Nations, 1989) is a very important piece of legislation in the signatory countries, whether the focus on children's rights has translated into children having agency and a recognized "voice" in how the curriculum in their educational setting is enacted is paradoxically another issue. This chapter will begin with some definitions of curriculum and how they enable understandings of democracy in the early years. Then some of the ideological underpinnings of curriculum are explored in relation to curriculum approaches. The different types of curriculum are critiqued to see whether they are fit for purpose for supporting children to become democratic citizens of a society. Finally, the chapter explores the notion of children's rights in relation to democracy and what is needed to enable children to become democratic citizens in the early years.

Definitions of curriculum and democracy

It is arguably useful to start any discussion of early childhood curriculum and democracy with some analysis of what a democracy is and how it might be related to early childhood settings. Sources used are both seminal and contemporary, drawing on internationally published scholarly theoretical and empirical research and some commissioned reports, such as those by OECD. A search for relevant literature in library databases and gray literature was conducted using the search terms curriculum, curricula, democracy and early childhood education. Where possible, evidence-based sources have been drawn on. Because the author is a New Zealander, some examples are drawn from New Zealand to illustrate key points.

Democracy

In a seminal argument, Cohen (1971) proposed that a democracy is a community in which everybody has the opportunity to influence decisions that are of importance to them. He argued that the ways citizens participate, inhabit, negotiate, and challenge the given conditions for participation indicate how democratic their situation is. According to Cohen (1971), different ways of participating, both quantitatively and qualitatively, indicate the breadth, depth, and range of democracy. Democracy can be expanded by widening community input into these factors. For early childhood communities, this notion would therefore concern how many people influence decision making, what level of involvement different individuals—including children—have, and who has authority to make decisions. Participation is the key concept, and it is one which becomes problematic when children are not defined as citizens with agency in the early childhood setting (Einarsdottir and Wagner, 2006; MacNaughton, 2007).

However, while Cohen (1971) relates democracy primarily to participation, he does not problematize the values that are interwoven in the ways of participation. Biesta (2011) argues that democratic participation needs reference to democratic values, and this is a point of connection with some early childhood curricular internationally (Einarsdottir et al., 2015). Moss (2011) provides a helpful twofold characterization of democracy and education as inseparably interconnected: democracy as a basic value and

practice in education; and education as a means to strengthen and sustain democracy. This is clearly something that can be promoted through engagement with young children in early childhood contexts.

For Held (2006, p. 1), democracy *prima facie* means “a form of government in which, in contradistinction to monarchies and aristocracies, the people rule. Democracy entails a political community in which there is some form of political equality among the people”. However, he also acknowledges the complexity of the topic and the potential for disagreement over definitions (Sousa and Oxley, 2019).

Although democracy is not the principal reason that early childhood education is advocated in many countries, the ability to be a good citizen is. Indeed, developing contributing members of society is one of the key arguments for investing in early childhood (Heckman, 2012; Heckman and Masterov, 2007). The arguments around this stem primarily from an economic analysis, not a values based one, but the outcome of having well educated members of a society is intertwined and investment provides strongest outcomes for the most disadvantaged children. As the noted economist, Heckman (2012) argues:

The highest rate of return in early childhood development comes from investing as early as possible, from birth through age five, in disadvantaged families. Starting at age three or four is too little too late, as it fails to recognize that skills beget skills in a complementary and dynamic way. Efforts should focus on the first years for the greatest efficiency and effectiveness. The best investment is in quality early childhood development from birth to five for disadvantaged children and their families.

Many nations have understood the need for economic investment in their youngest children, as part of increasing equitable outcomes for all children and thereby demonstrating principles of democracy. For example, the policy document that drove major change in early childhood education in New Zealand “*Education to be more*” (Department of Education, 1988) cites the outcomes of early childhood education as increased educational achievement and reduced need for special education, increased likelihood of employment, reduced delinquency and teenage pregnancy and an economic outcome of paying for itself by saving the taxpayer \$400–700 for every \$100 spent on early childhood. The rationale underpinning these documents was therefore twofold: early childhood education was defined in economic terms as having long term benefits for children, families; and ensuring the state’s goals for the economy were achieved. The government had two goals: to equip New Zealanders with 21st century skills, including democratic citizenship; and to reduce systematic underachievement in education; sometimes known by the catch phrase “raise achievement and reduce disparity” (Adams and Bethell, 2005, p. 144). However, achieving these outcomes came with a price tag: the introduction of a “charter” for licensed early childhood services, which mandated increased accountability and government control via licensing through the newly formed Ministry of Education and review via its quality assurance arm, the Education Review Office.

John Dewey is often cited as a key advocate of democracy as an outcome of education. Dewey (1916) argued that children learn about being a citizen through engagement with projects which challenge them to understand diverse perspectives. Central to Dewey’s educational thinking was the idea that project-based enquiry could engage learners with the knowledge needed to sustain a process of reconstruction (Dewey, 1916). He proposed that inquiry-based learning would help children develop and to overcome the dualism that may exist between the child and the curriculum and the school and society. In such a way, the curriculum is made relevant and meaningful to children, their previous experiences, and their life contexts. As Dewey proposed “the teacher should be occupied not with subject matter in itself but in its interaction with the pupils’ present needs and capacities” (Dewey, 1916, p. 191). Dewey argued that by involving children in learning things that are of interest to them and giving them choices about what they learn helped children to learn the link between education and a democratic and participative way of life. This principle of allowing children to select the focus of their learning and the pace and sequence of early childhood curriculum is part of many early childhood curriculum and is explored next.

Definitions of curriculum in relation to democracy in early childhood education

Curriculum is a complex concept influenced by various factors that reflect a society’s political, policy and technical consensus on the goals and values regarding children’s learning, development and well-being (Litjens and Taguma, 2010). According to Scott (2008), “A curriculum may refer to a system, as in a national curriculum; an institution, as in a school curriculum; or even to an individual school, as in the school geography curriculum” (pp. 19–20). According to Scott (2008), a curriculum can be organized specifically to include four dimensions:

aims or objectives, contents or subject matter, methods or procedures, and evaluation and assessment. The first dimension refers to the reasons for including specific items in the curriculum and excluding others. The second dimension is content or subject matter and this refers to knowledge, skills or dispositions which are implicit in the choice of items, and the way that they are arranged. Objectives may be understood as broad general justifications for including particular items and particular pedagogical processes in the curriculum; or as clearly defined and closely delineated outcomes or behaviours; or as a set of appropriate procedures or experiences. The third dimension is methods or procedures and this refers to pedagogy and is determined by choices made about the first two dimensions. The fourth dimension is assessment or evaluation and this refers to the means for determining whether the curriculum has been successfully implemented (pp. 19–20).

Thus, the four crucial elements which apply to curriculum in relation to building democracy include the following:

1. *Aims, goals, objectives or outcome statements*—what do we want this curriculum to achieve in terms of children and democracy, what would we expect to be the outcomes in terms of democracy as a result of participating in the implementation of this curriculum?
2. *Content, domains or subject matter*—what will we include or exclude from our curriculum to understandings of democracy and democratic practice?
3. *Methods or procedures*—what teaching methods or approaches will we use to achieve these goals and values of democracy?
4. *Evaluation and assessment*—how will we know when we have achieved them?

Scott's (2008) definition is obviously broad, and it equally applies to many levels of the education system. However, what a community expects its education system to achieve is a strong force in determining curriculum content. As the OECD (2018) state, a curriculum framework is an over-arching document that articulates the broad vision of curricula within the broader context of the education system. As the OECD (2018) argue, not all curricula or curriculum frameworks present a clear theoretical foundation of a particular pedagogical approach. Rather, some are broad and can embrace varied pedagogies. Curriculum is a crucial agent of education policy as it deliberately specifies what is valued for a particular level of education, potentially including learning objectives, content, methods (including assessment) and materials for teaching and learning, as well as arrangements for teacher training and professional development (OECD, 2018; Sylva et al., 2016). Therefore, in terms of promoting democracy through education, a country's community and its government will make key decisions about what role democracy plays in the curriculum of young children in local, regional and national communities, although there may be marked differences among communities within a nation.

According to the OECD (2018), curriculum is "a political, policy and technical agreement among the various institutions and stakeholders, from both inside and outside the education system, on why, what, how, when and where to educate and learn" (in OECD, 2019, p. 10). Within this definition, curriculum can be described as "intended", "implemented" or "achieved". The intended—also official or planned—curriculum specifies what authorities and society expect students will learn at school and how the outcomes of the teaching and learning process will be assessed. The implemented curriculum, also defined as the "taught curriculum" or "curriculum in action", is the "actual teaching and learning activities taking place in schools through interaction between learners and teachers as well as among learners" (OECD, 2019, p. 10). Finally, the achieved curriculum may differ from the intended and implemented curriculum as it "indicates what learners actually acquire as a result of teaching and learning, assessed through different means and/or demonstrated in practice" (OECD, 2019, p. 10).

In addition to the notions of intended, implemented, and achieved curriculum, Bernstein (1996) argues that there are essentially two models of curriculum and that all curricula fall into one or the other category. He states that curriculum is oriented to either *performance* or *competence* and that *performance* models of curriculum are the most dominant around the world. The *performance* model has its origin in the behavioral objectives' movement. It is a model that clearly emphasizes marked subject boundaries, traditional forms of knowledge, explicit rules for pedagogic practice and strong boundaries between different types of students (Scott, 2008, p. 4). In contrast, the *competence* model suggests that learners have some control over the selection, pacing and sequencing of the curriculum, echoing Dewey's (1916) notions of children being able to make choices to learn about what interests them. Competence models have been more common in early childhood education, such as in the early childhood curriculum of Australia and New Zealand (DEEWR, 2009; Ministry of Education, 1996, 2017), while some curriculum models also clearly promote notions of agency and democracy as part of early childhood curricula (Einarsdottir and Wagner, 2006). However, the dominance of competence models of curriculum in early childhood education is changing in some countries; the *Foundation Stage* curriculum in the UK (Aubrey, 2004; Standards and Testing Agency, 2016), for example, has clearly defined curriculum outcomes and achievement standards for very young children.

Using Bernstein's (1977) notion of classrooms as message systems—consisting of curriculum, pedagogy and evaluation—Munns and Martin (2005) argue that such message systems operate within "discourses of power", especially discourses around knowledge, ability, control, place and voice. They argue that these issues influence the ways in which teachers and learners view themselves in the classroom. They propose that the following questions need to be answered about the curriculum:

- What counts as knowledge and who has access to really useful knowledge?
- Who has ability?
- Who controls the teaching space?
- Who is valued as an individual and a learner?
- Whose voice is given credence within that space?

Munns and Martin (2005) argue that historically students from poor backgrounds have received disengaging messages about their knowledge and ability, their control of their learning, their place in school and community and their rights to a voice. Classroom processes and assessment practices have convinced them that they are not capable of achieving (ability). Their classrooms have been characterized by daily struggles over discipline and behavior (control). They have not seen themselves as valued as individuals and learners (place). There has been little opportunity for them to have a say in their learning processes and achievements (voice). This has resulted in their often not gaining classroom access to powerful and contextualized knowledge. All of these issues have clear implications for children's sense of agency and subsequent ability to feel like they can contribute as equal citizens, a central tenet of democracy.

Ideologies and curriculum

Another important dimension of curriculum is the ideology that underpins its design. Schiro (2013) argues that there are four dominant ideologies which shape curriculum practice in the United States, which have equally shaped ideology about curriculum and the potential place of democracy in curriculum in other countries too. These four ideologies are:

1. Scholar academic ideology
2. Social efficiency ideology
3. Learner-centred ideology
4. Social reconstruction ideology.

The *scholar academic ideology* is based on the notion that each culture has accumulated a body of knowledge over the centuries, which is organized into academic subject disciplines, which is needed to take part in contemporary society. The OECD (2006) classified any ECE curriculum that follows this ideology as a “pre-primary approach”, because the content of the curriculum mirrors traditional subject teaching in primary schools. The child is viewed as incomplete in this ideology—a “neophyte” or immature member of the discipline—who is capable of developing intellect, memory, and reasoning, shaped by the discipline (Schiro, 2013). This immature, “blank slate” or “empty vessel” view of the child, drawing on the notions of Aristotle and John Locke (Cleverly and Phillips, 1995), suggests that the role of the school or early childhood setting is to fill the child with predetermined knowledge and skills, a passive model of child development and learning. Learning is thus viewed as a function of teaching: the teacher is a transmitter, and the child is a receiver. De Castro (2012) argues that in such models, children labor in the present for coming into a social existence in the future and are thereby set aside from relevant social interactions and opportunities because they are not yet deemed prepared to take part in them. In a critique of such models, de Castro (2012) argues that the focus is on what children cannot do, rather than what they can do, and the child is constructed as non-adult and therefore irrational, vulnerable, immature, and dependent. In this model, the child is therefore the receiver of pre-existing models of democratic understanding over time and it is unlikely that the child will be involved in democratic decision making until seen to be at an appropriate age and stage for such learning within the curriculum. As Bubikova-Moan et al. (2019) argue, such approaches to curriculum are didactic and focused on explicit outcomes, such as school readiness. In this model, children are unlikely to readily experience or gain understandings of democracy.

The *social efficiency ideology* is based on the writings of Bobbitt (1918), and his book *The Curriculum*, and further developed by Ralph Tyler in 1949 in *Basic Principles of Curriculum and Instruction*, in which both authors declared that educators should learn to use the scientific techniques of production developed by industry to improve the outcomes of education. In this ideology, the school is the equivalent of a factory, and the child is the raw material: an adult is the finished product; the teacher is the factory worker, who runs the machines. A teacher in this model must determine what the finished product needs to be and then determine the most efficient way to achieve that outcome. Bobbitt (1918) proclaimed that the educational “task preceding all others is the determination of ... a scientific technique” of curriculum design (in Schiro 2013, p. 57).

If a democratic citizen was the desired outcome, then teachers would devise curriculum which would enable this learning to happen. Social efficiency educators devote much time to developing objectives in behavioral terms: as observable skills, capabilities for action, activities to be performed—demonstrable things people can do. For example, a behavioral objective related to democracy in early childhood might include social competency: in this case, the ability to take turns during an organized activity. The actions are clearly specified—taking turns in an organized game—and the child’s ability to take turns can be clearly observed and measured during the activity. An efficient curriculum in this ideology is one in which children experience a controlled learning environment, in which content is carefully sequenced and efficiently organized. This is often called a programmed curriculum. For example, the curriculum would be organized so that teachers could plan their teaching programs for incremental success of children becoming democratic. The curriculum developer’s role is that of “behavioral engineer” (Holland, 1960).

The assumptions underpinning this ideology involve the assumption that learning is in response to direct stimuli and that it’s a teacher’s role to organize stimuli and provide increasingly more complex organized activities to promote learning, which will stimulate children to practice the desired behaviors, which will result in the desired learning outcome. This approach to curriculum development supports the teaching method of planning for observable and directly measurable objectives. In many ways the Early Years Foundation Stage Profile (Standards and Testing Agency, 2016) with its 17 specified learning goals and achievement standards parallels this notion. The High/Scope® curriculum, although focused on children being active learners, is another example, as the curriculum includes 58 key experiences in child development for the preschool years and a wide range of practical strategies for promoting these key experiences (OECD, 2004). Goal 3 Initiative and Social Relations (to make plans, decisions, solve problems encountered in play, express feelings, be sensitive to others) is related to loosely related to notions of being a good citizen, but the notion of being agentic is missing from this and other goals. Education is therefore seen as a shaping process in this ideology, preparing children to create a better society and this notion is consistent with the definitions of democracy proposed by Cohen (1971), argued earlier. Children are seen as potential members of the adult society and as workers capable of putting energy into the education endeavor. The teacher is the “manager of the conditions of learning” (Gagne, 1970, p. 324) and teachers plan curriculum for average students, with adjustments for some students. Success in this curriculum model is defined as achievement of the learning objectives, which could include displaying democratic behaviors.

Learner-centred ideology is based on the concept of an ideal school, or what Dewey (1916) called “the school of tomorrow”. Current models which typify this ideology include Montessori preschools and Reggio Emilia schools (OECD, 2004). In learner-centred ideology, the needs of the individual child dominate, rather than the needs of teachers, school subjects, school

administrators or community groups. Within this ideology, a curriculum is focused on children having hands-on experiences and therefore involves children being actively involved in democratic practices themselves (MacNaughton, 2007; Myskiw, 2019). Activity is a key term—physical, social, verbal, and emotional (Schiro, 2013)—and takes place both inside and outside the classroom. The history of these ideas can be seen in the ideas of several key thinkers: Comenius (1592–1670), who argued that learning moves from concrete experience to abstract thought; Jean-Jacques Rousseau (1712–78), who believed that the purpose of education was to nurture children’s innate goodness and that learning should result from direct experiences; Johann Heinrich Pestalozzi (1746–1827), who argued that children should be free to pursue their own interests and advocated for the freedom of the “whole child”; and Friedrich Froebel (1782–1852), who invented the kindergarten, which was based on the education of children’s senses and the use of “gifts” and “occupations” that allowed children to perceive the order and beauty of nature as a result of interacting with it. Also associated with this ideology are G. Stanley Hall (1844–1924), the founder of the child study movement, in addition to John Dewey (1859–1952), who argued that “children learn by doing”.

A curriculum based on these ideas uses an integrated approach, based on principles of educating of the “whole child”, which integrates children’s home and school lives. It uses interdisciplinary approaches to academic disciplines and assumes that children will actively integrate their own knowledge in their own way and that there are multiple pathways to learning. The learner-centred ideology is humanist in orientation—the fundamental principles are about human growth, the potential for self-actualization, and the need for personal autonomy (Cleverly and Phillips, 1995). Learning is seen to occur through the child’s interaction with the environment and the dominant learning theory used is constructivism, often using Jean Piaget’s theory (see Bredekamp, 1987). Piaget’s (1936) theory proposed that children are active learners, and their interaction with the physical world provides the main constraints on and contributions to the development of their intelligence: children learn and construct meaning as they act upon objects in space and time.

More recently, social constructivism, based on Lev Vygotsky’s (1978) theories, has also been expressed within this ideology (see Copple and Bredekamp, 2009). Project approaches to learning are typical in this ideology—teachers help children scope and sequence learning and enable children to move from personal, concrete, and physical learning experiences to abstract, verbal and intellectual understandings (Schiro, 2013). In this ideology the teacher is not a transmitter, but carefully observes the children and sets up environments to support learning, including notions of democracy. The teacher may also intervene between the child and the environment to assist learning. Teachers working within this learner-centred ideology will develop curriculum in ways which integrate knowledge with children’s growing sense of themselves as democratic members of their society. Te Whāriki (Ministry of Education, 1996, 2017) and the Early Years Learning Framework (DEEWR, 2009) are good examples of this type of curriculum, which promotes the notion of preparing children to become good citizens of the early childhood community and wider society.

The *social reconstruction ideology* is premised on the idea that our society is threatened by many problems and that education is the cure. Giroux’s (2006) writing typifies this approach to curriculum:

Educators need to assume the role of leaders in the struggle for social and economic justice ... Educators must connect what they teach and write to the dynamics of public life ... and concern for ... democracy (p. 9).

According to Giroux, democracy is grounded in the notion of the citizen as an active and critical agent supporting principles of “sociality and community” (1989, p. 34). Giroux’s definition of democracy suggests he is a proponent of “idealist” democratic theory, which is concerned with both specific processes of governance, but also with producing certain types of public-spirited citizens (Arthur and Sawyer, 2009). From this standpoint on democracy itself, Giroux then argues that the aim of education in a democracy is equity and social justice, creating relations with communities and improving the “quality of public life” (1989, p. 19). Education for Giroux is therefore both a process and a product in itself: education is a moral, civic, public, and social enterprise aimed at creating the kind of citizen who might fit well into the idealist conception of democracy. One key strategy is linking students’ in-school experiences to those of their wider communities—in Giroux’s view, thus enabling a “voice” (1989, p. 200) and a sense of critical agency; he argues that schools are best understood as a locus of citizenship (Giroux, 1989, p. 201). Certain pedagogies permit agency to manifest itself by allowing students to use their own social and cultural capital (Arthur and Sawyer, 2009); thus, enabling students to feel like they belong, they are active agents in their own learning, able to make choices and co-construct meanings and identities (Dahlberg et al., 2007). These notions are increasingly reflected in curriculum for compulsory schooling and early childhood education.

Critical pedagogy, which the philosophy of education promoted by critical theorists such as Giroux (1989), Freire (1972), Apple (1979, 1982), and Foucault (1977) is underpinned by the belief that schools and curriculum are designed to legitimate some ideas and suppress others. Critical pedagogy is thus designed to disrupt and undermine the conventional forms of understanding, which serve to reproduce undemocratic, racist, sexist and unequal social relations. Implicit in this ideology are ideas of student-centredness, student empowerment and emancipatory citizenship as means of subverting the traditional social reproduction of schools, power relationships, hidden curriculum and inequalities produced by curriculum.

In relation to young children in this world view, ECE settings are often children’s first experiences with the wider community and are increasingly being recognized as a place where they can experience democracy and “citizenship” (MacNaughton, 2007). MacNaughton has demonstrated that young children have “powerful, insightful, valid and valuable” views on a range of issues relevant

to their lives (MacNaughton, 2007, p. 460). Accordingly, children's interests in issues such as the environment and social justice are starting to be included in curricula (Edwards and Cutter McKenzie, 2011) and in research about how children learn to become agentic citizens through their engagement in community projects, such as Rameka's research in New Zealand about culturally sustainable practices in ECE settings (Paul-Burke and Rameka, 2015). Teaching and learning in this ideology in ECE are based on the physical and social context and what the issues are in their community.

The principles of teaching children how to be good, democratic citizens is seen most clearly in the Nordic countries (Jensen, 2009). For instance, in the Swedish early childhood curriculum children are not just exposed to some of these big ideas, such as human rights; they are actively engaged as learners with issues—for example, one of the curriculum goals for children is “an understanding of democratic principles and the ability to cooperate and make decisions in accordance with them” (Swedish National Agency for Education, 2019, p. 17). Learning is based on experiences that intellectually and emotionally involve the children. Group discussion is a key instructional method, and values, feelings, injustices, alternatives and solutions are explored. The teacher is seen as a colleague, who has experiences to share with the children, rather than as an expert. Centers are thus seen as institutions of change, in which children are prepared to change society for the better. Children are not seen as children, but as potential contributing members of society who can aid reconstruction (Schiro, 2013); clearly demonstrating notions of social reconstruction ideology. Sweden's national curriculum is aimed to make day-care “take care of and support children in developing their ability to feel responsibility and social readiness to act in order to establish solidarity and tolerance” (Swedish National Agency for Education, 2006, p. 3). Furthermore, the objective is to “encourage and strengthen child's sympathy and empathy for other the situation of others” (Swedish National Agency for Education, 2006, p. 3) and “no child should be subjected to discrimination at pre-school based on gender, ethnic belonging, religion or perception of other belief, or sexual orientation” (pp. 3–4). An important point is that ECE sets the grounds for lifelong learning and that ECE must present possibilities for children's play, concentration and interplay between child and surroundings in order to develop children's societal understanding and understanding for mutual democratic values. Issues of social inequality and differences between children is addressed implicitly by writing that no child should be discriminated against.

Einarsdóttir et al. (2015) argue that a high priority is given to the values of social inclusion within the Nordic model, in which every citizen should have equal opportunities to participate in society (Kristjánsson, 2006). The Nordic ideal of child-centeredness specifically focusses on democracy as an outcome of participation in ECE, as it includes key notions of the naturalness of childhood, equality and egalitarianism, democracy, freedom, emancipation, warm and cooperative social relationships, and solidarity with Nordicness (Wagner and Einarsdóttir, 2008; Wagner, 2006). Nordic countries often also oppose introducing standards for learning too early in children's lives because of the risk that such standards may limit the children's free development (Jensen, 2009). Notions of school readiness are similarly opposed, as it could clash with a Nordic value of the good childhood, which means that ECE have to practice democracy in which children have to be active participants and have a voice in their own everyday life (Committee of the Rights of the Child 1989). For example, in Finland, the national curriculum revolves around cognitive, ethical, social, religious and aesthetic objectives (Ministry of Social Affairs and Health Finland, 2004, p. 24) and the teacher has to work toward these objectives, rather than fixed learning standards. Early childhood is seen as part of the overall objective of lifelong learning, as it is in Sweden. In this model, ECE is part of a seamless whole, with special emphasis on treating children as equals. “Early childhood education in Finland is based on a view of the child as a human being whose growth, development and learning happens in formal and informal ways in different environments” (Niikko, 2006, p. 142).

As this analysis suggests, it is possible to build principles and pedagogies for democracy into early childhood curricula, as the Nordic governments has done, but this reflects the cultural values of the society, as the OECD (2019) proposed. To ensure that children learn how to be democratic, the principle needs to be embedded in the learning goals of the curriculum and then enacted by teachers in the ways in which they involve children in the decision making in each early childhood setting. As Myskw (2019) demonstrates, this can be done within the Australian early childhood curriculum context, but it requires teachers to set up processes for engagement with children about how they want the center to run.

As this section has explored, it is possible to think about curriculum and democracy for young children, but this discussion can't be completed without examining notions of children's rights in relation to the delivery of curriculum. These issues will be explored next.

Children's rights: part of democratic practice?

One of the most interesting aspects of the debates around children and curriculum concerns the United Nations Convention on the Rights of the Child (UNCROC) (United Nations, 1989). De Castro (2012) argues that the situation of children changed in the late decades of the twentieth century once they became the objects of legislation of specific rights via the UNCROC agreement. The UNCROC agreement was supported by all world state nations, except the United States and Somalia. From this followed national statutes and bills which could secure and enhance the international legislation, creating a new subject position for children as “subjects of rights”. It is important to note that the granting of children's rights have not been the result of children's own dispute over their social conditions and therefore were not sustained by children's actions to assert their rights to certain social conditions. As de Castro (2012) points out, adults spoke in children's “best interests” and therefore it was unclear what relationships and acts of oppression were creating challenges for children. The moral dignity of children came hand in hand with the legal need to be protected, cared for, and educated by adults. The upgrading of children's positions to that of upholders of rights has enhanced their

visibility as objects of public policy, rather than as actors with agency in their own society as Giroux (1989) proposed. Paradoxably, although the UNCROC agreement makes children more visible to the public and a matter for public policy and investment, it equally positions children as vulnerable, lacking capacity and in need of adult protection. The right to participate, according to the agreement should only be “given” to children on the condition that they demonstrate beforehand their cognitive and linguistic competence to participate (De Castro, 2012). This position on children’s competence is so vague, that it engenders notions of protecting children, rather than seeing them as holding legitimate opinions on what their lives should be like.

De Castro (2012) argues that the scientific study of children in different disciplinary fields has largely treated children and politics as two distinct and unrelated domains. Furthermore, when children have been considered in relation to politics, she argues that the discussion avoids any consideration of children as political actors with agency, as Giroux (1989) argued. Instead, “child politics” refers to the gradual appearance of children as objects of public policies, culminating in the ratification of UNCROC by almost all nations. Currently, 196 countries are party to it, including every member of the United Nations, except the United States. De Castro concludes that because of the different political regimes, different levels of economic development and different interests by nations in international conventions and social issues, the ways this convention is enacted by countries varies considerably. It has also meant that the focus is broadly on protection of children and their rights, rather than a focus on children’s significance and contribution to political life, democracy or indeed to society. This outcome means that children are perceived as in need of protection, rather than also being actors in their own lives.

Taylor (2009) has similarly argued that children are viewed as an object of adults’ concerns, as projects or persons in the making, and as objects of society’s cost/benefit calculation of how much investment they should receive. In either perspective, as Taylor argues, children not perceived as agentic or a participating member of a democracy; they are simply an adult concern. Many countries enactment of the UN Convention endorse this approach, suggesting that the focus is on ensuring children simply have access to health and education, rather than democratic opportunities to offer opinions on their lived experiences. De Castro (2012) argues that this positioning of children as objects needing attention by adults needs re-evaluation. Viewed as “rights holders” a more participative role for children in society should be possible. There is a taken for granted aspect to children as subjects in need of protection that is clearly challenged in some national curriculum, particularly in Nordic countries (e.g. Einarsdottir, 2008; Jensen, 2009), and this is an issue that needs re-evaluation in many countries which are signatory to the UNCROC agreement. Although designed to ensure that children have rights to health and to education, the right for children to have a say in their own lives is a matter which is less clear in most curriculum (Jensen, 2009; Einarsdottir and Wagner, 2006). The reference to the United Nations Convention on the Rights of the Child (UNCROC) in the revised New Zealand early childhood curriculum, *Te Whāriki* (Ministry of Education, 2017) also signals the obligation on teachers to support children’s rights to health and education. The curriculum states: “Kaiaako (educators) work to uphold and protect children’s rights, interests, and points of view from the earliest ages” (p. 61). Notably, although New Zealand was signatory to the UNCROC agreement in 1993, there was no reference to it in the 1996 version of the curriculum. This change suggests a greater consciousness of the need to recognize that children can and should have a “voice” (Giroux, 1989) in early childhood settings.

Grindheim (2017) argues that the move away from an adult-centred view of citizenship allows for recognizing different forms of participation in both formal and informal settings. She argues that research has delineated how an adult-centred understanding of where and how democratic participation is to be exercised may not accord with how children participate, inhabit, negotiate, and challenge the given conditions for participation. She proposes that educators need to explore the notion of “citizenship as practice” (Lister, 2008), which is understood as the way children exercise their rights within ECE settings and is conceptualized as different ways of participating within their given context. Grindheim (2017) further argues that play is understood as activities that frame ways of participating, which are culturally framed and have included limited opportunities for freedom. Nevertheless, children as “child-citizens” do challenge, change and interrupt the given conditions, using a variety of participation forms. She suggests that through play, children can learn about citizenship and democratic participation.

Sousa and Oxley (2019) found that democracy is elastic, flexible and distinctive within specific educational contexts, rather than a self-evident and universal consensual truth. They argue democracy needs critical discussion and analysis in order to “survive” within the contexts of ECE practice; it cannot be taken for granted. Each of the three ECE settings in Portugal that Sousa and Oxley examined (one public, one private, one religious not for profit) had a distinctive range of representations of democracy, with foci varying from the individual to the group. Rather than one version of democracy, Sousa and Oxley found that democracy was represented differently in each of the three ECE settings, presenting new forms of active and critical democratic citizenship with the potential to contribute to the reduction of social inequality.

Myskiw (2019), who designed her early childhood curriculum in Australia on democratic principles to build children’s sense of agency, argues that a key feature is making sure that children’s voices can be heard. Although she acknowledges that the United Nations Convention on the Rights of the Child (United Nations, 1989) states that “the views of the child ... (shall be) given due weight in accordance with the age and maturity of the child” (article 12, S1), she argues that the question of “age and maturity” for early childhood practitioners is more about how a child’s voice should be heard, not whether they shall be heard. As she states, children’s voices and views about matters affecting them, no matter how large (consent, privacy) or small (the size of vegetables in their meals) should be a considered and supported democratic and ethical process. As she states:

It is radical thinking to set up a forum for debate that promotes the discussion of rights; it is even more radical to respect, listen to, and enact the outcome of the discussion. That’s when we enter a real dialogue and practice of enacting children’s rights in early childhood settings. Sometimes the

meetings go for a long time and sometimes they fizzle out after five minutes, but it is important to honour the process and turn up every time, because the children's image of us as leaders is now changing too" (p. 16).

Within this setting a "Children's Council" was established in which children and teachers negotiated the matters that concerned them. Myskiw (2019, p. 16) states that the following features of democracy are crucial to the success of such a venture:

- Strong relationships with children embedded in to practice
- Whole community approach
- Building a foundation of rights
- Creating a movement: lobbying and children as active citizens
- Image of the child
- Pedagogy of listening
- Rights education

Fleer (2003) similarly proposed that the way forward for ECE is to build a new "community of practice", using Wenger's (1998) terminology, in which teachers use a range of cultural tools to work with children and their families to build on the social and cultural strengths that they bring to their learning. One such example is the Philosophy for Children (P4C) movement, which aims to give voice to childhood thinking through philosophical practice. The curriculum is based on the construction of Communities of Inquiry (CoI), which promote reasoning and argumentative skills, foster creative imagination and develop caring dispositions (Di Masi and Santi, 2016). Founded by Professor Matthew Lipman in the 1970s in New York, P4C comprises a curriculum that emphasizes the formative dimension of philosophy with a CoI, which promotes "an approach that puts the child at the center and emphasizes learning by discovery and experiment, and the construction of knowledge" (Vansieleghem, 2005, p. 19). It is a general aim of Philosophy for Children to include the voices of every member of the community on the grounds that the more voices are heard, the greater will be the possibilities of reaching a general and appropriately representative consensus. Philosophy for Children is based on the proposition that critical thinking and dialogue, based on Bakhtin's (1981) notions of "voice" and Wenger's (1998) notions of "communities of practice" are the necessary conditions for emancipating children from determination and for transforming them into democratic, free citizens (Vansieleghem, 2005, p. 20); dialogues based on child participation in shared reasoning and argument evaluation critical autonomous thinking and prevents indoctrination according to this curriculum approach. Using a dialogic approach is argued to create the opportunities for children to become democratic citizens.

Summary and conclusions

An exploration of the concepts of curriculum and democracy have been explored in relation to the types of curriculum that have been used over time to support young children's learning. Using Scott's (2008) useful definitions of curriculum and Schiro's (2013) explanations of the key ideologies underpinning approaches to curriculum, this chapter has examined the different types of curriculum and whether they are fit for purpose for supporting children to become democratic citizens of a society. As this exploration showed, the only curriculum that are specifically designed to support children to develop a sense of agency within a democracy are based on social efficiency ideologies, in which the goals for children to develop their own "voice" and to actively participate in democratic processes are clearly outlined. These types of goals are more common in the Nordic countries (Jensen, 2009) but are also emerging in revisions of other curricula, such as in New Zealand (e.g. Ministry of Education, 2017). This chapter has also explored the tensions between the positioning of children in the United Nations Conventions on the Rights of the Child (1989) and curriculum, as children are typically positioned as in need of protection by more learned others, rather than being agentic in their own right. Arguably, many nations need to reflect more fully on what children's rights to health and education mean in relation to the early childhood curriculum. There is clearly a need for teachers, providers of initial teacher education and policy makers to reflect more fully on what it means to support children to become members of a democracy.

References

- Adams, P., Bethell, K., 2005. Shifting conceptions of education and educational institutions. In: Adams, P., Vossler, K., Scrivens, C. (Eds.), *Teachers Work in Aotearoa New Zealand*. Thomson/Dunmore Press, Southbank, Vic, pp. 132–148.
- Apple, M., 1982. *Education and Power*. Routledge, New York, NY.
- Apple, M., 1979. *Ideology and the Curriculum*. Routledge & Kegan Paul, Boston, MA.
- Arthur, L., Sawyer, W., 2009. Robust hope, democracy and early childhood education. *Early Years* 29 (2), 163–175. <https://doi.org/10.1080/09575140802628776>.
- Aubrey, C., 2004. Implementing the foundation stage in reception classes. *Br. Educ. Res. J.* 30 (5), 633–656.
- Bakhtin, M.M., 1981. *The Dialogic Imagination: Four Essays*. University of Texas Press, Austin, TX.
- Bernstein, B., 1996. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Taylor & Francis, London.
- Bernstein, B., 1977. *Class, Codes and Control*. Routledge & Kegan Paul, London.
- Biesta, G., 2011. The ignorant citizen: Mouffe, Rancière, and the subject of democratic education. *Stud. Philos. Educ.* 30, 141–153. <https://doi.org/10.1007/s11217-011-9220-4>.
- Bobbitt, F., 1918. *The Curriculum*. Houghton Mifflin, Boston, MA.

- Bredenkamp, S., 1987. Developmentally Appropriate Practice in Early Childhood Education Programs: Serving Children from Birth through Aged 8. National Association for the Education of Young Children (NAEYC), Washington DC.
- Bubikova-Moan, J., Hjetland, H.N., Wollscheid, S., 2019. ECE teachers' views on play-based learning: a systematic review. *Eur. Early Child. Educ. Res. J.* 27 (6), 776–800. <https://doi.org/10.1080/1350293X.2019.1678717>.
- Carlson, M., Earls, F., Foncsea, C., 2002. Children and Democracy. *ReVista*, Harvard Review of Latin America. Retrieved from. <https://revista.drclis.harvard.edu/book/children-and-democracy>.
- Clevery, J., Phillips, D.C., 1995. *Visions of Childhood: Influential Models from Locke to Spock*. Teachers College Press, New York, NY.
- Cohen, C., 1971. *Democracy*. University of Georgia Press, Athens, GA.
- Copple, C., Bredenkamp, S., 2009. Developmentally Appropriate Practice in Early Childhood Programs Serving Children from Birth through Age 8. NAEYC, Washington, DC. <https://www.naeyc.org/resources/pubs/books/developmentally-appropriate-practice-earlychildhood-programs-serving-children>.
- Dahlberg, G., Moss, P., Pence, A., 2007. *Beyond Quality in Early Childhood Education and Care: Languages of Evaluation*, second ed. Falmer, London.
- De Castro, L.R., 2012. Children— democracy and emancipation. *Altern. Glob. Local. Polit.* 37 (2), 165–177.
- DEEWR, 2009. *Belonging, Being and Becoming. The Early Years Learning Framework for Australia*. Commonwealth of Australia, Canberra, Australia.
- Department of Education, 1988. *Education to be More*. Government Print, Wellington, New Zealand.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan, New York, NY.
- Di Masi, D., Santi, M., 2016. Learning democratic thinking: a curriculum to philosophy for children as citizens. *J. Curric. Stud.* 48 (1), 136–150. <https://doi.org/10.1080/00220272.2015.1088064>.
- Edwards, S., Cutter-Mackenzie, A., 2011. Environmentalising early childhood education curriculum through pedagogies of play. *Aust. J. Early Child.* 36 (1), 51–59.
- Einarsdottir, J., Puroila, A., Johansson, E., Broström, S., Emilson, A., 2015. Democracy, caring and competence: values perspectives in ECEC curricula in the Nordic countries. *Int. J. Early Years Educ.* 23 (1), 97–114. <https://doi.org/10.1080/09669760.2014.970521>.
- Einarsdottir, J., Wagner, J.T., 2006. *Nordic Childhoods and Early Education: Philosophy, Research, Policy and Practice in Denmark, Finland, Iceland, Norway, and Sweden*. Information Age Publishing, Greenwich CT.
- Einarsdottir, J., 2008. Children's and parents' perspectives on the purposes of playschool in Iceland. *Int. J. Educ. Res.* 47 (5), 265–269.
- Fleer, M., 2003. Early childhood education as an evolving "community of practice" or as lived "social reproduction": researching the "taken for granted". *Contemp. Issues Early Child.* 4 (1), 64–79.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison*. Vintage, New York, NY.
- Freire, P., 1972. *Pedagogy of the Oppressed*. Penguin Books, Harmondsworth, Middlesex.
- Gagne, R.M., 1970. *The Conditions of Learning*, second ed. Holt, Rinehart and Winston, New York, NY.
- Giroux, H., 2006. *America on the Edge*. Palgrave Macmillan, New York, NY.
- Giroux, H., 1989. *Schooling for Democracy: Critical Pedagogy in the Modern Age*. Routledge, London.
- Grindheim, L.T., 2017. Children as playing citizens. *Eur. Early Child. Educ. Res. J.* 25 (4), 624–636. <https://doi.org/10.1080/1350293X.2017.1331076>.
- Heckman, J.J., 2012. Invest in Early Childhood Development: Reduce Deficits, Strengthen the Economy. <https://heckmanequation.org/>.
- Heckman, J., Masterov, D., 2007. The productivity argument for investing in young children. *Appl. Econ. Perspect. Pol.* 29 (3), 446–493. <https://doi.org/10.1111/j.1467-9353.2007.00359.x>.
- Held, D., 2006. *Models of Democracy*, third ed. Polity Press, Cambridge.
- Holland, J.G., 1960. Teaching machines: an application of principles from the laboratory. *J. Exp. Anal. Behav.* 3, 275–287.
- Jensen, B., 2009. A Nordic approach to Early Childhood Education (ECE) and socially endangered children. *Eur. Early Child. Educ. Res. J.* 17 (1), 7–21. <https://doi.org/10.1080/13502930802688980>.
- Kristjánsson, B., 2006. The making of Nordic childhoods. In: Einarsdottir, J., Wagner, J.T. (Eds.), *Nordic Childhoods and Early Education*. Information Age Publishing, Greenwich, CT, pp. 13–42.
- Lister, R., 2008. Unpacking children's citizenship. In: Invernizzi, A., Williams, J. (Eds.), *Children and Citizenship*. Sage Publications Ltd, London, pp. 10–19.
- Litjens, I., Taguma, M., 2010. Revised Literature Overview for the 7th Meeting of the Network on Early Childhood Education and Care, Official Document EDU/EDPC/ECEC(2010)3/REV1. OECD Publishing, Paris. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/EDPC/ECEC\(2010\)3/REV1&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/EDPC/ECEC(2010)3/REV1&doclanguage=en).
- MacNaughton, G., 2007. Young children's rights and public policy: practices and possibilities for citizenship in the early years. *Child. Soc.* 21, 458–469.
- Ministry of Education, 1996. Te Whāriki. He whāriki matauranga mo nga mokop-una o Aotearoa: Early Childhood Curriculum. Learning Media, Wellington. Retrieved from. <http://www.educate.ec.govt.nz/learning/curriculumAndLearning/TeWhariki.aspxm>.
- Ministry of Education, 2017. Te Whāriki. He whāriki matauranga mo nga mokop-una o Aotearoa: Early Childhood Curriculum. Retrieved from. <https://tewhariki.tki.org.nz/>.
- Ministry of Social Affairs and Health, Finland (Sosiaali- ja terveystieteidenministeriö), 2004. National Curriculum Guidelines on Early Childhood Education and Care in Finland. Stakes, Helsinki.
- Moss, P., 2011. Democracy as first practice in early childhood education and care. In: Bennett, J. (Ed.), *Child Care—Early Childhood Education and Care—Encyclopedia on Early Childhood Development*. Retrieved from. <http://www.childencyclopedia.com/child-care/early-childhood-education-and-care/according-experts/democracy-first-practice-early>.
- Munns, G., Martin, A., 2005. It's all about MeE: a motivation and engagement framework. In: Paper Presented at the Annual Conference of the Australian Association for Research in Education. Sydney, December. <http://www.aare.edu.au/05pap/mun05400.pdf>.
- Myskiw, S., 2019. "Take it to the council": democracy in action in early childhood settings. *Victoria. Connect Guardian Childcare Educ.* 240, 13–16.
- Niikko, A., 2006. Finnish daycare: caring, education, and instruction. In: Einarsdottir, J., Wagner, J.T. (Eds.), *Nordic Childhoods and Early Education – Philosophy, Research, Policy, and Practice in Denmark, Finland, Iceland, Norway, and Sweden*. Information Age Publishing, Greenwich, CT, pp. 133–159.
- Organisation for Economic Co-operation and Development (OECD), 2004. *Starting Strong: Curricula and Pedagogies in Early Childhood Education and Care*. OECD, Paris.
- Organisation for Economic Co-operation and Development (OECD), 2006. *Starting Strong II. Early Childhood Education and Care*. Organisation for Economic Co-operation and Development (OECD), Paris.
- Organisation for Economic Co-operation and Development (OECD), 2018. *Engaging Young Children: Lessons from Research about Quality in Early Childhood Education and Care, Starting Strong*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264085145-en>.
- Organisation for Economic Co-operation and Development (OECD), 2019. Curriculum Alignment and Progression between Early Childhood Education and Care and Primary School: A Brief Review and Case Studies. OECD Education Working Paper No. 193. <http://www.oecd.org/edu/workingpapers>.
- Piaget, J., 1936. *Origins of Intelligence in the Child*. Routledge & Kegan Paul, London.
- Paul-Burke, K., Rameka, L., 2015. Kaitiakitanga—active guardianship, responsibilities and relationships with the world: towards a bio-cultural future in early childhood education. In: Peters, M. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore. https://doi.org/10.1007/978-981-287-532-7_54-1.
- Schiro, M.S., 2013. *Curriculum Theory: Conflicting Visions and Enduring Concerns*, second ed. Sage, Los Angeles, CA.
- Scott, D., 2008. *Critical Essays on Major Curriculum Theorists*. Routledge, London.
- Sousa, D., Oxley, L., 2019. Democracy and early childhood: diverse representations of democratic education in post-dictatorship Portugal. *Educ. Rev.* 1–17. <https://doi.org/10.1080/00131911.2018.1556609>.
- Standards and Testing Agency, 2016. National Curriculum Assessments: Early Years Foundation Stage. Retrieved from. <https://www.gov.uk/government/collections/national-curriculum-assessments-early-years-foundation-stage>.
- Swedish National Agency for Education, 2006. Curriculum for the preschool: Lpfö 98. Fritzes, Stockholm.
- Swedish National Agency for Education, 2019. Curriculum for the preschool: Lpfö 98. Fritzes, Stockholm.

- Sylva, K., Pastori, G., Lerkkanen, M., Ereky-Stevens, K., Slot, P., 2016. D2.4: integrative report on a culture sensitive quality and curriculum framework. In: Curriculum Quality Analysis and Impact Review of European ECEC. Retrieved from: https://ecec-care.org/fileadmin/careproject/Publications/reports/D2_4_Integrative_Report_wp2_FINAL.pdf.
- Taylor, R., 2009. Children as projects and persons: a liberal antinomy. *Soc. Theor. Pract.* 35 (4), 555–576.
- Tyler, R., 1949. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, Chicago.
- United Nations, 1989. *United Nations Convention on the Rights of the Child*. United Nations, Geneva.
- Vansieleghem, N., 2005. Philosophy for children as the wind of thinking. *J. Philos. Educ.* 39, 19–35.
- Vygotsky, L., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge MA.
- Wagner, J.T., 2006. An outsider's perspective: childhoods and early education in the nordic countries. In: Einarsdóttir, J., Wagner, J.T. (Eds.), *Nordic Childhoods and Early Education*. Information Age Publishing, Greenwich, CT, pp. 280–306.
- Wagner, J.T., Einarsdóttir, J., 2008. The good childhood: Nordic ideals and educational practice. *Int. J. Educ. Res.* 47 (5), 265–269. <https://doi.org/10.1016/j.ijer.2008.12.005>.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, Cambridge.

Curriculum and LGBTQ+ inclusion

Steven Camicia, School of Teacher Education and Leadership, Utah State University, Logan, UT, United States

© 2023 Elsevier Ltd. All rights reserved.

Power relations, curriculum, and recognition	435
LGBTQ+ inclusive curriculum	435
Critical democratic education and LGBTQ+ inclusive curriculum	436
Greeting	436
Rhetoric and discourse	436
Narrative and counternarrative	437
Conclusion	438
References	438

The school curriculum communicates what we choose to remember about our past, what we believe about the present, what we hope for the future.
[Pinar \(2004, p. 20\).](#)

"Queer," to me, refers to a politics that values the ways in which meanings and institutions can be at loose ends with each other, crossing all kinds of boundaries rather than reinforcing them. What if the most productive junctures weren't the ones where everything means the same thing?
[Sedgwick \(2011, p. 200\).](#)

The curriculum communicates many things about our past, present, and future. If LGBTQ+ people and issues are not recognized in the curriculum, students are left with an inaccurate understanding of their community's past, present, and future. Recognition in the curriculum signifies the individuals, groups, social issues, and events that are of value to the community. Conversely, a lack of recognition within the curriculum signifies what individuals, groups, social issues, and events are of little value to the community or worse. When a dominant group is presented as the norm, this presents a picture of a standard in which all other groups are judged. It is as if this dominant understanding is fully human, and marginalized groups as lacking or less than human. Throughout history, the erasure and mistreatment of LGBTQ+ individuals have illustrated how dominant groups treat them as less valuable or less than human. Dehumanization is a tool used by dominant groups to justify mistreatment and violence against LGBTQ+ individuals and groups. This highlights inequitable power relations because a lack of recognition is framed by dominant groups, so absences and misrepresentations bolster inequalities. Because of omission from the curriculum, the marginalization of LGBTQ+ individuals and issues continues, and dominant sexualities and gender identities are reinforced.

As matters of equity, democracy, and social justice, inclusion is a necessity, and student inquiry into social inequalities needs to be a central theme that guides the curriculum.

This chapter presents a summary of my research and ways to conceptualize, design, and implement LGBTQ+-inclusive curriculum within the context of critical democratic education. The curriculum sends powerful messages to students and society about what experiences, knowledge, and identities are valued by communities. I first examine curriculum as an indicator of power relations between individuals and groups within society. Second, I present a heuristic for understanding how to increase LGBTQ+ inclusion through situated knowledge. Next, I connect the first and second parts of the chapter through the concept of critical education for democracy. The legitimacy of democratic communities relies upon the recognition of the voices and perspectives of individuals and groups within the community. This applies to LGBTQ+-inclusive curriculum in democratic communities as perspectives in the curriculum more accurately reflect the range of perspectives within a community.

With the exception of a few states ([GLSEN, 2019](#)), LGBTQ+ individuals, histories, and issues have not been recognized in the curriculum throughout the United States. State curriculum standards are exclusionary ([Camicia and Zhu, 2019](#)) and driven by anti-gay curriculum laws ([Rosky, 2017](#)). This lack of recognition is harmful to LGBTQ+ students who do not see themselves as represented or valued in the curriculum ([Burdge et al., 2012, 2013](#); [Kosciw et al., 2012, 2014](#); [Mayo, 2014](#); [Snapp et al., 2015](#)).

A lack of LGBTQ+ recognition and inclusion in the curriculum also has implications for how students learn about democratic governance ([Camicia, 2016](#)). Since democratic governance relies upon the voices and perspectives of the governed, a lack of recognition and inclusion of LGBTQ+ individuals and issues sends messages to students and the community about the relationship between democracy and inclusion. Students might believe that they are participating in democratic communities while the exclusion of LGBTQ+ individuals and issues undermines democratic governance. In other words, students are presented with competing messages in the curriculum. One message extols the importance of democracy while another message suppresses the inclusion of LGBTQ+ individuals in the curriculum. A democracy with limited inclusion is a weak form of democracy ([Camicia, 2009](#)). This chapter examines the relationships between curriculum, LGBTQ+ inclusion, and democratic communities.

Power relations, curriculum, and recognition

The curriculum is not designed and enacted in a decontextualized vacuum. Instead, the curriculum reflects the interests of dominant groups and norms (Apple, 2018). As such, the curriculum often fails to reflect the interests of marginalized and minoritized groups (Au, 2012; Au et al., 2016). Numerous studies have illustrated how textbooks (Brown and Brown, 2010; Camicia, 2007) and state curriculum standards (Anderson and Metzger, 2011; Busey and Walker, 2017; Eargle, 2016; Journell, 2009; Shear et al., 2015) do not recognize or misrecognize the histories of marginalized groups. Curriculum standards and instructional materials reflect an official stance of the state, so the messages that they send function to reify existing power relations within society. Because LGBTQ+ individuals and issues have been marginalized by society, the curriculum usually reflects this marginalization. In this way, the curriculum is a reflection of inequitable power relations.

With exception of a few states, LGBTQ+ individuals have been missing from the history curriculum across the United States. At the time of their data collection in 2016, Camicia and Zhu (2019) found that only seven states mentioned the terms LGBTQ in their K-12 social studies curriculum standards, and of the seven, only California included LGBTQ individuals and issues in any substantial way. Because social studies education in public schools comprises subject areas such as civics, history, economics, and geography, the absences or misrepresentations of LGBTQ+ individuals and issues in the subject areas are especially harmful to LGBTQ+ individuals and the community. The absences in state standards are an indication that LGBTQ+ curriculum is a null curriculum (Eisner, 2002) because there is a substantial absence of recognition. In addition, when LGBTQ+ individuals are recognized in the curriculum, the curriculum is often overrepresented by white, gay men. Bickford (2017, 2018) found that LGBTQ-based literature for elementary and intermediate level students followed this pattern, and transgender individuals were rarely recognized. This lack of recognition reflects dominant patterns of inequitable power relations within communities. The marginalization of LGBTQ+ individuals in society does not diminish when students enter schools. The marginalization is usually mirrored in the curriculum. As matters of social justice and human rights, the curriculum needs to include LGBTQ+ individuals and challenge inequalities that are perpetuated through the curriculum.

LGBTQ+ inclusive curriculum

Within the context of LGBTQ+ exclusion in the curriculum, educators can design and implement curriculum and instruction that is inclusive. This involves teacher education (Lilienthal et al., 2018; Murray, 2014), supportive administration (Hernandez and Fraynd, 2014; Mayo, 2014), and inclusive curriculum (e.g., Helmer, 2016; Logan et al., 2016; Moorhead and Jimenez, 2021; Sadowski, 2016; Souto-Manning and Lanza, 2019). One of the main strategies for increasing inclusion can be guided by queer theorizing where, as Sedgwick (2011) writes, we create opportunities for “crossing all kinds of boundaries rather than reinforcing them” (p. 200). In the area of curriculum, this means creating opportunities for students to question norms that reify the oppressive boundaries related to genders and sexualities.

In addition to the powerful messages about sexuality and gender that students receive from dominant groups in society, schools are sites where dominant norms are reinforced by the curriculum. Related to gender, pervasive binary constructions of gender are ubiquitous in the curriculum. This is often exacerbated by school activities that group students into categories of boys and girls, social events where binary gender and gender stereotypes are policed, and instructional materials that perpetuate fixed qualities for gender categories.

For the curriculum to be more inclusive, it needs to challenge exclusionary language by expanding perspectives and challenging gender stereotypes and binary gender. For example, rather than referring to boys and girls while communicating with students, educators can use gender-neutral terms. Souto-Manning and Lanza (2019) provide illustrations from an elementary school where students problematize the labels of boys and girls bathrooms and deconstruct gender stereotypes through the use of trade books. These books help students identify the types of labels that categorize people as a way to police gender norms. Students can also learn about historical and contemporary figures who do not fit within binary gender construction or identify with a gender other than the one that they were assigned at birth. In these ways, the curriculum expands inclusion by opening possibilities for gender expressions.

For the curriculum to be more inclusive of sexualities, it can provide spaces for students to challenge heteronormativity. Dominant norms of heterosexuality serve to exclude consideration and recognition of diverse sexualities. Helmer (2016) describes a high school English classroom where students deconstruct literary texts to identify hidden assumptions and norms related to sexuality. Helmer writes that the teacher “directly challenged her students’ heteronormative reader expectations and a ‘straight’ reading of the text” (p. 40). By challenging heteronormativity in interpretations of literature, the curriculum can increase inclusion. Students can also challenge the stereotypes of minoritized sexualities by identifying stereotypes and providing counternarratives that dismantle the stereotypes. I discuss counternarratives later in this chapter.

Given the proliferation of media that students experience, another area for increasing inclusion is in student deconstruction of media images and texts. Students can ask: What stereotypes of genders and sexualities do the texts communicate? What assumptions underly the messages in media such as commercials? When LGBTQ+ characters are included, what roles do they play, and how are they portrayed? These questions can also be applied to instructional materials across the curriculum in areas such as history, language arts, mathematics, art, and science. Rather than being restricted to literature, students can learn about people from multiple disciplines whose sexualities or genders have been erased in the curriculum.

Students also need to understand that there are multiple forms of oppression related to social inequalities. LGBTQ+ individuals have overlapping identities that position them within society (Moorhead and Jimenez, 2021). Oppressive ideologies and discourses of racism, sexism, classism, ableism, and exclusionary nationalism overlap the identities of LGBTQ+ individuals. By helping students understand these overlapping forms of oppression, the curriculum can increase inclusion.

In the next section, I present ways for students to identify and address exclusion as a way to increase inclusion and make their communities more democratic and socially just. The curriculum is an expression of how we understand the past and present, as well as what we hope for the future (Pinar, 2004). I turn to democratic theory as a way to envision the types of communities that we offer as choices for students. Is it a community based upon democratic inclusion of LGBTQ+ individuals or a community based upon authoritarianism and exclusion?

Critical democratic education and LGBTQ+ inclusive curriculum

One of the ways to increase our understanding of LGBTQ+ inclusion in the curriculum is to conceptualize inclusion as it relates to democracy (Camici, 2016). This provides a strong rationale for increasing the perspectives of LGBTQ+ individuals and social issues. Although an inclusive curriculum is good for the mental, physical, and academic health of LGBTQ+ students (Kosciw et al., 2020), all students and communities are strengthened by increased recognition. This is based upon the assumption that communities value democratic rather than authoritarian governance. If we value democratic governance, then the legitimacy of the curriculum relies upon the inclusion of all who are subject to the curriculum. This means that during student inquiry and discussion, the perspectives of LGBTQ+ individuals are included in the curriculum. This necessarily involves an understanding of how power functions to privilege some voices while marginalizing other voices because LGBTQ+ individuals have been and continue to be silenced and worse within society. Young's (2002) elements of inclusion within democracies provide guidelines that transfer to increasing LGBTQ+ inclusion in the curriculum. These elements are termed greeting, rhetoric, and narrative.

Greeting

Young (2002) described greeting as an ethical agreement to recognize people for whom they are rather than placing them in exclusionary categories. As applied to curriculum, this means agreeing to search for meaning and being open to changing perspectives once another perspective is presented. Under greeting, there is an understanding that we experience the world differently, and those experiences create diverse ways of understanding social issues. A commitment to greeting is a commitment to democratic governance because the views of community members are valued when considering issues of importance to the community. Young wrote:

Greeting, which I shall also call public acknowledgment, names communicative political gestures through which those who have conflicts aim to solve problems, *recognize* others as included in the discussion, especially those with whom they differ in opinion, interest, or social locations. By such Sayings [sic] discussion participants acknowledge that the others they address are part of the process, and that we who address them must be accountable to them, as they to us. (p. 61, emphasis in original)

This recognition is central to LGBTQ+-inclusive curriculum because LGBTQ+ individuals have been and continue to be rendered unrecognizable by dominant groups who make heteronormative, gender, and binary assumptions. As has already been described above, the curriculum can include historical examples of this lack of recognition through student examination of LGBTQ+ individuals and issues that have not been part of the curriculum across subject areas.

In the curriculum, greeting and recognition can increase among students through activities such as identity maps (Moorhead and Jimenez, 2021) and writing about labels and identities (Souto-Manning and Lanza, 2019). By exploring and sharing about identities and lived experiences, students can develop an understanding that some of the dominant categories often render the perspectives and identities of LGBTQ+ individuals unrecognizable. By presenting historical and contemporary examples of heteronormative and gender assumptions, the curriculum can open student conversations about an ethical agreement to recognize each other in the diverse ways that we understand ourselves and our identities. This is a vital component of education for democracy because recognition of diverse sexualities and gender identities can expand perspectives on social issues. A commitment to greeting is a commitment to including the perspectives of those subject to decisions in the decision-making process. As such, LGBTQ+ inclusion is an important component of a healthy democratic community.

Rhetoric and discourse

Young (2002) presented rhetoric as a second element for increasing inclusion in democracy. When discussing social issues in public spaces, norms of conversation and discourses influence what can and can't be said. This relates to the experiences of LGBTQ+ individuals because there is a long history of banning conversations about sexuality and gender. Young wrote;

The ideal of disembodied and disembedded reason that it presupposes is a fiction. What such privileging takes to be neutral, universal, and dispassionate expression actually carries the rhetorical nuances of particular situated social positions and relations, which social conventions do now mark as rhetorical and particular in the same way that they notice others. (p. 63)

The element of rhetoric can focus student inquiry and conversation on the ways that certain ways of talking about social issues are accepted by dominant groups. These groups define the boundaries of what is acceptable to talk about in public. Because LGBTQ+ issues have been regarded by many as inappropriate to discuss in schools (Rosky, 2017), it is important for students to examine the reasons for these prohibitions. Students can examine these *anti-LGBTQ+* laws and policies as texts for their classroom inquiry and discussion. They can ask, “Whose interests do these laws and policies serve?”

While Young (2002) did not explicitly refer to discourse, the connection can be helpful in using concepts of discourse analysis to help students better understand democratic inclusion and rhetoric. Hall (2001) wrote that discourse:

Defines and produces the objects of our knowledge. It governs the way that a topic can be meaningfully talked about and reasoned about. It also influences how ideas are put into practice and used to regulate the conduct of others. Just as a discourse ‘rules in’ certain ways of talking about a topic, defining an acceptable and intelligible way to talk, write, or conduct oneself, so also, by definition, it ‘rules out’, limits and restricts other ways of talking, of conducting ourselves in relation to the topic or constructing knowledge about it. (p. 72)

Dominant discourses in society and schools shape what can and can’t be said about sexualities and genders. This can have a profound impact on the ways that students can understand their communities, LGBTQ+ individuals, and LGBTQ+ issues. The curriculum can be more inclusive by examining the ways that discourses of transphobia, homophobia, and heteronormativity have influenced the degree of inclusion that LGBTQ+ individuals experience in and out of schools. While identifying these discourses, student inquiry can examine the ways that inequitable power relations perpetuate these discourses in the interest of dominant groups. With greeting and rhetoric, students can start understanding how to increase the democratic inclusion of LGBTQ+ individuals and foster democratic discourses.

Narrative and counternarrative

The third element of Young’s (2002) model of making democratic spaces more inclusive involves valuing the lived experiences of individuals and groups. When shared in classroom inquiry and discussion, these experiences can provide the text and examples that students need to understand how knowledge is situated and constrained by inequitable power relations. Young described the role that the element of narrative can play in the curriculum:

By adopting narrative as a pedagogical tool, teachers and students open up possibilities to make links between their experiences and those of strangers in distant places, including links between these strangers’ struggles for justice and their own. Narratives used in this way can be used to advance justice and human rights through education. (p. 49)

By opening the curriculum to a variety of experiences, students can better understand the experiences of LGBTQ+ individuals. These experiences often communicate a counternarrative to the narratives of dominant social groups. Because the curriculum has acted historically to bolster dominant narratives, the experiences of LGBTQ+ people have been excluded from the curriculum. Students can examine how historical narratives from a variety of LGBTQ+ individuals and groups were subject to heteronormative, homophobic, and transphobic discourses (e.g., Bronski, 2019; Stern, 2009). Educators and students can refer to resources such as GLSEN (2021), Lambda Literary (2021), Learning for Justice (2021), and Social Justice Books (2021) to help them build counternarratives and a more accurate portrayal of history and social issues. Counternarratives that challenge oppressive discourses can increase LGBTQ+ inclusion in the curriculum through building the influence of democratic discourses.

In addition to examining the counternarratives of LGBTQ+ individuals in history, students can develop their personal narratives in ways that situate their experiences in larger societal patterns of exclusion and inclusion. Methods of autoethnography can be used by students, educators, and community members to increase LGBTQ+ inclusion in the curriculum. Adams et al. (2017) described this method:

Autoethnography is a research method that uses personal experience (“auto”) to describe and interpret (“graphy”) cultural texts, experiences, beliefs, and practices (“ethno”). Autoethnographers believe that personal experience is infused with political/cultural norms and expectations, and they engage in rigorous self-reflection—typically referred to as “reflexivity”—in order to identify and interrogate the intersections between the self and social life.

This increased awareness of connections between personal experiences and larger political/cultural norms can be accomplished in conjunction with Young’s (2002) elements of greeting and rhetoric. By connecting their narratives and experiences with larger discourses in society, students can work toward a better understanding of LGBTQ+ individuals, experiences, and events. There are multiple ways for students to express their narratives. For example, students can create collages, poems, videos, and essays. These provide the material for them to develop their autoethnographies and understand the multiple narratives and situated knowledge within their classroom. In this way, greeting, rhetoric, and narrative are conceptually and pragmatically linked in ways that can increase LGBTQ+ inclusion in the curriculum.

Conclusion

The marginalization and exclusion of LGBTQ+ individuals, issues, and events from the curriculum is pervasive, as evidenced in curriculum standards and instructional materials. As a reflection of the larger society, the curriculum has reflected the interest of dominant groups while diminishing the recognition of marginalized and minoritized groups. This has had devastating consequences for the physical, mental, and academic health of LGBTQ+ students. The exclusion and marginalization of LGBTQ+ individuals and issues have also functioned to undermine the legitimacy and quality of democracy because democratic governance relies upon the voices and perspectives of the governed in decision-making. This makes inclusion in the curriculum a matter of social justice in many ways. The lessons learned from inclusion can be used to create more democratic communities that work against authoritarianism, whether in the form of governance or in the curriculum.

This chapter has presented ways to increase inclusion by designing and implementing a curriculum that challenges dominant narratives and discourses of transphobia, heteronormativity, and homophobia. Critical education for democracy is a way to increase LGBTQ+ inclusion and work against the social injustices that marginalize LGBTQ+ students (Camicia, 2016; Camicia and Knowles, 2021). This focuses upon the themes building ethical commitments of recognizing each other in our differences and using tools of discourse analysis and counternarrative to express these differences. Student inquiry and discussion can place the context of inequitable power relations as a central consideration for how some voices are valued more than other voices. By interpreting the ways that power functions to exclude LGBTQ+ individuals and issues, students can work to dismantle anti-democratic discourses in the service of building democratic discourses. While students learn about these discourses, they can begin to make choices about the types of communities that they want to build and maintain. Are they communities that are exclusionary or inclusionary? Are they democratic or anti-democratic? These choices indicate their approach to equity and social justice.

References

- Adams, T., Ellis, C., Jones, S.H., 2017. Autoethnography. In: Matthes, J., Davis, C.S., Potter, R.F. (Eds.), *The International Encyclopedia of Communication Research Methods*. John Wiley & Sons, Inc, pp. 1–11. <https://doi.org/10.1002/9781118901731.iecrm0011>.
- Anderson, C.B., Metzger, S.A., 2011. Slavery, the Civil War era, and African American representation in U.S. history: an analysis of four states' academic standards. *Theor. Res. Soc. Educ.* 39 (3), 393–415.
- Apple, M.W., 2018. *Ideology and Curriculum*, fourth ed. RoutledgeFalmer.
- Au, W., Brown, A.L., Calderón, D., 2016. *Reclaiming the Multicultural Roots of U.S. Curriculum: Communities of Color and Official Knowledge in Education*. Teachers College Press.
- Au, W., 2012. *Critical Curriculum Studies: Education, Consciousness and the Politics of Knowing*. Routledge.
- Bickford, J.H., 2017. The representations of LGBTQ themes and individuals in non-fiction young adult literature. *Soc. Stud. Res. Pract.* 12 (2), 182–198. <https://doi.org/10.1108/SSRP-05-2017-0021>.
- Bickford, J.H., 2018. Examining LGBTQ-based literature intended for primary and intermediate elementary students. *Elem. Sch. J.* 118 (3), 409–425.
- Bronski, M., 2019. In: Chavat, R. (Ed.), *A Queer History of the United States for Young People*. Beacon Press.
- Brown, K.D., Brown, A.L., 2010. Silenced memories: an examination of the sociocultural knowledge on race and racial violence in official school curriculum. *Equity Excell. Educ.* 43, 139–154. <https://doi.org/10.1080/10665681003719590>.
- Burdge, H., Sinclair, K., Laub, C., Russell, S.T., 2012. Lessons that Matter: LGBTQ Inclusivity and School Safety. Brief No. 14. https://www.gsanetwork.org/files/aboutus/PSH%20Report%206_2012.pdf.
- Burdge, H., Sinclair, K., Laub, C., Russell, S.T., Moody, R., 2013. *Implementing Lessons that Matter: The Impact of LGBTQ-Inclusive Curriculum on Student Safety, Well-Being, and Achievement*.
- Busey, C.L., Walker, I., 2017. A dream and a bus: black critical patriotism in elementary social studies standards. *Theor. Res. Soc. Educ.* 45 (4), 456–488. <https://doi.org/10.1080/00933104.2017.1320251>.
- Camicia, S.P., Knowles, R.T., 2021. *Education for Democracy: A Renewed Approach to Civic Inquiries for Social Justice*. Information Age Publishing.
- Camicia, S.P., Zhu, J., 2019. LGBTQ inclusion and exclusion in state social studies standards: implications for critical democratic education. *Curric. Teach. Dialog.* 21 (1 & 2), 7–20.
- Camicia, S.P., 2007. Deliberating immigration policy: locating instructional materials within global and multicultural perspectives. *Theor. Res. Soc. Educ.* 35 (1), 96–111.
- Camicia, S.P., 2009. Identifying soft democratic education: uncovering the range of civic and cultural choices in instructional materials. *Soc. Stud.* 100 (3), 136–142.
- Camicia, S.P., 2016. *Critical Democratic Education and LGBTQ-Inclusive Curriculum: Opportunities and Constraints*. Routledge.
- Eargle, J.C., 2016. The dominant narrative of slavery in South Carolina's history standards. *J. Soc. Stud. Res.* 40, 295–307.
- Eisner, E.W., 2002. *The Educational Imagination: On the Design and Evaluation of School Programs*, third ed. Prentice Hall.
- GLSEN, 2019. State-by-State Research. <https://www.glsen.org/research/state-by-state-research-2019-state-snapshots-national-school-climate-survey>.
- GLSEN, 2021. Advocate for Inclusive and Affirming Curriculum. <https://www.glsen.org/inclusive-curriculum>.
- Hall, S., 2001. The spectacle of the other. In: Wetherell, M., Taylor, S., Yates, S.J. (Eds.), *Discourse Theory and Practice: A Reader*. SAGE Publications, pp. 324–344.
- Helmer, K., 2016. Gay and lesbian literature disrupting the heteronormative space of the high school English classroom. *Sex. Educ.* 16 (1), 35–48. <https://doi.org/10.1080/14681811.2015.1042574>.
- Hernandez, F., Fraynd, D.J., 2014. Leadership's role in inclusive LGBTQ-supportive schools. *Theory Pract.* 53, 115–122. <https://doi.org/10.1080/00405841.2014.885811>.
- Journell, W., 2009. An incomplete history: representation of American Indians in state social studies standards. *J. Am. Indian Educ.* 48 (2), 18–32.
- Kosciw, J.G., Greytak, E.A., Bartkiewicz, M.J., Boesen, M.J., Palmer, N.A., 2012. The 2011 National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual and Transgender Youth in Our Nation's Schools. <https://www.glsen.org/sites/default/files/2011%20National%20School%20Climate%20Survey%20Full%20Report.pdf>.
- Kosciw, J.G., Greytak, E.A., Palmer, N.A., Boesen, M.J., 2014. The 2013 National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual and Transgender Youth in Our Nation's Schools. https://www.glsen.org/sites/default/files/2013%20National%20School%20Climate%20Survey%20Full%20Report_0.pdf.
- Kosciw, J.G., Clark, C.M., Truong, N.L., Zongrone, A.D., 2020. National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual, Transgender, and Queer Youth in Our Nation's Schools. GLSEN. https://www.glsen.org/sites/default/files/2021-04/NSCS19-FullReport-032421-Web_0.pdf.
- Lambda Literary, 2021. Resources. <https://www.lambdaliterary.org/resources/>.
- Learning for Justice, 2021. Gender and Sexual Identity. <https://www.learningforjustice.org/topics/gender-sexual-identity>.
- Lilienthal, L.K., Matyo-Cepero, J.A., Messinger, L., Mims, M., 2018. Creating an inviting classroom for all students: inviting teachers to learn about LGBTQ diversity. *Delta Kappa Gamma Bull.* 85 (1), 13–19.

- Logan, S.R., Watson, D.C., Hood, Y., Lasswell, T.A., 2016. Multicultural inclusion of lesbian and gay literature themes in elementary classrooms. *Equity Excell. Educ.* 49 (3). <https://doi.org/10.1080/10665684.2016.1194239>.
- Mayo, C., 2014. *LGBTQ Youth and Education: Policies and Practices*. Teachers College Press.
- Moorhead, L., Jimenez, J., 2021. "This is me": expressions of intersecting identity in an LGBTQ+ ethnic studies course. *J. Soc. Stud. Res.* 45, 35–57. <https://doi.org/10.1016/j.jssr.2020.04.003>.
- Murray, O.J., 2014. *Queer Inclusion in Teacher Education: Bridging Theory, Research, and Practice*. Routledge.
- Pinar, W.F., 2004. *What is Curriculum Theory?* Lawrence Erlbaum Associates, Inc. Publishers.
- Rosky, C., 2017. Anti-gay curriculum laws. *Columbia Law Rev.* 117, 1461–1541.
- Sadowski, M., 2016. *Safe is Not Enough: Better Schools for LGBTQ Students*. Harvard Education Press.
- Sedgwick, E.K., 2011. *Weather in Proust*. Duke University Press.
- Shear, S.B., Knowles, R.T., Soden, G.J., Castro, A.J., 2015. Manifesting destiny: re/presenting of indigenous peoples in K-12 U.S. history standards. *Theor. Res. Soc. Educ.* 43 (1), 68–101.
- Snapp, S.D., Burdge, H., Licon, A.C., Moody, R.L., Russell, S.T., 2015. Students' perspectives on LGBTQ-inclusive curriculum. *Equity Excell. Educ.* 48 (2), 249–265.
- Social Justice Books, 2021. *LGBTQ+*. <https://socialjusticebooks.org/booklists/lgbtq/>.
- Souto-Manning, M., Lanza, A., 2019. Pedagogical third spaces: inclusion and re-presentation of LGBTQ communities in and through teaching as a matter of justice. *Theory Practice* 58, 39–50. <https://doi.org/10.1080/00405841.2018.1536921>.
- Stern, K. (Ed.), 2009. *Queers in History: The Comprehensive Encyclopedia of Historical Gays, Lesbians, Bisexuals, and Transgenders*. BenBella Books, Inc.
- Young, I.M., 2002. *Inclusion and Democracy*. Oxford University Press.

Curriculum and inclusive education: universal design for learning as a “traveling” phenomenon

Christopher Johnstone and Hayley Niad, University of Minnesota Minneapolis, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: universal design for learning	440
The origins of universal design for learning	440
Universal design and access to curriculum	441
Interpreting UDL in practice	442
Policy borrowing and UDL	442
Policy travel, UDL, and education for all	442
National case studies	443
Conclusions	444
References	445

Introduction: universal design for learning

Curriculum is related to inclusive education in two ways. The first intersection between inclusive education and curriculum is in relation to its content. Particular topics, how children are asked to engage with those topics, and the degree to which discriminatory messages are present may all impact the degree to which children feel they belong in school. The second intersection of curriculum and inclusive education relates to access to the curriculum for children with diverse abilities and preferences. No matter what content is presented to students, if it is inaccessible there will be barriers to learning.

This article focuses on the second intersection of curriculum and inclusive education through the lens of accessibility. Specifically, we will provide an overview of Universal Design for Learning (UDL), a curricular approach that has gained global popularity over the past two or three decades. We will then examine Schwanke, Smith, and Edyburn's A3 model as a framework for linking UDL with the principle of accessibility. After tracing the origins of UDL from the United States, we then argue that UDL is undergoing a period of “policy borrowing” as its principles are slowly traveling worldwide and being incorporated into national education systems. We examine this phenomenon through Armenia and Ghana, two disparate nations that have experienced the import of UDL frameworks into their national education policies or curricula. We conclude by reflecting on the limitations of UDL to seamlessly travel across contexts and offer ideas for future enactments of the pedagogical approach.

The origins of universal design for learning

Universal Design for Learning is an educational framework that draws its roots from Universal Design, a set of architectural guidelines originally coined by Ronald Mace from North Carolina State University to support the development of built environments that can be used by all individuals without the need for special adaptation. Universally Designed environments are characterized by seven core principles, among which are equitable, flexible, and intuitive use ([Center for Universal Design, 2008](#)). One example of Universal Design in practice is the construction of ramps as an alternative to stairs. Ramps can support people using wheelchairs, just as they can be accessed by parents pushing strollers, people carrying heavy groceries, or older people walking with canes. In other words, there is no need for special or adapted environments when they are constructed in consideration of all people's needs in the first instance.

In the same light, Universal Design for Learning promotes diverse learning opportunities for all students that respond to their diverse learning needs. The Center for Applied Special Technology (CAST) began applying Universal Design principles to curriculum frameworks in the early 1990s as a way of designing educational approaches that are inherently inclusive of all learners. UDL principles are linked to research in neuroscience, which recognizes that there is variability in the way all people learn, and that learning is context-dependent, such that the same person might learn in different ways depending on the subject and situation ([Meyer et al., 2014](#)). Drawing from three of the networks that guide learning in the brain—(1) affective, (2) recognition, and (3) strategic—UDL as a curriculum framework suggests that learning environments should provide all students with (1) multiple means of engagement, (2) multiple means of representation, and (3) multiple means of action and expression ([Meyer et al., 2014](#)). At the core of UDL is the assertion that learners will benefit from multiple avenues and options to participate in their education.

Just as UDL recognizes the benefits afforded by diversity and variability in learning, so too are the strategies to apply UDL in instruction diverse and variable. The principle of multiple means of engagement emphasizes learner autonomy and choice; authentic and relevant content that respects culture, age, and other identities; and scaffolding to learners to monitor their progress, set goals, and receive feedback. Multiple means of representation requires the provision of content through multiple modalities such as hearing, vision, and touch; the use of multimedia such as charts, diagrams, and illustrations; and other cues to facilitate

learning such as mnemonic devices, checklists and organizers, or concept maps. Finally, learners using multiple means of action and expression should have a choice in how they show their understanding, such as physical movement, speaking, writing, or drawing; and alternative technologies such as adapted keyboards, text-to-speech software, or graphing calculators (CAST, 2018).

As a means of delivering flexibility and variability in instruction, assistive technologies and digital media serve as useful resources in implementing UDL, where available. This includes assistive and augmentative communication that supports learners with and without disabilities, such as those who use computer or tablet apps to express themselves orally, or any number of electronic books, podcasts, presentation software, or games and stories available digitally. There is also a substantial research base demonstrating the utility of assistive technologies for students with learning disabilities, such as word processing, multimedia, and hypertext interventions (Perelmutter et al., 2017). However, many of these technologies require a level of resourcing—such as internet connectivity and reliable electricity—that are not available to all learners across the globe. Nonetheless, UDL has also been used in environments with few material resources, through numerous strategies such as small group work, the use of no-cost manipulatives found in the environment, or games and songs that promote learner engagement (Hayes et al., 2018).

Universal design and access to curriculum

The linkage of a set of pedagogical practices to the democratization of education can be explained through Schwanke, Smith, and Edyburn’s A3 model. The A3 model was used to predict the politics and practice of inclusive education, especially as it relates to children with disabilities. The model itself follows educational development through three phases: advocacy, accommodations, and eventually accessibility. Schwanke et al. (2001) argued that the democratization of knowledge begins when advocacy efforts are present. In the case of children with disabilities, advocacy for greater access to the national or local curriculum has been led by parents of children with disabilities, organizations of persons with disabilities (OPDs), and children themselves. Schwanke et al. (2001) and later Edyburn (2010) observed that an outgrowth of advocacy work is that the needs of children with disabilities are identified in schools and individualized accommodations are made for them. Accommodations are designed to provide access to curriculum, by providing individual students with a degree of flexibility in terms of timing, response, how content is presented, length of assignments, or even the location where particular educational activities take place. For example, a student may receive extended time to complete an assignment, may be allowed to use a scribe if writing is a challenge, may have assignments read to them (by humans or through technology) if decoding print is impossible, or students may be allowed to sit in a quiet space if particular classroom activities are over-stimulating.

The final stage of Schwanke et al.’s (2001) A3 model is “accessibility.” As noted in the paragraphs above, accessibility is a general principle derived from architectural access that has been applied to education. Educational accessibility seeks to provide all students with opportunities to engage with their national curriculum, with minimized needs for stigmatizing accommodations. Within a universal design framework, individualized accommodations do not disappear, but become less noticeable because many of the changes described above become available to all students, not just students with disabilities. Therefore, curriculum and access to it are of central importance in a universal design framework, rather than an educational focus on the tasks associated with curricular learning. Fig. 1 (below) provides an overview of how Schwanke et al.’s model helps to explain the development of UDL.

Accessibility, as the end goal of advocacy work and inclusion through individualized accommodations, is typically synonymous with a UDL approach. For example, according to the US Higher Education Opportunity Act (2008: 102–24):

The term Universal Design for Learning means a scientifically valid framework for learning that (A) provides flexibility in the way information is presented, in the ways students respond to demonstrate knowledge and skills, and in the way students are engaged; (B) reduces barriers in instruction, provides appropriate accommodations supports, and challenges, and maintains high achievement for all students, including students with disabilities.

UDL is not the only way to provide “accessibility” of the curriculum for students but has become a popular technical solution for engaging students within a broader movement in the field of education. As narratives of “all means all” continue to emerge in educational circles (UNESCO, 2020), some scholars have begun to further associate the concept of accessibility as an antidote to

Advocacy	Allies of children with disabilities politically fight for the right to access to an inclusive curriculum.
Accommodations	If advocacy efforts are successful, disabilities may dictate that children need individualized adjustments or accommodations to allow for full participation.
Accessibility	As teachers and fellow students become comfortable with the diversity of engagement with the curriculum, accessibility is built into everyday teaching and learning activities, which may benefit all children.

Fig. 1 Development of UDL and access to curriculum. Adapted from Schwanke et al. (2001).

systemic “ableism”—a form of discrimination inflicted by people without disabilities onto people with disabilities (Hehir, 2007; Powell, 2013). UDL, then, has become an approach that can be championed for curricular and broader societal goals within educational settings.

Interpreting UDL in practice

As UDL’s popularity grew in Northern countries, however, so did emerging debates about how to specifically enact it in schools. For example, organizations like CAST frequently distribute guidance on how UDL can be implemented in classrooms, examples of curricular interventions, and underlying psychological research that validates how accessibility approaches can work to support children (CAST, 2018). Other scholars, however, view UDL as more of an educational philosophy meant to democratize access to the goods of education and critique the use of checklists, toolkits, and how-to manuals. For example, Disability Studies scholar Jay Dolmage (2015) critiques the use of prescribed checklists and encourages educators to engage with UDL critically, examining structures that inhibit student learning, and relying on both teacher reflection and collective idea generation as ways to promote accessibility. And, as the phenomenon of UDL expands globally, others caution that UDL should not be simply equated to the provision of an inclusive education, as additional supports for specific learners may still be needed in a universally designed classroom environment (International Disability Alliance, 2021).

The debates, promises, limitations, and checklists associated with Universal Design have primarily emerged from Northern countries. Schwanke et al.’s (2001) work, for example, is based in the United States, as was the setting for Rose and Meyer’s (2002) groundbreaking conceptualization of UDL. The above-mentioned work by Powell took place in a higher education institution in Europe, and Dolmage’s vision of a critical UDL approach was framed from a Canadian perspective.

The conceptualization of UDL as an outgrowth of architectural concepts also began in the United States, as did original policies that named UDL as an approach to meeting legal demands for accessible education of children with disabilities. For example, Universal Design for Learning (UDL) emerged in the United States as a policy initiative in the late 1990s and early 2000s. In 1997, the Individuals with Disabilities Education Act (IDEA) was reauthorized, 22 years after the original IDEA was signed by the US Congress. The 1997 version presented a stronger focus on student access to the general curriculum. The word “access” became an important pivot point for advocates of children with disabilities, and the concept of UDL began to be used by policymakers and education scholars (Orkwis, 1999). The US policy focus on “access” for children with disabilities later aligned with an increased policy focus on accountability through the No Child Left Behind Act of 2001. To meet the demands of these policy initiatives, a research agenda focused on the role of educational technologies to support curricular access was undertaken by scholars interested in inclusion and UDL in the United States (Edyburn, 2005; Pisha and Coyne, 2001; Rose and Meyer, 2002).

Policy borrowing and UDL

Over the two decades following the adoption of UDL frameworks in American educational policies, the concept of UDL began to “travel” around the world, through a process of borrowing and incorporation of principles and practice. Around the same time, Ochs and Phillips (2004) attempted to conceptualize the process of “policy borrowing.” In an article in the journal *Comparative Education*, the authors acknowledged that the concept of policy borrowing was controversial, and recognized that the concept is sometimes characterized by a process of “lending” (Steiner-Khamsi, 2004), as at times governmental and other actors actively import educational approaches while at times external actors “lend” policies through provision of grants, technical assistance, or other conditional loans. In the following section, we will draw upon the concept of policy borrowing to help explain how UDL has traveled worldwide as part of a broader global movement toward enhancing inclusive education for all students through the provision of an accessible curriculum.

Ochs and Phillips (2004) conceptualization of policy borrowing is useful because the authors present four phases as a theoretical framework designed to analyze how the process occurs. The authors argue that there is often an “attraction” between two societies that stimulates the flow of educational ideas. This attraction, according to Phillips and Ochs, might be in relation to the perceived ability of the new policy or practice to meet local needs. It is likely that the policy/curricular attraction of UDL could fill many gaps, as it at once relates to a philosophy (equitable opportunity), an ambition (accessibility), and an educational approach (multiple means of classroom engagement). Phillips and Ochs argue that once the policy attraction is developed, there is a decision to adopt, a period of early implementation of the policy/innovation as it looks in the “lender” society, then an eventual “indigenization” of approaches as policies/innovations are attempted and scrutinized within particular national contexts (Ochs and Phillips, 2004).

Policy travel, UDL, and education for all

The policy attraction of UDL is related to the concept of inclusive education. The concept of inclusive education is for all children, but specifically serves as a strategy for educating children who are marginalized in education systems, including children with disabilities. Inclusive education was first identified as a practice that would help increase the chances of success of Education for All movements that began in the early 1990s. Delegates at the Salamanca Conference on Special Needs Education identified that the most appropriate way to educate a child with special educational needs was to provide access and support in regular education environments (UNESCO, 1994). The concept of inclusive education gained traction both because it aligned with existing

educational initiatives (i.e., Education for All), but was also framed as a more cost-effective option for governments than building separate facilities for children with specific educational needs.

Although the concept of inclusion as outlined in UNESCO’s *Salamanca Statement and Framework on Special Needs Education* (1994) mainly focused on access to schooling, increased focus on access to curriculum once children arrived in school soon followed. In the years that followed the *Salamanca Statement*, scholars began to identify the substantive changes that would be required for inclusive education to occur. Lena Saleh (1999), for example, stated that holistic changes must occur in order to meet the recommendations of the *Framework*, including “genuine changes in all aspects of the school’s programs—curriculum, pedagogy, organization, assessment, staffing, school ethos and school support” (1999: 209). Similar calls for educational change occurred after the Convention on the Rights of Persons with Disabilities (CRPD) (United Nations General Assembly, 2006). CRPD, which, at the time this article was written, was signed by 182 countries, called for access to the general education curriculum for children with disabilities. It wasn’t until General Comment 4 on CRPD, which was negotiated by ministerial and civil society representatives in 2016, however, that Universal Design for Learning was foregrounded as a pedagogical approach for ensuring access to curriculum. The Committee on the Rights of Persons with Disabilities encouraged States parties to

apply the Universal Design for Learning (UDL) approach. UDL is a set of principles, providing teachers and other staff with a structure to create adaptable learning environments and develop instruction to meet the diverse needs of all learners. It recognizes that each student learns in a unique manner and involves developing flexible ways to learn: creating an engaging classroom environment; maintaining high expectations for all students, while allowing multiple ways to meet expectations; empowering teachers to think differently about their own teaching; and focusing on educational outcomes for all, including those with disabilities. Curricula must be conceived, designed and applied to meet and adjust to the requirements of every student, and providing appropriate educational responses. Standardized assessments must be replaced by flexible and multiple forms of assessments and recognition of individual progress towards broad goals that provide alternative routes for learning (2016: 9).

Explicit mention in United Nations proclamations and treaties does not necessarily guarantee that an educational concept is borrowed, lent, or transferred, but the practicality of UDL as a means for meeting broader-level societal goals of access to the general curriculum for students with disabilities soon after CRPD began to be chronicled in national examples in Europe, Asia, South America, and Africa and was framed as a new “paradigm” for learning (Ostroff, 2011). This new paradigm appears to either be developing as a result of newly made commitments to inclusion and accessibility or as an updated strategy to previous (yet similar) initiatives such as Child Friendly Schools (UNICEF, 2009), student-centered learning, and adaptive learning. The extent to which UDL has been appropriated into existing initiatives is unknown, but because UDL has a specific context from which it started and has specific terminology, analysis of policy travel is transparent. In the following paragraphs two national examples are provided based on this article’s authors’ experiences. The cases of Armenia and Ghana show different processes for the adaptation and adoption of UDL as examples of how Ochs and Phillips (2004) concepts are influenced by contextual realities such as political trends, resources, and external partnerships.

National case studies

In the paragraphs below, we present two national examples of the borrowing of UDL. Although the Republic and Armenia and Ghana operate education systems in two very different contexts, they are united in their attempts to apply UDL in educational frameworks for learners with diverse needs. After describing the case of each country, we will analyze the linkage between UDL and policy borrowing.

The Republic of Armenia, like many countries in Central and Eastern Europe, has had a long history of institutionalization of students with disabilities, which gradually evolved into a system of special schools available to parents seeking therapeutic interventions. Recent changes in Armenian law demonstrate that the push toward inclusive education is slow and measured, with a policy reluctance to adopt full inclusion. Armenia’s 2015 Law on Education, for example, does not even mention “inclusion” but rather gives parents the choice of where and how to educate their child with a disability, stating that “The education of children in need of special conditions for education may be provided upon the choice of parents both at common institutions of general education and special institutions through special programs” (Republic of Armenia, 2015: Article 19.3). Despite the slow policy adoption of inclusive education at the national level, there has been a large degree of “lending” and reinterpretation about inclusive education, UDL, and access to curriculum by international agencies (e.g., UNICEF), international non-governmental organizations (NGOs) (e.g., Open Society Foundation), and their Armenian counterparts (e.g., Education and Research Foundation, Armenian Center for Democratic Education). These organizations have pushed for greater inclusion of children with disabilities through inclusive pedagogy and accessible curriculum.

The Armenian State Pedagogical University, for example, recently adopted a textbook and coursework for training its preservice teachers that includes UDL (Tichá et al., 2019). A forthcoming article (Avagyan et al., 2021) contextualizes UDL in Armenia, framing it as a mechanism to provide all children with access to content related to Armenia’s history and civic traditions. The case of Armenia demonstrates that despite a policy attraction to international treaties (the Republic of Armenia signed the CRPD in 2010), histories and interpretations of inclusive education vary, as indicated by language of “choice” in Armenia’s laws. Further, the concept of UDL, based on emerging evidence, appears to be one embraced by civil society and non-governmental groups, including Teach for

Armenia (Miller, 2019), but is not a core governmental educational initiative. The journey of UDL, then, could be seen as one that has traveled to Armenia, is being both promoted by international non-governmental organizations and adopted by national NGOs, but has yet to, and may never, achieve policy significance in Armenia.

In Ghana, the recent implementation of UDL on a pilot basis in select schools is underpinned by decades of policy and programmatic change in special and inclusive education. This includes the 1992 Constitution of Ghana, 1998 Children’s Act, and the 2006 Persons with Disability Law, which all recognize the right of children with disabilities to an education (Avoke and Hayford, 2017). In practice, children with “severe” disabilities—categorized as those who are deaf, blind, or have intellectual disability—continue to be educated in segregated schools, a practice that neither the more recent 2015 Inclusive Education Policy nor the 2010–2020 Education Strategic Plan have explicitly challenged (Avoke and Hayford, 2017; Government of Ghana, 2012, 2015). The policy-level reinforcements of segregation for some learners, along with unspecific strategic planning on the transformation of mainstream education environments into truly inclusive settings, has been accompanied by an observation that these plans have not thoroughly considered the changes needed in teaching approaches to effectively include all learners (Nketsia, 2019). This challenge is compounded by a national system of identifying disabilities that remains “confusing, ineffective, and inefficient,” which is in turn linked to the rare practice of formally diagnosing a child with a disability in Ghana (Avoke and Hayford, 2017, p. 360).

One programmatic initiative responding to this concern about teaching approaches was UNICEF’s import—or arguably, borrowing—of its global child-friendly schools model to Ghana in the early 2000s (Chabbott, 2004). The principle of child-friendly education was then reflected in both the 2015 Inclusive Education Policy and the 2010–2020 Education Strategic Plan. Further to this support to child-friendly schools, UNICEF collaborated with the Ghana Education Service to pilot inclusive education activities across the nation, which has helped to produce over 2000 mainstream schools practicing inclusive education in Ghana (Ministry of Education, 2015).

Beginning in 2012, Ghanaian Professor Samuel Hayford presented the concept of UDL in Ghana, after he had learned about it through his Ph.D. studies in England. He also embedded it into the graduate education coursework taught at the University of Education, Winneba (personal communication, 2020). Ultimately, through support from Dr. Hayford, UNICEF, the Ghana Education Service, and others, the 2015 Inclusive Education Policy proclaimed the overarching goal of the policy is to “respond to the diverse needs of all learners within the framework of Universal Design for Learning and Child Friendly School Concept” (Government of Ghana, 2015, p. 3). An inclusive education framework founded in the principles of UDL has laid way to the piloting of a UDL pedagogical approach by 15 Ghanaian primary school teachers over the 2019–2020 schoolyear (Hayes, 2019; Niad, 2019). The Ghana Education Service and its National Teaching Council have since expanded the initiative on a national basis through the embedding of UDL into existing in-service training programs, an approach that may support sustainability as compared to treating UDL as a standalone topic.

The examples from Armenia and Ghana illustrate that educational policies (or strategies) such as UDL are often first brought in by external actors. In these cases, UNICEF was instrumental in establishing inclusive education ideals and the strategies that follow. In the case of Armenia, local civil society organizations picked up on examples, adapted them to local educational conceptions of public value, and implemented through training and outreach. Further, university faculty engaged in international projects began to adapt and adopt UDL content in pre-service teacher training. In Ghana, the combination of an external actor (UNICEF) and a local champion with international experience made possible the dialogs on UDL that eventually led to policy shifts that were scaled up across the country.

Conclusions

The case of UDL as a mechanism for providing widespread access to curriculum and education appears to follow some of the same patterns identified by Ochs and Phillips (2004). These patterns, according to the authors, are predictable. A global movement toward inclusiveness and access to education for children with disabilities is underway and the concept of UDL appears to be at once a guiding philosophy, ambition, strategy, enabling structure, attractive process, and new technique that has traveled well. The cases of Armenia and Ghana, however, highlight the limitations of UDL as a borrowed concept.

In both nations, there are gaps in policy in relation to supporting access to national curriculum for students with particular types of disabilities. In this case, the concept of inclusion for “all” still has asterisks for children with particular disability types. It is quite possible that the continued exclusion of children with particular disabilities from inclusive schools is also a relic of policy borrowing, either from the hyper-medicalized and segregated education system of special education brought to Central and Eastern European countries by the Soviet Union, or the charitable (and equally medicalized) system of segregated education brought to Ghana by European missionaries.

In many ways, UDL is a sign of the times. It provides a nicely packaged set of principles that can be adapted across educational settings to provide access to curriculum for children with disabilities alongside all other peers from their communities. Where UDL has fallen short in Armenia, however, is its potential for transforming curricular access for children with significant disabilities and the degree to which it has reached significant scale in its new host countries. In theory, accessibility should facilitate for inclusive classrooms that allow for any child to experience schooling in a way that is adaptable, flexible, and meets their needs. In practice, many of the early examples for UDL focused on children with reading and attentional disabilities (see Rose and Meyer, 2002), so when it traveled it did not have an impact on children with more significant learning challenges. Therefore, Ochs and Phillips (2004) concept of “synthesis” of UDL into existing systems of Armenia and Ghana, but also failed to create transformative change

for children with the most significant disabilities. This limitation may be the result of the borrowing process itself or lack of political will within countries to attempt inclusion for all students.

The story of UDL and policy borrowing for Armenia demonstrates both its promise and shortcomings. Readers may ask how children with significant disabilities can then be included if the concept of UDL has not yet facilitated such inclusion. The answer may lie in Schwanke et al.'s (2001) A3 model. While children with mild to moderate disabilities are now experiencing a universally designed curriculum in many areas of the world, a strategy for children who have not yet had such opportunity may be to return to “advocacy”.

According to Schwanke et al. the very first step in supporting accessible curriculum is to advocate for the inclusion of all children. This work has already begun in Armenia by parents and organizations of persons with disabilities. If and when these advocates are successful, there will likely be a period when students are accommodated individually as teachers adjust to expanded conceptualizations of what it means to be a member of a classroom community. With practice and support, it is feasible that teachers may then begin to think about accessibility in broader and more flexible ways than ever before. If such a learning revolution occurs in this country, the policies and practices employed may then be borrowed elsewhere, establishing a new circular flow of global policy borrowing, accessibility, and universal design for learning.

In Ghana, the day-to-day aspects of UDL implementation are slightly different than Armenia, but land in a similar place in terms of the continued need for advocacy for children with significant disabilities. Disability screening and identification is in a nascent stage in Ghana right now, as new policies and practices emerge which are aimed at finding and supporting children with disabilities. In this way, UDL has worked well as an approach to supporting all children who are “casually integrated” (see Miles, 1997) into Ghana’s mainstream education system. Unlike foundational UDL logics based in the US that aimed to support inclusive education for children with mild to moderate disabilities through creating accessible classrooms, in Ghana accessible classrooms are often a starting point for teachers who may have little diagnostic information on the children in their classrooms, including a common absence of basic hearing and vision screening. In many ways, this lack of identification allows for a leapfrogging in Ghanaian classrooms, where accessibility can flourish without a need for disability labels. In the A3 model, a system without labels for those with less significant disabilities might go directly to “accessibility” without needing prerequisite advocacy or accommodations. Such leapfrogging at once creates a model case for accessibility, but, as mentioned above, may be limited by traditional models of segregated education or lack of access to schooling at all for children with significant disabilities. In this case, advocacy and accommodations would benefit by supporting the increased entry of such children into schools, who could later be supported by UDL and other relevant strategies.

In conclusion, inclusive education and curriculum are two tightly wound initiatives. No system can claim to serve the public unless it is accessible to all of its members. UDL is one approach to ensuring that education can be accessible to all participants in an educational setting. It has shown great promise in the United States by equipping mainstream education environments with strategies to promote inclusion of learners with disabilities, and has subsequently traveled around the world, including two countries with relatively little in common—Armenia and Ghana. Our review of UDL and the borrowing of its concepts reveals two considerations for further research. First, UDL appears to be an intuitive approach to creating accessibility and can be easily applied in both systems with robust identification of children with disabilities and systems in which children’s needs are not specifically labeled, but may be recognized by teachers. In either setting, UDL can work, but the path to accessibility may be different. The second consideration for UDL as it is borrowed and travels worldwide is that more comparative data is needed on how UDL can be leveraged for access to mainstream curriculum for children with more significant sensory or cognitive disabilities.

Rao, Smith, and Lowry, for example, lamented that “Despite the fact that UDL is acknowledged as a means to further inclusionary practices for all students, UDL—and its potential outcomes for students with (intellectual disability)—has yet to be systematically considered and applied within current UDL reform efforts” (2017: 46). In this way, UDL’s promise has traveled, been locally adopted, and has provided promising outcomes worldwide. At the same time, its limitations have traveled, creating space for a new wave of localized, contextualized research on how a concept like UDL can be applied for new populations. Such research may “borrow” and “expand” UDL or may have an entirely new set of terms and applications. Either way, “all” still does not mean “all” in 2022 as children with some forms of disabilities continue to be excluded from their national curriculum.

References

- Avagyan, A., Johnstone, C., Asatryan, O., Khachatryan, L., Shagafyan, A., 2021. Localizing a universal claim: applying universal design strategies to support inclusion in Armenia. In: Schuelka, M., Carrington, S. (Eds.), *New Directions in Inclusive Education*. Routledge, London.
- Avoke, S.K., Hayford, S.K., 2017. Republic of Ghana. In: Wehmeyer, M.L., Patton, J.R. (Eds.), *The Praeger International Handbook of Special Education*. Praeger, Santa Barbara, CA, pp. 352–367.
- CAST, 2018. Universal Design for Learning Guidelines Version 2.2. Retrieved from. <http://udlguidelines.cast.org>.
- Center for Universal Design, 2008. Universal Design Principles. Retrieved from. https://projects.ncsu.edu/ncsu/design/cud/about_ud/udprinciples.htm.
- Chabbott, C., 2004. UNICEF’s Child-Friendly Schools Framework: A Desk Review. UNICEF. Retrieved from. https://www.researchgate.net/publication/292116843_UNICEF's_Child-Friendly_Schools_A_desk_review.
- Dolmage, J., 2015. Universal design: places to start. *Disabil. Stud. Q.* 35 (2). Retrieved from. <https://dsq-sds.org/article/view/4632/3946>.
- Eddyburn, D.L., 2005. Universal design for learning. *Spec. Educ. Technol. Pract.* 7 (5), 16–22.
- Eddyburn, D.L., 2010. Would you recognize universal design for learning if you saw it? Ten propositions for new directions for the second decade of UDL. *Learn. Disabil. Q.* 33 (1), 33–41.

- Government of Ghana, 2012. Education Strategic Plan 2010–2020. Retrieved from. <https://globalpartnership.org/sites/default/files/2013-Ghana-Education-Strategic-Plan-2010-2020-%20Vol.%201.pdf>.
- Government of Ghana, 2015. Inclusive Education Policy. Retrieved from. <https://sapghana.com/data/documents/Inclusive-Education-Policy-official-document.pdf>.
- Hayes, A., Turnbull, A., Moran, M., 2018. Universal Design for Learning to Help all Children Read: Promoting Literacy for Learners With Disabilities. USAID, Washington, D.C.
- Hayes, A.M., 2019. Identifying Students With Disabilities in Ghana: Challenges and Opportunities. Retrieved from. <https://www.inclusivedevpartners.com/contact-us-identifying-students-with-disabilities-in-ghana-challenges-and-opportunities>.
- Hehir, T., 2007. Confronting ableism. *Educ. Leadersh.* 64 (5), 8–14.
- International Disability Alliance, 2021. Universal Design for Learning and Its Role in Ensuring Access to Inclusive Education for All.
- Meyer, A., Rose, D.H., Gordon, D., 2014. Universal Design for Learning: Theory and Practice. CAST, Wakefield, MA.
- Miles, M., 1997. Disabled learners in South Asia: lessons from the past for educational exporters. *Int. J. Disabil. Dev. Educ.* 44 (2), 97–104.
- Miller, K.J., 2019. An American professor in Armenia: reflections on special education, speech-language pathology, and university teaching. *Commun. Disord. Q.* 40 (3), 156–166.
- Ministry of Education, 2015. Education Sector Performance Report. Retrieved from. [https://new-ndpc-static1.s3.amazonaws.com/CACHES/PUBLICATIONS/2016/03/22/Education+Sector+Performance+Report+\(ESPR\)+2015_Final.pdf](https://new-ndpc-static1.s3.amazonaws.com/CACHES/PUBLICATIONS/2016/03/22/Education+Sector+Performance+Report+(ESPR)+2015_Final.pdf).
- Niad, H., 2019. Ghana's Readiness for Inclusive Education. Retrieved from. <https://www.inclusivedevpartners.com/ghana-s-readiness-for-inclusive-education>.
- Nketsia, W., 2019. Inclusive education policy and practice in ghana: air castle or realistic goal? In: Pather, S., Slee, R. (Eds.), *Challenging Inclusive Education Policy and Practice in Africa*. Brill Sense, Leiden, pp. 69–86.
- Ochs, K., Phillips, D., 2004. Processes of Educational Borrowing in Historical Context, pp. 9–10. Na.
- Orkwis, R., 1999. Curriculum Access and Universal Design for Learning. ERIC/OSEP Digest# E586.
- Ostroff, E., 2011. Universal design: an evolving paradigm. In: Preiser, W.F.E., Smith, K.H. (Eds.), *Universal Design Handbook*, second ed. McGraw Hill, New York, pp. 34–42.
- Perelmutter, B., McGregor, K.K., Gordon, K.R., 2017. Assistive technology interventions for adolescents and adults with learning disabilities: an evidence-based systematic review and meta-analysis. *Comput. Educ.* 114, 139–163.
- Pisha, B., Coyne, P., 2001. Smart from the start: the promise of universal design for learning. *Remedial Spec. Educ.* 22 (4), 197–203.
- Powell, J.J., 2013. From ableism to accessibility in the universal design university. *Rev. Disabil. Stud.* 8 (4), 33–45.
- Rao, K., Smith, S.J., Lowrey, K.A., 2017. UDL and intellectual disability: what do we know and where do we go? *Intellect. Dev. Disabil.* 55 (1), 37–47.
- Republic of Armenia, 2015. Law on Education (Official English Translation). Ministry of Justice, Yerevan. Retrieved from. <http://www.anqa.am/en/about-us/legal-field/laws/law-of-the-republic-of-armenia-on-education/>.
- Rose, D.H., Meyer, A., 2002. Teaching Every Student in the Digital Age: Universal Design for Learning. Association for Supervision and Curriculum Development, Alexandria, VA.
- Saleh, L., 1999. The rights of children with special needs: from rights to obligations and responsibilities. *Prospects* 29 (2), 203–216.
- Schwanke, T.D., Smith, R.O., Edyburn, D.L., 2001. A3 model diagram developed as accessibility and universal design instructional tool. In: RESNA 2001 Annual Conference Proceedings, vol. 21. RESNA Press, Washington, DC, pp. 205–207. No. 1.
- Steiner-Khamsi, G. (Ed.), 2004. *The Global Politics of Educational Borrowing and Lending*. Teachers College Press, New York.
- Minneapolis, MN, USA: University of Minnesota. In: Tichá, R., Abery, B., Johnstone, C., Poghosyan, A., Hunt, P. (Eds.), 2019. *Inclusive Education Strategies: A Textbook*. UNICEF Armenia & Armenian State Pedagogical University, Yerevan, Armenia.
- UN CRPD Committee, 2016. General Comment. No. 4. UN Doc CRPD/C/GC/4.
- UNESCO, 1994. The Salamanca Statement and Framework for Action on Special Needs Education: Adopted by the World Conference on Special Needs Education; Access and Quality. UNESCO, Salamanca, Spain, pp. 7–10. June 1994.
- UNESCO, 2020. Inclusion in Education: All Means All. Global Education Monitoring Report 2020. UNESCO.
- UNICEF, 2009. Child Friendly Schools Manual. Author, New York. Retrieved from: <https://www.unicef.org/media/66486/file/Child-Friendly-Schools-Manual.pdf>.
- United Nations General Assembly, 2006. Convention on the Rights of Persons With Disabilities. United Nations.

FOURTH EDITION

International Encyclopedia of Education

Higher Education Policy,
Governance and Practice

VOLUME
8

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Chika Sehooole
Everard Weber



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 8

Higher Education Policy, Governance, and Practice

VOLUME 8 EDITORS

Chika Schoole and Everard Weber



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

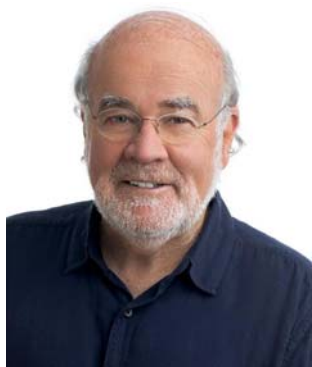
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



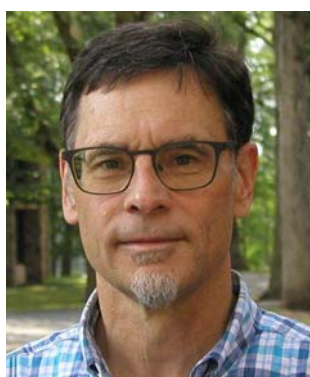
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 8: Higher Education Management, Policy, and Practice

Chika Schoole and Everard Weber

© 2023 Elsevier Ltd. All rights reserved.

Introduction

In recent years, several encyclopedias and handbooks such as the third edition of this encyclopedia have been published, addressing the emerging forms, practices, and issues of higher education. Most of these works, however, focus on shifts in higher education policy, management, and practice in the Global North. In so far as the descriptions of the systems of higher education in the Global South are concerned, much of where the interest lies is in economically advanced countries such as Singapore, South Korea, and Hong Kong and in the rapidly developing system of higher education in China. Little attention is paid in these encyclopedias and handbooks on the less developed systems, and when it is, it is invariably articulated in the language of challenges and barriers to reform, and their under-developed condition. This volume of the *International Encyclopedia of Education* is constructed deliberately to overcome this gap in the literature. It is edited from South Africa, with a number of papers included in it exploring issues of higher education as they are faced by various countries in Africa. Many other papers moreover examine the broader global issues in higher education from an African perspective, while others still could be viewed as an invitation to consider how global developments are interpreted and affect policy development in systems of higher education systems in the Global South, and their contribution to higher education's global architecture.

Higher education has for centuries been associated with the imperatives of societal development through educating individuals at a higher level, as well as through knowledge creation and advancements in the sciences and humanities. Before World War II, the forces of industrialization and modernization stimulated the establishment of higher education institutions (HEIs) to meet the requirements of emerging societies, not only in the Global North but also some parts of the South, where HEIs were created to serve colonial interests. After World War II, however, higher education in the Global South has widely been viewed as a major driver toward national development, supported by United Nations institutions and multilateral development banks. This period has also seen the emergence of a series of ideas and institutions in the Global North and on how the Global North could bring more development to the Global South. According to Aydin (2014) this changing relationship led to the emergence of the field of international development, as higher education and the ideas of social and economic development became intractably linked. Higher education has not been used only for national development; but after World War II, with the emergence of the Cold War, International Relations became shaped by a confrontation between the United States and Soviet Union, an East-West confrontation framed by divergent models of their respective political systems, economic systems, and ideologies. In this context international aid programs became a major part of the strategy to win over the "minds" and "allegiance" and to reproduce or transform foreign dominant groups (Tsvetkova, 2008).

After World War II, the western knowledge systems continued to dominate higher education policies and practices in the Global South, as a strategy to forge the "minds" of local populations capable of working toward development objectives. There is critique, for example, by Serian et al. (this volume) of the use of English as the unproblematic language of instruction, with a suggestion that if indeed an equitable, societally beneficial, and high-quality "education for all" is to be a major goal of higher education, it needs to be embedded within a larger processes of global knowledge systems in which English plays a hegemonic role. In more recent decades, the colonial assumptions persist even as aspirations of decolonization have acquired considerable currency. Within the Global South, the discourses of internationalization have become popular, but carry with them a range of dilemmas and contradictions.

It is with a deep awareness of the dilemmas and contradictions associated with the competing agendas of internationalization and localization that the papers in this volume examine issues of policy governance and leadership; internationalization; curriculum, teaching, and learning; diversity, inclusion, and equity; and labor markets and vocational education and training. The papers on curriculum, diversity, and labor markets are located within the broader discussion regarding, for example, globalization and decolonization. The latter, in other words, contextualize the former and we hope deepen analyses at the institutional level and at the level of practice. By this we do not mean to suggest the local developments are a function of historical or international policy change. Case studies in particular countries are instructive in throwing into relief how global developments are reconstituted and rendered complex on the ground. When all is said and done it is the people located at different levels and their actions, or inactions and interrelationships that shape higher education.

Higher education has multiple purposes, some relating to the broader global concerns, others dealing with highly localized issues and challenges. At the broader level, the contribution of higher education, research, and innovation to International Relations is increasing in importance and complexity, with the role of higher education in International Relations viewed through a "knowledge diplomacy" lens. Knowledge diplomacy is defined by Knight (in this volume?) as fundamentally different from cultural diplomacy, science diplomacy, and soft power. She identifies eight primary characteristics of knowledge diplomacy, with illustrative case studies from around the world. The examples of knowledge diplomacy focus on how through cooperation, reciprocity, and mutuality of benefits international higher education can strengthen relations between and among countries to address current global issues. Knowledge diplomacy is not, however, without its challenges and demands further reflection and research.

Internationalization of higher education has also found expression through different forms such as student mobility. In recent decades, the number of globally mobile students in higher education has grown rapidly, arguably as a direct outcome of the

increasing commodification of international education. Using Australian higher education as an illustrative case, Rizvi (in this volume) suggests that this commodification is best viewed as involving set of policy processes relating to the ways in which HEIs are now funded and governed. It shows how the commodification of social mobility has thus reconstituted the character of higher education, becoming a site for the formation of an emerging transnational elite class. Globalization and the internationalization of higher education have resulted in greater cross-border student mobility and increased international integration. These developments have influenced the social relations between the locals and the internationals at universities.

Historically, higher education has been an instrument in the battles that have been fought between the East and the West, the South and the North. Chasi (this volume) opines that the perspective of how Global North could bring about development in the Global South is informed by levels of inequalities between the bipolar worlds which are a result of colonialism and the underdevelopment of the South by the North. It assumes the dependency of the South on the North for its development. In this volume the North–South partnerships in the context of higher education internationalization are addressed from the vantage point of the Global South. It reflects on the past, present, and future of North–South engagements in higher education with reference to the notions of colonization, recolonization, and decolonization. A decolonial lens is used to critically reflect on the challenging dynamics of North–South partnerships, particularly regarding continued dependencies and inequalities, and to challenge existing partnerships practices in the broader context of higher education internationalization. Furthermore, it serves as a tool to conceptualize and empower positive change by reimagining North–South partnerships.

As has been noted, this volume is assembled from a Global South perspective that intends to highlight developments in higher education in the Global South in general, and the African perspective. These have not been given attention in the previous editions of this series and therefore make this volume particularly unique. As [Mouhabedai \(2014\)](#) highlights, centers of higher learning have existed in Africa for several centuries, well before the arrival of Europeans. Examples of these are the University of al-Karawiyyin in Fez, Morocco; Al-Azhar University in Cairo, Egypt; and the University of Timbuktu in Mali, the latter being attributed as the oldest in the world. However, modern higher education in Africa has its roots in university colleges that were created by and affiliated with European universities during the colonial period. They were staffed by Europeans or Africans trained in Europe, and their major objective was training the workforce for the public sector and teachers for the rapidly expanding secondary education sector. These institutions were patterned on the European higher education system; and after the colonies achieved independence in the 1960s, the colleges became universities. Because of underfunding that has been associated with the underdevelopment of African countries and former colonies, higher education development in these regions failed to catch up with that in the Global North. However, as some papers in this volume will demonstrate, there have been significant developments in higher education in African albeit not on the same level as those in the Global North, that make a significant contribution to continental development.

The role of higher education in development has been combined with the rate of return on investment both for the public and the individual has contributed to the increased demand for access to higher education. Many countries responded by expansion of higher education provision which led to the massification of higher education. As it will be demonstrated later in this volume, there has been an increase in universal access to higher education over the past two decades in all the regions. Between 2000 and 2018, the global higher education gross enrollment ratio (GER) increased from 19% to 38%. For the same period, the male GER increased from 19% to 36%, while that for females went from 19% to 41%. This demonstrates that women have benefited more from increased access to tertiary education. Currently higher education services 220 million students in 20,000 HEIs globally ([WHED, 2021](#); [World Bank, 2021](#)). Access to higher education can be understood in a diversity of ways, as can the accountabilities that are attached to this notion. The issue of access is addressed in this volume and a distinction between different conceptualizations of access is articulated on how these relate to accountability in the contemporary conditions of higher education, with a focus on university systems, institutions, and practices.

This increased access has not been without challenges but has been characterized by barriers such as poverty, unaffordability because of high tuition fees, entrance examinations, geographical mobility, and discrimination. In this globalized context of numerical student growth, access to higher education is often taken as the opportunity to gain formal entry to a higher education study pathway (Burke this volume). Commonly, this involves a focus on counting bodies of people from social groups that policymakers recognize as historically excluded from participation at tertiary level of their formal education system. Beyond this though, Burke (this volume) argues that if we explicitly move to consider how inequality becomes structured into institutions and practices, the notion of access can also help to explain whose and which knowledge matters in higher education. Drawing on sociological accounts of the character and limits of knowing and knowledge, the term “epistemic access” has been used to explore this dimension of accessibility. Efforts toward “widening” access to higher education before moving to the issue of epistemic access is explored with reference made to the relationship of access to accountability, including the tensions between these terms.

Academic freedom and institutional autonomy receive attention in this volume and the threat they are under. In his paper, Oanda shows how the exercise of academic freedom and social responsibility in African universities has been undermined over time by internal constituencies with the acquiescence of the academics. He further argues that the constriction of spaces for academic freedom seems to be deteriorating more as governments (the traditional inhibitors to academic freedom) have ceded governance oversight to the institutions and legalized staff unions to be part of the governance and management structures of the institutions. On the other hand Hutchens points out that political systems treat academic values differently and the same protections are not afforded globally. In his paper he explores the history of academic freedom and how it is viewed and protected worldwide. Of late, governmental protections have been precarious in some nations and scholars have been reprimanded in countries that encroach onto the autonomy of higher education.

The assault on higher education is not only through a threat to academic freedom but also through the encroaching of market forces in higher education. The paper by Suspitsyna and Pasque examine the current roles of European states, markets, and universities in knowledge production and review the critique of academic entrepreneurialism and neoliberal market influences on European higher education. On the other hand Magalhães and Veiga (this volume) identified the impacts of governance reforms in Portuguese higher education initiated in 2007. Under the influence of the New Public Management perspectives, the changes were driven by the inducement of a quasi-market environment enhancing institutional autonomy and promoting the inter-institutional competition. The reforms have transformed the relationship between the state and higher education system and the power relationships between HEIs' constituencies.

Yonezawa, Shimauchi, and Liu (this volume) examine the transformation of university missions from nation-building and social development, toward global competitiveness and engagement around common global challenges. At universities in East Asia, the translation of knowledge tends to be implemented within a context of hegemonic power, rather than through voluntary academic activities. This paper raises concerns over "losses in translation" that may lead to misunderstandings around the essential characteristics of universities, such as academic freedom and university autonomy.

As part of the pushback on the assault on higher education, Giroux and Filippakou (this volume) argue that the educational force of the wider culture in the age of a raging pandemic, economic populism, and authoritarian nationalism has become a force for promoting a fascist politics, a widespread assault on the welfare state, and a hatred of those groups marginalized by race, class, and religion. At stake here is a further assault on civic culture, civic literacy, and the recognition that a democracy cannot exist without informed, engaged, and critical citizens. They argue that educators, especially those working in higher education, need a new language for thinking about how democracy and politics should provide a new way of thinking about the centrality of critical pedagogy to the production of the ideals of democracy, critical agents, economic equality, and social justice. Their paper concludes by calling for a pedagogy of hope, one infused with a new vision of the future, support for young people who are leading the struggle for economic and social justice, and the widespread adoption of pedagogical practices that will create the agents and the conditions that will enable collective struggles and emancipatory transformative changes.

The use of technology in higher education, digitalization, and green transitions which affect the way we work and learn which have implications for student support and funding of higher education are some of the issues that are explored in this volume. According to Weiland (this volume) digitalization has also meant revision of how information—in journals, books, and other resources—is gained, used, and maintained. For higher education these have given rise to unanticipated forms of roles for scholarly communication, now understood as digital system of exchange of ideas and practices across the many disciplines of inquiry. In this regard, open access, which is another form of the expression of digital transformation, is transforming science and scholarship by removing financial constraints on access to journals and books. The use of technology regarding its impact on curricula is also explored.

A crucial dimension of the globalization and internationalization of higher education has been the neoliberal national, regional, and international contexts within which curriculum change has taken place locally. These have been shaped by what is transpiring within a range of HEIs. We need to know more how men and women in the North and South, and within and across different systems, on a day-to-day basis, receive and work with macro-development and how it articulates with teaching and learning. We suspect that in surprising ways there may be little impact at all. Bologna is not only about "harmonizing" diverse systems, but it also has symbolic meaning for the creation of a more united world in higher education and thus in what is or is not taught in lecture halls and/or online. But none of this is problem-free. Macro-developments such as neoliberal influences are often contested and reconstituted in hybrid ways.

The role and impact of technological development has been far-reaching. The debates about a 4th Industrial Revolution (4IR) pose critical questions about technology and curricula. More specifically, it has redefined, for example, assessment and performance measurements and teaching and learning outcomes and is presently reshaping distance education. These developments have been accelerated by the COVID-19 pandemic. The distinctions between the delivery of online and contact curricula have become blurred in ways which at this point we do not fully grasp.

The massification of higher education is further determining the nature of diversity at universities and colleges and raising more sharply questions about inclusion and equity. Thus, as more students from disadvantaged groups have been included, familiar issues about their alienating experiences within predominant and historically evolved cultures have been posed. How these are evolving should be researched. Questions of differences by race, class, and gender are reproduced within social contexts where countervailing forces pushing groups together. These as we show below are played out in a variety of ways and across diverse settings. They also include and overlap with difference by national origin as a result of recent migrations (Weber and Joubert in this volume).

Global and technological developments provide the background for how young people are prepared for labor markets and the world of work. Familiar concepts such as pathways become relevant in new ways. Thus, for instance, the question is posed about how curricula in schools prepare students for changing workplaces where teamwork driven by the needs for innovation and competition have replaced hierarchy. Men and women, blacks and whites, nationals and recently-arrived internationals must now work together more closely. We focus particularly on the development of professional identities of lecturers in this regard and highlight the importance of the wider social context in which teaching at universities occurs. And here most especially the difference between North and South is evident. For example, in their paper Sichombe et al. (this volume) discuss early career academics and researchers in Namibia require skills and competencies their counterparts in the North take for granted. This determines their struggles to conduct research, publish, and develop networks and thus their professional identities.

Curricula are delivered within overarching national and international contexts. Kenneth Saltman writes that in the United States neoliberalism as economic policy and as ideology has influenced education change from K12 to higher education. What specifically have these restructuring initiatives entailed? The profit motive underpins policy and public debate he says. A critical approach is required to dissecting what is becoming a common language and way of thinking about educational change. In the literature curricula are often abstracted from wider social developments. This volume addresses this problem through this paper. Additionally, what is novel in the analysis by Saltman are the natural and unnatural disasters within the discussions of neoliberalism and globalization. Environmentalism as it relates to curricula is one of the focus areas that regrettably we have neglected.

O'Leary states that research is important in improving higher education pedagogy. His view is informed by longitudinal studies of teachers and students in the UK. These studies as well as the paper by Cook-Sather draw attention to value of qualitative analyses of lecturer-student relations, particularly the emphasis on context and complexity. The existing scholarship highlights multiple difficulties and challenges within pedagogical relations of reciprocity. The compact teachers enter in their engagements with students defines the values, diversity, and specificity of place as key components of the social dimensions of curriculum.

Bajinath, Motala, and Menon show how the idea of a 4IR has been promoted and criticized in recent years. They argue that 4IR lies at the intersections of technological progress, policy, and development, and curricula, and teaching and learning. Against this background and with reference to a middle-income country, South Africa, the paper examines the conditions, challenges, and stumbling blocks that could lie in the way of the implementation of 4IR. But before we talk about implementation the skeptics state, we should examine the previous industrial revolutions and its consequences for the disadvantaged and marginalized. Do we wish more of the same? And what are the implications institutionally in the increased use of digital technology? Mass datasets must be managed and analyzed and this requires the development of institutional capacity. How will the data be used in assessment, for example, and what are the pedagogical implications of e-teaching and e-learning? And the social relations of lecturers and their students discussed earlier? And overall, Bajinath et al. ask, as the critics warn, will another digitalized layer of inequality be grafted onto the inequalities of previous industrial revolutions? History is relevant globally and in the context of historical development in South Africa since the discovery of gold and diamonds in the late 19th century.

Linda van Reyneveld writes that new development in the delivery of teaching and learning programs has been in distance education where the traditional distinction between contact and blended approaches, using technology, is now less clear. The paper poses questions such as: Can we develop frameworks for completely online in distance at residential universities? What will inform and underpin such frameworks and what by way of examples they outline, might these frameworks look like? What institutional changes might be required? The introduction of online distance programs at what used to be traditionally residential places of higher learning raises broader questions about the social relations between teachers and students. The latter as we have seen have been central to curriculum studies in higher education. Shifts toward blended approaches and to online programs, even if confined to distance education at residential universities, have implications for higher education, particularly to the development of curriculum studies.

COVID-19 pandemic has impacted the higher education sector across the globe in unprecedented ways, necessitating a range of extraordinary responses by governments and education institutions, from travel restrictions to campus shutdowns. International travel was negatively impacted with serious implications for international student mobility. HEIs retained students and maintained their access to learning through online courses. Sehoole, Ikwemesim, and Agbaje (this volume) as well as Ayoo (this volume) have examined the higher education landscape in Africa, discussed the key educational processes affected, and how some African higher education systems dealt with pandemic with differential levels of success in sustaining teaching and learning during the pandemic. On the other hand, Global North-based institutions have been able to navigate the pandemic with lesser impact on the delivery of teaching and learning. The ability to introduce and sustain teaching and learning during the pandemic also depended on the level of the use of technological and hybrid methodologies before the dawn of the pandemic.

Ronghuai Huang et al. discuss how the COVID-19 pandemic has highlighted what has been called Online merge Offline (OMO) programs. Based upon hybrid infrastructure, open education merges online and offline teaching and learning. Face-to-face classrooms merge seamlessly in theory with online classrooms. The writers argue that the implementation of OMO should be research-based. This paper is based on four international case studies. The findings draw attention to how the learning environment is constructed, the relevance of technology, communication, and teacher training. They write that the use of technology has limitations. Teachers should reflect on how OMO courses are designed, how they teach, and how learning communities are created and consolidated. Understanding students' experiences and what it means to them is crucial.

A great deal of recent attention has been on the outcomes and assessment in curricula. The paper by Matthew Wawrzynski in this volume focuses on various dimensions of assessment. It is useful he writes to start by trying to define commonly used terms such as assessment and learning outcomes and presenting a step-by-step process whereby the learning process can be determined. This serves as an introduction to the following paper by Dirk Ifenthaler on learning analytics. The adoption of analytics requires institutional change regarding the systemic use of appropriate technology and systems, policy regulations and guidelines, professional training and research (Ifenthaler). The degree of alignment between these factors is likely to be important. Additionally, the importance of changing mindsets and behaviors should not be underestimated and, more broadly, what historically has shaped them.

Learning analytics are illustrative of a broader development in higher education: using performance measurements in assessing learning outcomes. These performance measurements are often digitalized and claim to measure critical thinking which arguably

has been reduced to a competency. The art of reasoning is now called online reasoning. It is against this background that in this volume a call is made by Zlatkin-Troitschanskaia et al. for the critical evaluation methodology of performance measurement. The collection, processing, and analysis of data specific to this field are not problem-free. There is a relevant literature here and future research will be important (Zlatkin-Troitschanskaia et al.). This approach enables us to move beyond mere technical and mechanical implementation of digitalization.

Class size is often viewed as a key determinant of learning outcomes and academic performance. The intricacies of class size make its meaning far from straightforward. It depends, for example, on what one is talking about—the size of a class, the number of students in a program, the incoming cohort? And relatedly who are the constituencies within the university that might have an interest in determining class size? Academics, the budget staff, etc. These are usually the subject of protracted negotiations within institutions and can determine class sizes ultimately decided upon (Okerson in this volume).

Having assessed students, what credits are awarded? The globalization and the internationalization of higher education mean we can interrogate credit validity and recognition beyond national borders. While there have been attempts to synchronize credits across borders, notably, the Bologna Process, most credits are issued locally. Bologna represents a major initiative at the “harmonization” of different, nationally based higher education systems. It is instructive to examine the existing literature and to consider different ways in which cross-border credentializing might be achieved and what problems have arisen (Jason Stone in this volume).

There has been a huge global growth in student numbers in recent times across both the Global North and South. Access to higher education has come to be defined in various ways as write Burke et al. in this volume. These are related to questions about accountability. It has become widespread practice to count the numbers of students and faculty, especially of underrepresented groups and those who have traditionally been excluded from university. This raises broader and more fundamental questions about how inequality becomes embedded in institutions and how it is reproduced. How one defines access can explain what knowledge and whose knowledge is privileged. The idea of “epistemic access” in relation to and in tension with accountability has been used (Burke et al.). This discussion sets the scene for our section on the diversities in higher education.

As higher education has massified, the disadvantaged have been included, but have not always succeeded. Completion rates have been problematic. Conceptualizations of completion have favored students coming from advantaged groups. Comparisons of diverse institutions and systems are also not problem-free and understanding them is important. But despite the concern for completion, what does participation in higher education mean for underrepresented groups? This means focusing attention on the institution and its role in serving what are perceived to be national goals. Those traditional and current interests might not coincide with beliefs and values of the “others” now attending HEIs (Gale and Parker in this volume).

We specifically look at diversity from the position of black women. Venitha Pillay identifies and states her position as a black woman in the South. Women have been included in higher education but remain powerless, she argues. Their exclusion is not overt but embedded in gendered institutional practices and histories. There is a global dimension to this question. Pillay states that white women from the North are more privileged than black women in the South. Being a black woman in academia is compounded by being an academic mother. It is impossible in this context to “balance” life and work. The COVID pandemic has had terrible consequences on women in higher education and on academic mothers. However, there are strategies that can be used to negotiate the subordinate position of women in academia.

It is also useful to focus specifically on individuals at elite universities in a particular country. In Australian higher education students from low-income groups are underrepresented and often the first in their families to attend university. They enter an alienating institutional culture. Gender and class intersect in the stories of two students. Where they come from interacts with where they find themselves and what they wish to become (Stahl and McDonald in this volume).

Pacific minority students have faced difficulties in Aotearoa New Zealand. These challenges regarding the experiences of minoritization, academic performance, and the support or lack thereof can be compared with similar situations in other parts of the world (Cherie Chu-Fuluifaga in this volume). We also therefore focus on the perceptions of learning of adult working-class women students from the North. They were about to transition to the labor market. Their expectations of work and experiences in higher education were classed and gendered. These women experienced the reproduction of inequality (Barbara Merrill and Scott Revers). This volume also unpacks whiteness and building more inclusive cultures. The study by Karkouki draws on case study data at a private institution in the Northeast, United States. The findings can assist academic administrators build diversity and inclusivity.

Historically Black Colleges and Universities in the United States provide an alternative home to the alienating experiences of minority students in predominantly white institutions. The paper in this volume examines the history of these colleges with reference to, for example, the civil rights movement and discusses their unique contribution to US society today. It concludes with notes about their future (Marybeth Gasman).

We wish to draw attention to the many ways in which diversity and student experience can be defined. In the paper by Basit the spotlight is on a minority group of the student population, a particular cohort of students, those enrolled for doctoral study at a university in the United Kingdom. It surveys and analyzes their experiences of teaching, learning, support, and socialization using Bourdieu’s notion of academic habitus. It makes policy recommendations that have implications for future practices.

Pathways are commonly used to understand transitions of young people to the world of work. Harris and Guthrie write that Australian Vocational Education and Training is largely postsecondary and concentrates on lifelong learning and careers. Research highlights the importance of context in the analysis of pathways. They propose recommendations to administrators to better serve student learning in preparations for work. However, as Billet in this volume notes constant change in the workplace means that preparing students for the prescriptions of occupations is not easy. Graduates must translate and adapt what they have learned to the fluidity of the conditions and nature of work and 21st century reskilling.

Bayerlein writes that there are different types of learning outcomes and expectations in a variety of workplaces. This paper explores developments in work in relation to Work-Integrated learning programs. Recently skills in information technology, communication, and the dynamics of working in teams have become important within specific contexts and types of work. Internships have positive impacts and can play significant roles in the future (Bayerlein in this volume).

The growth of the importance of career pathways has become important globally. The paper by Bragg examines the role in the United States of different agencies in programs for youths and adults. It discusses a framework for the evaluation and assessment of career pathways, based on research, including the largest federal investment in career pathways. It concludes with a review of gaps in what is known about career pathways of underserved students.

Erica Smith writes that service-learning is expanding to many countries. It is a type of community-university engagement which is geared toward action-based learning and bridges the divide between on-campus theory and community-based practice with transformative goals. Smith surveys 30 years of competency-based (CBT) in Australian vocational education and training (VET). The goals and content of CBT are criticized as having narrowed over the years. She makes recommendations for improvement.

Against the background of international developments have been concerns regarding student success and demands on poorly equipped lecturers to address these issues. Professional development plays a significant role in teaching and learning despite its systemic neglect. The paper in this volume (Precious Sipuka and Shireen Motala) pays attention to the development of better teachers in South African higher education. It highlights the complexities in constructing and developing lecturers' identities including the impact of social and political contexts. Thus, alternative approaches to professional development in higher education are suggested. The study by Sichombe et al. examined the experiences, career development, and difficulties encountered by early career researchers in Namibia, a developing country located in the South. The study found that early career academics needed mentorship and guidance in research methods, project development and implementation, writing, networking, collaboration, and publications. There was a gap between experienced researchers and their younger colleagues that must be narrowed through closer collaboration with senior and more experienced academics.

In conclusion, the papers in the volume have demonstrated the crucial importance of viewing shifts in higher education policy and practice from the perspective of the histories, cultures, and experiences of and the diversities within various societies. They have shown how the focus on the developments within the Global South is of special significance. They have pointed to the need to understand the world of higher education more from the realities, individuals, and constituencies on the ground and to use them as the point of departure for the construction of further theoretical perspectives. The growing scholarship on postcolonialism has problematized representations of the South from the point of view of the North (see Weber and Joubert in this volume). The implications for future research, including issues related to research methodologies, are crucial and wide-ranging and have the potential to redefine the field of higher education.

References

- Ayidin, O.T., 2014. Current developments and trends in higher education. *J. Business Econ. Fin.* 3 (4), 470–489.
- Mohamedbhai, G., 2014. Massification in higher education institutions in Africa: causes, consequences and responses. *Int. J. Afr. Higher Educ.* 1 (1). <https://doi.org/10.6017/ijah.v1i1.5644>.
- Tsvetkova, N., 2008. International education during the cold war: Soviet social transformation and American social reproduction. *Comp. Educ. Rev.* 52 (2), 199–217.
- WHED, 2021. About – WHED – IAU's World Higher Education Database. World Higher Education Database. <https://www.whed.net/About.html>.
- World Bank, 2021. Tertiary Education. <https://www.worldbank.org/en/topic/tertiaryeducation>.

CONTRIBUTORS TO VOLUME 8

Kolawole Samuel Adeyemo

*Department of Education Management and Policy
Studies, University of Pretoria, Pretoria, South Africa;
and Center for Studies in Religion and Society,
University of Victoria, Victoria, BC, Canada*

Olaide Agbashe

University of Johannesburg, Johannesburg, South Africa

Meltem Akbulut Yıldırım

Technical University of Munich, Munich, Germany

Walter R Allen

*School of Education and Information Studies, University
of California, Los Angeles, CA, United States*

Hanne Andersen

*Section for History and Philosophy of Science,
Department of Science Education, University of
Copenhagen, Copenhagen, Denmark*

Peter Arena

New York University, New York, NY, United States

Philip O Ayoo

*School of Computing and Engineering, Uganda
Technology and Management University (UTAMU),
Kampala, Uganda*

Narend Baijnath

Tehmina N Basit

*School of Life Sciences and Education, Staffordshire
University, Stoke-on-Trent, United Kingdom*

Leopold Bayerlein

*UNE Business School, University of New England,
Armidale, NSW, Australia*

Stephen Billett

Griffith University, Brisbane, QLD, Australia

Kiana Billot-Vasquez

*School of Education and Information Studies,
University of California, Los Angeles, CA,
United States*

Ana-Maria Bliuc

*Department of Psychology, School of Social Sciences, The
University of Dundee, United Kingdom*

Debra D Bragg

*Bragg & Associates, Inc., Emeritus Faculty at the
University of Illinois at Urbana-Champaign,
Champaign, IL, United States*

Henry I Braun

*Boston College, Lynch School of Education and Human
Development, Chestnut Hill, MA, United States*

Robert G Bringle

*Indiana University Purdue University Indianapolis,
Indianapolis, IN, United States*

Bruno Broucker

*KU Leuven, Public Governance Institute, Leuven,
Belgium*

Bailey A Brown

*Department of Sociology and Anthropology, Spelman
College, Atlanta, GA, United States*

Matthew Bunn

*The University of Newcastle, University Drive,
Callaghan, NSW, Australia*

Daniel Burgos

*Research Institute for Innovation & Technology in
Education (UNIR iTED), Universidad Internacional de
La Rioja (UNIR), Logroño, Spain*

Penny Jane Burke

*The University of Newcastle, University Drive,
Callaghan, NSW, Australia*

Tim Carroll

New York University, New York, NY, United States

Lillian Vega Castaneda

*Professor Emerita Education California State University
Channel Islands, Camarillo, CA, United States*

Mario E Castaneda

*California State University, Los Angeles, CA, United
States*

Gloria Castrillón

*Center for Academic Planning and Quality Promotion,
Faculty of Education, University of Johannesburg,
Johannesburg, South Africa*

TingWen Chang

*Smart Learning Institute of Beijing Normal University,
Beijing, China*

Samia Chasi

*International Education Association of South Africa
(IEASA), Pretoria, South Africa; and University of the
Free State, Bloemfontein, South Africa*

Miriam Chitiga

Warner Pacific University, Portland, OR, United States

Cherie Chu-Fuluifaga

*Te Herenga Waka, Victoria University of Wellington,
Wellington, New Zealand*

Patti H Clayton

*Indiana University Purdue University Indianapolis,
Indianapolis, IN, United States; and PHC Ventures,
Cary, NC, United States*

Alison Cook-Sather

*Education Program, Bryn Mawr College, Bryn Mawr,
PA, United States*

Bri Damacio

*School of Education and Information Studies, University
of California, Los Angeles, CA, United States*

Dieter Dohmen

*FiBS Research Institute for the Economics of Education
and Social Affairs, CEO FiBS RILL Research Institute
on Lifelong Learning, Berlin, Germany*

Macey L Edmondson

*The University of Mississippi, Department of Higher
Education, University, MS, United States*

Robert A Ellis

*Department of Learning & Teaching, Griffith University,
Brisbane, QLD, Australia*

Serap Emil

Middle East Technical University, Ankara, Turkey

Johanna Ennser-Kananen

*Department of Language and Communication Studies,
University of Jyväskylä, Jyväskylä, Finland*

Jennifer M Eustis

*University of Massachusetts Amherst, Amherst, MA,
United States*

Frank J Fernandez

*The University of Florida, College of Education,
Gainesville, FL, United States*

Ourania Filippakou

Brunel University, London, United Kingdom

Jennifer Fischer

*Johannes Gutenberg University Mainz, Department of
Law and Economics, Chair of Business and Economics
Education, Mainz, Germany*

Sarah R Fitzgerald

*University of Massachusetts Amherst, Amherst, MA,
United States*

Trevor Gale

University of Glasgow, United Kingdom

Marybeth Gasman

*Samuel DeWitt Proctor Endowed Chair & Distinguished
Professor Rutgers University, Graduate School of
Education, New Brunswick, NJ, United States*

Henry A Giroux

McMaster University, Hamilton, ON, Canada

Hugh B Guthrie

*Melbourne Graduate School of Education, University of
Melbourne, Melbourne, VIC, Australia*

Linda Serra Hagedorn, PHD

Iowa State University, Ames, IA, United States

Feifei Han

*Institute for Learning Sciences & Teacher Education,
Australian Catholic University, Sydney, NSW, Australia*

Roger McL Harris

*Centre for Research in Educational and Social Inclusion,
UniSA Education Futures, University of South Australia,
Mawson Lakes, SA, Australia*

Ronghuai Huang

*Smart Learning Institute of Beijing Normal University,
Beijing, China*

Neal H Hutchens

*The University of Mississippi, Department of Higher
Education, University, MS, United States*

Dirk Ifenthaler

*University of Mannheim, Mannheim, Germany; and
Curtin University, Perth, Australia*

Lydia Ikwuemesim

University of Pretoria, Pretoria, South Africa

Ben Jongbloed

*Center for Higher Education Policy Studies (CHEPS),
Faculty of Behavioural, Management and Social
Sciences (BMS), University of Twente, Enschede, the
Netherlands*

Venise Joubert
SARCHI Teaching and Learning, University of Johannesburg, Johannesburg, South Africa

Ibrahim M Karkouti
The American University in Cairo, New Cairo, Egypt

Jane Knight
Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Betty R Leask
School of Education, La Trobe University, Melbourne, VIC, Australia

Jing Liu
Graduate School of Education, Tohoku University, Sendai, Japan

Matt Lumb
The University of Newcastle, University Drive, Callaghan, NSW, Australia

Suzanne Lyons
New York University, New York, NY, United States

António Magalhães
Centre for Research on Higher Education Policies (CIPES), Faculty of Psychology and Education Sciences, University of Porto, Matosinhos, Portugal

Mukovhe Masutha
Ali Mazrui Centre for Higher Education Studies, University of Johannesburg, Ackland Park, Johannesburg

Sarah McDonald
Education Futures, University of South Australia, Adelaide, SA, Australia

Kirti Menon
Division for Teaching Excellence, DST/NRF South African Research, Teaching and Learning (PSET), Faculty of Education, University of Johannesburg, Johannesburg, South Africa

Barbara Merrill
Centre for Lifelong Learning, University of Warwick, Coventry, United Kingdom

Shireen Motala
University of Johannesburg, Johannesburg, South Africa

Richard Muench
Zeppelin University, Friedrichshafen, Germany; and University of Bamberg, Bamberg, Germany

Muhammad Yasir Mustafa
Smart Learning Institute of Beijing Normal University, Beijing, China

Alina K Niipare
National Institute for Educational Development, Okahandja, Namibia

Ibrahim Ogachi Oanda
Council for the Development of Social Science Research in Africa (CODESRIA), Dakar, Senegal

Justine R Okerson
East Carolina University, Greenville, NC, United States

Matt O'Leary
Birmingham City University, CSPACE Research Centre, Faculty of Health, Education and Life Sciences, Birmingham, United Kingdom

Stephen Parker
University of Glasgow, United Kingdom

Penny A Pasque
The Ohio State University, Columbus, OH, United States

Venitha Pillay
University of Pretoria, Faculty of Humanities, Department of Sociology, Pretoria, South Africa

Scott Revers
Centre for Lifelong Learning, University of Warwick, Coventry, United Kingdom

Fazal Rizvi
University of Melbourne Australia, Melbourne, VIC, Australia

Taina Saarinen
Finnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland

Kenneth J Saltman
University of Illinois at Chicago, Chicago, IL, United States

Fatma Nevra Seggie
Boğaziçi University, İstanbul, Turkey

Chika Sehoole
University of Pretoria, Pretoria, South Africa

Richard J Shavelson
Stanford University, Graduate School of Education, Stanford, CA, United States

Sae Shimauchi
International Center, Tokyo Metropolitan University, Hachioji, Japan

Beatrice S Sichombe
Namibian College of Open Learning, Katutura, Namibia

Precious Bupe Sipuka
University of Johannesburg, Johannesburg, South Africa

Erica Smith
*School of Education, Federation University Australia,
Ballarat, VIC, Australia*

Garth Stahl
*School of Education, University of Queensland,
Brisbane, QLD, Australia*

Jason E Stone
*Educational Leadership and Policy Studies, Oklahoma
State University, Oklahoma City, OK, United States*

Tatiana Suspitsyna
*The Ohio State University, Columbus, OH, United
States*

Azeb Tadesse
*School of Education and Information Studies, University
of California, Los Angeles, CA, United States*

Ahmed Tlili
*Smart Learning Institute of Beijing Normal University,
Beijing, China*

Jafet S Uugwanga
*National Institute for Educational Development,
Okahandja, Namibia*

Linda van Ryneveld
*Graduate Centre, Hatfield Campus, University of
Pretoria, Pretoria, Gauteng, South Africa*

Amélia Veiga
*Centre for Research on Higher Education Policies
(CIPES), Faculty of Psychology and Education Sciences,
University of Porto, Matosinhos, Portugal; and Centre*

*for Research and Intervention in Education (CIEE),
Faculty of Psychology and Education Sciences, University
of Porto, Matosinhos, Portugal*

Carmen Wade
*Community College of Aurora, Aurora, CO, United
States*

Matthew R Wawrzynski
*Michigan State University, East Lansing, MI, United
States*

Everard Weber
*SARCHI Teaching and Learning, University of
Johannesburg, Johannesburg, South Africa*

Steven Weiland
*Michigan State University, East Lansing, MI, United
States*

Oliver Wieczorek
Zeppelin University, Friedrichshafen, Germany

Akiyoshi Yonezawa
*International Strategy Office, Tohoku University,
Sendai, Japan*

Liang Zhang
New York University, New York, NY, United States

Rongxia Zhuang
*Smart Learning Institute of Beijing Normal University,
Beijing, China*

Olga Zlatkin-Troitschanskaia
*Johannes Gutenberg University Mainz, Department of
Law and Economics, Chair of Business and Economics
Education, Mainz, Germany*

CONTENTS OF VOLUME 8

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 8: Higher Education Management, Policy, and Practice</i>	<i>xix</i>
<i>Contributors to Volume 8</i>	<i>xxv</i>

DEVELOPMENTS IN HIGHER EDUCATION

Higher education strategy formations in shifting landscape of higher education <i>Meltem Akbulut Yıldırım, Serap Emil, and Fatma Nevra Seggie</i>	1
Knowledge production in Europe: actors, policies and critiques <i>Tatiana Suspitsyna and Penny A Pasque</i>	12
Transformation of university missions—from nation-building to a focus on common global challenges reflections from East Asia, under the risk of being “lost in translation” <i>Akiyoshi Yonezawa, Sae Shimauchi, and Jing Liu</i>	21
Economic impact of higher education on students, corporations and society <i>Peter Arena, Tim Carroll, Suzanne Lyons, and Liang Zhang</i>	29
Academic capitalism and market thinking in higher education <i>Oliver Wieczorek and Richard Muench</i>	37
“Reschooling” higher education: the dynamics of curriculum in South Africa <i>Kirti Menon and Gloria Castrillón</i>	48
New pathways in education <i>Dieter Dohmen</i>	57
Open access and the unfinished transformation of scholarly communications <i>Steven Weiland</i>	63
Innovations in university student learning research into blended course designs: incorporating self-report and observational data <i>Robert A Ellis, Feifei Han, and Ana-Maria Bliuc</i>	74

Uses of technology in navigating the pandemic: the case of African higher education: Chika Schoole, Lydia Ikwuemesim, Olaide Agbashe <i>Chika Schoole, Lydia Ikwuemesim, and Olaide Agbashe</i>	83
The impact of COVID-19 on higher education in Africa <i>Philip O Ayoo</i>	91
Issues of production and utilization of knowledge: multiple uses of knowledge <i>Azeb Tadesse, Bri Damacio, Kiana Billot-Vasquez, and Walter R Allen</i>	102
Regimes of research and development funding in higher education <i>Bailey A Brown</i>	108
The hollowing out of academic freedom from African universities; from external capture to internal implosion <i>Ibrahim Ogachi Oanda</i>	114
The changing functions of libraries in higher education <i>Sarah R Fitzgerald and Jennifer M Eustis</i>	124
Uses and abuses of social media in higher education <i>Lillian Vega Castaneda and Mario E Castaneda</i>	131

POLICY, GOVERNANCE AND LEADERSHIP

Higher education, critical pedagogy, and the challenge of resistance <i>Henry A Giroux and Ourania Filippakou</i>	139
About the rise, the characteristics and future of New Public Management in Higher Education <i>Bruno Broucker</i>	145
Governance, institutional autonomy and leadership in Portuguese higher education <i>António Magalhães and Amélia Veiga</i>	152
Publish or perish <i>Hanne Andersen</i>	158
Internationalization in English? Problematizing the role of language in higher education internationalization <i>Johanna Ennser-Kananen and Taina Saarinen</i>	167
Academic freedom protections in national and international law <i>Neal H Hutchens, Macey L Edmondson, and Frank J Fernandez</i>	174
The funding of higher education: international perspectives <i>Ben Jongbloed</i>	183
Faculty compensation and salary equity <i>Linda Serra Hagedorn PHD</i>	191

INTERNATIONALISATION

Knowledge diplomacy: the role of international higher education, research and innovation in international relations. <i>Jane Knight</i>	202
North–South partnerships in higher education <i>Samia Chasi</i>	210

Reimagining internationalization of the curriculum <i>Betty R Leask</i>	220
International higher education and the commodification of student mobility <i>Fazal Rizvi</i>	233
Postcolonialism in the movies and higher education: a case study of the narratives of South African and international students at the University of Pretoria <i>Everard Weber and Venise Joubert</i>	241
CURRICULUM, TEACHING AND LEARNING	
Neoliberal education: from capitalizing on disaster to critical pedagogy <i>Kenneth J Saltman</i>	250
Linking research and teaching to benefit student learning in higher education: complexity, collaboration and a tailored approach <i>Matt O'Leary</i>	258
Reciprocal teaching and learning through student-faculty pedagogical partnership <i>Alison Cook-Sather</i>	268
Implications of the Fourth Industrial Revolution in the vortex of higher education in South Africa <i>Narend Baijnath, Shireen Motala, and Kirti Menon</i>	279
Toward a framework for the implementation of fully online distance education at a residential higher education institution <i>Linda van Ryneveld</i>	287
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis <i>Ronghuai Huang, Muhammad Yasir Mustafa, Ahmed Tlili, Rongxia Zhuang, TingWen Chang, and Daniel Burgos</i>	296
Assessing student learning outcomes <i>Matthew R Wawrzynski</i>	305
Advantages and challenges of performance assessment of student learning in higher education <i>Olga Zlatkin-Troitschanskaia, Jennifer Fischer, Henry I Braun, and Richard J Shavelson</i>	312
Adoption of learning analytics <i>Dirk Ifenthaler</i>	331
The intricacies of class size in higher education <i>Justine R Okerson</i>	336
Credit for experience, credit recognition and transfer <i>Jason E Stone</i>	343
The disruption of Black history and critical pedagogy in international higher education curriculum <i>Miriam Chitiga, Kolawole Samuel Adeyemo, and Carmen Wade</i>	350
Non-completion amongst black working-class students in South African universities: the dangers of a single story <i>Mukovhe Masutha</i>	356

DIVERSITY, INCLUSION, AND EQUITY

Tensions in access and accountability <i>Penny Jane Burke, Matthew Bunn, and Matt Lumb</i>	367
The problem with higher education student completions in the global north <i>Trevor Gale and Stephen Parker</i>	374
Women in higher education: exclusion, inclusion and seclusion <i>Venitha Pillay</i>	381
The first-in-family experience: gendered and classed subjectivities in elite higher education <i>Garth Stahl and Sarah McDonald</i>	392
Minority learners, diversity, and educational outcomes: the case of Aotearoa, New Zealand <i>Cherie Chu-Fuluiifaga</i>	398
Experiencing class and gender: mature women students and their journeys from HE to the labor market <i>Barbara Merrill and Scott Revers</i>	406
Deconstructing predominantly White cultures: creating inclusive learning environments in the United States <i>Ibrahim M Karkouti</i>	417
Historically Black colleges and universities: historical and contemporary issues <i>Marybeth Gasman</i>	423
"It was a life enhancing activity": the academic habitus of PhD students <i>Tehmina N Basit</i>	429

LABOUR MARKETS AND VOCATIONAL EDUCATION AND TRAINING

Pathways, labor markets and tertiary education systems <i>Roger McL Harris and Hugh B Guthrie</i>	438
Learning through work: augmenting vocational educational experiences to meet 21st-century work requirements <i>Stephen Billett</i>	452
The future of internships and workplace-based learning <i>Leopold Bayerlein</i>	462
Career pathways <i>Debra D Bragg</i>	469
Higher education: service-learning as pedagogy, partnership, institutional organization, and change strategy <i>Robert G Bringle and Patti H Clayton</i>	476
Thirty years of competency-based training: how Australia painted itself into a curriculum corner in vocational education and training <i>Erica Smith</i>	491
Exploring the professional identity of university academics: "becoming" a university teacher <i>Precious Bupe Sipuka and Shireen Motala</i>	504
Early career academics in Namibia: experiences, prospects and challenges <i>Beatrice S Sichombe, Alina K Niipare, and Jafet S Uugwanga</i>	513

Higher education strategy formations in shifting landscape of higher education

Meltem Akbulut Yıldırım^a, Serap Emil^b, and Fatma Nevra Seggie^c, ^a Technical University of Munich, Munich, Germany; ^b Middle East Technical University, Ankara, Turkey; and ^c Boğaziçi University, İstanbul, Turkey

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Trends in the higher education strategy formation	1
Strategy formation as a double-edged process: two perspectives	1
Strategy formation across world regions	3
Essence of the modern university and strategy formation: a mélange	7
References	9
Relevant websites	11

Introduction

Universities being one of the oldest and resilient institutions have been continuously evolving with the developments and turning points of each era in the history (Scott, 2006; Young and Pinherio, 2022). There have been some changes moving institutions forward, while there are some developments limiting the capacity of higher education. Thus, it is almost urgent to examine the nature of the shift in the higher education landscape to characterize the strategy formation in today's world. The core functions of higher education have become convoluted due to rankings, grants, measurable performance indicators, business management perspectives, and most importantly, its contribution to economic development. On the one hand, higher education institutions (HEIs) are still responsible for producing scientific knowledge, educating students, and serving the community to uphold their core functions. On the other hand, glancing over the strategy formation with a critical eye, it is seen that worldwide competition fetish is existent in higher education systems directed by the shamans—macro or micro actors—as Naidoo (2018) defined. These actors are governments, international organizations, or political actors with the agenda of harsh competition for HEIs (Naidoo, 2018). Yet, we strongly agree with UNESCO Assistant Director-General for Education Stefania Giannini's statement from the most recent UNESCO World Higher Education Conference 2022: "competition is important when it is about improving our institution, but it is not about putting in the rankings and having a competitive approach to the common public good" (O'Malley and Warden, 2022, para. 18).

Framed from this perspective, our argument creates possibilities of thinking about the conventional wisdom that universities strive for the truth, and that's their *raison d'être*. What is gaining ground, though, is that the characteristics of today's strategy-making have put them on a trajectory wherein they are becoming estranged from their criticality in the search for and production of truth. The intricacy of this issue lies in its direct bearing on how HEIs reestablish their core functions while living at the heart of and growing out of a variety of social, political, cultural, and economic factors. With this in mind, we analyze how higher education strategy formation processes operate in the higher education landscape and their implications on the criticality of HEIs in the search for and production of truth through cross-regional triangulation of strategy development practices. Based on these lines of analytical framing, in what follows, we first present the background of our argument by examining the nature, scope, and scale of the trends that operate in the formation of higher education strategies. Second, we discuss the nexus between these trends and the essence of the modern university before suggesting the study of the essence of the modern university as a mode of analysis for future research.

Trends in the higher education strategy formation

The central question driving our literature review is twofold: where, why, and how strategies originate in higher education and who is involved? We conduct this inquiry from the standpoint that calls into question the essence of the modern university. We aim to explore the existing complexities and countervailing patterns that shape the higher education strategy formation processes. To this end, in the remainder of this section, we unfold in broad terms the relationship between higher education and economic growth and consider in more operational terms how strategy formation in higher education manifests itself across the world.

Strategy formation as a double-edged process: two perspectives

Strategy, traditionally defined as explicit, intentionally developed in advance of specific decisions, is a set of guidelines determining the action of a given institution. It is crucial to note that strategies can be *a priori* or *a posteriori* affecting decision-making (Mintzberg, 1978). Yet, the strategy formation in today's world is not a straightforward process, for the nature of HEIs is multi-layered and complex. Resilience in universities (Young and Pinherio, 2022) allows HEIs to protect their essential function and identity while they are engaged in these complex and ever evolving change initiatives in relation to external forces. This is a rather chaotic process

where the pressure of external forces is challenged by the fundamental characteristics of the university that have been shaped throughout history.

These external forces attributed to HEIs, such as being entrepreneurial and innovative or producing high employability and reputation, are bound to complicate how to reinscribe the role of the university in societies. Some of the underlying factors for these trends can be stated as globalization, knowledge economy, or internationalization, though, in essence, it is the link between economic growth and education, as claimed by many. Accordingly, based on the extant literature, two main perspectives are at issue in the debate over higher education (Schwartzman, 2020). One of them—mentioned above—can be called from an economic perspective focusing on economic development around the globe. The other one focuses on education—the true essence of the university—“complex processes of social transformation and institutionalization” (Schwartzman, 2020, p. 10). In order to contextualize some of the practices adopted in the higher education strategy formation process that we explore in the second part of this section, we now attempt an overview of these two perspectives.

The economically oriented perspective, which has received sustained attention from economists and international agencies, analyzes higher education in terms of its possible contribution to economic growth and development. There are different ways HEIs can contribute to this perspective through research and technological innovations and through teaching graduates with capabilities for employment and to increase productivity (Oketch et al., 2014). There is also the human capital theory dominating the discussion on higher education outcomes since the 1960s (Marginson, 2017). By general agreement, Shultz (1963) and (Becker, 1993) mark the inauguration of a renewed interest in education as an investment in capital, encouraging the development of an economic orientation toward higher education in the 1960s and ensuing decades. This human capital theory, with its focus on “activities that influence future monetary and psychic income by increasing the resources in people” (Becker, 1993, p. 11), has had a substantial impact on education and economic growth relationship, more so on higher education with the advent of mass higher education in the 1960s, along with some suspicions on the concept of the academy (e.g., Becker, 1993) reaching to the present day (e.g., Scott et al., 2017). Education and training were considered the most essential investment in human capital (Drucker, 1993) which also is defined as the economic approach to human behavior (Becker, 1993) assuming an educational policy consensus (Lauder et al., 2006, pp. 8–10). The link between education and development for developing countries is strongly advocated (Ayers, 2000) and the human capital theory, within this context, becomes instrumental (Eide and Showalter, 2010; Oliver, 2004).

Birnbaum (1999), however, criticizes that there is little evidence to support this view (see also Enders, 2010) and stresses the importance of a collective initiative of higher education leaders in reorienting discussions about the purposes of higher education. According to him, developing new narratives or revitalizing the old ones demands connecting to the public. This point of view is not far removed from the suggestions in critical education policy scholarship. For example, Ozga and Lingard (2007) suggest the creation of research communities to resist the centrality of the all-pervasive “research steering” in knowledge production (pp. 78–79). Rizvi (2006) makes a plea to researchers to collectively “develop an alternative educational policy agenda,” “an alternative social imaginary committed to strong democracy and global justice” (p. 203) against the narrative supremacy of the “neo-liberal imaginary” (p. 200). According to him, the processes of (1) international cooperation, (2) global circulation of ideas, (3) international conventions, and (4) coercive strategies (pp. 201–204) work in collaboration to feed neo-liberal policies, attributing transnational, international, and global dimensions to national educational policies. As researchers, in his view, we need to analyze these processes to understand how to locate local and national educational policies within this multidimensional approach to education. Various reconfigurations of these processes and hence diffusion of authority at all levels are evident in the landscape of higher education strategy making today. Examples of current processes are discussed in the second part.

In the second perspective, education is sociological in orientation and approaches the role higher education plays in societies from the vantage point of its relation to social transformations, its educational and socially progressive prospects rather than from the restricted domain of economic development. Higher education, in its many varieties—private institutions, universities, vocational schools, and similar bodies—is perforce in the borderlands between being perceived as a credentialing endeavor and being admired as a stronghold of the intellectual progression by societies with which they evolve in tandem. Even as it has taken various institutional forms since the foundation of the first modern university in 1088, the University of Bologna, this chasm—if such it is—is still kept alive for the most part by the complexities ingrained in this long historical process of bridging relations with wider society (see also Rüegg, 2004). In 1983, a pioneering work by Burton Clark casted some light on the connections between higher education systems and change, while on the other hand, it has encouraged further exploration of the governance of higher education (e.g., Salazar and Leihy, 2013). As Lang (2017) argues, one can also use Clark’s Triangle of Coordination including three forces—“state authority,” “academic oligarchy,” and “market” (Clark, 1983, p. 143)—as a comparative tool for understanding institutional change within a system compared to other systems.

A crucial moment in understanding the links between higher education and social transformations came in 1988 with the Magna Charta Universitatum. The Magna Charta, endorsed then by 400 representatives from around the world, stands as the first universally accepted charter of the university as an institution. Acknowledging the challenge of conciliating the views of institutions with diverse geographies, histories, and cultures, Stephen Lay, in his *The Interpretation of the Magna Charta Universitatum and its Principles*, suggests adding a civic role to the principles of the Magna Charta to guide the university administration (Lay, 2004, pp. 107–111). In his view, this is a vital task of the university to get a foothold in the deep-seated challenge of engaging with society given the transformations such as globalization, knowledge economy/society, internationalization, and the like (Lay, 2004).

The Magna Charta Universitatum has recently been reviewed to sustain the positive contribution of higher education to social transformations. As of October 18, 2021, 339 out of 947 universities from 65 countries have signed the new version of the Magna

Charta. While remaining committed to its original fundamental principles—institutional autonomy, freedom of inquiry, and indivisibility of teaching and research—the Magna Charta Universitatum 2020 now includes a clear statement on the civic role and responsibility of the university under “Principles, Values and Responsibilities” as follows:

Universities have a civic role and responsibility. They are part of global, collegial networks of scientific enquiry and scholarship, building on shared bodies of knowledge and contributing to their further development. They also are embedded in local cultures and crucially relevant to their future and enrichment. While they are immersed in and connected with global developments, they engage fully with and assume leading roles in local communities and ecosystems (p. 2).

Supporting this civic role of higher education, Paul Ashwin, in his recently published book, *Transforming University Education: A Manifesto*, exactly pulls the audience’s attention from higher education as a commodity to an institution contributing to the development of students from an educational sciences perspective. He rebukes the dominant narrative as economic arguments and invites us to revisit the role of the university in transforming the students’ lives (Ashwin, 2020). He, then, elaborates the purpose of university education as student development through the transformational relationship with disciplinary or professional knowledge.

What the prospects might be going forward with these educational perspectives remain yet to be seen. One can argue that the unfolding of the two perspectives on higher education points to a double-edged relationship between higher education and economic growth which looms large in the modern university strategy formation, complicating the roles of the university and charging them with a diverse set of tasks. Perhaps what is declared under Values, Principles, and Responsibilities in the Magna Charta is a call to galvanize support for proactive ownership so that we can cut a familiar path across.

Strategy formation across world regions

To comprehend higher education in the modern world, examining the national, international, (bi-) regional, and transnational networks or associations would allow scholars to reframe the strategies for HEIs. To this end, this section provides examples of the trends in higher education strategy-making from five world regions and also inter-regional agencies in no particular order, while imperfect, following the UN designations: Africa, Americas, Asia, Europe, and Oceania.

Higher education in North America is closely linked up with the North American Free Trade Agreement (NAFTA), now known as the United States-Mexico-Canada Trade Agreement (USMCA). NAFTA, signed between Canada, Mexico, and the United States in 1992, came into effect in 1994. The trilateral trade agreement created a free trade and investment region among the three countries of North America through the elimination of most tariffs on products and also aspired to protect intellectual property rights and to foster fair competition (SICE the OAS Foreign Trade Information System, 2021). In 2020, three member countries accounted for 28% of the world’s gross domestic product (GDP), with a combined GDP of US\$23,6 trillion and a combined population of more than 496 million peoples (The World Bank and World Development Indicators, 2020a,b). Within the context of higher education, the three-country accord aimed to play a bridging role not only in stimulating cross-border academic collaboration, cooperation, and exchange, but also in eliminating obstacles to such cooperation and exchange via a joint leadership of the three countries. To this end, a Trilateral Steering Committee (TSC) was established in 1991 to guide the regional cooperation process in higher education (Mallea et al., 1997). Another ripple effect of NAFTA is the development of trilateral associations such as the Consortium for North American Higher Education Collaboration (CONAHEC) and the North American Institute (NAMI). Illustratively, CONAHEC, established in 1994, seeks to advance the human capital formation in the North American region and currently has around 180 members. While NAFTA partners form its core membership, institutions from around the world may also join the network based on the defined criteria. Among its activities are organizing conferences, coordinating academic exchange and mobility programs, helping members build international partnerships as their liaison, and publishing on matters of concern.

Barrow et al. (2003) argue that the TSC, CONAHEC, and NAMI are indicative of the institutionalization of NAFTA’s economic orientation toward higher education. In their view, economic concerns of the deal such as workforce development and building global competitive advantage over the European Union, Asia, Japan, and Australia through, say, technology transfer and flow of academics and students across the region have gained ground in the region thanks to the efforts of these three institutions while efforts targeted at spurring inter-cultural exchange has remained to a large extent meager. That economic concern is also evident in the 2000 report of the three national associations of universities—The American Council on Education (ACE); la Asociación Nacional de Universidades e Instituciones de Educación Superior [the National Association of Universities and Institutions of Higher Education] (ANUIES); the Association of Universities and Colleges of Canada (AUCC) (renamed itself as Universities Canada in 2015).

The report, *Increasing North American Higher Education Cooperation, Collaboration, and Exchange Through a Trilateral Education “Marketplace,”* prepared with support from their respective governments and submitted to the Trilateral Steering Committee, defines “a virtual marketplace” for higher education, goal of which is “to serve as ‘one-stop shopping’ for North American institutions interested in collaboration and exchange in higher education” (Collins, 2000a, p. 1, emphasis in original). For, “a virtual marketplace” is “the most cost-effective, productive, and sustainable approach to fostering North American interchange” (Collins, 2000b, p. 3). Consequently, it can be argued that, in the years since NAFTA, higher education’s relation to the Northern American region has mostly evolved from an instrumental lens. NAFTA was restructured recently and, building on NAFTA, the new United

States-Mexico-Canada Trade Agreement (USMCA) took effect on July 1, 2020. The ripple effects of the renewed deal remain to be seen, yet the emerging documents, with their economic argument relying on higher education for workforce, skills, and technology development, point toward a continuation of the market-based strategy-making trend in higher education within the USMCA frame (e.g., Santos and Wilson, 2021).

Australia and New Zealand, on the other hand, carry an outsized weight in higher education strategy formation in Oceania. Besides geographic isolation that impacts the policy settings of all the Oceanian countries, demographic disadvantages, low affordability, often poor education quality and educational facilities at all levels, lack of employment opportunities, threats of natural disasters, and dependencies on certain sectors—fisheries, agriculture, and tourism—form policy fundamentals within other Oceanian countries according to the 2021 *UNESCO Science Report* (Scott-Kemmis et al., 2021, pp. 707–708). To address these pressing policy issues, the Pacific Islands Forum, one of the ten intergovernmental organizations constituting the Council of Regional Organizations of the Pacific (CROP) and the body focused on economic and political concerns, endorsed the Framework for Pacific Regionalism in 2014 with its 18 members including Australia and New Zealand (Pacific Islands Forum Secretariat, 2014).

Yet, whether the implementation of the Framework has been effective is highly contested in the context of higher education. For instance, the 2021 *UNESCO Science Report* highlights that no legal guidelines for research exist in about half of Pacific Island countries and projects a continuity in limited budget allocation to research capacity building in many Pacific Island countries (Scott-Kemmis et al., 2021, p. 708). In line with this finding, Jacob et al. (2018) argue underdeveloped higher education opportunities as one of the root causes of migration from these Oceanian countries to Australia and New Zealand. Furthermore, Asia as a region of interest weighs heavy on the policy agendas of both Australia and New Zealand. The signing of the Regional Comprehensive Economic Partnership (RCEP) between the ten member states of the Association of Southeast Asian Nations (ASEAN) (Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, Vietnam) and its five free trade agreement (FTA) partners (Australia, China, Japan, New Zealand, and Republic of Korea) on November 15, 2020 connects about 30% of the world's population and GDP, forming one of the world's largest free-trade pacts. Whether the new agreement will enhance the existing disparities in higher education in Oceania will hinge on how both nations balance the implementation of RCEP and the Framework for Pacific Regionalism. Nonetheless, the recent research hints at the possibility of increasing the incentives for Asia engagement.

To exemplify, a key higher education strategy development trend in Australia concerns how to navigate higher education reform in light of the rise of Asia in the region and beyond. The Australian Government's White Paper, *Australia in the Asian Century*, is the most recent comprehensive report elucidating Australia's Asia engagement vision and the ways in which Australia's geographic location and connectivity with Asia affect Australia's future prosperity in the strategically contested Indo-Pacific region (Australian Government, 2012). Accordingly, the report situates this shift of power to the region as fundamental to Australia's policymaking and provides an account of how to build Asia-relevant capabilities. Placing higher education among strategic areas in Australia's vision of Asia engagement, the report suggests "every Australian university to have a presence in Asia" (p. 174) through, for example, encouraging in-Asia study by Australian students; developing partnerships and exchange arrangements; and encouraging the take-up of Asian studies and some priority Asian languages, such as Chinese, Japanese, and Indonesian, by improving the curriculum and pedagogy (pp. 171–175). At this point, however, it is worth noting that regardless of the appeals of the Australian government to the whole of Australia to increase Asia proficiency, deficiency of Asia capabilities is still a salient policymaking challenge in all sectors of Australia (e.g., Henry, 2016; Department of the Prime Minister and Cabinet, 2019). Henry (2016) points to the narrow lens of international competitiveness as to why Australia has continued to underinvest in Asian capabilities. As an example, Rizvi (2021), based on the perceptions of Chinese and Indian students in Australia, argues that a market-oriented strategy-making in universities leaves higher education in the human capital formation conundrum, subordinating the exchanges among students and formation of transborder communities. Hence, one can conclude that as higher education in Australia and New Zealand are disproportionately far ahead in the structuring and distribution of capabilities and power, the strategy development processes of the two countries are more likely to steer the course of the functioning of higher education in Oceania.

Regional partnerships are of strategic importance in higher education in Europe. The Bologna Process, which came into effect in 1999 with a targeted initiative to create a common higher education area and market—the European Higher Education Area (EHEA)—by 2010, is a notable example in this regard. Viewed from this perspective, inter-regionalism in European higher education strategy formation functions as a "strategically selective terrain," "bypassing existing scalar fixes, such as those of the nation state and its Westphalian claim to territory and sovereignty" (Robertson, 2010, p. 34). The EHEA operates today in 49 countries reaching far beyond the European Union (EU) member states and is a voluntary international exercise involving the European Commission as a member, yet being party to the European Cultural Convention is a precondition for the EHEA membership. A thorny question about this interregional agreement is the extent to which a voluntary approach benefits the implementation of the objectives set out by the EHEA Bologna Process, for progress to this end is slow and requires the development of new strategies (Zgaga, 2020). According to Bergan (2015), a perceived decline of political interest is likely to explain the low engagement on the side of the EHEA stakeholders. Indeed, what seems to be high on the political agenda of European leaders are two other regional initiatives—the European Universities Initiative and the European Education Area (EEA).

The European Universities Initiative is an emergent flagship initiative developed to build an EEA limited to EU member states. A close look into the strategy-making processes of these initiatives might serve to glean the Union's role within the context of higher education. In the March 2017 *White Paper*, the European Commission positioned education policies as critical to Europe's competitiveness and resilience at a global scale (European Commission, 2017a). In the *Strengthening European Identity through Education* report presented to European leaders in the November 2017 Gothenburg Summit, the Commission with a vision to build

a European Education Area (EEA) by 2025 recommended the building of “world-class European Universities” (European Commission, 2017b, p. 11). As a follow-up, in the December 2017 Conclusions of the Gothenburg Summit, European leaders set an objective to create “some twenty ‘European Universities’” by 2024 (European Council, 2017, p. 3, emphasis in original). In response, the Commission co-developed the concept of European Universities with HEIs, students, and member states as “ambitious transnational alliances of higher education institutions developing long-term structural and strategic cooperation” (European Commission, 2019, p. 1). To date, 41 European University alliances have been selected. The fact of the matter is that the idea of a European university is a source of deep-seated discussion, as also pointed out by Gunn (2020), and it still is not clear how well the Europeanization-inspired reforms speak to the national contexts.

Beyond the EU as a regional global actor, there are numerous national/international associations of universities within Europe. And some are relevant players not only in negotiating interests in strategy-making but also in turning strategies into policies through regions. In this respect, four illustrative examples are the European Association of Institutions in Higher Education (EURASHE), European University Association (EUA), Association of the Carpathian Region Universities (ACRU), and Universities UK (UUK).

Founded in 1990, EURASHE currently represents 67 HEIs of every type and size from 40 countries within and beyond the EHEA. Mainly concerned with issues related to professional higher education and positioned to build strategic alliances between HEIs and the labor market, the Association undertakes activities, including policy development, participation in joint research projects, organization of conferences and networking events, and advocacy of members’ interests in their relations with European and national policymakers. It works closely with other stakeholders in the EHEA, pursues the development of the EHEA and the European Research Area (ERA), and its membership is open to those who can commit to the EURASHE policy principles and promote the EURASHE working agenda that accepts the EHEA framework. Another example of comprehensive associations of HEIs with policy-influencing powers is EUA, whose members are individual European universities of any type, national associations of universities, university alliances, national rectors’ conferences, and institutions with common interests. Founded in 2001, the Association represents more than 850 members from 48 European countries in, like EURASHE, various policy settings such as European institutions, the Bologna Process, and the Council of Europe. It also engages in activities similar to those of EURASHE, is in support of the European Universities Initiative, and operates within the EHEA framework. Different from EURASHE and EUA, The Association of Carpathian Region Universities.

(ACRU) is more of an intra-regional association of universities. Established in 1994, the Association comprises six countries of the Carpathian region across East-Central Europe—Hungary, Poland, Romania, Serbia, Slovakia, and Ukraine—and now has 23 members (one from Hungary, one from Serbia, two from Slovakia, four from Poland, seven from Romania, eight from Ukraine). Any university beyond the Carpathian region may join its membership whose composition involves associate and cooperating partners from Prague, the United States, and the United Kingdom. ACRU is distinctive in that it seeks to promote the development of the Carpathian region and its cultures and languages through, for example, academic mobility, forums and workshops, and cooperation with local authorities and nongovernmental organizations (NGOs). UUK, on the other hand, is representative of a comprehensive national university association with international standing. Established at the meeting of 22 university leaders in 1918, UUK now represents 140 universities across the UK Universities Wales and Universities Scotland are part of UUK and autonomous in their operations. The Association contributes to the issues such as quality, funding, research and innovation, and internationalization through the Universities UK International (UUKI) team. Its core activities are raising the profile of the universities, providing guidance, influencing policy decisions through political advocacy, and advancing UK higher education at all levels.

In addition to the well-established higher education systems examined above, there are numerous associations and cooperation, including the universities in other continents. For instance, in Latin America and Caribbean higher education, the World Bank report from 1997 highlighted the importance of HEIs not only for economic developments but also for promoting democracy and space for critical thinking (de Moura Castro and Levy, 1997). The institutions in the region back then were reported as changing too little or too slowly while they were struggling with meeting the needs of students or the job market. Coming into 2000s, the increasing student enrollment in line with national economic restructuring was observed (Holm-Nielsen et al., 2004). They stated that similar to Africa, the Latin American and Caribbean countries have experienced a higher demand for higher education in the 21st century and been challenged by the restrained resources. Between 1985 and 2002, the development of private higher education has been observed in numerous countries like Brazil, Chile, Colombia, Peru and El Salvador, Argentina and so on (Holm-Nielsen et al., 2004).

The Latin American and Caribbean universities, gathered, for example, under the Association of Universities of Latin America (UDUAL) established in 1949, and the Caribbean was added later in 2004. They support academic exchange, mobility, recognition, and transferability and ensuring the quality of higher education. The mission of the association is “to drive academic and public politics that encourage the higher education institutes’ essential activities, considering their key connection with different regional realities, in order to be a strategic factor in the sustainable, democratic and inclusive development” (UDUAL, n.d., p. 4). UDUAL being the largest network of the HEIs in the region positions itself as the interlocutor between HEIs, governments, and society. Different from the other bodies mentioned above, it aims to facilitate an effective dialog for public politics. Even though the similar trends such as international cooperation, or HEI transversal perspectives drive the actions of the association, they tend to concentrate on some core issues—interculturality, gender equality, online education, digital society—linked with each university in their network. It is also important to note that they defend the university autonomy and freedom in academic, research, and institutional management. Some of the strategic projects of UDUAL are The Council for International Assessment and Accreditation (CEAI in Spanish); The Academic Program of Student Mobility (PAME in Spanish); The Common Space for Online Higher Education (ECE-SELL, in Spanish); Master’s Degrees, Massive Open Online Courses (MOOCs) and certification courses.

As a result of the Alfa PUENTES project, beginning from 2011, The Latin American Higher Education Area collaborated with European and Latin American University Associations by a bi-regional management team. A total of six European Associations and 14 Latin American university associations came together to deepen and widen the internationalization of higher education. The joint activities among the universities from Latin America and Europe worked toward quality assurance, transferability of degrees, and academic mobility (Kind and Colucci, 2014).

As an example of the economic argument, we can take a look at the Asian Development Bank report (2011), which presented the issues and strategies for Asian universities. The main roles of the HEIs mostly circle around the need for technical and administrative personnel in addition to its services as the incubator of innovation. The report involves target countries (Cambodia, Lao PDR, Mongolia, Viet Nam, Indonesia, Malaysia, Philippines, Sri Lanka, Thailand, India, and PRC) and the reference economies (Hong Kong, China, Singapore, Japan, and Republic of Korea). The challenges are defined more in line with the rest of the world, such as increasing enrollment, demand for research, and financial limitations over and above the decline in HE quality. The resolutions to these pressures are grouped within five areas as internal efficiency, external efficiency, cost and financing, administration and governance, finally access and equity that covers all the existing issues of higher education. However, the report states that the biggest needs in higher education are less capital intensive and more about the university leaders, faculty members, and national and institutional level policies with greater collaboration among those.

ASEAN, the Association of Southeast Asian Nations, with 10 member states formed in 1967 in Thailand, is an important regional cooperation to have an impact on economic growth, social progress and cultural development. While all aspects of economic, political, and socio-cultural development are considered by the association, higher education is also given particular attention. One of the key documents, ASEAN Declaration on Higher Education (n.d.), defines higher education as the catalyst of development through the following principles. Higher education being accountable for higher quality contributes to the wealth and well-being of people through synergy created among governments, industry, community, and academia. It also fosters entrepreneurial endeavors, mobilities, and visibility of the academic community in the global arena. Finally, in the declaration, higher education is given the responsibility of supporting sustainable development via its curriculum, research, community service, and stakeholder engagement.

Asian Universities Alliance (AUA)—a higher education body—was formed with 15 universities in 2016 by the proposal of Tsinghua University for the purpose of student mobility, research collaboration, strategy and policy and annual evaluation of Asian higher education. These works were designed around three themes—Artificial Intelligence & Data Science, Asian Civilizations, and Energy & Sustainability—representing the regional and global challenges. The AUA aims to be the platform for the “innovative collaboration ecosystem” among academia, government, and industry. Another body called Association of Asian Universities was established in 2013 with 29 universities from 8 countries. Similar to the AUA, this association lands itself into the creation of the joint Asian educational space, which focuses mostly on diversity and internationalization aspects of higher education. Thus, it regards the exchange and mobility programs among the member universities. Association of Universities of Asia and the Pacific (AUAP), on the other hand, is the oldest union established in 1995 initiated by the Suranaree University of Technology, Thailand. The main mission is similar to the others, promoting the culture of quality, innovation, and research in the Asia and the Pacific region. With the widest network with 161 member universities, the AUAP have been focusing on student and faculty mobility, faculty professional development, internationalization, global leadership through which they intend to learn and share the best practices with other universities.

Looking at the African continent, we see the Association of African Universities (AAU), located in Ghana, has been defined as “a platform for research, reflection, consultation, debates, cooperation, and collaboration on issues pertaining to higher education” (AAU, 2021, para. 4). It was formed by 34 public universities in 1967 in Rabat and has 320 members now. The Association aims to increase the quality of higher education in Africa and build networks among member institutions. Looking at their strategic goals, it advocates for an effective voice in national, regional and international levels. Ensuring the quality of HEIs in Africa seems to be the main focus, yet there are not many details on how or guidelines provided as a roadmap. Fredua-Kwarteng (2020), on the other hand, specifically discusses the strategic planning work in African higher education systems stating that the national context is necessary to consider even though strategic plans are answers to the challenges in the external environment. He raises an important question for African higher education as “How is strategy for an institution conceptualized or perceived in the African cultural context?” (Fredua-Kwarteng, 2020, p. xx).

Dei et al. (2019) utilized the philosophical approach of methodic doubt to examine the challenges and opportunities of African universities in the 21st century. Similar to any other higher education system, tertiary institutions are in the center for the development of Africa’s human capital and the enhancement of the African identity. As a result of the long history of colonialism, the dominance of the western model of higher education (Teferri and Altbach, 2004), Dei et al. (2019) state that the western philosophies of education can be blended with the indigenous education system. However, there are doubts about this in terms of African higher education producing its own knowledge systems to be able to compensate for the effects of colonialism.

Latin America, together with Africa, are the two continents identified as having the fewest top quality research-intensive universities in accordance with the world university rankings (Altbach and Salmi, 2021). This, by itself, may be seen as the basis of new innovative practices in HEIs in these regions. Looking at the suggestions made by Altbach and Salmi (2021), the second Cordoba reform is hoped to bring the Latin American universities’ level to the World-class universities. Increased public funding aligned with accountability, commitment to national and social development, modern governance systems, less student population, interdisciplinary and international dimensions to the university systems are some of those suggestions to increase the number of top research-intensive universities (Altbach and Salmi, 2021).

The multinational bloc—BRICS (Brazil, Russia, India, China, and South Africa)—also involves in the pact between higher education and society as Schwartzman et al. (2015) call. Altbach and Bassett (2014) stated that BRICS may make sense in terms of economic development but it is an irrelevant grouping to consider the higher education systems of these countries. Even though the academic traditions of these countries are quite different and there is little collaboration, BRICS countries also experience higher demand for higher education all around the world and are also experiencing this increase together with limited financial resources (Tilak, 2013). While the expansion is happening in the private sector, similar to most higher education systems, the public sector is treated with cost-recovery strategies and narrowing down consequently. The expansion has been focusing more on market-driven areas of study such as engineering or management. In the end, Tilak (2013) sees this process in line with the commoditization of higher education. The quality debate in this process is associated with the number of universities in world university rankings—the elite institutions. While the pact has led to economic competitiveness, the success of higher education has heavily been defined by the number of world-class universities from the BRICS countries. There is also a particular emphasis on science and technology that universities in respected countries are aligned with in terms of their research and curriculum (Chang et al., 2018).

We should not also forget these macro level actors—World Bank, OECD, IMF, UN, UNICEF, UNESCO—shaping influential policies on higher education, particularly for developing countries and fund related initiatives in relation to these policies. They generate key statistics and reports on some of the very critical issues of higher education in accordance with the regional developments. These institutions not only present the data on the various aspects of higher education systems but also try to influence the policymaking and strategy formation processes. While these works are valuable for the systems, they are also coming from one source with a particular understanding on the purpose of higher education. The context-bound factors are regarded yet not sufficiently and they end up in an isomorphic perspective.

Finally, it is worth mentioning the world university rankings impacting the strategy formation through their regional rankings and summits. For instance, the Times Higher Education (THE) ranking company is using this method very effectively. THE Latin American Universities Summit, THE Asia Universities Rankings and Summit, Japan University Rankings, Emerging Economies Rankings and Summit, Teaching Excellence Universities Ranking and Summit, Arab University Rankings and Summit (THE Official Website, 2021) can be given as examples. Similar structures can be observed in other world university ranking systems. The imposed upon reality in these ranking systems are measurable performance indicators (through publications, patents, research projects) that HEIs produce rather than a total of added value and impact (through teaching or community service) universities generate within their given society. There are mission statements of universities, which explicitly spell out its role as being ranked in top 100 universities around the world. Yet, this “global competition is [primarily] vectored by research capacity” (Marginson, 2006, p. 1). It is measurable and can be easily linked to innovation (Hazelkorn, 2012), which is the “engine of development in the new world economy” (Castells, 1996, p. 14). They are now “part of the central nervous system of societies ... with an entry point in the higher education quality assessment regimes” (Robertson, 2012, p. 3). The question raised by Hazelkorn (2012)—“Do they encourage a perverse incentive toward the creation of world-class universities at the expense of broader educational goals?” (p. 17)—is a legitimate one to ask in terms of reminding the very essence of HEIs in response to the popularity of these ranking systems affecting the strategy formation. Therefore, the strategy formation of HEIs is heavily steered in numerical and economic terms, connecting to the discussion in the first part of this section.

Essence of the modern university and strategy formation: a *mélange*

The modern university has been going through three transformations since the 1800s (Wittrock, 1993): the resurrection of the university, the rise of the modern-oriented university, and finally, the expansion and fragmentation phase. What we described in this work might call for a new phase of transformation, which stands as an arduous task in front of us. The two perspectives on higher education and operationalization of strategy formation practices cross-regionally indicate that the essence of the modern university has been in the midst of a severe contraction since the 2000s. While it is a truism that governments and international agencies have acknowledged the university’s power of transformative impact on societies, it is ironic that strategies shaped by (inter) governmental organizations and national, international, (bi-) regional, and transnational networks or associations have gained actorness and taken the lead over the actorness of the university itself. In other words, universities have been battling this push and pull between responding and protecting their autonomy as institutions with peculiar self-identities at the same time. We observe two crisscrossing processes—multipolarity and waning criticality—in the five regions’ trajectories that undergird universities’ predicament and find that to the extent that higher education systems with their stakeholders embrace these processes, they inform and shape higher education strategies.

One salient process shaping the development of higher education strategies is multipolarity. This multipolar process involves a wide range of (inter) governmental actors who project their power and influence by means of regional institutions (e.g., CON-AHEC, NAMI, European Union, the Pacific Islands Forum, Asian Development Bank, World Bank, UN, THE ranking company), associations (e.g., AAU, ACURU, ASEAN, UDUAL, UUK), and alliances (e.g., European Universities Initiative). While regionalism through institutionalization seems to lie at the epicenter of the multipolar world of higher education, advancement of economic growth via, say, human capital formation (e.g., Rizvi, 2021), innovation (e.g., Asian Development Bank, 2011), and technology (e.g., Santos and Wilson, 2021) can for the most part be argued to be the driving force behind higher education’s engagement in the multipolar context of the international system. In this new configuration, HEIs’ engagement and strategy formation processes stem mainly from a market-oriented narrative, and regional partnerships serve to formulate and actualize higher education strategies. Yet, the national capacity of a country, even the institutional capacities within the same countries, to embrace this version of

higher education engagement and strategy formation influences the extent to which strategies turn into policies and practices. This is evident in endeavors like free trade agreements (e.g., USMCA, RCEP), common higher education areas (e.g., EHEA, the Latin American Higher Education Area), and the BRICS bloc.

What further complicates this multipolar scene is the geopolitical appeal of such deals. We observe a clear orientation toward enhancing capabilities by allying with relevant actors in a region of interest who can help to weaken the global positioning of other challenger regions, a point also highlighted by Robertson (2010) in the context of European regional higher education governance and by Naidoo (2018) in the higher education “geopolitical rivalry” nexus (p. 607). Scott (2022) also argues that higher education policies, strategies, practices, and purposes have become predominantly affixed to a type of political economy that almost exclusively regards higher education as a “transactional space”—one fraught with “targets, metrics, deliverables and performance with much else missing” (para. 16). The strong association of higher education with this market-based type of political economy narrows, according to him, HEIs’ exercise of their core functions, especially their educative dimensions, to the terms of the market. In this sense, higher education strategy-making as part of geopolitics underpins a new phase of the transformation of the modern university that seems to have attained the feature of a geopolitically strategic global institution as the new oil of geopolitics.

Not surprisingly, there is little discussion of implications for the essence of the modern university in these regional arrangements, which are more prone to marry considerations of geopolitics to considerations to the real experiences of the modern university due to their national circumstances. To illustrate, it is observed that almost all regions have some form of “financial squeeze on HE” as the Asian Development Bank puts (2011, p. 5). The financial constraints are accepted as a natural consequence of the current economic system, but very few question why that is the case. As a result, the higher education landscape shifts from the public good to a private one and it is not being critically evaluated. For example, there is a high student enrollment rate all around the world; the limited financial and human resources are evidential; yet, the shamans expect higher education systems to function magically to accomplish all the goals attributed to them that require investment in human capital. There are no strategies on how to resolve the challenges, such as limited resources or the recruitment of higher education faculty to meet the needs of increasing enrollment. If we keep framing the role of the university in economic terms; then, our narrative is shaped by higher education as merchandise to be produced with the lowest cost but utmost profit. Even if there are available financial resources, it is already determined who will use them on what kinds of research or education by the shamanic actors again. We, therefore, argue that efforts to entangle geopolitics with the essence of the modern university can only offer a putative solution, leading only to mask the challenges confronting systems of higher education.

The multipolar process appears a big shock to the essence of the modern university, ushering in the second process, waning criticality. One interesting observation about this new world of universities is the transitory yet impactful nature of strategies. First off, strategies are overwhelmingly thematic around technology, accountability, transparency, innovation, entrepreneurship, private vs. public good, internationalization, transferability, access, world rankings, partnership/cooperation/collaboration, exchange/mobility, quality assurance, sustainability, and enrollment. Second, they inspire policymaking processes and are modified or renewed accordingly. Examples include the updated template for NAFTA, USMCA; Australia’s Asia engagement vision; the European Education Area; the Alfa PUENTES project; strategic plans in Africa; and the second Cordoba reform in Latin America. What prevails the accompanying strategy documents is the addressing of universities as being conventional into the bargain and the formation of new or modified strategies as a preventing-the-multipolar-system-from-clogging-up-mechanism. The upshot, however, is increased tendencies at the macro-level to glorify the challenges emanating from these processes by new or modified forms of strategies.

A macro-level strategy development process interwoven with economics and geopolitics leaves little room for the critical nature of universities to maneuver for engagement, although the exercise of criticality is how they exercise power and influence. The centralized decision-making in this process privileges the participation of senior higher education leaders, leaving faculty knowledge and expertise and input from students disengaged in the formation of strategies. That said, as the local, national, international, regional, and global institutions, universities bear the brunt of restructuring at the expense of a weakening in their critical essence. New roadmap released at the UNESCO World Higher Education Conference 2022, *Beyond Limits: New ways to reinvent higher education*, recognizes the need for a more inclusive approach and highlights that “students and teachers, ... HEIs, policymakers, international organizations, development partners, the private sector, civil society, youth organizations, social movements, local communities” are all needed if higher education is to contribute to overcome the local and global challenges (UNESCO, 2022, p. 27). The 2030 roadmap outlines six principles—(1) inclusion, equity and pluralism; (2) academic freedom and participation; (3) inquiry, critical thinking, and creativity; (4) integrity and ethics; (5) commitment to sustainability and social responsibility; (6) academic excellence through collaboration rather than excellence—to be followed by all the concerned parties throughout the reinvention process (UNESCO, 2022, p. 23). While these six principles are valuable to rethink the future of higher education strategy formations toward 2030 and beyond, we argue that a step forward to this end should also consider the challenge of macro-level strategy development to the reinventing higher education process. It is also worth noting that country-specific governance regimes impact how HEIs exercise their core functions—research, education, and service. As shown in the 2021 Democracy Index report by the Economist Intelligence Unit [EIU], democratic governance around the world is not on a positive trajectory: Only 6.4% of the world’s population—21 out of 167 countries—lives in a “full democracy” whereas 37.1%—59 out of 167 countries—lives under “authoritarian regimes” (EIU, 2022, p. 6, for a definition of full democracies and authoritarian regimes, see p. 68). That is, many HEIs suffer a decline in democracy, thus face intrusion into their institutional autonomy and academic freedom and significant constraints on their criticality in the search for and production of truth.

What arises from our cross-regional triangulation of higher education strategy formation practices is the reshaping of fundamental functions—research, education, and community service—of the modern university mainly based on a preponderance of

economic urges. We, instead, suggest this new phase in the modern university's transformation be approached as a *mélange*. A *mélange*, according to the [Online Etymology Dictionary \(2021\)](#), is "a mixture ... an uncombined mingling on elements, objects, or individuals." In doing so, we aim to encourage rethinking the modern university's engagement in the strategy-making processes under conditions of multipolarity. Anchored in this analytical approach, we heed [Ashwin's \(2020\)](#), [Birnbaum's \(1999\)](#), [Ozga and Lingard's \(2007\)](#), and [Rizvi's \(2021\)](#) calls to challenge the centrality of the economically oriented perspective in the development of higher education strategies and make the case for the study of the essence of the modern university as a unit of analysis for future research. In line with these scholars, we argue that fixation of strategy formation practices along market lines benefits from a system that condones universities as institutions of human capital formation. In response, we suggest the scholarship of the essence of the modern university renounce feeding this system, be concentrated on educational outcomes, civic role, and socially transformative sides of universities, and couched in educational sciences and higher education studies.

Building this type of scholarship is of dire necessity for the university as an institution to reclaim and promote its fundamental values—institutional autonomy, freedom of inquiry, and indivisibility of teaching and research—and to make sense of diversity in the context of multipolarity. In the absence of sustained countervailing response, a restraining alternative mechanism, the economic orientation toward higher education will continue framing strategies for educational outcomes, civic role, and socially transformative sides of universities, setting the tone for the scope, nature, and scale of a scholarly act for (wider) community engagement. A scholarship of this type requires more than educationally nuanced research efforts but rather proactively promotes engagement within communities of universities, with students, staff, academics, other stakeholders, and societies from a socially responsive standpoint. Formulating our research questions in this manner is apt to disturb the balance of the dominant narrative through acting as a power constraint mechanism. Perhaps this is how we can build proactive ownership around the challenges that confront us as communities of universities. Some questions that might open new windows of opportunity in this regard are as follows:

1. What constitutes knowledge that universities produce?
2. How do academics/university leaders perceive and reason through their roles in the supremacy of market orientation in higher education strategy making?
3. How has the identity of academics evolved out of their experiences within the multipolar world of universities?
4. Who gets to define and govern the identities of academics/students/university leaders?
5. What defines a scholarly act that nurtures the bonds of community engagement in the multipolar world of universities?
6. In what lines do students argue the purposes of university education?
7. How do communities of universities create proactive and collaborative ownership of the purposes of higher education?

Without proactive ownership of academic (scientific) values and interests and collective action, the fact is that the essence of the modern university will go on being denigrated by a global system encumbered with inequalities. The COVID-19 pandemic forced visibility of the existing inequalities within this system when the least developed countries were hit hardest by its dire effects; this situation is by no means over today. It also stressed the delicacy of the system and how narrowly defined its limits are when the economically more developed countries turned inward, focused their efforts on image building and geopolitical positioning. As one remedy, they reframed and/or announced strategic regional trade agreements to shore up their power and global position via higher education. The ongoing global health crisis presents an opportunity for higher education systems to contemplate the true meaning of higher education. Indeed, it is high time that communities of universities said enough.

References

- AAU (Association of African Universities), 2021. Our history. Available at: <https://aau.org/about/> (Accessed September 2021).
- Altbach, P.G., Bassett, R.M., 2014. Nix the BRICS—at least for higher education debate. *Int. High. Educ.* 77, 2–5.
- Altbach, P.G., Salmi, J., 2021. Why Latin America needs world-class universities. *Int. High. Educ.* 108, 20–21. <https://ejournals.bc.edu/index.php/ihe/article/view/14451>.
- Ashwin, P., 2020. *Transforming University Education: A Manifesto*. Bloomsbury, UK.
- Asian Development Bank, 2011. *Higher education across Asia: an overview of issues and strategies*. Asian Development Bank, Philippines. <https://www.adb.org/publications/higher-education-across-asia-overview-issues-and-strategies>.
- Australian Government, 2012. *Australia in the Asian Century*. Commonwealth of Australia, Canberra.
- Ayers, D.M., 2000. Tradition, modernity and the development of education in Cambodia. *Educ. Rev.* 44 (4), 440–463.
- Barrow, C.W., Didou-Aupetit, S., Mallea, J. (Eds.), 2003. *Globalisation, Trade Liberalisation, and Higher Education in North America: The Emergence of a New Market Under NAFTA?*. Springer Science+Business Media, Netherlands.
- Becker, G.S., 1993. *Human Capital: A Theoretical and Empirical Analysis With Special Reference to Education*, third ed. University of Chicago Press, Chicago, IL.
- Bergan, S., 2015. The EHEA at the cross-roads: the Bologna Process and the future of higher education [Overview paper]. In: Curaj, A., Matei, L., Pricopie, R., Salmi, J., Scott, P. (Eds.), *The European Higher Education Area: Between Critical Reflections and Future Policies*. Springer, pp. 727–742 eBook.
- Birnbaum, R., 1999. Academic leadership at the millennium: politics or porcelain? *J. Acad.* 85 (3), 14–19.
- Castells, M., 1996. *The Rise of the Network Society, the Information Age: Economy, Society and Culture Vol. I*. Blackwell Publishers, Massachusetts and Oxford.
- Chang, V., Chen, Y., Xiong, C., 2018. Dynamic interaction between higher education and economic progress: a comparative analysis of BRICS countries. *Inf. Discov. Deliv.* 46 (4), 225–238.
- Clark, B.R., 1983. *The Higher Education System: Academic Organization in Cross-National Perspective*. University of California Press, Berkeley, CA.
- Collins, N., 2000a. A virtual marketplace for North American collaboration and exchange. In: *Increasing North American Higher Education Cooperation, Collaboration, and Exchange Through a Collaboration "Marketplace"*. ACE—AUCC—ANUIES, pp. 1–8. Available at: <https://conahec.org/>. (Accessed October 2021).
- Collins, N., 2000b. Executive summary. In: *Increasing North American Higher Education Cooperation, Collaboration, and Exchange Through a Collaboration "Marketplace"*. ACE—AUCC—ANUIES, pp. 1–3. Available at: <https://conahec.org/>. (Accessed October 2021).
- de Moura Castro, C., Levy, D.C., 1997. *Higher education in Latin America and the Caribbean*. Strategy paper. Inter-American Development Bank, Washington, DC.

- Dei, D., Osei-Bonsu, R., Amponsah, S., 2019. A philosophical outlook on Africa's higher education in the twenty-first century: challenges and prospects. In: IntechOpen Book Series. <https://doi.org/10.5772/intechopen.86885>.
- Department of the Prime Minister and Cabinet, 2019. Our Public Service, Our Future: Independent Review of the Australian Public Service. Commonwealth of Australia, Canberra.
- Drucker, P.F., 1993. The rise of the knowledge society. *Wilson Q.* 17 (2), 52–71.
- Economic Intelligence Unit, 2022. Democracy Index 2021: The China Challenge [White Paper]. Economist Intelligence, London. Available at: <https://www.eiu.com/n/content/white-papers/>.
- Eide, E.R., Showalter, M.H., 2010. Human capital. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education: Higher Education*, third ed. Elsevier, Oxford, pp. 282–287.
- Enders, J., 2010. Higher education and the labor market. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education: Higher Education*, third ed. Elsevier, Oxford, pp. 365–369.
- European Commission, 2017a. White Paper on the Future of Europe: Reflections and Scenarios for the EU27 by 2025. European Commission, Brussels.
- European Commission, 2017b. Strengthening European Identity Through Education and Culture: The European Commission's Contribution to the Leaders' Meeting in Gothenburg, 17 November 2017. European Commission, Brussels.
- European Commission, 2019. European Universities: A Key Pillar of the European Education Area. European Commission, Brussels [Factsheet].
- European Council, 2017. European Council Meeting (14 November 2017)—Conclusions. European Commission, Brussels.
- Fredua-Kwarteng, E., 2020. Rethinking Strategic Planning in African Universities. *University World News*. <https://www.universityworldnews.com/post.php?story=2020100303173555>.
- Gunn, A., 2020. The European Universities Initiative: a study of alliance formation in higher education. In: Curaj, A., Deca, L., Pricopie, R. (Eds.), *European Higher Education Area: Challenges for a New Decade*. Springer, pp. 13–30. <https://doi.org/10.1007/978-3-030-56316-5> eBook.
- Hazekorn, E., 2012. Striving for "world class excellence": rankings and emerging societies. In: Araya, D., Marber, P. (Eds.), *Higher Education in the Global Age: Universities, Interconnections and Emerging Societies*. Routledge.
- Henry, K., 2016. Australia in the Asian century. *Asia Pac. Pol. Stud.* 3 (2), 132–139.
- Holm-Nielsen, L., Brunner, J.J., Balán, J., Thorn, K., 2004. Higher Education in Latin America and the Caribbean: Challenges and Prospects. UNESCO Report, pp. 1–28.
- Jacob, W.J., Neubauer, D., Ye, H., 2018. Financing trends in Southeast Asia and Oceania: meeting the demands of regional higher education growth. *Int. J. Educ. Dev.* 58, 47–63.
- Kind, N., Colucci, E., 2014. ALFA PUENTES University Associations in Latin America: A Snapshot. European University Association, Belgium. <https://eua.eu/resources/publications.html>.
- Lang, D.W., 2017. Fiscal incentives, Clark's triangle, and the shape and shaping of higher education systems. *Pol. Rev. High. Educ.* 1 (2), 112–138.
- Lauder, H., Brown, P., Dillabough, J.A., Halsey, A.H., 2006. Introduction: the prospects for education: individualization, globalization, and social change. In: Lauder, H., Brown, P., Dillabough, J.A., Halsey, A.H. (Eds.), *Education, Globalization, and Social Change*. Oxford University Press, Oxford, pp. 1–70.
- Lay, S., 2004. The Interpretation of the Magna Charta Universitatum and Its Principles. Bononia University Press, Bologna.
- Mallea, J., Malo, S., Pendergrast, D., 1997. The Vancouver Communiqué revisited: an assessment. In: *Understanding the Differences: A Working Paper Series on Higher Education in Mexico, Canada, and the United States* (Working Paper No. 8, Publication No. 2A315). Western Interstate Commission for Higher Education, USA, pp. 1–37. Available at: <https://conahcec.org/>. (Accessed October 2021).
- Marginson, S., 2006. Dynamics of national and global competition in higher education. *High. Educ.* 52 (1), 1–39.
- Marginson, S., 2017. Limitations of human capital theory. *Stud. High. Educ.* 44 (3), 1–15.
- Mintzberg, H., 1978. Patterns in strategy formation. *Manag. Sci.* 24 (9), 934–948.
- Naidoo, R., 2018. The competition fetish in higher education: shamans, mind snares and consequences. *Eur. Educ. Res. J.* 17 (5), 605–620.
- Observatory Magna Charta Universitatum, 2020. Magna Charta Universitatum 2020. Available at: <http://www.magna-charta.org/magna-charta-universitatum/mcu-2020>. (Accessed October 2021).
- Oketch, M., McCowan, T., Schendel, R., 2014. The Impact of Tertiary Education on Development: A Rigorous Literature Review. Department of International Development, UK.
- Oliver, D., 2004. Human capital theory and higher education in developing countries. *J. Thought* 39 (1), 119–130.
- O'Malley, B., Warden, R., 2022. Global "Roadmap to 2030" for Higher Education Announced. *University World News*. <https://www.universityworldnews.com/post.php?story=20220525095453173>.
- Online Etymology Dictionary, 2021. Mélange. Etymonline.com Dictionary. Available at: <https://www.etymonline.com/search?q=melange>. (Accessed October 2021).
- Ozga, J., Lingard, R., 2007. Globalisation, education policy and politics. In: Lingard, R., Ozga, J. (Eds.), *The RoutledgeFalmer Reader in Education Policy and Politics*. Routledge, London, pp. 65–82.
- Pacific Islands Forum Secretariat, 2014. The Framework for Pacific Regionalism. Pacific Islands Forum Secretariat, Suva, Fiji.
- Rizvi, F., 2006. Imagination and the globalisation of educational policy research. *Glob. Soc. Educ.* 4 (2), 193–205.
- Rizvi, F., 2021. International higher education and the formation of business diasporas. *Br. J. Educ. Sci.* 1–20.
- Robertson, S.L., 2010. The EU, "regulatory state regionalism" and new modes of higher education governance. *Glob. Soc. Educ.* 8 (1), 23–37.
- Robertson, S.L., 2012. World-Class Higher Education (For Whom?). Viewpoints/Controversies. Springer. <https://doi.org/10.1007/s11125-012-9236-8>.
- Rüegg, W. (Ed.), 2004. A History of the University in Europe, vols. 1–4. Cambridge University Press, Cambridge.
- Salazar, J., Leihy, P., 2013. Keeping up with the coordination: from Clark's triangle to microcosmography. *Stud. High. Educ.* 38 (1), 53–70.
- Santos, A., Wilson, C., 2021. Trade, Economy, and Work: A Shared Agenda for a Stronger Economic Future. Publication No. 2378. Georgetown Law Faculty Publications and Other Works, USA, pp. 1–12. Available at: <https://scholarship.law.georgetown.edu/>. (Accessed October 2021).
- Schwartzman, S., 2020. Higher education and the challenges of the twenty-first century: an introduction. In: Schwartzman, S. (Ed.), *Higher Education in Latin America and the Challenges of the 21st Century*. Springer Nature, Switzerland, pp. 1–18.
- Schwartzman, S., Pinheiro, R., Pillay, P., 2015. Education in the BRICS Countries: Investigating the Pact Between Higher Education and Society. Springer, Dordrecht.
- Scott, J.C., 2006. The mission of the university: medieval to postmodern transformations. *J. High. Educ.* 77 (1), 1–39.
- Scott, P., 2022. The Space for a New Saga for Higher Education is Emerging. *University World News*. <https://www.universityworldnews.com/post.php?story=20220525141306100>.
- Scott, P., Gallacher, J., Parry, G. (Eds.), 2017. *New Languages and Landscapes of Higher Education*. Oxford University Press, Oxford.
- Scott-Kemmis, D., Intarakummer, P., Rasiah, R., Amaradasa, R., 2021. Southeast Asia and Oceania. In: UNESCO Science Report: The Race against Time for Smarter Development. UNESCO, France, pp. 675–715.
- Shultz, T., 1963. *The Economic Value of Education*. Columbia University Press, New York.
- SICE the OAS Foreign Trade Information System, 2021. North American Free Trade Agreement (NAFTA), Chapter One: Objectives. Available at: <http://www.sice.oas.org/trade/nafta/naftatce.asp>. (Accessed October 2021).
- Teferri, D., Altbach, P.G., 2004. African higher education: challenges for the 21st century. *High. Educ.* 47, 21–50. <https://doi.org/10.1023/B:HIGH.0000009822.49980.30>.
- THE Official Website, 2021. World University Rankings. Available at: <https://www.timeshighereducation.com/content/world-university-rankings>. (Accessed September 2021).
- The World Bank, World Development Indicators, 2020a. GDP (Current US\$). Available at: <https://data.worldbank.org/indicator>. (Accessed October 2021).
- The World Bank, World Development Indicators, 2020b. Population, Total. Available at: <https://data.worldbank.org/indicator>. (Accessed October 2021).
- Tilak, J.B.G., 2013. Higher education in the BRIC member-countries: comparative patterns and policies. *Econ. Polit. Wkly.* 48 (14), 41–47.

- UNESCO, 2022. Beyond limits. New ways to reinvent higher education. In: Working Document for the World Higher Education Conference. 18–20 May 2022. UNESCO, Barcelona, Spain. Available at: <https://www.whhec2022.org/EN/homepage>.
- Wittrock, B., 1993. The modern university: the three transformations. In: Sheldon, R., Wittrock, B. (Eds.), *The European and American University Since 1800s: Historical and Sociological Essays*. Cambridge University Press, UK, pp. 303–362.
- Young, M., Pinheiro, R., 2022. The post-entrepreneurial university: the case for resilience in higher education. In: Pinheiro, R., Frigotto, M.L., Young, M. (Eds.), *Towards Resilient Organizations and Societies: A Cross-Sectoral and Multi-disciplinary Perspective*. Palgrave Macmillan, pp. 173–193. https://doi.org/10.1007/978-3-030-82072-5_7. Public Sector Organizations. eBook.
- Zgaga, P., 2020. Introduction: EHEA: the future of the past. In: Curaj, A., Deca, L., Pricopie, R. (Eds.), *European Higher Education Area: Challenges for a New Decade*. Springer, pp. XXXVIII–XLII. <https://doi.org/10.1007/978-3-030-56316-5> eBook.

Relevant websites

- ASEAN Socio Cultural Community, <https://asean.org/our-communities/asean-sociocultural-community#education>.
- Association of the Carpathian Region Universities (ACRU), <https://acru.uvlf.sk/home.html>.
- Association of the Universities of Latin America and the Caribbean (UDUAL), <https://www.udual.org/principal/wp-content/uploads/2020/12/Folleto-UDUAL-ingles.pdf>.
- The Association of Southeast Asian Nations (ASEAN), <https://asean.org/>.
- Council of Regional Organizations of the Pacific (CROP), <https://www.forumsec.org/council-of-regional-organisations-of-the-pacific/>.
- Delegation of the European Union to ASEAN, https://eeas.europa.eu/delegations/association-southeast-asian-nations-asean/905/about-eu-delegation-asean_en.
- European Education Area, The European Universities Initiative, https://ec.europa.eu/education/node_en.
- EU-ASEAN Socio-Cultural Cooperation, <https://euinasean.eu/cooperation/socio-cultural-cooperation/>.
- EU-ASEAN Strategic Partnership, <https://euinasean.eu/>.
- EU-LAC, <https://eulacfoundation.org/en/who-are/about-us>.
- EU-Mercosur, <https://www.mercosur.int/en/about-mercorsur/mercorsur-in-brief/>.
- EU-Mercosur Trade Agreement, <https://ec.europa.eu/trade/policy/in-focus/eu-mercorsur-association-agreement/>.
- European Association of Institutions in Higher Education (EURASHE), <https://www.eurashe.eu/>.
- European University Association (EUA), <https://eua.eu/>.
- la Asociación Nacional de Universidades e Instituciones de Educación Superior [The National Association of Universities and Institutions of Higher Education] (ANUIES), <http://www.anui.es/>.
- Observatory Magna Carta Universitatum, <http://www.magna-charta.org/magna-charta-universitatum>.
- Matrix of objectives and strategic lines for the construction of the Common Area of Higher Education EU-LAC, https://eulacfoundation.org/sites/default/files/documents/Matrix_AcadCoop_EU_LAC_06.2021_final.pdf.
- Pacific Islands Forum, <https://www.forumsec.org/>.
- Bologna Process, European Higher Education Area (EHEA), <http://www.ehea.info/>.
- See "SICE the OAS Foreign Trade System" for the full NAFTA chapters, <http://www.sice.oas.org/trade/nafta/naftatce.asp>.
- The American Council on Higher Education (ACE), <https://www.acenet.edu/Pages/default.aspx>.
- The Association of Asian Universities (AAU), <http://www.aauuniv.org/files/documents/00000013.pdf>.
- The Consortium for North American Higher Education (CONAHEC), <https://conahec.org/>.
- The EU-ASEAN, <https://euinasean.eu/the-eu-asean/>.
- The North American Institute (NAMI), <https://www.northamericaninstitute.org/>.
- The Pacific Data Hub (PDH), <https://pacificdata.org/>.
- The Regional Comprehensive Economic Partnership (RCEP), <https://rcepsec.org/>.
- The United States-Mexico-Canada Trade Agreement (USMCA), <https://www.trade.gov/usmca>.
- United Nations, Department of Economic and Social Affairs, Statistics Division, Standard country or area codes for statistical use (M49), Geographic regions, <https://unstats.un.org/unsd/methodology/m49/>.
- United Nations, Standard Country or Area Codes for Statistical Use, Geographic Regions, <https://unstats.un.org/unsd/methodology/m49/>.
- Universities Australia, <https://www.universitiesaustralia.edu.au/>.
- Universities Canada, formerly the Association of Universities and Colleges of Canada (AUCC), <https://www.univcan.ca/>.
- Universities New Zealand, <https://www.universitiesnz.ac.nz/>.
- Universities UK (UUK), <https://www.universitiesuk.ac.uk/>.
- University Associations in Latin America, <https://eua.eu/downloads/publications/alfa%20puentes%20publications%20university%20associations%20in%20latin%20america%20-%20a%20snapshot.pdf>.
- WHEC 2022 UNESCO World Higher Education Conference, <https://www.whhec2022.org/EN/homepage>.

Knowledge production in Europe: actors, policies and critiques

Tatiana Suspitsyna and Penny A. Pasque, The Ohio State University, Columbus, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

States and supranational organizations	13
Bologna process	13
Lisbon strategy	13
Industry	14
Science parks	14
Technology transfer offices	14
Entrepreneurial universities	15
Critiques of knowledge production in Europe	15
Academic capitalism	16
Entrepreneurialism and Europeanization as governmentality	16
Epistemic injustice	17
Concluding remarks	17
References	18

Glossary

- Academic Capitalism
- Accountability Systems
- Bologna Process
- Contributory Injustice
- Covid-19
- Epistemic Injustice
- Entrepreneurial Universities
- European Union
- Gender
- Hermeneutical Injustice
- Industry
- Lisbon Strategy
- Privatization
- Technology Transfer
- Testimonial Injustice
- Race
- Science Parks
- States and Supranational Organizations

Whether it was an institutional transformation (Schulze-Cleven, 2016) or a fully-fledged academic revolution (Etzkovits, 2019), scholars agree that the massification of higher education (Tight, 2019), internationalization of universities in Western (Hudson, 2016) and post-Soviet Central and Eastern Europe (Dakowska and Harmsen, 2015), a wide-spread adoption of European Union (EU) initiatives (Vukasovic, 2013), as well as the growing demands of knowledge-based economies on research in the early 21st century (Sørensen et al., 2015) have shaped the current context of knowledge production in Europe. Produced in spaces that bring together the interests and capacities of states, industries, and entrepreneurial universities (Etzkovits and Zhou, 2018; Etzkovits, 2019), knowledge and innovation in Europe are not only key factors of the competitive advantage and economic integration of European nations, but also a pathway to creating a “European knowledge society” (Hoekman et al., 2008).

While purely university-based scientific production is largely a mode of the past, most research and innovation in Europe follows Mode 2 of knowledge generation that is “socially distributed, application-oriented, trans-disciplinary, and subject to multiple accountabilities” (Nowotny et al., 2003, p. 179). The interactions of multiple academic, governmental, and market actors form a triple helix of scientific innovation that is now legitimated by its impact on economic development along with the more traditional expectations of contributing to health, military, and cultural advancements (Etzkovits and Leydersdorff, 2000). Recently, scholars have added a fourth helix to knowledge creation and transfer. Civil society is seen as a key participant in the

diffusion of knowledge and a source of a bottom-up approach to democratic and accountable scientific practices (Carayannis and Campbell, 2012).

In this article we examine the current roles of European states, markets, and universities in knowledge production and review the critiques of academic entrepreneurialism and neoliberal market influences on European higher education.

States and supranational organizations

As the political power of the state in Europe gradually moved to supranational structures like the EU (Lavdas et al., 2006), institutions of higher learning were strategizing to respond to higher enrollments, growing student mobility across national borders, and increasing global competition from knowledge-led economies in North America and Asia (Sursock and Smidt, 2010). The European universities' answer to these pressures was to europeanize (i.e., adopt EU policies and initiatives directed at national higher education sectors) (Vukasovic, 2013) and to harmonize (i.e., align their national research systems with labor market regulations, property right laws, and academic degree nomenclatures) (Hoekman et al., 2008). Arguably, the two main European policies driving that response are the Bologna Process and the Lisbon Strategy (Keeling, 2006; Sursock and Smidt, 2010).

Bologna process

If the Lisbon Strategy is more of the call to arms, the Bologna Process is how to take said action. The European Bologna Process is a set of strategies to align national study programs and regulations across European countries in order to develop European Higher Education Area (EHEA) (Huisman and van Vught, 2009), a competitive European market, ease of transfer across institutions, and to attract international students (Naidoo, 2011). This strategic alliance seeks to increase Europe's overall higher education market share and attract the best intellectual minds to the area (Robertson, 2010). While the Bologna Process does not operate under the auspices of the European Union, it was the original 29 ministers who came together in Bologna, Italy in June 1999 and created the European Higher Education Area by 2010. The goal was to increase the mobility, employability of Europe's higher education system and the international resurgence of the Continent (Gaston, 2010) through the assurance of more compatible and coherent higher education systems.

The original Bologna Process sought to accomplish six broad action items: (1) the creation of easily readable and comparable degrees documented through diploma supplements provided to every student; (2) the organization of higher education (undergraduate, masters, doctoral; also known as the three cycles of higher education qualifications); (3) the management of education credentials through a recognized system of credits; (4) the encouragement of educational mobility through a recognized system of credits; (5) the encouragement of educational mobility through the expansion of cross-border opportunities for studying and teaching; (6) the development of a stronger commitment and promotion of dimensions of curricular development, interinstitutional cooperation, mobility schemes and integrated programs of student, training and research. (Bologna, 1999; Gaston, 2010, pp. 35–36). The Ministerial Conferences, have held every 2–3 years, assess the progress made within the EHEA and to decide on the new steps to be taken. In 2020, 48 European (and other) countries were involved in this process (EHEA, 2020) where the process addressed research at the supranational level (van Vught, 2009).

Since 1999, the Bologna Process acquired political and economic dimensions along the way (Tomusk, 2004). Politically, it allowed the European Commission to exert more influence over national higher education systems and added legitimacy to the ambitions of former communist states that wished to join the European space (Tomusk, 2004). Economically, it charged European universities with contributing to the creation of European knowledge society. The framing of the Bologna process in the discourse of the knowledge economy and society enforced the nation states' obligation to the process and enabled them to commence higher education reforms (Ravinet, 2008).

Lisbon strategy

The Bologna Process and Lisbon Strategy are often discussed together. In relation to knowledge production, Lisbon is related directly to national economic competitiveness, rather than the whole of workforce development and, as such, we discuss it here.

The Lisbon Recognition Convention is an agreement on the recognition of educational qualifications in Europe that were adopted at the COE/UNESCO meeting in April 1997. It encourages consistent, convenient assessment of educational credentials earned in any European country. The Lisbon Recognition Convention and Bologna Process share a priority of higher education and are often confused (Gaston, 2010). The Bologna Process is a part of the overarching Lisbon Strategy. Envisioned at the meeting of the EU heads of states in Lisbon in 2000, the intention of the Lisbon Strategy is to enable Europe to increase its prosperity and, in turn, to enhance its social models (Rich, 2010). To achieve this, the European Council called for the EU to become the most competitive and knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion. This ambitious vision aimed to transform not only the European economy but also the social welfare and education systems in European nations: "Key elements of the strategy include budgetary sustainability, better regulation, improved tax and benefit systems, and improvements in education and training systems, and investment in research to maintain comparative advantage over other parts of the world" (Rich, 2010, p. 566).

Thus, the Lisbon Strategy explicitly linked research to economic development and set a goal of transforming the European Union into the most globally competitive region (Keeling, 2006). The implementation of the Strategy faced some setbacks, and the re-launch of the initiative in 2005 was criticized for its departure from the social goals of societal cohesion and inclusion in favor of an economically driven pursuit of job creation (Dieckhoff and Gallier, 2007). Nevertheless, one element of the Strategy that

proved most popular among the universities—funding of research—earned the initiative considerable legitimacy among academics and university administrators (Keeling, 2006).

The diverse legal frameworks, market regulations, and higher education systems of European nations presented a further challenge to European knowledge production and transfer. Removing barriers to scientific collaboration within the EU through “harmonization of national research systems” included the alignment of property rights laws, labor market practices and regulations, and diploma systems (Hoekman et al., 2008, p. 736). The harmonization efforts, as critics points out, made it possible to discuss a European higher education system rather than systems, yet the new configuration is marred by geographic disparities (Tomusk, 2004). Elite and capital regions with best resources and top scholars collaborate within themselves, excluding the periphery (Hoekman et al., 2008). In Central and Eastern Europe knowledge production remains linked to legacies of command-driven economies and to transitions: as these nations transitioned to capitalist market economies, and societies to post-communist welfare states, their higher education and research systems inherited the issues of access, funding, and governance of their predecessors (Kwiek, 2012).

Industry

In addition to securing funding for their research activities, universities are motivated to collaborate with industry to improve the quality of their scientists and see their innovations applied to real-life world (Muscio, 2010). In the 1980s and 1990s, the experience of high-technology-based economic growth around Cambridge in the UK, Twente in the Netherlands, and in Finland became exemplars of successful university-industry collaboration and gave impetus to a range of government supported initiatives such as science parks and technology transfer offices (Kitagawa, 2005). These initiatives were conceived as important elements of the European nations’ innovation systems that consist of “competitive firms and a competitive environment, an effective educational system, strong university research, a legal system with property rights, and a capital market that includes venture capital” (Link and Scott, 2007, p. 669).

Science parks

Depending on the accepted terminology in a specific European regional and national context, science parks are also known as hi-tech parks, innovation centers, technopoles, and campuses (Ng et al., 2019). Defined as “a real estate or an area development, managed by an on-site management company [and] home to knowledge organizations, such as research institutes, HEI, and firms in all business development phases” (Ng et al., 2019, p. 726), science parks successfully broke apart the monopoly of universities on the production of knowledge in Europe (Etztowitz, 2003). Over half of all European science parks (55%) are publicly owned by local and regional governments and higher education institutions, 14% are privately owned by private universities and businesses, and the remaining third (30%) have a mixed private-public control (Kriemadis, 2018).

The explicit assumption underlying the foundation of science parks is that they contribute to national competitiveness through a two-way flow of innovation between universities and industry (Link and Scott, 2007). On the basis of the purposes and characteristics of resident organizations, European science parks can be classified into three groups: incubators, research locations, and cooperative locations (Ng et al., 2019). The goal of science parks is to produce value through their knowledge and business ecosystems, and universities and research institutes play key roles in that task at research and cooperative locations (Ng et al., 2019). When located in cities, science parks form knowledge community precincts, contributing to innovative urban development and renewal (Olçay and Bulu, 2016) and become vital part of urban knowledge economies (Teixeira et al., 2019).

Talent recruitment and management is an important function of science parks that complete for best graduates and maintain networks with the alumni of tenant higher education institutions (Löfsten et al., 2020). As stakeholders in the success of science parks, governments can invest in the promotion and internationalization of their parks’ brand for greater contribution to their nations’ knowledge economies (Cadorin et al., 2019). Skilled migrants in Europe also make a positive impact on the growth of published academic articles and patented innovations in industry (Bosetti et al., 2015).

Technology transfer offices

Along with a more traditional commercialization of intellectual property through individual patents and licensing (Thursby and Thursby, 2007), technology transfer offices offer universities an opportunity to commercialize innovation by focusing on selected high-quality inventions for reputation building (Siegel et al., 2007). In transition economies of Central and Eastern Europe commercialization of university research occurs directly as a result of the direct industrial funding rather than through government-created technology transfer offices (TTOs) (Belitski et al., 2019).

Based on the assessment of potential profitability of the inventions, TTOs as owners of universities’ intellectual property typically choose one of the two strategies: licensing the inventions to existing companies or creating spin-offs, “new companies founded to exploit a piece of intellectual property created in an academic institution” (Shane quoted in Algieri et al., 2013, p. 383). In addition to the inventions’ potential, the decision about licensing or spin-off depends on the business expertise of the TTO staff and the university’s intent at the local level as well as the country’s intellectual property rights legal framework and the government funding system at the national level (Bengtsson, 2017). In all cases, however, the TTO’s hold a substantial advantage over individual academics because they have access to networks among firms and in industry to make commercialization successful (Siegel and Wright, 2015).

Knowledge and technology transfer offices typically cover three areas of activity: they support the intellectual property of academic researchers by protecting their research results and negotiating licensing contracts; they offer research support by overseeing researchers’

collaborations and contracts; and they assist with the business planning and funding of spin-offs (Brescia et al., 2016). Thus, to operate successfully TTOs must claim expertise both in science and in business. This dual identity may lead to legitimacy problems with academics and managers alike and require organizational efforts to meet normative expectations of both parties (O’Kane et al., 2015). At the same time, academic departments’ reputation for producing high-quality research and universities’ research ranking positively affect the use of TTOs: on the one hand, high-output departments make a more frequent use of TTOs than their less productive counterparts, on the other, companies are likely to contact those TTOs that are located at well-ranked universities (Muscio, 2010).

There are reported gender differences in knowledge and technology transfer in Europe. For example, research groups led by men are better positioned in the collaborative networks with industry and they are more likely to express interest in collaboration with external companies (Calvo et al., 2019). Furthermore, there is evidence that male scientists, more so than their female colleagues, are deterred by the difficulty of coming back to the academic position if their spin-off enterprise fails, while female scientists are more constrained by a lack of human resources and exchange with peers than their male counterparts (Sinell et al., 2018).

Entrepreneurial universities

Since taking hold in the mostly European and North American academic literature in the 1980s, the concepts of entrepreneurial university and academic entrepreneurialism have developed side by side to describe, predict, and critique the universities’ involvement in commercialization of research and creation of spin-off business-oriented ventures (Mascarenhas et al., 2017). In addition to its clear business connotations, entrepreneurial university is also described as “adaptive,” “proactive,” and “innovative” (Clark as cited in Schattock, 2005, p. 17), financially self-reliant and independent of the state (Schattock, 2005). The entrepreneurial university is also risk-taking, both intellectually and economically (Barnett, 2005). Because entrepreneurial universities are more responsive to the market, entrepreneurialism leads to diversification of institutional forms and missions and subjects universities to more global pressures (Schattock, 2005).

Unlike the US entrepreneurial university model that is based on the research mission, entrepreneurship of the European universities traditionally centered around their teaching mission (Etzkowitz, 2003). At the same time, eager to build up their knowledge-based economies, many European governments now task their higher education institutions with a third mission of service in the name of national and regional economic development (Williams and Kitaev, 2005). This third mission has acquired a particular strategic importance since the launch of the Lisbon Strategy. Entrepreneurial universities not only bring in technology and innovation to their locales but also serve as “reputational shopping malls for regional markets” that compete globally against each other and against the United States (Rinne and Koivula, 2005, p. 106).

Academic entrepreneurialism in Europe takes different forms in open and controlled markets, depending on the roles of national governments in the regulation of higher education systems. Barnett (2005) differentiates between soft and hard entrepreneurialism. Institutions’ explicit pursuit of profit through tuition price setting, revenue-generating services, program offerings, etc. characterizes a hard approach, which can take the form of unbridled entrepreneurialism in open markets and is curtailed in controlled markets. By contrast, soft entrepreneurialism is diffused in the culture and operations of universities and university departments. The soft approach, encompassing civic entrepreneurialism in open markets and hesitant entrepreneurialism in controlled markets, is grounded in the conceptualization of higher education as a public good and aims at anticipating and meeting the needs of the society through innovation in research and instruction (Barnett, 2005).

There are some notable disciplinary differences in the European professoriate’s perceptions of entrepreneurialism and knowledge transfer. More academics in natural sciences think of their departments as highly entrepreneurial than do their colleagues in social sciences (Kalar and Antonicic, 2015). Furthermore, those who view their departments as highly entrepreneurial tend to view knowledge transfer as beneficial rather than harmful to the academe and are themselves more likely to engage in entrepreneurial activities (Kalar and Antonicic, 2015).

Nevertheless, many see entrepreneurialism as a double-edged sword: on the one hand, it is beneficial as it leads to creativity and academic innovation and greater institutional autonomy; on the other, it is constraining because it leads to commercialization of research and instruction and less academic freedom for faculty to pursue their research, especially if their agendas have little market value. This tension is well captured by Rinne and Koivula, who observed that “the entrepreneurial university has been seen as a natural continuation of the development of the university ... but also as a deformed child of the research university” (Rinne and Koivula, 2005, p. 104).

Critiques of knowledge production in Europe

Scholars critical of European and North American knowledge production point at its embeddedness in the neoliberal market ideology and question its Euro-centrism. They observe that knowledge production is linked with knowledge legitimation and the “socially destructive effects of marginalizing forms of knowledge production” and contend that the designation of “developed” and “developing” assumes inevitability and appropriateness of converging on the Western market models of modern development (McGovern, 2012, p. 5). Other critics, approaching knowledge production from a Foucaultian poststructuralist perspective, view higher education policies and reforms as a form of governmentality, a political rationality aimed at directing the development of institutions and individuals (e.g., Ozga, 2008; Laalo et al., 2019; Day and Pirrie, 2021). Feminist, postcolonial, and Critical Race Theory scholars of higher education take a stand against a systematic disadvantage of women and people of color in the

production of knowledge that is dominated by white male academics and Western academic traditions (e.g., Blackmore, 2021; Joseph-Salisbury, 2019; Pasque and Carducci, 2015). In the following discussion we will present these lines of critique in some detail: Academic capitalism, including accountability and privatization; governmentality; and epistemic injustice.

Academic capitalism

Universities are often important cultural centers, spaces of the most cutting-edge research and development, and places of independent and critical thought (Altbach, 2016). University libraries contribute significantly to the knowledge economy as major repositories of knowledge and a greater amount of information is available open access through the internet, google scholar, Wikipedia, MOOCs, and global universities coordinated from a local institution. Yet, a global, knowledge-based economy is fundamentally different than other tenets of economics because knowledge behaves differently than other commodities and is particularly more difficult to measure (Neff, 1999). Knowledge is often driven by the economy and imperialism (Naidoo, 2011), pursuit of profit (Chomsky, 1999; Giroux, 2002; Giroux and Giroux, 2004), grants (Posselt, 2009), patents, copyrights commodifying the core academic function (Slaughter and Rhoades, 2004), corporate sponsorship (Bush, 2009) and General Agreement on Trade in Services (GATS), of which Australian and UK universities have led the way (Rhoades and Liu, 2009). The global economy is the most powerful force impacting universities today (Rhoades and Liu, 2009). This is what Apple (2006a) calls the era of conservative modernization or “a new hegemonic bloc ... a very creative articulation of themes that resonate deeply with the experiences, fears, hopes, and dreams of people as they go about their daily lives ... integrated ... with racist nativist discourses, within the economic dominant forms of understanding, and within a problematic sense of tradition” (pp. 27–28). This European-American socio-political context connects neoliberalism, neoconservatism, authoritarian populist religious conservatism, and managerialism (Apple, 2000, 2006a,b). This era of conservative modernization seeks to undermine certain conceptualizations of freedom and democracy and supplant them with conceptualizations that serve the private interests of the ruling class (Gildersleeve et al., 2010).

Accountability systems have become entrenched focusing on performance funding (in which state allocations are linked by formula to performance on an indicator or set of indicators), performance budgeting (in which performance data are used in budgeting but without a formula), performance reporting (in which there is no explicit link between the data reported and the budget process), or a combination (Zumeta and Kinne, 2011, p. 178). European institutions have justified this accountability and institutional autonomy logic through competitive logic. This logic breeds performance management, accreditation and rankings (Erkkilä and Piironen, 2014). International rankings such as the Times Higher Education ranking, the Shanghai ranking and Newsweek (a mix of Times and Shanghai) differ considerably and there is a concern that continental Europe is falling behind (Winckler, 2008).

Private, individual good Institutions of higher education have transformed from a focus on the public, social good to that of the private, individual good (McMahon, 2009; Pasque, 2010/2014). In Europe as well as in many countries, a shift of financial burden is being made from the financing of the taxpayer to the student (via tuition and fees) and from the public to the private sector (Newby, 2008). Private education is one of the fastest growing and dynamic areas of higher and postsecondary education throughout Europe and across the globe as 30% of global higher education enrollment is private (Altbach, 2016). Corporate sponsorship has benefited both the public and private institution, with conditions favoring the corporation including increasing the size of the potential labor pool, securing qualified employees for their own use, and negotiating special tuition rates (Bush, 2009). Ownership, copyright and profits matter as the commodification of higher education has taken hold in recent years. Benefits to the university may include increased revenue, lower cost, increased communication with corporate stakeholders, and increased access to potential employers. Corporate sponsorship colonization is defined as benefiting the advantaged institution and exploiting the less advantaged institution (Bush, 2009; also see Dunning, 2000). This has been termed alliance capitalism, where alliances are built between institutions and different stakeholders such as a private firm, public institution, regional and national government, interest groups, businesses, and beyond (Dunning, 2000).

Entrepreneurialism and Europeanization as governmentality

Poststructuralist critics of entrepreneurial universities and Europeanisation processes employ Foucault’s concept of governmentality to uncover political rationalities that govern European higher education policy making and the techniques of self that frame the behaviors of students, academics, and managers as “economically self-interested subjects” (Laalo et al., 2019, p. 96). From this perspective, the Bologna Process not only streamlines national systems of higher education for easier student mobility but also directs higher education institutions, academic communities, and student populations to become globally competitive entrepreneurial actors for their own private benefit and the benefit of the EU market (Nokkala, 2015). To put it differently, the creation of European knowledge society and the Europeanization measures aim at generating knowledge not only for industrial innovation but also—particularly in human sciences—for managing education and welfare of populations and as such are enmeshed in political technologies of European states and supranational organizations (Ozga, 2008). Framed by the ethics of neoliberalism, the EU discourses that promote entrepreneurship in higher education “to make stronger Europe” create a normative ideal of what a competitive university and an innovative academic look like (Laalo et al., 2019, p. 95). Emphasizing the market reasoning, that ideal explicitly neglects non-economic purposes of universities, such as the development of critical thinking and contestation of democratic ideas (Laalo et al., 2019). The role of university instructors is then transformed into managers of their own intellectual outputs and student satisfaction (Day and Pirrie, 2021).

Epistemic injustice

Epistemic injustice is a regulatory form of knowledge production that starts with the concept that social power is the capacity humans have to act and influence the social world, in this case, in the context of a European university (Fricker, 2008 also see www.oxfordscholarship.com). Power exists when put into action (Foucault, 1982) and *epistemic injustice* occurs when a wrong done to a person specifically in their capacity as a knower by a hearer (Fricker, 2008). This epistemic side of colonialism is the “devastating effect of the ‘disappearing’ of knowledge, where local or provincial knowledge is dismissed due to privileging of alternative, often Western, epistemic practices” (Dotson, 2012a, p. 236). There are three types of epistemic injustice: testimonial, hermeneutical, and contributory.

Testimonial injustice, a form of epistemic injustice, occurs when prejudice causes a hearer to give a reduced degree of credibility to a speaker’s words based on the speaker’s perceived gender, race, sexual orientation, class and/or other identity. An example of testimonial injustice is when a campus police officer does not believe you because you are a Black faculty; or a faculty wrongfully deflates your perspective as a student because you are a woman in the sciences. The speaker in these cases is wronged. *Hermeneutical injustice* occurs when there is a gap in collective or social-interpretive resources based on identity, and one is structurally prejudiced. Fricker offers an example of a new mother in the 1960s who, after participating in a small group feminist discussion, comes to understand that her depression was not her fault nor what her husband blamed her for previously. In this case, hermeneutical injustice is based on a structural definition of new mothers during that time period, and one could have this definition inflicted on her (by her husband) or she could inflict it on herself (also known as *internal oppression*, Lorde, 1984/2007). Hermeneutical oppression puts someone at a prejudicial disadvantage when it comes to making sense of their social experiences and systemic understanding, in this case of a new mother, and this sense-making changed over time to be defined as postpartum depression. Hermeneutical injustice at a European university exists when one suffers sexual harassment in a campus work culture that still lacks understanding of that critical concept. In this case when hermeneutical injustice is present, it is not the speaker’s lack of intelligibility but it is due to the deprived interpretive resources available to the speaker (Fricker, 2008). Dotson (2012a) extends understandings of epistemic injustice to include *contributory injustice*, which arises from a willful denial of epistemologies by the hearer. It occurs when diverse voices and experiences often counter to the dominant paradigms are offered but immediately or preemptively rejected. In this case, the hearer is unwilling to acknowledge the diverse ways of knowing (Dotson, 2012b), such as a campus climate that structurally supports harassment even after the #metoo movement (Founder T. Burke) and campus awareness.

With its focus on the economic outputs, the neoliberal entrepreneurial university inflicts epistemic injustice on women faculty and administrators by systematically disregarding the structural and cultural barriers to their advancement (Blackmore, 2021). Knowledge production in the EU remains dominated by men, with room for growth among women and transgender people. Elsevier’s (2020) report “The Researcher Journal through a Gender Lens” examines knowledge production between men and women regarding research, career trajectories and perceptions across the EU (referred to as the EU28) as well as 15 countries in 26 fields and disciplines. The report lays out disparities with women having a smaller footprint in the research landscape, grants and patent applications than men. In research authorship, women publish for nearly as long as men over the course of their career however the ratio of women to men as authors has decreased over time. The gender balance in the EU is slowly improving, with approximately 1/3 of all researchers identifying as a woman (p. 49). For more on transgender equality see the European Commission (2020).

Institutionalized in university curricula, organizational structures and cultures as well as scientific ethos, Eurocentrism and whiteness continues to dominate spaces of learning and knowledge production in Europe (Joseph-Salisbury, 2019). Engrained in the fabric of academic disciplines, Eurocentrism perpetrates what Alcoff (2017) calls a transcendental delusion, “a belief that thought can be separated from its specific, embodied, and geo-historical source,” claims universality and de-legitimizes other epistemologies that come from non-Western geo-historical sources (p. 397). Thus, Eurocentrism is predicated on epistemic injustice. Globally, epistemic injustice is manifested in the creation of “scientific ghettos” in the Global South where knowledge production is often directed by the interests and publication venues of the Global North, of which the European Union is a prominent actor (Dübgen, 2020).

Concluding remarks

The European Union and colleges and universities throughout the EU face a double pandemic of Covid-19 and racism. Covid-19 has caused EU higher education institutions to focus on physical health of their students, administrators and staff; environmental health factors such as cleaning, classroom chair locations, limiting visits to campus, and beyond; request faculty switch to distance learning, to name a few. “The other pandemic: Systematic racism and its consequences” (Elizondo-Urrestarazu, 2020, p. 1) has been a global initiative of awareness, protest, and a re-construction of knowledge around race where education serves as one of the main foci (along with housing, banking, criminal justice, public and health) of internalized, institutional and systemic oppression.

Myriad organizations such as Equinet, the European Network of Equality Bodies, have focused on the Covid-19 response and systemic racism. The European Network Against Racism (ENAR) has worked toward educating people about racism, including securing equality data, racial profiling, and policing. Further, the EU Fundamental Rights Agency (FRA) (2019) published a report focusing on discrimination faced by Black people in the EU which, we argue, continues to limit knowledge production throughout the Continent.

As institutions, states, supranational organizations including the EHEA, partnerships and cross-collaborations change in this new era, the production of knowledge will continue to change. Covid-19 and racial injustice have had an impact on the Bologna Process, science parks, technology transfer, and the like. In the field of higher education specifically, time is ripe to contest the post-colonial heritage of European higher education as a “powerful system of knowledge, constructed in contrast to, and in conflict with, non-European ‘intellectual traditions’” and to examine its neoliberal global ambitions (Ploner and Dana, 2020, p. 375). Challenging the racialized knowledge production in the historically white universities of Europe and the testimonial injustice that non-Western academics face in establishing their scholarly credibility in Europe (Thapar-Björkert and Farahani, 2019) should contribute to more equitable, inclusive, and productive research partnerships and collaborations.

References

- Alcoff, L.M., 2017. Philosophy and philosophical practice: Eurocentrism as an epistemology of ignorance. In: Kidd, I.J., Medina, J., Pohlhaus Jr., G. (Eds.), *The Routledge Handbook of Epistemic Injustice*. Routledge, pp. 397–408.
- Algieri, B., Aquino, A., Succurro, M., 2013. Technology transfer offices and academic spin-off creating: the case of Italy. *J. Technol. Transf.* 38, 381–400.
- Altbach, P.G., 2016. *Global Perspectives on Higher Education*. Johns Hopkins University Press, Baltimore, MD.
- Apple, M., 2000. *Official Knowledge: Democratic Education in a Conservative Age*, second ed. Routledge, New York, NY.
- Apple, M., 2006a. Interrupting the right: on doing critical educational work in conservative times. In: Ladson-Billings, G., Tate, W.F. (Eds.), *Education Research in the Public Interest: Social Justice, Action, and Policy*. Teachers College Press, New York, NY, pp. 27–45.
- Apple, M., 2006b. Educating the “Right” Way: Markets, Standards, God, and Inequality, second ed. Routledge, New York, NY.
- Barnett, R., 2005. Convergence in higher education: the strange case of “Entrepreneurialism”. *High Educ. Manag. Pol.* 17 (3), 51–64.
- Belitski, M., Aginskaja, A., Marozau, R., 2019. Commercializing university research in transition economies: technology transfer offices or direct industrial funding? *Res. Pol.* 48 (3), 601–615.
- Bengtsson, L., 2017. A comparison of university technology transfer offices’ commercialization strategies in the Scandinavian countries. *Sci. Publ. Pol.* 44 (4), 567–577.
- Blackmore, J., 2021. Governing knowledge in the entrepreneurial university: a feminist account of structural, cultural and political epistemic injustice. *Crit. Stud. Educ.* <https://doi.org/10.1080/17508487.2020.1858912>.
- Bologna, 1999. Overarching Framework of Qualifications of the EHEA. Retrieved from: <http://www.ond.vlaanderen.be/hogeronderwijs/bologna/qf/overarching.asp>.
- Bosetti, V., Cattaneo, C., Verdolini, E., 2015. Migration of skilled workers and innovation: a European perspective. *J. Int. Econ.* 96 (2), 301–322.
- Brescia, F., Colombo, G., Landoni, P., 2016. Organizational structures of knowledge transfer offices: an analysis of the world’s top-ranked universities. *J. Technol. Tran.* 41, 132–151.
- Bush, J., 2009. Corporations, colonization, and higher education: the “benefits” of institutional collaboration. In: Gibertson, T., Gibertson, G. (Eds.), *The Knowledge Economy Academic and the Commodification of Higher Education*. Hampton Press, Cresskill, NJ, pp. 185–192.
- Cadorin, E., Klofsen, M., Löfsten, H., 2019. Science parks, talent attraction, and stakeholder involvement: an international study. *J. Technol. Tran.* <https://doi.org/10.1007/s10961-019-09753-w>.
- Calvo, N., Fernández-López, S., Rodeiro-Pasos, D., 2019. Is university-industry collaboration biased by sex criteria? *Knowl. Manag. Res. Pract.* 17 (4), 408–420.
- Carayannis, E.G., Campbell, D.F.J., 2012. Mode 3 Knowledge Production in Quadruple Helix Innovation Systems: 21st Century Democracy, Innovation, and Entrepreneurship for Development. Springer, New York.
- Chomsky, N., 1999. *Profit Over People: Neoliberalism and Global Order*. Seven Stories Press, New York, NY.
- Dakowska, D., Harmsen, R., 2015. Laboratories of reform? Europeanization and internationalization of higher education in central and Eastern Europe. *Eur. J. High Educ.* 5 (1), 4–17.
- Day, S., Pirrie, A., 2021. Tales from the matrix: student satisfaction as a form of governmentality. In: Murphy, M., Burke, C., Costa, C., Raaper, R. (Eds.), *Social Theory and the Politics of Higher Education: Critical Perspectives on Institutional Research*. Bloomsbury Academic, pp. 45–64.
- Dieckhoff, M., Gallie, D., 2007. The renewed Lisbon Strategy and social exclusion policy. *Ind. Relat. J.* 38 (6), 480–502.
- Dotson, K., 2012a. Tracking epistemic violence, tracking practices of silencing. *Hypatia* 26 (2), 236–257.
- Dotson, K., 2012b. A cautionary tale: on limiting epistemic oppression. *Frontiers* 33 (1), 24–47.
- Dübgen, F., 2020. Scientific ghettos and beyond. Epistemic injustice in academia and its effects on researching poverty. In: Beck, V., Hahn, H., Lepenies, R. (Eds.), *Dimensions of Poverty. Philosophy and Poverty*, vol. 2. Springer. https://doi.org/10.1007/978-3-030-31711-9_5.
- Dunning, J.H., 2000. Regions, globalization, and the knowledge economy: the issues stated. In: Dunning, J.H. (Ed.), *Regions, Globalization, and the Knowledge-Based Economy*. Oxford University Press, Oxford, England, pp. 7–41.
- Elizondo-Urrestarazu, J., 2020. The Other Pandemic: Systemic Racism and its Consequences. Equinet European Network of Equality Bodies. Retrieved on September 11, 2020 from: <https://equineteurope.org/2020/the-other-pandemic-systemic-racism-and-its-consequences/>.
- Erkilä, T., Piironen, O., 2014. Shifting fundamentals of European higher education governance: competition, ranking, autonomy and accountability. *Comp. Educ.* 50 (2), 177–191. <https://doi.org/10.1080/03050068.2013.807643/>.
- Etzkovits, H., Zhou, C., 2018. *The Triple Helix: University-Industry-Government Innovation and Entrepreneurship*. Routledge, New York.
- Etzkovits, H., 2003. The European entrepreneurial university: an alternative to the U.S. model. *Ind. High. Educ.* 17 (5), 325–335.
- Etzkovits, H., 2019. The second academic revolution: antecedents and consequences of academic entrepreneurship. In: Varga, A., Erdős, K. (Eds.), *Handbook of Universities and Regional Development*, pp. 29–43. <https://doi.org/10.4337/9781784715717>.
- Etzkovits, H., Leydersdorff, L., 2000. The dynamics of innovation: from national systems and “Mode 2” to a triple helix of university-industry-government relations. *Res. Pol.* 29 (2), 109–123.
- European Commission, 2020. Lesbian, Gay, Bi, Trans and Intersex Equality. Retrieved September 13, 2020 from: https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/lesbian-gay-bi-trans-and-intersex-equality_en.
- European Higher Education Area, 2020. European Higher Education Area and Bologna Process. Retrieved on September 7, 2020. <http://ehea.info/index.php>.
- European Union Fundamental Rights Agency (FRA), 2019. Being Black in the EU: Second European Union Minorities and Discrimination Survey. Publication office of the European Union, Vienna, Austria.
- Foucault, M., 1982. Afterword. In: Sawyer, L., Dreyfus, H.L., Rabinow, P. (Eds.), *Michel Foucault: Beyond Structuralism and Hermeneutics*. Harvester Press, Hemel Hempstead.
- Fricker, M., 2008. Forum on Miranda Fricker’s “epistemic injustice: power and ethics of knowing”: précis. *Int. J. Theor. Hist. Found. Sci.* 23 (1), 69–71.
- Gaston, P.L., 2010. *The Challenge of Bologna. What United States Higher Education Has to Learn from Europe and Why it Matters that We Learn it*. Stylus, Sterling, VA.
- Gildersleeve, E.R., Kuntz, A., Pasque, P.A., Carducci, R., 2010. The role of critical inquiry in (re)constructing the public agenda for higher education: confronting the conservative modernization of the academy. *Rev. High. Educ.* 34 (1), 85–121.
- Giroux, H.A., Giroux, S.S., 2004. *Take Back Higher Education: Race, Youth, and the Crisis of Democracy in the Post-civil Rights Era*. Palgrave-Macmillan, New York, NY.

- Giroux, H.A., 2002. Neoliberalism, corporate culture, and the promise of higher education: the university as a democratic public sphere. *Harv. Educ. Rev.* 72, 425–464.
- Hoekman, J., Frenken, K., van Oort, F., 2008. The geography of collaborative knowledge production in Europe. *Ann. Reg. Sci.* 43, 721–738.
- Hudson, R., 2016. Dominated by economics? Evidence of changing drivers of internationalization and its funding within higher education institutions in Europe. *High Educ. Pol.* 29, 1–19.
- Huisman, J., van Vught, F., 2009. Diversity in European higher education: historical trends and current policies. In: van Vught, F.A. (Ed.), *Higher Education Dynamics, Mapping the Higher Education Landscape: Towards a European Classification of Higher Education*, vol. 28. Springer, New York, NY, pp. 17–37.
- Joseph-Salisbury, R., 2019. Institutionalized whiteness, racial microaggressions and black bodies out of place in Higher Education. *White. Educ.* 41 (1), 1–17. <https://doi.org/10.1080/23793406.2019.1620629>.
- Kalar, B., Antoncic, B., 2015. The entrepreneurial university, academic activities and knowledge transfer in four European countries. *Technovation* 36–37, 1–11.
- Keeling, R., 2006. The Bologna Process and the Lisbon Research agenda: the European Commission's expanding role in higher education discourse. *Eur. J. Educ.* 41 (2), 203–223.
- Kitagawa, F., 2005. Entrepreneurial universities and the development of regional societies: a special view of the Europe of knowledge. *High Educ. Manag. Pol.* 17 (3), 65–88.
- Kriemadis, T., 2018. Innovation, creativity, new product development processes and the role of science parks. In: Vrontis, D., Weber, Y., Thrassou, A., Shams, S., Tsoukatos, E. (Eds.), *Innovation and Capacity Building*. Palgrave, pp. 231–254.
- Kwiek, M., 2012. Universities and knowledge production in central Europe. *Eur. Educ. Res. J.* 11 (1), 111–127.
- Laalo, H., Kinnari, H., Silvennoinen, H., 2019. Setting new standards for homo academicus: entrepreneurial university graduates on the EU agenda. *Eur. Educ.* 51 (2), 93–110. <https://doi.org/10.1080/10564934.2018.1489729>.
- Lavdas, K.A., Papadakis, N.E., Gidaraku, M., 2006. Policies and networks in the construction of the European higher education area. *High Educ. Manag. Pol.* 18 (1), 121–132.
- Link, A.N., Scott, J.T., 2007. The economics of university research parks. *Oxf. Rev. Econ. Pol.* 23 (4), 661–674.
- Löfsten, H., Klotfsten, M., Cadorin, E., 2020. Science parks and talent attraction management: university students as a strategic resource for innovation and entrepreneurship. *Eur. Plann. Stud.* <https://doi.org/10.1080/09654313.2020.1722986>.
- Lorde, A., 1984/2007. Age, race, class, and sex: women redefining difference. In: Lorde, A. (Ed.), *Sister Outsider*. Crossing Press, New York, NY, pp. 114–123.
- Mascarenhas, C., Marques, C.S., Galvão, A.R., Santos, S., 2017. Entrepreneurial university: toward a better understanding of past trends and future directions. *J. Enterprising Communities People Places Glob. Econ.* 11 (3), 316–338.
- McGovern, S., 2012. *Education, Modern Development, and Indigenous Knowledge: An Analysis of Academic Knowledge Production*. Routledge.
- McMahon, W.W., 2009. *Higher Learning, Greater Good: The Private and Social Benefits of Higher Education*. John Hopkins University Press, Baltimore, MA.
- Muscio, A., 2010. What drives the university use of technology transfer offices? *J. Technol. Transf.* 35, 181–202.
- Naidoo, R., 2011. Rethinking development: higher education and the new imperialism. In: King, R., Marginson, S., Naidoo, R. (Eds.), *Handbook on Globalization and Higher Education*. Edward Elgar Publishing Limited, Cheltenham, UK, pp. 40–58.
- Neff, D., 1999. *A Little Knowledge is a Dangerous Thing: Understanding Our Global Knowledge Economy*. Butterworth-Heinemann, Woburn, MA.
- Newby, H., 2008. The challenge to European universities in the emerging global marketplace. In: Weber, L.E., Duderstadt, J.J. (Eds.), *The Globalization of Higher Education*. Economica, London, England, pp. 55–63.
- Ng, W.K.B., Appel-Meulenbroek, R., Cloodt, M., Arentze, T., 2019. Towards a segmentation of science parks: a typology study on science parks in Europe. *Res. Pol.* 48 (3), 719–732.
- Nokkala, T., 2015. The Bologna process and the role of higher education: discursive construction of the European higher education area. In: Enders, J., Jongbloed, B. (Eds.), *Public-Private Dynamics in Higher Education*. Transcript Verlag, pp. 221–246.
- Nowotny, H., Scott, P., Gibbons, M., 2003. "Mode 2" revisited: the new production of knowledge. *Minerva* 41 (3), 179–194.
- O'Kane, C., Mangematin, V., Geoghegan, W., Fitzgerald, C., 2015. University technology transfer offices: the search for identity to build legitimacy. *Res. Pol.* 44 (2), 421–437.
- Olcay, G.A., Bulu, M., 2016. Technoparks and technology transfer offices as drivers of an innovation economy: lessons from Istanbul innovation spaces. *J. Urban Technol.* 23 (1), 71–93.
- Ozga, J., 2008. Governing knowledge: research steering and research quality. *Eur. Educ. Res. J.* 7 (3), 261–272.
- Pasque, P.A., Carducci, R., 2015. Critical advocacy perspectives on organization in higher education. In: Paulsen, M.B. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. 29. Springer, New York, NY, pp. 275–333, 30.
- Pasque, P.A., 2010/2014. *American Higher Education, Leadership, and Policy: Critical Issues and the Public Good*. Palgrave Macmillan, New York, NY.
- Ploner, J., Dana, C., 2020. International student migration and the postcolonial heritage of European higher education: perspectives from Portugal and the UK. *High Educ.* 80, 373–389.
- Posselt, J.R., 2009. The rise and fall of need-based grants: a critical review of presidential discourses on higher education. In: *Higher Education: Handbook of Theory and Research*, vol. 24, pp. 183–226.
- Ravinet, P., 2008. From voluntary participation to monitored coordination: why European countries feel increasingly bound by their commitment to the Bologna Process. *Eur. J. Educ.* 43 (3), 354–369.
- Rhoads, R.A., Liu, A., 2009. Globalization, social movements, and the American University: implications for research and practice. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XXIV. Springer, New York, NY, pp. 273–315.
- Rich, D., 2010. The Bologna process in European higher education. In: *The International Encyclopedia of Education*, third ed., pp. 566–572. <https://www.sciencedirect.com/science/article/pii/B9780080448947008484>.
- Rinne, R., Koivula, J., 2005. The changing place of the university and a clash of values: the entrepreneurial university in the European knowledge society. A review of the literature. *High. Educ. Policy Manag.* 17 (3), 91–123.
- Robertson, S.L., 2010. The EU, "regulatory state regionalism" and the new modes of higher education governance. *Glob. Soc. Educ.* 8 (1), 23–37.
- Schattock, M., 2005. European universities for entrepreneurship: their role in the Europe of knowledge. The theoretical context. *High. Educ. Policy Manag.* 17 (3), 13–25.
- Schulze-Cleven, T., 2016. *Higher Education Reloaded: Studying the Politics of Institutional Change*, vol. 2. Europe Now. Available at: <https://www.europenowjournal.org/2016/11/30/higher-education-reloaded-studying-the-politics-of-institutional-change/>.
- Siegel, D.S., Wright, M., 2015. University technology transfer offices, licensing, and start-ups. In: Link, A.N., Siegel, D.S., Wright, M. (Eds.), *Chicago Handbook of University Technology Transfer and Entrepreneurship*. The University of Chicago Press, Chicago, IL, pp. 1–40.
- Siegel, D.S., Veugelers, R., Wright, M., 2007. Technology transfer offices and commercialization of university intellectual property: performance and policy implications. *Oxf. Rev. Econ. Pol.* 23 (4), 640–660.
- Sinell, A., Müller-Wieland, R., Muschner, A., 2018. Gender-specific constraints on academic entrepreneurship and engagement in knowledge and technology transfer. *Technol. Innov. Manag. Rev.* 8 (2), 15–26.
- Slaughter, S., Rhoades, G., 2004. *Academic Capitalism and the New Economy: Markets, State and Higher Education*. Johns Hopkins University Press, Baltimore, MD.
- Sørensen, M.P., Bloch, C., Young, M., 2015. Excellence in the knowledge-based economy: from scientific to research excellence. *Eur. J. High Educ.* 6 (3), 217–236.
- Sursock, A., Smidt, H., 2010. *Trends 2010: A Decade of Change in European Higher Education*. European University Association.
- Teixeira, S.J., Veiga, P.M., Fernandes, C.A., 2019. The knowledge transfer and cooperation between universities and enterprises. *Knowl. Manag. Res. Pract.* 17, 449–460.
- Thapar-Björkert, S., Farahani, F., 2019. Epistemic modalities of racialized knowledge production in the Swedish academy. *Ethn. Racial Stud.* 42 (16), 1–21.
- Thursby, J.G., Thursby, M.C., 2007. University licensing. *Oxf. Rev. Econ. Pol.* 23 (4), 620–639.
- Tight, M., 2019. Mass higher education and massification. *High Educ. Pol.* 32, 93–108.
- Tomusk, V., 2004. Three bolognas and a pizza pie: notes on institutionalization of the European higher education system. *Int. Stud. Sociol. Educ.* 14 (1), 75–95.

- van Vught, F.A., 2009. The Europe of knowledge. In: Dill, D.D., van Vught, F.A. (Eds.), *National Innovation and the Academic Research Enterprise*. Johns Hopkins University Press, Baltimore, MD.
- Vukasovic, M., 2013. Change of higher education in response to European pressures: conceptualization and operationalization of Europeanization of higher education. *High Educ.* 66, 311–324.
- Williams, G., Kitaev, I., 2005. Overview of national policy contexts for entrepreneurialism in higher education institutions. *High. Educ. Policy Manag.* 17 (3), 125–141.
- Winckler, G., 2008. Comprehensive universities in continental Europe: falling behind? In: Weber, L.E., Duderstadt, J.J. (Eds.), *The Globalization of Higher Education*. Economica, London, England, pp. 65–75.
- Zumeta, W., Kinne, A., 2011. Accountability policies: directions old and new. In: Heller, D.E. (Ed.), *The States and Public Higher Education Policy: Affordability, Access and Accountability*, second ed. Johns Hopkins University Press, Baltimore, MD, pp. 173–199.

Transformation of university missions—from nation-building to a focus on common global challenges reflections from East Asia, under the risk of being “lost in translation”

Akiyoshi Yonezawa^a, Sae Shimauchi^b, and Jing Liu^c, ^a International Strategy Office, Tohoku University, Sendai, Japan; ^b International Center, Tokyo Metropolitan University, Hachioji, Japan; and ^c Graduate School of Education, Tohoku University, Sendai, Japan

© 2023 Elsevier Ltd. All rights reserved.

Introduction	21
Implantation and development of the modern university, through translation (from the mid-19th century to the 1980s)	22
Japan: society based on translation	22
(Greater) China: seeking a hybrid identity	23
Korea: challenge to its unique identity	24
Transformation of the concept of “university” in East Asia through globalization (1990s–present)	25
Japan: growing interest in Asia	25
China (greater China): challenge to the “world class”	26
Korea: active ambition toward globalization	26
Conclusion and discussion	27
References	27

Introduction

Today, most universities—especially those claiming to be “world-class”—tend to profess missions focused on leading in the solution of common global challenges, such as the United Nations Sustainable Development Goals and Industry 4.0 (Stensaker et al., 2019). Many such institutions serve as “flagship” universities, taking responsibility for the development of the nation-states wherein they are based (Douglass, 2016). Under the new geopolitics of higher education, these world-class universities play an increasing role in promoting the global public good (Marginson, 2018). Two ideas underlie this exploration: first, that universities (or at least world-class universities) share a common identity as academic institutions, globally; and, second, that they serve beyond their national borders.

Yet universities throughout the world have frequently faced crises, as interventions from both religious and secular powers have challenged their academic freedom and autonomy. This situation continues even today (Neave, 2001). In regions with non-European traditional civilizations, such as East Asia, modern university systems were launched, developed, and transformed in particularly close association with nation-building after the mid-19th century. The articulation of Western imported ideas around universities—alongside Eastern intellectual traditions and identities—has constantly been a central issue throughout this process (Chan et al., 2017).

Recent decades have seen the dramatic rise of East Asian national economies, based on science and technology. This has driven national leaders and higher education researchers in the region to seek identities for their own universities and higher education systems that are different from those of Western universities—and that are both similar and diverse within the region (Marginson et al., 2011). The emergence of globally top-ranked universities in East Asia has accelerated this trend (Yonezawa, 2021). For example, the achievements of Singapore’s highly ranked top universities indicate that it is possible to establish world-class universities on a strong national basis. Yet a clear consensus on academic freedom and university autonomy is still lacking, while substantial changes to implement more decentralized university governance are ongoing.

The translation activities at a university can be defined as (1) the translation of the texts it provides and uses, (2) the translation for the lectures and academic dialog, and (3) the development of the university system and the concept of the university through translation. The modern need for a more complex global “translation” of the identity and concept of a university constitutes a departure from the medieval European universities, where the main functions played by translation were limited to (1) the translation of text from Arabic and other languages, into Latin as the common language; and (2) lectures and dialogs, among academics from various linguistic backgrounds, carried out in Latin. While medieval European universities had varied backgrounds and characteristics—such as the universities born in Bologna and Paris—there was a basic consensus on the system and concept of universities as autonomous academic communities, authorized by secular and/or sacred powers (Neave, 2001).

The translation and definition of the contemporary university in East Asia, in contrast, comprises all three perspectives noted above. East Asian universities’ concepts and identities have been established and transformed amid diverse translations—both to and from the multiple languages and cultures existing in the region. The concept of the university in East Asia has thus been formed and transformed in close association with the various societies of the region, through the path dependence mechanism where concepts introduced through translation in the past could be alive and transformed internally.

This article aims to examine how concepts around the “university” were formed and developed in East Asia—with a particular focus on translation as a function for the transmission, consolidation, and sharing of knowledge. From the late 19th century onwards, the region’s universities were formed—following, as models, the universities developed by the modern nation-states of the Western world. Under the strong influence of nationalism, East Asian universities were created using each country’s own national language. This article analyses how the formation and development of the “university” as a concept have been influenced by changing international and national social contexts—first, from the construction of modern university systems to the end of the Cold War; and then continuing on into the globalization age, up to today.

Implantation and development of the modern university, through translation (from the mid-19th century to the 1980s)

Japan: society based on translation

The use of the term “*Daigaku* (大学)”—or “university”—in Japan’s educational history is basically limited to its history as a modern nation-state; a Western-style government was established through the Meiji Restoration in 1868. Amano (2009), however, argues that Japan’s original university initiative, at the beginning of the Meiji Restoration, had the characteristics of a “restoration,” one that aimed to revive the “*Daigaku-Ryo*”—a higher education and research institute existed in eighth century, which strove to introduce advanced civilization from China—as a symbol of Japan’s own inherited intellectual history. Under the intervention of the new modern government, the building of modern universities and schools proceeded through various rather conflicting initiatives and interaction among different intellectual groups. These included: (1) *Yogaku* (Western study), which had been developed through the limited contact with and importation of Western civilization during the Edo period; (2) *Kangaku* (Chinese study), based on the continuous intellectual input received from China since ancient times; and (3) *Kokugaku* (Japanese study), representing Japan’s own intellectual heritage (Yonezawa et al., 2017). The concept of the modern Japanese “university” had thus been created and verbalized in the Japanese system, prior to modernization. As an East Asian tradition, this included the training of bureaucrats and technocrats—which was strongly expected and emphasized as a function of universities and higher education institutions.

Still, Yoshimi (2011) sees the concept of *Daigaku* as one based on the Humboldt philosophy of research and education, developed in Germany since the beginning of the 19th century—a major university model that influenced Japan. The philosophy and reality of the university were diverse, especially among private institutions. Joseph Hardy Neesima and Umeko Tsuda, who were originally born in Japan and studied at liberal arts colleges in the United States, started their own schools—which they first opened as English language schools—in Japan after the Meiji Restoration. Many private higher education institutions started out as professional training schools in various fields (e.g., law, medicine, and languages) and were later transformed into comprehensive universities (e.g., Doshisha University and Tsuda University).

The new government, established in 1868, founded and developed its Western-style education system through a series of ordinances that were mainly developed through trial-and-error. The Education System Ordinance of 1872 planned to establish eight *Daigakko* (大学校; *Grandes Écoles*) in Japan; yet these new higher education institutions were not clearly defined and the timeline for their establishment was not stated. The Education Ordinance of 1879, which replaced the Education System Ordinance, contained a *pro forma* legal definition of the *Daigakko* as an institution where specialized courses in law, science, medicine, literature, etc., were to be taught.

In 1877, the Tokyo *Daigaku* (東京大学; the University of Tokyo), was established as the first national university in Japan. It was a Western-style comprehensive university offering faculties of law, science, letters, and medicine. The term *Daigaku* became an official word for “university,” thereafter. Merging Imperial College of Engineering as the faculty of engineering, the University of Tokyo was renamed “Imperial University” in 1886. Its name was changed again—this time, to “Imperial University of Tokyo”—when Kyoto Imperial University was launched as the country’s second imperial university in 1897. According to the Imperial University Ordinance, enacted in 1886, these imperial universities aimed to offer instruction in the arts and sciences, which were indispensable requirements for the nation, and to conduct in-depth research. The 1918 University Ordinance officially authorized private universities, reiterating the aim of such universities (as in that of the imperial universities, above) and specifying that universities should pay attention to the cultivation of personality and national thought. This was also applied to Keijo Imperial University (in what is now Seoul, Korea) and Taipei Imperial University (in what is now Taipei, Taiwan), under Japanese colonization.

After World War II, Japan’s educational system was comprehensively reformed. Under the Allied occupation and with the arrival of the US Education Mission, a new university system was developed in Japan—with the accompanying “translation” between English and Japanese languages and traditions (Tsuchimochi, 1993). The 1947 School Education Act states: “the university aims to educate and research a wide range of specialized arts and professional knowledge, and to develop intellectual, moral and applied abilities, as well as to widely impart knowledge as a center of (the) academy” (Article 52). There are clearly significant changes to the rules governing the relationship between universities and the state, before and after the war.

After World War II and up to the end of the Cold War, the discussion around higher education research and policy focused on differentiating functions to provide mass and universal higher or tertiary education. Japan was ahead of its neighboring countries in East Asia at that time, in terms of broadening participation in higher education—with private universities and higher education institutions playing a major role in the expansion of the country’s higher education system. Until the end of the Cold War, Japan thus made almost no serious reference to models from neighboring East Asia countries to indicate the direction for Japanese universities and higher education.

In the 1970s, the discussion around higher education in Japan was oriented toward reaching the masses. Professional training colleges were launched in Japan in 1976, as non-university post-secondary education institutions—using the trend toward promoting the non-university education sector in Europe and the United States as a reference. Junior colleges, in contrast, have been constantly recognized as part of the university sector in Japan, while their de facto mission became short term higher education for female students. More recently, in 2019, the Japanese educational sector launched a new category of “professional university,” to meet needs for professional and vocation-oriented university education. Along with these, various *Daigakko* serve educational needs linked to ministries and agencies other than the Ministry of Education, Culture, Sports, Science and Technology (MEXT), and are affiliated with said institutions (e.g., the National Defense Academy falls under the Ministry of Defense).

(Greater) China: seeking a hybrid identity

In contrast with Japan, the conceptualization and formation of modern universities in China reflect various aspects of China’s unique context. First, in the Qing Dynasty of the 17th to the beginning of 20th century, the Manchu were the dominant ethnic group. Although the group’s native language was Manchu, administrative Chinese (the official language of Beijing) was widely used as a common language. Academic institutions’ activities were conducted in Chinese, as were the Chinese imperial examinations. Literature, knowledge, and other aspects of culture were also widely distributed to neighboring countries in Chinese.

Second, some elements of Western literature and knowledge had already been translated and transplanted into Chinese. Literature on Western societies that had been edited in Chinese—such as the 1842 “*Haiguo Tuzhi* (Atlas and Description of the Countries beyond the Seas)” —made a great impact in Japan at the end of the Edo era (Agawa, 2011). These translations were different from those translated into Japanese, which had the character of “transplanting advanced civilization to the local society.” The Chinese-language translations, instead, were based on Chinese thought. Their stance was open to the concept of “the application of Western learning.” Here, Western science and technology were intended to be coupled with Confucian thought—as well as with the Chinese imperial examination system and civil servant training, which had an impact on meritocracy through examinations in Western society. For example, the *Beijing Guozijian* (北京国子监), the Dynasty’s higher education institution—which functioned in close relationship with the national civil servant examinations and training—was reorganized into the Imperial University of Peking in 1898. It then became Peking University in 1912, after the Xinhai Revolution. A reform to strengthen the features of the modernized university was promoted, after Cai Yuanpei became President in 1917 (Yang, 2003). Along with the impact of Western universities, China’s establishment of the Imperial University of Peking was also influenced by Japan’s establishment of modern universities (Zhao, 1999). In other words, the *Daigaku* 大学 was first formulated as a modern concept of “university” in Japan. The *Dà Xué* 大学 in China referred to this and was defined as a comprehensive university. More than three such institutes were established under the regime of the Republic of China before World War II (Liu and Xu, 2004). Japanese intellectuals like Okuma (1922)—the founder of Waseda University, a top private university in Japan—intended to maintain a harmony between the East and the West as two equal civilizations. This is consistent with the positive acceptance of international students from Asia, especially from China, by Japanese universities and higher education institutions (Li and Liu, 2015).

Third, since China is a nation-state that covers a vast expanse of land and diverse regions—including, especially in its border regions, a variety of multi-ethnic communities—it experiences institutional transplant due to direct and indirect interference, as well as influence from foreign countries (non-Chinese language-speaking areas). This has also had an impact on the establishment of modern universities in China. Examples include a group of universities in Hong Kong that were established after the Opium War—as well as Kenkoku University, founded in Manchuria. The Tsinghua Xuetang, as the predecessor of Tsinghua University in contemporary China, was also established as a preparatory school for study at American universities. The establishment is a result of diplomatic negotiations with the United States (Hayhoe, 1996; Saito, 1990; Zhu, 2013). Private universities in China were likewise established with support by the US, Europe and Japan. This spurred a diversification of Chinese universities, up to the first half of the 20th century.

Fourth, Greater China—as an area with a common Chinese language and civilization—has been spreading across multiple countries and administrations. Accordingly, it encompasses various interpretations of the concept of “university.” From the Xinhai Revolution in 1911 to the founding of the People’s Republic of China in 1949, universities in China experienced rapid changes due to different localities’ separation and merger at the hands of warlords, civil war, and the Sino-Japanese war.

In the communist party-controlled areas of the People’s Republic of China, universities responsible for fostering leaders and secretariats in the liberated areas became the mainstream universities, together with universities supported by churches and private universities. There was a reorganization of domestic universities and higher education institutions.

In 1949, the Chinese “People’s Political Consultative Conference Common Program” was announced. Based on the experience of the areas liberated by the communist party, the new democracy education would absorb a significant experience of the education under the communist regime, taking advantage of the Soviet experience. A large number of supervisors from the Soviet Union were then dispatched to universities in China. The number of comprehensive universities that aimed at fostering experts decreased, while the number of specialized colleges increased. The government’s educational department was also centralized and institutions were adjusted to put a stronger emphasis on science and engineering (Otsuka, 1996; Yang, 2003).

The development of China’s universities in 1950s and 1960s became more dependent on the changes in its domestic and international situation. In the era of the Great Leap Forward, starting in 1958, the administrative authority of many higher education institutions was delegated to local governments and each provincial government established a comprehensive university as the center of each province. This aimed to promote the expansion of higher education opportunities. With the Great Cultural

Revolution, starting in 1966, the idea of “class struggle” was brought into universities; many university teachers and intellectuals were sent to re-education institutions in rural areas. As in Soviet Union-style higher education, Chinese education reform, which had emphasized “expertize” in the 1950s, was now totally denied. Instead, the government promoted higher education institutions that could provide curriculum and educational activities aiming to lead to greater labor productivity (Zhang, 2009; Otsuka, 1996).

In 1978, following the end of the Cultural Revolution, it was declared at the National Education Conference that the goal of educational development was to develop the human resources necessary for the “modernization” of the country. Disciplines deemed necessary for economic construction were expanded, in contrast to the previous “Science and Engineering centered” system. Adult higher education institutions, a self-study examination system, and satellite TV education were developed—aiming to accommodate the demand for higher education among those who were unable to go to universities and colleges during the Cultural Revolution. In 1978, the “key university” system with direct and prioritized control and support by the central government was also reopened (Nanbu, 2009).

It should be noted that, throughout the Cold War and in mainland China under the communist regime, universities and academic activities were conducted in Chinese. This also applies to Taiwan, which was under the *Kuomintang* (KMT; National Party) administration. Taiwan’s higher education was affected by the United States and other pre-war Chinese universities under the control of the KMT, the post-war US influence, and the long-standing influence of Japanese colonial control (Hou et al., 2021). As indicated by Wu et al. (1989), Taiwan’s citizens did not understand Chinese (Mandarin); directly after their return to China, they required extensive tutoring to learn the Mandarin required for study at Chinese universities. Yet, according to Ogawa and Nanbu (2008), Taiwan’s education system after 1950—under the administration of Chiang Kai-shek—had basic features such as a centralized education system, the usage of Mandarin (national language), and the “Three Principles of the People” of Sun Yat-sen. There was powerful government control over management and operation, including student recruitment, rights of curriculum making, teacher recruitment, and personnel management. The authors also note that “university autonomy” and “academic freedom” were limited, under the state of emergency.

In Hong Kong, under UK control (and especially at the Chinese University of Hong Kong, where the medium of instruction was Chinese)—as well as in Singapore, where Nanyang University existed temporarily—countries and administrative areas with large Chinese diaspora communities sought to construct Chinese language-based universities.

Korea: challenge to its unique identity

In the 19th century, the Korean Peninsula was under the control of the Joseon Dynasty. The Korean Empire was established as a modern nation-state in 1897, although it was soon merged with Japan in 1910. After World War II, the Korean War resulted in the territory’s separation into the Republic of Korea and the Democratic People’s Republic of Korea. Both countries have developed universities and higher education systems. Given the limitations on literature from North Korea, this paper focuses only on South Korea.

Umakoshi (1995) found that the development of the modern university concept in the Korean Peninsula—from the Joseon Dynasty of the enlightenment period, to the Japanese regime, and up to the modern period through the US military administration—was a history created by the Korean acceptance and resistance of external factors. These include the influences of the Qing Dynasty (China), Japan, and the United States; and the exploration of a unique identity for Korea.

Sungkyunkwan has a long history from the end of 14th century in taking responsibility for fostering bureaucracy—based on the imperial examination system, starting in the Korea Dynasty. In the enlightenment period of the 19th century, this institution transformed itself into one that taught the “new modern subjects,” such as geography and arithmetic. After Korea was colonized by Japan in 1910, Sungkyunkwan was reorganized into an academic research institute and focused on the study of Confucianism. There were also many traditional Korean private schools and modern private schools, many of which had been established by Christian missionaries from the United States and Europe. The Japanese colonizers adopted a policy of restraining the establishment of higher education institutions in Korean Peninsula. There were only two officially designated professional colleges: Yonhi College, which later became Yonsei University; and Busong College, which later became Korea University. Apart from Keijo Imperial University, no other university was established in Korea until the end of its post-war liberation (Kim, 2018a). In the interim, there was a spread of Korean-medium education and knowledge production in Korean language inside and outside of Korea. For example, Christian missionaries privately published Hangul-language books; these later became the foundation for Korean Studies in Europe and the United States (Asami and An, 2012).

Keijo Imperial University was the only *Daigaku* (university) on the Korean Peninsula, under Japanese occupation. Based on Japan’s policy principle of “Integration of Korea and Japan,” the educational curriculum—as well as the faculty members and teaching language (Japanese)—were all transplanted from Japan. Opportunities for higher education were extremely limited, as the share of Korean students at the Keijo Imperial Universities was 40% in maximum (Umakoshi, 1995).

After the founding of the Republic of Korea in 1948, an educational law was enacted in 1949 under the Rhee Syngman administration. The law clearly stipulates the development of “the spirit of patriotism” and the “succession and uplifting of ethnic culture,” as the national educational policy. The Keijo Imperial University was officially abolished and reconstructed into the Seoul National University (*Seoul Daehakgyo* 서울대학교), by integrating it with several former governmental professional colleges. The first president of Seoul National University was Harry B. Ansted, an American. In the process of “De-Japanization”, exchanges of faculty members and financial supports mainly from the US were implemented; many faculty members had received their degrees in the US. As a result, South Korea’s own universities were formed under the influence of the US system of higher education.

Meanwhile, one out of every four Korean graduates of Keijo Imperial University became a university professor (Jung, 2002) after the liberation, and Japanese literature was widely read until around 1970 (Kim, 2018b).

Along with the 4-year comprehensive universities, with their various faculties, a number of colleges were established to focus on single disciplines. Due to the demands of industry and to Koreans' strong aspirations for higher learning, many colleges became comprehensive universities in the 1980s as part of the process of providing mass higher education. Vocational schools were also upgraded to junior colleges and industrial universities. Many of these 2–4-year junior colleges and industrial universities have now been officially named *Daehakgyo* (universities).

Importantly, Korea's Constitution Revision of 1987 included a guarantee of universities' autonomy. Based on Article 9 of the Fundamental Law of Education, Article 28 of the Higher Education Act implemented in 1998 defines the purpose of the university as: "(to) foster a good personality, to contribute to the nation and human society by teaching and studying deep academic theories and their application methods necessary for the development of the nation and human society." Under this law, in addition to its (regular) universities, South Korea offers industrial universities, education universities, junior colleges, distance universities, and technical universities. Graduate schools have been classified into three categories: (1) general graduate schools for academic research and major educational purposes; (2) professional graduate schools providing the practical theory, research, and development required for the cultivation of human resources in professional fields; and (3) special graduate schools, whose main purpose is to offer vocational and continuing education.

Transformation of the concept of "university" in East Asia through globalization (1990s–present)

This article analyzes two time periods: before and after the end of the Cold War. According to Zhong (2018), this division of time periods corresponds to the rise of the neo-liberal alternative, which shifted thinking away from modernization as the premise of the nation-state and away from the theory of development. During the period of globalization after the Cold War, there were structural changes in higher education in Asia. Many Asians studied at US universities, then returned to Asia as faculty members. There was also an expansion of the international academic network, in which English was accepted as a common language. The structural changes can be summarized as below.

First, exchanges between China and the western world have significantly expanded. China's shift to reform and to opening up, starting at the end of the 1970s; its introduction of the market economy; and its restoration of diplomatic relations with the United States and Japan have all served to rapidly expand mutual understanding and networking between China and the West, including the promotion of study abroad and research exchanges. Ever-increasing globalization, after the collapse of the Soviet Union, has also led to a significant integration between the socialist and capitalist spheres—albeit with many twists and turns along the way.

Second, and in connection with the above, the concept of "East Asian models" for universities and higher education became crystalized. The World Bank (1993) reported that the role of education in human resource training was an important factor in social and economic development in the region, based on its analysis of Japan and of New Industrial Economies (NIEs) like South Korea and Taiwan. In Japan, research around the key features of the country's universities and education also progressed. Significant progress was likewise made, in research unveiling the common characteristics of higher education across East Asia. These include the value placed on a hierarchical system of higher education; and the high levels of investments made, along with the high value placed on educational qualifications, by families and by society at large (for example, Marginson et al., 2011).

Third, there has been a change to the prestige structure of universities in East Asia. Japan's top universities received world-wide attention at a very early stage and built dominant positions in Asia in the late 1990s. Universities in Hong Kong, South Korea, Taiwan, Singapore, and mainland China have subsequently risen rapidly in various types of university rankings. Conversely, Japanese universities have recently received less positive attention, against a backdrop of stagnant investment in higher education and delays in internationalization (Yonezawa, 2021).

Fourth, the social system has changed in East Asian countries and administrative regions, impacting higher education. Since the late 1980s, the democratization of Korea has had a major impact on expanding its higher education opportunities and systemic flexibility. Taiwan and Japan have also experienced changes in government, provoking a reconsideration of each country's path toward increased higher education opportunities. With the handing-over of Hong Kong and Macau to China, opportunities for higher education have expanded and educational systems and curriculum have changed.

Japan: growing interest in Asia

Against the background of changes in its economic growth and international situation, Japan has actively taken initiatives to internationalize higher education. The government launched a plan to accept 100,000 international students in 1983 and 300,000 international students in 2008, which was realized in 2019, just before the sustain of student mobility by the Covid-19 from 2020. To boost the international competitiveness of Japanese universities around science and technology, Japan has enacted the Basic Act of Science and Technology and promoted a project for capacity development at its research universities. The country's national and local public universities were incorporated in 2004. In 2013, the Japanese government set a policy goal to ensure that 10 Japanese universities were ranked among the top 100 in the world; this Top Global Universities project is being promoted as a 10-year project until 2023. Under the Democratic Party of Japan, from 2009 to 2012—proposed the joint promotion of quality assurance and student exchanges between Japan, Korea, and China. This Campus Asia project was later developed as part of the Inter-

University Exchange Project by the Japanese government, aiming at strategic cooperation with various countries and regions of the world. Yet Campus Asia, at present, does not function—in concept or in practice—as a framework for regional Higher Education, as does the European Higher Education Arena, while the efforts to develop regional higher education consortium covering East Asia and South East Asia are ongoing.

Japan's legal definition of a university remains as stated before in the School Education Law. However, the revision of the Fundamental Law of Education in 2006 included a new legal definition of “university”: “University as a center of academia would cultivate well-educated people and professional ability as well as creating new knowledge and, through providing these outcomes to the society, eventually contribute to the society” (Chapter 7 Article). Here, in addition to the functions of academic development and cultivating individuals, the university's relationship and contribution to society is given more emphasis.

Amid the process of transforming Japanese universities, there is a growing interest in Asian higher education—both in the region and in the world. Since the start of the 21st century, Japan has also begun to actively reference the higher education policies and practices carried out by its Asian neighbors (Jung et al., 2018).

China (greater China): challenge to the “world class”

In 1985, China announced the “Chinese Communist Party's Decision on Education System Reform.” The diversification of higher education institutions was implemented, and the main administrative system for universities shifted from the national level to the local level. The management autonomy of higher education institutions was also expanded; their capacity to adapt to the needs of economy and society was strengthened; private capital began to be permitted in higher education sector; the university curriculum was reformed; and teaching methods and promotion of research, based on “the president responsibility system,” were promoted. The expansion of exchanges with the United States and other Western countries has had a significant impact on the debate around the role of universities (Kaneko and Xu, 2006; Yang, 2003). In the 1990s, as part of “college and department adjustment,” specialized universities were merged and the establishment of comprehensive universities was promoted. In 1993, the central government launched the “211 project,” a policy of building key universities and key departments. It aimed at building 100 world-class universities over the 21st century. In 1998, the “985 Project” was announced. The project concentrates resources and investments on a smaller number of top universities, aiming to establish them as world-class (Huang, 2005; Yang, 2003). In 2017, the government launched a further policy: the “Double world-class university” project aims to build both world-class universities and world-class specialized fields (Yonezawa and Huang, 2021).

Simultaneously, the government is consistently making efforts to expand higher education opportunities and narrow regional disparities; collaboration with overseas universities and the inbound and outbound exchange of students are also being actively promoted. Chinese universities are also actively taking initiatives toward international expansion—such as the Confucius Institutes, which were undertaken through collaboration with universities. There is also the Belt and Road Initiative, linking China and the world through international cooperation (Jiang and Zhang, 2017).

International expansion and structural changes are also progressing in the Greater China region. In Taiwan, from 2011, the acceptance of a limited number of students from mainland China has been initiated (Wang, 2013). Universities in Hong Kong have also matched China's standards, to change their undergraduate programs from three years to four years. They have also introduced liberal arts education for the first year of undergraduate programs, per the advice of experts from American liberal arts colleges (Mok and Cheung, 2011). In China, “Shuyuan,” which is derived from the Confucian tradition, has been newly established; the initiative aims to develop a specific general education that was originally born in China. In addition, Chinese universities have developed branch campuses and programs in Southeast Asia, especially in places like Singapore and Malaysia that host a large Chinese diaspora. An exploration of the identity of Chinese universities and higher education—allowing for such innovations as “hybrid” universities, which are formulated through the mutual impact between the East and the West and apply an independent methodology—is not limited to the domestic level. Such explorations are also being developed at the global and Asian regional levels (Yan, 2016).

Still, “university” is not yet clearly defined in China's amended Higher Education Act of 2015. Moreover, its dual governance system—in which the University of China's Principal (President) and the Party Secretary share this responsibility—is maintained (CCP and the State Council, 2017). Furthermore, in the Greater China region, universities coexist with each other while maintaining diversity in a number of different political, social systems, and higher education systems (Lo, 2016).

Korea: active ambition toward globalization

Under Park Chung-hee's administration, South Korea's universities were controlled as institutions for nationalism and anti-communism. In the 1980—with the massification of higher education and social democratization—there was a diversification and expansion of higher education institutions, including the establishment of new universities and the upgrading of colleges to universities. Open universities were also established and educational universities were promoted to become four-year universities. Kim Young-sam's administration later launched the Korean version of globalization. During Kim Dae-jung's administration, the government implemented the Brain Korea 21 (BK21) Project. This aimed to build a high-quality human resources training system, by creating both world-class research-focused universities and universities with local characteristics. In recognition of the success of BK21 in terms of global competitiveness, the second stage of BK21 continued even under Roh Moo-hyun's administration.

Korea's dependence on overseas contacts (particularly in the United States) to foster researchers for the top universities has continued. Yet this also links Korean researchers with an international knowledge network (Yonezawa et al., 2016). The Korea Project has been implemented, aiming to accept international students and thus aid in the international dissemination of Korean studies and civilization. The project includes graduate schools for international studies, which carry out education and research in English; undergraduate programs, which are modeled after the liberal arts colleges of the United States; and education programs that allow international students to conduct research on Korean Studies in the Korean language (Shimauchi, 2016). Korea's presence in Asia has been strengthened—when it comes to both research and the meaningful exchange of experts, researchers, and students.

Conclusion and discussion

In East Asian universities and societies, the transmission, consolidation, and sharing of knowledge through translation have been implemented in patterns that are very different from those in medieval Europe (with Latin as the lingua franca) or in contemporary Europe (with English as a globally shared academic language). Modern universities in East Asia were established in close relation with each country's nation-building activities, and in line with a frequently changing hegemonic power context and geopolitics. The languages in use, and the patterns of translation, are closely connected to national identity—which also appears in the concepts and identities developed around universities in the region.

Under these diverse and dynamic conditions, East Asian societies have not yet been successful in forming a common understanding of the concepts and identities surrounding their own universities. There are always concerns over issues being “lost in translation”—which may lead to miscomprehension around the essential characteristics of universities, such as academic freedom and universities' autonomy.

In this context, it would be naïve to expect a seamless and straightforward transformation of the missions of East Asia's universities—from the nation-building of the past, toward a much-needed worldwide collaboration to tackle the global challenges of the present and future.

Let us recall that universities once closed, under the French Revolution—a fact that dramatically impacted the country for generations. History indicates that a national demand for science and technology, and for the development of high-skilled human resources, does not always result in supportive attitudes toward universities as autonomous academic communities. The global development of universities has stimulated an “arms race in knowledge,” and the close link between national governments and industry has tended to strongly link academic exchange and collaboration with national interests. Even in Japan—where the national constitution guarantees academic freedom—legal actions to limit international collaboration with researchers from specific countries are not uncommon.

The time has come for universities around the world to initiate mutual dialogs, once again seeking a common idea of a university in the contemporary context. These dialogs must be based not on a center-periphery structure, but on a mutual respect for diversity and on the need to address common global and regional challenges—challenges that are fast becoming impossible to ignore. Dialogs around today's idea of universities could be linked with postcolonial discourses. Yet, more importantly, they should be led by academics—across nations and institutions—who engage in self-reflection. Universities around the world can share an effective and meaningful common identity, in future, only if academics exert a strong will to create such an identity today.

References

- Agawa, S., 2011. Kaikoku Zushi to nihon [Haiguo Tuzhi and Japan] Gengo to Bunka 23, 1–15.
- Amano, I., 2009. Daigaku no Tanjo [The Birth of the University]. Chuko Shinsho, Tokyo.
- Asami, M., An, E., 2012. Kankoku to Kirisutokyo [Korea and Christianity]. Chuko Shinsho, Tokyo.
- Chan, S.J., Lee, M.N., Yang, R., 2017. The hybrid university in East Asia: searching for the new paradigm. *Stud. High Educ.* 42 (10), 1803–1808.
- CPC Central Committee and State Council, 2017. Guānyú jiāqiáng hé gāijin xīn xíngshì xià gāoxiào sixiǎng zhèngzhì gōngzuò de yìjiàn [Opinions on Strengthening and Improving Ideological and Political Work in Colleges and Universities Under New Situation]. People's Republic of China Government, Beijing. Available at: http://www.gov.cn/xinwen/2017-02/27/content_5182502.htm. (Accessed 3 January 2022).
- Douglass, J.A. (Ed.), 2016. *The New Flagship University: Changing the Paradigm From Global Ranking to National Relevancy*. Palgrave Macmillan, New York.
- Hayhoe, R., 1996. *China's Universities 1895–1995: A Century of Cultural Conflict*. Garland Publishing, New York.
- Hou, A.Y.C., Chiang, T., Chan, S., 2021. *Higher Education in Taiwan: Global, Political and Social Challenges and Future Trends*. Springer, Singapore.
- Huang, F., 2005. 1990 Nendai Iko no Chugoku Koto Kyoiku no Kaikaku to Kadai [Reforms and Challenges in Chinese Higher Education Since the 1990s]. Higashi-Hiroshima: Research Institute for Higher Education, Hiroshima University.
- Jiang, K., Zhang, J.F., 2017. Zhōngguó gāoděng jiàoyù duìwài kāifàng de jīběn tèdiǎn [Basic characteristics of opening-up of higher education in China] *Tsinghua J. Educ.* 38 (6), 7–15.
- Jung, J., Horta, H., Yonezawa, A. (Eds.), 2018. *Researching Higher Education in Asia: History, Development and Future*. Springer, Dordrecht.
- Jung, S.I., 2002. Gyeongseong jegug daehag-yeongu [Research of Keijo Imperial University]. Moonumsa Publishing, Seoul.
- Kaneko, M., Xu, G.X., 2006. Gāoděng jiàoyù fāzhǎn de zhōngguó móshì: Láizi rìbēn de guānchá [The Chinese model of higher education development: observations from Japan] *Res. Educ. Dev.* 9, 24–28.
- Kim, J.I., 2018a. Daehaggwa Gwonlyeog: Hangug Daehag 100 Nyeon-Ui Yeogsa [University and Power: 100 Years of Korean University History]. Humanist, Seoul.
- Kim, T., 2018b. Forming the Academic Profession in East Asia: A Comparative Analysis. Routledge, London.
- [International Students of Waseda: A Context Area of Knowledge in Modern Japan]. In: Li, C.S., Liu, J. (Eds.), 2015. *Ryugakusei no waseda: kindai nihon no chi no sesyoku ryoiki*. Waseda University Press, Tokyo.

- Liu, J.S., Xu, H., 2004. Zhōngguó jīndài gāoděng jiàoyù fāzhǎn lìchéng huígù (xià) [A review of the development of China's modern higher education (2)] J. Southeast Univ. 6 (2), 115–121.
- Lo, W.Y.W., 2016. The concept of greater China in higher education: adoptions, dynamics and implications. *Comp. Educ.* 52 (1), 26–43.
- Marginson, S., Kaur, S., Sawir, E. (Eds.), 2011. *Higher Education in the Asia-Pacific: Strategic Responses to Globalization*. Springer, Dordrecht.
- Marginson, S., 2018. *The New Geo-Politics of Higher Education*. Centre for Global Higher Education, Oxford.
- Mok, K.H., Cheung, A.B.L., 2011. Global aspirations and strategising for world-class status: new form of politics in higher education governance in Hong Kong. *J. High Educ. Pol. Manag.* 33 (3), 231–251.
- Nanbu, H., 2009. Chūgoku kōtō kyōiku dokugaku shiken seido no tenkai [Development of Higher Education Examination System for Self-Taught Learners in China]. Toshindo, Tokyo.
- Neave, G., 2001. The European dimension in higher education: an excursion into the modern use of historical analogues. In: Huisman, J., Maassen, P., Neave, G. (Eds.), *Higher Education and the Nation State: The International Dimension of Higher Education*. Pergamon, Oxford, pp. 13–73.
- Ogawa, Y., Nanbu, H., 2008. Taiwan no Koto Kyoiku: Genjo to Kaikaku Doko [Higher Education in Taiwan: Reform and Change in the 21st Century]. Hiagahi-Hiroshima: Research Institute for higher education, Hiroshima University.
- Okuma, S., 1922. *Tozai Bunmei no Chowa* [Reconciling Eastern and Western Civilisations]. Waseda University Press, Tokyo.
- Otsuka, Y., 1996. *Gendai Tyugoku Kotokyouku no Setsuritsu* [The Establishment of Higher Education in Modern China]. Tamagawa University Press, Tokyo.
- Saito, T., 1990. "Manshū-Koku" Kenkoku Daigaku no Sōsetsu to Tenkai [Foundation and development of "Kenkoku University"] *Occas. Pap.* 30, 110–132.
- Shimauchi, S., 2016. Higashijia ni okeru Ryūgakusei idō no paradaimu tankan [Paradigm Shift in the Movement of International Students in East Asia]. Toshindo, Tokyo.
- Stensaker, B., Lee, J.J., Rhoades, G., Ghosh, S., Castiello-Gutiérrez, S., Vance, H., Çalıkoğlu, A., Kramer, V., Liu, S., Marei, M.S., O'Toole, L., Pavlyutkin, I., Peel, C., 2019. Stratified university strategies: the shaping of institutional legitimacy in a global perspective. *J. High Educ.* 90 (4), 539–562.
- Tsuchimochi, G.H., 1993. *Education Reform in Postwar Japan: The 1946 US Education Mission*. University of Tokyo Press, Tokyo.
- Umakoshi, T., 1995. Kankoku Kindai Daigaku no Seiritsu to Tenkai: Daigaku Model no Denpa Kenkyu [The Establishment and Development of the Modern Korean Universities: A Study of the Propagation of the University Model]. Nagoya University Press, Nagoya.
- Wang, Z.H., 2013. Ping lùshēng lái tái jī qí zhēngyì [Commenting on the controversy over the arrival of Mainland students in Taiwan] *Taiwan Educ. Rev. Monthly* 2 (1), 29–31.
- World Bank, 1993. *The East Asian Miracle: Economic Growth and Public Policy*. Oxford University Press, Oxford.
- Wu, W., Chen, S., Wu, C., 1989. The Development of higher education in Taiwan. In: Altbach, P.G., Selvaratnam, V. (Eds.), *From Dependent to Autonomy: The Development of Asian Universities*. Kluwer Academic Publishers, Dordrecht, pp. 257–276.
- Yan, X., 2016. Wōguó gāoxiào jīngwài bànxué de dòngyīn fēnxī hé duìcè jiànyì [Motivation analysis and suggestions of Chinese Universities to go abroad to run schools] *J. High Educ. Manag.* 10 (3), 66–70.
- Yang, D.P., 2003. *Jiānnán de rì chū* [Hard Sunrise]. Wenhui Press, Shanghai.
- Yonezawa, A., Huang, F., 2021. World-class university policies and rankings in transition: a comparative study of the People's Republic of China and Japan. In: Liu, N., Wu, Y., Wang, Q. (Eds.), *World-Class Universities: Global Trends and Institutional Models*. Brill, Leiden, pp. 140–153.
- Yonezawa, A., Horta, H., Osawa, A., 2016. Mobility, formation and development of the academic profession in science, technology, engineering and mathematics in East and South East Asia. *Comp. Educ.* 52 (1), 44–61.
- Yonezawa, A., Hoshino, A., Shimauchi, S., 2017. Inter-and intra-regional dynamics on the idea of universities in East Asia: perspectives from Japan. *Stud. High Educ.* 42 (10), 1839–1852.
- Yonezawa, A., 2021. Reimagining university identities through rankings in Japan: the transformation of national policies and university behaviours in the broader East Asian context. In: Hazelkorn, E., Mihut, G. (Eds.), *Research Handbook on University Rankings: Theory, Methodology, Influence and Impact*. Edward Elgar, Cheltenham, pp. 231–246.
- Yoshimi, S., 2011. *Daigaku toha Nani ka* [What is University?]. Iwanami Shinsho, Tokyo.
- Zhang, Y., 2009. Zhōngdú 1958 nián gāoděng jiàoyù dà yuèjīn [Rereviewing the 1958 Great Leap forward in higher education] *J. East China Normal Univ.* 27 (1), 78–87.
- Zhao, J.M., 1999. Wūrúlún fù rì kǎochá yǔ zhōngguó xuézhì jīndài huà [Wu Rulun's Study tour to Japan and the modernization of the Chinese school system] *Archiv. Hist.* 5, 40–46.
- Zhong, Y., 2018. Magarikado no hito bungaku chi to nihon'nodaigaku no gurōbaru-ka [Humanities knowledge at a turning point and the globalisation of Japanese universities]. In: Tsuboi, H. (Ed.), *Baburu to ushinawareta 20-nen. Rinsen Book*, Kyoto, pp. 61–91.
- Zhu, L.L., 2013. Cōng qīnghuà xuétáng de biànciān kàn "gēngfú xīngxué" [Looking at the "Gengfu Xingxue" from the changes of Tsinghua Academy] *J. Puyang Vocat. Tech. Coll.* 26 (5), 53–56.

Economic impact of higher education on students, corporations and society

Peter Arena, Tim Carroll, Suzanne Lyons, and Liang Zhang, New York University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Impacts on students	30
Costs	30
Benefits	30
Variations in benefits across students	31
Impacts on corporations	32
Skilled and diverse workforce	32
Strategic relationships	33
Future research	33
Impacts on society	34
References	35

Economic analyses of investments in and returns on higher education are rooted in [Becker's \(2009\)](#) human capital framework. The central theme of human capital is that members of the labor force can acquire knowledge and skills by attending school or engaging in on-the-job training, and these investments lead to benefits. As such, students, corporations, and society weigh the costs and benefits of potential investments in higher education before making investment decisions. These decisions, in turn, shape how higher education is financed.

In many countries around the world, higher education is financed by a mixture of public and private sources. Through tax benefits and direct funding to institutions, and financial aid to students, public funding covers a substantial proportion of the costs of higher education. However, over the last 30 years or so, the private portion of costs in the form of tuition and fees has been on the rise in many countries, coinciding with increasing private returns of higher education. Recent decades have also witnessed the rise of corporate funding of academic research and grants for students. Using human capital as a backdrop, this article describes the economic impacts of higher education on students, corporations, and society with a focus on economic benefits rather than costs. Whereas higher education yields both economic and non-economic benefits, most researchers who have assessed the impacts of higher education have focused on more quantifiable economic benefits. The same caveat applies here.

The basic premise of human capital theory is that students and their families weigh the costs and benefits of potential investments in higher education before they make enrollment decisions. Economic costs for individuals include tuition and fees, books, room and board, as well as potential earnings foregone while attending college. Economic benefits for individuals accrue primarily in the form of earnings increases associated with higher education. While the cost side is relatively straightforward, the benefit side is much more complicated because (a) it occurs over one's lifetime; (b) the causal effect is difficult to measure; and (c) considerable variations exist across institutions, majors, and individual characteristics. These variations are discussed in this article.

Research and knowledge generation is another important function of higher education. Traditionally, research and knowledge was viewed primarily as a public good. For example, the federal government has played a dominant role in supporting research and development in the United States. Increasingly, however, research and knowledge are construed as private goods and higher education is shifting from a "public good knowledge regime" to an "academic capitalist knowledge regime" ([Slaughter and Rhoades, 2004](#), p. 28). As public funding for higher education has stagnated or declined in recent decades, corporate funding for academic research has increased globally, which has facilitated the commercialization of academic research while raising considerable concerns about academic freedom.

Higher education benefits not only individuals, but also society at large. Higher education institutions and the research they produce contribute to economic development at the societal level. College graduates pay higher taxes on higher earnings; they are also less likely to rely on government programs (e.g., health care, social security) or to be incarcerated. Because higher education is considered a public good, the level of demand for higher education by individual students in the absence of public funding would be less than optimal from a social perspective.

This article provides an overview of the economic impact of higher education on students, corporations, and society. Among these three subjects, the economic impact on students has been studied most extensively; thus variations across a range of factors including undergraduate majors, institutions, gender, and race/ethnicity are also discussed. While this article draws on empirical findings from a wide range of countries, the majority of studies included are based on data from the United States and OECD countries. Finally, education systems around the world use different terminologies (e.g., higher education, college, university, undergraduate, etc.) to describe postsecondary education; these terms are used interchangeably in this article.

Impacts on students

Human capital theory posits that the decision to pursue a postsecondary degree is based on the comparison of individual costs and benefits. This section begins with a brief overview of the immediate and deferred costs incurred by university students and their families, proceeds to an account of earnings and employment premiums for university graduates coupled with estimates of private returns to higher education, and concludes with a discussion of institutional, academic, and demographic variation in the economic impacts of postsecondary education.

Analyses of the individual economic impacts of higher education are complicated by selection issues. Compared to their peers, students who enroll in an institution of higher education may be more academically prepared, have access to more financial resources, or have greater motivation and/or aptitude in skills tied to postsecondary success, all of which may contribute to later success in the labor market. Researchers have attempted to mitigate selection issues by using quasi-experimental methods or incorporating extensive statistical controls. The best available evidence indicates that university attendance has positive economic impacts for most students, with benefits outweighing costs across a wide range of national contexts, fields of study, and student demographics. However, evidence also indicates persistent racial, ethnic, and gender disparities in the economic outcomes of university graduates.

Costs

Tuition costs for university students vary widely around the globe. In a third of OECD countries, the costs of a bachelor's degree are covered by government subsidies and other revenue sources, with no tuition for students (OECD, 2020). Costs to students in the next third of countries are relatively low, averaging less than \$2000 USD per year, while at the high end, students in Chile, the United States, and the United Kingdom pay an average of \$8000–12,000 USD annually while enrolled in bachelor's degree programs at public (or, in the UK, government-dependent private) institutions. Tuition for independent private institutions is generally much higher than for public or government-dependent private institutions—approximately three times as much in Spain, Italy, and the United States. Tuition also can vary significantly by field; for example, average costs of a bachelor's degree in education typically are lower than the costs of a health degree in a given country (OECD, 2020). Beyond tuition and fees, students incur costs for room and board, which may increase the total cost of university attendance substantially.

Meanwhile, the total direct costs of higher education may be reduced for some students through grants and scholarships; during AY 2017–2018, students at four-year (i.e., bachelor's degree-granting) colleges and universities in the US received, on average, approximately \$13,000 in scholarship and grant aid from governmental or institutional sources (de Brey et al., 2021). The availability of scholarship and grant aid varies substantially across OECD countries and across US states and institutions, with some prioritizing scholarships based on academic qualifications (e.g., secondary grades and/or standardized test scores) and others directing more funds to need-based grant aid, eligibility for which is inversely related to a student's financial resources (NASSGAP, 2020; OECD, 2020).

Students and their families may defer some costs by taking out loans, especially if the costs of university attendance are substantial relative to their financial resources. Many OECD countries offer means-tested public student loans, while some countries (e.g. Chile, Japan, and South Korea) use academic achievement to determine loan eligibility (OECD, 2020). The issue of student borrowing remains especially prominent in the United States, where nearly all US students are eligible to borrow a government student loan and over three fifths of full-time students incur debt before completing a bachelor's degree, with average debt at bachelor's degree completion increasing precipitously during the 1990s and 2000s and plateauing at just under \$30,000 in current dollars between 2012 and 2016, the most recent year for which detailed data is available (de Brey et al., 2021; Schak et al., 2020).

Borrowing can be a strategic move to alleviate credit constraints facing many students and their families, deferring the costs of university attendance to later years when a student anticipates higher earnings. In the United States, where postsecondary costs are generally high, borrowing is a necessity for many students whose families have limited financial resources. However, student debt has negative ripple effects on borrowers' financial health, limiting flexibility and shaping future decision-making and economic opportunities (Elliott and Nam, 2013). For example, higher student debt negatively affects borrowers' credit scores and rates of home ownership, a key avenue of wealth-building (Mezza et al., 2020). The financial consequences are even greater for borrowers who are unable to repay their loans and default on them, an outcome that is much more common in the United States now than in previous generations, and which disproportionately affects Black and Hispanic/Latino/Latinx students, individuals from low-income families, and those who attend for-profit universities (Hillman, 2014; Schak et al., 2020).

Benefits

For individuals, the economic benefits of higher education are primarily in the form of higher earnings and higher likelihood of employment, along with other benefits such as increased occupational prestige and better health outcomes, which are partly linked to employment-based access to insurance (Mayhew et al., 2016; McMahon, 2009; OECD, 2020). A recent report from the OECD (2020) indicates that in 2019, the average employment rate across OECD countries for 25- to 34-year-olds with a tertiary education was 85% compared to 78% for those with an upper secondary or postsecondary non-tertiary education and 61% for those without upper secondary education. Substantial advantages in the likelihood of employment have also been observed in the United States (Ma et al., 2019) and South Africa, where the employment probability premium rose substantially between the mid-1990s and

2010 (Branson and Leibbrandt, 2013). Importantly, a direct comparison of employment outcomes cannot fully capture the impact of a university education on access to employment, because one cannot assume that university graduates, *if they had not attended university*, would have the same economic trajectory as those who actually did not attend, given potential differences in unobserved characteristics. Nevertheless, these comparisons confirm that individuals with higher educational attainment have greater access to employment and financial security, on average.

In addition to a higher likelihood of employment, university students earn more than their peers, on average, across a wide range of national contexts. In Latin America and the Caribbean, workers with a postsecondary degree enjoy monthly earnings that are, on average, double the earnings of those with only a secondary degree (Urzúa, 2017), similar to the earnings advantage for both men and women in South Africa (Branson and Leibbrandt, 2013). The OECD countries with the highest earnings premiums for postsecondary education are all in Latin America, namely Brazil, Chile, Colombia, and Costa Rica. Nonetheless, premiums are substantial across OECD countries, with a 54% average earnings advantage for full-time workers with a tertiary degree versus those with an upper secondary degree, and an earnings premium of at least 20% in every OECD country except Norway (OECD, 2020). Education-based earnings premiums tend to widen with increasing age and work experience (Branson and Leibbrandt, 2013; OECD, 2020; Ma et al., 2019), and for some graduates are partly explained by the fact that they work longer hours than their peers (Zhang, 2008).

One way of assessing whether the financial benefits of higher education outweigh the costs is to estimate the age at which the cumulative earnings of an average university graduate surpass those of an average secondary school graduate. Assuming that an individual who does not enroll in higher education would enter the workforce full-time at age 18, Ma et al. (2019) calculated that both associate's degree and bachelor's degree completers in the United States reach the break-even point in their mid-30s, slightly earlier for those attending lower-cost institutions or receiving more scholarship and grant aid, and slightly later for those who take longer to complete their degree.

Another way of quantifying the economic advantage of university attendance is the private internal rate of return (IRR), i.e., the rate at which the present value of costs associated with an investment equal the present value of benefits. A review of rigorous analyses from the 2000s suggests a rate of return of approximately 12–14% for a four-year university education in the United States (Mayhew et al., 2016), and that return likely has increased over time (Long, 2010). Toutkoushian et al. (2013), whose IRR analysis incorporated the risk of dropout (a fairly common outcome in the United States and many OECD countries) and variation in the timing of both dropout and graduation, estimated an IRR of 15% for bachelor's degree graduates at private colleges and universities and 20% for graduates of (less expensive) public colleges and universities in the United States. Both high-achieving students and students whose academic performance is marginal see substantial returns, with some evidence that returns for the latter group are similar to or greater than those for the former (Brand and Xie, 2010; Oreopoulos and Petronijevic, 2013).

Notably, the economic benefits associated with higher education are not exclusive to university graduates: students who attend university but do not complete a degree tend to have better economic outcomes than those who do not attend university at all. In the United States, workers who are 25 and older with some postsecondary experience but no degree average approximately 13% more in post-tax income compared to those with only a secondary school degree (Ma et al., 2019), with evidence of especially strong returns in less credential-intensive fields such as protective services (e.g., firefighting and law enforcement) and skilled trades (Bahr, 2019). Toutkoushian et al. (2013) estimated the private IRR for students who attend a four-year college or university but do not graduate at 4.3% or 5.4% for public and private institutions, respectively. University attendance without degree completion is similarly associated with positive returns across Latin America and the Caribbean, although as in the United States, the returns within each country are substantially lower for non-completers than for degree completers (Urzúa, 2017). On the other hand, incurring student debt without the higher payoff associated with degree completion can be a major financial liability, and non-completers in the United States experience student loan default at substantially higher rates than associate's or bachelor's degree graduates (Schak et al., 2020).

Variations in benefits across students

Despite increasing costs for many students, the economic benefits of higher education for individuals outweigh the costs, on average, across a wide range of contexts. Yet these benefits do not accrue equally to all, with substantial variations across fields of study and institutions, as well as racial/ethnic and gender-based disparities. Generally, the postsecondary earnings advantage is higher in fields such as science, engineering, and business, and lower in fields such as the humanities, arts, and education (OECD, 2020). In fact, Urzúa (2017) found negative returns for some education and humanities graduates in Chile and Peru, while a review of US-based analyses from the early 2000s shows that estimates of the earnings premium associated with majoring in STEM and health sciences fields versus majoring in education range from approximately 25%–55% (Mayhew et al., 2016). An analysis of highly educated urban workers in Asia found variation in field-specific earnings premiums by city, with a significant earnings advantage for both men and women with degrees in the natural sciences (vs. the humanities) in Seoul and Shanghai, but not in Jakarta, Kuala Lumpur, Delhi/Mumbai, Hanoi/Ho Chi Minh, or Bangkok (Takeuchi, 2021). Meanwhile, an analysis of university graduates in Ireland shows that field-specific returns to a university education are greater at the lower end of the earnings distribution (Kelly et al., 2010).

Institutional differences, including funding, prestige, and/or quality (often measured in terms of selectivity, entrance exam scores, or external ratings systems; Mayhew et al., 2016; Zhang and Thomas, 2005), also shape private returns to higher education in many countries. For example, researchers using methodologies allowing for causal inference (e.g., regression discontinuity and

propensity score analyses) have found that eligibility for or attendance at an elite university increases wages by 5–10% in both China (Jia and Li, 2021) and the UK (Chevalier, 2014). In the United States, the wage returns associated with postsecondary quality have increased over time, but the relationship between quality and earnings varies considerably by race, gender, and parental education, among other factors (Long, 2010; Mayhew et al., 2016; Zhang and Thomas, 2005). Additionally, graduates of elite universities in the United States tend to work longer hours and have lower job satisfaction than their peers once earnings are controlled (Liu et al., 2010; Zhang, 2008).

In many countries, gender disparities in earnings persist across all levels of educational attainment (Ma et al., 2019; OECD, 2020; Urzúa, 2017), and are partly driven by disparities in degree attainment in fields of study associated with more lucrative employment. However, there is some variation globally in gender-based returns to higher education. Across 18 countries in Latin America and the Caribbean, the average gender gap in the return on a higher education degree is 15% in favor of men, but country-specific gaps range widely, from 92% to 108% gaps in favor of men in Mexico and Guatemala, respectively, to a 54% gap in favor of women in the Dominican Republic (Urzúa, 2017). Women experience higher returns on postsecondary degree attainment in South Africa and in the United States (Branson and Leibbrandt, 2013; Mayhew et al., 2016).

Researchers have found substantial returns to higher education for marginalized racial/ethnic groups in South Africa and the United States (Branson and Leibbrandt, 2013; Ma et al., 2019). In the United States, within-race earnings advantages for Black and Hispanic university graduates (vs. secondary school graduates) are generally similar to those for white university graduates (Ma et al., 2019). Hispanic graduates of Hispanic Serving Institutions (HSIs) have outcomes similar to their peers at comparable non-HSIs (Park et al., 2018), while Black graduates of Historically Black Colleges and Universities (HBCUs) have stronger long-term labor market outcomes than non-HBCU graduates (Price et al., 2011). In South Africa, earnings premiums for Black African university graduates are lower than for the population as a whole, but have increased at a faster rate than the population average between the mid-1990s and 2010 (Branson and Leibbrandt, 2013).

Yet educational attainment alone cannot erase deep racial inequities in economic opportunity. In South Africa, there is evidence that racial gaps in likelihood of employment among university graduates are partially driven by systematic disparities in the institutions to which students have access, including disparities in institutional capacity and resources (van Broekhuizen, 2016). In the United States, Black and Hispanic men with bachelor's degrees have similar average earnings to white men with associate's degrees, and Black and Hispanic men with associate's degrees have similar earnings to white men with only a secondary education. Likewise, at every level of educational attainment white women have higher wages than their Black and Hispanic peers, although the magnitude of the disparity is less pronounced for women in part because the within-race gender pay gap is wider for white men and women than for other racial/ethnic groups (Ma et al., 2019).

Impacts on corporations

The proliferation of higher education, especially since WWII, has led to a larger percentage of the workforce graduating with degrees from post-secondary institutions. These graduates have significant impacts on the quality, composition, and culture of the corporate workforce. Furthermore, corporations continue to partner with universities and faculty to advance research and develop new technologies. This relationship between higher education and corporations continues to develop globally as investments in higher education are associated with expectations of positive impacts on corporations.

Skilled and diverse workforce

Productivity is a principal concern for corporations, as it is a significant driver of profit generation and cost advantages over competitors. Given the increasing pace of technological innovation, many workers globally and in the United States must pursue educational opportunities to address gaps in their skills. Firms are therefore incentivized to improve the human capital of the existing workforce, which can be accomplished by hiring new graduates or training existing employees. While some of the demand for reskilling can be filled internally through corporate learning and development programs, corporations often outsource training to higher education institutions. In a recent study, Osterman (2020) found that 56.5% of employees had received formal training from their employers during the previous 12 months, although employer investments in training varied greatly across firms and employees. For example, full-time employees with more education who worked in large firms were more likely to have received formal training, whereas racial minorities including African American, Hispanic, and Asian workers were less likely to receive employer investments in training. Corporations can also rely on higher education institutions to upskill their workers. Community colleges work closely with regional corporations and establish certificate programs to provide training and workforce development (Osterman, 2020). Within emerging markets, improvements in higher education quality has been linked to improved labor markets. One study suggested that large-scale reform to the Senegalese higher education sector led to a 9% increase in employment gain for younger, high-skilled workers relative to older workers (Boccanfuso et al., 2015). It is likely that some benefits of education are passed on directly to corporations. In China, one study estimated that an additional year of schooling raised the marginal output of a given firm by 31%, while the marginal increase to worker wage is far lower (Fleisher et al., 2011). As corporations in emerging markets demand newer skills to compete globally, higher education institutions in those countries will play a critical role to ensure graduates are equipped with the necessary skills to supply the workforce. In each of these cases, higher education institutions have developed symbiotic relationships with corporations to meet the demand for skilled labor.

Given the productivity benefits of hiring talented employees, corporations compete to attract skilled graduates in a variety of ways. In some instances, recruitment methods reflect potential employees' experiences at higher education institutions. One relatively new development in recruitment is an emphasis on Corporate Social Responsibility to attract the most qualified recent graduates. These initiatives are intended to demonstrate that a company's values match the values of talented graduates, who are significantly influenced by their experiences in higher education institutions. Indeed, several studies have demonstrated an increasing global prevalence of social responsibility within higher education curricula (Vázquez et al., 2013), which in turn has encouraged the promotion of Corporate Social Responsibility in the competition for talent. As student preferences and higher education curricula evolve, corporate recruiting strategies are likely to respond.

Research has demonstrated that diverse organizations lead to greater innovation and better respond to the demands of a diverse customer base (Hunt et al., 2018). It is therefore a critical function of higher education institutions to supply the labor market with a pool of talent that is both skilled and diverse. In the United States, there is strong evidence that lower socioeconomic status is associated with a lower likelihood of degree completion (Baker et al., 2018). The result therefore is a bias toward a relatively homogenous graduate labor force derived from higher socioeconomic status, which in turn increases social inequality. As a consequence, corporations have supported admissions policies that promote diversity, arguing that education within a diverse environment better prepares potential employees for work in a global environment (Fisher, 2013). Such practices will benefit corporations by increasing the diversity and global awareness of potential employees.

In the United States, higher education institutions play an important role in supplying a diverse composition of graduates to the labor market. In Europe, institutions have emphasized diversity as it relates to a wider variety of skills and competencies (Santoalha et al., 2018). State policies have been developed to ensure institutions are sufficiently different so as to provide the labor market with a diverse skill set. One example of this practice is a focus on creative human capital. International policies have articulated the importance of strengthening the interface between economics, culture, and technology, with a particular focus on services and creative content (UNDP, 2013). Historically, most creative graduates from higher education institutions have moved into arts and culture-related professions, but more recently creative degree programs have broadened to provide critical talent to the corporate world. In the UK, universities have emerged as creative knowledge centers within regional economies that provide creative human capital to local corporations. This has led to a correlation between a city's creativity and the number of creative graduates it employs (Comunian et al., 2015). In any country, having successful corporations across a variety of sectors requires a workforce with diverse skills. Higher education plays a critical role in supplying this workforce to corporations.

Strategic relationships

Corporations also have benefitted from higher education's pursuit of new knowledge by entering into strategic alliances with institutions. Such relationships are driven by opportunities to both reduce research development costs for institutions and increase the potential of technology transfer for corporations (Elmuti et al., 2005). Through strategic alliances, corporations provide funding for research and learning opportunities for students. As corporations continue to grapple with the rapidly accelerating complexity of technology, it is often beneficial from both a cost and innovation perspective to supplement their internal research and development with the faculty expertise and research infrastructure provided by universities. As these strategic relationships continue to grow, both corporations and universities will become more dependent on each other's resources. In the case of universities, private sources of funding will exert greater influence on the type of research produced, while corporations will become more reliant on external talent for research. Effectively managing these relationships will be a critical factor for corporations to remain competitive.

The presence of higher education institutions has been shown to influence the locations of new and regional corporate offices due to "spillover" effects of knowledge production and human capital in regions surrounding universities (Acosta et al., 2011). Although much produced knowledge is widely published, closeness, proximity to the knowledge source can provide additional access to a given field of research. This access is especially important in new, complex technological fields where gaps between commercial knowledge and academic knowledge are widest. Furthermore, strategic partnerships between universities and corporations hold great potential to facilitate knowledge production. In addition to knowledge, graduates, who constitute a form of human capital, are likely to remain in closeness, proximity to a given university. These graduates not only supply talent to nearby corporations, but also can serve as a channel for disseminating knowledge. Corporations or new businesses are incentivized to establish locations in closeness, proximity to universities to absorb either the knowledge produced through strategic relationships or human capital by recruiting recent graduates.

Future research

In general, there is limited research on the quantitative economic impacts of higher education on corporations, particularly at the macro level. While there is substantial research linking diverse workforces to improved productivity and innovation, there is limited quantitative research linking the effect to improved diversity within higher education. Furthermore, given the opacity of strategic relationships between higher education institutions and corporations, there is limited research on the economic value created by such arrangements. Future lines of inquiry could attempt to quantify the return on investment for corporations or estimate the economic value of technology transfers or spillover effects.

Impacts on society

Higher education's contributions to social, economic, and cultural development are often referred to as "public goods," and serve as the primary justification for public support of institutions. Scholars point to the development of land grant colleges in the United States and more recent efforts in Europe and Africa as examples of intentional efforts to align higher education with the economic interests of nation-states (Compagnucci and Spigarelli, 2020; Hazelkorn and Gibson, 2019), although many social benefits also accrue through the production of knowledge and technology.

As investments in higher education grow, institutions around the globe are facing increasing demands for accountability to demonstrate the efficient use of funds and high quality outcomes. According to the UNESCO Institute for Statistics (2020), the average government expenditure on education in 2017 was 4.5% of GDP, with rates ranging from 1.5% to 7.9% globally. Some argue that social benefits constitute nearly half of higher education's overall benefits (Trostel, 2015); however, measuring them can be challenging.

Three main approaches have been used to measure the social benefits of higher education: the social rate of return, the national income growth model, and economic impact studies. First, the *social rate of return* (SRR) examines the discounted future earnings of graduates (before taxes) compared to the sum of the private costs and public investments in higher education. Because SRR includes both private and public expenditures on higher education, it is often estimated at a lower rate than the individual rate of return. In the United States, SRR estimates range from 11.6% to 12.1%, with remarkable stability over time (Leslie and Brinkman, 1988; McMahon, 2009). Graduates of bachelor's degree programs in the United States contribute approximately \$381,000 more in taxes than the associated government expenditures, including public subsidies and financial aid (APLU, 2016). A decennial review of global literature (Psacharopoulos and Patrinos, 2018) notes that the "social returns follow the well-known pattern of falling by level of development and level of education" (p. 453), with post-2000 social rates of return for higher education averaging 10.6% overall and 16.4% in developing countries, particularly in Africa where the social rate of return to higher education is 20–35% in some regions. Whereas the SRR is an easy way to calculate the social benefits of higher education, it fails to take into account important benefits beyond graduate salaries and therefore tends to underestimate overall social benefits.

A second common approach to quantifying the social benefits of higher education is the *national income growth* model, which involves examining the portion of national economic growth attributable to higher education based on land, labor, capital, and the "famous residual," which is the unexplained portion of growth considered attributable to knowledge and technology generated by university research. The dynamic growth model suggests that each new phase of social benefits builds upon previous social gains. Subsequently, the benefits of higher education accumulate over time, with up to a 20-year lag between investment and the realization of external benefits. Studies based on this approach have revealed that education contributes to approximately 15–20% of income growth, and about a quarter of this contribution can be attributed specifically to higher education (McMahon, 2009). More recent research specific to Central China shows education overall contributed 23–29% to the GDP depending on province from 2003 to 2014, with higher education contributing less than 5% across all six provinces (Zhu et al., 2018).

Higher education's social benefits also can be estimated in *economic impact studies*. The most common economic impact studies take the economic base approach, focused on business volume and job creation at the local level. A multiplier is extracted to reflect the amount of economic activity generated by higher education expenditures. Estimates from studies in the United States yield a multiplier of approximately 1.6 per dollar of operating budget and a job creation rate of 59 jobs per \$1 million spent (McMahon, 2009). In European countries, these studies most often take the form of input-output models with Type I multipliers indicating direct and indirect effects, and Type II multipliers indicating direct, indirect, and induced effects. Studies have yielded Type I multipliers close to 1.3 and Type II multipliers close to 3 in the United Kingdom (Hermannsson et al., 2015; Zhang et al., 2017).

Taking a different economic impact approach, a study of data from 78 countries found that an increase in the number of universities between 1950 and 2010 was positively associated with the future growth of GDP per capita, with the strongest associations at the subnational level including spillover effects from nearby regions (Valero and Van Reenen, 2016). Recognizing that universities vary in size and quality, a separate study of universities in 169 countries between 1996 and 2013 used research output as a proxy for quality and found a positive impact on economic growth in both developed and developing countries, with the strongest effects in developed countries (Solarin and Yen 2016).

In an effort to leverage anticipated economic impacts, some regions across the globe have implemented place-making (aka precinct-development) strategies using institutions of higher education as community anchors based on the triple helix model of government-industry-university. Knowledge and innovation precincts are more prominent in centrally located cities with greater density of higher education institutions such as Barcelona, Helsinki, New York, Sydney & Singapore (Pancholi et al., 2018), though recent research examines the potential and unique challenges of similar place-making attempts in less dense regions and towns such as Iceland (Yigitcanlar et al., 2017). Given the evolution of these strategies over the last two decades, much of the literature is qualitative in nature, but evidence suggests universities can play a central and significant role in knowledge based urban and regional development (Yigitcanlar et al., 2017) leading to increased attention to this strategy by government and funders alike across the globe including in South Africa where the Ford Foundation has supported university city-building efforts (Bank, 2018).

Although the aforementioned methods enable social benefits to be quantified in relation to outputs, separate lines of inquiry have explored the non-market social benefits of higher education. Some have conceptualized this difference in benefit measures as a utilitarian view, exemplified by the previous three methods, versus a focus on the broader societal contributions of higher education (Hazelkorn and Gibson, 2019; Schofer et al., 2021; Trostel, 2015). Non-market benefits related to democratization and equity highlight the social benefits of higher education at a societal level. Democratization includes fair elections, lack of

domination, the right to organize opposition parties, and similar characteristics (McMahon, 2009), and relies on the training of legal and political leaders and an educated citizenry who can turn opinions into shared action and organizational structures (Schofer et al., 2021). The critical thinking and global/cultural awareness associated with higher education further supports human rights and social movements to address issues of equality and equity (McMahon, 2009). These societal benefits can present both a challenge and benefit to industry, where human rights campaigns may protect vulnerable workers thereby slowing and/or increasing the cost of economic growth while equal opportunity and treatment may increase productivity (Schofer et al., 2021).

Aside from the traditional question of how to measure concepts such as democratization and globalization, there is also the question of how to address the interdependence and multi-directional flow between higher education and associated social benefits (McCowan, 2018). Additionally, not all individual benefits can be aggregated, and most social benefits cannot be neatly disaggregated (Schofer et al., 2021). Finally, an omnipresent question remains internationally about whether higher education is an engine of social mobility or if it simultaneously perpetuates stratification and inequality.

References

- Acosta, M., Coronado, D., Flores, E., 2011. University spillovers and new business location in high-technology sectors: Spanish evidence. *Small Bus. Econ.* 36 (3), 365–376.
- Association of Public Land-grant Universities, 2016. How Do College Graduates Benefit Society at Large? <https://www.aplu.org/PublicValues>.
- Bahr, P.R., 2019. The labor market returns to a community college education for noncompleting students. *J. High Educ.* 90 (2), 210–243.
- Baker, R., Klasik, D., Reardon, S.F., 2018. Race and stratification in college enrollment over time. *AERA Open* 4 (1), 1–28.
- Bank, L.J., 2018. Editorial: putting universities in their place. *Dev. South Afr.* 35 (5), 609–610.
- Becker, G.S., 2009. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.
- Boccanfuso, D., Larouche, A., Trandafir, M., 2015. Quality of higher education and the labor market in developing countries: evidence from an education reform in Senegal. *World Dev.* 74, 412–424. <https://doi.org/10.1016/j.worlddev.2015.05.007>.
- Brand, J.E., Xie, Y., 2010. Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *Am. Socio. Rev.* 75 (2), 273–302.
- Branson, N., Leibbrandt, M., 2013. Educational Attainment and Labour Market Outcomes in South Africa, 1994–2010. <https://doi.org/10.1787/5k4c0wbw0q-en>.
- Chevalier, A., 2014. Does higher education quality matter in the UK? In: Polacheck, S.W., Tatsiramos, K. (Eds.), *Factors Affecting Worker Well-Being: The Impact of Change in the Labor Market*, vol. 40. Emerald Group Publishing Limited, pp. 257–292.
- Compagnucci, L., Spigarelli, F., 2020. The third mission of the university: a systematic literature review on potentials and constraints. *Technol. Forecast. Soc. Change* 161.
- Comunian, R., Gilmore, A., Jacobi, S., 2015. Higher education and the creative economy: creative graduates, knowledge transfer and regional impact debates. *Geography Compass* 9 (7), 371–383.
- de Brey, C., Snyder, T.D., Zhang, A., Dillow, S.A., 2021. Digest of Education Statistics, 2019. NCES No. 2021009. National Center for Education Statistics, Institute of Education Sciences, US Department of Education. Retrieved May 6, 2021, from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021009>.
- Elliott, W., Nam, I., 2013. Is student debt jeopardizing the short-term financial health of US households? *Fed. Reserv. Bank St. Louis Rev.* 95 (5), 405–424.
- Elmuti, D., Abebe, M., Nicolosi, M., 2005. An overview of strategic alliances between universities and corporations. *J. Workplace Learn.* 17 (1/2), 115–129. <https://doi.org/10.1108/13665620510574504>.
- Fisher, v., 2013. Brief of Fortune-100 and Other Leading American Businesses as Amici Curiae in Support of Respondents, University of Texas, 570 US 297, (No. 14-981).
- Fleisher, B.M., Hu, Y., Li, H., Kim, S., 2011. Economic transition, higher education and worker productivity in China. *J. Dev. Econ.* 94 (1), 86–94. <https://doi.org/10.1016/j.jdevco.2010.01.001>.
- Hazelkorn, E., Gibson, A., 2019. Public goods & public policy: what is public good and who decides? *High Educ.* 78, 257–271.
- Hermansson, K., Lisenkova, K., McGregor, P.G., Swales, J.K., 2015. The expenditure impacts of London's higher education institutions: the role of diverse income sources. *Stud. High Educ.* 40 (9), 1641–1659.
- Hillman, N.W., 2014. College on credit: a multilevel analysis of student loan default. *Rev. High. Educ.* 37 (2), 169–195.
- Hunt, V., Prince, S., Dixon-Fyle, S., Yee, L., 2018. *Delivering through Diversity*. McKinsey & Company.
- Jia, R., Li, H., 2021. Just Above the Exam Cutoff Score: Elite College Admission and Wages in China (Working Paper No. 28450). National Bureau of Economic Research.
- Kelly, E., O'Connell, P.J., Smyth, E., 2010. The economic returns to field of study and competencies among higher education graduates in Ireland. *Econ. Educ. Rev.* 29 (4), 650–657.
- Leslie, L.L., Brinkman, P.T., 1988. *The Economic Value of Higher Education*. American Council on Education; Macmillan.
- Liu, X., Thomas, S., Zhang, L., 2010. College quality, earnings, and job satisfaction: evidence from recent college graduates. *J. Labor Res.* 31 (2), 183–201.
- Long, M.C., 2010. Changes in the returns to education and college quality. *Econ. Educ. Rev.* 29 (3), 338–347.
- Ma, J., Pender, M., Welch, M., 2019. *Education Pays 2019: The Benefits of Higher Education for Individuals and Society* (Trends in Higher Education Series, p. 44). College Board.
- Mayhew, M.J., Rockenbach, A.N., Bowman, N.A., Seifert, T.A.D., Wolniak, G.C., Pascarella, E.T., Terenzini, P.T., 2016. *How College Affects Students: 21st Century Evidence that Higher Education Works*. John Wiley & Sons, Inc.
- McCowan, T., 2018. Five perils of the impact agenda in higher education. *Lond. Rev. Educ.* 16 (2), 279–295.
- McMahon, W.W., 2009. *Higher Learning, Greater Good: The Private and Social Benefits of Higher Education*. JHU Press.
- Mezza, A., Ringo, D., Sherlund, S., Sommer, K., 2020. Student loans and homeownership. *J. Labor Econ.* 38 (1), 215–260.
- National Association of State Student Grant and Aid Programs (NASSGAP), 2020. 50th Annual Survey Report on State-Sponsored Student Financial Aid: 2018–2019 Academic Year. https://www.nassgapsurvey.com/survey_reports.aspx.
- OECD, 2020. *Education at a Glance 2020: OECD Indicators*. OECD Publishing. <https://www.oecd.org/education/education-at-a-glance/>.
- Oreopoulos, P., Petronijevic, U., 2013. Making College Worth it: A Review of Research on the Returns to Higher Education (Working Paper No. 19053). National Bureau of Economic Research.
- Osterman, P., 2020. *Skill Training for Adults* (Research Brief No. RB08-2020). MIT Work of the Future.
- Pancholi, S., Yigitcanlar, T., Guaralda, M., 2018. Attributes of successful place-making in knowledge and innovation spaces: evidence from Brisbane's Diamantina knowledge precinct. *J. Urban Des.* 23 (5), 693–711.
- Park, T.J., Flores, S.M., Ryan, C.J., 2018. Labor market returns for graduates of Hispanic-Serving Institutions. *Res. High. Educ.* 59 (1), 29–53.
- Price, G.N., Spriggs, W., Swinton, O.H., 2011. The relative returns to graduating from a Historically Black College/University: propensity score matching estimates from the National Survey of Black Americans. *Rev. Black Polit. Econ.* 38 (2), 103–130.
- Psacharopoulos, G., Patrinos, H.A., 2018. Returns to investment in education: a decennial review of global literature. *Educ. Econ.* 26 (5), 445–458.

- Santoalha, A., Biscaia, R., Teixeira, P., 2018. Higher education and its contribution to a diverse regional supply of human capital: does the binary/unitary divide matters? *High Educ.* 75 (2), 209–230.
- Schak, J.O., Wong, N., Fung, A., Ahlman, L., 2020. Student Debt and the Class of 2019 (Project on Student Debt). The Institute for College Access & Success.
- Schofer, E., Ramirez, F.O., Meyer, J.W., 2021. The societal consequences of higher education. *Sociol. Educ.* 94 (1), 1–19.
- Slaughter, S., Rhoades, G., 2004. *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. JHU Press.
- Solarin, S.A., Yen, Y.Y., 2016. A global analysis of the impact of research output on economic growth. *Scientometrics* 108, 855–874.
- Takeuchi, M., 2021. Determinants of earnings and gender earnings gaps among highly educated workers within major cities in Asian countries. *Asian Econ. J.* 35 (2), 166–186. <https://doi.org/10.1111/asej.12242>.
- Toutkoushian, R.K., Shafiq, M.N., Trivette, M.J., 2013. Accounting for risk of non-completion in private and social rates of return to higher education. *J. Educ. Finance* 39 (1), 73–95.
- Trostel, P., 2015. *It's Not Just the Money: The Benefits of College Education to Individuals and to Society*. Lumina Foundation.
- UNESCO Institute for Statistics, 2020. Government Expenditures on Education, Total (% of GDP). Retrieved May 6, 2021, from: https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS?most_recent_year_desc=false.
- United Nations Development Program (UNDP), 2013. *Widening Local Development Pathways (Creative Economy Report 2013 Special Edition)*. United Nations. <https://en.unesco.org/creativity/screen/report-widening-local-development-pathways>.
- Urzúa, S., 2017. The economic impact of higher education. In: Ferreyra, M.M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., Urzúa, S. (Eds.), *At a Crossroads: Higher Education in Latin America and the Caribbean*. The World Bank.
- Valero, A., Van Reenen, J., 2016. The Economic Impact of Universities: Evidence from across the Globe. National Bureau of Economic Research. Working Paper 22501.
- van Broekhuizen, H., 2016. Graduate Unemployment and Higher Education Institutions in South Africa (Stellenbosch Economic Working Papers). Stellenbosch University.
- Vázquez, J.L., Lanero, A., Licandro, O., 2013. Corporate social responsibility and higher education: Uruguay university students' perceptions. *Econ. Sociol.* 6 (2), 145.
- Yigitcanlar, T., Edvardsson, I.R., Johannesson, H., Kamruzzaman, M., Ioppolo, G., Pancholi, S., 2017. Knowledge-based development dynamics in less favoured regions: insights from Australian & Icelandic university towns. *Eur. Plann. Stud.* 25 (12), 2272–2292.
- Zhang, L., Thomas, S.L., 2005. Investments in human capital: sources of variation in the return to college quality. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. 20. Springer-Verlag, pp. 241–306.
- Zhang, Q., Larkin, C., Lucey, B.M., 2017. The economic impact of higher education institutions in Ireland: evidence from disaggregated input–output tables. *Stud. High Educ.* 42 (9), 1601–1623.
- Zhang, L., 2008. The way to wealth and the way to leisure: the impact of college education on graduates' earnings and hours of work. *Res. High. Educ.* 49 (3), 199–213.
- Zhu, T., Peng, H., Zhang, Y., 2018. The influence of higher education development on economic growth: evidence from Central China. *High Educ. Pol.* 31, 139–157.

Academic capitalism and market thinking in higher education

Oliver Wieczorek^a and Richard Muench^{a,b}, ^a Zeppelin University, Friedrichshafen, Germany; and ^b University of Bamberg, Bamberg, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	37
Macro level: the superimposition of institutional competition for the accumulation of economic and symbolic capital on individual competition for progress of knowledge and recognition	37
The entrepreneurial university	38
Oligopolistic structures of knowledge production	38
Oligopolistic structures of knowledge transfer	39
Meso level: the superimposition of the managerial control of research on its control by peer review	40
Limited scope for creative deviance	40
Consolidating established lines of research and nimble knowledge production	41
Micro level: the superimposition of compliance strategies for academic survival on the curious search for progress in knowledge	41
The bias against novelty	42
Looking for funding	42
The split between elite and rank-and-file habitus: self-conscious autonomy vs. anxious adaptation to the mainstream	43
The counterproductive side of academic meritocracy	44
Conclusion and outlook	44
References	44

Introduction

The sector of higher education and research has undergone a fundamental transformation over the past three decades. Academic capitalism and market thinking have become central features of the sector. Emerging in the wake of New Public Management (NPM) in the anglophone world of higher education they have spread all around the world to build “varieties” of academic capitalism and market thinking (cf. Schulze-Cleven and Olson, 2017). “Academic capitalism” means understanding universities as enterprises which compete for the ongoing accumulation of economic and symbolic capital. Economic capital comprises research funding, tuition, sponsors, scholars, and talented students. Symbolic capital includes prizes, honors, memberships in prestigious associations, citations, generally: prestige. At the macro level, this competition overrides competition for the best possible education and research. Without a competitive order, this competition tends to lead to the formation of oligopolies and an increasing gap between elite and rank-and-file universities in higher education and research on the macro level of the academic field. Elite and rank-and-file universities are therefore differently affected by academic capitalism. Therefore, they differ in their strategies of coping with competition on the macro, meso and micro levels. While elite universities dispose of great autonomy in defining their mission on the macro level and how to implement it on the meso and micro levels, non-elite universities are much more compelled to meet the prevailing expectations on the macro level and to turn them into the closely controlling management of research and teaching on the meso level and socialization of junior scientists on the micro level. In this chapter, we outline the state of research on the three levels of academic structural change in the context of spreading academic capitalism and market thinking as well as the different strategies of elite and non-elite universities to cope with increased competition for the accumulation of economic and symbolic capital. The focus is on the question of what effect this structural change has on the promotion of the best possible education and particularly knowledge advancement. We primarily deal with research because it is at the center of academic capitalism and therefore the primary topic of the respective literature. Nevertheless, teaching is part of this new feature of higher education institutions. With regard to research, universities are in permanent competition for the accumulation of economic and symbolic capital through raising funds, recruiting staff, commercializing findings through patents and disseminating expert knowledge. With regard to teaching, they compete for earning tuition fees, enrolling the brightest students, attracting foreign students as “cash cows” (Cantwell, 2015a), granting certificates for highest possible tuition fees, employing cheap part-time contingent faculty doing the job of undergraduate instruction, and placing graduates on the labor market.

Macro level: the superimposition of institutional competition for the accumulation of economic and symbolic capital on individual competition for progress of knowledge and recognition

Academic capitalism and market thinking at the macro level of the academic field turn universities into enterprises. Consequently, they superimpose competition between universities for the accumulation of economic and symbolic capital on competition

between scholars for best possible teaching and knowledge growth. With the accumulation of competitive advantages based on previous competitive advantages, this competition results in oligopolistic structures of knowledge production and knowledge transfer, privileging elite universities and disadvantaging non-elite universities.

The entrepreneurial university

Owen-Smith (2003) has observed increasing linkage of academic and commercial achievement at the most research-intensive US universities. In the knowledge economy, universities play a central role in generating economic growth. They are the dynamic force in the innovation network of science, business, and politics (cf. Jessop, 2017). According to Gibbons et al. (1994), this is mode 2 of knowledge production, where basic and applied research are not separated as in mode 1 but are closely intertwined. Etzkowitz (2003) and Etzkowitz and Zhou (2017) have coined this new mode of knowledge production the “triple helix of science, business, and innovation policy” (cf. Leydesdorff, 2012, 2018). According to Berman (2012), universities have become “economic engines” in this context by being able to retain ownership of patents generated with federal funds under the US Congress’ Bayh-Dole Act of 1980 (cf. Holloway and Herder, 2019; Shibayana et al., 2012). Strategic collaborations between universities and companies play a significant role. One basis for such collaborations is interlocking board memberships (Slaughter et al., 2014). Slaughter and Leslie (1997) and Slaughter and Rhoades (2004) have coined the term “academic capitalism” to describe this arrangement. According to Slaughter and Rhoades (2004: 1–34), academic capitalism manifests itself (1) in a close intertwining of knowledge generation with the production of economic goods, as seen in universities’ efforts to generate patents and copyrights of teaching materials; (2) in interstitial organizations such as technology transfer offices or intellectual copyright offices which directly link the university to the business world; (3) in the emergence of intermediate organizations such as the Association of University Technology Managers (AUTM) or the Business Higher Education Forum in the United States which mediate between academia and industry/business; (4) in the development of universities’ managerial capacity and internal control capabilities over their business processes that increase their strategic capabilities to position themselves in the competition for funding, acknowledgment, visibility, and market power.

In this process, universities are turned into businesses that aim at increasing their resources and translate them into economic as well as symbolic profits. Resources and reputation are multiplied in a spiraling accumulation process (Cantwell, 2016). Competition between universities for the accumulation of economic and symbolic capital increasingly overlaps with competition between researchers for the advancement of knowledge and recognition granted for contributions to knowledge growth by the scientific community (cf. Musselin, 2018; Pederzini, 2018). While in the latter competition the norms of collegiality and disinterestedness, as well as intellectual communism, as defined by Merton (1973 [1942]), ensure that the successful researchers do not completely crowd out the less successful ones, in academic capitalism there are no such barriers and there is cut-throat competition between universities aimed at recruiting outstanding scholars and the best students. There is a brain gain at the top and a brain drain across the board (cf. Khan, 2021). According to the Matthew effect (Merton, 1968), the winners enter an upward spiral by converting existing competitive advantage into further competitive advantage, while the losers continue to lose resources and reputation (cf. Cantwell and Marginson, 2018). This is why Clark (1998) was able to observe, as a matter of course, that universities which adopted an entrepreneurial strategy early on excelled in accumulating economic and symbolic capital compared to other universities (cf. Cantwell, 2016; Geiger, 2004; Rikap and Harari-Kermadec, 2020). Competitive advantages are concentrated at the top. Therefore, copying success models—for example investments in technology transfer centers—by rank-and-file universities often turns out to be a flop because the necessary resources in terms of economic and symbolic capital and a potent industrial environment in the neighborhood are lacking.

The concentration of resources on a small circle of top scientists also involves a growing number of people doing research and teaching under the direction of these “capital owners” and/or replacing them in teaching by contingent faculty on temporary contract. The result is a rising number of researchers and teachers who constitute a new academic proletariat. They are exploited to help produce all those merits which are appropriated by the capital owners (cf. McCarthy et al., 2017; Reay, 2014; Wheaton, 2020).

Protagonists of market thinking in higher education mostly think of ideal markets where equal conditions apply to every participant in the competition. This is, however, not the reality of higher-education markets. The market is widely shaped by inequalities of market power. To address this distorted nature of markets in higher education, a conflict-theoretical analysis of academic capitalism is called for. Most appropriate is drawing on Bourdieu’s theory of the academic field, academic habitus formation, academic capital accumulation and the embedding of the academic field in the field of power (cf. Bourdieu, 1975, 1988, 1996; Collyer, 2015; Mendoza et al., 2012; Münch, 2014; Beyer, 2021; Wieczorek et al., 2017; Wieczorek, 2021).

Oligopolistic structures of knowledge production

The result of the cutthroat competition unleashed by academic capitalism throughout the academic field are oligopoly structures that steer the knowledge process in the direction set by the successful elite universities and the so-called high impact journals that are closely intertwined with them through editors and authors (Hodgson and Rothman, 1999). The academic field is differentiated worldwide into a center occupied significantly by the US elite universities, a semi-periphery of older European and younger East Asian universities, and a periphery of universities in emerging and developing countries (cf. Beyer, 2021; Münch, 2014, pp. 13–37; Münch and Schäfer, 2014; Paasi, 2015). One consequence of the oligopolistic structure of the academic field is overinvesting

attention, recognition, and resources at the top while underinvesting across the board. The academically rich get richer, while the academically more modestly endowed stagnate in their resources or become even poorer (cf. [Mendoza et al., 2012](#)). An inverted u-shaped curvilinear relationship emerges between the investment of research funds and returns, measured in terms of publications and the progress of knowledge reported therein. That is, beyond an optimal mean endowment of research funds, a decreasing marginal utility of each additional amount of money invested is observed. Beyond the optimum, coordination expenses and administrative tasks grow to such an extent that they divert too much attention from actual research ([Münch, 2014](#), pp. 223–233; [Münch and Baier, 2012](#)).

In a purely economic sense, this is an inefficient allocation of research funds. The oligopolistically structured field allows members of the oligopoly to extract monopoly rents from their position. Monopoly rents are extra revenues in terms of economic and symbolic capital, generated through a monopoly position, in excess to the revenues which would be possible under perfect market competition (cf. [Welsh, 2021](#)). Two mechanisms of elite formation identified by [Burris \(2004\)](#) following [Weber \(1976, pp. 534–539\)](#) come into play here. The elite recruits its next generation from within its own circles through appointment networks, and it is wealthy enough to establish lavishly endowed research centers that have global appeal. Therefore, the academic field does not function according to the logic of economic efficiency, but according to the logic of academic distinction. And distinction has always included the privilege of the nobility, namely, to be able to display luxuries that make potential competitors green with envy.

If we consider the embedding of the academic field in the field of power (cf. [Bourdieu, 1996](#)), in which the representatives of the various fields, classes and groups struggle to shape society, we see here an extension of the oligopolistic structure of the academic field into the field of power and, via this field, into political consultation and influence on political decision-making processes. A prominent position is occupied by the US elite universities, not only in the field of power and in political consultation in the United States, but in the influence on the dissemination of scientific knowledge worldwide ([Wieczorek, 2021](#)). That a small group of extremely wealthy universities dominates the generation and implementation of scientific knowledge is the academic side of a kind of plutocracy that has emerged in the United States and shapes the development of science and the dissemination of scientific knowledge worldwide (cf. [Marginson, 2016](#)). Foundations play a leading role in this process, no longer practicing arbitrary philanthropy with their funds, but pursuing a global agenda of shaping the world, as exemplified by the Gates Foundation and its health as well as education agenda (cf. [Amarante, 2018](#); [McGoey, 2012](#)). The oligopolistic symbolic domination of knowledge and expertise by US elite universities extends far beyond the United States and applies to the entire world. International university rankings consecrate this oligopoly of knowledge generation and dissemination and cement its global symbolic domination ([Espeland and Sauder, 2007](#); [Erkkilä and Piironen, 2018](#); [Hazelkorn, 2015](#); [Kauppi, 2018](#)). Only a few universities outside the US—such as University of Oxford, University of Cambridge, University College London, ETH Zurich—still manage to claim some share of this symbolic domination. Recognition of European scientists, for example, depends on their international collaboration with colleagues from the center of global science. In this way they share somewhat in the center's symbolic capital. There is an increasing split between highly paid cosmopolitans and poorly paid locals (cf. [Kwiek, 2019](#)). Strictly speaking, a “cosmopolitan” outlook does, however, not mean “cosmopolitan” in the comprehensive sense but rather “closely linked to the center of global science.” The great mass of universities all over the world, like the mid- and lower-ranking US universities, play only a subordinate role in global knowledge production and dissemination. This constellation in the relationship between knowledge and power results in a narrowing down of knowledge, expertise, and political decision-making space. The possibilities of humankind to recognize and overcome challenges are thus restricted to a narrow spectrum of conceivable alternatives.

Oligopolistic structures of knowledge transfer

The example of the interconnectedness of pharmaceutical corporations, universities, and startups in biotechnology provides a good illustration of the oligopolistic structure of academic knowledge production and its close interconnection with business and politics. It is also an example of university-industry collaboration which transcends national borders and involves universities and corporations which operate on a transnational level so that we can speak of transnational academic capitalism ([Kauppinen, 2012, 2015](#)). A sharp controversy was triggered in 1998 by the \$25 million grant to the Department of Plant and Microbial Biology at UC Berkeley by the Swiss pharmaceutical company Novartis. In return, Novartis was allowed to send two representatives to the five-member research committee and to exploit all of the department's research findings, including those that emerge from research funded by public money such as the university's regular state budget and federal funds raised. Critics saw two problems in particular: first, the direct influence of a private company on public research, and second, the private exploitation of research funded with public money (cf. [Washburn, 2005](#); [Rudy and Coppin, 2007](#)).

The biotechnology cluster surrounding Harvard University and MIT in Cambridge/Boston, Massachusetts, is of outstanding importance in biotechnology research ([Birkinshaw et al., 2018](#); [Fochler, 2016](#); [Owen-Smith and Powell, 2004](#); [Stevens and Newman, 2019](#)). Here, there is an oligopoly structure in scientific knowledge production concentrated in a few universities directly reaching out to the world of business. This oligopoly structure is supported by a homologous oligopoly structure in the publication system. [Young et al. \(2008\)](#) demonstrate for biomedicine that the enforcement of the journal impact factor (JIF) as a quality standard for scientific publications, enforced by rankings, has created a winner take all market (cf. [Frank and Cook, 2010](#)). Only a tiny fraction of the research findings available per se are noticed and attract all attention. As with art auctions, this results in a kind of “winner's curse” for recipients. They invest so much attention and capital in processing the articles of high-impact journals that little is left for further, widely differentiated research with possibly surprising results. The knowledge value of these articles does not correspond to the high impact factor of the journals in the sense that many other publications question their results. This hinders the

progress of knowledge. Ioannidis (2005) has drawn attention to the fact that most published research findings, even those published in high impact journals, are not confirmed by further research. Publications in these journals paint a glossy picture of the research situation because their extremely high rejection rate implies that only highly significant and perfectly clean results are published. However, this does not correspond at all to the research reality which presents itself as much less clear and clean. Moreover, the oligopoly structure of the publication market has the effect of narrowing research to a few broadly investigated topics. Therefore, there is a lack of diversity of topics from which only new knowledge can emerge.

Finally, the intertwining of the oligopoly structure of academic research and publication outlets with the oligopoly of the world's leading pharmaceutical corporations means a further narrowing of the translation of research into technology. This also affects the approval of drugs for distribution in the drug market. In the United States, the Food and Drug Agency (FDA) is responsible for approving the dissemination of drugs. This is where the nexus of research, industry, bureaucracy, and politics is concentrated to ultimately distribute scientifically generated, politically desired, bureaucratically controlled, and economically profitable products on the market (cf. Carpenter, 2004; Hawthorne, 2010). The concentration of the relevant field-specific capitals among a few actors means that the tension between scientific truth-seeking, political goal setting, bureaucratic control, and economic profitability tends to be resolved in favor of economic profitability. Hearings often bring together the same experts and stakeholders who are in any case mutually committed to each other in close cooperation and form the center of power (Wieczorek, 2021). This significantly narrows the spectrum of alternative voices. The search for truth is then no longer open enough to sufficiently explore all possible risks and side effects of scientifically generated products. The consequence of the formation of oligopolies of research in the context of academic capitalism is accordingly a narrowing down of scientific and technological development.

Meso level: the superimposition of the managerial control of research on its control by peer review

At the meso level of the organization of universities, academic capitalism means that universities engaged in intensified competition for the accumulation of economic and symbolic capital rely on a strong, strategically operating management. For the purpose of best possible positioning in competition, this management internally exercises strong control over research and teaching by profiling it through research clusters and by regularly monitoring its research output and its teaching. There is a shift of power from the faculty to the administrative level. Ginsberg (2011), therefore, speaks of "the rise of the all-administrative university" (cf. Delucchi et al., 2021; Donoghue, 2008/2018; Engelen et al., 2014; McClure, 2016; Lorenz, 2012; Readings, 1996; Tuchman, 2009; Welsh, 2020). In this context, orientation to prevailing research themes, paradigms and methodological standards gains the upper hand over the free pursuit of research goals based solely on curiosity in the process of discovery, because this is the only way to avoid failure and ensure reception of one's research (cf. Hallonsten, 2021). This type of coping with competition affects middle-class universities particularly, which are struggling for recognition with scarce resources and cannot afford failure resulting from risky investments. Elite universities are much more able to invest in risky research aiming at novel findings because of their economic and symbolic wealth. At the very top, a culture of academic luxury, excellence and autonomy is celebrated while among the rank and file a culture of ubiquitous accountability and pressure to conform to established standards prevails (Gonzales et al., 2014; Mendoza et al., 2012). Elite universities can be proud of simply selecting the "best" and then let them do what they want to, the rank-and-file universities, in contrast, cannot recruit the "best" and feel more compelled to monitor in detail what their faculty is doing. They are strongly subject to the audit culture of New Public Management (cf. Power, 1997). As on the macro level, elite and non-elite universities are very differently affected by academic capitalism on the meso level.

Limited scope for creative deviance

Speaking with Goffman (1961), the monitoring tools, the visualization of every step taken in the research process, and ubiquitous accountability only allow for research autonomy in an unobserved underlife of the institution that leaves researchers the space they need to generate innovations. The latter require a high risk/high reward strategy that cannot be chosen when minimizing risk dominates research.

Studies from management and organizational research show that the possibility to deviate from organizational goals and directives by superiors promotes innovativeness and thus contributes to the competitiveness of organizations (Tenzer and Yang, 2019). This behavior is referred to as "creative deviance" (Mainemelis, 2010), i.e., the deliberate deviation from directives to pursue creative and new ideas on one's own authority. It needs free spaces in which these ideas can mature over a considerable period of time and also autonomy in allocating resources to projects. Within companies, this kind of project work requires a management that grants sufficient freedom and does not punish creative deviance, but rather encourages it, in order to address problems, for which no solutions have yet been found (Lin et al., 2016). Therefore, it is not surprising that radical innovation is listed as a positive effect of such behavior, while wasting resources and violating stakeholder trust are cited as negative effects.

For example, Lin et al. (2016) used 233 employee-management dyads to examine how managers' response to creative deviance in organizations is related to employee performance. They emphasize that employees' performance and contribution to problem solving increase if their independent creative work is recognized, whereas punishment and manipulative behavior management by supervisors hinder employees' performance. Moreover, Liu and Zhou (2020) show that research staff who feel privileged and "different" or "unique" are more likely to pursue their own creative ideas. They are also more likely to do so in the face of resistance

from superiors and other stakeholders to whom they are accountable. However, if they are granted the necessary freedom, they respond with a higher affective commitment to the organization and perform better.

As researchers are placed in multiple dependencies in the wake of New Public Management reforms, the space for creative deviance is increasingly limited. Specifically, this freedom is diminished by (1) the increasing obligation to document one's own activities, (2) by third-party funding competition and dependence on funders, (3) by monitoring through university managers, and (4) by visibility competitions expressed in rankings and citation indices. In this way, innovation is hindered because of lacking space for exploring new terrain (cf. [Anderson et al., 2014](#)).

Thus, if we apply the findings from creative deviance research to the academic field, we expect to see significant differences in the opportunity to pursue creative ideas between elite universities and the broad mass of middle and lower ranking universities. First, this is because researchers at elite universities are trusted more and are granted greater latitude for deviant behavior compared to researchers at non-elite universities, simply because of their institution's reputation. It is precisely these researchers who ensure that universities do well in rankings, accumulate scientific awards, attract external funding, and achieve high citation numbers. This in turn legitimizes the freedom that researchers has been granted. Second, elite universities dispose on the necessary resources to enable researchers to explore new terrain with uncertain outcome so that resources may be invested in vain. In contrast, researchers at non-elite universities must go the safe way to bring about visible results regularly. This can be seen both in the inverted-u-shaped relationship between resources deployed and publication output at elite universities, and in the conformity of topic choices to the mainstream of researchers at non-elite universities ([Wieczorek et al., 2017](#)). Thus, universities that dispose of sufficient resources of their own, exert less managerial control and respect researcher autonomy have a higher potential for creative deviance and for pursuing new lines of research. They allow a "high risk, high reward" strategy in the advancement of scientific knowledge (cf. [Fleck, 2013](#); [Uzzi et al., 2013](#); [Wang et al., 2017](#)).

Consolidating established lines of research and nimble knowledge production

In contrast to exploring new ground under conditions of autonomy, limited scope for deviance in research because of increased financialized managerial control in academic capitalism results in what can be called consolidation of established lines of research and nimble knowledge production (cf. [Engelen et al., 2014](#); [Hoffman, 2020](#)).

A first feature of favoring research along established lines by strong university management is the management's interest in large-scale funding through establishing interdisciplinary research clusters providing a widely visible research profile. This implies increasing size of collaborating teams. Research shows, however, that it is small teams of two to eight members, all actively engaged in research, with secure funding and little administrative work which are crucial for establishing new lines of research and achieving breakthroughs, while large-scale clusters are extending and consolidating the established lines of research ([Heinze et al., 2009](#); [Holingsworth, 2006](#); [Wu et al., 2019](#)).

A second feature of increased managerial control on research is favoring flexible adaptation to changes in funding opportunities. Research teams at sites that suffer from a permanent lack of funding or recognition by peers, or work on topics that are marginal to the respective discipline, are particularly affected by this economic logic. As [Hoffman \(2020\)](#) demonstrates in a field study on a computer science laboratory, the principals of such units adopt a strategy of "nimble knowledge production." This strategy involves (1) quickly finding provisional solutions to avoid unnecessary costs; (2) excluding complex theoretical problems in favor of "real world problems" outside of academia; and (3) quickly targeting new topics if hurdles or problems arise that run counter to the commercialization of findings. This focus may be applied by researchers out of themselves, but it may also be a result of governments intervening in the funding landscape. This strategy was observed, for example, with renewable energy research in the United States under the Trump administration, where researchers used previous findings and approaches to research to commercialize highly profitable fuels ([Jeon, 2019](#)). In such cases—emanating from researchers themselves and triggered by pressure from external actors—a purposeful ignorance of unprofitable lines of research is cultivated. In addition, there emerges ...

a fleeting attention to the long-term analytic, conceptual, and methodological problems of an academic field of research. Where this form takes hold, then, we might worry that sustained engagements with the thorniest of scientific problems can wither away as the champions of academic capitalism tell stories about how quick and clever they have become.

[Hoffman \(2020, p. 30\)](#).

Micro level: the superimposition of compliance strategies for academic survival on the curious search for progress in knowledge

At the micro level of researchers, academic capitalism translated into comprehensive monitoring of research and its achievements has the effect that researchers learn to aim at easily attainable research goals and topics to avoid failure and to get on a safe career path already in the process of their academic socialization. This calls for a conformist and adaptive habitus (cf. [Morris, 2020](#)). These are the disciplining effects of closely monitored research and teaching fueled by the omnipresence of rankings (cf. [Sauder and Espeland, 2009](#)). A difference between elite and mass is also noticeable at this level. Doctoral students at elite universities, on the one hand, have attained so much self-confidence and self-assurance, as a result of being chosen through a rigorous selection process, that they are confident to be able to break new ground in their research. They must do so if they want to have a chance of

being appointed at an elite department after having graduated. Of course, only a minority of them will make it in the end; the rest will have to make do with an appointment at a middle- or lower-ranking department. For the doctoral students at the middle- and lower-ranking departments, on the other hand, it means taking the safe path of standardized mainstream research right from the start so that their career continues on the middle-respectively lower-ranking level.

The bias against novelty

Certainly, the vast majority of about 80% of the researchers have always been engaged in mainstream research of normal science in the sense of Kuhn (1962), while it is left to about 20% to venture into new topics and methods. Even before the spread of academic capitalism, there was a “bias against novelty” in science (Wang et al., 2017), which can be explained by the academic community’s critical attitude toward uncertain knowledge not generated according to the established methodological standards. What is new about the translation of academic capitalism into comprehensive monitoring of the research process is the separation between research that is free in this respect at elite departments and controlled research in the mainstream becoming so entrenched that the broad mass of researchers at mid- and lower-ranked universities are excluded once and for all from exploring uncharted territory. This, in turn, has the effect of shrinking the potential for innovation overall, because only what sees the light of day at elite departments and only what is published in high-impact journals closely intertwined with those departments gets attention.

Incremental innovation and normal science are rewarded, and only carefully crafted articles with significant results are published in high-impact journals. This is the starting point for knowledge generation. Research on so-called sleeping beauties (van Raan, 2004) suggests that many scientific innovations are not even noticed and recognized by a broad scholarly audience. The rationale for this phenomenon is that innovative research with high novelty value is often initially published only in marginal journals, has little connectivity, and takes many years (if not decades) to be recognized (Dey et al., 2017; Yeung and Ho, 2018). To get them to awaken, they need to spark conventional follow-up research, which is more likely to be noticed the higher the rank of the journals in which the findings of this follow-up research are published and the higher the reputation of the researchers conducting follow-up research (Chai and Menon, 2019; van Raan, 2015). Scholarly awards confer greater legitimacy on both new lines of research and the scholars who help to further advance those lines of research and bring them into the scientific mainstream (Farys and Wolbring, 2021).

Looking for funding

Although autonomy and long-term funding have been variously proved to promote creativity, innovativeness, and research breakthroughs (Heinze et al., 2009, 2020; Whitley et al., 2018), nevertheless, the competition for external funding with relatively short funding periods—as a main feature of academic capitalism—has become ever-present since the 1980s. Looking for funding is everyday work of principal investigators (cf. Cantwell, 2015b). Attached to the increased competitive allocation of funds is the promise that the best scientists will be promoted, innovation will be favored, and competitive advantages of scientific locations will be generated. In fact, having to look permanently for external funding allocated through peer review according to established standards interferes with the opportunities of scholars to establish new lines of research and drive radical innovation. Greater dependence on external funding again subjects scholars at rank-and-file departments more to this effect of disciplining and normalization than their colleagues at elite departments. However, as Feyerabend (1970/1993) in particular has emphasized, the breakthrough to new knowledge requires exactly deviation from established methodological standards. This is all the easier to do the less dependent researchers are on the granting of research funds based on peer review. This may be the reason why Evans (2010, p. 444) observed more opportunities for industry-funded than government funded research to produce new insights. In contrast, as he also found, government funded research allows more penetrating exploration of issues in all their breadth and depth. Presumably, particularly distinguished researchers receive industry funding without strict peer review, so they are freer to decide on theory, methods, and subject matter than is the case with the usual publicly funded research that is subject to strict peer review. The latter, in turn, obviously appreciates methodologically standardized research within the framework of normal science and the incremental innovations achievable within this framework.

Various studies generally indicate an inverted u-shaped relationship between the degree of novelty of research and the likelihood of funding (Bourdeau et al., 2016; Gallo et al., 2016, 2018). The rationale for this finding is that the added value of the research is not recognized, but that too much proximity to previously addressed topics leads either to expert panels viewing applicants as competitors or their contribution as uninteresting. While in the former case the innovation opportunity tends to go unrecognized and the connectivity to previous research is not given, applications that are not very innovative are also not worthy of funding because they are considered trivial and not very insightful.

For example, looking at 150 research proposals in the field of biomedicine, Boudreau et al. (2016) observe that proposals with high and low novelty value are evaluated as not eligible and rejected. Project success, as measured by publication numbers, scientific impact, and societal impact, is considered as uncertain. This uncertainty, and ergo the risk of failing to meet the funding targets, leads to a lower scientific quality being attributed to the respective proposals. Gallo et al. (2016) highlight for the case of biomedicine that applicant seniority, reviewer’s expertise in a given field of research, and a priori knowledge on the applications increase the chances of a project to get funded. Additionally, they uncover a commensuration bias insofar, as reviewers overtly focus on methodological weaknesses of research proposals (cf. Lee, 2015). At the same time, there is little agreement on which applications are innovative. This is why reviews tend to be based on easily measurable rejection criteria. Furthermore, the findings of Li (2017) and

Howell (2017) suggest that an applicant's familiarity with expert panels implies a higher chance of approval. They most likely share common standards.

However, this feature of peer review is not limited to scientific bodies, but at the same time can also be observed in R&D departments in knowledge-intensive companies. Here, the study by Criscuolo et al. (2017) is instructive. Studying funding panels in an international engineering company, these authors suggest that there is not only an inverse u-shaped relationship between novelty of proposals and probability of funding (and amount of funding), but also that this curvilinear effect is amplified by the workload of expert panels. For example, reviewers who are overloaded with applications tend to be far more likely to fund projects with low novelty value than reviewers who have fewer applications to process. Thus, overcharging scholars with doing reviews—a feature of the monitoring regime of academic capitalism—works against funding research of high novelty value. At the same time, there is an interaction effect of the affiliation of reviewer and project applicant to an organizational unit. While shared affiliation has a positive effect on approval, it has negative consequences for approval of innovative projects. This may be due to the reviewers' desire to prevent innovations from within their own department that make their work look less up to date.

Taking the example of Japanese natural sciences, Wang et al. (2018) found that seniority and status effects are associated with the chance to acquire external funding and to perform innovative research. Using a combination of bibliometric analysis and standardized survey, they show that third-party funding for innovative projects is allocated to researchers who are employed at high-ranking universities, tend to be male, and already hold professorial titles. In addition, the age and networking of applicants, as measured by the number of co-authors per paper, are positively associated with funding for innovative research. Finally, Gallo et al. (2018) observe that there are large differences between reviewers and project applicants in their view of the rationale for accepting or rejecting research proposals. Based on a survey of reviewers and applicants in the life sciences, the authors find that a large proportion of reviewers attached great importance to innovativeness and risk in terms of potential failure of the project, as well as to the applicants' publication record, when evaluating projects. In contrast, only a minority of applicants reported that these three issues were addressed in the reviews. According to their reading of the reviews, there rather was criticism of the choice of methods and methodology, which is more concrete and thus easier to state than is the case with the innovative potential of research. It is simply less risky for reviewers to recommend funding based on methodological rigor than based on innovative potential.

Summarizing these findings, the peer review process in the allocation of external funding under the regime of academic capitalism rewards methodologically rigorous, extra-scientifically useful and thus at the same time "accountable" normal science. Consequently, this tends to hinder radical innovation (cf. Heinze and Münch, 2016).

The split between elite and rank-and-file habitus: self-conscious autonomy vs. anxious adaptation to the mainstream

The socialization of young scholars into the academic community represents one of the core tasks of universities (Parsons and Platt, 1973/2013, pp. 103–162). This socialization process involves both the internalization of academic values and norms as well as situationally appropriate action sets and objectives (e.g., research for the sake of research vs. research to obtain external funding). Related to this feature is the academic habitus as a set of interrelated dispositions that enable researchers to navigate the various academic contexts and situations and perform the tasks associated with the role of scholar, researcher, and teacher. Ideally, researchers are guided in this process by the academic norms of intellectual communism, universalism, disinterestedness, and organized skepticism articulated by Merton (1973 [1942]). However, the changes induced by academic capitalism do not stop at the internalization and interpretation of academic norms during the formation of an academic habitus and thus researcher personality. This concerns the formation of (1) selective ignorance of ineligible topics and, based on this, the equalization of relevant and eligible topics, (2) the increasing relevance of performance indicators, (3) the interpretation of research autonomy, and (4) the definition of research as a means to the end of gaining external funding and extra-academic attention.

Doctoral students increasingly learn during their training that it is strategically wise and a virtue to neglect research topics that do not attract government (Jeon, 2019) or industry funding (Hoffman, 2020). On the contrary, they learn that purposefully ignoring non-fashion topics is beneficial to their careers and, at the organizational level, helps sustain research units that must fund themselves primarily through external grants, and to conduct basic and applied research simultaneously (Rushforth et al., 2019). The research approaches that prevail in the competition for external funding in the context of project-based calls are associated with good research by junior scientists. As a result, competition for knowledge advancement is overshadowed by competition for external funding.

The capacity for selective ignorance and rapid redirection, coupled with the pressure to accumulate funds and permanently demonstrate one's performance, has at the same time implications for the understanding of research autonomy. The pressure to pursue research topics specified by third-party funders is particularly high among early career researchers at mid-level universities, where the perceived and formally assured autonomy of researchers is also seen as high (Wieczorek et al., 2017; Woelert et al., 2020). This is because the need to (re)align with topics specified by companies or ministries is increasingly judged to be a normal state and thus part of the belief system of the academic field. Such a normal state cannot reduce the perceived autonomy for those researchers who have already proven themselves and have internalized the competition for third-party funding (cf. Mendoza, 2007). Scholars learn how to accommodate their commitment to authenticity in doing research with the demands of being successful in the field (cf. Brunila and Valero, 2018; Cannizzo, 2018; Hall, 2018; Shields and Watermeyer, 2020).

Eventually, a most problematic effect of the omnipresent pressure to demonstrate performance measured in "international" publications is the growing market of so-called predatory publishing. Scholars in the periphery lack access to the oligopoly of respected academic publishing dominated by a small group of global publishing houses. To fulfill the expectation of publishing

in international journals they feel compelled to publish in one of the innumerable low standard journals which because of their English language are included in the duopoly of the bibliometric platforms Web of Science and Scopus in contrast to many journals publishing in languages other than English (Stöckelová and Vostal, 2017).

The counterproductive side of academic meritocracy

One could argue that the very unequal distribution of autonomy and resources strengthened by academic capitalism is most efficient, because in this case autonomy and resources are not allocated in an indiscriminate manner but according to creative and productive potential. Those scholars dispose of most freedom and resources who can make most of them, while those who cannot keep up do more modest normal science research. Following Evans (2010) as well as Wu et al. (2019), this may be a productive division of labor between research aimed at extending and consolidating established knowledge and research exploring new ground. And allocation of resources according to achievement always entails the incentive for all—the elite and the rank and file—to make every effort to improve one's performance. This argument is a kind of meritocratic theory of allocating rewards according to produced value, talent and its scarcity (Cole and Cole, 1973). In the knowledge society, knowledge growth is of highest value, it needs creative talent which is scarce, therefore, most autonomy and most resources are allocated to the most talented researchers. From a critical point of view, this meritocratic argument is blind regarding the unequal conditions under which achievement occurs and the inequality perpetuating consequences of the unequal allocation of rewards (cf. Young, 1958; McNamee and Miller, 2009; Markovits, 2019). And we can add that cemented inequality and particularly extreme inequality in disposing of autonomy, talent, and resources as in the case of a stratification of the academic field into a small group of extremely rich elite universities and the broad mass of modestly equipped rank and file universities fuels the inverted u-shaped relationship between investment and returns (Münch, 2014, pp. 223–233). Most of the rank-and-file universities possess too little autonomy, talent, and resources to contribute effectively to the renewal of scientific knowledge beyond normal science. The small group of extremely rich elite universities attracts talent and resources in such abundance that cannot all be turned into innovative knowledge growth because of excessive coordination costs, demonstration of luxury for its own sake and for securing prestige and unchallengeable competitive advantage. They build an oligopoly which is not only a source of knowledge renewal but also a device of establishing paradigms which resist attempts at change just because their protagonists occupy the most powerful positions in academies, disciplinary associations, scientific journals, and advisory committees.

Talent may arise everywhere and produce new knowledge if reasonably equipped. And it doesn't matter where new knowledge emerges, as long as it is a public good (cf. Callon, 1994). What academic capitalism is strengthening, however, is the very unequal distribution of talent and resources between universities involving brain gain for the rich ones and brain drain for the modestly and poorly equipped ones. Additionally, there is very unequal attention to knowledge production implying its concentration on the elite institutions according to the Matthew principle effective in winner take all markets (Frank and Cook, 2010; Merton, 1968). The result is limited instead of extended potential of knowledge renewal and growth.

Conclusion and outlook

Academic capitalism and market thinking have subjected higher education and research at universities to fundamental changes. The New Public Management reforms with the aim of increasing the performance of universities in research and teaching have been crucial for these changes. The corresponding changes are taking place at three levels: the macro level of the academic field, the meso level of the university as an organization, and the micro level of individual scholars. We have focused on developments that indicate counterproductive effects with respect to the improvement of education and the growth of knowledge in particular. At the macro level, oligopoly structures limit opportunities for broad-based renewal of scientific knowledge. At the meso level, increased managerial control over research and teaching, especially at the most competition-subjected mid-level universities, means an orientation toward the mainstream and a suppression of creative deviance. At the micro level, junior scholars acquire a habitus of compliance to eligible teaching and research formats, especially at mid-ranking universities. Future research should address these questions more broadly and deeply. At stake are the questions of how far academic capitalism and market thinking increase inequalities at the macro level and how far this limits the spectrum of possible renewal of scientific knowledge, how far increased managerial control of research and teaching at the meso level limits the possibilities for creative deviance, and how far academic socialization shapes the habitus of researchers and teachers at the micro level to the extent that they are capable of adapting to prevailing streams of research and teaching but do not follow the curiosity of their own accord.

References

- Amarante, E.F., 2018. The perils of philanthrocapitalism. *Md. Law Rev.* 78 (1). Article 2.
- Anderson, N., Potočník, K., Zhou, J., 2014. Innovation and creativity in organizations: a state-of-the-science review, prospective commentary, and guiding framework. *J. Manag.* 40 (5), 1297–1333.
- Berman, E.P., 2012. *Creating the Market University: How Academic Science Became an Economic Engine*. Princeton University Press, Princeton, NJ.

- Beyer, S., 2021. The Social Construction of the US Academic Elite. A Mixed Methods Study of Two Disciplines. Routledge, London and New York.
- Birkinshaw, J., Visnjic, I., Best, S., 2018. Responding to a potentially disruptive technology: how big pharma embraced biotechnology. *Calif. Manag. Rev.* 60 (4), 74–100.
- Boudreau, K.J., Guinan, E.C., Lakhani, K.R., Riedl, C., 2016. Looking across and looking beyond the knowledge frontier: intellectual distance, novelty, and resource allocation in science. *Manag. Sci.* 62 (10), 2765–2783. <https://doi.org/10.1287/mnsc.2015.2285>.
- Bourdieu, P., 1975. The specificity of the scientific field and the social conditions of the progress of reason. *Soc. Sci. Inf.* 14 (6), 19–47.
- Bourdieu, P., 1988. *Homo Academicus*. Stanford University Press, Stanford.
- Bourdieu, P., 1996. *State Nobility: Elite Schools in the Field of Power*. Polity, New York.
- Brunila, K., Valero, P., 2018. Anxiety and the making of research(ing) subjects in neoliberal academia. *Subjectivity* 11 (1), 74–89.
- Burris, V., 2004. The academic caste system: prestige hierarchies in PhD exchange networks. *Am. Socio. Rev.* 69 (2), 239–264.
- Callon, M., 1994. Is science a public good? *Sci. Technol. Hum. Val.* 19 (4), 395–424.
- Cannizzo, F., 2018. “You’ve got to love what you do”: academic labour in a culture of authenticity. *Socio. Rev.* 66 (1), 91–106.
- Cantwell, B., 2015a. Are international students cash cows? Examining the relationship between new international undergraduate enrollments and institutional revenue at public colleges and universities in the US. *J. Int. Stud.* 5 (4), 512–525.
- Cantwell, B., 2015b. Laboratory management, academic production, and the building blocks of academic capitalism. *High. Educ.* 70 (3), 487–502.
- Cantwell, B., 2016. The new prudent man: financial-academic capitalism and inequality in higher education. In: Slaughter, S., Cantwell, B. (Eds.), *Higher Education, Stratification, and Workforce Development*. Springer, Cham, pp. 173–192.
- Cantwell, B., Marginson, S., 2018. Vertical stratification. In: Cantwell, B., Marginson, B., Smolentseva, A. (Eds.), *High Participation Systems of Higher Education*. Oxford University Press, Oxford, pp. 125–150.
- Carpenter, D.P., 2004. The political economy of FDA drug review: processing, politics, and lessons for policy. *Health Aff.* 23 (1), 52–63.
- Chai, S., Menon, A., 2019. Breakthrough recognition: bias against novelty and competition for attention. *Res. Pol.* 48 (3), 733–747.
- Clark, B.R., 1998. *Creating Entrepreneurial Universities: Organizational Pathways of Transformation*. Pergamon Press, Oxford, UK.
- Cole, J.R., Cole, S., 1973. *Social Stratification in Science*. University of Chicago Press, Chicago, IL.
- Collyer, F.M., 2015. Practices of conformity and resistance in the marketisation of the academy: Bourdieu, professionalism and academic capitalism. *Crit. Stud. Educ.* 56 (3), 315–331.
- Criscuolo, P., Dahlander, L., Grohsjean, T., Salter, A., 2017. Evaluating novelty: the role of panels in the selection of R&D projects. *Acad. Manag. J.* 60 (2), 433–460.
- Delucchi, M., Dadzie, R.B., Dean, E., Pham, X., 2021. What’s that smell? Bullshit jobs in higher education. *Rev. Soc. Econ.* <https://doi.org/10.1080/00346764.2021.1940255>.
- Dey, R., Roy, A., Chakraborty, T., Ghosh, S., 2017. Sleeping beauties in computer science: characterization and early identification. *Scientometrics* 113 (3), 1645–1663.
- Donoghue, F., 2008/2018. *The Last Professors*. Fordham University Press, New York.
- Engelen, E., Fernandez, R., Hendrikse, R., 2014. How finance penetrates its other: a cautionary tale on the financialization of a Dutch university. *Antipode* 46 (4), 1072–1091.
- Erkkilä, T., Piironen, O., 2018. *Rankings and Global Knowledge Governance: Higher Education, Innovation, and Competitiveness*. Springer, Cham, Switzerland.
- Espeland, W.N., Sauder, M., 2007. Rankings and reactivity: how public measures recreate social worlds. *Am. J. Sociol.* 113 (1), 1–40.
- Etzkowitz, H., 2003. Innovation in innovation: the triple helix of university–industry–government relations. *Soc. Sci. Inf.* 42 (3), 293–337.
- Etzkowitz, H., Zhou, C., 2017. *The Triple Helix: University–Industry–Government Innovation and Entrepreneurship*, second ed. Routledge, London, UK.
- Evans, J.A., 2010. Industry induces academic science to know less about more. *Am. J. Sociol.* 116 (2), 389–452.
- Fayrs, R., Wolbring, T., 2021. Matthew effects in science and the serial diffusion of ideas testing old ideas with new methods. *Quant. Sci. Stud.* 2 (2), 505–526.
- Feyerabend, P., 1970/1993. *Against Method*. Verso, London.
- Fleck, C., 2013. The impact factor fetishism. *Eur. J. Sociol.* 54 (2), 327–356.
- Fochler, M., 2016. Variants of epistemic capitalism: knowledge production and the accumulation of worth in commercial biotechnology and the academic life sciences. *Sci. Technol. Hum. Val.* 41 (5), 922–948.
- Frank, R.H., Cook, P.J., 2010. *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us*. Random House, New York, NY.
- Gallo, S.A., Sullivan, J.H., Glisson, S.R., 2016. The influence of peer reviewer expertise on the evaluation of research funding applications. *PLoS One* 11 (10), e0165147.
- Gallo, S., Thompson, L., Schmalig, K., Glisson, S., 2018. Risk evaluation in peer review of grant applications. *Environ. Syst. Decis.* 38 (2), 216–229.
- Geiger, R.L., 2004. *Knowledge and Money: Research Universities and the Paradox of the Marketplace*. Stanford University Press, Stanford, CA.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzmann, S., Scott, P., Trow, M., 1994. *The New Production of Knowledge*. SAGE, London, UK.
- Ginsberg, B., 2011. *The Fall of the Faculty: The Rise of the All-Administrative University and Why It Matters*. Oxford University Press, New York, NY.
- Goffman, E., 1961. *Asylums*. Doubleday, New York.
- Gonzales, L.D., Martinez, E., Ordu, C., 2014. Exploring faculty experiences in a striving university through the lens of academic capitalism. *Stud. High. Educ.* 39 (7), 1097–1115.
- Hall, R., 2018. *The Alienated Academic: The Struggle for Autonomy Inside the University*. Springer, Cham.
- Hallonsten, O., 2021. Stop evaluating science: a historical-sociological argument. *Soc. Sci. Inf.* 60 (1), 7–26.
- Hawthorne, F., 2010. *Inside the FDA: The Business and Politics Behind the Drugs We Take and the Food We Eat*. John Wiley & Sons, New York.
- Hazelkorn, E., 2015. *Rankings and the Reshaping of Higher Education: The Battle for World Class Excellence*, second ed. Palgrave Macmillan, Basingstoke, UK.
- Heinze, T., Münch, R., 2016. *Innovation in Science and Organizational Renewal: Historical and Sociological Perspectives*. Springer, Cham.
- Heinze, T., Shapira, P., Rogers, J.D., Senker, J.M., 2009. Organizational and institutional influences on creativity in science. *Res. Pol.* 38 (4), 610–623.
- Heinze, T., von der Heyden, M., Pitman, D., 2020. Institutional environments and breakthroughs in science. Comparison of France, Germany, the United Kingdom, and the United States. *PLoS One* 15 (9), e0239805. <https://doi.org/10.1371/journal.pone.0239805>.
- Hodgson, G.M., Rothman, H., 1999. The editors and authors of economics journals: a case of institutional oligopoly? *Econ. J.* 109 (453), 165–186.
- Hoffman, S.G., 2020. A story of nimble knowledge production in an era of academic capitalism. *Theor. Soc.* 1–35.
- Hollingsworth, R., 2006. The dynamics of American science: an institutional and organizational perspective on major discoveries. In: Beckert, J., Ebbinghaus, B., Hassel, A., Manow, P. (Eds.), *Transformationen des Kapitalismus*. Campus, Frankfurt/New York, pp. 361–380.
- Holloway, K., Herder, M., 2019. A responsibility to commercialize? Tracing academic researchers’ evolving engagement with the commercialization of biomedical research. *J. Responsible Innov.* 6 (3), 263–283.
- Howell, S.T., 2017. Financing innovation: evidence from R&D grants. *Am. Econ. Rev.* 107 (4), 1136–1164.
- Ioannidis, J.P.A., 2005. Why most published research findings are false. *PLoS Medicine* 2 (8), e124. <https://doi.org/10.1371/journal.pmed.0020124>.
- Jeon, J., 2019. Invisibilizing politics: accepting and legitimating ignorance in environmental sciences. *Soc. Stud. Sci.* 49 (6), 839–862.
- Jessop, B., 2017. Varieties of academic capitalism and entrepreneurial universities. *High. Educ.* 73 (6), 853–870.
- Kauppi, N., 2018. The global ranking game: narrowing academic excellence through numerical objectification. *Stud. High. Educ.* 43 (10), 1750–1762.
- Kauppinen, I., 2012. Towards transnational academic capitalism. *High. Educ.* 64 (4), 543–556.
- Kauppinen, I., 2015. Towards a theory of transnational academic capitalism. *Br. J. Sociol. Educ.* 36 (2), 336–353.
- Khan, J., 2021. European academic brain drain: a meta-synthesis. *Eur. J. Educ.* 56 (2), 265–278.
- Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Kwiek, M., 2019. Social stratification in higher education: what it means at the micro-level of the individual academic scientist. *High. Educ. Q.* 73 (4), 419–444.
- Lee, C.J., 2015. Commensuration bias in peer review. *Philos. Sci.* 82 (5), 1272–1283.

- Leydesdorff, L., 2012. The triple helix, quadruple helix, ..., and an N-tuple of helices: explanatory models for analyzing the knowledge-based economy? *J. Knowl. Econ.* 3 (1), 25–35.
- Leydesdorff, L., 2018. Synergy in knowledge-based innovation systems at national and regional levels: the triple-helix model and the fourth industrial revolution. *J. Open Innov.* 4 (2), 16. <https://doi.org/10.3390/joitmc4020016>.
- Li, D., 2017. Expertise versus bias in evaluation: evidence from the NIH. *Am. Econ. J. Appl. Econ.* 9 (2), 60–92.
- Lin, B., Mainemelis, C., Kark, R., 2016. Leaders' responses to creative deviance: differential effects on subsequent creative deviance and creative performance. *Leader. Q.* 27 (4), 537–556.
- Liu, F., Zhou, K., 2020. Idiosyncratic Deals and Creative Deviance: The Mediating Role of Psychological Entitlement. *R&D Management*. <https://doi.org/10.1111/radm.12430>.
- Lorenz, C., 2012. If you're so smart, why are you under surveillance? Universities, neoliberalism, and New Public Management. *Crit. Inq.* 38 (3), 599–629.
- Mainemelis, C., 2010. Stealing fire: creative deviance in the evolution of new ideas. *Acad. Manag. Rev.* 35 (4), 558–578.
- Marginson, S., 2016. Global stratification in higher education. In: Slaughter, S., Taylor, B. (Eds.), *Higher Education, Stratification, and Workforce Development*. Springer, Cham, pp. 13–34.
- Markovits, D., 2019. *The Meritocracy Trap*. Penguin UK, London.
- McCarthy, G., Song, X., Jayasuriya, K., 2017. The proletarianisation of academic labour in Australia. *High. Educ. Res. Dev.* 36 (5), 1017–1030.
- McClure, K.R., 2016. Building the innovative and entrepreneurial university: an institutional case study of administrative academic capitalism. *J. High. Educ.* 87 (4), 516–543.
- McGoey, L., 2012. Philanthrocapitalism and its critics. *Poetics* 40 (2), 185–199.
- McNamee, S., Miller, R., 2009. *The Meritocracy Myth*. Rowman & Littlefield, New York.
- Mendoza, P., 2007. Academic capitalism and doctoral student socialization: a case study. *J. High. Educ.* 78 (1), 71–96.
- Mendoza, P., Kuntz, A.M., Berger, J.B., 2012. Bourdieu and academic capitalism: faculty "habitus" in materials science and engineering. *J. High. Educ.* 83 (4), 558–581.
- Merton, R.K., 1968. The Matthew effect in science. *Science* 159 (3810), 56–63.
- Merton, R.K., 1973 [1942]. The normative structure of science. In: Merton, R.K. (Ed.), *The Sociology of Science*. University of Chicago Press, Chicago, IL, pp. 267–278.
- Morrish, L., 2020. Academic freedom and the disciplinary regime in the neoliberal university. In: Dawes, S., Lenormand, M. (Eds.), *Neoliberalism in Context*. Palgrave Macmillan, Cham, pp. 235–253.
- Münch, R., 2014. *Academic Capitalism: Universities in the Global Struggle for Excellence*. Routledge, London and New York.
- Münch, R., Baier, C., 2012. Institutional struggles for recognition in the academic field: the case of university departments in German chemistry. *Minerva* 50 (1), 97–126.
- Münch, R., Schäfer, L.O., 2014. Rankings, diversity, and the power of renewal in science: a comparison between Germany, the UK and the US. *Eur. J. Educ.* 49 (1), 60–76.
- Musselin, C., 2018. New forms of competition in higher education. *Soc. Econ. Rev.* 16 (3), 657–683.
- Owen-Smith, J., 2003. From separate systems to a hybrid order: accumulative advantage across public and private science at Research One universities. *Res. Pol.* 32 (6), 1081–1104.
- Owen-Smith, J., Powell, W.W., 2004. Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community. *Organ. Sci.* 15 (1), 5–21.
- Paasi, A., 2015. Academic capitalism and the geopolitics of knowledge. In: Agnew, J.A., Mamadouh, V., Secor, A.J., Sharp, J.P. (Eds.), *The Wiley Blackwell Companion to Political Geography*. Wiley Blackwell, Hoboken, N.J., pp. 507–523.
- Parsons, T., Platt, G.M., 1973/2013. *The American University*. Harvard University Press.
- Pederzini, G.D.A., 2018. Neoliberal awakenings: a case study of university leaders' competitive advantage sensemaking. *High. Educ. Pol.* 31 (3), 405–422.
- Power, M., 1997. *The Audit Society: History, Institutions, and Social Analysis*. Princeton University Press, Princeton, NJ.
- Readings, B., 1996. *The University in Ruins*. Harvard University Press, Cambridge, MA.
- Reay, D., 2014. From academic freedom to academic capitalism. *Discover Soc.* 5, 15.
- Rikap, C., Harari-Kermadec, H., 2020. The direct subordination of universities to the accumulation of capital. *Cap. Cl.* 44 (3), 371–400.
- Rudy, A.P., Coppin, D., 2007. *Universities in the Age of Corporate Science: The UC Berkeley—Novartis Controversy*. Temple University Press, Philadelphia, PA.
- Rushforth, A., Franssen, T., de Rijcke, S., 2019. Portfolios of worth: capitalizing on basic and clinical problems in biomedical research groups. *Sci. Technol. Hum. Val.* 44 (2), 209–236. <https://doi.org/10.1177/0162243918786431>.
- Sauder, M., Espeland, W.N., 2009. The discipline of rankings: tight coupling and organizational change. *Am. Socio. Rev.* 74 (1), 63–82.
- Schulze-Cleven, T., Olson, J.R., 2017. Worlds of higher education transformed: toward varieties of academic capitalism. *High. Educ.* 73 (6), 813–831.
- Shibayama, S., Walsh, J.P., Baba, Y., 2012. Academic entrepreneurship and exchange of scientific resources: material transfer in life and materials sciences in Japanese universities. *Am. Socio. Rev.* 77 (5), 804–830.
- Shields, R., Watermeyer, R., 2020. Competing institutional logics in universities in the United Kingdom: schism in the church of reason. *Stud. High. Educ.* 45 (1), 3–17.
- Slaughter, S., Leslie, L.L., 1997. *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*. The Johns Hopkins University Press, Baltimore, MD.
- Slaughter, S., Rhoades, G., 2004. *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. The Johns Hopkins University Press, Baltimore, MD.
- Slaughter, S., Thomas, S.L., Johnson, D.R., Barringer, S.N., 2014. Institutional conflict of interest: the role of interlocking directorates in the scientific relationships between universities and the corporate sector. *J. High. Educ.* 85 (1), 1–35.
- Stevens, T., Newman, S., 2019. *Biotech Juggernaut: Hope, Hype, and Hidden Agendas of Entrepreneurial Bioscience*. Routledge, New York, NY.
- Stöckelová, T., Vostal, F., 2017. Academic stratospheres-cum-underworlds: when highs and lows of publication cultures meet. *Aslib J. Inf. Manag.* 69 (5), 516–528.
- Tenzer, H., Yang, P., 2019. Personality, values, or attitudes? Individual-level antecedents to creative deviance. *Int. J. Innovat. Manag.* 23 (02), 1950009.
- Tuchman, G., 2009. *Wannabe U: Inside the Corporate University*. University of Chicago Press, Chicago, IL.
- Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical combinations and scientific impact. *Science* 342 (6157), 468–472. <https://doi.org/10.1126/science.1240474>.
- van Raan, A.F.J., 2004. Sleeping beauties in science. *Scientometrics* 59 (3), 467–472.
- van Raan, A.F.J., 2015. Dormitory of physical and engineering sciences: sleeping beauties may be sleeping innovations. *PLoS One* 10 (10), e0139786. <https://doi.org/10.1371/journal.pone.0139786>.
- Wang, J., Veugelers, R., Stephan, P., 2017. Bias against novelty in science: a cautionary tale for users of bibliometric indicators. *Res. Pol.* 46 (8), 1416–1436.
- Wang, J., Lee, Y.-N., Walsh, J.P., 2018. Funding model and creativity in science: competitive versus block funding and status contingency effects. *Res. Pol.* 47 (6), 1070–1083.
- Washburn, J., 2005. *University, Inc.: The Corporate Corruption of Higher Education*. Basic Books, New York.
- Weber, M., 1976. *Wirtschaft und Gesellschaft*. Mohr Siebeck, Tübingen, Germany.
- Welsh, J., 2020. Dispossessing academics: the shift to "appropriation" in the governing of academic life. *Eur. J. Soc. Theor.* 23 (3), 350–369.
- Welsh, J., 2021. Tolling academics: rent-seeking and gatekeeping in the university space. *Cap. Cl.* 45 (1), 93–121.
- Wheaton, A., 2020. Shift happens; moving from the ivory tower to the mushroom factory. *High. Educ. Res. Dev.* 39 (1), 67–80.
- Whitley, R., Gläser, J., Laudel, G., 2018. The impact of changing funding and authority relationships on scientific innovations. *Minerva* 56 (1), 109–134. <https://doi.org/10.1007/s11024-018-9343-7>.
- Wieczorek, O., 2021. *Die Universität im Feld der Macht: Zur Konstruktion von Expertise in der US-amerikanischen Wissenschaftsgesellschaft*. Diss. rer. pol. University of Bamberg.
- Wieczorek, O., Beyer, S., Münch, R., 2017. Fief and benefice feudalism. Two types of academic autonomy in US chemistry. *High. Educ.* 73 (6), 887–907.
- Woelert, P., Lewis, J.M., Le, A.T., 2020. Formally alive yet practically complex: an exploration of academics' perceptions of their autonomy as researchers. *High. Educ. Pol.* <https://doi.org/10.1057/s41307-020-00190-1>.
- Wu, L., Wang, D., Evans, J.A., 2019. Large teams develop and small teams disrupt science and technology. *Nature* 566 (7744), 378–382.

Yeung, A.W.K., Ho, Y.-S., 2018. Identification and analysis of classic articles and sleeping beauties in neurosciences. *Curr. Sci.* 114 (10), 2039.

Young, M., 1958. *The Rise of the Meritocracy*. Transaction, Piscataway, NJ.

Young, N.S., Ioannidis, J.P., Al-Ubaydli, O., 2008. Why current publication practices may distort science. *PLoS Med.* 5 (10), e201. This page intentionally left blank

“Reschooling” higher education: the dynamics of curriculum in South Africa

Kirti Menon^a and Gloria Castrillón^b, ^a Division for Teaching Excellence, DST/NRF South African Research, Teaching and Learning (PSET), Faculty of Education, University of Johannesburg, Johannesburg, South Africa; and ^b Center for Academic Planning and Quality Promotion, Faculty of Education, University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	48
About curriculum	48
De- and “re-schooling” higher education systems	49
The politicization of curriculum	50
A critique of the outcomes-based approach in SA education	51
Strategic levers and curriculum	52
Supercomplexity as a driver of curriculum change	53
Meeting the needs of decolonization	53
The 4IR and planning for the 21st century	54
Conclusion	55
References	55

Introduction

With the advent of the “knowledge economy”, higher education (HE) has increasingly been viewed as the vehicle for transformative socio-economic change. Economists Maneejuk and Yamaka conclude that while “secondary enrollment rates positively affect economic growth” it is HE which is “the key to future growth and sustainability” (2021: 520). It is no surprise therefore that research into curriculum theory has expanded as a field in the last half-century. South Africa (SA) is no exception, and HE curricula have increasingly been under pressure not only to transform the economic lives of its populace, but also to “reschool” South African society on a wide range of social issues. Examples include diversity and multiculturalism, HIV and AIDS, and internationalization, to name a few. Added to the specifics of South Africa’s realities are the pressures of globalization, a weakening economy, and an ever-increasing gap between the haves and the have nots. The result is a heady mix of demands on the higher education system which, coupled with decreasing resources, results in yet more contestation over curricula in the system.

Higher education systems across the world have had to adjust to many of the same pressures, leading to a field of scholarship on curriculum which reflects on the conceptual, ideological and pedagogical issues that arise. As Lindén, Annala and Coate argue, “[a]lthough curricula are commonly investigated in HE research the field lacks a coherent theoretical approach to explaining the meaning and the importance of curriculum, particularly as part of the special role played by HE in society” (2017: 137). The incoherence results from the continuously increasing influence exerted by external stakeholders, the proliferation of the “knowledge economy”, and the pressure for institutions to compete to prove “excellence”. These are but some of the factors influencing how curricula are conceptualized, designed and articulated.

This chapter argues for the reschooling of higher education and focuses on a reconceptualization of the role of curriculum as central to this process. It raises, critically, the need for a society-wide reorientation of universities to respond to societal needs and challenges while addressing the competing ideologies that work to shape the decisions made in the sector. Curriculum theory is used to make sense of the dynamics at stake in pivotal shifts of the contemporary movements, such as decolonization, the Fourth Industrial Revolution (4IR) and the pressure to “produce” future-fit graduates. The focus in this chapter is on the taught curricula in universities and not on the structural educational reforms that occurred, such as the introduction of qualification frameworks, quality councils, and so on. The South African context is not necessarily unique however, and there may be lessons to be learned by and from other countries, whether rich, middle-income or poor.

About curriculum

Short (1993) frames the curriculum as the disciplinary and mission-oriented or practical knowledges that are held and transmitted by the university. Practical knowledge develops in situations where human activities and behaviors “pose questions that must be answered” (Short, 1993, p. 2) before any kind of concrete or appropriate action can be taken. Disciplinary knowledges are the structures in which meta-theories of reality and knowledge are organized and articulated, and through which practical knowledge is translated from abstract to concrete, via research, teaching and observed practice. These knowledges go beyond official curriculum content to include particular sets of ethics, codes and subjective experiences that occur through program exposure (an easy example would be for example, such as in medicine and law). There are critiques of the substance of particular curricula and the particular

meanings of curriculum as a concept. Barnett (Hicks, 2018) provides perhaps the most salient of these: curriculum is a vehicle for the process of knowing and understanding, teaching content, skills and forming personhood.

Baker (1996) suggests disciplinary knowledge must be situated within the socio-historical context it emerges from and engages with. It is the socio-historical context which will inform explanatory theories that make sense of and speak back to the world. Baker (1996) points out that much of the early critical work on curriculum in the 20th century focused on the relationship between socio-historical conditions and the spheres of social and specialized knowledge (Baker, 1996). While this relationship itself has not changed, the breadth of differentiation and pace of change within and outside of institutions has created what Barnett and Coate (2005) call “supercomplexity”. In a supercomplex environment institutions are under siege and are expected to remain responsive and accountable. Hicks (2018) suggests that there is a “new kind of chaos” to the usual pressures in higher education, with greater influence exerted by external variables, and a wider array of challenges confronting universities than they traditionally have had to accommodate. Developing a comparative understanding of how curricula function and of how they are contextually located has become a crucial priority to support navigating complexity.

Barnett and Coate (2005) argue that there is a dire need for curricula to be systematically reviewed to function optimally in a supercomplex world, reshaping and disrupting spaces for learning. Without this, there is an increased risk of ad hoc decision-making and implementation, with universities caught in a never-ending game of “catch-up” in a world that changes too rapidly for existing systems to respond effectively. Definitions of curriculum vary but focus primarily on the planned dimensions of learning and not on the unintended experiences, outcomes and discourses.

Research into curriculum continues to grow as higher education faces a further paradigm shift in the meaning and nature of teaching and learning. The globalization of international markets has contributed to the rise of comparative higher education studies which treat curriculum as a phenomenon of theoretical and practical inquiry. This includes how new forms of blended learning trouble the traditional “boundaries” of the curriculum, altering the role of lecturer and student, and creating new mediators of access and knowledge through the devices and information sources that extend beyond the classroom. The intertwining of curriculum and pedagogy comes to the fore as there is a shift toward a more parochial discourse on curriculum as a driver of economic and technological professionalization. The degree to which universities are able to respond to these pressures, however, is at least in part determined or affected by their national and regional contexts, as examples from South Africa, India and Ghana show.

Defining curriculum is complex as Lindén et al. (2017) argue that the field of curriculum studies lacks a coherent theoretical approach which explains the meaning and importance of curriculum and its role. There is a tendency to treat curriculum as either the sum of its constituent elements (such as assessment strategies and modules) or as the content that gives it meaning. The authors follow Young (2013) in arguing that a “curriculum crisis” has taken hold, the solution to which has been to emphasize the importance of aspects outside of curriculum, such as national qualification frameworks, holding on to learning outcomes as essential indicators of quality and relevance, and shifting from a disciplinary focus to more practical, generalized curriculum design issues. The lack of a coherent curriculum theory has been subsumed by ideological issues, fuzzily defined in terms such as “quality”, “relevance”, “excellence” and “efficiency” (Lindén et al., 2017). What is learned and how is not treated as an instructional question with a theoretical foundation which continues to be a challenge faced in higher education systems around the world. Higher education occupies an important place in the state’s strategic planning for employment, job creation and skills specialization, and curricula too are now vital dimensions of institutional strategic planning processes and not just mechanisms of knowledge transfer (Lindén et al., 2017). The problem is that curriculum as a phenomenon is essentialized to its technical elements to render it legible for planning processes which diminishes the question of the knowledge that is at the core of what a curriculum is (Barnett and Coate, 2005). The focus shifts toward tighter policymaking and designing learning experiences toward predetermined and clearly defined outcomes. The South African experience, it can be argued, demonstrates how a focus on outcomes may detract from a more broadly conceptualized need for higher education to shape society. In simple terms, there is no cogent evidence to suggest that the shift has had the significant impact policymakers assumed it would.

De- and “re-schooling” higher education systems

The question of the knowledge underpinning university curriculum touches at the heart of the politics of knowledge, and what one defines as knowledge and the different value tags attached to it. Pathways to epistemological and epistemic access have become a major curriculum turning point in the last decade in South Africa. Universities have been challenged continually to address the exclusions and silences within canons of knowledge that continue to act against deep social change (Sayed et al., 2019). In assessing the history of education in the South from colonial times through to the post-apartheid era, Soudien defines what he sees as the fundamental driver of curriculum in the apartheid state: “social difference, as opposed to, say, pedagogical reform” (2010: 20). His work demonstrates that curriculum development processes in former colonial countries required “a forceful incorporation into the dominant ideological structures of the world” (Soudien, 2010: 20). In South Africa, this social difference was encapsulated in the system of race, which was not critically interrogated after 1994 and as a result continued to act as a “definitional boundary” in the “new” South Africa (Soudien, 2010: 43): “The point to note in this analysis is how comprehensively the new curriculum has catalyzed and even amplified the major vectors of discrimination inherited from the past”.

The concepts of de-schooling and re-schooling are useful for making sense of the contradictory processes effected at the turn to democracy. “De-schooling”, as coined by Illich (1970), is a critique of schools as institutions that are ineffective learning sites which alienate and dehumanize learners and the process of learning. Illich’s work called for more open-ended, networked and holistic sites

where learning could take place, such as libraries and museums. The focus on “outcomes-based education” shares this appeal in its emphasis on unique and learner-centered experiences. However, there is also a conservative application of de-schooling as a framework which seeks to question the public value of education as a whole, shifting focus to whether schooling delivers on private outcomes such as higher salaries and greater social mobility, and disaggregating certifications to technical competencies (Sukarieh and Tannock, 2019).

Re-schooling is defined by Hartley (1997) as the *reinvention* of schools or education systems toward visions, purposes and practices that are reflective of and responsive to changing social conditions. The dual focus on decolonization and the 4IR represents one example of how re-schooling efforts are being made in higher education in South Africa currently. This chapter develops a working definition of re-schooling as changing the ways in which policy seeks to address core issues which are educational *as a result* rather than as a cause—in this case, society-level issues or challenges that cannot be resolved only at the level of curriculum. This also requires addressing the temporal challenge in curriculum design, which Soudien defines as the crafting of higher education policy for the society that the post-apartheid South African state *wanted* rather than the society it *had* (Soudien, 2010). This vision-setting approach belied the cleavages and inequities that the system inherited, deepening existing inequalities over time. The discourses of decolonization and 4IR are also somewhat competing visions for the system, but it is essential to understand both in order to make sense of the demands of reschooling society for more socially just and responsive higher education.

The challenge that emerges is how to synthesize a range of normative ideals and priorities—from sustainable development to nation-building, decolonization and the 4IR—in ways that produce cogent learning experiences within the broad scope of the disciplines in the university. That all of this must occur within the existing curriculum policy architecture in South Africa means that institutions will approach these priorities differently based on their own situational dynamics, constantly renegotiating the balance between policy and how this is implemented. The challenges to curriculum development that have most impacted higher education in the post-1994 period have been *#RhodesMustFall* and the articulation of the Fourth Industrial Revolution (4IR) as a critical epoch in global development. The encapsulation of the rejection of ostensible transformation saw the *#RhodesMustFall* student protests spread across South Africa and internationally. Student activists highlighted the lack of transformation of universities, and their resulting alienation from both the institutions in which they studied, and the epistemologies to which they were denied access. More recently, COVID-19 provided fertile testing ground for how South African higher education institutions have developed to address the complexities of the 21st century, notwithstanding the complex social and economic strictures within which they work.

The politicization of curriculum

The politicization of curriculum is not unique to any country, especially as decision-making around canons of knowledge remains inherently political. Its particular form in South Africa as well as other examples from India, Kyrgyzstan, and Ghana are instructive. Recently, the University of Delhi became a site of struggle over curriculum in four disciplines. English, Political Science, History and Sociology have fallen foul of the Indian government’s need to control universities at the cost of the academic freedom. Khattar (2019) shows how the Hindu nationalist party, *Bharatiya Janata* (BJP), is attempting to establish a Hindu majoritarian state by subverting the existence and histories of cultural, religious, and political minorities. The conflict is clearly politically motivated and demands relating to curriculum align directly to the interests of the ruling party. The University of Delhi has rebutted these requests, nonetheless the national conflict affected elections for students’ and staff’s representative bodies. This is an unambiguous example of the manipulation of curriculum by political powers.

Broadly speaking in the African context, Tade Akin Aina notes “the recognition that Africa’s stories of the development process are narratives of blocked transitions, routes not taken, itineraries hijacked, highways populated by robbers and bandits, and journeys led by shifty and unreliable scouts and guides (2010: 22). Against this backdrop, he argues, the “terrain of African higher education continues to resemble a thick forest of institutions, systems, and practices lacking clear and distinct tracks, values, and goals, or a mission and vision that connect the institutions and systems sufficiently to the major challenges of their contexts (whether global or local)” (2010: 23). Despite this, African universities remain important role players in the economic and social constitution of society, albeit constrained by the “blueprints” of institutions used to accredit universities by regulatory bodies (2010: 36). These institutions will need drastically to change their approaches to knowledge production and to the partnerships they select if they are to become “equal players in the global arena while remaining relevant nationally and locally” (2001: 36).

Shadymanova and Amsler (2018) offer a fascinating account of the Kyrgyzstan education system post the collapse of the USSR. Higher education institutions, despite being subject to government regulation, market forces, “new legal and political frameworks, changes in the relationship between higher education and labor”, nonetheless expressed their own institutional “strategies and responses to environmental changes” (2018: 230). One year post independence from the USSR, the Kyrgyzstan government enacted a law “On Education” which sought to reform the system for the political and economic context. Significantly, the law seeks to develop diverse educational programmes, “new learning forms and technologies”, and new funding sources (2018: 236). The state retained responsibility for basic funding, for standard setting for the different levels, for the approval of priorities in curriculum development, among other aspects. HEIs in Kyrgyzstan are required to adhere formally to quality assurance requirements and must meet the requirements of the state to operate (2018: 236), resulting in a significant increase in student numbers across the system.

In a less overt example, Gyamera and Burke (2017: 14) argue that transformation in Ghanaian universities has been perceived of in terms of “Western-oriented standards and neoliberal discourses, where there is a privileging of the economic”. The authors

comment on the failure to consider questions about the interplay between curriculum, knowledge, and power. As in South Africa, there can be no reliance on a “common sense” approach given the implications of this, and its historic marginalizing of local and indigenous knowledge forms.

The apartheid legacy of defining and dividing institution types by race or jurisdiction wrought permanent damage on the foundations of South Africa’s educational landscape. The political turbulence of the early 1990s with its emphasis on restructuring South African society along equitable lines saw the introduction of swathes of legislation designed to effect fundamental social and political change (for example, the *Employment Equity Act*, passed in 1998; the *Labor Relations Act*, passed in 1995; and the interim and final constitutions, passed in 1993 and 1996 respectively). With the transition to democracy in 1994, and against this highly iniquitous historical backdrop, the new state looked to the hotly contested HE arena as a critical lever for social, political and, significantly and as in the Ghanaian context, economic change. Opening access to higher education to all, regardless of race or other grounds would, it was argued, set the foundations for addressing apartheid’s negative structural and social legacy. The simplicity of the connection between increased access to education and employability and the resulting economic (and social) transformation was, however, thoroughly under-interrogated.

The complexities of such assumptions for higher education are replete. In addition to the demands of social context, the purpose of a qualification, its location in a discipline, the teaching and learning culture of an institution and the interplay between institutional, national and global contextual forces, universities were and are seen as responsible for articulating this change at the level of the learning program. At the level of the program of learning, and what takes place in a university, is the module, which is in the hands of the individual lecturer, or group of lecturers. Applying the imperatives to “curriculum” is thus almost intangible, and impervious to measurement. In addition, and as an added complexity, it is only recently that universities have begun to think about the pedagogies underpinning curricula. As universities change, and regardless of what these changes are, they translate these re-defined purposes and actions into their society. They feed into and from the challenges and complexities they are surrounded by and are required somehow to wrestle these into their curriculum planning processes (Barnett and Coate, 2004; Cross and Atinde, 2015).

Ramrathan argues that, despite the degree and extent of change in the SA HE system, “fundamental curriculum transformation has not yet taken place” (2016: 2). Using the Bachelor of Education as an example, he argues that the lack of transformation is because “what is kept and what is left out [of the curriculum is] based on what political decisions govern their actions” and programmes that did meet the demands of the Department of Higher Education were not approved for offering (2016: 7). Deep transformation of the curriculum, he argues, will be effected by obtaining new insights into curriculum, by focusing on “situated and propositional ontologies, innovative and transformative ways of inquiry” and by “embracing complexities and uncertainties” (2016: 7).

The paradox lies in that the architecture of a curriculum depends on decisions being made about how teaching and learning will take place. Lecturers use the content of their disciplines to create and teach the knowledges necessary to “graduate” the student with a set of intellectual and practical assets recognized as a qualification. But, as Bernstein (1996, p. 47) points out, it is through curriculum that society “selects, classifies, distributes transmits and evaluates” what it considers “to be public, [which] reflects both the distribution of power and the principles of social control”. This means that curricula are complex compositions of value-judgments intended to translate into social and individual value through graduates. The question of what is to be valued, and how it is to be articulated, remains a contested issue between different stakeholders in the higher education terrain, with reschooling offered as one process through which this question may be resolved.

Clearly, and as discussed elsewhere in this chapter, there is no uniformity in defining curricula, and it is also not easy to measure. Even from a policy perspective, there is no shared definition of what it is that the various imperatives should impact on. Using the state’s own definitions, it is the university’s responsibility to give effect to the policy directives through learning programmes and qualifications. The nature, form and translation of change occurs in a particular learning space beyond the reach of a bureaucracy to find. If the past decade in South African higher education taught universities anything, it is that what universities thought about themselves and what students thought about them were leagues apart, and it remains unclear whether it will be curricula alone that bridge that gap. Re-schooling offers an opportunity to orient interlocking systems and mechanisms of social reproduction toward more robustly determined developmental and ethical goals. The next section considers this in light of the continued prevalence of outcomes-based thinking in South African higher education.

A critique of the outcomes-based approach in SA education

There is widespread critique of outcomes-based education (OBE) in SA which, although having been withdrawn for use in schools, remains the model for higher education. As in Australia, SA regulators adopted OBE given its links to constructivist approaches to learning, focusing on the uniqueness of the learning experience for different students (which acquires more emphatic meaning given South Africa’s history), and the need to relate learning to “real-world” experiences. However, as Donnelly points out, a flaw in constructivism “is that it takes a common-sense observation, that learning is actively constructed, and that each person internalizes learning in a unique way and distorts it to such an extent that it becomes counter-productive” (2007: 12). Jansen (1998: 321) identifies other weaknesses, posing the following questions:

Do outcomes in fact deliver what they claim? How do outcomes play out in a resource-poor context? Can outcomes survive their psychological roots in behaviourism? Do outcomes in different contexts mean the same thing, e.g., are outcomes specified for education equivalent to those identified for training?

Having progressed through an OBE-driven system, and as a result of a lack of clarity on what the learning outcomes are or should be, coupled with a failure to assess these, students end up with “pseudo-knowledge, pseudo-skills, pseudo-attitudes and pseudo-values” (Malan, 2000: 22). According to Spreen and Vally (2010) the challenges SA faced post-apartheid led to the implementation of OBE as a means of differentiating the “new” SA from the old. In the new South Africa, it was anticipated that the “skills, knowledge and values [needed] for social justice, equality and development” could be provided through higher education (Spreen and Vally, 2010: 40). Critical to the regulators also (Sayed et al., 2019) was the perceived link between economic development and OBE. Problematically, however, education alone could not address the “social inequalities that were the legacy of apartheid education” and OBE “mostly failed and, in fact, resulted in exacerbating rather than overcoming inequality” (Spreen and Vally, 2010: 40). It is possible that the transformative and educational burden placed on OBE was simply too much. It is clear that in practice it did not work, although this is certainly more tangibly defined in the schooling sector and more intangible in higher education.

Most critiques of OBE have focused on school education, and it is fair to argue that OBE has, in large part, simply been ignored by the higher education sector, aside from where it matters: in engagements with the regulatory bodies. As the events of 2015 highlighted, universities had largely continued as they always have, with decisions regarding modules and curricula remaining in the domain of the individual lecturer. The OBE clothing simply enrobed the same body higher education had always had. In addition, neither OBE nor other systemic interventions in higher education have been able to resolve the socio-political, economic and other woes that bedevil SA society. This has meant that the re-schooling of both basic and higher education that was expected to occur through post-apartheid policymaking has largely failed through a failure to address the root causes of continued educational inequalities.

These root causes were sharply defined by the #RhodesMustFall student movement beginning in 2015. Students protested against untransformed university curricula, institutional cultures and teaching practices, as well as the material barriers to participation for poor students which included the high cost of fees. They issued a call for the decolonization of higher education including the removal of barriers to access to an important lever of social mobility in a society still grappling with the inequities of its apartheid past. While in the aftermath, universities have engaged with these concerns at both the curriculum and material levels, the need to fundamentally re-school society was placed on the backburner—despite students’ demands indicating that this would be essential to realizing the goal of decolonized higher education (Jansen, 2019).

Strategic levers and curriculum

There is no question that the history of education, and not only HE in SA, is brimming with rhetoric. Currently, the policy rhetoric is that curriculum change in HE takes place according to a set of “strategic levers”. By way of illustration, since 1994 higher education has had overtly to respond, as a matter of national policy, to a multiplicity of imperatives, often ill-defined and vague, with few attempts to arrive at a shared understanding of their meanings. The notion of “transformation” for example, remains largely undefined so that despite having “transformed” their staffing and student demographic profiles, for example, numerous universities across the country faced the accusation of not having transformed their curricula, as the events of 2015 and #RhodesMustFall so clearly indicated.

Other policy imperatives since 1994 have included HIV-AIDS “awareness”, entrepreneurship, employability, equity and access, and student success. These are stacked on top of the introduction of the National Qualifications Framework (NQF) (designed to facilitate articulation across the sector), OBE discussed above, a variety of transformation reports (for example, RSA, 2008), the realignment of qualifications to meet the National Plan and a new framework for higher education, the 2015–2016 discourses of decolonization, the post-2019 discourse of the 4IR, Sustainable Development Goals, and the current focus on the “21st century graduate”. The result is that all of these imperatives have been supposed to play out in the hotly contested space of what is termed curriculum. What is not in question is the enormous pressure exerted on academic curricula to comply. It is also not clear to what extent the meanings of these imperatives or levers are shared across 26 public universities, the private higher education sector, thousands of academics, tens of thousands of qualifications, and countless curricula, as misnomers and misimpressions result.

It is not within the scope of this chapter to state with any degree of certainty whether these “imperatives”—or which of them—has resulted in deep, meaningful curriculum change across the sector. Even where change can be said to take place, it cannot be argued that this has occurred with any evenness across or even within institutions. At the level of curriculum, a simple example is that the insertion of a module on African thought into a qualification does not mean that “decolonization” has occurred. The same is true for entrepreneurship, as Nabi et al. (2018) demonstrate: entrepreneurship education can “foster but also decrease entrepreneurial intentions”. Where increased, students have a “theoretical understanding of the start-up process ... combined with a strong inspiration”; where it decreases, however, “students are developed as individuals” but have a “more realistic and practical perspective on entrepreneurship” (Nabi et al., 2018: 463). As a consequence, aspects of curricula amendments—whether to programmes, modules or the content, may tend to become mere ciphers or symbols of the change they are meant to achieve. There is, equally, no measure at present to determine whether change has taken place or, if it has, that it has been in the particular direction envisaged.

Part of the difficulty is the nature of measurement and what would be measured even if there were such a measuring tool. As a result, assessments of the impact have tended to remain anecdotal at best.

Supercomplexity as a driver of curriculum change

The above pressures on higher education may be seen as discrete, but as the SA context demonstrates, this is not the case. This is simply an example of how supercomplexity impacts curriculum decision-making in a single higher education system (Barnett and Coate, 2005). There are the additional contextual and regional pressures pertaining to the global South, especially as it seeks its own solutions distinct from those proffered by the global North. Added to this, the (super)complexities impacting on curricula are quite literally the lived realities of the majority of South Africa’s students—whether climate change, poverty and economic austerity, or globalization, to name a few. In this way, the politics of knowledge continues to frame the robust question of what is to be learned, which within the contested arena of SA higher education remains no simple feat. This includes the making of curriculum choices that are specifically extraverted in nature, geared as they are toward producing graduates for a global marketplace (Tyler in Lindén et al., 2017). From the student perspective, there is pressure to be “job-ready” and be able to find or create employment and having the skills to do so. In this way, issues such as unemployment form part of the sphere of influences that drive changes to the curriculum through both state policy and ground-level pressures. The official South African unemployment rate of 32.5% is concerning, with the youth unemployment rate for ages 15–24 years and 25–34 years at 63.2% and 41.2%, respectively (Stats SA, 2021). Given the complexity of the South African contexts of higher education, the increasing pressures on curriculum are not likely to ease, and yet a better approach is clearly needed.

Employability continues to be a central issue of curriculum development in many African universities (Gyamera and Burke, 2017). Gyamera and Burke (2017) argue that the desire for “world-class”, “excellent” universities, defined in terms of the higher education systems of wealthy Northern countries, exerts significant influence on decision-making in curriculum planning in order to streamline graduates into an international labor market and knowledge economy. University leaders in their research spoke directly to making curriculum decisions that foregrounded expertise or experiences in the global North as a means of attracting funding or partnerships, or being pressured to do so in order to access donor finance (Gyamera and Burke, 2017), with the result that curricula are being amended on an often ad hoc basis to take on new perspectives and theories of change, while at the same time being standardized into the image of global curriculum planning that may be at odds with the contextual needs of students and staff. This results in, for example, the rationalization of micro-economics and other “orthodox” branches of Economics at a Ghanaian university in favor of courses that offer more practical and “competitive” skills—representing a break with critical forms of disciplinary knowledge that inform research and other high-prestige activities in the global North.

In addition, universities have to account for a myriad of considerations from stakeholders in addition to government and civil society, and which include labor, business and market institutions. The focus in South Africa on entrepreneurship and employment—to address the aims of development and redress—is a normative position that further shapes the tendency toward demonstrable and accessible “packages” of learning outcomes, as indicated in the linkages to OBE above. The competing pressure for the curriculum to factor in issues such as globalization, interdisciplinarity and preparation for 21st century conditions is also evident. This introduces key attributes such as lifelong learning, critical thinking skills and social responsibility which also require adaptation and orientation into competencies that can be defined and measured effectively.

Universities are also required to provide learning opportunities for students to become data and media-literate in a world of changing representations. This means offering students the knowledge and social relations needed to enable them to critically read—not only cultural texts and how these are regulated by various discursive codes—but also how such texts express and represent different ideological interests (Giroux, 1992:135). The complex array of skills and attributes curricula are meant to address places pressures on lecturers and students alike who are required to translate this ever-changing scope into meaningful, practical, intellectual and personal skills and competencies despite learning gaps deriving from specific lived realities. Some of these skills, attributes and competencies acquire a political flavor when understood in the context of macro-political events and trends, as the section below describes.

Meeting the needs of decolonization

The current focus on cultivating standardized and transferable experiences is counterintuitive in the context of renewed calls for the decolonization of HE. The term “epistemological access” is regularly used in South Africa to explain the extent to which students’ prior knowledge and experiences are recognized and incorporated into higher education. In South Africa, this produced a reckoning with the reality that many of the curricula in universities continued to be informed by colonial epistemologies even after the transition to democracy (Le Grange, 2016) which drove the curriculum debates during the 2015 #RhodesMustFall movement and after. Masenya (2021) writes that the issues raised by students gave rise to the interrogation of curriculum decision-making and learning content, and the relationship between the two. At the forefront of the debates that raged was the deliberate and sometimes unconscious exclusion of epistemologies associated with black contemporary and indigenous thought, as well as literature and critique by black scholars.

The decolonization movement in South Africa also touched on the material experiences of students and the question of the relevance of their studies, specifically in the context of a difficult job market and high expectations of returns to higher education. Gya-mera and Burke’s (2017) findings in relation to the Ghanaian case are instructive in this regard. There is a continuous challenge in balancing relevance to labor market outcomes with the need for a curriculum orientation grounded in Africa, drawing on its knowledge systems and responsive to the needs of states located in the global South.

The 4IR and planning for the 21st century

The 4IR signals a further shift in the imperatives underpinning curriculum in South Africa, as well as the linkages between curriculum and social or economic outcomes. Explicit decisions to include 4IR in agenda-setting for curriculum policy has given rise to the inclusion of new sets of key competencies, skills and capacities, as well as the personal attributes required by 4IR-ready graduates. The 4IR is characterized by the expansion and integration of exponential technologies and the reliance on knowledge and big datasets to inform computing, analytics and real-time decision-making (Schwab, 2016). On one hand, it brings into focus the growing influence of artificial intelligence, machine learning and the Internet of Things on changes to the nature of work and the economy. However, it also heralds a new understanding of how technology and data are shaping the everyday experiences of human life. Literature on the 4IR and the incorporation of the 4IR into education policy generally begins from the standpoint that what universities are currently producing in graduates is ineffective and insufficient for the current technological and social shifts underway (Menon and Castrillon, 2019).

The greater challenge at work in the implementation of the 4IR in the curriculum has to do with the fundamental axial shift implied by the transition to a new set of conceptual and practical tools—what Bruno-Jofré and Zaldivar (2012) call the shift from the page to the screen. Within this new technological normativity, there is a rupture implied where the notion of self is reconstructed, alongside consciousness, social relations and relationship to the world (including the environment). This occurs both through the content and substance of what is to be learned under this new paradigm, as well as the technologies and tools to be used to facilitate learning.

The future of work will demand graduates that are able to engage critically at the interface of technology and society, having developed the specialized disciplinary knowledge and skills to translate the conditions of the future in the contexts where they find themselves. This should not occur in a vacuum from other questions such as what it means to be a graduate of and in Africa, and with an African identity. The tendency to focus on the 4IR at the level of technological penetration has yet to grapple with the social change processes that this shift will catalyze, or the contemporary forces that continue to shape existing imperatives in curriculum formulation. Universities are still required to determine priorities within a contextually grounded position that emerges from their location along specific social, geographical and political axes. It is therefore possible to bring the social and technological shifts promised by the 4IR into dialogue with coterminous visions for higher education, including decolonization and social responsibility.

It is important to underscore the positioning of humanity and the physical world at the core of how digital technologies are to be integrated into society. The question of what knowledge is to be valued in the curriculum unearths a second question—that of which vessel, human or machine, will that knowledge be processed, filtered and transferred by and to what end. Ensuring that this industrial revolution harnesses the creative benefits of new technologies in support of a renewed social vision also holds the potential to unlock disciplinary knowledge from its silos, focusing instead on broad social questions and the coordination of different knowledges and sets of skills toward addressing them.

The question of how to deal with the politics of knowledge is universal. What is defined as knowledge and the value attached to the knowledge is related to questions of power and to structural and other constraints. In the words of Paolo Freire, “The starting point for organizing the program content of education or political action must be the *present, existential, concrete* situation, reflecting the aspirations of the people” (2005: 95, emphasis added) without which there is the risk of “preaching in the desert” (2005: 96). Soudien’s evaluation of the history of education in South Africa from colonial times through to the post-apartheid era echoes Freire’s words. Through a detailed historiography, Soudien argues that social difference in the region always “congeals around race” (2011: 35) which has significant ramifications for South African education which remains as the definitional boundary of post-Apartheid South Africa.

Despite the change of government in SA, at the core, the neo/liberal view of the human remains: as a rational, “conscious and deliberative individual whose subjectivity is derived from his or her engagement with the world of meaning in a fully responsible way”. What such a conception fails to address, however as Soudien points out is “the extent to which subjectivity in South Africa is a raced, cultured, gendered, and classed experience” (2010:41). It is a moot point that South Africans’ lived experiences remain racially, geographically and class bound. Policy does not, however appear to acknowledge this in its sublimation of the liberal view of the person and, as a result, fails to “engage with the sociological reality of the everyday” (Soudien, 2010: 41). In this way, rather than transforming society, there has in fact been a deepening of these inherited differentiators. In Soudien’s view, education policy acted as though the South African landscape “was already the society they wanted it to be, a largely middle-class one” (2010: 43). As a result, policymakers did and do not accurately understand the “subjects its curriculum is talking to” and so “[r]ace sits almost tangibly on the post-Apartheid landscape” (2010: 44).

Higher education forms part of the critical dimensions of Sustainable Development Goal (SDG) 4: inclusive and equitable quality education and lifelong learning opportunities for all. Higher education also forms part of the strategies designed to meet

a number of other SDGs, such as those relating to poverty, health, scientific innovation and economic growth (Kioupi and Voulvouli, 2020). While this highlights the important status of higher education in contributing to and shaping changes in society, it represents equally a new pressure exerted on curricula to engage with and articulate sustainability in ways that can be measured to satisfy reporting on SDG progress. The challenge that emerges is how exactly to synthesize and articulate this range of normative ideals and priorities—from sustainable development to nation-building, decolonization and the 4IR—in ways that produce cogent learning experiences within the broad scope of disciplines in the university. That all of this must occur within the existing curriculum policy architecture in SA means that institutions will approach these priorities differently based on their own situational dynamics, having constantly to renegotiate the balance between policy and how it is implemented.

Fundamentally, and as the South African, Ghanaian and Indian examples so aptly illustrate, however, is how curricula and the underlying power and other contextual issues remain unchallenged. Problematically, “discussions about the transformation of higher education after 1994 did not include any significant deliberations about the curriculum and Eurocentric hegemony” (Heleta, 2018: 54). Universities were consequently able to continue as is, until the 2015 #Feesmustfall movement which raised valid issues of equity of access and affordability of higher education. #RhodesMUstFall focused on the slowness of transformation in the SA higher education system including curricula, teaching and learning as well as the lived experiences on campuses. Writing on the US, Macedo refers to the failure to interrogate race, ethics and ideology “the foundation for ... the pedagogy of big lies” (in Heleta, 2018: 55).

There is no question that the continued and relentless pressure on curriculum as the vehicle for change places universities in the invidious position of trying constantly to achieve diverse, often national and institutional imperatives, while at the same time presenting a clear and coherent take on what its “core business” actually entails (Swartz et al., 2019: 568). What is clear is that universities have now, as part of their “core business”, come to adopt a role in addressing inequalities not only in relation to curriculum, but more directly in terms of social and economic inequality, seeking actively to improve the position of students while also enabling “contextually grounded teaching and knowledge production” (Swartz et al., 2019: 578). To go further than this would be to use universities in the reschooling of society, re-locating them as vital social centers through which communities are able to participate in engagements, processes and activities that share the knowledge and innovation created in these institutions. This would not only revitalize the sector, drawing it closer to contextual issues and demands; it would also ensure the full use and reach of universities’ resources. In school contexts in New Zealand, Brown (2020) finds that this can include use of campus meeting and learning spaces, on-site medical facilities, technology and art/design centers. Issues of access in higher education have foregrounded the student experience without acknowledging that universities are contextually embedded institutions that can serve important social functions. Recreating the “ivory tower” into a social resource is one avenue through which this can be achieved.

Conclusion

Curricula must be anchored in reality. For a curriculum to take effect, it must have the requisite clarity and consistency of a sound disciplinary, theoretical or vocational background needed to ensure appropriate academic activity (Winberg and Winberg, 2017). Curricula must be determined by context and pedagogical approaches. The changes needed are multi-layered and require more than top-down policy statements about the imperatives or levers for change faced by the body politic. In addition, speaking through curricula to the contexts of the students means that explicit attention to being an African graduate and to having a Pan-African identity or responsibility is essential. The global challenges encapsulated in the SDGs are arguably key to ensuring graduates able to thrive in these contexts. Clearly some professional orientations lend themselves more readily to this than others, but a requirement of curriculum must surely build the future graduates’ ability to recognize and address their contexts.

If a university is committed to a conscious process of transformation as a result of which students will be prepared for contemporary, supercomplex (Barnett and Coate, 2005) contexts, such change can happen only at the level at which curricula decision-making takes place. This is a crucial point—curricula must not only be cognizant of the students, but about where decisions about curriculum are made and by whom. Curriculum cannot be determined at the level of national policy nor even in most cases at institutional policy-level. As Luckett and Shay have argued, the “expansion of the HE system itself has not led to equity—and we would add—nor is it possible for an HE system to achieve this in isolation from radical transformation of the national and global economic systems that structure it” (2020: 3).

Teaching and learning is not divorced from the social location of a university, its students and faculty. As universities transform and evolve, new societal purposes, functions, challenges, complexities and pressures arise to be grappled with in curriculum planning (Barnett and Coate, 2004; Cross and Atinde, 2015). Reschooling HE in SA is a supercomplex phenomenon, with consensus that curricula are the battlefield in a perpetual site of struggle.

References

- Aina, T.A., 2010. Beyond reforms: the politics of higher education transformation in Africa. *Afr. Stud. Rev.* 53 (1), 21–40.
- Baker, A., 1996. Major disciplinary violations in a junior high school: an exploratory study. *Res. Middle Level Educ. Q.* 19 (3), 1–20.
- Barnett, R., Coate, K., 2004. *EBOOK: Engaging the Curriculum*. McGraw-Hill Education, UK.

- Barnett, R., Coate, K., 2005. *Engaging the Curriculum in Higher Education*. SRHE, Berkshire.
- Bernstein, B., 1996. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Taylor and Francis, London.
- Brown, G.T.L., 2020. Schooling Beyond COVID-19: An Unevenly Distributed Future. *Front. Educ.* 5 (82), 1–4. <https://doi.org/10.3389/educ.2020.00082>. <https://www.frontiersin.org/articles/10.3389/educ.2020.00082/full>. (Accessed 20 May 2022).
- Bruno-Jofré, R., Zaldivar, J.I., 2012. Ivan Illich's late critique of Deschooling Society: "I was largely barking up the wrong tree". *Educ. Theor.* 62 (5), 573–592.
- Cross, M., Atinde, V., 2015. The pedagogy of the marginalized: understanding how historically disadvantaged students negotiate their epistemic access in a diverse university environment. *Rev. Educ. Pedagog. Cult. Stud.* 37 (4), 308–325.
- Freire, P., 2005. *Pedagogy of the Oppressed*. Continuum International Publishing Group Inc., New York.
- Giroux, H.A., 1992. Language, difference, and curriculum theory: beyond the politics of clarity. *Theory Pract.* 31 (3), 219–227.
- Gyamera, G.O., Burke, P.J., 2017. Neoliberalism and curriculum in higher education: a post-colonial analyses. *Teach. High. Educ.* 1–8. <https://doi.org/10.1080/13562517.2017.1414782>. <https://www.tandfonline.com/doi/full/10.1080/13562517.2017.1414782>. (Accessed 10 May 2022).
- Hartley, D., 1997. *Re-schooling Society*. Falmer Press.
- Heleta, S., 2018. Decolonizing knowledge in South Africa: dismantling the 'pedagogy of big lies'. *Ufahamu A J. Afr. Stud.* 40 (2).
- Hicks, O., 2018. Curriculum in higher education: confusion, complexity and currency. *HERDSA Rev. Higher Educ.* 5, 5–30.
- Illich, I., 1970. *Deschooling Society*. Harper and Row, NY.
- Jansen, J.D., 1998. Curriculum reform in South Africa: a critical analysis of outcomes-based education. *Camb. J. Educ.* 28 (3), 321–331.
- Jansen, J., 2019. *Decolonisation in Universities: The Politics of Knowledge*. Wits University Press, Johannesburg.
- Khattar, D., 2019. Adjusted Syllabus: The Political Struggle Over the Teaching Curriculum in the University of Delhi. *The Activist History Review*. <https://activisthistory.com/2019/09/30/adjusted-syllabus-the-political-struggle-over-the-teaching-curriculum-in-the-university-of-delhi/>. (Accessed 27 May 2022).
- Kioupi, V., Voulvoulis, N., 2020. Sustainable development goals (SDGs): assessing the contribution of higher education programmes. *Sustainability* 12 (17), 6701.
- Le Grange, L., 2016. Decolonising the university curriculum. *S. Afr. J. High Educ.* 30 (2), 1–12.
- Luckett, K., Shay, S., 2020. Reframing the curriculum: a transformative approach. *Crit. Stud. Educ.* 61 (1), 5065.
- Lindén, J., Annala, J., Coate, K., 2017. The role of curriculum theory in contemporary higher education research and practice. In: *Theory and Method in Higher Education Research*. Emerald Publishing Limited.
- Malan, S.P.T., 2000. The 'new paradigm' of outcomes-based education in perspective. *J. Fam. Ecol. Consum. Sci.* 28 (1), 22–28.
- Maneejuk, P., Yamaka, W., 2021. The impact of higher education on economic growth in ASEAN-5 countries. *Sustainability* 13 (2), 520.
- Masanya, M.J., 2021. Toward a relevant de-colonized curriculum in South Africa: suggestions for a way forward. *Sage Open* 11 (4), 21582440211052559.
- Menon, K., Castrillón, N., 2019. "Making windows where there were once walls": transformation in higher education curricula. *S. Afr. J. High Educ.* 33 (3), 26–44.
- Nabi, G., Walmsley, A., Liñán, F., Akhtar, I., Neame, C., 2018. Does entrepreneurship education in the first year of higher education develop entrepreneurial intentions? The role of learning and inspiration. *Stud. High Educ.* 43 (3), 452–467.
- Ramrathan, L., 2016. Beyond counting the numbers: shifting higher education transformation into curriculum spaces. *Transform. Higher Educ.* 1 (1).
- Republic of South Africa, 2008. Report of the Ministerial Committee on Transformation and Social Cohesion and the Elimination of Discrimination in Public Higher Education Institutions (Chapter 6), 30 November. Government Printer, Pretoria.
- Sayed, Y., de Kock, T., Motala, S., 2019. Between higher and basic education in South Africa: what does decolonisation mean for teacher education? *Decol. Univ.* 156–180.
- Schwab, K., 2016. *The Fourth Industrial Revolution: What It Means, How to Respond*.
- Shadymanova, J., Amsler, S., 2018. Institutional strategies of higher education reform in Post-Soviet Kyrgyzstan: differentiating to survive between state and market. In: *25 Years of Transformations of Higher Education Systems in Post-Soviet Countries*, pp. 229–257.
- Short, E.C., 1993. *Knowledge and the Educative Functions of a University: Designing the Curriculum of Higher Education*. Paper Presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12–16, 1993).
- Soudien, C., 2010. "What to teach the natives": a historiography of the curriculum dilemma in South Africa. In: *Curriculum Studies in South Africa*. Palgrave Macmillan, New York, pp. 19–49.
- Soudien, C., 2011. The arhythmic pulse of transformation in South African higher education. *Alternation* 18 (2), 15–34.
- Spreen, C.A., Vally, S., 2010. Prospects and pitfalls: a review of post-apartheid education policy research and analysis in South Africa. *Comp. Educ.* 46 (4), 429–448.
- Sukarieh, M., Tannock, S., 2019. Subcontracting academia: alienation, exploitation and disillusionment in the UK overseas Syrian refugee research industry. *Antipode* 51 (2), 664–680.
- Swartz, R., Ivancheva, M., Czerniewicz, L., et al., 2019. Between a rock and a hard place: dilemmas regarding the purpose of public universities in South Africa. *High Educ.* 77, 567–583. <https://doi.org/10.1007/s10734-018-0291-9>.
- Statistics South Africa (Statistics SA), 2021. Quarterly Labour Force Survey (QLFS)—Q3:2021. <https://www.statssa.gov.za/?p=149570>. (Accessed 5 April 2022).
- Winberg, S., Winberg, C., April 2017. Using a social justice approach to decolonize an engineering curriculum. In: *2017 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, pp. 248–254.
- Young, M., 2013. Overcoming the crisis in curriculum theory: a knowledge-based approach. *J. Curric. Stud.* 45, 101–118.

New pathways in education

Dieter Dohmen, FiBS Research Institute for the Economics of Education and Social Affairs, CEO FiBS RILL Research Institute on Lifelong Learning, Berlin, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	57
Future learning pathways	58
Model 1—Tamagotchi: higher education for a good start in employment life	58
Model 2—Jenga: higher education as a solid foundation for further development	59
Model 3—Lego: higher education as a kit	59
Model 4—Transformer: higher education as an opportunity for change	60
Implications on framework conditions	61
References	62

Introduction

Education pathways into and within higher education are still more or less straightforward, by and large: young adults graduating from upper secondary education enroll mainly directly into university, which they leave with a bachelor or master's degree. Depending on the country and its educational tradition, they may enter the employment system with the bachelor's degree and return later or they continue directly to the master program. While university education is the most important pathway into employment in various countries, many countries have a vocational education and training (VET) system, which offers similar life and employment chances than the university education. However, in many countries VET is not yet reality, particularly not in relation to high-qualified jobs. And the attractiveness of VET depends to some extent on the opportunity to move on the university education.

Following these pathways, initial education ends commonly at the early to mid-twenties either with a university degree or a vocational qualification. While further or adult education plays some role already around the globe, it is likely that it will play an even stronger/more important role in the future. Moreover, the profession or occupation learnt during initial education may no longer serve as foundational core area for future pathways when moving up the career ladder to the extent it did in the past. Instead, people may learn different professions or occupations and shift from one to another and/or progress horizontally and/or vertically.¹

While traditional pathways remain important and relevant for large parts of the cohort, the share of individuals following a more flexible pathway is increasing. On the one hand, this concerns those entering the first or second stage of university education, following vocational qualifications and/or parental leave or a lower university qualification. On the other hand, the university system provides additional qualifications, such as, for example, MBAs, which are not necessarily formally recognized within the National Qualification Framework, but by the economy and employers. Vocational streams also offer advanced qualifications, such as, for example, master craftsmen, etc. Eventually, professional specialisations play another key role for those who plan to advance their career in one way or another.

Cultural and traditional structures play a key role in this respect, while the university system is more or less the only route to follow for those with higher educational aspirations in the vast majority of countries, other countries provide different options with more or less comparable employment opportunities. The German speaking countries, Germany (DE), Austria (AT) and Switzerland (CH), for example, are well established and recognized examples for the later. Yet, it should not be overlooked that the dual apprenticeship training system has a history of almost a millennium (Bank et al., 2015).

Finally, the education system itself is changing strongly: While education policy is focusing the formal education sector, e.g. school or university education, new media, such as, for example, LinkedIn and YouTube are becoming part of the learning ecosystem. This same applies to other EdTech companies, whose ties are more or less closely tight to the core education system (see for example: Rivas, 2021). As for many other sectors, the education sector has started – and will proceed – to re-invent itself in the light of ongoing developments: key drivers for future directions are the digital and green transformation which have started to change the economy, the society as well as our private lives already a lot over the last decades, and will continue to do so over the coming decades (Langthaler et al., 2021; UNESCO, 2020, 2018, 2017). Digitalization and automation – as well as climate change, sustainability and green skills – are about to redesign the world of work, its business models and – correspondingly – jobs as bundles of tasks as well as tasks (World Economic Forum, 2016). While early studies on the impact of digitalization on jobs suggested that many jobs will disappear, more recent research indicates that – while still large numbers of jobs will disappear – even

¹Vertical progression means, for example, to enter into higher education after graduation from VET or from (initial) VET to more advanced VET (e.g. master craftsman), while horizontal progression means to enter into another qualification at the same level, e.g. from one VET qualification into another VET qualification or from one BA-programme to another.

more will arise newly, with the major impact of jobs changing in the light of digitalization and/or green skills (e.g. [Cinco, 2021](#); [Orr et al., 2019](#); [Effenberger et al., 2019](#); [Frey and Osborne, 2017, 2013](#)).

Another driver of new pathways of learning is the prolongation of lifetime, which will possibly go hand in hand with an extension of working lifetime: if people will even more frequently turn 90 or even 100 years of age, they may also be supposed to work till they are 70, 75 or even 80 years of age. This would mean that people are working for more than 50 years, maybe even more than 60 years, if starting early to work after leaving school with a lower secondary qualification.

Digitalization does not only change the way we live and work, but also how – and what – we learn as well as the competences needed. It has become even more common to use the internet for any form of learning, particularly in times of the COVID-pandemic, when schools and universities were closed for some time and courses and classes were delivered only at a distance (see for example: [Munoz-Najar et al., 2022](#)). In addition, youtube and other comparable programmes are a natural way to learn (even during leisure time), particularly for young people. In recent years, it has become much easier to live, for example, in Germany and enroll in a study program or a course module at an US university – without moving to the US. Eventually, various platforms, such as LinkedIn etc., act as learning providers and offer content for premium clients. These new forms of learning do not replace the more traditional forms of presence learning in the classroom but complement it ([Costa et al., 2021](#); [Head et al., 2015](#)).

These different factors suggest that the education systems will have to reinvent themselves and complement the traditional structures and learning opportunities, with new forms of learning as well as with opportunities for new learning pathways, incl. more flexible options.

This view on changing learning and education pathways complements the common view on the implications of digitalization on learning, which is focused a lot on technology, infrastructure as well as on learning itself, incl. the opportunities of artificial intelligence.

Future learning pathways

Adding up the present situation, the share of people gaining higher qualifications is growing, time span in (initial) education is much longer than in the past, and the share of adults engaged in learning activities after initial education is growing. Depending on the definition of adult learning, up to 100% of adults are engaged in different types of learning. Yet, different from primary to tertiary education, adult learning is not necessarily focused on (formal) qualifications and certificates, though this also depends on the education system, its set-up and, particularly, its tradition ([Ramirez, 2018](#); [CEDEFOP, 2015](#)).

The trends briefly mentioned at the beginning of the article demand not only higher – or more – skills and competences, but also to keep aligned with quickly changing ongoing developments as well as to update or refresh skills and competences more frequently. In addition, employers may be less focused on formal qualifications, particularly in high-in-demand professions. Each single aspect and even more the combination is likely to lead to a diversification of learning pathways, which is also driven by the role of digitalization in education – and particularly in higher education.

The following models aim to symbolize the different options for the future in a simplified way²:

Model 1—Tamagotchi: higher education for a good start in employment life

Tamagotchi describes a learning model that is focused on the traditional learning path, where learning takes mainly place at the beginning of life (till the early or mid-20s), with limited learning later in live. In this model, knowledge and competences are at a high risk to become outdated, sooner or later. Alike many other goods, knowledge and competences are deteriorating and are to be depreciated. As for the Tamagotchi which “died” if not fed any longer and enough.

In this model, *students* are at the beginning of their careers and university education takes place immediately after graduation from upper secondary education. Students study full-time until the end of the study program after three or five years, depending on whether they are aiming for a Bachelor or Master degree and study as foreseen in the curricula. The purpose of the higher education studies is to acquire a formal university degree and to create a knowledge base that enables a successful transition into employment. Once the graduate left university, short programmes and courses are the models of further learning, usually non-formal.

Model 1 – Tamagotchi (Status quo plus)



This path is a continuation of the school system previously accessed, where secondary and tertiary education is ideally well coordinated, so that the transfer to higher education takes place without major discontinuities. The concept supports academic orientation on the one hand and a certain degree of self-organization and independent learning on the other.

²The following sections are built upon [Orr et al. \(2019\)](#).

In this model, control and coordination, as well as the design of the degree program, are the responsibility of the respective university.

The Tamagotchi model follows the traditional concept of higher education, which is built upon the assumption that the knowledge and skills acquired at the university give learners a future-proof competence profile and enable them to adapt flexibly to future requirements.

It will have to be seen to what extent the present approach, where learning takes place in cohorts which are relatively homogeneous in relation to age and educational background, which is more or less a pre-condition for successful graduation.

Model 2—Jenga: higher education as a solid foundation for further development

In the Jenga model, higher education institutions offer a foundation of knowledge and competences, which are extended by learners through shorter study blocks in their further curricula.

Alike the Tamagotchi model, students enter universities mainly immediately after graduation from upper secondary education. Commonly, students study full-time over a couple of years gaining basic knowledge and skills. The initial university period is shorter than in the Tamagotchi model, which may be more attractive for non-traditional students, who often feel that four to five years of study are too long and/or who preferably are looking for a more application-oriented university education. From the very beginning, the model intends that the learner can and will expand his/her knowledge and competence profile through additional modules in their CVs and after interruptions. One major advantage of this approach though is that it is better targeted to *upskilling* or *side-skilling*. The former may be, but must not be, in the same or a similar direction, while the latter provides an easily accessible pathway to enter into new areas at the same level of qualification. Another option might be to gain a more “broadband” set of competences, which are deepened during a second phase or one or more additional modules.

The Jenga model addresses the concern that traditional, Tamagotchi-style, university studies are not flexible and integrative enough to be future-proof in a very dynamic environment. In an initial study phase, the basics are taught in order to supplement them later. For this model, a decision has to be taken as to what is the foundation that is needed as the basis for the beginning of a career and what should, or could, follow in later phases. Whether these basic principles are general or transversal competences or specific basic knowledge will be answered differently from discipline to discipline and from university to university.



Didactically, a basic study phase supports self-learning and self-organization, where students’ learning and personal development are oriented toward a clear learning path with defined steps and clear results. During the initial phase of study, learning takes place mainly on campus with the support of global communication networks, simulations and augmented reality techniques to extend the learning environment through online experiences. Internships, maker spaces and fablabs, help to establish early connections to the later working world. After successfully completing their studies, the students first leave the university and enter the acquisition phase. The second study phase consists of a number of learning units chosen by the learners themselves, accounting, for example, for the changing competence needs of their job. In this regard, modules can be taken at various training providers, combining on campus and/or online learning in one form or another.

While the formal recognition of the first learning phase is assured as it follows common standards, the recognition of the other learning units depends on how recognition is organized more generally within the tertiary education landscape. For learners there may be the possibility to make results-based agreements with individual higher education institutions covering both the initial learning block and the additional units. This approach is flexible as to whether learning phase 1 and learning phase 2 are integrated into a uniform study program or whether the two phases are assigned independently of each other.

The Jenga model consistently responds to the needs of students and the labor market for a study design that is better able to prepare and respond to new needs from the world of work without replacing the basic structure of a university course of study completely.

Model 3—Lego: higher education as a kit

In the Lego-model, the course of study is not completed as a compact unit as before but consists of individually combined modules of different sizes.

In this model, students are self-reliant and follow an individual, non-standardized learning path that is fully oriented toward their learning needs and interests. For this purpose, various learning units can be freely selected by the students, whether they are online and/or on-campus by different universities and new education providers. The student decides where to enroll for a module and what the topic and content of this module is, based on his or her specific needs and interests. The chain of these

learning units forms their personal course of study. In addition, the frequent alternation between phases of gainful employment and learning is characteristic of this model.



It can be left open whether there is a group of learners who study jointly and who assemble their own individual studies according to their needs or whether individual learners enroll in the modules offered. This will succeed at least for the time being in occupations in which professional qualifications are less in demand than specific skills, e.g. software development. The primary aim is to acquire knowledge and skills that can be used directly for personal purposes.

Thus, the students pull together their own study program individually from different learning units. They may be supported by employers, representatives of occupational groups (who set occupational standards) or (if available) higher education institutions (and other learning providers) who offer modules and/or learning pathways. The learners themselves are not necessarily enrolled at their higher education institution, as this is the case in the traditional system.

Ideally, the learning units account for the practical experience and the informal and non-formal learning of the students, since the time in which the students are not in formal learning environments has a significant influence on their learning behavior.

The recognition of learning modules depends on the general organization of recognition within the higher education landscape and students can enter into a learning agreement with a single institution based on learning outcomes that combines different learning units. Alternatively, it is also possible to combine learning achievements into an academic degree and have them recognized subsequently.

The Lego model closes an important gap in the traditional educational programmes of universities which, due to the dynamics of social change, are not covered by classic Bachelor programmes. The small-scale combination of different courses makes it possible to react to short-term demand and to acquire very individual qualifications.

Model 4—Transformer: higher education as an opportunity for change

The transformer model focusses on students that do not enter directly into the university as school leavers, but once they have already acquired their own professional identity and life experience, which they contribute to their studies. Thus, the time spent in school and in initial education (possibly including higher, but more commonly vocational education) is some time ago. Learners may enter into higher education for the first time or return to it either to acquire new basic knowledge and skills (*side-skilling*) or to improve the level of their formal education (*upskilling*), e.g. because of a need to change careers or the intention to acquire higher qualifications. Learners study relatively intensively over a certain period, which may have a duration of three to five years to complete their higher education studies with the expectation of returning to or re-entering the labor market. Nevertheless, it may also be possible to study for shorter periods, not necessarily heading for full qualifications. The Transformer model meets the demand of learners for possibilities of “transforming” their respective knowledge and qualification profiles.



The central idea of this model is that there should be a possibility for everyone to leave the chosen path of life and to change course. In this respect, another option for a transformation may be to enroll for an advanced vocational qualification, after a university qualification. While the latter may be less common in many countries, it may offer interesting opportunities in other countries – possibly leading to a more balanced approach between vocational and university education.

The major concept behind the Transformer model is to support learning and personal development with clearly defined steps and results. As learners do begin their university education only long after leaving the formal education system, considerable support is needed. At the same time, learners have acquired knowledge, skills and experience through their previous biographies, which they can apply to their learning. Therefore, a careful balance between academic support, guidance and independent learning with individual goals is required.

The coordination as well as the design of the study program is the responsibility of the respective universities, while accounting for the knowledge, competences and experience of the learners. However, it may be an open question for the time-being whether and to what extent previous achievements are credited and recognized. The more is credited and recognized the shorter the study courses can be. In the course of their studies, however, more and more control over their own learning path is shifted to the students, so that after an initial phase of study, the proportion of *self-regulated learning* increases.

Learning takes place mainly on campus with the support of global communication networks, simulations and augmented reality techniques to extend the learning environment through online experiences. Further learning spaces can also be integrated into the learning experience through internships, maker spaces and fablabs. Compatibility with the profession is achieved above all by extending the standard period of study and by online course units.

Change in and advancement of the labor market are essential drivers for the Transformer model, which demand either to expand competences and knowledge or to seek new fields of activity. Ultimately, a basic, occupation-oriented course of study is offered here that meets the needs of an older target group due to its flexible mediation and special didactics.

Implications on framework conditions

It appears reasonable to assume that the different new pathways presented in the previous sections may demand different ecosystems, particularly in relation to funding models as well as student support, since most models so far refer to standard types of full- or part-time studying. While fee systems are often responsive to part-time students, e.g. by adjusting the fee rates, more or less, this is less the case for student support schemes. Yet, the more flexible study pathways and the demand for learning modules are, the more flexible funding models for universities and student support need to be, provided that the state plays a certain role in it. Or: as far as state funding is concerned.

When looking at funding models for universities one should distinguish between state funding to universities and individual student contributions, i.e. commonly fees. As far as we can establish, present public funding schemes for universities are either supply-side and/or different types of demand-led funding. Often, funding still largely refers to inputs, staff and other investment and recurrent expenses, addressing the fact that those have to be covered. In other cases, lump-sum payments or block grants account for certain shares of the costs, demanding from the university to generate additional funds to cover the full costs of operation, including fees. Performance-based funding addresses different types of input and/or output indicators for learning and research etc. to stimulate better performance.

Whatsoever the specific type of funding, they address different learning pathways more or less, commonly less – though this depends also on the specific approach to studying. If supply-side funding is employed covering the total costs of operation and not accounting for different types of students, e.g. part- vs. full-time or incoming and outgoing students etc., it is not very important what students do in practice, i.e. whether they attend courses more or less frequently or not, unless this does not overstretch the courses offered and the number of students per course foreseen.

More or less, the same applies to block grants and performance-based funding, if concrete individual demand is not accounted for. Only if funding depends on concrete demand, e.g. work load by number of courses, hours of course attendance or study credits demanded (e.g. based on ECTS).

From the students' point of view, the same applies, s/he can attend a certain number of courses or run for study credits etc. However, the amount of fees to be paid may matter, if fees are responsive to full-or part-time study, etc.

Yet, the funding regime might be challenged in case of the Lego model, where students enroll on a module or course basis and intend to demand only a small number which is below the official figures for part-time study. S/he may prefer to pay less, while the university may be reluctant to respond to this request. Whether or not the university may be responsive depends on different factors, e.g. the nature and design of public funding and the type of incentives implemented. Incentives to be responsive may be stronger in case that funding follows student and the number of modules or courses³.

A funding model that is fully responsive to very flexible individual study modes has been designed by Dohmen (Dohmen, 2003a,b). It aims to be neutral to different types of individual study modes from the viewpoint of the students, while providing an incentive for universities to enroll as many students as possible, till maximum capacity is reached⁴.

This funding model is based on the number of European Credit Transfer System (ECTS), which have to be acquired for a Bachelor or Master's degree. By the time of drafting, 180 ECTS were the standard norm for a Bachelor degree and another 120 for a Master degree, at least with regard to Germany. Thus, the student's learning account had 180 or 120 ECTS (plus x for some repetition or additional modules, if applicable). If the student enrolled for a module, the corresponding number of ECTS would be deducted (for example, half at the beginning of the module, the rest at the end) and the university received the corresponding amount of money, which should be responsive to full-costs coverage for a specific, standardized number of students per program or module⁵. The funds allocated were recommended to be responsive to the average costs of the programmes, i.e. commonly higher for STEM or medicine than for Arts and humanities, economics etc. This approach meant neutrality in relation to subjects as well as study behavior – allowing also for competition between the departments of the university or between universities. The learning account enabled the student to enroll in one German university for Module A and in another German or foreign university for module B and to spread the study program over as many semesters s/he wanted. In this sense, it would be fully responsive to all study models and learning pathways presented above – it would be particularly applicable to the Lego model, which is the most flexible one.

More precisely, the advantages of this approach were: (a) students are free to decide when and where to enroll, to study (temporarily) full-time and/or (temporarily) part-time and to what extent they wanted to study part-time, (b) the model was neutral and allocative efficient in relation to total public funding and (c) provided an incentive for the university to exploit its full capacity.

³The core discussion on „money follows students“ refers to the number of students at a university, but less to the amount of modules or any other type of workload (Musset, 2012).

⁴It might be worth, though, to point to the issue that overall upfront and step-fixed costs play a strong role in education, while marginal costs are often very low, if not even nil (Dohmen, 2015).

⁵It should be noted that a trade-off exists between a more general and a very specific funding approach and that universities have to recoup the full-costs of each program, module as well as at full-scale, at least, to avoid loss and bankruptcy (Dohmen, 2015). This suggests that certain lump-sum figures should apply, which allow full-cost coverage, even though this may not be fair for any single case.

While the model had been initially designed for the German university system in two states, it was eventually implemented in Switzerland to finance the Universities of Applied Sciences.

One may discuss whether student support schemes should be similarly flexible or not, and to what extent this is necessary to target students from lower socio-economic backgrounds. While full-time studies would have to be linked to a proper student support scheme, independent from age, this is less clear in relation to part-time studies, which may go along with part-time employment. However, as part-time employment does not necessarily result in an income that allows to cover the costs of living, particularly if the student has children or a financially-dependent partner, it seems advisable to make student support schemes fit for part-time students. It can be means-tested and/or be provided as an interest-bearing loan.

A key requirement for student support schemes is that they are less age restricted as presently the case (FIBS/DIE, 2013). Most student support schemes target more or less exclusively young adults, though different age limits apply. The more such support schemes address the Tamagotchi model, the more they target very young students, while they are less responsive to the Jenga or even Lego-model. More precisely, the target group of most support schemes are people till the mid or late twenties, studying full-time. Examples in this respect are the Dutch or some South-European models, but also the German BAfoeG, even though the latter is open to students enrolling for the first time at age 30 for a Bachelor program or even 35 for a master program. These models address equality of opportunity and help students from low-income families to overcome funding bottlenecks of full-time studies.

While these models may even be satisfying the needs of the Jenga- or Transformers model, provided people start studying within certain time or meet the age restrictions frames, they are particularly not supportive to the Lego model, as this is far more flexible.

While some funding schemes as well as student support are open to older students, as for example the Higher Education Contribution Scheme in Australia as well as the funding model in Sweden, they still are not fully responsive to the very flexible study model of Lego (FIBS/DIE, 2013).

References

- Bank, V., Dohmen, D., Frommberger, D., Busemeyer, M.R., 2015. Governance und Finanzierung kooperativer Berufsbildung: die Rolle von privaten Akteuren und Verbänden stärken. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Bonn.
- CEDEFOP, 2015. Job-related Adult Learning and Continuing Vocational Training in Europe: A Statistical Picture. Publications Office, Luxembourg. <https://doi.org/10.2801/392276>. Cedefop research paper; No 48.
- Cinco, S.J., 2021. Companion or substitution? Automation and digitisation in the workplace. J. Entrep. Innov. Emerg. Econ. 7 (2), 263–268. Entrepreneurship Development Institute of India.
- Costa, S.A., Kavouras, I., Cohen, N., Huang, T.T.K., 2021. Moving education online during the COVID-19 pandemic: thinking back and looking ahead. Front. Public Health 9, 751685. <https://doi.org/10.3389/fpubh.2021.751685>.
- Dohmen, D., 2003a. Eckpunkte eines StudienCredit-Modells zur Finanzierung der Hochschulen in Baden-Württemberg. Gutachten im Auftrag der Fraktion Bündnis 90/Die Grünen im Landtag von Baden-Württemberg. FIBS-Forum, Köln, p. 18.
- Dohmen, D., 2003b. Eckpunkte eines Studienkontenmodells zur Finanzierung der Hochschulen im Land Berlin. Gutachten im Auftrag der Senatsverwaltung für Wissenschaft, Forschung und Kultur des Landes Berlin. FIBS-Forum, Köln, p. 19.
- Dohmen, D., 2015. Finanzierung private Akteure in der beruflichen Bildung. In: Bank, V., Busemeyer, M., Dohmen, D., Frommberger, D. (Eds.), Governance und Finanzierung kooperativer Berufsbildung. Die Rolle von privaten Akteuren und Verbänden stärken. GIZ.
- Effenberger, A., Garloff, A., Würzburg, H., 2019. Beschäftigungseffekte der Digitalisierung - Forschungsansätze und Ergebnisse (Bundesministerium für Wirtschaft und Energie. Diskussionspapier, 07) Berlin.
- FIBS/DIE, 2013. Developing the Adult Learning Sector. Lot2: Financing the adult learning sector. http://ill.mon.bg/uploaded_files/financingannex_en.pdf.
- Frey, C.B., Osborne, M.A., 2013. The future of employment: how susceptible are jobs to computerisation? Technol. Forecast. Soc. Change 114, 1–79.
- Frey, C.B., Osborne, M.A., 2017. The future of employment: how susceptible are Jobs to Computerisation? Technol. Forecast. Soc. Change 114, 254–280.
- Head, A., Van Hoeck, M., Garson, D., 2015. Lifelong learning in the digital age: a content analysis of recent research on participation. Clin. Hemorheol. and Microcirc. 20 (2). <https://doi.org/10.5210/fm.v20i2.5857>.
- Langthaler, M., McGrath, S., Ramsarup, P., 2021. Skills for Green and Just Transitions: Reflecting on the Role of Vocational Education and Training for Sustainable Development. ÖFSE Briefing Paper, No. 30. Austrian Foundation for Development Research (ÖFSE).
- Munoz-Najar, A., Sanzana, G., Grace, A., Amer, H., Romani, C., Cristobal, J., Azevedo, J., Wagner, D.P., Akmal, M., 2022. Remote Learning during COVID-19 : Lessons from Today, Principles for Tomorrow. World Bank Group, Washington, D.C.
- Musset, P., 2012. School Choice and Equity: Current Policies in OECD Countries and a Literature Review. Organisation for Economic Co-Operation and Development (OECD).
- Orr, D., Luebcke, M., Schmidt, P., Ebner, M., Wannemacher, K., Ebner, M., Dohmen, D., 2019. AHEAD Internationales Horizon-Scanning: Trendanalyse zu einer Hochschullandschaft in 2030. Hauptbericht der AHEAD-Studie. Arbeitspapier Nr. 42. Hochschulforum Digitalisierung, Berlin. <https://doi.org/10.5281/zenodo.2677655>.
- Ramirez, R., 2018. Kompetenzentwicklung und Weiterbildung am Arbeitsplatz - Eine Analyse der European Skills and Jobs Survey, 23.1, Michaelkirchstraße 17/18, VoREFFI-WB project closing event. Slides, Berlin. available online: https://www.fibs.eu/fileadmin/user_upload/Literatur/Lectures/Praesentation_Kompetenzentwicklung.pdf.
- Rivas, A., UNESCO, 2021. The platformization of education: a framework to map the new directions of hybrid education systems. In: In-Progress Reflection No. 4, on Current and Critical Issues in Curriculum, Learning and Assessment.
- UNESCO, 2017. Education for Sustainable Development Goals. Learning Objectives. UNESCO Publishing, Paris.
- UNESCO, 2018. Building Tomorrow's Digital Skills – what Conclusions Can We Draw from the International Comparative Indicators? Working Papers on Education Policy UNESCO Publishing, Paris.
- UNESCO, 2020. Humanistic Futures of Learning. Perspectives from UNESCO Chairs and UNITWIN Networks. Unesco Publishing, Paris.
- World Economic Forum, 2016. The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution. World Economic Forum, Geneva (Switzerland).

Open access and the unfinished transformation of scholarly communications

Steven Weiland, Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Open access “Takes Flight”	63
Transformation by declaration	64
arXiv and after	65
Faculty responses to OA	66
The global gap	68
How much publishing is enough?	69
Conclusion: the “shock of the new,” again and again	71
References	71

Colleges and universities contribute to the preservation of the past and the invention of the future. The classroom and the library, actual and digital, are still the essential locations for postsecondary education. Indeed, for centuries the library supported knowledge seeking and making as its essence, a place where the world and experience could be studied with the print record that spanned centuries. But it is the rare scientist or scholar who visits the library today except on a screen. Digitalization has meant the near total revision of how information—in journals, books, and other resources—is gained, used, maintained, and circulated. Accordingly, as colleges and universities find new ways to advance their educational missions, there are unanticipated forms and roles for scholarly communications, now understood as a global digital system for the exchange of ideas and practices across the many disciplines of academic inquiry. Indeed, with due regard for the impact of technology on teaching, scholarly communications has been characterized as higher education’s “real digital change agent” to encourage recognition of its likely long-term and profound influence on the university as an institution (Mittell, 2013).

Open access “Takes Flight”

Scientists and scholars have recognized in recent decades exponential growth in academic publishing. There is no exact figure to represent the explosion of knowledge but an authoritative estimate claims that about 3 million academic articles appear annually in about 42,000 journals, more than three-quarters of them in English (Johnson et al., 2018). It is no surprise that there is so much interest in the problem of information overload, though it would be a mistake to think that it is unique to our time (Blair, 2010). Indeed, a taste for knowledge means attention to scholarly communications. It may not be a subject for commencement addresses but a recent film—*Paywall: The Business of Scholarship* (freely available at paywallthemovie.com)—shows the impact of the high cost and thus exclusionary subscription system of academic journals. Access to what scientists and scholars wish to read is essential to the faculty’s productivity and to professional identities.

The faculty and the library have been allies in their interest in the steady expansion of science and scholarship. Alas, scholarly publishers, chiefly a small group of highly profitable corporations who supply most academic journals, all in Europe and North America, have turned out to be antagonists to both, principally in what they can demand in online subscriptions. After all, each journal is unique and likely indispensable to some number of scientists and scholars on campus. The library cannot go down the street in search of a better price. What came to be known late in the last century as the “serials crisis,” or the nearly annual increases in subscription costs, signified the fragile financial status of libraries at institutions of all kinds. The strategy known as the “Big Deal,” devised by publishers, only exacerbated the problem, forcing libraries to buy many more journals than they desired in order to gain a better price on the most necessary. And negotiating on their own with publishers left libraries vulnerable and without the prospect of achieving cost effective ways for the faculty to gain access to the journals they relied on.

There were other widely recognized problems with the system. The traditional subscription structure left struggling and new institutions in poor parts of world with only the most limited access to the journal literature at a time when global academic aspirations were increasing. But a fortuitous gathering in late 2001 (as below) took a fresh look—philosophically, institutionally, and economically—at the conditions of scholarly communications and recent digital publishing experiments. They proposed in “Open Access” (OA) a new format for dramatically improving online availability of what scientists and scholars (and others interested in academic work) want to read. Two decades of proposals, initiatives, and new experiments followed. In 2021 UNESCO made its enthusiastic “Recommendation on Open Science” (endorsed by all 193 member states) and the influential journal *Science* reported that the Open Access movement had “taken flight” (Brainard, 2021). It cited the European Commission’s *Plan S* project, with its broad OA mandate, as primary evidence (coalition-s.org; Else, 2021). Not all nations joined the *Plan*. But the steady growth in OA journals and articles, if not yet in books, signified enough for *Science* to speak confidently of hopes for “upending” decades long traditions in scientific and scholarly publishing. Indeed, in 2017, for the first time, most new papers (as usual the vast majority were in science), were published as OA.

Open Access is now a permanent part of the higher education and scholarly communications landscape boosted, to a degree by the COVID-19 pandemic and the urgency it has meant for access to the latest research on vaccines, drugs, and the social and economic dimensions of the global health crisis. Still, as a 2018 statement by a group including participants in the institutional origins of OA makes plain, there are significant obstacles remaining before it becomes the default format in scholarly communications (Shockley et al., 2018). The response in publishing to the pandemic has had a positive if still uncertain lasting meaning for OA (Lane and Lifshitz-Assaf, 2022). There is recognition of the potential for a pandemic inspired “new normal” (Brainard, 2021). But even as an unfinished project in higher education OA is having a significant impact on publishers, libraries, and research funders, and on scientists and scholars hoping to make the most of the digital transformation of inquiry.

The goal of subsequent sections of this chapter is to add to the success story told by *Science* with recognition of what still requires attention if “open” is indeed to become the common condition in science and scholarship everywhere. The chapter begins with an account of the beginnings of OA including the allied phenomenon of repositories, now viewed as indispensable to the OA movement. But as the advent of Plan S suggests, the long term fate of OA will reflect the uneven participation of authors, the subject of the next part of the chapter. There follows attention to an important obstacle to OA’s success, the limits of its impact in the Global South, particularly Africa. The next part looks at the meanings, in the context of the rise of OA, of the steadily rising volume of scientific and scholarly publishing—an unwelcome development for some critics, including what it says about the allocation of faculty resources for teaching. Thus, the future of OA, as the chapter’s conclusion indicates, is by no means assured, at least as it was originally envisioned, and most of all outside the well resourced Global North. OA may have “taken flight” but its trajectory is still incompletely understood. It is an unfinished project.

Transformation by declaration

Open Access as a movement in scholarly communications began in late 2001 with a meeting in Budapest of sixteen librarians, scholars, foundation and academic leaders, technology professionals, and others. The Budapest Open Access Initiative (BOAI) was sponsored by the Open Society Institute (headquartered in Budapest) which became an important source of support. The brief statement, or “Declaration” (www.budapestopenaccessinitiative.org/read) announcing the goals of OA that came from the unusual meeting is little more than a page but it is rich in ideas about higher education and journal publishing, particularly what makes access to articles so costly and what adjustments must be made in the global scholarly communications system to gain more accessibility and to maintain high standards of inquiry.

Peter Suber, who has headed Harvard University’s *Open Access Project* (cyber.harvard.edu) for many years, participated in the Budapest meeting and (according to others) wrote the Declaration. There is an artful opening sentence: “An old tradition and new technology have converged to make possible an unprecedented public good.” In other words, the habits of scholars who freely contribute their work to journals are to be joined to what is made possible with low cost digital publishing. The first paragraph of the Declaration also reflects the idealism that still characterizes the OA movement. Thus, “Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge” (BOAI, 2002).

The Declaration offers this as the core of its definition of OA: “By ‘open access’ to the [scientific and scholarly] literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.” The last, alas, remains a problem in some parts of the world even twenty years later. Suber (2012) himself now often uses a more compact definition of OA: “Open Access literature is digital, online, free of charge, and free of most copyright and licensing restrictions” (p. 4).

Subsequent meetings (in Berlin and then Bethesda, Maryland) produced allied declarations. But the differences were reconcilable enough for Suber to sometimes refer to the “BBB Definition of OA” (p. 7). His most recent contribution to the OA movement is leadership of the project (in progress in 2022) to identify “Good Practices for University Open-Access Policies” (cyber.harvard.edu/hoap/Good_practices_for_university_open-access_policies). The phrase “good practices” rather than the ubiquitous (in higher education) “best practices,” signifies that on many points in the guide there are differing views among those contributing to it. Thus, “there can be wider agreement on which practices are good than on those which are best.” And “Good practices change as circumstances change” (Schieber and Suber, 2021).

But it is not only “practices” that diverge. At the Budapest meeting participants disagreed about what might be expected of the different interests in OA reform. It appeared that the group was seeking to speak mainly on behalf of scientists and scholars and prospects for improving communications among them and with others. But publishers, librarians, and professional associations (many with major investments in their journals) also had needed attention. At the 15 year anniversary of the BOAI Jean-Claude Guédon (2017), a Canadian scholar of information science who participated in the Budapest meeting, surveyed the OA landscape and found problems faced by the movement as important as its notable achievements. His extensive survey was published (at its website) by the BOAI, signifying its role in realizing the goals over the long term (other recent surveys include Pinfield et al., 2020 and Eve and Gray, 2020).

Guédon restates Suber's now well known image: "Open Access is simply a way to express the cross-fertilization of the very essence of science with new technologies to create the optimal communication system science needs" (p. 2). For Guédon, in the time since the Budapest Declaration a configuration of forces has kept the "optimal" at a distance. Chief among them is the oligarchical behavior of the leading commercial publishers (Lariviere et al., 2015). In effect, seeing academic journals as merely "commodities" they have focused on business models with the highest yield to investors. These came to include the notorious "Big Deals" with academic libraries and, most recently, the system of Article Processing Charges (APCs), a strategy designed to represent the publishers' version of commitment to OA while protecting their financial interests. Publishers have stood by the costly (to authors and their institutional sponsors) and journal-based "Gold OA," leaving a competing way of managing OA, the repository-based and generally free "Green OA," to non-profits like libraries and professional organizations.

In effect, policy making about OA, and scholarly preferences and practices, have organized around such a binary, one that had in fact been anticipated in the Budapest Declaration. But the differences between the two, complicated by both faculty inattention and, for some, loyalty to the traditional journal system and its measures of achievement (the well known "Impact Factor") have left publishers with sizable financial advantages. Academic libraries and repositories have suffered in recent decades along with their host institutions as budgets for higher education, apart from the highest ranking universities, have declined or barely kept pace with inflation.

In Guédon's analysis, genuine global OA, however well timed from a scientific and intellectual point of view—or as a force for the "public good"—fought an admirable battle against international conglomerates. Progress in OA has been substantial. By 2022 the *Directory of Open Access Journals* (DOAJ) had listings of over 17,000 journals (from over 130 countries and in 80 languages) with over 12,000 without APCs. Article records are far beyond seven million. But the competition for legitimacy in OA goes on. APCs, and enough compliance from scientists and scholars with their institutions, have allowed the publishers' business models to continue to forestall any alternate set of values for the dissemination of inquiry. Thus, the present system of journals, particularly those using Gold OA, obscures their profit making habits. "Together, these journals are engaged in an endless system of competition that is supposed to energize research globally, and serve its intrinsic objectives. In actuality, the system acts as the pressure system that, in rich countries, keeps armies of scientists in line" (Guédon, p. 38).

For Guédon, science and scholarship themselves must be at the very center of communications practices displaying the powerful effects of digital culture and the strength of networks. "In no case should economic interests be allowed to interfere with the full potential of a free communicating system designed and destined to help humanity—the whole of humanity—grow knowledge" (p. 38). In effect, OA is presented in this recent account of it as an educational project, as new knowledge is negotiated among scientists, scholars, and students. In the new digital configuration the library, still constrained by commercial publishing, plays its traditional educational role in providing many of the necessary resources for study as well as helping to advance the goals and practices of OA.

arXiv and after

As a cross disciplinary digital initiative Open Access is justifiably recognized as a groundbreaking innovation. But an important predecessor was already a decade old when OA was launched in Budapest. The universally popular (among scientists) website arXiv began in 1991 at the Los Alamos National Laboratory in New Mexico and since 2001 has been at Cornell University. The best known of the pre-print servers, arXiv represents how academic communications changed with the advent of the Internet, and how scientists adapted to new conditions of communications with widespread participation and enthusiasm. The OA site now serves physicists, astronomers, mathematicians, computer scientists, quantitative biologists, statisticians, and others. MedRxiv (founded in 2019 by the famed Cold Spring Harbor Laboratory in New York) aims (with others) OA resources at the health sciences, while additional new disciplinary repositories serve scientists in the prompt circulation of their research in the new format. MedRxiv became indispensable in the years of COVID-19.

In effect, seeking to keep their fields as up-to-date as possible, many users of sites like arXiv and medRxiv now bypass at least temporarily the conventional system of scientific journals to communicate directly with one another. They do so with what were at first called "e-prints" and now "pre-prints," or pre-publication versions of journal essays. For physicists and others pre-prints are as important as what might appear later in formal publications. And a significant number do ultimately appear in journals. But many valuable contributions to physics and other fields do not—their permanent research home is the freely accessible platform. At the site arXiv refers to itself as a "highly automated electronic archive and distribution server for research articles." Such a simple statement may obscure what has been termed "transformative" about the OA pre-print movement in scholarly communications (Chiarelli et al., 2019).

Submitting pre-prints to arXiv and similar sites requires "endorsement" by a disciplinary colleague. The site is "moderated" only to guarantee that submissions are relevant and their research methods responsible. But what appears at the site is not peer reviewed in the traditional sense. Thus, arXiv is an early example of the approach called "publish first, filter later" according to principles of open access. By 2022 arXiv had registered over two million submissions and downloads had passed two billion. The success of arXiv has prompted many institutions to establish similar sites, often called "repositories," a term also used by

countless colleges and universities around the world to name an institutional site, typically managed at the library, for housing digital documents. Ideally it features articles by faculty members but it can also include syllabi and many other kinds of resources of potential interest to scientists, scholars, students, and the public. In 2022 over 6000 OA repositories were registered at OpenDOAR (the Directory of Open Access Repositories; v2.sherpa.ac.uk/pendoar). As a global phenomenon they represented institutions, associations, and organizations of many kinds and different ideas about what is worth collecting, preserving, and disseminating (Arlitch and Grant, 2018).

The communications situation has been slow to change in fields other than scientific ones. For example, until recently in literary and language studies scholars had nothing like arXiv to centralize their efforts and to speed the pace of interactions among authors and readers. They could submit work to the comprehensive Social Science Research Network (now owned by the publishing giant Elsevier) a location used by scholars in over twenty fields. In 2016, however, the Modern Language Association (MLA) opened the Humanities Commons (hcommons.org) to provide a more specialized space for self-archiving, as some like to call it, based in literary studies but including allied fields. It is now based at Michigan State University in the United States. An account of OA by the Brookings Institution (US) reflects hopes for the expansion of the format into all disciplines (Lane and Lifshitz, 2022).

Some might say that speed, a frequently cited feature of OA publishing and pre-print servers in particular, is hardly vital to philosophers, art historians, scholars of French literature and others in the humanities. A former scientific director at arXiv has said that the site “accelerates the pace of science by allowing researchers to get material out there for others to see and build upon right away. It’s the go-to source, the core of an ecosystem, and it fills critical research needs for researchers around the world everyday” (Cornell Chronicle, 2015). The new field of Digital Humanities represents a change in the pace of research and thus the desire for OA. But the humanities generally, and a good share of the social sciences, have different digital needs, reflecting different information practices, than do the fields represented at arXiv. It took some time in the humanities to find a suitable structure. A distinguishing feature of the Humanities Commons is that it is a non-profit OA resource and social networking platform run by scholars themselves. It is utterly unlike the venture capital supported and for profit repositories and social media platforms academia.edu and [ResearchGate](https://ResearchGate.net).

Whatever its role in launching a revolution in how a field can organize its communications, arXiv demonstrates what more might be done. For its founder, physicist Paul Ginsparg—still an important voice in the transformation of scholarly communications—there was a time when faculty work was part of the pioneering spirit of the Internet, new developments being “naturally adapted” to research. In 2006 Ginsparg thought it likely that within a decade “more research communities will join some form of global unified archive system without the current partitioning and access restrictions familiar from the paper medium, for the simple reason that it is the best way to communicate knowledge and hence to create new knowledge” (p. 9607). But, for Ginsparg, to the degree that commerce and entertainment came to dominate the Web, academic interests (like OA) were lagging in innovation. Oddly enough, the unadorned arXiv in 2022 looks hopelessly out-of-date as an Internet resource. Many first time visitors are surprised and pleased by its simplicity or disciplined functionality. But, as a pioneer in OA it still stands for an approach to research and communications that couldn’t be more timely or user friendly. Hence its utility as a model during the COVID-19 pandemic and, for some observers, far beyond.

Nonetheless there are those who wonder about whether the highly decentralized or even fragmented institutional repository scene isn’t an obstacle to realizing the global goals of open access. Many authors are reluctant to participate, preferring, despite the cost, to have the advantages in visibility and reputation of the Gold OA system with its APCs. Mandates for OA from government agencies and private funders are nominally neutral about where work appears. But for many scientists and scholars the stakes are too high to rely on repositories, still awaiting interoperability and other features of the traditional system of dissemination offering inclusion in leading indices and thus visibility, or discovery (in the vocabulary of the library) and what follows in citations and then academic rewards. Thus, while the disciplinary cultures served by arXiv can capitalize on its reach, reflecting high levels of participation, contributors to the Humanities Commons must await confidence in discovery and impact. The problems are compounded by competition from institutional repositories (at postsecondary institutions) which offer even weaker prospects for both. Important early supporters of such repositories see them facing an “existential crisis” in the years ahead (Poynder, 2016). Some academic librarians favor a national level repository but do not see their colleagues as willing enough to collaborate to make the most of “network effects” (Arlitch and Grant, 2018). The same might be said for other problems and opportunities facing higher education.

Faculty responses to OA

Judging from its high rates of submissions and downloads (reported daily at its site) arXiv is part of the workday of a sizable number if not most physicists around the world. But for scholars in many fields there is not yet the habit of capitalizing on OA in any of its formats. As the latest Ithaka S+R faculty survey showed there is increasing enthusiasm for the conversion of the publishing system to OA reflecting the advantages survey respondents find as readers. But, thinking as authors about what they face at present in the academic rewards system, only 40% of respondents said that they pay any attention to OA in considering where to submit their work (Blankstein and Wolff-Eisenberg, 2019). How much the pandemic may change such views is something that future studies will record.

Open access has its faculty champions, those hoping to make “open” central to the evolving academic vocations and postsecondary educational professionalism. Martin Weller has been associated for many years with Britain’s Open University, the pioneering public institution of its kind. He has impressive experience in the design and assessment of online teaching and learning, particularly according to emerging norms in “open” education, or online degree programs organized for the greatest possible access. Weller moved beyond the instructional dimension of electronic change in education in his book of 2011, *The Digital Scholar: How Technology is Transforming Scholarly Practice*. He embraces the newest technologies and adapts innovations for broadening communications in order to build new communities of inquiry. Weller focuses on three conditions of “transformative practice” in scholarship and science: It is digital, networked, and open.

For Weller, a “digital scholar” is above all a social one, giving priority in his or her work to the professional connectivity provided by the Internet, and, in the spirit of OA, to “participation” by anyone with views about the topics in one or another field of science and scholarship. Expert knowledge will certainly have a role in participation in many subjects. But participation itself, and thus the democratization of knowledge, remains a core principle in the world of the digital scholar.

In the decade since *The Digital Scholar* OA has continued to expand its reach as is evident in the record of the DOAJ and OpenDOAR. While it is still far from having fully altered the influence of the major commercial publishers, convinced all scholars and scientists of its virtues, and helped to close the gap between rich and poor countries in scholarly communications, OA has indeed, as *Science* asserts, “taken flight.” And there has been a steady increase in citation rates for OA articles. By 2018 they were being cited 18% more often than paywalled ones (Piwowar et al., 2018).

Still, according to Guéron (and others) scholars and scientists themselves have played only minor roles in the conceptual and organizational development of OA. Plainly researchers are at the heart of scholarly communications. They provide their articles and books, and evaluate the work of others, as part of the publishing process. They communicate with one another but it is in a communications system they do not control. For many that means they pay little attention to its structure and debates, including the steps that have been taken among publishers, libraries, and funders to respond to the BOAI of 2002 (and later declarations). “The shaping of the future of science communication was not always in the hands of those who needed it most to communicate and work” (p. 8). Guéron believes that whatever the system does to support the “grand conversations” of the disciplines, as has long been said, it also “acts like a powerful tool to manage careers and reputations” often indifferent to the wishes of participants.

Thus, scientists and scholars are still prompted to focus on the misleading “impact factor” as the essential metric, the competition among journals for prestige, and the academic rewards to individuals that flow from publishing in the highest ranked outlets. OA is often perceived as lacking in the status most want for career advancement. And, alas, there is cooperation with the ultra-competitive publishing system by academic leaders, including colleagues serving in evaluation roles for promotion, tenure, and annual review who defer to the most conventional standards of research productivity. Moreover, despite the number of repositories there has been only limited faculty participation even with institutional and government mandates (Tmava and Miksa, 2017). Guéron concludes that “most researchers, although deeply dependent upon the system of scientific communication, are probably among the people that understand it most superficially—certainly more superficially than librarians and publishers” (p. 29).

True enough, the scale of contributions to OA megajournals like PLoS ONE is impressive even with declines in submissions (Brainard, 2019). But recent studies of faculty attitudes toward OA in all formats—when they are even aware of it—confirm the Ithaka S + R findings and demonstrate that the movement will require perseverance if it is to take hold broadly among scientists and scholars (Fitzgerald and Jiang, 2020; Dalton et al., 2020). What will convince more faculty to publish in OA journals and to submit their work to repositories? All authors want their work read and used. And the early record for the transition to OA publishing is promising as journals themselves, from the influential commercial giants (building the case for APCs) to the newest entrants in scholarly communications support the format. Academic libraries seek to educate the faculty, largely via their websites, about obligations (in mandates) and opportunities across the different forms of OA. Postsecondary institutions have roles, particularly in the making of policies for tenure and promotion and other academic rewards. But these rarely address the changes introduced by digital publishing, including the growth of OA.

For learned societies the new terrain of scholarly communications has prompted both enthusiasm for OA and anxiety about its impact on association sponsored journals (Wise and Estelle, 2019; Harrington, 2019). Thus, new organizations, like Transitioning Society Publications to Open Access (tspoa.org), are guiding stakeholders toward a level of transformation that suitably represents (in the vocabulary of the Budapest Declaration) the old and the new. Inevitably Plan S, with its far reaching mandate and urgency, has complicated the transition and prompted some learned societies (like the American Historical Association) to protest its impact on publishing in the humanities. In effect, scholars and scientists having to respond to the phenomenon of OA must navigate a system being reconfigured by interests that are not always complementary.

Journals, of course, do not exhaust the publishing plans of scholars. Still, this chapter, as does most work on OA, features its impact on serial publication, particularly in the sciences. The Directory of Open Access Books (www.doabooks.org) shows steady growth in recent years with (by 2022) over 51,000 titles from over 654 publishers. A study of OA books showed that authors are uniformly supportive of them for the visibility they give their work. The format lacks only greater institutional support and publisher’s receptivity to make it routine (Pyne et al., 2019).

There are now underway allied projects seeking to aid libraries and university presses in OA initiatives making books in the humanities and social sciences discoverable and useful. *Knowledge Unlatched* (knowledgeunlatched.org) uses funds contributed by libraries to bring OA monographs into collections in many nations. *Toward an Open Monograph Ecosystem* (TOME; www.openmonographs.org) is a collaborative book subsidy project of the Association of American Universities, Association of Research Libraries, and Association of University Presses. And the Andrew Mellon Foundation supports the development of platforms for multi-media OA monographs at US university presses (e.g., [Kasprzak and Smyre, 2017](#)). An account of new forms of inquiry in the humanities reflecting such projects displays satisfaction in OA for what it can yield for the scholarly vitality of academic authors ([Scarborough, 2021](#)). Similarly, a recent report on *The State of Open Monographs* urging the expansion of the OA landscape foresees the innovation's fruitful influence on the form of the monograph and the "ironclad" role of the traditional book in the academic reward system ([Grimme et al., 2019](#)).

The global gap

As *Science* suggests, a satisfying story can now be told about Open Access. But how far does it extend beyond European and North American research toward making "bibliodiversity" an essential feature of OA ([Berger, 2021](#))? Latin America offers a promising recent history in OA. In contrast to the European and North American deference to the major publishing corporations organizations in Latin America have built a non-commercial OA infrastructure. Called an OA "ecosystem," five publishing platforms provide software applications, interoperability, visibility, and discovery for Latin American research output ([Lopez and Garcia, 2019](#)). The group has opposed Plan S and the prospect of APCs, registering its OA experience: "When opening access is decontextualized from its historical and political roots, it has the potential to become as exploitative and oppressive as the system that it seeks to replace."

As part of a recent forum on global trends in OA, Vrushali Dandawate, an Indian librarian and also the DOAJ Ambassador for India and Pakistan, reported that while there are over 3600 Asian journals listed in the Directory and over 1200 Asian repositories registered in OpenDOAR, government and faculty indifference, particularly to mandates, has limited their utility ([Harris, 2021](#)). That has made the work of individual OA advocates crucial but the effort is fragmented, with limited regional interaction and collaboration. Indeed, only a small number of countries in the Global South (four in the classification presented by Prost and Schöpfel in [Herb and Schöpfel, 2018](#)) produce enough research or support enough repositories to make likely partners in international OA projects.

Scholars have for two decades lamented the great gap between OA resources and practices in the Global North and Global South. Early in this century the World Bank and the World Health Organization mounted projects supplying online access to journals, however much the lack of suitable digital infrastructure limits their uses. And even when private philanthropy helps with OA in Africa it can display what can be described as "neocolonial" habits. Thus, Florence [Piron \(2018\)](#) reports on a visit to the library at a Sub-Saharan university which had been awarded support for OA by an international organization. But the library was denied the choice of the journals it could access, being limited to what the project specified. And, as Piron shows, the local infrastructure was hardly improved: "The program had installed two computers providing access to its journals. However, to prevent unauthorized access, these computers were protected by passwords that changed every month! At the time of my visit, the librarians had lost track of these changes and, as a result, [the] new computers were unusable" (p. 122).

Other observers of the operations of OA in Africa see a pattern of neocolonialism, or the "corporate capture" of research with far reaching consequences (e.g., [Meagher, 2021](#)). Cameroonian scientist Thomas Nkoudou, following Dandawate in the online forum, acknowledged that the OA situation in Africa was improving but he was precise in specifying practical problems: "The challenges open access is facing in Africa are physical barriers with electricity issues, lack of access to equipment, internet connectivity, and the lack of infrastructure. We also have cognitive barriers, because a lot of people still lack skills related to open access, even sometimes librarians, and people are still alienated by the Impact Factor and university ranking" ([Harris, 2021](#)).

For Nkoudou two metaphors represent the fate of OA in the Global South. One is familiar. Thus, Europe and North America "sit at the center of the system, and countries of the Global South, are placed at the periphery." They are allocated resources and services accordingly, more often made objects of "knowledge philanthropy" reflecting how Western institutions and corporations understand their needs rather than being invited to collaborate in attention to local interests ([Sengupta, 2021](#)). Stating what he sees more dramatically, [Nkoudou \(2020\)](#) offers another binary, based on the classical Greek *pharmakon*, a medical ritual that sometimes cures and other times acts as a poison, (pp. 25–26). He fears widespread "alienation" among scientists and scholars from local cognitive traditions and research habits.

The pace of digital change puts fresh pressure on all stakeholders in the OA movement. The change is at a scale that requires cultural as well as digital transformation. In her account of the interactions of the two, long time publishing consultant Lettie [Conrad \(2021\)](#) focuses on the rapidly evolving OA activities of major publishers who "have been metamorphosing before our eyes, shedding old ways of doing business and taking up broad initiatives to reinvent what publishers do and why they matter." But in recent years, scholars of the Global South, including those who live and work there (like Nkoudou), have asked if there has been sufficient change in attitude to go with the redesign for OA of the international publishing system. According to Conrad "digital innovations are more of a human science than a computer science." She welcomes the claim from one of the large German

journal publishers: “The consumer and their information experience is at the heart of all decisions.” Such a stance presumably protects publishers, as Conrad puts it, from a “disconnect between the overarching [industry wide] narrative and the reality of what is being built.”

Looking back at the Budapest Declaration, and the many similar expressions of OA solidarity that have followed, including most recently from heretofore resistant but still increasingly profitable publishers, Nkoudou finds such a “disconnect” between “good intentions” and the “realities” of OA adoption in Africa. Since 2002 “OA has evolved positively but has also been deeply perverted,” largely by the introduction of APCs in Gold OA, as a “vehicle of continued exclusion” (pp. 28–29). They are far beyond the reach of African authors and institutions, even when (at times) partial waivers are applied. Conveniently enough for publishers, APCs “consolidate their market strategy” at the expense of attention to ideals of truly open knowledge dissemination. Worse yet, the current system, with its ill effects on African science and learning, yields “epistemic alienation,” or “the distortion of one’s native way of thinking, and of seeing and speaking of one’s own reality” (p. 32).

Nkoudou has practical and philosophical steps in mind. For Africans, “Failing to embrace OA would mean missing a significant opportunity to improve dissemination, visibility, and the impact of research findings from the African continent” (p. 35). There is more to the dissemination of knowledge than the traditional, published journal article (and Gold OA with its high APCs). Instead, Green OA, as in the uses of repositories, can support the rich “gray literature” like theses, dissertations, and research reports that are available in Northern databases (see also [Adam and Kaur, 2021](#)). There has been some progress. *African Journals Online* ([ajol.info](#)) now shows 306 online journals among the 583 at its database. But more than half of the nations with journals at the aggregator contribute three or fewer. Can they increase OA participation? Perhaps the best path, barring transformational reform among the major publishers, is adopting the organization of OA in Latin America, where scholars, scientists, and institutions themselves, with support from national governments, design and manage OA initiatives.

Anthropologist Mahmood Mandami, who has spent more than a decade at Makerere University in Uganda, sees the significance of OA in Africa historically. Colonial powers believed that the only valid learning was created in the West, needing only distribution to other regions. “It is time that we put a stop to this idea that Makerere and other African universities were created as centers for knowledge dissemination and not knowledge production” ([Agaba, 2022](#)). OA makes all nations potential distributors of knowledge. And the conditions of digital publishing, particularly OA, means opportunities in fields often neglected in graduate education and research in the developing world. According to Mandami “Our threat is a leadership which is infatuated with technology and which thinks that you do not need the humanities and social sciences, that all you need are engineers and doctors. We need to put aside the question of narrow specialization [because] it leads to a dead end [by] focusing on means and methods and not on goals.” Of course, OA also contributes to information overload. Speaking for scholarly communities preparing for participation in the emerging open system Mandami reminds them that “Not everything that anybody says or thinks should be published.” While it builds local OA participation, African scholars and institutions should be providing for debate as well as distribution.

To be sure, the problem remains of increasing recognition from scientists and scholars of the value of Green OA, including what uses of it can contribute to a more inclusive science. Seen this way, OA is a resource in the “quest for epistemic freedom or the right to think, to theorize, and to interpret the world; to develop one’s own methodologies, and to write from where one is located, unencumbered by Eurocentrism” (Nkoudou, p. 36; see also [Herb and Schöpfel \(2018\)](#) for a timely collection of articles on OA written in this spirit).

John Willinsky (2006) in his influential early account of OA, also acknowledged efforts around the time of the Budapest Declaration to improve the provision of OA in Africa. He spoke for “a far more distributed journal culture to spread through the academic community, against an otherwise centralized model” (p. 108). But he also anticipated Nkoudou and others looking at OA in the Global South today in recognizing “risks” in simply delivering more and more Western OA research. Willinsky asked: “Can technology indeed help rewire not only older patterns in the circulation of knowledge but the spread of education and the growth of research capacities?” (p. 109). If it can, then a strong local educational system can produce knowledge representing its own interests and methods. Willinsky urges recognition of what scholars in the Global North can gain from visibility (via OA) of the work of colleagues in Africa and elsewhere (see also [Bwalya and Akandelwa, 2021](#)). In *The Access Principle* (2006) he cites the Sri Lankan literature professor Suresh [Canagarajah \(2002\)](#) who says that “an engagement with local knowledge from peripheral contexts can help enrich, expand, and reconstruct mainstream discourses and knowledge” (p. 107).

How much publishing is enough?

However much the major publishing companies are leaving Open Access as unfinished in some parts of the world they have responded robustly to OA within the Western system, largely by becoming advocates for it, if on their own terms. Thus, national and local mandates are matched by new business models (e.g., “transformative agreements”), largely leaving in place the high profit margins the publishers have enjoyed for decades ([Hinchliffe, 2020](#)).

Despite the geographical “disconnect,” the reputation of OA is stronger now than ever. It is helped along by the fact that the volume of published work, from excellent to mediocre and worse, continues to increase at a high rate ([Génova et al., 2016](#)). Of course, a good deal of that reflects the ease in establishing electronic journals, at least in comparison to the demands of traditional print.

Inevitably, critics of academic publishing, finding too much of it, take the values of “publish or perish” and the expansion in opportunities it has prompted as the primary causes. Philip Altbach (2019), long an influential scholar of higher education, recently offered (with Hans de Wit) a brief but powerful statement of the problem, claiming that “the publishing system is out of control.” But even in the early days of the application of digital technology to scholarship there were those who were worried about its role in unnecessary “productivity.” The great increase in journals now “born online” (most are OA), has meant publishing opportunities that did not exist two decades ago. Many are worthwhile—there is always room for new ways to represent knowledge. But the journals and articles are piling up and it is hard to believe that it is all necessary—apart from the professional needs of authors. And, alas, the problem of “predatory publishing,” or the schemes of spurious journals to gain APCs from inexperienced authors, can add to skepticism about OA although the DOAJ itself claims that the problem is not as extensive as some assert (Beall, 2012; Olijhoek and Tennant, 2018).

By a decade into this century there were expressions of dismay about the publishing “avalanche” or “glut” and what it signified. “While brilliant and progressive research continues apace here and there,” academic observers say, “the amount of redundant, inconsequential, and outright poor research has swelled in recent decades” (Bauerlein et al., 2010). The best evidence for that, these critics note, is that fewer than half of the scientific articles published are cited even once. Thus, productivity may be high, defined as the number of articles published, but impact and utility are low. Indeed, the “productivity climate” has this effect according to literary studies and education scholar Mark Bauerlein and his co-authors: “Aspiring researchers are turned into publish-or-perish entrepreneurs, often becoming more or less cynical about the higher ideals of the pursuit of knowledge.” There are consequences for people and for budgets, whatever the low cost of digital and OA publishing: “We need to understand that there is such a thing as over publication, and that pushing thousands of researchers to issue mediocre, forgettable arguments and findings is a terrible misuse of human, as well fiscal, capital.”

There are legitimate questions about what peer review now means in a time of mega journals, the pressures on finding qualified and willing reviewers, and the steady growth of preprint servers and repositories. Can an article that appears only on preprint server count as “published”? Likely not for most tenure and promotion communities. But for many contributors to arXiv appearing there is enough, with the next steps in the formal journal based publication process unnecessary. Indeed, a report from the International Science Council finds in “The Normalization of Preprints” a sign of recognition that conventional peer review no longer carries primary authority in establishing research legitimacy (Drury, 2022). The preprint represents the advantages of the digital system to authors and readers alike, and to the benefits of timely uses of research. According to the ISC universities should adapt the academic reward system to acknowledge the change and new habits of inquiry. With work from a reputable OA source already in the hands (or on the screens) of readers months or even years before it would have reached them in a traditional journal it is time to go on to the next article. Of course many scientists and scholars still prefer the traditional system, from habit or the belief that it contributes to the rigor and durability of research, and to any case made for career advancement.

The sizable annual increases in the number of scientific and scholarly articles appearing in formats of all kinds can obscure questions being asked about research and publishing itself. Thus, an account of the rate of publishing in sociology makes plain that the decline in opportunities for tenure track faculty work has meant extraordinary pressure on young scholars to publish as much (and as quickly) as possible (Warren, 2019). Has the result been a period of exceptional research productivity or does it simply represent a “bubble” or “glut” reflecting the new professional conditions of academic work including the proliferation of OA journals? (Trimble et al., 2010; Génova et al., 2016).

Philosopher of education Nicholas Burbules asked in 2020 if scholars and institutions have been neglectful and even complacent about what has happened in recent decades. He recognizes the several good reasons to write and publish, including extending human knowledge, communicating with colleagues, and even fortifying individual identities as scientists and scholars. But the effects of the “publishing treadmill” should prompt serious rethinking about the system, including what Open Access has contributed to the “explosive availability of online publishing opportunities” (p. 662). The combination of the incentive system for authors and publishers alike, if reflecting different motives, and the digital reorganization of publishing opportunities, have left us with more and more unread work. As decentralized and complicated as scientific and scholarly publishing is, however, with what the OA movement has added to it, Burbules finds it hard to see beyond academic humility as a way to slow or stop the treadmill. Of course, as Burbules acknowledges, new habits of academic self-promotion via social media are driving things in the opposite direction, reflecting the overcrowded digital marketplace. More modesty, including self-imposed limits on publishing are what chemist David Muddiman (2016) believes will yield higher quality work and more satisfying careers. The goal would be to abandon the current habit of seeking a place for the “minimal publishable unit” (the MPU) and focusing instead on the “highest publishable unit” (HPU).

For historian of education David Labaree (2018) no external rationale is necessary for explaining the expansion of academic publishing in our time. It is an essential feature of faculty work. In a recent account of it he acknowledges that in a surplus of writing only a small number of articles are “worth reading” or “make a substantive contribution to their field.” And Labaree says that “It’s sad to see academic scholarship fall into a state where the accumulation of lines on a CV matters more than producing quality work.” He appears to agree with Altbach and De Wit. But no, according to this speculative formula: “Maybe it’s worth tolerating a high level of dross in the effort to produce scholarly gold.” Thus, for Labaree and seemingly most scientists and scholars the “incentive system for faculty publication still provides net benefits for both the academy and society.”

To the degree that the OA movement, with its intellectual and social idealism, has had a role in the steady growth of publishing, it is now part of a system increasingly accountable for its seemingly limitless expansion. Paradoxical as it may sound, looking ahead in 2006 Willinsky proposed that OA advocates have roles in skepticism about the looming outburst in journal publishing. "This overstated critique only raises the need for a more fundamental calling to account of higher education" (p. 137). For some scholars that is what is happening. Indeed, recent arguments for more "public engagement" by scientists and scholars, including seeking out OA opportunities, reflects disappointment with the conventional academic preoccupation with publishing as primarily a resource for professional advancement (Hoffman, 2021). And such arguments remind us that access should mean more than availability. As audiences expand, in adjacent disciplines and among non-academic readers, scientists and scholars can strive to be understood in ways that educate without excessive specialization. In effect, OA can reflect a form of academic "generosity," academic writing being an invitation to learning for all as well as a medium of professional knowledge making (Fitzpatrick, 2019).

Conclusion: the "shock of the new," again and again

A decade after the Budapest Declaration its author acknowledged that the changes in scholarly communications it urged had steadily gained ground. Open Access could be seen as "a solution to serious problems." But there was still "the need for the sheer passage of time to overcome the shock of the new" (Suber, 2012, p. 167). The example of arXiv had set the stage for OA principles. In the 1990s the preprint offered an alternative image of publishing. It is still thriving, with allied innovations in OA (Drury, 2022). Indeed, the COVID-19 crisis demonstrated how a purposeful if informal global alliance (or "scientific globalism") could demonstrate the virtues of OA in a scientific emergency (Horbach, 2020; Lee and Haupt, 2021).

Still, a new development, the so-called "transformative agreements" between the publishing oligarchs and major academic libraries, can prompt the view, among competing ones, that the lessons of OA are all about business models and sustaining the traditional system (Hinchliffe, 2020; Pool, 2021). But for the Organization for Economic Co-operation and Development's Toby Green (2019), troubled as he is by the success of Sci-Hub, the notorious outlaw dissemination service (Bohannon, 2016), OA will never solve the problems of cost, at least in the Gold version many academic authors accept. What is needed is the "transformation" (as Green uses the term) of the entire system of scholarly communication. Pre-prints and repositories would have primary roles and major journals maintained only to showcase a very small percentage of the best work. "Subscribe to Open," launched in 2019, is another OA experiment now underway but initial results are hardly transformational. By 2021 it had not gained participation from any of the major corporate publishers (subscribetoopencommunity.org; see also Harrington, 2021).

"Transformation," as in the title of this chapter, can have yet another meaning. The OED specifies "changes in form, shape, or appearance." For an influential group of scholars of journals in the decades ahead there will be new rounds of the "shock of the new," as experience and technology reveal fresh ways to design and present research (Herman et al., 2020). From text enhanced with media, to the megajournal, to the uses of "plain language" summaries, to new forms of social media based professional communications, to public formats for peer review, and more, authors and readers will encounter many tests of their scientific and scholarly habits. But so too will they face organizational reforms also intended to be transformational. Technological change in higher education, as in other domains, is often initiated with such a claim (e.g., Floridi, 2014).

Communications reflects the goals and activities of authors, from the publishing formats we know, to those we are experimenting with, to the unknown innovations that lie ahead, all aimed at establishing and maintaining relations with others dedicated to inquiry, colleagues and students alike. These are, as Suber and Willinsky insist, activities that should now be guided by "open" as the principle animating faculty work. Designating OA as an "unfinished" transformational project means that it is necessary to view it against its original goals, questions of scholarly communications it directs us to, and the evolving conditions of knowledge dissemination. But we can likely count on the continuing withdrawal of paywalls for science and scholarship as higher education welcomes new expressive opportunities.

References

- Adam, U., Kaur, K., 2021. Institutional repositories in Africa: regaining direction. *Inf. Dev.* 1–13.
- Agaba, J., 2022. Mamdani Talks About His Research Legacy and Work at Makerere. *University World News: Africa Edition*. March. Access. <https://www.universityworldnews.com/post.php?story=20220329112608625>.
- Altbach, P., De Wit, H., 2019. Too much academic research is being published. *Int. Higher Educ.* 96 (Winter), 2–4.
- Arlitch, K., Grant, C., 2018. Why so many repositories? Examining the limitations and possibilities of the institutional repositories landscape. *J. Libr. Adm.* 58, 264–281.
- Bauerlein, M., Gad-el-Hak, M., Grody, W., McKelvey, B., Tribble, S., 2010. We Must Stop the Avalanche of Low Quality Research. *Chronicle of Higher Education*. June 13.
- Beall, J., 2012. Predatory publishers are corrupting open access. *Nature* 489, 179.
- Berger, M., 2021. Bibliodiversity at the center: decolonizing open access. *Dev. Change* 52 (2), 383–404.
- Blair, A., 2010. *Too Much to Know: managing Scholarly Information Before the Modern Age*. Yale University Press, New Haven, CT.
- Blankstein, M., Wolff-Eisenberg, C., 2019. *Ithaka S+R Faculty Survey 2018*. Ithaka S+R, New York.

- Bohannon, J., 2016. Who's downloading pirated papers? Everyone. *Science* 352 (6285), 508–512.
- Brainard, J., 2019. Open-access megajournals lose momentum as the publishing model matures. *Science* 365 (6458), 1067–1068.
- Brainard, J., 2021a. Open Access takes flight. *Science* 371 (6524), 16–20.
- Brainard, J., 2021b. No revolution: COVID-19 boosted open access but pre-prints are only a fraction of pandemic papers. *Science* 373 (6560), 1182–1183.
- Bwalya, T., Akakandelwa, A., 2021. Open access and the future of scholarly communication in sub-saharan Africa. In: Jain, et al. (Eds.), *Open Access Implications for Sustainable Social, Political, and Economic Development*, pp. 205–220.
- Canagarajah, A.S., 2002. *The Geopolitics of Academic Writing*. University of Pittsburgh Press, Pittsburgh, PA.
- Chiarelli, A., Johnson, R., Pinfield, S., Richens, E., 2019. Accelerating Scholarly Communication: The Transformative Role of Preprints. Zenodo (CERN). <https://zenodo.org/record/3357727#.Y03Wt0wpC70>.
- Conrad, L., 2021. Digital Transformation Requires Cultural Evolution. *Scholarly Kitchen*. July 14. Access: <https://scholarlykitchen.sspnet.org/2021/07/14/>.
- Cornell Chronicle, 2015. Research Repository arXiv Hits 1 Million Submissions. January 15. Access: <https://news.cornell.edu/stories/2015/01/research-repository-arxiv-hits-1-million-submissions>.
- Dalton, E., Tenopir, C., Bjork, B.-C., 2020. Attitudes of North American academics toward open access scholarly journals. *Portal Libr. Acad.* 20 (1), 73–1000.
- Drury, L., 2022. The Normalization of Preprints. *International Science Council*, Paris, France. Access: <https://council.science/publications/normalization-preprints/>.
- Else, H., 2021. A guide to Plan S: the open-access initiative shaking up science publishing. *Nature*. April 8. <https://doi.org/10.1038/d41586-021-00883-6>.
- Eve, M., Gray, J. (Eds.), 2020. *Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access*. MIT Press, Cambridge, MA.
- Fitzgerald, S., Jiang, Z., 2020. Scholarly publishing at a crossroads: scholarly perspectives on open access. *Innovat. High. Educ.* 45, 457–469.
- Fitzpatrick, K., 2019. *Generous Thinking: A Radical Approach to Saving the University*. Johns Hopkins University Press, Baltimore, MD.
- Floridi, L., 2014. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press, Oxford.
- Génova, G., Astudillo, H., Fraga, A., 2016. The scientometric bubble considered harmful. *Sci. Eng. Ethics* 22, 227–235.
- Green, T., 2019. Is Open Access Affordable? Why Current Models Do Not Work and Why We Need Internet-Era Transformation of Scholarly Communications, vol. 32. *Learned Publishing*, pp. 13–25.
- Grimme, S., Holland, C., Potter, P., Taylor, M., Watkinson, C., 2019. *The State of Open Monographs: An Analysis of the Open Access Monograph Landscape and Its Integration Into the Digital Scholarly Network*. Digital Science, London.
- Guéron, J.-C., 2017. Open Access: Toward the Internet of the Mind. Access: <https://www.budapestopenaccessinitiative.org/open-access-toward-the-internet-of-the-mind>.
- Harrington, R., 2019. Why Scholarly Societies Are Vitally Important to the Academic Ecosystem. *Scholarly Kitchen*. October 3. Access: <https://scholarlykitchen.sspnet.org/2019/10/03/why-scholarly-societies-are-vitally-important-to-the-academic-ecosystem/>.
- Harrington, R., 2021. Subscribe to Open: An Interview Post in Two Parts. *Scholarly Kitchen*. July 28 and July 29. Access: <https://scholarlykitchen.sspnet.org/2021/07/29/subscribe-to-open-s2o-an-interview-post-in-two-parts-part-2/>.
- Harris, S., Garcia, A., Nkoudou, T., Dandawate, V., 2021. Global Trends in Open Access: Themes From Africa, Asia, and Latin America. *Scholarly Kitchen*. May 13. Access: <https://scholarlykitchen.sspnet.org/2021/05/13/global-trends-in-open-access-africa-asia-and-latin-america/>.
- Herb, U., Schöpfel, J. (Eds.), 2018. *Open Divide: Critical Studies on Open Access*. Library Juice Press, Sacramento, CA.
- Herman, E., Akeroyd, J., Bequet, G., Nicholas, D., Watkinson, A., 2020. The Changed—and Changing—Landscape of Serials Publishing: Review of the Literature on Emerging Models, vol. 33. *Learned Publishing*, pp. 213–229.
- Hinchliffe, L., 2020. Revisiting—Transformative Agreements: A Primer. *The Scholarly Kitchen*. April 23. Access: <https://scholarlykitchen.sspnet.org/2020/02/06/revisiting-transformative-agreements-a-primer/>.
- Hoffman, A., 2021. *The Engaged Scholar: Expanding the Impact of Academic Research in Today's World*. Stanford University Press, Stanford, CA.
- Horbach, S., 2020. Pandemic publishing: medical journals strongly speed up their publication process for COVID-19. *Quant. Sci. Stud.* 1 (3), 1056–1067.
- Johnson, R., Watkinson, A., Mabe, M., 2018. *The STM Report: An Overview of Scientific and Scholarly Publishing, 1968–2018*, fifth ed. International Association of Scientific, Technical and Medical Publishers, The Hague.
- Kasprzak, D., Smyre, T., 2017. Forerunners and Manifold: a case study in iterative publishing. *J. Sch. Publish.* 48 (2), 90–98.
- Labaree, D., 2018. Gold Among the Dross. *Aeon*. December 18. Access: <https://www.aneon.com/2018/12/18/gold-among-the-dross/>.
- Lane, J., Lifshitz-Assaf, H., 2022. Dismantling the Ivory Tower's Knowledge Boundaries: A Call for Open Access as the New Normal in the Social Sciences Post-COVID. *Brookings: Reimagining Modern-Day Markets and Regulations*. May 3. Access: <https://www.brookings.edu/series/reimagining-modern-day-markets-and-regulations/>.
- Larivière, V., Haustein, S., Mongeon, P., 2015. The oligopoly of academic publishers in the digital era. *PLoS One*. June 10. Access: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0127502>.
- Lee, J., Haupt, J., 2021. Scientific globalism during a global crisis: research collaboration and open access publication. *High Educ.* 81, 949–966.
- Lopez, E., Garcia, A., 2019. Latin America's Longstanding Open Access Ecosystem Could Be Undermined by Proposals From the Global North. *London School of Economics Impact Blog*. November 6. Access: <https://blogs.lse.ac.uk/latamcaribbean/2019/11/06/>.
- Meagher, K., 2021. The politics of open access—decolonizing research or corporate capture? *Dev. Change* 52 (2), 340–358.
- Mittell, J., 2013. The Real Digital Change Agent. *Chronicle of Higher Education*. March 4.
- Muddiman, D., 2016. What if you could only publish 50 papers your entire career? *Anal. Bioanal. Chem.* 408, 663–664.
- Nkoudou, T., 2020. Epistemic alienation in African scholarly communications: open access as a *pharmakon*. In: Eve, M., Gray, J. (Eds.), *Reassembling Scholarly Communications*, pp. 25–40.
- Olijhoek, T., Tennant, J., 2018. The “Problem” of Predatory Publishing Remains a Relatively Small One and Should Not Be Allowed to Defame Open Access. *LSE Impact of the Social Science Blog*. Access: <https://blogs.lse.ac.uk/impactofsocialsciences/2018/09/25/the-problem-of-predatory-publishing-remains-a-relatively-small-one-and-should-not-be-allowed-to-defame-open-access/>.
- Pinfield, S., Wakeling, S., Bawden, D., Robinson, L., 2020. *Open Access in Theory and Practice*. Routledge, London.
- Piron, F., 2018. Postcolonial open access. In: Herb, U., Schöpfel, J. (Eds.), *Open Divide: Critical Studies on Open Access*, pp. 117–128.
- Piowar, H., Priem, J., Larivière, V., Alperin, J.P., Matthias, L., Norlander, B., Fairley, A., West, J., Haustein, S., 2018. The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*. 6, e4375.
- Pool, R., 2021. Transformational OA agreements: help or hindrance? *Res. Inf.* #115 (August/September), 4–9.
- Poynder, R., 2016. Time to Rethink the Institutional Repository? *Coalition for Networked Information*. Access: www.cni.org/topics/publishing/rethink-ir.
- Pyne, R., Emery, C., Lucraft, M., Pinck, S., 2019. *The Future of Open Access Books: Findings From a Global Survey of Academic Book Authors*. SpringerNature, London.
- Scarborough, S., 2021. Turning the Page. *Emory University News*. Access: <https://news.emory.edu/features/2021/01/digital-publishing/index.html>.
- Sengupta, P., 2021. Open access publication: academic colonialism or knowledge philanthropy? *Geoforum* 118, 203–206.
- Shieber, S., Suber, P., 2021. Good Practices for University Open-Access Policies. Access: https://cyber.harvard.edu/hoap/Good_practices_for_university_open-access_policies.
- Shockley, N., Joseph, H., Hagemann, M., 2018. BOAI 15 Survey Report. *LIS Scholarship Archive*. April 12.
- Suber, P., 2012. *Open Access*. MIT Press, Cambridge, MA.
- Tmava, A.M., Miksa, S.D., 2017. Factors influencing faculty attitudes towards open access institutional repositories. *Proc. Assoc. Inf. Sci. Technol.* 54 (1), 519–522. Access: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pza2.2017.14505401061>.

- Trimble, S., Grody, W., McKelvey, B., Gad-el-Hak, M., 2010. The glut of academic publishing: a call for a new culture. *Acad. Quest.* 23, 276–286.
- Warren, J., 2019. How much do you have to publish to get a job in a top sociology department? Or to get tenure? Trends over a generation. *Sociol. Sci.* 6, 172–196.
- Willinsky, J., 2006. *The Access Principle: The Case for Open Access to Research and Scholarship*. MIT Press, Cambridge, MA.
- Wise, A., Estelle, L., 2019. Learned Societies, the Key to Realizing an Open Access Future? LSE Blog Impact of the Social Sciences. June 24. Access: <https://blogs.lse.ac.uk/impactofsocialsciences/2019/06/24/learned-societies-the-key-to-realising-an-open-access-future/>.

Innovations in university student learning research into blended course designs: incorporating self-report and observational data

Robert A. Ellis^a, Feifei Han^b, and Ana-Maria Bliuc^c, ^a Department of Learning & Teaching, Griffith University, Brisbane, QLD, Australia; ^b Institute for Learning Sciences & Teacher Education, Australian Catholic University, Sydney, NSW, Australia; and ^c Department of Psychology, School of Social Sciences, The University of Dundee, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	74
The challenge	74
Research into student approaches to learning	75
Quantitative approaches to identifying qualitative variation in learning	76
Materiality data used to complement self-report data as evidence of learning	78
Entanglement	78
Emergence	79
Entanglement in the 3P model	79
Combining self-report and observational data in investigations into blended course designs	79
Conclusions	81
References	82

Introduction

The quality of university student learning is an ongoing concern for those leading universities, for students, teachers and parents alike. Innovations in pedagogy, changes in the remit of universities to widen access and advances in Internet and computer-based technologies have increased the growth of alternative configurations of university credentials, structures of the student learning experience and even the learning spaces themselves, creating an urgent need for entirely new ways of conceptualizing and assessing the learning experience of students (Ellis and Goodyear, 2019; Wong et al., 2019). Research has repeatedly found positive impacts of various forms of innovations in learning, such as learning embedded in groupwork, inquiry-based learning, problem-based learning, social networking sites, web-conferencing for peer learning, e-portfolios, virtual spaces and second worlds (e.g., Anderson, 2016; Ngan et al., 2018). The higher education sector's rapid development in the use of computer and web-based technologies has motivated and shaped innovations in the experiences of learning of university students worldwide. These developments have created significant theoretical, methodological and philosophical challenges for researchers and those concerned about the quality of learning experiences at university. How do you evaluate the innovations and the quality of learning when part of the experience is in class and part is online? What role do learning technologies have in a quality learning experience? How do students collaborate in blended learning contexts? How should teachers design courses which require students to go back and forth between in-class and online contexts as they pursue their learning outcomes?

The challenge

This chapter considers these questions in the context of 50 years of research into university student learning, most commonly referred to as Student Approaches to Learning (SAL) (Pintrich, 2004; Prosser and Trigwell, 1999; Trigwell and Prosser, 2020). SAL is a relational model of the student experience of learning in which student prior characteristics of learning, the learning environment of their course and the perceptions the students hold about it, their approaches to, and conceptions of, learning and learning outcomes are inextricably entwined and contribute in different ways to the quality of the experience. Fig. 1 provides a visual presentation of the model.

SAL has been an important research program into the quality of university student learning. It has clarified the relationality of key aspects of the university student learning experience and how those aspects are related to each other in the context of qualitatively different outcomes. In summary, deep approaches to learning are related to positive perceptions of the learning context, cohesive conceptions of learning and relatively higher academic outcomes. In contrast, and significantly, surface approaches to learning are related to negative perceptions of the learning context, fragmented conceptions of the learning and relatively lower academic outcomes. It is the relationality amongst these aspects that allow institutional leaders, teachers and students to improve learning outcomes. An adjustment in how a student perceives the learning context can lead to deeper approaches to learning. A clarification of the concept of the object of learning in an experience can lead to a better perception of factors that shape the quality of learning outcomes. Five decades of research suggest that these qualitative shifts can lead to relatively higher academic outcomes for students.

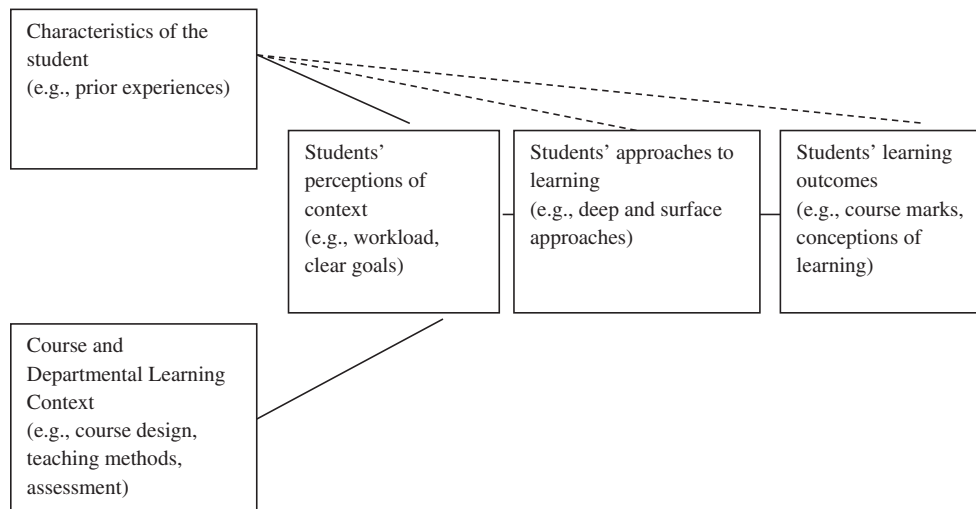


Fig. 1 The student experience of learning (3P model). Adapted from Prosser and Trigwell (1999).

While SAL has much to offer in understanding contemporary university student experiences of learning, it was not designed to deal with observational forms of evidence such as how students use learning technologies, nor with measurements of student group and teamwork. Understanding variation in conceptions, approaches and perceptions of learning has been shown to be essential if we want to understand variation in the students' ability to achieve learning outcomes. However, these self-report variations do not reveal granular descriptions of different strategies when hundreds or thousands of students use learning technologies in tens of thousands of different ways and sequences. To understand how material elements in the university student experience of learning are related to conceptions, approaches and perceptions of learning, it is illuminating to consider where they might be located in the 3P model of the learning experience in Fig. 1.

When looking at Fig. 1, it is possible to imagine how material elements might be conceived of in the course and departmental context, and how a student's use of technologies might be in approaches to learning and consequently in their perceptions of learning technologies in the course (Ellis and Bliuc, 2019). However, there are many additional challenges in studying the complexity of experiences of blended course designs (Ma and Lee, 2021), including students' choices of social interactions in learning, such as with whom and how to collaborate (Claros et al., 2015); and the purposes for which they use learning technologies in both physical and virtual learning spaces (Ellis and Goodyear, 2013; Laurillard, 2002). As such, it is valuable to evaluate university student learning experiences of blended course designs with research methods that move beyond self-report data to include observational data. By combining different sources of data and appropriate accompanying methodologies, researchers are able to develop a more holistic understanding of the contributions of learners and material things involved in learning to qualitatively different levels of academic performance, (Ellis and Goodyear, 2013; López-Pérez et al., 2011; Wu et al., 2010). To understand how observational evidence can elaborate the value of self-report evidence of student learning, the following begins with SAL methodologies and the insights it provides, and then considers methodologies that draw on observational and self-report data offering complementary yet alternate perspectives on qualitatively different outcomes.

Research into student approaches to learning

Early seminal SAL studies described variation in student approaches to studying (Marton and Säljö, 1976; Svensson, 1977). Approaches to studying were found to largely fall into one of two categories: deep approaches, which were closely related to understanding, to reflective and evidence-based thinking; or surface approaches, which were more mechanistic in nature, tending to simply reproduce without seeking the underlying meaning of the activity or outcome. These associations were found in student approaches to studying and examination performance (Svensson, 1977) and in student approaches to studying and the way they read passages. The latter study designed the questions for three separate reading tasks and discovered that the design of the questions about the readings if poorly designed, were more likely to induce a surface approach to studying.

Key contributions of SAL have been the identification of qualitatively different variables of approaches to learning and conceptions of learning, adopted by multiple well-cited researchers in the field (Biggs, 1987; Biggs and Tang, 2011; Entwistle and Tait, 1990; Prosser and Trigwell, 1999; Trigwell and Prosser, 2020; Ramsden, 1991, 2003). The following provides some examples of qualitatively different experiences of learning through discussions in a Social Work course (Ellis, Goodyear, Prosser, & O'Hara, 2006).

The examples in Table 1 present fundamental differences in the ways students conceived of learning through their discussions in a course that graded discussions in class and online as part of the assessment framework. The most holistic conceptions of

Table 1 Qualitatively different examples of conceptions of learning through discussions.

Category	Description	Representative quotation
A Challenging ideas	Discussions as a way of challenging ideas and beliefs in order to arrive at a more complete understanding	It (discussing) challenges my beliefs, which is always good ... because a belief is something that is based on knowledge and experience and your understanding of the world, and if it is being challenged you are testing it ... If my beliefs are challenged, I believe that my understanding of concepts is more complete
D Checking ideas	Discussions as a way of collecting ideas	Getting the teacher's point of view ... it's good being able to talk and make sure you are really learning what you are supposed to be learning. It is just sort of reassuring

Adapted from Ellis et al. (2006).

discussions reported that discussions were a way of testing one's beliefs and understanding. The accompanying interviews showed that it was the very process of challenging ideas through iterative conversation back and forth that clarified misconceptions and confirmed facts and evidence upon which beliefs were built. In contrast, some conceptions of discussions were focused on formulaic and mechanistic strategies that emphasized reproducing the ideas of the teacher, rather than developing personal understanding. In these conceptions, students thought that if they could reproduce the teacher's ideas, then they would be confident and succeed.

In the same study, students reported qualitatively different approaches to learning through discussions.

Table 2 describes qualitatively different approaches to learning through discussions in the same course as the conceptions in **Table 1**. The first thing to notice is that while a conception of learning through discussions encompasses both those in class and those online, the description of an approach to learning through discussions is described separately in order to reveal the intent and strategies adopted in class and those strategies adopted online. This is because not all strategies that work in class will necessarily suit an online context and vice versa. Deep approaches to learning through discussions in class emphasize comparison between one's own ideas and the others heard in discussions and use them as a way to receive feedback on ideas. Online, deep approaches emphasize reflection enabled in part because the exchanges are often asynchronous, with plenty of time to think in between posts. Significant reflection often involves considerations of how to apply the ideas on which are being reflected. In contrast, students reporting surface approaches to discussion do not approach them as a way of learning, but rather as a way of being exposed to different ideas so you can string them together. When pressed, discussions seem to be about fulfilling task requirements. This includes discussions online which may involve strategies of reading other postings in order to get a good mark or avoid repetition.

Quantitative approaches to identifying qualitative variation in learning

SAL has a long history of using quantitative methodologies in order to investigate and distinguish between the quality of the learning experiences of groups of students in the same class. Key instruments used in the research area are Biggs Study Process Questionnaire (Biggs, 1987; Biggs et al., 2001), Entwistle's Approaches to Study Inventory (Entwistle et al., 2000), Ramsden's perceptions subscales (Ramsden, 1991), Trigwell and Prosser's teaching inventory (Trigwell and Prosser, 2004).

As an example informed by these approaches, one study looked at engineering students' approaches to learning through discussions in an undergraduate class. In the final phase of the study, quantitative scales identified the following qualitatively different groups using cluster analysis.

In a third-year undergraduate engineering course, students were studying how to set up an e-commerce business. As part of their assessment, they were graded on their discussions in their laboratory work and online fora in relation to how well they demonstrated an understanding of the stages of setting up a business. The cluster analysis in **Table 3** identified two groups of students: one group who recorded high scores on fragmented conceptions, surface face-to-face and online approaches with low scores on cohesive conceptions, deep face-to-face approaches, and course marks. In contrast, the second group recorded relatively high scores on cohesive conceptions, deep face-to-face approaches and course marks and low scores on fragmented conceptions, surface face-to-face and online approaches and course marks.

These results provide outcomes consistent with the 3P model and previous research in the area. The results show the difficulties teachers face in achieving alignment with the learning outcomes of a course, particularly when including online parts into course design which require students to pursue their learning outcomes back and forth between in-class and online contexts. The results indicate the persistence of difficulties in helping students to adopt deeper approaches to learning and relatively more cohesive conceptions of learning and have been replicated and developed further across tens of thousands of students in many different universities on multiple continents (Biggs and Tang, 2011; Ellis and Goodyear, 2013; Prosser and Trigwell, 1999; Ramsden, 2003; Richardson, 2000).

The studies reported above and others in the SAL field have highlighted key challenges in ensuring the quality of the university student experience of learning. The evidence from the self-reported studies by students suggest that not all students experience the course design in the same way; some have misconceptions about what they should be learning, others have poor approaches to their learning both in class and online. Others report clear conceptions of the purpose of the course and learning outcomes; and useful,

Table 2 Qualitatively different examples of approaches to learning through discussions.

Category	Description	Representative quotation
A	Engaging in face-to-face discussions to analyze experiences and opinions through feedback	<i>Face-to-face</i> As far as learning, I like other people's ideas. I am never presumptuous enough to think that anything I say is the absolute truth and um for all I know I could be completely wrong ... I like to see how other people absorb my information. It is important to me that I have explained it well enough so other people can learn off it. I have always thought that if you can explain something to someone else, for example my girlfriend is in Year 12 right now and I really didn't do well in chemistry and after tutoring her I am like I could do my HSC now. To me it is important that I can explain my ideas so it is good to be able to let other people in and plus I get their feedback on it as well immediately.
	Engaging in online discussions to evaluate postings to reflect on key ideas	<i>Online</i> It just makes me think, like the ideas, like someone today when a post I made this morning and one two days ago, which was about men in counseling and about how they just feel society pretty much pushed them out of it. It wasn't really that original but it was something which I hadn't thought of before. So I mean, I didn't respond to it because I didn't have much to say. It was just something for me to think about it as I come to University, and it will be something that I have already thought about and probably be able to apply it.
D	Engaging in face-to-face discussions to fulfill task requirements	<i>Face-to-face</i> I don't think there is that much learning, the most amount of learning would happen in lectures... I don't think there is that much learning, I think it's just embracing different ideas and I don't think you are actually learning something, I think what's in the readings it's not what you actually learn it's just stringing them all together and engaging, yeah.
	Engaging in online discussions to read postings to avoid repetition	<i>Online</i> I tend to read all of them first. Because I tend to want to write something a bit different from all of them and sort of stand out a little bit because I thought I would get good marks for that, but that's not the point. But then if I read all of them as well I can reply to some of them, that's the thing like if I agree with someone's posting, then I can reply which is always good and then extend on from what they have just said. So yeah I tend to read all of them first yeah.

Adapted from Ellis et al. (2006).

Table 3 Cluster analysis of parts of the student learning experience and achievement.

Variable	Cluster 1 = 57	Cluster 2 = 31	p
Conceptions of learning through discussions			
Cohesive conceptions	-0.293	0.539	0.00
Fragmented conceptions	0.424	-0.779	0.00
Face-to-face approaches to learning through discussions			
Deep face-to-face approach	-0.161	0.297	0.01
Surface face-to-face approach	0.430	-0.791	0.00
On-line approaches to learning through discussions			
Deep on-line approach	-0.130	0.238	not sig.
Surface on-line approach	0.307	-0.565	0.00
Achievement			
Course mark	-0.246	0.451	0.00

Adapted from Ellis et al. (2008).

illuminative and deep approaches to learning. Furthermore, a key significance of these findings is the relatedness of key variables, such as conceptions and approaches to the quality of outcomes. It seems that teachers who succeed in helping students to move from surface approaches to deeper approaches, will also encourage students to develop more holistic and meaningful conceptions of what they are learning. Similarly, if they are able to clarify the purpose and goal of what the students should be learning in ways that help students to move away from fragmented conceptions toward cohesive ones, then the research suggests that students are more likely to adopt deeper approaches to their learning.

The outcomes from SAL research have a strength in that they use both qualitative and quantitative methodologies in order to reveal the qualitative variation across the student experience of learning. A combination of methodologies helps to provide a more thorough understanding of the student experience and offers actionable knowledge to teachers and education leaders who seek evidence upon which to design the program profile of a university. However, SAL research has been criticized for its reliance on self-report data including interviews only. There is an opportunity in higher education research to make use of the increasing availability of observational data from learning technologies in the form of learning analytics. For this reason, we discuss how a use of observational data in research designs into learning can complement, extend and enhance our understanding of the qualitative variation in the student experience.

Materiality data used to complement self-report data as evidence of learning

To understand how an evaluation of materiality in the student experience of learning can help improve student learning research and the actionable knowledge it provides, it is valuable to consider the role of materiality in learning, including the observational evidence it can offer to improve learning and the suggested pathways for research its combination with self-report data offers the field in the decades ahead. The examples described in the following capture methodological developments in research into university student learning from a series of closely related investigations by the authors and colleagues.

Materiality in the university student experience of learning is most ubiquitously understood in terms of the learning technologies that most students are using in contemporary course designs. It can also apply to other material aspects of the student experience such as learning spaces, but for the purposes of this chapter we use learning technologies as an example. A crucial concept in considering how learning technologies may be investigated for their role in contributing to qualitative variation in the learning experience is “entanglement”.

Entanglement

The idea of entanglement helps to conceptualize how non-human elements are bound up within university student learning processes (Fenwick, 2012). Contemporary course designs in university programs typically involve learning technologies in activities such as communication, research, reading and writing. In any activity it is common for a student to use a tool to support a process of, for example, writing, skim-reading, listening, talking, thinking, simulating or searching. Usually technology-enabled processes are an integrative and iterative experience as the student traverses back and forth and hopefully integrates ideas across resources, technologies, physical and virtual spaces and discussions. In practice it is difficult to know where one part of the process ends and another begins as the student engages in integrated processes that help them move toward their learning outcomes. It is in this sense that material aspects of the learning environment become entangled within the processes of the student experience of learning.

This idea of entanglement is very helpful to researchers seeking to understand how hundreds if not thousands of students might use learning technologies in their university experiences. Machine learning, algorithms, structural equation modeling, cluster analyses and other techniques can be used for descriptive, diagnostic, predictive purposes on large sets of observational data from the digital traces left behind by students engaging with the learning technologies, particularly online and networked tools and platforms. When these research methods draw on observational data, they can be used to complement and strengthen the findings of self-report data sources drawn from the same learning experiences. By providing outcomes from methodologies using both self-report and observational data, the understanding and usefulness of the knowledge arising from the research studies is likely to be more holistic and actionable.

The outcome of these methodologies involving large datasets are insights about the patterns of learning that can be discerned. While these are potentially very valuable, the area of learning analytics has received criticism that it does not frame the methods sufficiently in theory and that a lack of context of the intent behind the students’ use of the tools means that the interpretation of the results may be indeterminate at best and erroneous at worst (Boyd and Crawford, 2012; Selwyn, 2019). The combination of observational data sources with self-report data and methods helps to address this criticism, particularly if the self-report data emphasizes the theoretical context and intent of the learners. The self-report data described in the first part of this chapter for example, incorporates a use of extended interviews, carefully transcribed and analyzed and used to develop quantitative instruments within the SAL framework that can survey larger population samples. The point made here is that this type of self-report methodology, along with observational data and methods drawn on from within the same theoretical framework will provide relatively more holistic and complementary evidence about the qualitative differentiation of student experiences of learning. We are not suggesting that any one particular method is better than another, but rather a carefully designed and complementary use of self-report and observational data and methodologies will help provide knowledge which offers more insight into the actions that may lead to

improved experiences of learning. The benefits of combining self-report and observational sources of data are not, however, just limited to relatively more holistic research outcomes. They might also help us to discover how particular configurations of students, activity designs and technologies might enable a greater understanding by the way these elements combine, something that can be referred to as “emergence”.

Emergence

In addition to providing more useful results about learning involving materiality, the idea of “emergence” will help those concerned with the quality of university student learning not to miss any potential advantages for learning outcomes that may arise when non-human elements are part of the design of the learning experience. The idea of emergence probably first came from philosophical writing of Lewes ([Lewes, 1877](#)) who distinguished between resultant effects and emergent effects. The former can be measured simply by adding or subtracting the causes of the effects. In contrast, emergent effects are qualitatively more than the sum of the parts that led to the effects.

In the area of meteorology, a vortex is an example of emergence. Broken down to simple elements, a vortex may be made up of sticks, soil, stones, wind and other matter. However, when combined in a certain way, tornados and hurricanes create their own energy. In learning, there is potential for combinations of human and non-human elements like learning technologies to add to the understanding of students if they are used in certain ways in designs by teachers who understand how to help students develop deep approaches, cohesive conceptions and positive perceptions of their learning experience. In order to investigate this potential, research methodologies and the designs of research projects need to move beyond just looking at individuals as the unit of research, but individuals plus the non-human elements they use as they seek an outcome. The challenge in combining the unit of research to include both human and non-human elements is to identify appropriate self-report and observational data sources which are related and complementary.

Entanglement in the 3P model

We saw in the first part of the chapter that students adopt approaches to learning through discussions whose strategies are distinct to those they may adopt in class, some more effective than others. So it is clear that entanglement can occur in approaches to learning. Another fruitful area for understanding qualitative variation in the student experience of learning that arises from entanglement is their perceptions of the learning environment ([Han and Ellis, 2020a](#); [Ramsden, 1991](#)). The relationality of the 3P model suggests that students with positive perceptions of their experience, such as their perceptions about the quality of teaching, the clarity of the goals and standards of the course, the appropriateness of the assessment and overall satisfaction, tend to approach their learning in deep ways, hold cohesive conceptions and perform academically higher. Those with negative perceptions of these elements tend to adopt surface approaches, hold fragmented conceptions and tend to perform academically lower.

The relationality of perceptions to the quality of *material* affordances of their learning environment to the quality of outcomes has shown consistent associations across a number of emerging studies ([Ellis, 2016](#); [Han and Ellis, 2020b](#)). [Table 4](#) presents some example items investigating student perceptions of how the online part of their experience is related to their whole course experience.

Recent research has identified that students’ perceptions of a number of interesting aspects of their online experience are logically and positively associated with the quality of their learning outcomes ([Han and Ellis, 2020b](#)). These include how well integrated the online part of their experience is within their whole course design, their perceptions of the value of the online contributions from other students and their perceptions of the workload online in relation to the whole workload for the course. In short, students who adopt deep approaches to learning in blended course designs, tend to hold positive perceptions about these aspects and tend to perform relatively academically higher than those who hold negative perceptions of these aspects. So when considering approaches to online discussions from earlier in the chapter, and perceptions of the online part of their course experience, it is possible to see how entanglement can be conceived as a key dimension of the students’ whole experience of learning. However, these observations are built around self-report data sources and methodologies. How do we consider materiality in the learning experience methodologically using observational data? The next section considers this question with some examples.

Combining self-report and observational data in investigations into blended course designs

To improve the insight from research into student experience of learning in blended course designs, a combined use of self-report and observational data that includes dimensions of the materiality of the experiences has the potential to offer elaborated and relatively more holistic results to inform the actionable knowledge discovered from the research. As an example, the following summarizes the outcomes from a study into a first-year undergraduate engineering degree ([Ellis et al., 2017](#)). Students were expected to design an electronic system to solve an engineering problem, demonstrate how it related to the theoretical framework of the course, including basic programming constructs, write a report on the process and results and engage in a team-based process, which made significant use of the online environment. In the online environment, they were required to interact with the subject matter of the

Table 4 Perception subscales of key aspects of online environments in university courses.

<i>Subscale</i>	<i>Item</i>
Perceptions of integration between face-to-face and online	The online activities helped me to understand the lectures in my course. The online activities seemed to be well integrated with the assessment. The resources on the course website helped me to understand ideas in class. Studying online materials helped with the assessment in the course. I found the resources on the course website very useful for my learning.
Perceptions of online contributions	The ideas online in this course related to the ideas in class. Other students' online contributions helped me understand my ideas from a new perspective. Online contributions from others prompted me to reflect more on the ideas in this course. Online contributions from other students helped develop my understanding of particular topics. Online contributions by students in this course motivated me to think about things more. Online contributions by students in this course prompted me to engage more.
Perceptions of online workload	I needed more time to do my online activities. By the time I finished the online activities, I was running behind. The online learning in this course took too much time. The workload for the online activities was too heavy. A better balance between the online activities and the other tasks would help my workload.

Adapted from Han and Ellis (2020b).

course and engineering problem, engage with the online videos and multiple-choice questions, interact with the formative assessment elements and submit them and use feedback about their participation to gauge the level of their effort.

In that study, the approach subscales used were from the revised Study Process Questionnaire (Biggs et al., 2001). These were the self-report variables. The observational variables were the frequency of use of the online tools that enabled the students to engage with the resources, videos, multiple-choice questions, and feedback information from the dashboard. Table 5 presents the two-cluster solution of the experience. All variables were statistically significant in this study except for time spent online (duration).

Table 5 presents two qualitatively different groupings of students in the engineering course. One group of 43 students, who reported positive, statistically significant scores on the deep approaches to study variable, all the observational variables and the academic performance variable as indicated by the course mark. In contrast, a second group of 102 students reported a positive statistically significant score on the surface approaches to study variable, and negative scores on all the observational variables and a negative score on the academic performance variables as measured by the course mark. In brief, students reporting a deep approach to study, objectively and relatively engaged more with the online environment in statistically significant ways and performed a relatively higher level of academic achievement; whereas students reporting a surface approach to study, objectively and relatively engaged less with the online environment and performed at a relatively lower level of academic achievement.

Another aim of the above-mentioned study (Ellis et al., 2017) was to examine what contributions the observational variables made to the research design and results. To achieve this aim, two multiple regression analyses were performed. The first model involved the self-report variables only and the second model includes both the self-report and observational variables. Table 6 presents the results of the regression analysis. The self-report variable that was significant (surface approach to study) accounted for 9% of the variance in the outcomes. When the regression comprised both the self-report and observational variables, the model was able to explain 34% of the variance.

These results have been replicated with various levels of additional variance explained in Model 2 in a number of studies (Han et al., 2020; Pardo et al., 2017). Some of the designs of those studies have adopted machine learning methodologies to deal with data at scale. The innovation in the research methodology being emphasized here is the value of including observational data variables to complement and extend the results of the self-report variables within the same theoretical framework. By combining the data sources in the research design, the depth of insight into reasons for the variation is relatively more detailed and offers relatively more evidence for the development of actionable knowledge.

Table 5 Summary statistics of the two-cluster solution of the engineering student experience.

Variables	Deep cluster (n = 43) Mean (SD)	Surface cluster (n = 102) Mean (SD)	F	p	η^2
Self-report variables					
Deep approaches	0.93 (0.73)	-0.39 (0.83)	82.11	0.00	0.37
Surface approaches	-0.52 (1.18)	0.22 (0.82)	18.74	0.00	0.12
Observational variables					
Dashboard view	0.32 (1.20)	-0.14 (0.87)	6.69	0.01	0.05
Col-Exp	0.27 (1.02)	-0.11 (0.97)	4.46	0.04	0.03
Resource	0.33 (1.08)	-0.14 (0.93)	7.10	0.01	0.05
Video	0.26 (1.39)	-0.11 (0.76)	4.19	0.04	0.03
MCQ	0.61 (1.54)	-0.15 (0.35)	7.57	0.01	0.05
VMCQ	0.27 (1.77)	-0.11 (0.28)	4.72	0.04	0.03
Exercise	0.44 (1.25)	-0.18 (0.81)	12.58	0.00	0.08
Time					
Duration	0.20 (1.09)	-0.08 (0.95)	2.42	0.12	0.02
Academic Performance					
CM	0.53 (0.99)	-0.23 (0.82)	19.77	0.00	0.12

Notes: Deep = deep approaches to study, Surface = surface approaches to study, Dashboard view = student views of a dashboard with information about the engagement with the weekly activities, Col-Exp = collapse and expand view, Resource = student views of any page of the course notes, video = course videos, MCQ = multiple-choice questions, VMCQ = multiple-choice questions in videos, exercise = student answers to summative assessment exercises, Duration = time spent online, CM = course mark.

Adapted from Ellis et al. (2017).

Table 6 Results of multiple regression analysis of the engineering student experience.

Variables	B	SE B	B	t	Adjusted R ²	p	f ²
Model 1					0.09	0.00	0.10
Self-report variable							
Surface approaches	-0.31	0.08	-0.31**	-3.84		0.00	
Model 2					0.34	0.00	0.52
Self-report variable							
Surface approaches	-0.29	0.07	-0.29**	-4.16		0.00	
Observational variables							
Dashboard view	0.03	0.08	0.03	0.31		0.75	
Collapse and expand view	-0.34	0.19	-0.34	-1.84		0.07	
Resource	0.88	0.21	0.88**	4.28		0.00	
Multiple choice questions	0.29	0.12	0.29*	2.40		0.02	
Exercise	-0.26	0.18	-0.26	-1.40		0.16	

Notes: ** $p < 0.01$, * $p < 0.05$, Surface approaches = surface approaches to study, Dashboard view = student views of a dashboard with information about the engagement with the weekly activities, Resource = student views of any page of the course notes, Exercise = student answers to summative assessment exercises.

Adapted from Ellis et al. (2017).

Conclusions

The contemporary context of university student learning is characterized by an unprecedented amount of change. These changes are occurring in the areas of credential design, pedagogical approaches, learning technologies, and activity design. To understand the impact of these changes on the quality of the student experience of learning; robust, valid and reliable data is required from different perspectives on the experience so that findings can be extended, elaborated and enhanced. This can be achieved if self-report data is combined with observational data from the same learning experience and investigated within a strong theoretical framework like the SAL framework.

A key challenge in investigating university student learning in contemporary contexts is how to appropriately account for the contribution of materiality in the experience. Within the SAL framework, the ideas of entanglement and emergence help to conceptualize an integrated and embedded use of learning technologies that is consistent with the real-world use of learning technologies by students in their learning. Some students do adopt relatively deep approaches to learning, including when they use the technologies, and these tend to be related to cohesive conceptions about their subject matter, positive perceptions of the blended learning environment and relatively higher academic outcomes. Similarly and in contrast, other students adopt relatively surface approaches to learning, which are related to ideas about the subject matter which do not quite get the whole picture, but miss key aspects of what is being studied. These experiences tend to be related to negative perceptions of the learning environment and relatively lower academic outcomes.

The university student experience of learning is likely to continue to face ongoing change in the context in which it is occurring over the next decade. Consequently, those responsible for learning in universities, including the program directors, course convenors, teachers and even the students themselves, will benefit from evidence that provides real-work actionable knowledge to improve how, what, why and where students learn successfully.

References

- Anderson, T., 2016. Theories for learning with emerging technologies. *Emerg. Innov. Digital Learn.* 1, 35–50.
- Biggs, J., 1987. Study Process Questionnaire Manual. Student Approaches to Learning and Studying. Australian Council for Educational Research, Hawthorn.
- Biggs, J., Kember, D., Leung, D.Y., 2001. The revised two-factor study process questionnaire: r-SPQ-2F. *Br. J. Educ. Psychol.* 71 (1), 133–149.
- Biggs, J., Tang, C., 2011. Teaching for Quality Learning at University. McGraw-Hill Education, London.
- Boyd, D., Crawford, K., 2012. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inf. Commun. Soc.* 15 (5), 662–679.
- Claros, I., Cobos, R., Collazos, C.A., 2015. An approach based on social network analysis applied to a collaborative learning experience. *IEEE Trans. Learn. Technol.* 9 (2), 190–195.
- Ellis, R.A., 2016. Qualitatively different university student experiences of inquiry—associations amongst approaches to inquiry, technologies and perceptions of the learning environment. *Act. Learn. High. Educ.* 17 (1), 13–23.
- Ellis, R.A., Bliuc, A.M., 2019. Exploring new elements of the student approaches to learning framework—the role of online learning technologies in student learning. *Act. Learn. High. Educ.* 20 (1), 11–24.
- Ellis, R.A., Goodyear, P., 2013. Students' Experiences of E-Learning in Higher Education: The Ecology of Sustainable Innovation. Routledge, New York, NY.
- Ellis, R.A., Goodyear, P., 2019. The Education Ecology of Universities: Integrating Learning, Strategy and the Academy. Routledge, New York, NY.
- Ellis, R.A., Goodyear, P., Calvo, R.A., Prosser, M., 2008. Engineering students' conceptions of and approaches to learning through discussions in face-to-face and online contexts. *Learn. InStruct.* 18 (3), 267–282.
- Ellis, R.A., Goodyear, P., Prosser, M., O'Hara, A., 2006. How and what university students learn through online and face-to-face discussion: conceptions, intentions and approaches. *J. Comput. Assist. Learn.* 22 (4), 244–256.
- Ellis, R.A., Han, F., Pardo, A., 2017. Improving learning analytics—combining observational and self-report data on student learning. *Educ. Technol. Soc.* 2 (3), 158–169.
- Entwistle, N., Tait, H., 1990. Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *High Educ.* 19 (2), 169–194.
- Entwistle, N., Tait, H., McCune, V., 2000. Patterns of response to an approaches to studying inventory across contrasting groups and contexts. *Eur. J. Psychol. Educ.* 15 (1), 33–48.
- Fenwick, T., 2012. Mattering of knowing and doing: sociomaterial approaches to understanding practice. In: Hager, P., Lee, A., Reich, A. (Eds.), *Practice, Learning and Change*. Springer, Dordrecht, pp. 67–83.
- Han, F., Ellis, R.A., 2020a. Combining self-reported and observational measures to assess university student academic performance in blended course designs. *Australas. J. Educ. Technol.* 36 (6), 1–14.
- Han, F., Ellis, R.A., 2020b. Initial development and validation of the perceptions of the blended learning environment questionnaire. *J. Psychoeduc. Assess.* 38 (2), 168–181.
- Han, F., Pardo, A., Ellis, R.A., 2020. Students' self-report and observed learning orientations in blended university course design: how are they related to each other and to academic performance? *J. Comput. Assist. Learn.* 36 (6), 969–980.
- Laurillard, D., 2002. Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies. Routledge, New York, NY.
- Lewes, G.H., 1877. Problems of Life and Mind. Trübner & Company, London.
- López-Pérez, M.V., Pérez-López, M.C., Rodríguez-Ariza, L., 2011. Blended learning in higher education: students' perceptions and their relation to outcomes. *Comput. Educ.* 56 (3), 818–826.
- Ma, L., Lee, C.S., 2021. Evaluating the effectiveness of blended learning using the ARCS model. *J. Comput. Assist. Learn.* 37 (5), 1397–1408.
- Marton, F., Säljö, R., 1976. On qualitative differences in learning: I—outcome and process. *Br. J. Educ. Psychol.* 46 (1), 4–11.
- Ngan, O.M.Y., Tang, T.L.H., Chan, A.K.Y., Chen, D.M., Tang, M.K., 2018. Blended learning in anatomy teaching for non-medical students: an innovative approach to the health professions education. *Health Prof. Educ.* 4 (2), 149–158.
- Pardo, A., Han, F., Ellis, R.A., 2017. Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. *IEEE Trans. Learn. Technol.* 10 (1), 82–92.
- Pintrich, P.R., 2004. A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ. Psychol. Rev.* 16 (4), 385–407.
- Prosser, M., Trigwell, K., 1999. Understanding Learning and Teaching: The Experience in Higher Education. McGraw-Hill Education, London.
- Ramsden, P., 1991. A performance indicator of teaching quality in higher education: the course experience questionnaire. *Stud. High Educ.* 16 (2), 129–150.
- Ramsden, P., 2003. Learning to Teach in Higher Education. Routledge, New York, NY.
- Richardson, J.T., 2000. Researching Student Learning: Approaches to Studying in Campus-Based and Distance Education. Open University Press, Buckingham.
- Selwyn, N., 2019. What's the problem with learning analytics? *J. Learn. Anal.* 6 (3), 11–19.
- Svensson, L., 1977. Symposium: learning processes and strategies—iii: on qualitative differences in learning: iii—study skill and learning. *Br. J. Educ. Psychol.* 47 (3), 233–243.
- Trigwell, K., Prosser, M., 2004. Development and use of the approaches to teaching inventory. *Educ. Psychol. Rev.* 16 (4), 409–424.
- Trigwell, K., Prosser, M., 2020. Exploring University Teaching and Learning: Experience and Context. Palgrave Macmillan, London.
- Wu, J.H., Tennyson, R.D., Hsia, T.L., 2010. A study of student satisfaction in a blended e-learning system environment. *Comput. Educ.* 55 (1), 155–164.
- Wong, J., Baars, M., Koning, B.B.D., Zee, T.V.D., Davis, D., Khalil, M., Paas, F., 2019. Educational theories and learning analytics: from data to knowledge. In: *Utilizing Learning Analytics to Support Study Success*. Springer, Cham, pp. 3–25.

Uses of technology in navigating the pandemic: the case of African higher education: Chika Sehoole, Lydia Ikwuemesim, Olaide Agbashe

Chika Sehoole^a, Lydia Ikwuemesim^a, and Olaide Agbashe^b, ^a University of Pretoria, Pretoria, South Africa; and ^b University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	83
Egypt	84
South Africa	85
Nigeria	85
Kenya	86
Cases of selected African countries	86
Implications the pandemic for the future planning	87
Advancing meaningful partnerships	87
Advancing technological infrastructure	87
Innovative modes of education delivery	88
Addressing issues inequality, equity, and access in educational systems	88
Conclusion	88
References	89

Introduction

Pandemics are known as worldwide epidemics of excessive number of sicknesses that cause deaths, social disruption, high inflation and economic downturn (Nicola et al., 2020; Karabulut et al., 2020; Piret and Boivin, 2021). Over the years, there has been recorded history of pandemics in human history (Huremović, 2019; Karabulut et al., 2020; Piret and Boivin, 2021). Some of such historical pandemics include: the Spanish flu in 1918–1920, which was the first global pandemic, Asian flu in 1957, Hong Kong flu in 1968 and swine flu in 2009; the smallpox outbreak in the former Yugoslavia in 1972, HIV/AIDs which started in the United States of America in 1980s, Ebola outbreak of 2014–2016 that ravaged mostly the Central and West Africa (Huremović, 2019; Akin and Gözel, 2020; Bai et al., 2020; Lyngdoh, 2020; Owusu-Fordjour et al., 2020; Zhai and Du, 2020).

This chapter will focus on the most recent COVID-19 pandemic by providing an analysis of the uses of technology in navigating the pandemic in selected African countries.

The report of the outbreak of the novel COVID-19 in China at the end of December 2019 took the world by surprise. What started as a potentially dangerous but yet less seriously taken outbreak evolved to become one of the most important global occurrences of the 21st century with such devastating effects. Within a period of 4 months the virus had mutated from an epidemic to a pandemic that had engulfed the whole world. By mid-April UNESCO (2020) reported that 192 countries had closed all schools and universities, affecting more than 90% of the world's learners: almost 1.6 billion children and young people. With the massive closures of all education institutions, societies had to deal with unprecedented disruptions of teaching and learning in schools, colleges and universities which many sectors of society were unprepared for. The pandemic further triggered the closure of borders and restricted human travel to all, regardless of social, political or economic class, and it in this way became the "greater equalizer" in the midst of an unequal world (Emeto et al., 2021; Memish et al., 2021).

The pandemic further impacted the economies of many countries especially in Africa leading to the shrinking of their revenue bases (Gondwe, 2020; Obayelu et al., 2020). This forced the redistribution of and reprioritization of the use of meagre resources to more urgent needs and sectors, such as agriculture, food security and COVID-19-related healthcare. In some cases, the reprioritization process led to the postponement or even cancelling of commitments, for instance, to capital development, research and innovation and hiring staff, among others. Furthermore, societal lockdowns and spending to combat COVID-19 led to substantial and increasing economic challenges (Mhlanga and Ndhlovu, 2020). These may have an even greater impact in low- and middle-income countries. This happened in a context where many African countries were already dealing with issues of declining funding in real terms, government regulation of tuition fees, rising student debt, weak economic growth and weakening and volatile currencies.

Literature on COVID-19 focuses on the challenges and opportunities associated with online learning during COVID-19 (Ade-doyin and Soykan, 2020), the global impact of this pandemic on social, economic and health sectors of society (Ciotti et al., 2020; Adedoyin and Soykan, 2020; Daniel, 2020; Pokhrel and Chhetri, 2021). A study of COVID-19 pandemic in Africa by Lone and Ahmad (2020), focused on the vulnerability, preparedness and economic impact of COVID-19 in the African continent. There have not been enough studies that focused on the uses of technology in African higher education institutions in navigating the pandemic.

The purpose of this chapter is to discuss and analyze the challenges and implications of COVID-19 for higher education in Africa and the opportunities that arise therefrom. It will argue that while the COVID pandemic has accentuated the inequalities between African higher education systems and the rest of the world, it has also provided opportunities for innovation that may inform policy for the delivery of quality higher education.

Many education systems on the African continent, just like the rest of the world were, unprepared to deal with the pandemic. This calls for national governments to invest in their education systems to be robust and resilient throughout the 12 years of basic education, and for the higher education system to be agile in dealing with such disruptions. The unequal access to technological devices, data networks and reliable electricity supply impacted not only schools and universities but also the preparedness of educators, parents and the learners themselves to move rapidly to remote teaching and learning. These also revealed the glaring differences between rich and poor communities and nations. COVID-19 put on a mirror in front of many societies of the world to see and realize both their strengths and resilience on the one hand but also exposed the vulnerabilities of weaker communities especially in developing countries, on the other.

The closure of educational institutions across 165 countries, which affected 1.5 billion learners forced higher education institutions to think of new ways of ensuring learning continuity outside the four walls of the classroom. Since the outbreak of the Covid-19 pandemic, several studies (Hodges et al., 2020; El Said, 2021; Le Grange, 2020; Heyang and Martin, 2021; Menon and Motala, 2021; Ngwacho, 2020) have been conducted to explore how higher institutions around the world navigated through the unprecedented change in traditional teaching and learning. The most popular response to the crisis was a move to online delivery of lessons on diverse digital platforms United Nations (2020); even though early indications had suggested that online instruction was a poor substitute for face-to-face instruction (Domina et al., 2021). According to Bond et al. (2020), good number of higher education institutions (HEIs) adopted what was termed the “Emergency Remote Teaching—ERT” to mitigate the effect of the pandemic on education delivery. However, Hodges et al. (2020) and Le Grange (2020) argued that a shift to online teaching and learning in response to a crisis is not as authentic and as meaningful as a well-thought-out plan, as a sudden shift could be shocking to academics and students. Hodges et al. (2020) further argue that an ideal time to plan, prepare and develop a full university online course is between six to nine months and so, it would be impossible for lecturers to suddenly become experts in online teaching in a time of crisis or disaster. That notwithstanding, the timely effort of some African countries and HEIs in responding to the pandemic, mitigating its adverse effects on education and promptly leveraging technology for remote learning is laudable. The emergency remote teaching was implemented with various levels of success in many African countries. The following sections examine how some African countries and HEIs used technology to mitigate the effects of the pandemic to ensure continuous education delivery. In order to get a sense of how of the pandemic was dealt with and responded to by African higher education institutions, a selection of four countries representing the four major regions of the continent namely the North (Egypt), South (South Africa), East (Kenya) and West (Nigeria) was conducted to examine the use of technology in mitigating the effects of the pandemic. Additional cases of other African countries drawn from recent literature were also analyzed to provide more insights on how some of the countries used technology to navigate the pandemic.

Egypt

Egypt has the largest education system in the Middle East as well as North Africa with 44 universities and 830 other higher education institutions, enrolling 3 million students (El Said, 2021). As in other African countries, the Egyptian government made a swift move to close down all education institutions at the onset of the pandemic in March 2020. To ensure the continuation of educational processes, the Egyptian Ministry of Higher Education and Scientific Research urged all higher education institutions to adopt the modality of distance education under lockdown (El Said, 2021). Online platforms were then set up, in cooperation with Microsoft International and Blue Cloud for Software Development, (World University News). The platform allowed both faculty members and students to apply social distancing measures to combat the spread of the novel coronavirus (COVID-19), as well as allow for easy communication during virtual classes.

To measure the satisfaction of higher education stakeholders on the distance education mode, El Said (2021) reported that four-nation wide surveys were administered to students, education leaders, professors, and administrators in May 2020. A total of 106,897 participants completed the survey; and the findings showed internet connectivity and weak IT skills ranked the most difficult aspect of online learning in Egypt while recorded lectures ranked the most commendable way of online education delivery. Further findings revealed that the majority of education leaders recommended online distance education; and this led the Egyptian government to decide on incorporating both face-to-face and online learning into all future higher education plans. This, is believed, will make the country better positioned to deal with emergency situations like Covid-19. This will most importantly achieve crucial education goals such as reducing student density on campus, making the best use of professors’ expertise and IT infrastructure while gradually transforming students to lifelong learners. It was also reported that the move from contact sessions to online platforms did not lead to significant devastating learning experiences for students in Egypt. For example, an investigation carried out by El Said (2021) in one university in Egypt compared the academic grades of 376 business students who completed a face-to-face course in the spring of 2019 and the grades of 372 who completed the same course in the spring of 2020, but via online distance learning and under lockdown. The findings showed that there was no significant difference in the grades of students in the two groups and that the sudden move to online learning did not result in poor learning experience outcomes.

South Africa

On the March 5, 2020 South Africa recorded its first COVID-19 case. Within three weeks, the country was in total shut-down level five that brought everything to a standstill, not only in South Africa but also globally. South Africa was one of the countries that introduced the strictest country-wide lockdown actions beginning from March 26, 2020. The country also turned to online lessons to save the academic year. In the context of South Africa, online teaching and learning was not a novel idea. Before Covid-19, many universities had engaged in some form of hybrid or blended learning that combined face-to-face models with online models (Le Grange, 2020). Students and lecturers were familiar with the use of Learning Management Systems (LMS) and possessed the basic skills to at least upload and download course materials (Le Grange, 2020; Du Preez and Le Grange, 2020), access module information, continuous assessment, and exams marks, etc. For example, the University of Pretoria adopted the use of the online system known as Blackboard and branded as ClickUp while the University of Johannesburg makes use of Blackboard module. These functions of the LMSs were expanded to include live online lectures delivery (which could also be watched or downloaded at a later time), pre-loaded recorded lectures and online assessments, in order to aid the transition from face-to-face to virtual sessions. The new context of the pandemic required of both students and academics to have advanced technological skills and digital literacy levels for a full migration to virtual sessions during the pandemic. As Le Grange (2020) notes, some South African universities exhibited zeal, commitment, and passion to save the academic year by embracing the slogan “no student left behind” as their organizing and rallying theme. Innovative strategies to support students in ensuring that none of them is left behind include the following:

Universities purchased and loaned laptops to indigent students; made available data vouchers so that students can be able to connect to online services, those that did not have access to or reliable electricity supply were allowed to return to campus and residences so that they could continue with their online academic activities. The Department of Higher Education training (DHET) also came on board to support universities by negotiating with telecommunication companies to provide affordable pre-paid data packages and zero-rated data to students and staff. The zero-rated data enabled students and staff to access universities’ major websites, libraries websites and Learning Management Systems. However, Du Preez and Le Grange (2020) raised two concerns about online teaching and learning. The first was the promotion and support of unequal access to data, connectivity and technology in an already unequal society as South Africa, thereby raising the issue of distributive justice. Secondly, they raised concerns about epistemological access, as access to technology does not guarantee that all students have access to the knowledge that the university distributes. Students and student leaders also raised concerns about marginalizing students who lived in areas with poor internet connectivity and lack of electricity that would hinder them from effectively participating in online lessons. In response, to the concerns of students (and some staff members) and coupled with the commitment universities made to ensure that no students would be left behind, various universities utilized various means to address this problem. For example, the University of Cape Town, disbursed printed learning materials and USB drives to students who had no internet access at all (Petersen, 2020). The University of Pretoria couriered paper based learning materials, at huge cost, to students who were based in areas without internet services and electricity; and appointed tele-tutors to support those students. Tele-tutors made weekly telephone calls to students to support check on the progress of their studies.

Furthermore, Mphahlele et al. (2021) found out that some of the students in South Africa higher education had limited digital knowledge in Learning Management System. This impacted on the ability to meaningfully participate in the teaching and learning activities.

Nigeria

Nigeria operates a federal system of government with 36 states and the Federal Capital Territory of Abuja. The 36 states are made up of 744 local governments and education is managed by the federal, state and local governments. However, the Federal Ministry of Education is responsible for overall policy formation and ensuring quality control, but primarily involved with higher education. The higher education system consists of the university, polytechnics, monotechnics, and colleges of education (Jacob et al., 2020).

Although the first case of COVID -19 case was confirmed in Lagos, Nigeria, on 27 February 2020, it was only in March 2020 that Nigeria followed the global trend of shutdown of schools and universities in order to curb the spread of COVID -19 pandemic (Jegede, 2020).

Several studies that explored the impact of COVID-19 include positive responses from private university lecturers in navigating the use of technology in online learning (Ojo Joseph et al., 2021). According to Jacob et al. (2020), COVID-19 pandemic caused reduction of international education, disruption of academic calendar, cancellation of local and international conferences, teaching and learning gap, loss of workforce, and cut in budget of higher education in Nigeria.

The COVID-19 situation in Nigeria was worsened by the strike action of lecturers at public universities that was led by the Academic Staff Union of Universities (ASUU) which began on March 23, 2020, a few days before the closure of schools and universities nationwide. Unfortunately, the strike action lasted nine months, which led to the loss of the 2020 academic year (Uchendu et al., 2021).

Unlike countries such as South Africa and Egypt, many higher education institutions in Nigeria, especially public or government-owned universities, could not migrate to online teaching and learning during shutdown in March 2020. This is because hybrid learning and the incorporation of Information and Communication Technology into teaching and learning in public universities

had neither been popular nor successful. In addition, Egielewa et al. (2022) reported that due to poor electricity connectivity and internet accessibility in Nigeria, online learning during the COVID-19 pandemic was ineffective. This showed that Nigerian higher education students preferred the traditional classroom to online learning.

In another study conducted by Jegede (2020) the findings suggest that online learning can be an alternative measure to face to face learning in future pandemic situations.

However, Yunusa et al. (2021) reported challenges in using technology during online learning. These include lack of skills among personnel and students and poor or lack of digital technology infrastructure.

Yet in another study, Abosede et al. (2021) recommend that government should provide more technological infrastructural development in higher institutions while existing ones should be upgraded to meet the global standards.

Kenya

In Kenya the government ordered the closure of all schools and universities as a measure to mitigate the spread of the virus, just like other countries had done. Between March 15 and July 17, 2020, learning continued remotely and through home schooling. Most institutions leveraged television, radio broadcast, distance learning, online learning and gamification, to continue with the academic year with view to save the academic year (Kathula, 2020). Other commendable initiatives, applauded by university stakeholders were also taken by some universities to ensure the academic year was not lost and that teaching and learning reached a wider range of students. For example, the University of Nairobi made thousands of learning resources available remotely through electronic sources (Kathula, 2020). The Ministry of education also launched few strategies to ensure learning continuity, which include the broadcast of educational programmes via television, radio and YouTube (Ngwacho, 2020). The Kenyan government also partnered with Kenya Publishers Association to make free textbooks available to all students on the Kenya Education Cloud. In addition, the Kenya Civil Aviation Authority, on behalf of the Kenya government, partnered with Telkom Kenya to deploy Google's loon balloons of 4G base stations over Kenya's airspace. In spite of these laudable efforts to ensure no student was left behind during the pandemic, access of all students to these services remained a problem. As in many of African and developing countries, the deep-rooted issues of inequality and access in the society were brought to bare (Ngwacho, 2020). As a result, and coupled with the rising number of infections in the country, the Kenyan ministry of education announced the forfeiture of the 2020 academic year as a way of containing the virus. This had an adverse effect on the education of over 17 million students in Kenya.

Although the higher education in Kenya adapted to e-learning during COVID -19 pandemic, navigating through the pandemic was met with several challenges. Some of the challenge include: inadequate Information and Communication Technology (ICT) infrastructure, lack of technical and pedagogical competencies and training for e-tutors and e-learners, negative perceptions toward e-Learning, and domination of e-Learning aims by technology (Kibuku et al., 2020). Kathula et al. (2021), point out how students in remote and rural areas had challenges participating in online learning due to poor internet connection. Their challenges and grievances include excluding their marginalized peers with little or no access to internet services and those with no money to buy laptops (Wekullo et al., 2022).

Vegas and Winthrop notes that less than 25% of low-income countries in Africa had the capacity for some form of remote pedagogy, the majority of which were television and radio; compared to about 90% of high-income countries that had the capacity for remote pedagogy (Vegas, 2020). Many countries in Africa, even those who had no capacity for any form of internet-based teaching learning were also innovative in their approaches. They engaged some offline technology such as text messages and basic phone calls to ensure that learners in rural areas were not excluded. For instance, a national survey conducted about three weeks after leaning institutions were closed down in Senegal showed that school closure reduced learning and increased inequalities. Most learners could not continue their education via any technological platform, mainly TV and radio, but relied on parental support to continue learning (Le Nestour and Moscoviz, 2020).

Cases of selected African countries

The Democratic Republic of Congo (DRC) is one of the countries in central Africa that felt the devastating impact of COVID-19 pandemic in higher education (Sasidharan and Dhillon, 2020; Jacobs et al., 2021). It is on record that DRC experienced Ebola outbreak in 2014 and 2019–2020 (Lee et al., 2020). The outbreak of COVID-19 in 2020 was yet another blow on the country's higher education especially in their medical institutions (Échochard et al., 2021).

Sasidharan and Dhillon (2020) noted that as higher institution switched to online classrooms and blended learning approach globally, the medical institutions in DRC changed to the "Webside of Telemedicine", and Competency Based Medical Education (CBME). Although Karemere and Makali (2022) argue that distance learning determine pedagogical continuity during pandemic but cannot be a sustainable and generalized form of teaching. Sasidharan and Dhillon (2020) also reiterate that adapting to this "new—normal", of integration of technology in the medical institution negatively impacted the use of human skills like compassion and empathy which form the core cultural values of DRC and are critical in patient care.

In Ethiopia, Tamrat (2021) noted that navigating through the pandemic with the use of technology was very difficult especially where there is limited or no internet access. "Poor internet access, cost, availability of computers and related technology, little preparation, and the use of information and communication technology stood out as the most prevalent problems" (Tamrat, 2021, p.

66). Meanwhile in Sub-Saharan Africa [Aborode et al. \(2020\)](#) reported the impact of the pandemic in higher education as most of the students do not have access to home computers for online learning.

Meanwhile in Ghana, [Owusu-Fordjour et al. \(2020\)](#), reiterate the lack of the technical knowhow in the use of these technological devices by most Ghanaian students and the inability of most students to study independently as their parents could not assist them in their studies as the challenges experienced during the pandemic.

It is evident from the above analysis that while some African countries and learning institutions successfully transitioned to online mode of education delivery, others struggled, given the digital divide and underdeveloped technological infrastructure in Africa. The next section discusses some the implications brought about by the pandemic which which can be harnessed to build a stronger education system post Covid-19.

Implications the pandemic for the future planning

The innovative ways with which countries around the world, especially under-resourced African countries engaged their educational systems during the pandemic strongly suggest has implications for how to better plan for future pandemics. It is true that the pandemic has been tough, having adverse effects on individuals, families, societies and almost all sectors, including education. Nevertheless, we must brace up for life post-Covid 19, look beyond the difficulties and embrace the opportunities therein. The [United Nations \(2020, p. 4\)](#) rightly advises that, “the massive efforts made in a short time to respond to the shocks to education systems remind us that change is possible. We should seize the opportunity to find new ways to address the learning crisis and bring about a set of solutions previously considered difficult or impossible to implement.” In the sections that follow, we discuss the implications the pandemic has which could be harnessed to accelerate sustainable educational systems in Africa and globally.

Advancing meaningful partnerships

There have been calls for stronger and more meaningful partnerships amongst educational institutions, governments and private organizations ([Le Grange, 2020](#); [Ngwacho, 2020](#); [Vegas and Winthrop, 2020](#)). These partnerships offer some unique opportunities for all stakeholders to provide continuous and sustainable support to learners to access quality education, and not only at a time of crisis. These could include the provision of free devices, internet connectivity, textbooks and learning resources to students. This would make universities to be better equipped to cater for the educational needs of various students from different contexts. The move toward partnership is exemplified by the [UNESCO's \(2020\)](#), “Save Our Future” campaign, a global education coalition, that advocates for sustained education funding by international donors, for low and middle income countries. Multilateral partners such as the World Bank and the International Labor Organization, private companies such as Microsoft and Google, non-profit organizations such as Dubai Cares and Sesame Street, and media outlets such as the British Broadcasting Cooperation (BBC) world service have all joined the coalition to alleviate the adverse effects of school closure, especially on the most disadvantaged. Similarly, [Vegas and Winthrop \(2020\)](#), note that the closure of learning institutions witnessed partnerships between teachers and parents (the pandemic has proven that parents are strong allies in their children's education), community health workers and social welfare organizations, media companies and education leaders, technology companies and governments, non-profit organizations and businesses, all worked together to assist students like never before.

Advancing technological infrastructure

A study conducted by [Magesa and Josua \(2022\)](#) recommend more investments to widen access to information communication technology (ICT), and to promote inclusivity and massification in higher education despite internet challenges in Africa.

Covid-19 brought about the largest online teaching and learning movement in the history of mankind. There are strong indications that online teaching and learning would outlive this pandemic and constitute a significant part of educational programmes at universities post Covid-19. As discussed earlier, the responses of institutions around the world, and in some African countries, to ensure that the 2020 academic year was not thwarted were laudable. The pandemic has proven that Africa needs to develop its own technological infrastructure to be better positioned for not just a time of crisis, but also as a normal practice in teaching and learning ([Ngwacho, 2020](#)). [Vegas and Winthrop \(2020\)](#) argue for the smart use of technology and data, without replacing analog approaches. They further argue that the innovative use of technology, especially mobile phones—as an example of strategies that have emerged during the pandemic, could complement teaching and learning if sustained.

[Le Grange \(2020\)](#) notes that the Covid-19 pandemic could herald a new ranking game that would assess universities based on how well they are positioned for online teaching and learning. He further sees the pandemic as an opportunity for universities to increase investments in online education delivery, and that he foresees historically advantaged institutions leading the race while others play a catch-up game. Meanwhile, [Hodges et al. \(2020\)](#) warn against a low quality of courses offered because of swift adoption of online mode of education. Rather, online teaching and learning should be well-planned and meaningful to harness its full benefits. That said, technological efforts might not be sustainable if educators and learners are not supported to embrace the benefits of technology.

Innovative modes of education delivery

Du Preez and Le Grange (2020) and Vegas and Winthrop (2020) have called for an emphasis on the heart of the teaching and learning process, which Vegas and Winthrop (2020, para. 25) refer to as the “instructional or pedagogical core”. This implies that the way educators engage with students, educational technology and educational materials is crucial for learning, considering that educators are the most important role players in determining student learning. Simply put, instructional core is the interaction between educator, learners and educational materials. For example, high quality learning materials will not on their own elicit quality learning experience, until educators effectively engage them for teaching. According to Vegas and Winthrop (2020), one of the reasons many education innovations failed, despite painstaking effort, is because inadequate attention is not given to instructional core.

Addressing issues inequality, equity, and access in educational systems

The pandemic accentuated the already existing issues of equality, distributive justice and epistemic access in African societies. In relation to epistemic access, Du Preez and Le Grange (2020) argue that ensuring epistemic access in the time of emergency remote teaching (ERT) demand a great deal of work from academics. Lecturers need to bear in mind the various contexts in which students are learning. To therefore ensure epistemic access during and after Covid-19, “an expended notion of emergency remote teaching and learning” (Du Preez and Le Grange, 2020, p. 12) must be considered. This will involve employing other forms of learning to complement online teaching and learning such as interactive print materials and repository texts.

Du Preez and Le Grange (2020), clarify that face-to face teaching and learning does not guarantee epistemic access, neither does it remove historical disadvantages; nevertheless, it at least provides a levelling ground for all students to have the same access to the resources provided by a university, which might be harder to achieve through online teaching and learning. Universities should therefore devise strategies to ensure that students’ epistemic access during online teaching and learning match up with what is obtained during face-to-face sessions, if it cannot surpass it. Vegas and Winthrop (2020) do expect that education systems would be transformed in a way that they address inequality and enable all young people to acquire the skills needed to build a better future for themselves and their communities. Many proponents of education also have high hopes that a more equal education would emerge out of the pandemic.

Conclusion

This chapter has analyzed how African countries have navigated the pandemic and the implications for how the countries around the world can better plan for possible future pandemics. The shift from face-to-face contact to online learning allowed higher institutions to pursue the aspects of technology-based learning they have found most useful (Ali, 2020; Daniel, 2020). Covid-19 did not only show how higher education navigated the pandemic with the use of technology, but it also demonstrated the inadequacies and unpreparedness of the higher education systems in Africa to deal with the pandemic (Mutinda & Liu, 2021; Bolumole, 2020; Demuyakor, 2020). However, Kopp et al. (2019) believe the environment is optimal for digital transformation of higher education.

Muftahu (2020) also noted that the challenges caused by the Covid-19 pandemic have influenced opportunities to practice online and distance learning, with immediate and long-term changes that can transform the overall education system beyond the Covid-19 pandemic for the better. Although this chapter noted some challenges in access to technology and usage in higher education in Africa, it appears that the social, and mental health issues of students associated with the pandemic were some of the contributing factors to their inability to use the available technology (Owusu-Fordjour et al., 2020; Zhai and Du, 2020).

African higher education systems have come through a complex historical path of challenges and transformations and these dynamics seems to have created an imperfect Africa (Woldegiorgis, 2021). Covid-19 has forced innovations into the core functions of almost all education systems (Vegas and Winthrop, 2020). Reimers and Schleicher (2020), based on their survey of educators and education administrators in 59 countries, equally submit that the pandemic has revealed the massive possibilities for innovation which were previously dormant in the education systems. However, we wonder how technological innovation will be possible in Africa with so much diversity and inequalities. In this regard, Toure (2021) suggests that applying a diversity, equity, and inclusion lens to all policies and programs at national and regional levels will assist in the embracing of technological innovation and making it a reality. In addition, with the advancing of the Fourth Industrial Revolution (4IR) and developing environment in which technologies are changing how we live and work (Kayembe and Nel, 2019), Mahlaba (2020) advocates the need for self-directed learning during pandemics.

Similarly, Tamrat (2021) added that although there are evidently long-term effects of COVID-19 on higher education, there are some challenges in access to technology and usage in higher education in Africa. These developments have implications for and provide lessons for how governments should think about future plans for the development of their societies in general, and education in particular, in order to be better prepared for deal with future pandemics. In a world that is characterized by inequalities, in inequities and injustices; and is vulnerable to global hazards such climate change and regional wars, there is a need for more commitment to United Nation’s sustainable development goals (SDGs) in order to bring about a better world that is more equal and just. In this was, we will be better able to deal with future pandemics and associated world disasters.

References

- Aborode, A., Anifowoshe, O., Ayodele, T.I., Iretayo, A.R., David, O.O., 2020. Impact of COVID-19 on Education in Sub-Saharan Africa.
- Abosede, C., Oresajo, O., Akintola, O., 2021. COVID-19 and higher education in Nigeria: the present and the future. *Interdiscip. J. of Educ.* 4, 100–113.
- Adedoyin, O.B., Soykan, E., 2020. Covid-19 pandemic and online learning: the challenges and opportunities. *Interact. Learn. Environ.* 1–13.
- Akin, L., Gözel, M.G., 2020. Understanding dynamics of pandemics. *Turk. J. Med. Sci.* 50, 515–519.
- Bai, H.M., Zaid, A., Catrin, S., Ahmed, K., Ahmed, A., 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int. J. Surg.* 8, 8–17.
- Bond, M., Bedenlier, S., Marín, V. I., & Händel, M., 2021. Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 18(1), 1–24.
- Ciotti, M., Ciccocozzi, M., Terrinoni, A., Jiang, W.-C., Wang, C.-B., Bernardini, S., 2020. The COVID-19 pandemic. *Crit. Rev. Clin. Lab Sci.* 57, 365–388.
- Daniel, S.J., 2020. Education and the COVID-19 pandemic. *Prospects* 49, 91–96.
- Domina, T., Renzulli, L., Murray, B., Garza, A.N., Perez, L., 2021. Remote or removed: predicting successful engagement with online learning during COVID-19. *Socius* 7, 2378023120988200.
- Du Preez, P., Le Grange, L., 2020. The COVID-19 Pandemic, Online Teaching/Learning, the Digital Divide and Epistemological Access. Unpublished Paper.
- Écochard, R., Wimba, P., Bengheya, J., Katchunga, P.B., Lugwarha, S., Oyimangirwe, M., Bazebo, J.-A., Tshilolo, L., Longo-Mbenza, B., Rabilloud, M., 2021. The COVID-19 pandemic is deepening the health crisis in South Kivu, Democratic Republic of Congo. *Int. J. Infect. Dis.* 105, 716–720.
- Egielewa, P., Idogho, P.O., Iyalomhe, F.O., Cirella, G.T., 2022. COVID-19 and digitized education: analysis of online learning in Nigerian higher education. *E-Learn. Digital Media* 19, 19–35.
- El Said, G.R., 2021. How did the COVID-19 pandemic affect higher education learning experience? An empirical investigation of learners' academic performance at a university in a developing country. *Adv. Hum. Comp. Interact.* 2021.
- Emeto, T.I., Alele, F.O., Ilesanmi, O.S., 2021. Evaluation of the effect of border closure on COVID-19 incidence rates across nine African countries: an interrupted time series study. *Trans. Roy. Soc. Trop. Med. Hyg.* 115, 1174–1183.
- Gondwe, G., 2020. Assessing the Impact of COVID-19 on Africa's Economic Development. United Nations Conference on Trade and Development.
- Heyang, T., Martin, R., 2021. A reimagined world: international tertiary dance education in light of COVID-19. *Res. Dance Educ.* 1–15.
- Hodges, C.B., Moore, S., Lockee, B.B., Trust, T., Bond, M.A., 2020. The Difference Between Emergency Remote Teaching and Online Learning.
- Huremović, D., 2019. Brief history of pandemics (pandemics throughout history). In: Huremović, D. (Ed.), *Psychiatry of Pandemics: A Mental Health Response to Infection Outbreak*. Springer International Publishing, Cham.
- Jacob, O.N., Abigeal, I., Lydia, A., 2020. Impact of COVID-19 on the higher institutions development in Nigeria. *Electr. Res. J. Soc. Sci. Humanit.* 2, 126–135.
- Jacobs, C., Kyamusugulwa, P.M., Katembera, R.S., Kintuntu, H., 2021. Containing COVID-19 in the democratic republic of the Congo: government measures and women's compliance. *J. Afr. Law* 65, 287–310.
- Jegade, D., 2020. Perception of undergraduate students on the impact of COVID-19 pandemic on higher institutions development in Federal Capital Territory Abuja, Nigeria. *Electr. Res. J. Soc. Sci. Humanit.* 2, 211–222.
- Karabulut, G., Bilgin, M.H., Demir, E., Doker, A.C., 2020. How pandemics affect tourism: international evidence. *Ann. Tour. Res.* 84, 102991.
- Karemere, H., Makali, S., 2022. Distance learning in times of Covid-19: importance, strengths and limitations - case of the regional school of public health of the catholic University of Bukavu in Democratic Republic of Congo. *Acta Sci. Med. Sci.* 2, 40–47.
- Kathula, D.N., 2020. Effect of COVID-19 pandemic on the education system in Kenya. *J. Educ.* 3, 31–52.
- Kayembe, C., Nel, D., 2019. Challenges and opportunities for education in the fourth industrial revolution. *Afr. J. Public Aff.* 11, 79–94.
- Kibuku, R.N., Ochieng, D.O., Wausi, A.N., 2020. E-learning challenges faced by universities in Kenya: a literature review. *Electron. J. eLearn.* 18, 150–161.
- Kopp, M., Gröblinger, O., Adams, S., 2019. Five common assumptions that prevent digital transformation at higher education institutions. *INTED2019 Proc.* 1, 1448–1457.
- Le Grange, L., 2020. Could the Covid-19 pandemic accelerate the uberification of the university? *S. Afr. J. High Educ.* 34, 1–10.
- Le Nestour, A., Moscoviz, L., 2020. Five Findings From a New Phone Survey in Senegal. Center for Global Development. April 24, 2020. <https://www.cgdev.org/blog/five-findings-new-phone-survey-senegal>.
- Lee, H., Moon, S.J., Ndombi, G.O., Kim, K.-N., Berhe, H., Nam, E.W., 2020. COVID-19 perception, knowledge, and preventive practice: comparison between South Korea, Ethiopia, and Democratic Republic of Congo. *Afr. J. Reprod. Health* 24, 66–77.
- Lone, S.A., Ahmad, A., 2020. COVID-19 pandemic – an African perspective. *Emerg. Microb. Infect.* 9, 1300–1308.
- Lyngdoh, M., 2020. The parallels between COVID-19 and the Spanish flu of 1918. *Int. J. Community Med. Public Health* 7, 4687–4689.
- Magesa, E., Josua, L.M., 2022. Use of technology to Morph teaching and learning in higher education: post COVID-19 Era. *Creativ. Educ.* 13, 846–853.
- Mahlaba, S., 2020. Reasons why self-directed learning is important in South Africa during the Covid-19 pandemic. *S. Afr. J. High Educ.* 34, 120–136.
- Memish, Z.A., Alharthy, A., Alqahtani, S.A., Karakitsos, D., 2021. COVID-19 air travel restrictions and vaccine passports: an ongoing debate. *Travel Med. Infect. Dis.* 42, 102049.
- Menon, K., Motala, S., 2021. Pandemic leadership in higher education: new horizons, risks and complexities. *Educ. Change* 25, 1–19.
- Mhlanga, D., Ndhlovu, E., 2020. Socio-economic Implications of the COVID-19 Pandemic on Smallholder Livelihoods in Zimbabwe.
- Mphahlele, R., Seeletso, M., Muleya, G., Simui, F., 2021. The influence of COVID-19 on students' learning: access and participation in higher education in Southern Africa. *J. Learn. Dev.* 8, 501–515.
- Muftah, M., 2020. Higher education and Covid-19 pandemic: matters arising and the challenges of sustaining academic programs in developing African universities. *Int. J. Educ. Res. Rev.* 5, 417–423.
- Mutinda, G. & Liu, Z., 2021. Perceptions on the implications of the COVID-19 pandemic on university students' wellbeing in Kenya – a thematic analysis approach. *Higher Education Research & Development*, 1–15.
- Ngwacho, A.G., 2020. COVID-19 pandemic impact on Kenyan education sector: learner challenges and mitigations. *J. Res. Innov. Implicat. Educ.* 4, 128–139.
- Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., Agha, R., 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int. J. Surg.* 78, 185–193.
- Obayelu, A.E., Edewor, S.E., Ogbe, A.O., 2020. "Trade effects, policy responses and opportunities of COVID-19 outbreak in Africa," *Journal of Chinese Economic and Foreign Trade Studies*, Emerald Group Publishing, vol. 14(1), 44–59.
- Ojo Joseph, I., Femi Barnabas, A., Olukemi Grace, A., Mathew, R.E., Henry, N.-L.C., Tunde, A., Isola, L.A., 2021. COVID-19 pandemic: Nigerian University lecturers' response to virtual orientation. *Cogent Arts Humanit.* 8, 1932041.
- Owusu-Fordjour, C., Koomson, C.K., Hanson, D., 2020. The impact of Covid-19 on learning-the perspective of the Ghanaian student. *Eur. J. Educ.* Vol 7, pp 89–101.
- Petersen, T., 2020. UCT Provides Pre-paid Data at No Cost to Students for Online Learning. News24. Available at: <https://www.news24.com/SouthAfrica/News/uct-provides-pre-paid-data-at-no-cost-to-students-for-online-learning-20200421>.
- Piret, J., Boivin, G., 2021. Pandemics throughout history. *Front. Microbiol.* 11.
- Pokhrel, S., Chhetri, R., 2021. A literature review on impact of COVID-19 pandemic on teaching and learning. *High. Educ. Future* 8, 133–141.
- Reimers, F., Schleicher, A., 2020. Schooling disrupted, schooling rethought. In: *How the COVID-19 Pandemic is Changing Education*. Retrieved December, 14, 2020.
- Sasidharan, S., Dhillon, H., 2020. Medical training for undergraduates in the time of COVID-19: education cannot wait in DR Congo. *Rev. Chil. Anest.* 49, 770–773.
- Tamrat, W., 2021. Enduring the impacts of COVID-19: experiences of the private higher education sector in Ethiopia. *Stud. High Educ.* 46, 59–74.

- Toure, K., 2021. The Fourth Industrial Revolution (4IR) in Africa: A Dialogue on Policy Perspectives and Investment Options.
- Uchendu, E., Nnabueze, A., Onogwu, E., 2021. Locked down during the lockdown. In: Alliance for African Partnership Perspectives, vol. 1, pp. 101–106.
- UNESCO, 2020. Rallies International Organizations, Civil Society and Private Sector Partners in a Broad Coalition to Ensure #LearningNeverStops. UNESCO. March 26, 2020. <https://en.unesco.org/news/unesco-rallies-international-organizations-civil-society-and-private-sector-partners-broad>.
- United Nations, 2020. Policy Brief: Education During Covid-19 and Beyond. United Nations Sustainable Development Group (UNSDG). August 2020. <https://unsdg.un.org/resources/policy-brief-education-during-covid-19-and-beyond>.
- Vegas, E., Winthrop, R., 2020. Beyond Reopening Schools: How Education Can Emerge Stronger Than Before COVID-19. Brookings Institution. September 8, 2020. <https://www.brookings.edu/research/beyond-reopening-schools-how-education-can-emerge-stronger-than-before-covid-19/>.
- Vegas, E., 2020. School Closures, Government Responses, and Learning Inequality Around the World During COVID-19. Brookings Institution. April 14, 2020. <https://www.brookings.edu/research/school-closures-government-responses-and-learning-inequality-around-the-world-during-covid-19/>.
- Wekullo, C., Shiundu, J., Ouda, J., Mutevane, A., 2022. COVID-19 Crisis and the Future of Higher Education: Perspectives From Kenya. In: STAR Scholar Book Series, pp. 73–91.
- Woldegiorgis, E.T., 2021. Configurations of progress and the historical trajectory of the future in African higher education. *Educ. Philos. Theor.* 1–15.
- Yunusa, A.A., Sanusi, I.T., Dada, O.A., Oyelere, S.S., Agbo, F.J., Obaido, G., Aruleba, K., 2021. The impact of the COVID-19 pandemic on higher education in Nigeria: university lecturers' perspectives. *Int. J. Educ. Dev.* 17, 43–66.
- Zhai, Y., Du, X., 2020. Addressing collegiate mental health amid COVID-19 pandemic. *Psychiatr. Res.* 288, 113003.

The impact of COVID-19 on higher education in Africa

Philip O. Ayoo, School of Computing and Engineering, Uganda Technology and Management University (UTAMU), Kampala, Uganda

© 2023 Elsevier Ltd. All rights reserved.

Introduction	91
Higher education in Africa	91
Pandemics	92
Impact of COVID-19 pandemic	93
Responses to the COVID-19 pandemic	93
Effect of responses on higher education institutions	94
Impact on core missions of universities	94
Teaching and learning	94
Research	95
Outreach	95
Economic impact	95
Effect on inequalities within faculty and student demographics	96
Strategies to mitigate COVID-19 effects	97
Conclusion and lessons learned	98
References	99

Introduction

Pandemics—defined as large-scale outbreaks of infectious diseases that can greatly increase morbidity and mortality over a wide geographic area and cause significant economic, social, and political disruption—have increased in number over the past century because of increased global travel and integration, urbanization, changes in land use, and greater exploitation of the natural environment (Madhav et al., 2017). Since December 2019, COVID-19, a highly infectious disease caused by a new virus, has had an increasing impact on countries and regions around the world. Starting initially as a health crisis, the COVID-19 pandemic has quickly shown a socio-economic dimension (PWC, 2020). Consequently, schools, universities and other higher education institutions have been forced to suspend campus-based operations, such as teaching and research, with significant impact on countries' competitiveness and their ability to foster advanced skills and regional development (Arnhold et al., 2020).

Among the measures targeted at limiting the spread of the virus have been: bans on large gatherings, remote working and schooling, prescribing hygiene protocols at home, in hospitals and public places, through the use of masks and protection equipment, and cleaning of public spaces (PWC, 2020). A large number of schools and educational systems have had to adopt internet-based methodologies to support their students via distance education solutions during the period of school closures (WHO, 2020a), making disparities in distance education more evident in its wake. The risk of growing inequality is emerging also from the results of regional analyses, with higher education institutions in Africa reporting more difficulties and negative effects than higher education institutions in other regions. These divides, laid bare by COVID-19, are a major obstacle in the drive for inclusive and equitable access to higher education by 2030, which is a target of Sustainable Development Goal 4 (ACU, 2020).

In this article, we take a look at the higher education landscape in Africa, examining disruptions to education caused by the COVID-19 pandemic outbreak, and the key educational processes affected, including higher education business. We do this through a desktop analysis of secondary data wherein systematic literature, policy documents, as well as related models between COVID-19 and higher education are reviewed. With this review, we aim to contribute to the discussion, by recording how the various stakeholders within higher education responded to the pandemic and highlighting the real and estimated effects of the pandemic on higher education institutions, consequences and lessons of the responses to the pandemic.

Higher education in Africa

Higher education serves as the principal venue for knowledge creation and economic and social development (Aina, 2010) and occupies a very special place in any society's development due to its function of fostering the capacity of individuals and communities to promote sustainable development. As aptly described by Teffera and Altbach (2004, p. 3), higher education in Africa is "as old as the pyramids of Egypt, the obelisks of Ethiopia, and the Kingdom of Timbuktu." Higher education in Africa plays a critical capacity building and professional training role not only in national development but also in support of all the Millennium

Development Goals (MDGs) (Materu, 2007) and subsequently, Sustainable Development Goals (SDGs). Its central role in the economic and political development of the continent in an increasingly globalizing knowledge society is widely acknowledged (Cloete et al., 2011).

Globally, the pressure on higher education systems to provide access will continue unabated in the years to come due to the rapid growth of the sector, particularly in Asia and Africa (Bloom et al., 2005). Major transitions and reforms in African education systems have taken place focusing on access, equity, quality, affordability and relevance to the national philosophy and needs. Various policies and legislations have been put in place to enhance the expansion of higher education institutions, through both public and private means, to meet the aspirations of rapidly growing national populations and economies.

Pandemics

The late 20th and early 21st centuries have seen the rapid transmission and difficulty of containing various diseases. There have been five global health emergencies since 2005 when the declaration was formalized: swine flu in 2009, polio in 2014, Ebola in 2014, Zika in 2016 and Ebola again in 2019 (Mhlanga and Moloi, 2020). According to Bloom et al. (2020), many epidemics and pandemics are ongoing, claiming millions of lives per year (such as HIV/AIDS, malaria, and tuberculosis); some are on the rise in terms of infections and deaths (such as dengue fever and various forms of hepatitis); some have nearly been eradicated but with a dangerous prospect of re-emerging should containment efforts be relaxed (such as measles and poliomyelitis); some have abated or were successfully contained (such as SARS, Ebola, H1N1 influenza, and Zika); and some have frequent outbreaks (such as cholera).

But what exactly is a pandemic, and how does it differ from other terms usually used in health cycles—epidemic, endemic, and outbreak? According to the World Health Organization (WHO, 2020a), an epidemic is a disease that affects a large number of people within a community, population, or region. A pandemic, on the other hand, is an epidemic that has spread over multiple countries or continents, defined by the Merriam-Webster dictionary as a global outbreak of a disease from a new virus. An outbreak is defined by WHO (2020a) as a greater-than-anticipated increase in the number of endemic cases, an endemic being something that belongs to a particular people or country (WHO, 2020a). A pandemic, therefore, is an epidemic occurring over a very wide area, crossing international boundaries, and usually affecting a large number of people. Accordingly, pandemics are identified by their geographic scale rather than the severity of illness, and the term is used to describe a disease that affects a whole country or the entire world (Centers for Disease Control and Prevention, 2020; Madhav et al., 2017; WHO, 2020a). For example, in contrast to annual seasonal influenza epidemics, pandemic influenza is when a new influenza virus emerges and spreads around the world, and most people do not have immunity (WHO, 2010). Pandemics, therefore, are large-scale outbreaks of infectious disease that can greatly increase morbidity and mortality over a wide geographic area and cause significant economic, social, and political disruption.

Multiple outbreaks have exposed gaps related to the timely detection of disease, availability of basic care, tracing of contacts, quarantine and isolation procedures, and preparedness outside the health sector, including global coordination and response mobilization (Moon et al., 2015; Pathmanathan et al., 2014). These gaps are especially evident in resource-limited settings and have posed challenges during relatively localized epidemics, with dire implications for what may happen during a full-fledged global pandemic (Madhav et al., 2017). Evidence suggests that the likelihood of pandemics has increased over the past century because of increased global travel and integration, urbanization, changes in land use, and greater exploitation of the natural environment (Jones et al., 2008; Morse, 1995). These trends are likely to continue and will intensify (Madhav et al., 2017). Significant policy attention has focused on the need to identify and limit emerging outbreaks that might lead to pandemics and to expand and sustain investment to build preparedness and health capacity (Smolinsky et al., 2003).

The emergence and rapid transmission in the late 2019 and early 2020 of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the steep increase in the prevalence of the disease it causes (coronavirus disease 2019 or COVID-19) forcefully reminded scientists, policymakers, and the wider public of the world's vulnerability to infectious disease pandemics (Bloom et al., 2020). Within a few weeks, we moved from a few reported cases of a new pulmonary disease in the city of Wuhan in China to a state in which almost all countries reported infections and deaths, billions of people could only leave their homes for essential purposes, economic activity nosedived, and unemployment figures skyrocketed.

However, the international community has made significant progress toward preparing for, and mitigating, the impacts of pandemics. Crawford (2020) characterized four phases of pandemic response: rapid adaptation, improvement, consolidation, and restoration. The goal of the first phase is "to rapidly adapt core business for the new context"; the second phase is "to optimize the adapted core business to improve quality and begin to consider non-core activities"; the third phase is "to evaluate pre-pandemic measures of social, economic, and environmental success"; and, the fourth phase is "to determine what a return to business-as-usual looks like, and how it can occur" (Crawford, 2020, n.p.). It is recognized that institutions will likely go through these stages at different speeds, and some institutions may move fluidly between phases. Likewise, it is possible for an institution to exist inside of two phases, depending on their method of engagement and organizational strategy (Butler-Henderson et al., 2020).

Impact of COVID-19 pandemic

Since COVID-19 is not just a medical pandemic, but a social event that is disrupting our social order, tackling it has presented a humanitarian and development crisis that is threatening to leave deep social, economic and political scars for years to come (Teti et al., 2020). The education sector has not been immune from the impact of COVID-19, as it has affected all levels of global education systems from the pre-school to the university and has caused cancellation or postponement of academic conferences (Oyedotun, 2020). Several individuals, organizations and national entities have conducted surveys on the impact of COVID-19 on higher education at the institutional, national, regional and global levels. Other studies have been undertaken by academic researchers to understand the economic impact of this pandemic. Bloom et al. (2020) discuss and review literature on the macro-economic effects of epidemics and pandemics since the late 20th century. Philipson (2000), Malani and Laxminarayan (2011), Klein et al. (2007), Gersovitz (2011) and Rasul (2020) have all conducted surveys on the economics of infectious diseases with more focused treatments of the earlier literature. Brodeur et al. (2020) provide a survey of the emerging economic literature on COVID-19, and a wealth of research has been published as part of the Center for Economic Policy Research's "COVID Economics" series of vetted and real-time papers (Wyplosz, 2020). From all these studies, key issues have emerged that may be used to evaluate the exact impact of the COVID-19 pandemic on the various aspects of higher education.

Responses to the COVID-19 pandemic

The COVID-19 pandemic has highlighted the importance of rapid and well-coordinated responses to a global situation that has put pressure on every aspect of the health system, education, research, and practice (Lackie et al., 2020). Poor emergency responses were seen in the beginning of the outbreak (Muftahu, 2020), with some national governments underplaying the severity of the pandemic. In terms of policy responses, countries tended to limit themselves to three areas: (a) administrative measures to safeguard the operation of the system; (b) financial resources; and (c) the provision of resources to continue education activities (UNESCO-IESALC, 2020). Immediate economic responses implemented by countries can be highlighted in four categories, namely, restriction of movement, health measures, social measures and economic measures taken to reduce the socioeconomic burden on citizens while paving the way for economic recovery (AUC, 2020). From the beginning, institutional responses have covered different areas: the strictly health front, the adjustment of calendars, research and development contribution to mitigate the pandemic, the guarantee of continuity of teaching activities through distance education, bibliographic and technological resources, as well as socio-emotional support to the university community (UNESCO-IESALC, 2020).

Immediate responses to the COVID-19 outbreak within most countries have been developed primarily to control and curb the spread of the infection, and have led to the closure of entire countries (Nkhoma, 2020), as well as measures of social distancing, widespread lockdowns, and restrictions on traveling, movements, and gatherings (WHO, 2020a). From the responses, it is clear that the strategy for dealing with COVID-19 is context-specific as every country and institution is different (Nkhoma, 2020). In their report, the African Union Commission (AUC, 2020) found some striking differences on how COVID-19 is behaving in different countries. For instance, in Italy, there had been strong curtailing of social interactions yet COVID-19 mortality still remained high. In contrast, Japan had some of the earlier cases, but the mortality remained low despite not having adopted some of the more restrictive social isolation measurements. These puzzling differences have been adjudicated to different cultural norms as well as differences in medical care standards. To some extent, therefore, the effectiveness of the governmental responses could explain the difference in the magnitude of the impact of the pandemic between countries. Therefore, individual countries and/or institutions need to develop their own ways to deal with the pandemic (AUC, 2020), including clear social justice principles of accessibility, inclusivity, flexibility, equitability and connectivity (Zezeza, 2020).

Considering the outbreak does not respect borders, the inability of one country to fight and contain the virus will negatively impact its neighbors over time, thus threatening livelihoods and the growth of economies (AUC, 2020). Hence, it was necessary that Africa quickly and effectively develops a coordinated response to collectively fight the impact of the pandemic on the continent. Consequently, in addition to the policy actions at the national level to contain the spread of COVID-19 and mitigate its socioeconomic impact, the African Union has been at the forefront of the fight against the pandemic by coordinating the activities and raising resources to strengthen the continental response to COVID-19, maintain economic activities and revive African economies.

There have also been significant contribution of sub-regional, continental and international bodies in supporting HEIs and governments to respond to the impact of the pandemic through training, collecting data and monitoring the impacts of COVID-19 on higher education, hosting resource pages, recommending policies, and providing support materials for governments and other higher education stakeholders (Agyapong et al., 2020, p. 16). These include the Councils of Tertiary Education, the Inter-University Council for East Africa (IUCEA), the Association of African Universities, Association for the Development of Education in Africa (ADEA), the Regional Universities' Forum for Capacity Building in Agriculture (RUFORUM), the Southern African Regional Universities Association (SARUA), UNESCO, the International Association of Universities (IAU), the World Bank, among others. Furthermore, private companies, particularly telecom companies, are increasingly partnering with TELs and government to overcome the critical challenges by negotiating zero-rated access to specific educational and information websites. For example, in Senegal, dominant ICT player Sonatel has moved to address this issue, giving students the chance to activate a free 1-GB education pass (Agyapong et al., 2020, p. 17). Similar partnership agreements have been made between telecom companies and universities in countries such as Ghana, South Africa and Kenya.

Effect of responses on higher education institutions

In Africa, the policy measures taken by African Union member states have resulted in the disruptions of economic and social activities and the educational system was severely affected due to the total closure of learning institutions, schools, colleges and universities by governments across the continent. By end of March 2020, all African countries but one had countrywide closures of education institutions (AUC, 2020), forcing students out of their physical learning spaces, and confining lecturers and professors to their homes (Lues et al., 2020). The first immediate response to the lockdowns by HEIs has logically been to try to substitute all mobility with virtual activities. Site and exposure visits, training and education, workshops and consultancies have been replaced overnight by Zoom meetings, webinars, distance education and endless Skype calls (Aarts, 2020). Still, the general experience is that virtual cooperation also has many limitations. Moreover, precisely those groups that are often targeted in capacity building efforts, such as socio-economically vulnerable groups, women, girls, youth from poorer backgrounds, etc, tend to be less virtually connected. Now, therefore, is the time for Africa to build internationalization-at-home, by crafting relevant policies and programs, since the COVID-19 crisis calls for enhancing of intra-African mobility as well as internationalizing higher education curricula to attract international students globally (Waruru, 2020a).

While remote work was a feature of multinational firms and open source communities among others, the pandemic effectively forced “all remote all the time”, resulting in a scramble to adapt to remote collaboration and its technological infrastructure (George et al., 2020). The conditions under which higher education all-of-a-sudden had to perform research and what is now often referred to as “emergency online education,” demanded the reskilling of lecturers and instructors in modes of online teaching, assessment and mentorship (Marinoni et al., 2020). Perhaps on a positive note, though, the pandemic has forced universities to devise strategies and solutions to ensure quality provision, and students and higher education institutions have learned to adjust and ensure continued teaching and learning (Makoni, 2020).

Impact on core missions of universities

The traditional mandates of universities revolve around teaching/learning, research and community engagement or outreach. The important question, therefore, is how the COVID-19 pandemic has altered the main processes in education, including academic recruitment, academic management, teaching and learning, study and advancement, student life and welfare; and on the other hand, how COVID-19 has compelled the institutions of higher education to implement new approaches, and to let go of their current teaching practices (Abdullah et al., 2020). With the pandemic pushing education entirely online, universities must acknowledge that the game has changed and try new ways to progress with virtual learning.

Teaching and learning

Teaching and learning have been impacted by the COVID-19 crisis in the most direct way, especially due to the absence of options for face-to-face interactions for an extended period (Arnhold et al., 2020). Before the outbreak, there were only a few fully online universities in Sub Saharan Africa, including the African Virtual University (AVU), the Kenyatta Digital School of Virtual Learning, Laweh Open University, the Open University of Tanzania, and the University of Rwanda's e-learning platform (Agyapong et al., 2020, p. 15). According to the World Bank, these institutions have the existing capacity and experience to offer online programs, but they are mostly targeted toward students who want to upgrade their skills while they are already employed, though their mandates continue to expand (World Bank, 2020).

Whereas some universities that had adequate digital infrastructure have been swift in transitioning to online teaching and learning (Agyapong et al., 2020, p. 14), switching to the online mode of teaching and learning has revealed several global and local challenges that universities are facing right now. Among global challenges, the following are related to the online infrastructure: weakness of internet connection and internet speed in many countries; high prices for a good internet connection; absence of computers and other gadgets that support online teaching and learning; and many online instruments, platforms, and websites crashed when an unexpectedly high number of clients connected to them. With governments and HEIs adopting remote teaching to support distance learning and online education delivered through ICT, learners from poor, vulnerable and marginalized households have no access to these mediums of learning, further worsening inequality in access and quality of education (Areba, 2020). In order to mitigate this, most faculty had to do significantly more work to move courses online, rethink curriculums and assessment, learn to use software and technology for teaching and learning, and make sure that all their students were included and were coping.

On the other hand, the COVID-19 outbreak accelerated and forced the adoption of digital solutions at an unprecedented pace, creating unforeseen opportunities for scaling up alternative approaches to social and economic life (Hantrais et al., 2020), despite the digital risks and threats (Giordani and Rullani, 2020). The intense exploration of digital tools and their applications caused by COVID-19 represented a leap that would have taken years to be realized in normal times (George et al., 2020), due to experimenting without any safety net. Schwab and Malleret (2020) acknowledge that the world was facing a “defining moment” as the pandemic precipitated the fusion of technologies, enabling digital technologies to extend their reach, almost uncontrolled, into every aspect of life.

The transition to online learning has laid bare the digital divide within the African continent—between those countries that have better ICT infrastructure than others; between higher education institutions within the same country, with some being far better equipped and experienced than others; and between students within the same institution; the rich who live in urban areas; and the poor in rural areas who can barely afford to access the internet, when and if it is available (Mohamedbhai, 2020). To mitigate

the challenges that come with the digital divide, some African HEIs in countries such as Ghana, South Africa, and Rwanda among others, partnered with Telco's to zero-rate these platforms in some instances facilitating universities to make data packages and laptops available to some students to improve access (Agyapong et al., 2020, pp. 14–15).

Research

The most common impact of COVID-19 has been the stoppage of many scientific research projects, canceling of international travel and postponement of scientific conferences (Marinoni and de Wit, 2020). The pandemic has particularly affected early career researchers (ECRs) for whom precarity has manifested in many ways (Hooper and Djerassimovic, 2020)—from uncertainty regarding how they will support themselves when they do not complete their doctorate within prescribed timeframes, to the insecurity of undertaking multiple fixed-term contracts as postdocs (McKie, 2020). A recent survey (Vitae, 2020) highlighted that around two-thirds of ECRs are very worried about their future and finances, with only 10% of those whose contracts would end in 2020 having been granted extensions in the context of the pandemic, and only 12% of final-year doctoral students reporting that their institution had provided an option to extend their studies.

One of the few positives, however, has been the speed at which many global partners joined to battle the disease (Ferdig et al., 2020), with academic researchers, and even academic journals, joining in the fight. For instance, in addition to giving open access to articles, many medical journals switched to a speedier review to be able to quickly publish promising results. Therefore as researchers were making early discoveries, they had a way to bypass a traditionally longer review and publication process to give hope, share building blocks, and encourage collaboration.

Outreach

For the large majority of HEIs, COVID-19 has impacted their community engagement initiatives, but largely in positive ways—the crisis has increased HEIs' community engagement (Wachira and Ombati, 2020). African universities have responded to the pandemic with many innovations to solve local problems, with many public universities directly helping to fight the pandemic in different innovative ways—producing much-needed hand sanitizers and face masks, developing a prototype ventilator that could be affordably mass produced, undertaking joint vaccine research with other partners, developing a variety of COVID-19 fighting tools including a portable solar-powered ventilator, a contact tracing mobile application and an automatic solar-powered hand-washing machine. These universities are being true to their visions and missions and doing exactly what society expects them to do. On the e-learning front, some public universities have been doing fairly well (Wachira and Ombati, 2020), with a number of them already having active online and distance learning departments and were not necessarily starting from scratch during the pandemic. Many had already invested in e-library resources where they subscribed to various e-books and e-journals that are especially useful to their postgraduate students.

Regarding partnerships, many universities around the world have traditionally supported, in one way or another, partner institutions in poorer countries and regions, often in the Global South. Usually the aim is to strengthen the capacity of these partners to offer higher education of good quality, to underscore local development agendas and, along the way, contribute to the global sustainable development agenda. However, the COVID-19 crisis has forced universities everywhere into crisis mode overnight and meant that they have had to focus on keeping their own education and research running in the first instance, and capacity building efforts supporting far-away partners tended not to be part of the core concern (Aarts, 2020). In addition, funding for such efforts became less reliable as the upcoming economic crisis has affected international cooperation budgets. There is now recognition of the need to strengthen both local and regional collaborations, while leveraging the existing, strong international collaborations that many institutions have (Lues et al., 2020). The multi-, inter- and trans-disciplinary approach should foster closer collaboration between academia and university business entities and support the representation of academia on wider societal platforms. Further, partnership projects between universities and the private sector and civil society organizations are encouraged.

Economic impact

What started as a health crisis has quickly precipitated into a worst economic, social and humanitarian crisis, disrupting millions of people's livelihoods, affecting poor households and damaging formal and informal businesses and government revenues around the world (AUC, 2020). According to SPAU (2020), the socio-economic impact of the COVID-19 pandemic operates through two distinct channels: first are the direct and indirect effects of the sickness, which results from when an income-earner in the household falls ill, and the ratio of active members to dependents falls; second are aversion behavior effects resulting from the fear of catching the virus, which in turn leads to a fear of association with others and reduces labor force participation, closes places of employment, disrupts transportation, motivates some governments to close borders and restrict entry of citizens from afflicted countries, and motivates private decision makers to disrupt trade, travel, and commerce by canceling scheduled commercial flights and reducing shipping and cargo services.

The African Union Commission report (AUC, 2020) projected that the global economic activity would contract by 6% in 2020 (optimistic scenario), and 7.6% (a second wave of infection before the end of that year, according to OECD), while WEO estimated a 4.6% fall, much worse than during the 2008/2009 financial crisis, as per the most recent estimates by the IMF World Economic Outlook for June 2020. For HEIs, the pandemic would have immense financial consequences, as universities are faced with massive losses in revenues after shifting learning online, with queries arising over whether or not all students would return

from their home countries after the pandemic (Halterbeck et al., 2020; Maslen, 2020). For one, it became unclear whether universities could continue charging normal fees for programs that had shifted online, especially for those who had paid for face-to-face learning. At the same time, universities had already suffered very significant revenue losses in respect of accommodation, conferences, and events activity. Furthermore, due to the pandemic, the ability of parents to pay fees as before was likely to be lower, affecting universities' finances.

While public universities might have a strategy for accessing support to deal with reopening, the private institutions are less likely to have access to such funds. However, budget cuts are still to be expected for higher education funding moving forward particularly for public universities mainly because a number of economies are still recovering from the past global financial crisis (Jegede, 2020). Consequently, there is a growing need for universities to reduce their dependence on external funding and international philanthropy and pursue sustainable, locally-focused funding strategies. There is need for universities in the region to closely examine the tenets of "Africapitalism", an economic philosophy developed by Elumelu (Dell, 2020) predicated on the belief that Africa's private sector can and must play a leading role in the continent's development.

For many universities and countries, international student tuition fees are a critical source of revenue (AUC, 2020). However, under the pressure of the pandemic, promoting global mobility has become an emerging issue in higher education settings. Although various methods of enhancing student mobility have been implemented, little research has as yet confirmed the pandemic challenges for students (Wu et al., 2020). Furthermore, the COVID-19 pandemic has come with a lot of uncertainties for the higher education sector and for internationalization and mobility (Waruru, 2020b), including bringing in new conditions such as COVID-19 vaccination compliance certificates targeted at Africa, possibility of racial discrimination (e.g., as seen in the "black lives matter" movement), with learners from poorer parts of the world that mainly act as sources of international students being seen as "unsafe".

The immediate impact of COVID-19 on income and financial circumstances of students shows how many students are living in a fairly precarious financial situations, where it only takes a small change for many of them to suffer extreme financial hardship (O'Malley, 2020). In addition to financial challenges and concerns about degree completion, many students are experiencing stress and anxiety related to the cost of living, loss of income and increasing demands as parents, children and community members (Jaramillo and Stephenson, 2020). For international students, "taking a semester off" has visa implications and may require leaving the country of their studies. Moreover, for students who live in university housing and residences, taking a break would jeopardize their housing situation.

Effect on inequalities within faculty and student demographics

The COVID-19 pandemic is transforming society in profound ways, often exacerbating social and economic disparities and inequalities in its wake (Engzell et al., 2020), and revealing, reinforcing, and catalyzing new social and cultural relations, laying bare inequalities and anxieties, discrimination and division (Leach et al., 2020). The pandemic has exposed "fault lines of inequalities", both in society and in higher education institutions (Waruru, 2020b), revealing the glaring inequalities of access along the differentiations of class, gender, ethnicity, religion, age, location and other social markers (Zezeza, 2020). An equality agenda therefore needs to be supported everywhere in order for the innovation needed to rebuild our society is done by and with all, and not only by those who will be able to afford it (Marinoni et al., 2020).

As HEIs prepare to resume activities fully, they must take into account the fact that increased vulnerability to COVID-19 may pose a challenge for some of their older staff members (Odhiambo, 2020a). According to the United States Centers for Disease Control and Prevention (2020), older adults (60 years and above) are at a higher risk of falling victim to COVID-19 and suffering more serious symptoms. However, throughout the world universities are characterized by significant numbers of older people—senior professors, deans, principals and vice-chancellors—all of whom are recognized as some of the most important people in higher education institutions by virtue of their disciplinary expertise and experience. Strategies must be identified to ensure that the knowledge and experience of the older faculty is not lost, while at the same time ensuring their safety.

On the digital migration front, universities around the world were largely unprepared for a sudden switch to online learning prior to the COVID-19 pandemic—with only a third providing full online courses in some form, according to data collected for the 2020 edition of "U-Multirank." The data showed 60% of higher education institutions reporting online learning provisions in their strategic planning prior to COVID-19, but only one-third appearing to provide full online courses in some form. Some universities, especially those well-endowed financially, were more able to transition to online learning thanks to their superior infrastructure (Mitchell, 2020).

Specifically, Africans are paying some of the highest rates in the world for internet access as a proportion of income, according to an October 2019 "Affordability Report" produced by the Alliance for Affordable Internet. While mobile phone penetration in Africa is very good, including with internet connectivity, there is little accessibility to fixed broadband (MacGregor, 2020), even in Africa's most advanced countries. Many students are still unable to take part in online learning due to a lack of gadgets and electricity, and due to a lack of access to the internet (Waruru, 2020a). These vulnerable students, such as for instance those from less advantaged backgrounds, are especially likely to fall behind during the pandemic period, as they are less likely to have access to relevant learning digital resources (e.g., laptop/computer, broadband internet connection) and less likely to have a suitable home learning environment (Di Pietro et al., 2020). A lot more focus should therefore be given to bridging the digital divide, to ensure that no young African is left behind.

It is projected that the hampering of international student and academic staff mobility by COVID-19 will most likely diminish interest in global rankings. At the same time, national and specialized (by subject) regional ranking are projected to gain importance (Siwinski and Bilanow, 2020). For the very same reason, indicators linked to internationalization of the student body and academic staff may lose some of their weight in the methodology used by rankings, and instead, new indicators related to university presence on the internet will appear. Here, Isidro Aguillo from Spain, who published the first internet-based “Webometrics” ranking as far back as 2004, stands out as a forerunner. According to the “IREG Observatory on Academic Ranking and Excellence,” perception of rankings will likely be shifting in various directions. For instance, prospective students, double hit once by travel restrictions and then by a deteriorating economic situation, will be inclined to pay more attention to national and regional rankings than global.

On the gender front, the UNESCO Global Education Monitoring Report 2020 titled “Inclusion and Education: All means all,” indicates that just 11% of African countries had formulated comprehensive tertiary education equity and inclusion strategies. Noteworthy in the time of COVID-19 is that about 40% of low- and lower middle-income countries had not taken any measures to support learners at risk of exclusion during the crisis (Ligami, 2020). From the experience of the gender impacts of the Ebola virus disease in West Africa 2014–2016, the COVID-19 pandemic would affect women negatively and disproportionately both directly and indirectly due to gender power relations in decision-making (SPAUI, 2020). According to Siwinski and Bilanow (2020), the COVID-19 pandemic has placed young women researchers in a particularly disadvantaged position as female scientists, and especially those with kids, must share time between research and family duties. Furthermore, women often are active in areas of work that are the first to fall away in economic crises (AUC, 2020). As the Ebola outbreak demonstrated, when women have to give up work during times of crises, they will be confronted with more challenges than men when trying to spring back after the crisis. In addition, the COVID-19 pandemic worsened the issue of gender-based violence, which increased by 25% as lockdown measures were implemented (AUC, 2020).

Prolonged school closure and home confinement during a disease outbreak might have negative effects on children’s physical and mental health (Brazendale et al., 2017; Brooks et al., 2020), which might be exacerbated by the lack of peer support and alternatives for mitigation of risks (SPAUI, 2020). Evidence suggests that when children are out of school (e.g., weekends and summer holidays), they are physically less active, have much longer screen time, irregular sleep patterns, and less favorable diets, resulting in weight gain and a loss of cardiorespiratory fitness (Wang et al., 2019). Such negative effects on health are likely to be much worse when children are confined to their homes without outdoor activities and interaction with same aged friends during the outbreak. Furthermore, mental health issues for faculty cannot be ignored as significant numbers of faculty within the region have experienced stress and anxiety brought about by confined lives, increased workload, and job uncertainties (Agyapong et al., 2020, p. 8).

O’Malley (2020) observes that even before the crisis created by the COVID-19 outbreak, data was showing that access and equity remained big challenges in higher education, especially in relation to wealth distribution globally, with Sub-Saharan Africa falling further behind other regions. In 2018, access to higher education in Sub-Saharan Africa averaged about 9% compared to 70% in Europe and Central Asia. The COVID-19 pandemic has increased inequality in African higher education, with digital divide posing even a greater threat to equitable access to education and lifelong learning in Africa (Makoni, 2020). It is generally believed that enrollment, which is already low in African member states, will be further affected with the recognition that the longer children and young people stay away from school, with prolonged school closures, the less likely they are ever to return to school after school reopening (AUC, 2020).

To better understand the effects of epidemics on economic prosperity, knowing the channels by which health affects economic growth in general and the extent to which these channels differ between rich and poor countries is important. A particular challenge for policymakers is therefore to reduce the inequality-increasing tendencies of epidemics when designing policies to cope with epidemic-induced economic downturns (Bloom et al., 2020). In low-income countries, a lack of general access to quality health care, widespread undernourishment, and the prevalence of competing infectious diseases, exacerbate the impact of epidemics. In high-income countries, by contrast, access to effective treatments and a well-functioning health care sector can often mitigate negative health shocks.

Strategies to mitigate COVID-19 effects

As the pandemic continues to ease, it is necessary to assess the magnitude of its socio-economic impact on African countries and identify bold, urgent, medium-term and long-term measures to continue saving human lives and restoring African economies, in order to build a more endogenous and resilient economy to ensure Africa’s self-reliance (AUC, 2020). The educational system has been solid throughout history, continuously functioning in defiance of wars, epidemics, economic hardships and geographical burdens (Al Lily et al., 2020), and this system must now consider roundabout ways to ensure its continuity during, and in the aftermath of, the pandemic. The system does not normally tolerate its discontinuity and abolition, and the tolerance takes place within the sphere of searching for substitutes. Higher education studies (e.g., Butler-Henderson et al., 2020) have emphasized how students and staff are responding to rapid adaptation, how institutions are managing their new commitments and new service landscape, and how institutional responses differ.

Following the move to online or distance learning, availability of resources such as quality IT platforms, fast internet connection, and other computer facilities and equipment are essential (Abdullah et al., 2020). Nevertheless, the forced move to distance teaching and learning offers important opportunities to propose more flexible learning possibilities, explore blended or hybrid

learning and to mix synchronous learning with asynchronous learning (Marinoni et al., 2020). Therefore, universities need to be encouraged and facilitated to offer e-learning pedagogy as an integral part of career-long training for their academics (MacGregor, 2020). Furthermore, COVID-19 has thrown into stark relief the imperatives to bridge the digital divide and to share resources across Africa, and higher education has a key role to play in advancing both of those goals, including by teaching digital skills and collaborating in postgraduate training and research. There is now opportunity to radically revise the curriculum of higher education institutions from every corner to suit remote, distance, online, and digital forms of delivery (Butler-Henderson et al., 2020).

In the outbreak of the pandemic, staff had not been trained to deliver remote learning (O'Malley, 2020) so as to enable them prepare lectures in the context of online classes (Makoni, 2020). To mitigate this, the Commonwealth of Learning (COL, 2020) has developed the "Guidelines on Distance Education during COVID-19" to help governments, educational institutions, regulatory agencies, parents, and students to effectively deal with this crisis so that no one is left behind. Furthermore, COL has initiated several joint activities aimed at building the capacity of staff in distance and online provision and harnessing the potential of Open Educational Resources (OERs) to offer quality content. Other plans could include developing a comprehensive remote education policy, instituting strategies to incorporate critical aspects of the COVID-19 experience in Education Management Information System (EMIS) to inform education sector reform, and leveraging the lessons learned from the successful and failed partnerships and collaborations during the COVID-19 crisis to improve future stakeholder engagement in times of crisis (ADEA, 2020).

There has been an exponential increase in global research efforts to understand and control the virus (The Economist, 2020b), and some journals in the aforementioned disciplines have started to fast-track the peer-review process to accelerate the publication process (The Economist, 2020a). While under normal circumstances, publishing in scientific journals usually takes years, speed has become paramount and journals "have squeezed their normal processes down to days or weeks" (The Economist, 2020b, n.p.), thus providing physicians, policymakers, HEIs, and heads of state with the latest science in order to make far-reaching decisions (The Economist, 2020b). Other journals have lowered their paywalls to make research about SARS-CoV-2 more widely available (The Economist, 2020a).

The relative dearth of systematic and macro-level research has led to the exploration of creating a database that shares the research on higher education openly, such as the "COVID-19 in Higher Education Literature Database (CHELD) V1.0." This database makes the access to timely knowledge easier, with an aim to promote knowledge-sharing behaviors in higher education providers during the pandemic (Butler-Henderson et al., 2020). Now, there is need for university leaders, telecommunications companies, global employers and students to come together to develop a common agenda for the future of digital higher education (Waruru, 2020a).

The financial resources can be a critical component in handling and recovering from acute crises (De Carvalho et al., 2016). Additional financial support to the education sector for managing the disruption caused by the COVID-19 pandemic has largely been from the national stimulus or special funds as well as from external sources, particularly from major international and regional development organizations such as the World Bank Group, the African Development Bank Group, and the Global Partnership for Education, among others. A major gap exists, however, in financing and sustaining a solid mechanism for the education sector in Africa to manage crises like COVID-19 and to guide remote education on a mid-to long-term perspective. Availability of emergency relief funds in the form of soft loans or private donations will help students with daily expenses in order to continue their studies.

In an unprecedented situation as a result of COVID-19 pandemic, social support system may provide a valuable relief to those severely affected by the crisis (Abdullah et al., 2020). Some of those affected groups may require psychological help or counseling in order to relieve them from mental or emotional crisis that may lead to severe stress and depression. It is particularly important for various parties involved in the crisis to react resiliently and develop suitable solutions to overcome the undesired situation. Stress in teaching online using technology (or technostress) is stress caused by the use of ICT (Christian et al., 2020). Accordingly, Sawalha (2015) has called for an organizational culture which is open, trustful, and learning-oriented.

Conclusion and lessons learned

According to Giordani and Rullani (2020) the world lasts three periods: period one is to be interpreted as the pre-COVID phase; period two in which the COVID-19 shock occurs, and the non-digital market temporarily disappears; and period three is intended as the post-COVID era. COVID-19 has revealed that healthcare and public health systems were unprepared to meet a pandemic challenge due to underfunding, poor supply chain access, and disregard for healthcare personnel, school nurses included. During the pandemic, the things we all took for granted no longer became an option. Education is encountering the greatest disruption, marking a radical departure from historical modes of delivery. Whereas distance education is often touted for reasons such as geographical isolation, disability and wars, it is now being used as a tool for dealing with a medical tragedy (Al Lily et al., 2020).

Although there is a positive move in using new online and offline learning tools, platforms, and materials, it benefits only a small proportion of the learner population (ADEA, 2017; Oyedotun, 2020). The sudden transition to online pedagogy as a result of COVID-19 in developing countries has exposed some inequalities and challenges, as well as benefits. These challenges and inequalities have now become the new realities in the educational sector of developing countries. Adapting to the changes in the education

sector could benefit HEIs and staff in the long run; for instance once COVID disappears, the online provision will stay and be a solid element of teaching and learning (Siwinski and Bilanow, 2020). Online technology will not only change the teaching process and make it more competitive (for lecturers), but as a disruptive force will create opportunity to accelerate profound changes in the organization and introduction of innovation. There is, therefore, need to rethink how to deliver quality education equitably and inclusively, and the role of technology as one of the key enablers.

References

- Aarts, H., 2020. COVID-19 Hits HE Capacity Building Ties with Global South. University World News. Retrieved 26/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200603075317313>.
- Abdullah, M., Husin, N.A., Haider, A., 2020. Development of post-pandemic COVID-19 higher education resilience framework in Malaysia. Arch. Business Res. 8 (5), 201–210.
- ACU, 2020. Higher Education During COVID-19: A Snapshot of Digital Engagement in Commonwealth Universities. ACU Policy Brief. Retrieved 30/04 2021 from. www.acu.ac.uk/news/higher-education-during-covid-19-a-snapshot-of-digital-engagement-in-commonwealth-universities.
- ADEA, 2020. Impact of COVID-19 on Africa's Education: Reflecting on Promising Interventions and Challenges, Towards a New Normal. Association for the Development of Education in Africa, ADEA.
- Agyapong, S., Asare, S., Essah, P., et al., 2020. Learning in Crisis: COVID-19 Pandemic Response and Lessons for Students, Faculty, and Vice Chancellors in Sub-Saharan Africa. Education Sub-Saharan Africa (ESSA), Accra.
- Aina, T.A., 2010. Beyond reforms: the politics of higher education transformation in Africa. Afr. Stud. Rev. 53 (1), 21–40.
- Al Lily, A.E., Ismail, A.F., Abunasser, F.M., et al., 2020. Distance education as a response to pandemics: coronavirus and Arab culture. Technol. Soc. 10 (1016).
- Areba, G.N., 2020. COVID-19 pandemic impact on Kenyan education sector: learner challenges and mitigations. J. Res. Innov. Implicat. Educ. 2, 128–139.
- Arnhold, N., Brajkovic, L., Nikolaev, D., et al., 2020. Tertiary Education and COVID-19: Impact and Mitigation Strategies in Europe and Central Asia. Retrieved 24/11/2020 from. <http://pubdocs.worldbank.org/en/783451590702592897/ECA-TE-and-COVID-19-long-FINAL-25May20.pdf>.
- AUC, 2020. Socioeconomic Impact of COVID-19 on Africa: An Opportunity to Build More Endogenous and Resilient Emerging Economies. African Union Commission, Addis Ababa, Ethiopia.
- Bloom, D., Canning, D., Chan, K., 2005. Higher Education and Economic Development in Africa. International Bank for Reconstruction and Development/World Bank, Washington, DC.
- Bloom, D.E., Kuhn, M., Prettnner, K., 2020. Modern Infectious Diseases: Macroeconomic Impacts and Policy Responses. NBER Working Paper No. 27757.
- Brazendale, K., Beets, M.W., Weaver, R.G., et al., 2017. Understanding differences between summer vs. school obesogenic behaviors of children: the structured days hypothesis. Int. J. Behav. Nutr. Phys. Activ. 14, 100.
- Brodeur, A., Gray, D.M., Islam, A., et al., 2020. A Literature Review of the Economics of Covid-19. GLO Discussion Paper, No. 601. Global Labor Organization (GLO), Essen.
- Brooks, S.K., Webster, R.K., Smith, L.E., et al., 2020. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet. [https://doi.org/10.1016/S2215-0366\(20\)30077-30078](https://doi.org/10.1016/S2215-0366(20)30077-30078).
- Butler-Henderson, K., Crawford, J., Rudolph, J., et al., 2020. COVID-19 in higher education literature database (CHELD V1): an open access systematic literature review database with coding rules. J. Appl. Learn. Teach. 3 (2).
- Centers for Disease Control and Prevention, 2020. Lesson 1: introduction to epidemiology. Section 11: epidemic disease occurrence. Level Dis. Updated May 18, 2012.
- Christian, M., Purwanto, E., Wibowo, S., 2020. Technostress creators on teaching performance of private universities in Jakarta during Covid-19 pandemic. Technol. Rep. Kansai Univ. 62 (6), 2799–2809.
- COL, 2020. Guidelines on Distance Education During COVID-19. COL, Burnaby.
- Crawford, J., 2020. COVID-19 and Higher Education: A Pandemic Response Model from Rapid Adaption to Consolidation and Restoration. In press.
- De Carvalho, A.O., Ribeiro, I., Cirani, C.B.S., et al., 2016. Organizational resilience: a comparative study between innovative and non-innovative companies based on the financial performance analysis. Int. J. Innov. 4.
- Dell, S., 2020. Donor Funding—COVID-19 Will Change Everything. University World News. Retrieved on 26/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200716113159558>.
- Di Pietro, G., Biagi, F., Costa, P., et al., 2020. The Likely Impact of COVID-19 on Education: Reflections Based on the Existing Literature and Recent International Datasets. Publications Office of the European Union.
- Engzell, P., Arun, F., Verhagen, M., 2020. Learning Inequality During the COVID-19 Pandemic. University of Oxford, Mimeo.
- Ferdig, R.E., Baumgartner, E., Hartshorne, R., et al., 2020. Teaching, Technology, and Teacher Education During the COVID-19 Pandemic: Stories from the Field. Association for the Advancement of Computing in Education (AACE). Retrieved on 15/9/2020 from. <https://www.learnlib.org/p/216903/>.
- George, G., Lakhani, K.R., Puranam, P., 2020. What has changed? The impact of COVID pandemic on the technology and innovation management research agenda. J. Manag. Stud. 57 (8). <https://doi.org/10.1111/joms.12634>.
- Gersovitz, M., 2011. The economics of infection control. Annu. Rev. Resour. Econ. 3 (1), 277–296. <https://doi.org/10.1146/annurev-resource-083110-120052>. Retrieved on 03/05/2021 from. <https://ssrn.com/abstract=1928380>.
- Giordani, P.E., Rullani, F., 2020. The Digital Revolution and COVID-19. Working Paper Series n. 6/2020. Department of Management, Università Ca' Foscari Venezia.
- Halterbeck, M., Conlon, G., Williams, R., et al., 2020. Impact of Covid-19 on University Finances: Report for the University and College Union. London Economics, London.
- Hantrais, L., et al., 2020. Covid-19 and the digital revolution. Contemp. Soc. Sci. <https://doi.org/10.1080/21582041.2020.1833234>.
- Hooper, O., Djerassimovic, S., 2020. The coronavirus crisis: what impact for early career researchers? Res. Intell. 144.
- International Association of Universities, 2020. The Impact of Covid-19 on Higher Education Worldwide. Resources for Higher Education Institutions. Retrieved from. https://www.iauau.net/IMG/pdf/covid-19_and_he_resources.pdf. (Accessed 11 May 2020).
- Jaramillo, D.B., Stephenson, G.K., 2020. Interconnected Impacts of COVID-19 on Graduate Students. University World News. Retrieved on 27/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200615085637162>.
- Jegede, D., 2020. Perception of undergraduate students on the impact of Covid-19 pandemic on higher institutions development in federal capital territory Abuja, Nigeria. Electron. Res. J. Soc. Sci. Humanit. 2 (2), 211–222.
- Klein, A.M., Vaissiere, B.E., Cane, J.H., et al., 2007. Importance of pollinators in changing landscapes for world crops. Proc. Biol. Sci. 274 (1608), 303–313.
- Lackie, K., et al., 2020. Interprofessional education and collaborative practice research during the COVID-19 pandemic: considerations to advance the field. J. Interprof. Care 34 (5), 583–586. <https://doi.org/10.1080/13561820.2020.1807481>.
- Ligami, C., 2020. Country Responses to Equity Still a Challenge—UNESCO Global Report. University World News. Retrieved on 27/11/2020 from. <https://www.universityworldnews.com/post.php?story=202011210935281>.

- Lues, R., Padayachee, A., de Jager, H., 2020. Universities of Technology in the Post-COVID-19 Landscape. University World News. Retrieved on 27/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200713153430109>.
- MacGregor, K., 2020. Urgent Need for HE to Bridge Digital Divide and Share Resources. University World News. Retrieved from. <https://www.universityworldnews.com/post.php?story=20200716105729578>.
- Madhav, N., Oppenheim, B., Gallivan, M., et al., 2017. Pandemics: risks, impacts, and mitigation. In: Disease Control Priorities: Improving Health Reducing Poverty. <https://doi.org/10.1596/978-1-4648-0527-1/pt.5.ch17>.
- Makoni, M., 2020. COVID-19 is Stoking Inequality but Also Innovation in HE. University World News. Retrieved on 28/11/2020 from. <https://www.universityworldnews.com/post.php?story=20201121081243530>.
- Malani, A., Laxminarayan, R., 2011. Incentives for reporting infectious disease outbreaks. J. Hum. Resour. 46 (1), 176–202.
- Marinoni, G., de Wit, H., 2020. A Severe Risk of Growing Inequality Between Universities. University World News. Retrieved on 28/11/2020 from. <https://www.universityworldnews.com/post.php?story=2020060815405140>.
- Marinoni, G., van't Land, H., Jensen, T., 2020. The Impact of COVID-19 on Higher Education Around the World: IAU Global Survey Report. International Association of Universities, Paris.
- Maslen, G., 2020. 50% Drop in Foreign Student Enrolment by 2021 Predicted. University World News. Retrieved from. <https://www.universityworldnews.com/post.php?story=20200601150411139>.
- Materu, P., 2007. Higher Education Quality Assurance in Sub-Saharan Africa: Status, Challenges, Opportunities and Promising Practices. World Bank, Washington DC.
- McKie, A., 2020. Covid-19: Universities Treating Staff in “Vastly Different Ways”. Retrieved on 24/11/2020 from. www.timeshighereducation.com/news/covid-19-universities-treatingstaff-vastly-different-ways.
- Mhlanga, D., Moloi, T., 2020. COVID-19 and the digital transformation of education: what are we learning on 4IR in South Africa? Educ. Sci. 10 (7), 180. Retrieved 01/05/2021 from file:///G:/Downloads/education-10-00180.pdf.
- Mitchell, N., 2020. Universities Not Ready for Online Learning—U-Multirank. University World News. Retrieved on 28/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200609183303614>.
- Mohamedbhai, G., 2020. COVID-19: What Consequences for Higher Education? University World News. Retrieved on 16/04/2021 from. <https://www.universityworldnews.com/post.php?story=20200407064850279>.
- Muftah, M., 2020. Higher education and Covid-19 pandemic: matters arising and the challenges of sustaining academic programs in developing African universities. Int. J. Educ. Res. Rev. 5 (4), 417–423.
- Nkhoma, N.M., 2020. The Government Should Open Universities with Immediate Effect. University World News. Retrieved on 24/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200714144233896>.
- Odhiambo, W., 2020a. COVID-19—Managing the Risk to Older University Staff Members. University World News. Retrieved on 25/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200716062706283>.
- Odhiambo, W., 2020b. Sport Scholarship Students Suffer Under COVID-19 Lockdown. University World News. Retrieved on 25/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200609072143783>.
- O'Malley, B., 2020. Digital Divide “Catastrophic” for Many Students—World Bank. University World News. Retrieved on 27/11/2020 from. <https://www.universityworldnews.com/post.php?story=20201121084401266>.
- Oyedotun, T.D., 2020. Sudden change of pedagogy in education driven by COVID-19: perspectives and evaluation from a developing country. Res. Glob. 2. ISSN 2590–051X.
- Pathmanathan, I., O'Connor, K.A., Adams, M.L., et al., 2014. Rapid assessment of Ebola infection prevention and control needs—six districts, Sierra Leone. MMWR Morb. Mortal. Wkly. Rep. 63 (49), 1172–1174.
- Philpott, T., 2000. Economic epidemiology and infectious diseases. In: Culyer, A.J., Newhouse, J.P. (Eds.), Handbook of Health Economics, vol. 1. Elsevier, pp. 1761–1799.
- PWC, 2020. Social and Economic Impact Assessment of COVID-19 in Republic of Moldova: Deliverable 1—Initial Impact Assessment Report. Retrieved 26 Nov 2020 from. https://www.md.undp.org/content/moldova/en/home/library/inclusive_growth/social-and-economic-impact-assessment-of-covid-19-in-the-republ.html.
- Rasul, I., 2020. The economics of viral outbreaks. AEA Papers Proc. 110, 265–268. <https://doi.org/10.1257/pandp.20201016>.
- Sawalha, I.H.S., 2015. Managing adversity: understanding some dimensions of organizational resilience. Manag. Res. Rev. 38, 346–366.
- Schwab, M., Malleret, T., 2020. Covid-19: The Great Reset. World Economic Forum, Geneva.
- Siwinski, W., Bilanow, K., 2020. Will University Rankings Adjust to Our Uncertain Times? University World News. Retrieved on 27/11/2020 from. <https://www.universityworldnews.com/post.php?story=20201105133824428>.
- SPAU, 2020. Articulating the Pathways of the Socio-economic Impact of the Coronavirus (COVID-19) Pandemic on the Kenyan Economy. Policy Brief, Issue No: 4/2020.
- Tefera, D., Altbach, P.G., 2004. African higher education: challenges for the 21st century. High Educ. 47 (1), 21–50. <https://doi.org/10.1023/B:HIGH.0000009822.49980.30>.
- Teti, M., Schatz, E., Liebenberg, L., 2020. Methods in the time of COVID-19: the vital role of qualitative inquiries. Int. J. Qual. Methods 19. <https://doi.org/10.1177/1609406920920962>. Retrieved 04/05/2021 from. <https://journals.sagepub.com/doi/full/10.1177/1609406920920962>.
- The Economist, 2020a. Coronavirus Research is Being Published at a Furious Pace. <https://www.economist.com/graphic-detail/2020/03/20/coronavirus-research-is-being-published-at-a-furious-pace>.
- The Economist, 2020b. Scientific Research on the Coronavirus is Being Released in a Torrent. <https://www.economist.com/science-and-technology/2020/05/07/scientific-research-on-the-coronavirus-is-being-released-in-a-torrent>.
- UNESCO-IIESALC, 2020. COVID-19 and Higher Education: Today and Tomorrow. Retrieved 27/11/2020 from. <http://www.iesalc.unesco.org/en/wp-content/uploads/2020/04/COVID-19-EN-090420-2.pdf>.
- Vitae, 2020. Release of Initial Findings to Sector Following Response to Covid-19 Survey. Retrieved on 24/11/2020 from. www.vitae.ac.uk/news/vitae-news-2020/release-of-initialfindings-to-sector-following-response-to-covid-19-survey.
- Wachira, K., Ombati, R., 2020. E-Learning in Public Universities—The Only Way is Forwards. University World News. Retrieved on 24/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200511100801516>.
- Wang, G., Zhang, J., Lam, S.P., et al., 2019. Ten-year secular trends in sleep/wake patterns in Shanghai and Hong Kong school-aged children: a tale of two cities. J. Clin. Sleep Med. 15, 1495–1502.
- Waruru, M., 2020a. Digital Education Should Be a Sustainable Development Goal. University World News. Retrieved on 30/11/2020 from. <https://www.universityworldnews.com/post.php?story=20201111023337237>.
- Waruru, M., 2020b. Internationalisation and COVID-19—Challenges and Lessons. University World News. Retrieved on 17/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200715110850412>.
- World Bank, 2020. COVID-19's Immense Impact on Equity in Tertiary Education. <https://blogs.worldbank.org/education/covid-19s-immense-impact-equity-tertiary-education>.
- World Health Organization, 2020a. WHO Director-General's Opening Remarks at the Mission Briefing on COVID-19, 2020. Retrieved on 24/11/2020 from. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-mission-briefing-on-covid-19>.

- World Health Organization, 2020b. Coronavirus Disease 2019 (COVID-19) Situation Report-73. Retrieved on 01/05/2021 from. <https://www.who.int/docs/default-source/coronavirus/situation-reports/20200402-sitrep-73-covid-19.pdf>.
- Wu, S.J., Chang, D.F., Sun, F.R., 2020. Exploring college student's perspectives on global mobility during the COVID-19 pandemic recovery. *Educ. Sci.* 10, 218.
- Wyplosz, C., 2020. The Euro Area After COVID-19. Publication for the Committee on Economic and Monetary Affairs, Policy Department for Economic, Scientific and Quality of Life Policies. European Parliament, Luxembourg.
- Zezeza, P.T., 2020. COVID-19 and African Universities—AAP Dialogue Takeaways. *University World News*. Retrieved on 24/11/2020 from. <https://www.universityworldnews.com/post.php?story=20200714080656408>.

Issues of production and utilization of knowledge: multiple uses of knowledge

Azeb Tadesse, Bri Damacio, Kiana Billot-Vasquez, and Walter R. Allen, School of Education and Information Studies, University of California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	102
Academic perspectives	102
Global South perspectives	103
Indigenous knowledge(s)	104
Democratization of knowledge	105
Practitioner's knowledge	105
Conclusion	106
References	106

Introduction

Knowledge is tied to power (Foucault, 1977/1995). The power to produce, disseminate, and legitimize it looms large. Although knowledge can be produced anywhere by anyone, the nexus of power relations in society, acting as a gatekeeper in the interest of maintaining control, determines what is accepted and the truth (Foucault, 1980; Gramsci, 1971). Hence, the production and use of knowledge is ultimately linked to the nexus of what Foucault called power/knowledge. Foucault emphasizes that power is reinforced and maintained through acceptance of certain forms of knowledge and truths (Rainbow, 1991). This interlinks between power and knowledge has dominated our understanding of what constitutes legitimate knowledge production and utilization.

Historically, Eurocentric perspectives and narratives have dominated knowledge production. “The ‘Western university’ and intellectual property regimes were not initially built for all thinkers and knowers, nor for inclusive human advancement” (Wilder, 2013). These traditional, presiding forms of knowledge have worked to exclude other forms of knowledge with worldviews that differ from those in dominant cultures. These exclusions often censored knowledge originating from the Global South and indigenous cultures. By shutting out voices from the Global South, perhaps unwittingly, knowledge production enters the contested space of power and epistemology, and the enduring legacy of colonialism. Traditional academic perspectives reflect European and American knowledge hierarchies. Bhambra (2014) examines how decolonial thinking has the “radical potential” to force us to rethink traditional forms of knowledge production. In thinking about that “radical potential,” Bhambra (2014) writes, “Postcolonialism and decoloniality are only made necessary as a consequence of the depredations of colonialism, but in their intellectual resistance to associated forms of epistemological dominance they offer more than simple opposition. They offer, in the words of María Lugones, the possibility of a new geopolitics of knowledge” (p. 120).

Over the last decades, the way knowledge is generated and circulated has been transformed. It is our responsibility to understand how those changes have enabled other forms of knowledge to become more accessible and thus, easier to render visible. However, in some instances these changes made non-mainstream knowledge invisible, denying access to alternative world views. We offer a critical perspective of how knowledge is created and deployed with the ultimate goal to give voice to the voiceless and form to the unseen.

This review examines 5 topics associated with the production, use and privileges of traditionally accepted forms of knowledge. First, we review how knowledge circulates in higher education and how whiteness has dominated academic perspectives and knowledge. After understanding how knowledge has traditionally been circulated in academia, we turn to knowledge from the Global South in order to better understand the relevance of locality. Next, we consider how Indigenous forms of knowledge have been devalued over time and the relationship between imperialism, colonialism and knowledge. From there, we explore the democratization of knowledge and delve into how the emergence of the digital age has taken the democratization of knowledge further than at any time before. Lastly, we discuss practitioner knowledge, which has been consistently disregarded in favor of so-called “expert,” or traditionally legitimized forms of knowledge.

Academic perspectives

Delgado Bernal and Villalpando (2002) argue there is an “apartheid of knowledge” in the context of higher education in the United States. This phenomenon is maintained through what Scheurich and Young (1997) describe as epistemological racism that legitimizes Western ways of knowing in the academia and limits the possibility of non-Western (non-white) epistemologies from being considered a part of the mainstream research community. This means research produced by scholars of color that utilize theoretical

frameworks and methodologies different from established positivist traditions, is considered deficient, biased, and non-rigorous (Delgado Bernal and Villalpando, 2002; Solórzano and Yosso, 2002).

Patton (2016) insists we look at the composition of faculty members in the US in order to better understand how/why certain perspectives are dominant. She argues the lack of diversity in higher education is commonplace. According to the Pew Research Center, faculty members, despite the racial and ethnic changes in the past two decades, continue to be majority white; the Center notes that white faculty members make up about 76% of all college and university faculty members compared to 55% of total undergraduate students. Conversely, they found Latino, Black/African Americans, and Asians make up only 5%, 6%, and 11% of faculty members, respectively. Further, it is important to note that faculty of color are more likely than white faculty to be in non-tenured, lower, and lesser prestigious academic ranks such as lecturers and/or instructors; be concentrated in institutions of lesser prestige and with fewer resources; and be in departments that are considered less prominent and prestigious within higher education, namely the humanities, social sciences, ethnic studies, women's studies, and education (Astin et al., 1997; Garza, 1993; Allen et al., 2000). Garza (1993) describes these patterns as the "ghettoization" and "barrioization" of faculty of color and their scholarship.

Scholarship produced by scholars of color is often dismissed, marginalized, footnoted, or unacknowledged (Delgado Bernal and Villalpando, 2002). Stanley (2007) notes many publishing and journal editorial boards continue to favor mainstream research. While presumed to be fair and impartial, this process in fact erects barriers that reify, privilege, and confer status only to knowledge which has been legitimized or formalized by the academy. The ultimate result is to reproduce an incestuous cycle within higher education where a "small group of White faculty repeatedly cite their own work contributing to their status" (Patton, 2016). All of which results in "opportunity hoarding" and the concentration of power for white faculty. Faculty of color have limited status, tenure rates, and "voice." As Trinh (1989) reminds us, too often the only way to "science" is through whiteness; colonized and subjugated knowledge struggles to be seen as legitimate outside Western interpretation, analysis, and mastery. Ultimately, knowledge legitimized through the academy falsely normalizes the way(s) of knowing and knowledge created and imposed by colonizers. Dominant white, Western, and Global North world views, about what is legitimate knowledge, permeate every institutional structure, policy, and practice to sustain the status quo in higher education (Sefa Dei, 1999).

Global South perspectives

Current knowledge is framed and communicated through a Western, Global North lens. UNESCO's Global Education Coalition (GEC) is a timely initiative for an era of ever-increasing conflicts, xenophobia, and other forms of discrimination. However, it has encountered difficulty in gaining traction in some parts of the globe. This difficulty in local relevance is, in no small measure, due to its narrow definition of values and principles. Put simply, GEC lacks adaptability to local realities because it marginalizes important principles, values and world views of Global South societies. Perhaps the most conspicuous omission is the absence of voices from the Global South and the failure to connect important key concepts such as *citizenship*, *human rights*, *democracy*, and *peace*, (highly political terms) to localized concepts and values. Moreover, the framework for GEC is based largely on documents that, while globally ratified, are still locally controversial, such as the "Universal Declaration of Human Rights" and "the World Programme for Human Rights Education" (UNESCO, n.d.).

By overlooking important voices on its central issues from the Global South, GEC unwittingly enters the contested spaces of power, epistemology, and the enduring legacy of colonialism. Post-colonial scholar Boaventura de Sousa Santos defines colonialism as "a system of naturalizing differences in such a way that the hierarchies that justify domination and oppression are considered the product of the inferiority of certain peoples and not the cause of their so-called inferiority" (de Souza Santos, 2017, p. 18). In addition to realigning global paradigms of power and legitimacy, colonialism was also complicit in what Sousa Santos calls epistemicide "the destruction of the knowledge and cultures of [oppressed] populations, of their memories and ancestral links and their manner of relating to others and to nature" (de Souza Santos, 2017). Given this history and the continuing imposition of external agendas in the post-colonial era, societies in the Global South are justifiably suspicious and apprehensive of "international" initiatives that lack connection, relevance, and alignment with local realities and national priorities.

The goals and objectives of GEC, along with its resource documents, are largely communicated and conceptualized through a Global North lens—specifically a Western worldview. Although there are universally accepted and acknowledged ideals, one encounters difficulties in interpretation and placement of these universal ideals in local context. GEC appears to have overlooked the fact that the echoes of history and geopolitical alignments are a constant restraint in the effort to reach a consensus on workable, universal definitions of ideals, values, and goals. An illustration of this difficult task is evident in the titles of GEC resources offered by UNESCO with titles containing words such as "prevention of violent extremism," "reinforcing human rights," and "countering radicalization" (UNESCO, n.d.). The underlying intention of these resource packs is the near universal agreement on the need to protect civilians and to recognize dangerous viewpoints that lead to violence. However, the global and local divide on the understanding of these concepts is conditional on terminology and political outlook. Utilizing politically sensitive terms like "extremism" can be more detrimental than helpful to the GEC's purposes in developing these materials (Appiah-Mensah, 2007).

Creating resonance at the local level, requires first, deconstructing the primary concepts of GEC, followed by consciously contextualizing and incorporating epistemologies of the Global South. Key to this deconstructing and localizing is recognizing that by stating that these values need to be 'instilled' in a member of any given society, and in the process offering an externally formed definition for these values, GEC is implying that these humanistic values are locally absent.

Moreover, GEC is laden with politically charged terms such as citizenship and human rights that are strictly circumscribed, inhabiting adaptation to local equivalence concepts which may facilitate incorporation into the curriculum. For instance, although GEC may define citizenship in broad terms such as individual expression, creativity, and awareness. For a large part of the world, citizenship is a concept that encompasses contradictions, such as “belonging vs. exclusion,” “granting of rights vs. denying rights,” and “access vs. expulsion.” Additionally, it is important to develop a historical understanding of how these values and ideals were developed. For example, the current understanding of Human Rights results from reinterpretation of previous European colonial standards of who is “civilized” to one, which sees the US as Western civilization ideal. Hence, Human Rights standards are “universal” only inasmuch as these are a barometer for civilization and who is defined as civilized-as confirmed by the prevailing power championing the diffusion of these ideals and standards (Mignolo, 2012, p. 296).

For GEC to resonate locally, it must first accommodate Global South perspectives and secondly, reject imposition of its overly prescriptive language, concepts, values, and ideals and thirdly, allow for local interpretation and incorporation. For instance, the stark contrast in the discourse on colonialism and post-colonial era illustrates the fissure in perspective and interpretation between the Global North and South, as well as the compartmentalization of non-Western academics and authors. Works of Global South academics and authors such as Said (1979a), Rodney (1974), Fanon (1961), and Achebe (1960), that offer the intimate, personal, and social impact of oppression and offer contextualized interpretations are often framed as niche and regulated to specialized courses. Meanwhile, the canons are such works as Conrad (1996), Herskovits (1938, 1941, 2006), Forster (2005), and Coleman (1955, 1971, 1995), which are drafted through a Eurocentric lens and at times perpetuate Western superiority and ideology. This disparity in legitimacy and the detraction from the locally centered narrative for an established flawed Global narrative aptly illustrates the shortcomings of not only the GEC but also the wider academic and political discourse. As witnessed numerous times in history, the Global is ill equipped to define the priorities and conditions of the local because as Edward Said (1979) warned in *Orientalism*, “cultures, and histories cannot seriously be understood or studied without their force, or more precisely their configurations of power” (Said, 1979b, p. 5).

Indigenous knowledge(s)

Indigenous knowledge has been defined as “the foundation of Indigenous governance, ecological stewardship, social, ethical, linguistic, spiritual, medical, food, and economic systems, so that the continual production and reproduction of local, land-based knowledge is the basis of Indigenous identity and sense of place in the world, as well as of Indigenous groups’ very survival as distinct peoples” (Grey, 2014). Arguably more important than Indigenous Knowledge (IK) functioning as a survival mechanism, is its power to help free Indigenous people from the constraints of colonial rule and thought. In discussing the impact of Indigenous Knowledge for indigenous peoples, Simpson (2004) writes, “Recovering and maintaining Indigenous worldviews, philosophies, and ways of knowing and applying those teachings in a contemporary context represents a web of liberation strategies Indigenous Peoples can employ to disentangle themselves from the oppressive control of colonizing state governments”. Furthermore, Bruchac (2014) notes that those from outside Indigenous communities trying to understand indigenous knowledge should always consult with “local indigenous knowledge-bearers” to ensure accuracy, ethical practice, and avoid unnecessary harm.

The methods we use also have power in the construction of knowledge. Patel’s (2016) *Decolonizing Educational Research* explores the ways in which colonialism seeps into modern processes of knowledge formation and educational research. In “Educational Research as a Site of Coloniality,” Patel states:

Education research, through both meaning and matter, has played a deleterious role in perpetuating and refreshing colonial relationships among people, practices, and land. This is not to say this is the only purpose that educational research has performed, nor that all educational researchers are colonizers themselves. Rather, it means that logics of coloniality, which are connected to property and stratifications of society, are problematically enlivened through educational research (p. 12).

By turning to Indigenous knowledge, we can be equipped to understand something free of the colonial perspective—an imbalanced and miss referenced partial view. The lack of exposure to Indigenous knowledge in higher education and journals works to reproduce colonial logics and confer status to Western epistemologies. This absence renders Indigenous knowledge and cultures invisible. In addition to research’s potential to refresh colonial logics and render certain realities invisible, Patel (2016) also explores the difference in temporarily improving experiences versus systematically addressing and understanding them. She argues that, “When educational research focuses on these strata without addressing the societal design that creates the strata, it becomes complicit in the larger project. It may temporarily improve a small set of experiences, often conceptualized with gaining footholds on a slanted wall, rather than reconsidering the entire structure” (p. 18). Indigenous knowledge has been threatened by assimilation, industrialization, and outmigration. Furthermore, Grey (2014) states that “massive depopulation due to epidemics and malnutrition, immediately upon and for hundreds of years following the arrival of colonial powers in traditional territories, resulted in an incalculable loss of Indigenous knowledge” (p. 3231).

Democratization of knowledge

The emergence of the digital age has taken the democratization of knowledge further than at any time before. The invention of the printing press advanced the dissemination of knowledge not only to the broader populace but also faster than in previous eras. Similarly, digital technology has not only propelled knowledge further but has also democratized the creation of knowledge. Today, the proliferation of digital platforms has transformed how we create, disseminate, and legitimize knowledge. Not only has it increased the speed and spread of knowledge and information, it has also raised questions of who is an “expert” and qualified to produce knowledge challenging traditional epistemology.

The democratization of knowledge occurs through inclusivity and diversity in the production and dissemination of knowledge. The dominance of European perspectives and prevalence of “gatekeeping” have been, and continue to be, significant barriers to broadening representation in knowledge production and dissemination. However, the Internet and technological innovation have forever altered the landscape and created a platform to challenge traditional epistemology.

The Internet has launched a significant global revolution in knowledge production and access. Today digital technology is present in all aspects of knowledge production, from instruction to collaboration to publishing to research (Altbach et al., 2009). For instance, EdTech, which is at the intersection of technology and education, integrates innovations and advances in technology to support, supplement, customize, personalize, and diversify learning (EdTech 2.0, 2017). Innovations in EdTech gave rise to online learning, which employs computers and the Internet as a delivery mechanism for courses and instruction (Moore, 2013). With each stage of innovation, the goal has been to create a learning environment that diminishes the distance between instructor and student as well as the student and the material. Digital innovations have expanded EdTech from simple computers to interactive displays and simulations, increasing interaction between instructor and learner (EdTech Review, 2013). Although access to both the Internet and knowledge dissemination channels has improved, equity remains elusive with non-western countries facing significant connectivity challenges, which has implications for fully leveraging technology in education. However, with concerted efforts to increase affordability and access, EdTech can potentially expand learning, collaboration, and research, thus amplifying underrepresented voices and eventually diversifying epistemology.

College and university research libraries have been the most significant beneficiaries of digital technology innovation. Libraries no longer just curate books and materials for those who physically enter the space. They support e-learning students and provide mobility for digitally accessing materials and services. The digital age college library is both digital and physical, where students can check out a book or access it online from their personal computers. Beyond curating books and materials, it has become a hub for training and education on innovative technology, from coding to data management tools that support learning and research. Libraries are places where students go to access multimedia materials, high-tech solutions such as 3-D printers and laser cutters, and data visualization walls.

An example of a new era of libraries made possible by leveraging technology is the HathiTrust Digital Library, a digital repository collaborative of research libraries across the US, Canada, and Europe. Although a “library”, the HTD has no physical space or collection. It only has a digital presence, and all its resources are located online and managed by its member libraries. HTD increased digital access to books that were once in limited collections. It currently houses 17 million digitized volumes, 95% scanned by Google, that exist entirely online. All 9.5 billion pages are indexed with metadata for each volume and shared across partner universities. The breadth of its digital collections allowed the library’s research center to create the *Bookworm Visualization Tool* that “explores the evolution of words over time; an *Extracted Features Dataset* that tallies up words and parts of speech” (Bookworm Project, n.d.). These innovative research tools utilize AI to compare and learn patterns and provide new data and approaches to analysis, critical support, and resources for researchers.

The abovementioned instances highlight the progress in accessing existing knowledge owing to technological advancements. However, technology has also altered a critical component in the democratization of knowledge, such as the revolution in the legitimization of knowledge (Isife and Nneka, 2016). Although still fraught with challenges, the abundance and breadth of information available on the Internet through various platforms weakens the hold of “gatekeepers” as “we are all equally authoritative with respect to any issue” (Kitcher, 2017, p. 3). The reliance on “expert” opinions or select groups no longer holds sway when individuals become “cognitively more autonomous” by filtering and processing the vast array of information available on the Internet (Kitcher, 2017, p. 3). Though in some cases this leads to acceptance of misinformation, broad access to diverse knowledge and information is ultimately a net benefit for society and democracy.

Practitioner’s knowledge

Again, it is important to ask ourselves, “who produces knowledge?” Much of our current knowledge is obtained by way of some sort of empirical research process. Whether a case study, longitudinal survey, demographic analysis, or experiment, the research, and therefore knowledge we have unceasingly validated, is often linked to something that can be observed and measured. So, what happens when other forms of knowledge that do not follow conventional frames arise?

Practitioners’ knowledge values the knowledge that already exists within people and communities. With concern for the undervaluing of practitioner knowledge within studies of student success, Bensimon (2007) writes “When I say that practitioners are missing, I am referring to the lack of scholarly and practical attention toward understanding how the practitioner—her knowledge, beliefs, experiences, education, sense of self-efficacy, etc.—affects how students experience their education.” She is concerned that

the knowledge that lives within faculty, that they gained over the years is overlooked. Bensimon (2007) goes on to discuss how “practitioners for the most part are likely not aware of what knowledge or experiences constitute their sense-making and how the judgments they make about a phenomenon [...] are shaped by that sense-making (p. 446). In other words, the knowledge that lives within practitioner’s may not always be visible to the practitioner themselves, and it is therefore imperative that we empower practitioners to locate that knowledge within themselves.

The Capacity Building Center at UCLA recognizes the importance of practitioner knowledge and is utilizing it in one of our current projects (<https://www.uclacbc.gseis.ucla.edu/>). The Capacity Assessment Platform (CAP) is a diagnostic instrument administered through a consultative process to evaluate academic, administrative, and service operations of educational institutions. CAP is designed to support data-driven capacity building in areas such as management, processes, teaching, research, and student and faculty services. The tool consists of quantitative and qualitative approaches to develop a complete appraisal of the institutions and identify areas for improvement and intervention. We emphasize participatory processes and targeted investments that contribute to the development and sustainability of institutions, communities, and nations. In aligning with the value of practitioner knowledge, we aim to consciously promote culturally sensitive innovation and help advance local solutions to local problems expressed by those affected by the problems.

Our evaluations are meant to help identify and improve areas within institutions, with the input of the practitioners of that institution. It is critical for us that at each level of assessment, we are communicating with people from that institution who can provide context for initial findings and share their understanding of the assessment. This project largely relies on improving practice that is already in place. We intend to help show the value and importance of practitioner’s knowledge through our process and evaluation. The project is collaborative and demands a constant line of communication between both parties. This evaluation centers the knowledge of the practitioners at the institution to implement the strategies that would make most sense for their university. This is not only important for longevity of projects, but because it gives power back to the people being “studied.”

Conclusion

Knowledge is inextricably linked to power and wealth. Historically, Eurocentric worldviews, narratives, and priorities have controlled the production, dissemination, and validation of knowledge. By excluding voices and perspectives from the largely non-white, Global South, Eurocentric production/dissemination/utilization of knowledge enters the contested space of power, epistemology, and the enduring legacy of imperialism.

We offer a critical perspective on how knowledge is generated and deployed; by who and to what ends? We examine knowledge in higher education; knowledge in the Global South; devaluation of Indigenous knowledge; democratization of knowledge production/access in the digital age; and the importance of “Practitioner knowledge” alongside “Expert knowledge.” Our larger goal is to challenge and correct established hierarchies for the production and utilization of knowledge. Ultimately, the quest for knowledge-and effective applications to improve our lives- is best served by approaches that maximize diversity and inclusion and that empower “numerous voices at the table.”

References

- Achebe C. 1960. No longer at ease. Heinemann Educational.
- Allen, W.R., Eeps, E.G., Guillory, E., Suh, S., Bonous-Hammarth, M., Stassen, M., 2000. Outsiders within: race, gender, and faculty status in the US higher education. In: Smith, W., Altbach, P., Lomotey, K. (Eds.), *The Racial Crisis in American Higher Education*, second ed. State University of New York Press, New York, pp. 189–220.
- Altbach, P., Reisberg, L., Rumbley, L., 2009. Trends in Global Higher Education: Tracking an Academic Revolution. UNESCO 2009 World Conference on Higher Education.
- Appiah-Mensah, P., 2007. Training as a Tool for Performance Improvement in the Modern Banking Industry: A Case Study of the Rural Banks in the Brong-Ahafo Region. Doctoral Dissertation. University of Cape Coast.
- Astin, H., Antonio, A., Cress, C., Astin, A.W., 1997. Race and Ethnicity in the American Professoriate, 1995–96. Paper Presented at the UCLA Higher Education Research Institution, Los Angeles, CA.
- Bensimon, E.M., 2007. The underestimated significance of practitioner knowledge in the scholarship on student success. *Rev. High. Educ.* 30 (4), 441–469.
- Bhambra, G.K., 2014. Postcolonial and decolonial dialogues. *Postcolonial Stud.* 17 (2), 115–121.
- Bruchac, M., 2014. Indigenous knowledge and traditional knowledge. In: Smith, C. (Ed.), *Encyclopedia of Global Archaeology*. Springer, New York, pp. 3814–3824.
- Coleman, J.S., 1955. *The Problem of Political Integration in Emergent Africa*. University of London.
- Coleman, J.S., 1971. *Nigeria: Background to Nationalism*. University of California Press, Berkeley.
- Coleman, J.S., 1995. *Nationalism and Development in Africa: Selected Essays*. University of California, Berkeley, Calif.
- Conrad, J., 1996. *Heart of Darkness*. University of Virginia Library, Charlottesville, VA.
- de Souza Santos, B., 2017. Beyond abyssal thought: from global lines to an ecology of knowledge. *Review XXX* (1).
- Delgado Bernal, D., Villalpando, O., 2002. An apartheid of knowledge in academia: the struggle over the “legitimate” knowledge of faculty of color. *Equity Excell. Educ.* 35, 169–180.
- EdTech 2.0, June 4, 2017. Edtech 2.0: The Ultimate Guide to Edtech. Retrieved October 2020, from Medium. <https://medium.com/@edtech2/the-ultimate-guide-to-edtech-88b9f984ada0>.
- EdTech Review, February 15, 2013. Retrieved October 18, 2019, from. <https://edtechreview.in/dictionary/119-what-is-edtech>.
- Fanon, F., 1961. *The Wretched of the Earth*. Grove Press, New York.
- Forster, E.M., 2005. In: Mishra, P. (Ed.), *A Passage to India*. Penguin Classics.
- Foucault, M., 1980. *Power/Knowledge: Selected Interviews and Other Writings 1972–1977*, first ed. Vintage.
- Garza, H., 1993. Second-class academics: Chicano/Latino faculty in US universities. *N. Dir. Teach. Learn.* (53), 33–41.

- Gramsci, A., 1971 (Trans. Quintin, H., Geoffrey, N.S.). In: *Selections From the Prison Notebooks*. International Publishers, New York.
- Grey, S., 2014. Indigenous knowledge. In: *Encyclopedia of Quality of Life and Well-Being Research*, pp. 3229–3232.
- Herskovits, M., 1938. Dahomey: an ancient west African kingdom. Augustin (2 Vols.).
- Herskovits, M.J., 1941. *The Myth of the Negro Past*. Harpers.
- Herskovits, M.J., 2006. *The Human Factor in Changing Africa*. Routledge, London.
- Isife, C.T., Nneka, V., 2016. Curriculum implementation and culture in tertiary institutions as a tool for sustainable development. *J. Philos. Cult. Relig.* 19, 20–25.
- Kitcher, N.M., 2017. Knowledge, democracy, and the internet. *Minerva* 55 (1), 1–24.
- Mignolo, W., 2012. *Local Histories/Global Designs: Coloniality, Subaltern Knowledge, and Border Thinking*. Princeton University Press.
- Moore, M.G. (Ed.), 2013. *Handbook of Distance Education*. Routledge, New York.
- Naidoo, R., 2003. Repositioning higher education as a global commodity: opportunities and challenges for future sociology of education work. *Br. J. Sociol. Educ.* 24 (2), 249–259.
- Patel, L., 2016. *Decolonizing Educational Research: From Ownership to Answerability*. Routledge, New York, NY.
- Patton, L.D., 2016. Disrupting postsecondary prose: toward a critical race theory of higher education. *Urban Educ.* 51 (3).
- Rainbow, P., 1991. *The Foucault Reader: An Introduction to Foucault's Thought*. Penguin, London.
- Rodney, W., 1974. *How Europe Underdeveloped Africa*. Howard University Press, Washington, D.C.
- Said, E.W., 1979a. *Orientalism*. Pantheon Books, New York.
- Said, E.W., 1979b. Reflections on recent American "Left" literary criticism. *Boundary 2*, 11–30.
- Scheurich, J.J., Young, M.D., 1997. Coloring epistemologies: are our research epistemologies racially biased? *Educ. Res.* 26 (4), 4–16.
- Sefa Dei, G.J., 1999. Knowledge and politics of social change: the implication of anti-racism. *Br. J. Sociol. Educ.* 20, 395–409.
- Simpson, L., 2004. Anticolonial strategies for the recovery and maintenance of Indigenous knowledge. *Am. Indian Q.* 28 (3&4), 373–384.
- Solórzano, D.G., Yosso, T.J., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1), 23–44.
- Stanley, C., 2007. When counternarratives meet master narratives in the journal editorial-review process. *Educ. Res.* 36, 14–24.
- Trinh, T.M., 1989. *Woman, Native, Other: Writing Postcoloniality, and Feminism*. State University Press, Bloomington, IN.
- Wilder, C.S., 2013. *Ebony and Ivy: Race, Slavery, and the Troubled History of America's Universities*. Bloomsbury Publishing, New York.

Regimes of research and development funding in higher education

Bailey A. Brown, Department of Sociology and Anthropology, Spelman College, Atlanta, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

History of research funding in higher education	108
Theory and the distribution of funding models in higher education	108
Higher education funding across countries	109
Trends in research funding and current challenges	110
Future implications in higher education grant funding	111
References	112

History of research funding in higher education

Higher education institutions have evolved substantially over the last century with significant advancements in funding in the last several decades. Grant funding in higher education is often used to implement projects and support innovative research. The demand for higher education increased during the 1940s through the 1960s and pushed university systems to enhance knowledge production through the development of research centers and programs. Higher education enrollment is linked to a variety of advances including economic growth, social development, and political stability across the world (Buckner and Khoramshahi, 2021).

The demand for research funding has increased alongside the demand for public education leading to a greater supply of funding from public and private entities. Federal government spending toward research has increased since World War II (Benner and Sandström, 2000). During the last 50 years, the increase in funding led to a rise in the number of published articles, issued patents and faculty member salaries (Payne and Siow, 2003). Federal governments are often incentivized to provide grants to stimulate economic growth, advance prosperity and wealth, improve social and economic wellbeing, fund ground-breaking research, and maintain a competitive global advantage. Total funding has increased substantially across countries since the middle of the nineteenth century (Aagaard et al., 2019).

Over time, investments in higher education research have driven broader economic development and progress at the local, state and national levels (Atkinson and Stewart, 2011). Grants generally fund a variety of social justice, law enforcement, social services, and health/medical initiatives. Gains in scientific discoveries and research achieve national objectives to drive economic growth. Universities are able to conduct early-stage research which can promote long run scientific development which the private sector typically draws on for ideas and innovation for the greater public.

Over the last two decades research and development as a percentage of GDP has doubled in the world's wealthiest countries (Bloch and Sørensen, 2015). In order to maintain an advantage in a globalized economy, countries rely on universities to create and disseminate knowledge to the greater public (Himanen et al., 2009). Research performed at universities is essential to supporting production within the private sector. Over the last three decades, as private corporations have had to close down research laboratories, US universities have driven American global innovation and leadership. Grants to higher education institutions play a critical role in funding public and private service projects and discoveries that have a broad impact on global economic competitiveness.

Research and Development funding to higher education institutions in the United States has led to significant technological and communication advancements, positioning the United States as a global leader (Singer, 2014; Ulferts et al., 2021). The US has invested over \$53 billion in academic science research in recent times accounting for nearly 60% of total research funding to universities (Lanahan et al., 2016; Payne, 2001). A large portion of federal spending in the United States has been geared toward higher education grants since the late 1970s. From 2000 to 2005 the annual mean size of awards from the National Science Foundation has increased by 41% (Wagner and Alexander, 2013).

Theory and the distribution of funding models in higher education

In recent times, universities rely on financial incentives to promote research development. A large number of public and private entities provide grants each year. In the United States the Bill and Melinda Gates Foundation, the Ford Foundation, the Andrew W. Mellon Foundation, the Spencer Foundation, and the National Science Foundation are some of the top organizations and programs providing grant support. The development of these organizations in the United States and similar organizations abroad are tied to important theories around economic growth and research and development capacity. Postwar research conceptualized the relationship between research funding in higher education and innovation outcomes (Szarowská, 2017).

The Solow-Swan model argues that economic growth and industrial expansion requires up-to-date technological advances in research (Guerrini, 2006). From this conceptual framework, capital growth, labor growth, and productivity all operate according to levels of technological progress. According to the Solow-Swan model, research and development funding is central to economic policy. Studies have found that research and development expenditures were correlated with economic growth in both the United States and the European Union during the 1990s and early 2000s. Global GDP growth then reflects the adoption of new technological ideas and the diffusion of technological change across borders.

Several important funding theories outline the role of funding models in research and the relationship between higher education research and state agendas (Himanen et al., 2009). Some scholarship argues that in the context of higher education, two dichotomous models of state steering occur. The relationship between the state and higher education funding is either tightly or loosely linked (Gornitzka and Maassen, 2000; Neave and Vught, 1991). Other findings suggest that in order to compare and analyze how the funding relationship between states and universities shifts over time, additional models of steering relationships are necessary. During the late 1980s Olsen theorized that four models including the sovereign state, institutional state, corporate-pluralist state, and classical liberal state/state supermarket model more accurately described the ever-changing relationship between the state and the university (Olsen, 1988).

The sovereign state model suggests that higher education is an instrument that the national government uses to achieve broad political, economic or social goals. In contrast, the institutional steering model follows that universities exercise autonomy and must protect academic findings and traditions so that higher education institutions are not at the whim of national governments. The corporate-pluralist model seeks to recognize the multiple stakeholders that exercise authority and control in higher education including faculty, staff, and students. Lastly, the classical liberal state or supermarket steering model follows that the state plays a relatively minimal role in higher education and that the primary objective of the university is to conduct teaching and research.

More recent theory points to differences in the allocation of research funding based on specific characteristics of the research. Theory suggests that the properties of funding arrangements either discourage or promote different types of research (Laudel and Gläser, 2014). Researchers' practices of producing new knowledge depend on the political and governmental conditions for research. The epistemic and institutional properties of research projects shift depending on funding parameters. Laudel and Gläser (2014) find that relational properties, particularly the planned impact of the research project and the local properties of the individual characteristics of the project work together to alter the chances that a project receives grant funding. Research funding in higher education is dependent on the match between the epistemic properties of the research and the institutional conditions under which funding can be made. Research and development expenditures are also tied to conceptions of economic growth and the relationship between higher education institutions and the state.

Higher education funding across countries

The global context of research and development is central to understanding regimes of higher education funding. A substantial majority of research on funding regimes has centered on the US and Canada and parts of Europe. In countries outside of the United States, particularly countries in Europe, research and development expenditures have grown significantly. For example, universities in the UK and Finland have had some of the most competitive research funding environments (Auranen and Nieminen, 2010). Research has also become more highly regulated across countries outside of the United States. For instance, the British government evaluates university research to ensure an equitable distribution of support. In this case, research funding is allocated at the national level according to numbers of faculty and students. In other countries like Australia, Norway and Finland, there are fewer regulations and more incentives for universities to compete for funding (Auranen and Nieminen, 2010).

Countries outside of the United States have increased their commitment to funding scientific development. In the United States, funding is at the discretion of Congress and in recent years the United States has faced challenges with meeting federal funding needs (Atkinson and Stewart, 2011). Decreases in funding at the federal level in the United States may have indirect consequences on research investments from non-federal sources. In other contexts, like China for instance, research funding makes up a larger share of GDP growth at 59% (Atkinson and Stewart, 2011). China has become a leader in research and development based on the number of scientists engaged in research and the level of investment (Boeing, 2016). Similarly, Korea, Ireland, and the UK all increased research funding amounts over 30%, nearly double the rate of research funding growth in the United States (Atkinson and Stewart, 2011). In the UK two-thirds of private sector productivity during the early 2000s can be attributed to innovation in higher education. Australia and Norway for instance have seen a steady increase in state funding for research and development since the 1990s. The United States' decrease in funding from the federal government risks American global competitiveness in the long run (Atkinson and Foote, 2019). Other developed and developing countries invest a larger percentage of GDP in university research leading to more innovation-based jobs and larger economic returns. In fact, western European countries have consistently ranked among the top ten countries in terms of research and development spending (Dvouletý et al., 2020).

Funding agencies across countries also vary in the evaluation criteria used to determine the potential of scientific research and the allocation of large and small grants. Some national funding agencies emphasize societal relevance and broader implications, as is the case in Australia, Canada, France, Portugal, Netherlands, UK, and the United States (Meirmans et al., 2019). In these countries funding agencies require researchers to translate the science to the public policy arena. In other countries like Switzerland, Germany, and Sweden the emphasis on the societal impact is less necessary and rarely required. Funding distribution also varies across countries. For example, France, parts of the UK, Switzerland, Australia, and Norway often provide research to large teams of researchers

rather than individuals (Lewis and Ross, 2011). In contrast Canada, Germany, and the Netherlands provide grants to individual-led projects often in which a single applicant is preferred (Meirmans et al., 2019).

Despite important differences in allocation and broader research impact, countries across the world similarly provide grant funding that lasts for one to five years with most countries also accommodating flexibility in budget allocations. Countries across Europe and the United States also typically review grant proposals using a panel of experts over one or two rounds of review. In most countries, the panel includes external reviewers in addition to a panel of experts. Applications are also similarly judged and evaluated across countries. Evaluations typically focus on the quality of the project and of the researcher or research team. Some countries look at the novelty of the research like in Sweden, or the track record of the researcher like in Switzerland.

There is substantial variation in the expected length and structure of grant proposals across countries which creates different administrative burdens for researchers and research teams. In Australia, France, Portugal, and the United States, the burden to submit a proposal is fairly high with proposals averaging near 20–25 pages with some countries involving multiple rounds of evaluation (Meirmans et al., 2019). Funding agencies across countries also vary in their funding of diverse and mobile projects. Some funding agencies like those in Switzerland, Germany, and Sweden provide funding for use outside of the home country. In contrast, nations like the United States, France, and Australia encourage within-country research. Differences in mobility allowances can significantly shape scientific breakthroughs and discoveries as research advances when disciplines are crossed and when a variety of international experience can be combined. Overall important variation exists in how governments allocate funding and how researchers gain access to funding.

Trends in research funding and current challenges

Several trends over the last two decades have influenced funding at the university level. Universities have developed an increasing number of centers and large programs that are supported by large-scale project grants (Bloch and Sørensen, 2015). Over the last three decades research funding models have shifted from block funding for universities toward project funding arrangements (Lepori et al., 2007). The shift in funding toward centers has meant that a greater share of research funds is concentrated among a smaller number of researchers within a given center (Bloch and Sørensen, 2015). The project funding arrangement has implications for the organization of research and has increased the level of authority that university management teams and research group leaders have relative to individual scholars. Rather than supporting individual projects, grants are now more cooperative, supporting a number of researchers under the same center (Bozeman and Boardman, 2003).

Another important trend has been the share of funding from federal and non-federal sources. Direct government funding often provides the core funding that is allocated to assist universities in maintaining the basic infrastructure necessary to perform research activities. Federal funding has made up a large part of this increase in funding nearing 60% of total funding given to universities in the United States during the 1990s (Hicks and Katz, 2011). In recent decades, the share of direct government funding has decreased especially in the United States, and the share of external funding, especially funding from private industries has increased (Auranen and Nieminen, 2010). Federal funding ensures that research investment from private corporations and non-profit organizations also increases. While academic institutions gain resources from a myriad of sources, federal investment plays an important role, so much so that decreases in federal research funding result in decreases in research investment from other private and non-profit sources (Muscio et al., 2013). In past decades, federal funding helped ensure larger research outputs but with recent decreases in federal funding, universities rely on private companies to make up the difference.

Research finds that federal funding crowds in investments from other private and industry sources (Lanahan et al., 2016). Federal funding is an important complement for additional investments. Although the impact of federal funding varies considerably depending on the academic discipline, the form of institutional control, and the capacity of the field research, federal dollars help bring in non-federal funding to achieve broader societal objectives. Federal and non-federal funders work in a complementary relationship to ensure the development of scientific knowledge and discovery. Federal investments help to advance national directives and drive economic expansion and non-federal investments drive commercial innovation and provide a broader societal benefit for non-profit organizations and local economies.

Government funding toward research seeks to promote research efficiency. Government funding is often broadly distributed to maximize economic development and to enhance diversity and participation in research (Hicks and Katz, 2011). In recent times, global competition has also pushed the concentration of research funding in the hopes of driving technological advancements over other countries. As funding becomes increasingly concentrated in the hands of a single researcher or a group of researchers, these individuals and teams gain a cumulative advantage in which each new success is leveraged based on past accomplishments.

Across countries, researchers with substantial early funding will continue to outperform researchers who did not have access to this earlier funding source. Continuing to reward past top research performers may limit the funding potential for future top research performers (Bloch and Sørensen, 2015). Instead of concentrating research funding, studies suggest that dispersal of funding can reduce cumulative advantages (Berg, 2012; Fang and Casadevall, 2016). A concentration of funding for a single individual or within a particular discipline or for a particular center may decrease the level of diversity that could be achieved if funding supported multiple smaller projects (Horta et al., 2008). This can be particularly detrimental for researchers in smaller countries who may narrow their disciplinary focus in line with specific research initiatives. In this case, rather than funding areas with the greatest potential, research funding is concentrated in a very specific field. Research concentration in this form can lead to more economic inefficiency (Bloch and Sørensen, 2015; Hicks and Katz, 2011; Mongeon et al., 2016; Nag et al., 2013; Sousa, 2008).

Arguments in favor of the concentration of research claim that research is more efficient once a critical mass of funding is obtained. Once a critical mass is reached, researchers have greater access to human resources, equipment, and infrastructure. Pooling of such resources can allow for the development of large-scale research projects that would be impossible without a large amount of funding. Under this premise, the resources contained within centers or under largescale research projects are necessary for promoting rigorous scientific research.

Evidence also suggests that grants of greater size also create more research visibility which can help communicate benefits to the greater public and create a more stable funding base. Concentrated funding is believed to promote researcher interaction and mutual learning and promote increased productivity (Lee and Bozeman, 2005). Under this arrangement resources are pooled to achieve greater strides in teaching and administration that would not be possible under individual grants. Fewer grants concentrated within a smaller number of centers may also result in a lessened administrative burden with fewer barriers for achieving scientific excellence (Bloch et al., 2016; Bloch and Sørensen, 2015; Breschi and Malerba, 2011; Hellström et al., 2018; Hicks and Katz, 2011).

While some studies favor research concentration, there is little evidence that increasing grant size necessarily leads to increased research output and productivity (Aagaard et al., 2019). Research also finds diminishing marginal returns as grant sizes increase above a certain threshold. Research output per researcher decreases even with increasing amounts of funding (Asonuma and Urata, 2015). As an example, while grant funding over time has led to increases in publications, recent research finds that funding success is only moderately associated with gains in publication and citation rates (Fedderke and Goldschmidt, 2015). Other research finds no association between the level of funding a center receives and research output (Gaughan and Bozeman, 2002). Across most studies, grant volume only slightly improves performance (Laudel, 2005).

Evidence suggests that an increased dispersal of funding has a positive impact on research productivity. For these reasons, findings tend to support a greater dispersal of research funding and a more efficient allocation of resources. Dispersal also ensures that knowledge is broadened and that there is greater research breadth (Fortin and Currie, 2013; Katz and Matter, 2017; Vaesen and Katzav, 2017). In contrast, a concentration of research funding can lead to greater marginal costs for researchers with diminishing marginal returns (Berg, 2012; Breschi and Malerba, 2011; Lorsch, 2015).

In this vein, a diverse allocation of research funding increases the chances of innovative breakthroughs (Fang and Casadevall, 2016). If research is spread across many individuals and multiple centers, then the chance for scientific developments increases. At the same time diversification of research funds also minimizes the risk that research funding is wasted. If one research group fails, more allocated funding could ensure that another researcher would meet success. Research funding dispersal also keeps students and researchers active and engaged in new developments. Allocated funding ensures that early and mid-career researchers are funded and that student researchers gain access to funding opportunities.

As funding sources have become more concentrated under large projects and centers, administrative burdens increase. Large grants must be appropriately managed and regulated and rather than focusing on research development, grant winners may be inundated with more bureaucratic requirements. Rather than promoting research innovation, researchers focus substantially more attention on grant fundraising and grant management (Alberts, 1985; Bloch and Sørensen, 2015). Large funding incentives also create significantly more competition which can in turn affect research productivity. For instance, a researcher competing regularly for funding may in fact spend more time and energy engaged in the grant writing process than in the research development process. Funding incentives can also unintentionally weaken the quality of research if the prestige of gaining quantitatively more grants outweighs the benefit of stronger research development (Butler, 2003; Langford et al., 2006). Overall, the allocation of research and development funding continues to be critical for the future of research productivity across higher education institutions.

Future implications in higher education grant funding

As research development and spending increases in countries outside of the United States, future research should center on country-specific and regional-specific evaluations of funding incentives and research productivity. To date, most studies of funding regimes center on North America, Europe, and Australia. Research funding is largely controlled by developed countries at the expense of developing countries (Harris, 2004; Kitzman et al., 2010). Global unequal relations of power constrain the research capabilities of developing countries. Future research should investigate methods for improving the potential research base in developing countries. Further, research collaborations should center on building and fostering scientific development in countries with fewer resources. Studies that examine global trends in research funding and that enhance the scientific capacity of developing countries will substantially enhance knowledge around research productivity. More research is needed to understand the funding models that countries use to allocate funding and how countries use federal and non-federal funding as investments.

Research funding in recent times faces challenges related to research concentration and inequality in grant dispersion. Future research should expand knowledge around funding concentration and how the size of grants shapes research performance and other outcome measures (Butler, 2010). Current research on funding concentration is fragmented and inconsistent. Future research should also investigate what a concentration of research funding means for gender and racial imbalances. Some research finds that an increased focus on large research grants in centers may exacerbate gender inequalities (Aksnes et al., 2012). Future research should determine how to distribute research funding more equitably to ascertain social and economic impacts and to address any adverse effects on equity. We currently have a limited understanding of the internal mechanisms and the external policies

that have resulted in the concentration of research funding. If evidence suggests that the concentration of funding can limit the diversity of research, then future research should determine how to best redistribute research funds.

Current evidence also suggests that public research and development funding has been declining over time. Future research should investigate comparative levels of federal and non-federal funding and trends over time. This strain of research might also suggest potential policy reforms that would ensure that government funding toward research and development is appropriately allocated. Private research and development funding is also an important area of future research. Future studies should determine the precise role of industry funding in higher education research and development.

In all, worldwide funding limitations still constrain the ability of federal and non-federal agencies to support scientific discovery. Future research is needed to ascertain the most effective way to allocate funding to support scientific exploration and to achieve the related economic and societal advances. Funding scarcity often creates unnecessary competition for limited research grants. Future studies should investigate and compare funding allocation practices and various approaches to the distribution of funding. Given that it is unlikely that available governmental resources will radically increase over time, funding agencies have the critical responsibility of determining the most efficient ways to allocate funding that will generate the greatest returns. Current research points to more funding allocation for small grants to multiple researchers and the funding of both center-based and individual-led projects. Allocation efficiency that produces positive outcomes may also improve if more long-term projects are funded, if there is more flexibility in funding schemes, and if there are reductions in the administrative burden of applying for funding.

As universities continue to rely on funding agencies for research support and as funding agencies grow, scholarship should continue to evaluate how social, economic, industrial and political interests are integrated into the evaluation and performance of research in higher education institutions. Performance-based systems have driven grant funding in the last decades and how performance measures are regulated will shape the future of research innovation and funding opportunities.

Given the COVID-19 impact of 2020 and the surge in research and development funding toward vaccine production, future research will need to assess short-term and long-term funding toward research development in other academic areas within higher education. The impact on research development spending is still unclear and policymakers and research funders are unsure of whether or not similar biomedical funding levels can be sustained in the long-term.

Overwhelming evidence demonstrates that investment in research and development is essential for productivity, growth, and innovation on a global scale. The loss of research and development infrastructure can have detrimental societal effects. Continued evaluations of research productivity and the impact of sustained long-term investments in research will be valuable for understanding future trends in scientific funding arrangements.

References

- Aagaard, K., Kladakis, A., Nielsen, M.W., 2019. Concentration or dispersal of research funding? *Q. Sci. Stud.* 1 (1), 117–149. https://doi.org/10.1162/qss_a_00002.
- Aksnes, D.W., Schneider, J.W., Gunnarsson, M., 2012. Ranking National Research Systems by Citation Indicators. A comparative analysis using whole and fractionalized counting methods. *J. Informetr.* 6 (1), 36–43.
- Alberts, B.M., 1985. Limits to growth: in biology, small science is good science. *Cell* 41 (2), 337–338.
- Asonuma, A., Urata, H., 2015. Academic funding and allocation of research money. In: *The Changing Academic Profession in Japan*. Springer, pp. 57–77.
- Atkinson, R.D., Foote, C., 2019. US Funding for University Research Continues to Slide. Information Technology and Innovation Foundation.
- Atkinson, R.D., Stewart, L.A., 2011. *University Research Funding: The United States is Behind and Falling*, vol. 18.
- Auranen, O., Nieminen, M., 2010. University research funding and publication performance—an international comparison. *Res. Pol.* 39, 822–834. <https://doi.org/10.1016/j.respol.2010.03.003>.
- Benner, M., Sandström, U., 2000. Institutionalizing the triple helix: research funding and norms in the academic system. *Res. Pol.* 29 (2), 291–301. [https://doi.org/10.1016/S0048-7333\(99\)00067-0](https://doi.org/10.1016/S0048-7333(99)00067-0).
- Berg, J.M., 2012. Well-funded investigators should receive extra scrutiny. *Nature* 489 (7415), 203.
- Bloch, C., Sørensen, M.P., 2015. The size of research funding: trends and implications. *Sci. Publ. Pol.* 42 (1), 30–43. <https://doi.org/10.1093/scipol/scu019>.
- Bloch, C., Schneider, J.W., Sinkjær, T., 2016. Size, accumulation and performance for research grants: examining the role of size for centres of excellence. *PLoS One* 11 (2), e0147726.
- Boeing, P., 2016. The allocation and effectiveness of China's R&D subsidies-evidence from listed firms. *Res. Pol.* 45 (9), 1774–1789.
- Bozeman, B., Boardman, P.C., 2003. *Managing the New Multipurpose, Multidiscipline University Research*. IBM Center for the Business of Government.
- Breschi, S., Malerba, F., 2011. Assessing the scientific and technological output of EU framework programs: evidence from the FP6 projects in the ICT field. *Scientometrics* 88 (1), 239–257.
- Buckner, E., Khoramshahi, C., 2021. Does the private sector expand access to higher education? A cross-national analysis, 1999–2017. *Int. J. Educ. Dev.* 84, 102410.
- Butler, L., 2003. Modifying publication practices in response to funding formulas. *Res. Eval.* 12 (1), 39–46.
- Butler, L., 2010. Impacts of Performance-Based Research Funding Systems: A Review of the Concerns and the Evidence, pp. 127–165. <https://doi.org/10.1787/9789264094611-7-en>.
- Dvouléty, O., Srhoj, S., Pantea, S., 2020. Public SME grants and firm performance in European Union: a systematic review of empirical evidence. *Small Bus. Econ.* 1–21.
- Fang, F.C., Casadevall, A., 2016. Research funding: the case for a modified lottery. *Am. Soc. Microbiol.* 7 (2), e00422.
- Fedderke, J.W., Goldschmidt, M., 2015. Does massive funding support of researchers work? Evaluating the impact of the South African research chair funding initiative. *Res. Pol.* 44 (2), 467–482.
- Fortin, J.-M., Currie, D.J., 2013. Big science vs. little science: how scientific impact scales with funding. *PLoS One* 8 (6), e65263.
- Gaughan, M., Bozeman, B., 2002. Using curriculum vitae to compare some impacts of NSF research grants with research center funding. *Res. Eval.* 11 (1), 17–26.
- Gornitzka, V.A., Maassen, P., 2000. Hybrid steering approaches with respect to European higher education. *High Educ. Pol.* 13 (3), 267–285.
- Guerrini, L., 2006. The solow–swan model with a bounded population growth rate. *J. Math. Econ.* 42 (1), 14–21.
- Harris, E., 2004. Building scientific capacity in developing countries: simply transferring knowledge and instrumentation is not enough to help developing countries build their own research base. Such efforts must be tied to national and local needs to create trust and services for society in the long term. *EMBO Rep.* 5 (1), 7–11.
- Hellström, T., Jabrane, L., Brattström, E., 2018. Center of excellence funding: connecting organizational capacities and epistemic effects. *Res. Eval.* 27 (2), 73–81.

- Hicks, D., Katz, J.S., 2011. Equity and excellence in research funding. *Minerva* 49 (2), 137–151. <https://doi.org/10.1007/s11024-011-9170-6>.
- Himanen, L., Auranen, O., Puuska, H.-M., Nieminen, M., 2009. Influence of research funding and science policy on university research performance: a comparison of five countries. *Sci. Publ. Pol.* 36 (6), 419–430. <https://doi.org/10.3152/030234209X461006>.
- Horta, H., Huisman, J., Heitor, M., 2008. Does competitive research funding encourage diversity in higher education? *Sci. Publ. Pol.* 35 (3), 146–158. <https://doi.org/10.3152/030234208X299044>.
- Katz, Y., Matter, U., 2017. On the Biomedical Elite: Inequality and Stasis in Scientific Knowledge Production. Berkman Klein Center Research Publication (2017–5).
- Klitzman, R., Chin, L.J., Rifai-Bishjawish, H., Kleinert, K., Leu, C.-S., 2010. Disclosures of funding sources and conflicts of interest in published HIV/AIDS research conducted in developing countries. *J. Med. Ethics* 36 (8), 505–510.
- Lanahan, L., Graddy-Reed, A., Feldman, M.P., 2016. The domino effects of federal research funding. *PLoS One* 11 (6), e0157325. <https://doi.org/10.1371/journal.pone.0157325>.
- Langford, C.H., Hall, J., Josty, P., Matos, S., Jacobson, A., 2006. Indicators and outcomes of Canadian University Research: proxies becoming goals? *Res. Pol.* 35 (10), 1586–1598.
- Laudel, G., Gläser, J., 2014. Beyond breakthrough research: epistemic properties of research and their consequences for research funding. *Res. Pol.* 43 (7), 1204–1216. <https://doi.org/10.1016/j.respol.2014.02.006>.
- Laudel, G., 2005. Is external research funding a valid indicator for research performance? *Res. Eval.* 14 (1), 27–34. <https://doi.org/10.3152/147154405781776300>.
- Lee, S., Bozeman, B., 2005. The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* 35 (5), 673–702.
- Lepori, B., Van den Besselaar, P., Dinges, M., Van der Meulen, B., Poti, B., Reale, E., Slipersaeter, S., Jean, T., 2007. Indicators for comparative analysis of public project funding: concepts, implementation and evaluation. *Res. Eval.* 16 (4), 243–255.
- Lewis, J.M., Ross, S., 2011. Research funding systems in Australia, New Zealand and the UK: policy settings and perceived effects. *Pol. Polit.* 39 (3), 379–398. <https://doi.org/10.1332/030557310X520270>.
- Lorsch, J.R., 2015. Maximizing the return on taxpayers' investments in fundamental biomedical research. *Mol. Biol. Cell* 26 (9), 1578–1582.
- Meermans, S., Butlin, R.K., Charmantier, A., Engelstädter, J., Groot, A.T., King, K.C., Hanna, K., Reid, J.M., Neiman, M., 2019. Science policies: how should science funding be allocated? An evolutionary biologists' perspective. *J. Evol. Biol.* 32 (8), 754–768.
- Mongeon, P., Brodeur, C., Beaudry, C., Larivière, V., 2016. Concentration of research funding leads to decreasing marginal returns. *Res. Eval.* 25 (4), 396–404.
- Muscio, A., Quagliione, D., Vallanti, G., 2013. Does government funding complement or substitute private research funding to universities? *Res. Pol.* 42 (1), 63–75. <https://doi.org/10.1016/j.respol.2012.04.010>.
- Nag, S., Yang, H., Buccola, S., Ervin, D., 2013. Productivity and financial support in academic bioscience. *Appl. Econ.* 45 (19), 2817–2826.
- Neave, G.R., van Vught, F.A., 1991. *Prometheus Bound: The Changing Relationship Between Government and Higher Education in Western Europe*. Pergamon Press.
- Olsen, J.P., 1988. Administrative reform and theories of organization. *Organ. Gov.* 233, 254.
- Payne, A.A., Slow, A., 2003. Does federal research funding increase university research output? *B. E. J. Econ. Anal. Pol.* 3 (1). <https://doi.org/10.2202/1538-0637.1018>.
- Payne, A.A., 2001. Measuring the effect of federal research funding on private donations at research universities: is federal research funding more than a substitute for private donations? *Int. Tax Publ. Finance* 8 (5), 731–751. <https://doi.org/10.1023/A:1012843227003>.
- Singer, P.L., 2014. *Federally Supported Innovations: 22 Examples of Major Technology Advances that Stem From Federal Research Support*, vol. 12. Information Technology and Innovation Foundation.
- Sousa, R., 2008. Research funding: less should be more. *Science* 322 (5906), 1324–1325.
- Szarowská, I., 2017. Does public R&D expenditure matter for economic growth? *J. Int. Stud.* 10 (2).
- Ulferts, G.W., Cannon, N.J., Howard, T.L., 2021. The changing landscape for funding research and development. *Res. Higher Educ. J.* 40.
- Vaesens, K., Katzav, J., 2017. How much would each researcher receive if competitive government research funding were distributed equally among researchers? *PLoS One* 12 (9), e0183967.
- Wagner, C.S., Alexander, J., 2013. Evaluating transformative research programs: a case study of the NSF small grants for exploratory research program. *Res. Eval.* 22 (3), 187–197.

The hollowing out of academic freedom from African universities; from external capture to internal implosion

Ibrahim Ogachi Oanda, Council for the Development of Social Science Research in Africa (CODESRIA), Dakar, Senegal

© 2023 Elsevier Ltd. All rights reserved.

Introduction	114
Establishment of national universities in Africa and the quality of academic freedom in practice	115
1980–2000: corporatization of universities and undermining of conditions for academic freedom	118
Post-2000 developments: intensity and quality of academic freedom in African universities	120
Conclusion	122
References	122

Introduction

The conditions for the exercise of institutional autonomy and academic freedom have improved in many African universities over the last two decades, since the eve of the 21st century. To a greater extent, these improvements have been an indirect outcome of the fiscal crisis of the state in Africa and the imposition of structural adjustment program (SAPs) in the 1980s and 1990s. Implementation of SAPs in the midst of a fiscal crisis was used as a carrot to persuade African countries undertake political and economic reforms, part of which required restructuring in the governance structures of universities in ways that made direct state involvement in the management of the institutions untenable. Politically, one party-governments that in the past undermined institutional autonomy and academic freedom in the universities gave way to multiparty governments, through constitutional reforms, with the new constitutions embracing a bill of rights, including academic freedom. Economically, the declining public funding to support universities forced governments to accord a greater degree of institutional autonomy to universities through the creation of new governance bodies that would allow universities greater latitude in generating own operational resources. Recent surveys on the state of academic freedom in the continent show that by 2016, 14 (25.45%) of the 55 African countries explicitly recognize academic freedom in their constitutions, 8 (12.7%) of the countries implicitly recognize the exercise of certain aspects of academic freedom, such as scientific research as a right in their constitutions, while the remaining 34 countries indirectly acknowledge the exercise of academic freedom in their constitutions (Kwadwo et al., 2016).

The constitutional provision and/or recognition of the right for the exercise of academic freedom does create an expectation that conditions for academics to exercise academic freedom largely prevail across the continent and within the institutions. However, recent surveys on the quality of academic freedom in the continent show mixed results, that tend towards an increasing abrogation of academic freedom, both directly and indirectly in most of the countries. The 2007 report by Education International, based on the EI barometer does suggest widespread limitation of academic freedom throughout the continent. Education International undertakes regular analyses of trends in human and trade union rights in the education sector across several issues, including academic freedom. The report shows that of the 50 African countries whose reports were analyzed, 15 (30%) cited no violations or existence of formal restrictions; 15 (30%) of countries had no formal restrictions but institutional practices which limited academic freedom were prevalent; 16 (32%) of countries had formal restrictions on academic freedom through widespread legislations; 1 (2%) had severe formal restrictions while 3 (6%) showed recent evidence of improvement in academic freedom (Educational International, 2007). Aggregating together the countries that had incidences of academic freedom violations despite no formal restrictions, with those that reported formal restrictions through legislation and those with severe cases would suggest that by 2007, violations to academic freedom was prevalent in 32 African countries (64%). This situation was attributed to the emergence of new challenges to the traditional ideal of academic freedom, coupled with new technologies, restructured traditional universities, an expanded private higher education sector that includes for-profit institutions, a privatized public higher education sector, and the emergence of corporatized research in the universities (Altbach, 2007: 7).

More recently, a March 2021 report by the Global Public Policy Institute and Scholars at Risk Network (Kinzelbach et al., 2021), does suggest that in a number of African countries new threats to academic freedom are emerging, while the old ones are still in existence, reflecting a largely a global trend. The report measures the degree of academic freedom using five indicators; freedom to research and teach; freedom of academic exchange and dissemination; institutional autonomy; campus integrity; and freedom of academic and cultural expression. Of the 44 countries whose data was available, 9 countries (20%) had an academic freedom index suggesting on average they were fully free. On the other hand, 13 (30%), 11 (25%), 5 (11%), and 6 (14%) were mostly free, moderately restricted, severely restricted and completely restricted (Kinzelbach et al., 2021). While these data does suggest some improvement, it also shows that in a majority of countries, some restrictions on academic freedom do exist despite the greater constitutional recognition.

Historical literature on academic freedom in Africa largely focuses on two indicators; the freedom of an academic to teach and undertake scientific work and the institutional autonomy to guarantee the right environment for teaching and research. Thinking of

academic freedom together with the other notions of intellectual social responsibility and institutional autonomy as issues that are supposed to co-exist simultaneously however switches the focus from external restraints to academic freedom and institutional autonomy to examining ways in which individual academics and institutions do in fact undermine academic freedom and institutional autonomy. In other words, academic freedom is only tenable when coupled with intellectual social responsibility and institutional autonomy, while institutional autonomy is justified if only to guarantee conditions for socially responsible intellectual engagements by academics. So, while in the past academic freedom was framed in the context of the relationship between universities and authoritarian states (Grimm and Saliba, 2017), this article will argue that the increased constitutional recognition of academic freedom at different dimensions and the yielding of the state to some degrees of institutional autonomy throughout Africa has not resulted to conditions for real academic freedom. Rather new challenges have emerged from within the institutions, and often with the culpability of the academics themselves, that have hollowed out the capacity of the institutions and the academics to exercise academic freedom and social responsibility. The management of the institutions and the academics have yielded the space vacated by the state to newer actors in the name of responding to new managerial imperatives in ways that raise new questions about the autonomy of academics and institutions from the new forces that claim to support the universities, at least financially. In the name of responding to the needs of industry for example, academic programs have been introduced with little input from academics while the participation of university senates and faculty in the academic affairs of the institutions have been limited to basic reporting and approval practices (Fulda et al., 2021). The trends accord academics little ownership of the academic issues of the institutions as would be expected in situations where real academic freedom and intellectual responsibility are exercised. The UNESCO (1998) declaration on academic freedom does emphasize a greater degree of self-governance by academics in matters to do with academic work and accountability by academics on the quality of research and teaching. The weakening of the oversight role of academics and university senates on the academic processes of the institutions does seem to have weakened the academic integrity of the institutions, with instances of academic fraud and poor quality of academic programs emerging as a major concern affecting universities across the continent, post-2000 (Zezeza, 2021).

The subsequent sections of this chapter draw from historical and contemporary analysis of trends in academic freedom and social responsibility in African universities, both in theory and practice. The article first discusses conceptions of academic freedom and its limited practice in African universities within the first two decades of independence. The second part discusses the contestations over academic freedom in the era of what has been referred to as “corporate” universities. Lastly the nature of academic freedom under new institutional governance structures and its implications are discussed.

Establishment of national universities in Africa and the quality of academic freedom in practice

The notion of academic freedom and the university as a space where such a freedom has to be exercised has remained problematic in Africa, at least at the conceptual level. Recent calls for the decolonization of aspects of university spaces and the content of courses points to this unfinished historical business. Although some literature does suggest that the establishment of national universities in Africa at the onset of independence heralded the era of contestations between the state and academics on issues academic freedom (African Watch, 1991), such literature presents a partial picture of what has been otherwise a complicated relationship. Academic freedom and what this would constitute in the African context remained problematic to the African political leadership even before independence. This is because the notion of academic freedom and the associated concepts of institutional autonomy and social responsibility derived from the package proposed by the Asquith commission for the establishment of universities in British colonial Africa of what would constitute a “university” then (Report of the Commission of Higher Education in the Colonies, 1945; Hargreaves, 1973). The Asquith Commission Report was clear that the envisaged university institutions to be established in the colonies needed to enjoy the same level of institutional autonomy and academic freedom similar to what existed then in British universities. The Report for example did recommended that academics in the new universities needed to focus on fundamental research devoid of any utilitarian considerations, and placed the obligations for providing the funding and sufficient climate within the institutions for the furtherance of these objectives on the governing bodies of the universities (Report of the Commission of Higher Education in the Colonies, 1945: 27). In practice however, especially regarding governance and the academic freedom of African staff, the new institutions lacked sufficient autonomy as either the secretary of state or the governor had more powers in appointing members of the governance bodies, with less representation from faculty members (Dlamini, 2002).

Between 1947 and 1970, universities were established in British colonial Africa, following the model proposed by the Asquith Commission, with, strong desire to preserve the traditions of academic freedom and institutional autonomy (Ashby and Anderson, 1966). The special relationship the universities had with the University of London influenced writing into the instruments for the constitution of the new universities’ frameworks meant to ensure autonomy of the institutions and the freedom of the academic communities, including the intellectual values of research, but exactly in the manner this obtained in British universities (Dlamini, 2002). Ashby and Anderson (1966) summarized the expectations of academic freedom and institutional autonomy to entail the freedom of academic staff to undertake the following functions without any external direction, but based on rules and traditions developed by the universities: the admission and examination of students, the development of curricula for courses of study, the appointment and tenure of office of academic staff, the allocation of income among the different categories of expenditure and the freedom of individual academics to teach according to their conscience and convictions and to publish views on their own subject or on any other subject permitted by law.

Though from time to time, transplanted notions of institutional autonomy and academic freedom and their suitability to Africa were questioned and cautioned (Hargreaves, 1973; Heron, 1964; Ashby and Anderson, 1966), these early conceptualizations have remained the pivot around which academic freedom and institutional autonomy have been invoked, demanded and defended. The entrenched government involvement in constituting the governance bodies of the institutions for example was copied by African governments at independence and would become the main contention on issues of academic freedom and institutional autonomy between academics and governments in the subsequent decades. The notions of academic freedom and institutional autonomy as framed then were not wholly agreed upon, especially by African academics in the new universities. There were those who thought academic freedom and institutional autonomy as framed then overlooked the context of the university in Africa, and were at most colonial importations of universal values and ideas into African institutions to serve the needs of the small expatriate staff at the expense of the needs of African institutions in the 20th century (Hargreaves, 1973).

Besides the issue of context, two other contentious issues arose during this period, and which would characterize contestations about academic freedom and institutional autonomy in Africa in the coming decades. The first was whether academic freedom would be secured in a context devoid of institutional autonomy, or if the reverse would be possible; autonomous institutions would trample on the academic freedom of academics (Ashby and Anderson, 1966; Crawford et al., 1967; Mamdani, 1993; Kori, 2016). The second issue was whether academic freedom was an absolute right that would be secured as an end or whether it had to be exercised only as a means to something greater; some form of academic responsibility to society (Ashby and Anderson, 1966; Hargreaves, 1973; Coleman, 1977; Crawford et al., 1967; Mamdani, 1993; Chachage, 2009).

Neither are the issues of political interference in the running of the universities that have been documented as largely responsible for the dysfunctions between state-university relations and the limitations of academic freedom a post-independence development. Literature documents the fact that in establishing universities in Africa during the colonial period, the colonial governments did allocate itself greater latitude in the governance and academic life of the institutions (Oloka-Onyango, 1992). In the case of Makerere university College, for example, The 1930 Makerere College Ordinance, the first legal instrument that set guidelines for the governance of the college provided for broad state control and regulation of the operations of the college by the Secretary of State through a chairperson appointed by the conference of East African Governors (Oloka-Onyango, 1992). A subsequent amendment to this ordinance passed in 1941 made the office of the Secretary of State central in the formulation of the terms of service, appointment and dismissal of members of the college staff, and gave the Secretary of State power to appoint and remove the principal of the college with recommendations from the Council of Governors (Oloka-Onyango, 1992). Needless to say, the situation in Makerere mirrored the kind of governance and academic structures that obtained in the other Asquith colleges. The literature that portrays governments in post-Independence Africa and their involvement in the governance of universities as the genesis of the abrogation of academic freedom and institutional autonomy in institutions that were otherwise devoid of a culture of political intrusion is not correct (See for example Sifuna, 1998; Munene, 1997). The situation represented more of a continuity such that in place of the secretary of State, presidents of the independent countries took over the power that had been bestowed on the Secretary of State and went ahead to populate the governing bodies with government appointees. It can be argued that in trying to take a steering role in directing the development and academic life of the universities, the political leadership in post-independence Africa were trying to follow a script laid out by the colonial administration that founded the universities. If the colonial government directed the nature of institutional life and academic freedom in the pursuance of the colonial project, the political leadership were in a sense following the same script in the pursuance of Africanization and national development.

Immediately after African countries gained independence, the political leadership started re-establishing the colonial universities into national institutions. Accompanying the nationalization of universities were attempts to have clarity on what a university meant within the African context and what the national expectations were from the perspectives of academic freedom and institutional autonomy. The political leadership looked upon universities and the few university graduates then available to help deliver the promises of independence. Academic freedom and institutional autonomy would only be invoked and granted only to the extent that it served the needs of “developmentalism” as articulated and understood by the political leaders. In the climate of “Africanization” of the time, the political leadership was in a sense seeking to enculturate both the university as an institution and academic freedom into an African context. African intellectuals and politicians had mobilized together in the struggle for independence under the banner of African nationalism and Pan-Africanism (Hyden, 1967; Mnguni, 1987; Mkandawire, 2002). The unity between the politicians and the intellectuals in the nascent universities however didn’t last long as cracks emerged regarding the political class’s approach to development and the role of the university and academics in that process. The re-establishment of the colonial universities anew as regional and national universities in the 1960s and 1970s, was accompanied by an admonition from the political leadership of a blanket expectation of academic freedom and institutional autonomy by the academics in the institutions (Dlamini, 2002; Crawford et al., 1967). Throughout the continent, the establishment of the new universities was accompanied by contradicting declarations from the Chancellors, who happened to be the presidents of the countries of the kind of universities and the nature of academic freedom expected. In Ghana, for example, Kwame Nkrumah, while addressing graduates on 24th 1963 noted:

We know that the objectives of a University cannot be achieved without scrupulous respect for academic freedom, for without academic freedom there can be no University ... Scholars must be free to pursue the truth and to publish the results of their researches without fear, for true scholarship fears nothing ... We know that without respect for academic freedom, in this sense, there can be no higher education worthy of the name, and, therefore, no intellectual progress, no flowering of the nation’s mind

But he noted:

We know that academic freedom can be perverted and even abused. It can also become a dangerous cloak for activities outside the academic interests and preoccupations of the community or of the University. Where this has happened, a grave disservice is done to everything for which knowledge and truth really stand. True academic freedom—the intellectual freedom of the University—is everything fully compatible service to the community; for the University is, and must always remain, a living thing and serving part of the community to which it belongs.

Quoted in Kwapong (1966: 546–547).

Elsewhere in East Africa, Julius Nyerere, President of Tanzania, was invoking similar sentiments about the expectations of politicians from the university, while being installed as the first Chancellor of the university of East Africa. He noted:

The University Colleges which comprise this University cannot be islands filled with people who live in a world of their own, looking on with academic objectivity or indifference at the activities of those outside. East Africa cannot spend millions of pounds, cannot beg and borrow for the university, unless it plays a full and active part in the urgent tasks of East Africa. Even if it were desirable, we are too poor in money and educated manpower to support an ivory tower existence for an intellectual elite. Our problems will not wait. We must, and do, demand that this University take an active part in the social revolution we are engineering.

The Reporter (1963: 29).

Nyerere returned to the specifics of what he thought of a university in Africa and the purpose of academic freedom, while inaugurating the University of Dar es Salaam in 1970. He said:

Having made clear why we are establishing it, and what we expect from it, and having done our best to select administrators and teachers capable of fulfilling our intentions, we have then to trust those we employ and those we select to attend it. We can watch and warn. We can demand that they should explain what they are doing and why—and we can tell them to change if that is necessary! We can instruct the staff to examine themselves and their work every year—to conduct “post-mortems” with the students at the end of every course ... But we should be stupid to try to bind the University staff hand and foot, move them like puppets. The University must be allowed to experiment, to try new courses and new methods. The staff must be encouraged to challenge the students and the society with new arguments, and to put forward new suggestions about how to deal with the problems of building a socialist Tanzania based on human equality and dignity. Further they must be allowed, and indeed expected, to challenge orthodox thinking on scientific and other aspects of knowledge ... Only by allowing this kind of freedom to our University staff will we have a University worth its name in Tanzania ... For University of Dar-es-Salaam will be able to serve out socialist purposes only we accept that those whom we are paying to teach students to think, must themselves be allowed to think and speak their thoughts

Ranger (1981: 17–18).

While Nyerere’s speech sounded accommodative to some degree of academic freedom, his later statements signaled what would later turn to be the political understanding and demarcation of the limits to academic freedom in the universities. During the inauguration of the University of Dar es Salaam in August 1970 Nyerere explained that “the aim of the university of Dar es Salaam must be service to the needs of developing socialist Tanzania. This purpose must determine the subjects taught, the content of the courses, the method of teaching, and the manner in which the university is organized, as well as its relations with the community at large” (Schutte, 1972: 75). Henceforth, Nyerere sought to have firm political control of the university to drive his socialist political vision for Tanzania. The university whose establishment had started from the headquarters of the Tanzania African National Union (TANU) in 1961 (Tanisha, 2007), increasingly experienced entrenched political control of its academic and administrative processes especially after its inauguration as a national university in 1970. The first vice-chancellor of the university, Pius Msekwa, was a former executive secretary of TANU, despite the fact that he lacked the academic experience for such an appointment and in a context where there were qualified academics deserving of such an appointment (Omari, 1991). The confrontation between the academics and students of the university on one hand, and the political elite on the other in the 1970–1980 decade emanated from a desire from academics and students to resist encroachment on their academic freedom by a political elite that wanted to create a subservient university and academic community that followed and obeyed political party and government policies without overt contradictions (Nkulu, 2005).

Nyerere was not alone in these deepening of the state control of universities as a strategy to realize the kind of “development” they dreamed of the new nations. In Ghana, Kwame Nkrumah, bent on using the university to achieve his socialist vision, and suspicious that the University of Ghana would not play along, established a new institute at Winneba to facilitate the control of intellectual life by him as president, and by extension the Convention Peoples’ Party, CPP (Serra and Gerits, 2019). The institute at Winneba housed the Kwame Nkrumah ideological Institute (KNII) with university level status despite the protestations from the National Council for Higher Education advising that the KNII did not fall within either UNESCO’s definition of higher education or the Ghana Education Act and that it therefore should be ranked as a technical school (Serra and Gerits, 2019: 413).

It was after Nkrumah had been overthrown as president that academics and university administrators had the courage to come out and lament how politics had been allowed to override the core tenets of academic freedom at the universities. The excerpts below from the address by the vice-chancellor, of the University of Ghana, Dr. Alexander Kwapong, to the Congregation of the

University of Ghana on March 26, 1966, does show the extent the political establishment had defiled and limited the quality of academic freedom at the university. He said:

... this institution has been subjected to the most merciless and persistent attacks by the regime of Kwame Nkrumah ... In this university, with its substantial body of scholars recruited from all over the world and dedicated to the pursuit of truth, scholarship and research, unfettered by any foreign ideology or creed and enjoying academic freedom, Kwame Nkrumah and his henchmen, it was quite obvious, saw one of the most powerful obstacles to the realization of his despotic ambitions ... It was said by the critics of this university that academic freedom was a luxury which Ghana in its bid for rapid development could ill afford ... a shibboleth behind which imperialist, colonialist and neo-colonialist agents and their stooges sheltered, and in the rapid march towards the new socialist paradise and African unity, there was no place for this outworn bourgeois concept ... You may recall that the Institute of Education was arbitrarily removed from this university by the government without any substantial consultation of the university, against the responsible and reasoned objections of the academic body of this university, subsequently vindicated by events ... Not that enough damage had not been done by this time ... In 1964, this university was shaken by the deportation of six of its senior expatriate members (one of whom was subsequently invited back by Nkrumah to be a presidential professor!) and by the detention of a Ghanaian senior member and of a student and the invasion of the campus by a horde of C.P.P hooligans. The attempt, however, to turn this place into a C.P.P university and an instrument of Nkrumah's will was to have reached its climax this academic year when, with a dedicated party member and intellectual confidant of the ex-President installed as pro-vice-chancellor, a determined effort was to have been made to storm, once and for all, the Legon Bastille.

Excerpts from speech of the vice-chancellor, University of Ghana, 26th March 1966.

The nature and status of academic freedom in the universities could not however be simplistically dichotomized as one that pitied the political class on one side and the academics and students on the other. Nor were the interests at play nuanced and categorized along compacted lines, between those that were exercised by the politicians and representing the nationalist and popular interests of the African people and what was portrayed as "ivory-tower" interests of academics articulated in the language of academic freedom. Often the interests of politicians and the nascent African academics converged, especially among a group of African academics that sought political patronage to seek promotion and take over academic and administrative leadership positions from expatriates. Politicians, and especially the country presidents on the other hand, used their positions as Chancellors of the new universities to manipulate expatriate academics to support positions that eroded the quality of academic freedom especially of African academics that were considered recalcitrant. The appointment of an expatriate academic to the position of presidential professor by Nkrumah, as quoted above was an example of such a manipulation. The intimate manipulated relationship that emerged between expatriate academics who sought security of their contracts and African political regimes on one hand, and on the other nascent African academics posturing as nationalist intellectuals in exchange for political patronage as a strategy for academic promotions and appointment to university administrative positions set in motion a process of progressive and substantial erosion of university autonomy (Coleman, 1977). While security of tenure and deserved academic promotions are considered essential to academic freedom, political regulation and interpretation of what security of tenure and academic qualifications entailed started to be increasingly applied by co-opted university administrators to force compliance by expatriate staff and their African accomplices to curve in to the governments designs of the university (Crawford et al., 1967). Evidence suggests that a number of African academics supported the designs by the state, accepting limited conditions for academic freedom and institutional autonomy in preference for greater state involvement in institutional governance when this suited their immediate objectives (Mamdani, 1993).

The period 1960–1980, witnessed the nationalization of universities and the borrowing of a template from the colonial universities by independent governments to progressively deepen the erosion of academic freedom and institutional autonomy. The increased steering role of independent African governments, which often constituted what has been captured as the threats to academic freedom and institutional autonomy started as a genuine desire by the political elite to drive their vision of national independence and African development through the universities and the creation of an African intellectual elite to support this process. Certainly, the understanding of the political class during the 1970–1980 decade was that political infringement on academic freedom was the necessary political direction that universities needed as part of national planning, and which was considered primary, while the degrees of academic freedom that the universities required remained a matter of discretion.

1980–2000: corporatization of universities and undermining of conditions for academic freedom

Corporatization in regard to higher education in Africa has been loosely used to refer to the period particularly after 1990 when governments' direct involvement in the governance of universities begun to recede and was replaced with new bodies and business management logics and practices. Corporatization begun in the mid-1980s when the World Bank, in response to the economic crisis of the African state urged governments to implement economic structural adjustment interventions to improve the quality of public services, including universities through the deployment of new management practices, including commodifying and commercializing the spaces as a strategy to improve efficiency and effectiveness. The logic of the "developmental university" as established in 1970 was replaced with institutions that were referred variously as "corporate", "entrepreneurial" or "Transformational" universities (Rhodes, 1994; Pryor et al., 2014). The universities, with new governance bodies, were encouraged to explore alternative sources of funding their operations as governments gradually reduced public resources into the universities. These new arrangements, as will be discussed in the last section of this article, also redefined conceptualizations of academic freedom and institutional autonomy, at least from the core constituencies-academics, students and the management of higher education institutions.

The 1980–2000 decades when the governments began implementing the Structural Adjustment Programs witnessed extreme violations of academic freedom throughout the continent, compared to the previous decade. The violations peaked in reality from 1990 as the institutions commenced domestication of the structural conditions provoking resistance from academics and students. In 1991, a Report on academic freedom in Africa by *African Watch* (1991), documented about 158 extreme violations of academic freedom targeting academics and students, including the detention without trial of students and academics across 16 African countries. The nature of violations included arbitrary abrogation of the rights of academics and students to assemble and articulate grievances against university authorities or the government, withdrawal of travel documents from academics, withdrawal of citizenship and resident permits from foreign academics, violent breakup of peaceful demonstrations resulting in deaths of academic staff and students, monitoring of teaching spaces and denial of research permits and a gross amendment to the Acts governing universities without due consideration to the views of students and staff. In Nigeria alone, about 2000 academics from 7 universities were sacked by 1999, accused of teaching what they were not paid to teach (*Committee for Academic Freedom in Africa CAFA*, 2002). Between 1986 and 2001, about 139 students and 6 lecturers from various universities had been killed from police brutality and about 400 students remained expelled from their studies or rusticated because of participating in various acts of resistance against the government's undermining of university education (*CAFA*, 2002). The Nigerian situation largely mirrored what was happening elsewhere in Africa. Between 1999 and 2001, acts of student struggles against structural adjustment, often resulting in the killing of students, closure of institutions, abolition of student's organizations, arrest and detention without trial of students' leaders and the academics who supported them, were recorded in 26 universities across the continent (*CAFA*, 2003). Cumulatively, these actions greatly exposed the independence of the Universities to government interference and allowed vice-chancellors as government appointees to enforce the will of the governments within the institutions with little regard to the intellectual rights of students and staff. Academic staff unions, where they existed and operated in countries like Kenya and Nigeria were banned as a way of dealing with what was considered as dissent from mass organizations that had the capacity to mobilize opposition to the World Bank reforms (*African Watch*, 1991).

Why did governments clampdown on universities and academic freedom in Africa intensify and turn more violent in the 1980–2000 decades compared to the previous 1970–1980 decade? The explanations lie in the structural adjustment economic reforms that governments in Africa were forced to implement by the World Bank and other donor countries. The reforms offered double edged swords both to the governments and universities. Amid declining resources to support universities at a time demands for access to universities peaked all-over the continent, governments sought to force unpopular reforms in the universities, including reduction of the annual intake of students to the institutions. This was given the business-language of rationalization. The resistance to the reforms from better organized urban elites that included university academics and students, also included pressures for liberalization of politics in ways that threatened the one-party states in Africa and the incumbent regimes. Hence, while implementing the reforms assured cash-strapped governments of credit lines from the World Bank and other donors in the short-term, implementation of the reforms in the medium to the long-term created a situation where the governments authored the path of their exit from power. Such double-edged policies required, and indeed did manifest a higher degree of authoritarianism in their enforcement especially toward organized groups that fronted the greatest resistance to their implementation (*Ibhawoh*, 1999). Academics were seen as constituting political resistance beyond the call of academic freedom and institutional autonomy, and university spaces were seen as convenient for covert political organizing in an environment where political organizing in public spaces was highly restricted and monitored. Governments had to deploy violence and gross violation of human rights to implement the World Bank conditions and meet conditions for the promised loans while at the same time resisting pressures for political liberalization.

The suspicion of African governments that academics were engaging in politics aimed at what came to be referred as "regime change" and not raising issues of academic freedom was heightened by the support that academic freedom organizing in the continent began to receive from external neo-liberal academic organizations. These included, for example, the Committee for Academic Freedom in Africa (CAFA) established in 1980, but activated in 1990 to respond to the crisis occasioned by SAPs in African universities; The Association of Concerned African Scholars; the American Association for the Advancement of Science's Committee on Scientific Freedom and Responsibility; Human Rights Watch and African Watch. Jointly or individually, the organizations supported the quest for academic freedom in Africa, including financial support for establishing academic staff unions, highlighting the adverse effects of SAPs to higher education in the continent and exposing the extent of state violence on academics and students who were resisting the state's misadventures in higher education (*African Watch*, 1991; *Federici and Caffentzis*, 1991).

On the positive side, and despite the extreme academic and human rights violations, the 1990s witnessed the realization of an attempt for a decolonized articulation of what constituted academic freedom and institutional autonomy in African universities. In the face of rising abuses of academic freedom across the continent, two documents were produced that served as a guide to the struggle for better spaces for academic freedom as well as a nuanced understanding of the responsibilities and expansiveness of the rights and responsibilities of academics in Africa. The first was the Dar es Salaam Declaration on Academic Freedom and the Social Responsibility of Academics, which was adopted on April 19, 1990, when 12 delegates from autonomous staff associations from higher education institutions in Tanzania convened in Dar es Salaam. The second was the Kampala Declaration on Intellectual Freedom and Social Responsibility that resulted from a Symposium on Academic Freedom and Social Responsibility of Intellectuals convened by the Council for the Development of Social Science Research in Africa (CODESRIA) in Kampala, Uganda on November 29, 1990. The two declarations, while addressing the rights and freedoms of academics to teach and research, also point out to the responsibilities and duties that should accompany such freedoms and for which academics owe to society at large. In articulating

this position, the declarations emphasize the point that universities are not only sites for knowledge production but also of social commitment (Chachage, 2006), within which academics have a duty to contribute toward redressing historical and contemporary inequalities in society based on differences of class, beliefs, gender, race, nationality, region and economic conditions' (Dar es Salaam Declaration, 1990). The declarations do caution academics against hiding behind the veil of academic freedom to engage in unethical behavior and urge the basis of social responsibility to be justified by a commitment to objective science and the promotion of the compact between academic work and society. Article 19 of the Kampala Declaration notes that ... members of the intellectual community are obliged to discharge their roles and functions with competence, integrity and to the best of their abilities. They should perform their duties in accordance with ethical and highest scientific standards. As will be discussed in the last section of the article, this call to ethical and social commitment for academics in Africa means contextualizing the exercise of academic freedom within the context of African struggles for intellectual colonization and subverting the hierarchies of region, ethnicity, gender and race (Badat, 2009).

The expansive conceptualization of academic freedom for Africa as articulated by the two declarations has become untenable to realize. As shall be discussed in the last section of these article, the neo-liberal conditions implemented in African universities have resulted to the emergence of university spaces where the social commitment of academics is compromised by practices of managerialism, commitment to profits and individual material well-being. Rather than promoting inclusiveness and equity, universities have become centers for the reproduction of all sorts of social inequalities, with the full acquiescence of academics.

Embracing support from Western Neo-liberal institutions and supporting part of the World Bank sanctioned conditions in the higher education sector also had its underside to the evolution of genuine spaces for academic freedom in African universities. The reforms proposed by the World Bank and backed by western neo-liberal thinking including introduction of user fees in the financing of universities, reduction of unit costs, and shrinking of public sector participation in favor of the private sector (World Bank, 1988) contributed to eroding gains that had been realized in the development of the university sector in the 1970–1980 decade. Underfunding of the institutions through reduced public support meant declining terms of service, and deterioration of working conditions in ways that legalization of staff and students unions would not help alleviate. Hence, and as highlighted in the introduction to this chapter, though staff and student unions were legalized and the state's direct involvement in the governance of the institutions reduced, the conditions for academic work deteriorated.

As literature has documented, the World Bank reforms produced contradictory outcomes that by the end of the 1990s had negatively affected the institutions as academic institutions. On one hand, universities privatized some sections, including student streams which increased university revenues and staff remunerations. On the other hand, students from poor families were excluded from the institutions unless on dwindling government subventions. This fractured the traditional solidarity between students and their lecturers and among lecturers themselves as priority was put on market forces that considered education as a commodity offered to individual consumers who could afford it, thus moving the bases of solidarity from academic collegiality to business-like customer-client relations. There was also an increasing view of academics as wage-laborers whose wages depended on the financial status of the institutions and were thus obliged to compromise their freedom for a salary and research funding, leading to the general deterioration in the quality of teaching and learning in the institutions (Federici et al., 2000; Sawyer, 2004; Oanda, 2011; Ochwa-Echel, 2013; Yann and Oanda, 2020).

Post-2000 developments: intensity and quality of academic freedom in African universities

As discussed in the first section of this article, the restraints to the exercise of academic freedom have largely been eliminated in most of the countries. Academic freedom is recognized as a right in the constitutions of most countries, student and staff organizations have been allowed to operate, and in some countries, they are included as part of the governance structure of the institutions. Certainly, some of the old aspects of restraint to academic freedom still occur in a number of countries. In Cameroon, for example, academics have been entangled in the political tension between Anglophone and Francophone regions to an extent that academics' liberties to undertake research, teach or communicate ideas or facts that are inconvenient to the political establishment lose their jobs and are sometimes detained (Ngoh, 2021). On April 20, 2019, the minister of higher education of Cameroon caused authorities of the University of Buea to suspend classes of Agbor-Balla, a law lecture alleging the lecturer had breached the university's code of ethics and conduct (Human Rights Watch, 2019). According to the lecturer, the reason for the cancellation was because of an assignment he had given students requiring the students to explore the reasons behind the Anglophone crisis, to gauge how critical the students would engage with the assignment (Human Rights Watch, 2019). In Egypt, the military coup in 2013 did trigger a situation where universities became militarized as the military stifled any dissent from the universities (Saliba and Grimm, 2016). The Association for Freedom of Thought and Expression (AFTE) in Egypt documented 2138 violations of students' rights between 2013 and 2016, including 21 extrajudicial killings and 1181 student arrests (Saliba and Grimm, 2016). The Arab Spring also did trigger widespread political control over and censorship in higher education and research throughout North Africa's universities (Saliba and Grimm, 2016). At Makerere University, Uganda, 45 members of the academic staff were dismissed in December of 2019 over what university authorities said was an enforcement of discipline, but which the academics thought was a case of university authorities being used to silence academic thought to be opposed to the

administration of the country's president (<https://www.dw.com/en/ugandas-makerere-university-at-center-of-row-over-academic-freedom/a-47310962>). In Rwanda, the government continues to restrict academic freedom through the enforcement of particular views on the genocide and questions of political rights forcing academics and students to engage in self-censorship to avoid punitive reactions from the government (Freedom House, 2021). On August 14, 2020, the University of Rwanda (UR) dismissed professor Aimable Karasira in apparent retaliation for public expressions critical of the government (Scholars at Risk, 2020).

Despite the continued existence of the old external restrictions to academic freedom, the most worrisome abuses are coming from within the universities and involve its core constituencies, academics, students and management of the institutions. One adverse consequence of the World Bank conditions implemented in the universities was that by the end of the 1990s, reduced government investments in the universities had resulted in the collapse of the basic infrastructure for academic work making a mockery of the space for academic freedom that had been extended to the academics. Academic freedom and academic responsibilities are rights that can only be exercised in a context where conditions for good academic work and social engagement exist. The logic of the new governance frameworks that emerged in the institutions was however undermined by the realization of these conditions.

The new higher education governance frameworks that replaced governments had to design market like strategies for generating institutional revenues in ways that brought new challenges to the exercise of academic freedom. The reforms were driven by financial considerations and the logics of the market which reduced public funding to the institutions and undermined public accountability of higher education systems, and the economic prerequisites for academic freedom and social responsibility (Mama, 2006). Henceforth, issues of academic freedom begun to be articulated at the level of economic interests of both students and academics, and how the sharing of dividends from commercialized university activities would be shared between the institutions and academics and the management of institutions. Tensions between academics and university management have been limited to strikes over low salaries or non-payment of the same, or disagreement of how internal revenues generated by the universities have to be shared between the institutions and academics within the understanding of the market principles of profit sharing and the relationship between clients and customers.

Lack of investments in renewing teaching-learning infrastructures has also led to a neglect of an important component of academic life-the quality of students' services in the institutions. Implementation of the World bank reforms was supposed to lead to better learning environments. However, with university planning heavily borrowing from business models, little resources were spent to set up new facilities and improve existing ones. This is despite the fact that most institutions in countries that had introduced tuition fees resorted to increasing the intake of students as a strategy to generate more revenues for the institutions. By 2002, CAFA documented the growth of cults among students across Nigerian Universities shrouded in religious practices that resulted in the death of students and abandonment of academic life. Studies have documented the poor conditions of student facilities including overcrowding of, limited skills and capabilities of facility management departments, and lack of financial resourcing for better facilities across public universities in Africa to the extent that quality of the learning and teaching environment at the institutions has collapsed (Chikafalimani et al., 2021). The post-2000 period has witnessed the total triumph of the market logic, individualization of student grievances, fractured the collegiality between students and academics, as academics have opted for calm to make extra income from "unyielding clients (students), and the overall criminalization of any student grievances by university authorities" (Chikafalimani et al., 2021).

On the other hand, the frequent strikes by academics over salaries and closure of institutions also continue to accentuate the suffering and financial demands on students and families. Frequent university closures lengthen the time to graduation for students. There has also been an increase on the number of unethical behavior on the part of academics reported across the continent, including academics engaged in "degree mills" as a source of alternative income, selling grades for sex and other forms of unethical behaviour that cannot be defended by invoking the safeguards to academic freedom, although academic staff unions have invoked such rights for their members (Hallak and Poisson, 2007; Kirya, 2021). Often in resorting to strike action in furtherance of their quest for better salaries and terms of service, academics have ended trampling on the very right to academic freedom as the right of academics to strike action undermines the right of students to learn (Mamdani, 2019). Universities, operating on market logics may have also resorted to employing to full-time positions junior academics with only masters' degree qualifications, a category that should be on staff training and mentorship. Employing a large number of such academics and giving them full-time academic responsibilities may have compromised the standards and quality of academic freedom as such academics have not matured to appreciate the full burden of what academic freedom and social responsibility entail. By 2020, it was estimated that the proportion of academic staff having a Ph.D. in African universities rarely exceeded 40% and, depending on the country and the subject, the ratio would be as lower as 20% (Mohamedbhai, 2020).

How would the persisting and in some instances the deterioration of academic freedom and social responsibility of academics be explained in the context where most of the limitations imposed by governments have been removed and the right to academic freedom enshrined in most constitutions? In the first instance, it should be noted that governments vacated their governance roles together with the public funding responsibilities that secures the conditions for the exercise of academic freedom. Granting academics, the constitutional right to academic freedom, but abdicating the responsibilities for securing the conditions for the exercise of such freedoms to the vagaries of the market amounted to what Isaiah Berlin (Berlin, 1969) has conceptualized as the negative dimension of academic freedom. In this respect, constraints to the exercise of academic freedom such censorship, non-interference in teaching and research, the freedom of academics to take part in public discourses and the freedom to form and belong to

academic unions are legally guaranteed in most African universities. But the institutional mechanisms or enablers that facilitate academics to exercise the choices were hollowed out by the reforms governments implemented and the abdication of their realization to market forces. Academics would for example access research funds if only they tweaked their research engagements to address certain market needs outside considerations of objective knowledge and needs of society. The deepening of the neoliberal structures in the institutions and the reconfiguration of internal institutional organs of governance has resulted to situations where the governance organs have become antithetical to accommodation of academic freedom and institutional autonomy as traditionally understood. There is also a continuing surrender by academics of their intellectual authority to market forces, on issues of deciding what to teach and surrendering this to be determined by bodies outside the universities. The recourse to engaging a high number of academics who have not fully attained academic ranks has accentuated trends toward widespread unethical behavior and subversion of the conditions for academic freedom from within the institutions. The full exercise of academic freedom requires not only the absence of certain social relations of prevention but the presence of abilities and capabilities (Sen, 1993).

Conclusion

The excessive interference of governments or government functionaries in the academic affairs of universities is often cited in the literature as an example of what went wrong with universities in Africa from the 1990s. This article, by analyzing some historical literature has shown that government control of the institutions was an aspect that was built into the administrative logic of the institutions at a time they were established as university colleges during the colonial period. In the case of British colonial Africa, the secretary of state through the governors of the various colonies exercised this control. In 1970, when the universities were transformed into national institutions, the presidents of various African countries took over the levers of control and sought to use the institutions as vehicles for decolonization narrowly defined. The last in this approach is South Africa which post-1994 has tried to exert government control over the universities to achieve “transformation”. The worst excesses in academic freedom abuses in Africa were however witnessed from 1990 as African governments were arm-twisted to enforce the will of the World Bank in restructuring the universities. From 2000 to-date, literature does suggest that market conditions, pliant university management and an academic staff who have acquiesced to the logic of the market have largely contributed to hollow out requisite conditions for the exercise of academic freedom in African universities.

References

- African Watch, 1991. Academic Freedom and Human Rights Abuses in Africa. Human Rights Watch, New York/Washington DC.
- Altbach, P.G., 2007. Academic Freedom in a Global Context: 21st Century Challenges. The NEA 2007 Almanac 2 of Higher Education.
- Ashby, E., Anderson, M., 1966. Autonomy and academic freedom in Britain and in English-speaking countries of tropical Africa. *Minerva* 4 (3), 317–364. Springer.
- Badat, S., 2009. The role of higher education in society: valuing higher education. In: Paper Presented at the HER-SA Academy, University of Cape Town Graduate School of Business. https://www.ru.ac.za/vicechancellor/speechespresentations/name_58754_en.html.
- Berlin, I., 1969. Two concepts of liberty. In: Berlin, I. (Ed.), *Four Essays on Liberty*, new ed. in Berlin 2002. Oxford University Press, London.
- Chachage, S.L.C., 2006. The university as a site of knowledge: the role of basic research. *J. High Educ. Afr.* 4 (2), 47–67.
- Chachage, S.L.C., 2009. Academic Freedom and the Social Responsibilities of Academics in Tanzania. CODESRIA, Dakar.
- Chikafalimani, S.H.P., Kibwami, N., Moyo, S., 2021. Management of facilities at public universities in Africa: current challenges and the way forward. *Real Estate Manag. Valuat.* 29 (1), 21–29.
- Coleman, J., 1977. The academic freedom and responsibilities of foreign scholars in African universities. *Afr. Iss.* 7 (2), 14–32. <https://doi.org/10.1017/S1548450500000834>.
- Committee for Academic Freedom in Africa, CAFA, 1991. Academic Freedom in Africa Under Siege. Newsletter No 1. Spring.
- Committee for Academic Freedom in Africa, CAFA, 2002. Emigration, Campus Cults and Linguistic Imperialism: The Consequences of Structural Adjustment in African Universities. Newsletter, No 17.
- Committee for Academic Freedom in Africa, CAFA, 2003. Accountability and Myth. Newsletter, No 18.
- Council for the Development of Social Science Research in Africa, CODESRIA, 1990. The Kampala Declaration on Intellectual Freedom and Social Responsibility. Kampala, Uganda.
- Crawford, M., Mazrui, A.A., Tandon, Y., 1967. Government and the Universities in East Africa. *Minerva* 5 (3(Spring)), 376–386.
- Dar es Salaam Declaration on Academic Freedom, 1990. Declaration on Academic Freedom and Social Responsibility of Academics. <http://www1.umn.edu/humanrts/africa/DARDOK.htm>.
- Dlamini, C.R.M., 2002. University autonomy and academic freedom in Africa: ex Africa semper aliquid novi? *Comp. Int. Law J. South Afr.* 35 (1), 77–98.
- El Barometer, 2007. El Barometer of Human and Trade Union Rights. Education International, Brussels.
- Federici, S., Caffentzis, G., 1991. Academic Freedom in Africa under Siege. *Committee for Academic Freedom in Africa Newsletter*, NI. Spring.
- Federici, S., Caffentzis, G., Aidou, O. (Eds.), 2000. *A Thousand Flowers—Social Struggles Against Structural Adjustment in African Universities*. Africa World Press, Trenton.
- Freedom House, 2021. Rwanda. <https://freedomhouse.org/country/rwanda/freedom-world/2021>.
- Fulda, A., Heathershaw, J., Chubb, A., 2021. Involve Faculty in International Higher Education Governance From Start. University World News. <https://www.universityworldnews.com/fullsearch.php?mode=search&writer=John+Heathershaw>.
- Grimm, J., Saliba, I., 2017. Free research in fearful times. *Interdiscip. Polit. Stud.* 3 (1), 41–75. <https://doi.org/10.1285/i20398573v3n1p41>.
- Hallak, Poisson, 2007. *Corrupt School, Corrupt Universities: What Can Be Done?* UNESCO Press.
- Hargreaves, J.D., 1973. The idea of a colonial university. *Afr. Aff.* 72 (286), 26–36.
- Heron, D.B., 1964. Some reflections on academic freedom. *Makerere J.* (9).
- Human Rights Watch, 2019. No Room for Debate in Cameroon Classrooms. <https://www.hrw.org/news/2020/05/01/no-room-debate-cameroon-classrooms>.
- Hyden, G., 1967. The failure of Africa's first intellectuals. *Transition* (28), 14–18.
- Ibhawoh, B., 1999. Structural adjustment, authoritarianism and human rights in Africa comparative studies of South Asia. *Afr. Middle East* XIX (1).
- Kinzelbach, K., Saliba, I., Spannagel, J., Quinn, R., 2021. Free Universities; Putting the Academic Freedom Index Into Action. Global Public Policy Institute/Scholars at Risk Report.

- Kirya, M., 2021. Curbing Corruption in Higher Education. Sector Reform Experience and Strategies. CurbingCorruption.com. <https://curbingcorruption.com/wp-content/uploads/2021/06/210618-Curbing-Corruption-in-Higher-Education.pdf>.
- Kori, E., 2016. Challenges to academic freedom and institutional autonomy in South African universities. *Int. J. Teach. Educ.* IV (1), 45–53. <https://doi.org/10.52950/TE.2016.4.1.004>.
- Kwadwo, A.-A., Klaus, D.B., Terence, K., 2016. A review of academic freedom in African universities through the prism of the 1997 ILO/UNESCO recommendation. *J. Acad. Freedom* 7.
- Kwapong, A., 1966. Address by the vice-chancellor to the congregation of the University of Ghana, 26 March 1966. *Minerva* 4 (4), 542–554.
- Mama, A., 2006. Towards academic freedom for Africa in the 21st century. *J. High Educ. Afr.* 4 (3), 1–32.
- Mamdani, M., 1993. University crisis and reform: a reflection on the African experience. *Rev. Afr. Polit. Econ.* (58), 7–19.
- Mamdani, M., 2019. Has MUASA Become a Political Party? An Insider's View. *Monitor*. <https://www.monitor.co.ug/uganda/oped/commentary/has-muasa-become-a-political-party-an-insider-s-view-part-2-1808388>.
- Mkandawire, T., 2002. African intellectuals, political culture and development. *Austrian J. Dev. Stud.* XVIII, 31–48.
- Mnguni, H.M., 1987. African intellectuals and the development of African political thought in the twentieth century. *Presence Afr.* 3 (143), 113–121.
- Mohamedbhai, G., 2020. Quality Assurance of Doctoral Education is Now Urgent. *University World News*. <https://www.universityworldnews.com/post.php?story=20200609091837168>.
- Munene, I.I., 1997. The struggle for faculty unionism in a stalled democracy: lessons from Kenya's public universities. *J. Third World Stud.* 4 (1), 91–114.
- Ngoh, N.N., 2021. The status of academic freedom in Cameroon. In: Ceglie, R., Thompson, S. (Eds.), *Academic Freedom: Autonomy, Challenges and Conformation*. Emerald Publishing Limited, Bingley, pp. 111–123.
- Nkulu, K.L., 2005. Serving the Common Good: A Postcolonial African Perspective on Higher Education. Peter Lang Publishing, Inc, New York, pp. 52–53.
- Oanda, I.O., 2011. Neo-liberalism and the subversion of academic freedom from within: money, corporate cultures and “captured intellectuals in African public universities”. *J. High Educ. Afr.* 1 & 2.
- Ochwa-Echel, J.R., 2013. Neoliberalism and University Education in Sub-Saharan Africa. *SAGE Open*, pp. 1–8. July–September.
- Oloka-Onyango, J., 1992. The legal control of tertiary institutions in East Africa: the case of Makerere University. *Afr. Dev.* 17 (4), 47–66.
- Omari, I.M., 1991. Innovation and change in higher education in developing countries: experience from Tanzania. *Comp. Educ.* 27 (2), 186.
- Pryor, M.G., Odom, R.Y., Toombs, L.A., 2014. Organizational implosion; a threat to long-term viability. *Acad. Strat. Manag. J.* 13 (2).
- Ranger, T.O., 1981. Towards a Radical Practice of Academic Freedom: The Experience of East and Central Africa. University of Cape Town, Cape Town.
- Report of the Commission of Higher Education in the Colonies, Asquith Report, 1945. Presented by the Secretary of State for the Colonies to Parliament by Command of His Majesty. His Majesty's Stationery Office, London.
- Rhodes, R.A.W., 1994. The hollowing out of the state: the changing nature of the public service in Britain. *Polit. Q.* 65, 138–151.
- Saliba, I., Grimm, J., 2016. Deteriorating conditions for academic research in MENA. *APSA MENA Newsl.* 1 (1), 10–12.
- Sawyer, A., 2004. Challenges facing African universities: selected issues. *Afr. Stud. Rev.* 47 (1), 1–59.
- Scholars at Risk, 2020. Rwanda. <https://www.scholarsatrisk.org/regions/eastern-africa/rwanda/>.
- Schutte, D.G.W., 1972. The University of Dar es Salaam: a socialist enterprise. In: Niblett, W.R., Freeman Butts, R., Holmes, B. (Eds.), *World Yearbook of Education 1972/3. Universities Facing the Future*. Routledge, Oxon, pp. 75–96.
- Sen, A., 1993. Capability and well-being. In: Nussbaum, M., Sen, A. (Eds.), *The Quality of Life*. Clarendon Press, Oxford.
- Serra, G., Gerits, F., 2019. The politics of socialist education in Ghana. The Kwame Nkrumah ideological institute. *J. Afr. Hist.* 60 (3), 407–428.
- Sifuna, D.N., 1998. The governance of Kenyan public universities. *Res. Post Compuls. Educ.* 3 (2), 175–212. <https://doi.org/10.1080/13596749800200030>.
- Tanisha, D., 2007. The History of University of Dar es Salaam: Chronicling the Importance of the Student Voice. Unpublished Thesis. Justice and Peace Certificate Thesis Project, Dar es salaam.
- The Reporter, 1963. *East African*, p. 29.
- Uganda's Makerere University at Centre of Row Over Academic Freedom, 2019. <https://www.dw.com/en/ugandas-makerere-university-at-center-of-row-over-academic-freedom/a-47310962>.
- UNESCO, 1998. Recommendation Concerning the Status of Higher-Education Teaching Personnel. Adopted by the General Conference at Its Twenty-Ninth Session, Paris, 21 October–12 November 1997. UNESCO, Paris.
- World Bank, 1988. *Education in Sub-Saharan Africa*. Washington, D.C.
- Yann, L., Oanda, I.O., 2020. Higher Education Expansion and Social Inequalities in Sub-Saharan Africa: Conceptual and Empirical Perspectives. United Nations Research Institute for Social Development (UNRISD Working Paper 2020-10).
- Zezeza, P.T., 2021. Challenges and Opportunities for African Universities in a Post-COVID-19 World. The Elephant. <https://www.theelephant.info/long-reads/2021/06/07/challenges-and-opportunities-for-african-universities-in-a-post-covid-19-world/?print=pdf>.

The changing functions of libraries in higher education

Sarah R. Fitzgerald and Jennifer M. Eustis, University of Massachusetts Amherst, Amherst, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	124
Organization of libraries	124
What are public services?	124
What are technical services?	125
Library personnel	125
Library collections	126
Economics of information	127
Library usage	128
Movement toward networked systems	128
Linked data	128
Open source	129
Library services platforms	129
Diversity and inclusion	130
References	130

Introduction

Academic libraries, like higher education institutions, have multiple missions. They preserve information, provide that information to members of the institutions they belong to, and help people find and use that information. They serve both the highly specialized research needs of advanced scholars and the large volume of routine learning needs of undergraduate students. They also provide study spaces for students and access to hardware and software for learning and scholarship. While the missions continue, the mechanisms for executing the library's missions have expanded in the digital age. Preservation involves not just caring for aging paper and maintaining technology for using various audio-visual formats, but storing digital copies of information resources in various file types. Access entails both shelf maintenance and online resource architecture management. Instruction can be both face-to-face and virtual in the form of websites, video, and online chat.

Organization of libraries

Library Deans or Directors often report to the university provost. Where the library is seen as a peer to the teaching colleges, the title is often Dean, whereas libraries viewed as support services often have Directors (Hoffman, 2017). A Dean may be supported by Associate Deans and department chairs and a Director may be supported by department heads. The functions of academic libraries have traditionally been divided into the public services division and the technical services division.

What are public services?

The public services divisions of libraries are the personnel who interact with patrons. This includes instructional librarians, circulation staff, and staff to aid with technology such as printers and audiovisual equipment. Public service librarians serve patrons, teach students about information use, work with faculty members seeking information, and in large libraries, they often have specific subject specialties. They used to have large collection development responsibilities, but the rise of journal subscription packages and databases has moved many collection development decisions away from individual subject experts and toward centralized acquisitions librarians (Uziel, 2017). Instead, public services librarians spend more time teaching and reaching out to departments to keep them apprised of new resources. At large libraries, librarians no longer spend much time on reference desks, although they still answer questions referred to them by support staff or which come in virtually through email or chat. There is a trend toward consolidating reference and circulation desks so that simple queries and technology issues can be handled by circulation staff or student workers and librarians are called in only for complex questions (Forbes and Keeran, 2017).

There is decreasing demand for circulation of physical materials and for reference services, but increasing demand for library instruction. The area of library services where physical materials and reference support is still in highest demand is in special collections and archives, where many rare or unusual collections reside. Demand for print also continues in many developing nations where internet access is limited. Media production is a growing area of public service for academic libraries. Many academic libraries provide equipment and support for video and sound recording projects. Research data management and analysis are also growing areas of public services.

What are technical services?

Technical services units cover a wide range of activities that serve patrons either directly or indirectly. Broadly speaking, these functions consist of acquiring materials, licensing, managing electronic resources, managing finances, resource sharing, systems, metadata and cataloging, processing, or preservation and conservation. Common technical services unit names in the library include acquisitions, metadata, project management, negotiation, resource management, document delivery and interlibrary loan, or conservation. These core functions serve as the mostly hidden infrastructure that supports and facilitates the overall work of the academic library (Ciota and Knight, 2020).

Technical service functions continue to change as academic libraries evolve. The demand for electronic resources continues to grow and presents ongoing challenges. There are a number of models to acquire these resources through on-demand acquisition, purchase plans, firm orders, or leased materials. Each provider and publisher has a variety of models and ways to access electronic resources. This variety combined with sometimes predatory practices on the part of providers and publishers make for an often confusing user experience (Beall, 2009). In the realm of cataloging, the trend is to speak more of metadata. This highlights technical work with data of all formats that is used to contextualize resources for users. Metadata form the basis from which systems exchange and share information for users and by which users discover, search, find, and identify resources. The major change here is the re-focusing of work away from records and managing documents to metadata that is shared, exchanged, and managed in network environments (Long, 2020).

Library personnel

Library personnel are a diverse group, at least as diverse as faculty. They serve all the disciplines faculty serve, requiring specialist knowledge in every subject. They are not united by common duties of teaching or research as the teaching faculty are. They may belong to different unions, for example, with subject liaison librarians in a faculty union, computer programmers in a professional staff union, and circulation clerks in a clerical staff union. Libraries in North America and Europe increasingly prefer to hire PhDs for specialist librarian positions beyond or instead of a Master of Library Science (Brunner, 2010), and library personnel may have a variety of other degrees in subjects such as computer science, business, or law. Librarians in Asia, Africa, and South America sometimes hold bachelor's degrees rather than master's or doctoral degrees.

Library personnel numbers are decreasing and there is a trend toward hiring more librarians and other professional personnel and fewer clerical staff and student assistants (ARL, 2021) as can be seen Fig. 1. This is due to the increasing needs for library services like electronic resource management and information literacy instruction and the decreasing use of print materials which previously necessitated a lot of physical processing and maintenance. The overall decline in personnel numbers is at least in part a result of the decrease of library funding as a proportion of university spending. Growing librarian roles include data management, scholarly communications, digital humanities, geographic information systems, user experience, and assessment. The library profession is overwhelmingly white and the underrepresentation of black and Hispanic people in librarianship is problematic for serving an increasingly diverse student population (Brunner and Osario, 2017).

Librarians at some institutions have faculty status (Hoffman, 2017). This means they have scholarship and service requirements as well as their professional duties to attend to. Many go through a tenure or equivalent process. This sometimes creates tension at universities because librarians have both the demand for scholarly innovation like teaching faculty and the expectation of

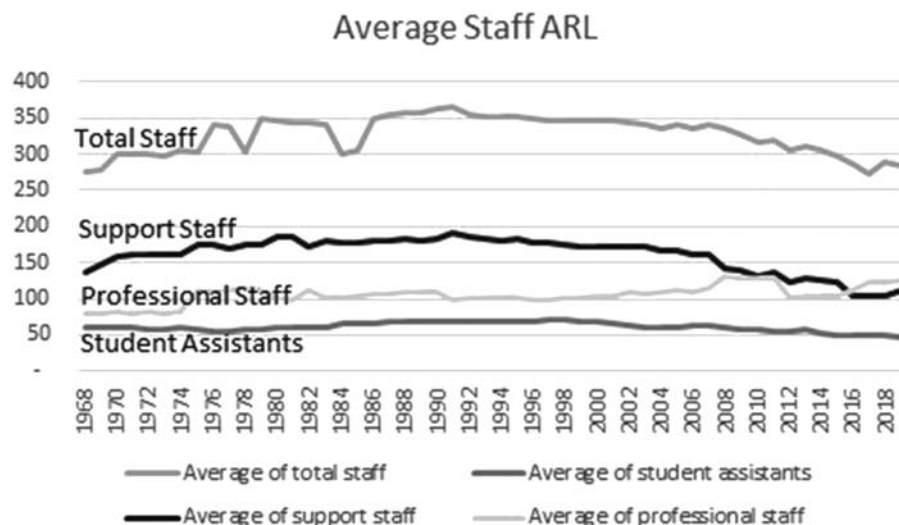


Fig. 1 Changes in average staffing of ARL libraries.

completing their professional duties 12 months a year, unlike teaching faculty. It is easy to overlook the research productivity of librarians since they also have practical responsibilities to support their institutions. The large number of academics searching for tenure positions means some PhDs find roles in the library which answer their career goals and employ their expertise as subject specialists to serve universities beyond the teaching departments.

Library collections

What academic libraries collect changes with what scholars value. As academia shifts its values, library holdings shift to continue to serve their patrons. This means more resources reflecting minority voices, more digital resources, greater emphasis on articles than monographs, and collections which lean toward the subjects best supported by grant funding. It means library collections have expanded to include not just published articles, but conference presentations, data sets, and research software.

Individual institutional mission affects the aims of the library serving it. For example, at research intensive institutions, the demand tends to be for more journal packages and database subscriptions, while at liberal arts institutions the monographs prized by the humanities disciplines tend to be in higher demand (Budd, 2018). At teaching focused institutions, instructional materials are in higher demand than the newest research publications.

There has been a shift from collecting physical materials to providing access to both physical and digital materials, but both of these parts of the library mission remain important. Although much newly produced information is available in digital formats, need remains for information resources produced in physical formats, so libraries must balance multiple format demands. The challenge for libraries moving forward is to balance continued access to historical information resources with access to the current and emerging resources. This is too big a task for any single library, even the largest and most affluent. The way forward for libraries is collaboration and resource sharing. Libraries must share the burdens of housing and preserving resources and streamline the pathways for sharing resources beyond their home institutions.

The graph in Fig. 2 illustrates the proportion of resource usage by type for libraries serving doctoral granting universities in the Association of College and Research Libraries (ACRL, 2021). The usage displayed is an average for a group of over 300 US libraries. Notice that print circulation makes up only a small portion of usage for these libraries. Much more library usage comes from article downloads through library subscriptions and institutional repositories. Ebook usage is also a significant part of library use. In this case, ebook usage is represented as ebook article or section usage, so it is not straightforward to compare with print book checkout. While visiting patrons can use print and digital material while on campus, most circulation, ebook use, and subscription journal use come from patrons affiliated with the library's institution. In contrast, the articles and resources shared in an institutional repository (IR) are readily available to users both inside and outside the institution. In contrast to the traditional model of providing resources produced by scholars worldwide to a local audience, IRs reverse this model by providing resources produced by scholars at the local institution to a worldwide audience.

The globalization enabled by the web is in some ways leading to more localization in the work of libraries. When it is impossible and unnecessary to purchase or house the entire scholarly record, libraries must each concentrate on preserving and curating the scholarly record produced by the experts at their institution and making those resources available globally.

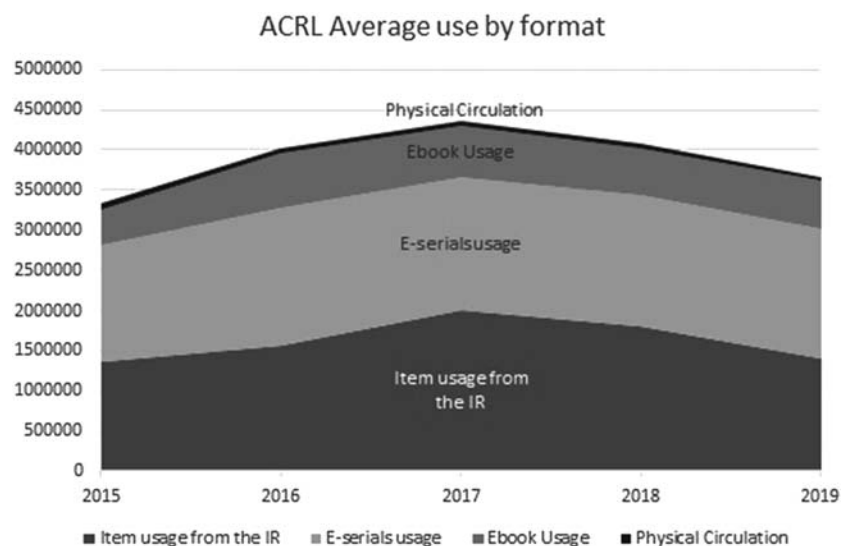


Fig. 2 Average usage by format at ACRL libraries.

Economics of information

Libraries are a form of collective ownership. Collective ownership of information facilitates more economical storage, use, and management of resources than individual ownership of information. Affordability of information resources is particularly important in the current higher education environment because of the growing costs of higher education. With efforts to make higher education accessible to everyone regardless of their economic circumstances, providing affordable access to the information resources required to succeed in higher education is increasingly important. The exponentially growing body of reputable scholarly knowledge is expensive to access. The body of free information available on the web is also growing rapidly, but consists of resources of varying quality. This makes the work of librarians and library personnel vital to higher education. The spectrum of higher education patrons from specialized scholars to beginning college students depends on the services of the library for access to and differentiation between information resources.

The market for scholarly publications is troubled. Libraries face stagnant budgets (Rauf, 2017) and the costs of providing access to scholarly information are increasing. Demand for scholarly information is inelastic. Scholars do not pay for the information they use, so their demand does not decrease with increasing costs. Library and university reputations depend on access to information, so for-profit publishers have few limits on how much they can charge. Studies have shown “the large commercial publishers routinely make a 35% to 45% profit margin on the scientific journal operations” (Lewis, 2016, p. 50). This is due in part to what is known as the “big deals” in which publishers bundle titles together, knowing that the majority of demand is for a minority of journal titles, leading libraries to pay for titles they don’t need to get titles they do need. Publishers argue that cost per title has not increased dramatically, while librarians point out that libraries and their patrons do not want or use most of the titles they pay for. Rising prices from the “Big 5” for-profit academic publishers has forced many libraries to sacrifice titles from independent publishers and limit staffing to reduce salary expenditures. Because much scholarly humanities literature is published by society presses or university presses, while literature in the sciences is often published by commercial publishers, the swing of funding toward the big commercial publishers particularly hurts the humanities.

The graph in Fig. 3 shows the proportion of spending on one time materials such as books, DVDs, and CDs, versus ongoing expenditures such as subscriptions to databases and journals (ACRL, 2021). The graph shows average collection spending per ACRL library at doctoral granting universities. Notice that ongoing expenditures far exceed one time expenditures.

Traditionally, libraries have not provided textbooks to students except on course reserves for a few hours. However, with the rise of online courses and ebooks, libraries are increasingly making efforts to provide digital textbooks for course support. When print publication was the norm, each volume produced and purchased had a cost. In the digital age, it costs the same to produce one copy of a document as it does to produce copies for everyone who wants one. However, in many cases, publishers continue to sell ebooks with digital restrictions to allow only one person or a few people to read them at a time instead of providing for unlimited users from an institution. This causes difficulty in universities where many students in a course are assigned to read the same text at the same time. Some libraries are involved in efforts to create Open Educational Resources which are free for unlimited students to use in their coursework (Owen, 2017).

Scholarly knowledge is a public good whose value goes beyond its use to academics to practitioners and policy makers in a wide variety of fields. To serve the public good, the scholars who produce the information are trying to disrupt the expensive market for scholarly knowledge. Particularly in the pure sciences, such as physics and biology, there are efforts to make access to scholarly research free. In addition to being part of the movement to disrupt the market for information resources for student learning, libraries are also a piece of the open access movement, helping scholars share their research through institutional repositories

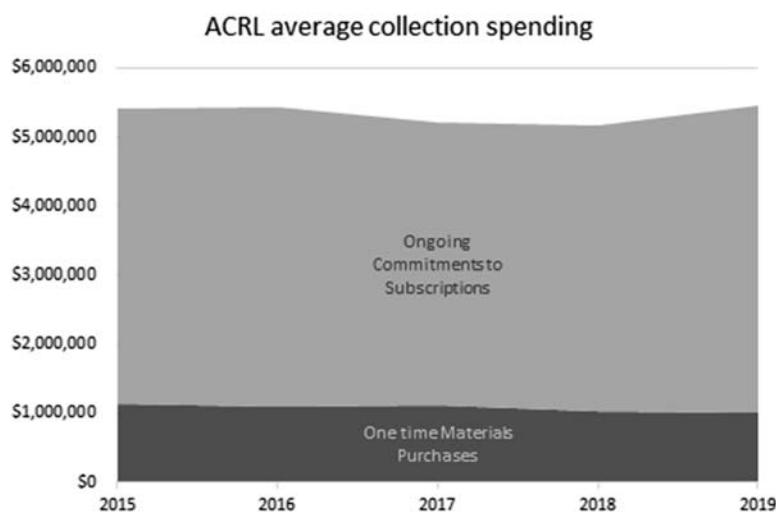


Fig. 3 Average collection spending at ACRL libraries.

(Owen, 2017). Libraries also make the increasing amount of information available freely on the web more discoverable for scholars at their institutions (Faulkner, 2017).

To decrease the cost per use of materials, libraries have adopted more patron driven acquisitions, in which ebook purchases are made when patrons download, print, or read an ebook instead of items being selected and purchased by librarians before they are requested by a patron. Libraries also acquire print books when requests are made by scholars at their institutions, though this process takes more time to complete.

Library usage

Decreasing areas of demand in libraries include print circulation and reference. Increasing areas of demand include instruction, electronic article downloads, and ebook usage. The increased effort to provide instruction to students may be part of the reduction in reference questions. The role of library discovery systems (an interconnected search system that patrons use to search various library resources) as a finding tool is high but decreasing while the use of libraries for access to collectively purchased content is increasing. The graph in Fig. 4 shows the decreasing average searches of electronic library resources and the increasing average of article downloads at libraries that are members of the Association of Research Libraries, a group of over 100 research libraries in North America (ARL, 2021). The decrease in searching using library discovery systems suggests growing reliance on search engines for discovery of resources funded by libraries for their patrons.

The trend toward using search engines to discover resources funded by the library may lead to a move away from “indexing databases” which provide discovery avenues for patrons without providing full text access to resources. As prices for information increase and library budgets are not increased at the same rate, this is one area where libraries may make spending cuts. However, since search engines tend to be much less robust in search functionality than library discovery systems, those cost savings would mean a loss of discovery capabilities.

Movement toward networked systems

Libraries are a somewhat unusual industry in that they are not in competition with one another, but all benefit from improvements at any individual library. Librarians at different institutions often collaborate to contribute to enhancing information availability and access globally. The largest collaborations involve sharing information on resources. This information is referred to as metadata, or descriptions that contextualize resources, allowing users to search, find, identify, and browse library resources. These metadata are shared and rely on large collections such as OCLC’s WorldCat for descriptions on resources that can be reused by libraries. This fundamental collaborative effort has led not only to networked systems but a host of best practices and national standards to facilitate metadata interoperability and exchange and being able to relationally link seemingly disparate metadata.

Linked data

The change from documents to data in cataloging and metadata units reflects the larger changes with the Internet. This is the dream of the “semantic web” where structure is applied to meaningful content in such a way that computer systems can carry out complex tasks (Berners-Lee et al., 2001). These are metadata that are shared, exchanged, and managed in network environments or linked data. The appeal of linked data for academic libraries is to improve the user experience and take advantage of the rich metadata

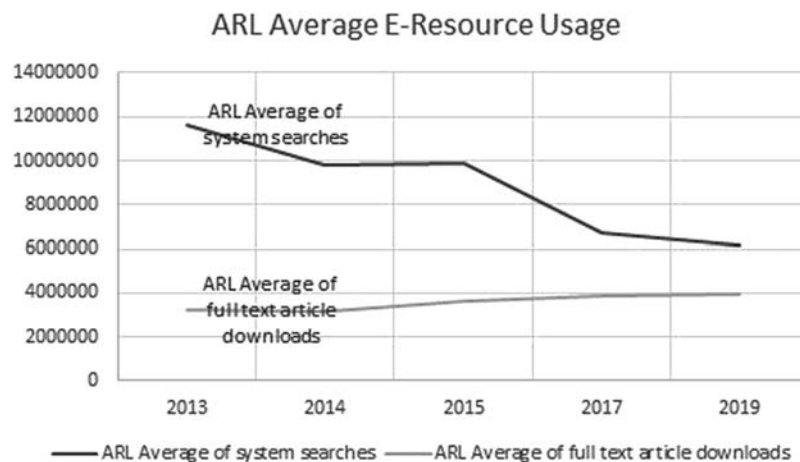


Fig. 4 Change in average e-resource usage at ARL libraries.

available on the Internet. However, the reality of linked data and the semantic web are still a dream. The work to ensure that metadata can be shared, exchanged, and managed between a panoply of different systems is daunting. To accomplish this, metadata must be structured in such a way that heterogeneous systems can process that information. This requires a shared understanding of how data are structured and shared that crosses barriers of language, cultures, and proprietary systems. Fundamentally, this work requires that people come together to realize that dream and ensure that metadata can flow between systems. Further, this work depends on the quality of the underlying metadata (Banerjee, 2020) which requires human resources to ensure that minimal quality. The continual trend of decreasing library staff particularly in technical services makes reaching this dream even more difficult. In short, linked data solves technical issues, not human problems (Banerjee, 2020).

Thanks to the transition to linked data and the need to solve human problems, technical services are often reinventing themselves. This is ever more true in light of budgetary challenges, the need to find efficiencies in workflows, the move toward more electronic resources, and shifts in staff. Efforts to reduce costs and increase collaborative efforts can often mean a chance to reinvent technical services and refresh staff toolkits (Cerbo, 2011). Academic libraries need to work to ensure that personnel reductions do not affect the human resources needed to maintain and update library infrastructure (Hill, 2008). The pressure to maintain staff levels coupled with rising costs of resources, predatory publishing practices, and a more complex networked information infrastructure result in technical services units that are often overwhelmed.

Open source

Open source refers most frequently to software code that is openly available. In librarianship, open source goes beyond this definition to include community, collaboration, innovation, shared principles, and often a transparent governance structure (Engard, 2010). Open source in librarianship and in particular technical services includes open source software systems such as those needed to perform core operations including circulation, cataloging, acquisitions, and tools needed to manipulate and manage library metadata. Open source solutions have been used for decades in technical services. A good example, often referred to as the swiss army knife of working with library data, is the software developed by Terry Reese called MarcEdit, originally created in 1999 (Reese, 2018). MarcEdit provides a set of tools for working with large sets of library data and was originally created as a way to automate large cleanup projects at the Oregon State University Libraries. Since that time, it is the preferred solution for working with library data all over the world. Thanks to its creator and maintainer Terry Reese, this tool now goes beyond library data and provides many more features to work with library data in different standards and formats. In the realm of core functional systems sometimes referred to as “integrated library systems” or more commonly now, “library service platforms”, an example is Evergreen, an open source integrated library system developed in 2006 by the Georgia Public Library System (Evergreen.org, 2017). Evergreen is now the library system used by more than 2000 libraries to manage data critical to daily operations such as circulation, acquisitions, or cataloging. These tools and library services platforms are intimately associated with their communities where interested parties strive to improve not just the open source software but also how the community and software is governed, developed, and shared. While there are a number of open source tools for those working in technical services, the availability of open source library service platforms remains very small.

Library services platforms

The move to open source library service platforms started early, as we saw with Evergreen. Koha, another open source integrated library solution, was developed around the same time as Evergreen in 1999 (Koha.org). Though these open source systems have been implemented worldwide, their adoption has been limited. Evergreen and Koha play a small role in the already very limited pool of options for library service platforms (Breeding, 2020). One reason for this is the rise of large companies that have multiple arms that serve both content and software to libraries. These large companies often acquire smaller or less diverse ones to gain a larger share of the market (Breeding, 2020). Further, these large conglomerates provide the core functions needed by libraries with varying degrees of customization allowed, and sometimes even with features that extend into the realm of managing scholarly output or digital resources. Because of this complexity in the core systems used by the library for its central functions, and the content that often comes from that same vendor, these systems are costly, can be hard to maintain, and are often difficult to update. This complexity often translates as high costs to acquiring and maintaining these core systems. This coupled with sometimes predatory publishing practices put many libraries in difficult situations, often having to make decisions on systems and content that can be purchased, leased, or rented. As a response to this and the small playing field of vendor solutions, more open source library service platforms are being considered. While Evergreen and Koha tend to be implemented in public libraries, FOLIO, or the Future of Libraries is Open, is being developed as an open source platform and community for academic institutions (FOLIO.org).

In spite of the above considerations, implementing open source should not be equated with acquiring tools or complex software like library service platforms free of charge (Morgan, 2004). Whether an open source or a vended solution is implemented, there is a cost for the core infrastructure of libraries in carrying out core services and managing the multiple types of resources that libraries manage for their patrons. In the open source community, this translates to being involved in the development and maintenance of the software, training, documentation, and/or governance. It is often said that a community is only as strong as its members.

Diversity and inclusion

In order to improve access to information for readers who are not connected to well-resourced libraries, many research libraries provide an institutional repository where scholars can upload preprints or postprints of their publications. This research is accessible free of charge around the world. Libraries also promote affordable access to educational materials to help students afford the information they need for their coursework. One way libraries provide affordable access to educational materials is by supporting the creation and dissemination of Open Educational Resources.

Advocating for inclusive access to research materials relies on both making materials affordable and making resources discoverable. It is crucial to recognize that inclusive access hinges on library users' being able to find resources in a way that embraces diversity and reflects their search habits. Hence, the role of library technical services in acquiring, managing, and serving up resources and content for library patrons must take into account the ways in which our metadata and systems are biased. Like many institutions, the information used to describe resources and often the tools and software used in technical services need to be improved. The standards and practices used in technical services were created by individuals with biases and blind spots. The recent example of trying to update the Library of Congress authorized subject term, "Illegal alien", to a more inclusive term captures not only the bias inherent in existing systems, but also the challenges faced when moving the community toward more inclusive standards and practices (Aguilera, 2016). Another example of combatting bias is the story of Dorothy Porter, named inaugural curator of the Howard University's Negro Collection or Mooreland Foundation in 1930. She unlearned her library training to develop a classification system inclusive to Howard's black students and black audiences in general (Helton, 2019). To avoid biases, errors, or oversights, the process of developing and maintaining library infrastructure should not be invisible or go unnoticed (Helton, 2019). In recent years, a surge of scholarship and practices has emerged around critically evaluating how information is used to describe and make resources available through library systems.

References

- ACRL, 2021. Summary Results. Retrieved from. <https://www.acrlmetrics.com>.
- Aguilera, J., 2016. Another word for 'Illegal Alien' at the library of congress: contentious. The New York Times. <https://www.nytimes.com/2016/07/23/us/another-word-for-illegal-alien-at-the-library-of-congress-contentious.html>.
- ARL, 2021. Data Repository. Retrieved from. <https://www.arlstatistics.org/>.
- Banerjee, K., 2020. The Linked Data Myth. Library Journal. Retrieved February 15, 2021, from. <https://www.libraryjournal.com/?detailStory=the-linked-data-myth>.
- Beall, J., 2009. The Charleston Advisor Bentham Open, vol. 11. Charleston.publisher.ingentaconnect.com, p. 1. Retrieved February 15, 2021, from. <https://charleston.publisher.ingentaconnect.com/contentone/charleston/chadv/2009/00000011/00000001/art00008#>.
- Berners-Lee, T., Hendler, J., Lassila, O., 2001. The semantic web. Sci. Am. 284 (5), 34–43. <https://doi.org/10.1038/scientificamerican0501-34>.
- Breeding, M., 2020. 2020 Library systems report: fresh opportunities amid consolidation. American Libraries. <https://americanlibrariesmagazine.org/2020/05/01/2020-library-systems-report/>.
- Brunner, M., Osario, J., 2017. Recruitment, retention, diversity, and professional development. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 143–160.
- Brunner, M.L., 2010. PhD holders in the academic library: the CLIR postdoctoral fellowship program. In: Walker, S., Williams, K. (Eds.), The Expert Library: Staffing, Sustaining, and Advancing the Academic Library in the 21st Century. Association of College and Research Libraries, Chicago, pp. 158–188.
- Budd, J.M., 2018. The Changing Academic Library: Operations, Culture, Environments, third ed. Association of College & Research Libraries.
- Cerbo II, M.A., 2011. Is There a Future for Library Catalogers. Technical Services Department Faculty Publications, p. 41. http://digitalcommons.uri.edu/lib_ts_pubs/41.
- Ciota, R., Knight, R., 2020. Introduction: the challenges facing library technical services: a dialogue. In: Marien, S. (Ed.), Library Technical Services: Adapting to a Changing Environment. Purdue University Press, West Lafayette, Indiana, pp. 1–14. <https://doi.org/10.2307/j.ctvs1g8h5.4>.
- Engard, N.C., 2010. Practical open source software for libraries. Chandos Publishing. <https://doi.org/10.1016/B978-1-84334-585-5.50001-1>.
- Faulkner, A., 2017. Challenges facing technical services. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 115–128.
- Forbes, C., Keeran, P., 2017. Reference, instruction, and outreach: current methods and models. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 85–100.
- Helton, L., 2019. On decimals, catalogs, and racial imaginaries of reading. PMLA/Publications of the Modern Language Association of America 134 (1), 99–120. <https://doi.org/10.1632/pmla.2019.134.1.99>.
- Hill, J.S., 2008. Is it worth it? Management issues related to database quality. Catal. Classif. 46 (1). <https://doi.org/10.1080/01639370802182885>.
- Hoffman, S., 2017. In: Gilman, T. (Ed.), Organization, Administration, Management, and Planning. Academic Librarianship Today. Rowman & Littlefield, pp. 29–46.
- Lewis, D.W., 2016. Reimagining the Academic Library [electronic resource]. Rowman & Littlefield.
- Long, K., 2020. From records to data: the new purposes of cataloging. In: Marien, S. (Ed.), Library Technical Services: Adapting to a Changing Environment. Purdue University Press, West Lafayette, Indiana, pp. 17–34. <https://doi.org/10.2307/j.ctvs1g8h5.5>.
- Morgan, E.L., 2004. Open source software in libraries. <http://infomotions.com/musings/biblioacid/>.
- Owen, B., 2017. Open access, institutional repositories, e-sciences, and data curation, and preservation. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 197–210.
- Rauf, T., 2017. Funding and budgeting. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 47–68.
- Reese, T., 2018. About MarcEdit. MarcEdit Development. Retrieved March 10, 2021, from. <https://marcedit.reeset.net/about-marcedit>.
- Uziel, L., 2017. Collection development: general, special, print and digital, and resource sharing. In: Gilman, T. (Ed.), Academic Librarianship Today. Rowman & Littlefield, pp. 101–114.

Uses and abuses of social media in higher education

Lillian Vega Castaneda^a and Mario E. Castaneda^b, ^a Professor Emerita Education California State University Channel Islands, Camarillo, CA, United States; and ^b California State University, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Learning management systems and social media in teaching and learning	132
Social media incorporated into LMS	133
When social media and universities collide	134
Challenges of higher education and social media	136
Privacy issues	136
Access issues	136
Closing thoughts	137
References	138

Over 3.6 billion people use social media globally. This is projected to increase to about 4.41 billion by 2025. China has the highest numbers of social media users across the globe with 999.95 million users. The remaining top five countries with the highest numbers of social media users in 2021 include (in millions): India (639.47 users), United States (295.48 users), Indonesia (193.43 users) and Brazil (159.01 users) (Statista, 2021).

Universities have embraced the use of social media around the world. Higher education has incorporated social media and its' use as an online platform that supports formal and informal communication. Education has embraced the use of social media globally (Chugh and Ruhi, 2017). Social media in higher education is used in several ways (1) university communication with current students and prospective students; (2) as a support or main "platform" for teaching and learning; (3) as a way of sharing information from university academic based centers We focus on the use of social media in teaching and learning and the associated benefits and challenges.

Institutions of Higher Education (IHE) communicate regular information with students and faculty via their official university Facebook, Twitter, Instagram or TikTok, etc., accounts. Different campus offices often have their own social media accounts, e.g., student affairs, President's Office and academic programs, Library, Sports, Health, etc. These offices provide a variety of information such as keeping students apprised of emergency information, campus closures, announcements, events, etc.

When the Covid-19 Pandemic hit different countries, the need for social media became more prevalent. For example, when Covid-19 arrived in the United States, emergency information was sent via telephone messages, text-messages, Email, LMS messages and via social-media that various campuses would move to 100 % remote instruction. The United States serves many international students. Chinese students represent 33.7 % of the international student population in the US. They are followed by Indian students at 18.4% of the international students in the US. The COVID-19 pandemic left many international students unable to return to their US campus (Times Higher Education, 2020). Across the globe, many countries began to rely on social media use. According to the United Nations:

UNESCO estimates that over 1.5 billion students in 165 countries are out of school due to the COVID-19 pandemic. The pandemic has forced the global academic community to explore new ways of teaching and learning including distance and online education. This has proven challenging for both students and educators who must deal with the emotional, physical and economic difficulties posed by the illness while doing their part to help curb the spread of the virus. The future is uncertain for everyone, particularly for millions of students scheduled to graduate this year who will face a world crippled economically by the pandemic.

(United Nations, Accessed November 22, 2021).

With the onset of COVID-19, many universities around the world were forced to close their doors to students, faculty and staff and immediately transferred to remote learning or work:

...education officials have been forced to cancel classes and close the doors

To campuses around the world in response to the growing Coronavirus

Outbreak.

(Times Higher Education, 2020).

It is safe to say that over the course of the COVID-19 Pandemic, universities across the globe saw an increase in social media use. By asynchronous, we refer to 100% online instruction on an LMS or social media faculty created LMS. This means that all instruction occurs online and there is no face-to-face meeting with the instructor via Zoom or some other video platform. Synchronous instruction during COVID-19 includes the use of LMS and social media and/or social media as an LMS created by faculty. This includes meeting face-to-face via Zoom or Teams or some other type of video meeting. Indeed, different higher education institutions across the world moved from face-to-face to online, remote, synchronous, and asynchronous education because of the COVID-19 Pandemic. Some educational institutions were able to move into online instruction with some ease, in that, they had established LMS structures in-place. In some developing countries, this was not the case. These universities did not have access to LMS and turned to the use of platforms such as Google Classroom and Zoom (Sobaih et al., 2020).

In the case of developing countries, there was a move to incorporate free social media sites into the remote teaching and learning environments. In some countries, faculty were challenged to identify social media sites to due to the lack of LMS. Sobaih et al. (2020) note:

However, public universities in many underdeveloped nations, e.g., Egypt, are

Suffering from a lack of technological platforms and formal online learning

Management systems (LMS) for communication with students or with their

Faculty members. Hence, they do not have the full capabilities to support the

Online learning process (page 1).

It should be noted that to commemorate Spanish Language Day and to consider the impact of.

COVID-19 on IHE's, the United Nations Academic Input (UNAI) "hosted a Spanish language webinar on April 23." Over seventy universities from Latin America were represented. Other universities in attendance were from Latin America. Other universities in attendance were from Japan, France, and Spain. (UNESCO, Report, United Nations, n.d.). The participants of the Spanish Language Day Webinar came together to discuss the impact of the pandemic in higher education, and as Christine Gallach, Secretary of State for Foreign Affairs (Spain) noted: "We have to take this opportunity to deepen this concept of shared mankind as we share not only a pandemic but a planet and ourselves" (Page 2, UNESCO, Report, United Nations, n.d.).

Gallach called for "academic cooperation and multilateral assistance to overcome the difficulties posed by the pandemic and contribute to the common good" (Page 2).

During the UNAI webinar, universities were called upon to collaborate and share new ways of sharing knowledge. Milenko Skoknic (Permanent Representative of Chile to the United Nations) stated that:

...solidarity was embedded in higher education and universities could facilitate the sharing of knowledge thru open courses, joint research, and new creative teaching methods and educational alternatives (Page 2).

In the next section we consider the different ways universities around the world began to deliver 100% online instruction thru a combination of LMS and social media and social media (only) as a type of informal, instructor designed "LMS."

Learning management systems and social media in teaching and learning

Faculty use social media to enhance instruction, or as an easy way to communicate with students (e.g., post announcements, assignments). Social media is often used as a support to a Learning Management System (LMS) used at a given university. Faculty use a combination of social media apps to support an LMS like Canvas, Blackboard or Moodle. Other LMS platforms include Google Classroom, which represents 39% of the LMS market. Other LMS used world-wide include: Mindflash, Skyprep, Proprofs, iSpring Learn, Talentme, Cocebo and Litmos (Software Testing Help, 2021). As of April 2020, 43% of educational institutions invested in new online learning resources (Workplace Learning Report, LinkedIn, 2021). In some cases, social media may be used as a faculty-designed LMS, especially given the last 18 months of the Covid-19 Pandemic which saw most (if not all) instruction move to an

on-line experience (asynchronous and synchronous). Notably, the software market that “loomed after COVID-19, LMS might be the biggest.” (Trustradius, 2021) (<https://www.trustradius.com/vendor-blog/lms-statistics-trends>). By the start of the Covid-19 Pandemic, higher education institutions moved online. This created a critical need for LMS software. As a result of the pandemic: “The LMS market exploded in 2020 and will continue growing well past 2021.” (www.trustradius.com). Indeed, as many universities have welcomed some students back to the campus, many faculties and students continue to meet asynchronously or synchronously with the support of social media.

Social media incorporated into LMS

Faculty incorporate social media to support teaching and learning outside of a formal Learning Management Platform using social media like a Facebook (closed) Group (where the instructor can set the class up as a group that is closed to the public and consists only of members of the class. The course instructor gives permission for the students to join the group.) that may focus on course content or be used by the instructor and students to communicate (Vega-Castaneda and Castaneda, 2019). Some LMS like Canvas, Blackboard or Moodle incorporate social media. For example, Canvas incorporates Twitter and YouTube. Most LMS’ allow the instructor to add a URL link for social media use, e.g., Twitter, YouTube, Facebook. Social media is also incorporated into course assignments and for sharing content. Facebook and YouTube allow the instructor to post lectures, information, discussions, upload files, etc. Students may use the same apps to respond to discussions or make presentations. Often, social media is incorporated into assignments. For example, in a history social studies course that I (Lillian) teach, students are required to create a Pinterest page that focuses on some aspect of the (content) of the course. In a course on first and second language acquisition, students present their projects on YouTube or Facebook video or Voice-Thread. Additionally, Facebook is used as a teaching tool and information tool. As part of the first and second language acquisition/bilingual course that I teach, students post bi-weekly language and culture related news stories. We discuss the social, political, and cultural aspects of language acquisition on social media. We also discuss the treatment on social media of individuals who are Emergent Bilingual/English Learners and immigrants and the challenges they face coming to a foreign country and learning a new language and culture. This makes it very easy and accessible for students and instructor alike.

Institutions of Higher Education (IHE) have seen an increased use of social media to support LMS or to serve as the main LMS. For example, some faculty have set up Google Classroom which is free. It allows the instructor to upload assignments, tests, and other course content. It also allows the instructor to upload video’s, PDFs, Google Docs, and various URL links. During the pandemic, instructors created their personalized and free self-designed learning management systems when their university did not have a formal LMS. These instructors created their personal “LMS” using social media apps, e.g., Google Classroom, YouTube, Facebook, Instagram, Twitter, etc. All of these are free social media apps and have aided in the creation of content and tools for communication between the instructor and student, and perhaps most importantly, promoted student to student interaction.

Facebook has been one of the main social media sites used as an educational tool. The benefits include all the aspects of sharing information and creating instances for posting readings and discussion that have been previously mentioned. One way to contrast the use of social media with a traditional LMS is how discussions occur. In comparison, a discussion on social media creates a less traditional way of communicating, in that, posts can be created, responded to information (including links to other sites) can be readily and instantly shared for everyone in the class to see. Facebook has multiple benefits that relate to increased interaction between instructor and student, and student and student. It promotes the exchange of academic content and has seen increased engagement and a “convenience of learning” in an informal (yet academic) context. Facebook has been found to enable connectivity, communication, and collaboration (Chugh and Ruhi, 2017). Indeed, the uses of social media are many:

Social media can be utilized to deliver teaching material, educational information,

Updates and facilitate communication and collaboration. Facebook’s social

Communication aspects benefit both educators and students by creating an online

Class group and increasing interactions between teacher-student and student-student

(Muñoz and Towner, 2011) and promoting collaborative learning (Menzie et al., 2015).

Chugh and Ruhi (2017).

Of note, due to the COVID-19 pandemic, there was an increase in students who began to use social media and created study groups and worked on assignments together. Al-Rahmi et al., 2020, studied students in Pakistan and focused on their use of social media for student interaction during the COVID-19 epidemic. This occurred due to the closure of university campuses, libraries, including the closure of non-university meeting places like coffee-shops or other meeting places. Students were forced to move on-line not only for courses, but for peer interaction. Students met on Facebook, Zoom, and like social media networks. Students used Google Docs to collaborate on academic work that called for collaboration.

Students in some developing countries view some social media as more interactive and user friendly than other free online platforms. Sobaih et al. carried-out research in Egypt and the study consisted of nine public colleges that grant bachelor's degrees in Tourism and Hotel.

Students have more patterns and practices in social media than their faculty member who just focuses on formal teaching and learning. Students used social media for connecting with their faculty member, support each other and building an online community.

Sobaih et al. (2020, p. 14).

When social media and universities collide

Most universities house a variety of centers that focus on some aspect of the academy. In the United States there are centers that address issues of hate and extremism. Other centers focus on specific areas of content that address content in teaching K-12 students. Here are some examples of academic centers that communicate and provide information (on their centers) via social media. For example, California State University, San Bernardino houses the Center for Hate and Extremism <https://www.csusb.edu/hate-and-extremism-center>. Professor Brian Levin is the director of the center. The Center for the Study of Hate and Extremism looks at the following: Anti-Semitic hate, Anti-Asian hate, Anti-Asian Prejudice, World-wide Threats to the Homeland (Professor Brian Levin testifies to the US House of Representative Committee on Homeland Security (September 17, 2020 & September 19, 2019).

The Center for the Study of Hate and Extremism has a Twitter and Facebook account. @proflevin is the Twitter account of Professor Brian Levin; the Facebook page is called "Center for the Study of Hate & Extremism." The Center has also produced different video's including a video on the Radicalization of CA Police Training Video; Black Lives Matter Video – BLM/Justice for Wyandotte (KS); First AME Church; Professor Levin has appeared on MSNBC, CNN, Bill Mahr, American Jewish Committee, and his reach is at the national level. This is just one example of a university-based center that addresses issues covered in the mainstream and social media.

Just as there are many advantages of using social media for communication and teaching and learning, social media in higher education can be mis-used in ways that may impact the university or student, faculty, or staff. Social media can be mis-used by having users post negative or harmful information that may be posted in response to the original information that was shared. On Twitter, for example, at the non-institutional level, something could be posted by a private Twitter user and another Twitter user could post a negative or harmful or threatening response. In this instance, the user may block that individual from their Twitter feed. Potentially, if a university post on Twitter or Facebook or Instagram account receives a negative response, there is the potential for the University to be accused of not respecting the students first amendment right to free speech—if the response is removed, or if the individual is blocked. Additionally, if a threatening post is made on a course social media page, the instructor must follow university policy and, in most cases report the incident to the University Police or Student Affairs office (or equivalent).

There are some recent events/instances where social media has been used to attack individuals in higher education; Here are some examples from the United States, in the post-Trump era. Recently, the author of the 1619 Project, Nikole Hannah-Jones was denied tenure at the University of North Carolina Chapel Hill (Robertson, 2021). Members of the Trustees overturned the recommendation for tenure made by the faculty. It is believed that the major reason the Trustees denied Hannah-Jones' tenure is the content she writes about in the 1619 Project, which covers the history of slavery in the United States. The 1619 Project has been used as a rallying cry for those who do not want to see the history of slavery and treatment of blacks told from a non-White perspective. It is believed that Hannah-Jones' tenure negative tenure decision occurred because the Trustee's did not "agree" with her 1619 writings. The 1619 Project and Hannah-Jones' tenure denial and ultimate offer of Full Professor and tenure at Howard University were covered across many different social media platforms. Some social media posts supported Hannah-Jones', but others saw individuals posting racist tropes. Hannah-Jones used the negative portrayal of the 1619 Project and turned it into a positive by announcing her acceptance of Full Professor with Tenure at Howard. The announcement was posted on major social media sites. In part, she states:

I will be taking a position as the inaugural Knight Chair in Race and Reporting at

Howard University, founded in 1867 to serve the formerly enslaved and their descendants

There, I will be creating a new initiative aimed at training aspiring journalists to cover the

Crisis of our democracy and bolstering journalism programs at historically Black

Colleges and universities across the country. I have already helped secure \$15 million for

effort, called the Center for Journalism and Democracy....

Hannah-Jones (2021).

The main author of the 1619 project now a Professor at Howard University (a Historically Black Research University), and the history of slavery will be covered across some universities and schools un the US. Some K – 12 school districts have incorporated the 1619 project into their curriculum. According to Education Next (Riley et al., 2020): “School districts in Chicago; Newark, N.J; Buffalo, N.Y., and Washington, DC all announced 1619 Project related events.”

The debate continues to occur via social media and shared by news media on social media, e.g., MSNBC, CNN, PBS, Fox, NBC, ABC, national and local news on their social media sites. One only must read the comments from the public to get an idea of the divide between teaching about the history of racism and slavery in the United States and those who argue that it doesn’t need to be taught, as if it never happened. For specific information on the 1619 Project please follow this link:

https://pulitzercenter.org/sites/default/files/full_issue_of_the_1619_project.pdf.

Universities and K-12 schools have recently been a subject of debate on various social media sites. The topic of debate – Critical Race Theory CRT) has been a topic of discussion on social media and has overlapped into whether it should be taught in K-12 schools (it isn’t). Universities have held different discussions on CRT with scholars as a way of countering the misinformation surrounding CRT. Critical Race Theory is taught in higher education, usually in law school. It can also be taught in master’s and Doctoral programs that address issues of race and racism. CRT is a way of looking at racial relations, particularly within the United States. It is looked at from the historical background of race and racism. CRT began in the mid-seventies, as several people in the legal profession began to worry about the slow rate at which laws were changing to promote racial equality, and that many of the early victories of the Civil Rights Movement were being eroded. One of the key elements of CRT is a way of looking critically at race relations in the United States, examining everyday interactions, and finding the racial component in them, that moves beyond the “color blind” approach to race relations. Although critical race theory began with the legal profession and legal professor Derrick Bell (easily the most important thinker within the movement) it has since spread to many other disciplines.

Educators may find CRT very important to their understanding of classroom dynamics, testing, tracking students, including curriculum bias. Individuals involved in the political sphere may find critical race theory useful in understanding voting discrepancies, race-based campaigning, and other issues. To be sure, it is not a surprise that higher education in the US has been at the forefront of the debate. Perhaps one of the most decisive arguments of CRT as shown on social media, is the questioning of the normative acceptance of what it is to be white. According to “Words that Wound: Critical Race Theory, Assaultive Speech, and the First Amendment (Matsuda et al., 1993, 2019):

Critical race theory recognizes that racism is endemic to American life.

Thus, the question for us is not so much whether or how racial discrimination

Can be eliminated while maintaining the integrity of other interests implicated

In the status quo such as federalism, privacy, traditional values, or established

Property interests. Instead, we ask how these traditional interests and values serve

As vessels of racial subordination.

Since CRT is taught in universities across the United States, the recent social media debate has re-emerged as a leading talking point within the political arena. CRT is most likely being questioned as it expresses skepticism toward dominant legal claims of neutrality, objectivity, color blindness and meritocracy. Because CRT's origins are in law, it insists on the recognition of experiential knowledge of people of color and our communities of origin, while analyzing law and society. To be clear, CRT is not taught in K-12 educational settings, yet the rhetoric on social media arguing against the teaching of CRT is present on social media and this connects back to higher education. To be clear, CRT is not taught in K-12 US schools, and it originated in the preparation of lawyers.

Challenges of higher education and social media

There are challenges associated with using social media as a teaching resource.

One of the main drawbacks is the lack of real time, face to face communication between the instructor and student and student to student. In an asynchronous setting, with an LMS and social media, or an instructor developed LMS consisting of social media apps, the lack of interaction can be described as removed. Although interaction on social media sites, such as Facebook can encourage interaction, written and in some cases, by video, the lack of a given meeting time on a platform like Zoom can create a distanced and removed communication approach. Conversely.

"Social media reduces boundaries between university faculty, students and audiences with access to social media platforms." (Ngai et al., 2015). While it is true that boundaries can be reduced, it is also true that the social distance between the instructor and students can also become removed. This clearly presents a challenge, and it is the instructor's role to ensure that communication and interaction on social media does not become static, and that communication and interaction occurs from professor to student, student to student and student to professor.

Privacy issues

Does the university have rules or guidelines for social media use? In some global universities, there is a challenge about using social media. Not all universities around the world support the use of social media (Waller et al., 2019).

If the instructor creates a closed group, for example a closed Facebook group or a closed Instagram group, some degree of privacy (between the group members) can be assumed. However, in the case of Facebook, each member's FB profile page can be viewed by everyone in the class.

With social media access to forums such as Facebook, both professors and students can see information, professional pictures, and communications about one another bringing about issues with professionalism.

Shelton (2016).

Some faculty members have identified challenges of using social media. Some of the barriers involve concerns with the digital divide, a sense of security, and loss of control. Some faculty do not think social media is appropriate for academic use (Sobiah et al., 2020, p. 3; Sobiah et al., 2016; Manca and Ranieri, 2016).

Waller et al. (2019) also identify educational concerns. How does the instructor which content to include on social media? What is considered acceptable and relevant to the subject matter of the course? How do students know what to share? What is considered acceptable and relevant? Many universities have policies that specify social media "behavior" when representing a university. Some universities require instructors to present a "plan" of how they will use or integrate social media into their courses. Instructors also include "social media etiquette" guidelines in their course syllabi.

Some faculty are hesitant to use social media (or LMS)—this can lead to a lack of organization of a remote classroom experience; Instructors need to be well-versed in technology, using social media and LMS (Waller et al., 2019).

Access issues

Student and faculty access issues to an LMS & social media on a computer, tablet, smartphone etc., also present a challenge. Do they have access to a smartphone or a computer? Does the student or faculty member have access to the internet? Does the university have the technological infrastructure to support remote teaching and learning? Do students and/or faculty know how to use social media and/or LMS and social media? Do students feel free to initiate posts? As noted earlier, during the Spanish Language Day Webinar hosted by the U.N. Academic Input (UNAI), participants discussed the lack of internet access for many students and professors. Francesco Pedró, director of the UNESCO International Institute of Higher Education in Latin America and the Caribbean noted:

While institutions of higher education have to continue their teaching responsibilities, at least a third of students and an equal percentage of professors do not have access to the internet, hampering the success of a temporary transition to virtual education.

United Nations (n.d.).

Issues of social media use in higher education and potential disruption of teaching and learning are not necessarily new to higher education (Digital Disruptions: Handling Social Media Misuse.

By Students and Educators. Fisher-Philips, n.d.): "However, use of social media by students and educators outside the classroom can have profoundly negative impact on the learning environment." Reference is made to Harvard University who, in 2017 learned that "dozens of members of the Class of 2021 participated in a private messaging group that included messages and images insulting people based on their national origin, contained sexually explicit material, and mocked the Holocaust, among other things (Page 2)."

Harvard University ended up rescinding the admission offers to these individuals (at least 10). This incident is considered one of the first where a private university reacted to what was posted by potential students on social media.

Public universities must ensure that "any student discipline does not run afoul of First Amendment free speech rights. Under longstanding constitutional law, a student receives protection if the offending speech addresses a matter of public concern and is not sufficiently disruptive to the learning atmosphere in the school" (Fisher-Phillips, n.d., p. 3).

There are also requirements for how faculty conduct their selves on social media. Social media policy should focus on how faculty and students are to engage and represent the university:

Instead of focusing on the delivery method, your social media policy should focus on the nature and effect of the expression by prohibiting specific conduct that would likely violate your ethics, harassment, or other applicable policies.

Fisher-Philips (n.d., p. 3).

Because social media is considered (somewhat) a new frontier in higher education and teaching and learning, the use of social media and policies that guide it, are still at the emerging stage. Yet, there are some court cases that we can look to, including *Pickering v. Board of Education* (1968) which involved a teacher who was fired for writing a letter to the editor in which the criticism, aimed at the superintendent and school board, focused on the high expenditures for sports as opposed to academics. The termination of the teacher was found unlawful and the "Supreme Court stated that government employers must arrive at a balance between the interests of the teacher, as a citizen, in commenting upon matters of public concern and the interests of the State, as an employer, in promoting the efficiency of the public services it performs through its employees" (Fisher-Philips, n.d., p. 7).

Closing thoughts

The use of social media in higher education is a new occurrence. Some instructors have used social media in their coursework prior to the Pandemic – in my (Lillian's) case, for ten years or more, where I have linked a Closed Group Facebook page and used it in conjunction with on-campus teaching and learning and an LMS. Social media like Facebook is easily integrated into coursework and is often attached to an LMS and a combination of face-to-face teaching and learning. The arrival of the COVID-19 pandemic threw many universities across the globe into an immediate movement of all face-to-face instruction to online, virtual instruction. The move to virtual instruction, due to the pandemic also included an increased use of traditional LMS like Canvas, Moodle, Blackboard, Google Classroom, etc., and in many cases, included the use of social media, e.g., Facebook, Twitter, Instagram, Pinterest, YouTube, and others. Some faculty incorporated social media into their LMS as a way of encouraging easier communication and information sharing, content, and exchange of ideas. Other universities, including those who did not have a traditional LMS relied on using social media as the "LMS" for delivering instruction, including teaching and learning. We refer to the 100% faculty created LMS as "faculty designed."

With the arrival of the COVID-19 Pandemic, universities across the globe moved almost instantly from on-campus, face-to-face instruction to all classes moving to an online/virtual environment. The impact of the Pandemic will not be known for some time, as the world is still in the midst-of the COVID-19 Pandemic, even though several vaccines have been adopted and used across the world. During the summer of 2021, the Delta variant arrived, and this led to the uptick in COVID-19 positivity rates and fatalities.

The delta variant, which was first detected in India last October, has been found in more than 130 countries globally, according to the World Health Organization.

Lee (2021).

Currently, in the Fall of 2021, the world is seeing the emergence of another new variant in "Southern Africa, Botswana, Belgium, Hong Kong and Israel."

The World Health Organization gave the new version of the virus the name Omicron and called it a "variant of concern, its most serious category." This variant has many mutations, some of which are concerning, the W.H.O. said in its official description. "Preliminary evidence suggests an increased risk of reinfection with this variant."

(Horowitz et al., November 26, 2021, NY Times).

Universities continue to use social media as a component of a formal LMS or as a faculty designed LMS consisting of social media apps. Some universities around the world have re-opened on-campus instruction with protocols in place to screen for COVID-19. Faculty and students continue to explore the role of social media in higher education with a variety of results. Depending on the faculty and student experience in using social media, the results of its use will be varied.

Clearly, there are challenges to incorporating social media into university teaching and learning, but we believe the benefits outweigh the drawbacks. The move to online instruction because of the pandemic has opened a different way of communicating and using knowledge across the globe and this will be a subject of continued study.

References

- Al-Rahmi, W.M., Yahaya, N., Alturki, U., et al., 2020. Social media – based collaborative learning: the effect on learning success with the moderating role of cyberstalking and cyberbullying. *Interact. Learn. Environ.* 1–14.
- Chugh, R., Ruhi, U., 2017. Social media in higher education: a literature review of Facebook. *Educ. Inf. Technol.* 23 (2), 605–616.
- Hannah-Jones, N., 2021. Nikole Hannah-Jones issues statement on decision to decline tenure offer at university of North Carolina-Chapel Hill and to accept Knight Chair appointment at howard university. NAACP Legal Defense and Educational Fund. Available from: <https://www.naacpldf.org/press-release/nikole-hannah-jones-issues-statement-on-decision-to-decline-tenure-offer-at-university-of-north-carolina-chapel-hill-and-to-accept-knight-chair-appointment-at-howard-university/>. (Accessed 26 November 2021).
- Horowitz, J., Chutel, L., Ives, M., 2021. W.H.O. says new variant in South Africa is 'of concern' as countries impose travel restrictions. *The New York Times*. Available from: <https://www.nytimes.com/2021/11/26/world/covid-variant-travel-restrictions.html>. (Accessed 26 November 2021).
- Khan, M.N., Ashraf, M.A., Seinen, D., et al., 2021. Social media for knowledge acquisition and dissemination: the impact of the COVID-19 pandemic on collaborative learning driven social media adoption. *Front. Psychol.* 12.
- Lee, Y.N., 2021. 3 charts show how far Covid delta variant has spread around the world. *CNBC*. Available from: <https://www.cnbc.com/2021/08/06/covid-charts-show-how-far-delta-variant-has-spread-around-the-world.html>. (Accessed 26 November 2021).
- Levin, B. (n.d.a). Global terrorism: threats to the Homeland, part I. Available from <https://www.csusb.edu/sites/default/files/GLOBAL%20TERRORISM-%20cong%20BL2%2091019.pdf> [Accessed 26 November 2021].
- Levin, B. (n.d.b). Worldwide threats to the homeland, part I. Available from [https://www.csusb.edu/sites/default/files/PDF-HSC-Statement of Brian Levin_0.pdf](https://www.csusb.edu/sites/default/files/PDF-HSC-Statement%20of%20Brian%20Levin_0.pdf) [Accessed 26 November 2021].
- Manca, S., Ranieri, M., 2016. Facebook and the others. Potentials and obstacles of Social Media for teaching in higher education. *Comput. Educ.* 95, 216–230.
- Matsuda, M.J., Lawrence, C.R., Delgado, R., et al., 2019. *Words that Wound: Critical Race Theory, Assaultive Speech, and the First Amendment*. Routledge, New York, NY.
- Menzies, R., Petrie, K., Zarb, M., 2015. A case study of Facebook use: outlining a multi-layer strategy for higher education. *Educ. Inf. Technol.* 22 (1), 39–53.
- Muñoz, C.L., Towner, T., 2011. Back to the “wall”: how to use Facebook in the college classroom. *Clin. Hemorheol. Microcirc.*
- Ngai, E.W., Moon, K.-L.K., Lam, S., et al., 2015. Social media models, technologies, and applications. *Ind. Manag. Data Syst.* 115 (5), 769–802.
- Phillips, F. (n.d.). Digital disruptions: handling social media misuse by students and educators. *Fisher Phillips*. Available from <https://www.fisherphillips.com/news-insights/digital-disruptions-handling-social-media-misuse-by-students-and-educators.html> [Accessed 19 November 2021].
- Riley, N.S., et al., 2020. “The 1619 project” enters American classrooms. *Education Next*. Available from: <https://www.educationnext.org/1619-project-enters-american-classrooms-adding-new-sizzle-slavery-significant-cost/>. (Accessed 26 November 2021).
- Robertson, K., 2021. Nikole Hannah-Jones denied tenure at university of North Carolina. Available from: <https://www.nytimes.com/2021/05/19/business/media/nikole-hannah-jones-unc.html>.
- Shelton, C., 2016. Giving up technology and social media: why university lecturers stop using technology in teaching. *Technol. Pedagog. Educ.* 26 (3), 303–321.
- Sobaih, A.E.E., Moustafa, M.A., Ghandforoush, P., et al., 2016. To use or not to use? Social media in higher education in developing countries. *Comput. Hum. Behav.* 58, 296–305.
- Sobaih, A.E.E., Hasanein, A.M., Elnasr, A.E.A., 2020. Responses to COVID-19 in higher education: social media usage for sustaining formal academic communication in developing countries. *Sustainability* 12 (16), 6520.
- Software Testing Help, 2021. 15 Best Learning Management Systems (LMS of the Year 2021). Available from: <https://www.softwaretestinghelp.com/learning-management-system/>. (Accessed 20 November 2021).
- Statista, 2021. Topic: Social media use during coronavirus (COVID-19) worldwide. Statista. Available from: <https://www.statista.com/topics/7863/social-media-use-during-coronavirus-covid-19-worldwide/>. (Accessed 23 November 2021).
- Times Higher Education (THE), 2020. The impact of coronavirus on higher education. Available from: <https://www.timeshighereducation.com/hub/keystone-academic-solutions/p/impact-coronavirus-higher-education>. (Accessed 25 November 2021).
- Trustradius, 2021. 49 LMS statistics and trends for a Post-COVID world. Available from: <https://www.trustradius.com/vendor-blog/lms-statistics-trends>. (Accessed 26 November 2021).
- United nations. (n.d.). COVID-19 and higher education: education and science as a vaccine for the pandemic. Available from <https://www.un.org/en/academic-impact/covid-19-and-higher-education-education-and-science-vaccine-pandemic>.
- Vega-Castaneda, L., Castaneda, M., 2019. *Teaching and Learning about Difference through Social Media: Reflection, Engagement, and Self-Assessment*. Routledge, New York, NY.
- Waller, R.E., Lemoine, P.A., Richardson, M.D., 2019. Instruction using social media in global higher education. *J. Educ. Dev.* 3 (1), 28.
- Workplace Learning Report 2020: LinkedIn Learning, 2021. Available from: <https://learning.linkedin.com/resources/workplace-learning-report-2020>. (Accessed 20 November 2021).

Higher education, critical pedagogy, and the challenge of resistance

Henry A. Giroux^a and Ourania Filippakou^b, ^aMcMaster University, Hamilton, ON, Canada; and ^bBrunel University, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

References

143

Across the globe, democratic institutions such as the independent media, schools, the legal system, certain financial institutions, and higher education are under siege. Such repressive forces are at work in Turkey, Hungary, Brazil, and a number of other countries. The promise, if not ideals, of democracy are receding as the barbarians who breathe new life into a fascist past are once again on the move subverting language, values, courage, vision and a critical consciousness. Education has increasingly become a tool of domination as right-wing pedagogical apparatuses controlled by the entrepreneurs of hate attack workers, the poor, people of color, refugees, immigrants from the south and others considered disposable. In the midst of an era when an older social order is crumbling and a new one is struggling to define itself, there emerges a time of confusion, danger, and moments of great restlessness.

In Turkey, dissident intellectuals including academics and journalists have been either imprisoned or expelled (Hansen, 2019). In Brazil, the reactionary President Jair Bolsonaro has launched a full-fledged campaign against the work of Paulo Freire (Woods, 2020). In Hungary, President Orbán shut down the liberal Central European University (Walker, 2018). President Xi Jinping has crushed dissent in Hong Kong silenced many in China's vibrant intellectual circles. What has become clear is that the present moment is once again at a historical juncture in which the structures of liberation and authoritarianism are vying for shaping a future that appears to be either an unthinkable nightmare or a realizable dream.

We have arrived at a moment in which two worlds are colliding and a history of the present is poised at a point when "possibilities are either realized or rejected but never disappear completely" (Thompson, 2013). First, there are the harsh and crumbling worlds of neoliberal globalization and its mobilizing passions that fuel different strands of fascism across the globe, touching countries on almost all continents (Hall, 2012; Harvey, 2005; Wolin, 2008; Brown, 2015; Eubanks, 2017; Monbiot, 2017; Giroux, 2018). Power is now enamored of amassing profits and capital and is increasingly addicted to a politics of white nationalism and racial cleansing. Second, there are growing counter movements, especially among young people, with their search for a new politics that can rethink, reclaim and reinvent a new understanding of democratic socialism, untainted by capitalism (Derber, 2017). Within the last few years massive protests have taken place among young people in the United States, Chile, Peru, and South Africa. In many countries such as South Africa, Tunisia, and the United States these protests have been met with massive violence. What is not in doubt is that something sinister and horrifying is happening to liberal democracies all over the globe. The global thrust toward democratization that emerged after World War II is giving way once again to authoritarian tyrannies. As alarming as the signs may be, the public cannot look away and allow the terrors of the unforeseen to be given free rein. For those who believe in democratic socialism we cannot allow the power of dreams and militant hopes to turn into ashes.

We now live in a world that resembles a dystopian novel. The COVID-19 crisis created a surrealist nightmare that floods our screens and media with images of fear. People can no longer shake hands, embrace our friends, use public transportation, sit in a coffee shop, or walk down the street without experiencing anxiety and fear. While vaccines have been developed to assuage such fears, the unequal distribution of such vaccines between the developed and undeveloped nations suggests that it will be some time before the virus is brought under control. Under such circumstances, what must be acknowledged is that the pandemic is more than a medical concept. It also refers to ideological and political plagues that emerged as a result of the irresponsible response of the US and other countries such as Brazil, the United Kingdom, and India to the Covid-19 crisis. Surprisingly, in spite of Africa not having adequate supplies of the vaccine and resources to fight the disease, a catastrophic outbreak has not yet materialized as badly as expected. Marked by inept leadership rooted in a distrust of science and reason and a blind allegiance to market forces, what emerged over time was unimaginable suffering, massive deaths, and a further legitimization of lies and right-wing violence. The horror of the pandemic often blinds us to the fact that anti-democratic economic and political forces that have prioritized profits over human needs have grinded away at the social order for the last 40 years.

A form of predatory capitalism has waged war on the welfare state, public sphere and the common good since the 1970s. As a form of predatory capitalism, neoliberalism believes that the market should govern the economy and all aspects of society. It concentrates wealth in the hands of a financial elite and elevates untrammelled self-interest, self-help, deregulation, and privatization as the governing principles of society. Under neoliberalism, everything is for sale and the only obligation of citizenship is consumerism. At the same time, it ignores basic human needs such as healthcare, food, decent wages, and quality education. Neoliberalism views government as the enemy of the market, limits society to the realm of the family and individuals, embraces a fixed hedonism, and challenges the very idea of the public good.

We live in an age when economic activity is divorced from social costs, while policies that produce racial cleansing, militarism, and staggering inequality have become defining features of everyday life and established modes of governance especially in the United States. Clearly, there is a need to reclaim a notion of democratic socialism in which matters of justice, equity, and equality become the central features of a substantive democracy. The good news is that the demonstrations taking place across the globe suggest that the spirit of democratic socialism is in the air. Yet, the Russian invasion of the Ukraine has undermined such hopes.

The pandemic revealed in all its ugliness, the death producing mechanisms of systemic inequality, deregulation, a culture of cruelty, and an increasingly dangerous assault on the environment. It has also made visible an anti-intellectual culture that derides any notion of critical education, that is an education that equips individuals to think critically, engage in thoughtful dialog, appropriate the lessons of history, and learn how to govern rather than be governed. At the same time, the claims of neoliberal capitalism have been undermined as a result of the economic failures and medical horrors let loose by the pandemic. What was once unthinkable is now said in public by demonstrators across the globe are protesting police violence, climate devastation, and the brutality of economic inequality. Young people are reaching out across borders to address their counterparts in Nigeria, Sweden, the United States, Brazil, and other countries. They are calling for a new narrative to repair the safety net, provide free healthcare, childcare, elder care and free quality public schools for everyone. There are loud calls to address state violence, and the plagues of poverty, homelessness, and the pollution of the planet.

The pandemic is a crisis that cannot be allowed to turn into a catastrophe in which all hope is lost. While this pandemic threatens democracy's ability to breathe, it offers the possibility to rethink politics and the habits of critical education, human agency, and elements of social responsibility crucial to any viable notion of what life would be like in a democratic socialist society. Matters of inequality, globalization, and climate change are objects of massive protests in counties that include but are not limited to Africa, the UK, Brazil, Uganda, Pakistan, and Ecuador. At the current moment, Russian brutality in Ukraine is the subject of mass protests across the globe.

Put differently, amid the corpses produced by neoliberal capitalism and COVID-19, there are flashes of hope, a chance to move beyond contemporary resurgences of authoritarianism. Paulo Freire understood that such a politics is rooted in a pedagogy of hope, one that integrated a critical reading of the world with an attempt to put into practice modes of struggle based on the principles of social and economic equality and human freedom.

It is hard to imagine a more urgent moment for taking seriously Paulo Freire's ongoing attempts to make education central to politics. At stake for Freire was the notion that education was a social concept, one rooted the goal of emancipation for all people. Moreover, this is an education that encourages human agency, one that is not content to enable people to only be critical thinkers, but also engaged individuals and social agents. This is a pedagogy that calls us beyond ourselves, and engages the ethical imperative to care for others, dismantle structures of domination, and to become subjects rather than objects of history, politics, and power. If we are going to develop a politics capable of awakening our critical, imaginative, and historical sensibilities, it is crucial for educators and others to remember Freire's ongoing project of literacy. Freire's work builds upon the work of many other theorists that include but are not limited to Antonio Gramsci, Raymond Williams, Nelson Mandela, and Frantz Fanon. Moreover, the influence of Freire's work was not only obvious in the United States, UK, and Brazil but also South Africa, especially given its influence in the Black Consciousness Movement often aligned with Steve Biko. This was particularly obvious given his early work in several former Portuguese colonies such as Guinea-Bissau and Mozambique. In the 1970s and 80s, Freire work was often associated with the United Democratic Front in South Africa. At the present moment, Freire's work is taken up in a number of ongoing projects that extend from Geneva, Switzerland and Berlin to Portugal and the University of KwaZulu-Natal, South Africa.

What many of these theorists shared was a political project in which civic literacy infused with a language of critique and possibility addressed the notion that there is no democracy without knowledgeable and civically literate citizens. Such a language is necessary to enable the conditions to forge a collective international resistance among educators, youth, artists, and other cultural workers in defense of public goods. This was especially true with regards to the role of literacy and consciousness raising in shaping trade unions and grassroot struggles. Such a movement is important to resist and overcome the tyrannical fascist nightmares and white supremacists movements that have descended upon the United States, Brazil, India, South Africa, and a number of other countries. In an age of social isolation, information overflow, a culture of immediacy, consumer glut, and spectacularized violence, it is all the more crucial to take seriously the notion that a democracy cannot exist or be defended without civically literate and critically engaged citizens and the institutions that make it possible.

Education both in its symbolic and institutional forms has a central role to play in fighting the resurgence of fascist cultures, mythic historical narratives, and the emerging ideologies of white supremacy and white nationalism. Moreover, as fascists across the globe are disseminating toxic racist and ultra-nationalist images of the past, it is essential to reclaim education as a form of historical consciousness and moral witnessing. This is especially true at a time when historical and social amnesia have become a national pastime, particularly in the United States, matched only by the masculinization of the public sphere and the increasing normalization of a fascist politics that thrives on ignorance, fear, the suppression of dissent, and hate. Education as a form of cultural work extends far beyond the classroom and its pedagogical influence, though often imperceptible, is crucial to challenging and resisting the rise of fascist pedagogical formations and their rehabilitation of fascist principles and ideas (Mayer, 2019).

The pedagogical lesson here is that fascism begins with hateful words, the demonization of others considered disposable, and moves to an attack on ideas, the burning of books, the disappearance of intellectuals, and the emergence of the carceral state and the horrors of detention jails and camps. As a form of cultural politics, critical pedagogy provides the promise of a protected space within which to think against the grain of received opinion, a space to question and challenge, to imagine the world from different standpoints and perspectives, to reflect upon ourselves in relation to others and, in so doing to understand what it means to "assume a sense of political and social responsibility" (Nixon, 2015).

Cultural politics in the last twenty years has turned toxic as ruling elites across the globe increasingly gain control of commanding cultural apparatuses turning them into pedagogical disimagination machines that serve the forces of ethical tranquilization by producing and legitimating endless degrading and humiliating images of refugees, immigrants, Muslims, and others considered excess, or wasted lives doomed to terminal exclusion. The capitalist dream machine is back with huge profits for the ultra-rich, hedge fund managers, and major players in the financial service industries. In these new landscapes of wealth, fraud, and social atomization, a savage and fanatical capitalism promotes a winner-take-all ethos, normalizes massive inequalities in wealth and power, and aggressively undermines the welfare state while pushing millions into hardship and misery. The geographies of moral and political decadence have become the organizing standard of the dream worlds of consumption, privatization, surveillance, and deregulation. Within this increasingly fascist landscape, public spheres are replaced by zones of social abandonment and thrive on the energies of the walking dead and avatars of cruelty and misery.

The assault on education has been particularly evident in the West and the emphasis we give the West in this paper is not one that clings to a notion of ethnocentrism or indulges a colonial legacy. The emphasis here is on a West is crucial because it operates within a powerful colonial framework that still imposes its ideologies, educational visions, and notions of politics on the rest of the world. Education within the last three decades has diminished rapidly in its capacities to educate young people and others to be reflective, critical, and socially engaged agents. Under neoliberal regimes, the utopian possibilities formerly associated with public and higher education as a public good capable of promoting social equality and supporting democracy have become too dangerous for the apostles of authoritarianism. Increasingly public schools are subject to the toxic forces of privatization and mindless standardized curricula while teachers are deskilled and subject to intolerable labor conditions. Higher education now mimics a business culture run by a managerial army of bureaucrats, enamored of market values, who resemble the high priests of a deadening instrumental rationality. The commanding visions of democracy are in exile at all levels of education. In many states books are being banned, critical race theory is under attack, and a war is being waged on LGBTQ youth.

Critical thought and the imaginings of a better world present a direct threat to neoliberal rationality in which the future must always replicate the present in an endless circle in which capital and the identities that it legitimates merge with each other into what might be called a dead zone of the imagination and pedagogies of repression. This dystopian impulse thrives on producing myriad forms of inequality and violence—encompassing both the symbolic and the structural—as part of a broader attempt to define education in purely instrumental, privatized, and anti-intellectual terms. What is clear is that neoliberal modes of education attempt to mold students in the market driven mantras of self-interest, harsh competition, unchecked individualism, and the ethos of consumerism. Young people are now told to invest in their careers, pack their resumes, and achieve success at any cost. It is precisely this replacement of educated hope with an aggressive dystopian neoliberal project and cultural politics that now characterizes the current assault on public and higher education in various parts of the globe. Under neoliberalism, the mantra of privatization, deregulation, the whitewashing of history, and the destruction of the public good is matched by a toxic merging of inequality, greed, and the nativist language of borders, walls, and camps.

It is crucial for educators to remember that language is not simply an instrument of fear, violence, and intimidation, it is also a vehicle for critique, civic courage, resistance, and engaged and informed agency. We live at a time when the language of democracy has been pillaged, stripped of its promises and hopes. For instance, authoritarians such as Jair Bolsonaro in Brazil, Viktor Orbán in Hungary, President Abdel Fattah al-Sisi in Egypt, and President Abdelmadjid Tebboune in Algeria, the degradation of language reinforces Umberto Eco's remark that education is an organizing principle of fascism. According to Eco, one of the central features of what he called "Ur-Fascism" was its undermining of civic literacy "through Fascist schoolbooks [that] made use of an impoverished vocabulary, and an elementary syntax, in order to limit the instruments for complex and critical reasoning" (Eco, 1995).

If fascism is to be defeated, there is a need to make education an organizing principle of politics and, in part, this can be done with a language that exposes and unravels falsehoods, systems of oppression, and corrupt relations of power while making clear that an alternative future is possible. Hannah Arendt was right in arguing that language is crucial in highlighting the often hidden "crystallized elements" that make fascism likely (Arendt, 2001). Language is a powerful tool in the search for truth and the condemnation of falsehoods and injustices. Moreover, it is through language that the history of fascism can be remembered and the lessons of the conditions that created the plague of genocide can provide the recognition that fascism does not reside solely in the past and that its traces are always dormant, even in the strongest democracies. Paul Gilroy argues correctly that it is crucial in the current historical moment to re-engage with fascism in order to address how it has crystallized in different forms and in doing so "work toward redeeming the term from its trivialization and restoring it to a proper place in discussions of the moral and political limits of what is acceptable" (Paul, 2000).

Gilroy provides one more reason for educators to make the political more pedagogical and the pedagogical more political. The latter is crucial in order to recognize, as Freire reminds us, that pedagogy is always a struggle over agency, identities, desires, and values while also acknowledging that it has a crucial role to play in addressing important social issues and in defending public and higher education as democratic public spheres. Making the political pedagogical in this instance suggests producing modes of knowledge and social practices that not only affirm oppositional cultural work and pedagogical practices but also offer opportunities to mobilize instances of collective outrage coupled with direct mass action, against a ruthless casino capitalism and an emerging fascist politics. Such mobilization must oppose the glaring material inequities and the growing cynical belief that democracy and capitalism are synonymous. At the very least, critical pedagogy proposes that education is a form of political intervention in the world and that it is capable of creating the possibilities for individual and social transformation.

Ignorance now rules many countries across the globe. Not the simple, if somewhat innocent ignorance that comes from an absence of knowledge, but a malicious ignorance forged in the arrogance of refusing to think hard about an issue, to engage language in the pursuit of justice. James Baldwin was certainly right in issuing the stern warning in *No Name in the Street* that

"Ignorance, allied with power, is the most ferocious enemy justice can have." Thinking is now viewed as an act of stupidity, and thoughtlessness is considered a virtue. All traces of critical thought appear only at the margins of the culture, as ignorance becomes the primary organizing principle of American society. As is well known, former President Trump's ignorance was on display daily. Not only is he a serial liar but his ignorance also serves as a tool of power to prevent power from being held accountable. In addition, ignorance is the enemy of critical thinking, engaged intellectuals, and emancipatory forms of education. Ignorance is not innocent, especially when it proclaims the space of commonsense and labels thinking dangerous while exhibiting a disdain for truth, scientific evidence, and rational judgments. However, there is more at stake here than the production of a toxic form of illiteracy celebrated as commonsense, the normalization of fake news, and the shrinking of political horizons. There is also the closing of the horizons of the political coupled with explicit expressions of cruelty and a "widely sanctioned ruthlessness" (Mishra, 2018).

The very conditions that enable people to make informed decisions are under siege as schools are defunded, media becomes more corporatized, oppositional journalists are killed, and reality TV becomes the model for mass entertainment. We now live in a new age of cruelty in which we are told that the central mark of our agency is to be at war with others, unleash our most ruthless and competitive side, and learn how to survive in what Naomi Klein calls the "cut-throat jungle of late capitalism."

Under such circumstances, there is a full-scale attack on thoughtful reasoning, empathy, collective resistance, and the compassionate imagination. In some ways, the dictatorship of ignorance resembles what the writer John Berger calls "ethicide": and Joshua Sperling defines as "The blunting of the senses; the hollowing out of language; the erasure of connection with the past, the dead, place, the land, the soil; possibly, too, the erasure even of certain emotions, whether pity, compassion, consoling, mourning or hoping" (Sperling and Appignanesi, 2019). Words such as love, trust, freedom, responsibility, and choice have been deformed by a market logic that narrows their meaning to either a relationship to a commodity or a reductive notion of self-interest. Freedom now means removing oneself from any sense of social responsibility so one can retreat into privatized orbits of self-indulgence. And so it goes. The new form of illiteracy does not simply constitute an absence of learning, ideas, or knowledge. Nor can it be solely attributed to what has been called the "smartphone society" (Aschoff, 2015). On the contrary, it is a willful practice and goal used to actively depoliticize people and make them complicit with the forces that impose misery and suffering upon their lives.

Given the current crisis of politics, agency, history, and memory educators need a new political and pedagogical language for addressing the changing contexts and issues facing a world in which capital draws upon an unprecedented convergence of resources—financial, cultural, political, economic, scientific, military, and technological—to exercise powerful and diverse forms of control. If educators and others are to counter global capitalism's increased ability to separate the traditional sphere of politics from the now transnational reach of power, it is crucial to develop educational approaches that reject a collapse of the distinction between market liberties and civil liberties, a market economy and a market society. Resistance does not begin with reforming capitalism but abolishing it. In this instance, critical pedagogy becomes a political and moral practice in the fight to revive civic literacy, civic culture, and a notion of shared citizenship. Politics loses its emancipatory possibilities if it cannot provide the educational conditions for enabling students and others to think against the grain and where students realize themselves as informed, critical, and engaged citizens. There is no radical politics without a pedagogy capable of awakening consciousness, challenging common sense, and creating modes of analysis in which people discover a moment of recognition that enables them to rethink the conditions that shape their lives. This is the moment of hope in which as Ruth Levitas points out the sense of "something missing can be read in every trace of how it might be otherwise, how the ever-present sense of lack might be [tempered]" (Levitas, 2003).

As a rule, educators should do more than create the conditions for critical thinking and nourishing a sense of hope for their students. They also need to responsibly assume the role of civic educators within broader social contexts and be willing to share their ideas with other educators and the wider public by making use of new media technologies. Communicating to a variety of public audiences suggests using opportunities for writing, public talks, and media interviews offered by the radio, Internet, alternative magazines, and teaching young people and adults in alternative schools to name only a few. Capitalizing on their role as public intellectuals, faculty can speak to more general audiences in a language that is clear, accessible, and rigorous. More importantly, as teachers organize to assert both the importance of their role as citizen-educators and that of education in a democracy, they can forge new alliances and connections to develop social movements that include and expand beyond working with unions.

In the current moment, Freire's legacy is all the more crucial in its call to embrace critical pedagogy as a political and moral practice in order to revive civic literacy, civic culture, and a notion of shared citizenship.

Education operates as a crucial site of power in the modern world. If teachers are truly concerned about safeguarding education, they will have to take seriously how pedagogy functions on local and global levels. Critical pedagogy has an important role to play in both understanding and challenging how power, knowledge, and values are deployed, affirmed, and resisted within and outside of traditional discourses and cultural spheres. In a local context, critical pedagogy becomes an important theoretical tool for understanding the institutional conditions that place constraints on the production of knowledge, learning, academic labor, social relations, and democracy itself. Critical pedagogy also provides a discourse for engaging and challenging the construction of social hierarchies, identities, and ideologies as they traverse local and national borders. In addition, pedagogy as a form of production and critique offers a discourse of possibility—a way of providing students with the opportunity to link understanding to commitment, and social transformation to seeking the greatest possible justice.

This suggests that one of the most serious challenges facing teachers, artists, journalists, writers, and other cultural workers is the task of developing discourses and pedagogical practices that connect a critical reading of both the word and the world in ways that enhance the creative capacities of young people and provide the conditions for them to become critical agents. In taking up this project, educators and others should attempt to create the conditions that give students the opportunity to acquire the knowledge, values, and civic courage that enables them to struggle in order to make desolation and cynicism unconvincing and hope practical.

Needless to say, such a task is related to the different histories, cultures, and politics that define many of the countries still willing to connect education to a form of political, social, and cultural empowerment. What these countries can share is a notion of hope that is removed from the fantasy of an idealism that is unaware of the constraints facing the struggle for a radical democratic society. Educated hope is not a call to overlook the difficult conditions that shape both schools and the larger social order nor is it a blueprint removed from specific contexts and struggles. On the contrary, it is the precondition for imagining a future that does not replicate the nightmares of the present, for not making the present the future.

Educated hope provides the basis for dignifying the labor of teachers; it offers up critical knowledge linked to democratic social change, affirms shared responsibilities, and encourages teachers and students to recognize ambivalence and uncertainty as fundamental dimensions of learning. Such hope offers the possibility of thinking beyond the given. As difficult as this task may seem to educators, if not to a larger public, it is a struggle worth waging.

In an age of predatory capitalism and an emerging fascist politics, educators, students, and other concerned citizens face the challenge of providing a language that embraces a militant utopianism while constantly being attentive to those forces that seek to turn such hope into a new slogan or to punish and dismiss those who dare to look beyond the horizon of the given. Fascism breeds cynicism and is the enemy of a militant and social hope. Hope must be tempered by the complex reality of the times and viewed as a project and condition for providing a sense of collective agency, opposition, political imagination, and engaged participation. Without hope, even in the most dire times, there is no possibility for resistance, dissent, and struggle. Agency is the condition of struggle, and hope is the condition of agency. Hope expands the space of the possible and becomes a way of recognizing and naming the incomplete nature of the present.

Hope is the affective and intellectual precondition for individual and social struggle. Hope, not despair, is the precondition that encourages critique on the part of intellectuals in and outside of the academy who use the resources of theory to address pressing social problems. Hope is also at the root of the civic courage that translates critique into political practice. Hope as the desire for a future that offers more than the present becomes most acute when one's life can no longer be taken for granted. Only by holding on to both critique and hope in such contexts will resistance make concrete the possibility for transforming politics into an ethical space and a public act. Building a better future than the one we now expect to unfold will require nothing less than confronting the flow of everyday experience and the weight of social suffering with the force of individual and collective resistance and the unending project of democratic social transformation. At the same time, in order for resistance to take on the challenges posed by the rise of a fascist politics, it will have to develop an awakening of desire. This form of educated desire is rooted in the dream of a collective consciousness and imagination fueled by the struggle for new forms of community that affirm the value of the social, economic equality, the social contract, and democratic values and social relations.

The current fight against a nascent fascism across the globe is not only a struggle over economic structures or the commanding heights of corporate power. It is also a struggle over visions, ideas, consciousness, and the power to shift the culture itself. It is also as Arendt points out a struggle against "a widespread fear of judging" (Arendt, 2003). Without the ability to judge, it becomes impossible to recover words that have meaning, imagine a future that does not mimic the dark times in which we live, and create a language that changes how we think about ourselves and our relationship to others. Any struggle for a radical democratic socialist order will not take place if "the lessons from our dark past [cannot] be learned and transformed into constructive resolutions" and solutions for struggling for and creating a post-capitalist society (Bertoldi, 2018).

In the end, there is no democracy without informed citizens and no justice without a language critical of injustice. Democracy begins to fail, and political life becomes impoverished in the absence of those vital public spheres such as public and higher education in which civic values, public scholarship, and social engagement allow for a more imaginative grasp of a future that takes seriously the demands of justice, equity, and civic courage. Democracy should be a way of thinking about education, one that thrives on connecting pedagogy to the practice of freedom, learning to ethics, and agency to the imperatives of social responsibility and the public good (Giroux, 2019). Neoliberal capitalism strips hope of its utopian possibilities and thrives on the notion that we live in an era of foreclosed hope, and that any attempt to think otherwise will result in a nightmare. Yet, the fact remains that without hope there is no agency and without collective agents, there is no hope of resistance. In the age of nascent fascism, it is not enough to connect education with the defense of reason, informed judgment, and critical agency; it must also be aligned with the power and potential of collective resistance. We live in dangerous times. Consequently, there is an urgent need for more individuals, institutions and social movements to come together in the belief that the current regimes of tyranny can be resisted, that alternative futures are possible and that acting on these beliefs through collective resistance will make radical change happen.

References

- Arendt, H., 2001. *Origins of Totalitarianism*, new ed. Harcourt Trade Publishers, New York.
- Arendt, H., 2003. Personal responsibility under dictatorship. In: Kohn, J. (Ed.), *Responsibility and Judgement*. Schocken Books, NY. <https://grattoncourses.files.wordpress.com/2016/08/responsibility-under-a-dictatorship-arendt.pdf>.
- Aschoff, N., 2015. The smartphone society. *Jacobin Mag*. <https://www.jacobinmag.com/2015/03/smartphone-usage-technology-aschoff/>.
- Bertoldi, N., 2018. Are We Living through a New "Weimar Era"? Constructive Resolutions for Our Future. *OpenDemocracy*. <https://us1.campaign-archive.com/?e=d77f123300&u=9c663f765f28cdb71116aa9ac&id=367a142d39>.
- Brown, W., 2015. *Undoing the Demos: Neoliberalism's Stealth Revolution*. Zone Books, New York.
- Derber, C., 2017. Welcome to the revolution: universalizing resistance for social justice and democracy in perilous times. New York: Routledge, 2017. In: Geiselberger, H. (Ed.), *The Great Regression*. Polity, London.

- Eco, U., 1995. *Ur-fascism*. N. Y. Rev. Books. <http://www.nybooks.com/articles/1995/06/22/ur-fascism/?pagination=false&printpage=true>.
- Eubanks, V., 2017. *Automating Inequality*. St. Martin's Press.
- Giroux, H.A., 2018. *American Nightmare: Facing the Challenge of Fascism*. City Lights.
- Giroux, H.A., 2019. *The Terror of the Unforeseen*. Los Angeles Review of Books, Los Angeles.
- Hall, S., 2012. The neoliberal revolution. Chapter 1. In: Rutherford, J., Davison, S. (Eds.), *The Neoliberal Crisis*. Lawrence Wishart, London. http://wh.agh.edu.pl/other/materialy/678_2015_04_21_22_04_51_The_Neoliberal_Crisis_Book.pdf.
- Hansen, S., 2019. "The era of people like you is over": how Turkey purged its intellectuals. N. Y. Times Mag. <https://www.nytimes.com/2019/07/24/magazine/the-era-of-people-like-you-is-over-how-turkey-purged-its-intellectuals.html>.
- Harvey, D., 2005. *A Brief History of Neoliberalism*. Oxford University Press, New York.
- Levitas, R., 2003. Introduction: the elusive idea of utopia. *Hist. Hum. Sci.* 16 (1), 4.
- Mayer, J., 2019. The making of the fox news white house. New Yorker. <https://www.newyorker.com/magazine/2019/03/11/the-making-of-the-fox-news-white-house>.
- Mishra, P., 2018. A Gandhian stand against the culture of cruelty. N. Y. Rev. Books. <http://www.nybooks.com/daily/2018/05/22/the-culture-of-cruelty/>.
- Monbiot, G., 2017. *Out of the Wreckage*. Verso Press.
- Nixon, J., 2015. Hannah Arendt: thinking versus evil. *High Educ.* <https://www.timeshighereducation.co.uk/features/hannah-arendt-thinking-versus-evil/2018664.article?page=0%2C0>.
- Paul, G., 2000. Against race: imagining political culture beyond the color line. Chapter 4. In: *Hitler in Khakis: Icons, Propaganda, and Aesthetic Politics*. The Belknap Press of Harvard University Press, Cambridge, pp. 144–145, 146.
- Sperling, J., Appignanesi, L., 2019. Berger's ways of being. N. Y. Rev. Books. https://www.nybooks.com/articles/2019/05/09/john-berger-ways-of-being/?utm_medium=email&utm_campaign=NYP%20Tintoretto%20Berger%20Mueller&utm_content=NYP%20Tintoretto%20Berger%20Mueller+CID_22999ee4b377a478a5ed6d4ef5021162&utm_source=Newsletter&utm_term=John%20Bergers%20Ways%20of%20Being.
- Thompson, P., 2013. The Frankfurt School, Part 5: Walter Benjamin, fascism and the future. *Guardian*. <https://www.theguardian.com/commentisfree/belief/2013/apr/22/frankfurt-school-walter-benjamin-fascism-future>.
- Walker, S., 2018. 'Dark day for freedom': soros-affiliated university quits Hungary. *Guardian*. <https://www.theguardian.com/world/2018/dec/03/dark-day-freedom-george-soros-affiliated-central-european-university-quits-hungary>.
- Wolin, S.S., 2008. *Democracy Incorporated: Managed Democracy and the Specter of Inverted Totalitarianism*. Princeton University Press.
- Woods, A., 2020. Why is the Brazilian Right Afraid of Paulo Freire? *OpenDemocracy*. <https://www.opendemocracy.net/en/democraciaabierta/why-is-the-brazilian-right-afraid-of-paulo-freire/>.

About the rise, the characteristics and future of New Public Management in Higher Education

Bruno Broucker, KU Leuven, Public Governance Institute, Leuven, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	145
New Public Management and its rise within Higher Education	145
New Public Management and Higher Education	146
Concerns that stimulated Higher Education reform	146
The operationalization of New Public Management in Higher Education	146
An evaluation of NPM within the Higher Education sector	147
What about the future?	148
The role and limitations of theories for HE policy	148
Maintaining benefits from NPM, contextualizing policy and broadening the perspective	149
Some conclusions	150
References	150

Introduction

In many OECD countries Higher Education (HE) was able to attract significant investment increases over the last 20 years. This is largely due to a higher enrollment in Higher Education, increasing cost and priorities set by government and policymakers to invest in skills, research and innovation (OECD, 2020). This conclusion, made by the OECD (2020), demonstrates not only the recognized importance that Higher Education has within society, but also the expectations that are attributed to the sector, and to the overall management and governance of Higher Education systems and institutions. This managerial approach, i.e., that Higher Education should deliver socio-economic outcomes in relation to the incomes the sector has, has become more important since the 1980s, under the influence of New Public Management. Since then, it can not be denied that New Public Management has largely shaped the way Higher Education is managed and perceived nowadays. Indeed, when going through the literature on Higher Education management and governance, or on Higher Education (HE) reform, there is no doubt that the concept of New Public Management (NPM) has been enormously influential within the Higher Education sector. It has not only been a multi-layered concept used by researches or academics to describe the reality, but it has definitely been a practice—and an explicit policy—that has been applied within the sector. That concept is at the core of this article.

The purpose of this article is to provide a brief, yet comprehensive insight of the concept of New Public Management in general, and what it means within the sector of Higher Education in particular. In a first paragraph the rise of the concept within the Higher Education sector will be discussed. In a second paragraph the characteristics that NPM, as a conceptual “basket”, has, will be explained. In the final part an overall evaluation of NPM will be provided, and some insight into how Higher Education might develop over the next decades.

New Public Management and its rise within Higher Education

The concept of New Public Management has not been invented for the Higher Education sector only. In fact it is an umbrella concept to describe the gradual wave of reforms that were introduced since the 1970s in public sector organizations and administrations across the globe. Those public sector reforms were triggered by the then identified flaws and weaknesses of the public sector as a whole and traditional bureaucracy in particular. The economic, social and political context of many countries, characterized by financial crises, an overall increasing level of distrust vis-à-vis the public sector, the perceived heaviness and slowness of bureaucracy and administrative procedures, and the burden of red tape, created for many governments the momentum to introduce major public sector reforms (Pollitt et al., 2007). Those public sector reforms were said to be necessary to increase the efficiency and effectiveness of public organizations, to enhance their overall performance, to improve the quality of their services, and to better meet the expectations and needs of citizens. Looking backwards, it can be stated that this was a quite revolutionary way to look at the public sector: from the 1970s onwards it was increasingly perceived as a sector that needed to become more client-oriented, customer-friendly and efficiently working.

A key characteristic of this wave of reforms was the inspiration it took from the private sector. Despite different, yet ideal typical, organizational targets—private organizations aiming for profit vs. public organizations aiming to contribute to the public good—there was a general belief that private sector practices could improve the functioning of the public sector, by copying and applying managerial techniques that were used in the private sector: (1) input/output control mechanisms; (2) performance management,

measurement and associated audits; (3) a focus on the cost management of public organizations; (4) marketization and competition with the use of contracts for resource allocation and service delivery within public services; (5) a preference for smaller and more result oriented organizations; (6) the presence of hands-on management and the organizational distance between policy implementation and policy making; (7) entrepreneurial leadership; (8) strategic human resource management (Ferlie et al., 2008; Gruening, 2001; Osborne, 2006; Pollitt and Bouckaert, 2000; Hood, 1995; Dunleavy et al., 2005; Broucker and De Wit, 2015; Scott, 2015). Since the start of this reform wave, many public sector organizations and administrations have been modernizing according to (some of) those principles, and it has been investigated enormously with a huge amount of scholarly work as result. One of the conclusions of this research is that the basic idea of the reforms was not only highly comparable, but that it also created a kind of uniformity in managerial practice and public sector policy (Broucker and De Wit, 2015).

The list of private sector techniques demonstrates that it is no wonder that underneath the conceptual umbrella of NPM various management styles were introduced (Hood, 1995). Broucker and De Wit (2015) stated that NPM was a “toolbox” wherefrom governments seem to choose instruments as they find them appropriate or necessary (Broucker and De Wit, 2015). In this sense, the implementation of NPM has taken many forms and has been implemented with various levels of intensity and at different periods (Pollitt and Bouckaert, 2000). This was also argued by Pollitt et al. (2007, p. 4) stating that within Europe, the implementation of NPM can be compared to a “chameleon, constantly changing its appearance to blend with the local context”. And even in 1995 it was already assumed that NPM would take numerous forms within the OECD (Hood, 1995).

Given the magnitude of the impact that NPM had within the public sector, it was obvious that it would also impact the Higher Education sector—even though it would sometimes be described differently: NPM has within the HE sector been adopted under the term of the “neoliberal university”, though other names have been used as well, such as “entrepreneurial”, “enterprise”, and “innovation” university (Marginson, 2009; Peters, 2013; Scott, 2015; Frances, 2015; Ehrenstorfer et al., 2015).

New Public Management and Higher Education

Concerns that stimulated Higher Education reform

The Higher Education sector is not an isolated sector within society. The sector influences, directly or indirectly, many aspects of it: it contributes on the skills development of individuals, it fuels the labor market, it supports the increase of social welfare, it provides scientific advice to policymakers in many different policy fields. That perspective already makes it obvious that the Higher Education sector could not remain unaffected by the New Public Management wave that took place within the overall public sector. But of course, reforming for the sake of reforming is rarely done. The Higher Education sector has also been affected by many concerns, pushing policymakers to introduce reform plans in Higher Education. Bleiklie (2018) summarizes the concerns that have triggered reforms over the last decades as follows: (1) the massification of Higher Education in terms of student numbers and number of institutions; (2) the increasing complexity, expenses, and visibility of HE systems; (3) the increasing importance that is attributed to HE and research for economic growth and social welfare (Bleiklie, 2018).

The operationalization of New Public Management in Higher Education

In a time when budgetary restrictions were a challenge, reforms that could help the increase of the productivity, efficiency, and relevance of academic activities were extremely important, which is why they were launched and progressively implemented since the 1980s (Bleiklie, 2018; Amaral, 2008). The Higher Education sector, wherein universities and university colleges were traditionally managed, was suddenly challenged and influenced by NPM to be steered differently. Similar NPM-instruments as for the overall public sector were introduced within the Higher Education sector by HE policy makers, such as an increased level of marketization with the aim to increase competition and excellence, state budget cuts, institutional reforms (De Vries and Nemec, 2013; Scott, 2015). Peters (2013) states that, because of the wave of neoliberalism (conceptually closely linked to NPM), universities had to diversify their funding resources, herewith increasing competition (Enders et al., 2011; Marginson, 2014) and transforming universities into “corporate actors” (De Boer et al., 2008; Marginson, 2009). It is within this context that some universities (with the Anglo-Saxon universities often cited as example) have been charging students with larger tuition fees, resulting in large debts after graduation (Frances, 2015; Scott, 2015). Due to performance funding, universities have also attempted to expand their national audience by attracting international students, or by establishing campuses in other countries (Wilkins and Huisman, 2012). This trend has been intensified by the numerous ranking systems that stimulate competition, mainly between research-intensive universities (Pritchard et al., 2015; Scott, 2015). Universities aimed also to expand knowledge production by creating research parks, establishing public-private co-operation for research, and commercialization of the results of research and education (Marginson, 2009; Scott, 2015). The expansion of universities went hand-in-hand with the growth of the university administration and management, resulting in more bureaucratization (Marginson, 2009; Broucker and De Wit, 2015). In a more recent article, Bleiklie (2018) summarizes five major reform trends that have been implemented in Higher Education, under the influence of the NPM-ideology: (1) the development of a systemic integration whereby not only universities fall under the HE sector, but all institutions contributing to tertiary education; (2) a change in decision making model whereby the management of universities has moved much more toward a business-like model with a top-down chain of command; (3) a fundamental shift in funding mechanisms whereby HEIs are increasingly dependent on output-funding mechanisms and external resources; (4) an increased importance

attributed to evaluations and quality assurance mechanisms; (5) a fundamental shift in work organization where researchers aren't working individually but in transdisciplinary teams and/research groups.

More concretely, several authors have tried to summarize the characteristics NPM would have within HE. [Marginson \(2009\)](#) emphasizes corporatization reform, growth in student fee-charging, an expansion of the role of private institutions, encouragement of commercial business activity in research, the creation of competition for parcels of government-provided resources, and output modeling.

The OECD accentuates leadership principles, incentives, and competition between public sector agencies and private entities to enhance the outcomes and cost efficiency of public services ([Hénard and Mitterle, 2006](#)).

[Bleiklie and Michelsen \(2013\)](#) stress hierarchization (leadership and management), budgetary constraints, the formalization of evaluation, and the increased autonomy for institutions.

[Ferlie et al. \(2008\)](#) arrive at a list of ten characteristics, including market-based reforms such as the stimulation of competition for students and funding and the encouragement of private sector providers; the development of real prices for student fees and research contracts; the development of audit and checking systems; and vertical steering with stronger and more overt managerialism.

From this overview, it is clear that NPM—under a comparable logic—contains a large diversity of elements. This implies that in the actual implementation of reforms not every aspect needs to be implemented to the same extent to be able to call it an NPM reform ([de Boer et al., 2008](#)). As stated above, the same was true for the overall NPM-based public sector reforms. [Broucker et al. \(2016\)](#) compared the NPM-reforms of the Higher Education systems of 10 countries, ranging from early adopters to mid-adopters of NPM, and from varying welfare systems and different geographical regions. The authors came to the conclusion that all of them could be classified under the NPM-umbrella, despite differences related to the level of marketization, budgetary reforms, levels of organizational autonomy and management/leadership style. The authors also came to the conclusion that substantial differences between the HE systems could be identified, despite the comparable logic. NPM was indeed used as a set of instruments that a government has at its disposal and from which it can choose which instruments it will use, at different times, with a different focus, taking into account topical policy goals ([Broucker et al., 2016](#)). From that perspective the authors labeled the NPM implementation in Higher Education as widespread, but uneven. This was also confirmed some years later by a comparative study of 20 years Bologna Reform within the European Higher Education Area that stated that New Public Management has largely influenced Higher Education System reform throughout Europe, despite the many differences between systems—and sometimes even within systems ([Broucker et al., 2019](#)).

An evaluation of NPM within the Higher Education sector

In general, NPM has been criticized because of its intra-governmental focus in an increasingly pluralist world ([Osborne, 2006](#)), and because it has led to unwanted side-effects as fragmentation or diminished coordination ([Hammerschmid et al., 2013](#)). This criticism can to some extent be summarized to one issue: the fact that NPM was said to be too much focused on efficiency, effectiveness and economy. In other words: the public sector also deals with other purposes than the narrow focus that was attributed to the NPM-logic. The same criticism can be found in the HE literature. As [Marginson \(2014, p. 31\)](#) states, “the one-sided fixation with market competition [...] has obscured the rich potential [of HE] for public good.” Indeed, it was said that many social drives of HE are losing their resonance in the rush to make HEIs accountable and responsive ([Singh, 2001](#)).

Over the years, and when the impact of NPM-based reforms in HE became clearer, other scholars have confirmed that NPM has not always been to the advantage of HE. [Giroux \(2010\)](#) for instance, concludes that HE has been subjected to “a corporate-based ideology that embraces standardizing the curriculum, supporting top-down management, and reducing all levels of education to job-training sites” and he pleads instead for “shared governance between faculty and administrators, educating students as critical citizens rather than potential employees of Walmart, or affirming faculty as scholars and public intellectuals ...” ([Giroux, 2010, p. 185](#)). He proposes a system that has a legitimate place in society as a public good and that protects the equality of educational opportunities. Other examples of criticism include [Carvalho \(2015\)](#) account of the reduced autonomy of universities and grade inflation in private schools in Portugal, and [Gaethgens \(2015\)](#) assessment of the German “Excellence Initiative” that has delivered benefits for only a few HEIs. [Lim and Lund \(2016\)](#) question the persisting role of individual performance for academic careers. [Frances \(2015\)](#) criticizes the high tuition fees in the USA, which is confirmed for the UK by [Strike \(2015\)](#) where students are now facing large debts (see also [Marginson, 2014](#)). [De Boer et al. \(2008\)](#) point at “the competitive pressure” in the UK and its negative consequences. [Shattock \(2008, p. 191\)](#) summarizes the UK HE policy as follows, “tighten controls, budget cuts, freeze new appointments, campaign against waste (value of money), squeeze the system”. [Sutton \(2017\)](#) states that the university has, because of its neo-liberal focus, lost its moral norm of education as a public good.

[Broucker et al. \(2018\)](#) have stated that the criticism on NPM in HE runs parallel to the criticisms on NPM in the larger public sector. They summarize the concerns as follow: first, within Higher Education there is an increasing demand for more partnership and trust, less competition and pressure and a less narrow focus on performance alone. Second, little is known about the outcomes of NPM reforms and, though the authors above highlight to a certain extent their criteria to assess the success rate of a reform, the evaluation criteria remain vague, somehow normative and lacks objective data. One of the reasons is that it is difficult to prove that HE has become cheaper, more efficient or effective because of the NPM reform. Of course, this is also due to the nature of NPM itself: conceptually NPM is very broad and in practice NPM has been implemented very differently, which make it hard to measure in its totality. The [OECD \(2020\)](#), while also referring to the potential benefits of NPM, largely confirmed this conclusion (p. 11–12): “Public authorities in OECD member and partner countries regularly need to make and justify decisions about how to mobilize,

allocate and manage financial and human resources in Higher Education. This can be challenging, as not only are the effects of individual policy choices difficult to predict and prove, but the core objectives of Higher Education—[...] are hard to measure [...]. Governments have often sought to reshape Higher Education systems through promoting expansion and diversification of Higher Education provision or the concentration of Higher Education activities [...] in a smaller number of institutions to build critical mass, promote excellence or achieve efficiency savings. The development of new institutional and program types, as well as the expansion of private sector provision and online learning have facilitated large-scale widening of access to Higher Education in many systems. In contrast, concentration policies, including complementary specialization, institutional collaboration, alliances and mergers, have often been pursued with the aim of increasing quality and lowering costs. The success of these policies is challenging to measure.”

The fact that it is hard to prove that NPM might not have delivered the expected outcomes, makes it equally hard to develop something new for the HE sector. Broucker et al. (2018) stated that to overcome this challenge, HE as a whole should be placed back within its broader context. To study that viewpoint, Broucker et al. (2021) have compared the Higher Education systems of Flanders, Italy and Finland to understand what the effects were of the implementation of (comparable) NPM-based reforms, and why the effects were met or not. They concluded that the NPM-model was indeed dominant in those systems, despite the fact that none of the cases introduced it in its extreme ideal-typical form. From their comparative overview the authors identified similarity between HE governance instruments used, policy expectations and rationale of the (NPM-based) reform. Yet, they did see quite some differences between the effects of similar policies in different contexts. One of the conclusions of the research was that—although policy goals are generally well-defined by governments -, the goal (un-)clarity used during the reform and the (lack of) support were potentially detrimental for the implementation and the impact of the introduced policy. This conclusion clarified the lack of desired effects from the NPM-reforms, but also the difficulty in assessing the real impact of the use of NPM-related instruments within Higher Education.

What about the future?

The role and limitations of theories for HE policy

On the one hand New Public Management has probably been the dominant policy and reform model for Higher Education over almost half a century. But on the other hand the criticism is clear: too much emphasis on an economic rationale, too much pressure on output, a risk in deducing academic freedom which is deemed to be necessary for innovation, and so on. The first question therefore is: why is it still then so present within Higher Education? Tight (2019) summarizes it as follows. First, Higher Education is not an island within society. Nations worldwide have adopted neoliberalism within their overall economic and social policies. It is no wonder that it has affected Higher Education, and that Higher Education cannot stand apart from this. The contrary would not only be strange to witness, it would also be counter-productive for the general policy. Indeed, as stated above already, Higher Education contributes to social welfare and the labor market. Policies developed for those two policy fields can not be completely detached from Higher Education. It is likely to say that a neoliberal policy in key domains will influence other policy domains as well.

Second, it seems as if policymakers don't see a valid alternative to neoliberalism and New Public Management. This acceptance by the persons who are responsible for Higher Education results into an adoption of NPM-principles across the sector. Moreover, as stated above, as long as it remains hard to evaluate thoroughly the effects, it will be hard to develop an alternative—because the question about whether an alternative would be necessary remains unanswered.

Third, the large number of characteristics of NPM in general, and the whole range of connected ideas and initiatives, have a multiplying effect on the Higher Education sector. It has such a large impact on Higher Education in general, and on its stakeholders (researchers, students, administrative staff, ...) that it cannot just simply be reversed. On the contrary, staff for instance will remain expected to be excellent in what it does, and herewith it confirms the principles that have influenced Higher Education over the last decades.

The second question to be addressed is: how can Higher Education maintain the potential benefits it has gained from New Public Management, and simultaneously address its identified weaknesses? Because this is what policy should do and normally does: iteratively improve what can be improved. Within the literature a large number of alternatives can be found already. In their article, Broucker et al. (2017) have discussed those NPM-related theories. The authors identified three large categories of theories. A first group of theories that are in fact an application of NPM (i.e., (new) managerialism and academic capitalism). A second group of theories that propose adaptations to the NPM-model (network governance, new public governance and public value): those theories address the issues that NPM has been blamed to neglect, such as collaboration and interdependency. They start from the observation that HE can at best be a quasi-market, that government steering remains strong, and that universities, the state, and other actors external to the university are interdependent and connected in a network structure. And a third group of theories that are seen as a reaction to NPM (shared governance, flexible governance). Those theories take a step away from the NPM framework and managerialism, and focus on the internal working of universities. They stress the interdependence and need for cooperation between the internal stakeholders. Broucker et al. (2017) conclude that theories in this cluster are a way of rethinking NPM, and re-emphasizing collaboration within institutions as an important factor, be it for an *effective* and an *efficient* policy.

The complexity of those theories is that even within each category of theories different applications can be found and that between categories gray zones exist. Moreover, it is also no wonder that when it comes to theories related to Higher Education that they are linked to ideology, political ideas about the role of HE in society, or normative impressions about the labor market, the welfare state, or others. This is what Tight (2019) referred to when stating that Higher Education cannot be detached from

overall policy and its implications in the economic and social context. Despite the usefulness or the inspiration other theoretical models might bring, their real value are equally hard to assess. Therefore the authors state in another article that it is crucial to look at the overall context of a Higher Education System to be able to fully understand what can be considered as a flaw and what not; similarly that context will better highlight whether NPM as a model is useful or not, and how large the number of identified flaws is (Broucker et al., 2018). Doing so, Higher Education will much more benefit from theories and models.

Maintaining benefits from NPM, contextualizing policy and broadening the perspective

It is the role that HE is expected to play within society that should be at the core of the debate when talking about HE policy and benefits. What do stakeholders want from Higher Education? This role will be differently defined according to the ideology that guides the opinion of stakeholders (Broucker et al., 2018). The discussion about the benefits of HE should be equally broad. Next to efficiency and performance (NPM-values), other benefits and aims should be considered, by looking at all objectives defined by the relevant stakeholders, so that can be assessed whether reforms are aimed at building a HE that is functional for the achievement of those objectives (Broucker et al., 2018). When discussing NPM, several important nuances need to be made.

First, it's not all black and white: when Higher Education is built in a country that puts large emphasis on equal opportunity, it is logic that the entrance barrier to Higher Education will be kept as low as possible. But it also implies that there might be a large number of students failing in the first year, because they all have the opportunity to try it out. From a NPM perspective one can argue that this is inefficient. But from the ideology of democratization it is logic that this inefficiency is seen as necessary, and even expected. The question is whether on the longer run society will be better off. And if so, then one can ask whether that is still inefficient or not? In other words: what seems to be ineffective on the short term, can become effective on the longer term. A lot depends on the perspective one looks at a system. The contrary can also be witnessed: in a country that puts liberalism as a model much more upfront, it is logic that the individual will be much more responsible in entering Higher Education. This can result into a situation where less students are able to enter Higher Education, with a lower level of study failure in the first year, and thus with a higher level of efficiency. But then the question is what that effect will be on the labor market and social welfare in general? Of course, reality is much more complex than what those two examples seem to indicate: Higher Education systems that put democratization high on the agenda quite often build rules to avoid that students repeatedly fail in Higher Education, because otherwise the financial burden would become too heavy. And vice versa: Higher Education systems that are built as elite systems build in measures to attract bright students who do not have the necessary financial means to study. This means that in both cases NPM-measures can be applied, to a different degree, but at the same time counterbalanced by other measures to take out the potential sharp edges of the consequences of an ideal-typical implementation. How much is a theory than worth if it is only exceptionally witnessed in its "pure" form in reality? How much "NPM" has then been implemented—and how can NPM then be assessed, or even compared?

Second, it's not only about the policy or the instruments. It's about the context and the customization of principles and policy instruments to that context. Howlett (2019) came to a similar conclusion when stating that researchers should pay more attention to the appropriateness and preconditions of policy instruments rather than on the implemented instruments and their goals. Here-with the argument is that policy implementation needs to be put again into the policy process. Thus: within the context. Broucker et al. (2021) highlighted in their research that emphasis should be put on the implementation *feasibility* of each governance instrument, instead on the instrument as such. In other words: studying NPM, how it has been implemented and which characteristics were used, is fine; but studying under which circumstances it has been introduced in a system might provide more insight into its usefulness and into the reasons why it has been successful or not. Indeed, this will increase insight into successful governance approaches, and allow policy makers to better take into account organizational path dependency and contextual/cultural characteristics. The way NPM- (or other-) instruments have been adapted to the specific political, social and economic context and the organizational/national values and culture are in this respect more crucial for successful implementation than the choice of certain instruments for specific expectations. And this will much better fuel or evaluation of NPM. Obviously, it needs to be acknowledged that this customization already happens to some extent: it was already highlighted that only very few countries adopted the NPM-governance model in its most extreme form, though customization needs to be improved.

Third, nothing is fixed in stones. Reform often takes quite some time, and over time principles might change and managerialist approaches might also shift over time. This implies that at different stages, over time, the definition of how a governance model is understood, can differ. The difference between policy formulation and policy implementation is also crucial to take into account. Important is also to acknowledge that governments tend to first define strategic goals and then to implement the respective reforms for the achievement of those goals. Broucker et al. (2021) highlighted that *along the way* goals do not necessarily change, but policy instruments do. Governance mechanisms are then adapted to increase reform acceptancy, and to leave out potential sharp edges of certain (in this case market-oriented) reforms. Hence, it would make more sense to first determine established, supported and unambiguous goals—in order to in a second stage define an adapted (mixture) of governance instruments that would allow the achievement of those goals. This helps to understand which governance techniques work under which circumstances, and which not. Policy implementation studies could benefit to further elaborate on this. Another issue is that the role of HE in society is not value-free, and research in that domain always risks to reflect certain preferences regarding that role in society.

Fourth, there is no need to throw away the baby out with the bathwater. As the NPM approach has a neoliberal ideological background, it is tempting, even for scholars, to be for or against that ideology and approach. But it has to be acknowledged that the dominant NPM approach has had not only negative but also positive effects. Even when arguing that efficiency and effectiveness are too narrow as a goal for HE, it cannot be denied that in many ways efficiency and effectiveness have been increased and this has

been to the benefit of HE (Broucker et al., 2017). Broucker et al. (2017) have therefore proposed a public value model that does not replace NPM for Higher Education, but that puts HE back into the arena of dialog between stakeholders—and where other than economic principles are present. A holistic view on Higher Education and its role in society, and a better insight into the different logics that are applied to Higher Education will provide both researchers and policymakers a better understanding of what Higher Education is, and what it should do. It will also provide a better idea of the needed adaptations, and the created results. And finally, it will avoid to narrow the debate to the implementation and criticism on NPM. Instead, it will help to rediscuss and reposition the role of Higher Education within society and consider NPM—or any other theory driven policy instruments—as a mean to that explicit end.

Some conclusions

There are several conclusions to be made from this article. First of all, New Public Management has become a dominant model for the public sector and the Higher Education sector. Its managerialist approach has pervaded Higher Education systems and institutions across the globe, and it has infiltrated in every aspect of Higher Education (Shepherd, 2018).

Second, it has been studied extensively in many different ways: in literature an enormous amount of case studies about NPM-implementation can be found, and the same goes for research reports on the way NPM has characterized Higher Education in its funding, performance management or other organizational features. Likewise, the outcomes of NPM have been discussed, though they are not easy to assess (see also OECD, 2020), the criticism on the model is known, and alternative theories have been developed. Though NPM remains an important organizational model, it is said to be replaced by New Public Governance (Shepherd, 2018)—though Broucker et al. (2018) have highlighted that this model is merely an adaptation of NPM, and not a replacement. This also implies that NPM will not just simply disappear; it has influenced other approaches, but several basis principles will remain—be it in a less “extreme” form.

Third, and linked to the previous: the basic principles of NPM are not disputed. As already stated: most Higher Education systems want to be efficient—as this is to the benefit of society. But what is disputed or criticized is its sometimes narrow focus, be it on the organization of HE, be it on its objectives, be it on how it has been implemented. It cannot be neglected that many countries have followed the NPM route, as there probably is no valid alternative up until now that is capable to fully counter the identified deficiencies of NPM: either alternative theories do not integrate the importance of other values and other stakeholders; either they minimize the importance of external partners and their influence; either the theories elaborate on NPM but are hard to operationalize. This highlights that NPM as a concept is in itself perhaps too vague, even though this can also be seen as a strength: HE systems are free to use it as they see fit. Moreover, none of the alternative theories are against efficiency or effectiveness. Most theories are against the fact that these would be the only objectives, but at the same time, even NPM is not against collaboration or interdependency, as long as it is for the benefit of efficiency.

Fourth, what cannot be neglected is that NPM *in practice* is the operationalization of an ideology into a policy. There is no doubt that the criticism first starts with what went wrong in the development or implementation of the policy, and that the critique is then in a second stage transferred to the ideology itself. But the fundamental issue however remains: the way HE policy is described and studied runs the risk of being either one-sided, or conceptually (too) complex. What is needed is a more comprehensive theory or a multi-dimension theory that not only take into account the ins and outs of Higher Education, the way it should be organized to achieve the stipulated results, or the role of management in Higher Education institutions. What is needed, for both researchers and policymakers, is a more detailed picture about Higher Education, and a model that is both perspicuous and able to catch the complexity of HE policy in itself, its role within society, the values it is built on, and the stakeholders it involves. NPM was in the beginning an answer to a range of problems that were identified in many countries, and what is needed now is to take the debate further, i.e., to rediscuss the specific customization of principles to the societal context, just because contexts change and new challenges emerge. And in order to be able to do it is necessary to reflect on the role of Higher Education in society nowadays. That question could put the debate on HE back into its social, political and economic context, herewith broadening the debate on what HE should do (efficiency vs. socio-economic development)? The empirical investigation of the role of HE in society could provide inspiration for the theoretical approach toward HE. In a second stage, more international, multidisciplinary and comprehensive research is needed on the impact that NPM-driven HE reforms had. The answer on that question would provide insight into the effects of concepts and theories when they are used as an ideological basis for reforms, and thus provide insight into their value for HE research and policy. A focus on the effects of past reforms and the role of HE in society could open up the debate on what governance of HE for the next decades could be, while respecting the policy options of the past, and could be the ground for integrating the different theoretical paradigms into a coherent framework, allowing to grasp HE to its full extent.

References

- Amaral, A., 2008. Transforming higher education. In: Amaral, A., Bleiklie, I., Musselin, C. (Eds.), *From Governance to Identity*. Festschrift for Mary Henkel. Springer, Dordrecht, pp. 81–94.
- Bleiklie, I., Michelsen, S., 2013. Comparing HE policies in Europe. Structures and reform outputs in eight countries. *High Educ.* 65, 113–133.
- Bleiklie, I., 2018. New public management or neoliberalism, higher education. In: *Encyclopedia of International Higher Education Systems and Institutions*, pp. 1–6.

- Broucker, B., De Wit, K., 2015. New public management in higher education. In: Huisman, J., de Boer, H., Dill, D.D., Souto-Otero, M. (Eds.), *Handbook of Higher Education Policy*. Springer, pp. 57–74.
- Broucker, B., De Wit, K., Leisyte, L., 2016. Higher education system reform: a systematic comparison of ten countries from a new public management perspective. Chapter 2. In: Pritchard, R., Pausits, A., Williams, J. (Eds.), *Positioning Higher Education Institutions. From Here to There*. Sense Publishers, ISBN 978-94-6300-659-0, pp. 19–40. https://doi.org/10.1007/978-94-6300-660-6_2.
- Broucker, B., De Wit, K., Verhoeven, J., 2017. Higher education research: looking beyond new public management. Chapter 2. In: Huisman, J., Tight, M. (Eds.), *Theory and Method in Higher Education Research*. Emerald, Bingley, ISBN 978-1-78743-223-9, pp. 21–38.
- Broucker, B., De Wit, K., Verhoeven, J., 2018. Higher education for public value: taking the debate beyond new public management. *High Educ. Res. Dev.* 37 (2). <https://doi.org/10.1080/07294360.2017.1370441>.
- Broucker, B., De Wit, K., Verhoeven, J., Leisyte, L. (Eds.), 2019. *Higher Education System Reform. An International Comparison after Twenty Years of Bologna*. Brill, Leiden, ISBN 978-90-04-40010-8.
- Broucker, B., Ursin, J., Dal Molin, M., De Wit, K., 2021. To Achieve the Same as the Others? Policy Preconditions for Successful Higher Education Governance. *Int. J. Public Adm.* 1–10.
- Carvalho, L., 2015. Merit and student selection: views of academics at the University of Porto. In: Pritchard, R.M., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education. Can the Challenges Be Reconciled?* Sense Publishers, Rotterdam/Taipei, pp. 91–108.
- De Boer, H.F., Enders, J., Schimank, U., 2008. Comparing higher education governance systems in four European countries. In: Soguel, N.C., Jaccard, P. (Eds.), *Governance and Performance of Education Systems*. Springer, Dordrecht, pp. 35–54.
- De Vries, M., Nemec, J., 2013. Public sector reform: an overview of recent literature and research on NPM and alternative paths. *Int. J. Public Sect. Manag.* 26 (1), 4–16.
- Dunleavy, P., Margetts, H., Bastow, S., Tinkler, J., 2005. New public management is dead—long live digital-era governance. *J. Publ. Adm. Res. Theor.* 16, 467–494.
- Ehrenstorfer, B., Sterrer, S., Preymann, S., Aichinger, R., Gaisch, M., 2015. Multi-tasking talents? Roles and competences of middle-level manager-academics at two Austrian higher education institutions. In: Pritchard, R.M., Teichler, U., Klumpp, M. (Eds.), *Diversity and Excellence in Higher Education*. Sense Publishers, Rotterdam/Boston/Taipei, pp. 175–200.
- Enders, J., De Boer, H., File, J., Jongbloed, B., Westerheijden, D., 2011. Reform of higher education in Europe. In: Enders, J., De Boer, H.F., Westerheijden, D.F. (Eds.), *Reform of Higher Education in Europe*. Sense Publishers, Rotterdam/Boston/Taipei, pp. 1–10.
- Ferlie, E., Musselin, C., Andresani, G., 2008. The steering of higher education systems—a public management perspective. *High Educ.* 56 (3), 325–348.
- Frances, C., 2015. The dangerous role of economics in shaping American higher education policy: the disasters they are causing that Europeans should hasten to avoid. In: Pritchard, R.M., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education. Can the Challenges Be Reconciled?* Sense Publishers, Rotterdam/Boston/Taipei, pp. 53–70.
- Gaethgens, C., 2015. Does size matter? The example of the “Excellence Initiative” and its impact on smaller universities in Germany. In: Pritchard, R.M., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education. Can the Challenges Be Reconciled?* Sense Publishers, Rotterdam/Boston/Taipei, pp. 19–30.
- Giroux, H.A., 2010. Bare pedagogy and the scourge of neoliberalism: rethinking higher education as a democratic public sphere. *Educ. Forum* 74 (3), 184–196.
- Gruening, G., 2001. Origin and theoretical basis of new public management. *Int. Publ. Manag. J.* 4, 1–25.
- Hammerschmid, G., Van de Walle, S., Oprisor, A., Stimac, V., 2013. Trends and Impact of Public Administration Reforms in Europe: Views and Experiences from Senior Public Sector Executives. European Policy Brief. European Commission, COCOPS, Brussels. Retrieved from: <http://www.cocops.eu/archives/1345>.
- Hénard, F., Mitterle, A., 2006. Governance and Quality Guidelines in Higher Education. A Review on Governance Arrangements and Quality Assurance Guidelines. OECD, Paris.
- Hood, C., 1995. The “new public management” in the 1980s: variations on a theme. *Account. Organ. Soc.* 20 (2/3), 93–109.
- Howlett, M., 2019. Moving policy implementation theory forward: a multiple streams/critical juncture approach. *Publ. Pol. Adm.* 34 (4), 405–430.
- Lim, M.A., Lund, R., 2016. Passion, Robustness and Perseverance. *Higher Education Research and Development*.
- Marginson, S., 2009. The Limits of Market Reform in Higher Education. Simon Marginson Hiroshima University RIHE 17 August 2009. Research Institute for Higher Education (RIHE), Hiroshima University, Japan, Hiroshima, p. 18. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.664.8458>.
- Marginson, S., 2014. Higher Education as a public good in a marketized East Asian environment. In: Yonezawa, A., Kitamura, Y., Meerman, A., Kuroda, K. (Eds.), *Emerging International Dimensions in East Asian Higher Education*. Springer Science+Business Media, Dordrecht, pp. 15–33. https://doi.org/10.1007/978-94-017-8822-9_2.
- OECD, 2020. *Resourcing Higher Education: Challenges, Choices and Consequences*, Higher Education. OECD Publishing, Paris. <https://doi.org/10.1787/735e1f44-en>.
- Osborne, S., 2006. The new public governance? *Publ. Manag. Rev.* 8 (3), 377–387.
- Peters, M.A., 2013. Managerialism and the neoliberal university: prospects for new forms of “open management” in higher education. *Contemp. Read. Law Soc. Justice* 5 (1), 11–26. Retrieved from: <http://heinonline.org>.
- Pollitt, C., Bouckaert, G., 2000. *Public Management Reform: A Comparative Analysis*. Oxford University Press, Oxford.
- Pollitt, C., Van Thiel, S., Homburg, V., 2007. “New Public Management in Europe”, *Management Online Review*. <http://www.morexpertise.com/view.php?id=78>. (Accessed 27 January 2015).
- Pritchard, R.M., Klumpp, M., Teichler, U., 2015. The challenges of diversity and excellence. In: Pritchard, R.M., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education. Can the Challenges Be Reconciled?* Sense Publishers, Rotterdam/Boston/Taipei, pp. IX–XXI.
- Scott, P., 2015. Markets and managerialism. Enhancing diversity or promoting conformity. In: Pritchard, R.O., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education*. Sense Publishers, Rotterdam/Boston/Taipei, pp. 3–17.
- Shattock, M., 2008. The change from private to public governance of British higher education: its consequences for higher education policy making 1980–2006. *High Educ. Q.* 62 (3), 181–203.
- Shepherd, S., 2018. Managerialism: an ideal type. *Stud. High Educ.* 43 (9), 1668–1678.
- Singh, M., 2001. Reinserting the “public good” into higher education transformation. *KAGISANO High. Educ. Discuss.* 1, 8–18.
- Strike, T., 2015. Improving access to postgraduate study in England. In: Pritchard, R.M., Klumpp, M., Teichler, U. (Eds.), *Diversity and Excellence in Higher Education. Can the Challenges Be Reconciled?* Sense Publishers, Rotterdam/Boston/Taipei, pp. 133–152.
- Sutton, P., 2017. Lost souls? The demoralization of academic labor in the measured university. *High Educ. Res. Dev.* 36 (3), 625–636.
- Tight, M., 2019. The neoliberal turn in higher education. *High Educ. Q.* 73 (3), 273–284.
- Wilkins, S., Huisman, J., 2012. The international branch campus as transnational strategy in higher education. *High Educ.* 64, 627–645.

Governance, institutional autonomy and leadership in Portuguese higher education

António Magalhães^a and Amélia Veiga^{a,b}, ^a Centre for Research on Higher Education Policies (CIPES), Faculty of Psychology and Education Sciences, University of Porto, Matosinhos, Portugal; and ^b Centre for Research and Intervention in Education (CIE), Faculty of Psychology and Education Sciences, University of Porto, Matosinhos, Portugal

© 2023 Elsevier Ltd. All rights reserved.

Introduction	152
Changing state regulation	152
Higher education governance reform in Portugal	154
Institutional governance and management structures	155
Reframing institutional leadership	156
Conclusion	156
Acknowledgment	157
References	157

Introduction

To understand governance and institutional autonomy in Portuguese higher education, the regulation impact of the changing governance models is of key importance. The Portuguese higher education system is binary comprising universities and polytechnics, both public and private. Higher education institutions are under the supervision of the Ministry of Science, Technology and Higher Education. Universities aim at guaranteeing scientific, cultural and technical education, and create knowledge, develop innovation, critical and professional competences. Polytechnic institutions focus on applied research and vocational and advanced technological training. Since their inception, they were expected to respond to regional needs and lifelong learning demands. In Portugal, there are 14 public universities and 27 polytechnics, and 23 private universities and 50 polytechnics.

New Public Management (NPM) drove Portuguese higher education governance reforms, however elements stemming from other governance approaches (e.g., Network Governance and New Governance) also emerged mainly as a reaction to the implementation of NPM reforms. The regulation efforts developed at the national level echo a trend toward meta-governance (Peters, 2010) that is also visible at the institutional level. Higher education institutions and their leadership need to ensure that their organizational objectives are to be attained in a context where actors and governance and management bodies enjoy pedagogic and scientific autonomy.

Institutional autonomy was widely assumed as a regulation instrument to politically steer public systems and institutions, aiming at implementing governance models that might adequately respond to competitive and ever-changing environments. The consensus about awarding autonomy to higher education institutions and the need to increase self-governance creates the need that governments make sure that their goals are actually taken into account, enhancing the instrumentality of regulation mechanisms. While there is a wide consensus about the attribution of autonomy to higher education institutions and the need to increase self-governance, the pursuit of public goals as represented by the governments leads to the reinforcement of their regulation frames. Similarly, institutional leaderships need to guarantee that the goals and strategies set out at the central level are actually pursued at other organizational levels, regardless of the fact that they are aligned horizontally or vertically in the decision-making processes (Magalhães et al., 2012).

The Portuguese case illustrates how governance reforms impact on institutions and how these reforms correspond to meta-governance efforts in the aftermath of the introduction in 2007 of a new legal framework. Common organizational behaviors directed toward competitiveness, accountability and quality promoted by the dissemination of a common grammar on funding and quality issues were induced (Magalhães et al., 2012).

In the first part of the article, we argue that the prevalence of governance, both as a concept and political practice, over governing corresponds to the reconfiguration of Portuguese state regulation in the field of higher education. This had implications on autonomy and on the framework within which leadership evolves at the institutional level. In the second part, the focus is on the impact of governance reforms on internal governance and management structures and on the reframing of institutional leadership.

Changing state regulation

The state control model implied a cleavage between governing and governance. The former was specially directed at tracing social objectives and goals, and the latter focused on managing the processes and the instruments for their achievement. State officials and

bureaucrats, as professionals of administration, were in charge of translating governmental objectives into norms and regulations, and applying them (Peters, 2001, 2010). Under this model, governance appeared, at least formally, as an articulated set of human and material resources, and organizational structures to serve governments' objectives and goals. In this sense, and in spite of the "technicality" and specialization of the professionals in charge of governance activities, there was supremacy of government over governance. In the 1980s, across Europe the model of state control was replaced by the model of state supervision in which the central government administration downsized its interference in the daily decisions of institutions. In line with this, institutional autonomy enhanced the self-regulatory capacities of the system and of higher education institutions to promote the public good.

The fragmentation of decision-making processes in public affairs and the increasing influence of non-state actors brought to the fore the issue of how to politically steer autonomous and potentially self-governing organizations. The meta-governance perspectives (Peters, 2010) underline the need to ensure the fulfillment of the policy objectives traced at the central level. The literature on regulatory governance (Baldwin and Cave, 1999) has already been looking at the interaction between the regulator's discretion and institutions' autonomy. Lodge (2003, p. 159) has underlined constraints inherent to the regulatory frames and "the institutional factors that structure relationships between the policy domain and its environments". Actually, the absence of formal authority and legitimacy, central to traditional settings of government, started being dealt with governance structures and processes that cannot be reduced to a mere delegation of competencies.

In Portugal, the influence of New Public Management (NPM) on the 2007 Legal Framework for Higher Education Institutions (RJIES) (Portugal Ministério da Ciência e do Ensino Superior, 2007a) reshaped the relationships between governing, governance and management. Although there is not a consensual definition of NPM, it represents the idea that governments should maintain and preserve only a minimum of public services while emphasizing the need to attain the objectives and the results aimed at. Efficiency and effectiveness are to be achieved through management instruments used in the private sector, specifying goals, emphasizing competition for clients, performance measurement, de-centralization of decision-making and the use of markets to regulate public services (Meek, 2003).

Detailed control instruments have been set up "over and above that control and oversight traditionally exercised through legislative and legal procedures" (Neave, 2008, p. 10). Neave argues that these instruments were the main driver for the reforms undertaken in higher education. Under the influence of NPM, in the last decades the narratives on the enhancement of institutional self-regulation have reinterpreted the meaning of institutional autonomy in the field of public policies. This trend promoted the technicality of the governance and management instruments enhancing the managerial structures and processes within European higher education institutions (Bleiklie and Kogan, 2007), Portugal included.

A tension emerged between the reconfiguration of state regulation and the instruments stemming from it with impact on the relationship between governance and management. Governance is about political management of rule systems, both formal and informal, that drive values and norms affecting actors and constellations of actors' behaviors and attitudes (Hall and Taylor, 1996; Kjaer, 2010). In turn, management refers to the implementation of the objectives set out by the governance structures on the basis of established rules and it is concerned with the efficiency, effectiveness and quality of services provided to internal and external stakeholders (Eurydice, 2008).

The tendency to subsume governance to management is visible in the Portuguese higher education context (Magalhães et al., 2017). This trend evolved under the fixation of meanings attributed to governing, governance and management. Governing refers to the actions taken by the government to politically steer social systems. In turn, governance, in a narrower sense, refers to the actions taken for the achievement of policy goals in dealing with increasing fragmentation of public decision-making and the degree of interdependence between state and non-state actors of civil society. Referring to this fragmentation, Peters (2010) underlines the need of meta-governance:

I shall be using the term "meta-governance" to describe the process of steering devolved governance processes. In other terms, it is the "governance of governance". The notion of meta-governance is that a number of organisations and processes within the public sector have attained a substantial degree of autonomy — a condition better described as governance — and that there may be a need to impose some control over those components of governing

(Peters, 2010, p. 37).

The regulation efforts undertaken at the European and national levels reflect a trend toward coordination of devolved governance processes by enhancing the State regulation frames and meta-governance efforts. Research has underlined that when higher education institutions are forced to compete among themselves, particularly in a context of financial stringency, they tend to pursue their "own good" rather than the "common good" as represented by the governments' objectives and goals (Amaral and Magalhães, 2007). The need to coordinate organizational subsystems and actors also raises the issue of meta-governance at the institutional level, making more complex the relationship between governance and management structures and actors.

Autonomy, accountability and quality assessment became keywords in the last decades of policy-making in higher education (Magalhães, 2004) and its governance arrangements. Institutional autonomy is supposed to increase organizational awareness about changes in the environment and is assumed as a fundamental political instrument in market-driven regulation contexts. The organizational awareness of external transformations is particularly substantiated in the presence of external stakeholders (Magalhães and Amaral, 2000) in the governance structures of higher education institutions. The concept of representation of external interests in higher education in the governance of institutions simultaneously transforms internal actors into stakeholders.

Awarding autonomy reshapes the relationship between state and higher education institutions. In political terms, this is visible on a contractual relationship “governed” by the state that does not weaken or diminishes in extension and intensity its regulation. Neave (2008, p. 10) argues:

One of the most significant achievements — indeed, an indispensable condition for re-defining the relationship between Nation and university these two decades past, has been the setting up of a powerful, relatively precise and detailed instrumentality over and above that control and oversight traditionally exercised through legislative and legal procedures.

Neave also argues that this increased instrumentality in higher education was the main driver for the reforms undertaken in the sector. The transformation of the regulation relationships replaced a priori control by a posteriori control referred to institutions’ outputs. This induced the elaboration of instruments based on performance indicators that spread all over Europe (e.g. financing while incentive, in Denmark and financing referred to performance in Sweden). In this context, collegial governance is being challenged by the enhancement of managerial governance and reconfiguring the meaning of institutional leadership, as illustrated by the Portuguese case.

Higher education governance reform in Portugal

Since the mid-1990s, debates in Portugal about institutional autonomy have highlighted a tension between institutional autonomy granted to higher education institutions by the 1988 Autonomy Act and the need for stronger regulation by governments, and more efficient decision-making structures and processes.

The implementation of the governance reforms of Portuguese higher education has been developed under the influence of NPM resulting in the loss of collegial governance and enhancement of managerial self-regulation. The NPM perspective is visible in discourses stimulating competition for students; in the hardening of budgetary constraints; in the vertical steering of the system by setting explicit targets and using performance contracts; in the market-based research funding; in the development of a “management must manage” perspective; in the emphasis on stronger managerial roles of rectors, deans, heads of department; in the focus on efficiency and value for money; in the development of strong central administrations and in the weakening of the representation of faculty in higher education management (Paradeise et al., 2009). Managerialism has been driving the changes incorporated in the new legislation (Magalhães and Santiago, 2012). Among the political drivers of these reforms, the most relevant are the move to appointed rather than elected posts; weakening in the representation of the constituencies of higher education institutions; stronger managerial roles of rectors, deans and heads of departments and private-sector style of human resources management (Amaral et al., 2013). These drivers resulted in an increasing centralization of power at the institutional top, suppression or weakening of the collegial decision-making bodies and the strengthening of the presence of external stakeholders in the institutions’ governance structures.

The need for meta-governance of the higher education system and institutions to monitor institutional performance against policy objectives has become evident. In examining the impact of the implementation of the legal framework for higher education, established in 2007, previous research (Magalhães et al., 2013) has revealed that the New Governance perspective allows a broader interpretation of the governance reform outcomes. From this perspective, Salamon (2002) identifies the governance approach to public issues by pointing out characteristics such as: (a) focus on governance techniques rather than on state programmes; (b) focus on networks rather than on hierarchy; (c) shift from public versus private to public plus private; (d) shift from command and control to negotiation and persuasion; (e) shift from management skills to enablement skills. In the Portuguese context, the shift from management to enablement skills have the potential of identifying specific features in the development of the process of reframing institutional leadership in the Portuguese context.

By 2009, Portuguese higher education institutions had already adapted their statutes to RJIES (2007). As already mentioned, existing research reveals that this legal framework strengthened organizational rationales such as empowering managerial bodies to the detriment of collegial bodies, centralization of decision-making processes and the presence of external stakeholders at the central and faculty/department/school level (Amaral et al., 2013). This legal framework also offered the possibility to adopt a foundational model enabling institutions to rule themselves by private law. This model enhances output-based contracts, accountability, individual responsibility and performance. Five universities and one polytechnic have joined this model and others may be pursuing the same purpose.

Institutional autonomy allows higher education institutions to declare their mission, define their statutes, manage their resources and carry out their activities within a self-regulation rationale, which did not substantially change under RJIES. However, the legal framework for the development of quality in higher education, based on the compulsory accreditation of higher education institutions and of degree programmes reduced to some extent pedagogic autonomy. RJIES reinforced the notion that the autonomy of higher education institutions does not preclude either regulation by the Ministry, or accreditation and external evaluation under the responsibility of the National Agency for Quality Assurance and Accreditation (A3ES). The legal framework for quality evaluation and accreditation passed in 2007 (Portugal Ministério da Ciência e do Ensino Superior, 2007b) determined that higher education institutions should develop a quality assurance policy for their programmes, a culture of quality, quality assurance mechanisms and a strategy for their continuous improvement.

Next, we will look at the internal governance and management structures resulting from the interpretation of RJIES by higher education institutions.

Institutional governance and management structures

Under the influence of NPM, collegiality as reflecting social capital (Dill, 1995) and empowering academics and their networks is decreasing in favor of the prioritization and development of control mechanisms by institutional leadership. The RJIES gives higher education institutions some leeway to create diverse [optional] organizational structures depending on individual institutional contexts. For instance, institutions may opt for being a public institute or a public foundation, the latter seen as a stronger expression of privatization. On the one hand, they may counteract a bolder emergence of NPM by creating optional structures that promote the participation of higher education constituencies in institutional governance. Therefore, the ways higher education institutions implemented RJIES may also be diverse in combining NPM with elements stemming from other governance perspectives. In fact, other governance perspectives have been developed, interacting with NPM, to “make sense of the rise of horizontal rearrangements within higher education systems, in relation with the emergence of new actors in new arenas” (Paradeise et al., 2009, p. 243). This is the case of New Governance approaches that emphasize a shift from hierarchy to networks, from command and control to negotiation, and from management skills to enablement skills (Salamon, 2002).

In 2008 and 2009, Portuguese higher education institutions adapted their organizational and governance structures to RJIES. According to the legal framework, higher education institutions are free to draw up their own governance bodies, but it specifies the main governance structures — general board, rector (or president) and management council. In the case of universities, the law gives leeway for the possibility of adding advisory bodies. The senate, which was an academic structure under the Law of Autonomy (1988) with decision-making capacity, either no longer exists or became an advisory body that should be consulted by the rector in terms to be specified in the statutes of individual institutions. The law makes compulsory the participation of internal and external stakeholders in the general board. In the foundations, the governance bodies are the board of trustees and the auditor. The government appoints the board of trustees, which is composed of external stakeholders recognized as highly qualified and with relevant professional experience, following the proposal of the general board of the higher education institution.

The role and power of academics and students in the governance bodies have changed challenging collegial governance. The centralization of power at the top governance bodies and reduced membership affected the participation of academic staff, students and non-academic staff members. The participation of external stakeholders representing 30% of board members is potentially influential in contributing to the strategy and performance in the core business of higher education institutions. This change is noticeable, for instance, in the limit set by law of 25 members total in the membership of scientific or pedagogic councils, irrespective of the size of higher education institutions. To mitigate this effect and to counterbalance the underrepresentation of faculties/schools/departments in the governance bodies, some higher education institutions have created advisory bodies that can be seen as elements of a focus on the enhancement of academic networks (Magalhães et al., 2018). From the New Governance perspective, building academic advisory bodies can be seen as a guarantee that academics could be involved in decision-making processes.

The rector/president is elected by the general council, but directors at the middle management level are elected by a collegial board (optional by law, though) based on individual action programmes. This means that while the legitimacy of the rector/president among academics is weakening, at the middle management level, the empowerment of directors potentially increases. In other words, decision-making takes place at the various levels but the rector/president reports to the general council and not directly to the academic community.

Higher education institutions have room of maneuver to develop their own statutes, including the definition of the structures at the middle management level when the sub-units (Faculties/Schools/Departments) have self-governing bodies. This brings to the fore the role of institutional leadership in determining the internal governance structures of their institutions. Curiously enough, at the middle management level collegial governance has, to a certain extent, been maintained as the institutions’ board of representatives elects the director of the Faculty/School. The tension between the elements of NPM and those stemming from collegial governance is made visible in the redistribution of academic power in decision-making processes. Actually, research has been showing different weights of managerialism in the governance processes and structures substantiating the relationship between hard and soft managerialism (Magalhães et al., 2017).

Centralization of power at the top of higher education institutions may not be as effective as it might be expected by central administration as decision-making processes are fragmented across the central administration and middle management levels. This calls for meta-governance efforts influenced by top-down managerial perspectives. A hierarchical relationship is established between the governance bodies at middle management level and the governance bodies at the central level. This relationship is aimed at guaranteeing the alignment of strategies, accountability and the intervention of central-level governance bodies, thus signaling institutional top-down control.

The RJIES deviates from a legitimacy principle based on traditional democratic elections but without imposing a pure top-down system of appointments. However, in spite of the increase of the nomination as constitution principle of the governing bodies, election remained not favoring a clear top-down command chain. Additionally, as RJIES determines that “The General Board members do not represent groups or sector interests and are independent in the performance of their duties”, elected professors, researchers and students, after election, no longer represent the interests of those who elected them, thus counteracting democratic principles.

The mixed character of the selection procedures of some members of the governance bodies, together with the basis of democratic legitimacy of elected individuals, is not conducive to see either command-and-control or negotiation-and-persuasion prevalence. This might create governance problems as legitimacy is fragmented between representativeness and the hierarchical line of decision-making and implementation. For instance, and as already mentioned, in some higher education institutions a tension can be found as directors, at the middle-management level, are elected by a collegial board. Directors or presidents of faculties/

schools report to the council of representatives or to the faculty council, which might create difficulties in developing strategic action plans approved at the faculty level as they might not fit into the institutions plan (Magalhães et al., 2013).

The mixed character of RJES allows for the development of horizontal power relations and addresses an important dimension underlined by the New Governance approach. However, collaboration and negotiation will depend on the characteristics of institutional leadership. This raises the question of management skills versus enablement skills.

Reframing institutional leadership

In Portugal, the RJES differentiated the purposes of universities and polytechnics and institutional autonomy gives centrality to the role of the general boards. They approve medium-term strategic plans presented by the rector/president, his/her action plan for four years, the scientific and pedagogic guidelines and the budget proposal. The influence of external stakeholders might reinforce the local and national strategic priorities easing the establishment of active partnerships with the society (Magalhães et al., 2013).

An important aspect to be underlined within meta-governance efforts at the institutional level relates to the need to specify operational objectives, the relevant activities to fulfill those objectives and scorecard instruments to measure the degree of their achievement. The emphasis on institutional meta-governance stresses the enhanced capacity for institutional self-adjustment and potentially accelerates the dynamic of system change from within. Managerial perspectives influence the development of these instruments stemming from the need to demonstrate results to external stakeholders, such as the ministry, and those coming from the corporation and business world. The adoption of a management jargon such as “strategic maps”, “indicators”, “metrics”, “action plans” and “monitorization” is being disseminated as a translation of what higher education institutions are about, increasing the instrumentality of their mission statements.

While contracts may improve the capacity to act strategically they also potentially reduce the autonomy of higher education institutions reflecting the paradox of this type of regulation (Amaral and Magalhães, 2001). As stressed by Neave (2008), the development of governance instruments to coordinate institutional autonomy appears as a step toward stronger and potentially more intrusive relationship between state and higher education institutions. This is also valid for the foundational institutions that under private law enjoy a broader room of maneuver at administrative and financial levels. The foundational model broadens the possibilities of strategic management, since the government funding is based on a 5-year contract program. In addition, the erosion of collegiality, the promotion of the participation of external stakeholders and the centralization of power at the top of the institution are key to understand the reframing of leadership at the institutional level.

Additionally, the emphasis on management and leadership can also be identified in the Portuguese governing and governance of higher education. Governance and institutional leadership can be pointed out as key factors in Portuguese governance reforms featured by the enhancement of managerial self-regulation governance, decrease of collegial governance and strengthening of the role of external stakeholders in the governance bodies (Veiga et al., 2015). Collegial bodies are being re-configured, in a tension, as stressed by Neave (2012, p. 177), “between the values associated with ‘collegiality’, on the one hand, and ‘managerial’ rigor, on the other”.

From the perspective of New Governance, the shift from management to enablement skills is at the core of the reframing institutional leadership. This shift is required to engage partners arrayed horizontally in networks and to bring in multiple stakeholders together for a common end in a situation of interdependence (Salamon, 2002). While the NPM approach also moved the emphasis from control to monitoring the performance, its central concern remains in the functioning of internal management based on management skills.

Signs of the shift from management skills to enablement skills are to be found in the role of the rector/president that, according to the RJES, he/she is expected to elaborate and submit the strategic plans, the guiding principles of the institution focusing on scientific and pedagogical issues, plans and report plans to the general board for approval. The limitation and decrease of representativeness of all constituencies in the structures governing higher education institutions is more visible in the approval of strategic plans. The general board is expected to approve the strategic plans of each individual faculty/school, asking for reformulation, or even rejecting those plans that do not align with the strategic plan of the higher education institution. Coordination difficulties may not allow for aligning strategic action at middle-management level with the central administration. This creates tensions between structures and processes and it is difficult to see a clear shift to enablement skills.

The relationships between the rector/president and the general board, together with the underrepresentation of the constituencies in the governance bodies, make the analysis of the shift from management to enablement skills also highly complex. Apparently, the command-and-control governance features promoting management skills do not match with the organizational specificities of higher education institutions. The shift must go along with the move from a hierarchic organization to a horizontally networked organization that is still embryonic in the Portuguese context (Magalhães et al., 2018).

Conclusion

In Portugal, the features of NPM were softened in the interpretation of the legal framework by higher education institutions by ingredients stemming from other governance perspectives. This corresponds to a European trend as pointed out by existing research (Paradeise et al., 2009, p. 243).

The Portuguese case also shows an embryonic trend underlining networks rather than hierarchy, being the characteristics of the institutional leadership also shaped by the need and, in some cases, the demand to enhance enablement skills. The multiplication of managerial bodies induces the existence of a wide range of mechanisms that while fragmenting the decision-making processes gives leeway to collegial bodies to make a stance in institutional governance.

Finally, the analysis of the Portuguese governance reforms from the perspective of New Governance opened vistas about new features developed together with NPM as it emphasizes the importance of promoting enablement skills to cope with management challenges on the basis of interdependencies and networks. This also goes beyond the Portuguese context to the degree that enhancement of institutional autonomy is a core issue in European governance reforms.

Acknowledgment

This work was funded by national funds by the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the framework of the project UIDB/00757/2020.

References

- Amaral, A., Magalhães, A., 2001. On markets, autonomy and regulation the Janus head revisited. *High Educ. Pol.* 14 (1), 7–20.
- Amaral, A., Magalhães, A., 2007. Market competition, public good and institutional governance: analyses of Portugal's experience. *High Educ. Manag. Pol.* 19 (1), 63–76.
- Amaral, A., Tavares, O., Santos, C., 2013. Higher education reform in Portugal: a historical and comparative perspective of the new legal framework for public universities. *High Educ. Pol.* 26, 5–24.
- Baldwin, R., Cave, M., 1999. *Understanding Regulation*. Oxford University Press, Oxford.
- Bleiklie, I., Kogan, M., 2007. Organization and governance of universities. *High Educ. Pol.* 20 (4), 477–494.
- Dill, D., 1995. Through Deming's eyes: a cross-national analysis of quality assurance policies in higher education. *Qual. High Educ.* 1 (2), 95–110.
- Eurydice, 2008. *Higher Education Governance in Europe: Policies, Structures, Funding and Academic Staff*. Eurydice European Unit, Belgium.
- Hall, P., Taylor, R., 1996. Political science and the three institutionalisms. *Polit. Stud.* 44 (5), 936–957.
- Kjaer, A., 2010. *Governance*. Polity, Cambridge.
- Magalhães, A., Amaral, A., 2000. Portuguese higher education and the imaginary friend: the stakeholders role in institutional governance. *Eur. J. Educ.* 35 (4), 437–446.
- Magalhães, A., Santiago, R., 2012. Governance, public management and administration of higher education in Portugal. In: Neave, G., Amaral, A. (Eds.), *Higher Education in Portugal 1974–2009: A Nation, a Generation*. Springer, Dordrecht, pp. 227–248.
- Magalhães, A., Veiga, A., Ribeiro, F.M., Sousa, S., Santiago, R., 2012. Creating a common grammar for European higher education governance. *High Educ.* 1–18.
- Magalhães, A., Veiga, A., Amaral, A., Sousa, S., e Ribeiro, F., 2013. Governance of governance in higher education: practices and lessons drawn from the Portuguese case. *High Educ. Q.* 67 (3), 295–311.
- Magalhães, A., Veiga, A., e Videira, P., 2017. Hard and soft managerialism in Portuguese higher education. In: Eggins, H., Deem, R. (Eds.), *The University as a Critical Institution?* Sense Publishers, Rotterdam, pp. 39–53.
- Magalhães, A., Veiga, A., Amaral, A., 2018a. The changing role of external stakeholders: from imaginary friends to effective actors or non-interfering friends. *Stud. High Educ.* 43 (4), 737–753.
- Magalhães, A., Veiga, A., Videira, P., 2018b. Reconfiguring power in Portuguese higher education. *Educ. Pol. Anal. Arch.* 26 (135).
- Magalhães, A., 2004. *A identidade do ensino superior - política, conhecimento e educação numa época de transição*. Fundação Calouste Gulbenkian/Fundação para a Ciência e Tecnologia, Lisbon.
- Meek, V.L., 2003. Introduction. In: Amaral, A., Meek, V.L., Larsen, I.M. (Eds.), *The Higher Education Managerial Revolution?* Kluwers Academic Publishers, Dordrecht, pp. 1–29.
- Neave, G., 2008. From guardian to overseer: trends in institutional autonomy, governance and leadership. In: Amaral, A. (Ed.), *Reforma do ensino superior: quatro temas em debate*. Conselho Nacional de Educação, Lisboa.
- Neave, G., 2012. *The Prince and His Pleasure: The Evaluative State, Institutional Autonomy and Reengineering Higher Education in Western Europe*. Palgrave, Basingstoke.
- Paradeise, C., Reale, E., Gostellec, G., Bleiklie, I., 2009. Universities steering between stories and history. In: Paradeise, C., Reale, E., Bleiklie, I., Ferlie, E. (Eds.), *University Governance: Western European Comparative Perspectives*. Springer, Dordrecht, pp. 227–246.
- Peters, G., 2001. *The Future of Governing*. University Press of Kansas, Lawrence.
- Peters, G., 2010. Meta-governance and public management. In: Osborne, S. (Ed.), *The New Public Governance? Emerging Perspectives on the Theory and Practice of Public Governance*. Routledge, London, pp. 36–51.
- Portugal, Ministério da Ciência e do Ensino Superior, 2007a. Regime jurídico das instituições do ensino superior. Law n.º 62/2007.
- Portugal, Ministério da Ciência e do Ensino Superior, 2007b. Regime jurídico da avaliação do ensino superior. Law n.º 38/2007.
- Salamon, L., 2002. The new governance and the tools of public action: an introduction. In: Salamon, L. (Ed.), *The Tools of Government: A Guide to the New Governance*. Oxford University Press, Oxford, pp. 1–47.
- Veiga, A., Magalhães, A., Amaral, A., 2015. From collegial governance to boardism: reconfiguring governance in higher education. In: Huisman, J., de Boer, H., Dill, D., Souto-Otero, M. (Eds.), *The Palgrave International Handbook of Higher Education Policy and Governance*. Palgrave Macmillan, London, pp. 398–416.

Publish or perish

Hanne Andersen, Section for History and Philosophy of Science, Department of Science Education, University of Copenhagen, Copenhagen, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction	158
The role of publications in academic careers	158
Sociological studies of the academic reward system	159
The role of bibliometrics in research assessment	160
Productivity	160
Impact	161
Aggregate indicators	161
Publication venues	162
Metrics as a product	163
Recommendations for good practice	163
Adverse effects of a quantified view on “publish or perish”	165
References	165

Introduction

The phrase “publish or perish” has traditionally been linked to a particular understanding of academia as a primarily research-oriented meritocracy. At many academic institutions around the globe, merit in research tend to be perceived as more important than merit in teaching and other academic duties. Promotions, pay raises, research grants, academy memberships and other forms of academic rewards are very often granted primarily based on the recipient’s research achievements. This reward structure also has a tendency to be self-reinforcing: the more you have published, the more likely is it that you can get additional research funding, and the more additional publications will you be able to produce. The academic reward structure is therefore often perceived as producing a very strong pressure on academics to publish as much as possible. The phrase “Publish or perish” expresses this pressure.

The phrase has been used at least since the first half of the 20th century (see e.g., [Case, 1927](#); [Wilson, 1942](#) for early examples). In academic novels, such as, for example, Kingsley Amis’ (1954) *Lucky Jim*, or James Hynes’ (1997) *Publish and Perish: Three Tales of Tenure and Terror* it is a frequently depicted theme how success in academia is tied closely to publishing, and vice versa, that failure to publish is the road to academic perdition (see e.g., [Moseley, 2007](#); [Tierney, 2004](#) on the publish or perish theme in academic novels).

At many institutions around the globe, the focus on achievements has gradually widened to include not only on the amount of research that a scholar has published, but also the impact that this research has had in the academic community. Reflecting this developments, the phrase “publish or perish” has been supplemented with a number of additional phrases that express the need not only to publish, but also to produce publications that will be read and cited by others, such as for example the expressions “impact or perish” ([Biagioli, 2016](#); [Biagioli and Lippman, 2020](#)), “be discoverable or die” ([Doyle and Cuthill, 2015](#)), or “get visible or vanish” ([Doyle and Cuthill, 2015](#)).

This entry will provide a brief overview of several facets of the “publish or perish” idea. First, it will describe the role that publications play as one factor among others in academic careers. Further, it will describe how the strong emphasis on publications in the academic reward system has been vindicated by early work in sociology of science. Next, it will review some of the ways in which scholars’ research achievements are assessed today. Finally, it will describe some of the adverse effects of current assessment practices, and it will discuss some of the points of criticism that are today raised against the publish or perish doctrine.

The role of publications in academic careers

Globally, there is a great deal of variation in what is considered important for succeeding in an academic career ([Altbach, 1996, 2016](#); [Boyer and Altbach, 1997](#); [Boyer, 1990](#)). Not only are there important differences on a global scale, for example between the Global North and the Global South (see e.g., [Cloete and Maassen, 2015](#)); there are also important regional differences, for example between the individual European countries, or between, for example, research-intensive and teaching-intensive institutions in the US ([Boyer, 1990](#)). It is therefore difficult to give a unified treatment of the role of publications in academic careers that gives justice to all local variations.

Rooted in the idea of the research university as it developed in Europe and North America during the 19th and 20th centuries, there has been a strong tendency to base hiring and promotion decisions primarily on research achievements, especially at research-

intense universities. At the same time, it has varied widely between institutions, between countries, and between fields, how this focus on research achievements has been implemented. In some fields and in some countries, regular publication of specialized monographs is necessary in order to be seen as a serious academic—even if at the same time it is discussed whether this version of the publish or perish doctrine leads to an overproduction of monographs (Lewis, 2002; Waters, 2001). In other fields and in other countries, journal publications are at focus. In some fields, authors on the by-line of a publication are ordered according to their contributions, where typically the first author has been a driving force in producing the research reported, while the last author has had a supervisory or overseeing role. In fields that follow this convention, it is typically expected that junior researchers develop their careers by getting more first-authorships, while senior researchers develop their careers by getting more last-authorships. Further, other kinds of research achievements are often included as well when assessing a scholar, such as, for example, the ability to attract research funding, or the amount of invitations to give talks.

However, critics have argued that different universities serve different goals and that a narrow focus on academic publications can be counterproductive for the role that a university serves locally or globally (Chou and Chan, 2017; Confraria and Godinho, 2015; Confraria et al., 2017, 2018, 2020; Mirnezami and Beaudry, 2021). Similarly, with the development of the mass university, it has been questioned whether it is adequate to focus exclusively research achievements when assessing the merits of university professors. For example, Ernest L. Boyer has argued in the report *Scholarship Reconsidered: Priorities of the Professoriate* (1990) published by the Carnegie Foundation for the Advancement of Teaching that the traditional focus on research is too narrow a conception of the role of professors. Instead, he found that the scholarship of the professoriate consists not only of cutting-edge inquiry, but also of the ability to place discoveries into a larger context, the ability to apply new knowledge, and the ability to transmit new knowledge to new generations. Similarly, the Career Framework for University Teaching, developed by Ruth Graham in collaboration with universities from around the world, has been “designed to guide and support the career progression of academics on the basis of their contribution to teaching and learning” (<https://teachingframework.com/>). This framework describes how decisions on promotion can be based on contributions that enhance excellence in teaching and learning within and beyond the assessed scholar’s institution, and that such contributions can be documented, for example, through professional activities such as leadership roles in policy development, or by design and delivery of high-impact educational material.

Sociological studies of the academic reward system

The link between publications and academic reward is an integral part of the modern research university. Up until the mid-20th century, the link was primarily described qualitatively, such as, for example, the account offered by dean Logan Wilson of Tulane University in his 1942 monograph *The Academic Man* that explored academic hierarchies, academic status and academic processes and functions at North American universities. Here, Wilson described both how the individual researchers’ possibilities for “horizontal or vertical mobility” is dependent on visibility and therefore by publications, and the precarious situation of the young instructors who needed to pay lip service to teaching while at the same time dealing with departments that emphasized quantity rather than quality in publication.

Quantitative studies of the academic reward system and of how stratification works in academia became popular during the 1960s and 1970s, after the emergence of the Science Citation Index made it possible to track citations. This commercial service made it possible to examine correlations between, for example, on the one hand scholars’ numbers of publications and citations, and on the other hand the positions and awards that they achieved. Such studies became an important part of the sociology of science that emerged with the work of scholars such as Robert Merton, Harriett Zuckermann, Jonathan Cole, Stephen Cole, Diana Crane, Warren Hagstrom, Daryl Chubin, and Derek de Solla Price.

Work from this era has produced several important qualifications to the “publish or perish” doctrine. First, studies of scholars from specific disciplines revealed that awards and prestigious positions were more strongly correlated with citation than with the number of publications (Cole and Cole, 1967, 1974). These studies indicated that it is not the act of publishing per se that is important for recognition, but rather the degree to which the produced publications influence other scholars. Second, these empirical studies also showed that while there is a certain correlation between citations and recognition, this correlation is not perfect. Admittedly, it is rare that a scientist with little impact receive significant rewards; yet, many scientists with high impact do not receive the same recognition as what has been bestowed to other peers with a similar impact (Merton, 1968). Third, the empirical studies also showed that reward is determined by many other factors than productivity and impact. For example, comparing scholars from within the US it was shown that scholars gain more recognition from their high productivity when they are employed on a major university rather than a less known university (Crane, 1965).

Studies of the correlation between publications, citations and reward have therefore also been used to point to structural inequalities in the scientific reward system. These studies were initiated with Merton’s work on the Matthew effect (Merton, 1968, 1988). Based on a combination of his own historical studies and the quantitative studies performed by a number of colleagues, most notably Zuckermann, Cole and Cole from his own research group, Merton argued that prominent scientists accrued greater increments of recognition for similar contributions than their less prominent peers and vice versa, and that this produced a cumulative advantage to the benefit of the more prominent scholars (Merton, 1968, 57). Later, historian of science Margaret Rossiter (1993) added new perspectives to this work, arguing that many women in the history of science had not received the recognition that they deserved. This phenomenon of systematic under-recognition of women’s achievements she named the Matilda

effect, after the nineteenth-century American sociologist of knowledge Matilda Joselyn Cane.¹ In recent years, there has been an increasing awareness of structural inequalities on a global scale (Nielsen and Andersen, 2021). For example, analyses of the academic impact of scholars from the Global South indicate that researchers who collaborate with partners from the Global North are more frequently cited than researchers without such collaborations (Confraria et al., 2017, 2018; Mirnezami and Beaudry, 2021).

The role of bibliometrics in research assessment

Over the last decades, it has become increasingly popular to draw on bibliometric indicators to assess the merits of individual scholars when they apply for positions or research grants. Multiple factors have contributed to this development. First, the companies that produce citation-tracking databases have advertised how their products can be used for producing objective evaluations (Garfield and Malin, 1968; Garfield, 1970). Second, the volume of research activity has increased immensely over the last century, and so has the number of scholars competing for positions and funding. In this situation, agencies and institutions that are faced with the time-consuming task of sorting between more and more applicants often see an interest in indicators that can establish a ranking between them more easily than through traditional review. Similarly, many academics are interested in keeping track of their position relative to their peers in the scholarly hierarchy, and bibliometric indicators that can be found in public databases can be seen as an easily tool for that purpose. Third, as public funding of research has increased, so has the demand for documenting what is gained from this expenditure. Further, institutions at all levels, from the individual departments over universities, to large foundations and governments agencies are increasingly interested in benchmarking. Hence, bibliometric indicators are increasingly being used not only to assess individual scholars, but also to compare between institutions or to document the return of investment in research (Godin, 2004).

With the increasing demand for bibliometrics, the number of bibliometric indicators have proliferated and so has the academic study of scientometrics. The literature is vast, and following section will focus only on those bibliometric indicators that are currently most popular for assessing individual scholars.

Productivity

The number of publications that a scholar has authored is often used as a measure of his or hers productivity. However, this is a very crude indicator, which obviously benefits senior scholars over junior scholars, simply because they have had more time available. If instead one wants to understand productivity as related to the efficiency in producing a given output, the number of publications need to be seen in relation to the time and resources that have been involved in producing these publications.

This suggests, first, that information about an author's number of publications should be supplemented with information about the time span during which they have been produced. Knowing both the number of publications that a scholar has authored, and the time span during which they have been authored makes it easier to compare scholars at different career stages. In principle, it makes it possible to take leaves of absence, parental duties etc. into account when assessing productivity. However, practices for whether or not to consider such aspects vary considerably between institutions and between countries and cultures. Similarly, it makes it possible to adjust for differences in working conditions, for example between research-intensive and teaching-intensive institutions, but this possibility is rarely pursued in practice.

Second, understanding productivity as efficiency also suggests that information about an author's number of publications should be supplemented with information about the amount of collaboration involved in the production of these publications. To compensate for the role of co-authors, the number of publications can, for example, be replaced with a basic fractional counting of authorship that only adds the number $1/m$ for a publication with m authors.² However, basic fractional counting of authorships assumes that all co-authors have shared the work equally among them, and this is only rarely the case. To compensate for that, it has been suggested instead to compute a weighted fractional output by weighing each author's basic fraction with a factor that depends on their order on the byline and on their institutional distribution (Abramo et al., 2013). Still, this depends on specific assumptions about the distribution of labor between, for example, first and last authors. In reality, a realistic assessment of the workload that a particular author has contributed to a given publication can only be assessed if the publication contains a contribution statement that offers detailed information about the contributions of each individual author. While more and more journals have started to require such detailed contribution statements, the categories used vary considerably from statement to statement, and there is currently no general procedure for transforming such statements into a quantified measure for the degree of authorship.

¹For an overview of recent empirical studies on this effect, see e.g., Larivière et al. (2013); for an estimate of the development over time and the prospect of a harmonization between the genders, see e.g., Bendels et al. (2018).

²Considerations such as these are included in the program Harzing's Publish or Perish (available from <https://harzing.com/resources/publish-or-perish>) that computes the average number of authors on a given scholar's publications as well as the normalized number of papers per author computed as the sum of the normalized author count for each of the papers that the author has co-authored (see also Harzing, 2010).

Impact

Productivity alone of course says little about the quality of the publications produced. Traditionally, the quality of research publications has been assessed by experts who have read them. However, as also mentioned above, as more and more scientists apply for positions and compete for research grants, it is becoming more and more time consuming to conduct such assessments. Many institutions and agencies therefore have a strong desire for quantitative indicators that can be used as a short cut instead of the time-consuming reviews.

Often, it is assumed that publications that provide important advances within their fields will also be widely read and cited, in other words, that research of a high quality will also have a high impact that can be traced through citations. Based on this assumption, the number of citations that a publication receives is seen as an indicator of its impact and quality.

Various databases exist that track citations. For example, the database Scopus, offered by the publishing company Elsevier, tracks how many citations are made to each publication in the database from other publications in the database and offers users information of the total number of citations, their development over time, and gives an assessment of how the citation count for a given publication compares to the average in the field. In addition, the service SciVal, also offered by Elsevier, uses the Scopus database to provide statistics of how large a share of a scholars publications that are among the top 10% and top 1% for citations within their field worldwide. Similarly, the database Web of Science, offered by Clarivate, also tracks citations between publications in the database. Google Scholar traces citations in scholarly literature more widely, including theses, abstracts, preprints and documents from academic online repositories and university web pages. For scholars who have created a personal Google Scholar profile, their personal profile page reports the aggregate number of citations. For scholars who do not have a personal Google Scholar profile, programs such as Harzing's Publish or Perish can be used to track their citation in Google Scholar as well as in a number of other databases.

Because these databases differ in which journals, languages, fields, and non-journal types of publications they cover, citation numbers for the same publication or the same author may vary from database to database. Further, citations depend on many factors, including local citation practices in the field, the number of other scholars working in the field, the language and the accessibility of the publication, etc. Citation numbers are therefore difficult to compare, especially across fields and across languages.

There is an extensive literature in scientometrics on how citations scores can be corrected for differences between fields; what it typically referred to as field-normalization. Field-normalizing citation scores requires a clear definition of scientific fields. Some field-normalization procedures draw on the field classification of journals provided by the databases Web of Science or Scopus, while others are based on how publications cluster together (Waltman and van Eck, 2019).

In addition, Leydesdorff and colleagues have argued that a distinction should be made between short-term citations that indicate the degree to which a publication is involved in current academic discourse, and long-term citations that indicate the degree to which a publication is incorporated into a generally accepted body of knowledge (Leydesdorff et al., 2016).

Aggregate indicators

In many decisions on academic reward, the aim is to honor those who are very efficient in producing publications of a high quality. It has therefore become popular to use aggregate indicators that combine productivity and impact. The currently most popular aggregate indicator is the so-called H index that was developed by the physicist Hirsch (2005). A scholar has an H index of n if he or she has produced n publications that have each been cited at least n times.

A crucial shortcoming of the H index is that it does not distinguish between, for example, a scholar who has published twenty papers that have each received little more than twenty citations, and a scholar who has published twenty papers that have each received hundreds or even thousands of citations. As argued by Gingras (2016), it is a typical shortcoming of aggregate indicators such of the H index that they fail to be sensitive to important variation in the phenomenon that they are supposed to measure.

Due to this shortcoming of the H index, a plethora of alternative indicators have been suggested to represent particular aspects of productivity and impact. Especially, the g and e indexes have been introduced to compensate for the problem that a scholar who has published ten papers with each ten citations has the same h index as a scholar who has published ten papers with each one hundred citations. For a scholar with an H index of n , the e -index is computed as the square root of the number of citations that exceeds n^2 (Zhang, 2009), while the g index is computed as the largest number such that the top g articles in a set of articles ranked in decreasing number of citations received (together) at least g^2 citations (Egghe, 2006). Comparing two scholars with similar numbers of publications and citations, where the one scholar has some publications that have received more citations than any of the publications authored by the other scholar, while the other scholar has fewer publications than the first that receive only few citations, the former will have higher e and g indexes, while the latter will have a higher h index. This is illustrated in Fig. 1, where the scholar with the higher peak has higher e and g indexes, while the scholar with the fatter tail will have a higher h index. There is no simple answer to the question which of the two scholars is "better." For example, a high peak may be due to the production of significant scientific advances, or it may be due to the production of widely applicable reviews of the research of others. Similarly, a fat tail may be due to a steady and solid research activity on a focused topic, or it may be due to participation in a "citation cartel" where a group of authors always cites each other's work. Only a close scrutiny of the actual publications can provide a more detailed assessment of the two compared to each other.

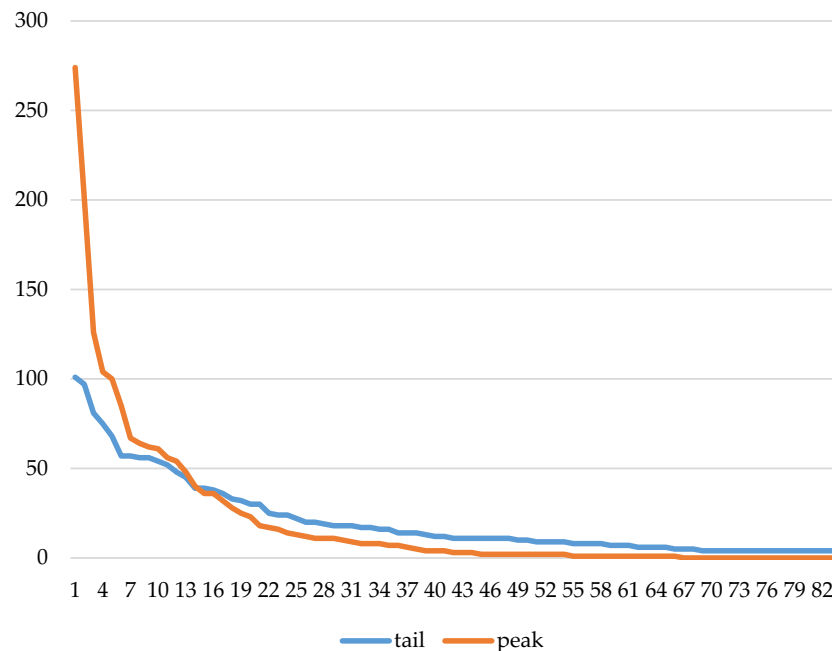


Fig. 1 Number of citations (y-axis) per paper for two different researchers with comparable number of publications and citations. The one with the higher peak has higher e and g indexes, while the one with the fatter tail has a higher h index.

Another point of criticism is that the H index is cumulative and therefore favors researchers who have been publishing over a long period of time. To adjust for academic age, the m quotions divides the H index by the number of years since the publication of the scholar's first publication, while the hIa-index adjusts for the duration of the career as well as for co-authorships by normalizing citations for each paper by dividing the number of citations by the number of authors for that paper, calculating the hI norm as the h-index from these normalized citation numbers, and then dividing by the number of years that the author has been active (Harzing, 2010). One the one hand, the m index and the hIa index in each their different ways adjust for the length of a scholar's career, but in doing so they also penalize authors who happened to have published their first paper particularly early in their career.

Commenting on the steadily increasing number of different bibliometric measures, Jappe et al. (2018) have argued that evaluative citation analyses does currently not form a traditional academic field with established standards of research competence and methodology, and where experts in the field exert reputational control of the field's development. Without an organized scientific community to give authoritative recommendations, no standard practices have developed.

Publication venues

Another popular proxy indicator for the quality of research publications is the publication venue. Assuming that particular journals primarily publish papers of a very high quality, publications published in these particular journals are expected to be of a high quality, while the same expectation does not hold for publications published in other journals. Many scholars therefore flag if their work is published in what they perceive as high quality publication venues.

There are different ways in which to identify high quality journals. Traditionally, it has been part of the shared knowledge in an academic field to know which journals were seen as central to the field, or in which journals it was considered especially hard to get a manuscript accepted. With the emergence of citation databases, this has been supplemented with quantitative indicators based on average citation rates. Web of Science computes a Journal Impact Factor (IF) for journals included in the database as the average number of citations that articles in a given journal receives over a two-year period. Similarly, the SCImago Lab uses data from Scopus to compute a SCImago Journal Rank (SJR) that deviates from the Web of Science' impact factor by using a 3-year citation window and by weighing citations according to the prestige of the journal from which they come.

However, the publication venue is a highly fallible indicator for the quality of individual publications. It is important to be aware that journal impact factors were originally developed to assist libraries with limited budgets to decide on subscriptions, not as an indicator of the quality of the individual papers (Garfield, 1972). For example, although it may be relatively rare that poor research is published in a highly competitive journal, high quality research can easily be published in a mediocre journal. Hence, despite information on the average impact of the journal in which a paper has been published, its quality can only be assessed by reading it, or as a proxy by assessing its impact by counting citations to the paper itself.

Metrics as a product

It is easy for scholars to keep track of their publications and to document their productivity through publication lists that specify the bibliographic information for each publication. In contrast, it can be much more difficult to get an overview of how often their publications have been cited. Citations can be made by many different people and in many different channels, hence, there is no easy way of knowing all citations to a particular publication or to a particular author.

Several companies therefore offer databases that track citations between academic publications. The first such service was the Scientific Citation Index that was created in the late 1950s and introduced as a commercial product in the mid 1960s. Originally, it covered only the natural and medical sciences. Later, the Arts and Humanities Index and the Social Science Index was added, and the indexes were merged into what is now called the Web of Science. To be included in the index, journals have to live up to a number of quality criteria and impact criteria. These include such issues as significance of the content, presence of policies for research ethics and for peer review, functionality of the website, and timeliness in production. Over the last decades, various alternatives to the Web of Science have emerged. The database Scopus produced by Elsevier is similar to the Web of Science in providing a broad coverage that includes not only the natural and health sciences but also the social sciences and humanities. Most other databases focus on particular academic areas, such as, for example, the biomedical database PubMed.

Different databases may differ substantially in their coverage of the literature. Such differences may include how far back the database collects information, the journals or publishers from which it collects information, or whether it includes books, conference series and other types of information. Therefore, the same author or the same publication may be listed in different databases with different citation counts. Some databases change their coverage over time, for example by adding new types of material or by extending its historical coverage further back in time. When additional material is added, citations will be added as well. Hence, as databases improve their coverage, scholars' citation numbers increase simply because more old citations are being counted.

While databases such as Web of Science or Scopus are curated and have detailed policies for which publications can be included and for the editorial processes through which it is decided, Google Scholar instead uses automated software to index publications found on websites, including academic repositories provided by universities and scholarly societies. In general, Google Scholar therefore covers much more literature than the curated databases, but it is also less transparent what is being covered. Citation counts drawn from Google Scholar therefore tend to be substantially higher than citation counts found in the curated databases. Although some studies indicate that citation analyses conducted on data from Web of Science, Scopus and Google Scholar provide equivalent results (Harzing and Alakangas, 2016; Martín-Martín et al., 2018), it is an issue of debate whether the excess citations in Google Scholar compared to Web of Science or Scopus should carry the same weight as the academic citations counted in the curated databases.

Recommendations for good practice

While the use of bibliometric indicators for the assessment of academics has become widespread, this use has also become subject of criticism, especially from scholars from scientometrics (Hicks et al., 2015; Stephan et al., 2017), statistics (Adler et al., 2009) and history and philosophy of science (Gingras, 2016).

To accommodate some of the points of criticism, some researchers have created guidelines for good practice in research evaluation. The Leiden Manifesto for Research Metrics (Hicks et al., 2015) lists the following ten principles for best practice in metrics-based research assessment³:

1. Quantitative evaluation should support qualitative, expert assessment
2. Measure performance against the research missions of the institution, group or researcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical processes open, transparent and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgment of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognize the systemic effects of assessment and indicators.
10. Scrutinize indicators regularly and update them.

Similarly, the San Francisco Declaration on Research Assessment (DORA; <https://sfdora.org>) presents a set of recommendations for agencies, institutions and researchers for how metrics should be used in research assessments. The initiative for these recommendations comes from a group of editors and publishers of scholarly journals, and an important focus of these recommendations is to eliminate the use of journal-based metrics such as journal impact factors when assessing researchers in relation to funding, appointment or promotion decisions. Thus, a central theme in these recommendations is the need to assess research on its own merits rather than based on the venue in which it has been published.

³A video version as well as translations into several languages can be found at <http://www.leidenmanifesto.org/>.

San Francisco Declaration on Research Assessment**General recommendation**

- Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.

For funding agencies

- Be explicit about the criteria used in evaluating the scientific productivity of grant applicants and clearly highlight, especially for early-stage investigators, that the scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published.
- For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.

For institutions

- Be explicit about the criteria used to reach hiring, tenure, and promotion decisions, clearly highlighting, especially for early-stage investigators, that the scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published.
- For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.

For publishers

- Greatly reduce emphasis on the journal impact factor as a promotional tool, ideally by ceasing to promote the impact factor or by presenting the metric in the context of a variety of journal-based metrics (e.g., 5-year impact factor, EigenFactor [8], SCImago [9], h-index, editorial and publication times, etc.) that provide a richer view of journal performance.
- Make available a range of article-level metrics to encourage a shift toward assessment based on the scientific content of an article rather than publication metrics of the journal in which it was published.
- Encourage responsible authorship practices and the provision of information about the specific contributions of each author.
- Whether a journal is open-access or subscription-based, remove all reuse limitations on reference lists in research articles and make them available under the Creative Commons Public Domain Dedication [10].
- Remove or reduce the constraints on the number of references in research articles, and, where appropriate, mandate the citation of primary literature in favor of reviews in order to give credit to the group(s) who first reported a finding.

For organizations that supply metrics

- Be open and transparent by providing data and methods used to calculate all metrics.
- Provide the data under a license that allows unrestricted reuse, and provide computational access to data, where possible.
- Be clear that inappropriate manipulation of metrics will not be tolerated; be explicit about what constitutes inappropriate manipulation and what measures will be taken to combat this.
- Account for the variation in article types (e.g., reviews vs. research articles), and in different subject areas when metrics are used, aggregated, or compared.

For researchers

- When involved in committees making decisions about funding, hiring, tenure, or promotion, make assessments based on scientific content rather than publication metrics.
- Wherever appropriate, cite primary literature in which observations are first reported rather than reviews in order to give credit where credit is due.
- Use a range of article metrics and indicators on personal/supporting statements, as evidence of the impact of individual published articles and other research outputs.
- Challenge research assessment practices that rely inappropriately on Journal Impact Factors and promote and teach best practice that focuses on the value and influence of specific research outputs.

Finally, the Research Quality Plus Assessment Framework (RQ+) developed by Canada's International Development Research Center (IDRC) aims at evaluating the quality of translational research, especially research that is grounded in and applied to local experience (Lebel and McLean, 2018). This framework assists evaluators, managers and funders in identifying constraining factors of a research projects related to:

1. Maturity of the research field
2. Research capacity strengthening
3. Risk in the data environment
4. Risk in the research environment
5. Risk in the political environment

As well as dimensions of research quality related to.

1. Research integrity

2. Research legitimacy
3. Research importance
4. Positioning for use

RQ+ provides an evaluative rubric with a language for describing key influencing factors and for specifying criteria for judging quality. Hence, the framework is aimed at including “important aspects of the context in which the research has been conducted, and the manner in which it has been managed” (Ofir et al., 2016, 12).

Adverse effects of a quantified view on “publish or perish”

Strong incentive structures often have adverse effects. The “publish or perish” doctrine’s strong emphasis on the need to publish or the need to be visible have several adverse effects.

First, scholars may boost their publication lists by publishing the same result several times. If the redundancy is openly declared, it can be fully legitimate, for example if the parallel publications serves to communicate to a different audience than the original publication. However, if the redundancy is concealed, the duplicate publication is seen as questionable research practice. Similarly, scholars may boost their publication list by slicing a given result into as many “salami” publications as possible, instead of publishing it all at once.

Second, scholars may boost their publication list by becoming co-authors on paper where they do not really deserve it. Sometimes, co-authors are included on the byline of a paper although their contribution to the publication has been only marginal. This is usually referred to as “gift authorships.” There can be many reasons why scholars grant such gift authorships to a co-author. Sometimes, it is difficult to draw the line of whom to include on the byline and gift authorships may be granted out of courtesy. Sometimes, they are granted to prominent scholars in the field in order to increase the visibility of a paper. Sometimes, they are demanded by senior scholars of power, and they are then referred to as coercive authorship. In all of the mentioned cases, a gift authorship confers credit to people who have not earned this credit, and it gives the reader a wrong impression of who bears responsibility for the research reported in the paper. It is therefore considered questionable research practice to accept as well as to grant gift authorships.

To promote good authorship practices, the International Committee of Medical Journal Editors has issued a set of guidelines for authorship that gradually become accepted by more and more journals across a broad range of fields. These guidelines, often referred to as the Vancouver guidelines, stipulate that authors to a publication should fulfill the following four criteria:

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- Drafting the work or revising it critically for important intellectual content; AND
- Final approval of the version to be published; AND
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

However, although the guidelines are widely adopted, empirical studies indicate that many authorship are still granted for minor activities that should instead have been accredited in the acknowledgments (Patience et al., 2019; Macfarlane, 2017; Martinson et al., 2005).

The strong emphasis on bibliometric indicators such as the H index can also lead to adverse effects. For example, due to the strong emphasis on high impact, scholars may boost their citation counts by forming citation cartels whose members continuously cite each other. Similarly, referees may in their reviews request citations of their own work, or editors may in their decision letters request citations of papers in their journal. The same holds for indicators at aggregate levels. For example, editors may boost their journal’s impact factor by carefully distributing articles across the year in such a way that papers for which the editor expects many citations are published at the beginning of the year and therefore stays for a longer time in the citation window that is used to compute the impact factor, while papers that are expected to receive only few citations are published at the end of the year.

Finally, it has also been argued that the overall publication pressure in academia makes some researchers cut corners and rush their findings and that this has contributed to the ongoing reproducibility crisis (Baker, 2016; Grimes et al., 2018; Ioannidis, 2005; Young et al., 2008).

Hence, as argued by Elton (2004), examples such as these vindicate Goodhart’s Law that says that an indicator typically ceases to be a good indicator when it becomes a target in itself. This is because indicators only serve as partial proxies for actual performance. Hence, when indicators are used for control, they change behavior to optimize what is being measured and neglect that is not being measured.

References

- Abramo, G., Andrea D’Angelo, C., Rosati, F., 2013. The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences. *J. Informetr.* 7 (1), 198–208.

- Adler, R., Ewing, J., Taylor, P., 2009. Citation statistics: a report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Stat. Sci.* 24 (1), 1–14.
- Altbach, P.G., 1996. *The International Academic Profession*. Carnegie Foundation for the Advancement of Teaching, Princeton, NJ.
- Altbach, P.G., 2016. *Global Perspectives on Higher Education*. JHU Press.
- Baker, M., 2016. Reproducibility crisis. *Nature* 533 (26), 353–366.
- Bendels, M.H.K., Müller, R., Brueggmann, D., Groneberg, D.A., 2018. Gender disparities in high-quality research revealed by Nature Index journals. *PLoS One* 13 (1), e0189136.
- Biagioli, M., Lippman, A., 2020. *Gaming the Metrics: Misconduct and Manipulation in Academic Research*. MIT Press.
- Biagioli, M., 2016. Watch out for cheats in citation game. *Nature* 535 (7611), 201.
- Boyer, E., Altbach, P.G., 1997. *The International Academic Profession*. Carnegies Foundation for the Advancement of Teaching, New York.
- Boyer, E.L., 1990. *Scholarship Reconsidered: Priorities of the Professoriate*. Carnegies Foundation for the Advancement of Teaching, New York.
- Case, C.M., 1927. Scholarship in sociology. *Sociol. Soc. Res.* 12, 323–340.
- Chou, C.P., Chan, C.-F., 2017. Governance and academic culture in higher education: under the influence of the SSCI syndrome. *J. Int. Comp. Educ.* 6 (2), 63–75.
- Cloete, N., Maassen, P., 2015. The role of universities and the African context. In: Cloete, N., Maassen, P., Bailey, T. (Eds.), *Knowledge Production and Contradictory Functions in African Higher Education*. African Minds, Cape Town, pp. 1–17.
- Cole, S., Cole, J.R., 1967. Scientific output and recognition: a study in the operation of the reward system in science. *Am. Sociol. Rev.* 32 (3), 377–390.
- Cole, J.R., Cole, S., 1974. Social stratification in science. *Am. J. Phys.* 42 (10), 923–924.
- Confraria, H., Godinho, M.M., 2015. The impact of African science: a bibliometric analysis. *Scientometrics* 102 (2), 1241–1268.
- Confraria, H., Godinho, M.M., Wang, L., 2017. Determinants of citation impact: a comparative analysis of the Global South vs. the Global North. *Res. Pol.* 46 (1), 265–279.
- Confraria, H., Blanckenberg, J., Swart, C., 2018. The characteristics of highly cited researchers in Africa. *Res. Eval.* 27 (3), 222–237.
- Confraria, H., Blanckenberg, J., Swart, C., 2020. Which factors influence international research collaboration in Africa?. In: *Africa and the Sustainable Development Goals*. Springer, pp. 243–255.
- Crane, D., 1965. Scientists at major and minor universities: a study of productivity and recognition. *Am. Sociol. Rev.* 30 (5), 699–714.
- Doyle, J., Cuthill, M., 2015. Does “get visible or vanish” herald the end of “publish or perish”. *J. High. Educ. Res.* 34 (3), 671–674.
- Egghe, L., 2006. An improvement of the h-index: the g-index. *ISSI Newsl.* 2 (1), 8–9.
- Elton, L., 2004. Goodhart’s Law and performance indicators in higher education. *J. Eval. Res. Educ.* 18 (1–2), 120–128.
- Garfield, E., Malin, M.V., 1968. Can Nobel Prize Winners be Predicted. 135th meetings of the American Association for the Advancement of Science, Dallas, TX.
- Garfield, E., 1970. Citation indexing for studying science. *Nature* 227 (5259), 669–671.
- Garfield, E., 1972. Citation analysis as a tool in journal evaluation. *Science* 178 (4060), 471–479.
- Gingras, Y., 2016. *Bibliometrics and Research Evaluation: Uses and Abuses*. MIT Press.
- Godin, B., 2004. *Measurement and Statistics on Science and Technology: 1920 to the Present*. Routledge.
- Grimes, D.R., Bauch, C.T., Ioannidis, J.P.A., 2018. Modelling science trustworthiness under publish or perish pressure. *R. Soc. Open Sci.* 5 (1), 171511.
- Harzing, A.-W., Alakangas, S., 2016. Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics* 106 (2), 787–804.
- Harzing, A.W., 2010. *The Publish or Perish Book: Your Guide to Effective and Responsible Citation Analysis*.
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., Rafols, I., 2015. Bibliometrics: the Leiden Manifesto for research metrics. *Nature* 520 (7548), 429.
- Hirsch, J.E., 2005. An index to quantify an individual’s scientific research output. *Proc. Natl. Acad. Sci. U. S. A.* 102 (46), 16569–16572.
- Ioannidis, J.P.A., 2005. Why most published research findings are false. *PLoS Med.* 2 (8), e124.
- Jappe, A., Pithan, D., Heinze, T., 2018. Does bibliometric research confer legitimacy to research assessment practice? A sociological study of reputational control, 1972–2016. *PLoS One* 13 (6), e0199031.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto, C.R., 2013. Bibliometrics: global gender disparities in science. *Nature* 504 (7479), 211.
- Lebel, J., McLean, R., 2018. A better measure of research from the global south. *Nature* 559, 23–25.
- Lewis, P., 2002. Is monographic tyranny the problem? *Publ. Mod. Lang. Assoc. Am.* 117 (5), 1222–1224.
- Leydesdorff, L., Bornmann, L., Comins, J.A., Milojević, S., 2016. Citations: indicators of quality? The impact fallacy. *Front. Res. Metr. Anal.* 1, 1.
- Macfarlane, B., 2017. The ethics of multiple authorship: power, performativity and the gift economy. *Stud. High. Educ.* 42 (7), 1194–1210.
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., López-Cózar, E.D., 2018. Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. *J. Informetr.* 12 (4), 1160–1177.
- Martinson, B.C., Anderson, M.S., De Vries, R., 2005. Scientists behaving badly. *Nature* 435, 737–738.
- Merton, R.K., 1968. The Matthew effect in science: the reward and communication systems of science are considered. *Science* 159 (3810), 56–63.
- Merton, R.K., 1988. The Matthew effect in science, II: cumulative advantage and the symbolism of intellectual property. *Isis* 79 (4), 606–623.
- Mirnezami, S.R., Beaudry, C., 2021. Does experiencing international research collaboration permanently affect the impact of scientific production? Evidence from Africa. *J. Afr. Econ.* <https://doi.org/10.1093/jae/ejab008>.
- Moseley, M., 2007. Types of academic fiction. In: *The Academic Novel: New Classical Essays*, pp. 99–113.
- Nielsen, M.W., Andersen, J.P., 2021. Global citation inequality is on the rise. *Proc. Natl. Acad. Sci. U. S. A.* 118 (7).
- Ofir, Z., Schwandt, T., Duggan, C., McLean, R., 2016. RQ+. Research Quality Plus. A Holistic Approach to Evaluating Research. IDCRC, Ottawa.
- Patience, G.S., Galli, F., Patience, P.A., Boffito, D.C., 2019. Intellectual contributions meriting authorship: survey results from the top cited authors across all science categories. *PLoS One* 14 (1), e0198117.
- Rosser, M.W., 1993. The Matthew Matilda effect in science. *Soc. Stud. Sci.* 23 (2), 325–341.
- Stephan, P., Veugelers, R., Wang, J., 2017. Reviewers are blinkered by bibliometrics. *Nature* 544 (7651), 411.
- Tierney, W.G., 2004. Academic freedom and tenure: between fiction and reality. *J. High. Educ.* 75 (2), 161–177.
- Waltman, L., van Eck, N.J., 2019. Field normalization of scientometric indicators. In: *Springer Handbook of Science and Technology Indicators*. Springer, pp. 281–300.
- Waters, L., 2001. Rescue tenure from the tyranny of the monograph. *Chron. High. Educ.* 47 (32), B7–B10.
- Wilson, L., 1942. *The Academic Man*. Oxford University Press, New York.
- Young, N.S., Ioannidis, J.P.A., Al-Ubaydli, O., 2008. Why current publication practices may distort science. *PLoS Med.* 5 (10), e201.
- Zhang, C.-T., 2009. The e-index, complementing the h-index for excess citations. *PLoS One* 4 (5), e5429.

Internationalization in English? Problematizing the role of language in higher education internationalization

Johanna Ennser-Kananen^a and Taina Saarinen^b, ^a Department of Language and Communication Studies, University of Jyväskylä, Jyväskylä, Finland; and ^b Finnish Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	167
Ideologies of internationalization	167
Language and internationalization: the case of English	168
The multilayered dominance of English and anglophone contexts	169
English in non-anglophone contexts	169
English, internationalization and Empire	170
English for the (first) world? The pitfalls of ELF (and other) approaches	170
Conclusion: ways forward	171
References	172

Introduction

In recent higher education policy debates, internationalization tends to be a relatively unquestioned policy phenomenon that is described by a seemingly straight-forward adding of an “international, intercultural or global dimension into the purpose, functions or delivery of postsecondary education” (Knight, 2003). Building on this definition by Jane Knight, de Wit & Hunter (2015, p.3) tied internationalization to a societal purpose and defined it as “the *intentional* process of integrating an international, intercultural or global dimension into the purpose, functions and delivery of post-secondary education, *in order to enhance the quality of education and research for all students and staff, and to make a meaningful contribution to society* [emphasis in the original]”. Ultimately, both definitions fail to acknowledge the ideological nature of internationalization and the layers of historical, societal, and cultural power structures embedded in it (see for instance Brooks and Waters 2011). As pointed out by Jokila (2020), despite the recent increase in activities that are tagged as “international” and the political measures undergirding those activities, internationalization is still depicted as a neutral concept, policy, and activity.

Internationalization of higher education and (student) mobility tend to be used interchangeably (Knight, 2012), as internationalization policies have often been operationalized as mobility and co-operation programs. Newer debates on internationalization and mobility have, however, demonstrated the unsuitability of homogeneous typologizations of separate categories of student mobility, labor market mobility, family based mobility, refugee status mobility etc., as these do not take into account the varied and hybrid roles and identities that mobile individuals and groups have over time (see for instance Robertson, 2013, for a discussion of the complexity of the student - migrant categories). Consequently, the shortcomings of current conceptualizations of the role of language in internationalization relate to un(der)addressed issues in internationalization debates that this article hopes to bring to the fore.

Ideologies of internationalization

Motivations for internationalization have shifted since the Second World War as internationalization gradually started to become a systematic part of the activities of higher education institutions. Post- World War II internationalization has been described as “benevolent laissez-faire” (Baron, 1993; cited by de Wit and Merckx, 2012), implying benevolence and unregulated mobility for those with access and resources to be mobile. In the post-war period, internationalization had strong peacekeeping undertones particularly in Europe that was recovering not just from a devastating war but also from genocides and unprecedented refugee numbers (see also de Wit, 2020).

Toward the 1960s, the ideologies behind internationalization became more complex. Post-war humanitarian idealism increasingly gave way to polarized ideologies and an explicit politicization of internationalization during the Cold War (Altbach and de Wit, 2015). In addition, students from recently independent countries such as India or several West-African countries began to increasingly enter higher education, often in the countries of their former colonizing powers. (Ploner and Nada, 2020). A further factor influencing post-war internationalizations was the active US cultural foreign policy as a part of the post-war “Pax Americana”, or post-World War US policies aimed at solidifying its dominant position, resting on US economic and military dominance as well as the ideological appeal of the US and the international institutions built by it since 1945 (Layne, 2012). This policy emerged particularly in Western Europe and Eastern Asia in the form of exchange programs such as Fulbright (see for instance Haberland, 2009). As a result, English started to take over as a language of internationalization.

During the 1970s and 1980s, economic motivations began to drive the systematic development of higher education internationalization, particularly in Europe but also elsewhere (de Wit and Merckx, 2012), strengthening the understanding of higher education as an economic commodity (de Wit 2020). Demands for public accountability and external evaluation since the global economic depression of the 1970s and the ensuing cuts in public funding, coupled with policies that aimed to ensure HE “quality” and “comparability”, set the tone also for internationalization policies. Internationalization became a more systematic activity, and was largely motivated by educational and labor market mobilities particularly in European integration (see for instance de Wit and Hunter, 2015; Teichler, 2004). This development further emphasized the position of English as the language of the post-nationalist global economy (Pujolar, 2017).

The global political turn in the 2000s toward new nationalism and populism, combined with protectionist agendas of creating barriers to international (student) mobility also changed the ways in which “internationalization of higher education” is understood and enacted. One example of this are the emerging concerns for the national languages that are presented as being threatened by increasing use of English (see Saarinen, 2020; Soler, 2019). As national authorities have started criticizing the extensive use of English, its use has been anticipated to be “reaching a ceiling” in at least European higher education, (Altbach and de Wit, 2018). Another example is presented by Mathies and Weimer (2018) who discuss the hierarchization of international students in new nationalism. Discussing Chouhada et al. (2012) taxonomy (cited by Mathies and Weimer 2018, 149–150), they discuss the role of neo-racism (i.e. racism based on a proxy such as language or nationality; Lee and Rice, 2007; Lee et al., 2016) in international student mobility. The dimensions of international students’ academic credentials of preparation (high/low) and the financial resources at their disposal (high/low) can be used to categorize students into four types. Traditionally, the thinking has been that the “high flyers” in the high academic preparation/high financial resources category have not faced much difficulty in international mobility. However, adding a third dimension of high/low desirability based on origin makes visible the racialized nature of international student categorizations. (New) nationalist factors (leading to racist categorizations and policies) impact international student mobility more than higher education research and policy making have previously recognized.

In addition to the macro trends in internationalization, depicted above, also more short-term developments impact internationalization trends globally and regionally. Ecological and health factors (natural disasters, pandemics, etc.), cultural factors (for instance developments in “cultural foreign policy” such as the US Fulbright or Chinese Confucius Institute schemes), demographics, and academic rationales (which are by no means detached from the others) all have their influence on internationalization policies (de Wit and Altbach, 2021). These various, historically and politically layered motivations of internationalization emphasize the need to unpack the seemingly benevolent or neutral purposes of internationalization and ask, whose benefit we are ultimately talking about.

An example of the complexity of benefits of mobilities is the concept of “internationalization at home” that has been coined to refer to international interaction in one’s own environment (Nilsson, 2000). The concept of internationalization at home aims at offering an international learning experience to a wider group of students who remain at their home campuses by internationalizing the curriculum in different ways. As the expectation is that at least a part of internalization at home requires interaction with incoming staff and students, it still hinges on someone else’s mobility, thus raising important questions not only about “who is able to be mobile?” but also “who has to be mobile?”

As mentioned above, the term “international” assumes “national” units, and is hierarchical in nature, as made evident by the discourses of migrant versus expatriate or mobility versus migration (see Kunz, 2016, for a review of privileged categories of migration). Internationalization is, thus, conceptually linked with the nation state and its interests, and carries recycled discourses and the ensuing baggage of “nation” with it (Saarinen, 2020). This national(ist) nature of international is also reflected in unproblematised monolingual assumptions about mobility and internationalization, as we invariably discuss “foreign” students participating in “international” mobility in a “host” country in a particular language, thus homogenizing internationalization.

The current policies and practices of internationalization as a part of global higher education trends are the result of a variety of entangled processes that mix neoliberal ideologies and new nationalist protectionism with (allegedly) humanitarian values of cosmopolitanism, multiculturalism, and diversity. The normative and normalizing assumptions of internationalization as a humanitarian activity have been entangled with the values of competition and market dynamics of higher education. This, in turn, normalizes the meritocratic view of internationalization as individual and systemic competition, deflecting our view from the inequities embedded in the system (Bamberger et al., 2019, p. 211–212.).

Language and internationalization: the case of English

In internationalization of higher education, language may appear as a non-issue, as English tends to be taken for granted in many contexts (Jenkins and Mauranen, 2019). This *internationalist* view is often linked with discourses of national competition in international markets, as nation states develop systematic internationalization of their higher education systems to promote their international competitiveness (Hultgren et al., 2014; Saarinen, 2020). The *culturalist* view, in turn, focusses on safeguarding the national culture and national language, often from alleged threats brought on by global cooperation and specifically the use of English (Hultgren et al., 2014; Phillipson, 2015; Salö, 2014). Interestingly, the culturalist discourses are produced either from right wing (populist new nationalism) or left wing (critique of globalist threats to welfare state) approaches (see Salö, 2014).

As language is a gatekeeper for international students, who are often required to take standardized language tests to demonstrate adequate language skills, it is important to examine the ostensibly neutral or unaddressed ways language and language skills occupy

in internationalization contexts. For example, the ideological nature of language in internationalization context surfaces in different English language requirements that are based on political rather than linguistic criteria (e.g., Whose English is considered adequate? - see [Saarinen and Nikula, 2013](#)) or in the motivations for so-called accent reduction or accent modification courses offered to international students (and clearly geared at particular populations) in the United States ([Enns-Kananen et al., 2021](#)).

In this section, we will focus on the role of English in internationalization of higher education, while recognizing that neither English nor the linguistic landscape of higher education are homogeneous.

The multilayered dominance of English and anglophone contexts

English appears as an unproblematised language in internationalization on at least three levels.

First, economically developed English-dominant countries such as the USA, the UK, Australia, Canada, and New Zealand have long been core destinations for foreign students pursuing higher education abroad ([Barnett et al., 2016](#); [Gümüş et al., 2020](#)). A vast majority of students from China, the largest group of mobile students, choose an English speaking country for their international studies ([Barnett et al., 2016](#)). Mobility to tuition/fee-charging countries is also an economic issue: International students are a major source of revenue in for instance Australia and other Anglophone countries, and are subject to tuition in some systems, even if locals are not ([Cantwell, 2015](#)). Also in Nordic countries, where study fees are a relatively new phenomenon, the discourse of commodification of higher education has accompanied the introduction of study fees for international students ([Saarinen 2020](#)).

Receiving hubs of international students are also emerging in non-Anglophone Eastern Asian and European contexts, where either English Medium Instruction (EMI) is set up to attract international students, or where languages other than English are used for purposes of mobility ([Chen and Barnett, 2000](#); [Barnett et al., 2016](#); [Galloway et al., 2020](#)). In contrast, in terms of numbers of received students, African and West Asian countries remain minor destinations in internationalization networks. As [Chen and Barnett \(2000, 435\)](#), have pointed out, this suggests that hegemonies of internationalization are run along the same lines as global economic and political hegemonies. In all these contexts, language plays an important role in creating networks and directing student flows to major receivers for instance from countries who are political allies or former colonies ([Barnett et al., 2016](#)).

Second, research on international student mobility seems to center around Anglophone researchers and contexts ([Gümüş et al., 2020](#)), which creates a conceptual bias. This, combined with the dominant status that countries such as the UK, US, Australia, and (Anglophone) Canada hold in the field of international mobility ([Barnett et al., 2016](#)) normalizes the hegemonic position of Anglophone countries in internationalization processes. In other words, both the activity of internationalization and the research on it are largely rooted in so-called “western” and specifically Anglophone contexts (see [de Wit and Hunter, 2015](#); [Jokila, 2020](#); [Teichler, 2004](#)).

Third, and related to the previous point, also research metrics and databases such as Web of Science that are used to study the conceptual basis of internationalization are biased toward English language research ([Gümüş et al., 2020](#); [Kuzhabekova et al., 2015](#)). In other words, internationalization research is likely overlooking studies on internationalization and mobility that are published in other languages than English and/or empirically conducted in non-English dominant areas. In short, the Anglophone bias of our view likely distorts our view both empirically and conceptually.

English in non-anglophone contexts

As mentioned above, in the higher education context of many non-Anglophone-dominant countries, English Medium Instruction (EMI) programs have been created with the goal of attracting international students. For instance in Europe and Eastern Asia, universities have been politically encouraged to promote policies of internationalization, and have consequently added EMI programs to their curricula ([Galloway et al., 2020](#)).

English medium instruction (EMI) is driven by a variety of motivations in non-Anglophone higher education contexts (see articles in [Dimova et al., 2015](#)). For instance, EMI may be offered in order to attract international students for so-called “brain gain”, defined as the problematic notion of importing in-demand knowledge and skills. This policy relates to the ongoing emphasis in higher education on public accountability and so-called “higher education excellence” (see section on ideologies of internationalization above).

Another motivation for offering EMI is to offer international experiences for local students (*internationalization at home*; see section on ideologies of internationalization above) and, relatedly, help prevent “brain drain” by helping keep local students in their home countries and institutions ([Soler, 2019](#)).

A third motivation, particularly in contexts where student cohorts are waning, is to compensate for a decrease in local student numbers with an increase in international (possibly fee-paying) ones, as is for instance the case in Estonia ([Soler, 2019](#)) and Latvia ([Kibbermann, 2017](#)). In the context of European integration, an increase in EMI programs has gone hand in hand with tuition waivers for EU students but introducing or increasing tuition for others, which creates inequities among different groups of international students.

In the contexts where English has been introduced as language of internationalization, it has sparked a variety of discourses, ranging from benevolent “language of internationalization” to English as “killer language” that poses a threat to other academic languages ([Bull, 2012](#)). This reflects on the societal, political, and cultural role of English as far more than “just a language that can be taught” (see [Pennycook, 2017](#) for an extensive discussion). Although the *internationalist* and *culturalist* (see section on

language and Internationalization above) discourses of internationalization tend to revolve around the role of English and its relationship with the local language(s), it is important to note that such discussions about the role of English in internationalization cannot be reduced to the debate of local languages versus English alone (see Salö, 2014, on the Swedish situation or Saarinen 2020 on the Finnish one), as many socio-political, historical, and linguistic factors shape them.

Although we cannot disentangle the discourses of English in internationalization of higher education in depth in this chapter, we contend that in order for this discussion to be fruitful and productive, both internationalization and English need to shed their cloaks of innocence, educational quality, or global benevolence (Bruthiaux, 2003). Instead, we have to recognize the power dynamics, especially in the shape of nationalism, neoliberalism, and Empire (see below), that are intertwined in these processes.

English, internationalization and Empire

While the unique and ambivalent status of English (or rather, Englishes) in our societies has to be recognized (Canagarajah, 1999; Saarinen and Ennser-Kananen, 2020), it is important to critically examine the equation of English with “international” and the celebration of English as a worldwide lingua franca, beginning with the recognition that the global spread of English is not a coincidence but rather the outcome of a colonial project. Important work in this area has begun to gain traction in TESOL (Teaching English to speakers of other languages), which is closely related and partly overlaps with EMI contexts. Motha (2014) has examined the ongoing involvement of the field of TESOL in processes of coloniality through a lens of what she calls Empire.

Although scholars from all disciplines readily acknowledge colonialism as a historical fact, its ongoing processes and its intertwinedness with English in educational contexts are far from consensual. In her description of colonialism, Motha (2014) stresses that, “(t)he insertion of English into the mouths and hearts of the (colonized) population was often an inherent part of the project”, with the consequence that in the field of TESOL, “a colonial imprint is stamped into our profession and remains there indelibly even in a modern-day context in which English is presented as intensely desirable through global media and the English-teaching industry.” (p. 27). To speak about this continued form of colonialism, Motha draws on Negri and Hardt (2000) in her use of *Empire* (with capital E) to denote a contemporary form of colonial activity, which is strongly intertwined with neoliberalism and may be a “less willfully intentional” (p. 31) and thus less recognizable force than historical colonialism. Importantly, Empire works through “constructing the desirability of ‘Western-ness’ and English and (keeps) legitimate ownership of English associated with Whiteness and ‘Western’ identity” (p. 31). It is through this lens we suggest looking at the intersection of internationalization and English.

Many others have pointed to Empire in the context of language and particularly English education. For instance, in his 1998 book *English and the Discourses of Colonialism*, Pennycook (1998) shows how colonial discourses are by no means restricted to the historical contexts of their origins but rather very much ongoing processes that form the foundation of so-called “Western” thinking and being. For instance, in an effort to reject notions of English as neutral global lingua franca (see the next section on English as Lingua Franca), he draws comparisons between colonial and modern descriptions of English as a superior language, showing and refuting discourses of supremacy (e.g. linguistic purity, richness, openness, and versatility) of both the language and, by extension, its speakers. For today’s professionals and students in English learning/teaching contexts, it is important to understand that such discourses are not based on linguistic realities but rather an (not always intentional) enactment of Empire that justifies itself by constructing an ideology of superiority of some languages and varieties and its speakers over others. Resisting English altogether is not always a feasible or beneficial solution (Pennycook, 1998), but the promotion and establishment of counter discourses to Empire is one productive avenue to pursue.

In all, a large amount of research (Hsu, 2017) has explicated the entrenchedness of Empire and global capitalism in the learning and teaching of English, and how these dynamics are still ongoing and palpable in today (Canagarajah, 1999; Kachru, 1992). To engage in the topic of English and internationalization in a meaningful way, we therefore suggest that rather than taking English for granted as *the* international language, it is more helpful to understand this equation as a discourse. As Kubota (2012) puts it,

The idea that English is an international language can be viewed as a discourse rather than an absolute fact, shaping people’s consciousness, social practice and institutional policies, including the current heavy focus on teaching English worldwide.

(Kubota, 2012, p. 55).

This discourse can then be examined for its ideological underpinnings and implications, including the participation of “English as the international language” in processes of coloniality and discrimination.

English for the (first) world? The pitfalls of ELF (and other) approaches

Several noteworthy attempts have been undertaken to recognize the unique position of English globally while overcoming its tendency to hierarchize speakers by variety, accent, or proximity to so-called “nativeness”. In the newer edition of his book (2017), Pennycook criticizes conceptualizations of English as “international language”, in particular English as a lingua franca (ELF), which conceptualizes English among so-called non-native speakers as a variety in its own right (Seidlhofer, 2005), World Englishes, which recognizes varieties of English in many countries worldwide, and theories of linguistic imperialism. In particular, he takes issue with the failure of ELF and World English approaches “to address power and inequality in any adequate way” and

with the inability of linguistic imperialism approaches “to construct a plausible model of language” (p. viii). In the same vein, Kubota’s (2012) critique of (allegedly) anti-normative approaches to English as an international language underlines the failure of World English approaches to shed their nation-statist paradigm and cultural essentialism; the risk of ELF to perpetuate the marginalization and homogenization of non-native speakers; the insensitivity of linguistic imperialism toward social identities; and the shortcomings of approaches that promote the hybridity and fluidity of identities while ignoring that stability and fluidity coexist and are shaped by dominant power dynamics.

English as a lingua franca (ELF) has been recognized for its intention to shed native-speakerism and uplift those who learn English as an additional language. However, ELF as variety of English and related pedagogies have also been met with strong criticism. For instance, Park and Wee (2011) have called for a more power-sensitive approach, explaining that

simply creating an egalitarian market where everyone has equal access to valued resources or claims of legitimacy is not possible; and for this reason, endeavors such as the ELF project need to be greatly more sensitive to the political consequences of formulating and promoting such alternative markets (Park and Wee, 2011, p. 366).

Additionally, O’Regan’s “immanent critique” (2014, p. 533) of the ELF concept points to inconsistencies in its theorization and critiques its reliance on neoliberal ideologies that undergird a naive approach to globalization, which “has brought about fluidity and hybridity in language and where in global communication flows everybody is accommodating to everybody else in a playful postmodern manner” (O’Regan, 2014, p. 543). Although this critique has been responded to thoroughly (Baker et al., 2015) and although, as Baker et al. (2015) point out, issues of power have been in the focus of several ELF scholars, the authors do not speak to the ability or willingness of ELF approaches to acknowledge injustices that are rooted in sociohistorical processes linked to coloniality. For instance, the native-non-native speaker divide that still plagues much of language education and interaction also in HE contexts is a problem against which ELF has been called into existence. A proclamation of equity among all English users can only be substantiated if the roots of the “native speaker” ideal in colonial processes and thinking (Pennycook, 1998) are acknowledged, and language scholars actively participate in dismantling such ongoing processes. Particularly in combination with the problematic concept of “internationalization”, in connection to which English is often uncritically positioned as “international language”, an uncovering of the socio-historical and socio-political power dynamics and their entanglement with Empire are paramount.

Conclusion: ways forward

We started by pointing out that Knight’s (2003) and de Wit and Hunter’s (2015) definition of internationalization fail to take into account systemic inequities in higher education. We conclude by offering some suggestions for addressing these inequities in relation to language and internationalization efforts. In our suggestions, we keep in mind – or maybe naively assume – that equitable, societally beneficial, and high quality education for all remains the goal of higher education and thereby internationalization. We thus come back to de Wit and Hunter (2015) and their proclaimed goal of internationalization to “*enhance the quality of education and research for all students and staff, and to make a meaningful contribution to society*” (p. 3), and argue that ideas and initiatives exist that are better equipped to move higher education toward these goals than internationalization in its current form.

Ethnic studies, one promising attempt of centering student populations that have historically been in the margins of education (Banks, 2012; Sleeter, 2011), has proven to be an effective way not only to center those students and their experiences, but also to increase attendance as well as educational achievement (Dee and Penner, 2017). Centering the histories, experiences, and identities of minoritized (and often racialized and colonized) populations is not merely an initiative for those groups. Such programs offer the development of crucial life skills, including those promoted by internationalization, also to members of dominant groups, including the opportunity to develop socio-historical awareness, anti-oppressive consciousness and activism, and a critical, reflexive stance toward themselves, their communities, and society at large. The goals that have (allegedly) been those of internationalization could be pursued via ethnic studies.

Second, research from TESOL contexts (Hsu, 2017) that is applicable to higher education studies helps identify strategies for challenging the perpetuation of the coloniality of English, starting with a theoretical reconceptualization of the field, which includes a pushing back against discourses of English as neutral language and its color-evasiveness. We suggest that such self-reflexive practice also centers and critiques racialization discourses and raciolinguistic ideologies (Flores and Rosa, 2015) that have sustained and normalized colonial practices in the academy. In addition, we amplify Hsu’s call for critically employing methods and curriculum with the goal of building on local and communal knowledges, including translanguaging (García and Wei, 2014), and legitimize learners from all backgrounds as rightful knowers and multilingual speakers. Specifically, there is a need in TESOL as well as higher education research to address the colonial legacy of the fields and build on this historical awareness to develop and practice anti-normative, anti-colonial pedagogies and practices.

Concrete ideas exist to work toward such pedagogies and practices, for example by Kubota (2012). Her concept of “border-crossing communication” (p. 63) aims to promote students’ sensitivity toward questions of power and privilege and fosters equitable encounters across different languages and contexts. Similarly, we find Motha’s (2014) call for provincialized English

useful, which forefronts the effects of Empire on learning and education. In her words, “(t)o provincialize English would mean that inherent in the learning of English would be an intense awareness of the effects on English’s colonial and racial history on current-day language, economic, political, and social practices”. She continues by explaining that such an approach “would furthermore examine and critique the mechanisms that sustain the invisibility of race and empire in English language teaching and would explore possibilities for transformation and agency” (p. 129). Applied to higher education studies, such an approach could transform the areas and processes that are currently known as internationalization.

In the end, we come back to the purpose of internationalization. As a concept that gained traction in the 19th Century industrializing, modernizing, Western emerging nation states (see Holmila and Ihalainen, 2022), we argue that internationalization in its current form as a normative, ostensibly neutral and ultimately nation state-centered activity has reached its expiration date. In order to challenge inequities, internationalization cannot be treated as an isolated policy activity in higher education but instead needs to be embedded in a larger process of improving higher education for all.

References

- Altbach, P.G., de Wit, H., 2015. Internationalization and global tension: Lessons from history. *J. Stud. Int. Educ.* 19 (1), 4–10.
- Altbach, P.G., de Wit, H., 2018. The Challenge to Higher Education Internationalization, vol. 27. University World News.
- Baker, W., Jenkins, J., Baird, R., 2015. ELF researchers take issue with ‘English as a lingua franca: an immanent critique’. *Appl. Linguist.* 36 (1), 121–123.
- Bamberger, A., Morris, P., Yemini, M., 2019. Neoliberalism, internationalisation and higher education: Connections, contradictions and alternatives. *Discourse* 40 (2), 203–216.
- Banks, J.A., 2012. Ethnic studies, citizenship education, and the public good. *Intercult. Educ.* 23 (6), 467–473.
- Barnett, G.A., Lee, M., Jiang, K., Park, H.W., 2016. The flow of international students from a macro perspective: a network analysis. *Compare* 46 (4), 533–559. <https://doi.org/10.1080/03057925.2015.1015965>.
- Baron, B., 1993. The politics of academic mobility in Western Europe. *High. Educ. Pol.* 6 (3), 50–54.
- Brooks, R., Waters, J., 2011. Student Mobilities, Migration and the Internationalization of Higher Education. Palgrave, Basingstoke.
- Bruthiaux, P., 2003. Squaring the circles: issues in modeling English worldwide. *Int. J. Appl. Ling.* 13 (2), 159–178.
- Bull, T., 2012. Against the mainstream: universities with an alternative language policy. *Int. J. Sociol. Lang.* 2012 (216), 55–73.
- Canagarajah, S., 1999. Resisting linguistic imperialism in English teaching. Oxford University Press, Oxford.
- Cantwell, B., 2015. Are international students cash cows? Examining the relationship between new international undergraduate enrollments and institutional revenue at public colleges and universities in the US. *J. Int. Stud.* 5 (4), 512–525.
- Chen, T., Barnett, G.A., 2000. Research on international student flows from a macro perspective: a network analysis of 1985, 1989 and 1995. *High. Educ.* 39 (4), 435–453.
- Choudhary, R., Orosz, K., Chang, L., 2012. Not all international students are the same: understanding segments, mapping behavior. World Education Services, New York.
- Dee, T.S., Penner, E.K., 2017. The causal effects of cultural relevance: Evidence from an ethnic studies curriculum. *Am. Educ. Res. J.* 54 (1), 127–166.
- de Wit, H., Altbach, P.G., 2021. Internationalization in higher education: global trends and recommendations for its future. *Policy Rev. High. Educ.* 5 (1), 28–46.
- de Wit, H., Hunter, F., 2015. The future of internationalization of higher education in Europe. *Int. High. Educ.* 83, 2–3.
- de Wit, H., Merckx, G., 2012. The History of Internationalization of Higher Education. The SAGE handbook of international higher education, Thousand Oaks, pp. 43–59.
- de Wit, H., 2020. Internationalization of higher education: the need for a more Ethical and qualitative approach. *J. Int. Stud.* 10, i–iv.
- Dimova, S., Hultgren, A.K., Jensen, C. (Eds.), 2015. English-medium Instruction in European Higher Education. Walter de Gruyter, Berlin.
- Enns-Kananen, J., Halonen, M., Saarinen, T., 2021. “Come join us and lose your accent!” Accent modification courses as hierarchization of international students. *J. Int. Stud.* 11 (2), 322–340.
- Flores, N., Rosa, J., 2015. Undoing appropriateness: raciolinguistic ideologies and language diversity in education. *Harv. Educ. Rev.* 85 (2), 149–171.
- Galloway, N., Numajiri, T., Rees, N., 2020. The ‘internationalisation’, or ‘Englishisation’, of higher education in East Asia. *High. Educ.* 80 (3), 395–414.
- García, O., Wei, L., 2014. Translanguaging: Language, Bilingualism and Education. Palgrave, Basingstoke.
- Gümüş, S., Gök, E., Esen, M., 2020. A review of research on international student mobility: Science mapping the existing knowledge base. *J. Stud. Int. Educ.* 24 (5), 495–517.
- Haberland, H., 2009. English—the language of globalism. *Rask: Internationalt Tidsskrift for Sprog Og Kommunikation* 30, 17–45.
- Holmila, A. and Ihalainen, P., 2022. Debating internationalisms: contexts, concepts and historiography. In: P. Ihalainen & A. Holmila (eds.), *Nationalism and Internationalism Intertwined: A European History of Concepts Beyond the Nation State*. Berghahn, New York, pp. 1–19.
- Hsu, F., 2017. Resisting the coloniality of English: a research review of strategies. *CATESOL J.* 29 (1), 111–132.
- Hultgren, A.K., Gregersen, F., Thøgersen, J., 2014. English at Nordic universities: ideologies and practices. In: Hultgren, A.K., Gregersen, F., Thøgersen, J. (Eds.), *English in Nordic Universities. Ideologies and Practices*. John Benjamins, Amsterdam, pp. 1–25.
- Jenkins, J., Mauranen, A. (Eds.), 2019. Linguistic Diversity on the EMI Campus: Insider Accounts of the Use of English and Other Languages in Universities within Asia, Australasia, and Europe. Routledge.
- Jokila, S., 2020. Re/configuring National Interests in the Global Field of Higher Education. International Student Recruitment Policies and Practises in Finland and China. University of Turku, Turku.
- Kachru, B.B., 1992. World Englishes: approaches, issues and resources. *Lang. Teach.* 25 (1), 1–14.
- Kibbermann, K., 2017. Responses to the internationalisation of higher education in language policies of Estonia and Latvia. *Eesti ja soome-ugri keeleteaduse ajakiri. J. Est. Finno-Ugrian Linguist.* 8 (1), 97–113.
- Knight, J., 2003. Updating the definition of internationalization. *Int. High. Educ.* 33 (6), 2–3.
- Knight, J., 2012. Student mobility and internationalization: trends and tribulations. *Res. Comp. Int. Educ.* 7 (1), 20–33.
- Kubota, R., 2012. The politics of EIL: toward border-crossing communication in and beyond English. In: Matsuda, A. (Ed.), *Principles and Practices for Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 55–69.
- Kunz, S., 2016. Privileged mobilities: Locating the expatriate in migration scholarship. *Geogr. Compass* 10 (3), 89–101.
- Kuzhabekova, A., Hendel, D.D., Chapman, D.W., 2015. Mapping global research on international higher education. *Res. High. Educ.* 56 (8), 861–882.
- Layne, C., 2012. This time it’s real: the end of unipolarity and the Pax Americana. *Int. Stud. Q.* 56 (1), 203–213.
- Lee, J.J., Jon, J., Byun, K., 2016. Neo-racism and neo-nationalism within East Asia: the experiences of international students in South Korea. *J. Stud. Int. Educ.* 21 (2), 136–155.
- Lee, J.J., Rice, C., 2007. Welcome to America? International student perceptions of discrimination. *High. Educ.* 53 (3), 381–409.
- Mathies, C., Weimer, L., 2018. A changing narrative for international students? : the potential influence of Brexit and Trump. In: Proctor, D., Rumbley, L. (Eds.), *The Future Agenda for Internationalization in Higher Education: Next Generation Insights into Research Policy and Practice*. Routledge, Oxon, pp. 144–152.
- Motha, S., 2014. *Race, Empire, and English Language Teaching: Creating Responsible and Ethical Anti-racist Practice*. Teachers College Press, New York.
- Negri, A., Hardt, M., 2000. *Empire*. Harvard University Press, Harvard.

- Nilsson, B., 2000. Internationalising the curriculum. In: Joris, M., Otten, M., Nilsson, B., Teekens, H., Wächter, B. (Eds.), *Internationalisation at Home: A Position Paper*. EAIE, Amsterdam, pp. 21–27.
- O'Regan, J., 2014. English as a lingua franca: an immanent critique. *Appl. Linguist.* 35 (5), 533–552.
- Park, J.S.Y., Wee, L., 2011. A practice-based critique of English as a Lingua Franca. *World Englishes* 30 (3), 360–374.
- Pennycook, A., 1998. *English and the Discourses of Colonialism*. Routledge, London.
- Pennycook, A., 2017. *The Cultural Politics of English as an International Language*. Routledge, London.
- Phillipson, R., 2015. English as threat or opportunity in European higher education. In: Dimova, S., Hultgren, A.K., Jensen, C. (Eds.), *English-medium Instruction in Higher Education in Europe*. Mouton de Gruyter, Berlin, pp. 19–42.
- Ploner, J., Nada, C., 2020. International student migration and the postcolonial heritage of European higher education: perspectives from Portugal and the UK. *High. Educ.* 80, 373–389.
- Pujolar, J., 2017. Post-nationalism and language commodification. In: Tollefson, J., Pérez-Milans, M. (Eds.), *The Oxford Handbook of Language Policy and Planning*, Oxford University Press, Oxford.
- Robertson, S., 2013. *Transnational Student-Migrants and the State: The Education-Migration Nexus*. Palgrave, Basingstoke.
- Saarinen, T., Ennser-Kananen, J., 2020. Ambivalent English: what we talk about when we think we talk about language. *NJES* 19 (3), 115–129.
- Saarinen, T., 2020. *Higher Education, Language and New Nationalism. Recycled Histories*. Palgrave, Cham.
- Saarinen, T., Nikula, T., 2013. Implicit policy, invisible language: policies and practices of international degree programmes in Finnish higher education. In: Doiz, A., Lasagabaster, D., Sierra, J. (Eds.), *English-medium Instruction at Universities: Global Challenges*. Multilingual Matters, Bristol, pp. 131–150.
- Saló, L., 2014. Language ideology and shifting representations of linguistic threats. In: Hultgren, A.K., Gregersen, F., Thøgersen, J. (Eds.), *English in Nordic Universities. Ideologies and Practices*. John Benjamins, Amsterdam, pp. 83–109.
- Seidlhofer, B., 2005. English as a lingua franca. *ELT J* 59 (4), 39–341. <https://doi.org/10.1093/elt/cci064>.
- Sleeter, C.E., 2011. *The Academic and Social Value of Ethnic Studies: A Research Review*. National Education Association Research Department, Washington, DC.
- Soler, J., 2019. *Language Policy and the Internationalization of Universities: A Focus on Estonian Higher Education*. Walter de Gruyter, Berlin.
- Teichler, U., 2004. The changing debate on internationalisation of higher education. *High. Educ.* 48 (1), 5–26.

Academic freedom protections in national and international law

Neal H. Hutchens^a, Macey L. Edmondson^a, and Frank J. Fernandez^b, ^aThe University of Mississippi, Department of Higher Education, University, MS, United States; and ^bThe University of Florida, College of Education, Gainesville, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Contemporary roots of academic freedom	175
US universities and academic freedom—tenure and constitutional protections	175
From promising start to legal morass—academic freedom and the US Constitution	176
Increasingly Fragile State of tenure in United States	177
Academic freedom as a human right	178
Challenges and opportunities for academic freedom	179
Conclusion	181
References	181

Formulating a consistent definition for academic freedom, even from within the perspective of a single nation or academic tradition, presents a difficult task. Despite challenges in delineating a singular academic freedom definition, the issue of providing scholars with sufficient intellectual independence to carry out their professional duties in the realms of research, teaching, and service to their students, colleagues, and institutions and to society more broadly has been identified as a salient concern for colleges and universities throughout the world (e.g., Kinzelbach et al., 2021; Altbach, 2013). For example, academic freedom has been identified as a necessary ingredient for the building of high quality higher education institutions, with a “high degree of academic freedom ... particularly important for the social sciences and humanities, but all fields benefit [ing] from freedom of inquiry and a sense that the university is committed to the free expression of ideas” (Altbach, 2013, p. 137). The Scholars at Risk Network and Global Public Policy Institute have described academic freedom as encompassing a “universal right and essential to quality education, teaching and research. It is a driver of innovation, enhances the capacity of scholars and students to acquire and generate knowledge, and thereby protects societies’ capacity for self-reflection” (Kinzelbach et al., 2021, p. 4).

A 2020 United Nations report by David Kaye while serving as the U.N. Special Rapporteur for Freedom of Opinion and Expression espoused the importance of academic freedom as helping to maintain “democratic life,” “scientific progress,” and “human development” (Kaye, 2020, p. 4). Kaye’s report looked to language from the 1997 United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation Concerning the Status of Higher Education Teaching Personnel as to the contours of academic freedom:

Members of the academic community, individually or collectively, are free to pursue, develop and transmit knowledge and ideas, through research, teaching, study, discussion, documentation, production, creation or writing. Academic freedom includes the liberty of individuals to express freely opinions about the institution or system in which they work, to fulfil their functions without discrimination or fear of repression by the State or any other actor, to participate in professional or representative academic bodies, and to enjoy all the internationally recognized human rights applicable to other individuals in the same jurisdiction.

Quinn and Levine (2014) offer a similar assessment of the role of academic freedom and independent higher education institutions, including as a space for the exercise of democratic values, stating:

At their best, higher education communities are models and teachers of democratic values. They model and pass into society the skills and knowledge necessary for democratic value systems to function properly, most notably a democratic “knowledge-over-force” principle that rejects violence and force as determinants of outcomes, in favour of process, evidence, reasoned discourse and quality (p. 214).

They also note that higher education institutions do not operate in a values vacuum and that institutions and members of academic communities, including as citizens in general, must be able to exist in overall political and social environments that are conducive to supporting and maintaining academic freedom.

Beiter et al. (2016) describe how academic freedom provides an avenue to support the advancement of knowledge and what they describe as the “discovery of truth” (p. 128). Furthermore, the authors point out the distinctive role of universities in advancing knowledge, an outcome that requires sufficient individual academic freedom and institutional autonomy. At the individual level, the authors note the role of higher education personnel in promoting “intellectual independence” in support of “ethical individualism” (Beiter et al., 2016, p. 111). For the concept of ethical individualism, they look to Ronald Dworkin, whose work dealt with the idea of individuals having the freedom to follow personal conviction in determining the ingredients

and purposes of a successful life (Beiter et al., 2016, p. 130). For the authors, the “ultimate” rationale to safeguard academic freedom emanates from the key role of higher education institutions and personnel in supporting and advancing human dignity (Beiter et al., 2016, p. 132).

This entry reviews the concept of academic freedom as a human right, with special attention paid to international agreements and proclamations, such as through the United Nations. Coverage is also given to academic freedom in the United States and the use of tenure as an employment arrangement intended to safeguard faculty academic freedom (Sun and Hutchens, 2019). The development of US research universities provides a useful example of institutions where academic freedom came to be viewed as a core value of academic life and institutional functioning (see, e.g., Marginson, 2014; Altbach, 2013). However, as covered in the section, rhetoric from academic leaders or language in institutional mission statements do not always match with reality. The entry also examines current challenges to academic freedom around the globe.

Contemporary roots of academic freedom

Simon Marginson (2014) argues that the development of “academic freedom is best understood not as an abstract universal principle or an ideal state of being but as concrete university practices nested in specific relational environments” (p. 24). The institutional practices that underpin academic freedom are institutional autonomy and the professionalization of the professoriate, which trace back in nascent forms to Islamic madrasas and palace schools in the Ottoman empire (Kreckel, 2017). Sociologist David Baker argues that medieval universities in Western Europe developed institutional autonomy by positioning themselves between the Church and state. External institutions sought “to instrumentalize the university,” but “in struggles with first the Church, then European absolutist rulers, then the modern nation-state, and perhaps lastly with large-scale capitalism, the basic model of universal inquiry has become institutionally stronger within the university” (Baker, 2014, p. 70). As medieval universities escaped ideological capture by religious authorities, monarchs, and governments, they cultivated legitimacy for themselves and their operators.

Beiter et al. (2016) trace the development of academic freedom in contemporary higher education settings to developments in the Enlightenment that found special expression in Prussia, including the importance of the ideas of “freedom of teaching (*Lehrfreiheit*), freedom of learning (*Lernfreiheit*), [and] the unity of teaching and research (*Einheit von Lehre und Forschung*) (in terms of which both students and professors are involved in a joint endeavor of discovering the truth)” (p. 114) (see also, e.g., Altbach, 2013; Sun and Hutchens, 2019). Practices and ideas around academic freedom in Prussia and then a unified German state would influence developments in other European countries, with Beiter et al. (2016) noting that the concept of “freedom of science” would especially gain traction in other European nations. While related to academic freedom, the authors describe freedom of science as distinct from the more encompassing term of academic freedom. The authors point to the United Kingdom and the United States as the places where the contemporary “concept” of academic freedom most originated and developed. In relation to the United Kingdom, they observe that UK law, except in a few instances, is “essentially silent” on academic freedom. According to Beiter et al. (2016), “To a large extent, the protection of academic freedom and institutional autonomy in the UK has been and remains a matter of convention and practice” (p. 117). For the United States, as taken up in the next section, efforts to protect academic freedom developed along two main tracks—(1) reliance on professionalism standards and on tenure as a means to secure economic security for professors and (2) legal protections for academic freedom derived from the First Amendment of the US Constitution.

US universities and academic freedom—tenure and constitutional protections

The establishment of the research university in the United States played a notable role in the development of contemporary concepts of academic freedom (Marginson, 2014). The US experience is thus informative in assessing global trends, both historical and contemporary, related to academic freedom. The overall success of US higher education and espoused allegiance to academic freedom by higher education institutions throughout the United States also provides a cautionary tale. Despite appearing as firmly entrenched and accepted in US higher education, in recent decades, the arrangements arrived at to safeguard academic freedom in the United States are increasingly under stress, even as the system continues to struggle to better mirror the nation’s racial and gender diversity in its faculty ranks and in higher education overall.

As noted, the development of higher education in the later part of the nineteenth century in the United States would be influenced by events in Germany related to freedom, teaching, and research (Sun and Hutchens, 2019). Prior to this period, colleges in the United States were not conceived as institutions to advance knowledge. Instead, colleges were largely viewed as places to preserve established truths and knowledge and to inculcate students with these established truths and to prepare them for service as ministers and as social and political leaders (Thelin, 2019).

As faculty members in the United States began to adopt and adapt practices imported from Germany related to research and academic freedom, they encountered institutional structures and social and political norms that were not conducive to the exercise of faculty independence in teaching and scholarship (Cain, 2018). Faculty members who offended powerful and influential members of society or institutional leadership could and did find themselves dismissed from academic positions. In a response to such incidents, efforts emerged to establish protections for the academic freedom of faculty members at US colleges and

universities. An important culmination in these efforts came with the formation of the American Association of University Professors (AAUP) in the early twentieth century. In 1915, the AAUP issued a Declaration of Principles on Academic Freedom and Tenure. The document specifically referenced the concepts of *Lehrfreiheit* and *Lernfreiheit* and described academic freedom as “compris[ing] three elements: freedom of inquiry and research; freedom of teaching within the university or college; and freedom of extramural utterance and action” (AAUP, 2015, p. 3).

A primary mechanism looked to by the AAUP to safeguard faculty academic freedom came in the call for the establishment of tenure for faculty members after a probationary period, so as to provide ongoing employment security for professors (Sun and Hutchens, 2019). In doing so, the AAUP sought to establish academic freedom as a professional norm in US higher education with economic security serving as the primary means to uphold the academic freedom and independence of faculty members. In the opening decades of the twentieth century, most US colleges and universities would move to accept the AAUP principles, resulting in the 1940 Statement of Principles of Academic Freedom and Tenure, which was endorsed by the AAUP and a prominent national association of colleges (Cain, 2018). It is important to highlight that the protections of tenure and academic freedom were distributed in widely inequitable ways from the start. Many individuals, notably women and men of color, were generally excluded from tenure-stream appointments at US higher education institutions. Today, serious challenges continue to remain in diversifying the US professoriate in terms of those individuals employed in tenure-stream positions.

Throughout the middle decades of the twentieth century, many US colleges and universities would adopt tenure as a standard practice, comprising part of what has been referred to as the “academic revolution” and the ascendance to prominence of US scholars and researchers (Jencks and Riesman, 1968). As a professional norm backed by formal employment arrangements, legal protections for academic freedom were grounded in the contractual relationship between a professor and the employing college or university (Sun and Hutchens, 2019). Under such a contractual relationship, a faculty member and the employing college are engaged in a continuing employment relationship that can only be ended by the institution for acceptable reasons, such as malfeasance, incompetence, or financial emergency. For public colleges and universities, the contractual relationship for tenured professors has come to have a constitutional aspect, as the US Supreme Court has interpreted an ongoing faculty employment agreement as providing a protected property interest that cannot be taken away by a public college or university without an adequate, lawful basis (e.g., *Board of Regents of State Colleges*, 1972).

From promising start to legal morass—academic freedom and the US Constitution

While tenure and academic freedom in the United States were conceived on the basis of professional standards and an ongoing contractual arrangement, a line of legal cases would result in an interpretation of the First Amendment to the US Constitution as an independent source of legal protection for faculty academic freedom (Kaplin et al., 2019; Sun and Hutchens, 2019). The development of a constitutional dimension to academic freedom arose during what is commonly known as the McCarthy era—a period in which fear of communism propelled efforts, often based on paranoia and unfounded accusations, to root out communists and communist influences in US society, including in educational settings. The US Supreme Court would grapple with issues of academic freedom in this environment.

In a 1952 case (*Adler v. Board of Education*), the US Supreme Court, in approving the firing of a teacher, Irving Adler, upheld a New York law that prohibited public employment, including in educational institutions, from individuals belonging to groups determined to be “subversive.” In a dissenting opinion in the case—that is, an opinion opposing the Court’s decision upholding the New York law and the dismissal of Adler—Justice William O. Douglas provided the initial mention of academic freedom in a US Supreme court case. Douglas warned against intrusions into educational settings and the disruption of academic independence. In that same year, Justices Felix Frankfurter and William O. Douglas in a concurring opinion in another case—that is, an opinion in agreement with a majority of the court on outcome but that seeks to further delineate particular points or to offer a substitute grounds for the decision—also looked to principles of academic freedom. The case involved faculty members in Oklahoma who successfully challenged a state-mandated loyalty oath that did not permit differentiation between “innocent” and “knowing” membership in an organization deemed as subversive (*Wieman v. Updegraff*, 1952). Frankfurter and Douglas offered the following sentiments in the concurring opinion:

To regard teachers—in our entire educational system, from the primary grades to the university—as the priests of our democracy is therefore not to indulge in hyperbole. It is the special task of teachers to foster those habits of open-mindedness and critical inquiry which alone make for responsible citizens, who, in turn, make possible an enlightened and effective public opinion They cannot carry out their noble task if the conditions for the practice of a responsible and critical mind are denied to them.

(*Wieman v. Updegraff*, 1952, pp. 196–197)

In 1957, The US Supreme Court decided a case with important academic freedom implications, *Sweezy v. New Hampshire*. Paul Sweezy, a Marxian economist, refused to answer questions during an interrogation by a governmental official, including ones related to lectures that he had given at the University of New Hampshire and his writings. As a result, Sweezy was held in contempt of court. A majority of the US Supreme Court, in overturning Sweezy’s conviction, relied partially on principles of academic freedom but did not squarely base its ruling on academic freedom grounds. In a well-known concurring opinion in the case, Justice Felix

Frankfurter, looking to a statement from a group of South African scholars, offered the “four essential freedoms” of a university: “to determine for itself on academic grounds who may teach, what may be taught, how it shall be taught, and who may be admitted to study” (*Sweezy v. New Hampshire*, 1957, p. 263).

The US Supreme Court’s support for academic freedom would hit a legal highwater mark in 1967 in *Keyishian v. Board of Regents*. Again considering the New York law at issue in *Adler* (1952), this time the court invalidated the provision on academic freedom grounds, with the court stating:

Our Nation is deeply committed to safeguarding academic freedom, which is of transcendent value to all of us and not merely to the teachers concerned. That freedom is therefore a special concern of the First Amendment, which does not tolerate laws that cast a pall of orthodoxy over the classroom. (*Keyishian v. Board of Regents*, 1967, p. 603, p. 603)

In subsequent legal decisions, such as in approving the use of race as a factor in higher education admissions (see, e.g., *Grutter v. Bollinger*, 2003), the US Supreme Court has periodically touted the importance of academic freedom, but its decisions and those from lower courts have also introduced uncertainty regarding the extent to which academic freedom is protected by the First Amendment to the US Constitution. For example, in a case involving the dismissal of a student for poor academic performance, the Supreme Court noted the importance of deference to higher education institutions as an exercise of academic freedom (*University of Michigan v. Ewing*, 1985). But, in a footnote in the case, the Court hinted at ambiguity over whether First Amendment academic freedom rights accrue to individual faculty members or to the institution or to both: “Academic freedom thrives not only on the independent and uninhibited exchange of ideas among teachers and students, ... but also, and somewhat inconsistently, on autonomous decisionmaking by the academy itself” (*University of Michigan v. Ewing*, 1985, p. 222, fn. 7).

Rather than developing clear legal rules grounded in academic freedom principles and matching the rhetoric found in *Keyishian v. Board of Regents* (1967), courts have largely looked to other legal standards in deciding cases that implicate faculty academic freedom (Kaplin et al., 2019; Sun and Hutchens, 2019). For example, courts have often turned to the legal standards dealing with the free speech rights of public employees in general in deciding claims by public university faculty that raise academic freedom concerns. However, a 2006 US Supreme Court decision, *Garcetti v. Ceballos*, threw this arrangement into doubt. In *Garcetti*, the Court ruled that a public employee does not possess First Amendment protection when engaged in carrying out their official employment duties. Since the *Garcetti* decision, lower courts have reached conflicting decisions concerning whether the standard announced in the case applies to faculty members in public higher education when issues of academic freedom are at stake, such as in research or teaching (Kaplin et al., 2019; Sun and Hutchens, 2019). This state of affairs exemplifies a general lack of clarity over academic freedom protections under the First Amendment of the US Constitution. In sum, the status of a federal constitutional dimension to academic freedom protections in the United States is presently enveloped by substantial legal disagreement and uncertainty.

Increasingly Fragile State of tenure in United States

The precarious legal status of First Amendment protections for academic freedom in the United States coincides with another trend with significant implications for the status of academic freedom in the United States—the erosion of tenure as a faculty employment arrangement at both public and private colleges and universities. The percentage of the US professoriate in tenure-stream positions has declined markedly in the past several decades. Now, overall, only about 25% of faculty members in the United States are employed in tenure-stream positions (Tolley, 2018). About another quarter of faculty members are not in tenure-stream positions but are employed full-time with employment contracts of varying lengths. Important variations exist in the employment arrangements for non-tenure-track faculty members in full-time positions. Some are employed on year-to-year work agreements and face annual uncertainty as to the renewal of their position, while some full-time, non-tenure-track faculty members work on multi-year contracts and have meaningful employment protections. In contrast to the economic security afforded by tenure or multi-year contracts, around half of professors in the United States are employed in part-time adjunct teaching positions. Some institutions rely heavily on adjunct professors. These positions often lack employment benefits, such as health insurance, and are precarious in terms of reappointment. In essence, these adjunct professors are excluded from the economic arrangements of tenure originally designed to safeguard the academic freedom of individual faculty members in the United States (see, e.g., Cain, 2018).

Academic freedom remains an espoused cornerstone of US higher education and continues, to varying degrees, an important value for many colleges and universities. Yet, the erosion of tenure and legal doubt over First Amendment protections for academic freedom create considerable uncertainty over the future of academic freedom in the United States. Rather than grounded in contract law or the First Amendment, protection for academic freedom is increasingly subject to the preferences and whims of institutional leaders and higher education governing boards. Recent efforts in some states to block the teaching of critical race theory, including in colleges and universities, highlight a system that is under some level of stress in relation to the future of academic freedom (Flaherty, 2021). Partisan attacks on higher education and on individual professors have only exacerbated that status of academic freedom in US higher education (Mele, 2016). With Covid-19, for instance, efforts by elected officials, including former President Donald Trump, to downplay or discredit the work of scientists, including at colleges and universities, represented a serious threat to the role of academic freedom in the United States (Graham, 2020; Tollefson, 2020). In a similar vein, climate change skeptics in the

United States, including elected officials, have sought to discount the overwhelming body of scientific evidence related to climatic change resulting from human behavior and the severe threats posed to the planet (Jackman, 2014). While academic freedom continues to occupy a central role in US higher education, the legal protections and political and social climates required to uphold academic freedom are under strain. As considered later, such types of threats to academic freedom are far from confined to the United States.

Academic freedom as a human right

In considering academic freedom from an international perspective, as Marginson (2014) states, it is important to recognize that the “research university is not an identical institution everywhere, any more than political cultures and other forms of social organization can be considered uniform” (p. 25). Accordingly, rather than a one-size-fits all approach, considerations of the independence faculty members need to accomplish their roles as teachers, scholars, and members of civil society must be viewed as nested in unique social and political contexts. For example, Marginson states that the exercise of academic freedom and scholarly independence in China are more likely to occur behind the doors of the academy rather than to be asserted in a public manner. He contrasts this approach with the regular conflation in US higher education of academic freedom protections with the general free speech rights that may apply when faculty members are speaking on matters beyond their area of scholarly expertise.

Even while keeping in mind the need to account for distinctive iterations of academic freedom among nations, there are some aspects that appear to be in common. In the push and pull of state authority and institutional autonomy, higher education institutions need the requisite autonomy to create conditions in which scholars possess sufficient independence to carry out their work, as do individual members of the academic community in relation to state authority. And within higher education organizations, institutional authority must not unduly encroach on the professional expertise and independence of members of the academic community, professors and students, to carry forward with their scholarly work. Whether at the state or institutional level, this kind of scholarly or academic freedom encompasses what Marginson (2014) has called “freedom from constraint or coercion” (p. 39).

Quinn and Levine (2014) observe how the concept of academic freedom remains underdeveloped in international law:

Human rights claims arising ... if they are brought, typically do not include claims for violations of academic freedom. (Here “claims” are understood broadly as including but not limited to complaints, petitions and other advocacy before United Nations, regional, sub-regional and national institutions competent to address alleged violations of human rights, including courts, tribunals, committees, commissions, ombudspersons and other regular or special procedures and mechanisms.) As a result the standards are underdeveloped relative to other human rights claims and the scope of academic freedom violations globally. This may be in part due to lack of familiarity with issues of academic freedom and the available standards for protection among human rights advocates (a fact which becomes negatively reinforcing), and in part to the fact that attacks on academic freedom often manifest as violations of other rights under which claims are brought (p. 211).

They assert that international human rights law provides a basis for recognition of academic freedom as an independent right grounded in the rights of freedom of opinion and expression and the right to education (Quinn and Levine, 2014). While acknowledging the difficulty of arriving at a singular definition of academic freedom, the authors, looking to UNESCO’s 1997 declaration on higher education teaching personnel, offer the “core” of academic freedom as

scholars’ right, without constriction by prescribed doctrine, to freedom of teaching and discussion, freedom in carrying out research and disseminating and publishing the results thereof, freedom to express freely their opinion about the institution or system in which they work, freedom from institutional censorship and freedom to participate in professional or representative academic bodies. (Quinn and Levine, 2014, p. 212) (quoting from UNESCO 1997 Recommendation Concerning the Status of Higher Education Teaching Personnel)

Additionally, Quinn and Levine (2014) contend that for academic freedom to flourish in a practical sense, it has to be viewed as connected to related values, including access to higher education, sufficient autonomy of institutions to support the work and activities of individual members of an academic community, physical safety, and social responsibility (p. 213). Rather than depict academic freedom solely as an internal dynamic for higher education, the authors also view research and scholarship as linked to the larger society in which scholars are located. From this perspective, support for academic freedom is nested within and tied to overall “democratic values” and involves the responsibility of “scholars to engage with society outside of the classroom, laboratory or campus, and in return providing some measure of protection for what otherwise might be construed as extramural activities” (Quinn and Levine, 2014, p. 213). Even if an activity is more appropriately labeled as “personal” or “political” in nature as opposed to academic, the authors note that other human rights protections can still apply (Quinn and Levine, 2014, p. 213).

In their examination of academic freedom, Quinn and Levine (2014) consider the mission and work of Scholars at Risk (SAR), a group committed to furthering academic freedom protections worldwide. They describe how SAR views protection for academic freedom as encompassing three, overlapping areas, (1) issues related to “core and socially-engaged understandings of academic

freedom,” (2) that higher education teaching personnel also require the civil, political, social, and cultural protections afforded to all citizens (quoting from the [UNESCO 1997](#) declaration on higher education teaching personnel), and (3) and that violations of core rights impede academic freedom based on individuals directly affected and by “third parties intimidated by attacks on others” ([Quinn and Levine, 2014](#), p. 213).

In terms of a specific grounding for academic freedom rights, [Quinn and Levine \(2014\)](#) contend that academic freedom should constitute an independent human rights claim in proceedings brought before national and international bodies, such as courts or tribunals. They view academic freedom as emanating from the rights of opinion and expression and right to education under Article 19 of the International Covenant on Civil and Political Rights and Article 13 of the International Covenant on Economic, Social and Cultural Rights, as well as similar declarations of rights from other U.N. sources and from “major regional human rights systems” ([Quinn and Levine, 2014](#), p. 216). Additionally, the authors point out that infringements on other rights can be used to undercut academic freedom. The authors consider how arbitrary arrests and deprivations of liberty, restrictions placed on “travel and movement,” dismissal or loss of position as a form of retaliation or on illegitimate grounds, and censorship and restricting access to research findings as ways in which academic freedom rights can be impeded or denied.

Similar to [Quinn and Levine \(2014\)](#), [Beiter et al. \(2016\)](#) argue for the recognition of academic freedom as a human right. They discuss how academic freedom is not specifically mentioned in two of the major international U.N. human rights agreements, the International Covenant on Civil and Political Rights (ICCPR) or the International Covenant on Economic, Social and Cultural Rights (ICESCR). However, they point out that “there seems to exist consensus that the Covenants should be considered to protect that right” ([Beiter et al., 2016](#), p. 110). The authors, like [Quinn and Levine \(2014\)](#), look to multiple international provisions as protective of academic freedom rights, such Articles 19 (right to freedom of expression) and 15 (right to freedom of scientific research) of the ICCPR. The authors point especially to Article 13 of the ICESCR (the right to education) as “the provision which assembles all aspects of academic freedom under “a single roof” and whose normative context provides the proper framework for interpretation” ([Beiter et al., 2016](#), p. 111).

[Beiter et al. \(2016\)](#) offer the following as to their interpretation of Article 13 of the ICESCR as providing the most appropriate home for academic freedom in current international human rights law:

[A]cademic freedom entails much more than free speech rights as covered by the right to freedom of expression—a right, in fact, accruing to each and every person. Academic freedom also covers rights of “free action,” and the free speech rights covered are, in fact, *special* speech rights, conditioned by the dictates of learning, teaching, and research. An approach based on Article 13 also best accommodates the special responsibilities of institutions of higher education and their academic staff in discovering new knowledge and facilitating ethical individualism, adequately emphasizes the special vulnerability of teachers and researchers in higher education to undue pressures, and correctly recognizes that academic freedom must also serve as an organizational principle in structuring the activities of institutions of higher education. The “ordinary” right to freedom of scientific research, as protected in Article 15 of the ICESCR, accrues to each and every person doing research. Although it is highly relevant in this regard too, of course, it lacks a specific reference to the higher education context, where the claim to the protection of free inquiry is the most acute (p. 112).

Alongside international human rights laws and agreements, in their examination of academic freedom, [Quinn and Levine \(2014\)](#) consider how laws and standards in some nations provide academic freedom protections, either through explicit language or indirectly so that academic freedom constitutes an explicit or implied right. Countries may also, they note, incorporate into their laws and constitutions provisions from standards in international human rights treaties. The authors reported in 2012 that some “20 state constitutions include explicit protections, 99 direct protections, and 77 indirect protections” for academic freedom ([Quinn and Levine, 2014](#), p. 233). According to the authors, more recent constitutions are more likely to provide specific language related to academic freedom.

Challenges and opportunities for academic freedom

While keeping in mind the need to avoid a narrow view of academic freedom, there are clear indications that academic freedom is under stress in many countries, a situation that is not new. Writing twenty years ago, [Altbach \(2001\)](#) stated, “In many parts of the world, academic freedom is under attack,” and he further observed that “[a]cademic freedom is by no means secure worldwide” (p. 205). At the same time, he held out “considerable cause for optimism” in pointing to the states that were part of the former Soviet Union and nations in Central and Eastern Europe as having attained “reasonable levels” of academic freedom ([Altbach, 2001](#), p. 25). Improving academic freedom in areas that were part of the Soviet Union is important because the USSR’s oversight of scientists and universities led it to go against the global trend of increasing scholarly publications during the later 20th century ([Powell et al., 2017](#)). While reasons for optimism exist in the second decade of the 21st century, there are many challenges to academic freedom around the world, with academic freedom still “under attack in many places” ([Kinzelbach et al., 2021](#), p. 4). In its 2020 report on academic freedom, *Scholars at Risk*, a New York-based organization that monitors and advocates for academic freedom around the world and was mentioned earlier, reported that from September 1, 2019, to August 31, 2020, it documented 341 incidents worldwide that were attacks on academic freedom, including 124 acts of killing, violence, or disappearance (p. 4). Another 148 individuals faced imprisonment or prosecution.

Events in specific countries or regions can help illustrate the types of challenges to academic freedom that exist around the world. In Hungary, for instance, the government has enacted legislation in recent years that encroaches on the autonomy of higher education institutions and academic freedom (Corbett and Gordon, 2018). In writing about Hungary, Anne Corbett and Claire Gordon (2018) also comment more broadly on emerging challenges to academic freedom in Central and Eastern Europe from governmental action, such as “influencing or determining key university leadership appointments, controlling and directing funding sources, controlling or influencing research agendas, vetting research and publication, and revoking a university’s license” (p. 470). Still, the authors point to some reasons for optimism regarding forces to advance academic freedom in Europe, including Central and Eastern Europe. These include potential support for academic freedom coming from the participants in the Bologna Process and “voluntary but widely accepted European lines of support for academic freedom” (Corbett and Gordon, 2018, p. 472).

To give another example of challenges to academic freedom, in recent years, academic freedom and institutional autonomy appear increasingly under strain in India. Writing of post-colonial Indian higher education, Sundar (2018) stated that three major commissions on higher education in the nation had all pointed to the importance of institutional autonomy and “freedom of thought, expression, and research” (p. 51). Despite these pronouncements, the author warns of growing threats to freedom of teaching and research in India, such as restrictions on curricula (see also, Lau, 2019). Scholars at Risk also highlighted in the aforementioned 2020 report alarming events in India related to academic freedom, such as violent attacks on and retaliation against students in Kashmir.

In Hong Kong, the government has increasingly cracked down on dissent, including that arising from universities (McLaughlin, 2021). In June 2020, China implemented a National Security Law that resulted in the arrest of scholars in Hong Kong for their participation in protests, the canceling of campus events, and the suppression of student speech (O’Malley, 2021). Turkey has shown a consistent willingness to stifle the views of academics (Redlawski, 2021). Russia has also grown increasingly authoritarian in its treatment of faculty members (Schiermeier, 2021). Some African countries have made notable progress in protections for academic freedom, but many nations in the region are “lagging behind” in building and sustaining robust systems for the protection of academic freedom and institutional autonomy (Appiagyei-Atua et al., 2016, p. 110). During the Covid-19 pandemic, scholars from multiple nations have faced retaliation, including arrests, for Covid-19 research and related activities (Scholars at Risk, 2020). The greater use of online teaching modalities during the Covid-19 pandemic also meant that governments had an enhanced ability of oversight and coercion of classroom environments and course content (see, e.g., Bergan et al., 2021). Threats to academic freedom have also been identified based on continuing marketization of higher education (Altbach, 2013). As covered previously, the future foundations of academic freedom in the United States are uncertain given the erosion of tenure as a faculty employment arrangement.

In relation to China, Qiang Zha and Wenqin Shen (2018) consider tensions between China’s efforts to establish world-class research universities and the traditional role of the state in regulating higher education institutions, which reflects the legacy of the “Chinese Confucian knowledge tradition” (p. 448). Within this tradition, the authors note an emphasis on “collective well-being” and serving and upholding the social order and the government. Nonetheless, they point to possible avenues for enhanced academic freedom in China, which could especially prove beneficial to scholars in the humanities and social sciences. Specifically, influences from Neo-Confucianism and Western-based concepts of academic freedom “have enabled more and more Chinese scholars to seek intellectual freedom” (Zha and Shen, 2018, p. 452). Additionally, Zha and Shen (2018) emphasize that

... [A]cademic freedom in China does not have to exactly mirror Western notions. Academic freedom is not absolute, and like freedom itself, it is often conditioned by particular historical and social contexts. In this sense, a kind of Chinese-flavored academic freedom could emerge to allow Chinese institutions of higher education to become the world-class institutions to which they aspire (p. 452).

The introduction in 2020 of the Academic Freedom Index (AFi) added a helpful tool to further understanding of the status of academic freedom globally. The AFi was developed by researchers at FAU Erlangen-Nürnberg, a German research university, the Global Public Policy Institute, a think tank based in Berlin, and Scholars at Risk, an organization covered earlier (Redden, 2020). The index provides a calculation of the current status of academic freedom in specific countries based on the following criteria: “(1) freedom to research and teach; (2) freedom of academic exchange and dissemination; (3) institutional autonomy; (4) campus integrity; and (5) freedom of academic and cultural expression” (Kinzelbach et al., 2021, p. 6). These criteria are described as *de facto* factors and evaluated by “independent experts” for each country using a “pre-defined scale” (p. 6). Additionally, the index is informed by “some additional, factual indicators, assessing states’ *de jure* commitments to academic freedom at (6) constitutional and (7) international levels, as well as (8) whether universities have ever existed in a given country” (p. 6).

The developers of the AFi urge its use by a variety of stakeholders, such as governments, international organizations, university leaders, and advocacy organizations (Kinzelbach et al., 2021). Organizations that provide funding to higher education are encouraged to use the index in efforts to recognize countries where students or staff could be more at risk, as a criterion in funding decisions, and to assess the need for specialized preparation or training based on conditions in specific locations.

One intriguing suggestion is for groups that conduct prominent university rankings to take academic freedom scores into account in compiling such rankings so that “[i]f a certain country scores badly on the AFi, then all universities in that country should be subject to a reduction of their score in a ranking” (Kinzelbach et al., 2021, p. 25). Developers of the AFi state that inclusion of academic freedom in university ranking systems “would create strong incentives for university rectors and governments who value

and act on these prestigious rankings to take steps to strengthen academic freedom. Universities that want to be ranked could also be asked to file a written commitment to academic freedom as a precondition for being assessed" (Kinzelbach et al., 2021, p. 25). For nations and higher education institutions motivated by such rankings, inclusion of an academic freedom criterion could help promote the development of standards and practices supportive of academic freedom.

The most recent report on the AFI highlights the need for innovative strategies and continuing efforts to advance academic freedom around the world. The majority of the 175 countries covered in the second iteration of the AFI do not protect academic freedom at the highest level, with the countries grouped in the index into A (highest score), B, C, D, and E (lowest score) categories using a scaled score from 0 to 1 (Kinzelbach et al., 2021). According to the report, Brazil, Columbia, Hong Kong, Nicaragua, Turkey, and Zambia experienced the sharpest declines in academic freedom over the past five years (Kinzelbach et al., 2021, p. 15). In contrast, scores over the past five years improved the most in Kazakhstan, Maldives, North Macedonia, Sudan, and The Gambia. As covered in the report, the existence of academic freedom provisions in constitutions or laws does not necessarily equate with actual strong protections of academic freedom, with "more than one-third of countries with the worst performances in 2020 (a "D" or "E" status, that is AFI < 0.4) hav[ing] such *de jure* guarantees in place while extensive violations of academic freedom occur in practice" (Kinzelbach et al., 2021, p. 9).

The Academic Freedom Index indicates that academic freedom is strongly to moderately protected in many countries, but it also shows this is not the case in many nations. Realization of academic freedom as a global right represents a work in progress in many places. And, as shown in the AFI report and from the country-specific examples included in this entry, protections for academic freedom are subject to erosion or to new threats.

Conclusion

While the contours of academic freedom are flexible and adaptable based on national context and history, wide agreement exists that academic freedom is crucial in higher education. Academic freedom is not intended to accrue to individual scholars for their personal benefit, but, instead, is meant to allow them to perform their roles as teachers, researchers, and members of civil society. Academic freedom also envelops students and the larger society and contributes to the overall vitality of a nation. An integral facet to the academic freedom of scholars is the independence to carry out their work without undue governmental encroachment or limitations. As shown in the Academic Freedom Index covered in this entry, there are numerous nations in which the current status of academic freedom appears strong. However, there are many countries, as indicated by the index, where academic freedom protections are marginal to seemingly non-existent. As considered in this entry, even for a nation like the United States—which receives a strong score on the Academic Freedom Index—arrangements designed to safeguard academic freedom, notably tenure, are under increasing stress. The uneven distribution of academic freedom protections and increasing threats to academic freedom in some nations and regions reveal ongoing challenges to academic freedom from an international perspective. Crucial challenges exist for researchers, higher education institutions, policymakers, governments, and advocacy groups to protect and advance academic freedom protections around the world.

References

- Adler v. Board of Education, 342 U.S. 485, 1952.
- Altbach, P.G., 2001. Academic freedom: international realities and challenges. *High Educ.* 205–219. <https://doi.org/10.1023/A:1026791518365>. Available at: (Accessed 9 July 2021).
- Altbach, P.G., 2013. The International Imperative in Higher Education. In: *International Perspectives on Education and Society Series*, vol. 27. SensePublishers, Rotterdam.
- American Association of University Professors, 2015. *Policy Documents and Reports*, eleventh ed. Johns Hopkins University press, Baltimore, Md.
- Appiagyei-Atua, K., Beiter, K., Karran, T., 2016. A review of academic freedom in Africa through the prism of UNESCO's 1997 recommendation. *J. High Educ. Afr.* 14 (1), 85–117. Available at: <https://www.jstor.org/stable/10.2307/90016101>. (Accessed 10 July 2021).
- Baker, D.P., 2014. *The Schooled Society: The Educational Transformation of Global Culture*. Stanford University Press, Stanford, California.
- Beiter, K.D., Karran, T., Appiagyei-Atua, K., 2016. "Measuring" the erosion of academic freedom as an international human right: a report on the legal protection of academic freedom in Europe. *Vanderbilt J. Transnatl. Law* 49 (3), 597–691.
- Bergan, S., Gallagher, T., Harkavy, I., Munck, R., van't Land, H. (Eds.), 2021. *Higher Education's Response to the Covid-19 Pandemic: Building a More Sustainable and Democratic Future* (Council of Europe Higher Education Series No. 25). Council of Europe Publishing.
- Board of Regents of State Colleges v. Roth, 408 U.S. 564, 1972.
- Cain, T.R., 2018. Friendly public sentiment' and the threats to academic freedom. *Hist. Educ. Q.* 58 (3), 428–435. <https://doi.org/10.1017/heq.2018.19>. Available at: doi. (Accessed 9 July 2021).
- Corbett, A., Gordon, C., 2018. Academic freedom in Europe: the central European university affair and the wider lessons. *Hist. Educ. Q.* 58 (3), 467–474. <https://doi.org/10.1017/heq.2018.25>. Available at: (Accessed 9 July 2021).
- Flaherty, C., 2021. Legislating against critical race theory, Inside Higher Ed. Available at: <https://www.insidehighered.com/news/2021/06/09/legislating-against-critical-race-theory-curricular-implications-some-states>. (Accessed 26 October 2021).
- Garcetti v. Ceballos, 547 U.S. 410, 2006.
- Graham, D.A., 2020. What a direct attack on free speech looks like. *Atlantic*. Available at: <https://www.theatlantic.com/ideas/archive/2020/07/trump-universities/614038/>. (Accessed 9 July 2021).
- Grutter v. Bollinger, 539 U.S. 306, 2003.
- Jackman, T., 2014. Va. Supreme Court rules for U-Va. in global warming FOIA case. *Wash. Post*. Available at: <https://www.washingtonpost.com/blogs/local/wp/2014/04/17/va-supreme-court-rules-for-u-va-in-global-warming-foia-case/>. (Accessed 9 July 2021).

- Jencks, C., Riesman, D., 1968. *The Academic Revolution*. Doubleday, Garden City, N.Y.
- Kaplin, W.A., Lee, B.A., Hutchens, N.H., Rooksby, J., 2019. *The Law of Higher Education*, sixth ed. Jossey-Bass, San Francisco, Calif.
- Kaye, D., 2020. Report of the [U.N.] Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression [special report on academic freedom protections presented to U.N. General Assembly]. Available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/197/86/PDF/N2019786.pdf?OpenElement>. (Accessed 9 July 2021).
- Keyishian v. Board of Regents, 385 U.S. 589, 1967.
- Kinzelbach, K., Saliba, I., Spannagel, J., Quinn, R., 2021. Free Universities: Putting the Academic Freedom Index into Action. Global Public Policy Institute and Scholars at Risk Network. Available at: <https://www.gppi.net/2021/03/11/free-universities>. (Accessed 9 July 2021).
- Kreckel, R., 2017. On academic freedom and elite education in historical perspective: medieval Christian universities, Islamic madrasas, Ottoman palace schools, French grandes écoles, and "modern world class research universities.". Paper available through Social Science Research Network (SSRN). Available at: <https://ssrn.com/abstract=3081429>. (Accessed 9 July 2021).
- Lau, J., 2019. India Warned on Academic Freedom: lack of respect for international norms will detract from efforts to build excellence. Inside Higher Ed. Available at: <https://www.insidehighered.com/news/2019/10/10/india-warned-lack-academic-freedom-will-hurt-drive-research-excellence>. (Accessed 9 July 2021).
- Marginson, S., 2014. Academic freedom: a global comparative approach. *Front. Educ. China* 9 (1), 24–41. Available at: <https://link.springer.com/article/10.1007/BF03397000>. (Accessed 9 July 2021).
- Mele, C., 2016. Professor watchlist is seen as threat to academic freedom. *N. Y. Times*. Available at: <https://www.nytimes.com/2016/11/28/us/professor-watchlist-is-seen-as-threat-to-academic-freedom.html>. (Accessed 9 July 2021).
- McLaughlin, T., 2021. How academic freedom ends. *Atlantic*. Available at: <https://www.theatlantic.com/international/archive/2021/06/china-hong-kong-freedom/619088/>. (Accessed 9 July 2021).
- O'Malley, B., 2021. New threats to academic freedom emerge from pandemic. *University World News*. Available at: <https://www.universityworldnews.com/post.php?story=20210313080914479>. (Accessed 26 October 2021).
- Powell, J.J.W., Baker, D.P., Fernandez, F. (Eds.), 2017. *The Century of Science: The Global Triumph of the Research University*. International Perspectives on Education and Society Series, vol. 33. Emerald Publishing.
- Quinn, R., Levine, J., 2014. Intellectual-human rights defenders and claims for academic freedom under human rights law. *Int. Hum. Right. Law Rev.* 3 (2), 209–247.
- Redden, E., 2020. Ranking academic freedom globally. Inside Higher Ed. Available at: <https://www.insidehighered.com/news/2020/03/30/new-index-rates-countries-degree-freedom-scholars>. (Accessed 9 July 2021).
- Redlawski, D.P., 2021. Academic freedom under attack in Turkey: 2019 presidential address, international society of political psychology. *Polit. Psychol.* <https://doi.org/10.1111/pops.12724>. Available at: (Accessed 9 July 2021).
- Schiermeier, Q., 2021. Russian academics decry law change that threatens scientific outreach. *Nature*. Available at: <https://www.nature.com/articles/d41586-021-00385-5>. (Accessed 9 July 2021).
- Scholars at Risk, 2020. Free to think: report of the Scholars at Risk academic freedom monitoring report. Available at: <https://www.scholarsatrisk.org/wp-content/uploads/2020/11/Scholars-at-Risk-Free-to-Think-2020.pdf>. (Accessed 9 July 2021).
- Sun, J.C., Hutchens, N.H., 2019. Faculty expressions and legal implications. In: Bon, S.C., Nguyen, D.H.K., Rippner, J.A. (Eds.), *Contemporary Issues in Higher Education Law*, fourth ed. Education Law Association, Cleveland, Ohio, pp. 108–128.
- Sundar, N., 2018. Academic freedom and Indian universities. *Econ. Polit. Wkly.* 53 (24), 48–57. Available at: <https://www.epw.in/journal/2018/24/special-articles/academic-freedom-and-indian-universities.html>. (Accessed 9 July 2021).
- Sweezy v. New Hampshire, 354 U.S. 234, 1957.
- Thelin, J.R., 2019. *A History of American Higher Education*, third ed. Johns Hopkins University Press, Baltimore, Md.
- Tolley, K., 2018. *Professors in the Gig Economy: Unionizing Adjunct Faculty in America*. Johns Hopkins University Press, Baltimore.
- Tollefson, J., 2020. How Trump damaged science — and why it could take decades to recover. *Nature*. Available at: <https://www.nature.com/articles/d41586-020-02800-9>. (Accessed 9 July 2021).
- University of Michigan v. Ewing, 474 U.S. 214, 1985.
- United Nations Educational, Scientific and Cultural Organization, 1997. Recommendation concerning the status of Higher Education Teaching Personnel. Available at: http://portal.unesco.org/en/ev.php-URL_ID=13144&URL_DO=DO_TOPIC&URL_SECTION=201.html. (Accessed 9 July 2021).
- Wieman v. Updegraff, 344 U.S. 183, 1952.
- Zha, Q., Shen, W., 2018. The Paradox of Academic Freedom in the Chinese context. *Hist. Educ. Q.* 58 (3), 447–452. <https://doi.org/10.1017/heq.2018.22>. Available at: (Accessed 9 July 2021).

The funding of higher education: international perspectives

Ben Jongbloed, Center for Higher Education Policy Studies (CHEPS), Faculty of Behavioural, Management and Social Sciences (BMS), University of Twente, Enschede, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Funding higher education: the key issues	183
How much is spent on higher education? By whom?	184
The funding environment for higher education institutions	185
How are public funds made available to higher education institutions?	186
Performance-based funding	188
Final observations	189
References	189

Funding higher education: the key issues

With higher education tuition fees and student debt rising and government budgets allocated to universities and colleges showing little signs of decreasing, the financing of higher education is an often debated issue. Questions are raised in political debates, policy reviews and academic discussions about issues such as escalating cost burdens (Clotfelter, 1996), improving efficiency (Massy, 2004), the rates of return on investment in higher education (Blöndal et al., 2002; Boarini and Straus, 2007; Psacharopoulos, 1981) and finding adequate ways of resourcing higher (or tertiary) education (OECD, 2020). In this contribution we will present an overview of some of the literature on the critical issues around higher education funding. This overview will be presented around a few key questions that are introduced below. In addition, we will also show some of the differences that exist between countries in how they have addressed the key funding-related issues.

Funding is never an end in itself. It is always a means to an end. For governments and students alike, the resources that are dedicated to higher education are regarded as an investment—a means to encourage the building up of knowledge and skills through (higher) education. Through supporting their research universities, governments promote research—basic, applied, strategic research; they encourage the pursuit of knowledge for its own sake as well as for supporting the technological and social innovations in today's knowledge-driven societies (Salmi, 2003). Thus, countries, states and regional governments invest in higher education to promote economic growth (Nelson and Phelps, 1966) and innovation (Acemoglu, 1997; Redding, 1996). Higher education, therefore, is a tool for creating individual and national competitive advantage—a means to contribute to the well-being of nations (Healy and Côté, 2001).

In deciding on the funds invested in higher education, the governments and citizens of countries need to make important choices. They face trade-offs in using scarce resources to achieve multiple goals and will need to keep an eye on the instrumental objectives of efficiency (getting the most we can from the scarce resources) and equity (distributing the benefits in a fair way among the members of society). In many ways, these are choices of a political-economic nature (Garritzmann, 2016). The choice of goals to pursue is very much an outcome of a political process that can be informed by academic research and expert advice on the benefits and costs of particular policy options. Some of these policy evaluations—the ones directly related to funding reforms—will be mentioned in the remainder of this text.

In this entry to the encyclopedia we will discuss higher education funding through the following key questions that are relevant for all higher education systems:

1. How much is spent on higher education (or spent per student, per graduate, or per measure of research activity)?
2. What is the share of private spending (by students; households; companies) on higher education as compared to the public (i.e., government) spending?
3. How are public funds made available to higher education institutions?

Answers to these questions are very much driven by policymakers' choices and what they regard as constituting an efficient and fair system of financing higher education. In the following sections, we will touch upon these three key questions, starting with the first two in the next section.

Before addressing the third question, we will present a broad overview of the different revenues of higher education institutions—showing the variety of their revenue sources. We then continue with a typology of the mechanisms that determine the core funding that higher education institutions receive from the government.

In the final sections of this article we touch upon the question of whether the type of funding matters for the functioning of the higher education system—its diversity, and some of the dimensions related to the performance of the system. To encourage that performance, many governments over the years have introduced reforms in the way they fund their higher education institutions. The emergence of performance-based funding and competitive funding are examples that will be discussed here. Both examples are

manifestations of an increased reliance on market-type steering approaches in higher education—a trend also known as new public management.

How much is spent on higher education? By whom?

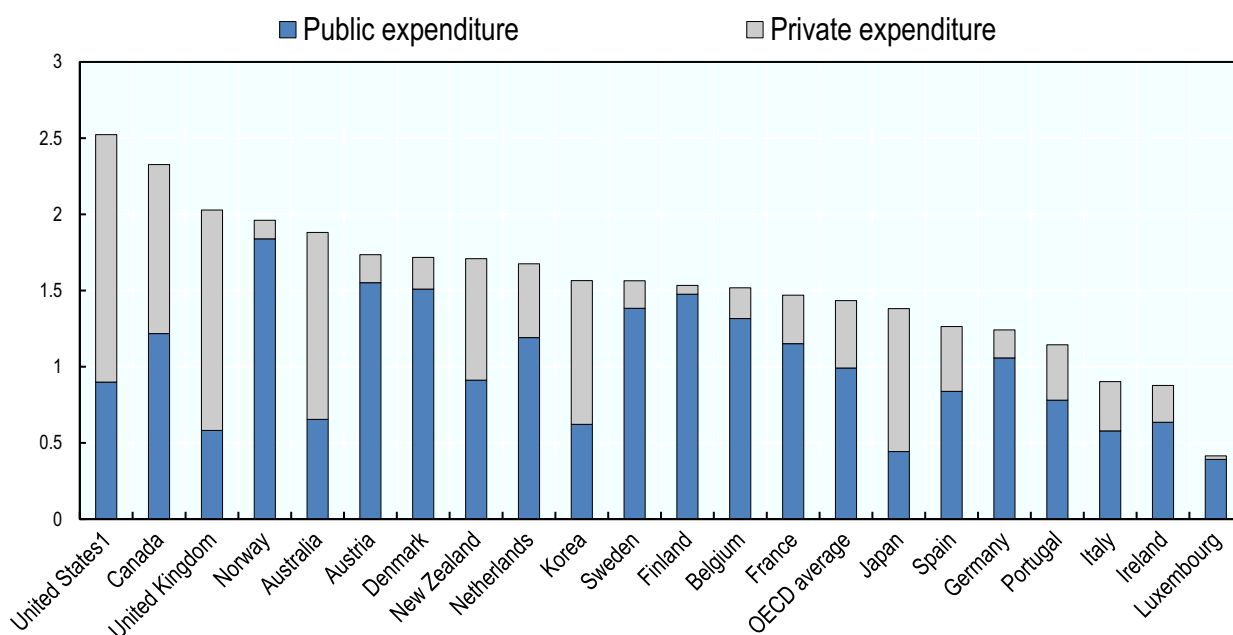
When it comes to the volume of resources spent on higher education, the broadest measure is total public and private spending on higher education. Here, we see large differences not just between the developed and less-well developed countries in the world, but also between countries within these two broad groups. The wealth of countries is one reason for these differences, but so are the priorities attached to higher education. Even among OECD countries—the richest countries in the world—we see quite substantial differences. **Fig. 1**, based on the OECD's Education at a Glance database (OECD, 2021) ranks 22 OECD countries on the basis of their higher education expenditures.

Public spending includes not just the higher education institutions' operating grants for the purpose of education, research and other activities (capital expenditures; community services), but also the payments that governments make to students as part of the student support system. In 2018, the OECD countries shown in **Fig. 1** on average spend about 1.4% of GDP on higher education from public and private sources - two-thirds of it on average coming from public and one-third from private sources (OECD, 2021).

Fig. 1 also reveals wide differences in the share of public spending versus that of private spending. The Nordic countries (Norway, Denmark, Sweden, Finland) stand out with generous public spending, whereas Anglo-Saxon countries such as the UK and the US spend much less from the public purse. Western European countries invest quite considerably in higher education, while Eastern Europe and some Southern European countries score in the middle of the distribution. **Fig. 1** also shows that private expenditure on higher education is particularly high in the Anglo-Saxon countries, especially in the US. The same is true for both Korea and Japan. In these countries, the level of tuition fees is relatively high.

The increasing demand for higher education, the rising per-student costs, and the budgetary restrictions facing many governments have triggered the introduction of tuition fees. Cost-sharing is the term often used for this trend (Johnstone and Marcucci, 2010). Cost sharing reflects the shift in the higher education cost burden from governments (or taxpayers) to parents and students. It encompasses the introduction of (or sharp increases in) student tuition fees, student user charges, and substituting student grants for student loans.

Several countries have introduced cost-sharing policies (Teixeira et al., 2006). At the same time, and to ensure that (potential) students are not deterred by the financial burden erected by tuition fees, governments have started to provide financial support to students. Providing income-contingent loans (ICLs) to students who cannot or do not want to pay the tuition fees and wish to defer



Note: 1. Figures are for net student loans rather than gross, thereby underestimating public transfers.

Fig. 1 Total expenditure on higher education institutions as a % of GDP, by source of funds for the year 2018. Note: 1. Figures are for net student loans rather than gross, thereby underestimating public transfers. Based on OECD (2021), Figure C2.2 (<https://stat.link/n2rbd1>).

their payments is an innovative method of supporting students in meeting the costs of higher education. ICLs were introduced in Australia (Chapman, 2006) and have become more widespread over time.

Clearly the discussion on the balance between public and private contributions and how to pay for higher education is a politically sensitive issue, with countries showing a variety of approaches and models (Cantwell et al., 2018; Garritzmann, 2016), depending on the value they place on the public benefits (i.e., externalities) of higher education versus the private (say, individual) benefits. Clearly, this issue is linked to questions around efficiency and equity.

Following the OECD, we can divide countries into four groups, depending on tuition fee levels and student support availability (e.g., OECD, 2008, p. 275). Table 1 places countries in these four groups, using a share of 50% of the students benefiting from student support and a tuition fee level of (roughly) USD 4000 as the cut-off points for making distinctions between groups. The Nordic countries have a student support system that is relatively generous and their students pay no fees. All of them are social-democratic regimes that have a progressive tax structure and comparatively high income tax rates. In many countries in continental Europe, students are charged low or moderate fees (with the exception of the Netherlands). In the more liberal welfare states (in the bottom of the table), governments stress the private benefits of investing in higher education and expect their students to pay a higher share of the cost burden; expecting the majority to use student loans to cover education costs.

The funding environment for higher education institutions

In light of the evidence that investments in (higher) education (say, human capital) are beneficial to economic and social development (Healy and Côté, 2001), the public budgets for higher education have grown considerably over time (Johnstone, 2004). And with higher education being such a large part of the public sector (see indicator C4 in OECD, 2021), there is increasing scrutiny on how public resources for higher education are allocated and used—for education, research, student support, infrastructure, staffing, campus development, etcetera. At the national (i.e., federal, country, state) level, reforms in educational financing are frequently debated in policy circles to identify the funding mechanisms that produce the best outcomes in terms of guaranteeing access for students, high-quality education, and high-quality research, as well as connecting this education and research to the needs of society.

Before looking at those funding reforms, let us first discuss the different channels through which higher education institutions receive their resources. Fig. 2 provides an overview of the funding environment of higher education institutions. It shows the different funding streams and their sources, and illustrates the different levels of decision-making dealing with the allocation of financial resources. It also displays how funding is the outcome of a complex and nested set of processes that involves multiple authorities, agencies and entities. Each of these actors will have its own goals, preferences, rules and instruments and their decisions on resource allocation and revenue generation are embedded in a broader political and budgetary context.

Core funding is in the center of the picture. It represents the basic (i.e., annual, recurrent) funding allocated by the government from its higher education and research budget for the general functioning of the higher education institutions. This core budget is then redistributed by the higher education institution to its various departments and organizational units.

Student support is provided to students in various ways. It can consist of grants, loans, vouchers and support for the students' parents. Part of this will be used by students to pay for the tuition fees charged by higher education institutions. The type and the level of the fees will depend on a country's regulatory system and the legal status of the institution (i.e., public, private, for-profit, not-for-profit).

A third funding stream originating from the public sector consists of research grants awarded by research councils operating on behalf of (but largely independent of) the public authorities. This funding stream directly flows to the researchers and research groups in the university. In many cases, these grants are awarded selectively, with researchers submitting project proposals to research councils that, after a peer review process, decide which research projects receive funding. The funding is targeted to a particular project and awarded for a specific period of time. Allocations aimed at creating centers of excellence constitute another example of the indirect, non-core funding of higher education institutions.

Again, countries will differ in terms of the shares of competitive funding in their system and the degree of competition experienced by researchers. Countries also show structural differences in the number and functions of the public research organizations that exist next to the research universities. The size of this sector also will determine the research budget of the university sector.

Table 1 Student support and tuition fees in world regions (situation: 2017/18).

	<i>Less than 50% of students benefit from public loans AND/OR scholarships/grants</i>	<i>More than 50% of students benefit from public loans AND/OR scholarships/grants</i>
Below average (or zero-fees) charged by public HEIs	Most of continental Europe	Nordic countries
Above average tuition fees charged by public HEIs	East Asia	Anglo-Saxon countries

Author, based on OECD's Education at a Glance OECD (2021).

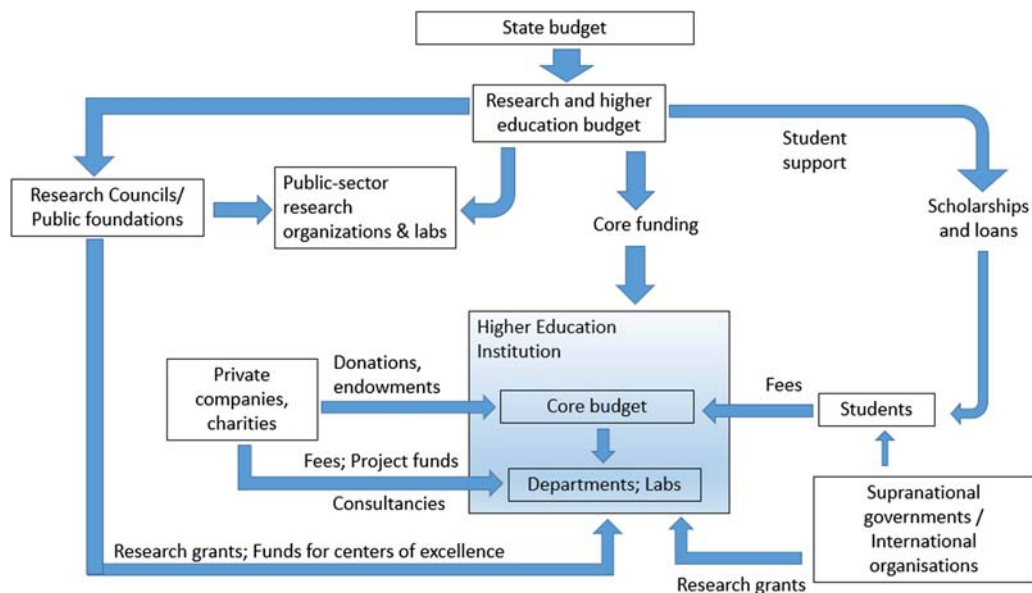


Fig. 2 The funding environment of higher education institutions. Extended from Lepori and Jongbloed (2018).

Shown on the bottom right hand side of Fig. 2, supranational governments, such as the European Commission, also award project funding and research grants to researchers. The European Commission also provides financial support to students and researchers (e.g., Ph.D. candidates) to encourage international mobility.

Research funds from private sector organizations (both for-profits and non-profit organizations) and their payments for education and research services constitute another part of the non-core funds of higher education institutions. In some countries, these private organizations provide gifts and donations to higher education institutions—for endowed professorships, for infrastructural facilities (e.g., libraries, labs) or to support the general funds of the institution.

In most countries, the core funds represent the largest share of the higher education institution's budget. These institutional funds provide the institution with some stability in the support of their education and research activities. In most cases the core funds are allocated as a lump sum payment—one block grant for the carrying out of education and research. This allows the institution to decide on the use of the funds and set its own priorities. As part of this, higher education institutions act strategically (Bonaccorsi and Daraio, 2007), building up their strengths and capabilities, and competing for project funds and other third party funding. In this context, higher education institutions are diversifying their funding sources (De Dominicis et al., 2011) acting entrepreneurial in an almost businesses-like way, competing on markets for resources, students and academic staff (Deiaco et al., 2010; Teixeira et al., 2004). Some authors (e.g., Slaughter and Rhoades, 2004) have used the term academic capitalism to describe this development. Academic capitalism describes the phenomenon that universities are turned into knowledge producing enterprises, competing for capital accumulation and trying to turn their findings into patents and profitable commodities.

How are public funds made available to higher education institutions?

Now that we have looked at the funding environment for higher education institutions, let us consider the different approaches—allocation mechanisms—that national (or state) governments use to determine the volume of core funds allocated to their higher education institutions. In many countries—certainly in Europe—this core funding represents the largest share of the revenues of higher education institutions. In times of multiple claims made on public budgets, there is increasing scrutiny on how public resources for higher education are allocated to higher education institutions and how these resources are used.

In Fig. 3 we classify the different mechanisms that can be employed. The figure distinguishes formula-based and negotiations-based approaches, as well as approaches where the previous years' allocations drive the core funding (and where allocations are often adjusted in an incremental way). These approaches intersect with funding methods that focus on performance, respectively inputs.

In formula funding, the core funds are calculated through one or more mathematical formulas, based on a set of predefined parameters and indicators. The formula will typically include volume measures (e.g., student enrollments per degree program) and price components (i.e., unit costs—or prices), where the latter are set with a view to the cost of delivering a particular unit in terms of education or research.

The main aim of a formula is to treat all higher education equally—according to the same formula, thus providing transparency in the funding decisions. The variables and parameters in the formula—e.g., the funding basis and the funding weights—will be

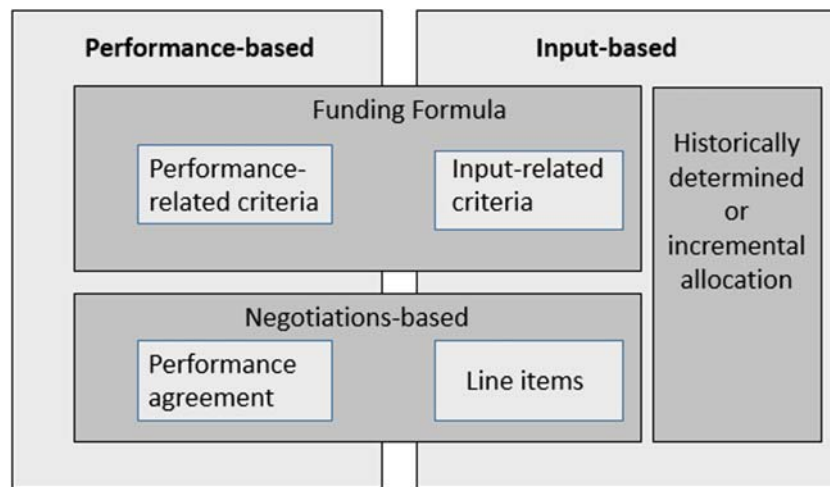


Fig. 3 Core funding mechanisms. Author inspired by [Claeys-Kulik and Estermann \(2015\)](#).

specified by the funding authorities in such a way that they align closely to the broader higher education policy priorities of the government, allowing for changes to be made to the parameters over time ([OECD, 2020](#)).

Funding formulas often include a mix of input and output (i.e., performance) indicators, along with their weights to stress higher resource needs or higher priorities attached to particular activities in education or research. For instance, a higher weight might be applied to indicators related to the number of students from a particular group (e.g., students with disabilities, or from a disadvantaged background). In this way, funding authorities have a way to express policy priorities.

In the United States, funding formulas have been in use since the 1960s ([Lasher and Greene, 1993](#); [Lasher and Sullivan, 2004](#)). In Europe, its introduction is of a more recent date ([Kaiser et al., 2001](#)). Initially, funding formulas predominantly included education-related indicators and parameters. Simple input indicators were used, such as the number of students and the cost of instruction weighted by disciplinary field. Another example of an input measure is the staff volume of an institution or its floor space.

With the introduction of New Public Management approaches in the public sector in the 1990s ([Ferlie et al., 1996](#)), many countries have reformed the ways in which their higher education institutions are funded and have started to introduce measures of the institutions' performance in the funding formula ([Jongbloed and Vossensteyn, 2016](#)), thus creating a quasi-market environment ([Herbst, 2007](#)). Through performance-based funding ([Burke, 2002](#)), institutions are rewarded for delivering particular outputs; they are incentivized to pay more attention to degree completion, their students' time-to-degree, the amount of credits accumulated by students, research productivity, citation rates, et cetera.

In negotiated funding systems, shown in the bottom-half of [Fig. 3](#), the amount of core funding allocated to an institution is an agreed budget negotiated between funding authorities and the individual institution. Traditionally, these negotiations were about staff (e.g., the number of professors), infrastructure needs (e.g., lecture halls, laboratories) and students to be admitted to the institution. All of these are input-related criteria. In the traditional budget approaches—shown on the right-hand side of [Fig. 3](#)—those negotiations dealt with a large number of separate line items, such as salaries, materials or infrastructure costs. Line items then were adjusted using parameters such as the rate of inflation, and next year's core budget would be heavily based on the previous year's line items.

Over time, such historical and primarily input-based approaches have been replaced by more transparent, indicator-based systems and funding formulas. In addition, and once again inspired by the market-based approaches aligned with new public management, funding authorities have started to introduce performance criteria in the negotiations-based approaches. In some countries this resulted in a system of core funding that is—at least partly—based on “performance agreements” ([de Boer et al., 2015](#)). In the case of a performance agreement, the funding authority and each individual higher education provider agree on the goals for the institution to achieve and the plans it aims to carry out for the years ahead. A part of the institution's core funds is tied to the realization of those goals. The goals can be described in qualitative terms (e.g., improve equal access of men and women to senior academic positions, improve access for disadvantaged students), but they can also be specified using quantitative indicators (e.g., targets in terms of an increase of completion rates, the number of female professors, or the intake of Ph.D. students). The ways and details in which performance agreements are laid out will depend on the rationale for the agreements and the goals that the national authorities have for their higher education system.

The agreement also specifies the “rules of the game,” in terms of the evidence that needs to be delivered by the institution to the funding authorities and the positive or negative financial consequences connected to meeting the goals. Performance agreements thus are expected to encourage higher education institutions to improve their performance and to act strategically—in line with their ambitions, strengths and weaknesses.

Performance-based funding

The way in which funds are allocated matters, because it affects the behavior of those (i.e., the higher education institutions, students, researchers) receiving them. Performance-based approaches are increasingly employed in Europe (Claeys-Kulik and Estermann, 2015) and North America (e.g., Dougherty et al., 2016; Jongbloed and Vossensteyn, 2016) as a leverage to achieve national goals. In many countries, improving completion rates and limiting drop-out are key objectives underlying funding reforms and the introduction of performance-based funding.

Jongbloed et al. (2022) describe that sixteen of the 27 European Union member states nowadays include an indicator related to completion in their funding formulas or performance agreements, and several European countries include measures of research performance in their formulas for awarding core funding to universities. A classic example of the latter is the formula used in the United Kingdom for determining research funding allocations on the basis of peer review scores awarded in research assessments (Barker, 2007).

Performance agreements have been implemented—under several names (e.g., performance contracts, development contracts, strategic mandate agreements) and in various forms—in European countries (e.g., Austria, Croatia, Denmark, Finland, Germany, Ireland, the Netherlands, Scotland), Asia (Hong Kong, Japan), Australia, Canada, and the USA (de Boer et al., 2015; Jongbloed and Vossensteyn, 2016; Jongbloed et al., 2022).

Performance agreements allow institutions to select and negotiate their goals with an eye upon their individual context, strengths and key stakeholders. Therefore, performance agreements allow for a more tailored approach to higher education funding compared to the more one-size-fits-all formula-based approaches that make use of a uniform formula and a small set of indicators applied to all institutions in the higher education sector. Because of this, the performance agreements can help to increase mission diversity in the higher education system (Van Vught, 2008), next to their main goal of increasing performance.

Policymakers often see performance-based funding as a useful tool in connecting funding to measurable indicators. However, there is a lack of conclusive evidence that performance funding results in significant improvements in student outcomes (Hillman et al., 2015; Ziskin et al., 2018). While there is substantial research on whether performance-based research funding systems affect research performance (e.g., Hicks, 2012), much less research is available on the impact of performance-based education funding. The available research mainly focuses on the experience of the USA (Hillman and Corral, 2017; Kelchen, 2018), although some work has been done on the Nordic countries (Aagaard, 2015; Kivisto et al., 2017; Mouritzen and Opstrup, 2020) and the Netherlands (Jongbloed et al., 2019). For the United States and Denmark there are indications that performance-based education funding raised the instructional efforts of the institutions, encouraging them to make improvements in their programs and services (Dougherty et al., 2016; OECD, 2020, p. 90).

Evidence showing improvements in completion and graduation rates is, however, limited. Existing empirical studies are restricted to specific (country) cases and limited aspects of performance, making it difficult to generalize their findings. In a recent study on European countries, it was concluded that performance-based funding does seem to contribute to a higher degree of goal orientation and helps increase transparency and accountability around public spending on higher education (Jongbloed et al., 2022). In the USA and Europe, performance-based funding has been linked with higher research productivity (Dougherty and Natow, 2019), e.g., in terms of publication patterns. Yet, a recent study (Ward and Ost, 2021) on its effects in two US states (i.e., Louisiana and Tennessee) found no evidence that these programs improve key academic outcomes.

Performance-based funding also has led to debates about unintended negative consequences. In particular in case mitigation measures are not adequately accounted for in the design of the funding system. An often-heard concern is that institutions may be tempted to lower the quality of student learning, placing quantity above quality (de Boer et al., 2015). On the same note, institutions may start reducing the intake of students from disadvantaged socio-economic backgrounds in the belief that these students have a lower chance of completing their program successfully (Dougherty et al., 2016). This may have consequences for student access; making higher education institutions place a lower priority on their equity objectives.

In Slovakia, performance-based funding has been linked to researchers publishing in lower-quality “predatory” journals and increasing their citations of close colleagues (Pisár et al., 2019). In Poland, the use of research outputs in the funding formula has led to insufficiently productive staff being reassigned to a teaching path in economics universities (Jelonek and Mazur, 2020).

Furthermore the use of performance elements in funding formulas potentially increases the homogeneity in the higher education system, as it may lead to risk-avoiding behavior among institutions and faculty. A uniform funding formula may induce institutions to focus their education and research activities on the more easily achievable outputs (OECD, 2020). In the USA, performance-based funding was associated with more short-term certifications being awarded, a higher selectivity in student recruitment, and a lower academic quality (Li, 2018). In Italy, performance-based funding is linked to a greater stratification in the higher education landscape (Mateos-González and Boliver, 2019). These negative effects have sometimes been countered by changes in the funding system, in the form of weight adjustments, a redefinition of indicators, and/or the use of regulatory instruments (Bennetot-Pruvot et al., 2015).

The question whether funding systems—be they input- or performance-based—ultimately have an effect on the behavior of higher education institutions is difficult to answer. Funding systems are embedded in the broader (national and international) governance context for higher education. This makes it difficult to isolate the effect of performance-based funding from the effects of other institutional arrangements and incentive structures. Furthermore, performance-based systems come in various shapes and sizes. For instance, the share of core funding tied to performance differs substantially across countries (e.g., Jongbloed et al., 2022), and so does the share of competitive (e.g., project-based) funding in the “funding mix” (Lepori and Jongbloed, 2018). Finally, there

are various dimensions of performance—as part of the institutions’ education, research and societal missions. In other words, every case is in some way unique and general conclusions are difficult to make. The available overview studies and meta reviews that synthesize the existing research on the impact of performance-based funding (e.g., Hillman et al., 2015; Kivistö and Kohtamäki, 2016; Ortagus et al., 2020) often have difficulty drawing firm conclusions on the matter, or at best find no effects at all.

Final observations

This contribution has shown that the topic of funding of higher education leads to many questions. We looked at questions of “how much, who pays, what types and what conditions of funding?” It became clear that the answers to the questions differ across the higher education systems in the world, depending on a country’s wealth, its political priorities and its approach to the governance of government-dependent institutions. Funding arrangements are adjusted regularly—to address the needs of society and to incorporate lessons from experience—both at home and abroad. Governments and funding authorities will constantly be looking for the most cost-efficient, yet least intrusive type of funding system to steer their universities and colleges in delivering on their missions.

The multiple objectives and dilemmas surrounding resource allocation in higher education touch upon issues of the balance between (1) public and private investments in higher education; (2) equity and efficiency; (3) input-based and performance-based approaches for the funding of higher education, (4) the room for competition between educational institutions as opposed to their demands for stability and collaboration (5) core funding and project funding, and (6) formula-based and negotiations-based methods. In addition, there is the question on the implications of all this on the efficiency, quality and equity in the higher education system. Indeed, there is a large literature on how countries go about in addressing these trade-offs and dilemmas and how all that is reflected in their resourcing of higher education (OECD, 2020).

We still know relatively little about the impact of the different types of funding arrangements on the performance of the higher education system—their students, lecturers and researchers. Firstly, because performance (or impact) itself is a multi-dimensional concept (OECD, 2019) and, secondly, because of the wide variety in resourcing approaches that exist and the different contexts in which they are embedded. Finding the best funding system for higher education and getting the incentives right (Canton et al., 2001) is a near impossible endeavor, but learning about the parts of the system that work and the parts that do not is possible, thanks to the accumulation of knowledge through policy evaluations and experiments. And, as so often (for example, in higher education), the process of learning may be more important than the outcome.

References

- Aagaard, K., 2015. How incentives trickle down. *Sci. Publ. Pol.* 42 (5), 725–737.
- Acemoglu, D., 1997. Training and innovation in an imperfect labor market. *Rev. Econ. Stud.* 64 (3), 445–464.
- Barker, K., 2007. The UK Research Assessment Exercise: the evolution of a national research evaluation system. *Res. Eval.* 16 (1), 3–12.
- Bennetot-Pruvot, E., Claeys-Kulik, A.L., Estermann, T., 2015. Strategies for efficient funding of universities in Europe. In: Curaj, A., Matei, L., Pricopie, R. (Eds.), *The European Higher Education Area*. Springer, Cham, pp. 153–168.
- de Boer, H., Jongbloed, B., Benneworth, P., 2015. Performance-Based Funding and Performance Agreements in Fourteen Higher Education Systems: Report for the Dutch Ministry of Education.
- Blöndal, S., Field, S., Girouard, N., 2002. Investment in human capital through upper-secondary and tertiary education. *OECD Econ. Stud.* 34 (1), 41–90.
- Boarini, R., Strauss, H., 2007. The Private Internal Rates of Return to Tertiary Education: New Estimates for 21 OECD Countries. *OECD Economics Department Working Papers*, No. 591. OECD, Paris.
- Bonaccorsi, A., Daraio, C. (Eds.), 2007. *Universities and Strategic Knowledge Creation: Specialization and Performance in Europe*. Edward Elgar Publishing.
- Burke, J.C. (Ed.), 2002. *Funding Public Colleges and Universities for Performance: Popularity, Problems, and Prospects*. State University of New York Press, Albany.
- Canton, E., Venniker, R., Jongbloed, B.W.A., Koelman, J., van der Meer, P., Vossensteyn, J.J., 2001. *Higher Education Reform: Getting the Incentives Right*. SDU, The Hague.
- Cantwell, B., Coates, H., King, R. (Eds.), 2018. *Handbook on the Politics of Higher Education*. Edward Elgar, Cheltenham, UK, and Northampton, MA.
- Chapman, B., 2006. Income contingent loans for higher education: international reforms. In: Hanushek, E.A., Welch, F. (Eds.), *Handbook of the Economics of Education*, vol. 2. Elsevier, Amsterdam, pp. 1435–1503.
- Claeys-Kulik, A., Estermann, T., 2015. *DEFINE Thematic Report: Performance-Based Funding of Universities in Europe*. European University Association, Brussels.
- Clotfelter, C.T., 1996. *Buying the Best: Cost Escalation in Elite Higher Education*. Princeton University Press, Princeton, NJ.
- Deiaco, E., Holmén, M., McKelvey, M., 2010. What does it mean conceptually that universities compete? In: McKelvey, M., Holmén, M. (Eds.), *Learning to Compete in European Universities*. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA, pp. 300–328.
- de Dominicis, L., Pérez, S.E., Fernández-Zubieta, A., 2011. *European University Funding and Financial Autonomy: A Study on the Degree of Diversification of University Budget and the Share of Competitive Funding*. Publications Office of the European Union, Luxembourg.
- Dougherty, K.J., Jones, S.M., Lahr, H., Natow, R.S., Pheatt, L., Reddy, V., 2016. *Performance Funding for Higher Education*. Johns Hopkins University Press, Baltimore.
- Dougherty, K.J., Natow, R.S., 2019. Performance-based funding for higher education: how well does neoliberal theory capture neoliberal practice? *High Educ.* 80 (3), 1–22.
- Ferlie, E., Ashburner, L., Fitzgerald, L., Pettigrew, A., 1996. *The New Public Management in Action*. Oxford University Press, Oxford.
- Garrizmann, J.L., 2016. *The Political Economy of Higher Education Finance. The Politics of Tuition Fees and Subsidies in OECD Countries, 1945–2015*. Palgrave Macmillan, Basingstoke.
- Healy, T., Côté, S., 2001. *The Well-Being of Nations: The Role of Human and Social Capital*. Education and Skills. OECD, Paris.
- Herbst, M., 2007. *Financing Public Universities: The Case of Performance Funding*. Kluwer, Dordrecht.
- Hicks, D., 2012. Performance-based university research funding systems. *Res. Pol.* 41 (2), 251–261.
- Hillman, N., Corral, D., 2017. The equity implications of paying for performance in higher education. *Am. Behav. Sci.* 61 (14), 1757–1772.
- Hillman, N., Tandberg, D., Fryar, A., 2015. Evaluating the impacts of “new” performance funding in higher education. *Educ. Eval. Pol. Anal.* 37 (4), 501–519.
- Jelonek, M., Mazur, S., 2020. Necessary changes, adverse effects? The institutional patterns of adaptation of economics universities to changes prompted by the reform of Poland’s science and higher education system. *Manag. Learn.* 51 (4), 472–490.

- Johnstone, D.B., 2004. The economics and politics of cost sharing in higher education: comparative perspectives. *Econ. Educ. Rev.* 23 (4), 403–410.
- Johnstone, D.B., Marcucci, P.M., 2010. *Financing Higher Education Worldwide: Who Pays? Who Should Pay?* Johns Hopkins University Press, Baltimore.
- Jongbloed, B., Kaiser, F., Westerheijden, D.F., 2019. Improving study success and diversity in Dutch higher education using performance agreements. *Tert. Educ. Manag.* 26 (3), 329–343.
- Jongbloed, B., de Boer, H., McGrath, C., de Gayardon, A., 2022. Study on the State and Effectiveness of National Funding Systems of Higher Education to Support the European Universities Initiative. DG-EAC, Brussels.
- Jongbloed, B.W.A., Vossensteyn, J.J., 2016. University funding and student funding: international comparisons. *Oxf. Rev. Econ. Pol.* 32 (4), 576–595.
- Kaiser, F., Vossensteyn, H., Koelman, J., 2001. *Public Funding of Higher Education: A Comparative Study of Funding Mechanisms in Ten Countries*. Ministerie van Onderwijs, Cultuur en Wetenschappen, Zoetermeer.
- Kelchen, R., 2018. Do performance-based funding policies affect underrepresented student enrolment? *J. High Educ.* 89 (5), 702–727.
- Kivistö, J., Pekkola, E., Lyytinen, A., 2017. The influence of performance-based management on teaching and research performance of Finnish senior academics. *Tert. Educ. Manag.* 23 (3), 260–275.
- Kivistö, J., Kohtamäki, V., 2016. Does performance-based funding work? Reviewing the impacts of performance-based funding on higher education institutions. In: Pritchard, R.M.O., Pausits, A., Williams, J. (Eds.), *Positioning Higher Education Institutions*. Sense Publishers, Rotterdam, pp. 215–226.
- Lasher, W., Greene, D., 1993. College and university budgeting: what do we know? What do we need to know?. In: Smart, J. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. IX. Agathon Press, New York, pp. 428–469.
- Lasher, W.F., Sullivan, C.A., 2004. Follow the money: the changing world of budgeting in higher education. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XIX. Springer, Dordrecht, pp. 197–240.
- Lepori, B., Jongbloed, B., 2018. National resource allocation decisions in higher education: objectives and dilemmas. In: Cantwell, B., Coates, H., King, R. (Eds.), *Handbook on the Politics of Higher Education*. Edward Elgar, pp. 211–228.
- Li, A., 2018. *Lessons Learned: A Case Study of Performance Funding in Higher Education*. Research Gate. Available online at: <https://www.researchgate.net>. (Accessed 13 February 2022).
- Massy, W.F., 2004. Markets in higher education: do they promote internal efficiency? In: Teixeira, P., Jongbloed, B., Dill, D.D., Amaral, A. (Eds.), *Markets in Higher Education: Rhetoric or Reality?* Springer, Dordrecht, pp. 13–35, 2006.
- Mateos-González, J.L., Boliver, V., 2019. Performance-based university funding and the drive towards “institutional meritocracy” in Italy. *Br. J. Sociol. Educ.* 40 (2), 145–158.
- Mouritzen, P.E., Opstrup, N., 2020. *Performance Management at Universities*. Springer, Dordrecht.
- Nelson, R., Phelps, E., 1966. Investment in humans, technological diffusion and economic growth. *Am. Econ. Rev.* 56 (1/2), 69–75.
- OECD, 2008. *Education at a Glance 2008: OECD Indicators*. OECD Publishing, Paris.
- OECD, 2019. *Benchmarking Higher Education System Performance*. OECD Publishing.
- OECD, 2020. *Resourcing Higher Education: Challenges, Choices and Consequences*. OECD Publishing, Paris.
- OECD, 2021. *Education at a Glance 2021: OECD Indicators*. OECD Publishing, Paris.
- Ortagus, J.C., Kelchen, R., Rosinger, K., Voorhees, N., 2020. Performance-based funding in American higher education: a systematic synthesis of the intended and unintended consequences. *Educ. Eval. Pol. Anal.* 42 (4), 520–550.
- Pisár, P., Šipikal, M., Jahoda, R., Špaček, D., 2019. Performance based funding of universities: Czech Republic and Slovakia. In: De Vries, M.S., Nemec, J., Špaček, D. (Eds.), *Performance-Based Budgeting in the Public Sector*. Palgrave Macmillan, Cham, pp. 237–254.
- Psacharopoulos, G., 1981. Returns to education: an updated international comparison. *Comp. Educ.* 17 (3), 321–341.
- Redding, S., 1996. The low-skill, low-quality trap: strategic complementarities between human capital and R & D. *Econ. J.* 106 (435), 458–470.
- Salmi, J., 2003. Constructing knowledge societies: new challenges for tertiary education. *High Educ. Eur.* 28 (1), 65–69.
- Slaughter, S., Rhoades, G., 2004. *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. Johns Hopkins University Press, Baltimore.
- Teixeira, P., Jongbloed, B., Dill, D.D., Amaral, A., 2004. *Markets in Higher Education. Rhetoric or Reality?* Kluwer Academic Publishers, Dordrecht.
- Teixeira, P.N., Johnstone, D.B., Rosa, M.J., Vossensteyn, H. (Eds.), 2006. *Cost-sharing and Accessibility in Higher Education: A Fairer Deal?*. Springer, Dordrecht.
- Van Vught, F., 2008. Mission diversity and reputation in higher education. *High Educ. Pol.* 21 (2), 151–174.
- Ward, J., Ost, B., 2021. The effect of large-scale performance-based funding in higher education. *Educ. Finance Policy* 16 (1), 92–124.
- Ziskin, M., Rabourn, K.E., Hossler, D., 2018. Performance-based funding of higher education: analyses of policy discourse across four case studies. *J. Crit. Educ. Pol. Stud.* 16 (2), 164–215.

Faculty compensation and salary equity

Linda Serra Hagedorn PHD, Iowa State University, Ames, IA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	191
Faculty salary levels	192
Constituents of faculty salary	192
Institutional control	192
Classification systems	192
Rank	192
Full/part-time status	193
Tenure	193
Geographic region within the country	193
Disciplinary differences	193
Union affiliation	194
Bonuses and subsidies	194
Factors for incremental salary increase	195
Cost of living adjustments (COLA)	195
Merit, performance, and productivity	195
Summer compensation	195
Non-monetary compensation (benefits)	195
Medical insurance	196
Life insurance	196
Retirement and savings contributions	196
Tuition assistance	196
Faculty development support	196
Miscellaneous benefits	196
The full spectrum of academic remuneration	196
Negotiation	196
Augmenting salary by non-conventional tactics	197
Moonlighting	197
Salary parity, discrimination, and equity	197
Salary differentials by gender	197
Salary differentials by race	198
Faculty compensation around the globe	198
Conclusions	199
References	200

Introduction

Across the globe, college and university faculty remuneration is multi-faceted and complex. While the overarching question of this article is “what determines the level of faculty compensation,” a secondary question is “what are the non-monetary aspects in faculty remuneration?” Also included will be the critical aspect of salary involving parity and equity. This entry will use the US model as its framework and provide global comparisons of faculty conditions and compensation.

The remuneration of educators has a long history tracing back to ancient China, where the students of Confucius paid for his instruction with whatever they could afford, perhaps just a scrap of meat (Shen, 2014). Practices of educators deriving income from contributions or honorariums were standard across medieval Europe and elsewhere (Xiaojie, 2017). There is, however, evidence that the practice of awarding salaries to faculty emerged sometime between the 13th and 14th centuries; the amount based on the notoriety and reputation of the scholar (Xiaojie, 2017). The modern faculty compensation systems in colleges and universities have evolved to a complex array.

An important note is that the academic profession competes in a global market. When hiring faculty, universities rarely limit their reach to local citizens. Instead, universities seek a global international faculty to convey quality and worldwide prominence. Hence, while the derivation of faculty salary differs across the globe, there are similarities and patterns to attract an international pool of academics.

Faculty salary levels

Salary is computed differently in many countries and may consist of a complex assortment of basic pay plus bonuses, special allowances, supplements, and subsidies (Altbach et al., 2012). However, with some exceptions, perhaps most visibly in Russia, university faculty are respected and receive a level of compensation to support a middle-class lifestyle. Using the US as an example, the US Bureau of Labor Statistics (US Bureau of Labor Statistics, 2020) reports a postsecondary faculty member's 2019 median annual salary was \$79,540. Significant variations can be seen by the wages at the 10th percentile, \$53,000, compared to that at the 90th percentile, \$164,000 (Payscale, 2021) representing a three-fold difference.

Salary ranges tend to be broad globally. For example, the difference between the lowest entry-level ranks and the very top ranks of the professoriate in China is more than four-fold (Altbach et al., 2012). In Italy, the top rank exceeds the entry-level by a factor of 2.6 (Altbach et al., 2012).

Constituents of faculty salary

This section explains the monetary variations in faculty salaries using the US as the primary example and framework. Although the components are examined individually, it is imperative to understand that faculty salary results from the complex interaction of multiple components.

Institutional control

Compensation levels by control follow prestige. In the US, holding all variables equal, private independent institutions generally provide higher compensation, followed by religiously affiliated, followed by public institutions (AAUP, 2020). In Russia, China, and typically throughout the far East, public institutions are afforded the highest prestige and pay much higher salaries than private or proprietary institutions.

Classification systems

In most countries, there is a hierarchy or system of university and college classification stratified by prestige and research prominence. Typically, research universities occupy the top tier, while 2-year vocational or technical colleges can be found at or near the bottom rungs. With some variation, faculty salaries are highest within the most prestigious institutions. China has a three-tier structure for universities. The top tier is universities in the 985 sector, followed by those in the 221 sector, and lastly by private universities (Ma and Wen, 2012). The classification does not recognize China's vocational and less than 4-year institutions.

In the US, institutions are classified by the Carnegie Classification system. The major divisions are doctoral universities, master's colleges and universities, baccalaureate colleges, baccalaureate/associate's colleges, associate's colleges, special focus institutions, and tribal colleges (Carnegie Classification of Institutions of Higher Education, 2021). While there is no general formula or tiers of salaries assigned to classification, on average, salaries are highest at doctoral universities, followed by master's, baccalaureate, and associates (AAUP, 2020). Tribal colleges provide the lowest average salaries of all Carnegie types (Chronicle of Higher Education, 2019). Some colleges, typically community or Associate level, have their salary structure in line with a step process defined by experience and education.

Rank

Globally, rank has evolved with diverse tiers of different levels and subranks. However, no matter how it is structured, rank remains a potent determinant and driver of academic salary. This is especially true for Norway, a Democratic social society, where the faculty salary structure pays little attention to differences by discipline, research productivity, or even seniority leaving rank as the prime determinant (Kyvik, 2012). In the United Kingdom, the salary structure is complex and defined as a "continuous spine" with 51 points from lecturer to professor as a prime salary predictor (Kitagawa, 2012). Similarly, Saudi Arabian faculty salaries depend heavily on a ranking system consisting of 15 steps per rank, each bringing a salary increase (Al-Ohali and Al-Mehrej, 2012). Rank can carry more than financial implications. For example, failure to achieve promotion from lecturer to associate professor generally means dismissal in China.

In the US, rank is the most potent predictor of salary after discipline differences. Table 1 provides the US average full-time salary by rank (Professor, Associate, Assistant) using 2020–2021 dollars. Within almost all US institutions, holding teaching discipline constant, a professor will be paid more than an associate professor, and an associate professor will be paid more than an assistant professor. Promotion from one rank to another typically includes a sizable raise in salary. Violations of the salary level by rank create an adverse condition referred to as salary compression.

Table 1 Average salary in 2020–2021 by rank by AAUP and Carnegie classification.

<i>AAUP category</i>	<i>Professor</i>	<i>Associate professor</i>	<i>Assistant professor</i>
Doctoral	\$159,919	\$104,482	\$91,408
Master	\$104,389	\$83,958	\$73,492
Baccalaureate	\$107,432	\$82,664	\$69,814
Associate (with rank)	\$91,196	\$75,550	\$64,823

Adapted from AAUP Faculty Compensation Survey Results, Table 1.

Full/part-time status

Regardless of country, full-time faculty are almost always paid significantly more than their part-time counterparts (sometimes referred to as adjuncts). While this comparison may be obvious, it is essential to note that a half-time faculty member will likely be paid much less than one-half of the salary allotted to an equivalent full-time member and will typically not receive any benefits. The use of adjunct, hourly, or part-time faculty may bring economic benefits to a university but struggle for faculty who can only find part-time employment. It is not uncommon for a part-time faculty member to work at several universities or colleges to cobble together a liveable wage. Leading the world in the proportion and use of adjunct faculty, Mexico hires most faculty on an hourly basis; 87% at private institutions and 55% at public institutions ([Maldonado-Maldonado, 2012](#)).

In the US, the AAUP Faculty Survey ([2021](#)) reports that the average amount paid to a part-time faculty member to teach a standard course was \$3556. However, the same report documents the wide range as \$548 to \$47,588. [Table 2](#) provides the average compensation for teaching a standard course by institutional type and control.

Tenure

The award of tenure is widely practiced in the US, Canada, and western Europe but has been recently abandoned in The United Kingdom, Australia, and New Zealand. Some countries, including China, use the terms “tenure” or “tenured professor” as a designation for exemplary performance rather than a protection of academic freedom. The practice of awarding tenure in the US is declining and often meets the public’s displeasure. While many US tenure-providing institutions will grant tenure at the promotion from assistant to associate level, the salary increase is due to higher rank and not the receipt of tenure.

Geographic region within the country

Faculty salaries may differ with respect to the location within a given country. To illustrate, in the US, there is significant institutional variation; however salaries tend to be highest in the New England area, followed by the Pacific, middle Atlantic, south Atlantic, northeast central, southwest Atlantic, northwest Central, southeast central, and Mountain region ([AAUP, 2020](#)). [Table 3](#) provides average salaries by geographic region by rank.

Disciplinary differences

In many countries, faculty salaries depend on the discipline. This is especially true in America, where teaching discipline is one of the most significant determinants of differences in faculty salary. Disciplinary salary differences are highly correlated with salaries in the general workforce. For example, a law professor will be paid significantly more than a professor of elementary education because lawyers earn higher salaries than elementary school teachers outside of academe. The lowest salaries are generally in the disciplines of criminal justice, social work, and education. The highest salaries are usually in law, medicine, engineering, accounting, and economics. [Table 4](#) provides an average salary comparison by discipline using a sample of 10 land grant universities for nine-month tenured and tenure-track faculty for academic year 2019–2020 using data collected through the Association of American Universities Data Exchange (AUDE) and the Oklahoma State University Faculty Salary Survey ([Iowa State University, Office of Institutional Research, 2020](#)). The percentages were calculated using the overall average of all ranks and all disciplines for the peer group to a standard (\$125,715).

Table 2 Average compensation for part-time faculty for a standard course.

	<i>Doctoral</i>	<i>Master's</i>	<i>Baccalaureate</i>	<i>Associate colleges (with ranks)</i>	<i>Associate colleges (without ranks)</i>
Public	\$4134	\$3200	\$3708	\$3019	\$2611
Private	\$5050	\$3473	\$4871	No data	No data
Religiously associated	\$5760	\$3254	\$3108	No data	No data

Adapted from AAUP Faculty Compensation Survey Results, Table 15.

Table 3 Average salaries for full-time faculty members by geographic region and rank.

<i>Region</i>	<i>Professor</i>	<i>Associate professor</i>	<i>Assistant professor</i>
New the United Kingdom (CT, ME, MA, NH, RI, VT)	\$168,701	\$108,397	\$92,738
Middle Atlantic (NJ, JY, PA)	\$151,297	\$101,063	\$85,602
East North Central (IL, IN, MI, OH, WI)	\$133,217	\$91,938	\$81,602
West North Central (IA, KS, MN, MO, NE, ND, SD)	\$118,196	\$85,743	\$75,682
East South Central (AL, KY, MS, TN)	\$118,736	\$84,896	\$74,182
West South Central (AR, LA, OK, TX)	\$134,720	\$92,494	\$84,076
South Atlantic (DE, DC, FL, GA, MD, NC, PR, SC, VI, VA, WV)	\$134,236	\$93,543	\$84,076
Mountain (AZ, CO, ID, MT, NV, NM, UT, WY)	\$120,066	\$90,300	\$77,301
Pacific (AK, CA, HI, GUM, OR, WA)	\$156,420	\$107,255	\$93,155

Adapted from AAUP Faculty Compensation Survey Results, Table 4.

Table 4 Average faculty salary comparisons of full-time tenure/tenure-tracked faculty by select disciplines for a peer of land grant universities.

	<i>Percentage of the average of all ranks and disciplines and ranks—(\$125,715)</i>
Finance	196.5%
Accounting	177.4%
Marketing	160.0%
Electrical and Computer Engineering	118.8%
Chemistry	112.5%
Aerospace Engineering	109.2%
Political Science	108.0%
Economics	106.4%
Genetics	100.5%
Psychology	99.3%
Civil, Construction, and Environmental Engineering	98.3%
Mathematics	94.0%
History	84.8%
Kinesiology	84.3%
Architecture	80.7%
English	78.2%
Music and Theater	74.2%
Education	72.8%
World languages	70.7%

Institutions: University of California-Davis, University of Illinois- Urbana, Michigan State University, University of Minnesota, North Caroline State University, Ohio State University, Pennsylvania State University, Purdue University, Texas A&M University, University of Wisconsin-Madison.

Union affiliation

Unions and collective bargaining play various roles in the computation of faculty salary globally due to diverse definitions in different countries. In Israel and Australia, salary levels and salary differentiation rely heavily on collective wage bargaining agreements; and much less on discipline or faculty productivity ([Guri-Rosenblit, 2012](#); [Welch, 2012](#)).

In the US, union affiliation is confounded by its prevalence only in public institutions and lower Carnegie levels ([Ehrenberg et al., 2016](#)). Union affiliation is not a dependable predictor of salary level. While some studies have identified a minimal positive effect, others have found no effect or even a negative one.

Bonuses and subsidies

Bonuses take on a multitude of forms. Dutch faculty can count on an end-of-the-year bonus plus a holiday bonus paid in December and May, respectively. Combined, these two bonuses account for 14% of the total faculty compensation ([Jongbloed, 2012](#)).

In China, bottom-line salaries are determined by a complex array of base wages and subsidies. Subsidies can include food and housing and even items such as laundry, telephone, and transportation ([Ma and Wen, 2012](#)). Similarly, in India, the salary of an academic includes components for transportation and house rent ([Jayaram, 2012](#)).

Bonuses can be awarded for any number of reasons including, teaching an extra course, taking on additional students, chairing committees, or anything seen to be above the typical assignment. An administrative stipend may be added to salary in some institutions if the faculty member serves as department chair, faculty leader, or head of an important committee.

Factors for incremental salary increase

In addition to the factors identified, faculty salaries are also subject to cyclic or time-based increases. These salary changes are neither guaranteed nor fixed in amount but rather depend on circumstances and institutional will.

Cost of living adjustments (COLA)

In many countries, there is a process for salary increases with time. It may depend on union negotiation or just governmental policy when academics are civil servants. One of India's five components of academic salaries consists of a "dearness allowance," an amount provided to compensate for inflation (Jayaram, 2012). In Nigeria, faculty salaries are reviewed every three years and can be increased to reflect inflation and remain competitive with other African nations (Bamiro, 2012). In the US, university salary structures generally allow for a cost-of-living adjustment (COLA) that is neither guaranteed nor automatic and available only when the university is in a financial state to afford it. It is not unusual for colleges and universities to entirely forego COLA due to financial concerns. For example, the Coronavirus Pandemic of 2020–2021 produced low enrollments that led to tuition freezes and economic disruption that created hiring freezes, layoffs, and no raises for faculty and staff in many institutions.

Merit, performance, and productivity

Merit raises, or pay for productivity, are not automatic. Many research universities reward faculty productivity, generally defined as publications, patents, international awards, and other scholarly products with salary increases. Merit raises are as much of a retention device as a reward. At the extreme, faculty who have received international recognition like a Nobel Prize or a Pulitzer Award will undoubtedly be rewarded monetarily and sought by other universities. In the US, one of the highest honors is membership to one of the National Academies of Science. In Mexico, the highest national honor is membership in the National System of Researchers (Sandoval-Romero and Larivier, 2020).

Merit is generally calculated through the annual evaluation process conducted at virtually all US colleges and universities. The annual performance process can be very detailed, requiring faculty to fill out detailed evaluation rubrics, assemble a portfolio of annual work, or write descriptive narratives describing their work over the academic year. The evaluation criteria are generally in line with the institutional mission and identify what the institution most values. Hence, research universities will generally award merit increases for scholarly activity (publications, grants, patents, etc.), while institutions that emphasize teaching may provide increases based on teaching evaluations and service. Faculty in the performance, fine, or visual arts will be evaluated on creative activities (exhibitions, recitals, productions, etc.).

Summer compensation

Academic years are calculated differently across the globe. In the US, Canada, and some parts of Europe, faculty are hired on a 9- or 10-month contract and may be able to augment salary through teaching and/or research during the off-contract time (summer). Summer compensation is varied, limited, and defined by university policy across institutions.

During the summer term, compensation for teaching is typically based on the number of courses or credits taught and the faculty member's base salary. In some institutions, the amount per course is based on faculty rank. The maximum allocation is normally one-third of the 9-month base salary.

Faculty may also be able to derive summer compensation through participation in research. Summer research support is often funded through external contracts and grants. In other words, the funder pays the institution to employ the faculty member to perform research as proposed through the provisions of the grant.

Non-monetary compensation (benefits)

Compensation is not solely measured by salary. In comprehending any compensation package, one must consider other aspects often referred to as "benefits." For some faculty, the importance and value of these benefits may far outweigh the salary amount. It is impossible to discuss all of the variations in benefits across the globe. In Mexico, for example, benefits may include food coupons, child care assistance, and even funeral expenses (Maldonado-Maldonado, 2012).

The following sections represent some of the more common benefits but do not represent an exclusive list. Like salary amounts, benefits vary drastically across colleges and universities based on institutional type, control, location, size, collective bargaining unit, status (full time, tenured, etc.), and classification.

Medical insurance

Medical insurance is often tied to employment. In the Czech Republic, health benefits are obligatory and paid to all employees by every employer (Sebkova, 2012). In Nigeria, faculty receive free medical treatments under their National Health Insurance (Bamiro, 2012). In other countries, like Japan, France, and Germany, faculty are considered civil servants and receive health insurance in concert with all other civil servants (Shima, 2012). Even Russia provides faculty health care (Androushchak and Yudkevich, 2012).

In the US, medical insurance is reserved only for full-time faculty and represents a highly sought benefit, especially for faculty with dependents. Although the institution may invite faculty to join the institutional plan, faculty almost always pay a significant portion of the premiums. Packages may include single coverage for the faculty member with options to include immediate family members. The plans may span the options of HMO, PPO, or other structures and may consist of coverage in part for dental, psychological, and or healthy lifestyle programs.

Life insurance

Life insurance is not standard from an international perspective. However, in the US, many colleges and universities will offer faculty a basic term life insurance policy and a long-term disability policy covering the faculty member while employed at the institution. There may be options to upgrade to higher levels of coverage at the faculty member's expense.

Retirement and savings contributions

A pension of some sort is generally awarded to academics in countries where faculty are deemed civil servants. In the US, retirement funds typically allow a faculty member to invest a specified percentage of salary to a tax-deferred account to which the institution will also contribute. These contributions grow tax-deferred and are designed to create a revenue flow post-retirement. According to the latest AAUP report (2020), close to 97% of US full-time faculty have retirement programs with funds contributed in part by their employing institution. However, the options within retirement and savings contributions are highly variable from institution to institution. While some institutions do not offer an option, recent evidence indicates that some institutions are either cutting their contributions or suspending them entirely (Flaherty, 2020). Among public institutions, there may be an option to enroll in a state public employee's system.

Tuition assistance

In many countries, full-time faculty members and sometimes their immediate dependents can enroll in courses at a reduced or free cost. In the US, about one in four institutions will offer tuition waivers (full or part) through a consortium of institutions (AAUP, 2020).

Faculty development support

The area of faculty development is globally broad and diverse. Faculty development may include institutional support (full or partial) to include expenses related to attending a conference in a faculty member's discipline. Or it could be support to attend training on diversity, equity, teaching, or another aspect of faculty responsibilities.

Miscellaneous benefits

Depending greatly on institutional type, other forms of faculty benefits may also be available. Examples include a reduced teaching load, funded graduate assistants, laboratories, and supplies, technology, including computers and other equipment. At onboarding, relocation expenses may be offered.

The full spectrum of academic remuneration

Faculty remuneration is as varied as the institutions themselves. There are also other and diverse ways that faculty salaries may be fixed and/or augmented.

Negotiation

In countries where salaries are not fixed by law, negotiations are an essential part of salary level. It is important to note that the starting salary accepted by a new faculty member affects the level of remuneration for as long as that faculty member remains employed by the institution and beyond. In some countries, retirement benefits rely on the salary earned. Typically merit and COLA increases are calculated as a percentage of the base salary, and if the base wage starts low, it will likely remain so. However,

in countries where negotiations are possible, offers can and probably should be negotiated. Moreover, it is not only salary level, but also other benefits that may be negotiated including, but not limited to, start-up funds, conference-travel expenses, moving expenses, and even a position for a spouse or partner (Kelsky, 2014).

Augmenting salary by non-conventional tactics

Some faculty members dissatisfied with their remuneration seek relief by seeking a position or an offer from another institution. In the US, it is not uncommon for a faculty member to get an offer from another college/university that includes a better compensation package. When the offer is revealed to the current institution, it may trigger a retention offer that matches or exceeds the offer. A word of caution here, some faculty members have unsuccessfully pursued the “competing offer” for the sole purpose of getting the offer matched at their current institution without sincere intention to accept it. This tactic has been known to backfire when the current institution bids a warm farewell rather than the desired retention package.

Moonlighting

Although technically not a part of faculty salary, moonlighting or adding part-time work elsewhere is commonplace and even expected in some countries. It is essential to augment low faculty salaries in Russia as low academic wages are insufficient to support the academic and family. In China, India, and Nigeria, it is common for faculty to add teaching an evening class at a private college (Bamiro, 2012; Ma and Wen, 2012).

Salary parity, discrimination, and equity

Unfortunately, disparities regarding pay parity and discrimination in the workplace exist in virtually all industries; and academia is no exception. In the US, wage discrimination based on gender, race, sexual orientation, age, or disability is prohibited under the Equal Pay Act, Title VII, The Americans with Disabilities Act, and the Age Discrimination in Employment Act. Faculty members who identify discrimination or can make a case that others with similar qualifications and experience receive higher compensation may begin with a discussion with their department chair and/or Dean to request investigation and possible adjustment. There have been many documented cases of discrimination and subsequent settlements involving faculty remuneration. For example, a 2018 case involving the University of Denver resulted in an order to pay a \$2.6 million settlement to seven female law professors who proved that they were being paid less than their male counterparts (Flaherty, 2018).

Salary differentials by gender

Holding all other variables constant, on average, female faculty earn less than their male counterparts worldwide. Of course, gender gaps vary widely by country. In Europe, the gender gap is especially wide in Estonia, the Czech Republic, Israel, and Portugal, reaching levels as high as 35% (European Commission, 2007). In Canada, a country that aggressively addresses discrimination, gender gaps were identified in all but three of its large research institutions, the greatest being more than 15% (Cummings, 2020). Although high-rank female representation within elite university faculties is below parity, Australia boasts small gender pay gaps due to union-negotiated agreements (Bailey et al., 2016). However, Bailey et al. point out that Australian base salary comparisons mask gender differentials due to the widespread use of bonuses (2016).

Specific faculty gender gaps in China, Japan, Korea, and India are not routinely published, but the broad evidence suggests that gender gaps exist and are likely comprehensive.

Studies of salary parity by gender for college and university faculty, specific to the US, date back to the early 1970s. Hagedorn (1996, 1998), using data from the 1989 Carnegie Foundation for the Advancement of Teaching national survey of faculty, reported a mean differential of \$8840, which decreased to \$2456 after applying a method to isolate only discriminatory differences. Barbezat and Hughes (2005), using national data from 1999, reported males earning 20.7% more than females with a parity gap of between 17% and 23% depending on the decomposition process used. Table 5 provides salary data from AAUP's most current study (2020).

Table 5 Average salary for full-time men and women (all AAUP/Carnegie categories combined) 2020–21.

Rank	Male average salary	Female average salary
Professor	\$147,042	\$127,819
Associate	\$98,995	\$92,208
Assistant	\$87,530	\$79,621
Instructor	\$65,544	\$61,034
Lecturer	\$69,857	\$64,187
No rank	\$78,340	\$69,828

Adapted from AAUP Faculty Compensation Survey Results Table 3.

In addition to the differentials, the AAUP also reports that female faculty have less security (2020). In addition, numerous universities have performed internal analyses, with many finding some evidence of inequities by gender.

Acker (1989, 1990, 2006) identifies universities as gendered organizations that manifest an uneven distribution of power and position, rendering males more powerful and entitled. Studying science and engineering, Johnson and Taylor (2019) identified one of the major culprits to gender salary inequity as “masculine norms and male majority” (p. 21). They used data from the National Science Foundation, the NSF Survey of Earned Doctorates, and the Integrated Postsecondary Data Systems (IPEDS) to find a gender gap at all ranks averaging \$5211 (2019).

Salary differentials by race

Many countries have a relatively homogeneous faculty population regarding race, leaving racial differences in compensation unstudied. However, based on the significant evidence of racial salary gaps in most professions in multi-racial countries, such as the United Kingdom, Australia, and the US, it is safe to assume that differences by race extend to academe. South Africa, a country where academic compensation is exceptionally high, presents a picture of racial differentials that present less directly than a comparison of White and non-White faculty salaries. Although apartheid has been outlawed, colleges and universities previously designated for Black students remain underfunded. These universities cannot compete in the hiring nor compensating of star faculty (Sehoole, 2012).

For a more detailed look at racial differences, we delve into the conditions in the U.S. Discrimination by race is less evident by salary amount than by representation. Simply stated, people of color are underrepresented in academia. Only 22% of all full-time faculty are people of color; a percentage that declines if disaggregated by rank or administrative duty. Only 16% of department chairs are people of color (Sziron, 2020). However, when analyzing the wages paid to faculty of color, the picture is not simple. Unfortunately, the U.S. Department of Education (IPEDS) does not collect salary data disaggregated by race or gender. An excellent project by Li and Koedel (2017) analyzed data from 40 selective public universities using a regression approach. After controlling for university effects, field effects, the prestige of Ph.D. granting institution, experience, and faculty h-index, a negative wage gap remained only for Asian faculty (\$−1409.96). Black and Hispanic faculty had small but positive wage gaps of \$3316.14 and \$1201.72, respectively. However, these small differences were explained by further analyses controlling for academic field, experience, and research productivity, thus concluding that there is neither evidence of wage discrimination nor wage premiums for faculty of color. Interpretation is that for those faculty of color who have navigated to faculty positions, there may not be a wage gap. However, one must remember that the controls used in Li and Koedel’s research could not uncover discrimination or bias in promotions or barriers within the structures of academic disciplines.

To better understand faculty salary by race/ethnicity in the absence of national data requires reliance on individual university or university systems. Following are four large and representative studies that are typical of the type run by universities.

- The University of Arizona conducted a comprehensive study of base salary for tenured and tenure-eligible faculty and reported no systemic disparities by gender, race or ethnicity (Arizona Board of Regents, 2021). A subsequent review at the unit level did reveal “correctable salary variations,” which were addressed.
- UCLA’s Faculty Equity Report of July 2020, analyzing 2018–19 data across four university units, reported lower salaries for under-represented minority (URM) faculty compared to their White counterparts. After controlling for a set of variables, the report concluded that race differentials were no longer statistically significant (UCLA Office of the Executive Vice Chancellor and Provost, 2020).
- The University of Oregon hired outside consultants to investigate salary inequities by gender and race. The report concluded that no protected group was systematically receiving lower than predicted salaries (Berkeley Research Group, 2019).
- The University of North Carolina reported that URM faculty were less likely to be tenured and more likely to rank below professor than their White counterparts. However, after controlling for the usual variables, reported no statistical salary differences (Office of Institutional Research & Assessment of the University of North Carolina at Chapel Hill, 2017).

In conclusion, universities report few or no statistically significant differences by race or ethnicity, based on available institutional data. However, inequities do exist when viewing the number of faculty of people of color. Moreover, there is also evidence that race may play an inappropriate role in promotions of rank and the granting of tenure.

Faculty compensation around the globe

It is an accepted fact that the strength and prosperity of a nation rely on an educated citizenry where universities play a dominant role. It follows, therefore, that the quality of the university lies squarely on the qualifications and quality of the faculty it employs. Many countries are experiencing an increase in students seeking postsecondary degrees and credentials, thus promoting growth. Moreover, a positive step to becoming an internationally respected university requires attracting an international faculty.

Realistically speaking, university structure and remuneration vary vastly from country to country and are difficult to impossible to compare directly. However, to get a general idea of the level of compensation provided to faculty, I provide a comparison using the US as the standard and convert all salaries to U.S. dollars. Table 6 provides the annual mean salaries of a university professor in

Table 6 International average salaries (local and USD currencies) of university professors and cost of living.

	Country	Mean annual base salary in local currency	Mean salary in \$(USD)	Cost of living index
North America	Canada	108,183 (CAD)	\$89,586.34	70.08
	Mexico	393,849 (MXN)	\$19,727.90	35.05
	USA	104,716 (USD)	\$104,716.00	71.92
Central America	Costa Rica	18,540,650 (CRC)	\$29,850.45	50.64
	Guatemala	399,186 (GTQ)	\$18,137.82	44.76
South America	Argentina	1,100,112 (ARS)	\$11,606.18	37.87
	Brazil	132,879 (BRL)	\$26,323.33	31.82
	Chile	26,257,047 (CLP)	\$36,523.55	49.33
	Peru	56,176 (PEN)	\$14,628.23	36.26
Europe	France	60,457 (EUR)	\$73,564.08	80.62
	Germany	75,037 (EUR)	\$91,305.02	70.62
	Italy	56,446 (EUR)	\$68,683.49	73.11
	Poland	139,284 (PLN)	\$37,996.68	42.04
	Spain	47,754 (EUR)	\$58,107.07	59.09
Asia	China	275,109 (CNY)	\$42,999.54	42.54
	India	9,48,017 (INR)	\$13,006.79	25.16
	Indonesia	350,891,733 (IDR)	\$24,537.86	37.44
	Japan	9,552,467 (JPY)	\$87,194.92	87.77
	South Korea	55,336,184 (KRW)	\$49,835.77	81.2
	Pakistan	2,045,493 (PKR)	\$13,111.61	21.59
	Vietnam	483,390,908 (VND)	\$20,969.50	38.05
Mideast	Egypt	192,897 (EGP)	\$12,254.75	30.1
	Iraq	33,459,296 (IQD)	\$22,916.27	32.72
	Israel	227,077 (ILS)	\$69,803.47	86.63
	Saudi Arabia	196,736 (SAR)	\$52,449.82	49.95
	Turkey	105,529 (TRY)	\$12,156.94	31.48
	Russia	1,115,433 (RUB)	\$15,292.59	33.17
Scandinavia	Denmark	572,347 (DKK)	\$93,635.97	91.67
	Finland	70,348 (EUR)	\$85,606.48	77.46
	Iceland	10,517,856 (ISK)	\$86,909.04	96.77
	Norway	776,707 (NOK)	\$93,515.52	106.09
	Sweden	566,762 (SEK)	\$68,521.53	79.17
Oceania	Australia	125,591 (AUD)	\$97,219.99	84.14
	New Zealand	114,819 (NZD)	\$82,738.57	79.14
Africa	South Africa	664,086 (ZAR)	\$49,427.92	41.69
United Kingdom	The United Kingdom	51,884 (GBP)	\$73,504.06	71.03
	Scotland	50,467 (GBP)	\$71,496.60	71.03

37 countries. Like all averages, these numbers mask significant differences in regions and institutional types. As reported by Salary Expert, these annual averages are powered by the Economic Research Institute ([Economic Research Institute, 2021](#)). Conversions to USD were done through Google and reflect the conversion rates on June 6, 2021. Notably, the highest salary is paid in the US. To further provide context, [Table 6](#) also includes a cost-of-living indicator as recorded for 2021 by Numbeo ([Salary Expert, 2021](#)). These indices are calculated relative to New York City, receiving an index of 100. A broad interpretation can be understood from taking the example of Argentina, which is the lowest salary in the chart (\$11,606.18) but is in an environment where the cost of living is only 38% of that of New York City. Mathematically, this salary would equate to earning \$30,542.58 while living in New York.

Conclusions

Some countries do not extend autonomy to assign faculty salaries to the institutions; instead, colleges and universities must adhere to federal regulations. In most aspects of faculty salary and benefits, there is wide variation both inter and intra-country. Rank and its relation, longevity, are strong predictors across the globe. While merit is a predictor of salary in many locales, it is not universal.

The faculty salary structures reflect national and institutional values. It is not by accident that education faculty will typically earn less than business faculty. Nor is it by chance that female faculty salaries on average are less than their male counterparts. Moreover, salary differences by discipline may reflect profound cultural devaluation of women through inequitable salaries for faculty in fields predominantly female such as education, social work, and consumer sciences ([Bellas, 1994](#)). Universities are microcosms of the countries in which they reside and reflect both the positives and the negatives cultural norms.

Discrimination is devious and can escape scrutiny controlling for the salary building blocks. For example, faculty salary studies controlling for a long list of variables, including rank, mask any discriminatory practices in awarding rank.

This discussion of academic compensation did not include the costs and investment in time as well as opportunity costs by the faculty themselves. To be an academic requires extensive graduate education, most often culminating with a Ph.D. or equivalent. Many faculty had to rely on student loans to pay for their education and must pay these loans back over many years with interest. After receipt of the terminal degree, some future faculty must first accept a very low-paying post-doc position for a year or more in preparation for the actual role of faculty. While a faculty member's job certainly has its advantages, many faculty could quickly point out the disadvantages as well. Most faculty work long hours and balance high expectations in teaching, research, and service. Disentangling the sacrifices and expenses alongside the benefits was not a question considered in this article. However, in most parts of the world, those who have persevered and have achieved a high academic rank are compensated relatively well compared to their country's average compensations.

References

- AAUP, 2020. The annual report on the economic status of the profession, 2019–2020. *Academe* 106 (3), 21–49.
- AAUP, 2021. Faculty Compensation Survey Results. Retrieved from AAUP. <https://www.aaup.org/2020-21-faculty-compensation-survey-results>.
- Acker, J., 1989. Doing Comparable Worth: Gender, Class and Pay Equity. Temple University Press, Philadelphia, PA.
- Acker, J., 1990. Hierarchies, jobs, bodies: a theory of gendered organizations. *Gen. Soc.* 4, 139–158.
- Acker, J., 2006. *Class Questions: Feminist Answers*. Rowman and Littlefield, Lanham, MD.
- Al-Ohali, M., Al-Mehrej, H., 2012. Faculty salary and remuneration in the Kingdom of Saudi Arabia. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 279–287.
- Altbach, P.G., Reisberg, L., Pacheco, I.F., 2012a. Academic remuneration and contracts. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate: A Global Comparison of Compensation and Contracts*. Routledge, New York, pp. 3–20.
- Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F., 2012b. *Paying the Professoriate: A Global Comparison of Compensation and Contracts*. Routledge, New York.
- Androushchak, G., Yudkevich, M., 2012. Russian higher education. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 265–278.
- Arizona Board of Regents, 2021. Report on 2020 Salary Equity Review for Tenure/Tenure Eligible Faculty. The University of Arizona. Retrieved from. <https://provost.arizona.edu/content/report-2020-salary-equity-review-tenure-tenure-eligible-faculty>.
- Bailey, J., Peetz, D., Strachan, G., Whitehouse, G., Broadbent, K., 2016. Academic pay loadings and gender in Australian universities. *J. Ind. Relat.* 58 (5), 647–668.
- Bamiro, O.A., 2012. Nigeria: toward an open market. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 245–254.
- Barbezat, D.A., Hughes, J.W., 2005. Salary structure effects and the gender pay gap in academia. *Res. High. Educ.* 46 (6), 621–640.
- Bellas, M.L., 1994. Comparable worth in academia: the effects on faculty salaries of the sex composition and labor-market conditions of academic disciplines. *Am. Socio. Rev.* 59, 807–821.
- Berkeley Research Group, 2019. Analyses of 2018 Total Salary Differences by Gender and by Race/Ethnicity. Berkeley Research Group, Berkeley, CA. Retrieved May 2021, from. https://provost.uoregon.edu/files/consultant_report_for_web.pdf.
- Carnegie Classification of Institutions of Higher Education, 2021. 2018 Update Facts & Figures. The Carnegie Classification of Institutions of Higher Education. Retrieved 2021, from. <https://carnegieclassifications.iu.edu/downloads/CCIHE2018-FactsFigures.pdf>.
- Chronicle of Higher Education, 2019. Chronicle Data. Retrieved from Chronicle Data: <https://data.chronicle.com/category/ccbasic/16/faculty-salaries/>.
- Cummings, M., 2020. Pay Gap Between Male and Female Professors Continues to Plague Canadian Universities. Edmonton, Alberta, Canada. Retrieved September 12, 2021, from. <https://www.cbc.ca/news/canada/edmonton/gender-pay-gap-persists-at-canadian-universities-1.5739466>.
- Economic Research Institute, 2021. Accurate Compensation Data at Your Fingertips. Economic Research Institute. Retrieved from. https://www.eri.com/?__hssc=33573502.d54323959eccc308fccc8e8505e21aa.1622571176103.1622574718523.1622585880430.3&__hssc=33573502.5.1622585880430&__hsfp=3792209412.
- Ehrenberg, R.G., Klaff, D.B., Kezsbom, A.T., Nacowski, M.P., 2016. Collective bargaining in American Higher Education. In: Ehrenberg, R.G. (Ed.), *Governing Academia: Who is in Charge at the Modern University?* Cornell University Press, Ithaca, NY, pp. 209–232.
- European Commission, 2007. *Remuneration of Researchers in the Public and Private Sectors*. Office of SDME, Brussels.
- Flaherty, C., 2018. Closing the Pay Gap. *Inside Higher Ed.* May 18, 2018. U. Of Denver Settles With EEOC, Agreeing to Pay \$2.66 Million to Seven Female Law Professors Who Alleged Gender-Based Pay Bias. [insidehighered.com](https://www.insidehighered.com).
- Flaherty, C., 2020. Colleges Lower the Boom on Retirement Plans. *Inside HigherEd.* May 21, 2020. More Institutions are Suspending or Cutting Retirement Plan Contributions. [insidehighered.com](https://www.insidehighered.com).
- Guri-Rosenblit, S., 2012. Israel academic salaries and remuneration. In: Altbach, P.G., Reisberg, L.Y., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 166–175.
- Hagedorn, L.S., 1996. Wage equity and female faculty job-satisfaction: the role of wage differentials in a job satisfaction causal model. *Res. High. Educ.* 37 (5), 569–598.
- Hagedorn, L.S., 1998. Implications to postsecondary faculty of alternative calculation methods of gender-based wage differentials. *Res. High. Educ.* 39 (2), 143–162.
- Iowa State University, Office of Institutional Research, 2020. Peer Land Grant Universities Faculty Salary Comparisons 2009–2020 Through 2019–2020. Internal report, Ames, Iowa.
- Jayaram, N., 2012. Academic salaries and career advancement: tuning the professoriate for a knowledge economy in India. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 155–165.
- Johnson, J.A., Taylor, B.J., 2019. Academic capitalism and the faculty salary gap. *Innovat. High. Educ.* (44), 21–35.
- Jongbloed, B., 2012. Introducing market forces in academic remuneration. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 234–244.
- Kelsky, K., 2014. The Professor is in: OK, Let's Talk About Negotiating Salary. Chronicle of Higher Education. Retrieved January 5, 2021, from. <https://community.chronicle.com/news/400-the-professor-is-in-ok-let-s-talk-about-negotiating-salary>.
- Kitagawa, F., 2012. Academic salary in the United Kingdom: marketization and national policy development. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 308–317.
- Kyvik, S., 2012. Academic salaries in Norway: increasing emphasis on research achievement. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 255–264.

- Li, D., Koedel, C., 2017. Representation and salary gaps by race-ethnicity and gender at selective public universities. *Educ. Res.* 46 (7), 343–354.
- Ma, W., Wen, J., 2012. A study on academic salary and remunerations in China. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 94–103.
- Maldonado-Maldonado, A., 2012. Mexican faculty salaries today. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 225–233.
- Office of Institutional Research, Assessment of the University of North Carolina at Chapel Hill, 2017. Faculty Salary Equity Report. The University of North Carolina at Chapel Hill, Office of Institutional Research & Assessment, Chapel Hill. Retrieved May 2021, from: <https://facultygov.unc.edu/wp-content/uploads/sites/261/2019/03/2016-17-Faculty-Salary-Equity-Report-to-Provost-Final-March-2017.pdf>.
- Payscale, 2021. Average Professor, Postsecondary/Higher Education Salary. Payscale. Retrieved 2021, from: https://www.payscale.com/research/US/Job=Professor%2C_Postsecondary_%2F_Higher_Education/Salary.
- Salary Expert, 2021. Research Salaries. Retrieved from Salary Expert Powered by ERI. <https://www.salaryexpert.com/salary>.
- Sandoval-Romero, V., Larivier, V., 2020. The national system of researchers in Mexico: implications of publication incentives for researchers in social sciences. *Scientometrics* 122, 99–126.
- Sebkova, H., 2012. The Czech Republic: high estimation for the academic profession. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 114–124.
- Sehoole, C.T., 2012. The unequal playing field. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 288–296.
- Shen, V., 2014. *Dao Companion to Classical Confucian Philosophy*. Springer, Netherlands. <https://doi.org/10.1007/978-90-481-2936-2>.
- Shima, K., 2012. Working conditions and salaries of the academic profession in Japan. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 185–195.
- Sziron, M., 2020. 2019–20 Trends in Faculty Hiring and Compensation. *HigherEdJobs*. <https://www.higheredjobs.com/articles/articleDisplay.cfm?ID=2184>.
- U.S. Bureau of Labor Statistics, 2020. Education, Training, and Library Occupations. Retrieved 2021, from Occupational Outlook Handbook. www.bls.gov/ooh/education-training-and-library/home.htm.
- UCLA Office of the Executive Vice Chancellor and Provost, 2020. UCLA Faculty Salary Equity Report for the Fiscal Year 2018–2019. UCLA, Los Angeles.
- Welch, A., 2012. Academic salaries, massification, and the rise of an underclass in Australia. In: Altbach, P.G., Reisberg, L., Yudkevich, M., Androushchak, G., Pacheco, I.F. (Eds.), *Paying the Professoriate*. Routledge, New York, pp. 61–71.
- Xiaojie, Z., 2017. The faculty salary system of the Medieval University. *Chin. Stud. Hist.* 1 (50), 14–23.

Knowledge diplomacy: the role of international higher education, research and innovation in international relations.

Jane Knight, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

The expanding and changing role of international higher education in international relations	202
Toward a working definition of knowledge diplomacy	203
Why study knowledge diplomacy	204
Differentiating knowledge diplomacy from cultural, science, education diplomacy	204
Misconceptions about knowledge diplomacy	204
Knowledge diplomacy is broader than cultural diplomacy	204
Knowledge diplomacy is inclusive of science diplomacy but more comprehensive	204
Why not use the term education diplomacy?	205
Knowledge diplomacy is different from soft power	205
A soft power approach juxtaposed to a diplomatic approach	205
Key characteristics of knowledge diplomacy	206
Diversity of Actors and Partners	206
Focus on Higher education, research and innovation	206
Recognition of different needs and collective use of resources of actors	206
Reciprocity -mutual but different benefits	207
Based on negotiation, collaboration, cooperation	207
Different levels of collaboration	207
Commonality of issues – addressing regional and global issues	207
Build and strengthen relations between and among countries	207
Examples of knowledge diplomacy initiatives	207
Issues and challenges	209
References	209

The expanding and changing role of international higher education in international relations

International higher education (IHE) has a long and rich history in relation to its contribution to relations between and among countries. But in today's more complex, interdependent and globalized world there are new rationales, opportunities, benefits and risks attached to the role and contribution of higher education and research to international relations. Examining the role of IHE in building relations between and among countries is not new. However, using a knowledge diplomacy framework for analysis rather than the traditional lens of cultural diplomacy and soft power is new.

The purpose of this discussion is to promote a greater awareness and understanding of the role and dimensions of knowledge diplomacy. The first section examines the importance of studying the role of higher education in international relations followed by a discussion on defining knowledge diplomacy and clarifying the fundamental differences between cultural diplomacy, science diplomacy, soft power and knowledge diplomacy. The next section identifies and examines eight primary characteristics of knowledge diplomacy which leads to brief summaries of some case studies which demonstrate knowledge diplomacy in action. The examples of knowledge diplomacy focus on how through cooperation, reciprocity and mutuality of benefits international higher education can strengthen relations between and among countries to address current global issues using international higher education, research and innovation (IHERI) as key strategies. The last section identifies some issues and challenges which need to be addressed as knowledge diplomacy becomes better understood, researched and operationalized.

Four concepts form the backbone of this analysis of knowledge diplomacy. They are higher education, research, innovation and international relations. It is important to explain how these concepts are used as they are not formal scholarly definitions per se but are more descriptive and aim to orient the reader as to how they are used.

- Higher education refers to the processes of teaching and learning whether it be in a formal or informal manner using a broad range of strategies. Training is included
- Research includes the gathering of information on a subject, investigation, or experimentation aimed at producing new knowledge
- Innovation refers to the application of research findings and new knowledge to produce change or new ideas
- International relations refers to the bilateral or multi-lateral relations between and among countries
- Knowledge, in the term knowledge diplomacy, is used in a representational and inclusive sense not literally. It refers to the combination of higher education, research and innovation

Toward a working definition of knowledge diplomacy

Diplomacy has a very long history. As a result, there are a myriad of interpretations of diplomacy. Many scholars and professionals believe that there is not one universally accepted definition of diplomacy as it is often defined in terms of issues, interests, approaches and strategies used. But the concept most fundamental to diplomacy and most frequently cited is the “the management of relations between and among countries” (Kerr and Wiseman, 2013).

The traditional way of understanding diplomacy focuses on the role of the government, usually the Department of Foreign Affairs and their representatives in embassies around the world. Primary tasks included representation of the country’s interests, negotiation of agreements and treaties, as well as mediation of conflict (Pigman, 2010). There is no question that projection and protection of self-interests are front and center, but diplomacy concentrates on finding solutions through negotiation, mediation and compromise. Security and economic development have been the central issues in traditional diplomacy and still are. But, during the last three decades, diplomacy has changed significantly (Bjola and Kornprobst, 2013).

In contemporary diplomacy, there are a broader set of actors, a greater set of issues and challenges to be addressed, as well as new approaches and strategies being used. It is no longer only the government and their representatives who are seen as the major agents, there are a diversity of new actors including civil society, multinationals and government departments other than foreign affairs (Cooper et al., 2013). Theme-based diplomacies such as health, environmental, trade, refugee diplomacy have grown in importance. New approaches and instruments such as track-two, digital, summit and public diplomacy are being used.

Knowledge diplomacy is one of the new types of diplomacy emerging in contemporary times. An early interpretation of the term related to the negotiation of intellectual and property rights by the World Trade Organization members (Ryan, 1998). However, defining knowledge diplomacy in relation to intellectual property rights in trade agreements was a very narrow and a short-lived approach. A more recent conceptualization of knowledge diplomacy has broadened the focus to “the role of international higher education, research and innovation (IHERI) in building/strengthening relations between and among countries” (Knight, 2022). This acknowledges the growing importance of higher education in the more inter-connected and interdependent world in which we live and how it contributes to international relations and the challenges inherent in addressing current global issues.

Knowledge diplomacy can be understood as a two-way process as depicted in Fig. 1. Knowledge diplomacy includes the role of international higher education, research and innovation (IHERI) in managing and strengthening international relations (IR) AND conversely, the role of international relations in enhancing international higher education, research and innovation. The focus of this discussion is on the role of IHERI in managing and strengthening international relations.

Important to note is that Knight’s proposed working definition of knowledge diplomacy does not include or delineate the actors, rationales, strategies, and outcomes. These aspects are deliberately not included because they can differ dramatically according to the issues being addressed, the strategies being used, and the primary actors and countries involved. Furthermore, it must be recognized that this definition is not neutral as it infers a positive outcome. While the intention is to strengthen international relations as well as enhance international higher education, research and innovation, there can be unexpected negative outcomes depending on the issues and actors. This is the nature of international relations.

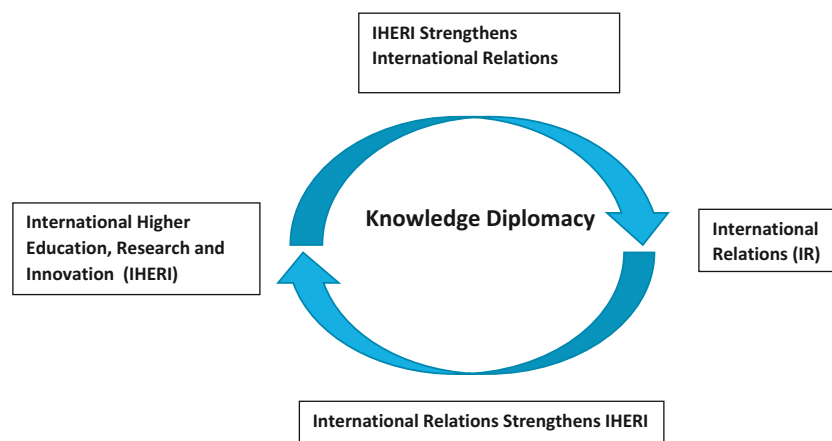


Fig. 1 Knowledge diplomacy as a two way process.

Why study knowledge diplomacy

Why study knowledge diplomacy is a complex question which begs a far more comprehensive analysis than this discussion can provide. However, there are a number of current issues and critical developments which necessitate a deeper understanding of the two way relationship between higher education, research and innovation and international relations. These include:

- The number of world issues, some call them crises, which cannot be solved by one nation alone. A bilateral or multilateral approach is imperative for addressing issues and threats such as climate change, epidemics, refugees and migration, water, food security, biodiversity
- The evolution of diplomacy such that contemporary diplomacy includes new actors, new strategies, new issue based diplomacies and new approaches
- The changing landscape of international higher education. In addition to the traditional elements of IHE such as student/scholar mobility, joint research projects and institutional agreements, there are new development such as knowledge hubs, international joint universities, regionalization of higher education and policies, and transformation of university networks into regional universities.
- The escalation of national public/private partnerships between universities and industry to the international and regional levels.
- The growing interdependence and interconnectivity among nations in spite of more recent threats of growing nationalism
- The increasing importance and emphasis on the knowledge society as a broader concept than knowledge economy
- The combination of soft power (attraction and persuasion) with hard power (military and economic sanctions) into smart power.
- The need for reliable research, verifiable evidence and knowledge sharing in the emerging “post-truth era.”

Differentiating knowledge diplomacy from cultural, science, education diplomacy

Misconceptions about knowledge diplomacy

Knowledge diplomacy presents a new approach to exploring the relationship between IHERI and international relations. As such the term is interpreted and used in a variety of ways as it becomes part of the higher education and international relations discourse. Knowledge diplomacy is being used interchangeably and confused with terms such as soft power, cultural and science diplomacy (Knight, 2015). There are instances where knowledge diplomacy has been inappropriately described as a way to lobby for more national and regional government funding for IHE. The term knowledge diplomacy has also been incorrectly used as a synonym for the commercialization of IHERI in the knowledge economy.

Important to understand is that knowledge diplomacy is a more comprehensive process than individual traditional higher education internationalization activities and that it is different from soft power and other forms of diplomacy such as cultural, science and education diplomacy. Understanding the difference among these different terms involves analyzing two aspects. The first is understanding inherent values attached to the different approaches and secondly analyzing the nature of the strategies and activities.

Knowledge diplomacy is broader than cultural diplomacy

Cultural diplomacy has been a popular term for decades. While the meaning and activities have evolved it is primarily oriented to international exchanges, exhibitions and events in all fields of the arts, music, theater, literature, film, media, architecture as well as sports and other cultural expressions (Goff, 2013). The goal of cultural diplomacy is primarily to enhance cross-cultural awareness, trust and relations between and among countries (Gienow-Hecht and Donfried, 2010). When higher education is referred to as part of cultural diplomacy the most common activities cited are student/scholar exchanges; language learning, international sports/debating/competitions; and cultural events (Pajtinka, 2015). While cultural diplomacy can include a wide range of people-to-people education and cultural exchanges, it typically is not seen as being broad enough to include all the central elements of higher education particularly those involved in scientific research and subsequent innovations. Furthermore, the emphasis is primarily on people mobility and the recent increase in the scope and scale of universities and other higher education providers moving across borders to offer foreign education programs and qualifications in the students’ home country has to date, not been accommodated in the notion of cultural diplomacy (Knight, 2016a).

Knowledge diplomacy is inclusive of science diplomacy but more comprehensive

The increasing importance of science diplomacy as evidenced in both national government science policies and international summits begs the question as to whether science and knowledge diplomacy are not one in the same. This is a question worthy of consideration and depends on how broadly science is being defined and used. If science is broadly interpreted to mean knowledge then there is a close relationship. But, traditionally science diplomacy has been seen in terms of hard sciences and more recently it has been placed within the broader framework of science, technology and innovation (Miremadi, 2016). There is no doubt that this reflects the centrality of science and technology in today’s knowledge economy. However, the focus on science and technology excludes, to a large extent, other sectors, issues and disciplines related to the social sciences and humanities. For instance, it is highly unlikely that science diplomacy initiatives or negotiations would include issues such as refugee, gender, poverty or human rights initiatives. Thus, while full acknowledgment is given to the importance and role of science diplomacy it does not rule out the necessity of the process of knowledge diplomacy which is a more inclusive concept in terms of education and the production and application of knowledge.

Why not use the term education diplomacy?

The term education diplomacy is usually applied to basic education and is linked closely to advocacy. The Association for Childhood Education International has adopted this term and believes that “education diplomacy uses the skills of diplomacy grounded in human rights principles to advance education as a driver for human development” (Hone, 2016). This raises the question as to whether the term education diplomacy or knowledge diplomacy is more appropriate for higher education. It is a question worthy of consideration. In addition to education and training, knowledge diplomacy includes research and the use of research and new knowledge for innovation, two areas not usually associated with basic education. Furthermore, the drivers and outcomes differ. Education diplomacy is oriented to human development while knowledge diplomacy has a broader mandate and in many cases focuses on addressing and solving common global issues which face countries in all regions of the world.

Knowledge diplomacy is different from soft power

In the mid-nineties the term “soft power” was introduced as a distinct and viable alternative to hard power. Soft power was described as “the ability to influence others and achieve national self-interests through attraction and persuasion” (Nye, 2004). It is different than hard power which uses coercion through military force and economic sanctions to achieve national interests. The concept of smart power emerged later (Nye, 2009) and is a calibrated combination of soft and hard power strategies to achieve a country’s goals. The latest term in the ever-changing lexicon is sharp power. This term was first introduced by Walker and Ludwig (2017) as the use of manipulative diplomatic policies by one country to influence the political system of a target country. A fuller description of sharp power suggests that it uses aggressive and subversive policies by authoritarian governments as a projection of state interests in democratic countries (Walker, 2018).

Fascination with the term soft power resulted in its application to many sectors and types of bilateral/multilateral relationships. Higher education was no exception. The debate and articles about higher education as an instrument of soft power sky rocketed in the early 2000s-especially from Asia (Li, 2018). Even long standing collaborative academic programs, capacity building and development projects were framed as soft power initiatives. International higher education was solidly positioned as part of the competitiveness and power agenda (Wojciuk, 2018). The popularized term of soft power was used to inappropriately describe higher education partnerships as being about competitiveness, dominance and a power dynamic (Nye, 2005) while in reality many were rooted in the notion of cooperation, reciprocity and mutuality of benefits.

In many ways, the term knowledge diplomacy emerged as a counter balance or alternative to soft power. Important to point out is that while some international higher education initiatives may be based on competitiveness and dominance and are correctly labeled as soft power (Chou and Spangler, 2018); not all are. It is imperative that the difference between the role of IHERI in a knowledge diplomacy approach to international relations is distinguished between the role of IHERI in a soft power approach.

A soft power approach juxtaposed to a diplomatic approach

While both soft power and knowledge diplomacy initiatives exist in higher education they differ significantly according to factors such as intentions, nature of relationships, strategies and underlying values. In short there is a difference between a soft power approach and a diplomatic approach as illustrated in Table 1. This section discusses the main attributes of soft power and diplomatic frameworks. Hard power, which is predicated on the use of military force or sanctions is not usually used in higher education although there are major implications for higher education when it is applied to a conflict situation between countries.

Table 1 A Soft Power Approach juxtaposed to a Diplomatic Approach.

	<i>Soft power approach</i>	<i>Diplomatic approach</i>
Self-interest	National self-interests dominate at the expense of others	National self-interests leading to mutual benefits
Modes	Attraction Persuasion Compliance Domination	Negotiation Mediation Communication Conflict resolution Conciliation Collaboration
Values	Authoritarianism Competition Cooption	Reciprocity Mutuality Common ground
Outcomes	Win -lose Zero-sum game	Win-win with mutual but different benefits
Nature of relationships	Vertical Top down	Horizontal Collaborative

- *Approach* - The common motivations behind soft power include self-interest, influence and dominance and thus a vertical or top-down approach is used to achieve desired outcomes. With diplomacy the intentions are closely linked to addressing self-interests and common issues in a collaborative fashion ensuring that there are benefits, albeit different benefits, for all partners. National self-interests are still at play, it would be naïve to deny this, but a more horizontal type relationship exists among partners/actors in diplomacy.
- *Strategies*- in a soft power approach, attraction and persuasion leading to compliance are the primary strategies used whereas diplomatic strategies are usually characterized as negotiation, communication, conciliation, cooperation and mediation thus supporting the collaborative horizontal approach.
- *Values*- the values underpinning a soft power paradigm are domination, authoritarianism, and competition to achieve national goals and self-interests. These differ significantly from diplomacy which builds on values of reciprocity, mutuality, and compromise to meet the different needs and interests of partners toward reaching a common end while still recognizing the centrality of national self-interests which differ among countries.
- *Outcomes*- in terms of outcomes the soft power approach is often characterized as one-sided in terms of benefits as opposed to a diplomatic approach which tries to achieve a win-win outcome with mutual but different benefits for all major players.

Thus, a soft power approach versus a diplomatic approach in international higher education, research and innovative is based on fundamentally different motives, modes, values, and outcomes. In the next section knowledge diplomacy is analyzed using the diplomatic framework. The characteristics or key elements of knowledge diplomacy are discussed at both a conceptual and pragmatic level.

Key characteristics of knowledge diplomacy

An important aspect of defining and understanding knowledge diplomacy is the identification of the fundamental or common characteristics. These include guiding values or principles, different ways of interacting, types of relationships and activities among the range of actors, and the spectrum of benefits and outcomes. The following list of eight characteristics identifies important and strategic dimensions of knowledge diplomacy and aligns closely with the diplomatic framework described above.

Diversity of Actors and Partners

In knowledge diplomacy a diversity of actors is actively engaged in collaborative research, knowledge production, training, and innovation projects. While universities and colleges are key players there is usually a range of other higher education actors involved. These include national, regional or international centers of excellence, research institutions, private and public foundations, think tanks, professional associations, non-governmental organization related to education, and governmental departments/agencies at all levels. In the majority of cases the higher education actors are also working with other sectors and/or disciplines depending on the nature of the initiative. Common partners include industry, civil society groups, foundations, governmental agencies. Higher education actors working together with other sector partners is a key feature of knowledge diplomacy and often differentiates knowledge diplomacy from higher education internationalization.

Focus on Higher education, research and innovation

Knowledge diplomacy builds on the fundamental functions of higher education (teaching/learning, research, knowledge production and innovation, and service to society). Normally, the process of knowledge diplomacy involves multiple forms of IHERI as dictated by the nature and complexity of the issue being addressed. Individual IHE activities (i.e., student mobility, scholar exchange, joint conference) are part of the larger networked series of activities involving multiple actors and strategies but as stand-alone activities they normally do not constitute knowledge diplomacy. As stand alone activities they are more accurately described as internationalization strategies and while they have multiple benefits they are not specifically designed to contribute to sustainable international relations between countries although that can be a by-product.

Recognition of different needs and collective use of resources of actors

Because knowledge diplomacy brings together a network of different partners from various sectors to address common global issues there will be different rationales and implications for the individual countries and actors involved. This means that actors will bring different needs, priorities and resources to the partnership. These differences will need to be respected and negotiated to ensure that the strengths and opportunities for each partner are optimized. This is done through a horizontal collaborative type of relationship that acknowledges the different but collective needs and resources of the group of partners. Leadership is critical but not in the form of dominance, authoritarianism or compliance as in a soft power relationship.

Reciprocity -mutual but different benefits

Different needs and resources of actors will result in different benefits (and potential risks) for partners. Mutuality of benefits does not mean that all actors/countries will receive the same benefits. It does mean, however that the principle of mutuality and reciprocity of benefits will guide the process and there will be collective and different benefits accrued for different actors and countries.

Based on negotiation, collaboration, cooperation

Knowledge diplomacy is based on horizontal relationships between and among major actors and countries and focuses on cooperation, collaboration, negotiation and compromise to ensure that the overall goals are met and there are benefits for all. This is founded on a win-win approach which is fundamental to knowledge diplomacy and as noted in table one differs from soft power.

Different levels of collaboration

Knowledge diplomacy includes different levels of bilateral and multi-lateral cooperation at regional and international levels.

Commonality of issues – addressing regional and global issues

It is acknowledged that there may be varied motivations driving knowledge diplomacy. However, addressing common issues, as well as specific national interests, requires cooperation among higher education actors and other partners from different countries. This necessitates an agenda which highlights common issues and concerns among countries who believe that a collaborative and co-ordinated effort will lead to results that could not be achieved by one country alone. As discussed, while the issues of concern will be common the different needs, priorities, resources, benefits and implication may vary among actors and countries. Succinctly put, it means that the “whole is greater than the sum of the parts” but that each part brings different strengths and may receive different benefits.

Build and strengthen relations between and among countries

Central to the notion of knowledge diplomacy is that the cooperation among the different actors and partners depends on and further strengthens positive and productive relations between and among countries. This builds on but goes beyond the contribution that bilateral and multi-lateral agreements between higher education institutions make. Clearly, there is a sliding scale with regards to the breadth and depth of contributions that knowledge diplomacy can make to relations between and among countries but working to address pressing global issues that impact each and every country is an important way forward.

Important to emphasize is that these knowledge diplomacy characteristics can overlap with different types of internationalization activities, but it is a mistake to suggest that the processes of internationalization and knowledge diplomacy are the same. Knowledge diplomacy is clearly oriented to strengthening relations between and among countries and uses IHERI activities as a means to this end. However, a commonly accepted definition of higher education internationalization “the process of integrating international, intercultural and global perspectives into the primary functions (teaching/learning, research and service to society) and delivery of higher education institutions and systems” (Knight, 2016b) has a narrower and more academic focus.

Examples of knowledge diplomacy initiatives

The collaborative knowledge diplomacy approach is being explored as an alternative to a more one-sided soft power approach. It is interesting and helpful to look at some contemporary projects that exhibit key characteristics of knowledge diplomacy. These examples of knowledge diplomacy have been chosen with the following factors in mind:

- come from all regions of the world,
- include bilateral and multilateral initiatives,
- address a variety of global issues,
- involve HEIs as key participants among a broader group of actors and sectors
- represent long term successful projects and newer initiatives,
- include all aspects of higher education in terms of teaching/learning, research and innovation, and service to society.
- contribute to the building and strengthening of relations between and among countries of the world

These initiatives were carefully chosen to illustrate that knowledge diplomacy includes but goes far beyond typical international higher education activities such as student/scholar mobility, joint research projects and conferences.

The Pan African University. The African Union in cooperation with other African and international organizations has created five different research institutes each dedicated to research and teaching in a specific set of disciplines and are located in five different regions of the continent. The research institutes are: (1) Basic Sciences, Technology and Innovation located at Jomo Kenyatta University of Agriculture and Technology in Kenya; (2) Life and Earth Sciences, including Health and Agriculture, located at University of

Ibadan Nigeria; (3) Governance, Humanities and Social Sciences located at University of Yaoundé II in Cameroon; (4) Water and Energy Sciences located at the Abou Bakr University of Tlemcen in Algeria; (5) Space Sciences Institute located at Cape Peninsula University of Technology in South Africa which is still under development. Each institute provides graduate programs and serves as a network hub for research collaboration with other universities and research partners (industry, government, NGOs) in the region. It is a multi-lateral knowledge diplomacy initiative where IHERI serves to strengthen African regionalization and where collaboration among countries and actors in the continent serves to advance IHERI. <https://pau-au.africa/>

Humanitarian Relief Initiatives. Brown University from the USA works bilaterally with countries in the Global South to promote more research and training on humanitarian aid. It is at the cutting edge of working in local settings to create and distribute new knowledge to serve better policy making and management practices in disaster management at both the community and national levels. This knowledge diplomacy initiative addresses a critical global issue by working collaboratively with local and national universities, governments, and industry in host countries while training young academics both at Brown and beyond to undertake research with multiple disciplines and sectors on this critical issue. <https://watson.brown.edu/biari/about/history>.

Women and Gender Research Institute. Located at Granada University in Spain but closely networked with other universities, research centers and NGOs across Europe and beyond, this research institute and related graduate programs put women and gender issues at the center of interdisciplinary research, policy development, and training of new scholars. It is a self-funded multi-lateral knowledge diplomacy project which has developed over a couple of decades and demonstrates how cooperating with other higher education and NGO partners nationally, regionally and internationally serves to advance research, knowledge exchange and advocacy on this critical issue. In addition, it demonstrates how IHERI is effective in building closer relationships within Europe to advance the regionalization agenda. https://www.ugr.university/pages/research_innovation_transfer/research_institutes_centres/the_institute_of_womens_and_gender_studies.

International Joint Universities (IJU) - the German-Jordanian University. The establishment of internationally co-founded and developed joint universities by two or more institutions/countries is a fascinating development which contributes to strengthening IHERI as well as bilateral relations between countries. As of 2021, there are 22 operating around the world with several others in the planning stage. The German-Jordanian University, established in 2005, exemplifies how the strategic use of joint academic programs and collaborative research between academics and industry in both countries yields benefits for all, and how student and scholarly exchanges builds closer cultural and scientific relationships between the two countries. <http://www.gju.edu.jo/>

ZIKA Rapid Response Project. The outbreak of the Zika virus in Brazil and the threat of it spreading to other countries resulted in a collaborative research project that originally involved researchers from Brazil and the UK but soon included universities and research centers from around the world. The urgency for research on treatment and prevention of Zika resulted in an abbreviated time period between the application and awarding of the funding by three funding agencies in Britain and a foundation in Brazil. While this knowledge diplomacy project is primarily oriented to research it illustrates the importance of collaboration between expert researchers and the advantages of involving different sectors and disciplines to undertake the research and share the new knowledge. <https://webarchive.nationalarchives.gov.uk/20200923143229/https://mrc.ukri.org/funding/browse/zika-rapid-response-initiative-parent/zika-rapid-response-initiative/>.

Australia-India Research Fund: Since 2006, the Australia-India Strategic Research Fund (AISRF) has strengthened relations between Australia and India by supporting scientific research that are jointly undertaken by researchers from both the public and private sectors in both Australia and India. Studies must address mutual priority areas for the two nations such as: agriculture, astronomy and astrophysics, biomedical devices and implants, clean energy technologies, food and water security, information and communication technology, marine sciences, nanotechnology, stem cells and vaccines. The AISRF grants can be used for research, conferences, workshops and fellowships for early-career researchers thus focusing on several the key strategies of knowledge diplomacy. <https://www.industry.gov.au/funding-and-incentives/collaborating-with-india-on-science-and-research>.

Sustainable Development Solutions Network (SDSN)- Climate Change: This successful multilateral knowledge diplomacy initiative was established by the United Nations in 2012 to mobilize global scientific and technological expertise to promote practical solutions for sustainable development especially in the area of climate change. SDSN work is guided at the global level by a group of internationally recognized experts. Universities play a major role in research and innovation projects in 32 countries involving researchers from across different disciplines and sectors as well as the theme based groups. In addition, a Universities Partnership Network develops curriculum based on the new knowledge developed through research. All SDSN projects aim to close the gap between emerging research and the necessary policy development to effect change especially regarding climate crisis. <http://unsdsn.org/about-us/vision-and-organization/>.

RENKEI- Research and Education Network for Knowledge Economy Initiatives –Japan and UK. RENKEI which means collaboration in Japanese is a university research network and knowledge diplomacy initiative between Japanese and UK universities and the private sector. RENKEI was founded in 2012 with the goal of strengthening relationships between the two nations by developing academic-industry research collaborations that would address major societal issues. The network includes six universities in Japan and eight universities in the UK, as well as dozens of research partners from industry, business and civil society. Between 2012 and 2018, RENKEI's multi-actor working groups organized joint research projects, workshop, and conferences which addressed pertinent issues such as Sustainable Energy, War, Slavery, Aerospace Engineering, Renaissance Entrepreneurship, and Living with an Aging Society. <https://www.britishcouncil.jp/en/programmes/higher-education/university-industry-partnership/renkei>.

This selection of examples individually and collectively illustrates the key characteristics of knowledge diplomacy. They come from different regions of the world, represent partnerships between higher education actors and partners from other sectors and disciplines, are based on reciprocity and mutuality of benefits, and address a variety of global issues.

Issues and challenges

Knowledge diplomacy is not without its challenges. First is the issue of values. Values play a central role in diplomacy and explain why the contribution of international higher education and research to international relations and vice versa is conceptualized in a diplomatic framework and not a soft power paradigm. Knowledge diplomacy recognizes the diversity of priorities and resources among countries and that interests and benefits will differ among partners. However, there is the reality and risk that knowledge itself can be used as an instrument of power to enhance self-interest, competitiveness and dominance by one country. This is why values and principles are important and it is critical to distinguish between a knowledge diplomacy approach and a soft power approach. Unintended consequences are always present. While foresight can help mitigate risks, it is only hindsight that tells the story of impact. The values of collaboration and mutuality which underpin knowledge diplomacy can be easily eroded. There is the potential risk that education, research, and innovation will be used to widen the knowledge divide among countries instead of being a bridge to address global challenges through collaboration, exchange and trust. Knowledge diplomacy can easily become a buzzword to camouflage national and regional ambitions to promote self-interests at the expense of mutual interests and benefits. As the concept of knowledge diplomacy becomes more common place, unrealistic expectations can be made about its role and contributions. Knowledge diplomacy is not a silver bullet. Expectations of its contribution to international relations need to be managed to avoid early misunderstandings or dismissal of its value and potential.

There are many unanswered questions about the concept of knowledge diplomacy. Will universities and individual scholars take a collaborative and mutual benefit approach when competitiveness is becoming more and prominent in higher education? Will politicians appreciate knowledge diplomacy as an international relation's instrument that can advance the interests of some nations without limiting the prospects of others? Can knowledge diplomacy be operationalized in light of competing priorities within and between countries/regions? Can the contribution and impact of knowledge diplomacy be measured? Is it feasible to develop mechanisms where education, research and innovation complement each other to achieve goals that each could not accomplish as stand alone activities? Will knowledge diplomacy be seen as a two-way process whereby strong relations between and among countries will help to strengthen higher education and research? These are but a few of the questions that need to be explored.

Developing a framework, strategies and commitment to knowledge diplomacy cannot be done without facing the harsh realities of international politics and the challenges of the more competitive and turbulent world in which we live. However, the question must be asked whether we can afford to ignore the potential of knowledge diplomacy to address and contribute to the resolution of current national, regional and global challenges. Further reflection, analysis and research is called for to better understand and further the contribution that international higher education, research and innovation makes to strengthening relations between and among countries and to addressing issues that know no borders.

N.B. This paper has been excerpted with permission from the British Council publication by Knight, J. (2019) *Knowledge Diplomacy in Action*. Discussion Paper. British Council, London, United Kingdom

References

- Bjola, C., Kornprobst, M., 2013. *Understanding International Diplomacy: Theory, Practice, and Ethics*. Routledge, Oxon, UK.
- Chou, C.P., Spangler, J., 2018. *Cultural and Education Exchanges between Rival Societies- Cooperation and Competition in an Interdependent World*. Springer Nature, Singapore.
- Cooper, A., Heine, J., Thakur, R., 2013. *The Oxford Handbook of Modern Diplomacy*. Oxford University Press, Oxford, UK.
- Gienow-Hecht, J., Donfried, M., 2010. *Searching for a Cultural Diplomacy*. Berghahn Books, New York, USA.
- Goff, P., 2013. *Cultural diplomacy*. In: Cooper, A., Heine, J., Thakur, R. (Eds.), *The Oxford Handbook of Modern Diplomacy*. Oxford University Press, Oxford, UK.
- Hone, K., 2016. *Education Diplomacy: Negotiating and Implementing the Sustainable Development Goals- Looking Back and Looking Ahead*. Centre for Education Diplomacy and Leadership, Childhood Education International, USA.
- Kerr, P., Wiseman, G. (Eds.), 2013. *Diplomacy in a Globalizing World: Theories and Practices*. Oxford University Press, New York.
- Knight, J., 2022. *Knowledge Diplomacy in International Relations and Higher Education*. Springer Nature.
- Knight, J., 2016a. *Transnational Education Remodeled: Towards a common TNE framework and definitions* in. *J. Stud. Int. Educ.* 20 (1), 34–47.
- Knight, J., 2016b. *Meaning, rationales and tensions in internationalization of higher education*. In: McGrath, S., Gu, Q. (Eds.), *Routledge Handbook on International Education and Development*. Taylor Francis, pp. 325–339.
- Knight, J., 2015. *The Potential of Knowledge Diplomacy: Higher Education and International Relations*. The European Association for International Education, Amsterdam, The Netherlands.
- Li, J., 2018. *Conceptualizing Soft Power of Higher Education: Globalization and Universities in China and the World*. Springer Nature Publishers.
- Miremadi, T., 2016. *A model for Science and Technology Diplomacy: how to align rationales of foreign policy and research*. SSRN Electron. J.
- Nye, J.S., 2009. *Get smart: combining hard and soft power*. *Foreign Aff.* 88, 160.
- Nye, J.S., 2005. *Higher Education and Soft Power*. *Forum Futures* 2005, 11–14. Educause, Cambridge MA.
- Nye, J.S., 2004. *Soft Power: The Means to Success in World Politics*. Public Affairs, New York, New York.
- Pajtinka, E., 2015. *Cultural Diplomacy in the Theory and Practice of Contemporary International Relations*. Politické Vedy.
- Pigman, G., 2010. *Contemporary Diplomacy. Representation and Communication in a Globalized World*. Polity Press, Cambridge, England.
- Ryan, M., 1998. *Knowledge Diplomacy: Global Competition and the Politics of Intellectual Property*. Brookings Institution Press, Washington, D.C., USA.
- Walker, C., 2018. *What is "sharp power"?* *J. Democr.* 29, 9–23.
- Walker, C., Ludwig, J., 2017. *Sharp Power: Rising Authoritarian Influence*. National Endowment for Democracy, Washington, DC, USA.
- Wojciuk, A., 2018. *Empires of Knowledge in International Relations. Education and Science as Sources of Power for the State*. Routledge, New York, USA.

North–South partnerships in higher education

Samia Chasi, International Education Association of South Africa (IEASA), Pretoria, South Africa and University of the Free State, Bloemfontein, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	210
International higher education partnerships	210
Global North–Global South	211
Internationalization and transformation in South African higher education	211
Colonization: historical North–South links in higher education	212
Recolonization: continued Northern domination	213
Imbalances in North–South partnerships	213
Decoloniality	214
A southern critique of internationalization	214
Internationalization as recolonization	215
Decolonization: toward a new paradigm of North–South partnerships	216
Conclusions	217
References	218

Introduction

This article addresses North–South partnerships in the context of higher education internationalization from the perspective of the Global South. It reflects on the past, present and future of North–South engagements in higher education with reference to the notions of colonization, recolonization and decolonization. In doing so, a decolonial lens is used to gain a deeper understanding of the challenging dynamics of North–South partnerships, particularly regarding continued dependencies and inequalities, and to challenge existing partnerships practices in the broader context of higher education internationalization. At the same time, a decolonial approach serves as a tool to conceptualize and empower positive change by imagining North–South partnerships differently. Before exploring this in more detail, an introductory discussion briefly touches on international higher education partnerships, the meaning of *North–South* and the relevance of the South African higher education sector in this context.

International higher education partnerships

In the 21st century, partnerships are a key element of internationalization activities of universities around the globe, as “institutions that want to be international players have to work with partners, be that in education, in government, in the NGO sector or in the private sector” (Foskett, 2010, p. 139). Playing a central role in higher education internationalization, “the rationale for developing key, strategic, international-education alliances at both the national and institutional level is not so much an end unto itself but a means to achieving academic, scientific, economic, technological, or cultural objectives” (Knight, 2004, p. 27). Partnerships are therefore important for all core functions of universities, including teaching and learning, research and community engagement. One of the most common models of international partnerships are exchanges established to facilitate student and staff mobility (Sutton et al., 2012).

Partnerships for mobility and other purposes such as research collaboration have become a crucial characteristic of internationalization. They are increasingly used as internationalization vehicles (Sutton et al., 2012). For example, strategic international partnerships are a popular and growing tool used by universities to enhance their international reputation (Ford and McMullin, 2016; Maringe and De Wit, 2016). Strategic university alliances mainly aim to strengthen institutional capacity by sharing resources and facilities as well as to support research across disciplines and national borders, especially in response to global challenges (Ford and McMullin, 2016). Global research partnerships facilitate university collaboration to find mutually beneficial and sustainable solutions to the challenges faced by humanity at large (Rensburg, 2016). With no region in the world able to “fully develop its people, let alone its knowledge, without collaboration” (Rensburg, 2016, p. 315), partnerships are also pursued in the context of development and capacity building, particularly in the North–South context. North–South partnerships emerged during the 1990s as the prevailing model for international development assistance (Koehn and Obamba, 2012). At the heart of development cooperation between partners in the Global North and the Global South often lies an assumption that also characterizes industry–university partnerships, which entails that “one partner has the right answers and the other must adopt their way of working” (Foskett, 2010, p. 141). Before exploring the underlying dynamics of North–South partnerships further, it is useful to clarify the meaning of *North–South*.

Global North–Global South

The compound *North–South* is often used in geographical or economic terms. In case of the former, it indicates a geographical divide in terms of countries that are located to the north and south of the equator. In case of the latter, it refers to a wealth divide, with rich nations located in the North and poor nations located in the South (Maringe and De Wit, 2016). While geographic location and economic status can coincide, there is no direct correlation between the two.

In the context of this article, the Global South comprises countries that “are often poor and tend to have a previous history of colonial occupation” (Maringe and De Wit, 2016, p. 304). North–South has a broader meaning in this regard, as it “bespeaks a relation, not a thing in or for itself” (Comaroff and Comaroff, 2012, p. 127). It is used to denote countries that are “differentiated in terms of power, wealth, and especially in relation to previous history of colonial occupation and domination” (Maringe and De Wit, 2016, p. 304). In this sense, it highlights a relational aspect, where the terms *North* and *South* are used to “emphasize relations—authority, exclusion and inclusion, hegemony, partnerships, sponsorship, appropriation—between intellectuals and institutions in the metropole and those in the world periphery” (Connell, 2007, pp. viii–ix). North–South partnerships are understood to be partnerships between higher education institutions in the Global North and their Southern counterparts. In the area of research, for example, North–South partnerships are shaped by colonial legacies in terms of research agendas and collaboration patterns (Boshoff, 2010). Such partnerships are, with reference to the African continent, sometimes also referred to as historical (Teferra, 2012) or traditional (Jowi et al., 2013).

It is important to point out that the notion of North–South is dynamic and can be expanded in a number of ways. *North* does not only refer to the former colonizing nations in Europe but also to the United States of America (USA). Referring to research collaborations involving South African partners, for example, it is noted that colonial linkages could explain why 29 per cent of the country’s co-publications are with partners in the United Kingdom (UK). However, the USA ranks first with almost 32 per cent of all South African co-publications (Boshoff, 2010, p. 486), which underscores the global dominance of the USA and the imperial relationship it has with many other countries in the world. These overall trends are, more recently, confirmed for health sciences research of a South African university, where the USA and the UK have remained the top two international partner countries for co-publications between 1999 and 2015 (De Jager et al., 2020).

Furthermore, North–South relations can exist outside direct colonial or imperial links. In the case of Irish–South African school linkages, for example, “the colonial patterns that lie at the heart of DE [development education] in the UK are replaced in Ireland by a pattern established by Irish Catholic missions” (Gallwey and Wilgus, 2014, p. 524). In this case, dominance is exerted on religious grounds. Further complexity is added when the notion of North–South is used even more broadly as a metaphor for inequality, linked to the above mentioned economic divide between the rich and the poor. It is then possible to point to variations within the North and the South, speaking of “north of the south, south of the north: Detroit as south in the US, Johannesburg as north in Africa, and so on” (Menon, 2018, p. 37). In this sense, Northern partners can be “South-like”, and vice versa.

With all this in mind, the perspective of this article is distinctly Southern. It considers the idea of the Global South an “invitation to imagine the world afresh” (Menon, 2018, p. 40). It speaks from “an ‘ex-centric’ location” and from a position that is “an outside to Euro-America” (Comaroff and Comaroff, 2012, p. 127). This vantage point is illustrated by South Africa, a post-colonial and post-apartheid society in transformation. The country is used as an example and to provide a textured understanding of relevant issues. South Africa is geographically South because it is located south of the equator, on the southernmost tip of the African continent. Economically, the country is typically referred to as an emerging economy. However, while South Africans have become less poor over the course of two decades since the country became a parliamentary democracy, inequality has been on the rise (Gaventa and Runciman, 2016). In relational terms, South Africa is a representative of the Global South due to its shared colonial history with two European countries. In addition, black South Africans endured further domination under apartheid, a system of institutionalized racial segregation enforced between 1948 and 1994, “tied up with white supremacy, capitalist exploitation, and deliberate oppression” (Biko, 1987, p. 28). Apartheid has been described as a special type of colonialism, where the colonizer (white people) and the colonized (non-White people) were not geographically distant but lived within the same borders (Nkwinti, 2016). Non-white South Africans were discriminated against in all spheres of life, including higher education. The legacies of colonialism and apartheid are still felt in the South African higher education sector today and need to be taken into account in any discussion of internationalization.

Internationalization and transformation in South African higher education

For South African higher education, the importance of internationalization has grown since apartheid came to an end in 1994, which is a “reflection of globalization as well as of South Africa’s return to the international community” (DHET, 2013, p. 39). Internationalization is pursued by institutions operating in a sector that is generally underfunded, with government subsidies for public universities declining in real terms. Additional challenges arise due to the sector’s differentiated nature. As a legacy of colonialism and apartheid, inequality remains a key characteristic of higher education in South Africa. Historically black universities were established and maintained to train black people, serving the needs of the colony as well as the apartheid state (Heleta, 2016). The distinction between historically advantaged institutions (HAIs) and historically disadvantaged institutions (HDIs) reflects the lingering influence of colonialism and apartheid, which saw different institutions receive different kinds of status and treatment, along racial lines. Such differentiation mirrors a North–South divide in terms of inequality, with HAIs symbolically assuming a Northern position. To this day, South African universities “continue to ail under the crippling influence of the legacy of apartheid,

with statistics of performance, participation, persistence, dropping out, and quality of educational outcomes still quite traceable along racial lines” (Wits, 2014, p. 10). HAs generally fair better than HDIs in terms of institutional infrastructure, resources, capacity and performance.

Against this backdrop, inequality is not only a characteristic of the higher education system in general but also impacts internationalization, which has not been advanced evenly across the country’s universities (IEASA, 2020). South Africa’s first national internationalization policy framework, whose adoption was formally announced in November 2020, recognizes existing inequalities in terms of internationalization benefits, where HDIs are especially disadvantaged due to low levels of international relations (DHET, 2019). Universities’ internationalization efforts should therefore “be permeated by an agenda that focuses on inclusion and social justice” (IEASA, 2020, p. 4). If applied uncritically, within dominant paradigms, internationalization can further entrench existing inequalities. To help mitigate this potential, internationalization in South Africa cannot be a goal in itself, but needs to be aligned to relevant institutional and national objectives, contributing to transformation, nation building, and the public good.

Addressing and overcoming past inequalities is a key concern of South Africa’s transformation agenda, and transformation goals permeate government policy in higher education as much as in other spheres. The internationalization policy framework developed by the Department of Higher Education and Training (DHET) recommends that universities design their internationalization activities in such a way that first priority is given to the needs and interests of South Africa. After that, “where possible and relevant, the following order of priority focus should be observed in terms of interests: the SADC states; the rest of the African continent; BRICS; the Global South and emerging economies; and the world beyond” (DHET, 2019, p. 22). Such a priority focus is an example of moving the “center”, which is traditionally understood as a synonym for Euro-America or the West. Centering South Africa and Africa signals a shift away from colonialism, which placed Europe at the center of attention and engagement, where Europe was considered the norm or “the gold standard against which the rest of the world was positioned” (Martin and Griffiths, 2012, p. 910). Before exploring in more detail how transformation in today’s South African higher education sector is linked to processes of decolonization, it is important to first briefly outline the history of higher education linkages between South Africa and Europe.

Colonization: historical North–South links in higher education

The idea of higher education and universities in Africa was not introduced by Europeans. Their origins can be traced to three distinct traditions involving the Alexandria Museum and Library, early Christian monasteries as well as Islamic mosques, going back as far as the third century BC as well as the third and first centuries AD, respectively (Zeze, 2006). However, “While Africa can claim an ancient academic tradition, the fact is that traditional centers of higher learning in Africa have all but disappeared or were destroyed by colonialism” (Teferra and Altbach, 2004, p. 23). In addition, universities on the continent were, with only one possible exception of Ethiopia, “shaped by colonialism and organized according to the European model” (Teferra and Altbach, 2004, p. 23), with most “Western” style institutions being established in the 19th century.

While colonialism was effected through political and economic power, “it was also accomplished through cultural subjugation—for instance, through control of the education system” (Wa Thiong’o, 2009, p. 108). Despite differences in the European models, colonial higher education policies of colonizing countries had some elements in common, including limited access, academic freedom and autonomy, the use of colonial languages as well as limited curricula and academic offerings. These features are linked to the main purpose of education in the colonies, which was to support the administrative systems of colonialism (Teferra and Altbach, 2004).

The current South African higher education system traces its origins to the 19th century, and its development has been shaped by European models from the start (Sehoolo, 2006). The origins of some of the country’s oldest and most renowned universities such as the University of Cape Town and Stellenbosch University lie in colleges that were initially set up in 1829 and 1866 to cater for the needs of English and Afrikaner settlers, respectively (Zeze, 2006). Apart from being rooted in a history of British and Dutch colonialism, which saw the setting up of institutions for Europeans in South Africa and modeling them on European examples, the South African higher education system has also been impacted by apartheid. To state an earlier point differently, “South Africa became the last colony to decolonize in 1994 from an internal form of white colonialism called apartheid” (Ndlovu-Gatsheni, 2013, p. 13).

After gaining independence, African countries have generally maintained very strong ties to the countries that represent their former colonizers. In this context, the “impact of the colonial past and of the continuing impact of the former colonial powers remains crucial in any analysis of African higher education” (Teferra and Altbach, 2004, p. 24). Higher education institutions in Africa continue in the way they were set up, following colonial examples. Their programs and curricula continue to replicate colonial European models, and European languages continue to be used as the main medium of instruction (Singh, 2010). European languages, and particularly English, are not only the dominant languages of instruction at higher education institutions in Africa, but they also dominate scholarship:

African universities rely on the knowledge system that has been conceived, developed, and organized based on Western languages. The Western world produces the majority of knowledge conveyed in those languages. African universities do not have the capacity to generate enough knowledge of their own, nor do they have the capacity and infrastructure to process and translate existing ones virtually from the Western world—yet.

Teferra and Altbach (2004, p. 46).

Regarding partnerships, African universities have traditionally partnered with institutions in the Global North (Jowi et al., 2013), especially in European countries or in the USA, and such partnerships have expanded in the first decades of the 21st century (Obamba et al., 2013). The reasons why Southern institutions pursue partnerships with Northern partners are manifold. They are often framed in the context of development. As has been noted in that regard, “There seems to be a cruel irony in the inverse relationship between the size of the development challenges that nations face and the capacity of their university systems to rise to meet them” (Ramphele, 2004, p. 17). Institutions enter into partnerships for reasons relating to promoting innovation, sustainability, competitiveness, capacity building and quality enhancement (Anderson and Maharaso, 2002). Partnerships are also expected to help address challenges such as brain drain, inadequate numbers of academic staff with PhDs, inadequate curricula and poor infrastructure (Fischer and Lindow, 2008). In light of dwindling resources, partnerships help Southern institutions gain access to much needed funds (Malaza, 2011).

African institutions have a long tradition of partnering with their counterparts in the Global North and continue to do so for many reasons. However, North–South partnerships are often understood as problematic, particularly from the perspective of the Global South, as they tend to be imbalanced and disproportionately benefit the Northern partners. In the sections that follow, these imbalances will be discussed in more detail, with specific reference to resources, power and knowledge.

Recolonization: continued Northern domination

This section discusses the relevance of the notion of recolonization in the context of higher education internationalization. After providing a brief overview of the imbalanced nature of North–South partnerships, it explores how a decolonial approach can help understand the complexities and underlying power relations of such partnerships. It ends with a critique of internationalization from the perspective of the Global South, which frames North–South partnerships as an instrument of continued Northern domination and Southern dependencies and, therefore, as a form of recolonization.

Imbalances in North–South partnerships

Generally, “International partnerships can be complex, contested, and embedded with multiple forms of asymmetries” (Obamba et al., 2013, p. 172). Partnerships between institutions in Africa and in developed countries are particularly challenging in that regard, as they are often characterized by inequalities and a dominance of the Global North (Jowi et al., 2013). The imbalanced nature of partnerships involving African universities and their Euro-American counterparts is “due primarily to the gross and sometimes obscene disparity in the financial resources of the partners” (Hagenmeier et al., 2017, p. 87).

Resource imbalances are closely linked to power imbalances, which manifest in different ways. For example, a South African institution is described as having “minority-status” in terms of their engagements with Northern partners. In this dynamic, the Northern partners are perceived to be in a position where they can dictate the terms of the partnership, exerting pressure on the South African institution to conform (Tedrow and Mabokela, 2007). As far as leadership roles are concerned, the North is observed to be involved in more strategic and knowledge-based leadership, while the South’s participation is more at the operational level (Maringe and De Wit, 2016). In other words, the Northern partners are, generally speaking, in the driver’s seat, while the Southern partners assume the role of passengers coming “along for the ride” (Gallwey and Wilgus, 2014, p. 530). Regarding the question of who sets the tone in international research collaborations, the following is noted:

... for many decades the more affluent countries of the world (the political “North”) have been driving—or at least significantly influencing—the science and technology (S&T) and research agendas of poorer countries (the political “South”). The North’s impress on the S&T and research agendas of the South can be explained in terms of a number of factors, the most obvious being the North’s ability to unleash significant amounts of funding for global science development.

Boshoff (2010, p. 482).

North–South partnerships have been described as “highly extractive relationships”, which tend to disproportionately benefit the Northern partners (Maringe and De Wit, 2016, p. 308), especially in terms of facilitating talent migration. Increasing flows of people and finances from the Global South to the Global North is a key characteristic of the global higher education space (Maringe and Foskett, 2010). By recruiting the brightest talent, Northern institutions contribute to the continued depletion of human resources from other countries (Beck, 2013). Regarding knowledge production, the value of partnerships “is not evenly distributed between the partner in the North and the South, with knowledge outlets and authorship arrangements favoring the North” (Beck, 2013). Imbalances in reputation building as gauged, for example, by international university rankings, also favor Northern institutions. Furthermore, North–South partnerships are seen as “opportunities for rubber stamping the hegemony of the West, which exert its authority using Western knowledge and discourses, money and financial incentives to export solutions to an unsuspecting Global South” (Maringe and De Wit, 2016, p. 308). Given their imbalanced nature, North–South partnerships can be critiqued as replicating patterns of economic, geopolitical and other dependencies. Can such partnerships therefore be understood as a form of recolonization? To tackle this question, it is helpful to engage with the notions of coloniality and decoloniality as a point of departure.

Decoloniality

A decolonial position, as assumed in this article, has at its core an interest in exploring how colonialism has shaped the world we live in today. It is also interested in undoing such influence and in driving change. Regarding ways in which colonialism lingers on, participants in study visits between the UK and countries in the Global South have been observed as generally focusing on “inequality in the world *today*, without necessarily understanding the influence of the *past*, (that is, the former colonial relationship between the UK and Southern countries), which potentially denies an understanding of how that inequality came about” (Martin and Griffiths, 2012, p. 909). It is, therefore, imperative to “reconsider the colonial encounter and its continuing impact from the perspective of formerly colonized countries, regions and peoples, but within the context of contemporary globalization” (Crossley and Tikly, 2004, p. 148). Furthermore, “Such critiques will include debates about the effects of post-colonial and globalizing conditions, and the role of global institutions, foreign governments and development aid, seen against the relative levels of influence of African institutions in shaping policy choices” (Wits, 2010, p. 13).

Despite the formal end of different types of colonialism, a decolonial approach recognizes that underlying principles of colonialism still apply, especially in the context of relations between the former colonizers and the former colonized. In other words, “if colonialism is a way of maintaining an unequal international relation of economic and political power [...], then no doubt we have not yet fully transcended the colonial” (Williams and Chrisman, 1993, p. 4). In South Africa, the legacies of all types of colonialism mean that “many structural imbalances, inequalities and injustices remain stumbling blocks for the emancipation of black South Africans” (Heleta, 2016, p. 1). This understanding is linked to the concepts of coloniality and decoloniality, which are closely interlinked.

Coloniality, on the one hand, is understood “as a global power structure that sustains asymmetrical power relations between the Euro-American world and the Global South (Ndlovu-Gatsheni, 2013, p. 11). Furthermore:

Coloniality ... refers to long-standing patterns of power that emerged as a result of colonialism, but that define culture, labour, intersubjectivity relations, and knowledge production well beyond the strict limits of colonial administrations. Thus, coloniality survives colonialism. It is maintained alive in books, in the criteria for academic performance, in cultural patterns, in common sense, in the self-image of peoples, in aspirations of self, and so many other aspects of our modern experience.

Ndlovu-Gatsheni (2013, p. 13).

Decoloniality, on the other hand, refers to “the dismantling of relations of power and conceptions of knowledge that foment the reproduction of racial, gender, and the geo-political hierarchies that came into being or found new and more powerful forms of expression in the modern/colonial world” (Ndlovu-Gatsheni, 2013, p. 14). A “more massive and possibly more profound shift away from modernization towards decoloniality as an unfinished project took place in the twentieth century and is still unfolding now” (Maldonado-Torres, 2011, pp. 1–2). This project involves the “task of the very decolonization of knowledge, power and being, including institutions such as the university” (Ndlovu-Gatsheni, 2013, p. 14). It also needs to be valued as “liberatory thought that gestures towards the possibility of another world and knowledge” (Ndlovu-Gatsheni, 2013, p. 15). Both these aspects are taken up again during the discussion of recolonization and decolonization, respectively.

At this point, it is reiterated that the historical context needs to be taken into account when discussing concepts such as higher education internationalization and partnerships in Africa, as due to colonial legacies the impact of globalization is not felt evenly across the globe, and global networks are, first and foremost, controlled by the North (Tikly, 2001, p. 156). Highlighting how today's internationalization practices are being shaped by colonial histories entails a critical engagement with issues of power, knowledge and being, particularly from an ex-centric position. When critically engaging with internationalization, one cannot help but observe that inequality is a key characteristic of internationalization in general insofar as it is heavily dominated by Northern perspectives in terms of definitions, needs, practices and scholarship (Chasi, 2019). A critical analysis of internationalization is necessitated by the experiences and lived realities of people and institutions in the South.

A southern critique of internationalization

Decoloniality provides ex-colonized peoples with a space to judge Euro-America (Ndlovu-Gatsheni, 2013) and to challenge legitimating narratives (Appiah, 1991). Such critical engagement is in the main driven by those that “are judged to be less interesting, less significant, and more parochial than *those who hold the power to institutionalize their views* judge themselves to be” (Moya, 2011, p. 84). From a Southern perspective, the concept of internationalization can be critiqued at various levels. Internationalization stands accused of cognitive, cultural and technological imperialism (Ndlovu-Gatsheni, 2021). It can therefore not be theorized without paying attention to how it is influenced by globalization in economic, political and cultural ways (Beck, 2013). Traditional internationalization activities such as student and staff exchanges, research collaboration, credit harmonization and offshore programs, to name a few, “follow the logics of academic and intellectual structural adjustments to the dictates of market fundamentalisms” (Ndlovu-Gatsheni, 2021, p. 93).

Models and practices of higher education internationalization are mostly driven by Northern interests and needs. For example, the internationalization drive of the US is seen as an attempt to strengthen the country's quest for domination in political, cultural and social spheres, which is particularly noticeable in the context of study abroad programs (Martin and Griffiths, 2012). Regarding

projects focusing on harmonization of higher education, the Bologna model illustrates continued European dominance (Singh, 2010). This influence is direct, for example, in cases where qualification structures and curricula of African universities are aligned with the Bologna model, or indirect, in terms of borrowing policy models, aiming for greater regional cooperation and articulation or establishing systems for recognition of qualifications as well as quality assurance. The Bologna process is therefore seen as an instrument through which Europe exerts influence on higher education in Africa and other regions in the world (Singh, 2010).

A decolonial perspective also facilitates a critique of commonly used terminology, questioning the very definition of internationalization as a “process of integrating an international and intercultural dimension into the teaching, research and service functions of the institution (J. Knight, as cited in Knight, 2004, p. 9). Viewed from the South, this definition reflects perspectives of the colonizer, not the colonized. Higher education internationalization “cannot be a simplistic process of integration of ‘the rest of us’ into the ‘European’ game” (Ndlovu-Gatsheni, 2021, p. 85). Furthermore, if it is assumed that universities in Africa are essentially (still) colonial, then it stands to reason that they already are international. In that sense, the question arises whether universities in Africa need to de-internationalize to internationalize (Jowi, 2012), with de-internationalization, in this context, understood as synonymous with decolonization. Furthermore, Knight’s definition has become widely accepted and adopted by universities around the world. In a way, its analytical framework serves as a global template for analyzing internationalization in diverse regional and national settings (Singh, 2010). Such universalism is criticized in a decolonial context, where knowledge is situated, where different identities and experiences will shape different assumptions, questions and approaches (Moya, 2011). With all of this in mind, it is possible to interpret internationalization as a form of recolonization in the field of higher education.

Internationalization as recolonization

Drawing such a parallel recognizes that colonialism lingers on and shapes the world we live in today. Recolonization is understood here as “a new phase of ‘colonization’” (Saul, 2008, p. 73), one that is much less visible than colonization (Martin and Griffiths, 2012). Recolonization refers to relations where “the South continues to be the suppressed underside of the North” (Comaroff and Comaroff, 2012, p. 115). It captures “a process of quite self-conscious and unapologetic resubordination of much of the Global South ... to a western-centred capitalism ... to be carried out in the name of and under the guise of ‘globalization’” (Saul, 2008, p. 70). In this context, the nature of relationships in internationalization is interrogated from the perspective of power and how it operates (Beck, 2013). It is argued that internationalization, which is complexly linked to globalization, can be understood as a form of recolonization, particularly in light of continued patterns of Northern dominance and Southern subordination.

Recolonization, linked to “the hegemony of finance” and “market-driven and capital-driven economic priorities” (Saul, 2008, p. 73), manifests in higher education as much as in other spheres. In higher education, “The continuing dependence on Western universities and Western education further advances and entrenches former colonial influences” (Beck, 2013, p. 54). Regarding neo-liberal approaches to education, it is noted that “From a postcolonial perspective, a free market approach to both trade and education is seen as a continuation of those old policies of plunder that characterized colonial times” (Martin and Griffiths, 2012, p. 909). While economic dependencies play a key role, dependencies also continue in other ways, particularly regarding the domination of the North in knowledge production and dissemination. For example, powerful Northern institutions such as the World Bank and the European Union “have allocated themselves privileged roles of not only dictating whose knowledge matters, but more importantly, what knowledge is worthwhile and to whom” (Ndofirepi and Cross, 2014, p. 295). Furthermore, patterns of knowledge production and dissemination generally favor the North (Maringe and De Wit, 2016), and the North–South divide in knowledge production is illustrated as follows:

European and North American scholars are only too aware that they can ignore with impunity what is done in peripheral sites like the African continent, while any African scholar who similarly ignores western scholarship puts his or her professional competence at issue.

Nyamnjoh (2012, p. 16).

Linked to such examples of knowledge dependency, global university rankings are an instrument that perpetuates Northern dominance, as such rankings generally favor Northern institutions in terms of their positions in the league tables as well as the methodologies applied in the ranking exercise. Under the banner of international competitiveness, African universities, which are under enormous pressure to “perform or perish” (Rensburg, 2016, p. 307), subject themselves to international rankings and “push lecturers to publish in international journals yet do little to promote journals on the continent” (Nyamnjoh, 2012, p. 9).

Internationalization implicates universities in furthering relations that entrench and maintain a hierarchical world order (Beck, 2013). In that regard, “knowledge and education from Europe and North America are being globalized while that of the Global South is being pushed to the margins” (Ndlovu-Gatsheni, 2021, p. 80). With benefits of internationalization not evenly shared between institutions in the Global North and their counterparts in the Global South, internationalization is considered an “instrument of continued Northern domination in the field of higher education, which finds clear manifestation in North–South partnerships” (Chasi, 2019, p. 105). Facilitated by a decolonial lens, this assertion emerges from a deeper understanding of and critical engagement with the imbalances of resources, power and knowledge that are inherent in such partnerships, as outlined earlier. It takes the underlying power dynamics of these partnerships as well as their extractive nature into account. To illustrate this,

parallels are drawn between exploitative practices—from colonial relations focusing on the extraction of raw materials to the extraction of raw talent and data through internationalization in the present day (Chasi, 2019).

Acknowledging that traditional North–South partnerships have the capacity to entrench and perpetuate inequalities and asymmetries, future partnerships should be transformational in character and partners should engage on equal terms (De Wit, 2015). Applying a decolonial lens is useful in that regard, as it opens up the possibility of considering partnerships an opportunity to bring about change in North–South relations by decolonizing future interactions and by rehumanizing the world (Chasi, 2019).

Decolonization: toward a new paradigm of North–South partnerships

The decolonization agenda is, in general terms, driven by “those situated on the downside of some corresponding and relatively intractable power relation that serves to structure our global society” (Moya, 2011, p. 79). In the context of higher education internationalization, there is a need to counter currently dominant partnerships paradigms and trends and to envision a different future. *Agenda 2063*, the vision of the African Union (AU) for the continent, foresees “a self-re-empowerment of Africa; Africa determining its own future, leading its agenda and developing effective partnership frameworks with its external partners” (AU, 2015, pp. 42). Furthermore, the “new strategic role and place of Africa in the global arena will be based on mutually beneficial partnerships with the outside world” (AU, 2015, pp. 42).

At a Global Dialog on higher education internationalization hosted by the International Education Association of South Africa (IEASA) in 2014, representatives of several national, regional and other organizations working in the field highlighted the “importance of decision-making and practices in the development of internationalization activities that are imbued with ethical considerations and inclusivity” in the *Nelson Mandela Bay Global Dialog Declaration* (IEASA, 2014, p. 2). Furthermore, “internationalization must be based on mutual benefit and development for entities and individuals in the developed, emerging and developing countries” (IEASA, 2014, p. 2). Partnerships have an important role to play in that regard. Focusing on their equalizing role, partners are proposed to operate in a way that aligns with the concept of a “global commons” (Jooste, 2015). Considering the lack of equality a key threat to partnerships, the importance of substantive equality for mutually beneficial and reciprocal partnerships is highlighted (Hagenmeier, 2015). African higher education institutions should continue to develop partnerships with institutions in other world regions “whilst firmly invoking their rights as equal partners in them” (Hagenmeier et al., 2017, p. 88). Illustrating some practical implications of this, the following is observed:

Perhaps not all money should be accepted as a stimulant for development. More importantly, institutions in the Global South should become more centrally involved in determining the agenda behind partnerships, working to ensure that the flows of talent and the mobility of knowledge favors its developmental needs in the same way they do those of the institutions in the Global North.

Maringe and De Wit (2016, p. 311).

In South Africa, decolonization is one of the most pertinent issues affecting the country’s universities. It has most recently come to prominence in the wake of university student protests of 2015 and 2016, which formed part of what is now commonly referred to as the #FeesMustFall and #RhodesMustFall movements. The latter has gained prominence in and outside the country. It highlights the complexities of colonial legacies in South African higher education by addressing a variety of dimensions that urgently require transformation. These include statues, such as the one of Cecil John Rhodes on the campus of the University of Cape Town, which gave the movement its name, as well as symbols and names. In addition, it speaks, more broadly, to issues of access, financing and relevance, with a particular focus on institutional racism, white privilege and black pain.

For higher education institutions as sites of knowledge production, decolonization is particularly concerned with epistemological issues. It speaks to privileging “locally situated ways of seeing, thinking and knowing” (Frassinelli, 2018, p. 5) and involves “seeing the world from where we are, not from the perspective of the West” (Frassinelli, 2018, p. 6). It recognizes that the curriculum “remains largely Eurocentric and continues to reinforce white and Western dominance and privilege while at the same time being full of stereotypes, prejudices and patronizing views about Africa and its people” (Heleta, 2016, p. 2). A key element of decolonization is “the need to move the center from its assumed location in the West to a multiplicity of spheres in all the cultures of the world” (Wa Thiong’o, 1993, p. 16). On the African continent, decolonization is therefore inherently linked to Africanization, which focuses on increasing relevance to local contexts and needs by moving Africa to the center of attention. This entails, for example, placing Africa and African languages at the center of the curriculum, highlighting the contributions of Africa to world affairs (Maake, 2004).

Regarding the reconstitution of the literature department at the University of Nairobi in the late 1960s, a new ordering principle was applied focusing on “a study of Kenyan and East African literature, African literature, third world literature and literature from the rest of the world” (Wa Thiong’o, 1987, p. 94) to “establish the centrality of Africa in the department” (Wa Thiong’o, 1987, p. 94) without excluding literatures from other parts of the world. This organizing principle, which puts Africa at the center, is mirrored in South Africa’s national internationalization policy, where it is proposed that in their pursuit of internationalization South African universities must first give priority to the interests of South Africa, followed by those of the SADC states, the rest of Africa, the Global South and then the world beyond (DHET, 2019).

Decolonization has a role to play in bringing about change in internationalization. It has been argued that “there is no genuinely international higher education without decolonization of knowledge and education” (Ndlovu-Gatsheni, 2021, p. 78) and that decolonial internationalization is “predicated on a recognition of the diverse ways through which different people view and make sense of the world” (Ndlovu-Gatsheni, 2021, p. 79). Regarding the curriculum, internationalization in the African context seeks to develop a curriculum that “situates the notion of Africanness as a key lever for engaging with the global world in solving its developmental challenges and in seeking to position itself as a competitive entity in the globally competitive HE context” (Wits, 2014, p. 11). This is an expression of how “Africa and the African come first, before looking to the exportation of the same to the global arena” (Ndofirepi and Cross, 2014, p. 297), in alignment with objectives of *Agenda 2063*. What this means, for example, for internationalization objectives of preparing graduates to be global citizens, is illustrated as follows:

The primary aim of the discourse on Africanisation is to support the needs, interests and aspirations of people living on the African continent. The focus is on raising the awareness of students regarding the special and peculiar needs of their own societies, while simultaneously preparing those students for global citizenship. We shall call the process of creating continental and global awareness “dynamic symbiosis”—a process which will be running concurrently, infusing and buttressing each other.

Hagenmeier et al. (2017, p. 87).

In this sense, Africanization, as a contextual modality of decolonization, offers an opportunity to connect the local with the global, and vice versa, addressing the concern that “the international dimensions of higher education today may have become too disconnected from the local context” (Knight and De Wit, 2018, p. 3). Africanization provides an approach to not only engage with internationalization but to bring distinctly African perspectives to it. For example, African cultures can contribute to Western knowledge, as “one can extract from our indigenous cultures a lot of positive virtues which should teach the Westerner a lesson or two. The oneness of community for instance is at the heart of our culture” (Biko, 1987, p. 31). Instead of focusing on either Euro-America or Africa as the center, the world is re-imagined “with a multiplicity of centers” (Frassinelli, 2018, p. 7) and a “garden of many-colored flowers”, noting that the “floweriness” of the different flowers is expressed in their very diversity” (Wa Thiong’o, 1993, p. 42). There is “cross-fertilization between them. And what is more they all contain in themselves the seeds of a new tomorrow” (Wa Thiong’o, 1993, p. 42). Cross-fertilization between different centers can be facilitated by internationalization. What is important in that regard is that such internationalization confronts and subverts long-standing hierarchies of race, gender and class “for the purposes of re-membering/re-humanizing people and the production of relevant knowledge and just societies” (Ndlovu-Gatsheni, 2021, p. 81).

As a lens, decolonization can be used to address hierarchies and inequalities in higher education internationalization and partnerships. Regarding knowledge imbalances, for example, a decolonial, integrative perspective takes “the form of a genuine synthesis of all existing endogenous and exogenous knowledges, as a paradigm shift into a newly reconstructed and reconstituted space in which no cultural or social group’s knowledge wields more power than others” (Ndofirepi and Cross, 2014, p. 296). A South African case study (Chasi, 2019) reveals how the traditional paradigm of North–South partnerships in higher education can be changed toward more equitable and balanced engagements, facilitated by a decolonial approach. This requires all partners to deliberately and deeply engage with existing imbalances of resources, power and knowledge. It also requires all partners to be committed to mutuality, particularly regarding partnership outputs and benefits.

Transforming the nature of North–South partnerships toward more equitable, mutual engagements forms part of imagining the world differently. It contributes to internationalization that is “a liberatory and rehumanizing project engaging colonialism and dislocating it” (Ndlovu-Gatsheni, 2021, p. 94). Reimagining and transforming the nature of partnerships and internationalization begins with consciousness, with a process of “decolonizing the mind”, as in the title of Ngũgĩ wa Thiong’o’s book (1987). Decolonizing the mind has implications for both Northern and Southern partners. In general, institutions in the South should be less complicit and less accepting of inequalities in their relations with partners in the North. Particularly in early stages of partnerships development, they need to act with increased agency and self-determination. For example, partnerships aims, objectives and expected outcomes need to be negotiated with a certain amount of confidence and assertiveness to ensure that they meet the needs and interests of Southern partners as much as those of their Northern counterparts (Chasi, 2019).

While the imperative for change is generally driven by the South, Northern partners also have a role to play in shifting partnerships paradigms, particularly regarding a commitment to humility, mutuality and flexibility (Chasi, 2019). Humility includes an awareness of positions of power and privilege that Northern partners may enjoy as well as an openness to the possibility of learning from others. Mutuality involves active engagement with Southern partners to ensure that partnership activities create benefit and value for all involved. For this, Northern partners need to be empathetic toward their Southern partners and willing to see the world through Southern eyes. Finally, partnerships can benefit from greater flexibility on behalf of Northern institutions in terms of the conceptualization and implementation of partnerships, leaving room for contingency and the ability to respond to local needs and realities (Chasi, 2019). If partners in both the Global South and the Global North are willing to change their “terms of engagement” in such ways, it stands to reason that North–South partnerships will ultimately be more balanced, equitable, inclusive and socially just.

Conclusions

From a Southern perspective, partnerships between higher education institutions in the Global North and the Global South are often inherently problematic. This is mainly due to their imbalanced nature, which sees Southern partners generally benefitting

significantly less than their Northern counterparts. Through a decolonial lens, such partnerships are, in the broader framework of higher education internationalization, seen as an instrument of continued Northern domination and Southern dependencies. Against the backdrop of colonial links, the power dynamics of such partnerships amount to a form of recolonization.

At the same time, however, North–South partnerships provide us with an opportunity to shift the dominant colonial paradigm toward more equitable, inclusive and mutually beneficial engagements, thus opening up the possibility of a better future. With the Global South serving as a point of departure, decolonization can be used as a tool to critically reflect on existing inequalities in higher education internationalization and partnerships, with a view toward bringing about positive change. More specifically, Africanization, which is concerned with moving Africa to the center as a starting point for contextually relevant engagements, facilitates reflection on how partnerships can be transformed and enriched by African perspectives and contributions.

In line with the concept of convivial scholarship, which “sees the local in the global and the global in the local” and “brings them into informed conversations, conscious of the hierarchies and power relations at play at both the micro and macro levels of being and becoming” (Nyamnjoh, 2018, para. 20), North–South partnerships can be examples of convivial practice. In doing so, they allow for hopeful discourse on internationalization from the perspective of the Global South and contribute to moving the proverbial mountain by transforming partnerships in higher education and other spheres, thus laying new foundations for a different and more equitable future and for a global society that is socially more just.

References

- African Union [AU], 2015. Agenda 2063—The Africa We Want: Framework Document. Retrieved from: https://au.int/sites/default/files/documents/33126-doc-framework_document_book.pdf.
- Anderson, B., Maharasoa, M., 2002. Internationalization of higher education: facilitating partnerships between universities. *S. Afr. J. High Educ.* 16 (1), 15–21.
- Appiah, K.A., 1991. Is the post- in postmodernism the post- in postcolonial? *Crit. Inq.* 17 (2), 336–357.
- Beck, K., 2013. Making sense of internationalization—a critical analysis. In: Hébert, Y., Abdi, A.A. (Eds.), *Critical Perspectives on International Education*. Sense Publishers, Rotterdam/Taipei, pp. 43–59.
- Biko, S., 1987. *I Write What I like*. Heinemann, Oxford.
- Boshoff, N., 2010. South–South research collaboration of countries in the Southern African Development Community (SADC). *Scientometrics* 84 (2), 481–503.
- Chasi, S., 2019. North–South Partnerships in Public Higher Education: A Selected South African Case Study. Doctoral dissertation. University of the Witwatersrand, Johannesburg. Wits Institutional Repository Environment on DSpace. Retrieved from: <http://wiredspace.wits.ac.za/handle/10539/28064>.
- Comaroff, J., Comaroff, J.L., 2012. Theory from the South: or, how Euro-America is evolving toward Africa. *Anthropol. Forum* 22 (2), 113–131.
- Connell, R., 2007. *Southern Theory. The Global Dynamics of Knowledge in Social Science*. Polity, Cambridge.
- Crossley, M., Tikly, L., 2004. Postcolonial perspectives and comparative and international research in education: a critical introduction. *Comp. Educ.* 40 (2), 147–156.
- De Jager, K., Chimhundu, C., Harley, Y.X.R., Douglas, T.S., 2020. Collaboration and citation impact: trends in health sciences research at the University of Cape Town. *South Afr. J. Sci.* 116 (3/4), 1–8.
- De Wit, H., 2015. Partnerships for the future: trends, challenges and opportunities. In: Jooste, N., De Wit, H., Heleta, S. (Eds.), *Higher Education Partnerships for the Future*. Unit for Higher Education Internationalisation in the Developing World, Port Elizabeth, pp. 95–101.
- Department of Higher Education and Training [DHET], 2013. White Paper for Post-school Education and Training—Building an Expanded, Effective and Integrated Post-school System. Retrieved from: <http://www.dhet.gov.za/SiteAssets/Latest%20News/White%20paper%20for%20post-school%20education%20and%20training.pdf#search=internationalisation>.
- DHET, 2019. Policy Framework for Internationalization of Higher Education in South Africa. Retrieved from: <https://www.dhet.gov.za/Policy%20and%20Development%20Support/Policy%20Framework%20for%20Internationalisation%20of%20Higher%20Education%20in%20South%20Africa.pdf>.
- Fischer, K., Lindow, M., 2008. Africa Attracts Renewed Attention from American Universities. *The Chronicle of Higher Education*. Retrieved from: <https://www.chronicle.com/article/Africa-Attracts-Renewed/26527>.
- Ford, C., McMullin, J., 2016. Internationalization of higher education institutions: challenges and opportunities. In: Côté, J.E., Furlong, A. (Eds.), *Routledge Handbook of the Sociology of Higher Education*. Routledge, New York, NY, pp. 289–298.
- Foskett, R., 2010. Capacity building for demography in Southern Africa: international collaboration in action. In: Maringe, F., Foskett, N. (Eds.), *Globalization and Internationalization in Higher Education. Theoretical, Strategic and Management Perspectives*. Continuum, London, pp. 139–152.
- Frassinelli, P.P., 2018. Decolonization: what it is and what research has to do with it. In: Tomaselli, K.G. (Ed.), *Making Sense of Research*, first ed. Van Schaik, Pretoria, pp. 3–9.
- Gallwey, S.K., Wilgus, G., 2014. Equitable partnerships for mutual learning or perpetuator of North–South power imbalances? Ireland–South Africa school links. *Compare* 44 (4), 522–544.
- Gaventa, J., Runciman, C., 2016. Untangling economic and political inequality: the case of South Africa. In: United Nations Educational, Scientific and Cultural Organization, World Social Science Report 2016—Challenging Inequalities: Pathways to a Just World. UNESCO, Paris, pp. 70–73.
- Hagenmeier, C.C.A., Lansink, A., Vukor-Quarshie, G.N.K., 2017. Internationalization and African intellectual metissage: capacity-enhancement through higher education in Africa. *S. Afr. J. High Educ.* 31 (1), 81–103.
- Hagenmeier, C., 2015. Ensuring equality in higher education partnerships involving unequal universities in divergent contexts. In: Jooste, N., De Wit, H., Heleta, S. (Eds.), *Higher Education Partnerships for the Future*. Unit for Higher Education Internationalisation in the Developing World, Port Elizabeth, pp. 41–46.
- Heleta, S., 2016. Decolonization of higher education: dismantling epistemic violence and Eurocentrism in South Africa. *Transform. High. Educ.* 1 (1), 1–8.
- IEASA, 2020. Higher Education Internationalization in South Africa in the Context of the COVID-19 Pandemic—Impact on International Students. Retrieved from: https://ieasa.studyasa.org/wp-content/uploads/2020/05/IEASA-position-paper_COVID-19_May2020_Final.pdf. (Accessed 27 January 2021).
- International Education Association of South Africa [IEASA], 2014. Nelson Mandela Bay Declaration on the Future of Internationalization of Higher Education. IEASA, Port Elizabeth. Retrieved from: https://www.aieaworld.org/assets/docs/Press_Releases/nelson_declaration.pdf.
- Jooste, N., 2015. Higher education partnerships for the future: a view from the South. In: Jooste, N., De Wit, H., Heleta, S. (Eds.), *Higher Education Partnerships for the Future*. Unit for Higher Education Internationalisation in the Developing World, Port Elizabeth, pp. 11–21.
- Jowi, J.O., Knight, J., Sehoole, C., 2013. Internationalization of African higher education: status, challenges and issues. In: Sehoole, C., Knight, J. (Eds.), *Internationalization of African Higher Education—Towards Achieving the MDGs*. Sense, Rotterdam, pp. 11–31.
- Jowi, J.O., 2012. Should Africa De-internationalize to Internationalize? Issue 213 *University World News*. Retrieved from: <http://www.universityworldnews.com/article.php?story=20120318085122612>.
- Knight, J., De Wit, H., 2018. Internationalization of higher education: past and future. *Int. High. Educ.* 2–4.
- Knight, J., 2004. Internationalization remodeled: definition, approaches, and rationales. *J. Stud. Int. Educ.* 8 (1), 5–31.

- Koehn, P.H., Obamba, M.O., 2012. Transnational research and development partnerships in higher education. In: Deardorff, D.K., De Wit, H., Heyl, J.D., Adams, T. (Eds.), *The SAGE Handbook of International Higher Education*. Sage, Los Angeles, pp. 359–378.
- Maake, N., 2004. Conception of an African university. In: Seepe, S. (Ed.), *Towards an African Identity of Higher Education*. Visa University and Skotaville Media, Pretoria, pp. 157–170.
- Malaza, D., 2011. The contribution of South African higher education in Africa's development. In: International Education Association of South Africa. *Study South Africa—The Guide to South African Higher Education: Africa and the Knowledge Economy*, eleventh ed. IEASA, Durban, pp. 6–7.
- Maldonado-Torres, N., 2011. Thinking through the decolonial turn: post-continental intervention in theory, philosophy, and critique—an introduction. *Transmodernity* 1 (2), 1–15.
- Maringe, F., De Wit, H., 2016. Global higher education partnerships: equity and epistemic concerns with distribution and flows of intellectual capital. In: Côté, J.E., Furlong, A. (Eds.), *Routledge Handbook of the Sociology of Higher Education*. Routledge, New York, NY, pp. 299–314.
- Maringe, F., Foskett, N., 2010. Introduction: globalization and universities. In: Maringe, F., Foskett, N. (Eds.), *Globalization and Internationalization in Higher Education. Theoretical, Strategic and Management Perspectives*. Continuum, London, pp. 1–13.
- Martin, F., Griffiths, H., 2012. Power and representation: a postcolonial reading of global partnerships and teacher development through North–South study visits. *Br. Educ. Res. J.* 38 (6), 907–927.
- Menon, D., 2018. Thinking about the Global South. Affinity and knowledge. In: West-Pavlov, R. (Ed.), *The Global South and Literature*. Cambridge University Press, Cambridge, pp. 34–44.
- Moya, P.M.L., 2011. Who we are and from where we speak. *Transmodernity* 1 (2), 79–94.
- Ndlovu-Gatsheni, S.J., 2013. Why decoloniality in the 21st century? *Thinker* 48, 10–15.
- Ndlovu-Gatsheni, S.J., 2021. Internationalization of higher education for plurality: a decolonial reflection. *J. Br. Acad.* 9 (s1), 77–98.
- Ndofirepi, A.P., Cross, M., 2014. Transforming epistemologies in the postcolonial African university? The challenge of the politics of knowledge. *J. Educ. Learn.* 8 (4), 291–298.
- Nkwinti, G., 2016. Colonialism of a Special Type Lives On. Politics Web Online. Retrieved from: <http://www.politicsweb.co.za/opinion/colonialism-of-a-special-type-lives-on>.
- Nyamnjoh, F.B., 2012. Potted plants in greenhouses: a critical reflection on the resilience of colonial education in Africa. *J. Asian Afr. Stud.* 26.
- Nyamnjoh, F.B., 2018. Thinking, Researching and Writing Africa: Insights from Nigeria's Tutuola. *The Conversation*. Retrieved from: <https://theconversation.com/thinking-researching-and-writing-africa-insights-from-nigerias-tutuola-90236>.
- Obamba, M.O., Kimbwarata, J., Riechi, A.R., 2013. Development impacts of international partnerships: a Kenyan case study. In: Sehoole, C., Knight, J. (Eds.), *Internationalization of African Higher Education. Towards Achieving the MDGs*. Sense, Rotterdam, pp. 151–175.
- Ramphele, M., 2004. The university as an actor in development: new perspectives and demands. *J. High Educ. Afr.* 2 (1), 1–34.
- Rensburg, I., 2016. Reinventing greatness: responding to urgent global-level responsibilities and critical university-level priorities. In: Weber, L.E., Duderstadt, J.J. (Eds.), *University Priorities and Constraints*. Economica, London, pp. 307–318.
- Saul, J.S., 2008. *Decolonization and Empire. Contesting the Rhetoric and Reality of Resubordination in Southern Africa and Beyond*. Wits University Press, Johannesburg.
- Sehoole, C., 2006. Internationalization of higher education in South Africa: a historical review. *Perspect. Educ.* 24 (4), 1–13.
- Singh, M., 2010. Re-orienting internationalization in African higher education. *Glob. Soc. Educ.* 8 (2), 269–282.
- Sutton, S.B., Egginton, E., Favela, R., 2012. Collaborating on the future: strategic partnerships and linkages. In: Deardorff, D.K., De Wit, H., Heyl, J.D., Adams, T. (Eds.), *The SAGE Handbook of International Higher Education*. Sage, Los Angeles, pp. 147–165.
- Tedrow, B.M., Mabokela, R.O., 2007. An analysis of international partnership programs: the case of an historically disadvantaged institution in South Africa. *High Educ.* 54 (2), 159–179.
- Teferra, D., Altbach, P.G., 2004. African higher education: challenges for the 21st century. *High Educ.* 47 (1), 21–50.
- Teferra, D., 2012. Partnerships in Africa in the new era of internationalization. *Int. High. Educ.* 19–21.
- Tikly, L., 2001. Globalization and education in the postcolonial world: towards a conceptual framework. *Comp. Educ.* 37 (2), 151–171.
- University of the Witwatersrand [Wits], 2010. *Pan-African Universities Partnership*. In: *Draft Conceptual Framework*, vol. 6.
- Wa Thiong'o, N., 1987. *Decolonizing the Mind: The Politics of Language in African Literature*. Zimbabwe Publishing House, Harare.
- Wa Thiong'o, N., 1993. *Moving the centre*. In: *The Struggle for Cultural Freedoms*. East African Educational Publishers, Nairobi.
- Wa Thiong'o, N., 2009. *Something Torn and New. An African Renaissance*. Basic Civitas Books, New York.
- Williams, P., Chrisman, L., 1993. *Colonial Discourse and Postcolonial Theory. A Reader*. Harvester Wheatsheaf, New York.
- Wits, 2014. *Faculty of Humanities—Draft Framed Guidance for Curriculum Internationalization. Scholarship for Afro-Global Excellence (SAGE)*.
- Zezeza, P.T., 2006. Beyond Afropessimism: Historical Accounting of African Universities. *Pambazuka News*. Retrieved from: <https://www.pambazuka.org/governance/beyond-afropessimism-historical-accounting-african-universities>.

Reimagining internationalization of the curriculum

Betty R. Leask, School of Education, La Trobe University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	220
An emerging paradigm of internationalization of the curriculum “for all students”	222
Rethinking internationalization of the curriculum	227
Characteristics of best practice	228
Best practice principles for internationalization of the curriculum	229
Principle 1: Inclusive, holistic and values driven	229
Principle 2: Purposeful, planned, ongoing	229
Principle 3: Systematic, system-wide and context relevant	230
Conclusion	230
References	231

Introduction

Internationalization of the curriculum is a concept nested within the broader context of the internationalization of higher education and cannot be understood in isolation from that phenomenon. Both concepts have evolved over the past 25 years in response to critical commentary, research and changes in the world in which they are enacted.

In its simplest form the internationalization of higher education is a process related to researching, educating, communicating and trading across national borders. In 2015 de Wit, Hunter et al. updated Knight's (2004) commonly accepted working definition of the internationalization of higher education to the following (*italics indicate changes made to Knight's (2004) definition*):

the *intentional* process of integrating an international, intercultural or global dimension into the purpose, functions and delivery of post-secondary education, *in order to enhance the quality of education and research for all students and staff, and to make a meaningful contribution to society*
de Wit et al. (2015, p. 29).

The research informing the revised definition incorporated perspectives from the Global South as well as the Global North. Hence it provides a valid foundation for a discussion of best practice in internationalization in universities in different regions of the world today.

The revised definition maintains the emphasis on the importance of addressing “the intersection of international and intercultural” (Knight, 2004, p. 49) but in addition emphasizes the importance of *intentionality* and *purpose*. Intentionality implies coordinated, purpose driven planning and purpose is specifically linked to improving quality, engaging all students and staff and providing benefit to the wider society.

The study by de Wit et al. (2015) also recognized that there are global, regional, national and institutional commonalities and differences in the development of internationalization, but that common goals and objectives can also be observed. The 2015 definition they developed is also consistent with definitions of internationalization of the curriculum (Leask, 2009, 2015) and internationalization at home (Beelen and Jones, 2015) both of which focus on the development of all students' international/intercultural skills and understandings through active engagement with diverse people and ideas in class, on campus and in the local community.

Internationalization of the curriculum in higher education has been critiqued in the past for being poorly understood. Leask (2015) makes a clear distinction between the process of internationalization of the curriculum sometimes shortened to IoC) and its product, an internationalized curriculum, as well as the importance of understanding both.

Internationalisation of the curriculum is the incorporation of international, intercultural and/or global dimensions into the content of the curriculum as well as the learning outcomes, assessment tasks, teaching methods, and support services of a program of study
Leask (2015, p. 9).

The focus on a *program of study* highlights the need to plan and scaffold opportunities for all students to develop deep knowledge and advanced skills and hence move beyond approaches to internationalizing the curriculum based on isolated, optional experiences and activities for a minority.

An internationalised curriculum will engage students with internationally informed research and cultural and linguistic diversity and purposefully develop their international and intercultural perspectives as global professionals and citizens

Leask (2015, p. 10).

The focus on *engaging students purposefully* emphasizes the need for carefully planning and assessment of all students' international/intercultural learning across the program of study.

At the same time as IoC was being discussed in Australia, Canada and the UK, Internationalization at Home (IaH) was emerging as an idea in Europe, initially as "anything other than mobility" and later as:

the purposeful integration of international and intercultural dimensions into the formal and informal curriculum for all students within domestic learning environments

Beelen and Jones (2015a, p. 69).

Beelen and Jones (2015) summarize the common features of the two concepts:

They both aim to reach 100% of students, focus on the intercultural as well as the international; are embedded within the core formal and informal curriculum, not simply in the elective elements; are delivered through internationalised learning outcomes and assessment; do not depend on the presence of international students or staff and do not assume that their presence will automatically 'internationalise' the student experience; do not depend on teaching in English; and are specific to individual programmes of study and the academics who deliver them (p. 8).

In summary, the main difference between IoC and IaH is that IaH specifically and deliberately excludes mobility experiences, whereas IoC does not exclude it but sees it as insufficient on its own, emphasizing the importance of integrating the learning from those experiences into the curriculum at home for all students (Leask and Green, 2020).

Some have addressed the relationship between IoC and IaH by combining the two terms as "internationalization of the curriculum at home". In Europe and Latin America, both concepts have been used to focus internationalization on a broader range of activities than outbound and inbound mobility and programs taught in English. In Australia, the UK and Canada they have been used to draw attention to the importance of higher education institutions preparing all students to live and work, as citizens and professionals who understand the deep and complex relationships between the "local" and the "global" in their professional and personal lives.

The above discussion highlights just some of the complexity of internationalization of the curriculum, often masked by its ubiquitous nature. Some of this complexity is due to the fact that the terms *internationalization* and *curriculum* are themselves variously understood. The study of curriculum itself has much to offer both IaH and IoC. The curriculum has been less well researched in the higher education sector than in the school sector in the past, although it has received more attention in recent years, particularly in Australia and the surrounding region. Hicks (2018) notes that this literature uses a broad definition of curriculum as "a process for connecting and integrating university learning and teaching into coherent and meaningful experiences for students" (p. 1). He also noted an increase in research into internationalization of the curriculum and higher education and in the last decade.

This research has been informed by the work of various educational and social philosophers including but not limited to John Dewey (learning as an active, social activity), Jack Mezirow (transformative learning), Paulo Freire (the link between individual and social transformation); Pierre Bourdieu (the link between education and culture), Michel Foucault (the relationship between power, knowledge and social change) and Kwame Anthony Appiah (cosmopolitanism). The influence of these and other theorists on the practice of internationalizing the curriculum is evident in a focus on active learning and on learning outcomes and learning design that foster the development of students' critical skills and international/intercultural understanding in multiple settings over time. For example, within the formal degree program, but also in the co-curriculum, through engagement with diversity in class, on campus and in the community. They have prompted increased emphasis on the ways in which the personal transformation of students engaged in an internationalized curriculum will contribute toward social change, and specifically the creation of a fairer, more equal and just society—locally and globally. They have also prompted concerns about the elitist nature of approaches to internationalization focused on a minority of mobile students traveling primarily within the Global North.

With reference to a growing body of relevant literature this article discusses some characteristics of an emerging paradigm of the practice of internationalization of the curriculum and identifies three good practice principles for curriculum internationalization in the 21st Century. It is important to note that this discussion is a continual work in progress, as more research is conducted and as practice evolves.

An emerging paradigm of internationalization of the curriculum “for all students”

Six characteristics of this emerging paradigm of internationalization of the curriculum at home for all students are summarized here and described in more detail below, with reference to relevant literature.

First, conceptually and practically this emerging paradigm is based on an understanding of the “*curriculum*” as a *system*—a system in which the many parts work together to restrict or extend students’ learning in class, on campus and in the community, at home and abroad. A system that will directly influence the contributions graduates make to their local communities and to the world.

Second, *preparing all students for life and work in a globalized world*, and the attributes that should therefore be developed in graduates of an internationalized curriculum have assumed increasing importance in the discussion. This amounts to a shift in focus toward learning outcomes (of all students) rather than *inputs* such as mobility experiences (for an elite minority of students). There are a number of ongoing discussions concerning the specific nature of these attributes, how to achieve them and what to call them in an internationalized curriculum.

Third, connections between *decolonization of the curriculum and cognitive justice* and the internationalization of teaching and learning are emerging as increasingly important in discussions of internationalization of the curriculum. In some areas this is interpreted as programs of study making space for students to critically examine dominant knowledge paradigms, their strengths and their weaknesses. It may result in reading lists that are inclusive of a broad range of views of the world, rather than an exclusive range that reflects the global distribution of power.

Fourth, *interculturalization* of the curriculum is assuming greater prominence in this new paradigm. The intercultural and international elements of the curriculum have traditionally sat beside each other. Until recently the international orientation has dominated, always in first place (e.g., international/intercultural). However, this emphasis is shifting, with the intercultural assuming greater significance as a result of an increased focus on the development of international and intercultural learning outcomes for all students, regardless of their ability to have a mobility experience as part of their degree.

Fifth, *active learning* is increasingly recognized as at the heart of an internationalized curriculum, at home as well as abroad. For academic and student services staff, active professional learning in communities of practice is also recognized as an important part of both the initial and ongoing process of internationalization of the curriculum.

Sixth, in this emerging paradigm, internationalization of the curriculum is approached as a *planned, purposeful, ongoing process* of review and quality improvement.

(1) Curriculum as a system

Educational literature has defined the *curriculum* as being inclusive of the stated *purpose* of a course or program of study (including intended learning outcomes), the *teaching and learning processes*, what is *assessed* and *how* it is assessed, as well as *the students’ experience of learning beyond the classroom* through activities organized by the university on campus and in the community (Barnett, 2000). Mestenhauser (1998) identified the importance of moving away from “fragmented, unintegrated and enclosed” (p. 7) approaches to internationalization of the curriculum—fragmented by disciplinary thinking, approached as an add-on, enclosed within individual projects and courses. Such an approach denies the potential that the complexity of curriculum as a system offers. Curriculum is indeed complex having multiple dimensions, each comprised of a network of interconnected parts. Leask (2015) identifies three dimensions of the curriculum. The formal, informal and hidden curriculum. She describes the *formal curriculum* as the content described in the syllabus and the orderly, planned schedule of experiences and activities that students must undertake as part of their degree program. The *informal curriculum* is the learning that happens out of class, including for example support services and additional activities organized by the university and its affiliated associations, clubs and societies. These activities are usually optional and not assessed, although they may support assessed learning within the formal curriculum. Examples include peer mentoring programs, peer assisted study sessions and cultural festivals organized by clubs and societies. These activities are an important part of the overall student experience, and they can contribute to the creation of campus culture which supports international and intercultural learning.

The third important dimension of the curriculum is the *hidden curriculum*, rarely discussed and largely invisible to many. Studies of the hidden curriculum in higher education are also more uncommon in higher education than in primary and secondary education (Margolis et al., 2001). The hidden curriculum comprises the various unintended, implicit and hidden messages including the norms, values and belief systems that are embedded in both the formal and informal curriculum. For example, the textbooks that are selected send a “hidden” message concerning *whose knowledge counts* in the curriculum and by implication, whose does not (see Leask, 2015, 7-15). Attention to the hidden curriculum invites consideration of issues around decolonization of the curriculum and cognitive justice (discussed below). Hidden messages are also conveyed through the informal curriculum when for example all international students are required to complete cross-cultural skills training prior to the commencement of classes but the same is not required of domestic or home students. Is this because the domestic students have the required skills? Or perhaps these skills are not important for domestic students because it is up to international students to “fit in” and make adjustments? Are these the intended messages, part of a socialization process that indeed assumes that international students need to become “more like us”?

The formal, informal and hidden elements of the curriculum are connected, rather than discrete—a dynamic system of parts working together as a complex whole. Intended learning outcomes, pedagogy, assessment, disciplinary knowledge systems, generic skills development, content and experiences in and out of the classroom all shape the lived experience of all students

and contribute significantly to their international and intercultural learning throughout their degree program. The point where all three elements of the curriculum connect is a dynamic space, offering rich opportunities for learning for all students.

While all three dimensions of the curriculum are identifiable and important, when internationalization and interculturalization of the curriculum are discussed, it is the formal and informal curriculum that is most often the object of discussion. However, it is equally important to pay attention to the hidden curriculum.

(2) Preparing all students for life and work in a globalized world

Increasingly we see a move toward approaches to internationalization of the curriculum at home focused on the development of graduate attributes related to global learning and global citizenship. Graduate attributes (sometimes called “graduate qualities”) are used in universities predominantly in the US, UK, Canada, Australia and parts of Europe to describe the unique set of soft skills, values and attitudes that will characterize graduates from a particular university. The purpose of internationalization of the curriculum to prepare graduates for “the highly interdependent and multicultural world in which they live and (will) have to function in the future” (Harari, 1992, 53) is facilitated by the description of graduate attributes that are specifically international or intercultural in nature at institutional, program and course level (Leask, 2001).

Internationalization is also connected with the preparation of students for life and work in a global economy (Robson, 2011) and therefore with the employability agenda (Jones, 2013). Graduates it is argued will need to be employable internationally and/or be able to work in multicultural/multinational teams and/or should be committed to ethical action and social responsibility on a global scale.

Global citizenship, described as a combination of knowledge, skills and attitudes that will prepare students to live and work in a globalized world, has increasingly been used as a graduate attribute of an internationalized curriculum. For example, Ogude (2007) argued that internationalization of the curriculum in South Africa should not only generate new knowledge but should also be connected to preparing graduates to be globally competitive. In 2009, the Association of Universities and Colleges of Canada suggested that an internationalized curriculum is “a means for Canadian students to develop global perspectives and skills at home” (AUCC, 2009, p. 5). Global citizenship has become “part of the internationalization discourse in higher education around the world,” (Deardorff and Jones, 2012, p. 295).

A trend to focus specifically on how to provide equal opportunities for *all* students to develop the awareness, skills, knowledge and attitudes required to be responsible global citizens is also evident in university strategies, such as that of Florida International University <https://goglobal.fiu.edu/> whose Global Learning program is focused on providing opportunities for “every student to engage with others to find connections among divergent perspectives ... to create innovative, equitable, and sustainable solutions for our interconnected human and natural communities”.

The term *global citizenship* is not, however, a benign concept. It is problematic through its association with neo-liberal agendas in higher education. There is a danger that narrow notions of global citizenship, focused only on the development of students as economic beings, for example focused on developing students’ ability to work for multi-nationals who trade for advantage over less economically developed groups, is more likely to exacerbate rather than ease existing inequalities produced by colonialism and the dominance of the Global North. It can also be resisted by academic staff who connect the ideal of global citizenship with a capitalist society and its Western heritage (Clifford and Montgomery, 2014). In this regard the very idea of global citizenship is inconsistent with decolonization of the curriculum and cognitive justice also characteristics of the emerging paradigm described here (and discussed in more detail below).

Discussions of cosmopolitanism offer the possibility to resolve this tension. Rizvi (2009) suggests the need for a stronger focus in education programs on cosmopolitanism driven by recognition of increasing global interconnectedness and interdependence, and the need for global solutions to global problems—often cited as rationales for developing students as global citizens. Cosmopolitanism is interpreted in different ways, including as a transnational lifestyle and an ethical orientation toward globalization (Appiah, 2006). Rizvi (2009) argues that cosmopolitanism is only worth pursuing “if we are able to use it as an instrument of critical understanding and moral improvement” (p. 263). Rizvi’s work has influenced discussions of internationalization of the curriculum in significant ways. It has for example resulted in increased emphasis on the development of students’ criticality as part of internationalization of the curriculum. Rizvi (2009) argues that cosmopolitan learning is less about “imparting knowledge and developing attitudes and skills for understanding other cultures per se” (p. 266), and more about helping students to understand the ways in which global processes create inequality, how they may unwittingly be contributing to that inequality and how they might challenge dominant ways of thinking about global connectivity to create a more equal world. This shift is important if we consider higher education institutions as active players in contributing to social change and the creation of a better world (Escrigas et al., 2014).

These ideas have prompted discussion of the importance of developing students’ capacity to imagine difference, question their own assumptions, think as the “other”, and walk in others’ shoes (Lilley et al., 2014). They have also stimulated other discussions around the importance of the ideal global graduate having a personal ethic which is both local and global in scope and includes acceptance of personal accountability for local actions that will have an impact on the lives of people in other, faraway places. Principled decision-making, solidarity across humanity (Schattle, 2009) and a commitment to the collective well-being (Rizvi and Lingard, 2010) are also identified as important for “responsible” global citizenship (Leask, 2015). These discussions continue with some agreement that an internationalized education should develop graduates who have a particular ethical stance in relation to the world.

A focus on providing opportunities for all graduates to develop global mindsets, an understanding of inequalities in the world, and a willingness to accept a form of global social responsibility at home is increasingly the focus of curriculum internationalization. This requires a holistic view of learning and the development of students' global selves (Killick, 2015), a curriculum that addresses the complex, contested and dynamic nature of knowledge and ensures that the scope of whose knowledge counts in the curriculum is broad. Institutional approaches that recognize internationalization as a powerful force for change at the personal level as well as the national and global level are also required. This suggests strategies that connect internationalization of the curriculum at home for all students with advancement of the global common good (Marginson, 2011) for example through programs in the formal and informal curriculum related to the achievement of the United Nations Sustainable Development Goals (Gegersen-Hermans, 2021).

(3) A focus on decolonization and cognitive justice

One of the results of colonization has been that Western ways of knowing have shaped our educational frameworks and our dominant political and economic systems. Decolonization of the curriculum is a critical view of traditional ways of knowing and learning (Stein and Andreotti, 2016) that stimulates the exploration and nurturing of new possibilities for content and pedagogy, and broadens the foundation of knowledge and values on which the curriculum is based.

There is no one way to decolonize the curriculum. We might, for example, have students critically examine the ways in which dominant approaches to knowledge production and distribution have shaped international markets and economies and the positive and negative impact of that on the local market and economy. We might also consider if and how this contributes to reproducing and strengthening social problems from generation to generation, perpetuating and exacerbating inequalities; and explore alternative approaches and what they might offer (Escrigas et al., 2014). Decolonization of the curriculum might also include imagining and envisioning how alternative cultural views of the world and diverse knowledge systems can be acknowledged, examined, valued and included for the perspectives they offer.

In terms of practical examples, in areas such as medicine, physics, nutrition, and geology a curriculum based solely on commercial research funded by companies in the Global North may distort the community of knowledge and result in social injustice on a global scale (McArthur, 2013). In this case decolonization of the curriculum might involve students examining how the commercial funding of the research on which the curriculum is based might have had an impact on the questions researchers asked, where research was undertaken and who benefited from it. They could also consider how commercial funding has in some areas resulted in secrecy and restricted access on the open sharing of ideas, competition and economic self-interest replacing the common good of humanity.

In the humanities, Breit et al. (2013) discuss de-westernization of the journalism curriculum. However, while positive steps are evident in general education and in specific disciplines such as sociology (Hack, 2020) this area is still often overlooked.

Cognitive justice refers to the right of different knowledges to coexist so long as they sustain the life, livelihoods, and life chances of people. A cognitively just curriculum will recognize that different ways of knowing and understanding co-exist in the world, each intimately connected to the cultures, practices and beliefs of groups of people; and that all societies, not just the post-industrial, are knowledge societies. It will also recognize that higher education privileges some knowledges over others relegating that which is tribal, Indigenous and traditional to obsolescence—and will take steps to address this situation (Visvanathan, 2021).

There is no doubt that in both the formal and the informal curriculum in higher education in some countries in the Global South local knowledge traditions and paradigms are valued less than those coming from the Global North. This is in part because colonization has provided a vehicle for internationalization and modernization through the introduction of international and knowledge and values into the curriculum. The key issue here is the extent to which those perspectives dominate, the status they are afforded, and how they are used. Are imported knowledge and values presented as superior? Have they replaced traditional knowledge in the curriculum completely? If so, then paradoxically internationalization of the curriculum at home through the lens of decolonization and cognitive justice is akin to curriculum localization. It might for example, include creating spaces for Indigenous knowledge and approaches to learning in the curriculum. For example, there may be a requirement that all academic programs include Indigenous knowledge or perspectives in consultation with local Indigenous people, and that all commencing students complete a compulsory co-curricular module exploring Indigenous Australian history, culture and customs and the foundations of Indigenous knowledge. Such programs simultaneously communicate the extent to which the university values Indigenous knowledge, develops in students a broader graduate capability of cultural literacy by requiring students to engage in critical reflection about the cultural foundations of knowledge and their own attitudes, values and beliefs (La Trobe University, 2016).

A range of other approaches to reimagining internationalization of the formal and informal curriculum through decolonization of the curriculum and cognitive justice are described in Robson (2011), de Wit and Leask (2017), Green and Whitsed (2015) and Montgomery (2019).

(4) Intercultural first, then international

The *intercultural* has long been associated with the *international* in discussions of internationalization of the curriculum. Many different definitions of intercultural competence have been applied within the field. They include: "the ability to communicate effectively and appropriately in intercultural situations based on one's intercultural knowledge, skills and attitudes" (Deardorff, 2006, p. 247), "the understanding of others' world views" (Deardorff, 2006, p. 248), knowledge of others and of oneself and

skills to interpret, relate, discover, interact as well as a valuing of others' values, beliefs, and behaviors (Byram, 1997), "the culture-specific and culture general knowledge, skills, and attitudes required for effective communication and interaction with individuals from other cultures" (Paige et al., 2003, p. 177). In their model of "Global Literacy" Killick and Foster (2021) include "cross-cultural capabilities", which "incline and enable students to interact successfully with people they perceive to be different: socially, civically and professionally" (p. 10). Discipline specific examples and models also exist, such as that of Freeman et al. (2009), who developed a taxonomy of intercultural competence designed to assist academic staff to map existing opportunities, as well as design and incorporate new opportunities, for students to become interculturally competent in their study program.

Intercultural competence is clearly multi-faceted and requires the development of complex skills, knowledge and attitudes over time. However, discussions of pedagogies to develop intercultural competence that have been tested in diverse discipline contexts have until recently been quite limited. This is in part because intercultural learning was for a long time wrongly assumed to be an automatic outcome and benefit of intercultural contact on campus, in class and abroad, more intense and almost instantly "transformative" when it occurred abroad. However, multiple studies have shown that this is rarely the case, that intervention is required in classrooms at home and while students are abroad if intercultural competence is to be developed in a program of study (see Leask and Carroll, 2011 for a summary of multiple studies up to that date). There has also been interesting research conducted into pedagogies to develop intercultural awareness, capacities, capabilities and competencies in diverse classrooms at home. Guidance, case studies and examples are described by, for example, Carroll (2015), De Louw and Bulnes (2021), Killick and Foster (2021), Leask and Carroll (2013) and Townsin and Walsh (2016). There is also useful guidance on the sorts of interventions that have worked to improve intercultural learning through connecting learning abroad with learning at home in study abroad programs. See for example, Leask and Green (2020), Gothard et al. (2012) and Weber-Bosley (2010).

However, while the importance of developing all students' global mindsets, skillsets and also heartsets, through an internationalized curriculum has been recognized, until recently the focus has still been very much on international activities such as study abroad and exchange, only accessible to a minority of privileged students worldwide. An increased focus on internationalization of the curriculum *at home* as an institutional strategy has seen some changes in thinking. The discourse has started to shift, to highlight the importance of *interculturality* in the development of students as global citizens and professionals (see for example Jiang, 2011; Jones, 2016, 2019). This is evident in institutional strategies that refer directly to *interculturalization of the curriculum* and provide training for faculty in the design of intercultural learning outcomes and inclusive pedagogy in order to prepare students as effective professionals and citizens in increasingly complex and diverse global and local contexts (Queen's University, 2022; Thompson Rivers University, 2022).

Interculturality is similar to and different from intercultural competence. Interculturality refers specifically to mutually respectful interactions and dialog between people from different cultures, that promote mutual understanding while preserving the cultural identity of each individual. Importantly, the concept of interculturality not only refers to the relations that develop between individuals belonging to different countries or regions of the world, it also takes into account, individuals who are within the same community but who have different characteristics from an ethnic, social or other point of view (Killick, 2017; Killick and Foster, 2021). It is therefore a valuable concept to guide internationalization of the curriculum at home, for all students, because globalization has resulted in increasing diversity in local communities and the massification of higher education has also resulted in increased socio-cultural diversity in classrooms.

Bringing interculturalization of the curriculum to the forefront is a means of engaging a majority, if not all students, in a higher education program that develops the skills and understandings they will need as graduates to contribute to a complex, interconnected, global world. It is a means of recalibrating internationalization to achieve a better balance between activities focused on international/intercultural engagement abroad for a minority of students with activities focused on developing all students' intercultural skills and understandings at home. This balance is important because the terms *interculturalization* and *internationalization* have been closely connected in definitions of the internationalization of higher education since 2004 (Knight, 2004).

Curriculum, teaching and learning have been identified as at the heart of the internationalization process (Hudzik, 2011). Bringing the intercultural to the forefront of program design and delivery usefully broadens the range of theoretical concepts and ideas influencing content and pedagogy to include for example, intercultural theory. When interculturalization is included in a university strategy it positions active engagement with cultural diversity in local classrooms and communities as being at least as important for intercultural/international learning as engagement with cultural diversity abroad.

(5) Active learning

It is increasingly recognized that *active learning* whether in class, on campus, in local communities or abroad is the heartbeat of an internationalized curriculum for students. Active learning includes the use of pedagogies that support students' participation in a range of authentic intercultural/international learning experiences within their program of study, as well as in the informal curriculum. Active learning increases student motivation and engagement, develops critical thinking skills and promotes deep learning. Nussbaum (2010) argues that it is important that students are challenged and supported to become 'thoughtfully critical in a complex world' (p.18). For many, mobility programs are the gold standard of active learning in an internationalized curriculum. However, active, experiential learning is also possible vitally important for those students who cannot participate in a mobility experience. It is also possible to provide all students with opportunities for active, critical intercultural and

international engagement if cultural diversity at home is recognized, and celebrated and students are supported to engage with the “other” in class, on campus, in the community and online. Approached in this way active learning may also assist the integration of international students into local classrooms (Marrone et al., 2018).

Active, experiential learning has also been found to increase the engagement of academic and professional staff in internationalizing their curricula in communities of practice (Green and Whitsed, 2015; Zou et al., 2020). Breit et al. (2013) also found that a diverse community of practice—one that included staff from different cultural, linguistic and disciplinary backgrounds—learned a great deal when they were encouraged and supported to challenge their assumptions about each other, their students and the curriculum and to work together to internationalize their curriculum design and delivery.

(6) A purpose driven, ongoing process

The development of international and intercultural knowledge, skills and attitudes in across a program of study requires purposeful planning and ongoing review of impact on student learning outcomes. The development of skills such as language capability and intercultural competence may need to be embedded in a number of courses at different levels. A range of strategies to assist all students to achieve desired learning outcomes by the end of the program may be required. These might include strategies which mobilize and utilize student services and the informal curriculum in supporting the work undertaken in the formal curriculum.

The internationalization of *higher education* and of *the curriculum* are related and co-dependent processes. They are ongoing, continuous and in some ways at least contextually unique. Context in relation to internationalization of the curriculum is multi-dimensional, including the disciplinary as well as the institutional context. Thus, an internationalized curriculum in nursing will look quite different from one in teaching, engineering, the arts or the sciences. However, it also includes national, regional and global contexts and the ways those contexts interact with each other (see Fig. 1:1, A conceptual framework of internationalisation of the curriculum). Increasingly that context includes the global as it interacts with the local in the development of students as social, human and economic beings.

The framework reflects the “supercomplex” world in which we live—one in which the very frameworks by which we orient ourselves to the world are themselves changing and contested (Barnett, 2000, 257). This world requires regular review and reconstituting of the curriculum as priorities in the different layers of context (represented by the lower half of the framework) shift and change, interdependently. Similarly, the framework allows for variation in the level of influence exerted by different

A conceptual framework of internationalisation of the curriculum



Fig. 1 A conceptual framework of internationalization of the curriculum (Leask, 2015, p. 27).

layers of context in different regions or nations. For example, in some regions the national context may be more important than the regional. Research has found that overwhelmingly academic staff assign primary importance to the disciplinary context (Clifford, 2009). Curriculum decisions usually reflect *dominant paradigms* in the discipline, in international education and in learning design (represented in the top half of the framework) but in the process of internationalization *emerging paradigms* should also be considered.

The staged process of internationalization of the curriculum represented in Fig. 2 is essentially a traditional action research approach applied to the review and revision of curricula. However, there is one key difference between the *Process of IoC* and traditional approaches to curriculum review. This is the inclusion of an *Imagine* stage. In this stage those designing the curriculum are invited to explore new possibilities and ideas, challenge dominant paradigms and think in new ways, taking the perspectives of others into account. This is important because Bartell (2003) found that “some disciplines tend to adopt a relatively narrow focus, impoverished by an absence of intercultural and international perspectives” (p. 49). The process of internationalisation of the curriculum (described in more detail in Leask, 2013, 2015) prompts staff to push the boundaries of possibility in relation to the curriculum by challenging dominant disciplinary paradigms of content, teaching and learning, assessment and support services. The process is applicable to the internationalization of the informal curriculum as well as the formal curriculum. In the review and design of the informal curriculum, student services staff and students are usually the principal actors, whereas in the formal curriculum it is academic staff who are the principal decision-makers.

Another key element of the process of IoC in this emerging modern paradigm of internationalization is working with students as partners in the process of internationalizing the curriculum. Using students as partners is “a process through which all participants have the opportunity to contribute equally, although not necessarily in the same ways, to the conceptualization, decision-making, and practice of teaching, learning and assessment” (Cook-Sather et al., 2014, 6–7). Green’s (2018) research into the engagement of students in the process of internationalization found that using students as equal partners, rather than casual informants, challenged naturalized institutional practices concerning access and equity, outcomes and process, and power and privilege. Her research suggests that approaching students as partners opens up new possibilities and may be a powerful tool for those who seek to create a more inclusive and equitable form of IoC.

Rethinking internationalization of the curriculum

The concept and practice of IoC have clearly evolved over time in response to changes in the world and changes in thinking based on research and scholarly practice. These changes include but are not limited to increasing recognition of the need for all students to understand the relationship between the local and the global, research on approaches to IoC in different disciplines and programs and the use of different theoretical frameworks to shape thinking and practice.

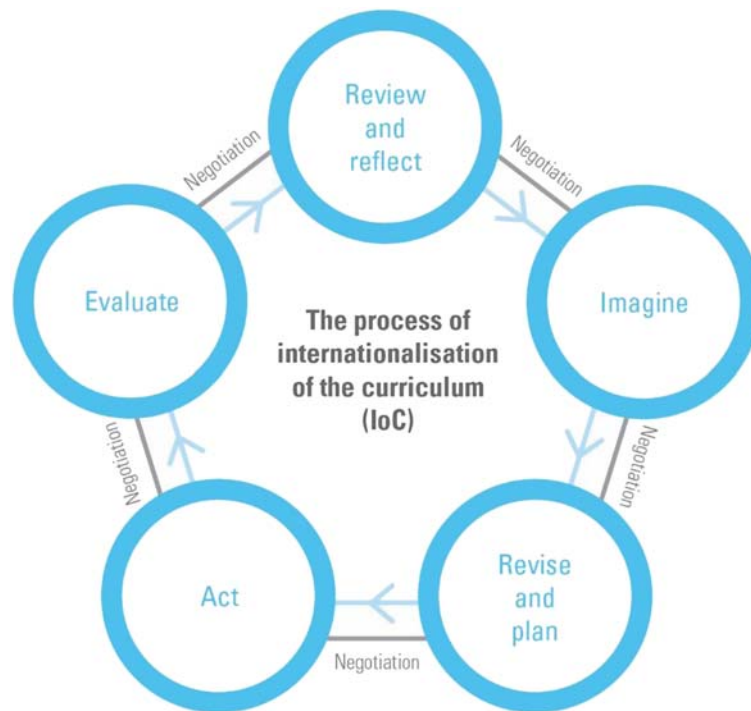


Fig. 2 The process of internationalization of the curriculum (Leask, 2015, p. 42).

The characteristics of an emerging paradigm of internationalization of the curriculum described above have implications for how internationalization of the curriculum is defined and described, which will also influence how we define best practice in IoC. The following “rethinking” of internationalization of the curriculum attempts to capture the salient features described above as the foundation for the discussion of best practice in IoC that follows in the next section.

Interculturalisation/internationalisation of the curriculum is the strategic incorporation of intercultural, international, and global dimensions into the content of the curriculum as well as the learning outcomes, assessment tasks, teaching methods and support services of a program of study. It is a collaborative, scholarly process of review and renewal of the formal, informal and hidden curriculum that respectfully incorporates diverse points of view. Inclusive pedagogies will purposefully engage all students in active intercultural/ international learning at home in class, on campus and in their local communities. The ultimate goal is to prepare all students to make valuable contributions to society as ethical, caring citizens and professionals in an interconnected rapidly changing world

Based on [Leask \(2015, 9\)](#).

This description of internationalization of the curriculum brings the intercultural to the forefront and highlights the importance of the values underpinning the internationalization of higher education. It emphasizes the need to engage *all students* actively in an intercultural/international curriculum. It highlights that the purpose of an internationalized/interculturalized curriculum is to ensure that all graduates are able to make valuable contributions to local and global society as human, social and economic beings.

This means that all students will require:

- intercultural capabilities including the capacity to further improve and develop their intercultural skills and knowledge in their professional and personal lives;
- an understanding of the ways in which what happens locally is and may be influenced by international and global tensions and possibilities;
- an appreciation of the cultural foundations of knowledge, and how international power relations have influenced the disciplines and professional practice;
- an intercultural critical consciousness that both “enables and disposes” globally diverse students to work together to see the complexity of an issue, analyze it and take action based on their collaborative analysis [Killick and Foster, 2021](#).

A curriculum that will do the above requires careful, critical course *and* program design (so that for example, learning of complex intercultural skills can be scaffolded across different year levels). It requires attention to describing *and* assessing intercultural/international goals in the informal as well as the formal curriculum while also identifying the impact of the hidden curriculum. It is concerned with the impact of the system of curriculum design on the total student experience, and to the ways in which international, intercultural and global relations impact everyone’s life and work.

Characteristics of best practice

In simple terms best practice is a method or approach that is superior to alternatives because it produces better results in a particular context. What is best practice in one context may not be best practice in another. Hence *best practice can only be measured against a purpose or desired outcome that makes sense in the context in which it occurs*.

Curriculum design is a cultural statement. Internationalized curricula are the result of complex decisions influenced by multiple layers of context and the values of those who design the curriculum. The culture of the discipline, the program of study and the institutional context will all influence the decisions made by key decision-makers. Decisions such as whose knowledge and points of view will be included, what skills students will develop and how these will be taught and assessed are often informed by ideology. All decisions, by their very nature, exclude other possibilities, sometimes intentionally, and at other times, unintentionally. Disciplines have different approaches to knowledge production and dissemination and to inducting students into their culture and preparing them for professional practice and life as graduates. This has resulted in different interpretations of what internationalization of the curriculum means across disciplines and professions ([Clifford, 2009](#); [Stohl, 2007](#)).

Because approaches to internationalization are also influenced by regional and national higher education policies, best practice in internationalization of the curriculum can only be fully understood and nurtured in context. The conjunction of disciplinary, institutional, regional, national and global contexts will shape both the goals for an internationalized curriculum and the means of achieving them. As contexts are fluid, *approaching IoC as a cyclical process of curriculum review and renewal* is a defining feature of best practice.

As well as fitness for purpose and context, and a process approach, another defining feature of best practice in IoC, is *evidence-based practice*. There is a growing body of research into intercultural teaching and learning and internationalization of the curriculum. Best practice will be informed by this constantly evolving international and interdisciplinary evidence-base, which like other contextual layers, is constantly changing as new research is conducted and new knowledge is created.

Evidence-based practice in IoC also includes scholarly professional practice. Aside from formal research studies, practice in IoC should also be informed by small, local studies, what is often referred to as scholarly professional practice. This requires the *systematic collection and analysis of data and its use to improve educational practices and outcomes within a program or institution*.

In summary, best practice in internationalization of the curriculum will have the following characteristics:

1. have a defined purpose or desired outcome that makes sense in the context in which it occurs.
2. be fit for purpose within that context.
3. be approached as an ongoing cyclical process of curriculum review and renewal.
4. be informed by relevant international interdisciplinary research and methods, as well as locally collected research and evaluation data.

Still, some of these characteristics are difficult to create. The following best practice principles capture core ideas from recent research and developments in thinking concerning internationalization of the curriculum. They are specific enough to guide practice and flexible enough to accommodate application in a variety of different contexts. This is important because research, frames of reference, approaches and discourses of internationalization of the curriculum are dominated by the Global North and as such “privilege the needs and realities of universities in wealthy, industrialized countries in North America, Europe and Asia” (Majee and Ress, 2020, p. 475).

Best practice principles for internationalization of the curriculum

Best practice principles are particularly important for internationalization of the curriculum because of its sensitivity to context in both design and teaching. The three principles proposed here are consistent with the characteristics of best practice identified above.

Principle 1: Inclusive, holistic and values driven

As all graduates will live and work, in an increasingly complex, unpredictable and interconnected world, best practice in internationalization of the curriculum will consider what this means for *all students'* intercultural and international learning, in the design of the formal and informal curriculum.

An internationalized curriculum is a composition that has many connected parts. Outcomes of an internationalized curriculum will be holistic in that together, internationalization of the formal and informal curriculum will have an impact on students as human, social and economic beings.

Discourse is a powerful and often overlooked source of marginalization and privilege for individuals and groups. Discourses of internationalization and interculturalization that are inclusive of the complex realities of those with different histories, for example those affected by the legacies of colonial expansion, will result in different approaches to internationalization that push the boundaries of best practice. Approaching internationalization of the curriculum as a *systematic, purpose-driven ongoing process* that includes critical, inclusive and complicated conversations opens up opportunities to reshape the discourse in ways that would otherwise be impossible.

International and intercultural learning outcomes, content, pedagogy and assessment services are important parts of internationalization of the formal curriculum. Course and program design decisions will often be values driven. Institutions may choose for example, to use the Sustainable Development Goals or Global Citizenship to express their values through a university-wide institutional commitment to internationalization of the curriculum. Such approaches invite and allow many different but related activities to connect with and contribute to a larger whole.

Extra-curricular programs that are holistic and values driven can also be adapted to different contexts. Macquarie University's Global Leadership Program (GLP), launched in the Australian university sector in 2005, has since been adopted and adapted by many other universities in and beyond Australia (Macquarie University, 2021). The program focusses on students' responsibilities and opportunities as global citizens, while increasing their employability and preparing them for a globalized workforce in the 21st century. Skills and knowledge are primarily developed in informal/co-curriculum, and students are actively engaged in the creation of a unique program tailored to their interests which they complete at their own pace. Such programs signal a university's commitment through internationalization to the global common good as well as the private benefit of students.

Student and academic staff interactions with professional associations and community organizations, colleagues and fellow students in other parts of the world, and from different cultural and linguistic backgrounds (at home and abroad) are also an important part of an inclusive and holistic approach.

Principle 2: Purposeful, planned, ongoing

Purpose is expressed in different ways within an institution. It is important that the rationale for and purpose of interculturalization/internationalization of the curriculum is consistent across institutional mission statements, goals and plans and program and course statements and is focused on both the common and the personal good.

Carefully formulated plans and actions directed to achieving intercultural/international learning goals consistent with the mission and values of the institution are required in the short and the long term, for the formal and informal curriculum. Regular

monitoring of achievements against goals, informed by local quantitative and qualitative data will ensure that resources are allocated/reallocated efficiently and effectively.

Learning design at the program and course level can provide all students with opportunities to develop their international and global understanding, intercultural awareness and lifelong learning skills across a program of study. Senior academic leaders in programs of study who make critical decisions about content, employ teachers and induct them into the culture of the School and Program are key actors in the process of IoC.

University leaders, academic leaders, teachers, educational developers, learning designers and students all have important roles to play in ensuring that IoC is implemented in purposeful ways, toward shared intercultural and international goals. IoC is a long-term project, a cyclical process of review and reflection leading to continuous quality improvement, with no defined endpoint.

Active, purposeful engagement of academic staff and students in the process of internationalisation is critical to its success. Best practice will include ongoing professional learning programs for staff and a range of programs in the formal and informal curriculum that support students to develop critical intercultural skills.

Principle 3: Systematic, system-wide and context relevant

An institutional culture that openly and obviously values diversity in its actions as well as its words is an important indicator of best practice consistent with this principle. An institutional culture is built by people (what they say and what they do) and the systems that support them. Institutional leaders should strive to create a culture that openly and obviously values diversity and has a clear *educational* purpose for internationalization of relevance to all students.

The internationalization of higher education is important beyond the academy over the long term. Decolonization of the curriculum and cognitive justice require system-wide changes in thinking and complex conversations within higher education institutions, but also nationally and globally. The process of internationalization of the curriculum provides the space for conversations about these concepts within institutions. Tolerance of diversity and difference, and interdisciplinary encounters are not enough. What is needed is a plurality that actively recognizes the need for diversity and a dialog of knowledges at the local and the global level. In some contexts that will mean engagement with Indigenous people and their knowledge systems, in others it may mean something different.

Higher education institutions have a responsibility to commence and sustain complicated conversations locally, internationally and globally.

Conclusion

The curriculum has the potential not only to give learners access to “different worlds” (Goodson, 1995, p. 27) but also to restrict learners if too narrowly focused. It is the means by which institutions can reach all students and make a meaningful contribution to society by ensuring that the students of today graduate ready and willing to make a positive difference in the world of tomorrow. Internationalization of the curriculum, founded on education and social theories, offers opportunities for the development of graduates as human, social and economic beings committed to the global common good. However, achieving this requires short-term, medium-term and long-term thinking and action on the part of institutional leaders as well as those involved in planning and teaching the curriculum. In the short term, a commitment to thinking differently and imagining new possibilities and approaches to internationalization for both personal and common good are necessary. In the medium term, actions focused on reviewing and rethinking policies, strategies and taken-for-granted approaches to internationalization, including whose knowledge counts in the curriculum, will be necessary to ensure alignment between vision, values and strategy. In the short, medium and long term there are many actions that can be taken to ensure internationalization of the curriculum improves the quality of education for all students and makes a positive contribution to local and global communities now and in the future.

For around 25 years internationalization of the curriculum has been recognized as an essential component of higher education. In this time both the concept and practice have evolved, in response to changes in the context within which universities work and informed by research and scholarly practice in the field. Changes have included increasing recognition of the need for the knowledge base of the curriculum to be diversified and pedagogies and practices to be more inclusive and equitable. The focus has moved to learning outcomes for all students and interculturalizing as well as internationalizing the different dimensions of the curriculum. A range of theoretical frameworks underpin the concept and practice of internationalization of the curriculum in general terms and within and across disciplines. To some extent practice has depended on how deep the institution, individual professors, teachers, staff and students have wanted to go. For some students and staff the result has been transformational, but for most it has not and there is still much work to do to ensure that curricula are inclusive and critical in theory and practice.

Perhaps the biggest change has been increasing recognition that connecting the purpose of internationalization with the notion of developing graduates who can make a “meaningful” contribution to society requires rethinking traditional approaches to internationalizing the curriculum focused on mobility programmes, teaching in English or international student recruitment.

With reference to extant literature and practice in different parts of the world, this chapter has described six characteristics of a paradigm of *holistic and inclusive internationalization of the curriculum*. This paradigm has the potential to significantly increase impact, open up new worlds to all students, and through them extend the benefits of internationalization to local and global communities. An extended and revised description of internationalization of the curriculum and three principles of best practice

have been discussed to facilitate the exploration of this new paradigm of internationalization in diverse contexts. The implications for best practice of this modern paradigm of internationalization of the curriculum have been summarized as three principles, that can be applied in different contexts.

Best practice principles provide flexibility and guidance without being prescriptive. They are useful for those seeking to reimagine internationalization of the curriculum. Ultimately, the pursuit of best practice in internationalization of the curriculum is an ongoing process for institutions everywhere as contexts shift and change, and new ideas and approaches emerge.

A version of this article was published in Spanish as 'Reimaginar las mejores prácticas en la internacionalización del currículo capítulo introductorio'. In Leask, B., Torres-Hernandez, A.M., Bustos-Aguirre, M. L. & de Wit, H. (eds). (2022) *Reimaginar La Internacionalización Del Currículo: Mejores prácticas y posibilidades prometedoras*, pp. 21-58. Universidad de Guadalajara.

References

- Appiah, K., 2006. *Cosmopolitanism: Ethics in a World of Strangers*. W.W. Norton, New York.
- AUCC, 2009. *Internationalisation of the Curriculum: A Practical Guide to Support Canadian Universities' Efforts*. Association of Universities and Colleges of Canada, Ottawa, Canada.
- Barnett, R., 2000. *Realising the University in an Age of Supercomplexity*. The Society for Higher Education and OUP, Ballmoor, Bucks.
- Bartell, M., 2003. Internationalisation of universities: a university culture-based framework. *High Educ.* 45 (1), 43–70.
- Beelen, J., Jones, E., 2015a. Redefining internationalisation at home. In: Curai, A., Matei, L., Pricopie, R., Salmi, J., Scott, P. (Eds.), *The European Higher Education Area: Between Critical Reflections and Future Policies*. Springer, Dordrecht, pp. 67–80.
- Beelen, J., Jones, E., 2015b. Looking back at 15 years of internationalisation at home. In: Forum, Winter 2015. European Association for International Education, Amsterdam, pp. 6–8.
- Breit, R., Obijiofor, L., Fitzgerald, R., 2013. Internationalisation as de-westernization of the curriculum: the case of journalism at an Australian University. *J. Stud. Int. Educ.* 17 (2), 119–135.
- Byram, M., 1997. *Teaching and Assessing Intercultural Communicative Competence*. Multilingual Matters, Clevedon.
- Carroll, J., 2015. *Tools for Teaching in an Educationally Mobile World*. Routledge.
- Clifford, V., 2009. Engaging the disciplines in internationalising the curriculum. *Int. J. Acad. Dev.* 14 (2), 133–143.
- Clifford, V., Montgomery, C., 2014. Challenging conceptions of western higher education and promoting graduates as global citizens. *High Educ. Q.* 68 (1), 28–45. <https://doi.org/10.1111/hequ.12029>.
- Cook-Sather, A., Bovill, C., Felten, P., 2014. *Engaging Students as Partners in Learning and Teaching: A Guide for Faculty*. Jossey-Bass, San Francisco.
- Deardorff, D., 2006. Identification and assessment of intercultural competence as a student outcome of internationalization. *J. Stud. Int. Educ.* 10 (3), 241–266.
- Deardorff, D., Jones, E., 2012. Intercultural competence: an emerging focus in international higher education. In: Deardorff, D., de Wit, H., Heyl, J., Adams, T. (Eds.), *The SAGE Handbook of International Education*. SAGE, California, pp. 283–303.
- De Louw, E., Bulnes, C., 2021. Interculturality: making global education work in local contexts. In: Nixon, P., Dennen, V., Rawal, R. (Eds.), *Reshaping International Teaching and Learning in Higher Education: Universities in the Information Age*. Routledge.
- De Wit, H., Leask, B., 2017. Re-Imagining the curriculum for the 21st century. In: Grau, F.X., Escrigas, C., Goddard, J., Hall, B., Hazelkorn, E., Tandon, R. (Eds.), *Higher Education in the World 6: Towards a Socially Responsible University: Balancing the Global with the Local*. Global University Network for Innovation (GUNi), pp. 222–235. <http://www.guninetwork.org/report/higher-education-world-6>.
- De Wit, H., Hunter, F., Egron-Polak, E., Howard, L. (Eds.), 2015. *Internationalisation of Higher Education. A Study for the European Parliament*. European Parliament, Brussels.
- Escrigas, C., Sanchez, J., Hall, B., Tandon, R., 2014. Editors' introduction: knowledge, engagement and higher education: contributing to social change. In: *Higher Education in the World 5: Knowledge, Engagement and Higher Education: Contributing to Social Change*. Palgrave Macmillan, pp. xxxi–xxxix.
- Freeman, M., Treleaven, L., Simpson, L., Ridings, S., Ramburuth, P., Leask, B., Caulfield, N., Sykes, C., 2009. *Embedding the Development of Intercultural Competence in Business Education*. ALTC Project Report. Available from: <http://hdl.voced.edu.au/10707/279607>.
- Goodson, I., 1995. *The Making of Curriculum: Collected Essays*, second ed. Falmer, London.
- Gothard, J., Downey, G., Gray, T., 2012. Bringing the Learning Home: Re-entry Programs to Enhance Study Abroad Programs in Australian Universities. <http://www.olt.gov.au/project-bringing-learning-home-murdoch-2010>.
- Green, W., 2018. Engaging 'students as partners' in global learning: some possibilities and provocations. *J. Stud. Int. Educ.* 24 (1), 1–20.
- Green, W., Whitsed, C., 2015. Introducing critical perspectives on internationalising the curriculum. In: Green, W., Whitsed, C. (Eds.), *Critical Perspectives on Internationalising the Curriculum in the Disciplines: Reflective narrative accounts from Business, Education and Health*. Sense Publishers, Amsterdam, pp. 3–22. Chapter 1.
- Gregersen-Hermans, J., 2021. Toward a curriculum for the future: synthesizing education for sustainable development and internationalisation of the curriculum. *J. Stud. Int. Educ.* 25 (4), 461–481. <https://doi.org/10.1177/10283153211031033>.
- Hack, K., May 26, 2020. Decolonisation of the Curriculum - A Conversation. *Advance HE*. <https://www.advance-he.ac.uk/news-and-views/decolonisation-curriculum-conversation>.
- Harari, M., 1992. The internationalisation of the curriculum. In: Klasek, C. (Ed.), *Bridges to the Future: Strategies for Internationalising Higher Education*. Association of International Education Administrators, Carbondale, IL, pp. 52–79.
- Hicks, O., 2018. Curriculum in higher education: confusion, complexity and currency. *HERDSA Rev. High. Educ.* 5. www.hersa.org.au/herdsa-review-higher-education-vol-5/5-30.
- Hudzik, J., 2011. *Comprehensive Internationalization: From Concept to Action*. NAFSA: Association of International Educators, Washington, District of Columbia.
- Jiang, X., 2011. Why interculturalization? A neo-Marxist approach to accommodate cultural diversity in higher education. *Educ. Philos. Theor.* 43 (4), 387–399.
- Jones, E., 2013. Internationalisation and employability: the role of intercultural experiences in the development of transferable skills. *Publ. Money Manag.* 33 (2), 95–104.
- Jones, E., 2016. Mobility, graduate employability and local internationalisation. In: Jones, E., Coelen, R., Beelen, J., de Wit, H. (Eds.), *Global and Local Internationalization*. Sense Publishers, pp. 107–116.
- Jones, E., 2019. Is it time to pursue "Interculturalisation"? *High Educ.* <https://www.timeshighereducation.com/opinion/it-time-pursue-interculturalisation>.
- Killick, D., 2015. *Developing the Global Student – Higher Education in an Era of Globalization*. Routledge, Abingdon.
- Knight, J., 2004. Internationalisation remodeled: definitions, approaches and rationales. *J. Stud. Int. Educ.* 8, 5–31.
- Killick, 2017. *Internationalisation and Diversity in Higher Education: Implications for Teaching, Learning and Assessment*. Palgrave Macmillan.
- Killick, D., Foster, M., 2021. *Learner Relationships in Global Higher Education: Critical Pedagogy in a Multicultural World*. Routledge.
- La Trobe University, 2016. *Wominjeka La Trobe*. Melbourne, Australia. Retrieved from: <https://www.latrobe.edu.au/students/subjects/current/abs0wom-wominjeka-la-trobe>.
- Leask, B., 2001. Bridging the gap: internationalising university curricula. *J. Stud. Int. Educ.* 5 (2), 100–115. <http://jsi.sagepub.com/cgi/content/abstract/5/2/100>.
- Leask, B., 2009. Using formal and informal curricula to improve interactions between home and international students. *J. Stud. Int. Educ.* 13 (2), 205–221. <http://jsi.sagepub.com/cgi/content/abstract/13/2/205> (1552-7808).
- Leask, B., 2013a. Internationalising the curriculum in the disciplines – imagining new possibilities. *J. Stud. Int. Educ.* 17 (2), 103–118. <https://doi.org/10.1177/1028315312475090>. <http://jsi.sagepub.com/content/17/2/103>.

- Leask, B., 2013b. Internationalisation of the curriculum and the disciplines – current perspectives and directions for the future. *J. Stud. Int. Educ.* 17 (2), 99–102. <https://doi.org/10.1177/1028315313486228>. <http://jsi.sagepub.com/content/17/2/99>.
- Leask, B., 2015. *Internationalising the Curriculum*. Routledge, Abingdon.
- Leask, B., Carroll, J., 2011. Moving beyond 'wishing and hoping': internationalisation and student experiences of inclusion and engagement. *High. Educ. Res. Dev.* 30 (5), 647–659. Special Issue: Internationalising the Home Student.
- Leask, B., Carroll, J., 2013. *Learning and Teaching across Cultures: Good Practice Principles and Quick Guides*. IEAA, Melbourne, Australia.
- Leask, B., Green, W., 2020. Curriculum integration: maximizing the impact of education abroad for all students (Chapter 11). In: Ogden, A.C., Streitwieser, B., Van Mol, C. (Eds.), *Education Abroad: Bridging Scholarship and Practice*. Routledge, pp. 174–183.
- Lilley, K., Barker, M., Harris, N., 2014. Exploring the process of global citizen learning and the student mindset. *J. Stud. Int. High. Educ.* 19 (3), 225–245.
- Macquarie University, 2021. Global Leadership Program. Macquarie University, Australia. <https://www.mq.edu.au/partner/contribute-to-education/global-leadership-program>.
- Majee, U.S., Ress, S.B., 2020. Colonial legacies in internationalisation of higher education: racial justice and geopolitical redress in South Africa and Brazil. *Compare* 50 (4), 463–481.
- Marginson, S., 2011. Higher education and public good. *High Educ. Q.* 65 (4), 411–433. <https://doi.org/10.1111/j.1468-2273.2011.00496.x>.
- Margolis, E., Soldatenko, M., Acker, S., Gair, M., 2001. Peekaboo: hiding and outing the curriculum. In: Margolis, M. (Ed.), *The Hidden Curriculum in Higher Education*. Routledge, pp. 1–20.
- Marrone, Taylor, Hammerle, 2018. Do international students appreciate active learning in lectures?. In: *Research on Educational Technologies International Students & Active Learning*, vol. 21. Available from: https://www.researchgate.net/publication/323805969_Do_International_Students_Appreciate_Active_Learning_in_Lectures. (Accessed 17 April 2021).
- Mestenhauser, J., 1998. Portraits of an international curriculum: an uncommon multidimensional perspective. In: Mestenhauser, J., Ellingboe, B. (Eds.), *Reforming the Higher Education Curriculum: Internationalising the Campus*. Oryx Press, Phoenix, Arizona, pp. 3–38.
- McArthur, J., 2013. *Rethinking Knowledge within Higher Education: Adorno and Social Justice*. Bloomsbury, London.
- Montgomery, C., 2019. Surfacing "southern" perspectives on student engagement with internationalization: doctoral theses as alternative forms of knowledge. *J. Stud. Int. Educ.* 23 (1), 123–138.
- Nussbaum, M., 2010. *Not for Profit: Why Democracy Needs the Humanities*. Princeton University Press, Princeton, NJ.
- Ogude, N., 2007. Internationalising the curriculum. In: Jooste, N., Neale-Shutte, M. (Eds.), *Volume 2: Internationalising the Curriculum*. Nelson Mandela University, Port Elizabeth.
- Paige, R., Jorstad, J., Siaya, L., Klein, F., Colby, J., 2003. Culture learning in language education: a review of the literature. In: Lange, D., Paige, R. (Eds.), *Culture as the Core: Integrating Culture into the Language Education*. Information Age, Greenwich, CT, pp. 173–236.
- Queen's University, 2022. https://www.queensu.ca/teachingandlearning/modules/globalization/32_sec_04_interculturalizing_curriculum.html.
- Rizvi, F., 2009. Towards cosmopolitan learning. *Stud. Cult. Polit. Educ.* 30 (3), 253–268.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, Abingdon, UK.
- Robson, S., 2011. Internationalization: a transformative agenda for higher education? *Teach. Teach.* 17 (6), 619–630.
- Schattle, H., 2009. Global citizenship in theory and practice. In: Lewin, R. (Ed.), *The Handbook of Practice and Research in Study Abroad: Higher Education and the Quest for Global Citizenship*. Routledge, New York, pp. 3–30.
- Stein, S., Andreotti, V., 2016. Decolonization and higher education. In: Peters, M. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Science + Business Media, Singapore. https://doi.org/10.1007/978-981-287-532-7_479-1.
- Stohl, M., 2007. We have met the enemy and he is us: the role of the faculty in the internationalization of higher education in the coming decade. *J. Stud. Int. Educ.* 11 (3/4), 359–372.
- Thompson Rivers University, 2022. *Interculturalizing the Curriculum*. Faculty PD Program. <https://www.tru.ca/intercultural/faculty-staff/interculturalizing-curriculum.html>.
- Townsin, L., Walsh, C., 2016. A new border pedagogy: rethinking outbound mobility programs in the Asian Century. In: Valliarus, D., Coleman-George, D. (Eds.), *Handbook of Research on Study Abroad Programs and Outbound Mobility*. IGI Global, Pennsylvania, USA, pp. 215–247.
- Visvanathan, S., 2021. 'The Search for Cognitive Justice' in Dialogues of Difference Association of Commonwealth Universities. <https://www.acu.ac.uk/the-acu-review/the-search-for-cognitive-justice/>.
- Weber-Bosley, G., 2010. Beyond immersion: global engagement and transformation through intervention via student reflection in long-term study abroad. In: Jones, E. (Ed.), *Internationalisation and the Student Voice*. Routledge, New York.
- Zou, T.X.P., Chu, B.C.B., Law, L.Y.N., Lin, V., Ko, T., Yu, M., Mok, P.Y.C., 2020. University teachers' conceptions of internationalisation of the curriculum: a phenomenographic study. *High Educ.* 61 (1). <https://doi.org/10.1007/s10734-019-00461-w>.

International higher education and the commodification of student mobility

Fazal Rizvi, University of Melbourne Australia, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	233
Processes of commodification	234
Selling mobility for higher education	235
Understanding student demand	236
New class formations	237
Beyond the pandemic	238
References	239

Introduction

A focus on the global mobility of students has seemingly become a permanent feature of higher education institutions (HEIs). In high-income countries, the ability to attract globally mobile students is now regarded as a marker of status and prestige of HEIs, as well as a major source of their income. Not surprisingly therefore HEIs have developed complex operational systems to recruit international students, underpinned by an ideological discourse that regards internationalization of higher education to be decisive in preparing students to meet the requirements of the global economy. The efforts of HEIs have enjoyed great success, resulting in rapid growth in the number of internationally mobile students in higher education over the past four decades. The number of international students enrolled in a program at a higher education institution abroad from which they expect to receive a formal qualification grew from 2 million in 2000 to 5.3 million in 2017 (UIS, 2019). Once the study abroad and exchange students are added to this figure, the number of globally mobile students could have approached 10 million. Before Covid-19 complicated the trend, rising annual rates of growth of international students were simply assumed in the planning processes in many HEIs.

Although such numbers were unimaginable barely 40 years ago, there is however nothing entirely new about the global mobility of students and scholars. Scholars have always traveled to centers of higher learning beyond their own communities in search of new knowledge. The global circulation and exchange of knowledge across cultures was an important aspect of academic life in the ancient academies of India, China, and the Middle East (Guruz, 2008). During the colonial era, student mobility across borders became linked to the logic of colonialism. It was assumed that the graduates of European universities would return to the colonies, appropriately socialized in “modern dispositions,” indebted to their colonial masters (Seth, 2007). They would become part of the administrative apparatus of the colonial state.

When the colonial era came to an end, student mobility became associated with the projects of nation-building of the newly independent countries. Aid Programs such as the Fulbright Program in the United States and the Colombo Plan in British Commonwealth countries—and similar programs in the Soviet Union—provided talented students scholarships to acquire advanced technical, scientific, and administrative training at universities abroad. The so-called “developed” countries saw it as their moral responsibility to assist the poorer nations that they had once occupied, helping them to realize their national aspirations of social and economic development. Such programs of educational aid however simultaneously served the strategic interests of the economically advanced countries within the broader geopolitics of the Cold War. Global mobility of students thus became a major instrument of public diplomacy, an exercise in “soft power” (Nye, 2005).

Since the mid-1980s, the logic of student mobility has once again acquired a new logic. While the older discourses of foreign aid and public diplomacy have not entirely vanished, the new understanding of student mobility is now tied to commercial concerns. It is increasingly viewed as a source of revenue for HEIs, which, under fiscal pressures of their own, no longer feel obligated to financially support international students from the “developing countries.” Nor are governments any longer reluctant to reduce the number of educational scholarships they offer in the form of foreign aid (Harman, 2004). Instead, international students are viewed as a potential source of revenue. To help them generate income, governments now help universities with a complex administrative technology to recruit fee-paying international students, so long as they meet basic entry requirements. Jokila et al. (2019), for example, show how the Finnish government has sought to convert student mobility into an opportunity to resolve the financial crisis facing its systems of education.

Global mobility of students has thus become commodified, subject to selling and buying. Like tourism, international education is now regarded as a service, an export commodity. Commodification, however, is not an end-state itself but involves a range of dynamic complex processes, including strategies of accumulation, transforming students into consumers, and making market forces relevant to the governance of higher education. In this paper, I use the case of Australian higher education to illustrate how these processes of commodification associated with the global mobility of students have been developed and involve a particular discourse to justify it, as well as a range of commercial practices to sell student mobility and cultivate a culture of demand. I then explore how such commodification has not only transformed the traditional character of higher education but is also performing a role in the formation of a new transnational elite class. At the same time, I argue how the logic underlying this mode of thinking

about global mobility of students has been unsettled by the pandemic, opening the possibilities of radically reimagining international higher education. These possibilities, I argue however, will remain constrained unless ways are found to tame the relentless commodification processes.

Processes of commodification

According to Leys (2012), “the dominant process underlying the transformation of life in all societies, since at least the mid-nineteenth century is the conversion of things and activities into commodities or commodification.” The idea of commodification thus refers to the transformation of social relationships formerly unaffected by the considerations of commerce into business relationships, relationships of exchange, of buying and selling. An ever-increasing number of goods that were once owned by the public have, under modern practices of capitalism, been commodified. While goods have always been commodified, in a globalizing economy, the rate at which services are also being commodified has also been remarkable (Roberts, 2015). In recent decades, services such as health and education have been broken up into commodities through which profits are now derived.

It was Karl Marx who first provided a most comprehensive account of the processes of commodification in the mid-nineteenth century. Commodity, he argued, is “an external object, a thing which through its qualities satisfies human needs of whatever kind” (Marx, 2011, section 125) but which can be exchanged for something else, namely money. Marx believed that a commodity has both use and exchange value, that is, it can be used both to satisfy our needs and potentially have monetary value. It can thus be subordinated to the logic of capitalism since it could be exchanged in the market. A commodified service has value, which implies that it represents a quantity of human labor. Under the conditions of a non-commodity economy, Marx suggested, the general human labor does not manifest itself as value, which means that there is no basic contradiction between use-value and exchange-value, concrete labor and abstract labor, social labor, and individual labor. Under capitalism, however, all labor can be converted into monetary value. A factory worker, for example, need no longer be seen as a human being but as an entry in a business ledger, Marx noted, turning human relations into economic transactions.

Since Marx made these powerful observations, ever-increasing aspects of human activity have been converted into commodities. Marxist thinking shows how capitalism expands by commodifying social relations, a process whereby commodity production replaces pre- or non-capitalist forms of production. Capitalism does not, however, only commodify products it also seeks to extend the market reach of those products, potentially around the world. Under changing conditions, it finds new opportunities, while adjusting itself to the particularities of the local contexts. For example, historically, colonialism enabled the reach of capitalism to be expanded across the globe, converting local goods and services that were never subject to commercial exchange into commodities, developing new cultural tastes in the colonies and establishing new circuits of ownership, production, distribution, and consumption (Pieterse, 2010). At the same time, automation has enabled capitalism to create new products for sale. Developments in means of transport have over the decades also opened opportunities for goods to be transported across vast distances.

What these examples demonstrate is that the processes of commodification are dynamic and change in line with shifting conditions. So, for example, as public sentiments and state regulations change, it becomes possible for many services that once belonged to the realm of social responsibilities and obligations, such as age care, to be turned into commodities. Over the past few decades, new practices of commodification have become possible within the new transnational phase in the development of the global capitalist system. Driven largely by technological revolutions, particularly in communication, transportation, marketing and management, globalization has created the material basis of a more extensive globally networked society, marked by transnational cultural and political processes, opening the possibilities of the global integration of social life. As modes of accumulation have become globalized, national production systems are fragmented, making economic growth increasingly dependent on globalized circuits of production and consumption, extending the processes of commodification. As Robinson (2004, p. 6) has noted, globalization has greatly “extended outwards to areas that previously were outside the system of commodity production”; and as it has deepened, it has ensured that the “human activities that previously were outside the logic of capitalist production are brought into this logic.”

However, such a march toward commodification has not occurred in any uniform and universal manner but has involved diverse modes of transformation, both discursive and material, through the intervention, for example, of states and corporations and international organizations. Despite this diversity, contemporary forms of commodification have largely been underpinned by a particular understanding of human activities, relationships, and institutions. This understanding is widely described as “neoliberalism” (Rizvi and Lingard, 2010). It refers to a set of loosely connected notions that celebrate the role of the market in reshaping society, with an emphasis on individual enterprise. Focusing on the importance of competition, it seeks to minimize the role of the state, emphasizing policies of deregulation, privatization and corporatization. The hegemonic dominance of these notions has led Brown (2015, p. 36) to refer to it as “neoliberal rationality,” a mode of thinking that assumes that human beings are above all market actors. In neoliberal reason, she insists, “competition replaces exchange as the market’s root principle and basic good.”

While the ideas associated with neoliberal rationality are not entirely consistent and have been articulated and justified in a range of different ways, what cannot be doubted is that as an ideology it has become enormously influential globally (Harvey, 2005). It has moreover shaped the processes of commodification, giving them their contemporary form. Within neoliberal rationality, it has become possible to create new products for sale, justify attempts to commodify traditionally public goods, and encourage the search of new markets both within and across national borders. Arguably, global mobility of students is one such example of commodification, promoted through a range of policy initiatives and institutional practices intended to transform the ways in which students themselves begin to think about their mobility as an economic investment from which they can expect to gain tangible monetary

and other benefits. Such a transformation of their subjectivity is accomplished by shifting discursive claims about the purposes and governance of higher education. In what follows, focusing on the illustrative case of Australia, I discuss the ways in which global mobility of students has been commodified through the policies and practices of not only HEIs but also of governments, professional associations, international organizations such as the OECD, and various other actors who have been active in the business of encouraging, recruiting and supporting students to study abroad.

Selling mobility for higher education

Until the mid-1980s, higher education in Australia was widely regarded as a public good (Marginson, 2011). The Australian Government met most of the costs of running universities and the students were not required to pay tuition fees. In 1984, however, the government took a range of highly consequential decisions to “reform” the Australian economy by opening it to global competition. It reduced tariffs and floated the Australian dollar, for example. While these economic policies did not refer specifically to higher education, they had a major impact on reimagining its purposes and governance, justifying an end to the era of free tuition, asking HEIs to consider ways of generating income to grow their operations, and encouraging them to become globally competitive. Within the context of this drift toward neoliberalism, an intense debate took place about the continuing need to support the global mobility of students through various programs of educational aid. In 1986, the Government published two major reports with contrasting recommendations. One of these, the *Review the Australian Overseas Aid Program* (the Jackson Report) supported the “trade” view of higher education, while the *Review of Private Overseas Student Policy* (the Goldring Report) rejected the need to shift to a market-based approach. In the end, the Jackson view prevailed, sowing the seeds of a new way of thinking about international students (Rizvi, 2011).

Over the next decade, Australian universities embraced this shift from “aid” to “trade” in higher education services more enthusiastically than was expected. They began to consider ways of recruiting international students, expanding their global reach, in search of additional revenues to manage the declining fiscal support from the public purse. Effectively, they rearticulated the idea of international education in market terms. They did not however quite abandon the older rationales for internationalization (Knight, 2004), which had stressed the role of higher education in assisting the developing countries toward their “modernization,” helping them to develop social and economically through capacity-building, and in promoting international understanding and intercultural relations. These notions were instead incorporated within the broader canons of neoliberal rationality (see Lipura and Collins, 2020). This rationality implied a new discourse of international higher education, expressed in such policy imperatives as revenue generation and institutional growth as a way of building profile and reputation and diversifying the campus. In relation to curriculum, it suggested the need to develop human resources for a globalizing economy and, with respect to research it stressed the importance of the commercialization of knowledge.

In Australia and elsewhere, this view of internationalization has now become part of a new “commonsense” (Apple, 2000). It interprets the global mobility of students in commodified terms, involving two contrasting lines of thinking: one focused on the need to integrate an “international perspective” into the primary functions of teaching, research and service, and to promote international activities for “mutually beneficial relationship,” and the other pointing to the opportunities that now exist to monetize international student mobility, enabling the formation of “an export industry” in which universities compete for students, funds and status. It is the latter view that has transformed the core character of higher education around the world, unleashing a culture of entrepreneurialism that would have been inconceivable in the mid-1980s. International student recruitment has now become an increasingly dominant feature of the governance of HEIs, with much of their decision-making filtered through the logic of commodity exchange.

As an industry, the idea of internationalization has spawned an administrative technology of recruitment, with certain rules of operation based on an expertise that incorporates knowledge of market segments and specificities, as well as a symbolic language that highlights the multiple benefits of internationalization. Globally, an aligned industry has emerged around the testing of English to meet enrollment requirements, as have Academic English preparation courses, often run by the universities themselves with financial returns as their major motivation (Kedzierski, 2016). Developed also are targeted advertising programs conducted through the media, as well as through educational Expos and market-orientated conferences. While other aspects of internationalization, such as teaching and learning, are not entirely overlooked, market concerns disproportionately attract the attention of senior university administrators, as they struggle to balance their budgets within the context of declining public funds. Indeed, the success of HEIs is often measured in terms of the number of fee-paying international students they are able to recruit. Their financial success is celebrated by government agencies and various ranking systems alike (Hazelkorn, 2018).

To extend their market reach, universities now enter into complex articulation and twinning arrangements with overseas educational providers to ensure a steady flow of in-coming students. A highly complex system now exists in the use of local recruitment agents who are often the first point of contact between potential students and a university. Also established is a vast array of transnational programs in which academics and support staff travel far and wide (Zigurus and McBurnie, 2006). In the image of other service industries, various franchise arrangements are also developed with varying degrees of success. Several Australian, British, and American universities have now also established campuses in strategically advantageous locations, partly to promote brand recognition, but also to expand their commercial operations (McBurnie, 2015). Kondakci et al. (2018) have used social network analysis to demonstrate the emergence of “regional hubs” around which international student mobility is increasingly structured.

Globally, these initiatives have resulted in a spectacular growth in the number of globally mobile students, with shifting patterns of mobility over the past three decades (Hou and Du, 2020). However, this growth could not have been achieved without the role played by governments whose policy settings are highly supportive of entrepreneurial activities, often allowing recruitment practices to take place through their diplomatic missions. While the work of governments in casting higher education in trade terms has been crucial, international organizations, such as the OECD, too have played an important role in legitimizing a market discourse in higher education, well beyond a view of mobility that was already familiar to them through such programs as Erasmus and various other scholarship schemes. Governments have also participated actively in negotiations over the General Agreement in Trade of Services (GATS) at the World Trade Organization (WTO), aimed at determining a globally agreed set of rules for trade in educational services (Sidhu, 2007). In Australia, a close nexus has also been forged between student mobility for higher education and immigration policies (Robertson, 2013), which other countries such as Canada and the United Kingdom have sought to emulate, within the context of what Mok et al. (2020) refer to as the “quest for global talent.”

Understanding student demand

These initiatives have enjoyed considerable success in institutionalizing a commodified view of the global mobility of students. Recognizing the potential, HEIs have created an administrative technology of recruitment designed to build supply, but, as is the case with other service industries, these technologies have also had the effect of building demand, encouraging potential students to view themselves as consumers able to imagine the various benefits of higher education abroad. Traditionally, the motivations of international students were quite diverse. They included such factors as the lack of opportunities at home; perceptions of better curriculum and pedagogy; the prestige associated with international education; following the family tradition and social networks; interest in travel and a cosmopolitan life; greater freedom and independence abroad; possibilities of immigration or permanent residence; and so on. The new neoliberal logic consists in several additional, and perhaps more weighty, factors, most notably those relating to the returns on educational investment and better employment prospects, as well as certain beliefs about the value of international education in the global labor market. In this way, it is not only HEIs that have embraced the neoliberal imaginary, most international students have too, with their shifting subjectivities (Mathews and Sidhu, 2005).

As the number of international students on their campuses has increased, it has made Australian HEIs more attractive to potential applicants. In a highly mediatized space, stories of agreeable and meaningful experiences have been circulated by HEIs and their recruitment agents, often endorsed by students themselves. Alumni networks have also been cultivated to promote the many benefits of international education. As HEIs have become culturally diversified their appeal to students has also intensified. These factors have had the effect of greatly increasing demand for global mobility. The perceptions that international graduates have a better chance of obtaining prestigious jobs in multinational corporations have also contributed to building demand, as has the growing attraction of cosmopolitan lifestyles, most notably in Asia’s rapidly increasing middle class, where international education has become a marker of social status and prestige. These supply and demand factors have thus been drawn into synergy in support of the processes of commodification.

In the 1990s, Australia was a global trendsetter in prompting this commodified view of global mobility. However, its success has been emulated by other systems of higher education. Recognizing the shifting motivations of international students, they too have begun to embrace a similar discourse about the importance of global mobility, focusing similarly on the commercial opportunities associated with “the export of higher education.” To attract international students, universities in Germany, France and Netherlands have, for example, developed specific academic programs in high demand areas such as Business and Technology. Countries such as Egypt, South Africa and Singapore also have sought to become major regional hubs of commercialized international education (de Wit et al., 2017). Convinced that English is now the *lingua franca* of the global economy, most of these universities have begun to offer courses in English to attract international students. Much of the focus of their recruitment efforts has been on Asia, based on the recognition of strong desire for cosmopolitan experiences that now exists in its rapidly growing middle class.

This desire has indubitably served to create and strengthen what Sklair (2001) refers to as a “transnational elite class.” International students and their families are arguably either drawn from this class or seek to join it. In Asia, for example, many wealthy parents often write a script for their children’s future education as soon as they are born. This script now includes higher education abroad. The rapid rise in the number of international schools in Asia, which have as one of their principal goals the preparation of students for tertiary education abroad, reflects this pattern of desire for international education (Kenway et al., 2017). The growing popularity of the International Baccalaureate (IB) as an entrance examination for accessing high-status universities abroad is also another manifestation of this trend. This shows how while the supply of international higher education has increased rapidly so has the demand.

It is important however to understand the changing nature of this demand. Like other markets, the global market in higher education is not static but has to respond to the shifting perceptions, as well as consumer concerns. There are, for example, some legitimate concerns in Australia about issues of quality, capacity and support provided to international students (see for example, Marginson et al., 2010). There is also considerable evidence of exploitative practices within international education markets that are inadequately regulated. It is also clear that the fields of Business Studies, Engineering and Computer Education remain dominant, with more than 60% of international students around the world enrolled in these disciplinary areas. English appears to have become institutionalized as the language of international education, confirming the knowledge asymmetry that had always existed, but under a market orientation has been reinforced. Furthermore, a large proportion of students appear to

have accepted the view that international higher education is a private good, and that, within its logic, they are consumers with some ability to define its content and mode of governance. They view international education as an investment toward a prosperous future in which their cosmopolitan aspirations and experiences are assumed to be major assets.

New class formations

In this way, global mobility of education contributes to the emergence of a new class politics both within and across nation-states. While the earlier forms of global mobility of students also reflected patterns of global inequalities, the idea of international education embedded within neoliberal rationality performs class relations in several new ways, located within a wider transnational space. Retaining the traces of “colonial modernity” (Mignolo and Escobar, 2013), the idea of international higher education is now aligned to new thinking about global conditions, opportunities and aspirations. Theoretically, then, this shows how the recent transformations in higher education allow us to think about contemporary class formations differently. Since these transformations are located within both national and transnational spaces, international education no longer involves attempts solely to produce a national elite, but a “transnational capitalist class” and its various fractions, for example, “globalizing bureaucrats,” “globalizing entrepreneurs” and “globalizing professionals” (Sklair, 2001, p. 7).

Such class fractions emerge through the works of transnational corporations, media and other institutions of consumption, which now include higher education. Sklair (2001, p. 4) has shown how “a new class is emerging and how it pursues people and resources all over the world in its insatiable desire for private profit and external accumulation.” This transnational capitalist class, he has noted, consists of “corporate executives, globalizing bureaucrats and politicians, globalizing professionals and consumerist elites.” Robinson (2011) has similarly maintained that the rise of the new capitalist class not only works through global circuits of investment and production but also through institutions, such as higher education, that are now entirely comfortable ideologically with the relations of class they represent. (Carroll & Sapinski 2010) have mapped the changing fields of power generated by such elite relations in the interests of the world’s largest corporations and related political organizations.

Global mobility of students is located squarely within this field of power. In so far as it is increasingly catering to the needs of students with global aspirations, embedded in the “culture-ideology of consumerism” (Sklair, 2001, p. 4), it is contributing to new class formations, no longer tied to national interests, but to a global system of accumulation. This system has given rise to new borders and practices of class, encouraging cultural practices, classed behaviors and identities located in a transnational space. Bamberger (2020) has shown cosmopolitan and ethnic identity capital is increasingly accumulated through international student mobility. It is attempting to develop in students a sense of themselves and others like them as cosmopolitan citizens and consumers, linked to each other in social and business networks that span the globe. Illustrative of this trend are the ways in which HEIs around the world now promote a particular notion of “global citizenship education,” focused on the development of a range of personal attributes for a successful life in a globalizing economy and culture. This “soft” view of global citizenship assumes the key predicates of neoliberal rationality, obscuring issues of emergent class inequalities, of “whether and how to address the economic and cultural roots of the inequalities in power and wealth/labor distribution in a global complex and uncertain system” (Andriotti, 2006, P. 5).

As compelling as this characterization of the relationship between globalization and new class formations, it is limited in several important respects. In recent years, the complex relationship between globalization and economic liberalization and emerging class politics has been examined by several scholars. In most cases, this examination has been supported by rich ethnographic research. Drawing upon their research in India, Ganguly-Scrase and Scrase (2007), for example, challenge the notion of homogenous class analysis, such as that proposed by Sklair and Robinson. They show that while many in India have clearly benefitted greatly from the country’s drift toward neoliberal economic policies, and readily pursue the benefits of commodified international higher education, many others among the Indian elite are skeptical of the advantages secured from these reforms. They are deeply conscious of the contradictions associated with the neoliberal rationality that defines India’s liberalization program (Dreze and Sen, 2015), which includes a commitment to the internationalization of its systems of higher education, encouraging the recruitment of fee-paying international students.

What is clear is that the contemporary policies and practices around global mobility of students have generated a range of complex and social tensions, transforming the traditional view of citizens into consumers of education, amid the complex interactions between culture and political economy. This has transformed the dynamics of sociality and has shifted the nature of the politics of higher education as it relates to the processes of class formation. According to Hsin-Huang (2018), the upwardly mobile middle classes throughout Asia increasingly look to the transnational space in which to locate their identities, giving expression to new social formations and aspirations, modes of consumption and ways of being in contemporary urban space, even when their global aspirations are filtered through a deep commitment to local religious, linguistic, and cultural traditions. Systems of higher education are not unaware of this tension and have sought to modify their discourses and practices accordingly, while continuing to view international higher education in market terms.

Traditionally, descriptions of social class have either specified social stratification located within the borders of a nation-state or presented it in terms of the global relations of power, as is the case with Sklair’s analysis of the transnational capitalist class. This binary is misleading. Student aspirations are clearly affected by global forces, while their global experiences and connections potentially reshape their class imaginations and aspirations. However, these shifts are not entirely disconnected from colonial history and local cultural practices and priorities. For most globally mobile students, national traditions continue to matter as they continue to

interpret their sense of belonging through a national cultural imaginary. At the same time, however, they regard global mobility for higher education as a valued route toward global experiences and economic success within the transnational space.

The local is clearly resilient. The power of global capitalism is not as hegemonic as Sklair and others assume. It is a mistake therefore to essentialize the global as somehow separate, external and hegemonic in the shaping of the local's destiny and character. But nor is it irrelevant in the new formations of class relations. As Massey (1994) argued more than 25 years ago, we need a progressive sense of the local, where boundaries are open and social actors are continuously engaged in constructing their sense of the place, out of both internal factors and influences that are produced in their entanglements with places elsewhere. Global mobility of students meshes local and global processes in producing shifting class relations through the continuing particularities of place despite global influences.

Beyond the pandemic

Such tensions surrounding issues of global mobility of students have become abundantly clear during the pandemic. The assumptions that were once made by HEIs in Australia and elsewhere about the continuing rates of growth in student numbers can no longer be sustained. The COVID crisis hit Australian universities hard, with their projections of growth in student numbers in the Australian Government's (2016) "National Strategy" turned upside down. As the lockdown was announced in March 2020, HEIs moved rapidly to online teaching, while many of their Chinese and other international students returned home. Almost overnight, the global dynamics of international higher education was transformed. It remains unclear as to how long the crisis surrounding COVID 19 might last, what order of decline in the number of international students Australian HEIs should expect, how they should plan for recovery if the bodily mobility of students in previous numbers can no longer be assumed.

In Australia, there is a great deal of talk about recovery from the Covid crisis, what the "post-Covid normal" might look like. What is clear, however, is much of this talk continues to valorize cross-border mobility for the purposes of commercial gain. The value of international students continues to be measured in terms of their contribution to the financial sustainability of universities, rather than the transformative possibilities that the pandemic might have opened. In the public discourse, international education continues to be portrayed as a "successful export industry" that is now threatened. This business model still rejoices in the success that Australian HEIs have had in generating revenue from tuition paid by international students, and hence they fear their financial sustainability without that source of revenue. The talk of recovery does not question the underlying logic of international education, and hence the solutions are presented in a "language of arresting the decline in student numbers, expanding and diversifying the market beyond China, rethinking the technologies of student recruitment, meeting the needs of the student consumer, and putting pressure on the government to change the student visa conditions in order to make Australia a more attractive educational destination" (Rizvi, 2020).

Little effort has been made into rethinking international education in a more radical register, looking at the ways in cross-border mobility of ideas, knowledge and media might nonetheless be possible without the movement of physical bodies, using communication technologies, for example. While the pedagogic possibilities of online learning have been widely examined, these too have been linked to commercial considerations, as a possible substitution for the physical mobility of students. Largely ignored is any consideration of how online pedagogy might help to "overhaul the nature of student engagement and student-teacher relations, how the processes of knowledge ownership, creation, distribution and utilization might be reimagined, and how the idea of learning itself might be reconceptualized" (Rizvi, 2020). Little attention has been paid to the potential of online tools for internationalizing higher education by creating new spaces and cultures of reasoning and understanding. As Kalantzis and Cope (2020) have argued, "online education should not be viewed simply as an alternative delivery system," but as a pedagogic approach that opens up the possibilities of "a new perspectives on, for example, what 'communication' means in the classroom and what it is for."

More broadly, it is abundantly clear that new more critical narratives of global mobility and international education are needed, consistent with the emerging dynamics and possibilities of transnationality beyond the pandemic, in terms that are not solely economic but more importantly ethical and political. It is now increasingly clear that the global context within which international education takes place is characterized by multiple ties and interactions linking people and institutions across the borders of nation-states—not always mediated by inter-national relations, but defined by systems that demand reciprocity and mutual benefit. As higher education systems around the world rethink internationalization, they need to consider the new possibilities of collaboration and networking among institutions dealing with knowledge production and dissemination (Vertovec, 2009)—but in terms other than those that are solely commercial.

An opportunity now exists for the post-pandemic "transnational" dynamics of higher education to challenge the asymmetrical power relations that have resulted in unidirectional flows of students—from the Rest to the West. The changing political architecture of the world demands new ways of thinking about internationalization, especially against a growing assertion of knowledge traditions other than colonial modernity (Stein, 2019) and western rationalism (Connell, 2007), as well as the recognition of the importance of non-economic values, such as those relating to the issues of climate change. At the same time, developments in technology have eroded to a considerable extent the distinction between knowledge production and dissemination and have given rise to new pedagogic possibilities of intercultural communication and learning. Major shifts in youth cultures are accompanied by new practices of global networking, transforming the ways in which internationalization might now be envisaged.

These and other developments have highlighted the importance of transnational collaborations in higher education, ahead of a focus on educational markets. They suggest the possibilities of symmetrical transnational links inherent in the emerging distributive systems of knowledge development and dissemination. They indicate the need to create transnational bilateral and multilateral teaching and research networks among universities and industries, as a way of developing new modes of sharing income, resources and effort. This emphasis on transnational collaborations implies rethinking the purposes and governance of higher education.

This new possibilities on internationalization can only be realized by re-examining the traditional curriculum, challenged by the claims of “other” knowledge traditions, developing new pedagogies that are more responsive to recent innovations in social media and the ubiquitous technologies of communication. Beyond the focus on educational markets, the potential of the idea of internationalization lies in the possibility that HEIs can indeed provide leadership in transnational processes of knowledge production and dissemination in ways that are not only economically efficient but also ethically and socially productive. As Appadurai (2001, p. 5) has noted, this demands “deparochialisation” of the dominant modes of research and knowledge production in universities everywhere. Global mobility of students should remain an important goal of universities around the world, but only if it can be rescued from the clutches of the neoliberal rationality.

References

- Andriotti, V., 2006. Soft versus critical global citizenship education. In: *Development Education: Policy and Practice*, vol. 3.
- Appadurai, A. (Ed.), 2001. *Globalization*. Duke University Press, London UK.
- Apple, M., 2000. The politics of common-sense: why the right is winning. In: Apple, M. (Ed.), *Official Knowledge: Democratic Education in the Conservative Age*. Routledge, London UK.
- Australian Government, 2016. National Strategy for International Education 2025. <https://protectau.mimecast.com/s/qohgCgZolkFAwXQDoF328J6?domain=nsie.education.gov.au>.
- Bamberger, A., 2020. Accumulating cosmopolitan and ethnic identity capital through international student mobility. *Stud. High. Educ.* 45 (7), 1367–1379.
- Brown, W., 2015. *Undoing the Demos: Neoliberalism's Stealth Revolution*. MIT Press, Cambridge MA.
- Connell, R., 2007. *Southern Theory: Social Science and the Global Dynamics of Knowledge*. Polity, Cambridge UK.
- De Wit, H., Gacel-Avila, J., Jones, E. (Eds.), 2017. *Globalization of Internationalization: Emerging Voices and Perspectives*. Routledge, London UK.
- Dreze, J., Sen, A., 2015. *An Uncertain Glory: India and Its Contradictions*. Princeton University Press, Princeton NJ.
- Ganguly-Scrase, R., Scrase, T., 2008. *Globalisation and the Middle Classes in India: The Social and Cultural Impact of Neoliberal Reforms*. Routledge, London UK.
- Gurur, K., 2008. *Higher Education and International Student Mobility in the Global Knowledge Economy*. University of New York Press, Albany NY.
- Harman, G., 2004. New directions in internationalising higher education: Australia's development as an exporter of higher education services. *High. Educ. Pol.* 17, 101–120.
- Harvey, D., 2005. *A Brief History of Neoliberalism*. Oxford University Press, London UK.
- Hazelkorn, E., 2018. Reshaping the world order of higher education: the role and impact of rankings on national and global systems. *Pol. Rev. High. Educ.* 2 (1), 4–31.
- Hou, C., Du, D., 2020. The changing patterns of international student mobility: a network perspective. *J. Ethnic Migrat. Stud.* 1–25.
- Hsin-Huang, M.H., 2018. *Middle Class, Civil Society and Democracy in Asia*. x Routledge, Albany NY.
- Jokila, S., Kallo, J., Mikkilä-Erdmann, M., 2019. From crisis to opportunities: justifying and persuading national policy for international student recruitment. *Eur. J. High. Educ.* 9 (4), 393–411.
- Kalantzis, M., Cope, W., 2020. After the COVID-19 Crisis: Why Higher Education May (And Perhaps Should) Never Be the Same. https://cgscholar.com/community/community_profiles/newlearning/community_updates/114650.
- Kedzierski, M., 2016. English as a medium of instruction in East Asia's higher education sector: a critical realist Cultural Political Economy analysis of underlying logics. *Comp. Educ.* 52 (3), 375–391.
- Kenway, J., Fahey, J., Epstein, D., Koh, A., McCarthy, C., Rizvi, F., 2017. *Class Choreographies: Elite Schools and Globalization*. Palgrave, London.
- Knight, J., 2004. Internationalization remodelled: definitions, approaches and rationales. *J. Stud. Int. Educ.* 8 (1), 5–31.
- Kondakci, Y., Bedenlier, S., Zawacki-Richter, O., 2018. Social network analysis of international student mobility: uncovering the rise of regional hubs. *High. Educ.* 75, 517–535. <https://doi.org/10.1007/s10734-017-0154-9>.
- Leys, C., 2012. Commodification: The Essence of Our Time. *Open Democracy*. <https://www.opendemocracy.net/en/opendemocracyuk/commodification-essence-of-our-time/>.
- Lipura, S.J., Collins, F.L., 2020. Towards an integrative understanding of contemporary educational mobilities: a critical agenda for international student mobilities research. *Glob. Soc. Educ.* 18 (3), 343–359.
- Marginson, S., Nyland, C., Sawir, E., Forbes-Mewatt, H., 2010. *International Student Security*. Cambridge University Press, London UK.
- Marginson, S., 2011. Higher education and public good. *High. Educ. Q.* 65 (4), 411–433. <https://doi.org/10.1111/j.1468-2273.2011.00496.x>.
- Marx, K., 2011. *Das Kapital: A Critique of Political Economy*. translated by Samuel Moore. Pacific Publishing Studies, Madison Park.
- Massey, M., 1994. *Space, Place, and Gender*. University of Minnesota Press, Minneapolis MN.
- Mathews, J., Sidhu, R., 2005. Desperately seeking the global subject: international education, citizenship and cosmopolitanism. *Glob. Soc. Educ.* 3 (1), 49–66.
- McBurnie, G., 2015. International branch campuses: Australian case studies. *Int. High. Educ.* (29).
- Mignolo, W., Escobar, A., 2013. *Globalization and the Decolonial Option*. Routledge, New York NY.
- Mok, K.H., Lang, S., Xiao, H., 2020. The quest for global talent for changing economic needs: a study of student mobility and job prospects for returnees in China. *Glob. Soc. Educ.* 18 (1), 79–96.
- Nye, J., 2005. *Soft Power: The Means to Success in World Success*. Perseus Books Group, New York NY.
- Pieterse, J.N., 2010. *Development Theory*, second ed. Sage, Thousand Oaks CA.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Rizvi, F., 2011. Student mobility and the shifting dynamics of globalization. In: Davis, D., McIntosh, B. (Eds.), *International Education in Australia: Celebrating Twenty Five Years of Achievement*. University of New South Wales Press, Sydney.
- Rizvi, F., 2020. Reimagining recovery for a more robust internationalization. *High. Educ. Res. Dev.* 39 (7), 1313–1316.
- Roberts, J., 2015. Globalization of services. In: Bryson, J., Daniels, P. (Eds.), *Handbook of Service Business: Management, Marketing, Innovation and Internationalisation*. Edward Elgar, Cheltenham UK.
- Robertson, S., 2013. *Transnational Student-Migrants and the State: The Education-Migration Nexus*. Palgrave, London UK.
- Robinson, W., 2004. *A Theory of Global Capitalism: Production, Class and the State in a Transnational World*. John Hopkins Press, Baltimore MD.
- Robinson, W., 2011. Global capitalism theory and the emergence of transnational elites. *Crit. Sociol.* 38 (3), 1–23.
- Seth, S., 2007. *Subject Lessons: The Western Education of Colonial India*. Duke University Press, Durham NC.

- Sidhu, R., 2007. GATS and the new developmentalism: governing transnational education. *Comp. Educ. Rev.* 51 (2), 203–227.
- Sklair, L., 2001. *The Transnational Capitalist Class*. Wiley-Blackwell, London.
- Stein, S., 2019. Navigating different theories of change for higher education in volatile times. *Educ. Stud.* 55 (6), 667–688.
- UNESCO Institute of Statistics, 2019. Global Flows of Tertiary Level Students. <http://uis.unesco.org/en/uis-student-flow>.
- Vertovec, S., 2009. *Transnationalism*. Routledge, London UK.
- Zigurs, C., McBurnie, G., 2006. *Transnational Education: Issues and Trends in Offshore Higher Education*. Routledge, London UK.

Postcolonialism in the movies and higher education: a case study of the narratives of South African and international students at the University of Pretoria

Everard Weber and Venise Joubert, SARCHI Teaching and Learning, University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	241
Representations of the other and voice	241
Politics	244
Hybridity and change	246
Conclusion	248
References	248

Introduction

In recent decades historical developments that have made the world a smaller place have intensified at breath-taking pace. Time, space and place have been compacted to an unprecedented degree. The use of English spread during the colonial and neo-colonial periods of the 19th and 20th centuries and has now become even more widespread through the internet. Economies have become inter-dependent regionally and internationally as the relations of hegemony and inequality have persisted and grown. Knowledge and innovation now travel across the world matching the speed of capital and financial flows. Higher education has become “massified.” Cross-border growth in the numbers of international students and faculty who study and teach in countries other than the ones in which they were born has burgeoned. Against this background our interest is in the social relations of integration, separation and conflict which have become entwined within institutions of higher learning. How can we make sense of unfolding developments in the sociology of higher education? We wish to show the usefulness of postcolonialism as portrayed in film in analyzing the experiences of postgraduate university students in South Africa.

There is a genre of writing that focuses top-down on higher education as a system. Thus mega-data are used quantitatively to inform analyses of the size, shape and trends in higher education. Its ideological roots can be traced to the commodification of knowledge. Data are collected to solve what are the pressing problems of the day. There is a business-like concern for “efficiency” and instrumental outcomes encapsulated in practical suggestions. The humanistic perspectives of the humanities, like the academic disciplines associated with it, are devalued. In contrast, the approach adopted here throws into relief the sociology of higher education on the ground. This article follows the publication of a book (Weber, 2017) which interpreted interviews collected at residences at the University of Pretoria, South Africa. These residences housed the largest numbers of postgraduate international students and their South African peers. One way to bring humanity to social study is through the media of novels, film and the arts and through listening with care to the voices of those directly involved.

Representations of the other and voice

Said's (1978) landmark study, *Orientalism*, critiques Western scholarship as portraying the Orient as secretive, exotic, and sensual as opposed to a rational, objective, scientific West. This particular Orient he argues is not a portrait of reality. It derives from and reflects the political and economic power of the West over the East. There is therefore for Said a relationship between the forms and content of knowledge and the social structures of hegemony.

In Lin and Ritchie's (2019) movie, *Aladdin*, we are told, “People see what they want to see.” A “street rat” is insulted thus: “You were born worthless and will die worthless. Only your fleas will mourn you.” The movie nevertheless asks that we identify with the urchin. On one occasion he says, “If only they'd look closer, they'll find the real me.” He succeeds in showing the world his real self, actualizes his potential and in the end marries the beautiful princess. More soberly, in *David Copperfield* by Charles Dickens, the character, Wilkens Micawber famously said, “Something will come up”. The promise that something will come up for street people under capitalism serves to mock them. The national lottery in South Africa uses an advertisement that states, “One day is one day ...”

Black South Africans fought against racial othering and discrimination under apartheid. And yet, in post-apartheid society there have been recurring xenophobic outbursts of violence against foreign nationals from other African countries. The historical victims of prejudice now practice discrimination against fellow-Africans. A student said:

South Africans ... when they see people coming from other African countries the question is, “Why did you leave your country to come and study in South Africa? Were there no universities in your country?” ...

"We don't want you in this country. You should study in your own country." And sometimes when you are finishing your degree the question [becomes] "When are you leaving?" ...

That's ... very rude ... And when you start telling them, "I'm not leaving I'm staying here, I have a future in this country, I have studied, I have improved myself, if there is a job that ties with the degree I have done, I can work here what is the problem with that?" they get very insulted, they get very infuriated. ...

South Africa is for South Africans and not for people from other races [i.e. international students from other African countries]. "The earlier you get out of here the better for us; you people are using our resources."...

quoted in Weber (2017, pp. 96–97).

Similarly, black South African workers who clean university residences were reported by postgraduate students to find the idea repulsive of their daughters marrying Africans from other countries. And there was the somewhat bizarre idea among certain South African students that while African students "from the north" might be black but they did not have to be *that* black. There were degrees of blackness with South Africans (despite the evidence!) perceived to be lighter-skinned than the rest of the Africans. These others were all lumped together as coming from the alien north which one student described as "up there."

The subordinate position and thus silencing of international students becomes manifest in many ways. It derives from, for example, their disempowerment because of their financial struggles to survive and complete their studies in a foreign country that often, albeit not always does not welcome them. In the interview extract above one might ask: Who has power over whom? How? And who consequently is able to insult and humiliate whom? Power and domination are asserted in and through acts of discrimination. Black South African students who discriminate against international students are themselves subjected to discrimination, especially the large numbers who presently study at what were former whites-only institutions. Students spoke about the separation between white and black South Africans in lecture halls. The racial groups also congregated in separate social spaces and restaurants on the campuses. Race relations were described as "not harmonic" and fraught with latent and open conflict. A black South African student made the point that she had participated in a group assignment in which white and black students had been required to work together. She stated that the whites dominated the discussions and that the blacks, herself included, had been passive and quiet. The blacks, faced with, the assertion of white power had in this instance "allowed" it. Nationally, there have been demands for greater recognition and use of African languages and the abandonment of Afrikaans, associated with apartheid, as an official language of certain universities and its use as a medium of instruction. Language is inseparable from the articulation of voice and identity and feeling that one is at home at the university. The number of blacks has increased at historically white universities. However, transforming institutional cultures and curricula which are integral to language have proved to be stubborn. (see for example, Moguerane, 2007; Prah, 2015; Msila, 2009; Thiong'o, 1986; Department of Education, 2008). One the one hand therefore access in terms of numbers has increased for black South Africans; on the other hand once inside they are pushed out because of their experiences at these institutions.

Aladdin represents a Western image of benign masculinity. In the opening scenes of the 2019 film (Lin and Ritchie, 2019) oriental women in a market are attracted to him, the playboy westerner. He seduces them with clever repartee and they are eager to be seduced by him. The scene might be spirited but its playfulness hides the seriousness of its message. Aladdin's persona, like his sexuality, knows no Freudian dark side. His poverty does not bother him. The image is contrasted with the "real" East embodied in the evil Jafar. Jafar is the embodiment of overwrought sexual repression. While Aladdin might smile endearingly; Jafar is constantly angry, sinister and deceitful. His face remains taut throughout the movie. He moreover possesses devilish powers. He casts spells over people. Orientals like him cannot be trusted. Jafar is more and more appropriately dressed as the Genie transforms him into his wish to become the most powerful sorcerer in the world. He progressively wears black with shades of red to the reddish-black attire at the height of his power. The red symbolizes the blood that will flow during his rule at home and in foreign lands. Aladdin enters the city as a rich sultan dressed in white and gold. Jafar's repressed sexuality takes on perverse forms. Soon after Jasmine articulates her refusal in song: "You can't keep me silent ... I cannot be broken ... I won't be speechless," Jafar asserts his mystical powers over her father and forces Jasmine to marry him. Jafar's political ambitions to become the dictator of Agrabah and militarily expand his reign territorially can be compared to his intentions to conquer Jasmine and dominate her. She will no doubt join his harem as his concubine. Her wedding dress is purple-red. Jafar's plans are skuppered when Jasmine escapes to her Aladdin in whose arms she will enjoy self-actualization, freedom of speech and democracy (Lin and Ritchie, 2019).

Discriminatory oppression entails the suppression of voice and myriad misrepresentations of the other. Voice is related to representation and discourse, disdain for and humiliation of the other, and the exercise of power over them. This is also an epistemological question, as Said (1979) argues. Western scholarship speaks for the East in the same way that men might speak for women, the rich for the poor or in which whites might wish to speak for blacks. One can look at it politically and historically. Under legislation passed in 1936 in South Africa, black Africans in the Cape province lost the right to the common-roll franchise. This was replaced with a vote for three white members of parliament and two in the provincial council. All Africans in the country (around more than 75% of the total population) could elect four white senators. The whites would represent the blacks and speak on their

behalf. Before 1936 and decades after similar arrangements until the attainment of majority rule in 1994 were implemented the goal of which was to deny blacks political rights and to disempower them.

Feige and Coogler's (2018) movie *Black Panther* challenges black voicelessness, inferiority and oppression. We are invited to imagine an African country, Wakanda, that is technologically advanced and more developed than the West. Here Africans govern Africans. The film raises but does not engage with portraits of a conflict-free, unchanging, romanticized pre-colonial Africa. Such scenes of the African past are portrayed in *Black Panther*. Consider, for example, the Garden-of-Eden-in-Africa which appears when T'Challa dreams about his dead father and speaks to him. The Wakandan nation repeatedly genuflects before its ancestors. From an idyllic past its latter day elites can draw spiritual inspiration. "I call upon the ancestors ... Heal him" (Feige and Coogler, 2018). Film clips of tranquil rurality hide the modern architectural structures of the country, its wealth and high-tech wonders. Nevertheless, *Black Panther* does disrupt racist perceptions of black people and backward Africa. It presents a vision of what African development might have been had the continent been free of colonialism and neo-colonialism. It de-centres narratives about African and Southern development and history as a function of European history and development. Wakanda, we are told "does not engage in international trade or accept aid." Put another way:

"Here's another [wounded] white boy for us to fix [medically]. This is going to be fun." As the patient wakes up and recovers from his treatment, he confusedly asks, "Is this Wakanda?" The reply: "No its Kansas". He is nick-named "colonizer"

Feige and Coogler (2018).

All the heroes in the movie are black men and women, albeit well-off black men and women whose class position is unproblematised. They can fight in the best traditions of the white heroes in the other Marvel action films. What white people can do black people can do better. The women in the movie are "on top" (Davis 1978). They are strong in word and deed and, like the best brothers in the hood in their kick-ass *attitude*. One is reminded of Richard Roundtree's cool portrayal of a black detective in the *Shaft* movies of the early 1970s. Black attitude jars and stabs at the white status quo. It is in-your-face demonstration of black empowerment rooted in visceral experiences of blackness. Attitude mirrors the (albeit computer-generated) fighting of the women. *Black Panther* is as anti-sexist as the *Shaft* movies were not. The soundtrack was an instant hit, but the lyrics went,

Who's the black private dick

That's a sex machine to all the chicks?

(Shaft!)

You're damn right

Who is the man

That would risk his neck for his brother man?

(Shaft!)

...

He's a complicated man

But no one understands him but his woman

(John Shaft)

Hayes (1971).

The picture of Wakanda is in several ways that of a typical African country and the idea that Africa is one place. Typical African countries are supposedly divided along tribal lines. Centuries ago Wakanda consisted of 5 “tribes”, 4 of which eventually united, except the menacing Jabari who fled to the mountains. Their leader at one point threatens the humans in front of him, “I will feed you to my children” and then laughs, not to worry, “We are vegetarians!” Wakandan elites have trans-historically and across the ages been at war with one another. Who is to be the next ruler-king-dictator? Who will control the valuable and powerful vibranium deposits buried beneath Wakandan soil? The tribal myths about the singularity of African history are good for present day Wakanda – despite its technological advancement. Wakanda, in other words, remains trapped in the imagined continuities of a particularized African past. The physicality of the heroes fighting not only resembles their white counterparts in other Marvel movies, but takes on what are perceived to be primordial forms. *Black Panther* (Feige and Coogler, 2018) repeats colonial and current myths about Africa, its histories and peoples. Similarly, in *Aladdin* (Lin and Rithie, 2019) the precolonial Orient has been colonized by contemporary Hollywood, in the story-line of the poor boy who eventually makes it to the top and the pop orchestral soundtrack deemed appropriate for a tale set in Agrabah. The Bollywood dancing sequences portray happy-happy Orientals. Agrabah, the Orient, India—north, south and central—are one and the same, just like the whole of Africa north of the Limpopo is same—but according to some of our respondents apparently blacker than the south.

As we have argued, these representations are inextricably tied to questions of power, voice and the silencing of the subaltern. I speak from a position of authority over you when I tell you, explicitly or otherwise, that I am superior to you. Spivak (1988, 2016) discusses the problem of whether the lowest and most oppressed groups in society can speak and who can speak on their behalf. There is therefore a degree of continuity and conceptual complementarity between Said and Spivak. It concerns the development of a political consciousness among the poor and marginalized so that they can represent themselves and speak for themselves and politically act by themselves. The African writer Ngũgĩ, as is well-known has raised the question of what language the subaltern might speak and for whom, by himself refusing to write in English. In a different way this question is familiar to anthropologists and qualitative researchers. It is not uncommon for white middle class men from “the university” to “go native” in foreign lands or poor black communities and write reports about them, typically in English. This raises at least 2 other questions. First, certain colonials educated in leading metropolitan universities have led revolutionary movements, despite their formal Western education. They committed what Cabral, himself an example of this group, called “class suicide” by identifying with the oppressed and leading their struggles. On balance these organic intellectuals are exceptional and cannot be said to represent their class. Second what is called Western scholarship is of course far from a unified, timeless body of knowledge. It has been characterized in no small measure by critique and if not an acceptance, a tolerance for critique. While we are born into and socialized by society we also have the ability to critically reflect upon it and ourselves.

Speaking in this context is assumed to be the articulation of thought through the medium of language, education and socialization. It is therefore not only a question of whether or not the subaltern can speak, who speaks on her behalf, and in which language she does so, but also a question of whether thought is written and published or spoken. The subaltern, based on experience of and in the world, have independent insight and representations of reality. The problem as noted above is when qualitative researchers—like ourselves—seek to rewrite and interpret it and present it in a different form and language. From the late 1970s Ngũgĩ challenged the idea that theater was meant to be about actors performing for silent audiences. He wrote plays with which ordinary folk could identify. *Ngaahika Ndeenda* was a play co-written and developed with Thiong'o and Mirii, 1980 at the open-air Kamiriithu Cultural Center at Limuru, Kenya. It proved to be popular through its use of song and the engagement with and active participation of the audience. It attacked colonialism, religion and the postcolonial African elite, portrayed as avaricious and disconnected from the majority of the population. The play was banned and the playwrights were subsequently detained. The content and subject matter of the play and its populist form had irked the government. The specificity of the content mattered for the Kenyan government as it matters in what the academe chooses to write or research about and what it chooses to ignore and remains silent about (see Pillay, 2016).

Politics

In *Black Panther* (Feige and Coogler, 2018) the plot focuses on the in-fighting among the Wakandan well-off who are also its rulers. The populous are faceless and silent, except for chanting and stamping their spears as the princes engage in ritual fights. Politics is further influenced, but to a far lesser degree on the position of the military who swear allegiance to whoever occupies the throne, and what looks like a council of wise elders. But these constituencies play second fiddle to the comings and goings of the princes. The masses do not have voices or roles to play in determining the course of the country. Politically they neither think nor act.

What does the fighting between the main protagonists, Erik Killmonger and T'Challa, represent? T'Challa wants to maintain Wakanda's independence and isolation; Killmonger wants it to use its resources and liberate blacks internationally. He speaks for Pan-African unity and struggle. The movie treats T'Challa sympathetically while Killmonger is the villain, a kill-monger as his name suggests.

Killmonger, addressing the Wakandan elders:

Y'all sitting up here comfortable. Must feel good.

There are 2 billion people all over the world that look like us. But their lives are a lot harder than yours. Wakanda has to liberate them all
Feige and Coogler (2018).

The movie makers interpret the black liberation represented by Killmonger as akin to Wakandan imperialism. Similarly revolution in contemporary colloquial language is associated with the epithet "bloody". Killmonger is a cold-blooded mercenary and revolutionary: "... [he is] hell bent on killing for killing's sake – indeed, his body is marked with a scar for every kill he has made" (Lebron, 2018). Violence and revolution in this sense are devoid of political organization, programmes, ideas or political consciousness. Thought and knowledge cannot be born in acts of revolt. "Your heart is so full of hatred you are not fit to be king (Feige and Coogler, 2018)." Killmonger disrupts middle class liberal sensibilities by his anger and militancy, the product of the colonialism the movie critiques, but which it ultimately eschews. It urges us to view T'Challa sympathetically. He ends up killing Killmonger. In T'Challa we are presented with the wise, humane king who does not want "to wage war" because, "It is not our way." "He is a good man with a good heart." Wakanda must remain isolated and "safe" from the rest of the black world. He cannot reply when Killmonger asks him about the rest of the continent. "Are they not our people too?" (Feige and Coogler, 2018). The solution to Africa's problems is a narrative in which the good black Africans (in alliance with an affable CIA spy!) defeat the bad black Africans. "Perhaps Killmonger's main dream to free black people everywhere decisively earns him the fate of death" (Lebron, 2018).

As Killmonger dies, T'Challa offers to heal him. He declines saying that he will then be imprisoned for his rebellion. He poignantly recalls how black slaves who were transported on ships to America chose to commit suicide rather than live in captivity. This would have been a strong ending. It would have invested Killmonger's character with a measure of humanity (see Lebron, 2018).

Black Panther was a well-known political party active in the United States in the late 1960s and 1970s. It did not reject violence as did other tamer civil rights organisations. It was originally called the Black Panther Party for Self-defense (against police brutality).

... [the Party] mobilized ... many people to fight white supremacy and capitalist exploitation. The Party generated an escalating cycle of insurgency by linking disruptive actions to community programs as part of a coherent anti-imperialist politics ... [which] forged common cause with a wide array of domestic and international constituencies

Bloom and Martin (2016, p. xvi).

As the credits of *Black Panther* roll, the liberal critique of Black Panther-ism and Killmonger resumes. T'Challa buys several blocks of urban ghetto-land to build the first "International Wakanda Outreach Center". He addresses the United Nations (in the presence of the approving CIA-spy!). Wakanda will henceforth share its resources and break its historical isolation. The Black Panther king's message is one of goodwill and peace. Y'all must love one another.

The South African student rebellion of 2015–16 was a national revolt which linked problems in higher education to the wider society. Its politics were different from politics of *Black Panther* because here the masses not only shout, they act, march, organize and formulate demands. They unsettle those in power and contest their hegemony. The most prominent of the *FeesMustFall* struggles of 2015–16 revolved around the racially discriminatory cultures at universities and colleges discussed above, the affordability of higher education, the predominance of English and Afrikaans as languages of instruction, demands that curricula be decolonized and Africanized, and a critique of top-down leadership in university management. More broadly the movement challenged the social agreement between corporate business, the apartheid government and the liberation movement that was supposed to lead to "peace and reconciliation" in a "new" South Africa after the first democratic elections in 1994 (Booyesen et al., 2010).

It is instructive to compare the politics of the *FeesMustFall* protests not only to the politics of the movie makers of *Black Panther*, but to the narratives of the international students we interviewed. They stood aloof from these struggles which were spearheaded by South African students. International students' apparent withdrawal from politics this time and in this country did not mean that they might not participate in future social movements either here or in their home countries. An Iranian student and a Palestinian student spoke about the politics in their respective countries and how they followed unfolding events. The Palestinian related the impact upon his uprooted family which was dispersed across several countries. He was also affected, and deeply so, by the injustice of the situation in Gaza. It was almost impossible to say how these students would react or what they would do when they returned home – in the case of the Palestinian assuming he was permitted entry to his homeland. Similarly, it is difficult to predict how international students in South Africa would act if the xenophobic violence that has hitherto been confined to the townships spills over onto university campuses.

The continuities and contradictions of political change and of political consciousness are manifest nationally and across time. During the 1960s there was little protest in South Africa after the wave of resistance following the Second World War had been

crushed in the aftermath of the 1960 Sharpeville massacre. The Soweto student revolt in 1976 broke the 1960s' political silence and sense of defeat and culminated in the attainment of majority rule in 1994. If the politics of quietism predominated during the 1960s, its opposite, open, mass rebellion predominated during the latter part of the 1970s and especially during the 1980s. Protest politics therefore assume different forms and are contingent upon the times and social circumstances. Many of the tensions in social relations between black South African students and international students occur during inter-personal encounters. These relations could lay the bases for future developments. In the case of South African students on-campus racial discrimination contributed to the protests that occurred in 2015–16. On the other hand our interviews show that the issue that unites most students, irrespective of their country of birth or whether they come from middle class or working class homes, rural or urban areas, are the fees they are required but cannot afford to pay.

Hybridity and change

Said's work is characterized by binaries, between East and West for example. Bhabha (1994) has developed a more nuanced conceptualization in his writing on the relationship between colonizer and colonized. The encounter is described by hybridity, in-betweenness and ambivalence, rather than either-or-isms. This arguably is the opposite of structuralist explanations of how knowledge or certain ideologies gain currency in society. Gramsci stated that the subaltern is not only ruled through absolute power but with their consent. Foucault was interested in how power/knowledge is constructed and disseminated in society. There is a distinction between the explicit use of force to impose ideology as in a police state and its more subtle, stealthy and more sinister dissemination under say liberal democracies. Both are mechanisms of control and are linked to how power can be exercised and how it can be contested.

Cameron and Landau's *Avatar* (2009) is set in the year 2154. Earth faces an energy crisis because all its natural resources have been depleted. There are, however, rich deposits of unobtainium on the planet Pandora, inhabited by the Na'vi. The Na'vi's skins are blue. Humans cannot breathe the atmosphere on Pandora. Their exploration of Pandora is made possible by using Na'vi-human hybrids or "avatars." Jake Sully, the hero of the movie, is such an avatar.

The Na'vi inhabits a sublime forest-paradise which they have to defend against the imperialist humans, or Sky People. The earthlings' ultimate wish is to exploit Pandora's unobtainium deposits. The Blue People, at one point derogatively called "blue monkeys," come to represent oppressed indigenous groups worldwide and the activist causes of environmentalists. The head of security explains that the Na'vi, who, significantly, stand 3m tall above the humanoids, are "fond of arrows, [which] dipped in neurotoxins can stop your heart in a minute." Moreover, "They are very hard to kill" (Cameron and Landau, 2009). The story moves from futile attempts at persuasion to the colonizers' realization of the necessity to use military force. Put the Na'vi and the Sky People together and we have the makings of a "third space," or the embodiment of Dr. Jekyll and Mr. Hyde at the same time and place. The movie does not really explore this hybridity. However the dynamics of the encounters between the Blue People and the Sky people are given some attention. For example, the Na'vi do not kill the "dream-walker", Jake Sully, but allow him in to learn more about their language and culture in the hope that he might be cured of his "madness" (ignorance). In contrast the Na'vi are studied from a strategic and military and a scientific point of view with the goal to eventually conquer them. Jake Sully is initially sent to spy on them. "Just find out what the blue monkeys want," Sully is told (Cameron and Landau, 2009).

The liberal-minded scientists are provided with resources to study the Na'vi. It is presumed that these scientists are sympathetic, perhaps too sympathetic toward them. They were supposed to be "winning hearts and minds" through "education, medicine and teaching them English." In particular they were supposed to find a "diplomatic solution" for the removal of the Na'vi from Hometree, the sacred place beneath which are the richest deposits of unobtainium. Instead relations with the Na'vi become progressively more hostile (Cameron and Landau, 2009).

We are left wondering about the nature of the social relations between the Na'vi and the Sky people: what did the Na'vi learn and assimilate from the humans and vice versa? What are the specificities, complexities and contradictions of this space? How are we to understand the process? This applies to the avatar Scully as well. He is simply shown to be now humanoid, now Blue. Bhabha has theorized about the point of encounter between colonized and colonizer. He is concerned with strategies the colonizer adopts in order to survive but crucially also identity formation and political action: "dealing with or living with and through contradiction and then using that process for social agency" (Bhabha, 1994). The third space constitutes the ways in which oppressed "mimics" what is expected of him or her. This means understanding the ways of the oppressor, his mindset and craftily hiding your real goals in regard to resistance. Hybrid identities and persona are formed in such situations.

As the story in *Avatar* unfolds Scully and the liberals switch their allegiance and any ambivalence they may have had about loyalty to their race and join the Na'vi cause militarily. They defeat the almighty Sky imperialists in battle and send them back to earth. In the movie "going native" is transformative. It means in the end siding with and joining the natives in battle. Scully turns blue this time permanently.

A mixed in-between identity is an unstable identity. Colonial rule might conquer indigenous peoples, impose their culture on them and through education systems socialize the colonized, but this hardly ever occurs in a "pure" form. Indigenous languages, for example, are often not eradicated. The old and the new can become hybridized. Thiong'o, (1986, p. 7) critiques African writers' use of English. As opposed to using their mother tongue when writing, African writers tended to "[enrich] foreign languages by injecting ... "black blood" into their rusty joints". African writers use of English was not "pure", but had been made to "carry the weight of ... African experience by, for instance, making them "prey" on African proverbs and other peculiarities of African speech and folklore".

If we think of hybridization and what Bhabha (1994) calls a “double consciousness” as unstable we can imagine change over time and possibility with traces of the original forms surviving in new, diverse ways. In the end however Jake Scully turned Blue. Prior to this was the act and either-or-ism of war which facilitated the birth of the new. The military encounter of the earthlings and the blue people has political significance because there is nothing indeterminate about war. In this respect war is similar to periods of mass anti-apartheid revolt like the 1980s in South Africa. The double consciousness in identity might be contingent upon political and social conjuncture and it might lead to transformation but might also persist or re-surface beyond the act or event of transformation. Looking into the future, what is the depth and permanence of Scully’s blueness?

Hybrid and ambivalent social relations are evident in the consciousness of black students who come from rural areas and the townships as well as international students who come from other African countries to study in South African towns. This is evident in how students constantly think about home as they live and study away from home. There were South African students from other parts of the country and from rural areas who were now studying at the University of Pretoria who found it to be a place as alienating as the international students from other African countries. The architecture of the University of Pretoria, its buildings, high-tech lecture halls and blended teaching and learning speaks to a modernity and class privilege that black students coming from townships and rural areas experience as bewilderingly strange. Some of them were angered by the clothes and cars certain white students possessed and wore. They typically had attended overcrowded classrooms in schools without adequate water and sanitation and were taught by stressed teachers battling to understand new curricula. These students who were there to succeed and join the middle class were denied the possessions white South African students displayed.

Several students spoke of changing social relations on campuses and in residences:

I think it’s changing because when I ... [arrived] here in 2011 it was in rare cases ... [that you saw] blacks and whites and Indians together ... that place I’m telling you [about on campus] ... it’s for rich people. It’s normally Indians who ... sit there, but nowadays you see Indians, blacks [and] whites.

Although this sentiment was not shared by everyone certain students spoke of gaining an understanding of difference and diversity. For some this understanding came from the shared experience of being students and with the acknowledgment that academic assistance could be given by anyone, not just “people from your geographic region.” It is not just support that fosters understanding but also knowledge:

... because of being with different people ... I think there are no culture shocks anymore. If someone is doing something ... I know ... it’s not to hurt me ... it’s because according to [their] culture [it] is normal. But according to my culture it’s not normal.

And:

You forget about differences, coming together ... in class and being one thing ... [sharing] those notes, those question papers, asking for that help, [encountering] somebody who is asking [for] help from you who is not ... [the] same colour of ... skin ... [T]hat brings the change and you get to interact. You just see people now ... you don’t see [colour] ...

Other students spoke about a willingness to “engage” with each other and the appeal of learning from one another, learning to cook local dishes and learning new languages:

People are beginning to accept differences ... People want to learn more about other countries, they want to visit other countries. And this ... I see [and] I hear in conversations on campus ...

It is in the meeting within the metaphorical third space, on campuses and in classrooms and residences that the possibility of change becomes evident. Foucault (1984) argued that power allows for the possibility of action and resistance, and subsequently, change. We have the ability to recognize and question normalized and accepted belief systems. Belief systems are redefined and reinforced through the flux of ideologies. Belief systems do not delineate an absolute truth that can be discovered and accepted, but instead, represents “a battle about the rules according to which the true and false are separated and specific effects of power are attached to the true” ... a battle about “the status of truth and the economic and political role it plays” (Foucault, in Rabinow, 1991: 74).

A student said:

I’ve [definitely] changed ... I’m more open minded. I came when I was very religious. I think I’ve lost my religion ... [not] in the sense that I’ve lost the values that come ... with my Christianity ... [but] I’m questioning more of the things about religion ... [and] culture ... I think [it’s] good ... my mum would say otherwise [and] my aunt would pray for me right now ... [but] I don’t want to be cramped into one box ... I don’t want to be like someone

who thinks and believes in only one thing you know because I believe I'm more now, more acceptive of different traditions and different ways of doing things because if you look at them, they've also helped other societies you know ... I think I'm more acceptive and a more willing to listen to what someone else has to say regardless of [what] ... I initially believed ...

[F]or instance, I'm more tolerant to ... [and accepting of] gay people now ... [B]ack home it's a ... no go. It's crazy, they kill gay people ... [Y]ou question why you have this anger ... and when you ... look at it ... it's because of what I've heard. I've not really seen the person or interacted with the person and realised that it is their choice ... [B]ack home for instance ... the reason ... I think ... I was so against [gay people] is because there was ... talk that they go into schools and ... force children to become gay and I think that's what many people like me thought ... [M]y cousin has a company which I help him with and they ... [employ] many gay people and these are people who are working and adding value and very good people in society ...

The student's account speaks of a reimagining of the self. The student is almost ashamed of previous biases and prejudices. She questions the foundation of her beliefs and religion, knowing that her family would disapprove of the progressive and liberal views. Change is intermittent and varies from person to person. Many students did not perceive any changing social relations at the university and did not acknowledge change within themselves. It is unrealistic to expect everyone's experiences to have transformative or cathartic effects like this student or like Scully. It is within the locus of struggle and resistance of the third space that we can imagine the possibility of change, especially in times of heightened social and political transformation.

Conclusion

Globalization and the internationalization of African higher education have resulted in large scale cross-border mobility. This has influenced social relations as local and internationals engage one another. Sociologists and historians commonly studied South African society in terms of race, class and perhaps gender as well. Today nationality has become increasingly important as the numbers of foreign nationals from the rest of Africa has grown. Economic, financial, and political integration have influenced our social consciousness, how we view the world and act upon it. It has had a profound impact on our social imaginations (Appadurai, 1996). The cultural and sociological dimensions of globalization are significant and deserve study in their own right.

We have discussed several of the basic tenets of postcolonialism through a review of films and interview extracts. Postcolonialism provides a rationale and a theoretical tool for the study of the sociology of higher education. Said (1979) was concerned with the relationship between East and West and how that relationship flowed from unequal relationships between the colonized and the colonisers. Implicit within this thesis is the absence and suppression of an authentic colonial voice. This has been the starting point of Spivak's work. Voice in this context does not only concern the articulation of thought. Voice is central to politics and political action. Can others speak for the under classes? For women? What happens when they do? Qualitative researchers will be familiar with this problem from a methodological and scholarly point of view. Historians have to grapple with it as well as they seek to portray periods, individuals and groups. The question is also relevant to the creation of knowledge. Today and historically few writings and publications on and about Africa originate from within Africa by Africans.

Postcolonialism emphasizes decentering or looking at the world from the South rather than Eurocentricism, groundedness, othering in social experience and in epistemology, the suppression of subaltern voice, and the complexity of hybrid identities. These concepts can be found in the foundational writings of Said, Spivak Bhabha and Ngũgĩ. They are key themes in the recent movies *Aladdin*, *Avatar* and *Black Panther*. Sequels are planned for the latter 2 films. We await them with keen interest.

We have integrated problems commonly discussed in the literature, such as globalization and the internationalization of higher education through the lens of postcolonialism and film. The latter can be taken as an example of cultural studies in general and can also include the fine arts. Consistent with the humanistic traditions of qualitative research and of interdisciplinarity, it has entailed working with different types of data and trying to integrate them into a coherent story. Postcolonialism is ideally suited to this type of analysis. It can enrich and at the same time elucidate the subject of enquiry. Thus, we have juxtaposed the abstractions of content, and of style of Homi Bhabha with the moviemakers and action scenes of *Avatar*. One of the common criticisms of number-crunching in social science is that we do not encounter and are invited to engage with people and individuals where it matters most – on the ground and in real life. This highlights a familiar problem which we leave for discussion on another occasion about the relationship between the data one collects and the written form in which it is presented in print. In our view too many academic journals, following the natural sciences have cast form into repetitive stone. A further advantage of postcolonialism and popular movies is that it facilitates thinking about national questions and developments such as those unfolding at South African tertiary institutions, internationally. When all is said and done, the overarching topic of our writing on international higher education is well summed up by Alain Touraine (2000), "Can we live together?"

References

- Appadurai, A., 1996. *Modernity at Large: Cultural Dimensions of Globalization*. University of Minnesota Press, Minneapolis, MN.
 Bhabha, H., 1994. *The Location of Culture*. Routledge, New York.

- Bloom, J., E Martin, W., 2016. *Black against Empire: The History and Politics of the Black Panther Party*. University of California Press, Oakland, CA.
- Boysen, S., Godesll, G., Chikane, R., Mofu-Walsh, S., Ntshingila, O., Lepere, R., et al., 2010. *Student Revolt, Decolonization and Governance in South Africa*. Wits University Press, Johannesburg. muse.jhu.edu/book/5052.
- Cameron, J., Landau, D., 2009. [Motion Picture] 20th Century Fox. *Avatar*.
- Davis, N.Z., 1978. Women on Top: symbolic sexual inversion and political disorder in early modern Europe. In: Babcock, B. (Ed.), *The Reversible World: Symbolic Inversion in Art and Society*. Cornell University Press, Ithaca.
- Department of Education, 2008. *Final Report of the Ministerial Committee on Transformation and Social Cohesion and the Elimination of Discrimination in Public Higher Education Institutions*. Department of Education, Republic of South Africa, Pretoria.
- Feige, K., Coogler, R., 2018. *Black Panther*. Motion picture. Marvel Studios.
- Foucault, M., 1984. Truth and power. In: Rabinow, P. (Ed.), *The Foucault Reader*. Pantheon Books, New York, pp. 51–75.
- Hayes, I., 1971. *Shaft!* [Song Writer and Producer], Stax Recording Studios. Memphis, Tennessee.
- Lebron, C., 2018. "Black Panther" Is Not the Movie We Deserve. *Boston Review*, 17 February.
- Lin, D., Ritchie, G., 2019. *Aladdin*. Motion picture. Walt Disney Pictures.
- Moguerane, K., 2007. Post-Apartheid Transformation Race in a Student Community in Pretoria. *Saarbrücken Vdm Verlag Dr. Muller*.
- Msila, V., 2009. Africanisation of education and the search for relevance and context. *Educ. Res. Rev.* 4 (6), 310–315.
- Pillay, S.R., 2016. Silence is violence: (critical) psychology in an era of Rhodes Must Fall. Editorial. *S. Afr. J. Psychol.* 46 (2).
- Prah, K., 2015. Has Rhodes Fallen? Decolonizing the Humanities in Africa and Constructing Intellectual Sovereignty. *The Academy of Science of South Africa (ASSAF) Inaugural Humanities Lecture*. HSRC, Pretoria.
- Rabinow, P., 1991. *The Foucault Reader: An introduction to Foucault's Thought*. Penguin, Harmondsworth.
- Said, E., 1978. *Orientalism*. Pantheon Books, Cop, New York.
- Spivak, G.C., 1988. Can the subaltern speak? In: Nelson, C., Grossberg, L. (Eds.), *Marxism and the Interpretation of Culture*. Macmillan Education, Basingstoke, UK, pp. 271–313.
- Spivak, G.C., 2016. *Critical Intimacy: An Interview with Gayatri Chakravorty Spivak by Steve Paulson*. Retrieved 13 November from. <https://lareviewofbooks.org/article/critical-interview-gayatri-chakravorty-spivak/>.
- Touraine, A., 2000. *Can We Live Together? Equality and Difference*. Univ. Press, Stanford, Calif.
- Thiong'o, N. wa, Mirii, N. wa, 1980. *Ngaahika Ndeenda I Will Marry When I Want*. Heinemann Educational Books.
- Thiong'o, N. wa, 1986. *Decolonising the Mind: The Politics of Language in African Literature*. J. Currey, London.
- Weber, E., 2017. *"A Home Away from Home": A Community of International and South African University Students*. Nova Science Publishers, New York.

Neoliberal education: from capitalizing on disaster to critical pedagogy

Kenneth J. Saltman, University of Illinois at Chicago, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Neoliberal educational restructuring	250
Pursuit of profit	251
Ideological motives	252
Still capitalizing on disaster	253
From schooling in disaster capitalism to critical pedagogy	255
Conclusion	256
References	256

Neoliberal educational restructuring

In the United States corporate school reform or neoliberal educational restructuring has, since the 1980's, largely saturated educational practice, policy, pedagogy, reform, organization, school culture, and curriculum. Neoliberals promote market models of reform including, particularly privatization, deregulation, business language, logic, and assumptions, and the ideology of corporate culture. Neoliberals position public schools as businesses, school districts as private markets, parents and students as consumers of a private service, and knowledge as a commodity to be bought and exchanged. Privatizers contract out management of schools to privately managed charters or for-profit educational management organizations (Byrnes, 2009; Miron, 2011; Miron and Urschel, 2010; Molnar et al., 2010; Peterson and Chingos, 2009). They seek to expand voucher schemes or neo-voucher scholarship tax credits (Welner, 2008), school commercialism initiatives, corporate-style "turn-arounds" on schools and faculty. These schemes fire entire teacher work forces and administrations, bust unions, and drain educational resources. Overwhelming evidence has established the failure of these reforms on their own terms of test score improvement as well as other terms. The standards and accountability movement works hand in hand with privatization to narrow curriculum and pedagogy to that which is numerically quantifiable and testable. These reforms have been a boon to commercialism in the multi-billion dollar test and textbook publishing industries and the technology sector. Additional corporate school reform schemes include making so-called "portfolio districts" that are charter expansion schemes that portray districts as a stock portfolio and schools as stock investments (Saltman, 2010a, 2010b), privatizing teacher education and educational leadership, gutting advanced degrees and certification in teaching and administration in favor of pay for test performance such as value-added assessment (Baker et al., 2010; Saltman, 2010c).

Higher education has been subject to corporate models of reform (Giroux, 2014; Newfield, 2011, 2016). These range from corporate models of management and the disciplining of the labor force to cutting costs and increasing revenue regardless of the relationship to the mission of these not for profit institutions. These also include managerialism and an emphasis on the quantified measurement of all aspects of the university, clientism and consumerism – the transformation of campuses to look like shopping malls, and student retention initiatives organized around entertainment, consumer demand driven program development. Corporatization also includes the vast expansion of a mid level strata of "deanlet" managers and investment in physical sites at the expense of affordable tuition, faculty pay, tenure track faculty (Ginsberg, 2013). Corporatization also involves the usurping of faculty primary responsibilities of governance over curriculum and personnel as the strata of midlevel administration expands. The emphasis on revenue generation, the abandonment of funding by states and the shifting of the cost burden onto students has been a boon to banks but it has created a vast debt bubble and a perverse incentive for students and parents to eschew subjects of study that do not have the most direct commercial implications. Hence, the STEM fields, Career and Technical Education, and business expand as the abstract sciences, the humanities, and the arts are eviscerated. The disciplines which create the conditions for societies to imagine alternative futures and support a collective conscience are at risk of extinction. Corporatization of higher education and the instrumentalism and vocationalism it promotes revises the public purposes of higher education as a public sphere committed to serving the public good and addressing the most pressing public problems.

In the past decades corporate school reformers in K12 have been aided by venture philanthropists and philanthrocapitalists such as the Gates, Walton, Dell, and Broad Foundations, Chan Zuckerberg Initiative and other LLCs, privatization lobbying outfits such as The Alliance for School Choice, and a string of right wing think tanks, consultancies, and supra-national organizations like the OECD, WB, WTO, and WEF (Saltman and Means, 2018). Privatization has been increasingly turning to both "innovative" finance schemes like Social Impact Bonds and impact investing that promote privatization and public resource skimming under the guise of performance accountability metrics and a false promise of cost savings (Saltman, 2018). As well, educational privatization has gone increasingly high tech as online programs, adaptive learning technology, and AI have been sold to districts under the aegis of technoutopian discourse that calls for massive public spending on untested technology products under a promise of innovation and hence "excellence." The new frontier of educational privatization involves the vast expansion of technology products and purveyors selling AI software, curriculum programs, adaptive learning technologies, online Social Emotional Learning programs, and

biometric teaching and tracking and digital surveillance products (Boninger et al., 2019). Tech companies seek profits through contracting and data commerce (Sadowski, 2019). The Covid 19 pandemic turbo charged this industry allowing technology companies such as google Alphabet, Microsoft, Apple, as well as educational test and textbook companies like Pearson, Kaplan, and Houghton-Mifflin to vastly expand their contracting and commercialism presence in schools.

In what follows here I address two of the primary forces driving the neoliberal restructuring of public education: profit motives and ideology. The first and second sections respectively cover these topics. The third section discusses the longstanding uses of disaster to further neoliberal aims. The fourth section contends that critical pedagogy provides a hopeful public and democratic educational alternative to the hegemonic trajectory.

Pursuit of profit

Much of the neoliberal restructuring of public education has been propelled by the pursuit of profit by for-profit corporations and investors. This is exemplified by the multimillion-dollar for-profit education management industry (Miron et al., 2012), the steady push to expand school vouchers that use public money for private school consumption, the expansion of scholarship tax credits that function like vouchers incentivizing parents to opt out of public schools, the deregulation and privatization of teacher education and educational leadership programs, and the shifting of ownership and control over curriculum and pedagogy from teachers and districts to large technology companies and billionaire-controlled corporate foundations. Such market-based reforms intersect with profit interests in the multibillion-dollar test and textbook publishing industries, the information technology and database tracking industries, the for-profit higher education sector, and the contracting industries. Test and textbook publishing companies, especially the largest such as NCS Pearson, Houghton-Mifflin, and ETS, profit from promoting the expansion of standardized testing and standardization of curriculum. These companies invest millions in lobbying for standardized curriculum and assessments at the state and national level which yields billions in revenue. Publishers pursue a variety of educational for-profit ventures, including toys, online services, special education remediation, and teacher certification. Increasingly, for-profit educational companies are becoming media companies and for-profit media companies are becoming education companies. This is evident as media executives participate in neoliberal restructuring efforts and a revolving door between school districts and media corporations. These executives take aim at public schools to grab public tax dollars with for-profit education technology. Similarly, corporate consultants such as McKinsey and ALEC influence, direct, and draft corporate school policy and reform, promoting and profiting from public-private partnership schemes. Despite the neoliberal rhetoric of freeing education from the entanglements of public bureaucracy, corporate reformers and ideologues have developed a multi-dimensional web of bureaucratic infrastructure that uses nonprofit organizations to promote privatization.

On a global scale, the World Trade Organization, International Monetary Fund, and World Bank frame public education as a private consumable, creating the conditions to redistribute governance from public to private sectors by imposing corporate trade rules on nation-states and promoting private for profit forms of educational development in poor nations and discouraging universal free public education.

The business sector has for decades framed public education in the United States as a \$600 billion per year industry, primed for the taking ("Special Report Education a New Push to Privatize 2002") or as slated for destruction and replacement by private schools. Corporate school reform aims to transform public schooling into a private industry nationally by replacing public schools with privately managed charter schools, voucher schemes, and tax credit scholarships for private schooling. The closure of neighborhood schools and massive expansion of de-unionized, nonprofit, privately managed charter schools with short-term contracts is a step toward the declaration of their failure and replacement by the for-profit industry in EMOs. EMOs extract profit by cutting teacher pay and educational resources while relying on high teacher turnover and labor precarity. Charter schools have steadily expanded annually in the past two decades. In Michigan over 80% of charter schools are operated by for profit companies that can skim out educational resources as owner profit. Trump appointed to Secretary of Education a Michigan billionaire lobbyist for school privatization who contributed to both for profit charter and voucher expansion in her state. In her term she promoted the use of public funding for private, for profit, and religious education while bailing out the for profit college industry.

Profit motives through neoliberal restructuring in education involve not just privatization but also consideration of the role of public education in social and cultural reproduction. Social and cultural reproduction involves the ways that education teaches skills and know how wrapped in class-based ideologies that are crucial for capital accumulation and the maintenance of the racialized class hierarchy. These ideologies teach social relationships and identifications that prepare students to take their class-prescribed roles in labor markets. Since the early 1980's the social and cultural reproduction role of schools has changed. Prior to the advent of the neoliberal education era social and cultural reproduction was characterized by what critical sociologists referred to as the "hidden curriculum" of schools. The hidden curriculum was the implicit teaching of skills, know-how, dispositions, and ideologies conducive to reproducing the capitalist economy. In the industrial economy the primary means of profit for capitalists came from paying workers less than the value of their labor. Working class students learned not just basic skills for the industrial economy but especially also dispositions of obedience to authority in the classroom to prepare students to submit to the boss in the workplace. Professional class students largely learned dispositions of dialogue, debate, curiosity, and investigation for them to take leadership roles in the public and private sectors. These different ideologies and dispositions created the social conditions for the replication of class hierarchy. In schools a number of mechanisms sanctified and naturalized the unequal distribution of life chances including grades, testing, sorting and sifting of students, as well as the unequal distribution of cultural capital (socially valued

knowledge, tastes, and dispositions) and the means of acquiring it. Profit was sought by capital through social and cultural reproduction by creating the conditions for labor exploitation – paying workers less for their labor than it is worth. For example, submitting to the authority of the teacher prepares the student to later submit to the authority of the boss.

In the neoliberal era beginning in the late 1970's forms of social and self control began to take different forms to reproduce the economy. The neoliberal transformation of the economy resulted in deindustrialization and the turn to financialization and service sector labor. As Nancy Fraser pointed out, the time and labor-intensive reproductive mechanisms of the industrial economy involving learned self-control such as psychological therapy, rehabilitation for prisoners, and learned dispositions of disciplinary self control for students gave way to more direct controls over the body. The supermax prison and solitary confinement, the epidemic shift to prescription mood and behavior modifications (from Xanax to Ritalin), and the roll out of the increasing militarization and prisonization of schools, the expansion of repressive pedagogies, and the focus on greater corporeal controls. As well, the mode of profit seeking through commercialism and privatization decreasingly depended on making docile and disciplined subjects for the future workforce. In the neoliberal era the long term investment in exploitable labor gives way to the short term profit in contracting. This intersection of short term profit seeking through contracting has increasingly converged with the turn to repressive pedagogical approaches such as charter models that are highly prescriptive, standardized, and homogenous. As well, the intersections of neoliberal restructuring and repression appear as biometric pedagogies and behaviorist grit pedagogies gut dialogic forms of teaching and disregarding thinking and questioning in favor of teaching to the body itself (Saltman, 2016).

In higher education profit motives are driving “partnerships” in which universities serve private business to do worker training. Students end up paying to work uncompensated for businesses. The commercial logic that is now epidemic in higher education has resulted in commercial motives driving curriculum development, program development, administrative mandates, and faculty hiring. Humanities, social sciences, and abstract sciences are threatened as disciplines with dire social consequences. As well, teacher education and educational leadership programs at the nexus of higher education and K12 schooling have been subject to privatization.

The kind of schooling envisioned by privatization advocates aims to transform a current dual system of public schooling into another dual system of public schooling. In the current dual system, elite public schools in rich, predominantly white communities prepare managers, leaders, and professionals for the top of the economy and the state while the underfunded public schools in poor, working-class, and predominantly nonwhite communities prepare the docile, disciplined workforce for the bad jobs at the bottom of the economy and for exclusion from the economy and containment (i.e., disposability). Despite the ceaseless neoliberal and liberal rhetoric of crisis and failure, the public schools (as Freire, Bourdieu, Ollman, and others recognize) do exactly what they are supposed to do: they produce a stratified workforce while sanctifying inequality as a matter of individual merit or talent. The neoliberal privatization reforms maintain the dual system, leaving in place the elite public schools but targeting poor schools and predominantly students of color to turn them into short-term profit opportunities in numerous ways: the contracting, testing, and tutoring schemes but also for-profit management and all the ancillary profits that can be generated through privatization—including the public funds that will pour into marketing charter schools to prospective “customers” through advertising and public relations, and the lucrative real estate deals needed to establish charter schools. At present, the lower end of the dual system provides a deferred investment in low-pay, low-skill disciplined workers and fodder for the for-profit prison industry and the military.

Privatization targets the low end of the dual system and pillages the public sector for short-term profits, benefitting mostly the ruling class and professional class (“poverty pimping”) while doing nothing to transform the dual system of public education into a single system as good as its best parts throughout. As William I. Robinson points out, as capital has exhausted the places for exploitation through colonization and conquest and through the neoliberal privatization of the public sphere, increasingly capital accumulation turns to pillaging the lifeworld and subjectivity and creating new investment opportunities through violence (Robinson, 2020). The steady growth of capitalizing on disaster in education discussed below (New Orleans, Detroit, Chicago, Iraq, Puerto Rico, etc.) exemplifies this trend of using and making disasters in education to create investment opportunities for capital (Saltman, 2016). Privatization increasingly takes form in education as the targeting of affect (Social Emotional Learning) and the aim of fostering “resilience,” the targeting of teachers labor by corporations (replacement of teachers with adaptive learning technology, avatars, and other AI and digital privatizations), and especially contracting arrangements for curriculum and technology products, school management services, and so on (Saltman, 2018). For investors in privatization, the benefits are double: Money can be made in the short run by draining public tax revenue, while the future exploitable workforce can still be produced for the long run. And as the investors are benefitting twice, they can feel good that they are giving poor students “every opportunity” to benefit themselves. Of course, class mobility is declining, economic inequality is expanding, and the onus to remedy social exclusion is shifting onto school kids who will need to be, in the words of Thomas Friedman, more entrepreneurial to get and keep scarcer jobs (Friedman, 2009). Such school reform ideology displaces the violence of the capitalist economy onto the weakest and most vulnerable citizens: children.

Ideological motives

The neoliberal privatization of education is propelled not only by profit motive but also by ideology. Neoliberal ideology positions education not as a public good ideally serving a democratic society or fostering humanistic ideals but as a private good for preparing the economic activity of work and consumption. Neoliberals have utilized a number of key market framing terms to cast education through the lens of the private sphere and business. These include positioning parents and students as consumers endowed with

“choice” of schools, contending that schools must “compete” against each other for consumers, suggesting that some schools should be allowed to “fail,” “just like business,” invoking business clichés of “creative destruction,” “innovation,” and “disruption.” In this narrative, the social is comprehended through radical individualism and a social Darwinian ethos. Neoliberalism sees the individual foremost as an economic actor competing against others for scarce resources. In this view, education creates the conditions for educational competition in preparation for economic competition, initially against others in the nation and then for competition against other nations. As neoliberalism suggests that there is no alternative to the market (the TINA thesis), a primary means of agency for youth is to learn to be savvy competitors and choosers of private services. Knowledge and curriculum in this view are portrayed as consumable commodities that are delivered and consumed by students. In such a view efficacious delivery becomes of paramount concern as opposed to the politics of knowledge, matters of interpretation and judgment, or investigation into the relationship between claims to truth and social authority.

Neoliberal education responds to public problems in private sector fashion, not only through the contracting out of schools and services, but by taking aim at unions, and basing schooling and administration on corporate culture. Neoliberals join metaphors of fiscal discipline with gendered and racialized demands for the disciplining of students’ bodies. In this view the natural discipline of the invisible hand can subject soft, lazy female teachers and indolent black and brown students to the necessary pressures of market competition and choice. Neoliberalism tends toward repression by modeling the school on the hierarchical and anti-democratic corporation and by denying the constitutive antagonisms that structure curriculum. Education becomes about enforcement of the right knowledge as democratic values for debate, dissent, and dialogue become casualties of the neoliberal logic of apolitical post-ideological managerialism.

The neoliberal education movement is funded and promoted by a number of super-rich individuals and owners of large businesses. These people largely believe that because the private sector has served them, the private sector should be a model for public institutions and specifically public schools. Most of the supporters of corporate school reform, despite not owning businesses or having a financial stake in privatization schemes, have been educated to believe the framing and assumptions put forward by the funders and promoters.

On their own terms of increasing standardized tests or cutting costs the corporate reforms have failed (in fact around 2010 as this was becoming apparent these reformers changed the measure of success to a vague unmeasured aim called “college and career readiness”)(Saltman, 2012). However, the corporate reforms have succeeded in reframing public discourse about education: transforming public schooling into an individualized responsibility, upwardly redistributing educational resources, gutting public schooling for the poor, making kids into commodities by privatizing schools, and radically undermining the autonomy of teachers to link knowledge to public and critical issues and to help students theorize particular experiences and contexts in ways that would facilitate political agency and public life.

Neoliberal educational restructuring has no way of dealing with several crucial problems historically facing public schooling. Privatization such as chartering and vouchers do not challenge racial segregation and white flight in urban schooling. Instead they exacerbate segregation and naturalize an apartheid system of schooling and real estate markets. Public school reform instead ought to aim to dismantle racial, ethnic, and linguistic segregation as well as class segregation as part of a broader commitment to inclusive and equal democracy. In addition, neoliberal reforms offer nothing to challenge the legacy of radically unequal funding among public school districts. In fact, the neoliberal emphasis on “performance outcomes” not only narrows the curriculum and limits the intellectual inquiry characteristic of serious learning, but, when tied to funding as it has been in the past two decades, also punishes the poor and rewards the rich.

Title I funding that has historically been used to equalize funding was reinvented through George W. Bush’s No Child Left Behind to punish those schools and children not making standardized test score improvements. Neoliberalism’s tendency for over-emphasizing quantifiable outcomes fosters the anti-intellectualism of delinking knowledge from broader contexts, historical debates, and interpretations, while it also has a natural affinity for cultural conservative agendas that emphasize a common core of the “right” knowledge that students are to learn as dogma.

Like No Child Left Behind Obama’s Race to the Top demonstrated a perverted socialism for the rich along the lines of the 2008 financial bailouts. Billions of public dollars were dangled in front of states to induce them to expand privatized and managerialist school reform including charter schools, cash for grades, turnarounds, and other schemes. This strategy of so-called “leveraging” by the U.S. Department of Education follows the corporate playbook of the venture philanthropists and was been implemented by former staff of the Gates Foundation. Such revisions imagine historically neglected schools as private enterprises that need to be subjected to the “creative destruction” of private markets. The Trump administration bailed out corrupt for profit colleges and universities and promoted K12 privatization in the form of voucher, scholarship tax credit and charter expansion.

Still capitalizing on disaster

What is at stake in corporate school reform debates is a basic commitment to public education or a basic commitment to destroy it. Privatization advocates have made up their minds that public education has failed. The declaration of the failure of public schooling forms the backbone of the privatization agenda.

Terms such as failure, choice, and competition (as well as consumers, efficiency, and monopoly) are part of a broader long-standing neoliberal agenda that extends far beyond education: misrepresenting public goods as private consumables, replacing the collective purpose of general welfare with the misguided terminology of profit accumulation, portraying citizenship as

consumerism, and alleging that public schools need to be allowed to fail just like businesses (we should suppose this “allowed to fail” logic does not apply to banking or automobile industries). The educational policy debates remain trapped in this framing, making it difficult to assert crucial democratic educational values including equality, fairness, justice, care, intellect, and the public good, to name a few.

In 2004, Bill Gates, who has championed charter schools far more aggressively than anyone else, appeared before the National Governors Association and gave a speech, a version of which was reprinted in multiple newspaper op-ed columns. Gates stated that, “Our high schools are obsolete. By obsolete, I don’t just mean that they’re broken, flawed or under-funded, although I could not argue with any of those descriptions. What I mean is that ... even when they work exactly as designed, our high schools cannot teach our kids what they need to know.” “This is an economic disaster,” he concluded, one that is ruining children’s lives and “is offensive to our values” (Colvin, 2005). Likewise, Pete DuPont of the Lynde and Harry F. Bradley Foundation (a major funder of voucher projects) described the public school system as “awful,” the worst thing the government does in America, and as “collectivism” that could be remedied by creating a market in education, and again by treating students and parents as consumers of private education. Similarly, Bruno Manno, who has served on the board of Fordham and as a senior associate at the Casey Foundation, wrote in the Hoover Institution-published Primer on America’s Schools that “the present school enterprise is not just doing poorly, but is incapable of doing much better because it’s intellectually misguided, ideologically wrong-headed, and organizationally dysfunctional” (Cohen, 2007).

Despite the frequently heard claims about evidence-based reforms, the goal of the corporate school movement is not to improve public education; rather, it is to replace public education with a privatized national system of schools competing for scarce public dollars, regularly going out of business and allowing other profit-seekers to try their hand at running schools for profit. The privatization advocates call this “creative destruction” or “churn” and make little secret of their long-term vision.

Gates’ position has changed little. At the outset of the Covid-19 pandemic Bill Gates and Andrew Cuomo took advantage of the disaster to call for an end to brick and mortar schooling and to “reimagine education” (by handing it over to the tech sector) (Strauss, 2020). The use of the pandemic to call for market based tech reform to public education follows a longstanding pattern of using disaster to impose radical neoliberal reforms without public involvement and oversight.

In my 2007 book *Capitalizing on Disaster: Taking and Breaking Public Schools*, I showed how natural and human-made disasters are being used to implement school privatization policies that could not be put in place through standard political means. I detailed the ways that Hurricane Katrina was used to put in place preformulated plans for no-bid contracting, to expand turnaround consulting, to create the largest experiment in vouchers to date, to require data reporting from the public sector to be used by the private sector, and, most significantly, to create the largest and most aggressive experiment in dismantling an entire public school district—firing all the teachers, concentrating control over hiring of teachers and administrators under a single CEO, and replacing the former public district with a largely privatized network of charter schools.

I argued in *Capitalizing on Disaster* that what was afoot in the post-Katrina Gulf Coast needed to be understood as part of a broader trend to declare educational institutions as “failed” or as “disaster areas” to justify radical unproven experiments—specifically, radical experiments with a business approach to school reform. Rather than seeing such approaches as a marginal phenomenon, we ought to understand this as a dominant trend. The declaration of “disaster” as a reason for selling off public schools to private companies was at the core of Arne Duncan’s corporate-led Chicago school reform “Renaissance 2010” (Saltman, 2007). Renaissance 2010 closed sixty schools and opened 100 privatized, de-unionized charters. This plan was continued in Chicago since Duncan but more importantly nationalized under Obama’s administration with Duncan as Secretary of Education. As well, the educational reconstruction in post-invasion Iraq in 2003 attempted to install charter schools there with multimillion-dollar profits for educational contractor Creative Associates International Inc. These plans share a two-part process of radical disinvestment or destruction of public schooling followed by investment in privatization.

Within the neoliberal view of education, declarations of “failure” have more than one function. As a rhetorical strategy, they make it seem as though the fault for low scores has to do with the low merits of students and the underperformance of teachers rather than excessive standardization, financial pressures on school systems, overstrained parents, economic disadvantages, racial discrimination, or the over bureaucratization of knowledge and unequal distribution of cultural capital.

The neoliberal declaration of a “failed system,” which relies on the metaphor of business failure, is selectively deployed and is racially and class coded. It is not leveled explicitly against rich, predominantly white communities and public schools for which high levels of historical investment and the benefits of cultural capital (affirmation in school of the socially valued class-based knowledge tastes and dispositions first learned in the home) have resulted in high achievement, traditionally defined. Rather, the declaration of system failure is leveled against the working class and the poor, predominantly nonwhite urban communities and schools. As such, it misattributes educational inequalities and shortcomings to the public sector rather than to the private sector. In the United States, the private sector bears responsibility for school underfunding as corporations capture public wealth through tax policies while not contributing to public school funding that provides their future workforce. The neoliberal solutions actively deny the ways in which the public system has a long history of business-led engineered failure, not to mention a history of unequal resource distribution by being tied to property wealth (Shipps, 2006).

When neoliberal ideologues declare the “failure” of public schooling as potentially causing a broader economic crisis for the United States in the world, they are calling for a turn to the private sector to redress the problems that too much private sector involvement in education created in the first place. Moreover, these business tycoons and ideologues mis-represent their desire for an educated workforce where workers would compete in the global economy as a universally valuable vision rather than a class-specific one that largely benefits those who own and control capital. Such logic promotes that the public sector should

subsidize job training and job preparation for the corporate sector, but that the corporate sector should not have to pay for it through taxes. Subjugating the public purposes of public schooling to primarily that of making competitive workers for the global economy presumes that the public interest is principally served by engaging in the global race to the bottom, fostered by the neoliberal vision of trade deregulation and public sector privatization. Privatizers openly talk about US students ideally becoming workers who will compete for scarce jobs against workers from poorer nations. Values of worker discipline, docility, and submission to authority are injected into the corporate school vision as they represent the ideal of the disciplined, docile, and submissive bottom tier of the workforce. This view of the national education system, serving the interests of capital in a global economy, is at odds with the global public interest that would be better served by a critical pedagogy (Giroux).

From schooling in disaster capitalism to critical pedagogy

The US invasion of Iraq in 2003 and the chartering initiatives there, post-Katrina New Orleans and the privatizations there in 2006, Chicago's radical neoliberal restructuring and the Bush administration's signature education policy were connected by the tendency of profiteers and rightist ideologues to use disaster to declare educational failure and then impose market-based solutions such as chartering, vouchers, and no-bid contracting. Since these examples took place schooling in disaster capitalism has continued in stunningly similar form. Two of the most visible examples have been the privatizations of the Detroit and Puerto Rican public schools. Both locales suffered human engineered disaster, brutal disinvestment in social services followed by crony-connected profiteering schemes. Predominantly working class African American and Latinx citizens of Detroit, as in Chicago and New Orleans were on the receiving end of a business led plan to dispossess people of their communities and schools. Education privatizers in Puerto Rico took advantage of hurricanes Maria and Irma to close public schools and open for profit charter schools while shifting governance to private hands. The continuity of schooling in disaster capitalism owes in large part to the persistence of some of the key factors that propelled it in the past. Some of these factors include crises of capital accumulation that are being resolved through force (what Robinson calls "militarized accumulation"), neoliberal ideology, and related worsening class inequality. The privatization and corporatization of public schools is part of a sustained assault on public culture and forms of life that can nurture and sustain critical consciousness and democratic social relations. School privatizations facilitated through disaster capitalism have not only worsened racial and class segregation, they have tended to expand repressive pedagogical approaches and banking education that are antithetical to critical pedagogy and democratic educational ideals. Critical approaches to teaching and learning provide students the tools to interpret knowledge and experience in relation to broader social, cultural, economic, and political antagonisms and offer dialogic forms of teaching in which knowledge can form the basis for acts of interpretation to become acts of social intervention, agency, and collective self governance. Critical pedagogy can make schooling in disaster capitalism an object of critical analysis by comprehending the material and symbolic interests at stake in it. Critical pedagogy fosters in students the tools of social criticism and development as critical intellectual citizens who can learn to govern themselves and others, and who can deliberate on shared values beyond the shallow imperatives of consumerism and the desperate scramble for a shrinking pool of jobs.

One of the most significant aspects of disaster capitalism has been the growing existential urgency of the climate crisis and the failure of business and political elites to address it. Along with rebounding nuclear proliferation the climate catastrophe stands as the gravest threat to the continued existence of the species and life on the planet. However, the devastating effects of climate inaction will not be experienced equally. The poor who have historically contributed the least to global warming stand to be the most gravely effected by sea level rise, land displacement, and extreme weather events while the rich stand to profit from the destruction. Critical pedagogy can recognize the ongoing tendencies of elites to use disaster to achieve what are largely unpopular educational projects such as school privatization schemes, public school closures and disinvestments, mass firings of teachers and union busting, and public failure declarations. As well, critical pedagogies can link the broader politics and economics of global warming to the desperate need to transform consciousness about the urgency of environmentally sustainable practices. Critical pedagogies can relate environmental sustainability to comprehension of the utter inseparability of global warming from the capitalist imperative for profit and growth at any cost as well as the racialized dimension of the class project driving planetary destruction.

Industrial era uses of public education for capital involved a long term investment in reproducing the labor force by teaching students skills and know how wrapped in class based ideologies. This form of reproduction has partly given way to capital's use of privatization and contracting from controlling students' bodies in more direct forms (Saltman, 2016). This post-fordist trend of short term profit taking through coercion and control has become more pronounced as new technologies including adaptive learning, biometric pedagogy and AI have been used by education, media companies, and corporate philanthropies to commodify students and use them as data production engines (Saltman, 2018). The big business in repressive pedagogies is wrapped up with contracting profits as "resilience," "grit," quantified measures of "social emotional learning" and revived techniques of behaviorism, and high tech tracking and behavior modification devices continue to proliferate in schools (Saltman, 2016). The charter and philanthrocapitalist profiteers and ideologues such as KIPP, CZI, and Gates have played no small part in promoting repressive pedagogies. As well, the high stakes testing trend has been the basis for profiteering not only by test and textbook companies but by the pharmaceutical industry that sold teachers, students, and parents ADHD medications to game the tests. These new technologies of control are utterly antithetical to the aspirations of critical pedagogy. They treat learning as a performance (typified by edTPA and biometric webcam pedagogies) measured by an effect on the body rather than by fostering forms of learning through dialogue, thinking, theorizing, and mediating experience. Knowledge in these repressive pedagogies is not comprehended as meaningful for its social and political uses – the means of interpreting and acting on the social world and public problems. Rather it is

decontextualized from the social relations and broader social forces that make it meaningful and seen instead as a means of academic exchange. A crucial imperative for critical pedagogy today is to develop strategies to counter the ways that the time, space, curriculum, and pedagogy of schools has been captured by the technologies of repression that have been facilitated by neoliberal restructuring. This requires denaturalizing and politicizing knowledge, curriculum, and pedagogy in schools as well as technologies and the ideologies that they promote that often appear beyond question by being conveyed through machines (e.g. biometrics), algorithms (e.g. adaptive learning), and pseudo-scientific lingo (e.g. quantification of social emotional learning). It is imperative that critical pedagogical efforts be linked to the growing broader movements by teachers and communities to reclaim teacher autonomy and reject corporate school reform.

Another significant development in disaster capitalism has been the ways that around the world right-wing populists and authoritarian leaders have scapegoated racial, ethnic, religious, and sexual minorities and women for the structural and systemic causes of the destruction wrought by decades of neoliberal globalization. Authoritarians have revived cynical essentialist views of identity falsely equating race, ethnicity, and gender with natural essence, reviving white supremacy and retrograde stereotypes. Such identity essentialism trades in false guarantees of certainty in a time of economic and political precarity. Critical pedagogy provides a hopeful way for education to be understood as the practice of cutting through ideological mystifications that falsely position oppressed as oppressor and oppressor as oppressed. Both liberal and conservative educational discourse have created the conditions for authoritarians to seduce people to reactionary ideologies, to delegitimize science and educational institutions, and to propagate post truth relativism in which truth is only grounded in the authority of the claimant (is this not the logic of the standardized test with the maker of the test and the underlying assumptions beyond question?). By rejecting educational theory and the politics of education liberals and conservatives create the conditions for authoritarian rather than democratic culture. Liberal and conservative approaches to education reject the need for theorizing practice and instead embrace depoliticized and radically empiricist approaches to teaching that misrepresent knowledge as neutral, apolitical, and having nothing to do with broader material and symbolic antagonisms or subjective experience. In this view knowledge is not produced through dialogue, debate, investigation and curiosity but rather transmitted from “the ones who know” and deposited in the ones who don’t. On the contrary, critical pedagogy allows students to theorize and interpret not only the self as a social product, the social as constituted by class and cultural antagonism, but also knowledge and learning in relation to the social and political forces and struggles that make them meaningful. As such critical pedagogy creates the conditions for knowledge to be a source of political agency (Giroux, 2011). Despite the continuity of schooling in disaster capitalism, the development of new technologies of repression, and the resurgence of reactionary ideologies and political movements there is always the possibility of hope even in uncertain times.

Conclusion

All educational work is a form of cultural production that inevitably makes social meanings and contributes to affirming or contesting existing meanings and broader public discourses. In that sense uncertainty at the core of signification represents not a threat but an opportunity in every meaning-making encounter to contextualize truth claims in terms of broader social relations and interests, to theorize experience, and to work for education to become the basis for opposition to disaster capitalism and for broader struggles for democratic social relations in all institutions and the basis for shared learning and labor for shared benefit.

Struggles against neoliberal restructuring of education at all levels have to be comprehended as more than representational. They must be comprehended as part of a broader global struggle to expand democratic social relations throughout culture, economy, and politics and to expand the universal provision of vital public goods and services including education (ideally in critical forms), healthcare, housing, water, and food. Such struggles cannot expect or depend upon the beneficence or charity of the very people who stand to benefit most materially from the maintenance of the neoliberal order such as venture philanthropists and technology corporations and their foundations. Struggles for justice entail linking critical educational practices inside and outside of formal educative institutions and taking seriously the pedagogical and political force of cultural institutions. It also demands linking such projects to struggles for justice, equality, and freedom expressed in social movements for labor and economic redistribution, racial, gender, ecological, and other struggles for justice.

References

- Baker, E.L., Barton, P.E., Darling-Hammond, L., Haertel, E., Ladd, H.F., Linn, R.L., Ravitch, D., Rothstein, R., Shavelson, R.J., Shepard, L.A., 2010. Problems with the Use of Student Test Scores to Evaluate Teachers (EPI Briefing Paper #278). Retrieved from: <http://www.epi.org/publication/bp278/>.
- Boninger, F., Molnar, A., Saldana, C., 2019. Personalized Learning and the Digital Privatization of Curriculum and Teaching April 30. National Education Policy Center available at: <https://nepc.colorado.edu/publication/personalized-learning>.
- Burch, P., 2009. Hidden Markets: The New Education Privatization. Routledge, New York, NY.
- Byrnes, V., 2009. Getting a feel for the market: the use of privatized school management in philadelphia. *Am. J. Educ.* 115, 437–455.
- Cohen, R., 2007. Strategic Grantmaking: Foundations and the School Privatization Movement. Retrieved from: http://www.ncrp.org/index.php?option=com_oxocart&Itemid=41&p=product&id=4&parent=3.
- Colvin, R.L., 2005. A new generation of philanthropists and their great ambitions. In: Hess, F. (Ed.), *With the Best of Intentions*. Harvard Education Press, Cambridge, MA, p. 21.
- Dillon, S., August 9, 2010. Inexperienced companies chase U.S. School funds. *N. Y. Times*. Retrieved from: http://www.nytimes.com/2010/08/10/education/10schools.html?_r=0.
- Friedman, T.L., October 20, 2009. The new untouchables. *N. Y. Times* A31.
- Ginsberg, B., 2013. *The Fall of the Faculty*. Oxford University Press, Oxford.

- Giroux, H.A., 2011. *On Critical Pedagogy*. Continuum, New York.
- Giroux, H.A., 2014. *Neoliberalism's War on Higher Education*. Haymarket Books, Chicago.
- Miron, G., 2011. Review of Charter Schools: A Report on Rethinking the Federal Role in Education. National Education Policy Center, Boulder, CO. Retrieved from. <http://nepc.colorado.edu/thinktank/review-charter-federal>.
- Miron, G., Urschel, J.L., 2010. Profiles of Nonprofit Education Management Organizations: 2009–2010. National Education Policy Center, Boulder, CO. Retrieved from. <http://nepc.colorado.edu/publication/EMO-NP-09-10>.
- Miron, G., Urschel, J., Aguilar, M.A.Y., Dailey, B., 2012. Profiles of For-Profit and Nonprofit Education Management Organizations: Thirteenth Annual Report - 2010–2011. Retrieved from. <http://nepc.colorado.edu/publication/EMO-profiles-10-11>.
- Molnar, A., Miron, G., Urschel, J.L., 2010. Profiles of For-Profit Education Management Organizations: Twelfth Annual Report, 2009–2010. National Education Policy Center, Boulder, CO. Retrieved from. <http://nepc.colorado.edu/publication/EMO-FP-09-10>.
- Murray, C., May 4, 2010. Op-ed: why charter schools fail the test. N. Y. Times. Retrieved from. <http://www.nytimes.com/2010/05/05/opinion/05murray.html>.
- Newfield, C., 2011. *Unmaking the Public University*. Harvard University Press, Cambridge.
- Newfield, C., 2016. *The Great Mistake*. Johns Hopkins University Press, Baltimore.
- Peterson, P.E., Chingos, M.M., 2009. Impact of For-Profit and Nonprofit Management on Student Achievement: The Philadelphia Intervention 2002–2008 (Working Paper PEPG 09–02). Harvard University, Program on Education Policy and Governance, Cambridge, MA.
- Robertson, S.L., 2006. Globalisation, GATS, and trading in education services. In: Kall, J., Rinne, R. (Eds.), *Supranational Regimes and National Education Policies—Encountering Challenge*. Finnish Education Research Association, Helsinki, Finland.
- Robertson, S.L., Dale, R., 2009. The World Bank, the IMF, and the possibilities of critical education. In: Apple, M.W., Au, W., Gandin, L.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, NY, pp. 23–35.
- Robinson, W.I., 2020. *The Global Police State*. Pluto Press, London.
- Sadowski, J., January 17, 2019. When data is capital: datafication, accumulation, and extraction. *Big Data Soc.*
- Saltman, K.J., 2007. *Capitalizing on Disaster: Taking and Breaking Public Schools*. Routledge, New York.
- Saltman, K.J., 2010a. *The Gift of Education: Public Education and Venture Philanthropy*. Palgrave Macmillan, New York, NY.
- Saltman, K.J., 2010b. Urban School Decentralization and the Growth of 'Portfolio Districts'. The Great Lakes Center for Education Research and Practice. Retrieved from. http://www.greatlakescenter.org/docs/Policy_Briefs/Saltman_PortfolioDistricts.pdf.
- Saltman, K.J., 2010c. "Value Added" Assessment: Tool for Improvement or Educational "Nuclear Option"? Retrieved from. <http://truthout.org/archive/component/k2/item/91775: valueadded-assessment- tool-for-improvement-or-educational-nuclear-option>.
- Saltman, K.J., 2012. *The Failure of Corporate School Reform*. Routledge, New York.
- Saltman, K.J., 2016. *Scripted Bodies: Corporate Power, Smart Technologies, and the Undoing of Public Education*. Routledge, New York.
- Saltman, K.J., 2018. *The Swindle of Innovative Educational Finance*. University of Minnesota Press, Minneapolis.
- Saltman, K.J., Means, A.J., 2018. *The Wiley Handbook of Global Education Reform*. Wiley Blackwell, New York.
- Shipp, D., 2006. *School Reform Corporate Style: Chicago 1880–2000*. University Press of Kansas, Lawrence, KS.
- Smarick, A. (2010). The turnaround fallacy. *Education next*, 10, 1. "Special Report education a new push to privatize." (2002, January 14). *Bus. Week*.
- Strauss, V., 2020. Cuomo questions why school buildings still exist – and says New York will work with Bill Gates to 'reimagine education'. *Wash. Post*. May 6.
- The Cost of a Soldier Deployed in Afghanistan, 2011. *Marketplace Morning Report*. Retrieved from. <http://marketplace.publicradio.org/display/web/2011/02/22/am-the-cost-of-a-soldier-deployed-in-afghanistan/>.
- Tooley, J., 2009. *The Beautiful Tree*. Cato Institute Press, Washington, DC.
- Toppo, G., October 1, 2009. Union claims Filipino teachers held in "Virtual servitude" in Louisiana. *USA Today*. Retrieved from. http://usatoday30.usatoday.com/news/education/2009-10-01-filipino-teachers_N.htm.
- Wehner, K.G., 2008. *Neo-Vouchers: The Emergence of Tuition Tax Credits for Private Schooling*. Rowman and Littlefield, Lanham, MD.

Linking research and teaching to benefit student learning in higher education: complexity, collaboration and a tailored approach

Matt O'Leary, Birmingham City University, CSPACE Research Centre, Faculty of Health, Education and Life Sciences, Birmingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	258
Understanding the importance of complexity and context in teaching and learning	259
Research context and methodology	260
Project 1 – Cycle of Peer Observation (CoPO)	261
Restoring professional agency through CoPO	262
Project 2 – Improving teaching and learning through collaborative observation	263
Developing classroom consciousness between students and staff	264
Building collaborative relationships between students and staff	265
Concluding thoughts and future directions	266
Acknowledgment	267
References	267

Introduction

The quality of teaching has only recently emerged as a policy priority in higher education, with some governments such as the UK even identifying it as a weakness (e.g. [Gill 2015](#)). Historically, teaching has occupied the peripheries of the policy agenda in higher education, with research the priority focus and the dominant benchmark against which to measure the quality of academics' work. High quality research and the prestige associated with it has long been regarded as the gold standard for recognizing and rewarding academic excellence, as well as having an impact on the recruitment of staff and students internationally (e.g. [Blackmore et al., 2016](#)).

In recent decades, the economization, marketization and massification of higher education have all played their part in creating an increasingly competitive environment, with institutions forced to compete against one another for students. It is in the context of this environment that the quality of teaching has surfaced as an area of policy focus, with research into teaching and students' learning having an increasingly important role to play in informing the debate, particularly in terms of research evidence being used to support teachers and their students in better understanding and ultimately improving their teaching and learning. While the vast majority of research in higher education focuses on investigating discipline/subject specific matters and is rarely concerned with researching the teaching of that discipline/subject, education research is very much an outlier in that sense. In other words, investigating the teaching of the subject as well as the subject itself is one of the key lines of inquiry of education research.

To date, politicians and education policy makers have tended to favor positivist models of experimental research design that typically consist of large-scale quantitative studies underpinned by metrics-based approaches to data collection and analysis. This has particularly been the case among policy makers globally in the context of improving the effectiveness of teaching in schools and the wider global education reform movement (e.g. [Meyer and Benavot 2013](#); [Sahlberg, 2011](#)). Such approaches invariably conceptualize teaching and learning through an instrumentalist, technicized lens, viewing them reductively as separate, linear processes with observable and measurable correlations between the input (teaching) and the output (learning). These approaches also tend to generate and/or test out performance criteria which are used to evaluate and codify the quality of teaching and learning accordingly, with an overreliance on linear systems of quantitative performance data that atomize the complexity of the teaching-learning interrelationship in a way that is not only reductionist but also decontextualized.

It is perhaps unsurprising that politicians and policy makers have favored quantitative research studies in education, as they generate data that feed models of governance and regulation that are based on what [Bamber \(2020\)](#) refers to as "metrication," which she describes as the "systemic use of metrics primarily for managerial and accountability purposes." Such systems have dominated higher education policy in recent years, where ease of management and measurement of data have become the key drivers. [Steven-son \(2017, 538\)](#) succinctly captures the current climate of education policy making and the dominance of metrics-based approaches to understanding and improving educational provision when he comments that "the measurement of everything is central to the modern educational experience." However, the fact that such approaches have become normalized and embraced by education systems worldwide neither validates their effectiveness in enhancing knowledge and practice in the field of teaching and learning nor justifies policy makers' overreliance on them. In his critique of what he refers to as the "what works" evidence-based practice movement in education, [Biesta \(2017, 316\)](#) asserts that:

While the performance of many aspects of the education system is measured in much detail and with much precision and statistical sophistication, the question that remains is whether this brings us any closer to an understanding of the value of the processes and practices that are being measured.

Building on Biesta's thinking, I would argue that not only is the question of value important to consider but so too is that of the efficacy of positivist quantitative research studies in education and their capacity to advance understanding and applied practice. The ontological, epistemological and methodological positioning adopted in this article is based on the premise that teaching and learning are not separate processes but interconnected and inform one another in a complex, nuanced and iterative way. This contrasts with positivist approaches that conceptualize the processes of teaching and learning as individualistic, performance products that can be itemized into quantifiable units.

One of the consequences of this dependence on quantitative metrics in education and indeed other public services is that qualitative data is often side-lined, seen as an additional source to the main quantitative data sets or even inferior, which not only exacerbates this reliance on quantitative metrics but actually undermines the value of qualitative data in its own right. This exemplifies what Fricker (2007) refers to as "epistemic injustice." It also reinforces Ball's argument that (2015, 299), "we are subject to numbers and numbered subjects" as a result of what he refers to as the "tyranny of numbers." It is not within the scope of this article to present an in-depth critique of the shortcomings of positivist quantitative approaches to pedagogic research and the improvement of teaching and learning. Instead, I want to focus on exploring the value of educational research as a collaborative, discursive process that seeks to illuminate our understanding of the complexity of teaching and learning in situated contexts.

As a counter narrative to the hegemony of positivist quantitative education research, this article draws on qualitative data from university teachers and students involved in two distinct yet interrelated projects focused on improving teaching and learning in higher education. The key findings from these data help to support my call for the need to move away from reductionist approaches that seek to conceptualize effective teaching and learning as linear, technical processes to be mastered and instead embrace more collaborative approaches between staff and students that acknowledge the contingent and emergent nature of the teaching-learning interrelationship.

The article begins by considering the importance of complexity and context in understanding teaching and learning and the relationship between the two. It then moves on to discussing a longitudinal body of work from two recent, interrelated research projects in higher education that adopted an innovative, inquiry-based approach to using observation as a platform for university teachers and students to collaborate on improving their teaching and learning. The article concludes by reflecting on how such collaborative approaches to researching teaching and learning can open up new opportunities for developing and informing collective understanding between teachers and students about what effective teaching and teaching looks like and how such understanding can feed into improving the experience of both.

Understanding the importance of complexity and context in teaching and learning

Complexity theory can provide a useful lens through which to view and conceptualize the teaching-learning interrelationship. Complexity theory argues that many natural and social systems cannot be captured and represented in sequential, linear relationships, as they comprise complex processes that are not only difficult to predict but equally do not lend themselves to reductive analysis or simplistic representation. In addition, complex systems are open (Cilliers, 1998), which means that their boundaries are permeable, with interactions extending beyond the confines of the sites in which they occur (e.g. classrooms, lecture theaters, virtual workspaces) and involving the wider environment. It is therefore unsurprising that teaching and learning interactions share many of the characteristics associated with complex systems. For example, the permeable boundaries of a complex system are all too apparent in teaching-learning scenarios, in person or virtually. The boundaries of real or virtual classrooms, for instance, extend beyond their spatial and temporal constituents. Besides, a multitude of external factors can potentially impact on students' engagement with learning, as indeed they can on educators' teaching.

Just as teaching and learning are part of a complex interrelationship that is made up of multiple elements, so too are they socially situated, intellectual activities (Lave, 1993). What this then means is that if we are to develop a rich understanding of teaching and learning in higher education, any such understanding needs to embody this complexity to be able to derive true meaning. As Lave's work (1988, 1993) demonstrates, meaning is inherently relational, generated in and through the interaction between teachers and their students in teaching and learning encounters. In higher education, experiences and understandings of teaching and learning tend to be shaped by disciplinary cultures and practices (Neumann, 2001). Although disciplinary knowledge is created through collective intellectual work, one of the biggest influences on academics' pedagogic knowledge and skills is often their own personal educational experiences.

Teaching in higher education is a multi-layered concept that has moved a substantial way from the traditional lecturing mode that has been associated with the sector for so long. It now typically involves a multitude of differing scenarios, sites, interactions and agents, particularly in the case of practice-based courses that incorporate work-based placements and are invariably influenced by the requirements of professional regulatory bodies. It can also be viewed and understood through multiple lenses e.g. the performance, the process and the scholarship. This is also reflected in how the relationships between teachers, students and the teaching-learning interactions orient themselves (e.g. Su and Wood, 2012). Previous research has identified traits of teaching excellence (e.g.

Skelton, 2004; Shephard et al., 2010), the most commonly occurring of which include the complex interactions and negotiations between aspects of technical, emotional and ethical deliberations. Whatever lens or focus we choose to adopt when discussing teaching, it is clear that we are dealing with a complex process that does not naturally lend itself to being neatly categorized according to an inherently reductive term like “excellence,” as it can only serve to dilute and simplify that complexity rather than attempt to capture or understand it.

In recent years, mass participation in higher education has meant that students come from a more diverse range of socioeconomic, cultural and educational backgrounds, with extensive needs, thus resulting in a highly complex teaching and learning environment. These students also bring with them their lifetime experiences of learning and their understandings of teaching and learning, which increases the importance of the need for staff to engage in ongoing development of their teaching.

To date, the way in which many universities have sought to capture students’ views on teaching and learning has tended to focus on a customer satisfaction approach of the product, rather than allowing for substantive involvement in discussions about the process and ways of improving teaching and learning. In contrast to the marketized conceptualization of students as consumers and staff as providers, which has become commonplace in higher education in recent years, this article argues that students and staff need to be viewed as members of their course community who have agency and are active participants in understanding and shaping teaching and learning in their respective communities. As Bowden and Marton (2004, 14–15) have discussed, an understanding between students and staff based on a common frame of reference of teaching and learning is fundamental to building a collective consciousness of learning in the context of their program:

Learning from other people means that we become aware of their ways of seeing things, regardless of whether or not we are convinced by, or appropriate, their ways of seeing ... this means that not only do students have to learn from teachers but teachers have to learn from students as well ... Our views of a certain phenomenon can therefore be shared or they can be complementary. Combining differing views implies richer, more powerful, ways of understanding a phenomenon or a situation and is likely to offer more options for handling varying conditions.

Viewing subject specialist learning through others’ eyes can help us to develop a mutual awareness and understanding, which in turn encourages us to appreciate, challenge and further our individual and collective understandings. Teaching and learning are social practices that require the key agents (i.e. teachers and students) to engage in a process of reflexivity. Examining our understandings and experiences of teaching and learning by cross referencing these assumptions and opening them up to dialogic exchange enables us to become aware of the strengths and areas for development in our practices. Instead of generating and cascading feedback to each other about teaching and learning, as is the norm in approaches where evaluation is the key driver, this article argues that by creating shared spaces for teachers and their students to engage in reflexive dialog, this leads to collective sense making and meaning (Fielding, 2004), as exemplified by the approach discussed below.

Research context and methodology

This article draws on a longitudinal body of work from two discrete projects, both situated in a modern English university and with a focus on improving teaching and learning. The first was an internally funded project involving academic staff working in collaborative pairs/small groups as part of a new *Cycle of Peer Observation* (CoPO) implemented in 2016/17 and which was still in place at the time of writing this article. The second was an externally funded project, *Improving teaching and learning through collaborative observation*, undertaken from 2016 to 2018 and funded by the Higher Education Funding Council for England. It comprised five case studies of undergraduate programmes across a single faculty, with each case study including small groups of academic staff and student participants. What links these two projects and their accompanying body of work is the application of observation as a shared frame of reference for participants to explore their teaching and learning experiences, to identify and discuss the critical aspects from these experiences and to reciprocally reflect on their understanding of them. In the case of the first project, this was between academic staff and in the case of the second project, this was between academic staff and their students.

Recent research in the field of observation has revealed how performative, assessment-based approaches to observation that are designed to monitor and measure teaching performance often result in counterproductive consequences in relation to improving the quality of teaching and learning and often act as a deterrent to developing innovations in practice (e.g. O’Leary, 2020; O’Leary and Wood, 2017). Learning from the drawbacks of how observation has been misappropriated as a performance management tool therefore provided an important platform for shaping the conceptual and methodological framework of both projects discussed in this article.

The observation approach used in both projects was designed to prioritize collaborative inquiry, collective reflection, professional dialog and collegial development. The starting point for this design involved a significant reconceptualization and reconfiguration of the way in which the participants would engage with observation as a method for studying and enhancing teaching and learning. A key element to overcoming some of the limitations of conventional approaches to observation was severing the link between observation and its use as a method of assessing teaching and teacher performance, as I was convinced that unless observation was removed from the assessment context, this would jeopardize efforts to capture situated examples of authentic teaching and learning and in turn to create a safe, trusting and collaborative environment for reflection and dialog between staff and between staff and students in both projects.

What follows are vignettes from each project of the conceptual and methodological design, along with illustrative examples of key themes to emerge.

Project 1 – Cycle of Peer Observation (CoPO)

Observation of teaching practice was conceptualized and applied as a common touchpoint and catalyst for dialog between pairs of teaching staff (observer and observee). The Cycle of Peer Observation (CoPO, see Fig. 1) was developed as a process to facilitate effective dialogic interaction between observer and observee. Observers were instructed to compile a descriptive log/field notes of what they observed along with associated questions and/or reflective comments to discuss during the professional dialog stage rather than make evaluative judgments. It was made clear to all staff that observers were not there to provide answers or solutions to all their questions but to support them as peers to reflect on their practice.

All observers undertook an intensive one-day observation training program and were required to organize a reciprocal observation with a fellow observer as a means of practising their skills before formally undertaking their roles. The first part of the training focused on staff discussing and critically reflecting on their understandings, perceptions and prior experiences of observation. We have since come to articulate this stage of the training as a process of *conceptual catharsis*, having learned that unless academic staff are allowed to divest themselves of prior experiences and associations with observation in an assessment context, then expecting them to engage with a reconceptualized approach is likely to be more problematic and ultimately less successful. In short, encouraging teaching staff to detach observation from assessment and embrace it as an exploratory tool requires a significant shift individually and collectively, and it is important to integrate time for this in the training.

The second part of the training provided staff with opportunities to develop an understanding of the epistemological and methodological underpinnings of CoPO, to discuss the processes of the cycle, as well as practice some basic coaching and observation skills through a series of interactive tasks. For example, they practised carrying out non-judgmental observations via a selection of video clips of teaching and compared their observation notes and reflections with their peers. This was to help to develop an awareness of what different observers notice when they observe, reflect on their values of learning and teaching and how these values inform their observations and reflections. They also practiced conducting conversations using coaching questions.

As Fig. 1 shows, there are seven stages to CoPO. Stage 1 provides the observee with an opportunity to think holistically about their teaching through self-reflective writing, encouraging them to reflect on their strengths and identify areas for further professional learning. This self-reflection is shared with the observer and forms a catalyst for professional dialog in the pre-observation meeting (Stage 2), where they have the opportunity to establish a professional rapport and explore themes articulated by the observee in their self-reflection. Where appropriate, the observer uses coaching questions to establish an ethos of inquiry as the observee articulates their chosen observation focus, and both agree the logistical details of the observation.

Stage 3 is the observation, which typically lasts about an hour. In contrast to assessment-based models of observation, observers are not required to use a proforma when taking notes as there is no checklist of behaviors or assessment criteria. They are simply asked to record what they see, with the understanding that they should avoid judgment and evaluation. Their goal is to act as “detached awareness raisers” (Whitmore, 2002, 42), a lens through which the observee can gain an additional view of the events of the observation and to use the observer’s notes as a springboard for collegial dialog.

Following the observation, individual reflections (Stage 4) are carried out by both observer and observee. The observer further develops their observation notes to add questions and comments in preparation for the post-observation meeting (Stage 5). Both bring their observation notes and Stage 4 reflections to the Stage 5 discussion and relate these back to what was discussed in Stage 2

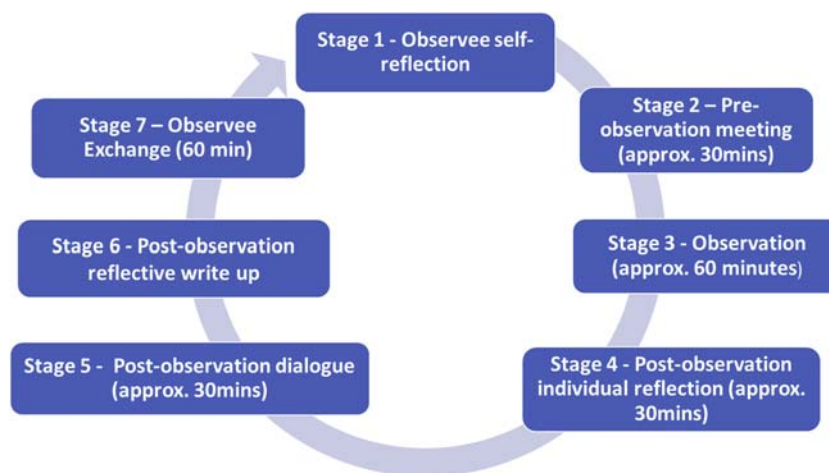


Fig. 1 Cycle of Peer Observation (CoPO).

in terms of the agreed focus of the session. The observer draws upon non-judgmental coaching skills as they discuss the learning and teaching they experienced and observed, culminating in the observee's formulation of teaching objectives which encapsulate the learning they have taken from the observation process and steps for developing their practice.

Stage 6 requires both observer and observee to complete their reflective write-ups independently. The aim is to capture the holistic experience of both parties of the cycle, making connections between this observation experience and their past learning and teaching experiences. The observee is also expected to document the objectives identified in the Stage 5 meeting and reflect on how they intend to move their teaching forward, along with the resources/support they might need to enable them to do so. Several months after the completion of their observations, both observers and observees are invited to a focus group with their peers (Stage 7). The purpose is to move the focus away from the observation as an episodic event by creating further opportunities for dialogic professional learning as they share the insights gained and the impact these have had on both their teaching and their students' learning.

Restoring professional agency through CoPO

Establishing rapport between observer and observee is essential to create an atmosphere of trust and confidence to openly discuss and reflect on practice, particularly for colleagues who may not know each other well. Observees overwhelmingly felt that the observers achieved this; the words they used repeatedly to describe their observer's approach were "supportive," "friendly" and "professional." Echoing [Buskist, Ismail, and Groccia's \(2014, 50\)](#) emphasis on the importance of structure and planning for "intentional discussion of pedagogy," both observers and observees remarked on how CoPO provided a useful framework for focused discussion. In particular, they underlined the "vital role" of the pre-observation meeting (Stage 2) in the process. It allows the observer to clarify the purpose of the observation and to dispel any myths that may exist about the performative nature of the process. An observer from radiotherapy commented:

The pre-observation meeting has been so important in agreeing and clarifying the focus and unpicking the observee's thinking and rationale for their focus. It's been interesting how deep some of those discussions about teaching have been too.

Equally, from the observees' perspective, the pre-observation meeting "set the tone" by encouraging them to take the lead in the discussion and its focus, while also helping to establish the roles, responsibilities and expectations of each party in the process.

Observees valued the professional trust given to them. As one teacher education lecturer commented, "the fact that we have the opportunity to choose our focus and not have to pander to others' agendas is very empowering to me." Comments like this were common among observees across different data sets, reinforcing the importance of professional agency in the process and how this is valued by academics (e.g. [McMahon, Barrett, and O'Neill, 2007](#)).

The following astute remark by a nursing lecturer encapsulates the success of this project:

I specifically chose a session I wasn't comfortable with and I found this process to be much more supportive than previous experiences ... my observer wasn't imposing their view of what they consider to be effective teaching on me and I really appreciated that because I think I'm probably the best person to know how to change [my teaching] ... I just needed somebody to help me find the best way to do it.

She went on to describe how she used the observation as an opportunity to focus on an area of her teaching that she felt was "too didactic," which had previously emerged as a problem based on student feedback and a lack of student engagement during sessions. As a result of her observation discussion, she trialled more interactive teaching approaches, which were very well received by students. She subsequently went on to change all her teaching sessions to promote more interaction on the basis of her CoPO experience.

Another spoke of being in a small team who all observed each other and that "the observations generated many conversations outside of this process ... The whole ethos of the course has changed and the observation process was the catalyst for that."

Like the staff in [Tenenberg's study \(2016\)](#), observers found the process of observation, reflection and professional dialog was also beneficial to their own practice: "It was really quite rewarding as an observer; I learned more from the observee than they ever would have learned from me if I were forced into the role of trying to advise." Another said "being an observee made me reflect upon my own practice with greater breadth and question some things I do that have become unquestioned habits."

To conclude, CoPO completed its second full iteration at the end of the academic year 2017/18, with approximately 300 academic staff across the faculty having undertaken at least one full cycle. Survey results revealed high levels of engagement and satisfaction with the new approach, with 93% of observee respondents (strongly) agreeing with the statement, "The observation cycle helped me reflect on my teaching and student learning." When asked to use a slider scale to answer the question, "Was the observation process more informative/developmental or performative/judgemental?" the average rating was 92/100 in favor of informative/developmental. This was echoed in the qualitative comments of staff, with one lecturer describing it as a "genuinely democratic process that values individuals for the strengths they have, rather than an attempt to identify and criticize perceived

weaknesses” and another that it was an “incredibly positive and supportive approach to observing and reflecting on our teaching and students’ learning.”

Project 2 – Improving teaching and learning through collaborative observation

The project comprised five case studies of undergraduate programmes across a single faculty, with each case study including at least two academic staff and two student participants for each observation cycle. The five case studies included in the project were: (1) Adult Nursing; (2) Child Nursing; (3) Early Childhood Studies; (4) Primary Education and (5) Radiotherapy. As part of the Cycle of Collaborative Observation (CoCO), students and staff were all encouraged to take an active role in critically reflecting on their practices, viewing the “same” classroom experience from their perspective and exchanging their observations and reflections with each other. The project methodology was designed to draw individual perspectives together to observe teaching and learning at a course-specific level rather than focusing attention on an individual’s practice or a one-off session. Central to the project’s philosophy of improving teaching and learning was the need for students and teaching staff to take shared responsibility for developing mutual understanding, using a shared frame of reference from which to generate new understandings of situated teaching and learning situations.

The key driver for each of the project’s five case studies was to provide a platform for adopting a collaborative, inquiry-based approach for students and staff to co-investigate cognate issues relating to their respective programmes. Bowden and Marton’s (2004) “collective consciousness” was a helpful concept to consider the learning that takes places for individuals and for the case study as a collective. This concept shares the critical social constructivist stand we adopted, emphasizing how teaching and learning should be built on an ethos that challenges our taken-for-granted views and practices in order to develop new understandings and subsequently make improvements (Fig. 2).

CoCO began with two separate training sessions for staff and students, each session lasting 5 hours in total. Staff training focused on their previous experience of observations, developing an understanding of the project’s collaborative approach and discussing the stages of CoCO. They worked collaboratively on a range of immersive exercises to facilitate their familiarization with the methodology and the activities involved. While there were many similarities with the student observer training, there was a shift in focus insomuch as the students’ session was concerned with helping them to think about their position as students through the lens of learning, to reflect on how they engaged with teaching and learning and how that connected to their knowledge as a whole, as well as their experience as a student on their course.

Stage 1 of the cycle involved staff and students working with a set of prompts to compile individual reflective accounts of their learning/teaching autobiographies. Students were prompted to reflect on their prior experiences of learning and how they went about making sense of their experience as learners on their courses. Staff were prompted to reflect on their teaching and learning journeys by focusing on what their philosophies of teaching were, how these impacted on the preparation, delivery and evaluation of teaching and learning, how these connected to their philosophies of learning, and what they considered to be the relationship between teaching and learning on their respective courses.

Stage 2 *Pre-observation meetings*, students and staff first met in their peer groups to discuss their individual reflections from Stage 1 and worked together on developing their learning objectives and inquiries for their case study. Aligned with our conceptualization of learning, we adopted the term “coaching” and underpinned this with Bokeno’s (2009) work, who suggests that organizations focus on “learning relationships,” which are based upon an ethos of inquiry and a willingness to explore unasked questions. Using a coaching ethos and techniques, members in each peer group facilitated each other’s development through the use of open and

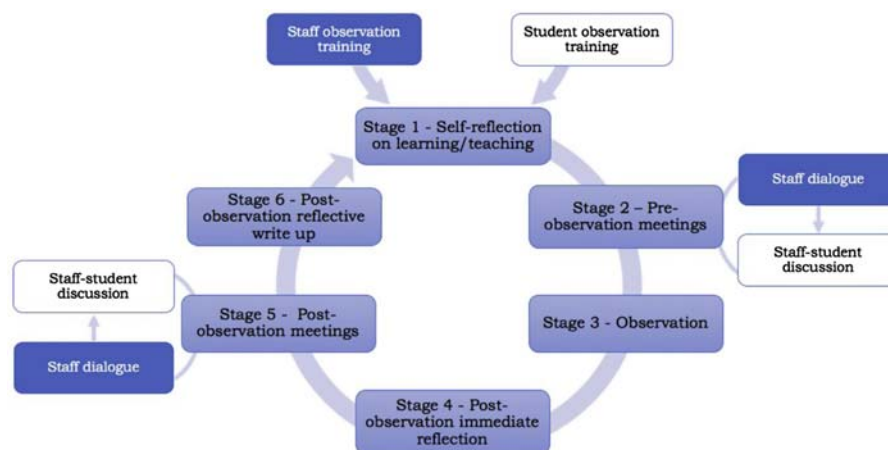


Fig. 2 Cycle of Collaborative Observation.

thought-provoking questions. The use of questions aimed to create a working ethos of enquiry and support participants in taking responsibility for their own learning.

After peer discussion, staff and students came together to share and discuss their individual reflections and to connect these reflections with teaching and learning on their course collectively. They identified the inquiries they wished to investigate together, agreed the focus and dates of the observed sessions, the roles of each member during the course of the cycle and the expectations of both parties. The focus of the discussion was on teaching and learning rather than the teachers/students. This was where, for the first time during CoCO, the two sets of lenses came together to co-interrogate teaching and learning.

The observation (Stage 3) involved students working in pairs to observe part of the session from their perspective as learners. A lecturer led the session while the other observed. Students and staff agreed their focus prior to the observations. To ensure that they were able to carry out high quality observations with limited distractions to their learning during the session, it was recommended that each individual student limit their observation time to 30 minutes. Observers took handwritten notes with no prescribed criteria on what to observe. Following the observation, *immediate reflections (Stage 4)* were carried out by all participants individually. Each observer then developed their notes, supplementing them with questions and comments in preparation for their post-observation meetings.

Stage 5 post-observation meetings took a similar format to stage 2. Students and staff met in their peer groups to discuss their observation experience and relate what they had observed to their own learning/teaching experience on the program. When meeting as a group, all participants brought their notes and reflections to the discussion and linked these back to the stage 2 discussion and the agreed learning objectives and focus. Each individual's perspective came together to form a collective understanding of the shared classroom teaching and learning they had observed and experienced.

Stage 6 of the cycle culminated in a *reflective write up* by each participant independently. The aim was to capture the holistic experience of both parties of the cycle. What's more, each participant was prompted to make connections between this observation experience with their past teaching and learning experiences and their current teaching and learning experiences across the course.

Over the course of the two iterations of CoCO, each case study produced a large volume of rich data, which included initial reflections from all student and staff participants, recordings of the pre- and post-observation meetings, observation field notes, individual reflective write-ups and recordings of evaluation meetings with the project researchers at the end of each cycle of observation. The thematic discussion that follows is drawn from an analysis of these data.

Developing classroom consciousness between students and staff

In Cycle 1, each case study focused on understanding the impact of teaching on student learning, exploring key aspects of teaching such as pedagogical approaches, the classroom environment, resources etc. The lecturers decided the focus of the observations and led the discussions during the pre- and post-observation meetings. One of the outcomes to emerge from Cycle 1 across each case study was the development of students' understanding of the relationship between learning and teaching. While Cycle 1 observations focused largely on teaching, the experience of collectively observing, discussing and reflecting on learning and teaching on their respective courses resulted in the unearthing of their classroom consciousness, as the following extract from Katie's (Radiotherapy student) end of cycle reflection illuminates:

The student observer process so far has been extremely insightful and made me think in more depth why lecturers do things in the techniques they do. It has made me understand that one technique is not suitable for all aspects of my course; for example being taught anatomy purely through PowerPoints and books would be extremely difficult to absorb and would not put the knowledge in to context however the human models and practical tasks gave a more natural way of learning.

One of the significant developments to occur across all case studies moving from Cycle 1 to Cycle 2 was the collective decision to shift the observation focus from teaching to learning. This shift in focus emerged as a suggestion during a project review day involving staff participants at the end of Cycle 1. After reading and discussing reflections by students, staff were keen to empower their students to take more responsibility for deciding the foci of Cycle 2 observations. This was then later agreed between staff and students in their respective case study meetings before starting Cycle 2.

Compared to Cycle 1, this shift in focus marked a key milestone in the project, leading to a greater depth of reciprocal awareness and understanding of classroom learning and teaching among students and staff. Across the case studies, it gave rise to a more even balance in the discursive interaction between participants. This was not only reflected in an increase in student talking time in Cycle 2 meetings compared to Cycle 1, but also in the depth of their critical thinking and reflections on learning and teaching.

In her work on the scholarship of teaching and teaching excellence, Kreber (2002) refers to a conception of teaching that is "learning oriented." We found this a helpful term and concept to explore the significance of this shift in focus from Cycle 1 to Cycle 2. Learning was placed at the forefront of all discussions in Cycle 2 with students encouraged to reflect on their experiences, views and feelings. As Jay (Adult Nursing student) commented, "this cycle is a lot better, because it's just us, isn't it? So, we can just focus on us, our learning and if everybody individually focuses on themselves, we've got something to put together and collaborate on." Both Jay and his peer Oliver agreed that focusing on their own learning in Cycle 2 "felt more natural" compared to having to discuss and evaluate their lecturers' teaching (as they did in Cycle 1 observation), which Jay maintained they were "not qualified to judge."

Jay and Oliver's perceptions were representative of their peers in the other case studies inasmuch as the shift to a learning-oriented focus also helped to remove some of the apprehension they felt in their roles as observers of their lecturers' teaching. For the lecturers, because the focus was driven by students in Cycle 2, this provided them with rich and timely insights into their students' learning. For example, in Radiotherapy where the Cycle 2 focus was on classroom dynamics, Nick (lecturer) highlighted how his participation in CoCO had made him reflect on the "social aspect of the classroom," triggering novel perspectives on his students' learning:

The most useful thing about this observation project to me is just working out the social aspect of the classroom. I've never had time to stop and think about it before... I don't remember having conversations about it when I was taught how to teach.

While the observation training for both students and staff emphasized how the observer's role was not to evaluate the teaching but to use observation as a tool of collaborative inquiry through which to explore teaching and learning, the focus on teaching in Cycle 1 meant that discussions often gravitated toward assessing the effectiveness of teaching. With the focus on learning, both parties were able to derive value that fed into their wider classroom consciousness of their learning/teaching and the relationship between the two. Stacey (Child Nursing student) remarked: "I think that was a good thing because it's made me reflect more on myself and where I'm learning and what my learning is and am I learning?" Nick (Radiotherapy lecturer) commented: "The feedback I want is 'does this make your practice more relevant?' It's not the process of being observed, it's the whole context of being observed."

Analysis of the transcripts from the meetings across both cycles revealed a development of confidence and fluency among the student participants in their discursive exchanges and the manner in which they were able to articulate their thoughts. However, it is important to acknowledge that this may also have been partly due to the development of the student-teacher relationship as each group got to know each other better over the duration of the project.

Building collaborative relationships between students and staff

The relationships between the participants in each case study were critical to their collaborative learning. What brought the participants together were their interests in learning and teaching and a desire to work together to develop individual and collective understandings of learning and teaching on their respective courses. The mutual interests in pedagogy and learning in each case study team were what formed this meaningful bond. Oliver (Adult Nursing Student) in his end of Cycle 2 reflection commented on what this pedagogical connection with his lecturers and his peers meant to him:

The major benefit I've taken away is the chance to have an informal meeting with the people providing my education and an entire hour set aside to talk about my experience. I was both contributing to and benefiting from the best part of a new model of student-teacher collaboration. It felt like a team-based approach, and I had Jay as my constant peer and colleague to give me a chance to listen to the way other students experience, perceive and evaluate the same things I have.

The social relationships across case studies included the formal roles and commitments participants had as students and lecturers in their respective courses, as well as their personal commitments to their collaborative partnership. Each participant took their formal role and commitment seriously in observing, questioning, critiquing and developing situated pedagogical practices in their classrooms. Besides, the nature of the collaborative partnerships seemed to yield more substantive personal connections between them. For the students, this was a new and different kind of relationship with their teachers, as Oliver explains during the pre-observation meeting with his lecturers and fellow student:

I know that you are obviously human, but when you're a student and you're sitting in front of a lecturer, sometimes it's that barrier of "I'm a student they're a lecturer" whereas all of a sudden when there's personal stories that come into it and their own little phrases and the way that they explain things, that adds to the idea that this is more of a peer relationship. We get more from that relationship than you would from just lecturer and students.

Students and staff clearly valued this personal connection with each other. Following on from the learning oriented lens (Kreber, 2002) discussed in the previous section, it is useful to draw on Rogers' work (2002) on interpersonal relationships in education to delve deeper into the nature of such personal connections. Rogers' conceptualization of learner-centered education provides a useful frame of reference for situating the discussion about the relationship and orientation between learning and teaching.

From the beginning of CoCO, a gradual process of familiarization unfolded between the participants. Cycle 1 interactions between staff and students started by focusing largely on procedural and technical aspects of observation. As the relationships developed, staff participants in each case study began more openly to share their insights on teaching (e.g. their pedagogical reasoning) and their assumptions of learning. This openness enabled the establishment of authenticity in the pedagogical relationships

between participants. As Rogers (2002) observed, this means that teacher and learner are able to enter interactions being themselves, with a heightened awareness of their feelings and views. This is clearly reflected in Oliver's comment above and similar remarks from other students. In particular, this meant students were able to see teaching not just as a job their lecturers do as a "curriculum instructor" (Rogers, 2002) but as decisions and actions carried out consciously and deliberately by a person who has shared educational interests with them. As the lecturers became more open about their practices and views, this was reciprocated by the students who were candid in sharing their opinions and feelings. This authenticity provided a platform for students and staff to engage in an open inquiry of learning and teaching. As discussed above, one of the principles of CoCO is that participants exchange their views and listen to others without judgment. Building and maintaining the realness of the pedagogical relationship between students and lecturers meant both parties were disposed to exchanging personal feelings and views, acknowledging the different perspectives of others and learning from each other, leading to a collective enhancement of classroom consciousness.

Rogers (2002) highlighted how this kind of understanding and acceptance differs markedly from evaluative approaches that focus on a "what works" methodology (Biesta, 2017) in teaching and learning. For Rogers (2002, 30), a sensitive empathy develops in pedagogical relationships that are built on realness and acceptance, which then enable students to feel "at last someone understands how it feels and seems to be me without wanting to analyze me or judge me. Now I can blossom and grow and learn." Rogers' work focused on the impact teaching had on learning and he saw the nurturing of such relationships as the responsibility of the teacher. In our project, both students and lecturers embraced this responsibility. As a result, a shared empathy emerged where students and lecturers developed a reciprocal awareness and understanding of the processes of learning and teaching in their classroom and how these are experienced from each other's perspectives. Reflections by Aneesa (student) and Ilana (lecturer) from Child Nursing illustrate the power of such shared understanding in developing their classroom consciousness and pedagogical practices:

This reflective cycle has enabled me as a student to work in collaboration with Ilana and Nathalie whereby we were all equal participants in research. This relationship has given me an incomparable insight into how much preparation and time goes into teaching and has encouraged me to make the most out of the lectures. I now have a better understanding of how I can benefit from different methods of teaching (Aneesa).

Participating in the process has impacted my teaching on a practical level. More importantly I've gained a deeper, shared understanding about learning and teaching as a result of the collaborative meetings with the students. This observation unlocked the potential for us to understand our students and thereby allowing us to meet their learning needs on a deeper level. However, learning still does depend on the students' willingness to participate in the process. By allowing students to become more aware of their own learning as a result of their participation in observation, consequently led to them being empowered to overcome self-identified barriers and leading to a better and more satisfactory teaching and learning experience for all. By enabling students to participate in the collaborative observation cycle, change does not only happen on one-dimensional level "teacher growth," but becomes a multi-dimensional change by enabling "student growth" and "teacher-student growth"

(Ilana).

CoCO played a critical role in facilitating this dialectic relationship development. For each participant, the individual reflections at the beginning and end of the cycle gave them the space to develop their autonomous voices. In their reflections, each participant talked openly about their experiences and reflected on them which triggered their questions on teaching and learning. When the participants came together at the pre- and post-observation meetings, they used their voices to negotiate and build the interconnecting part of their collaborative inquiry. In our project, as a shared point of reference, CoCO created the conditions for this dialectic relationship to flourish.

Concluding thoughts and future directions

It is highly likely that by the time this article is published, the world will have succeeded in controlling the Covid-19 pandemic and a sense of normality will have been restored to all walks of public life. In the case of education, however, it would be an opportunity missed if we were simply to return to the linear systems of quantitative performance data and reductionist approaches to learning about and improving teaching and learning that have dominated education internationally in recent years. Such systems and approaches serve to perpetuate the status quo of the "what works" movement, which in turn closes down alternative ways of knowing and doing. Context clearly matters and it is important to acknowledge the implications of this from a methodological perspective and the subsequent application of any research findings.

What the two projects discussed in this article reveal is that effective teaching and learning occurs as a result of sustained, personal and professional learning that involves critical reflection and ongoing negotiation of meaning and understanding between those involved in this interrelationship. This process is often unique to individuals and as such requires a tailored approach. Besides, teaching and learning and their interrelationship are inherently qualitative in nature and as such require appropriate methodologies and methods to explore, interrogate and make sense of these complex phenomena. As Strom and Viesca (2020, 13) argue, "research activity needs to be situated and local, with a focus on the particular."

An important finding to emerge from this body of research is that teachers and students learn about teaching and learning by interacting with their peers and with each other, by sharing their insights and experiences in collaborative, cooperative forums. In other words, teaching and learning are inherently socially situated practices. It therefore makes sense that any attempt to enhance understanding of and improve these practices is best served by allowing its key participants to be part of a collective community in which they are encouraged to engage in a process of dialectic pedagogical knowledge and relationship development. Creating the conditions in which teachers and students have the opportunity to examine their understanding and experiences of teaching and learning and open them up to dialogic exchange is fundamental to developing greater awareness of the strengths and areas for development in their practices. In the context of teaching and learning, approaches like CoPO and CoCO can help to create the shared spaces in which teachers can come together with their peers and equally with their students to engage in reflexive pedagogical dialog on their classroom teaching and learning. In turn, this can lead to collective sense making, which can make a valuable contribution to both students and their teachers in understanding and developing their practices.

Acknowledgment

The project *Improving teaching and learning through collaborative observation* was supported by the Higher Education Funding Council for England: [grant number CL2016/20a].

References

- Ball, S.J., 2015. Education, governance and the tyranny of numbers. *J. Educ. Pol.* 30 (3), 299–301.
- Bamber, R., 2020. Our Days Are Numbered – How Metrics Are Changing Academic Development. WONKHE, 18/12/2020. Online blog published 18th December 2020. Available at: <https://wonkhe.com/blogs/our-days-are-numbered-how-metrics-are-changing-academic-development/>. (Accessed 20 December 2020).
- Biesta, G., 2017. Education, Measurement and the Professions: reclaiming a space for democratic professionalism in education. *Educ. Philos. Theor.* 49 (4), 315–330.
- Blackmore, P., Blackwell, R., Edmondson, M., 2016. Tackling Wicked Issues: Prestige and Employment Outcomes in the Teaching Excellence Framework. HEPI Occasional Paper 14. Available at: <http://www.hepi.ac.uk/2016/09/08/tackling-wicked-issues-prestige-employment-outcomes-teaching-excellence-framework/#comments>. (Accessed 12 January 2020). Accessed.
- Bokeno, R., 2009. Genus of learning relationships: mentoring and coaching as communicative interaction. *Dev. Learn. Organ.* 23 (1), 5–8.
- Bowden, J., Marton, F., 2004. *The University of Learning: Beyond Quality and Competence*. Routledge, London.
- Buskist, W., Ismail, E.A., Groccia, J.E., 2014. A practical model for conducting helpful peer review of teaching. In: Sachs, J., Parsell, M. (Eds.), *Peer Review of Learning and Teaching in Higher Education: International Perspectives*. Springer, Dordrecht, pp. 33–52.
- Cilliers, P., 1998. Complexity and Postmodernism: Understanding Complex Systems. Routledge, London.
- Fielding, M., 2004. 'New Wave' student voice and the renewal of civic society. *Lond. Rev. Educ.* 2 (3), 197–217.
- Fricker, M., 2007. *Epistemic Injustice: Power & the Ethics of Knowing*. Oxford University Press, Oxford.
- Gill, J., 2015. David Willetts Interview: 'What I Did Was in the Interests of Young People'. *Times Higher Education*. Article published online 18th June 2015. Available at: <https://www.timeshighereducation.com/david-willetts-what-i-did-was-in-the-interests-of-young-people>. (Accessed 12 April 2020). Accessed.
- Kreber, C., 2002. Teaching excellence, teaching expertise, and the scholarship of teaching. *Innovat. High. Educ.* 27, 5–23.
- Lave, J., 1988. *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge University Press, Cambridge.
- Lave, J., 1993. The practice of learning. In: Chaiklin, S., Lave, J. (Eds.), *Understanding Practice*. Cambridge University Press, Cambridge, pp. 3–32.
- McMahon, T., Barrett, T., O'Neill, G., 2007. Using observation of teaching to improve quality: finding your way through the muddle of competing conceptions, confusion of practice and mutually exclusive intentions. *Teach. High. Educ.* 12 (4), 499–511.
- Meyer, H.-D., Benavot, A. (Eds.), 2013. *PISA, Power and Policy: The Emergence of Global Educational Governance*. Symposium Books, Oxford.
- Neumann, R., 2001. Disciplinary differences and university teaching. *Stud. High Educ.* 26, 135–146.
- O'Leary, M., 2020. *Classroom Observation: A Guide to the Effective Observation of Teaching and Learning*, second ed. Routledge, London.
- O'Leary, M., Wood, P., 2017. Performance over professional learning and the complexity puzzle: lesson observation in England's further education sector. *Prof. Dev. Educ.* 43 (4), 573–591.
- Rogers, C., 2002. The interpersonal relationship in the facilitation of learning. In: Harrison, R., Reeve, F., Hanson, A., Julia, C. (Eds.), *Perspectives on Learning*. RoutledgeFalmer, London.
- Sahlberg, P., 2011. *Finnish Lessons: What Can the World Learn from Educational Change in Finland?* Teachers College Press, New York.
- Shephard, K., Harland, T., Stein, S., Tidswell, T., 2010. Preparing an application for a higher-education teaching-excellence award: whose foot fits Cinderella's shoe? *J. High Educ. Pol. Manag.* 33 (1), 47–56.
- Skelton, A., 2004. Understanding teaching excellence in higher education: a critical evaluation of the national teaching fellowships scheme. *Stud. High Educ.* 29 (4), 451–468.
- Su, F., Wood, M., 2012. What makes a good university lecturer? Students' perceptions of teaching excellence. *J. Appl. Res. High Educ.* 4 (2), 142–155.
- Stevenson, H., 2017. The "datafication" of teaching: can teachers speak back to the numbers? *Peabody J. Educ.* 92 (4), 537–557.
- Strom, K.J., Mitchell Viesca, K., 2020. Towards a complex framework of teacher learning-practice. *Prof. Dev. Educ.* <https://doi.org/10.1080/19415257.2020.1827449>.
- Tenenberg, J., 2016. Learning through observing peers in practice. *Stud. High Educ.* 41 (4), 756–773.
- Whitmore, J., 2002. *Coaching for Performance: Growing People, Performance and Purpose*, third ed. Nicholas Brealey, London.

Reciprocal teaching and learning through student-faculty pedagogical partnership

Alison Cook-Sather, Education Program, Bryn Mawr College, Bryn Mawr, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: student-faculty pedagogical partnership	268
Consistent outcomes of student-faculty pedagogical partnership	269
Partnership values	270
National, institutional, disciplinary, and individual influences on approaches to partnership	271
Models of partnership facilitation	272
Center-based forms of facilitation	272
The creation of teams of students, faculty, and teaching and learning center staff	272
Bringing together practitioners of student-faculty partnership	272
Whole-of-class facilitation	272
Approaches to facilitation of partnership approaches supported by liminally positioned student partners	272
Facilitation by a faculty or staff member	272
Co-facilitation by a pair of faculty/staff members	273
Co-facilitation by pairs or teams of staff members and students	273
Co-facilitation by a faculty/staff member and post-baccalaureate fellow	273
Facilitation by a single student	273
Challenges and ways of addressing them	273
Future directions	275
Approaches to and analyses of educational practice	275
Conceptions of educational development	275
Scholarship of teaching and learning	275
Acknowledgments	275
References	276

Introduction: student-faculty pedagogical partnership

While teaching and learning are often conceptualized as distinct processes undertaken by differently positioned participants in higher education, student-faculty pedagogical partnership enacts and supports reciprocity in teaching and learning—positioning both students and faculty as both teachers and learners. A widely cited definition of pedagogical partnership is “a collaborative, reciprocal process” through which “all participants have the opportunity to contribute equally, although not necessarily in the same ways, to curricular or pedagogical conceptualization, decision making, implementation, investigation, or analysis” (Cook-Sather et al., 2014, pp. 6–7). Key to this definition is the phrase “not necessarily in the same ways”; indeed, it is the different experiences and perspectives each participant brings and the different ways in which participants contribute, all equally valued, that position faculty and students in student-faculty partnership as both teachers and learners.

Scholars have developed multiple typologies to classify forms of partnership work, all of which focus directly or indirectly on teaching and learning. A frequently referenced typology focuses on arenas of engagement: (1) learning, teaching, and assessment; (2) subject-based research and inquiry; (3) scholarship of teaching and learning; and (4) curriculum design and pedagogic consultancy (Healey et al., 2016). Another typology classifies partnership work in terms of student partner role: (1) consultant: sharing and discussing valuable perspectives on learning and teaching; (2) co-researcher: collaborating meaningfully on teaching and learning research or subject-based research projects with faculty and staff; (3) pedagogical co-designer: sharing responsibility for designing learning, teaching and assessment; and (4) representative: ensuring student voices contribute to decisions in a range of university settings (Bovill et al., 2016). Yet another typology offers a participation matrix keyed to level of student involvement across different activities traditionally considered the sole purview of faculty (such as course design or project evaluation): inform, consult, participate, partner, and control (Bovill, 2019; see also Bovill, 2017).

Reciprocal teaching and learning through student-faculty partnerships directly focused on classroom practice can be distinguished by location of the student partner in relation to the teaching and learning practice: as enrolled or involved in it, or as positioned in a liminal role outside it. The former, which Bovill (2020) calls whole-of-class co-creation, is often undertaken with the goal of democratizing classroom practices (Bergmark and Westman, 2016; Cook-Sather et al., 2014) and “may be inherently more inclusive of students” (Bovill, 2020, p. 1023) since all students are involved. The latter kind of partnership work positions students outside the standard teacher-student power dynamic (Cook-Sather and Alter, 2011), fosters engagement in respectful dialog (Cook-Sather, 2020) and generation of more candid feedback (Cook-Sather, 2009), and affords students a rare opportunity to mobilize their

identities to foster change for equity, inclusion, and culturally sustaining practices (Bell et al., 2020; Cook-Sather and Agu, 2013; Cook-Sather et al., 2019a,b,c) from less central, but therefore sometimes more powerful, positions (Cook-Sather, in press).

Student-faculty pedagogical partnership work that both enacts and fosters reciprocal teaching and learning is rapidly expanding around the world. There are established programs in Aotearoa/New Zealand, Australia, Canada, England, Israel, Pakistan, Scotland, Sweden, and the United States through which teaching and learning centers or institutes support student-faculty pedagogical partnership across the institution. There have been pilot efforts in Grenada, Hong Kong, Italy, Malaysia, Norway, and Singapore through which individual faculty members have invited students to contribute to curricular or pedagogical conceptualization, decision making, implementation, investigation, or analysis. And there is growing interest in China, Germany, and The Netherlands in such work. For instance, a scoping review of partnership work in Asia reports on the literature from China, Hong Kong, Singapore, and Malaysia through the lens of the power dynamics that inform student-teacher relationships, particularly in Asian contexts (Liang and Matthews, 2021).

This entry provides an overview of various aspects of reciprocal teaching and learning through student-faculty pedagogical partnership work in higher education. It first reviews the reported outcomes of this work. Next, it reviews the core values that guide pedagogical partnership work, emphasizing that how those values are enacted depends on context and evolves as partnership work spreads around the globe. Turning to a practical aspect of supporting student-faculty partnership, the entry summarizes multiple approaches to facilitating partnership work. Then, stepping back from outcomes, values, and approaches to facilitation, the entry identifies the challenges of partnership work. It concludes by suggesting future directions for reciprocal teaching and learning not only in classroom-focused student-faculty partnership but also in conceptions of educational development and scholarship of teaching and learning.

Consistent outcomes of student-faculty pedagogical partnership

There have been numerous, cross-context analyses of the outcomes of student-faculty pedagogical partnership that look back over a decade of such partnership (e.g., Cook-Sather et al., 2019a,b,c, 2014; Mercer-Mapstone et al., 2017). The positive outcomes of this work are remarkably similar across context. Students consistently experience:

- Heightened engagement in, motivation for, and ownership of their learning
- Increased confidence, self-efficacy, and sense of agency
- Deepened capacity to articulate their own perspectives
- Deeper understanding of others' experiences (both teachers' and students')
- Enhanced relationships with or trust of faculty and other students
- Deeper understanding of themselves as learners and their own learning processes (metacognitive learning, self-evaluation, self-awareness)
- Raised awareness and enactment of graduate attributes/employability skills/career development
- Increased sense of belonging to university or discipline or community
- Improved student content/discipline learning (actual or perceived)
- Positively shifted identity as student/learner/person/professional
- Deeper understanding of and appreciation for the difficult work of teaching and greater empathy for teachers

Revealing some of the same experiences as those reported by students and some different experiences, faculty consistently report:

- Accelerated processes of acclimating to campus culture and unfamiliar students
- Deepened confidence and clarity about their pedagogical commitments
- Courage to follow through on their pedagogical convictions and responsibilities
- Gaining a perspective that they cannot achieve on their own
- Recognizing and making intentional good pedagogical practices
- Discovering the benefits of sharing power—and responsibility—with students
- Enhanced relationships with or trust of students and other faculty
- Development of new or better teaching or curriculum materials
- Deeper understanding of others' experiences
- New beliefs about teaching and learning that change practices for the better
- Re-conceptualization of teaching as a collaborative process to foster learning

There are also negative outcomes of student-faculty pedagogical partnership, though these are less common and also less frequently reported. Among the negative outcomes for students are feeling:

- Unproductively vulnerable to faculty partners or traumatized by a partnership
- Frustrated with non-partnership frames and practices in other contexts (after having experienced partnership) and in partnerships that do not realize their potential
- Hyper-responsible as a result of increased awareness and capacity
- Disenchanted with non-engaged work

Among the negative outcomes for faculty are:

- Feeling unproductively vulnerable to students and others—at risk or exposed
- Experiencing initial resistance and dismissal
- Rejecting the partnership approach

A growing body of scholarship that brings an equity frame to the processes and outcomes of partnership work emphasizes that if partnership is not carefully conceptualized, structured, and revised as needed, it can exacerbate the epistemic, affective, and ontological harms many students experience in higher education (de Bie et al., 2019, 2021). Such work documents the harm that indigenous students (Yahlnaaw, 2019), mature-aged students (Vayada, 2021), international students (Colón García, 2017), students of color (Ntem and Cook-Sather, 2018), and students from other equity-denied groups can experience in partnership when certain, particularly Western, notions of who has and can create knowledge about teaching and learning (Bindra et al., 2018) and overly goal-driven (rather than emergent) approaches are privileged.

The outcomes for students and faculty of student-faculty pedagogical partnership affirm the reciprocal nature of the teaching and learning fostered by this work. The clear pattern of each participant gaining deeper insight into the other's experience and contributing to the learning of the other, as well as the potential for unproductive vulnerability and harm, highlight two sides of such relational work. Through enacting and supporting reciprocal teaching and learning, partnerships can lead to deeply meaningful experiences from which to learn more about one's self, others, and teaching and learning.

Partnership values

As partnership work was initially emerging, primarily in the United States and the United Kingdom, scholars identified three core underlying values of partnership practices: respect, reciprocity, and shared responsibility (Cook-Sather et al., 2014). Respect is an attitude: a valuing of others' experiences and perspectives. Reciprocity is both an ethic and a mode of engaging—a “process of balanced give-and-take not of commodities but rather of contributions: perspectives, insights, forms of participation” (Cook-Sather and Felten, 2017, p. 181). And shared responsibility is both a process (during partnership) and an outcome of partnership. To these values, Healey et al. (2014) added authenticity, honesty, inclusivity, empowerment, trust, courage, and plurality.

As partnership work has expanded to other regions of the world, scholars and practitioners have recast and revised these core values. For instance, feminist scholars in the United States have recast the values of respect, reciprocity, and shared responsibility as agency, accountability, and affinity (Cates et al., 2018), and drawing on partnership work in England, Aquarone et al. (2020) emphasize the value of recognition. Student partners in Australia have added vulnerability to the list (Matthews, 2021; Vayada et al., 2020), and student partners in the United States highlight the value of affirmation (Ntem, 2016; Jonsson, 2020). A cross-context set of perspectives of 20 students, faculty, professional staff and administrators in Aotearoa/New Zealand, Australia, Canada, England, Pakistan, Singapore, South Africa, the United States, and Wales offered in response to the question “What does successful partnership mean to you?” surfaced the following values that are consistent with arguments offered in partnership literature: promotes shared ownership, decision-making, and mutual benefits; joins distinct but complementary perspectives and attends to unique needs of partners; pushes comfort zone, challenges norms, and asks new questions; differentiates between process and outcome, which requires ongoing sustained work; and promotes adaptability and responsiveness (Cook-Sather et al., 2021b).

As partnership work expands beyond the largely Western contexts in which it originated, scholars and practitioners are wrestling with how the originally identified core values, often evoked to launch partnership initiatives, intersect with core values in a range of contexts. For instance, in relation to Asian contexts, Liang and Matthews (forthcoming) argue that “Confucian educational values,” which have sometimes been evoked as a barrier to partnership in Asian contexts, are “compatible with the contemporary values evoked as the guiding ethos of student partnership” and “can inform the implementation of [partnership work] in China and Asian universities.” In striving to reconcile and revise Western partnership values with non-Western values, scholars have explored cultural phenomena such as: negotiating student deference to authority and a focus on grades rather than growth and learning (Seow, 2019), issues of face (Sim, 2019), and conceptions of students as clients, customers, objects, products but not (yet) partners (Chng, 2019) in Singapore; dichotomies between teachers and students that value student deference to faculty authority in Malaysia (Kaur, 2020); valuing confidence and active participation in a Pakistani context in which students do not typically have a voice (Ansari, 2021; Iftikhar, 2021; Waqar and Asad, 2020); and cultural norms across contexts according to which students engaging in dialog about teaching and learning with faculty is simply neither an established value nor a practice (Liang and Matthews, 2020; Felten et al., 2019; Kaur and Yong Bing, 2020; Moorhouse and Oh, 2019).

The definition of partnership as supporting participants in contributing “equally, although not necessarily in the same ways, to curricular or pedagogical conceptualization, decision making, implementation, investigation, or analysis” (Cook-Sather et al., 2014, p. 7) highlights that partnership practices are defined by reciprocity. Such practices require and inspire an ongoing commitment to identifying and re-examining values that underpin the partnership work guided by an awareness of the potential for all the desired benefits and equally of the risk of conflict and harm.

National, institutional, disciplinary, and individual influences on approaches to partnership

The nested contexts of and participants in student-faculty pedagogical partnership work influence how it is conceptualized and how it unfolds (Cook-Sather and Wilson, 2020; Healey and Healey, 2018). The “Ako in Action” program at Victoria University of Wellington, for instance, offers an example at the intersection of national and institutional considerations and participant identities. Aotearoa/New Zealand is a bicultural country, and it is home to Māori, settlers of European descent, and peoples of Asian descent. Victoria University of Wellington, a research-intensive university with approximately 22,000 students in the country’s capital city, houses a student-faculty pedagogical partnership program that was developed in the context of “a new Learning and Teaching Strategy that embedded a bicultural approach to learning and teaching” adopted by the university in 2017. This approach was built on the principle of “*‘ākoranga*,” a Māori term that means both to teach and to learn.” This principle is enacted at the university as “a ‘collective responsibility for learning,’” and the partnership program at the university was “co-conceived, designed, and coordinated by students and academic (faculty) developers” (Leota and Sutherland, 2020, p. 93.)

The student-faculty pedagogical partnership program at Kaye Academic College in Be’er Sheva, Israel, provides another example of the intersection of national and institutional considerations and participant identities. Kaye Academic College focuses on the preparation of primary and secondary teachers and serves approximately 2000 students from both the Jewish and Bedouin sectors of the Negev region of the country. Many of the students “are the first generation of their families to attend higher education ...; many come from marginalized populations ...; very few have English as their mother tongue, indeed for many students in the English teacher-training department, where English is the medium of instruction, English is often their third or fourth language” (Narkiss and Naaman, 2020, pp. 42–43). Considering how to name one of the key commitments of partnership, respect, is complicated in this program because the word “respect” can evoke hierarchy rather than collaboration. In a context like Kaye as an institution and Be’er-Sheva more broadly, with many different ethnic groups who have differing words for and understanding of words like respect, they use the term “*kavod*” to express several English concepts including honor, respect, dignity and esteem (Judith Zamin, Personal Communications, October 24 and October 25, 2019). More challenging still is structuring partnership such that it feels welcoming for Bedouin students and is responsive to their ways of being. Many of these students live according to clan values and structures based on blood and extended family connections and strict loyalties (informal conversation with colleagues at Kaye). These do not easily fit with the cross-role, reciprocal, and critical dimensions of pedagogical partnership as it is typically conceptualized in Western contexts.

Stories of pedagogical partnerships from around the world reveal how partnership develops not only in national and institutional contexts but also at the intersection of disciplinary norms and participant identities, drawing in reciprocal ways on each partner’s identities and commitments. For instance, a female, Nigerian, graduate student in gender studies and feminist research working in partnership with a female, white, faculty member in engineering in a Canadian university drew on her own family’s background in engineering as well as her belief in the potential of “radical pedagogy that centers students from minoritized groups” to design a new engineering course on Inclusion in the Engineering Workplace (Chukwu and Jones, 2020, p. 14). These partners reflected on how, through the partnership, they both “learned and grew, and benefited from one another’s disciplinary expertise” (Chukwu and Jones, 2020, p. 20) and how, in particular, working across disciplines “made it possible to circumvent the faculty-student power dynamic to some extent” (p. 16) and create “a productive and truly equitable partnership” (p. 20).

As another example, the Medical Humanities curriculum, a compulsory part of the Bachelor of Medicine, Bachelor of Surgery program at the University of Hong Kong, was developed to support doctors who are working in “technologically advanced, difficult and emotionally charged conditions” learn to be “more humanistic towards both their patients and themselves” (Chen and Ho, 2020, p. 21). The student and faculty pair who worked on developing this curriculum brought both personal and professional perspectives to bear—the loss of a parent, the established emphasis on interpersonal and communication skills in the curriculum. Through this partnership, the faculty member became much more aware “of the interrelationship between formal curricular initiatives and the learning environment, culture, or social setting in which students are immersed,” and the student partner found that “pedagogical partnership encourages mutual ownership and encourages active contribution” that can contribute to “satisfactory and sustainable changes and learning outcomes” (Chen and Ho, 2020, p. 27).

As yet another example, a student and faculty partnership in Malaysia developed through the collaboration of a female, Indian-born, visiting faculty member and a male student from a minority community in Malaysia. These two worked together on a psychology course to “transform traditional one-way and hierarchical pedagogical methods” typically embraced in the Malaysian context and “pave ways for pedagogical methods that are co-constructed with students” (Kaur and Yong Bing, 2020, p. 70). The student partner’s contributions provided his faculty partner with “a window into students’ hearts and minds” and brought to her attention things that she “might have ignored otherwise” (Kaur and Yong Bing, 2020, p. 67). At the same time, the partnership work built the student partner’s confidence and helped him “come to believe that a teaching and learning paradigm shift could be achieved ... through teacher-student partnership” in his cultural context (Kaur and Yong Bing, 2020, p. 68).

This is a small sample of the growing number of partnership efforts around the world. All of these efforts unfold within and beyond power dynamics and prescribed roles that are structured into institutions, even as they strive to be attentive to context and responsive to participants. All are doing so through the “as-if” spaces that partnerships create (Cook-Sather and Felten, 2017), providing participants liminal spaces within which to explore, try out, and carry forward reciprocal teaching and learning.

Models of partnership facilitation

There are many approaches to supporting reciprocal teaching and learning through student-faculty pedagogical partnership. The best documented of these include teaching and learning center-based forms of facilitation; the integration into the pedagogy of whole-of-class student-faculty partnership practices; and various approaches to facilitation of partnership approaches supported by liminally positioned student partners.

Center-based forms of facilitation

Academic developers can play a key role in supporting the complex, relational processes of partnership; fostering and advocating for institutional or structural change to address resistance; and focusing on the potential of partnership to make institutions more equitable and empowering spaces (Matthews et al., 2019). Existing forms of center-based approaches include the creation of teams of students, faculty, and teaching and learning center staff who focus on revision of a course and related pedagogical practices, and the bringing together of larger groups of practitioners of student-faculty partnership work to share practices.

The creation of teams of students, faculty, and teaching and learning center staff

This is the approach taken through Elon University's Center for the Advancement of Teaching and Learning. The first iteration of this approach provided the model: two faculty members, the center director, and seven students participated in a Course Redesign Design Team to redesign an education course (Delpish et al., 2010). This approach "not only produced an effective new course (as seen in the higher evaluations than in previous terms), but also produced significant learning for everyone on the team" (Delpish et al., 2010, p. 126)—reciprocal teaching and learning in team-based curriculum redesign. Other institutions, such as University College London in England (Quy et al., 2020), have developed variations on this approach not only for courses but also for entire departments and programs. Such an approach helps "access multiple perspectives" (Quy et al., 2020, p. 82) in developing and revising educational structures and practices.

Bringing together practitioners of student-faculty partnership

A kind of "middle-out" model brings practitioners together to share the diverse range of partnership practices in which those practitioners engage. Such an approach builds community and affirms and further develops practices. For instance, at the University of Queensland (UQ) in Australia, the UQ Students as Partners Community of Practice "creates space for staff [faculty] and students to explore, share, and improve partnership practices together to positively transform teaching, learning, and the student/staff experience" (Khoury et al., 2017, p. 3). This program is open to all UQ students and faculty, and participants meet every six weeks with the goals of sharing experiences, advice, and resources about partnership work; supporting each other in undertaking and enhancing partnership work; networking and fostering collaboration between community members; and influencing policy and practice at UQ. This model might be useful in contexts in which participants do not have structured programs or funds to support those who are keen to practice partnership both formally and informally.

Whole-of-class facilitation

Facilitation of whole-of-class student-faculty pedagogical partnership typically involves the faculty member communicating with students prior to or early in the semester about the reciprocal teaching and learning approach they plan to take in the course, inviting student input on that approach, and structuring student engagement in that approach (Bovill, 2020; Bryson and Callaghan, 2021). In this model, facilitation is part of the pedagogy of the course. Faculty have invited students to: co-create curriculum with the goals of democratizing class practice (Bergmark and Westman, 2016) and catalyzing new professional partnerships that in turn contribute to refreshed learning experiences and communities (Cates et al., 2018); co-create assessment with the goals of challenging power dynamics in assessment and developing assessment literacy (Deeley and Bovill, 2017; Deeley and Brown, 2014; Goldbold et al., 2021; Monsen et al., 2017); and embracing and enacting anti-racist pedagogical approaches (Wilson and Cook-Sather, in press).

Approaches to facilitation of partnership approaches supported by liminally positioned student partners

Across different contexts and institutions, varying models of facilitation of partnership approaches supported by liminally positioned student partners—students positioned in partnership roles outside the standard teacher-student roles, power dynamics, and institutional relationship—have been employed. These models include: facilitation by a single faculty/staff member; co-facilitation by a pair of faculty/staff members; co-facilitation by a faculty/staff member and a post-baccalaureate fellow; co-facilitation by staff members and student partners; and facilitation by a single student.

Facilitation by a faculty or staff member

One of the most common approaches to facilitation has a faculty, staff member, or center director serve as facilitator. In this case the role may be compensated by pay or release time from teaching for a faculty, pay or reassignment of responsibilities for a staff member, and assignment as part of a center director's responsibilities. This is the approach taken at a wide range of institutions, including Lahore University of Management Sciences in Pakistan, Kaye Academic College in Israel, and numerous small private colleges in the United States, including Bryn Mawr and Haverford, Reed, and Smith Colleges. Those choosing this model of

facilitation are advised to consider how to implement it with sensitivity to context since this approach can either fit within or build beyond existing institutional structures (see [Cook-Sather et al., 2019a,b,c](#) for detailed guidelines for this approach.)

Co-facilitation by a pair of faculty/staff members

Another approach involves two faculty or staff members co-facilitating. This is the approach taken at Florida Gulf Coast University in the southern region of the United States, a public university and Hispanic-Serving Institution, at Yale University, a private research institution in the northeastern region of the United States, and at Amherst and Ursinus Colleges, two private liberal arts college in the northeastern and Mid-Atlantic regions, respectively, of the United States. The different institutional positions and disciplinary backgrounds such co-facilitators bring to facilitation contribute different insights and lead to different emphases ([Reynolds et al., 2019](#)). This approach affords both facilitators and student partners the benefit of having two experienced educators who bring different perspectives to their facilitation and who model reciprocal teaching and learning as facilitators.

Co-facilitation by pairs or teams of staff members and students

At several institutions, support for pedagogical partnership work is provided by pairs or teams of staff members and students. This approach has been implemented with staff and undergraduate students as facilitators at Vassar, a small liberal arts college in New York ([Bala, 2021](#); [Bala and Kahn, 2021](#)), and Tufts, a private research university in Massachusetts ([Impastato, 2021](#); [Topper, 2021](#)), both in the United States; with a staff member and a graduate student in one stream of a larger partnership program at McMaster, a large, research institution in Canada ([Foran et al., 2020](#)); and with a staff member and an undergraduate student at Victoria University of Wellington, a research-intensive university in Aotearoa/New Zealand. About her experience of co-facilitating with a staff member, an undergraduate student wrote that one of the most important lessons she learned was “that pedagogical partnership work is not about knowing but *seeking*.” She continued: “I realized that genuinely engaging and exchanging points of view with others through open-ended questions encourages growth and learning and builds cherished relationships in academia” ([Bala, 2021](#) p. 5). [Leota and Sutherland \(2020\)](#) describe facilitation of their partnership program in Aotearoa/New Zealand in reciprocal terms as well: “we ask everyone participating to think of themselves in partnership; it is not just students as partners, but also academics as partners and academic developers as partners” (p. 95).

Co-facilitation by a faculty/staff member and post-baccalaureate fellow

Post-baccalaureate (post-bac) fellows who support pedagogical partnership work at several institutions are recent graduates of colleges at which they participated as undergraduate student partners. Such fellows bring lived experience and knowledge of pedagogical partnership work that is invaluable in supporting current student partners as well as that complements and informs faculty/staff directors’ knowledge and expertise. Trinity University in the southern region of the United States was the first to pilot this approach, and the post-bac fellow role has remained part of their permanent facilitation structure ([Cook-Sather et al., 2019a,b,c](#)). At Berea College, also in the southern United States, student partners enroll in a course for academic credit, and for the first 2 years of the program, as part of its launch, this course was co-facilitated by the director of the partnership program and a post-bac fellow ([Ortquist-Ahrens, 2021](#)). This approach provides person-power when directors might find it difficult to devote time and energy to developing partnership programs on their own, given their other responsibilities; offers experience with or knowledge of partnership work to those who lack them, as do most beginning directors; and lends credibility to those who have a commitment to modeling and enacting partnership in developing partnership ([Ortquist-Ahrens, 2021](#)). A post-bac position was also conceptualized as part of planning for a partnership program in Grenada, West Indies.

Facilitation by a single student

A newly emergent approach positions an undergraduate with experience in pedagogical partnership work as the facilitator of student-faculty partnership at other undergraduate institutions. A student who had been the co-creator of the pedagogical partnership program at Vassar College in New York took on a similar role for Emmanuel College in Massachusetts, both private colleges ([Bala, 2021](#)). The reciprocity in such a situation goes in multiple directions—between this student facilitator and student partners, between the student facilitator and faculty partners, and between the student facilitator and her collaborator at Bryn Mawr and Haverford Colleges, who also supported the launch of both Vassar’s and Emmanuel’s partnership program. This multi-layered, cross-role, inter-institutional partnership work to support reciprocal teaching and learning positioned all participants as teachers and learners.

Challenges and ways of addressing them

There are complex and overlapping challenges to engaging in and fostering reciprocal teaching and learning through pedagogical partnership. In keeping with the spirit of partnership, challenges can be conceptualized as opportunities for critical reflection, bringing to bear multiple perspectives on any given issue or problem, and embracing a diversity of possible approaches to addressing the challenges.

One key challenge to engaging in and fostering reciprocal teaching and learning through pedagogical partnership is resistance to co-creation ([Bovill et al., 2016](#); [Cook-Sather and Matthews, 2021](#); [Matthews, 2019](#); [Ntem and Cook-Sather, 2018](#)). Some of the sources of resistance have been conceptualized as threshold concepts in pedagogical partnership—such as that students have

valuable knowledge of and important perspectives on teaching and learning (see Cook-Sather et al., 2019a,b,c, for a full list of threshold concepts in pedagogical partnership). If these are not addressed, they can not only impede the realization of partnerships' potential but also cause harm. While resistance may be understandable, given the counter-cultural nature of partnership work, faculty and staff committed to reciprocal teaching and learning through pedagogical partnership have described striving to clearly communicate intentions to open up dialog and give people time to grapple with the paradigm shift partnership requires and embodies (Cook-Sather and Matthews, 2021). When student partners have encountered resistance, they have described "striving to build trust and relationships; taking further steps toward their faculty partners (or a kind of leaning in) in an effort to realize the transformative potential of partnership; and, conversely, withdrawing (or taking a step back) out of self-protection" (Ntem and Cook-Sather, 2018, p. 83).

A second challenge to engaging in and fostering reciprocal teaching and learning through pedagogical partnership is navigating institutional structures, practices, and norms. These might include curriculum constraints and accreditation or institutional reporting requirements as well as overall culture. While an institutional assessment framework or a set of metrics might not be changed, easily or at all, it is possible to focus on what can be co-created; faculty and students can work together to decide how to meet institutional requirements (Cook-Sather and Matthews, 2021). In terms of overall institutional culture, when faculty and students embrace empathy for and productive vulnerability to another, they take action to make change that can contribute to culture shifts (Cook-Sather and Mejia, 2018; Gibson and Cook-Sather, 2020; Hernandez, 2015). Positioning equity-denied students in partnership with faculty supports those students in mobilizing their identities to foster change for equity, inclusion, and culturally sustaining practices (Bell et al., 2020; Cook-Sather and Agu, 2013; Cook-Sather et al., 2019a,b,c). Finally, time is often identified as a challenge, although investing time up front in developing partnership practices can save time later, since learning experiences/insights/outcomes will likely be more meaningful and impactful (Cook-Sather et al., 2014; Cook-Sather and Matthews, 2021).

A third challenge is inclusivity. The whole-of-class approach to student-faculty pedagogical partnership includes all students (Bergmark and Westman, 2016; Bovill, 2020), although that may not mean that it is in fact realizing inclusivity. For student-faculty pedagogical partnership work with liminally positioned student partners, considerations such as which students to invite into partnership and why, who has access to partnership and who does not feel they do, and how partnership challenges or reinforces existing inequities are all questions of inclusivity (Ameyaa et al., 2021; Bindra et al., 2018; Bovill et al., 2016; Cook-Sather, 2020; de Bie et al., 2021; Cook-Sather and Matthews, 2021; Marquis et al., 2018; Mercer-Mapstone and Bovill, 2019; Mercer-Mapstone et al., 2019). Guidelines developed for launching and supporting partnership work with liminally positioned student partners address inclusivity—as well as equity and justice—in the following way (the text in regular font is from Cook-Sather et al., 2019a,b,c) reframed in equity terms (the text in italics is from de Bie et al., 2021, p. 106):

- Consider why you want to develop a pedagogical partnership program and what might get in the way, *paying particular attention to who has been well- and under-served by your institution, who stands to benefit from a partnership program or project (e.g., the institution, those least served), and what potential harm a partnership initiative might inadvertently cause.*
- Decide what kind of partnership program is right for your context and how, *in particular, such a program or project might redress harms and support the pursuit of greater equity.*
- Think through how to situate and structure the program, get started, and plan for sustainability *such that the needs and interests of those who have been least well served by the institution continue to be addressed in the program or project.*
- Decide on the shared responsibilities of facilitating pedagogical partnerships, *paying particular attention to the forms of and whose emotional as well as epistemic labor might be involved and how to resist emotional and epistemic exploitation.*
- Create an approach to inviting and supporting participants, *particularly attending to creating structures and approaches that welcome students from equity-seeking groups and recognize differences in experiences rather than assuming a homogeneous student perspective or experience.*
- Prepare for managing logistical and emotional challenges of partnership *and how differently positioned people will face different challenges, understanding that the challenges may not be fairly distributed and that participants will be differently impacted or affected by them.*
- Develop ways to assess pedagogical partnership work *that focus in particular on the potential for partnership to exacerbate or redress epistemic, affective, and ontological harms.*

A fourth challenge, which is implicit in the challenges noted above, is how to navigate power dynamics. Power dynamics are embedded in daily pedagogical partnership processes and relationships, even as those endeavor to "disrupt traditional student-faculty power relations" (Kehler et al., 2017, p. 4). In whole-of-class student-faculty partnerships, faculty members' "dismantling of the classroom power structure" (Cates et al., 2018, p. 41) can lead to students feeling empowered, but faculty sharing power can also create both internal conflict (within individual students) and interpersonal conflict (among students and between students and the teacher), which can be both challenging and generative (Godbold et al., 2021). In partnership work, many faculty can feel "wary of giving too much power to students" (Murphy et al., 2017, p. 7), while others can embrace an approach to "working together to take full advantage" of the expertise of students alongside that of faculty and staff, which can lead to understanding "the true meaning and importance of shared power through collaboration" (Mihans et al., 2008, p. 5). The latter reflects an understanding of "sharing power with [student partners], not exerting it over them" (Delpish et al., 2010, p. 98).

Future directions

Reciprocal teaching and learning through student-faculty partnership has informed not only approaches to and analyses of educational practice but also conceptions of educational development and scholarship of teaching and learning. Future directions in this work might explore how reciprocal teaching and learning can more fully permeate these and other areas, particularly as an integral component of expanding commitments to equity.

Approaches to and analyses of educational practice

Classrooms are “locations of possibility” (Cates et al., 2018; hooks, 1994), and both locations of and possibilities for partnership in teaching and learning have expanded as a result of the Covid-19 pandemic (Ntem et al., 2020). These spaces of learning and teaching might be further reconceptualized through sharing responsibility for the co-creation of learning environments, curriculum, assessment, and more. A focus on equity in classroom practice is increasingly recognized as essential, and student-faculty partnership can raise awareness of a range of equity issues and take those as the focus of partnership work (Cook-Sather, 2020; Marquis et al., 2021). Exploring the role of culture in partnership work by linking macro level to everyday/micro work of the partnership can make the role of culture more visible and present and challenge assumptions that Western culture is primary and all others are “other” (Zhang et al., forthcoming). This questioning of the centrality of “Western” culture in partnership work/scholarship is a vital issue for ongoing consideration.

Conceptions of educational development

Several scholars have argued for a reconceptualization of educational development that both enacts and supports reciprocal teaching and learning through student-faculty partnership (Cook-Sather et al., 2019a,b,c; Felten et al., 2019; Marquis et al., 2019). Positioning undergraduate students as educational developers challenges power dynamics within educational development and supports reciprocal learning about teaching through partnership (including co-created educational research), which in turn can support reciprocal teaching and learning in classrooms. Supporting the development of such work is challenging in the Global North (Cook-Sather et al., 2019a,b,c) and even more so in the Global South (Cook-Sather et al., forthcoming). Future research might further explore what enhances outcomes and what reduces inhibitors to the engagement of students and faculty in learning and teaching partnerships (Matthews et al., 2019) across countries, cultures, and contexts.

Scholarship of teaching and learning

Reciprocal teaching and learning through student-faculty partnership in scholarship of teaching and learning (SoTL) might shift toward valuing not only dominant, Western ways of knowing but also long-standing, indigenous ways (Yahlinaaw, 2019) and recognizing all students as authors, reviewers, advisors, and editors (Cook-Sather et al., 2021a). SoTL projects that embrace student-faculty partnership enact one of Felten’s (2013) five principles of good practice in SoTL—“that inquiry into learning be conducted in partnership with students” (p. 123)—and valuing a diversity of perspectives and forms of writing will be critical to the future of SoTL if it wants to strive for inclusion and equity. Partnership in SoTL can generate enhanced research that in turn informs and constitutes a form of educational development with the potential to enhance teaching and learning.

These forms of reciprocal teaching and learning through student-faculty partnership—approaches to and analyses of educational practice, conceptions of educational development, and scholarship of teaching and learning—constitute current, enacted, educational practices, support of and growth in conceptions of and approaches to educational practice, and theories and analyses of both. Each informs the other, and, through this dynamic, institutions of higher education might move to create cultures of partnership through which reciprocal teaching and learning become expected and supported rather than the exception perpetually in search of greater support. Such a culture would support the realization of higher education’s espoused aspirations toward equity and inclusion. Essential to creating such a culture is the embrace of a “partnership mindset” (Paseta et al., 2020). In the same way that partnerships create “as-if” spaces (Cook-Sather and Felten, 2017) in which to develop practices that can then be carried beyond the partnerships, partnership mindsets can shape how partners work, act, and live outside of higher education—precisely the orientation that individuals and communities need to work on the complex issues facing our neighborhoods, cities, countries, and world.

Acknowledgments

Many thanks to Peter Felten, Beth Marquis, Kelly E. Matthews, and Elliott Shore for their feedback and suggestions on drafts of this entry.

References

- Ameyaa, R.A., Cook-Sather, A., Ramo, K., Tohfa, H., 2021. Undergraduate students partnering with faculty to develop trauma-informed, anti-racist pedagogical approaches: intersecting experiences of three student partners. *J. Innov. Partnersh. Change* 7 (1). <https://journals.studentengagement.org.uk/index.php/studentchangeagents/article/view/1020>.
- Ansari, H., 2021. Building bridges: conceptualizing a student-teacher partnership program to improve pedagogical practices. *Teach. Learn. Together High. Educ.* 33. <https://repository.brynmawr.edu/tithe/vol1/iss33/8>.
- Aquarone, F., Nehéz-Posony, L., Anwar, P.R., Salam, S., Koutsouri, E., Kim, M., Suh, S., Mayomi, T., Pilarska, J., Houghton, E., Boodai, Y., 2020. "We're Trying to Do Things Differently": the Challenges of Relationships and Recognition in Higher Education. Centre for Public Policy Research (CPPR), King's College London. <https://www.kcl.ac.uk/ecs/assets/were-trying-to-do-things-differently-ebook.pdf>.
- Bala, N., Kahn, J., 2021. Exploring belonging in pedagogical partnership programs. *Int. J. Stud. Partners* 5 (2), 134–140. <https://doi.org/10.15173/ijasp.v5i2.4563>.
- Bala, N., 2021. A STEPP into uncertainty: pursuing passions to embrace pedagogical risks. *Teach. Learn. Together High. Educ.* 33. <https://repository.brynmawr.edu/tithe/vol1/iss33/4>.
- Bell, A., Barahona, S., Beg, G., Coulson, S., Eymont, R., Hartman, J., Hubble, T., Leung, N., McDonnell, M., Ni, J., Peseta, T., Sakhaee, E., Uptin, J., 2020. Students and academics working in partnership to embed cultural competence as a graduate quality. In: *Cultural Competence and the Higher Education Sector: Australian Perspectives, Policies and Practice*. Springer, 9789811553615.
- Bergmark, U., Westman, S., 2016. Co-creating curriculum in higher education: promoting democratic values and a multidimensional view on learning. *Int. J. Acad. Dev.* 21 (1), 28–40. <https://doi.org/10.1080/1360144X.2015.1120734>.
- Bindra, G., Easwaran, K., Firasta, L., Hirsch, M., Kapoor, A., Sosnowski, A., Stec-Marksmann, T., Vatansever, G., 2018. Increasing representation and equity in students as partners initiatives. *Int. J. Stud. Partners* 2 (2), 10–15. <https://doi.org/10.15173/ijasp.v2i2.3536>.
- Bovill, C., Cook-Sather, A., Felten, P., Millard, L., Moore-Cherry, N., 2016. Addressing potential challenges in co-creating learning and teaching: overcoming resistance, navigating institutional norms and ensuring inclusivity in student-staff partnerships. *High. Educ.* 71 (2), 195–208. <https://link.springer.com/article/10.1007/s10734-015-9896-4>.
- Bovill, C., 2017. A framework to explore roles within student-staff partnerships in higher education: which students are partners, when, and in what ways? *Int. J. Stud. Partners* 1 (1). <https://doi.org/10.15173/ijasp.v1i1.3062>.
- Bovill, C., 2019. A co-creation of learning and teaching typology: what kind of cocreation are you planning or doing? *Int. J. Stud. Partners* 3 (2), 91–98. <https://doi.org/10.15173/ijasp.v3i2.3953>.
- Bovill, C., 2020. Co-creation in learning and teaching: the case for a whole-class approach in higher education. *High. Educ.* 79, 1023–1037. <https://doi.org/10.1007/s10734-019-00453-w>.
- Bryson, C., Callaghan, L., 2021. A whole cohort approach to working in partnership between students and staff: problematising the issues and evaluating the outcomes. *Stud. Engagem. High. Educ.* J. 3 (2), 176–196. <https://sehej.raise-network.com/raise/article/view/1023>.
- Cates, R.M., Madigan, M.R., Reitenauer, V.L., 2018. "Locations of possibility": critical perspectives on partnership. *Int. J. Stud. Partners* 2 (1), 33–46. <https://doi.org/10.15173/ijasp.v2i1.3341>.
- Chen, J., Ho, J., 2020. A medical humanities curriculum in medical school: unexpected partnerships and unintended consequences. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books, Lanham, MD.
- Chng, H.H., 2019. The possibilities of students as partners—a perspective from Singapore. *Teach. Learn. Together High. Educ.* 27. <https://repository.brynmawr.edu/tithe/vol1/iss27/3>.
- Chukwu, A., Jones, K., 2020. Feminist interventions in engineering: co-creating across disciplines and identities. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Pedagogical Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books.
- Colón García, A., 2017. Building a sense of belonging through pedagogical partnership. *Teach. Learn. Together High. Educ.* 22. <http://repository.brynmawr.edu/tithe/vol1/iss22/2>.
- Cook-Sather, A., Agu, P., 2013. Students of color and faculty members working together toward culturally sustaining pedagogy. In: Groccia, J.E., Cruz, L. (Eds.), *To Improve the Academy: Resources for Faculty, Instructional, and Organizational Development*, vol. 32. Jossey-Bass, San Francisco, pp. 271–285. <https://quod.lib.umich.edu/t/tia/17063888.0032.020/-16-student-consultants-of-color-and-faculty-members-working?rgn=main;view=fulltext;q1=Agu>.
- Cook-Sather, A., Alter, Z., 2011. What is and what can be: how a liminal position can change learning and teaching in higher education. *Anthropol. Educ. Q.* 42 (1), 37–53.
- Cook-Sather, A., Felten, P., 2017. Ethics of academic leadership: guiding learning and teaching. In: Wu, F., Wood, M. (Eds.), *Cosmopolitan Perspectives on Academic Leadership in Higher Education*. Bloomsbury Academic, London, pp. 175–191.
- Cook-Sather, A., Matthews, K.E., 2021. Pedagogical partnership: engaging with students as co-creators of curriculum, assessment, and knowledge. In: Hunt, L., Chalmers, D. (Eds.), *University Teaching in Focus: A Learning-Centred Approach*, second ed. Routledge, pp. 243–259.
- Cook-Sather, A., Mejia, Y., 2018. Students Experience Empowerment and Empathy Through Pedagogical Partnership. <https://www.bera.ac.uk/blog/students-experience-empowerment-and-empathy-through-pedagogical-partnership>.
- Cook-Sather, A., Bovill, C., Felten, P., 2014. *Engaging Students as Partners in Learning & Teaching: A Guide for Faculty*. Jossey-Bass, San Francisco.
- Cook-Sather, A., Bahti, M., Ntem, A., 2019a. Pedagogical Partnerships: A How-to Guide for Faculty, Students, and Academic Developers in Higher Education. Elon University Center for Engaged Learning. <https://doi.org/10.36284/celelon.0a1>.
- Cook-Sather, A., Krishna Prasad, S., Marquis, E., Ntem, A., 2019b. Mobilizing a culture shift on campus: underrepresented students as educational developers. *N. Dir. Teach. Learn.* 159, 21–30. Retrieved from: <https://onlinelibrary.wiley.com/toc/15360768/2019/2019/159>.
- Cook-Sather, A., Matthews, K.E., Bell, A., 2019c. Transforming curriculum development through co-creation with students. In: Quinn, L. (Ed.), *Reimagining Curriculum: Spaces for Disruption*. AFRICAN SUN MeDIA.
- Cook-Sather, A., Wilson, C. (Eds.), 2020. *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books.
- Cook-Sather, A., Healey, M., Matthews, K.E., 2021a. Recognizing students' expertise and insights in expanding forms of academic writing and publishing about learning and teaching. *Int. J. Stud. Partners* 5 (1), 1–7. <https://doi.org/10.15173/ijasp.v5i1.4626>.
- Cook-Sather, A., Slaters, S., Allen, M., Baskaran, R., Crombie, P., Derounian, J.G., Dianati, S., Dudson, M., Grayson, D., Haroon, M.A., Fraser, C., Jardine, H., Millmore, A., Milton, E., Mori, Y., Morgan, A., Pereira, J.V.-B., Salman, M., Scharff, L., Storey, J., Sum, K., Tisdell, C.C., 2021b. The many manifestations of successful partnership. *Int. J. Stud. Partners* 5 (2), 221–232. <https://doi.org/10.15173/ijasp.v5i2.4911>.
- Cook-Sather, A., Ho, L., Kaur, A., Tamim, T. (forthcoming). Translating pedagogical partnership in/to academic staff development in the Global South. In: Chitanand, N., Rathilal, S. (Ed.), *Academic Staff Development: Disruptions, Complexities, Change*. African Sun Media.
- Cook-Sather, A., 2009. From traditional accountability to shared responsibility: the benefits and challenges of student consultants gathering midcourse feedback in college classrooms. *Assess. Eval. High. Educ.* 34 (2), 231–241.
- Cook-Sather, A., 2020. Respecting voices: how the co-creation of teaching and learning can support academic staff, underrepresented students, and equitable practices. *High. Educ.* 79 (5), 885–901. <https://link.springer.com/article/10.1007/s10734-019-00445-w> <https://doi.org/10.1007/s10734-019-00445-w>.
- Cook-Sather, A. (in press). *Co-Creating Equitable Teaching and Learning: Structuring Student Voice into Higher Education*. Boston, Harvard Education Press.

- de Bie, A., Marquis, E., Cook-Sather, A., Luqueño, L., 2019. Valuing knowledge(s) and cultivating confidence: contributions of student-faculty pedagogical partnerships to epistemic justice. In: Hoffman, J., Blessinger, P., Makhanya, M. (Eds.), *Strategies for Fostering Inclusive Classrooms in Higher Education: International Perspectives on Equity and Inclusion* (Innovations in Higher Education Teaching and Learning, Vol. 16). Emerald Publishing Limited, pp. 35–48. <https://doi.org/10.1108/S2055-364120190000016004>.
- de Bie, A., Marquis, E., Cook-Sather, A., Luqueño, L.P., 2021. *Promoting Equity and Justice Through Pedagogical Partnership*. Stylus Publishers, Virginia. <https://www.centerforengagedlearning.org/books/promoting-equity-and-justice-through-pedagogical-partnership/>.
- Deeley, S.J., Bovill, C., 2017. Staff student partnership in assessment: enhancing assessment literacy through democratic practices. *Assess. Eval. High. Educ.* 42 (3), 463–477. <https://doi.org/10.1080/02602938.2015.1126551>.
- Deeley, S.J., Brown, R.A., 2014. Learning through partnership in assessment. *Teach. Learn. Together High. Educ.* 13. <http://repository.brynmawr.edu/tithe/vol1/iss13/3>.
- Delpish, A., Holmes, A., Knight-McKenna, M., Mihans, R., Darby, A., King, K., Felten, P., 2010. Equalizing voices: student-faculty partnership in course design. In: Werder, C., Otis, M.M. (Eds.), *Engaging Student Voices in the Study of Teaching and Learning*. Stylus, Sterling, VA, pp. 96–114.
- Felten, P., Abbot, S., Kirkwood, J., Long, A., Lubicz-Nawrocka, T., Mercer-Mapstone, L., Verwoord, R., 2019. Reimagining the place of students in academic development. *Int. J. Acad. Dev.* 24 (2), 192–203. <https://doi.org/10.1080/1360144X.2019.1594235>.
- Felten, P., 2013. Principles of good practice in the scholarship of teaching and learning. *Teach. Learn. Inq.* 1 (1), 121–125. <https://doi.org/10.2979/teachlearninq.1.1.121>.
- Foran, G., Knorr, K., Taylor, R.L., 2020. Evaluating factors contributing to positive partnership work in a students-as-consultants partnership program. *Int. J. Stud. Partners* 4 (2), 27–44. <https://doi.org/10.15173/ijasp.v4i2.4095>.
- Gibson, S., Cook-Sather, A., 2020. Politicised compassion and pedagogical partnership: a discourse and practice for social justice in the inclusive academy. *Int. J. Stud. Partners* 4 (1). <https://mulpress.mcmaster.ca/ijasp/article/view/3996>.
- Godbold, N., Hung, T.Y., Matthews, K.E., 2021. Exploring the role of conflict in co-creation of curriculum through engaging students as partners in the classroom. *High. Educ. Res. Dev.* 1–15.
- Healey, M., Healey, R.L., 2018. "It depends": exploring the context-dependent nature of students as partners' practices and policies. *Int. J. Stud. Partners* 2 (1). <https://doi.org/10.15173/ijasp.v2i1.3472>.
- Healey, M., Flint, A., Harrington, K., 2014. *Students as Partners in Learning and Teaching in Higher Education*. Higher Education Academy, York. <https://www.heacademy.ac.uk/knowledge-hub/engagement-through-partnership-students-partners-learning-and-teaching-higher>.
- Healey, M., Flint, A., Harrington, K., 2016. Students as partners: reflections on a conceptual model. *Teach. Learn. Inq.* 4 (2). <https://files.eric.ed.gov/fulltext/EJ1148481.pdf>.
- Hernandez, P., 2015. *The Pedagogy of Real Talk: Engaging, Teaching, and Connecting with Students at Risk*. Corwin.
- hooks, b., 1994. *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, New York.
- Ifitkhar, F., 2021. The importance of trust in student-faculty partnerships: my journey from a student partner to a co-lead of a pedagogical partnership program. *Teach. Learn. Together High. Educ.* 33. <https://repository.brynmawr.edu/tithe/vol1/iss33/9>.
- Impastato, J., 2021. "Very nervous and very excited": my experience developing a pedagogical partnership program at tufts university. *Teach. Learn. Together High. Educ.* 33. <https://repository.brynmawr.edu/tithe/vol1/iss33/5>.
- Jonsson, M., 2020. Five reasons why working as a student partner is energizing. *Int. J. Stud. Partners* 4 (2), 150–154. <https://doi.org/10.15173/ijasp.v4i2.4336>.
- Kaur, A., Yong Bing, T., 2020. Untangling the power dynamics in forging student-faculty collaboration. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books, Lanham, MD.
- Kaur, A., 2020. Students as partners: challenges and opportunities in the Asian context. *Int. J. Stud. Partners* 4 (20), 145–149. <https://doi.org/10.15173/ijasp.v4i2.4366>.
- Kehler, A., Verwoord, R., Smith, H., 2017. We are the process: reflections on the underestimation of power in students as partners in practice. *Int. J. Stud. Partners* 1 (1). <https://doi.org/10.15173/ijasp.v1i1.3176>.
- Khouri, F., Oberhollenzer, Y., Matthews, K., 2017. *Co-designing a Community of Practice for Students as Partners in Partnership*. University of Queensland, Brisbane. <https://core.ac.uk/download/pdf/84596576.pdf>.
- Leota, A., Sutherland, K., 2020. "With your basket of knowledge and my basket of knowledge, the people will prosper": learning and leading in a student-staff partnership program. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books.
- Liang, Y., Matthews, K.E., 2021. Students as partners practices and theorisations in Asia: a scoping review. *High. Educ. Res. Dev.* 40 (3), 552–566. <https://www.tandfonline.com/doi/abs/10.1080/07294360.2020.1773771>.
- Liang, Y., Matthews, K.E., forthcoming. Are Confucian educational values a barrier to engaging students as partners in Chinese universities?
- Marquis, E., Jayaratnam, A., Mishra, A., Rybina, K., 2018. "I feel like some students are better connected": students' perspectives on applying for extracurricular partnership opportunities. *Int. J. Stud. Partners* 2 (1). <https://doi.org/10.15173/ijasp.v2i1.3300>.
- Marquis, E., Carrasco-Acosta, E., de Bie, A., Krishna Prasad, A., Wadhvani, S., Woolmer, S., 2019. *Pedagogical Partnerships and Equity in the Classroom: Insights From One Partnership Program*. Presentation at the 2019 Symposium on Scholarship of Teaching and Learning. Banff, AB., November 7–9.
- Marquis, E., de Bie, A., Cook-Sather, A., Krishna Prasad, S., Luqueño, L., Ntem, A., 2021. "I saw a change": enhancing classroom equity through student-faculty pedagogical partnership. *Can. J. Scholarsh. Teach. Learn.* 12 (1). <https://doi.org/10.5206/cjsotlrceae.2021.1.10814>.
- Matthews, K.E., Mercer-Mapstone, L., Dvorakova, S.L., Acai, A., Cook-Sather, F.P., Healey, R.L., Healey, M., Marquis, E., 2019. Enhancing outcomes and reducing inhibitors to the engagement of students and staff in learning and teaching partnerships: implications for academic development. *Int. J. Acad. Dev.* 24 (3), 246–259. <https://doi.org/10.1080/1360144X.2018.1545233>.
- Matthews, K.E., 2019. Rethinking the problem of faculty resistance to engaging with students as partners in learning and teaching in higher education. *Int. J. Scholarsh. Teach. Learn.* 13 (2). <https://doi.org/10.20429/ijstl.2019.130202>. Article 2.
- Matthews, K.E., 2021. *Energizing Institutional Assessment Activities With Students as Partners*. IUPUI Assessment Institute.
- Mercer-Mapstone, L., Bovill, C., 2019. Equity and diversity in institutional approaches to student-staff partnership schemes in higher education. *Stud. High. Educ.* <https://doi.org/10.1080/03075079.2019.1620721>.
- Mercer-Mapstone, L., Dvorakova, S.L., Matthews, K.E., Abbot, S., Cheng, B., Felten, P., Knorr, K., Marquis, E., Shammass, R., Swaim, K., 2017. A systematic literature review of students as partners in higher education. *Int. J. Stud. Partners* 1 (1). <https://mulpress.mcmaster.ca/ijasp/issue/view/306>.
- Mercer-Mapstone, L., Islam, M., Reid, T., 2019. Are we just engaging "the usual suspects"? Challenges in and practical strategies for supporting equity and diversity in student-staff partnership initiatives. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2019.1655396>.
- Mihans, R., Long, D., Felten, P., 2008. Power and expertise: student-faculty collaboration in course design and the scholarship of teaching and learning. *Int. J. Scholarsh. Teach. Learn.* 2 (2). <https://doi.org/10.20429/ijstl.2008.020216>. Article 16.
- Monsen, S., Cook, S., Hannant, L., 2017. Students as partners in negotiated assessment in a teacher education course. *Teach. Learn. Together High. Educ.* 21. <http://repository.brynmawr.edu/tithe/vol1/iss21/2>.
- Moorhouse, B.L., Oh, M.H.M., 2019. Students as partners beyond formal education: a mentoring partnership in the first year of teaching. *Int. J. Stud. Partners* 3 (2), 156–162. <https://doi.org/10.15173/ijasp.v3i2.3893>.
- Murphy, R., Nixon, S., Brooman, S., Fearon, D., 2017. "I am wary of giving too much power to students:" addressing the "but" in the principle of staff-student partnership. *Int. J. Stud. Partners* 1 (1). <https://doi.org/10.15173/ijasp.v1i1.3055>.
- Narkiss, D., Naaman, I., 2020. Voicing and reflecting in a pedagogical partnership. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books, Lanham, MD.

- Ntem, A., Cook-Sather, A., 2018. Resistances and resiliencies in pedagogical partnership: student partners' perspectives. *Int. J. Stud. Partners* 2 (1). <https://doi.org/10.15173/ijasp.v2i1.3372>.
- Ntem, A., Nguyen, E., Rafferty, C., Kwan, C., Benlahcene, A., 2020. Students as partners in crisis? Student co-editors' perspectives on COVID-19, values, and the shift to virtual spaces. *Int. J. Stud. Partners* 4 (2), 1–8. <https://doi.org/10.15173/ijasp.v4i2.4432>.
- Ntem, A., 2016. Learning to be brave within and beyond partnership. *Teach. Learn. Together High. Educ.* 18. <https://repository.brynmawr.edu/tithe/vol1/iss18/6>.
- Ortquist-Ahrens, L., 2021. Building partnership through partnership: reflections on the role of post-baccalaureate fellows in program development. *Int. J. Stud. Partners* 5 (2), 191–197. <https://doi.org/10.15173/ijasp.v5i2.4618>.
- Peseta, T., Pizzica, J., Beathe, A., Lynch, R., Manthos, M., Nguyen, K., Raza, H., 2020. A partnership mindset. In: Mercer-Mapstone, L., Abbot, S. (Eds.), *The Power of Student-Staff Partnership: Revolutionizing Higher Education*. Elon University Center for Engaged Learning Open Access Series. <https://www.centerforengagedlearning.org/books/power-of-partnership/section-two/chapter-6/>.
- Quy, K., Fuller, A., Kar, A., Baldwin, M.T., Hallab, O., 2020. Learning through cooperation: interdisciplinary insights into students' experiences in a developing program. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching through Student-Faculty Partnership: Stories from across Contexts and Arenas of Practice*. Lexington Books, Lanham, MD.
- Reynolds, W., Greene, J., Halpern, C., 2019. Student-Faculty Partnerships: Assessing a New Program. [Conference presentation], Pittsburgh, PA.
- Seow, Y., 2019. Taking a small step toward partnership. *Teach. Learn. Together High. Educ.* 27. <https://repository.brynmawr.edu/tithe/vol1/iss27/5/>.
- Sim, J.Y.H., 2019. The “face” barriers to partnership. *Teach. Learn. Together High. Educ.* 27. <https://repository.brynmawr.edu/tithe/vol1/iss27/4/>.
- Topper, L., 2021. Building partnership through partnership. *Teach. Learn. Together High. Educ.* 33. <https://repository.brynmawr.edu/tithe/vol1/iss33/6>.
- Vayada, P., Liang, Y., Matthews, K.E., 2020. We are all in this together. *High. Educ. Res. Dev. Soc. Australas. Connect* 42 (3), 15. <https://www.hersa.org.au/sites/default/files/HERDSA%20CONNECT%20-%20spring%202020.pdf>.
- Vayada, P., 2021. Being a mature-aged student partner: joy, alienation, and disappointment. In: Simmons, N., Podrebarac, J. (Eds.), *Adults in the Academy: Voices of Lifelong Learners*, pp. 27–34.
- Waqar, Y., Asad, A.M., 2020. Student as co-designer: processes of planning and teaching with the student in mind. In: Cook-Sather, A., Wilson, C. (Eds.), *Building Courage, Confidence, and Capacity in Learning and Teaching Through Student-Faculty Partnership: Stories From Across Contexts and Arenas of Practice*. Lexington Books, Lanham, MD.
- Wilson, C., Cook-Sather, A. (in press). Rippling the patterns of power: enacting anti-racist pedagogy with students as co-teachers. In: Neuhaus, J. (Ed.), *Picture a Professor: Intersectional Teaching Strategies for Interrupting Bias About Faculty and Increasing Student Learning*. West Virginia University Press.
- Yahlnaaw, 2019. T'aats'iigang: stuffing a jar full. *Int. J. Stud. Partners* 3 (2), 6–10. <https://doi.org/10.15173/ijasp.v3i2.4081>.
- Zhang, M., Matthews, K.E., Lui, S. (in press) Recognising cultural and symbolic capital through shared meaning-making in cross-cultural partnership practices. *IJSaP* 6 (2).

Implications of the Fourth Industrial Revolution in the vortex of higher education in South Africa

Narend Bajinath, Shireen Motala, and Kirti Menon

© 2023 Elsevier Ltd. All rights reserved.

Introduction	279
What is the Fourth Industrial Revolution?	279
Socioeconomic realities confronting SA higher education	280
Supportive infrastructure, access to devices and cost-effective connectivity in higher education	281
The potential for higher education, teaching and learning	281
Enabling conditions	283
Implications for curriculum	283
Imperatives for change flowing from the 4IR	284
Conclusion	285
References	285

Introduction

The emerging Fourth Industrial Revolution (4IR) has gained widespread recognition, as has the imperative to ensure informed and proactive responses at state, sectoral and institutional levels. This paper aspires to assist our understanding of this phenomenon in South Africa, mindful of the many dimensions of the 4IR and its promise for national development and the higher education sector. It further considers interventions to leverage the potential of the 4IR in higher education while seeking to understand wider policy, planning, and development implications. Finally, challenges and obstacles are identified, many of which relate directly to South Africa's socioeconomic conditions.

According to the World Bank, South Africa is one of eight upper-middle-income economies in South Africa. Despite this, and since 1994, the socio-economic status of much of South Africa's population has not changed. The Gini coefficient of per capita household income is usually well above 0.6.6 according to [Statistics South Africa \(2019\)](#). On a global basis therefore, this puts South Africa among the "top 5" most unequal countries. Despite disagreement over the actual levels of inequality, the experts agree that inequality in South Africa is still extremely high and has remained unchanged since 1993 ([Statistics South Africa, 2019](#)). Consumption disparities or inequalities persist over time due to high wealth inequality, which inhibits intergenerational economic mobility ([World Bank, 2022](#)). In the labor market, having a post-secondary or university degree is essential for both finding work and earning higher income. The significant inequality in land ownership continues to contribute to the perpetuation of historically high levels of income inequality in Namibia and South Africa ([World Bank, 2022](#)). These two countries have far greater levels of inequality than Botswana, Eswatini, Lesotho who also have wealth inequalities ([World Bank, 2022](#)). In South Africa, the richest 10% possess 86% of all the wealth, with the top 0.1% owning nearly a third ([Chatterjee et al., 2022](#)). More than 15% of household net worth is concentrated in the top 0.01% of the distribution (3500 individuals), compared to the bottom 90% as a whole ([Chatterjee et al., 2022](#)). Since the end of apartheid, there has been little sign of decreasing poverty ([Chatterjee et al., 2022](#)). Despite progress toward equality and redress in post-apartheid South Africa, the reality is that an estimated 48% of the population lives on less than \$2 per day, according to the [National Planning Commission \(2011\)](#).

As in other countries, wealth inequality in South Africa affects both basic and higher education. The historical and current racial, economic, and intellectual conditions result in a two-tier system, with historically advantaged institutions desegregating by attracting middle-class black students ([Motala et al., 2021](#)). Historically disadvantaged institutions' students are largely black working-class ([Motala et al., 2021](#)). [Mamdani \(2016\)](#) argues that the gap between higher education institutions (HEIs) in South Africa reflects the continuous divide between quality and relevance, as previously advantaged institutions rely on excellence to maintain their advantage, networks, and academic autonomy. It seems there is no other country with the unique distinctions between historically advantaged and disadvantaged institutions ([Mamdani, 2016](#)).

With these ongoing disparities in South Africa and across higher education, the 4IR could exacerbate inequality. It is unclear if historically disadvantaged institutions have the infrastructure to embrace the 4IR. It is not all doom and gloom. The 4IR provides a glimmer of hope in an education system with high disparities. There is a possibility that it could increase the rates of access and success in higher education. Conditions need to be put in place so that the introduction of 4IR in all HEIs be done in such a way that it does not widen existing gaps.

What is the Fourth Industrial Revolution?

[Schwab \(2015\)](#) is the leading and arguably most influential proponent of the idea of the 4IR, or Industry 4.0 ([Marwala, 2020a](#), p. 25). However, there are disagreements. Industry 4.0 has been seen as part of the 4IR with its roots in German manufacturing in the

early 2000s (Gleason, 2018, p. 2), suggesting that developments associated with the 4IR have occurred over time rather than in a revolutionary moment. They should appositely be labeled *evolutionary*, not *revolutionary*.

Schwab foresees that the 4IR “will fundamentally alter the way we live, work and relate to one another. In its scale, scope and complexity, the transformation will be unlike anything humankind has experienced before” (Schwab, 2015, p. 1). Gastrow elaborates that humanity is “undergoing multiple changes: increased longevity, augmentation of human bodies and minds, the changing nature of work, the changing nature of learning, changes to human connection and connectivity, and changes to identities, among others” (Gastrow, 2018, p. 10). Schwab’s signaling of the far-reaching and unprecedented nature of anticipated changes constitutes the main argument for the 4IR as a *revolution*. Innovation stemming from 4IR technologies, impels leaders to understand and respond to the changing environment. Schwab imagines citizen engagement with governments using emerging technologies, predicting that governments will harness technology to exercise greater control and surveillance (Schwab, 2015, pp. 1–6).

Risking (2016) argues that the 4IR is merely continued digitalization, not fundamentally different from the 3IR (in Jivan, 2020, p. 7). Marwala (2020a,b) distinguishes between the 4IR and the 3IR, or digital revolution, maintaining that the 4IR is focused with the making of intelligent digital infrastructure that seamlessly use data to improve, integrate and advance economies and societies.

Schwab’s response is that while features of the digital revolution have provided the foundations on which advances in the 4IR have been built, the 4IR’s scope, speed and system-wide impact is qualitatively different (Jivan, 2020, p. 7). Beyond changing the nature of work and productive systems, it signals critical changes to how knowledge is used, developed, disseminated and shared. Universities play a vital role in these processes. From an industry perspective, Deloitte adds that the need is apparent “to nurture a culture of lifelong learning, equipping their workforces with the skills necessary to succeed in the future” (Deloitte, 2020, p. 2). Automation will impact on jobs. It is projected that 50% of jobs in agriculture, forestry, fishing, and hunting, representing 328 million workers, are automatable. Figures for manufacturing are 237.4 million; for retail 187 million; for China 395 million; for India 235 million; for the USA 60.6 million (Gleason, 2018, p. 4, quoting a McKinsey Global Institute Report of 2017). While automation will benefit production, catering for an ever-growing world population, the socio-political impact of the loss of livelihoods demands attention, raising the question of the interplay between technological and social change (Gastrow, 2018, p. 23). Deloitte further proposes that organizations begin to gather data that will help them develop a deeper insight into the future so that their strategies related to Industry 4.0 are based on firmer foundations (Deloitte, 2020, pp. 5–7).

In South Africa, responses include initiatives driven by the Presidency and a Ministerial Task Team (MTT) on the 4IR established by the Minister of Higher Education and Training. The South African Presidential Commission on the Fourth Industrial Revolution (PC4IR) reported in October 2020 with a vision of 4IR as potentially enabling attainment of the National Development Plan (NDP) goals (PC4IR, p. 137). Policy interventions such as the PC4IR (2020), and MTT (DHET, 2020b), acknowledge South Africa’s need for a coordinated and development-orientated approach, with the 4IR in a potentially catalytic role. Gastrow, like others, (Gleason, 2018; Marwala, 2020a), has argued for the momentous and imminent significance of the 4IR, noting its potential developmental benefits. He contends that being prepared means positioning institutions so that “the 4IR is harnessed for the benefit of human wellbeing, and in support of national and international social and economic benefits” (Gastrow, 2018, p. 10), a view in accordance with the NDP (National Planning Commission, 2011), and PC4IR (2020).

Socioeconomic realities confronting SA higher education

This positioning of the 4IR occurs in concert with the socio-economic challenges that are currently urgent priorities in national development in South Africa. Having inherited structural inequalities from the apartheid regime, the country continues to be among one of the most unequal in the world. In the labor market, having a post-secondary or university degree is essential for both finding work and earning higher income.

A snapshot of the degree of inequality in the country can be described by the following statistics. At the top end, all types of assets, including homes, pension funds, and financial assets, can account for such levels of inequality. Since the end of apartheid, there has been little sign of decreasing poverty with 86% of wealth in the country belonging to 10% of the population (Chatterjee et al., 2022). These inequalities have affected every sphere of society in South Africa, including basic and higher education, which are regularly held as examples of the lasting impact of apartheid social engineering on the shaping of progress and opportunity.

The racial, economic, and intellectual conditions in South Africa’s higher education have resulted in a two-tier system affecting every aspect of how institutions operate. Historically white institutions continue to be research-intensive, attract significant funding, and attract high-quality staff which in turn allows for greater selectivity of students (Motala et al., 2021). Historically black and disadvantaged institutions, on the other hand, continue to cater to working-class black students, are heavily reliant on (reducing) government funding and financial aid, and lack the enabling conditions to support research due to heavy teaching loads and academic support requirements (Mouton et al., 2015, 2021), which has consequences for throughput, graduation, and post-university prospects. These were just some of the cleavages contributing to the eruption of the #RhodesMustFall and #FeesMustFall student movement in 2015.

While the 4IR could exacerbate inequalities, it offers the prospect of increasing rates of access and success in the system through targeted interventions using the enhanced technologies on offer. South Africa is not a unique case to higher education systems around the world, especially in post-colonial settings, which have had to address access and equity to unlock opportunities for marginalized groups. However, given the high degrees of inequality in the country, intimately coupled with a history of racial

segregation, the country represents a unique case for understanding both the urgency of social justice and the complex challenges that emerge in the course of addressing it—which may perhaps be further complicated by the prospects and possibilities of the 4IR.

Supportive infrastructure, access to devices and cost-effective connectivity in higher education

National infrastructure to support the 4IR in South Africa is critical to success. Marwala poses the dilemma: “How can we possibly focus on the 4IR when we are still struggling with the second industrial revolution?” (2020a, p. 58). South Africa is in the early stages of 4IR preparedness, including in the higher education sector (Marwala, 2020a, p. 25). Encouragingly, a high percentage of African institutions indicate that student enrollment is undertaken online: 53% fully online, 37% partly so (Jensen, 2019, p. 26). This is unexpectedly high given the resourcing and bandwidth challenges in many African countries.

Reinitz (2020) cautions against assuming that putting courses online or using a technological solution is digital transformation. Developing Risking’s position, Reinitz distinguishes between *digitization*, constituting transition from analog to digital form, entailing organization of vast amount of online information; *digitalization*, using digital technologies to transform institutional operations, and *digital transformation* (Reinitz, 2020, p. 2). *Digital transformation* comprises “an institutional strategy to transform the strategic direction or value proposition of the college or university,” leading to institutional transformation (Reinitz, 2020, p. 2). This is perhaps the most promising approach to leveraging the affordances of the 4IR. From this vantage point, true transformation is thoroughgoing and coordinated, scrutinizing the strategy, value proposition, and business model of the organization (Reinitz, 2020, p. 1).

There are stark inequalities in the African context. In Africa, 24.4 individuals per 100 used the internet in 2018, compared to 45.3 in the developing world and 80.9 in the developed world. Only 7% in Africa considered internet infrastructure very satisfactory, with 39% indicating that it was good in big cities but poor in rural areas (Jensen, 2019, pp. 16–17). These figures suggest a huge obstacle to harnessing technology and taking advantage of the 4IR. In South Africa, the DHET survey revealed that most students own smartphones and at least 60% own a laptop. Devices have been acquired in a variety of ways, with at least 44% paid for through NSFAS or other bursary funds (DHET, 2020a). Only 50% of students reported good connectivity where they are studying. Most bought data bundles from service providers (62%), while 46% utilized free access through the university. Other alternatives included public Wi-Fi (8%); home facilities (16%); monthly SIM contracts (9%); and hot spotting off someone else (23%) (DHET, 2020a, pp. 28–31). The dependence of a large proportion of students on their universities to provide Wi-Fi access is significant. Cost of data in South Africa, among the highest in the world, remains prohibitive for most students. This constitutes a barrier to students using online modalities for learning, communication, and support.

The IAU study reveals that 24% of institutions in Africa report using learning management systems fully compared to a global average of 35%. A further 46% indicate that they are using them somewhat, 19% report it as being discussed and 11% indicate it is not on the radar (Jensen, 2019, p. 27). The management of the COVID-19 pandemic in higher education in South Africa demonstrated that institutions held varying capacity to respond to the multipronged needs of students and staff in the context of increased crisis (Menon and Motala, 2021).

The 4IR Task Team identified challenges to tertiary and vocational education that required addressing in the process of optimizing the sector. Greater flexibility is needed in admissions, curriculum design, teaching, learning and assessment, with more responsive quality assurance and support to undergird these processes (DHET, 2020b, p. 40). The necessity of flexibility and for institutions to respond dexterously, unfettered by bureaucratic oversight and regulation, cannot be overemphasized. Considering all this, the PC4IR Report states that the 4IR provides “an opportunity to redesign, streamline and align the education system through a coordinated, robust, multi-stakeholder process. The purpose of the next version of our skills ecosystem will be to leapfrog our youth into productive work and reskill current workers for job retention and ongoing productive work in the economy” (PC4IR, 2020, p. 173).

The potential for higher education, teaching and learning

Universities are changing. Williamson and Hogan contend that the ideal is shifting to the “‘data-intensive’ and ‘digital-first universities’ ... that creates valuable new digital knowledge and develops digital data skills to support emerging capitalist data economies.” (Williamson and Hogan, 2021, p. 13, referencing Berry, 2020). Badat (2020) has contested Marwala’s (2020b) upbeat vision of a 4IR future on a similar basis, and cautioned against pursuing a “technocratic utopia,” drawing attention to the restricted and uneven benefits of previous industrial revolutions. Marwala argues that a novel feature is that it “is based not on a single technology, but on the confluence of multiple developments and technologies,” changing life, work, and communication and “reshaping government, education, healthcare, and commerce.”

Badat (2020) acknowledged that “New developments in knowledge, science, technology and ways of economic production are the probable trajectory of the future” but warned of the insidious effects “manifest in rampant inequality and the massive accumulation of wealth and power in the hands of a tiny minority.” Central to his critique is that the 4IR is viewed as entirely scientific and technological, “not as a social or human matter,” especially in a society as deeply fragmented as South Africa. Indeed, increased unemployment and other regressive outcomes are entirely possible, instead of the anticipated developmental surge. Paradoxically,

the 4IR poses huge threats to a society struggling to shake off its grievous inheritance of poverty, inequality, unemployment, and lack of opportunities.

Gillwald, observing the relentless advance of 4IR technologies, warns about consequences. Their rise, she says, will disrupt economies and societies unevenly: "Like other industrial revolutions, this one will be characterized by evolution as much as by disruption. Unless something dramatically different is done, one of the continuities will be the perpetuation of inequality. And the primary determinant of inclusion is education and digital skilling" (Gillwald, 2019, p. 1). Indeed, a social justice perspective obliges consideration of how the already disadvantaged and marginalized may benefit from the 4IR, failing which inequality is likely to increase.

While a good start at state level would be progressive policies, solutions focusing on skills development may not be entirely relevant. Vally and Motala (2014, p. 4) argue that the skills discourse "has 'persuaded' policy makers to ignore the combination of social, economic, political and cultural factors which make the acquisition of knowledge and skills difficult for working-class and poor communities." Like Fataar, they contend that it "downplays the impact of poor access to educational institutions and costs, locally-based educational facilities and infrastructure, pedagogical barriers arising from being taught in unfamiliar languages together with the relative powerlessness to intervene in ways that can counteract the weight of existing relations of power" (Vally and Motala, 2014, p. 4; Fataar, 2020, p. 4). Clearly expectations surrounding the 4IR must be tempered by consideration of underlying inequality, and how the underprivileged may be impacted as the 4IR advances.

Presciently, Jensen warns of the cost of neglecting social justice, and the responsibility that therefore devolves on HEIs. He argues that "knowledge divides," excluding the underprivileged from active roles in society, are more significant than "digital divides." "Higher education institutions are at the heart of knowledge creation and dissemination; it is therefore only natural that higher education takes an active part in shaping a knowledge society" (Jensen, 2019, p. 53). As institutions with immense capacity to inform, support and advocate social justice, HEIs will be failing in these duties should they not attend to access to and availability of technology to the poor and marginalized, as well as shape teaching and innovation toward the pursuit of more inclusive societies. This foregrounds both their intellectual and developmental role in society (Mamdani, 2016).

Technology in higher education has become an increasingly strategic priority for governments and universities. The International Association of Universities (IAU), with approximately 650 members, anticipated that technology would have implications for the future development of higher education and society. During 2018 it embarked on developing a policy statement on the subject with consultations held with higher education leaders from nearly 700 institutions from 106 countries (Jensen, 2019, pp. 8–9). Jensen describes how HEIs need to respond to a technologically transforming world which impacts on "the everyday lives of citizens, on how societies are developing, on the skills and competences required to take part in society; and most importantly, on how to access information and knowledge." Although transformations are multifaceted, "one factor common to HEIs is that they are all confronted with the question of how to adapt and shape higher education in an increasingly digital world" (Jensen, 2019, p. 51). Given the obstacles related to infrastructure, data, and access to devices which have emerged in analyses such as those of the NDP and PC4IR, the IAU's findings that technology is being substantially integrated into teaching are encouraging. For African universities in the study, 38% of respondents indicated this is happening very much, while 49% indicated that it is happening to some extent (Jensen, 2019, p. 29). There is some way to go, but progress is noteworthy.

To illustrate, in South Africa, reforms aimed at harnessing the 4IR to national development goals are signaled in the PC4IR Report (PC4IR, 2020, p. 161). Driving the national 4IR initiative will be a Digital and Future Skills Forum which will probably provide guidelines for curriculum reform and projections of the skills and competencies required for future employment, cultural and intellectual enrichment and learning beyond the formal curriculum. Curricula will have to be geared to developing critical digital skills. Fundamental questions about the relevance, adequacy, and responsiveness of current curricula will doubtless loom large.

More sophisticated use of digital technology in tertiary education is required, starting with the currently unevenly spread 3IR technologies which continue to be a restraint to the 4IR. Online data and information management; data aggregation; and analyses of data for student success and support is needed to increase success and throughput. Possibilities emerge for more frequent assessment and for online learning management. In line with these possibilities, du Preez and Sinha argue for post school education to at least "adapt to the changing nature of work, and ideally, lead this evolution" (2021, p. 13). South African HEIs are addressing these imperatives, hastened by the COVID-19 pandemic that has shifted most teaching and learning online.

Most HEIs have responded to technological change within their capacities. Czerniewicz et al. (2021, p. 12) state that large-scale changes are refashioning higher education: "these changes are producing emergent forms of teaching and learning and question academic roles with regard [to] who has ownership and control over the teaching and learning process." They also argue that marketization (entrepreneurial competition), digitization (educational provision mediated by digital technology) and unbundling (disaggregation of educational provision through technology-based partnerships with private providers) are impelling new forms of teaching and learning (2021, pp. 1–2). They say that "there are polarized views (in the contexts of South Africa and the UK) about the value of digitally mediated forms of provision" (2021, p. 10). These findings signal the need to sift realistic and advantageous approaches from the sometimes near-evangelical promotion of resource-intensive and ever more complex technological solutions.

In higher education, more sophisticated use of digital technology for administration, teaching, learning support, and e-learning could have a discernible impact on quality, efficiency, and success. Online data and information management; aggregation of data; and analyses of data for student success and support have impacted success and throughput rates positively. In the final analysis, the use of ICTs must lead to improved student engagement and learning, as Mishra and Panda (2020, p. 4) and as Marwala advocates,

emphasizing “real-world problem-solving abilities and a multi-disciplinary approach to curricula that are more interactive” (Marwala, 2020a, p. 158).

The spread, reach and impact of technologies and services aimed at higher education are clearly underway. Indeed, they are ubiquitous: “The global industry of educational technologies and data services has grown to encompass every aspect or ‘market segment’ of HE activity” (Williamson and Hogan, 2021, p. 15). The benefits in efficiency, reach in recruitment, simpler business processes and enhanced learning support, among others, are clear, even as more pernicious effects emerge. These include increased unemployment, a widening digital and knowledge divide, and greater intrusion on privacy as data on individuals is collected and mined by state and commerce.

Enabling conditions

It has been noted in this discussion that the 4IR may have negative employment consequences. Marwala characterizes the future as “post-work ... because the need for humans in the workforce will be severely curtailed, effectively changing the face of labor.” Nonetheless, opting out of the 4IR is not advisable: “Either we participate, or as a country we are economically obliterated, relegated to the ‘dustbin of history’” (Marwala, 2020a, p. 130).

The 4IR has immense promise, but various enabling conditions are needed at national and institutional levels before tangible benefits can be realized. First is the imperative to harness the benefits of the 3IR, many of which are preconditions for the 4IR. The COVID-19 pandemic has had a discernible effect on access and attainment due to technological barriers, which emphasizes the urgency of the need for digital integration (Atherton, 2020, pp. 1–2).

Measuring success in the use of and student receptiveness to learning technology is important even before considering devices and connectivity. COVID-19 has exposed many students to the advantages of online instruction and the DHET survey indicates some satisfaction with it. A significant proportion of respondents indicated that learning materials provided clear learning outcomes (65%); gave access to relevant content (69%); explained new content and concepts carefully (60%); was well organized and sequenced (61%); and provided activities enabling critical engagement (61%) (DHET, 2020a,b, p. 34). Capacity notwithstanding, these findings signal the readiness of most public HEIs to shift decisively to online learning and support, and growing receptiveness among students.

The impact of online learning is yet to be established, but the DHET survey has shown that many students are utilizing it for various purposes. Ninety-two percent indicated they submit assignments online; 78% for downloading materials; 76% for writing tests; 73% for email communications; 71% for accessing learning materials; and 70% for discussions with classmates. A further 43% attended virtual lectures (DHET, 2020a, p. 29). Academics and learning support practitioners carry a similar responsibility to become proficient in digital tools and working with students in adapting to new learning modalities. Menon and Castrillon observe that “If HE is to prepare students for the challenges of the 4IR, a more rigorous engagement is required by all academics across the board—cutting through silos—innovative pedagogies that enable students to become independent learners outside the confines of the lecture halls” (2019, p. 11). This requires major shifts to the process of curriculum development and the way that teaching and learning grapple with the 4IR as a reality within a promising yet challenging and highly unequal social context.

Implications for curriculum

A central purpose of higher education is cultivating “capacity for academic achievement and retention of knowledge among graduates to prepare them for a productive life” (Gleason, 2018, p. 5). Caution is needed to ensure that economic imperatives do not overshadow the social and the environmental. How curricula change in response presents its own challenges. Responsiveness to alterations in knowledge and work, to equipping students with digital skills, and to often narrow industrial needs may lead to greater instrumentalization of higher education, especially when coupled with new approaches to credentialing, such as stackable qualifications and micro-credentialing. As proposals become policy, universities must guard against unintended consequences detracting from their role and purpose. Potential threats are greater vocationalization of curricula, and the development of tiered higher education, with higher order knowledge for elites and vocationally orientated knowledge for the masses.

Fataar identifies the fundamental faultline in the conventional curriculum, with its emphasis on “generic skills and student learning, and its de-emphasis of disciplinary learning, concepts, and content knowledge.” It therefore “falls into the trap of what Maton (2014) has called knowledge blindness” (Fataar, 2020, p. 16) rooted in a failure to provide students with concepts which form the foundations for ongoing acquisition of the more comprehensive knowledge being offered in its stead. In response, multi-disciplinarity is gaining momentum as a resolution to complex challenges precipitated by the 4IR (Marwala, 2020a, pp. 156–157).

The PC4IR Report also considers the curriculum, signaling what might require attention to promote sectoral alignment with South Africa’s national developmental goals. In the short-term urgent action is required to redesign, streamline and align the education system. It also envisions “stackable competencies which are micro-credentialed, industry-aligned and allow people to enter and exit the system at multiple points as part of a lifelong learning process” (PC4IR, 2020, p. 173). The MTT on the 4IR concluded similarly (DHET, 2020b, p. 10). The MTT also examined the changing work landscape, with new technology-related capabilities, skills, and roles in the workplace (DHET, 2020b, p. 24). The MTT proposed a new PSET vision, with core education and training programs aligned with the changing needs of South African society, ensuring high-quality education meeting the urgent demand for digital

skills, providing short courses for the unemployed and underemployed and preparing people to cope with accelerating social and economic change.

Vittore and Govern argue that “[I]nstitutions, students and employers are placing increasing attention on transversal, in addition to subject-specific competences, particularly as the employment market becomes more fluid. As the labor market changes, so do expectations regarding the type of transversal skills needed by graduates” (Vittore and Govern, 2020, p. 4). The capacity and appetite of the PSET sector to realize the 4IR’s potential is inadequate according to the MTT: “Current curricula, programs, and courses are misaligned with labor market demands, while mechanisms to review and update programs and curricula are highly bureaucratic and operate in long, slow cycles” (DHET, 2020b, p. 35). This signals the need for sustained effort to reach an appropriate state of responsiveness and capability, beginning with a review of the perceived misalignment, and devising approaches to narrow the gap.

Imperatives for change flowing from the 4IR

As discussed, in Gastrow’s words, we will probably experience “technological change as a broad and accelerating process” (2018, p. 17). Consequences follow for strategic, policy and systemic change. Du Preez and Sinha remind us how inadequate higher education systems were during previous industrial revolutions, when the need to develop new disciplines was incontrovertible. They indicate that we are in a similar situation and caution us to respond rapidly to 4IR disruptions and to recognize the need for change at all levels (2021, p. 14).

Digital technologies will be prominent in this area. This preoccupies policy makers and thinkers in the developed and developing world. In Europe, Jørgensen contends that dependence on American or Chinese technology is seen as a threat to sovereignty. If it “has too few citizens with digital skills, particularly those that develop new technologies, it would be completely dependent on imported technologies” (Jørgensen, 2019, p. 4). The same concerns apply in developing countries such as South Africa.

Adaptation in the context of wider transformation is echoed in the MTT report. The intersection of technological revolution, societal transformation and radically new business and governance methods makes 4IR disruptive and transformative. South Africa therefore needs to identify, and excel in, targeted technical and scientific disciplines that best align with the country’s development agenda (DHET, 2020b, p. 21). While the emphasis on STEM disciplines is understandable, this should not happen at the expense of the social sciences and humanities.

Investment in 4IR related infrastructure, IT solutions, capacity development and revised business processes must accompany student access to devices and economical internet access. Student skills development must enhance internet use for learning purposes, enriching lives culturally and intellectually and broadening learning beyond the formal curriculum. Gastrow advocates strengthening interaction between educational institutions and employers. This requires:

- Cultivating the capacity of post-school education institutions to engage with employers and understand their changing skills requirements.
- Shortening the curriculum change cycle to respond to changing technologies.
- Cultivating a research agenda that responds nimbly to technological change (Gastrow, 2018, p. 26).

The orthodoxies currently shaping higher education, including program and qualification structures and regulations, quality assurance, how learning credits are accumulated and how learning occurs, need fundamental rethinking. Barber argues that technology and the resulting competition will threaten many aspects of traditional universities, implicitly warning them to become efficient, economical, and appealing (Barber et al., 2014, p. 32). Gleason agrees, noting that “from micro-credentials, Education 3.0, nano-degrees, adaptive learning, micro learning, upskilling, to the idea of preparing for just-in-time education, the message is that we must all keep learning” (Gleason, 2018, p. 6).

As argued earlier, many obstacles need overcoming if 4IR potential is to be realized and the required curriculum reform implemented. The MTT report identifies a critical barrier, “the rigid, slow and arduous processes of quality assurance and program accreditation at institutional and national levels. The policy environment is at best unreceptive, and at worst hostile, to a more agile credentialing system” (DHET, 2020b, p. 34). Relevant agencies such as SAQA and the CHE must be responsive and adaptive. Menon and Castrillon are cutting:

This compliance-heavy system, coupled with the assumption that education means skills which correlate to programme “outcomes” at a certain defined level, means that a thorough reversal or re-imagining of the nature and form of higher education may be needed to prime the system for the 4IR and make a discernible impact on the employability of students in the system. The net result of current regulatory and quality assurance processes is to frustrate institutions trying to respond dexterously to emerging societal needs (2019, pp. 7–8).

Attention to these issues must be tempered by the realization that the interventions are not a panacea for resolving development challenges. Vally and Motala caution that education does not guarantee full employment and that an instrumentalist view is unhelpful especially as it is always based on a raft of unjustified claims about the outcomes of education and skills in capitalist societies; that education and training is not simply a handmaiden for resolving the problems of low economic output; and that

a wide range of exogenous factors and social relations ... circumscribe the potential value of education and training (2014, pp. 31–32).

A constantly shifting 4IR landscape makes it difficult to predict the skills and capabilities demanded from graduates when they start work. Jørgensen suggests that universities should proactively “change the content of their teaching (including lifelong learning), adding new study programs and new curricula, including data literacy. This can include common courses in basic digital literacy for all students” (2019, p. 6).

Finally, the challenges of change and adaptation fall on individuals who also need to adapt and become resilient in a changing, complex and uncertain future. Barber et al. argue that citizens must take responsibility for themselves and their world and constantly learn and relearn, applying their knowledge to present and future problems (2013). Marwala also argues compellingly for individual and institutional responsibility: “The education system at all levels must promote problem-solving skills, computational thinking, multi-disciplinary skills and systems-thinking, while equipping our students to master the social, economic and political worlds” (2020a, p. 27). This increased drive to develop “individual responsibility” should not obscure the importance of enhancing collective social outcomes and using the 4IR to support, and not hinder, improved livelihoods and social justice.

Conclusion

A developing economy and relatively young democracy with enormous socio-economic challenges, South Africa is still on a steep developmental trajectory. How it harnesses technologies, especially those of the 4IR, to embrace new possibilities for development, will determine whether it leaps toward being a developed economy; plods along lethargically; or, at worst, stagnates or regresses. Addressing the numerous policy, regulatory, funding and planning challenges at national level is therefore crucial to the success of its developmental aspirations. The COVID-19 pandemic's negative effects on the economy, livelihoods and society are yet to be fully measured. It is clear though that it would be short-sighted not to give the required attention to the potential and perils of the 4IR for our society and economy.

While HEIs grapple with equipping graduates, employers have the related challenge of remaining competitive while harnessing emerging technologies, which may well have negative consequences for jobseekers (Peters, 2019, p. 1). It is apparent that there is a real possibility that the 4IR could lead to increased unemployment and poverty, deepening inequalities, more state control, and pervasive economic exploitation. These possibilities contrast with the upsurge in development, better economic prospects and achievement of development goals stemming from the 4IR, as detailed in the PC4IR (2020). Thoughtful negotiation of these tensions is necessary to ensure that universities are able to be responsive to a changing world in all its challenges and opportunities.

References

- Atherton, G., 2020. University Access, Student Success and COVID-19 in a Global Context. The Sutton Trust.
- Badat, M.S., 2020. The 4IR Super-Highway: A Dangerously Technocratic Utopia. www.dailymaverick.co.za/opinionista/2020-06-01-the-4ir-super-highway-a-dangerously-technocratic-utopia/.
- Barber, M., Donnelly, K., Rizvi, S., 2014. An avalanche is coming: higher education and the revolution ahead. *Int. J. Pedagog. Innov. New Technol.* 1 (1), 77–80. Available at: <https://doi.org/10.5604/23920092.1134803>.
- Berry, D., 2020. The Ideal of a Digital-First University. *Stunlaw*. Available at: <http://stunlaw.blogspot.com/2020/03/the-idea-of-digital-first-university.html>.
- Chatterjee, A., Czajka, L., Gethin, A., 2022. Wealth inequality in South Africa, 1993–2017. *World Bank Econ. Rev.* 36 (1), 19–36.
- Czerniewicz, L., Mogliacci, R., Walji, S., Cliff, A., Swinnerton, B., Morris, N., 2021. Academics teaching and learning at the nexus: unbundling, marketisation and digitisation in higher education. In: *Teaching in Higher Education: Critical Perspectives*. Routledge. Available at: <https://doi.org/10.1080/13562517.2021.1876019>.
- Deloitte, 2020. The Fourth Industrial Revolution: At the Intersection of Readiness and Responsibility. Deloitte. Available at: <https://view.deloitte.nl/Industry-40-Readiness-report-2020.html>.
- DHET, 2020a. Students' Access to and Use of Learning Materials, Survey Report 2020. Department of Higher Education and Training. Available at: <https://www.siyaphumelela.org.za/documents/600ff2719d9c1.pdf>.
- DHET, 2020b. Report of a Ministerial Task Team on the Implications of the 4IR for the Post School Education and Training System. Department of Higher Education and Training. [https://www.dhet.gov.za/Social%20Inclusion/Report%20of%20the%20MTT%20on%204IR%20Implications%20for%20PSET_Final%20\(1\).pdf](https://www.dhet.gov.za/Social%20Inclusion/Report%20of%20the%20MTT%20on%204IR%20Implications%20for%20PSET_Final%20(1).pdf).
- Du Preez, J., Sinha, S., 2021. A paradigm shift in higher education in the context of the Fourth Industrial Revolution. *IEEE Potentials* 40 (2), 13–18.
- Fataar, A., 2020. The emergence of an education policy dispositif in South Africa: an analysis of educational discourses associated with the fourth industrial revolution. *J. Educ.* (80), 5–24. Available at: <https://doi.org/10.17159/2520-9868/180a01>.
- Gastrow, M., 2018, 10–12 December. Policy options framework for the Fourth Industrial Revolution in South Africa. In: Paper Presented at the SA–EU Strategic Partnership Dialogue Conference on Disruptive Technologies and Public Policy in the Age of the Fourth Industrial Revolution. HSRC, Pretoria.
- Gillwald, A., 2019. SA Must Be Wary of Big Promises Made About 4IR. *Tech Central*. Available at: <https://techcentral.co.za/sa-must-be-wary-of-big-promises-made-about-4ir/93052/>.
- Gleason, N.W. (Ed.), 2018. *Higher Education in the Era of the Fourth Industrial Revolution*. Palgrave Macmillan. Available at: <https://doi.org/10.1007/978-981-13-0194-0>.
- Jensen, T., 2019. Higher Education in the Digital Era: The Current State of Transformation Around the World. International Association of Universities. Available at: https://iau-aiu.net/IMG/pdf/technology_report_2019.pdf.
- Jivan, A., 2020. HR's Place in the Fourth Industrial Revolution. Fact Sheet Number 2020/01. SA Board for People Practices.
- Jørgensen, T., 2019. Digital Skills: Where Universities Matter. Working Paper: Learning and Teaching Paper No. 7. European University Association, Brussels.
- Mamdani, M., 2016. Between the public intellectual and the scholar: decolonization and some post-independence initiatives in African higher education. *Inter Asia Cult. Stud.* 17 (1), 68–83.
- Marwala, T., 2020a. Closing the Gap: The Fourth Industrial Revolution in Africa. Macmillan.
- Marwala, T., 2020b. COVID-19 Has Forced Us Into the Fast Lane of the 4IR Super-Highway. *Daily Maverick*. Available at: <https://www.dailymaverick.co.za/opinionista/2020-05-28-covid-19-has-forced-us-into-the-fast-lane-of-the-4ir-super-highway/>.

- Maton, K., 2014. *Knowledge and Knowers: Towards a Realist Sociology of Education*. Routledge.
- Menon, K., Castrillon, G., 2019. Reimagining curricula for the Fourth Industrial Revolution. *Indep. J. Teach. Learn.* 14 (2), 2–16.
- Menon, K., Motala, S., 2021. Pandemic leadership in higher education: new horizons, risks and complexities. *Educ. Change* 25 (1), 1–19.
- Mishra, S., Panda, S. (Eds.), 2020. *Technology-Enabled Learning: Policy, Pedagogy and Practice*. Commonwealth of Learning.
- Motala, S., Sayed, Y., de Kock, T., 2021. Epistemic decolonisation in reconstituting higher education pedagogy in South Africa: the student perspective. *Teach. High. Educ.* 26 (7–8), 1002–1018. <https://doi.org/10.1080/13562517.2021.1947225>.
- Mouton, J., van Lill, M., Botha, J., Boshoff, N., Valentine, A., Cloete, N., Sheppard, C., 2015. A Study on the Retention, Completion and Progress Rates of South African Postgraduate Students. Department of Science & Technology.
- National Planning Commission, 2011. *National Development Plan 2030: Our Future—Make It Work*. The Presidency. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-workr.pdf.
- Peters, M.A., 2019. In Age of AI, Universities Will Need to Rethink Their Purpose. *University World News*. Available at: <https://www.universityworldnews.com/post.php?story=20190116094437946>.
- Presidential Commission on the Fourth Industrial Revolution, 2020. *Summary Report and Recommendations: Presidential Commission on the Fourth Industrial Revolution*. Government Gazette, Pretoria, pp. 121–347.
- Reinitz, B., 2020. Consider the three Ds when talking about digital transformation. *Educause Review*. Available at: <https://er.educause.edu/blogs/2020/6/consider-the-three-ds-when-talking-about-digital-transformation>.
- Schwab, K., 2015. The fourth industrial revolution: what it means and how to respond. *Foreign Affairs Newsletter*. Available at: <https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution>.
- Statistics South Africa, 2019. *Inequality Trends in South Africa: A Multidimensional Diagnostic of Inequality*. Pretoria, South Africa.
- The World Bank, 2022. *New World Bank Report Assesses Sources of Inequality in Five Countries in Southern Africa*. World Bank Group.
- Vally, S., Motala, E. (Eds.), 2014. *Education, Economy and Society*. Unisa Press.
- Vettori, O., Gover, O., 2020. *Curriculum Design: Thematic Peer Group Report*. Working Paper: Learning and Teaching Paper No. 8. European University Association.
- Williamson, B., Hogan, A., 2021. *Pandemic Privatisation in Higher Education: EdTech & University Reform*. Education International Research.

Toward a framework for the implementation of fully online distance education at a residential higher education institution

Linda van Ryneveld, Graduate Centre, Hatfield Campus, University of Pretoria, Pretoria, Gauteng, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	287
Background	287
Fully online distance education	288
Method	289
Features of the fully online distance education framework	290
Marketing	291
Onboarding process	291
Academic offerings	292
Support structures	293
Graduation	294
Conclusion	294
References	294

Introduction

The line between the various delivery modalities of teaching and learning has blurred in recent times (Lockee, 2021). Defining what constitutes concepts such as traditional face-to-face residential teaching, blended learning, flexible learning, HyFlex learning, distance education and fully online learning has become increasingly complex. The traditional divide between contact and distance education modes in higher education institutions (HEIs) has become less obvious as many institutions opt for a more nuanced approach that allows for the inclusion of various combinations of delivery modalities in a single program. This trend is not limited to residential institutions only, as traditionally distance education institutions also seem to embrace a variety of modes, for example in offering face-to-face tutor sessions, or on-site examinations (UNISA, 2022).

Even before the COVID-19 pandemic forced HEIs to convert their traditional face-to-face offerings into emergency remote teaching (Hodges et al., 2020), many institutions considered offering academic programs in a more flexible delivery mode. In the process, the distinct differences between traditional residential contact teaching and learning on the one side of the spectrum, and fully online distance education on the other, became nebulous.

This article explores the various elements to be considered when a traditionally residential HEI moves into the fully online distance education arena. A conceptual framework is proposed for the introduction of fully online distance education at higher education institutions.

Background

Fully online distance education refers to a modality of teaching and learning where students enroll for a formal academic qualification at a higher education institution and where the program that they enroll for is offered virtually. This means the student never need to travel to a physical campus to attend class or practical sessions, or to write exams. There are several reasons why fully online distance education has been regarded as an attractive option in recent years.

Because of the widening gap in wages and salaries for workers with different levels of post-school qualifications, more and more school leavers in South Africa are applying to enter university programs in an attempt to increase their social mobility (Li, 2013). This need for equity, social justice and democracy to deal with the injustices brought about by the apartheid system, coupled with the need to equip people with high-level skills to meet the employment needs of the public and private sectors, caused an alarming increase in the demand for higher education in South Africa (DHET, 2021). Certain programs are specifically high in demand, albeit because of the social aspirations of parents and caregivers, or because of perceptions that certain types of qualifications are more meaningful than others (Li, 2013).

Whatever their reasons, approximately 1.2 million school leavers enrolled at public institutions of higher learning in 2019 (DHET, 2021). In the South African context, the country's 26 public higher education institutions have struggled for many years to provide access to this ever-increasing number of school leavers who aspire to obtain a qualification at the university level (DHET, 2022a; Regenesys Business School, 2019). However, the high demand for access to universities has not been accommodated for various reasons, including limited physical space in lecture halls and laboratories (Malele, 2011; Saaid et al., 2018). Besides the physical constraint in terms of academic spaces, universities also struggle to keep up with the demands for formal and informal student accommodation in residences or near campus (DHET, 2011). In 2022, the University Student Housing Survey found only a total capacity of 287 507 beds at universities around the country (Cloete, 2022).

In some cases, the demand for access is focused on a particular institution, for instance, where the institution's brand is specifically desirable due to its rankings on a particular world ranking system (Bridgestock, 2022), or simply because of its geographic proximity. Other reasons for high demand may include the cost of the tuition fees for an academic program (NSFAS, 2022) or its attainable entry requirements as measured in an admission point score (APS) (SABursaries, 2022). Still, other prospective students prefer to pursue their studies at a particular institution due to the reputation of a specific academic department, program or even a distinguished professor (Bhardwa, 2017).

Since the demand for the available spaces in popular higher education programs typically outweighs the supply (Li, 2013), universities are notoriously bad at marketing their academic offerings. There seems to be a general belief that prospective students will automatically be drawn to an institutional website where they will find all the relevant information about an academic offering. The belief also seems to be that prospective students will manage to find their way through the maze of an unfriendly institutional student management system, such as PeopleSoft (Oracle PeopleSoft Applications, 2022) and ITS (Govender, 2022), to apply for the program and residence of their choice. Apart from generic marketing and awareness drives (Van der Walde and Bonnema, 2007), such as informational websites, information brochures, open days and school career fairs, very few academic program coordinators ever feel the need to advertise their programs since the demand for spaces normally outweighs those which are available.

Academic program coordinators often put a cap on student numbers to be accommodated in their programs in any given year, as they have to juggle all the variables, including the number of academic staff members available to teach the modules, the administrative and professional staff to support the students throughout their studies, the number of seats available in lecture halls, the capacity of the computer and other laboratories, and the availability of hard copy text books in the library, to name but a few (Saaid et al., 2018; Malele, 2011; Regenesys Business School, 2019). As such, there is seldom a need for additional marketing initiatives, as program coordinators can pick and choose from the available candidates. To ensure that it isn't simply a matter of first come, first serve, academic program coordinators often put specific selection mechanisms in place to ensure that the process results in selecting students who are most likely to succeed. As a result, many deserving students who qualify for access may miss out on being selected for the few coveted spots in a particular program.

Fully online distance education

To accommodate more students, some South African universities have started to consider fully online distance education as an addition to the normal residential offerings (University of Johannesburg, 2022; University of Pretoria, 2022; University of the Free State, 2022; University of the Witwatersrand, 2022). With advances in technology and increased availability of computing devices such as laptops, and Internet access, fully online distance education has now become an option for universities strapped for space (Manichander, 2020; Graham, 2006). It is obvious that not all academic programs lend themselves to a fully online distance delivery modality, but for those that do, this type of offering may just bring much-needed relief (Altbach and De Wit, 2020).

Studying fully online means a student never need access to a physical campus or other learning facilities for teaching, learning, or assessment purposes (Vaughan, 2010). A fully online student only needs a reliable computer and stable internet access. Since these students don't need to attend physical lectures on campus, the HEI can support significantly more students without investing in new buildings, research facilities and larger auditoria (Mukhtar et al., 2020). Fully online distance education effectively eliminates the additional strain on the available facilities and physical environment that extra residential students would have entailed (Rolheiser et al., 2019).

Along with increased access for local students, the fully online distance education modality also allows HEIs to expand their geographical reach nationally to other provinces and internationally to other countries (CHE, 2004; Lederman, 2020). Normally, when international students choose to study abroad at one of the South African HEIs, they need to resign from their full-time jobs in their country of origin, obtain study visas and medical insurance, leave their families behind, and move to a foreign country. There, they have to support themselves financially as full-time students, and pay their tuition fees and associated costs, such as textbooks and software (Khanal and Gaulee, 2019). Fully online studies, however, mean that international students can remain in their country of origin, retain their full-time jobs, and remain in touch with their loved ones, while still furthering their studies at a university of their choice.

As many traditional universities in South Africa have an overt ambition to be recognized as research-intensive (BusinessTech, 2021; University of Pretoria, 2011), institutions of higher learning typically want to be able to accommodate a larger number of students in their postgraduate programs. As postgraduate studies target an older student population (DHET, 2021), these students are often working adults with established careers and family commitments. As such, the fully online mode of delivery is potentially a good fit as students can study asynchronously in the evenings or over weekends, despite working a full 40-h work week (Arghode et al., 2017). Fully online distance education eliminates the need for daily residential-class attendance. It also eliminates the need for students to attend evening, or block classes, which may have entailed long hours on the road, and may have required the student to take a leave of absence from their work (Appanna, 2008).

Because fully online distance education opens up the way for more international and postgraduate students to further their studies, the potential for larger numbers in these two categories may influence the ranking of an HEI. In higher education, rankings are important as prospective students often use them as a decision-making tool when choosing an institution to further their studies at (Bridgestock, 2022; Lubacz, 2022). The visibility of an HE institution's brand increases with a good ranking, and therefore institutions often use their ranking for marketing and publicity purposes. Favourable rankings furthermore increase the likelihood of

partnerships and collaboration with other prestigious institutions abroad, thus further improving an institution's reputation and ability to attract top students and staff alike (Lubacz, 2022).

However, this need must be carefully balanced with the workload of an already burdened staff component (Braat, 2021). Academic staff are notoriously under pressure due to the demands to publish, supervise postgraduate students, teach large groups of undergraduate students, and be involved in their local communities (Karma et al., 2021). As such, any changes in the usual way of doing things are experienced as time-consuming and stressful, and require a careful change management strategy to ensure buy-in and commitment.

With the potential for increased student numbers, fully online distance education offers institutions the ability to increase their revenue in tuition fees and state subsidies. Even though the current Department of Higher Education and Training (DHET) funding formulae dictate that input subsidies for distance education are only half that of residential studies (DHET, 2022b), distance education programs can accommodate a larger student population tends to balance it out. It is foreseen, though, that the latest blurring of the modality lines may require legislative bodies such as DHET and the Council on Higher Education to review and renew their funding and accreditation policies.

As many South African institutions are still reeling financially after the #FeesMustFall drive in 2016, the recent COVID-19 pandemic, global wars, and the effect thereof on the ever-increasing cost of living, HEIs welcome any initiatives that could potentially generate an additional stream of income. As such, the revenue generated by fully online distance education programs is an attractive by-product, even if generating funds is not the driving force behind taking academic programs fully online. Since fully online distance education is particularly suited to fit the lifestyle of postgraduate students, an increase in these student numbers is especially welcome, seeing that postgraduate programs generate more subsidies than those for undergraduate programs (DHET, 2022b).

Finally, it is important to consider the needs and context of the student body. Most HEIs claim to be student-centered in their academic offerings (Weimer, 2002; Wright, 2011). The flexibility that fully online distance education offers means that students can fit their studies into their normal daily routines without having to deal with the normal disruptions if they needed to attend classes on a physical campus (Naidu, 2017). The flexibility to learn whenever and wherever it suits the student must be one of the most attractive attributes of a fully online academic program (Tait and Mills, 2002).

Due to the reasons discussed above, it is clear why so many traditional HEIs choose to include fully online distance education programs in their program qualification mix (PQM) (Nel, 2014). However, with only one distance education institution in the country, it is evident that there is not a lot of expertise available to draw from when a traditionally residential institution wishes to embark on a fully online distance education journey. Furthermore, the systems that work well in a face-to-face contact environment may not be as efficient and effective if transferred to a fully online distance environment (Kumari et al., 2021). As such, the question arises as to whether a framework is available to guide the implementation of fully online distance education programs in a traditionally residential institution.

Method

Following a design thinking process (Brown, 2008; Liedtka, 2018; Owen, 2007; Rowe, 1987), the study was undertaken at a traditional residential university in South Africa that had recently decided to branch out into offering fully online distance education programs as well. As a traditional residential university, neither the academic nor the professional staff (administrative and support) fully grasped the radical differences between normal contact offerings and those in a distance education mode. The research team had to set aside their assumptions about the world (Sharlip, 2019) and actively seek insight into the needs of fully online students, as well as the staff that were going to be involved at each of the touch points in these students' academic journeys.

The design thinking process that was followed included six distinct phases (see Fig. 1). After an extensive review of existing literature and an intense benchmarking exercise (empathize phase) that included exploratory visits to other higher education institutions nationally and internationally, a special project was launched. The project had to investigate, design and pilot a system to allow the university to offer fully online distance programs in a way that was aligned with the values and standards of the normal residential programs. A number of internal stakeholders were identified (define and ideate phases), and they represented several areas in the university, including: student administration, finances, information technology, marketing, academic planning and academic faculties. Multiple interviews, brainstorming sessions, and workshops were held with these stakeholders to ensure that the problems were captured and defined (define phase), after which possible solutions and required elements were identified (ideate phase).

After much investigation and back-and-forth scrutiny, the proposed framework for fully online distance education was drafted (prototype phase). It was then submitted to several test and revision runs (test phase). Once all the obvious issues were ironed out, the framework was formally introduced when the first two fully online programs went live (implementation phase).

The non-linear, iterative, and sometimes chaotic design thinking process that was followed by the team allowed for a better understanding of the problem (Braha and Reich, 2003; The Interaction Design Foundation, 2022). It also allowed for the redefinition of the problem and the provision of creative solutions. The analytic and creative design thinking process that was followed allowed the team to experiment and come up with a design for a prototype framework, and to gather feedback, redesign and modify where it was necessary (Razzouk and Shute, 2012; Rowe, 1987). The methodology suited the complex problem well, as it is a natural fit with qualitative research. The design thinking nature of the methodology, as a human-centered and empirical research method

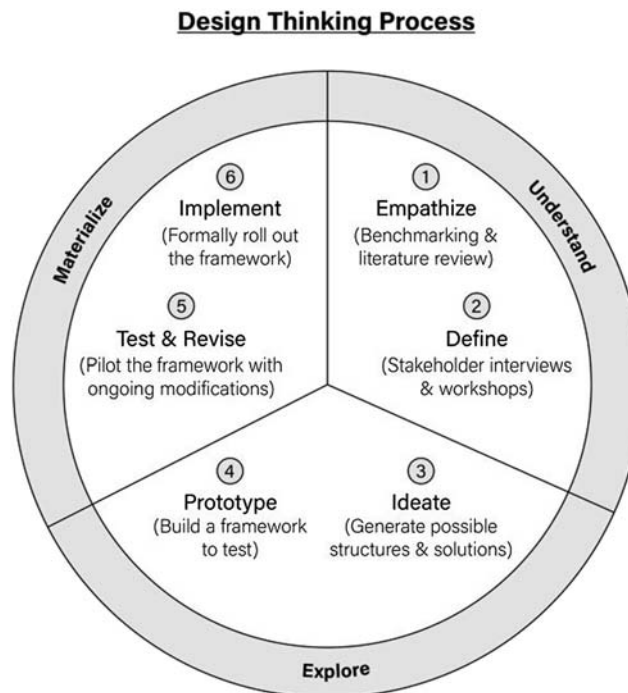


Fig. 1 Design thinking process.

that employed interviews, workshops, brainstorming sessions, and focus groups, proved effective in coming up with a framework for fully online distance education (Brown 2008; Moran, 2021). What started as a cloudy idea of what the framework should look like (Dorner, 1999), with time, crystallized and transformed into a clear image of the framework presented below.

Features of the fully online distance education framework

The proposed framework for fully online distance education at a higher education institution is characterized by several distinct features (see Fig. 2). These features constitute the backbone of the framework, and all work together to provide the student with the best learning experience possible. The ultimate aim is to ensure student success.

Firstly, it is important to recognize that the discussion is not about fully online short courses, but fully-fledged accredited academic programs. In the South African context, it means that these programs are approved by the Department for Higher Education (DHET), accredited by the Council for Higher Education (CHE), and registered by the South African Quality Authority (SAQA).

Secondly, these programs are offered fully online, implying that a student in one of these programs never needs to attend classes or practical sessions in a physical location, such as a campus or a designated meeting space. All assessment activities are also taking place online, and as such, it is never a requirement for a student to show up in person at a physical venue, not even for graduation purposes. Students get access to their study material, all learning activities and assessment opportunities in an online virtual environment, such as a learning management system (LMS). The benefits of this feature are clear, as working adults, those with financial or time constraints, and international students alike, appreciate that traveling and synchronous contact time is no longer a prerequisite.

The third feature entails multiple intakes per year. This means that students are no longer bound to wait till the beginning of an academic year (January in the South African context), but can start their studies at various points throughout the year. In this study, the system was set up to accommodate six intakes per year. The implication is that prospective students can theoretically start their first module within two months after applying.

Even though there are exceptions to the rule, modules are offered linearly, one after the other, and are facilitated over eight weeks. The first seven weeks are meant for academic engagement and assessment, and the final week is a rest period for the students, whereas the lecturing staff finalize the grading of assessments and calculation of final marks. Students also use this week to register and pay for their next module in the cycle. This fourth feature allows students to focus all their study time on a single module, without the distraction of clashing tests and assignment due dates as would have been the case if they had to focus on multiple modules running concurrently.

While it is assumed that working adults will choose the option to study part-time and therefore take only one module at a time, nothing prevents a department from offering a fully online program as a full-time option as well. In such a scenario, students may then be required to register for two or more modules concurrently.

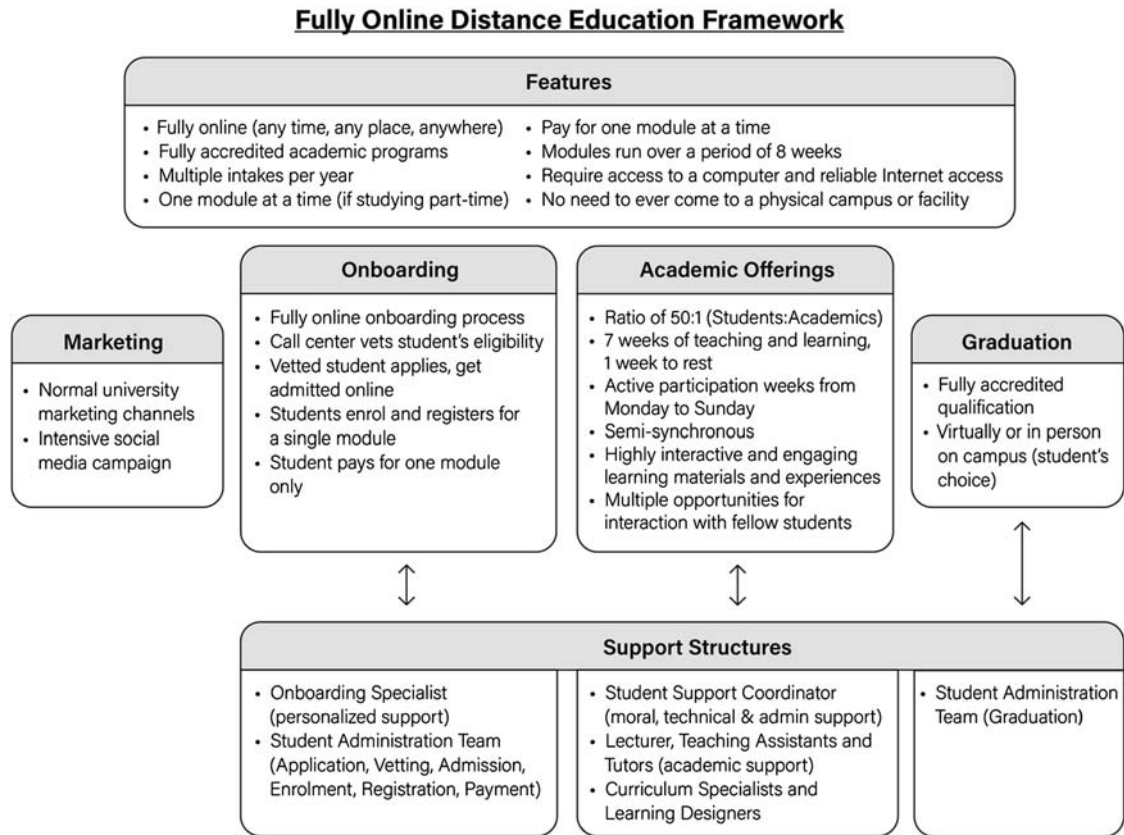


Fig. 2 Fully online distance education framework.

Because bursars often choose to fund students for their first degrees only (NSFAS, 2022), tuition fees can become an insurmountable barrier for prospective students. In the proposed model, students only pay for the module they are currently registering for. As the modules run over eight weeks, they can put money away to pay for their further online studies. This feature also benefits the university, as students who haven't paid for a module are not given access, thus ensuring there is no bad student debt.

The final feature of the proposed fully online distance education framework is that prospective students must have access to a computer and a reliable Internet connection during the time available to spend on their studies. While some learning materials and activities can be accessed on devices like mobile phones and tablets; most academic modules will require students to submit written assignments which would require access to a computer. Some modules may include synchronous virtual sessions, but synchronicity is not encouraged in most cases due to differing time zones and work hours.

The framework identifies four stages: marketing, onboarding, academic offering, and graduation, taking the life cycle of a fully online distance program into account. These stages are unpacked below to highlight the practices experienced at the university where the study was undertaken.

Marketing

Universities all have various marketing initiatives, including websites, brochures, booklets, open days, and talks at schools and career fairs (Kieu et al., 2020; Van der Walldt and Bonnema, 2007). In the proposed framework for fully online distance education, these activities were still utilized to include the new fully online programs. Great care was taken to distinguish the residential offerings from those that did not require students to come to campus. A rather aggressive marketing campaign was also launched via various social media outlets, including LinkedIn and Facebook, as it was assumed that users of these platforms were already comfortable using technology in their everyday lives.

Onboarding process

The university in question had already transitioned to an online application process several years ago. During the pandemic, this early adoption of the available technologies came in handy since most systems were already set up to accommodate online

onboarding. However, since there were significant differences between established processes for the normal residential programs and the new ones for the fully online programs, a major revamp of the information management system was required. With six intakes per year and a payment model where students only pay for one module at a time, the programming and IT development team had to develop several new systems from scratch. In the end, the new application process for the fully online students accommodated all the onboarding processes online, with no need for students to come and queue at the administrative offices on campus. Even the contract that students sign with the university was adjusted to avoid having to submit the physical signed copy. As far as possible, though, the onboarding process for fully online students aligned with those involved in getting the residential students onboard.

The normal application process was, for example, revisited to include a step where an Onboarding Specialist called prospective students as soon as they showed interest in studying online. During these calls, the Onboarding Specialist did an initial vetting of the prospective student to ensure that they understood that the program they were interested in was offered online only and that they, at least at face value, complied with the minimum requirements to gain entry into the program.

After a prospective student passed the initial vetting, they received an email with all the necessary application instructions. Students would then apply and upload all the evidence and proof that they complied with the minimum requirements. Once the Student Administration team checked the documents for legibility and completeness, a list of vetted applicants would be referred to the Academic Program Coordinator, usually the lead lecturer in the program, to select those students who, according to them, complied with the requirements and were to be admitted to the program. Because physical space was no longer an issue, all qualifying students were admitted. The only other concern regarding large class sizes is the lecturer-to-student ratios, which will be discussed later in this article. At this point, the Student Administration Team would then admit the successful students and inform them by email on how to enroll on the program and how to pay and register for the first module.

At the time of this study, the university was not yet offering any fully online distance programs on a full-time basis. As such, students were only allowed to register for one module at a time. The model also worked well financially since students only needed to pay for the module they intended to take next. The costs for each module were not relative to the number of credits, but rather standardized to allow working adults to plan their budgets effectively. The standardization was done by dividing the total cost for the program equally into the number of modules.

Since each module was offered over a fixed period of eight weeks, employed students typically had at least two paychecks before they had to pay for the next module. For many, this eliminated the need to approach a financial institution for a study loan. Some students were fortunate enough to be in a position to pay in full or had their studies subsidized by their employers or other bursars. For those students, the option to pay the total amount for the entire program upfront still existed.

The university required that all students pay in full for the module they intended to take before or on the day the module started. If the tuition fees did not appear in the student's account after five days, their access to the module was blocked. At that time, the Onboarding Specialist would follow up with the student, telephonically or by email, requesting them to upload their proof of payment. Once the proof was provided, the student's access was then reinstated. If the tuition fees still did not reflect in the student's account by day ten, the student was dropped from the module altogether. The student could then save up the necessary fees, and once they paid in full for at least one module, they could join the next module, which always started a maximum of eight weeks later.

Academic offerings

Because lecturing staff at traditional residential universities are not necessarily familiar with the pedagogical underpinnings required for fully online distance teaching and learning, each online module was developed by a team consisting of subject matter experts who brought the discipline knowledge and a learning designer who focused on the andragogical needs of adult learners who study over a distance. In some cases, a curriculum design expert, a graphic designer, an intern, and/or a language editor were also added to the development team. Developing the online version of the modules typically ran over a minimum of six months, indicating the depth and scope of the development work required to produce a solid online learning experience. The end products, though, were highly interactive, well-structured, and engaging online modules that provided multiple opportunities for interaction with fellow students, tutors, teaching assistants, the lecturing staff, and guest presenters.

Each module was structured to run over seven full weeks, Monday to Sunday, followed by a one-week break before the next module commenced. On the Monday of each week, the week's learning outcomes were released, together with the learning activities and assignments that needed to be completed that week. Students could then access their module on a Monday and plan their workload according to their own schedules, resulting in students working during the early hours of the morning, late at night, or at differing hours over the weekends. The due dates for assignments were typically scheduled for midnight on the Sunday evening before the next week's materials and activities were released the next day. As such, the modules could be described as 'semi-synchronous' in nature, as students were all engaging with the learning outcomes of a particular week at the same time (synchronous), but focusing on their studies at different times during that week (asynchronous).

Some lecturers opted to facilitate a synchronous contact session online to kick off the work each week (typically on a Monday evening). During that time, they briefly introduced the topics covered that week and reviewed their expectations regarding the assignments. These sessions were never compulsory, though, seeing that there could be a multitude of reasons why students would not be able to attend, for example, being in another time zone or having to participate in a parent evening at school. For that reason,

the sessions were always recorded and made available online immediately afterward, so that students who missed the session could still work through them whenever they had the time. Most lecturers however pre-recorded their introductory lectures and then provided students with the opportunity to ask questions in an online discussion space that became known as the 'Emergency Room'.

All modules had an Emergency Room built in as a standard feature. The understanding was that this space was to be used when students had questions regarding the content, module design, assessment activities, or any other challenges they could not solve among themselves. The lecturers, tutors and teaching assistants all checked into the emergency room at least twice a day to ensure that students received high-quality feedback as soon as possible, preferably within 12 h of posting their questions.

To encourage the principle of 'Ask three, then me', a 'Coffee Shop' discussion forum was also created in all online modules, where students could converse with their peers in the module. The students were encouraged to discuss any topic of their choice and use the Coffee Shop as a space where they could network, exchange ideas, and freely share their points of view. The members of the lecturing team did not have a presence in this discussion forum. It was intended for students to meet one another online and share their experiences, ask each other questions, and provide peer support. This discussion forum aimed to promote the formation of a supportive online community of practice (Wenger-Trayner and Wenger-Trayner, 2015). It was, furthermore, intended to provide students with a safe place where they could vent their frustrations and celebrate their achievements with the group.

Since one of the reasons the university opted to introduce fully online distance programs was to grow student numbers in their postgraduate programs, it was not preferred to cap the number of students in any of the fully online programs. The main concern, however, was always the ratio of students to lecturers, mainly due to the workload associated with the grading of assignments and learning activities, and the short turn-around time that the weekly structure of the modules and the seven-week time frame allowed for this purpose. In such a tight, condensed schedule, students needed quick and regular formative feedback on their assignments and work, seeing that they needed to learn from their mistakes to make good progress in the following weeks. As such, there was pressure on the lecturing staff to grade assignments and learning activities in the week after they were submitted. For that reason, the university decided to work on a ratio of 50 students to one lecturer. Once the 50 mark was reached, a lecturer had the option to appoint either tutors or teaching assistants to help with various tasks, including grading, responding to student questions, and providing guidance when it became obvious, for example, in the online discussions, that students were on the wrong track.

Seed funding was made available to appoint these tutors and teaching assistants for the first year that a fully online distance program ran to ensure that the quality of the learning experience was not hampered by the lack of time on the lecturers' side. All the learning assignments and experiences were, therefore, consciously designed to ensure that the lecturing team members would each be able to complete all the grading of at least 50 students within a week. After the first year, the tuition income generated by the program was normally sufficient to ensure that the program itself could carry the costs for the teaching assistants and tutors.

Some assessment activities were designed to be marked automatically by the computer. In these types of assessments, great care was taken to ensure that the automated feedback provided by the computer was thorough and formative by nature. The time-consuming nature of developing such detailed feedback contributed to the lengthy module development time as indicated earlier.

Support structures

As distance education can be an isolating and lonely experience (Rush, 2015; Stewart and Lowenthal, 2022), it was important to build in structures for administrative, emotional, academic, and technical support. Prospective students often find the application processes of higher education institutions awkward, inconvenient and unfriendly. When they get stuck on a particular section of the application form, there is usually not an easy just-in-time contact person on standby to assist. Therefore, in an attempt to ensure that all students who showed interest and qualified for the fully online distance programs could apply with ease, the online system was streamlined to be student-centered, as far as the rather rigid parameters of the student management system allowed.

However, there was an obvious need for additional human support in case a prospective student was unsure about the next step, or had technical difficulties uploading a particular document. As such, the Onboarding Specialist was made available, and their contact details were clearly displayed for easy contact. The Onboarding Specialist could typically respond to all the prospective students' questions about the fully online program and the onboarding process. This person also had access to the students' application data online and could see if and where they got bogged down in the process. The Onboarding Specialist were tasked to contact students who, according to the system, started an application but did not complete it, and offer them guidance and support. It was found that this simple touch of 'customer service' went a long way in ensuring that the realization rate of leads (in this case prospective students who had shown an interest in the fully online program), increased.

Once a student was admitted and officially enrolled for the fully online program, they were then handed over to a Student Success Coordinator. This person's role was to support the student throughout their studies regarding any queries they may have. Having set up the online system to flag students who did not log into the learning management systems for a couple of days, the Student Success Coordinator would typically contact students who fall behind to find out what challenges they experienced. If the problems were academic, they would be referred to the lecturing team. However, the Student Success Coordinator often simply provided an empathetic ear to students who struggled with time management, or other personal matters in their day-to-day lives. They also offer various support services, including entry-level technical support, informational support, e.g., how to pay and register for their next module, and, as mentioned, a motivational role of encouragement when students struggled emotionally.

The lecturing team consisted of the main lecturer (the facilitator), typically a seasoned subject-matter expert. As soon as the number of students exceeded 50 (and in increments of 50), additional tutors and/or teaching assistants were appointed to assist with responding to student queries and grading assignments. The lecturing team underwent some training before a module went live for the first time and was made aware of the significant differences between teaching in a normal residential program as opposed to a fully online distance one.

The lecturing team often needed support in the first six months when the fully online module had to be developed and during the first couple of times that a module was offered. In this capacity, the Learning Designers had a vital role to play. Providing pedagogical and andragogical guidance to the subject matter experts, who often had limited teaching experience in an online environment, was essential during the development phase. They often also needed the input and guidance of a curriculum development expert during the development phase. Once the modules went live, lecturers often encountered unforeseen technical difficulties, for which the Learning Designers had to be on standby. From time to time, learning activities didn't work as well as intended, and lecturers needed to fix things in real-time. The Learning Designers then provided moral and technical support to the lecturers, and provided educational advice. They also reminded the lecturing team of the importance of having an online presence and the value of high-quality, just-in-time feedback to student queries.

Graduation

Once a student had successfully completed all the required academic modules, they graduated with a fully recognized and accredited higher education qualification. Successful students were then allowed to either attend the formal graduation ceremony on campus, or receive their degree certificate via the post. Interestingly, despite their choice to study remotely, most students in the first graduating cohort elected to attend the graduation ceremony in person.

Conclusion

As the lines between the different delivery modes for teaching and learning blended with time, this article explored the various elements that need to be considered when a traditional residential higher education institution moves into the fully online distance education space. A conceptual framework was proposed to introduce fully online distance programs at a higher education institution. The framework highlighted the features to be considered and described the details of the processes involved as they played out at the institution where the research was undertaken. Higher Education institutions in South Africa will be looking forward in anticipation to see how legislative bodies such as DHET and the CHE respond to the recent shifts in teaching and learning modalities going forward.

References

- Altbach, P.G., De Wit, H., 2020. Are we at a transformative moment for online learning? *University World News* 2. Available at: <https://ads.universityworldnews.com/bannerclick.php?id=uaeresdec2021aldr>. (Accessed 15 June 2022).
- Appanna, S., 2008. A review of benefits and limitations of online learning in the context of the student, the instructor and the tenured faculty. *Int. J. e Learn.* 7 (1), 5–22.
- Argnode, V., Brieger, E.W., McLean, G.N., 2017. Adult learning theories: implications for online instruction. *Europ. J. Train. Develop.* 41 (7), 593–609. <https://doi.org/10.1108/ejtd-02-2017-0014>.
- Bhardwa, S., 2017. Why do students go to university and how do they choose which one? *Student*. Available at: <https://www.timeshighereducation.com/student/news/why-do-students-go-university-and-how-do-they-choose-which-one>. (Accessed 15 June 2022).
- Braat, M., 2021. High workload among academics is nothing new. *DUB*. Available at: <https://dub.uu.nl/en/depth/high-workload-among-academics-nothing-new>. (Accessed 8 June 2022).
- Braha, D., Reich, Y., 2003. Topological structures for modeling engineering design processes. *Res. Eng. Des.* 14, 185–199. <https://doi.org/10.1007/s00163-003-0035-3>.
- Bridgestock, L., 2022. World University Ranking Methodologies Compared. *Top Universities*. Available at: <https://www.topuniversities.com/university-rankings-articles/world-university-rankings/world-university-ranking-methodologies-compared>. (Accessed 30 May 2022).
- Brown, T., 2008. Design thinking. *Harv. Bus. Rev.* 86 (6), 85–92.
- BusinessTech, 2021. These Are the 11 Best Universities in South Africa. Available at: <https://businesstech.co.za/news/trending/518044/these-are-the-11-best-universities-in-south-africa/>. (Accessed 1 June 2022).
- Council on Higher Education (CHE), 2004. Enhancing the Contribution of Distance Higher Education in South Africa. Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.ch.ac.za/sites/default/files/publications/d000070_Distance_Education_Sept2004.pdf. (Accessed 10 June 2022).
- Cloete, R., 2022. Universities Still Lack Capacity to House Students. Available at: <https://www.skillsportal.co.za/content/universities-still-lack-capacity-house-students>. (Accessed 15 June 2022).
- Department of Higher Education and Training, 2011. Report on the Ministerial Committee for the Review of the Provision of Student Housing at South African Universities. Pretoria. Available at: https://www.researchgate.net/publication/326508920_Importance_level_of_on-campus_Student_Housing_Facility_Spaces_Perception_of_Postgraduate_Students. (Accessed 15 June 2022).
- Department of Higher Education and Training, 2021. Statistics on Post-School Education and Training in South Africa: 2019.
- Department of Higher Education and Training, 2022a. Statistics on Post-School Education and Training in South Africa: 2020.
- Department of Higher Education and Training, 2022b. Ministerial Statement on University Funding: 2022/23 and 2023/24.
- Domer, D., 1999. Approaching design thinking research. *Des. Stud.* 20, 407–415. [https://doi.org/10.1016/S0142-694X\(99\)00023-X](https://doi.org/10.1016/S0142-694X(99)00023-X).

- Govender, P., 2022. Higher education management: why an ITS integrator is the answer - adapt IT education. Adapt IT education. Available at: <https://education.adaptit.tech/blog/higher-education-management-why-an-its-integrator-is-the-answer/#:~:text=ITS Integrator was used to,help management make better decisions.> .
- Graham, C.R., 2006. Blended learning systems. In: Bonk, C.J., Graham (Eds.), *The Handbook of Blended Learning*. Wiley E Sons, Hoboken, NJ, pp. 3–21, 2012.
- Hodges, C., Moore, S., Lockee, B., Trust, T., Bond, M.A., 2020. The Difference between Emergency Remote Teaching and Online Learning. Virginia Tech. <http://hdl.handle.net/10919/104648>.
- Karma, Y., Kezang, S., Pema, C., Sangay, P., Sonam, D., 2021. Well-being and academic workload: perceptions of Science and technology students. *Educ. Res. Rev.* 16 (11), 418–427.
- Khanal, J., Gaullee, U., 2019. Challenges of international students from pre-departure to post-study: a literature review. *J. Int. Stud.* 9 (2), 563–565.
- Kieu, T., Mogaji, E., Mwebesa, C., Sarofim, S., Soetan, T., Vululleh, S., 2020. Marketing higher education in Africa. In: Mogaji, E., Maringe, F., Hinson, R.E. (Eds.), *Strategic Marketing of Higher Education in Africa*. Routledge, Abingdon Oxfordshire, pp. 221–237.
- Kumari, S., Gautam, H., Nityadarshini, N., Das, B.K., Chaudhry, R., 2021. Online classes versus traditional classes? Comparison during COVID-19. *J. Educ. Health Promot.* 10, 457. https://doi.org/10.4103/jehp.jehp_317_21.
- Lederman, D., 2020. How the Shift to Remote Learning Might Affect Students, Instructors and Colleges. Available at: <https://www.insidehighered.com/digital-learning/article/2020/03/25/how-shift-remote-learning-might-affect-students-instructors-and>. (Accessed 25 April 2020).
- Li, H., 2013. *The Rising Cost of Higher Education: A Supply & Demand Analysis*. Honours dissertation at the Undergraduate College, Leonard N. Stern School of Business, New York University.
- Liedtka, J., 2018. Why design thinking works. *Harv. Bus. Rev.* 96 (5), 72–79.
- Lockee, B.B., 2021. Online education in the post-COVID era. *Nat. Electron.* 4, 5–6.
- Lubacz, M., 2022. Seven Reasons Why Your Institution Should Participate in the World University Rankings. Times Higher Education. Available at: <https://www.timeshighereducation.com/world-university-rankings/seven-reasons-why-your-institution-should-participate-world-university>. (Accessed 3 June 2022).
- Malele, I., 2011. Access to Higher Education: Challenges. Higher Education SA Briefing Held on 7 February 2011. Available at: <https://pmg.org.za/committee-meeting/12495/>. (Accessed 17 June 2017).
- Manichander, T., 2020. *Innovative Technology in Teacher Education*. Lulu Publication, Morrisville, NC.
- Moran, K., 2021. Design Thinking: Study Guide. Nielsen Norman Group. Available at: <https://www.nngroup.com/articles/design-thinking-study-guide/>. (Accessed 13 June 2022).
- Mukhtar, K., Javed, K., Arooj, M., Sethi, A., 2020. Advantages, Limitations and Recommendations for online learning during COVID-19 pandemic era. *Pakistan J. Med. Sci.* 36 (COVID19–S4), S27.
- Naidu, S., 2017. How flexible is flexible learning, who is to decide and what are its implications? *Dist. Educ.* 38 (3), 269–272. <https://doi.org/10.1080/01587919.2017.1371831>.
- Nel, W.P., 2014. Universities and decision-making: program and qualification mix - four learning pathways. *J. S. Afr. Inst. Min. Metall* 114 (5), 01.
- National Student Financial Aid Scheme (NSFAS), 2022. The DHET Bursary Scheme. Available at: <https://www.nsfas.org.za/content/bursary-scheme.html>. (Accessed 3 June 2022).
- Oracle, 2022. Oracle PeopleSoft Applications. Available at: <https://www.oracle.com/za/applications/peoplesoft/>. (Accessed 8 June 2022).
- Owen, C., 2007. Design thinking: notes on its nature and use. *Des. Res. Quart.* 2 (1), 16–27.
- Razzouk, R., Shute, V., 2012. What is design thinking and why is it important? *Rev. Educ. Res.* 82 (3), 330–348.
- Regenesys Business School, 2019. Poor Access to Higher Education in South Africa Is a Disaster in Waiting. Accessed at: <https://regenesys.net/reginsights/poor-access-to-higher-education-in-south-africa-is-a-disaster-in-waiting/>. (Accessed 15 June 2022).
- Rolheiser, C., Olmstead, K., Gordon, K., 2019. Goodbye Lecture Halls, Hello Active Learning Spaces: Strategies to Transform Your Classroom into a Lively, Student-Centered Learning Environment. Available at: <https://hbsp.harvard.edu/inspiring-minds/goodbye-lecture-halls-hello-active-learning-spaces>. (Accessed 15 July 2022).
- Rowe, P., 1987. *Design Thinking*. MIT Press, Cambridge, MA.
- Rush, P., 2015. Isolation and connection: the experience of distance education. *Int. J. E-Learn. Dist. Educ.* 30 (2). Available at: <https://www.ijede.ca/index.php/jde/article/view/936>. (Accessed 10 June 2022).
- Saaied, M.N.F., Ayob, M.A., Yunus, A.I.A., Razali, A.W., Maarof, K.A., 2018. The challenges of space management in higher education institution. In: Presented at the IOP Conf. Ser. Mater. Sci. Eng., 429, p. 012102. <https://doi.org/10.1088/1757-899x/429/1/012102> (1).
- Sabursaries, 2022. What Is Your APS Score? Bursaries South Africa. Available at: <https://bursaries-southafrica.co.za/what-is-your-aps-score/>. (Accessed 30 May 2022).
- Sharlip, J., 2019. Applying the Design Thinking Process in Qualitative Research. ORCA. Available at: <https://www.qrca.org/blogpost/1488356/315846/Applying-the-Design-Thinking-Process-in-Qualitative-Research>. (Accessed 4 June 2022).
- Stewart, W.H., Lowenthal, P.R., 2022. Distance education under duress: a case study of exchange students' experience with online learning during the COVID-19 pandemic in the Republic of Korea. *J. Res. Technol. Educ.* 54 (Suppl. 1), S273–S287.
- Tait, A., Mills, R., 2002. The convergence of distance and conventional education: patterns of flexibility for the individual learner. In: *The Convergence of Distance and Conventional Education*. Routledge, Abingdon-on-Thames, Oxfordshire United Kingdom, pp. 13–16.
- The Interaction Design Foundation, 2022. What Is Design Thinking? Available at: <https://www.interaction-design.org/literature/topics/design-thinking#:~:text=Design%20thinking%20is%20a%20non,are%20ill%2Ddefined%20or%20unknown.> (Accessed 15 June 2022).
- University of Free State, 2022. Online Advanced Certificates. University of Free State. Available at: <https://onlinelearning.ufs.ac.za/>. (Accessed 8 June 2022).
- University of Johannesburg, 2022. Study Online Degrees and Programs. UJ Courses. Available at: <https://online.uj.ac.za/>. (Accessed 8 June 2022).
- University of Pretoria, 2011. Strategic Plan. The Vision, Mission and Plan of the University for 2025 [ebook] Available at: https://www.up.ac.za/media/shared/1/ZP_Files/About%20UP/strategic_plan_v11_13feb.zp168988.pdf. (Accessed 3 June 2022).
- University of Pretoria, 2022. UPOnline. Available at: <https://www.up.ac.za/uponline>. (Accessed 2 March 2022).
- University of South Africa, 2022. Unisa Rules for Students. Unisa, City of Tshwane. Available at: https://www.unisa.ac.za/static/corporate_web/Content/Register%20to%20study%20through%20Unisa/Documents/Unisa-rules-for%20students-2017-1.pdf. (Accessed 13 July 2022).
- University of the Witwatersrand, 2022. Online Course, Program & Degree | Best Distance Learning. University of the Witwatersrand. Available at: <https://online.wits.ac.za/>. (Accessed 8 June 2022).
- Van der Walddt, D., Bonnema, J., 2007. Marketing to prospective students at higher education: information needs and source preferences. In: *First European Conference on Higher Education Marketing Academy of Marketing*. Budapest, Hungary. First European Conference on Higher Education Marketing.
- Vaughan, N.D., 2010. Blended learning. In: Cleveland-Innes, M.F., Garrison, D.R. (Eds.), *An Introduction to Distance Education: Understanding Teaching and Learning in a New Era*. Oxfordshire United Kingdom: Taylor & Francis, Abingdon-on-Thames, p. 165.
- Weimer, M., 2002. *Learner-centered Teaching: Five Key Changes to Practice*. John Wiley and Sons, Hoboken, NJ.
- Wenger-Trayner, E., Wenger-Trayner, B., 2015. Introduction to Communities of Practice: A Brief Overview of the Concept and its Uses. Wenger-trayner.com. Available at: <https://wenger-trayner.com/introduction-to-communities-of-practice/>. (Accessed 9 June 2022).
- Wright, G.B., 2011. Student-centered learning in higher education. *Int. J. Teach. Learn. High. Educ.* 23 (1), 92–97.

Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis

Ronghuai Huang^a, Muhammad Yasir Mustafa^a, Ahmed Tlili^a, Rongxia Zhuang^a, TingWen Chang^a, and Daniel Burgos^b, ^a Smart Learning Institute of Beijing Normal University, Beijing, China; and ^b Research Institute for Innovation & Technology in Education (UNIR ITED), Universidad Internacional de La Rioja (UNIR), Logroño, Spain

© 2023 Elsevier Ltd. All rights reserved.

Introduction	296
Education in the post-Covid-19 era	296
Research motivations and objectives	297
Case study collection	297
Case study 1	297
Case study 2	298
Case study 3	298
Case study 4	299
Method	299
Results	299
What types of learning strategies have been implemented in OMO learning?	299
What are the functions of technology in OMO learning?	300
How OMO learning spaces are designed?	301
Discussion	302
Conclusion and implications	303
References	303

Introduction

Education in the post-Covid-19 era

The COVID-19 pandemic has shone a bright light on humanity's flaws and obstacles. It has further emphasized the current inequalities in different fields including education—the most important of which is to think of future education in emergencies and uncertainties, due to that the COVID-19 has disrupted the learning of more than 1.6 billion children (UNESCO, 2021; UNICEF, 2021). The world is now steadily returning from the global pandemic, where students are increasingly returning to campus in Europe, while online courses continue to be offered (Birnbaum, 2020). However due to travel constraints, international students are stranded either in their destination country or their home country, taking online courses to resume their studies (China, 2020). Most colleges in the United States have announced that online courses will continue to be available after 2020, with fewer in-person options. China, on the other hand, has reopened schools and universities (Time, 2021) while India (Express, 2021) has resumed in-person classes for senior students. Schools and colleges are taking numerous precautions concerning these in-person courses. For example, classroom seatings were spread out to keep social distancing, and class enrollment capacity were decreased (Guthrie et al., 2021).

COVID-19s impact on education has created several issues in the post-COVID-19 environment. Now, teachers have to offer the same course two times: face to face and online for those students who cannot attend face-to-face classes (TUoEB, 2021). However, separating students into offline and online classrooms to study independently may limit class interaction and efficiency, which might have a detrimental influence on academic performance (Keržič et al., 2021; Shanta, 2022).

Recently, the Online merge Offline (OMO) Learning (Huang et al., 2021) concept has emerged and opened new doors for more immersive and flexible learning experiences by overcoming many of the educational challenges in the post- COVID-19 era. OMO learning is “a learning model that relies on hybrid infrastructure and open educational practices to merge online and offline (i.e., physical classrooms) learning spaces together in real-time while simultaneously seamlessly teaching students in both the physical classroom and online” (Huang et al., 2021, p. 17).

As Fig. 1 shows, the typical OMO learning environment should have (Huang et al., 2021): (1) internet connection to connect both offline and online students together. For instance, ClassinX, an educational platform, was used to connect both students together; (2) An interactive smart screen, allowing the teacher to simultaneously instruct online and offline students together; (3) cameras to capture, record, and transmit the classroom activities for the online participants; (4) sound system to transmit the voice of instructor and students; and, (5) tablets which are optional and can be used for a variety of instructional activities.

In OMO learning, the students share the learning process by combining two distinct but interconnected environments (physical and online) with the help of innovative technologies and instructional approaches. In addition, for improved learning outcomes in OMO learning, the instructor must design learning activities that encourage interaction, as well as collaboration between online and offline students together (Huang et al., 2021). This ensures that technology and open educational tools not only aid teaching and learning processes, but also promote collaboration and engagement among teachers and students (both online and offline), even if

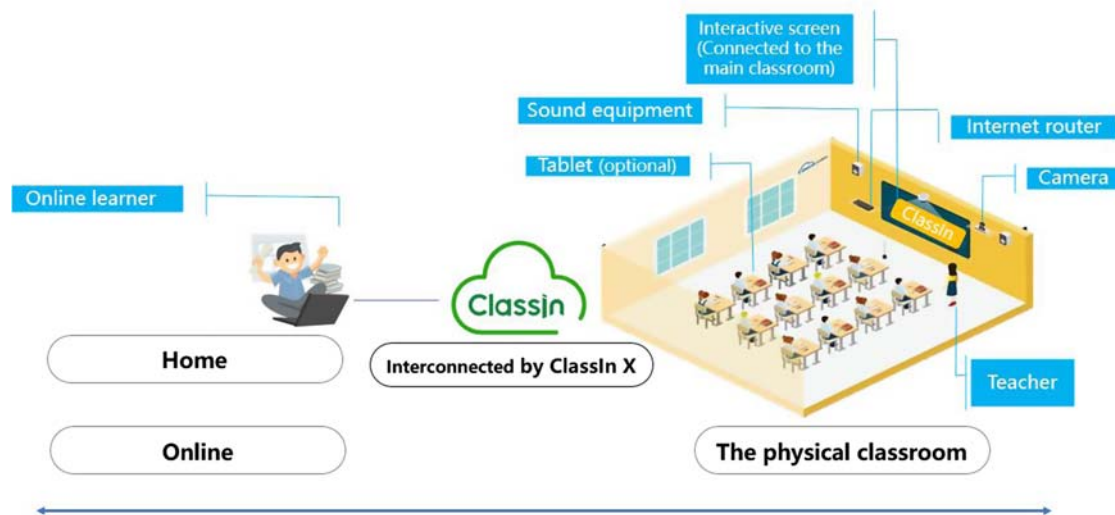


Fig. 1 OMO classroom (Huang et al., 2021).

they are in separate learning spaces (Huang et al., 2021). The objective is to have an immersive and authentic learning opportunities even for those students who are online.

Research motivations and objectives

Traditional methods of online teaching are evolving with the design of more interactive learning spaces and guidelines that facilitate student discussion and collaboration (Wang et al., 2020). Bower et al. (2017) found that evidence-based collaborative technologies will become invisible that students and teachers interacting from different locations will feel that they are in the same room. However, the changes in instructional approaches and the use of new technologies need further planning and coordination (Bower et al., 2015; Wiles and Ball, 2013). A systematic design of the learning environment is therefore needed by identifying relevant areas and success factors that support the design process of these environments. Within such framework, educationalists will be able to create a meaningful OMO learning, based on conceptual, human-centered, and didactic principles. Similarly, Wolff (2020) indicated that careful selection of instructional approaches, facilities, and technology is required for ensuring long-term education. For this, more attention should be paid for designing OMO learning environments to ensure an effective teaching and learning processes. In this context, UNESCO (2020) has emphasized the need of investigating future learning environments to ensure the continuity of education in times of uncertainties. Specifically, this study answers the following research questions:

RQ1. What types of learning strategies have been implemented in OMO Learning?

RQ 2. What are the functions of technology in OMO learning?

RQ3. How OMO learning spaces are designed?

Case study collection

To answer the above research questions, several OMO learning case studies have been collected from different contexts in the post-COVID-19 era. The collection process is detailed in the Method section.

Case study 1

In the first case study shown in Fig. 2A course name "Design and Learning" was taught jointly within 11 universities in China. An OMO (Online merge offline) model was used where 75 (50 online and 25 offline) students participated in this course.

Group discussion between both online and offline students, as well as inquiry-based pedagogy were used by incorporating different technologies, such as cameras, interactive screen, ClassInX (Teaching assistant software), courseware, content editing software, customize blackboard layout, IPad, and teaching assistant software.

The classroom was equipped with an interactive screen to design learning activities for both online and offline participants (in the physical classroom) where they work simultaneously by making groups. Regardless of the learning area, this approach can facilitate the learning interaction between online and offline students at the same time. The online students can also watch the learning activity live and receive oral feedback from their teacher and their offline and online classmates at the same time. The classroom was also equipped with ClassInX, a smart tool, which contains an open educational resource repository that teachers may utilize when



Fig. 2 OMO classroom environment for “Design and Learning” course.

preparing their teaching materials. The students, on the other hand, can work on the teaching materials and update them, attribute an open license for them, and save them on the cloud. This allows these revised teaching materials to be revised by the students themselves, their peers or teachers.

Case study 2

In the second case study, Mr. Ahmed (a pseudonym) taught English communication skills as a subject in one of the universities in Pakistan, as shown in Fig. 3.

There was a total of 60 students (30 online and 30 offline) attended the course. Several technologies like ZOOM, large touch screen, Multi-finger gesture, small blackboard, courseware display, and breakup rooms were incorporated in the teaching and learning processes. For synchronized online discussion and communication between online and offline students, teachers used Zoom. They also used a large touch screen where the Zoom tool was running. This screen was visible to both online and offline students simultaneously. A Group of face-face and online students were asked to discuss and perform group activities by using “Breakout Rooms” in ZOOM. The teachers used Google classroom as a courseware and an Learning Management System to provide online content and learning material to both online and offline students. The teacher used Google cloud and the PADLET (A web Application) to perform collaborative activities among students at both ends (online and offline). A sound system was installed to provide clear audio, hence facilitate the interaction and communication between the teacher and the students (online and offline).

Case study 3

In the third case study, Mr. Chuhan (a pseudonym) an instructor from one of the prestigious institutes in Singapore, taught Mathematics as a subject to the 9th and 10th-grade students by using VOOV meetings. Total 170 students (100 online and 70 offline) attended those subjects.



Fig. 3 OMO classroom environment to teach English communication skills.

The groups of online peering with offline students were formed for collaborative activities by using PADLET a web application. 45 participants attended this session. The teacher made groups of offline and online students for discussion. The other technologies e.g., Digital Cameras were also installed to capture the movement of instructor and offline students activities from online students. The VOOV which is an online meeting app was running on an Interactive touch screen from offline students, GeoGebra (online application) was integrated to teach mathematics-related concepts to both online and offline students at the same time. The learning analytics plugin was enabled to track the collaborative activities of online and offline students.

Case study 4

In the final case study, the instructor in Spain taught social studies subjects to the seventh-grade students using ZOOM and Google meet to interact with online students. There were a total of 45 (25 online and 20 offline) students who attended the class. The other technologies and apps like JAM BOARD, Digital cameras, Sound system, PADLET, EDMODO (A web based Social Learning management system), and video recording apps were integrated into the learning environment.

Students were asked to work on assigned projects by making groups of 3–4 online students peering with offline students. JAM BOARD was shared among online and offline students at the same time. The teacher also used the PADLET as an alternative to JAM BOARD for interactive and collaborative activities and group discussion. Additionally, the teacher created breakout rooms in Google meet for collaborative assignments.

Method

This study used a collective case study technique, which followed conventional case study data gathering and analysis technique (Yin, 2009). Particularly, a call for case studies about educational institutes that have implemented OMO learning has been shared through the professional network of the authors. In this call, each educator had to describe his/her case study aligned with the 5 OMO learning characteristics presented by Huang et al. (2021): (1) Hybrid infrastructure (2) Extended learning space regardless of students (3) Suitable and interactive (open) resources; (4) Open and flexible educational practice and (5) A set of comprehensive competencies of teachers and students to design OMO learning infrastructure). A case study is considered valid if it is: (1) aligned with the OMO definition by Huang et al. (2021); (2) empirical, (3) used for teaching or learning; and (3) well detailed for further analysis. Four case studies were finally selected for analysis and evaluation. Initial shortlisting of the cases was done primarily based on their validity criteria.

Analysis of each case study was conducted iteratively based on each of the aforementioned research questions (Bower et al., 2015). The coded data were then evaluated and cross-checked by the authors to ensure that the study's themes were consistent. Table 1 presents the extracted data from each case study.

Results

What types of learning strategies have been implemented in OMO learning?

The teachers applied different learning strategies to provide an immersive and interactive OMO learning experiences. The following extracted learning strategies were classified according to the Technology Integration Matrix (TIM) along with ISTE Standards (Winkelman, 2020), namely: Active, Authentic, Collaborative, goal-directed and Constructive learning.

1. *Active Learning*: To promote engagements and active learning, the teachers attempted to design student-centered activities by, for instance, asking students to enrich the provided learning materials with more examples from their personal understanding of the discussed learning concept. This made students more engaged to learn, rather than being passive consumers of the information.
2. *Collaborative Learning*: In the OMO learning, collaborative learning is one of the strategies applied by teachers. Collaborative learning is a teaching method that includes students working in groups to improve their learning. For instance, in the first case study, the teacher made breakout rooms where both offline and online student can join these rooms to finish different learning activities.
3. *Authentic Learning*: The Authentic feature highlights the link between new learning and the real world (Winkelman, 2021). In OMO Learning, teachers used storytelling techniques to make learning meaningful. Additionally, the use of different cameras and interaction methods between offline and online students made learning authentic even for those who are online (i.e., they felt that they are with their classmates offline).
4. *Goal-Directed Learning*: The Goal-Directed characteristic describes how technology is used to set goals, plan activities, monitor progress, and evaluate results (Florida Center for Instructional, 2021). This examines how well technology aids, enables or supports meaningful reflection as well as metacognition. In one of the case studies in the OMO environment, Learning Analytics (LA) has been used to monitor and track students learning behavior. Although teachers generally appreciated the benefits of integrating technologies in the OMO learning environment, they frequently find integrating new technology difficult.

Table 1 Summary of cases and participants.

No. of cases	Learning strategy			Spaces	
	Subject area	Instruction/learning activity	Technology	No. of online participants	No. of offline participants
1	Multiple courses	- Group projects/group discussions, peering with online and offline students - Inquiry-based learning	- CLASSIN - WeChat - The camera captures the image of the teacher in class - Teach on iPad - Courseware displays - Editing tools - Teaching tools to empower engaging learning - Customize blackboard layout - Collaborative interactive whiteboard - Teaching assistant Software	50	25
2	English communication skills	- Collaborative activities - Group discussion - Face to face communication	- EDMODO - ZOOM APP - Touch large screen - PADLET - POWERPOINT - Multi-finger gestures - Screen switching - Small blackboard - Courseware displays - BREAKOUT classrooms - Video editing softwares - GOOGLE SHEETS	30	30
3	Mathematics	- Group discussion - Face to face	- VOOV MEETING - Camera - Interactive touch screen - Sound system - PADLET - GEOGEBRA - Learning Analytics - GRAMMARLY APP	100	70
4	Social studies	- Group discussion - Project-based learning	- ZOOM - GOOGLE MEET - JAMBOARD - PADLET - EDMODO - PREZI - Interactive screens - Digital cameras - Breakup rooms - Sound system - Video recordings - GOOGLE FORM - GOOGLE DOC	25	20

What are the functions of technology in OMO learning?

After analyzing the four case studies, it was found that multiple technologies were used and incorporated in the OMO learning environment. For instance, online and offline students used desktop computers, laptops, and smart phones as basic technologies to learn in an OMO environment. Internet was available to ensure connectivity between online and offline students. Functions of technologies were categorized with the lens of the SAMR Model (Aldosemani, 2019). SAMR is designed to analyze how ed-tech is used in a classroom. SAMR is divided into two domains, and each domain has two levels.

The SAMR model was introduced by Puentedura (2014). SAMR stands for Substitution, Augmentation, Modification, and Redefinition. In the first two levels, technology is used to enhance a task and in the second two, technology is used to transform the task. Much like the TIM, it is not assumed that teachers are going to work their way up to the “highest level” and then teach at that level consistently forever. Rather, it is understood that there’s an appropriate place for lessons at every level. With the Technology

Integration Matrix (TIM), we emphasize that entry or adoption levels are often appropriate when introducing a new unit or technology before lessons are conducted at the higher levels. The same concept holds in the SAMR model where “SAMR ladders” are constructed to increase the level of tasks throughout a series of lessons.

1. *Substitution*: In OMO learning, the different functions of technologies were used as a substitute to the traditional ones to facilitate the teaching and learning processes. For instance, in the collected case studies the teachers have used Classin, Zoom, WeChat and Google Meet for online communication as a substitute to the face-to-face interaction in the classroom. The opportunity to share screens became a common substitute for the whiteboard/blackboard. The Microsoft Word documents were used instead of filling out a worksheet with pen/pencil. In the second case, the teacher mentioned that he used electronic or web-based versions of the documents instead of using hard copies. He also used presentation software (like PowerPoint or Prezi) to construct a presentation providing information about the topic. In the third and fourth studies, the digital versions of the books (E-books) were provided to the students. In the first case study, the teacher mentioned that she asked her students to read online articles and then discuss the findings via WeChat and WhatsApp. As mentioned in Table 1, the teacher in the first case study used Google form and in the second case study used LMS embedded quiz for quiz assignments. Interactive screens were used as a substitute to the white/blackboard for online learners and breakout rooms in Classin were used for group assignments and discussion.
2. *Augmentation*: Augmentation is when technology acts as a direct tool substitute that adds functional improvement to the activity. For example, in the first and third case studies, students received help from Grammarly application and Google translator to write the assignments. As mentioned in the third and fourth case studies, teachers provided web links for YouTube video tutorials to learn different concepts. Courseware was also used to provide web-based content. The Jam board was incorporated in the teaching processes to teach mathematics and drawing concepts.
3. *Modification*: Technology is used to design interactive and dynamic tasks that go beyond the limitations of a traditional classroom. For example, in the fourth case study, students used shared documents using Google doc for collaborative work in large groups, allowing for seamless collaboration and knowledge sharing. In the first case study, a collaborative interactive whiteboard was used where students can save and re-edit multiple pages of documents. In the first case study, editing tools in the CLASSES classroom helped the teacher to mark the key points for further improvement. In the second case study, teachers mentioned that they used EDOMO to share a post and get students feedback to improve the quality of documents. Some teachers used a video editing software to create video content by incorporating photos, images, text, and audio, as a group project. Teachers also used Google sheets and directed students to online examples and supplementary learning materials for areas that they might struggle with.
4. *Redefinition*: Technology allows for the new form of previously impossible activities. For example, in the third case study, students were asked to make a documentary video answering a key question about various core ideas.

How OMO learning spaces are designed?

A sound architecture that caters to hybrid teaching and learning based on open instructional standards and tools is one of the main aspects of OMO learning. Therefore, to facilitate online activities such as video lectures and forums, secure networks and digital devices are needed. However, ergonomic classroom furniture is critical for offline students. After analyzing the four case studies we found that the learning spaces were extended by facilitating both offline and online students at the same time. The social distancing and safety measures were kept in view.

The main physical classrooms used by the offline students were equipped with both technology and ergonomic classroom furniture. Availability of classroom furniture, such as movable chairs and tables, was consistent everywhere. However, the technology varies in every classroom. For instance, the offline students came with their own laptop/mobile phones (having required software and application) by following BYOD (Bring your own devices) policy and some of the teachers used smart computer labs equipped with digital interactive screens, desktop computers installed with basic required software and applications, cameras, sound systems. Students in the remote/offline mode were equipped with laptop/mobile phone/desktop computers and internet connectivity.

For clear voice communication, microphones were often used to ensure that those in remote/online learning settings could hear their offline classmates, even though they were dispersed around the classroom for social distancing purposes. Similarly, external cameras were used to assist in-class students in sharing their work with remote/online peers without the need to pass around a laptop/desktop or cell phone. The customize blackboard layout e.g., camera/courseware/blackboard writing were combined and placed freely.

For presentation, sharing, and editing, Microsoft PowerPoint, Prezi, and Microsoft Word were used to present and edit digital content. Cloud drives like Google Drive was used to share digital content. Learning management systems like Blackboard, Moodle, and Edmodo peering with WhatsApp were also used to transfer web links, videos, documents, and recorded information. Some of the teachers used online coursewares for data sharing and presentation. After opening the courseware, the teacher can maximize, minimize or adjust the size of the courseware window. Additionally, some of the teachers used YouTube tutorials to deliver video content and video editing e.g., online Video editing tools.

For group discussion and collaborative activities, the Chat Rooms in ZOOM and GOOGLE MEET and the CLASSIN were used where online and offline students can work collaboratively on the same project/Activity. Teaching tools integrated into the CLASSIN e.g., SELECTOR, TIMER, DICE, SMALL BLACKBOARD were used to boost learning engagement. The Collaborative interactive whiteboard e.g., JAM BOARD and PADLET were incorporated for online group discussion, collaboration, and recording key points.

Discussion

The OMO learning mode was addressed in this study to meet the needs of students and teachers in the post-COVID-19 era internationally. This research analyzed the collected cross-case studies to design an OMO learning infrastructure to facilitate teaching and learning in the future. Specifically, this study discussed OMO learning from three perspectives, namely: learning strategies, functions of technology and space design requirements. Fig. 4 summarizes the obtained results.

In OMO Learning, technology is one of the important components that plays a crucial role to connect and facilitate both online and offline students at the same time. To effectively utilize the resources and constantly tailor material to fit functionality, users must grasp the essential capabilities of different technologies and devices integrated with OMO or e-learning software (McKenney and Reeves, 2016). For instance, the teachers should be aware of what type of technology will best complement their instruction. Functionality refers to the capabilities contained in a device or program that enable it to achieve its job (Gros, 2016). The display capability, audio and video capabilities, multi-language capability, video conferencing, online interaction, and the quality of the operating platform are all examples of functionality. It must be kept in mind that using or integrating technology and instruction in the OMO environment is different as compared to traditional/physical, online, or blended learning environments.

Based on the collected case studies, we found that substitute technologies used in OMO learning environments can facilitate teachers and learners in their learning processes. First at the substitution level, according to Puentedura (2014), teachers are largely performing the same things they were doing in their classrooms before the digital integration. Just the usage of technology may not have a substantial influence on learning outcomes at this stage (Puentedura, 2014). Puentedura (2014) used substitute technology to complete work more efficiently and effectively. Making slide shows, producing papers, conducting research, evaluating data, and communicating with others are all examples of this. Microsoft Word application has been introduced as a substitute for handwritten work. The unique opportunity to share screens between online and offline students has also become a common substitute for the whiteboard/blackboard in the OMO environment. The substitute technologies used in OMO learning have the potential to make learning processes more convenient and efficient in an emergency just like a pandemic.

Some technologies enhanced the learning activity and transformed it. For instance, the use of video conferencing tools combining with cameras to monitor teacher's movement and sound system enabled the entire classroom to be connected with online students for collaboration and interaction. Having learning in an OMO environment was a new experience that was not possible in an online or even blended learning environment. The learning that takes place in these scenarios would not be possible without the ubiquitous integration of technology. Students' interactions with each other via linked networks to cooperate, explore, assess, and reflect on learning are examples of this type of learning.

This study also found that some technologies helped teachers for the functional improvement in the activities during online and offline classes. It helped teachers and students to make learning more authentic. It may also be beneficial to have both synchronous and asynchronous copies of the given lectures. If you're using a videoconferencing tool like Zoom or Skype to host class sessions, it is possible to make a recording for students who can't attend. You can also make educational videos that students can watch at their leisure.

Learning analytics has been used to monitor and track students learning behavior rather than teachers who work hard to keep a record of students' behaviors and their progress. For instance, learning analytics helped a scholar from Northeastern University to realize students' learning habits in online classrooms during the pandemic (James, 2021). This is one of the key distinguishing

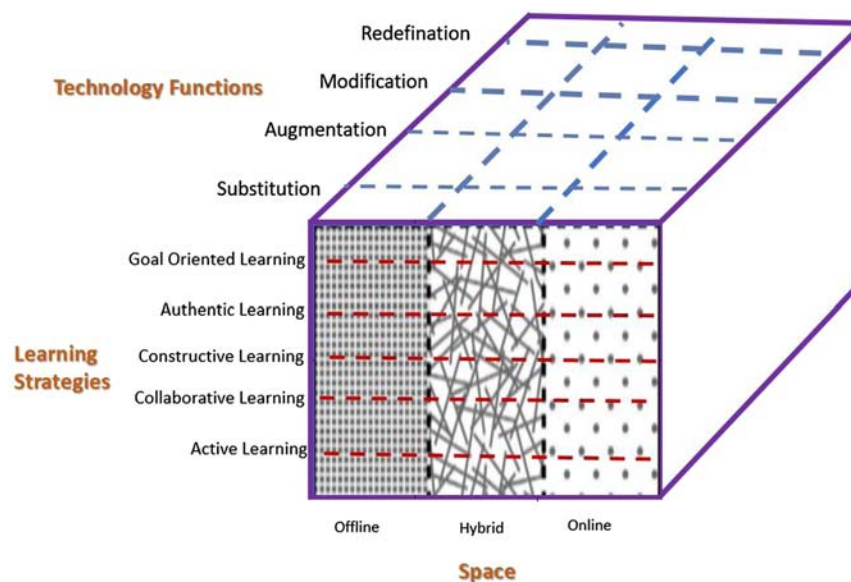


Fig. 4 OMO learning environment model.

features to make OMO teaching and learning processes easier and focus should be paid to incorporating learning analytics (LA) and building LA dashboards, however, its effectiveness needs to be explored further.

In the OMO learning model as shown in the substitution and augmentation phases, teachers will often use automation to complete typical assignments. Teachers modify the assignments and the tasks extend into the confines of the classroom at the next two stages (modification and redefinition). However, not all tasks must be at the redefinition phase. Romrell et al. (2014) looked at tasks in various levels of the SAMR. They discovered that every task, for instance, at the redefinition level was “personalized situated and linked,” but not the tasks at the lower level of the SAMR model (Romrell et al., 2014). They conclude that customized, location-based, and related experiences can rethink and change learning (Romrell et al., 2014).

However, lack of adequate training regarding technology integration in OMO learning was one of the key findings related to the challenges reported by teachers. These findings are also aligned with a pilot study conducted by the authors, where they suggested that OMO learning requires teachers to develop different digital teaching competencies, including operating online platforms, communicating with online and offline participants, and utilizing resources. To realize effective technology integration in OMO learning, school administrators should seek assistance to identify and provide ongoing trainings. Furthermore, using the guidelines offered by ISTE (2021), schools/universities should define professional development services for OMO teaching and learning that better meet their needs. The second major issue reported by teachers was data protection and privacy in the OMO learning environment. The subject of student data privacy has never been more relevant, important, and stress-inducing for school administrators (Huang et al., 2020). While the privacy has been discussed since the 1970s, when schools started gathering electronic data, nothing has changed since the days of analog technology and cassette tapes.

In OMO learning, technologies are likely to have a more central role because they permit active student learning activities in which the teacher serves as a facilitator of the learning process. This study found that OMO learning uses various tools and platforms that allow a collaborative learning environment that makes teaching practices more flexible, engaging, and open. Huang et al. (2021) found that pedagogies in OMO learning are more open and flexible. Additionally, OMO learning also encourages active learning even in times of crises and emergencies. In this context, Huang et al. (2021) point out that open educational practices are an effective and innovative way to enable and maintain education from home during times of crisis and difficult situations.

As previously stated, the OMO modality effectively integrates or enhances learning with the right mix of technology and instruction to allow rich real-time engagement, teamwork, and connectivity between teachers and students (Toor, 2021). The OMO approach involves transparent and flexible teaching and learning activities, allowing students and teachers to engage more diversely. As a result, students will study at their own pace, for example, by reconsidering previously published lessons under an open license. Teachers must undergo a paradigm change from the teacher-centered classroom to the student-centered classroom to make efficient use of computer-based devices (Benlahcene et al., 2020; Ellis and Bliuc, 2019; Gu, 2021; Morel, 2021; Wong, 2021). The increasing adoption of constructivist learning theories, along with intelligent learning innovations provide new possibilities to solve individual discrepancies of the pupil, one of the emphases of contemporary educational instruction.

Conclusion and implications

The four OMO learning cases investigated in this study underlined the significance of planning for active learning, selecting technology-based on representational and communication needs, preparing spaces and teachers for OMO learning. The rich-media content, students and teachers interaction from different locations, open and collaborative technologies will become so invisible and common that students and teachers communicating from different places will feel as if they are in the same room. Meanwhile, current technologies continue to pose substantial obstacles in OMO settings, necessitating teachers to carefully consider their designs and use instructional techniques and workarounds to enhance learning, community development, and the student experience.

In addition, it would also be useful to look at ways to improve the level of engagement among online and offline students. During the teaching process, several online students were bored and walked away from their computers. Furthermore, it is essential to analyze ways to track students involvement and engagement, as their presence in video lectures does not imply sometimes that they are following the learning process.

Despite the solid findings about OMO learning environments in the post-COVID-19 era, it is important to acknowledge that this study contains several limitations. For example, this study analyzed a limited number of case studies (only 4 case studies) and it was mainly based on qualitative analysis. As a result, the findings cannot be generalized, and further quantitative analysis with more case studies should be carried out in the future to make more detailed insights regarding the design of Online-Merge-Offline (OMO) learning environments.

References

- Aldosemani, T., 2019. Inservice teachers' perceptions of a professional development plan based on SAMR model: a case study. *Turkish Online J. Educ. Technol.*
- Benlahcene, A., Lashari, S.A., Lashari, T.A., Shehzad, M.W., Deli, W., 2020. Exploring the perception of students using student-centered learning approach in a Malaysian public university. *Int. J. High. Educ.* <https://doi.org/10.5430/ijhe.v9n1p204>.
- Birnbaum, 2020. Reopened schools in Europe and Asia have largely avoided coronavirus outbreaks. They have lessons for the US. *The Washington Post*. Retrieved April 9, 2021, from: https://www.washingtonpost.com/world/europe/schools-reopening-coronavirus/2020/07/10/865fb3e6-c122-11ea-8908-68a2b9eae9e0_story.html.

- Bower, M., Dalgarno, B., Kennedy, G.E., Lee, M.J.W., Kenney, J., 2015. Design and implementation factors in blended synchronous learning environments: outcomes from a cross-case analysis. *Comput. Educ.* <https://doi.org/10.1016/j.compedu.2015.03.006>.
- Bower, M., Lee, M.J.W., Dalgarno, B., 2017. Collaborative learning across physical and virtual worlds: factors supporting and constraining learners in a blended reality environment. *Br. J. Educ. Technol.* <https://doi.org/10.1111/bjet.12435>.
- Civil Aviation Administration of China, 2020. Preventing Spread of Coronavirus Disease 2019 (COVID-19) Guideline for Airlines. Retrieved April 9, 2021, from: <http://www.caac.gov.cn/en/XWZX/202006/W020200626738640795089.pdf>.
- Ellis, R.A., Blüch, A.M., 2019. Exploring new elements of the student approaches to learning framework: the role of online learning technologies in student learning. *Act. Learn. High. Educ.* <https://doi.org/10.1177/1469787417721384>.
- Express, I., 2021. How and When Will Schools Reopen. Retrieved April 9, 2021, from: <https://indianexpress.com/article/explained/how-andwhen-%0Awill-schools-reopen-what-will-change-for-students-6704619/>.
- Florida Center for Instructional, 2021. The Technology Integration Matrix: A Project of the Florida Center for Instructional Technology. Retrieved May 16, 2021, from: <https://fcit.usf.edu/matrix/matrix/>.
- Gros, B., 2016. The design of smart educational environments. *Smart Learn. Environ.* <https://doi.org/10.1186/s40561-016-0039-x>.
- Gu, X., 2021. Response to "A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it": a design perspective. *Educ. Technol. Res. Dev.* <https://doi.org/10.1007/s11423-020-09896-w>.
- Guthrie, B.L., Tordoff, D.M., Meisner, J., Tolentino, L., Jiang, W., Fuller, S., Green, D., Loudon, D., Ross, J.M., 2021. Summary of School Re-opening Models and Implementation Approaches During the COVID 19 Pandemic. Retrieved May 23, 2021, from: <https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/20200706-SchoolsSummary.pdf>.
- Huang, R.H., Liu, D.J., Zhu, L.X., Chen, H.Y., Yang, J.F., Tili, A., Fang, H.G., Wang, S.F., 2020. Personal Data and Privacy Protection in Online Learning: Guidance for Students, Teachers and Parents, vol. 109. Retrieved from: <http://creativecommons.org/licenses/by-sa/3.0/igo/>.
- Huang, R., Tili, A., Wang, H., Shi, Y., Bonk, C.J., Yang, J., Burgos, D., 2021. Emergence of the online-merge-offline (OMO) learning wave in the post-COVID-19 era: a pilot study. *Sustainability*. <https://doi.org/10.3390/su13063512>.
- ISTE, 2021. ISTE Standards for Educators. Retrieved May 23, 2021, from: <https://www.iste.org/standards-for-educators>.
- Keržič, D., Alex, J.K., Alvarado, R.P.B., da Silva Bezerra, D., Cheraghi, M., Dobrowolska, B., et al., 2021. Academic student satisfaction and perceived performance in the e-learning environment during the COVID-19 pandemic: evidence across ten countries. *PLoS One*. <https://doi.org/10.1371/journal.pone.0258807>.
- McKenney, S., Reeves, T.C., 2016. Educational design and construction: processes and technologies. In: *Lecture Notes in Educational Technology*. https://doi.org/10.1007/978-3-662-47724-3_8.
- Morel, G.M., 2021. Student-centered learning: context needed. *Educ. Technol. Res. Dev.* <https://doi.org/10.1007/s11423-021-09951-0>.
- Puentedura, R., 2014. Find Out How You Can Use Technology to Engage Students in Rich Learning Experiences. Retrieved May 30, 2021, from: <https://www.common Sense.org/education/articles/samr-and-blooms-taxonomy-assembling-the-puzzle/>.
- Romrell, D., Kidder, L.C., Wood, E., 2014. The SAMR model as a framework for evaluating mLearning. *J. Asynchronous Learn. Netw.* <https://doi.org/10.24059/olj.v18i2.435>.
- Shanta, S.A., 2022. Student performance evaluation between offline and online pedagogy: a critical analysis with possible suggestions. *J. Inf. Manag.* <https://doi.org/10.31580/jmi.v8i4.2080>.
- Time, T.N.Y., 2021. How China Brought Nearly 200 Million Students Back to School. Retrieved December 21, 2021, from: <https://www.nytimes.com/2020/09/12/world/asia/china-schools-reopening.html>.
- Toor, R., 2021. Turns Out You Can Build Community in a Zoom Classroom, the Chronicle of Higher Education. Retrieved May 17, 2021, from: <https://www.chronicle.com/article/Turns-Out-You-Can-Build/249038>.
- TUoEB, 2021. Information About Semester 2. Retrieved May 29, 2021, from: <https://www.ed.ac.uk/news/covid-19/currentstudents/%0Asemester-2>.
- UNESCO, 2021. UNESCO Launches Global Consultation on the Futures of Education Initiative. Retrieved April 9, 2021, from: <https://en.unesco.org/futuresofeducation/initiative>.
- UNICEF, 2021. Keeping the World's Children Learning Through COVID-19. Retrieved April 9, 2021, from: <https://www.unicef.org/coronavirus/keeping-worlds-children-learning-through-covid-19>.
- Wang, C., Fang, T., Gu, Y., 2020. Learning performance and behavioral patterns of online collaborative learning: impact of cognitive load and affordances of different multimedia. *Comput. Educ.* <https://doi.org/10.1016/j.compedu.2019.103683>.
- Winkelman, R., 2020. ISTE Standards and the Technology Integration Matrix. Retrieved May 30, 2021, from: <https://fcit.usf.edu/matrix/iste-standards/>.
- Winkelman, R., 2021. Authentic Learning: Mirroring the Real World. Retrieved May 16, 2021, from: <https://fcit.usf.edu/matrix/authentic-learning-mirroring-the-real-world/>.
- Wong, K.M., 2021. "A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it" by Lee and Hannafin (2016): an international perspective. *Educ. Technol. Res. Dev.* <https://doi.org/10.1007/s11423-020-09842-w>.
- Yin, R.K., 2009. Case Study Research: Design and Methods, Applied Social Research Methods Series. Book.

Assessing student learning outcomes

Matthew R. Wawrzynski, Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Assessment and student learning	305
Shared responsibility for student learning	305
Creating student learning outcomes	306
Benefits of assessing student learning outcomes	307
Steps in assessing student learning outcomes	308
Examples	309
National Peer Educator Study	309
Nelson Mandela University Student Learning Outcomes Initiative	310
Concluding thoughts	310
References	311

The purpose of this entry is to explore the topic of assessing student learning outcomes. I first start by offering a definition of what I mean by assessment and student learning outcomes. Then, I discuss who has responsibility for student learning, key characteristics of student learning outcome statements, the benefits of using student learning outcomes, and steps for assessing student learning outcomes. I also include two examples of student learning outcome assessment efforts.

Assessment and student learning

Assessment has become a ubiquitous concept in higher education but often gets caught in a tangled web of terms, particularly with definitions of research and evaluation. Although one can easily be concerned with the nuances of each of these terms and definitions, for clarity and for the purpose of this entry, I define assessment as the process of collecting, analyzing, and interpreting information to improve effective change for students.

Similarly, the definition of student learning outcomes is also likely to vary among the authors who write about them. However, for the purposes of this entry, I define them as outcomes to explain, demonstrate, or understand what students are expected to learn and what they are expected to be able to do, as a result of attending and participating in a class, experience, or opportunity at a higher education institution (HEIs). If student learning is truly a desirable outcome for HEIs, then it must be assessed as such and not seen as a by-product for students enrolled in the systems of postsecondary or tertiary education. Although student learning outcomes should be developed and assessed for the multifaceted experiences and opportunities that exist at universities, this entry specifically focuses on those experiences that occur in the cocurriculum (i.e., the part of the college experience where students are involved in out-of-class activities, often referred to as clubs, organizations, and societies). Taken together, assessing student learning outcomes then is the intentional process of collecting, analyzing, and interpreting what students are expected to learn and what they are expected to be able to do to enhance their learning as a result of their involvement in cocurricular experiences.

Shared responsibility for student learning

Increased demands by the public, legislators, parents, and students, among others, for greater accountability and continuous improvement in HEIs have permeated all sectors of higher education regardless of stature, location, or global ranking. Central to these demands is the rising costs of attending higher education; escalating student loan debt; and as more graduates are unemployed or underemployed, concerns about the ability to translate what students learn to what employers need and want from employees. Increased attention on institutional assessment efforts to address these issues and demonstrate return on investment are met with mixed results. More specifically, assessment is inextricably tied to the accreditation or accountability process, which creates issues for establishing a culture of institutional assessment because of the continued start and stop nature of these processes. For example, when student learning assessment efforts focus primarily on the accreditation process, once it is finished, the desire to assess is likely to subside until the next accreditation cycle when the assessment process starts all over again. Similarly, when accountability is the purpose for assessment efforts, student learning is not the focus, but rather the focus becomes the return on investment or what graduates earn. Unfortunately, when the philosophy for conducting assessment of student learning is primarily focused on accreditation and accountability processes, then the learning students derive from their curricular and co-curricular experiences becomes

ancillary. Although many Institutional leaders engage in assessment for accreditation and accountability, others pursue assessment efforts better to understand the growth and learning of students. I argue assessing student learning outcomes does not need to be an either/or dilemma, but rather a both/and opportunity, and assessing student learning needs to occur in both the academic and cocurriculum.

One could reasonably argue assessing student learning outcome efforts has gone beyond accreditation and accountability measures to include how students spend their time and what they gain from attending college (Kuh, 1993). Yet, much of this focus is on more cognitive dimensions of learning from the college experience that is tied to the academic curriculum. Indeed, over the last 20 years, assessing student learning in the academic curriculum (e.g., assessing what students should learn and be able to do as a result of attending a class or completing coursework) has received greater attention in Western and International contexts using adapted versions of the National Survey of Student Engagement (NSSE, 2021), that is, the South African Survey of Student Engagement (SASSE) (Strydom and Mentz, 2010); Australian Survey of Student Engagement (AUSSE) (Coates, 2010); and the Chinese College Student Survey (CCSS) (Shi et al., 2014). Indeed, measuring academic learning outcomes have become increasingly important in international higher education (Coates, 2016) because of the quest for greater recognition often associated with climbing the global ranking ladder.

Despite these assessment initiatives, their primary focus remains on the academic curriculum, without much attention given to other aspects of the college experience, particularly in the cocurriculum. Moreover, cocurricular assessment activities, unfortunately, are often not well integrated with other campus assessment activities—in part, because academic affairs still conduct much of the work on student learning outcomes. This, then, leads other parts of the college experience unassessed. When only one part of the college experience is explored it fails to take into account the overall contributions of the many facets of involvement and limits the extent to which strategies for improving student learning and success can be implemented.

Cocurricular assessment efforts are often referred to as a missing element in the day-to-day work of those who work with cocurricular programs and services. However, when professionals whose primary responsibility for the growth and development of college student experiences outside of the classroom engage in assessment strategies, their assessment efforts are spent more on student satisfaction. Putting more assessment effort into student satisfaction usually equates to far less time assessing what students are learning from the many programs, workshops, seminars, and events offered under their purview. Moreover, since cocurricular experiences help us to understand the contributions they make to student learning, additional efforts are needed that focus on the cocurriculum as this voice is too often absent from campus-wide discussions on assessment of student learning.

Creating student learning outcomes

Creating student learning outcomes is a systematic process accomplished through essential steps requiring educators to identify and articulate specific learning goals students achieve through participation in learning experiences and opportunities. To be most effective, student learning outcomes must be clearly communicated such that they identify a target audience (i.e., who should be assessed), the type of learning behavior to result from the experience (e.g., what should be learned), the learning statement that identifies the expected condition or action (e.g., how should the learning occur), and the degree to when the learning might occur (e.g., when will the learning occur). Since the focus of this entry is student learning outcomes for students in postsecondary or tertiary education, students then are the target audience or the subject of the learning outcome.

In the next step, educators must be careful to identify specified learning behaviors or action(s) that are observable for the students involved in the experiences and opportunities and not on what the program intends to offer. Making student learning outcomes observable is the foundation in this step and is accomplished by using operational and action verbs as a core element in the learning outcome. Moreover, a key component of clearly communicating the student learning outcome then, is choosing the correct verb, as the verb is the fulcrum of the student learning outcome, meaning that it balances the audience and the learning statement.

The use of action verbs derives from the hierarchical taxonomy of complex processes advanced by educational psychologist Benjamin Bloom and his colleagues. Bloom's taxonomy has 6 meta categories of educational objectives: knowledge, comprehension, application, synthesis, and evaluation (Bloom et al., 1956). Each meta category in turn has a number of action verbs to represent elements of cognitive development. Higher cognitive levels of understanding incorporate and build upon elements of lower levels of cognitive learning. Bloom's taxonomy (Bloom et al., 1956) and revised taxonomy (Krathwohl, 2002) are valuable resources as they offer useful language to be included in student learning outcome statements. In fact, the *Council for the Advancement of Standards* (CAS) (Council for the Advancement of Standards in Higher Education, 2019), *Learning Reconsidered* (Keeling, 2004), and *Learning Reconsidered 2* (Keeling, 2006) are but a few examples of how Bloom's taxonomy is integrated into frameworks that guide student learning outcome development. Moreover, since Bloom's taxonomy is one of the oldest and advanced, it has received the greatest attention and support in the literature as being illustrative and useful in providing an established sequential model of learning actions. Incorporated into each meta category of the model is a list of action verbs for writing learning outcomes such as cite, compile, create, design, identify, plan, select, utilize, demonstrate, prepare, use, compute, discuss, explain, assess, compare, critique, outline, and evaluate that characterize the ability to demonstrate cognitive thought processes. For example, as a way of

students demonstrating development in the knowledge category of the taxonomy, the lowest level of complexity, students can cite key facts and concepts, but in the highest level, they need to engage in interpretation by assessing the quality of information and provide evidence to support their claims.

Another valuable resource to utilize when identifying observable student learning outcomes is a departure from the traditional action verbs found in Bloom's meta-categories. This more nontraditional perspective is advanced by Adelman (2015), who advocates for the inclusion of Natural Language Processing (NLP) (Levin, 1993) to guide 20 categories of operational verbs. Although the methodologies and categories of NLP are beyond the scope of this entry, Adelman's argument for the deviation from the traditional action verb language is important to consider because of the value of language and syntax in creating student learning outcomes, particularly in international contexts. Adelman contends that many of the action verbs found in Bloom's meta-categories lack measurability, whereas, an operational verb references "student actions that are directly observed in external contexts and subject to judgment" (p. 7). Moreover, Adelman argues for the exclusion of the verb "able" and the noun "ability" in developing student learning outcomes and groups the categories of operational verbs by their governing functions (e.g., modes of communication, ways students process data, ways students certify gathered information). Among the 20 categories are verbs like access, collect, retrieve, obtain, frame, classify, describe, define, estimate, assemble, illustrate, identify, specify, organize, compare, gather, examine, construct, argue, defend, adapt, and negotiate.

Although readers will quickly identify some overlap between action and operational verbs, I encourage readers to review the work advanced by Bloom and Adelman and then consider both the advantages and disadvantages to using either body of work. The ability to actually observe and measure student's actions must be taken into account to assess student learning outcomes. The emphasis I offer in discussions about Bloom and Adelman is that equally important to using verbs that describe exactly what the learners will be able to do is avoiding verbs that lack clarity in the context of student learning outcomes (e.g., learn, appreciate, understand) because they lack observability, which then affects their ability to be measured. Regardless of the body of work one chooses, this step in the process, allows educators to identify strategies to assess and measure the degree to which student's actions are observable by being able to describe, demonstrate, perform, analyze, or generate the desired knowledge or skill as a result of the experience and opportunity.

When student learning outcome statements incorporate these characteristics, they can then be measured and assessed to determine growth and development. For example, a sample learning outcome statement could be, "As a result of students attending this career workshop, they will identify key elements of an interview." In this example, students are the intended audience, and the observable behavior and outcome is identifying key elements of an interview; the how is that they will learn how to identify, and when is attending the career workshop.

Furthermore, when these steps are used as a guiding framework, then assessment becomes an integral component of creating the student learning outcomes process. Furthermore, this process is most effective when educators take time to reflect on the essence of the experience and opportunity they offer to students and then describe what students will achieve as a result of their involvement. Student learning outcomes should also be connected to a larger set of desired outcomes or competencies such as those in the mission of a department or institution or in professional fields. Although assessment of student learning outcomes can be a formalized process overseen by an institutional assessment office, which is often the case with accreditation, it does not need to be so. In fact, I argue assessment must be integrated in all components of the college experience where there is potential for student learning to occur. I encourage institutions to engage in a process of mapping their learning environments where educators identify curricular and cocurricular experiences and opportunities that are likely to lead to desired student learning outcomes. Mapping the variety of these occurrences then provides opportunities for anyone with responsibility for student learning to engage in creating specific outcomes for those experiences.

Benefits of assessing student learning outcomes

When there is a thoughtful and intentional institution-wide approach taken to creating and implementing student learning outcomes, benefits to assessing them can be realized: six of which are noted here. First, when assessment of student learning extends to all aspects of the student experience, there is greater potential to identify learning areas where students gain the most, and least, from their college experiences. Because students are far less likely to create a bifurcated system for their college experience, even though such separation is often created in HEIs by those who see their primary responsibility as either in the academic curriculum or the cocurriculum, being able to identify learning outcome areas can lead to campus-wide discussions and efforts to develop strategies to improve deficient areas of student growth and development. Similarly, knowing where learning is taking place also provides the opportunity for greater synergy and connections to help students have a seamless experience between what they are learning and doing in the classroom to what they are learning and doing in cocurricular experiences.

Second, assessing students on the intended learning outcomes for the various types of experiences they participate in conveys a message on the role and importance that learning plays in those experiences. Then, students can use this assessment information to better understand their role and responsibility in identifying the learning they desire to gain from their experiences. Students should be encouraged to develop an action plan that takes into account knowledge they have gained and knowledge they still hope to gain while in college. So, when students are more aware of what they should learn and do learn, they can be active participants in their own goal setting by seeking out those experiences that are likely to lead to greater success. Similarly, sharing student

learning outcome assessment results with students and encouraging them to reflect on their learning helps them to share responsibility for their learning. When students know what they gain from their college experiences, they may be more likely to invest the time and energy in other activities where they identify what they will learn.

Third, assessing student learning outcomes affords greater opportunity to showcase how student learning outcomes are being fulfilled by a department, division, or institution's mission. Knowing if students gain intended learning outcomes provides opportunities for departments or offices to showcase how the programs and events they offer align with the institution's mission or that they need to reconsider their offerings to create better alignment. When HEI budgets become considerably tighter and constrained, being able to show how student learning from various programs contributes to the overall institutional mission is essential particularly when there are implications for potential resource allocation.

Fourth, assessing student learning helps those designing curricula and programs the opportunity to determine whether the intended outcomes from the experiences are being reported by the students who participate in them. Additionally, assessing helps determine if the intended learning outcomes of the programs and services they implement are having the types of intended influence for students.

Fifth, assessing student learning outcomes and analyzing results allows for going beyond hunches and anecdotal evidence. Assessment is an effective way to become experts on student learning. When those responsible for student learning believe they share the responsibility of knowing and supporting students personally and holistically to persist and be successful in college (Sedlacek, 2004), assessment efforts will more likely be integrated by multiple people and not only an institution's accreditation office.

Finally, if student learning is a shared responsibility by campus stakeholders, then assessing student learning outcomes provides the opportunity to see how experiences and learning opportunities are connected so students can be encouraged to reflect on their experiences. When students are encouraged to reflect and make connections and meaning, their learning becomes transformative and applied experiences (Wawrzynski and Lemon, 2019), which allows them to also see how their gained knowledge might be applied in other settings and experiences.

Steps in assessing student learning outcomes

The intrinsic value of engaging in assessment to understand what and if students are learning can no longer be taken for granted in systems of higher education. Indeed, many who are responsible for student learning contend they lack the time and the knowledge base to engage and conduct assessment efforts. However, when the steps presented earlier in this entry are used as a guiding framework to create student learning outcomes, assessment challenges and barriers are more likely to be overcome. I offer the following eight steps as a process for assessing student learning outcomes. A key assumption when developing these steps is that the purpose of engaging in assessing student learning outcomes is an intentional effort to collect, analyze, and interpret what students are expected to learn and what they are able to do as a result of their involvement in an experience. Remember, the specific focus I make is on involvement in cocurricular experiences (Table 1).

Table 1 Steps to assessing student learning outcomes.

1. **Determine what experience you want to assess.**
 - Are you assessing student learning outcomes from a specific experience?
2. **Determine the purpose for the assessment.**
 - What goals are you attempting to achieve?
 - What will you do with the data once it is collected?
3. **Determine what students should know and be able to do as a result of the experience.**
 - What do you want students to know as a result of participation in an experience, class, workshop, etc?
4. **Develop observable learning outcomes using action and operational verbs.**
 - Develop measurable student learning outcomes.
5. **Determine the appropriate assessment method.**
 - Are qualitative or quantitative strategies most appropriate for what you want to know?
 - How will you collect the data? Will you use focus groups, individual interviews, surveys (online or paper and pencil)?
 - Are there pre-existing instruments to use to accomplish your assessment goals?
 - Who should collect the data? Do those collecting data have experience and knowledge to do so?
6. **Determine how you will analyze and interpret the data.**
 - Is there someone who has experience and knowledge analyzing data given your chosen collection method?
 - How do you reduce bias in your interpretation? Is there a colleague who can provide feedback about the interpretation process?
7. **Generate implications for policy and practice.**
 - What did you learn from the study?
 - What changes can you suggest from what you learned?
8. **Develop a plan to report results.**
 - Which stakeholders do you need to share information with?
 - Who is likely to benefit from your results and what format is best to disseminate findings and implications?
 - Be sure to share results with participants who provided data for your assessment.

Note. Each step represents important considerations when preparing to engage in assessment.

Examples

I offer a brief overview of two assessment projects to demonstrate ways to engage in assessing student learning outcomes. Both projects started with a review of the larger extant literature associated with the general thematic nature of the project to gain a better understanding of what is known about the topic. From that initial review of the literature, both a qualitative and quantitative study were developed as part of each study's overall attempt to gather data and assess student learning outcomes. Although the first study is a larger multi-institutional study, the process is very similar to conducting assessment of student learning at a single institution.

National Peer Educator Study

The National Peer Educator Study is an ongoing study first started in 2004 with support from the BACCHUS Peer Education Network; now, the Health, Wellness, and Safety Initiatives of the National Association of Student Personnel Administrators; using both qualitative and quantitative assessment strategies (Wawrzynski et al., 2011). The purpose of this project is to examine learning outcomes of students who participated in specific types of leadership roles (i.e., peer educators).

University administrators value the pivotal role peers play in a student's growth and development while in college, and, as a result, utilize students selected to serve as peer educators in a number of cocurricular experiences. Students in these roles are called upon to mentor and facilitate dialog that encourage peers to consider, talk honestly about, and develop healthy lifestyle choices. Because of their ability to communicate with their peers about sensitive issues, the result has been an exponential growth of peer education programs. The strength of the peer influence has several benefits on the outcomes of student learning, attitudes, and behaviors, which explains why peer education programs have experienced growth across many higher education sectors. However, despite substantial use of students as peer educators to positively influence and enhance the undergraduate student growth and development of their peers, little knowledge existed, beyond anecdotal evidence, about what benefits or learning outcomes peer educators derive, if any at all, from their experiences in these roles. Recognizing this gap in the literature, the National Peer Educator Study was launched.

The first stage of the National Peer Educator Study consisted of a review of the literature to guide the development of a qualitative assessment protocol developed for use in focus group interviews. Over 100 undergraduate peer educators attending a national peer education conference participated in one of a dozen focus groups to share their personal learning and outcomes attributable to their peer education involvement (Wawrzynski et al., 2004). Themes constructed from these data, among others, included peer educators' gains in intrapersonal development (e.g., value clarification and self-confidence) and practical competence (e.g., public speaking and presentation skills) (Wawrzynski and Cox, 2005).

In the project's second stage, the focus group data helped inform the development of the *National Peer Education Survey (NPES)*, which is used to gather data and advance the evidence-based research on the learning outcomes peer educators report as a result of their experiences (see Wawrzynski et al., 2011 for more information about the study and survey). The NPES was developed through a multi-step process starting with a review of the literature by college health officials and those directing the project. This review specifically included identifying desired outcomes for college peer educators and student outcomes noted in the literature that correspond to the seven learning domains in *Learning Reconsidered* (National Association of Student Personnel Administrators [NASPA] and American College Personnel Association [ACPA], Keeling, 2004). In the first iteration of the NPES, Bloom's Taxonomy (1956) was also utilized because it offered language consistent with student learning outcomes identified in the literature and desired peer educator outcomes of interest. *Learning Reconsidered* served as a conceptual framework for the study because it provided language for the development of student learning outcomes, as its domains aligned with student learning outcomes, we were interested in observing. Specifically, we explored six learning domains:

- Cognitive complexity (develop solutions, analyze situations, apply prior experiences)
- Practical competence (effectively present programs, sense of purpose, provide feedback)
- Intrapersonal competence (recognize strengths/weaknesses, clarifying values and beliefs)
- Interpersonal competence (effectively work with others, manage conflicts, lead a group)
- Knowledge acquisition, construction, integration and application (re-evaluate previous assumptions, listen to and consider others' viewpoints, formulating approaches and solutions)
- Humanitarianism and civic engagement (have conversations with students different from oneself, responsible for helping to improve campus community)

Although the NPES has received slight modifications over the years, most notably with the release of *Learning Reconsidered 2* (Keeling, 2006) and substantive revisions in the *Council for the Advancement of Standards [CAS]* Professional Standards for Higher Education (CAS, 2019) it has remained a relatively valid and reliable instrument that certainly advances our understanding of gains peer educators derive from their experiences. In fact, since its debut, peer educators at over 400 institutions have participated in the administration of the NPES of which over 70 institutions do so annually. Institutions who participate in the study receive an institution specific report of the results, which allows them to assess the learning of peer educators in their programs. Although the

National Peer Educator Study specifically focuses on college peer educators who work with health and wellness initiatives, the learning outcomes are easily applicable to peer educators in other campus services and initiatives (e.g., resident assistants, peer tutors, orientation assistants).

Nelson Mandela University Student Learning Outcomes Initiative

The second assessment project, the Nelson Mandela University Student Learning Outcomes Initiative, is more than a decade old and part of a larger institution-wide student learning outcomes assessment effort undertaken at Nelson Mandela University (NMU), in Port Elizabeth, South Africa. The motivation for the development of this project was two-fold. First, the project explores larger aspects of how students engage in and benefit from the college experience in the re-structured, post-apartheid higher education system in South Africa (Wawrzynski et al., 2012). More specifically, particular attention in this project is on the efforts to understand the influence and contributions of cocurricular involvement to student learning, since a majority of a student's time is spent outside the classroom. Indeed, assessing student learning outcomes from other parts of the collegiate experience is now more important than ever, as high unemployment rates in South Africa, acerbated by the lack of traditional academic ways (e.g., internships) students gain skills that employers seek (e.g., collaborating with other members of a group, comparing and contrasting different points of view, and evaluating other points of view to come to a workable solution) presents unique opportunities to assess other experiences where students do gain them.

The second component of the project dovetailed the first aspect of the project to assess student learning outcomes from cocurricular involvement through the development of institutional learning outcomes and indicators. In order to gain support across campus, various campus stakeholders (e.g., students, faculty, student affairs staff) with an interest in student success were recruited to be part of the University Cocurricular Forum Committee (UCFC). Through a consultative process, the UCFC was tasked with developing a set of institutional student learning outcomes specific to cocurricular experiences. Members reviewed both national and international literature including culturally relevant literature (e.g., noncognitive variables, student learning domains from the Council for the Advancement of Standards in Higher Education (CAS, 2009), and CAS's Frameworks for Assessing Learning and Development Outcomes (Strayhorn, 2006) to adapt and create a set of 15 institutional learning outcomes with 66 developmental indicators for students participating in NMU cocurricular activities.

These institutional student learning outcomes and developmental indicators were then used as a conceptual framework to guide the creation of a cocurricular transcript and qualitative and quantitative assessment strategies. Using this framework, one quantitative assessment included the design of, and data collection with, the Student Life Survey (SLS). The Student Life Survey was developed through an iterative process with the UCFC, which was also consistent with survey design strategies (see Wawrzynski et al., 2012). For example, using the 66 developmental indicators, the SLS is a 39-item questionnaire and asks questions like, "Student life activities I participate in at NMU help me acquire skills that I can use after leaving the University." Additionally, since the SLS is administered anonymously, students are asked to provide demographic information, their participation in officially recognized societies and clubs, and other questions that measure sense of belonging and tied to student success.

The SLS is administered to NMU students as part of an annual data collection process assessing cocurricular learning. Students are solicited to participate in the SLS in one of two ways: (a) an email invitation is sent to students, and (b) student volunteers are recruited to canvas strategic locations across campus (e.g., taxi stops, public transportation areas, library, and residence dining halls) asking students to complete the online SLS using a tablet computer. These two assessment strategies were employed because many students do not have access to computers once they leave the university for the day. Although the SLS has received slight modifications over the years, it has shown to have excellent psychometrics.

Data analysis identified learning outcomes associated with collaboration, facilitating group discussions, cooperating with others to achieve a common purpose, appreciating human and cultural differences, seeking involvement with people who hold different points of view, understanding values and ethics, commitment to personal morals, greater self-awareness, and pursuing individual goals as a result of their involvement in cocurricular experiences. These results are disseminated to campus stakeholders so they can identify the areas of greatest and smallest reported growth for students who participate in various types of cocurricular experiences. Additionally, analysis results help guide larger institutional conversations on the value of student learning outcomes from cocurricular involvement.

Concluding thoughts

Certainly, there are benefits to developing thoughtful and intentional efforts to assess what students are learning from their college experience. Moreover, given the current climate of accountability, being able to provide empirical data that can highlight the contributions of cocurricular programs and events to students' growth and development has gained interest in higher education institutions globally. Collecting these data must start with developing observable and measurable student learning outcomes, which then can be coupled with effective assessment strategies to more adequately identify the learning students gain from college experiences.

References

- Adelman, C., 2015. To Imagine a Verb: The Language and Syntax of Learning Outcomes Statements. (Occasional Paper No. 24). University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NIOA), Urbana, IL.
- Bloom, B., Englehart, M., Furst, E., Hill, W., Krathwohl, D., 1956. Taxonomy of educational objectives: the classification of educational goals. In: *Handbook I: Cognitive Domain*. Longman, White Plains, NY.
- Coates, H., 2010. Development of the Australasian survey of student engagement (AUSSE). *High. Educ.* 60, 1–17.
- Coates, H., 2016. Assessing student learning outcomes internationally: insights and frontiers. *Assess. Eval. High. Educ.* 41 (5), 662–676.
- Council for the Advancement of Standards in Higher Education, 2019. *CAS Learning and Development Outcomes*, tenth ed. Author, Washington, DC.
- Council for the Advancement of Standards, 2009. *CAS Professional Standards for Higher Education*, seventh ed. Author, Washington, DC.
- Keeling, R.P. (Ed.), 2004. *Learning Reconsidered. A Campus-Wide Focus on the Student Experience*. American College Personnel Association and National Association of Student Personnel Administrators, Washington, DC.
- Keeling, R.P. (Ed.), 2006. *Learning Reconsidered 2: A Practical Guide to Implementing a Campus-Wide Focus on the Student Experience*. American College Personnel Association, Association of College and Housing Officers–International, Association of College Unions–International, National Academic Advising Association, National Association of Campus Activities, National Association of Student Personnel Administrators, & National Intramural–Recreational Sports Association, Washington, DC.
- Krathwohl, D., 2002. A revision of Bloom's taxonomy: an overview. *Theor. Pract.* 41 (4), 212–218.
- Kuh, G.D., 1993. In their own words: what students learn outside of the classroom. *Am. Educ. Res. J.* 30 (2), 277–304.
- Levin, B., 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago.
- National Survey of Student Engagement (NSSE), 2021. Happy 20th, NSSE! Entering NSSE's Third Decade: Retrospective and Prospective [Online]. Available at: https://nsse.indiana.edu/news-events/_news/2021-04-01.html. (Accessed 28 May 2021).
- Sedlacek, W.E., 2004. *Beyond the Big Test: Noncognitive Assessment in Higher Education*. Jossey-Bass, San Francisco, CA.
- Shi, J., Li, W., Chu, J., 2014. China College Student Survey (ccss): breaking open the black box of the process of learning. *Int. J. Chin. Educ.* 3, 132–159.
- Strayhorn, T., 2006. *Frameworks for Assessing Learning and Development Outcomes*. Council for the Advancement of Standards in Higher Education, Washington, DC. Author.
- Strydom, J.F., Mentz, M., 2010. *South African Survey of Student Engagement: Focusing the Student Experience on Success Through Student Engagement*. Council on Higher Education, Pretoria, South Africa.
- Wawrzynski, M.R., Cox, J., 2005. Identifying Outcomes of Peer Education on Peer Educators. Paper presented at the US Department of Education National Meeting, Indianapolis, IN.
- Wawrzynski, M.R., Lemon, J.D., 2019. Understanding student learning outcomes of peer educators. In: Trolan, T.L., Jach, E. (Eds.), *Applied Learning in Higher Education: Curricular and Co-curricular Experiences That Improve Student Learning*. New Directions for Higher Education. Wiley, San Francisco, CA, pp. 61–69.
- Wawrzynski, M.R., Steinke, K., Radford, A., 2004. Peer Education: A View Through the Lenses of Peer Educators. Paper presented at the American College Personnel Association, Philadelphia, PA.
- Wawrzynski, M.R., LoConte, C.L., Straker, E.J., 2011. The National Peer Educator Study: assessing outcomes of peer educators. L. B. Williams. In: *New Directions for Student Services*, vol. 133. Jossey Bass, San Francisco, CA, pp. 17–27.
- Wawrzynski, M.R., Heck, A., Remley, C., 2012. An exploration of student outcomes in a South African context. *J. Coll. Student Dev.* 53, 106–123.

Advantages and challenges of performance assessment of student learning in higher education

Olga Zlatkin-Troitschanskaia^a, Jennifer Fischer^a, Henry I. Braun^b, and Richard J. Shavelson^c, ^a Johannes Gutenberg University Mainz, Department of Law and Economics, Chair of Business and Economics Education, Mainz, Germany; ^b Boston College, Lynch School of Education and Human Development, Chestnut Hill, MA, United States; and ^c Stanford University, Graduate School of Education, Stanford, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Performance assessments in higher education	312
The Internet as an information and learning environment in higher education	312
Need for digital performance assessments in higher education	313
Assessment frameworks and formats for measuring student learning in higher education	313
Computer-based performance assessments in higher education	314
Digital performance assessments	314
Conceptual-theoretical and methodological foundation of performance assessments	315
Performance assessment framework	315
Measurement design	316
Assessment task format	316
Scoring	317
A systematic literature review	317
Current state of research that includes (digital) performance assessments in higher education	317
Advantages and challenges of (digital) performance assessments in higher education	319
Potential for educational practice	319
Challenges of implementing and using performance assessments in higher education	320
Conclusion	321
Acknowledgment	321
References	321
Appendix A	324

Performance assessments in higher education

The Internet as an information and learning environment in higher education

In the era of digitalization and globalization, society, and higher education in particular, is faced with the challenge of how to deal with a new, rapidly changing information and learning environment. The Internet presents a flood of complex, potentially conflicting and competing information of highly diverse quality and from media sources of varying credibility. New information and communication technologies (ICTs) provide countless benefits and opportunities for higher education and research, particularly with regard to new forms of knowledge generation, accumulation, and dissemination through (new) digital media. Established systems for knowledge generation, including education, research, and traditional media, have been complemented by new, predominantly web-based media and dissemination channels. The latter are emerging at an ever-increasing rate, while the quality and sources of information are becoming harder to track and monitor (Brooks, 2016; Rott and Garvin, 2015). Erroneous information, deliberately misleading information, and breaches of privacy are problems in and of themselves (Ciampaglia, 2018). Recent studies indicate both positive and negative effects of the use of digital media on student learning. For instance, these can include student information overload, when they cannot handle diverse knowledge sources of varying quality (Hahnel et al., 2019) and/or are unable to process and efficiently organize the information provided (for current meta-studies, see Molerov et al., 2020; Zlatkin-Troitschanskaia et al., 2021).

Today, the Internet is among the most widely used sources for information. Students increasingly learn with the help of Internet-based sources, including mass and social media, even more frequently than via traditional sources for learning, like textbooks (Brooks, 2016; Maurer et al., 2020). The online environment ostensibly presents unlimited and diverse information and sources of knowledge with highly diverse (e.g., domain-specific, sociocultural) attitudes and values. Not least during the COVID-19 pandemic it became evident, that there are various ways in which “information” can be understood. For instance, personal values significantly shape the way in which students access and process information. Therefore, it is not enough to present students with what is considered adequate or appropriate information, if students then apply different values for processing this information. Evidentially, with low access barriers, deficient quality control in publishing, and easy dissemination, especially online mass and social media often contain incomplete or biased information (Nagel et al., 2020; Machete and Turpin, 2020) that can lead to the acquisition of false domain-specific and generic knowledge and/or misconceptions among students (Rott and Garvin, 2015; Shahibi and Rusli, 2017; Weber et al., 2018a).

Need for digital performance assessments in higher education

Most recently, the COVID-19 pandemic has led to an historically significant change in learning practices and educational assessment. Distance learning in universities is being expanded, thus demanding better implementation of ICTs as well as integrated digital assessments. However, substantial deficits in both students' digital learning and digital assessments in higher education were recently observed (Marks et al., 2021; Raaheim et al., 2019). Thereby, two connected but distinct current developments should be distinguished: First, steadily increasing formal (and informal) uses of digital media in formal learning settings. Second, the role of digital assessment technologies, which are increasingly used to capture and measure students' information processing and online learning but rarely used as (formative or summative) assessments in formal educational settings. Since the pandemic will further lead to substantial and permanent changes in the media and learning environments (although the scope and impact of these changes remain unknown), this represents both an opportunity and a challenge for formal learning and assessment in higher education, i.e., the integration of (digital) assessment and learning.

The valid measurement of student competencies and learning outcomes in higher education, which include key 21st century skills like critical online reasoning is becoming more important than ever. Today, the competent use of (online) information is key to the functioning of educational (and societal) systems. The demand to validly measure student learning and student learning outcomes in digital and non-digital settings in higher education has gained momentum worldwide. As illustrated by the largest international comparative study in higher education, the Assessment of Higher Education Learning Outcomes (AHELOs), there have been both advances and challenges in the development of valid and reliable assessments of student learning outcomes and, in particular, in the use of (digital) PAs in higher education. AHELO examined the feasibility of validly measuring student learning outcomes internationally, including performance assessments (PAs) for measuring generic skills (OECD, 2012, 2013).

To date, many types of assessments have been developed to measure student learning outcomes, which include self-report surveys of learning and multiple-choice and short-answer tests (Braun, 2019; Liu et al., 2016; Sendurur, 2018; Sparks et al., 2016). However, so far only few formats have been tailored specifically to reliably and validly measure key student learning outcomes like critical reasoning in online media use (Goldman and Brand-Gruwel, 2018). In light of the challenges that higher education is facing in the context of a flood of (digital) information, current research stresses how (digital) PAs can drive the introduction of 21st century skills and deeper learning principles into educational practice (Darling-Hammond and Adamson, 2014). Such practices can foster much-needed skills for learning in digital settings, as students learn to reliably and autonomously build knowledge, while navigating the ocean of online information (McGrew et al., 2018; Weber et al., 2018a).

To investigate how the competence set required for successful learning in digital and non-digital settings can be effectively promoted in higher education, the underlying skills must be measured using valid, authentic, digital PAs.

Assessment frameworks and formats for measuring student learning in higher education

There is a consensus in current educational research that complex cognitive higher-order procedural skills and multifaceted competencies like critical online reasoning can hardly be validly measured with traditional standardized tests. For instance, multiple-choice knowledge tests usually only assess declarative and conceptual/factual knowledge (Braun, 2019). Closed test formats generally do not assess the complex, multidimensional competence constructs and instead focus on students' skills and knowledge facets (e.g., pedagogical content knowledge) (Davey et al., 2015).

As recent research reports indicate (Liu et al., 2016; Molerov et al., 2020), there are many closed-format tests for assessing complex competencies like critical thinking. One of the main shortcomings of these tests is their limited authenticity and ecological/"face" validity. Moreover, these instruments demonstrate a strong correlation with tests for general cognitive skills like intelligence tests or the Scholastic Assessment Test (SAT, Suckow et al., 2015). Ultimately, they fail to measure more specific procedural/functional skills, for instance the competent use of (online) information required for successful learning in higher education and beyond (Shavelson et al., 2019; Davey et al., 2015; Sparks et al., 2016). For instance, the well-established assessments of critical thinking skills are based on standard-setting research (Braun, 2019) but usually use multiple-choice formats and brief situational contexts (Liu et al., 2014). They mainly assess generic cognitive and inferencing abilities. For example, the most well-known test is the Watson-Glaser Critical Thinking Appraisal (Watson and Glaser, 2002) that comprises tasks on inferences, recognition of assumptions, deduction, interpretation, and evaluation of arguments. It is questionable to what extent the tasks are ecologically valid and whether students can transfer the measured skills to real, more complex requirement situations.

At the other end of the assessment spectrum are traditional essay prompts with open responses. Their suitability for validly assessing complex student competencies like critical thinking is generally limited by challenges in objective and reliable scoring and the brevity/openness of the task prompt.

Compared to the established assessment approaches (for an overview, see e.g., (Liu et al., 2014), students' complex, multifaceted competencies can be more adequately measured through PAs that present realistic scenarios and tasks (Davey et al., 2015; Braun, 2019). A PA typically includes one or more tasks, at least one of which must be a performance task and other tasks may be included such as short answer (for details, see Sections **Conceptual-theoretical and methodological foundation of performance assessments** and **Performance assessment framework**).

Digital PAs are better suited to reflect current learning contexts, conditions, and requirements both in and outside of higher education (Shavelson et al., 2019; Davey et al., 2015; Oliveri and Mislevy, 2019). They allow for higher education students to interact with a simulated digital environment to perform complex tasks and demonstrate skills they would use in real life situations.

Computer-based performance assessments in higher education

So far, few projects and assessments focus on measuring higher education students' performance using real-world tasks. The CAE sponsored the Collegiate Learning Assessment (CLA) developed by Klein, Shavelson, and Benjamin (Shavelson, 2010), which was also used in the AHELO study, and has launched a refined performance test of critical thinking, the CLA+. The revision comprises shorter performance tasks with an open-ended response format and is complemented by 25 selected-response questions on separate item stems so as to produce individual student scores (Zahner and Ciolfi, 2018). The CLA+ is available internationally for various countries (Wolf et al., 2015). It has been used in the US and was also adapted and used in Italy, the UK and Finland (Zahner and Ciolfi, 2018; Kleemola et al., 2021). It has undergone preliminary validation in Germany (Zlatkin-Troitschanskaia et al., 2018). This computer-based assessment consists of a performance task in which students are confronted with a complex case scenario, for instance focusing on health or ecological topics. Students are presented with a collection of documents with additional data, which simulate the information environment students in higher education are usually confronted with. The task prompt asks students to carry out a task, i.e., to critically evaluate the case including given documents and, based on this information, to decide on a course of action, i.e., write a reasoned decision. The CAE claims that the performance task component of the CLA+ measures the following abilities: problem-solving and analysis, writing effectiveness and mechanics, reasoning scientifically and quantitatively, reading critically and evaluatively, and critiquing an argument (Wolf et al., 2015).

Drawing on this research, the International Performance Assessment of Learning (iPAL) collaborative focuses on the measurement of 21st century skills, that include critical thinking, analytic reasoning, problem solving, and perspective taking. It employs complex tasks that simulate real-life decision-making and judgment situations (for details, Shavelson et al., 2018, 2019; Braun et al., 2020). To conduct cutting-edge, fundamental research and related practical development implementations of PAs and the corresponding teaching-and-learning tools in higher education, iPAL brings together researchers from multiple countries. In particular, iPAL develops valid PAs of university students' critical thinking in different countries and also adapts these assessments from one national context to another. In this manner, iPAL contributes to enhancing the comparability, validity, and fairness in international comparisons of higher-order cognitive skills in higher education as currently required (e.g., Suter, 2019; Berman et al., 2020). This work leverages the current state of technology and research in related domains and focuses on 21st century generic skills, like critical thinking, that university students are expected to develop in order to successfully address life's everyday decisions and challenges, as well as to carry out the responsibilities of engaged, democratic citizenship (for iPAL contributions, see Zlatkin-Troitschanskaia and Shavelson, 2019). The PAs developed to measure critical thinking skills go well beyond established task formats since they employ high fidelity "real-life" decision-making contexts and challenges.

Other recently developed assessments for higher education like HEIghten by ETS (Liu et al., 2016) can be considered knowledge-based, closed response format tests of analytical thinking as they do not encompass any performance tasks (Braun, 2019; Machete and Turpin, 2020).

Digital performance assessments

Research on learning in digital or digitally supported environments and on online learning outside of institutionally controlled environments, for instance by means of mass and social media, is still scarce (for overviews, Molerov et al., 2020; Maurer et al., 2020; Zlatkin-Troitschanskaia et al., 2021). Despite the urgency of this topic for higher education, digital PA tools have so far been developed by only few projects internationally, as also our literature review presented here demonstrates.

In the following, we focus on digital PAs of generic skills, in particular critical thinking and related constructs for two reasons: (1) valid assessments of such domain-independent skills are of particular importance to life-long learning in the Internet age, and (2) they are applicable not only across disciplines and domains but also across institutions and nations. Due to this focus, however, we cannot address the full spectrum of higher education and learning topics to which digital PAs could be applied (for more general use of PAs, Braun, 2019).

Generally, digital PAs allow for measuring new complex competencies (e.g., digital literacy) or established constructs (e.g., critical thinking) in different ways. Most recently, one digital assessment of "civic online reasoning" was developed and validated in the US by Wineburg et al. (2016). In the large-scale US study, students' critical use of digital information was examined in a realistic online environment, i.e., the tasks use real websites that focus on civic, consumer, and general knowledge topics. Students are asked to perform an unrestricted online search to solve the tasks. In addition, this digital assessment was successfully implemented as a training tool to significantly promote civic online reasoning among US undergraduates in an intervention study (McGrew et al., 2018, 2019). The assessment was adapted and validated for use in German higher education (Molerov et al., 2020). The newly developed German assessment, CORA (Critical Online Reasoning Assessment), measures students' searching for, critically evaluating, selecting, processing and use of online sources and information more comprehensively. In the short performance tasks, for instance students evaluate the credibility of websites by performing an open web search, judging online information, and writing a short open-ended response for each task. The particular difficulty lies in evaluating a source in a short period of time while recognizing hidden interests and biased information representations (for details on CORA, Molerov et al., 2020).

Recently, other digital PAs were developed for higher education (for an overview, see Table A.1 in the appendix). Many newly developed digital tests focus on multiple document use/comprehension as a particularly important competence of higher education students in the Information age (for examples, Grossnickle Peterson and Alexander, 2020; Schoor et al., 2020). As new ICTs have led to an exponential increase in the data volume available in daily life (Bonikowska et al., 2019), other digital PAs for higher education

target related domains like data literacy skills, which are also considered essential for participating in the (academic and professional) world of the 21st century. Data-literate individuals for instance need to identify issues and challenges associated with this overflow of information (Leighton et al., 2021).

To provide a more systematic overview, the results of a current literature review on digital PA is presented next. It particularly focuses on students' online information processing and critical use of digital information as key competencies for successfully studying in today's learning environments. After the description of the conceptual-theoretical and methodological foundations of PA, we present the approach and central results of the literature review. The research synopsis includes a critical evaluation and international comparison of assessments and analysis approaches, their frontiers, advances, and perspectives for future research.

Conceptual-theoretical and methodological foundation of performance assessments

Over the last decade, the emphasis in educational assessment has shifted from "analytic models" of particular components of knowledge and specific skills, for which multiple-choice tests are well suited, to holistic, integrative models of performance, for which PAs are well suited (Shavelson et al., 2019). "A performance assessment [...] is an activity or set of activities that requires test takers, either individually or in groups, to generate products or performances in response to a complex, most often real-world task. These products and performances provide observable evidence bearing on test takers' knowledge, skills, and abilities—their competencies—in completing the assessment [...]" (Davey et al., 2015, p. 10). PAs, then, model and measure the outcome of learning processes not only at the level of particular learners' dispositions and abilities (e.g., domain-specific knowledge) but also at the level of their actual performance (i.e., manifested action) (for examples, see Table A.1 in the appendix).

A set of various competencies is generally required to master tasks in everyday academic and/or professional practice. A criterion-sampling, holistic approach is fundamental to the PA design that recognizes the complex multifaceted nature of 21st century skills to be measured (e.g., critical thinking) and conceives the performance as being more than the sum of its component parts, i.e., underlying individual skills (Shavelson, 2013). Real-world situations demanding these skills can rarely be segmented in ways that require only discrete skill components. Obviously, subsets of such skills are employed in sequence over the course of addressing situation-specific challenges (Corno et al., 2002). Based on Snow's (Snow, 1996) concept of aptitude, Corno et al. (2002, p. 3) refer to the concept of "aptitudes," which he defines as "all ... [those] characteristics (e.g., experience, ability, knowledge, motivation, and regulatory processes) that an individual brings to and cobbles together to perform in a particular situation." These "aptitudes" are considered dynamic since their characteristics can change depending on amendments in the (task) environment (Shavelson et al., 2002).

A holistic approach to assessment measures competencies in an environment that resembles real-life situations as closely as possible and therefore encompasses all skills and abilities that are required for meeting academic and/or professional demands needed in these situations (task scenarios) (Weinert, 2001). The holistic approach is based on the "criterion-sampling" approach to measurement (McClelland, 1973), which asserts that solving complex performance tasks requires an integrated set of abilities that cannot be validly assessed when divided into and measured as individual components. The key tenet of this approach is as follows: identifying what a learner knows and can do (e.g., search with Google), sampling real-life tasks from the domain in which that learner has to act (e.g., the Internet), directly observing participants actual performance (in this domain or an authentic simulation), and from (multiple) observations, inferring their competence and learning progress. As this is a complex process, a learners' actual ability, for instance to search with Google, cannot be assessed validly using a multiple-choice test, which would only be able to measure, for instance declarative knowledge about search operators or Internet-based beliefs. To validly measure students' actual abilities in searching, a test for searching the Internet that incorporates a sample of typical authentic searching tasks is required. Based on this sample of performance, valid inferences about a student's more general search performance can be drawn (Mislevy, 2018). However, the external validity of the measurement and generalizability of the results can be limited, for instance due to strong interactions with the scenario content; and sufficient reliability requires multiple observations, which may be difficult to realize due to the long testing time (Davey et al., 2015).

Consequently, instead of developing a "unidimensional" measurement for each particular competence sub-component (e.g., knowledge facet), typical situations are sampled from real-world contexts and are adapted for use in an assessment following a detailed design framework that provides definition of construct, organization of test, and a basis for scoring students' responses to holistically assess performance (for details, see next Section). Therefore, the PA design framework recognizes that properly addressing the situation-specific demands requires the application of combinations of knowledge and skills measured in the PA (Shavelson et al., 2019).

Performance assessment framework

A performance assessment framework specifies the construct definition, corresponding characteristics of the scenarios used, the nature and format of the materials used, and the nature of the required task-solving processes and responses to the task, i.e., performance, and scoring system (for an example, see Shavelson et al., 2019; Braun et al., 2020).

Measurement design

Most frameworks of current PAs, including the iPAL framework, are based on Mislevy's Evidence-Centered Design (ECD). With the ECD strategy, three connected models are differentiated and further specified: (1) a student model that represents the abilities that students are to exhibit in carrying out a task; (2) a task model that specifies task features that are supposed to elicit behaviors that allow inferences about these abilities; and (3) an evidence or "interpretative model" that both specifies the features to be extracted from the response and describes the way in which the test task scores are related to (latent) student abilities. The interpretative model includes the measurement model, i.e. the scoring rubrics and the statistical method for obtaining test score(s) (for additional models and more details on the ECD, [Mislevy et al., 2003, 2012](#)).

Accordingly, the PA design is based on the definition of a construct (e.g., students' abilities and mental processes) to be measured ("construct definition") as well as on a "test definition" that describes how this construct can be operationalized in order to be tested by PAs. Therefore, the following question arises: What kinds of situations (task prompts, psychological stimuli) are suited to validly measure students' abilities and mental processes in accordance with the construct definition? (for details and discussion, see [Mislevy, 2018](#)).

In the first step, a construct definition for the skills to be measured via the PA, for instance critical online reasoning, must be established. Next, a real-world situation or problem that requires these skills is selected. For instance, a defining characteristic for PAs tapping "information problem solving" and related constructs, is that elements are "manipulated" in the problem or (task) scenario that are most likely to evoke critical thinking, i.e., "the process of conceptualizing, analyzing or synthesizing, and evaluating and applying information to solve a problem or decide on a course of action or to find an answer or reach a conclusion" ([Shavelson et al., 2018](#), p. 203).

Following the criterion-sampling approach, the PA's designer samples typical, everyday situations demanding such skills (e.g., searching Google for study purposes) to create authentic performance task scenarios and to identify appropriate information sources. Documents and other types of media provided in the PA task are sampled according to three criteria: (1) the credibility of the sources, (2) the relevance of the information for the particular problem or case, and (3) the information's susceptibility to reasoning bias (e.g., authority bias, [Paul and Elder, 2005](#)) and judgmental errors, including decision-making heuristics (e.g., rational thought, [Kahneman, 2011](#); [Stanovich, 2009, 2016](#)), which are typical when students use information to make judgments and/or decisions in the real world. Thereby, different types of heuristics are incorporated into the PAs that simulate everyday situations with high fidelity.

The PA tasks are typically complex and do not offer obvious paths to solutions, decisions, or actions. They are authentic as they represent current, everyday challenges that learners face as university students, graduates, professionals and, more generally, as citizens. For instance, the PA can present situations in which deliberation is required to reason among alternative decisions or courses of action, with due regard to their consequences ([Shavelson et al., 2019](#)). The participant generates an answer to a task question, develops a product or is observed while performing a task. This type of testing enables observing the actual performance of participants from which the underlying test taker's knowledge, skills, and abilities in a professional domain are inferred ([Davey et al., 2015](#)).

For instance, the iPAL's PAs of critical thinking introduce students to a situation based on a real-world event (storyline) and ask them to take on a specific role (e.g., member of an advisory board). Students are provided with an accompanying document library that includes a selected set of documents with various perspectives and arguments from various information sources including government reports, newspaper articles, web blogs, YouTube videos, research reports, journal articles and the like. The information varies in its trustworthiness, relevance, and susceptibility to bias and judgmental errors ([Braun et al., 2020](#)). The PAs ask students to evaluate the situation they are presented with and to justify their decisions and recommendations using the information provided. The complex PA tasks generally do not admit a single solution path. Different reasoning and argumentation approaches are expected and also evident ([Shavelson et al., 2019](#); [Zlatkin-Troitschanskaia et al., 2019, 2020](#)).

As an example, one of these PA tasks, "Wind Turbine", was developed to assess higher education students' critical thinking. It consists of a realistic short-frame scenario and incorporates a document library for additional background information on the case including 22 snippets and sources of information that vary in their relevance and trustworthiness. The question at hand: Should the town approve the construction of a farm of wind turbines? Students are assigned a role in a communal real-world decision-making scenario in a German town and are asked to complete a review and to use the information provided to write a policy recommendation for a course of action, i.e., whether or not wind turbines should be built in their small town. The total test time for the PA task is 60 min. The participants are required to decide, based on relevance, which sources and arguments to select and review more thoroughly, as there is not enough time to focus on all of them ([Shavelson et al., 2019](#); [Zlatkin-Troitschanskaia et al., 2019, 2020](#)). The type of communicative action to be taken adds further formal requirements. Students receive higher scores on the PA tasks when they, for instance, (1) recognize and use credible information and avoid less-than-credible or peripheral information (facet "use of sources and arguments"); (2) avoid judgmental and decision-making "traps" and biases (facet "avoidance of, or critical reflection on, heuristics and specific distractor information"); (3) consider alternative courses of action to the one proposed and indicate why they are giving a certain recommendation; show consideration of "pro"/"con" arguments, including an explanation of different points of view in the task scenario (for details, [Zlatkin-Troitschanskaia et al., 2019, 2020](#)).

Assessment task format

Generally, PAs have an open-ended format and are typically administered on a computer platform and, increasingly often, via the Internet. Digital assessment approaches provide several significant advantages, for instance, in terms of administering tasks, collecting response process and performance data as well as in their similarity to the real world (information and learning environment

that higher education students use) they are to represent, i.e., fidelity to the criterion situation being simulated. The Internet-based assessment environment or its computer-based simulation (Intranet) enable the use of audio, video, and also computer-based simulations including virtual and augmented reality to enhance the fidelity of the task and collect students' written and/or oral responses. However, computer-based and digital assessments can be technically demanding, cost-intensive, and have specific requirements, for instance, in terms of data security and privacy.

The PAs are often accompanied by additional tasks in other formats, like short, self-contained tasks with either short answers or selected responses (e.g., multiple-choice questions), to measure students' abilities and skills theoretically related to the test construct, for instance, general intellectual abilities or other so-called conventional skills (e.g., reading skills). The use of accompanying closed test tasks can improve the reliability of the full assessment that includes the PA.

Scoring

To evaluate the students' constructed responses, analytical multidimensional scoring rubrics need to be developed based on the construct definition. Scoring rubrics should be based on all construct facets measured in PAs and include fine-grained differentiations of score levels aligned to the construct definition. This is also important in terms of the scalability of (digital) PAs. For instance, rubrics should recognize students' use of credible or untrustworthy and relevant or irrelevant information, as well as their consideration of differing arguments and perspectives. The rubric should also consider whether students tap in to, or avoid, heuristics that could lead to errors in judgment and decision-making, whether they use information provided when justifying decisions, solving problems and/or recommending certain actions, including argumentation, and the appropriate use of evidence to support their arguments (for a scoring scheme of a PA, see Zlatkin-Troitschanskaia et al., 2019; Bloom and Deyrup, 2015).

In digital PAs, the real online assessment environment constitutes a crucial aspect to be considered in a reliable scoring scheme of task performance (Molero et al., 2020). In a simulated, closed information environment with a selection of sources, an evaluation of all given information is possible. However, in an Internet-based assessment environment, an indefinitely large number of sources are available. Therefore, processing and solving PA tasks in an open environment requires a different set of skills. This substantial difference in the *task model* should be appropriately reflected in the *student model* and the corresponding scoring scheme (for examples of a scoring scheme for closed information environment, Zlatkin-Troitschanskaia et al., 2019; Brand-Gruwel et al., 2017; for examples of a scoring scheme for open assessment environment, Molero et al., 2020; Powers, 2019). The evaluation should be carried out by at least two trained scorers per task, and a high level of interrater agreement must be achieved.

A systematic literature review

A systematic literature and database research was conducted in the first half of 2020 to generate a comprehensive structured overview of digital PAs used in higher education research that focus on students' search for, selection and evaluation of online information. As a first step, key central terms, including relevant concepts and synonyms as well as data collection and analysis methods, were defined. These defined terms and operators were used to carry out a systematic database search. Two internationally relevant search engines, Education Resources Information Center (ERIC) and Google Scholar, were primarily utilized to find literature on this topic. For additional research, Research Gate was used. First, all sources published in English that address the use of digital PAs and/or online surveys of students in higher education were considered. In a first and second search phase, in which search phrases and their combinations were systematically increased and changed step-by-step, over 500 titles and abstracts that sounded relevant were identified. In the next step, numerous selection criteria were defined and applied. For instance, only sources published since 2009 were included. In addition, only empirical studies were considered for further analysis. Next, all publications that did not specifically include higher education students in their sample, as well as all studies that did not use digital PAs, were excluded. After this multi-step selection process including the screening of about 100 abstracts and the full texts of 55 publications, 19 publications remained that fully met all selection criteria and therefore were included in further qualitative analysis (for a detailed overview of the included studies, see Table A.1 in the appendix).

Current state of research that includes (digital) performance assessments in higher education

All 19 publications employed digital performance tasks to measure students' online information processing (and/or related constructs such as "information problem solving" (Brand-Gruwel et al., 2017), "critical online reasoning" (Molero et al., 2020) and students' actual use of online media for study purposes). However, their specific research foci and scope of the analysis differed greatly. For instance, some digital PAs capture the complete search process (Bloom and Deyrup, 2015; Hargittai et al., 2010), while others focus on the (critical) evaluation and/or selection of online sources (Leeder and Shah, 2016) or the use of specific (domain-specific) databases (Dahlen et al., 2020; Furi and Balog, 2016).

The topics varied as well. Their scope ranges from the domains of medicine, economics, and linguistics (Judd and Kennedy, 2011; Raido, 2013) to less domain-specific contexts, for instance societal and health topics (Nagel et al., 2020; Quinn et al., 2017). Every study included had a local, regional, or national, but never cross-national scope. About half of the studies ($n = 9$)

were conducted in the USA, four studies were conducted in Germany, and one study each was conducted in the Netherlands, Croatia, UK, New Zealand, Australia, and Turkey.

The digital PAs were mediated by a computer and/or video recordings. Log file data were collected frequently ($n = 9$) as well as audio (voice) recordings ($n = 1$), and/or think-aloud approaches ($n = 7$) to collect students' response processes and performance data. Two studies collected data using eye-tracking (for an overview, see [Table 1](#), for more detailed information on the studies, see [Table A.1](#) in the appendix). In addition to quantitative approaches, qualitative analyses were often conducted to evaluate the collected data. In all studies, the students' answers and in some cases their online behavior while processing the PA task was rated using (original) scoring schemes (e.g., [Nagel et al., 2020](#)). Quantitative analyses were carried out, for instance when examining the processing time required and/or the number of clicks for task completion. The collection of log data ($n = 7$) allows the use of complex analysis methods, including process mining ([Schmidt et al., 2020](#)).

In addition to the collection of screen and/or video recording and log file data, self-reporting was deployed in all studies to obtain further information, for instance, regarding the participants' Internet-related beliefs ([Salehi et al., 2018](#)). Selected-response items with Likert scaling were mainly used for this purpose. Moreover, socio-demographic data (e.g., gender, age), were usually collected.

The majority of the 19 studies (17) have a cross-sectional design (some with a treatment and a control group). The other two are intervention studies which used a pre-post-design ([Rosman et al., 2015](#); [Weber et al., 2018b](#)). Apart from this, there was no study with a longitudinal design.

The studies show a wide range regarding the number of participants (minimum = 4; maximum = 842). Most of the studies report a sample of up to 55 participants ($n = 15$). Notable are the studies by [Judd and Kennedy \(2011\)](#) with 842 and the study by [Hahnel et al. \(2018\)](#) with 310 participants.

In most studies, students from different study semesters were surveyed. Across all studies, both beginning and advanced students and graduates were assessed (for an overview, see [Table A.1](#) in the appendix).

The studies illustrate various PA approaches to measuring students' online information use and related skills. They do not focus on single skills and/or knowledge facets but rather address multidimensional competence constructs needed for solving the PA tasks. A majority of the studies ($n = 11$) applied PA tasks with computer-based simulations, i.e., created a closed, controlled assessment environment in which the participants were asked to conduct research on a given topic. However, no studies could be found that apply virtual or augmented reality (though such studies are available in school education). While some studies used several tasks that differed in their level of complexity and difficulty, many others relied on but one or a few tasks with equal levels of difficulty. Some studies provided a particular pre-selected website in a PA task and requested the participants to assess it in terms of credibility and trustworthiness. Approximately half of the studies focused on a free web search that students had to conduct to solve the PA task. In these, the comparability of participants' test results, however, is limited since each student's search created a unique task environment. At the same time, this type of assessment is more authentic and ecologically valid, offering a much more detailed representation of students' actual performances that supports its validity (for details and discussion, see [Molero et al., 2020](#)). Notably, in four studies, participants were asked to investigate their own research topic or a topic of interest. This may positively influence test motivation and therefore test performance in comparison to PA tasks with a given topic. However, again, the comparability of students' performance is limited in these cases because of topic variation.

Table 1 Used methods.

<i>Study</i>	<i>Screen-casting</i>	<i>Think-aloud</i>	<i>Eye-tracking</i>	<i>Log-files</i>	<i>Voice-recording</i>
Bloom and Deyrup (2015)	X	X			X
Brand-Gruwel et al. (2017)		X	X	X	
Dahlen et al. (2020)	X				
Furi and Balog (2016)	X	X			
Hahnel et al. (2018)				X	
Hargittai et al. (2010)	X	X			
Holman (2011)	X	X			
Judd and Kennedy (2011)				X	
Leeder and Shah (2016)				X	
Liu et al. (2018)					
Nagel et al. (2020)				X	
Powers (2019)	X	X			
Quinn et al. (2017)	X			X	
Raido (2013)					
Rosman et al. (2015)					
Salehi et al. (2018)	X				
Sendurur (2018)		X	X		
Tylor (2015)	X				
Weber et al. (2018b)				X	
Total	9	7	2	7	1

In two studies, participants were not given specific tasks. Instead, the students were observed, and their research process was recorded via screencast as they worked and searched the Internet to prepare documents relevant for their classes (Bloom and Deyrup, 2015; Tylor, 2015). For instance, the participants of one study had to translate a popular-science text from Spanish into English for a class and were asked to use the Internet as an external resource of consultation (Raido, 2013). Overall, since all PA tasks reviewed can be considered to be complex, they differ considerably, for instance, in terms of real (i.e., Internet) or simulated assessment environment, including psychological stimuli such as potentially conflicting information of diverse quality and scope.

Advantages and challenges of (digital) performance assessments in higher education

Through examination of these 19 studies, an emerging multidisciplinary field of research on digital PAs in higher education becomes evident. It specifically focuses on students' online information processing and (critical) use of online information in study-related contexts. The more critical review of current research studies reveals that existing research using digital PAs is characterized by both various advantages and shortcomings. In general, digital PAs enable measuring of new complex competency constructs (e.g., critical online reasoning) or established constructs (e.g., multiple source use) in a more objective and valid way. Most studies are characterized by innovative PA designs, and in many studies, the collected data were used for innovative analyses (e.g., process mining), which provide novel insights into student learning in various (digital) contexts and student learning outcomes.

However, one of the shortcomings of the current state of research is that process data were collected and analyzed in less than half of the studies. In this respect, the particular potential of digital PAs not only for measurement purposes but also in terms of (improving) student learning in the higher education practice remains untapped. This is because particularly important information about the learning processes of students can be obtained on the basis of response process data (Kane and Mislevy, 2017; Zumbo and Hubley, 2017). By means of learning analytics, student learning can be supported quite significantly through the use of digital assessments (Drachler and Goldhammer, 2020; Jivet et al., 2020).

Another major limitation of current research is its highly exploratory and descriptive character. Many studies did not employ theory-based research hypotheses, and the results obtained using digital PAs are mainly considered in terms of the extent and level of the students' skills investigated. The development of these skills and the possible connections to university study contexts and actual student learning were explicitly considered in but two intervention studies. In this context, there is also a lack of longitudinal studies that capture and analyze the development of students' skills through the course of study using digital PAs.

The complexity of the analyzed studies is also reflected in the scope and amount of collected response process and performance data. For instance, interview data from think-aloud protocols or eye-tracking data were analyzed quantitatively, qualitatively, and via more innovative mixed-methods approaches (see Table 1). However, as the use of qualitative methods is typically very time- and resource-consuming, samples are often small and range from 10 to 50 participants in most studies. Since current and ongoing technological advantages and more complex analytical methods enable the evaluation of "big" learning data, the collection and analysis of larger samples must be extended and improved so that future research in this field can provide more valid and generalizable findings (Mislevy, 2018).

Another limitation is the local or regional scope of the research identified. None of the studies pursued an explicitly comparative focus, for instance, between students from different content domains, study groups or universities. Nor could any internationally comparative study be found. In a digitalized world, however, cross-national comparative studies, especially in higher education, are highly important to provide more evidence on student learning in the global information and learning environment, i.e., the Internet. The use of digital PAs, in particular, provides the opportunity for international deployment. For instance, performance and process data (e.g., log files) of test takers can be collected from anywhere and anytime (e.g., in different time zones simultaneously). However, in contrast to these and further opportunities, digital testing also poses specific technical challenges. For instance, the methods used to collect data often require specific technical conditions and/or additional on-site support (Newman et al., 2021). Moreover, additional challenges in terms of translation and adaptation and the comparability of PAs arise (e.g., Ronderos et al., 2021). As recent studies indicate, it is not enough to simply deploy digital assessment across countries. Rather, test administration procedures and analyses need to be comparable to ensure the comparability of the test results. This includes the implementation of similar processes across countries and institutions for pretesting, sampling, task administration, incentivization, validity and comparability analyses (ITC, 2017; for more details on the challenges of international assessment, see e.g., Suter, 2019).

Many of these strengths and limitations lie in the nature of digital PAs, and they are outlined in more detail in the following by highlighting specific advantages and challenges for higher education and student learning practice and research.

Potential for educational practice

After establishing the reliability and validity of the digital PA, these tools can be used effectively by universities nationally and cross-nationally (1) to produce valid summative measures and (2) formatively evaluate individual students and/or groups of students (Shavelson et al., 2019). The potential for the formative use of PAs in classrooms targeting students' skills in real-world situations is especially promising and twofold. The PAs can be used to measure students' learning in higher education and to train students (and teachers) in the assessed skills. In this way, and in the sense of the so-called "Curriculum-Instruction-Assessment Triad"

(Pellegrino, 2002), the digital assessment can improve both teacher instructions and student learning and therefore contribute to better learning outcomes in higher education and beyond.

In addition, by using PAs, universities can generate evidence in terms of the efficacy of courses of study and co-curricular/instructional activities (Braun, 2019) to ultimately improve teaching and learning in higher education. Therefore, the most important use for PAs is in the classroom, where tasks, e.g., with computer-based simulations (possibly including virtual and augmented reality) can provide the basis for problem-based and/or simulation-based learning models (Chernikova et al., 2020) and for innovative pedagogical interventions (Berliner, 2020). The problem-based learning (PBL) model proposes that well-designed and authentically presented real-life problems can motivate students to solve more challenging tasks than they would otherwise be inclined to do. Such tasks can encourage critical questioning and elaboration by presenting authentic and interesting situations and real-life simulations (Barrows, 1996). Research on interventions based on PBL indicate that such approaches can support deeper learning and transfer (Gijbels et al., 2005; Strobel and van Barneveld, 2009) as required for and in higher education.

Digital PAs can drive innovation of assessment practices and online learning, even beyond formalized higher education contexts (for a discussion of more general use of PAs, see Braun, 2019). For instance, the case-based PA task format is suited for reproduction and use in further technological contexts, e.g., virtual and augmented reality. It is likely to spark the development of both assessment tasks and accompanying teaching-and-learning materials, such as online (self-)assessments, training materials, and video courses. For instance, the Stanford History Education Group, successfully transferred research with PAs into a series of online training videos aimed at the average consumer (Crash Course Navigating Digital Information). Thus, based on (digital) PAs and their results, instructional interventions in higher education can be designed, for instance, to train students how to competently use and reason from online information and equip them with the skillset necessary for acquiring high-quality knowledge using the Internet (for examples in teacher, law, and medicine training, see Zlatkin-Troitschanskaia et al., 2021). This can help universities to fulfill their educational mandate in the context of growing global competition stemming from the Internet.

Challenges of implementing and using performance assessments in higher education

PAs provide numerous opportunities to holistically measure students' complex, multifaceted competencies and learning outcomes by using authentic real-life and professional tasks. The PAs present specific (task) situations that require students (individually or in groups, e.g., in classrooms) to create performances and/or products in response. These performances and/or products provide evidence on the basis of which inferences of the students' complex competencies and higher-order cognitive skills can be made — within or across academic and/or professional domains.

This type of assessment is accompanied by specific challenges that need to be considered in educational research and practice. First, as the considerations on PA design demonstrate (see Section **Performance assessment framework**), PA tasks are challenging to create and implement, and they require intensive pre-testing and validation work to relate students' test performances to a score by using a differentiated rating scheme (e.g., Molerov et al., 2020).

Second, the more general challenge lies in the nature of PAs, which involve complex, real-world tasks (i.e., simulations) and corresponding, open real-life responses that are usually complex, relatively long, often nonstandardized and limited due to (test) time and cost factors. This leads to the issues of limited reliability, validity, and scalability of (digital) PAs (see Section **Performance assessment framework** and below).

More specifically, there are numerous methodological and psychometric challenges, as discussed, for example, in Davey et al. (2015), Shavelson et al. (2018), and Mislevy (2018). First, most commonly used measurement and psychometric models were developed for closed, standardized, selected response questions (e.g., multiple-choice tasks). In these assessments, several discrete test tasks are scored dichotomously and often organized into particular scales that measure one distinctly defined (unidimensional) construct (e.g., fluid intelligence). That is usually not the case in the PAs, which holistically measure student's ability to demonstrate multiple skills, so that the interpretation of the test results and inferring students' abilities from actual performance are demanding (Shavelson et al., 2019; Zlatkin-Troitschanskaia et al., 2019).

Second, the ordinarily limited number of observations leads to additional specific challenges since standard measurement and psychometric models, especially those modeling and assessing complex, multifaceted, latent constructs, typically require a large number of observations of performance to infer the assessed construct (e.g., critical thinking) to achieve an acceptable model fit (e.g., Zlatkin-Troitschanskaia et al., 2019). However, (test) time and costs for PAs, for instance, regarding scoring of open responses, lead to a practical limitation in terms of the possible number of observations. As the literature review also illustrates, most studies only used one or a few PA tasks and only a limited number of students (see Table A.1 in the appendix).

In this respect, reliable scoring poses a significant challenge and limitation, as responses to a PA cannot generally be scored dichotomously but typically require the application of multiple rubrics, resulting in multiple, varied, and more unusual score scales of different types (for examples, Bloom and Deyrup, 2015; Zlatkin-Troitschanskaia et al., 2019). In cases where response process data (e.g., log files) are collected in addition to performance data, further process-related indicators, such as response time, number of clicks, etc., need to be used (e.g., Nagel et al., 2020; Schmidt et al., 2020).

A corresponding challenge, therefore, lies in human rating and the associated limitations in terms of reliability and validity of scoring. Student responses and/or observations in PAs are usually scored by trained raters, according to defined scoring rubrics. Despite high-level intensive training, limitations resulting from rater inaccuracy and rater variation (e.g., interrater reliability) cannot be fully ruled out and can lead to significant measurement errors. Although the use of PAs in the classroom promises many pedagogical advantages, the evaluation and analysis of these data is particularly challenging, as the responses can be

influenced by all students in a group. In this context, one particularly promising perspective can be the development of computer models for (semi)automated scoring based on students' written responses (for successful examples of computer scoring in educational assessment research, e.g., [Wahlen et al., 2020](#)). Computer-based scoring can provide the basis for learning analytics-based tutoring systems with automated, real-time feedback sensitive to online context, content, and learning process (e.g., alerting students to suitable verification steps based on textual cues or timing in web behavior). However, valid and reliable computer scoring of open-ended responses pose numerous specific challenges (for examples, [Horbach and Zesch, 2019](#); [Wind et al., 2018](#)).

Conclusion

The assessment of student learning outcomes and learning development has gained increasing importance in recent years. The selection of suitable measuring instruments is crucial, as each instrument has its strengths and limitations. In addition to traditional on-site PAs, there are assessments conducted in a digital format, which are increasing in number ([Braun, 2019](#); [Oliveri and Mislevy, 2019](#)). This trend has amplified, especially in recent years and due to the Covid-19 pandemic.

This chapter has provided a systematic, structured overview of different types of (digital) PAs, their application in educational research and practice, including the opportunities for improving student learning as well as challenges arising from their creation and use. For this purpose, a comprehensive literature review was conducted. To create comparability between the assessments presented, the research was further focused on digital PA when measuring students use of online information, critical thinking, and related constructs. Consequently, this review could not address the full spectrum of higher education and student learning topics to which (digital) PA could be applied (for more general uses of PAs, [Braun, 2019](#)).

Today, students' learning using the Internet has become one of the most vital prerequisites for academic success in higher education. Prior research indicated that Internet-based learning includes collecting, comparing, managing, evaluating, and applying information in a critical manner to make effective decisions in knowledge-based and technologically rich environments. Therefore, this literature review focusses on the current state of research using increasingly popular digital PAs to measure student learning progress and outcomes.

PAs are complex tasks and attempt to simulate the real world as closely as possible. These authentic tasks are drawn from real-life situations and follow a holistic approach to measure complex, multifaceted competences important for meeting specific (e.g., academic and/or professional) demands. These competencies cannot be neatly divided analytically into single components and are therefore analyzed by observing the distinct processes/steps the students go through when handling a PA task and their overall performance ([Shavelson et al., 2019](#); [Schmidt et al., 2020](#)).

The high level of authenticity is particularly attractive when measuring professional and/or academic skills. For instance, at the end of a student's course, the PA can provide evidence on how competent the student is in the course-related subject, as well as evidence on which job-related knowledge and skills the student has acquired. The use of digital PAs offers ground-breaking possibilities in educational teaching and assessment practice. Particularly, this enables the valid measurement of complex competencies constructs like critical thinking, for instance due to the high ecological/"face" validity and the collection of response process data (e.g., log files), which allows for complex subsequent analyses (e.g., process mining), learning analytics-based feedback and more individualized instructional interventions. Ultimately, in this way, the digital PAs can substantially contribute to improving educational practice and student learning in digital contexts.

Although PAs offer many opportunities, test developers, researchers and educators are faced with a number of conceptual, technical, and methodological challenges when creating and using them in research and practice (see Section **Challenges of implementing and using performance assessments in higher education**). For instance, the scoring and (qualitative) analysis of response process and performance data can be very time- and resource-consuming ([Dolan and Burling, 2018](#)). This results in small samples and cross-sectional (not longitudinal) designs, as revealed in the literature review presented here and which go along with various limitations, for instance, regarding the internal and external validity and limited generalizability of results. Moreover, no international study using digital PAs in higher education could be found.

Overall, by providing a comparison of different PAs at an international level this literature review particularly contributes to a greater transparency regarding existing research, its interrelationships, and overlaps. We hope that our critical discussions and suggestions for further research motivates much needed cross-national comparative studies in this field (for a critical review on the development of large-scale international comparative studies, see [Suter, 2019](#)).

Acknowledgment

This study is part of the PLATO project, which is funded by the German federal state of Rhineland-Palatinate. We would like to thank our student assistant Jan Krstev, who supported us in conducting the literature review.

References

- Barrows, H.S., 1996. Problem-based learning in medicine and beyond: a brief overview. *N. Dir. Teach. Learn.* 3–12.
- Berliner, D., 2020. The role of modeling for 'seeking truth' in an educational policy classroom. In: Zlatkin-Troitschanskaia, O. (Ed.), *Frontiers and Advances in Positive Learning in the Age of Information*. Springer, Wiesbaden, pp. 21–39. https://doi.org/10.1007/978-3-030-26578-6_3.
- Berman, A.I., Haertel, E.H., Pellegrino, J.W., 2020. Comparability of Large-Scale Educational Assessments: Issues and Recommendations. National Academy of Education.

- Bloom, B., Deyrup, M.M., 2015. The SHU research logs: student online search behaviors trans-scripted. *J. Acad. Librarian* 41 (5), 593–601. <https://doi.org/10.1016/j.acalib.2015.07.002>.
- Bonikowska, A., Sanmartin, C., Frenette, M., 2019. Data Literacy: What it Is and How to Measure it in the Public Service. Statistics Canada. <https://www150.statcan.gc.ca/n1/en/pub/11-633-x/11-633-x2019003-eng.pdf?st=BoMlagz>.
- Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., van Gog, T., 2017. Source evaluation of domain experts and novices during web search. *J. Comput. Assist. Learn.* 33 (3), 234–251. <https://doi.org/10.1111/jcal.12162>.
- Braun, H.I., 2019. Performance assessment and standardization in higher education: a problematic conjunction? *Br. J. Educ. Psychol.* 89 (3), 429–440. <https://doi.org/10.1111/bjep.12274>.
- Braun, H.I., Shavelson, R.J., Zlatkin-Troitschanskaia, O., Borowiec, K., 2020. Performance assessment of critical thinking: conceptualization, design, and implementation. *Front. Educ.* 5, 156. <https://doi.org/10.3389/feduc.2020.00156>.
- Brooks, D.C., 2016. ECAR Study of Undergraduate Students and Information Technology, 2016. EDUCAUSE.
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., Fischer, F., 2020. Simulation-based learning in higher education: a meta-analysis. *Rev. Educ. Res.* 90 (4), 499–541.
- Ciampaglia, G.L., 2018. The digital misinformation pipeline. In: Zlatkin-Troitschanskaia, O., Wittum, G., Dengel, A. (Eds.), *Positive Learning in the Age of Information*. Springer, Wiesbaden, pp. 413–421. https://doi.org/10.1007/978-3-658-19567-0_25.
- Corno, L., Cronbach, L.J., Kupermintz, H., Lohman, D.F., Mandinach, E.B., Porteus, A.W., Talbert, J.E., 2002. *Remaking the Concept of Aptitude: Extending the Legacy of Richard E. Snow*. Routledge.
- Dahlen, S.P., Haeger, H., Hanson, K., Montellano, M., 2020. Almost in the wild: student search behaviors when librarians aren't looking. *J. Acad. Librarian* 46 (1), 102096. <https://doi.org/10.1016/j.acalib.2019.102096>.
- Darling-Hammond, L., Adamson, F., 2014. *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*. Jossey-Bass. <https://doi.org/10.1002/9781119210863>.
- Davey, T., Ferrara, S., Shavelson, R., Holland, P., Webb, N., Wise, L., 2015. Psychometric Considerations for the Next Generation of Performance Assessment. Center for K-12 Assessment & Performance Management, Educational Testing Service, Washington, DC. Retrieved from: https://www.ets.org/Media/Research/pdf/psychometric_considerations_white_paper.pdf.
- Dolan, R., Burling, K., 2018. Computer-based testing in higher education. In: Charles, S., Brian Denison, D. (Eds.), *Handbook on Measurement, Assessment, and Evaluation in Higher Education*, second ed. Routledge, New York, NY, pp. 370–384.
- Drachler, H., Goldhammer, F., 2020. Learning analytics and eAssessment—towards computational psychometrics by combining psychometrics with learning analytics. In: Burgos, D. (Ed.), *Radical Solutions and Learning Analytics. Lecture Notes in Educational Technology*. Springer, Singapore.
- Furi, I., Balog, K.P., 2016. Information seeking behaviour in the digital environment: information science vs. Non-information science students. *Knjižnica* 60 (1), 61–82.
- Gijbels, D., Dochy, F., Van den Bossche, P., Segers, M., 2005. Effects of problem-based learning: a meta-analysis from the angle of assessment. *Rev. Educ. Res.* 75 (1), 27–61.
- Goldman, S.R., Brand-Gruwel, S., 2018. Learning from multiple sources in a digital society. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, New York, NY, pp. 86–95.
- Grossnickle Peterson, E., Alexander, P.A., 2020. Navigating print and digital sources: students' selection, use, and integration of multiple sources across mediums. *J. Exp. Educ.* 88, 27–46. <https://doi.org/10.1080/00220973.2018.1496058>.
- Hahnel, C., Goldhammer, F., Kröhne, U., Naumann, J., 2018. The role of reading skills in the evaluation of online information gathered from search engine environments. *Comput. Hum. Behav.* 78, 223–234.
- Hahnel, C., Schoor, C., Kröhne, U., Goldhammer, F., Mahlow, N., Artelt, C., 2019. The role of cognitive load for university students' comprehension of multiple documents. *Z. Padagog. Psychol.* 33 (2). <https://doi.org/10.1024/1010-0652/a000238>.
- Hargittai, E., Fullerton, L., Menchen-Trevino, E., Yates Thomas, K., 2010. Trust online: young adults' evaluation of web content. *Int. J. Commun.* 4, 468–494, 1932–8036/20100468.
- Holman, L., 2011. Millennial students' mental models of search: implications for academic librarians and database developers. *J. Acad. Librarian* 37 (1), 19–27. <https://doi.org/10.1016/j.acalib.2010.10.003>.
- Horbach, A., Zesch, T., 2019. The influence of variance in learner answers on automatic content scoring. *Front. Educ.* 4, 28. <https://doi.org/10.3389/feduc.2019.00028>.
- International Test Commission (ITC), 2017. *The ITC Guidelines for Translating and Adapting Tests*, second ed. Available from: www.InTestCom.org.
- Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., Drachler, H., 2020. From students with love: an empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *J. Internet High. Educ.* <https://doi.org/10.1016/j.jheduc.2020.100758>.
- Judd, T., Kennedy, G., 2011. Expediency-based practice? Medical students' reliance on Google and Wikipedia for biomedical inquiries. *Br. J. Educ. Technol.* 42 (2), 351–360. <https://doi.org/10.1111/j.1467-8535.2009.01019>.
- Kahneman, D., 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York, NY.
- Kane, M., Mislevy, R., 2017. Validating score interpretations based on response processes. In: Ercikan, K., Pellegrino, J.W. (Eds.), *Validation of Score Meaning for the Next Generation of Assessments*. Routledge, pp. 11–24.
- Kleemola, K., Hytinen, H., Toom, A., 2021. Exploring internal structure of a performance-based critical thinking assessment for new students in higher education. *Assess. Eval. High Educ.* <https://doi.org/10.1080/02602938.2021.1946482>.
- Leeder, C., Shah, C., 2016. Practicing critical evaluation of online sources improves student search behavior. *J. Acad. Librarian* 42 (4), 459–468. <https://doi.org/10.1016/j.acalib.2016.04.001>.
- Leighton, J.P., Cui, Y., Cutumisu, M., 2021. Key information processes for thinking critically in data-rich environments. *Front. Educ.* <https://doi.org/10.3389/feduc.2021.561847>.
- Liu, O.L., Frankel, L., Roehr, K.C., 2014. Assessing critical thinking in higher education: current state and directions for next-generation assessment. *ETS Res. Rep. Ser.* 2014. <https://doi.org/10.1002/ets2.12009>.
- Liu, O.L., Mao, L., Frankel, L., Xu, J., 2016. Assessing critical thinking in higher education: the HEIghten™ approach and preliminary validity evidence. *Assess. Eval. High Educ.* 41 (5).
- Liu, J., Zamir, H., Li, Y., Hastings, S.K., 2018. Search systems and their features: what college students use to find and save information. *Libr. Inf. Sci. Res.* 40 (2), 118–124. <https://doi.org/10.1016/j.lisr.2018.06.002>.
- Machete, P., Turpin, M., 2020. The use of critical thinking to identify fake news: a systematic literature review. In: Hattingh, M., et al. (Eds.), *Responsible Design, Implementation and Use of ICT*. https://doi.org/10.1007/978-3-030-45002-1_20.
- Marks, A., AL-Ali, M., Attasi, R., Abu Elkishk, A., Rezgui, Y., 2021. Digital transformation in higher education: maturity and challenges. In: Rocha, A., Ferrás, C., López-López, P.C., Guarda, T. (Eds.), *Information Technology and Systems, ICITS 2021*, p. 1.
- Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O., Jitomirski, J., 2020. Positive and negative media effects on university students' learning: preliminary findings and a research program. In: Zlatkin-Troitschanskaia, O. (Ed.), *Frontiers and Advances in Positive Learning in the Age of Information*. Springer, New York, NY.
- McClelland, D.C., 1973. Testing for competence rather than for "intelligence". *Am. Psychol.* 28 (1). <https://doi.org/10.1037/h0034092>.
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., Wineburg, S., 2018. Can students evaluate online sources? *Theor. Res. Soc. Educ.* 46 (2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>.
- McGrew, S., Smith, M., Breakstone, J., Ortega, T., Wineburg, S., 2019. Improving university students' web savvy: an intervention study. *Br. J. Educ. Psychol.* 89, 485–500.
- Mislevy, R.J., 2018. *Socio-cognitive Foundations of Educational Measurement*. Routledge, New York, NY.

- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A brief introduction to evidence-centered design. ETS Res. Rep. Ser. 2003. <https://doi.org/10.1002/j.2333-8504.2003.tb01908.x>.
- Mislevy, R.J., Behrens, J.T., DiCerbo, K.E., Levy, R., 2012. Design and discovery in educational assessment: evidence-centered design, psychometrics, and educational data mining. *J. Educ. Data Min.* 4, 11–48. Available from: <https://jedm.educationaldatamining.org/index.php/JEDM/article/view/22>.
- Molero, D., Zlatkin-Troitschanskaia, O., Nagel, M.-T., Brückner, S., Schmidt, S., Shavelson, R.J., 2020. Assessing university students' critical online reasoning ability: a conceptual and assessment framework with preliminary evidence. *Front. Educ.* <https://doi.org/10.3389/educ.2020.577843>.
- Nagel, M.-T., Schäfer, S., Zlatkin-Troitschanskaia, O., Schemer, C., Maurer, M., Molero, D., Schmidt, S., Brückner, S., 2020. How do university students' web search behavior, website characteristics, and the interaction of both influence students' critical online reasoning? *Front. Educ.* <https://doi.org/10.3389/educ.2020.565062>.
- Newman, A., Bavi, Y.L., Mount, M., Shao, B., 2021. Data collection via online platforms: challenges and recommendations for future research. *Appl. Psychol.* 70 (3), 1380–1402. <https://doi.org/10.1111/apps.12302>.
- OECD, 2012. Assessment of Higher Education Learning Outcomes. Feasibility Study Report: Volume 1 – Design and Implementation. Available from: <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>.
- OECD, 2013. In: Assessment of Higher Education Learning Outcomes. AHELO Feasibility Study Report, vol. 2. Paris.
- Oliveri, M.E., Mislevy, R.J., 2019. Introduction to “challenges and opportunities in the design of ‘next-generation assessments of 21st century skills’”. *Int. J. Test.* 19.
- Paul, R., Elder, L., 2005. A Guide for Educators to Critical Thinking Competency Standards, Principles, Performance Indicators, and Outcomes with a Critical Thinking Master Rubric. Foundation for Critical Thinking.
- Pellegrino, J.W., 2002. Understanding how students learn and inferring what they know: implications for the design of curriculum, instruction and assessment. In: NSF K-12 Mathematics and science curriculum and implementation centers conference proceedings. National Science Foundation and American Geological Institute, Washington, DC, pp. 76–92.
- Powers, E.M., 2019. How students access, filter and evaluate digital news: choices that shape what they consume and the implications for news literacy education (doctoral dissertation). *J. Lit. Technol.* 20 (3).
- Quinn, S., Bond, R., Nugent, C., 2017. Quantifying health literacy and eHealth literacy using existing instruments and browser-based software for tracking online health information seeking behavior. *Comput. Hum. Behav.* 69, 256–267. <https://doi.org/10.1016/j.chb.2016.12.032>.
- Raaheim, A., Mathiassen, K., Moen, V., Lona, I., Gynild, V., Ringlund Bunæs, B., Trygve Hasle, E., 2019. Digital assessment – how does it challenge local practices and national law? A Norwegian case study. *Eur. J. High Educ.* 9 (2), 219–231. <https://doi.org/10.1080/21568235.2018.1541420>.
- Raido, V.E., 2013. Investigating the Web Search Behaviors of Translation Students: An Exploratory and Multiple-Case Study (PhD thesis). Universitat Ramon Llull.
- Ronderos, N., Shavelson, R.J., Holtsch, D., Zlatkin-Troitschanskaia, O., Solano-Flores, G., 2021. International performance assessment of critical thinking: framework for translation and adaptation. *J. Supranatl. policies educ.* 13, 62–87.
- Rosman, T., Mayer, A.K., Krampen, G., 2015. Combining self-assessments and achievement tests in information literacy assessment: empirical results and recommendations for practice. *Assess. Eval. High. Educ.* 40 (5), 740–754. <https://doi.org/10.1080/02602938.2014.950554>.
- Rott, S., Garvin, B., 2015. Comprehending and learning from Internet sources: a conceptual replication study of Goldman, Braasch, Wiley, Greaser and Brodowska (2012). *CALICO Journal* 32 (2), 323–354.
- Salehi, S., Du, J.T., Ashman, H., 2018. Use of web search engines and personalisation in information searching for educational purposes. *Inf. Res.* 23 (2).
- Schoor, C., Hahnel, C., Mahlow, N., Klagges, J., Kroehne, U., Goldhammer, F., Artelt, C., 2020. Multiple document comprehension of university students: test development and relations to person and process characteristics. In: Zlatkin-Troitschanskaia, O., Pant, H.A. (Eds.), *Student Learning Outcomes in Higher Education*. Wiesbaden Springer, pp. 223–240.
- Schmidt, S., Zlatkin-Troitschanskaia, O., Roeper, J., Klose, V., Weber, M., Bültmann, A.-K., Brückner, S., 2020. Undergraduate Students' Critical Online Reasoning—Process Mining Analysis. *Front. Psychol.* 11 (576273). <https://doi.org/10.3389/fpsyg.2020.576273>.
- Sendurur, E., 2018. Students as information consumers: a focus on online decision making process. *Educ. Inf. Technol.* 23 (6), 3007–3027. <https://doi.org/10.1007/s10639-018-9756-9>.
- Shahibi, M.S., Rusli, K.N.K.K., 2017. The influence of Internet usage on student's academic performance. *Int. J. Acad. Res. Bus. Soc. Sci.* 7 (8), 873–887.
- Shavelson, R.J., 2010. *Measuring College Learning Responsibly: Accountability in a New Era*. Stanford University Press, Stanford, CA.
- Shavelson, R.J., 2013. On an approach to testing and modeling competence. *Educ. Psychol.* 48 (2), 73–86.
- Shavelson, R.J., Roeser, R.W., Kupermintz, H., Lau, S., Ayala, C., Haydel, A., Schultz, S., Quihuis, G., Gallagher, L., 2002. Richard E. Snow's remaking of the concept of aptitude and multidimensional test validity: introduction to the special issue. *Educ. Assess.* 8 (2), 77–100.
- Shavelson, R.J., Zlatkin-Troitschanskaia, O., Marino, J.P., 2018. International performance assessment of learning in higher education (IPAL): research and development. In: Zlatkin-Troitschanskaia, O., Toepper, M., Pant, H.A., Lautenbach, C., Kuhn, C. (Eds.), *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*. Springer, New York, NY.
- Shavelson, R.J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., Marino, J.P., 2019. Assessment of university students' critical thinking: next generation performance assessment. *Int. J. Test.* 19 (4), 337–362. <https://doi.org/10.1080/15305058.2018.1543309>.
- Snow, R.E., 1996. Aptitude development and education. *Psychol. Publ. Pol. Law* 2 (3/4), 536–560.
- Sparks, J.R., Katz, I.R., Belle, P.M., 2016. Assessing digital information literacy in higher education: a review of existing frameworks and assessments with recommendations for next-generation assessment. ETS Res. Rep. Ser. <https://doi.org/10.1002/ets2.12118>.
- Stanovich, K.E., 2009. What Intelligence Test Miss: The Psychology of Rational Thought.
- Stanovich, K.E., 2016. The comprehensive assessment of rational thinking. *Educ. Psychol.* 1–12. <https://doi.org/10.1080/00461520.2015.1125787>.
- Strobel, J., Van Barneveld, A., 2009. When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdiscip. J. Probl. Based Learn.* 3, 44–58.
- Suckow, D.W., Braher, C.J., Donahoe-Fillmore, B., Fisher, M.I., Anloague, P.A., 2015. The Association between Critical Thinking and Scholastic Aptitude on First-Time Pass Rate of the National Physical Therapy Examination. Physical Therapy Faculty Publications, p. 27.
- Suter, L.E., 2019. Growth and development of large-scale international comparative studies and their influence on comparative education thinking. In: Suter, L.E., Smith, E., Denman, B.D. (Eds.), *The SAGE Handbook of Comparative Studies in Education*. Sage, London.
- Taylor, J., 2015. An examination of how student journalists seek information and evaluate online sources during the newsgathering process. *New Media Soc.* 17 (8), 1277–1298. <https://doi.org/10.1177/1461444814523079>.
- Wahlen, A., Kuhn, C., Zlatkin-Troitschanskaia, O., Gold, C., Zesch, T., Horbach, A., 2020. Automated scoring of teachers' pedagogical content knowledge – a comparison between human and machine scoring. *Front. Educ.* 5, 149. <https://doi.org/10.3389/educ.2020.00149>.
- Watson, G., Glaser, E., 2002. *Critical Thinking Appraisal*. Pearson Assessment, London.
- Weber, H., Becker, D., Hillmert, S., 2018a. Information-seeking behavior and academic success in higher education: which search strategies matter for grade differences among university students and how does this relevance differ by field of study? *Int. J. High. Educ. Res.* 77 (4), 657–678. <https://doi.org/10.1007/s10734-018-0296-4>.
- Weber, H., Hillmert, S., Rott, K.J., 2018b. Can digital information literacy among undergraduates be improved? Evidence from an experimental study. *Teach. High. Educ.* 23 (8), 909–926. <https://doi.org/10.1080/13562517.2018.1449740>.
- Weinert, F.E., 2001. Concept of competence: a conceptual clarification. In: Rychen, D.S., Salganik, L.H. (Eds.), *Defining and Selecting Key Competencies*. Hogrefe & Huber, Seattle, pp. 45–65.
- Wind, S.A., Wolfe, E.W., Engelhard, G.J., Foltz, P., Rosenstein, M., 2018. The influence of rater effects in training sets on the psychometric quality of automated scoring for writing assessments. *Int. J. Test.* 18, 27–49. <https://doi.org/10.1080/15305058.2017.1361426>.

- Wineburg, S., Breakstone, J., McGrew, S., Ortega, T., 2016. Evaluating Information: The Cornerstone of Civic Online Reasoning. Stanford History Education Group. Available from: <https://stacks.stanford.edu/file/druid:tv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf>.
- Wolf, R., Zahner, D., Benjamin, R., 2015. Methodological challenges in international comparative post-secondary assessment programs: lessons learned and the road ahead. *Stud. High. Educ.* 40 (3).
- Zahner, D., Ciolfi, A., 2018. International comparison of a performance-based assessment in higher education. In: Zlatkin-Troitschanskaia, O., Toepper, M., Pant, H.A., Lautenbach, C., Kuhn, C. (Eds.), *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*. Springer, New York, NY, pp. 215–244.
- Zlatkin-Troitschanskaia, O., Shavelson, R.J., 2019. Advantages and challenges of performance assessment of student learning in higher education. *Br. J. Educ. Psychol.* 89 (3).
- Zlatkin-Troitschanskaia, O., Toepper, M., Molerov, D., Buske, R., Brückner, S., Pant, H.A., Hofmann, S., Hansen-Schirra, S., 2018. Adapting and validating the collegiate learning assessment to measure generic academic skills of students in Germany – implications for international assessment studies in higher education. In: Zlatkin-Troitschanskaia, O., Toepper, M., Pant, H.A., Lautenbach, C., Kuhn, C. (Eds.), *Assessment of Learning Outcomes in Higher Education: Cross-National Comparisons and Perspectives*. Springer, New York, NY, pp. 245–266.
- Zlatkin-Troitschanskaia, O., Shavelson, R.J., Schmidt, S., Beck, K., 2019. On the complementarity of holistic and analytic approaches to performance assessment scoring. *Br. J. Educ. Psychol.* <https://doi.org/10.1111/bjep.12286>.
- Zlatkin-Troitschanskaia, O., Beck, K., Fischer, J., Braunheim, D., Schmidt, S., Shavelson, R.J., 2020. The role of students' beliefs when critically reasoning from multiple contradictory sources of information in performance assessments. *Front. Psychol.* 11, 2192. <https://doi.org/10.3389/fpsyg.2020.02192>.
- Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., Krstev, J., 2021. Students' online information use and learning progress in higher education – a critical literature review. *Stud. High. Educ.* 1–26.
- Zumbo, B.D., Hubley, A.M. (Eds.), 2017. *Understanding and Investigating Response Processes in Validation Research*, vol 26. Springer International Publishing, Cham.

Appendix A

Table A.1 Selected studies.

<i>Number, authors, title, year</i>	<i>Focus of study</i>	<i>Method</i>	<i>Sample</i>	<i>Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)</i>	<i>Study design</i>	<i>Scope</i>
1. Bloom and Deyrup (2015)	Online research behaviors as well as emotive and affective responses	Task: Self-selected online research tasks by using the software (OpenHallway) Screen casting and voice recording while conducting online class research Think aloud Post-questionnaire on students' perceptions of research history and level of success, feelings about and experience doing scholarly research	42 undergraduate students (sophomores, juniors, and seniors) enrolled in courses that required a sizable research project	QUAL: Coding of student responses (coding schema with eight categories) QUAN	Cross-section analysis	United States
2. Brand-Gruwel et al. (2017)	Evaluating and selecting online sources	Tasks: two information problem-solving tasks; selecting results best suited with a Google-like-SERP consisting of a selected mix of existing websites as a basis for writing an article (offline setting) Thinking aloud Cued retrospective reporting instruction Eye tracking Log-file analysis Questionnaire on the use and functioning of search engines Written free recall task on prior topic knowledge	19 domain novices (15 first semester Psychology students; 4 prospective students) and 16 domain experts (university Psychology teachers)	QUAL: Scoring task performance (coding scheme for evaluation search results, websites and information) QUAN: eye-movement data analysis, log-file analysis	Cross-section analysis	Netherlands

Table A.1 Selected studies.—cont'd

<i>Number, authors, title, year</i>	<i>Focus of study</i>	<i>Method</i>	<i>Sample</i>	<i>Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)</i>	<i>Study design</i>	<i>Scope</i>
3. Dahlen et al. (2020)	Research using two categories of search tools Use of search features; search facets	Tasks: two search tasks; both on finding two suitable articles on a topic for writing a paper. Useable search tools: EBSCO's Social Sciences Abstracts and ProQuest's Summon Screencasting Semi-structured interviews (ex post)	50 juniors and seniors majoring in Social and Behavioral Sciences 26% upper-level majors	QUAL: Coding student search behavior and use of search features (i.e., discovery systems and traditional library databases); Scoring articles used by participants analyzing student interviews (impressions of search tool, reason for selecting articles, and level of extent participant considered publication date, credibility, and relevance) QUAN: time needed, number of clicks for task completion, number of search steps, sequence of steps	Cross-section analysis	United States
4. Furi and Balog (2016)	Actual information-seeking behavior	Tasks: three simulated search tasks to be solved by searching the Internet, the second by some of the scholarly databases, and the third by using the library's OPAC Think-aloud Video recording Questionnaire on demographics, students' information literacy, educational background, and self-perception of their information seeking expertize	6 social science graduate students ($n = 3$ Information Science, $n = 1$ Psychology, $n = 1$ Economics, $n = 1$ Cultural Management)	QUAL: sources used, emotions during task, success at completion of task QUAN: time needed, number of clicks for task completion, number of search steps, sequence of steps	Cross-section analysis	Croatia
5. Hahnel et al. (2018)	Multiple document comprehension	Task: Computer-based multiple document comprehension test - Simulated search task with corresponding search engine result page(s) (SERPs) Log-file analysis Questionnaire on demographics	310 German university students enrolled in several bachelor's or master's programs in the Social Sciences and Humanities 70% bachelor and 30% master students	QUAL: Scoring students' selection of the most informative and credible SERP link (dichotomous scoring scheme) QUAN: generalized linear mixed models used for statistical analyses carried out in R	Cross-section analysis	Germany
6. Hargittai et al. (2010)	Finding and evaluating Web content	Task: 12 open web searches, ranging from locating information about current events to emailing House Judiciary Committee, accessing literature, finding information on	Observation and interview: – 102 first-year students from a first-year writing program at an urban public research university	QUAL: Coding data obtained via audio recording QUAN: Respondents' Internet use experiences and questions about the credibility (using a Likert scale)	Cross-section analysis	United States

(Continued)

Table A.1 Selected studies.—cont'd

<i>Number, authors, title, year</i>	<i>Focus of study</i>	<i>Method</i>	<i>Sample</i>	<i>Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)</i>	<i>Study design</i>	<i>Scope</i>
		health matters, finding best practices in poster layout and information on Microsoft Word settings, assessing person's gender based on first name, find authoritative source for writing resume Think aloud and Screencasting Paper-pencil questionnaire on students' Internet use and assessment of online materials Follow-up in-person observations and interviews on students' level of understanding of 27 Internet-related terms	– 23% Health Sciences, 15% Science, 13% Business, 10% Social Science, 11% Math/Computer Science/ Engineering, 8% Nursing, 7% Arts/ Humanities, 7% Architecture/Design, 5% Education			
7. Holman (2011)	Information-seeking behavior and search strategies Choice of search tools and search engine queries	Task: Conducting research in chosen search engines or article databases in a university usability lab for about 1 h over a 3-week period Think aloud Video screening Semi-structured interview (ex post) Students' visual diagram of the relationship between their search terms and the items retrieved by the search tool	21 first-year undergraduate students	QUAL: Analyzing think aloud protocols, video recordings & captured searches (use of search engines and databases, use of Boolean operators, use of punctuation, use of keywords and synonyms)	Cross-section analysis	United States
8. Judd and Kennedy (2011)	Locating and accessing resources+	Students' sessional Internet usage was collected over a 4-week period in a large open-access computing laboratory Log-file analysis Self-developed usage and evaluation questionnaire (5-point Likert scales)	842 undergraduate (1st-, 2nd- and 3rd-year) medical students	QUAL: Rating of reliability of sources located and accessed QUAN: students' perception of usefulness and reliability of five targeted sites chosen by researchers (Likert scale)	Cross-section analysis	United States
9. Leeder and Shah (2016)	Online search behavior and results Evaluating online sources and information	Task and topic: Searching online for information; Evaluating the quality of 15 given URLs of	32 undergraduate university students from several academic departments 1 freshman, 11	QUAL: Coding productivity along three characteristics (goodness of queries, goodness of the found	Cross-section analysis with treatment and control condition	United States

Table A.1 Selected studies.—cont'd

<i>Number, authors, title, year</i>	<i>Focus of study</i>	<i>Method</i>	<i>Sample</i>	<i>Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)</i>	<i>Study design</i>	<i>Scope</i>
		online sources before conducting own research using provided criteria (treatment condition); Reviewing 15 given URLs of online sources (without specific direction or prompts; control condition) Log file analysis	sophomores, 10 juniors and 10 seniors	information, overall productivity); analyzing of open-ended text responses (questions about search process and further strategies in continuing search) QUAN: descriptive statistics (total number of pages viewed, unique pages viewed, pages bookmarked, distinct queries and search engine result pages (SERPs), most frequent words used in search process (shared, control and treatment condition))		
10. Liu et al. (2018)	Finding and saving online information <i>Topic:</i> University-course-related work	2 self-identified search tasks (recall of a recently finished real task and completion of a self-identified task) Pre- and post-questionnaire on interest, topic knowledge, perceived difficulty, and helpful system features	32 college students from randomly selected academic departments, such as Accounting, Biomedical Engineering, Chemistry, English, Journalism, Social Work, and Media Arts 2 freshmen, 4 sophomores, 8 juniors, 9 seniors and 9 graduate students	QUAN: use of sources and corresponding task type	Cross-section analysis	United States
11. Nagel et al. (2020)	Critical online reasoning Web search behavior Quality of websites	Tasks: 5 search/evaluation tasks with written responses in an open-ended format; participants must judge (a) whether a given website or tweet is a reliable source of information (b) whether a given claim is true or untrue by performing a time-restricted open web search Log-file analysis Socio-demographic data and media use behavior via standardized questionnaire	45 economics students Majority at the beginning of their studies	QUAL: Rating URLs accessed by participants, Scoring written responses QUAN: Analyzing number of URLs and subpages of URLs	Cross-section analysis	Germany

(Continued)

Table A.1 Selected studies.—cont'd

Number, authors, title, year	Focus of study	Method	Sample	Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)	Study design	Scope
12. Powers (2019)	Accessing, filtering and evaluating online news	Search tasks: two tasks; Finding credible news (text, audio, or video) on a topic of interest by starting at the website most often/ second-most often used when beginning a news search in 5 min Concurrent think-aloud protocols Screencasting Semi-structured interview	37 undergraduate college students with a range of academic backgrounds Most often underclassmen Merely 8.6% majoring in Journalism	QUAL: Coding think- aloud and subsequent semi-structured interviews (original coding scheme)	Cross-section analysis	United States
13. Quinn et al. (2017)	Online health information seeking behavior	Tasks: six search tasks asking to answer specific health questions by using the Internet (if needed, answers without conducting online search were permitted) Health literacy instrument eHealth literacy instrument Screencasting Log-file analysis	54 undergraduate and postgraduate university students and university staff The majority (62.9%) had achieved an undergraduate or postgraduate degree Mostly Computer Science majors (70.3%)	QUAN: bivariate analysis; test score analysis; log file analysis (search queries entered, web pages visited and duration of each visit)	Cross-section analysis	United Kingdom
14. Raido (2013)	Translation and online search behavior	Search task: Students using the Web as an external resource of consultation for translation problem solving; Translating four different source texts Written "Online search report" Semi-structured interview (ex post)	4 translation students (BA in Spanish (major) and French or Latin American Studies (minor)) 2 translators with different levels of professional and teaching experience	QUAL: Coding data (type of search needs, search goals, initial search actions, information sources, among others) QUAN: descriptive statistics to provide overview of participants' Web search behavior	Cross-section analysis	New Zealand
15. Rosman et al. (2015)	Finding and selecting online sources	Task: Information search tasks consisted of 3 increasingly difficult tasks (e.g., Finding two longitudinal studies) accompanied by question about participants' course of action while searching (e.g., Where did you search, and which keywords did you use?) Self-efficacy scale for information searching	81 undergraduate (67%) and master (33%) psychology students	QUAN: Analyzing intercorrelations of study variables (i.e., information literacy via Self-Efficacy Scale for Information Search Behavior (SES-IB-16)	Inter-vention study	Germany

Table A.1 Selected studies.—cont'd

Number, authors, title, year	Focus of study	Method	Sample	Method of evaluationQualitative data analysis (QUAL);Quantitative data analysis (QUAN)	Study design	Scope
		behavior (5-point Likert scale) Post-questionnaire on students' course of action while searching Information literacy test (situational judgment (or scenario-based) test format; 5-point Likert scale)				
16. Salehi et al. (2018)	Information-seeking behavior and preferences for the purpose of completing academic tasks	Study 1: Questionnaire 5-point Likert scales Study 2: Experimental design with five Web search tasks (switching between Google personalized and non- personalized searches; Reviewing the entire first search engine result page; reviewing the most interesting and relevant results) Screencasting	Study 1: – 120 university students from a variety of subjects and attendance statuses – 58% undergraduate students – 42% research and coursework postgraduate students Study 2: – 28 university students from the Schools of Education and Information Technology – 46% undergraduate students – 54% postgraduate students	QUAL: Analyzing observational notes from web search tasks, Analyzing screen casting results QUAN: Analyzing self- reports on students' information-seeking behavior (5-point Likert scale)	Cross-section analysis (with treatment and control condition)	Australia
17. Sendurur (2018)	Finding and selecting online information	2 paper-based scenarios and tasks asking to search the web as in daily life (Finding the best three accommodation options; Finding the three best resources for a topic) Think aloud Eye-tracking	15 university students from a variety of departments 11 undergraduate and 4 graduate students	QUAL: Content analysis, coding data from think-aloud and explanation QUAN: Descriptive statistics, eye movement analysis → Mixed methods approach	Cross-section analysis	Turkey
18. Tylor (2015)	Online newsgathering process Evaluating the credibility of websites	Study 1: Quasi- experimental design (searching) Video recording while conducting online research for a specific hard-news story topic that students had been assigned to work on for class Study 2: Subsequent semi-structured phone interviews	38 Journalism students Mostly sophomores or juniors	QUAL: Analyzing of self- reported behavior from semi-structured interviews; analyzing the web search behavior (e.g., type of web source, types of search engines used, hierarchy of search engine results clicked, type of search query, contact page visits) QUAN: Coding observed behavior from video recordings	Cross-section analysis	United States

(Continued)

Table A.1 Selected studies.—cont'd

Number, authors, title, year	Focus of study	Method	Sample	Method of evaluation	Study design	Scope
				Qualitative data analysis (QUAL);Quantitative data analysis (QUAN)		
19. Weber et al. (2018b)	Testing the effect of an information literacy workshop Students' search strategies (usage of academic databases, sophistication of search queries)	Tasks: Three search tasks (Using the Internet to search for sources on a topic and name three sources you would cite in your thesis) Log-file analysis	38 third-year students of Sociology or Educational Science who were about to write their final bachelor's thesis	QUAL: Categorizing and Rating data from search queries (academic reputation and relevance of content)	Inter-vention study with treatment and control condition	Germany

Adoption of learning analytics

Dirk Ifenthaler, University of Mannheim, Mannheim, Germany; and Curtin University, Perth, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	331
Overview on learning analytics	331
Adoption of learning analytics	332
Change management	333
Conclusions	334
References	334

Introduction

The use of static and dynamic information about learners, teachers, and learning environments for real-time modeling, prediction, and optimization of learning processes, learning environments, as well as educational decision-making is often referred to as learning analytics (Ifenthaler, 2015). Hence, learning analytics draw on an eclectic set of data and methodologies to provide summative, real-time, and predictive insights for improving learning, teaching, and organizational efficiency (Lockyer et al., 2013; Long and Siemens, 2011). While the field of learning analytics is receiving a lot of attention for its capacity to provide lead indicators of student failure (Sclater and Mullan, 2017) or student success (Ifenthaler et al., 2019), it has focused to date on individual courses in isolation, rather than the capabilities of educational organizations in general (Gašević et al., 2015). However, the implementation of learning analytics at educational organizations may have broad implications for the organization (e.g., technological infrastructure, policies and regulations) and its stakeholders (e.g., students, academic staff, administrators) including changes in learning culture (Ifenthaler, 2020).

The adoption of learning analytics within an educational organization is depending on several factors including technological systems, regulations and guidelines, as well as professional development of involved stakeholders (Ifenthaler, 2017). Learning analytics ready technological systems are required to process vast amounts of available educational multimodal data, flexible data mining tools, and advanced statistical methods including machine learning algorithms (Berland et al., 2014). Learning analytics ready regulations observe data privacy legislation and ethical consideration of (semi-)automated recommendations (Prinsloo and Slade, 2014). Learning analytics ready stakeholders demonstrate educational data literacy, understood as the ethically responsible collection, management, analysis, comprehension, interpretation, and application of data from educational contexts.

Accordingly, the degree of adoption of learning analytics within an educational organization is a measure of the number of technological systems and organizational regulations implemented as well as the number of stakeholders using learning analytics or who have altered their practice because of learning analytics. This article provides an overview on learning analytics and its benefits for educational organizations. It further reflects on the factors which influence the adoption of learning analytics including change management processes.

Overview on learning analytics

There are several concepts closely linked to processing, analyzing and visualizing educational information. However, these concepts are often confused and lack universally agreed as well as applied definitions. Educational data mining refers to the process of extracting useful information out of a large collection of complex educational datasets (Berland et al., 2014). Academic analytics are the identification of meaningful patterns in educational data in order to inform academic issues (e.g., retention, success rates) and produce actionable strategies (e.g., budgeting, human resources) (Long and Siemens, 2011). Teaching analytics focus on data from classroom management and interactions between teachers and students (Sergis and Sampson, 2017). School analytics capture organizational data including information about management, decision-making and staff capabilities (Sergis and Sampson, 2016). Assessment or measurement analytics emphasize the importance of the diagnostic opportunities of educational data. Hence, the concept of learning analytics has been used in various contexts and with different perspectives, resulting in a lack of clarity and precise definition (Wong et al., 2019).

Current learning analytics research and practice are spearheaded in Australia, USA, and United Kingdom confirming the added value for addressing issues related to study success and identification of students-at-risk (Sclater and Mullan, 2017) as well as for monitoring and improving organizational capabilities (Ifenthaler et al., 2019). Wong (2017) presents several case studies utilizing learning analytics for (a) improving student retention, (b) supporting informed decision making, (c) increasing cost-effectiveness, (d) helping to understand learning behavior, (e) providing personalized assistance, and (f) delivering feedback and interventions. Another work by Ifenthaler et al. (2019) focusses on utilizing learning analytics for study success which has been a global concern for many years (Tinto, 2005). Further research has concentrated on ethics and privacy issues linked to learning analytics, discussing

ethical dilemmas of intervention strategies (West et al., 2016), the ethics of categorizing at-risk students (Scholes, 2016), student perceptions and expectations of data privacy (Ifenthaler and Schumacher, 2016), and ethical oversight of student data (Willis et al., 2016). A generally accepted finding includes the requirement of defining who has access to which data, where and how long the data will be stored, and which procedures and algorithms to implement for further use of the available data (Ifenthaler and Schumacher, 2019; Pardo and Siemens, 2014). Dashboards and visualizations have been identified as another major focus area in learning analytics research (Sahin and Ifenthaler, 2021). Hence, the still emerging field of learning analytics has introduced new frameworks, methodological approaches, and empirical investigations into educational research; for example, novel methods in educational research including machine learning, network analyses, and empirical approaches based on computational modeling experiments (Chatti and Muslim, 2019; Dyckhoff et al., 2012; Gašević et al., 2019a).

From a holistic perspective, educational organizations and involved stakeholders can derive multiple benefits from learning analytics by using different data analytics strategies to produce summative, real-time and predictive insights and recommendations which can be associated with four levels of stakeholders: mega-level (governance), macro-level (organization), meso-level (curriculum, learning design, teacher/tutor), and micro-level (learner) (Ifenthaler, 2015). The *mega-level* facilitates cross-organizational analytics by incorporating data from all levels of learning analytics initiatives. Based on common data standards and open data schemas, rich datasets may enable the identification and validation of patterns within and across organizations and therefore provide valuable insights for informing educational policy making. The *macro-level* enables organization-wide analytics for better understanding learner cohorts to optimize associated processes and allocate critical resources for reducing dropout or increasing retention as well as success rates. The aforementioned levels can be associated with the emerging domain of Educational Leadership Data Analytics (ELDA) which is centered at the intersection of education leadership and education data analytics (Bowers et al., 2019). The *meso-* and *micro-level* provide analytics insights within entities of the organization (Ifenthaler and Widanapathirana, 2014). At the meso-level, learning designers may use analytics insights in order to improve learning artifacts or the sequencing of learning tasks as well as overall curricular design (Ifenthaler et al., 2018). The micro-level focusses on learning-processes of individual students adapting to their current needs in the learning process and helping them to reflect and act on recommendations based on multiple analytics insights (Gašević et al., 2017).

An essential prerequisite of learning analytics benefits, however, is the perspective of data and analytics involved: (1) summative and descriptive; (2) formative and (near) real-time; (3) predictive and prescriptive. The *summative and descriptive* perspective provides detailed insights after completion of a learning phase (e.g., study period, semester, final degree), often compared against previously defined reference points or benchmarks. The *formative and (near) real-time* perspective uses ongoing information for improving processes through direct interventions on-the-fly. The *predictive and prescriptive perspective* is applied for forecasting the probability of outcomes in order to plan for future interventions, strategies, and actions.

In order to fully develop the multiple benefits of learning analytics, a better understanding about reliable and valid indicators for supporting learning analytics is needed. Yau and Ifenthaler (2020) reflect on the effectiveness of learning analytics indicators. In summary, student study history such as GPA or evaluating the learners' learning progress from assessment results are the most successful key indicators. However, for entry-level courses or first-year students, historic academic performance or study history may be unavailable (Mah and Ifenthaler, 2017, 2020). Hence, quiz results in the first week can be used to obtain this information to make the appropriate analysis. In MOOCs, clickstream data seem to be the key indicators, as performance and demographic data may not be available due to limited data collection and privacy issues (Hlosta et al., 2017). Regarding the distribution of indicators mapped on three data profiles (student, learning, curriculum) (Ifenthaler and Widanapathirana, 2014) it becomes apparent that indicators related to the curriculum profile are underrepresented. As such curriculum related indicators may function as benchmarks for formative feedback and (near) real-time scaffolds as well as for improving the learning design, it is suggested to further investigate the benefits and usability as well as the validity of such indicators (Yau and Ifenthaler, 2020).

Adoption of learning analytics

Since a decade, the potentials of learning analytics for educational organizations have been discussed (Pistilli and Arnold, 2010), however, there remains a lack of organization-wide examples demonstrating a systematic and holistic adoption of learning analytics (Gašević et al., 2019b).

In a recent systematic review including over 6000 initial studies on learning analytics, Ifenthaler and Yau (2020) indicate that a wider adoption of organization-wide learning analytics systems is needed. While standards for data models and data collection, such as xAPI (Experience API), exist (Kevan and Ryan, 2016), learning analytics research and development needs to clearly define standards for reliable and valid indicators, informative visualizations, and design guidelines for pedagogically effective learning analytics interventions (Seufert et al., 2019; Yau and Ifenthaler, 2020).

Accordingly, the adoption of learning analytics within an organization involves an organizational diffusion process including communication channels between potential actors, who then follow a well-known decision-making process (i.e., moving from unawareness to knowledge, then being persuaded that learning more will be beneficial, then trying out the change, then moving into routine implementation, and finally creating confirmatory data and helping others) as the stakeholders individually adopt or ignore the innovation of learning analytics (Rogers, 1962; Kotter, 2007). Influencers who wish to promote the adoption of learning analytics must ensure and secure open communication channels and encourage people to start and stay on the journey of decision-making from awareness to action (Drachler and Greller, 2016; Colvin et al., 2015; Ifenthaler and Gibson, 2020).

However, an international study on the adoption of learning analytics revealed that there is a lack of staff available for learning analytics projects (Ifenthaler, 2017). Specialized staff with a strong background in learning and teaching as well as data science are in short supply. For example, only 25% of the organizations included in the study report that they employ learning analytics specialists. Similarly, the findings clearly indicate that educational organizations do not have the necessary technology available to implement valid learning analytics frameworks. For example, only 19% of the organizations report that they have implemented a working data warehouse. Accordingly, only a small number of educational organizations meet the high staff and technology requirements for learning analytics frameworks (Kevan and Ryan, 2016).

Nonetheless, more recent research focusing on the adoption of learning analytics reports high interest among higher education organizations (Ifenthaler and Yau, 2019; Tsai and Gasević, 2017). On the same time, the maturity level of available organization-wide learning analytics systems is still low, i.e., organizations are aware of learning analytics and start experimenting with dashboards for students and teachers, however, they are far from organizational transformation (Siemens et al., 2014; Viberg et al., 2018). Accordingly, learning analytics remains an interest for educational organizations rather than a major priority (Lester et al., 2017; Ifenthaler and Yau, 2019).

Therefore, actionable frameworks and adoption models focusing on learning analytics are required for successful integration of learning analytics systems into educational organizations (Buckingham Shum and McKay, 2018; Dyckhoff et al., 2012). However, these models may vary across different organizations due to contextual idiosyncrasies as well as in different countries due to cultural differences (Klasen and Ifenthaler, 2019; Schumacher et al., 2019). The current state of adoption and implementation of learning analytics in educational organizations may be a paradoxical exercise. On the one hand, world-class experts in data science, information systems, management, educational leadership and learning science may be part of the research organization. They all could contribute to relevant developments of robust learning analytics systems. On the other hand, these experts may only have loose insights into the requirements of learning analytics when it comes to questions of implementation or the administration is not interested in their expertise for organizational development of the institution. Further, as bureaucracy takes over, these experts may not be interested in facilitating change processes at their home organization. In order to overcome such organizational barriers, Buckingham Shum and McKay (2018) suggest organizational architectures including pros and cons: (a) The IT Service Center Model assumes that the learning analytics system is developed and implemented through IT services of the organization. The pros of this model are clearly on the technical side, as all systems (e.g., learning management system, data warehouse, student information system, etc.) are managed by this unit. Faculty may not have the opportunity to guide the development and implementation of the learning analytics system. In addition, an evidence-centered approach driven by academic faculty is unlikely. (b) The Faculty Academics Model assumes that academics conduct research and drive the implementation of a learning analytics system. Through evidence-centered research and development, immediate and robust findings are produced with regard to acceptance, effectiveness, and challenges involved when implementing learning analytics systems. (c) The Innovation Center Model assumes an autonomous entity of the educational organization which is well connected to faculty and administrative services. Partnerships among all stakeholders are vital for the success of this model. When overall accepted, this model combines the strengths of the aforementioned models (Buckingham Shum and McKay, 2018).

Change management

Alongside the evolving social and organizational context, educational organizations must transform its processes to accommodate new conceptualizations of student capability and success (Vey et al., 2017). In particular, higher education institutions must reconsider five important dimensions of learning and teaching operations (Henry et al., 2020): (1) finding and selecting students; (2) knowing learners and their expectations; (3) just-in-time services, content, mentoring and support; (4) anytime, anywhere accessibility; and (5) global connectivity. These operation can only be reached through implementing change management processes which involves all stakeholders and considers organizational, technological, and staff readiness for change (Kotter, 2007).

While organizations choose to invest into advancing learning analytics capabilities, they will encounter several adoption challenges (Drachsler and Greller, 2016; Leitner et al., 2019; Tsai and Gasević, 2017). Accordingly, a comprehensive change management strategy for implementing actionable learning analytics seems to be pivotal. As a first step, a readiness assessment is conducted while referring to the above mentioned benefits of learning analytics introduced by Ifenthaler (2015). Based on the initial step of identifying the to be achieved benefits of learning analytics, the educational organization needs to carefully build a strategy based on an in-depth review of existing practices, procedures, and capabilities. Hence, the strategy includes decisions which benefits and specific features of learning analytics to be included as well as which infrastructure for a successful learning analytics implementation is required. The second step includes the readiness assessment using standardized instruments. Data for the readiness assessment is collected on organizational level, i.e., organizational readiness (e.g., existing policy, data protection regulation), with a specific focus on necessary technology, i.e., technical readiness (e.g., data warehouse, system integration), and with regard to capabilities staff members of the education institutions possess, i.e., staff readiness (e.g., educational data literacy) (Schumacher et al., 2019; Ifenthaler, 2017).

Based on the results of the in-depth readiness assessment, the resulting implementation strategy covers all areas of change. Further, the stepwise implementation of learning analytics needs to be monitored and evaluated with regard to predefined and measurable KPIs (Key Performance Indicators). More importantly, the return on investment, defined as the expected gains (returns) per unit of cost (investment) of a variety of potential actions need to be monitored closely (Psacharopoulos, 2014). The return on

investment may not only be conceptualized as monetary returns, but may also be returns conceptualized as retaining students, or improving staff and student satisfaction (Gibson et al., 2018). Clearly, the change management process of an educational organization is a time-consuming process and requires acceptance among all involved stakeholders. A change management strategy for implementing actionable learning analytics in educational organizations could be guided by the following principles (Leitner et al., 2019; Tsai and Gasević, 2017; Ifenthaler, 2020):

- Definition of the learning analytics vision and objectives (e.g., using the above mentioned benefits matrix) and align them with the organizations mission and learning culture;
- Identification of organizational, political, or technological factors that will affect the implementation;
- Involvement and continuous information of all stakeholders including students, teachers, administrators, etc.;
- Development of (and continuously update) a strategic plan focusing on short-term and long-term wins, including a needs and risk analysis as well as a clear timeline outlining responsibilities of involved stakeholders;
- Allocation of resources and identification of expertise (inside and outside of the organization) for achieving the learning analytics objectives;
- Undertake a robust formative and summative evaluation of the learning analytics initiative to further refine the overall implementation and organizational change process.

Conclusions

The field of learning analytics is generating growing interest in data and computer science as well as educational science, hence, becoming an important aspect of modern digital learning environments (Ifenthaler and Widanapathirana, 2014). Despite the high interest, the adoption of learning analytics at educational organizations requires capabilities not yet fully developed (Ifenthaler, 2017). For example, stakeholders seem to be overwhelmed with available data and cannot make use of it for supporting learning and teaching as well as educational decision-making (Papamitsiou et al., 2021).

International perspectives on adoption models (Nouri et al., 2019; Ifenthaler and Gibson, 2020) as well as on policy recommendations (Ifenthaler and Yau, 2019) may help to move the pioneering efforts on learning analytics forward. An international Delphi study identified four themes to be addressed by researcher, policymaker and practitioner dialogs (Ifenthaler et al., 2021): (1) Development of data literacy in all stakeholders, (2) Updating of guiding principles and policies of educational data, (3) Standards needed for ethical practices with data quality assurance, and (4) Flexible user-centered design for a variety of users of analytics, starting with learners and ensuring that learners and learning is not harmed.

At the same time, on-going rigorous research is needed focusing on the effectiveness of learning analytics systems, successful strategies of adopting learning analytics, and creating multi-institutional perspectives and international comparisons on the implementation of learning analytics systems.

To conclude, more educational data does not always mean better educational data. Learning analytics have obvious limitations, and data collected from personal and educational sources (can) have multiple meanings. Identifying the validity of learning analytics adoption processes and confirming sustainable benefits for learning-processes and improving learning environments will further encourage the adoption of learning analytics.

References

- Berland, M., Baker, R.S., Bilkstein, P., 2014. Educational data mining and learning analytics: applications to constructionist research. *Technol. Knowl. Learn.* 19, 205–220.
- Bowers, A.J., Bang, A., Pan, Y., Graves, K.E., 2019. Education Leadership Data Analytics (ELDA): A White Paper Report on the 2018 ELDA Summit. Teachers College, Columbia University, New York, NY.
- Buckingham Shum, S., McKay, T.A., 2018. Architecting for learning analytics. *Innovating for sustainable impact. Educ. Rev.* 53, 25–37.
- Chatti, M.A., Muslim, A., 2019. The PERLA framework: blending personalization and learning analytics. *Int. Rev. Res. Open Dist. Learn.* 20.
- Colvin, C., Rodgers, T., Wade, A., Dawson, S., Gasevic, D., Buckingham Shum, S., Nelson, K.J., Alexander, S., Lockyer, L., Kennedy, G., Corrin, L., Fisher, J., 2015. Student Retention and Learning Analytics: A Snapshot of Australian Practices and a Framework for Advancement. Australian Government Office for Learning and Teaching, Canberra, Act.
- Drachler, H., Greller, W., 2016. Privacy and Analytics—It's a DELICATE Issue. A Checklist for Trusted Learning Analytics. Sixth International Conference on Learning Analytics & Knowledge. Edinburgh, UK.
- Dyckhoff, A.L., Zielke, D., Bültmann, M., Chatti, M.A., Schroeder, U., 2012. Design and implementation of a learning analytics toolkit for teachers. *Educ. Technol. Soc.* 15, 58–76.
- Gasević, D., Dawson, S., Siemens, G., 2015. Let's not forget: learning analytics are about learning. *TechTrends* 59, 64–71.
- Gasević, D., Jovanović, J., Pardo, A., Dawson, S., 2017. Detecting learning strategies with analytics: links with self-reported measures and academic performance. *J. Learn. Anal.* 4, 113–128.
- Gasević, D., Joksimović, S., Egan, B.R., Shaffer, D.W., 2019a. SENS: network analytics to combine social and cognitive perspectives of collaborative learning. *Comput. Hum. Behav.* 92, 562–577.
- Gasević, D., Tsai, Y.-S., Dawson, S., Pardo, A., 2019b. How do we start? An approach to learning analytics adoption in higher education. *Int. J. Inf. Educ. Technol.* 36, 342–353.
- Gibson, D.C., Huband, S., Ifenthaler, D., Parkin, E., 2018. Return on Investment in Higher Education Retention: Systematic Focus on Actionable Information From Data Analytics. Ascilite Conference, Geelong, VIC, Australia. (Accessed 25 November 2018).
- Henry, M., Gibson, D.C., Flodin, C., Ifenthaler, D., 2020. Learning innovations for identifying and developing talent for university. In: Parrish, D., Joyce-McCoach, J. (Eds.), *Innovations in Higher Education—Cases on Transforming and Advancing Practice*. IntechOpen, London, UK.
- Hlosta, M., Zdrahal, Z., Zendulka, J., 2017. Ouroboros: Early Identification of At-Risk Students Without Models Based on Legacy Data. Seventh International Conference on Learning Analytics & Knowledge. Vancouver, Canada.

- Ifenthaler, D., Gibson, D.C. (Eds.), 2020. *Adoption of Data Analytics in Higher Education Learning and Teaching*. Springer, Cham.
- Ifenthaler, D., Schumacher, C., 2016. Student perceptions of privacy principles for learning analytics. *Educ. Technol. Res. Dev.* 64, 923–938.
- Ifenthaler, D., Schumacher, C., 2019. Releasing personal information within learning analytics systems. In: Sampson, D.G., Spector, J.M., Ifenthaler, D., Isaias, P., Sergis, S. (Eds.), *Learning Technologies for Transforming Teaching, Learning and Assessment at Large Scale*. Springer, Cham.
- Ifenthaler, D., Widanapathirana, C., 2014. Development and validation of a learning analytics framework: two case studies using support vector machines. *Technol. Knowl. Learn.* 19, 221–240.
- Ifenthaler, D., Yau, J., 2019. Higher education stakeholders' views on learning analytics policy recommendations for supporting study success. *Int. J. Learn. Anal. Artif. Intell. Educ.* 1, 28–42.
- Ifenthaler, D., Yau, J.Y.-K., 2020. Utilizing learning analytics to support study success in higher education: a systematic review. *Educ. Technol. Res. Dev.* 68, 1961–1990.
- Ifenthaler, D., Gibson, D.C., Dobozy, E., 2018. Informing learning design through analytics: applying network graph analysis. *Australas. J. Educ. Technol.* 34, 117–132.
- Ifenthaler, D., Yau, J.Y.-K., Mah, D.-K. (Eds.), 2019. *Utilizing Learning Analytics to Support Study Success*. Springer, New York, NY.
- Ifenthaler, D., Gibson, D.C., Prasse, D., Shimada, A., Yamada, M., 2021. Putting learning back into learning analytics: actions for policy makers, researchers, and practitioners. *Educ. Technol. Res. Dev.* 69, 2131–2150.
- Ifenthaler, D., 2015. Learning analytics. In: Spector, J.M. (Ed.), *The SAGE Encyclopedia of Educational Technology*. Sage, Thousand Oaks, CA.
- Ifenthaler, D., 2017. Are higher education institutions prepared for learning analytics? *TechTrends* 61, 366–371.
- Ifenthaler, D., 2020. Change management for learning analytics. In: Pinkwart, N., Liu, S. (Eds.), *Artificial Intelligence Supported Educational Technologies*. Springer, Cham.
- Kevan, J.M., Ryan, P.R., 2016. Experience API: flexible, decentralized and activity-centric data collection. *Technol. Knowl. Learn.* 21, 143–149.
- Klasen, D., Ifenthaler, D., 2019. Implementing learning analytics into existing higher education legacy systems. In: Ifenthaler, D., Yau, J.Y.-K., Mah, D.-K. (Eds.), *Utilizing Learning Analytics to Support Study Success*. Springer, New York, NY.
- Kotter, J.P., 2007. Leading change: why transformation efforts fail. *Harv. Bus. Rev.* 96–103.
- Leitner, P., Ebner, M., Ebner, M., 2019. Learning analytics challenges to overcome in higher education institutions. In: Ifenthaler, D., Yau, J.Y.-K., Mah, D.-K. (Eds.), *Utilizing Learning Analytics to Support Study Success*. Springer, Cham.
- Lester, J., Klein, C., Rangwala, H., Johri, A., 2017. *Learning Analytics in Higher Education*. Wiley, Malden, MA.
- Lockyer, L., Heathcote, E., Dawson, S., 2013. Informing pedagogical action: aligning learning analytics with learning design. *Am. Behav. Sci.* 57, 1439–1459.
- Long, P.D., Siemens, G., 2011. Penetrating the fog: analytics in learning and education. *Educ. Rev.* 46, 31–40.
- Mah, D.-K., Ifenthaler, D., 2017. Academic staff perspectives on first-year students' academic competencies. *J. Appl. Res. High Educ.* 9, 630–640.
- Mah, D.-K., Ifenthaler, D., 2020. What do first-year students need? Digital badges for academic support to enhance student retention. *J. Appl. Res. High Educ.* 12, 86–96.
- Nouri, J., Ebner, M., Ifenthaler, D., Saqr, M., Malmberg, J., Khalil, M., Viberg, O., Bruun, J., Conde González, M.A., Papamitsiou, Z., Berthelsen, U.D., 2019. Efforts in Europe for data-driven improvement of education: a review of learning analytics research in seven countries. *Int. J. Learn. Anal. Artif. Intell. Educ.* 1, 8–27.
- Papamitsiou, Z., Filippakis, M., Poulou, M., Sampson, D.G., Ifenthaler, D., Giannakos, M., 2021. Towards an educational data literacy framework: enhancing the profiles of instructional designers and e-tutors of online and blended courses with new competences. *Smart Learn. Environ.* 8, 18.
- Pardo, A., Siemens, G., 2014. Ethical and privacy principles for learning analytics. *Br. J. Educ. Technol.* 45, 438–450.
- Pistilli, M.D., Arnold, K.E., 2010. Purdue Signals: mining real-time academic data to enhance student success. In: *About Campus: Enriching the Student Learning Experience*, vol. 15, pp. 22–24.
- Prinsloo, P., Slade, S., 2014. Student data privacy and institutional accountability in an age of surveillance. In: Menon, M.E., Terkla, D.G., Gibbs, P. (Eds.), *Using Data to Improve Higher Education. Research, Policy and Practice*. Sense Publishers, Rotterdam.
- Psacharopoulos, G., 2014. The returns to investment in higher education. In: Menon, M.E., Terkla, D.G., Gibbs, P. (Eds.), *Using Data to Improve Higher Education. Global Perspectives on Higher Education*. Sense Publishers, Rotterdam.
- Rogers, E.M., 1962. *Diffusion of Innovations*. Free Press of Glencoe, New York, NY.
- Sahin, M., Ifenthaler, D., 2021. Visualizations and dashboards for learning analytics: challenges and future directions. In: Sahin, M., Ifenthaler, D. (Eds.), *Visualizations and Dashboards for Learning Analytics*. Springer, Cham.
- Scholes, V., 2016. The ethics of using learning analytics to categorize students on risk. *Educ. Technol. Res. Dev.* 64, 939–955.
- Schumacher, C., Klasen, D., Ifenthaler, D., 2019. Implementation of a learning analytics system in a productive higher education environment. In: Khine, M.S. (Ed.), *Emerging Trends in Learning Analytics*. Brill, Leiden NL.
- Sclater, N., Mullan, J., 2017. *Learning Analytics and Student Success—Assessing the Evidence*. JISC, Bristol.
- Sergis, S., Sampson, D.G., 2016. School analytics: a framework for supporting school complexity leadership. In: Spector, J.M., Ifenthaler, D., Sampson, D.G., Isaias, P. (Eds.), *Competencies in Teaching, Learning and Educational Leadership in the Digital Age*. Springer, Cham.
- Sergis, S., Sampson, D.G., 2017. Teaching and learning analytics to support teacher inquiry: a systematic literature review. In: Peña-Ayala, A. (Ed.), *Learning Analytics: Fundamentals, Applications, and Trends*. Springer, Cham.
- Seufert, S., Meier, C., Soellner, M., Rietsche, R., 2019. A pedagogical perspective on big data and learning analytics: a conceptual model for digital learning support. *Technol. Knowl. Learn.* 24, 599–619.
- Siemens, G., Dawson, S., Lynch, G., 2014. Improving the Quality and Productivity of the Higher Education Sector—Policy and Strategy for Systems-Level Deployment of Learning Analytics. Office of Learning and Teaching, Australian Government, Canberra, Australia. Available: http://solaresearch.org/Policy_Strategy_Analytics.pdf.
- Tinto, V., 2005. Reflections on student retention and persistence: moving to a theory of institutional action on behalf of student success. *Stud. Learn. Eval. Innov. Dev. Psychol.* 2, 89–97.
- Tsai, Y.-S., Gasević, D., 2017. *Learning Analytics in Higher Education—Challenges and Policies: A Review of Eight Learning Analytics Policies*. Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada.
- Vey, K., Fandel-Meyer, T., Zipp, J.S., Schneider, C., 2017. Learning & development in times of digital transformation: facilitating a culture of change and innovation. *Int. J. Adv. Corp. Learn.* 10, 22–32.
- Viberg, O., Hatakka, M., Bälter, O., Mavroudi, A., 2018. The current landscape of learning analytics in higher education. *Comput. Hum. Behav.* 89, 98–110.
- West, D., Huijser, H., Heath, D., 2016. Putting an ethical lens on learning analytics. *Educ. Technol. Res. Dev.* 64, 903–922.
- Willis, I.J.E., Slade, S., Prinsloo, P., 2016. Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educ. Technol. Res. Dev.* 64, 881–901.
- Wong, J., Baars, M., De Koning, B.B., Van Der Zee, T., Davis, D., Khalil, M., Houben, G.-J., Paas, F.G., 2019. Educational theories and learning analytics: from data to knowledge. In: Ifenthaler, D., Yau, J.Y.-K., Mah, D.-K. (Eds.), *Utilizing Learning Analytics to Support Study Success*. Springer, Cham.
- Wong, B.T.M., 2017. Learning analytics in higher education: an analysis of case studies. *Asian Assoc. Open Univ. J.* 12, 21–40.
- Yau, J., Ifenthaler, D., 2020. Reflections on different learning analytics indicators for supporting study success. *Int. J. Learn. Anal. Artif. Intell. Educ.* 2, 4–23.

The intricacies of class size in higher education

Justine R. Okerson, East Carolina University, Greenville, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	336
Admission office perspective	337
Selectivity	337
Recruitment	337
The campus visit	338
Retention	338
Transfer students	339
Variations across types of institutions	339
Disparities by country/socioeconomic status	339
Costs	340
Student success	340
Registrar	341
Pressures affecting class size	341
Changing demographics	341
Final thoughts	341
References	342

Overview

In higher education there are multiple classifications for stakeholders within the organization including faculty, staff, and students. When researching a college or university, most institutions start to distinguish themselves by sharing a multitude of defining numbers with the general public and their prospective student audience. A school profile or about page on an institutions' Website often hosts a list of these numerical divisions, for example the number of undergraduate and graduate students, the number of majors and minors offered, the number of graduate degree programs offered, if any, the student to faculty ratio, the retention and graduation rate, and the school's average class size. One of the most essential terms when considering admission and selection to a university is class size. The term class size is defined as serving two purposes: the number of students in an individual course, or the average number of students being taught in a specific educational system. Class size is also an instrumental measurement for a registrar's office to use for classroom management and creating sections for courses (Rooney, 2020).

Class size in relation to the American four-year residential college model also has two meanings: both the size of individual academic courses and the size of an incoming cohort of students (e.g., the freshman class) (Bettinger et al., 2017). Class size can be referred to when recruiting students into certain cohort models for scholarships, graduate programs, or early assurance programs to law school, medical school, etc. However, class size is a critical component to attracting prospective high school students to attend a specific institution. Class size also affects how admission officers craft an incoming cohort (Duffy and Goldberg, 2014). Class size is relevant to prospective college students, college rankings, admission offices, and institutional priorities.

To better distinguish each definition and usage of class size, the type, stakeholders, and definition of each variation of class size is defined, as seen in Table 1.

Table 1 Class size types and definitions.

<i>Class size type</i>	<i>Stakeholders</i>	<i>Definition</i>
Class size (academic)	Department-specific	Number of students in an individual course
Class size (comparative)	Rankings	Average number of students being taught in a specific educational system/institution
Class size (enrollment)	Admission office; scholarships, specific departments	Size of an incoming cohort of students or the total number of students enrolled at an institution
Class size (section)	Registrar's office; individual academic departments	A metric for dividing courses into smaller units for classroom management
Class size (selectivity)	Prospective students	An indicator of a university's dedication to undergraduate student teaching and mentorship, as well as the accessibility level of faculty

This chapter will examine the various meanings and definitions of the term class size as they relate to the admission office perspective, selectivity, recruitment, the campus visit, retention, transfer students, variations across institutions, disparities by country/socioeconomic status, costs, student success, registrar, pressures facing class size, and changing demographics in higher education.

Admission office perspective

Perhaps the first mention of class size (enrollment) occurs while high school students are progressing through their college choice process (Hossler and Palmer, 2008; Perna, 2006). In our rapidly changing economy and world, universities are quickly transforming and adopting business practices to attract students (Rossi, 2014; Slaughter and Rhoades, 2004). The college admission process is becoming a frenzy, as institutions of higher education seek to compete for students and increase their tuition revenues as the operating costs for higher education continue to rise (Borin, 2014) and state support for public universities drop (Baum et al., 2013). High school students receive an overwhelming amount of information geared to help them make decisions about their post-graduation plans. They collect brochures regarding four-year, two-year, and professional school options; high school students are also often subscribed to social media campaigns run by admission offices through blogs, Facebook, Twitter, and Instagram promoting day in the life campaigns, college tours, and college fairs (Gregory, 2014; Johnston, 2010). Mailings include college viewbooks, emails, letters, and packets that all promote the college admission process.

One of the first questions school counselors and family members ask a rising high school senior about attending college is regarding institution size and whether or not the student hopes to attend a small or large institution, to determine the student's ideal class size (enrollment) (Johnston, 2010). In the barrage of advertisements received, high school students are left to make meaning of institutions and their own potential interest in attending based on facts and figures regarding average class sizes (comparative) and student-to-faculty ratios (Borin, 2014). Many prospective students and families compare colleges and universities to determine their best academic fit based on these precise details regarding class size (comparative), both in the classroom and in their cohort (Mattern et al., 2010). Class size (comparative) is a factor in various ways regarding the admission process and selectivity of an institution; there are also variations in pressures that affect class size (enrollment) across different types of institutions (Hossler and Palmer, 2008).

Selectivity

The official reporting of class size (selectivity) is one of many factors that contribute to college rankings and ultimately impact the admission process (Kim, 2018). Given the importance of the college choice decision by students and the desire of selective institutions to maintain their reputation and rankings, college enrollment managers are paying increased attention to all elements of institutional communication that may influence applicants' decisions (Hossler and Palmer, 2008). Because students are applying to more schools on average than previously, it becomes even more challenging for colleges and universities to predict enrollment and thus accurately predict acceptance and yield numbers (Clinedinst and Koranteng, 2017). If acceptance numbers rise in a given year, the institution's selectivity rating will be affected in national rankings, which ultimately impacts the next year's prospective applicants due to changes in reputation of perceived selectivity (Clinedinst and Koranteng, 2017). For students seeking out a selective higher education institution, a dip in the rankings even by a few points may impact a student's perception of academic reputation, campus resources, and career prospects after graduation.

From this vantage point, class size (selectivity) is viewed as an indicator of a university's dedication to undergraduate student teaching, faculty/student mentorship, and accessibility of faculty members for office hours, research opportunities, and physical presence in the classroom. At institutions with larger average class sizes (academic) and fewer professors per student, many prospective students will assume a dependence on teaching assistants, an abundance of large and impersonal lecture hall classes, or an institutional-based emphasis and priority toward graduate level coursework/research (Duffy and Goldberg, 2014).

Recruitment

Class size (enrollment) is important when recruiting students into cohort models for merit scholarships or early assurance programs to graduate programs. Often times while the admission office may be looking to fill a larger class size (enrollment) for the incoming class of students that number can be divided even further into multiple sub-groups and populations. A merit scholarship office has the responsibility to fill one or multiple merit scholarship spaces with a certain number of students and varying demographic requirements. The admission and financial aid office often have one-off scholarships as well, often funded by donors with strict regulations, for example a student from a specific county or a student with a specific academic or extracurricular interest (Dynarski, 2004).

From the admission office perspective, class size (enrollment) is paramount in concern for building a balanced class and meeting the needs of a variety of institutional stakeholders (Borin, 2014). One of the many concerns for an institution is shaping a class that is well-suited for academic and personal success at the specific institution, yet also creating a group with a broad range of diverse

backgrounds and experiences (e.g., geographic diversity, racial/ethnic diversity, socio-economic status diversity, sexual/gender diversity, political diversity) to create globally engaged and informed citizens.

As a method of increasing yield among highly desirable students there may also be consideration for a small stipend, admission to the honors college, selection for gap year funding, or guaranteed/early admission to business school, medical school, or law school. The athletic and arts departments are also attempting to fill any open spots on their teams or within their departments. Some institutions also make use of recruitment of first-year students to study abroad or gap year programs to fill any potential gaps into the class size (enrollment) later. These students may come in as second-semester freshmen or sophomores to increase retention and graduation rates, despite other students withdrawing, leaving, or failing after the first semester. Often it is first-year housing that makes over-enrolling so complicated, so bringing in students to attend in-person a semester later or even come to campus as a sophomore gives colleges more options when building their class (Duffy and Goldberg, 2014). The making of a class also means the making of many sub-classes within that class. One shift here or there, or lack of yield in a certain area affects all stakeholders involved in the process.

The campus visit

While institutional characteristics such as size, location, prestige, cost, diversity, admission rate, and college rankings vary between colleges, each characteristic may influence students in making decisions about whether or not to visit, apply, or enroll in a specific institution (Perna, 2010). Regardless of whether or not a student is seeking a large or small school, the campus visit allows a student to experience what the size of the institution truly “feels” like (Okerson, 2016). Research investigated the role of habitus in addition to the role of community, campus, and policy makers on college choice (Perna, 2006). Students are seeking a sense of belonging, and the most influential part of the campus visit occurs on an individualized level (Perna, 2006). Students decide whether and where to attend college, but also decide what to study, whether to persist and graduate all based on the amount of human, financial, social, and cultural capital available to the student throughout their postsecondary educational experience (Perna, 2006). While high school students can read about the average class size (comparative) and student to faculty ratio, the campus visit allows a prospective student to evaluate that size first-hand in order to determine fit (Mattern et al., 2010).

Previous research has shown that the aesthetics of campus make a significant impact on a student’s decision of where to attend college, and within aesthetics is the size of the campus (Okerson, 2016). Size of the college campus was an important factor for students when deciding where to apply and how the students felt they connected to the campus once they visited (Okerson, 2016). Even though some students initially stated a preference for a smaller or larger college environment, most students applied to a mixture of sizes. During the campus tour, students seemed to have a better grasp of whether the size was a good fit for them or not (Okerson, 2016; Mattern et al., 2010).

While not all campus tours allow students to view a live class or classroom, colleges and universities often provide an experience for a prospective student to sit in on a class or shadow a current student for a day (Okerson, 2016). With COVID-19, the future of seeing a classroom in action is also in question, but it is critical as it allows prospective students to evaluate the academic rigor of an institution, engage and observe current students in the classroom, and evaluate the size of future academic classes through their own personal lens. The campus visit is essential for a prospective college student to better understand the size of an institution as well as get a glimpse of what their future class sizes (academic) would be on campus (Okerson, 2016).

Retention

Colleges and universities are often graded and ranked on metrics such as retention and graduation rates as measures of their success by stakeholders, policymakers, and accrediting boards (Millea et al., 2018). Incoming students and prospective families are interested in those numbers as a way to assess how many students are “happy” with the product of their college education. They assume that if the majority of students are happy, that their student will be as well. It is extremely difficult to replace students who leave before completing their degrees, and thus institutions are motivated to find ways to increase consumer happiness, occasionally through class size (academic) (Millea et al., 2018). Some universities limit and reduce class sizes (academic) to increase accountability for students to encourage them to attend classes, engage with professors, and participate in their college experience; it is also a helpful method for student affairs officers to identify students who may be falling through the cracks.

Research shows that students who have more shared experiences in common with other students through courses and student life programming are more likely to persist through their college experience (Sapelli and Illanes, 2016). And while large class sizes (academic) may not negatively impact a student’s grades, they can negatively impact a student’s level of satisfaction with their institution in addition to their final evaluation of their instructors/professors (Sapelli and Illanes, 2016). For many institutions that seek to increase their retention and graduation rates over time, class size (academic) is an important factor to consider investing in to improve student success and satisfaction, and ultimately their future number of graduates (Millea et al., 2018).

Transfer students

Retention and graduation rates go hand in hand with the idea of class size (enrollment) post-first year in college. As incoming first-year students admitted to a college or university finish their first year of study, many do not return. Whether a student transfers, takes a gap year, pursues a job or internship opportunity, or requires a medical leave of absence, each student that leaves the institution ultimately impacts class size (enrollment) for the institution (Glynn, 2019). Generally, the first-year cohort will always be the largest at the majority of institutions, as there is a new applicant pool to replenish it with each year, but it is challenging for four-year institutions to recruit and yield transfer students. Enrolling transfer students into a predetermined class size (enrollment) is yet another challenge as institutional size and type impact how many transfer students are accepted, how many in-state, out-of-state, or international students are admitted, and the amount of financial aid the institution can support.

For many years transfer students were looked to as an incentive and a way to replenish funding streams for the university (Taylor and Jain, 2017). Transfer student admission programs are rarely as robust or slick as the resources, time, and materials that go into recruiting first-year students and less time and attention are often dedicated to transfer orientation and success programs on college campuses (Glynn, 2019). Yet, transfer students are often a great method for an institution to bring in extra students to campus, who have demonstrated success in college-level courses and may have a more mature understanding of what they are seeking in their desired college fit. When transfer students transition well and are successful, they can also be instrumental in recruiting and connecting with prospective transfer students. Putting time and effort into recruiting transfer students and orienting them to campus is necessary for institutions trying to improve their graduation and retention rates. Transfer applicants are an essential component of understanding class size (enrollment) and the impact of retention in higher education.

Variations across types of institutions

Class size also varies depending on the context. Online universities will view class size (academic) using a different lens, as research shows class size (academic) has little impact in online settings (Bettinger et al., 2017). Students may choose to take online classes through their home institution or receive transfer credit toward their home institution. Community colleges are also known for typically having smaller academic class sizes (academic) across the board. Class size (academic) stipulations also vary widely in international sectors, based on whether higher education is viewed as a public good or a private benefit (Lambert, 2014). As a public good, class sizes (enrollment) are likely larger to accommodate more students, and their counterparts at expensive institutions that offer private benefits often tout small class sizes (academic) in addition to more attention and resources for students (Lambert, 2014).

Cohort and academic class sizes fluctuate at various institutions based on a number of factors including institution size. Class size (academic) differs by funding type (i.e., whether the institution is private or public, for-profit or non-profit), in addition to whether the institution is more narrowly or broadly focused (e.g., liberal arts focused or strictly engineering). Individual class sizes (academic) also vary due to the competitiveness of an institution, and the method of delivery. Colleges or universities who are more selective tend to have larger endowments and can afford additional faculty and additional class offerings to support smaller average class sizes (academic) and smaller faculty to student ratios (Rossi, 2014). In many ways, more competitive institutions must continue dedicating funds to ensure their individual class sizes (academic) remain the same in order to remain selective in the rankings.

Disparities by country/socioeconomic status

Another important distinction for class size (academic) is context. Class size (academic) varies by country and is also impacted by socioeconomic status (SES). In the United States, class size (academic) affects students' achievement and psychological well-being in relation to group composition (immigrant background, gender, SES, and achievement) and size (Alivernini et al., 2020). Students who come from a lower SES background have an elevated risk of lower levels of well-being and higher chances of developing psychosocial stress (Currie et al., 2012). Class size (academic) also has a significant impact on how a student feels about school (Rjosk et al., 2015). These risks make studying class size (academic) important to researchers around the world, as they desire to provide the greatest probabilities of academic success and positive mental health for all students.

Class size (academic) reduction is touted by many in the United States as a method of improving learning outcomes and behavior (Alivernini et al., 2020; Rjosk et al., 2015; Rossi, 2014), but a South African study chose to take on this research as it is a country whose considerably larger class sizes are found in low SES areas of the country (Köhler, 2022). In this study, there were 20 or more learners in the average classroom of the poorest public schools in the country (Köhler, 2022). But higher-class size in this study was not associated with negative learning outcomes in low SES areas; instead, negative learning outcomes were only found in wealthier schools with larger class sizes, which means other factors are distinctly at play (Köhler, 2022).

Flores-Mendoza et al. (2021) performed research in Argentina to study SES in the classroom and found significant ties between larger class sizes and low SES areas as they attempted to study the influence of a school's resources on a student's ability and academic performance. Ultimately this study found that individual differences, not SES variables were most at play for differences

in achievement (Flores-Mendoza et al., 2021). Artificial Intelligence allowed for the study of 110,627 students, to understand academic achievement in public high schools in a European Union County, and the study found that class size had no impact on academic achievement in that context (Cruz-Jesus et al., 2020).

Another study specifically analyzing European countries found Portugal as an outlier as it was the only country that does not see an impact on class size for students' reading scores; Croatia and Italy also stood out as the most important variable for math scores in both countries was class size (Bonacini et al., 2021). Overall, the study demonstrated that the gaps in test scores by county varied most due to the quality of the school, and by class size, rather than varying by social or economic status (Bonacini et al., 2021).

Context is extremely important in evaluating class size (academic), as each country, region, and school has a different environment that results in disparate academic achievement markers. Class size (academic) is affected by SES, school resources, and teacher quality. Even in studies of just one country, finding definitive answers on class size proved challenging. Continued research regarding class size and implications is needed in each specific context for researchers to better understand the phenomenon in their country, region, or continent.

Costs

Class size is also associated with cost for higher education. Institutions have to strike the right balance to provide for students and also invest for the future. While small class sizes (academic) have historically been revered in the United States, these same idyllic classes come with a hefty price tag. In relation to academic courses on campus the smaller the class size (academic), the higher the number of total courses offered has to be, which in turn means a department needs to hire and pay a greater number of instructors, and all of that adds up for an institution, especially if tenure is involved and the institution is research-focused (Baum et al., 2013). Each class size (academic) also comes with facility costs for heating and cooling lecture halls, classrooms, and labs. If class sizes (enrollment) are too large, institutions have capacity concerns. The economies of scale are influential for every institution. Finding the right balance of small classes where it matters in addition to larger class sizes (academic) allows institutions more flexibility and additional cost savings.

In admission, on the other hand, coming in short on the incoming class size (enrollment) means the possibility of not having enough tuition or revenue to support salaries, buildings, and existing programs on campus. Of course, many institutions utilize waitlists in order to combat summer melt and committed students withdrawing after being admitted to other institutions, and that can help to ensure the incoming class size (enrollment) is on target. The more selective a college or university is, the smaller the class size (enrollment) and student to faculty ratio. But often there are factors that are not within an admission director's control. If a nearby competing campus over enrolls that can impact enrollment for another institution.

During COVID-19, some students opted to take a year off from school to work or chose to enroll at local community colleges to receive college credits online for a fraction of the price of their home college or university. Many institutions are grappling with how to price online education to stay competitive in the academic market and still bring in enough revenue to keep their programs afloat (Baum et al., 2013; Bettinger et al., 2017). As the cost of higher education continues to rise, students and families struggle when considering which institution of higher education provides the best education, campus life, campus amenities, and location; however, these elements must be weighed against the total cost of that institution (Lambert, 2014).

Student success

Class size (academic) and student success have always been intertwined, as many have suggested that smaller class sizes (academic) indicate quality teaching, instruction, feedback, and learning and thus provide better avenues for demonstrating potential student success, but the accuracy of these claims is also often debated (Ake-Little et al., 2020; Kuh, 2008). While research advocates for small first-year seminars as a high-impact experience for students to gain critical thinking skills and develop intercultural competency, the research does not demonstrate that class size (academic) impacts future success or retention (Kuh, 2008).

Recent research from Temple University drew a robust sample size to test whether or not smaller class sizes (academic) accomplish the intended goal of greater student success (Ake-Little et al., 2020). Many limitations existed with previous research and did not control for instructor experience or attempt to ensure that race, gender, and academic disciplines were all factors being controlled for during previous studies (Ake-Little et al., 2020). But in this particular case, researchers looked at all factors across several years of general education courses to see if there is indeed a correlation between success and class size (academic) (Ake-Little et al., 2020).

The results indicated that class size (academic) effects are much more nuanced than previously imagined. While instructors may prefer smaller class sizes (academic) because it allows them to get to know their student's better and use a larger variety of pedagogical tools, not all learning occurs within the classroom itself, and thus the researchers propose that social group theory, and the impact of a student's peers also is a factor of the impact of class size (academic) (Ake-Little et al., 2020; Sapelli and Illanes, 2016). When examining both race and gender there are also distinct nuances regarding impact that shift based on the academic discipline (Ake-Little et al., 2020). Surprisingly no student group seemed to be affected negatively or positively by large class sizes (academic) (Ake-Little et al., 2020). The researchers go on to explain that in a larger class, an instructor must shift the assignments and objectives based on the size of the classroom and thus the experience and expectations are different in larger class size

(academic) sections of a course (Ake-Little et al., 2020). But one limitation of the research is that the researchers used student grades as the measurement output of student success, which while it is a comparable piece of data from every class regardless of class size (academic), it does not always accurately measure student success or student learning. Also, there is not research to indicate the balance of varying class sizes (academic) in a student's course load. But class size (academic) is not dictated solely by the institution itself, as accreditation bodies also examine class size (academic) as a method of reviewing academic quality and standards.

Registrar

The primary role of a university registrar is to maintain course records and student data. Often this includes tasks like scheduling space and time for class, maintaining course catalogs, managing registration, directing curriculum management, and fulfilling transcript requests (Rooney, 2020). Registrars are closely aligned with class size (section) as they are constantly dealing with the question of increasing capacity without affecting academic integrity. Especially during the COVID-19 pandemic, university registrars have had to reimagine and attend to best practices for online class sizes (section) (Rooney, 2020). They also have to assess the data after the fact, examining how many students withdrew from the course, the distribution of grades, and the course evaluations to determine if the ratio of students to instructor were on or off for a particular course. Registrars are constantly examining the idea of ideal class size (section) for different types of courses, departments, and college-level students to find the right balance to enhance student success and the overall academic experience on campus (Rooney, 2020).

Pressures affecting class size

Institutions are concerned with bringing students together who have different academic and extracurricular interests. The challenge for admission offices is to balance external pressures for funding and rankings with institutional needs and relationships with on-campus partners (e.g., President/Provost's office, athletics, arts departments, university advancement) in order to define the incoming class size (enrollment). There is often a great deal of pressure from the university administration to finish the admission cycle with a very specific target. As most universities are highly dependent on tuition revenues, meeting enrollment targets is critical. This precise number depends on admission using prediction models based on previous admission cycles and an analysis of the decision-making of prospective students. A higher number of incoming students can negatively impact housing availability in addition to student happiness and retention. A lower number of incoming students affects the institution's budget. Many offices and resources on campus depend on a specific target number being met.

In order to maintain control over size, many colleges and universities utilize a waitlist as a method of managing these precise targets. By enrolling just under the number of students desired in the next incoming class, a strategic admission office is able to fill the spots one by one with candidates off of the waitlist to end up at their precise target without over-enrolling. When an institution increases their class size (enrollment) for the next incoming year, prediction models often fail, and many institutions may over or under enroll at a significant rate.

Changing demographics

High school class sizes (enrollment) are another element impacting the future of higher education today, as changing demographics will make filling classrooms and cohorts even more difficult as we look into the future (Grawe, 2018). Forecasters suggest a significant drop in the number of high school graduates based on birth rates affected by the economic downturn in 2008 (Grawe, 2018). Also, in the years to come there will be important differences in the makeup of the college-going population with an expected boom of first-generation, underrepresented, and rural students (Grawe, 2018).

Many colleges and universities are working now to attract non-traditional adult learners to their programs. Institutions are expanding online education opportunities that they discovered helped to increase access during the COVID-19 pandemic, increasing outreach to different parts of the country, focusing on diversity, and being intentional about outreach to first-generation underrepresented students (Grawe, 2018). These changing demographics require admission offices and marketing staff adjust strategies, plan for the future, and be creative in new ways to attract students to apply and enroll in order to adapt their class sizes (enrollment) to the changing demographics of the United States (Grawe, 2018).

Final thoughts

Class size is a critical element of higher education to be understood through the lens of prospective students, admission officers, and institutional stakeholders. It is not just the number of students in one particular academic class or cohort; it is also the framing of what the community is for a particular college or university. The number of students in a particular environment and the representation of inclusion among various types of diversity on campus within that number are critical for many prospective students, as well as for the staff and faculty. The definitions and the context in this chapter provide additional context on the study and

implications of class size for various stakeholders in higher education. Class size is important to understand when evaluating admission and enrollment, the role of the registrar, and the reputation and accreditation of an institution. For many institutions, much is at stake in the decisions that drive both the number of students on campus, in cohorts, and in class years in addition to those pre-determined class sizes (academic) and offerings.

References

- Ake-Little, E., von der Embse, N., Dawson, D., 2020. Does class size matter in the university setting? *Educ. Res.* 49 (8), 595–605.
- Alivernini, F., Cavicchiolo, E., Manganelli, S., Chirico, A., Lucidi, F., 2020. Students' psychological well-being and its multilevel relationship with immigrant background, gender, socioeconomic status, achievement, and class size. *Sch. Effect. Sch. Improv.* 31 (2), 172–191.
- Baum, S., Ma, J., Payea, K., 2013. *Education Pays 2013*. College Board. Retrieved from: <https://files.eric.ed.gov/fulltext/ED572537.pdf>.
- Bettinger, E., Doss, C., Loeb, S., Rogers, A., Taylor, E., 2017. The effects of class size in online college courses: experimental evidence. *Econ. Educ. Rev.* 58, 68–85.
- Bonacini, L., Brunetti, I., Gallo, G., 2021. Choose the School, Choose the Performance. New Evidence on the Determinants of Student Performance in Eight European Countries (No. 905). Retrieved from: <https://www.econstor.eu/bitstream/10419/236625/1/GLO-DP-0905.pdf>.
- Borin, J., 2014. How to Manage the College Admissions Frenzy. *College Connections*. Retrieved from: <http://www.college-connections.com/blog/manage-college-admissions-frenzy/>.
- Clinedinst, M., Koranteng, A.M., 2017. State of College Admission. Retrieved from: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/83148/StateofCollegeAdmission.pdf?sequence=1&isAllowed=y>.
- Currie, C., Zanotti, C., Morgan, A., Currie, D., de Looze, M., Roberts, C., Samdal, O., Smith, O., Barnekow, V. (Eds.), 2012. Social Determinants of Health and Well-Being Among Young People. Health Behaviour in School-Aged Children (HBSC) Study: International Report From the 2009/2010 Survey. Health Policy for Children and Adolescents, No. 6. WHO Regional Office for Europe, Copenhagen. Retrieved from: https://www.euro.who.int/__data/assets/pdf_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf.
- Cruz-Jesus, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M., Rosa-Louro, A., 2020. Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon* 6 (6), e04081.
- Duffy, E.A., Goldberg, I., 2014. *Crafting a Class: College Admissions and Financial Aid, 1955–1994*, vol. 77. Princeton University Press.
- Dynarski, S., 2004. The new merit aid. In: *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*. University of Chicago Press, pp. 63–100.
- Flores-Mendoza, C., Ardila, R., Gallegos, M., Reategui-Colareta, N., 2021. General intelligence and socioeconomic status as strong predictors of student performance in Latin American schools: evidence from PISA items. *Front. Educ.* vol. 6, 45.
- Glynn, J., 2019. Persistence: The Success of Students Who Transfer From Community Colleges to Selective Four-Year Institutions. Retrieved from: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/90744/PersistenceTransferFourYears.pdf?sequence=1&isAllowed=y>.
- Gregory, J., 2014. 10 Ways Students Search College's Today- and How to Adapt: Understand How Students Choose a School and Recalibrate Your Approach for Recruitment Success, vol. 17. University Business. Retrieved from: <https://universitybusiness.com/10-ways-students-search-colleges-today-and-how-to-adapt/>.
- Grave, N.D., 2018. Demographics and the Demand for Higher Education. JHU Press.
- Hossler, D., Palmer, M., 2008. Why understand research on college choice? In: *National Association of College Admissions Counselors (Ed.), Fundamentals of College Admissions Counseling*. Kendall Hunt Publishers, Dubuque, IA, pp. 42–53.
- Johnston, T.C., 2010. Who and what influences choice of university: student and university perceptions. *Am. J. Bus. Educ.* 3 (10), 15–24.
- Kim, J., 2018. Prestige, rankings and the competition for status. In: *Controversies on Campus: Debating the Issues Confronting American Universities in the 21st Century*, pp. 99–133.
- Köhler, T., 2022. Class size and learner outcomes in South African schools: the role of school socioeconomic status. *Dev. South Afr.* 39 (2), 126–150.
- Kuh, G.D., 2008. High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter. Report from the Association of American Colleges and Universities.
- Lambert, M.T., 2014. *Privatization and the Public Good: Public Universities in the Balance*. Harvard Education Press, Cambridge, MA.
- Mattern, K.D., Woo, S.E., Hossler, D., Wyatt, J., 2010. Use of student-institution fit in college admissions: do applicants really know what is good for them? *Coll. Univ.* 85 (4), 18–26.
- Millea, M., Wills, R., Elder, A., Molina, D., 2018. What matters in college student success? Determinants of college retention and graduation rates. *Education* 138 (4), 309–322.
- Okerson, J.R., 2016. Beyond the Campus Tour: College Choice and the Campus Visit. Ph.D. Dissertation. William & Mary. Available at: <https://scholarworks.wm.edu/etd/1463413085/>.
- Perna, L.W., 2000. Differences in college enrollment among African Americans, Hispanics, and whites. *J. High. Educ.* 71, 117–141.
- Perna, L.W., 2006. Studying college access and choice: a proposed conceptual model. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XXI. Springer, New York, NY, pp. 99–157.
- Perna, L.W., 2010. Toward a more complete understanding of the role of financial aid in promoting college enrollment: the importance of context. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XXV. Springer, New York, NY, pp. 129–179.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Stanat, P., 2015. Classroom composition and language minority students' motivation in language lessons. *J. Educ. Psychol.* 107 (4), 1171–1185. <https://doi.org/10.1037/edu0000035>.
- Rooney, C., 2020. The registrar's office as a storytelling organization. *Coll. Univ.* vol. 95 (3), 31–34.
- Rossi, A., 2014. How American Universities Turned Into Corporations. Time. Retrieved from: <http://time.com/108311/how-american-universities-are-ripping-off-your-education/>.
- Sapelli, C., Illanes, G., 2016. Class size and teacher effects in higher education. *Econ. Educ. Rev.* 52, 19–28.
- Slaughter, S., Rhoades, G., 2004. *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. Johns Hopkins University Press, Baltimore, MD.
- Taylor, J.L., Jain, D., 2017. The multiple dimensions of transfer: examining the transfer function in American higher education. *Community Coll. Rev.* 45 (4), 273–293.

Credit for experience, credit recognition and transfer

Jason E. Stone, Educational Leadership and Policy Studies, Oklahoma State University, Oklahoma City, OK, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	343
Credit hours	343
Credit for experience	344
Independent study	344
Work experience	344
Portfolio demonstration	344
Credit recognition	344
Cambridge International Examinations	345
Excelsior College exams	345
Completion of other learning opportunities	345
Non-credit adult classes	345
Transfer credit	345
Europe	346
Asia/Pacific	346
Australia	347
Africa	347
Central, South, and Latin America	348
North America	348
Conclusion	348
References	349
Relevant websites	349

Introduction

Most universities throughout international higher education offer alternative arrangements for earning college credit outside of classrooms. Three popular mechanisms for awarding and documenting that learning are credit for experience, credit from recognition, and transfer credit. Faculty members award credit for experience when they are convinced that a student has engaged in meaningful educational experiences that were similar to or equivalent to the curricular pathway offered through formal matriculation. An example might be the owner of a printing company who returns to college to complete a degree in small business management. The institution might be inclined to award the student credit for life experiences that the student already navigated as a small business owner. The second mechanism for awarding credit for learning that occurs outside a formal curriculum is credit recognition. Several agencies and organizations measure learning experiences and competencies outside of a college credit matriculation regime. Sometimes universities recognize those learning opportunities and credential them as such. The final alternative credit mechanism is transferring credits earned at another University or country to a destination university.

Credit hours

Universities afford students the opportunity to learn curriculum that teaches them general knowledge, marketable skills, and the opportunity to signal to employers that they are qualified and ready for work in their organizations (Tomlinson, 2018). Typically, progress and completion of a path of matriculation to a degree or certificate is measured by the completion of credit hours (Prashun, 2014). The modern credit hour evolved from a desire on the part of universities across North America to standardize high school curriculums in “Latin, history, or algebra” (National Education Association, 1894, p. 17). Some educational historians have used different years to mark the different periods of the maturation of the credit hour (Prashun, 2014). However, at least two significant writings agree that around 1870 through the early 1900s colleges and universities around the United States were dissatisfied with the dramatically uneven college preparation of incoming first-year students. These recent high school graduates had inconsistent preparation for their collegiate studies. This uneven preparation was a source of considerable dissatisfaction for both students and faculty alike. In 1906 the Carnegie Foundation was established. By 1909, the Carnegie Foundation had built on work accomplished by various accrediting agencies. Further, the provision of a potential monthly pension for teachers by the Carnegie Foundation provided encouragement for high schools and colleges to use the Carnegie Unit to measure high school time on task and as a potential unit to measure how much instruction a potential college student received. Both high schools and colleges throughout much of the United States adopted the Carnegie unit to describe, contrast, and compare high school instruction was widely adopted by 1910 (Prashun, 2014). This period from 1895 through 1920 is defined by some educational historians as the age of “standards”, due to

the fact that so many uncertain and unstandardized variables were coordinated and clarified during this period (Hawkins, 1997). Interested readers are encouraged to consult Hawkins for a more detailed sequence of events, including organizations, motivations, and their leaders.

Out of the age of “standards” arose the widespread adoption of the “credit hour”. Universities attempted to formalize work assignments, course work, and semester duration. The “credit hour” sought to ensure that every college graduate engaged in 60 year-long credits. The formalization of two semesters (fall and spring) led to the breaking down of the 60 year-long credits to 120 semester-long credits (Shedd, 2003). The credit hour continues to factor heavily in discussions of the future of higher education both in the United States and throughout much of the world. Discussion of transferability, international access, globalization of higher education, student choice, and the difference between competency-based and credit-hour based matriculation are great examples of how the credit hour continues to shape the educational landscape. In fact, most of the alternative credit awarding systems below exist as an attempt to demonstrate an equivalence to the North American credit hour system.

Credit for experience

The United States Department of Education asserts that, students should beware institutions that seek to offer a degree based solely or mostly on life experience. Legitimate institutions offering “credit for life experience” use a combination of methods to determine how much credit to award. Those methods include examining the results of standardized tests, a prior learning portfolio of work examples, oral exams, past college credit, and professional certifications. While this information assumes a context within the United States, many universities worldwide describe their processes in similar terms. The US Department of Education urges caution from students when engaging with diploma mills who award credit for experience without also using any of the above listed methods. Most reputable universities award some credit for experience. No reputable university would award an entire degree, or most of a degree, based on previous experiences without the demonstration of additional credentials.

Independent study

Faculty members occasionally or frequently teach students with independent studies. These types of courses take many forms. Students engage in extensive study or designed experiences. The students are then asked to write and/or speak about their study, experience, or work. Increasingly these courses are used in service learning, community engagement, home-schooled students, and for other high school students who want to complete learning outside of the formal curriculum (Rafoth, 2003). Over the years independent study at various universities has included the preceding types of experiential courses, as well as more traditional internships, directed readings, honor’s scholar course projects, and thesis or dissertation hours.

Work experience

If a student has worked in an industry or held managerial positions, then that student maybe awarded credit for certain aspects of their “work experience”. These arrangements are increasingly rare. The development and wide-spread availability of tests and standardized curriculum means that students are more routinely awarded credit because they are able to demonstrate a competency, not because they worked in a certain position for a length of time. A student is much more likely to earn college credit by successfully completing the final exam for a marketing course and demonstrating work examples in a portfolio than by being awarded the credit simply because they managed some marketing functions at their previous job. Most universities will award a small number of credits for introductory courses if they award any credit for work experience at all.

Portfolio demonstration

Students can also use a portfolio to demonstrate that previous course work, life experiences, or work experience deserve evaluation for credit. Obviously, a portfolio with tangible or virtual examples makes discussing previous work products easier. Portfolio demonstrations are often required as part of a “creative” degree pathway in art, dance, or painting. Portfolios are often occasionally required for job placement in coding, graphic design, and video production or editing. When students are seeking credit for work accomplished outside of traditional learning pathways, a portfolio can provide examples of work product.

Credit recognition

Many higher education institutions are offering more non-credit adult courses. Some examples of non-credit bearing courses are high school equivalency courses, co-parenting classes for divorcing parents, non-academic certificates and micro-credentials (e.g., badges) in areas like web page development, floral design, phone application development, or safety courses. These types of courses may or may not involve some element of the university’s community outreach.

These courses encompass broad areas of interest and are often associated with community relevance in assisting life-long learners. These types of courses are often much shorter than traditional academic semester or even trimesters. Additionally, the course maybe designed for the primary purpose of transmitting information via video or text without much student assessment or with limited generative assignments. Examples of such courses are the University of Oklahoma’s “Chemistry of Beer”, The University of Madrid’s Calligraphy “The Geography of the stroke” course, or the University of Tasmania’s “Understanding Multiple Sclerosis” course. Much of this non-credit bearing, community-oriented, life-long learning is happening within Massively Open Online Courses or MOOCs. While most of the preceding courses are not typically offered for credit, they are representative of an alternative

educational resource available to students free of charge. These courses are tracked and occasionally certificates of completion are issued, but these courses are rarely if ever treated as equivalent to on-site learning offered at reputable institutions. Many of these classes are more oriented to community service and engagement rather than transcript-eligible credit bearing courses. While there are multiple instances of MOOCs that are taught for college credit (Stone, 2016), those arrangements are increasing rapidly through mechanisms such as the European MOOC Consortium, co-funded by the EU, and Open Universities of Australia. For nations and universities willing to cooperate the opportunity to recognize the education and instruction happening at other universities grows exponentially each day. Some well-recognized examples of credit recognition are briefly summarized below.

Cambridge International Examinations

According to the Cambridge Assessment website, Cambridge International Examinations is a subsidiary of Cambridge University Press and Assessment. The organization claims to be the 'largest provider of international education programmes and qualifications for 5–19 years'. The aforementioned "credit hour" problem and lack of primary and secondary school standardization was handled differently by the faculty and Cambridge. Rather than creating a standardized curriculum at the K-12 level, the faculty at Cambridge set standards for admission and empaneled a group of local examiners to qualify students seeking admission. In 1858, 370 school candidates sat for the examine. Today, over 8 million candidates in 160 countries have sought endorsement of their learning through Cambridge examinations and assessments.

Excelsior College exams

Excelsior College was founded in 1971 by the New York State Board of Regents. The College was originally formed as the Regents External Degree Program with generous grants from the Ford Foundation and the Carnegie Corporation. In 1998 the College reorganized as a private, non-profit, independent institution. In 2001, Regents College changed its name to Excelsior College. During the institution's affiliation with the New York State system for higher education the College provided high quality independent studies and correspondence courses to working and/or place bound adults. Excelsior College provides online and correspondence options for earning college credit in over 55 different courses. Further Excelsior College offers a distance-based independent study in nursing. Excelsior College provides Applied Associates through Master's degrees in nursing. All students graduating from these courses are still subject to state and national accreditation examines such as NCLEX. Certain components of advanced nursing courses require in-person or web-proctored demonstration of skills by the students.

Completion of other learning opportunities

The American Council of Education (ACE) is a trade/service organization that represents the university presidents of member institutions. ACE reviews a variety of educational opportunities provided outside of formal university matriculation pathways and recommends how academic departments and university registrars should regard those learning opportunities. A cursory examination of ACE's "National Guide" reveals a wide array of educational experiences for which ACE recommends college credit. The National Guide is produced with input from faculty members or former faculty members. The National Guide shows five different pathways for credit recommendations: for colleges and universities, for learners, and for training providers, for apprenticeships, and for the military. Each pathway engages different types of learning and different recommendations for earning college credit through that pathway. ACE is specific to education provided in the United States and in a US context. However, because many non-U.S. residents are trained by the military, at colleges and universities, and through corporate learning opportunities, ACE credit recommendations potentially impact learners world-wide. The ACE refers to some of those sources include trainings conducted by the United States military, corporate trainings, executive leadership programs.

Non-credit adult classes

Over 220 million learners world-wide have participated in at least one Massively Open Online Course or MOOC. These non-credit adult education classes are offered for free through various national and international providers. Two of the top three MOOC providers are based in the US Coursera and EdX. The third is FutureLearn, which is based in the UK. More than 950 colleges and universities have participated in or taught a MOOC. To date, around 1500 learners have earned micro-credentials and around 70 learners have earned actual degrees through these "non-credit" course pathways. The overwhelming majority of content available on MOOCs is for free and doesn't involve any formal credentialing of learning. Micro-credentials and formal degrees are akin to "in-app" purchases. The course content exists as an application. To upgrade for credit or for faculty evaluation and faculty directed and supervised assessment, then the student must upgrade to receive a higher level of service and credentials. MOOCs are a part of internationalization and are because they operate internationally they are helping to accelerate trends toward internationalization and harmonization of education (Stallivieri and Boal, 2015).

Transfer credit

Transfer credit remains the most recognized and dependable method of porting credits from one college or program to another. Transfer credit arrangements are either predicated on a shared expectation of contact and teaching hours, rigor, and content or on shared outcome based competencies. Typically, transfer credit has been predicated on the credit hour. College and universities granted credits to transfer students based on recognizing courses already completed by the student. Students earned full or partial credit in furtherance of their degree program based on course work completed at other educational institutions. Increasingly,

colleges and universities are attempting to coordinate with other universities in their country and throughout their continents. The General Agreement of Trade and Tariffs 2003 trade negotiations focused on portability of education. Out of those negotiations sprang the Bologna Process. The Bologna Process is explained in more detail in the next section.

Europe

In 1987 the European Union formed the Erasmus program to enable the mobility of students throughout Europe (Roper, 2007). The program used bilateral and institution to institution contracts to allow 'study abroad' educational opportunities to appear on transcripts as credit counted toward graduation (Froment, 2003). The Erasmus program was replaced by Erasmus University Charter in 2003. Erasmus membership entitles participating European Universities to secure funding that assists in harmonization of content, standards, and assessments across multiple institutions. As of 2021, the efforts of these universities have enabled more than 9 million people to seek instruction at different universities. In addition to serving the needs of millions of Europeans, the Erasmus program has served as a functional model upon which additional programs and efforts to support international harmonization have been based (Roper, 2007).

In 1992 a 6-year pilot program called the European Credit Transfer System was created (ECTS). The ECTS was a project of the Erasmus program. The ECTS adopted a credit-based system that focuses on required workloads to achieve objectives, learning outcomes achieved, and competencies to be developed. The ECTS is based on students earning 60 credits for full-time enrollment in a typical academic year. The ECTS credit system counts time on task away from class. The US credit system counts contact hours between professor and student. The ECTS count all time on task to achieve a certain level of knowledge or proficiency (de Wit, 2007).

Bologna Declaration/Process was signed June 19, 1999. The Bologna Declaration is primarily a European document but does include non-European members. The principle undergirding the Process is that the best way to ensure credit transferability and portability throughout the European Union was to standardize the duration of academic terms and encouraging similar curriculum. Each of these different components have been adopted by signatory nations. Some nations have done more to implement the accord than others. Many European Union accords and or processes are driven more by the EU or by implementing supra-national agencies; however, the Bologna Process is meant to be a standardization that is encouraged by common deadlines and objectives and less by penalties and restrictions (Westerheijden, 2003).

The Bologna process involves four parts. First, signatory nations are to adopt a 3-2-3 structured system. All units assume academic years. The 3-2-3 system means that European students should be able to earn a bachelor's degree in 3 years; a Master's degree in 2 years; and a doctorate in 3 years. Theoretically, up to 80% of the courses and curriculum would be transferable and portable across all of the European Higher Education Area. Second, the Process will further define for employers what type of degree is needed for a specific occupation. Does that job involve understanding and skills that are more associated with a bachelor or master's degree? Third, the Process will rely heavily upon the European Credit Transfer System to encourage and enable transferability of credits and portability of completed degrees. Fourth, the process encourages transnational cooperation in quality assurance (Roper, 2007).

In addition to the Bologna Process, the Erasmus program, and ECTS, Europe also boasts OpenUpEd. Formed in 2013 as a response to for-profit Massively Open Online Courses provider Coursera, OpenUpEd is a not-for-profit European consortium of Open Universities (Schuwer et al., 2015). Because OpenUpEd is affiliated with and provides a service for a number of European universities, the credit earned on OpenUpEd can be more easily transported to universities in other European countries (de Langen, 2018). The curriculum is designed with an international consumption in mind (van Dijck and Poell, 2015). This consortium also involves partnerships with other universities that are not officially considered Open. OpenUpEd now includes almost 20 partner and affiliate universities throughout Europe, and a handful of universities in North America, Africa, and the Middle East as partners (OpenUpEd "Partners", 2022). OpenUpEd is sponsored and hosted by funds from The European Association of Distance Teaching Universities. Europe's efforts at harmonization have served as a model for other continents (Schuwer et al., 2015; van Dijck and Poell, 2015).

Asia/Pacific

Educational harmonization efforts in Asia date back to 1994. Three different credit transfer evaluation and portability systems emerged since that time. University Mobility Asia was an international consortium of governments and universities. This system was largely limited as was not participated in widely. In 2008, the second system developed from a group called "Asia Cooperation Dialog". That consortium involved 18 countries from East Asia, Southeast Asia, South Asia, and the Middle East (Asia Minor). This group was inspired by the success of the ECTS. They proposed an Asian Credit Transfer System. Finally, in 2009 the Association of Southeast Asian Nations put forth the ASEAN Credit Transfer System, now referred to as the AUN-ACTS (Asian University Network—Asian Credit Transfer System (Hotta, 2020).

Each of these three systems involves different scope, signatories, members, and sponsorship. Japan is included in the ASEAN group. Notably, Japan, South Korea, and China account for more than 3000 of Asia's 7800 institutions of higher education. Many regional solutions that fail to include any one of these countries remain partial and incomplete. Beginning in 1997 the ASEAN community began creating organizations that include China, Japan and South Korea, in the "ASEAN Plus Three" framework. Asia's efforts at harmonization have been directed to making study at other Asian universities more attractive and profitable. Asia's educational harmonization efforts have yet to address inconsistent durations of terms and variable definitions of credit hour/contact hours across different countries. The duration of terms varies across Asia with China preferring 17 week terms and Laos using

13 week terms. Recently the ASEAN Credit Transfer System has been focused on articulating with the EU and the ECTS. ASEAN Plus Six efforts include China, Japan, South Korea, Australia, India, and New Zealand. Some educational theorists believe that the ASEAN Plus Six service area is too large, unwieldy, and culturally diverse to be effectively harmonized (Hawkins, 2012).

Australia

Australia remains a destination for international students. The fact that the instruction is in English and Australian Universities have a good reputation internationally encourages students to study at Australian universities (Hotta, 2020). Australia has done much to harmonize its own educational offerings in the past decade. Much of that consolidation has been accomplished with Open Universities and Open2Study. Open2Study was a catchall consortium for Open Universities Australia. Open Universities of Australia now provides Massively Open Online Courses from one clearinghouse of Australian MOOCs. These courses are sponsored by 21 different Australian universities and provide a tremendous amount of flexibility and portability for their students.

In 2021, Australia reaffirmed its integration with ASEAN by signing a comprehensive trade and cooperation framework. Additionally in 2020, Australia formalized member in ASEAN's Regional Comprehensive Economic Partnership. This free trade agreement includes ASEAN member nations plus China, Japan, South Korea, and Australia. The signatories to this free trade agreement account for 30% of global population and 30% of global gross domestic product. Theoretically, Australia's multiple levels of strategic partnership with ASEAN lays the groundwork for an Australia system of higher education that is more harmonized with Asia than any other region. Australian higher education frequently attracts many students from throughout Asia. Australia has also cooperated with universities in China and Japan to create "twin" campuses in both those countries (Yenchken et al., 2021).

Africa

Efforts at higher educational harmonization in Africa have a long, rich history and are ongoing (Woldegiorgis, 2017). Some scholars assert that 'harmonization' in Africa is often used interchangeably with "regionalization" (Knight, 2017). Colonial rule, uneven national educational systems, inconsistent educational opportunities, and linguistic diversity are all challenges to educational harmonization efforts in Africa. Higher levels of priority to harmonize educational systems between and among neighboring and/or friendly Africa states dates back to the post-colonial era of the early 1960s (Mazrui, 1980). A second era of robust harmonization followed the changing of the Organization for African Unity (OAU) into the African Union (Woldegiorgis, 2017). The first intergovernmental conference of African Ministers of Education was held in Ethiopia in 1961. The conference was well attended by governments and nongovernmental agencies. The Ministers crafted and endorsed the Educational Development Plan, a 20-year plan to improve education on the continent. The plan also set practical goals to increasing intra-African educational cooperation, providing a higher education for up to 20% of Africa's high school graduates, and producing more primary and secondary teachers (Woldegiorgis, 2017). Africa produced almost 300,000 teachers in the following 15 years. The following year in Madagascar, the second intergovernmental dialogue was held in September of 1962. This meeting focused on higher education in Africa. The primary emphasis following this conference was on decolonization of the African university and producing a system of higher education that worked for Africans rather than recreating or perpetuating the university systems of western colonizers (Knight and Woldegiorgis, 2017). In 1967, Africa's Education Ministers created the Association of African Universities, the AAU. The AAU would become an important organization promoting African educational harmonization. In 1981 at the Intergovernmental Education Ministers meeting in Arusha Tanzania, the Ministers agreed to a UNESCO recommendation to adopt a "credit transfer" system. Sadly, despite over 20 African countries signing the accord, as of 2016, no country has ratified the Arusha credit transfer accord. Article 2 of the legal instrument articulates criteria for portability of certificates, credits, degrees, and diplomas (Woldegiorgis, 2017). The African Ministers expressed strong agreement with the proposed framework. The Arusha Accords remain salient in discussions of African educational harmonization. The Accords has been revised at least six times since the beginning of the 21st century. The African Education Ministers had another watershed meeting in 1982 in Zimbabwe in 1982. The result of this meeting was an acknowledgment that despite the a nine-fold increase in attendance at African universities and enthusiastic agreement among national Ministers of education, that little actual educational cooperation had yet to occur (Woldegiorgis, 2017). At this meeting the Ministers endorsed a framework of "regional" cooperation that still resonates throughout many African efforts at educational harmonization. Cooperation with close neighbors was considered more culturally, linguistically, and practically effective. Sharing information, creating educational exchanges, and close cooperation is easier with immediate neighbors. Africa has rich linguistic diversity. Linguists estimate that there are 1500–2000 African languages (Heine and Heine, 2000). The linguistic differences between to neighboring countries are less challenging to overcome than the linguistic differences of two countries on different sides of a continent (Woldegiorgis, 2017).

The second robust effort at educational harmonization occurred after the formation of the African Union in the post-2000 era. The primary stated purpose of the OAU was to decolonize African political governance and to eliminate residual forms of imperialism that diminished self-determination and freedom of Africans. The OAU transformed into the African Union (AU) in July of 2002. The AU is primarily interested in creating a more comprehensive unity including economic trade, competitiveness with the global economy, and creation of knowledge-based economies, and creation of African Economic Areas (Adejo, 2001). One of the central organizing principles of the AU is promotion of education. The African Union created a program to establish and support the development of partnerships in Africa called the New Partnership for Africa's Development (NEPAD). NEPAD has set strategic priorities and designated some universities as centers for excellence in certain disciplines. NEPAD has worked to facilitate exchanges and cooperation among bilateral partners. NEPAD has worked to improve access to internet communication technology. Despite significant meeting and communication, most AU and NEPAD harmonization efforts have not been

fully implemented and remain conceptual with little actual cooperation and sharing to date (Woldergiorgis, 2017). Some scholars and administrators alike are calling for an African Secretariat of MOOCs to help with harmonization efforts (Oyo and Kalema, 2014).

Central, South, and Latin America

Harmonization initiatives throughout Central, South, and Latin America are proceeding similarly to efforts in both Africa and Europe. While harmonization with other providers in the Americas remains underdeveloped, some universities are robustly cooperating in exchange and credit transfer arrangements with universities throughout the United States and Europe. Prior to the 1990s most Central and South American universities were state chartered and sponsored educational institutions. Economic structural adjustment regimens undertaken at the behest of the World Bank and International Monetary Fund gripped most of Latin America in the 1970, 1980s and 1990s. Those developments coupled with the United States abandoning the “gold standard” in the early 1970s lead to a time of extreme governmental spending austerity. As a result of decreased spending, many Latin American universities struggled with inadequate funding (Noel, 2009). Subsequently, most new universities chartered in Latin America since the early 1990s have been private institutions (Teodoro, 2015). Many of those private college were religiously affiliated. In recent years, many of those newer universities have been criticized for inconsistent learning and a lack of rigor. Compared to each of the other continents and/or regions discussed here, Latin America has the least amount of resources devoted to research. Individual universities throughout Latin America appear to have more connections to the European Common Educational Area than to other universities in Latin America.

Despite resource limitations, Central, South and Latin America are proceeding with regional harmonization and in cooperation with the European countries. One significant cooperative development is the launch of the Alfa Puentes. Alfa Puentes was a digital platform that provides a public portal for interface with the public and general public, a closed Intranet for internal collaboration between users, mini-sites and pages for the organization of events and special projects. The Alfa Puentes site was a three-year long pilot project to jump start academic cooperation between Latin America and Europe (Clark, 2014). The site also hosted on-board support for multi-languages. The three languages supported by the site were English, Spanish, and Portuguese. Professors in Central America shared syllabi, a degree sheets, proposed course descriptions, a reading lists, assignments, and course content on the site. Group members were able to evaluate policies, share documents, and engage in textual communication with collaborators. The site has six formal partners in Europe, including the EUA, and man partners throughout Latin American. The sunset and concluding reports of the cooperative efforts are available for review.

Several other attempts at regional credit harmonization, regional assumptions for quality assurance and shared criteria for faculty qualifications have been attempted throughout Central, South, and Latin America for the past 35 years (Noel, 2009). The Adean Community, Mercosur, and the Ibero-American Network for the Accreditation and Quality of Higher Education are prominent examples. Initiatives aligned with or cooperating with the EU or UNESCO appear to produce the most results and/or demonstrate the most staying power.

Finally, while regional harmonization is less robust in Central, South, and Latin America, bilateral agreements and cooperation is comparatively more developed. As much as 20% of Central America’s scientific investment is provided by bilateral cooperative agreement with other countries (Svenson, 2013).

North America

From a credit harmonization perspective North America does will continue to be a destination country with as many as 130,000 international students expected to study at universities in the United States by the year 2025 (Macgregor et al., 2017). The top exporter of students to North America is still China. However, after 2025, India is projected to be the largest source of international students. Canada also attracts a large number of students from China, Australia, and France. Canada is the second largest destination market. Canada is the preferred destination of several countries for post-bachelor’s graduate studies (Yencken et al., 2021).

Universities within the United States continue to harmonize with one another and cooperate vigorously with other campuses on other continents, especially in China, Europe, and Latin America. This “twin” campus phenomenon has gained momentum. This arrangement accurately describes the international cooperation between several North American universities and their Chinese “twin” counterpart campus. Credits earned at one institution typically transfer to the other. Recruiting for faculty takes place primarily in North America. Scores of Chinese universities have direct bilateral affiliations with American universities.

Conclusion

Students in both national, regional, and international contexts have many opportunities to earn college credit through extra-institutional arrangements and through credit transfer. The Bologna Declaration and the development of so many designated Higher Education Areas throughout Europe, Asia, Africa, and Latin America are important watersheds in the mobility of student credit transfer regimes. Because ethnic, cultural, and geographic diversity add richness of experience to all educational pursuits, there is little doubt that transnational cooperation agreements, designated cooperation areas, MOOCs, and the desire to create greater potential for mobility within Higher Education will shape how and where students learn in the future.

References

- Adejo, A.M., 2001. From OAU to AU: new wine in old bottles? *Afr. J. Int. Aff.* 4 (1–2).
- Boal, H., Stallivieri, L., 2015. The Influence of Massive Online Open Courses (MOOCs) on the Internationalization Process of Higher Education, vol. 21. Digital Universities.
- Clark, N., 2014. The Bologna Process as a Model for Regional Integration. *World Education News and Reviews*. In: <https://wenr.wes.org/2014/06/the-bologna-process-as-a-model-for-regional-integration>.
- de Langen, F., 2018. Sustainability of open education through collaboration. *Int. Rev. Res. Open Dist. Learn.* 19 (5).
- de Wit, H., 2007. European integration in higher education: the Bologna process towards a European higher education area. In: *International Handbook of Higher Education*. Springer Netherlands, Dordrecht, pp. 461–482.
- Froment, E., 2003. The European higher education area: a new framework for the development of higher education. In: *Higher education in Europe. Enseignement supérieur en Europe. Vysshee obrazovanie v Evrope*, vol. 28, pp. 27–31 (1).
- Hawkins, H., 1997. Toward system. In: Goodchild, L.F., Wechsler, H.S. (Eds.), *The History of Higher Education*, second ed. Simon & Schuster Custom Publications, Needham Heights, MA, pp. 318–332.
- Hawkins, J.N., 2012. Regionalization and harmonization of higher education in Asia. *Asian Educ. Dev. Stud.* 1 (1), 96–108.
- Heine, B., Heine, B., 2000. *African Languages: An Introduction*. Cambridge University Press, Cambridge.
- Hotta, T., 2020. The development of “Asian academic credits” as an aligned credit transfer system in Asian higher education. *J. Stud. Int. Educ.* 24 (2), 167–189.
- Knight, J., 2013. Towards African higher education regionalization and harmonization: functional, organizational, and political approaches. In: Wiseman, A.W., Wolhuter, C.C. (Eds.), *Development of Higher Education in Africa: Prospects and Challenges*. Emerald Publishing Limited, 2013.
- Knight, J., 2017. The Concept and process of higher education regionalization in Jane Knight, and Emnet Tadesse Woldegiorgis. In: *Regionalization of African Higher Education: Progress and Prospects*, pp. 9–28. Brill, 2017.
- Knight, J., Woldegiorgis, E.T., 2017. *Regionalization of African Higher Education: Progress and Prospects*, first ed. Sense Publishers, Rotterdam, Netherlands.
- MacGregor, K., Mihut, G., et al., 2017. 39. Global: Global Postgraduate student mobility trends to 2024. *Understanding Higher Education Internationalization: Insights from Key Global Publications*, pp. 183–183. SensePublishers, Rotterdam, Netherlands. In press.
- Mazrui, A., 1980. *The African Condition: A Political Diagnosis*. Cambridge University Press, Cambridge.
- National Education Association, 1894. Report of the Committee of Ten on Secondary School Studies, with the Reports of Conferences Arranged by the Committee. American Book Company, New York, p. 17, 1891.
- Noel, B., 2009. Education Reform in Latin America: equal educational opportunity? *GIST Educ. Learn. Res. J.* 3, 134–157.
- Oyo, B., Kalema, B.M., 2014. Massive open online courses for Africa by Africa. *Int. Rev. Res. Open Dist. Learn.* 15 (6). <https://doi.org/10.19173/irrodl.v15i6.1889>.
- Prasuhn, F.C., 2014. Credit hours with no set time: a study of credit policies in asynchronous online education. *Am. J. Dist. Educ.* 28 (1), 4–13.
- Rafoth, M.A., 2003. Independent study. In: Guthrie, J.W. (Ed.), *Encyclopedia of Education*, second ed. Macmillan Reference USA, New York, NY.
- Roper, S.D., 2007. European education reform and its impact on curriculum and admissions: Implications of the Bologna process on United States education. *J. Polit. Sci. Educ.* 3 (1), 51–60.
- Schuer, R., et al., 2015. Opportunities and threats of the MOOC movement for higher education: the European perspective. *Int. Rev. Res. Open Dist. Learn.* 16 (6), 20–38.
- Shedd, J.M., 2003. The history of the student credit hour. *N. Dir. High. Educ.* 2003 (122), 5–12.
- Stone, J.E., 2016. Awarding college credit for MOOCs: the role of the American Council on Education. *Educ. Pol. Anal. Arch.* 24 (1).
- Svenson, N., 2013. Central America: the value of international academic cooperation. *Int. High. Educ.* 71.
- Teodoro, A., 2015. European and Latin American higher education between mirrors. Designing possible futures. *Rev. Lusófona Educ.* 31 (31).
- Tomlinson, M., 2018. Conceptions of the value of higher education in a measured market. *High Educ.* 75 (4), 711–727.
- van Dijk, J., Poell, T., 2015. Higher Education in a networked World: European Responses to U.S. MOOCs. *International Journal of Communication* 9, 2674–2692. <https://ssrn.com/abstract=2645629>.
- Westerheijden, D.F., 2003. Accreditation in western Europe: adequate reactions to Bologna declaration and the general agreement on trade in services? *J. Stud. Int. Educ.* 7 (3), 277–302.
- Woldegiorgis, E.T., 2017. Historical and political perspectives in Jane Knight, and Emnet Tadesse Woldegiorgis. In: *Regionalization of African Higher Education: Progress and Prospects*, pp. 29–46. Brill, 2017.
- Yencken, E., et al., 2021. Transnational education provision in a time of disruption: perspectives from Australia. *Int. J. Chin. Educ.* 10 (3), 221258682110691.

Relevant websites

Cambridge International Examinations, <http://www.cambridgeassessment.org.uk/>.

Excelsior College Exams, <http://www.excelsior.edu/exams/>.

Information About ECTS and Erasmus, <http://eua.eu/>.

MOOC Data, <http://www.classcentral.com/>.

New Partnership for Africa's Development, <http://www.nepad.org/>.

U.S. Department of Education, <http://www.ed.gov/>.

The disruption of Black history and critical pedagogy in international higher education curriculum

Miriam Chitiga^a, Kolawole Samuel Adeyemo^{b,c}, and Carmen Wade^d, ^aWarner Pacific University, Portland, OR, United States; ^bDepartment of Education Management and Policy Studies, University of Pretoria, Pretoria, South Africa; ^cCenter for Studies in Religion and Society, University of Victoria, Victoria, BC, Canada; and ^dCommunity College of Aurora, Aurora, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	350
Decolonial intersectionality	351
Instructional leadership in decolonizing the curriculum	352
Curriculum and Blacks in higher education	353
Concluding thoughts	354
References	354

Introduction

Curriculum reform, which began in the late 1960s as a systemic change method, targets both subject integration and classroom pedagogy in higher education systems. In Africa, the decolonization of the higher education curriculum, described as opposing domination of the curriculum in which a particular knowledge paradigm exists and is perpetuated, has been widely discussed in the broader literature (Cebekhulu et al., 2006, p. 88; Etieyibo, 2016, p. 404; Masaka, 2019). To interrupt this practice, it is necessary to acknowledge and accept there are competent and authentic knowers and knowledge elsewhere that exists beyond the dominant perspective, in which the traditional curriculum is saturated (Keet, 2014, p. 25). This dominant ideology leads to the attainment of epistemic justice. On the contrary, epistemic injustice is a result of doubting persons to be reliable knowers, which is typically done on prejudicial grounds (Masaka, 2019). Masaka (2019) presents the argument of epistemic justice which ought to surpass the parameter which leads the way in awakening agency of learners in how they contribute to the determination of their own destiny. In Africa, it is suggested that higher education has suffered so much that it surpasses the notion of exclusion of the indigenous people's knowledge paradigms which could be used in responding to prevailing circumstances (Masaka, 2019).

The topics of curriculum change took dimensions of colonization, decolonization, neoliberalism, and black studies in higher education in West Asia, North Africa, and Tunisia (Mullin, 2018). In relation to neoliberal policy, Uganda like many other countries has reformed its curriculum focusing on what students learn and this approach is mainly based on the outcome of knowledge (Alti-nyelken, 2010). The outcome-based approach to curriculum reform has failed learners in Africa (Gumede and Biyase, 2016) because it ignored both the process and contexts in the curriculum transformation processes. Chigbu et al. (2017) evaluated land governance degree programs in universities in Botswana, Ghana, Nigeria, Rwanda, Kenya, Mauritius, Uganda, Namibia, South Africa, and Tanzania and found capacity building as the key to achieving curricular reform in Africa. The capacity building perhaps has helped many Chinese and Asian universities recently to internationalize their higher education curriculum to be among the top universities in the world compared to Latin American universities (Zezele, 2012). Aside from capacity building, the internationalization of higher education curriculum and curriculum reform should fit into the contexts and address issues of epistemic injustice.

A more substantive transformation and decolonization of the higher education curriculum in Africa, for example, ought to go further, to focus on empowering the learners to take a lead in offering solutions to the problems and challenges that countries in Africa face. This means that the mere recognition and acceptance of previously unacknowledged epistemological paradigms—such as that of the indigenous people of Africa—into the epistemological canon are a necessary step to epistemic justice, but not the ultimate goal. Restoration of epistemic status to the paradigms that have previously been considered as lacking epistemic credibility may not lead to authentic epistemic justice (Masaka, 2019). This new curriculum was to be student centered and focused with an emphasis on an investment of students in their learning, instilling a sense of “full, shared ownership” of the curriculum through “respectful, thoughtful, and open-minded listening” (Kuh, 2001). The basic tenets of the new curriculum were connection and infusion. For instance, the history of curriculum changes in South Africa and Germany is different because of the different social issues that the reforms targeted (Chisholm, 2015). In South Africa and the United States, the focus of curriculum change was to address issues of racism, inequality, and epistemic injustice in Historically Black and White Institutions or Universities (Luckett and Shay, 2020). The critical question for Germany is whether the curriculum reform that has taken place between 1990 and 2010 has been able to redress the inequality of the Nazi racial curriculum that taught racism in schools and fostered hatred for Jewish, for instance. For African and Historically Black Universities and Colleges in the United States, the reforms are expected to recognize ways of knowing for different minority and racial groups.

In this chapter, curriculum reform is defined as the process of developing new courses and curricula across all disciplines based on a critical analysis of existing knowledge on various intersecting identities, such as racism, class, gender, disability, religion, sexual orientation, gender expression, and nationality (Schmitz, 2012). There is idea that the new curriculum should place gender and

diversity at the core of the academic program by integrating race, ethnicity, gender, and sexuality into existing courses (Kuh, 2001). Higher education should be exploring the multitude of barriers that are counterproductive to student success and are actively exploring effective strategies to overcome these challenges (Chitiga et al., 2019). As a result of the challenge, a new curriculum was explained in the context of higher education and within the decolonial intersectionality framework.

Decolonial intersectionality

The work discussed in this chapter is grounded in the theoretical framework of decolonial intersectionality. Intersectionality is grounded in the Black Women Feminist Movement and discusses and serves as a key paradigm for the analyses of identity, oppression, and privilege. According to many scholars and researchers, intersectionality has its roots in the Combahee River Statement of 1977 and Kimberlé Crenshaw's analysis of structural and political dimensions of intersectional oppression (Harris and Watson-Vandiver, 2020). The Combahee River Statement is an essay that was written by Black feminists that challenged the White feminism and the Civil Rights Movements as they did not see themselves reflected in those struggles, and as a result, did not promote the liberation of Black women (Harris and Watson-Vandiver, 2020). In the Combahee River Statement of 1977, the collective of Black feminists identified Black feminism as the "logical political movement to combat the manifold and simultaneous oppressions that all women of color face" (p. 264).

Within the "colonial curriculum," there lies the presence of domination which limits the capacity and potential of liberation. Decolonial theorists describe the conventional appropriations of intersectionality as rooted in the coloniality of power, knowledge, and being. Kurtis and Adams (2016) discusses the modern global order, rooted in the harmful legacy of racialized colonial violence that allowed Euro-American global domination, which perpetuates the traditional curriculum. According to Kurtis and Adams (2016), there is a need to disrupt the current domination of the marginalized many for the benefit of the privileged few. A decolonized intersectionality calls upon and elevates the voice of the silenced perspectives, knowing, and being of people in Majority-World spaces to "propose sustainable ways of being consistent with global social justice" (Kurtis and Adams, 2016, p. 47).

In the context of social justice, Chitiga (2019) discussed the purpose and consequences of Historically Black Colleges and Universities HBCUs as it relates to the miseducation of Black communities with emphasis on governance, management, leadership, the academic core, curricula exposure, student development, socialization, and cultural exposure. According to Carter G. Woodson, "African Americans have been educated away from their own culture and traditions, attached to the fringes of European culture, and dislocated from themselves.... African Americans often valorize European culture to the detriment of their own heritage" (Asante, 2003, p. 37). The White, Anglo-Saxon, Protestant, oppression-cleansed narrative is propagated in higher education curriculums while non-White perspectives are marginalized, disrupted, or removed (Harris and Watson-Vandiver, 2020). Henry A. Giroux, *In Pedagogy and the politics of hope*, argues that historical illiteracy is both an "engrained and intentional phenomenon that is a by-product of knowledge control and that engagement with pre-colonial history is transformative in our own self-actualization experiences" (Harris and Watson-Vandiver, 2020, p. 4). Historical illiteracy is counterproductive to the self-liberation and self-actualization of Black people as discussed by Carter G. Woodson (originally published in 1933) in *The Miseducation of the Negro*.

If you can control a man's thinking you do not have to worry about his action. When you determine what a man shall think you do not have to concern yourself about what he will do. If you make a man feel that he is inferior, you do not have to compel him to accept an inferior status, for he will seek it himself. If you make a man think that he is justly an outcast, you do not have to order him to the back door. He will go without being told; and if there is no back door, his very nature will demand one (p. 192).

Because much of African history has not been told or told from a White perspective and overshadowed by Western thought, these stories have been omitted and underrepresented in HBCU curriculums. Chitiga (2014) argues that "many people are not equipped with the necessary knowledge of American political history, democratic institutions, processes, and civic life needed to allow them to become active, effective, responsible, and empowered citizens and leaders of the future (p. 59)". Weber (2010) asserts that "full citizenship rights, for inclusion in the political process, and thus for an equal voice in public policies has been long and difficult" (p. 121). Furthermore, Weber (2010) discusses that "participation in the political institutions of society defines groups' relationships to the policy enforcement institutions (law, criminal justice, the police, and the military)" (p. 121). "Control over the political domain also increases control in the ideological and economic domains by increasing power over the production of controlling ideologies about subordinate groups and by increasing access to greater economic resources" (Weber, 2010, p. 121).

In America for example, because of the overrepresentation of white, heterosexual males in Congress, this group is able to pass legislation that supports the economic dominance (tax legislation, business legislation, etc.) of this particular group. Resources are directed to support this group, along with their constituents. As a result of miseducation, omission, and underrepresentation, there are other domains that are affected that may seem less obvious, such as health. Weber (2010) argues that because resources are typically directed to those that have an elevated voice in decision making, such as in healthcare, the life expectancy and disease rates of communities tend to be higher in those communities of color where there are less people of color on city council. As students learn more about themselves and "participate in developing the pedagogy of their liberation" (p. 48), with emphasis on the democratic proposals of problem-posing education where "men and women develop their power to perceive critically the way they exist in the

world with which and in which they find themselves; will they then come to see the world not as a static reality but as a reality in the process of transformation (p. 12)" in which they can bring about social change (Freire, 1996).

Navigating Whiteness and White patriarchy is a skill that must be a part of the higher education curriculum for students of color to fully understand the world around them and their place in that world, which has led to a decline in Black historical consciousness. Decolonizing the higher education curriculum calls for a form of academic resistance into the White narrative that perpetuates much of higher education institutions, and as a result, the learning experience of many students of color as the counternarrative. Furthermore, it is necessary for higher education globally to examine their own implicit biases, deep-rooted assumptions, and beliefs as it relates to its history and commitment to critical Black pedagogy.

Instructional leadership in decolonizing the curriculum

Transformational leadership is vital to the longevity of public and private institutions of higher learning, specifically, HBCUs in fulfilling their missions and upholding their historical values (Chitiga, 2018). Furthermore, transformational leadership in organizations is necessary to promote, support, and ensure equity, inclusion, and social justice in the interrogation of past policies, practices, and ways of knowing and being. Transformational leadership is at the forefront of driving this change (Chitiga, 2018). According to Neophytou and Valiandes (2013), the curriculum should not be seen as a document that is preplanned with content, outcomes, learning objectives, assignments, activities, and the plan to get there. On the contrary, the curriculum is a living document that serves as a lived experience of the people infused with their meaning and interpretation of that experience as it relates to their knowing and being (Aoki, 1991, 14, cited in Joseph, 2011a,b).

Furthermore, this curriculum should be intentionally, carefully, and thoughtfully created, to include a specific set of ideas that elevate the voice and agency of the knowers. As a result, a curriculum cannot be implemented as a mass of ideas or content, but must be carried out in a diminutive manner, centering the voice of the authentic knowers. The "old curriculum" may be thought of as a cooking recipe where the ingredients, instructions, and therefore, the final product, are predetermined. However, transforming and decolonizing the curriculum requires a co-creation of learning where the knowledge and capacity of the knowers emerges in strength-based approaches to set the path for learning. The higher education then provides the platform for critical Black pedagogy as the Black student then develops the agency to create the learning necessary for liberation. This requires a shifting of power where the student has agency in the ingredients and the instructions of this recipe, which means a different product (of liberation) possibly for each student.

Transformation leadership is an impetus for decolonizing the curriculum. Critical literacy in higher education is necessary for effective transformational leadership in decolonizing the curriculum. Critical literacy supports a holistic meaning of the world in which we live. As an example, texts in critical literacy are designed to facilitate productive dialog as "vehicles through which individuals communicate with one another using the codes and conventions of society" (Robinson and Robinson, 2003, p. 413). However, the challenge then becomes applying critical literacy in a field where there is a saturation of anti-critical monologue (Robinson and Robinson, 2003). Furthermore, critical literacy empowers students and teachers to learn together in a co-creation of knowledge. There is no longer a "sage on the stage." As a result, social constructs are reconstructed with new meaning. Lefkios (2013) describes a critically literate individual as one who has a solid competency in power dynamics and relationships and most importantly, is an intellectual leader who initiates the transformation of a society for liberation of a better future. Leithwood et al. (2004), Wahlstrom et al. (2010), and Hallinger (2011) discuss the importance of defining the school's mission, providing the necessary supports to faculty and staff to accomplish high quality teaching, organizing the school appropriately, and managing the instructional program (Brazier and Bauer, 2013, p. 650).

It is important that higher education institutions are supported in this transformation of leadership as critically literate instructional leaders. It must be acknowledged that many higher education institutions have been socialized according to the "old curriculum" and there is a "renewing of the mind" necessary for effective implementation of this "new curriculum" discussed here. There must be space provided for reflexive practice to hone the skills required for an instructional leader that is facilitating the Black consciousness of students for social change through critical Black pedagogy. Heifetz and Laurie (1997) as cited in Chitiga (2018), posits that transformational leaders strive "to engage in confronting the challenge, adjusting their values, changing perspectives, and learning new habits" (p. 134). Brazier and Bauer (2013) discussed the need for a reimagining of leadership preparation for instructional leaders on the basis of pedagogical content knowledge, educational connoisseurship and educational criticism in the context of leadership.

As noted above, students, teachers, and teacher trainers have already been deeply socialized by prior "banking" models (Shor, 1999). In the banking model, students are asked to uncritically accept information by the individual who possesses the knowledge, the "sage on the stage." The opposite of the banking model is the "problem posing" concept, discussed by critical theorist Paul Freire in *Pedagogy of the Oppressed*, where students are presented with a problem affecting society and are challenged to critically evaluate the problem from a cultural and social perspective (Chitiga, 2014).

In the current climate that many higher education institutions are experiencing, an increase in demands, but a lack of capacity, it is imperative that spaces and opportunities be created for sharing and collaboration among colleagues to improve the quality, vitality, and relevance of transforming curriculums in demand-driven spaces. Because some higher education institutions are required to teach more, research more, and administer more, the World Wide Web (WWW) also offers a plethora of support, research, and resources to assist teachers in their development as instructional leaders equipped for transformational leadership

that the decolonization of the curriculum requires (Engel, 2016). Engel (2016) suggests creating online spaces (e.g. Zoom rooms, Google Meet, Canvas Conferences, Facebook Hangouts) for teachers to share helpful resources, such as subject outlines, lectures, tutorial ideas, simulations, games, readings, and other helpful information. Contributions could be made under a creative common license to allow for modifications and accommodations for those that contribute a required amount of resources, as decided by the group, based on the teacher, institution, and student needs and requirements (p. 564).

According to Chitiga (2018), “transformational leadership is central to empowering leaders and subordinates to mutually engage with each other to ensure the transformation of organizations” (p. 60). For teachers to have the capacity and know-how to empower their students, teachers must also be empowered by their trainers. Those that have the responsibility and opportunity of training teachers must also be acknowledged as transformative leaders of transformative leaders that are leading other transformative leaders, their students (Neophytou and Valiandes, 2013).

Curriculum transformation is directly related to transformational leadership. Transformational leaders support the holistic development of students and are committed to the positive learning experience of students that ensure positive student outcomes, specifically, increased student achievement, which is an impetus for the growth of the institution (Chitiga, 2018). For transformational leadership to decolonize the curriculum, it is necessary for instructional leaders to transition from the banking model, anti-critical monolithic model of education, to a more self-authored, self-defined, student-centered, problem-based learning approach, where students’ agency is increased as the authentic knowers of knowledge based on their lived experiences which will disrupt the dominant narrative found in many curriculums at higher education institutions (Brazier and Bauer, 2013, p. 680). Instructional leadership goes beyond passing tests and achieving minimum standards, but requires critical reflection, the interrogation of pedagogy and practices, leadership knowledge, skills, dispositions, and courage, conducive for transformational leadership. This means continuing down an iterative path of inquiry and continuous improvement as it relates to teaching and learning (Brazier and Bauer, 2013, p. 647).

Curriculum and Blacks in higher education

According to Neophytou and Valiandes (2013), “it (curriculum transformation) is primarily a multi-layered socio-cultural, political, and ethical process (Joseph, 2011a,b) requiring active comprehension and artistry. It is human oriented and human necessitating: it aims toward people and needs people to gain its substance” (p. 415). Engel (2016) suggests adding to the “old curriculum” by adapting current offerings to the “new curriculum” with a focus on student-centeredness as the student as consumer.

As a case, this chapter contextualizes the HBCU within a crucial historical epoch—the 1940s–1960s Civil rights and Black consciousness movement that led to significant legislative changes that were intended to advance the rights of Black people in the United States. It argues that sometime after the Civil Rights Act of 1964, most HBCUs lost sight of their role in preparing critically conscious students. Given the current 2020 revitalization and re-examination of the “civil rights” fight, the chapter provides insights for HBCUs to reinvestigate their relevance by proposing some bold changes to the total curricula.

The majority of HBCUs were founded after the US Civil War (1861–65). While most private HBCUs were formed by Christian church organizations, many were chartered by states led by developed by White politicians, with the support of business owners who largely sought to use education as a tool to socialize African Americans to be submissive and docile workers. As HBCUs began to evolve, they became sites of empowerment for African Americans that produced Civil Rights, Black Power, and Pan African leaders. Until the Reconstruction period (1865–1877), African Americans were prohibited from attaining an education (Weber, 2010, p. 164). Historically Black Colleges and Universities (HBCUs) were created to promote liberation of African Americans through education.

Currently, HBCUs “graduate 25% of all African Americans earning bachelor’s degrees, produce over 30% of Black engineers, 50% of Black elected members of Congress, and 80% of Black federal judges. Furthermore, one-third of Black Americans with a doctorate in science, technology, engineering, or math (STEM) fields earned their undergraduate degrees from HBCUs” (Strayhorn, 2020, p. 7). According to the [Hundred-Seven.org](#), there are currently 107 (56 private/51 public) HBCUs in the United States of America, serving approximately 228,000 students. In *A Pledge of Allegiance to America’s Historically Black Colleges and Universities: Key Priorities of the Biden-Harris Education Agenda*, Dr. Terrell Strayhorn (2020) asserts that “HBCUs have long since been physical manifestations of Black excellence, cultural repositories of the Black lived experience in America” (p. 7).

HBCUs have supported the social and economic mobility of African Americans through education. Weber (2010) posits “the institution of education supports both the political and economic structures of society, but at its core it is an ideological institution, intended to create and transmit the ideas on which society is organized and which will support its continuation” (p. 145). George Kuh asserts that higher education remains more important than ever as social and political issues become increasingly complex (Kuh, 2001). According to Harris and Watson-Vandiver (2020), decolonial education is a force of activism with education’s role serving as an equalizer for change. Dr. Hakim Lucas, President and Chief Executive (CEO) of Virginia Union University proclaims that HBCUs offer students’ academic programs and experiences where they grow and develop cultural, spiritual, economic, and political sensibility which sets them apart as global leaders (Strayhorn, 2021, p. 6). As such, colleges, HBCUs, are “critical in the reproduction of social class and racial advantage” (Weber, 2010, p. 180) and the critical Black pedagogy. The responsibility of HBCUs to support the development of the Black consciousness movement through critical Black pedagogy by decolonizing the higher education landscape for the empowerment of 21st century African American students can be found in mission statements of HBCUs from across the country.

Spelman College, located in Atlanta, Georgia, was founded April 11, 1881, as an Atlanta Baptist Female Seminary by Sophia B. Packard and Harriet E. Giles. According to its mission statement, Spelman College asserts to empower its students to engage the many cultures of the world while inspiring a commitment to positive social change. Similarly, Morehouse College, also in Atlanta, Georgia, founded February 14, 1867 by Reverend William Jefferson White, Reverend Richard C. Coulter, and Reverend Edmund Turney, declares their mission is to assume special responsibility for teaching the history and culture of black people. Xavier University of Louisiana, located in New Orleans, Louisiana, founded October 6, 1925 by Saint Katharine Drexel and the Sisters of the Blessed Sacrament avows its mission to contribute to the promotion of a more just and humane society by preparing its students to assume roles of leadership and service in a global society. Tuskegee University, located in Tuskegee, Alabama, founded July 4, 1881, by George Campbell and Lewis Adams, states their mission is to produce public-spirited graduates who are both competent and morally committed to public service with integrity and excellence. Although not necessarily specifically stated in terms of Black students, Tuskegee reiterates its commitment to producing graduates that are prepared to serve with “sensitivity” and uplift their country through the high leveled training received at the institution. Hampton University, located in Hampton, VA, founded September 17, 1861 by Benjamin Butler and Mary Peake, emphasizes its focus on the liberal arts in the development of character. North Carolina A & T, located in Greensboro, North Carolina, founded March 9, 1891 by the North Carolina General Assembly, discusses in its mission statement its commitment to develop students that have the capacity and know-how to create innovative solutions that address the challenges and economic needs of North Carolina, the nation and the world. Howard University, located in Washington, DC, founded March 2, 1867, by General Oliver O. Howard, specifically states their mission to provide an educational experience of exceptional quality with a focus on educational opportunities for Black students. In addition, Howard University states their commitment to developing “historically aware” students that are equipped to solve human problems in the United States and throughout the world. Claflin University, located in Orangeburg, South Carolina, founded in 1869 by Methodist missionaries, Dr. Alonzo Webster (the university was named after Lee and William Claflin), states its dedication to providing a student-centered, liberal arts education grounded in cutting-edge research, experiential learning, state-of-the art technology, community service, and life-long personal and professional fulfillment.

To reinvigorate their fading relevance, HBCUs must recommit to raising the level of Black consciousness in preparing students to be change agents through critical Black pedagogy. This requires bold changes to the total curricula which involves a decolonization of the curriculum, where culture, student-agency, and understanding of the broader socio-political context that is directly connected to Black history emerges from the institutional curriculum, not a single course. This must be the focus so that students are prepared to be the public-service, social change agents described in the HBCU mission statements above. Black students must see themselves in the world as they help to solve problems of the world for their own liberation. Demonstrated in the mission statements, the emerging themes of raising the consciousness of Black students through the education of Black history rings true throughout many of our HBCUs across the country. With HBCUs housing many Black students, in decolonizing curriculums, our HBCUs have a great opportunity to deliberately empower its students through critical Black pedagogy that raises Black consciousness to promote and empower social change.

Concluding thoughts

The idea in this chapter is that the curriculum should not just focus on skills (McCarthy, 2015) or outcomes but should also promote inclusivity in multiracial societies (Banks, 2015) and close gender gap (DiPrete and Buchmann, 2013) in higher education. Thus, curriculum transformation in higher education should call for the development of inclusive perspectives within interdisciplinary fields such as women’s studies, ethnic studies, queer studies, deaf studies, and studies of multiple different abilities. Teaching these studies in higher education curriculum, throughout academic programs, ensures that all students graduate with a comprehensive knowledge of diversity, inclusion, equity, and social change across disciplines. To support the learning, identity development, self-actualization, and liberation of marginalized communities, culturally relevant critical Black pedagogy, that includes a curriculum that recognizes, appreciates, and celebrates the culture, histories, and aspirations of all groups, is vital for 21st century learners, to the development of Black critical consciousness. Holistic student development and academic engagement depends on a decolonized curriculum where students can see themselves in the world through a problem-posed education. Diverse students, teachers, and faculty are needed to deconstruct power in the classroom so that learning can be co-created through productive dialog.

References

- Altinyelken, H.K., 2010. Curriculum change in Uganda: teacher perspectives on the new thematic curriculum. *Int. J. Educ. Dev.* 30 (2), 151–161.
- Aoki, T.T., 1991. Inspiring Curriculum and Pedagogy Talks to Teachers.
- Asante, M.K., 2003. The Afrocentric idea in education. In: *Afrocentricity and the Academy: Essays on Theory and Practice*, pp. 37–49.
- Banks, J.A., 2015. *Cultural Diversity and Education: Foundations, Curriculum, and Teaching*. Routledge.
- Brazer, S.D., Bauer, S.C., 2013. Preparing instructional leaders: a model. *Educ. Adm. Q.* 49 (4), 645–684.
- Cebekhulu, E., Mantzaris, E.A., Cebekhulu, E.N., 2006. Not yet Uhuru! power struggles in a neo-liberal university. *Alternation* 13 (2), 153–173.
- Chigbu, U.E., Tenadu, K.W.A.M.E., Mwasumbi, A.G.N.E.S., 2017. Curriculum reform in land governance education: the need for transforming existing curricula in Africa. In: *World Bank Conference on Land and Poverty*. March, Washington DC, pp. 20–24.
- Chisholm, L., 2015. Curriculum transition in Germany and South Africa: 1990–2010. *Comp. Educ.* 51 (3), 401–418.

- Chitiga, M., Kaniuka, T., Ombonga, M., 2019. How do millennials learn?: implications for higher education pedagogy. *Int. J. Inf. Commun. Technol. Educ.* 15 (1), 29–41.
- Chitiga, M., 2014. Performing arts for effective civic engagement: developing creative civically engaged student leaders. *Int. J. Civ. Engagem. Soc. Change* 1 (3), 59–74.
- Chitiga, M., 2018. An exploration of the implications of transformational leadership behaviors on empowering employees and transforming organizations. *Afr. J. Publ. Aff.* 10 (1), 60–82.
- Chitiga, M., 2019. What is Curriculum Transformation: How Does Curriculum Transformation Benefit the Student? Curriculum Transformation Seminar, University of South Africa.
- DiPrete, T.A., Buchmann, C., 2013. *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools*. Russell Sage Foundation.
- Engel, S., 2016. Curriculum reform: a transformation or consumption model for politics and international relations? *Aust. J. Polit. Sci.* 51 (3), 555–567.
- Etieyibo, E., 2016. Why ought the philosophy curriculum in universities in Africa be Africanised? *S. Afr. J. Philos.* 35 (4), 404–417.
- Freire, P., 1996. *Pedagogy of the Oppressed*. M. Bergman Ramos, Trans.. Penguin Books, London. England.
- Gumede, V., Biyase, M., 2016. Educational reforms and curriculum transformation in post-apartheid South Africa. *Environ. Econ.* 7 (2), 69–76.
- Hallinger, P., 2011. Leadership for learning: lessons from 40 years of empirical research. *J. Educ. Adm.* 49 (2).
- Harris, L., Watson-Vandiver, M., 2020. Decolonizing race and gender intersectionality in education: a collaborative critical autoethnography of hope, healing and justice. *J. Cult. Anal. Soc. Change* 5 (2), 1–16.
- Joseph, P.B. (Ed.), 2011a. *Cultures of Curriculum*. Routledge.
- Joseph, P.B., 2011b. Conceptualizing curriculum. In: *Cultures of Curriculum*. Routledge, pp. 19–38.
- Keet, A., 2014. Epistemic “othering” and the decolonization of knowledge. *Afr. Insight* 44 (1), 23–37.
- Kuh, G.D., 2001. Assessing what really matters to student learning inside the national survey of student engagement. *Change* 33 (3), 10–17.
- Kurtis, T., Adams, G., 2016. Decolonial Intersectionality. In: *Intersectional Pedagogy: Complicating Identity and Social Justice*, p. 46.
- Leithwood, K., Louis, K.S., Anderson, S., Wahlstrom, K., 2004. *How Leadership Influences Student Learning. Review of Research*. Wallace Foundation, The.
- Luckett, K., Shay, S., 2020. Reframing the curriculum: a transformative approach. *Crit. Stud. Educ.* 61 (1), 50–65.
- Masaka, D., 2019. Attaining epistemic justice through transformation and decolonization of education curriculum in Africa. *Afr. Ident.* 17 (3–4), 298–309.
- McCarthy, M., 2015. Reflections on the evolution of educational leadership preparation programs in the United States and challenges ahead. *J. Educ. Adm.* 53 (3).
- Mullin, C., 2018. *Resisting the Slow Violence of the North African and West. Asian University*.
- Neophytou, L., Valiandes, S., 2013. Critical Literacy needs teachers as transformative leaders. Reflections on teacher training for the introduction of the (new) Modern Greek language curriculum in Cyprus. *Curric. J.* 24 (3), 412–426.
- Robinson, E., Robinson, S., 2003. *What Does It Mean?: Discourse, Text, Culture-An Introduction*. McGraw-Hill.
- Schmitz, B., 2012. Curriculum transformation, higher education. In: Banks, J.A. (Ed.), *Encyclopedia of Diversity in Education*, vol. 1. SAGE Publications, Inc., pp. 590–593. <https://www.doi.org/10.4135/9781452218533.n184>
- Shor, I., 1999. What is critical literacy? In: Shor, I., Pari, C. (Eds.), *Critical Literacy in Action: Writing Words, Changing Worlds*. Heinemann, Portsmouth, NH, pp. 1–30.
- Strayhorn, T.L., 2020. Sense of belonging predicts persistence intentions among diverse dental education students: a multi-institutional investigation. *J. Dent. Educ.* 84 (10), 1136–1142.
- Strayhorn, T.L., 2021. A Pledge of Allegiance to America’s Historically Black Colleges and Universities: Key Priorities of the Biden-Harris Education Agenda. Center for the Study of HBCUs, Richmond. Available at: <https://www.luminafoundation.org/wp-content/uploads/2021/02/a-pledge-of-allegiance.pdf>. (Accessed 3 May 2021).
- Wahlstrom, K.L., Louis, K.S., Leithwood, K., Anderson, S.E., 2010. *Learning From Leadership: Investigating the Links to Improved Student Learning*. The Informed Educator Series. Educational Research Service.
- Weber, L., 2010. *Race, Class, Gender, and Sexuality: A Conceptual Framework*.
- Zezeza, P.T., 2012. *Internationalization in Higher Education: Opportunities and Challenges for the Knowledge Project in the Global South*.

Non-completion amongst black working-class students in South African universities: the dangers of a single story

Mukovhe Masutha, Ali Mazrui Centre for Higher Education Studies, University of Johannesburg, Ackland Park, Johannesburg

© 2023 Elsevier Ltd. All rights reserved.

Introduction	356
Chimamanda Ngozi Adichie: the danger of a single story	357
Carol Bacchi's WPR approach	358
Method	358
Findings	359
The <i>untold sides</i> of the <i>single story</i> representations of the dropout challenge in HE	359
Linda's story	359
Homeless and new in town	359
Thrown into the deep end	359
The language constraint: "... it was a different experience for me having to hear someone trying to teach me in English throughout the lecture"	360
Lost in a crowd	360
Persistence: down but not out	360
Let down by the National Students Financial Aid Scheme	361
Wadzi's story	361
Teachers' viewpoints: "NSFAS needs to go beyond just providing funds. NSFAS needs to look at the broader context"	362
Applying Bacchi's WPR approach to representations of the dropout problem amongst financial aid funded students	363
Conclusion	364
References	365

Introduction

Low rates of retention and completion amongst students from traditionally excluded communities has been a long standing challenge for higher education (HE) globally (Aina et al., 2018; Love, 1993; Vignoles and Powdthavee, 2009). In the United Kingdom, black students are more than 50% more likely to drop out when compared to their white and Asian counterparts (Bulman, 2017). A report by the University Partnerships Program Foundation (Brabner, 2017) also found worrying rates of non-completion amongst those marginalized students targeted by widening participation initiatives in British Universities. A similar trend can be observed in the United States where the Hechinger Report (Sanchez and Kolodner, 2021) found white students to be far more likely to graduate than their African American and Hispanic students in public universities. In Nigeria, a protracted labor relations dispute between academics and the federal government has pushed many students out of the university system (Salako, 2022). Moreover, the disproportionate impact of the Covid-19 pandemic on black communities is thus likely to worsen what has already but an untenable condition for marginalized communities in HE.

In order to mitigate against the exclusion of students from traditionally marginalized communities, governments have identified various student financial aid models to facilitate equitable HE participation with mixed results. The American Pell Grant (Ishitani, 2020), the UK's income-contingent student loan model (Callender and Jackson, 2004) and Kenya's Higher Education Loans Board (HELB) (Mobegi and Kara, 2022) are examples of such interventions. Prominent amongst these financial aid models has been the cost-sharing principle of income contingent student loans (ICLs), a form of student loan that enables fee-free access to HE at point of access with deferred repayment due once graduates begin earning. In South Africa, the democratic government established the National Students Financial Aid Scheme (NSFAS), an ICL scheme aimed at improving the participation and representation of poor and working class students in HE.

Nearly three decades since the end of apartheid, despite bold widening participation targets, by the year 2020, only 5.9% of South Africans between the ages of 25 and 64 had achieved a bachelor's degree from a university (DHET, 2021). When assessed in terms of participation by race, the South African picture remains far from equal, with the university participation rate amongst African youth at 18% compared to 56% amongst their white counterparts (DHET, 2021). This level of inequality deepens when one looks at which universities and degree programmes remain (in)accessible to which categories of students and graduates from which university. The picture gets bleaker when we look at the question of who graduates? And who graduates from where? (Masutha and Naidoo, 2021).

Despite NSFAS's strides, low student retention and completion rates remains a major challenge to South African higher education's pursuit for just and equitable educational experiences and outcomes (Angu, 2019; Bernard, 2018; de Villiers, 2012; Kelly-Laubscher et al., 2018; Masutha and Naidoo, 2021). When access to HE is understood to mean more than the demographic

composition of the student population, regrettably, and of particular concern in this article, the growth of financial aid funded students has occurred alongside very high non-completion rates amongst this group of students. The first ministerial review of the NSFAS scheme revealed that by 2010 (DHET, 2010, p. 51) “67% of 656,000 financial aid-funded students were no longer in higher education; 72% of this 67% dropped out without completing their studies; and only 28% graduated” (DHET, 2010, p. 51). In their 2018 study on factors affecting the success of First-Generation students at a South African university (Kelly-Laubscher et al., 2018), observed how patterns of access, retention and success in South African higher education continue to reflect the country’s patterns of class and racial inequality. They found that:

... even for those participating in higher education, the throughput of African and Coloured students is lower than their white counterparts, with only 20 and 24 per cent of African and coloured students, respectively, graduating within the regulation time for their diplomas/degrees, compared to 44 per cent of white students ... It is clear from these statistics that increased participation by these groups has not resulted in a corresponding increase in graduation rates and throughput.

Kelly-Laubscher et al. (2018, p. 98).

While the challenge of non-completion remains prevalent amongst students from all walks of life, it has particularly paralyzed the HE aspirations and experiences of financial aid funded students, all of whom come from poor and working-class households. In this article, “poor and working class student” refers to a student from a household with an annual income of up to R350 000.00 (21 000 USD) as designated by the South Africa’s National Students Financial Aid Scheme (NSFAS, 2019). These are children of families of unemployed parents, those surviving on social security grants, farm workers, mineworkers, security guards, and entry level public servants (DHET, 2021).

Student alienation, Poverty, financial pressures, mental health issues, a student debt crisis, mismatch between students’ expectations and their subsequent reality of HE are some of the common factors attributed to persistent disparities in student retention and completion rates in HE (Greenfield, 2021; Letseka and Breier, 2008; Moodley and Singh, 2015). However important these factors are, the picture remains incomplete without the voice and narrative accounts of those students who “dropped out” and those who experienced teaching this group of students. Given that students who dropout often disappear without a trace, conventional theories of student non-completion tend to lack the voice and lived realities of students who dropout and academic staff who have experienced teaching this group of students. We know little about the side of the story of those who “dropped out” of HE and those who experienced teaching this group of students. Current representations of the dropout problem are thus incomplete and mainly from the perspective of dominant groups in HE. This article contributes toward addressing this gap in knowledge.

Drawing on data from narrative interviews with financial-aid funded students (who dropped out of university), academic and support staff (who shared their experiences to teaching and interacting with this group of students), this study contributes a counter narrative to the dominant and unjust single story prevalent in common conceptions of non-completion amongst underprivileged students in HE (Flint, 2021; Mallman et al., 2021; Mills and Gale, 2011; Owusu-Kwarteng, 2020). Indeed, continuous critical exploration of this challenge in teaching and learning has its basis and justification on the personal, family, institutional and community-wide costs involved in the educational journeys and aspirations of students from marginalized households and communities being cut short prior to attaining their qualifications. The revolving door of black working-class non-completion in HE also undermines the very idea of (higher) education as a vehicle that should contribute to more just and less stratified futures.

The next section of the article outlines of how Carol Bacchi’s WPR (What’s the Problem Represented to be) approach (Bacchi, 2012) and Chimamanda Ngozi Adichie’s notion of *Danger of a single story* (Adichie, 2009) inspire an alternative and arguably more just approach in disrupting fragmentary and unjust ways in which the dropout “problem” amongst financial aid funded students in South African universities has been *represented* in mainstream narratives, and the adverse implications of such *single story problematization* on marginalized students’ experiences of teaching and learning in HE. This is followed by a thematic analysis (Braun and Clarke, 2006) of participants’ narrative accounts and the lessons that flow from a close up analysis of participants’ first person accounts. For the rest of the article, in addition to illuminating the limits of financial aid in insulating poor and working-class students from unequal and marginalizing HE experiences, this article contends with the increasingly dominant and largely taken-for-granted equal-opportunity theories of teaching and learning that inform the *single story problematization* of low retention and completion rates amongst marginalized students. To the extent that power is the position to not only tell the story of another person, but to make it the definitive story of that person, this article takes back that narrative power and foreground the voices and lived experiences that often go unheard as valid, valuable and critical in theorizing students’ experiences of non-completion in HE.

Chimamanda Ngozi Adichie: the danger of a single story

... to create a single story, show a people as one thing, as only one thing, over and over again, and that is what they become ... It is impossible to talk about the single story without talking about power. Like our economic and political worlds, stories too are defined by the principle of nkali (inequality): how they are told, who tells them, when they’re told, how many stories are told, are really dependent on power ... Power is the ability not just to tell the story of another person, but to make it the definitive story of that person ... The consequence of the single story is this: It robs people of dignity ... I would like to end with this thought: that when we reject the single story, when we realize that there is never a single story about any place, we regain a kind of paradise.

Adichie (2009).

The excerpt above comes from *The Danger of a Single Story*, a TEDTalk presentation that Nigerian novelist Chimamanda Ngozi Adichie delivered in 2009 (Adichie, 2009). Adichie's TedTalk has since been watched 10 million times on Youtube.

Adichie skillfully affirms those who have long embraced counter storytelling "as both a method of telling the story of those experiences that have not been silenced and as a tool for analyzing and challenging the privileging of the majoritarian narratives and making explicit the extent to which they "act as a camouflage for the self-interest, power, and privilege of dominant groups" (Solórzano and Yosso, 2002a). Adichie urges us to get closer to the lived realities of those whose stories are often distorted and silenced in order to understand the dangers of certain people and their experiences being packaged into a single story by those in positions of power. In naming their reality, counter-storytellers provide those in the margins of society with "psychic and emotional barriers against the damage caused by majoritarian stories that, for instance, blame people who are targeted by racism for their own subordination" (Love, 2004).

Carol Bacchi's WPR approach

Bacchi's WPR approach is premised on the idea that what one advances as a solution to certain challenges reveals what that person or institution thinks is the problem or needs to change, and that "policies and policy proposals contain *implicit* representations of what is considered to be the 'problem' ('problem representations')" (Bacchi, 2012, p. 21). For (Bacchi, 2009) "there are no problems separate from the proposals purported to address them", and that in our analysis of various "problem representations" should pose the following questions: (a) *What's the problem represented to be in a specific policy or policy proposal?*; (b) *How has this representation of the problem come about?* (c) *what are the rationales and deep seated assumptions behind the representation of the "problem"*; (d) *What is left unproblematic in this problem representation? Where are the silences? Can the "problem" be conceptualized differently?*; (e) *What effects are produced by this representation of the "problem"?* (Bacchi, 2012, p. 21).

The end goal is to re-problematize "problem representations" inherent in certain policy statements and to subject them to critical scrutiny. Given that certain "problem representations benefit the members of some groups at the expense of others", our task as researchers is "to challenge problem representations that have these deleterious effects, and to suggest that issues could be thought about in ways that might avoid at least some of these effects (Bacchi, 2009). As a researcher, I am committed to this goal. I believe the manner in which the challenge of non-completion amongst financial aid students has been represented is unjust and contains such deleterious effects.

Despite South African HE still operating under the shadow of its colonial self, the following excerpts by South African policy-makers get us close to *the seat of the trouble* (Woodson, 1969), they exemplify the victim-blaming and deficit *single story* through which the "problem" of non-completion amongst financial students is currently "represented" in South African Universities:

In responding to the dropout challenge, the Minister of Higher Education and Training expressed his failure "*understand why people would want to fail twice or even more*" given the "*huge number of pupils who passed matric with flying colors, but sadly cannot afford to go to university*. Is that fair?" he asked (Ncwane, 2020). More recently, his department announced that it was looking to amend the funding progression rules wherein financial aid funded students would be required to pass at least 75% of their modules in order to qualify for continued financial aid (Mkhwanazi, 2021). Echoing the Minister, Professor Adam Habib, former Vice Chancellor and Principal of University of the Witwatersrand, characterized the dropout rate as "*a shocking waste of human talent*" (News, 2016).

The above assertions share rationales and deep seated assumptions that have informed the *single story* through which students who dropout of HE have been theorized as "early leavers from HE" (Davies and Elias, 2003) due to their "course dropping behavior" (McKinney et al., 2019). As (Mallman, 2017) observed, this victim-blaming not only looks at half an already distorted picture, it also induces feelings of inadequacy and illegitimacy, leading some students to view their failure as inevitable. It is therefore important that we critically examine the single story because of its roots in *structures of power* in higher education institutions (HEIs) and its influence in policy making circles (Bacchi, 2000). In doing so, it is important that we remain alive to the ways in which the single story is sometimes deployed to inform and justify punitive forms of pedagogy that obscures the exclusion of vulnerable groups of students (Allen, 2020).

Method

This article reports on a section of my doctoral study that examined narrative accounts of working-class students' experiences of completion and noncompletion across three different status universities in South Africa (a traditionally white university, a traditionally black university and a post-apartheid merger university), and how insights from such experiences can help move understandings of (under)achievement rates amongst financial aid funded BWS beyond the single story problematization. By adopting a qualitative narrative approach to studying the challenge of student non-completion in HE, this study further advances a growing body of researchers who foreground the voice and lived experiences of the unheard as valid, valuable and critical in theorizing of persistent inequality in education (Arbouin, 2018). Stories are not just stories, they are fundamental to illuminating key dimensions of our lived experiences (Hinchman and Hinchman, 1997). It is through storytelling that we as educators come to know what they

know about teaching and learning experiences and this in turn makes narratives one of the most effective methods of understanding and representing educational experiences (Connelly and Clandinin, 1990).

A total of forty-six in-depth narrative interviews with twenty-four financial aid funded black working class graduates and drop-outs and twenty-two academic and support staff were conducted at three different status universities (Traditionally Black University, Traditionally White University and Merger University). A combination of purposeful sampling, maximum variation sampling and snow-balling techniques were employed (Neuman, 2000).

Findings

The untold sides of the single story representations of the dropout challenge in HE

Narratives included in this article are not representative of the entire dataset, I seek not to generalize participants' experiences but rather to illuminate key dimensions of their narrative accounts as they relate to representations of the problem of non-completion amongst financial aid students. Linda and Wadzi are financial aid funded dropouts from Merger University, an urban university based in South Africa's economic hub. I lifted their stories because they highlight the many sides of dropouts that are not captured and in many ways silenced in dominant conceptions of teaching and learning in HE. Narrative accounts of Professor Kigali (Ivory Tower University) and Professor Ngoye (Merger University) were also purposefully selected for analysis because they worked closely with undergraduate financial aid students. Participiant's names have been replaced with pseudonyms throughout this article.

Linda's story

Linda was born and raised in a rural village in Kwa-Zulu Natal Province. In his region there is hardly any economic or industrial activity taking place, unemployment levels are high, and the vast majority of households are dependent on subsistence farming and government's social security grants for pensioners and mothers with children under the age of 18. Neither his parents nor any of his six sisters went to university, making him the first in his family. By the time Linda was six, his mother left him with his six sisters to work as a domestic worker in Johannesburg, about 600 km from their village. His mother sadly passed away in his first year of study at Merger University, leaving him and his sisters orphaned.

Even though Linda knew no one who had ever been to university and very little about what actually happens at the university, he always *"wanted to go to varsity so that obviously the only picture you have in your head is to change your family situation, to change those houses that are made of mud and everything and just, you want a better life ... You want to make your parents proud. And that was the main reason obviously"* he said.

Homeless and new in town

Linda experienced a turbulent transition and hard landing into HE which was due to a variety of hindering factors. His first hindrance had to do with his uncertain sources of funding and lack of familiarity with the funding opportunities available to people of his socio-economic status. He admits that after completing his secondary education he had no idea what to do next. Due to his funding situation, hardly halfway into the first semester of his first year of study, Linda found himself homeless in a town he knew no one. He explains:

LINDA ... we {him and his secondary school mates} had no idea on what to do, how to get to varsity. The only thing we were given were application forms There were no bursaries and I will tell you now that I was one of the students who didn't know how to apply for NSFAS or if there are any other bursaries except NSFAS.

It's funny because I went there (University), I knew no-one, I had no siblings, I had applied for accommodation but I had to pay some fees, a certain amount of money which I felt like I can't pay because I have to pay monthly. Since I haven't got funding, I can't do that. So, I went there and I have no place to stay. So, I think I spent a week or so living at the taxi rank and going to varsity. It was in the early days of varsity still dealing with the registration.

Thrown into the deep end

His lack of familiarity with the university and its institutional cultures and practices was worsened by his non-existent portfolio of social capital. The excerpt below illustrates how Linda experienced the first few weeks upon arrival at the university:

MUKOVHE: ... *When you first arrived at University, what did you know about the University?*

LINDA: To tell you the truth, I had no idea what was expected of me. I mean, I spent the whole first semester without a tutor. Most students were assigned to tutors and I wasn't because I had no idea there are people called tutors ... To tell you the truth, I think the reason why I didn't know about tutors and everything was because I missed out on most of my classes because I was getting lost on campus trying to find classes and everything.

I spent like first two weeks without a timetable because you had to make one for yourself. They say you first have to look into a computer and then ... I had no idea. So, that's two weeks wasted. I have no idea. I just go there to varsity, look around, geez, I don't know what's going on, until some of the students helped me create one ... that was the most painful part, knowing that someone could have helped me with other things, but I had no idea.

The language constraint: "... it was a different experience for me having to hear someone trying to teach me in English throughout the lecture"

Inside the classroom, Linda's struggle with the English language hindered both his confidence and ability to engage with what was being taught. It was particularly challenging for him to transition from having been taught mainly in his home language in school to the university where lectures were delivered exclusively in English. He explained this frustration:

LINDA: And now that I got to varsity, my first lecture was in English. There I was struggling to understand, even scared to ask questions because obviously I was nervous ... it was a different experience for me having to hear someone trying to teach me in English throughout the lecture. I was there trying to hold those words that were difficult to me, trying to jot them down so that I can later on get some sort of explanation and definitions on that.

The English problem then weakened Linda's ability to mobilise and accumulate much needed social capital and access student support services, a reality that further weakened his transition and integration period.

LINDA: It took me some time to make friends and by then I feel like it was already late because I had missed out on a lot of things. And I think University was just confusing for me. I mean, there were buses that I didn't even know about them, the bus is taking students from varsity to town ... I would have saved a lot {of time and money} but I didn't know about that. I only caught up on the second semester.

Lost in a crowd

While struggling to find his foot on campus, Linda battled with an institutional culture he found to be "strict", "rough" and "confusing". Thrown in the deep end, he felt "university has no time" for someone with his slow pace of figuring things out, pace of learning, weak knowledge of technology, etc.

LINDA: I would say that university is so strict and it's even confusing if you're a first-time student in varsity. It's so confusing. Like you get lost and no-one will help you. It's like you are there to study, so you got to figure it out yourself. They don't care if you know the University life or not, you're just there to study. You have to know everything fast and it wasn't easy for me.

So, I was still struggling using emails, I remember that was the first time of having an email and I didn't know that the communication in varsity is done through emails, then even if you're going to have a test, you're going to be told via email and everything is like that. I created my email, but I hadn't checked it and only to find out that when I asked, when I was talking to some of the kids that I was in class with, they were like, "we were writing a test yesterday", and I'm like where and when, how? They told me that they told us through email and that was the first time I opened my emails and only to find out that I've missed two of the tests because I wasn't able to use email and email was just a nuisance for me.

Linda found the university to be impatient and quite isolating to struggling students like him. He constantly sought to highlight to me the evident gap between his learner and cultural identity and the university's expectations of him:

LINDA: And that was, it was so difficult. I mean ... they have no time for excuses, they have no time to hear your side of the story. They want you to be there, they want you to be serious. In fact, when you go to varsity it's like you need to grow up. I don't know, they expect you to be this responsible individual who knows it all, who's capable of catching up fast and who's able to, I don't know, to be responsible enough to be there.

Persistence: down but not out

When Linda's financial aid was disrupted, he was forced to leave the university at the end of his first year of study. He did not quit his higher education altogether, instead he went back to work as a security guard in Johannesburg, saved some money and enrolled at a University the following year.

LINDA: And then I ran out of money. I spent a week not going to school until I called my mom and told her, no, let us discuss, I can't do it, I have to come back because I've run out of money, and she was like, "okay, cool, you can do that because I cannot support you either". Then this was the only option to drop out and I can't live here where I have no family and I have no money, so it's better to be in Johannesburg closer to my family and go to a nearer University.

Let down by the National Students Financial Aid Scheme

Having enrolled for a degree in Public Management and Governance at Merger University, Linda was failed by the National Students Financial aid scheme. With his mother late, Linda's household income was now zero. NSFAS' administrative deficiencies over the past years has resulted in the scheme approving more students that it had funding available, and Linda became one of the victims of this ineptitude. He cries about how his NSFAS funding application status remained "provisionally funded" throughout the year and yet his tuition debt remained on his student account at Merger University. He explains:

I was provisionally funded and by then it was around March. It was around this time when I went to ask them, and they told me that I was provisionally funded, but then because of the fact that the funding was not enough, they couldn't fund us. But it was provisionally funded, we had already qualified for funding, but we can't get it because there's not enough funds.

As a result, Linda's higher education journey crash landed due to NSFAS's administrative deficiencies. Despite his funding constraints, Linda continued attending classes at Merger University under the hope that NSFAS would eventually come through. It did not. At the end of the academic year Linda's academic records were withheld and he was denied registration the following academic year due to "*outstanding fees*". And that is how Linda was pushed out of the field of higher education.

Toward the end of our interview, I asked Linda to look back at his higher education experience and identify that which he wishes he knew about university prior to admission. He wished he known about the significance of "*orientation week*" in aiding transition into higher education, he wished he had made friends earlier and he wished he had known about the help offered by university tutors.

LINDA: I wish I managed to attend the orientation whereby I was going to be able to understand the campus better. But I couldn't be there because I was from far and obviously, I had to find a place to stay, which I failed to get, and I missed out on the orientation. But I wish I had understood the campus better, I had known most of the things about the campus life and what was expected of me as a student. And I wish I was able to make friends earlier and I wish I had a tutor.

Linda's humble beginnings, his grossly under-resourced schooling experiences, turbulent transition and hard landings on campus are consistent with all dropouts in this study (i.e Martin, Naledi, Tali and Wadzi). In this case, their stories were very similar.

Linda's story is also unique and does not fit into the stereotype of unteachable students who are "underprepared" for HE. Each chapter in his life story illuminates extraordinary dimensions of his journey to and through HE, particularly his possession of what (Yosso, 2005) calls aspirational, navigational and resistant capital. His ability and willingness to persist with his studies despite what appeared to be insurmountable barriers is inconsistent with the victim-blaming single story that is central to dominant theories of working-class non-completion in HE. His experience is wholly inconsistent with the simplistic and deficit *single story* representations of the "problem" of working-class experiences of non-completion in HE. Not only did he have to transform himself in order to be enculturated into the field of HE, he was let down by NSFAS when he needed it the most. The deficit single story representation of those who dropout of HE as failing financial aid students who need exit out of the system to make space for others is not only devoid of the extraordinary circumstances he endured, it is particularly silent on the largely inequitable institutional environment and its lack of preparedness to serve the educational needs of a diverse student population as some of the drivers of hindrances that Linda experienced on campus.

Wadzi's story

Although born and raised in very different parts of the country, Wadzi's story is very similar to that of Linda. Wadzi was born and raised in Hillbrow in a largely degenerated and working-class section of the city of Johannesburg. efore her passing, Wadzi's mother worked as a security guard while her father sold food to manual labourers for a living in order to pay rent and send Wadzi to school.

Once registered at Merger University, Wadzi's transition into higher education continued to be haunted by her outstanding NSFAS funding status. She could not measure her academic progress because the university, as part of their standard practice, withheld the grades of students whose tuition fee was outstanding. Toward the end of her first academic year at Merger University, Wadzi and the rest of NSFAS provisionally funded students were informed that NSFAS at Merger University has run out of funds and that they must try the following year. This effectively pushed Wadzi out of the university. The following year her NSFAS funding was once

again approved but she was prevented from registering until she cleared “her historic debt”. Unable to come up with the money, Wadzi was forced to drop out of Merger University. Until this day she has no idea how she performed academically as her grades remain withheld due to outstanding fees.

The University’s refusal for Wadzi to register for her second academic year left her father particularly devastated. *“So my dad was sick and tired of explaining to people. So he was quite disappointed. At some point I think he felt like he’s failing.”*

Their disappointment was worsened by how, whilst in ques for help at NSFAS offices on campus, Wadzi recalled witnessing fraud, favouritism and general disorganisation. A class mate informed her of a guy who was helping students secure NSFAS funding provided she brought a R1 000 bribe. She failed to raise the bribe amount and the girl who paid the bribe was sorted ... *“she went there and then now she’s studying; she’s doing her third year, she got NSFAS”*.

Today, given her experience with NSFAS at Merger University, Wadzi is not just disappointed that her educational journey crash-landed, she wants nothing to do with NSFAS.

You can’t even count on NSFAS. Don’t trust it ... I discovered that my problems were not even big problems. When I go to talk to people, people thought of bad things, some people even think of committing suicide because they’re failing. So one student’s mum was working as a domestic worker and then she failed, she failed dismally, and then at some point she thought of taking her life because she can’t face the disappointment. We all have our own stories to tell.

(Wadzi, Dropout, Merger University).

Teachers’ viewpoints: “NSFAS needs to go beyond just providing funds. NSFAS needs to look at the broader context”

Professor Kigali, a Lecturer at Ivory Tower University, reflected on the irony of the institutional violence directed at students like Linda who are excluded from the university despite their willingness to persist with their studies:

So, for me that {pushed dropout} presupposes certain things. It presupposes that this person is not teachable. That is the bottom line. And then the question becomes, who is not teachable? It presupposes that we have done everything that we can do as an institution to teach this person and when have we ever done that? So, fundamentally I would be opposed to this idea of exclusions because those two things are not evident, that a person is not teachable and that we have done everything that we can to teach but funny enough, even that exclusion thing is violence. That’s it. It is symbolic violence that results in real violence because you’ve conferred an identity on this person; you are not teachable. I mean, conferring that identity you are sentencing them to something, to a life where they’re not able then to meet the demands of the market society that we exist in.

(Prof Kigali, Ivory Tower).

Professor Kigali added that:

I think we also need to be careful about the values that are in our curriculum, the unspoken values in the curriculum that we take for granted I think even when we assess students, we are not just assessing how much they have memorised the contents, we are assessing how well they are able to produce it in the language that we want ... I think we take for granted that our pedagogy is not just the readings that we prescribe. Our pedagogy is the values that we promote.

(Prof Kigali, Ivory Tower).

Professor Ngoye sought to highlight the side of the financial aid student non-completion that is often silenced and thus not forming part of the single story representation of the problem of student non-completion. As a Lecturer in both historically black universities (HBUs) and historically white universities (HWUs), Prof Ngoye reflected on the varied ways (inadequacies) the NSFAS scheme was failing its beneficiaries and thus achieving the opposite of its stated objectives. HE felt that the scheme was not comprehensive enough:

if you look at the {financial aid} package as well, of what they approve ... If you travel from a rural province like your Limpopo, your Eastern Cape, your KZN, your Free State, you arrive in Johannesburg, you need accommodation, you also need food for you to survive. You need stationary, you need books, you need all sorts of things. You also need an iPad because our Universities require that by the time you arrive here you will be able to type assignments and those types of things ... NSFAS needs to go beyond just providing funds. NSFAS needs to look at the broader context. For various learners come from various backgrounds. They don’t just need funding, but they also need advice and guidance. And that is what NSFAS is not seeing. A roadshow in my view can be very useful. All NSFAS students are given funding and also being given some kind of orientation into the University life”.

(Prof Ngoye, Head of School, Merger University).

Prof Ngoye further identified NSFAS’s administrative deficiencies and mismanagement to adversely financial aid students’ transition, adjustment and integration at higher education institutions. Given that the academic year in South Africa commences in February, the delayed payment of financial aid packages was highlighted as some of the factors adversely affecting students’ transition and odds of completion.

With regard to the dropout rates of NSFAS recipients ... generally NSFAS applications are processed very late. At this point in time when we have commenced term two but you still have students who were supposed to have registered in January, NSFAS only approved their funding right now. It's better if the funding is secured and confirmed at least at the beginning of the year already, so that you start the term ... some people are from rural areas, they still need to learn how to use Blackboard (university's online platform), they still need to learn how to type their assignments and most of those interventions happen at the very beginning of the year. So, if you come now, you're assisted by a tutor to catch up. You've already missed the term's work and that really creates problems. If they can allocate the resident fee at least while they are processing the rest ... Its April now and there are some people who have just gotten an approval from NSFAS a week ago. They have not registered. They were not in the system. They will catch up with all the term's work

(Prof Ngoye, Head of School, Merger University).

Participants' (Linda, Wadzi, Prof Kigali and Prof Ngoye) narrative accounts present a composite counterstory that illuminate some of the key issues that are often silenced, omitted or ignored in dominant single story representations of the "problem" of high rates of non-completion amongst financial aid students in South African HE. The thematic thread highlighted across all these stories is consistent with the findings of the first Ministerial Review Report (The Balintulo Report, 2010) of the NSFAS Scheme during the year 2010. In addition to an array of administrative inefficiencies, mismanagement and corruption, the Balintulo Report found that:

NSFAS's major shortcoming is that the amount of funding available, despite steady annual increases, falls far short of demand. Current estimates are that NSFAS has only half of the funds it would need to meet the demand for student financial aid from qualifying applicants ... Some institutions choose to enhance the principle of access by awarding some NSFAS funding to all qualifying students resulting in students having to settle the difference between the award and the actual cost of studying.

DHET (2010, p. 73).

This practice, also known as "*top slicing*", resulted in many students incurring historic debt and therefore prevented from registering in order to proceed with their studies. Top-slicing also maintained financial disparities between expensive historically white universities (HWUs) and under-resourced and working-class serving historically black universities (HBUs) whose cost is far lower than the former. And yet, such critical reflections on the limitations of the scheme are often left out in the colorblind and meritocratic theories of the challenge of student non-completion in HE.

Applying Bacchi's WPR approach to representations of the dropout problem amongst financial aid funded students

When some of the key questions in Bacchi's WPR approach are applied to the student dropout problem, we get closer to seeing the limitations of the stereotypical *single story* through which the problem of students' non-completion has largely been represented. In advancing his failure "*to understand why people would want to fail twice or even more*" and bemoaning such expectations as *unfair*, we identify what the problem is represented to be. The Minister represents the problem as "*people {who} would want to fail twice or even more ... people who do not want to complete their courses and graduate in record time*". The proposed solution, as seen in the Minister's statements above, is that students who fail to pass a specified number of modules (e.g 75% of their modules in an academic year) should be academically excluded in order to make space for others to benefit and to make the system "fair".

Here we can see that the Minister's statement is anchored on a flawed and increasingly taken-for-granted idea of universities as equal-opportunity and meritocratic spaces (Gillborn, 2019; Ladson-Billings, 2006). The punitive side of the Minister's rationale (i.e. forced dropouts) is consistent with what (Drake, 2022) referred to as the criminalization of failure in education spaces. The rationales and deep seated assumptions behind this representation of the problem of non-completion has been thoroughly and correctly critiqued and discredited as mythical and the basis of perennial patterns of inequality in HE attainment (Arbouin, 2018; Boliver et al., 2021; Gallagher, 2003; Littler, 2017; Lucido et al., 2019). Writing for *The Conversation* (Lucido et al., 2019), identified the "the myth of meritocracy in higher education" as a key factor behind persistent inequality of access and outcomes to higher education. They added:

The most damaging myth in American higher education is that college admissions is about merit, and that merit is about striving for – and earning – academic excellence. This myth is often used as a weapon against policies like affirmative action that offer minor admissions advantages to low-income students and racial and ethnic minorities. The deck is stacked in favor of affluent parents who use their privilege and exploit these institutional needs to find their children a way into elite colleges.

Lucido et al. (2019).

As Ladson-Billings (2015) cautions, we often have plenty of verifiable facts about the extent of racial and class disparities in educational experiences and outcomes, but little in the way of the context in which we can make real sense of those facts. The colour-blind meritocracy single story serves to explain *what is wrong with financial aid funded students and their interaction (non-completion)*

with south african universities while silencing the question *what is wrong with the South African field of university education and its interactions with financial aid funded?* (Jackson and Moore, 2008). The role of South Africa's largely untransformed educational institutions with well known histories of excluding black working-class students (Naidoo, 2004) is left *unproblematic* and *silenced*.

As foregrounded in Wadzi and Linda's story, the student non-completion problem can be *conceptualized differently*. Are South African universities able to serve students from diverse social origins and with diverse social and educational needs? If not, what can universities improve their teaching and learning methods to address this gap in pedagogic relations? By looking beyond what this group of students is perceived to lack, this article makes visible the power imbalances and injustices that result from reducing marginalized groups through a single story and stereotypes (Adichie, 2009).

With much of the debate around inequality of access having focused on who gets to access HE in general (Boliver, 2013), suggest that we expand the scope of the debate from whom goes to university to who goes to which university? Who studies what at which university? This is particularly important given that inequality of access has been particularly striking in highly selective higher education institutions, with elite universities largely remaining a natural educational destination and preserve for children of the privileged middle and upper classes (Reay et al., 2001). As seen in participation's counterstories, what is also left unproblematic, despite being evidently problematic, is the impact of NSFAS's underfunding, administrative deficiencies, mismanagement and the adverse consequences of these institutional barriers on teaching and learning.

Conclusion

In this article, inspired by Carol Bacchi's WPR (What's the Problem Represented to be) approach and Chimamanda Ngozi Adichie's notion of *Danger of a single story*, i set out to challenge and hopefully disrupt a dangerous deficit, victim-blaming, colourblind and meritocratic *single story* through which the "problem" of high non-completion rates amongst financial aid funded students in South African HE has been *represented* amongst scholars, media and policymaking circles, and the implications of such *single story problematization* on experiences of teaching and learning in HE. Consistent with Adichie (2009), I conclude that how the stories of student non-completion are told, who tells them, when they're told, how many stories are told, are really dependent on unequal power relations prevalent within the field of HE.

Power is indeed the ability not just to tell the story of another person, but to make it the definitive story of that person (Adichie, 2009). Consistent with (Arbouin, 2018; Love, 2004; Yosso *, 2005), I call on researchers and policymakers working in unequal and complex contexts to look beyond the single story through which marginalized students are framed in majoritarian narratives. By centering Wadzi and Lida's story, this article takes back this power and foregrounds the voices and lived experiences that often go unheard as valid, valuable and critical in theorizing students' experiences of non-completion in HE.

Consistent with (Boliver et al., 2021; Carnevale et al., 2020; Littler, 2017) and others, I further advance that both formal and epistemic justice will remain elusive for as long as dominant notions of teaching and learning remain anchored on the mythical equal opportunity, victim-blaming, colourblind, meritocratic master narrative. As Wadzi and Lida's stories show, their lived realities in the field of HE were far from neutral or equal. New voices are needed to tell multiple versions of the stories of those who remain in the margins of HE's institutional logics (Griffin et al., 2014; Mills and Gale, 2011; Solórzano and Yosso, 2002b). This article is part of their counter-story.

In the words of Chimamanda Adichie:

Start the story with the arrows of the Native Americans, and not with the arrival of the British, and you have an entirely different story. Start the story with the failure of the African state, and not with the colonial creation of the African state, and you have an entirely different story ... I've always felt that it is impossible to engage properly with a place or a person without engaging with all of the stories of that place and that person.

The approach adopted in this study shows the value of listening to stories of those who often disappear without a trace from our campuses. Their stories are rich with lessons for educational researchers, policymakers and practitioners involved in transforming participation in South African HE and elsewhere in the world. Indeed this should fundamentally further problematising narrow and exclusive theories of low working-class retention and completion in HE that are devoid of their voice and those who have experienced teaching them. Further research needs to go into the social and education destinations of students who are forced out of the HE system despite their willingness to continue with their studies (Mathebula, 2019). makes a case for HE researchers and educators to recognize poor black youth as both givers and takes of knowledge or "epistemic contributors", and that doing so constitutes an ethical response to the structural inequalities that limit equitable university access and participation. Given the challenging circumstances embedded in participants' stories, I hope that participating in this study was at least of therapeutic value.

As shown in the work of (Brabner, 2017; Bulman, 2017; Sanchez and Kolodner, 2021), Linda and Wadzi's experiences are not unique to South Africa or developing countries. Marginalized educational transitions and intersectional disparities in rates of HE attainment are a global phenomenon. Given the acceleration of global inequality across class, gender and ethnic lines, South Africa being one of the most unequal societies in the world, can be seen as an extreme case of global patterns, and thus offers important insights for other contexts.

References

- Adichie, C.N., 2009. The danger of a single story. TED Talk. Available at: <https://sch.rcschools.net/ourpages/auto/2017/7/24/35784355/danger%20of%20a%20single%20story.pdf>.
- Aina, C., Baici, E., Casalone, G., Pastore, F., 2018. The economics of university dropouts and delayed graduation: a survey. Available at: <http://ftp.iza.org> <https://papers.ssrn.com/abstract=3153385>.
- Allen, K., 2020. "When does hot become cold": why we should be disrupting narrow and exclusive discourses of success in higher education. Available at: <http://hdl.handle.net/1959.13/1419047>. uon:37406; ISSN:2203-8841.
- Angu, P.E., 2019. Understanding voices from the margins: social injustice and agency in first-year students' literacy narratives. *J. Furth. High. Educ.* 43 (8), 1152-1162.
- Arbouin, A., 2018. *Black British Graduates: Untold Stories*. Trentham Books. Trentham Books.
- Bacchi, C., 2000. Policy as discourse: what does it mean? Where does it get us? *Discourse: Studies in the Cultural Politics of Education* 21 (1), 45-57.
- Bacchi, C., 2012. Introducing the 'What's the problem represented to be?' approach. In: *Engaging with Carol Bacchi: Strategic Interventions and Exchanges*, pp. 21-24.
- Bacchi, C.L., 2009. Analysing policy: what's the problem represented to be? Pearson.
- Bernard, T., 2018. In: Bangeini, B., Kapp, R. (Eds.), 2017. *Negotiating Learning and Identity in Higher Education: Access, Persistence and Retention*. Bloomsbury, London. Available at: <https://doi.org/10.24085/jsaa.v6i1.3074>. *Journal of Student Affairs in Africa*, 6(1).
- Boliver, V., 2013. How fair is access to more prestigious UK universities? *Br. J. Sociol.* 64 (2), 344-364.
- Boliver, V., Banerjee, P., Gorard, S., Powell, M., 2021. Reconceptualising fair access to highly academically selective universities. *High Educ.* Available at: <https://doi.org/10.1007/s10734-021-00755-y>.
- Brabner, R., 2017, July 19. Study finds Ethnic minority and poorer students more likely to drop out of English universities. UPP Foundation. Available at: <https://upp-foundation.org/study-finds-ethnic-minority-poorer-students-likely-drop-english-universities/>.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qual. Res.* 3 (2), 77-101.
- Bulman, M., 2017, July 19. Black students 50% more likely to drop out of university, new figures reveal. Independent. Available at: <https://www.independent.co.uk/news/uk/home-news/black-students-drop-out-university-figures-a7847731.html>.
- Callender, C., Jackson, J., 2004. *Fear of Debt and Higher Education Participation*. London South Bank University.
- Carnevale, A.P., Schmidt, P., Strohl, J., 2020. *The merit myth: how our colleges favor the rich and divide America*. The New Press.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2-14.
- Davies, R., Elias, P., 2003. *Dropping Out: A Study of Early Leavers from Higher Education*. Citeseer.
- DHET, 2010. Report of the Ministerial Committee on the Review of the National Student Financial Aid Scheme. Pretoria.
- DHET, 2021. FACT SHEET: The Highest Level of Education Attainment in South Africa. Pretoria.
- de Villiers, P., 2012. The poor and access to higher education in South Africa. Available at: https://www.researchgate.net/publication/267519547_The_poor_and_access_to_higher_education_in_South_Africa.
- Drake, S.J., 2022. Academic apartheid. Available at: <https://doi.org/10.1525/9780520381384>.
- Flint, M.A., 2021. Racialized retellings: (Un)ma(r)king space and place on college campuses. *Crit. Stud. Educ.* 1-16.
- Gallagher, C.A., 2003. Color-blind privilege: the social and political functions of erasing the color line in post race America. *Race, Gender Class.* Available at: <https://www.jstor.org/stable/41675099>.
- Gillborn, D., 2019. Hiding in plain sight: understanding and addressing whiteness and color-blind ideology in education. *Kappa Delta Pi Rec.* 55 (3), 112-117.
- Greenfield, N., 2021, November 6. The Crisis in Black University Enrolment and Graduation. University World News. Available at: <https://www.universityworldnews.com/post.php?story=20211105110142162>.
- Griffin, R.A., Ward, L., Phillips, A.R., 2014. Still flies in buttermilk: black male faculty, critical race theory, and composite counterstorytelling. *Int. J. Qual. Stud. Educ.: QSE* 27 (10), 1354-1375.
- Ishitani, T., 2020. Graduation Rates of Pell Grant Recipients at Public and Private Four-Year Institutions, vol. 95. College and University, pp. 10-21.
- Jackson, J.F.L., Moore, J.L., 2008. Introduction: the african American male crisis in education: a popular media infatuation or needed public policy response? *Am. Behav. Sci.* 51 (7), 847-853.
- Kelly-Laubscher, R., Paxton, M., Majombozi, Z., Mashele, S.S., 2018. Factors affecting the success of first generation university students at a South African university. In: *Understanding Experiences of First-Generation University Students: Culturally Responsive and Sustaining Methodologies*.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in U.S. Schools. *Educ. Res.* 35 (7), 3-12.
- Letseka, M., Breier, B., 2008. Higher Education Dropout and Poverty. In *Education and Poverty Reduction Strategies: Issues of Policy Coherence, Colloquium Proceedings*, pp. 83-101 unknown.
- Littler, J., 2017. *Against meritocracy: culture, power and myths of mobility*. Routledge.
- Love, B.J., 1993. Issues and problems in the retention of black students in predominantly white institutions of higher education. *Equity & excellence in education. Univ. Mass. Sch. Educ.* J. 26 (1), 27-36.
- Love, B.J., 2004. Brown plus 50 counter-storytelling: a critical race theory analysis of the "majoritarian achievement gap" story. *Equity & excellence in education. Univ. Mass. Sch. Educ.* J. 37 (3), 227-246.
- Lucido, J.A., Posselt, J.R., Polikoff, M., 2019. Why meritocracy is a myth in college admissions. *The Conversation*. Available at: <http://theconversation.com/why-meritocracy-is-a-myth-in-college-admissions-113620>.
- Mallman, M., 2017. The perceived inherent vice of working-class university students. *Socio. Rev.* 65 (2), 235-250.
- Mallman, M., Harvey, A., Szalkowicz, G., Moran, A., 2021. Campus convivialities: everyday cross-cultural interactions and symbolic boundaries of belonging in higher education. *Stud. High Educ.* 46 (7), 1449-1461.
- Masutha, M., Naidoo, R., 2021. Stories from the margins. *Knowledge Beyond Colour Lines: Towards Repurposing Knowledge Generation in South African Higher Education* 59.
- Mathebula, M., 2019. Recognising poor black youth from rural communities in South Africa as epistemic contributors. *Crit. Stud. Teach. Learn.* Available at: <https://www.ajol.info/index.php/cristal/article/view/190561>.
- McKinney, L., Novak, H., Hagedorn, L.S., Luna-Torres, M., 2019. Giving up on a course: an analysis of course dropping behaviors among community college students. *Res. High. Educ.* 60 (2), 184-202.
- Mills, C., Gale, T., 2011. Reasserting the place of context in explaining student (under)achievement. *Br. J. Sociol. Educ.* 32 (2), 239-256.
- Mkhwanazi, S., 2021, December 1. Blade Nzimande defends 75% eligibility criteria for NSFAS students. IOL | News That Connects South Africans. Available at: <https://www.iol.co.za/news/politics/blade-nzimande-defends-75-eligibility-criteria-for-nsfas-students-3d764dd4-ae93-4a7e-8b9c-864696aa8a4f>.
- Mobegi, F.O., Kara, A.M., 2022. Methods of financing university education and their implications on access, quality and completion rates: an experience shared from undergraduate education students, Maasai Mara University. Kenya. *European Journal of Education and Pedagogy* 3 (3), 280-287.
- Moodley, P., Singh, R.J., 2015. Addressing student dropout rates at South African universities. *Alternation*. Available at: <https://openscholar.dut.ac.za/handle/10321/1648>.
- Naidoo, R., 2004. Fields and institutional strategy: Bourdieu on the relationship between higher education, inequality and society. *Br. J. Sociol. Educ.* 25 (4), 457-471.
- Ncwane, N., 2020, May 25. Blade Nzimande comes out guns blazing in defence of N+2 rule. *The South African*. Available at: <https://www.thesouthafrican.com/studentlife-c/blade-nzimande-n2rule-nsfas/>.

- News, 2016, April 12. Higher education not reducing inequality – Habib. news24. Available at: <https://www.news24.com/News24/higher-education-not-reducing-inequality-habib-20160412>.
- Owusu-Kwarteng, L., 2020. "Studying in this England is wahala (trouble)": analysing the experiences of West African students in a UK higher education institution. *Stud. High Educ.* 1–12.
- Reay, D., Davies, J., David, M., Ball, S.J., 2001. Choices of degree or degrees of choice? Class, "race" and the higher education choice process. *Sociology* 35 (4), 855–874.
- Salako, P., 2022. "We are stuck": months of university strikes leave young Nigerians' lives on hold. *Guardian*. Available at: <https://amp.theguardian.com/global-development/2022/jun/08/months-of-university-strikes-leave-young-nigerian-lives-on-hold>.
- Sanchez, O., Kolodner, M., 2021, October. Why white students are 250% more likely to graduate than Black students at public universities. Available at: <https://hechingerreport.org/%e2%80%8b%e2%80%8bwhy-white-students-are-250-more-likely-to-graduate-than-black-students-at-public-universities/>.
- Solórzano, D.G., Yosso, T.J., 2002a. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1), 23–44.
- Solórzano, D.G., Yosso, T.J., 2002b. A critical race counterstory of race, racism, and affirmative action. *Equity & Excellence in Education. Uni. Mass. Sch. Educ. J.* 35 (2), 155–168.
- Vignoles, A.F., Powdthavee, N., 2009. The socioeconomic gap in university dropouts. *B E J. Econ. Anal. Pol.* 9 (1). Available at: <https://doi.org/10.2202/1935-1682.2051>.
- Woodson, C.G., 1969. *The Mis-education of the Negro*. ReadaClassic.com.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.

Tensions in access and accountability

Penny Jane Burke, Matthew Bunn, and Matt Lumb, The University of Newcastle, University Drive, Callaghan, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

“Widening” access	367
Fees and funding	368
Accounting for access	369
Access on the peripheries (of institutions and systems)	370
Governance and evaluation	370
Epistemic access and accountability	371
Conclusion	372
References	372

“Widening” access

Efforts to widen access to and build participation in higher education (HE) are largely concerned with redressing the historical and persistent underrepresentation of certain social groups in HE. The particular social groups recognized by policies have often been explicitly defined, although this is not consistently the case. For example, in Australia there are currently six “equity groups” identified by the Australian Federal government,¹ with a primary focus on students from low socio-economic status backgrounds; in China, policy has largely focused on rural and remote groups; in the United States a key policy emphasis has been people of color; in England it has been lower socio-economic groups, low-participation neighborhoods and people with disabilities (Ding, 2007; Gale and Tranter, 2011; Baptiste and Villa, 2012; Burke and Kuo, 2015); and in South Africa recent decades have seen significant change in policy and practice since the end of apartheid, including new forms of access for historically excluded groups, yet also much ongoing challenge and contestation (Nyamnjoh, 2012; Lockett, 2016). Yet across these countries, questions over who should be targeted varies and there are ongoing struggles over which groups become visible or invisible through different policy discourses.

Policy and practice relating to widening access are often associated with social justice but is commonly underpinned by meritocratic notions, for example the premise that initiatives should provide access to higher education to all who can benefit from it, regardless of social background. Research on HE has shown that it is a changing landscape reshaped by globalization, diversification and neoliberalism (Burke et al., 2017). Projects of access and accountability are increasingly shaped by these forces, with neoliberal commitments in particular creating the new conditions of HE differently in each social context. This ‘neoliberalisation’ of HE is an ongoing intensification of a strong valuing of individualization, choice, market security and minimal government (Slaughter and Leslie, 1997). It is also an intensification of embedding forms of power guiding people to orient themselves toward entrepreneurial and competitive dispositions. These orientations then support efforts to privatize and deregulate, rebuilding and blurring notions of ‘the state’, ‘the market’, ‘public’ and ‘private’ (Ball, 2012). The intensification of neoliberalism is important because, despite a perceived increase in institutional autonomy, these commitments to open markets, free trade, reduced state intervention and the value of increased competition work directly to shape approaches to access and accountability taken up in contemporary HE (Savage et al., 2013).

HE policies across many national contexts now prioritize the development of human capital and a highly skilled workforce to create competitive knowledge economies in the wake of deindustrialization in the global north. Access becomes linked to the imperative of national economies as its primary aim (Olsen and Peters, 2005). With the massification of HE, huge levels of growth in participation over many decades and a shift away from an elite system, access and the attendant accountabilities are reconstructed as about “the numbers”. As more students enter higher education, there is greater attention to financial considerations which has led to increasing student fees. As this has happened, HE has been seen largely in terms of individual investment, rather than broader notions of HE for the public good, and this has placed emphasis on individual choice, responsibility and reward. The idea of students as consumers has grown to shape the HE policy and practice landscape in many contexts, with students from disadvantaged social groups positioned as vulnerable or weak consumers in need of special support via commodified systems. This has tended to rationalize student fees in terms of the financial return to the individual graduate, where participation in HE is often portrayed as primarily a project of self-investment. This is used to justify moves toward charging student fees as an important aspect of providing a form of access, with claims often made of positive ramifications for social justice. Increasing student numbers, and therefore growing the market of higher education, is also seen as generally desirable by some policymakers as the funding of HE has shifted from public investment to provision through individual student fees (Carpentier, 2012).

¹The six equity groups are: People who identify as Aboriginal and/or Torres Strait Islander; People who are from low SES backgrounds; People with a disability; People from non-English speaking backgrounds; People from regional and remote areas; and Women in non-traditional areas. These groups are currently under review.

Fees and funding

Access agendas are entangled with wider HE reforms that play out differently in varied national contexts, with different implications for questions of accountability. However, the overarching forces driving change across different national contexts tend to be characterized by globalization and neoliberalism, with significant shifts toward market-based approaches.

Strong market-based systems blur the lines of accountability for access, locating it largely in the “choice” of the individual. It has been argued that the market system of HE operates most strongly in the United States (US). Financial aid is a major strand of policy approaches relating to access and participation in the US with extensive numbers of undergraduates requiring scholarships to afford the high cost of HE study (Mohrman, 2009; Zaloom, 2019). The preference for scholarships is driven by moves toward a market system and the race for HE institutions to be positioned at the higher ranks of university league tables. Merit-based financial aid has been identified as a strategy to support institutions to attract those students with strong academic records or those seen to have specific athletic or musical abilities. Students tend to opt for the institution offering the highest amount of financial aid even if that university or college is not their top choice. The process escalates as students seek larger scholarships and institutions spend more in the competition to select those applicants seen as most desirable or most deserving.

In the UK context, discourses relating to widening access were explicitly drawn on to create support for the reform of student fees from 2012 to 13 onwards. Those HE institutions wishing to charge home/EU students more than £6000 per annum became locked into a system for the control, accountability, and regulation of access-widening activity. A tie-in between permission to charge fees above the £6000 baseline and efforts to widen access and boost participation constituted a financial control mechanism for national government to steer HE at a distance, with the capacity to determine, through regulation, how access was enacted. In the UK, access-related activity is tied to government performance indicators, creating a form of accountability at the institutional level. Responsibility for strategies relating to access and participation funding in England is with the Department of Education through what was the Higher Education Funding Council for England (HEFCE) and the Office for Fair Access (OFFA). In 2018, HEFCE and OFFA became the Office for Students (OfS), with the remit to “put students at the heart of higher education”. The Department of Education as the ministry responsible for HE, sets high level policy and decides on access and participation funding allocations, which are then distributed by the OfS. In addition, HE institutions are required to use a portion of fees charged over the baseline to support student access, success and progression by developing dedicated infrastructure and activities. This needed to be evidenced within the institution’s “Access Agreement”—detailed institutional plans approved by the non-departmental body responsible for safeguarding “fair” access to higher education in England. As a regulatory body, OfS has powers to fine a university or suspend part of its grant if it deems an Access Agreement unacceptable. A refusal to approve an Access Agreement would prevent the institution from charging above the baseline fee. Overall, institutional expenditure on this area has been shown to vary greatly, as does the design and implementation of access-related activity. A new but similar regulatory system in the UK has recently come into place with “Access and Participation Plans”, which are monitored by the OfS to ensure that HE providers meet the commitments made to students in their plans. The plans must set out how the HE provider will improve equality of opportunities for underrepresented groups, including a vision for change, and the investment that will be made to reach this desired change with specific targets identified.

There is an ongoing policy tension between HE institutional autonomy and demands from policy and funding bodies to develop forms of accountability around access agendas. For example, there is increased pressure to develop “better” monitoring, evaluation and regulation of activities designed to expressly widen access to HE in the UK context. For instance, measures were taken in 2014 to ensure that HE institutions moved away from using additional fee income to provide financial support directly to students in the form of bursaries and/or scholarships, and toward investment in infrastructure and activity designed to support student access, success, and progression through an emphasis on the entire student lifecycle. This was justified as the deferral and income-contingent character of student loan repayment with the aim that no-one is excluded by virtue of financial circumstances from participation in HE. Rather, efforts would be directed at providing quality advice, guidance and information to potential students from underrepresented groups and then supporting them practically to progress and succeed through and from HE. Since 2014, there has also been greater attention to evaluating and demonstrating evidence of outcomes. Initiatives such as the HEAT (Higher Education Access Tracker) and the NERUPI partnership (Network for Evaluating and Researching University Participation Interventions) are products of such attention. At times, tense debates circulate regarding which approaches and methods are best suited to understanding these educational dynamics, leading to a challenging environment for notions of accountability around this policy and funding to hold together in any coherent way.

In Australia, Denise Bradley delivered a major federal review of HE in 2008 with an explicit policy focus on setting the target of 20% of the student body to be composed from low socioeconomic status backgrounds and 40% of 25–34 year olds to hold a bachelor-level qualification by 2020. Funding was provided to support the deregulation of the HE system with the aim to create more undergraduate places. To complement this new “Demand Driven System”, the Higher Education Participation and Partnership Program (HEPPP) was introduced in 2010 as an annual fund given to universities to increase the participation of students from low socioeconomic status backgrounds, and to develop partnerships with local schools, vocational and educational training (VET) providers and other universities to develop initiatives that supported the intention of the policy. A National Priority Pool (NPP) was also part of this policy suite, offering grants to support the development and implementation of initiatives and to develop an evidence base to guide the emerging field. Research in the Australian context has uncovered

how the responsibility and benefits of access and participation related work has shifted toward a market-based system as in the US and UK. Australian student fees for home/domestic students vary across HE institutions but are largely subsidized through the Higher Education Loan Program (HECS-HELP) for what are known as “Commonwealth Supported Places” or CSPs. To be eligible for HECS-HELP a person must be studying in a CSP as an Australian citizen who meets the residency requirements or a New Zealand Special Category Visa (SCV) holder or permanent humanitarian visa holder and meet the residency requirements.

Recent years in Australia have seen increasing uncertainty in relation to access and participation agendas as a conservative Liberal/National Party coalition government introduces new national HE reforms, making the field a precarious one for those tasked with managing this area of university responsibility. In late 2020, the “Job-Ready Graduates” package passed federal parliament. Australian HE was already amid a crisis brought on by the global coronavirus pandemic. With the sharp decline in international student enrollment in the Australian sector, a major income stream for universities, Australian HEIs were already facing dramatic restructuring to make up what amounted to a significant financial shortfall. The job ready graduate reforms were an extensive re-shaping of the way particular degree areas would be funded, attempting to direct student participation toward areas of nation-state productivity. An important aspect of these reforms is a move toward what has been termed the IRLSAF (Indigenous, Regional and Low SES Attainment Fund) which will fund universities to support Indigenous students and students from low SES and regional backgrounds. The establishment of this fund, which will incorporate HEPPP, is part of a growing focus on regional, rural and remote students as part of the access and participation agenda in Australian HE. The reforms have experienced extensive critique from many analysts of HE. One area of concern is to what extent Australian “Enabling Programs” (non-traditional, open access pathway programs) will be funded over the long term, despite this providing a major accessible entry route for students from historically under-represented communities.

Accounting for access

Despite the dynamic and contested character of access-related policymaking outlined above, the funding explicitly identified for access and widening participation requires clear lines of accountability to ensure appropriate use of these precious resources. Access and participation activity occurs however at the nexus of a complicated mix of social processes, governmental responsibilities, and value-systems that are guided, often implicitly, by the framing of the purpose of HE participation. This has led to contestation over the different perspectives regarding the purpose of HE participation, framed differently as they are by commitments to social justice, social diversity, social mobility, meritocracy and neoliberal perspectives, among others (Whitty et al., 2015). The framing of the reason to invest in access and participation comprises markedly different activities related to the macro and micro contexts in which the activities are designed. Examples include building partnerships between local communities (often with a focus on schools) and the university, with activities designed to “raise aspiration”, outreach programs that focus on advice, guidance and information about how to access HE, various scholarship schemes and student support activities such as counseling, study skills and peer mentoring (Bennett et al., 2015). Over time, and with some critique of the focus of these activities on “the problem” of access and participation as being outside of the HE context, there have been increasing efforts to create inclusive curricula and teaching practices with a focus on supporting student retention, completion and success. An increasing concern has been to invest funding in larger scale and more diverse programs of evaluation to ensure that the use of funding and resources are appropriate, effective, with efforts to understand “what works”. Yet this particular aspect of what might be considered accountability has largely focused on dominant perspectives of the purpose of HE participation, with evaluation often ignoring marginal value-systems altogether. Instead, evaluation is often designed to consider only that which is easy to measure (Harrison and Waller, 2017). This primary accountability focus of evaluating the “measurable” has been critiqued as decontextualizing the programs or community members involved, and/or ignoring the complexities involved in understanding why something “works”, for whom, over what duration, and why (Burke and Lumb, 2018).

If accountability in this context is to be taken seriously, it is important that increased attention is given to the methods by which resources and opportunities can be differently targeted to those who have suffered long-standing forms of social, educational and economic inequality. This has been described in research as a key opportunity of access and participation efforts through processes of redistribution, pushing beyond dominant forms of accountability toward an ethical responsibility of HE to redress historical exclusions and injustices. Research in this area has pointed to a tension requiring careful consideration when developing approaches to targeting resources toward groups or individuals. This tension lies in the requirement to identify the “appropriate” target groups (i.e., those who have experienced social disadvantage and inequality) while simultaneously avoiding constructing classifications that reproduce derogatory depictions of groups through deficit discourses (including flawed assumptions that the problems lie in a form of lack within those groups being targeted by access and participation activities). Some have referred to this as a process of misrecognition that distorts the framing of the challenge of access and participation through a lens of deficit (Burke, 2012). Misrecognition contributes to processes that objectify, dehumanize and/or homogenize people and groups through instruments of classification and categorization. Finally, if we are to grapple with the notion of accountability in relation to access and participation, we must consider questions relating to social power, influence and decision-making. Representation of the issues at stake by those in

positions of authoritative power lead to a distorted focus on individual attitudes and flawed assumptions of what those targeted by policies lack, rather than on redressing historical, systemic, structural, economic and cultural inequalities that reproduce educational disadvantages despite efforts to widen access and participation. Thus, attention must be paid to the interconnections between redistribution, recognition and representation when considering access and accountability.

Access on the peripheries (of institutions and systems)

Research has shown that work within institutions to build new forms of student access is often peripheral, and that the access constructed is on the peripheries of HE systems themselves, providing limited participation to more elite degrees and/or organizations (Burke, 2012). Concentrated work attempting to build access for historically underrepresented groups is not commonly seen as “core business” and tends to be located at the peripheries of institutions rather than the project of access being embedded in institutional structures and practices. In countries where there is dedicated national funding, specialized units, centers or departments emerge at HE institutions with professional staff responsible for developing new strategies, infrastructure and activities that aim to facilitate the intended access and participation outcomes. In this way, access and participation can be relatively detached from the main focus of universities (Jones and Thomas, 2005). This dramatically reduces the potential for institutional, structural, and cultural change. Furthermore, the new programs and practices designed are often different in each HE institution with vastly variable staffing profiles, initiatives, and general resources. Clearly it is important for initiatives seeking to engage and redress historical underrepresentation to address the specific local conditions in which the work is situated, yet the structures in place have implications for the representation of access and participation issues across the institution, the status of initiatives and the persons and groups held accountable. This has implications too for the quality of material and opportunity made available to students and communities. Units located more centrally within institutional structures and hierarchies carry status and authority in terms of strategy and investment in agendas for creating greater equity of access and participation, compared to units with little influence on senior level decision-making. Furthermore, the peripheral positioning of access units and programs, and the notion that the responsibility for access and participation efforts lies with units outside of academic faculties, can mean that practices of teaching and research are disconnected from these efforts. This has implications for developing new knowledge and broader expertise across the domains of higher education (e.g. research, teaching, curriculum development, assessment, admissions, student support), limiting opportunities to build embedded institutional approaches whereby agendas relating to equity, access and participation are held as a responsibility of all staff members. This kind of ‘embedded’ approach requires however close attention to institutional structures which are historically and socio-politically formed, and it is an approach which must include capacity to support high quality professional development across the institution for understanding of the issues relating to organizational and community contexts. This approach also requires carefully developed governance structures and evaluation methodologies.

Governance and evaluation

Access and participation agendas have produced significant government investment and national campaigns, yet there remains a lack of research investigating governance of these agendas and related questions of responsibility and accountability. Scholars of HE have argued that the limitations of current governance and management necessitates developing more effective structures and practices, particularly in when it comes to establishing more equitable systems of tertiary education. Countries such as Australia and England for example have established policy frameworks yet stark disparities of access and participation endure within and across nation state systems. In many contexts there are calls for rigorous evaluation approaches to understand the impact of different access and participation structures and activities, and whether current governance arrangements are appropriate. For example, scholars in the South African context have advocated for more sophisticated frameworks of evaluation in higher education that are capable of engaging deeply embedded historical inequality structures (Bozalek and Leibowitz, 2012). Many institutions consistently fail to meet benchmarks despite regulatory mechanisms and qualitative research highlights that the structures, activities and support that seek to redress inaccessibility are highly variable in terms of quality across institutions, regions and localities.

Due to the paucity of research regarding how and with what effect access and participation efforts are governed, structured and supported, precisely how allocation procedures operate and what forms of resource accountability there are at the institutional level is largely unknown across the sector. It could be argued that there is insufficient oversight into questions of to whom or to what institutional policymakers are accountable when deploying these public resources, although this varies across different nation-state contexts. There is a scarcity of analysis in terms of, for example, how the stratification of HE institutions maps onto policy and practice and how that differently (and potentially unequally) impacts on local and/or regional underrepresented communities across the wider sector. Relevant data is collected through different bodies but rarely combined, analyzed or presented in a coherent manner. This points to a need to develop ethically formed and contextualized research and evaluation methodologies that can generate richer understanding of access and participation patterns to inform structures and practices across local, regional, and national scales.

Surrounding access and participation agendas there is a problematic “politics of knowledge” that tends to endlessly foreground the historical prominence of dominant social groups. This dynamic plays out in a myriad of ways, some of which is explored in the next section of this entry relating to “Epistemic Access and Accountability”. One way this problem plays out though is in a medical, clinical, and experimental mode of knowing that has come to colonize evaluative research in relation to educational effectiveness. Randomized control trials (RCTs) have in many contexts been turned to as a potential evaluation approach in the context of understanding “what works” when it comes to access and participation investments and initiatives. This has however been extensively critiqued as not even remotely fit for purpose (Gale, 2017). RCT research design pays scant attention to context and is ill-suited to understanding how social inequalities operate in relation to education due to a distorting emphasis on cause and effect. RCTs are unable to shed light on how students make choices and decisions, develop their sense of aspiration and capability, experience themselves as learners, feel inspired, excluded or valued. Students do not choose to study simply through “rational choices” but are guided by an array of different circumstances and experiences within their lives that escape simple metrics and measurements (Brynin, 2012). Lived experience of inequalities and exclusions cannot be measured, nor can the pleasure of learning or the sense of belonging. These require methodologies that can generate insight beyond that which is readily observable and/or reducible. Evaluative research must therefore attend to relational, dynamic, contextual and embodied dimensions of education and engage the complex formations that develop the potential for parity of participation in HE. Without methodologies that enable a form of critical analysis, the reduction of evaluation surrounding access and participation agendas to methods of measurement is likely to reproduce representation of inequalities through instruments that effectively divide, measure and classify. If the notion of access is to hold the possibility of transforming HE systems and institutions toward more inclusive and just patterns and arrangements, evaluative research needs to be led by the perspectives of those whose judgments and values have historically been disregarded or under-represented. This will ensure that the assessment of what is valuable about access and participation is not simply an ongoing reflection of those in relatively privileged social positions.

Epistemic access and accountability

The concept of access is primarily defined to mean “formal access”, i.e., being granted means to formally enter a higher education study pathway leading toward a formally recognized qualification. These qualifications follow national guidelines, and thus curriculum are composed in accordance with national regulations, policy and expert bodies. A student granted access will meet a series of academic basic requirements that will have been attained in previous education. However, HE scholars have continued to draw attention to forms of access beyond this rigid definition. The formalized means of assessing potential HE candidates upholds a restrictive understanding of personal ability and aptitude. This produces extremely narrow conditions for entry that can be leveraged through privileged access to social networks and services to effectively “game” HE access.

In contrast, many potential students have not had the same level of access to resources, including money, tutors, networks, support and time for building the basis to respond to traditional access criteria. Even though globally there have been shifts in the make-up of student populations on the basis of previously ‘under-represented’ groups (such as along ethnic, gender and class lines), there has been little shift in what constitutes legitimate knowledge, skills and attributes to be granted access. The denial of formal access for disadvantaged groups is thus on the basis of lacking the narrow forms of “cultural capital” deemed appropriate for competence within the forms of knowledge that are seen as legitimate within higher education. Even countries with a sustained agenda for widening formal access have largely ignored questions of the criteria used to determine who has the appropriate academic capability. Moreover, as mentioned above little mapping has been completed to understand the relationship between access criteria and assessment and the distribution of access to elite institutions.

A further and connected issue of access is what is often referred to as epistemological or epistemic access. Starting with Morrow’s (2009) concept of “epistemological access”, HE scholarship has extended its focus from formal access to consider what forms of knowledge and knowledge production are being accessed. Higher education disciplines contain what can be called “systematically organized knowledge”. Epistemological access is seen to be divided by contextual knowledge and theoretical knowledge. Contextual knowledge is understanding the facts of knowledge as they exist and pertain to specific real-world situations. Theoretical knowledge refers to understanding the way that information and data has been collected and collated, and the often-unwritten normative rules that allow for recognizing historical processes of producing knowledge and for assessing new claims to knowledge (Muller, 2014: 262). This then provides access to the production of knowledge, and to assessing claims to knowledge specific to a discipline.

Epistemic access has been signaled as a concern in the movement of higher education toward outcome and employment as the basis of curriculum design. Many authors have argued that policy emphasizing outcomes at the expense of theoretical knowledge is changing what is being accessed within higher education (Shay, 2013). Outcome focused HE (such as in policy like the Australian Job-Ready Graduates Package, DESE, 2020) risks producing new divides in HE between those who are trained only to put knowledge into practice, and those who can make claims to knowledge. This is an important way in which inequalities are perpetuated, as epistemic access has historically run across gender, ethnic and class lines (Wheelahan, 2012).

This leads into a broader question of epistemic monism and pluralism. Higher education contains a series of disciplines that are tightly defined ways to produce knowledge. These typically maintain a monopoly over the legitimate knowledge pertaining to an area of study. These disciplines have a colonial history that stems from Europe and now extends globally in the recognition of its claim to legitimacy. Because of this, many knowledges and ways of knowing have not been sanctioned as legitimate higher

education knowledges. This raises the question of the legacy of colonialism on the consecration and codification of Western knowledges while denying knowledges, or even producing “epistemicide” (Santos, 2014) of knowledges from Indigenous groups and the global South.

Calls for change in who is given epistemic validity, and how this is sanctioned and valued, have become increasingly prominent. For example, in South Africa the Rhodes Must Fall Movement has underlined the need to adopt different curricula, pedagogy, symbols and institutional structures that incorporate a broader array of different epistemologies. These calls see claims to diversity and inclusivity through practices like representation as achieving very little in providing access for other ways of knowing. Providing a decolonial approach to epistemology and curriculum requires the dismantling of the universal/monopolizing claims to knowledge that higher education disciplines maintain, opening the space for an epistemological plurality in how forms of knowledges are legitimated (see Mignolo, 2007, 2011; Escobar, 2010).

Conclusion

Widening access and participation has been a key theme of higher education policy over many decades, with an explicit policy focus in some parts of the world. Although the concept of access is contested, the policy and practice of widening access and participation is increasingly driven by the forces of globalization, neoliberalism and marketization. Furthermore, policy interventions have involved both fiscal and human investments and have raised complex challenges for issues of accountability. Who determines how funding is allocated? What methods are used for the distribution of resources and who decides? Who is accountable for the methods by which funding is allocated? Who is held accountable for the effectiveness of the use of funding? Who gains epistemic access and whose knowledges count? These questions are challenging for higher education institutions and governments and have been considered in different ways, including through evaluation and measurement instruments. Increasingly, where there is specialist funding attached to widening access and participation, higher education institutions are being required to evaluate the effectiveness of their activities as a regulatory measure of accountability.

References

- Baptiste, H.P., Villa, E.Q., 2012. Affirmative action. In: Banks, J.A. (Ed.), *Encyclopedia of Diversity in Education*. SAGE Publications, Thousand Oaks, CA, pp. 35–40.
- Ball, S.J., 2012. *Politics and Policy Making in Education*. Taylor & Francis, Hoboken, NJ.
- Bennett, A., Naylor, R., Mellor, K., et al., 2015. *Equity Initiatives in Australian Higher Education: A Review of Evidence of Impact*. Centre of Excellence for Equity in Higher Education, University of Newcastle, Newcastle. Available at: https://www.newcastle.edu.au/__data/assets/pdf_file/0016/261124/REPORT-FINAL.pdf.
- Bozalek, V., Leibowitz, B., 2012. An evaluative framework for a socially just institution. In: Leibowitz, B. (Ed.), *Higher Education for the Public Good: Views From the South*. African Sun Media, Stellenbosch, pp. 59–72. <https://doi.org/10.18820/9781928357056/05>.
- Brynin, M., 2012. Individual choice and risk: the case of higher education. *Sociology* 47 (2), 284–300.
- Burke, P.J., Kuo, Y.C., 2015. Widening participation in higher education: regimes and globalizing discourses policy. In: *The Palgrave International Handbook of Higher Education Policy and Governance*. Palgrave Macmillan, Hampshire and New York, pp. 547–568.
- Burke, P.J., Crozier, G., Misiaszek, L., 2017. *Changing Pedagogical Spaces in Higher Education*. Routledge, London & New York.
- Burke, P.J., 2012. *The Right to Higher Education: Beyond Widening Participation*. Routledge, London and New York.
- Burke, P.J., Lumb, M., 2018. Researching and evaluating equity and widening participation: praxis-based frameworks. In: Burke, P.J., Hayton, A., Stevenson, J. (Eds.), *Evaluating Equity and Widening Participation in Higher Education*. Trentham Books, London.
- Carpentier, V., 2012. Public-private substitution in higher education: has cost-sharing gone too far? *High Educ. Q.* 66 (4), 363–390.
- Ding, X., 2007. Expansion and equality of access to higher education in China. *Front. Educ. China* 2 (2), 151–162.
- Escobar, A., 2010. “Worlds and knowledges otherwise”: the Latin American modernity/coloniality research program. In: Mignolo, W., Escobar, A. (Eds.), *Globalization and the Decolonial Option*. Routledge, Abingdon, pp. 179–210.
- Gale, T., Tranter, D., 2011. Social justice in Australian higher education policy: an historical and conceptual account of student participation. *Crit. Stud. Educ.* 52 (1), 29–46.
- Gale, T., 2017. What’s not to like about RCTs in education? In: Childs, A., Menter, I. (Eds.), *Mobilising Teacher Researchers: Challenging Educational Inequality*. Routledge, London.
- Harrison, N., Waller, R., 2017. Success and impact in widening participation policy: what works and how do we know? *High Educ. Pol.* 30 (2), 141–160. <http://www.doi.org/10.1057/s41307-016-0020-x>.
- Jones, R., Thomas, L., 2005. The 2003 UK government higher education white paper: a critical assessment of its implications for the access and widening participation agenda. *J. Educ. Pol.* 20 (5), 615–630.
- Luckett, K., 2016. Curriculum contestation in a post-colonial context: a view from the South. *Teach. High. Educ.* 21 (4), 415–428. <https://doi.org/10.1080/13562517.2016.1155547>.
- Mignolo, W., 2007. Delinking: the rhetoric of modernity, the logic of coloniality and the grammar of de-coloniality. *Cult. Stud.* 21 (2–3), 449–514. <https://doi.org/10.1080/09502380601162647>.
- Mignolo, W., 2011. Epistemic disobedience and the decolonial option: a manifesto. *Transmodernity* 1, 3–23.
- Morrow, W.E., 2009. *Bounds of Democracy: Epistemological Access in Higher Education*. Hsr Press.
- Mohrman, K., 2009. What world-class universities should not adopt from the American higher education model. In: Sadlak, J., Cai, L.N. (Eds.), *The World-Class University as Part of a New Higher Education Paradigm: From Institutional Qualities to Systemic Excellence*. UNESCO-CEPES, Bucharest, pp. 97–118.
- Muller, J., 2014. Every picture tells a story: epistemological access and knowledge. *Educ. Change* 18 (2), 255–269.
- Nyamnjoh, F.B., 2012. “Potted plants in greenhouses”: a critical reflection on the resilience of colonial education in Africa. *J. Asian Afr. Stud.* 47 (2), 129–154. <https://doi.org/10.1177/0021909611417240>.
- Olssen, M., Peters, M., 2005. Neoliberalism, higher education and the knowledge economy: from the free market to knowledge capitalism. *J. Educ. Pol.* 20 (3), 313–345.
- Santos, B., 2014. *Epistemologies of the South: Justice Against Epistemicide*. Routledge, London & New York.
- Savage, G., Sellar, S., Gorur, R., 2013. Equity and marketisation: emerging policies and practices in Australian education. *Discourse* 34 (2), 161–169.

- Shay, S., 2013. Conceptualizing curriculum differentiation in higher education: a sociology of knowledge point of view. *Br. J. Sociol. Educ.* 34 (4), 563–582. <https://doi.org/10.1080/01425692.2012.722285>.
- Slaughter, S., Leslie, L.L., 1997. *Academic Capitalism: Politics, Policies and the Entrepreneurial University*. John Hopkins University Press, Baltimore.
- Wheelahan, L., 2012. *Whose Knowledge Matters in the Curriculum: A Social Realist Argument*. Routledge, London & New York.
- Whitty, G., Hayton, A., Tang, S., 2015. Who you know, what you know and knowing the ropes: a review of evidence about access to higher education institutions in England. *Rev. Educ.* 3 (1), 27–67.
- Zaloom, C., 2019. *Indebted: How Families Make College Work at Any Cost*. Princeton University Press, New Jersey.

The problem with higher education student completions in the global north

Trevor Gale and Stephen Parker, University of Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	374
The completion problem for particular students and institutions	375
The completion problem for particular higher education systems	376
Reimagining the problem	378
Acknowledgments	379
References	379

Introduction

Concern about student completion rates is relatively recent in the history of higher education (HE) and tends to be associated with an expansion in undergraduate student participation. As the proportion of a nation's population studying for a bachelor degree increases, so too does concern about suboptimal rates of completion, or what some pejoratively refer to as "drop out." It has become a particular concern for OECD nations (OECD, 2019a). The context is important for understanding why.

Since the mid twentieth century, a defining feature of higher education systems of the global north¹ has been the steady increase in their student numbers, particularly undergraduates. Trow (1974, 2006) characterizes the phenomenon as a shift in student participation from elite to mass to universal levels. For example, in Germany there were 836,002 students studying for a university undergraduate degree in 1975, increasing to 1.74 m in 1990 and to 2.9 m in 2020 (Óhidy, 2018; Statista, 2021). Illustrated in these figures is that the rate of HE's expansion across most of the global north gathered pace early in the 21st century. For example, in Australia the number of undergraduate students in the 10 years from 2010 to 2019 increased by about 30% or an additional 183,410 students (DESE, 2020).

Initially—in the mid to late twentieth century—much of the increase was prompted by an "unmet demand" for HE by "baby boomers" who were predominantly middle class (or aspiring) and while they qualified for university entry, there were insufficient places and/or the funds at their disposal were inadequate to secure their participation (Gale and Tranter, 2011). Over time, governments in the global north have sought to rebalance HE demand and supply through the introduction of a range of measures, including: financial support to enable participation by students with minimal financial means; increased funding for universities to grow the number of student places; building new universities and repurposing other tertiary institutions as universities, thus creating more university places; and extending HE provision to include bachelor and associate degrees at vocational tertiary institutions, although uptake of these has been relatively slow (Knight et al., 2021).

Justification for implementing these measures has often been couched in terms of social justice and "fairness" (Rawls, 1999). At first, at least in some jurisdictions, these were coupled with a Keynesian economic logic; that is, meeting current social demands, financed on the expectation of future economic growth. Later, the national cost of expanding HE provision was offset against expectations of "social mobility" and other individual benefits associated with obtaining a bachelor degree, for which students were obliged to financially contribute, often through schemes that deferred payment until graduates secured paid employment (Gale and Parker, 2018). The latter was framed by a neoliberal logic, which also informed the ambition of nations in the global north to become leaders in a competitive global knowledge economy (requiring an investment in human capital), to rival the global industrial economy in which their dominance had begun to dissipate. Thus, while social justice often featured in the rationale for HE's expansion, it was nonetheless secondary to economic and political considerations.

This was evident even in programs ostensibly aimed at increasing participation by under-represented groups. Widening Participation (WP)—as the strategy became known in the UK and which is now the term that tends to be used more globally—became more prevalent mid to late in HE's period of expansion (1960s–2020s), resulting in increased numbers of students from low socioeconomic-status (SES) backgrounds and from other marginalized groups taking up university study. In some jurisdictions, the proportion of these students studying HE also increased. For example, in England in 2014/15 there were 159,015 students enrolled in HE from the most deprived backgrounds, representing 17% of all undergraduate enrollments, which increased to 191,305 or 19% of enrollments by 2019/21 (HESA, 2020, 2021). Similarly, in Australia, from 2010 to 2019 the number of university students from low SES backgrounds increased by 33% (33,165 students), compared with the 30% increase in all university students (DESE, 2020).

¹In a similar manner to how Connell (2007) uses "global north" to signal the dominance or centrality of particular (academic) cultures in the realm of knowledge, we use the term to refer to nations that are more centrally located in global fields of economic and political power, rather than to their respective geographical locations. For example, while Australia is geographically located in the global south, it is more closely aligned with the economic and political interests and influence of the global north.

However, the driver for this growth in the proportion of HE students from under-represented groups was predominantly economic. This was most evident in Australian HE following the Global Financial Crisis (GFC) of 2007/2008, when the government introduced what became known as a “demand-driven system” (Australian Government, 2009). Australia was one of the few global-north nations not to experience a GFC-induced recession, placing it in a strong position to invest in HE provision (among other things). Such investment ushered in the nation’s “fourth wave” (Gale, 2021) of WP differed from previous government interventions in that instead of responding to unmet demand, it “sought to create demand [particularly among under-represented groups] to meet ambitious (and eventually unachieved) HE participation targets” (26). Pre-GFC, other global-north countries held similar aims (e.g. in England, a target of 50% was set for 18–30 year old HE participation by 2010)—with WP seen as a major strategy to achieve this—with the view that HE’s “contribution to the economic and social well-being of the nation is of vital importance” (DfES, 2003: 10).

It is this context that frames student completion rates in higher education in the global north. As more students have entered higher education, including more from disadvantaged backgrounds, completion rates—or, more specifically, rates of non-completion—have increasingly been described as a problem. In one illustration of this, in 2013, Jan Truszczyński, Director-General of the European Commission’s Directorate-General for Education and Culture, observed that:

Currently too many students in the EU drop out before they complete their higher education degree. Students from a lower socio-economic background and other disadvantaged groups are the most likely to drop out. This is a problem across the EU.

Truszczyński (2013: 7).

Responding to such assessments, the following problematizes simple accounts of “dropping out” as problem. This is not to suggest that non-completion is nothing to worry about, but rather we identify problems with: the characterization of non-completion, in ways that tend to benefit some institutions and social groups more than others; the location of the problem within student groups with higher rates of non-completion, rather than in institutions and systems; the assumption that 100% completion is the only positive outcome so that all non-completion is regarded as negative for students, institutions and systems; and the limits on comparisons of student completion rates across institutions and systems, given their different definitions, parameters, options and societal expectations of HE student completion. As well as identifying these problems with the problem (of non-completion), we conclude that non-completion might not be the problem imagined if less emphasis is placed on what institutions and systems lose by way of non-completion and more emphasis is placed on the benefits of HE from the perspective of students.

We begin, though, with how “drop-out” has been conceived as a problem and, because of its construction, the solutions that it suggests.

The completion problem for particular students and institutions

With HE student numbers rising, particularly as a proportion of the wider population and particularly including students from disadvantaged backgrounds, two concerns have been raised by conservatives. Reasons for their concern is a perceived threat to HE’s quality (i.e. an assumed lowering of academic standards in order to facilitate the entry and progression of the less able) and/or the financial and emotional distress experienced by students (especially from groups most likely to drop out), who fail to meet HE’s high standards and subsequently fail to complete, which could have been avoided if they had not been granted access in the first place.

These concerns tend to be most prominent among high-status higher education institutions (HEIs) and conservative media. The general tenor of the complaints is that it is unfair to enroll students who performed poorly in school, i.e. typically students from disadvantaged backgrounds. To do so is “setting them up for failure” (Cervini, 2012, in Gale and Parker, 2017), causing unjustified heartbreak as well as creating additional expense for institutions (Gale and Parker, 2017). The argument is that it would be unrealistic to expect these students to be sufficiently prepared or capable for the rigors of university study: “He or she didn’t do well at school, so why is higher education going to be any different?” (Hare, 2015).

These are claims based on an apparent correlation or association between students from low SES backgrounds and their comparatively high rates of non-completion or attrition. Yet closer examination of the data often reveals that in HEIs where the retention of students from low SES backgrounds is lower, retention for all students is also lower, suggesting that the assertion *low SES = high attrition* is more complex than it is often made out to be (Gale and Parker, 2017). With the evidence not matching the claims, resistance to higher levels of low SES enrollments might be interpreted as prestigious institutions seeking to maintain their status and distinction (cf. Gale, 2021), to differentiate themselves from other HEIs with higher rates of low-SES participation. This was exemplified in 2009 when the then Executive Director of Australia’s “Group-of-Eight” universities, Michael Gallagher, branded recommendations to expand participation in HE as “a drift to mediocrity,” and reasserted the need to maintain a stratified system in which elite universities can focus on “achievement of research excellence” (Gallagher, 2009, in Gale, 2011: 13).

A second set of explanations for the non-completion problem has led to a different set of solutions. Faced with what HEIs perceive as an ever-growing cohort of under-prepared students, equity practitioners have devised a range of support measures to assist students with their “transition” (cf. Gale and Parker, 2014) into and beyond the 1st year of HE study. A key premise of

this approach is that “access [to HE] without support is not opportunity” (Tinto, 2008). That is, it is not sufficient to merely expand access to HE to disadvantaged and other marginalized groups; there needs to be a high level of coordinated support to ensure that students can appropriately adjust to university life and be inducted into institutional ways of studying and being (Gale and Parker, 2014). Indeed, institutional support to aid retention and completion is increasingly seen by the European Commission (2015) as essential.

There are several broad groups identified as needing support in the 1st year and beyond:

- Those from disadvantaged groups (such as low SES students and minority ethnic groups) that are traditionally the focus of widening participation initiatives; with students from low SES groups regarded as the most likely to discontinue their studies (Quinn, 2013).
- Students from diverse groups who are not necessarily associated with WP. Indeed, “drop-out” is not always neatly associated with WP policies in European nations, instead often attributed to a “lack of attention to diverse student needs” (Quinn, 2013: 58; European Commission, 2015), reflecting a “more general experience of difficulties in making sense of HE” (Ulriksen et al., 2017: 56).
- First year students in general (e.g. Kift et al., 2010). Some consider that simply being in the 1st year of HE study equates to membership of an “equity group.”

What these have in common is the perception that HE students associated with population groups most likely to drop out, need extra support not provided by traditional HE pedagogies and curricula. Throughout the UK, Europe, the USA, and Australia, academics and equity practitioners champion and implement a range of support structures aimed at enhancing student engagement, to help integrate students into, what is for them, an alien culture of higher education and to ease their transition (Zepke, 2021; Tinto, 1993; Gale and Parker, 2014). The nature of the support and engagement varies from context to context, but includes efforts to increase the cultural capital of students from non-traditional backgrounds (e.g. Thomas, 2002; Zepke, 2021), providing institutional information, and offering additional tuition to enhance academic skills (Thomas, 2021).

These approaches to support and engagement are largely “bolt on” strategies, in that they are in addition and separate from mainstream HE curriculum, pedagogy and assessment. Equity practitioners and sympathetic academics have also developed what they term a “whole of institution” approach to improving retention and completion, particularly among those from non-traditional backgrounds. In these approaches, the institution is required to provide embedded, consistent and targeted support that engages with the needs of a diverse cohort.

This whole-of-institution approach includes changes to admission processes, modified modes of course delivery, flexible organizational structures, and re-timetabling of courses (Thomas, 2021). In terms of strategies aimed directly at students, the so-called “transformative” approach includes more student-centered pedagogies granting better access to the curriculum, direct learning support, co-curricula activities, making explicit the hidden curriculum of HE (e.g. Thomas, 2021).

However, despite being described as “transformative” and “whole-of-institution” (Thomas, 2021; Kift et al., 2010), these approaches tend to leave HE’s existing ways of learning, dominant forms of knowledge and epistemic hierarchies unquestioned.² In short, they aim to “mediate diversity”; i.e. compensate for or overcome students’ differences from the norm. There is little to no recognition of students’ backgrounds and living conditions as sources of valuable knowledge rather than mere impediments’ (Rios-Aguilar and Kiyama, 2018: 5; see also Moll et al., 1992) to learning and success in HE.

The completion problem for particular higher education systems

As well as locating the completion problem in particular student groups and HEIs, supranational organizations (such as the EU and the OECD) also attribute completion problems to particular HE systems. But there are problems with such comparisons, even with the creation of measures for comparison such as a degree’s “theoretical duration” (TD). Table 1, derived from recent OECD (2019a,b) reports, is a good example of the differences and the problems such data present for meaningful comparisons.

Reporting 2017 data in columns two and three, the Table purports to provide a comparison of HE degree completion rates in OECD countries.³ The problem with the data becomes apparent the more one delves into their points of comparison. For example, included in the Table are all graduations from cohorts that commenced a bachelor degree, even if the qualification obtained was not the same as the bachelor degree they commenced.⁴ These can still be regarded as completions given that the students were neither still studying nor did they withdraw. Contrary to concerns noted above of the problems of under-prepared students and their potential for non-completion, the more general point is that non-completion—whether that relates to the degree in which a student was

²For critiques of these, see respectively Gale et al. (2017), Connell (2007) and Bernstein (1999).

³Note that the OECD (2019a) report does not include all OECD countries, only those with available completion rate data.

⁴This includes “short cycle tertiary” programs that extends to those transferred to/graduated from master’s or doctoral programs, as well as those that transferred to a “master’s long first degree” (e.g. an undergraduate masters) (OECD, 2019a: 227). These amount to a very small proportion of students and but also represent system differences in terms of what is transfers between programs are possible and permitted in a given system (e.g. undergraduate masters programs do not exist in Australia, while transfer to an Further Education college is possible in England).

Table 1 Degree completion rates in selected OECD countries, compared with PISA 2018 scores in reading (by social background) and the score difference between least and most disadvantaged students.

Country	Completion rate by the Theoretical Duration (TD) 2017	Completion rate by the Theoretical Duration plus 3 years (TD+3) 2017	PISA 2018 equity in reading by social background classification	PISA 2018 reading score difference between advantaged (top quarter) and disadvantaged (bottom quarter ESCS)
United Kingdom	72	85	Around OECD average	80
Ireland	63	81	Better than OECD average	75
Lithuania	61	65	Around OECD average	89
Israel	60	83	Not as good as OECD average	121
Canada	48	n/a	Better than OECD average	68
Norway	44	72	Better than OECD average	73
Finland	43	73	Better than OECD average	79
Sweden	42	56	Around OECD average	89
France	41	67	Not as good as OECD average	107
Average	39	67	—	89
Switzerland	39	81	Not as good as OECD average	104
United States	38	69	Around OECD average	99
Iceland	36	69	Better than OECD average	72
New Zealand	35	77	Around OECD average	96
Australia	34	70	Around OECD average	89
Estonia	34	59	Better than OECD average	61
Brazil	33	50	Around OECD average	97
Flemish Comm. (Belgium)	33	67	Not as good as OECD average (Belgium whole)	109 (Belgium whole)
Portugal	30	65	Around OECD average	95
Netherlands	28	70	Around OECD average	88
French Comm. (Belgium)	27	54	Not as good as OECD average (Belgium whole)	109 (Belgium whole)
Austria	26	58	Around OECD average	93
Slovenia	24	53	Above OECD average	80
Chile	16	54	Around OECD average	87

Source: compiled by the authors, derived from OECD (2019a,b) data.

initially enrolled or that no qualification was awarded at all—can result in positive outcomes for students. Representing all non-completion as a problem effectively imposes institutional and system interests onto students.

An examination of student completions in UK⁵ HE systems provides a further example. These data show that the UK appears to fair quite well in terms of student completions when compared with other OECD nations, with 72% of the cohort graduating within the TD of their undergraduate degrees (3 years in England, Wales and Northern Ireland; four in Scotland). Whereas Switzerland appears to have significantly below average student completion rates. However, when the time taken by HE students to complete their degree is extended to the TD of their degree plus 3 years, the UK (with an 85% TD+3 completion rate) is much more on a par with other nations. By the same metric, Switzerland has an 81% TD+3 completion rate even though its TD student completion rate is comparatively low. The data suggest that when provided with more time, most students within a system do graduate. Perhaps more time spent in HE allows for more support although the reasons are likely to be more complex.

Indeed, explanations for differences between completion rates within TD and TD+3 are difficult to pin down, partly due to the enormous variation between HE systems in OECD nations. For example, because of privacy laws, German universities do not track the progress of individual students but only record progress at a cohort, year or degree level. Thus, they are less able to identify students in need of support before they drop out. A less specific reason for TD/TD+3 differences in Table 1 countries might be that less flexible systems—for example, those that mandate completion within a set period, or those that limit the amount of “time out” that can be taken as leave from study—might require students to complete sooner. Similarly, nations that charge relatively high tuition fees, or those that have dramatically increased their fees, incentivize students to complete sooner. This may go some way to explain why the UK has a very high completion rate within a program’s TD, with annual tuition fees of around £9000 for English universities. The change in approach to fees in England—before 1998 no tuition fees were charged; at their introduction they were a modest £1000 increasing to £3000 in 2006, and then rising dramatically to the present £9000 in 2012 (Riddell et al., 2018)—might influence how students now perceive the costs of prolonging their studies or of not completing at all. Yet, high

⁵The UK does not have a unified HE system but four systems, one in each of its four nations, which is not reflected in the OECD data. Granted, the systems in Wales and Northern Ireland are small and tend not to vary greatly from what occurs in England. HE in Scotland probably differs the most: with four-year undergraduate degrees the norm and fee-free student tuition, it is larger than Wales and Northern Ireland combined, and older, with several of its universities predating most in England. Grouping them together in the data as one system might be convenient, but it is also illustrative of the larger problem of making country comparisons.

tuition fees and student loans in the USA (as onerous as they are) have endured for decades, without a comparable effect on student completion rates (see [Table 1](#) for USA TD and TD+3 rates). And, if tuition fees do influence student completions within a degree program's TD, what is the effect on the accuracy of reported UK completion rates, of combining England and Scotland HE systems when Scotland does not charge fees for undergraduate courses?

Another possible explanation for differences in HE student completion rates relates to the extent to which OECD nations' schooling systems are equitable. The OECD's PISA (Program for International Student Assessment) measure—of 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges—uses the term “equitable” to refer to the distance between the highest performing 15-year-olds and the lowest performing, within a schooling system. The smaller the score difference the greater the equity in the system; the greater the score difference the greater the inequity of a system.⁶ We might surmise that in more equitable systems, school students who perform at the lower end of recorded PISA scores have acquired more of the “cultural capital” privileged by schooling ([Bourdieu, 1986](#)) than students within less equitable school systems, and this is thus more likely to translate to higher student completion rates in HE. In columns four and five, [Table 1](#) includes this PISA data for *reading* by students' social background classification, for each of the country HE systems compared by the OECD. Also included are PISA score differences between the top (most advantaged) quartile, and the bottom (least advantaged) quartile. It is 2018 data, illustrative of PISA assessments generally recorded in these countries in the 2010s. According to the data, the least equitable schooling system in [Table 1](#) is Israel, with 121 points difference between its highest and lowest PISA scores, along with a below average equity score in reading. Yet its HE system has relatively high student completion rates in HE. Conversely, Estonia has a more modest 59% HE student completion rate (even after TD+3) but has the most equitable schooling system of all nations listed in the Table (with a 61-point difference between the most and least advantaged students). In short, PISA data do not appear to be predictive of HE student completion rates and do not appear to be a good indicator of school students' levels of accumulated cultural capital, which would contribute to their HE completion.

Reimagining the problem

What stands out in the “solutions” above is the superficial, narrow, reductive constructions of the completion problem. These are largely built on comparisons and assumptions about students, degrees, institutions and systems informed by limited and sometimes questionable data about “the other,” particularly data that appear to be comparable (e.g. when the same terms are used but they mean different things to different people in different contexts or when variables are standardized in order to shoehorn data into a form that allows for their comparison). As we have suggested, the reality of non-completion can be quite different and far more complex.

To reiterate, non-completion can be a positive outcome for the student who decides mid-stream to switch from a teaching to a journalism degree, or a necessity for the student whose personal circumstances require them to commit their time to doing other things (e.g. secure paid employment; care for a family member; etc.) and so on. Students are not all the same; one set of circumstances and how students respond to them are not representative of all. Even when non-completion constitutes “failure” in HE terms, that is not always a bad thing; it can lead to considerable reflection and personal growth. It becomes more of a problem when particular kinds of students—i.e. population groups—are over-represented in non-completion statistics. It is exacerbated when the problem is attributed to the population group rather than to institutions and systems.

Similarly, degrees do not differ just in relation to their duration. They also vary in the transfer options available to students and in the fees students are charged for undertaking a degree. Perhaps more importantly, degrees vary in form and function ([Gale, 2021](#)): what is valued within the degree, how that is represented, what participation looks like, and what the degree's purposes are. To complicate this further, the form and function of degrees vary *within* disciplines, institutions and systems as well as *between* them. Degrees are not all the same and the student experience of undertaking a degree is not uniform. Consider, for example, the cohort mentality (e.g. “the class of 2025”) culturally instilled in HE students in USA HEIs, particularly in prestigious Ivy League universities and the Association of American Universities, which creates a “completion ideation”: i.e. a “psychological contract” ([Argyris, 1960](#); [Rousseau, 1989](#)), a commitment to one another, facilitating completion.⁷ This is quite a different way of being a HE student compared with some other HE systems. For example, compare the different expectations and experiences of HE students in England and Australia. In the transition from school to university, the majority of Australian HE students tend to continue living in the family home, commuting to their HEI on a daily basis. Whereas in England, especially for the middle classes, part of the university experience is to leave the family home and set up accommodation close to one's HEI, sharing house or university student “dorms” with one's peers. It is a rite of passage that goes hand in hand with the school-university transition. These comparisons contribute to a broader understanding of HE students “navigating change” ([Gale and Parker, 2014](#)), though not always in the same way.

In a further comparison, juxtapose the psychological with the “social contract” ([Rawolle, 2013](#)) of HE, between student and institution/system/nations, which purports to offer students who complete their degree with certain benefits: e.g. social mobility; better employment opportunities and conditions, including higher than average wages; better than average health and life

⁶These are relative assessments of system equity rather than an absolute measure of equity comparable across nations.

⁷We are indebted to Lizzie Knight for her observations of cohort mentality, psychological contracts and completion ideation, and how these play out in different HEIs and systems.

expectancy; and so on. Indeed, several HE systems justify their student fees based on these promised benefits (e.g. the rationale for differential fees in Australian education until 2020; Gale and Parker, 2018). Increasingly, though—especially for “late-comers” (i.e. marginalized groups)—the “democratization” of HE (i.e. WP) delivers a “blighted hope or frustrated promise” (Bourdieu, 1984: 150). Whereas students once entered university with “high hopes of what the knowledge economy has to offer, fueled by the expansion of higher education ... the labor market cannot keep pace with social expectations of work, rewards and status” (Brown, 2003: 150). There is no longer certainty that a simple undergraduate degree offers a positional advantage. With the social contract between HE student and institutions/systems/nations undermined, incentives to complete are diminished.

Yet even to broaden out understandings of HE completion, to complicate things, to include what it means for students (not just for institutions, systems, nations), does not delve deep enough into why non-completion has become a problem in the first place. The place to begin that investigation is from the perspectives of students, especially those from disadvantaged, marginalized groups. Instead of explanations for why some students drop out and what can be done, we should be asking why students would *want* to participate in HE, what is in it for them not just what is in the national interest, and what HEIs and systems can do about that. This is more than just including explicit accounts of how to engage with HEIs, how to negotiate their systems and cultures, what to do to be successful, to complete. Those are supports essentially framed by the interests of HEIs and their systems.

The disjuncture between what HEIs want (high completion rates) and what some students experience (an alienating HE) could be likened to what Connell (2007) refers to as power relations in the realm of knowledge; center-periphery relations between global north and south. Even with all the support, with a whole-of-institution approach, students from marginalized groups can feel as though they are not wanted, that to stay means to reject their own cultures and histories. Indeed, for these very reasons some young people decide not to participate in HE at all. Drawing on Sen (1999), Watts and Bridges (2006: 287) observe that some young people “reject HE because it does not enhance their freedom to lead lives that they value and have reason to value.” That is a reasonable position to take but it also conveniently leaves HEIs and their systems without the need to change very much. Instead, what if HE *did* enhance students’ freedoms to lead lives that they value and have reason to value, not just for mainstream students but also those from disadvantaged backgrounds? What if HE afforded them a compelling reason to participate? That would require HEIs to focus on what students are participating in; what the HE equity literature describes as a concern for “access to what” (Gale, 2011).

If we were to focus on that, perhaps HE student completion wouldn’t present as such an issue.

Acknowledgments

Thanks to Laela Adamson for chasing up the data for the Table and her related contributions to our thinking. Thanks too to Lizzie Knight for her reflections on psychological contracts in the context of HE student completions. Of course, responsibility for any shortcomings in the essay remains with us.

References

- Australian Government, 2009. Transforming Australia’s Higher Education System. DEEWR, Canberra.
- Argyris, C., 1960. Understanding Organizational Behavior. Dorsey Press, Homewood, IL.
- Bernstein, B., 1999. Vertical and horizontal discourse: an essay. *Br. J. Sociol. Educ.* 20 (2), 157–173. <https://doi.org/10.1080/01425699995380>.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, Westport, Connecticut, pp. 241–258.
- Bourdieu, P., 1984. *Distinction: A Social Critique of the Judgement of Taste*. R. Nice, Trans.. Routledge & Kegan Paul, London.
- Brown, P., 2003. The opportunity trap: education and employment in a global economy. *Eur. Educ. Res. J.* 2 (1), 141–179. <https://doi.org/10.2304/eeerj.2003.2.1.4>.
- Connell, R.W., 2007. Southern Theory: The Global Dynamics of Knowledge in Social Science. Routledge, Abingdon.
- Cervini, E., 2012. More Students, But Are They Doomed to Fail? The Age, 13 March.
- DfES [Department for Education and Skills], 2003. The Future of Higher Education. Cm 5735. HMSO, Norwich.
- DESE [Department of Education, Skills and Employment], 2020. Selected Higher Education Statistics—2019 Student Data: Equity Groups from. <https://www.dese.gov.au/higher-education-statistics/student-data/selected-higher-education-statistics-2019-student-data>. (Accessed 17 November 2021).
- European Commission, 2015. Dropout and Completion in Higher Education in Europe: Main Report. European Commission, Brussels.
- Gale, T., 2011. Student equity’s starring role in Australian higher education: not yet centre field. *Aust. Educ. Res.* 38 (1), 5–23. <https://doi.org/10.1007/s13384-010-0007-3>.
- Gale, T., 2021. Higher vocational education as a work of art. In: Knight, E., Bathmaker, A.-M., Moodie, G., Orr, K., Webb, S., Wheelahan, L. (Eds.), *Equity and Access to High Skills Through Higher Vocational Education*. Palgrave.
- Gale, T., Parker, S., 2014. Navigating change: a typology of student transition in higher education. *Stud. High. Educ.* 39 (5), 734–753. <https://doi.org/10.1080/03075079.2012.721351>.
- Gale, T., Parker, S., 2017. Retaining students in Australian higher education: cultural capital, field distinction. *Eur. Educ. Res. J.* 16 (1), 80–96. <https://doi.org/10.1177/1474904116678004>.
- Gale, T., Parker, S., 2018. Student tuition fees in Australian higher education: a litany of public issues and private troubles. In: Riddell, S., Minty, S., Wheedon, E., Whittaker, S. (Eds.), *Higher Education Funding and Access in International Perspective*. Emerald, Bingley, pp. 223–240.
- Gale, T., Tranter, D., 2011. Social justice in Australian higher education policy: an historical and conceptual account of student participation. *Crit. Stud. Educ.* 52 (1), 29–46.
- Gale, T., Mills, C., Cross, R., 2017. Socially inclusive teaching: belief, design, action as pedagogic work. *J. Teach. Educ.* 68 (3), 345–356. <https://doi.org/10.1177/0022487116685754>.
- Gallagher, M., 2009. Bradley’s Flawed Vision. *The Australian*, 18 February.
- Hare, J., 2015. Annual ATAR row erupts anew in an arena still shrouded by opacity. *The Australian*, 21 January. <https://www.pressreader.com/australia/the-australian/20150121/282741995193622>.

- Higher Education Statistics Agency [HESA], 2020. Higher Education Student Statistics: UK, 2018/19—Student Numbers and Characteristics. <https://www.hesa.ac.uk/news/16-01-2020/sb255-higher-education-student-statistics/numbers>.
- Higher Education Statistics Agency [HESA], 2021. Figure 6—UK Domiciled Full-Time HE Undergraduate Student Enrolments by Participation Characteristics 2018/19 to 2019/20. <https://www.hesa.ac.uk/data-and-analysis/sb258/figure-6>.
- Kift, S., Nelson, K., Clarke, J., 2010. Transition pedagogy: a third generation approach to FYE—a case study of policy and practice for the higher education sector. *Int. J. First Year High. Educ.* 1 (1), 1–20.
- Knight, E., Bathmaker, A.-M., Moodie, G., Orr, K., Webb, S., Wheelahan, L. (Eds.), 2021. *Equity and Access to High Skills Through Higher Vocational Education*. Palgrave.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141. <https://doi.org/10.1080/00405849209543534>.
- Óhidy, A., 2018. Widening participation in higher education: policies and outcomes in Germany. In: Riddell, S., Minty, S., Wheedon, E., Whittaker, S. (Eds.), *Higher Education Funding and Access in International Perspective*. Emerald, Bingley, pp. 163–183.
- Organisation for Economic Co-operation and Development [OECD], 2019a. *Education at a Glance 2019: OECD Indicators*. OECD Publishing, Paris. <https://doi.org/10.1787/f8d7880d-en>.
- Organisation for Economic Co-operation and Development [OECD], 2019b. *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OECD Publishing, Paris. <https://doi.org/10.1787/b5fd1b8f-en>.
- Quinn, J., 2013. Drop-Out and Completion in Higher Education in Europe Among Students From Under-Represented Groups. Report for the EU Commission.
- Rawls, J., 1999. *A Theory of Justice*, revised ed. The Belknap Press of Harvard University Press, Cambridge, Mass.
- Rawolle, S., 2013. Understanding equity as an asset to national interest: developing a social contract analysis of policy. *Discourse* 34 (2), 231–244. <https://doi.org/10.1080/01596306.2013.770249>.
- Riddell, S., Minty, S., Wheedon, E., Whittaker, S., 2018. Introduction: higher education funding and access in international perspective. In: Riddell, S., Minty, S., Wheedon, E., Whittaker, S. (Eds.), *Higher Education Funding and Access in International Perspective*. Emerald, Bingley, pp. 1–11.
- Rios-Aguilar, C., Kiyama, J.M., 2018. Introduction: the need for a funds of knowledge approach in higher education contexts. In: Kiyama, J.M., Rios-Aguilar, C. (Eds.), *Funds of Knowledge in Higher Education: Honoring Students' Cultural Experiences and Resources as Strengths*. Routledge, New York, pp. 3–6.
- Rousseau, D.M., 1989. Psychological and implied contracts in organizations. *Empl. Responsib. Rights J.* 2 (2), 121–139. <https://doi.org/10.1007/BF01384942>.
- Sen, A., 1999. *Development as Freedom*. Knopf, New York.
- Statista, 2021. Number of Students in Universities in Germany During Winter Semesters From 2002/2003 to 2020/2021 from. <https://www.statista.com/statistics/584061/university-student-numbers-winter-semesters-germany/>. (Accessed 17 November 2021).
- Thomas, L., 2002. Student retention in higher education: the role of institutional habitus. *J. Educ. Pol.* 17 (4), 423–442.
- Thomas, L., 2021. Student retention: the need for institutional change. In: Shah, M., Kift, S., Thomas, L. (Eds.), *Student Retention and Success in Higher Education: Institutional Change for the 21st Century*. Springer International Publishing, Cham, pp. 41–61.
- Tinto, V., 1993. *Leaving College: Rethinking the Causes and Cures of Student Attrition*, second ed. University of Chicago Press.
- Tinto, V., 2008. Access Without Support is Not Opportunity. Paper Presented at the 36th Annual Institute for Chief Academic Officers, The Council of Independent Colleges, November, in Seattle.
- Trow, M., 1974. Problems in the transition from elite to mass higher education. In: *Policies for Higher Education*. OECD, Paris, pp. 51–101.
- Trow, M., 2006. Reflections on the transition from elite to mass to universal access: forms and phases of higher education in modern societies since WWII. In: Forrest, J.J.F., Altbach, P.G. (Eds.), *International Handbook of Higher Education, Part One: Global Themes and Contemporary Challenges*. Springer, Dordrecht, pp. 243–280.
- Truszczyński, J., 2013. Forward. In: Quinn, J. (Ed.), *Drop-Out and Completion in Higher Education in Europe Among Students From Under-represented Groups*. Report for the EU Commission, p. 7.
- Ulriksen, L., Madsen, L.M., Holmegaard, H.T., 2017. The first-year experience of non-traditional students in Danish science and engineering university programmes. *Eur. Educ. Res. J.* 16 (1), 45–61. <https://doi.org/10.1177/1474904116678628>.
- Watts, M., Bridges, D., 2006. The value of non-participation in higher education. *J. Educ. Pol.* 21 (3), 267–290. <https://doi.org/10.1080/02680930600600267>.
- Zepke, N., 2021. Student engagement: key to retaining students. In: Shah, M., Kift, S., Thomas, L. (Eds.), *Student Retention and Success in Higher Education: Institutional Change for the 21st Century*. Springer International Publishing, Cham, pp. 63–87.

Women in higher education: exclusion, inclusion and seclusion

Venitha Pillay, University of Pretoria, Faculty of Humanities, Department of Sociology, Pretoria, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	381
Structural framework	381
The politics of being a feminist researcher	382
Conceptual clarity	383
Looking back	384
In but still out	384
The mirage of work/life balance	386
The daily bread of being black in academia	387
The curveball of Covid-19	388
Winning ways	389
Conclusion	390
References	390

Introduction

My original intention when I began this chapter was to use Butler's (1990) notion of performativity to interrogate the ways in which women academics perform gender in higher education institutions. In *Gender Trouble*, Butler (1990) questions whether, "if gender is constructed, it could be constructed differently, or whether its constructedness implies some form of social determinism, foreclosing the possibility of agency and transformations" (pp. 10–11). The question she poses rests on the recognition that "there is no gender beyond the expressions of gender" (Butler, 1990, p. 34); that identity is constituted through performativity, that is, "a set of repeated acts within a highly rigid regulatory frame" (Butler, 1990, p. 45). My intention was to use the data collected from interviews with twenty South African women academics to examine how, if at all, women academics were performing gender in ways that sought to champion and/or change their epistemological, ontological and gendered occupation of higher education institutions. But as I worked through an extensive literature review, I found myself embedded in narratives of exclusion, inclusion and seclusion that permeated the journeys of women's presence in higher education institutions. Given the demands of a handbook to be all-embracing, I opted instead to pursue a literature review that focused on the leitmotifs of exclusion, inclusion and seclusion that repeatedly rear their heads in the literature. In making this decision, I doff an apologetic hat to Butler¹ and the women I interviewed, whose stories I shall tell in a different setting.

Structural framework

I begin this review with an affirmation of my political positionality as a feminist black² woman researcher from the Global South. I identify standpoint theory and intersectional feminisms as foundational heuristics that guide my positionality. In this context I make an unapologetic decision to foreground literature about women academics from the Global South. I use the term Global South as a metaphor for underdevelopment which "marks a shift from a focus on development or cultural difference to an emphasis on geopolitical power relations" (Dados and Connell, 2012, p. 12).

I then provide a brief description of the concepts of inclusion, exclusion, and seclusion which serve as broad thematic frames in the literature on women in academia. I show that inclusion and exclusion are used to illustrate that while women are unequivocally *in* higher education, they remain largely *out* of positions of power. I show that acts of seclusion are used by women and institutions as strategies to blur or conceal aspects of women's gendered identity, for example of motherhood, and gendered experiences, for example of sexual harassment, at universities.

Guided by the broad frames of inclusion, exclusion, and seclusion, I identify three prominent themes in the literature on women in academia. The first is that women academics have staked an unshakable and permanent place in academia. The second is that

¹In doffing my hat to Butler, I pay respect to the importance of Butler's notions of performativity but also recognize that it will not adequately serve the purpose of this article. Instead, I plan to return to Butler when I write the stories of the women I interviewed.

²I purposefully use the term black women and women of color interchangeably as a linguistic strategy of inclusiveness. The choice of identifying as a black woman or as a woman of color is arguably rooted in one's own political positionality and context. Given the broad scope of this review, I do not choose a single descriptive term. Suffice to say that I self-identify as a black woman, knowing full well that other women may not perceive me as such. My upbringing and identification in the anti-apartheid struggle as a black woman is a political position that is irrevocable. I also use the terms interchangeably as a form of respect to the writing being referenced at any given point in the review. Simply, I echo the term that author being reviewed uses. I acknowledge that both terms offer limited markers of marginalization and that women's experiences of marginalization is complex, multi-layered and contextually nuanced.

their absences in positions of leadership, in research, in the professoriate and in selected disciplines, persist. The third is the competing demand of work and family responsibility. It is arguable that these themes have validity for all women academics, irrespective of race, class or location in the Global South or Global North, albeit to varying extents and ways. However, being a black woman in academia is imbued with its own challenges. I therefore focus on some of the specific challenges experienced by black³ women academics. In the wake of Covid 19, I examine women's responses to the inescapable condition of working from home. In the final section I provide brief examples of women academics' strategies to build their careers and claim professional authority.

The politics of being a feminist researcher

This review is guided by two mutually supportive heuristic frames, that of standpoint theory and intersectional feminism. Fundamentally, Harding's (2004) standpoint theory is rooted in the idea that the standpoint of women is not simply a different perspective from that of a normative masculine world, but that it is a perspective that is imbued with particular moral and political commitments and that everyday experiential accounts of historically marginalized people are a legitimate site of scientific knowledge. Collins (2004) fortified the notion of marginalization through her focus on women of color, albeit, in the main, a focus on African American women. Intersectional feminism extends standpoint theory in that it acknowledges and addresses the differences between women. Davis (2008) argues that the strength of intersectional feminism is that it has been "devoted to understanding the effects of race, class, and gender on women's identities, and struggles for empowerment. It has especially been concerned about the marginalization of poor women and women of color within white, Western feminist thought." Furthermore, intersectional feminism's embrace of postmodernism provides the opportunity to welcome the notion of "shifting and multiple identities" (Davis, 2008, p. 71). In sum, my political positions shaped by, inter alia, my location as a black feminist from the Global South, are etched in this review, consciously or unconsciously, informing the inclusions, exclusions and seclusions I opt for.

As a feminist researcher I accept few certainties. One of them is that research is never innocent or apolitical. As an intersectional feminist and a black from the Global South, my research will always be driven by these immutables. Simply, I rely on Davis's argument that one's feminist location(s) is an analytic resource, not just a social marker. (Davis, 2008, p. 72), I contend that my location in the Global South is an analytic resource that is instructive of how I tell the story of women's exclusion, inclusion and seclusion in higher education.

As an intersectional feminist from the Global South, I am painfully conscious that, globally, research about women frequently tacitly and explicitly excludes women of color and women from the Global South. These are not innocent mistakes or accidental exclusions. I do not expect all research about women to be inclusive of all women, nor do I suggest that this is feasible. However, in my experience, I find that the research produced and published in the Global North, with a few exceptions, makes little to no effort to include research from and about the Global South. Those of us in the Global South have to find our own ways of talking and thinking about how to find a fit for ourselves in this world that consistently excludes us. For example, I was excited when I saw the title of a doctoral thesis: "Women being and becoming academics: exploring gendered career journeys and their implications for career development."⁴ It seemed to have strong resonance with my research and I was eager to read a study about women's academic journeys, only to discover that the study was of five women academics in the United Kingdom. The lure of inclusion implied in the title took me to the same old place of exclusion. The racial identity of participants is not offered which left me believing that they may well be white, given the established normativeness of whiteness and the otherness of being black.⁵

In their study of articles of women in higher education in 15 selected education journals published between 2000 and 2013, in Australia, the United Kingdom, Europe and North America, Cama and Jorge (2016) identified four themes. The first was aimed at analyzing sex differences in faculty tenure and promotion, the second examined women in higher education leadership and management, the third was related to gender earning gaps, and the fourth was linked to policies aimed at promoting gender equity. The journals selected were exclusively from the Global North and provided no recognition of the challenges experienced by women from the Global South. Despite the probability that there may well be several points of resonance and recognition between women from the Global South and North, the disparity in geopolitical power relations and knowledge production meant that women from the Global South were left to fend for themselves to find resonances in a study that explicitly excluded them.

Over the years experiences of exclusion have become a familiar and I have to constantly caution myself to expect less. Simply, unless research consciously signals that it is inclusive of women from the Global South or about women of color, it is rarely about us.

It is in this context that I make a clear political decision as I write this chapter: that is, to make every effort to foreground research produced by the Global South about the Global South. I believe it is time for the Global North to make its own efforts to find resonances with the Global South. As an intersectional feminist, it is my responsibility to disrupt established, invidious and entrenched norms of disempowerment. I argue that research that fails to pay even scant attention to writers of the Global South is a conscious act of disempowerment. In so doing, I contend that the scope of intersectionality should extend beyond the commonly held

³I use black women here to refer to all women of color. I note that not all women of color describe themselves as black women.

⁴Clayton, S. 2015. Women being and becoming academics: exploring gendered career journeys and their implications for career development. University of Sussex Center for Higher Education and Equality Research.

⁵See Spivak, G. C. 1996. *The Spivak Reader: selected works of Gayatri Chakravorty Spivak*. New York Routledge.

precepts of race, class, age, gender, sexuality and religion to include the vast differences in access to knowledge and knowledge dissemination between the Global South and the Global North. Therefore, this review is written with a disruptive imperative in mind and seeks to denounce the seeming innocence of the literature review as a guileless description of relevant writings.

Having said that, I am keenly aware of the limitations of the political imperatives I have set for myself. The literature from the Global South referred to in this chapter is predominantly African with a disproportionate focus on South Africa. Despite an extensive search of relevant journals published on the African continent, the literature about women's successes, resistances and challenges in higher education is scant. For example, the last 10 issues of the *International Journal of African Higher Education* do not contain a single article about women academics on the continent or in the African Diaspora.⁶ Indeed, there is a remarkable absence of research about gender in journals that do not, by definition, focus on gender, implying that gender is a "separate" issue meant only for journals about gender.⁷ I rely extensively on the *South African Journal of Higher Education* (SAJHE), owing to its education focus and the inclusion of research about gender from across the continent, albeit dominated by research about South Africa. Publications emanating from non-English-speaking countries in the Global South were not accessible to me.

The review below is therefore undertaken with the "political and moral commitment" (Harding, 2004, p. 1) to foreground the literature from the Global South. In so doing, it pays homage to intersectional feminism (see Davis, 2008) alongside standpoint theory (see Harding, 2004).

Conceptual clarity

Inclusion and exclusion are familiar concepts that have been used extensively to illustrate that while women are unequivocally *in* higher education (inclusion), they remain largely *out* of positions of power (exclusion). The notion of seclusion is one that I introduce⁸ here. I use seclusion to refer to the ways in which women choose to hide or offer only partial glimpses of their coping strategies and/or gendered identities as they perform and/or disrupt the gendered behavior patterns expected of them in universities. Expressions of seclusion are evident when women academics choose to hide their family responsibilities in the workplace, in their silences about the trade-offs they find themselves making as they try to be caregivers and professionals, and when they pretend they don't have children, or stop breastfeeding too soon, or never say no when given unreasonable workloads relative to the men at the same level (see Bassett, 2005; Connelly and Ghodsee, 2011; Moe, 2003; Ward and Wolf-Wendel, 2012). Hudock's furtive way of surviving academic motherhood is a telling expression of seclusion: "The ivory tower seems to have no room for children or grocery shopping or laundry. I remember the feeling of being caught by students in the act of grocery shopping, and they were always surprised that I eat. Imagine if they saw me breastfeeding" (2009, p. 63).⁹ Seclusion of her gendered identity as a mother and the expression of discomfort at being "caught" in the act of a being a 'normal' person, reveals the extent to which women hide/seclude themselves in order to survive in academia.

A different example of seclusion is illustrated by a common theme in the study of academic mothers in India, Brazil and South Africa (Pillay et al., 2017). Almost all the women interviewed were at pains to distance themselves from aspirations to climb the academic ladder or even to name themselves as academics. Instead, they foregrounded their motherhood and presented their academic selves as secondary, incidental or accidental.

The stories of these academic mothers suggest that adopting a demure stance towards scholarship, while relishing its potential for empowerment, is a coping strategy and a means towards legitimizing their enjoyment of work, especially in the face of social expectations that they should be mothers above all else ... they downplay their academic aspirations, suggesting that they are not "really academics" and frequently expressing their desire to be good mothers. Muting their academic selves, while expressing the importance of their motherhood, may serve as a way to lessen the guilt and potential social censure

(Pillay et al., 2017, pp. 176–177).

The women in the study chose seclusion, consciously or unconsciously, of aspects of themselves as a strategy to manage the gendered demands of home and work, and the social and cultural expectations implicit in their contexts. In Pillay et al. (2017) the social and cultural expectations of women from India, Brazil and South Africa is given significant attention. Despite localized nuances, there was broad recognition that women were expected to primarily to be caregivers, homemakers, and mothers, that professional responsibilities were secondary to these, that she was expected to marry, to be actively engaged in extended family events and to observe traditional responsibilities allocated to women. Smith (2011) describes similar cultural expectations of Latina women. It is arguable that in the context of these social and cultural expectations women academics obscure/seclude their professional ambitions while vocalizing their family responsibilities, as strategic devices that deflect attention from their professional selves. Such deflection also provides space to enact their professional selves without drawing attention.

⁶(see <https://ejournals.bc.edu/index.php/ijah/issue/archive>).

⁷I undertook an extensive review of *Agenda*, the first feminist journal in Africa, and *Feminist Africa*, another feminist journal on the continent, but found no recent articles relevant to this chapter.

⁸To the best of my knowledge the concept has not been used to describe women's engagement with higher education institutions.

⁹See also Robinson and Clardy (2010), Smith Silva (2011), Connelly and Ghodsee (2011), Ward and Wolf-Wendel (2012).

Institutions, too, perform acts of seclusion when they hide salary disparities between male and female academics or when they pretend that the playing field for woman and male academics is even or when they refuse to admit that academic fathers do not go home to the same responsibilities that academic mothers do (American Council on Education, 2018; Cama and Jorge, 2016; She Figures 2021; UNESCO, 2021).

Institutions practice acts of seclusion each time they hide acts of sexual harassment and violence against female staff and students. Focusing on the European context, Bondestam and Lundqvist show that “precarious working conditions, hierarchical organizations, a normalization of gender-based violence, toxic academic masculinities, a culture of silence and the lack of active leadership are all key features enabling sexual harassment” (2020, p. 397). In an online study of over 16,000 university students in the United States, Wood et al. (2018) show that 19% had experienced sexual harassment from staff and 30% from peer perpetrated harassment. In their research that focused on public affairs education programs, Young and Wiley (2021) show that 13% of women experienced sexual harassment at the hands of faculty. In April 2016 the #EndRapeCulture protest spread across South Africa in response to, inter alia, repeated and increasing levels of sexual violence against women at universities (Gouws, 2018; Orth et al., 2020). A common theme across these articles is that universities continue to remain complicit in silencing, secluding and obscuring acts of sexual harassment and violence.

I suggest that exclusion, inclusion and seclusion are not mutually exclusive but that they function as mutually constitutive, occurring simultaneously in the lives of women academics. I therefore do not structure the chapter into neat sections on exclusion, inclusion and seclusion. Instead, I suggest that the fluidity of these concepts, alongside the experiences of the individual woman academic, should encourage women academics to construct their own meanings, interpretations and instances of exclusion, inclusion, and seclusion.

Looking back

Within the above political framework of foregrounding the Global South that I set for myself, albeit imperfect, I opted to first provide a brief overview of the trajectory of the literature about women in academia. I begin this section by establishing the progress women have achieved in their representation in higher education. I then focus on two key pervasive challenges: their absence from positions of leadership, and the challenges implicit in the competing demands of family and work. I rely on the research I have conducted and my extensive engagement with a literature base that extends beyond the Global North, to identify the key issues, albeit not exhaustive, discussed in this review. I then examine the additional forms of exclusion that black women in academia experience. I follow this with a brief overview of the impact of Covid-19 on women academics. In the final section of this literature review, I identify research about women’s resistance to established and entrenched tropes of disempowerment.

In but still out

The global literature of the last five decades on women in academia offers a traceable track, arguably kicked off by a focus on the absence of women academics in the historically male-dominated and masculine environments that define universities. From the groundbreaking research produced by Acker in the 1980s and ’90s and Simeone,¹⁰ there is now a plethora of literature that tracks the presence of women in universities that have for centuries been an enclave of men, where knowledge production and dissemination actively excluded women (see Elsevier, 2020; She Figures, 2021; Squazzoni et al., 2020; UNESCO, 2021).

Today, it is unequivocal that women have staked their place in higher education: in 2018, women represented 43% of teachers in tertiary education globally (UNESCO, 2021, p. 25). However, despite their significant numeric presence in universities, women comprise only 30% of the world’s researchers at universities (UNESCO, 2021). Table 1 shows the regional distribution of women researchers in 2017.

The lower number of women researchers correlate with their lower representation in the professoriate given the higher value placed on research for promotion. Lower research outputs may be attributed to women academics’ dominance in teaching forward positions that overwhelm the space and time for research. The high value placed on research and the devaluing of teaching for promotion, suggests that the returns to women academics for the extensive time devoted to teaching, is small, and makes no significant contribution toward their career development (Krebsbach, 2022). That outcomes for women academics, in the form of promotion is improved when teaching and research carry equal value for promotion is borne out by research conducted by Subbaya and Vithal (2017). In a study of promotions at a South African university where teaching and research have been equally evaluated, they found that “although more men applied for promotion¹¹ ... the overall success rates for women were higher, [and that] there were no statistically significant differences in the success rates between males and females”. They also show that “there was an

¹⁰See Acker (1980), Acker (1994), Simeone (1987).

¹¹In research not yet published I found that women academics are often hesitant to apply for promotion and doubt their promotability. A female colleague that I mentored was surprised when I suggested she apply for promotion and noted that none of her line managers had suggested she apply. She applied and was successful. The literature on the value of mentoring for women academics is extensive. For example, see Vongalis-Macrow (2014) Career moves: mentoring for women advancing their career and leadership in academia.

Table 1 Women researchers by region, 2017

<i>Region</i>	<i>% of women researchers</i>
Central Asia	48.5
Latin America and the Caribbean	45.8
Arab States	40.9
Central and Eastern Europe	39.0
North America and Western Europe	32.9
Sub-Saharan Africa	31.1
East Asia and the Pacific	25.0
South And West Asia	23.1

Extracted from [UNESCO \(2021\)](#), p. 22).

overall equitable outcome at the level of full professor where excellence in teaching and research were compulsory for success” (944–945).

The UNESCO publication (2021) referenced above also shows that women are underrepresented in senior positions; are over-represented among teaching staff; that STEM (science, technology, engineering and mathematics) fields are heavily skewed in favor of male academics; and that, on average, men publish more articles than women, especially in high impact journals.¹²

There is clear consensus in the literature that although the number of women academics in higher education has increased significantly in the last two decades, women remain a minority at universities, and they continue to populate the lower levels of the academic ladder. There is also consensus that the masculine monopolization of academic leadership and professorship persists.

In South Africa, although the number of women academics in the system has increased substantively in the last five years, they constitute the largest portion of temporary staff at the lowest levels of the academic ladder while professorships remain largely male (*VitalStats*, 2016).

In 2019, there were twice the number of male professors than female professors (down from three times the number in 2010). The difference in the number of male (4018) and female (4668) lecturers in 2019 is not significant, suggesting that the number of males and females at the starting block of lectureship may be similar but men reach the “finish line” of professorship more speedily than women. So, while entry into academia does not appear to be the stumbling block for women academics, progress through the system is. Of the 26 public universities in South Africa, four are headed by women, one of which, the University of Cape Town, is the highest ranked institution on the continent (*Mail and Guardian*, 20 August 2020). In 2019, 46% of professors in Brazil were women, and in India 27.3%, while in Australia only 33.9% of people holding positions above senior lecturer were women. In Canada, women are near parity at all academic levels except professorship (*UNESCO*, 2021, pp. 25–27).

According to *She Figures* (2021), which tracks gendered data in the European Union (EU), despite targeted policy initiatives “the under-representation of women in senior academic and decision-making positions in the EU continues to be a significant issue” (p. 7). According to *She Figures*, in 2019 women comprised 42.3% of all academic staff, 26.2% occupied senior positions and 47.1% were in junior positions.¹³ Put differently, the percentage of women academics at the lowest levels is almost twice that of women academics at the highest levels. The report shows that women comprise only 26% of Grade A positions, that is, the most senior academic position (p. 176). In STEM fields the number of women at Grade A level is 19% (see also *Morley*, 2013; *Redmond et al.*, 2017; *Subbaye and Vithal*, 2017; *UNESCO*, 2021).

In a study commissioned by the British Council, *Morely and Crossouard* (2015) investigated the presence/absence of women in positions of leadership at universities in South Asia (Afghanistan, India, Nepal, Sri Lanka, Pakistan, and Bangladesh). Not surprisingly, the study found that although there is a vast increase in the number of women graduating from higher education institutions, the number of women in academic leadership positions remains miniscule. The study found that socio-cultural issues like family responsibility and cultural norms about the role of women was a key factor that contributed to the absence of women leaders. In a study that included 40 participants from South and East Asia, the Middle East, North Africa, Australasia, and Europe, *Morely* (2014) points to the perception among women academics that women, if they are leaders, are often responsible for internal affairs like teaching support and men are better suited to managing external affairs. So, even when women are leaders, what they lead may also be determined by perceived gender differences.

The global data with respect to women academics in senior and leadership positions are rarely racially disaggregated. For example, *She Figures* (2021) and the *UNESCO* (2021) report offer gender disaggregation of the data but no race profiles of the data provided. Given the geographic focus of these reports, it is likely the data with respect to women refers mainly to white women. Data on the inclusion and exclusion of black women in leadership is usually to be found in publications that focus specifically on black women. For example, the *American Council on Education* (2018) report, “Voice from the Field: Women of Color Presidents in

¹²See also the *Elsevier* (2020). *The researcher journey through a gender lens* for a more detailed gendered analysis of, inter alia, research output, grants, scholarships and patents. It should be noted that although the publication is signaled as a global report, no African countries are included and Japan is the only Asian country included. These exclusions are by virtue of the methodological decision to focus on countries had a defined number of active authors in the selected time period 1999–2003 and 2014–8. The presentation of the report as a global report is therefore misleading.

¹³*She figures*. <https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/61564e1f-d55e-11eb-895a-01aa75ed71a1>.

Higher Education”, indicates that in 2016 only 5% of college and university presidents were black women. In 2014, 26% of college and university presidents in the United States were women (of all races) (see <https://www.forbes.com/sites/lucielapovsky/2014/04/13/why-so-few-women-college-presidents/?sh=593aa2466bcf>).

Women students’ access to higher education echoes similar patterns to those of women academics. Globally, the number of female undergraduate students has exceeded that of males since 2002. (UNESCO, p. 13).¹⁴ However, this phenomenon of the “female advantage” excludes STEM fields and ceases at the level of postgraduate study. Except for Central Asia, globally there are far fewer females than males at doctoral level. In South Africa, in 2019, only 3% of all African women enrolled in higher education were registered for Ph.D. study, as opposed to 4% of all white women. The fact that African women comprise more than 40% of the South African population and white women only 5% of the population illustrates the stark racial divide in female student achievement in higher education in the country.¹⁵

The American Council on Education reports that since 2006, women have earned more than 50% of all doctoral degrees in the country but in 2015 comprised only 32% of full professorships at degree-granting postsecondary institutions.

In sum, although women have staked an unshakeable claim to a place in higher education, their access to senior academic and leadership positions remains small. So, while women are unequivocally *in* higher education, they remain largely *out* of positions of leadership and power. Additionally, unless stated otherwise, discourses about women in higher education is implicitly about white women from the Global North. Put differently, white women from the Global North are more *in* than black women and women from the Global South. The challenges of entering and staying in higher education are different for black and white women. So, too, is their challenge in striving for and succeeding in leadership.¹⁶

The mirage of work/life balance

Alongside the recognition of the increased presence of women at universities and their absence in leadership positions, the literature on the challenges and obstacles experienced by women in academia is also substantial. Perhaps the most commonly cited challenge is the impossibility of balancing work and family demands, what Morley (2013) refers to as being “caught between two greedy institutions – the extended family and the university” (p. 122). The limitations on career development imposed by family demands, especially, but not only, in culturally restrictive contexts and for women of color, has begun to take a firm seat in the literature on women in higher education (see Hunter and Leahey, 2010; Jay et al., 2010; Motherhood Initiative for Research and Community Involvement (MIRCI), 2003; Mudaly, 2015; Obers, 2014; Pillay et al., 2017; Toffoletti and Starr, 2016; Wolfinger et al., 2008; Zulu, 2011).

The impact of family demands on women’s academic careers is deep seated. In a study of 14 women academics at Rhodes University in South Africa, Obers found that “the most productive male authors far exceed the most productive women in terms of quantity of peer reviewed article input” (2014, p. 1109), a trend that is also evident in other parts of the world (UNESCO, 2021; She Figures, 2021). Obers shows that, not surprisingly, the women in the study were adamant they *chose* to prioritize their family responsibilities over career development. The women involved in Pillay’s study (Pillay et al., 2017) made similar assertions. Toffoletti and Starr (2016) point to the myth of free choices that “women academics may cling to in an effort to minimize their discomforts and frustrations”. Specifically, they point to the

... power of work-life balance discourses to craft female academics as self-realizing agents who shoulder the responsibility for successfully managing the domains of work and life. [These discourses] also expose the discursive power of the concept to reinforce structurally embedded gender inequalities by masking the institutional forces (such as increased workplace demands and performance expectations ...) that impact the seemingly free “choices” work-life balance policies “empower” women to make

(Toffoletti and Starr, 2016, p. 490).

Toffoletti and Starr argue that discourses surrounding work-life balance are deeply gendered and that, in the main, women academics find this balance impossible to achieve. The literature is replete with research that bears testimony to the inevitability of failure as women academics persist in the futile search for balance between work and family responsibilities while simultaneously feeling crushed by guilt for not giving enough to either (see Connelly and Ghodsee 2011; Evans and Grant, 2009; Pillay et al., 2017; Ward and Wolf-Wendel, 2012). In the conclusion of *Academic Mothers*, Pillay asserts that “feminism needs to renounce the very idea of balance It [predicts] an evaluative frame within which we measure ourselves daily and constantly and inevitably fail. For as long as we describe women’s efforts at mothering and working as a balancing act, we are complicit in sublimating our efforts to a discourse of dysfunctionality” (2007, p. 148). Following this, the plethora of literature that focus on the impossibility of balancing work and family life reaffirms the assertion that the search for balance is a key culprit in perceptions of ‘failure’ and guilt that all too often accompany women to and from work every day.

¹⁴The sub-Saharan (SSA) region, where there are 73 females enrolled for every 100 males, is the only exception.

¹⁵Data extracted from South African Higher Education Management Information Systems (HEMIS) obtained from the national Department of Higher Education.

¹⁶For an evocative exploration into this difference, see Smith and Nkomo (2021).

In their study of women in postgraduate study in STEM subjects in 17 African countries, Fisher et al. (2020) showed that getting married during the Ph.D. reduced women's publication productivity but increased that of men, and that becoming a parent during the Ph.D. was a key reason that women took longer to complete the Ph.D. Similarly, Zulu asserts that "the gender productivity gap still remains because of the time woman would have lost in their early years of childbearing. Male academics' productivity is largely not affected by child-rearing and related responsibilities" (2013, p. 758). A recent study in Rwanda exposed the challenges of pregnancy, family and motherhood for women academics and students. For some, pregnancy signaled an end to future education, while others chose to hide their pregnancies or the fact that they were mothers (<https://www.universityworldnews.com/post.php?story=20210627165955830>).¹⁷

The last decade has seen an increasing focus on the competing demands of motherhood in the Global North. Two persistent challenges are pervasive in North American publications on the theme of motherhood and work. The first is the challenge of child-care (Bassett, 2005; Connelly and Ghodsee, 2011; Moe, 2003; Ward and Wolf-Wendel, 2012). The Organization for Economic Cooperation and Development (OECD) statistics indicate that the United States is the only high-income country in the world that does not mandate paid maternity leave,¹⁸ nor does the federal government support any form of public childcare options beyond limited tax credits (Miller, 2021). Miller (2021) points out that in Denmark, a two-year-old is guaranteed a place in a daycare center and parents will pay a maximum of 25% of the cost, whereas in the US parents have no guarantees and will pay full cost, estimated at around \$1100 per month. Connelly and Ghodsee (2011) observe that many women academics find themselves making a choice between opting out, childlessness or delaying childbirth to the point of infertility: "Although women have been making steady advances in the academy in the last three decades, scholarly life in the United States is still a deeply masculinist culture; there is a continued expectation that professors are not the primary caregivers to their children" (2011, p. 5). *Mama Ph.D.* (Evans and Grant, 2009) is replete with stories of women who opted out of academia, largely as a result of the impossibility of managing work and family responsibilities.

The second, not unrelated challenge, is the grueling path to tenure. Ward and Wolf-Wendel draw attention to the tenure clock, established as the first seven years of academic life, which is "tuned to men who are presumed, ideally, to have no outside interests or responsibilities such as family, and is often in direct conflict with female clocks" (2012, p. 41). The clash between the tenure clock and the biological clock of childbirth is one that Connelly and Ghodsee (2011) also address as they interrogate decisions to delay or opt out of childbirth.

What is evident in this section is that family responsibilities, childbirth, childcare, domestic responsibilities alongside work responsibilities is a global challenge for women academics. While national, local, cultural and traditional contexts have an impact on how this challenge plays out in each woman's life and career, the primary battle between two "greedy" institutions, the family and the university (Morley, 2013, p. 122), is common.

The daily bread of being black in academia

For women of color at universities the challenges are exponentially greater as they battle entrenched and insidious forms of racism alongside entrenched and insidious forms of sexism (Madileng, 2014; Nkambule, 2014). The SAJHE published a special edition on black women academics in one university in South Africa (2014, 28 (6)). In reading these stories alongside those of other black women in higher education, I focus on two commonalities that may be observed. The first is that the pathway to becoming women academics takes many twists and turns, often as teachers in schools, then doctoral students, perhaps back to being school teachers, taking time off for motherhood and then perhaps returning to study or work. In the SAJHE special edition, none had actually planned to become academics. In exploring the experiences of women administrators in sub-Saharan Africa Johnson observes "[e]ach woman, except for one (a nurse) began her career as a teacher" (2014, p. 839). In Pillay et al. study (2017) which spans India, Brazil and South Africa, the constant refrain of having come to academia accidentally is voluble. Similar stories are evident in a study conducted by the American Council on Education (2018), which focused on women of color who are university presidents. The autoethnographic stories told in *Tedious Journeys: Autoethnography by Women of Color in Academe* (2010) recount similar tales of becoming accidental academics after taking convoluted side streets that skirted the borders of academia. The research on black women academics is clear: planning a career is not a luxury available to many black women, especially, but not only, those from the Global South.

A second commonality in the research on black women in higher education is the overt and covert forms of racism and sexism they encounter in their everyday interactions at universities. In the SAJHE special edition, the stories recounted stark and deeply damaging incidents of racism which often took the form of undermining and undervaluing language and research skills, or suggestions of a poor work ethic or even questioning of their intellectual ability, often euphemistically referred to as "not being academically ready". Madileng (2014) talks of her head of department who advised her to pursue studies in Setswana and not English even though she held a degree in Applied Linguistics (English). The fact that her head of department was male constituted a double-barreled attack. Nkambule (2014) recounts how a white female tutor asked her "are you sure you want to do psychology, considering your background?" (p. 2005), and another tutor who asked, "English is a difficult course, are you sure you want to major in it,

¹⁷See the work conducted by the Forum for Women Educationalists (FAWE) a pan-African women's organization with the mission, inter alia, to support girls and women to get an education.

¹⁸See: https://www.oecd.org/els/soc/PF2_1_Parental_leave_systems.pdf.

you can major in Xhosa?" (p. 2005). At postgraduate level she was told "students from disadvantaged backgrounds usually have problems with research because of language" (p. 2006). Botha (2014) says she was told after joining an all-white, mainly female department at a university that "[n]o black person stayed for more than three months" (p. 98). The expectation that she would fail and not cope with the academic rigor that was required, was palpable. The fact that racial denigration of black women academics is widespread is indisputable. Young and Anderson (2021), both women academics at one university in the United States, one white and not in a position of leadership, the other black and in a position of leadership, use the term "microaggression" to describe "brief and commonplace daily verbal, behavioral, or environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative racial slights and insults toward people of color" (p. 80). (see also American Council on Education, 2018; Robinson and Clardy, 2010).

Alongside the research on black women academics is the research on the experiences of black women Ph.D. students. Loots et al. (2016) show that although the number of black women registering for PhDs at South African universities is increasing, their rate of participation relative to population size is small.¹⁹ The global literature on black women in Ph.D. study offers a core, common acknowledgment: that experiences of racism, alongside the deeply entrenched expectation that black women are nurturers, makes success in doctoral study for black women extraordinarily challenging (see Connelly and Ghodsee, 2011; Evans and Grant 2009; Magano, 2011; Prozesky, 2008; Wolfinger et al., 2008).

The research by black women academics and doctoral students about their experiences in academia is small but growing. Whether they name themselves as intersectional feminists or not, their narratives of racial othering and exclusions add credible weight to the political project of intersectional feminism to interrogate race, class and gender as mutually constitutive sites of marginalization.

Finally, a common thread among women academics, arguably one that transcends race, is the determination to make deliberate choices to use research approaches and methodologies that speak to the ways in which women want to be heard and known, and to use writing strategies that sit comfortably with their sense of self. In 2005, in a reflexive article that used as a point of departure the use of the first person in my Ph.D. thesis, I made the point that the politics of resistance and transformation meant that I did not have the option to find refuge in the anonymity embedded in the term "the researcher" but rather had to confront and center the self if I was to produce trustworthy research (Pillay, 2005; see also Lather, 1996). Today, as I read the stories of women's experiences in higher education, I find strong echoes of similar stances. In the preface to *COVID Diaries: Women's Experience of the Pandemic*, Gouws and Ezeobi (2021) record their indignation at a male colleague who was conducting a survey of women's experiences during Covid-19 and asked "how would *statistics*²⁰ express what we feel?", an indignation also cast upon the absurdity of "men interpreting our experiences for us" (2021, p. vi). A few paragraphs later they make the critical observation that "'giving voice' is a cardinal feminist ideal" (2021, p. vi). I contend that today women academics have authoritatively taken hold of the ideal of giving voice. Research about women, by women and for women bears testimony to this ideal, the foundations of which were laid by the groundbreaking writings of, inter alia, Ribbens and Edwards (1998), Lather (1991, 2007), and Robinson and Clardy (2010).

The curveball of Covid-19

In the context of Covid-19, the legacy of guilt experienced by academic mothers, linked largely to the competing demands of family and work and their absence from home, was violently thrown off course. Guilt associated with not spending enough time with their children and family went out the window. Women academics found themselves at home, with their children, spouse/partner and other family members. The separation of home and work no longer existed²¹; in the absence of childcare and schooling, the increasing burden of household chores and home schooling fell largely on women. Claassens confesses, "I love my daughter ... But I really did not want to be her schoolteacher" (2020, p. 172). In an instant, women became caregivers, cleaners, cooks, and professionals all at the same time and in the same space.

Women also lost the battle for claiming physical workspaces in the home, as men asserted stronger claims over such spaces, often by virtue of their higher incomes, or perceptions that a man's job was more important than a woman's, or just by virtue of being male. Ezeobi says "there were days I sat in my car to grade essays, because I could not do it in my home" (Gouws and Ezeobi, 2021, p. 157). Ironically, while the house and home arena has traditionally been perceived as the domain of women, the onset of Covid-19 saw men usurping this position and staking an assertive claim to a space that historically they have been at pains to distance themselves from. But this prerogative extended only so far as the right to priority over the physical spaces in the house, and not to taking care of the children or other domestic responsibilities (Chitsamatanga, 2021; Stanford University, 2021).²²

Not surprisingly, the impact of Covid-19 on women academics was far greater than that on men. With respect to research output, numerous sources point to the reduced output from women academics, while that of men remained steadfast and, in many

¹⁹See HEMIS data provided on page 22.

²⁰Italics in original.

²¹At the end of 2020, I was part of a team that interviewed a number of woman students who had applied for a scholarship. One of the questions asked was how she had coped during lockdown. One woman's response was striking. She said that she had always thought that working from home was her ideal. She would be able to take care of her children, work and not feel guilty. She realized soon after lockdown began that leaving home and going to work was important for her own well-being and that she rarely paid attention to what would be good for *her*.

²²Stanford University published a series of articles on the impact of Covid-19 on women academics. <https://facultydevelopment.stanford.edu/diversity-engagement/faculty-networks/faculty-womens-forum/gendered-covid-19-faculty-experiences>.

instances, increased. In a study conducted by Squazzoni et al. (2020), which reviewed all articles and manuscripts submitted to all Elsevier journals between May 2018–20, they found that women submitted proportionally fewer articles than men during the months of the lockdown. Between February and May 2019, women submitted 877,978 and men 1,927,984 articles; in 2020, women submitted 144,978 articles and men 2,519,734. Accordingly, in 2019 men submitted 1,050,006 more articles than women and in 2020, 2,374,756 more. Put differently, the difference between men's and women's submissions pre- and post-pandemic more than doubled. The data collected also showed that the deficit between women's and men's submissions was more pronounced in cohorts of younger women. This finding perhaps signaled that younger women, who were more likely to have young children, bore the brunt of reduced research outputs.²³ In a study conducted in Poland, also part of a geographic region, namely Eastern Europe, that is excluded from "global" literature on women and higher education,²⁴ one male academic whose wife was also an academic said, "my wife's work was compromised ... whereas mine was not affected at all. It even got better, and I could work longer. So my work was definitely prioritized over my wife's" (Gorska et al., 2021, p.11). In the same study, a woman academic without children said that since Covid-19, her research output was "booming", while academics who were mothers not only bore the brunt of responsibility for caring for their own children at home, but also took on the care of their students' well-being. The study reaffirmed the research cited in Gorska et al. (2021) that in times of crisis women take on additional household duties and that men's paid work is prioritized.

The publication edited by Gouws and Ezeobi (2021) is unmistakably a cathartic journey for the women whose voices people the pages of the book. The effects of Covid-19 on cultural, family and individual practices and contexts recounted here signal untold intimate insights that have been rarely heard. The proud cry of "*wathint' abafazi, wathint' imbokodo!*"²⁵ famously ingrained in the nation's history when the women of South Africa marched to the seat of government to protest the imposition of passes²⁶ on African women, took on new meaning for Kunene (2021). She observes "[t]his is something that many black women in South Africa have to do: be rocks and keep it moving. We have to be ... unbreakable rocks that can take incredible amounts of trauma". Kunene realizes that the burden of always being hard and strong, of "packing away her anxieties, trauma ... and keep it moving" was a burden she did not want to bear, and she gave herself permission to do what was good for her, in this instance to be alone. While Kunene chose to be alone, Mpanza observed that social distancing, key to survival during Covid-19, was anathema to her community where social connections were the foundation to survival in apartheid South Africa. She says that "isolation had produced loneliness even in a home full of people" (Mpanza, 2021, p. 80). The stories in this anthology suggest that performances of seclusion, perhaps in the form of avoiding mental health vulnerabilities, or pretending to always be a rock, no longer worked. The women found themselves confronting themselves in ways that they may have previously assiduously avoided.

While the long-term impact of Covid-19 on the careers of women academics is yet to be fully evaluated, there is no doubt that their careers will reflect the burden of this pandemic far more than those of men.

Winning ways

As indicated in the introduction of this chapter, there is extensive literature on the obstacles and challenges facing women in academia, but there is also a small but growing body of literature that speaks to the way in which women academics have succeeded in building their professional careers. I offer two brief examples here.

The first example is that Mudaly's (2015) recounting of her academic journey. She embarks on an intense and exacting process of self-reflection to reconstruct and reshape her professional identity from that of a high school teacher to a researcher at a leading university in South Africa. Like many other black women academics, Mudaly found herself overburdened by a massive teaching load which was exhausting but which played little if any role in transitioning her from a junior academic to a more senior position.²⁷ In other words, teaching, which held little currency for promotion, was her main responsibility, leaving little room for research. In self-critiquing her stance of complicity embedded in her teaching strategies, she began to craft other ways of being and to, inter alia, take a transformative approach to her teaching and to thinking about her career path. She says she consciously discarded pedagogy that was rooted in submission and maintenance of the status quo, a stance that science teachers were trained in, and adopted a transformative and emancipatory approach to her teaching. The confidence she gained as she explored these approaches encouraged her to request a formal meeting with the Dean of Research and to state her case for why she would be a good researcher. She was subsequently invited to serve on a panel that would review research proposals. Her career as a researcher began to take shape. Five years after the publication of the article in which she reflected on her academic journey (Mudaly, 2015), with the encouragement of supportive colleagues, Mudaly was promoted to associate professor.²⁸ In the conclusion to the article she makes two key observations: external recognition of her value was important but so was her own recognition of and confidence in

²³See also: <https://www.theguardian.com/education/2020/may/12/womens-research-plummets-during-lockdown-but-articles-from-men-increase>.

²⁴The authors make the point that literature about women in academia is usually about the United States and Western Europe.

²⁵When you strike a woman, you strike a rock.

²⁶A pass was a document that had to be signed by an employer or other person in authority (usually white) granting permission for an African person to travel from one place to another inside the country. If caught without a pass, African people were jailed. Until 1956, the law applied only to men. The women's protest march to government succeeded in stopping the imposition of the pass law for women.

²⁷See also Subbaye and Vithal (2017) for other examples of women academics similarly overburdened.

²⁸Communication with Mudaly.

her capabilities. In denying practices of seclusion and accepting her aspirations she gave herself permission to advance her academic career.

Using a capabilities framework,²⁹ Loots et al. (2016) identify a number of factors that support black women's³⁰ success in higher education. Key among these is "the importance of acting as an active agent, taking control, standing up for oneself and having a voice" (p. 114). Loots et al. (2016) show that the women in the study began to identify personal capabilities that inspired them to move ahead, to build resilience against institutionalized racism and find value in themselves. They had also built their own communities of support in the face of entrenched institutional obstacles. In commenting on her recent appointment as the first female black Dean of Engineering at the Tshwane University of Technology, Kanakana-Kutamba, born in the Limpopo province where electricity and connectivity are rare, identifies the capability of innovation as a key ingredient of success. She had not planned on academia as a career path but found herself switching from the private sector to a university when she discovered that she enjoyed contributing to the development of young people. In other words, she chose to embrace the activities that gave her joy and to translate this into a career path.³¹

Conclusion

In conclusion, I have attempted in this review to provide a global overview of women academics, academic mothers and students in higher education. As an intersectional feminist I made the decision to foreground literature from the Global South without excluding literature from the Global North. The fact that exclusions from masculine, white, Global North epistemologies and ontologies endemic to higher education institutions are repeatedly performed by women (and men), who continue to perpetuate the normativeness of their places of privilege by the simple, yet powerful acts of exclusion, is disappointing. This review is a small disruption of this practice. I have also made a conscious decision to offer a louder voice to black women, a rare occurrence in the literature that claims to be global.

As a start, I have also introduced the notion of seclusion to, inter alia, understand the ways in which academic women, whether from the Global North or South, adopt practices that keep gendered performativity alive by, inter alia, secluding from ourselves, and possibly from communities and contexts in which we live and work, our own aspirations, capabilities and strengths. Confronting acts of seclusion, evidenced in the frequent silence of universities to sexual violence against women in their institutional spaces, may also be a useful tool for universities to engage with their complicity in repression and oppression.

References

- Acker, S., 1980. Women, the other academics. *Wom. Stud. Int. Forum* 1 (1), 81–91.
- Acker, S., 1994. *Gendered Education: Reflections on Women, Teaching and Feminisms*. Open University Press, Buckingham, Philadelphia.
- American Council on Education, 2018. *Voice from the Field: Women of Color Presidents in Higher Education*. American Council on Higher Education TIAA Institute, Washington, DC.
- Bassett, R. (Ed.), 2005. *Parenting and Professing: Balancing Family Work with an Academic Career*. Vanderbilt University Press, Nashville.
- Bondestam, F., Lundqvist, M., 2020. Sexual harassment in higher education—a systematic review. *Eur. J. High Educ.* 10 (4), 397–419.
- Botha, E.M., 2014. What were they all about? Two questions that provoked different reactions and feelings. *S. Afr. J. High Educ.* 28 (6), 1972–1985.
- Butler, J., 1990. *Gender Trouble*. Routledge Classics, New York and London.
- Cama, M.G., Jorge, M.L., 2016. Gender differences between faculty members in higher education: a literature review of selected higher education journals. *Educ. Res. Rev.* 18, 58–69.
- Chitsamatanga, B.B., 2021. Coronavirus (COVID-19) and women in higher education in a South African university: academic and social implications. *Afr. J. Gend. Soc. Dev.* 10 (1), 7–34.
- Claassens, L.J., 2020. Very tired superwoman. In: *Covid Diaries: Women's Experience of the Pandemic*. Imbali Publishers, Cape Town, pp. 172–178.
- Collins, P., 2004. Learning from the outsider within: the sociological significance of black feminist thought. In: Harding, S. (Ed.), *The Feminist Standpoint Theory Reader*. Routledge, New York, pp. 103–126.
- Connelly, R., Ghodsee, K., 2011. *Professor Mommy*. Rowman & Littlefield, United Kingdom.
- Dados, N., Connell, R., 2012. The global south. *Contexts* 11 (1), 12–13.
- Davis, K., 2008. Intersectionality as a buzzword: a sociology of science perspective on what makes a feminist theory successful. *Fem. Theor.* 9 (67), 67–85.
- Elsevier, 2020. *The Researcher Journey through a Gendered Lens*. Elsevier Publishing, USA.
- Evans, E., Grant, C., 2009. *Mama PhD: Women Write about Motherhood and Academic Life*. Rutgers University Press, New Jersey.
- Fisher, M., Osiru, M., Nyabaro, V., 2020. Making it to the PhD: Gender and Student Performance in Sub-Saharan Africa. *PLoS One*. <https://journals.plos.org/plosone/article/metrics?id=10.1371/journal.pone.0241915>. (Accessed 15 July 2021).
- Gaspar, D., 2007. What is the Capability Approach? Its core, rationale, partners and dangers. *J. Soc. Econ.* 36, 335–359.
- Gorska, A.M., Kulicka, K., Stanisewska, Z., Dobija, D., 2021. Deepening inequalities: what did COVID-19 reveal about the gendered nature of academic work? *Gend. Work Organ* 28 (4), 1546–1561.
- Gouws, A., 2018. #EndRapeCulture campaign in South Africa: resisting sexual violence through protest and the politics of experience. *SA J of Political Studies* 45 (1), 3–15.

²⁹See Gasper 2007 for a discussion on a capabilities approach (CA).

³⁰The study specifically focused on seven black women doctoral studies. Loots et al. make the point that white women still hold considerable advantage over black women at universities in South Africa. See the data provided earlier in this chapter, showing that, although a miniscule part of the female population in South Africa, there are more white women registered for PhDs (4% of all white female students) than African women, and that only 3% of African women in higher education are registered for PhDs even though they comprise more than 40% of the female population. See also Johnson-Bailey's (2004) research on black women at universities in the United States.

³¹University World News Africa Edition. May 2021. <https://www.universityworldnews.com/post.php?story=20210504042258474>.

- Gouws, A., Ezeobi, O. (Eds.), 2021. *Covid Diaries: Women's Experience of the Pandemic*. Imbali Publishers, Cape Town.
- Harding, S. (Ed.), 2004. *The Feminist Standpoint Theory Reader*. Routledge, New York.
- Hudock, A., 2009. First day of school. In: Evans, E., Grant, C. (Eds.), *Mama PhD: Women Write about Motherhood and Academic Life*. Rutgers University Press, New Jersey, pp. 63–65.
- Hunter, L.A., Leahey, E., 2010. Parenting and research productivity: new evidence and methods. *Soc. Stud. Sci.* 40 (3), 431–451.
- Jay, M.L., Packer-Williams, C.L., Jackson, T.O., 2010. Are we change agents or pawns? Reflecting on the experiences of three African American junior faculty. In: Robinson, C., Clardy, P. (Eds.), *Tedious Journeys: Autoethnography by Women of Color in Academe*. Peter Lang, New York, pp. 97–118.
- Johnson-Bailey, J., 2004. Hitting and climbing the proverbial wall: participation and retention issues for black graduate women. *Race Ethn. Educ.* 7 (4), 331–349.
- Krebsbach, J.M., 2022. The impact of academic publishing: inequality for women in behavior analytic journals. *Frontiers in Sociology*, 7, pp. 1–10.
- Kunene, A., 2021. I followed the fire. In: Gouws, A., Ezeobi, O. (Eds.), *Covid Diaries: Women's Experience of the Pandemic*. Imbali Publishers, Cape Town, pp. 93–98.
- Lather, P., 1991. *Getting Smart: Feminist Research and Pedagogy With/in the Postmodern*. Routledge, London.
- Lather, P., 1996. Issues of validity in openly ideological research; between a rock and a soft place. *Interchange* 17 (4), 63–84.
- Lather, P., 2007. *Getting Lost: Feminist Efforts Toward a Double(d) Science*. State University of New York Press, New York.
- Loots, S., Ts'ephe, L., Walker, M., 2016. Evaluating black women's participation, development and success in doctoral studies: a capabilities perspective. *SA J. High. Educ.* 30 (2), 110–128.
- Madleng, M.M., 2014. Critical reflections: experiencing discrimination, disrespect and disregard – forming a professional identity. *SA J. High. Educ.* 28 (6), 2027–2040.
- Magano, M., 2011. The social and emotional challenges of female postgraduate students in South Africa. *J. Soc. Sci.* 29 (3), 205–212.
- Mail and Guardian. Retrieved from <https://mg.co.za/education/2020-08-20-few-women-leaders-in-academia/>.
- Miller, C.C., 2021. How Other Nations Pay for Childcare: The US Is an Outlier. *The New York Times*. Retrieved from: <https://www.nytimes.com/2021/10/06/upshot/child-care-biden.html>.
- Moe, K.S. (Ed.), 2003. *Women, Family, and Work: Writings on the Economics of Gender*. Blackwell, United Kingdom.
- Morely, L., 2014. Lost leaders: women in the global academy. *High Educ. Res. Dev.* 33 (1), 114–128.
- Morely, L., Crossouard, B., 2015. *Women in Higher Education Leadership in South Asia: Rejection, Refusal, Reluctance, Revisioning*. University of Sussex Centre for Higher Education and Equality Research.
- Morley, L., 2013. The rules of the game: women and the leaderist turn in higher education. *Gend. Educ.* 25 (1), 116–131.
- Motherhood Initiative for Research and Community Involvement (MIRCI), 2003. *Mothering in the academy*. *J. Assoc. Res. Mothering* (5), 2.
- Mpanza, J., 2021. To be aged, black and terrified: the times of Covid-19. In: Gouws, A., Ezeobi, O. (Eds.), *Covid Diaries: Women's Experience of the Pandemic*. Imbali Publishers, Cape Town, pp. 75–83.
- Mudaly, R., 2015. Creating my academic self and space: autoethnographic reflections on transcending barriers in higher education. *J. Educ.* 62, 35–57.
- Nkambule, T., 2014. Against all odds: the role of 'community cultural wealth' in overcoming challenges as a black African woman. *SA J. High. Educ.* 28 (6), 1999–2012.
- Obers, N., 2014. Career success for women academics in higher education: choices and challenges. *SA J. High. Educ.* 28 (6), 1107–1122.
- Orth, Z., van Wyk, B., Andipatin, M., 2020. "what does the university have to do with it?" Perceptions of rape culture on campus and the role of university authorities. *SA J. High. Educ.* 34 (2), 191–209.
- Pillay, V., 2005. Narrative style; the inseparability of self, style and text. *Reflective Pract.* 6 (4), 539–549.
- Pillay, V., 2007. *Academic Mothers*. Trentham Books, Stoke on Trent, UK, Sterling, USA. UNISA Press, Pretoria, South Africa.
- Pillay, V., Mitra vom Berg, N., Bhana, D., Lúcia Guimarães de Mattos, C., Almeida de Castro, P., 2017. *Academic Mothers in the Developing World: Stories from India, Brazil and South Africa*. Africa World Press and Red Sea Press, Trenton, NJ.
- Prozesky, H., 2008. A career-history analysis of gender differences in publication productivity among South African academics. *Sci. Stud.* 21 (2), 47–67.
- Redmond, P., Gutke, H., Galligan, L., Howard, A., Newman, T., 2017. Becoming a female leader in higher education: investigations from a regional university. *Gend. Educ.* 29 (3), 332–351.
- Ribbens, J., Edwards, E. (Eds.), 1998. *Feminist Dilemmas in Qualitative Research*. Sage Publications, London.
- Robinson, C., Clardy, P. (Eds.), 2010. *Tedious Journeys: Autoethnography by Women of Color in Academe*. Peter Lang, New York.
- She Figures, 2021 from: <https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/67d5a207-4da1-11ec-91ac-01aa75ed71a1>. (Accessed 27 January 2022).
- Simeone, A., 1987. *Academic Women: Working Toward Equality*. Bergin and Garvey, South Hadley, MA.
- Smith, E.B., Nkomo, S.M., 2021. *Our Separate Ways: Black and White Women and the Struggle for Professional Identity*. Harvard Business Review Press, Massachusetts.
- Smith Silva, D. (Ed.), 2011. *Latina/Chicana Mothering*. Demeter Press, Bradford, Canada.
- Squazzoni, F., Bravo, G., Grimaldo, F., Garcia-Costa, D., Farjam, M., Mehmani, B., 2020. Only second-class tickets for women in the COVID-19 race: a study on manuscript submissions and reviews in 2329 Elsevier Journals. *PLoS One* 16 (10), e0257919. <https://doi.org/10.1371/journal.pone.0257919> 10.2139/ssrn.3712813.
- Stanford University, 2021. Gendered Covid -19 Faculty Experiences. <https://facultydevelopment.stanford.edu/diversity-engagement/faculty-networks/faculty-womens-forum/gendered-covid-19-faculty-experiences>. (Accessed 1 February 2022).
- Subbaye, R., Vithal, R., 2017. Gender, education and academic promotions in higher education. *Gend. Educ.* 7, 926–951.
- Toffoletti, K., Starr, K., 2016. Women academics and work–life balance: gendered discourses of work and care. *Gend. Work. Organ.* 23 (5), 489–504.
- UNESCO, 2021. Women in Higher Education: Has the Female Advantage Put an End to Gender Inequalities? <https://unesdoc.unesco.org/ark:/48223/pf0000377182>. (Accessed 15 July 2021).
- University World News Africa Edition, May 2021. <https://www.universityworldnews.com/post.php?story=20210504042258474>.
- VitalStats, 2016. *Publication Higher Education 2016*. Council on Higher Education, South Africa. Pretoria.
- Vongalis-Macrow, A. (Ed.), 2014. *Career Moves: Mentoring for Women Advancing Their Career and Leadership in Academia*. Sense Publishers, Rotterdam.
- Ward, K., Wolf-Wendel, L., 2012. *Academic Motherhood: How Faculty Manage Work and Family*. Rutgers University Press, New Jersey.
- Wolfiger, N.H., Mason, M.A., Goulden, M., 2008. Problems in the pipeline: gender, marriage, and fertility in the ivory tower. *J. High. Educ.* 79 (4), 388–405.
- Wood, L., Hoefler, S., Kammer-Kerwick, M., Parra-Cardona, J., Busch-Armendariz, N., 2018. Sexual harassment at institutions of higher education: prevalence, risk and extent. *J. Interpers. Violence*, 1–25.
- Young, K., Anderson, M., 2021. Hierarchical microaggressive intersectionalities; small stories of women of color in higher education. *Metropolitan Universities* 32 (1), 78–103.
- Young, S., Wiley, K.K., 2021. Erased: why faculty sexual misconduct is prevalent and how we could prevent it. *J. Publ. Aff. Educ.* 27 (3), 276–300.
- Zulu, C., 2011. Women leaders' construction of leadership and management of the academic department. *SA J. High. Educ.* 25 (4), 838–852.
- Zulu, C., 2013. Women academics research productivity at one university campus: an analysis of dominant discourses. *SA J. High. Educ.* 27 (3), 750–767.

The first-in-family experience: gendered and classed subjectivities in elite higher education

Garth Stahl^a and Sarah McDonald^b, ^a School of Education, University of Queensland, Brisbane, QLD, Australia; and ^b Education Futures, University of South Australia, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	392
Classed and gendered subjectivities during a transitional time	392
Changes to subjectivities in higher education	393
Kate	393
Vuong	394
Discussion	395
Conclusion	396
Funding	396
References	396

Introduction

While it may differ country to country, we are witness to striking class and gender inequalities across all sectors of education. Our longitudinal research, the *First-in-Family Project*, documented the experiences and identity work of first-in-family students in Australia, a country which has seen recent policy directives aiming to widen participation (Bradley et al., 2008). Despite such efforts, students from non-traditional backgrounds remain largely underrepresented. In the study of equity and social justice in higher education, students who are the first in their family to attend university remain a “critical group for understanding social inequality in higher education and processes of social mobility” (Beattie, 2018, p. 172). As a group, they can often serve as a barometer for the nation’s equity agenda. Social theorists interested in class disadvantage continue to document how first-in-family students access and experience higher education. However, what remains largely unexplored is the gendering of such experiences and how these experiences are informed by institutional cultures and wider societal discourses (see Stahl and McDonald, 2021). As first-in-family students transition to university, we are interested in how gender and class intersect and influence their experience, contributing to the maintenance of their aspirations.

For those who are first-in-family/first-generation, the underlying logic is that the journey through their education will lead to increased opportunities and, thus, allow them to become more socially mobile. However, much of the research concerns the ways in which first-in-family students struggle during the transition into university in comparison to their middle and upper-class counterparts (Pike and Kuh, 2005; O’Shea et al., 2018). Without adequate preparation and support, many first-in-family students may struggle in higher education (Stahl, 2021) and research indicates a high probability of dropping out (Harvey et al., 2017). Recently we have seen a shift in research on first-in-family students which problematizes the first-in-family experience (see Lehmann, 2013) where there are significant variations that need to be acknowledged or what Patfield et al. (2020) call “degrees of being first.” Additionally, research has highlighted how overcoming barriers and traversing higher education as a first-in-family student can increase “grit” and determination (Hodge et al., 2018). In an effort to capture the first-in-family experience, this article focuses on a rather undeveloped aspect of research on first-in-family students, that of gender identity, and how experiences with gender and certain expectations around gender contribute to first-in-family students feeling like outsiders.

Classed and gendered subjectivities during a transitional time

Education systems play a significant role in how students come to understand themselves as gendered (Weiner and Arnot, 1987; Mac an Ghaill, 1994; Francis, 2000) and classed individuals (Reay, 2002; Archer et al., 2010; Stahl, 2014). Studies focusing on class inequality within education draw heavily on sociological theories of social and cultural capital, particularly research exploring how students negotiate pathways and choices within higher education (Kurlaender and Hibel, 2018). Limits in these forms of capital can heavily impact how students from low-socioeconomic backgrounds experience university, which can—and often does—have a significant influence on their sense of identity. Research on how class disadvantage is experienced within higher education remains prominent (Reay, 1998; Spiegler and Bednarek, 2013), yet within the past decade, there has arguably been a decline in similar research into gendered subjectivities in higher education. We do not dispute the ways in which social class structures opportunities and contributes to what is possible (see Stahl, 2020); however, for our purposes, we are interested in how the interplay between class and gender structures the first-in-family experience. Understanding how young first-in-family men and women transition from secondary school primarily requires documenting their shifting, discursive, and intersecting masculine and feminine

subjectivities. We are interested in how such transitions are gendered and classed as well as how they contribute to the formation of aspirations. We now outline recent gender theorizations which have informed how social theorists have come to understand masculinities and femininities in relation to the student experience.

We understand categories of male and female to be binary while gender categories of masculine and feminine are diverse, multiple and overlapping and formed through multiple spaces which are discursive, material, and relational (Harwood et al., 2018). Drawing on post-structural thinking, the subject works toward a particular gendered identity which is continually positioned against, and positioned by, understandings of traditional masculine traits (strong, brave, a leader, stoic) and feminine traits (nurturing, emotionally open, mild-mannered). Bearing this in mind, our research seeks to explore how gendered norms as “socially and culturally produced patterns of language, known as discourses, construct people and the power relationships among them in particular ways” (Francis, 2011, p. 78). These patterns occur both within wider social structures and at a local level. In considering the discursive and relational aspects of gender, Paechter (2002, 2003) contends that we perform gender based on shared understandings of how gendered performances are positioned within different social groupings, referred to as “communities of practice.” According to Paechter (2002), gender is performed through participation in these communities, where it becomes validated. Adding a layer of complexity, Paechter (2002) notes that most people are “simultaneously members of a number of communities of masculine and feminine practice in the different contexts of their lives” (p. 72). During the transition to higher education, many of these localized communities are naturally subject to significant changes where both young men and women—especially those from disadvantaged backgrounds—engage with a wider array of masculinities and femininities.

Adding another dimension, gender scholarship continues to foreground how young people negotiate “the tensions between continuities and social changes” (Arnot and Mac an Ghaill, 2006, p. 3), contributing to expanded understandings of, and engagement in, performances of gender (Davies, 1993). It is through movement within new spaces that young people encounter diverse gender discourses, which may be experienced as a liminal time of self-evaluation as they consider their own gendered performances in relation to the communities they come from (Stahl and McDonald, 2021). Gender, therefore, becomes a site of reaffirmations as well as rebuttals where the conventions of femininity and masculinity are reinscribed in new ways, contributing to new expectations (see Adkins, 2002; Kenway and Kelly, 2000). To understand how gender is experienced in relation to social change, Risman (2004) posits three dimensions of gender: individual (gendered selves), interactional (gendered expectations during daily interactions), and institutional proposed (organizational practices and regulations).

In all aspects of the life course, identities are often constructed and reconstructed according to experiences within multiple structures, including familial, institutional, and other collectives. Transitions to higher education for all students are never linear and are often fragmented as they come to adapt to what is expected by the institution. Research on widening participation continues to document the motivations of non-traditional students and how motivations change over time, influenced by their experiences. Our research engages in a critical consideration of the contemporary university experience where, as Pötschulat et al. (2020) note, the so-called “student experience” is problematic as it has “implications for how we come to understand a complex set of practices” (p. 2). Their scholarship calls attention to how students engage in negotiations of contemporary studenthood where everyday practices and circumstances are evaluated and positioned according to socially accepted norms and expectations (Stahl and McDonald, 2021).

Drawing from our qualitative interviews with young working-class men and women participating in higher education, our interest is in how our participants’ gendered subjectivities shift over time as well as their shifting perceptions of self. To illustrate our thinking around gendered subjectivities in transition, we present two case studies of first-in-family people studying in elite spaces. Their topics of chosen study are both STEM which carries a connotation of an elite profession and, therefore, both participants could be considered an example of what Southgate et al. (2017) call “extreme social mobility.” This is followed by a discussion drawing on theories of gendered subjectivities, primarily Risman (2004) and Paechter (2003), to explore the journeys of Kate and Vuong as learners navigating communities of practice where they adopt, and rebuff, gendered performances aligned with the norms of the discursive space of the university.

Changes to subjectivities in higher education

Kate

Prior to entering university, working-class girls may experience difficult identity negotiations in order to become academically successful. Recent research on how girls construct learner identities in terms of their gendered subjectivities — such as the “successful girl” who is effortlessly successful both socially and academically — holds salience as an example of normative and idealized femininities (Allen, 2016; Bowers-Brown, 2018; Renold and Allan, 2006). Academically inclined girls from working-class backgrounds may, either consciously or unconsciously, engage in work on the self in order to enact performances aligned with notions of middle-class girlhood. These performances are often the result of extensive labor as girls negotiate a “can do” or successful girl subjectivity (see Harris, 2004; Pomerantz and Raby, 2017). Furthermore, researchers continue to highlight how educationally successful girls from working-class backgrounds—where success carries with it the expectation of social mobility—may come to disidentify with their classed backgrounds through displays of both conformity and resistance (Lucey et al., 2003; Walkerdine, 2003).

We first met Kate soon after she finished her final year of secondary school. Kate, who is Anglo-Australian, is an only child whose mother is a stay-at-home-mum while her father works in packaging. Kate was confident and chatty with aspirations for herself, and she was looking forward to her enrollment in an engineering degree at a prestigious university. In terms of her experiences with

feminine identities, Kate spoke about how most of her friends were male, “the majority of people I get along with are guys,” and this was largely due to a common interest in video gaming. Furthermore, Kate dressed in a masculine style, preferring men’s jeans and shirts because she found them more comfortable than women’s clothing; she often wore accessories associated with her hobby of gaming, such as a Pokemon t-shirt or a Doctor Who baseball cap. Kate was aware that her personal style positioned her outside of feminine norms, “I’d be the really odd one if you looked at a table of girls.” In many ways, it seemed engineering, a male dominated degree, would be a natural fit for Kate, given how she felt comfortable within masculine communities of practice and the male dominated nature of the STEM field (see [DeWitt and Archer, 2015](#)). In research on girls from non-dominant backgrounds, [Calabrese Barton et al. \(2013\)](#) draw attention to how the identities of young women are shaped by how others “recognize and respond” to their self-representation as well as their social and learning practices.

The transition to university can be a time of significant change largely in response to the exposure to new experiences and new people. Young women’s identities are constructed in relation to the possibilities, positioning and performances of femininities largely in relation to the academic experiences informing their aspirations and subjectivities ([Stahl et al., 2020](#); [McDonald, 2021](#)). University, as a performative space which carries the expectation of capital accrual, is a site for the (re)making of identities. When we met with Kate during her first semester of her chemical engineering degree, she spoke to us about her experiences with the Women in STEM community, describing an event occurring on International Women’s Day:

And they were like, what are your [female] role models in STEM, and I was like, I don’t have any. And they were like, “Oh, is that ‘cause women are underrepresented in STEM?” No, because I haven’t looked, and I don’t care.

Kate’s words suggested a frustration regarding discussions of gender inequality in STEM and she struggled to see how it related to her. However, as she engaged more in various overlapping communities of practice where she was exposed to ways of being not only a scientist *but a female scientist* ([Calabrese Barton et al., 2013](#)), Kate’s identity and her opinions began to shift. By the end of the year, Kate was beginning to speak about these issues in a different manner, telling us how one of her lecturers often assumed she and other girls in the class did not understand certain concepts, “he kept assuming we had done it wrong.” The following year, Kate had joined the Women in STEM group which she had initially dismissed and was able to speak at length about how she was less likely to get a job in engineering due to her gender.

Focusing on the intersection of gender and class and contributing to motivations and the maintenance of aspirations, Kate’s experience in an elite setting had also raised her awareness regarding how her class background placed her as an outsider. She recounted an anecdote where class intersected with gender, where Kate highlighted the need to dress in a particular way—dresses, high heels, and makeup—during networking events. Kate, who was used to wearing men’s jeans and shirts, conformed to expectations in terms of feminine clothing, demonstrating how she felt compelled to perform a particular gendered and classed identity within STEM in order to both experience acceptance and to position herself advantageously in the job market. Such an experience led Kate to reflect on where she fit within the university STEM culture. In dressing up, she noted, “I’m like, but why? All it does is show off some social class status that I’m not even a part of.” We are reminded here of how gender and class become sites of struggle where one needs to know the expectations and how, for first-in-family students, this is an additional stressor on top of academic work at university.

Vuong

Prior to entering university, working-class boys contend with various struggles in their education and many of these struggles are identity-based. There exist significant limits within the discursive space of schooling, and wider society, which hinder the ways in which studious working-class boys navigate their journey as learners. Research continues to suggest that boys—and most often working-class boys—struggle to be socially accepted as “cool” by their peer group while also maintaining their academic progress (see [Reay, 2002](#)). We assume when working-class young men move into higher education, the logics around gender alter and what was once pathologized may now be held in esteem. In terms of masculine performances, there may be more possibilities, though this depends on the course of study.

Boys who choose to study mathematics have long been linked to the construction of “a social image of being a ‘nerd’ or a non-popular boy within the peer group” ([Diez-Palomar and Mara, 2020](#), p. 55). In [Mendick’s \(2006\)](#) study of masculinities in mathematics, she calls attention to the “working-class class pretenders” (p. 75) noting the classing and gendering of different disciplines can be an uncomfortable fit for working-class boys. Such experiences of feeling out of place can potentially undermine their confidence and consequently their academic progress. There exist few examples where students can be the all-rounder ([Skelton and Francis, 2011](#)), or where boys can bring a “sporty” and mathematically inclined learner identity together (see [Pomeroy, 2015](#)). When we first encountered Vuong, a Vietnamese-Australian, he presented himself as a confident mathematics student with high aspirations. He complained often about the pressures he put himself under, though he was determined to be successful and proud to be one of the few in his secondary school to attend university. When asked explicitly about being perceived as a geek, he said:

To me I don’t really feel like a nerd or a geek because, although I study and revise a lot, I still go outside with friends, I still have that type of social life that the stereotypical geek and nerd might not have usually.

Vuong, who transitioned to an elite university in the city, enrolled in an advanced mathematics program only available to a select group of students. Therefore, his daily activities were located in an elite space within an elite space. While academically he was successful at university, socially Vuong did struggle to acclimate in what he described as a hyper-competitive environment where many of the boys he knew were involved in elite sport activities such as rowing and lacrosse (these are sports commonly associated with elite private schools in Australia). To give a sense of Vuong's class background, he had never left the Australian state in which he lives. In focusing on place and gender in terms of academics, Vuong described his mathematics course as composed of predominantly men and, more to the point, a very different type of male student than he had encountered before.

Vuong: But, yeah so there is a bit of a difference between just the average guy in my school, high school, towards the average guy in my degree.

Researcher: Yeah, it's just different types of people.

Vuong: Yeah. Actually, there was a case where one of the guys in my degree challenged my masculinity. It was very annoying.

Researcher: How so? How'd that come about?

Vuong: Yeah, he was like, "Do you even play any sports or anything?" And I just had this twinge of annoyance in me that I went, "Excuse me, but you remember that I told you that I play volleyball and everything." And then like, "Oh yeah, you do, don't you?" And then I was going ... He's just very testosterone driven ... Yeah. I put that guy in his place when he challenged my masculinity. Very annoying ... Yeah, because he was trying to make an example.

Vuong's experience works to subordinate his masculine identity and reinforce the gender hierarchy where, in this instance, sports participation functions as a proxy influencing what is deemed acceptable in terms of one's gender identity. Boundaries are being formed and what is considered normative is delineated. As [Paechter \(2003, p. 71\)](#) notes, how we come to understand masculinities and femininities takes place in relation to "a social configuration" where there are "shared practices in pursuit of the common goal of sustaining particular localized masculine and feminine identities." Furthermore, Vuong's words echo research which shows how boys are compelled to balance social acceptability while also maintaining their academic profile, an endeavor which is a continual struggle (see [Reay, 2002](#)).

Discussion

[Risman's \(2004\)](#) posits three dimensions of gender: individual, interactional (gendered expectations), and institutional proposed (organizational practices). We see through the case studies of Kate and Vuong how they experience changes to their gendered identities across these three dimensions and how these changes occur in relation to social class. While Kate is confident in her gender identity on an individual level, recognizing how she transgresses expectations in terms of interactional gendered norms, she is compelled to engage in performances of a feminine identity which is more closely aligned with the institutional expectations of the university. For Kate, these new performances represent a shift in her gendered identity as she comes to understand how the expectations and practices of gendered norms contribute to her long-term goals. Her experience further highlights how individuals are able to, and are oftentimes compelled to, engage in differing and overlapping performances of gender.

For Vuong, his acclimatization is defined by not having, or performing, the right sort of sporty masculinity. While Vuong does in fact participate in sports, he is stereotyped as the "Asian geek" ([Huynh and Woo, 2014](#)) who is un-sporty and, therefore, assumed not to be masculine enough. Through his interaction with the "testosterone-driven" student, Vuong's expectations about what it means to be a male are confirmed. Furthermore, these expectations are confirmed through university marketing materials, what Risman refers to as "organizational practices," where such materials foreground an individual who is sporty, confident, socially engaged and academically inclined, the so-called "all-rounder" ([Skelton and Francis, 2011](#)). We can assume this informs how a gendered sense of self becomes regulated. Furthermore, in contrast to Kate, Vuong is unable to engage in performances of a macho masculine identity expected within the mathematics course. It is unclear how resisting certain gendered performances, how speaking back to false assumptions, may contribute to how Vuong comes to experience university though we suggest it may contribute to him continuing to feel like an outsider.

[Paechter's \(2003\)](#) work on communities of practice highlights how both location and membership within communities bring with them particular ways of performing gendered identities. For both Kate and Vuong, there is a sense of a certain level of comfort within their gendered performances in secondary school. Vuong was confident in his masculine identity where he did not identify

with particular configurations such a geek while Kate was comfortable within her feminine identity where she transgressed gendered norms and expectations, but very much identified, and expected to be identified, as a girl. Yet as these two students transitioned into university, and into new and varying communities of practice, they were both challenged in terms of their gendered identities. As Kate began her engineering degree, she did not locate herself within her degree in terms of a gender binary, and in fact was skeptical of being drawn into the Women in STEM community, a localized community of practice which did not align with her gendered subjectivity. Yet, as she progressed through her degree, it is how she was defined and located within the degree by others, for example as a girl who did not understand the content, that impacted on and mediated her understandings of self as well as her understanding of the performances she must engage in to be successful. Similarly, Vuong's masculinity is positioned by other young men in terms of the logic of a localized community of practice as he recognized the different masculine performances which contrasted with those of his secondary school. In contrast to Kate, Vuong did not appear to engage in new performances of masculinity but rather attempted to position himself within the community of practice by highlighting his involvement in sport and deflecting challenges to his masculine identity. As Kate and Vuong transition from their working-class secondary schools, we see how the space of university is a site of (re)construction of gender, where certain forms of masculinity and femininity are privileged.

Conclusion

While research on gender in higher education has declined in recent decades, our study considers contemporary experiences within education as remaining highly gendered and, when coupled with class, reinforcing of social hierarchies (Stahl and McDonald, 2021). The experiences of Kate and Vuong have implications not only for how they come to be at university but also for how first-in-family students come to understand themselves as socially mobile individuals where belonging is integral. Their stories speak to the student experience as a "complex set of practices" (Pötschulat et al., 2020, p. 2) informed by institutional and societal discourses. Furthermore, the stories of Kate and Vuong highlight how students from non-traditional backgrounds may find the social acclimatization to elite spaces particularly confronting (Reay, 1998), adding an additional barrier to the first-in-family experience. Their experiences compel us to reflect on what is means to embody the ideal university student or, more specifically, what a failure to embody idealized versions of masculinity or femininity could mean. While Kate and Vuong feel comfortable recounting these stories, we are not privy to the gender biases that they were perhaps not conscious of or the various micro-aggressions that could have informed their everyday experiences. While universities in Australia have become inclusive and more students from non-traditional backgrounds attend than ever before, we cannot help but think how the journeys of Kate and Vuong would be different if they could have just been themselves and, more to the point, felt comfortable being themselves.

Funding

Australian Research Council (DE170100510).

References

- Adkins, L., 2002. Revisions: Gender and Sexuality in Late Modernity. Open University Press, Buckingham, UK.
- Allen, K., 2016. Top girls navigating austere times: interrogating youth transitions since the "crisis". *J. Youth Stud.* 19 (6), 805–820.
- Archer, L., Hollingworth, S., Mendick, H., 2010. *Urban Youth and Schooling*. Open University Press, Berkshire.
- Arnot, M., Mac an Ghaill, M., 2006. (Re)contextualizing gender studies in education: schooling in late modernity. In: Arnot, M., Mac an Ghaill, M. (Eds.), *The RoutledgeFalmer Reader in Gender and Education*. Routledge, London, pp. 1–14.
- Beattie, I.R., 2018. Sociological perspectives on first-generation college students. In: Schneider, B. (Ed.), *Handbook of the Sociology of Education in the 21st Century*. Springer, Cham, Switzerland, pp. 171–191.
- Bowers-Brown, T., 2018. It was noticeable so I changed: supergirls, aspirations and Bourdieu. In: Stahl, G., Wallace, D., Burke, C., Threadgold, S. (Eds.), *International Perspectives on Theorizing Aspirations: Applying Bourdieu's Tools*. Bloomsbury, London, pp. 145–160.
- Bradley, D., Noonan, P., Nugent, H., Scales, B., 2008. Review of Australian Higher Education: Final Report. Commonwealth of Australia, Canberra.
- Calabrese Barton, A., Kang, H., Tan, E., O'Neill, T.B., Bautista-Guerra, J., Brecklin, C., 2013. Crafting a future in science: tracing middle school girls' identity work over time and space. *Am. Educ. Res. J.* 50 (1), 37–57.
- Davies, B., 1993. *Shards of Glass: Children, Reading and Writing Beyond Gendered Identities*. Unwin and Allen, Sydney.
- DeWitt, J., Archer, L., 2015. Who aspires to a science career? A comparison of survey responses from primary and secondary school students. *Int. J. Sci. Educ.* 37 (12), 2170–2190.
- Diez-Palomar, J., Mara, L.C., 2020. Mathematics makes you feel attractive. Empowering new alternative masculinities in the context of mathematics classrooms. *Masculinities Soc. Change* 9 (1), 53–84.
- Francis, B., 2000. Boys, Girls and Achievement: Addressing the Classroom Issues. Routledge/Falmer, London.
- Francis, B., 2011. Postmodern and poststructural theories. In: Bank, B.J. (Ed.), *Gender and Higher Education*. JHU Press, Baltimore, MD, pp. 78–85.
- Harris, A., 2004. *Future Girl: Young Women in the Twenty-First Century*. Routledge, London, England.
- Harvey, A., Szalkowicz, G., Luckman, M., 2017. The Re-recruitment of Students Who Have Withdrawn From Australian Higher Education: Report for the Australian Government Department of Education and Training. Centre for Higher Education Equity and Diversity Research, La Trobe University, Melbourne.
- Harwood, V., Hickey-Moody, A., McMahon, S., O'Shea, S., 2018. *The Politics of Widening Participation and University Access for Young People: Making Educational Futures*. Routledge, Abingdon, Oxon.
- Hodge, B., Wright, B., Bennett, P., 2018. The role of grit in determining engagement and academic outcomes for university students. *Res. High. Educ.* 59 (4), 448–460.

- Huynh, K., Woo, B., 2014. "Asian fail": Chinese Canadian men talk about race, masculinity, and the nerd stereotype. *Soc. Ident.* 20 (4–5), 363–378.
- Kenway, J., Kelly, P.J., 2000. Local/global labor markets & the restructuring of gender, schooling and work. In: Stromquist, N.P., Monkman, K. (Eds.), *Globalization & Education: Integration & Contestation Across Cultures*. Rowman & Littlefield Publishers, Maryland, pp. 173–197.
- Kurlaender, M., Hibell, J., 2018. Students' educational pathways: aspirations, decisions, and constrained choices along the education lifecourse. In: Schneider, B. (Ed.), *Handbook of the Sociology of Education in the 21st Century*. Springer, Singapore, pp. 361–384.
- Lehmann, W., 2013. Habitus transformation and hidden injuries: successful working-class university students. *Sociol. Educ.* 87 (1), 1–15.
- Lucey, H., Melody, J., Walkerdine, V., 2003. Uneasy hybrids: psychosocial aspects of becoming educationally successful for working-class young women. *Gend. Educ.* 15 (3), 285–299.
- Mac an Ghaill, M., 1994. *The Making of Men: Masculinities, Sexualities and Schooling*. Open University Press, Buckingham.
- McDonald, S., 2021. She's like, "you're a uni student now": the influence of mother–daughter relationships on the constructions of learner identities of first-in-family girls. In: Brookes, R., O'Shea, S. (Eds.), *Reimagining the Higher Education Student*. Routledge, London, pp. 27–44.
- Mendick, H., 2006. *Masculinities in Mathematics*. Open University Press, Buckingham, UK.
- O'Shea, S., Stone, C., Delahunty, J., May, J., 2018. Discourses of betterment and opportunity: exploring the privileging of university attendance for first-in-family learners. *Stud. High Educ.* 43 (6), 1020–1033.
- Paechter, C., 2002. Masculinities and femininities as communities of practice. *Women Stud. Int. Forum* 26 (1), 69–77.
- Paechter, C., 2003. Masculinities and femininities as communities of practice. *Women Stud. Int. Forum* 26 (1), 69–77.
- Patfield, S., Gore, J., Fray, L., 2020. Degrees of "being first": toward a nuanced understanding of first-generation entrants to higher education. *Educ. Rev.* 1–20. <https://doi.org/10.1080/00131911.2020.1740172>.
- Pike, G., Kuh, G., 2005. First- and second-generation college students: a comparison of their engagement and intellectual development. *J. High. Educ.* 76 (3), 276–300.
- Pomerantz, S., Raby, R., 2017. *Smart Girls: Success, School, and the Myth of Post-feminism*. University of California Press, Berkeley, California.
- Pomeroy, D., 2015. Ethnicity, socio-economic status, and non-mathematical masculinities. *Res. Math. Educ.* 17 (3), 244–246.
- Pötschulat, M., Moran, M., Jones, P., 2020. "The student experience" and the remaking of contemporary studenthood: a critical intervention. *Socio. Rev.* 69 (1), 3–20.
- Reay, D., 1998. "Always knowing" and "never being sure": familial and institutional habituses and higher education choice. *J. Educ. Pol.* 13 (4), 519–529.
- Reay, D., 2002. Shaun's story: troubling discourses on white working-class masculinities. *Gend. Educ.* 14 (3), 221–234.
- Renold, E., Allan, A., 2006. Bright and beautiful: high achieving girls, ambivalent femininities, and the feminization of success in the primary school. *Discourse* 27 (4), 457–473. <https://doi.org/10.1080/01596300600988606>.
- Risman, B.J., 2004. Gender as a social structure. *Theory wrestling with activism*. *Gend. Soc.* 18 (4), 429–450.
- Skelton, C., Francis, B., 2011. Successful boys and literacy: are "literate boys" challenging or repackaging hegemonic masculinity? *Curric. Inq.* 41 (4), 456–479.
- Southgate, E., Brosnan, C., Lempp, H., Kelly, B., Wright, S., Outram, S., Bennett, A., 2017. Travels in extreme social mobility: how first-in-family students find their way into and through medical education. *Crit. Stud. Educ.* 58, 242–260.
- Spiegler, T., Bednarek, A., 2013. First-generation students: what we ask, what we know and what it means: an international review of the state of research. *Int. Stud. Sociol. Educ.* 23 (4), 318–337.
- Stahl, G., McDonald, S., 2021. *Gendering the First-in-Family Experience: Transitions, Liminality and Performativity*. Routledge, London.
- Stahl, G., McDonald, S., Loeser, C., 2020. Feminine subjectivities and aspirational learner identities: Asian–Australian young women navigating possible selves in the first year of university. *Gend. Educ.* <https://doi.org/10.1080/09540253.2020.1786013>.
- Stahl, G., 2014. White working-class male narratives of "loyalty to self" in discourses of aspiration. *Br. J. Sociol. Educ.* 37 (5), 663–683.
- Stahl, G., 2020. Class differences and impact on student access and outcomes. In: David, M., Amey, M.J. (Eds.), *The SAGE Encyclopedia of Higher Education*. SAGE, London, pp. 244–246.
- Stahl, G., 2021. *Working-Class Masculinities in Australian Higher Education: Policies, Pathways and Progress*. Routledge, London.
- Walkerdine, V., 2003. Reclassifying upward mobility: femininity and the neo-liberal subject. *Gend. Educ.* 15 (3), 237–248.
- Weiner, G., Arnot, M., 1987. Teachers and gender politics. In: Arnot, M., Weiner, G. (Eds.), *Gender and the Politics of Schooling*. Open University Press, Boston, pp. 354–370.

Minority learners, diversity, and educational outcomes: the case of Aotearoa, New Zealand

Cherie Chu-Fuluifaga, Te Herenga Waka, Victoria University of Wellington, Wellington, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	398
Negative and “color-blind” discourses on minority learners	399
Minority learner experiences in higher education	400
Pacific people in Aotearoa New Zealand	400
Social relationships in higher education	400
Conclusion	403
References	404

Introduction

Minority learners include diverse groups. Even when minority learners are defined by race, culture, or ethnicity, there are multiple layers of identity and complexity among any minority learner group. Minority learners often feel a sense of “in betweenness,” which can create a poor sense of entitlement and tension about their rights to occupy education spaces in majority cultures. Minority learners occupy a “third space” reflecting a hybridity which has emerged from the experiences of living between cultures (Aujla-Bhullar, 2016). It has been reported that minority learners do not achieve at the same rate as students from majority cultures (Vang, 2005) and gaps between these groups of learners have persisted, despite considerable efforts to address disparities. Minority learners within school systems often access fewer educational resources, encounter lower educator expectations, and through encouragement to drop out or take lower level courses, have poorer achievement rates (Stuart et al., 2011). These experiences contribute to many minority learners having a negative view of the school system, particularly secondary and higher education. Minority learners are often the first person in their family to pursue higher education and so are more reliant on lecturers’ support to guide their learning, encouragement which is often withheld (Stuart et al., 2011). Jackson and Moore (2008) describe a “doom and gloom” trajectory for black male college students who face a number of barriers within higher education, such as low expectations of academic staff and feeling overwhelmed by course content. Harper (2009) states that in the United States more than two-thirds (67.9%) of black men who start college do not graduate within six years, which is the lowest college completion rate among both sexes and racial/ethnic groups.

Minority learners who also face economic and social barriers experience particular disadvantage in education systems, because these factors impact on academic attainment. Despite admission to higher education having increased for students of color in the United States, minority learners from impoverished areas are less likely to attend higher education than their peers (Reddick et al., 2011). In New Zealand, Pacific children who come from some of the least well off families have the lowest rates of early childhood enrollments, with 86.7% of Pasifika children enrolled in 2012 compared to 98% of European children. A total of 68.1% of Pasifika students left school with the National Certificate of Educational Achievement (NCEA) level 2 or above in 2012, compared with 82.1% of European students. NCEA is the official secondary-school qualification. Level 2 indicates that students have a minimum of 60 credits and have met literacy and numeracy requirements at level 1. Only 11.5% of the Pasifika population have a Bachelor’s degree or above, in contrast to 18.6% of the European population (Marriott and Sim, 2014).

In particular, minority learners from indigenous populations, living in postcolonial Western countries, face the biggest disadvantage. Their lower education participation rates are evident across the United States, Canada, New Zealand, and Australia. Poor educational success and low secondary education completion rates have a negative impact on job and earning prospects, trapping future children of indigenous students in cycles of poverty. Indigenous students, across jurisdictions, struggle in primary and secondary schools which do not reflect their knowledge base or cultural styles. Even at university they report feeling isolated and frustrated, facing interpersonal discrimination and poor interaction with nonindigenous students (Bailey, 2016).

It is apparent from research that across jurisdictions the poorest educational outcomes are experienced by indigenous learners. In New Zealand, Māori early childhood education rates lag behind European rates, with 90.9% enrollments compared to 98% of European children (Marriott and Sim, 2014). Māori have lower secondary school retention, 73.7% of Māori youth at 16 years attend schools compared to 91.8% of non-Māori (Te Puni Kokiri, 2012). For 18- and 19-year-olds enrolled in higher education in 2003, 52% had completed their qualification by 2010 compared with 67% of all 18- and 19-year-olds (Te Puni Kokiri, 2012). Within the Canadian education context, in 2006, only 9% of First Nations women had obtained a university degree, compared with 23% of women in the general population (Bingham et al., 2014). First Nation students need to adapt and shift to the culture of educational institutions, a process which they compare to historic colonial experiences. They experience negative stereotypes and prejudice from other students and inaccurate portrayals of their people in the curriculum (Gallop and Baestein, 2016). In Hawaii, 28% of public school students identify as Native Hawaiian or part Hawaiian. These students encounter a range of struggles,

achieving at a lower level in standardized tests, which generate low rates of enrollment in higher education (Yamamoto and Black, 2015). For Alaskan students, Campbell (2007) estimates 29% drop out of high school, compared to 16% of their white counterparts. Alaskan students make up only 1% of students enrolled in postsecondary institutions and only 8% of those have higher degrees (Campbell, 2007). There are similar educational disparities between Aboriginal students and the rest of the population. Despite an increase on Aboriginal student participation in high school, from 35% in 1996 to 44% in 2006, this has not occurred in higher education settings, with a large decline in indigenous student participation after compulsory schooling (O'Shea et al., 2013).

Negative and “color-blind” discourses on minority learners

A negative and “color-blind” discourse exists within the literature on minority learners, which views these poor education outcomes as the result of economic deprivation, housing problems, poor family backgrounds and cultural deficits. Failure of minority learners is explained by inadequate family systems and community structures, which are seen to cause behavioral problems among students. Cultural deficits are also viewed as a significant factor in explaining the socio-economic status and educational failure of minority learners. George (2000), for example, claims black students belong to a culture infected with an anti-intellectual strain which teaches them not to embrace school work. These ideas are often transmitted through educator education programs, where trainee educators learn to encourage minority students to reduce family and community affiliations, and adopt dominant white middle-class cultural practices in order to succeed (Bishop et al., 2003). A similar negative discourse attributes the poor educational outcomes of international students to cultural distance, with a focus on cultural assimilation as a means to improve student integration and success (Vu, 2013).

Assumed lack of family involvement is a significant part of negative discourses concerning minority learner underachievement. In a study of educators' discursive positioning in relation to factors which influence New Zealand Māori achievement, most viewed Māori students themselves or their home contexts as contributing to poor educational outcomes pathologizing their lived experiences (Bishop et al., 2003). Perrot (2015) declares that there is also a perception that Pasifika parents' lack of understanding about their children's education and disconnect between home and school life limits their ability to help their children. Participants in Fletcher et al. (2009) study contextualized poor oral language environments at home with the socio-economic demands in Pasifika families, meaning there is less time for parents to develop rich oral environments to support learning. Zalaquett and Lopez (2007) report that a lack of strong adult supervision is seen as a contributing factor to the struggle of Latinx students within the education system. While Latinx parents are thought to understand the value of education, they do not have experience to assist their children, due to language barriers, low income attainment and insufficient knowledge of the higher education application process (Zalaquett and Lopez, 2007).

This negative discourse is pervasive and is reinforced when educators bring similar negative preconceptions about students' abilities to their teaching practice. In the classroom, they can develop low expectations of minority learners, viewing them as poor listeners, who talk out of turn or speak too softly. Lower educator expectations are noted for Māori students at all levels impacting on their achievement, particularly those moving from Māori medium to English medium schools (Oag.govt.nz, 2012). Nakhid (2003) argues similarly that the true identities of Pacific students are undervalued and that they are not satisfied with the perceptions that their educators have of them. O'Shea et al. (2013) also describes a pervasive “mindset” within the Australian school system, which seems to accept Aboriginal underachievement as normal, based on poor student attendance, a lack of systematic analysis of indigenous student performance and underdevelopment of strong indigenous educational leadership (O'Shea et al., 2013).

This negative discourse is “colorblind” to the impact of culture and racism on the education of minority learners. Yamamoto and Black (2015) argue that Native Hawaiian students struggle to succeed in Western-based education and economic systems. For them to succeed there is the expectation that they take on manners and values related to success in a Western system. These students then struggle with identity negotiation in the school system, instructional methods and measures of accountability, which contest their own values (Yamamoto and Black, 2015).

This explains why disparities have persisted despite education reforms and attempts to develop bicultural and multicultural approaches to mainstream education. Reforms are ineffective because they are developed within a neo-colonial paradigm, based on Western knowledge that has racism embedded in the privileging of the dominant cultural practices and processes. It also imposes sense-making and knowledge-generating approaches from outside minority cultural groups, which in New Zealand, for example, fails to acknowledge Mataranga Māori (Māori ways of knowing) (Berryman and Eley, 2017). Hemara (2004) considers that gaps between Māori and non-Māori students need to be considered within the context of community values and argues that the important disparity is Māori aspirations and their actual achievements, arguing that traditional values can be transitioned into contemporary contexts. Mahuika et al. (2011) also argue for acknowledgment of the uniqueness of Māori learners, and for assessments which reflect better the cultural needs of these students.

Whether Pacific learners are minority learners remains a question to be debated. We are resistant to labeling them, for fear of deficit theorizing.

Minority learner experiences in higher education

Pacific people in Aotearoa New Zealand

Who are Pacific people in Aotearoa New Zealand? This is a question that is constantly discussed in educational spaces. What term is most relevant? What term is most suitable? Actually, there is no one appropriate response. The umbrella term “Pacific” has been used by the New Zealand government, agencies, educational institutions, and academics to describe the ethnic make-up of people migrating from the Pacific Islands to Aotearoa New Zealand (NZ) (Cook et al., 1999). Bedford and Didham (2001) state that the term “Pacific” has been commonly and widely utilized at all levels of society including educators, policy makers, community workers, the media and institutions. The use of the term has often led to broad generalizations about a group of people who have diverse cultural identities.

A foundational document for research titled the “Pasifika Education Research Guidelines” (Anae et al., 2001), developed for the Ministry of Education in Aotearoa New Zealand, provided one definition of Pacific peoples. At the time of development, it made reference to the six Pacific nations of Samoa, Tonga, Niue, Cook Islands, Tokelau, and Fiji. In this context, “Pacific people” is exclusive of Indigenous Māori and in the broadest sense covers peoples from the Island Nations in the South Pacific, and in its narrowest sense, Pacific peoples in Aotearoa New Zealand. The research guidelines go on to clarify the issue of Pacific people as being a heterogeneous group with different inter- and intra-ethnic variations in the cultures. Variations include New Zealand-born/raised, island-born/raised Pacific people and being recognized as diverse groups. For the purposes of this article “Pacific” has been utilized, rather than “Pacific Island.” “Pacific” refers to people who can “trace descent to and/or are citizens of any of the territories commonly understood to be part of the Pacific (i.e., Melanesia, Micronesia, Polynesia)” (Fairbairn-Dunlop et al., 2008, p. 11).

The way Pacific learners are positioned and theorized about is interesting for educators. Pacific learners are often theorized as minority learners or diverse learners. But what does this really mean. The process of colonization has been one of the most fundamental disruptions of Pacific Islands’ systems. Colonization has impacted on Pacific peoples’ values, beliefs, language, philosophies, religions, and social, cultural and economic systems. There are challenges, narratives and areas of attention which have encouraged Pacific people to reclaim indigenous identities. Decolonization in the Pacific Islands region has been focused on communities in the Pacific region for strengthening and developing the well-being of young Pacific people in higher education (Chu, 2009).

Decolonization for Pacific people has been a gradual process of discussions and attitudinal changes through talanoa/discussions and informal educational transformations. The diverse Pacific region has contributed to a more fluid initiation into decolonization for varying groups. The year 2000 marked a significant decolonization shift in education for the Pacific region. Pacific people initiated the process of identifying challenges, needs, and areas of attention for their countries. The Rethinking Pacific Education Initiative (RPEI) was created as a cause movement by a group of Pacific educators and leaders who wanted to see deeper systemic changes occur for Pacific learners and their communities (Chu, 2009). The decolonization movement has allowed Pacific people to reclaim indigenous Oceanic perspectives, knowledge and wisdom that have been devalued and oppressed. Activist and thought leader, Thaman (2003) has provided much of the impetus to enable Pacific people to reflect critically on the nature, scope, and processes of colonialism in Oceania. The challenges have been thrust at the dominance of Western philosophy and pedagogy. Thaman’s early criticism of colonization has fueled the process of decolonizing the field of Pacific studies. The impact of colonialism on people’s minds has impacted in their ways of knowing, their views of who and what they are, and what they consider worthwhile to educate and to learn. The premise is that Pacific learners should be supported in higher education to live and learn with their Pacific cultural identities intact. Pacific students do not have to leave their cultural identities at the door to become somebody else or to have to assimilate into the dominant cultural identities of higher education.

Social relationships in higher education

In the South Pacific region, higher education is also called post-secondary education, and is any level of education pursued beyond high school, including undergraduate and graduate credentials. These credentials encompass certificates, diplomas or academic degrees in Aotearoa New Zealand. An arm of the New Zealand government that overlooks higher education is the Tertiary Education Commission (TEC). The TEC has continually stipulated through the past two decades that one of its priorities is to increase the outcomes of Pacific students achieving at higher levels. Further, TEC believes that the NZ higher education sector has a key role to play in meeting the development needs and aspirations of Pacific peoples. New Zealand research suggests institutions have to create environments that engage students with institutional culture by offering support in a variety of ways. The premise is that the New Zealand education system does not adequately meet the needs of Pacific learners. There is evidence of varying support programs that have been developed to enhance the pastoral care of students, but without any thought to the unique circumstances of Pacific learners’ academic needs. Others have been superficially adapted with respect to program materials, but without a deeper consideration of the myriad programmatic, organizational and external factors that exist. Higher staff are eager for guidance to more meaningfully adapt or develop teaching and learning methods that meet the needs of these students. Rather than focusing on these negative statistics, educators in higher education must shift the whole paradigm to a strengths-based approach. This is an urgent shift. A strengths-based approach focuses on developing assets that are known protective factors, such as strong relationships, life skills and school connectedness.

As a response to colonization, one question can be asked, what is success for Pacific people in higher education? Pacific scholars have tried to address this issue of success for Pacific people in all forms of education (Perrot, 2015). While there are signs that Pacific achievement is trending up and has made incremental improvements (Southwick et al., 2017), comparable to other groups, Pacific learners are still behind, statistically (Southwick et al., 2017). Research has shown that achievement for Pacific people in education should be focused on the authentic development of self (Perrot, 2015; Chu et al., 2013), where education and learning environments develop students as strong positive people who have confidence in their ability and connected with quality education that builds quality nurturing relationships with staff, educators and the school. And also when environments are inclusive, which reflect and legitimize Pacific culture, connections and values, the result can be successful academic achievement. In the context of decolonization, achievement and educational outcomes will have a different meaning for communities. This meaning is based on the values of the community. For example, education for diverse Pacific communities is about building and fostering relationships, rather than solely focusing on the measurement of grades as success. Racism is revealed to be an issue at the tertiary level by Mayeda et al. (2014). In addition to experiencing university as a “white” place as a result of a Eurocentric curriculum, the research describes the way on-campus racism reinforces a sense of alienation for Māori and Pacific students. This can take the form of assumptions by other students of Māori and Pacific privilege regarding entry to tertiary education. This implies a lack of effort and ability. A sense that majority-group students require others to forsake their identity in order to fit in can be felt by Pacific students. The adoption of a low profile can be the result of repeated discrimination. The researchers suggest that, although some students succeed in converting discrimination to motivation, this responsibility should not rest with students. Consequently, a recommendation is that management teams should not only address perpetrators of discrimination, but also work with Māori and Pacific students to warn them of the issues and encourage them in ways of channeling the negative emotions they are likely to encounter as a result of everyday racism and colonialism in education.

If educators apply the law of “indirect effort” by steering their focus on the development of the person through authentic cultural learning environments and the building of quality relationships with students, achievement that is based on peoples’ aspirations and ideals will follow. Success can be explored through a holistic lens of student identity and learning needs, this point resonates with the “Third Space” (Bhabha, 2004), which posits the theory of hybridity to include identity, community and education. The Third Space allows for recognition of the uniqueness of each person and/or context in their lived experiences. It is a place where hybrid identifications are possible and where dialogs about contestation, conflict and politics between cultures progress. Borders can separate cultures and societies. In today’s globalized world, cultures are blending increasingly. As traditional colonial knowledge is mixed with peoples’ own indigenous knowledge the opportunity exists for creative heterogeneity. The result of this encounter is the emergence of a hybrid culture that can no longer be traced back to the roots of either community. Bhabha asserts that nations and cultures must be understood as narrative constructions rising out of the hybrid relations of competing constituencies. There is a continual process of interpretation and reinterpretation. In education this is evidenced by an individual’s characteristics evolving. Bhabha studied the space constituted around the encounters between the colonizers and the colonized. Within the Third Space hybrid cultures are created as a fusion of the colonizer and the colonized. As one focal point, the middle ground between Pacific communities and the State has provoked an ongoing debate about whose values are represented in the NZ educational system, and the absence of Pacific peoples’ values from the classroom environment (Chu et al., 2013).

To take this notion of the Third Space for Pacific learners in Aotearoa New Zealand, it is evident that the Third Space is a complex ground. When educators put the person at the center of learning, that person becomes self-determined for achievement and not the other way around. Educators who put achievement over personal development are in for an arduous task if they desire success for their Pacific students. Person-centered approaches (such as educational mentoring) embody a strong sense of self can help to develop a strong identity for young Pacific people. When Pacific students are strong, they embrace their own story; when they embody self-awareness, then their personal learning and achievement are almost predictable. The terms of learning and achievement are defined by the students. In the spaces of the classroom, skilled educators draw out the cultural identities of students with the provision of appropriate content, assessment and curriculum. This process must be one that is reclaimed and defined by the students and not the institution. What is also rewarding is seeing the shift that happens within students when they finally feel successful (Perrot, 2015). Personal development for students has positive effects. In higher education, mentoring has been described as a one-to-one learning relationship with the mentor being older than the student, perhaps a faculty member, staff member or student. The process of the relationship, which can be informal or formal, is based on modeling behavior and extended dialog (Johnson, 1989). However, Jacobi (1991) found that 67% of undergraduate students at a large West Coast university in the United States reported that it was challenging to find a mentor or role model. It is further suggested by Jacobi (1991) that the prevalence of natural or informal mentoring relationships in educational contexts, including undergraduate and graduate education, is unknown. Since the 1980s there has been an increase in mentoring of undergraduate students as a retention strategy rather than the focus on the apprentice model of graduate education in the United States (Jacobi, 1991). For instance, using undergraduate students as peer mentors discounts the importance of age differences between mentors and protégés (Rice and Brown, 1990). Knox and McGovern (1988) found that 51% of mentoring relationships were developed during graduate school and another study showed that graduate school protégés who had strong collaborative relationships with their mentors tended to subsequently achieve an academic position (Long and McGinnis, 1985). For minority ethnic groups, Johnson (1989) discovered that mentoring for undergraduate women and students of color was relatively rare, whereas Blackwell (1989) found that one in eight African-American students had a true mentor. However, at the graduate level, mentoring appeared to be more common (Jacobi, 1991). Due to a variety of factors of mentoring in higher education, the understandings and definitions vary (Merriam, 1983). The variance in factors is attributed to different student populations (students of different ethnic backgrounds, men vs. women); different academic levels

(undergraduates vs. graduate students); different institutional characteristics (small arts colleges vs. large public institutions); and different fields of study. In Aotearoa New Zealand there has been a growing collection of educational research on Pacific learners with a greater emphasis on understanding the theoretical concepts of what works for Pacific learners as they go about their educational journeys. Much of the established research discusses in general what people or organizations provide to enable Pacific learners. The concept of the learning village looks at the importance of the contextual and institutional factors identified from the literature that enhanced positive outcomes for Pacific students. Overall these factors enable students to feel safe as Pacific individuals in education. [Benseman et al. \(2006\)](#) identified contextual factors that would help increase Pacific student retention. Central to the educational success of students is the connection and communication between the student, the community, and the institution. These include:

- Presence of Pacific academics and staff allied with services and programs designed to specifically support Pacific students.
- Pacific presence in institutions such as provision of dedicated space, holding Pacific events, and creation of Pacific associations and a supportive environment for Pacific students.
- Provision of role models such as educators, parents, older siblings, or cousins.
- Pedagogical practices which are student-centered, engaging both learners and lecturer, committed to high expectation and achievement for all students, availability of academic support, access to resources, accessible pastoral care and a “staircase” environment that provides links for one level of qualification to the next.
- Accurate, appropriate, and timely information made accessible and available to all students.
- Presence of Pacific staff in leadership roles in higher institutions, as indicated by [Peterson et al. \(2006\)](#).

This provision of caring practice is aligned with [Tinto’s \(1993\)](#) “Model of Institutional Departure” states that students need integration into formal (academic performance) and informal (faculty/staff interactions) academic systems and formal (extracurricular activities) and informal (peer-group interactions) social systems. Through deliberate practice, staff are committed to the students they serve, and that they put student welfare ahead of other institutional goals. This means that institutional goals should always have a direct or indirect relationship to student success and achievement and be committed to the education of all, not just some, of their students. Furthermore, Tinto emphasizes that students must feel that they are valued and full members of the community. That also means that the expectations of achievement and behavior are the same for all students. The interaction student attrition model of [Tinto \(1975, 1998\)](#) explains that students have a variety of educational experiences, skills and competencies, family and community backgrounds and values. These social and individual factors influence how a student integrates into their studies at universities. Several programs in Aotearoa New Zealand have implemented inclusive environments to support Pacific students in their journey through higher education. As some higher education success examples which embedded social and individual values of learners, Te Rōpū Āwhina was a program at Victoria University of Wellington that produced Māori and Pacific professionals who could contribute towards community development and leadership. The principle behind the initiative was the creation of an inclusive off and on campus whanau/family environment imbued with positive values of high expectations, achievement, collective success and reciprocity. Based on the quantitative analysis of student achievement and student surveys, the results show consistency with the improvement in achievement and retention of Māori and Pacific students ([Wilson, 2011](#)). It can be inferred that such a mentoring program for ethnic minority students at a university leads to a change in the mainstream culture. Mentoring programs provide significant changes for education, and this leads to a process where Pacific peoples are able to recreate a new narrative, rejecting the negative narrative of society. Where the goal is to bring about change by shifting the culture of the organization, additional degrees of complexity and resistance are expected.

Cultures within higher education need to be examined in ways that embrace ethnic minority groups of students. [Tierney’s \(1988\)](#) work provides a framework for higher education culture which includes six major components: environment, mission, socialization, information, strategy, and leadership. The socialization element represents one aspect in which mentoring can contribute additional information. This can be illustrated with [Tierney’s \(1988\)](#) framework, where he asserts that socialization takes into account answering such questions as “How do members become socialized? How is it articulated? What do we need to know to survive/excel in this organization?” The educational organization reflects the norms, values and practices of the wider society and can present significant challenges for diverse communities whose values are not represented. Cultural alignment is relevant to educational access for Pacific students in NZ because, where educational organizations listen to and understand what matters to families and communities, students are more likely to be supported within inclusive relationships. However, where relationships between schools and communities are poorly configured, Pacific student success depends on the skills of unsupported students to negotiate between conflicting cultural demands or choices. An unbridged distance between home and school can operate to make access to all aspects of education problematic for Pacific students since it ignores the collective basis of Pacific cultures.

[Millward et al. \(2012\)](#) examined the retention and success rates of students at three different campuses which were similar despite the differences in ethnicity, socio-economic background and previous academic success. The underrepresented groups in two out of the three campuses, mainly Māori and Pacific, entered the university via the special admissions pathway and passed their courses at similar rates as the regular entry students. This links into the study of [Bunce et al. \(2021\)](#) and the experiences of British university students from black and minority ethnic (BME) backgrounds who were less likely to achieve a “good” degree classification than white students. The researchers used the self-determination theory which proposes that achieving one’s full potential for learning, alongside experience of well-being, is supported by environments that help individuals to meet their needs for relatedness, competence and autonomy. They found barriers which impaired BME students’ experiences situated in the university structures.

Accordingly, Davidson-Toumu'a and Dunbar (2009) argued that education and culture should go together and that educators should be aware of how the two concepts work. They added that Pacific students should be viewed as a whole person whose world is filled with cultural roles and familial obligations other than just being a student. Pacific success in higher education is evidenced when Pacific students build quality relationships with academic and nonacademic staff and other students within their programs. Quality relationships are defined as those that are reciprocal in nature and are intentional in their intent towards growth and development for Pacific individuals both as people and students.

Maxwell (2007, p. 116) states that "... people don't care how much you know until they know how much you care." Evidence in Aotearoa New Zealand suggests that when educators primarily focused on the care of Pacific students, facilitated through the development of authentic reciprocal relationships, the by-product was increased Pacific academic achievement in higher studies, as defined by the learners themselves (Southwick et al., 2017; Chu et al., 2013). Students' reflections revealed that their success had more to do with the quality relationships they developed with their tutors and/or lecturers than the quality or quantity of content students received (Southwick et al., 2017; Chu et al., 2013). The realization of equitable academic achievement for Pacific as Pacific requires academic staff to shift their philosophy of teaching to be primarily about relationship building than content delivery. Conceptually, the "relationship" developed has to be seen as the vehicle through which knowledge, learning, and content are funneled. But without a strong relationship this can be a barrier. In stark contrast, a study by Tran (2013) explored ethnic minorities' experiences in education in Vietnam. They found that when teachers had low expectations or neglected students, this had negative impacts. The abandoned students felt alienated from the school, hardly participated in the class activities, had poor educational performance, felt discriminated against, and eventually dropped out. The problems between students and teachers existed due to the physical and social distance among teachers and students, differences in cultural sensitivity, lack of extracurricular activities, and lack of encouragement and praise from teachers.

Chu et al. (2013) identified that Pacific students across multiple institutions felt that achievement as students came about through relationships described as respectful and nurturing. The word nurturing in the Oxford University Press Dictionary (2020, para 1) is explained as to "Care for and protect (someone or something) while they are growing." This care and protection are evidenced by educators that go outside of their job description for Pacific students. They demonstrated an eagerness to put in extra one-on-one time for them, and demanded high expectations and standards for them as students. Educators held the unyielding belief that they could be successful as students, coaching and mentoring them to pursue higher level qualifications and set goals. Further, educators who put in work to ensure their teaching strategies and design were student centered met the needs and multiple literacies of students. Students described staff as flexible and passionate, and that they made an effort to connect with students' communities and families. Educators also helped students "edge walk" the multiple worlds they existed in and understood or also took the time to understand Pacific students' worlds. In this process they made students feel proud and strong about their multiple identities and cultures. The educators also enabled students to critically reflect on how their success as students was a form of social justice to an education system that seemed structured in such a way as to prevent their success (Southwick et al., 2017; Chu et al., 2013; Marshall et al., 2008). This level of care from educators ultimately protected students, and for many it enabled them for the first time to feel safe, cared for, liberated, secure, and strong as Pacific people in education. This notion of edge-walking can draw some similarities to the Native American situation where higher education organizations have reported retention challenges for Native American students. Agbo (2001) suggests that high dropout rates and low levels of academic achievement of Indigenous students are not due to the inability to adjust, cultural impoverishment, or cultural differences, but rest in cultural discontinuity. The main argument of cultural discontinuity is the assertion that Indigenous peoples experience reality in a completely different way to the dominant culture. As a result, the extent of the incongruence of realities Indigenous peoples often experience within the dominant culture can be devastating to the Indigenous sense of self and experience of being in the world. As a response to the lived realities, Shield (2004) provided a framework called the Medicine Wheel where the aspects of community, cultural values, Indigenous identity, and spirituality are encompassed. These components are necessary for the development and well-being of Indigenous learners in their educational experience. The model is Indigenous in substance and content, and it is generated and grounded in Indigenous thought and view. The model relies on the foundation of spirituality and solidification of Indigenous identity throughout all aspects of its design.

Conclusion

Social relationships for minority ethnic learners in higher education can enable learners to enjoy academic and personal success while studying. Minority learners often experience challenges in higher education that is beyond their immediate control. Institutions and educators have a responsibility to provide teaching and learning that is conducive to the cultural needs of learners. The process of positive learning can be evidenced as educational reform as a way to strengthen the lives and cultures of learners and their communities. The importance of relationships for Pacific students in education is crucial. If Pacific students are to succeed as Pacific in higher education, academic and nonacademic/professional staff have to foster quality relationships with students. Student have to feel cared for and protected by the people that are in education and supporting them. Staff as well as institutions also have to foster learning environments that are collaborative and enable aiga (family) feelings of well-being among students. Relationship building cannot be an optional extra for educators. It has to be part of their core function. The narrative about Pacific learners and other minority ethnic groups must be centered on strengths-based approaches as the populations of learners become more

diverse in Western institutions. Instead of inflicting negative discourses about diverse learners, the policies, teaching, and resources should reflect the identities and needs of ethnic learners so that they have a sense of belonging.

References

- Agbo, S.A., 2001. Enhancing success in American Indian students: participatory research at Akwesasne as part of the development of a culturally relevant curriculum. *J. Am. Indian Educ.* 40 (1), 31–55.
- Anae, M., Anderson, H., Benseman, J., Coxon, E., 2001. *Pacific Peoples and Tertiary Education: Issues of Participation*. Auckland: Auckland Uniservices.
- Aujla-Bhullar, S., 2016. Breaking silence and opening windows: insider stories of South Asian Canadian women. *Cult. Pedagogical Inq.* 8 (2), 23–31. Retrieved from. <https://journals.library.ualberta.ca/cpi/index.php/cpi/article/view/28672/20964>.
- Bailey, K.A., 2016. Racism within the Canadian University: indigenous students' experiences. *Ethn. Racial Stud.* 39 (7), 1261–1279. <https://doi.org/10.1080/01419870.2015.1081961>.
- Bedford, R., Didham, R., 2001. Who are the "Pacific Peoples"? Ethnic identification and the New Zealand census. In: Macpherson, C., Spoonley, P., Anae, M. (Eds.), *Tangata O Te Moana Nui. The Evolving Identities of Pacific Peoples in Aotearoa/New Zealand*. Dunmore Press, Palmerston North, pp. 21–43.
- Benseman, J., Coxon, E., Anderson, H., Anae, M., 2006. Retaining non-traditional students: lessons learnt from Pasifika students in New Zealand. *High Educ. Res. Dev.* 25 (2), 147–162.
- Berryman, M., Eley, E., 2017. Succeeding as Māori: Māori students' views on our stepping up to the Ka Hikitia challenge. *N. Z. J. Educ. Stud.* 52 (1), 93–107. <https://doi.org/10.1007/s40841-017-0076-1>.
- Bhabha, H.K., 2004. *The Location of Culture*. Routledge, Abingdon, p. 55.
- Bingham, J., et al., 2014. Indigenous women college students' perspectives on college, work, and family. *J. Coll. Student Dev.* 55, 615–631. <https://doi.org/10.1353/csd.2014.0055>.
- Bishop, R., Berryman, M., Tiakiwai, S., Richardson, C., 2003. *Te Kōtahitanga Phase: The Experiences of Year 9 and 10 Māori Students in Mainstream Classrooms*. Retrieved from. https://www.educationcounts.govt.nz/_data/assets/pdf_file/0017/7532/te-kotahitanga.pdf.
- Blackwell, J., 1989. Mentoring: an action strategy for increasing minority faculty. *Academe* 75 (5), 8–14.
- Bunce, L., King, N., Saran, S., Talib, N., 2021. Experiences of black and minority ethnic (BME) students in higher education: applying self-determination theory to understand the BME attainment gap. *Stud. High Educ.* 46 (3), 534–547. <https://doi.org/10.1080/03075079.2019.1643305>.
- Campbell, A.E., 2007. Retaining American Indian/Alaskan native students in higher education: a case study of one partnership between the Tohono O'odham Nation and Pima Community College, Tucson, Arizona. *J. Am. Indian Educ.* 46 (2), 19–41.
- Chu, C., Abella, I., Paurini, S., 2013. *Educational Practices that Benefit Pacific Learners in Tertiary Education*. Ako Aotearoa, Wellington. Available at. <https://ako.ac.nz/assets/Knowledge-centre/NPF-10-001A-Pasifika-Learners-and-Success-in-Tertiary-Education/RESEARCH-REPORT-Educational-Practices-that-Benefit-Pacific-Learners-in-Tertiary-Education.pdf>.
- Chu, C.M., 2009. *Mentoring for Leadership in Pacific Education: A Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Education*. Thesis PhD. Victoria University of Wellington.
- Cook, L.W., Didham, R., Khawaja, M.A., 1999. On the demography of Pacific people in New Zealand. *Demogr. Trends* 19–41.
- Davidson-Toumu'a, M.R., Dunbar, K., 2009. Understanding the experiences of Pacific students and facilitating socio-cultural adjustment into higher education in Aotearoa, New Zealand. *J. Aust. N. Z. Stud. Serv. Assoc.* 33, 69–88.
- Fairbairn-Dunlop, P., et al., 2008. *Inventory of Pacific Research at Victoria University of Wellington 1999–2005*.
- Fletcher, J., Parkhill, F., Fa'afai, A., Taleni, L.T., O'Regan, B., 2009. Pasifika students: teachers and parents voice their perceptions of what provides supports and barriers to Pasifika students' achievement in literacy and learning. *Teach. Teach. Educ.* 25 (1), 24–33. <https://doi.org/10.1016/j.tate.2008.06.002>.
- Gallop, C.J., Baestein, N., 2016. Supporting success: aboriginal students in higher education. *Can. J. High. Educ.* 46 (2), 206–224.
- George, J.M., 2000. Emotions and leadership: the role of emotional intelligence. *Hum. Relat.* 53 (8), 1027–1055. <https://doi.org/10.1177/0018726700538001>.
- Harper, S.R., 2009. Race, interest convergence, and transfer outcomes for black male student athletes. *N. Dir. Community Colleges* (147), 29–37. <https://doi.org/10.1002/cc.375>.
- Hemara, W., 2004. *Whakapapa as Curricular and Pedagogical Practice: An Historical, Social and Philosophical Discourse on Māori Engagement With Traditional and Contemporary Education*. Victoria University of Wellington.
- Jackson, J.F.L., Moore, J.L., 2008. Beyond Brown: new approaches to addressing inequities in education for African American males. *Am. Behav. Sci.* 51 (7), 847–853. <https://doi.org/10.1177/0002764207311992>.
- Jacobi, M., 1991. Mentoring and undergraduate academic success: a literature review. *Rev. Educ. Res.* 61 (4), 505–532.
- Johnson, C.S., 1989. Mentoring programs. In: Upcraft, M.L., Gardner, J. (Eds.), *The Freshman Year Experience: Helping Students Survive and Succeed in College*. Jossey-Bass, San Francisco, CA, pp. 118–128.
- Knox, P.I., McGovern, T.V., 1988. Mentoring women in academia. *Teach. Psychol.* 15 (1), 39–41.
- Kōkiri, T.P., 2012. *Māori Youth in Education and Employment*. Retrieved from. [youth-in-education-and-employment](https://www.mta.govt.nz/youth-in-education-and-employment).
- Long, J.S., McGinnis, R., 1985. The effects of the mentor on the academic career. *Scientometrics* 8 (6), 255–280.
- Mahuika, R., Berryman, M., Bishop, R., 2011. Issues of culture and assessment in New Zealand education pertaining to Māori students. *Assess. Matters* 3, 183–198. <https://doi.org/10.18296/am.0093>.
- Mariott, L., Sim, D., 2014. *Indicators of Inequality for Māori and Pacific People*. Retrieved from. https://www.wgtn.ac.nz/saci/centres-and-chairs/cpi/publications/working-papers/WP09_2014_Indicators-of-Inequality.pdf.
- Marshall, J., Baldwin, K., Peach, R., 2008. *Te Rau Awhina the Guiding Leaf: Good Practice Examples of Māori and Pasifika Private Training Establishments*. New Zealand Qualifications Authority, Wellington, N.Z.
- Maxwell, D., 2007. Improving and learning from practice. *J. Pharm. Pract. Res.* 37 (1), 38–41. <https://doi.org/10.1002/j.2055-2335.2007.tb00656.x>.
- Mayeda, D.T., Keil, M., Dutton, H.D., Ofamo'oni, I., 2014. "You've gotta set a precedent": Māori and Pacific voices on student success in higher education. *Alternative* 10 (2), 165–179.
- Merriam, S., 1983. Mentors and protégés: a critical review of the literature. *Adult Educ.* 33 (3), 161–173. <https://doi.org/10.1177/074171368303300304>.
- Millward, P., Turner, R., Linden, D., 2012. The retention and success of under-represented groups in a Bachelor of Education program in New Zealand. *Asia-Pac. J. Teach. Educ.* 40, 171–183. <https://doi.org/10.1080/1359866X.2012.670873>.
- Nakhid, C., 2003. Intercultural perceptions, academic achievement, and the identifying process of Pacific Island students in New Zealand schools. *J. Negro Educ.* 72 (3), 297–317.
- Office of the Auditor General, 2012. *Education for Māori*. Context for Our Proposed Audit Work Until 2017. Retrieved from. <https://oag.parliament.nz/2012/education-for-Māori>.
- O'Shea, S., Harwood, V., Kervin, L., Humphry, N., 2013. Connection, challenge, and change: the narratives of university students mentoring young indigenous Australians. *Mentor. Tutoring* 21 (4), 392–411. <https://doi.org/10.1080/13611267.2013.855863>.
- Oxford Learner's Dictionaries, 2020. *Nurture*. Retrieved from. https://www.oxfordlearnersdictionaries.com/definition/english/nurture_1.

- Perrot, A.R.D., 2015. Overcoming Challenges: Pacific Students' Experiences of Being Resilient Through Tertiary Education. Available at. <http://researcharchive.vuw.ac.nz/handle/10063/5187>. (Accessed 13 May 2021).
- Peterson, G., Mitchell, D., Oettli, P., White, H., Kalavite, T., Harry, K.S., 2006. Engagement of Key Stakeholder Groups With the Tertiary Education Providers. Ministry of Education, Wellington, New Zealand.
- Reddick, R.J., et al., 2011. Stories of success: high minority, high poverty public school graduate narratives on accessing higher education. *J. Adv. Acad.* 22 (4), 594–618. <https://doi.org/10.1177/1932202X11414133>.
- Rice, M.B., Brown, R.D., 1990. Developmental factors associated with self-perceptions of mentoring competence and mentoring needs. *J. Coll. Student Dev.* 31, 293–299.
- Southwick, M., Scott, W., Mitaera, J., Nimarota, T., Falepau, L., 2017. Articulating a Pedagogy of Success for Pacific Students in Tertiary Education. Retrieved from Wellington. <https://ako.ac.nz/assets/Knowledge-centre/NPF-14-004-Articulating-and-implementing-a-pedagogy-of-success-for-Pacific-students-in-tertiary-education/5d213fbf60/RESEARCH-REPORT-Articulating-a-Pedagogy-of-Success-for-Pacific-Students-in-Tertiary-Education.pdf>.
- Stuart, M., Lido, C., Morgan, J., 2011. Personal stories: how students' social and cultural life histories interact with the field of higher education. *Int. J. Lifelong Educ.* 30 (4), 489–508. <https://doi.org/10.1080/02601370.2011.588463>.
- Thaman, K.H., 2003. Decolonizing Pacific studies: indigenous perspectives, knowledge, and wisdom in higher education. *Contemp. Pac.* 15 (1), 1–17. <https://doi.org/10.1353/cp.2003.0032>.
- Tierney, W.J., 1988. Organizational culture in higher education: defining the essentials. *J. High. Educ.* 59, 2–21. <https://doi.org/10.2307/1981868>.
- Tinto, V., 1993. Leaving College: Rethinking the Causes and Cures of Student Attrition. Second Edition. University of Chicago Press.
- Tinto, V., 1975. Dropout from higher education: a theoretical synthesis of recent research. *Educ. Eval. Pol. Anal.* 45, 89–125. <https://doi.org/10.2307/1170024>.
- Tinto, V., 1998. Colleges as communities: taking research on student persistence seriously. *Rev. High. Educ.* 21, 167–177.
- Tran, T.T., 2013. Is the learning approach of students from the confucian heritage culture problematic? *Educ. Res. Pol. Pract.* 12 (1), 57–65.
- Vang, C.Y., 2005. The Lao: gender, power, and livelihood. *Contemp. Sociol.* 34 (1), 28–30. <https://doi.org/10.1177/009430610503400117>.
- Vu, H.T., 2013. Vietnamese International Students' Experiences in a New Zealand University: A Narrative Study. Victoria University of Wellington. Retrieved from. <https://researcharchive.vuw.ac.nz/xmlui/handle/10063/2809>.
- Wilson, K., 2011. Social Work: An Introduction to Contemporary Practice.
- Yamamoto, K.K., Black, R.S., 2015. Standing behind and listening to native Hawaiian students in the transition process. In: *Career Development and Transition for Exceptional Individuals*, vol. 38, pp. 50–60. <https://doi.org/10.1177/2165143413498412>, 1.
- Zalaquett, C.P., Lopez, A.D., 2007. Learning from the stories of successful undergraduate Latina/Latino students: the importance of mentoring. *Mentor. Tutoring* 14 (3), 337–353. <https://doi.org/10.1080/13611260600635563>.

Experiencing class and gender: mature women students and their journeys from HE to the labor market

Barbara Merrill and Scott Revers, Centre for Lifelong Learning, University of Warwick, Coventry, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Background: higher education and the labor market	406
Definitions of employability	406
Student experience	407
The research	408
Gender and class	408
Class	409
A note on methodology: biographical approaches	410
Experiencing HE: gender and class	410
Early life experiences	411
Experiencing university life	412
Gender and class	412
Transitions and the graduate labor market	413
Summary	415
References	415

Background: higher education and the labor market

For many decades higher educational institutions in both the UK and globally have been changing in the light of policies influenced by marketization and neoliberalism (Molesworth et al., 2011). With the erosion of the Keynesian welfare state in the 1970s, the rise of neoliberalism has gradually transformed higher education (HE) into a market in which universities are expected to compete for funding while students play their part in shaping the system through their respective choices (Troschitz, 2017: 139). As a result the political discourse around higher education institutions has shifted away from HE as a publicly sourced and protected resource, a “public good”, with a critical role in defining and engaging with issues pertaining to society and culture, toward more of a business performance model in which issues like quality, competition and student satisfaction have become paramount (Molesworth et al., 2011). Such institutions cannot assume independence from the market but rather must justify their existence by satisfying relevant performance indicators to legitimate their status (Walters, 2012: 62).

The marketization of higher education has been accompanied in the UK by a government-backed drive toward a knowledge-based economy that has sought to broaden the scope and appeal of a university education. One of the outcomes of this transition has been that higher education is no longer simply the preserve of a privileged elite but has become a mass-based system (Trow, 1989). Since the 1990s the numbers of students, universities and colleges have multiplied. Furthermore, the student body has expanded in terms of class, gender, ethnicity, disability and age (Merrill, 2015), thereby capturing a more diverse swathe of the population including under-represented groups.

During this period the perception of the role of higher education has evolved into an environment concerned with delivering lifelong learning and employability skills for the labor market (Molesworth et al., 2011; Barnett, 2017). Higher educational institutions have met this challenge by advancing a series of measures designed to enhance student employability. These include more vocational modules, department and institutional initiatives. Work-based learning together with links with external organizations and entrepreneurial activities are also being encouraged.

The employability narrative has also been advanced in the wake of universities being monitored to gauge how well they perform in making their students employable. Graduation rates are published annually (O’Leary, 2020) as well as league tables and world-wide surveys into graduate recruitment to highlight the institutions that are most successful in promoting student employability. This identification of the role of universities to deliver employability reflects the concerns of students in the wake of rising education fees and increasing competition in the labor market. In the pursuit of employment a degree is no longer enough (Tomlinson, 2008) and the ability to highlight extra-curricular experiences and skills development alongside credentials are needed to stand out from the crowd (Tholen, 2013).

Definitions of employability

Employability has been conceptualized in a multitude of different ways (Forrier and Sels, 2003). As Tymon identifies, “employability is complex and multidimensional” (2013: 842). Appreciating the differing motivations of the respective actors involved

has a part to play in understanding the variegated complexity of this term, whether from students, higher educational institutions or employers (Hugh-Jones et al., 2006). Nevertheless, the discourse around employability can be understood to refer to preparing students for the labor market with skills that appropriately reflect and are relevant to their qualifications and intellect.

The ability to find employment, stay in employment and obtain new employment as and when required has typically been thought to depend on a student's human capital (Hillage and Pollard, 1998; Thijssen et al., 2008). An employable person is portrayed as successful at achieving his or her goals within employment. Employability here is understood in the context of universities developing and fine-tuning student skills to meet the needs of the labor market. This "human capital" view of the relationship between education and economic growth continues to be the heart of government policy in the UK (Troschitz, 2017).

This definition of employability is often informed by employer accounts of skill requirements. These tend to lack specificity, as generic identifications of skill requirements are difficult to establish. Further points to observe include that employers are not restricted in their assessment to simply job requirements but may also be influenced by the supply of graduates at any time. What universities actually teach is also not necessarily at the forefront of employer concerns. Several empirical studies draw attention to the point that selection and hiring practices are not simply merit-based (Jackson, 2007) nor do they reward "typical" graduate skills (Wilton, 2011). In this regard, Archer and Davidson (2008) revealed when looking at employer's perspectives that soft skills like communication, teamwork and integrity are increasingly valued as much as knowledge/subject skills.

However, the concept of employability together with its reliance on human capital theory has been subject to wide-ranging critique. The theoretical dimensions have been challenged (Tomlinson, 2010), as has the lack of attention to subjective experience (Holmes, 2001). In broader ideological terms, the concept of employability emerges in the context of a shift from an engagement with the question of employment from a structural perspective to one concerned with an agency-led address of what the prerequisites are to be employed in the current labor market. Its individualistic perspective as Brown et al. (2002) highlight fails to acknowledge that "employability is primarily determined by the labor market rather than the capabilities of individuals" (2002: 9) and thus conforms to a neoliberal view of market individualism.

In acknowledgment of the structural component of labor demand, Brown and Hesketh introduce the notion of the "duality of employability". This duality entails both an "absolute dimension" in which individual skills, knowledge and commitment and a "relative dimension" in which this is gauged in relation to "how one stands relative to others within a hierarchy of job seekers" (Brown and Hesketh, 2004: 25). In other words, the relative dimension addresses the question of supply and demand, an important consideration in relation to an environment like the UK in which the growth in supply of graduates does not match the demand in terms of graduate level jobs. Employability from their perspective should therefore be defined "as the relative chances of finding and maintaining different kinds of employment" (Brown et al., 2002: 11).

As Brown and Hesketh identify, the relative position of graduates is always susceptible to change irrespective of their ability to satisfy skill requirements. In a market in which demand outstrips supply and the reification of skills is regarded as the measure the struggle for positional advantage only intensifies further and other factors that serve to make the difference become even more significant. An example of this is where British students take opportunities to "fast-track" graduate jobs through internships to differentiate themselves from other students. Such trajectories only serve to reinforce structural inequalities since these paths are more available to students from privileged backgrounds at elite institutions.

The framing of the employability agenda in terms of university institutions, employers and students tends to overlook how gender, ethnicity, social class and disability play a part in disclosing labor market opportunities (Morley, 2001; Finnegan et al., 2020). These factors allied to inequalities relating to social and cultural capital are often lacking from policy debates on employability and educational reform. From the perspective of educational institutions, efforts to align the learning experience with the perceived demands of employers are driven by the need to evidence high rates of employment outcomes for students. These concerns however underestimate the structural inequalities reinforced in the context of promoting these ends as well as more directly within the labor market itself. As many studies show (Thijssen et al., 2008; Finnegan et al., 2020) the emphasis on skills and employment outcomes typically ignores institutional policy, social context or the macroeconomic conditions.

Student experience

In *The Student Experience*, Haselgrove (1994) highlighted the need for institutional thinking and engagement to respond to the challenges of a more diverse student body and an acknowledgment that class remains a fundamentally important reference point for understanding what it means to experience being a student. In the UK there is a growing amount of research that examines student experiences across the university sector (Archer and Hutchings, 2000; Harvey et al., 2006) and within that material that looks at working-class experiences (Tett, 2000; Reay, 2003; Reay et al., 2010; Bradley and Ingram, 2012). Bradley and Ingram (2012) found that working-class students experience more challenges and constraints than their middle-class colleagues. Finnegan et al. (2020) similarly identified that working-class students reported that experiences in HE did not cater sufficiently to non-traditional students and the outcomes in terms of employability were not equivalent, encountering high-flier employers that targeted specific universities and specific groups within those universities that tended to favor traditional students. The UK education system (including HE) reflects the broader society in being highly stratified and those from private schools, Oxford and Cambridge continue to get the top jobs, often through the "old boy's network". As the *Social Mobility and Child Poverty Commission* (2015) found recruitment to top jobs is now favoring "poshness". Non-traditional adult graduates, therefore, experience barriers in the graduate labor market.

The research

This article focuses on the learning experiences of women adult students studying for an undergraduate degree in higher education in the UK and their expectations for the future in relation to the graduate labor market. Issues of gender and class are central to this discussion as the women in our research were working-class women. We draw on the voices of women adult students from two European Union (EU) funded research projects. The first project—Access and Retention: Experiences of Non-traditional Learners in Higher Education (RANLHE)—(see Finnegan et al., 2014) involved eight partners from England (2), Germany, Ireland, Poland, Scotland, Spain and Sweden. The overall aim of this project was to look at the experiences of adult students in relation to access and retention and why some were able to “keep on going on” and succeed while others from the same socio-economic background chose to leave university without completing. The second project—Enhancing the Employability of Non-traditional Students in HE (EMPLOY)—aimed to promote the enhancement of the employability of students in higher education from a non-traditional background through improving the efficiency of transitions into the graduate labor market. The research highlighted the inequalities (class, gender, race, age) which adult students faced in transitioning from higher education to the labor market. Importantly the project took a critical stance on the concept of employability as the traditional perspective of employability is framed within a neo-liberal, marketized and individualized view of higher education and the labor market. University—employer links are now viewed as good thing for the employability of graduate students and for the needs of the labor market. In doing so the social purposes of the university is lost. This project was also funded by the EU and involved six partners from England, Ireland, Poland, Portugal, Spain, and Sweden.

Both research projects used the term non-traditional student. The research team defined non-traditional student in the following way:

Students who are under-represented in higher education and whose participation in HE is constrained by structural factors. This includes, for example, students whose family members have not been to university before, students from low income families, students from minority-ethnic groups, students living in what have been traditionally low participation areas as well as mature age students and students with disabilities.

At the core of our research was the employment of biographical methods. This will be elaborated on later but using such a method enabled the women students to tell their stories and talk about their experiences in-depth. And central to the stories told by women were issues of gender and class inequalities.

Gender and class

The widening access to higher education in the UK since the 1980s has encouraged adult women (and men) to return to education and the majority of adult students in HE have been and continue to be women. Studying for a university degree can be a powerful biographical experience for non-traditional students as it opens up a new way of looking at the world (Alheit and Dausien, 2002; West, 1996). At the same time it can also be a risky business (Reay, 2003; Barnett, 2017). The biographies, as highlighted later in this article, reveal the determination, but also the struggles of working-class women adult learners to transform their lives in some way (Skeggs, 1997; Merrill 1999, 2015; Tett, 2000), not only for themselves but also for their children. Structural constraints were lessened through the use of agency enabling the majority of women to complete their degrees in the social science disciplines.

Feminism was pivotal in making public the structural inequalities and oppression which women experience in society. Feminism stimulated the development of feminist theory and research in sociology in the 1970s and later in adult education and in doing so critiqued and challenged traditional “malestream” sociology (Delamont, 2003). It questioned who has the power to construct knowledge. Feminist research brought women out of obscurity emphasizing that the “personal is political”. As Reinharz explains:

Making the invisible visible, bringing the margin to the centre, rendering the trivial important, putting the spotlight on women as competent actors, understanding women as subjects in their own right rather than objects for men.

Reinharz (1992: 248).

From the 1970s onwards in the UK feminist adult educators focused on women’s education for social purpose, social justice and transformation through a curriculum of women’s studies which examined the nature of women’s oppression through “really useful knowledge” (Johnson, 1988). Feminist popular adult education is seen as a site of struggle (Barr, 1999; Thompson, 2000), working with women in oppressed communities while research largely focused on working-class women’s experiences in HE. Recently research has explored the intersection between gendered and classed experiences in both further and higher education (Skeggs, 1997; Tett, 2000; Reay, 2003; Merrill, 2015).

Gender cannot be looked at in isolation as it is not the only form of inequality experienced by the women we interviewed. The women did not see themselves as being just women as they described and defined themselves as working-class women. Being

a woman and being working-class cannot be separated. Class as well as gender, and for some race, were subjectively and emotionally perceived as defining their identity and who they are. This resonates with the women students studying in further education who Beverly Skeggs interviewed. As Skeggs asserts “... the category ‘woman’ is always produced through processes which include class and classifying produces very real effects which are lived on a daily basis” (1997: 2). She continues “The women never see themselves as just women: it is always read through class” (1997: 91). People, therefore, experience multiple forms of inequality in society. The concept of intersectionality introduced by Crenshaw (1989) which first linked gender and race to understand black women’s lives was later extended to include other forms of inequality such as class. Anthias summarizes this in the following way by stating that: “classes are always gendered and racialized and gender is always classed and racialized and so on” (2005: 33). As the women’s stories reveal gender and class are inextricably linked in terms of their identity and daily lives.

Class

Class and class inequality are fundamental to understanding UK society, including, and importantly, the education system in which class inequalities are reproduced. Class structures the everyday lives of people through economic, social and cultural inequalities and access to resources. As Andrew Sayer reminds us class “affects how others value us and respond to us, which in turn affects our sense of self-worth” (2005: 1). The study of social class—once dominant in UK sociology—went into retreat during the 1980s and 1990s due to several factors. Sociological discourse increasingly focused on individualism and also on postmodernism and its dismissal of grand narratives. Challenges also came from feminists who criticized the use of male occupation in a family as a way of defining a person’s social class. While sociologists like Pahl (1989) dismissed class as an outdated concept for characterizing people’s lives other sociologists began to re-introduce and re-assert class into the discipline such as Skeggs (1997), Savage (2000), Sayer (2005) and within adult education through the work of Jane Thompson, Ian Martin, Rennie Johnston and others. As Skeggs poignantly stated:

To abandon class as a theoretical tool does not mean that it does not exist any more: only that some theorists do not value it ... Retreatists either ignore class or argue that “class is an increasingly redundant issue”.

Skeggs (1997: 6,7).

The new debates and theories on class took into account the changed socio-political situation to include social, cultural, and symbolic dimensions of class (Bourdieu, 1984) while recognizing that ownership of the means of production and economic capital still continue to be important aspects of social power. Feminists also contributed to the discussion on the changing approach to class by pointing out that gender interacts with class modifying class inequalities.

Class is both a subjective and objective experience located in a historical specificity. While social, economic, political and cultural structures identify a person’s social class people, individually and collectively, are subjectively conscious of class in everyday interaction. Through agency a person may act upon their objective position to make changes to their lives. In our research we were interested in how both class and gender shapes identities in the context of higher education and how “class shapes embodied experience and practical sense” (Finnegan and Merrill, 2017: 309).

In our analysis of class and adult students’ experiences in HE we draw on the work of Bourdieu while also acknowledging that other scholars have used Bourdieu in similar research contexts such as Skeggs (1997) and Reay (2003). Bourdieu’s work is useful and inciteful to us because of his work on the social reproduction of class and identity, habitus, capitals, dispositions and fields (in this case university) as well as his study on universities—*Homo Academicus* (1988). The concept of habitus allowed us to explore how the social and cultural capitals of working-class women students enables them to cope and adjust to or not to the symbolic and intellectual capitals of the university. The context of the research for both projects was an elite university whose culture was different from that of the women students. Middle-class students because of their higher levels of cultural and social capitals, as Bourdieu elicits, are more likely to be “fish in water” while working-class students may find themselves as being “fish out of water”. (And some of the women in our studies did refer to themselves as being fish out of water). A person’s habitus is a culture which shapes how they view the world, their behavior, language and lifestyle. Habitus is linked to disposition whereby through socialization in the family and school, for example, a person learns how to present themselves within a particular habitus in everyday life. As other researchers in our project explain:

Bourdieu’s ideas about forms of capital and the ways they operate in specific fields offer a tool for understanding the enduring impact of social inequality in students’ lives. Stories are clearly formed by lived experiences of social power.

West et al. (2013: 122).

Habitus and the levels of cultural and social capitals locates a person’s position or “a sense of one’s place” (Bourdieu, 1990: 131) in relation to others and the social structure. While habitus reproduces and determines an individual’s class position and social trajectory Bourdieu acknowledged that it is not necessarily static and can be transformed through agency:

Habitus is not the fate that some people read into it. Being the product of history, it is an open system of dispositions that is constantly subjected to experiences and therefore constantly affected by them in a way that either reinforces or modifies its structures.

Bourdieu and Wacquant (1992: 133).

A note on methodology: biographical approaches

Biographical approaches were at the heart of our research (for all partners) as, in the tradition of feminism, we wanted to place the women at the center of our research by giving them voice, enabling them to tell their stories as:

... such methods offer rich insights into the dynamic interplay of individuals and history, inner and outer worlds, self and other. We use the word "dynamic" to convey the of human beings as active agents in making their lives rather than being simply determined by historical and social forces.

Merrill and West (2009: 1).

Biographical research importantly encompasses a humanistic and subjective approach to research (Plummer, 2001). Within European adult education research biographical methods are popular (West et al., 2007) fueled by the "biographical turn" (Chamberlayne et al., 2000). It marked a turn away from positivism and quantitative research which reduced and dehumanized people to numbers and statistics. The UK team drew heavily on feminist approaches to biographical research as well as symbolic interactionism.

Feminist research challenged malestream sociology by stressing that women's lives are also worthy of research (Smith, 1987) and in doing so questioned who has the power to construct knowledge. For Reinharz "biographical work ... draws women out of obscurity, repairs the historical record ..." (1992: 126). Challenging gender inequality is at the core of feminist research. Biographical research in general connects the past, present and future, the private and the public and the macro, meso and micro. We embrace the work of C. Wright Mills (*The Sociological Imagination*, 1957, Reprint in 1971) in which he links, biography, history, social structures and individual agency through the concept of the sociological imagination.

A key contribution by feminists to biographical research lies in their approach to interviewing. In giving voice to women interviews have the potential to be powerful and empowering enabling their story to be heard and for someone to take interest in it. Such interviews are both dialogical and creative. Importantly biographical interviews are a form of learning and a social process. Oakley (1981) stresses that interviews should be more like a conversation thus breaking down, as far as possible, the power relations between the researcher and the researched, and lead to building trust and a democratic relationship. The interview process is both a subjective and inter-subjective one so that the researcher is also part of the story. As Stanley and Wise state: "Personhood cannot be left out of the research process ... We see the presence of the researcher's self as central in all research" (1993: 161).

For us biographical interviews generate thick description or a "good story". "By good stories we mean narrative material that is both rich in detail but also experientially inclusive and reflective in character" (Merrill and West, 2009: 113). At first sight a biographical interview appears to be individual but in reading all the stories shared experiences of gender and class, for example, are highlighted as well as the collectivities (Merrill, 2007). Biographies also reveal the interaction of structure and agency in people's lives although at certain times one may be more dominant than the other. The stories of the working-class women we interviewed in HE illuminate the role of agency and structure in shaping their learning experiences and identities. Agency was used in the act of deciding to study at university but once in the system constraining factors, for some, such as lack of finance took over impeding their studies. Experiencing, coping and drawing on one's biographical experiences is a biographical learning experience, or as Alheit and Dausien term it "biographicity". We learn from our past biography and learning is part of our biography so as Alheit and Dausien assert "without biography there can be no learning, without learning there is no biography" (2002: 15). Telling life stories are a powerful tool for raising critical consciousness about the role of inequalities such as gender and class. The next section explores and shares the stories of the working-class women we interviewed and the centrality of the intersection of gender and class in their biographies. In order to understand present lives and hopes for the future it is essential to look at early life influences such as family and school in order to understand why they chose to return to education as an adult. We then explore their experiences of learning in an elite university and their transition into the graduate labor market and their expectations, emotions and hopes tied up with these processes.

In the RANLHE project different cohorts of students were interviewed: those in their final year, those who left but later returned, those that dropped out as well as following through a cohort from their first to final year of study (100 students in total). With the EMPLOY project a smaller sample of adult students were interviewed—40 in total. Thirty of these were undergraduate students and the other ten were graduate students. A small sample of the undergraduate students were interviewed at the end of their degree course and some for a third time a year after they left university. Most women were in their thirties or forties. All except one woman were studying on a Social Studies degree course designed for local adult students.

Experiencing HE: gender and class

Class and gender are dominant aspects of a person's identity of how they see their self and their self in relation to others and in society. Some sociologists have argued that collective identity, (an aspect of the UK working class), has been replaced by

an individual identity (Beck, 1992; Giddens, 1991). Yet others argue that there is always a collective dimension to identity (Ferguson, 2009; Jenkins, 2000) as it is shaped by personal, social and institutional processes. Berger and Luckmann also asserted that “identity is a phenomenon that emerges from the dialectic between individual and society” (1973: 195). The stories we collected, although individual, illuminated shared classed and gendered experiences which in turn shaped identities of themselves as non-traditional students. A person’s identity is linked to biography as it is shaped and constructed through past and present experiences (Ferguson, 2009). Working-class students imagine a future of who they want to be and become. A university is a transitional space between past and future lives. It is also a temporary and bounded space and place:

A transitional space, therefore, implies being in-between in terms of identity whereby an individual lets go of part or all of their old identity and changes to assume a new or modified identity.

Merrill (2012: 28).

Early life experiences

In both research studies many of the women were of the generation who experienced a very gendered education system at school. Working-class girls were expected to leave school at the earliest possible age and transition into a job before getting married and having a family (Spender, 1989). Some were resentful with their parents and teachers for not encouraging girls with their education:

... it wasn’t until later that I felt quite resentful about the experiences I had at school ... I just felt that there was a lot of potential that I had that was totally wasted because assumptions were made about me. Too young at the time to know but I do feel it came back to my background and my family and where I lived and that influenced how they treated me and that’s why college never got mentioned.

Paula.

Cathy despised her father for treating her differently to her brothers in relation to education. She said “I am still bitter toward my father for not giving me the opportunity because both my brothers were”. Sarah had wanted to stay on at school after the age of 16 (first possible leaving age) but accepted her fate as a girl:

My dad got me a job in a hairdresser’s and that’s where I sort of stayed because I’d wanted to be a hairdresser when I was seven so that means—you want to be a hairdresser for the rest of your life.

Soon after leaving school she had a child and continued working as a hairdresser but experienced a restless desire to return to education and achieve.

Going to university was an opportunity to prove to themselves that they could achieve educationally, and move away from the gender and class restrictions of their childhood. This was important to Sue, who as a single parent, wanted to obtain a better life for her and her children:

I did my degree to prove something to myself. To get a degree that I should have got when I was younger. So to prove to myself that I can get that qualification and for my children as well so that they would think it’s the norm.

Other women, such as Mary, were bored with doing jobs which were unskilled and unfulfilling. Mary worked in a chip shop and could have had promotion to chief fryer. This was a turning point moment (Denzin, 1971) as she thought “no thanks and I thought I need to leave this place because otherwise I’m going to be here when I’m their age” (other women workers).

For some women returning to education was an escape from an unequal relationship in the home and domesticity. Entering HE was a way out of this situation, doing something for themselves and the act of doing this led, in some cases to divorce. The decision to return to education involved reflecting back on their past and present lives and actively wanting to change their lives for the better:

Through agency they were reflecting, constructing and re-constructing their biographies to adjust to transitions and the changing realities of their private and public lives and the structural constraints experienced in communities. It is both a personal journey and sometimes a collective one with other women, in which they reflect upon their lives as women in the public and private domains struggling for change and transformation.

Merrill (2003: 135).

Experiencing university life

The women were studying at an elite university and as adult working-class women they found themselves in a minority. Class consciousness was raised through interaction with lecturers, middle-class students, academic language and knowledge and institutional ethos and culture. In elite universities working-class students “stick out” and many become aware that they are the “other” so that: “Working-class students are, therefore, very often arriving at quite traditional universities and possess different cultural capital and habitus to those of middle-class students and institutions” (Finnegan and Merrill, 2017: 14). One woman used Bourdieu’s term and described herself as “being fish out of water” while Julia stressed that in seminars “I definitely felt like I was in a different class to most of the people I was learning with—that was quite clear”.

However, the fact that it was an elite university was one of the reasons why they applied to the university (there is another university in the city) as they felt its status would help them in the graduate labor market. Sharon, for example, applied to study law and on being accepted stated:

Yeah it was brilliant, it was the best thing ever. I was just so happy that I got into X because I didn’t want to go anywhere else and I just felt so grateful. I was like “oh my God I can’t believe I’ve got into the top five universities from coming from nothing”.

Sharon had also applied to the other university in the city as a fall back: “I didn’t want to go to Y anyway but I thought if I don’t get in to X I need somewhere to go”. She had been brought up by foster parents since the age of twelve after her father murdered her mother but became homeless at the age of 16 and a single mother at the age of 17. She began her degree in her late thirties and before then worked in various jobs and also studied to get qualifications including an Access course (an Access course prepares adults for entry to university) as she was determined to change her life. The choice of a law degree stemmed from her childhood experiences.

Language and vocabulary were key issues in relation to class and the experience of learning. Anne felt herself to be different from the younger middle-class students.

I find it hard and I’m really sort of aware of the way—how different I speak. They’ve got amazing vocabularies and you just think, God you sound brilliant and then I just go red. I just can’t get my words out. I just lack confidence I suppose.

For Jenny seminars were also a situation where she lacked confidence but devised a strategy to try to deal with it:

I cannot do that and made worse by the fact that these young people in the class were very intimidating and I’m absolutely sure they didn’t mean to be. They were upper class. I used to sit and look at them and think well you know I’ve got so much experience that you haven’t got.

Gender and class

Experiences at university were also gendered and classed, as Mary’s quote illustrates, the lives of the younger and mature students were completely different:

They come from different backgrounds and that’s when I found I couldn’t speak. In the module Politics and Food, we (the mature students) were talking about school dinners and were saying that they had to be good because this is the only hot meal they get and they couldn’t understand that. Daddy had bought them ponies and daddy had got them this car.

Unlike male adult students the women had to juggle and cope with undertaking multiple roles which often made finding time to study difficult (Merrill, 1999). They had to try to balance the different roles and this sometimes meant not being able to put themselves first. Some felt guilty about not being “a good mother”. Others studied late at night after the children had gone to bed. Mary a single parent (divorced) with two children explains:

Whenever I can. I study in the evening because my oldest is always on his Xbox and my younger lad plays out ... they’re good kids and at weekends they go to their dad’s every second weekend so that’s good because I get a complete weekend if I want to go to the library or ...

Sue also talked about the problem with time since beginning her studies:

It’s always a challenge with a little one (child) ... I had to work out when it was best for me to study and just all the other stuff that comes with being an adult with education ... Not when do I spend but what time do I spend with Pete and Emma and how much time do I put in, how much time do I put to washing up and ironing and hovering?

Although wanting to do a degree she had, like others, self-doubt about her ability to study at this level:

It's that whole fear of not being able to actually do it. I can't do a degree. I'm not clever enough and can I really go without a full wage now? I guess just all the fears when you get older and realize what you could lose and what you could gain and then obviously with family as well. I was six months on (pregnant) when I started the degree so it was like if I'm going to do this it has to be now. I had to know that I had the support of Pete like a 150% to get through this. I can quite definitely say if I hadn't had his support I probably wouldn't still be sat here right now. There have been times when I've gone "I just can't do this anymore".

Sue.

Sue was unusual in having the support of her husband. Being a student and a single parent brought other difficulties and obstacles to completing a degree:

Being a lone parent is very, very difficult—not having the support, the family support at home and finances again very, very difficult. If you're worrying about silly things like whether you've got enough petrol to drive into university or stuff like that then that really hinders your learning so much. It did get to the stage where I was struggling to even pay the utility bills and buy the food and it seemed ridiculous. I was having an education, having a university education but I just couldn't afford to.

Julie.

Age is also a factor which results in mature students experiencing university life differently as Sue found in her sociology seminars but not always in a negative way:

It was a bit weird at first because obviously I'm the old one in the room and I think some people occasionally thought I was the lecturer rather than a student but actually just through seminars they were actually really encouraging to me. The amount of people who were like you're a wife, and a mother and a student, how on earth do you do this?

Studying at university can distance them as an individual from their family and friends in terms of class and how others view them as they see that person as "getting above their station". Julie who lived on a council estate (social housing) in a deprived area of the city encountered disapproval from her family and friends:

I couldn't talk about university to other mums as I walked to school because I felt that they had snubbed me—"Look at you with your big briefcase". It was really tricky to explain to people why I wanted to progress myself. I withdrew from my friends on the estate a little because I found it hard to explain to them why I wanted something different but I understood why they didn't or why they were happy in the situations they were in.

Julie's family also distanced her as they could not understand why she wanted an education. In UK working-class terms she was "getting above herself" and being disloyal to her class in wanting to "better herself". This reaction by her family was very emotional to her. As Hoggart (1957) explains it is an isolating experience. Although she was "getting out and getting away" (Lawler, 1999) Julie continued to subjectively define herself as being working-class even after she obtained her degree. She had no desire to become middle-class as the women in Lawler's study yet she was caught between two cultures—her working-class one and her academic one. Being working-class was embodied in her identity. She had "got out" but only a little bit. While she had observed the middle-class world at university she did want to embrace it.

Transitions and the graduate labor market

Transitioning to the labor market brought about a different set of experiences for the women making them aware of inequalities in relation to gender, class and age. However, they did feel that having a degree from a top university gave them an edge over adults from other universities as Sharon stated earlier. In her second interview she felt it was a reason that she got the job as a PA in a law firm although it was a lower level job than Sharon had hoped (she got a low level mark for her degree):

Putting X on my CV got me the job because the first thing they said in the interview was "X—that's a very good university. How did you get in there? What did you do? And that basically got me the job."

She elaborated further that she was accepted by the solicitors at the firm because of the university she went to as she was asked

What university did you go to? And whenever you say X they gave an exclamation of surprise ... "Okay", it's as if you are in a different light all of a sudden. Because they had all been educated at Eton and Oxford and Cambridge so they felt I fitted in because I'm from X. If I had told them I had been to Y (the other university in the city) I'd have been cut out.

Although she was accepted there was also an element of surprise that a working-class women student had studied at X.

Increasingly employers are expecting students to gain work experience before employment. Most placements are unpaid and take place during the university vacation. For the women in our study this was not possible because of childcare responsibilities or the need to earn money and often the placements are not local:

Most of the graduate schemes that are available are for people without responsibilities. I can't go travelling and leave my son at home ... There doesn't seem to be any niche for graduate schemes that are solely based locally. The jobs I've looked at, they're looking for people that are flexible and are willing to work all the hours god sends and I have commitments but that isn't taken into account for the mature student in the job market

Jane.

Gaining relevant work experience is, therefore, difficult and this puts women at a disadvantage on the job market:

And I find with some places now, especially where I'm looking at, they want the qualifications but they also want the experience to go with it and they're saying you need ... How on earth do you try and get both at the same time as well as everything else you've got to do and that's what I find highly frustrating sometimes.

Sue.

There was a concern also identified that employers are more likely to favor younger graduates:

Yet at the end of this year I will have a degree but looking at my age and looking at the students—so many young students with the same degree as me when it comes to employment. Employers—maybe they will say “Yes you have your degree expertise but your age” and would rather be looking at someone younger than me.

Another student identified that within the legal profession examples of social and cultural capital were pervasive in securing employment:

Some of the students I've spoken to their parents are partners in solicitors so obviously they're going to walk into a job aren't they? Definitely down the barrister route—it's definitely about what private school you've been to. I think money definitely because if you haven't got money you're just not going to the Bar really are you. It costs too much money. Their course is £18,000 ... it's just not an option

Sharon.

In the context of these challenges students were keen to stress the need for employers to appropriately value the qualities that they bring as mature learners to an organization:

So I think a lot of employers must realize that non-traditional students bring something different to the table, so we've got life experiences as well.

These broader “life experiences” are often complemented by work experience that evidences their commitment and reliability:

I think the bonus about being a non-traditional student—well for me anyway—I've had the experience of work for quite a number of years so already on my CV it looks like employers can see that I am hard working and I have worked so I think that's already a bonus.

Sue.

However, the commitment to family and place restricted many students from being able to offer the so-called “flexibility” that many employers seem to value:

Friends of mine have done English degrees and (one of them went to work in the English quarter in America he's) travelled the world on the basis of that degree ... Non-traditional students don't have the luxury to do things like that. It's like the internships and stuff like that. I could have got a job with UPS on an internship, fast-tracked through to management but that means doing 6 months say in Northumberland and 6 months in Scotland and then maybe a month in America and I can't do that

Jane.

Summary

As stated earlier higher education is a transitional and temporary space in which identities are re-worked and re-shaped as a result of the educational experience. Despite the feeling of being the other in the middle-class university environment there was also a sense of achievement even though there were struggles on the journey but also belonging in some way:

I can't explain that feeling of not just pride but about being part of the University, that you feel part of something ... I'd actually come to the end of something quite brilliant and I should feel really, really proud.

Julie.

Often the process of “becoming educated”—of adapting and succeeding—can later raise issues of belonging in relation to the self, family and friends’ (Finnegan and Merrill, 2017, 320).

The women’s voices above offer an insight into their experiences of learning as a working-class woman in HE and beyond into the labor market and how this was shaped by issues of gender and class. Our research indicates that studying does bring personal, social, economic and academic benefits and can change the lives of women in some way, despite being sometimes a risky business (Reay, 2003). Importantly universities act as a transitional space enabling women to reflect upon their identity as a working-class woman and reshape and transform themselves. The motivations for entering HE centered around “doing something for myself”, finishing their education and proving to the self that they are capable of studying. There is also the desire for a better life in material and employment terms. This was particularly the case with the lone parents in the research as they wanted to improve the quality of life for their children and move out of the deprived areas of the city. The women wanted more interesting jobs which will engage and fulfill them. Their stories, while individual, were also collective highlighting similar experiences as well as illustrating the significance of the interaction of agency and structure in everyday life.

Although universities are now more diverse universities are not a level playing field as they are a site where inequalities are experienced and reproduced. Significant cultural, social and financial differences are experienced by adult students. Their cultural capital and habitus is different from that of middle-class students and the institution (Bourdieu, 1984). Working-class women have to learn to adopt and cope to avoid becoming a “fish out of water”. Issues of inequalities persist into the graduate labor market where they experience gender, class and age discrimination even if they have a degree from a top university. Despite these inequalities the women valued the education they received and had no regrets in studying for a degree. While it may not have guaranteed them upward mobility (Tomlinson, 2008) the experience had nonetheless enriched them. But one aspect of their lives remained constant—defining themselves as working-class women. Universities and employers, however, need to recognize and address the inequalities which adult students face in order to offer a more level playing field.

References

- Alheit, P., Dausien, B., 2002. The double face of lifelong learning: two analytical perspectives on a silent revolution. *Stud. Educ. Adults* 34 (1), 3–22.
- Anthias, F., 2005. Social stratification and social inequality: models of intersectionality and identity. In: Devine, F., Savage, M., Scott, J., Crompton, R. (Eds.), *Rethinking Class: Cultures, Identities and Lifestyle*. Palgrave Macmillan, New York, pp. 22–45.
- Archer, W., Davison, J., 2008. Graduate Employability: The Views of Employers. Council for Industry and Higher Education, London.
- Archer, L., Hutchings, M., 2000. “Bettering Yourself?” Discourses of risk, cost and benefit in ethnically diverse, young working class non-participants’ constructions of higher education. *Br. J. Sociol. Educ.* 21 (4), 555–574.
- Barnett, R., 2017. *Thinking and Rethinking the University*. Routledge, London.
- Barr, J., 1999. *Liberating Knowledge: Research, Feminism and Adult Education*. NIACE, Leicester.
- Beck, U., 1992. *Risk Society: Towards a New Modernity*. Sage, London.
- Berger, P.L., Luckmann, T., 1973. *The Social Construction of Reality*. Penguin, Harmondsworth.
- Bourdieu, P., Wacquant, L.J.D., 1992. *An Invitation to Reflexive Sociology*. University of Chicago Press, Chicago.
- Bourdieu, P., 1984. The forms of capital. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, New York, pp. 241–258.
- Bourdieu, P., 1988. *Homo Academicus*. Policy Press, Cambridge.
- Bourdieu, P., 1990. In *Other Words: Essays Towards a Reflexive Sociology*. Polity, Cambridge.
- Bradley, H., Ingram, N., 2012. Banking on the future: choices, aspirations and economic hardship in working-class student experience. In: Atkinson, W., Roberts, S., Savage, M. (Eds.), *Class Inequality in Austerity Britain: Power, Difference and Suffering*. Palgrave Macmillan, Basingstoke, UK, pp. 51–69.
- Brown, P., Hesketh, A., 2004. *The Mismanagement of Talent: Employability and Jobs in the Knowledge Economy*. Oxford University Press, Oxford.
- Brown, P., Hesketh, A., Williams, S., 2002. *Employability in a Knowledge-Driven Economy*. Working Paper Series, Paper 26. Cardiff University.
- Chamberlayne, P., Bornat, J., Wengraf, T., 2000. *The Turn to Biographical Methods in Social Science: Comparative Issues and Examples*. Routledge, London.
- Crenshaw, K., 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics, vol. 14. *University of Chicago Legal Forum*, pp. 538–554.
- Delamont, S., 2003. *Feminist Sociology*. Sage Publications, London.
- Denzin, N.K., 1971. The logic of naturalistic inquiry. *Soc. Forces* 50.
- Ferguson, H., 2009. *Self-Identity and Everyday Life*. Routledge, London.
- Finnegan, F., Merrill, B., 2017. “We’re as good as anybody else”: a comparative study of working-class university students’ experiences in England and Ireland. *Br. J. Sociol. Educ.* 38 (3), 307–324.
- Finnegan, F., Merrill, B., Thunborg, C., 2014. In: *Student Voices on Inequalities in European Higher Education: Challenges for Theory, Policy and Practice in a Time of Change*. Routledge, London.

- Finnegan, F., Merrill, B., Revers, S., 2020. "When it comes to what employers are looking for, I don't think I'm it for a lot of them": class and capitals in, and after, higher education. *Stud. High Educ.* 45 (1), 163–175.
- Forrier, A., Sels, L., 2003. The concept employability: a complex mosaic. *Int. J. Hum. Resour. Dev. Manag.* 3 (2), 102–124.
- Giddens, A., 1991. *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Polity, Cambridge.
- Harvey, L., Drew, S., Smith, M., 2006. *The First Year Experience: A Review of the Literature for the Higher Education Academy*. HEA, York.
- Haselgrove, S. (Ed.), 1994. *The Student Experience*. Society for Research into Higher Education and Open University Press, Buckingham.
- Hillage, J., Pollard, E., 1998. *Employability: Developing a Framework for Policy Analysis*, Research Brief RR85. Department for Education and Employment, Nottingham.
- Hoggart, R., 1957. *The Uses of Literacy*. Chatto & Windus, London.
- Holmes, L., 2001. Reconsidering graduate employability: the "graduate identity" approach. *Qual. High Educ.* 7 (2), 111–119.
- Hugh-Jones, S., Sutherland, E., Cross, A., 2006. *The Graduate: Are We Giving Employers What They Want?* Paper Presented at the Teaching and Learning Conference, January, 6, Leeds.
- Jackson, M., 2007. How far merit selection? Social stratification and the labor market. *Br. J. Sociol.* 58 (3), 367–390.
- Jenkins, R., 2000. Categorization: identity, social process and epistemology. *Sociology* 48 (3), 7–25.
- Johnson, R., 1988. "Really useful knowledge" 1790–1850: memories for education in the 1980s. In: Lovett, T. (Ed.), *Radical Approaches to Adult Education: A Reader*. Croom Helm, London.
- Lawler, S., 1999. Getting out and getting away: women's narratives of class mobility. *Fem. Rev.* 63, 3–24.
- Merrill, B., West, L., 2009. *Using Biographical Methods in Social Research*. Sage, London.
- Merrill, B., 1999. *Gender, Change and Identity: Mature Women Students in Universities*. Ashgate, Aldershot.
- Merrill, B., 2003. Women's lives and learning: struggling for transformation. In: Dybbroe, B., Ollagnier, E. (Eds.), *Challenging Gender in Lifelong Learning: European Perspectives*. Roskilde University/ESREA, Roskilde.
- Merrill, B., 2007. Recovering class and the collective in the stories of adult learners. In: West, L., Alheit, P., Andersen, A.S., Merrill, B. (Eds.), *Using Biographical and Life History Approaches in the Study of Adult and Lifelong Learning: European Perspectives*. Peter Lang, Frankfurt-am-Main.
- Merrill, B., 2012. Learning to become an adult student: experiences in a UK university. *Int. J. Contin. Educ. Lifelong Learn.* 5 (1), 21–42.
- Merrill, B., 2015. Determined to stay or determined to leave? A tale of learner identities, biographies and non-traditional students in higher education. *Stud. High. Educ.* 40 (3), 1859–1871.
- Molesworth, M., Scullion, R., Nixon, E., 2011. *The Marketization of Higher Education and the Student as Consumer*. Routledge, London.
- Morley, L., 2001. Producing new workers: quality, equality and employability in higher education. *Qual. High Educ.* 7 (2), 131–138.
- Oakley, A., 1981. Women interviewing women: a contradiction in terms. In: Roberts, H. (Ed.), *Doing Feminist Research*. Routledge, London, pp. 30–62.
- O'Leary, J., 2020. *The Times Good University Guide 2021*. Times Books, London.
- Pahl, R.E., 1989. Is the emperor naked? *Int. J. Urban Reg. Res.* 13, 711–720.
- Plummer, K., 2001. *Documents of Life 2: An Invitation to a Critical Humanism*. Sage Publications, London.
- Reay, D., Crozier, G., Clayton, J., 2010. "Fitting in" or "standing out": working-class students in UK higher education. *Br. Educ. Res. J.* 36 (1), 107–124.
- Reay, D., 2003. "A Risky business" mature working-class women students and access to higher education. *Gend. Educ.* 15 (3), 301–317.
- Reinharz, S., 1992. *Feminist Methods in Social Research*. MacMillan, London.
- Savage, M., 2000. *Class Analysis and Social Transformation*. Open University Press, Buckingham.
- Sayer, A., 2005. *The Moral Significance of Class*. Cambridge University Press, Cambridge.
- Skeggs, B., 1997. *Formations of Class and Gender*. Sage, London.
- Smith, D., 1987. Women's perspectives as a radical critique of sociology. In: Harding, S. (Ed.), *Feminism and Social Theory*. Open University Press, Milton Keynes, pp. 84–96.
- Social Mobility and Child Poverty Commission, 2015. *State of the Nation 2015 Report*. HMSO, London.
- Spender, D., 1989. *Invisible Women: The Schooling Scandal*. The Women's Press, London.
- Stanley, L., Wise, S., 1993. *Breaking Out Again: Feminist Ontology and Epistemology*. Dover Publications, New York.
- Tett, L., 2000. "I'm Working class and proud of it"—gendered experiences of non-traditional students in higher education. *Gend. Educ.* 12 (2), 183–194.
- Thijssen, J.G.L., Van der Heijden, B.I.J.M., Rocco, T., 2008. Toward the employability link model: current employment transition to future. *Hum. Resour. Dev. Rev.* 7, 165–183.
- Tholen, G., 2013. The social construction of competition for graduate jobs: a comparison between Great Britain and The Netherlands'. *Sociology* 47 (2), 267–283.
- Thompson, J., 2000. *Women, Class and Education*. Routledge, London.
- Tomlinson, M., 2008. The degree is not enough: students' perception of the role of higher education credentials for graduate work and employability. *Br. J. Sociol. Educ.* 29 (1), 49–61.
- Tomlinson, M., 2010. Investing in the self: structure, agency and identity in graduates' employability. *Educ. Knowl. Econ.* 4 (2), 73–88.
- Troschitz, R., 2017. *Higher Education and the Student: From the Welfare State to Neoliberalism*. Routledge, London.
- Trow, M., 1989. The Robbins trap: British attitudes and the limits of expansion. *High Educ. Q.* 43 (1), 35–75.
- Tymon, A., 2013. The student perspective on employability. *Stud. High Educ.* 38 (6), 841–856.
- Walters, W., 2012. *Governmentality: Critical Encounters*. Routledge, London.
- West, L., Alheit, P., Andersen, A.S., Merrill, B. (Eds.), 2007. *Using Biographical and Life History Approaches in the Study of Adult and Lifelong Learning: European Perspectives*. Peter Lang, Frankfurt-am-Main.
- West, L., Fleming, T., Finnegan, F., 2013. Connecting Bourdieu, Winnicott and Honneth: understanding the experiences of non-traditional learners through an interdisciplinary lens. *Stud. Educ. Adults* 45 (2), 119–136.
- West, L., 1996. *Beyond Fragments: Adults. Motivation and Higher Education*. Taylor & Francis, London.
- Wilton, N., 2011. Do employability skills really matter in the graduate labor market? The case of business and management graduates. *Work. Employ. Soc.* 25 (1), 85–100.
- Wright Mills, C., 1957. *The Sociological Imagination*. Penguin, Harmondsworth.

Deconstructing predominantly White cultures: creating inclusive learning environments in the United States

Ibrahim M. Karkouti, The American University in Cairo, New Cairo, Egypt

© 2023 Elsevier Ltd. All rights reserved.

Deconstructing predominantly white cultures	417
Background of the study	418
Conceptual framework	418
Research questions	418
Methodology	419
Participants and setting	419
Data analysis	419
Limitations	420
Representative preliminary findings	420
Understanding culture: acknowledging identity, values, and beliefs	420
Learning about other cultures: A small step brings you closer	420
Recognizing and deconstructing white culture: overcoming oneself breaks cultural barriers	420
Recognizing the legitimacy of multiple cultures: Acceptance fosters coexistence	421
Developing a multicultural outlook: rethinking worldview nurtures democracy	421
Conclusion and recommendations	421
Significance of the study	422
Acknowledgments	422
References	422

Deconstructing predominantly white cultures

Student affairs practitioners face many hurdles as they try to create inclusive learning environments conducive to student academic success and psychological wellbeing. These barriers include role strain (Goode, 1960), macroaggressions (Sue et al., 2009), and *onlyness*, which means being the only individual performing the job from the one racial and ethnic group (Harper, 2013). In order to cope with these challenges, student affairs practitioners find themselves caught between a rock and a hard place to meet their job expectations in predominantly White cultures. Specifically, student affairs practitioners wear a white mask as a survival mechanism in order to meet pre-existing professional standards in predominantly White universities (Nguyen and Duran, 2018). Nevertheless, whiteness or masquerading as White is emotionally burdening for people of color trying to enhance diversity on campus (Matias, 2015).

According to Nguyen and Duran (2018), whiteness is defined as “an ideology wherein society considers dominant cultural norms as neutral, and anything outside of those standards is *othered* – seen as deviant, strange, or unusual” (p. 114). This means that whiteness is an already existing social construct that refers to a privileged group who has power and influence over others in terms of social, political, and economic decision making (Cabrera, 2014). In today’s workplace, whiteness is deeply rooted and manifested covertly (Nguyen and Duran, 2018).

In order to combat whiteness and deconstruct predominantly White cultures, Nguyen and Duran (2018) recommended a number of strategies that include the following:

1. Creating an inclusive community: Student affairs practitioners should establish common ground with culturally different others to process the impact of whiteness. In doing so, stakeholders of diverse racial and ethnic backgrounds learn about different cultures and meet halfway.
2. Promoting dialogue: Student affairs practitioners should encourage people of color to share their experiences with their White counterparts to adjust existing preconceived notions about minority groups. This form of dialogue gives voice to marginalized stakeholders through counter-story-telling that not only acknowledge the legitimacy and authenticity of people of color, but also helps them “reframe the dominant narrative about professionalism and whiteness in student affairs” (p. 119).
3. Encouraging cross-racial friendships: In addition to engaging people of color in dialogue, student affairs practitioners are required to promote a culture of friendship whereby all students create a healing space of validation and understanding that ultimately enhances racial climate and diversity on campus.

Next, a presentation of the preliminary findings of a qualitative, phenomenological case study that explored student affairs practitioners’ perceptions of the strategies they implement to deconstruct predominantly White cultures and create racially inclusive environments that promote equal representation of minority students and enhance their learning.

Background of the study

The role of student affairs practitioners is to create diverse learning environments that provide students with cross-racial engagement opportunities, enhance their learning experience, and expose them to new learning paradigms that could interrupt long-standing segregation trends in the society. In an attempt to improve student affairs practitioners' multicultural competence and promote equal representation of minority students, many universities and colleges developed and embarked upon multiple diversity programs and awareness initiatives. Nevertheless, these initiatives were not successful in transforming segregated learning environments into inclusive ones because institutions of higher learning often fail to transform their espoused values into enacted ones (Karkouti, 2016; Manning and Coleman-Boatwright, 1991). According to Baxter-Magolda (2003), predominant White cultures explain why student affairs practitioners' efforts often fall short in transforming their institutional cultural climate. These cultures are persistent in that they serve as a pseudo-transformational force that intensifies discrimination and racism, undermines public and private spheres including postsecondary education, and widens the gap between different social classes.

White culture is believed to dominate all facets of campus life by defining what constitutes acceptable practices, behaviors, and principles (Karkouti, 2016; Manning and Coleman-Boatwright, 1991). For example, personal and professional qualities that differ from the established norm are rejected and considered unusual, intrusive, and unacceptable. Therefore, minority students become subject to alienation in predominantly White institutions (PWIs) where their way of being, identity, and culture is disregarded in the institutional mission (Heath, 1983). Further, research indicates that privileged and White predominant cultures create hostile or anxiety-laden learning environments where racial tensions among students, faculty, and staff prevail (King et al., 2011). Therefore, in order to help educators enhance diversity on campus, this study explored student affairs practitioners' perceptions of the strategies they implement to deconstruct predominantly White cultures and create racially inclusive environments that promote equal representation of minority students and enhance their learning.

Conceptual framework

In an effort to enhance diversity and multiculturalism on campus, Ortiz and Rhoads (2000) supplied a five-step multicultural education framework that helps student affairs practitioners deconstruct White cultures, create inclusive learning environments, and promote equal representation of minority students. In addition to student affairs practitioners, the multicultural education framework also helps students understand the concept of predominant cultures and their role in its obliteration or development (Baxter-Magolda, 2003). According to Baxter-Magolda, students participating in Ortiz and Rhoads's multicultural education framework learn "who they are becoming [which] is central to success in learning, career decisions, understanding diversity, and interacting peacefully with others" (p. 244). Table 1 provides a comprehensive summary of the framework's components.

Research questions

The primary research questions that guided this study are:

1. What are student affairs practitioners' perceptions of the strategies that deconstruct predominantly White cultures on campus?
2. What are student affairs practitioners' perceptions of the strategies that create inclusive learning environments and promote equal representation of minority students?

Table 1 Multicultural education framework (Ortiz and Rhoads, 2000).

<i>Transformational steps</i>	<i>Description</i>
Understanding culture	As a first step, students learn more about their culture, explore their identity and practice how to interact with diverse populations.
Learning about other cultures	This stage teaches students how to acknowledge culturally different others. It shows students that cultural differences are authentic and helps them understand others' behaviors, traditions, and virtues.
Recognizing and deconstructing white culture	In this stage, students explore how white culture defines acceptable practices, behaviors, and principles. Regarding deconstructing white culture, students should engage in dialogue to understand, explain, and show how white culture affects their lives. White students could be uncomfortable in this stage because it teaches them how to confront their own racial identity.
Recognizing the legitimacy of multiple cultures	After deconstructing white cultures, students start acknowledging other cultures. This will help students establish a common culture that values coexistence and promotes diversity.
Developing a multicultural outlook	During this stage, students learn how culture, society and individuals are interrelated and interconnected. Specifically, they learn how diversity and multiculturalism create pluralistic societies where all people are equally represented.

Methodology

Phenomenological research helps interpret and explain sophisticated phenomena through understanding people's lived experiences (Sofaer, 1999). Therefore, a phenomenological approach was deemed appropriate for this study that explored student affairs practitioners' perceptions of the strategies they implement to deconstruct predominantly White cultures and create racially inclusive environments that promote equal representation of minority students and enhance their learning. Willis (2007) explained that phenomenological research helps understand a specific phenomenon by drawing on people's personal experience, and qualitative interviews serve this purpose. According to Patton (2002) interviews "find out what is in and on someone else's mind" (p. 341). Therefore, in order to establish their understanding of the strategies that create inclusive learning environments, semi-structured interviews with 10 student affairs practitioners were conducted at a single private higher education institution in the Northeast United States.

Participants and setting

Polkinghorne (1989) explained that interviewing 5 to 25 participants with direct experience of the investigated phenomenon suffices a qualitative study. Given there is no single rule regarding sample size in qualitative research, the researchers interviewed 10 student affairs practitioners, six of whom were female administrators. This study was conducted at a single mid-sized, racially diverse private higher education institution in the Northeast United States where its largest minority group on campus constitutes 28% of the student body with a diversity index of 0.73. Table 2 provides a brief summary of the sample's demographic data.

Data analysis

Interview data were thematically analyzed in the context of the main research questions which involved "breaking down, examining, comparing, conceptualizing, and categorizing data" (Strauss and Corbin, 1990, p. 61). Thematic analysis helped identify common themes, patterns, and connections between datasets. To enhance the rigor of the research process and establish trustworthiness, study participants reviewed their responses and verified the accuracy of the results. The findings derived from data analysis are presented in the form of narrative passages in order to represent participants' perceptions of the explored phenomenon.

Table 2 Participants' Characteristics ($N = 10$).

Characteristics	<i>N</i>	%
Nationality		
American	5	50%
Canadian	1	10%
Mexican	2	20%
South african	1	10%
Turkish	1	10%
Race and ethnicity		
African american	2	20%
Asian american	1	10%
Native american	1	10%
Hispanic	3	30%
Kurdish	1	10%
White american	2	20%
Social status		
Married	4	40%
Native language		
English	7	70%
Spanish	2	20%
Turkish	1	10%
Fluency in english		
Fluent	10	100%
Degree awarding nation		
USA	8	80%
Canada	1	10%
Turkey	1	10%
Higher education experience		
Five years and above	10	100%

Limitations

As with any study, this research has its own limitations. Given the qualitative nature of the study, the findings are not generalizable but provide valuable insights into the strategies that deconstruct predominantly White cultures and create racially inclusive learning environments. The limited sample size also restricts the application of results to a limited population of student affairs practitioners serving in the field of higher education in the Northeast United States; however, the author believes the findings that are presented in the next section may find applicability among higher education administrators more broadly.

Representative preliminary findings

Student affairs practitioners are uniquely situated to raise awareness of how White culture often determines the degree to which non-White students experience anxiety. In addition, they are also able to disrupt white students' current beliefs and values. The findings will identify approaches and positions that will further inform the field for increased equity and inclusiveness in higher education. Next, representative findings are presented in five sections.

Understanding culture: acknowledging identity, values, and beliefs

According to [Ortiz and Rhoads \(2000\)](#), students must know their own culture and acknowledge their identity prior to learning how to socially interact with diverse populations. In an effort to help White students understand themselves as social agents and identify their social and cultural background, student affairs practitioners work closely with faculty and staff to create professional development programs that teach students how culture shapes the daily aspects of their lives. In support of this argument, one participant explained:

[Along] with university professors and social workers, each semester, I meet with new students, graduate and undergrads, and teach them about their culture and how it is supposed to help them construct identity and how this relates to diversity, pluralism, and democracy.

Another participant mentioned, "What matters for me the most is showing new and existing students how their beliefs and values welcome diversity and how they support admiring and acknowledging different others ... I conduct workshops on this each semester."

Learning about other cultures: A small step brings you closer

[Ortiz and Rhoads \(2000\)](#) explained that learning about other cultures is the second step toward deconstructing predominantly White cultures. Specifically, the authors argue that students at this level start creating an understanding of other cultures, traditions, and beliefs. Findings from this study indicate that student affairs practitioners provide students with cross-racial engagement opportunities where they get to know students of other cultures and ethnicities, which in turn help them recognize the legitimacy of culturally different others. Speaking about the importance of this stage, one student affairs practitioner stated, "We make sure that sororities and fraternities are as diverse as possible and that they reach out to minority students to engage them in various university and community-based activities that teach all students they are one." Another student affairs practitioner explained, "I make sure that I host students of different backgrounds, colors, and ethnicities every other week on campus where they all have free lunch and I make sure that self-segregation does not exist in the room."

Recognizing and deconstructing white culture: overcoming oneself breaks cultural barriers

This step is best described as the process through which students learn how to confront their own identity and belief system in order to enhance learning and success. It is usually done by facilitating dialogues between students, faculty, and staff ([Ortiz and Rhoads, 2000](#)). According to study participants, engaging all students in a dialogue regarding the importance of diversity and multiculturalism helps students adjust common misconceptions about different races and ethnicities, aids students overcome possible existing resentment against others, enhances campus racial climate, and sends a clear message to the public that diversity matters and is not brought to light for public display only. In support of this argument, one student affairs practitioners mentioned:

When we constantly stress the importance of diversity in everything we do and when we talk and consider students' input regarding creating inclusive learning environments communicates our values as a racially diverse school. Therefore, we tend to create extra-curricular activities and co-curricular programs conducive to inclusive learning in order to leave no one behind ... This is a great chance for all of them to interact, share ideas, and engage in change.

Another student affairs practitioner said:

My colleagues and I make sure that we engage all stakeholders in constructive and effective conversations that transforms our community to a melting pot and promotes minority students' interests on campus ... This helps us create policies that enhance the representation of minority students.

Recognizing the legitimacy of multiple cultures: Acceptance fosters coexistence

The main purpose of this stage is to alter existing structures that value a common culture to one that promotes diversity and coexistence (Ortiz and Rhoads, 2000). According to study participants, transforming the organizational culture starts at this stage when they proactively examine their institutional culture in order to determine the factors that desegregate their campuses, reduce racial hostility, and construct racially affirmative environments. Speaking about the importance of this stage, a student affairs practitioner mentioned:

What we usually do at this stage is preparing a comprehensive report to the administration after assessing our campus climate and make recommendations to improve our admission policies, increase need-based financial aid programs, integrate diversity in all academic programs, and increase hiring and promoting faculty of color ... This, of course, cannot be done without senior leadership commitment. We are lucky that the Board of Trustees is on our side and what we do informs policymaking and tells everyone that they matter here.

Developing a multicultural outlook: rethinking worldview nurtures democracy

This final stage is best described as the process through which students learn about the power dynamics of diversity and how different cultures build democratic and pluralistic societies (Ortiz and Rhoads, 2000). According to one student affairs practitioner, "My colleagues and I never spare an opportunity to explain how different cultures are interrelated and interconnected and how each one complements and adds to others. This is really helpful as we try to show where America gains its strength." Another participant explained, "When students of a particular group or race learn about others, their virtues and values, they become more aware of themselves and see the bigger picture, pluralism."

Conclusion and recommendations

Throughout this study, student affairs practitioners have consistently expressed their belief in their role to change campus life no single culture is represented to the exclusion of all others. The informants also maintained that their perceptions of a campus inclusive of all cultures is also one that reduces anxiety about academic and social life. At the same time, they project and hold the belief that the campus values diversity ideally and in practice.

Student affairs practitioners are uniquely situated to understand the perspectives of the student body at any given institution. Moreover, they are in a position of strength to disrupt the thinking that leads to one culture overshadowing and excluding other cultures from campus life. The role of professors, administrators, and other campus stakeholders should not be understated. However, the persons most able to form robust perceptions of the anxieties, fears, and also successes that come with greater understanding of diverse student population enrolled at university and who can expect that their own cultures are a significant part of campus life.

Today, more than ever, student affairs practitioners should show a genuine concern for all students if they are to deconstruct predominantly White cultures. Creating a safe and welcoming learning environment that recognizes the legitimacy and authenticity of each student not only enhances campus racial climate, but also sends a clear message to the university community that diversity truly matters. Student affairs practitioners can achieve this by:

1. Establishing a sincere student-to-administrator/teacher relationship that alleviate student anxiety and stress.
2. Teaching faculty, administrators, and students how to wear hats of humor, patience, understanding, and love.
3. Building a coalition of trust between university stakeholders and garner their support for institutional and policy-level diversity programs.
4. Endorsing openness, transparency, and accountability in each decision they make.
5. Establishing an open door policy whereby students feel welcome to voice their concerns anytime.
6. Bridging the learning gap between different students through differentiated instruction that meets student learning styles and needs.
7. Creating diversity pre-requisite courses whereby all students have to attend before starting their concentration courses.
8. Formulating strategies that increase retention and persistence rates through offering students with all types of support they need during their course of study.

Failing to respond to diversity issues in a timely manner has many implications on students' academic performance and psychological wellbeing. Therefore, it is imperative for student affairs practitioners to strengthen minority students' sense of belonging

through creating a diverse campus community that fosters cross-racial engagement, promotes dialogue across all organizational levels, and welcomes and acknowledges cultural, racial, and ethnic difference. Student affairs professionals are also required to offer professional development programs for faculty and staff if they are to improve cultural sensitivity and ensure diversity responsive teaching that accounts for minority students' learning needs.

Finally, diversity in higher education should not be neglected or brought to light only for public display. Instead, institutional leaders should prioritize multiculturalism, proactively examine their campus racial climate, and constantly endorse social cohesion and national unity programs if they are to create civic cultures inclusive of diversity. As a first step toward accomplishing this ambitious goal, higher education administrators are required to transform espoused institutional values into enacted ones and reconfigure education to include pluralism, secularism, and diversity in all aspects of teaching and learning.

Significance of the study

This study adds to the literature available on diversity and educational equity, specifically, the strategies that deconstruct predominantly White cultures and promote equal representation of minority students. Although data were collected from student affairs practitioners in the Northeast United States, the findings could be of interest to local and international audiences striving to improve the representation of underserved populations in higher education. Further, this study tests the viability of Ortiz and Rhoads's (2000) multicultural education framework in a new context, a racially diverse learning environment in the United States. Finally, findings from this study could serve as a steppingstone for conducting similar research on diversity and multiculturalism at public and private postsecondary institutions in other regions and continents.

Acknowledgments

The author would like to thank Dr. Everard Weber for his valuable input.

References

- Baxter-Magolda, M., 2003. Identity and learning: student affairs' role in transforming higher education. *J. Coll. Student Dev.* 44 (1), 231–247.
- Cabrera, N.L., 2014. Exposing whiteness in higher education: white male college students minimizing racism, claiming victimization, and recreating white supremacy. *Race Ethn. Educ.* 17 (1), 30–55.
- Goode, W.J., 1960. A theory of role strain. *Am. Socio. Rev.* 25, 483–496.
- Harper, S.R., 2013. Am I my brother's teacher? Black undergraduates, racial socialization, and peer pedagogies in predominantly White postsecondary contexts. *Rev. Res. Educ.* 37 (1), 183–211.
- Heath, S., 1983. *Ways with Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, New York, NY.
- Karkouti, I.M., 2016. Black students' educational experiences in predominantly White universities: a review of the related literature. *Coll. Student J.* 50 (1), 59–70.
- King, P.M., Baxter-Magolda, M.B., Massé, J.C., 2011. Maximizing learning from engaging across difference: the role of anxiety and meaning making. *Equity & Excell. Educ.* 44 (4), 468–487.
- Manning, K., Coleman-Boatwright, P., 1991. Student affairs initiatives toward a multicultural university. *J. Coll. Student Dev.* 32 (4), 367–374.
- Matias, C.E., 2015. "I ain't your doc student." the overwhelming presence of whiteness and pain at the academic neoplatation. In: Varner, K., Albert, K., Mitchell, R., Allen, C. (Eds.), *Racial Battle Fatigue in Higher Education*. Rowman & Littlefield Publishers, New York, NY, pp. 59–68.
- Nguyen, C., Duran, L., 2018. Performing and deconstructing whiteness in student affairs. *Vt. Connect.* 39 (17), 113–121.
- Ortiz, A.M., Rhoads, R.A., 2000. Deconstructing Whiteness as part of a multicultural education framework: from theory to practice. *J. Coll. Student Dev.* 41 (1), 81–93.
- Patton, M.Q., 2002. *Qualitative Evaluation and Research Methods*, third ed. Sage, Newbury Park, CA.
- Polkinghorne, D.E., 1989. Phenomenological research methods. In: Valle, R.S., Halling, S. (Eds.), *Existential-phenomenological Perspectives in Psychology*. Plenum, New York, NY, pp. 41–60. https://doi.org/10.1007/978-1-4615-6989-3_3.
- Sofaer, S., 1999. Qualitative methods: what are they and why use them? *Health Serv. Res.* 34 (5), 1101–1118.
- Strauss, A., Corbin, J.M., 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage, Newbury Park, CA.
- Sue, D.W., Lin, A.I., Torino, G.C., Capodilupo, C.M., Rivera, D.P., 2009. Racial macroaggression and difficult dialogues on race in the classroom. *Cult. Divers Ethnic Minor. Psychol.* 15 (2), 183–190.
- Willis, J., 2007. *Foundations of Qualitative Research: Interpretive and Critical Approaches*. Sage Publications, Thousand Oaks, CA.

Historically Black colleges and universities: historical and contemporary issues

Marybeth Gasman, Samuel DeWitt Proctor Endowed Chair & Distinguished Professor Rutgers University, Graduate School of Education, New Brunswick, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Historical background	423
The rise of industrial philanthropy	423
Prominent legal cases and HBCUs	424
The 1960s and civil rights	425
Federal recognition of HBCUs	425
Black colleges today and their environments	425
HBCU contributions in STEM	426
Diversity in HBCU faculty	426
HBCUs and economic mobility	426
The future of HBCUs	426
References	427

Historical background

From their arrival on the shores of the United States, Black people craved knowledge and understood literacy and education as the path to their freedom. These enslaved people pursued many forms of education despite laws, in all Southern states, forbidding them from learning to read and write (Williams, 2005). Several Black colleges were created just before the Civil War, such as Lincoln and Cheyney Universities in Pennsylvania and Wilberforce University in Ohio. These institutions were established through collaboration between free Black ministers affiliated with the African Methodist Episcopal Church in the North and Quakers who supported the abolition of slavery (Drewry and Doermann, 2003).

At the end of the Civil War, the large task of providing education to more than 4 million formerly enslaved Black individuals was coordinated by both the federal government, through the Freedman's Bureau, and many northern church missionaries – of both Black and White denominations (Brazzell, 1992). Beginning in 1865, the government's Freedmen's Bureau began creating Black colleges, resulting in principals and teachers with military backgrounds. During the postbellum period, most Black colleges were so in name only; these institutions, with a few aberrations, provided primary and secondary education. Of note, this approach was also a feature at most historically White colleges and universities—starting with Harvard University—during the first decades of their existence. Missionary organizations—mainly affiliated with northern White denominations such as the Baptists and Congregationalists and some linked with Black churches such as the African Methodist Episcopal and the African Methodist Episcopal Zion—were actively working with the Freedmen's Bureau. Two of the most well-known White organizations were the American Baptist Home Mission Society and the American Missionary Association, but there were many others. White northern missionary societies founded Black colleges such as Fisk University in Nashville, Tennessee, and Spelman College in Atlanta, Georgia. The benevolence of the missionaries was tinged with self-interest and sometimes racism (Anderson, 1988). Their goals in establishing these colleges were to bring Christianity to the freedmen and to rid the country of the “menace” of uneducated Black people. Among the many colleges founded by Black denominations were Morris Brown College in Georgia, Paul Quinn College in Texas, and Allen University in South Carolina. Unique among American colleges, these institutions were founded by Black people for Black people and had an expressed mission to affirm and uplift Black people. Because these institutions relied on less support from White missionaries and White philanthropists, they were able to craft their own curricula without outside influence and control. However, they also were more susceptible to financial instability (Watkins, 2001).

With the passage of the second Morrill Act in 1890, the federal government again took an interest in Black education, establishing public Black colleges (Thelin, 2011). This act stipulated that those states practicing segregation in their public colleges and universities would forgo federal financial support unless they established agricultural and mechanical institutions for the Black population. Despite the wording of the Morrill Act, which called for the equitable distribution of federal monies, these new institutions received fewer dollars than their White counterparts and as a result, had inferior facilities. Among the 17 new “land grant” colleges were institutions such as Alcorn State University and Kentucky State University.

The rise of industrial philanthropy

At the close of the 19th century, private Black colleges had exhausted funding from missionary sources. At the same time, new support emerged, that of White, northern, industrial philanthropy. Among the industry titans who initiated this type of support were John D. Rockefeller, Andrew Carnegie, Julius Rosenwald, and John Slater. These men were motivated by Christian

benevolence, but more so by a desire to control all forms of industry. The organization making the largest contribution to Black education was the General Education Board (GEB), a conglomeration of northern White philanthropists, established by John D. Rockefeller Sr. but led by his son John D. Rockefeller Jr. Between 1903 and 1964, the GEB gave more than \$63,000,000 to Black colleges, a substantial figure, but only a fraction of what they donated to White institutions. Regardless of their individual motivations, the funding system that these industrial moguls established demonstrated a strong desire to control Black education for their own benefit—to produce graduates who were skilled in the trades that served their own enterprises; this type of education was known as industrial education (Anderson, 1988; Watkins, 2001). Above all, the colleges that they supported were very careful not to upset the segregationist power structure that dominated the American South by the 1890s. Black colleges such as Tuskegee and Hampton were showcases of industrial education (Watkins, 2001). It was here that students learned how to care for horses, make dresses, cook, and clean under the leadership of individuals like Samuel Chapman Armstrong (Hampton) and Booker T. Washington (Tuskegee). Learning these types of skills was silly in many ways as Black people already had these skills due to the confines of slavery (Gasman and McMickens, 2010).

The White philanthropists' support of industrial education was in direct conflict with many Black scholars and intellectuals who favored a liberal arts curriculum because it stimulated critical thinking skills (Anderson, 1988; Gasman, 2007). Institutions such as Dillard, Howard, Fisk, Morehouse and Spelman were more focused on the liberal arts curriculum favored by W. E. B. Du Bois than on Booker T. Washington's emphasis on advancement through labor and self-sufficiency (Anderson, 1988; Bieze and Gasman, 2012; Gasman, 2007; Watkins, 2001). Whatever the philosophical disagreements may have been between Washington and Du Bois, the two educational giants did share a goal of educating African Americans and uplifting their race (Bieze and Gasman, 2012). Their differing approaches might be summarized as follows: Washington favored educating blacks in the industrial arts so they might become self-sufficient as individuals, whereas Du Bois wanted to create an intellectual elite in the top ten percent of the Black population (the Black "talented tenth") to lead the race as a whole toward self-determination (Aldridge, 2018; Brazzell, 1992; Gasman, 2002).

In 1915, there was a slight change in the attitude of the White, industrial philanthropists; they started to turn their attention to the Black colleges that emphasized the liberal arts (Anderson, 1988). Realizing that industrial education could exist in collaboration with a liberal arts curriculum, the philanthropists chose to distribute their money and influence throughout the Black educational system. The pervasive influence of White industrial philanthropy in the early 20th century created a conservative environment on many Black college campuses. This environment tended to tolerate only those administrators (typically White men) who accommodated segregation (Anderson, 1988; Gasman, 2007; Watkins, 2001). However, attention from the White philanthropists was not necessarily welcomed by institutions like Fisk University, where protests ensued against autocratic presidents who were thought by students to be mere puppets of the White philanthropists. Despite these conflicts, industrial philanthropists provided significant support for private black colleges up until the late 1930s (Anderson, 1988; Gasman, 2007).

At this time, the White, industrial philanthropists turned their attention elsewhere. In response, Frederick D. Patterson, then president of the Tuskegee Institute, suggested that the nation's private Black colleges come together around their fundraising efforts. As a result, in 1944, the leaders of 29 Black colleges established the United Negro College Fund (UNCF) (Drezner, 2008; Gasman, 2004, 2007; Tucker, 2002; Williams, 1980). The UNCF began solely as a fundraising entity but eventually took on a large advocacy role for Black colleges as well.

Prominent legal cases and HBCUs

Until the *Brown v. Board of Education* decision in 1954, both public and private Black colleges in the South remained segregated by law and were the only educational options for Black people. Although most institutions of higher education did not experience the same fallout from the *Brown* decision as southern public schools, they were greatly affected by the decision (Anderson, 1988; Walker, 1996). The Supreme Court's landmark ruling meant that Black colleges would be placed in competition with White institutions in their efforts to recruit Black students. With the triumph of integration in terms of the law, many began to call Black colleges into question and point to them as vestiges of segregation. Desegregation proved glacially slow, with public Black colleges maintaining their racial makeup into the current day. In the state of Mississippi, for example, the *Fordice* case was held up in the courts for almost 25 years, with a firm and final decision handed down in 2004. The *Fordice* case, which reached the United States Supreme Court, asked whether the State of Mississippi had met its affirmative responsibility under the 14th Amendment's Equal Protection Clause to dismantle its prior dual university educational system. Despite ample evidence to the contrary, the high court decided that the answer was yes. Although the *Fordice* case applied only to those public institutions within the fifth District, it had a rippling effect within most southern states, resulting in sluggish funding for public Black colleges and limited access for African Americans into predominantly White institutions.

Following the *Brown* decision, private Black colleges, which have always been willing to accept students from all racial and ethnic backgrounds if the state laws would allow, struggled to defend issues of quality in a climate that labeled anything all Black as inferior. Many Black colleges also suffered from "brain drain," as Predominately White Institutions in the North and some in the South made efforts to attract the top 10% of their students to their institutions once racial diversity was embraced by higher education.

The 1960s and civil rights

Black colleges began to change significantly in the 1960s and were quite different than those of the 1920s. The leadership switched from White to Black and, because Black people had more control over institutional funding, there was increased tolerance for Black dissent and self-determination. On many private and public Black college campuses, students were participating in sit-ins and protesting against segregation. Most prominent among these sit-in movements were the four Black college students from North Carolina A & T University, who refused to leave a segregated Woolworth lunch counter in 1960 (Morgan and Davies, 2002; Williamson-Lott, 2018).

During the 1960s, the federal government began to have a greater interest in Black colleges. In an effort to provide clarity, the 1965 Higher Education Act defined a Black college as an institution whose primary mission was the education of African Americans. The recognition of the uniqueness of Black colleges implied in this definition led to increased federal funding for these historic institutions (Gasman, 2007; Thelin, 2011).

Federal recognition of HBCUs

The federal government intervened again on behalf of Black colleges in 1980 when President Jimmy Carter authorized and signed Executive Order 12,232, which established a national program to alleviate the impact of systemic racism and strengthened and expanded Black colleges in their efforts to provide quality education. Since this time, every US president has provided funding to Black colleges through this program. President George H. W. Bush followed up on Carter's initiative in 1989, signing Executive Order 12,677, which established the Presidential Advisory Board on Historically Black Colleges and Universities to advise the president and the Secretary of Education on the future of these institutions.

Black colleges today and their environments

As mentioned earlier, the first HBCUs were established in the United States just prior to and just after the end of the Civil War. They were birthed in a time of deep segregation, at a time when most institutions systematically excluded Black people. This makes them unlike any other universities or colleges in the United States. The only peer is the system set up for the non-White population by the South African Apartheid government in 1959. But while there were once nearly 200 HBCUs in the United States, South Africa only created five, including some designed for "Coloureds" and Indians as well as Africans. Now most of South Africa's 'Black' colleges have been merged with nearby historically White institutions (with the exception of the University of the Western Cape, which was established for the "Colored" population) (Green, 2004). In the United States, although there have been a few mergers, and talk of a few more, the majority of HBCUs have maintained the infrastructure of their origins and remain majority African American. As with their South African counterparts, HBCUs always operated under unequal and structurally inferior conditions. Even today, as research demonstrates that some public HBCUs are forced to educate their students with only half the funding given to counterpart Primarily White Institutions (PWIs) in the same state (Boland and Gasman, 2014).

Currently, more than 300,000 students attend the nation's 105 historically black colleges (40 public four-year, 11 public two-year, 49 private four-year, and 5 private 2-year institutions). This amounts to 8% of all African American college students. Overall, the parents of Black students at Black colleges have much lower family incomes than those of parents of Black students who attend Predominantly White Institutions. Many researchers who study Black colleges, however, have found that African Americans who attend Black colleges have higher levels of self-esteem and find their educational experiences more nurturing (Conrad and Gasman, 2015). Moreover, graduates of Black colleges are more likely to continue their education and pursue graduate degrees than their counterparts at Predominantly White Institutions. Despite the fact that only 8% of African American college students attend Black colleges, these institutions produce significant numbers of our nation's African American professors, politicians, lawyers, doctors, scientists, and teachers (Gasman et al., 2017; Gasman and Nguyen, 2019; Perna et al., 2009).

Black colleges in the 21st century are remarkably diverse and serve varied populations (Gasman and Nguyen, 2015; Jewell, 2002; Lundy-Wagner, 2015; Maramba et al., 2015; Shorrette and Arroyo, 2015). Although most of these institutions maintain their historically Black traditions, on average 13% of their students are white, two percent are Latino, and two percent are Asian American (Gasman, 2013). Because of their common mission (that of racial uplift), they are often lumped together and treated as a monolithic group, causing them to be unfairly judged by researchers, the media, and policymakers (Gasman, 2007; Williams et al., 2019). Just as Predominantly White Institutions are varied in their mission and quality, so are the nation's Black colleges. Today, the leading Black colleges cater to those students who could excel at any top tier institution regardless of racial makeup. Other institutions operate with the needs of Black students in the surrounding region in mind. And some maintain an open enrollment policy, reaching out to those students who would have few options elsewhere in the higher education system (Gasman et al., 2010).

Black colleges have particular strengths that other institutions could learn from. These include their ability to propel student into successful careers in the STEM fields, their faculty diversity, and their role in fostering economic mobility for their students. The sections below delve more deeply into these areas.

HBCU contributions in STEM

According to the National Science Foundation, HBCUs confer nearly 15% of all bachelor's degrees (16% in the sciences) awarded to Black students despite making up only 3% of all postsecondary institutions in the US. Xavier University of Louisiana now produces more alumni who go on to graduate from medical schools than any other institution in the country. In fact, along with Xavier, according to the Association for American Medical Colleges, four HBCUs – Spelman College, Morehouse University, Howard University and Florida A&M University – other HBCUs are ranked in the top-20 for supplying the largest percentage of Black students applying to medical school. When we consider that these HBCUs enroll 2–10,000 students and are competing against much larger and more resourced institutions such as the University of Florida (52,218 students) or the University of Texas at Austin (51,832 students), we cannot ignore that HBCUs are doing something profound, which any university committed to diversity and creating a positive learning culture for all students should consider (Gasman and Nguyen, 2019; National Academies of Sciences, Engineering, and Medicine, 2019).

HBCUs generally value collaboration over competition, especially in the STEM fields. At most colleges and universities in the United States, the sciences are highly competitive, deeply individualistic and are built on the idea of “weeding out” students in early classes or being exclusive. As a result of this embedded culture, students often see classmates as competition and don't want to help each other. Such an atmosphere can serve as a deterrent to African Americans, since it tends to reinforce feelings of isolation that these students already experience in classes filled with mainly White men (Gasman and Nguyen, 2019; McGee and Martin, 2011; Ong et al., 2011). At HBCUs, however, students are encouraged to work together and to understand how their success is tied to each other's. They are urged to make sure that their classmates understand the material, to share notes, to work with peers who might be taking longer to understand the material and to socialize with each other in order to build trust in the classroom and around group projects (Conrad and Gasman, 2015; Gasman and Nguyen, 2019).

Diversity in HBCU faculty

HBCUs also provide same-race role models for Black students. Whereas the faculty makeup at PWIs is 76% White, HBCUs have a faculty that is 60% Black with the remaining 40% being highly diverse across the racial and ethnic spectrum (Gasman and Nguyen, 2019; McGee and Martin, 2011). Research shows that when African Americans see someone like them in the classroom and having an opportunity to forge friendships with them, they are more likely to persist in college and in their particular science major (McGee and Martin, 2011; National Academies of Science, Engineering and Medicine, 2019). Gender adds an additional dimension to this. Figures from 2016 indicate that Black women made up only 3% of professors in the US, but 18% of them work in HBCUs. They can therefore act as crucial role models for their Black women students (Gasman and Nguyen, 2019).

HBCUs and economic mobility

Black colleges built the Black middle class in the United States. Although they educate fewer African Americans than they did prior to the end of segregation, these institutions continue to have a profound impact on the students who attend them. Research shows that HBCUs contribute disproportionately to the upward mobility of their students, especially given that the majority (90%) of their students are from low-income households and 71% are eligible for Pell Grants, making them some of the poorest students in the nation (Natheson, Castro Samayoa and Gasman, 2019). Nearly 70% of Black students who enter HBCU from low-income homes attain middle-class incomes after graduation. HBCUs that perform in this manner at a particularly high rate include Tuskegee University, Bennett College, Florida A&M University, Dillard University and Clark Atlanta University. These institutions have some of the greatest percentages of low-income students among HBCUs but still move them to the middle class. The institutions that are able to move students to the highest levels in terms of socio-economic status – meaning moving students from the lowest levels to the highest levels are Xavier University of Louisiana, Dillard University, Prairie View A&M University and Tuskegee University. Of note, some of the most highly ranked HBCUs, such as Spelman College, Howard University, and Morehouse College, propel more than one-third of their low-income students into the top-fifth of income earners. Overall, research demonstrates that African Americans who graduate from HBCUs are less likely to move downward in terms of economic mobility than their peers at Predominantly White Institutions (Nathenson et al., 2019).

The future of HBCUs

Despite a storied history and ample contributions to the continued success of African Americans, HBCUs are constantly questioned about their relevance and importance in the 21st Century (Gasman et al., 2019; Kimbrough, 2011). Those without a deep understanding of their history and current day contributions, make predictions about their demise and often refer to them in derogatory ways, even stating that they “retard” students (Brown, 2013; Butrymowicz, 2014; Carr, 2013; Carter, 2019; Fryer and Greenstone, 2007; Riley, 2010; Smith-Barrow, 2019; Vedder, 2019; Whitford, 2021). However, much like Black people overall, HBCUs have proven resilient with even those HBCUs that have faced seemingly insurmountable problems coming back from the brink. In recent

years, these HBCUs have included Paul Quinn College in Dallas, Texas, Morris Brown College in Atlanta, Georgia, and Wilberforce University in Wilberforce, Ohio. Interestingly, each of these particular HBCUs is affiliated with the African Methodist Episcopal Church and were founded by African Americans for African Americans.

In recent years, HBCUs have garnered more positive attention from media outlets as well as philanthropists, including MacKenzie Scott (Amazon) and Reed Hastings (Netflix) (Anderson and Lumpkin, 2020; Beras, 2020). HBCU leaders have become more vocal in terms of sharing their successes and making sure that the larger nation understands their contributions. In addition, racial turmoil and discrimination has led more students and their parents to look to HBCUs for a college education as they have always ensured that Black lives matter on their campuses. Moreover, many of the most prominent and fresh political leaders in the United States were educated at HBCUs – including Vice President Kamala Harris, voting rights activist Stacey Abrams, and Georgia Senator Raphael Warnock.

References

- Aldridge, D.P., 2018. *The Educational Thought of WEB Du Bois: An Intellectual History*. Teachers College Press, New York, NY.
- Anderson, J.D., 1988. *The Education of Blacks in the South, 1860–1935*. University of North Carolina Press, Chapel Hill, NC.
- Anderson, N., Lumpkin, L., 2020. 'Transformational': MacKenzie Scott's gifts to HBCUs, other colleges surpass \$800 million. Wash. Post. https://www.washingtonpost.com/local/education/mackenzie-scott-hbcu-donations/2020/12/17/Oce9ef5a-406f-11eb-8db8-395dedaaa036_story.html.
- Beras, E., 2020. Historically Black Colleges Get a \$120 Million Gift. They Need It. <https://www.marketplace.org/2020/06/18/historically-black-colleges-get-a-120-million-gift-they-need-it>.
- Bieze, M.S., Gasman, M. (Eds.), 2012. *Booker T. Washington Rediscovered*. Johns Hopkins University Press, Baltimore, MD.
- Boland, W., Gasman, M., 2014. *America's Public HBCUs: A Four State Comparison of Institutional Capacity and State Funding Priorities*. Penn Center for Minority Serving Institutions.
- Brazzell, J.C., 1992. Bricks without straw: missionary-sponsored Black higher education in the post-emancipation era. *J. High Educ.* 63 (1), 26–49.
- Brown, M.C., 2013. The declining significance of historically black colleges and universities: relevance, reputation, and reality in obamamerica. *J. Negro Educ.* 82 (1), 3–19.
- Butrymowicz, S., 2014. Historically black colleges are becoming more white. Time. <https://time.com/2907332/historically-black-colleges-increasingly-serve-white-students>.
- Carr, S., 2013. Black Colleges Face Uphill Battle to Survive. Hechinger Report. <https://hechingerreport.org/black-colleges-face-uphill-battle-to-survive>.
- Carter, J., 2019. Defunct HBCUs Must Stop Defrauding Blacks. Florida Courier. <https://flcourier.com/defunct-hbcus-must-stop-defrauding-blacks>.
- Conrad, C., Gasman, M., 2015. *Educating a Diverse Nation*. Harvard University Press, Cambridge, MA.
- Drewry, H.N., Doermann, H., 2003. *Stand and Prosper: Private Black Colleges and Their Students*. Princeton University Press, Princeton, NJ.
- Drezner, N.D., 2008. For alma mater and the fund: the United Negro College Fund's National Pre-Alumni Council and the creation of the next generation of donors. In: Gasman, M., Tudico, C. (Eds.), *Historically Black Colleges and Universities*. Palgrave Macmillan, New York, NY, pp. 15–26.
- Fryer, R., Greenstone, M., 2007. The Causes and Consequences of Attending Historically Black Colleges and Universities. National Bureau of Economic Research, Working Paper 13-36. <https://www.nber.org/papers/w13036>.
- Gasman, M., 2002. WEB Du Bois and Charles S. Johnson: differing views on the role of philanthropy in higher education. *Hist. Educ. Q.* 42 (4), 493–516.
- Gasman, M., 2007. *Envisioning Black Colleges: A History of the United Negro College Fund*. Johns Hopkins University Press, Baltimore, MD.
- Gasman, M., 2004. Rhetoric vs. reality: the fundraising messages of the United Negro College Fund in the immediate aftermath of the Brown decision. *Hist. Educ. Q.* 44 (1), 70–94.
- Gasman, M., 2013. *The Changing Face of Historically Black Colleges and Universities*. Penn Center for Minority Serving Institutions, Philadelphia, PA.
- Gasman, M., Toldson, I., Price, G., 2019. Covid Closures Could Hit HBCUs Particularly Hard. The Conversation. <https://theconversation.com/covid-19-closures-could-hit-historically-black-colleges-particularly-hard-134116>.
- Gasman, M., McMickens, T.L., 2010. Liberal or professional education?: the missions of public Black colleges and universities and their impact on the future of African Americans. *Souls* 12 (3), 286–305.
- Gasman, M., Nguyen, T.H., 2019. *Making Black Scientists: A Call to Action*. Harvard University Press, Cambridge, MA.
- Gasman, M., Nguyen, T.H., Conrad, C.F., Lundberg, T., Commodore, F., 2017. Black male success in STEM: a case study of Morehouse College. *J. Divers. High. Educ.* 10 (2), 181–200.
- Gasman, M., Nguyen, T.H., 2015. Myths dispelled: a historical account of diversity and inclusion at HBCUs. *N. Dir. High. Educ.* 2015 (170), 5–15.
- Gasman, M., Lundy-Wagner, V., Ransom, T., Bowman III, N., 2010. Unearthing promise and potential—our nation's historically black colleges and universities. *ASHE High. Educ. Rep.* 35 (5), 1–134.
- Green, P.E., 2004. A comparative examination of historically disadvantaged institutions in the republic of South Africa and historically black colleges and universities in the United States. *Race, Gender & Class* 11 (3), 153–176.
- Jewell, J.O., 2002. To set an example: the tradition of diversity at historically Black colleges and universities. *Urban Education* 37 (1), 7–21.
- Kimbrough, W., 2011. Black colleges still play a vital role in education. Chron. High Educ. <https://www.chronicle.com/article/black-colleges-still-play-a-vital-role-in-education>.
- Lundy-Wagner, V.C., 2015. Coming Out of the Shadows: Rethinking the Education Policy Agenda for Diversity and HBCUs. *New Directions for Higher Education*, pp. 91–101, 2015(170).
- Maramba, D.C., Palmer, R.T., Yull, D., Ozuna, T., 2015. A qualitative investigation of the college choice process for Asian Americans and Latina/os at a public HBCU. *J. Divers. High. Educ.* 8 (4), 258–271.
- McGee, E.O., Martin, D.B., 2011. "You would not believe what I have to go through to prove my intellectual value!" Stereotype management among academically successful Black mathematics and engineering students. *Am. Educ. Res. J.* 48 (6), 1347–1389.
- Morgan, I., Davies, P., 2012. *From Sit-Ins to SNCC: The Student Civil Rights Movement in the 1960s*. University Press of Florida, Gainesville, FL.
- Nathanson, R., Castro Samayoa, A., Gasman, M., 2019. Moving Upward and Onward: Income Mobility at Historically Black Colleges and Universities. Rutgers Center for Minority Serving Institutions. https://cmsi.rutgers.edu/sites/default/files/EMreport_R4_0.pdf.
- National Academies of Sciences, Engineering, and Medicine, 2019. *Minority Serving Institutions: America's Underutilized Resource for Strengthening the STEM Workforce*. National Academies Press.
- Ong, M., Wright, C., Espinosa, L., Orfield, G., 2011. Inside the double bind: a synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harv. Educ. Rev.* 81 (2), 172–209.
- Perna, L., Lundy-Wagner, V., Drezner, N.D., Gasman, M., Yoon, S., Bose, E., Gary, S., 2009. The contribution of HBCUs to the preparation of African American women for STEM careers: a case study. *Res. High. Educ.* 50 (1), 1–23.
- Riley, J., 2010. Black colleges need a new mission: once an essential response to racism, they are now academically inferior. Wall St. J. <http://www.nafeonation.org/hbcu-leaders-respond>.
- Shorette, C.R., Arroyo, A.T., 2015. A closer examination of White student enrollment at HBCUs. *N. Dir. High. Educ.* 2015 (170), 49–65.

- Smith-Barrow, D., 2019. H.B.C.U.s' sink-or-swim moment. N. Y. Times. <https://www.nytimes.com/2019/10/21/opinion/hbcu-college.html>.
- Thelin, J.R., 2011. *A History of American Higher Education*. Johns Hopkins University Press, Baltimore, MD.
- Tucker, S.K., 2002. The early years of the united Negro college fund, 1943-1960. *J. Afr. Am. Hist.* 87 (4), 416–432.
- Vedder, R., 2019. Is 'diversity' Destroying the HBCUs? *Forbes*. <https://www.forbes.com/sites/richardvedder/2019/11/04/is-diversity-destroying-the-hbcus/?sh=34ec38024195>.
- Walker, V.S., 1996. *Their Highest Potential: An African American School Community in the Segregated South*. University of North Carolina Press, Chapel Hill, NC.
- Watkins, W.H., 2001. *The White Architects of Black Education: Ideology and Power in America, 1865-1954*. Teachers College Press.
- Whitford, E., 2021. Back from 2 Decades on the Brink. *Inside Higher Education*. <https://www.insidehighered.com/news/2021/01/12/morris-brown's-accreditation-would-mean-resurrection-hbcu-long-brink-closure>.
- Williams, K.L., Burt, B.A., Clay, K.L., Bridges, B.K., 2019. Stories untold: counter-narratives to anti-Blackness and deficit-oriented discourse concerning HBCUs. *Am. Educ. Res. J.* 56 (2), 556–599.
- Williams, H.A., 2005. *Self-taught: African American Education in Slavery and Freedom*. University of North Carolina Press, Chapel Hill, NC.
- Williams, L.E., 1980. The United Negro College Fund in retrospect: a search for its true meaning. *J. Negro Educ.* 49 (4), 363–372.
- Williamson-Lott, J.A., 2018. *Jim Crow Campus: Higher Education and the Struggle for a New Southern Social Order*. Teachers College Press, New York, NY.

"It was a life enhancing activity": the academic habitus of PhD students

Tehmina N. Basit, School of Life Sciences and Education, Staffordshire University, Stoke-on-Trent, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	429
Methods	430
Teaching and learning	431
Getting acquainted with processes and procedures	431
PhD supervision	431
Managing their own learning	432
Peer relationships	433
Socialization	433
Support from peers	434
Employment prospects	434
Graduate attributes	434
Career opportunities	434
Conclusion	435
References	436

Introduction

Undertaking a doctorate can be a highly fulfilling experience, both personally and professionally, and students may wish to study for a PhD for a number of reasons. As a result of an online survey of 1901 Asian and African PhD students, [Fomunyam \(2020\)](#) reports the students' motivations to obtain a PhD to include intrinsic interest, employment considerations, and personal and professional development. Nevertheless, studying for a PhD can be a lonely experience ([Juniper et al., 2012](#)) and students can encounter multiple challenges during their period of study leading to feelings of stress and isolation ([Janta et al., 2014](#)). Students at UK universities may start at two or three different points in the academic year and unless concerted efforts are made by staff and students, the latter may feel isolated and unsupported. This is in contrast to the professional doctorates in which the students begin as a group by studying core and optional modules sequentially, and a community of practice quickly develops. Conversely, there may not be a structured curriculum for PhD study as it is a research degree. Also, much of PhD pedagogy is in the form of supervision by a team, usually comprising an experienced principal supervisor and a novice second supervisor, and the quality of relationships varies, which may impact on students' motivation and progression.

PhD students, therefore, need to be highly motivated and able to manage their time effectively, as well as being capable of completing a degree at doctoral level, while having an effective working relationship with their supervisors who are able and willing to scaffold students' learning on an on-going basis. PhD students may start the program as unconfident individuals and may continue to view their experience as challenging and demanding, with a tendency to worry about whether they will complete on time and attain the required standards. There can be issues related to the student's gender, time, finance, childcare and other caring responsibilities, confidence and so forth. Achieving a balance between home and academic life can be a source of stress, which may lead to dropping out of the program ([Castello et al., 2017](#)). These problems are even more acute in case of PhD students in an environment where studentships in the UK are few and far between and are extremely competitive.

[Maloshonok and Terentev \(2019\)](#), in their research on doctoral students in Russia, note various problems encountered by these students to include transition to a structured model of doctoral education; diffusion of doctoral education's goals; unpreparedness of Russian universities for expansion of PhD education; ineffective mechanisms of student selection; lack of funding and need for paid work; excessive dependence on supervisors; and insufficient study time and skills for meeting the requirement for publications before the date of defense. They acknowledge that some problems correlate with the global challenges, but some are unique to the Russian institutional context.

Yet, studying for a PhD is not a negative experience for all students. [Mowbray and Halse \(2010\)](#) argue that in contrast to the literature highlighting the traumatic and debilitating effects of the PhD, their qualitative study of 20 final year PhD students at an Australian university illustrates that struggles, setbacks and other negative experiences during the PhD can have positive impacts, as they improve the students' experiential knowledge and understanding of the world, making them creative and resourceful problem-solvers. Students in Iran value their student identity and the feeling of being a student despite experiencing everyday necessities and struggles of life and facing considerable challenges in the learning milieu ([Hemmati and Mahdie, 2020](#)).

However, there can be issues for international students in addition to those that are faced by all PhD students. [Son and Park \(2014\)](#) point to the limited availability of resources for international students in Australian universities, the lack of understanding of the challenges and difficulties faced by such students, to include language problems, cultural differences and personal issues, and their implications for the Australian higher education context. In the UK, university guidelines, though helpful, fail to fully address

the requirement to support international PhD students who come to study at a foreign university, having made substantial personal, professional and financial sacrifices (Cree, 2012). Research in the US indicates that most international students persist despite facing problems, as they are motivated by an intrinsic interest in research and teaching, high value placed on an earned PhD from a US university, and high emotional and social cost of withdrawing (Zhou, 2015).

A large percentage of PhD students may already be in employment when they embark on a doctorate, and hope to hone their personal and professional competencies further through doctoral study. Platow (2012) highlights the perceived acquisition of graduate attributes related to post-PhD productivity, and the role that PhD education has in developing the attributes which could be brought to fruition in employment through enhanced productivity. Generally, the gains to society appear to be higher than the gains to the PhD graduates themselves and to their employers. The traditional PhD degree is sometimes viewed as too narrow, the research findings of limited use beyond the student's specialism and does not necessarily guarantee an individual's employment and earning prospects. Sharmini and Spronken-Smith (2020) note that most PhD programs in the UK and Australasia focus on developing disciplinary knowledge and research skills, with the assessment narrowly concentrating on a thesis and viva.

Nevertheless, highly qualified people such as PhDs can enhance other people's abilities, contribute to a pool of knowledge from which other people can draw, and are better at gaining from that pool of knowledge themselves, thus contributing to societal good (Casey, 2009), as the majority of them may already be in employment and making an important contribution to society on an ongoing basis through their PhD research. Universities are able to provide all types of knowledge and skills which can be useful to employees in the workplace, though they also ought to ensure that employers understand the language of academic quality (Basit et al., 2015).

This article uses Bourdieu's theory of habitus as a lens to examine the PhD program at a university. Bourdieu (1990) considers habitus as both a system of production of practices, and perception and appreciation of practices. Habitus produces practices and representations which are available for classification, and are objectively differentiated, though they are perceived as such only by individuals who possess the classificatory schemes needed to understand their social meaning. Reay (2004) notes that the notion of habitus, which is at the heart of Bourdieu's theoretical framework, is complex as it takes many shapes and forms. Habitus is "a constellation of perceptions and attitudes that frame particular actions as possible in different situations. So, habitus acts as both a generative and restrictive mechanism to action such that some actions are deemed more appropriate in certain contexts based on an individual's class status and experiences. This then assists in determining what is valued and acceptable" (Gopaul, 2011:14).

Further, Bourdieu (1988) highlights the notion of academic habitus, which he applies to university teachers, and notes the reproduction of the university system which produces teachers possessing stable and homogeneous social and academic characteristics, and an academic habitus at every hierarchical level. It can be argued that Bourdieu's concept of academic habitus is equally relevant to PhD students, who are advised and guided, mainly, though not exclusively, through a one-to-one system of pedagogy. This helps them to develop new knowledge, independent thinking and reflective practices, leading to an academic habitus apposite to PhD students who will ultimately cascade this knowledge to wider groups and make an original contribution to their field of study for the benefit of future generations. Delamont et al. (1997) draw attention to the transmission of academic habitus in the natural sciences and contend that the enculturation of doctoral students is fundamental to intergenerational transfer of cultures within scientific disciplines through unique methods of social and intellectual organization. They note the strength of pedagogic continuity in the natural sciences through laboratory arrangements or research groups, which could be possible in humanities and in social sciences too, though in a different form. Their research on doctoral students indicates that pedagogic continuity is vital for understanding the intergenerational transfer of knowledge and the reproduction of academic habitus.

Methods

This article is based on a research project which sought to examine the perceptions of PhD students regarding their experiences at a post-1992 university in the UK (hereafter called the University). Traditionally, post-1992 universities, or "New" universities, differ from "Red-Brick" and/or "Russell Group" universities in certain ways. New universities attract more local students who live at home, part-time students who usually work fulltime while they study for a doctoral degree, and more home than international students. This is reflected in our research participants.

The main method of data collection was interviews, as "interviews are embedded in human experience and flagrantly draw on the participants' subjectivities. They seek the participants' perceptions of the social world as it is experienced and lived by them and those around them" (Basit, 2010:100). All PhD students in different schools within the four faculties at the University were approached and invited to participate in the research. Semi-structured interviews were carried out with the students: face-to-face or by Skype. A total of 23 PhD students were interviewed, including both male (10) and female (13); home (15), EU (2) and international (6) students; and those at different stages of their PhD, from across the University. The majority of them were part-time (18) students, with a few on the full-time (5) mode of study.

The following nine Schools from the University's four faculties participated in the study: Art and Design; Journalism, Humanities and Social Sciences; Business; Education; Law; Engineering; Sciences; Psychology, Sport and Exercise; and Nursing and Midwifery. For reasons of anonymity and confidentiality, each school has been identified by a letter. Further, students' names have been changed to pseudonyms. All participants have been given Western names so that international and minority ethnic students in different schools cannot be recognized. A number of issues were raised in the interviews and the data were analyzed by identifying themes related to these issues. These are discussed below under the headings of themes and sub-themes emanating from the data.

Teaching and learning

Getting acquainted with processes and procedures

Matos (2013) highlights the increased managerialism of the PhD process due to different funding priorities, which have changed the nature and objectives of doctoral study, though acknowledges improved PhD practices as a consequence of changes implemented by funding bodies. Changes include the annual monitoring reports of PhD students' progress. Mewburn et al. (2014) point to the confusion around the purpose and audience of these reports, which leads to inefficacious reporting by all parties, as countersigning requirements inhibit the frank reporting of progress, with reporting obligations going counter to institutional requirements. Nevertheless, most PhD students in the current study were satisfied with PhD processes and procedures:

My supervisor took me to different offices and to meet different people. Online enrolment was pretty easy. Transfer from [another university], where I had started with my supervisor and followed her here, was smooth as it was taken care of by my supervisor. Annual Progress Review is useful and very well organized.

Laura, School-S.

Despite such comments, some niggling issues were highlighted. It was queried why students were required to enroll every year (Adrian, School-R) and why they were not told at the beginning that they had to enroll annually (Dominic, School-X). Annual monitoring of students needed tightening (Emily, School-T), and some processes were slightly repetitive (Simon, School-S). Some procedures were a little confusing for distance learning students (Penelope, School-G), and there were too many papers to fill in (Melanie, School-S). One student reported a highly disappointing experience with regard to University policies and procedures:

There was a very lengthy delay between submission and viva. Examination team was changed several times by the Research Degrees Committee without clear rationale being given. Chairing of viva was poor, allowing external to dominate proceedings The University's bureaucratic processes in relation to PhD students, as governed by the Research Degrees Committee, are cumbersome and work to the disadvantage of the student.

Caroline, School-X.

Although this was an isolated comment, students, albeit rarely, had frustrating experiences of University procedures, when appropriate support was not provided expeditiously at specific stages. It appeared that the academic habitus of PhD students in the same university was not consistent across the university.

PhD supervision

Effective doctoral pedagogy is based on a mutually respectful relationship between student and supervisor, supported by a flexible learning structure, which enables modeling of scholarly practices, and opportunities for participation and reflection through scaffolding (Cotterall, 2011). Research in the US indicates that supervisory support leads to a strong sense of belonging and academic self-concept for both home and international doctoral students (Curtin et al., 2013). Students and supervisors have equally high expectations of the student-supervisor relationship which should be warm and supportive, and both groups should articulate their expectations clearly at the outset (Moxham et al., 2013).

Many supervisors expect written work from students in the first year, to present papers at conferences and publish a paper before or soon after the submission of thesis indicating a developmental relationship. Gruzdev et al. (2020) in their study on PhD students at Russian universities observe that students report the longest expected time to degree completion and the lowest level of satisfaction with supervisors whom they label "hands-off supervisors", as they provide no help to their students. On the other hand, students express the highest level of satisfaction and shortest expected time to PhD completion with supervisors whom they call "superheroes" and "mentors." However, Wisker and Robinson (2013) question the image of an enriching and developmental relationship between supervisor and doctoral student, and the latter's PhD experience as emotionally fulfilling and intellectually rewarding. Further, Cheng et al. (2016), in their qualitative study at two UK universities suggest that PhD students' satisfaction does not necessarily indicate quality education, since the degree of student satisfaction can be affected by students' expectations before commencing PhD study and their prior ideas about the value of a PhD.

The vast majority of students in the present study expressed their satisfaction with their supervisors, and saw them as outstanding (Dominic, School-X), excellent (Valerie, School-G), providing very good support and knowledge base (Timothy, School-G), who could not be faulted (Ralph, School-M) and who were there when needed (Sylvia, School-L). The quality of the supervision was viewed as timely, responsive, understanding, professional and with a clear common vision (Melanie,

School-S); and very good, very helpful, just what was needed (Olivia, School-R). The supervisors gave clear targets, with the freedom to explore; and the students never felt alone (Simon, School-S); they were proactive in asking students to do conference presentations and write for publication (Valerie, School-G). A full time EU student stated:

Can't complain; have never felt unsupported. I meet with supervisors once a week or once a fortnight. Also, a senior advisor spoke to me about methodology, for example, grounded theory, along with my supervisors.

Joyce, School-G.

And a part-time international student maintained:

My PhD is a credit to my supervisors. My principal supervisor was like my own mother. She gave me time, pep talk, and understood me. PhD is a long process. I might have left the program, but she kept me going. I had productive meetings with her and my second supervisor, face-to-face and by Skype, and regular contact by email.

Laura, School-S.

Such sentiments have been expressed in previous studies, too (see [Cree, 2012](#)). However, one student questioned the quality of supervision, which:

... would not bear close examination by any external agency interested in quality and standards. Second supervisor, who became principal, in name only after two years, never really understood what my thesis was trying to do and became invisible.

Caroline, School-X.

Such a disappointing experience of supervision was rare and was not reported by any other participant. Student satisfaction with their PhD experience is strongly related to the quality of supervision ([Harman, 2002](#)) and student expectations ([Cheng et al., 2016](#)). [Griffioen et al. \(2018\)](#) argue that differences in student ability may result in differences in student satisfaction, and educational stimulation and quality education can be conducive to students' educational fulfillment. This indicates that in ostensibly similar groups, each individual is likely to have his or her unique habitus ([Basit, 2012](#)). No two individual histories are alike, nor are two individual habitus identical, though there are classes of experiences and hence classes of habitus ([Bourdieu, 1993](#)).

As seen in the comment above, change of supervisors can be distressing for PhD students. This can also have ramifications for timely completion. Loss of supervisors through neglect or the supervisor transferring to another university can affect the student's confidence, ability and academic identity ([Wisker and Robinson, 2013](#)). Some students in the current study had problems with supervisors in the past but were happier with a new supervisory team (Abigail, School-Z). Others complained about the tardiness of feedback and felt the supervisors should step up (Emily, School-T).

Research highlights the kind of feedback wanted by doctoral students. This includes comments addressing their main idea, argumentation, clarity and information coverage, with written feedback comprising three aspects of function, focus, and presentation ([Can and Walker, 2014](#)). Also, there can be confusion about writing up as a separate or continuous endeavor, with many students waiting until after the completion of their research to start writing their thesis. Those encouraged to write continuously observe the merits of this approach, which would lead to timely completion ([Lindsay, 2015](#)). In their New Zealand study, [Wei et al. \(2019\)](#) highlight the liminality of the first writing submission and feedback interaction between doctoral students and their supervisors, involving a variety of emotions including anxiety, hope, awkwardness, excitement or dread. They caution that negative emotions can continue if the writing and feedback exchange is not managed carefully by supervisors. Evidently, writing is something that PhD supervisors at universities can advise their students to undertake early, and on a continuing basis, and provide timely, clear and detailed feedback on it to help students to progress in a confident manner.

Managing their own learning

For many students, the personally challenging nature of the PhD is central to the research process ([Owler, 2010](#)). While this may have implications for the wellbeing of the students, [Cotterall \(2011\)](#) argues that students ought to develop an understanding of the demanding nature of a PhD so that they can manage their own learning. Many students in the current study indicated that they could successfully manage their own learning since:

A PhD is, by its very nature, an exercise in demonstrating the ability to research. This means that the students must be able to manage their own learning. Students who have difficulty in this respect may also find it difficult to be independent researchers, and this can only be tested by doing it.

Callum, School-H.

Across the University, students were motivated by their supervisors to be creative and manage their own learning (Nicola, School-R); they felt empowered and encouraged (Joyce, School-G) by the expert guidance of their supervisors who, through active discussions, made them reflect critically and creatively on their work to develop their own learning (Melanie, School-S). This is contrary to research on doctoral students in Sweden (Brodin, 2018) which shows that their scholarly creativity is stifled, leading to a silence emanating from traditionally restrictive intellectual, political and economic institutional agendas.

Nevertheless, managing their own learning could be more difficult for those with fulltime careers, such as the University's own staff members; since as a PhD student and an academic staff member, you had to balance your PhD with your day job (Simon, School-S). Some felt that managing their own learning was an individual skill, which depended on the personality, motivation and creativity of the student (Dominic, School-X). An International student noted:

I had come from a system where you read and reproduce. My supervisors told me what was expected from a PhD student. Now at the end of my PhD, I can do research on my own. PhD has given me this confidence and I am now self-efficacious.

Laura, School-S.

Such confidence in their own abilities is significant, exemplifying an academic habitus, which can then lead to the students applying their knowledge to bring about positive changes in their occupational milieu. International doctoral students in the US highly value research and academic experiences, and despite encountering problems related to enculturation, a higher percentage complete its doctorates earlier than home students (Curtin et al., 2013). The reason for this may be because they want to gain competencies that can enable them to bring about change and improvement in higher education in their own country. Research (Rhoads et al., 2017) shows how Chinese doctoral students develop their methodological knowledge and skills with the aim to generate insights which would be potentially helpful to higher education reform in their home country, leading to the creation of world class research universities in China.

It was suggested that controls should be placed on students to manage their own learning and deliver on time as there was no support on this (Emily, School-T). Conversely, others were expected to manage their own learning almost entirely, which they found problematic (Abigail, School-Z). This illustrates that the academic habitus possessed by the students varied between disciplines and supervisors. While many students were taught to manage their own learning and were successfully able to do so, some others needed appropriate scaffolding to enable them to gradually become independent and creative learners. Literature shows that some students possess considerable academic habitus and are able to undertake doctoral studies without difficulty (Gopaul, 2011). This indicates that universities need to identify students who require additional support and provide it in a timely and appropriate manner.

Peer relationships

Socialization

Research on doctoral students in the departments of chemistry and history at two US universities (Gardner, 2007) highlights the importance of socialization in these students' success. Different kinds of information exchange between doctoral students: basic, administrative, professional, social and personal has been noted by Lee et al. (2017). Only a few students in our study mentioned opportunities for socializing, which included students within specific schools interacting with one another (Nicola, School-R); friendships forming among students within the school (Sylvia, School-L) and those supervised by the same supervisor (Olivia, School-R). In one school, students met once a week and one of them presented from their research (Ralph, School-M). Events such as the joint annual conference with another university were also seen as an opportunity for socializing with peers (Nicola, School-R). However, some students from different schools reported that they refrained from socializing with their peers because of lack of time (Emily, School-T; Joyce, School-G) or lack of inclination (Valerie, School-G; Adrian, School-R; Caroline, School-X), since the majority of students had full time jobs.

One faculty had arranged a successful welcome buffet for all new and current PhD students the year before, which was attended by supervisors as well as students, and hoped to make it an annual event. Another organized semi-social events where the staff got together with the students within the postgraduate research center for "a nibble and chat". Students were also able to socialize when PhD students in a faculty or school shared the same office or laboratory and thus a quasi-research-community gradually developed. Some students felt isolated as they were aware of social events at other faculties (Abigail, School-Z); and thought it would be nice to have occasional interaction with peers and discuss work (Timothy, School-G). Tactics used by doctoral students to combat loneliness and isolation include student-supervisor relationship; working spaces; social interaction with peers; networking at seminars, conferences and workshops; and online fora to provide emotional support to students (Janta et al. 2014). Those who wanted social interaction with their peers felt lonely and mentioned the lack of opportunities for socializing:

I have joined groups online, but no face-to-face interaction Need something like a Graduate School, as in other universities! A sense of belonging is needed.

Abigail, School-Z.

Support from peers

It has been argued that peer learning, located within the concept of communities of practice, could be an appropriate lens to view research degree education (Boud and Lee, 2005), and students view peer mentoring as a developmental role of peer relationships (Lee et al., 2017). There was a feeling among participants in the current study that PhD students in the University lacked a sense of community. The need for more peer support was emphasized (Nicola, School-R). While an online community was encouraged, it did not happen (Dominic, School-X). Minimal peer support, such as feedback at presentations, was reported by the students, though some had different strengths which enabled them to support each other (Ralph, School-M); and sharing a working space helped (Sylvia, School-L; Joyce, School-G); others received no support whatsoever from peers (Emily, School-T; Nicola, School-R). International and EU students had negligible peer support and chances to socialize with other PhD students; they did not seem to be aware of online fora, and suggested a website/portal or Facebook page for alumni and current students to interact with one another (Laura, School-S). Also:

Online seminars where a group of international students meet via video conference to present and discuss what they are working on under the supervision of an academic expert from the University. The seminars could take up a particular focus such as methodology or literature review and the international students could meet up online to present and discuss their work

Melanie, School-S.

Some received information from students who were ahead of them in their PhD study (Olivia, School-R), and it was felt that PhD students at an advanced stage of their research should act as peer-mentors to new students (Sylvia, School-L). This was already happening in some schools in the University, indicating the added value and transferability of the academic habitus of older students conveyed to newer students. This is something that other faculties can emulate by creating opportunities for PhD graduates and current PhD students to meet personally or virtually.

Employment prospects

Graduate attributes

Bryan and Guccione (2018), in their study of UK doctoral graduates, identify four areas of doctoral value, shaped by aspects during and after their doctorates. These include time since graduation; supervision; accumulated social connectivity; and value of the doctorate to the employer. On the basis of their national survey in France, Durette et al. (2016) built a framework of 111 competencies, leading to a set of "core" competencies found in doctoral graduates, regardless of discipline or date of graduation. The authors contend that PhDs develop a set of common competencies, which is not limited to knowledge and technical skills, but includes transferable competencies. These comprise competencies learned through practical experience, as well as courses, and include dispositions, behaviors and meta-competencies, giving PhD graduates the ability to adapt their competencies to different professional contexts. These competencies are also identified as "graduate attributes" in the literature.

Generic qualities acquired along with disciplinary knowledge in higher education, such as critical thinking, scholarship, initiative, team working and so forth, and jointly labeled as graduate attributes, are viewed as the *raison d'être* for higher education (Platow, 2012). PhD graduates report the acquisition of many graduate attributes before commencing their PhD, through education, debate, work experience, personal characteristics and ability, family training and responsibilities, socializing and so forth. They develop further graduate attributes during their PhD, such as in-depth knowledge of the field of study, analytical skills, critical judgment, and so on, because of supervision, interaction with and observation of peers and academics, feedback, conferences and workshops, and writing for publication (Manathunga et al., 2009).

The students in the current study acknowledged that they had developed a range of attributes during their PhD study. This included more confidence (Laura, School-S); sharper research skills and improved ability as a researcher (Dominic, School-X; Emily, School-T); analytical skills (Lee, School-T); meeting deadlines (Neville, School-T; Hazel, School-T); time management, project management, ability to take feedback (Nicola, School-R); personal and communication skills (Timothy, School-G), and how to write (Ralph, School-M). It taught them skills that led to knowledge enhancement (Adrian, School-R), organizing conferences and events, and teaching, and self-managed mentoring (Sylvia, School-L). In short, PhD was a life enhancing activity (Caroline, School-X). It is noteworthy that the final comment was from a student who had reported some negative experiences of supervision and examination, as noted above, yet, in essence, benefited from her PhD study and found it to be a productive and enriching experience.

Career opportunities

PhD competencies or graduate attributes constitute a common vocabulary between doctorate holders and employers; are regarded as a valuable tool for PhDs to impress recruiters; and highlight the suitability of doctoral training to meet the needs of the labor market (Durette et al., 2016). Adkins (2009) observes that PhD study involves the translation of knowledge to domains outside the university where it can be taken up and applied. Students in the current project believed that studying for a PhD had enhanced

their career prospects and chances of career progression, as they had gained additional employability skills (Jason, School-T), and the ability to communicate effectively with outside agencies (Callum, School-H). In particular, it had improved their critical thinking competence and in-depth research skills to prepare them for a career (Valerie, School-G; Joyce, School-G). Walker and Yoon (2017) contend that specific academic practices, attributes and behaviors, both existing and acquired, jointly form, what they regard as, doctoral capital that can lead to graduates successfully acquiring academic posts after the completion of their doctorates.

Literature shows that a high proportion of students favor university-industry links, and those funded by industry express a higher level of satisfaction for their PhD program, believing that industry funding could improve their career prospects (Harman, 2002). Some students in the current study already had links with industry, and some were jointly sponsored by the University and an external establishment:

I am a full time EU student ... The University pays half my fees and [an external organization] pays the rest. I get living expenses paid by the University. I work on two research projects which are linked to my PhD. I do up to six hours of teaching per week.

Joyce, School-G.

Such University-industry links were undoubtedly beneficial for the student, University and the industry partner. Some other faculties at the University had also forged links with industries connected to their discipline, thus enabling the students to extend their academic habitus to the wider social and employment contexts. Yet, Manathunga et al. (2009) note that while the majority of students in their study agree that skills developed during the PhD prepare them for the world of work, less than half believe that links with industry will improve their career options. These authors note the diverse fields that PhD graduates get employment in and highlight the need for the corresponding diversity of skills needed for these roles. This shows that stronger links between universities and a variety of employers can be of benefit to organizations, students and employees.

Conclusion

Reay (2004) supports the use of habitus as a conceptual tool to ensure that the research focus is wider than the specific issue under study. Listening to student voices on a range of issues is crucial to improve the curriculum, pedagogy and support mechanisms. An appreciation of the viewpoints and experiences of recent doctoral graduates is important in order to design and evaluate PhD programs more effectively on a regular basis (Manathunga et al., 2009). We argue that it is equally important to consider the perceptions of current PhD students, so that they, too, can benefit from any changes to policies and practices which result from the articulation of their insights about their doctoral program and develop their academic habitus at the same time.

Bryan and Guccione (2018) categorize four areas of doctoral value as time since graduation, supervision, social connectivity, and value to the employer. The vast majority of students in our study are highly satisfied with their supervision, with isolated cases of dissatisfaction, due to tardy feedback, or other issues, most of which were resolved by changing supervisory teams. Faculty Directors of the PhD Program and Chairs of Research Degree Committees should identify these isolated cases during the evaluation of PhD programs and act swiftly in case of problems between students and supervisors, among examination panels, or to address other issues raised by the students. Exercising transparency is crucial in such instances so that the student is fully cognizant with the reasons if a supervisor or examiner is suddenly changed due to unavoidable circumstances. A new supervisor or examiner should be au fait with the student's work before assuming the role so that the student is not disadvantaged. Feedback should be timely and helpful, to enable the student to move forward.

Starting to write early on is beneficial for doctoral students as it gives them the opportunity to critique previous research in the field and gives them confidence to begin to write at PhD level. It also leads to timely completion (Lindsay, 2015). Students acquire a range of graduate attributes, including writing at doctoral level, during their PhD, which are linked to their academic habitus. Yet, the possibility of gaining various attributes in different faculties and schools can be dissimilar. There should be opportunities for faculties to share best practice through a University-wide hub, such as a Graduate School, which is currently lacking at the University.

PhD students, in the main, can manage their own learning, though some of those studying part-time and working full-time can find it difficult. This is particularly problematic for University staff undertaking PhDs, as well as other students who have demanding jobs, when they are given a new responsibility or get promoted, leading to more time spent on work-related issues, with the PhD taking a lower priority. Such students need a combination of clear guidance and scaffolding, with well-defined short-term targets to fulfill bigger objectives to enable them to manage their own learning. In congruence with what has been reported in the literature (Castello et al., 2017), a number of mechanisms are in place at the University to ensure standards and timely completion of PhD to help students achieve a balance between personal, professional and academic life.

Completion of procedures can take a long time resulting in student frustration. Some students question the nature of policies and the duration of procedures, with some instances of extreme dissatisfaction and frustration. Progress reporting should be linked to substantive reviews of progress and embedded in the pedagogy and curriculum of doctoral programs (Mewburn et al., 2014). Completion of forms by students at various stages of the PhD should be signed off and evaluated swiftly by supervisors and

evaluators at the University. University processes and procedures need to be made simpler and students encountering problems with supervisors or examiners should receive prompt support. Students are largely satisfied with the resources and other academic support, though there is room for improvement, as there is lack of uniformity in provision of resources across faculties. Faculties can learn from the good practice of one another. More resources need to be made available to students as required, and academic support, other than supervision, can be improved in line with student expectations to enhance their academic habitus.

In our knowledge society and economy, doctoral education is increasingly considered as a means to produce workforce to meet the needs of the global labor market (Durette et al., 2016). The academic habitus of PhD students in our study, mainly developed through a combination of supervisory support and the acquisition of graduate attributes will enable them to succeed in the employment market. There is a clear and positive relationship between graduate attributes and productivity on PhD completion, emphasizing the value of these attributes and the role of the university milieu in providing attributes needed for employment (Platow, 2012).

Individuals undertake PhD study for both professional and personal development. Knowledge acquired during PhD study can be transmitted to realms outside the university (Adkins, 2009). The academic habitus acquired during their PhD program has enabled our research participants to develop not just as people and professionals, but they have also gained the ability and confidence to be catalysts of change not only in their own organizations, but also in their respective countries. Research on Chinese doctoral students (Rhoads et al., 2017) suggests these students develop knowledge and skills with the objective of facilitating higher education reform in their home country. The part-time, as well as fulltime, students in our research clearly have the potential to make a valuable contribution to the knowledge economy, if they are not already doing that, as most of them are in employment. This is due to their existing and acquired skills and attributes. According to Walker and Yoon (2017) these skills and attributes jointly constitute what they consider as doctoral capital. We believe this doctoral capital is the product of these students' academic habitus.

Literature (Gardner, 2007) emphasizes the significance of socialization in the success of doctoral students. Our study illustrates that the students have developed individual and collective habitus based on the academic, social and wider support in the University as well as their personal attributes. Yet, the academic habitus possessed by students varies between disciplines and supervisors and is not consistent across the university. Our research also highlights the added value and transferability of the academic habitus of older students transmitted to newer students, pointing to generational gain. Further, faculties who have forged links with industries related to their discipline have enabled their students to extend their academic habitus to wider social and employment contexts.

Delamont et al. (1997) argue that students' doctorates are required to extend the traditions of their discipline. Intellectual continuity is thus accomplished through one-to-one associations, with students supervised with continuity in the same tradition by the same supervisors. This tradition illustrates a specific academic habitus developed through such intellectual relationships and internalized by the students. A similar academic habitus is epitomized by PhD students in our study. All our students are either progressing well, are close to completion, or have recently submitted. These students exemplify a specific academic habitus, which can only be found in highly intellectual and successful individuals who have the potential to improve not only their own professional and personal lives, but also to make a significant intellectual contribution to societal good.

References

- Adkins, B., 2009. PhD pedagogy and the changing knowledge landscapes of universities. *High Educ. Res. Dev.* 28 (2), 165–177.
- Basit, T.N., Eardley, A., Borup, R., Shah, H., Slack, K., Hughes, A., 2015. Higher education institutions and work-based learning in the UK: employer engagement within a tripartite relationship. *High Educ.* 70 (6), 1003–1015.
- Basit, T.N., 2010. *Conducting Research in Educational Contexts*. Continuum, London.
- Basit, T.N., 2012. But that's just the stereotype: gender and ethnicity in transition to adulthood. *Race Ethn. Educ.* 15 (3), 405–423.
- Boud, D., Lee, A., 2005. "Peer learning" as pedagogic discourse for research education. *Stud. High Educ.* 30 (5), 501–516.
- Bourdieu, P., 1988. *Homo Academicus*. Trans. Peter Collier. Polity, Cambridge.
- Bourdieu, P., 1990. In *Other Words: Essays Towards a Reflexive Sociology*. Trans. Matthew Adamson. Polity, Cambridge.
- Bourdieu, P., 1993. *Sociology in Question*. Trans. Richard Nice. Sage, London.
- Brodin, E.M., 2018. The stifling silence around scholarly creativity in doctoral education: experiences of students and supervisors in four disciplines. *High Educ.* 75 (4), 655–673.
- Bryan, B., Guccione, K., 2018. Was it worth it? A qualitative exploration into graduate perceptions of doctoral value. *High Educ. Res. Dev.* 37 (6), 1124–1140.
- Can, G., Walker, A., 2014. Social Science doctoral students' needs and preferences for written feedback. *High Educ.* 68 (2), 303–318.
- Casey, B.H., 2009. The economic contribution of PhDs. *J. High Educ. Pol. Manag.* 31 (3), 219–227.
- Castello, M., Pardo, M., Sala-Bubare, A., Sune-Soler, N., 2017. Why do students consider dropping out of doctoral degrees? Institutional and personal factors. *High Educ.* 74 (6), 1053–1068.
- Cheng, M., Taylor, J., Williams, J., Tong, K., 2016. Student satisfaction and perceptions of quality: testing the linkages for PhD students. *High Educ. Res. Dev.* 35 (6), 1153–1166.
- Cotterall, S., 2011. Doctoral pedagogy: what do international PhD students in Australia think about it? *Pertanika J. Soc. Sci. Humanit.* 19 (2), 521–534.
- Cree, V.E., 2012. "I'd like to call you my mother": reflections on supervising international PhD students in social work. *Soc. Work. Educ.* 31 (4), 451–464.
- Curtin, N., Stewart, A.J., Ostrove, J.M., 2013. Fostering academic self-concept: advisor support and sense of belonging among international and domestic graduate students. *Am. Educ. Res. J.* 50 (1), 108–137.
- Delamont, S., Parry, O., Atkinson, P., 1997. Critical mass and pedagogic continuity: studies in academic habitus. *Br. J. Sociol. Educ.* 18 (4), 533–549.
- Durette, B., Fournier, M., Lafon, M., 2016. The core competencies of PhDs. *Stud. High Educ.* 41 (8), 1355–1370.
- Fomunyan, K.G., 2020. Theorizing intercontinental PhD students' experiences: the case of students from Africa and Asia. *Int. J. High. Educ.* 99 (3), 232–239.
- Gardner, S.K., 2007. "I heard it through the grapevine": doctoral student socialization in chemistry and history. *High Educ.* 54 (5), 723–740.
- Gopaul, B., 2011. Distinction in doctoral education: using Bourdieu's tools to assess the socialization of doctoral students. *Equity Excell. Educ.* 44 (1), 10–21.

- Griffioen, D.M.E., Doppenberg, J.J., Oostdam, R.J., 2018. Are more able students in higher education less easy to satisfy? *High Educ.* 75 (5), 891–907.
- Gruzdev, I., Terentev, E., Dzhafarova, Z., 2020. Superhero or hands-off supervisor? An empirical categorization of PhD supervision styles and student satisfaction in Russian universities. *High Educ.* 79 (5), 773–789.
- Harman, G., 2002. Producing PhD graduates in Australia for the knowledge economy. *High Educ. Res. Dev.* 21 (2), 179–190.
- Hemmati, R., Mahdie, A., 2020. Iranian PhD students' experiences of their learning environment and scholarly condition: a grounded theory study. *Stud. High Educ.* 45 (1), 187–208.
- Janta, H., Lugosi, P., Brown, L., 2014. Coping with loneliness: a netnographic study of doctoral students. *J. Furth. High. Educ.* 38 (4), 553–571.
- Juniper, B., Walsh, E., Richardson, A., Morley, B., 2012. A new approach to evaluating the wellbeing of PhD research students. *Assess. Eval. High Educ.* 37 (5), 563–576.
- Lee, J., Anderson, A., Burnett, G., 2017. Peer relationships and mentoring between LIS doctoral students: a qualitative approach. *J. Librarian. Inf. Sci.* 49 (1), 115–124.
- Lindsay, S., 2015. What works for doctoral students in completing their thesis? *Teach. High. Educ.* 20 (2), 183–196.
- Maloshonok, N., Terentev, E., 2019. National barriers to the completion of doctoral programs at Russian universities. *High Educ.* 77 (2), 195–211.
- Manathunga, C., Pitt, R., Critchley, C., 2009. Graduate attribute development and employment outcomes: tracking PhD graduates. *Assess. Eval. High Educ.* 34 (1), 91–103.
- Matos, F., 2013. PhD and the manager's dream: professionalizing the students, the degree and the supervisors? *J. High Educ. Pol. Manag.* 35 (6), 626–638.
- Mewburn, I., Tokareva, E., Cuthbert, D., Sinclair, J., Barnacle, R., 2014. “These are issues that should not be raised in black and white”: the culture of progress reporting and the doctorate. *High Educ. Res. Dev.* 33 (3), 510–522.
- Mowbray, S., Halse, C., 2010. The purpose of the PhD: theorizing the skills acquired by students. *High Educ. Res. Dev.* 29 (6), 653–664.
- Moxham, L., Dwyer, T., Reid-Searl, K., 2013. Articulating expectations for PhD candidature upon commencement: ensuring supervisor/student “best fit”. *J. High Educ. Pol. Manag.* 35 (4), 345–354.
- Owler, K., 2010. “A problem to be managed”? Completing a PhD in the arts and humanities. *Arts Humanit. High Educ.* 9 (3), 289–304.
- Platow, M.J., 2012. PhD experience and subsequent outcomes: a look at self-perceptions of acquired graduate attributes and supervisor support. *Stud. High Educ.* 37 (1), 103–118.
- Reay, D., 2004. “It's all becoming a habitus”: beyond the habitual use of habitus in educational research. *Br. J. Sociol. Educ.* 25 (4), 431–444.
- Rhoads, R.A., Zheng, M., Sun, X., 2017. The methodological socialization of social science doctoral students in China and the USA. *High Educ.* 73 (2), 335–351.
- Sharmini, S., Spronken-Smith, R., 2020. The PhD—is it out of alignment? *High Educ. Res. Dev.* 39 (4), 821–833.
- Son, J.-B., Park, S.-S., 2014. Academic experiences of international PhD students in Australian higher education: from an EAP program to a PhD program. *Int. J. Pedagog. Learn.* 9 (1), 26–37.
- Walker, J., Yoon, E.S., 2017. Becoming an academic: the role of doctoral capital in the field of education. *High Educ. Res. Dev.* 36 (2), 401–415.
- Wei, J., Carter, S., Laurs, D., 2019. Handling the loss of innocence: first-time exchange of writing and feedback in doctoral supervision. *High Educ. Res. Dev.* 38 (1), 157–169.
- Wisker, G., Robinson, G., 2013. Doctoral “orphans”: nurturing and supporting the success of postgraduates who have lost their supervisors. *High Educ. Res. Dev.* 32 (2), 300–313.
- Zhou, J., 2015. International students' motivation to pursue and complete a PhD in the US. *High Educ.* 69 (5), 719–733.

Pathways, labor markets and tertiary education systems

Roger McL. Harris^a and **Hugh B. Guthrie^b**, ^aCentre for Research in Educational and Social Inclusion, UniSA Education Futures, University of South Australia, Mawson Lakes, SA, Australia; and ^bMelbourne Graduate School of Education, University of Melbourne, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	438
Labor markets and skill formation systems	439
Examining tertiary education pathways	440
Typologies of educational pathways	441
Pathways between educational sectors	441
Pathways from education into the labor market	441
Pathways through the lifespan	441
A case study: Australia	443
The Australian labor market	443
Careers advice and its quality in Australia	444
The Australian tertiary system	444
Pathways into and within Australia's tertiary education	445
Pathways into tertiary education	445
Pathways within tertiary education	445
Issues in developing a more coherent Australian tertiary education system	446
Approaches to tertiary funding and effects on pathways	446
Perceptions of VET and higher education and effects on pathways	447
Approaches to revitalize Australian VET	447
Australian Qualifications Framework and effects on pathways	447
Conclusion	448
References	449

Introduction

Governments have for some time been extolling the virtues of “seamless” education systems. They perceive them as key organisms in their efforts to maximize learning opportunities within their twin agendas of skill development and social equity, and so there is keen interest in arrangements that either facilitate or hinder transfer between educational sectors and from these sectors into employment. Attainment of the reality, however, has proved more elusive than the rhetoric. Many studies have recognized that there are challenges, as “effective transitions and articulation with other sectors of education and training are often elusive” (OECD, 2014, p. 33). Young (2006, p. 5) even raises the question of whether “seamlessness is at the heart anti-educational because it fails to take account of the specificities of different types of learning and knowledge”. Others allude to the power-plays that can occur in moves to draw together more closely, for example, the academic and vocational or the secondary and tertiary sectors, as the culture of the respective academic and higher education (HE) sectors invariably predominates. Moreover, as Bathmaker (2017, p. 4) states, hybrid or “dual” routes at higher levels, while proving increasingly desirable to high achievers, tend to become colonized by those from advantaged backgrounds, squeezing out the disadvantaged and underserved. Yet attempts to blend “the worlds of education and of work via continued reform of vocational education and training (VET) systems has become a key priority of European Union policies, especially in the aftermath of the 2008 global crisis” (Pouliakas and Ranieri, 2018, p. 491) and more recently the Covid-19 pandemic which has presented unprecedented challenges to national VET systems (GOVET, 2022; OECD, 2021).

Prevalent in political discourse is the notion of educational pathways. Over the years, there have been common themes in policy documents and other literature concerned with people's transitions through their life and career course—the need for ease of movement and for pathways and cross-sectoral collaboration, as well as for career guidance to assist transition (Céreq International Mission, 2020; CareersNet, 2020; Cedefop, 2017; OECD, 2014, 2000). The OECD (2020) selected VET as its theme for its *Education at a glance 2020* because of the “renewed interest in VET in recent years as an effective means to open pathways for further learning and personal growth, and ease the transition of learners into the labor market” (p. 15). A popular approach has been the construction of national qualification frameworks, designed to take account of a range of academic and vocational credentials that would cater for individual learner needs and be credible and understood generally. Opening pathways in this way implies take-up of recognition of prior learning (RPL) and credit transfer as well as removal of “time-served” aspects in qualifications. Another policy response has been the deregulation of tertiary education. This has involved stimulation of a market model, where private providers proliferate and public educational institutions are encouraged to amalgamate and/or re-invent themselves. Nevertheless, such a commodification of education has its limitations, especially when barriers to market entry are set too low and not adequately

regulated and monitored (Toner, 2014, 2018). Yet competition has become a key ingredient in this educational maelstrom, with the anticipation of institutional responsiveness and learner choice. Other measures have included more flexible modes of study, such as part-time and modular arrangements, distance and blended learning, and competence-based approaches, to enable pathways to be more flexible and cater more adequately for needs of learners.

The OECD (2000, p. 13) identified six key elements of successful transition systems: a healthy economy; well-organized pathways that connect initial education with work and further study; opportunities to combine workplace experience with education; safety nets for those at risk; sound information and guidance; and effective institutions and processes. These ingredients represent a helpful litmus test for reflecting on situations in different countries, and the degree to which educational systems facilitate these criteria. Any discussion, therefore, involving pathways must inevitably also consider context—the nature of educational systems, particularly skills formation, in which they are embedded and of labor markets with which they are intricately linked.

In this chapter, the first section draws on the international literature to examine the main issues and debates in the areas of pathways and their twin contexts of labor markets and skill formation (education) systems. There exists a vast quantum of literature on pathways, and here only the most recent research is reviewed in order to furnish the most up-to-date portrayal. The second section then focuses on the Australian situation as its case study to illustrate types of pathways, how they operate in practice and related issues that impact upon them, as Australia can provide an informative and balancing case. For not only has the bulk of literature come from European perspectives (Caves et al., 2021), it has also focused on school-based VET systems and been preoccupied with young people's entrances into labor markets (Hanushek et al., 2017; Hodkinson and Bloomer, 2002). In contrast, the Australian skills system is largely post-secondary (though some lower level VET occurs in schools¹) and focused more on whole of life and careers. As post-secondary vocational education has been characterized by the OECD (2014) as a "hidden world" (p. 11) and "inadequately understood" (p. 22), focusing on pathways in Australia is worthy of examination.

Labor markets and skill formation systems

Pathways need to be considered within their context. However, the issue here is that in the literature there are various typologies of labor markets and skill formation systems depending upon what criteria are employed. At the broadest level, they are often distinguished as either liberal market economies (LMEs), such as the United States, United Kingdom, Canada and Australia, or coordinated market economies (CMEs), such as Germany, Austria and Switzerland (Fortwengel et al., 2019; Vossiek, 2018). In the view of Fortwengel et al. (2019), the key differentiating feature is the extent to which they experience difficulty in sustaining institutions which support apprenticeship training systems; companies in LMEs rely largely on recruiting skilled workers from external labor markets and/or training recruits in firm-specific skills on-the-job, while those in CMEs rely more on apprenticeships to recruit and train their workforce. Early work distinguished between high-skill and low-skill systems (Finegold and Soskice, 1988), and various writers appropriated different criteria for classifying systems, such as the role of the state in skill provision (Greiner, 1993), the predominant venue for training – institutions or workplaces (Lynch, 1994) or the specificity of skills provided (Hall and Soskice, 2001).

More recently, Bussemeyer and Trambusch (2012, pp. 11–12) identified four different skill formation systems: liberal (US, Ireland), segmentalist (Japan), statist (France, Sweden) and collective (Germany, Netherlands, Denmark, Austria, Switzerland). For these authors, the issue was the varying degrees of company and state involvement in training. A further refinement by Pilz (2016, 2017) adopted a multi-perspective approach, expanding on previous attempts by, firstly, combining skill formation, stratification and standardization approaches, and secondly, incorporating the micro-level criterion of the practice of learning. It thus attempts to add to the debate by countering the criticism made of typologies that criteria are restricted and often incoherent. Pilz's typology analyses the first seven countries in Table 1 and we have added others. Of particular relevance to this chapter and further discussed later, the Australian skill formation system is claimed to be industry-led, though in reality it is heavily government controlled and regulated, largely takes place in post-secondary institutes and independent providers (and at a lower level in secondary schools separated from general education subjects), is standardized by the Australian Qualifications Framework and Training Packages, and its teaching and learning is closely linked to vocational practice via competency-based training and assessments in workplaces.

From analyses such as these, one notes that the debate centers on the degree of "fit" between the educational system and the labor market—it is far tighter in CMEs than in LMEs. This holds significant implications for pathways, not only when individuals enter employment but through their careers. Where there is closer coupling between education and labor market, tertiary education more clearly regulates individuals' access to particular occupations. Where there is looser connection, this relationship is characterized by more intensive competition, lower employment tenure and market rules. In such contexts, there is greater potential for displacement between levels of education and occupational position, and individuals are likely to perceive a potential mismatch between their qualifications and returns in the job market (Tomlinson, 2012), discussed later in the chapter in the case of Australia. As Roberts (2020, p. 27) comments, governments in LMEs cannot impose transition regimes in total and in detail, and thus the types of education that governments will fund, including training provisions together with employment regulations, "set the context

¹The recent Australian review by the Education Council, *Looking to the future* (2020), concluded that "VET delivered to senior secondary students is of inconsistent quality, difficult to navigate, and not well integrated into senior secondary studies" (p. 76). It recommended that: "VET should only be provided at school where it can be done in a high-quality way as demonstrated through the explicit endorsement of local employers or industry bodies ..." (p. 21).

Table 1 Typology of VET systems.

<i>Countries</i>	<i>Skill formation</i>	<i>Stratification</i>	<i>Standardization</i>	<i>Practice of learning</i>
USA	Individualized	Low	Low	High
France	State	High	High	Low
Germany	State/company	High	High	High
Japan	Company	High	High	High
India	Individualized	High	Low	High
China	State	High	High	Low
Mexico	Individualized	High	Low	High
Australia	State/company	Low	High	High
Singapore	State	High	High	Low
UK	Individualized	Low	Low	High
Canada	Individualized	Low	Low	High

*Skill formation: the influence of the state on VET and the potential for activity by and influence of companies. *Stratification: issues of “tracking” and the separation of general training courses and vocational courses. *Standardization: how structures and processes underpinning VET are standardized and subject to regulation. *Practice of learning: relevance of teaching and learning processes to vocational practice or later roles in employment.

Adapted from Pilz (2016, 2017).

in which educational actors, employers, training providers and young people pursue their own agendas”. In LMEs links between qualifications and occupations are quite weak with the result that skills gaps and skills mismatches can exist (Hughes and Smith, 2020; Puckett et al., 2020; Karmel et al., 2008). The key issue is the relationship between educational systems and the labor market. It is often assumed that education is crucial for meeting the challenges concerning the future of work. But in the view of Wheelahan and Moodie (2017), variations in the role of vocational education in preparing graduates for employment and educational pathways mostly lie not in the nature of vocational education nor its qualifications, but more in the labor market structure and how employers use qualifications to select graduates and progress in employment.

In this respect, debate continues over the development of micro-credentials and short courses. Europe (Lantero et al., 2021), Canada (Colleges and Institutes Canada, 2021), Australia (Australian Government, 2021) and New Zealand (NZQA, 2022) are currently considering their value. They are seen as helpful for industry and the economy to recover from the Covid pandemic and respond to the needs of the digital age, with the increasing need for upskilling and reskilling the labor force, and the emphasis on student-centered learning and flexible learning paths. They can provide training that is accessible to vulnerable groups who have been under-represented in education, and their flexibility allows learners to track their own pathways; for example, arranging schedules to take into consideration childcare or eldercare as well as their work—such flexibility can be particularly advantageous for women learners. However, in a background paper commissioned by UNESCO, Buchanan et al. (2020) contends that focusing on reforms like micro-credentials and 21st century skills to achieve employment growth is actually a distraction from what successful education systems can achieve. In his view, education cannot make up for inadequacies in other policy domains that have caused and continue to cause declining job quality as well as mass unemployment and under-employment. His plea is for education to be brought back into the debate as important *per se*, rather than as an auxiliary policy aimed at fixing flaws of other policy domains.

Examining tertiary education pathways

Review of the literature reflects that much policy attention has focused on institutional architecture for facilitating movement, while there has been far less consideration of learners’ experiences in their educational journeys and how they make meaning of their transitions. Both structure and agency need to be considered for meaningful understanding of youth transitions (Roberts, 2020; Schoon and Lyons-Amos, 2017).

Interrogated from this perspective, such policy promotion is debatable in that it may be perceived as both positive and problematic. On one hand, increasing “seamlessness” and facilitating pathways can provide considerable choice for learners as they wrestle with challenges of learning and work. On the other, they can lead to a climate of uncertainty and indecision, where learners struggle to make meaning of the plethora of possibilities, grapple with decisions over appropriate ways ahead and discover later that less than optimal selections can be both financially draining and time inefficient. Policy that promotes unlimited choice has tremendous potential, yet necessarily makes assumptions about (a) individuals having the requisite “personal capital” (agency) to take advantage of alternatives, (b) the usefulness and timeliness of career information and services, and (c) the availability and

accessibility of courses and providers especially in rural locations and socio-economically disadvantaged areas. Making choices is thus viewed as an unproblematic process where students can engage freely and actively. This view emerges from the stable of rational choice theory, whereas prospect theory, on the other hand, can offer a more realistic alternative by recognizing that individuals' capacity for rationality is limited (Brown, 2017). Furthermore, such policy promotion may also be under-estimating the subtle though powerful influence of sectoral and institutional cultures, and their power relations that can work against smooth transitioning of learners (Bathmaker and Thomas, 2009) as well as inducements that may be offered to make particular choices (Toner, 2014). As Crump (2007) argued, "there is very little evidence-based research on choice and flexibility in further and higher education compared to a wealth of research about school choice", mainly because the option of choice beyond schooling has been "seen (falsely) as self-evident" (p. 4). Educational pathways of learners, therefore, are perhaps not and cannot be as clear or seamless as policy might desire.

Moreover, the very concept of pathway can be contentious as it reflects a policymaking perspective, focusing on structural considerations that are directly under the control of policy makers. It therefore needs to be balanced with the complementary concepts of "navigations", emphasizing more the perspectives of individuals as they strive to manage their own itineraries, and "trajectories", highlighting the social forces that shape transitions and that individuals making transitions vary in their resources (Evans and Furlong, 1997). While Raffe (2003) argues that pathway can be a useful organizing concept, he does conclude that official maps of pathways must (a) correspond to the ways in which they are actually experienced, (b) allow for different types of transitions, being flexibly designed to accommodate an array of events through life that affect transitions in and out of work and education, and (c) incorporate clear signposts for individuals learning how to navigate them.

Typologies of educational pathways

There have been a considerable number of studies researching educational pathways and "transition trajectories" (Keskiner, 2019, p. 107) in recent international literature—only a few examples can be mentioned here. Some have focused on movements between educational sectors, others on transitions from education into the labor market and still others on pathways through the lifespan.

Pathways between educational sectors

Harris and Rainey (2012) identified five pathways in Australia's tertiary education system in their study of two-way movement of students between the HE and VET sectors. These were: erratic pathway (pursued by interest-chasers); flowing pathway (developers); merging pathway (combiners); tangential pathway (forced learners); and parallel pathway (two trackers). Studying the post-compulsory educational pathways of young people in England who had spent some or all of their childhoods in local authority care, Cameron (2018) singled out six pathways, named as: Traditional A levelers, Alternative traditional, Ordinary young lives, Futures on hold, Home-based locals and Building a life. Only the first two were secure in their orientation to their own futures. For the others, structural factors such as access to citizenship or the right to work, or employment, or familial or individual factors such as health and caring commitments, acted as barriers to their ambitions.

Pathways from education into the labor market

By far the most studies focus on this transition. Recent research (Ranasinghe et al., 2019) on school-to-work transitions of Australian youth identified five pathways: HE and work; early entry to full-time work; mix of HE and VET; mixed and repeatedly disengaged; and mostly working part-time. These are similar to those of Roberts (2020, pp. 37–38) who identified that there were five main transition routes in the UK: university into elite careers; university into the lower middle class; exit education at 18 and enter the lower middle class, possibly via an apprenticeship; exit education at 18 and enter the upper working class, possibly via an apprenticeship; fail in education and join the precariat. Also in the UK, Schoon and Lyons-Amos (2017), drawing on data from the Longitudinal Study of Young People, identified six distinct pathways: an academic track, three pathways involving further education and training, a work-focused transition group and a group who over a long period were not in education or training. In Germany, Brzinsky-Fay and Solga (2016) studied school-to-work-transition patterns for five birth cohorts. While they found that traditional pathways (general education-tertiary education-into employment) increased over the cohorts, so too did non-linear pathways (status shifts, and some non-employment experience).

Pathways through the lifespan

Fry and Boulton (2013), drawing on *Household, Income and Labor Dynamics in Australia* (HILDA) survey data in Australia, identified 17 pathways through a lifespan. Different pathways can arise with key life events—for those aged 15–24 (five pathways), aged 25–39 (four pathways), aged 40–54 (four pathways) and aged 55–64 (four pathways).

There are a number of ongoing debates on pathway typologies. They may be clustered into two broad categories: the nature of pathways and the factors that influence them. In the first category, the nature of pathways, are embedded questions about whether patterns are linear (traditional) or divergent (alternative), and what they actually look like (e.g., Keskiner, 2019). The traditional

pattern is generally depicted as a smooth trajectory that follows institutionally predictable, linear pathways, while the divergent path is where learners experience periods of study in different sectors, unemployment, shift between different employment statuses or change direction. Inherent in this debate is the role of agency versus structure—whether shifting transitions are initiated by learners' desires to experience alternative pathways or by environmental conditions forcing them into precarious situations. Schoon and Lyons-Amos (2017) contend that study of structural constraints and agent roles are both essential.

In the second category are factors that affect pathways. Some researchers cite studying VET subjects at school, individual school achievement and socio-economic issues as key influences (e.g., Ranasinghe et al., 2019). Others report the substantive influences of parents and friends on the educational expectations of secondary school learners (Roth, 2017), although Billett et al. (2020) report that while parents and school-teachers may under-estimate their influence on learners' decision-making, their advice is primarily based on their own experiences, rather than more broadly informed and more impartial considerations. In this regard, tracking is an important issue that is often researched. In what Bol and van de Werfhorst (2013) call "a trade-off between equality and labor market preparation", they acknowledge that systems with a high level of tracking and vocational focus have been shown to improve the allocation of school-leavers in the labor market, but also to increase inequality of educational opportunity. Related to this debate is the relative value of vocational vis-à-vis general qualifications in transitions. Many researchers have found that vocational qualifications can help more than general qualifications in gaining more rapid entry into the labor market, but are soon overtaken by their counterparts with a general education (OECD, 2020; Middeldorp et al., 2019; Roosmaa et al., 2019), mainly due to occupation-specific skills making them less flexible. Hanushek et al. (2017) refer to this as another "trade-off", for which they found strong and robust support in analyzing data from 11 countries, where gains in youth employment can be offset by less adaptability and diminished employment later in life. Thus, potential short-term advantages of vocational qualifications are likely to be offset by disadvantages in later life, what Forster et al. (2016) label "the life-cycle effect" (p. 474) and found to be independent of the strength of the vocational system.

The evidence indicates that transitions are "problematic and protracted" (Alam and de Diego, 2019) and that youth integration into the labor market differs considerably across countries (e.g., Ryan, 2001; Wolbers, 2007). Based on their analysis of international practice from 105 reports, Hughes and Smith (2020, pp. 32–43) identified seven key guiding principles which policymakers should use when shaping policy and practice: start early; support teaching and learning; embed careers firmly in the curriculum and school-to-work transitions; encourage effective career and transition dialog; engage employers; utilize technology and labor market information, and importantly, make use of evidence to strengthen skills and create new pathways.

From this complexity at least three common and important threads may be discerned in the literature. One is the prevalence of non-linear pathways, even in highly stratified systems as in Germany (Roth, 2017; Brzinsky-Fay and Solga, 2016). Harris and Rainey (2012) characterized this non-linearity as crazy paving rather than orderly stepping-stones. Whether or not these non-linear pathways are inefficient, especially in national systems where links between education and labor market are loose, is debatable. Certainly, politicians are fond of expressing such sentiments (see examples in Vaughan, 2005; Karp, 2020). Many studies have drawn attention to the phenomenon of indirect transfer where movement of tertiary students involves several moves within and between institutions, sectors and workplaces. This phenomenon has been variously described as "milling" (OECD, 2000, p. 76), "churning" (Fry and Boulton, 2013, p. 18), "swirling" (de los Santos and Sutton, 2012, p. 967), "the community college shuffle" (Maxwell et al., 2002, p. 1) and, most recently, "meandering" (Roberts, 2020, p. 37). Considerable research in various countries including Australia (Harris and Rainey, 2009, 2012), the UK (Tomlinson, 2012) and Singapore (Harris and Ramos, 2012, 2013) has demonstrated high levels of learner movement between sectors, highlighting the complexity in the phenomenon of pathways both within and between tertiary education and employment. Research has also shown that patterns in learner movement vary markedly by sector and education field and that these variations can be explained by labor market and regulatory structures in occupations (Fredman, 2013). Yet many policy settings assume pathways occur between the same fields and that the internal mechanics of tertiary education alone can be adjusted to improve transfers. As Tomlinson (2012, p. 419) concludes:

The themes of risk and individualization map strongly onto the transition from [tertiary] education to the labor market ... The transition ... is perceived to be a potentially hazardous one that needs to be negotiated with more astute planning, preparation and foresight.

A second thread is the importance of individual wherewithal, or personal capital, so critical as to whether pathway opportunities are taken up and pursued. Navigating one's career usually involves moving between familiar and unfamiliar environments, often referred to as boundary-crossing. Such transitions, as argued through this chapter, can be complex. In the career navigation process, a key issue is where the locus of control lies – externally or internally? Global forces make for "boundaryless careers" (Arthur, 2014), characterized by uncertainty and flexibility, where individuals search for meaning in their continuously changing career trajectories. While the boundaryless career concept highlights the shifting nature of job markets external to the individual, the concept of "protean career" (Gubler et al., 2014) focuses more on the internal aspects where individuals strive to take charge of their careers and make sense of their environment without external guidance. A recent study in Australia, summarizing the 10-year transition pathways of 3186 individuals aged between 16 and 25, demonstrated that "young people's post-school pathways are diverse, individualized and complex, underscoring the importance of understanding youth transitions as a process" (Ranasinghe et al., 2019, p. 9). This suggests that learners with fewer personal resources and lower competence may be unlikely to be able to adjust satisfactorily. As Solomon (2005, p. 105) concluded:

... while boundary crossing between working and learning may be desirable in policy and rhetorical terms, it is neither uncomplicated nor unproblematic. Yet the conflation of work and learning is discussed as if it were natural and transparent. ... we can't ignore the complexities the conflation ... produces, nor ignore that fact that these complexities comprise intricate identity issues.

While pathways between education and training sectors may have improved over time, debates remain over the extent to which barriers still exist, such as funding, inadequate credit recognition and transfer arrangements, and administrative issues. Studies frequently focus on such matters as articulation and curriculum issues between institutions, blurring sectoral boundaries and differences in accreditation arrangements. It is unlikely, however, that the solution is as simple as resolving credit transfer and administrative processes. The reverse side of the coin is the extent to which individuals have the choice of exercising control over their own itineraries, or are influenced by the status of particular pathways; for example, school to higher education being privileged over school to vocational education studies and apprenticeships. Considerable research illustrates that such pathways tend to be erratic rather than linear or predictable, and rarely the products of rationally determined choice, as they are inextricably linked with other life experiences and personal circumstances. Our principal argument here is that the importance of personal agency in transitioning, what [Heinz \(2009, p. 391\)](#) calls “planful and adaptive competence”, has been under-emphasized given the preoccupation with institutional responses to government agendas.

A third thread is that transition could have been enhanced with targeted, accessible and accurate information and more career advice. Professional education and training occur at career crossroads, “leading on into different careers and further learning opportunities, [and] crossroads need good pathways and clear signposts” ([OECD, 2014, p. 89](#)). It can play such a critical role in helping young people navigate the different needs and pressures driving their decision-making process, raising awareness of options available and signposting new occupational opportunities constantly emerging in times of rapid change ([Agbenyo and Collett, 2014](#)). However, too often career guidance can be inaccurate and incomplete. For example, in the UK, research found that further education students considered the advice and guidance in their college was often too broad, generic and insufficiently tailored to individual needs ([Quyoun et al., 2021](#)).

In summary, one notes from the recent international literature many issues and ongoing debates. There are questions about the policy promotion of seamlessness, the nature of pathways—whether linear or divergent, the role of agency versus structure, what factors influence transitions, and the extent to which barriers remain. Other debates relate to the degree of fit between educational systems and labor markets, whether pathways are contributors to inefficiency, and the relative value of vocational vis-à-vis general qualifications in transitions and over a life-course. What is clear is that “new pathways to success are needed, as traditional routes to good livelihoods are no longer enough” ([Hughes and Smith, 2020, p. 8](#)). One of the key guiding principles when shaping policy and practice, according to [Hughes and Smith \(2020\)](#), is to make use of evidence to strengthen skills and create new pathways. Our main argument through this review of the key issues and debates has been that pathways and transitions are indeed complex, need to be analyzed within their context of skill formation systems and labor markets, and that in their interpretation there has been an over-emphasis on structure rather than agency. Given that much of the research evidence emanates from Europe and concentrates on young people’s entrances into labor markets, some different perspectives on pathways and their issues may be gleaned through examining the Australian situation where the skills formation system is largely post-secondary and focused more on whole of life and careers. Our discussion, therefore, turns to the Australian context to illustrate how the issues and debates highlighted in this review play out in that nation.

A case study: Australia

Australia is a federation comprising a national Government, six states and two territories. Its population is around 25.75 million. Many of the key issues related to its labor markets, approaches to career navigation and trajectories and the role of the tertiary sector in these have already been broadly highlighted in the literature cited in the first part of this paper. One of the issues in determining the effectiveness of pathways and their seamless use, therefore, is the way a country’s education sectors actually operate which, in turn, is affected by their design, funding arrangements and the relative status of their general and vocational education components. Another important one is the ways individuals actually use the education systems and the wide range of available transitions and pathways - their use of “agency”—a point strongly made earlier in this chapter drawing on the literature.

In this case study we examine the range of the enablers and barriers to pathways within and between labor market segments and most particularly into, out of and within Australia’s tertiary education sector given that post-secondary education is “where the action is”. It first considers the Australian labor market and the ways in which careers advice and other factors affect the availability and uptake of tertiary studies and pathways. This is followed by a consideration of Australia’s tertiary education system and the extent to which it enables or prevents effective responses to labor market needs and the ready availability of educational pathways. Another issue relates to the extent to which pathways are enabled, or not, by qualification frameworks (in this case, the Australian Qualifications Framework—AQF).

The Australian labor market

A 30-year perspective on the Australian labor market published by the Commonwealth Department of Employment, Skills, Small and Family Business ([Australian Government, 2019](#)), showed that, like other developed economies, significant structural change has occurred in its labor market, with a move away from production industries to more labor-intensive service industries and to higher level occupations with an increasing level of educational attainment being required and with a shift to a skills base that is not easily automatable. Such change has been accompanied by young people remaining in education for longer, increased female

participation in work and a move to flexible, part-time and more tenuous employment and an aging workforce with a rise in mature-aged participation even while transitioning to retirement.

Australia maintains skill shortage lists at both national and state and territory level. Persistent shortages are often addressed by a mixture of migration, education and retraining and by the use of data gathered by such agencies as the newly established National Skills Commission. The federal and state and territory governments attempt to encourage the take-up of education and training in response to these skill shortages, often through financial inducements such as fee-free VET courses. In June 2020, the (then) Australian Government announced a rise of 113% in the cost of studying humanities degrees and 28% for law and commerce, in order to fund an expansion of 39,000 university places and cheaper degrees for those studying courses in demand in the Australian economy such as teaching, nursing, maths, science and engineering (Karp, 2020). It is as yet unclear how successful this attempt at market engineering will be in effecting change in course choices and pathways. In the last few months, there has been a change of federal government, and it is still unclear how that will play out in terms of policy change and emphasis.

However, what is known is that tertiary study and course choice are affected by many factors: interest and passion; advice from significant others (e.g., parents, schools, careers advisors, workplaces and friends); quality of information about occupations and prospects; life circumstances; timetables, fees and affordability; possible sites of delivery or availability of flexible study options; and ease of access to the institution and chosen course of study (Brown, 2017). Thus, choice of career and chosen study programs can, in fact, be somewhat limited by circumstance and do not necessarily represent a rational or most desired decision but rather one that is more pragmatic, based on imperfect advice and information and the ready availability, or lack, of the desired course of study at decision time.

Careers advice and its quality in Australia

A range of Australian reports have called for well-targeted, accessible and accurate information and career advice, particularly for those moving into tertiary studies from secondary school, into and within the workforce or between VET and HE (e.g., BCA, 2017; Parliament of Victoria, 2018; Joyce, 2019; Education Council, 2020). The review by Joyce (2019) recommended the formation of a National Careers Institute. However, despite the important role careers advice and advisors play at this critical life period in connecting senior secondary students to further education and employment opportunities, many students feel they do not receive comprehensive and unbiased advice with many also feeling that they were pressured by schools and parents to pursue a university pathway, with VET being a less well-understood and preferred option (Shipley and Stubbley, 2018).

A recent review of post-school pathways recommended that students and their parents need to have access to earlier, more regular and intensive career advice informed by knowledge of the future labor market and delivered both face-to-face and online (Education Council, 2020). Thus, the focus of research in Australia has persisted on young people rather than those who are older, in mid-career or transitioning to other opportunities, either through personal ambition or circumstances such as redundancy or unemployment. Essentially, however, the process requires closer matching of skills to industry and individual needs and better information about emerging opportunities and trends as careers advice is important through an individual's life course and needs to be "lifelong and lifewide" (Guthrie and Nechvoglod, 2011, p. 16). Indeed, the Careers Industry Council of Australia had argued earlier that:

Well organized career development services have a significant contribution to make to the effective operation of Australia's education systems, to the efficient operation of its labor market, and to the effective development and use of the nation's talent and skills.

CICA (2007, p. 6).

Nevertheless, careers advice suffers from information asymmetry, and this is especially so for those in disadvantage (Bretherton, 2011). Guthrie and Nechvoglod (2011) advocated improvement in the quality of careers advice in Australia by promoting career management skills as a fundamental employability skill for all individuals; developing well-integrated, well-publicized, high quality, comprehensive, readily useable and accessible career information; providing a cascading level of service delivery particularly targeted at the most vulnerable "at risk" groups and providing genuine and comprehensive wrap-around services. In addition, it requires appropriate initial and ongoing training and professional development for career development professionals and others, better approaches to sharing resources and expertise, increased focus on employers and their role in the career development process and more effective evaluation of quality and impact of services.

The Australian tertiary system

In Australia, university education is generally seen as a national government responsibility in funding terms, while VET and schools are the principal funding responsibility of the respective state and territory governments. Australia's tertiary education system comprises just over 40 public and private universities and 127 non-university higher education providers (Norton et al., 2018) and a VET system with a total of around 3500 Registered Training Organizations (RTOs) (NCVER, 2021). The great majority of

RTOs are small, private providers, but there are over 20 large state- and territory-based and government-owned Technical and Further Education (TAFE) institutions, as well as six dual sector universities offering both VET and HE studies. In addition, there are enterprise-based RTOs, pathway colleges which specialize in diploma courses aimed at facilitating entry to university courses (Norton et al., 2018) and adult and community education institutions offering a wide range of general interest and preparatory courses as well as nationally accredited vocational programs. Schools also offer VET programs as part of upper secondary education. Most RTOs offer a limited range of courses, often in quite specific occupations and skill needs. However, some RTOs, particularly TAFEs, have a large number of course offerings – including at HE level – that are more substantive and address a wide range of occupational and skills needs.

Australia's National Centre for Vocational Education Research (NCVER, 2021) reported that, in 2020, there were a total of 3.9 million students enrolled in VET programs, with just over half enrolled in nationally recognized programs, mainly through the sector's national Training Packages "that are occupational skills standards against which training delivery and assessment of competency can take place" or accredited courses "which enable nationally accredited training in niche areas or in response to rapidly emerging industry needs" (Productivity Commission, 2020, p. 5.5). Much of the VET sector's education and training is delivered on a fee-for-service basis and an estimated 21.7% of the Australian resident population aged 15–64 years participate in a nationally recognized VET program (NCVER, 2021). Most students (around 72%) are enrolled in nationally recognized training at a private training provider, with 20% at a TAFE institute, 9.8% at a community education provider, and about 2–3% in each of enterprise providers, schools and universities (some enrollments span across provider types). It is to be noted, however, that because these data are for total VET effort,² many may be undertaking VET courses that are short and relatively specific rather than the more substantive programs that typically characterize HE studies.

According to the Federal Government's statistics for 2020, the HE sector has just over 1.62 million students, of whom around 90% are in public institutions, about 70% are studying full-time, 30.6% are postgraduates and the remainder undergraduates. Around 30% were international students studying in Australia, although numbers were affected when COVID-19 hit and borders were effectively closed for an extended period. Two clear differences are that HE students are generally younger with most aged under 24 while most VET students are aged between 25 and 64, and most in HE study full-time while most in VET are part-time.

Pathways into and within Australia's tertiary education

Pathways into tertiary education

The Longitudinal Survey of Australian Youth (LSAY, 2019) recently established that higher education is the preferred post-school destination for Australian Gen Zs with the majority (59%) planning to go to university or take a gap year (10%). Another 13% plan to get a job, and only 11% see an apprenticeship, traineeship or other VET study as their preferred option post school. In terms of actual destinations, just over two-thirds (69%) of those who planned to go to university are there, just over half (57%) of those who planned to work are working, while half (51%) of those who sought VET studies are studying in the sector. Furthermore, in terms of changing plans, 20% of those who planned to work are studying some form of VET and 35% of those who planned VET studies are working.

Shedding further light on this issue, another study into how Gen Z make decisions about their future reported that:

43% of young people have "no idea" what they want to do when they leave school. When survey respondents were asked why they believed this was, only 9% said "figuring out what I want to do is not a priority for me right now". According to a much larger proportion of the cohort, a lack of exposure to different careers (77%), life experience (68%), knowledge of all options (67%) and career advice/support (55%) were to blame for their confusion. (Shipley and Stubley (2018, p. 35).

The Education Council (2020) concluded that students and their parents need to have access to earlier, more regular and intensive career advice informed by bigdata analysis of tertiary pathways derived from the use of Australia's Unique Student Identifier (discussed below).

Pathways within tertiary education

Improving pathways for students between VET and HE has been a focus of sector attention and government policy since the early 1980s. Bradley et al. (2008, p. 21) noted that VET is a common pathway to HE for many people from under-represented groups. However, while the review saw that improving pathways is important, it did not believe that VET studies should not be considered a feeder for HE. Indeed, the reviewers also believed that VET's primary purposes must not be distorted by the need to increase HE participation.

In reality, it has been very difficult to determine movements within tertiary education and the associated study pathways because of the lack of comprehensive data. However, Australia has now developed a lifelong unique student identifier (USI), first introduced in 2015. It presently covers nationally recognized vocational training offered in both schools and post-secondary RTOs. From the

²From 2014, all RTOs, including private providers, unless granted exemption, must collect and report full data on all nationally accredited training, in accordance with the National VET Provider Collection Data Requirements Policy.

beginning of 2021, the USI was extended to HE to replace the present identifier (the Commonwealth Higher Education Student Support Number). Determining movements between sectors and the nature of qualifications undertaken could be achieved only through database matching in the past. A project to do this was undertaken by the former Commonwealth Department of Education, Skills and Employment.³ It estimated an annual flow rate of 14–16% of students from HE to VET and 0.4–1.3% from VET to HE, which it believed to be an underestimate. However, the size of these flows runs against policy and funding approaches that promote and privilege “onward and upward” pathways. Rather, pathways appear to be more complex as has been reported above. In addition, it found the majority of students were younger when attending HE with most (70%) aged under 24 compared with when they studied VET, where 76% of students were aged 25–64 years.

The study found that, of those moving from HE to VET, most had completed a bachelor’s degree (63%) while others had dropped out (26%) or were still enrolled (12%). Their most likely VET enrollments were short and applied non-AQF qualifications mainly in the Health field (e.g., first aid or occupational health and safety courses). These are often required by employers or volunteer organizations. On the other hand, those undertaking more substantive courses most often studied in the Management and Commerce or Society and Culture fields (e.g., psychology; law and criminal justice; community, social and support services; social sciences). Students transitioning from VET to HE had most often completed their VET qualification and then undertaken a degree in some field. Watson et al. (2013) found, however, that access to HE institutions from VET studies was variable and dependent on the institution to which they applied, mainly due to inconsistencies in admission policies and practices. Full implementation of the USI will mean that sound data on pathways and transitions, whether by articulation, credit transfer or RPL, will become a reality, but this will require a long-running dataset to allow for the time over which individuals move into and out of a variety of tertiary and less formal studies, and work.

Issues in developing a more coherent Australian tertiary education system

The significant issues in Australian tertiary education emerging from the academic literature, government and other stakeholder voices include the call for, and moves toward, a more connected sector with better articulation and credit transfer arrangements and more equitable and fair funding across the components of the sector, particularly for the VET’s publicly-owned provider TAFE. Rhetoric promoted by both the Australian Industry Group (AiG) and the Business Council of Australia (BCA), two of Australia’s peak employer bodies, and others, is that equal value should be given to both VET and HE (AiG, 2019; BCA, 2017; Expert Panel, 2019) so that they are a more coherent, equal and integral parts of a joined up and accessible tertiary education system with strong pathways between them and from school. However, as this chapter shows, such rhetoric seems not to be reflected in the current reality of career and study choices and related tertiary pathways.

Critically, there are differences in philosophy and perceived roles of the HE and VET sectors, and in approaches to qualification design and ratification and development of curricula. VET’s approaches are dominated by an “industry-led” and competency-based philosophy, which can act against ready credit transfer, articulation arrangements and student transitions to HE. Moreover, universities enjoy a level of independence and freedom of action, including self-accreditation, not afforded institutions in the rest of tertiary education.

Recently, AiG (2019) characterized tertiary education as a highly unbalanced binary model with no coherent policy and funding framework. Indeed, Dawkins et al. (2019) assert that the dichotomy between HE and VET in Australia has, if anything, become starker because of disjointed policy since a review of higher education in 2008 strongly advocated an inter-connected tertiary system (Bradley et al., 2008). Dawkins et al. (2019) consider that tertiary curricula, and particularly those in vocational education, need reform. This requires a rethinking of VET’s competency-based model which, they contend, has squeezed broader-based skills and capabilities to the margins of curriculum and hinders tertiary pathways, especially those between VET and HE. Universities also emphasize the different approaches to learning and assessment in the two sectors which represent a philosophical and pedagogical divide that affects the extent to which tertiary education can operate as a coherent educational sector with strong connections and pathways.

Approaches to tertiary funding and effects on pathways

Funding of tertiary education, particularly the entitlement of students to support, is also inconsistent across the sectors and the current entitlement funding approach privileges both HE (Hurley and Van Dyke, 2019, 2020) and upward pathways to higher level qualifications to the detriment of study and pathway options. The demand-driven funding of HE, which ceased in 2018, also distorted the tertiary education market by making it easier to study in HE because a deferred debt rather than an upfront fee was accrued (Burke, 2018). As early as 2008, Bradley et al. had recommended the introduction of a uniform tertiary entitlement funding model across HE and VET to ensure that student choices were not distorted by different levels of public subsidy for similar education and training depending on the tertiary education sector which delivered it (Burke, 2018). This view was supported by the Australian Industry Group (AiG, 2019). The Business Council of Australia proposed the development of a Lifelong Skills Account capped at a set number of years of accredited learning and a continuing system of associated income-contingent loans, along with developing

³See <https://www.education.gov.au/better-understanding-student-journeys>. Some sections of this present chapter have drawn on a redacted version of a more comprehensive report on this project supplied by the Department.

a culture of lifelong learning (BCA, 2017). Likewise, Dawkins et al. (2019) proposed a more comprehensive system of income contingent loans across VET and HE that would remove the up-front fees faced by many VET students. To date, the relative levels of tertiary funding remains contentious.

Perceptions of VET and higher education and effects on pathways

Australia's HE system has not experienced the same large-scale quality and regulatory issues and problems as the VET sector, brought about in VET's case by a strong move to a market-driven approach which enabled students to enroll in courses of their choice coupled with the introduction of a poorly conceived and executed income contingent loan scheme for its higher-level qualifications at Diploma level or above: VET FEE-HELP (Saccaro and Wright, 2018). Nevertheless, and more recently, HE has not avoided critique either. Reviews by the Productivity Commission (2019, 2017) raised concerns about the HE sector's declining admissions standards, increasing student attrition rates and declining graduate outcomes under the demand-driven system. Declining admission standards, coupled with parity of esteem issues and lack of knowledge about VET study options and pathways vis-à-vis university studies by parents and schools, opened the doors to university studies to a wider range of students, some of whom may have found VET studies and pathways more appropriate (and to which individuals subsequently turn, as has already been noted above). In addition, scandals associated with vocational education and lack of knowledge of the possibilities of VET study options among school students, schools, careers advisors and parents all acted to privilege university studies and pathways over those in VET immediately post school (Education Council, 2020; South Australia Department for Education, 2019; Shipley and Stubley, 2018).

Dawkins et al. (2019) proposed that entry requirements and pathways into and through tertiary education be liberalized and diversified to recognize better the diversity of knowledge, skills and capabilities students bring to tertiary education, especially if they have undertaken prior studies after leaving school. They, like others, believe that articulation arrangements into university courses need to be more transparent and attract more credit for prior studies. This has been difficult in the past due to relatively inflexible practices by individual universities (Watson et al., 2013) and interpretations of the "Qualifications Pathways Policy" of the Australian Qualifications Framework (AQFC, 2013).

Approaches to revitalize Australian VET

A number of recent reviews of Australian VET have been conducted at both federal (Joyce, 2019) and state levels (Macklin, 2021; Gonski and Shergold, 2021). Macklin's review in the state of Victoria recommended developing a new model for financing the VET system, including lifelong learning, and strengthening partnerships between the TAFE and training sectors, HE and adult and community education providers. This model would unblock the flow of students through segments of tertiary education to enable it to respond with more agility to the economy's needs by brokering better collaboration and pathways between component providers. Gonski and Shergold's review in the state of New South Wales proposed the development of a new type of tertiary institution: an Institute of Applied Technology with strong industry input and partnerships which would deliver fully-integrated theoretical and practical employability skills, provided through a number of constituent colleges, with curricula designed in collaboration with industry and focused on the state's emerging labor market needs.

A federal initiative is the VET Reform Roadmap, highlighted in a consultation draft prepared by the Skills Senior Officials Network (2020), which proposes a greater focus on high quality education, training and assessment focusing on excellence; speeding up and streamlining current training package arrangements and training products; developing stronger alignment and integration between VET and HE; and developing a new, simpler and more flexible model for national qualifications based around occupational clusters and across the full range of the revised Australian Qualifications Framework, including micro-credentials.

The most recent national review (Joyce, 2019) led to the development of a vision for VET and a Skills Package: "Delivering skills for today and tomorrow" (Australian Government, 2020). The Skills Package includes the establishment of (a) the National Skills Commission to provide centralized VET information and skills forecasts and develop approaches to VET pricing that balance consistency with local flexibility, and (b) the National Careers Institute to enable Australians to access accurate and up-to-date career information and support regardless of age and stage of study and work. Nevertheless, there have been issues with that review recently reported from a range of sources, including its failure to recognize and accommodate the need for cooperative federalism in VET and the adequacy of VET funding (Fowler, 2022). Finally, it has proved impossible—to date—to negotiate a new National Agreement for Skills and Workforce Development (NASWD). The recent change of federal government in 2022 may assist an agreement between the federal and state/territory governments being reached, however. Fowler (2022) noted the need for federal and state/territory governments to work collaboratively to find new constructive ways of working, or risk competitive stagnation of Australia's skilled workforce and labor productivity.

Australian Qualifications Framework and effects on pathways

The Australian Qualifications Framework (AQF), introduced in 1995, was one of the first developed in the world and has been a critical part of the Australian education and training system. It defines the essential characteristics of qualifications issued across the senior secondary education, VET and HE systems. However, the recent and widespread move toward micro-credentials, flexible delivery options and mechanisms to assist learners in constructing their own programs, sometimes across sectors to meet individual learning needs, has challenged the original AQF.

A Qualification Pathways Policy was developed (AQFC, 2013) with the intent to support students' lifelong learning. It also aimed to enhance student progression into and between AQF qualifications, recognize the multiple pathways that students take to gain AQF qualifications and that learning can be formal, non-formal or informal, and support the development of pathways in qualifications design (AQFC, 2013). However, the policy focused primarily on HE qualifications and did not have adequate regard for multi-directional or intra-sectoral pathways (Expert Panel, 2019).

A review of the AQF has now been completed (Expert Panel, 2019). It proposed more flexible and multi-directional learning pathways which are viewed as more suited to lifelong learning and rapid retraining to meet new technological challenges rather than the simple, hierarchical ones that characterized the original AQF. This new approach encourages RPL and workplace learning within or outside formal studies. At present, the necessary mechanisms are lacking to assign levels to "non-standard" learning options, or to regulate or quality assure credit toward formal qualifications for in-service or informal learning, or micro-credentials, in any way that ensures national or international consistency.

The review recommended that the Pathways Policy needed revision not only "to broaden guidelines for credit recognition across AQF qualifications and to define and provide for recognition of shorter form credentials, including micro-credentials, toward AQF qualifications" (Expert Panel, 2019, p. 9), but also "to provide equal emphasis to pathways from vocational education and training to higher education and from higher education to vocational education and training, and within sectors, by including examples of the pathways that are possible and any requirements associated with them" (p. 71). The development of a "prototype national credit points system" was also proposed for voluntary adoption so that:

Post-secondary education and training is conceived and redesigned as a diverse set of offerings, available through better linkages and pathways between the VET and higher education sectors. These linkages and pathways will no longer be linear and hierarchical; they will need to recognize that throughout adulthood, people need to develop new skills in different areas and at different levels. Central to this objective is reinvigorating the VET system and raising its standing.

Expert Panel (2019, p. 8).

Enabling courses (also called foundation, access, widening participation, pathway, bridging or tertiary preparation programs) are a feature of offerings in both vocational and community education in Australia, but are not currently recognized in the current early version of the AQF. Thus, students enrolled in an enabling course may not be able to transfer their learning to another institution due to lack of clear pathways or credit policy (Expert Panel, 2019). However, and as already noted, data show that the link is loose between formal training undertaken and destination occupations for the majority of courses. It therefore appears that the view of VET courses as occupationally specific is appropriate only for a few cases (Wheelahan et al., 2015; Wheelahan and Moodie, 2011; Karmel et al., 2008). This is also likely for those in HE, except when, as in vocational education, there are registration requirements linked to the attainment of particular qualifications (for example and across tertiary education: lawyers, doctors, electricians and plumbers).

Despite the several years since this review was undertaken and its recommendations accepted by the Federal Government, implementation has proved problematic and remains in the "too hard basket" because of the significant implications of its adoption, including national VET funding and financing approaches, qualification/credential re-design, flow-on quality and regulatory changes, plus disturbances to a multitude of legislative, industrial awards and professional standards all anchored to the current AQF (Fowler, 2022).

Conclusion

This chapter has examined the notion of tertiary education pathways in the contexts of labor markets and skills formation systems. It has drawn on recent international literature for an overview and then zoomed in on the Australian situation as a case study because of its largely post-school and whole of life focus.

Paying close attention to these research studies and different contexts contributes to our knowledge in at least two main ways. First, they enrich our understanding of the notion of learning pathway by nuancing its complexities and issues. They highlight several potential disjunctures between policy emphases and by better documenting the actual experiences of learners, areas ripe for further research. For example, these could include provision of various types of career guidance vis-à-vis actual use of them made by learners; between the policy emphasis on job-related training and the actual drivers for many learners who state that they study for personal interest or development reasons; and between policy expectation that pathways allow credit transfer and RPL and the actual amount of these applied for and granted. Analyzing the views of various stakeholders helps in developing our understanding of how phenomena are experienced by the actors themselves, thereby adding a valuable dimension to official interpretations from governments and media.

Second, they provide valuable insights into tertiary education practice. Deeper understanding of different types of pathways and how individuals use them can assist tertiary administrators in customizing initiatives to respond appropriately to changes in learners' lives, learning and work. It can raise significant issues concerning relationships between academic education and vocational education and can provide policymakers and institutional planners with insights into how best to fashion these relationships between sectors. For example, studies on learners' frequent moves between sectors and between sectors and employment unveil

how the sectors are playing complementary roles in providing for the needs of worker-learners at different stages through their careers. Moreover, an understanding of pathways can assist learners themselves to navigate through sectors with very different missions, cultures and practices. In Australia the tertiary sector is in a constant state of evolution, with a strong need to overcome parity of esteem and status issues for VET. It is a long way from understanding the dimensions and nature of pathways between study options and work.

Studies on transitions confirm that tertiary education is an important means for facilitating pathways into and through the labor market. Yet continuing vocational training has been labeled “the weakest link in lifelong learning” (Tessaring and Wannan, 2004, p. 6) and has been highlighted as one of five major priorities for promotion, along with tackling the large number of low-skilled people, increasing mobility to promote the development of the labor market, investing in quality VET systems, and ensuring high quality VET professionals. These areas represent fertile ground for continuing cultivation.

References

- Agbenyo, H., Collett, K., 2014. Career advice and guidance in a world where vocational skills matter. In: Arulmani, G., Bakshi, A., Leong, F., Watts, A. (Eds.), *Handbook of Career Development*. International and Cultural Psychology Series. Springer, New York.
- AIQ [Australian Industry Group], 2019. Realising Potential: Solving Australia's Tertiary Education Challenge. AIQ, Sydney.
- Alan, A., de Diego, M.E., 2019. Unpacking School-To-Work Transition: Data and Evidence Synthesis. UNICEF, New York.
- AQFC [Australian Qualifications Framework Council], 2013. Australian Qualifications Framework, second ed. Australian Qualifications Framework Council, Adelaide.
- Arthur, M.B., 2014. The boundaryless career at 20: where do we stand, and where can we go? *Career Dev. Int.* 19 (6), 627–640.
- Australian Government, 2019. Changes in the Australian Labour Market: A 30-year Perspective. Department of Employment, Skills, Small and Family Business, Canberra.
- Australian Government, 2020. Delivering Skills for Today and Tomorrow. Department of Education, Skills and Employment, Canberra.
- Australian Government, 2021. Vocational Education and Training (VET) Reform Roadmap Draft. Department of Education, Skills and Employment, Canberra. <https://www.dese.gov.au/expert-review-australias-vet-system/resources/vocational-education-and-training-vet-reform-roadmap-draft>.
- Bathmaker, A.-M., 2017. Post-secondary education and training, new vocational and hybrid pathways and questions of equity, inequality and social mobility. *J. Vocat. Educ. Train.* 69 (1), 1–9.
- Bathmaker, A.-M., Thomas, W., 2009. Positioning themselves: an exploration of the nature and meaning of transitions in the context of dual sector FE/HE institutions in England. *J. Furth. High. Educ.* 33 (2), 119–130.
- BCA [Business Council of Australia], 2017. Future-Proof, Protecting Australians Through Education and Skills. Business Council of Australia, Melbourne.
- Billett, S., Dymock, D., Hodge, S., Choy, S., Le, A.H., 2020. Shaping Young People's Decision-Making About Post-School Pathways: Institutional and Personal Factors, Book Chapter in the Professional and Practice-Based Learning Series. Springer. Reported in Newsletter for European Research in Learning and Work, August. www.news.wifo-gate.org.
- Bol, T., van de Werfhorst, H.G., 2013. Educational systems and the trade-off between labour market allocation and equality of educational opportunity. *Comp. Educ. Rev.* 57 (2), 285–308.
- Bradley, D., Noonan, P., Nugent, H., Scales, B., 2008. Review of Australian Higher Education. Department of Education, Employment and Workplace Relations, Canberra.
- Bretherton, T., 2011. Understanding the Undertow: Innovative Responses to Labour Market Disadvantage and VET. National Centre for Vocational Education Research, Adelaide.
- Brown, J., 2017. In Their Words: Student Choice in Training Markets: Victorian Examples. NCVER, Adelaide.
- Brzinsky-Fay, C., Solga, H., 2016. Compressed, postponed, or disadvantaged? School-to-work transition patterns and early occupational attainment in West Germany. *Res. Soc. Stratif. Mobil.* 46, 21–36.
- Buchanan, J., Allais, S., Anderson, M., Calvo, R.A., Peter, S., Pietsch, T., 2020. The Futures of Work: What Education Can and Can't Do. Background Paper for the Futures of Education Initiative. <https://unesdoc.unesco.org/ark:/48223/pf0000374435>.
- Burke, G., 2018. Changes in Funding in Australian Vocational Education and Their Effects. L.H. Martin Institute, University of Melbourne, Melbourne.
- Bussemeyer, M.R., Trampusch, C. (Eds.), 2012. The Political Economy of Collective Skill Formation. Oxford University Press, Oxford.
- Cameron, C.L., 2018. Changing fortunes? Aspiration and realisation for looked after young people's post-compulsory educational pathways in England. *Longitud. Life Course Stud.* 9 (1), 101–114.
- CareersNet, 2020. Inventory of Lifelong Guidance Systems and Practices. Cedefop. <https://www.cedefop.europa.eu/en/publications-and-resources/country-reports/inventory-lifelong-guidance-systems-and-practices>.
- Caves, K.M., Baumann, S., Renold, U., 2021. Getting there from here: a literature review on vocational education and training reform implementation. *J. Vocat. Educ. Train.* 73 (1), 95–126 published online: 4 December.
- Cedefop [European Centre for the Development of Vocational Training], 2017. Education and Labour Market Outcomes for Vocational Education and Training Graduates in Different Types of VET Systems in Europe. Cedefop, Thessaloniki, Greece.
- Céreq International Mission, 2020. About the School to Work Transition. In and Around, 1, March. <https://www.cereq.fr/en/about-school-work-transition>.
- CICA [Career Industry Council of Australia], 2007. The Public Benefits of Career Development Services. <https://cica.org.au/wp-content/uploads/The-Public-Benefits-of-Career-Development-Services-A-Position-Paper.pdf>.
- Colleges and Institutes Canada, 2021. The Status of Microcredentials in Canadian Colleges and Institutes: Environmental Scan Report. Colleges and Institutes Canada, Ottawa.
- Crump, S., 2007. Changing identities and performance of post-compulsory educational providers. *Aust. Educ. Res.* 34 (3), 1–14.
- Dawkins, P., Hurley, P., Noonan, P., 2019. Rethinking and Revitalising Tertiary Education in Australia. Mitchell Institute, Victoria University, Melbourne.
- de los Santos, A.G., Sutton, F., 2012. Swirling students: articulation between a major community college district and a state-supported research university. *Commun. Coll. J.* 36 (12), 967–981.
- Education Council, 2020. Looking to the Future: Report of the Review of Senior Secondary Pathways Into Work, Further Education and Training. Education Council, Carlton South, Victoria.
- Evans, K., Furlong, A., 1997. Metaphors of youth transitions: niches, pathways, trajectories or navigations. In: Bynner, J., Evans, K., Furlong, A. (Eds.), *Youth, Citizenship and Social Change in a European Context*. Ashgate, Aldershot.
- Expert Panel [Expert Panel for the Review of the Australian Qualifications Framework & Australia Department of Education], 2019. Review of the Australian Qualifications Framework: Final Report. Department of Education, Canberra.
- Fingold, D., Soskice, D., 1988. The failure of training in Britain: analysis and prescription. *Oxf. Rev. Econ. Pol.* 4 (3), 21–53.
- Forster, A.G., Bol, T., van de Werfhorst, H.G., 2016. Vocational education and employment over the life cycle. *Sociol. Sci.* 3, 473–494.
- Fortwengel, J., Gospel, H., Toner, P., 2019. Varieties of institutional renewal: the case of apprenticeship in the US, England and Australia. *J. Vocat. Educ. Train.* <https://doi.org/10.1080/13636820.2019.1688855>.

- Fowler, C., 2022. Will a new government undo knotted problems in the stalled reform of the VET National Agreement and the Australian Qualifications Framework? *Campus Rev.* <https://www.campusreview.com.au/2022/06/will-a-new-government-undo-knotted-problems-in-the-stalled-reform-of-the-vet-national-agreement-and-the-australian-qualifications-framework-opinion/rovide>.
- Fredman, N., 2013. Student movement: pathways, fields and links to work. *Int. J. Train. Res.* 11 (1), 5–16. <https://doi.org/10.5172/ijtr.2013.11.1.5>.
- Fry, J., Boulton, C., 2013. Prevalence of Transition Pathways in Australia. Productivity Commission Staff Working Paper, Canberra.
- Gonski, D., Shergold, P., 2021. In the Same Sentence: Bringing Higher and Vocational Education Together. Department of Premier and Cabinet, Sydney (viewed 6 June 2022). <https://education.nsw.gov.au/about-us/strategies-and-reports/our-reports-and-reviews/review-on-the-nsw-vocational-education-and-training-sector>.
- GOVET (German Office for International Cooperation in Vocational Education and Training) (Ed.), 2022. Effects of the COVID-19 Pandemic on Vocational Education and Training: International Perspectives of Policy Makers and Practitioners. Federal Institute for Vocational Education and Training, Bonn, 2022.
- Greinert, W.-D., 1993. Das 'deutsche system' der berufsausbildung: geschichte, organisation, perspektiven. studien zur vergleichenden. Berufspädagogik. Nomos, Baden-Baden.
- Gubler, M., Arnold, J., Coombs, C., 2014. Reassessing the protean career concept: empirical findings, conceptual components, and measurement. *J. Organ. Behav.* 35 (S1), S23–S40.
- Guthrie, H., Nechvoglod, L., 2011. Identification of Issues that Impact Upon the Provision of Effective Career Development Services for VET Learners. NCVER, Adelaide.
- Hall, P.A., Soskice, D. (Eds.), 2001. Varieties of Capitalism: The Institutional Foundations of Comparative Advantage. Oxford University Press, Oxford.
- Hanushek, E.A., Schwerdt, G., Woessmann, L., Zhang, L., 2017. General education, vocational education, and labour-market outcomes over the life cycle. *J. Hum. Resour.* 52 (1), 49–88.
- Harris, R., Rainey, L., 2009. Interrogating the notion of transition: learner experiences of multiple transitions between the vocational and higher education sectors. In: Brooks, R. (Ed.), *Transitions From Education to Work: New Perspectives From Europe and Beyond*. Palgrave Macmillan, Basingstoke, Hampshire, pp. 42–60.
- Harris, R., Rainey, L., 2012. Learning pathways between and within vocational and higher education – towards a typology? *Aust. Educ. Res.* 39 (1), 107–123.
- Harris, R., Ramos, C., 2020. Youth Transitions: Creating Pathways to Success. Education Development Trust, Reading, UK.
- Harris, R., Ramos, C., 2013. Building career capital through further study in Australia and Singapore. *Int. J. Lifelong Educ.* 32 (5), 620–638.
- Heinz, W.R., 2009. Structure and agency in transition research. *J. Educ. Work* 22 (5), 391–404.
- Hodkinson, P., Bloomer, M., 2002. Learning careers: conceptualising lifelong work-based learning. In: Evans, K., Hodkinson, P., Unwin, L. (Eds.), *Working to Learn*. Kogan, London.
- Hughes, D., Smith, G., 2020. Youth Transitions: Creating Pathways to Success. Education Development Trust, Reading, UK.
- Hurley, P., Van Dyke, N., 2019. Australian Investment of Education: Vocational Education and Training. Victoria University, Melbourne.
- Hurley, P., Van Dyke, N., 2020. Australian Investment of Education: Higher Education. Victoria University, Melbourne.
- Joyce, S., 2019. Strengthening Skills: Expert Review of Australia's Vocational Education and Training System [Joyce review]. Department of the Prime Minister and Cabinet, Canberra.
- Karmel, T., Mlotkowski, P., Awodeyi, T., 2008. Is VET Vocational? the Relevance of Training to the Occupations of Vocational Education and Training Graduates. NCVER, Adelaide.
- Karp, P., 2020. Australian university fees to double for some arts courses, but fall for stem subjects. *The Guardian*, 19 June. <https://www.theguardian.com/australia-news/2020/jun/19/australian-university-fees-arts-stem-science-maths-nursing-teaching-humanities#:~:text=Tehan%20says%20the%20student%20contribution,increase%20from%20%2434%2C000%20to%20%2443%2C500>.
- Keskiner, E., 2019. Youth Transitions Among Descendants of Turkish Immigrants in Amsterdam and Strasbourg: A Generation in Transition. Springer, Cham, Switzerland.
- Lantero, L., Finocchietti, C., Petrucci, E., 2021. Micro-credentials and Bologna Key Commitments: State of Play in the European Higher Education Area. https://microcredentials.eu/wp-content/uploads/sites/20/2021/02/Microbol_State-of-play-of-MCs-in-the-EHEA.pdf.
- LSAY [Longitudinal Survey of Australian Youth], 2019. Generation Z: Leaving School. Infographic. December. <https://www.lsay.edu.au/publications/search-for-lsay-publications/generation-z-leaving-school>.
- Lynch, L.M., 1994. Training and the Private Sector: international comparisons. University of Chicago Press, Chicago.
- Macklin, J., 2021. Future Skills for Victoria: Driving Collaboration and Innovation in Post-secondary Education and Training. Victorian Government, Melbourne.
- Maxwell, W., Hagedorn, L.S., Brocato, P., Hye, S.M., Perrakis, A., 2002. The Community College Shuffle: Student Patterns of Multiple Enrolments. Paper at Annual Meeting of the American Educational Research Association, New Orleans, April 1–5.
- Middeldorp, M.M., Edzes, A.J., van Dijk, J., 2019. Smoothness of the school-to-work transition: general versus vocational upper-secondary education. *Eur. Socio Rev.* 35 (1), 81–97.
- NCVER [National Centre for Vocational Education Research], 2021. Total VET Students and Courses 2020. NCVER, Adelaide.
- Norton, A., Cherastidham, I., Mackey, W., 2018. Mapping Australian Higher Education 2018. Grattan Institute, Melbourne.
- NZQA [New Zealand Qualifications Authority], 2022. Micro-credentials. NZQA, Wellington, NZ. <https://www.nzqa.govt.nz/providers-partners/approval-accreditation-and-registration/micro-credentials/>.
- OECD [Organisation for Economic Co-operation and Development], 2000. From Initial Education to Working Life: Making Transitions Work. OECD Publishing, Paris.
- OECD, 2014. Skills Beyond School. OECD Publishing, Paris.
- OECD, 2020. Education at a Glance 2020: OECD Indicators. OECD Publishing, Paris.
- OECD, 2021. Implications of the COVID-19 Pandemic for Vocational Education and Training. OECD Publishing, Paris. <https://doi.org/10.1787/55afea00-en>.
- Parliament of Victoria, 2018. Inquiry Into Career Advice Activities in Victorian Schools. Economic, Education, Jobs and Skills Committee, Melbourne.
- Pilz, M., 2016. Typologies in comparative vocational education: existing models and a new approach. *Vocat. Learn.* 9 (3), 295–314.
- Pilz, M., 2017. Policy borrowing in vocational education and training (VET) – VET system typologies and the '6P strategy' for transfer analysis. In: Pilz, M. (Ed.), *Vocational Education and Training in Times of Economic Crisis: Lessons From Around the World*. Springer, Cham, Switzerland, pp. 473–490.
- Pouliakas, K., Ranieri, A., 2018. Governance of labor market and skills intelligence as driver of VET reform. In: McGrath, S., Mulder, M., Papier, J., Suart, R. (Eds.), *Handbook of Vocational Education and Training*. Springer, Cham, Switzerland, pp. 491–511.
- Productivity Commission, 2017. Shifting the Dial: 5 Year Productivity Review. Productivity Commission, Canberra.
- Productivity Commission, 2019. The Demand Driven University System: A Mixed Report Card. Productivity Commission, Canberra.
- Productivity Commission, 2020. Report on Government Services 2020: Vocational Education and Training. Productivity Commission, Canberra.
- Puckett, J., Boutenko, V., Hoteit, L., Polunin, K., Perapeshka, S., Stepanenko, A., Loskareva, E., Bikkulova, G., 2020. Fixing the Global Skills Mismatch. <https://www.bcg.com/en-gb/publications/2020/fixing-global-skillsmismatch.aspx>. (Accessed August 2020).
- Quyoum, A., Hordósy, R., Mittelmeier, J., Jones, S., McCaldin, T., 2021. Career decisions of further education college students: where does higher education 'fit in'? *Widening Participat. Lifelong Learn.* 23 (1), 54–78.
- Raffe, D., 2003. Pathways linking education and work: a review of concepts, research and policy debates. *J. Youth Stud.* 6 (1), 3–19.
- Ranasinghe, R., Chew, E., Knight, G., Siekmann, G., 2019. School-to-Work Pathways. NCVER, Adelaide.
- Roberts, K., 2020. Regime change: education to work transitions in England, 1980s–2020s. *J. Appl. Youth Stud.* 3, 23–42.
- Roosmaa, E.-L., Martma, L., Saar, E., 2019. Vocational upper-secondary education and participation in non-formal education: a comparison of European countries. *Int. J. Lifelong Educ.* 38 (3), 268–286.
- Roth, T., 2017. Interpersonal influences on educational expectations: new evidence for Germany. *Res. Soc. Stratif. Mobil.* 48, 68–84.
- Ryan, P., 2001. The school-to-work transition: a cross-national perspective. *J. Econ. Lit.* 39 (1), 34–92.
- Saccaro, F., Wright, R., 2018. Vet Fee-Help: What Went Wrong? L.H. Martin Institute, University of Melbourne, Melbourne.

- Schoon, I., Lyons-Amos, M., 2017. A socio-ecological model of agency: the role of structure and agency in shaping education and employment transitions in England. *Longit. Life Course Stud.* 8 (1), 35–56.
- Shipley, B., Stubley, W., 2018. After the ATAR II: Understanding How Gen Z Make Decisions About Their Future. Year13, Australia. <https://year13.com.au/>.
- Skills Senior Officials Network, 2020. Vocational Education and Training Reform Roadmap. <https://docs.employment.gov.au/documents/vocational-education-and-training-vet-reform-roadmap-consultation-draft>.
- Solomon, N., 2005. Identity work and pedagogy: textually producing the learner-worker. *J. Vocat. Educ. Train.* 57 (1), 95–108.
- South Australia Department for Education, 2019. Vocational Pathways to Employment: A Review of VET for School Students. Department for Education, Adelaide. <https://www.education.sa.gov.au/sites/default/files/vet-issues-paper.pdf>.
- Tessaring, M., Wannan, J., 2004. Vocational Education and Training—Key to the Future. European Centre for the Development of Vocational Training, Thessaloniki [Cedefop].
- Tomlinson, M., 2012. Graduate employability: a review of conceptual and empirical themes. *High Educ. Pol.* 25, 407–431.
- Toner, P., 2014. Contracting out publicly funded vocational education: a transaction cost critique. *Econ. Lab. Relat. Rev.* 25 (2), 222–239.
- Toner, P., 2018. How Economics Explains Failure of the Publicly Funded Privately Delivered Training Market. L.H. Martin Institute, University of Melbourne, Melbourne.
- Vaughan, K., 2005. The pathways framework meets consumer culture: young people, careers and commitment. *J. Youth Stud.* 8 (2), 173–186.
- Vossiek, J., 2018. Collective Skill Formation in Liberal Market Economies? The Politics of Training Reforms in Australia, Ireland and the United Kingdom. Peter Lang, Bern.
- Watson, L., Hagel, P., Chesters, J., 2013. A Half-Open Door: Pathways for VET Award Holders Into Australian Universities. NCVER, Adelaide.
- Wheelahan, L., Moodie, G., 2011. Rethinking Skills in Vocational Education and Training: From Competencies to Capabilities. NSW Board of Vocational Education and Training, Sydney.
- Wheelahan, L., Moodie, G., 2017. Vocational education qualifications' roles in pathways to work in liberal market economies. *J. Vocat. Educ. Train.* 69 (1), 10–27.
- Wheelahan, L., Buchanan, J., Yu, S., 2015. Linking Qualifications and the Labour Market Through Capabilities and Vocational Streams. NCVER, Adelaide.
- Wolbers, M.H.J., 2007. Patterns of labour market entry: a comparative perspective on school-to-work transitions in 11 European Countries. *Acta Sociol.* 50 (3), 189–210.
- Young, M., 2006. Further and higher education: a seamless or differentiated future? *J. Furth. High. Educ.* 30 (1), 1–10.

Learning through work: augmenting vocational educational experiences to meet 21st-century work requirements

Stephen Billett, Griffith University, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Augmenting vocational educational experiences to meet 21st-century work requirements	452
Adaptability and vocational education	453
Key changes reconfiguring the goals for and processes of vocational education	454
Job readiness as well as occupational preparation	454
Educational goals: canonical, situational and adaptability	454
Securing “hard to learn” knowledge	455
Responsive vocational education curriculum and pedagogic practices	455
Organizing and providing workplace experiences	456
Intentionally and actively integrating students’ experiences	456
Institutional-based activities that incite authentic work experiences	457
Educational processes promoting adaptability	458
Securing “hard-to-learn” knowledge	459
Promoting learning agency	459
Developing a skillful and adaptable workforce	460
References	460

Augmenting vocational educational experiences to meet 21st-century work requirements

Constant changes in occupational and workplace performance requirements and increasingly dynamic working lives prompt fresh considerations about the goals for and processes of vocational education and training (VET), whether in schools or in specialized VET institutions. Central here are that these goals and processes are focused on generating adaptable learning outcomes: generating robust occupational knowledge in students. That is, VET must have the capacity to be applied beyond the circumstances of its initial learning and respond to changing occupational and workplace requirements. As part of meeting these needs in initial occupational preparation, VET students require workplace experiences to develop the capacities they require to (a) transition to effective work performance beyond graduation, and (b) respond productively to changing occupational and workplace performance requirements. Yet, there are important contributions from experiences provided through educational programs and processes and those from the workplace (Billett, 2009). Hence, the key consideration is for providing and augmenting these two sets of experiences so that the potential of the contributions of both can be realized. These premises shape what constitutes the purposes of vocational education in the contemporary era (i.e., adaptable occupational and/or workplace preparation), the kinds of curriculum and pedagogic practices (e.g., education experiences promoting adaptability) enacted to achieve those outcomes, and how students need to be positioned in initial and continuing vocational education (e.g., being adaptable workers).

This article proposes that to meet the requirements of the 21st-century workforce, VET provisions need to be directed toward educational processes that promote adaptability within the domain of occupational competence, premised on triadic foundations comprising (a) curriculum and (b) pedagogic practices directed toward generating student outcomes premised on foundations of occupational knowledge, and (c) engagement with and adaptive use of that knowledge by learners (i.e., students, workers). This requires accounting for the kinds of experiences provided to students in VET programs, and how those students should engage with and reconcile experiences provided by educational and workplace settings, as they each make specific and important contributions to both the development of occupational knowledge and the capacity to apply it adaptively.

In describing, justifying and discussing these three premises, this article seeks to explain how adaptive outcomes might be achieved through work-based experiences aligned with and augmented in vocational education programs. The conceptual premises are about the intersection among (a) experiences of different kinds being provided in educational and workplace settings through vocational educational programs, (b) the bases by which students come to engage in those experiences, and (c) the relations between the contributions of those two settings. The article outlines the changes requiring the reconfiguration of the goals for vocational education. These include (a) a growing focus on job readiness as well as occupational preparation, (b) an increasing requirement for conceptual and symbolic knowledge in contemporary work, and (c) the importance of being adaptive and innovative. Central to these changes is the key role that vocational education students and graduates play in being active and adaptive learners both in their initial occupational preparation and then in their ongoing development across lengthening working lives. Ultimately, the development of capacities required to be adaptable is something needing to be learned and enacted by individuals.

Having established these premises, models of vocational education curriculum that include workplace experiences are advanced. These emphasize the provision and integration of workplace or practice-based experiences into the overall curriculum. Broadly,

these responses are those associated with provisions of activities and interactions in vocational education institutions and workplaces that are structured to assist the development of adaptive occupational capacities to support graduates' transitions to initial employment and changing occupational requirements across working life through informed principles and practices. This all suggests the need to both provide and augment workplace experiences to support this learning. Beyond considerations about the sequencing, duration and means of integrating workplace or practice-based experiences are those about the selection and use of pedagogic practices to generate adaptive outcomes through conceptual, procedural and dispositional knowledge of the occupation, and something of its diverse manifestations in actual work settings. These pedagogic practices extend beyond those enacted by teachers and in classroom settings to include or replicate experiences in workplaces and/or educational settings likely to generate rich and adaptive occupational learning. Considerations of augmenting experiences in education and workplace settings are those associated with developing students' capacities to be active, effective and intentional learners at the commencement of, and across, their working lives.

Adaptability and vocational education

Being adaptable has always been important in responding to occupational and workplace challenges and has long been understood as a key quality of expert occupational performance (Ericsson and Lehmann, 1996). In the current era, responding to the constant changes in the requirements for occupational and workplace performance means that adaptability is becoming a core capacity for contemporary workers. Therefore, developing adaptable occupational capacities is a central goal for vocational education. Although adaptability is understood as comprising the ability to adapt existing knowledge to new challenges and novel circumstances (Stevenson, 1991), it is helpful to understand what constitutes adaptability and how contemporary views of it have emerged. Earlier, it was proposed that general problem-solving strategies could be developed that were applicable to any circumstances, suggesting some forms of knowledge are inherently adaptable (Bartlett, 1958). These capacities are not restricted to specific activities: they included literacy and numeracy that, it was claimed, can be applied across domains of activities, both within and outside of working life. Such claims led to a focus on general problem-solving strategies and interest by governments in key or core competencies. It was claimed that, once learned, these capacities could assist the person in responding to new challenges, regardless of the domain of knowledge in which the challenges were being confronted (Mayer, 1992; The Secretary's Commission on Achieving Necessary Skills, 1992). It has been suggested that engaging in gameplaying or science (Bantock, 1980), can lead to logical thinking and enhanced problem-solving that could have a wider application. When advising UNESCO about educational processes for the future, for instance, Faure et al. (1972) favored general problem-solving over specific occupational preparation. They argued that, because the requirements for occupations and work are constantly changing, students should learn problem-solving strategies rather than being prepared for specific occupations. However, both these claims and more recent manifestations promoting generic skills or competencies have found to be wanting. They ignored existing (Anderson, 1982) and an emerging body of findings claiming that responding to novel tasks or nonroutine problems is seen to reside within specific domains of activities (Chi et al., 1982; Sweller, 1990a,b): Cleverness (i.e., the ability to manipulate knowledge) is insufficient for nonroutine, domain-specific problem-solving (Chase and Simon, 1973), and situational and procedural factors challenge the efficacy of broadly transferable knowledge, even in maths (Bishop, 1991; Carraher et al., 1985). That is, without having knowledge of the domain in which the problems arise and are to be solved, general problem-solving strategies have very limited potency. Certainly, the importance of domain specificity (e.g., occupations) is repeatedly emphasized in studies of expertise (Glaser, 1984; Ericsson and Lehmann, 1996). Indeed, the hallmark of expertise is an ability to perform nonroutine problem-solving within that domain. Yet, more than being competent within an occupational domain, the requirements for, practice of and judgments about what constitutes effective performance are manifested situationally (Billett, 2001). That is, performance requirements for what constitutes expertise are situated in specific circumstances of the occupational practice. This means that domain-specific competence at both the occupational (i.e., canonical) and situational levels are necessary, as is adaptability across them. In short, as with earlier versions (i.e., key competences, SCANS), capacities posited as 21st-century skills and processes of adaptability need to be embedded in domains of occupational knowledge and its practice.

So, when referring to the development of occupational skills through vocational education, that development must be directed toward two kinds of domains: canonical occupational knowledge (i.e., requirements of that occupation) and the domain of activities within a specific workplace or practice (i.e., performance requirements within the work setting); that is, being adaptable within a domain of occupational knowledge manifested in the circumstances of practice (i.e., workplace/work practice; Billett et al., 2018). Rather than vocational education efforts being directed toward generic problem-solving strategies, the focus needs to be on developing the capacities that permit graduates to engage in effective problem-solving within specific domains of occupational knowledge, taking into account how that knowledge and its use are manifested in particular work situations (Billett, 2001). What this suggests is that there is no such entity as an occupational expert per se. Instead, expertise needs to be judged in terms of practitioners effectiveness in engaging in nonroutine problem-solving in specific work settings. This suggests that a foundation of occupational knowledge is required prior to efforts to promote adaptability (Alexander and Judy, 1988; Gelman and Greeno, 1989). The salience of this adaptability is also emphasized in the Program of International Assessment of Adult Competence (PIAAC) data that indicate all classes of contemporary workers engage in extensive and frequent nonroutine problem-solving and, therefore, adaptability is central to responding to nonroutine problem-solving (Organisation for Economic Co-operation and Development, 2013). All of this is consistent with the concepts of canonical and situational occupational knowledge (Billett, 2017). Responding to occupational challenges in workplaces, addressing clients' problems and needs, requires the capacities to adapt occupational knowledge

to meet those requirements (Ericsson and Lehmann, 1996). Consequently, more than vocational education being focused on developing occupational knowledge per se (i.e., national occupational standards), central here is generating student capacities that include that knowledge being applied in different ways and having the adaptive capacities to understand the circumstances of and requirements for adapting to emerging challenges and new situations.

An important element of those capacities is the active and agentic engagement by VET students. It is they who need to construct the knowledge (i.e., learn), apply it adaptively and interact with others in doing so. Importantly, developing those capacities and responding to these changes is often associated with interdependence; that is, through individuals working and learning with others and in ways that promote the accessing of an engagement with knowledge that is socially generated. Occupational knowledge arises through history, culture and situation (Sennett, 2008), which means individuals need to engage with the world around them to access and learn that knowledge through interacting with it: inter-psychologically (Wertsch, 1991). Through interaction with others, particularly those who are more experienced or informed, shared learning arises in ways that are reliant upon the contributions of others and objects (Newman et al., 1989). Yet, in terms of change, the interdependence between the workplace and the worker comes to the fore. The workplace requires workers to have the capacity to respond to new challenges and generate new procedures and practices to achieve workplace goals (Billett, 2014). Without the actions of the workers the workplace would become moribund and unable to advance. Yet, workers also need the workplace, the contributions it provides and its role in continuing to develop that knowledge that is so central to being able to respond to emerging and novel challenges.

The key points here are that vocational education's purposes are about developing occupational capacities, but in ways that are adaptable to specific workplace settings and novel tasks that arise in them, as well as the routine activities that constitute existing practice. These are all key elements that are required to respond to the kinds of workplace challenges that exist and are emerging for contemporary workers, seemingly in ever more demanding ways.

Key changes reconfiguring the goals for and processes of vocational education

The constant changes to occupational practice and workplace requirements outlined above have direct implications for the provision of vocational education. These include: (a) a focus on job readiness as well as occupational preparation; (b) addressing the specific needs of workplaces; and (c) accessing and learning the kinds of conceptual and symbolic knowledge required for processes. Each of these changes is discussed briefly to indicate implications for vocational education.

Job readiness as well as occupational preparation

In the modern era, the role of vocational education has been to prepare people for working life and specific occupations (Department of Innovation and Universities and Skills, 2008). This role emphasizes identifying occupational requirements and then preparing students to meet them. Statements of these requirements have been generated through professional associations or licensing bodies for many occupations. Consequently, external bodies such as peak industry bodies, including those of employers and unions, have played a significant role in identifying those requirements and advising on what content should be taught, the kinds of assessments students will be subjected to and how those assessments are conducted. Hence, there has been a great focus on national curriculum and occupational standards, and these are set as the goals toward which vocational education's processes and outcomes should be directed. However, increasingly, employers, governments, community and students themselves are now expecting that graduates from vocational education programs should be "job ready" (Billett, 2015a); that is, able to effectively perform in specific workplaces. This is a particularly tough goal for vocational education, and hugely demanding on students, as they do not know where graduates will find employment, and therefore can be prepared and acquitted as being job ready. There is also considerable difference between curriculum intents and experiences designed to meet the needs of occupational requirements and regulations, those associated with meeting the requirements of specific workplaces, and the need to be adaptable in both. So, this growing emphasis and expectation that students from vocational education will be able to move smoothly into employment in a specific workplace requires a different set of educational goals and processes than those associated with occupational readiness as the traditional goal for vocational education. This requirement extends to knowing something about the variations of occupational practice and the reasons for those variations and their consequences for work performance. So, this change warrants a significant reconsideration of educational goals and processes.

Educational goals: canonical, situational and adaptability

Education is an intentional process. It is guided and driven by specific kinds of intentions (e.g., goals, aims, objectives) generated by contributions and insights from a range of informants, balancing among these contributions and selecting and generating kinds of educational intents (Marsh, 2004). Very rarely is the development of educational programs and processes the product of individual teacher decision-making alone (Skilbeck, 1984). Instead, it reflects the involvement of a range of interests, contributors and perspectives. Also, the degree of specification for educational processes can vary from being wide and open to being highly specific. The latter often occurs within preparation for occupations because, increasingly, there are specific occupational requirements that need to be met. What electricians, pilots, nurses, builders and so forth are required to perform is regulated, and vocational education provisions need to address those regulations. Following from what has been proposed above, it is helpful to consider occupational

preparation in terms of it comprising canonical occupational knowledge (Billett, 2017), that is, the knowledge required by anybody practicing that occupation: what people need to know, do and value. Knowing refers to factual, conceptual, propositional and causal knowledge (Prawat, 1989). Being able to secure goals is achieved through procedural knowledge that ranges from highly specific procedures (to achieve individual tasks) through to strategic procedures required for planning and evaluating work activities (Anderson, 1982; Ericsson and Lehmann, 1996; Sun et al., 2001). There is also a consideration of values—the dispositional qualities of interest and intentionality that are central to how individuals undertake their work (Perkins et al., 1993). It is, together, these three forms of knowledge that underpin effective work performance, and what some refer to as expertize (Ericsson et al., 2018).

All of these forms of knowledge become the goals for vocational education at both the occupational and situational levels. Often, those goals are captured in national standards and statements of occupational competence and presented in syllabus documents, usually focused on the canonical occupational knowledge to be learned. However, there is also a need for goals associated with understanding situational requirements that permit job performance in a specific workplace/work practice (Billett, 1998). Added to this are those associated with developing the adaptability that permits what has been learned to be employed in other circumstances and tasks than those in which and through which they were learned. So, the educational intents need to include how students might come to learn something of the diverse kinds of workplace performance associated with their occupations. Central here is development of adaptability in students: extending what they have learned to be applied in circumstances other than those in which it was learned.

Consequently, the challenge for vocational education is for its students to understand something of the diversity of situational requirements and to expose students to instances of that diversity and allow them to appreciate the range of requirements arising from it. Engagements and partnerships with local employers and industry representatives and those from professional bodies are likely to clarify these requirements. Those kinds of engagements are also hallmarking for what constitutes mature vocational education systems (Billett, 2013; Deissinger, 1996), in particular, when those engagements include students spending time and engaging in workplace settings and engaging in purposeful and authentic tasks. A key emerging requirement for such engagements now extends to the deliberate development of conceptual and symbolic knowledge that underpins the use of much electronic technology that comprises 21st-century work, yet can be hard to learn.

Securing “hard to learn” knowledge

Much existing and “future work” is likely to be reliant on conceptual and symbolic knowledge (Barley and Batt, 1995; Hull, 1997), including what is held to comprise 21st-century skills, that is, the knowing that needs to be mediated through concepts and symbols, because this kind of knowledge cannot be directly expressed or experienced (Harteis, 2018b). Much of this learning is associated with understandings premised on knowledge that is opaque and difficult to access and, therefore, to learn (Harteis, 2018a). Across human history, and increasingly, this kind of knowledge has been used to represent meanings that cannot be easily stated or represented. The use of symbols to denote chemicals or factors such as those used in mathematical calculations or physics are instances of symbols that require those using them to understand their meaning and be able to manipulate them to achieve goals (Vosniadou et al., 2002). Examples here include the growing use of technologies in fields such as banking, clerical work, as well as healthcare (Hajkowicz et al., 2016). This kind of knowledge is often associated with changes in how work is conducted and how individuals come to mediate it. Scribner (1985), with great prescience, stated: “Hardly have we approached the problem of understanding the intellectual impact of the printing press and we are urged to confront the psychological implications of computerization” (p. 138). Scribner sought to understand these changes through a comparison between the use of manual lathes and those which are computer numerically controlled (CNC lathes; Martin and Scribner, 1991). She noted how these new kinds of lathes required a shift away from those operated through human sensory processes (i.e., vibrations, noise, smell, sight) to those that integrated traditional machining knowledge with symbolic knowledge and logical skills, enacted by computers (Martin and Scribner, 1991). These kinds of knowledge are difficult to learn because there are often too difficult with which to engage. However, they need to be constructed by individuals (i.e., learned) because they become personal tools to mediate work and learning. These forms of knowledge are not restricted to technology: they exist in different forms and ways of knowing for experts, comprising concepts and heuristics upon which they draw.

However, what we know is that these forms of knowledge can be difficult to learn because they are not easy to represent and are often abstracted from the circumstances of use. Therefore, they are difficult to engage with, remember and recall. For instance, science education has long struggled with identifying the most effective ways of representing this kind of knowledge for students to come to understand science and physics. Concerns about force, vectors, stress on physical components et cetera need to be understood through systems that require this kind of conceptual knowledge. Yet, there are practical ways of seeking to make these forms of knowledge accessible, and it is these that need to become part of the considerations for vocational education to address the requirements of what is referred to as 21st-century work.

Responsive vocational education curriculum and pedagogic practices

As foreshadowed, there is a growing need to be responsive in terms of curriculum and pedagogic practices to these challenges. This includes, for instance, providing authentic occupational experiences either in workplaces or through activities within vocational education institutions. Given the limitation of transferring what is learned in substitute environments to where it is to be applied, such experiences are increasingly seen to be important (Ericsson and Lehmann, 1996). A key source is the provision of workplace experiences such as practicums, clinical placements, internships and other forms of work experience (Billett, 2009). It is these

structured experiences that are central to identifying preferred occupations or specialisms, assisting in the development of occupational capacities, and intentionally developing adaptability. However, more than providing experiences in education institutions and workplaces, it is necessary to optimize the integration of specific kinds of learning that arise through the two sets of experiences for students and reconciling them to be generative of adaptable occupational knowledge (Cooper et al., 2010). These processes need to commence before students engage in workplace experiences, during, and then after them. As noted, the developing occupational adaptability likely requires establishing a platform of occupational concepts, procedures and associated values, and exposure to variations and diverse manifestations of those practices (Billett, 2015a). These stand as a means of assisting students and graduates to be aware that occupational knowledge needs to be adapted to the circumstances of practice. Equally, securing knowledge that is hard to learn—conceptual and symbolic knowledge—requires kinds of pedagogic practices and engagements by learners. Those forms of knowledge that are hard to directly experience require specific educational interventions. Such pedagogic practices are not restricted to securing knowledge per se, but also their active engagement by learners. The processes require developing initial occupational competence through reconciling experiences of different kinds through to the ongoing development of that knowledge prior to and during working life, demanding a level of agency by the learner. Learning strategies, rather than teaching, are likely to be required for the ongoing development of workers' knowledge across lengthening working lives; that is, not based as much on teaching, but on curriculum, pedagogic and personal practices associated with learning.

These considerations lead to questions about the kinds of curriculum and pedagogic practices that can support occupational competence, job readiness and adaptability, and that can sustain employability across working life. These issues are addressed in the following sections that offer six considerations: (a) organizing and providing workplace experiences, (b) intentionally and actively integrating students' experiences, (c) institutional-based activities that incite authentic work experiences, (d) educational processes promoting adaptability, (e) securing hard to learn knowledge, and (f) promoting learning agency. It is these six practices that are addressed in turn.

Organizing and providing workplace experiences

With emerging focuses for vocational education on workplace or job readiness and understanding variations of that knowledge, the imperative for having workplace experiences becomes pre-eminent. These experiences provide access to (a) authentic activities and interactions; (b) richly contextualized experiences that engage students in multisensory ways and provide clues and cues about how they need to engage in their work; (c) purposive goal-directed activities aligned with the kinds of knowledge students need to learn for their targeted occupations; (d) goal-directed activities requiring the resolution of problems and through them, learning; (e) securing episodic workplace experiences from which important causal and propositional links are developed; and (f) the ability to monitor the activities in which they engage (Billett, 2015a).

Securing these processes depend on the kinds and duration of workplace experiences, and how students are supported in workplace settings. Nevertheless, these kinds of experiences are not usually accessible within vocational education settings. These work-based experiences make pertinent contributions to students' learning; they are not just complementary experiences provided by educational institutions. Yet, there is a need to integrate the two sets of experiences.

Intentionally and actively integrating students' experiences

As noted, workplaces provide learning experiences and outcomes that are often quite different from those provided and secured through vocational education institutions. Consequently, there is a need to include and integrate these experiences within tertiary educational provisions (Cooper et al., 2010). As noted, the focus here is on organizing and providing these experiences in ways that assist the development of students' occupational capacities (i.e., canonical knowledge), while engaging them with particular instances of practice and the use of instructional strategies to integrate those experiences and learnings into the vocational education program (Orrell, 2011). So, again, there is need to consider both curriculum (i.e., the provision of experiences) and pedagogies (i.e., augmenting educationally those experiences). The findings of a project involving numerous tertiary education institutions on the integration of workplace experiences identified curriculum and pedagogic practices that need to be adopted (Billett, 2011). In Table 1, issues associated with the intended (i.e., what is designed and planned), enacted (i.e., what occurs with implementation) and experienced (i.e., what students experience and learn) curriculum are set out, drawing from those studies.

There are sets of considerations in Table 1 for how the intended curriculum (i.e., what is planned) needs to be organized to optimize the integration of students' experiences in work settings. Also, there is a set of practices associated with the enactment of the curriculum informing practices and priorities associated with this implementation. Then, and perhaps most importantly, are a set of considerations about how students will come to understand and engage with what has been implemented and experienced by them (i.e., the experienced curriculum).

Beyond these curriculum considerations are those associated with pedagogic practices to augment the curriculum experiences through processes that assist students to integrate and reconcile their experience in work settings. In Table 2 are some suggestions regarding educational interventions that might be enacted before, during and after workplace experiences. These comprise preparatory activities prior to students engaging in work settings, and interventions promoting student learning during their workplace experiences. Also, it was found that having interventions after the students have experienced work practice was particularly helpful as students could then compare, share and critically appraise their experiences with their peers (Cardell and Bialocerkowski, 2019). These activities provide bases to develop understandings about the occupational practice, and variations of that practice in action

Table 1 Curriculum considerations for integrating workplace experiences (Billett, 2015a).

<i>Intended curriculum—what is planned</i>	<i>Enacted curriculum—what is implemented</i>	<i>Experienced curriculum—what students experience and learn</i>
Being clear about what is to be learned through workplace experiences	Augmenting or maximizing available opportunities (e.g., appropriate settings)	Students' interest and readiness central to their engagement and learning in practice settings, and reconciling it with their coursework
Aligning experiences provided for students with the intended learning outcomes	Considering options other than supervised placements to secure experiences	Immediate concerns (e.g., performing in practicum) are focus of students' interest
Aligning the duration of experiences with educational purpose (e.g., orientation vs. skill development)	Accounting for students' readiness (e.g., interest, capacities, confidence) when selecting and enacting experiences	Early and staged engagement in practice settings boosts many students' confidence to re-engage and learn effectively
Intentionally sequencing preparatory experiences to secure, consolidate and reconcile learning from practice experiences	Additional or specific experiences may be needed for student cohorts (e.g., overseas students)	Challenges to personal confidence and competence can be redressed by effective group processes, including sharing of experiences

Table 2 Pedagogic strategies promoting integration of workplace learning experiences (Billett, 2015a).

<i>Before workplace experience</i>	<i>During workplace experience</i>	<i>After workplace experience</i>
Orient students to requirements for effectively engaging in work practices	Direct guidance by more experienced practitioners (i.e., proximal guidance)	Facilitate the sharing and drawing out of students' experiences
Clarify expectations about purposes of, support in and responsibilities of parties in practice settings etc.	Active engagement in pedagogically rich work activities or interactions (e.g., handovers)	Make explicit links to, and reconciliations between, what is taught (learned) in the academy, and what is experienced in practice settings
Prepare students to engage as agentic learners (e.g., importance of observations, engagement)	Effective peer interactions (i.e., students' collaborative learning)	Emphasize the active and selective qualities of students' learning through practice
Develop procedural capacities required for tasks in the workplace	Active and purposeful engagement by the students as learners in the workplace	Generate students' critical perspectives on work and learning processes
Prepare for contestations that might arise		

(Williams et al., 2019). This phase also has the potential for students to develop understandings about practice in action and to develop critical capacities to appraise work situations and the efficacy of kinds of practices they might enact, through processes such as reflection (Noble et al., 2019). Moreover, engaging students in this kind of thinking and acting provides the basis for their own personally managed ongoing learning once they engage in their occupational practice in work settings beyond graduation.

It follows from the project's findings that there are actions that can be taken in the design and implementation of vocational education programs that can assist in the effective engagement in and integration of students' experiences in workplaces. Perhaps most important, were the interventions occurring after students had completed their work experiences as these provided bases for students to develop knowledge from what they had experienced and engage critically, comparatively and in an informed way with peers (Billett, 2019).

The two sections above emphasize the importance of providing and integrating vocational education students with authentic experiences associated with their preferred occupations and, wherever possible, with access to some variations of those experiences. It has also been suggested that educational processes can reconcile experiences in work settings with those in the vocational education situations. However, it is not always possible to provide students with workplace experiences: Some countries have very strong traditions of providing workplace experiences, whereas others do not. So, there are also considerations for what can happen within vocational education institutions to provide authentic experiences.

Institutional-based activities that incite authentic work experiences

Not all vocational education students can access work experiences, perhaps for institutional and cultural reasons. So, while it is very common in Germany for vocational students to engage in significant work experiences, in the Netherlands there is not such a strong tradition of providing students with work experiences. Instead, vocational education institutions and applied science universities there have a tradition of providing what is referred to as hybrid experiences within those institutions (Zitter et al., 2017). Students are provided with simulation or project work intended to equate to the kinds of activities occurring in occupational practice. This is also the case in Singapore where many of the postsecondary education institutions (i.e., the polytechnics and the Institute for Technical Education) also provide project-type activities within the institutional setting. For instance, in one of the polytechnics, the information technology students engage in assisting other students with setting up their laptops and tablets with the institution's

electronic systems and engage in a helpdesk and troubleshooting when students have problems. That is, the information technology students are engaging in authentic activities associated with that domain of work.

There are two broad considerations here: curriculum and pedagogy. In terms of curriculum—the activities that are organized for students and their sequencing—consideration of shared or individual projects associated with the field of study, simulated activities (e.g., training restaurants, legal centers, broadcast and print media, healthcare clinics) can provide students with educational activities that are closely linked with their intended occupation (Zitter and Hoeve, 2012). This approach engages students in the kinds of goal-directed activities associated with their preferred occupation. They can have high levels of authenticity in terms of the knowledge targeted for them to learn. Also, these kind of projects and simulated activities engage students actively in educational processes, thereby providing experiences and realizing outcomes that more passive forms of education are unlikely to be able to achieve. That is because students are placed in the “driver’s seat”: they must make decisions and complete actions and monitor and evaluate the efficacy of those actions. This kind of engagement and of thinking and acting is necessary for developing the higher order capacities required for many occupations (Evans, 1991; Stevenson, 1991). So, these kinds of experiences can be made authentic in terms of the sorts of activities and interactions and can assist students in generating the kind of understandings that allow them to learn that knowledge in ways that demand knowledge utilization, stimulate recall, and are aligned with performing occupational tasks upon graduation.

In terms of pedagogic practice, establishing these projects and ensuring that the criteria for them and support for their enactment are adequate, are all important. As noted above, a central consideration for pedagogic practices is to organize and implement activities for students that engage them in the kinds of thinking and acting that are required for the occupational practice and in ways that progressively engage them in more demanding and effortful activities. For instance, when I was teaching in vocational education in the field of garment design and manufacturing, students would commence with designing and producing garments that were relatively easy to complete (e.g., skirts), before designing and producing more complex garments such as dresses, jackets, suits, et cetera. The pedagogic practices here are to carefully select and organize the activity, support students’ engagement in it, and then guide their thinking and acting as they progress with the activity. Added here are teachers’ ability to use narratives and storytelling and to verbalize the kinds of knowledge needing to be learned. When they use instances from practice it is likely to be particularly helpful for students to learn and recall what has been taught, as it provides a basis for indexicality: the means by which individuals organize and recall their knowledge, which is essential for expert performance (Ericsson and Lehmann, 1996). Activities that emphasize the utilization of the knowledge outside of the circumstances in which is being taught will assist students to learn about the applicability of that knowledge through providing experiences which are not able to be directly engaged by them. Hence, educational processes that promote adaptability are important.

Educational processes promoting adaptability

A fundamental goal for education is for the capacities that have been developed through it to be adapted to circumstances beyond where they have been learned, as noted. Consequently, a key concern for vocational education is for students to develop the capacities to adapt what they have learned to activities and circumstances beyond the institution. This assists in developing their ability to apply what they have learned to its application in work practices, and also across the length of their working lives, as they face changes in occupational and workplace practices. So, adaptability is foundational and fundamental, if not the defining quality of educational provisions. Yet, there are specific challenges associated with hard-to-learn knowledge and the development of conceptual and symbolic capacities required for 21st-century work that challenge such adaptability. These kinds of learning are not easily derived from direct teaching. Instead, other kinds of experiences are often required, albeit supported by teacherly practices. A key consideration for vocational education is making accessible and comprehensible the conceptual and symbolic knowledge. This kind of access can be provided using stories, analogies, explanations, and illustrations by teachers, drawing upon specific instances of practice that exemplify and illustrate the concepts being used. A key curriculum goal is, therefore, the development of adaptable understandings and practices (i.e., informed principles and practices) that promote adaptability. This comprises identifying and assisting students to develop understandings and procedures that optimize their ability to adapt what they know, can do and value to changing circumstances. These outcomes can be realized in a range of ways, one of which is the direct teaching of such principles and procedures, but most likely supported with authentic examples and instances of work. Then, mnemonics derived from practice can be used to assist with the development of procedural capacities (i.e., to assist recall of a comprehensive range of potential considerations). Another approach is for students to share with peers their different kinds of work experiences and engage in identifying what is common across the enactment of that occupation (i.e., the canonical knowledge of the occupation) and what is specific to work settings (i.e., situational requirements). This is an activity that can be undertaken once students have had workplace experiences. It is the combination of developing canonical occupational knowledge alongside the understandings of variations and situational requirements that provides the bases for using occupational knowledge adaptively. Engaging with other students to compare their experiences of work settings, albeit facilitated by teachers, can open up a range of options and possibilities that permit them to realize that there are variations in that practice, and for what reasons and how these might be accommodated within enactment of the occupation (Billett, 2015a). Consequently, providing opportunities for students to engage in sharing, discussion and dialog are likely to be important, as are, if possible, students rotating through different kinds of work settings and circumstances of occupational practice and having opportunities to consider what is common and distinct about performance requirements in these different circumstances. For instance, for a long time, nurse education was undertaken wholly within hospitals with trainee nurses rotating across different kinds of hospital wards. This array of experiences provides the opportunity to

understand the canonical knowledge of nursing, on the one hand, and on the other hand to identify specific nursing requirements in different kinds of wards. Even then, those experiences alone may be insufficient unless there is the opportunity to consider, discuss and extend knowledge about those practices.

This concern with adaptability is fundamental to the educational project. Educational institutions and their programs have been established not just for learning that is relevant to them: the learning needs to be applicable to circumstances and activities that are distinct from those enacted in educational institutions. Consequently, focusing on strategies that seek to extend the applicability of what students come to know, do and value is likely to be a crucial consideration in vocational teaching. This is particularly relevant in the contemporary era where change in workplace and occupational requirements occurs so frequently.

Securing “hard-to-learn” knowledge

A specific quality of the 21st-century work referred to above is the increased use of conceptual and symbolic knowledge as part of workplace performance requirements. This kind of knowledge has increased with the use of technology and is often what mediates work performance when using tools and artifacts. However, there are three concerns associated with learning conceptual and symbolic knowledge. Firstly, it can be difficult to capture in written form (e.g., force, electronic systems, information processing). Secondly, it can be difficult to access because there is no direct engagement with it (e.g., it cannot be directly experienced). Thirdly, these kinds of knowledge are often difficult to be taught, and therefore need to be learned. Science education has long struggled to address this problem in school classrooms: that is, how can students learn concepts that they cannot directly experience? Literature from that field suggest the importance of making that knowledge explicit in some way (through the use of examples, instances, representations; able to be experienced or engaged with or even visualized and then having students construct meaning from those explicit representations; [Billett, 2018](#)). One aspect of this is the lessening need for the direct use of tools that must be manipulated manually and the importance of ergonomic type capacities such as deftness of hand movements ([Harteis, 2018a](#)). As foreshadowed and as elaborated below, much of the conceptual and symbolic knowledge required for contemporary work cannot be experienced directly through physical sensation. Consequently, the means of it being learned may require kinds of interventions to promote its development ([Vosniadou et al., 2002](#)). It is necessary, therefore, to promote learners’ agency to be effective, outward looking and intentional in the activities and interactions from which they learn ([Goller, 2017](#)). It is these educational responses that are used to propose how vocational education can proceed with developing skillful and adaptable workers.

For vocational education, more broadly, the kinds of pedagogic divisions and practices required for developing conceptual knowledge are as follows. Firstly, consideration needs to be given to the kind of experiences and forms of learner engagement that are most likely to be able to represent that knowledge, on the one hand, and on the other, having processes that students come to engage with it. It has been suggested ([Brown et al., 1989](#)) that developing a conceptual model of the task is important as this provides learners with:

- an advanced organizer for attempting to execute the task;
- bases to utilize feedback, hints and corrections during interactions;
- an internalized guide for independent practice by successive approximations; and
- a conceptual model which can be updated ([Collins et al., 1989](#)).

These kinds of considerations can be helpful as students construct this “hard-to-learn” kind of knowledge. All this needs to be underpinned by the provision of authentic activities that provide concrete instances of how that knowledge needs to be utilized and for what purposes and these activities also assist with indexicality (i.e., the construction of knowledge in ways that permit its utilization and recall). So, it is these kinds of pedagogic practices that are likely to be helpful for developing this “difficult to access” knowledge that is becoming increasingly part of contemporary work life.

Promoting learning agency

Across the considerations advanced above is the enduring focus on learner engagement. We are reminded here that vocational education provisions are nothing more or less than invitations to change. How students take up these invitations is fundamental to the effectiveness of any educational provisions. So, student engagement and agency are central to effective vocational education, and to how graduates come to adapt their occupational or work-related knowledge to circumstances beyond the educational provision and engage in the kind of effortful learning required for productive working life ([Goller, 2017](#)). That agency shapes the focus, intensity and direction of vocational education students’ learning. However, there are some key premises associated with promoting learning agency that are outlined briefly here. Firstly, student readiness (i.e., their ability to engage in activities), and an awareness of that readiness, is an essential starting point for encouraging students to be proactive, interdependent and directed in their activities and learning ([Billett, 2015b](#)). Secondly, having goals and processes that warrant student engagement seems to be important. That is, having activities that are relevant to what needs to be learned, interesting enough to engage students and to be perceived by them to be sufficiently worthwhile to engage them in the kinds of effortful processes required to generate the linkages and associations among conceptual knowledge, and the associated considerations required for strategic procedural practices. Central to that engagement is developing a strong sense of subjectivity associated with the occupation that will encourage students to engage effortfully in its learning. That is, an important goal is to assist students in coming to recognize their occupation as their vocation. Fourthly, and consequently, selecting and providing experiences that engage learners and that they would be interested in engaging effortfully is

likely to be required for the kinds of learning needed to become adaptable occupational practitioners. Part of that curriculum and pedagogic process is to place learners in the circumstances where they must take responsibility for their thinking, acting and learning. This does not suggest or imply that there is no role for teachers. Instead, teachers' role is to guide and support vocational education students, but primarily, the focus is on the agency and interdependence of learners.

Practical consideration for developing these capacities is having students evaluate their peers' processes and outcomes and this kind of activity can develop the kinds of evaluative and critical capacities that people need to develop to effectively monitor and evaluate their own work practices.

Developing a skillful and adaptable workforce

It has been proposed in this article that changes in occupational workplace requirements and working life, and specifically the development of what is referred to as 21st-century skills, prompt a reappraisal of the goals and processes of vocational education. A broader view of curriculum and pedagogies need to be considered, engaged with and enacted to accommodate changes in educational goals for vocational education. These include addressing the specific requirements of workplaces as well as developing occupational competence; learning knowledge that is difficult to experience directly (i.e., conceptual and symbolic knowledge) but that is required for what is often referred to as knowledge work; developing adaptive occupational capacities because the requirements for work and work performance are constantly changing and because of the importance for students to become active and intentional learners for their initial preparation, but also that ongoing development across working life. It is proposed here that a way forward is to adopt curriculum and pedagogic practices that are aligned with achieving these kinds of outcomes. This includes a consideration of what constitutes effective educational experiences (both within educational institutions and workplaces), ordering and reconciling these two sets of experiences, the use of educational interventions that can generate the kinds of capacities within vocational education students and assisting students to become active and intentional learners across their working life. Educational interventions are likely to be required to address the growing elements of occupational hard-to-learn knowledge. Promoting learner agency and interdependence is likely to be an important educational outcome, not just for immediate employability, but for learning across working life.

To do this requires consideration of what constitutes the existing and emerging requirements for occupational and workplace performance and then aligning these with the kinds of curriculum and pedagogic practices that vocational education institutions need to advance and the kinds and quality of engagement that students need to adopt and practice. These concerns are not just about individuals' personal learning: they extend to the efficacy of work practices, workplaces and communities.

In all, educational goals and processes are needed for developing adaptive capacities on the occupational plane (e.g., informed principles and practices) and active and intentional learning on the personal epistemological plane. It is these kinds of curriculum, pedagogic and personal concerns that may be effective in securing the kinds of outcomes required for vocational education provisions, whether in schools or colleges, to be responsive to the emerging occupational and workplace requirements often referred to as 21st-century work requirements.

References

- Alexander, P.A., Judy, J.E., 1988. The interaction of domain specific and strategic knowledge in academic performance. *Rev. Educ. Res.* 58 (4), 375–404.
- Anderson, J.R., 1982. Acquisition of cognitive skill. *Psychol. Rev.* 89 (4), 369–406.
- Bantock, G.H., 1980. *Dilemmas of the Curriculum*. John Wiley & Sons, New York.
- Barley, S., Batt, R., 1995. *The New Crafts: The Rise of the Technical Labor Force and Its Implication for the Organization of Work*. University of Philadelphia, National Center on the Education Quality of the Workforce.
- Bartlett, F.C., 1958. *Thinking: An Experimental and Social Study*. Basic Books, New York.
- Billett, S., Harteis, C., Gruber, H., 2018. Developing occupational expertise through everyday work activities and interaction. In: Ericsson, K.A., Hoffman, R.R., Kozbelt, A. (Eds.), *Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, New York, pp. 105–126.
- Billett, S., 1998. Situation, social systems and learning. *J. Educ. Work* 11 (3), 255–274.
- Billett, S., 2001. Knowing in practice: re-conceptualizing vocational expertise. *Learn. Instruct.* 11 (6), 431–452.
- Billett, S., 2009. Realizing the educational worth of integrating work experiences in higher education. *Stud. High Educ.* 34 (7), 827–843.
- Billett, S., 2011. *Curriculum and Pedagogic Bases for Effectively Integrating Practice-Based Experiences*. Retrieved from Office of Teaching and Learning, Sydney.
- Billett, S., 2013. Towards a mature provision of vocational education. *Int. J. Train. Res.* 11 (2), 184–194.
- Billett, S., 2014. Interdependence on the boundaries between working and learning. In: Harteis, C., Rausch, A., Seifried, J. (Eds.), *Discourses on Professional Learning: On the Boundary Between Learning and Working*. Springer, Dordrecht, The Netherlands, pp. 369–385.
- Billett, S., 2015a. *Integrating Practice-Based Experiences Into Higher Education*. Springer, Dordrecht, The Netherlands.
- Billett, S., 2015b. Readiness and learning in healthcare education. *Clin. Teach.* 12, 1–6.
- Billett, S., 2017. Developing domains of occupational competence: workplaces and learner agency. In: Mulder, M. (Ed.), *Competence-Based Vocational and Professional Education: Bridging the Worlds of Work and Education*. Springer International Publishing, Cham, Switzerland, pp. 47–66.
- Billett, S., 2018. Accessing and securing conceptual and symbolic knowledge required for digital era work. In: Harteis, C. (Ed.), *The Impact of Digitization in the Workplace—An Educational View*. Springer, Dordrecht, The Netherlands, pp. 197–212.
- Billett, S., 2019. Augmenting post-practicum experiences: purposes and practices. In: Billett, S., Newton, J.M., Rogers, G.D., Noble, C. (Eds.), *Augmenting Health and Social Care Students' Clinical Learning Experiences*. Springer, Dordrecht, pp. 3–25.
- Bishop, A., 1991. *Mathematical Enculturation: A Cultural Perspective on Mathematics Education*, vol. 6. Springer Science & Business Media.
- Brown, J.S., Collins, A., Duguid, P., 1989. Situated cognition and the culture of learning. *Educ. Res.* 18 (1), 32–34.

- Cardell, E., Bialocerkowski, A., 2019. Bouncing forward: a post-practicum workshop to promote professional identity, self-efficacy, and resilience in Master of Speech Pathology students. In: Billett, S., Newton, J.M., Rogers, G.D., Noble, C. (Eds.), *Augmenting Health and Social Care Students' Clinical Learning Experiences: Outcomes and Processes*. Springer, Dordrecht, pp. 211–234.
- Carraher, T.N., Carraher, D.W., Schliemann, A.D., 1985. Mathematics in the streets and in schools. *Br. J. Dev. Psychol.* 3 (1), 21–29.
- Chase, W.G., Simon, H.A., 1973. The mind's eye in chess. In: Chase, W.G. (Ed.), *Visual information processing*. Academic Press, New York, pp. 215–281.
- Chi, M.T.H., Glaser, R., Rees, E., 1982. Problem-solving ability. In: Sternberg, R.J. (Ed.), *Advances in the Psychology of Human Intelligence*, vol. 1. Erlbaum, Hillsdale, NJ, pp. 7–76.
- Collins, A., Brown, J.S., Newman, S.E., 1989. Cognitive apprenticeship: teaching the crafts of reading, writing and mathematics. In: Resnick, L.B. (Ed.), *Knowing, Learning and Instruction: Essays in Honour of Robert Glaser*. Erlbaum & Associates, Hillsdale, NJ, pp. 453–494.
- Cooper, L., Orrel, J., Bowden, M., 2010. *Work Integrated Learning: A Guide to Effective Practice*. Routledge, London.
- Deissinger, T., 1996. Germany's Vocational Training Act: its function as an instrument of quality control within a tradition-based vocational training system. *Oxf. Rev. Educ.* 22 (3), 317–336.
- Department of Innovation, Universities and Skills, 2008. *Higher Education at Work: High Skills: High Value*. Department of Innovation, Universities and Skills, Sheffield.
- Ericsson, K.A., Lehmann, A.C., 1996. Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu. Rev. Psychol.* 47 (1), 273–305.
- Ericsson, K.A., Hoffman, R.R., Kozbelt, A. (Eds.), 2018. *Cambridge Handbook of Expertise and Expert Performance*, second ed. Cambridge University Press, New York.
- Evans, G., 1991. Lesson cognitive demands and student processing in upper secondary mathematics. In: Evans, G. (Ed.), *Learning and Teaching Cognitive Skills*. ACER, Melbourne, Australia, pp. 124–143.
- Faure, E., Herrera, F., Kaddoura, A., Lopes, H., Petrovsky, A., Rahnema, M., Ward, F., 1972. *Learning to Be*. Retrieved from Paris, OECD.
- Gelman, R., Greeno, J.G., 1989. On the nature of competence: principles for understanding in a domain. In: Resnick, L.B. (Ed.), *Knowing, Learning and Instruction: Essays in Honour of Robert Glaser*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 125–186.
- Glaser, R., 1984. Education and thinking—the role of knowledge. *Am. Psychol.* 39 (2), 93–104.
- Goller, M., 2017. *Human Agency at Work: An Active Approach Towards Expertise Development*. Springer Fachmedien, Wiesbaden, Germany.
- Hajkowicz, S., Reeson, A., Rudd, L., Bratanova, A., Hodggers, L., Mason, C., Boughen, N., 2016. *Tomorrow's Digitally Enabled Workforce: Megatrends and Scenarios for Jobs and Employment in Australia Over the Coming Twenty Years*. CSIRO, Canberra.
- Harteis, C., 2018a. Machines, change, work: an educational view on the digitalization of work. In: Harteis, C. (Ed.), *The Impact of Digitalization in the Workplace: An Educational View*. Springer, Cham, pp. 1–10.
- Harteis, C. (Ed.), 2018b. *The Impact of Digitalization in the Workplace*, vol. 21. Springer, Dordrecht.
- Hull, G., 1997. Preface and introduction. In: Hull, G. (Ed.), *Changing Work, Changing Workers: Critical Perspectives on Language, Literacy and Skills*. State University of New York Press, New York, pp. 3–39.
- Marsh, C.J., 2004. *Key Concepts for Understanding Curriculum*. RoutledgeFalmer, London.
- Martin, L.M.W., Scribner, S., 1991. Laboratory for cognitive studies of work: a case study of the intellectual implications of a new technology. *Teach. Coll. Rec.* 92 (4), 582–602.
- Mayer, E.C., 1992. *Employment-Related Key Competencies for Post-compulsory Education and Training—A Discussion Paper*. Australian Government Printing Service, Canberra, Australia.
- Newman, D., Griffin, P., Cole, M., 1989. *The Construction Zone: Working for Cognitive Change in Schools*. Cambridge University Press, Cambridge, UK.
- Noble, C., Armit, L., Collier, L., Sly, C., Hilder, J., Molloy, E., 2019. Enhancing feedback literacy in the workplace: a learner-centered approach. In: Billett, S., Newton, J.M., Rogers, G.D., Noble, C. (Eds.), *Augmenting Health and Social Care Students' Clinical Learning Experiences: Outcomes and Processes*. Springer, Dordrecht, pp. 283–308.
- Organisation for Economic Co-operation and Development, 2013. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD, Paris, France.
- Orrell, J., 2011. *Good Practice Report: Work Integrated Learning*. Office of Teaching and Learning, Sydney, Australia.
- Perkins, D., Jay, E., Tishman, S., 1993. Beyond abilities: a dispositional theory of thinking. *Merrill-Palmer Q.* 39 (1), 1–21.
- Prawat, R.S., 1989. Promoting access to knowledge, strategy, and dispositions in students: a research synthesis. *Rev. Educ. Res.* 59 (1), 1–41.
- Scribner, S., 1985. Knowledge at work. *Anthropol. Educ. Q.* 16, 199–206.
- Sennett, R., 2008. *The Craftsman*. Yale University Press, New Haven & London.
- Skilbeck, M., 1984. *School Based Curriculum Development*. Harper and Row, London.
- Stevenson, J.C., 1991. Cognitive structures for the teaching of adaptability in vocational education. In: Evans, G. (Ed.), *Learning and Teaching Cognitive Skills*. ACER, Victoria, Australia, pp. 144–163.
- Sun, R., Merrill, E., Peterson, T., 2001. From implicit skills to explicit knowledge: a bottom-up model of skill development. *Cogn. Sci.* 25, 203–244.
- Sweller, J., 1990a. Cognitive processes and instructional procedures. *Aust. J. Educ.* 34 (2), 125–130.
- Sweller, J., 1990b. On the limited evidence for the effectiveness of teaching general problem-solving strategies. *J. Res. Math. Educ.* 21 (5), 411–416.
- The Secretary's Commission on Achieving Necessary Skills (SCANS), 1992. *Learning a Living: A Blueprint for High Performance*. US Department of Labor, Washington, DC.
- Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., Papademetriou, E., 2002. Designing learning environments to promote conceptual change in science. *Learn. Instruct.* 11 (4–5), 381–419.
- Wertsch, J.W., 1991. A sociocultural approach to socially shared cognition. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC, pp. 85–100.
- Williams, L.M., Ross, L., Mitchell, L., Markwell, K., 2019. The reflective debrief: using students' placement experiences to enrich understandings of distinct kinds of nutrition and dietetics practice. In: Billett, S., Newton, J.M., Rogers, G.D., Noble, C. (Eds.), *Augmenting Health and Social Care Students' Clinical Learning Experiences: Outcomes and Processes*. Springer, Dordrecht, pp. 259–281.
- Zitter, I., Hoeve, A., 2012. *Hybrid Learning Environments: Merging Learning and Work Processes to Facilitate Knowledge Integration and Transitions*. OECD Publishing, Paris.
- Zitter, I., Hoeve, A., de Bruijn, E., 2017. A design perspective on the school-work boundary: a hybrid curriculum model. *Vocat. Learn.* 9 (1), 111–131.

The future of internships and workplace-based learning

Leopold Bayerlein, UNE Business School, University of New England, Armidale, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Definition of internships	462
Internships in education	462
Identification and attainment of learning outcomes	463
Contextual changes impacting internships	464
Emergence of technology-based internships	464
Expansion of internships to Work-Integrated Learning	465
Recommendations for future internship developments	465
Preparing participants for technology-based workplaces	466
Reducing participation barriers	466
Focusing experiences on workplace-readiness competencies	467
References	467

Definition of internships

Within an educational context, Internships, co-ops and work experiences (hereafter: Internships) represent workplace-based learning experiences that develop, extend and apply participants contextualized skills and knowledge. Internships are conceptually embedded within the larger field. Internships represent a sub-category of Work-Integrated Learning because they generally involve the “placement” of a participant within a placement organization for the purpose of undertaking a specific learning experience (Bayerlein et al., 2021b).

The scale (for example measured in hours of commitment per time period), duration and income generation potential of Internships may vary considerably between different placements, industries and contexts. Similarly, the nomenclature used to describe workplace-based learning experiences, including the terms Internships, co-ops and work experiences, may vary in their definition across different contexts (Bayerlein, 2019). To avoid confusion related to contextually dependent nomenclature, this work defines Internships through the educational “substance” that underpins particular workplace-based learning activities, rather than the “form” or “descriptor” through which these activities are operationalized. To be classed as Internships, activities require placement providers, participants and education providers to identify, and agree-upon, the learning outcomes of the Internship, the learning activities through which learning outcomes are to be achieved, and how learning within the Internship is to be assessed (also, see: Jackson, 2015). This focus on the “substance” of the activity, evidenced through the development, implementation and assessment of a structured learning experience for participants, represents the distinguishing feature of Internships, and separates Internships from other work activities where participant learning, as well as the assurance of that learning, is absent, incidental or unstructured.

Internships in education

Governments as well as employer organizations are increasingly expecting graduates of education programs to meet a range of workplace-competence or employability criteria. This includes transferable graduate skills, such as problem solving, workplace agility and the ability to undertake work independently with minimal direct supervision, as well as occupation specific skills related to frameworks, processes and systems that are used within a particular occupation. The growing importance of the graduate skills and knowledge that are associated with workplace-readiness has been recognized in the literature for some time (for example, see: Arum and Roksa, 2011; Jackson et al., 2014; Rosenberg et al., 2012). Despite this recognition, providers of education programs have received persistent criticism for producing graduates that appear inadequately prepared for the world of work (for example, see: Coll and Zegwaard, 2006; Fleming, 2008; Jackson et al., 2017).

Internships, and Work-Integrated Learning activities more generally, are regarded as one important mechanism through which graduate workplace-readiness can be improved (Knouse and Fontenot, 2008; Silva et al., 2018; Jackson and Bridgstock, 2021). Internships are, in principle, able to support the development of learning outcomes for students in almost all discipline areas (Candy and Crebert, 1991; Maertz et al., 2014). The specific learning outcomes achieved within a particular Internship depends on the design of the placement, as well as the context in which the placement is undertaken (Jackson, 2015). However, prior literature has demonstrated a vast array of potential benefits, ranging from the general preparation of participants for the demands of a particular workplace or industry (Guile and Griffiths, 2001; Jackson, 2015; Rosenberg et al., 2012; Wilton, 2012), and the application of theoretical frameworks, process and systems knowledge in practical settings (Jackson et al., 2014; Kavanagh and Drennan, 2008), to the development of transferable conceptual graduate skills, such as problem solving, communication and agility, which are applicable across a range of occupations (Maertz et al., 2014; Orrell, 2011; Smith and Worsfold, 2015).

Practical employment focused benefits of Internships that may arise in addition to the Internship learning outcomes achieved by participants relate to increased employment opportunities (Hergert, 2009), and the ability of participants to assess their own suitability for the employer organization, industry or work-environment in which the Internship takes place (Rothman and Sisman, 2016). While it is possible for interns to attain these additional benefits, they are not integral to the Internship experience, unless they relate to a specific Internship learning outcome. In addition, recent literature (Jackson and Bridgstock, 2021) has highlighted that the attainment of improved employment opportunities through Internships, which are generated through improvements in industry networks, requires the explicit integration of activities that support such an outcome in the Internship.

Internships that require participants to undertake real-world work placements within a placement organization, and physically attend a particular placement location for extensive face-to-face on-site interactions between Internship participants and other employees, have traditionally been the most common Internship format (Bayerlein and Jeske, 2018b). The success of such Internships is well documented, and their benefits are widely accepted by education providers and employers across numerous occupations (Maertz et al., 2014; Kraft et al., 2019). However, in recent years Internship practitioners and researchers have increasingly noted the challenges associated with traditional Internships, including the location-boundedness (Bayerlein and Jeske, 2018a,b), limited accessibility (Maertz et al., 2014; Kraft et al., 2019) and limited availability (Maertz et al., 2014) of internships, as well as the potential benefits resulting from the application of new technologies and innovative Internship concepts. As a result, alternative ways to structure and implement Internships, and Work-Integrated Learning activities more generally, have gained substantial interest in the literature (Kay et al., 2019; Silva et al., 2018; Bayerlein et al., 2021b). Assessments of innovative Internship concepts have highlighted that such Internship formats should principally be able to generate learning outcomes that are comparable to traditional Internships (Bayerlein and Jeske, 2018b; Zegwaard and Rowe, 2019), while addressing some of the major accessibility and relevance concerns surrounding traditional Internship activities (Perkiss et al., 2019; Bayerlein, 2015; Bayerlein and Jeske, 2018a). However, empirical evidence that evaluates the participant learning outcomes of different emerging Internship (for example see: Bayerlein, 2020) and Work-Integrated Learning (for example, see: Perkiss et al., 2020; Wersun et al., 2019) formats remain limited.

Identification and attainment of learning outcomes

Regardless of the format through which Internships and Work-Integrated Learning is delivered to participants, the mechanisms through which Internships create benefits for participants are similar in all contexts. The employed mechanisms revolve around the integration of participants in work environments, as well as the provision of deliberate support structures by education providers and internship providers to onboard, support and exit participants from the learning activity. The creation of learning outcomes for Internship participants, which represent the focus of the participant experience, may be viewed through a variety of lenses, which are linked to the intentions that underpin a particular Internship program or opportunity. All Internship activities require an inherent focus on attaining an identified, and agreed-upon, set of participant learning outcomes (Jackson, 2015). Despite this commonality, individual Internships may be underpinned by differing intentions, including the improvement of participants generalist workplace competencies (Prikshtat et al., 2019), or occupation, profession or employer organization specific employability skills (Gardner and Liu, 1997).

Prior literature examining the impact of Internships on participants has identified a range of positive aspects that are implicitly related to workplace competencies (for example, see: Maertz et al., 2014; Rosenberg et al., 2012; Smith and Worsfold, 2015), and employability (for example, see: Hergert, 2009; Rothman and Sisman, 2016), or both (for example, see: Bayerlein, 2015, 2020; Jackson, 2015). However, much of the prior literature has either used the concepts of workplace competencies and employability interchangeably, or been unclear about the conceptual lens(es) employed. While this lack of clarity about the intention(s) and conceptual lens(es) that underpin an individual Internship experience does not detract from the highly positive and far reaching impacts of Internships in general (Hergert, 2009), both research and practice are likely to benefit from a clear identification of the conceptual underpinnings and intentions of individual Internships.

Internships may, in principle, be divided into (1) experiences that transition participants into specific employment opportunities through the development of role-specific employability skills (Gardner and Liu, 1997), and (2) experiences that develop participants general/transferable workplace competencies (Prikshtat et al., 2018, 2019). Employability focused Internships are likely to be most successful if their learning outcomes are closely aligned with the requirements of a particular employment opportunity (Gardner and Liu, 1997), employer organization or industry. Practical participant benefits arising from employability focused Internships include the application of theoretical knowledge in specific practical settings (Bayerlein, 2015, 2020), increased employment opportunities (Hergert, 2009), and opportunities for participants to self-assess their suitability for their chosen career (Rothman and Sisman, 2016).

Workplace-competency focused Internships are likely to be most successful if their learning outcomes are focused on the integrated development of transferable competencies that are critically important in contemporary, as well as future, workplaces (Prikshtat et al., 2019). This focus differs from employability based Internships because it distinguishes the development of competencies, and the ability to consistently apply the developed competencies at a required level, from the context in which competencies are developed, practiced and assessed (Prikshtat et al., 2019). As a result, the learning outcomes of workplace-competency based Internships, such as problem solving, communication skills and workplace agility, arise out of identified, and agree-upon, transferable

skills, with their contextualized application being reflective of the specific environment, but not the learning outcomes, in which a specific Internship experience is situated.

The different mechanisms through which the impact of employability based and workplace-competency based Internships is created are also important from an experience design and assurance perspective. When designing (for example see: [Wiggins and McTighe, 1998](#)) or assuring (for example, see: [Biggs, 1996](#)) learning, the identification the specific learning outcomes to be achieved should act as the starting point for all future decisions. For employability-based Internships, such learning outcomes are likely to be based on (1) the needs of an individual employer organization or industry, and (2) the participants demonstrated ability to meet those needs. This focus, in turn, implies that learning outcomes are dictated by the workplace context in which the Internship occurs. For workplace-competency based Internships, learning outcomes are likely to be created through (1) the skills and knowledge that participants must possess to be successful in a variety of workplaces, and (2) the participants demonstrated ability to apply these skills and knowledge in one or more practical contexts. This focus, in turn, implies that learning outcomes are dictated by beliefs about transferable conceptual workplace skills, and that the context in which these skills are developed, applied and assessed is secondary to the participants learning.

Contextual changes impacting internships

Providers of education have a strong incentive to increase the volume and quality of Internship opportunities for participants ([Jackson et al., 2017](#)). At the same time, the breadth of Internship experiences available to participants is growing to meet the increasingly fragmented and specific needs of participants, education providers and placement providers ([Kay et al., 2019](#); [Kraft et al., 2019](#)). An important reason for the changing nature of Internships are the structural changes occurring in the world-of-work ([Bayerlein et al., 2021a](#)). Workplaces world-wide are undergoing a rapid transition toward technology-based structures, and it important that Internships, as well as Work-Integrated Learning activities in general, continue to evolve to reflect contemporary workplace requirements and practices ([Dean and Campbell, 2020](#)).

Two important workplace trends that create an immediate impact on Internships are the increasing application of, and reliance on, technology across virtually all workplaces ([Bayerlein et al., 2021a](#)), as well as the growing importance of transferable professional skills to meet increasingly generalist workplace expectations ([Hite and McDonald, 2020](#)). Although both trends have impacted workplaces for some time, the COVID-19 pandemic has changed the previously slow transition process into a rapid adoption activity ([Carroll and Conboy, 2020](#); [Bayerlein et al., 2021a](#)). Literature suggests that the COVID-19 pandemic will have long-term impacts on society ([Sneader and Singhal, 2020](#)), as well as workplaces and employment ([Hite and McDonald, 2020](#); [Yawson, 2020](#)). Of particular concern for Internships is emerging evidence that many of the workplace changes, including technology based workplace changes, caused by the COVID-19 pandemic are likely to persist into the future ([Yawson, 2020](#)).

Internships are well suited to prepare participants for emerging workplace structures and requirements because they embed participants into current workplace practices ([Jackson et al., 2017](#)). It is likely that the relative importance of Internships for education providers, Internship participants and Internship providers increases because providers of education have historically found the development of curricula focused on future workplace-competencies difficult ([Bayerlein et al., 2021a](#)). The increasing implementation of Internships in education programs addresses this concern by enabling participants to develop the skills, knowledge and competencies that are required in contemporary, as well as emerging, workplaces ([Dean and Campbell, 2020](#)) without the need to undertake extensive changes within the remaining educational curriculum.

Emergence of technology-based internships

The structures, processes and arrangements underpinning modern workplaces are in a constant state of change. Traditionally, the implementation of changes, including the uptake of technology as the basis upon which workplaces are organized, has been relatively slow. However, the COVID-19 pandemic has accelerated the rate of change, and in particular the rate of technology adoption across virtually all workplaces. The impact of this rapid (for example, see: [Carroll and Conboy, 2020](#)) and persistent (for example, see: [Yawson, 2020](#)) change of workplace arrangements is well captured by the application of the technology use continuum of [Schuster and Glavas \(2017\)](#).

[Schuster and Glavas \(2017\)](#) posit that technology application within a workplace falls within four broad categories: *technology-supported*, *technology-facilitated*, *technology-blended*, and *technology-based*. Workplaces generally have been transitioning from the lower technology reliance categories (technology-supported and technology-facilitated) to higher technology reliance categories (technology-blended and technology-based) for some time ([Bayerlein et al., 2021a](#)). This shift of workplace arrangements along the [Schuster and Glavas \(2017\)](#) technology use continuum impacts the skills, knowledge and competencies required in contemporary and emerging workplaces, and it increases the importance of appropriate technology use in all facets of future Internship arrangements ([Bayerlein et al., 2021a](#)).

The rate with which emerging technologies have been adopted in education generally ([Alexander et al., 2019](#)), and in the delivery of technology enabled Internships and Work-Integrated Learning programs specifically ([Kay et al., 2019](#)), has historically been relatively slow ([Bayerlein et al., 2021b](#)). Despite the relatively slow adoption rate, technology application in Internship arrangements have generally mirrored wider changes in the way workplaces are organized (for example, see: [Bayerlein et al., 2021a](#)), and technology-blended or technology-based Internships have become normalized ([Bayerlein et al., 2021b](#)). While the

world-wide impact of the COVID-19 pandemic has undoubtedly accelerated (Carroll and Conboy, 2020) the transition of Internships along the technology use continuum of Schuster and Glavas (2017), technology-based Internships have been recognized as a discrete Internship formats for some time (Van Dorp, 2008; Jeske and Axtell, 2014, 2016, 2018).

The adoption of technology in Internships has followed a similar trajectory to the technology adoption across workplaces generally, with Internships initially exhibiting features aligned with the technology-supported, technology-facilitated, and technology-blended categories of Schuster and Glavas (2017). The more recent transition to technology-based Internships, reflects a new level of technology use because the activities and learning outcomes of Internships, as well as the mechanisms to support and evaluate participant learning outcomes, are deliberately and entirely reliant on the application of technology (Jeske and Axtell, 2018).

The potential learning outcomes of technology-based Internships are similar to those of traditional Internships (Bayerlein et al., 2021b; Zegwaard and Rowe, 2019), because in both instances, learning outcomes are reliant on the meaningful, and structured, integration of a participant learning experience within a workplace context (Bayerlein and Jeske, 2018b). While the learning activities and support structures that are available in traditional and technology-based Internships differ, the cognitive, skill-based and affective psychological development outcomes (Kraiger et al., 1993) that are attainable in both settings are largely comparable (Bayerlein and Jeske, 2018b). In addition, technology-based Internships have the potential to reduce the accessibility (Bayerlein et al., 2021b, Kraft et al., 2019, Bayerlein and Jeske, 2018a; Kay et al., 2019) and resource allocation barriers (Irwin et al., 2012) of Internships considerably.

Expansion of internships to Work-Integrated Learning

The nature of contemporary workplaces is changing rapidly and permanently. While the adoption of technology across virtually all workplaces is one important source of change, the increasing generalist nature of workplaces (Hite and McDonald, 2020) represents another macro trend. The changing nature of workplaces, in turn, requires workplace participants to develop skills and knowledge that are substantially different from those required in traditional workplaces (Bayerlein et al., 2021a).

Changes between traditional workplaces and emerging workplaces are particularly pronounced in areas such as personal development, knowledge acquisition, and the development of workplace cultures (Bayerlein et al., 2021a). The changing nature of workplaces impacts these skill and knowledge areas because their development and application is largely socially constructed (Hora et al., 2018, 2019). In traditional Internships, the integration of participants into the social environment that surrounds their work and learning experience (Konradt and Schmook, 1999) enables the attainment of socially constructed learning outcomes that are specific to the placement setting. Technology-based Internships are principally able to achieve comparable outcomes, because both Internship modes are conceptually based on the notion of employability (Gardner and Liu, 1997), rather than the notion of workplace competencies (Prikshat et al., 2019).

Work-Integrated Learning represents a wide field of workplace-based learning activities that encompasses Internships as well as a range of well-established and emerging learning activities. Work-Integrated Learning is based on well-established frameworks and concepts (Kay et al., 2019), but the delivery mechanisms used to achieve learning outcomes for students have changed substantially in recent years. Changes to Work-Integrated Learning are the result of the changing nature of work, as well as the growing demand for learning activities that align participant experiences with the workplace competencies required in emerging workplaces (Bayerlein et al., 2021a,b; Dean and Campbell, 2020). Of particular note in this area are technology-based Work-Integrated Learning activities that enable a focus on the development of transferable workplace-competencies. Examples of such activities include simulated Internship experiences (for example, see: Bayerlein, 2015) or digital service learning (Perkiss et al., 2020), which combines the traditional notion of service learning as a learning activity that provides a service to a particular community, or society at large, with a technology-based, location un-bounded, learning environment. New forms of Work-Integrated Learning, such as digital service learning, are aligned with many facets of traditional Internships because participants develop a critical understanding of the context and substance of a problem, and develop professional judgment and reflective capabilities as part of the activity (Perkiss et al., 2019; Wersun et al., 2019; Dean et al., 2018). Simulated Internships are more closely aligned with traditional, as well as technology-based, Internship arrangement because the conceptual idea of a work-placement is retained (Bayerlein, 2015), and some of the challenges associated with traditional internships are also present in technology-based internship arrangements (Bayerlein, 2020). However, the ability to focus participant learning experiences and outcomes on workplace competencies that are critically important for industries or professions generally, rather than the learning activities and outcomes that can be attained within a particular placement, provides additional flexibility for education providers and participants.

Recommendations for future internship developments

Changes in the structure and technology use in virtually all workplaces have gathered momentum during the COVID-19 pandemic (Bayerlein et al., 2021a). While the challenges associated with the resulting workplace changes and career disruptions are not new (for example, see: Mishra and McDonald, 2017), ongoing structural changes to workplaces increase the importance of transferable generalist skills and competencies for existing and future workforce participants (Bayerlein et al., 2021a). In practice, this means that future workplaces are likely to become more generalist in nature, and workplace participants are likely to require a greater resilience to change (Hite and McDonald, 2020). Given that education providers are exposed to long-standing pressures from governments and employer organizations to ensure graduates meet a growing number of workplace-competence and employability criteria,

changes to education programs are required. This work has positioned Internships, and Work-Integrated Learning activities more generally, as one educational experience that is well suited to address these needs (also, see: [Bayerlein et al., 2021b](#)).

Internships, and Work-Integrated Learning more generally, allow education providers to extend existing curricula to prepare graduates for emerging workplaces. However, in order to achieve this outcome Internships are required to adapt to changes in (emerging) workplace environments to ensure that the skill and knowledge development of participants is aligned with workplace macro-trends. In addition, the systematic application of learning design (for example, see: [Wiggins and McTighe, 1998](#)) and assurance of learning (for example, see: [Biggs, 1996](#); [Vitale et al., 2020](#)) principles in Internships is likely to necessitate a re-development of traditional Internship structures, processes and learning activities to ensure alignment between Internship learning activities, Internship learning outcome and the workplace competencies expected by employers and the society at large.

Three macro-trends that are likely to impact future Internship arrangements substantially are (1) the proliferation of technology-based ways of working ([Bayerlein et al., 2021a](#)), (2) the opportunities for Internships participants and Internship providers that arise from a reduction of Internship participation barriers ([Kay et al., 2019](#); [Kraft et al., 2019](#)), and (3) the growing focus on transferable, and increasingly generalist, workplace-readiness competencies over employer specific, or employment opportunity specific, employability skills ([Bayerlein et al., 2021a](#)).

Preparing participants for technology-based workplaces

The increasing uptake of technology-based distributed work arrangements across workplaces world-wide ([Bayerlein et al., 2021a](#)), as well as the associated growth in technology-based Internship arrangements, places additional burdens and responsibilities on employees as well as Internship participants ([Anderson, 2020](#)). Within Internships, additional challenges for participants arise because technology-based distributed work arrangements require them to work independently ([Xu and Tracey, 2014](#)), and to develop personality traits that are critically important for successful online learning (for example, see: [Xiao, 2012](#)). Education providers must recognize these challenges in their overall curriculum structures, and ensure that participants are adequately prepared for technology-based internship environments, as well as technology-based workplaces in general.

The implementation of technology-based distributed Internships is likely to require collaborative efforts from education providers and placement providers. Such collaborations are necessary because providers of education may find the development and implementation of technology-based Internship solutions challenging (for example, see: [Anderson, 2020](#); [Dean and Campbell, 2020](#)). Challenges arise due to the limited experience of participants and education provider staff in organizing and implementing technology-based learning activities. Many placement providers, on the other hand, developed extensive technology-based remote-working processes and support structures during the COVID-19 pandemic ([De et al., 2020](#)), and the knowledge gained from these experiences is now available to support technology-based Internships. Providers of education are able to contribute their expertise in the scaffolded development of distance learning skills to technology-based and distributed Internships, and technology-based Internships are likely to be most successful for participants if education providers and placement providers collaborate in the design and implementation of the experience.

The identification of clear learning outcomes, support structures and learning outcome assurance activities is critically important for all Internships ([Jackson, 2015](#)). In technology-based Internships, the identification of explicit mechanisms through which participant learning outcomes are achieved, and the structures through which placement and education providers support participants in attaining the identified outcomes, is even more critical because traditional Internship support structures are largely unavailable in technology-based Internships ([Bayerlein and Jeske, 2018b](#)).

One potential challenge for technology-based Internships, and technology-based Work-Integrated Learning activities generally, is the historically low level of acceptance for such learning experiences by employers and providers of education ([Bayerlein and Jeske, 2018b](#)). While empirical post-pandemic evidence of the acceptance of technology-based workplace learning activities is not yet available, it appears likely that the COVID-19 pandemic may have changed the perceptions of employers, and education providers considerably.

Reducing participation barriers

Future Internship developments should explicitly consider structures and delivery modes that improve participation opportunities for societal groups that have previously found it difficult to engage with, or participate in, Internship activities ([Kay et al., 2019](#)). A continued focus on traditional—location-bounded—internships could inadvertently limit participation opportunities of societal groups with additional needs ([Bayerlein and Jeske, 2018a](#); [Kraft et al., 2019](#)), family or care-giving commitments ([Bayerlein and Jeske, 2018a](#)), or other factors preventing attendance at a specific internship location ([Moore et al., 2015](#)).

Two internship developments that are likely to improve participation opportunities are technology-based Internships, as well as the transition of workplace-based learning away from narrowly defined Internships, and toward a wider range of Work-Integrated Learning activities. While neither of these developments were explicitly designed to reduce participation barriers for specific participants or occupations, their application should, in principle, lead to positive impacts for participants and employer organization alike (for example, see: [Kraft et al., 2019](#)). Despite early evidence about the potential impact of these developments for participants and internship providers, empirical assessments of this impact are still largely absent from the literature.

Focusing experiences on workplace-readiness competencies

Internships, and Work-Integrated Learning activities generally, should enable participants to develop the skills and knowledge that are necessary to meet the demands of increasingly complex and constantly changing workplaces. To meet these demands, it is recommended that future workplace based learning activities should aim to develop participants integrated workplace-readiness competencies (Prikshtat et al., 2019), rather than more context specific employability requirements (Gardner and Liu, 1997).

Workplace competencies focused Internships, and Work Integrated Learning activities more generally, are likely to meet the demand for increasingly diverse (Kay et al., 2019) and generalist (Hite and McDonald, 2020) learning outcomes. However, a focus on activities that preference workplace-readiness competencies over employability may also result in a further decrease in the networking opportunities and employment prospects that are important features of traditional internships. Given the growing overall demand for Internships, it appears likely that Internships with different learning outcome focus areas will emerge. While such a diversification of learning opportunities has clear benefits for all stakeholders, the success of individual workplace-based learning activities is likely to depend on the clear identification and communication of learning outcomes available in each Internship (Jackson, 2015). Without clearly identified and communicated learning outcomes, increasing Internship diversity is unlikely to be beneficial for participants or internship providers.

References

- Alexander, B., Ashford-Rowes, K., Barajas-Murphy, N., Dobbin, G., Knott, J., McCormack, M., Pomerantz, J., Seilhamer, R., Weber, N., 2019. EDUCAUSE Horizons Report: 2019 Higher Education Edition. EDUCAUSE, Louisville, CO.
- Anderson, V., 2020. A digital pedagogy pivot: re-thinking higher education practice from an HRD perspective. *Hum. Resour. Dev. Int.* 23, 452–467.
- Arum, A., Roksa, J., 2011. *Academically Adrift: Limited Learning on College Campuses*. The University of Chicago Press, Chicago IL.
- Bayerlein, L., Jeske, D., 2018a. The potential of computer-mediated internships for higher education. *Int. J. Educ. Manag.* 32, 526–537.
- Bayerlein, L., Jeske, D., 2018b. Student learning opportunities in traditional and computer-mediated internships. *Educ. Train.* 60, 27–38.
- Bayerlein, L., Hora, M.T., Dean, B.A., Perkiss, S., 2021a. Developing skills in higher education for post-pandemic work. *Lab. Ind.* 418–429.
- Bayerlein, L., Perkiss, S., Dean, B., Jeske, D., 2021b. Technology platforms for WIL. In: Ferns, S., Rowe, A., Zegwaard, K. (Eds.), *Advances in Research, Theory and Practice in Work-Integrated Learning: Enhancing Employability for a Sustainable Future*. Routledge, New York.
- Bayerlein, L., 2015. Curriculum innovation in undergraduate accounting degree programmes through “virtual internships”. *Educ. Train.* 57, 673–684.
- Bayerlein, L., 2019. Internships and co-ops. In: Amey, M.J., David, M.E. (Eds.), *The SAGE Encyclopedia of Higher Education*. Sage, Thousand Oaks.
- Bayerlein, L., 2020. The Impact of Prior Work-Experience on Student Learning in Simulated Internships, vol. 17. *Journal of University Teaching & Learning Practice*.
- Biggs, J., 1996. Enhancing teaching through constructive alignment. *High Educ.* 32, 347–364.
- Candy, P.C., Crebert, R.G., 1991. Ivory tower to concrete jungle: the difficult transition from the academy to the workplace as learning environments. *J. High Educ.* 62, 570–592.
- Carroll, N., Conboy, K., 2020. Normalising the “new normal”: changing tech-driven work practices under pandemic time pressure. *Int. J. Inf. Manag.* 55.
- Coll, R., Zegwaard, K., 2006. Perceptions of desirable graduate competencies for science and technology new graduates. *Res. Sci. Technol. Educ.* 24, 29–58.
- De, R., Pandey, N., Pal, A., 2020. Impact of digital surge during Covid-19 pandemic: a viewpoint on research and practice. *Int. J. Inf. Manag.* 55.
- Dean, B.A., Campbell, M., 2020. Reshaping Work-Integrated Learning in a post-COVID-19 world of work. *Int. J. Work Integr. Learn.* 21, 356–364.
- Dean, B., Gibbons, B., Perkiss, S., 2018. An experiential learning activity for integrating the Sustainable Development Goals into business education. *Soc. Business* 8, 387–409.
- Fleming, D.L., 2008. Building bridges to connect the disconnects: an analysis of business program design processes. *Am. J. Bus. Educ.* 1, 21–49.
- Gardner, P.D., Liu, W.-Y., 1997. Prepared to perform? Employers rate workforce readiness of new grads. *J. Career Plann. Employ.* 57, 32–56.
- Guile, D., Griffiths, T., 2001. Learning through work experience. *J. Educ. Work* 14, 113–131.
- Hergert, M., 2009. Student perceptions of the value of internships in business education. *Am. J. Bus. Educ.* 2, 9–14.
- Hite, L.M., McDonald, K.S., 2020. Careers after COVID-19 challenges and changes. *Hum. Resour. Dev. Int.* 23, 427–437.
- Hora, M.T., Benbow, R.J., Smolarek, B.B., 2018. Re-thinking soft skills and student employability: a new paradigm for undergraduate education. *Change* 50, 30–37.
- Hora, M.T., Smolarek, B.B., Martin, K.N., 2019. Exploring the situated and cultural aspects of communication in the professions: implications for teaching, student employability, and equity in higher education. *Am. Educ. Res. J.* 56, 2221–2261.
- Irwin, P., Ellis, A., Graham, I., 2012. Virtual WIL: A Collaborative Approach to Work Integrated Learning Using a Virtual World. Australian Collaborative Education Network National Conference.
- Jackson, D., Bridgstock, R., 2021. What actually works to enhance graduate employability? The relative value of curricular, co-curricular, and extra-curricular learning and paid work. *High Educ.* 81, 723–739.
- Jackson, D., Sibson, R., Riebe, L., 2014. Undergraduate perceptions of the development of team-working skills. *Educ. Train.* 56, 7–20.
- Jackson, D., Rowbottom, D., Ferns, S., McLaren, D., 2017. Employer understanding of Work-Integrated Learning and the challenges of engaging in work placement opportunities. *Stud. Cont. Educ.* 39, 35–51.
- Jackson, D., 2015. Employability skill development in Work-Integrated Learning: barriers and best practice. *Stud. High Educ.* 40, 350–367.
- Jeske, D., Axtell, C.M., 2014. E-internships: prevalence, characteristics and role of student perspectives. *Internet Res.* 24, 457–473.
- Jeske, D., Axtell, C.M., 2016. How to run successful e-internships: a case for organizational learning. *Dev. Learn. Org. Int. J.* 30, 18–21.
- Jeske, D., Axtell, C.M., 2018. The nature of relationships in e-internships: a matter of psychological contract, communication and relational investment. *J. Work Organ. Psychol.* 34, 113–121.
- Kavanagh, M.H., Drennan, L., 2008. What skills and attributes does an accounting graduate need? Evidence from student perceptions and employer perceptions. *Account. Finance* 48, 279–300.
- Kay, J., Ferns, S., Russell, L., Smith, J., Winchester-Seeto, T., 2019. The emerging future: innovative models of Work-Integrated Learning. *Int. J. Work Integr. Learn.* 20, 401–413.
- Knouse, S.B., Fontenot, G., 2008. Benefits of the business college internship: a research review. *J. Employ. Counsel.* 45, 61–66.
- Konradt, U., Schmook, R., 1999. Telework: stress and strain in a longitudinal study. *Z. Arbeits Organisationspsychol.* 43, 143–150.
- Kraft, C., Jeske, D., Bayerlein, L., 2019. Seeking diversity? Consider virtual internships. *Strateg. HR Rev.* 18, 133–137.
- Kraiger, K., Ford, J., Salas, E., 1993. Application of cognitive, skill-based and affective theories of learning outcomes to new methods of training evaluation. *J. Appl. Psychol.* 78, 311–328.
- Maertz, C.P., Stoeberl, P.A., Marks, J., 2014. Building successful internships: lessons from the research for interns, schools, and employers. *Career Dev. Int.* 19, 123–142.
- Mishra, P., McDonald, K., 2017. Career resilience: an integrated review of the empirical literature. *Hum. Resour. Dev. Rev.* 16, 207–234.
- Moore, K., Ferns, S., Peach, D., 2015. The Australian collaborative education network student scholarship for Work-Integrated Learning. *Asia Pac. J. Cooper. Educ.* 16, 241–254.

- Orrell, J., 2011. Good Practice Report: Work-Integrated Learning. Australian Learning & Teaching Council, Strawberry Hills.
- Perkiss, S., Dean, B.A., Alejandra Gonzalez-Perez, M., Heithaus, T., Acosta, P., Mills, R., Gibbons, B., Anastasiadis, S., Bayerlein, L., Jun, H., Mesicek, R., Wersun, A., 2019. Learning experientially for corporate contribution to global sustainable development: International Applications of the WikiRate Project. In: Gonzalez-Perez, M.A., Lynden, K., Taras, V. (Eds.), *The Palgrave Handbook of Learning and Teaching International Business and Management*, first ed. Palgrave Macmillan, London.
- Perkiss, S., Anastasiadis, S., Bayerlein, L., Dean, B., Jun, H., Pilar, A., Gonzales-Perez, M.A., Wersun, A., Gibbons, B., 2020. Advancing sustainability education in business studies through digital service learning. *Am. Bus. Rev.* 23, 276–292.
- Prikshat, V., Nankervis, A., Burgess, J., Dhakal, S., 2018. Conceptualising graduate work-readiness: theories, concepts and implications for practice and research. In: Dhakal, S., Prikshat, V., Nankervis, A., Burgess, J. (Eds.), *The Transition From Graduation to Work: Challenges and Strategies in the Twenty-First Century Asia Pacific and Beyond*. Springer, Singapore.
- Prikshat, V., Kumar, S., Nankervis, A., 2019. Work-readiness integrated competence model: conceptualisation and scale development. *Educ. Train.* 61, 568–589.
- Rosenberg, S., Heimler, R., Morote, E.-S., 2012. Basic employability skills: a triangular design approach. *Educ. Train.* 54, 7–20.
- Rothman, M., Sisman, R., 2016. Internship impact on career consideration among business students. *Educ. Train.* 58, 1003–1013.
- Schuster, L., Glavas, C., 2017. Exploring the dimensions of electronic Work Integrated Learning (eWIL). *Educ. Res. Rev.* 21, 55–66.
- Silva, P., Lopes, B., Costa, M., Melo, A.I., Dias, G.P., Brito, E., Seabra, D., 2018. The million-dollar question: can internships boost employment? *Stud. High Educ.* 43, 2–21.
- Smith, C., Worsfold, K., 2015. Unpacking the learning-work nexus: “priming” as lever for high-quality learning outcomes in work-integrated learning curricula. *Stud. High Educ.* 40, 22–42.
- Sneider, K., Singhal, S., 2020. *Beyond Coronavirus: The Path to the Next Normal*. McKinsey & Company, New York.
- Van Dorp, K.J., 2008. A premier European platform for clearing e-internships. *Br. J. Educ. Technol.* 39, 175–176.
- Vitale, C., Bowyer, D., Bayerlein, L., 2020. Developing and presenting a framework for meeting industry, student and education expectations in university degrees. *E J. Business Educ. Sch. Teach.* 14, 59–67.
- Wersun, A., Dean, B., Mills, R., Perkiss, S., Acosta, P., Anastasiadis, S., Gibbons, B., Gonzalez-Perez, M.A., Heithaus, T., Jun, H., Mesicek, R., Bayerlein, L., 2019. An exploration of student learning for sustainability through the WikiRate student engagement project. *Int. J. Manag. Educ.* 17, 1–12.
- Wiggins, G., Mctighe, J.H., 1998. *Understanding by Design*. Merrill Prentice Hall, Columbus, Ohio.
- Wilton, M., 2012. The impact of work placements on skills development and labour market outcomes for business and management graduates. *Stud. High Educ.* 37, 603–620.
- Xiao, J., 2012. Successful and unsuccessful distance language learners: an “affective” perspective. *Open Learn.* 27, 121–136.
- Xu, H., Tracey, T.J.G., 2014. The role of ambiguity tolerance in career decision making. *J. Vocat. Behav.* 85, 18–26.
- Yawson, R., 2020. Strategic flexibility analysis of HRD research and practice post COVID-19 pandemic. *Hum. Resour. Dev. Int.* 23, 406–417.
- Zegwaard, K.E., Rowe, A.D., 2019. Research-informed curriculum and advancing innovative practices in work-integrated learning. *Int. J. Work Integr. Learn.* 20, 323–334.

Career pathways

Debra D. Bragg, Bragg & Associates, Inc., Emeritus Faculty at the University of Illinois at Urbana-Champaign, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Career pathways in the international context	469
Career pathways in the United States	470
Career pathways models and approaches	471
Federal investment in career pathways in the US	472
What works clearinghouse meta-analysis on career pathways	473
Research needed on career pathways	474
References	474

Career pathways in the international context

The focus of education on preparing workers to enter employment and progress in careers is a global concern. Rooted in long histories of vocational education and training (VET) that vary from country to country, systems and programs of career pathways through education and to employment are changing worldwide. Virtually all VET prepares graduates for employment but how this is done varies widely in the emphasis on vocational knowledge and skills relative to general education, the methods of delivery of VET systems and programs, and the emphasis placed on preparing youths to enter employment versus adults already in the workforce. VET systems also vary in their intentionality to meet the education and employment needs of populations with low participation rates in postsecondary education, further marginalizing them from professional employment that garners higher income and social mobility (Bathmaker, 2015).

Typically under the auspices of VET, the formulation of career pathways systems and programs also depends on past and present governmental policies and labor market economies. Wheelahan and Moodie (2017) assessed the patterns of relationships between education and employment in different national contexts to understand changes made to improve VET that have implications for how career pathway systems and programs are evolving worldwide. Their analysis shows Germany and the Scandinavian countries have market economies that coordinate VET through national and regional governmental policies, as well as business and labor that also exert strong influence. This approach to VET results in preparation for employment that follows distinct postsecondary education tracks that matches graduates to job vacancies. Because these countries have relatively stable labor markets, employers and employees are incentivized to invest in VET, including playing strong roles in identifying the qualifications for occupations they need to fill. As a result, these coordinated systems of VET give reliable signals to employers of the competencies graduates need to hold to enter and advance in employment. Similarly, graduates are assured that their investment in VET will lead to predictable employment, and sometimes also career progression.

In contrast, Anglophone countries with liberal market economies rely more heavily on labor markets and employers to match graduates and employment, without direct connections to or influence on education. Ironically, Wheelahan and Moodie (2017) argue VET in these economies tends to place higher priority on occupational-specific competencies or skills than northern European countries because the limited engagement of employers in these countries weakens their understanding that what is taught as relevant to the jobs they need to fill, and further reduces their commitment to hiring graduates. As a result, the cost of VET in these countries is born heavily by students rather than business and industry, except to the extent employers provide varying levels of training for entry-level workers.

One goal of career pathways systems and programs is to close this gap between education and employment by incentivizing governments, educational institutions, employers, and labor to work together to improve the alignment of academic and occupational knowledge and skills with employment. A common outcome of these efforts is a set of qualifications, including credentials, used to articulate more clearly what career pathways are attempting to do to prepare graduates for work. Though well intentioned, Wheelahan and Moodie point out tensions exist in countries with liberal market economies that adopt competency-based qualifications for specific jobs concerning their emphasis on general and academic education. Some educational experts observe the blurring of academic and vocational education to de-emphasize tracking, seen in Anglophone countries such as Australia, Canada, and the United Kingdom, emphasizes a more democratic form of education while other experts suggest the ambiguity in education characterizing these nations contributes to inequitable outcomes, further marginalizing racially minoritized and low-income populations. Contributing to this latter perspective is research suggesting students in Anglophone countries are less certain of the value of the qualifications than in countries that coordinate VET and employers and also that employers in these countries are less likely to see VET as a reliable signal that graduates are prepared for employment (Moodie, 2008).

Looking more deeply at the development of career pathways in Anglophone countries, we see the emergence of higher vocational education (HIVE). Policy and programmatic changes to VET in these countries include growing recognition that the knowledge and skills needed to work in the global economy require knowledge and skills offered by higher education. These arguments for career

pathways systems and programs also recognize the importance of providing education beyond compulsory age for underserved students who are far less likely to attend college and secure well-paying employment as a result of their education than students from more privileged backgrounds (Webb et al., 2017). Although, while VET systems in northern European countries are creating more options for students to participate in post-compulsory education (McKay et al., 2021) these systems continue to perpetuate differences in upward social mobility (Protsch and Solga, 2016).

The question of alignment between VET and employment has been especially active in Anglophile countries over the last decade or more, due to the loose coordination of education and employment. For example, that Australia and Canada provide limited alignment between VET, employment in specific industry sectors, and the labor market has been the subject of research by Wheelahan and Moodie (2017, p. 11). This study concluded graduates have diminished potential to secure well-paying jobs “when ‘skills mismatches’ between credentials and occupations are widespread.” Also, because governments of Anglophile countries tend not to intervene in liberal market economies, VET operates less effectively to match graduates to employment than in northern European countries. Due to longstanding policies and deeply rooted structures that produce mixed results, various reforms intended to improve VET, including career pathways systems and programs, are being pursued in almost all countries. Common features pervade these reforms according to Keating et al. (2002) who studied VET systems in nine countries. These similarities include:

- VET reform is the focus of more national governments.
- VET and general education are becoming increasingly more integrated.
- VET departments are becoming more integrated to broaden employment options.
- VET programs are broadening curriculum to include “soft” skills, core competencies, and related knowledge.
- VET is under pressure to cut costs.
- VET policy and funding is increasingly driven by industry and labor market demand than by governments.
- VET tends to be experimenting with developing education and training markets.
- VET has not been seen as sufficiently client-focused.
- VET has developed and continues to exist as an alternative to higher-status academic pathways.

Career pathways in the United States

Career pathways in the United States integrate secondary education and/or adult basic education (ABE) with workforce training and postsecondary education to help transition youths and adults into living wage employment and careers (U.S. Department of Education, 2015). Career pathway students prepare to enter and advance by mastering competencies aligned with industry-recognized and -college credentials, including associates and baccalaureate degrees (Bragg et al., 2019). By better integrating career pathway systems and programs more closely with the labor market, occupational preparation is expected to be more efficient, effective and equitable. Also, by sequencing competencies and credentials, student outcomes are expected to improve for diverse learners from the time they enter employment through their working lives.

Whereas career pathways are theorized to benefit all students in some way, this approach is thought most transformative for student populations who have been underserved by postsecondary education. By easing the transition into college and better aligning postsecondary studies with the labor market, career pathways are expected to enhance social mobility for students who have experienced barriers to postsecondary education participation. In the US where racially minoritized groups have been tracked away from course taking required to access and successfully navigate college attendance (Domina et al., 2019), career pathway systems and programs are expected to help improve student outcomes. Racially minoritized and other underserved populations are prepared for academic as well as vocational requirements to succeed in postsecondary education.

In the US, career pathways systems and programs prepare youths rising from high school to college and adults marginalized from postsecondary education and good jobs to enter living wage employment and advance in careers (Durham and Bragg, 2019; Sama and Strawn, 2018). By designing career pathways for youths and adults, this comprehensive approach to occupational preparation is thought to optimize opportunities for individuals to secure the qualifications needed for well-paying employment (Alssid et al., 2002; Hedge and Carter, 2020). This argument for career pathways is rooted in a belief that liberal economies such as the US and Anglophile countries with historically weak connections between education and the labor market need to strengthen these connections to assist more graduates to secure good jobs (Wheelahan and Moodie, 2017). Career pathways are especially keen on addressing this problem for students who have been underserved by postsecondary education. Students who are racially minoritized, first-generation college, immigrants and English Language Learners (ELL), individuals already working in low-wage and essential jobs, and other groups with limited resources to attend college are a primary target audience of career pathways.

Positioned as the gateway to postsecondary education in the US, community colleges are often designated as a preferred provider of career pathways because of their open-access mission, historic commitment to occupational preparation, and strengthening relationships with public workforce training (Eyster, 2020). Situated between high schools and universities, community colleges offer 2 years of full-time postsecondary study to prepare graduates for credentials leading up to the associates degree and increasingly baccalaureate degrees in occupational fields requiring these credentials. Community colleges also prepare students to transfer to

universities where they complete two more years of study to secure baccalaureate degrees. The option to exit postsecondary education to obtain employment is possible with any of these options but most optimal when students earn credentials recognized by employers as well aligned to jobs that need to be filled (Carnevale et al., 2020). Besides working with employers and public workforce training, community colleges may also work with K-12 (kindergarten through grade 12) education, community-based organizations and other social services agencies to design and deliver career pathways addressing local and regional economic and social needs (Jacobs and Worth, 2019).

The heuristic of ladders or lattices that enable students to advance from K-12 education or ABE to postsecondary education is a common feature of career pathways (Herhusky and Ludwin, 2021). By improving the alignment of curriculum and instruction leading to credentials (certificates, licensures and degrees), graduates are expected to secure knowledge and skills needed for well-paying jobs that lead to career advancement (Sarna and Strawn, 2018). Careful mapping of knowledge and skill requirements to credentials enables graduates to enter and advance in industry sectors such as healthcare, manufacturing, and information technology. Wheelahan and Moodie (2017) observe qualifications associated with career pathways operate as signals or screens to employers who seek to hire qualified workers. Signals are strongest in regulated occupations such as nursing, engineering, and teaching, and weakest in unregulated occupations associated with management, retail and the service sector.

Considerable debate appears in literature in the US and internationally about whether the credentials associated with career pathways actually lead to well-paying jobs. Advocates argue credentials offered through career pathways, including short-term training offered at relatively low-cost, expedite employment that is needed to secure income needed to continue education and advance in careers. However, more tangible evidence is sought to understand how credentials associated with career pathways benefit underserved populations any better than past efforts to enhance social mobility for racially minoritized and low-income students (Ositelu et al., 2021).

Career pathways models and approaches

Federal and state governments in the US have increased the public investment in career pathways over the last two decades but large-scale investments have also been made by states and philanthropic organizations (Bragg, 2015). This mix of public and private funding has played a substantial role in scaling up career pathway models across the US. These models are voluminous, so I have chosen to discuss two models targeting youths and two focusing on adults that are widely known and implemented. Targeting youths are the Pathways to Prosperity and New Skills for Youth (NSYF) models, and targeting adults are the Integrated Basic Education Skills Training (I-BEST) and Accelerating Opportunity (AO) models.

Pathways to Prosperity Network (PPN) was championed by Harvard University faculty and collaborators who theorized career pathways are needed to fill the middle-skills gap and address inequalities in education and employment in the US (Schwartz, 2016). This effort has been funded by numerous philanthropies to support leaders in 10 states (Arizona, California, Delaware, Hawaii, Illinois, Minnesota, New York, Ohio, Tennessee, and Texas) to establish career pathways for youths in grades 9–14. High-skill, high-demand (HSHD) sectors such as healthcare, manufacturing, and information technology are prioritized and expected to provide youths with competencies and credentials to secure well-paying employment without closing off postsecondary education options, which has been a limitation of some forms of vocational education in the US. A key goal of PPN is to integrate college and career options and create statewide career pathway systems that reach at least 40% of grade 9–14 students. Recognizing challenges to expanding career pathways to large numbers of US youths, this model advocates for enhanced employer engagement. Intentionally replicating elements of Switzerland's VET system, PPN recommends a common set of standards to guide the development of career pathways similar to the Swiss VET model. Community colleges are seen as the most logical postsecondary partner to anchor career pathways for youths, and states adopting PPN model are encouraged to reform VET [called career and technical education (CTE)] in the US into career pathways systems and programs.

New Skills for Youth (NSYF) is a multi-million investment by JPMorgan Chase and Company in career pathways for youths in 10 states (Delaware, Kentucky, Louisiana, Massachusetts, Nevada, Ohio, Oklahoma, Rhode Island, Tennessee, and Wisconsin). [Note, three of these states are also engaged in PPN (Delaware, Ohio, and Tennessee).] Over the last 6 years these states have built on existing CTE programs, including utilizing federal and state funding to enhance youth transition from secondary to postsecondary education that leads to employment in healthcare and manufacturing (Staklis et al., 2020). Recognizing widening inequalities, NSYF attempted to improve programs for underserved youths through improved partnerships with employers, enhanced postsecondary and workforce connections, and strengthened accountability systems. However, the completion rates of students participating in career pathways created through NSYF rose only modestly, from 12% in 2015–16 to 18% in 2018–19. Undeterred from its long-term investment, JPMorgan Chase recently announced it would continue to support career pathways and disseminate lessons learned to a larger audience of K-12 education (JPMorgan Chase and Co, 2020).

Turning our attention to adults, a plethora of career pathway models and approaches have evolved over the last two decades. The Integrated Basic Education and Skills Training (I-BEST) model is one of the most well-known and foundational to other adult-facing career pathways models and approaches in the US (Eyster, 2020). Developed by the Washington State Board of Community and Technical Colleges (SBCTC), with funding from the state and the Bill & Melinda Gates Foundation, I-BEST is a structured model for integrating CTE and contextualized basic skills primarily targeting adults who have not completed a formal high school diploma. I-BEST relies on the co-teaching of technical content, literacy and other content by pairing CTE and basic skills instructors. Supplemental instruction and comprehensive student support help adults secure the GED and enter postsecondary to secure credentials

and eventually employment in industry sectors offering well-paying jobs. Early evaluation of I-BEST from 2005–2006 to 2007–2008 found student participation had a positive relationship to earning college credit and a certificate or degree and achieving point gains on basic skills exams (Wachen et al., 2012). Later analysis showed the intention to treat students with I-BEST revealed a positive causal relationship between I-BEST and many of the same student outcomes evaluated by Wachen et al. (Zeidenberg et al., 2010).

Finally, a national initiative called Accelerating Opportunity (AO) evolved from the I-BEST model to provide adults, as well as underserved youths, with career pathways offered by community colleges. Taking lessons from the Breaking Through initiative as well as I-BEST (Anderson et al., 2017), AO allowed students scoring in the 6th-through 12th-grade functioning levels of National Reporting System to enroll in CTE courses concurrently with high school equivalency completion programs, including the GED. Career pathways offered courses and stackable, industry-recognized credentials within roughly 12 college credit hours. Team teaching patterned off the I-BEST model, amounting to at least 25% of college classes, as well as contextualized and accelerated instruction, support services, and connections with employers and workforce agencies are integral to the AO model. Rigorous research designs to track student outcomes through 12 quarters of initial enrollment found AO had a positive impact on the number of college-awarded credentials earned by almost all students (Anderson et al., 2017). In most cases, AO students earned more credentials than comparison students while taking fewer college credits, possibly indicating more efficient course-taking through accelerated instruction. However, other student outcomes varied by the states, reflecting differences in student populations, career pathway interventions, and regional economies where the programs were offered.

Federal investment in career pathways in the US

Federal investment in career pathways systems and programs has grown in the US over the last two decades (Mortrude, 2018). Twelve federal agencies signed on to a common definition of career pathways by 2012, moving federal leadership toward a shared commitment to integrated postsecondary education and workforce training for youths and adults. Federal agencies also initiated funding for career pathways systems and programs, as well as research on career pathways (United States Department of Education, Office of Vocational and Adult Education, 2012). These projects included large-scale evaluations funded by the USDOL, USDOEd, and USHHS. The USDOL and USDOEd also worked together to develop tools and information about career pathways, emphasizing cross-agency partnerships, employer engagement in HSHD industry sectors, funding and capacity-building resources; aligned public policies and delivery systems; and evaluation of system change, program and student outcomes (Bragg, 2015; Eyster, 2020).

These leadership and funding efforts by the federal government also served as models for state and local investments through ongoing federal sources such as the Workforce Investment and Opportunity Act (WIOA) and the Strengthening of Career & Technical Education for 21st Century Act. The braiding of federal, state and other funding was recommended to develop and implement career pathways and recruit students who could benefit from participation in these programs. Along with other key components of career pathways, the notion of bridging students from one system to another, for example, ABE to community college, is an important feature of career pathways funded through federal and state policy.

Following this trajectory of government support, one federal law shaped local implementation of career pathways on a particularly major way. Created during President Obama's administration, the Trade Adjustment Assistance Community College and Career Training (TAACCCT) grant program awarded grants to higher education institutions, primarily community colleges, to help address the severe unemployment and economic ramifications of the Great Recession. TAACCCT is arguably the most expansive federal law passed thus far supporting community colleges to reform and anchor career pathways systems and programs at the postsecondary education level (Bragg et al., 2019; McKay and Van Noy, 2021).

Between 2011 and 2018, the DOL funded 256 grants totaling nearly \$2 billion through four rounds of competition for the TAACCCT grant program. These grants were awarded to 630 (of approximately 1000) community and technical colleges across all 50 states, with community colleges making up 85% of all postsecondary institutions funded under TAACCCT (Cohen et al., 2017). To receive TAACCCT funding grantees were expected to create career pathways that would "provide workers with the education and skills to succeed in high-wage, high-skill occupations" (USDOL, 2016, p. 3). Drawing on lessons learned in earlier career pathways initiatives, including directly referencing I-BEST and AO, the TAACCCT grant program provided resources on a far greater scale to support community colleges engaged in this work than had ever occurred in the past.

Ultimately, TAACCCT developed or redesigned about 2700 programs that enrolled one-half million students who earned more than 350,000 credentials. Manufacturing and healthcare were the two sectors with the most grants, followed by energy, information technology, transportation and logistics, green technology, and agriculture (USDOL, 2016). Numerous research and evaluation studies showed the grants reached the primary target of adult students, with 31 being the average age of all grant participants (Durham et al., 2017). However, TAACCCT grant participants were not as diverse as the national population of community college students, with a higher proportion of men than women participants (60% compared to 40%, respectively). Neither were as many grant participants identified with racially minoritized groups than community college students generally (49% versus 59%, respectively). Also, the predominant enrollment pattern changed over the course of the grant, with the majority students enrolling full-time during the first two rounds beginning October 1, 2011 and 2012, but shifting to part-time during the next 2 years beginning October 1, 2013, and extending to October 1, 2014. This change shift in enrollment paralleled the nation's improving economy when unemployment rates dropped and community college enrollments began dropping precipitously, a trend that has not changed to the present time.

Required by TAACCCT, 215 third-party evaluations were completed in rounds two through four of the grant program to assess implementation and participant educational and employment outcomes and impacts (Eyster, 2018). Researchers at the Urban Institute and Abt Associates prepared a series of reports from these third-party evaluations and the national evaluations to draw lessons for future initiatives (Eyster et al., 2020; Kuehn and Eyster, 2020; Mikelson et al., 2017). These results revealed promising developments in how community colleges changed career pathways systems and programs to better support adult learners, though capacity challenges born by community colleges that receive modest public funding compared to other higher education institutions (e.g., public and private universities) hampered some of the goals and outcomes grantees hoped to achieve.

A separate meta-analysis of the TAACCCT third-party evaluations was conducted by a team of researchers affiliated with New America and the University of Washington (Blume et al., 2019; Bragg, 2019). This study conducted after the TAACCCT grant ended in 2017 included a systematic review of more than 200 third-party evaluation reports that were funded under the grant. Using a set of criteria to assess the rigor of the research design used in these studies, this meta-analysis focused on 36 quasi-experimental design (QED) evaluations to estimate overall effects sizes for education and employment outcomes. Results revealed a positive effect on both education and employment outcomes for TAACCCT participants, with a greater magnitude of effect on the education outcomes (program completion and credential attainment) than the employment outcomes (post-program employment and pre-to post-program wage change). This analysis found students in TAACCCT-funded programs were nearly twice as likely to complete a program or earn a credential than comparison students, and 27% more likely to have positive labor market outcomes (employment or wage gains) than comparison students.

Complementing the quantitative meta-analysis was a qualitative study designed to understand implementation of key program elements funded by TAACCCT. This analysis confirmed career pathways implemented widely by grantees included stacked and latticed credentials, prior learning assessment, comprehensive student supports, career advising and guidance, and employment supports. Partnerships between community colleges, workforce agencies, and employers were also developed through TAACCCT, and grantees also mentioned the importance of TAACCCT's extensive funding of third-party evaluations as an important way to improve the tracking of participant outcomes.

What works clearinghouse meta-analysis on career pathways

Commissioned by the What Works Clearinghouse (WWC) of the USDoEd, a panel of researchers and practitioners was formed to summarize rigorous studies that met the WWC standards and present results and recommendations to guide future career pathways implementation in the US community college context. An exhaustive review of published and unpublished literature located over 16,000 studies potentially relevant to the project, identifying 122 meeting the strictest eligibility requirements for review. Of these, 22 studies met the WWC standards with or without reservations, and most of these focused on career pathways entry and integrated education and training functions. Nearly all these 22 studies also focused on adult education, workforce training and non-degree postsecondary education, providing further evidence pertaining to how entry and integrated education and training manifest in career pathways. Results of the studies were synthesized around five themes that formed the basis for recommendations that represent the most systematic and exhaustive analyses of career pathway completed to date in the US (Cotner et al., 2021).

Applying the meta-analysis process required by the WWC, nine of the 22 studies showed overall statistically significant favorable results on college credit accumulation, industry-recognized credential attainment, short-term employment, short-term earnings, and non-degree credentials. Only two studies measured college degree attainment, highlighting a void in what is known about how career pathways help to advance participants beyond short- and medium-term education and credentialing. The panel's recommendation to intentionally design and structure career pathways with career progression in mind is intended to enable students to access further education and obtain and advance in employment over the course of their lives. This recommendation also recognized that, whereas all educational institutions enroll students in courses and programs, career pathways should prioritize the student's entire educational experience, from access and entry, to completion and credentialing, and on to employment and career advancement. This purposeful approach to supporting student attainment continues to be theorized to lead to positive student outcomes despite a dearth of longitudinal studies to assess these results.

Studies that met the WWC standards revealed findings on college credit accumulation, industry-recognized credential attainment, and postsecondary degree attainment. Though this group of studies was relatively small, results showed a consistent, positive impact on college credit accumulation and industry-recognized credential attainment. These results showed students' educational outcomes were improved and compelled the panel to recommend delivering contextualized or integrated basic skills instruction to accelerate students' entry into and successful completion of career pathways. Some of the studies also provided descriptive results on program implementation that may be useful to scaling up career pathways in the community college context.

Studies meeting the WWC standards were further scrutinized to understand results pertaining to five education outcomes: industry-recognized credential attainment, academic performance, technical skill proficiency, credit accumulation, and postsecondary degree attainment. As seen in the above mentioned discussion, results on industry-recognized credential attainment were statistically significant and favorable. However, results on academic performance, technical skill proficiency, and credit accumulation were indeterminate, meaning findings could be favorable or unfavorable but were not statistically significant. Moreover, nearly all the studies that met the WWC standards focused on non-degree credentials, providing limited information about results on postsecondary degree attainment. Given available evidence, the panel recommended community colleges should offer flexible

instructional delivery schedules and models to improve credit accumulation and enhance completion of non-degree credentials along career pathways.

The rigorous studies meeting WWC standards and measuring outcomes relative to comprehensive student supports focused on the education outcome domain only. Similar to the previous recommendation on flexible scheduling, findings on industry-recognized credential attainment were statistically significant and favorable. Findings on credit accumulation were also statistically significant and favorable, reinforcing the findings of Blume et al., 2019 suggesting holistic supports are related to favorable student outcomes. Drawing on these findings, the panel recommended that career pathways provide coordinated comprehensive student supports to improve credit accumulation and completion of non-degree credentials along career pathways.

Statistically significant and favorable outcomes were found in the WWC meta-analysis process for short-term employment, short-term earnings, and medium-term earnings. However, for medium-term employment, the meta-analysis results were indeterminate. To clarify, this does not mean employment-focused partnerships do not lead to labor market success; instead it points to a dearth of research that tracks students through training and into employment for long enough to measure long-term outcomes. The panel's recommendation related to partnerships encouraged community colleges to develop and leverage partnerships to prepare students and advance in the labor market. This recommendation specifically emphasizes labor market success, based on studies that examined the following four short- and medium-term labor market outcomes: short-term employment, short-term earnings, medium-term employment, and medium-term earnings.

Research needed on career pathways

Going forward, it is important to continue to evaluate the impact of career pathways on student outcomes using rigorous research designs that generate results for underserved student populations. Knowing that inequality is growing in the US by race, income, social class, age, and other demographic characteristics is critical to understanding how career pathways work in the aggregate, but more importantly how they work for student sub-groups. The vision of career pathways to address structural inequalities inherent in postsecondary education and workforce training programs remains unknown without systematic examination of outcomes by race, income, and other factors known to exacerbate inequality. Implementing and scaling improved career pathways systems and programs requires reaching youths and adults who are most disadvantaged by current dysfunctional policies and practices.

Moreover, additional research is needed on the key career pathways functions of entry, integrated education and training, and career progression. Each of these functions plays a specific role in the overall career pathways framework, and yet many research and evaluation studies fail to distinguish which functions are being employed in the career pathways they are examining and who these functions are intended to serve. With a more nuanced approach to research, it may be possible to more clearly understand how career pathways functions prepare students for employment and career advancement. Again, this research should disaggregate results by race, income, social class, age, and other demographics to reveal how career pathways may be benefiting student sub-groups differentially and inform implementation to avoid so systematic inequities.

References

- Allsid, J., Gruber, D., Jenkins, D., Mazzeo, C., Roberts, B., Stanback-Stroud, R., 2002. Building a Career Pathways System: Promising Practices in Community College-Centered Workforce Development. Workforce Strategy Center, Brooklyn, NY. <https://files.eric.ed.gov/fulltext/ED470457.pdf>.
- Anderson, T., Kuehn, D., Eyster, L., Barnow, B., Lerman, R., 2017. New Evidence on Integrated Career Pathways: Final Impact Report for Accelerating Opportunity. Urban Institute, Washington, DC. http://www.urban.org/sites/default/files/publication/91436/ao_final_impacts.pdf. (Accessed 10 November 2021).
- Bathmaker, A.M., 2015. Thinking with Bourdieu: thinking after Bourdieu. Using "field" to consider in/equalities in the changing field of English higher education. *Camb. J. Educ.* 45 (1), 61–80. <https://doi.org/10.1080/0305764X.2014.988683>.
- Blume, G., Meza, E.A., Bragg, D.D., Love, I., 2019. Estimating the Impact of the Nation's Largest Single Investment in Community Colleges: Lessons and Limitations of TAACCCT Evaluations. New America, Washington, DC. <https://www.newamerica.org/education-policy/reports/estimating-impact-taaccct/>.
- Bragg, D.D., 2015. President Obama's career-technical education and workforce development initiatives: will they matter? In: Zamani-Gallaher, E. (Ed.), *The Obama Administration and Educational Reform: Toward New 21st Century Schools for the Global Knowledge Economy*. Emerald Group Publishing Limited, Bingley, United Kingdom.
- Bragg, D.D., 2019. Lessons from TAACCCT: Recommendations for Future Evaluation Investments. New America, Washington, DC. Retrieved from: <https://www.newamerica.org/education-policy/policy-papers/lessons-taaccct/>.
- Bragg, D.D., Endel, B., Anderson, N., Sarconi, L., Acevedo, E., 2019. What Works for Adult Learners. Jobs for the Future, Boston, MA. <https://files.eric.ed.gov/fulltext/ED598339.pdf>. (Accessed 29 November 2021).
- Carnevale, A., Schmidt, P., Strohl, J., 2020. The Merit Myth: How Colleges Favor the Rich and Divide America. The New Press. <https://thenewpress.com/books/merit-myth>. (Accessed 30 November 2021).
- Cohen, E., Michelson, K.S., Durham, C., Eyster, L., 2017. TAACCCT Grantee Characteristics: The Trade Adjustment Assistance Community College and Career Training Grant Program Brief 2. Urban Institute, Washington, DC. <https://www.urban.org/research/publication/taaccct-grantee-characteristics>. (Accessed 10 October 2021).
- Cotner, H., Bragg, D., Goold, G., Heiser, E., Miller, D., Van Noy, M., 2021. Designing and Delivering Career Pathways at Community Colleges: A Practice Guide for Educators. <https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/WWC-PraxGuide-Career-Pathways-full-text-Final-508.pdf>. (Accessed 1 December 2021).
- Domina, T., McEachin, A., Handselman, P., Agarwal, P., Hwang, N., Lewis, R.W., 2019. Beyond tracking and detracking: the dimensions of organizational differentiation in schools. *Sociol. Educ.* 92 (3), 293–322. <https://doi.org/10.1177/0038040719851879>.
- Durham, C., Eyster, L., Mikelson, K. S., Cohen, E. 2017. Early Results of the TAACCCT Grants, Brief 4. Urban Institute, Washington DC. https://www.urban.org/sites/default/files/publication/89336/2017.02.08_taaccct_brief_4_final_v2.pdf

- Durham, B., Bragg, D.D., 2019. The contested evolution and future of vocational education in the United States. In: Guile, D., Unwin, L. (Eds.), *Wiley Handbook of Vocational Education and Training*. Wiley Blackwell, pp. 229–250.
- Eyster, L., 2018. Building Workforce Success. Urban Institute, Washington, DC. file:///Users/dbragg/Desktop/Back-up%2011.29.15/Current-Files-2015-11.29.15/Misc/CP-Book-Chapter/CP-Chapter/building_workforce_success_advancing_careers_with_multiple_postsecondary_credentials.pdf.
- Eyster, L., 2020. Pathways to success through career and technical education: developing and evaluating community college career pathways initiatives. In: Hedges, J.W., Carter, G.W. (Eds.), *Career Pathways: From School to Retirement*. Oxford University Press, pp. 36–61.
- Eyster, L., Briggs, A., Durham, C., Hafford, C., Spievack, N., 2020. Systems Change in Community Colleges: Lessons From a Synthesis of the Round 3 TAACCCT Third-Party Evaluation Findings (Research Report). Prepared for the US Department of Labor, Chief Evaluation Office. Urban Institute, Washington, DC.
- Hedge, J.W., Carter, G.W., 2020. Understanding the concept and promise of career pathways. In: Hedge, J.W., Carter, G.W. (Eds.), *Career Pathways: From School to Retirement*. Oxford Press.
- Herhusky, S., Ludwig, D., January 28, 2021. Career Pathways, Laddering, and Lattices: Getting From Point A to Point B. <https://www.careerforcemn.com/careerforce-blog/career-pathways-laddering-and-lattices-getting-point-a-to-point-b#:~:text=Career%20pathways%20encompass%20both%20career%20ladders%20and%20career,and%20new%20employment%20opportunities%20in%20a%20given%20field>. (Accessed 4 December 2021).
- Jacobs, J., Worth, J., 2019. The Evolving Mission of Workforce Development in the Community College. Community College Research Center, Teachers College. <https://files.eric.ed.gov/fulltext/ED594218.pdf>. (Accessed 1 December 2021).
- JPMorgan Chase & Co, February 12, 2020. JPMorgan Chase Investing \$75 Million Dollars to Prepare Young People for Future of Work. <https://www.jpmorganchase.com/news-stories/jpmc-investing-75mm-to-prepare-young-people-for-future-of-work>.
- Keating, J., Medrich, E., Volkoff, V., Perry, J., 2002. Comparative Study of Vocational Education and Training Systems: National Vocational Education and Training Systems Across Three Regions Under Pressure of Change. NCVET.
- Kuehn, D., Eyster, L., 2020. A Synthesis of Findings From the Round 3 Trade Adjustment Assistance Community College and Career Training Third-Party Impact Evaluations (Research Report). Prepared for the U.S. Department of Labor, Chief Evaluation Office. The Urban Institute, Washington, DC.
- McKay, H., Van Noy, M., 2021. The goals and context for TAACCCT. In: Van Noy, M., McKay, H., Bragg, D.D. (Eds.), *Large-scale change: lessons from TAACCCT, New Directions for Community Colleges*, vol. 193, pp. 15–22.
- McKay, H., Vinton, J., Boyle, J., Van Noy, M., September 2021. Lessons From European States: Policy and Practice in Career and Technical Education. Rutgers Education and Employment Research Center. <https://www.preparingtechnicians.org/wp-content/uploads/Lessons-from-the-European-States-with-Appendices.pdf>. (Accessed 6 December 2021).
- Mikelson, K.S., Eyster, L., Durham, C., Cohen, E., March 2017. TAACCCT Goals, Design, and Evaluation: The Trade Adjustment Assistance Community College and Career Training Grant Program Brief 1. Urban Institute, Washington, DC. <https://www.dol.gov/sites/dolgov/files/OASP/legacy/files/20170308-TAACCCT-Brief-1.pdf>.
- Moodie, G., 2008. From Vocational to Higher Education: An International Perspective. Open University Press.
- Mortrude, J., September 2018. Better Together: Career and Guided Pathways. Center for Law and Social Policy, Washington DC. <https://www.clasp.org/sites/default/files/publications/2018/09/2018careerandguidedpathways.pdf>.
- Osteltu, M.O., McCann, C., Laitanen, A., May 2021. The Short-Term Credentials Landscape. New America. <http://newamerica.org/education-policy/reports/the-short-term-credentials-landscape/>. (Accessed 6 November 2021).
- Protsch, P., Solga, H., 2016. The social stratification of the German VET system. *J. Educ. Work* 29 (6), 637–661. <https://doi.org/10.1080/13639080.2015.1024643>.
- Sarna, M., Strawn, J., 2018. Career Pathways Implementation Synthesis: Career Pathways Design Study. U.S. Department of Labor, Washington DC.
- Schwartz, R.B., 2016. The career pathways movement. *J. Soc. Issues* 72 (4), 740–759. <https://doi.org/10.1111/josi.12192>. <https://www.dol.gov/asp/evaluation/completed-studies/Career-Pathways-Design-Study/3-Career-Pathways-Implementation-Synthesis.pdf>.
- Staklis, S., Foster, L.R., Mikulka, D., Smith, C., 2020. Connecting college and the workplace through pathway development in Kansas. In: Hedges, J.W., Carter, G.W. (Eds.), *Career Pathways: From School to Retirement*. Oxford University Press.
- U.S. Department of Education, Office of Career, Technical, and Adult Education, 2015. The Evolution and Potential of Career Pathways. U.S. Department of Education, Office of Career, Technical, and Adult Education, Washington, DC. http://s3.amazonaws.com/PCRN/docs/Evolution_of_Career_Pathways_092315.pdf.
- U.S. Department of Education, Office of Vocational and Adult Education, 2012. Promoting College and Career Readiness: Bridge Programs for Low-Skill Adults. U.S. Department of Education, Office of Vocational and Adult Education, Washington, DC. <https://www2.ed.gov/about/offices/list/ovae/pi/cclo/brief-1-bridge-programs.pdf>.
- U.S. Department of Labor, 2016. Trade Adjustment Assistance Community College and Career Training Grant Program: Fiscal Years 2014–2016 Report to the Committee on Finance of the Senate and Committee on Ways and Means of the House of Representatives. U.S. Department of Labor, Washington, DC.
- Wachen, J., Jenkins, D., Belfield, C., Van Noy, M., 2012. Contextualized College Transition Strategies for Basic Skills Students: Learning From Washington State's I-BEST Model. Community College Research Center, Teachers College. <https://files.eric.ed.gov/fulltext/ED538999.pdf>. (Accessed 5 November 2021).
- Webb, S., Bathmaker, A.-M., Gale, T., Hodge, S., Parker, S., Rawolle, S., 2017. Higher vocational education and social mobility: educational participation in Australia and England. *J. Vocat. Educ. Train.* 69 (1), 147–167. <https://doi.org/10.1080/13636820.2016.1269359>.
- Wheelahan, L., Moodie, G., 2017. Vocational education qualifications' roles in pathways to work in liberal market economies. *J. Vocat. Educ. Train.* 69 (1), 10–27.
- Zeidenberg, M., Cho, S., Jenkins, D., 2010. Washington State's Integrated Basic Education and Skills Training Program (I-BEST): New Evidence of Effectiveness. Community College Research Center, Teachers College, Columbia University, New York. <https://ccrc.tc.columbia.edu/publications/i-best-new-evidence.html>.

Higher education: service-learning as pedagogy, partnership, institutional organization, and change strategy

Robert G. Bringle^a and Patti H. Clayton^{a,b}, ^a Indiana University Purdue University Indianapolis, Indianapolis, IN, United States; and ^b PHC Ventures, Cary, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Conceptualizing service-learning	477
Service-learning as community-campus engagement	477
Service-learning as experiential education	478
Defining service-learning	478
Empirical support for student learning through service-learning	481
Longitudinal studies	481
Authentic evidence	482
Course characteristics	482
Community-campus relationships and partnerships	482
Institutional organization to support service-learning	483
Institutional organization to support the development of relationships and partnerships	485
Conclusion	487
References	487
Relevant websites	490

Higher education must not only give solid skills for the present and the future world but must also contribute to the education of ethical citizens committed to the construction of peace, the defense of human rights and the values of democracy.

UNESCO (2009, p. 2).

Internal and external pressures are stimulating re-examination of the academy's primary role. Is the core purpose teaching the disciplines, cultivating critical thinking and problem-solving, preparing students for careers, promoting economic development, contributing to local communities, generating knowledge, or cultivating the civic capacities needed for a flourishing democracy? The easy, and not inaccurate, answer is that the mission of higher education encompasses all of these possibilities and others.

Bringle et al. (2014, p. 14).

Service-learning is a type of experiential education and a form of community-campus engagement that offers higher education what may be its best means of fulfilling the complex and vital charges articulated above. It is a key strategy for reclaiming and centering civic purposes—understood as the intertwining of learning, action, and change—for everyone in the academy and for institutions of higher education (Saltmarsh and Zlotkowski, 2011). From its roots in social justice, service-learning has evolved over the past 45+ years to be not only a mechanism for social change and a means of engaging students and instructors with community members and issues but also a pedagogy variously focused on education for democracy, disciplinary and interdisciplinary learning, and career development around the world. It is increasingly used as a strategy for democratic engagement, collective impact, and critical engagement with systems of oppression (Bringle et al., 2014; Clayton and Bandy, 2021; Mitchell, 2008). Service-learning in the southern hemisphere is explicitly oriented toward awareness of and action on issues associated with power, oppression, privilege, and justice—a framing that is often referred to there in terms of solidarity and that is known in the United States as critical service-learning (Mitchell, 2008; Tapia, 2012). It is understood by *Centro Latinoamericano de Aprendizaje y Servicio Solidario* (CLAYSS), for example, as an approach to co-learning from experience in which all participants “together commit themselves in the transformation of reality” (CLAYSS, no date, para 1).

Originating in the United States, service-learning has spread and grown around the world. Such organizations as the International Association for Research on Service-Learning and Community Engagement, the Asia-Pacific Regional Conference on Service-Learning, the Service-Learning Asia Network, the European Observatory of Service-Learning in Europe, the Latin American Center for Solidarity Service-Learning, Engagement Australia, and the South African Higher Education Community Engagement Forum have enabled the worldwide development of both practitioner networks and a growing base of service-learning research, theory, and best practices.

Key to the potential of service-learning is civic action as both a means and a desired outcome of education. Service-learning has the potential to “change everything” (Saltmarsh, 2010, p. 332) about higher education's functions. As a substantive and often counter-normative alternative to hierarchical and technocratic processes, service-learning—when designed to fulfill its democratic, critical, transformative potential—confronts the academy with an array of complexities, challenges, and possibilities. Thus, it can provide a basis for contributing to the development of individuals and institutions in ways called for in the opening quotes.

We examine the meanings, purposes, and roles of service-learning in higher education; empirical support for service-learning's contributions to students' academic learning, civic learning, and personal growth; the nature of the community-campus relationships at the heart of service-learning; institutional support for service-learning; and how that support can be directed toward enhancing community-campus relationships in service-learning and community engagement more generally. Our goal is to advance conversations and deepen practices in the academy and in communities related to the potential of service-learning to enhance learning and action, deepen relationships between individuals and organizations on campuses and in communities, and promote positive change. The stakes are both timely and enduring given the significant and unique contributions higher education can and must make with local to global communities to co-create "a world that is increasingly peaceful, compassionate, just, inclusive, and verdant" (Clayton et al., 2014, p. 6).

Conceptualizing service-learning

Service-learning is one example of processes through which campuses collaborate with broader communities that intertwine learning and change. The broader framework of community-campus engagement provides a basis for understanding service-learning as not only a pedagogy but also a means of developing relationships with community members and organizations in the context of which higher education institutions (re-)organize themselves to co-create change. In this section we locate service-learning within both community engagement and experiential education frameworks and explore key elements of a definition of service-learning in this light.

Service-learning as community-campus engagement

Launched in the United States in 2006 and increasingly of interest in other countries (it was piloted in Ireland; McIlrath et al., 2021), the Carnegie Foundation's Elective Classification for Community Engagement foregrounds service-learning as a central dimension of the engaged campus and defines community-campus engagement as the "partnership (of knowledge and resources) between colleges and universities and the public and private sectors" in order to "enrich scholarship, research, and creative activity; enhance curriculum, teaching and learning; prepare educated, engaged citizens; strengthen democratic values and civic responsibility; address critical societal issues; and contribute to the public good" (Carnegie Foundation for the Advancement of Teaching, 2020). A similar benchmarking tool for community engagement that has been developed in Australia (Garlick and Langworthy, 2008). Farnell and Ilic (2021) developed a Toward a European Framework for Community Engagement in Higher Education Toolkit that was piloted in four universities in four countries that also included community-engaged teaching and service-learning.

Fig. 1 provides a framework for conceptualizing community engagement in terms of faculty teaching, research, and service (Bringle et al., 1999). This framework of faculty work, which is from the United States, deviates somewhat from the European framework of teaching, research, and third mission; readers can adapt this conceptual analysis to their own contexts. The Venn diagram in Fig. 1 indicates that all three types of activities can occur both on campus and in communities. For example, a campus can deliver courses at community sites (included in Europe's third mission but included under teaching in Fig. 1) and online as well as on campus; researchers can collect data in communities as well as in classrooms and laboratories; and instructors and staff can provide professional services in communities as well as to their disciplines, departments, and institutions.



Fig. 1 Community engagement as faculty work in the community. Adapted from Bringle et al. (1999, p. 5).

Community engagement, however, is distinct from teaching, research, and service that take place *in* and *for* communities. Community engagement refers instead to activities undertaken *with* communities—with “students, faculty, and community members functioning as co-educators, co-learners, and co-generators of knowledge” (Jameson et al., 2011, pp. 261–262). Reciprocity encompasses respectful integration of multiple perspectives, epistemologies, and assets and is a defining aspect of community engagement and service-learning. From the perspective of faculty work conveyed in Fig. 1, service-learning as a form of community engagement may be seen as the intersection of service with teaching (and/or with research, although not expressed in this visual) in which faculty co-create community-engaged activities that intertwine learning, action, and change with community members, students, and professional staff to the benefit of learning and to the benefit of communities.

Service-learning as experiential education

Service-learning is also grounded in the pedagogical principles of active and experiential learning articulated by Dewey, Lewin, and Piaget, among others (Kolb, 2015). Bringle and Plater (2017) presented a framework of experiential education, modified slightly, that identified the following five components that are integrated in curricular experiential education:

- (a) learning goals/objectives,
- (b) content/body of thought (related to the learning goals/objectives),
- (c) experience/activity (related to the content and the learning goals/objectives),
- (d) critical reflection (on the experience and the content, oriented toward the learning goals/objectives), and
- (e) assessment and feedback (at least of the learning, including as generated through critical reflection on the experience and content).

This framework guides the intentional design of any form of experiential education, including service-learning, toward academically-grounded, assessable learning outcomes.

Service-learning (as with any form of experiential education), when carefully designed (Hatcher et al., 2017), is well-aligned with the attributes of teaching and learning that produce deeper learning. It also pushes the nature and understanding of teaching and learning further through, for example, demonstrating the powerful processes of co-inquiry and co-learning that include not only instructors and students but also members of broader communities. The science of learning provides information on how learning occurs and guidelines for how it can be improved. For example,

What neuroscience researchers have made clear is that “the one who does the work does the learning” (Doyle, 2008, p. 25). The more ways you engage with something that you are learning—such as listening, talking, reading, writing, reviewing, or thinking about the material or skill—the stronger the connections in your brain become and the more likely the new learning will become a more permanent memory.

Doyle and Zakrajsek (2013, p. 7).

Related to these findings is an extensive body of research in psychology on critical thinking and transfer of learning across settings (Fitch et al., 2013). As Halpern (2014) pointed out, “The whole enterprise of learning how to improve thinking is of little value if these skills are only used in the classroom or only on problems that are very similar to those presented in class” (p. 14). Providing learners with diverse and authentic contexts that include multiple opportunities to practice, obtain feedback on, and improve learning is a powerful approach to generating deeper learning. Furthermore, advances in pedagogy should reflect what research has delineated as instructional components that produce deeper learning: (a) active learning; (b) frequent feedback from others (e.g., instructors, other students, community members); (c) collaboration with others; (d) cognitive apprenticeship (i.e., mentored relationships in which students can discuss and learn generalization of principles, transfer of knowledge between theory and practice, and analysis of perplexing circumstances); and (e) practical applications that involve students in tasks that have consequences with a safety net as a buffer against high-stakes mistakes (Marchese, 1997). Well-designed service-learning reflects all of these attributes.

Defining service-learning

There are many forms of experiential education, some of which take place in communities (e.g., internships, clinical practice, field research, work-integrated experiences). However, most of these do not have all the characteristics that define service-learning. After examining nine definitions, De la Cerda et al. (2009) identified seven common features: “responding to social needs or carrying out an action to benefit the community, learning something, performing a service, having a meaningful experience, performing activities of reflection, collaborating with other social institutions, and contributing to training for citizenship” (Sotelino-Losada et al., 2021, p. 3). US scholars Furco and Norvell (2019) reviewed lists of essential elements of service-learning and found that “there are fundamental definitions, elements, and principles of service-learning that apply no matter what the situation or context;” they also noted, however, that the “cultural fibre of the societies in which service learning is practised will ultimately shape overall character of the service learning experience” (p. 32).

In the European context, service-learning is considered a “mechanism” through which “to achieve the ideals associated with the democratic citizen and the promotion of active citizenship” (McIlrath et al., 2019, p. 2), which have become a focus and priority of

higher education on the continent. Europe Engage, a 3-year project started in 2015 involved 12 European Union higher education institutions and was developed to broaden and deepen service-learning as a pedagogy throughout the region. Europe Engage (no date) identified 14 quality standards for service learning, including “defined goals,” “systematic reflection,” “support and coaching for students,” “evaluation and documentation,” and “sustainability” (Preradovic and Stark, 2019) and defined it as:

An innovative pedagogical approach that integrates meaningful community service or engagement into the curriculum and offers students academic credit for the learning that derives from active engagement within community and work on a real world problem. Reflection and experiential learning strategies underpin the learning process and the service is link [sic] to the academic discipline. Service-learning brings together students, academics and the community whereby all become teaching resources, problem solvers and partners. In addition to enhancing academic and real world learning, the overall purpose of service learning is to instill in students a sense of civic engagement and responsibility and work towards positive social change within society.

Europe Engage (n.d.).

We offer the following US.-based definition of service-learning as an example of one that integrates many of the essential elements and that offers specificity that is useful in course design and associated research:

Service-learning is both course-based, credit bearing pedagogy and change strategy that engages students, community members, and instructors/staff in co-creating relationships that integrate academic material, community-engaged activities, and critical reflection to advance public purposes and to achieve clearly articulated academic learning, civic learning, and personal growth goals.

Adapted from Bringle and Clayton (2012, p. 105).

This definition provides a basis for examining several important attributes of service-learning. First, it identifies service-learning as academic work, in contrast with volunteering and co-curricular civic education programs, although some definitions of service-learning do include co-curricular activities in which service is co-determined with community members, aligned with learning goals, and integrated with reflection.

Second, it insists that service-learning is about change as well as learning. The change aspired to may be on campus and can include developing democratic classrooms; recruiting and retaining diverse students, staff, and faculty; operationalizing initiatives for equity and inclusion; strengthening community-campus partnerships; connecting various parts of campus that interact with broader communities; expanding understandings of teaching and scholarship in promotion, tenure, and advancement policies; connecting the goals and activities of academic and student affairs; deepening public trust and accountability; and achieving re-accreditation, quality assurance, and various forms of recognition. Change aspired to in the community may include improving quality of life, strengthening nonprofit organizations and networks as well as community-campus partnerships, informing government agencies about social issues, disrupting systems that perpetuate injustice, increasing available resources, building evaluation capacities, organizing community residents for advocacy work, developing laws and policies, and contributing to environmental and economic sustainability. *Centro Latinoamericano de Aprendizaje y Servicio Solidario* (CLAYSS), which is based in Argentina and facilitates regional and international service-learning networks, conceptualizes service learning largely in terms of change, defining it as:

A pedagogical connection in which the educator and educated learn together from experience and together commit themselves in the transformation of reality. It implies action and reflection on the practice and establishment of ties that allow to act and to learn reciprocally with and from the community (CLAYSS, no date, para 1).

In an overview of service-learning in Africa, Pachó (2019) noted that the focus is often on “(local) developments for a sustainable economy and social welfare” (p. xiv).

Third, the definition we offer highlights the centrality of partnerships in the design, implementation, and assessment of service-learning courses. In the United States, partnerships in service-learning generally include relationships between and among any of the following constituencies of the SOFAR model: Students, community Organization staff, Faculty, institutional Administrators or professional staff on campus and Residents/clients (Bringle et al., 2009). Community organizations are often but not necessarily non-profits or non-governmental agencies. Governmental and municipal agencies, grassroots movements, neighborhood associations, funding organizations, and other collectives may also function as partners. Saltmarsh et al. (2009) noted and called for a shift from (a) community-campus interactions that incorporate the dominant model of academic-expert-centered, deficit-based, hierarchical relationships on campus and in communities toward (b) partnerships grounded in democratic values and commitments that position everyone involved as co-educators, co-learners, and co-generators of knowledge and practice and that foreground an asset-based orientation in which power and responsibility are shared by all partners. Democratic engagement is characterized by partnerships that transcend exchange-based transactions and aspire to transformational change (e.g., of self, others, community organizations, communities, educational institutions, systems, paradigms). In the same vein, the inclusion of *solidario* (solidarity) in the name of the pedagogy as CLAYSS defines it “underline[s] that we are not referring to a traditional, paternalistic, occasional, or superficial kind of service, but to a more horizontal bond between peoples and communities working together for the common good” (CLAYSS, no date, para 1).

Fourth, this definition indicates that partners undertake one or more community-engaged activities as part of service-learning. In all cases, activities must be designed to attend to the full range of goals collectively identified, including promoting student (and other partners') learning and advancing desired change in the broader community (and on campus) while minimizing unintended costs. Community-engaged activities may include direct service (i.e., supporting one another face-to-face), indirect service (i.e., working behind the scenes at a nonprofit organization to develop processes, resources, and/or materials), research (i.e., gathering, analyzing, sharing information with the community), advocacy and social change (i.e., increasing awareness, policy change), or relationship-building and cultural exchange activities (i.e., spending time getting to know one another as individuals and populations). The United Nations' Sustainable Development Goals (SDGs) provide a structure for organizing and determining important foci for service-learning activities in areas such as responsible production and consumption, education, gender equality, affordable and clean energy, environmental issues, and reduced inequality.

Fifth, critical reflection is highlighted in this definition as the necessary process of making meaning from the milieu of academic material and community-engaged activities. Indeed, it is the component of service-learning (and experiential education generally) that generates, deepens, and documents learning and improves the quality of practice, partnerships, and inquiry (Ash and Clayton, 2009). Reflection is best understood and undertaken as critical reflection—as a process of making meaning that is grounded in critical thinking and in the systems critique of critical theory. Reflection should occur before, during, and after community-engaged activities; can take written, oral, audio-visual, or embodied forms; can be conducted individually, with other students, with community members, or with instructors; and can be done in person or virtually (synchronously and asynchronously). Well-crafted critical reflection prompts guide the learner's attention to the presence (or absence) of aspects of their experiences that are related to learning goals and pushes their thinking beyond initial interpretations. The DEAL model provides a structure to guide the design of critical reflection prompts that are tightly aligned with desired learning objectives (Ash and Clayton, 2009). This customizable, research-grounded model supports learners in describing (D) their experiences, examining (E) their experiences using prompts linked to learning goals, and then articulating (A) learning (L) in a way that leads to enhanced future action and ongoing learning (specifically, in this last step, by answering four questions: What did I learn? How did I learn it? Why does this learning matter? What will I do in light of it?). Tools and rubrics affiliated with the DEAL model use Paul and Elder's (2001) standards of critical thinking (e.g., clarity, depth, breadth, fairness) and Bloom's (1956) Taxonomy—moving reasoning from the lower levels of identification, explanation, and application to the higher levels of analyzing, evaluating, and synthesizing—to structure feedback and to assess direct evidence of quality of learning.

Finally, this definition illustrates that service-learning is typically viewed as contributing to three primary domains of learning objectives: academic learning, personal growth, and civic learning. Sometimes professional development, metacognition, ethical inquiry, intercultural or global learning, and other categories of learning are added or are integrated into these three. Regardless of how these categories are conceptualized, the learning focus of service-learning is not limited to students, which is often the default presumption, but also includes learning by all partners (i.e., instructors, campus staff, community residents, community organization staff).

Emphasis around the world has, however, been on conceptualizing, designing for, and measuring student outcomes in particular for some aspects of all three of the primary categories of learning goals. This has been the case in the United States (e.g., Steinberg et al., 2011; Moely et al., 2002); Singapore (Choo et al., 2019), Hong Kong (Ma et al., 2019; Snell and Lau, 2020); Slovakia, Croatia, and Romania (Ilic et al., 2021), and Spain (Regueiro et al., 2021).

Academic learning objectives are often primary. Jameson et al. (2013) offered a framework for conceptualizing and organizing academic learning into foundational learning, higher-order learning, thinking from disciplinary/interdisciplinary perspectives, and critical thinking. Personal growth learning goals require learners to relate their experiences (e.g., observations, interactions) to their values, assumptions, behavioral tendencies, decision-making, strengths, and lives more generally. Brandenberger (2013) explored sub-categories of personal growth that are widely used in service-learning, including moral development and education, identity development, spirituality, self-efficacy, and self-awareness. Civic learning is a defining domain of learning goals of service-learning (as distinct from most other community-based pedagogies and classroom-based curricula) and is conceptualized in a wide variety of ways around the world, including, for example: social responsibility, knowledge and skills for democracy, values clarification, moral development, public good, citizenship, democratic competences, sense of solidarity, Ubuntu, commitment to and capacities to advance social justice, civic identity, cosmopolitan citizenship, life purpose, and global citizenship. Chapters in *Research on Student Civic Outcomes in Service Learning* (Hatcher et al., 2017) explored different meanings of, measures for, and approaches to cultivating and studying civic learning in the US, including from several disciplinary perspectives. Bringle and Clayton (2020) described approaches to designing service-learning courses to generate civic learning outcomes as conceptualized from a variety of national, regional, and disciplinary perspectives and proposed methods for assessing student achievement of civic learning outcomes.

What constitutes a civic learning goal in one context might be viewed as an academic or personal growth learning goal in another. These primary categories of learning goals in service-learning can also be conceptualized as closely intertwined (Hemer and Reason, 2017). The Council of Europe's (2016) Competences for Democratic Culture, for example, identified several domains that implicate personal growth and civic learning and, depending on the context, may constitute or overlap with academic learning as well: values such as human dignity, cultural diversity, and social justice; attitudes such as openness to cultural otherness, respect, and tolerance of ambiguity; and skills such as listening and observing, empathy, and conflict resolution; and knowledge related to world politics, law, human rights, cultures, religions, and history. Sotelino-Losada et al. (2021) highlight "social commitment that entails the development of awareness ethics" (p. 4)—which speaks to both civic learning and personal growth—as key to

distinguishing service-learning from other approaches to experiential education. The emphasis placed by CLAYSS (no date) on solidarity and action yields a vision of citizenship—with corollary, multi-faceted learning goals that can be conceptualized in terms of any or all of the three categories and their intersections—in terms of “not just the knowledge of norms and values and the diagnosis of political and socioeconomic problems” but also “the elaboration of proposals and ... the active participation [in] initiatives” as well as “taking responsibility and commitment in the construction of overcoming alternatives, and the work in articulation with the authorities and civil society organizations.”

Service-learning’s framework for organizing learning into academic, civic, and personal growth learning categories can be used to focus critical reflection toward sets of desired learning outcomes related to global citizenship, which are relevant in many if not most higher education contexts around the world. The Curriculum Framework (Osman et al., 2017) associated with the United Nations’ Social Development Goals, for example, posits such competencies as being able to imagine a better future, learning to question current belief systems, systemic thinking, learning to work together, empowering oneself and others, resilience, and commitment to justice. As a second example, UNESCO’s Framework for Global Citizenship (2015) includes understanding “connections between global, national, and local systems and processes”; appreciating “difference and multiple identities”; developing skills in “information technology,” “media literacy,” and “peace building”; recognizing the influence of beliefs on decision-making and on “perceptions of social justice”; caring for people and the environment; valuing fairness; and developing skills to “critically analyze inequalities based on gender, socio-economic status, culture, religion, age and other issue” (p. 16).

Empirical support for student learning through service-learning

Empirical findings provide a context for understanding how service-learning can contribute to and enrich learning outcomes in higher education. The vast majority of research on service-learning has focused on student outcomes (Clayton et al., 2013a,b; Hatcher et al., 2017), and we limit this section accordingly. Chapters in Clayton et al. (2013a,b) reviewed research, theory, and assessment strategies for inquiry into relationships, processes, and outcomes associated with service-learning for faculty, institutions, communities, and partnerships as well as for students.

Kuh (2008) identified service-learning as a high-impact practice based on students’ self-reported gains of deep learning, general learning, personal development, and practical competence. He determined that “historically underserved students tend to benefit more” from engaged pedagogies such as service-learning than majority students do, although these practices are “high-impact” for all types of students (Kuh, 2008, p. 17). Finley (2011) found that, relative to other high-impact practices studied, service-learning had “the greatest impact on each of the four outcomes [learning, general education, personal development, practical competence] measured” (p. 2).

Meta-analyses of research have documented the contributions of service-learning to student learning outcomes. Celio et al. (2011) found significant improvements associated with service-learning for attitudes toward self, attitudes toward school and learning, social skills, and academic performance among pre-collegiate and collegiate students. Yorio and Ye (2012) found similar results for understanding of social issues, personal insight, and cognitive development. According to Conway et al. (2009), service-learning was associated with positive outcomes in academic, personal, social, and citizenship domains among pre-collegiate, collegiate, and adult students. Novak et al. (2007) compared courses with and without a service-learning component using a quantitative measure of cognitive learning; the effect size in this meta-analysis that was attributable to service-learning was a 53% improvement in learning. Warren (2012) found a similar effect size with a somewhat different sample of research studies. Summarizing the findings of several major studies comparing students who did and did not participate in service-learning, Eyler (2010) found that service-learning increased political interest and efficacy, connectedness to community, social responsibility, future intent to participate in community life, and life skills.

Fitch et al. (2013) summarized research on cognitive outcomes that have been found to result from participation in service-learning, including critical thinking, intellectual development, problem solving, and transfer of learning. More generally, service-learning is associated with such academic success variables such as motivation to complete college and attitudes toward learning (Lockeman and Pelco, 2013; Reed et al., 2015). Research confirms that service-learning contributes to civic knowledge, civic skills, civic attitudes and values, civic behavior, and civic identity (e.g., Battistoni, 2013; Hatcher et al., 2017; Hemer and Reason, 2017). Related to personal growth, Holsapple (2012) documented outcomes including belief in the value of diversity, and Everhart (2016) documented enhanced empathy.

Longitudinal studies

Using a sample of 3450 students attending 42 institutions in the United States, Astin and Sax (1998) found that service-learning had a strong impact on reported academic learning, life skill development, and civic responsibility upon graduation, after controlling for pre-existing student characteristics. Astin et al. (2006) found that service-learning was associated with post-graduation civic leadership, charitable giving, and political engagement 10 years after entering the university. Kerrigan (2005) reported that students in a service-learning course demonstrated enhanced communication abilities, leadership skills, community involvement, understanding of diversity, and career development 3 years after graduation. In Hong Kong, after graduation, students who had participated in service-learning (compared to those who had not) reported higher civic responsibility (e.g., civic awareness, civic efficacy), greater community involvement, and more hours of service per month (Ma et al., 2016).

Authentic evidence

Although most of the research on student outcomes of service-learning has used self-report measures, a few studies have documented authentic evidence of learning that is independently evaluated. Markus et al. (1993) compared students in service-learning and non-service-learning sections of a large political science course and found that learning and course grades were higher for service-learning students. Eyster and Giles (1999) conducted interviews focused on analyzing social problems, with interview transcripts rated by blind reviewers. Relative to students in non-service-learning courses and poorly implemented service-learning courses, students in well-implemented service-learning courses exhibited improvements in causal complexity, solution complexity, knowledge application, and personal political strategy. Reeb et al. (1999) found that service-learning and non-service-learning students had similar levels of academic performance early in the semester but service-learning students showed increasingly superior performance subsequently. Ash et al. (2005) had raters evaluate written critical reflection products from service-learning courses using two types of rubrics and found improvements across the semester on higher order reasoning abilities and critical thinking within all three categories of academic learning, civic learning, and personal growth. Thus, the results that have been found with self-report measures of student outcomes have been supported for authentic, direct measures of learning.

Course characteristics

Knowing how the design of service-learning components (e.g., learning objectives, discipline, reflection, community-engaged service activities, partnerships) differs across courses can clarify which course characteristics are related to which outcomes. Service-learning is too often treated as a constant, homogeneous intervention in research into student outcomes. However, not all service-learning courses are designed according to principles of good practice, and not all well-designed service-learning courses are implemented according to their design.

One of the most central issues in optimal course design is the quality of reflection. Critical reflection can cultivate students' abilities to evaluate various perspectives and circumstances, which enhances their ability to think critically—about themselves (personal growth); about power dynamics, citizenship, policy-making, and systems change (civic learning); and about the content of their courses and the assumptions, methods, and theories of (inter)disciplines (academic learning) (Ash and Clayton, 2009). Outcomes are enhanced when the content of the reflection activities is aligned with the identified academic learning, civic learning, and personal growth outcomes (Astin et al., 2006; Dahan, 2016; Moely and Ilustre, 2014); and when reflection takes multiple forms (e.g., written, oral, collaborative) (Astin et al., 2006; Dahan, 2016). Hatcher et al. (2004) found that regular reflection, structured reflection, and reflection focused on personal values were independently related to the quality of the learning environment in service-learning courses (e.g., active learning, course satisfaction, peer interaction, personal relevance). Conway et al. (2009) meta-analysis found that structured reflection (compared to unstructured reflection) yielded larger effects on citizenship outcomes, social outcomes, and personal outcomes. Choo et al. (2019) found in a study in Singapore that students' perceived quality of reflection prompts was related to academic connections, career preparation, and changes in civic attitudes.

The selection or development of the community-engaged service activities also needs to be aligned with desired outcomes (e.g., for learning, for community change). Brown et al. (2016) found that direct service that facilitated autonomy-oriented helping (compared to indirect service and dependency-oriented helping) resulted in more positive views of social equality. Miller and Yen (2005) reported higher final exam scores for students involved in direct (compared to indirect) service. In Dahan's (2016) research, reflection, direct service, and service duration (>30 h of service) were related to positive change in civic attitudes. Choo et al. (2019) found that students in Singapore reported greater changes in civic outcomes, greater academic connection, and greater career preparation when they also thought they were well-prepared for and engaged in meaningful service activities.

Community-campus relationships and partnerships

Service-learning and community engagement explicitly place importance on community-campus relationships and partnerships, recognizing that they are at the heart of new paradigms for framing activities in the community (i.e., teaching, research, and service in and with communities) that can be both just and meaningful because they integrate knowledge, resources, experience, and questions from a variety of perspectives. Relationships and partnerships may be local, international, or virtual; and they may be with nonprofits, for-profits, government agencies, or schools (to name only a few possibilities). The differences within and among them implicate a wide variety of opportunities and challenges related to goals, culture, history, personalities, economics, politics, and place. The possibilities for developing, maintaining, deepening, and ending as well as for studying relationships and partnerships are rich and varied. In this section, a few ways to conceptualize partnerships, organize them, and investigate them are summarized, and a later section examines ways institutions can support partnerships.

An early framework conceptualizing the types of relationships and partnerships formed by communities and campuses identified five categories, differentiated by duration and complexity: (a) one-time interactions; (b) short-term placements; (c) ongoing, with mutual dependence; (d) core partnerships, with interdependence; and (e) joint creation, with transformation (Enos and Morton, 2003). Any given relationship may move among these categories over time; and any one campus or any one community organization (or community) may engage in more than one of these types of relationships with more than one community or campus partner, simultaneously and across time. More recently, Janke and Clayton (2015) conceptualized community-campus partnerships

into three types, each of them associated with different levels of institutional support (see more on this below): (a) partnerships grounded in individual agendas and interpersonal relationships; (b) partnerships that take the form of networks or collectives, and (c) multi-faceted, institution-level, inter-organizational partnerships. Regardless of the particular typology, clarity as to the nature of the relationship in question aids in establishing goals, designing activities, inquiring into processes and outcomes, and building appropriate mechanisms for communication and trust.

Bringle et al. (2009) drew a distinction between “relationship” and “partnership,” suggesting that partnerships comprise a subset of relationships based on three attributes: closeness, equity, and integrity. Relationships, some of which are partnerships, can be interpersonal (e.g., Bringle and Clayton, 2013) or interorganizational (Janke, 2013) and can take the form of dyads, triads, networks, or partnership entities (Kniffin et al., 2020). As discussed earlier, the SOFAR model parses “community” into representatives of community Organizations and Residents and “campus” into Students, Faculty, and Administrators or professional staff. Bringle et al. (2009) posited that SOFAR relationships can be characterized on an “E-T-T” continuum that ranges from Exploitative (in which someone is taken advantage of or harmed), through Transactional (in which value is exchanged) to Transformational (in which new value is generated and individuals, organizations, partnerships, and/or systems change). Relationships may exhibit characteristics associated with multiple points along the E-T-T continuum (e.g., they may be transactional in some aspects and transformational in others), and they may shift in either direction on the continuum across time and across contexts (Clayton et al., 2010).

The Transformational Relationship Evaluation Scale (TRES) was developed to operationalize the E-T-T continuum. TRES I focuses on the interpersonal relationships in dyadic SOFAR relationships (Bringle et al., 2009). TRES II explicitly broadens the focus of inquiry to include not only interpersonal relationships but also networks (e.g., multi-partner, inter-organizational) and partnership entities (Kniffin et al., 2020). Both TRES I and TRES II invite perceptions of the current and the desired status on each of several relationship characteristics (e.g., goals, decision-making, power, resources). These tools have proven useful in many contexts for studying relationships, structuring reflection on relationships, assessing some of the attributes of particular relationships, conducting research, and guiding practice (Kniffin et al., 2020).

Inter-organizational relationships and partnerships have the potential for enhanced durability (i.e., they are structured to persist despite changes in individual participants), reliability (i.e., they operate in consistent ways), and accountability (i.e., they have standards and guidelines) (Janke, 2013). Interorganizational partnerships in service-learning can evolve into entities with their own partnership identity, in which members are invested in the entity itself, not only in the organizations they represent. Janke’s research into partnership entities focused on “the extent to which and conditions under which partners developed a collective and coherent sense of who they were as a partnership distinct from their employing organizations” (2013, p. 585) and determined that “partnerships that develop an identity may be likely to last beyond the present project; endure changes in persons and resources; and transform how partners view each other, the partnership, and even their employing institutions” (p. 586).

Inquiry into interorganizational relationships and partnership entities is complicated because they involve interpersonal relationships as well, similar to the ways in which inquiry into interpersonal relationships and partnerships must take into account the influences of a variety of organizational, institutional, and other contexts. In addition, whether at the interpersonal or interorganizational level, relationships and partnerships are dynamic, changing as participants and external conditions change. In both practice and inquiry, it is essential to ask who is and is not “at the table” and what dynamics of power and privilege shape their influence on the relationships and partnerships at the heart of service-learning. Bandy et al. (2018) provided a framework and questions to critically evaluate the extent to which inquiry into partnerships aligns with the commitments to democratic engagement (e.g., full participation, co-creation). It includes an extensive critique of the work on TRES through this lens, which has informed both the ongoing refinement of those tools and approaches to assessing the qualities of relationships and partnerships more generally (Kniffin et al., 2020).

Institutional organization to support service-learning

Whereas some examples of service-learning result from the initiative of individual instructors, staff, and/or students, institutional support begins with institutional planning, decision-making, and investment. The impetus for those decisions might begin with senior administration, student leadership, faculty governance, faculty or staff champions, strategic planning, mission-driven changes, institutional and national policies, and accreditation/quality assurance processes. Early on, much effort was devoted to making service-learning a normative part of the work of higher education (i.e., to institutionalize it) by establishing its academic legitimacy. Thus, the types of institutional support included the commitment of permanent funds for centralized offices (preferably located in academic affairs) and the investment of resources (e.g., release time, conference travel, grants) for faculty development. Centralized units can hire and house professional staff (e.g., community-engagement professionals, service-learning specialists, community-campus liaisons, assessment specialists, communications professionals, development professionals), allocate resources to stimulate and sustain service-learning courses, determine and promote best practices in service-learning, assess and report on outcomes, receive inquiries from community organizations and students, and facilitate appropriate linkages of campus and community resources. Examining infrastructure for service-learning in Hong Kong, Ma and Mun (2019) concluded that there was “evidence... that faculty members found support from the OSL [Office of Service-Learning] to be essential and vital to their successful adoption of SL in their teaching” (p. 48). Ti et al. (2021) reached a similar

conclusion based on their evidence of rapid growth of service-learning when it was introduced as their Singaporean institution's signature pedagogy. From the European perspective, a centralized unit has also been endorsed as important to the growth of service-learning (Europe Engage, 2017).

One level of intervention is focused on helping instructors develop service-learning courses. Many departments, programs, and campuses use workshops to introduce instructors to service-learning (Welch and Plaxton-Moore, 2017). Bringle and Hatcher (1995) presented a curriculum for faculty workshops that includes partnerships as one topic. These workshops can be facilitated by professional staff, experienced service-learning instructors, disciplinary specialists, community partners, and students. Less frequently used (Welch and Plaxton-Moore, 2017), faculty learning communities or communities of practice engage a small cohort of instructors in a series of scaffolded workshops over a period of time (e.g., one semester, an academic year) during which they read, discuss, share ideas, and develop plans for implementing service-learning (e.g., Bringle et al., 2000; Jameson et al., 2012; Rice and Stacey, 1997). Mini-grants to instructors to develop service-learning courses have also been used to motivate faculty (Bringle et al., 2000; Bringle and Hatcher, 1995; Rice and Stacey, 1997). Embedded in all of these support programs can be the attributes that transformational and democratic partnerships aspire to. As just one example, professional development for faculty can be held at community sites and co-designed and co-facilitated by community partners (e.g., Studer et al., 2017).

Typically, service-learning occurs at the course level when an instructor collaborates with partners (e.g., students, community members, campus staff) to integrate community-engaged service activities and critical reflection activities into a course. It can also occur at the curricular level when designed to connect course clusters (e.g., themed learning communities), to deepen categories of courses (e.g., 1st year seminars), or to build developmentally across course sequences (e.g., minors, capstones, certificates). Service-learning thus can be—and needs to be—institutionally supported at multiple levels (e.g., individuals, courses, departments, schools, degree programs, campus as a whole). Therefore, strategies that develop leadership, advocacy, and support among multiple constituencies (e.g., executives, deans, chairs, faculty, students, community partners) are important.

Several frameworks and tools have been created over the past several decades to support campuses in such strategic advancement of institutionalization. Bringle and Hatcher's (1996, 2000) Comprehensive Action Plan for Service Learning (CAPSL) is a framework for developing centralized institutional support for service-learning and partnerships. CAPSL delineates four constituencies: institution, faculty, students, and communities and identifies for each the following steps toward institutionalizing service-learning: planning, increasing awareness, identifying prototypes, acquiring resources, initiating activities that result in expansion, providing recognition, monitoring courses, conducting evaluation, conducting research, institutionalization. Other approaches to advancing institutionalization of service-learning and community engagement involve the use of rubrics that describe levels of institutionalization for each of several domains (e.g., Furco, 2002; Holland, 2001).

However, institution type (e.g., 2-, vs. 3-, vs. 4-year vs. graduate; religious; private; public comprehensive; research intensive; metropolitan) and climate (Janke, 2013) as well as specifics of community, economic, political, historic, regional, and national contexts significantly influence the nature of potential institutional support and the processes for cultivating, maintaining, and extending it. Zlotkowski (1998) gathered case studies of approaches taken by different types of institutions in the US. Tijsma et al. (2021) studied institutionalization at a research university, a university of applied science, and a vocational education institute. They reported that the institutional priorities and conflicting structures were different at each institution and, therefore, steps taken to institutionalize service learning need to be tailored to each institution.

The Global University Network for Innovation (2014) overviewed work related to knowledge production, community engagement, and service-learning in Africa, Arab States, Asia and the Pacific, Europe, Canada and North America, and Latin America and the Caribbean. Institutions of higher education around the world are developing their own models of engagement that reflect unique mission statements, institutional traditions and structures, historical and cultural contexts, and community assets. These include examples in Europe (e.g., Aramburuzabala et al., 2019), China (Ma et al., 2018), Ireland (McIlrath et al., 2012), Eastern Europe (Regina and Ferrara, 2017), Singapore (Ti et al., 2021), Asia (Xing and Ma, 2010), Australia (Patrick et al., 2019; Sachs and Clark, 2017), South Africa (Stanton and Erasmus, 2013), and Hong Kong (Ma, 2018). Cross-national comparative studies have emphasized that although there are similarities in the goals to enhance student learning and develop the public purposes of higher education, differences in norms and traditions associated with civic involvement, economic differences, political context, and institutional environments all shape the implementation of service-learning (Bernardo et al., 2012; Hatcher and Erasmus, 2008; Stanton and Erasmus, 2013; Thomson et al., 2011).

Throughout more than four decades of development of service-learning in the United States, and during a shorter time around the world, practitioners have realized that further expansion and improvement of service-learning, and more broadly community engagement, can be facilitated by strategic institutional support (Bringle and Hatcher, 2000; Furco, 2002; Welch and Saltmarsh, 2013; Young et al., 2007). The type of support an institution provides for service-learning, however, depends in part on the goals. One key distinction is whether the intent is to "institutionalize" service-learning—embed it into existing structures—or to leverage service-learning to transform the institution to be more democratic in all aspects of its work. Much of the history of service-learning has embodied the tension between these two understandings of the relationship between service-learning and institutions, with the current conversation arguably evolving in the direction of institutional transformation and perhaps beyond. Zlotkowski's (1995) answer to the question of whether service-learning had a future was that "nothing less than a transformation of contemporary academic culture" would be required (p. 130). Several years ago, once again considering the meaning and role of institutional transformation, he acknowledged that while institutionalization had occurred in many ways, institutional transformation had not. He called attention beyond that distinction, noting that "Recent developments in the world beyond the academy compel us to ask: When does the need to act more forcefully in the greater community overshadow—not eliminate—other needs such as institutional

transformation?" (2015, p. 84). Thinking strategically about institutional support for service-learning highlights the importance of considering the mix of ultimate—perhaps intertwined, perhaps at odds—purposes: institutionalization, institutional transformation, or societal transformation.

With this historical perspective and the distinction between institutionalization and institutional transformation in mind, Saltmarsh et al. (2015) proposed a framework for the latter they called "DPI," standing for Deep, Pervasive, and Integrated. In order for service-learning and community engagement to transform a campus, they posited, they must embody second order change (i.e., fundamental culture change on campus). According to Eckel et al. (1998), on whose work they drew, deep change "profoundly ... affects behavior or alters structures ..., imply[ing] a shift in values and assumptions that underlie the usual way of doing business" (p. 4); and pervasive change "is far-reaching within the institution ... cross[ing] unit boundaries and touch[ing] different parts of the institution" (p. 4). However, even deep and pervasive change can be compartmentalized. Integrative change related to service-learning "foregrounds the synergies that can result from holistic and interdependent approaches to institutional priorities, highlighting the importance of aligning and intertwining [service-learning/community engagement] with ... campus initiatives such as access and inclusion, citizenship and leadership development, internationalization, and assessment" (Saltmarsh, Janke, and Clayton, pp. 124–125). The implications of institutional support for service-learning with institutional transformation as the goal are thus many. Saltmarsh, Janke, and Clayton explored questions of service-learning infrastructure through the lens of DPI:

Given that we seek deep or high quality (asset-based, reciprocal, critically reflective, impactful, continuously improving), pervasive (spread into all corners and dimensions of the life of the campus), and integrated (aligned and intertwined with the full range of campus priorities) SLCE [service-learning, community engagement], what questions do we need to ask in determining whether and how to create institution-wide centers or offices? ... How do we operationalize the interconnections between diversity and community engagement as we think about infrastructure? Do we continue to have offices with staff responsible for these various initiatives working in different parts of campus but in a more intentionally collaborative fashion? Or do we need to embed SLCE within these other offices and integrate SLCE principles and practices within existing structures? (p. 125).

Additional questions about infrastructure that are increasingly being posed at the intersection of institutionalization, institutional transformation, and societal transformation priorities include, for example: (a) whether physical space that houses service-learning and community engagement staff and resources ought to be located off-campus, amid partnering organizations; (b) whether community partners should be financially compensated for participating in professional development, mentoring students, and co-designing or co-facilitating courses; (c) whether to invest limited financial and time resources into faculty conference travel or international travel opportunities for students or directly into social justice initiatives; and (d) whether professional staff can appropriately dedicate work hours to engagement with community partner organizations. Thus, institutional support specifically in the context of service-learning relationships and partnerships, to which we now turn, may generate challenging questions about where it is most important and most urgent to focus institutional support given the ultimate purposes posited for higher education—purposes at the interface of learning, civic action, and change.

Institutional organization to support the development of relationships and partnerships

Institutional support consists of intentional and deliberate steps to provide some form of infrastructure, funding commitments, professional assistance, and policies to establish, support, and improve relationships and partnerships across all engagement activities, including service-learning. In all cases, promoting community-campus relationships that aspire to being transformational and democratic partnerships should focus on all constituencies in SOFAR (i.e., students, community organization staff, faculty, administrators/staff on campus, and community residents). Each of these can be in a position to advocate for and help develop democratic, transformative partnerships. This section shares examples related to community partners and faculty; other sources contain information for those interested in work related to community engagement professionals (e.g., Dostilio, 2017; Dostilio and Welch, 2019) and students (e.g., Zlotkowski et al., 2006) as supporters of all aspects of service-learning, including partnerships.

Engaging with community members to determine overall goals for service-learning and community engagement and associated infrastructure, funding, professional assistance, and policies to support relationships and partnerships is key and, in some quarters, increasing. Sandy and Holland (2006) asked community partners about their perspectives on service-learning partnerships and documented particular interest in:

how to engage more deeply with faculty; recognizing partners as co-educators of students, with all the planning and preparation that entails; committing the academic institution to ongoing action as an institutional citizen in larger-scale community development projects; and finding ways to develop appropriate institutional infrastructure that supports and facilitates these shifts in emphasis and practice (p. 40).

Similarly, and more recently, interviews with community partners related to their perceptions of the processes and outcomes of their partnerships with higher education institutions (Muse, 2018) revealed their conviction that democratic partnerships are desired outcomes in themselves and that they require foregrounding relationship quality. The community partners in the

Sandy and Holland (2006) study identified working effectively with faculty as the central challenge they face and expressed a desire for more direct and longer-term interaction, a more collaborative approach to course design, and deeper engagement of faculty in their organizations' work. They also indicated that they most valued opportunities to be co-educators of students. Writing about the value they placed on institutional policies and practices that support faculty in the work of democratic partnerships, community partners stated:

We were able to level the playing field to a great degree because of the ... faculty members' great commitment to learning to work in a more participatory process such as being trained in consensus decision-making.

Freeman et al. (2009, pp. 93–94).

Thus, designing institutional support for service-learning and community engagement should be deeply informed by—if not also undertaken collaboratively with—community partners.

As institutional support is initiated, developed, and strengthened, faculty are central to the development of relationships and partnerships in service-learning and community engagement.

Faculty development plays a key role (e.g., Bringle and Hatcher, 2000; Furco, 2002; Holland, 2001) and can be accomplished through such mechanisms as faculty/staff learning communities, workshops, presentations, discipline specific consultants, and informal one-on-one consultation with professional staff. Levkoe and Stack-Cutler (2018) provided an analysis of relationship brokers in Australia and the United Kingdom.

Another level of focus in institutional support for activities that include SOFAR relationships occurs at the meta-course level. Institutions can devote activities and resources to having departments and schools review their entire curriculum, consider where service-learning fits within it, and plan what steps might be taken to enhance its role across the curriculum (Kecskes, 2013). Such an approach encourages re-examination of unit-wide learning outcomes, including civic learning, and explicit mapping of courses and pedagogical approaches to those outcomes. In addition, it provides opportunities for members of an academic unit to consider how and with whom they are partnering in communities and how they can strategically improve SOFAR relationships in ways that are consistent with transformational and democratic partnerships (Messer and Kecskes, 2008). Service-learning courses and other community engagement activities can be part of interdisciplinary coalitions that are organized around a common theme (e.g., health, environment, food, poverty) or geographic area (Tijmsa et al., 2021). Campuses can then coordinate and support multiple courses with one or more community organizations.

Both for individual instructors and courses and for programs operating across multiple service-learning courses and community engagement activities, there are opportunities for professional staff in centers to support the development of democratic relationships, reciprocal relationships, and transformational partnerships. Bringle et al. (2009) suggested using TRES to deliberately make visible the perspective of each individual in the SOFAR model concerning a relationship's position on the E-T-T continuum, and professional staff can facilitate such examination. One strategy is to have each individual in a relationship complete TRES, indicating their perceptions of the current and desired status of each item, and then bring them together to compare and contrast their responses and envision the future of the relationship accordingly (Bringle et al., 2009; Bringle and Clayton, 2013). Such a process can result in determination of appropriate changes and investment in the capacity building needed to take concrete steps to move forward together.

Davis et al. (2017) described another approach to facilitating learning and change associated with service-learning relationships. Their approach has participants from both campus and community organizations map power and reciprocity followed by dialog on civic engagement and transformative learning. The presumption is that this process generates creative tensions among participants that allow them to “uncover premature consensus and insincere tolerance” (p. 46). Participants can then identify mutually agreed upon goals, make changes in goals and processes, and learn from each other about how the relationship should proceed.

A partnership between Museum Studies faculty and the staff of several local museums in Indianapolis was similarly supported by a service-learning specialist in using the visual reflection method of Collaborative Relationship Mapping (Price, 2016) to examine the ways the nature of their relationships—particularly in terms of closeness and integrity—influenced programmatic and organizational outcomes. Among the transformational results of this activity, the partners created new processes to educate employees and faculty about the history of the partnership and initiated longer-term joint planning cycles to enable deeper connections (Bandy et al., 2018).

As another example of institutional support for faculty development focused on enhancing relationships and partnerships, external consultants can be engaged to facilitate self-study. Miller-Young et al. (2015) reported on such a process, in which a group of faculty and staff examined a key tenet of SOFAR relationships, reciprocity, using a method from Decoding the Disciplines (Pace and Middendorf, 2004). Consisting of reflecting individually, interviewing one another, and discussing experiences and readings, the process provided a space in which participants deepened their understanding of reciprocity and identified associated actions to improve their practice. Both the participants in and the facilitators of the project experienced transformative learning about the nature of reciprocity. Anderson (2017) and Lloyd et al. (2017) also provided examples of centralized initiatives that operate across a campus and promote quality relationships across community-campus projects.

Conclusion

Service-learning is an evolving field. Some of the recent trends that are shaping the future include credentialing service-learning and civic engagement professionals (Dostilio, 2017; Dostilio and Welch, 2019); scaffolding service-learning across sequences of courses and across disciplines (Jameson et al., 2008; Ti et al., 2021); assessing the quality of components of service-learning courses and mapping components to outcomes (Bringle et al., 2017; Furco and Matthews, 2018); developing global service-learning with and without international travel and attending to implicated ethical issues (Green and Johnson, 2014; Hartman et al., 2018); integrating service-learning with other pedagogies, including internships and research (Bringle, 2017) and into course clusters (MacGregor, 2003); enhancing the democratic aspects of relationships, processes, and outcomes of service-learning and community-campus engagement (Bandy et al., 2018; Saltmarsh et al., 2009); integrating service-learning with distance education and technology (Bringle and Clayton, 2020; Waldner et al., 2012); integrating initiatives focused on diversity, equity, and inclusion with service-learning and community engagement (Sturm et al., 2011); and using the United Nations' Sustainable Development Goals as a basis for organizing and coordinating activities across service-learning and community engagement (e.g., Gonzalez Rabanal, 2021). With all of these recent and emerging developments and in light of its intertwining of learning, civic action, and change, service-learning is clearly at the forefront of the academy's revisioning around its public purposes and more authentic relationships with community constituencies.

References

- Anderson, L., 2017. Useful, usable, and used. *Gatew. Int. J. Community Res. Engagem.* 10, 58–77.
- Aramburuzabala, P., McIlrath, L., Opazo, H. (Eds.), 2019. *Embedding Service Learning in European Higher Education*. Routledge, New York.
- Ash, S.L., Clayton, P.H., 2009. Generating, deepening, and documenting learning: the power of critical reflection for applied learning. *J. Appl. Learn. High. Educ.* 1, 25–48.
- Ash, S.L., Clayton, P.H., Atkinson, M.P., 2005. Integrating reflection and assessment to improve and capture student learning. *Mich. J. Community Serv. Learn.* 11 (2), 49–60.
- Astin, A.W., Sax, L.J., 1998. How undergraduates are affected by service participation. *J. Coll. Student Dev.* 39, 251–263.
- Astin, A.W., Vogelgesang, L.J., Misa, K., Anderson, J., Denson, N., Jayakumar, J., Saenz, V., Yamamura, E., 2006. *Understanding the Effects of Service-Learning: A Study of Students and Faculty*. Report to the Atlantic Philanthropies. Higher Education Research Institute, Los Angeles.
- Bandy, J., Price, M.F., Clayton, P.H., Metzker, J., Nigro, G., Stanlick, S., Etheridge Woodson, S., Bartel, A., Gale, S., 2018. *Democratically Engaged Assessment: Reimagining the Purposes and Practices of Assessment in Community Engagement*. Imagining America, Davis.
- Battistoni, R.M., 2013. Civic learning through service learning: conceptual frameworks and research. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning*. Vol. 2A: Students and Faculty. Stylus, Sterling, pp. 111–132.
- Bernardo, M., Butcher, J., Howard, P., 2012. An international comparison of community engagement in higher education. *Int. J. Educ. Dev.* 32 (1), 187–192.
- Bloom, B.S., 1956. *Taxonomy of Educational Objectives: Handbook I. Cognitive Domain*. David McKay, Philadelphia.
- Brandenberger, J.W., 2013. Investigating personal development outcomes in service learning. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning: Conceptual Frameworks and Assessment* (Vol. 2A). Stylus, Sterling, pp. 133–156.
- Bringle, R.G., Clayton, P.H., 2012. Civic education through service-learning: what, how, and why? In: McIlrath, L., Lyons, A., Munck, R. (Eds.), *Higher Education and Civic Engagement: Comparative Perspectives*. Palgrave, New York, pp. 101–124.
- Bringle, R.G., Clayton, P.H., 2013. Conceptual frameworks for partnerships in service learning. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning: Conceptual Frameworks and Assessment* Vol. 2B: Communities, Institutions, and Partnerships. Stylus, Sterling, pp. 539–571.
- Bringle, R.G., Clayton, P.H., 2020. Integrating service learning and digital technologies: examining the challenge and the promise. *RIED* 23 (1), 43–65.
- Bringle, R.G., Hatcher, J.A., 1995. A service-learning curriculum for faculty. *Mich. J. Community Serv. Learn.* 2, 112–122.
- Bringle, R.G., Hatcher, J.A., 1996. Implementing service learning in higher education. *J. High. Educ.* 67, 221–239.
- Bringle, R.G., Hatcher, J.A., 2000. Institutionalization of service learning in higher education. *J. High. Educ.* 71, 273–290.
- Bringle, R.G., Plater, W.M., 2017. Reflections on the Macquarie experience. In: Sachs, J., Clark, L. (Eds.), *Learning through Community Engagement: Vision and Practice in Higher Education*. Springer, New York, pp. 301–319.
- Bringle, R.G., Games, R., Malloy, E.A. (Eds.), 1999. *Colleges and Universities as Citizens*. Allyn & Bacon, Boston.
- Bringle, R.G., Games, R., Ludlum, C., Osgood, R., Osborne, R., 2000. Faculty Fellows Program: enhancing integrated professional development through community service. *Am. Behav. Sci.* 43, 882–894.
- Bringle, R.G., Clayton, P.H., Price, M.F., 2009. Partnerships in service-learning and civic engagement. *Partnerships* 1 (1), 1–20.
- Bringle, R.G., Edwards, K.E., Clayton, P.H., 2014. The roots of service-learning as a means to advance the civic mission of community colleges. In: Traver, A., Katz, Z.P. (Eds.), *Service-Learning at the American Community College: Theoretical and Empirical Perspectives*. Palgrave Macmillan, New York, pp. 13–35.
- Bringle, R.G., 2017. Hybrid high-impact pedagogies: integrating service-learning with three other high-impact pedagogies. *Mich. J. Community Serv. Learn.* 24 (1), 49–63.
- Bringle, R.G., Hatcher, J.A., Hahn, T.W., 2017. Introduction to research on service learning and student civic outcomes. In: Hatcher, J.A., Bringle, R.G., Hahn, T.W. (Eds.), *Research on Student Civic Outcomes in Service Learning: Conceptual Frameworks and Methods*. Stylus, Sterling, pp. 3–24.
- Brown, M.A., Wymer, J.D., Cooper, C.S., 2016. The counter-normative effects of service-learning: fostering attitudes toward social equality through contact and autonomy. *Mich. J. Community Serv. Learn.* 23 (1), 37–44.
- Carnegie Foundation for the Advancement of Teaching, 2020. *Elective Community Engagement Classification: First-Time Classification Documentation Framework*. <https://drive.google.com/file/d/107oaW9gbV6k6UllUgZpcD9VgIHraWwvNMQ/view>.
- Celio, C.I., Durlak, J., Dymnicki, A., 2011. A meta-analysis of the impact of service-learning on students. *J. Exp. Educ.* 34, 164–181.
- Centro Latinoamericano de Aprendizaje y Servicio Solidario, (no date). What is "Service Learning"? http://www.clayss.org/english/servicelearning_school.html.
- Choo, J., Tan, Y.K., Ong, F., Shiluan, S.T., Nair, S., Ong, J., Chan, A., 2019. What works in service-learning? Achieving civic outcomes, academic connection, career preparation and personal growth in students at Ngee Ann Polytechnic. *Mich. J. Community Serv. Learn.* 25 (2), 95–132.
- Clayton, P.H., Bandy, J., 2021. Are We a Field or a Movement, and Why Should We Care? Revisiting Recurring Tensions. Paper Presented at the Gulf South Summit, Belmont University, Nashville.
- Clayton, P.H., Bringle, R.G., Senor, B., Huq, J., Morrison, M., 2010. Differentiating and assessing relationships in service-learning and civic engagement: exploitive, transactional, and transformational. *Mich. J. Community Serv. Learn.* 16 (2), 5–21.
- Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), 2013a. *Research on Service Learning: Conceptual Frameworks and Assessment*. Vol. 2A: Students and Faculty. Stylus, Sterling.

- Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), 2013b. *Research on Service Learning: Conceptual Frameworks and Assessment*. Vol. 2B: Communities, Institutions, and Partnerships. Stylus, Sterling.
- Clayton, P.H., Hess, G., Hartman, E., Edwards, K.E., Shackford-Bradley, J., Harrison, B., McLaughlin, K., 2014. Educating for democracy by walking the talk in experiential learning. *J. Appl. Learn. High. Educ.* 6, 3–36.
- Conway, J.M., Amel, E.L., Gerwien, D.P., 2009. Teaching and learning in the social context: a meta-analysis of service learning's effects on academic, personal, social, and citizenship outcomes. *Teach. Psychol.* 36, 233–245.
- Council of Europe, 2016. *Competences for Democratic Culture: Living Together as Equals in Culturally Diverse Democratic Societies*. Council of Europe Publishing, Strasbourg.
- Dahan, T.A., 2016. Revisiting pedagogical variations in service-learning and student outcomes. *Int. J. Res. Serv. Learn. Community Engagem.* 4 (1), 3–15.
- Davis, K.L., Klierer, B.W., Nicolaides, A., 2017. Power and reciprocity in partnerships: deliberative civic engagement and transformative learning in community-engaged scholarship. *J. High. Educ. Outreach Engagem.* 21 (1), 30–54.
- De la Cerda, M., Graell, M., Martián, X., Muñoz, A., Puig, J.M., 2009. Aprendizaje servicio: ejemplos y definiciones. In: Puig, J.M. (Ed.), *Aprendizaje Servicio (ApS)*. Educación y compromiso cívico. Graó, Barcelona, pp. 15–32.
- Dostilio, L.D., Welch, M., 2019. *The Community Engagement Professionals Guidebook*. Campus Compact, Boston.
- Dostilio, L.D. (Ed.), 2017. *The Community Engagement Professional in Higher Education: A Competency Model for an Emerging Field*. Campus Compact, Boston.
- Doyle, T., 2008. *Helping Students Learn in a Learner-Centered Environment: A Guide to Facilitating Learning in Higher Education*. Stylus Publishing, Sterling, VA.
- Doyle, T., Zakrajsek, T., 2013. *The New Science of Learning: How to Learn in Harmony With Your Brain*. Stylus, Sterling.
- Eckel, P., Hill, B., Green, M., 1998. *En Route to Transformation (On Change)*. American Council on Education, Washington, DC.
- Enos, S., Morton, K., 2003. Developing a theory and practice of campus community partnerships. In: Jacoby, B., Associates (Eds.), *Building Partnerships for Service-Learning*. Jossey-Bass, San Francisco, pp. 20–41.
- Europe Engage, 2017. *Europe Engage: Developing a Culture of Civic Engagement Through Service-Learning With Higher Education in Europe: Guidelines for Institutionalization of Service-Learning*. Universidad Autonoma de Madrid, Madrid.
- Europe Engage, (n.d.). Our Definition of Service-Learning, <https://europeengage.org/our-definition-of-service-learning/>.
- Everhart, R.S., 2016. Teaching tools to improve the development of empathy in service learning students. *J. High. Educ. Outreach Engagem.* 20 (2), 129–156.
- Eyler, J.S., Giles Jr., D.E., 1999. Where's the Learning in Service-Learning? Jossey-Bass, San Francisco.
- Eyler, J.S., 2010. What International Service Learning Research can learn from research on service learning. In: Bringle, R.G., Hatcher, J.A., G Jones, S. (Eds.), *International Service Learning: Conceptual Frameworks and Research*. Stylus, Sterling, pp. 225–242.
- Farnell, T., Ilic, B.C., 2021. Towards a European framework for community engagement in higher education. In: Hall, B., Rajesh, T. (Eds.), *Socially Responsible Higher Education: International Perspectives on Knowledge Democracy*. Brill, Leiden, pp. 253–264.
- Finley, A., 2011. Assessment of high-impact practices: using findings to drive change in the Compass Project. *Peer Rev.* 13 (2), 29–33.
- Fitch, P., Steinke, P., Hudson, T.D., 2013. Research and theoretical perspectives on cognitive outcomes of service learning. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning*. Vol. 2A: Students and Faculty. Stylus, Sterling, pp. 57–83.
- Freeman, E., Gust, S., Aloshe, D., 2009. Why faculty promotion and tenure matters to community partners. *Metrop. Univ.* 20 (2), 87–103.
- Furco, A., Matthews, P.H., 2018. Using a New, Research-Based Tool to Assess the Quality of Planning and Implementation of Service-Learning Courses. Paper Presented at the Conference of the International Association for Research on Service-Learning and Community Engagement, New Orleans, LA.
- Furco, A., Norvell, K., 2019. What is service learning? Making sense of the pedagogy and practice. In: Aramburuzabala, P., McIlrath, L., Opazo, H. (Eds.), *Embedding Service Learning in European Higher Education: Developing a Culture of Civic Engagement*. Routledge, New York, pp. 13–35.
- Furco, A., 2002. Institutionalizing service-learning in higher education. *J. Publ. Aff.* 6, 39–67.
- Garlick, S., Langworthy, A., 2008. Benchmarking university community engagement: developing a national approach in Australia. *High. Educ. Manag. Pol.* 20 (2), 105–117.
- Global University Network for Innovation, 2014. *Higher Education in the World 5: Knowledge, Engagement and Higher Education: Contributing to Social Change*. Palgrave MacMillan, New York.
- Gonzalez Rabanal, M.C., 2021. Indicadores para evaluar la eficiencia de programas sociales. Una propuesta a partir de un Proyecto de Aprendizaje-Servicio entre alumnos de la Facultad de Derecho de la UNED y el Ayuntamiento de Madrid. *Rev. Fom. Soc.* 76 (1), 29–68.
- Green, P.M., Johnson, M. (Eds.), 2014. *Crossing Boundaries: Tension and Transformation in International Service-Learning*. Stylus, Sterling.
- Halpern, D.F., 2014. *Thought and Knowledge: An Introduction to Critical Thinking*. Routledge, New York.
- Hartman, E., Kiely, R., Boettcher, C., Friedrichs, J., 2018. *Community-Based Global Education: The Theory and Practice of Ethical Engagement at Home and Abroad*. Stylus, Sterling.
- Hatcher, J.A., Erasmus, M.A., 2008. Service-learning in the United States and South Africa: a comparative analysis informed by John Dewey and Julius Nyerere. *Mich. J. Community Serv. Learn.* 15 (1), 49–61.
- Hatcher, J.A., Bringle, R.G., Muthiah, R., 2004. Designing effective reflection: what matters to service learning? *Mich. J. Community Serv. Learn.* 11 (1), 38–46.
- Hatcher, J.A., Bringle, R.G., Hahn, T.W. (Eds.), 2017. *Research on Student Civic Outcomes in Service Learning: Conceptual Frameworks and Methods*. Stylus, Sterling.
- Hemer, K.M., Reason, R.D., 2017. Student civic outcomes in higher education. In: Hatcher, J.A., Bringle, R.G., Hahn, T.W. (Eds.), *Research on Student Civic Outcomes in Service Learning: Conceptual Frameworks and Methods*. Stylus, Sterling, pp. 25–43.
- Holland, B.A., 2001. A comprehensive model for assessing service-learning and community-university partnerships. *N. Dir. High. Educ.* 114, 51–60.
- Holsapple, M.A., 2012. Service-learning and student diversity outcomes: existing evidence and future directions. *Mich. J. Community Serv. Learn.* 18 (2), 5–18.
- Ilic, B.C., Gregorova, A.B., Rusu, A.S., 2021. Service-learning as a novelty experience at central and eastern European universities: students' narratives of satisfaction and premises of change. *Front. Educ.* 6 (606356), 1–13.
- Jameson, J.K., Clayton, P.H., Bringle, R.G., 2008. Investigating student learning within and across linked service-learning courses. In: Bowden, M., Billig, S.A., Holland, B. (Eds.), *Scholarship for Sustaining Service-Learning and Civic Engagement*. Information Age, Greenwich, pp. 3–27.
- Jameson, J.K., Clayton, P.H., Jaeger, A.J., 2011. Community-engaged scholarship through mutually transformative partnerships. In: Harter, L.M., Hamel-Lambert, J., Milleisen, J.L. (Eds.), *Participatory Partnerships for Social Action and Research*. Kendall Hunt, Dubuque, pp. 259–278.
- Jameson, J., Clayton, P.H., Jaeger, A., 2012. Investigating faculty learning about community engaged scholarship. *Mich. J. Community Serv. Learn.* 18 (1), 40–55.
- Jameson, J.K., Clayton, P.H., Ash, S.L., 2013. Conceptualizing, assessing, and investigating academic learning in service learning. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning*. Vol. 2A: Students and Faculty. Stylus, Sterling, pp. 85–110.
- Janke, E.M., Clayton, P.H., 2015. Transitioning Toward Partnership Impact—Transitioning Through Partnership Approaches. *IUPUI Center for Service and Learning Newsletter*, 4(3). Indiana University Purdue University Indianapolis, Indianapolis.
- Janke, E.M., 2013. Organizational partnerships in service learning: advancing theory-based research. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning: Conceptual Frameworks and Assessment*. Vol. 2B: Communities, Institutions, and Partnerships. Stylus, Sterling, pp. 573–598.
- Kecskes, K., 2013. The engaged department: research, theory, and transformation of the academic unit. In: Clayton, P.H., Bringle, R.G., Hatcher, J.A. (Eds.), *Research on Service Learning*. Vol. 2B: Communities, Institutions, and Partnerships. Stylus, Sterling, pp. 471–503.
- Kerrigan, S., 2005. College graduates' perspectives on the effect of capstone service-learning courses. In: Martinez, M., Pasque, P.A., Bowman, N. (Eds.), *Multidisciplinary Perspectives on Higher Education for the Public Good*. National Forum on Higher Education for the Public Good, Ann Arbor, pp. 49–65.
- Kniffin, L., Camo-Biogradlija, J., Price, M.F., Kohl, E., Williams, J., Dickovick, A.D.C., Goodwin, J., Johnson, K.V., Clayton, P.H., Bringle, R.G., 2020. Relationships and partnerships in community-campus engagement: evolving inquiry and practice. *Int. J. Res. Serv. Learn. Community Engagem.* 8 (1), 1–15.
- Kolb, D.A., 2015. *Experiential Learning: Experience as the Source of Learning and Development*. Pearson, New York.

- Kuh, G.D., 2008. High-Impact Educational Practices. Association of American Colleges and Universities, Washington, DC.
- Levkoe, C.Z., Stack-Cutler, H., 2018. Brokering community-campus partnerships: an analytical framework. *Gatew. Int. J. Community Res. Engagem.* 11 (1), 18–35.
- Lloyd, K., Bilous, R., Clark, L., Hammersley, L., Barker, M., Coffey, E., Rawlings-Sanaei, F., 2017. Exploring the reciprocal benefits of community-university engagement through PACE. In: Sachs, J., Clark, L. (Eds.), *Learning Through Community Engagement: Vision and Practice in Higher Education*. Springer, Singapore, pp. 245–261.
- Lockeman, K., Pelco, L., 2013. The relationship between service learning and degree completion. *Mich. J. Community Serv. Learn.* 20 (1), 18–30.
- Ma, C.H.K., Mun, S.L.S., 2019. Faculty experience of service-learning pedagogy at a Hong Kong university. *J. High. Educ. Outreach Engagem.* 23 (3), 37–53.
- Ma, C.H.K., Wing-fung, C.C., Cheung-ming, A.C., 2016. The long-term impact of service-learning on graduates' civic engagement and career exploration in Hong Kong. *J. High. Educ. Outreach Engagem.* 20 (4), 37–55.
- Ma, C.H.K., Chan, A.C.M., Liu, A.C., Mak, F.M.F. (Eds.), 2018. *Service-Learning as a New Paradigm in Higher Education of China*. Michigan State University Press, Lansing.
- Ma, C.H.K., Fung, C.W., Hoi, I.T.P., 2019. A common measurement for service-learning in Hong Kong. *J. High. Educ. Outreach Engagem.* 23 (3), 3–19.
- Ma, C.H.K., 2018. Service-learning development in higher education in Hong Kong. In: *Hong Kong in 2017: Two Decades of Post-1997 Hong Kong Developments*. World Scientific Publishing, Hackensack, pp. 43–61.
- MacGregor, J., 2003. *Integrating Learning Communities With Service-Learning*. The Washington Center, Olympia.
- Marchese, T.J., 1997. The new conversation about learning: insights from neuroscience and anthropology, cognitive studies and work-place studies. In: *American Association for Higher Education. Assessing Impact: Evidence and Action*. American Association for Higher Education, Washington, DC, pp. 79–95.
- M Markus, G.B., Howard, J.P.F., King, D.C., 1993. Integrating community service and classroom instruction enhances learning: results from an experiment. *Educ. Eval. Pol. Anal.* 15, 410–419.
- McIlrath, L., Lyons, A., Munck, R. (Eds.), 2012. *Higher Education and Civic Engagement: Comparative Perspectives*. Palgrave, New York.
- McIlrath, L., Aramburuzabala, P., Opazo, H., 2019. Introduction. In: Aramburuzabala, P., McIlrath, L., Opazo, H. (Eds.), *Embedding Service Learning in European Higher Education*. Routledge, New York, pp. 1–9.
- McIlrath, L., Broderick, C., Naughton, M.M., Kelly, M., 2021. The Irish Carnegie Community Engagement Classification Pilot: a critical analysis on culture and context from a community of practice approach. *Gatew. Int. J. Community Res. Engagem.* 14 (1), 1–14.
- Messer, W.B., Kecskes, K., 2008. An anatomy of a community-university partnership: the structure of community collaboration. *J. High. Educ. Outreach Engagem.* 12 (3), 191–204.
- Miller, K.K., Yen, S., 2005. Group differences in academic achievement: service learning in a child psychology course. *Psychol. Teach.* 32, 56–58.
- Miller-Young, J., Dean, Y., Rathburn, M., Pettit, J., Underwood, M., Gleeson, J., Lexier, R., Calvert, V., Clayton, P.H., 2015. Decoding ourselves: an inquiry into faculty learning about reciprocity in service-learning. *Mich. J. Community Serv. Learn.* 22, 32–47.
- Mitchell, T.D., 2008. Traditional vs. critical service-learning: engaging the literature to differentiate two models. *Mich. J. Community Serv. Learn.* 14 (2), 50–65.
- Moely, B.E., Ilustre, V., 2014. The impact of service-learning course characteristics on university students' learning outcomes. *Mich. J. Community Serv. Learn.* 21 (1), 5–16.
- Moely, B.E., McFarland, Miron, D., Mercer, S., Ilustre, V., 2002. Changes in college students' attitudes and intentions for civic involvement as a function of service-learning experiences. *Mich. J. Community Serv. Learn.* 9, 18–26.
- Muse, S., 2018. *Exploring the Community Impact of Community-University Partnerships*. Doctoral dissertation. University of Denver. <https://digitalcommons.du.edu/etd/1415/>.
- Novak, J.M., Markey, V., Allen, M., 2007. Evaluating cognitive outcomes of service learning in higher education: a meta-analysis. *Commun. Res. Rep.* 24 (2), 149–157.
- Osman, A., Ladhani, S., Findlater, E., McKay, V., 2017. *Curriculum Framework for the Sustainable Development Goals*. Commonwealth Secretariat, London.
- Pace, D., Middendorf, J. (Eds.), 2004. *Decoding the Disciplines: Helping Students Learn Disciplinary Ways of Thinking*. Jossey-Bass, San Francisco.
- Pacheco, T., 2019. Service-learning: an innovative approach to education in Africa. In: Mugo, J.K., Namubiru-Ssentamu, P., Njihia, M. (Eds.), *The Good Education and Africa's Future: Concepts, Issues and Options*. Pauline's Publications Africa, Nairobi, pp. 232–259.
- Patrick, C., Valencia-Forrester, F., Backhaus, B., McGregor, R., Cain, G., Lloyd, K., 2019. The state of service-learning in Australia. *J. High. Educ. Outreach Engagem.* 23 (3), 185–198.
- Paul, W.L., Elder, L., 2001. *The Miniature Guide to Critical Thinking*. Foundation for Critical Thinking, Tomales. www.criticalthinking.org.
- Preradovic, N.M., Stark, W., 2019. Identified service learning practices in European higher education. In: Aramburuzabala, P., McIlrath, L., Opazo, H. (Eds.), *Embedding Service Learning in European Higher Education*. Routledge, New York, pp. 109–131.
- Price, M.F., 2016. Collaborative Relationship Mapping (CoRM): An Overview. Center for Service and Learning, Indiana University-Purdue University Indianapolis, Indianapolis. <https://csl.iupui.edu/teaching-research/curriculum/community-planning/coirm/index.html>.
- Reeb, R.N., Sammon, J.A., Isackson, N.L., 1999. Clinical application of the service-learning model in psychology: evidence of educational and clinical benefits. *Prev. Interv. Community* 18, 65–82.
- Reed, S.C., Rosenberg, H., Staham, A., Rosing, H., 2015. The effect of community service learning on undergraduate persistence in three institutional contexts. *Mich. J. Community Serv. Learn.* 21 (2), 22–36.
- Regina, C., Ferrara, C., 2017. *Service-Learning in Central and Eastern Europe: Handbook for Engaged Teachers and Students*. Centro Latinoamericano de Aprendizaje y Servicio Solidario, Ciudad de Buenos Aires.
- Regueiro, B., Rodriguez-Fernandez, J.E., Crespo, J., Pino-Juste, M.R., 2021. Design and validation of a questionnaire for university students' generic competencies (COMGAU). *Front. Educ.* 6 (606216), 1–12.
- Rice, D., Stacey, K., 1997. Small group dynamics as a catalyst for change: a faculty development model for academic service-learning. *Mich. J. Community Serv. Learn.* 4, 64–71.
- Sachs, J., Clark, L. (Eds.), 2017. *Learning Through Community Engagement: Vision and Practice in Higher Education*. Springer, New York.
- Saltmarsh, J., Zlotkowski, E., 2011. *Higher Education and Democracy: Essays on Service-Learning and Civic Engagement*. Temple University Press, Philadelphia.
- Saltmarsh, J., Hartley, M., Clayton, P.H., 2009. *Democratic Engagement White Paper*. New England Resource Center for Higher Education, Boston.
- Saltmarsh, J., Janke, E.M., Clayton, P.H., 2015. Transforming higher education through and for democratic civic engagement: a model for change. *Mich. J. Community Serv. Learn.* 22, 122–127.
- Saltmarsh, J., 2010. Changing pedagogies. In: Fitzgerald, H.E., Burack, C., Seifer, S.D. (Eds.), *Handbook of Engaged Scholarship: Contemporary Landscapes, Future Directions: Volume 1: Institutional Change*. Michigan State University Press, Lansing, pp. 331–352.
- Sandy, M., Holland, B.A., 2006. Different worlds and common ground: community partner perspectives on campus-community partnerships. *Mich. J. Community Serv. Learn.* 13 (1), 30–43.
- Snell, R.S., Lau, K.H., 2020. The development of a Service-Learning Outcomes Measure Scale (S-LOMS). *Metrop. Univ.* 31 (1), 44–77.
- Sotelino-Losada, A., Arbués-Radigales, E., García-Docampo, L., González-Geraldo, J.L., 2021. Service-learning in Europe. Dimensions and understanding from Academic Publication. *Front. Educ.* 6 (604825), 1–13.
- Stanton, T.K., Erasmus, M.A., 2013. Inside out, outside in: a comparative analysis of service-learning's development in the United States and South Africa. *J. High. Educ. Outreach Engagem.* 17 (1), 61–94.
- Steinberg, K., Hatcher, J.A., Bringle, R.G., 2011. A North star: civic-minded graduate. *Mich. J. Community Serv. Learn.* 17 (1), 19–33.
- Studer, M., Benton, M., Rogers, C., Quirke, M., 2017. We are all teachers: modeling democratic engagement in faculty development. *Mich. J. Community Serv. Learn.* 23, 150–153.
- Sturm, S., Eatman, T., Saltmarsh, J., Bush, A., 2011. *Full Participation: Building the Architecture for Diversity and Public Engagement in Higher Education (White paper)*. Columbia University Law School, Center for Institutional and Social Change, New York.

- Tapia, M.N., 2012. Academic excellence and community engagement: reflections on the Latin American experience. In: McIlrath, L., Lyons, A., Munck, R. (Eds.), *Higher Education and Civic Engagement: Comparative Perspectives*. Palgrave, New York, pp. 187–203.
- Thomson, A.M., Smith-Tolken, A., Naidoo, A.V., Bringle, R.G., 2011. Service learning and civic engagement: a comparison of three National Contexts. *Voluntas* 22, 214–237.
- Ti, C., Tang, J., Bringle, R.G., 2021. Initiating and extending institutionalization of service-learning. *J. High. Educ. Outreach Engagem.* 25 (2), 5–22.
- Tijmsma, G., Urias, E., Zweekhorst, M., 2021. A thematic approach to realize multidisciplinary community service-learning education to address complex societal problems: a-win-win situation? *Front. Educ.* 5, 617280.
- UNESCO, 2009. 2009 World Conference on Higher Education: The New Dynamics of Higher Education and Research for Societal Change and Development. UNESCO, Paris.
- UNESCO, 2015. *Global Citizenship Education: Topics and Learning Objectives*. UNESCO, Paris.
- Waldner, L.S., McGorry, S.Y., Widener, M.C., 2012. E-service-learning: the evolution of service-learning to engage a growing online student population. *J. High. Educ. Outreach Engagem.* 16 (2), 123–150.
- Warren, J.L., 2012. Does service-learning increase student learning? A meta-analysis. *Mich. J. Community Serv. Learn.* 18 (2), 56–61.
- Welch, M., Plaxton-Moore, S., 2017. Faculty development for advancing community engagement in higher education: current trends and future directions. *J. High. Educ. Outreach Engagem.* 21, 131–166.
- Welch, M., Saltmarsh, J., 2013. Current practice and infrastructures for campus centers of community engagement. *J. High. Educ. Outreach Engagem.* 17 (4), 25–55.
- Xing, J., Ma, C. (Eds.), 2010. *Service-Learning in Asia: Curricular Models and Practices*. Hong Kong University Press, Hong Kong.
- Yorio, P.L., Ye, F., 2012. A meta-analysis on the effects of service-learning on the social, personal, and cognitive outcomes of learning. *Acad. Manag. Learn. Educ.* 11 (1), 9–27.
- Young, C.A., Shinnar, R.S., Ackerman, R.L., Carruthers, C.P., Young, D.A., 2007. Implementing and sustaining service-learning at the institutional level. *J. Exp. Educ.* 29 (3), 344–365.
- Zlotkowski, E., Williams, J., Longo, N. (Eds.), 2006. *Students as Colleagues: Expanding the Circle of Service-Learning Leadership*. Campus Compact and Bolton: Anker, Providence.
- Zlotkowski, E., 1995. Does service-learning have a future? *Mich. J. Community Serv. Learn.* 2 (1), 123–133.
- Zlotkowski, E. (Ed.), 1998. *Successful Service-Learning Programs: New Models of Excellence in Higher Education*. Anker, Boston.

Relevant websites

- Ma'an Arab University Alliance for Civic Engagement, <https://www.facebook.com/MaanAlliance/>.
- Campus Compact, <https://compact.org/>.
- Canadian Alliance for Community Service-Learning, <https://mcconnellfoundation.ca/grant/canadian-alliance-for-community-service-learning-cacsl/>.
- Community-Campus Partnerships for Health, <https://ccphealth.org/>.
- European Observatory of Service Learning in Higher Education, <https://www.eoslhe.eu/>.
- Imagining America: Artists and Scholars in Public Life, <https://imaginingamerica.org/>.
- Service-Learning Asia Network, <https://su.edu.ph/slan/>.
- South African Higher Education Community Engagement Forum, <http://www.sahecef.ac.za/>.

Thirty years of competency-based training: how Australia painted itself into a curriculum corner in vocational education and training

Erica Smith, School of Education, Federation University Australia, Ballarat, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	491
Background on the Australian vocational education and training system	491
Method and evidence base for this chapter	492
The first decade of CBT	493
The second decade of CBT	494
The third decade of CBT	495
What is the current nature of CBT curriculum?	496
In what ways has the system painted itself into a corner?	497
Getting out of the corner: how can this situation be improved?	498
Conclusion	499
Appendix 1. Extracts from two units of competency (elements and performance criteria only) and assessment requirements (knowledge evidence only)	499
Unit AURTTJ003—"Remove and replace wheel and tyre assemblies"	499
Unit HLTINF001—"Comply with infection prevention and control policies and procedures"	500
References	502

Introduction

In this chapter, 30 years of competency-based training in Australia are recorded and analyzed. The paper updates two previous accounts by the author, of 10 and 20 years of competency-based training (CBT) by presenting a new analysis of the third decade of CBT. The paper analyses the major curriculum flaws in the system, and suggests modest improvements.

The chapter takes a specific and narrow approach to the meaning of the term CBT. It refers to the Australian/British meaning of the term (Smith, 2010) where VET systems are primarily based on industry derived national competency standards which apply nationally. In Australia these standards are contained within artifacts known as Training Packages. Australia is perhaps the most extreme exponent of this system, as in the UK system, other curriculum alternatives are available besides what are known there as National Vocational Qualifications. The Australian VET system has been competency-based since the mid-1980s, and as in many countries, competency-based training as a concept has been heavily critiqued on a number of fronts; Smith (2010) describes them as "philosophical," educational and practical arguments.

As was noted in Smith (2010) there are other understandings of the actual term "competence," which are generally broader, both in national systems and in academic debates. Even in Australia, early debates around competency-based training included discussions of more general competencies and extended beyond VET (e.g. Gonczi, 1994). These discussions no longer occur, and hence when CBT is discussed in Australia, it means the system as it currently stands.

One reason why Australia might be of particular interest internationally, apart from being an extreme example of a particular type of CBT, is that its elements of its system have been "exported" to other countries, especially in Asia (e.g. Sung, 2011, discussing Singapore). In these exports, CBT often contributes to formal national qualification structures, as in, for example, India's National Skills Qualification Framework (<https://nsda.gov.in/>). More broadly, occupational standards linked to national VET qualifications frameworks are common to several developed countries (Allais et al., 2014).

Background on the Australian vocational education and training system

Australia has a well-developed VET system. It is overseen by the national ("Commonwealth") government but the lower tier of government—eight States and two Territories—provide funding for VET courses, and employ VET teachers in the public system, TAFE (Technical and Further Education). In Australia, VET is usually understood to have two main purposes. One is to improve people's skills by training them before they enter the workforce, or by upskilling them after they enter the workforce, thus lifting productivity, reducing accidents, and making movement of workers among industries and occupations easier. The other purpose is to provide social benefit, improving individual people's life chances, confidence, and employment prospects. Currently, the benefits of VET in Australia tend to be expressed in economic rather than social terms; hence industry and employers feel they have ownership of VET. This is an important contributor to the nature of CBT in Australia.

VET has qualifications covering most industry areas, for occupations that are not regarded as needing university preparation. The Australian Qualifications Framework (www.aqf.edu.au) ranges from AQF 1 (pre-employment courses) to AQF 10 (university

doctorates). With some minor exceptions, VET covers AQF levels 1 to 6. The first four levels are Certificate I to Certificate IV, Level 5 is Diploma and Level 6 is Advanced Diploma. Most VET students are in Certificate II to Certificate IV qualifications; in 2019, almost half (47.2%) of students were in Certificate III programs (National Centre for Vocational Education Research [NCVER], 2020). Certificate III is the level of most apprenticeship programs and is regarded as the qualification for “skilled” non-supervisory work.

Most qualifications are contained in “Training Packages,” based around industry and/or occupational areas, all of which have large numbers of units of competency gathered into multiple qualifications. There are also a small number of other “accredited courses,” usually accredited by individual States. Training Packages were introduced initially in 1997 and are explained in more detail throughout this chapter. Many occupations did not have formal VET qualifications until the advent of Training Packages (Smith and Keating, 2003). The six Training Packages with the most government-funded enrollments pre-COVID were Community Services (which includes child-care training); Construction; Business; Tourism and Hospitality; Electrotechnology; and Health (NCVER, 2019, reporting 2018 Figures¹).

Training providers are known as “Registered Training Organizations” (RTOs). There is a public provider, TAFE (Technical and Further Education), and around 5000 other training providers, many of which are small and in niche areas, and most of which are private and for-profit. Often the term “RTO” is used to refer only to non-TAFE training providers. Marketization of the VET system (i.e. access by private providers to government funding for their students) has proceeded steadily during the past 30 years. This process has been assisted by competency-based training, as national competency standards have removed the necessity for providers to develop their own curriculum from first principles.

In 2019, there were just over one million students—1,011,000—enrolled in government-funded VET (NCVER, 2020), just over half in TAFE, from an Australian population of just over 25 million. In addition, some VET is delivered on an entirely fee-paying basis, including to overseas students in Australia, and, to a lesser extent, in other countries. Students were evenly distributed across age ranges, with the largest concentration in the group 25–44 years (NCVER, 2020). The number of government-funded students in VET has declined over the past 8 years, from a high of 1.54 million in 2012 (NCVER, 2019), so that there are now fewer government-funded VET students than higher education students in Australia. The decline has been partly due to a number of financial scandals, mainly involving misuse of government funding by private RTOs (Smith, 2019), but also due to withdrawal or reduction of funding for training for many occupations.

The financial scandals have led to a heavy focus on regulatory compliance, with strict rules put in place to try to prevent further problems. Regulation of the system has been carried out since 2011 by a national regulatory body for the VET sector, known as the Australian Skills Quality Authority (ASQA) (www.asqa.gov.au). This body registers and re-accredits RTOs, as well as carrying out special audits in qualifications known to be “at risk.” Prior to 2011, regulation was carried out by the State and Territory governments.

The web site www.training.gov.au enables all curriculum elements of the national competency-based system to be viewed by anybody. Later in this chapter, the section “the current nature of CBT curriculum” provides more details and refers the reader to **Appendix 1** which contains some examples.

Method and evidence base for this chapter

The chapter seeks to answer three research questions.

1. How has the Australian system of CBT evolved over the past 30 years?
2. In what ways has it painted itself into a corner?
3. How can this situation be improved?

As explained in Smith (1999), competency-based training as a curriculum approach was not unknown in Australia prior to the end of the 1980s in some industry areas and some training providers, but its national hegemony dates from the end of that decade. Hence the paper takes 1990 as the beginning of CBT as it exists today.

The chapter builds upon the two accounts (Smith, 1999, 2010) which reported and analyzed developments in the first ten and then the first 20 years of CBT as we currently know it, augmented by an analysis of developments since 2010. The earlier accounts and the current third decade update, as well as referring to the scholarly literature, are based on a number of sources. They draw on publicly available documents, including government announcements and consultation papers, about changes in the system. They also draw on the author’s understandings from leading several major research projects specifically on CBT, and on industry training and Vocational Education and Training (VET) more generally. Perhaps most importantly, the accounts are strengthened by the author’s engagement in and with the VET system as follows:

- Working in the development and review of Training Packages, in different capacities over a 20 year period:
- For 2 years as a manager of a State-based Industry Training Advisory Board, managing the development and maintenance of seven Training Packages in the retail and personal services area, and
- For substantial periods of time, as a member of relevant national committees overseeing the development, review and maintenance of the qualification for VET teachers and trainers.

¹These 2018 data are the latest figures which are readily available in the public domain. NCVER no longer publishes analyses of the data by Training Package.

- Constant dialog, as a policy-engaged researcher, with VET government officials, industry and stakeholder groups, in Australia and also in Europe and Asia;
- Teaching hundreds of practicing VET teachers over most of the period 1993–2020 in VET teacher-education programs, at three universities;
- Experience of embedding CBT qualifications within those VET teacher-education programs for many years, and thus constructing curriculum and assessment tasks according to, and compliant with, required Training Package “rules.”

Accounts of relevant developments in each of the three decades of CBT are now provided, each section beginning with an overview followed by more detail on major events and developments. The current system is then explained in more detail with specific reference to national “curriculum” documentation, i.e. Training Package components. An analysis follows of the ways in which the developments described, and the current system, have forced Australian VET into a “curriculum corner,” and how it could begin to emerge.

The phrase “painting oneself into a corner” means to create a situation from which extrication is very difficult. It implies a progressive “boxing in,” and also implies that the situation is one which a person or people had created for themselves. While phrases may differ from country to country, the concept is universal. An equivalent phrase in Spanish, for example is “Cuando uno se mete en un callejón sin salida”—turning into an alley from which there is no exit. Other English phrases are similar to “painting oneself into a corner,” such as “digging oneself into a hole” but the latter implies more finality; “painting oneself into a corner” offers a greater possibility of extrication, albeit messy.

The first decade of CBT

Overview: The “10-year” account (Smith, 1999) showed that the main features of the first decade of CBT—the 1990s—were the establishment of a national system, the penetration of CBT into all parts of formal VET, and, toward the end of the decade, the initial introduction of Training Packages designed to create broader coverage of the system and to remove accreditation of curriculum. From 1992 the system was overseen by the Australian National Training Authority, a statutory body set up to implement what was known as the Training Reform Agenda (Smith and Keating, 2003) which included aspects other than curriculum, including marketization of VET.

The current system has its basis in an agreement in the late 1980s between the Commonwealth (national) government and the trade union movement that the VET system needed to be reformed. The government wanted to upskill the workforce to meet perceived international economic competition, and to improved labor market outcomes for young people; and the trade unions were interested in skill-related career paths for workers (Phillimore, 1997). A meeting of Commonwealth (national) and State Ministers in 1989 decided to introduce a competency-based system, establishing a National Training Board to develop industry competency standards. In 1990, a “strategic framework” for CBT implementation was agreed, and a target was set for “substantial implementation” by the end of 1993.

While employers were not pivotal in the establishment of the system, by the early 1990s one of the three main peak employer bodies, the Australian Chamber of Commerce and Industry, was fully on side, thus cementing the “tripartite” (government, labor and industry) nature of the system. ACCI developed a definition of CBT as follows:

A way of approaching (vocational) training that places primary emphasis on what a person can do as a result of training (the outcome), and as such represents a shift away from an emphasis on the process involved in training (the inputs). It is concerned with training to industry specific standards rather than an individual’s achievement relative to others in the group.

Australian Chamber of Commerce and Industry, 1992; in Smith and Keating (2003: 102).

ACCI is now one of the major proponents of the system, along with the manufacturing and metal workers’ union, which was a major driver of competency-based training in the late 1980.

An early government-commissioned evaluation in 1990 (Thomson et al., 1990) found that only about 13% of courses in TAFE were competency-based. A further government-commissioned evaluation, which was led by the author (Smith et al., 1996) was undertaken in 1994, covering non-TAFE training providers as well. By this time implementation had spread, but still only 29% of TAFE courses and 39% of non-TAFE courses were found to have CBT features. One common discovery in both evaluation projects was an inability of stakeholders to agree on what CBT actually meant. In the Thomson et al. project a course was classified as CBT if it had any three of seven competency-based attributes; in the Smith et al. project, a categorization of “CBT-ness” was developed, with courses allocated to one of four bands: non-CBT, SC, SCIR and SCIRA. The bands were defined as follows (Smith et al., 1996: 64).

- **Band SC**—courses which were based on competency standards (S) or with formal industry consultation where there are no standards and where the course is written in CBT format (C). This band was taken as the “minimum” definition of CBT.
- **Band SCIR**—courses which were defined as meeting the minimum definition of CBT, and also involved industry in course monitoring (I) and operate recognition of prior learning (RPL) procedures (R).

- *Band SCIRA*—courses which, in addition to Band SCIR characteristics, had assessment available on demand and at least partly in the workplace (A). These were considered at that time to be the main attributes of competency-based assessment

Only 2% of TAFE courses and 11% of non-TAFE courses were in the SCIRA band. Adoption of the assessment features, which were not at that time part of government policy, were posited by the authors (Smith et al., 1996) to depend on a number of factors: institutional, industry area, the nature of the students, and the nature of the course. These decisions might, and could, be made at the point of writing of the curriculum document, the State level, the individual training provider, or by the teacher, and decisions could be conscious or could occur by default. The study concluded that some features would never be universally appropriate; for example, some students might not have access to a workplace and therefore assessment could not occur in the workplace.

The National Framework for the Recognition of Training (NFROT) was introduced in 1992 to be used by State accreditation bodies. Courses accredited under NFROT could then be nationally recognized. The courses were required, for accreditation, to present curriculum, including “learning outcomes” derived from the National Training Board competency standards; the “learning outcomes” guided assessment. Smith et al. (1996) found that courses that had been through the NFROT process were more likely to have more features of CBT.

A major government-commissioned review of the system (Allen Consulting, 1994) led to a restructuring of the CBT system and its associated structures. In 1997, NFROT was abolished and there was no longer any scrutiny of curriculum. Instead, training providers—now to be known as “Registered Training Organizations”—would be accredited to deliver “Training Packages.” Training Packages were to contain only competency standards, packaged into qualifications, together with assessment guidelines. The Training Package system began with the establishment of industry competency standards bodies (known as Industry Training Advisory Bodies), with government funding. Initially they were still assisted by the National Training Board.

CBT was contested from the beginning. As noted earlier, Australian arguments could be grouped into “philosophical,” educational and practical (Smith and Keating, 2003). West (2004), writing about the UK described similar groups of early arguments as “technical and moral.” In Australia, as in UK also (Burke, 1989) arguments were mounted around teacher capability and/or engagement. Teachers tended to be suspicious of CBT and unsure how to implement it (Robinson, 1993; Cornford, 1996), and little staff development was received by teachers.

The second decade of CBT

Overview: In this second decade, the 2000s, Smith (2010) shows that the Training Package system in Australia, which has been described as “second generation CBT” (Barratt-Pugh and Souter, 2002), became firmly established. The term “nationally-recognized training” came into common use, referring to qualification-based training based on Training Packages. The industry competency standards development bodies were replaced by Industry Skills Councils (Australian National Training Authority, 2004), but the system of Training Package development remained similar. At the time of the 20-year review by Smith (2010) there were 80 Training Packages, some of which were company-specific.

There were several key events during this decade. A major review of Training Packages was commissioned by ANTA, in 2003 (Schofield and McDonald, 2004). The review made some design recommendations and also advocated for better staff development for teachers so that they could deliver Training Packages more confidently. However, ANTA was disestablished in 2005, with some, but not all, of its functions moved to the relevant department in the national government, and the recommendations withered without a clear driver. There was a subsequent national project known as “VET Training Products for the 21st century” (Joint Steering Committee of the National Quality Council and the Council of Australian Governments, 2009). The term “training products” referred to the different “components” of CBT (Training Packages, qualifications, units of competency) and this term is still in use today. In the project, different approaches to CBT were considered, but were not supported by the stakeholders who were consulted.

Training Packages, then, seemed to have general support. The main advantages (as listed in Smith, 2010), compared with previous 1990s CBT curriculum models, were considered to be that Training Packages allowed more autonomy for training providers and teachers, that they could be adapted for specific uses such as training in companies, that they did include at least a cursory acknowledgment of what was called “underpinning knowledge”, and that assessment could be holistic, both within units of competency and, theoretically across units of competency. However, arguments against CBT continued to be made, including arguments about the dominance of assessment over learning (Gillis and Griffin, 2005; and, in the UK; Torrance, 2012).

During this second decade, grading in competency-based assessment, which was reasonably common in the early days of CBT and had even grown during the first decade (Smith and Keating, 2003), began to disappear. The problem with grading was that extremely explicit criteria were presented as being necessary for grading above the pass level of “competent” (Gillis and Griffin, 2005), a seemingly highly technical task that was beyond the capabilities of most RTOs.

During the 2000s, changes occurred to the bodies which developed Training Packages. The network of national and State ITABs collapsed first, as the Commonwealth government removed funding in 2002 from State ITABs, which had been co-funded by national and State governments. Then in 2003, national ITABs were replaced by Industry Skills Councils. There were only 11 of these, replacing 29 national ITABs in existence in 2003 (Smith and Keating, 2003). As the Industry Skills Councils (ISCs) necessarily covered a range of industries, each ISC established a number of industry sub-committees. The role of the ISCs was to develop, maintain and review national Training Packages, and to undertake a number of ancillary activities, including the development of what

were known as “Training Package support materials” which typically included teacher guides to delivering the material, and learner workbooks.

Initially, training providers were excluded from the governing bodies of VET and from Training Package development processes, with those responsible claiming that industry should own all processes, but this decision was reversed during the decade, with training providers allowed to participate in important committees and to comment on Training Packages. The Boards of ISCs included employer and trade union representatives, who were sometimes at odds over the content and format of Training Packages and qualifications, with trade unions, for example, seeking to protect industrial relations privileges including preferential arrangements for apprenticeships (Smith, 2010). Even after Training Packages had passed this barrier, further steps needed to be followed. Firstly, all State and Territory governments had to be consulted and then the new or revised Training Package was submitted to a national committee.

One feature of the 2000s was the introduction of skill sets (Mills et al., 2012). Since the beginning of the Training Package system, formal certificates called “statements of attainment” could always be issued by training providers for individual units of competency or groups of units. In this decade, the name of “skill sets” was conferred upon groups of units. Skill sets could be designated either by training providers or, more formally, by Industry Skills Councils along with Training Packages (Mills et al., 2012). Essentially these “bundles” of units of competency—what might now be called micro-credentials—were intended to meet specific industry or worker needs.

The third decade of CBT

Overview: In this third decade, there were major CBT-related policy developments: Industry Skills Councils were supplanted, in a battle over industry control of the process of developing competency standards; and a new regulatory body was introduced. In curriculum developments, the format of units of competency was changed, a decision whose ramifications were broader than might be expected as will be explained later; and there were changes to the way competency-based assessment was policed.

Industry Skills Councils (ISCs) were abolished after only 12 years of existence, and replaced in 2015 by a new model where Training Packages were developed by “Industry Reference Committees” which were “serviced” by “Skills Service Organizations” (SSOs). For example, the author is a member of the Industry Reference Committee for the Training Package for VET teachers, the Training and Education Training Package. While some former ISCs gained contracts to become SSOs, others did not. The reason for this change was not publicly stated, but a presentation by a senior official of an industry peak body, ACCI, at an international conference (Lambert, 2016) provides some insight. Lambert stated that some ISCs had become “self-interested” and had become “masters rather than servants of industry.” As stated earlier, ACCI was a proponent of CBT from the early 1990s. The new system was quickly found to be flawed (Joyce, 2019) and SSO contracts were only renewed until the end of 2021.

During this third decade, the 2010s, ASQA, the national regulatory body for VET, was established. As explained earlier, ASQA was set up in 2011 to provide a national auditing function for RTOs. “RTO Standards” were instituted for this purpose, with a revised set of standards in 2015; RTOs are obliged to adhere to the standards to gain or maintain registration. These standards do not contain direct reference to Training Packages, but there are guidelines for each standard which are quite prescriptive about, for example, what RTOs’ curriculum documents derived from Training Package qualifications (known as “Training and Assessment Strategies”) should look like. In these guidelines, Training Packages are explicitly referenced (e.g. <https://www.asqa.gov.au/standards/training-assessment/clauses-1.1-to-1.4-2>). ASQA auditors also expect to see “mapping documents”, which show how assessment processes address every item mentioned in the unit of competency and its associated “assessment requirements” document.

In the 2010s, Training Packages, and particularly units of competency, changed their appearance and content. In response to complaints about the length and the jargon-ridden nature of the units, during the first half of the decade they were “simplified,” which perhaps ironically led to a lengthy bureaucratic process as units were progressively “simplified.” A consequence of this change was that “assessment requirements” were decoupled from the units. This was reported to have been undertaken because some employers might want to deliver the content contained in the units, but not assess. An unfortunate result was that the “knowledge” section of units of competency, always skimpy, was removed from the unit and placed into the “assessment requirements” document. During this decade, commonly-occurring units of competency known as “Working effectively in the [e.g. retail] industry”, including knowledge about the industry as a whole, were also removed from Training Package qualifications, as, it was argued, they did not describe a work role. This further reduced the “knowledge” contained in Training Packages. Finally there was a successful attempt in the 2010s to reduce the number of Training Packages, as some were rarely utilized.

Assessment was identified during this third decade as a major problem with the VET system (e.g. Department of Education & Training [DET], 2016). Industry reportedly complained of a lack of confidence in assessment, and AQSA, the regulator, found assessment problems in key industry areas through strategic audits. Attempts were made in the 2015 RTO standards (Department of Education and Training, 2017) to tighten up the definition of assessment, even incorporating in the legislated standards some difficult-to-interpret “rules” and “principles” proposed by assessment “experts.” The perceived problem was so bad that “external assessments” were mooted as one possible “solution” to the problem, i.e. removing from RTOs the ability to assess their own students (DET, 2016). While this was not introduced, other measures were. During this decade immensely complicated “assessment tools” (ASQA, 2015) became the required norm in VET; these documents, often 20 pages or more, included minutiae of how to assess the units, including all the tasks, instructions for teachers, arrangements for keeping all the “evidence” for many years,

and so on. The 2015 standards also introduced a requirement for all RTOs to undertake an “ongoing systematic validation of assessment practices and judgments”; every 5 years all assessment tools had to be “validated,” and moderation of results undertaken. Again this process was highly complex, tightly specified within the standards, and imperfectly understood in the VET sector. Finally, assessment decision-making was identified as a major problem in the sector during this decade (Clayton, 2013), with teachers routinely reporting difficulty in deciding whether to pass students or not.

What is the current nature of CBT curriculum?

After these three decades, what does competency-based training look like in Australia? In 2020 there were 57 Training Packages (www.training.gov.au, viewed 27.9.20). These contained 1047 qualifications, 16,784 units of competency. 1439 formal skill sets had been endorsed. They are common in certain industries and for certain cohorts; for example 11 of them are in community Aboriginal health. The parallel system of accredited courses had generated 680 qualifications. Rules called the Standards for Training Packages (Department of Industry, 2012) govern Training Packages (<https://docs.employment.gov.au/documents/standards-training-packages>).

Each Training Package on the “National Register of VET” contains three components. This is explained on the web site of the Australian Skills Quality Authority (ASQA) thus:

- “Units of competency define the skills and knowledge needed, and how to apply them in a workplace context.
- A qualifications framework contains groups of units of competency used to develop learning outcomes. These groupings range from Certificate I to Graduate Diploma level.
- Assessment guidelines cover the qualifications required by assessors, the design of assessment processes and guidelines for assessment management. Assessment guidelines explain the industry’s preferred approach to assessment” (<https://www.asqa.gov.au/about/vet-sector/training-packages>).

As explained above, during the third decade of CBT, units of competency were divided into the units themselves and separate documents called “assessment requirements” for each unit. The content of both documents is shown on the web page for each unit of competency. The content and layout of Training Packages are themselves governed by their own national standards (Department of Education Skills & Employment, 2012). Units of competency have some sections of content which are compulsory, and a few which are non-compulsory. Non-compulsory sections include a listing of pre-requisite units and the “range of conditions,” which describes the context for the task described in the unit. The core of each unit is “elements” and “performance criteria.” According to the Department of Education Skills & Employment (2012),

Elements describe the essential outcomes; elements describe actions or outcomes that are demonstrable and assessable. Performance criteria describe the performance needed to demonstrate achievement of the element. 1.1 Performance criteria clearly relate to the element. 1.2 They are expressed as a standard. 1.3 They specify the required performance in relevant tasks, roles, and skills. 1.4 They reflect the applied knowledge that enables competent performance

In practice, performance criteria generally read more like sub-sections of the elements.

Appendix 1 illustrates the above requirements with extracts from two units of competency: one for changing wheels and tyres and one for infection control. Elements and performance criteria are presented in the Appendix, together with the “knowledge” section from the “assessment requirements” documents for the units. These items essentially represent the content to be taught; Training Packages are not allowed to provide any information about how to deliver the training.

The units selected for the Appendix include a very simple and short unit, “Remove and replace wheel and tyre assemblies” used in low-level automotive qualifications, and a larger unit, “Comply with infection prevention and control policies and procedures”. Both units are quite small. There are no specified hours of delivery, as this is forbidden in the Training Package standards, but units are assigned “nominal hours” by State governments for funding purposes, which are taken as a rough guide to appropriate hours for delivery. The two units in Appendix 2 are quite small, 10 and 25 nominal hours respectively; most units of competency, in fact, have between 40 and 60 “nominal hours”.

Appendix 1 illustrates the odd language that persists in units of competency, despite “simplification” of Training Packages. It will be noted that the elements and performance criteria are written in the passive voice and describe tasks performed in workplaces. They read more like detailed job descriptions—for example “Potential hazards and risks associated with task are identified and reported to workplace supervisor” than curriculum for training. The knowledge requirements sections, with their sections and sub-sections provide only a list of “bits” of knowledge, with no coherent explanation nor a picture of the overall body of knowledge that is to be imparted to students. Knowledge sections vary considerably among units, with some, in fact, containing virtually nothing.

The other vital part of the unit for curriculum development is the “assessment conditions” section in the “assessment requirements” document. This section sometimes provides some clues about delivery, for example about whether there should be a work

placement associated with the unit. However, there is still no guidance about how to determine whether a student is competent or not, despite assessment decision-making having been identified as a problem.

In what ways has the system painted itself into a corner?

From a policy point of view the following statement from a report by one of the key components in the system in the 2010s, the National Skills Standards Council, sums up the developments:

The history of the standards for the regulation of vocational education and training since the early 1990s illustrates how the key features of these initial regulatory and quality frameworks have endured. It also shows that the regulatory system has responded to prevailing policy concerns and perceptions of risk, and moved between being primarily a recognition system with certain prescriptions, to a system adopting an outcomes-based and continuous improvement approach, and reverting back to a system favoring compliance.

National Skills Standards Council [NSSC] (2013: 7).

The NSSC, in this excerpt, summarizes a move away from tight prescription in the 1990s and back again during the early 2010s (and which continued to the end of the 2010s). While the statement refers to policy and regulation, the same pattern can be detected in curriculum. Early, looser perceptions of CBT—“based on industry competency standards or developed with formal industry consultation and written in CBT format”—as described by [Smith et al. \(1996\)](#), quickly became tightened when NFROT began to accredit course curriculum in 1992. But this system lasted for only 5 years. Training Packages removed this accreditation, leaving only industry competency standards as the key component of competency-based training. Discussion of the resultant curriculum needed to deliver training from these standards was actively discouraged, although the need for curriculum was evidenced by the government funding of Training Package support materials in the 2010s, recognizing that many RTOs were not capable of developing curriculum from the industry competency standards.

At the same time, during the 2000s, the average qualification levels of VET teachers began to fall, due to the introduction of a new Training Package Certificate IV qualification for VET teachers. This was designed as a minimum qualification but quickly became the maximum required by RTOs ([Smith and Keating, 2003](#)). As [Hodge \(2014\)](#) points out, teachers lack the skills to work with Training Packages, being unable to interpret units of competency correctly. One commonly reported problem is a perception among teachers that it is forbidden to teach more than the content in the units, thus ensuring that the “taught” or “delivered” curriculum ([Glattthorn, 1987](#)) does not evolve in line with industry developments. The removal of approved curriculum has therefore been problematic. Teachers also worry about the language of “competence.” A CBT qualification does not mean that somebody is competent at work; it means that somebody has passed a course that teaches them certain skills and knowledge. This simple fact needs to be recognized, as it was in the early days of CBT ([Thomson et al., 1996](#)).

A combination of the removal of accreditation for curricula, the degradation of the expertise of the VET teaching workforce (augmented by the national government’s adoption of the term “trainers and assessors” instead of “VET teachers”), together with the financial scandals of the 2010s led to re-imposition of a “system favoring compliance” as the [NSSC \(2013\)](#) stated. The tight specification of assessment processes in the RTO standards of 2015 was a reflection of that. The Standards for RTOs ([DET, 2017](#), p. 8) define assessment as follows:

the process of collecting evidence and making judgments on whether competency has been achieved, to confirm that an individual can perform to the standard required in the workplace, as specified in a training package or VET accredited course.

This definition asks teachers (or “assessors” as they are referred to in this instance) to assert that students (or “candidates” as they are called in assessment jargon) are competent to workplace standards. Of course it is not possible for this assertion to be made, as teachers cannot always accompany students to workplaces (if they have one) and watch them work, even leaving aside the fact that all workplaces have different requirements and standards. Moreover, the skillful task of assessing students is reduced to the task of “collecting evidence.” Very little grading is now practiced in the sector, except for some higher level qualifications; teachers report that it is difficult to encourage high performance when no grading is available; but more worryingly, both students and teachers report that the emphasis on “evidence gathering” means that students often pass who should not.

At the same time as the doomed attempt to ensure assessment quality through tight prescription, there was the assertion of more industry control over curriculum. The “simplification” of Training Packages reduced the emphasis on knowledge and ensured that teachers focused only on skills performance. Industry demands for the abolition of Industry Skills Councils removed the curriculum expertise that had resided in those bodies, replacing them with a system that was intended to reassert industry control. In a move analogous to the introduction of Training Package support materials because curriculum documents were not developed and teachers lacked the skills to develop their own curriculum, the national government introduced a well-hidden “Training Package Quality Assurance Panel” of consultants, after the abolition of Industry Skills Councils, who could be hired to assist Skills Service Organizations in Training Package work. And in fact, it was industry pressure over so-called assessment problems that led to the assessment mis-steps of the 2010s.

Thus while Training Packages were intended to lead to flexibility, the system has become progressively less flexible. As problems began to appear in the system, “reforms” were introduced which painted competency-based training into a smaller and smaller corner of the VET “room”. Moreover, RTOs added extra layers of sticky paint to avert potential audit problems and from a lack of trust of teaching staff, either justified or unjustified. For example, in many large RTOs including TAFE institutes, managers

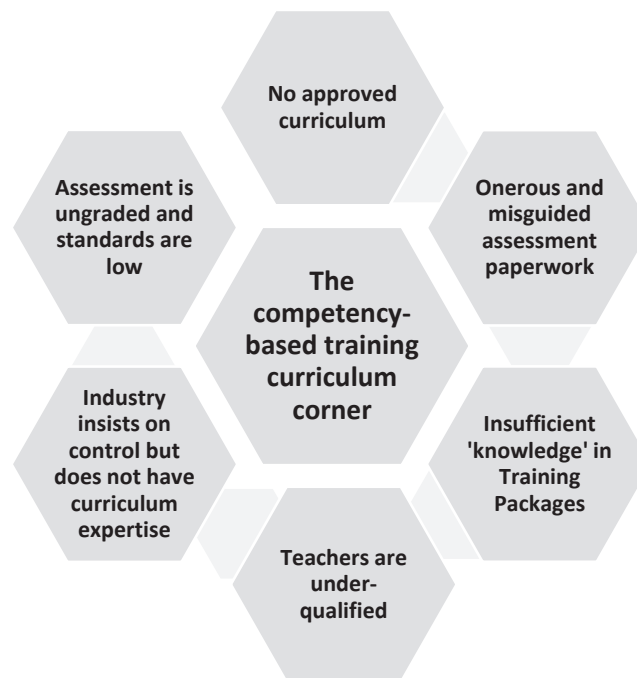


Fig. 1 How VET curriculum is constrained in Australia.

formerly known as “teaching and Learning managers” and concerned with improvement the quality of teaching delivery, progressively added “quality” to their title and by 2020 most were simply called “quality and compliance” managers. The paperwork required to generate “training and assessment strategies,” “assessment tools” and to carry out the “assessment validation” cycle utterly consumed them and their staff.

Fig. 1 below illustrates the features of the current competency-based system which are preventing the development of good curriculum, and hindering teachers’ ability to assess students confidently and accurately.

Getting out of the corner: how can this situation be improved?

How should the system get out of its corner and escape its constraints? Some researchers who are opponents of the current CBT system (e.g. [Wheelahan, 2016](#)) advocate a jump. They would prefer to see the whole system overturned; in terms of the metaphor they advocate leaving the room altogether. But they do not offer realistic alternatives that acknowledge the day-to-day operations of the VET system. The huge leap that would be required to abandon the current CBT system and enter a new, unpainted, “room” would not be achievable, at least unless it was absolutely clear that everyone found the current system unworkable. But this is not the case; there are many stakeholders in the VET system whose voices cannot be ignored, and many training providers—public and private—who would be adversely affected by needing to rewrite all of their teaching programs.

In this spirit, a small number of improvements are now suggested which would involve the CBT system in Australia tip-toeing out of its current corner. They would undoubtedly involve situations where stakeholders and practitioners would find their feet encountering sticky surfaces, but there would be a relatively light tread involved. The list of improvements is not intended to be exhaustive but provides a way out which would not abandon the CBT system entirely.

Firstly, it needs to be widely understood and promulgated that the content of units of competency is necessary but not necessarily sufficient, and that training providers are free to add extra content. This statement could be easily added into each unit of competency, with little disruption caused.

Secondly, more “knowledge” could readily be inserted into Training Packages. Specific “knowledge units” could be added to the core units for each qualifications. The existing “knowledge” sections for each unit (See [Appendix 1](#)) could immediately be moved from the “assessment requirements” into the unit of competency itself. This would be a “paperwork” exercise that would not cause great disruption, but that would ensure that the knowledge components of units were actually taught. Thirdly, grading in assessment can be instituted. It need not be complex and technical; teachers could be trusted to work within an assessment framework set by their employing RTO, as are university teachers.

Fourthly, a mid-way point is needed between units of competency, which are essentially lists of work activities, and the teaching-learning interaction. The current mandated “training and assessment strategies” are far too prescriptive. Teachers should be left to develop their own unit- and session-level curricula. But an RTO needs to decide an overall teaching strategy for its qualifications, of a similar nature to the “National Framework for the Recognition of Training” curriculum between 1992 and 1997.

Finally, a restoration of the education levels of VET teachers to former levels would better equip teachers to manage the dilemmas associated with competency-based training, and to augment curriculum as required. The low level of qualification for

VET teachers is perhaps a peculiarly Australian problem. Fewer than 25% of VET teachers are now qualified past Certificate IV level in pedagogy (Knight et al., 2020), and only 10% at university level. In the 1980s and 1990s, by contrast, all new full-time teachers in the public (TAFE) system had undertaken in-service teacher-education at universities (Harris, 2020). The gradual decline in the pedagogical qualifications of the VET workforce began, somewhat ironically, with the introduction of the competency-based Certificate IV qualification for VET teachers (Smith, 2021).

Conclusion

The chapter has described the changes in the Australian VET curriculum system which can be described as successive attempts to “fix” problems arising from an originally imperfect model. The competency-based system is outcomes-focused, which in itself is not problematic, but the implementation has been deeply flawed. In general the successive steps have been taken in good faith, and seem to flow logically, but have progressively improved options and reduced quality. Alternative solutions have not been considered because of the influence of powerful stakeholders including industry peak bodies and trade unions. VET researchers have tended to be oppositional rather than constructive, and teachers have been ill-equipped to argue for better solutions due to a lack of formal educational knowledge.

The chapter, based on research and extensive high-level experience in and with the sector, presents a picture of a system in some disarray, due to particular circumstances in the Australian system. The account may be helpful for other countries going through similar processes and debates, or considering reworking their curriculum systems. It is hardly news that curriculum is a political arena (e.g. Print, 1993), and VET has a unique set of strong political actors due to tripartite involvement in most countries (e.g. Deissinger and Gonon, 2015), with employers and trade unions alike claiming ownership of VET systems. Because VET curriculum is thus embedded within other political institutions, a gradual approach to change, including—as in this case—extricating itself from the corner into which the curriculum has become stranded, is the most likely to succeed.

Appendix 1. Extracts from two units of competency (elements and performance criteria only) and assessment requirements (knowledge evidence only)

Source: www.training.gov.au, the national register.

Unit AURTTJ003—“Remove and replace wheel and tyre assemblies”

Parent Training Package: Automotive Retail, Service and Repair. Found in only two qualifications—both at Certificate II level. 10 hours delivery specified in the Victorian Purchasing Guide: <https://www.education.vic.gov.au/training/providers/rto/Pages/purchasingguides.aspx>.

Note: In the absence of Training Package guidance on hours of training delivery, State Government Purchasing Guides are used as a proxy. The guides produced in Victoria are in wide use.

Elements	Performance criteria
Elements describe the essential outcomes	Performance criteria describe the performance needed to demonstrate achievement of the element. Where bold italicized text is used, further information is detailed in the range of conditions section
1. Prepare to remove and replace wheel and tyre assemblies	1.1 <i>Safety requirements</i> are sourced and interpreted 1.2 Task instruction is interpreted and vehicle wheel and tyre assemblies to be worked on are identified 1.3 Manufacturer specifications and workplace procedures for removing and replacing wheel and tyre assembly are sourced and interpreted 1.4 Potential hazards and risks associated with task are identified and reported to workplace supervisor 1.5 <i>Tools and equipment</i> required for wheel and tyre assembly removal and replacement are identified according to manufacturer specifications
2. Remove and inspect vehicle wheel and tyre assemblies	2.1 Tools, equipment and materials are selected and checked prior to use according to manufacturer specifications and safety requirements 2.2 <i>Vehicle</i> is prepared for wheel removal according to workplace procedures and safety requirements 2.3 <i>Wheel and tyre assemblies</i> are removed according to workplace procedures and manufacturer specifications, and without causing damage to components, tools or equipment 2.4 Wheel and tyre assemblies are identified and inspected according to manufacturer specifications 2.5 Wheel and tyre assembly inspection results are recorded

(Continued)

<i>Elements</i>	<i>Performance criteria</i>
3. Replace wheel and tyre assemblies	3.1 Wheel and tyre assemblies are prepared for replacement 3.2 Wheel and tyre assemblies are replaced according to workplace procedures, manufacturer specifications and safety requirements without causing damage to components, tools or equipment 3.3 Wheel and tyre assembly on vehicle is checked for correct operation
4. Complete work processes	4.1 Final inspection is made to ensure work meets task instruction and workplace standards, and vehicle wheels and tyres are presented ready for use or storage according to workplace procedures 4.2 Work area is cleaned, waste and non-recyclable materials are disposed of, and recyclable material is collected and stored according to environmental requirements and workplace procedures 4.3 Tools and equipment are checked and stored according to workplace procedures, or tagged and reported where necessary 4.4 Workplace documentation is processed according to workplace procedures

Knowledge evidence

Individuals must be able to demonstrate knowledge of:

- key aspects of work health and safety (WHS) and occupational health and safety (OHS) requirements, including:
 - use of personal protective equipment, including safety glasses, ear protection and safety footwear
 - use of hand tools and lifting equipment
 - procedures for handling wheel and tyre assemblies
- types and application of wheel assemblies, including:
 - stamped or pressed steel wheels
 - alloy wheels
 - wheel studs and nuts
 - tyres, including cross-ply and radial tyres
- inspection procedures for tyre and wheel assemblies, including wheel inspection and tyre wear
- removal and replacement procedures and precautions
- tyre air pressure setting and test procedures
- work area clean-up and maintenance requirements.

Unit HLTINF001—“Comply with infection prevention and control policies and procedures”

Parent Training Package: Health. Found in 38 qualifications: Certificate II to Diploma level. 25 hours delivery specified in the Victorian Purchasing Guide.

<i>Element</i>	<i>Performance criteria</i>
Elements define the essential outcomes	Performance criteria describe the performance needed to demonstrate achievement of the element
1. Follow standard and additional precautions for infection prevention and control	1.1 Follow hand hygiene practices in accordance with organizations policies and procedures 1.2 Implement hand care procedures and cover cuts and abrasions 1.3 Follow organization procedures for choice and use of personal protection equipment 1.4 Follow procedures for respiratory hygiene and cough etiquette 1.5 Follow procedures for environmental cleaning 1.6 Follow procedures for handling, transporting and processing of linen in a manner that controls the spread of infection 1.7 Follow procedures for disposal of contaminated waste 1.8 Follow procedures for handling and cleaning client equipment that prevents skin and mucous membrane exposures, contamination of clothing, and transfer of pathogens 1.9 Identify and respond to situations where additional precautions may be required to prevent transmission of infection

(Continued)

<i>Element</i>	<i>Performance criteria</i>
2. Identify infection hazards and assess risks	2.1 Identify infection hazards associated with own role and work environment 2.2 Identify own areas of responsibility in relation to infection prevention and control 2.3 Assess risk by determining the likelihood and severity of harm from identified hazards. 2.4 Document and report activities and tasks that put self, clients, visitors and/or other workers at risk 2.5 Identify appropriate control measures to minimize risk in accordance with organizations procedures
3. Follow procedures for managing risks associated with specific hazards	3.1 Follow protocols for care after exposure to blood or other body fluids as required 3.2 Place appropriate signs when and where appropriate 3.3 Remove spills in accordance with the policies and procedures of the organization 3.4 Minimize contamination of materials, equipment and instruments by aerosols and splatter 3.5 Identify, separate and maintain clean and contaminated zones 3.6 Confine records, materials and medicaments to a well-designated clean zone 3.7 Confine contaminated instruments and equipment to a well-designated contaminated zone

Knowledge evidence

The candidate must be able to demonstrate essential knowledge required to effectively complete tasks outlined in elements and performance criteria of this unit, manage tasks and manage contingencies in the context of the work role. This includes knowledge of:

- established guidelines for the prevention and control of infection, including those for:
 - personal and hand hygiene:
 - how to hand wash
 - how to hand rub
 - pre-surgical hand preparation
 - clinical moments when hand hygiene should be performed with soap and water rather than alcohol-based hand rub
 - non-clinical moments for hand hygiene hand care, including guidelines on maintaining intact skin, fingernails and jewellery/watches
 - use and scope of personal protective equipment guidelines for:
 - glove use
 - wearing gowns and waterproof aprons
 - wearing masks
 - wearing protective glasses
 - surface cleaning:
 - cleaning procedures and their specified times
 - routine surface cleaning
 - managing a blood or body fluid spill
 - sharps handling and disposal techniques
 - reprocessing procedures for equipment
- types of additional precautions and their relevance to particular areas of work or client groups
- types of hazards in the work environment and associated risks and control measures
- chain of infection:
 - source of infectious agent
 - mode of transmission
 - susceptible host
- basis of infection, including:
 - bacteria and bacterial spores
 - difference between harmless microorganisms and pathogens
 - difference between colonization, infection and disease

- fungi
- viruses
- key modes of disease transmission—contact, airborne and droplet:
 - paths of transmission including direct contact, aerosols and penetrating injuries
 - risk of acquisition
 - sources of infecting microorganisms including persons who are carriers, in the incubation phase of the disease or those who are acutely ill
- factors that increase the susceptibility to infection:
 - immune status
 - wounds or devices
 - medications and comorbidities
 - age

References

- Allais, S., Marlock, C., Molebatsi, P., 2014. The Development of Occupational Standards in English-Speaking Countries. A Report for the International Labor Organization, Moscow.
- Allen Consulting Group, 1994. Successful Reform. Allen Consulting Group, Melbourne.
- Australian National Training Authority (ANTA), 2004. Industry Skills Councils: Skills for Australian Industry. Out of Print. Available at: <http://hdl.voced.edu.au/10707/268511>.
- Australian Skills Quality Authority (ASQA), 2015. Guide to Developing Assessment Tools. ASQA, Melbourne. https://www.asqa.gov.au/sites/default/files/Guide_to_developing_assessment_tools.pdf.
- Barratt-Pugh, L., Soutar, G., 2002. Paradise Nearly Gained: Vol. 1 Evaluating the Frontline Management Initiative. NCVER, Adelaide.
- Burke, J., 1989. The implementation of NVQs. In: Burke, J. (Ed.), *Competency-Based Training*. Falmer, London, pp. 109–131.
- Clayton, B., 2013. Assessment in VET: An Exploration of the Issues. Presentation at University of Ballarat. October.
- Cornford, I., 1996. Experienced teachers' views of competency-based training in NSW TAFE. In: *Learning and Work: The Challenges Conference*, December, Griffith University, Gold Coast, December.
- Deissinger, T., Gonon, P., 2015. Stakeholders in the German and the Swiss VET system and their role in innovating apprenticeships. In: *Architectures for Apprenticeship: Achieving Economic and Social Goals*, 6th INAP (Network on Innovative Apprenticeship) Conference, Federation University Australia, Ballarat, 1–2 September.
- Department of Education & Training, 2016. Quality of Assessment in Vocational Education and Training. Discussion Paper. Australian Government, Canberra.
- Department of Education and Training, 2017. Standards for RTOs (2015), Including 2017 Amendments. Australian Government, Canberra. <https://www.legislation.gov.au/Details/F2017C00663/Download>.
- Department of Education, Skills & Employment, 2012. Standards for Training Packages. Commonwealth of Australia, Canberra. <https://docs.employment.gov.au/documents/standards-training-packages>.
- Department of Industry, 2012. Standards for Training Packages. Australian Government, Canberra. https://docs.education.gov.au/system/files/doc/other/standards_for_training_packages_2012.pdf.
- Gillis, S., Griffin, P., 2005. Principles underpinning graded and assessment in CBT: a critique of prevailing perceptions. *Int. J. Train. Res.* 3 (1), 53–78.
- Glatthorn, A., 1987. Curriculum Renewal. Association for Supervision and Curriculum Development, New York.
- Goncz, A., 1994. Competency-based assessment in the professions in Australia. *Assess. Educ.* 1, 27–44.
- Harris, R., 2020. Landmarks in the Development of the VET Workforce: An Overview. NCVER, Adelaide.
- Hodge, S., 2014. Interpreting Competencies in Australian VET: Challenges and Issues. NCVER, Adelaide.
- Joint Steering Committee of the National Quality Council and the Council of Australian Governments, 2009. VET Training Products for the 21st Century. Out of print. Available at: <https://www.voced.edu.au/content/ngv%3A5470>.
- Joyce, S., 2019. Strengthening Skills: Expert Review of Australia's Vocational Education and Training System. Department of Prime Minister and Cabinet, Canberra.
- Knight, G., White, I., Granfield, P., 2020. Understanding the Australian Vocational Education and Training Workforce. NCVER, Adelaide.
- Lambert, J., 2016. Australian vocational training arrangements—industry skills Committees. In: *Jobs, Skills and Strategies for the Future: Perspectives From Asia-Pacific*. ILO, Korean Ministry of Employment and Labor, and HRD-Korea, Seoul, 5–6 October 2016.
- Mills, J., Crean, D., Ranshaw, D., Bowman, K., 2012. Workforce Skills Development and Engagement in Training Through Skill Sets. NCVER, Adelaide.
- National Centre for Vocational Education Research (NCVER), 2020. Government-Funded Students and Courses 2019. NCVER, Adelaide.
- National Skills Standards Council, 2013. Standards Policy Framework: Improving Vocational Education and Training: The Australian Vocational Qualifications System. NATESE, Melbourne.
- NCVER, 2019. Government-Funded Students and Course 2018. NCVER, Adelaide.
- Phillimore, J., 1997. Trade unions and the National Training Reform agenda in Australia, 1983–1996. *Int. J. Train. Dev.* 1 (1), 34–48.
- Print, M., 1993. Curriculum Development and Design, second ed. Routledge, Abingdon, U.K.
- Robinson, P., 1993. Teachers Facing Change: A Small-Scale Study of Teachers Working With Competency-Based Training. NCVER, Adelaide.
- Schofield, K., McDonald, R., 2004. Moving on: Report of the High Level Review of Training Packages. ANTA, Melbourne.
- Smith, E., Keating, J., 2003. From Training Reform to Training Packages. Social Science Press, Tuggerah Lakes, NSW.
- Smith, E., Hill, D., Smith, A., Perry, P., Roberts, P., Bush, A., 1996. The Availability of Competency-Based Training in TAFE and Non-TAFE Settings in 1994. Australian Government Publishing Service, Canberra.
- Smith, E., 1999. Ten years of competency-based training: the experience of accredited training providers in Australia. *Int. J. Train. Dev.* 3 (2), 106–117.
- Smith, E., 2010. A review of twenty years of competency-based training in the Australian vocational education and training system. *Int. J. Train. Dev.* 14 (1), 54–64.
- Smith, E., 2019. Competency-based training Australian style: how it exposes VET to risk. In: “No Future for Old VET”: Researching for the Training System/s of Tomorrow: 22nd Annual Conference of AVETRA, Western Sydney University, Parramatta, June 17–18. <http://avetra.org.au/pages/conference-archives-2019.html>.
- Smith, E., 2021. Australian TVET teacher training—once flourishing but now neglected. In: Stolte, H. (Ed.), *TVET Teacher Education and Training in International and Developmental Co-operation—Models, Approaches and Trends*, Springer Series Technical and Vocational Education and Training: Issues, Concerns and Prospects. https://doi.org/10.1007/978-981-16-6474-8_26. Chapter 26.
- Sung, J., 2011. The Singapore Continuing Education and Training System. Labor Market Research. February 1–12.
- Thomson, P., Foyster, J., Tucker, J., Toop, L., Dean, S., 1990. Competency-Based Training in TAFE. TAFE National Centre for Research and Development, Adelaide.

- Thomson, P., Mathers, R., Quirk, R., 1996. The Grade Debate. NCVER, Adelaide.
- Torrance, H., 2012. Formative assessment at the crossroads: conformance, deformative and transformative assessment. *Oxf. Rev. Educ.* 38 (1m), 323–342.
- West, J., 2004. Dreams and Nightmares: The NVQ Experience. Working Paper No. 45. Centre for Labor Market Studies, Leicester.
- Wheelan, L., 2016. Patching bits won't fix vocational education in Australia—a new model is needed. *Int. J. Train. Res.* 14 (3), 180–196.
- This page intentionally left blank

Exploring the professional identity of university academics: “becoming” a university teacher

Precious Bupe Sipuka and Shireen Motala, University of Johannesburg, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	504
Professional identity in education	504
“Being” a good university teacher	505
Constructivism and social practice	506
Methods	506
Results	507
Personal dispositions and intrinsic factors	507
“Becoming” a good teacher	507
Professional development	508
Discussion	509
Conclusion	510
References	511

Introduction

The discourse on quality teaching as an enabler of student success has brought into focus issues around teacher identity in higher education. Although there has been considerable research on the identity of teachers in the schooling system (Beauchamp and Thomas, 2010; Akkerman and Meijer, 2011; Flores and Day, 2006; Stenberg, 2010), there have been only a few studies of lecturer identity in higher education (Archer, 2008; Clegg, 2008; Lea and Stierer, 2009). Global movements including massification, rising costs in higher education and advances in communication technology have brought the issue of quality teaching to center stage (Mok and Welch, 2003). The emergence of a knowledge-based economy requires that higher education re-imagines its character and purpose (Burbules and Torres, 2000; Jaffer et al., 2007; Scott et al., 2007; Bozalek et al., 2013). Diversity in the student population and increased access and participation rates require that higher education responds adequately to increasing demands.

Professional identity in education

Based on the social constructivist approach, identity is characterized as the social and cultural self, shaped by life experiences and through reflecting on these experiences with oneself and others (Swennen et al., 2008). In the study of “identity,” the level of consciousness about one’s identity and the extent of agency in its “construction” are essential questions to consider. Identity is not fixed but evolves over time (Beijaard et al., 2004). Identity building is inclusionary: each individual interacts in many socially and culturally created and distinct worlds which are continuously socially organized and replicated (Swennen et al., 2008). The formation of identity takes place in an inter-subjective field and is an ongoing process of self-interpretation and acknowledgment in a particular context (Gee, 2001). According to Wenger (1998), identity is created by involvement in various communities of practice, deliberately and unintentionally.

Professional identity as individual interaction is one element of the individual self-concept that is formed by a conversation between the self and outside reality in a social context (Hermans and Dimaggio, 2007). Thus, professional identity is always relational depending on the context and position in which one finds her/himself. By selecting and/or rejecting other possibilities, a person confirms affiliations in a variety of professional fields and distinguishes those which are an essential component of her career: her identity (Coldron and Smith, 1999). Kreber (2010) indicates that teacher identities are formed by the complex relationship between personal teaching theories and self-perceptions, both mediated by social and career background.

Globally, discourse on teacher identities has gained prominence as higher education evolves and as demands increase. There is more scrutiny of how shifting roles, pedagogy and values shape academics’ professional identities (McNaughton and Billot, 2016; Trautwein, 2018). Locally, academic developers such as Jawitz (2009) and Leibowitz et al. (2005) have raised concerns about the impact of global changes on the professional identity and development of university teachers, especially in the complex context of post-apartheid South Africa. This study aims to expand on the global discourse on teacher identity, using South Africa as a case study. Traditionally, literature on teacher identity tends to reflect views from the global north. It is important to engage in research

which contributes to the discourse on teacher identity from a perspective relevant to the conditions of the global south. For this purpose, the study explores the following timely and important questions:

1. How have the changing nature of higher education and increasing demands affected/shaped teacher/lecturer identity within higher education?
2. How has this identity shift impacted on teacher practices and ultimately student outcomes?

Research on teacher identity argues that current demands on university teachers have affected the ability to develop and build a clear professional identity (Thomas and Beauchamp, 2011; Jawitz, 2009). Formation of identity is an ongoing process of perception and reinterpretation of experiences (Beijaard et al., 2004, 2000) and is a complex process of learning which is continuous and affects personal history, social interactions and cultural and psychological factors (Lamote and Engels, 2010). MacLure (1993) argues that the self and identity are fundamental to teaching and professional growth suggesting that teachers use identities to explain, justify and understand themselves with others and to the world as a whole.

“Being” a good university teacher

Elements that affect professional identity formation are closely associated with conceptions of the ideal person as a professional and the principles on which the ideal person is based (Verhofstadt-Deneve, 2003). According to Laurila and Kukkonen (2003), the creation of the teacher’s identity can be interpreted as seeking a balance and resolving the contradictions between the ideal self and the real self. In understanding the dynamics of the teacher’s development process, it is therefore important to see identity from two perspectives: the idea of an “ideal” self and the real self (Beauchamp and Thomas, 2010).

An individual’s ideal identity consists of the ideal image of what she wants to be (Ruyter and Conroy, 2002). Each teacher has an image of ideal teachers and their values and skills (Arnon and Reichel, 2007). The ideal is a measure of the assessment of realistic activities of a teacher, personal as well as technical. It is an “ideal” that serves as a potential future and a sum of the desired qualities of a teacher. This idea is backed by possible-self theory (Markus and Nurius, 1986), explaining the significance and the dynamics of self-relevant, forward-looking concepts and how these contribute to motivation for present and future actions. These self-conceptions relate to how people think about their potential and future and may reflect a person’s expectations: the hopes, dreams, fears and threats that they foresee in the near or more distant future (Hamman et al., 2010).

In the view of Markus and Nurius (1986), potential roles themselves perform essential purposes leading to current behavior, reflect the “picture” of what one wants to become and serve as a “behavioral blueprint.” Individuals who are aware of what they want to become will be more likely to follow their interests and goals (Hamman et al., 2010). Holding an idea of your future self and understanding the complex forces that can impact the future can allow teachers to track their progress (Beauchamp and Thomas, 2010). Fletcher (2000) explains how the teacher’s values can be transformative both by conceiving a particular purpose as a potential teacher (for example: to be caring and efficient) and by the motivational value of finding and anticipating the nature of the teacher.

Increasingly, studies show the role that socio-cultural and political factors play in shaping the professional identity of university teachers (Lasky, 2005; Phyak and Baral, 2019). As Freire (in Giroux, 2010, 715) argues, the pedagogical foundations of higher education institutions are “a deeply civic, political, and moral practice.” In South Africa, transformation has been a key factor that has shaped the pedagogical orientation of higher education and university teacher identity has had to be reconstructed around the transformative agenda of post-apartheid South Africa. As previous studies show (Beauchamp and Thomas, 2010; Hamman et al., 2010; Arnon and Reichel, 2007), the interpretation of the good teacher is a significant part of the ideal teacher. Ruyter and Conroy (2002) found that, in their study, the majority of students with a Bachelor of Education qualification listed personal characteristics such as being polite, truthful, accessible and natural as the features of their ideal instructor. Likewise, Strage’s (2008) research found that the features most used by undergraduates to represent their ideal professor were: educated, fun, caring and student-oriented. A teacher’s personal experience of learning in the sense of her/his experiences (positive or negative) not only affects her/his image but also forms her/his cognition of the ideal school or university teacher.

In their research, Arnon and Reichel (2007) found that teacher’s expectations of an ideal teacher include professional personality consisting of both subject-matter expertise and didactic skills. Karjalainen and Nissilä (2008), in their work with university teachers, describe the qualities of an ideal university teacher across three key fields: teaching abilities, teacher personality and subject matter expertise. The three dimensions are shaped and negotiated through the socio-cultural and political context to construct the “ideal” university teacher.

Literature on teacher education (Loughran, 2002; Freese, 2006) indicates a strong relationship between reflection and identity development. According to Johns (1995), the process of reflection helps people to consider and assess their experiences. Korthagen and Vasalos (2005) understand reflection as a process focused on actions to create awareness, which is informed by the social reality and the intrinsic qualities of the subject. Farrell (2011) defines reflection as a key component of the self-concept because it takes implicit concepts to a level of consciousness. In other words, when there is no thinking, it is difficult to talk for yourself.

Lamote and Engels (2010) note that it is important to focus also on the kind of teachers that university teachers want to be in the course of the growth of their professional identities. Furthermore, teacher identity formation means that teachers explore potential

perspectives on their task and professional growth, known as anticipatory reflection (Beijaard et al., 2004). While retrospective thought is based on memory, anticipatory reflection contains an element of imagination which looks forward to an opportunity that the student-teacher may create (Beuchamp and Thomas, 2010). This study reflects on lecturer beliefs about what makes a good or “ideal” teacher in relation to quality teaching.

Constructivism and social practice

This study¹ employs Sin’s (2014) concept of the “policy object.” According to Sin, this describes not the actual policy text but its core dimensions or concepts; actors’ beliefs about the policy object will influence how they enact policy and impact policy outcomes. Sin’s “policy object” is drawn from two theories: Berger and Luckmann’s (1966) social construction of reality and the social practices perspectives of Bamber et al. (2009). Both the social construction of reality and the social practice approach acknowledge the significant role that actors and context play in the way that “shared knowledge and collective symbolic structures of knowledge and the way meaning is ascribed to the world enable certain interpretations of reality and practice” (Sin, 2014, 438).

In the social construction of reality, reality and knowledge are transmitted and maintained in social situations (Berger and Luckmann, 1966). The main tenet of social constructionism is the duality of reality: that the world is created not given and is both subjective and objective (Berger and Luckmann, 1966). In this case, reality is produced by social action as it appears to the individual as separate and independent from her/him, although it is the result of subjectively meaningful actions (Berger and Luckmann, 1966). Social practice theory views actions as practices that have been conditioned by the socially shared structures of knowledge which lead to certain interpretations of reality and behaviors (Bamber et al., 2009, 7–8). As Ball (1994, 22) states “discourse speaks to us” and shapes the policy positions that are taken.

In the academic context, social practice theory relates to a process where people working together develop understandings, values, attitudes and practices particular to their context (Bamber et al., 2009). The social practice approach adopts a socio-cultural lens to explain daily practices; thus “social practice” is used as the unit of analysis. This study, therefore, provides an important opportunity to explore how the implementation of teaching and learning policies shape lecturer enactment toward good teaching.

Methods

The Ph.D. research on which this paper is based was a multi-level and multi-case study, collaborative and reflective in nature. Five universities were selected according to the following criteria: rural, disadvantaged and teaching focused; institutional type (merged, comprehensive, traditional); location (rural or urban); focus (teaching, research or both). The criteria used in the study reflect the kinds of institutions in South Africa in relation to geography, access to resources, student population and focus. This allowed for a more nuanced understanding of how context shapes the enactment of teaching.

The Ph.D. study was qualitative in nature. Qualitative research provides a framework for understanding social or human problems based on building a complex, holistic picture, formed with words, reporting detailed views of informants and conducted in a natural setting (Creswell, 1998; Denzin and Lincoln, 2000). This ontological view of reality is in line with the theoretical approach adopted for the study which sees reality as being a result of social interactions. Qualitative research allows us to answer descriptive questions about “what is happening” and “what or how it is happening” (Shavelson and Towne, 2002, 99). Qualitative research, therefore, acknowledges the social way in which reality is created and made sense of and attempts to obtain insights into particular social process and practices within a context (Connolly, 1998).

A multi-case study approach was taken, informed by a number of case study methodologies. According to Yin (2003, 47) a multiple case study either “(a) predicts similar results (a literal replication) or (b) predicts contrasting results but for predictable reasons (a theoretical replication).” As Yin (1994, 31) indicates, “a previously developed theory is used as a template with which to compare the empirical results of the case study. If two or more cases are shown to support the same theory, replication may be claimed. Empirical results may be considered yet more potent if two or more cases support the same theory but do not support an equally plausible rival theory.” A case is defined as “a unit of human activity embedded in the real world; which can only be studied or understood in context; which exists in the here and now; that merges in with its context so that precise boundaries are difficult to draw” (Yin, 1994, 13). An argument for case studies in qualitative research is that they make it possible to consider contextual conditions (Şebnem, 2009); in this study, the case study design provided a framework for exploring how social processes and practices, in a specific context, shape the professional identity of university academics as teachers.

¹This paper is drawn from a doctoral study with a social constructivist and practice orientation. It was a sub-study within a larger National Research Foundation project which aimed to strengthen staff development in the institutional context using Margaret Archer’s social realist theory. Structure, culture and agency as important elements of social realist theory provided an important foundation for the Ph.D. study. Ethical clearance for the study was obtained from each of the participating institutions. All information identifying specific individuals was removed at the stage of transcription. The data was thereafter archived electronically. Only individuals who conducted the interviews know the identity of the interviewees.

Forty-nine semi-structured interviews were conducted with teaching academics selected according to a predetermined matrix of criteria: gender, faculty, “race,” seniority, nationality. Thematic analysis was employed as the analytical framework of the study. Data coding was carried out, with themes used as a guide for further analysis of data. The themes initially created assisted in the unifying and clarifying of ideas (Anderson et al., 2003).

Results

Personal dispositions and intrinsic factors

The study revealed that university teachers’ professional identity is linked to personal and intrinsic dispositions and attitudes. In particular, the positive orientation shaped by their internal locus of control toward teaching is an important element of their identity as teachers.

but I think, for me personally, it’s more of intrinsic motivation. I want to be a good lecturer because I love it. I love it when I see the students suddenly realize how something fits together and, you know, the sort of the penny drops, so to speak [L11].

I really enjoy what I do, I love working with students and I love the idea that, that I’m actually building skill that’s useful in the workplace [L49].

Personal values such as passion, commitment and love are described as key to shaping lecturer teaching practice. These are reflected in student-centered, transformative and developmental approaches to teaching.

So my orientation towards teaching is really about people. It’s about building relationships that will enable them to develop character, skills, transferable skills and also to develop a passion for the things that I love, which is chemistry [L14].

Intrinsic attributes contribute to how teaching is perceived and experienced and the commitment that lecturers have to good teaching practice. Hence, personal values shape teacher orientations toward teaching (interactive teaching, peer group learning) which are associated with effective teaching.

[H]ere is the natural chemistry between people, individuals and groups of people, but I really enjoy teaching. I really enjoy being in that classroom and interacting with the students. I also think that it’s a very creative space or can be and that I’ve often tried out new things and I’ve been quick to abandon things that don’t work well, and I’d say that although there has been a framework of continuity in my actual teaching approach; teaching materials and so on [L3].

The study revealed the key role that personal values play in teachers’ perceptions and experiences of teaching; and that values interact with other elements in the teaching environment to shape enactment.

“Becoming” a good teacher

Interestingly, “becoming” a good lecturer emerged as an important component of good teaching. Quality teaching appears to be understood as a process of “becoming” and/or “being” rather than a static event. Being a good lecturer appears to be understood as a developmental process and cumulative in nature. Some lecturers acknowledged the impact that time and history play in shaping one’s pedagogical practices.

I believe that teaching is a spiral process or teaching and learning is a spiral process so when one teaches you to teach at this level and then you come back and you teach the same thing but you’re going into deeper and deeper and deeper levels of it. So, it’s cumulative teaching but going in a spiral movement and I also think that I’m as much a learner as I am a lecturer because I learn from the students, every day. I learn from them [L9].

An inherent characteristic of being a good lecturer appears to be one’s ability to keep up with the demands of the evolving higher education context and, by extension, society.

I really enjoy being in that classroom and interacting with the students. I also think that it’s a very creative space or can be and that I’ve often tried out new things and I’ve been quick to abandon things that don’t work well, and I’d say that although there has been a framework of continuity in my actual teaching approach; teaching materials and so on—it’s been an experience of constant change as well [L3].

The ability to teach well is therefore defined not only by internal attributes but, more so, by external factors which demand continuous reflection and improvement. The professional identity of the lecturer is thus always evolving.

According to the findings from this study, beliefs about what it means to teach well or what it means to be a good teacher can be summarized through seven key concepts: transformative learning, student-centered approaches, pedagogy, curriculum, internal motivation and becoming a good lecturer. Underlying all these concepts is the significant impact that context has on shaping what it means to teach well. As a result, what emerges is the multi-dimensional nature of quality teaching. We cannot underplay the role that both lecturers and students play in ensuring that learning takes place. Thus, the relational nature of good teaching emerges as a crucial component.

What does this mean for higher education? It is clear that the political context of South Africa continues to shape higher education, with the transformative agenda remaining at the center of development. Pedagogical practices in higher education cannot be understood outside the civic, political and a moral imperatives (Freire, in Giroux, 2010), and the relational nature of good teaching demands that lecturers see teaching as more than a technical exercise.

Professional development

Perhaps not surprisingly, the study reveals professional development as an important mechanism for developing one's identity and skills as a university teacher. In particular, respondents identified workshops and seminars as important not only for reflecting on teaching practices but also creating a community of practice.

You know seminar sessions, things like the assessor's course, the teaching portfolio for me was huge in really getting me to take seriously a reflection you know to invest in it and then share that with the community that I work in [L18].

Peer networks as communities of practice were also identified as important spaces for sharing good practices. Lecturers said that peer networks allowed them to share experiences, learn about good practices and develop as academics by engaging critically on teaching and learning issues with fellow academics.

What I have been doing is attending conferences and presenting what I'm doing. Because what I have found, I learn better from sharing with other people. So, when people share their own experiences, then I can be able to actually see ... or are you doing it this way, how can I do it better. That's the only way that I feel that I learn best, by interacting with intellectuals, academics and other professionals with the same interests [L32].

The study shows the role that critical friends play in a lecturer's development as a university teacher. Peer interaction and critical feedback emerged as vital mechanisms for improving teaching practices. Communities of practice, in particular, were singled out as stimulating dialog and innovation in the classroom. A considerable number of lecturers mentioned that they engage with peers as critical friends in the classroom to get critical feedback to improve teaching practices.

I think speaking with my older colleagues in the department and asking them questions about particular things like, you know, how to set out the module and using older modules for similar courses that I was teaching, to try and get a sense of what it is that they had done [L19].

We co-teach for six months. And that is a fantastic experience because I like co-teaching, it helps one also to simplify. In humanities, we sit in on each other's lectures, that's one of the hallmarks of the course [L38].

Because critical friends appear to play a significant role in shaping teaching practices in the process of becoming a good lecturer, institutions are increasingly creating spaces and structures that facilitate peer interaction and dialog on teaching and learning.

Well, we do have, we do have those conversations in our staff meetings and we each year in our department we have a curriculum workshop that we have for the whole day. Trying to revise how the teaching takes place, how the modules are set up. But I think there could be more, I think more awareness about the issue. Because I can't change one way to benefit the student and someone else is doing something else. So, I think it definitely can improve [L17].

Interestingly, communities of practice appear to transcend departmental disciplinary and institutional boundaries. This is significant for creating a system with a shared understanding of good teaching practices.

But if I start looking at where the support comes from it first comes from, we have a really good national network within the profession. And good networks with other universities so and that's a good supportive environment Because it's a great team of people and there are really exceptional educators who I learn from all the time, the other lecturers that I work with, highly motivated. Also looking to improve, looking to do things better, looking for student feedback and student input so that we can sort of speak back to those things [L4].

The study signals the potential role that communities of practice may play in shaping lecturer identities as good teachers and in developing shared and common understandings of good teaching. This is important in a context where many lecturers find formal professional development activities to be ill-equipped, irrelevant and not responsive to their needs.

Something we don't do enough of is think critically about teaching beyond the introductory level, so all the workshops and discussions that happen at [the institution] tend to happen very much at the introductory level [L45].

Some lecturers reported that facilitators do not always have the expertise (qualifications and knowledge) to facilitate professional development activities.

I definitely, I think, I think you need the qualifications and obviously those qualifications must be given by people who are accredited and actually know what they're teaching. It can't just be a course for the sake of a course. It has to be thorough and that, you know peers and everyone can say yes, we believe in this course, we believe in the value of what's being taught [L16].

Apart from the lack of capacity in professional development, workload issues appear to negatively impact lecturers' ability to participate in professional development activities. Respondents mentioned that the timing of developmental activities often conflicted with teaching time.

Since the beginning of the year, the institution has been phenomenal in the number of trainings that they provide to lecturers to attend. The only problem is that they schedule it when I am teaching and then I am like, I am not going to give up my lecturing time with my students to go and train [L21].

because I think even the workload might affect how you embrace a lot of things, because if you are teaching a huge number, it tends to affect how you avail yourself to such opportunities [L27].

The findings reveal how structural and material conditions within institutions interact with personal values to affect the professional identity of university lecturers and enactment of good teaching.

Discussion

The multifaceted nature of teaching and learning demands that the identity of university teachers evolves with and responds to the current need to produce critical thinkers who are socially conscious for the 21st century labor market. “Ideals,” including identity, are complex and their roots and survival are shaped in a relevant context by personal interests, experiences and outcomes (Hamman et al., 2010). Identities evolve with the time and experience of teaching as professional identities grow. Consequently, at the heart of good teaching is the understanding that the ability to teach well is a process of “being” and “becoming.” The identity of the university lecturer, unlike teachers in the schooling system, has not been well understood (Beijaard et al., 2000; Beauchamp and Thomas, 2009). University lecturers are primarily appointed as disciplinary experts but are also expected to fulfill the teaching role. Therefore, the professional identity of the university lecturer is mainly formed and shaped in practice (in-service). The university lecturer “becomes” a teacher in practice. The process of professionalizing university teachers therefore involves not only what they know (discipline) and can do (skills) but also who they are becoming (Dall'alba, 2009). It does not happen in a vacuum but is mediated by the political, social and personal. Current professional development initiatives often neglect to reflect how university teachers navigate the political and socio-cultural influences which often shape how university academics develop as teachers and how they negotiate these in practice. Research reported by Thomas and Beauchamp (2007) stresses the significance of sociocultural factors in understanding how teachers negotiate their personal and professional selves in relation to the broader socio-political and institutional conditions. This gap is reflected in the development of one-size-fits-all, generic professional development programs. As Badat and Sayed (2014) advocate, professional development of university teachers should be done using rigorously conceptualized and designed academic development programs. Although the process of becoming is often facilitated by continuous professional development, one would question whether the current professional development models reflect the “becoming”

nature of good teachers. This study found that many lecturers experience professional development to be highly fragmented and not responsive to their needs. This negatively impacts identity formation of the academic as a teacher. Furthermore, the research and teaching nexus tension continues to constrain the development of academics as lecturers. Although institutional priorities vary, there is a concern that research is more highly valued than teaching. Chalmers' (2011), for example, found that university lecturers felt that their institutions prioritized research at the expense of teaching and learning.

Lecturers also argue that universities send out mixed messages especially when promotions, performance and publications are prioritized over teaching (Clarke et al., 2012, 10). Van Lankveld et al. (2017, 326) echo this by stating that undervaluation of teaching leads to tensions in academics' identity as lecturers and feelings of insecurity and reduced self-esteem. This is another example of how structural and material conditions within institutions constrain the promotion of good teaching. Herman and Cilliers (2008) indicate how the status of teaching has implications for academics' professional learning and the extent to which they seek to enhance their teaching practice. With increasing pressures and competing priorities, academics are more likely to engage in activities on which they are appraised and for which they are rewarded. Ball (2003) describes the current conundrum faced by university lecturers as one where they have to contend with numerous changes as professionals: “the struggle for the soul of the lecturer.”

Such research findings reflect the contested nature of how academics construct and re-construct their professional identity. However, Bathmaker and Avis (2005) argue that lecturer identity in itself is not the main issue but rather the way in which it shapes and contributes to how teaching and learning is enacted. It is therefore important to understand the elements that shape lecturer identity to understand how they shape conceptions and enactment of teaching.

Consistent with international practices, professional development emerges as an important activity for developing academics as lecturers. However, disciplinary expertise remains the main focus of professional development in South Africa. It may be argued that professional development is still shaped by the socio-political transformative imperatives of a post-democratic dispensation. Pedagogy remains on the periphery in higher education. This is in contrast to international practices increasingly recognizing the significant role that pedagogical training plays in shaping good teaching practices. Institutions continue to develop academics using traditional and formal methods: short courses and workshops by expert academic developers (Laat and Schreurs, 2013). This approach has been found to be less effective in developing academics as lecturers (Knapper, 2003). As observed in this study, traditional approaches to professional development have not been effective in developing lecturers because they tend to be designed with a one-size-fits-all approach, often lacking depth and expertise and conflicting with teaching time. As Smith and Kanuka (2018) have indicated, professional development activities tend to be generic. Its voluntary nature may also impact participation due to competing priorities within higher education.

Framing professional development as a social practice is important as studies have shown that the context within which academics work has a considerable influence on practice (Leibowitz et al., 2014). However, professional development approaches do not reflect current discourse which highlights the role that informal social networks play in how academics develop, learn, interpret, embrace, share, compile, contextualize and sustain new knowledge (Baker-Doyle and Yoon, 2011). Boud and Hager (2012) advance the notion that professional development models should be developed to locate learning in professional, day-to-day practice:

Learning is a normal part of working, and indeed [of] most other social activities. It occurs through practice in work settings from addressing the challenges and problems that arise. Most learning takes place not through formalized activities, but through the exigencies of practise with peers and others, drawing on expertise that is accessed in response to need. Problem-solving in which participants tackle challenges which progressively extend their existing capabilities and learn with and from each other appears to be a common and frequent form of naturalistic development (2012, 22).

In devising professional development models, we need to understand that learning in networks is beneficial for professionals (Harasim et al., 1995; Goodyear et al., 2004) and is a productive approach to professional development (Baker-Doyle and Yoon, 2011). However, learning communities can only be successful if they have a shared vision and institutional support. As Van der Linden et al. (2000) state, shared and supportive leadership, shared vision and values and shared learning collectively are important elements for learning communities. This has implications for policy design and signals the need for collaborative partnerships and engagement between a range of key stakeholders at different levels (Hudson et al., 2019). However, collaborative partnerships and engagement are complex and require that stakeholders with differing priorities establish common ground.

Conclusion

This study reflects on the need for greater insight into and scrutiny of the identity of university teachers. Increasing demands, technological advancements, funding, managerialism and the contested nature of higher education requires that we interrogate the (re) construction of teacher identities within higher education. It is important that we interrogate how university teachers navigate and negotiate the demands placed on them and how these shape and impact their practices. This will inform the development of responsive and contextually relevant professional development models.

The study contributes to current debates on teacher identity by expanding on the conceptual and methodological tools by using a constructivist and practice-based approach. It emphasizes the dynamic nature of teacher identity and requires us look at site-specific socio-cultural and political factors to understand how they construct or attempt to construct their professional identity; including what enables or constrains the development of a “healthy” identity.

In summary, this study reveals that teacher identities are always in a state of evolution, shaping and (re) shaping the nature and purpose of pedagogy. Locally, university teachers must negotiate, navigate and respond to the transformative agenda for a post-democratic dispensation, declining funding and technological advancements within a contested socio-political landscape. This pressure is compounded by the fact that university teachers are required to respond to low throughput rates and to improve student success. What does this mean for professional development? In South Africa, professional development models have reflected traditional and one-size-fits-all approaches. This study argues for a new way of theorizing professional development and the “university teacher” and the development of models which reflect the increasingly complex nature of teacher identity and the demands placed on them. The implications of this study have wider relevance for higher education across the world because of increasing changes and demands. The study contributes to a broader knowledge base and more nuanced understanding of teacher identity for the development of content, design and implementation of professional development models that are relevant and responsive to student success.

References

- Akkerman, S.F., Meijer, P.C., 2011. A dialogical approach to conceptualizing teacher identity. *Teach. Teach. Educ.* 27, 308–319.
- Anderson, M., Standen, P., Noon, J., 2003. Nurses' and doctors' perception of young people who engage in suicidal behavior: a contemporary grounded theory analysis. *Int. J. Nurs. Stud.* 40, 587–597. [https://doi.org/10.1016/S0020-7489\(03\)00054-3](https://doi.org/10.1016/S0020-7489(03)00054-3).
- Archer, L., 2008. The new neoliberal subjects? Young/er academics' constructions of professional identity. *J. Educ. Pol.* 23, 265–285.
- Aron, S., Reichel, N., 2007. Who is the ideal teacher? Am I? Similarity and difference in perception of students of education regarding the qualities of a good teacher and of their own qualities as teachers. *Teach. Teach. Theor. Pract.* 13 (5), 441–464.
- Badat, S., Sayed, Y., 2014. Post-1994 South African Education: the challenge of social justice. *Ann. Am. Acad. Polit. Soc. Sci.* 652, 27–148. <https://doi.org/10.1177/0002716213511188>.
- Baker-Doyle, K.J., Yoon, S.A., 2011. In search of practitioner-based social capital: a social network analysis tool for understanding and facilitating teacher collaboration in a US-based STEM professional development program. *Prof. Dev. Educ.* 37 (1), 75–93. Available from. <https://www.learnlib.org/p/52158/>.
- Ball, S.J., 1994. *Education Reform: A Critical and Post Structural Approach*. Open University, Bristol.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Policy* 18 (2), 215–228. <https://doi.org/10.1080/0268093022000043065>.
- Bamber, V., Trowler, P., Saunders, M. (Eds.), 2009. *Enhancing Learning, Teaching, Assessment and Curriculum in Higher Education Theory, Cases, Practice*. McGraw-Hill: Society for Research into Higher Education and Open University Press, Maidenhead, England.
- Bathmaker, A., Avis, J., 2005. Becoming a lecturer in further education in England: the construction of profession identity and the role of communities of practice. *J. Educ. Teach.* 31 (1), 47–62. <https://doi.org/10.1080/02607470500043771>.
- Beauchamp, C., Thomas, L., 2009. Understanding teacher identity: an overview of issues in the literature and implications for teacher education. *Camb. J. Educ.* 39 (2), 175–189. <https://doi.org/10.1080/03057640902902252>.
- Beauchamp, C., Thomas, L., 2010. Reflecting on an ideal: student teachers envision a future identity. *Reflective Pract.* 11 (5), 631–643.
- Beijaard, D., Meijer, P., Verloop, N., 2004. Reconsidering research on teachers' professional identity. *Teach. Teach. Educ.* 20, 107–128.
- Beijaard, D., Verloop, N., Vermunt, J., 2000. Teachers' perceptions of professional identity: an exploratory study from a personal knowledge perspective. *Teach. Teach. Educ.* 16, 749–764.
- Berger, P.L., Luckmann, T., 1966. *The Social Construction of Reality. A Treatise in the Sociology of Knowledge*. Doubleday, New York.
- Boud, D., Hager, P., 2012. Re-thinking continuing professional development through changing metaphors and location in professional practice. *Stud. Cont. Educ.* 34 (1), 17–30. <https://doi.org/10.1080/0158037X.2011.608656>.
- Bozalek, V., Ng'ambi, D., Gachago, D., 2013. Transforming teaching with emerging technologies: implications for higher education institutions. *S. Afr. J. High. Educ.* 27 (2), 419–436. Available from. <https://core.ac.uk/download/pdf/83123732.pdf>.
- Burbules, N.C., Torres, C.A., 2000. Globalization and education: an introduction. In: Torres, N.B. (Ed.), *Globalization and Education: Critical Perspectives*. Routledge, New York, pp. 1–26.
- Clegg, S., 2008. “Academic identities” under threat? *Br. Educ. Res. J.* 34, 329–345.
- Chalmers, D., 2011. Progress and challenges to the recognition and reward of the scholarship of teaching in higher education. *High. Educ. Res. Dev.* 30 (1), 25–38. <https://doi.org/10.1080/07294360.2011.536970>.
- Clarke, C., Knights, D., Jarvis, C., 2012. A labour of love? Academics in business schools. *Scand. J. Manag.* 28 (1), 5–15. <https://doi.org/10.1016/j.scaman.2011.12.003>.
- Coldron, J., Smith, R., 1999. Active location in teachers' construction of their professional identities. *J. Curric. Stud.* 31 (6), 711–726.
- Connolly, P., 1998. “Dancing to the wrong tune”: ethnography generalization and research on racism in schools. In: Connolly, P., Troyna, B. (Eds.), *Researching Racism in Education: Politics, Theory, and Practice*. Open University Press, Buckingham, UK, pp. 122–139.
- Creswell, J.W., 1998. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage, Thousand Oaks, CA.
- Dall'alba, G., 2009. Learning professional ways of being: ambiguities of becoming. *Educ. Philos. Theor.* 41 (1), 34–45. <https://doi.org/10.1111/j.1469-5812.2008.00475.x>.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2000. *The Handbook of Qualitative Research*. Sage Publications Ltd, Thousand Oaks, CA.
- Farrell, T.S.C., 2011. Exploring the professional role identities of experienced ESL teachers through reflective practice. *System* 39 (1), 54–62.
- Fletcher, S., 2000. A role of imagery in mentoring. *Career Dev. Int.* 5 (5), 235–243.
- Flores, M.A., Day, C., 2006. Contexts which shape and reshape new teachers' identities: a multi-perspective study. *Teach. Teach. Educ.* 22, 219–232.
- Freese, A., 2006. Reframing one's teaching: discovering our teacher selves through reflection and inquiry. *Teach. Teach. Educ.* 22, 110–119.
- Gee, J.P., 2001. Identity as an analytic lens for research in education. In: Secada, W.G. (Ed.), *Review of Research in Education*. American Educational Research Association, Washington, DC, pp. 99–125, 25.
- Giroux, H.A., 2010. Rethinking education as the practice of freedom: Paulo Freire and the promise of critical pedagogy. *Pol. Futures Educ. Internet* 8 (6), 715–721. <https://doi.org/10.2304/pfie.2010.8.6.715>.
- Goldhaber, D., Anthony, E., 2004. Can Teacher Quality Be Effectively Assessed? Available from. http://www.crpe.org/workingpapers/pdf/NBPTSquality_report.pdf.
- Goodyear, P., Banks, S., Hodgson, V., McConnell, D., 2004. *Advances in Research on Networked Learning*. Kluwer, Norwell, MA.

- Hamman, D., Gosselin, K., Romano, J., Rommel, B., 2010. Using possible-selves theory to understand the identity development of new teachers. *Teach. Teach. Educ.* 26, 1349–1361.
- Harasim, L., Hiltz, R., Teles, L., Turoff, M., 1995. *Learning networks: A field guide to teaching and learning online*. MIT Press, Cambridge, MA.
- Hermans, H., Dimaggio, G., 2007. Self, identity, and globalization in times of uncertainty: a dialogical analysis. *Rev. Gen. Psychol.* 11 (1), 31–61.
- Herman, N., Cilliers, F.J., 2008. Time+Research=Prestige+Money. Time+Teaching≠Prestige or Money. Why do academics bother to develop their teaching?. In: Paper presented at the Conference of the International Consortium for Educational Development (ICED), Salt Lake City, Utah.
- Hudson, B., Hunter, D., Peckham, S., 2019. Policy failure and the policy-implementation gap: can policy support programs help? *Pol. Des. Pract.* 2 (1), 1–14. <https://doi.org/10.1080/25741292.2018.1540378>.
- Jaffer, S., Ng'ambi, D., Czerniewicz, L., 2007. The role of ICTs in higher education in South Africa: one strategy for addressing teaching and learning challenges. *Int. J. Educ. Dev. ICT* 3 (4), 131–142. <https://doi.org/10.1080/13601440801962949>.
- Jawitz, J., 2009. Academic identities and communities of practice in a professional discipline. *Teach. High. Educ.* 14 (3), 241–251. <https://doi.org/10.1080/13562510902898817>.
- Johns, C., 1995. The value of reflective practice for nursing. *Clin. Nurse* 4, 23–30.
- Karjalainen, A., Nissilä, S.P., 2008. Designing and piloting 60-Ects credit teacher education program for university teachers. In: Gerber, K. (Ed.), *Personal—und organisationsentwicklung in Einrichtungen der Lehre und Forschung 5*. Universitätsverlag Webler, Bielefeld.
- Knapper, C., 2003. Three decades of educational development. *Int. J. Acad. Dev.* 8 (1–2), 5–9. <https://doi.org/10.1080/1360144042000277892>.
- Korthagen, F., Vasalos, A., 2005. Levels in reflection: core reflection as a means to enhance professional growth. *Teach. Teach.* 11, 47–71.
- Kreber, C., 2010. Academics' teacher identities, authenticity and pedagogy. *Stud. High. Educ.* 35 (2), 171–194.
- Laat, M., Schreurs, B., 2013. Visualizing informal professional development networks: building a case for learning analytics in the workplace. *Am. Behav. Sci.* 57 (10), 1421–1438. <https://doi.org/10.1177/0002764213479364>.
- Lamote, C., Engels, N., 2010. The development of student teachers' professional identity. *Eur. J. Teach. Educ.* 33 (1), 3–18.
- Lasky, S., 2005. A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teach. Teach. Educ.* 21 (8), 899–916. <https://doi.org/10.1016/j.tate.2005.06.003>.
- Laurila, T., Kukkonen, J., 2003. Evolving Professional Identity: Awareness of Self-Discrepancies as a Spur for Learning. Proceedings of the Conference "Teaching and Learning in Higher Education: New Trends and Innovations". University of Aveiro, April, 13–17.
- Lea, M.R., Stierer, B., 2009. Lecturers' everyday writing as professional practice in the university as workplace: new insights into academic identities. *Stud. High. Educ.* 34 (4), 417–428.
- Leibowitz, B., Bozalek, V., van Schalkwyk, S., Winberg, C., 2014. Institutional context matters: the professional development of academics as teachers in South African higher education. *High. Educ.* 69 (2), 315–330. <https://doi.org/10.1007/s10734-014-9777-2>.
- Leibowitz, B., Adendorff, H., Daniels, S., Loots, A., Nakasa, S., Ngxabazi, N., Van der Merwe, A., Van Deventer, I., 2005. The relationship between identity, language and teaching and learning in higher education in South Africa. *J. Lang. Learn.* 21 (2), 23–37. <https://doi.org/10.5785/21-2-74>.
- Loughran, J., 2002. Effective reflective practice: in search of meaning in learning about teaching. *J. Teach. Educ.* 53 (1), 33–43.
- MacLure, M., 1993. Arguing for yourself: identity as an organising principle in teachers' jobs and lives. *Br. Educ. Res. J.* 19 (4), 311–322.
- Markus, H., Nurius, P., 1986. Possible selves. *Am. Psychol.* 41 (9), 954–969.
- McNaughton, S., Billot, J., 2016. Negotiating academic teacher identity shifts during higher education contextual change. *Teach. High. Educ.* 1, 15. <https://doi.org/10.1080/13562517.2016.1163669>.
- Mok, K.H., Welch, A. (Eds.), 2003. *Globalization and Educational Restructuring in the Asia Pacific Region*. Palgrave Macmillan, Basingstoke, Hampshire.
- Phyak, P., Baral, R., 2019. Teacher identity in higher education: a phenomenological inquiry. *Batuk* 5 (2), 74–86. <https://doi.org/10.3126/batuk.v5i2.30119>.
- Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools, and academic achievement. *Econometrics* 73, 417–458. Available from: http://edpro.stanford.edu/Hanushek/files_det.asp?FileId=73.
- Ruyter, D., Conroy, J., 2002. The formation of identity: the importance of ideals. *Oxf. Rev. Educ.* 28 (4), 509–522.
- Scott, I., Yeld, N., Hendry, J., 2007. *Higher Education Monitor 6: A Case for Improving Learning and Teaching in South African Higher Education*. Council on Higher Education, Pretoria.
- Şebnem, S., 2009. The case study research method in translation studies. *Interpreter Transl. Train.* 3 (1), 37–56. <https://doi.org/10.1080/1750399X.2009.10798780>.
- Shavelson, R.J., Towne, L. (Eds.), 2002. *Scientific Research in Education*. National Academy Press, Washington, DC. Available from: <http://www.nap.edu/catalog/10236.html>.
- Sin, C., 2014. The policy object: a different perspective on policy enactment in higher education. *Int. J. High. Educ. Educ. Plan.* 68 (3), 435–448. <https://doi.org/10.1007/s10734-014-9721-5>.
- Smith, E.E., Kanuka, H., 2018. Transdisciplinary or pedagogically distinct? Disciplinary considerations for teaching certificates in higher education. *Int. J. Teach. Learn. High. Educ.* 30 (3), 388–401. Available from: <https://files.eric.ed.gov/fulltext/EJ1199278.pdf>.
- Stenberg, K., 2010. Identity work as a tool for promoting the professional development of student teachers. *Reflective Pract.* 11, 331–346.
- Strage, A., 2008. Traditional and non-traditional college students' descriptions of the "ideal" professor and the "ideal" course and perceived strengths and limitations. *Coll. Student J.* 42 (1), 225–231.
- Svennen, A., Volman, M., van Essen, M., 2008. The development of the professional identity of two teacher educators in the context of Dutch teacher education. *Eur. J. Teach. Educ.* 31 (2), 169–184.
- Thomas, L., Beauchamp, C., 2011. Understanding new teachers' professional identities through metaphor. *Teach. Teach. Educ.* 27 (4), 762–769.
- Trautwein, C., 2018. Academics' identity development as teachers. *Teach. High. Educ.* 23 (8), 995–1010. <https://doi.org/10.1080/13562517.2018.1449739>.
- Van der Linden, J., Erkens, G., Schmidt, H., Renshaw, P., 2000. Collaborative learning. In: *New Learning*. Springer Netherlands, pp. 37–54.
- Van Lankveld, T., Schoonenboom, J., Volman, M., Croiset, G., Beishuizen, J., 2017. Developing a teacher identity in the university context: a systematic review of the literature. *High. Educ. Res. Dev.* 36 (2), 325–342. <https://doi.org/10.1080/07294360.2016.1208154>.
- Verhofstadt-Deneve, L., 2003. The psychodramatical "social atom method": dialogical self in dialectical action. *J. Constr. Psychol.* 16, 183–212.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, New York.
- Yin, R.K., 1994. *Case Study Research: Design and Methods*, second ed. Sage Publications, Newbury Park, CA.
- Yin, R.K., 2003. *Case Study Research: Design and Methods*, third ed. Sage, Thousand Oaks, CA.

Early career academics in Namibia: experiences, prospects and challenges

Beatrice S. Sichombe^a, Alina K. Niipare^b, and Jafet S. Uugwanga^b, ^a Namibian College of Open Learning, Katutura, Namibia; and ^b National Institute for Educational Development, Okahandja, Namibia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	513
Early career researchers in context	513
Literature on early career researchers: experiences, challenges and prospects	514
Methodology	515
Namibian ECRs' experiences, prospects and challenges	515
Discussion of findings	517
Conclusion and recommendations	518
References	518

Introduction

Early Career Academics (ECAs) are not new to research. Countries like Australia, New Zealand, UK, Japan, the USA, Canada and South Africa, among others, are way ahead with research. Scholars have opened conversations on various aspects, ranging from support and guidance, capacity building, funding, workload, retention of academics, career development and challenges ECAs encounter. Entering the academic world is challenging to ECAs. Nevertheless, higher education institutions expect their academic staff to publish papers as it benefits them and the wider academy. The need to understand ECAs' experiences, prospects and challenges cannot be overemphasized. So much is written on ECAs in other countries and contexts, but little remains known about Namibian ECAs. There is a dearth of literature on the issue. Due to the scarcity of literature and conversations in Namibia on ECAs and reflecting on our journeys as Early Career Researchers (ECRs's), we decided to open up a conversation with Namibian ECAs. And to address this gap, we share in this article the Namibian ECRs' experiences concerning their development into experienced researchers, including their prospects of becoming experienced researchers, as well as challenges hindering them from becoming experienced researchers.

Drawing on the popular aphorism 'publish or perish, this article draws attention to the challenge of ECRs' capacity to publish research papers in peer-reviewed journals (Shinkafi, 2020) in order to develop experienced researchers (Akerlind, 2008). Scholars continue to converse on the kind of support ECRs need to research and publish. Scholars argue that ECRs are inexperienced researchers who require guidance, support, and mentoring (Shinkafi, 2020) to excel in academia.

The article begins with conceptualizing ECRs by putting it into context and distinguishing between two related terms, ECAs and ECRs. We use a narrative method embedded with quantitative data to tell the story of Namibian ECRs' journeys of becoming seasoned researchers. We then explain the methodology informing our research, and ECRs' experiences, prospects, and challenges are thereafter presented and discussed.

Early career researchers in context

There is no agreed definition of what constitutes an ECA (Price et al., 2015 in Masinga et al., 2016). McAlpine and Åkerlind (2010) further exacerbate the complexity of the term ECA in terms of what they do and how they develop as academics. Hemmings (2012) views ECAs as academics who are within their first five years as members of staff in Higher Education Institutions (HEIs) those with limited experience in academia (Price et al., 2015). Similarly, Vajoczki et al. (2011, p. 75) defines ECAs as those in the first five years of their academic appointment, the period they should develop an academic profile in teaching, research and administration. However, the nature of appointments differ from post-doctoral researchers or teaching fellows to probationary academics, pre-tenured faculty members, to Doctor of Philosophy (PhD) students with teaching responsibilities (Lesenyeho et al., 2018).

Researchers do not draw a clear distinction between the dyads ECAs and ECRs, and some tend to use the two terms interchangeably. However, Haddow and Hammarfelt (2019) indicate that ECAs are those 'heavily' involved in teaching after obtaining their PhDs; whereas, ECRs are the ones that spend 'most' of their time on research. This article is very specific on ECRs and focuses on post-doctoral researchers who are yet to develop their academic profile in research (Vajoczki et al., 2011).

We adopt Bazeley (2003, p. 274) definition of an ECR as one "who is currently within his/her first five years of academic or other research-related employment allowing uninterrupted, stable research development following completion of their postgraduate research training". Developing as a researcher is understood as a process comprising four categories, characterized by confidence, recognition, productivity and sophistication (Akerlind 2008). Becoming a confident researcher entails that one has acquired skills to do research successfully and the confidence of being on the right track with one's research. The category of becoming a recognized researcher, Akerlind (2008) explains, is marked by recognition by other scholars of a researcher's contribution in one's field of study and to the academic community at large. The next category, becoming a productive researcher is marked by increasing productivity

and accumulation of research over time, and this is measured by one's research output. The last category, becoming a sophisticated researcher is increasing one's knowledge horizons and theoretical understanding, and increased sophistication as a researcher.

A sophisticated researcher is an elite (Dalton et al., 1977) and a leading researcher (Woolley et al. 2016). When researchers reach the 'experienced researcher' stage, it means that they have reached the 'most complex' understanding and a great enhancement in doing research (Åkerlind, 2008). Reaching this stage means that a researcher has developed through the three stages of being a confident researcher, a recognized researcher, and a productive researcher (Åkerlind, 2008). Experienced researchers do better research and become research experts (Åkerlind, 2008). In addition, they shape the direction of the knowledge production of their community (Gläser and Laudel, 2015) and develop a strategic vision on the future of the research field (Woolley et al., 2016). Furthermore, experienced researchers contribute substantially to their research field (Woolley et al., 2016) since they reach a theoretical understanding and become aware of broader perspectives (Åkerlind, 2008). Finally, they publish and present influential research papers and books (Woolley et al., 2016).

Literature on early career researchers: experiences, challenges and prospects

ECRs irrespective of whether in Africa or developed countries experience challenges during their journeys of becoming seasoned researchers (Shinkafi, 2020). However, Africa is worse, as Shinkafi (2020) argues, institutions do not pay much attention to the development of ECRs. This could be attributed to the magnitude of challenges African ECRs face as compared to their counterparts in developed countries (Tumwiyukye et al., 2013).

ECRs in Africa struggle gaining funds even if they have a "solid range of knowledge and skills" (Bell et al., 2016, p. 3) in their disciplines. According to Shinkafi (2020), research funds for ECRs in Africa are limited and most funding agencies only fund senior researchers. According to Locke et al. (2018), choosing a career inside academia due to the high pressure to publish in high impact journals is challenging to young ECRs. Resultantly, some highly skilled ECRs who opt for a career outside of academia (Flinders and Anderson, 2019). Stiles and McCulloch (2017) state that collaborative work environments and flexibility are some of the reasons why ECRs prefer to pursue careers outside of academia. Therefore, a systemic change to research cultures within academia is needed in order to retain young academics in academia without being more competitive and exploitative (Bailey et al., 2020).

Brennan et al. (2020) claim that a number of universities worldwide adopt 'teaching, research and administration' approaches. For instance, in most universities in the United Kingdom and overseas, ECRs are expected to teach, conduct scientific research and demonstrate leadership as part of the universities' development (Bailey et al., 2020, p. 2). This approach does not benefit ECAs as they miss out on research opportunities due to their teaching workload (Graham, 2015). Thus, a UK based study revealed that in most universities teaching without integrating research adds no value into lecturers' promotion requirements and this frustrates young academics when they are unable to conduct research (Graham, 2015).

Career instability is one of the stressful aspects that ECRs are faced with (Heaney et al., 2019). According to Bozzon et al. (2016) and Christian et al. (2021) many postdoctoral fellows in academic settings encounter job insecurity. They also have difficulties in acquiring better permanent posts or contract posts even outside of academia. Bozzon et al. (2016) stress that a higher degree of job instability and uncertainty in academic context is not necessarily as a result of the inexperience of ECRs but mostly constituted by a "growing competition for permanent positions and, as a consequence, by a greatly increased pressure" (p. 3). Furthermore, Heaney et al. (2019) indicate that in many countries, ECRs are normally "employed in insecure forms of employment" and offered the so-called 'no tenured' on their first position and this comes with conditions (Nissen et al., 2020, p. 9). According to Heaney et al. (2019), some of the conditions are that ECRs are required to source research funds and publish a certain number of articles regardless of quality. And if they fail to meet the requirements, they can certainly lose their jobs. And with such conditions, many ECRs might see themselves as incompetent and "often feel hopeless" (Heaney et al., 2019, p. 10). As a result, ECRs are left with no option rather than withdraw from leaving their research careers (Christian et al., 2021) and search for new employment opportunities outside of academia.

Mobility of ECRs within academia is another challenge. Despite the availability of job opportunities, Guthrie et al. (2017, p. 5) indicate that "the desire to develop international research networks and collaborations, and to access specific expertise, resources or prestige that allow researchers to progress their careers" are some of the factors that drive ECRs to mobility in academia. In other words, job security, availability of research funding and grants are the major drivers of researcher mobility choices (Guthrie et al., 2017). According to Pugh and Thomas (2018), mobility raises up many concerns ranging from personal lives to research and professional lives. Thus, research mobility is lower in women as compared to men in academia because of family commitments. As a result, young female researchers find it difficult to navigate the globe in searching of career opportunities to broaden their research horizons (Pugh and Thomas, 2018).

Developing the skill of an independent researcher requires a good knowledge base and abilities (Laudel and Gläser, 2007) particularly good writing skills, yet a skill always lacking among postgraduates candidates, coupled with good writing skills is publishing (Almatarneh et al., 2018).

Publishing is considered an important activity for academics and universities as a measure for performance and a criterion for academic promotion and competitive research funding (Kamler, 2008). Hemmings and Kay, (2007) demonstrated that lecturers who publish are appointed in 'more' senior academic positions. Similarly, Kamler (2008) argues that unpublished research results deny the academic field opportunities of professional dialog and knowledge building. Despite the irreplaceable role publications play, publication outputs remain generally low (McGrail et al., 2006; Lee and Kamler, 2007) due to lack of momentum, motivation and confidence, and a lack of framework to support writing (McGrail et al., 2006), as well as barriers such as workload, lack of support, and an under-developed research culture (Hemmings et al., 2007).

The role networking plays in academia cannot be overemphasized. Scholars refer to networking within academia as ‘predictor of career success’ (Blickle et al., 2009), and a way of ‘getting ahead’ (Wolff and Moser, 2009) but not easily acquired by ECRs (Camedda et al., 2017). However, due to the well-researched need of networking skills to ECRs (de Janasz and Forret, 2008; Ansmann et al., 2014) the acquisition of these skills should be one of the priorities of ECRs. However, scholars argue that networking does not happen by itself; it develops through concerted effort and participation in dedicated-focused initiatives (Camedda et al., 2017).

Kirk and Lipscombe (2019) alluded that novice researchers are faced with many challenges during publications and co-authorship opportunities are regarded as the positive developmental strategies to support ECRs in their research careers. Broström (2019) stresses that for the ECRs to receive credit for the work published; they should rather co-author with respected supervisors who gained reputation on research work within the wider scientific community. However, co-authorship has its own challenges such as who should get an authorship credit for the published work between the ECRs and supervisors (Macfarlane et al., 2017). In many cases, “all persons designated as authors should qualify for authorship and all those who qualify should be listed” and share the authorship’s reputational credit (Albert and Wager, 2003, p. 34). In China for instance, ECRs keenly want to be the first or corresponding authors other than their supervisors, so that “their work can be recognized by their institutes and funders” (Nicholas et al., 2018, p. 208). And to be listed as the first or corresponding author; it means that such a person should take the full responsibility for the published research work (Albert and Wager, 2003). Moreover, Minhas, (2021) argues that there is a need to provide public awareness among the ECRs about the authorship ethics and criteria involved when ECRs and supervisors are co-authoring the research work for publications.

Methodology

The study used mixed methods research through a combination of qualitative and quantitative research approaches. The embedded mixed method design was used to gain an insight into ECRs’ personal experiences, prospects, as well as challenges faced on becoming experienced and seasoned researchers. This design embeds qualitative components within quantitative design (Cresswell, 2014) in a supportive secondary role which enables the researchers to make sense of the study as a whole (Almalki, 2016).

Data were collected through an on-line questionnaire and semi-structured interviews (McMillan and Schumacher, 2010). Semi-structured interviews were preferred as they allow for the modification of the questions during the interview (Ary et al., 2010). The data collection procedure was preceded by the piloting of the data collection instruments for validity purposes. The population consisted of Namibian postdoctoral candidates who completed their doctoral degrees within the past five years irrespective of the field of study and institution. Participants were either in HEIs or outside HEIs in both academic and non-academic environments.

Snowball sampling (Parker et al., 2019) was used to find qualifying participants. Three qualifying participants formed the initial contacts to which the questionnaire link was sent with a cover letter specifying ethical issues. These recommended other potential participants who were willing to be part of the study, and the process was replicated until saturation was reached (Guest et al., 2020). Saturation was determined when there were no more new respondents in two weeks. The questionnaire remained active for one month. Twenty-eight respondents completed the questionnaire. A database of respondents was kept and this assisted in identifying interviewees. Out of the total sample, 25% (7) were conveniently sampled (Etikan et al., 2016) on the basis of willingness and availability. The semi-structured interviews were done virtually via ZOOM as well as telephonic. Thematic content analysis was employed to analyze qualitative data; and Microsoft Forms, a web application, assisted with the analysis of quantitative data.

Namibian ECRs’ experiences, prospects and challenges

Research Profile: All participants were ECRs whose profiles ranged from one to eight years after graduation, with the newly graduates (1–3 years) being in the majority. Fields of study included Education, Curriculum Studies, Education Management, Humanities, Social Sciences, Public Health, Natural Sciences, and Mathematics Education. ECRs’ growth in research skills and knowledge since obtaining their doctoral degrees differed. Some were satisfied with their growth, while to others the growth was slow, or non-existent. Satisfactory growth was attributed to, among others, self-motivation and curiosity to learn, interest in research, and belief in continued growth beyond PhD. Those that experienced stagnation or slow growth attributed this to lack of incentives for research in Namibia, work overload particularly in academic institutions, lack of direction on how to proceed beyond PhD, lack of support and guidance, and lack of interest in research among others.

Publication status: As other scholars argue, publishing is considered an important activity for academics and universities; it is a measure for performance and a criterion for academic promotion and competitive research funding (Kamler, 2008). The majority (66.7%) of the participants’ dissertations were published, and only 33.3% were not. Out of the published theses, 11.1% published whole dissertations and the majority (89.9%) published articles from their dissertation.

Research milestones: Among milestones reached, publishing was flagged by many (27%); followed by student supervision (23%), and 22% collaborating with supervisors. The lowest rated (10%) milestone was the other category, for instance tutoring and assisting postgraduate students with research work, and thesis examination. In addition to the ECRs’ milestones reached after obtaining their doctorate degree, the researchers also wanted them to narrate their mentoring and supervision ability in relation to

their research capacity building. Some participants narrated their successful achievements of supervising student research at both undergraduate and postgraduate levels with 100% completion rate.

Furthermore, the researchers wanted to know whether participants were envisaging becoming research supervisors or mentors. All participants eagerly envisaged becoming research supervisors or to continue supervising students. Enhancement of research knowledge and skills was cited by many; others viewed this from the support and guidance to students' perspective as they felt that being a research supervisor or mentor you render the needed support and motivation to students.

Research support: To become an experienced researcher, supervisor or mentor, scholars (Locke et al., 2018; Shinkafi, 2020) argue for research support and guidance. Some ECRs attested that they had not been receiving any support and guidance. Lack of support and guidance was attributed to non-academic environments by others. They felt that non-academic environments do not provide research support and training unlike academic environments. In the absence of support, others had devoted themselves to self-support initiatives. Others attested that they had been receiving support whose quality was at varying degrees. A total of 60% of the participants rated the quality of research support they received as poor to very poor. However, the remaining 40% were satisfied and rated the support as good (25.7%) to quite good (14.3%).

Other research related skills: It seems that not only research support is needed by ECRs to become experienced researchers, research supervisors or mentors, but other research related skills too as part of their growth in academia. Research related skills included among others, writing research proposals for grants and funding, enhancement on qualitative and quantitative research methods, referencing styles, advances in Information and Communication Technology (ICT) use for data analysis, and various theories and practices within the research discipline. The need for refresher workshops was cited by many as a platform for the enhancement of research related skills. ECRs felt that through such workshops, researchers share research knowledge, skills and best practice, and funding opportunities.

Motivation to study for a doctoral degree: Participants who were motivated to pursue their doctoral degrees because they wanted to improve their career prospects to become researchers topped the list (20.8%); this was followed by a tie of those that were encouraged by their former academic supervisors and the prestige of becoming a doctor in an academic field (18.9%). 9.4% studied because of available funding, and 7.5% because the opportunity to do a doctoral degree was available to them. The remaining reasons, interest in the subject, improving prospects of becoming lecturer, considering it as a natural step, and current job requirements each scored 5.7%. Each was the motivating factor. However, the available opportunity to do a doctoral degree was 9.4%. In spite of the stated motivating aspects, only 1.9% of the participants had other motivating reasons.

Career aspirations: More than half (50%) of the respondents aspired to work in higher education institutions after completing their doctoral degree focusing mainly on teaching and research. Equally, respondents who were interested to work in the institutions of higher learning primarily for research only and those that aspired to become self-employed totaled to 14.8% each. And 18.5% of the respondents aspired to take up teaching positions in higher education institutions.

Current and future job prospects: A tie of 18.2% was on research jobs outside higher education and teaching either in or outside higher education. 13.6% of the respondents indicated that their job in higher education included both teaching and research; and those who were self-employed and who were pursuing post-doc studies counted 9.1% each. The lowest percentage (4.5%) was those whose primary job in higher education was research. None of the respondents was unemployed. Future prospects in the next five years included aspirations of becoming associate professors/professors, research consultants (edit research papers, assist ECRs to publish), research supervisors, and experts in research methods with the ability to mentor ECRs.

Presence in other scholars' work: Only 27% of the participants indicated that their research work has been cited by other scholars as compared to the majority (50%) whose research work was not cited, and 23% who could not attest whether their research work was cited by other scholars or not.

Collaboration status: The majority (60%) of the participants were involved in collaborative research, and 40% were not. When asked which collaborators, the majority (67%) collaborated with experienced researchers, mainly to enhance their research knowledge and skills, and the ability to write good publishable research papers in accredited journals. The remaining 33% who collaborated with fellow ECRs had no other choice as experienced researchers and professors were not willing to collaborate with ECRs; others believed in the exchange of research knowledge and skills; while some wanted to expand their research horizons and gain new perspectives on research.

Stages of development: Participants had to indicate the stage at which they were on their journey to become experienced researchers. The majority (65%) of the participants were still at the basic stage of apprentice whereby they needed a mentor to do research; and 35% considered themselves at master stage as they had been mentoring others on research; none of the participants were at colleague and elite stages.

Current research project: The majority (60%) of the respondents were not, at the time of the study, working on a research project, and 40% were working on a research project. When further asked in which research journal would they wish to publish their research work, generally, the majority (90%) wished to published in peer reviewed journals, others specified some journals: Elsevier, Journal for Language Teaching (JLT), Namibian Educational Research Association (NERA), Francis and Taylor, and 10% of the participants had not yet decided.

Number of articles published: Half (50%) of the participants had each published two articles each; 30% each published one article; and a tie of 10% was recorded on participants that published four and five articles. All of these articles were published in peer-reviewed journals.

Non-dissertation-based publications: The researchers were also interested to know whether participants had published articles not based on their PhD dissertation since graduation. Slightly above 50% of the participants had published articles not based on their

dissertations, 40% had not and 7% had submitted articles to journals for possible publication. Of the non-dissertation published articles, 75% were published in peer reviewed journals, and the remaining 25% in non-peer reviewed journals. The set standards and requirements as well as the high cost of publication charges were commonly cited as obstacles. Other reasons included lack of confidence in research knowledge and skills, and the belief that they had to start at that level and move up the ladder of publishing.

Challenges: About a quarter (25.9%) of the participants indicated that all the listed challenges applied to them, namely, carrying independent research, publishing in accredited journals, support and guidance to write publishable research papers, research resources and networking with other scholars. This was followed by 18.5% who pointed at networking with other scholars, 15.8% who identified publishing in accredited journals, 14.8% pointed at research resources; carrying independent research and writing quality research papers without supervision received the lowest responses. Further engagement with them in interviews revealed similar challenges, namely, lack of funding, data analysis, ICT related challenges, and publishing. However, in academia, credit and recognition is given to academic papers published in accredited and reputable journals. In spite of the prestige that comes with publications in accredited journals, Namibian ECRs faced challenges and attributed this to ineffective, or non-existence of mentorship programs, lack of research collaboration opportunities with experienced researchers, lack of guidance and support, lack of funding opportunities, the intensity of comments from reviewers, shortfalls in academic writing genre, lack of publishing skills, and limited publications for ECRs to benchmark on, and the fear of the unknown.

Discussion of findings

Ideally, doctoral dissertations published are/should be cited by other scholars. This study, however, revealed that only a few dissertations were cited by other scholars. In addition, many ECRs collaborated with experienced researchers to gain research knowledge and skills and also the ability to write good publishable research papers in accredited journals. According to McGloin (2018), collaboration is the best practice as it encourages teamwork to share required research. To become an experienced researcher, there are stages to go through. And many participants were still at the basic stage of apprentice whereby they needed a mentor to do research and only few were at the master stage (Laudel and Glaser, 2008). This might be the reason why most of the participants were not working on a research project; even though they wished to publish many of their research papers in peer reviewed journals. According to Akerlind (2008) and Shinkafi (2020), ECAs were aspiring to develop into experienced and sophisticated researchers but lacked capacity to write good research papers and publish in relevant peer-reviewed journals.

To become an experienced researcher requires someone to conduct research and publish few research articles. This concurs with Akerlind (2008) and Shinkafi (2020) who argue that for ECRs to develop into experienced and sophisticated researchers they should be able to research and publish. Thus, the majority of the participants published articles from their dissertations compared to those who published whole dissertations. And this was a milestone reached by ECRs after obtaining their doctoral degrees. Despite the publications, some ECRs had also supervised students, did collaboration work with their supervisors, tutored students and assisted postgraduate students with research work, and thesis examination. For collaboration, McGloin (2018) admits that in some cases, collaboration frustrates and discourages researchers, especially the young academics. In spite of what have been accomplished by ECRs, those in non-academic environments had not been supported and guided after obtaining their doctoral degrees other than self-support initiatives. Furthermore, ECRs indicated that they needed research related skills like writing research proposals for grants and funding, enhancement on qualitative and quantitative research methods, referencing styles, advances in ICT use for data analysis, and various theories and practices within the research discipline. Similarly, Hlengwa (2019) and Marsay (2020) denote that lack of research grants and funding opportunities prevent ECRs outside academia from conducting research. In other words, unavailability of research funding and grants are the major drivers of researcher immobility choices (Guthrie et al., 2017). Participants also needed refresher workshops as a platform to share research knowledge and skills and best practices.

Many ECRs pursued their doctoral degrees because they wanted to improve their career prospects compared to other motivating factors such as being encouraged by their former academic supervisors and the prestige of becoming a doctor and work in higher education. Guthrie et al. (2017, p. v) indicate that “the desire to develop international research networks and collaborations, and to access specific expertise, resources or prestige that allow researchers to progress their careers” inspired students to pursue their doctoral degrees. And all ECRs had future prospects, aspiring to become associate professors, research consultants, research supervisors, and/or experts in research methods with the ability to mentor the upcoming ECRs.

Becoming an ECR comes with challenges especially when one aspires to publish in peer reviewed journals. Thus, in academia, credit and recognition is given to academic papers published in accredited and reputable journals. In support, Hemmings and Kay (2007) demonstrate that lecturers who research and publish are more confident and can be appointed in ‘more’ senior academic positions than those who do not. However, the set standards, requirements and high cost of publication charges were cited by ECRs as obstacles to having their research work published in accredited journals. In spite of the prestige that comes with publications in accredited journals, Namibian ECRs faced challenges and attributed this to ineffective, or non-existence of mentorship programmes, lack of research collaboration opportunities with experienced researchers, lack of guidance and support, and lack of funding opportunities. In support, Bell et al. (2016) state that ECRs are faced with numerous academic challenges including gaining funds even if the ECRs have “solid range of knowledge and skills” in their disciplines (p. 3). In addition, Hlengwa (2019) and Marsay (2020) state that more teaching responsibilities with fewer resources and funding opportunities for research are some of the challenges that prevent ECRs from conducting research.

Conclusion and recommendations

References

- Akerlind, S., 2008. Growing and developing as a university researcher. *High Educ.* 55 (2), 241–254.
- Albert, T., Wager, E., 2003. How to Handle Authorship Disputes: A Guide for New Researchers. The COPE Report 2003. Committee on Publication Ethics.
- Almalki, S., 2016. Integrating Quantitative and Qualitative Data in Mixed Methods Research: Challenges and Beliefs. Taif University, Saudi Arabia.
- Almatameh, A.H.M., Rashid, R.A., Yunus, K., 2018. The academic writing experience of Jordanian postgraduate students at a University in Malaysia. *Arab World English Journal (AWEJ)*. 9 (3), 248–257. <https://doi.org/10.24093/awej/vol9no3.17>.
- Ansmann, L., Flickinger, T.E., Barello, S., Kunnean, M., Mantwill, S., Quilligan, S., Zanini, C., Aelbrecht, K., 2014. Career development for early career academics: benefits of networking and the role of professional societies. *Patient Educ. Couns.* 97 (1), 132–134.
- Ary, D., Jacobs, L.C., Razavieh, A., Sorensen, C.K., 2010. *Introduction to Research in Education*, 8 ed. Hult Rinchart & Wiston, New York, NY.
- Bailey, L., Behar, V., Bohmert, D., Järbur, C., Klee, D., 2020. Boost the Careers of Early Stage Researchers. CESAER, Leuven.
- Bazeley, P., 2003. Defining 'early career' in research. *High Educ.* 45 (3), 257–279.
- Bell, J., Ancillotti, M., Coathup, V., Coy, S., Rigter, T., Tatum, T., Grewal, J., Akcemes, F.B., Brkić, J., Causevic-Ramosevac, A., Milovanovic, G., Nobile, M., Pavlidis, C., Finlay, T., Kaye, J., 2016. Challenges and opportunities for ELSI early career researchers. *BMC Med. Ethics* 17 (37), 1–6.
- Blickle, G., Witzki, A., Schneider, P.B., 2009. Self-initiated mentoring and career success: a predictive field study. *J. Vocat. Behav.* 74 (1), 94–101.
- Bozzon, R., Murgia, A., Poggio, B., 2016. Supporting Early Career Researchers through Gender Action Plans: A Design and Methodological Toolkit (vol. 9). University of Trento.
- Brennan, N.M., Chapman, A., Zipin, L., Rudolph, S., Barron, R., Rogers, B., Woods, A., 2020. The Growing Urgency of Attending the State of Education Research in Australian Higher Education. Report to the AARE Executive and Members from the Working Party 'Protecting and Extending Research in Education in Australia'. Australian Association for Research in Education.
- Broström, A., 2019. Academic breeding grounds: home department conditions and early career performance of academic researchers. *Res. Pol.* 48 (7), 1647–1665.
- Camedda, D., Mirman-Flores, A., Ryan-Mangan, A., 2017. Young Researchers Need Help with Academic Networking.
- Christian, K., Johnstone, C., Larkins, J.A., Wright, W., Doran, M.R., 2021. Research Culture: a survey of early-career researchers in Australia. *Elife* 10, e60613.
- Cresswell, J.W., 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, fifth ed. Pearson Education Inc, Boston.
- Dalton, G., Thompson, P., Raymond, P., 1977. The four stages of professional careers - a new look at performance by professionals. *Organ. Dynam.* 6 (1), 19–42.
- de Janasz, S.C., Forret, M.L., 2008. Learning the art of networking: a critical skill for enhancing social capital and career success. *J. Manag. Educ.* 32 (5), 629–650.
- Etikan, I., Musa, S.A., Alkassim, R.S., 2016. Comparison of convenience sampling and purposive sampling. *Am. J. Theor. Appl. Stat.* 5 (1), 1–4.
- Flinders, M., Anderson, A., 2019. Fit for the Future? Researcher Development and Research Leadership in the Social Sciences. Economic & Social Research Council: University of Sheffield.
- Gläser, J., Laudel, G., 2015. Three Careers of Academics. Discussion paper. Retrieved from: http://www.tu-berlin.de/ztg/menue/publikationen/discussion_papers/. (Accessed 11 September 2021).
- Graham, R., 2015. Does Teaching Advance Your Academic Career? Perspectives of Promotion Procedures in UK Higher Education. Royal Academy of Engineering.
- Guest, G., Namey, E., Chen, M., 2020. A simple method to assess and report thematic saturation in qualitative research. *PLoS One* 15 (5), 1–17.
- Guthrie, S., Lichten, C., Corbett, J., Wooding, S., 2017. *International Mobility of Researchers. A Review of the Literature*. Rand Corporation.
- Haddow, G., Hammarfelt, B., 2019. Early career academics and evaluative metrics: ambivalence, resistance and strategies. In: *The Social Structures of Global Academia*. Routledge, pp. 123–143.
- Heaney, L., Wong, A., Martel, C., Rawlins, C., 2019. How to Transition from Early Career to Established. Retrieved from: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Heaney%2C+Wong%2C+Martel+%26+Rawlins+2019&btnG=-. (Accessed 25 July 2021).
- Hemmings, B.C., 2012. Affective solidarity: feminist reflexivity and political transformation. *Fem. Theor.* 13 (2), 147–161.
- Hemmings, B., Kay, R., 2007. I'm sure I can write! Writing confidence and other factors which influence academic output. In: *European College Teaching and Learning Conference*. Ljubljana, Slovenia.
- Hemmings, B.C., Rushbrook, P., Smith, E., 2007. Academics' views on publishing refereed works: a content analysis. *High Educ.* 54 (2), 307–332.
- Hlengwa, A., 2019. How are institutions developing the next generation of university teachers? *Crit. Stud. Teach. Learn.* 7 (1), 1–18.
- Kamler, B., 2008. Rethinking doctoral publication practices: writing from and beyond the thesis. *Stud. High Educ.* 33 (3), 283–294.
- Kirk, M., Lipscombe, K., 2019. When a postgraduate student becomes a novice researcher and a supervisor becomes a mentor: a journey of research identity development. *Stud. Teach. Educ.* 15 (2), 179–197.

- Laudel, G., Gläser, J., 2007. Interviewing scientists. *Sci. Technol. Innovat. Stud.* 3 (2).
- Laudel, G., Gläser, J., 2008. From apprentice to colleague: the metamorphosis of Early Career Researchers. *High Educ.* 55, 387–406.
- Lee, A., Kamler, B., 2007. Bringing pedagogy to doctoral publishing. *Teach. High. Educ.* 13 (5), 511–523.
- Lesenyehlo, D.L., Barkhuizen, N.E., Schutte, N.E., 2018. Factors relating to the attraction of talented early career academics in South African higher education institutions. *SA J. Hum. Resour. Manag.* 16 (1), 1–9.
- Locke, W., Freeman, R., Rose, A., 2018. Early Career Social Science Researchers: Experiences and Support Needs. Centre for Global Higher Education, UCL Institute of Education.
- Macfarlane, B., Devine, E., Drake, T., Gilbert, A., Robinson, M., White, I., 2017. Co-authorship in Humanities and the Social Sciences: A Global View.
- MacMillan, G.A., Falardeau, M., Girard, C., Dufour-Beauséjour, S., Lacombe-Bergeron, J., Menzies, A.K., Henri, D.A., 2019. Highlighting the potential of peer-led workshops in training early-career researchers for conducting research with Indigenous communities. *FACETS* 4, 275–292.
- Marsay, E., 2020. COVID-19 Report: Inequalities in Academia—Impact on Early Career Researchers. Retrieved from: <https://wbg.org.uk/wp-content/uploads/2020/11/Early-Career-Researchers-Covid-19-report-.pdf>. (Accessed 23 September 2021).
- Masinga, L., Myende, P., Marais, A., Singh-Pillay, A., Kortjass, M., Chirikure, T., Mwelli, P., 2016. "Hear our voices": a collective arts-based self-study of early-career academics on our learning and growth in a research-intensive university. *Educ. Res. Soc. Change* 5 (2), 117–135.
- McAlpine, L., Åkerlind, G.S., 2010. Academic Practice in a Changing International Landscape. *Becoming an academic: International Perspectives*, pp. 1–17.
- McGloin, S., 2018. Reconceptualising the supervisory role: border-crossings and bounded spaces. In: *Book: Spaces, Journeys and New Horizons for Postgraduate Supervision*. African Sun Media, p. 109.
- McGrail, M.R., Rickard, C.M., Jones, R., 2006. Publish or perish: a systematic review of interventions to increase academic publication rates. *High Educ. Res. Dev.* 25 (1), 19–35.
- McMillan, J.H., Schumacher, S., 2010. *Research in Education: Evidence-Based Inquiry*. MyEducationLab Series. Pearson.
- Minhas, N.M., 2021. Authorship ethics: An overview of research on the state of practice. In: *International Workshop on Ethics in Software Engineering Research and Practice*, pp. 31–38. <https://doi.org/10.1109/SEthics52569.2021.00013>.
- Nicholas, D., Abdullah, A., Boukacem-Zeghmouri, C., Clark, D., Herman, E., Bravo, B.R., Świgoń, M., Watkinson, A., Xu, J., 2018. Early Career Researchers: The Harbingers of Change? CIBER Research LTD.
- Nissen, S., Naepi, S., Powell, D., Baker, T., Bolton, A., Stewart, L., 2020. Early Career Researchers in Aotearoa: Safeguarding and Strengthening Opportunity after COVID-19. Early Career Researcher (ECR) Forum of Royal Society Te Apārangi.
- Parker, C., Scott, S., Geddes, A., 2019. *Snowball Sampling*. SAGE Research Methods Foundations.
- Price, R.B., Kuckertz, J.M., Siegle, G.J., Ladouceur, C.D., Silk, J.S., Ryan, N.D., Dahl, R.E., Amir, N., 2015. Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychol. Assess.* 27 (2), 365.
- Pugh, R., Thomas, E., 2018. Moving, settling and becoming: a conversation about mobility between two early career researchers. *Regions Magazine*.
- Shinkafi, T.S., 2020. Challenges experienced by early career researchers in Africa. *Future Science: OA* 1–4.
- Stiles, J., McCulloch, C., 2017. CADRE Early Career Guide: Tips for Early Career STEM Education Researchers. Retrieved from: <http://cadrek12.org/resources/cadre-early-career-guide-tips-early-career-stem-ed-researchers>. (Accessed 10 August 2021).
- Tumwijukye, H., Motevalli, M., Nambi, S., Kyeyune, I., 2013. Developing African novice researchers into career investigators: innovative options. *J. Res. Adm.* 44 (2), 57–73.
- Vajoczki, S., Biegas, T.C., Crenshaw, M., Healey, R.L., Osayomi, T., Bradford, M., Monk, J., 2011. Professional development in teaching and learning for early career academic geographers: contexts, practices and tensions. *J. Geogr. High Educ.* 35 (3), 395–408.
- Wolff, H.G., Moser, K., 2009. Effects of networking on career success: a longitudinal study. *J. Appl. Psychol.* 94 (1), 196.
- Woolley, R., Cañibano, C., Tesch, J., 2016. A Functional Review of Literature on Research Careers. Ingenio (CSIC-UPV) working paper series 2016-05.

FOURTH EDITION

International Encyclopedia of Education

Inclusive Education and Disability
Studies in Education

VOLUME
9

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editor

Roger Slee



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 9

Inclusive Education and Disability Studies in Education

VOLUME 9 EDITOR

Roger Slee



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

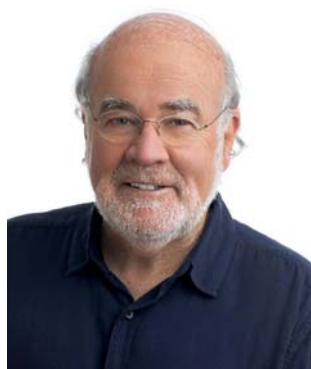
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



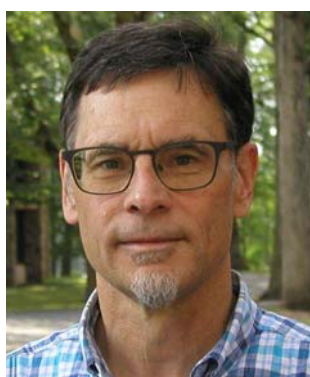
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 9: Inclusive Education and Disability Studies in Education

Roger Slee

© 2023 Elsevier Ltd. All rights reserved.

Dedication

This volume of the *International Encyclopedia of Education (4th Edition)* is dedicated to the memory of Dr Anne Hughson, Professor Emerita of the Cumming School of Medicine, The University of Calgary. Psychologist, educator, researcher, disability rights advocate, and community leader, Anne established the first postsecondary education initiative for Canadian disabled people in her home province of Alberta.

Introduction: Education for All in New Times

In *Inadvertent* Karl Ove Knausgaard, a Norwegian writer, engages with the simple question: Why do I write? He admits to “...sitting in front of my desk in southern Sweden for three days without making headway” (Knausgaard, 2018: 1). “Simplicity”, he tells us, “Is treacherous” (Knausgaard, 2018: 1). At first glance inclusive education would appear conceptually simple, elegantly compelling. After all, the right to education is enshrined in international agreements such as the United Nations Convention on the Rights of the Child (United Nations, 1989). Access to and participation in education is offered as a pledge to all children in Sustainable Development Goal 4 (SDG4) (United Nations, 2022a,b). Inclusive education, it follows, is the enabling and safeguarding of that pledge.

Achieving that pledge is some way off. In March 2020 the Director-General of UNESCO, Audrey Auzalay (2020) conservatively estimated that 290 million children were out of school. Quite obviously this figure represented the growing impact of school closures as governments sequentially responded to the global spread of the COVID-19 virus. Nevertheless, prior to the onset of the COVID-19 pandemic, the numbers of children out of school had been cause for concern. One such expression of concern came from the American economist Jeffrey Sachs who commenced the George F. Kneller Lecture to the Comparative and International Education Society in April 2019 to gently register his concern that the assembled educators, and educators in general, maintained such polite quietude when on that day some 250 million children were out of school. Sachs indicated the symbolism of delivering this lecture in San Francisco’s Herbst Auditorium where in 1945 the United Nations Charter was signed foreshadowing the signing 3 years later in 1948 of the “moral charter”: the *Universal Declaration of Human Rights* (United Nations, 1948), in which the right of all to education is inscribed. The problem of children not receiving an education could be resolved, he assured his listeners, through political will and material redistribution through fair taxation.

Who are these educationally excluded children? There are obvious barriers to schooling when “our lives are reeling in the maelstrom of multiple crises,” when “there are malignant winds blowing on this blue planet” (Castells, 2019). Unprecedented levels of population displacement through war; not so natural disasters that speak of accelerating environmental degradation; food insecurity and poverty; the inevitability of mass redundancy driven by the unrelenting global sourcing of cheap labor to reduce production costs and increase profit all erect barriers to education. Childhood labor and marriage; tribal, ethnic, and religious conflict; and the lottery of intersecting identity markers such as geographic location, gender, race, sexuality, language, and disability are at play in excluding children and young people from participation in education. The UNESCO Institute for Statistics (2022) reminds us that estimates of the numbers of children out of school are inadequate because of the invisibility of marginalized population cohorts.

It is important to observe that while the concentration of many of the transnational development organizations is fixed on the countries of the Global South, exclusion from school is a long-standing and resilient problem in the Global North where school remains an active agent in the reproduction of privilege and disadvantage (Connell, 1993; Mowatt, 2022; Picketty, 2022). Like Connell (1994) and Teese (2013) before them, Lamb et al. (2020) at the Mitchell Institute highlight the nonrandom distribution of academic underachievement according to postcode as a proxy for socioeconomic status. Pandemic has trained a spotlight on educational inequality and exclusion with education jurisdictions in countries of the Global North and South expressing concern over the missing or “ghost children” who remain out of school (UK Parliament, 2022).

Simply put, inclusive education may therefore refer to the mitigation or dismantling of barriers to access to and participation and success in education for all children and young people. Inclusive education is not simply a project for getting children into school; it also addresses the differential experiences of children in schools. The experience of schooling may be more or less inclusive: more or less enabling, more or less disabling. We are now beginning to encounter the complexity of this never to be completed project.

Recording a history of inclusive education reminds us of Knausgaard’s struggles with apparent simplicity. In a general sense inclusive education must refer to any initiatives to increase access, presence, participation, and success in education. From this point complexity insinuates itself into our discussion. A relatively modern invention, schooling was not originally intended or designed for everyone (Tait, 2019). Indeed, major changes in the duration of attendance and the substance of curriculum and pedagogy proceed from economic imperative or requirements for specific knowledge and skills in human capital (Apple, 1979). Over time we have seen moves to improve access to and the quality of education for those children and young people marginalized and excluded based on identity markers such as indigeneity, gender, class, race, ethnicity, language, poverty, geographic location,

and disability. The hard-won victories of the civil rights and women's movements demonstrating institutional racism and sexism as barriers to education have been important in the development of critiques of educational exclusion. Understanding formations and mobilizations of exclusion, who's in and who's out, requires acknowledgment of context and a reading of social theory and political economy (Slee, 2011).

As acknowledged, inclusive education is not just a question of access. What happens to children and young people inside schools? Here we enter the complex matrix of overlapping relationships between students and their peers, between students and their teachers, between students and how they are being taught, between students and what they are being taught, and between students and the rituals, routines, and cultures of schools. How do these dynamic relationships affect students? Do they engage or disengage, include or exclude? Understanding the organizations and cultures of schools are central to the theory and practice of inclusive education. Curriculum, pedagogy, assessment, student classification, and stratification are all at play in determining the quality and inclusiveness of educational experiences.

Different population cohorts shaped by geopolitical context and intersecting identity markers have sought to ensure educational access and inclusion. The progress of education theory making, policy-making, and practice has reflected this. Schooling reflects global population movements (Rizvi and Lingard, 2010) and in broad terms we now observe greater population diversity in schools accompanied by discourses of multiculturalism, diversity, and inclusion. Simultaneously, we see intransigent demarcations in and out of schooling according to the erstwhile fault lines of race, ethnicity, class, gender, and disability. Schooling, reflecting neoliberal political economy, is a highly regulated performative enterprise (Ball, 2017). Jurisdictions and individual schools compete against each other to secure higher positions on international and local ranking tables. The unit of measure, student test results, is the incubus for an enveloping ethic of competitive individualism. Students are the bearers of success and risk. In this environment schools, being risk averse, are increasingly selective (Gillborn and Youdell, 1999; Ravitch, 2014; Slee, 2018). As Bauman (2004) counsels, schools produce "surplus populations" who become "collateral casualties" of intensifying competition. Paradoxically, floating discourses of merit, standards, and inclusion are deployed to describe cultures of schooling that must inevitably sift and sort, categorize and regulate, and ultimately exclude students (Ball and Collet-Sabe, 2021).

Inclusive education, therefore, is not what it may appear to be—it is anything but neat. Yes, in many parts of the world there are obviously more children engaged in basic and secondary education than was formerly the case. This is cause for celebration and for intensifying our efforts. Many other children still fall through the cracks and remain invisible. Moreover, the experience of those in school may be less than an inclusive educational experience and this is reflected in data on increasing rates of exclusion and home schooling wherein vulnerable population cohorts loom large (Mowatt, 2022).

Representing Inclusive Education

In this volume of the Fourth Edition of the *International Encyclopedia of Education* I have searched for a way of first representing the work of Inclusive Education writ large and also of one of the strands within that field of education research, policy, and practice: Disability Studies in Education. To that end, I have adopted the catch cry of the 1990 World Conference on Education for All in Jomtien, Thailand (UNESCO, 1990). "Education for All" unambiguously encapsulates the value position of inclusive education—education is the right of all children. Education for All has continued to form the organizing motif for global education to ensure that vulnerable population cohorts such as the girl child, child soldiers and laborers, disabled children, and refugee children secure their right to attend school and participate in a high-quality and inclusive education.

The first group of chapters consider this broader challenge of inclusive education adopting both broader and specific concerns within the inclusive education agenda. The second section of this volume takes up and advances the challenge of "Disability Studies in Education" as a critical concern within the broader field of inclusive education. Many readers coming to this encyclopedia may be puzzled by nomenclature. You might be asking: "Why disability studies in education" and not "special educational needs"? This arose as an issue in communications with prospective contributors and is an ongoing debate in the field. Indeed, in many jurisdictions inclusive education is default vocabulary for special or disabled students, most frequently described as having "special educational needs" following Mary Warnock's (Department for Education and Science, 1978) administrative descriptor. In the education lexicon "special" has and will also represent a point of temporal and spatial separation for specified groups of children. Accordingly, chicanery lurks within the ubiquitous coupling of "inclusive" and "special". Drawing from the transdisciplinarity of critical disability studies, disability studies in education marks a separation from special education's ontological assumptions about impairment, ability, disability, and disablement.

My point is not to demonize special schools and special educators. It is important to return to the beginnings of special schooling to apply perspective. Historically, special schooling exists because of regular neighborhood schools' unwillingness and refusal to enroll and educate disabled children. Parents, medics, and maverick educators went against the grain to insist that "these children" who were considered defective, invalid, and ineducable could learn and be taught (Danforth, 2009). Special schooling at its inception was both a radical resistance to educational foundations and a conservative falling into line with educational structures. It embodies the principle of fairness in ensuring that children who would otherwise remain excluded from education got to attend a school and not remain hidden from the community and other children, but it also energizes ableist assumptions about the nature of impairment, ability, and disability. The systemic bifurcation of normal and special children was complete and stable for decades, each side of the institutional division growing under its own administrative and cultural domain. This deep divide has been unsettled for some time with demands for mainstreaming, integration, and inclusion.

In 1994 representatives from 92 governments and 25 international organizations gathered in the Spanish city of Salamanca and signed the “Salamanca Statement and Framework for Action on Special Needs Education.” It is worth reminding ourselves of two of its guiding principles:

- Those with special educational needs must have access to regular schools which will accommodate them within a child-centered pedagogy capable of meeting those needs.
- Regular schools with this inclusive orientation are the most effective means of combatting discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all ...

(UNESCO, 1994: viii & ix)

Reflecting on the creation and legacy of the “Salamanca Statement and Framework for Action on Special Needs Education” some 25 years later Ainscow et al. (2019) observe that:

The publication of the Salamanca Statement suggested a significant change in policy direction, not least for the education of disabled students. However, the document contains some ambiguities that have become apparent when used in the field. In particular, the deficit orientation of special education discourse continues to undermine more radical moves ... Indeed, this conservative discourse has protected resistance to the reforms it foreshadowed.

General comment No. 4 on Article 24 - the right to inclusive education of the Convention on the Rights of Persons with Disabilities (2016) addresses this by stipulating the requirements of an inclusive education:

It requires an in-depth transformation of education systems in legislation, policy and the mechanisms for financing, administering, designing, delivering and monitoring education.

United Nations (2016).

Herein lies a clarion call for educational transformation. Inclusive education is not an edict for emptying (or merging) special schools into regular schools. The recent public hearings of the Royal Commission into Violence, Abuse, Neglect and Exploitation of People with Disability in Australia provide ample evidence of the failure of neighborhood regular schools to provide an inclusive education for disabled students (Royal Commission into Violence, Abuse, Neglect and Exploitation of people with Disability, 2019, 2020). A key to this transformation is a reconsideration of the knowledge and discourse we employ for that task. The discourse of inclusion has been thoroughly appropriated and deployed to support the conservative interests of a codependent partnership between special and regular education (Tomlinson, 2017) behind a veneer of inclusion. This volume, in advancing disability studies in education, is testing the capacity of disability studies in education to provide critique and reconstruction of the “mechanisms for financing, administering, designing, delivering and monitoring education.”

Returning to the Desk in Southern Sweden

You will remember that I commenced my introduction to this volume with Karl Ove Knausgaard contemplating the simple question: Why do I write? I did so because I was impressed by his observation of the treachery in simplicity and by his need to write to resist “easy thinking and preconceived notions.” It reminded me that I have frequently used the Belgian surrealist painter, René Magritte’s iconic 1929 painting entitled *The Treachery of Images* in which a representation of a pipe is accompanied by the inscription, “Ceci n’est pas une pipe,” which is French for “This is not a pipe.” There are many representations of inclusion in education. All of them warrant authenticity checks.

Knausgaard also refers to an artist to explain why he must write. The artist he chose is the Norwegian painter Edvard Munch. Munch paints, says Knausgaard (2018:5), because he has lost trust in the world. This may well be the predicament of inclusive education. The research domain exists because we cannot trust education jurisdictions to be inclusive.

References

- Ainscow, M., Slee, R., Best, M., 2019. The Salamanca Statement 25 Years On (Editorial). *International Journal of Inclusive Education*, 23 (7-8), 671–676. <https://www.tandfonline.com/doi/full/10.1080/13603116.2019.1622800>.
- Apple, M., 1979. *Ideology and Curriculum*. Routledge and Kegan Paul, Boston.
- Azoulay, A., 2020. 290 Million Students Out of School Due to COVID-19: UNESCO Releases First Global Numbers and Mobilizes Response. UNESCO, Paris. Retrieved from. <https://en.unesco.org/news/290-million-students-out-school-due-covid-19-unesco-releases-first-global-numbers-and-mobilizes>.
- Ball, S.J., 2017. *The Education Debate*, third ed. The Policy Press, Bristol.
- Ball, S.J., Collet-Sabe, 2021. Against school: an epistemological critique. *Discourse*. <https://doi.org/10.1080/01596306.2021.1947780>.
- Bauman, Z., 2004. *Wasted Lives: Modernity and Its Outcasts*. Polity, Cambridge.
- Castells, M., 2019. *Rupture: The Crisis of Liberal Democracy*. Polity Books, Cambridge.

- Connell, R.W., 1993. *Schools and Social Justice*. Temple University Press, Philadelphia.
- Connell, R.W., 1994. Poverty and education. *Harv. Educ. Rev.* 64 (2), 125–150. <https://doi.org/10.17763/haer.64.2.m14947g30k1x5781>.
- Danforth, S., 2009. *The Incomplete Child*. Peter Lang, New York.
- Department for Education and Science, 1978. *Special Educational Needs. Report of the Committee of Enquiry Into the Education of Handicapped Children and Young People (The Warnock Report)*. Her Majesty's Stationery Office, London.
- Gillborn, D., Youdell, D., 1999. *Rationing Education: Policy, Practice, Reform, and Equity*. Open University Press, Buckingham.
- Knausgaard, K.O., 2018. *Inadvertent*. Yale University Press, London.
- Lamb, S., Huo, S., Walstab, A., Wade, A., Maire, Q., Doecke, E., Endekov, Z., 2020. *Educational Opportunity in Australia 2020: Who Succeeds and Who Misses Out*. Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute, Melbourne.
- Mowatt, J.G., 2022. *Building Community to Create Equitable, Inclusive and Compassionate Schools Through Relational Approaches*. Routledge, Abingdon.
- Picketty, T., 2022. *A Brief History of Equality*. The Belknap Press of Harvard University Press, Cambridge Mass.
- Ravitch, D., 2014. *Reign of Error: The Hoax of the Privatisation Movement and the Danger to America's Public Schools*. Vintage Books, New York.
- Rizvi, F., Lingard, 2010. *Globalizing Education Policy*. Routledge.
- Royal Commission Into Violence, Abuse, Neglect and Exploitation of people with Disability, 2019. *Education & Learning. Issues Paper*, 30th October, Canberra. <https://disability.royalcommission.gov.au/system/files/2020-07/Issues-paper-Education-Learning.pdf>.
- Royal Commission Into Violence, Abuse, Neglect and Exploitation of People With Disability, 2020. *Inclusive Education in Queensland – Preliminary Inquiry. Public Hearing Report*. <https://disability.royalcommission.gov.au/system/files/2020-10/Report%20-%20Public%20hearing%20-%20Inclusive%20education%20in%20Queensland%20-%20preliminary%20inquiry.pdf>.
- Slee, R., 2011. *The Irregular School*. Abingdon: Routledge.
- Slee, R., 2018. *Inclusive Education Isn't Dead, It Just Smells Funny*. Routledge, London.
- Tait, G., 2019. *Making Sense of Mass Education*, third ed. Cambridge University Press, Melbourne.
- Teese, R., 2013. *Academic Success and Social Power: Examinations and Inequality*. Scholarly Press, North Melbourne.
- Tomlinson, S., 2017. *A Sociology of Special and Inclusive Education. Exploring the Manufacture of Inability*. Routledge, Abingdon.
- UK Parliament, 2022. Government Must act on "Ghost Children" Missing From Education System. UK Parliament, Westminster. <https://committees.parliament.uk/committee/203/education-committee/news/171073/government-must-act-on-ghost-children-missing-from-education-system/>.
- United Nations, 1948. *Universal Declaration of Human Rights*. New York: United Nations.
- United Nations, 1989. *United Nations Convention on the Rights of the Child, (UNCRC). Resolution 44/25 of 20th November*. United Nations, New York (see: Articles 23, 28, 29). https://www.unicef.org.uk/wpcontent/uploads/2010/05/UNCRC_united_nations_convention_on_the_rights_of_the_child.pdf.
- United Nations, 2016. *General Comment No. 4 on Article 24 - The Right to Inclusive Education*. United Nations, New York. <https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-4-article-24-right-inclusive>.
- United Nations, 2022a. *Sustainable Development Goals: 4 Quality Education*. United Nations, New York. <https://www.un.org/sustainabledevelopment/education/>.
- UNESCO, 1990. *World declaration on education for all: framework for action to meet basic learning needs*. In: *World Conference on Education for All: Meeting Basic Learning Needs*, Jomtien, Thailand, March 1990. UNESCO, Paris.
- UNESCO, 1994. *Final Report. World Conference on Special Needs Education: Access and Quality. Salamanca Statement and Framework for Action on Special Needs Education*. UNESCO, Paris.
- UNESCO, 2022b. *Reaching Out to All Learners: A Resource Pack for Supporting Inclusion and Equity in Education*. <https://www.inclusive-education-in-action.org/resources/reaching-out-all-learners-resource-pack-supporting-inclusion-and-equity-education>.

CONTRIBUTORS TO VOLUME 9

Rachel Aberdein

Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Mel Ainscow

University of Glasgow, Glasgow, United Kingdom; and University of Manchester, Manchester, United Kingdom

Hadeel F Al-Nawab

Faculty of Education, Monash University, Melbourne, VIC, Australia

Julie Allan

School of Education, University of Birmingham, Birmingham, United Kingdom

Ines Alves

University of Glasgow, Glasgow, United Kingdom

Dimitris Anastasiou

Southern Illinois University Carbondale, Department of Counseling, Quantitative Methods and Special Education, Carbondale, IL, United States

David Armstrong

School of Education, RMIT University, Melbourne, VIC, Australia

Andrew Azzopardi

Faculty for Social Wellbeing, University of Malta, Malta

Bernadette Baker

University of Wisconsin-Madison, Madison, WI, United States

Angharad E Beckett

University of Leeds, Leeds, United Kingdom

Marnie Best

Centre for Research in Educational and Social Inclusion, Education Futures, University of South Australia, Adelaide, SA, Australia

Brent Bradford

Concordia University of Edmonton, Edmonton, AB, Canada

Sam Brain

Deloitte Access Economics, Melbourne, VIC, Australia

Suzanne Carrington

Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Tsitsi Chataika

University of Zimbabwe, Department of Educational Foundations, Harare, Zimbabwe

Kathy Cologon

Macquarie University, Sydney, NSW, Australia

David J Connor

Hunter College, City University of New York, New York, NY, United States

Bill Cope

University of Illinois Urbana-Champaign, Champaign, IL, United States

Tim Corcoran

School of Education, Deakin University, Melbourne, VIC, Australia

Abigail E Croydon

Southampton Education School, University of Southampton, Southampton, United Kingdom

Elisabeth De Schauwer

Department of Special Needs Education, Ghent University, Ghent, Belgium

Verity J Donnelly

European Agency for Special Needs and Inclusive Education, Odense, Denmark

Arden A Duncan Bonokoski

STEPS Forward Inclusive Post-Secondary Education Society, Chilliwack, BC, Canada

Miriam Edwards

Faculty of Business and Economics, The University of Melbourne, Parkville, VIC, Australia

Thomas T Engsig
University College of Northern Denmark, Aalborg, Denmark

Marion Felder
Inclusion and Rehabilitation, Department of Social Sciences, University of Applied Sciences, Koblenz, Germany

Domingos Fernandes
ISCTE—Instituto Universitário de Lisboa, Lisbon, Portugal

Deborah J Gallagher
University of Northern Iowa, Cedar Falls, IA, United States; and Trickum Middle School, Gwinnett County Public Schools, Lilburn, GA, United States

Dan Goodley
iHuman, School of Education, Sheffield, United Kingdom

Miro Griffiths
School of Sociology and Social Policy, University of Leeds, Leeds, United Kingdom

Bjørn F Hamre
History and Education, Department of Communication, University of Copenhagen, Copenhagen, Denmark

Whitney Hanley
Trickum Middle School, Gwinnett County Public Schools, Lilburn, GA, United States

Bella Harper-Hill
Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Keely Harper-Hill
Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Seamus Hegarty
University of Warwick, United Kingdom

David Hernández-Saca
University of Northern Iowa, Cedar Falls, IA, United States

Garry Hornby
University of Plymouth, Plymouth, United Kingdom

E Anne Hughson
Community Rehabilitation and Disability Studies, Community Health Science, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Christopher J Johnstone
University of Minnesota, Minneapolis, MN, United States

Mary Kalantzis
University of Illinois Urbana-Champaign, Champaign, IL, United States

Jenny E Kassen
The University of Western Ontario, Faculty of Education, London, ON, Canada

James M Kauffman
Department of Curriculum, Instruction and Special Education, University of Virginia, Afton, VA, United States

Jesper V Kragh
The SAXO Institute, University of Copenhagen, Copenhagen, Denmark

James G Ladwig
School of Education, University of Newcastle, Newcastle, NSW, Australia

Lucy Lake
CAMFED International, Cambridge, United Kingdom

Kamal Lamichhane
Disability Studies and Inclusive Education, Graduate School of Comprehensive Human Sciences, The University of Tsukuba, Tsukuba, Japan

Anastasia Liasidou
Faculty of Education, Humanities and Social Sciences, St Mary's University, Twickenham, London, UK

Kirsty Liddiard
iHuman, School of Education, Sheffield, United Kingdom

Levan Lim
Psychology and Child & Human Development Academic Group, National Institute of Education, Nanyang Technological University, Singapore

Joao Lopes
School of Psychology, University of Minho, Braga, Portugal

Tim Loreman
Concordia University of Edmonton, Edmonton, AB, Canada

Andrew Loxley
School of Education, Trinity College Dublin, Dublin, Ireland

Jude MacArthur
Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

Gunnlaugur Magnússon
Department of Education, Uppsala University, Uppsala, Sweden

Rachael Mariani
Victorian Department of Education and Training, Melbourne, VIC, Australia

Wayne J Martino
The University of Western Ontario, Faculty of Education, London, ON, Canada

Katerina Mavrou
Department of Education Sciences, European University Cyprus, Nicosia, Cyprus

Anne-Marie McIlroy
Te Pae Aronui Ministry of Education, Wellington, New Zealand

Amy McPherson
National School of Education, Australian Catholic University, Strathfield, NSW, Australia

Cor JW Meijer
European Agency for Special Needs and Inclusive Education, Odense, Denmark

Kyriaki Messiou
University of Southampton, Southampton, United Kingdom

Valentina Migliarini
School of Education, University of Birmingham, Birmingham, United Kingdom

Missy Morton
Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

Joan G Mowat
School of Education, University of Strathclyde, Glasgow, Scotland, United Kingdom

Erica Neeganagwedgin
Western University, London, ON, Canada

Melanie AE Nind
Southampton Education School, University of Southampton, Southampton, United Kingdom

Evdoxia Nteropoulou-Nterou
School of Education, Department of Early Childhood Education, National and Kapodistrian University of Athens, Athens, Greece

Kenan Omercajjic
The University of Western Ontario, Faculty of Education, London, ON, Canada

Rachel Outhred
Oxford MeasurEd, Oxford, United Kingdom

Olivia Penna
Melbourne Graduate School of Education, The University of Melbourne, Parkville, VIC, Australia

Shiralee Poed
School of Education, The University of Queensland, St Lucia, QLD, Australia

Rebecca L Porter
Center for Disability Studies, University of Leeds, Leeds, United Kingdom

Bethany M Rice
Towson University, Towson, MD, United States

Richard Rieser
World of Inclusion Ltd London, UK

Jonathan Rix
The Open University, Milton Keynes, United Kingdom; and Inland Norway University of Applied Sciences, Innlandet, Norway

Katherine Runswick-Cole
iHuman, School of Education, Sheffield, United Kingdom

Gill Rutherford
University of Otago College of Education, Dunedin, New Zealand

Beth Sagers
Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Matthew J Schuelka
University of Minnesota, Minneapolis, MN, United States

Umesh Sharma
Faculty of Education, Monash University, Melbourne, VIC, Australia

Nidhi Singal
Faculty of Education, University of Cambridge, United Kingdom

Josephine Sirotkin
Center for Disability Studies, University of Leeds, Leeds, United Kingdom

Roger Slee

Center for Disability Studies, School of Sociology & Social Policy, University of Leeds, Leeds, United Kingdom

Hannah Smith

Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

Katie J Smith

CAMFED International, Cambridge, United Kingdom

William C Smith

University of Edinburgh, Edinburgh, United Kingdom

Kristin Snoddon

School of Early Childhood Studies, Ryerson University, Toronto, ON, Canada

Pearl Subban

Faculty of Education, Monash University, Melbourne, VIC, Australia

Thana Thaver

Psychology and Child & Human Development Academic Group, National Institute of Education, Nanyang Technological University, Singapore

Gary Thomas

The School of Education, University of Birmingham, Birmingham, United Kingdom

Tanya Titchkosky

Department of Social Justice Education, OISE of the University of Toronto, Toronto, ON, Canada

Fergal Turner

Oxford MeasurEd, Oxford, United Kingdom

Anastasia O Tzirides

University of Illinois Urbana-Champaign, Champaign, IL, United States

Geert Van Hove

Department of Special Needs Education, Ghent University, Ghent, Belgium

Wayne Veck

Department for Education Studies and Liberal Arts, Faculty of Education, University of Winchester, Winchester, United Kingdom

Elizabeth Walton

School of Education, University of Nottingham, Nottingham, United Kingdom; and School of Education, University of Witwatersrand, Johannesburg, South Africa

Amanda Watkins

European Agency for Special Needs and Inclusive Education, Odense, Denmark

Ben Whitburn

University of Southampton, UK

Jodie Wilde

Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

CONTENTS OF VOLUME 9

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 9: Inclusive Education and Disability Studies in Education</i>	<i>xix</i>
<i>Contributors to Volume 9</i>	<i>xxiii</i>

INCLUSIVE EDUCATION & EDUCATION FOR ALL

Inclusion and equity: an agenda for educational reform <i>Mel Ainsco</i>	1
Re-examining the education for all agenda in “apocalyptic” times <i>Marnie Best</i>	12
Quality and inclusion in SDG 4 <i>William C Smith and Christopher J Johnstone</i>	19
Reform and purpose for inclusive values in complex education systems <i>Matthew J Schuelka</i>	28
Paying for inclusion and exclusion <i>Amanda Watkins, Verity J Donnelly, and Cor JW Meijer</i>	35
Economics of inclusive education <i>Kamal Lamichhane</i>	51
Legislation and policy in Europe <i>Valentina Migliarini, Julie Allan, Geert Van Hove, and Elisabeth De Schauwer</i>	58
Inclusive pedagogies, new learning, new times <i>Anastasia O Tzirides, Bill Cope, and Mary Kalantzis</i>	68
Within school stratification <i>James G Ladwig and Amy McPherson</i>	76
Working collaboratively with the school community to build inclusion for all <i>Joan G Mowat</i>	85

Nordic representations and paradoxes of inclusive education <i>Thomas T Engsig</i>	98
Utilizing a global social justice lens to explore indicators of inclusive education <i>Umesh Sharma and Pearl Subban</i>	104
Principles of inclusive practice in schools: what is COVID-19 teaching us? <i>Brent Bradford, Tim Loreman, and Umesh Sharma</i>	115
Delivering on the promises of tertiary education in low- and middle-income countries: risks, sustainability and inclusion <i>Rachel Outhred and Fergal Turner</i>	126
Changing knowledge & cultures in teacher education: autoethnography for critical reflection <i>Bethany M Rice</i>	136
The role of students' voices in promoting inclusive education <i>Kyriaki Messiou</i>	143
First Nations & Indigenous identities in education <i>Erica Neeganagwedgin</i>	151
Gender identities, sexuality and education <i>Wayne J Martino, Jenny E Kassen, and Kenan Omercajic</i>	159
Refugee student identities and educational exclusion <i>Wayne Veck</i>	169
Educational trends and exclusion <i>David Armstrong</i>	177
Building inclusive communities <i>Andrew Azzopardi</i>	183

ADVANCING DISABILITY STUDIES IN EDUCATION

"Disability and society: changing our frames of reference" <i>Tanya Titchkosky</i>	191
Disability studies in a posthuman age <i>Katherine Runswick-Cole, Dan Goodley, and Kirsty Liddiard</i>	200
From capacity to ability to automation: "Western" conceptions of the figure of man and ableist subjectivities <i>Bernadette Baker</i>	209
Following the money? Commodification, segregation and shopping <i>Andrew Loxley and Gary Thomas</i>	219
"First you have to see me as a human being": disabled students' experiences of schooling <i>Gill Rutherford</i>	228
The concept of belonging and other areas of interest in disability studies in education <i>David J Connor</i>	238
Disability studies in education and intersectionality <i>Elizabeth Walton</i>	249
Cross-cultural perspectives on the rise of special education policy discourses. <i>Anastasia Liasidou and Katerina Mavrou</i>	259

Recent debates in special & inclusive education <i>James M Kauffman, Dimitris Anastasiou, Marion Felder, Garry Hornby, and Joao Lopes</i>	269
Re-examining special needs—what could be better? <i>Jonathan Rix</i>	284
Developing inclusive education for disabled children and students: advocacy in hard times <i>Richard Rieser</i>	293
“But tell me, where do the children play?” Disabled children and the right to play <i>Angharad E Beckett</i>	306
Everybody’s got to learn sometime: disability youth activism and the pursuit for inclusive education <i>Miro Griffiths</i>	320
Attitudes toward inclusion <i>Tim Corcoran</i>	328
Considerations for implementing Universal Design for Learning: toward anti-oppressive pedagogies <i>Kathy Cologon</i>	334
Teachers with disabilities in mainstream schools: embodying and enacting inclusive education <i>Nidhi Singal</i>	341
Using disability studies in education to move from individualized assessment to partners in pedagogy <i>Missy Morton, Jude MacArthur, and Anne-Marie McIlroy</i>	352
Inclusive postsecondary education: vision, practice, higher education reimagined <i>E Anne Hughson and Arden A Duncan Bonokoski</i>	362
Researching curriculum and disability in higher education <i>Shiralee Poed, Miriam Edwards, Hadeel F Al-Nawab, and Olivia Penna</i>	368
Reviews of education for students with disability in Australia <i>Sam Brain and Rachael Mariani</i>	376
From Salamanca to Sweden: inclusive education as policy in transit <i>Gunnlaugur Magnússon</i>	386
Public policies in Portuguese education: the path to inclusion for all <i>Ines Alves and Domingos Fernandes</i>	397
Inclusive education policies and practices in Malawi and Zimbabwe <i>Tsitsi Chataika</i>	404
Crisis, austerity and the inclusive education agenda: the case of Greece <i>Evdoxia Nteropoulou-Nterou</i>	418
Inclusive education in the Republic of Singapore: a situated perspective <i>Levan Lim and Thana Thaver</i>	426
Writing the history of Nordic disability and mental disorders <i>Bjørn F Hamre and Jesper V Kragh</i>	435
The National Disability Insurance Scheme, education and funding to support the inclusion of students with disabilities across educational sectors <i>Ben Whitburn</i>	446

Disability and employment: finding a way to work <i>Abigail E Croydon and Melanie AE Nind</i>	453
Disability and poverty <i>Rebecca L Porter and Josephine Sirotkin</i>	462
Gender, poverty and disability—responding to a triple exclusion <i>Lucy Lake, Katie J Smith, and Seamus Hegarty</i>	470
Inclusive education and deaf learners <i>Kristin Snoddon</i>	480
Girls, autism identities, and education <i>Suzanne Carrington, Beth Saggars, Keely Harper-Hill, Rachel Aberdein, Jodie Wilde, Hannah Smith, and Bella Harper-Hill</i>	485
Looking for something outside of ourselves: the politics of researching race and dis/Ability and (special) education <i>Deborah J Gallagher, David Hernández-Saca, and Whitney Hanley</i>	497
Rethinking schooling for inclusion <i>Roger Slee</i>	506

Inclusion and equity: an agenda for educational reform

Mel Ainscow^{a,b}, ^a University of Glasgow, Glasgow, United Kingdom; and ^b University of Manchester, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

The international policy context	1
Promising developments	3
Other global trends	3
A whole-system approach	4
Inclusion and equity as principles	5
Use of evidence	5
School development	7
Involving the wider community	7
The role of administrative departments	8
Conclusion	9
References	10

This chapter focuses on a what is arguably the greatest challenge facing education systems around the world, that of finding ways of including all children in schools. In economically poorer countries this is mainly about the millions of children who are not able to attend formal education (UNESCO, 2015). Meanwhile, in wealthier countries many young people leave school with no worthwhile qualifications, while others are placed in special provision away from mainstream education and some choose to drop out since the lessons seem irrelevant (OECD, 2012). And, more recently, teachers have experienced unprecedented pressures in trying to find ways of ensuring quality education for all their students within the context of the COVID-19 pandemic. Faced with these challenges, there is evidence of an increased interest internationally in the idea of making education more inclusive and equitable. However, the field remains confused as to the actions needed in order to move policy and practice forward.

In the Chapter I draw on studies carried out over many years with colleagues in various parts of the world that point to some promising possibilities (for more detailed accounts of this research see: Ainscow, 2016a, 2020). In so doing, I provide a framework that has proved to be helpful in reviewing and developing national policies. I begin, however, by summarizing relevant international developments.

The international policy context

Over the last 30 years there have been major international efforts to encourage inclusive educational developments. In particular, the United Nation's Education for All (EFA) movement has worked to make quality basic education available to all learners. Agreed in 1990, the EFA Declaration sets out an overall vision, which is about being proactive in identifying the barriers some learners encounter in attempting to access educational opportunities (UNESCO, 1990). It also involves the identification of resources available at national and community level, and bringing them to bear on overcoming those barriers. This vision was reaffirmed by the World Education Forum meeting in Dakar (2000), held to review the progress made in the previous decade. The Forum declared that EFA must take particular account of the needs of the poor and the disadvantaged, including working children, remote rural dwellers and nomads, and ethnic and linguistic minorities, young people affected by conflict, HIV/AIDS, hunger and poor health, and those with special learning needs (UNESCO, 2000).

A major impetus for inclusive education was given by the World Conference on Special Needs Education in 1994. More than 300 participants representing 92 governments and 25 international organizations met in Salamanca, Spain to further the objective of Education for All by considering the fundamental policy shifts required to promote the approach of inclusive education, namely enabling schools to serve all children, particularly those defined as having special educational needs (UNESCO, 1994). Although the immediate focus of the Salamanca conference was on what was termed special needs education, its conclusion was that:

Special needs education – an issue of equal concern to countries of the North and of the South – cannot advance in isolation. It has to form part of an overall educational strategy and, indeed, of new social and economic policies. It calls for major reform of the ordinary school (pp. iii–iv).

The aim, then, is to reform education systems. This can only happen, it is argued, if mainstream schools become capable of educating all children in their local communities. In a much-quoted extract, the Salamanca Statement concluded that:

Regular schools with [an] inclusive orientation are the most effective means of combating discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all; moreover, they provide an effective education to the majority of children and improve the efficiency and ultimately the cost-effectiveness of the entire education system (p. ix).

As this key passage indicates, the move toward inclusive schools can be justified on a number of grounds. There is an *educational justification*: the requirement for inclusive schools to educate all children together means that they have to develop ways of teaching that respond to individual differences and that therefore benefit all children; a *social justification*: inclusive schools are able to change attitudes to difference by educating all children together, and form the basis for a just and non-discriminatory society; and an *economic justification*: it is likely to be less costly to establish and maintain schools which educate all children together than to set up a complex system of different types of school specializing in particular groups of children.

Further impetus to this movement was provided by the 48th session of the IBE-UNESCO International Conference on Education, with its theme “*Inclusive Education: The Way of the Future*,” held in 2008. The long-term objective of this event was to support UNESCO member states in providing the social and political conditions which every person needs in order to exercise their human right to access, take an active part in, and learn from educational opportunities (Operti et al., 2014). During the conference, ministers, government officials and representatives of voluntary organizations discussed the importance of broadening the concept of inclusion to reach all children, under the assumption that every learner matters equally and has the right to receive effective educational opportunities.

Moving forward, the year 2016 was particularly important in relation to the future of the EFA movement and, indeed, the legacy of Salamanca. Building on the Incheon Declaration agreed at the World Forum on Education in May 2015 (UNESCO, 2015), it saw the publication by UNESCO of the Education 2030 Framework for Action. This emphasizes inclusion and equity as laying the foundations for quality education. It also stresses the need to address all forms of exclusion and marginalization, disparities and inequalities in access, participation, and learning processes and outcomes. In this way, it is made clear that the international EFA agenda really has to be about “all.”

The importance of including disabled children is an essential strand within this new international policy agenda. This was stressed in the United Nations’ Convention on the Rights of Persons with Disabilities (United Nations, 2008), which states:

The right to inclusive education encompasses a transformation in culture, policy and practice in all educational environments to accommodate the differing requirements and identities of individual students, together with a commitment to remove the barriers that impede that possibility. (General Comment No 4).

The Convention defines non-inclusion, or segregation, as the education of students with disabilities in separate environments (i.e. in separate special schools, or in special education units located with regular schools). It commits to ending segregation within educational settings by ensuring inclusive classroom teaching in accessible learning environments with appropriate support. In essence, this means that education systems must provide a personalized educational response, rather than expecting the student to fit the system.

A new commitment reinforcing inclusion and equity in education was expressed at an International Forum, co-organized by UNESCO and the Ministry of Education of Colombia in September 2019 to celebrate the 25th anniversary of the Salamanca Declaration (UNESCO, 2019). As underlined by the theme “Every learner matters,” the Forum was an opportunity to revive the broadened notion of inclusion as a general guiding principle to strengthen equal access to quality learning opportunities for all learners.

Despite these developments, a recent Global Monitoring Report points out that an estimated 258 million children, adolescents and youth are still not in school (UNESCO, 2020). Furthermore, the UNESCO Institute for Statistics data show that more than 617 million children and adolescents are not achieving minimum proficiency levels in reading and mathematics (UIS, 2019). It is also important to note that only half of the world’s children receive pre-primary education, a failure that limits children’s futures and deepens inequities in later learning (UNICEF, 2019). Significantly, the disadvantaged are least served by quality early childhood care and education, although they benefit most from such interventions.

Meanwhile, there are high rates of student dropout globally, particularly at the secondary level, as well as challenges experienced by many marginalized groups during the transition to higher education (UNESCO, 2015). In developing countries, the low access to higher education of young people living in socioeconomically disadvantaged contexts and/or who are vulnerable for various reasons, is one of the most dramatic gaps in opportunities. This holds back the overall development of these nations.

In addition, it is vital to recognize that large gender gaps continue to exist in regards to access, learning achievement and continuation in education in many settings, most often at the expense of girls, although in some parts of the world boys are at a greater disadvantage. The evidence is that education systems often perpetuate rather than challenge gender inequalities (UNGEI, 2012).

While these concerns are most acute in the developing world, there are similar concerns in many wealthier countries, as noted by the OECD (2018), which reports that across its member countries, almost one of five students do not reach a basic minimum level of skills to function in today’s societies. It also states that students from low socio-economic background are twice as likely to be low performers, implying that personal or social circumstances are obstacles to achieving their educational potential.

Promising developments

Like all major policy changes, progress in relation to inclusion and equity requires an effective strategy for implementation. In particular, it requires new thinking that focuses attention on the *barriers* experienced by some children that lead them to become marginalized as a result of contextual factors, such as inappropriate curricula and forms of assessment, and inadequate teacher preparation and support. The implication is that overcoming such barriers is the most important means of development forms of education that are effective for all children. In this way, the focus on inclusion and equity becomes a way of achieving the overall improvement of education systems.

The situation across the world in relation to this challenging policy agenda is complex, with some countries making great strides, while others continue to have segregated provision of various forms of education for some groups of learners. There are, however, countries where there has been encouraging progress, such as:

- For more than 30 years, New Brunswick in Canada has pioneered the concept of inclusive education through legislation, local authority policies and professional guidelines (AuCoin et al., 2020). More recently, New Brunswick adopted a policy which defines the critical elements of an inclusive education system that supports students in common learning environments and provides supports for teachers. It sets clear requirements for school practice including procedures for the development of personalized learning plans, inclusive graduation, as well as strict guidelines when a variation of the common learning environment may be required.
- The Italian government passed a law in 1977 that closed all special schools, units and other non-inclusive forms of provision (Ianes et al., 2020). This legislation is still in force and more recent amendments have further strengthened the inclusive nature of the education system. Not only did this close segregated educational facilities but, starting with pre-schools, it removed the possibility of exclusion from school as a corrective sanction. While practice varies from place to place, there is no doubt that the principle of inclusion is widely accepted. There is, however, recent concern about an increase in students being withdrawn for periods outside the regular classroom for additional support.
- Having enacted legislation making disability discrimination within education unlawful, Portugal has gone much further in enacting an explicit legal framework for the inclusion of students with and without disabilities in education (Alves et al., 2020). Recent legislation requires that the provision of support for all students be determined, managed and provided at the regular school level, with local multidisciplinary teams responsible for determining what support is necessary to ensure all students (regardless of labels, categorization or a determination of disability) have access and the means to participate effectively in education, with a view to their full inclusion in society. It is also significant that Portugal has developed progressive assessment practices to support the achievement of all learners.
- In Finland, a country which regularly outperforms most other countries in terms of educational outcomes, education is viewed as a right and not as a privilege. The country's success is partly explained by the progress of the lowest performing quintile of students who out-perform those in other countries (Sabel et al., 2011). This has increasingly involved an emphasis on support for vulnerable students within mainstream schools, as opposed to in segregated provision. There is a particular focus on prevention of learning difficulties in Finland, and a high level of resources are directed at this in primary schools. For example, all Finnish schools are assigned specialists to support any student who requires additional help.

In drawing attention to these examples of policy development, it must be stressed that they are not seen as being perfect. Rather, they are countries where there are interesting developments from which to learn. They are also varied with respect to the approach being taken and what they have achieved. Consequently, while lessons can undoubtedly be learned from these developments, they must be replicated with care. There are many sources of inequity in education related to political, economic, social, cultural and institutional factors, and these vary across countries. This means that what works in one country may not work elsewhere. An emphasis on system change strategies being contextually sensitive is one of the pervading themes of the recommendations made later in this paper.

Other global trends

In some countries, increasing pressures to improve the rankings of education systems on global league tables is creating new barriers to progress in relation to inclusion and equity. This arises because of an increased emphasis on school autonomy, competition between schools and parental choice, and takes a variety of forms (Meyland-Smith and Evans, 2009). The schools involved have different titles, such as charter schools in the USA, free schools in Sweden, academies in England and independent public schools in parts of Australia. Implicit in these independent state-funded schools is an assumption that greater autonomy will allow space for the development of organizational arrangements, practices, and forms of management and leadership that will be more effective in promoting the learning of all students, particularly those from economically disadvantaged and minority backgrounds.

This global trend is a matter of considerable debate and there are varied views as to the extent to which it is leading to the desired outcomes. In particular, there is a concern that the development of education systems based on autonomy, coupled with high-stakes accountability and increased competition between schools, will further disadvantage learners from low-income and minority families (Salokangas and Ainscow, 2017). For example, parental choice and competition between schools has widened the gap between

schools that are seen to be more successful and those that are perceived to be less so in countries as varied as Chile, England and Sweden.

Recently New Zealand published a policy document based on the findings of a review of national policy carried out by an independent task force (Ministry of Education, 2019). This concludes that a key reason for the country's poor equity and achievement outcomes is that, since reforms introduced in 1989, schools have predominantly operated as autonomous, self-managing entities, loosely connected to each other and with a distant relationship with the center. This autonomy has left schools to operate largely on their own and without sufficient support. The document outlines the Government's strategy for the reform of the schooling system. This will involve a move toward a networked system that is more responsive to the needs of all learners, an approach I recommend later in this chapter.

In relation to these concerns, the OECD reports the success of certain education systems that rank highly on measures of both quality and equity (OECD, 2012). It argues:

The evidence is conclusive: equity in education pays off. The highest performing education systems across OECD countries are those that combine high quality and equity. In such education systems, the vast majority of students can attain high level skills and knowledge that depend on their ability and drive, more than on their socio-economic background (p. 14).

Further encouragement for this view is provided by the recent Report Card prepared for UNICEF by the Innocenti Center (UNICEF, 2018). The report focuses on educational inequalities in 41 of the world's richest countries, all of which are members of the Organization for Economic Co-operation and Development and/or the European Union. It argues that there is no systematic relationship between country income and indicators of equality in education. For example, it is notable that some of the economically poorer countries in the comparison, such as Latvia and Lithuania, achieve near-universal access to preschool learning and curb inequality in reading performance among both primary and secondary school students more successfully than countries that have far greater resources. The Report Card concludes that Finland, Latvia and Portugal have the most equal education systems.

The implication, then, is that it is possible for countries to develop education systems that are both excellent and equitable. The question is: how can this be achieved?

A whole-system approach

International experience has led me to formulate a framework for thinking about how to promote inclusion and equity within education systems (see Fig. 1). This forms the basis of UNESCO's *"Guide for Ensuring Inclusion and Equity in Education"* that I helped to develop with a team of international experts (UNESCO, 2017). Gradually amended through experience from an earlier version (Ainscow, 2005), the framework focuses attention on five interrelated factors, summarized in Fig. 1. These factors can help determine "levers for change" (Senge, 1989).

Senge (1989) sees "levers" as actions that can be taken in order to change the behavior of an organization and those individuals within it. He goes on to argue that those who wish to encourage change within an organization must be smart in determining where the high leverage lies. Too often, he suggests, approaches used to bring about large-scale changes in organizations are "low leverage." That is to say, they tend to change the way things look but not the way they work. Possible examples of low leverage activity in the education field include: the production of policy documents, conferences and in-service training courses. While such initiatives may

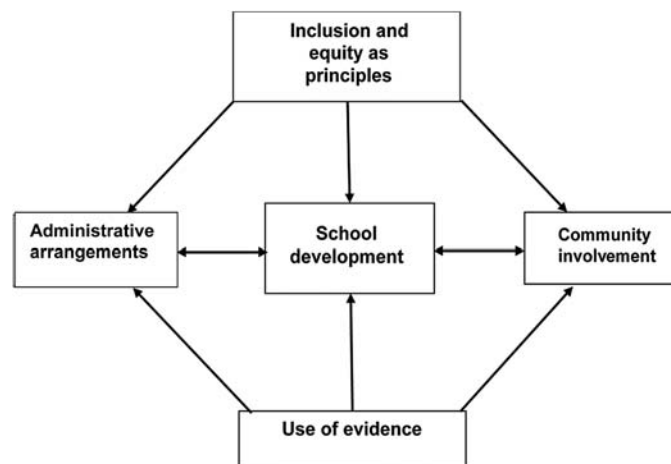


Fig. 1 A framework for promoting inclusion and equity within education systems.

make a contribution, by and large they do not lead to significant changes in thinking and practice. Our aim, therefore, must be to identify what may turn out to be more subtle, less obvious, higher leverage efforts to bring about change in education systems.

The framework places schools at the center of the analysis. This reinforces the point that moves toward inclusion must be focused on increasing the capacity of local neighborhood mainstream schools to support the participation and learning of an increasingly diverse range of learners. This is the paradigm shift implied by the Salamanca Statement, which I have described as an “inclusive turn” (Ainscow, 2007). It argues that moves toward inclusion are about the development of schools, rather simply involving attempts to integrate vulnerable groups of students into existing arrangements.

At the same time, the framework draws attention to a range of contextual influences that bear on the way schools carry out their work. These influences relate to: the principles that guide policy priorities within an education system; the views and actions of others within the local context, including members of the wider community that the schools serve and the staff of the departments that have responsibility for the administration of the school system; and the criteria that are used to evaluate the performance of schools. As explained below, these influences may provide support and encouragement to those in schools who are wishing to move in an inclusive direction. However, the same factors can act as obstacles to progress.

In what follows each of the five factors are explained, leading to a series of key ideas that need to be considered in developing future policies.

Inclusion and equity as principles

Terms such as “equity” and “inclusion” can be confusing since they may mean different things to different people. This is a particular problem when trying to move forward with other people – not least in schools, where everybody is so busy. Put simply, if there is not a shared understanding of the intended direction of travel, progress will be difficult. There is, therefore, a need for agreed definitions of these concepts.

In some countries, inclusive education is still thought of as an approach to serving children with disabilities within general education settings. Internationally, however, it is increasingly seen more broadly as a principle that supports and welcomes diversity among all learners. This presumes that the aim of is to eliminate social exclusion that is a consequence of attitudes and responses to diversity in race, social class, ethnicity, religion, gender and ability. As such, it starts from the belief that education is a basic human right and the foundation for a more just society. Hence, the emphasis on equity, which implies a concern with fairness. In the UNESCO Guide we sum this up as follows: “Every learner matters and matters equally” (UNESCO, 2017).

Our experience in many countries suggests that it is helpful to use a definition of inclusive education that involves the following elements:

Inclusion is a process. That is to say, inclusion has to be seen as a never-ending search to find better ways of responding to diversity. It is about learning how to live with difference, and, learning how to learn from difference. In this way differences come to be seen more positively as a stimulus for fostering learning, among children and adults.

Inclusion is concerned with the identification and removal of barriers. Here barriers may take different forms, some of which are to do with existing policies, the way schools are organized, relationships and the forms of teaching provided. Consequently, it is necessary to collect, collate and evaluate evidence about these factors in order to plan for improvements in policy and practice. This involves using evidence of various kinds to stimulate creativity and problem-solving.

Inclusion is about the presence, participation and achievement of all students. Here “presence” is concerned with where children are educated, and how reliably and punctually they attend; “participation” relates to the quality of their experiences while they are there and, therefore, must incorporate the views of the learners themselves; and “achievement” is about the outcomes of learning across the curriculum, not merely test or examination results.

Inclusion involves a particular emphasis on those groups of learners who may be at risk of marginalization, exclusion or under-achievement. This indicates the moral responsibility to ensure that those groups that are statistically most at risk are carefully monitored, and that, where necessary, steps are taken to ensure their presence, participation and achievement within the education system. At the same time, it is necessary to be vigilant in watching out for learners who may be overlooked.

We have found that a well-orchestrated debate about these elements can lead to a wider understanding of the principle of inclusion (Ainscow et al., 2006). Furthermore, such a debate, though by its nature slow and, possibly, never ending, can have leverage in respect to fostering the conditions within which schools can feel encouraged to move in a more inclusive direction. These debates must involve all stakeholders within communities, including families, political and religious leaders, and the media. They should also involve those within national and local district education offices. Recently, in countries such as Australia, Colombia, Mexico, Oman, Peru, Sweden and Uruguay, I have used the indicators provided by the UNESCO Guide for Ensuring Inclusion and Equity in Education to facilitate such discussions.

Use of evidence

In order to answer questions about inclusion and equity in education systems, it is important to know who is included, who is segregated and who is excluded from schooling. Without some form of statistical analysis in relation to these concerns, there

can be no accountability. However, when data collection efforts are only focused on particular categories of learner, there is a risk of promoting deficit views of those students who share certain characteristics, or come from similar backgrounds. This also distracts attention from more fundamental questions, such as: *Why are we failing some learners? What are the barriers experienced by some of our students?*

Previously I have argued that evidence is the life-blood of inclusive development (Ainscow, 2005, 2012). Therefore, deciding what kinds of evidence to collect and how to use it requires care, since, within education systems, “*what gets measured gets done.*” This is widely recognized as a double-edged sword precisely because it is such a potent lever for change. On the one hand, data are required in order to monitor the progress of children, evaluate the impact of interventions, review the effectiveness of policies and processes, plan new initiatives, and so on. On the other hand, if effectiveness is evaluated on the basis of narrow, even inappropriate, performance indicators, then the impact can be deeply damaging. While appearing to promote the causes of accountability and transparency, the use of data can, in practice: conceal more than they reveal; invite misinterpretation; and, worse of all, have a perverse effect on the behavior of professionals. The challenge is, therefore, to harness the potential of evidence as a lever for change, while avoiding these potential problems.

The starting point for making decisions about the evidence to collect at the system level should be with an agreed definition of inclusion. In other words, we must “measure what we value,” rather than is often the case, “valuing what we can measure.” In line with the suggestions made earlier, then, evidence collected within an education system needs to relate to the *presence, participation and achievement* of all students, with an emphasis placed on those groups of learners regarded to be “at risk of marginalization, exclusion or underachievement.”

This means that, in an education system based on the principles of inclusion and equity, all students should be assessed on an on-going basis in relation to their progress through the curriculum. This allows teachers to respond to a wide range of individual learners, bearing in mind that each learner is unique. It means that teachers and other professionals must be well informed about their students’ characteristics and attainments, while also assessing broader qualities, such as their capacity for cooperation. However, the ability to identify each student’s level of performance, or to enumerate certain student’s particular difficulties, is not enough. Teachers in inclusive systems need to gauge the effectiveness of their teaching for all of their students and should know what they need to do to enable each student to learn as well as possible.

An engagement with evidence is particularly crucial at the level of the school, where it can provide the stimulus for professional learning aimed at the development of inclusive thinking and practice. The starting point for developing inclusive practices is with the sharing of existing approaches through collaboration among staff, leading to experimentation with new practices that will reach out to all students (Ainscow, 2016b). This requires the development of a common language with which colleagues can talk to one another and, indeed, to themselves, about detailed aspects of their practice. Without such a language, teachers find it very difficult to experiment with new possibilities (Huberman, 1993).

A framework that can help in the promotion of an inclusive dialog within a school is provided by the *Index for Inclusion*, a review instrument developed originally for use in England but now available in many countries (Booth and Ainscow, 2002). It is intended to help draw on the knowledge and views of teachers, students, parents/carers and community representatives about barriers to participation that exist within the existing “*cultures, policies and practices*” of schools in order to identify priorities for change. In connecting inclusion with the detail of policy and practice, the Index encourages those who use it to build their own view of inclusion, related to their experience and values, as they work out which policies and practices they wish to promote or discourage.

The use of evidence to study teaching within a school can help in generating a language of practice (Ainscow et al., 2006). This, in turn, can foster the development of ways of working that are more effective in reaching hard to reach learners. Specifically, it can create space for rethinking by interrupting existing discourses. Particularly powerful techniques for generating evidence involve the use of mutual lesson observation, sometimes through video recordings, and comments collected from students about teaching and learning arrangements within a school (Ainscow and Messiou, 2017). Under certain conditions, such approaches provide *interruptions* that stimulate self-questioning, creativity and action. In so doing, they can lead to a reframing of perceived problems that, in turn, draws the teacher’s attention to overlooked possibilities for addressing barriers to participation and learning. In this way, differences among students, staff and schools become a catalyst for improvement.

Our recent research has explored teacher/student dialog as a means of making this happen. This was stimulated by a study carried out between 2011 and 2014 that had a major impact on our thinking (see Messiou et al., 2016, for more details). It involved collaborative action research carried out by teams of teachers in eight secondary schools in three countries (i.e. England, Portugal and Spain), with support from locally based university researchers.

Drawing on the lessons from this research, another study in primary schools trialled and evaluated a new model for developing learning and teaching that we call “Inclusive Inquiry” (Messiou and Ainscow, 2020). This uses dialog involving teachers and students in order to explore possibilities for making lessons more inclusive. Central to the model is the idea of engaging with the views of students, a process that permeates all the processes involved. The dialogs that this encourages are focused on learning and teaching. More specifically, differences among students and teachers are used to challenge existing thinking and practices in ways that are intended to encourage experimentation in order to foster more inclusive ways of working. This, in turn, sets out to break down barriers that are limiting the engagement of some learners.

The evidence from this study supports our earlier research which suggested that classroom-based research carried out by practitioners can be a powerful way of moving practice forward within a school. What is distinctive, however, is the added value that comes from engaging students themselves in the process. It is also clear that the use of the approach can present various

organizational challenges. In particular, it has sometimes proved to be challenging of the status quo within schools. Consequently, greater collaboration is needed among teachers in order to support the introduction of new practices. This requires organizational flexibility and the active support of senior staff, prepared to encourage and support processes of experimentation. Indeed, it implies the need for cultural change within a school, a theme I return to in the next section.

School development

There is not one single model of what an inclusive school looks like. What is common to highly inclusive schools, however, is that they are welcoming and supportive places for all of their students, not least for those with disabilities and others who experience difficulties (Dyson et al., 2004). This does not prevent these schools from also being committed to improving the achievements of all of their students. Indeed, they tend to have a range of strategies for strengthening achievement that are typical of those employed by all effective schools, and the emphasis on supporting vulnerable students does not appear to inhibit these strategies. As noted above, a key factor is the emphasis placed on tracking and supporting the progress of all of the students.

A problem reported from a number of countries, is that despite national policies emphasizing the equal rights of all children to attend mainstream settings, there has been evidence over many years of significant increases in the proportions of students being categorized in order that their schools can earn additional resources (Fulcher, 1989; Meijer and Watkins, 2019). This has led to dissatisfaction with progress toward integration and brought demands for more radical changes. One of the concerns of those who adopt this view is with the way in which students come to be designated as having special needs (e.g. Slee, 2010). They see this as a social process that needs to be continually challenged. More specifically, they argue that the continued use of what is sometimes referred to as a “medical model” of assessment – within which educational difficulties are explained solely in terms of a child’s deficits – prevents progress in the field, not least because it distracts attention from questions about why schools fail to teach so many children successfully.

The orientation I am adopting has been defined as a “social model.” Rather than focusing on the characteristics of learners, this involves an assessment of contextual barriers facing some learners and an identification of resources, particularly human resources, that can be mobilized in order to overcome these difficulties. The implication is that efforts to promote inclusion and equity should be seen as what my colleagues and I referred to as “school improvement with attitude” (Ainscow et al., 2006).

This means that the focus is on the reform of policies and practices in ways that will lead them to respond positively to student diversity – seeing individual differences not as problems to be fixed, but as opportunities for enriching learning. Within such a conceptualization, a consideration of difficulties experienced by students can provide an agenda for change and insights as to how such changes might be brought about. However, this kind of approach is more likely to be successful in contexts where there is a culture of collaboration that encourages and supports problem-solving (Ainscow, 2016; Skrtic, 1991).

These themes are further supported by a review of international research literature that examines the effectiveness of school actions in promoting inclusion (Dyson et al., 2004). In summary, this suggests that some schools are characterized by an “inclusive culture.” Within such schools, there is a degree of consensus among adults around values of respect for difference and a commitment to offering all students access to learning opportunities.

The extent to which such inclusive cultures lead directly and unproblematically to enhanced student participation is not clear. Some aspects of these cultures, however, can be seen as participatory by definition. For instance, respect for diversity from teachers may itself be understood as a form of participation by children within a school community. Moreover, schools with such cultures are also likely to be characterized by forms of organization (such as specialist support being made within the ordinary classroom, rather than by withdrawal for separate attention) and practices (such as cooperative group work) which could be regarded as participatory by definition.

All of this means that attempts to develop inclusive schools should pay attention to the building of consensus around inclusive values within school communities. This implies that school leaders should be selected in the light of their commitment to inclusive values and their capacity to lead in a participatory manner (Riehl, 2000). Finally, the external policy environment should be compatible with inclusive developments in order to support rather than undermine the efforts of schools.

Involving the wider community

In order to foster inclusion and equity in education, governments need to mobilize human and financial resources, some of which may not be under their direct control. Forming partnerships among key stakeholders who can support and own the process of change is therefore essential. These stakeholders include: parents/caregivers; teachers and other education professionals; teacher trainers and researchers; national, local and school-level administrators and managers; policy-makers and service providers in other sectors (e.g., health, child protection and social services); civic groups in the community; and members of minority groups that are at risk of exclusion.

Family involvement is particularly crucial. In some countries, parents and education authorities already cooperate closely in developing community-based programmes for certain groups of learners, such as those who are excluded because of their gender, social status or impairments (Miles, 2002). A logical next step is for these parents to become involved in supporting change for developing inclusion in schools.

Where parents lack the confidence and skills to participate in such developments, it might be necessary to engage and build capacity and networks. This could include the creation of parent support groups, training parents to work with their children, or building the advocacy skills of parents to negotiate with schools and authorities. Here, it is worth adding that we have evidence that the views of families, including children themselves, can be helpful in bringing new thinking to the efforts of schools to develop more inclusive ways of working (Calderón-Almendros et al., 2020).

All of this means changing how families and communities work, and enriching what they offer to children. In this respect there are many encouraging examples of what can happen when what schools do is aligned in a coherent strategy with the efforts of other local players – employers, community groups, universities and public services (Kerr et al., 2014). This does not necessarily mean schools doing more, but it does imply partnerships beyond the school, where partners multiply the impacts of each other's efforts.

This has implications for the various key stakeholders within education systems. In particular, teachers, especially those in senior positions, have to see themselves as having a wider responsibility for all children, not just those that attend their own schools. They also have to develop patterns of internal organization that enable them to have the flexibility to cooperate with other schools and with stakeholders beyond the school gate.

This approach draws on the principles underpinning the highly acclaimed Harlem Children's Zone in the USA (Whitehurst and Croft, 2010). These involve efforts to improve outcomes for children and young people in areas of disadvantage through an approach that they characterize as being "doubly holistic". That is to say, they seek to develop coordinated efforts to tackle the factors that disadvantage children and enhance the factors which support them, across all aspects of their lives, and across their life spans, from conception through to adulthood. Dobbie and Fryer (2009) describe the Children's Zone as "arguably the most ambitious social experiment to alleviate poverty of our time" (p. 1).

Another American initiative, Strive Together, acts as a central backbone organization for sites using similar ideas across the country that are locally tailored (Grossman, Lombard, and Fisher, 2014). These initiatives are guided by indicators that span young people's lives "from cradle to career," with progress determined using data at all stages.

The role of administrative departments

Policy is made at all levels of an education system, not least at the school and classroom levels (Ball, 2010). As such, the promotion of equity and inclusion is not simply a technical or organizational change – it is a movement in a clear philosophical direction. Moving to more inclusive ways of working therefore requires changes across an education system. These span from shifts in policy-makers' values and ways of thinking, which enable them to provide a vision shaping a culture of inclusion, to significant changes within schools and classrooms.

A culture of inclusion in education requires a shared set of assumptions and beliefs among policy-makers and senior staff at the national, district and school level that value differences, believe in collaboration, and are committed to offering educational opportunities to all students. However, we have found that changing the cultural norms that exist within an education system is difficult to achieve, particularly within a context that is faced with so many competing pressures and where practitioners tend to work alone in addressing the problems they face (Ainscow et al., 2020). Leaders at all levels, including those in civil society and other sectors, therefore have to be prepared to analyze their own situations, identify local barriers and facilitators, plan an appropriate development process, and provide leadership for inclusive practices and effective strategies for monitoring equity in education.

National and district administrators have important roles in promoting inclusive ways of managing schools and education processes. In particular, they have to establish the conditions for challenging discriminatory educational practices. They also need to establish the conditions that build consensus and commitment toward putting the universal values of inclusion and equity into practice. Certain forms of leadership are known to be effective in promoting inclusion in education (Riehl, 2000). These approaches focus attention on teaching and learning; they create strong supportive communities of students, teachers and parents; they nurture the understanding of a culture of education among families; and, they multi-agency support.

There is also evidence that school-to-school collaboration can strengthen the capacity of individual organizations to respond to learner diversity (Muijs et al., 2011). Specifically, collaboration between schools can help to reduce the polarization of schools, to the particular benefit of those students who are marginalized at the edges of the system. In addition, there is evidence that when schools seek to develop more collaborative ways of working, this can have an impact on how teachers perceive themselves and their work. Specifically, comparisons of practices in different schools can lead teachers to view underachieving students in a new light. In this way, learners who cannot easily be educated within the school's established routines are not seen as "having problems," but as challenging teachers to re-examine their practices in order to make them more responsive and flexible.

Certain conditions are necessary in order to make school-to-school collaboration effective (Chapman and Ainscow, 2021). In summary, these are as follows:

- The development of positive relationships among groups of schools, in some instances across the borders of local authorities;
- The presence of incentives that encourage key stake holders to explore the possibility that collaboration will be in their own interests;
- Senior staff in schools who are willing and skilled enough to drive collaboration forward toward collective responsibility, while coping with the inevitable uncertainties and turbulence;
- The creation of common improvement agendas that are seen to be relevant to a wide range of stakeholders; and

- Coherent external support provided by credible consultants/advisers (from the local authority or elsewhere) who have the confidence to learn alongside their school-based partners, exploring and developing new roles and relationships where necessary.

Local coordination is needed in order to encourage such forms of collaboration (Armstrong and Ainscow, 2018). Indeed, a recent report noted that four of the most successful national education systems - Singapore, Estonia, Finland, and Ontario – each has a coherent “middle tier,” regardless of their differing extents of school autonomy or devolution of decision-making (Bubb et al., 2019). In particular, they all had district level structures that offer a consistent view that, to maintain equity as well as excellence, there needs to be an authoritative co-ordinating influence with local accountability.

Having analyzed two relatively successful large-scale improvement initiatives, Andy Hargreaves and I have suggested a way of supporting local authorities in responding to these new demands (Hargreaves and Ainscow, 2015). We argue that, in taking on new roles, districts can provide a valuable focus for school improvement, be a means for efficient and effective use of research evidence and data analysis across schools, support schools in responding coherently to multiple external reform demands, and be champions for families and students, making sure everybody gets a fair deal.

The problem is, of course, that not all local school systems are strong. Therefore, a way to reduce variation among school districts is to promote collaboration among them so they share resources, ideas, and expertise, and exercise collective responsibility for student success. In adopting this “leading from the middle” approach, districts can become the collective drivers of change and improvement together.

Finally, there are implications here for national accountability and inspection systems. In particular, the research on the benefits of school-to-school cooperation suggests that it is now time for school evaluation to be carried out *by schools for schools*, in ways that can act as a stimulus for improvement. This echoes the recommendations of Fullan et al. (2015) who argue for a move away from a heavy reliance on external accountability toward an investment in the professional capital of teachers and school leaders. However, this has to be challenging and credible. In other words, it must not involve forms of collusion within which partner schools endorse one another in an acceptance of mediocrity. Here, again, evidence is an essential ingredient.

Conclusion

In summary, then, the ideas that have emerged from this review of international experiences, research and policy documents regarding ways of promoting equity and inclusion are as follows:

- Policies should be based on clear and widely understood definitions of what the terms inclusion and equity mean;
- Strategies should be informed by evidence regarding the impact of current practices on the presence, participation and achievement of all students;
- There should be an emphasis on whole-school approaches in which teachers are supported in developing inclusive practices;
- Policies should draw on the experience and expertise of everybody who has an involvement in the lives of students, including the children themselves; and
- Education departments must provide leadership in the promotion of inclusion and equity as principles that guide the work of teachers in all schools

Building on these key ideas, I recommend that those wishing to promote inclusion and equity within their national education system should focus on the following agenda for reform:

In terms of principles

- Are there agreed definitions of inclusion and equity in education?
- If so, are these definitions widely known and understood across the education system?

In terms of use of evidence

- Are systems in place to monitor the presence, participation and achievement of all school students?
- Is there awareness of which students are left behind or left out?
- Do schools use evidence to stimulate efforts to include all students?

In terms of school development

- Are teachers encouraged to develop practices that are intended to include all students?
- Do schools have effective professional development programmes related to inclusive practices?
- Do school leaders emphasize the importance of inclusive education?

In terms of community involvement

- To what extent are families seen as partners in supporting their children’s education?
- Do schools collaborate with other professionals, such as those involved in the health and social services?
- Do schools link with the efforts of other local players, e.g. employers, community groups, universities and public services?

In terms of the roles of education departments

- To what extent do national policy-makers encourage the development of inclusive practices?
- Do local district administrators take actions to promote inclusion and equity?

This reform agenda indicates that the promotion of inclusion and equity in education is less about the introduction of particular techniques or new organizational arrangements, and much more about processes of social learning within particular contexts. As I have argued, the use of evidence as a means of stimulating experimentation and collaboration is seen as a central strategy. As Copland (2003) suggests, inquiry can be the “engine” to enable the distribution of leadership that is needed in order to foster participation in learning, and the “glue” that can bind a community together around a common purpose.

All of this has major implications for leadership practice within schools and education systems. In particular, it calls for efforts to encourage coordinated and sustained efforts around the idea that changing outcomes for vulnerable groups of students is unlikely to be achieved unless there are changes in the thinking of adults. Consequently, the starting point must be with policy-makers and practitioners: in effect, enlarging their capacity to imagine what might be achieved, and increasing their sense of accountability for bringing this about. This may also involve tackling taken-for-granted assumptions, most often relating to expectations about certain groups of students, their capabilities and behaviors.

References

- Ainscow, M., 2005. Developing inclusive education systems: what are the levers for change? *J. Educ. Change* 6 (2), 109–124.
- Ainscow, M., 2007. Taking an inclusive turn. *J. Res. Spec. Educ. Needs* 7 (1), 3–7.
- Ainscow, M., 2016a. *Struggles for Equity in Education: The Selected Works of Mel Ainscow*. Routledge World Library of Educationalists series, London.
- Ainscow, M., 2016b. Collaboration as a strategy for promoting equity in education: possibilities and barriers. *J. Prof. Capit. Commun.* 1 (2), 159–172.
- Ainscow, M., 2020. Promoting inclusion and equity in education: lessons from international experiences. *Nord. J. Stud. Educ. Pol.* 6 (1), 7–16.
- Ainscow, M., Booth, T., Dyson, A., with Farrell, P., Frankham, J., Gallannaugh, F., Howes, A., Smith, R., 2006. *Improving Schools, Developing Inclusion*. Routledge, London.
- Ainscow, M., Chapman, C., Hadfield, M., 2020. *Changing Education Systems: A Research-Based Approach*. Routledge.
- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012. *Developing Equitable Education Systems*. Routledge, London.
- Ainscow, M., Messiou, K., 2017. Engaging with the views of students to promote inclusion in education. *J. Educ. Change* 19 (1), 1–17.
- Alves, I., Campos Pinto, P., Pinto, T.J., 2020. Developing inclusive education in Portugal: evidence and challenges. *Prospects*. <https://doi.org/10.1007/s11125-020-09504-y>.
- Armstrong, P., Ainscow, M., 2018. School-to-school support within a competitive education system: views from the inside. *Sch. Effect. Sch. Improv.* 29 (4), 614–633.
- AuCoin, A., Porter, G.L., Baker-Korotkov, K., 2020. New Brunswick's journey to inclusive education. *Prospects*. <https://doi.org/10.1007/s11125-020-09508-8>.
- Ball, S.J., 2010. New class inequalities in education. *Int. J. Sociol. Soc. Pol.* 30 (3/4), 155–166.
- Booth, T., Ainscow, M., 2002. *The Index for Inclusion*. Centre for Studies on Inclusive Education, Bristol.
- Bubb, S., Crossley-Holland, J., Cordiner, J., Cousin, S., Earley, P., 2019. *Understanding the Middle Tier: Comparative Costs of Academy and LA-maintained School Systems*. Sara Bubb Associates, London.
- Calderón-Almendros, I., Ainscow, M., Bersanelli, S., Molina-Toledo, P., 2020. Educational inclusion and equity in Latin America: an analysis of the challenges. *Prospects* 49, 169–186.
- Chapman, C., Ainscow, M., 2021. *Educational Equity: Pathways to Success*. Routledge, London.
- Copland, M.A., 2003. Leadership of inquiry: building and sustaining capacity for school improvement. *Educ. Eval. Pol. Anal.* 25 (4), 375–395.
- Dyson, A., Howes, A., Roberts, B., 2004. What do we really know about inclusive schools? a systematic review of the research evidence. In: Mitchell, D. (Ed.), *Special Educational Needs and Inclusive Education: Major Themes in Education*. Routledge, London.
- Fulcher, G., 1989. *Disabling Policies? A Comparative Approach to Education Policy and Disability*. Falmer, London.
- Fullan, M., Rincon-Gallardo, S., Hargreaves, A., 2015. Professional capital as accountability. *Educ. Pol. Anal. Arch.* 23 (15). <https://doi.org/10.14507/epaa.v23.1998>.
- Hargreaves, A., Ainscow, M., 2015. The top and bottom of leadership and change. *Phi Delta Kappan*.
- Huberman, M., 1993. The model of the independent artisan in teachers' professional relationships. In: Little, J.W., McLaughlin, M.W. (Eds.), *Teachers' Work: Individuals, Colleagues and Contexts*. Teachers College Press, New York.
- lanes, D., Demo, H., Dell'Anna, S., 2020. Inclusive education in Italy: historical steps, positive developments, and challenges. *Prospects*. <https://doi.org/10.1007/s11125-020-09509-7>.
- Kerr, K., Dyson, A., Raffo, C., 2014. *Education, Disadvantage and Place: Making the Local Matter*. Policy Press, Bristol.
- Lewis, C., Perry, R., Murata, A., 2006. How should research contribute to instructional improvement? The case of lesson study. *Educ. Res.* 35 (3), 3–14.
- Messiou, K., 2019. The missing voices: students as a catalyst for promoting inclusive education. *Int. J. Incl. Educ.* 23 (7/8), 768–781.
- Messiou, K., Ainscow, M., Echeita, G., Goldrick, S., Hope, M., Paes, I., et al., 2016. Learning from differences: a strategy for teacher development in respect to student diversity. *Sch. Effect. Sch. Improv.* 27 (1), 45–61.
- Messiou, K., Ainscow, M., 2015. Engaging with the views of students: a catalyst for powerful teacher development? *Teach. Teach. Educ.* 51 (2), 246–255.
- Meijer, C.J.W., Watkins, A., 2019. Financing special needs and inclusive education: from Salamanca to the present. *Int. J. Incl. Educ.* 23 (7/8), 705–721.
- Miles, S., 2002. *Family Action for Inclusion in Education*. Enabling Education Network, Manchester.
- Ministry of Education, 2019. *Supporting all schools to succeed*. Wellington, New Zealand.
- Muijs, D., Ainscow, M., Chapman, C., West, M., 2011. *Collaboration and Networking in Education*. Springer, London.
- OECD, 2012. *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. OECD Publishing, Paris.
- OECD, 2018. *Education at a Glance 2018: OECD Indicators*. OECD Publishing, Paris.
- Opertti, R., Walker, Z., Zhang, Y., 2014. Inclusive Education: from targeting groups and schools to achieving quality education as the core of EFA. In: Florian, L. (Ed.), *The SAGE Handbook of Special Education*, 2nd Revised Edition. SAGE, London.
- Riehl, C.J., 2000. The principal's role in creating inclusive schools for diverse students: a review of normative, empirical, and critical literature on the practice of educational administration. *Rev. Educ. Res.* 70 (1), 55–81.
- Sabel, C., Saxenian, A., Miettinen, R., Kristenson, P.H., Hautamaki, J., 2011. *Individualized Service Provision in the New Welfare State: Lessons from Special Education in Finland*. SITRA, Helsinki.
- Senge, P.M., 1989. *The Fifth Discipline: The Art and Practice of the Learning Organisation*. Century, London.
- Skrtic, T., 1991. *Behind Special Education: A Critical Analysis of Professional Culture and School Organization*. Love, Denver.
- Slee, R., 2010. *The Irregular School: Exclusion, Schooling, and Inclusive Education*. Routledge, London.
- Takala, M., Pirttimaa, R., Tormane, M., 2009. Inclusive special education: the role of special education teachers in Finland. *Br. J. Spec. Educ.* 36 (3), 162–172.
- UIS, 2019. Fact Sheet No. 56. UNESCO, Paris. <http://uis.unesco.org/sites/default/files/documents/new-methodology-shows-258-million-children-adolescents-and-youth-are-out-school.pdf>.
- UNGEI, 2012. *Gender analysis in education: a conceptual overview*. United Nations Girls' Education Initiative.

- UNESCO, 1990. World Declaration on Education for All. UNESCO, Paris.
- UNESCO, 1994. Final Report: World Conference on Special Needs Education: Access and Quality. UNESCO, Paris.
- UNESCO, 2000. World Education Forum, Dakar, Senegal, 26-28 April 2000: Final Report. UNESCO, Paris.
- UNESCO, 2015. Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4. UNESCO, Paris.
- UNESCO, 2017. A Guide for Ensuring Inclusion and Equity in Education. UNESCO, Paris.
- UNESCO, 2019. Cali Commitment to Equity and Inclusion in Education. UNESCO, Paris.
- UNESCO, 2020. Global Education Monitoring Report 2020: Inclusion and Education: All Means All. UNESCO, Paris.
- UNICEF, 2019. Fast Facts: 10 Facts Illustrating Why We Must #EndChildMarriage. UNICEF, New York. Available at: <https://www.unicef.org/press-releases/fast-facts-10-facts-illustrating-why-we-must-end-child-marriage>.
- United Nations, 2008. Convention on the Rights of Persons with Disabilities. United Nations, New York.

Re-examining the education for all agenda in “apocalyptic” times

Marnie Best, Centre for Research in Educational and Social Inclusion, Education Futures, University of South Australia, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	12
Educational inclusion	12
Toward an inclusive education for all	13
The education for all agenda	14
A sudden shift for teaching and learning in “apocalyptic” times	15
Conclusion	17
References	18

Introduction

The impact of the global pandemic has been far reaching. It has affected almost all aspects of our lives. It has disrupted schooling and the workforce more broadly, upending education and the psychological, physical, social, emotional and cognitive development, and wellbeing of many. While it has interrupted education and employment, it has also resulted in lost lives and livelihoods, isolation and separation from family and loved ones, and created a time of uncertainty and fear. What the pandemic has achieved is a simultaneous sense of togetherness and separation.

Despite the insecurity and trepidation the pandemic has brought about, such a significant global event has provided opportunity for alternative perspectives on how we approach and engage with education. The pandemic has afforded a unique time to re-examine and re-imagine more equitable societies and education systems (Arnone, 2020), particularly given that the consequences of the pandemic have impacted the most vulnerable the hardest (Schleicher, 2020). It is through both subtle and overt forms of educational and social exclusion that we are forced to consider more inclusive ways of thinking, knowing and doing.

The global pandemic is arguably the greatest challenge that education systems have ever experienced (Daniel, 2020). As Ainscow (1995, pp. 11–12) wrote some 25 years preceding the pandemic, “traditional school cultures, supported by fixed organizational arrangements and high levels of specialisms among staff who are geared to predetermined tasks, are in trouble when faced with unexpected circumstances.” It was precisely these unexpected circumstances that challenged traditional understandings of “traditional” school cultures and organizational arrangements. Put simply, most educational systems were ill-equipped and unprepared for a rapid shift in education from face-to-face delivery to online teaching and learning arrangements. It also signaled that schools provided far more than education; they provided a place of safety, social connectedness, connections to interagency supports, meals and access to learning resources (Stanistreet et al., 2020). For many learners, access to these basic provisions ceased during months of government-imposed lockdowns in an effort to control the spread of the virus. An estimated 90% of the global student population were directly affected by school closures due to the pandemic (Global Education Monitoring Report, 2020).

This chapter aims to re-examine the Education for All [EFA] agenda in what has been termed “apocalyptic” times: apocalyptic to capture the momentous and seemingly catastrophic times we find ourselves in. One in which there seems to be no end and few affordances for escape. While the global pandemic served to create barriers, exclusion and distance, it simultaneously served to bring unity and often unexpected change. These changes have permeated what appear to be all aspects of society and shifted how we conceptualize and enact education at local, national and global levels.

Educational inclusion

Inclusion cannot be considered independent of exclusion. In so doing, this chapter considers how the global pandemic has influenced how we understand inclusion and exclusion, from educational and societal perspectives, and the subsequent challenges and opportunities such an event created. In order to do so, there is a necessity to clarify and distinguish between key terms often misunderstood or used interchangeably within the field of inclusive education: equity, equality and inclusion.

Central to discussion regarding inclusion is the need to differentiate between equity and equality, both in terms of policy and practice. While each centers on fairness, they have vastly different meanings. In essence, to distinguish between the terms is to equate equity to fairness and equality to sameness. That is, equity means that every learner is afforded the kinds of opportunities they require, reflective of their specific circumstance. Equality, however, means that all students receive the same resources and opportunities, almost irrespective of their specific circumstance. Inclusion, however, is comparatively more difficult to define. Bearing similarity to equity, inclusion may be considered as “a process: actions and practices that embrace diversity and build a sense of belonging, rooted in the belief that every person has value and potential and should be respected” (Global Education Monitoring Report, 2020, p. 11).

Given the common blurring between these related terms, it is perhaps not surprising therefore that the field of inclusive education “remains confused about the actions that need to be taken to move policy and practice forward” (Ainscow and Miles, 2008, p. 16). While some countries situate inclusive education as an approach to serving children with disability within general or mainstream education settings, there is a shift toward conceptualizing inclusive education as a reform that supports and embraces the diversity of all learners. Within this notion is the premise that “education is a basic human right and the foundation of a more just society” (Ainscow and Miles, 2008, p. 16). And yet, less than 10% of countries have laws that aim to support full inclusion in education (Global Education Monitoring Report, 2020) and of the 68% of countries that have defined inclusive education, just 57% of definitions include multiple marginalized groups (Global Education Monitoring Report, 2020). When we consider an “education for all,” it seems apparent that there are many marginalized groups that appear to fall outside the understanding of “all.” Through stigmatizing, stereotyping and discrimination, millions of children and young people are alienated and excluded within classrooms and school settings (Global Education Monitoring Report, 2020).

As acknowledged elsewhere in this chapter, education is a far broader concept than the acquisition of skills and knowledge. If we continue to consider what inclusive education entails, we must appreciate that it aims to “promote democratic principles and a set of values and beliefs relating to equality and social justice so that all children can participate in teaching and learning” (Miles and Singal, 2010, p. 11). As Ainscow et al. (2019, pp. 675–676) have argued, the implication of many inclusive education policies and agendas is that “moving in an inclusive direction continues to involve struggles to find the most appropriate ways of moving forward ... the promotion of inclusion is not simply a technical or organizational change—it is a movement in a clear philosophical direction.” In critically examining the underpinnings of inclusive education, we are therefore afforded:

an opportunity for society to examine critically its social institutions and structures. It necessarily challenges didactic, teacher-centred teaching practices, such as rote learning, and so opens up opportunities for developing better pedagogy and greater competence. [Education For All] often fails to explore such broad issues.

Miles and Singal (2010, p. 12).

Therefore, and as Miles and Singal (2010, p. 12) have reasoned, “radical changes are required in education systems, and in the values and principles of the people involved in delivering education, if the world’s most vulnerable and disadvantaged children are to gain access to their local school.” In essence, inclusive education can be considered as both a philosophy and a practice which is underpinned through local, national and global policy ideals.

Toward an inclusive education for all

A commitment to in-school inclusion deviates from an historical tendency to exclude learners with disability from education or to segregate learners to specialized settings. Such a commitment clearly distinguishes inclusion from integration and can be seen in global policies that have made an intentional shift toward an inclusive education for all, acknowledging education as a basic human right.

While education was recognized as a human right in 1948, the call for inclusion came some time later. However, calls for inclusion are not new, with international declarations regarding the commitment to non-discrimination made since 1960 and to inclusion since 1990 (Global Education Monitoring Report, 2020). The 1994 Salamanca World Conference on Special Needs Education endorsed the notion of inclusive education (UNESCO, 1994), with the aim of furthering the Education for All objective by considering the fundamental policy shifts required to promote inclusion (Ainscow et al., 2019). As a result, the *Salamanca Statement and Framework for Action on Special Needs Education* is arguably the most significant international document to appear in the field of special education. It argues that regular or mainstream schools that have an inclusive orientation are “the most effective means of combating discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all” (United Nations Educational Scientific and Cultural Organization (UNESCO), 1994, p. 10).

The 2006 UN Convention on the Rights of Persons with Disabilities (CRPD) likewise advocated for the right to inclusive education. The General Comment No. 4 to CRPD Article 24 positioned inclusive education as “a process ... to provide all students ... with an equitable and participatory learning experience and environment that best corresponds to their requirements and preferences.” Such understanding is echoed within related global policies, including Education for All and the United Nations 2030 Agenda for Sustainable Development, with its fourth Sustainable Development Goal (SDG 4) committing to “ensur[ing] inclusive and equitable quality education and promote lifelong learning opportunities for all” (United Nations, 2020, p. 32).

However, exclusion persists. And perhaps, it is no timelier, than in the midst of a global pandemic, for this to be illustrated so starkly. In monitoring the progress of education targets adopted by UN Member States in the 2030 Agenda for Sustainable Development across 209 countries, 258 million children and young people were entirely excluded from education (Global Education Monitoring Report, 2020). For those who remained connected to education, there was a considerable population whom were inadequately supported. As Global Education Monitoring Report (2020) data has illustrated, 40% of the world’s poorest countries failed to provide adequate support to learners at risk during the pandemic, thereby further marginalizing some of the most vulnerable children and young people. Those who were most at risk of exclusion, exacerbated during the pandemic, included students living

remotely, in poverty, from linguistic and cultural minorities and learners with disability. In consideration of an “education for all,” how this was understood and transposed to “apocalyptic” times remained problematic.

The education for all agenda

The Education for All agenda emerged from the 1990 World Conference on Education for All, held in Jomtien, Thailand. Promoting the event through the slogan, “EFA by the year 2000”, the event brought together 155 countries to commit to achieving an education for all. Key to this commitment was a pledge to provide primary education “for all” by the year 2000 and to “ensure that children and adults would benefit from educational opportunities designed to meet their basic learning needs” (Bloom and Cohen, 2002, p. 86). The conference was particularly significant for two key reasons: firstly, it acknowledged the considerable population of vulnerable and marginalized groups of learners excluded from education systems worldwide and secondly, it outlined a vision of education as being far broader than just schooling.

Underpinning the Education for All agenda was an international commitment to ensuring that all children and adults receive a good quality education. In exploring the Agenda, there is a need to understand the justification behind such a movement. The provision of an “Education for All” has been reasoned on four grounds: humanitarian, sociological, political, and economic (Bloom and Cohen, 2002). As Bloom and Cohen (2002, pp. 87–88) describe, they are:

- humanitarian: education enables human beings to develop their capacities so that they can lead fulfilling and dignified lives;
- sociological: social and cultural capital, which are crucial ingredients in the development process and ones that enhance the operation of other development channels, can be greatly strengthened by education;
- political: education is popular among voters; and
- economic: education, or investment in people’s capacities, raises people’s productivity and provides a foundation for rapid technological change.

Building on these justifications, the Education for All agenda has presented a powerful force in positioning education as a rights issue, acknowledging that education is central to developing economies, fostering the disability movement, and presenting a stronger realization that education is essential for “global tolerance” (Miles and Singal, 2010, p. 11).

Despite the direction of the Jomtien Declaration, the call to hold the Salamanca World Conference on Special Needs Education only four years following appeared to imply that the phrase “education for all” may have really meant “almost all,” continuing the historical assumption that “a small percentage of children have to be seen as ‘outsiders’, whose education must be catered for by a separate, parallel system, usually known as special education” (Ainscow and Miles, 2008, p. 16).

Progress in meeting the goals underpinning the Education for All agenda were reviewed 10 years following the 1990 World Conference on Education for All, at the World Education Forum in Dakar, Senegal. The event was attended by 1100 participants representing 164 countries. By that time, considerable advances had been made, with some countries nearing the target of achieving universal primary education. As such, six key education goals, designed to meet the learning needs of all children, young people and adults by 2015, were identified. They included:

- Goal 1: Expand early childhood care and education
- Goal 2: Provide free and compulsory primary education for all
- Goal 3: Promote learning and life skills for young people and adults
- Goal 4: Increase adult literacy by 50%
- Goal 5: Achieve gender parity by 2005, gender equality by 2015
- Goal 6: Improve the quality of education (United Nations, 2000)

Such commitment to Education for All was reinforced through the eight Millennium Development Goals (The World Bank, 2020). The goals were visioned to contribute toward building a better world in the 21st century, with the aim to:

- Goal 1: Eradicate extreme poverty and hunger
- Goal 2: Achieve universal primary education
- Goal 3: Promote gender equality and empower women
- Goal 4: Reduce child mortality
- Goal 5: Improve maternal health
- Goal 6: Combat HIV/AIDS, malaria, and other diseases
- Goal 7: Ensure environmental sustainability
- Goal 8: Develop a global partnership for development (The World Bank, 2020)

Evident within the Millennium Development Goals was a commitment to eradicate extreme poverty and hunger and to achieve universal primary education (The World Bank, 2020). Inherent within the Goals was a pledge to intercept the cycle of exclusion from education which often contributes to poverty and broader social exclusion. However, concern enveloped the EFA agenda in that there appeared to be a narrowing to the second Millennium Development Goal, noting that an “education for all” had funneled toward a narrow focus on five years of compulsory schooling, from ages 6 to 11 years, inferring the view that 5 years of primary schooling will suffice (Miles and Singal, 2010). Such inference is therefore problematic to truly realizing an education

for all. We therefore need to problematize the relationship between the Education for All agenda and inclusive education. As Ainscow and Miles (2008, p. 20) have argued:

progress in relation to both the EFA agenda and inclusive education requires greater clarity about what becoming more inclusive involves. For us, it is a principled approach to education, which involves:

- The process of increasing the participation of students in, and reducing their exclusion from, the curricula, cultures and communities of local schools;
- Restructuring the cultures, policies and practices in schools so that they respond to the diversity of students in their locality; [and]
- The presence, participation and achievement of all students vulnerable to exclusionary pressures, not only those with impairments or those who are categorized as “having special educational needs.”

In considering Education for All and inclusive education as separate entities, we create “an unhelpful and wasteful polarization between EFA and inclusive education” (Miles and Singal, 2010, p. 1) in that inclusive education is often perceived in terms of overcoming barriers to learning for all children and young people, thereby “filling” the gap of EFA. Contributing to this, in the context of Southern countries, inclusive education tends to focus on students with disability. The challenge, therefore, for educators and educational systems, is to reconceptualize the relationship between the Education for All agenda and inclusive education through increasing collaboration and synergy between these interrelated agendas (Miles and Singal, 2010). Where the Education for All agenda and inclusive education converge however, is in the commitment to providing quality education for all children and young people.

In conceptualizing inclusive education in this way, it captures several key philosophical underpinnings. That is, inclusion:

- involves all learners in educational contexts;
- is centered on presence, participation and achievement;
- and exclusion are related, in so far that inclusion actively aims to respond to exclusion; and
- is an ever-revolving process. That is, a truly inclusive school is fluid (Ainscow and Miles, 2008) and responsive to given changes.

To advocate for inclusive education is to actively shift away from explanations of educational failure that focus on individual deficit to critically examining the barriers to education that are created by educational policies, practices and systems (Best, 2016). Therefore, in building on the legacy of the *Salamanca Statement and Framework for Action on Special Needs Education*, in order to progress an inclusive education, we must focus on developing inclusive education systems (Ainscow et al., 2019). In moving a commitment to an education for all forward in “apocalyptic” times, attention must shift to addressing the cultural and contextual fitness of educational programs which specifically focus on addressing educational and social inequity (Miles and Singal, 2010).

As this chapter has highlighted, the field of inclusive education is entrenched with “uncertainties, disputes and contradictions” (Ainscow and Miles, 2008, p. 20). Notwithstanding this, attempts have been made to respond to the educational needs for learners in more inclusive and equitable ways. The challenge now is to continue these movements in light of the dynamic challenges that a global pandemic has presented.

A sudden shift for teaching and learning in “apocalyptic” times

The pandemic has served to both highlight and deepen inequality and expose the seeming fragility of societies worldwide. In exposing such inequality and fragility, there is a “collective responsibility” (Global Education Monitoring Report, 2020, p. iii) to support those who are most vulnerable and disadvantaged. While the pandemic has much to account for, many of these inequities were apparent prior to COVID-19, with one in five children, adolescents and young people excluded from education in entirety (Global Education Monitoring Report, 2020). This worldwide circumstance has served to amplify the situation.

It is almost unquestionable that all learners have been affected, in one way or another, by school closures during the pandemic. While it is anticipated that learners have experienced a short-term direct loss of learning, there remains concern regarding longer lasting effects that may be indirectly caused by impending recessions that will steer millions of people closer to the poverty line (Global Education Monitoring Report, 2020). The need to address this in order to strive for an education for all is therefore commanding.

Government-imposed lockdowns due to the pandemic derailed economic activity and severely threatened the livelihoods of billions of people globally. Perhaps most impactful for education, and indeed, for teachers, students and caregivers, was the closure of schools and universities which lasted for at least 10 weeks in most OECD and partner countries (Schleicher, 2020). At the apparent peak of the closure period, occurring in April 2020, 91% of the global student population was affected in 194 countries (Global Education Monitoring Report, 2020). By the end of May 2020, national school closures remained in effect in 150 countries, affecting 68% of the global education population (Global Education Monitoring Report, 2020). Because of these

government-imposed lockdowns, educational policy makers and practitioners were confronted with tough decisions. What was evident was a distinct need to deliver education in non-traditional ways, including through technologically-mediated platforms.

Perhaps nowhere are the inequities emanating from the pandemic highlighted more patently than in education systems. As [Armove \(2020, p. 43\)](#) notes, “the individuals most marginalized and discriminated against have suffered the greatest from the closure of schools and the efforts to reach students with online instruction.” When we consider what an education for all in apocalyptic times means, we need to understand that education necessitated a difference in terms of delivery. For some students, this process was one premised on inclusion, while for others, it exacerbated exclusion. For example, 64% of some of the poorest, low-income countries used radio for primary education, as compared to 42% for upper-middle-income countries ([Global Education Monitoring Report, 2020](#)). Further, 74% of lower-middle-income countries relied on television to support the delivery of primary education, compared to 36% of low-income countries. However, when we consider how online platforms facilitated the delivery of learning experiences, they were used in approximately 55% of low-income, 73% of lower-middle-income and 93% of upper-middle-income countries ([Global Education Monitoring Report, 2020](#)).

The interruption caused by a quick shift to online learning was untested and unprecedented. It created disruption to many families around the world. While it inevitably impacted caregivers’ productivity, it likewise impacted learners’ social lives and learning ([Burgess and Sievertsen, 2020](#)).

Such a rapid transition was complex and comprised an intricate mix of urgency, readiness to engage with digital platforms, online teaching tools, digital and technological fluency, while compounded by a need to respond to emotions of fear and the distancing created through social isolation ([Agnoletto and Queiroz, 2020](#)).

Such a sudden and somewhat unexpected closure of schools and shift to online learning served to further marginalize and discriminate against some of the most vulnerable learners ([Armove, 2020](#)). In consideration of an education for all, such a radical shift created barriers for some students accessing education. That is, education suddenly became technologically-mediated and those who did not have access to a computer, the internet, or lived in remote areas with unreliable or unavailable electricity sources were excluded from accessing basic education ([Armove, 2020](#)). For those students who did have a computer and internet at home, many were competing with parents and caregivers for access, due to the swift deployment of work-from-home requirements. So, while some students required technology in order to access education, parents and caregivers required the same technology to undertake paid employment duties. This likewise extended to teachers, many of whom were juggling the home schooling of their own children while simultaneously teaching their own classes remotely. While educators can facilitate online learning from afar, it is often caregivers that are left to field questions and clarify content for learners in their care. For many learners, this meant relying on caregivers who had little or no education. This issue, albeit complex, seems to have been relatively unacknowledged by educational systems in their commitment to an education for all in apocalyptic times.

Enacting effective reform requires moving beyond just reimagining a new vision for education. As [Bloom and Cohen \(2002\)](#) claim, effective reform requires the appreciation of the different goals of education in different cultures, as well as developing the human and technological capacity requisite to meeting those goals. In many ways, the pandemic has signaled the opportunities that digital technologies can provide in enabling access and to and engagement with education ([Best and MacGregor, 2017](#); [Schleicher, 2020](#)). That is, it provides alternative ways of approaching what is learned, how learning occurs, and where and when. Employing digital technologies in an attempt to increase innovation and access enables a shift for the role of educators as knowledge leaders to working alongside students as co-creators, coaches, mentors, and evaluators ([Schleicher, 2020](#)).

Despite a diversity of circumstances that have placed students at risk of educational underachievement and attrition, the social and digital divides created by the pandemic have served to marginalize many of these students further. Global efforts to maintain continuity with learning further exacerbated exclusion. Although education systems pivoted to the online delivery of learning, most of these were considered inferior substitutes to face-to-face classroom instruction. Learners who required more specialized support or assistance were particularly disadvantaged through the delivery of online learning experiences. However, despite increasing reliance and accessibility on technology to mediate learning, there remains a distinct digital divide that illuminates the limitations of such an approach. That is, there are assumptions inherent with online learning: students and educators require access to adequate internet connection, technological equipment, and the skills and conditions to access, engage and utilize such platforms.

The shift to online learning was sudden, with educators, school administrators and educational systems needing to shift—almost overnight—to new tools to design, deliver and distribute learning and associated content and to communicate with students and their caregivers. Traditional approaches to face-to-face learning have required synchronous learning to typically occur in classroom settings. However, in shifting to remote forms of learning, schools and educators have been challenged to adapt to asynchronous approaches to teaching and learning. Asynchronous learning provides educators with increased flexibility in preparing learning materials, while enabling students time to navigate the demands of home and study ([Daniel, 2020](#)). This affords educators increased flexibility in relation to the timing of delivering materials (for example, through asynchronous methods such as wikis, blogs and email) and in the monitoring of students and their respective progress.

However, despite this, perhaps most concerning about considering an education for all in “apocalyptic” times are the implications for learners with disability and those in the early years of school. Research consistently evidences that learners with disability are at an increased risk of school suspension and exclusion ([Graham et al., 2020](#)). While many educators were able to adapt to online teaching and learning with minimal notice, for those who worked with students who had physical impairment, this was a different case. That is, while the technology and resources exist for vision and hearing-impaired learners, these were not always accessible or available for home learning situations. Moreover, learners with learning difficulties often struggle to engage with

independent learning. For many students however, the sheer absence of a daily school routine added a layer of challenge for those who were sensitive to change or unpredictability.

Likewise, it may be our youngest learners who were most negatively impacted during the pandemic. Given that early childhood is very much centered around “providing real-life sensory experiences, hands-on activities and meaningful face-to-face interactions” (Spiteri, 2021, p. 145), such experiences are extremely challenging to achieve in virtual teaching and learning contexts. In moving an “education for all” forward, consideration must be given to how this includes learners in the early years of school, for which technologically-mediated learning appears to have significant limitations.

While schools pivoted to online teaching and learning arrangements with minimal notice, the shift back to face-to-face delivery requires careful consideration. That is, “schools (or parts of schools) at the development extreme may be so over-confident of their innovative capacities that they take on too much too quickly, thus damaging the quality of what already exists” (Ainscow, 1995, p. 12). So how schools embrace new approaches in “educating for all” in unprecedented times, while balancing existing educational pedagogies and practice, requires a careful transition. As Audrey Azoulay, Director-General of UNESCO, has stated, “rethinking the future of education is all the more important following the COVID-19 pandemic, which further widened and put a spotlight on inequalities. Failure to act will hinder the progress of societies” (Global Education Monitoring Report, 2020, p. online).

The global pandemic has clearly evidenced that an education for all in apocalyptic times cannot be achieved through technological solutions alone. While online learning has gained much traction, there remains a population of learners for whom the basic infrastructure to enable engagement with online learning remains elusive. While online learning presents considerable opportunity for pedagogical innovation, this remains contingent on access to technology, internet and electricity. There is a desperate need to focus attention on those “left behind” as schools begin to reopen, in order to address issues of educational and social inclusion and to aim for an education for all.

Conclusion

As reasoned throughout this chapter, the pandemic has served to highlight the many inadequacies and inequities within and across our education systems, from basic access to more complex resourcing (Schleicher, 2020). The global pandemic has evidenced that there remains great provision to reconsider how we understand education in increasingly unpredictable and unstable times. Governments and educational systems need to scrutinize how inclusion and exclusion have been enacted in order to rebuild education systems that are fairer and more accessible for all learners to which they account. While the pandemic has created new layers and understandings of exclusion, primarily related to accessibility, it has also served to marginalize a population of learners through the necessitation of online learning opportunities.

There are two fundamentals if approaches to Education for All are to be inclusive and realized. That is, there must be a common sense of purpose in relation to inclusive education and a consistent use of language (Ainscow and Miles, 2008). Never has this been more opportune than in the midst of a pandemic. We do not need new techniques, but alternative ways of thinking, knowing and doing. There must be “collaboration within and between schools, closer links between schools and communities, networking across contexts, and the collection and use of contextually relevant evidence” (Ainscow and Miles, 2008, p. 15).

While preceding the pandemic, Ainscow’s (1995, p. 12) earlier thoughts captured the quandary that many educators, schools and educational systems now face: that is, “they cannot remain as they now are if they are to respond to new challenges, but at the same time they also need to maintain some continuity between their present and their previous practices.” The pandemic has therefore signaled a call to renew the commitment to the Sustainable Development Goals. How we reimagine education and indeed, an education for all in apocalyptic times, requires us to be malleable to new ideas and responsive to the unexpected in times of the unpredicted. This underwrites a moral and professional commitment to ensure “that all young people have the opportunity to succeed at school and develop the knowledge, skills, attitudes and values that will allow them to contribute to society” (Schleicher, 2020, p. 26). Such a view sits at the core of global education agendas and is the underpinning of education’s promise to learners and to future society more broadly. While the pandemic has challenged how education is understood and conceptualized amidst unprecedented disruption, it is how we move forward as a more informed and resilient society that is key to determining the future of education and schooling.

While this period of turbulence was intensified during the pandemic, it enabled schools to enact change and respond to shifting educational and social demands. Such times and turbulence provide an opportunity to actively and responsively rethink curricula, enabling a shift from the one-size-fits-all approach, toward more individualized approaches that are better tailored toward meeting the educational and social outcomes of all students and to become more authentically inclusive.

Therefore, if we are to strive for more inclusive forms of education, teachers are at the core. From this perspective, teachers’ beliefs, attitudes and actions shape the contexts in which learners engage (or not) with education (Best, 2017). It is particularly important then, to develop education systems that both support educators and their varied practices as they aim to respond to the diverse needs of learners within their care. This is no timelier than as we begin to emerge from a global pandemic.

“In thinking about what inclusive practice involves we also have to be sensitive to the complex nature of teaching” (Ainscow and Miles, 2008, p. 22). Therefore, in order to become more inclusive, we must reconsider how we think and talk, review and refine practice and attempt to create a more inclusive culture and this means that we must acknowledge that inclusion is intertwined with the cultural and social relations in which it develops. As Ainscow and Miles (2008, p. 26) have stated, “it is in the complex

interplay between individuals, and between groups and individuals, that shared beliefs and values exist and change, and it is impossible to separate those beliefs from the relationships in which they are embodied.”

Preceding the pandemic, there was little argument that new thinking on how we design and implement a high-quality education for all was needed (Bloom and Cohen, 2002). Now, intra-pandemic, that new thinking is even more necessary. There is no debate that while primary education is essential, in and of itself, it is insufficient. There remains a considerable population for whom secondary and higher education is not accessible and low standards are endemic in areas of poverty (Bloom and Cohen, 2002). Policy makers and practitioners are increasingly acknowledging systemic failures within education systems. Consequently, an “education for all” has not and cannot be achieved until we design, develop and implement new strategies, systems and determinations that reflect the ideals of the World Conference on Education for All. The pandemic presents opportunity to make a call to (re)examine, (re)consider and (re)design current local, national and global policies to be more inclusive by building unity to intersect disadvantage.

As the pandemic has served to illustrate, real change occurs through the realization of the deep crisis we were confronted with. While it is implausible that we will return to the previous iteration of “normal,” our new “normal,” however that may unfold, will ultimately be shaped through our “collective and systemic responses” (Schleicher, 2020, p. 26) to the pandemic. How we adjust and respond to the disruptions that have arisen in recent times should serve as a timely reminder of our need to continue to (re)examine the Education for All agenda, both in “apocalyptic” times and in moving forward.

References

- Agnoletto, R.N., Queiroz, V.C., 2020. COVID-19 and the challenges in education. *Centro de Estudos Sociedade e Tecnologia* 5 (2).
- Ainscow, M., 1995. Education for All: Making It Happen Keynote Address Presented at the International Special Education Congress (Birmingham, England).
- Ainscow, M., Miles, S., 2008. Making education for all inclusive: where next? *Prospects* 38, 15–34. <https://doi.org/10.1007/s11125-008-9055-0>.
- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676. <https://doi.org/10.1080/13603116.2019.1622800>.
- Arno, R.F., 2020. Imagining what education can be post-COVID-19. *Prospects* 49, 43–46. <https://doi.org/10.1007/s11125-020-09474-1>.
- Best, M., 2016. Wellbeing in alternative education. In: McCallum, F., Price, D. (Eds.), *Nurturing Wellbeing Development in Education: From Little Things, Big Things Grow*. Taylor and Francis, pp. 72–87.
- Best, M., 2017. Transforming pre-service teachers' beliefs and understandings about design and technologies. *Aust. J. Teach. Educ.* 42 (7), 47–65. <https://doi.org/10.14221/ajte.2017v42n7.4>.
- Best, M., MacGregor, D., 2017. Transitioning design and technology education from physical classrooms to virtual spaces: implications for pre-service teacher education. *Int. J. Technol. Des. Educ.* 27 (2), 201–213. <https://doi.org/10.1007/s10798-015-9350-z>.
- Bloom, D.E., Cohen, J.E., 2002. Education for all: an unfinished revolution. *Daedalus* 131 (3).
- Burgess, S., Sievertsen, H., 2020. Schools, Skills, and Learning: The Impact of COVID-19 on Education. <https://voxeu.org/article/impact-covid-19-education>.
- Daniel, S.J., 2020. Education and the COVID-19 pandemic. *Prospects* 1 (6). Advance online publication. <https://doi.org/10.1007/s11125-020-09464-3>.
- Global Education Monitoring Report, 2020. Global Education Monitoring Report, 2020: Inclusion and Education: All Means All - UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf00000373718>.
- Graham, L., McCarthy, T., Killingly, C., Tancredi, H., Poed, S., 2020. Inquiry Into Suspension, Exclusion and Expulsion Processes in South Australian Government Schools.
- Miles, S., Singal, N., 2010. The Education for all and inclusive education debate: conflict, contradiction or opportunity? *Int. J. Incl. Educ.* 14 (1), 1–15. <https://doi.org/10.1080/13603110802265125>.
- Schleicher, A., 2020. The Impact of COVID-19 on Education - Insights From Education at a Glance 2020.
- Spiteri, J., 2021. Quality early childhood education for all and the Covid-19 crisis: a viewpoint. *Prospects* 51, 143–148. <https://doi.org/10.1007/s11125-020-09528-4>.
- Stanistreet, P., Elfert, M., Atchoarena, D., 2020. Education in the age of COVID-19: understanding the consequences. *Int. Rev. Educ.* 66, 627–633. <https://doi.org/10.1007/s11159-020-09880-9>.
- The World Bank, 2020. Millennium Development Goals. World Bank Group. <https://www5.worldbank.org/mdgs/>.
- United Nations, 2000. United Nations Millennium Declaration. A/RES/55/2. https://unstats.un.org/unsd/mdg/Resources/Static/Products/GAResolutions/55_2/a_res55_2e.pdf.
- United Nations, 2020. The Sustainable Development Goals Report. <https://unstats.un.org/sdgs/report/2020/The-Sustainable-Development-Goals-Report-2020.pdf>.
- United Nations Educational Scientific and Cultural Organization, June 7–10, 1994. World Conference on Special Needs Education: Access and Quality (Salamanca, Spain).

Quality and inclusion in SDG 4

William C. Smith^a and Christopher J. Johnstone^b, ^a University of Edinburgh, Edinburgh, United Kingdom; and ^b University of Minnesota, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	19
Evolution of inclusive education	20
The creation of SDG 4	20
Operationalization of SDG 4	21
The utilitarian capture of quality	24
Competing agendas and the re-fragmentation of inclusive education	24
Conclusion	25
References	26

Introduction

Agreed upon by over 180 countries in December 2015, the Sustainable Development Goal for Education (SDG 4) marked an important moment in global education discourse and policy. Highlighted by some as a “quality turn” in education (Sayed and Moriarty, 2020), SDG 4 addressed gaps that emerged from both the Education for All and Millennium Development Goals, the two previous global goals (beginning in 1990 and 2000 respectively) that focused on participation in education. Two of these gaps included the over-emphasis on access to education (at times, at the cost of quality) and the failure for policies focused on “all” to produce meaningful access and outcomes data for historically marginalized populations worldwide.

When suggesting SDG 4 is a “quality turn,” it is important to ask whose definition of quality is being validated? And for whom is quality for? Discussions and definitions of quality are political and embedded in different ideological perspectives. These initial orientations lead to definitions of quality that see quality as universally applied or context specific, highlighting specific aspects of education or the system as a whole, and limiting quality to formal schooling or recognizing the contributions of non-formal and informal learning (Smith and Benavot, 2021).

Utilitarian and humanistic or transformative approaches are considered the two dominant perspectives in education today (Brissett and Mitter, 2017) and each have a different understanding of quality. The aim of education from a utilitarian approach is to prepare “children to work within an established socio-economic order with the ultimate goal of achieving economic growth” (Brissett and Mitter, 2017, p. 195). Associated with this aim is an understanding of quality limited to foundational skills, generally captured through narrow school-based learning outcomes (Benavot and Smith, 2020). In contrast, the humanistic approach sees education as an avenue to address “the inequalities and injustices that are embedded in the larger society” (Maclure et al., 2009, p. 367). This human rights based approach emphasizes equity and a broader conception of quality that includes inputs, processes, and outcomes (Benavot and Smith, 2020).

Inclusive education, and its recognition that all have the right to education, clearly fits into the humanistic camp. Proponents claim that “inclusive education ... is the foundation of an education system of good quality” (Antoninis et al., 2020, p. 5). They point to evidence that inclusive practices do not just benefit students with special needs, but benefits all learners (Rambla and Langthaler, 2016; UNESCO, 2009). This is clear in two recent reviews. Kefallinou et al. (2020), in reviewing literature between 2015 and 2020 found that “including learners with disabilities in mainstream settings had no negative impact on the outcome of their peers without disabilities” (p. 4). Results from Hehir et al. (2016), following a systematic review of 280 studies, was even more positive, finding “clear and consistent evidence that inclusive educational settings can confer substantial short- and long-term benefits for students with and without disabilities” (p. 1). The presence of inclusion as a necessary component of quality in the overall goal of SDG 4—to “ensure *inclusive* and equitable *quality* education and promote lifelong learning opportunities for all” (emphasis added)—was a significant improvement over the earlier global goals in education (Brissett and Mitter, 2017).

This chapter further examines the promise of inclusive quality education in SDG 4. Specifically, it looks at the construction, communication, and early implementation of the sustainable development goal for education and asks if inclusion is indeed seen as essential for quality or if these concepts are treated independently. The chapter starts by briefly examining the history of the inclusive education movement. This is followed by an exploration of key moments in the creation of SDG 4. As we inspect the operationalization of the sustainable development goal for education—from goals to targets to indicators—we continue to ask which approach to quality is valued and whether the principles of inclusive education are seen in the practices of applying SDG 4. The implications for our results—suggesting that quality has been captured by a narrow, utilitarian perspective—are then laid out before suggesting next steps needed to ensure a more holistic, rights-based view of educational quality.

Evolution of inclusive education

Prior to the 1970s, children with special needs were generally excluded from education (Smith, 2018). The first efforts to provide children with special needs access to regular schools were seen in Italy in 1971 and the United States in 1975 (Slee, 2018). The 1975 Education of All Handicapped Children Act in the United States (replaced in 1990 by the Individuals with Disabilities Education Act) focused on providing appropriate education for handicapped children.

Globally, the 1994 Salamanca Statement represents a key moment where 92 countries acknowledged that students with special needs are able to learn in general education settings (Florian, 2019) and increased attention was placed on integrating such students into national education systems (Johnstone et al., 2020). Just over a decade later the UN Convention on the Rights of Persons with Disabilities (CRPD) was ratified, with article 24 highlighting the removal of barriers that exclude persons with disabilities as a key aspect of inclusive education (Kefallinou et al., 2020).

Since Salamanca, inclusive education has transformed from a narrow focus on the integration of students with special needs into regular school to a broader understanding (Florian, 2019; Smith, 2018) of inclusive education as part of a “social movement against educational exclusion” (Slee and Allan, 2001, p. 177). The shift in focus has been identified by Ainscow (2007) as the “inclusive turn” where mere integration is no longer sufficient and transforming schools and systems to meet the needs of all learners is now considered necessary. This is perhaps best captured in general comment 4 of the CRPD. Released in 2016 it states that “... inclusion involves a process of systematic reform embodying changes and modifications in content, teaching methods, approaches, structures and strategies in education to overcome barriers within a vision serving to provide all students of the relevant age range with an equitable and participatory learning experience and environment that best corresponds to their requirements and preferences” (p. 4). The transition to the broad view of inclusive education was, in part, accomplished by the increased merging of language of inclusive education and special education (Slee, 2018) and is the result of the human rights movement and corresponding treaties that made it clear that inclusive education is morally and legally just (Kefallinou et al., 2020).

At present, the broad understanding is the dominant approach seen in global discourse. Within this understanding some overarching principles are clear: (1) all children have the right to a good quality education, (2) education needs to be flexibly shaped to meet the needs of all learners, (3) education should meet the learner where they are currently at and move them forward, (4) all teachers should be trained in inclusive pedagogy, and (5) inclusive practices are a central part of quality education and benefit all learners. Yet, while these principles represent an ideal, how inclusive education is specifically defined and implemented remains in question (Florian, 2019). As Forlin et al. (2013) note, the abstract nature of inclusive education “may contribute to some misconceptions and confused practice” (p. 6).

In an effort to better conceptualize inclusion, Johnstone et al. (2020) attempted to map the discourse of inclusion by drawing upon legal and political continua. For example, inclusion can either be framed as universal (narratives of “all”) or plural (as advocacy for particular groups’ rights to education). Similarly, inclusion can be characterized as social (an act of welcoming into existing systems) as well as relational (that which requires resource redistribution to diminish historical exclusion and marginalization). Based on these continua, quality inclusive education could be theoretically framed as:

- Universal-social: Inclusion for all by welcoming every child, adolescent, and adult into existing education systems (i.e., access).
- Universal-relational: Inclusion for all by welcoming every child, adolescent, and adult into existing education systems and distributing necessary additional resources for those in need (i.e., accommodations, scholarship support).
- Plural-social: Adjusting systems to be more responsive to diversity within existing systems without prioritizing needs of one group over another (i.e., multiculturalism, cultural responsiveness).
- Plural-relational: Adjusting systems to be more responsive to diversity within existing systems and diverting resources toward historically marginalized or excluded groups.

While at first glance these continua appear to replicate some of the narrow and broad views outlined above (with plural representing a narrow view and universal representing a broad view), this is not the case. Plural-social and plural-relational conceptualizations are not suggesting that all children, adolescents, and adults do not have the right to access education, but instead that when they access education their individual talents and needs are recognized. In short, a universal conception fails to recognize differences within the classroom while a plural approach embraces diversity. The other key difference between universal and plural is whether the existing system is maintained or adjusted. While the former aligns more with earlier attempts at integration the latter better represents the transformative desire of the humanistic perspective. Finally, distinctions between social and relational conceptions depend on whether needs of particular groups are recognized and addressed. Some may question whether the needs of historically marginalized groups can be met in systems that perpetuate their exclusion. Others may argue that the desire to have all learners in education may be at odds with tailored approaches to equity. This, however, need not be the case. As Schuelka et al. (2019) suggest, representational acknowledgment is one part of a broader vision of inclusive education that has “both a systems focus and a focus on specific marginalized and/or disadvantaged groups. They do not need to be mutually exclusive” (p. xxxvi).

The creation of SDG 4

The 1990 World Conference on Education for All in Jomtien, Thailand laid out the first of a series of global goals for education. The corresponding Education for All (EFA) movement highlighted the right to education for all, regardless of background and

contributed to broadening the view of inclusive education (Florian, 2019). The sustainable development goal for education, agreed upon in late 2015, is the latest iteration of the global goals. Unlike prior efforts, the creation of the SDGs was hailed as an open, collaborative process that included representatives from civil society, government, and the scientific community (Johnstone et al., 2020). Merging efforts from prior global development architectures (Millennium Development Goals and Rio Summit on the Environment and Development), the 4-year consultative process of the SDGs included meetings in over 60 countries and a My World survey, capturing public priorities (Smith, 2019). As Jordan Naidoo, the UNESCO Director of the Division for SDG Education Support and Coordination wrote in 2016, the lead up to the SDGs was "... arguably the most inclusive process of consultation in the history of the United Nations."

Over this 4-year consultation, there were at least eight key outputs that shaped the creation of SDG 4. Appointed by the UN Secretary General in 2012, a High Level Panel of Eminent Persons, chaired by leaders from Indonesia, Liberia, and the UK produced a report in May 2013 that largely reflected the goals of the UK's development agency (DFID). In April 2014, UNESCO's EFA Steering Committee released a position paper following their regional consultations with governments and NGOs. In part to challenge the EFA Steering Committee paper, the Sustainable Development Solutions Network (SDSN) produced a lobbying document in May 2014. Written by the leadership council of the network, high level political figures, leading academics, and representatives from business were included, but no leading figures from the EFA movement. In contrast, the Muscat Agreement largely reflected the position of the EFA Steering Committee. Led by the Director General of UNESCO, the agreement was the result of the May 2014 global Education for All Meeting which included national ministry officials, multi- and bi-lateral organizations, NGOs, and private sector leaders (Unterhalter, 2019).

In July 2014, the Open Working Group (OWG) published their progress report, suggesting what is now the 17 SDGs. Created to provide an alternate vision to the high level panel, the OWG consisted of 30 seats, each representing 3 countries, and heard input from multiple stakeholders, including teacher unions, NGOs, and indigenous peoples groups. The UN Synthesis Report, published in December 2014, attempted to bring the various tracks together and ultimately endorsed the view promoted by civil society and the EFA Steering Committee. The last two documents included the Incheon Declaration, which largely laid out SDG 4 as we know it today, and the Education 2030 Framework for Action, which provided an initial framework for how to implement the aims laid out in the Incheon Declaration. The declaration was the outcome document of the May 2015 World Education Forum that included representatives from government, civil society, and seven UN organizations. The framework was agreed upon by 184 countries at UNESCO headquarters in November 2015 (Unterhalter, 2019).

With many stakeholders present, clear, rival visions can be seen in the desired aims of education. Unterhalter (2019), in her review of these eight outputs, highlighted that those without a strong civil society or EFA presence tended to adopt a utilitarian view of quality education, emphasizing education's role in economic development. On the contrary, meetings with a diverse delegation that better represented humanity tended to adopt a holistic, humanistic view of quality education. Table 1 expands on the work of Unterhalter (see her work in the Definition of Quality row in Table 1) by examining how inclusive education is represented in these key documents.

The first observation from Table 1 is that inclusive education is rarely mentioned specifically. The only attempt to define inclusive education beyond general statements on inclusion or "all" is found in the Education 2030 Framework for Action. At the surface level it appears that inclusive education is considered part of quality in six of the eight documents. This is often done by repeating what is now (in slightly modified form) the overarching goal of SDG 4, which was first proposed in the EFA Steering Committee report, "ensure equitable and inclusive quality education and lifelong learning for all by 2030."

The two reports with a utilitarian perspective are clear by their general focus and noticeable absence of inclusion as part of quality education. The High Level Panel focuses on the learning crisis and rates of return on education and believes quality entails meeting "minimum standards" through measurable learning outcomes. In a similar vein, the SDSN report focuses on skill attainment and education as a means to an end, with quality associated with a (albeit wider) range of learning outcomes.

Mentions of "all" are present throughout the reports. What has been presented is only an example of what is included, with multiple more "all" statement found in the documents, notably the Incheon Declaration and Education 2030 Framework for Action. In reviewing "all" statements they tend to support access to or outcomes of education, rarely mentioning the process of education applied within the system. Once again, the Incheon Declaration and the Education 2030 Framework for Action are noticeable exceptions. Both recognize that changes are necessary to meet the needs of learners. The Incheon Declaration states that "inclusion and equity in and through education is the cornerstone of a transformative education agenda" and the framework points to changes in public policy and teacher training as necessary steps to reshape the system.

Operationalization of SDG 4

Drawing directly from the Incheon Declaration and as outlined in the Education 2030 Framework for Action, at first glance it appears that SDG 4 has embraced a humanistic view of education quality. Rightfully celebrated as an important addition to the overarching goal of SDG 4, the voice of inclusive education, and other advocates of the humanistic approach, diminished as SDG 4 was operationalized.

Under each overall goal in the SDGs, there are corresponding targets and indicators. The initial plan was to have four types of indicators that would complement each other and allow for national flexibility to recognize the different local realities (Brissett and Mitter, 2017). Global indicators would include a small number of emphasized indicators that capture the essence of the target and are comparable across countries. Thematic indicators are additional, cross-national comparable indicators that fit under each target.

Table 1 Inclusion and quality as understood in the key outputs preceding SDG 4.

	<i>High level panel (2013)</i>	<i>EFA steering committee (2014)</i>	<i>SDSN report (2014)</i>	<i>Muscat agreement (2014)</i>	<i>OWG SDG proposal (2014)</i>	<i>UN SDG synthesis report (2014)</i>	<i>Incheon declaration (2015)</i>	<i>Education 2030 framework for action (2015)</i>
General focus	Focus on learning crisis, minimal mentions of barriers, highlights rate of return on education	Human-rights based, focused on gender, some mention past learning outcomes to include global citizenship and sustainable development	Focus on skill attainment and livelihoods—includes social and cognitive skills, problem solving and critical thinking	Holistic and right-based approach. Education is fundamental right and essential for peace, economic growth, equality, & global citizenship	Focus on poverty eradication and sustainable, inclusive economic growth. More on SDG 4.7—gender equality, peace/non-violence, citizenship	End poverty while meeting human needs & respecting human rights. Promote sustainable & inclusive economic growth, development, & environmental protection	Renewed education agenda that is holistic, ambitious and aspirational, leaving no one behind. Ed as public good & fundamental right	Comprehensive, holistic, ambitious, aspirational & universal. “Rights-based & inspired by humanistic vision”
Definition of quality (Unterhalter, 2019)	Meeting “minimum standards” of reading, writing and counting at primary level; meeting measureable learning outcomes at junior secondary	Developing capacity to learn, meeting global benchmarks of learning outcomes, fully participation in society, access decent work; express values (global citizenship & sustainable development)	Wide range learning outcomes, reduce dropout in higher education, develop knowledge economics	No definition for early years, primary and secondary. For youth and adults’ increasing participation in society, enhancing peace, sustainable development and global citizenship	No definition of quality but associated with free & equitable schooling, affordable vocational education and relevant skills	No definition of quality but discussion links education with dignity, sustainability, gender equality, and enhancing know how	Very full definition of quality looking at practice, includes rights and gender equality	Links quality with the humanistic inclusive vision of SDGs and with learning processes as well as outcomes
Inclusive education mentioned	Yes, once in appendix	No	No	No	No	No	No	Yes, multiple. “Inclusive education for all should be ensured by designing and implementing transformative public policies to respond to learners’ diversity and needs, and to address the multiple forms of discrimination and of situations.” Teacher training, meet most marginalized, necessary to transform education system

Inclusive education as part of quality	No	“Ensure equitable and inclusive quality education and lifelong learning for all by 2030”	No	“Ensure equitable and inclusive quality education and lifelong learning for all by 2030”	“Ensure equitable and inclusive quality education and lifelong learning for all by 2030” & in target 4a on inclusive environments	“Ensure equitable and inclusive quality education and lifelong learning for all by 2030” & in target 4a on inclusive environments	“Ensure equitable and inclusive quality education and lifelong learning for all by 2030” “inclusion and equality in and through education is the cornerstone of a transformative education agenda”	“Ensure quality and inclusion and equality in and through education and address all forms of exclusion and marginalization, disparity, vulnerability and inequality in education access, participation, retention and completion and in learning outcomes.”
Mentions of “all”	“We need to ensure all children, regardless of circumstance, are able to enroll and complete a full course of primary and lower secondary education and, in most cases, meet minimum learning standards.”	“Equitable access to education for children, youth and adults should be provided for all and at all levels of education”; “All children complete free and compulsory quality basic education of at least 10 years and achieve relevant learning outcomes”; All learners acquire knowledge, skills, values and attitudes for global citizenship and sustainable development.”	“All girls and boys complete affordable and high-quality early childhood development programs, and primary and secondary education to prepare them for the challenges of modern life and decent livelihoods. All youth and adults have access to continuous lifelong learning to acquire functional literacy, numeracy, and skills to earn a living through decent employment or self-employment.”	All girls and boys complete free and compulsory quality basic education of at least 9 years and achieve relevant learning outcomes [target 2]”; “All learners acquire knowledge, skills, values and attitudes to establish sustainable and peaceful societies [target 5]	“All girls and boys complete free equitable and quality primary and secondary education leading to relevant and effective learning outcomes” [T 1]; “All girls and boys have access to quality early childhood development, care and pre-primary education” [T 2]; “All learners acquire the knowledge and skills needed to promote sustainable development” [T 7]	“All children and adolescents have a right to education and must have a safe environment in which learn.”	“No education target should be considered met unless met by all.” “All children have access quality early childhood development, care and education.” “We commit to promoting quality lifelong learning opportunities for all, in all settings and at all levels of education.”	Throughout document and mirroring current SDG4 targets
Identified target subgroups	None	Gender, most marginalized	Girls, children in hard to reach areas, income groups, disability	Girls, most marginalized	Gender, disability, indigenous people, children in vulnerable situations	Gender, age, race, ethnicity, migration status, geographic location, disability, unemployed	Most disadvantaged, those with disabilities	Gender, most vulnerable, person w/disability, indigenous peoples. Large list of demographics

Sources: HLPPEP (2013), UNESCO (2014a), SDSN (2014), UNESCO (2014b), UNESCO (2015), OWG (2014), UN (2014).

Regional indicators recognize that commonalities in educational aim and system are more likely at the regional level and as thus countries within similar regions were encouraged to establish indicators related to SDG 4 targets. Finally, national indicators provide an additional level of autonomy, with countries encouraged to adapt the SDGs to their local needs or suggest additional complimentary indicators that represent their national interest. In reality only global and thematic indicators are used with any regularity. For example, in [Smith's \(2021\)](#) analysis of 32 voluntary national reviews—submitted by countries to report their progress on the SDGs—he found that all global indicators were reported at some point across the 32 countries, but 45% of thematic indicators were not reported by a single country, while regional and national indicators were absent or nearly absent. The prominence of global indicators may be due to their obligatory reporting status and their appearance as often the only indicators included in global reports and on websites, both of which encourage states to target resources to global indicators ([Smith, 2019](#)).

It is disturbing, then, that the process of moving from SDG 4 targets to the associated global indicators have resulted in a narrowed down and simplified notion of education that fails to capture the essence of the overarching SDG 4 goal or the related target. While this is most apparent in SDG 4.1, where the global indicator fails to capture completion, equitable, relevant, effective, or free descriptors of education and quality is limited to test scores ([King, 2017](#)), the movement from relatively more humanistic targets to utilitarian global indicators is found across SDG 4 ([Unterhalter, 2019](#)).

The utilitarian capture of quality

A utilitarian perspective has captured the understanding of quality in the global indicators, the primary way progress on SDG 4 is communicated. Various factors help explain this capture. First, might civil society be fatigued? Civil society and EFA partners lobbied substantially in the build up to SDG 4. This is noticeable in the more humanistic approach taken in six out of the eight documents examined in [Table 1](#) and the structural changes demanded in the Muscat Agreement and the Incheon Declaration, such as the need for “transformative public policies to respond to learners’ diversity and needs, and to address the multiple forms of discrimination and of situations, including emergencies, which impede the fulfillment of the right to education” in the latter. In directing significant resources and energy to meet the moment in Incheon, is there substantial capacity left to continue the fight.

Second, the operationalization process itself tends to rely on a specific, technical skill set. Led by the UNESCO Institute of Statistics (UIS) and working with 27 national statistical offices, the indicator creation process included a starkly different group of stakeholders, relative to pre-SDG consultations. Additionally, the need for quantifiable, comparable indicators limit the potential metrics for global indicators.

Third, there is substantial, disproportionate investment in metrics that represent a simplified, utilitarian view of quality as learning outcomes. For donors, they are seen as easily communicated ways to measure impact ([Winthrop and Anderson Simmons, 2013](#)) and are part of a larger shift of donor support toward learning assessments and outcomes ([Benavot and Smith, 2020](#)), including earmarking contributions to UIS ([Wulff, 2020](#)). The Global Alliance to Monitor Learning (GAML) is a self-funded group under UIS, tasked with creating indicators for SDG 4.1, 4.2, 4.4, 4.6, and 4.7. The largest supporters of GAML are the UK’s DFID, Australia’s Department of Foreign Affairs and Trade (DFAT) and the Hewlett Foundation. Although charged with the creation of indicators for five SDG 4 targets, SDG 4.1 is prioritized. This is clear in GAML’s Results Framework, where two out of every three targets focus on 4.1, and in the advancement in the global indicator for SDG 4.1. Over the first 5 years of the SDGs, the global indicator from 4.1 developed from having no established methodology (tier 3) to being conceptually clear with an established methodology (tier 1). Progress of the global indicator for the more humanistic SDG 4.7 has stalled at tier 3.

Finally, no robust inclusive alternate definition of quality was present in the outputs leading to SDG 4. In reviewing [Table 1](#) there is no definition and little recognition of “inclusive education” as a valued approach to education that benefits all learners ([Anderson and Boyle, 2015](#)). Notions of inclusion, when incorporated, focus on integration into the system; learners accessing and adapting to the system instead of a system adapting to learners. This is illustrated in overreliance on top-level “all” statements that focus on access. [Johnstone et al. \(2020\)](#) identified a similar trend when they recognized that across the SDGs the term “inclusion” is included 40 times, but it is undefined and limited to notions of “all.”

Competing agendas and the re-fragmentation of inclusive education

The reductionist process that takes place in operationalizing SDG 4 echoes the work of some who have researched what happens to competing principles when they are attempted to be put into practice. While at the abstract principle level, a humanistic ideal of inclusion as quality is easy to envision, as multiple principles, or views on quality education, can be held “without compromise and without constructing an a priori hierarchy of principals” ([Clark et al., 1999](#), p. 173). In practice, however, principles are thrust into competition and thus will “always be partial and compromised” ([Clark et al., 1999](#), p. 173). The result is inclusion now having to compete with a version of quality, narrowly defined as learning outcomes.

As [Slee \(2013\)](#) pointed out, more humanistic values of education, such as inclusion and equity, are often “inaudible when located amid more strident educational discourses” (p. 6), such as testing regimes. As the views of quality are derived from different underlying ideological perspectives, it is challenging, if not impossible “for any educational policy initiative to equally serve such conflicting value-laden discourses” ([Brissett and Mitter, 2017](#), p. 182). In the end there is a real fear the inclusive education agenda will be overshadowed and drowned out.

Ultimately, however, it is not just that utilitarian definitions of quality take attention away from the inclusive education position, it is that the definition of quality held up by this perspective actively undermines efforts toward inclusive education (Allan, 2012). The capture of quality as testing is detrimental to the ideal of inclusion. The reliance on learning assessments creates structural barriers that limit participation in groups less likely to sit for cross-national assessments (Johnstone et al., 2020). The ranking, blaming, and shaming of individuals associated with test scores can increase the pressure felt by individual learners (Boeren, 2019), lead to labeling learners, and reinforcing a deficit view (Ainscow, 2020).

Positioning inclusion as separate from quality also promotes a reversal to an earlier narrow version of inclusive education. Instead of transforming education to meet the needs of all learners, in achievement-focused educational environments, access to the opportunity to demonstrate achievement is considered a form of inclusion. For example, Almond et al. (2002) argued nearly two decades ago that students with disabilities should participate in learning assessments for accountability purposes—but that participation must include appropriate accommodations. The authors' logic was that participation would highlight systemic gaps in instruction for this population. Likewise, LeRoy et al. (2019) and Smith and Douglas (2014) advocated for greater inclusion of children with disabilities in PISA testing, but acknowledged the sampling and accommodations limitations of such inclusion. In each of these cases, however, the participation of children with disabilities in learning assessments was argued to be a mechanism for revealing systemic failures.

The reduction of inclusive education to the disaggregation of data of a few identified groups is not without its benefits, although the fragmentation of the inclusive education coalition can lead to in-fighting for available resources. Johnstone et al. (2020) argued that the SDGs provided *some* indication or acknowledgment of historic exclusion and discrimination. By naming excluded populations, the SDGs took what Mégrét (2008) would call a plural rights approach. Further, in the call for addressing the specific needs of girls, low-income individuals, rural individuals, and persons with disabilities, the SDGs hinted at a “relational” approach to inclusion that acknowledged resource inequalities as a source of exclusion (see Gupta and Vegelin, 2016). According to Abualgehaib et al. (2019), the collection of good quality data is necessary for identifying cases of exclusion and “crucial to achieving inclusion” (p. 3091). Additionally, the disaggregation of data in SDG 4 highlights the intent of the global community to move past the “politics of silence” present in past global goals and recognize the needs of students with disabilities (Anastasiou and Keller, 2014). Useful data for such a task might include educational assessment data (LeRoy et al., 2019) or disability identification data (Abualgehaib et al., 2019). The Washington Group Question Sets are an example of the latter, to date at least one of their sets have been used by over 70 countries in their censuses or other national data collection instruments, such as those used in Education Management Information Systems (EMIS).

Conclusion

If continuing on its present trajectory, SDG 4 will fail to reach its aspirational goal of inclusive quality education. A humanistic view of quality as a transformed education system, meeting the needs of all learners is far off. Quality as testing reproduces a status quo that disadvantages some students and labels others as “less than.” While the overall goal of SDG 4 hints at a transformative agenda, the SDGs still fall short of the drastic redistribution recommendations either within governments or between high- and low-income countries. Further, if “quality” is defined as outcomes, the SDGs provide expectations for equality in outcomes among some groups (e.g., girls and boys in Target 4.1), but outcome targets are overlooked for children and youth experiencing intersectional exclusion, as persons with disabilities, indigenous peoples, and children in vulnerable situations are named, but Target 4.5 only requires states parties to “ensure equal access” of these groups and does not address quality.

The SDGs stand at a crossroads between universalist narratives and population-focused constructions of inclusion. Universal narratives of “all” fail to acknowledge that exclusion is a choice in policy and practice that must be addressed through affirmative action approaches. At the same time, plural definitions of inclusion are often still assumed to be a form of education focused on children with disabilities or special needs that are designed to address a perceived deficit or dysfunction (Slee, 2011). This latter approach was used to both provide a pathway to children's rights to education and as therapeutic and “normalizing” for children with presumed pathologies. In SDG 4, such inclusion is limited to the mere presence of particular children without any significant expectation of system change or outcomes for such children. This is insufficient and fails to account for the basic principles of the broad view of inclusive education.

Increased efforts are needed to challenge the definition of quality, positioning inclusion at the center. As Brissett and Mitter (2017) acknowledge, the “aims of education must be expanded such that ‘quality’ education is no longer strictly associated with standardization, efficiency, and employment, but instead viewed as a fundamental human right and a catalyst for social change” (p. 201). The SDGs do not do enough to shift the normative center of education systems that see difference and define it as “additional support needed.” Continuing with the status quo means that, even when broad goals of inclusive education are held by governments in principle, in practice exclusion rituals continue (Florian, 2019).

Part of transforming toward inclusive quality education is recognizing that every learner is unique. When difference is considered ordinary, education values diversity and inclusive approaches are commonplace and seen as the norm. As we all have multiple, intersecting identities (Antoninis et al., 2020), we must see individual differences “not as problems to be fixed but as opportunities to enrich learning” (Ainscow, 2020, p. 6).

Mechanical solutions that work around the periphery of education systems are not enough. Real systems change starts with a broad view of inclusive education and a recognition that systems need to change for learners, not learners changing to fit in

systems. The recommendations of the 2020 Global Education Monitoring (GEM) Report are a good first step. Their first recommendation urges countries to adopt a wide understanding of inclusive education, including everyone regardless of identity, background or ability. The adjustments necessary in such a system must be part of a coherent education plan at all levels. The 2020 GEM Report can be a valuable tool for inclusive education advocates as they work to reposition inclusion at the heart of quality. A renewed, collective effort from the inclusive education community is needed. Working together we need to push beyond access for all and discuss what inclusive education looks like in practice. Emphasis and additional evidence demonstrating the relationship between inclusive education and pedagogy and quality education for all is required to clearly communicate that quality is not possible without authentic inclusion.

References

- Abualghaib, O., Groce, N., Simeu, N., Carew, M.T., Mont, D., 2019. Making visible the invisible: why disability-disaggregated data is vital to "leave no-one behind". *Sustainability* 11 (11), 3091.
- Ainscow, M., 2007. Towards a more inclusive education system. Where next for special schools? Included or excluded? In: Cigman, R. (Ed.), *The Challenge of the Mainstream for Some SEN Children*. Routledge, London, pp. 128–139.
- Ainscow, M., 2020. Inclusion and equity in education: making sense of global challenges. *Prospects* 49 (3), 123–134.
- Allan, J., 2012. The inclusion challenge. In: Barow, T., Ostlund, D. (Eds.), *En pedagogisk utmaning [Education for Everyone! A Pedagogical Challenge]*. Hogskolan Kristianstad, Tryck, pp. 109–120.
- Almond, P.J., Lehr, C., Thurlow, M.L., Quenemoen, R., 2002. Participation in large-scale state assessment and accountability systems. In: Tindal, G., Haladyna, T.M. (Eds.), *Large-Scale Assessment Programs for All Students: Validity, Technical Adequacy, and Implementation*. Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 341–370.
- Anastasiou, D., Keller, C.E., 2014. Cross-national differences in special education coverage: an empirical analysis. *Except. Child.* 80 (3), 353–367.
- Anderson, J., Boyle, C., 2015. Inclusive education in Australia: rhetoric, reality and the road ahead. *Support Learn.* 30 (1), 4–22.
- Antoninis, M., April, D., Barakat, B., Bella, N., D'Addio, A.C., Eck, M., Endrizzi, F., Joshi, P., Kubacka, K., McWilliam, A., Murakami, Y., Smith, W., Stipanovic, L., Vidarte, R., Zekrya, L., 2020. All means all: an introduction to the 2020 global education monitoring report on inclusion. *Prospects* 49 (3), 103–109.
- Benavot, A., Smith, W.C., 2020. Reshaping quality and equity: global learning metrics as a ready-made solution to a manufactured crisis. In: Wulff, A. (Ed.), *Grading Goal Four: Tensions, Threats and Opportunities in the Sustainable Development Goal on Quality Education*. Brill, Leiden, pp. 238–261.
- Boeren, E., 2019. Understanding sustainable development goal (SDG) 4 on "quality education" from a micro, meso and macro perspective. *Int. Rev. Educ.* 65, 277–294.
- Brissett, N., Mitter, R., 2017. For function or transformation? A critical discourse analysis of education under the sustainable development goals. *J. Crit. Educ. Pol. Stud.* 15 (1), 181–204.
- Clark, C., Dyson, A., Millward, A., Robson, S., 1999. Theories of inclusion, theories of schools: deconstructing and reconstructing the "inclusive school". *Br. Educ. Res. J.* 25 (2), 157–176.
- Florian, L., 2019. On the necessary co-existence of special and inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 691–704.
- Forlin, C., Chambers, D., Loreman, T., Deppeler, J., Sharma, U., 2013. *Inclusive Education for Students With Disabilities. A Review of Best Evidence in Relation to Theory and Practice*. Australian Research Alliance for Children & Youth, Canberra.
- Gupta, J., Vegelin, C., 2016. Sustainable development goals and inclusive development. *Int. Environ. Agreements Polit. Law Econ.* 16 (3), 433–448.
- Hehir, T., Grindal, T., Freeman, B., Lamoreau, R., Borquaye, Y., Burke, S., 2016. *A Summary of the Evidence on Inclusive Education*. ABT Associates, Cambridge.
- HLPEP (High Level Panel of Eminent Persons), 2013. *A New Global Partnership. Report of the High Level Panel of Eminent Persons on the Post 2015 Development Agenda*. United Nations, New York.
- Johnstone, C.J., Schuelka, M.J., Swadek, G., 2020. Quality education for all? The promises and limitations of the SDG framework for inclusive education and students with disabilities. In: Wulff, A. (Ed.), *Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education*. Brill, Leiden, pp. 96–115.
- Kefallinou, A., Symeonidou, S., Meijer, C.J.W., 2020. Understanding the value of inclusive education and its implementation: a review of the literature. *Prospects* 49 (3), 135–152.
- King, K., 2017. Lost in translation? The challenges of translating the global education goals and targets into global indicators. *Compare* 47 (6), 801–817.
- LeRoy, B.W., Samuel, P., Deluca, M., Evans, P., 2019. Students with special educational needs within PISA. *Assess. Educ. Princ. Pol. Pract.* 26 (4), 386–396.
- Maclure, R., Sabbah, R., Lavan, D., 2009. Education and development: the perennial contradictions of policy discourse. In: Haslam, P., Schafer, J., Beaudet, P. (Eds.), *Introduction to International Development: Approaches, Actors, and Issues*. Oxford University Press, Oxford, pp. 367–383.
- Mégret, F., 2008. The disabilities convention: human rights of persons with disabilities or disability rights? *Hum. Right Q.* 494–516.
- Naidoo, J., 2016. Education by 2030: Ambitious Targets Require a Community United to Succeed. Deliver 2030. Blog Post Retrieved from: <https://www.webarchive.org.uk/wayback/archive/20180120160720/http://deliver2030.org/?p=7007>.
- OWG, 2014. Report of the Open Working Group of the General Assembly on Sustainable Development Goals (A/68/970). United Nations, New York. July.
- Rambla, X., Langthaler, M., 2016. The SDGs and Inclusive Education for All: From Special Education to Addressing Social Inequalities. Austrian Foundation for Development Research (OFSE). Briefing Paper, No. 14.
- Sayed, Y., Moriarty, K., 2020. SDG 4 and the "education quality turn": prospects, possibilities, and problems. In: Wulff, A. (Ed.), *Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education*. Brill, Leiden, pp. 194–213.
- Schuelka, M., Johnstone, C., Thomas, G., Artiles, A. (Eds.), 2019. *The Sage Handbook on Inclusion and Diversity in Education*. Sage, London.
- SDSN, 2014. *An Action Agenda for Sustainable Development*. Sustainable Development Solutions Network, New York.
- Slee, R., Allan, J., 2001. Excluding the included: a reconsideration of inclusive education. *Int. Stud. Sociol. Educ.* 11 (2), 173–192.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling, and Inclusive Education*. Routledge, London.
- Slee, R., 2013. Meeting some challenges of inclusive education in an age of exclusion. *Asian J. Incl. Educ.* 1 (2), 3–17.
- Slee, R., 2018. Defining the Scope of Inclusive Education. Think Piece Prepared for the 2020 Global Education Monitoring Report. UNESCO, Paris.
- Smith, E., Douglas, G., 2014. Special educational needs, disability and school accountability: an international perspective. *Int. J. Incl. Educ.* 18 (5), 443–458.
- Smith, W.C., Benavot, A., 2021. Quality education and global learning metrics. In: McGowan, T., Unterhalter, E. (Eds.), *Education and International Development*, second ed. Bloomsbury, London, pp. 189–206.
- Smith, W.C., 2018. Quality and inclusion in the SDGs: tension in principle and practice. In: Ydesen, C., Morin, A., Hamre, B. (Eds.), *Testing and Inclusive Schooling*. Routledge, Oxford, pp. 89–104.
- Smith, W.C., 2019. One indicator to rule them all: how SDG 4.1.1 dominates the conversation and what it means for the most marginalized. In: *Annual Review of Comparative and International Education*. Emerald, Bradford.
- Smith, W.C., 2021. An exploration of SDG 4 coverage in voluntary national reviews. In: Zajda, J. (Ed.), *Globalisation, Social Change and Education Reforms: New Paradigms and Ideologies*. Springer Publishing, New York, pp. 961–980.
- UN, 2014. *The Road to Dignity by 2030. Synthesis Report of the Secretary General on the Post 2015 Agenda*. United Nations, New York.

- UNESCO, 2009. Policy Guidelines on Inclusion in Education. UNESCO, Paris.
- UNESCO, 2014a. Joint Proposal of the EFA Steering Committee on Education Post 2015. Education for All Steering Committee, Paris.
- UNESCO, 2014b. "The Muscat Agreement" Final Statement of the Global Education for All Meeting, Muscat, Oman 12–14 May. UNESCO, Paris.
- UNESCO, 2015. Education 2030. Incheon Declaration and Framework for Action. UNESCO, Paris.
- Unterhalter, E., 2019. The many meanings of quality education: politics of targets and indicators in SDG4. *Glob. Pol.* 10 (1), 39–51.
- Winthrop, R., Anderson Simons, K., 2013. Can international large-scale assessments inform a global learning goal? Insights from the learning metrics task force. *Res. Comp. Int. Educ.* 8 (3), 279–293.
- Wulff, A. (Ed.), 2020. Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education. Brill, Leiden, The Netherlands.

Reform and purpose for inclusive values in complex education systems

Matthew J. Schuelka, University of Minnesota, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	28
Education as a complex system	28
Theoretical models for understanding complexity in education systems	29
Inclusion and diversity as an educational value in a complex system	31
The (complex) path forward	32
References	33

Introduction

Any seemingly logical reform in education policy and practice will always run up against the illogic of people. Actors in a social system will make decisions based not upon hard evidence, but more upon their own lived-experiences and the logic of the social system itself; the historical, cultural, and contextual norms that shape the production of institutional formal education. There is a wide range of support in academic literature that has firmly established the policy-to-practice gap. In truth, I would argue that highlighting this gap has become too clichéd to remain interesting, in that the work to establish this pattern is already definitive. Over 25 years ago, [Tyack and Cuban's \(1995\)](#) *Tinkering Toward Utopia* provided a landmark and comprehensive analysis on education policy and practice in the United States. In their analysis, Tyack and Cuban analyzed education policy cycles and broke them down into three phases: policy talk, policy action, and implementation. In each phase of a policy cycle, there is a constant negotiation between competing interests, values, and realities among actors and, in the end, education reform becomes a chimera of disparate interpretations. This is a similar phenomenon also found at the international level, in what [Steiner-Khamsi and Stolpe \(2006\)](#) label the “global speak” of education reform ideas that travel globally, but become something quite different in terms of policy action and implementation. As [Malen \(2006, 89\)](#) succinctly states, “policies unfold in nested contexts.”

In this chapter, I argue that reform in the inclusivity and diversity of education—so-called “inclusive and special education”—has too often made assumptions in the policy cycle of talk, action, and implementation. Arguing for greater reform and inclusion in education globally has never quite transcended the human rights discourse of “because it’s the right thing to do.” Inherent and fundamental human rights matter, of course, but are not necessarily effective in comprehensive education system reform. This is because education systems are complex systems. Aligning positive social values such as “inclusion” across an entire complex social system takes extraordinary coordination that is almost never achieved. Similarly, the advocacy for greater inclusivity and diversity in education also runs up against the problem that the complex system of education has been historically and culturally rooted in exclusion and elitism. These are, obviously, competing values to inclusion and diversity.

In this chapter, I will first explain what I mean when I assert that education is a complex system, particularly in the context of inclusion and diversity in education. Next, I will provide two recently developed theoretical models for further understanding and analyzing complex education systems. I will then provide a short case study and summary of my work in the Himalayan country of Bhutan that will illustrate the difficulty of aligning inclusive education values across a complex system. At the conclusion of the article, I will provide paths forward to better understand and recognize complexity in education systems.

Education as a complex system

By arguing that education is a complex social system, I borrow many ideas from complexity theory and chaos theory from the physical sciences as to how elements in a complex and chaotic system interact, the consideration of closed and open systems, and in the alignment or misalignment of systemic attributes. While there is not always a direct correlation between the oft-predictability of physical phenomenon and the oft-unpredictability of social phenomenon, there is an increasing awareness between the two poles of science that the inorganic world is much more dynamic and complex; and the organic world does not fit neatly into formulas and provable hypotheses. For an accessible definition of a complex system, I much prefer to cite [Amaral and Ottino \(2004, 148\)](#):

A complex system is a system with a large number of elements, building blocks or agents, capable of interacting with each other and with their environment. The interaction between elements may occur only with immediate neighbors or with distant ones; the agents can be all identical or different; they may move in space or occupy fixed positions, and can be in one of two states or of multiple states.

While Amaral and Ottino are describing complex systems in nature, I believe that this is an approachable means in which to understand the dynamics of an education system. First, any education system will contain a large number of elements—here, a plethora of education agents and actors such as children, teachers, paraprofessionals and other education support staff, parents, school leadership, school administration, district/local education authority staff, policy makers from local to international, and the numerous community persons and groups that enter a school every day, just to name a few. These actors all interact with the education system bringing their own understandings, world views, contexts, histories, and so on. To quote a previously-written article on the subject: “Complex systems are open and nested, meaning that there is little to no boundary between elements and surroundings, between inputs and outputs, and elements within the system are complex systems within complex systems ever unfolding like a fractal” (Schuelka and Engsig, 2020, 5).

The assertion that education is a complex system is not necessarily revolutionary, but is new in how it understands the dynamics of element-actors within an open, boundary-less, and self-reflexive space of constant and dynamic interaction between elements and surroundings. There certainly is an element of Practice Theory (Ortner, 2006) to this way of thinking, as well as other sociocultural means of policy analysis (Shore and Wright, 1997; Sutton and Levinson, 2001). Complexity—sometimes under a different term or phrase—provides the underlying assumption in many education studies and theoretical arguments (e.g., Davis and Sumara, 2006; Opfer and Pedder, 2011; Walton et al., 2020). In the section below, I will further elaborate upon education as a complex system, while also providing two theoretical models for understanding and analyzing the complexity within an education system.

Theoretical models for understanding complexity in education systems

Within a complex system, there is not necessarily an absolute beginning and ending to analysis. What we observe when we attempt to understand and analyze a particular dynamic or phenomenon within a complex system is only a brief snapshot in time. However, in order to understand even this brief snap-shot we must also recognize that it contains multi-dimensionality and endless subjective interpretation. That being said, in recognizing education policy and practice, it falls roughly upon the same phases as asserted by Tyack and Cuban (1995): policy talk, policy action, and policy implementation.

In terms of policy talk, for the purposes of this argument I mean the discourses across all elements of society, often influenced by a diversity of world-views, histories, contexts, cultures, and values. Policy talk is self-reflexive and complex in that it both shapes, and is shaped by, societal trends and discourses. To use the Tyack and Cuban (1995) definition, I also mean that policy talk is how we identify and frame societal problems and policy solutions. There is just as much importance to the problems that are “silent” as to the solutions that are ideological hammers “in search of a nail.”

For policy action, in my interpretation this involves the decision-making process of those that enact policy at all levels of implementation. This can mean at an international-national level, such as policy transfer and borrowing between countries (Phillips and Ochs, 2003; Rappleye, 2006). Policy action can also mean the decision-making and negotiation that occurs as an enactment of policy, in which “policy goals, practices, and outcomes become negotiable, context-informed, and dependent on the meanings that policy actors make in their own situated context” (Massouti, 2018, 185).

Implementation is distinct from policy action in that implementation involves the practice and production of education policy down to the micro-classroom level. In terms of understanding implementation, Honig (2006, 14) argues, “Whereas past implementation research generally revealed *that* policy, people, and places affected implementation, contemporary implementation research specifically aims to uncover their various dimensions and *how and why interaction among these dimensions* shape implementation in various ways.” The actual implementation of policy into practice is perhaps the most complex part of the complex system, if you will, in that it involves all elements in a boundless system. Put another way: “... we understand policy as a complex, ongoing social practice of normative cultural production constituted by diverse actors across diverse contexts” (Levinson et al., 2009, 770).

In a previous article, I and my co-author Schuelka and Engsig (2022) provide a theoretical model for the purpose of understanding and analyzing the multi-dimensionality of a complex education system. We term this as Complex Educational Systems Analysis (CESA) and it is visually represented by the Complex Educational Systems Analysis Cube (CESA³). The CESA³ can be seen in Fig. 1 below, and I will explain each dimension on the cube in turn. As in the original article, I will note here that the CESA³ is not a tool for quantitative measurement, as there is no actual volume within the cube itself. In other words, there is not a place where a datum point could be located between dimensional axes, as by definition a complex system is open, dynamic, and boundless.

On the multi-level dimension axis, there is a spectrum of focus from the micro-level to the macro-level. This means that any policy talk, action, or implementation occurs at any or all levels—from international discourses and initiatives such as from the United Nations agencies, down to the everyday interactions between the teacher and student in the classroom. The focus and understanding of multiple levels is influenced both from Bronfenbrenner’s (1979) ecological system theory, as well as comparative case study methodology (Bartlett and Vavrus, 2017; Schuelka, 2018a).

While policy can occur at and between different analytical levels, it can also be enacted across a wide variety of communities. There are essentially an infinite number of ways that we can organize community groupings and pairings for the purpose of analyzing an education phenomenon. Here in the CESA³, we propose six key communities that have analytical power. Policy communities are both formal and informal, explicit and implicit, and can be found along the levels of the previous axis. Formal and professionally-led teaching and learning communities are those structured, institutional arrangements between adult teachers and child learners that we often associate with education. Adult-child communities are less formal and occur outside of the formal

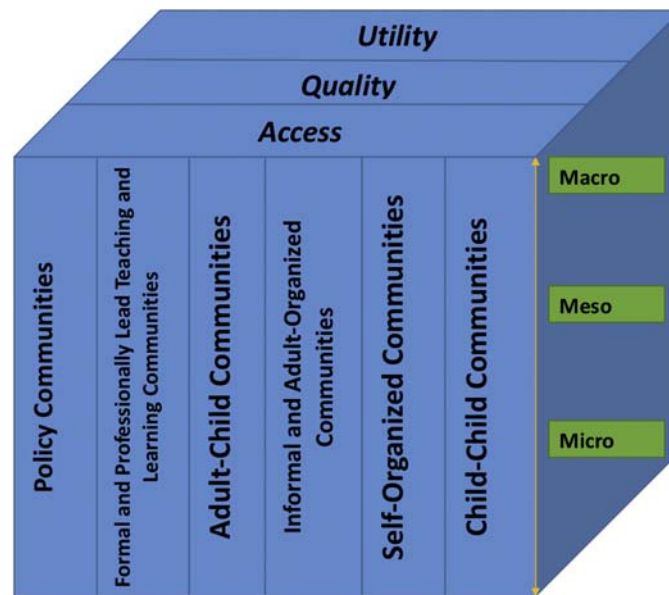


Fig. 1 The complex educational systems analysis cube (Schuelka and Engsig, 2022).

learning in the classroom, but still within the “place” of school such as the playground or the hallway. Informal and adult organized communities are still directed by adults, but occur both in and out of the “place” of school. For example, this could be a sports practice or an after-school club. Self-organized communities are child-led and do not involve adults. For example, this could be self-organized play during recess or interactions on social media that project specific identities and norms. Children may organize themselves around a shared appreciation for a particular online game, or music group, or entertainment universe. Child-child communities are interpersonal dynamics between two or a small group of students.

The third dimensional axis has to do with the attributes of education: access, quality, and utility. These attributes are what we argue to be the three main values of education in relation to the attributes and goals of society. The first attribute, access, is the question of “who goes to school and to whom does the school serve?” The second attribute, quality, is the question of “what do they do there?” and the third attribute, utility, is the question of “what is the purpose of going to school and what are the outcomes of going to school?” In responses to these three questions, complex education systems provide both explicit and implicit values. For example, there has always been a tension in public education between the notion of it serving the individual interests of the child (social mobility and achievement) and the collective interests of society (democratic equality and social efficiency) (Labaree, 1997). How we, as a complex system, believe in the purpose and utility of education ultimately affects who is embraced and served by the school, and the quality of the education provision.

In order to further analyze education values in applicable practice across all the dimensions of the CESA³, I and Kezang Sherab (2022) have provided a second theoretical framework that can be used to specifically evaluate various elements within a complex educational system. The Educational Values Evaluation and Design (EVED) Framework fits inside of the CESA³ in that any societal and educational value contains the multi-dimensionality of levels, communities, and attributes. The visual representation of the EVED framework can be seen in Fig. 2.

The EVED framework is an Ishikawa diagram that delineates various elements that form a particular, or inter-related, set of educational values. The use of this framework can specifically highlight various elements that either align or misalign and project certain educational values. The EVED framework can be used both to design a reform or value-change in an education system, or to evaluate the effectiveness of a reform or value-change after the fact. The nine elements identified—children, teachers, materials, curriculum, pedagogy, policy and governance, economy, socio-cultural factors, family and community—are a comprehensive, but not exhaustive, take on the elements that inform how a societal or educational value is produced and enacted through schooling. These are also broad categories which can be further dissected. For example, materials could include the school building, campus, books, experimental equipment, classroom furniture, technology, or even the resources available to the family and community outside of the classroom. None of the elements are mutually exclusive and, like the CESA³ and complexity in education in general, there is interdependence and dynamism between all elements. In terms of educational values in Fig. 2, these are just a few examples but certainly not an exhaustive list.

In the next section of this chapter, I will look specifically at inclusion and diversity in education as a value in a complex system, applying the CESA³ and EVED framework on the case of Bhutan. The case of Bhutan in the context of this chapter is a summation of my longitudinal work there since 2012 with the support of a wide variety of projects funded by the University of Minnesota, British Academy, UK Economic and Social Research Council, UKAid, the Toyota Foundation, and the European Union.

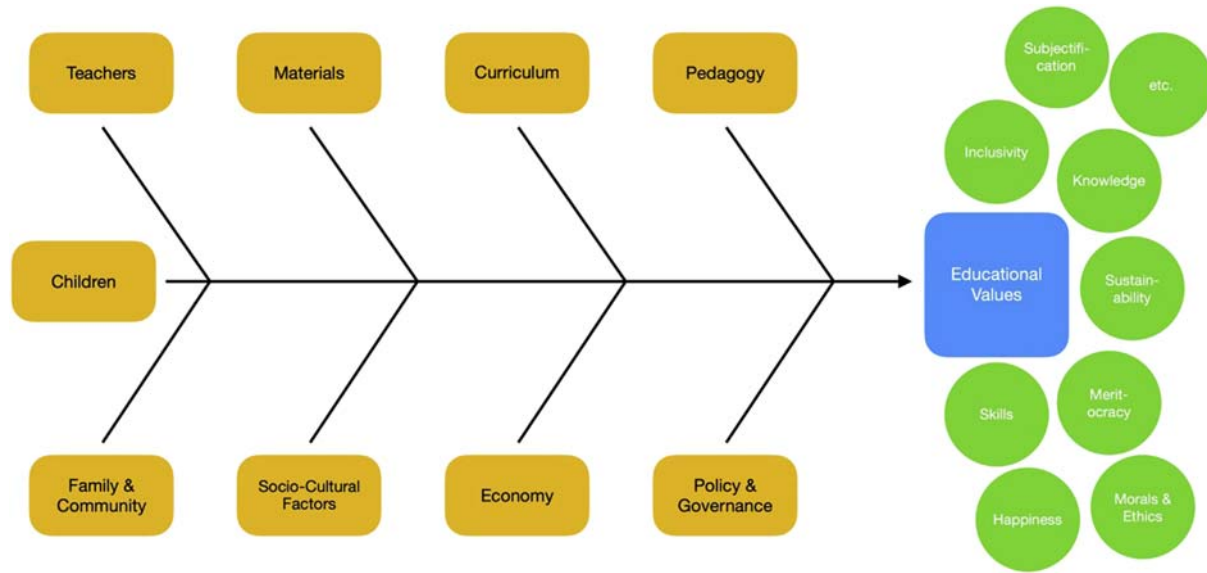


Fig. 2 The educational values evaluation and design framework (Schuelka and Kezang Sherab, 2022).

Inclusion and diversity as an educational value in a complex system

Having laid out the theoretical and analytical frameworks for education in a complex system, I now turn my focus to the particular case of inclusion and diversity as an educational value. As I argued in the introduction to this article, inclusion and diversity in education is too often framed as a human rights issue within the argument of “because it’s the right thing to do” along a narrow set of criteria and target population, without working toward reform of the entire complex education system. Since the Salamanca Statement in 1994 supporting inclusive education as a human right, there has been much effort in assuming this to be the case and then all of the academic and international development energy has been put into identifying the barriers to realizing this right. In my view, this has been myopic and a failure to recognize that an education system is complex. For example, in many ways the Education for All push in the early 21st century had a counter-productive effect in that the number of children given access to school was greatly increased—which is positive and notable—but there was very little in terms of increasing the quality of education provision, the quality of teachers, the quality of materials, the availability of resources, and the alignment of increased educational opportunity with increased (or not) economic opportunity. By providing more education access for those students that were traditionally marginalized by education systems—for example, students with disabilities—the classrooms and teachers did not subsequently adjust themselves to a more diverse and heterogeneous student population, nor was there adequate and compensatory resource allocation, training, and adjustment to the curriculum and the structure of schooling. A classic example of this is England, which embraced inclusive education early, but has seen a significant increase in specialized and segregated education provision for students with disabilities because government schools were overwhelmed and under-resourced (Black, 2019).

If we are to assume that inclusive education is a human right—as the United Nations Convention on the Rights of Persons with Disabilities would suggest—then we can view inclusion and diversity in education as a societal value and view it through the theoretical lens of Complex Educational Systems Analysis using the Educational Values Evaluation and Design Framework. In a generic sense, any study could conduct this kind of analysis by examining an educational value across elements, times, levels, and communities. In this article, I will use Bhutan as an illustrative case study to demonstrate the application of the theoretical frameworks. When discussing inclusive education in Bhutan, I am specifically referring to education for children with disabilities as other attributes such as language and ethnicity are not given policy consideration in Bhutan.

Bhutan is a small country about the size of Switzerland nestled between India and Tibetan China and is located entirely in the Eastern Himalaya. The Bhutanese population is approximately 780,000—compared with approximately 8.4 million in Switzerland for some comparison—and most Bhutanese are still primarily involved in small-scale agricultural production. The agrarian cultural heritage of Bhutan is rapidly changing, however, with increased youth urbanization in the capital, Thimphu, and the border town with India, Phuentsholing. While I will not take the space to provide a full historical and cultural context of Bhutan (see Schuelka and Maxwell, 2016 for a more comprehensive overview), I will very briefly state that formal secular education in Bhutan is relatively recent, having been initiated slowly in the 1960s with the explicit purpose of human capital development. The education system of Bhutan was greatly influenced by India during its early days. Today, nearly a quarter of the entire population of Bhutan attend some form of formal secular schooling.

In terms of inclusive education in Bhutan, this was an initiative that began in the 2000s with the creation and identification of Special Educational Needs (SEN) school sites in various locations across Bhutan (Rinchen Dorji and Schuelka, 2016). With the support of UNICEF and international consultants, some students with disabilities were “mainstreamed” and were “pushed in” and “pulled out” according to their needs. In 2019, the National Policy for Persons with Disabilities was enacted in Bhutan and included requirements for all schools to make reasonable accommodations, although also indicating support for separate and special provisions. This was a codification of a previous draft policy that has been in place since 2012, called the National Policy for Special Educational Needs.

In my original ethnographic fieldwork in Bhutan from 2012 to 2013, my research was primarily focused on understanding the discourses of disability and inclusive education as it flowed between micro, meso, and macro-levels. I also wanted to understand where inclusive education “came from” and entered into Bhutan as previously the idea of Education for All was not a primary concern for the Bhutanese education system. I found that there was a confusing and often-conflicting array of exogenous ideas brought into Bhutan, representing both a human rights-inclusion perspective and a medical-special perspective, that mixed together with endogenous Buddhism and the Bhutanese development philosophy of Gross National Happiness (Schuelka, 2015). The understanding of disability and inclusion morphed and changed as it passed through various actors and institutions, and was shaped by certain individual policy actor-“influencers.”

However, while Bhutan was providing greater access to education provision for a greater number of students, there was not a similar increase in the provision of teachers or the quality of education. During my original fieldwork, I found that various forms in the education system—physicality, curriculum, pedagogy, language—produced and constructed “students with disabilities” simply through the institution itself (Schuelka, 2018b). Re-looking at this project through the lens of the CESA, it is easy to see that disability and student difference was being enacted across a wide variety of community groupings and relationships. In terms of the attributes of education, Bhutan was working toward increasing access, but not enough on the quality of the education experience and adjusting the system to accommodate diversity in the classroom.

More recent work on educational values in Bhutan—Toyota Foundation, 2018–2022—uncovered significant misalignments of educational values across various elements within the complex system. This is where the EVED framework originated. To isolate just inclusivity as an educational value, we observed that curriculum was narrow, fixed, and oversubscribed. Teachers were forced to maintain certain modes of pedagogy—didactic and teacher-centered—in order to simply get through the required curriculum. The education system in Bhutan is examination-focused, with summative exams at the end of every grade and major exams at grade 10 and grade 12. At each examination point, student failure means that they do not progress. While teachers indicated to us in focus groups that they wanted to introduce variety into their teaching, it simply was not possible given the rigid curriculum and examination structure. This was counterproductive to the value of inclusion in the classroom, as there was no space for accommodation, flexibility, and adaptation. Educational material availability is a significant challenge in Bhutan as well, with a heavy reliance on textbooks distributed centrally.

Surrounding the education practices in schools is also the larger socio-cultural and economic context of a rapidly changing Bhutan. In what seems counterintuitive, the largest group of unemployed and underemployed persons with Bhutan are actually school graduates (NSB, 2018). This is an indicator that the education system is misaligned with the supply needs and availability demands within the economy. While access to school was a given as a human right, there is a diminishing sense that education provides enough utility to greater society. Within the CESA, this highlights the interconnectedness of the educational attributes dimension, as well as how they can be connected and yet misaligned. This misalignment can also further erode the value of inclusion in the education system as it lessens the incentives for education to gain social mobility. In other words, by performing a personal cost-benefit analysis an individual may decide that going to school does not actually give them much.

The case of Bhutan makes an important point for inclusive education in that the value of inclusion and diversity in education needs to be aligned with enough elements in a complex system in order to make a difference. Simply providing the legal and policy means of “access” are not enough to reform a system. A singular focus only on teachers and their pedagogical practices will not reform a system that has not made other systemic changes to curriculum, or the structure of the school day, or the physical space of classrooms, or the availability of resources, and so on. In Bhutan, as in all other countries, increasing the value of inclusion and diversity in education means working with all elements of the complex system simultaneously. This, of course, is not the easy path forward, but I would argue that it is the most effective.

The (complex) path forward

In summary of the above sections, I argued that reform in complex educational systems involves an alignment of elements across the system. Those that have advocated for greater inclusivity and diversity as an educational value have often pulled one lever in a system without accounting for the other levers that either affect, or are affected by, the first lever. The first suggestion I make through this article is that an important first step in education reform is to recognize the complexity of the education system itself.

The second suggestion I make through this article is that once you recognize the complexity of the system, you can begin to understand and analyze it. In this article, I have provided two tools in which to facilitate this: Complex Educational Systems Analysis (CESA) and its application, the Educational Values Evaluation and Design (EVED) Framework. However, I do not

propose these as yet another tool to identify “barriers to inclusive education,” or whatever other phenomenon you are studying. The whole reason that “Evaluation and Design” are in the EVED Framework is that I argue that through evaluation we can provide better design to achieve educational value goals. Like the Ishikawa diagram on which it is based, the intention of the EVED Framework is to intentionally work toward solutions by recognizing that interconnected elements inform the entire process.

For example, in the case of Bhutan, our most significant finding was that poor curriculum was pulling on both teachers and their pedagogical practices in that it was not incentivizing teachers to be more inclusive in their classrooms and encouraged complacency in terms of engaging with students. It can also be argued that old-fashioned pedagogical modes of delivery—namely the kinds of teacher-centered practices imported from India—have entrenched themselves in Bhutanese education and shaped the curriculum around that preferred mode of delivery. For the teachers, their identities mostly revolved around content delivery and student academic achievement. I strongly aver that it is not teacher attitudes that are a barrier to inclusive education, but it is in how other elements in a complex system incentivize or disincentivize teachers from reconceptualizing their role and their identity in a different way. If I was an education policy maker in Bhutan—which I am not—and I wanted to make the education system more inclusive, I would focus on how curriculum affected the other elements, and in how the other elements affected curriculum. Of course, the policy makers need to want to make the education system more inclusive, but that is a whole different issue.

The third suggestion I make through this article is that education reform is slow, often nonlinear, and uneven. This is the chaos of a complex system at work. However, just as in nature, sometimes there are moments of synchrony (Strogatz, 2003) when elements move together in a complex system. If elements move together, sustainable reforms can occur, but it involves deep and longitudinal work in a system that is fairly unglamorous. This is the problem with education reform—particularly as perpetuated by NGOs and private foundations—in that often it is more lucrative to an organization to strive for one, big, seductive reform that claims to be the panacea for all of the education systems’ ills. As complexity theory tells us, it is never just about one thing, and a new element in the system will have an untold number of effects across all other elements.

Finally, I suggest that reform in complex educational systems needs to pay greater attention to the underlying purpose of education and the values the education system wishes to inculcate. As I said above, these can be purposefully designed if a greater recognition of the complexity of the system is made. It is important to ask teleological questions of education, and not to assume that education has always had the purpose of being in human history—which it most certainly has not. Fortunately, education is a human-constructed system and it is certainly within our power to shape it any way we want to. Unfortunately, we have made the education system extraordinarily complex.

References

- Amaral, L.A.N., Ottino, J.M., 2004. Complex networks: augmenting the framework for the study of complex systems. *Eur. Phys. J. B* 38, 147–162. <https://doi.org/10.1140/epjb/e2004-00110-5>.
- Bartlett, L., Vavrus, F., 2017. *Rethinking Case Study Research: A Comparative Approach*. Routledge, New York.
- Black, A., 2019. A picture of special educational needs in England—an overview. *Front. Educ.* 4 (79). <https://doi.org/10.3389/feduc.2019.00079>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development: Experiments in Nature and Design*. Harvard University Press, Cambridge, MA.
- Davis, B., Sumara, D., 2006. Complexity and Education: Inquiries Into Learning, Teaching, and Research. Lawrence Erlbaum Associates, Mahwah, NJ.
- Dorji, R., Schuelka, M.J., 2016. Children with disabilities in Bhutan: transitioning from special educational needs to inclusive education. In: Schuelka, M.J., Maxwell, T.W. (Eds.), *Education in Bhutan: Culture, Schooling, and Gross National Happiness*. Springer, Singapore, pp. 181–198.
- Honig, M.I., 2006. Complexity and policy implementation: challenges and opportunities for the field. In: Honig, M.I. (Ed.), *New Directions in Education Policy Implementation: Confronting Complexity*. State University of New York Press, Albany, NY, pp. 1–24.
- Labaree, D.F., 1997. Public goods, private goods: the American struggle over educational goals. *Am. Educ. Res. J.* 34 (1), 39–81.
- Levinson, B.A.U., Sutton, M., Winstead, T., 2009. Education policy as a practice of power: theoretical tools, ethnographic methods, democratic options. *Educ. Pol.* 23, 767–795. <https://doi.org/10.1177/0895904808320676>.
- Malen, B., 2006. Revisiting policy implementation as a political phenomenon: the case of reconstitution policies. In: Honig, M.I. (Ed.), *New Directions in Education Policy Implementation: Confronting Complexity*. State University of New York Press, Albany, NY, pp. 83–104.
- Massouti, A., 2018. (Re)thinking the adoption of inclusive education policy in Ontario schools. *Can. J. Educ. Adm. Pol.* 185, 32–44.
- NSB [National Statistics Bureau], 2018. 2017 Population and Housing Census of Bhutan. National Statistics Bureau, Thimphu.
- Opfer, D., Pedder, D., 2011. Conceptualizing teacher professional learning. *Rev. Educ. Res.* 81 (3), 376–407. <https://doi.org/10.3102/0034654311413609>.
- Ortner, S.B., 2006. *Anthropology and Social Theory: Culture, Power, and the Acting Subject*. Duke University Press, Durham, NC.
- Phillips, D., Ochs, K., 2003. Processes of Policy Borrowing in Education: some explanatory and analytical devices. *Comp. Educ.* 39 (4), 451–461. <https://doi.org/10.1080/0305006032000162020>.
- Rapplee, J., 2006. Theorizing educational transfer: toward a conceptual map of the context of cross-national attraction. *Res. Comp. Int. Educ.* 1 (3), 223–240.
- Schuelka, M.J., Engsig, T.T., 2022. On the question of educational purpose: complex educational systems analysis for inclusion. *Int. J. Incl. Educ.* 26 (5), 448–465. <https://doi.org/10.1080/13603116.2019.1698062>.
- Schuelka, M.J., Maxwell, T.W. (Eds.), 2016. *Education in Bhutan: Culture, Schooling, and Gross National Happiness*. Springer, Singapore.
- Schuelka, M.J., Sherab, K., 2022. Educational values in complex systems: an introduction to the Educational Values Evaluation and Design framework and a case study of inclusivity in Bhutan. *Cur. Iss. Comp. Ed.* 24 (1), 114–131. <https://doi.org/10.52214/cice.v24i1.8773>.
- Schuelka, M.J., 2015. The evolving construction and conceptualization of disability in Bhutan. *Disabil. Soc.* 30 (6), 820–833.
- Schuelka, M.J., 2018a. Advancing a comparative case study approach towards education and disability research: an example from Bhutan. In: Singal, N., Lynch, P., Johansson, S. (Eds.), *Education and Disability in the Global South: New Perspectives From Asia and Africa*. Bloomsbury, London, pp. 89–106.
- Schuelka, M.J., 2018b. The cultural production of the “disabled” person: constructing student difference in Bhutanese schools. *Anthropol. Educ. Q.* 49 (2), 183–200.

- Shore, C., Wright, S., 1997. Policy: a new field of anthropology. In: Shore, C., Wright, S. (Eds.), *Anthropology of Policy: Critical Perspectives on Governance and Power*. Routledge, London, pp. 3–39.
- Steiner-Khamsi, G., Stolpe, I., 2006. *Educational Import: Local Encounters With Global Forces in Mongolia*. Palgrave Macmillan, New York.
- Strogatz, S.H., 2003. *Sync: The Emerging Science of Spontaneous Order*. Hyperion, New York.
- Sutton, M., Levinson, B.A.U. (Eds.), 2001. *Policy as Practice: Toward a Comparative Sociocultural Analysis of Educational Policy*. Ablex, Westport, CT.
- Tyack, D., Cuban, L., 1995. *Tinkering Towards Utopia: A Century of Public School Reform*. Harvard University Press, Cambridge, MA.
- Walton, E., McIntyre, J., Awidi, S.J., De Wet-Billings, N., Dixon, K., Madziva, R., Monk, D., Nyoni, C., Thondhlana, J., Wedekind, V., 2020. Compounded exclusion: education for disabled refugees in Sub-Saharan Africa. *Front. Educ.* 5 (47). <https://doi.org/10.3389/feduc.2020.00047>.

Paying for inclusion and exclusion

Amanda Watkins, Verity J. Donnelly, and Cor J.W. Meijer, European Agency for Special Needs and Inclusive Education, Odense, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction	36
Shifting the focus of financing within education systems	37
Financing inclusive education goes beyond financing general education	38
Financing provided by a range of stakeholders	38
Financing “non-educational” aspects of provision	38
Financing different sectors within the education system	38
From financing special needs education to financing inclusive education for all learners	38
Financing special needs education	38
Financing inclusive education systems	39
Moving toward financing that takes a preventative policy approach	40
Inclusive financing of whole education systems	41
Ensuring learners are effectively included in appropriate educational opportunities	42
Promoting a school-development approach for inclusive education	42
Providing innovative and flexible learning environments	42
Ensuring transparent and accountable systems of inclusive education	43
Reinforcing the need for prevention-based policy approaches	43
Including all learners in appropriate educational opportunities	44
Underpinning inclusive education by a school-development approach	44
Developing innovative and flexible learning environments	45
Introducing transparent systems of accountability for inclusive education	45
Concluding comments	46
Acknowledgments	47
References	47
Further reading	49
Relevant websites	49

Glossary

Capacity-building The process of building the capacity of mainstream schools to meet the needs of all learners, rather than distributing additional resources for learners with additional support needs. This process involves increasing the knowledge and skills of all education professionals (i.e., leaders, teachers and specialist staff), as well as enhancing the collaborative practices within schools and across the local communities, with the ultimate aim of transforming learning and teaching and improving learner outcomes ([European Agency, 2014](#))

Cross-sectoral co-operation Cross-sectoral cooperation means that multiple stakeholders from outside the education sector are engaged to create an alliance or network of different individuals, organizations and sectors with diverse perspectives and resources to address complex challenges more effectively ([Cross-sectoral co-operation](#)).

Financing Education financing: ... refers to governmental and organizational processes by which revenues are generated (through taxation, tuition, fees, and philanthropy), distributed, and expended for the operational and capital support of formal schooling ([Oxford Bibliographies, no date](#))

Financing system A financial system consists of institutional units and markets that interact, typically in a complex manner, for the purpose of mobilizing funds for investment, and providing facilities, including payment systems, for the financing of commercial activity ([OECD, 2005](#); Glossary of Statistical terms)

Funding Funding refers to “money provided, especially by an organization or government, for a particular purpose” ([Oxford Dictionaries, no date](#)). Within an educational context, financing is most often the responsibility of different ministries and/or sectors and funding is therefore provided in different ways

Governance “The exercise of political, economic and administrative authority” (p. 275). Governance refers to “the way that organizations or countries are managed at the highest level, and the systems for doing this” ([Cambridge Dictionary, no date](#)). It also refers to measures ensuring effective planning intending to develop appropriate and sustainable provision. The governance of system-level budgeting deals with: the role and responsibilities of the different actors involved in the budget process; the distribution of authority, particularly between the ministry of finance, line ministries and agencies; and the

institutional arrangements and relationships between these actors. Governance also touches upon the question of distribution of power between regional and central government, as well between the executive and legislature (Fakharzadeh)

Inclusive education system UNESCO stresses the need for a systemic, capacity-building approach to developing inclusive education systems: Implementing changes effectively and monitoring them for impact, recognizing that building inclusion and equity in education is an on-going process, rather than a one-time effort (UNESCO, 2017, p. 13). All European countries are committed to working toward ensuring more inclusive education systems. They do so in different ways, depending on their past and current contexts and histories. Inclusive education systems are seen as a vital component within the wider aspiration of more socially inclusive societies that all countries align themselves with, both ethically and politically. The ultimate vision for inclusive education systems is to ensure that all learners of any age are provided with meaningful, high-quality educational opportunities in their local community, alongside their friends and peers (European Agency, 2015, p. 1)

Input funding Input funding (also sometimes referred to as “needs based funding”; see Meijer, 1999; European Agency, 2016a) is when the funding is based, for example, on the identified need of each of the destination levels, such as the number of learners with special needs in a school, municipality or region. Inputs may also be defined in terms of referral rates, low achievement scores, the number of disadvantaged learners and so on. The key point is that funding is based on the (expressed or measured) needs (European Agency, 2018a)

Policy framework A policy framework brings together policies/policy elements that set out the requirements and processes for reaching policy goals in line with national/organizational values and principles. A policy framework outlines roles and responsibilities for policy development, stakeholder engagement, implementation, dissemination, monitoring/evaluation, governance and operational processes (European Agency working definition)

Throughput or discretionary funding Throughput, or discretionary, funding is based on the functions or tasks that have to be undertaken or developed. It is not based on needs, but rather on the services provided by a school, municipality or region. Finances are allocated on the condition that particular services will be developed or maintained. Schools, municipalities or regions are equally treated: funds are based on total enrollment or on other population indicators. Of course, certain output conditions can be included in this model, but funding itself is not based on outputs (or inputs) (Meijer, 1999; European Agency, 2018a)

Introduction

The UNESCO Policy Guidelines (2009) describe inclusive education as a process that strengthens the capacity of the education system to reach out to all learners. The policy guidelines, as well as the 2017 UNESCO Guide for ensuring inclusion and equity in education, argue that an inclusive education system can only be created if all schools are supported to be more effective at educating all children in their communities. This line of thinking underpins the work of the European Agency for Special Needs and Inclusive Education (European Agency) with 31 European countries, covering 35 jurisdictions. European Agency work is guided by the overarching principle that inclusion and equity should provide the foundation of *all* educational policies and practices.

The position of the European Agency is that inclusive education is a human-rights issue, and also a part of a wider societal process working toward more inclusive and equitable societies as argued by international organizations (such as UNESCO, 2017 and 2020b), as well as European Union organizations (Council of the European Union, 2018a,b; Council of Europe Commissioner for Human Rights, 2017; European Union Agency for Fundamental Rights, 2020; European Commission, 2020).

The position statement of the European Agency argues that: “The ultimate vision for inclusive education systems is to ensure that all learners of any age are provided with meaningful, high-quality educational opportunities in their local community, alongside their friends and peers” (European Agency, 2015, p. 1). This position has to be recognized as a vision for all countries across Europe and globally.

Ydo (2020) links issues raised in work conducted by UNESCO (2020a) with challenges for achieving the vision of inclusive education highlighted in the Global Education monitoring Report Inclusion and Education: All Means All (2020b). These challenges include: differing understandings of the word inclusion; a lack of teacher support; absence of data on those excluded from education; inappropriate infrastructure; persistence of parallel systems and special schools; lack of political will and community support; untargeted finance; uncoordinated governance; multiple but inconsistent laws; and policies that are not being followed through (Ydo, 2020, p. 100).

Such challenges mean that, rather than “achieving” a fully inclusive system, countries have a range of policies that can be considered inclusive or exclusionary to differing degrees (European Agency, 2020; UNESCO, 2020b) and Ainscow (2020) stresses the importance of ensuring contextually sensitive change policy development work.

Financing impacts upon all of these challenges. European Agency work with its member countries over more than two decades (1999, 2016a,b, 2018a,c) has clearly shown that the strategies and mechanisms used to finance education systems are possibly the single most critical lever in making those systems more inclusive. The crucial role of financing systems in developments toward inclusive education have been identified by researchers (Slee, 2018; Meijer and Watkins, 2019; Ebersold et al., 2020), as well as international organizations (UNICEF, 2015; UNESCO, 2017, 2020b).

Financing refers to the process of providing different types of funds and resources to ensure the provision of educational opportunities for all children, learners and adults (OECD, 2017a; European Agency, 2018a). Importantly, financing also relates to the different possibilities and degrees of freedom within different parts of the education systems to use available funds and resources in flexible and innovative ways to meet all learners' needs (Meijer and Watkins, 2019; European Agency, 2016a).

However, financing inclusive education involves far more than "spending" issues within the general education system (European Agency, 2018a). Allocating sufficient money is essential, but effective financing systems also require resourcing mechanisms and strategies that incorporate:

- Quality assurance processes designed to achieve, maintain and enhance quality in inclusive education. Such procedures should examine how educational organizations account for their activities, accept responsibility for them and share information on their results openly and transparently;
- Governance mechanisms that ensure accountability, transparency and responsiveness of the system for inclusive education;
- System capacity building strategies for developing and strengthening professionals' attitudes, skills and abilities. Such strategies should consider the support and resources that educational organizations and communities require to develop the necessary working procedures for inclusive education (European Agency, 2018b).

Slee (2011) argues that developments toward inclusive education must include a *practical understanding* of what attitudes, processes and structures are involved in exclusionary policies and practices. A further critical factor is whether *inclusion* or *exclusion* is the essential focus of financing within education systems (Ebersold et al., 2020).

The starting point for this paper is that efficient financing mechanisms aim to enable schools to provide flexible learning opportunities for all learners and support teachers to meet diverse educational needs through innovative teaching practices (Ebersold and Meijer, 2016; European Agency, 2016b,c, 2018c). Efficient and effective financing mechanisms are in effect geared toward *paying for inclusion*.

By contrast, inefficient financing mechanisms may act as a driver for maintaining or even promoting education approaches that reinforce the labeling of some learners and their placement in different types of provision, including separate settings such as special classes or special schools. Financial incentives maybe be "hidden in legislation" (Meijer, 1999) and departments from different ministries may use financing policies with conflicting incentives. It is important to reveal and analyze the ways in which financial incentives or disincentives act as a support for, or a barrier to, inclusive education. It is also important to see financing from a broader perspective than ministries of education (European Agency, 2018a,c).

These possible inefficient and ineffective financing mechanisms, geared toward *paying for exclusion*, are also examined within this paper.

The following sections aim to explore what paying for inclusion and exclusion means in relation to the development, implementation and outcomes of inclusive education policy frameworks. Two main issues are considered:

- How financing inclusive education needs to go beyond financing general education; and
- The challenges of moving from financing special needs education to financing inclusive education for all learners.

Shifting the focus of financing within education systems

How general education systems are financed has been covered extensively—for example, in research at the international level by OECD (2017a), UNICEF (2015), UNESCO (2020b). Across European countries, the work of the European Commission (EACEA/Eurydice, 2014) demonstrates " ... an overall picture of great variety across Europe with respect to funding systems. These systems have developed over many decades to meet the needs of individuals, wider society and the economy. The changing priorities of education systems have also shaped the way in which funding systems have developed" (p. 8).

Funding specifically dedicated to general education in many European countries encompasses frameworks—services, structures and organizations—for supporting the individual needs of learners, or the collective needs of an entire school community, municipality or region (EACEA/Eurydice, 2014; European Agency 2016a, 2018a,c). These resourcing frameworks often focus upon providing technical support or offering materials for the planning and implementation of local, regional or national support services. They may also offer advice and consultancy both at the system and at the individual learner level, as well as providing in-service training opportunities.

The European Commission (2014) argues that two specific factors within countries' systems result in challenges to identifying and understanding the funding methods and criteria that education authorities use when allocating resources within the general education system:

- The level of autonomy given to regional and/or local authorities and schools when making funding decisions; and
- The complexity of funding formulas and variables or indicators used for allocating funds and resources.

While this demonstrates the complex nature of general educational financing, the financing of inclusive education systems, involving additional elements becomes even more complex (European Agency, 2016a,b). These elements are discussed in the following sections.

Financing inclusive education goes beyond financing general education

In most European countries, the financing of inclusive education involves additional elements of multi-level and multi-stakeholder funding and resourcing that covers mainstream education and different forms of specialist provision (Ebersold, 2008, 2016; European Agency, 2016a,b).

Financing provided by a range of stakeholders

Within most countries' systems, education involves financing from a range of sources, including but not limited to ministries of education and finance (Eurydice, 2014). The wider demands of inclusive education mean that increasingly, financing from other ministries—such as those responsible for health, social, labor and ICT issues is required as well as resources from non-governmental organizations and the private sector (UNESCO, 2020b; European Agency, 2016a,b, 2018c).

Funding from non-educational sources may therefore cover health and welfare issues, or consist of financial support for learners and/or their families to cover costs linked to school materials, meals, transport or additional services and technology.

Financing “non-educational” aspects of provision

Systems for inclusive education need to specifically resource non-educational aspects that ensure all learners access to high-quality inclusive education. An example of a non-educational aspect is accessibility as defined by Article 9 of the UNCRPD (United Nations, 2006). This may involve physical access to school premises, accessibility of transport opportunities, as well as access to extra-curricular activities for all learners.

Financing different sectors within the education system

Within all European countries, resourcing of inclusive education may include special settings—either separate classes, units or schools—as a feature of provision within education systems (European Agency, 2019). All countries embed funding for these settings in their resource allocation frameworks and special schools and classes are part of a continuum of service provision offering educational opportunities for learners who experience the greatest challenges in their learning. In some countries, special settings act as resource centers that provide additional support for individual learners and/or enable mainstream schools and teachers to provide inclusive education. In others, they provide education to learners whose special needs and/or disabilities require educational approaches addressing pedagogical, psychological and rehabilitation issues (European Agency, 2019).

From financing special needs education to financing inclusive education for all learners

Across the globe, policy understandings and conceptualizations of inclusive education have grown out of discussions surrounding specialist segregated provision, integration and mainstreaming (Watkins, 2017; Ramberg and Watkins, 2020). Thinking around inclusive education based on a human rights perspective (Donnelly, 2010; Operti et al., 2014) has developed from a focus on the rights of some learners to access different sectors and opportunities within a country's education system toward a focus upon systemic approaches geared toward raising the achievements of all learners (Florian et al., 2017), as well as the overall system (European Agency, 2015; Watkins, 2017; Ramberg and Watkins, 2020).

Within many policy contexts, the understanding of the concepts behind policies for inclusive education have shifted from being mainly associated with learners with special educational needs and/or disabilities and in particular their educational placement, to meeting the needs of a far wider range of learners who may be vulnerable to exclusion. This shift in thinking is also increasingly evident—and necessary—in the shift in emphasis in financing models, mechanisms and strategies (Ebersold et al., 2020).

Changes and developments in policy for inclusive education need to be mirrored by a change in financing policies and mechanisms in order to involve a broad range of stakeholders, within and outside of education. A move from financing special needs education toward financing inclusive education does not necessarily mean spending more, or less, but it does mean using financial resources in different ways for different purposes (European Agency, 2018a,c). The financing of inclusive education can be contrasted with financing special needs education in a number of important ways. These are explored in the next sections.

Financing special needs education

The essential aim of special needs education is to provide additional, specialist support for learners who have diverse (and possible multiple) social, intellectual or physical needs. Special needs education is usually framed by a medical approach to disability and/or special educational needs focusing on learners' inabilities and/or deficiencies rather than on weaknesses within the educational system (Ebersold, 2008; Meijer and Watkins, 2019). Legislation may include defined categories of special educational needs—that is “within learner factors—that are often guided by policies that are administered by different ministries (e.g., health or welfare) or from within special services of the ministry of education (European Agency, 2016a). Special needs education emphasizes the idea of different possible placements in education: enrollment may be in special, separate settings, such as special classes, special schools and special residential schools or institutions, as well as in mainstream settings. It also raises the

possibility of educational provision outside of school settings: in learners' homes, in hospitals, or in other forms of alternative provision.

Financing of special needs education has often been built around locating "problems" within the learner and then allocating resources to provide additional provision to compensate for those problems in order to fit the learner into existing educational arrangements. This process is often associated with a formal assessment of learners' abilities, linked with predefined resourcing strategies, such as funding coefficients (European Agency, 2016a, 2018a). This approach can be characterized as an *input* model of funding—sometimes called needs-based funding. Learners' needs are identified and assessed and a form of official decision or statement is issued before an allocation of additional funding is made (Meijer, 1999, 2003; Meijer and Watkins, 2019). In effect, the higher the level of needs identified, the more funds are made available.

Input-based resource allocation mechanisms can be seen to have a number of potentially negative consequences, for individual learners, as well as the overall education system:

1. Input funding models provide incentives for identifying learners as having some form of special educational need as they then attract additional funds (Meijer, 1999). Such models may increase strategic behavior that leads to a rise in the identification and labeling of learners as having special educational needs.
2. Enrollment in special settings is also encouraged when funding mechanisms lead stakeholders from mainstream settings to regard the enrollment of learners with different learning profiles as a financial burden for them. They may also believe that experts from specialist provision are more able to cope with these learners (Meijer, 1999; Parrish, 2001, 2014).
3. Input funding mechanisms can also be seen to actively encourage mainstream school stakeholders to "outsource" the responsibility for meeting some learners' needs to other professionals, or other forms of dedicated, specialist provision. Outsourcing may result in learners being enrolled in special classes or schools, where such provision is specifically dedicated to meeting a type of need or supporting a particular disability (Ebersold and Meijer, 2016).
4. Crucially, such outsourcing can limit the perceived need, or actual opportunities for any form of development of the mainstream school's ability to more effectively meet diverse learning needs in practice (Ebersold et al., 2020).

Financing inclusive education systems

In contrast to special needs education, inclusive education aims to provide resources that support the adaptation of the education system and environment to the needs of all learners. It involves the identification and removal of barriers to high quality education for all learners by funding and resourcing the development of innovative teaching approaches and practice (Ainscow et al., 2012; OECD, 1999, 2004, 2012; European Agency, 2016b; Parrish, 2014).

The goal of an inclusive education system is to support all learners and provide effectively for each of them. An education system cannot be considered to be of high quality if some learners are out of school or not fully participating in learning opportunities available to their peers (UNESCO, 2009, 2017, 2020a,b).

Financing of inclusive education systems must therefore be guided by Slee's (2018) goal of dismantling "... barriers to education for all children so that they have access to, are present and participate in and achieve optimal academic and social outcomes from school" (p. 2). Funding and resourcing strategies must address the multidimensional aspects of presence, placement, participation and achievement for all learners (Ainscow, 2016; Norwich, 2013), not just those with some form of label or official decision of need.

Financing of inclusive education necessarily involves the provision of funding and resources to support changes in educational culture, as well as in teaching and support practices. Financing inclusive education requires moving away from a "one-size-fits-all" education and funding model, toward a tailored approach that aims to increase the ability of both the system and of individual schools to respond to learners' diverse needs without the need to allocate funds by categorizing and labeling learners.

Such additional funding may enable schools to provide different forms of intensified support for some learners who are at risk of exclusion from learning opportunities. These approaches have been termed as *throughput funding models* (Meijer, 1999; Ebersold and Meijer, 2016; European Agency, 2016b). The European Commission (2014) suggests such approaches can be characterized by the discretionary determination of resources, where the responsible funding authority has discretion in determining the resources to be allocated, based on an understanding of case by case, local resource needs.

Throughput funding mechanisms can have a number of benefits for both learners and the education system:

1. Crucially, throughput, or discretionary resource allocation approaches go some way to addressing equity issues that arise as a result of funding strategies that reinforce disparities between regions or geographical areas.
2. Throughput, or discretionary resource allocation focuses on services that enable schools to provide intensified support to learners who are facing difficulties. Such approaches to funding do not connect eligibility for additional provision or support with an official decision or labeling process, but with a needs identification process that is implemented at the school level. These funding approaches aim at reducing attainment gaps between individual or groups of learners who are at risk of exclusion. For example, funding and eligibility for support without having an official decision of SEN may be dedicated to specific programs meant for learners from disadvantaged backgrounds and/or for learners from a migrant background (European Agency, 2016a).

3. While some forms of specific (input-based) additional funding may be required to support a small number of learners in need of the most intensive support, more flexible (throughput) resource allocation methods can help schools to develop inclusive learning environments that work toward the *prevention* of learning problems by being universally accessible and adaptable to each learner's diverse needs (UNICEF, 2015; European Agency, 2016a, 2018a).

Despite the points raised about different funding approaches, both approaches are needed. However, decisions regarding funding should be made as late as possible within the financing process and be taken at the level closest to the learner involved; also involving teachers and families. Input-based funding decisions to support school needs should be taken at the level of local or regional authorities (Meijer, 1999, 2003) to ensure awareness of the local context.

The following section focuses on how financing mechanisms can support education policies that aim to take a prevention approach.

Moving toward financing that takes a preventative policy approach

Financing inclusive education involves critical actions that are closely related to a shift in funding from special needs education—in particular segregated settings schools or separate classes—to inclusive education in mainstream schools. The implementation of these actions may require a fundamental shift in thinking and understanding in terms of the *intentions underpinning* policy initiatives and strategies.

European Union work (European Commission, 2020) emphasizes the importance of a balance of policy approaches to reflect broader thinking about quality education that meets diverse learner needs.

The 2020 Eurydice report on equity in school education considers system level factors including influences and inter-relationships in 42 European countries' education systems. Key factors impacting upon equity include:

- Stratification (grouping learners based on ability and/or interests);
- Standardization (application of the same quality standards);
- Support measures (to promote equity and mitigate disadvantage); and
- Diversity in types of schools (resulting in an increase in academic segregation and decrease equity).

The report highlights differences in inclusive, equitable learning opportunities occur due to governance and funding mechanisms, as well as curriculum and structural features (e.g., types of school).

The European Education Area (EEA) targets (European Commission, 2020) to reduce low achievement and increase secondary education attainment build on four pillars:

- Monitoring (allowing tracking/targeted action);
- Prevention (in particular for groups at risk);
- Early intervention (for learners showing difficulties); and
- Compensation (for those who have already had poor results/need a second chance).

The concepts of Prevention-Intervention-Compensation policy approaches (set out in Council recommendation June 28, 2011 on policies to reduce Early School Leaving—OJ C 191 July 1, 2011) have been used by the European Agency (2018d) to assess the current policy situation in European countries and make recommendations about policy actions that enable progress by increasing preventative approaches.

Within its work with member countries, the following operational definitions have been agreed:

- *Prevention*: policies that aim at *avoiding* educational exclusion and longer-term social exclusion (for example, anti-discrimination legislation promoting a rights approach, avoidance of disabling policies that lead to gaps in provision, lack of qualifications, etc.).
- *Intervention*: policies *supporting* the effective implementation of inclusive education for example, the existence of clear policies leading to high-quality flexible support systems for mainstream education).
- *Compensation*: policies *addressing* existing difficulties that will lead to educational exclusion if they are not implemented (for example, separate segregated educational programs or provision, support for failing schools, second-chance educational programs) (European Agency, 2018d).

The goals of inclusive education (e.g., to remove barriers to learning, support all learners and provide effectively for each of them) are most effectively met when policy and practice *prevents* different forms of educational exclusion before they happen; *intervenes* to ensure that good quality inclusive education is available for all learners at all times; *compensates* with specific actions and provision when prevention and intervention are not enough to ensure learners' needs are adequately met in inclusive settings (European Agency, 2018d; European Agency, in press).

In the following section, the use of this approach in considering financing policies, strategies and mechanisms and determining the overall level of inclusiveness of countries education systems is discussed.

Inclusive financing of whole education systems

Inclusive education is a multi-faceted process requiring inputs from different groups, organizations and sectors with different professional identities and experiences. The effective implementation of inclusive education systems requires synergies among stakeholders that are promoted by clear financing mechanisms that support a move toward a prevention policy approach. Such policy mechanisms aim to fund inclusion and avoid funding exclusion (European Agency, 2018c).

Financing of inclusive education systems may or may not be framed as a single “stand-alone” policy. In many country contexts, the goals, objectives and strategies for the implementation of financing are embedded within other educational policies and actions (European Agency, 2016a, 2018c).

It is therefore important that a preventative financing approach is seen as operating at the interface between several different educational policy areas, as shown in Fig. 1 below.

All of these policy areas need to be considered in order to address the challenges of compensatory financing approaches that essentially pay for exclusion and move toward preventative financing approaches that aim to pay for inclusion.

Within a comprehensive policy framework for financing inclusive education systems, financing is not an end in itself. Rather, it is a tool for promoting and ensuring inclusive education systems that provide quality educational opportunities for all learners (European Agency, 2018a). Three principles relating to learners’ rights to inclusive education underpin the development and implementation of any preventative financing policies, frameworks and strategies that aim to pay for inclusion and not exclusion:

1. **Efficiency**, in terms of improving the cost-benefit relationships within systems;
2. **Effectiveness**, in terms of considering educational outcomes for learners, as well as other stakeholders in the system; and
3. **Equity**, in terms of ensuring equitable educational opportunities through respecting diversity and eliminating discrimination.

European Agency thematic project work with its member countries during 2014–2018 has highlighted that efficient cost-effective inclusive education systems can be directly connected with four underpinning resourcing issues and related critical resourcing factors that determine equitable, efficient cost-effective inclusive education. The factors are, in turn, linked to key funding drivers that are considered essential for implementing effective financing policies (European Agency, 2018a,c). Together, the issues, factors and drivers can be seen as a framework for inclusive financing of whole education systems.

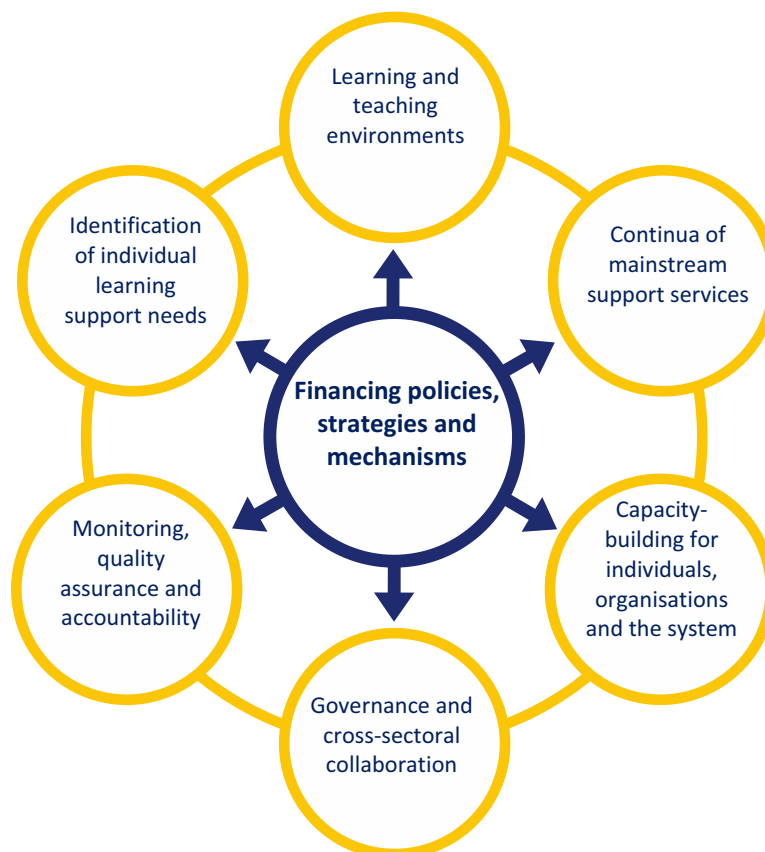


Fig. 1 Financing at the interface with other policy areas for inclusive education systems.

The four sections below describe the main cross-sectoral issues to be considered within preventative financing that aims to pay for inclusion.

Each section begins by presenting the *current challenges* to be addressed in relation to funding exclusion and then the *main goal* for financing that aims to take a prevention approach. The tables then describe the *critical resourcing factors* and *key drivers* associated with a more preventative approach, targeted toward funding inclusion. A consideration of these factors and key drivers aims to address the current challenges and support the development of a more preventative financing policy approach.

Ensuring learners are effectively included in appropriate educational opportunities

Current challenge: to prevent exclusionary strategies that deny learners their rights to and within education. These strategies may involve excluding learners from education, from participation in inclusive education, or from engagement in meaningful learning opportunities. Such strategies are often linked to the over-identification of learners who require an official decision of special educational needs and, potentially, to increasing expenditure related to segregation and/or individual learner support and provision.

Main goal: ensure all financing strategies are targeted toward educational inclusion, not exclusion.

<i>Critical resourcing factors</i>	<i>Key drivers</i>
There is cross-sectoral policy commitment to the right to inclusive education for all learners	The financial commitment toward inclusive education is identified and specified. The financial commitment toward high quality outcomes for all system stakeholders is clearly stated. The commitment toward the development of adequately resourced, diverse support measures for learners and stakeholders is clearly stated.
All resourcing mechanisms support the implementation of inclusive education within local contexts using a community-based approach	The resources required to implement inclusive education as a key task and area of responsibility at all decision-making levels are described. The resources required for schools to implement their social responsibility toward inclusive education are described. The specific and targeted resources required for schools to meet the full range of learners' diverse needs are described.
All resourcing mechanisms support school development and capacity-building for inclusive education	The implementation of a sustainable balance between whole-school (throughput) funding approaches and needs-based (input) funding approaches is outlined. The resources required for developing inclusive learning communities are identified and described.

Promoting a school-development approach for inclusive education

Current challenge: to avoid financing mechanisms that act as a disincentive for inclusive education. Flexible financing systems must ensure a school-development approach that builds learning communities through the development of innovative and flexible forms of teaching that combine high performance and equity.

Main goal: support school teams to take responsibility for meeting all learners' needs.

<i>Critical resourcing factors</i>	<i>Key drivers</i>
There are incentives for the provision of supportive learning environments in all schools	The targeted financial support for schools and learners at risk of under-achievement is specified. The provision of resources for effective working within learning networks is described.
School autonomy in taking responsibility for meeting all learners' needs is promoted and supported	The degrees of flexibility in the use of public funding are outlined. The degrees of schools' organizational flexibility over curricula, assessment and resource allocation are clearly stated.
The implementation of inclusive education is embedded within quality assurance mechanisms at school level	The resources to support schools to implement distributed leadership are outlined. The resourcing mechanisms necessary for supportive and innovative learning environments are described.

Providing innovative and flexible learning environments

Current challenge: to avoid funding mechanisms that act as an incentive for segregation and exclusion when teaching and support in mainstream settings are seen as inadequate for meeting learners' needs. This may lead stakeholders to perceive that special settings (i.e., separate schools and classes) provide better educational support for some learners.

Main message: effective funding mechanisms are an incentive to inclusive education when they promote capacity-building that empowers stakeholders to develop innovative and flexible mainstream learning environments for all learners.

<i>Critical resourcing factors</i>	<i>Key drivers</i>
All resourcing and support systems promote capacity-building strategies at different system levels	The resources required for local-level community-based capacity-building strategies are outlined. The resources required for school-level capacity-building strategies are outlined. The specific and targeted resources required for capacity-building strategies around the provision of external support for learners' needs are outlined.
Special settings are effectively supported to act as a resource for mainstream settings	The incentives for special settings to act as resource centers supporting the mainstream sector are specified. The resources required to ensure inclusive education issues are embedded within all pre- and in-service training of specialists working in separate, specialist settings are specified.
Inclusive education is embedded within all professional development opportunities	The resources required to embed inclusive education in all teacher training/education opportunities are identified and described. The resources required to promote leadership capabilities for developing inclusive schools are identified and described. The resources required to include parents in training/development opportunities are identified and described.

Ensuring transparent and accountable systems of inclusive education

Current challenge: to avoid resource allocation mechanisms that promote the labeling of learners as these are not cost-efficient or equitable in the long term. Rather, areas for development within educational support and provision should be identified. Ineffective cross-sectoral collaboration (i.e., with health and social protection services) can result in duplication of services and inconsistent approaches.

Main message: funding and resourcing systems that balance efficiency, effectiveness and equity issues are clearly linked to regulatory frameworks focusing on overall system governance, accountability and improvement.

<i>Critical resourcing factors</i>	<i>Key drivers</i>
There are multi-level and multi-stakeholder governance strategies that promote effectively co-ordinated systems for inclusive education	The governance and resourcing mechanisms necessary to promote and guide the work of networks of schools are outlined. The governance and resourcing mechanisms necessary to promote and guide local-level, horizontal collaboration, resource allocation and decision-making within and across local communities are outlined. The governance and resourcing mechanisms required to ensure an effective professional inter-disciplinary support framework are described. The resources required to ensure there is an effective and efficient inter-ministerial governance framework are identified and described.
A move from procedural control mechanisms to systems where key stakeholders are clearly held accountable for inclusive education is supported and implemented	The connection between the funding of inclusive education and evidence-based resource-planning is described. The resources required to implement monitoring mechanisms that link schools' use of resources to evidence of progress toward efficiency and equity goals are identified and outlined. The funding of inclusive education is clearly connected to and mapped against a framework of system goals and indicators for inclusive education. The resources required to embed inclusive education issues within all reporting and dissemination mechanisms are outlined.
The goals and objectives for inclusive education underpin the quality assurance framework	The resourcing mechanisms required to ensure inclusive education issues are embedded within all quality assurance systems are identified and outlined. The resources required to develop and implement inclusive education quality assurance tools and mechanisms are identified and outlined.

Reinforcing the need for prevention-based policy approaches

The COVID-19 pandemic has focused attention on the inadequacy of many of the traditional structures and processes in education systems—including how these structures and processes are financed. Many of these are the very same structures and processes that need to be transformed in order to make education systems more inclusive and ensure that all learners, in particular those from disadvantaged groups are included. The [United Nations \(2020\)](#) estimated that 23.8 million additional children and youth are at risk of dropping out of school due to the pandemic's economic impact.

This economic impact is likely to affect the funding allocated to education (World Bank, 2020) and will increase the need to consider more efficient and effective ways to increase quality of education for all learners within the system. For many countries an examination of the effectiveness and efficiency of system processes is considered overdue. According to OECD (2019) there has been no real overall improvement in learning outcomes for students in OECD countries even though expenditure on education has risen by more than 15% in the past decade.

Pursuing the economic argument, Di Pietro et al. (2020) note that the learning loss suffered by students during the COVID-19 crisis "... will translate into a reduction of available human capital, with negative effects on productivity growth, innovation and employment, including future lower earnings for the student cohorts directly affected by the lockdown" (p. 5). They stress that these numbers should be considered by policymakers when making decisions about investment to mitigate the negative effects of COVID-19 on education.

Overall, an important lesson from the pandemic is that countries need to "... [build] resilient education systems for equitable and sustainable development" in order to "reimagine education and accelerate change in teaching and learning" (United Nations, 2020, p. 3). This could be helped by funding prevention policy approaches, in particular attending to the needs of learners experiencing disadvantage to ensure early intervention. Sahlberg (2020) notes that such approaches can help children "catch up" during teaching and learning and effectively reduce performance gaps. Some compensatory policies will also be needed at least in the shorter term with funding for marginalized learners who have been left behind.

This section will set out some factors to consider in planning and funding responses to COVID-19 to ensure that the needs of all learners are taken into account and exclusionary mechanisms avoided. World Bank (2020) highlights that in all countries, the pandemic places a spotlight on the need to use resources as efficiently and as equitably as possible, addressing factors such as uneven distribution of teachers, fiscal transfer formulas that fail to take account of differences in student populations and weaknesses in public procurement systems.

The underpinning resourcing issues set out above are used here to consider the benefits of a move toward a more inclusive system as countries recover from the pandemic and avoid the risk that in the absence of appropriate policy measures, the short-term inequality caused by COVID-19 may persist or even grow over time, leading to more economic disparity in the future (Di Pietro et al., 2020).

Including all learners in appropriate educational opportunities

Poverty inevitably impacts on children's education. Van Lacker and Parolin (2020) note that food insecurity and housing instability impact on physical health and mental well-being—both associated with low educational attainment.

Dorn et al. (2020) point out that in many communities, schools are also the hub for supports such as school meals, mental-health counseling, and childcare. In many homes, especially for low-income families, students lack access to the internet, devices, and a dedicated, quiet place to study (p. 2). Lack of heating and of outdoor spaces and inability to speak the majority national language all compound the difficulties of home schooling. Further, for some children, schools provide protection for those who do not feel safe at home. Sectors need to work together, coordinated across all system levels to fund the development of resilient community services that support children in a more holistic way.

Such an approach to education can also better address learners' social and emotional as well as learning needs (Cerna, 2020). Cerna notes that: "Considerable joint efforts by school leaders, teachers, parents, students, educational and health care professionals, and communities will be needed to (re-) create schools as safe, supportive and inclusive places for all students". Funding such developments will support prevention policy approaches, reducing more costly exclusion from schools and the need for compensation in the longer term.

OECD (2020) state that, if it is true that schools account for approximately one-third of variability in student test results, then the best opportunities for improving the quality of educational outcomes lie in developing system level conditions that impact on lives of children. Key here will be a reduction in child poverty. The World Bank (2020) notes that COVID 19 will slow progress on poverty reduction and in some cases reverse previous gains. This could be compounded by a reduction in support from NGOs and charities who may also be affected by reduced funds and adds further weight to the need for cross sector collaborative working to use resources in the most efficient way to support learners and families in local communities (UNESCO, 2020b).

The role of parents also becomes more critical, in particular for younger learners and others with additional support needs who may find online learning a challenge. Di Pietro et al. (2020) suggest that parents should be involved in the design and implementation of learning strategies as they need to fully understand what is taught and why as well as the emotional challenges of online learning. Effective communication between parents, teachers, and the school will be essential to the success of online learning strategy with a range of benefits for every community.

Underpinning inclusive education by a school-development approach

OECD (2020) highlight the need for education systems that are built on trust-based professionalism. This, together with greater flexibility to allow schools the autonomy to adapt the curriculum to local needs and strengths and teachers to adjust teaching and learning to changing circumstances, builds greater resilience to cope in times of crisis. Skilled teachers who really know their learners can make the most of the available flexibility in schools/classroom organization and in curriculum to provide appropriate learning opportunities for everyone (European Agency, 2018e).

Dorn et al. (2020) recognize that a key factor fueling the disparity in achievement is disparity in teaching. To support learners and close the achievement gap, teachers should use evidence-based strategies, targeted to the students who need them most, leveraging the best formative assessments and early-warning systems to identify students at risk.

Sahlberg (2020) raised the need for more personalized and blended learning for all learners and notes that students who are more self-directed in their own teaching and learning often do better in both normal and crisis situations as such a strategy: “... *strengthens student engagement, brings about authentic learning and helps schools respond better to emerging inequalities*” (p. 7).

This approach requires resources to develop schools as inclusive learning communities with teachers able to use a range of strategies to support more diverse learning needs with the involvement of parents and others as discussed above. Greater stakeholder involvement, for example with curriculum can benefit staff and learners as well as the wider community (European Agency, 2018e).

A school development approach also depends on developing a system of support for leaders and teachers as well as learners (European Agency, 2018e). Such support may come from specialists working with the school, from peers in schools as well as more structured professional development opportunities.

Funding school development using the strategies outline above will create more innovative and flexible environments as set out in the next section. This enables the focus on preventative approaches to be increased which, together with appropriate interventions for learners who require additional support is likely to reduce the escalation of issues leading to drop out and/or exclusion and the reliance on compensatory approaches in the longer term.

Developing innovative and flexible learning environments

In the aftermath of COVID-19, the Council of Europe (2020) notes the need for innovation in learning and teaching, for example by blending face to face and online provision in new ways that may require innovative pedagogical approaches. COVID could therefore potentially serve as a “catalyst to create a new, more effective hybrid model of educating students in the future” (Kaden, 2020). However, as Schleicher suggests this crisis brings to the fore the need for effective leadership in education, at every level (2020).

Bassot (2020) notes some possible responses to issues raised by COVID, including the need for:

- Teacher professional development, including the sharing of expertise/resources;
- Improving the use of technology with the employment of extra technical staff;
- Flexible curricula that bridge formal non formal and informal learning;
- Considering the role of continuous formative assessment;
- Improving home school communication, cooperation with and support for parents; and
- Autonomy for those who thrive at home.

These all require resources to build school capacity to support learning, in particular for those experiencing disadvantage. While there is unlikely to be any single catch-up approach, evidence from the USA and UK suggests that small group tuition may be one effective response (Education Endowment Foundation, online).

Di Pietro et al. (2020) suggest the possibility of rotating attendance at schools as part of blended learning with an analysis of the curriculum to identify which parts are best suited to which approaches. They stress in particular the importance of identifying learners in need of emotional support and a focus on learners in transition to secondary school.

Given that online learning is a way to reduce potential learning losses and increase the resilience of education systems, it is essential to improve teachers’ digital competences, for example with workshops and training courses (Redecker, 2017). Supporting collaborative professional development between teachers (e.g., teacher networks) is also important, to support peer learning (Di Pietro et al., 2020).

Finally, regarding learners who need additional support, in particular those with disabilities, it is essential to provide assistive technologies that can improve communication and mobility and increase participation. Teachers and families may need support here from specialist provision or advisory services but making best use of such technology as part of mainstream provision is likely to reduce the impact of disability/disadvantage on learning. Here, involving students and their families in the assessment process and undertaking on going monitoring of progress is likely to ensure the most effective use of such resources.

The approaches above are likely to increase school capacity to meet the requirements of diverse learners in more innovative and flexible ways, providing appropriate levels of support and intervention—using resources to increase equity and inclusion.

Introducing transparent systems of accountability for inclusive education

Education is equitable when differences in outcomes are unrelated to learners’ background, economic or social circumstances over which they have no control (OECD, 2020). Council of Europe (2020) stress the need to ensure that no students are cut off from the benefits of education and from the possibility to develop their talents and potential to the full regardless of where they live, their gender, their ethnic origin, their beliefs and convictions, the socio-economic difficulties they face and the discouragement they face in their immediate environment.

The measures taken to control the COVID-19 pandemic have had the most impact on those already experiencing disadvantage. This reinforces the need to attend to equity. In particular, the shift to digital/remote learning has exposed and even increased inequities. While some learners, who receive additional support in school may lack the resources needed and struggle to learn remotely,

others will miss the support of their peers as well as access to resources such as libraries, gyms, laboratories etc that often extend opportunities. Such opportunity gaps may exacerbate achievement gaps in the longer term.

For some learners, remote learning has enabled them to thrive without stress caused by for example peer pressure, bullying etc and this raises the need to consider individual approaches and potentially different patterns of school attendance in more flexible learning communities. Above all schools need to adapt to learners not vice versa and teachers need to accept responsibility and respond to learners who need different approaches.

Globally, there is growing interest in “measuring” school performance, for example through standardized tests for the purpose of evaluation, and comparison of national and subnational education systems. This approach stresses outputs (e.g., performance scores, comparative rankings) over inputs (e.g., teacher preparation) in an attempt to quantify and measure education and pursue the idea that teachers can be “materially dis/incentivized to perform better and to increase student- and school-level performance results” (Holloway et al., 2017, p. 6).

The cancellation of such high-stakes examinations and tests due to COVID has raised the issue of reliance on such externally measured standards. Cairns (2020) questions whether exams, as a centuries old model of assessment are the right approach.

Kellaghan and Greaney (2020) acknowledge some advantages of high-stakes examinations including an association with equity and meritocracy; a way to open up education to those who might have been excluded by discriminatory traditions and a degree of impartiality with anonymity of learners excluding the possible influence of student-teacher relationships, political and family influence, or socio-economic background.

Regarding disadvantages, potentially damaging effects on curriculum, pedagogy and learner and teacher well-being are well-documented (e.g., Kelly, 2019; Lingard et al., 2016; OECD, 2017b; Volante, 2017). Muskin (2015) notes that teachers preparing learners for high stakes assessment may narrow the curriculum and the instructional process, allocating more time to subjects that are tested with less emphasis on formative feedback and the development of learners’ personal competencies.

A further challenge is noted by Kefallinou and Donnelly (2016) concerning in particular learners with disabilities who, not being able to fully participate, may be exempted from tests and examinations. While access can be improved by attention to, for example language issues in test papers and reasonable adjustments at the time of the test, for many learners the stress will mean that the exam may not be a true reflection of their performance.

While tests and examinations are widely perceived to be reliable, Rimfield et al. (2019) found that test scores correlate highly with the teacher assessments in the UK In response to the pandemic, they argue that: “... substituting high-stakes exams for teacher assessments might be a good thing, not just during the current COVID-19 crisis, but on a permanent basis” (Rimfield et al., 2020, para. 18).

Overall, the costs of development and marking of tests and examinations contribute primarily to evaluation rather than to serve children’s learning (Donaldson, 2015) while the development of teacher assessment, including moderation (e.g., through the use of exemplar materials, meetings and professional dialog) can have a positive impact on teachers’ assessment practice, as well as teaching and learning more widely. The Council of Europe (2020) notes that the way in which countries design, assess and recognize qualifications may contribute to early school leaving and the number of learners not in education or training.

OECD (2020) suggest that reducing reliance on high-stakes tests will support a more resilient system, taking account of the issues raised above and the need to fund more inclusive ways to recognize and accredit learning and avoid *paying for exclusion*.

Concluding comments

By considering recent work in countries, as well as their responses to the current COVID-19 crisis, it can be argued that inclusive education systems are most effectively supported by a complementary, balanced combination of prevention, intervention and compensation policy actions. Long-term, sustainable developments toward inclusive education systems can be seen as a combination of these three approaches in order to provide high-quality education for all learners, including those requiring additional support (European Agency, in press).

Prevention approaches can be seen as being cost effective and efficient in the long term while compensation approaches may provide short term “fixes”, but are over time cost inefficient and ineffective in terms of supporting system developments (Ebersold et al., 2020).

A country’s journey to an effective and equitable inclusive education system can be identified by movements away from mainly compensatory policy actions, toward more intervention-and prevention-focused policy actions.

On-going European Agency work with its member countries (European Agency, in press) indicates that most countries are moving toward a predominantly preventative approach in relation to their overall policy priorities for their education systems, as well as their specific goals for inclusive education.

In relation to financing policy approaches, five evaluative questions can be identified for reviewing developments:

1. Do financing policy initiatives take all learners into account?
2. Do financing policy initiatives safeguard the rights of all learners to high-quality inclusive education?
3. Do financing policy initiatives promote the active participation of learners and their families in decision-making that affects them?
4. Do financing policy initiatives monitor, evaluate and secure the effective implementation of an inclusive education system?

5. Do financing policy initiatives address barriers to the inclusive education system? (Adapted from [European Agency, 2018d](#)).

Negative responses to these questions will provide an indication of the degree to which financing mechanisms in countries may be paying for exclusion, rather than supporting the achievement of wider policy goals for inclusive education by *paying for inclusion*.

Acknowledgments

The authors gratefully acknowledge the contributions of all the European Agency country representatives to the activities this article uses as its foundation.

References

- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012. Making schools effective for all: rethinking the task. *Sch. Leader. Manag.* 32 (3), 1–17.
- Ainscow, M., 2016. Diversity and equity: a global education challenge, New Zealand. *J. Educ. Stud.* 51 (2), 143–155.
- Ainscow, M., 2020. Inclusion and equity in education: making sense of global challenges. *Prospects* 49, 123–134.
- Bassot, É., 2020. Ten Opportunities for Europe Post-coronavirus. Exploring Potential for Progress in EU Policy-Making. EPRS | European Parliamentary Research Service, Brussels.
- Cairns, R., 2020. Exams tested by COVID-19: an opportunity to rethink standardized senior secondary examinations. *Prospects*. <https://doi.org/10.1007/s11125-020-09515-9>.
- Cambridge Dictionary (no date) <https://dictionary.cambridge.org/dictionary/english/governance> Last accessed Feb 2022.
- Cerna, L., 2020. Coronavirus School Closures: What Do They Mean for Student Equity and Inclusion? <https://oecdutoday.com/coronavirus-school-closures-student-equity-inclusion/>.
- Commissioner for Human Rights, 2017. Fighting School Segregation in Europe Through Inclusive Education: A Position Paper. Council of Europe Publishing, Strasbourg.
- Council of Europe, 2020. Making the Right to Education a Reality in Times of COVID 19. A Roadmap for Action on the Council of Europe Education Response to COVID 19. Council of Europe Publishing, Strasbourg.
- Council of the European Union, 2011. Council Recommendation of 28 June 2011 on Policies to Reduce Early School Leaving 2011/C 191/01. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2011:191:0001:0006:en:PDF>.
- Council of the European Union, 2018a. Council Recommendation on Promoting Common Values, Inclusive Education and the European Dimension of Teaching. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0607\(01\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0607(01)&from=EN).
- Council of the European Union, 2018b. Council Conclusions on Moving Towards a Vision of a European Education Area. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018XG0607\(01\)&rid=5](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018XG0607(01)&rid=5).
- Cross-sectoral co-operation, ESPlus portal – <http://esplus.eu/glossary-phrase/cross-sectoral-cooperation> Last accessed February 2022.
- Di Pietro, G., Biagi, F., Costa, P., Karpinski, Z., Mazza, J., 2020. The Likely Impact of COVID-19 on Education. Reflections Based on the Existing Literature and Recent International Datasets. European Commission Joint Research Center Technical Report. Publications Office of the European Union, Luxembourg.
- Donaldson, G., 2015. Successful Futures. Independent Review of Curriculum and Assessment Arrangements in Wales. Available. <https://www.nasuwt.org.uk/asset/A788604C-3046-4005-A1EAOEAF023E0DD/>. (Accessed 5 September 2018).
- Donnelly, V.J. (Ed.), 2010. Inclusive Education in Action: Project Framework and Rationale. European Agency for Development in Special Needs Education, Odense, Denmark.
- Dorn, E., Hancock, B., Sarakatsnnis, J., Viruleg, E., 2020. COVID-19 and Learning Loss—Disparities Grow and Students Need Help. Mc Kinsey and Company, New York.
- Ebersold, S., Meijer, C.J.W., 2016. Financing inclusive education: policy challenges, issues and trends. In: Watkins, A., Meijer, C.J.W. (Eds.), *Implementing Inclusive Education: Issues in Bridging the Policy-Practice Gap (International Perspectives on Inclusive Education)*, vol. 8. Emerald Group Publishing Limited, Bingley.
- Ebersold, S., Óskarsdóttir, E., Watkins, A., 2020. Financing Mechanisms to Support Inclusive Practice. <https://doi.org/10.1093/acrefore/9780190264093.013.1259>.
- Ebersold, S., 2008. Adapting higher education to the needs of disabled students: developments, challenges and prospects in OECD member countries. In: *Higher Education to 2030*. OECD, Paris.
- Ebersold, S., 2016. Youth with disabilities and transition to adulthood. In: Devlieger, P., Rusch, F.R., Pfeiffer, D. (Eds.), *Rethinking Disability: The Emergence of New Definitions, Concepts and Communities*. Garant, Antwerpen-Apeldorn.
- European Agency for Special Needs and Inclusive Education, (2022). Country Policy Review and Analysis: Findings Report. (A. Watkins, V.J. Donnelly and S. Symeonidou, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education (2014) Organisation of Provision to Support Inclusive Education – Summary Report. Odense, Denmark: European Agency for Special Needs and Inclusive Education (Editors: Verity Donnelly and Mary Kyriazopoulou).
- European Agency for Special Needs and Inclusive Education, 2015. Agency Position on Inclusive Education Systems.
- European Agency for Special Needs and Inclusive Education, 2016a. Financing of Inclusive Education: Mapping Country Systems for Inclusive Education (S. Ebersold, Ed.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2016b. Financing of Inclusive Education: Background Information Report (A. Mangiaracina, Ed.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2016c. Raising the Achievement of All Learners in Inclusive Education—Literature Review (A. Kefallinou, Ed.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2018a. Financing Policies for Inclusive Education Systems: Resourcing Levers to Reduce Disparity in Education (S. Ebersold, E. Óskarsdóttir and A. Watkins, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2018b. Analysis Framework for Mapping Inclusive Education Policies (V. Soriano, A. Watkins, M. Kyriazopoulou, V. Donnelly, A. Kefallinou, S. Ebersold and G. Squires, eds.) Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2018c. Financing Policies for Inclusive Education Systems: Policy Guidance Framework (A. Watkins, E. Óskarsdóttir and S. Ebersold, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2018d. Country Policy Review and Analysis: Methodology Report—Revised 2018 (S. Symeonidou, V.J. Donnelly, V. Soriano and A. Watkins, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2018e. Key Actions for Raising Achievement: Guidance for Teachers and Leaders (V. Donnelly and A. Kefallinou, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2019. Changing Role of Specialist Provision in Supporting Inclusive Education: Mapping Specialist Provision Approaches in European Countries (S. Ebersold, M. Kyriazopoulou, A. Kefallinou and E. Rebollo Piriz, Eds.), Odense, Denmark.
- European Agency for Special Needs and Inclusive Education, 2020. European Agency Statistics on Inclusive Education: 2018 Dataset Cross-Country Report (J. Ramberg, A. Lénárt, and A. Watkins, Eds.), Odense, Denmark.

- European Commission, 2020. Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Achieving the European Education Area by 2025 (SWD 2020, 212 Final). European Commission, Brussels.
- European Commission/EACEA/Eurydice, 2014. Financing Schools in Europe: Mechanisms, Methods and Criteria in Public Funding. Eurydice Report. Publications Office of the European Union, Luxembourg. eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/170EN.pdf.
- European Union Agency for Fundamental Rights, 2020. Fundamental Rights Report 2020. Publications Office of the European Union.
- Florian, L., Black-Hawkins, K., Rouse, M., 2017. Achievement and Inclusion in Schools. Routledge, London.
- Holloway, J., Sørensen, T.B., Verger, A., 2017. Global perspectives on high-stakes teacher accountability policies: an introduction. *Educ. Pol. Anal. Arch.* 25 (85), 1–18. <https://doi.org/10.14507/epaa.25.3325>.
- Kaden, U., 2020. COVID-19 school closure-related changes to the professional life of a K–12 teacher. *Educ. Sci.* 10 (6), 165. <https://doi.org/10.3390/educsci10060165>.
- Kefallinou, A., Donnelly, V., 2016. Inclusive assessment: issues and challenges for policy and practice. In: Watkins, A., Meijer, C. (Eds.), *Implementing Inclusive Education: Issues in Bridging the Policy-Practice Gap*. Emerald, Bingley.
- Kellaghan, T., Greaney, V., 2020. Public Examinations Examined. World Bank, Washington, DC.
- Kelly, A., 2019. The High Stakes of Testing: Exploring Student Experience With Standardized Assessment Through Governmentality. Brill Sense, Leiden.
- Lingard, B., Thompson, G., Sellar, S., 2016. National Testing in Schools: An Australian Assessment. Routledge, London.
- Meijer, C.J.W., Watkins, A., 2019. Financing special needs and inclusive education—from Salamanca to the present. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2019.1623330>.
- Meijer, C.J.W. (Ed.), 1999. Financing of Special Needs Education. A Seventeen-Country Study of the Relationship Between Financing of Special Needs Education and Inclusion. European Agency for Development in Special Needs Education, Middelfart, Denmark.
- Meijer, C.J.W. (Ed.), 2003. Special Education Across Europe in 2003: Trends in Provision in 18 European Countries. European Agency for Development in Special Needs Education, Middelfart, Denmark.
- Muskin, J.A., 2015. Student Learning Assessment and the Curriculum: Issues and Implications for Policy, Design and Implementation. In-Progress Reflections No. 1 on Current and Critical Issues in the Curriculum and Learning. UNESCO, IBE.
- Norwich, B., 2013. Addressing Tensions and Dilemmas in Inclusive Education. *Living With Uncertainty*. Routledge, London.
- OECD, 1999. Inclusion at Work. OECD, Paris.
- OECD, 2004. Equity in Education. OECD, Paris.
- OECD (2005) <https://stats.oecd.org/glossary/detail.asp?ID=6189> Last accessed Feb 2022.
- OECD, 2012. Equity and Quality in Education: Supporting Disadvantaged Students and Schools. OECD Publishing, Paris.
- OECD, 2017a. The Funding of School Education: Connecting Resources and Learning. OECD Publishing, Paris. <http://www.oecd.org/governance/the-funding-of-school-education-9789264276147-en.htm>.
- OECD, 2017b. PISA 2015 Results: Students' Well-Being. OECD, Paris. https://read.oecd-ilibrary.org/education/pisa-2015-results-volume-iii_97892-64273-856-en.
- OECD (2019) Future of Education and Skills 2030 Conceptual learning framework. Student Agency. https://www.oecd.org/education/2030-project/teaching-and-learning/learning/student-agency/Student_Agency_for_2030_concept_note.pdf.
- OECD, 2020. Learning remotely when schools close: how well are students and schools prepared? Insights from PISA. In: OECD Policy Responses to Coronavirus (COVID-19). Retrieved from. <https://www.oecd.org/coronavirus/policy-responses/learning-remotely-when-schools-close-how-well-are-students-and-schools-prepared-insights-from-pisa-3bfda177/>.
- Opertti, R., Zachary, W., Zhang, Y., 2014. Inclusive education: from targeting groups and schools to achieving quality education as the core of EFA. In: Florian, L. (Ed.), *The SAGE Handbook of Special Education*, second ed., vol. 1. SAGE, London.
- Oxford Bibliographies (no date) <https://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0012.xml> Last accessed Feb 2022.
- Oxford Dictionaries (no date) <https://www.lexico.com/definition/funding> Last accessed Feb 2022.
- Parrish, T.B., 2001. Who's paying the rising cost of special education? *J. Spec. Educ. Leader* 14 (1), 4–12.
- Parrish, T.B., 2014. How should we pay for special education? In: Bateman, B., Lloyd, J.W., Tankersley, M. (Eds.), *Enduring Issues in Special Education: Personal Perspectives*. Routledge, New York.
- Ramberg, J., Watkins, A., 2020. Exploring inclusive education across Europe: some insights from the European Agency statistics on inclusive education. *Fire* 6 (1), 85–101.
- Redecker, C., 2017. European Framework for the Digital Competence of Educators: DigCompEdu. JRC Working Papers JRC107466. Joint Research Center, Seville.
- Rimfeld, K., Malanchini, M., Hannigan, L.J., Dale, P.S., Allen, R., Hart, S., Plomin, R., 2019. Teacher assessments during compulsory education are as reliable, stable and heritable as standardized test scores. *JCPP* 60 (12), 1278–1288.
- Rimfeld, K., Malanchini, M., Plomin, R., 2020. Don't worry about cancelled exams: research shows we should switch to teacher assessment permanently. *Conversation*. <https://theconversation.com/dont-worry-about-cancelled-exams-research-shows-we-should-switch-to-teacher-assessment-permanently-134709>.
- Sahlberg, P., 2020. Does the pandemic help us make education more equitable? *Educ. Res. Pol. Pract.* <https://doi.org/10.1007/s10671-020-09284-4>.
- Schleicher, A., 2020. How Can Teachers and School Systems Respond to the COVID-19 Pandemic? Some Lessons From TALIS. <https://oecdeditoday.com/how-teachers-school-systems-respond-coronavirus-talis/>.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Routledge, Abingdon.
- Slee, R., 2018. Defining the Scope of Inclusive Education. Paper Commissioned for the 2020 Global Education Monitoring Report, Inclusion and Education. UNESCO, Paris.
- UNESCO, 2009. Policy Guidelines on Inclusion in Education. UNESCO, Paris.
- UNESCO, 2017. A Guide for Ensuring Inclusion and Equity in Education. UNESCO, Paris. <http://unesdoc.unesco.org/images/0024/002482/248254e.pdf>.
- UNESCO, 2020a. Towards Inclusion in Education: Status, Trends and Challenges. The UNESCO Salamanca Statement 25 Years On. UNESCO, Paris. <https://unesdoc.unesco.org/ark:/48223/pf000-03742-46>.
- UNESCO, 2020b. Inclusion and Education: All Means All. Global Monitoring Report (GEMR). UNESCO, Paris. <https://en.unesco.org/gem-report/report/2020/inclusion>.
- UNICEF, 2015. The Investment Case for Education and Equity. UNICEF, New York, NY.
- United Nations, 2006. United Nations Convention on the Rights of Persons With Disabilities. United Nations, New York.
- United Nations, 2020. Policy Brief: Education During COVID-19 and Beyond. UN, New York, NY. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf.
- Van Lacker, W., Parolin, Z., 2020. COVID-19, school closures and child poverty; a social crisis in the making. *Lancet*. [https://doi.org/10.1016/S2468-2667\(20\)30084-0](https://doi.org/10.1016/S2468-2667(20)30084-0). Published Online April 7, 2020.
- Volante, L., 2017. *The PISA Effect on Global Educational Governance*. Routledge, London.
- Watkins, A., 2017. Inclusive education and European educational policy. In: *Oxford Research Encyclopaedia, Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.153>. education.oxfordre.com.
- World Bank Group, 2020. The Impact of the COVID-19 Pandemic on Education Financing. World Bank, Washington, DC.
- Ydo, Y., 2020. Inclusive education: global priority, collective responsibility. *Prospects* 49, 97–101. <https://doi.org/10.1007/s11125-020-09520-y>.

Further reading reports

- European Agency for Special Needs and Inclusive Education, 2018b. Financing Policies for Inclusive Education Systems: Resourcing Levers to Reduce Disparity In Education (S. Ebersold, E. Óskarsdóttir and A. Watkins, Eds.), Odense, Denmark. <https://www.european-agency.org/resources/publications/fpies-synthesis-report>.
- European Agency for Special Needs and Inclusive Education, 2016. Financing of Inclusive Education: Mapping Country Systems for Inclusive Education (S. Ebersold, Ed.), Odense, Denmark. <https://www.european-agency.org/resources/publications/financing-inclusive-education-mapping-country-systems-inclusive-education>.
- European Agency for Special Needs and Inclusive Education, 2018a. Financing Policies for Inclusive Education Systems: Policy Guidance Framework (A. Watkins, E. Óskarsdóttir and S. Ebersold, Eds.), Odense, Denmark. <https://www.european-agency.org/resources/publications/financing-policies-inclusive-education-systems-policy-guidance-framework>.
- European Commission/EACEA/Eurydice, 2016. Structural Indicators for Monitoring Education and Training Systems in Europe—2016. Eurydice Background Report to the Education and Training Monitor 2016. Eurydice Report. Publications Office of the European Union, Luxembourg. https://eacea.ec.europa.eu/national-policies/eurydice/content/structural-indicators-monitoring-education-and-training-systems-europe-2016_en.
- OECD, 2017. The Funding of School Education: Connecting Resources and Learning. OECD Publishing, Paris. https://www.oecd-ilibrary.org/education/the-funding-of-school-education_9789264276147-en.
- UNESCO, 2020. Global Education Monitoring Report 2020: Inclusion and Education: All Means All. UNESCO, Paris. <https://en.unesco.org/gem-report/report/2020/inclusion>.

Articles

- Ebersold, S., Óskarsdóttir, E., Watkins, A., 2020. Financing mechanisms to support inclusive practice. In: Oxford Research Encyclopedia of Education. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.1259>.
- Watkins, A., Ebersold, S., 2016. Efficiency, effectiveness and equity within inclusive education systems. In: Watkins, A., Meijer, C. (Eds.), International Perspectives on Inclusive Education: Issues in Bridging the Policy-Practice Gap (Implementing Inclusive Education), vol. 8. Emerald Group Publishing Limited, Bingley.

Relevant websites

UNESCO/UIS Education Finance

<http://uis.unesco.org/en/topic/education-finance>, The UNESCO Institute for Statistics works directly with countries to compile education expenditure data through an annual survey and helps to improve national statistical systems and adopt new methodologies, such as National Education Accounts (NEAs). Through this country-level support, the UIS is also seeking to reduce data gaps and provide a more complete picture of education spending. With good quality data, governments and donors can see trends in the overall funding of education. They can also see the extent to which different levels of government and families are contributing to education. In short, they see the full picture in striving to provide quality education for all.

The Governance of School Education Systems

<https://www.schooleducationgateway.eu/en/pub/resources/governance-of-school-edu.htm>, European Union education ministers have highlighted the need for modern approaches to teaching and learning and to the governance of school education systems in order to help schools respond to the changing educational demands of learners and society. In 2016, the Education and Training 2020 Working Group on Schools, made up of representatives from European education ministries and stakeholder organizations, was asked to critically reflect on the governance of school education and how systems can address the serious challenges of quality and equity in order to strive for better learning outcomes for all young people. As part of the European online School Education Gateway, a range of European-level resources for improving school governance is presented.

OECD School Resources Review

<http://www.oecd.org/education/school-resources-review/>, The School Resources Review aims to help countries make resource decisions that support quality, equity and efficiency in school education. The Review provides country-specific and comparative analysis on the use of financial, physical and human resources in school systems. It offers policy advice on how to govern, distribute and manage resources so that they contribute to achieving countries' educational objectives.

Financing of Inclusive Education—Mapping Country Systems for Inclusive Education

<https://www.european-agency.org/projects/financing-inclusive-education-mapping-country-systems-inclusive-education>, The Financing of Inclusive Education—Mapping Country Systems for Inclusive Education project activities focus on information collection via desk research and country data gathering in order to provide an update on the previous Agency work as well as build on the findings of other Agency and international work. The key question for the project was: what are the critical factors within financing that support learners' rights to inclusive education.

Financing Policies for Inclusive Education Systems

<https://www.european-agency.org/projects/financing-policies-inclusive-education-systems>, Policy-makers across Europe recognise that funding mechanisms are a critical lever in reducing disparity in education. They require more detailed information on the impact of funding mechanisms on inclusive education that can be used to guide their policy developments. The Financing Policies for Inclusive Education Systems (FPIES) project is a response to this identified policy need.

Country Policy Review and Analysis

<https://www.european-agency.org/projects/country-policy-review-and-analysis>, The Country Policy Review and Analysis (CPRA) work supports reflection regarding the development of country policies for inclusive education. CPRA aims to act as a tool for stimulating discussion in the country concerned. Its central focus is to analyse the available information about current country policy for inclusive education; the CPRA work does not in any way address the actual implementation of the policy being considered.

Inclusive Education in Action

<https://www.inclusive-education-in-action.org/>, IEA is a joint initiative by UNESCO and the Agency. It aims to be a unique online global resource base to facilitate information-sharing and enhance knowledge and understanding of inclusive education. It provides a comprehensive resource and knowledge base of documentary resources and in-depth case studies on inclusive approaches to legislation, policies, programs and practices from around the globe.

Pacific Indicators for Disability Inclusive Education (Pacific INDIE) Project

<https://www.monash.edu/education/research/projects/pacific-indie>, This international project aims to develop a set of contextually specific indicators for disability-inclusive education in the Pacific and guidelines for implementation. The project's goal was to support the implementation of disability-inclusive education in the Pacific Islands and, specifically, to work collaboratively with local and international partners. The indicators were developed to assist countries to evaluate their efforts and develop further plans and targets for providing quality education for children and young people with disabilities. Indicators and guidelines were developed to ensure three strategic objectives of access, quality, and effectiveness and efficiency.

Economics of inclusive education

Kamal Lamichhane, Disability Studies and Inclusive Education, Graduate School of Comprehensive Human Sciences, The University of Tsukuba, Tsukuba, Japan

© 2023 Elsevier Ltd. All rights reserved.

Disability and rate of returns on the investment in education	51
Disability biases and change in economic household decision in education	52
Poverty and additional determinants of schooling	54
Intervention strategies to overcome barriers	54
Disability related extra costs in education	55
Methodological challenges in evaluating the cost benefits of inclusive education	55
References	57

Disability and rate of returns on the investment in education

Although there are debates on what kind of education is needed for contemporary society, there is a common consensus on the need for greater investment to enjoy multifaceted benefits of education. Numerous studies have examined the cost of educating our children; simply put, the economics of education. In short, economic issues in education such as the demand for education, the financing of education, and the comparative efficiency of various educational programs and policies are studied in economics of education. Links between schooling and labor market outcomes for individuals is one of the central components for economics of education (Weil, 2009).

This consensus about the need for greater investment in education is not universally apparent when considering the education of individuals with disabilities. The prevailing assumption is, still, that even if persons with disabilities are educated, they are less likely to make use of the education, or that they will not be useful in the workforce. Though still rare, some studies have started empirically challenging this assumption in lower and middle income countries (Lamichhane and Watanabe, 2015; Mori and Yamagata, 2009; Lamichhane and Sawada, 2013). Adopting an economic analysis, using primary data collected by myself and utilizing secondary data in Nepal, I published a paper with Yasuyuki Sawada (Lamichhane and Sawada, 2013) to build a compelling justification for why Governments and other relevant authorities must increase their investment in the education of disabled children. Our study was the first to estimate the returns to the investment, commonly referred to as *yield*, in education for people with disabilities in developing countries. It gained wide attention from different agencies including the World Bank. A novel feature of this article is its use of micro-data collected on the timing of being impaired during school-age years as a means for identifying instrumental variables for years of schooling.

The data and evidence in this area reveal that after controlling for endogeneity bias arising from schooling decisions as well as sample selection bias due to endogenous labor participation, the estimated rate of returns to investment in education was found to be very high among persons with disabilities, ranging from 19.3% to 25.6% (see Table 1: Disability and returns to the investment in education (Nepal); also, Lamichhane and Sawada, 2013 for details). This yield is two or three times higher compared with roughly 10% for those without disabilities in different regions of the world (see Fig. 1: Global comparison of returns to education). One important point to mention is that although the rate of higher returns to the investment in education for PwDs is observed, it should not be interpreted in a way that PwDs out earn people without disabilities. As educational attainment correlates with the skill level of the occupations that people are able to enter, the higher returns for PwDs could be interpreted as the possibility of increasing wage earning more from what they would have been without education.

Regardless of disability status, poverty is one of the severe issues affecting the lives of individuals especially in low- and middle-income countries. Disabled people have a far higher likelihood of suffering both extreme and relative poverty compared to those without disabilities largely because they are less invested in education by families, communities and the Governments. Keeping discrimination, negative employer and community attitudes and erstwhile obstructions to access aside, it is significant that unemployment is so high even where the requisite education and skills for undertaking the work are low. Consequently, not only disabled people but also their families live in severe financial hardship. However, as stated here, multiple studies for multiple countries have confirmed that investing in education substantially reduces their risk of falling into the poverty trap. Research from Nepal (Lamichhane et al., 2014) demonstrated that the higher the education levels, the lower the likelihood of disabled people falling into poverty. This was even so for minority ethnic groups. Studies also suggest that for disabled individuals, like general population groups, the longer the years of schooling, the higher the probability of getting full-time white-collar jobs compared to those with less or no education (Mizunoya et al., 2018; Lamichhane, 2012, 2015; Lamichhane and Okubo, 2014; Mitra and Sambamoorthi, 2009). It is also important to mention that if proper support is made available, issue of productivity differential cannot be considered as an excuse for not hiring qualified disabled individuals.

Table 1 Disability and returns to the investment in education (Nepal).

	(1) <i>OLS</i>	(2) <i>OLS</i>	(3) <i>Tobit</i>	(4) <i>Tobit</i>	(5) <i>IV-Tobit</i>	(6) <i>IV-Tobit</i>
Years of schooling ^a	0.059 (0.027)*	0.065 (0.030)*	0.213 (0.062)***	0.193 (0.067)**	0.256 (0.103)*	0.248 (0.091)**
Dummy = 1 if the follow up survey	–0.704 (1.419)	–0.493 (1.385)	–3.935 (4.067)	–3.166 (4.004)	–2.084 (2.297)	–2.079 (2.233)
Severity of impairments	–0.226 (0.657)	–0.155 (0.643)	–0.538 (1.894)	–0.205 (1.865)	–0.017 (1.065)	–0.025 (1.035)
School Leaving Certificate (SLC) score	0.011 (0.02)	0.008 (0.021)	0.019 (0.044)	–0.014 (0.044)	–0.01 (0.025)	–0.01 (0.024)
Score of tests given during the survey	–0.389 (0.433)	–0.327 (0.429)	–0.941 (0.975)	–0.527 (0.964)	–0.543 (0.565)	–0.535 (0.548)
Years of work experience	–0.014 (0.06)	–0.011 (0.059)	0.09 (0.126)	0.124 (0.124)	0.082 (0.065)	0.081 (0.063)
Years of work experience squared	0.000 (0.001)	0.000 (0.001)	0.000 (0.002)	0.000 (0.002)	0.000 (0.001)	0.000 (0.001)
Dummy = 1 if female	–0.374 (0.265)	–0.314 (0.25)	–0.471 (0.557)	–0.38 (0.566)	–0.198 (0.322)	–0.202 (0.313)
Age	–0.021 (0.144)	–0.015 (0.146)	0.357 (0.239)	0.373 (0.236)	0.12 (0.119)	0.123 (0.114)
Age squared	0.001 (0.002)	0.000 (0.002)	–0.005 (0.004)	–0.006 (0.004)	–0.002 (0.002)	–0.002 (0.002)
Dummy = 1 if full-time worker	0.062 (0.273)	–0.026 (0.275)	7.645 (0.587)***	7.488 (0.609)***	4.42 (0.407)***	4.442 (0.382)***
Dummy = 1 if hearing impaired		–0.086 (0.276)		–1.98 (0.719)**	–0.993 (0.513)	–1.021 (0.480)*
Dummy = 1 if physically impaired		–0.479 (0.388)		–2.083 (0.728)**	–1.763 (0.441)**	–1.75 (0.424)***
Constant	3.693 (2.725)	3.514 (2.68)	–11.678 (5.620)*	–11.013 (5.540)*	–4.877 (3.01)	–4.839 (2.922)
<i>R</i> -Squared	0.073	0.086				
Number of observations	222	222	398	398	373	373

^aEndogenous variable.

*Significant at 10%.

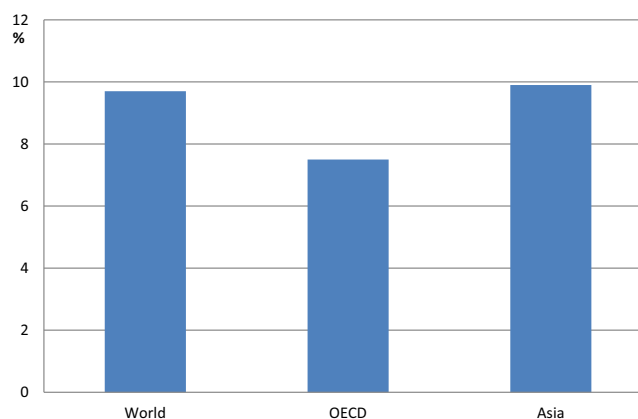
**Significant at 5%.

***Significant at 1%.

Source: Lamichhane and Sawada (2013).

Disability biases and change in economic household decision in education

Despite clearly measurable multifaceted benefits of education, and impactful yet unmeasurable positive effects, people with disabilities are deprived of first access to education itself and more clearly to receiving quality education due to several interconnected yet

**Fig. 1** Global comparison of returns to education source: Psacharopoulos and Patrinos (2004).

poorly understood factors. Recent studies indicate factors such as a fundamental misunderstanding of disability, and that poverty of the households is linked to household investment disparity between disabled and nondisabled children. Such investment biases obstruct access to and continuation of quality education for disabled children. When the resources of the family are limited, especially if people with disabilities are the head of the household, their impact on educating their disabled children seems small. Some of the earlier studies have supported this claim. Hoogeveen (2004) noted a significant education deficit in children in households in Uganda headed by disabled persons, linking it with limited or reduced ability of the household to afford school fees due to the direct costs of disability. On the one hand, poor families with a disabled family-member struggle with social stigma in the community and on the other hand, poverty may prevent them from investing in their children's education.

While there are intense debates surrounding the lower school enrollment and educational attainment of disabled people, except for Rosales-Rueda's (2014) research on children with mental health conditions in the USA and Lamichhane and Takeda's (2022) research on Bangladesh, studies are silent on household economic behavior particularly the intrahousehold resource allocation for education for children with disabilities. That said, it is important to note that some of the earlier studies have discussed discriminatory attitudes, parent's financial constraints and institutional barriers as some of the plausible reasons for the lower level of schooling of children with disabilities (Lamichhane, 2013, 2015; Lamichhane and Kawakatsu, 2015; Mizunoya et al., 2018; Takeda and Lamichhane, 2018). Questions arise:

- are these the only barriers to their education?
- Are there additional barriers and biases for intrahousehold resource allocation?
- Why are studies on household economic behavior in education rare or unavailable?
- Is this because research on this topic is considered less important for policy making?
- Is it simply because this sector is overlooked given the dearth of relevant and valid datasets?

On the one hand, the lack of rigorous studies directly causes gaps in education between people with and without disabilities and on the other hand, evidence-based policy making would be negatively affected and implications for the education of children with disabilities continue to be obscured and overlooked.

As their preliminary attempt, Lamichhane and Takeda (2022) conducted a study in Bangladesh to see whether there exists any disability bias in intrahousehold investment in education. In case of the existence of biases, their additional research question sought to locate such bias, either in the stage of enrollment or thereafter and identifying key factors influencing parental investment decisions in the education of children with disabilities. Despite growing attention regarding the quality of education for children with disabilities, Lamichhane and Takeda (2022) found bias in household investment in education mainly in the stage of enrollment decision-making. Moreover, disability bias is found in both primary and secondary levels, indicating that such bias exists in the stage of providing basic education (see Table 2: Difference of Educational Expenditure by Disability Status (Bangladesh); Lamichhane and Takeda, 2022).

According to Table 2: Difference of Educational Expenditure by Disability Status (Bangladesh), the sample that includes out-of-school children suggests that there is a significant difference in educational expenditure at the primary level and secondary level statistically at 5%. On the other hand, there is little difference in educational investment by disability status for enrolled children. The significant difference is detected at 10% level for Secondary age children. They have attributed this to school enrollment of children with disabilities costing more than for children without disabilities given inaccessibility of infrastructure for many children with disabilities, the absence of appropriate support systems. However, once enrolled, the results revealed only small differences

Table 2 Difference of educational expenditure by disability status (Bangladesh).

		Mean difference		
	Non-children with disabilities	Children with disabilities	(Non-children with disabilities–children with disabilities)	t-value
Educational expenditure for both enrolled and not enrolled children				
Primary level	1988.33	1015.08	973.24**	2.56
Secondary level	4190.56	2894.42	1296.14***	2.46
Higher level	5347.69	5154.21	193.48	0.13
Educational expenditure only for enrolled children				
Primary level	2574.50	1682.53	891.97*	1.66
Secondary level	4961.95	4475.4	486.55	0.71
Higher level	12,074.42	13,588.36	–1513.95	–0.46

Note. Each value is calculated by author using dataset from Bangladesh Household Income Expenditure Survey 2010 (HIES). Expenditure is measured in Taka. Gap by disability status is statistically significant.

*10% Significance level.

**5% Significance level.

***1% Significance level. The definition of Children with disabilities is children with disabilities aged below the age of 20 who acquire impairments such as visual, hearing. Physical, cognitive and communication.

Source: Lamichhane and Takeda (2022).

in educational expenditure. Parents' expectation on returns from education of children with disabilities is much lower than that of children without disabilities (i.e., expected contribution to household income is lower).

Again, it is important to mention that the education of children with disabilities may be partly driven by shifts in parental opportunistic investment strategies that likely prioritize non-disabled children. By investing heavily in children without disabilities, parents seem to attempt to provide them with a competitive advantage in the acquisition of both resources and friends. However, wealthier and economically stable parents opt for a more opportunistic strategy of educational investment in the required level of education that does not discriminate between offspring regardless of disability status.

Poverty and additional determinants of schooling

Using nationally representative data from various Asian and African countries, we have gained results that household poverty is one of the serious issues negatively affecting the education of disabled people. From the 2005 India Human Development Survey (IHDS), [Takeda and Lamichhane \(2018\)](#) examined the determinants of children's school progression and academic achievements. One of the important contributions of the IHDS is that it attempted to identify respondents' disability status with the adaptation of questions from the Washington Group on Disability Statistics. The questions mostly focused on activities of daily living, interrogating difficulties that respondents faced in seeing, hearing, walking, self-care, and speaking. However, they lacked a question on remembering and concentration. In our study, a person is considered to have a disability if he or she has difficulty in any of the domains. Results revealed important dynamics of disability and income, both negative in educational attainment, completion, dropout and academic achievements. The effect of father's education was small when it interacted with children's disability, whereas mother's education had a positive association in secondary school completion as well as with reading and writing skills. Similarly, using the nationally representative dataset of Bangladesh and utilizing the fixed effects and logit model techniques, we gained similar findings to India ([Lamichhane and Kawakatsu, 2015](#)). We examined the determinants of school participation between children with and without disabilities. In the analysis of total samples, it was observed that while disability is negatively correlated with school participation, once the sample was restricted to just disability, results showed that together with the household head's higher educational level, monthly expenditure and working age members are positively correlated with the probability of school participation.

Furthermore, despite initiatives such as Universal Primary Education (UPE) policy, a study carried out in Uganda ([Lamichhane and Tsujimoto, 2017](#)) found that disabled children's education is especially compromised in households experiencing poverty. More specifically, our findings indicate that those not completing primary school are from poor families. Our findings that disabled individuals who have not benefited from UPE can be attributed to several factors. As UPE is supposed to simply waive the tuition fees and not the additional cost required to accommodate the needs arising from disability; for example, making school infrastructure and roads from home to school accessible together with providing alternative learning resources including textbooks in Braille or assistive devices for different impairment groups, we are in fact unsurprized not to see the desired impact of UPE especially in poor households. Though not specifically researching disability and education, while reviewing UPE situation in Sub-Saharan Africa (SSA), [Bennell's, 2021](#) argument that only those sections of the poor who do not need additional support for primary education could have benefited from UPE also supports our finding that disabled children who might require additional support beyond tuition fee waiver are not positively affected by UPE in Uganda. Though not explicitly concerned with disability, [Kakuba et al. \(2021\)](#) found that only 22% of children aged 13–18 were enrolled in secondary school, and that demand and supply related predictors are moderated by the socio-economic status of households to influence schooling outcomes. Government was to provide free tuition to secondary school students starting with 300,000 primary school graduates in 2007. Parents, on the other hand, were to provide exercise books, accommodation, medical care, meals and other resources. In such circumstances, those in poverty would experience difficulties sending their children to school and obviously its consequences would be larger for disabled children with parents' opportunistic investment strategies in education coming to the foreground.

Intervention strategies to overcome barriers

In order to overcome barriers to education arising from household poverty level, interventions such as *conditional cash transfers* (CCT) and lunch in school are implemented as an effective strategy toward increasing access to education for the general population in various parts of the world. In some countries, programs such as providing lunch to the children in school are also offered to increase girls' access to education. As will be discussed in the subsequent section, the cost for education in general and for disabled children has fundamental differences. Therefore, other than conventional approaches and schemes such as CCT, additional strategies geared for programs with targeted subsidies to increase access to education by disabled children is required. Inaccessibility of infrastructure, including the proximity from home to school, requires additional investment to make them accessible to all. While quality basic education lies at the heart of inclusive development, and access to quality basic education is an ambitious goal to be achieved by 2030, it cannot succeed without identifying the difficulties children with disabilities and their parents face. Policies addressing disability-based discrimination, accessibility of the school environment and social protection schemes are necessary to improve the education of children with disabilities. Pursuant to providing quality education, 2015 Sustainable Development

Goal (SDG) 4 has clearly expressed commitment for quality lifelong education and SDG 4.5 emphasizes the accommodation of the needs of educating disabled people.

Disability related extra costs in education

The goal of a global education program is to equalize the ability of all to participate in society regardless of any differences and functional limitations which could act as a disability when efforts fail to accommodate needs related to impairment. Though all people incur costs for education, such costs may be significantly greater for those with disabilities. [Mont and Cote \(2020\)](#) claim that if all people have the full costs of living with a disability covered, they are on a more equal footing and more likely to be full members of the community. One of the issues that is sometimes difficult to grasp for policy makers and advocates alike is the cost required and the range of difference of costs based on type and severity of disability. The casual observation is that for most people with significant disabilities, extra cost would be required to cover pedagogy, extra assistant teachers, devices and educational materials. Sadly, allocated budget provided in low- and middle-income countries rarely covers basic consumption, let alone disability related costs that many consider to be higher than the measurable benefits resulting from education, especially for those with high support requirements.

In addition, in most low- and middle-income countries, there is less or no distinction between disability related extra cost and the comprehensive impact of education which should be beyond the benefit of rate of returns on education. For many, the measurement of benefit is restricted to wages. The traditional notion of understanding cost benefit analysis focusing only on wage earning disregarding other externalities attached to education is simply insufficient especially for marginalized people. For them, their inclusion in community carries greater significance. In the absence of complementary measures, resulting in part from not having a proper understanding of disability and education citing the scarcity of resources, budget allocated to disability and education generally only supports basic access but does not seriously emphasize the quality and continuation of an education. Nevertheless, considering the wide-ranging benefits of education, countries should not avoid confronting the need for greater investment in education for persons with disabilities.

Intra-household dynamics that could lead to a person with disability being deprived of education has several factors to consider especially in a poor household. Considering the diversity of disability related costs, [Mont and Cote \(2020\)](#) acknowledge that a complicating factor is that not all households with disabilities face the same extra costs. The size of extra costs depends on a variety of factors including not only the type and degree of disability, but also the local environment. Another complicating factor is that the extra costs of education can vary based on whether a school is well equipped and furnished or not and the extent to which their facilities are designed to meet the diverse needs of learners regardless of disability status. If the school has failed to do so, then they must be prepared for necessary investment to make their school and education program accessible.

Nepal initially implemented inclusive education as a form of integrated education. They are yet to identify the cost difference between special and inclusive education. Countries such as Nepal, India, Bangladesh and Uganda, have different levels of support for different levels of disability. This assumes that people with severe disabilities require adjustments in pedagogy and teachers with special training; both requiring higher levels of investment.

Instead of acknowledging the essence of inclusive education:

- achieving excellent education; and
- achieving community inclusion and economic independence,

many low- and middle-income countries have the incorrect understanding that inclusive education is simply cutting resources to provide opportunity for education. Consequently, policy makers and relevant authorities often overlook the investment required to be fully prepared for making schools inclusive with teaching and learning activities addressing the individualized educational needs of all learners.

Methodological challenges in evaluating the cost benefits of inclusive education

In education, one of the issues that often draws mixed attention is its cost. Many economists focus on cost-benefit analyses of education. Though cost is important, it is equally important to mention that all components should be considered. First and foremost, education is the fundamental right of all, regardless of disability. In relation to under-served population groups in particular, positive discussion of cost is often overlooked. Though cost plays an important role, it is equally noteworthy that all states have an obligation to provide basic education at least up to high school level education. This being the case, it is also understandable that the Governments and other relevant authorities in education regularly must make difficult decisions about how best to invest resources to serve their citizens. Yet systematic analysis of the relationship between costs and outcomes of education for people with disabilities is not common in disability studies in education research. This type of evaluation is a fundamental tool for policy makers to decide on the fair allocation of resources.

While discussing cost-benefit, it should consider factors such as economic rate of return (ERR) of education and the benefit-cost ratio. Sadly, discussion around these issues is at best weak or for the most part missing in disability studies in education. While emphasizing returns to the investment, one should be cautious that as education has multifaceted benefits which economists regard

as other externalities, simply measuring education from the viewpoint of rate of economic returns makes it more challenging to estimate cost benefit. Here the positive externalities refer to the benefits other than direct income gained because of the level of education. Enabling people with disabilities to acquire higher levels of education increases their presence in community, raises not only his or her output but also the output of those around him or her. Educated workers can bring new technologies, methods and information to the attention of others. They can teach things to others. The positive externalities of education include the effects of personal networks and the roles educated workers play in them (Münich and Psacharopoulos, 2018)

Other than the direct economic benefits of education, Lamichhane (2012) argued that disabled individuals employed in the workforce can socialize with their colleagues inside and beyond the workplace. The analysis of cost benefit could not capture this return on investment fairly using conventional approaches. Making friends and socializing is a basic human need. Widening one's social horizons leads to a deeper sense of inclusion. When employed persons with disabilities socialize with their colleagues, their increased visibility in the community contributes to reducing prejudice, misunderstanding and discrimination. Such impacts must not be overlooked while calculating the cost for years of schooling and cognitive skills only. If educated disabled individuals acquire skills required for the jobs and are employed, their families can mutually enjoy a sense of positive interdependence, rather than a dynamic of dependency and provision (Lamichhane, 2012).

Therefore, investment in the provision of infrastructure associated with inclusive education is indeed important. According to Levin and McEwan (2001) several factors including:

- **Personnel**—teachers, support staff, aides, administrators, cafeteria workers, maintenance staff, counselors, mentors, coaches, consultants, special education service providers, curriculum supervisors, union representatives, and those providing various services from the district.
- **Facilities**—classrooms, computer labs, shared spaces, buildings, grounds, heating/cooling, lighting, repairs, maintenance, and cleaning.
- **Equipment and Materials**—desks, chairs, white boards, curricular materials, information technology, books, and recreational equipment.
- **Other program inputs**—data protection, data analysts, program evaluators, consultants, project managers.

All required direct cost and that can be somehow estimated by gathering information on market prices (e.g., the hourly cost of personnel to implement a program or construct a database). However, there are many other less obvious resources such as “donated” or additional teacher time (i.e., the extra time teachers spend working outside their contract hours), parental involvement, and opportunity costs which should be equally considered as their value is typically outside of the economists’ calculations. These issues cannot be properly estimated in monetary gains or losses and are therefore complicated to analyze. Also, as their impact is often relatively invisible compared to the other benefits, they are often ignored in narrow cost-benefit analysis. Understanding how to include different intervention strategies such as making school buildings accessible is necessary. Accessible buildings are not just for individuals with disabilities and should not be considered as *additional* costs. It is likely that countries count the budget needed to make schools and infrastructure accessible as a budget for disabled people and thus they conclude that the cost for educating disabled children is significantly higher compared than for the general population.

Although it is difficult to calculate the exact cost benefit of educating disabled individuals, it is possible to measure progress on “inclusive quality education” and its impact on the lives of disabled individuals—their families and communities. First to measure the effect of education between those with and without disability and compare it internationally, we can use the Washington Group on Disability Statistics (WG) WG/UNICEF Child Functioning Module together with WG/UNICEF Inclusive Education Module. Measuring the impact of education or let's say inclusive education on the lives of disabled people, their family and the society itself is important. Building reliable datasets on disability is still a challenge, however some countries have started incorporating questions on disability into their planning cycles which will help us to undertake further studies in this area. We also need to continue lobbying with bilateral and multilateral agencies together with Governments to explicitly include disability related components in all the surveys so that research on various issues of disability can be undertaken and results can be valid.

Although providing quality education with a cost minimization approach is necessary especially in the resource poor countries, citing disability as additional costs to be avoided diminishes the returns from human capital benefits. The accumulation of human capital requires investments in disability. Studies on disability and the nexus of education and employment have indicated that undertaking education is investment in the acquisition of skills and knowledge that will increase earnings providing long-term benefits such as job stability, securing white collar jobs, economic independence and lifting people out of poverty. Despite the long-term benefits of quality education and principally, there being a consensus that education should be a top priority, the education of children with disabilities in low- and middle-income countries was compromised: it is funded by NGOs/INGOs even though in most countries school education is predominantly financed and provided by governments. Departing from NGOs/INGOs funded education of disabled people, in recent years, Governments have started taking responsibility for educating their citizens with disabilities and have increased funding for inclusive or special education.

Because of the collective efforts of Governments guided by several international initiatives such as Education for All (EFA), and respectively goal 2 and 4 of both Millennium Development Goals (MDGs) for 2000–15 and Sustainable Development Goals (SDGs) for 2015–30 for education, we have witnessed improvement for access to education of disabled people in recent years. However, research reveals that simply providing financial support for teachers’ salary and textbooks alone is inadequate. Governments, bilateral and multilateral donor agencies therefore need to increase their support, covering expenses required for the proper implementation of inclusive education considering the reasonable accommodations that disabled children require together with

resolving accessibility issues. Research reveals that simply providing financial support for teachers' salary and textbooks alone is inadequate. Governments, bilateral and multilateral donor agencies need to increase their support, covering expenses required for the proper implementation of inclusive education.

References

- Bennell, P., 2021. The political economy of attaining Universal Primary Education in sub-Saharan Africa: the politics of UPE implementation. *Int. J. Educ. Dev.* 80, 102317.
- Hoogeveen, J.G., 2004. Measuring welfare for small but vulnerable groups poverty and disability in Uganda. In: Social Protection Unit Human Development Network the World Bank, Social Protection Discussion Paper Series, No.0419.
- Kakuba, C., Nzabona, A., Asimwe, J.B., Tuyiragize, R., Mushomi, J., 2021. Who accesses secondary schooling in Uganda; was the Universal Secondary Education policy ubiquitously effective? *Int. J. Educ. Dev.* 83, 102370.
- Lamichhane, K., 2012. Employment situation and life changes for people with disabilities: evidence from Nepal. *Disabil. Soc.* 27 (4), 471–485.
- Lamichhane, K., 2013. Disability and barriers to education: evidence from Nepal. *Scand. J. Disabil. Res.* 15 (4), 311–324.
- Lamichhane, K., Ballabha, P.D., Kartika, D., 2014. Analysis of poverty between people with and without disabilities in Nepal. In: JICA Research Institute Working Paper, No.77.
- Lamichhane, K., 2015. Disability, Education and Employment in Developing Countries: From Charity to Investment. Cambridge University Press, New Delhi, India.
- Lamichhane, K., Okubo, T., 2014. The nexus between disability, education, and employment: evidence from Nepal. *Oxf. Dev. Stud.* 42 (3), 439–453.
- Lamichhane, K., Watanabe, T., 2015. The Effect of Disability and Gender on Returns to the Investment in Education: A Case from Metro Manila of the Philippines. JICA-RI Working Paper No. 103. JICA, Tokyo, Japan.
- Lamichhane, K., Kawakatsu, Y., 2015. Disability and determinants of schooling: a case from Bangladesh. *Int. J. Educ. Dev.* 40, 98–105.
- Lamichhane, K., Sawada, Y., 2013. Disability and returns to education in a developing country. *Econ. Educ. Rev.* 37 (C), 85–94.
- Lamichhane, K., Tsujimoto, T., 2017. Impact of Universal Primary Education Policy on Out-of-School Children in Uganda. JICA-RI Working Paper No. 153. JICA, Tokyo, Japan.
- Lamichhane, K., Takeda, T., 2022. Disability and intrahousehold investment decisions in education: empirical evidence from Bangladesh. *Asian Dev. Rev. Stud. Asian Pac. Econ. Issues* 39 (1), 201–238.
- Levin, H.M., McEwan, P.J., 2001. Cost-Effectiveness Analysis: Methods and Applications. SAGE Publications.
- Mitra, S., Sambamoorthi, U., 2009. Wage differential by disability status in an agrarian labor market in India. *Appl. Econ. Lett.* 16 (14), 1393–1398.
- Mizunoya, S., Mitra, S., Yamasaki, I., 2018. Disability and school attendance in 15 low- and middle-income countries. *World Dev.* 104 (C), 388–403.
- Mont, D., Cote, A., 2020. Inclusive Social Protection for Empowerment of Persons with Disabilities Considering the Disability Related Extra Costs in Social Protection. <https://www.social-protection.org/gimi/RessourcePDF.action?id=56925>.
- Mori, S., Yamagata, T., 2009. A note on income and poverty of persons with disabilities in Metro Manila. In: Mori, S., Reyes, C., Yamagata, T. (Eds.), *Poverty Reduction for the Disabled in the Philippines—Livelihood Analysis from the Data of PWDs in Metro Manila*. IDE-JETRO, Chiba, Japan, pp. 145–157.
- Münich, D., Psacharopoulos, G., 2018. Education externalities-What they are and what we know. In: European Expert Network on Economics of Education (EENEE), EENEE Analytical Report No.34. MergedFile. [education-economics.org](https://www.education-economics.org). (Accessed 1 July 2022).
- Psacharopoulos, G., Patrinos, H.A., 2004. Returns to investment in education: a further update. *Educ. Econ.* 12 (2), 111–134.
- Rosales-Rueda, M.F., 2014. Family investment responses to childhood health conditions: intrafamily allocation of resources. *J. Health Econ.* 37, 41–57.
- Takeda, T., Lamichhane, K., 2018. Determinants of schooling and academic achievements: comparison between children with and without disabilities in India. *Int. J. Educ. Dev.* 61, 184–195.
- Weil, D.N., 2009. *Economic Growth*, second ed. Pearson Addison-Wesley, Boston, ISBN 978-0-321-41662-9.
- Washington group on disability statistics: <https://www.washingtongroup-disability.com/question-sets/wgunicef-child-functioning-module-cfm/>.

Legislation and policy in Europe

Valentina Migliarini^a, Julie Allan^b, Geert Van Hove^c, and Elisabeth De Schauwer^c, ^a School of Education, University of Birmingham, Birmingham, United Kingdom; ^b School of Education, University of Birmingham, Birmingham, United Kingdom; and ^c Department of Special Needs Education, Ghent University, Ghent, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	58
Legislation and policy in Italy	59
Legal framework: a critical analysis	59
Current amendments to IEP and PDP	61
Legislation and policies in England	61
Legislation and policies in the Flanders (Belgium)	63
Principle 1: the M-Decree strengthens the right to enroll	64
Principle 2: the M-Decree embeds the right to reasonable adjustments in educational legislation	64
Principle 3: the M-Decree tightens supervision of the referral of students to special education	65
Conclusion	66
References	66

Introduction

Inclusive education policies and practices have been received with praise and enthusiasm by national governments within and outside Europe (Artiles and Kozleski, 2016). However, inclusive education has also been resisted and criticized by educational professionals, researchers and politicians (Slee, 2011). The debates around the benefits and possibilities of inclusive education, which developed in the 1970s and throughout the 1990s in Europe and globally have now become less vigorous. Currently, the idea of inclusive education seems to permeate policies, legislation and teacher training in most European and non-European countries. Despite the universality of the concept of inclusive education, there remains a significant gap in the literature, generated by the lack of critical discussion on the origins of inclusive education and of the paradoxes of its contemporary implementation. This is particularly evident as inclusive education intersects with other equity education reforms (Ferguson et al., 2003). This contribution considers the notion of inclusive education, as framed by the “Education for All” movement and it focuses on contemporary inclusive legislation and policies in three countries, which have adopted the Salamanca Statement and Framework of Action on Special Needs Education (UNESCO, 1994), namely Italy, England and Flanders (Belgium).

The origins of inclusive education policies are grounded in a social justice agenda. Its objectives were to transform education systems to eradicate educational inequalities that had been historically embedded in the socioeconomic and cultural conditions of nations (Sailor et al., 1993). Inclusion intends to disrupt assumptions, practices, tools and policies that privilege certain groups of students and disadvantage others, on the basis of their race, ability level, language and gender (Skrtic, 1995). The objective of inclusive education has always been to design and implement education systems that would be responsive to the needs and interests of students representing a wide spectrum of human variability (Artiles and Kozleski, 2016). Considering these premises, most scholars in Europe and beyond would agree that inclusive education intends to transform school culture to:

- increase access of all students, and not just multiply marginalized groups;
- enhance school personnel and students’ acceptance of all students;
- maximize students’ participation in various domains of activity; and
- increase the achievement of all students (Artiles et al., 2006b).

The original justice agenda of inclusive education has been framed as a matter of individual rights, and connected to access, participation and outcomes. Multiple discourses have circulated in policy and practice in the history of inclusion. These discourses favor certain paradigms that have implications for reform agendas and their implementation. Discourses of inclusive education have been concerned with its justification and implementation (Dyson, 1999). The justification of inclusion is rooted in rights and ethics perspectives. Inclusive education has been justified when special education did not show a positive impact on students’ development. The paramount justification for inclusive education was that access to educational opportunities for all students is paramount (Artiles et al., 2006a). The discourse on inclusion implementation revolves around arguments for engaging in political struggles as a way for change, and pragmatic considerations. In the first case, inclusion cannot be a reality unless political struggle takes place to dismantle exclusionary education systems. In the second case, there are more practical aspects at stake: curriculum, pedagogy, assessment and professional training.

Influenced by an intersectional and interdisciplinary lens such as Critical Disability Studies and Critical Race Studies in Education (DisCrit) (Annamma et al., 2013), this contribution analyses inclusive education policies and legislations in Italy, England and Flanders (Belgium) from a historical perspective. The analysis attempts to answer these questions:

- (1) What is the problem(s) that the policies are trying to address?
- (2) Whose interest are these policies responding to?
- (3) Who's excluded from these policies?
- (4) What does it take to shift these legislations toward equity?

The article starts with an analysis of policies and legislation that paved the way for inclusive education, and its practical tools in Italy. The analysis shows that policies and practices of inclusion are still deeply rooted in the medical model of disability. The medical model of disability is driven by the imperative to “healthy normalcy” (Watermeyer, 2013, p. 29), whose defining characteristic is the location of disability within the individual with biological impairments, ignoring macro-sociopolitical contexts of racism, ableism, and other intersecting systems of oppression (Artiles, 2013). The article then continues with an exploration of the Special Educational Needs Code of Practice for Disability in the UK, which substantially fails its intent of being an inclusive legislation. Lastly, the contribution focuses on the M-Decree policy in the Flanders (Belgium), as a first attempt to implement the UN Convention on the Rights of Persons with Disabilities (UNCRPD, 2006). In conclusion, the article emphasizes how the three countries consider that inclusive education policies and practices are still highly influenced by the medical or “deficit” model of disability. In addition, there seems to be a reformist desire to involve parents and families in the decision-making process related to assessment of and educational provision for disabled children.

Legislation and policy in Italy

This section presents a critical analysis of the Italian laws and policies that paved the way for inclusive education and the creation of supposedly inclusive teaching tools, such as the Individualized Educational Plan (IEP) and the Personalized Teaching Plan (*Piano Didattico Personalizzato*, or PDP). Influenced by the intersectional framework of DisCrit, which exposes the fault lines in ableist and deficit-oriented perspectives of disability to illustrate how disability interconnects with other socially constructed identities (e.g., class, gender and sexual diversity), we analyze the focus of such policies and tools, and their implementation in Italian schools. Finally, we explore the latest amendment to policies regulating inclusive education and its tools, and how these have changed in light of distance and blended learning due to the global COVID-19 pandemic.

Legal framework: a critical analysis

Despite Constitutional obligations, Italy officially introduced the right to school inclusion for disabled and marginalized children in 1971 with *Law 118*. The same right was recognized for adults in 1977 under *Law 517*. Reformist in their objectives, these laws continued to remain anchored to the medical model of disability, the centrality of the “deficit” perspective and of the categorization of disabled and marginalized children and adults (D'Alessio, 2014). The resilience of the medical perspective of disability is particularly evident in the language used by *Law 517/1977*. Here, the importance of categorizing students and distributing them to “special” teachers is reinforced, in spite of the pressure put by the 1975 law proposal wanting to overcome the delegation of disabled students to specialist figures such as “specialized” teachers or social workers and health workers.

Fifteen years after *Law 517*, the publication of the Framework Law on disability—*Law 104*, 1992—maintains the “deficit” perspective on disability and on the bodies of disabled students. From 1977 to 1992, and indeed in the following years, the notion of inclusive education remains embedded within ideas of a “norm”, with ableism as the criterion for functioning, a disability diagnosis by medical doctors, and the specialization of support teachers. In these years, as a consequence of the dominance of the medical model of disability, school practices based on rehabilitation and compensation for disabled students have been strengthened and stabilized (D'Alessio, 2011). A testimony of the centrality of the medical model of disability can also be found in the Presidential Decree published on February 24th, 1994, which specifies the tools for the diagnosis and evaluation of disabled students in inclusive classrooms. The 1994 Decree introduces the Functional Diagnosis, which describes the clinical-functional situation of disabled students. The second tool, the Dynamic Functional Profile, indicates students' development through the analysis of strengths and difficulties at multiple levels: cognitive, socio-emotional, linguistic and communication, sensory, neuropsychological, and learning autonomy. The purposes and implementation mechanisms of the IEP, focusing on learning interventions, social relations and autonomy, are carefully delineated by Article 4 of the Decree.

The publication of the International Classification of Functioning Disability and Health (ICF) by the World Health Organization (2001) has boosted the medical perspective of disability in the protocols for the writing and implementation of the IEP and reinforced the claims of scientific validity of the notion of “functioning”, as a way to establish what students can or cannot do in the classroom. The Italian Ministry of Education (MIUR) has adopted the ICF to frame the guidelines for the inclusion of disabled students in school, containing instructions for IEP writing and implementation (Italian Ministry of Education, 2009). It has also supported research intending to test the application of ICF at every grade level in public schools (Italian Ministry of Education, 2010). The first research sponsored by the Ministry of Education involved 500 schools across Italy, and among these only 95 were selected for testing ICF implementation and financed 1.6 million Euros. This research did not produce the expected results that would have pushed the Ministry of Education to apply the ICF model in all Italian schools, despite the government's numerous requests and published policies. A more recent investigation has been conducted in the province of Verona, in the North East of Italy, to explore teachers' knowledge and application of the ICF in the design of the IEP for disabled students (Pasqualotto and

Ghirotto, 2019). Results show teachers' superficial knowledge of the ICF and lack of its systemic application at school level. Only 20% of the teachers interviewed report that the IEP is actually influenced by the ICF, and only 12% affirmed that the diagnosis made by medical doctors is based on the functioning classification (Pasqualotto and Ghirotto, 2019).

In the guidelines for school integration of students with disabilities (Italian Ministry of Education, 2006) and in the more recent Law 96 of 2019, the writing and implementation process of the IEP is still influenced by the ICF model. However, a significantly low percentage of Italian public schools have adopted the IEP framed through the ICF. The majority of them maintain the design established by the 1994 Presidential Decree. This results in major discrepancies at national school level in relation to the design and implementation of the IEP. Both the 2006 guidelines and the 2019 decree law remain influenced by the medical model of disability emphasizing individual "deficit". Accordingly, the 2019 Decree Law presents significant contradictions. For example, Article 1 on "inclusive principles and purposes" is contradicted by Article 2, which affirms that support interventions are exclusively for those who are certified as disabled. The Decree Law has currently removed disabling premises and language, such as "certified disability". However, maintaining the utterance "with a documented disability to guarantee inclusion" of students creates two main issues. First, the Decree confirms both the categorization and the certification of disability and reinforces a problematic and ableist definition of inclusion for school professionals and education stakeholders. Second, while promoting "inclusion," the Decree remains silent on the presence of special schools and advanced classes for "overachieving" students. The reasonable accommodations supported by the Decree are not extended to all students, but they are considered as compensatory measures targeting students with disabilities.

With the creation of the macro-category of Special Educational Needs (SEN), through the Ministerial Decree of December 2012, and the following Circular of March 2013, the Italian Ministry of Education introduces the PDP for those students classified as SEN, but without a certified disability, which otherwise will not have access to the IEP. Interestingly, even in this case, the Ministry's articulation of the PDP within policies follows the ICF model discourse. As expressed in the 2012 Ministerial Directive and 2013 Circular, the PDP focuses on:

- Certified diagnosis for students with learning difficulties;
- Analysis of students' needs and social contexts through observations;
- Planning of educational interventions based on personalization and individualization;
- Compensatory tools in case of severe difficulties;
- Assessment tailored on students' needs and pathways; and
- Families and students' commitment in order to establish a collaboration (p. 3).

As all of the above-mentioned legal frameworks and policy documents show, the IEP and PDP differ by virtue of diagnostic categories (i.e., disability and learning difficulties), but they are structured similarly in terms of theoretical underpinnings and systemic implementation.

The IEP and PDP both reiterate the dominance of the medical model of disability, based on individual deficits, and the necessity of categorizing children according to what they can or cannot do. In addition, they reinforce teachers' pathologizing attitudes toward students whose identities are at the intersections of multiple differences. Consequently, the notion of inclusion presented by the Ministry of Education is misleading, and the intention of actually introducing a more radical inclusive pedagogy within the IEP and the PDP is extremely limited.

Our analysis reveals five critical points in the current design of IEP and PDP:

1. Risk of exclusion;
2. Standardization;
3. Centrality of teachers and support teachers from dominant backgrounds;
4. Atemporality;
5. Classroom atmosphere.

The following five paragraphs provide a consideration of each of these points:

1. *Risk of Exclusion*: By applying a critical intersectional lens, it is possible to understand that the IEP and PDP can become tools of exclusion for multiply-marginalized students. Leaving unchallenged the medical paradigm, which constitutes the existing tools' foundation. The IEP and PDP intend to include, but they actually control, manage and confirm a perception of diversity according to an individual deficit way of understanding students' differences. Practical examples of this can be found in the emphasis given to differential teaching and curriculum, specialized support teachers, and a different and separated space outside of the classroom for students classified as SEN or with disabilities.
2. *Standardization*: This critical aspect derives from a generic approach focused to the application of a list of tools: mediators, compensatory, and dispensary tools not linked to functions (e.g., reduction of cognitive load, attention and processing support) and the different situations in which these can be used (e.g., explanations, questions, class assignments, personal study). Standardization is therefore limited to the citation of the above tools, supposedly making them neutral and static, without a proper critical pedagogical framework to justify the use of these tools.
3. *Centrality of Teachers and Support Teachers*: This is closely related to standardization as it highlights that, often, the choice of teaching tools and mediators is made without reference to the specificity of the student. The choice of mediators and tools cannot, therefore, be present in classroom practice or in the teacher's didactic thinking, but it must take into account the

implications it may have on the specificities of learning and on the pupil's requests. Of course, there are teachers who take a positive direction, but in various situations the compensatory and dispensary tools are not activated despite being included, above all, in the PDP: for example, when writing notes with maps, summaries, or slides are not replaced. So, the question is not only in relation to the tools (i.e., "What do we use?"), but above all, in relation to its function in the teaching-learning processes of a particular school phase. This implies paying attention to how inclusive tools, such as the PDP or the IEP, can be adapted to follow the evolutionary learning trajectories of specific students.

4. *Atemporality*: This term refers to the risk that tools such as the IEP and PDP are not modified through the years, following students' learning progression and development, and curricular changes. This resistance to change these tools on a yearly basis, highlights how SEN students and students with disabilities are perceived as unable to progress.
5. *Classroom, Atmosphere*: The classroom cannot be represented in a static way, but must be seen in constant change. Classrooms constitute an amalgam of complicated and ever-changing relationships. Classroom atmosphere constitutes an indicator of the group's quality of life and is determined by the conditions that teachers and pupils create through their interactions. The positive classroom atmosphere, however, is achieved through the positive relational and communicative methods of teachers, as well as their attention and intervention aimed at enhancement of personal and group identities and their assumption of responsibility. A collaborative-oriented class atmosphere, developed by a teacher is predisposed to help all students, and not to regimes of sanction and control. In fact, such classrooms favor the use of support and mediation strategies in the management of the difficulties that arise during the teaching-learning process. On the contrary, in non-collaborative classrooms, competition, the accentuated evaluations of the students, the absence of a collaborative climate and mutual support, the lack of attention of teachers to socio-relational dynamics, lead students to use defensive mechanisms.

Current amendments to IEP and PDP

Following the new *Decree Law n. 66* of April 13, 2017, a series of legislative measures were passed in 2020 which established that IEPs would be designed based on the International Classification of Functioning, Disability and Health (WHO, 2001). The IEP based on the ICF model has become a unique tool at the national level to manage disability in schools, and the centrality of families' and disabled students' participation in the structuring and drafting of the IEP (and consequently in PDPs) has been recognized. In the new ICF model, support teachers are not the only professionals with a decision-making power during the functional diagnosis and planning of interventions. Power and decision-making are also bestowed to parents via the School Working Group on Inclusion (*Gruppo di Lavoro Operativo per l'Inclusione* [GLO]), the school body in charge of elaborating the IEP (Tramontano, 2020). Nevertheless, medical experts, such as neuropsychiatrists and psychologists, are identified as key decision makers in the planning and designing of IEPs.

Despite medical expertise being crucial to understanding students' functioning, school contexts cannot be compared to therapeutic settings, with inevitable consequences in terms of teaching and learning strategies. Teachers and school practitioners cannot play a secondary role. Thus, with the latest reforms, the IEP and PDP models remain anchored in the "functional" and "deficit" medical model of disability. Lastly, these reforms have said very little about the practical ways in which schools and teachers are going to engage families, community members and students, and how to monitor students' progress at the end of each year, especially during transitions onto different grades of schooling.

A critical analysis of new inclusive policies and IEP models demonstrates that, although the active participation of parents, community members, and disabled students is highly encouraged for the structuring of the IEP, its philosophy remains anchored in the "functional", "deficit" medical model of disability. Importantly, very little has been said about the practical ways in which schools and teachers are going to engage families, community members and students, and how to monitor students' progress at the end of each year, especially during transitions onto different grades of schooling.

The next section explores legislations and policies on inclusive education in contemporary England.

Legislation and policies in England

This section focuses on one current policy in England, namely the Special Educational Needs Code of Practice for and Disability: 0–25 years, 2014 (Department for Education/Department of Health, 2015). This came into force in September 2014, in response to what the House of Commons Education and Skills Committee (2006) identified as "serious flaws in the SEN system" and their declaration that "the SEN system is failing". The Education and Skills Committee hailed the Warnock report of 1978 a success in "radically changing the conceptualization of special educational needs" and bringing in, first integration and then inclusion into subsequent (1981) legislation and recognized the value of the 1997 Labor Government's Green Paper *Excellence For All Children Meeting Special Educational Needs*, which aligned with the 1994 UN position statement on special education, urging governments to commit to inclusive education.

In spite of this, the Education and Skills Committee declared that the Warnock framework was "struggling to remain fit for purpose" and noted the frustrations this brought to all concerned. Here, the Committee specified parents, children, teachers, local authorities, but not school leaders. The Audit Commission (2002) and Office for Standards in Education (2004) had been similarly critical, specifying that too many children were waiting too long to have their needs met; parents were lacking confidence in the

system, leading to pressure for statements (the formal assessment of special educational needs); and some children who could be taught in a mainstream school were being turned away.

The Code of Practice was presented, in the foreword from the Parliamentary Under-Secretary of State for Health and the Parliamentary Under-Secretary of State for Children and Families as following “extensive consultation” (Department for Education/Department of Health, 2015, p. 11) and as a response to the problems identified by the House of Commons Education and Skills Committee and the Audit Commission. By promising a system that was “less confrontational” (Department for Education/Department of Health, 2015, p. 11), the Code of Practice acknowledged that this had been a problem in the past. The two senior officials promised early identification and assessment of special educational needs, full involvement of children and young people and their parents or carers, and raised aspirations through an increased focus on life outcomes, including employment and greater independence. They did not promise inclusion. They did, however, declare that their vision for children with special educational needs and disabilities was the same as that for all children—that they should achieve well at all stages of education and “lead happy and fulfilled lives” (Department for Education/Department of Health, 2015, p. 11).

The absence of a promise of inclusion should not be surprising when we recall that some four years earlier, the then Conservative Prime Minister declared the party’s intention to “prevent the unnecessary closure of special schools and remove the bias toward inclusion” (HM Government, 2010, p. 29). This followed on from a substantial challenge to inclusive education from Mary Warnock (2005), the so-called architect of inclusion (CSIE, 2005), who has pronounced it “the most disastrous legacy of the 1978 report” (Warnock, 2005, p. 22). Warnock claimed that inclusion—her preferred term rather than inclusive education—was problematic and argued that inclusion could damage children. Using unreflexive language, she argued that the “Down’s Syndrome girl” (Warnock, 2005, p. 35); the child who suffers from such disabilities as [asthma or epilepsy]; and autistic children who “suffer all the pains of the permanent outsider” and makes a strong argument for special schools. At the time, David Cameron described Warnock’s pamphlet as a “stunning recantation” (Hansard, 2005, Col 825). Clearly, by 2010, Cameron had come to see inclusion as less than attractive.

The Code of Practice has status as legislation, creating statutory requirements; as policy, in specifying good practice; and as guidance, in suggesting how practice might be achieved. However, the shady mixture of “musts” and “shoulds” throughout the document make it difficult to distinguish these constituent parts. The document attempts to be helpful in this regard specifying that where the word “must” be used, there is a statutory requirement under primary legislation, regulations or case law on those named to undertake the specified actions:

The bodies listed in paragraph iv. *must* have regard to the Code of Practice. This means that whenever they are taking decisions, they *must* give consideration to what the Code says. They cannot ignore it. They *must* fulfill their statutory duties towards children and young people with SEN or disabilities in the light of the guidance set out in it. They *must* be able to demonstrate in their arrangements for children and young people with SEN or disabilities that they are fulfilling their statutory duty to have regard to the Code.

Department for Education/Department of Health (2015, p. 12; original emphasis, with “musts” in bold in text).

As if anticipating confusion of their own making from indicating “shoulds” alongside the “musts,” clarification is offered:

So, where the text uses the word “should” it means that the guidance contained in this Code **must** be considered and that those who **must** have regard to it will be expected to explain any departure from it.

Department for Education/Department of Health (2015, p. 12; original emphasis, with “musts” in bold in text).

The term inclusive education never appears in any of the 261 pages of the Code of Practice. There is mention of inclusive practice and indeed a nod to the social model of disability in a reference to the UK government’s commitment to the removal of barriers to learning and participation in education. The securing of the assumption of mainstreaming in law through the Code of Practice is underlined, with signposting more than once of guidance, in Chapter 1, relating to inclusive practice. Chapter 1, however, contains no more than a restatement of the presumption of mainstreaming and of “a focus on inclusive practice and removal of barriers to learning” (Department for Education/Department of Health, 2015, p. 25). Elsewhere in the document, where the word inclusion is mentioned, it is expressed vaguely and as something that “should”, rather than “must” be done: “When securing funded early education for 2-, 3- and 4-year-olds local authorities should promote equality and inclusion for children with disabilities or SEN” (Department for Education/Department of Health, 2015, p. 80). It does, however, invoke the Equality Act 2010 to insist that schools “must make reasonable adjustments ... for disabled children to prevent them being put at a substantial disadvantage” (Department for Education/Department of Health, 2015, p. 93).

Referring to the National Curriculum Inclusion Statement, the Code of Practice states that all children should have access to a broad and balanced curriculum and calls for teachers to set high expectations for all pupils and to undertake lesson planning. However, with the specter of exclusion for some never far away, the Code states: “In many cases, such planning will mean that pupils with SEN and disabilities will be able to study the full national curriculum” (Department for Education/Department of Health, 2015, p. 94).

The excessive use of the word “support” throughout the document—appearing 775 times—creates a sense of disabled people as docile and unproductive (Hodkinson and Burch, 2019). There is also frequent mention of “employment” (431 times) and

“independence” (122 times), leading Hodkinson and Burch to contend that the ideology of inclusion has been replaced by an ideology of employment and independence. However, they also observe the note of responsibility and concomitant “steering and regulating” (Rose et al., 2003, p. 324) that this new ideology brings.

The Code of Practice creates little prospect of inclusion and instead adds to what Tomlinson (2012, p. 267) has described as “an expanded and expensive SEN industry.” It also creates some new types of losers. The first of these are key individuals who are “ghosted” or erased from the Code of Practice (Allan and Youdell, 2017) and who might be considered to be the central characters—children with special educational needs. The Code of Practice does this by not naming them in anything other than in general terms. A second set of “losers” are those who have previously been marginal to assessment and provision and who are targeted for enforced inclusion—parents. In their new role as customers, holders of “personal budgets” and purchasers of provision, parents are expected to exercise judgment and responsibility and to accept the consequences of their own decision-making. Such enforced inclusion is achieved through a process of responsabilization (Peeters, 2013; Olmedo and Wilkins, 2017), whereby parents are marked out as both the cause of and the potential solution for the problem. This may ultimately be exclusionary because of the accountability that surrounds this way of engaging parents.

The third set of “losers” are the local authorities. They are at the heart of the “musts” directives throughout the document and are required to take responsibility for what the Code of Practice refers to as the “Local Offer,” and to coordinate agencies across health, education and social care. They are also expected (though not required) to socialize parents into their new roles as customer-purchasers of provision and at the same time manage the consequences of placing requests that may be financially punitive. Schools appear to be the final group of “losers.” They are directed clearly enough on their statutory obligations, but are given minimal guidance on how to ensure their provision is inclusive. They are also given no indication that there will be any accountability in relation to inclusion. The consequences here are inevitable and unsurprising. There appear to be no “winners” in the Code of Practice. Everyone, including the state, is caught in the obligation of ensuring a “better life for those deserving of rescue from the dreadful fate of not progressing” (Thomson, 2013, p. 173).

The Code of Practice fails spectacularly as a policy on inclusion, but it may not have been intended as such. Hodkinson and Burch (2019) note that the efforts to control the population through the promotion of the concepts of independence have depleted activist movements such as the Independent Living Movement. They contend that disabled people should resist the docility and dependency visited upon them by the Code of Practice. There would certainly appear to be opportunities arising from the ineptitudes of the Code of Practice for reforming educational policy and for “wriggling out from the pincer-like grips of modernity” (Goodley, 2009, p. 260). Perhaps the first step toward this would be to replace “extensive consultation” (Department for Education/Department of Health, 2015, p. 11) with stakeholders with a policy construction process which involves them in a meaningful way, and which also gives them control of the discourse production. The article continues with an analysis of inclusive policies in Flanders, one of Belgium’s federal states.

Legislation and policies in the Flanders (Belgium)

As a result of the development of a robust special education system in Flanders during the past 50 years (European Agency for Special Needs and Inclusive Education, 2018), pupils with disabilities attending education were mainly taken care of in a non-inclusive education system of segregated special schools. Table 1 provides a statistical overview of children in special schools.

In Flanders, the Flemish speaking part of Belgium (with 6.63 million inhabitants on January 01, 2020), the M-Decree (approved in the Flemish Parliament 12-03-2014) came into effect at the beginning of the 2015–16 school year (Struyf et al., 2015). This Decree can be considered as the first attempt to organize a legal/policy translation of the UN Convention on the Rights of Persons with a Disability (UNCPRD, 2006) and more specifically of Article 24, this article states that inclusive education is a fundamental human right for every child (with a disability) at all educational levels. The M-Decree can be regarded as a large-scale educational reform (Victoor, 2017). It is a reform developed by the government, but responsibility for practical implementation lies with the schools. Such a great responsibility necessitates support “from above”. Various actions were proposed to support pupils, teachers and schools. For example, in-service training courses were offered with the aim to broaden and deepen the basic competencies of teachers to work more through an inclusive lens. In addition, “competency counselors” were called in at the request of schools, to

Table 1 Statistics students in nursery, primary and secondary special schools in Flanders (Flemish Education in Figures 2018–2019; onderwijs.vlaanderen.be/onderwijsstatistieken).

	2013–14	2014–15	2015–16	2016–17	2017–18	2018–19
Nursery	2042	2000	1984	1967	1986	2140 (T:264575) (type5:+134)
Primary	28,307	27,483	25,578	24,645	24,380	24,784 (T:467059) (type5:+209)
Secondary	20,495	20,661	20,333	20,130	20,331	20,544 (T:448000) (type5:+447)

start giving support to mainstream school teams in their inclusive work. Staff from special schools were also released to support their colleagues in mainstream schools.

Flanders chose from the start of the M-Decree to preserve Special Education (and even create one extra type of special education for children with autism) "... for pupils who need it ...". The Government assumed that consolidating the dichotomy between mainstream and special education could be used for the time being as an in-between phase on the way to an inclusive education system. When the M-Decree was introduced, clear principles were given priority. These principles had to be regarded as impulses to realize inclusive education. We will critically discuss those principles consecutively, adding the critical view in italics:

Principle 1: the M-Decree strengthens the right to enroll (UNIA Interfederaal Gelijkekansencentrum, 2016)

The Decree ensures that every child can enroll in mainstream education. However, from its commencement, the Decree made a distinction between children based on the curriculum they followed. The right to enrollment is being eroded for children with an individual adapted curriculum (IAC) by the option to cancel the enrollment. This feeds the perception that following a common curriculum is a precondition for inclusion. The condition requiring a child to be able to follow the general curriculum seems to indicate that making exceptions to the general curriculum is not viewed as reasonable adjustment, and therefore is the reason for the referral of disabled children to special education (De Schauwer et al., 2019).

Allowing enrollment based on a "cancellation clause" raises the question a priori for some children whether regular education is a viable option for them. Within 60 days the school can decide whether it is possible to give a permanent enrollment or whether the reasonable accommodations are disproportional. Termination of a registration does not always take place based on a thorough assessment of needs and feasible reasonable adjustments. Often a correct assessment of proportionality is still lacking, and schools come no further than a capacity assessment of what they are willing to do. An assessment based on established and transparent criteria is not a common practice. Now it seems too much like a boundary is drawn for those children who have an individual adapted curriculum. In this way, children do not get equal rights. A lot of legal uncertainty arises for families.

Principle 2: the M-Decree embeds the right to reasonable adjustments in educational legislation (UNIA Interfederaal Gelijkekansencentrum, 2016)

Pupils with disabilities are entitled to "reasonable accommodations". That right is laid down in the [UNCPRD \(2006\)](#) and in a Flemish Decree of July 10, 2008 containing a policy framework for Equal Opportunities and Equal Treatment. The M-decree anchored that right in the primary education decree and in the codex secondary education. The protocol of July 19, 2007 on the concept of "reasonable accommodations" in Belgium clarifies how the concept of "reasonable accommodation" should be interpreted. An accommodation is defined as a concrete measure—material or immaterial—that can neutralize the restrictive effects of an inaccessible environment on the participation of a person with a disability. Whether an accommodation is reasonable must be assessed based on well-defined indicators, for example:

- The financial impact of the adjustment, taking into account any financial compensations/subsidies and the financial strength of the person/organization subject to the obligation of "reasonable accommodation";
- The organizational impact of the requested "reasonable accommodation";
- The likely frequency and duration of use of the reasonable accommodation by persons with disabilities;
- The impact on the quality of life of (an) actual or potential user(s) with a disability;
- The impact on the environment and other parties involved;
- The lack of equivalent alternatives;
- The failure to comply with obvious or legally required standards.

In contrast to the right to inclusive education, which can be progressively achieved in Flanders, the right to reasonable accommodation must be immediately guaranteed. In the educational context, the right to reasonable accommodations creates certain obligations for schools. When a child wants to enroll in a school, the school may not refuse the enrollment because the child needs certain adjustments. Registration must be completed immediately. This registration must be followed by an analysis of the accommodations the child needs. Consultation and negotiation with the parents about the concrete needs of their child is an essential step.

The decision about the reasonableness of accommodations in educational contexts in Flanders rests ultimately within the authority of the school. When a school judges that an adjustment is not reasonable, it must be able to demonstrate that it has gone through the process correctly and assessed its reasonableness against the imposed indicators. The school must also actively look for solutions and alternative proposals that are reasonable. It was the job of the competency and/or pedagogical counselors to assist schools in applying reasonable accommodations. However, these counselors themselves felt they were missing expertise in what inclusive education is about, experience with developing reasonable accommodations and using the indicators to judge the reasonableness of an adjustment.

Some obstacles to the build-up of expertise were soon noticeable. To build expertise there was a need for a powerful and stable support model, which was lacking. Regular schools continue to consider that support for children with special educational needs is

the responsibility of internal or external experts. They expect guidance to come from outside of the classroom and to be provided on an individual basis. Teachers downplay their own roles in achieving inclusive education and “are more likely to pass questions [to specialists] because they believe that specialists have” better answers (Van de Putte and De Schauwer, 2013). Indeed, there is more and more outsourcing in regular schools, and research indicates that such outsourcing increases social inequalities (Verschuere, 2016).

Some reasonable accommodations required the intervention of specific disciplines (e.g., speech therapy, occupational therapy ...) these interdisciplinary teams were not available in mainstream schools (some disciplines may also be required when performing medical acts; mainstream schools lack the legal framework for these acts). Schools quickly appear to use the term “disproportionality”. This goes hand in hand with the lack of “a basic care framework” in mainstream schools, as a result of which staff quickly started using “not feasible” or “unreasonable” as an evaluation.

The medical model of disability is still strongly present and dominates practices. School teams are mainly looking at how they can adapt the student to the curriculum and usual school practices and less attention is paid to the question of how to remove barriers in the learning environment. The medical view is inextricably linked to the rapid referral of students to special education. Despite the Decree, it remains clear that inclusion is only possible for students who require minimal adjustments, and who follow the common curriculum. The obligation to make reasonable accommodations moreover is not accompanied by financial measures and a general strategy to make the Flemish education system inclusive (de Beco, 2016).

Principle 3: the M-Decree tightens supervision of the referral of students to special education

The coexistence of two educational systems makes referral a realistic option. A long history of referrals to the special education system cannot simply be reversed by the introduction of a Decree. In principle, regular education was considered the first option of education for all pupils. Special education was only considered as an alternative. It soon became clear that many factors continue to influence referrals. For example, in secondary schools, a student with an Individual Adapted Curriculum (IAC) in a mainstream school can only obtain “certificates of acquired competences”. This situation leads to important consequences with regard to validation of studies. Such a certificate of acquired competencies for a student who follows an IAC in mainstream secondary education provides no entitlement to “professional integration trajectories” and the associated “integration benefits”; these allowances are (in contrast) available to pupils attending special secondary education. Accessibility is a fundamental characteristic of an inclusive education system and concerns the design of and access to buildings, transport to school, circulation within the learning environment, classroom equipment, information and communication, learning resources, and so on.

In the Flemish monitor concerning school buildings (AGION, 2020) we learn that although there is a clear improvement in the evaluation of accessibility compared to the measurement in 2013 (+10%), accessibility measures are often still absent in 2018–2019. Necessary measures and interventions are often still lacking. In 56% of the school buildings there is a barrier-free access and access path, 53% have adapted toilets for students with a disability, and 53% have an adapted parking space. Other measures appear to be present at less than half of the sites, such as the availability of adapted toilets for staff with disabilities, clear signage, an elevator and ramp for level differences. The third edition of the school building monitor also specifically examined whether there were measures being taken in the context of the M-decree. This question shows that 19% of the buildings have been adjusted in function of the M-decree. In 58% it is stated that there is no adjustment. The large number of respondents who are not aware of this is striking (23%).

Investments must be made in strengthening the pedagogical competencies of teachers so that they can cope with reasonable accommodations within the curriculum and school practices. This starts with strengthening their knowledge and understanding of the educational goals and increasing their skills to formulate alternatives to those goals and/or to make adjustments (UNIA Interfederal Gelijkekansencentrum, 2016, 18). We can observe that teachers can follow (after their Bachelor’s program to become Teacher for primary Education) a Bachelor after Bachelor training about special educational needs. The fact that this (separate) program exists proves that basic teacher training is not giving enough attention to care and special needs.

Cooperation between regular and special education is not self-evident. Regular and special education are competing systems in the Flemish educational market. Such a competitive atmosphere does not guarantee a smooth, serene and constructive growth toward an inclusive education system. There are also tensions about the organization of the collaboration. Who determines which expertise and experience is sent to regular education? Who directs the staff? The distribution of all support remains in the hands of special education. Regular education indicates that support is only available sparingly and that it is not working proactively enough. Sometimes paramedical personnel are sent to regular schools. Most of them lack pedagogical expertise. Moreover, it is not always the most experienced and knowledgeable staff who are sent to support regular schools. Their experience is sometimes characterized by a medical view of disability and a standard approach to certain “types” of disability. Sometimes they take that experience with them to regular education, rather than experience with inclusion and strengthening the competences of regular schoolteachers.

It is clear that the implementation of the M-Decree did not go smoothly.

A very good external illustration of this can be found in a media analysis of the M-Decree (Tavernier, 2017). A large number of articles can be found in the press that support the negative positioning of the Decree. For example, there are articles that only describe the Decree as a “cost-saving measure”. Other articles frame the Decree as an “unattainable utopia”. A final group of articles framed the Decree as “unclear and insufficiently provided with information”. Some articles positioned as counter frames mainly

point to “Good Practices”, or to the “unnecessary fear” (of the abolition of special education), in a final counter frame a lot of information is given about the M-Decree.

Conclusion

In each of the three contexts reported—Italy, England and the Flanders region of Belgium—we can see some striking similarities. The most significant of these is the continued presence and force of the medical model of disability. This is attributed in part to the robust legacy of special educational needs, but in the case of Italy, can be seen as directly influenced by the application of the ICF. This International framework has brought a new legitimacy to the classification and assessment and subsequent individualization and personalization. International influence has been felt in both Flanders and England, through the [UNCRPD \(2006\)](#), but whereas in England this has merely constituted an affirmation of the government’s commitment to inclusive education, in Flanders the [UNCRPD \(2006\)](#) enshrined the right to a reasonable adjustment within the M-decree. However, the possibilities for exceptions to be made on the grounds of being “not feasible” or “unreasonable” have become apparent. [Buchner et al.’s \(2020\)](#) study of seven European countries also highlights the extent of the conservation of well-established special education systems which, together with “persistent ableist ideology” (p. 13), leads to a failed implementation of inclusive education.

A desire to secure the involvement of parents and families in decision-making about assessment, IEPs and educational provision is evident in the legislation and policies of each of the three countries. However, it is compliance and cooperation, rather than empowerment and authority that seems to be sought from parents and families. There is little evidence of efforts to strengthen parents’ capacity to function on an equal footing with officials and a strong sense of their continued construction as recipients, along with their children, of provision.

Teachers are recognized in the Italian, English and Flanders legislation and policies, as being central to the success of inclusive education. At the same time, their lack of competence, especially in relation to curriculum adaptations and making differentiated decisions based on individual needs, is exposed. Another legacy is also exposed here—one of training teachers only to recognize deficits and to default to standardized approaches of teaching and assessment. This overemphasis on students’ impairments and the failure of teacher training to equip teachers with the skills they need is also noted by [Buchner and Proyer \(2020\)](#) who consider it to be one of the key barriers to progress in inclusive education. Teacher education (rather than training) would be a good place to begin the shift to a more equitable and inclusive education system in Italy, England and Flanders. This should include, as [Buchner and Proyer \(2020\)](#) suggest, inter-disciplinary co-operation, co-operative teaching, group work, curriculum modification, differentiation and a range of teaching strategies.

Inclusive education, we are reminded, has become a “global concept” ([Boyle and Sharma, 2015](#), p. 2; [Buchner et al., 2020](#), p. 1). But this status brings with it new pressures in the form of demands for universality in meaning and competition to be further ahead—more inclusive—than other countries. We might do well to consider inclusive education, as traveling theory, as suggested by Edward [Said \(1983](#), p. 226): “Like people and schools of criticism, ideas and theories travel—from person to person, from situation to situation, from one period to another.” During this differentiated direction and pace of travel, according to Said, there is a certain displacement of ideas which includes “acknowledged or unconscious influence, creative borrowing, or wholesale appropriation” (p. 226). Although this displacement can be transgressive, disruptive and potentially productive it is also, importantly, colonializing and reductive. In order to shift the legislations and policy in Italy, England and Flanders toward equity and to move closer to the realization of inclusive education, we need greater criticality, curiosity and suspicion of the legislative and policy processes. We also need enhanced capacity among stakeholders to enable them to become actively engaged in the construction of legislation and policy.

References

- AGION, 2020. Schoolgebouwenmonitor 2018–2019—Indicatoren voor de kwaliteit van de schoolgebouwen in Vlaanderen Agentschap voor Infrastructuur in het Onderwijs Koning Albert II-laan 15, 1210 Brussel. Available at: https://www.agion.be/sites/default/files/images/D_Schoolgebouwenmonitor2018_2019_MS.pdf.
- Allan, J., Youdell, D., 2017. Ghostings, materializations and flows in Britain’s special educational needs and disability assemblage. *Discourse* 38 (1), 70–82.
- Annamma, S.A., Connor, D., Ferri, B., 2013. Dis/ability critical race studies (DisCrit): theorizing at the intersections of race and dis/ability. *Race Ethn. Educ.* 16 (1), 1–31.
- Artiles, A., Dyson, A., 2005. Inclusive education in the globalization age: the promise of comparative cultural historical analysis. In: Mitchell, D. (Ed.), *Contextualizing Inclusive Education*. Routledge, pp. 37–62. London, Routledge.
- Artiles, A., Kozleski, E., 2016. Inclusive education’s promises and trajectories: critical notes about future research on a vulnerable idea. *Educ. Pol. Anal. Arch.* 24 (43), 1–29.
- Artiles, A., Harris-Murri, N., Rostenberg, D., 2006a. Inclusion as social justice: critical notes on discourses, assumptions and the road ahead. *Theory Pract.* 45, 260–268.
- Artiles, A., Kozleski, E., Dorn, S., Christensen, C., 2006b. Learning in inclusive education research: re-mediating theory and methods with a transformative agenda. *Rev. Res. Educ.* 30, 65–108.
- Artiles, A.J., 2013. Untangling the racialization of disabilities: an intersectionality critique across disability models. *Du Bois Rev.* 10 (2), 329–347. <https://doi.org/10.1017/S1742058X13000271>.
- Audit Commission, 2002. *Special Educational Needs: A Mainstream Issue*. Audit Commission, London.
- Boyle, C., Sharma, U., 2015. Inclusive education—worldly views? *Support Learn.* 30 (1), 1–3.
- Buchner, T., Proyer, M., 2020. From Special to Inclusive Education Policies in Austria—Developments and Implications for Schools and Teacher Education.
- Buchner, T., Shevlin, M., Donovan, M., Gercke, M., Goll, H., Šiška, J., Janyšková, K., Smogorzewska, J., Szumski, G., Vlachou, A., Demo, H., Feyerer, e., Corby, D., 2020. Same progress for all? Inclusive education, the United Nations Convention on the rights of persons with disabilities and students with intellectual disability in European countries. *J. Pol. Pract. Intellect. Disabil.* 18 (1), 7–22.

- Centre for Studies in Education, 2005. News Digest. June. Retrieved 3 March, 2006, Available at. <http://inclusion.uwe.ac.uk/csie/june05.htm>.
- D'Alessio, S., 2011. Inclusive Education in Italy. A Critical Analysis of the Policy of Integrazione Scolastica. Sense Publishers.
- D'Alessio, S., 2014. Le Normative sui Bisogni Educativi Speciali in Europa e in Italia. Verso un'Educazione Inclusiva? La Prospettiva dei Disability Studies. In: Gasperi, P. (Ed.), *Pedagogia Speciale e BES*, pp. 217–244. Anicia.
- de Beco, G., 2016. Transition to inclusive education systems according to the convention on the rights of persons with disabilities. *Nord. J. Hum. Rights* 34 (1), 40–59.
- De Schauwer, E., Van de Putte, I., de Beco, G., 2019. Inclusive education in Flanders, Belgium: a country with a long history of segregation. In: *The Right to Inclusive Education in International Human Rights Law*. Cambridge University Press, pp. 514–529.
- Department for Education/Department of Health, 2015. Special Educational Needs and Disability Code of Practice: 0–25 Years. Statutory Guidance for Organisations Who Work With and Support Children and Young People With Special Educational Needs and Disabilities. DfE/DoH, London.
- Dyson, A., 1999. Inclusion and inclusions: theories and discourses in inclusive education. In: Daniels, H., Gardner, P. (Eds.), *World Yearbook of Education 1999. Inclusive Education*, pp. 36–53. Kogan Page.
- European Agency for Special Needs and Inclusive Education, 2018. Country Information for Belgium (Flemish Community) Legislation and Policy. Available at. <https://www.european-agency.org/country-information/belgium-flemish-community/legislation-and-policy>.
- Ferguson, D., Kozleski, E., Smith, A., 2003. Transforming General and Special Education in Urban Schools. *Advances in Special Education*. Available at. [https://www.emerald.com/insight/content/doi/10.1016/S0270-4013\(03\)80033-8/full/html](https://www.emerald.com/insight/content/doi/10.1016/S0270-4013(03)80033-8/full/html).
- Goodley, D., 2009. Bringing the psyche back into disability studies: the case of the body with/out organs. *J. Lit. Cult. Disabil. Stud.* 3 (3), 257–272.
- Hansard, 2005. Debate on Special Schools and Special Educational Needs, 22 June. Retrieved 4 December, 2007, Available at. <http://www.publications.parliament.uk/pa/cm200506/cmhansrd/vo050622/debtext/50622-11.htm>.
- HM Government, 2010. The Coalition: Our Program for Government. Cabinet Office, London. Retrieved on 18 November 2020 Available at. <http://programmeforgovernment.hmg.gov.uk/files/2010/05/coalition-programme.pdf>.
- Hodkinson, A., Burch, L., 2019. The 2014 special educational needs and disability code of practice: old ideology into new policy contexts? *J. Educ. Pol.* 34 (2), 155–173. <https://doi.org/10.1080/02680939.2017.1412501>.
- House of Commons Education and Skills Select Committee, 2006. Special Educational Needs: Third Report of Session 2005, 2006 (Volume 1). Report, Together With Formal Minutes. Stationery Office, London. Retrieved on 18 November 2020 Available at. <https://publications.parliament.uk/pa/cm200506/cmselect/cmmeduski/478/47809.htm>.
- Italian Ministry of Education (MIUR), 2006. Linee Guida per l'Integrazione degli Alunni con Disabilità. Available at. http://www.edscuola.it/archivio/norme/circolari/nota_4_agosto_09.pdf.
- Italian Ministry of Education (MIUR), 2009. Nota n. 3121 26/11/2009 Rilevazione classi con un numero elevato di alunni disabili. Available at. <http://www.integrazionescolastica.it/article/912>.
- Italian Ministry of Education (MIUR), 2010. Progetto ICF. Available at. http://www.edscuola.eu/wordpress/wp-content/uploads/2015/12/progetto_icf.pdf.
- Office for Standards in Education, 2004. Special Educational Needs and Disability: Towards Inclusive Schools. Office for Standards in Education, London. Retrieved on 16 December 2004, Available at. <http://image.guardian.co.uk/sys-files/Education/documents/2004/10/12/Ofsted.pdf>.
- Olmedo, A., Wilkins, A., 2017. Governing through parents: a genealogical enquiry of education policy and the construction of neoliberal subjectivities in England. *Discourse* 38 (4), 573–589.
- Pasqualotto, L., Ghirotto, L., 2019. Implementation of the ICF in schools with Legislative Decree n. 66/2017: a study amongst teachers. *Integrazione Scolastica Soc.* 18 (2), 184–198.
- Peeters, R., 2013. Responsibilization on Government's terms: new welfare and the governance of responsibility and solidarity. *Soc. Pol. Soc.* 12 (4), 583–595.
- Rose, N., Rabinow, P., 2003. *Foucault Today: The Essential Foucault; Selections From the Essential Works of Foucault, 1954–1984*. New Press, New York.
- Said, E., 1983. *The World, the Text, and the Critic*. Harvard University Press, Cambridge MA.
- Sailor, W., Gee, K., Karasoff, P., 1993. Full inclusion and school restructuring. In: Snell, M. (Ed.), *Instructions of Students With Severe Disabilities*. MacMillan, pp. 1–30.
- Skrkt, T., 1995. Power, knowledge and pragmatism: a postmodern view of the professions. In: Skrtk, T. (Ed.), *Disability and Democracy: Reconstructing (Special Education for Postmodernity)*. Teachers College Press, pp. 25–62.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling, and Inclusive Education*. Routledge.
- Struyf, E., Bodvin, K., Jacobs, K., 2015. Toeleiding naar Zorgaanbod. Een onderzoek naar bestaande praktijk en verklarende factoren op kind-, gezins-, en schoolniveau in het gewoon en buitengewoon onderwijs in Vlaanderen. Eindrapport OBPWO, 13.01. download 17-11-2020. <file:///Users/Geert/Downloads/7751.pdf>.
- Tavernier, H., 2017. Een Media-Analyse Van Het M-Decreet. Masterproef Universiteit Gent, 2016–2017 (promotor Elisabeth de Schauwer, Prof.dr.).
- Thomson, P., 2013. Romancing the market: narrativizing equity in globalizing times. *Discourse* 34 (2), 170–184.
- Tomlinson, S., 2012. The irresistible rise of the SEN industry. *Oxf. Rev. Educ.* 38 (3), 267–286.
- Tramontano, L., 2020. Riforma sostegno, Gruppi di Lavoro Operativo per l'inclusione: composizione e compiti. *Orizzonte Scuola*. Available at. <https://www.orizzontescuola.it/riforma-sostegno-gruppi-di-lavoro-operativo-per-l-inclusione-composizione-e-compiti/>.
- United Nations Convention on the Rights of Persons with Disabilities [UNCRPD], 2006. Convention on the Rights of Persons With Disabilities. Available at. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- UNESCO, 1994. World Conference on Special Needs Education: Access and Quality. Final Report. UNESCO.
- UNIA Interfederaal Gelijkekansencentrum, 2016. Inclusief Onderwijs in Vlaanderen: Een Tussentijdse Analyse. Available at. https://www.unia.be/files/Documenten/Tussentijdse_analyse_inclusief_onderwijs-CRPD.pdf.
- Van de Putte, I., De Schauwer, E., 2013. Becoming a Different Teacher...: Teachers' Perspective on Inclusive Education. *Erdélyi Pszichológiai Szemle* (december), 245–265/1454.
- Verschueren, K., 2016. Buitenschoolse hulpverlening en zorg op school: Samenhang, afstemming en verklarende factoren (Eindrapport OBPWO 12.01, KUL/UA).
- Victoor, M., 2017. Condities voor de implementatie van het M-Decreet als grootschalige hervorming: perspectieven van pedagogisch begeleiders en schoolleiders. Masterproef UGent, 2016–2017 (Promotor: Prof.dr. Geert Devos).
- Warnock, M., 2005. *Special Educational Needs: A New Look*. Impact No 11. The Philosophy Society of Great Britain, London.
- Watermeyer, B., 2013. *Towards a Contextual Psychology of Disablism*. Routledge, New York, NY.
- World Health Organization [WHO], 2001. International Classification of Functioning, Disability and Health (ICF). Available at. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>.

Inclusive pedagogies, new learning, new times

Anastasia O. Tzirides, Bill Cope, and Mary Kalantzis, University of Illinois Urbana-Champaign, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	68
The dimensions of learner differences	69
Material conditions	69
Social class	69
Locale	69
Family	69
Corporeal attributes	69
Age	69
Race	69
Sex and sexuality	69
Physical and mental abilities	69
Symbolic differences	70
Language	70
Ethnos	70
Communities of commitment	70
Gendre	70
Categorization of differences for an inclusive strategy	70
The confounding factors of the categorization	70
Unmanageable lists of differences	70
Internal group variations	70
Intergroup relationality	70
Intersectionality	71
Change	71
The “lifeworld differences” framework	71
Affinities	71
Personae	71
Agencies	71
Narratives	71
Historical responses to learner differences	72
Exclusion	72
Assimilation	72
Inclusivity	72
Toward an inclusive pedagogy	72
Awareness of learner profiles	73
Differentiated curriculum	73
Recursive feedback	73
Amplified learner agency	73
Productive diversity	73
Measuring comparabilities	74
Concluding remarks	74
References	74

Introduction

Living sociably in a context of human diversity has come to be recognized as one of the critical social challenges of our times. As a site of socialization where differences meet, education is a fundamental site where issues of diversity need to be addressed. In this article, we develop an approach to mapping differences of learners, examine the historical approaches of diversity so far, and offer a framework for inclusive pedagogy that comprises selected elements of previous approaches, while having a different end goal, one that we term “productive diversity.”

The dimensions of learner differences

The starting point of an inclusive pedagogy, that accepts and harnesses diversity in theory and in practice, is the recognition of the varied dimensions of differences that learners present in educational settings. Kalantzis and Cope (2016) have suggested a set of factors that can be used to categorize the human and learner differences and can guide the definition of one's identity, either as a self-definition or as a definition by others.

There are three main categories in this typology of human differences:

- *Material conditions*, which include social class, locale, and family;
- *Corporeal Attributes*, which include age, race, sex and sexuality, physical and mental abilities; and
- *Symbolic differences*, which include language, ethnos, communities of commitment, and genre.

We will now analyze the meaning of each of these categories and subcategories.

Material conditions

Social class

The first subcategory of material conditions comprises social class. Social class refers to the separation of society in groups, based on unequally ordered social structures, that have different access and control to material resources and social power. Throughout history, terms like slave-owner and slave, serf and feudal lord, worker and capitalist have been used to describe differences in social class (Hindness and Hirst, 1975). The subcategory outlines the relationships of power and control of one group over the other. During modern times, material inequality generated by the differences of the social classes has been intensified (Piketty, 2014), despite the more complicated classification and transformation of the social classes (Wright, 1985). In education, the subcategory of social class has been correlated with success and failure in schools (Bernstein, 1971; Bowles and Gintis, 1976).

Locale

This subcategory refers to the various favorable or unfavorable circumstances that occur in different locations (neighborhoods and regions) affecting education, employment and access to social resources. The circumstances established in each location are correlated to a great extent with the social classes (Harvey, 1996). More broadly, locale can be conceived across post-colonial divides, between the Global South and Global North, or between what are euphemistically called "emerging" versus "developed" economies.

Family

The third subcategory of material differences refers to the notion of family. Family pertains to the various relationships of domesticity and cohabitation that form socialization settings. Family structure and practice is a significant determinant in the domain of education, as it influences both informal learning and the connection with formal learning institutions.

Corporeal attributes

Age

Age is considered an element to determine bodily and mental capacities (Piaget, 1923/2002). By extension, it also defines pertinent and proper methods and ways of learning and teaching.

Race

Race refers to the observable characteristics that different human populations have as a result of the interaction of their genotype with the environment. These characteristics include skin color, facial features, hair color and texture, height and physique, and their biological importance is insignificant. Learning scientists today agree that intelligence and ability to learn are not rooted in the racial characteristics of an individual (Fraser, 1995). Nevertheless, race remains a perilous subcategory of the typology, as it supports "racist" claims and uses deceitful biological claims to justify inequality (Miles, 1989; Roberts, 2011).

Sex and sexuality

Sex refers to the biologically genetic differences between males and females. Sexuality concerns the biological tendencies to form bonds of sexual intimacy. Nowadays, there is a broader range of sexualities acknowledged, such as intersex, homosexuality, bisexuality, transsexuality or asexuality (Herdt, 1994; Mayo, 2013).

Physical and mental abilities

There is a range of bodily form and cognitive capability that expands and refers to what is framed as "disabilities". These can refer to impairments or conditions that affect hearing, sight, speech, mobility, child and adult development, learning, psycho-emotional challenges, and chronic illness. However, disabilities can also be understood as the deficiency of social and physical context to accommodate appropriately all the dimensions of physical and mental abilities (Bowe, 1978). This way of framing disability and disablement is referred to as the "social model of disability" (Oliver, 1990).

Symbolic differences

Language

Language concerns oral and written words and phrases. In educational settings, learners are differentiated in classes and modes of teaching and learning depending on the factor of language, for example regarding their first and second language in relation to the language of instruction if they speak a dialect, or how their social language relates to the accepted academic discourse (Kalantzis and Cope, 2012a).

Ethnos

Ethnos is the cultural identity of a “people.” Nation-states are trying to create “nations” that are connected with the geopolitical region that the state/country is located in (Anderson, 1991; Gellner, 1983). Nevertheless, nations cannot maintain their autonomy and ethnic diversity appears often within them, going beyond the borders of countries and states. Similarly, the indigenous identities are also disrupting the idea of a “pure nation” (Nakata, 2001). The same happens with diaspora communities that bring their identities with them to new places due to migration (Glazer and Moynihan, 1975).

Communities of commitment

Religion and political orientation are commitments that bring people together and they traverse the subcategories of language and ethnos (Bourdieu, 1993). These commitments are often complemented with processes of affiliation and institutional structures. Similarly, even being areligious or apolitical comes with the commitment of rejecting others’ declared commitments.

Gendre

Kalantzis and Cope (2012b) have proposed the term “gendre” to outline the symbolic and cultural traits that are related to sex and sexuality. Gendre extends the idea of gender (considering only male and female) and captures the complex and hybrid realities of sex and sexuality in the modern world, as well as their symbolic manifestation as identities.

The categorization of human characteristics is both necessary and raises inevitable problems of oversimplification. In creating educational policies and practices, it is important to think about these categories analytically and use them with great care.

Categorization of differences for an inclusive strategy

In this section, we want to highlight the issues that arise by using the above categorization in a simplistic way to guide everyday interactions. We also want to suggest an alternative method of categorization, that of “lifeworld differences.”

The confounding factors of the categorization

Unmanageable lists of differences

The list of categories and subcategories outlined above encompasses a finite number of the human and learner differences. This is necessary in order for it to be feasible for stakeholders in the provision of social services for policy building and to manage diversity in their everyday practices. However, over-simplification inevitably occurs, glossing over important nuances of human reality and experience. Therefore, it becomes critical to create a categorization that is subject to continuous refinement. As a result, we create an ever-extending list of categories and subcategories. The more finely grained the account of differences, the less manageable it becomes for stakeholders.

Internal group variations

Within any demographic classifier, the internal differences cross an enormous range. In fact, it is frequently the case that the internal differences within a group are greater than the average difference between the groups captured by the umbrella categories. Large and oversimplified categories can contribute to stereotypical generalizations that may lead to discriminatory actions. In educational settings, diversity programs with proper intentions can end up having counterproductive effects. One example is the oversimplification of the deficit models of education based on “disadvantaged” background. However, there are circumstances when “disadvantaged” background becomes the stimulus for students’ success and resilience (McGinty, 1997).

Intergroup relationality

The groups of the categorization are not distinct, as the semantics may tend to indicate, but defined in relation to other groups. Therefore, groups are defined in terms of social relationships of comparative power, privilege and access to resources. Through time, the groups evolve based on these historical and continuous intergroup relationships. In educational settings, these relationships (e.g., racism, sexism, comparative socio-economic privilege) are present and influence the group dynamics, focusing more on the intergroup definitions than on the internal features of the group members.

Intersectionality

A person will inevitably be a member of many different categories and subcategories in the aforementioned typology. The moment a person is classified to a category, this characteristic of the category interacts with the person's other characteristics and, thus, the categories transform. They take a new meaning, a unique form, that describes the identity of that specific person. This phenomenon of the identity that has many layers and aspects is known as "intersectionality" (Crenshaw, 2015). Each categorization of a person forms the unified whole. Every human is a unique combination of dimensions of differences, that create a unique blend of community experiences, based on the transformed characteristics of the categories. Through this process of transformation, each feature of the categories, when it is ascribed to a person, means something different than its original definition, due to the interaction with the other variables that constitute this person as a whole.

Change

The final factor that confounds categorization is change. The groups named in the categorization are never stable. They are dynamic and people, as social actors, change them individually and collectively. These changes occur either from people inside the groups: e.g., a woman who wants to change the role of women, or from outsiders: e.g., educators who want to assist their learner's transformation and growth beyond the possibilities that are attached to the initial definition of a category.

As it can be inferred from the above analysis, the role of categorization has lost its initial operational legitimacy. Nevertheless, we should not disregard it. The categorization has historical and pragmatic meaning, as it is a proof of the struggles for justice, recognition and institutional accommodation for the various differences. The categories are critical to remind us of our past and how we evolved and progressed through time to arrive in the situation we are in today and the way we need to view differences in the future of society. While keeping the categorization as a reminder and call to action, we need to embrace a more nuanced perspective and attend to the realistic differences in the world, which are called "lifeworld attributes."

The "lifeworld differences" framework

The notion of "lifeworld" refers to the things we know without consciously thinking how we ended up knowing them (Bourdieu, 1993; Cope and Kalantzis, 2000; Habermas, 1981/1987; Husserl, 1954/1970; Kraus, 2014). It concerns the way we are without deliberately deciding to be that way. It involves a set of habits, values, behaviors, and interests that are given and taken for granted in a specific context. The lifeworld is the basis of the person's existence, what they have learned and what they continue to learn and experience in their everyday life. Thus, the lifeworld is dynamic, unpredictable and continuously changing.

In formal educational settings, the lifeworld encompasses the lived experience that learners bring with them and has shaped their identity and subjectivity. This includes the influences from their family, local community, their friends, their peers, and the culture that they identify in. A lifeworld differences framework reveals the issues that come with the demographic categorizations, like over-generalization, unhelpful or false categorization, and even counterproductive stereotypes.

In what follows, we present some of the characteristics that constitute a lifeworld attributes framework.

Affinities

Affinities refer to the associations that a person identifies in and their meaning, which together build what the person is. The notion may include formal membership connections to groups, relations in networks, as well as informal connections in common causes.

Personae

The concept of "persona" refers to the type of person we imagine ourselves to be and present ourselves like this to others. It is the idea we have of how others see us, as well as how we want others to see us. This can be conscious, semi-conscious or unconscious. One's persona can be expressed through gestures, ways of behavior, social intersubjectivity, modes of "presentation of self" (Goffman, 1959), fashion and person-to-person interactions.

Agencies

In the framework of lifeworld differences, a person has a greater ability to act and define oneself in respect to their symbolic qualities. Even when inequalities of access to material resources continue or deteriorate, the arena of symbolic differences fosters greater agency for self-making and self-definition.

Narratives

This idea refers to the trajectories that people follow in their lives. A person's life narrative includes their past, present, as well as their goals and aspirations for the future. It reflects who the person is and how they portray their being.

To sum up, the lifeworld differences framework provides greater agency to a person to define themselves regarding their symbolic and corporeal differences, in contrast to the previous demographic categorization that was administered top-down. Concerning the access to material resources, despite the fact that they continue to exist and there is also greater polarization, a person has again greater agency to define their position in this area. In a learning context, a student has the agency to achieve intergenerational upward mobility by using formal educational qualifications and get access to material resources. Similarly, a learner by becoming political, can offer support to reform agendas that diminish or eradicate material inequalities.

Through the framework of lifeworld differences in learning situations, the process of self-understanding and understanding one's identity leads us to a far more refined comprehension of learning needs and the realities of the learners than it was ever possible with demographic categorization alone. This framework concentrates on the narratives of a person and their inter-relations. As for learning, it leads to success or failure depending on the degree that it engages with the subjectivity of learners. Therefore, in education, it is fundamental to address the differences in a deeper level and with a more refined approach than demographic categorization. Real learners are defined, beyond the demographics, by what they have already learned and what experiences they bring with them, and education should account for these attributes in order to aim for improved learning outcomes.

Historical responses to learner differences

Historically, there has been a range of alternative paradigms for addressing and negotiating the differences of humans in societies and specifically learners in schools. These paradigms are exclusion, assimilation and inclusivity. Each uses in different ways the demographic categories that we have analyzed earlier in this chapter.

Exclusion

Exclusion addresses the diversity that humans present by setting strict boundaries and promoting norms of homogeneity for people in social spaces. In this paradigm, difference has no place. It cannot be accepted in any way and people that are different are renounced.

Assimilation

Assimilation is a process whereby differences are viewed as something that needs to be changed in order to fit the "normal" standards. The dominant groups set these rules of sameness as conditions of entry in a social space. If the process of change fails to fit the standards, then access to social spaces is denied and dominant groups pretend that these differences do not exist, as a consequence of which the diversity of groups of people is ignored.

Inclusivity

Inclusivity aims to address the historical inequalities and injustices that come with differences as a result of changeable access to material resources, the corporeal attributes and the symbolic meanings of a person (Kalantzis, 2000). Inclusivity is an aspirational process that requires constant struggle, since the demographic categorization and approaches of the previous two paradigms still persist in our contemporary society. There have been some intermediate steps toward inclusivity, such as the idea of tokenistic "multiculturalism" (Castles et al., 1992; Glazer, 1997; Taylor, 1994) that celebrates symbolic differences, but doesn't necessarily improve social access. Even when intermediate actions are ambitious, results are mostly disappointing.

Inclusivity aims to achieve recognition of symbolic differences, and through redistribution aims to address material differences by sharing them more equitably. Inclusivity in practice contains elements of the previous two paradigms, but with the variation that these practices are without prejudice toward differences. For example, when linguistic minorities speak the dominant language this resonates with ideas of assimilation, or when specialist schools are designed for specific ethnic and linguistic communities, this seems to point to exclusion. Nevertheless, these practices do not have the same intentions as in the previous paradigms. Their goal is to extend participation, to offer equitable access to resources and to support a balanced reciprocal flow of symbolic interactions and equitable access to material resources.

Toward an inclusive pedagogy

Contemporary schools frequently follow the traditional paradigm of education by promoting sameness. In such schools, teachers, textbooks and the syllabus possess and transmit the knowledge and students listen and try to commit this knowledge to memory. In the end, students undertake tests that will evaluate if they remember the instructed content. Discussions in classrooms are one-sided, as the teacher asks a question and a student responds with the correct answer, without leaving leeway for diverse responses. This architecture of the classroom operation is a process that promotes sameness. As we have articulated before, this architecture does not correspond to the reality of the lifeworld.

An inclusive pedagogy, by contrast recognizes diversity and works with the learner differences. It constructs equal opportunities for all students, not identical opportunities, but of the same kind and value, regarding access to material resources, such as employment opportunities, civic participation, senses of belonging to broad and localized communities (Kalantzis and Cope, 2012b). In this context, learning is a process of accepting as a base the lifeworld experience and the cultural repertoire of learners—extending these, without leaving them behind, in order to achieve enhanced outcomes.

Following, we present some of the elements of an inclusive pedagogy:

Awareness of learner profiles

In educational settings, it is essential to know your students; for example, what languages they speak, their socio-economic background, and so on, by addressing all the demographic categories and subcategories that were outlined earlier in this article. This is a starting point of knowing the students, but we should not stay there. For an inclusive pedagogy, knowing your learners is a continuous, dialogical and open-ended process that needs to be embedded in the curriculum. Drawing from the Learning by Design and Multiliteracies pedagogy of [Cope and Kalantzis \(2015\)](#), “experiencing the known” is a process where students bring their lifeworld experiences to school and work with them and then, they are “applying creatively” what they have learned to school into the real world, by evolving their current lifeworld. Therefore, in an inclusive pedagogy, lifeworld differences of learners and teachers need to be recognized and mutually evolved and extended. The ultimate goal of such a pedagogy is to expand the horizons of learners, having as a basis their current lifeworld.

Differentiated curriculum

In the context of inclusive pedagogy, the one-size-fits-all instruction has no place. The instruction should be adaptive, personalized or differentiated, having as the ultimate goal adjusting learning to individual learners ([Walkington, 2013](#); [Wolf, 2010](#)). In the contemporary society, with the massive use of new digital media, this process has become more straightforward and practicable. New educational media elements like dashboards, on-the-fly learning analytics, alternative navigation paths and adaptive learning mechanisms make the idea of differentiated instruction manageable. Students can do the same thing at their own pace, space and time or they can work on different things that correspond better to their interests and needs. Through this process we aim to achieve better outcomes by respecting our students’ identities and lifeworld.

Recursive feedback

In an inclusive pedagogy that accounts for learner differences, feedback needs to be provided to students according to their interests, needs and topical focus. Recursive feedback or formative assessment ([Black and Wiliam, 1998](#); [Wiliam, 2011](#)) is a practice that provides feedback to students based on where they are in the learning process and what they want to achieve. Nowadays, with the abundance of digital tools and the advancement of technology, machine feedback and machine-mediated social feedback can be easily produced and be part of all the learning environments ([Cope and Kalantzis, 2013](#); [Cope et al., 2011](#)).

Amplified learner agency

An inclusive pedagogy should offer greater learner agency and use the identity and subjectivity of students as a driving force in the learning process. Learners are considered knowledge producers and not knowledge consumers of already created facts, definitions, and procedures. The influx of new media in the contemporary society has introduced increased opportunities for learners to have an active role in communication spaces. The affordances of the new media have promoted a shift in the balance of agency, where learners are no longer recipients of information, but knowledge and content creators. This change should be harnessed in an inclusive pedagogy. Learners should navigate and evaluate critically the abundance of sources that are available online or in community experience. Learners should be positioned as disciplinary practitioners and evidence of their learning should be evaluated in the knowledge artifacts they have created, for example: student projects, online discussions, navigation paths that they followed through games, simulations, or intelligent tutors. This way, the world of knowledge is redesigned by active learners and according to their interests, identities and experience. This outcome is also the goal of the well-known project-based learning approach ([Kilpatrick, 1918](#); [Rugg and Shumaker, 1928](#); [Waks, 1997](#)) that is long-cherished in the domain of education.

Augmented learner agency also means that the curriculum should cater for different starting points and it should allow learners to go after their interests, engaging the characteristics of their unique identity. It must embrace different modes of meaning-making and expressive potentials. It should offer alternative pathways and destinations through the learning process. Following such practices, an inclusive pedagogy provides to learners the opportunity to move and alter their pre-defined destiny according to the demographic categories, and it opens their horizons of possibility, as well as letting them transform themselves to something outside of the foregone conclusion, according to the discriminatory historical paradigms.

Productive diversity

The utilization of learner agency as an integral part of the learning process creates the basis for a new way of collaborative learning. The diverse perspectives of learners when brought into a discussion prompt more meaningful and deeper arguments. The possibility to choose to work on the topic the learners want, and a subsequent process of peer feedback exposes them to a range of diverse perspectives and ways of thinking. Sharing work-in-progress and finished works that are on different topics provides a range of different knowledge aspects to learners. Thus, the diversity of students’ lifeworld and their consequent perspectives develop into

resources for learning. This way, knowledge creation and learning engage more productively with diversity and set differences of experience, perspective, paradigm and expertise as the basis of the learning ecology they operate in. These generative synergies of difference are called “productive diversity” (Cope and Kalantzis, 1997; Page, 2007). The educational objective of productive diversity is to reinforce the self-development of learners, by encouraging them to navigate change, negotiate diversity, create shared understandings and extend their cultural and knowledge repertoires.

Measuring comparabilities

Through an inclusive pedagogy, the intention is to achieve learner self-transformation, by recognizing and engaging with the student's differences in the process of knowledge construction and learning. In such a case, traditional summative assessment that tests the knowledge of students in the end of a course cannot work. Each learner, through an inclusive approach, is developed uniquely and not in the exact same way as the other learners. This situation calls for innovative assessment and evaluation practices, like portfolio evaluation, peer-to-peer reviews, self-reviews, and “big data” analyses of learner progress (Cope and Kalantzis, 2015). These types of assessment should not target to measure identical “standards,” but comparabilities. They should check for evidence in the knowledge representations and the learning activity that are comparable in terms of epistemic effort, intellectual results, cognitive or practical difficulty. The main idea is that there is no need to ask all learners and their products to be exactly the same in order to be equal.

Concluding remarks

Learners nowadays should be seen as complex human beings with changing and self-defined attributes that can be reflected only in the narratives of their history and experience. Reflecting on the elements that we suggest as part of an inclusive pedagogy, the fundamental idea is to set learners' differences in the center of the learning process and utilize them in ways that contribute to their self-transformation. This means that in an inclusive approach, differences are recognized, they are utilized in the knowledge construction, and institutions provide the necessary circumstances for these processes to occur. Toward the achievement of this purpose, the typology of demographic categorization is a starting point. Practices of assimilation and exclusion can be present, but their intention is to provide safe spaces and accommodations for learners that will lead them to thrive in the environments whose ultimate objective is inclusion. As a consequence, in an inclusive pedagogy approach, strategies from all the historical paradigms can be exploited, as long as their underlying objective is only that of an inclusive productive diversity.

References

- Anderson, B., 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso, London.
- Bernstein, B., 1971. *Class, Codes and Control: Theoretical Studies Towards a Sociology of Language*, vol. 1. Routledge & Kegan Paul, London.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess. Educ.* 5, 7–74.
- Bourdieu, P., 1993. *The field of cultural production*. Columbia University Press, New York, NY.
- Bowe, F., 1978. *Handicapping America: Barriers to Disabled People*. Harper & Row, New York, NY.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Education Reform and the Contradictions of Economic Life*. Basic Books, New York, NY.
- Castles, S., Cope, B., Kalantzis, M., Morrissey, M., 1992. *Mistaken Identity: Multiculturalism and the Demise of Nationalism in Australia*. Pluto Press, Sydney.
- Cope, B., Kalantzis, M., 1997. *Productive Diversity: A New Approach to Work and Management*. Pluto Press, Sydney.
- Cope, B., Kalantzis, M., 2000. Designs for social futures. In: Cope, B., Mary, K. (Eds.), *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, London, pp. 203–234.
- Cope, B., Kalantzis, M., 2013. Towards a new learning: the “scholar” social knowledge workspace, in theory and practice. *E-Learn. Dig. Media* 10, 334–358.
- Cope, B., Kalantzis, M., 2015. Sources of evidence-of-learning: learning and assessment in the era of big data. *Open Rev. Educ. Res.* 2, 194–217.
- Cope, B., Kalantzis, M., McCarthy, S., Vojak, C., Kline, S., 2011. Technology-mediated writing assessments: principles and processes. *Comput. Compos.* 28, 79–96.
- Crenshaw, K., 2015. *On Intersectionality*. New Press, New York, NY.
- Fraser, S., 1995. *The Bell Curve Wars: Race, Intelligence, and the Future of America*. Free Press, New York, NY.
- Gellner, E., 1983. *Nations and Nationalism*. Cornell University Press, Ithaca, NY.
- Glazer, N., Moynihan, D.P., 1975. *Ethnicity: Theory and Experience*. Harvard University Press, Cambridge, MA.
- Glazer, N., 1997. *We Are All Multiculturalists Now*. Harvard University Press, Cambridge, MA.
- Goffman, E., 1959. *The Presentation of Self in Everyday Life*. Doubleday Anchor, New York, NY.
- Habermas, J., 1981/1987. *The Theory of Communicative Action*. In: *Lifeworld and System: A Critique of Functionalist Reason*, vol. 2. Beacon Press, Boston, MA.
- Harvey, D., 1996. *Justice, Nature and the Geography of Difference*. Blackwell, Cambridge, MA.
- Herd, G., 1994. *Third Sex, Third Gender: Beyond Sexual Dimorphism in Culture and History*. Zone Books, New York, NY.
- Hindness, B., Hirst, P., 1975. *Pre-capitalist Modes of Production*. Routledge & Kegan Paul, London.
- Husserl, E., 1954/1970. *The Crisis of European Sciences and Transcendental Phenomenology*. Northwestern University Press, Evanston.
- Kalantzis, M., Cope, B., 2012a. *Literacies*. Cambridge University Press, Cambridge.
- Kalantzis, M., Cope, B., 2012b. *New Learning: Elements of a Science of Education*. Cambridge University Press, Cambridge.
- Kalantzis, M., Cope, B., 2016. Learner differences in theory and practice. *Open Rev. Educ. Res.* 3 (1), 85–132.
- Kalantzis, M., 2000. Multicultural citizenship. In: Hudson, W., Kane, J. (Eds.), *Rethinking Australian Citizenship*. Cambridge University Press, Melbourne, pp. 99–110.
- Kilpatrick, W., 1918. The project method: the use of the purposeful act in the educative process. *Teach. Coll. Rec.* 19, 319–335.
- Kraus, B., 2014. Introducing a model for analyzing the possibilities of power, help and control. *Soc. Work. Soc.* 12, 1–12.
- Mayo, C., 2013. *LGBTQ Youth and Education: Policies and Practices*. Teachers College Press, New York, NY.

- McGinty, S., 1997. *Resilience, Gender and Success at School*. Peter Lang, New York, NY.
- Miles, R., 1989. *Racism*. Routledge, London.
- Nakata, M., 2001. Indigenous peoples, racism and the United Nations. *Common Ground*, Melbourne.
- Oliver, M., 1990. *The politics of disablement*. Macmillan Education UK, London. <http://link.springer.com/10.1007/978-1-349-20895-1>.
- Page, S.E., 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton University Press, Princeton, NJ.
- Piaget, J., 1923/2002. *Language and Thought of the Child*. Routledge, London UK.
- Piketty, T., 2014. *Capital in the Twenty-First Century*. Harvard University Press, Cambridge, MA.
- Roberts, D., 2011. *Fatal Invention*. Free Press, New York, NY.
- Rugg, H., Shumaker, A., 1928. *The child-centered school: an appraisal of the new education*. World Book Company, Yonkers, NY.
- Taylor, C., 1994. *Multiculturalism: Examining the Politics of Recognition*. Princeton University Press, Princeton.
- Waks, L.J., 1997. The project method in postindustrial education. *J. Curric. Stud.* 29, 391–406.
- Walkington, C.A., 2013. Using adaptive learning technologies to personalize instruction to student interests: the impact of relevant contexts on performance and learning outcomes. *J. Educ. Psychol.* 105, 932–945.
- William, D., 2011. *Embedded Formative Assessment*. Solution Tree Press, Bloomington, IN.
- Wolf, M.A., 2010. *Innovate to Educate: System [Re]Design for Personalized Learning*, a Report From the 2010 Symposium. Software and Information Industry Association, Washington, DC.
- Wright, E.O., 1985. *Classes*. Verso, London.

Within school stratification

James G. Ladwig^a and Amy McPherson^b, ^a School of Education, University of Newcastle, Newcastle, NSW, Australia; and ^b National School of Education, Australian Catholic University, Strathfield, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	76
Early schooling and ability grouping	77
Transition into and out of primary schools	79
Curriculum stratification within secondary schooling	80
Basic issue and the comprehensive high school	80
Within-school stratification internationally	82
References	83

Introduction

The term “within-school stratification” emanates from a long lineage of research into the ways in which schools and school systems differentiate educational opportunities of students over the course of a child’s school life. The notion of stratification builds directly from sociological analyses of occupational and status differentiation within the broader society structured by “layers” of socio-economic-status that have characterized most economically developed nations since their origins. While the first sociological theories developed different theoretical lenses for understanding that development (from the classic triad of Marx, Durkheim, and Weber), the role of schooling in the social and economic reproduction of those differentiated social layers has been a specific focus since the 19th century beginnings of mass education systems. Within the sociology of education, debates about which theoretical framework best characterized the functioning of schooling continued at least until the late 1980s (see, for example: [Hurn, 1993](#)). What’s important to note about those theoretical debates is that there wasn’t really debate about whether or not schooling did differentiate students into different pathways, which in turn typically led to differentiated social opportunities. The debate was largely about whether or not the differentiation was legitimate or functioned in a manner claimed by common rationalizations, evident in the debates on “meritocracy”.

In both the US and UK (and beyond), the specific focus on the stratifying effects of schooling came into sharper focus after WWII, due to at least four major historical developments. First, systems of schooling with relatively small portions of the population attending secondary school (in mostly urban contexts) had previously based educational differentiation on different types of secondary schools, following (roughly) the split between academic and vocational schooling—itsself inherited from the turn of the century. But after the war, the creation of comprehensive high schools was a growing transformation as vertical educational expansion saw increases in the proportion of students entering public (government) secondary schooling on both sides of the Atlantic. This transformation of secondary schooling was motivated by contradictory rationales, from interests in maximizing what we now call human capital (a la [Conant, 1959](#)), to the second major historical development: political calls for equal educational opportunities based on moral commitments to social equality. Ultimately, the focus on equal opportunity was directly implied in the then new debates about “meritocracy” (as in, e.g., [Halsey, 1977](#)). Third, the growth of quantitative sociological analyses based on large scale measures of outcomes and effects of institutional arrangements ultimately led to the landmark US analysis of equal educational opportunity led by James Coleman ([Coleman et al., 1966](#)). Developed in the context of the US equal rights movement, that report was particularly notable for documenting the degree of *between* school racial and economic segregation *and* that variance in student achievement was far greater *within* schools than between them (see [Gamoran and Long, 2007](#)). Fourth, specifically within the sociology of education, growing recognition of the importance of within-school processes, of the need to examine inside “the black-box” of schools, motivated several qualitative accounts of schools and classroom, such as Colin [Lacey’s \(1970\)](#) case study analysis of an English grammar school, or Phil Jackson’s ([Jackson, 1968](#)) US-based *Life in Classrooms*. The seminal account of these developments, and the ultimate call for a “New Sociology of Education” (the sociology of curriculum), remains Karabel and Halsey’s introduction to their benchmark text ([Karabel and Halsey, 1977](#)). The term “New Sociology of Education” was specifically created by Young ([Young, 1971](#)) to direct attention to the role of curriculum within the larger context of social and cultural reproduction.

Since the focus on within-school stratification has been recognized as a significant factor in the unequal distribution of educational opportunity and outcomes, there has been a very substantial production of research generated in the attempt to understand and potentially alter the now well documented inequities of conventional school structures and processes.

Two more general points need to be kept in mind. First, there is substantial research and evidence which makes it very clear that between school differentiation plays a significant role in the social reproduction of educational inequity. The degree to which this is true varies both between and within nations—but it is almost always true to some degree, even in school systems of the remaining socialist nations (see e.g., [Griffiths and Arno, 2015](#)). Second, we need to acknowledge that the institutional educational conditions at both ends of the pathway on which we focus here are also implicated in educational inequity. At both the pre-school, early

education level and at the post-secondary school tertiary level, processes which produce educational inequities are at play. Here, however, we will be focusing only on within-school stratification.

To organize this overview of that research, we will present relevant research along the pathway students would travel from initial primary/elementary schooling through to the end of secondary school attainment. From within the more quantitative driven sociological research, in addition to the overall pathway framing, it is important to also note two different phenomena. Consider **Fig. 1**: Hypothesized pathway of within school stratification below: a heuristic model of the hierarchical structures typically found within primary, middle and secondary schools (labeled in simplified high-middle-low terms) and the plausible pathways between these structures (whereby students may “move up” or “move down” as they proceed through schooling). The structure of this model is meant to represent the various foci of research addressing within-school stratification. That is, we will discuss literature that:

- (1) has described or measured the effects of different “levels” of hierarchically structured learning opportunities at each level of schooling; and
- (2) the literature which addresses the relatively likelihood of students taking one of the hypothesized pathways available to them, often referred to as transition points by sociologists of social stratification.

Because the nature of the various phenomenon being studied, differential learning experiences versus likelihood of following different pathways, the research documenting the consequences of their object of study can be split into two forms: the first is typically studied by reference to an amount of learning or achievement, the latter by reference to subsequent placement within the next levels of a hierarchically structured educational system.

Two more general points need to be made about the vast array of literature related to within school stratification. First, because the larger forms of social inequality and stratification are related to large societal inequities such as those based on class, race, gender, there is a substantial amount of research that analyses how each of the social inequities relate to educational stratification. Among that very large body of literature, within school stratification is often identified as a central mechanism for the maintenance of social inequities. Second, wherever there has been a focus on formally recognized outcomes (achievement scores, credential acquisition, attainment measures), there is also a focus on non-academic questions—often analyzed as cultural and social capital, or dispositions, norms and values. Here, the literature on analyzing the “hidden curriculum” a term coined by [Jackson \(1968\)](#), and later linked to class reproduction and ideological complicity by [Apple \(1971\)](#).

Early schooling and ability grouping

In many senses, within school stratification begins before children attend any schools. School readiness has been noted to be impacted by both access to early childhood education programs and the nature of those programs. While this has been a generally accepted view for many decades, over recent decades economists have been especially interested in early childhood and these programs have been identified as improving both academic and social and emotional behaviors through the provisions of systematic learning activities for children (see e.g., [Magnuson et al., 2007](#)). Within this broad pattern, however, children from disadvantaged families (living in poverty or with less-educated parents) are often less likely to attend early childhood education programs in

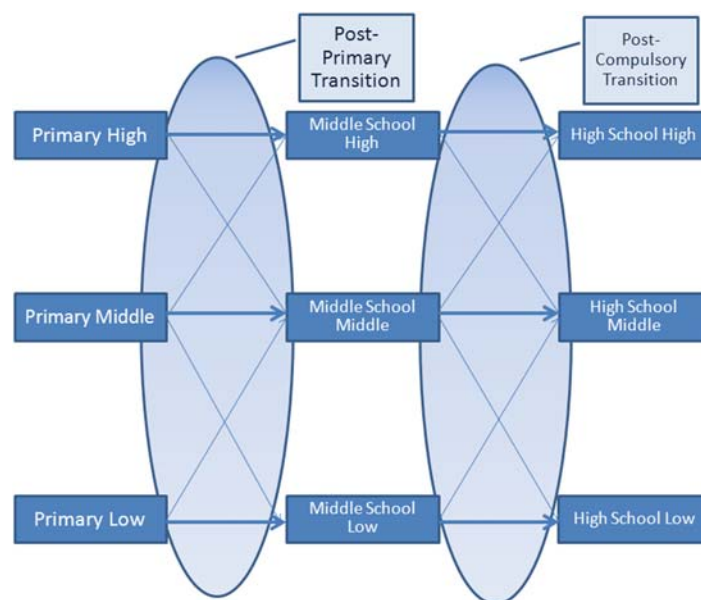


Fig. 1 Hypothesized pathway of within school stratification.

many national contexts, raising concerns that they are less prepared to start school than their more advantaged peers (see [OECD, 2020](#)). How schools respond to the known early childhood inequities becomes the basis of later within-school stratification processes.

For example, Curran, Little, Cohen Vogel & Domina (2020) examined Kindergarten readiness assessments used by schools to identify school and academic preparedness. Such information, they contend is utilized to support ability grouping practices. The basis of this concern stems from some of the earliest studies of ability group formation in primary schools where the interaction between school practices and students' unequal social backgrounds was brought into focus. The conceptual framework for this line of research was largely established well before the turn of this century. That is, although within-school stratification is typically associated with streaming or tracking in secondary schools, these studies of the earliest schooling have demonstrated processes that reinforce economic class differences, in the latter half of the 20th century, this was evident in the focus on the prevalent practice of grouping students by ability within classroom.

In the US, for example, the now classic studies of Ray Rist ([Rist, 1970, 1974](#)) documented how students in one urban elementary school were grouped in kindergarten and first grade by what the teacher understood to be ability—but that those groupings aligned with economic class differences between students. Students from lower socio-economic background were largely grouped into lower reading groups, whereas lower middle and upper-middle class students were assigned group membership into higher groups. The overall argument was that what the school and teacher understood to be a question of ability was directly related to background social class differences students had before they entered the school door. The main mechanism identified here was announced in the title of his 1970s article: the self-fulfilling prophecy, as reference to ([Rosenthal and Jacobson, 1968](#)). This important shift in Rist's analysis is that he showed how what was initially identified as a psychological phenomenon was a micro-level mechanism for larger macro-level social stratification. We should note that Rist's work built on earlier studies of primary school done within the Chicago Sociology tradition (e.g., [Becker, 1952](#); [Jackson, 1968](#)) as well as early US ethnographies of schooling ([Henry, 1955, 1963](#)). But this was not the only way of framing a corresponding analysis.

At the time when Rist's first studies appeared, students' social-linguistic differences were known to be a main means of recognizing social class differences, in the UK, the US and France. In the UK, the seminal work of Basil Bernstein and his colleagues had already documented class differences in language ([Bernstein, 1971, 1975a,b](#)), and in the US similar observation had been made in the early sociolinguistic work of ([Labov, 1966](#)). Bernstein and Labov were contemporaneous and later drawn into theoretical dispute on the degree with which either were committed to a so-called "deficit" view of language—but both had made clear links between differences in the use of English (both in cities where English was the dominant language) and economic class differences. There are two links within this work which should be noted. First, while language itself was the object of analysis in these studies, it was the underlying link to basic human cognition which made the connection to early school practices evident in the work of Aaron Cicourel and his colleagues ([Cicourel et al., 2014](#)). Second, this line of analysis led directly to an entire subfield of educational research. In the UK, the links between language and schooling sparked the early formulation of the New Sociology of Education ([Young, 1971](#)) and direct connections with the stratification of curriculum ([Bernstein, 1975b](#)). In the US this work was the early seeds of what became an immensely productive set of sociolinguistic analyses of school and the classroom, through the work of Courtney Cazden ([Cazden, 2017](#)), Bud Mehan ([Mehan, 1979](#)), and Shirley Brice Heath ([Heath and Press, 1983](#)). Methodologically, most of this work employed what became known as ethnomethodological studies ([Erickson et al., 1980](#)) wherein the connection between language and the larger social and political structures of society were evident ([Erickson, 1978](#)).

In France, similar studies relating to class reproduction and school had begun, most notably those led by Pierre Bourdieu. Translations of the work had begun around the time of their first publication (e.g., [Bourdieu, 1967](#)), but began to flourish in the Anglophone literature following the English translation of *Reproduction* ([Bourdieu and Passeron, 1977](#)) and continued well into the late 20th century ([Bourdieu et al., 1996](#)). Methodologically, these early French studies were largely what would now be called "mixed-methods", combining quantitative and qualitative empirical analyses. Theoretically they were very much informed by Bourdieu's own anthropological background, with an early focus on the socio-linguistic basis of cultural differentiation—in direct response to the heavy reliance on linguistic competence in the French educational system. Bourdieu's work on educational stratification was directed across all "levels" of systems—with a very significant focus on higher education, but here it is important to note that his earliest foray into analyzing educational stratification began from a focus on what linguistic differentiations reveals (see [Bourdieu et al., 1965](#)).

Beyond the focus on its sociolinguistic foundations, in the English language work, the interest in the impact of ability grouping in elementary schooling has also been a hotly contested terrain producing numerous studies—especially in reading and literacy. Major studies in this line of work often do not make direct links to wider questions of social stratification such as in the work Robert Slavin ([Slavin, 1987, 1993](#)), however the link between ability grouping was well noted as being in need of more study in England ([Francis et al., 2017](#)) just prior to Francis's ([Francis et al., 2019](#)) recent significant contribution to our overall understanding of the social impacts of ability-grouping. This recent work by Francis and her colleagues provides a nice summation of the various mechanisms identified in the literature on within-school stratification:

- misallocation to groups;
- lack of fluidity of groups;
- quality of teaching for different groups;
- teacher expectations of pupils;
- pedagogy, curriculum and assessment applied to different groups;

- pupil perception and experiences of “ability” grouping, and impact on their learner identities; and that,
- these different factors work together to cause a self-fulfilling prophecy (Francis et al., 2017 p. 4).

Notice these mechanisms correspond to three different domains: the structure of grouping students; the quality of what students experience within classrooms; and the subsequent psycho-social effects of these two main factors. This corresponds to the question of the effects of differential qualities within a given level of schooling and the issue of who gets stratified into what pathway at specific transition points.

The longevity of a research focus on understanding inadvertent ways that within group stratification makes its way into the early school years was examined by Burns and Mason (2002) who present findings about class formation in two groups/composite classes in elementary schools demonstrating that principals deliberately organized class arrangements for teaching purposes. The inadvertent side effect of such organizational practices was that “principals strategically robbed several single-grade classes of particular kinds of students to concentrate them in one combination class” (Burns and Mason, 2002, p. 227). Additionally, for children brought up outside the middle-class mainstream, differences between home and school are dramatic. “The conventions of the school, with its achievement orientation, its expectation that children will stay on task and work independently without close monitoring, its tight schedule of moving from lesson to lesson, its use of ‘network’ English, its insistence on punctuality, and its evaluation of children in terms of what they can do instead of who they are, all can be daunting” (Entwistle and Alexander, 1993, p. 405).

The recognition that organizational practices of schools play a role in how students develop from different social and economic backgrounds from the earliest years of schooling continues to motivate research to this day, with contemporary sociological analyses focusing on how parents’ own cultural capital plays a significant role in navigating student experiences of early schooling (Lareau, 2003; Lareau and Weininger, 2003) (inclusive of group placements), through to more technical questions of the impact ability grouping has on students from non-dominant cultural and linguistic backgrounds. The basic pattern of these more technical analyses (which grouping practices benefits whom) carries forward through the schooling years along two main axes. Overall the basic pattern suggests students from non-dominant backgrounds are more highly represented in lower “ability” grouping and that higher groups benefit from, or at least are not hindered by, that grouping; whereas lower group are typically hindered in terms of academic achievement and socio-emotional impacts (e.g., Lleras and Rangel, 2009; Catsambis and Buttar, 2012). It is important to note that this differs between subject areas of achievement and with specific second language students (Garrett and Hong, 2016; Robinson, 2008). It is also important to note that the impact of the grouping depends largely on how it is structured (size of group, for what length of time, what sort of pedagogical experience is offered within them) (see e.g., Hong et al., 2012). Interestingly, Stürup et al. (2017) have raised the prospect that socialization into stratified systems themselves is a consequence of some pre-school practices.

Transition into and out of primary schools

Key transition points in schooling have been significant in studies of within school stratification because of their important role in establishing achievement trajectories. Studies into school transition points also emphasize the complex interplay between children and young people’s home contexts and educational achievement, such as parent income, occupational status, education levels and educational expectations, as well as the levels of engagement between the family and the school.

The initial focus on early transitions into schooling has long been recognized as a crucial step in a much longer process of educational stratification, and within-class stratification has been among the many factors contributing to the formation of social strata. In a relatively early review of studies into school entry, Doris Entwistle and Karl Alexander point out that the logic of grouping students almost always is done with the intent of reducing heterogeneity within instructional groups, and that ‘instructional grouping is just one component of an elaborate apparatus for “educational sorting” that confronts children as they begin their school careers’ (Entwistle and Alexander, 1993, 413–414). Although with several out of school factors, the other school organizational practices identified by Entwistle and Alexander included sorting students into Special Education programs and holding students back (“grade retention”) before allowing them to advance beyond the first year of schooling. Through this mix of school organizational practices, according to Entwistle and Alexander, “educational stratification begins in earnest when children start their formal schooling or just before, at roughly ages 3–7. This is a ‘critical period’ in their lives, a special time, when for developmental/cognitive and social/contextual reasons they are especially responsive to schooling” (1992, p. 417).

Sociological studies in transition point allocations have traditionally focused on understanding the relative (and independent) contribution different background characteristics have on the probability of being allocated to one stratum or another (Kerckhoff, 2000; Lucas, 1999, 2001). The focus of such studies typically tries to tease out how much of that allocation depends on “educational” characteristics (most often prior achievement) relative to social demographic characteristics (most often race, economic advantage, gender). Studies in school retention or completion are the clearest examples of such allocation studies, of which there are numerous Australian exemplars at the post-compulsory level (see, e.g., Lamb, 2006). The differential pedagogical processes between strata are typically examined as potential contributors to subsequent achievement (or completion). In quantitative terms, the allocation processes are typically modeled in terms of the probability of specific allocation, and qualitative studies of allocations will focus on decision making, sense making, and choices made in the allocation process. Studies of within strata pedagogical processes ultimately lead to quantitative estimates of the effects of curriculum differentiation on future achievement or subsequent pathway allocations, and qualitative studies that link the nature of work, knowledge and subjectivity within strata with broader social stratifications beyond schooling. Current school research in Australia has focused almost exclusively on the last stage of these pathways and remains limited (as noted above).

The beginnings of various within-school stratification process such as negative evaluations and lower expectations from teachers, within and between class ability grouping, grade retention and special education classes all form part of educational experiences even in the early years of school. Entwistle and Alexander (1993) drew attention to educational stratification in the early years and transition school, identifying school transition as a “critical stage” in children’s development through the establishment of school achievement trajectories with school readiness identified in relation within school stratification processes in the early years. Duncan et al. (2007) highly cited work identified school readiness as having a lasting impact on school achievement and noted “math and reading skills at the point of school entry are consistently associated with higher levels of academic performance in later grades” (Duncan et al., 2007, p. 1444). Other work examining teachers’ perceptions of school readiness identifies a lack of academic skills as one of the central problems for children starting primary school by Rimm-Kaufman and Pianta (2000).

Dauber et al. (1996) in the first study to examine, middle school transition, noted that placements in higher tracked English and mathematics at the point of transition were better for children whose parents have higher levels of education and/or income (Dauber et al., 1996). The use of standardized test scores from the fifth grade were also a primary mechanism used for allocating students to a middle school track placement with test scores seen as an organizationally efficient method of making decisions regarding the class organization of students coming from several different feeder schools.

Curriculum stratification within secondary schooling

Basic issue and the comprehensive high school

The clearest example of within-school stratification occurs in secondary schools structured to deliver specific subject-based curriculum. The basic structure of using ability grouping within subject areas (e.g., having “top”, “middle” and “general” levels of mathematics classes) has been well established from the earliest secondary “colleges”. But that structure became a linchpin in the much more complicated organization structures created in the latter part of the 20th century, with the creation of the “comprehensive high school”. Calls for, and analyses of, the “comprehensive high school” date from the mid-1950s in both the US and UK (and beyond). Notable texts include the US post-WWII the call for comprehensive high schools to meet the expected expansion of secondary education—based on the promise of opportunity (Conant, 1959), and a later analysis led by the same Harvard University President, James Conant (Conant, 1967). In a parallel development, the UK was also developing and analyzing a similar expansion of secondary education, albeit from a rather different historical system of examinations for selection into secondary ‘public’ colleges and academically oriented “grammar schools” (see, e.g., Floud and Halsey, 1956).

The rationale for the development on both sides of the Atlantic (and throughout the Anglosphere), was constructed within the now well-known logic of meritocracy—opening entry into secondary schooling to all students with the promise of educational opportunities within a system designed to maximize the productivity of a full range of talents and skills, beyond traditional academic foci. The crucial question was when in a student’s life is it most efficient to provide differentiated educational pathways for different parts of the population. Where separate schools for education of students seen to be destined for vocational occupations were part of the larger systems since at least the turn of the 20th century, the comprehensive high school was designed to delay that differentiate by providing both academic and vocational education under one roof. As Floud and Halsey explained:

The most radical inference (of their analysis) is that English secondary education should be reorganised along the comprehensive lines. There is indeed something to be said for the view that the common secondary school is best suited to the needs of a technological society – least likely to stand in the way of free vocational choice and movement, most likely to produce the maximum supply of skilled and responsible individuals particularly in the middle ranges of the occupational structure.

Floud and Halsey (1956, p. 89).

The congruence of the comprehensive high school and the growth of policy rhetoric based in meritocracy was by no means contained to English speaking countries, as has been noted by a wide range of scholars in comparative education studying world systems of schooling. (e.g., Griffiths and Millei, 2012).

Thus, the terms of debates about within-school stratification were set from the inception of comprehensive high schools. What are within-classroom distinctions in primary schooling in the formation and use of ability-grouping became part and parcel of what became a dominant form of secondary schooling in the US and UK, at least until the late 1980s rise of school choice policies. That is, rather than grouping students within individual classrooms, secondary schools grouped students based on prior achievement (often conflated with an unexamined notion of “ability”), so that each class was seen to be relatively “homogenous” compared to the alternative mixed ability classes, known as “heterogenous grouping”.

In the US, debates on within-school stratification reached a watershed moment with the 1985 publication of Jeannie Oakes’ *Keeping Track: How Schools Structure Inequality* (Oakes, 1985). Building on data from a very large study of American secondary schools (Goodlad, 1984) in that work, and similar concurrent large studies of high schools (Boyer, 1983), Oakes presented a detailed and systematic analysis of how tracking practices related to larger dynamics of social and economic reproduction of the educational inequalities of economic class and race. Within the overall argument, Oakes focused on the major differences of educational experiences between tracks—paying particular attention to the lived experience of students in terms of what knowledge

each track taught, the forms of pedagogy taken within different tracks, the subsequent student attitudes, differences in classroom climate and the special case of vocational education. Effectively Oakes brought together both the selection mechanisms (or transition points) used to direct students into and out of tracks, and the internal quality of what student experience within tracks over time. One crucial point of this analysis was how individual decisions about small shifts or seeming details of management combine over time into a larger system creating the social inequality of the larger US society.

Two ongoing definitional debates lay behind the specifications of the dynamics Oakes revealed. First, how the specific forms of grouping can work in widely different ways, from entire year levels of students in a school being grouped into different class levels and then given entirely separate class schedules for each track or stream, through to ability grouping only applying to individual subjects and any subsequent grouping being due to students' selection or timetabling constraints. The distinction between these forms of tracking and streaming eventually developed its own technical nomenclature—in this case between tracking *de jure* v. *de facto* tracking (e.g., Lucas and Berends, 2002). Second there has been ongoing need to specify definitions as patterns of within school-stratification that have been identified with differing terminology—initially by the term “tracking” (mostly used in US nomenclature) and “streaming” (most commonly used in British studies).

Since *Keeping Track*, the overall trends of critical studies on streaming and tracking largely focus on two main lines of critique. On the one hand, there are many studies which question whether or not prior achievement is actually the main driver of which students get selected into which group, with many studies identifying disproportionate distribution of students based on social factors, including socio-economic status, race and gender (Oakes and Guiton, 1995). Studies of the probability of different social groups being more or less represented at different levels of within school systems of stratification are now widely globally dispersed, typically attempting to differentiate how much of those decisions are due to measurable prior achievement as opposed to other “exogenous” factors. Consequently, there are many studies debating the degree to which social inequality is implicated in any one system point. But when building from the insights of individual transitions into larger systemic analyses, the focus on divergent life pathways through schooling have been very clearly linked to socially stratified pathways into adulthood, both in the US (Kerckhoff, 1993) and the UK (Kerckhoff, 2019).

From an organizational perspective, the impact of tracking has been specified in greater detail. Gamoran (1992), for example, differentiated and tested the impacts of “four structural characteristics of tracking systems: selectivity, electivity, inclusiveness, and scope” (Gamoran, 1992, 812), as a means of explaining “variable effects” of tracking. That is, by the time of that analysis, studies into the impacts of tracking had revealed differing patterns from different studies, leading Gamoran to question which forms of tracking had which effects. Of prime focus for this earlier specification were differences that had been found between public and Catholic high schools in the US. Gamoran's (1992) analysis is especially notable because he not only specified patterns of effect, but also provided clear articulation of the plausible mechanisms leading to those effects. Consistent with much earlier work, the simple organizational design of tracking was seen to trigger several socio-psychological responses of students, akin to Rists's earlier elaboration of the “Pygmalion” effects that result from intra-group identification processes, where tracking was shown to create polarization between students in terms of their fit with academic demands of schooling. The crucial dimension here is to see that while “ability grouping” is a primary organization feature of tracking systems, the organization question is what form that ability grouping takes over time, and across subject areas (essentially the more static the grouping, the more notable its impacts).

This social-psychological phenomenon has since been the focus of a large body of research related to Herb Marsh's work on the generation of the academic self-concept of students and its later impacts. Formulated as the “Big-Fish-Small Pond” (BFSP) effect, Marsh and his colleagues have produced several studies demonstrating that students moving from smaller settings in which they are ranked academic high, into larger structures where they might no longer be top in class can negatively impact a student's academic self-concept (Marsh and Hau, 2003; Marsh et al., 2007). Importantly this work has been shown the BFSP effect to be multinational, and, in the more recent work of Parker, to be related to the degree of stratification within national systems (Parker et al., 2021, 2016, 2018).

The second line of analysis following *Keeping Track* that Gamoran (1992) began studying carefully focused more closely on the curricular experiences of students between transition points, within the classrooms structured by the stratifying effects of tracking. That is, in addition to the social-psychological impacts of tracking, Oakes' argued that significant differences develop in terms of the academic quality of teaching between tracks (Oakes, 1985), and Gamoran later confirmed this to be both true, and a factor in the noted consequent achievement gaps that develop between tracks (Gamoran, 1996; Gamoran and Mare, 1989; Gamoran et al., 1995). This line of analysis was informed by earlier studies of the ways in which teachers adjust teaching to specific contexts (such as in McNeil's *Contradictions of Control* (McNeil, 1986)), and its central tenet has been re-confirmed many times, (e.g., Carbonaro and Gamoran, 2002), as a partial explanation for the unequal achievement impacts of tracking.

In an overall analysis intended to inform policy, Gamoran has developed the argument that when judged in terms of overall achievement (average achievement per cohort), tracking can be shown to be less productive the more rigid and static the stratification of tracking becomes. This line of reasoning builds from the common finding that lower track placement have persistent negative impacts on students, while evidence of any benefits on top track placement has been much more mixed (Oakes et al., 1992). The question raised by an analysis of “productivity” is whether or not there are other, more effective means of grouping students (Gamoran, 2011).

This search for effective alternative systems of organizing high school instruction, in the US, has followed from the research documenting the negative consequences of tracking (Oakes and Lipton, 1992). That quest remains an ongoing question, and a source of frustration for many school reformers. After all, recognizing negative impacts of an organization is not the same as consciously

transforming organizational practices, and this later endeavor opens the larger attempts to reform or restructure schooling. For our purposes here, however, it is important to recognize that endeavor largely remains as yet unfulfilled (Gamoran, 2011).

Within-school stratification internationally

Outside of the voluminous address of within-school stratification research in the US, quite similar lines of analysis have developed in the UK. Dating to the early analyses of comprehensive high schools noted above, the specific interest in UK “streaming” was placed well and truly on the agenda by the early work of Stephen Ball. Ball (Ball, 1981) had developed a case study of one high school in an analysis that roughly parallels Oakes’ more generalized work from the same time period. Interestingly, however, it has only been in the last couple decades that strong empirical research has investigated within-school tracking in the UK in any further depth. Much of this UK work has been produced by Diane Reay (Reay, 1999a,b, 2004, 2007), and several other British colleagues (e.g., Connolly et al., 2019); including addressing the lack of impact of that research on policy (Francis et al., 2017). It is important to recognize that long term stratified student grouping practices do vary substantially between systems. The case of England makes this very clear where streaming practices have been documented to occur earlier than in many US systems (Taylor et al., 2022). Note here, that differences in nomenclature include the English use of the term “attainment” to refer to what is typically referred to as “achievement” in most other parts of the world. English attainments are levels of achievement on exams.

Like Oakes, much of the UK work, especially that by Reay, has come to the question of within-school stratification from a broader sociological interest in the lives and prospects of working-class families and children. This link between large societal level inequalities and academic organizational practices does remain contested by those who believe academic achievements can be readily separated using conventional statistical modeling. Examples of this latter concern largely attempting to differentiate prior student characteristics from the ostensible more justified notions of student cognitive ability (e.g., see Marks, 2013), as if statistically independent variables do represent independent social dynamics. That debate directly links to larger questions related to questions of meritocracy.

This conditioning of the impacts of within-school tracking/streaming is directly related to between school differences across several countries. In many parts of the world, including the UK, there are substantial school differences related to larger systems of student “choice” models, the uses of within school stratification varies substantially between schools (Taylor et al., 2022), and in many nations those between school differences can be related to differences in public and private systems, as they are in Australia (Perry and McConney, 2010; Perry and Southwell, 2014; Rowe and Perry, 2020). It is very important, also, to note that in many other parts of the world, streaming students into stratified groupings has largely gone uncontested until very recently (Johnston and Wildy, 2016; Liu et al., 2005; Wang, 2021). Whether or not any one system substantially shifts from the use of within-school stratification successfully remains to be seen.

Thus, while the amount of research emanating from the United States on within-school stratification is disproportionate to other nations, there is little reason to believe the general phenomenon is no less an issue wherever the practices are found. The inclusion of measures of differing types of student grouping has allowed substantial international work to develop, but the global picture remains less systematic than might be possible. To date, our capacity to develop such a more global model is, at least for conventional statistical models, limited simply because of a lack of measures. As Parker et al. (2021, p. 344) put it:

To our knowledge, no internationally comparable database allows for between country comparisons in between-school versus within-school tracking. As such, the relative relationship between ability stratification and tracking at these respective levels does not yet seem possible, but it is an important research topic when such data do become available.

That said, from the perspective of larger systems of educational policy and its link to broad social and economic impacts of education, the role of within-school stratification faces a dilemma:

‘Overall school tracking has an ambiguous effect in our sample of countries. On the one hand, and consistently with the previous literature, tracking has a detrimental impact on educational attainment, because it prevents some individuals from further progressing to the tertiary level of education (the diversion effect). On the other hand, the curricula offered in vocational schools seem more effective in promoting further training and adult competences (the specialization effect), thereby reducing the impact of parental background on these two outcomes. Thus, reducing the extent of student tracking, either by raising the age of first selection or by reducing the number of tracks available, may be appropriate for increasing inter-generational mobility in educational attainment, but may increase social exclusion for people from disadvantaged backgrounds.

Brunello and Checchi (2007).

This brief summary notes the special case of vocational education in the overall patterns by which schools reproduce existing economic class differences, a topic worthy of its own address. Here, however, it raises a larger question about just which outcomes schools are intended to serve. In the case of conventional academic achievement, however, there is little doubt that within-school stratification carries substantial negative impacts for those on the bottom of that hierarchy, and lower overall educational “productivity”.

References

- Apple, M.W., 1971. The hidden curriculum and the nature of conflict. *Interchange* 2 (4), 29–40.
- Ball, S., 1981. *Beachside Comprehensive: A Case Study of Secondary Schooling*. Cambridge University Press, Cambridge.
- Becker, H.S., 1952. Social-class variations in the teacher-pupil relationship. *J. Educ. Sociol.* 25 (8), 451–465.
- Bernstein, B., 1975. Class, Codes and Control. In: *Towards a Theory of Educational Transmission*, vol. III. Routledge & Kegan Paul, London.
- Bernstein, B.B., 1971. Class, Codes and Control: Theoretical Studies Towards a Sociology of Language. Routledge.
- Bernstein, B.B. (Ed.), 1975. Class, Codes and Control: Applied Studies Towards a Sociology of Language. Routledge.
- Bourdieu, P., 1967. Systems of education and systems of thought. *Int. Soc. Sci. J.* 19 (3), 338. [citeulike-article-id:6241181](https://doi.org/10.1080/00220226708839001).
- Bourdieu, P., Passeron, J.-C., de Saint Martin, M., 1965. *Rapport pedagogique et communication*. La Haye, Mouton, Paris.
- Bourdieu, P., Passeron, J.C., 1977. *Reproduction* (R. Nice, Trans.). In: *Education, Society and Culture*. Sage, London and Beverly Hills.
- Bourdieu, P., Passeron, J.C., de Saint Martin, M., 1996. *Academic Discourse: Linguistic Misunderstanding and Professorial Power* (R. Teese, Trans.). Stanford University Press, Palo Alto.
- Boyer, E.L., 1983. *High School: A Report on Secondary Education in America*. Harper & Row.
- Brunello, G., Checchi, D., 2007. Does school tracking affect equality of opportunity? New international evidence. *Econ. Pol.* 22 (52), 781–861. Retrieved from. <http://www.jstor.org.ezproxy.newcastle.edu.au/stable/4502215>.
- Burns, R.B., Mason, D.A., 2002. Class composition and student achievement in elementary schools. *Am. Educ. Res. J.* 39 (1), 207–233. <https://doi.org/10.3102/00028312039001207>.
- Carbonaro, W.J., Gamoran, A., 2002. The production of achievement inequality in high school English. *Am. Educ. Res. J.* 39 (4), 801–827. <https://doi.org/10.3102/00028312039004801>.
- Catsambis, S., Buttar, A., 2012. Revisiting “Kindergarten as academic boot camp”: a nationwide study of ability grouping and psycho-social development. *Soc. Psychol. Educ.* 15 (4), 483–515. <https://doi.org/10.1007/s11218-012-9196-0>.
- Cazden, C.B., 2017. *Communicative Competence, Classroom Interaction, and Educational Equity: The Selected Works of Courtney B. Taylor & Francis*, Cazden.
- Cicourel, A.V., Jennings, K.H., Jennings, S.H.M., 2014. *Language Use and School Performance*. Elsevier Science.
- Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.J., Weinfeld, F.D., York, R.L., 1966. *Equality of Educational Opportunities*. United States Government Printing Office, Washington D.C.
- Conant, J.B., 1959. *The American High School Today*. McGraw-Hill, New York.
- Conant, J.B., 1967. *The Comprehensive High School*. McGraw-Hill, New York.
- Connolly, P., Taylor, B., Francis, B., Archer, L., Hodgen, J., Mazenod, A., Tereshchenko, A., 2019. The misallocation of students to academic sets in maths: a study of secondary schools in England. *Br. Educ. Res. J.* 45 (4), 873–897. <https://doi.org/10.1002/berj.3530>.
- Dauber, S.L., Alexander, K.L., Entwisle, D.R., 1996. Tracking and transitions through the middle grades: channeling educational trajectories. *Sociol. Educ.* 69 (4), 290–307. <https://doi.org/10.2307/2112716>.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C., 2007. School readiness and later achievement. *Dev. Psychol.* 43 (6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>.
- Entwisle, D.R., Alexander, K.L., 1993. Entry into school: the beginning school transition and educational stratification in the United States. *Annu. Rev. Sociol.* 19 (1), 401–423. <https://doi.org/10.1146/annurev.so.19.080193.002153>.
- Erickson, F., 1978. *The Politics of Speaking: An Approach to Evaluating Bilingual-Bicultural Schools*. National Dissemination and Assessment Center, California State University, Los Angeles.
- Erickson, F., Florio-Ruane, S., Buschman, J., Institute for Research on Teaching, Michigan State University, 1980. *Fieldwork in Educational Research*. Institute for Research on Teaching, Michigan State University.
- Floud, J.E., Halsey, A.H., 1956. English secondary schools and the supply of labour. In: Halsey, A.H., Floud, J.E., Anderson, C.A. (Eds.), *Education, Economy and Society: A Reader in the Sociology of Education*. The Free Press, New York, pp. 80–92.
- Francis, B., Archer, L., Hodgen, J., Pepper, D., Taylor, B., Travers, M.-C., 2017. Exploring the relative lack of impact of research on ‘ability grouping’ in England: a discourse analytic account. *Camb. J. Educ.* 47 (1), 1–17. <https://doi.org/10.1080/0305764X.2015.1093095>.
- Francis, B., Taylor, B., Tereshchenko, A., 2019. Reassessing ‘Ability’ Grouping: Improving Practice for Equity and Attainment. Taylor & Francis.
- Gamoran, A., 1992. The variable effects of high school tracking. *Am. Socio. Rev.* 57 (6), 812–828. <https://doi.org/10.2307/2096125>.
- Gamoran, A., 1996. Curriculum standardization and equality of opportunity in scottish secondary education: 1984–90. *Sociol. Educ.* 69 (1), 1–21. <https://doi.org/10.2307/2112720>.
- Gamoran, A., 2011. Designing instruction and grouping students to enhance the learning of all: new hope or false promise? In: Hallinan, M.T. (Ed.), *Frontiers in Sociology of Education*. Springer Netherlands, Dordrecht, pp. 111–126.
- Gamoran, A., Mare, R.D., 1989. Secondary school tracking and educational inequality: compensation, reinforcement, or neutrality? *Am. J. Sociol.* 94 (5), 1146–1183. Retrieved from. <http://www.jstor.org/stable/2780469>.
- Gamoran, A., Long, D.A., 2007. Equality of educational opportunity a 40 year retrospective. In: Teese, R., Lamb, S., Duru-Bellat, M., Helme, S. (Eds.), *International Studies in Educational Inequality, Theory and Policy*. Springer Netherlands, Dordrecht, pp. 23–47.
- Gamoran, A., Nystrand, M., Berends, M., LePore, P.C., 1995. An organizational analysis of the effects of ability grouping. *Am. Educ. Res. J.* 32 (4), 687–715. <https://doi.org/10.2307/1163331>.
- Garrett, R., Hong, G., 2016. Impacts of grouping and time on the math learning of language minority kindergartners. *Educ. Eval. Pol. Anal.* 38 (2), 222–244. <https://doi.org/10.3102/0162373715611484>.
- Goodlad, J.I., 1984. *A Place Called School: Prospects for the Future*. McGraw-Hill Book Company.
- Griffiths, T.G., Arno, R.F., 2015. World culture in the capitalist world-system in transition. *Glob. Soc. Educ.* 13 (1), 88–108. <https://doi.org/10.1080/14767724.2014.967488>.
- Griffiths, T.G., Millei, Z., 2012. *Logics of Socialist Education: Engaging With Crisis, Insecurity and Uncertainty*. Springer, Netherlands.
- Halsey, A.H., 1977. *Heredity & Environment*. Methuen.
- Heath, S.B., Press, C.U., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press.
- Henry, J., 1955. Docility, or giving teacher what she wants. *J. Soc. Issues* 11 (2), 33–41. <https://doi.org/10.1111/j.1540-4560.1955.tb00313.x>.
- Henry, J., 1963. *Culture Against Man*. Random House, Inc., New York.
- Hong, G., Corter, C., Hong, Y., Pelletier, J., 2012. Differential effects of literacy instruction time and homogeneous ability grouping in kindergarten classrooms: who will benefit? Who will suffer? *Educ. Eval. Pol. Anal.* 34 (1), 69–88. Retrieved from. <http://www.jstor.org.ezproxy.newcastle.edu.au/stable/41413075>.
- Hurn, C.J., 1993. *The Limits and Possibilities of Schooling: An Introduction to the Sociology of Education*. Allyn and Bacon.
- Jackson, P.W., 1968. *Life in Classrooms*. Holt, Rinehart and Winston, New York.
- Johnston, O., Wildy, H., 2016. The effects of streaming in the secondary school on learning outcomes for Australian students - a review of the international literature. *Aust. J. Educ.* 60 (1), 42–59. <https://doi.org/10.1177/0004944115626522>.
- Karabel, J., Halsey, A.H. (Eds.), 1977. *Power and Ideology in Education*. Oxford University Press, New York.

- Kerckhoff, A.C., 2019. *Getting Started: Transition to Adulthood*. Taylor & Francis, Great Britain.
- Kerckhoff, A., 2000. *Generating Social Stratification: Toward a New Research Agenda*. Westview Press, Boulder, CO.
- Kerckhoff, A.C., 1993. *Diverging Pathways: Social Structure and Career Deflections*. Cambridge University Press.
- Labov, W., 1966. *The Social Stratification of English in New York City*. Center for Applied Linguistics.
- Lacey, C., 1970. *Hightown Grammar: The School as a Social System*. Manchester University Press.
- Lamb, S., 2006. *Longitudinal Surveys of Australian Youth: Variations in VET Provision across Australian Schools & Their Effects on Student Outcomes*. Australian Council for Educational Research, Camberwell, VIC.
- Lareau, A., 2003. *Unequal Childhoods: Class, Race, and Family Life*. University of California Press.
- Lareau, A., Weininger, E.B., 2003. Cultural capital in educational research: a critical assessment. *Theor. Soc.* 32 (5/6), 567–606. <https://doi.org/10.2307/3649652>.
- Liu, W.C., Wang, C.K.J., Parkins, E.J., 2005. A longitudinal study of students' academic self-concept in a streamed setting: the Singapore context. *Br. J. Educ. Psychol.* 75, 567–586. <https://doi.org/10.1348/000709905X42239>.
- Lleras, C., Rangel, C., 2009. Ability grouping practices in elementary school and African American/hispanic achievement. *Am. J. Educ.* 115 (2), 279–304. <https://doi.org/10.1086/595667>.
- Lucas, S.R., 1999. *Tracking Inequality: Stratification and Mobility in American High Schools*. Teachers College Press, New York, NY.
- Lucas, S.R., 2001. Effectively maintained inequality: education transitions, track mobility, and social background effects. *Am. J. Sociol.* 106 (6), 1642–1690.
- Lucas, S.R., Berends, M., 2002. Sociodemographic diversity, correlated achievement, and de facto tracking: a Magazine of Theory and Practice. *Sociol. Educ.* 75 (4), 328–348. <https://doi.org/10.2307/3090282>.
- Magnuson, K.A., Ruhm, C., Waldfogel, J., 2007. Does prekindergarten improve school preparation and performance? *Econ. Educ. Rev.* 26 (1), 33–51. <https://doi.org/10.1016/j.econedurev.2005.09.008>.
- Marks, G.N., 2013. *Education, Social Background and Cognitive Ability: The Decline of the Social*. Taylor & Francis Group, London, United Kingdom.
- Marsh, H., Hau, K.-T., 2003. Big-fish-little-pond effect on academic self-concept: a cross-cultural (26-country) test of the negative effects of academically selective schools. *Am. Psychol.* 58 (5), 364–376.
- Marsh, H.W., Trautwein, U., Lüdtke, O., Baumert, J., Köller, O., 2007. The big-fish-little-pond effect: persistent negative effects of selective high schools on self-concept after graduation. *Am. Educ. Res. J.* 44 (3), 631–669.
- McNeil, L.M., 1986. *Contradictions of Control: School Structure and School Knowledge*. Routledge & K. Paul.
- Mehan, H., 1979. *Learning Lessons: Social Organization in the Classroom*. Harvard University Press.
- Oakes, J., 1985. *Keeping Track: How Schools Structure Inequality*. Yale University Press, New Haven, CT.
- Oakes, J., Gamoran, A., Page, R., 1992. Curriculum differentiation: opportunities, outcomes, meanings. In: Jackson, P.W. (Ed.), *Handbook of Research on Curriculum*. Macmillan Pub. Co, New York, pp. 570–609.
- Oakes, J., Guiton, G., 1995. Matchmaking: the dynamics of high school tracking decisions. *Am. Educ. Res. J.* 32 (1), 3–33. <https://doi.org/10.3102/00028312032001003>.
- Oakes, J., Lipton, M., 1992. Detracking schools: early lessons from the field. *Phi Delta Kappan* 73 (6), 448. Retrieved from. <http://proquest.umi.com/pqdweb?did=29443434&Fmt=7&clientid=29744&RQT=309&VName=PQD>.
- OECD, 2020. *Early Childhood Education: Equity, Quality and Transitions*. Retrieved from Paris. <https://www.oecd.org/education/school/early-childhood-education-equity-quality-transitions-G20.pdf>.
- Parker, P., Dicke, T., Guo, J., Basarkod, G., Marsh, H., 2021. Ability stratification predicts the size of the big-fish-little-pond effect. *Educ. Res.* 50 (6), 334–344. <https://doi.org/10.3102/0013189X20986176>.
- Parker, P.D., Jerrim, J., Schoon, I., Marsh, H.W., 2016. A multinational study of socioeconomic inequality in expectations for progression to higher education: the role of between-school tracking and ability stratification. *Am. Educ. Res. J.* 53 (1), 6–32. <https://doi.org/10.3102/0002831215621786>.
- Parker, P.D., Marsh, H.W., Jerrim, J.P., Guo, J., Dicke, T., 2018. Inequity and excellence in academic performance: evidence from 27 countries. *Am. Educ. Res. J.* 55 (4), 836–858. <https://doi.org/10.3102/0002831218760213>.
- Perry, L.B., McConney, A., 2010. Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teach. Coll. Rec.* 112 (4), 1137–1162. <https://doi.org/10.1177/016146811011200401>.
- Perry, L.B., Southwell, L., 2014. Access to academic curriculum in Australian secondary schools: a case study of a highly marketised education system. *J. Educ. Pol.* 29 (4), 467–485. <https://doi.org/10.1080/02680939.2013.846414>.
- Reay, D., 1999a. "Class Acts": educational involvement and psycho-sociological class processes. *Fem. Psychol.* 9 (1), 89–106. <https://doi.org/10.1177/0959353599009001008>.
- Reay, D., 1999b. Linguistic capital and home-school relationships: mothers' interactions with their children's primary school teachers. *Acta Sociol.* 42 (2), 159–168. <https://doi.org/10.1177/000169939904200205>.
- Reay, D., 2004. Exclusivity, exclusion, and social class in urban education markets in the United Kingdom. *Urban Educ.* 39 (5), 537–560. <https://doi.org/10.1177/0042085904266925>.
- Reay, D., 2007. "Unruly places": inner-city comprehensives, middle-class imaginaries and working-class children. *Urban Stud.* 44 (7), 1191–1201. <https://doi.org/10.1080/00420980701302965>.
- Rimm-Kaufman, S., Pianta, R., 2000. An ecological perspective on the transition to kindergarten: a theoretical framework to guide empirical research. *J. Appl. Dev. Psychol.* 21 (5), 491–511.
- Rist, R., 1970. Student social class and teacher expectations: the self-fulfilling prophecy in ghetto education. *Harv. Educ. Rev.* 40 (3), 411–451.
- Rist, R.C., 1974. *The Urban School: A Factory for Failure*. MIT Press, Cambridge, MA.
- Robinson, J.P., 2008. Evidence of a differential effect of ability grouping on the reading achievement growth of language-minority hispanics. *Educ. Eval. Pol. Anal.* 30 (2), 141–180. Retrieved from. <http://www.jstor.org.ezproxy.newcastle.edu.au/stable/30128058>.
- Rosenthal, R., Jacobson, L., 1968. Pygmalion in the classroom. *Urban Rev.* 3 (1), 16–20. <https://doi.org/10.1007/BF02322211>.
- Rowe, E., Perry, L.B., 2020. Private financing in urban public schools: inequalities in a stratified education marketplace. *Aust. Educ. Res.* 47 (1), 19–37. <https://doi.org/10.1007/s13384-019-00328-0>.
- Slavin, R.E., 1987. Ability grouping and student achievement in elementary schools: a best-evidence synthesis. *Rev. Educ. Res.* 57 (3), 293–336. <https://doi.org/10.3102/00346543057003293>.
- Slavin, R.E., 1993. Ability grouping in the middle grades: achievement effects and alternatives. *Elem. Sch. J.* 93 (5), 535–552. Retrieved from. <http://www.jstor.org.ezproxy.newcastle.edu.au/stable/1001827>.
- Stirrup, J., Evans, J., Davies, B., 2017. Learning one's place and position through play: social class and educational opportunity in Early Years Education. *Int. J. Early Years Educ.* 25 (4), 343–360. <https://doi.org/10.1080/09669760.2017.1329712>.
- Taylor, B., Hodgen, J., Tereshchenko, A., Gutiérrez, G., 2022. Attainment grouping in English secondary schools: a national survey of current practices. *Res. Pap. Educ.* 37 (2), 199–220. <https://doi.org/10.1080/02671522.2020.1836517>.
- Wang, L.-Y., 2021. Levelling up academically low-performing students in student-centric education in Singapore: global trend, local policies and future directions. *Educ. Rev.* 73 (3), 374–390. <https://doi.org/10.1080/00131911.2019.1642304>.
- Young, M.F.D. (Ed.), 1971. *Knowledge and Control: New Directions for the Sociology of Education*. Collier Macmillan, London.

Working collaboratively with the school community to build inclusion for all

Joan G. Mowat, School of Education, University of Strathclyde, Glasgow, Scotland, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction: The context for inclusion for all	85
Globalization and the neo-liberal agenda	85
Children's rights and Education for All	86
The Salamanca Statement and Framework	86
United Nations Convention on the Rights of the Child (UNCRC) and Education for All	86
The Sustainable Development Goals (SDGs)	86
The impact of inequitable societies on children's wellbeing and life chances	86
The impact of the pandemic on children's education and mental health and wellbeing	87
A state of precarity	87
Covid-19: the learning and digital divide	87
The risks posed to children with disabilities	88
Covid-19 and child protection	88
Inclusion for all	88
Inclusion – a “slippery,” contentious concept	88
Inclusive education	89
Models of disability	89
Who's in? Who's out?	89
Inclusive education: a reconciliation	90
Building inclusion for all collaboratively	91
Examining the boundaries between inclusion and exclusion	91
A rationale for change	91
Frameworks for exploring policy and practice	92
Implications for school leadership	93
Reflections	93
The importance of a sense of belonging to school	93
Building community and empowering the school community	94
Transforming schools to build inclusion for all	94
Acknowledgment	95
References	95
Further Reading	97

Introduction: The context for inclusion for all

Inclusion for all has been a global aspiration, formalized through the Salamanca agreement (UNESCO, 1990). It aligns with the United Nations Convention on the Rights of the Child (UNICEF Office of Research, 1989) and with the United Nations Goals for Sustainable Development to be achieved by 2030 (UNICEF Office of Research, 2017; 2019b). It also aligns with a drive to ensure equity for all children, brought into sharp relief by the Covid-19 pandemic which has brought to the fore, and exacerbated, existing inequalities and inadequacies in the support for children with mental health issues and Special Educational Needs (SENs).

Globalization and the neo-liberal agenda

Within the complex and unpredictable world in which we live, it is crucial to understand the forces that shape and drive educational policy across the world, particularly as the locus for policy-making shifts from the national to the supranational (D'Agnese, 2018). Global organizations such as the Organization for Economic Cooperation and Development (OECD) and the World Bank play a key role in shaping the political agenda. It cannot be denied that much valuable work in addressing inequality across the world has been undertaken by such organisations. It could be argued, however, that, through the power invested in and resources available to them, they are able to set and dictate the agenda.

Increasingly, there are voices of dissent with many commentators critiquing the neo-liberal agenda that frames the values underpinning and context for action – the focus on performativity, marketization and the commodification of education (Ball, 2003; 2015; Connell, 2013; D'Agnese, 2018). This can be seen very clearly in England where there is a plethora of different forms of schooling (Comprehensive, Free, Faith, Grammar and Private Schools, Academy Trusts and City Technology Colleges). Schools compete with each other to attract students who are most likely to enhance their reputation, leading, almost inevitably, to

exclusionary practices for children who are deemed to not add value to the school (Ainscow, Booth and Dyson, 2006a; Slee, 2018b). Thus, children, themselves, become commodities.

This neo-liberal agenda is also evident in the OECD's global standardized tests (such as the Program for International Student Assessment – PISA), leading, ultimately, to a reification of performance in a narrow range of competencies (with little consideration of cultural context) and practices such as teaching to the test (Ball, 2003; 2015; Ball and Noddings, 2014; Clapham et al., 2016; D'Agnesse, 2018; Holloway and Brass, 2018; Moore and Clarke, 2016; Solomon and Lewin, 2016).

What emerges from OECD documentation is a narrative of risk, creating a mandate for change, in which key terms, such as excellence, are rarely defined and key concepts, such as social justice and inclusion, are understood in very restricted and reductionist ways (Mowat, 2018). There is a failure to recognize competing policy imperatives and to recognize that the neo-liberal agenda through which policy is both formulated and enacted and the emphasis on performativity inherent within this almost inevitably create the conditions under which some children will fall by the wayside (Apple, 2001; MacBeath et al., 2007), marginalizing children who are already disadvantaged.

Children's rights and Education for All

The Salamanca Statement and Framework

The *Salamanca Statement and Framework* (UNESCO, 1990) represents one of the key landmarks in the quest for inclusive education, bringing children's rights to the fore. It challenged the orthodoxy at the time that not all children should be educated in regular or mainstream schools; it called for major restructuring of the school system; and it raised consciousness of inclusive education (Flores, 2019), creating a mandate for change (Ainscow et al., 2019). Subsequent to Salamanca, a range of policy initiatives has taken this agenda forward globally from *Education for All* (UNESCO, 1990) to the recent reports emanating from UNESCO (Slee, 2018a; UNESCO, 2017).

United Nations Convention on the Rights of the Child (UNCRC) and Education for All

Two further critical developments have been the UNCRC which through 54 articles sets out the civil, political, economic, social and cultural rights for all children (UNICEF Office of Research, 1989) and the *Education for All* movement. *Education for All*, launched by UNESCO in 1990 and taken forward through the work of UNICEF and organizations such as the World Bank (*Learning for All*), is informed by a human rights perspective and is concerned with increasing access to and participation within education across the world. It seeks to address the two hundred-and-sixty-three million children and young people who are debarred from schooling, fueling intergenerational poverty and disadvantage and robbing their communities of economic growth (UNICEF Office of Research, 2019a).

The Sustainable Development Goals (SDGs)

Building on the earlier developments, the seventeen *Sustainable Development Goals* (UNICEF Office of Research, 2017; 2019b) provide a global framework for addressing key international imperatives including those most pertinent to children, such as ensuring access to quality education, good health and wellbeing and eradicating poverty and reducing inequalities.

The impact of inequitable societies on children's wellbeing and life chances

As has been ascertained, even prior to the pandemic, equitable access to schooling across the world is far from a reality for a vast number of children. Even in more advanced nations, where access to education is less problematic, inequities in the experience and outcomes of schooling abound. The poverty-related attainment gap persists over time, having a lasting impact on life opportunities, such as access to Further and Higher education (OECD, 2018; UNICEF Office of Research, 2010).

This is reflective of broader inequalities in society. Putnam (2015), in his exploration of the demise of the "American Dream," demonstrates an increasingly divided society where the social and economic distance between the "haves" and the "have nots" has widened over time. It is not just children from low-income families who are impacted negatively by inequality. On a range of measures relating to education, health and life satisfaction, the outcomes for children are poorer in more inequitable societies with unequal income distribution, irrespective of family wealth (UNICEF Office of Research, 2016; 2017; Wilkinson and Pickett, 2010).

Amidst rising concerns about the prevalence of mental health issues in society in general, increasingly, concerns are raised about the mental health and wellbeing of children and young people with, at times, devastating consequences (Bruckauf, 2017; Patel et al., 2018; World Health Organisation, 2016). There are particular concerns around the mental health of adolescent girls, a growing trend in many recent surveys (Bruckauf, 2017; OECD, 2017a; 2017b; Patalay and Fitzsimons, 2018). In the UK, over a two-year period, the subjective wellbeing of young people aged 10–15 decreased significantly (The Children's Society, 2020). Those children living in the poorest families fare worse and have the poorest health outcomes. This is more acute for adolescent girls than boys (UNICEF Office of Research, 2016).

There is an indirect relationship between poverty and poor subjective health related to structural inequalities in society which impact on family affluence, limiting children and young people's access to material resources and creating stressors in their lives (UNICEF Office of Research, 2016). Long-term poverty impacts negatively on infant and childhood mortality, poorer child health and incarceration (Cooper and Mulvey, 2015; Mowat, 2019a).

While much of the above is beyond the direct control of schools, these are the realities that schools and their leadership teams face on a daily basis and is crucial to our understanding of how to build sustainable, inclusive school communities.

The impact of the pandemic on children's education and mental health and wellbeing

As previously intimated, the pandemic has served to exacerbate existing inequalities. This impacts on communities, families and individual children.

A state of precarity

Precarity is understood in the social sciences and humanities as an effect that arises from neo-liberal practices, particularly related to work that is temporary, insecure, poorly paid and sporadic in nature. It not only impacts on levels of poverty within communities, but on the day-to-day life of people – their health and wellbeing, manifest in crime-rates, pupil indiscipline, mental health issues, gender inequalities and discrimination, among many other things – and it is these relationships that the concept seeks to capture (Kirk, 2020).

While one might imagine that the financial impact of the pandemic would be experienced similarly across the UK as many people have had to rely on furlough payments, grants and loans from the government and welfare benefits, a report from the financial firm Lane, Clark and Peacock LLP (Lane et al., 2021; Peachey, 2021) indicates that this is not the case. Indeed, six million employees have become “accidental savers,” squirreling away thousands of pounds that otherwise would have been spent on work-related expenses, holidays and leisure activities. Thus, the differential in disposable income between the “haves” and the “have nots” has widened, exacerbating existing inequalities and putting more families into poverty and debt.

In the UK, families who, prior to the pandemic, were on low incomes and who faced the greatest constraints and instability had fewer resources to draw on to help them to “weather the storm.” In particular, those families who were already facing significant disadvantage and who experienced sudden job loss, reduction in hours and had insufficient budget to cover additional costs were particularly affected (Hill and Webber, 2021).

In Scotland, families who were already reliant on welfare benefits reported that their financial situation had worsened and debt increased. 82% of these families had experienced deterioration in their mental health over the course of the pandemic. A key concern for them was support for their children, including helping them to deal with mental health issues (Incudem, 2020). The Child Poverty Action Group (CPAG) found that families were more anxious about finances than prior to the pandemic and 90% were spending more on household bills during lockdown when children were at home. Schools had sought to ameliorate the negative impact of Covid-19 on families, alleviating financial pressure, communicating with parents and children and young people and supporting home learning and the social engagement of children and young people (Child Poverty Action Group, 2021).

Covid-19: the learning and digital divide

Access to schooling

In the eleven months from March the 20th, 2020, more than 168 million children globally were not in attendance at school, representing an average of 95 days lost (United Nations Children's Fund, 2021a). During this time, 214 million children from pre-primary to upper secondary missed at least three-quarters of classroom instruction. The duration of school closures varied significantly between different countries (UNESCO, UNICEF, and the World Bank, 2020; United Nations Children's Fund, 2021). In the period July–October 2020, 31% of the days missed were accounted for by children living in low-income countries in comparison to 14% of children living in high-income countries. Low- and lower-middle income countries were significantly less likely to monitor student learning. When schools did re-open, high-income countries were more likely to adopt a form of blended learning (UNICEF and UN Women and Plan International, 2020).

In the earlier stages of the pandemic, significant differences emerged in the experiences of children attending private, as apart from state, schools in England and in the experiences of children coming from middle-class rather than working-class homes (Cullinane and Rebecca Montacute, 2020). This affected many aspects of children's learning, such as the frequency and regularity of communication with teachers; access to online learning; the degree of satisfaction with support from the school; the confidence of parents to support home learning (related to the level of education of parents); the amount of time devoted to learning; and the amount and quality of work received by teachers. The authors caution that, without measures taken to mitigate the impact of loss of learning for children in disadvantaged communities, the attainment gap will widen even further and the prospects for social mobility dim.

A range of approaches is being adopted globally to mitigate the negative impact of interruptions to schooling, among which are increase in class time and accelerated and remedial programs (U. W. a. P. I. United Nations Children's Fund and UN Women and Plan International, 2020).

The digital divide

Globally, those countries that had the longest period of school closures also tended to have the lowest proportion of children with a fixed internet connection which would have facilitated home schooling (United Nations Children's Fund, 2021). It has been established that 65% of children and young people, under the age of twenty-five, lack internet-access at home. This is dependent on locale and socio-economic circumstance (United Nations Children's Fund and International Telecommunication Union, 2020).

Children living in conflict areas, from rural communities and those children living in poverty were least able to access the technologies that enabled remote learning, exacerbating the impact of the existing digital divide (UNESCO et al., 2020).

Two out of three of the most affluent countries provided free or subsidized access to the internet, but this was not the case in less prosperous countries. Children living in the higher income band countries were also more likely to be in receipt of subsidized or free devices to support online learning. Not surprisingly, remote learning was considered to be much less effective than face-to-face teaching in low-income countries (only 20% considering it to be sufficient, in contrast to 86% in high-income countries) (UNESCO et al., 2020).

In Scotland, a year into the pandemic, 35% of families on low-income were still missing essential resources to support learning, such as laptops, tablets and printers. Even when devices were available, the software was not always compatible with school systems. Children in receipt of free school meals were more likely to have to share devices to access learning with family members and to use a mobile phone to complete schoolwork (Child Poverty Action Group, 2021).

The risks posed to children with disabilities

The pandemic has undoubtedly had a disproportionately negative impact on disabled children across the globe. In particular, the stigma and associated discrimination toward disabled children often leads to increased exposure to abuse and neglect which, together with restricted access to services (health and social care) brought about by Covid-19 restrictions, leads to the magnification of social inequalities. Over 80% of countries in Eastern Europe and Central Asia reported disruption in access to disability-related health services and 25% to social services. At least half had not put in place measures to facilitate learning for disabled children and disability-inclusive responses to the pandemic were largely inadequate (with a few exceptions) (United Nations Children's Fund, 2020a).

In the absence of support networks, the anxieties and depression induced by lockdown conditions may compound existing mental health issues. For some children with intellectual or psychosocial difficulties, coping with disrupted routines may be particularly challenging. The shielding or quarantining of disabled children may place increased pressures on family life, potentially exposing them to the increased risk of violence in the home. Home schooling for all parents and children can be difficult but, without the assistive devices and specialized curricula that support learning for disabled children, it can be particularly challenging (United Nations Children's Fund, 2020a).

Covid-19 and child protection

Globally, Covid-19 has posed significant challenges with regard to protecting children from neglect, abuse and violence. Increased risk may arise from loss of care for children due to family bereavement, illness or separation; the stressors on family life arising from financial insecurity (as previously discussed); and increased care-giving responsibilities at a time when family supports are reduced. This is compounded by disruption to normal child protection processes and procedures, leaving many families and children vulnerable. The lack of contact with families and children due to the imposition of lockdown has made it more difficult for teachers and those who come into contact with children to pick up the signs of neglect and abuse, leading to a reduction in referrals to child protection services (United Nations Children's Fund, 2020b).

Previous pandemics and world conflicts have led to disruption of services and Covid-19 has been no exception with welfare services (48%); violence prevention programs (49%); case management services/referral pathways (52%) and household visits to children and women at risk of abuse (53%) affected. The United Nations advocates that child protection services should be designated as essential work with governments prioritizing critical prevention and response services to protect children from violence (United Nations Children's Fund, 2020b).

Thus, it can be seen that, across the world, children and young people have been marginalized in many different ways with significant impacts on their lives and life-chances both pre- and post-pandemic.

Inclusion for all

The discussion to this point has set the context for the discussion to follow, highlighting the scale and complexity of the problem to be overcome. The next part of the discussion will focus on building inclusion for all, exploring initially how the concept is understood and the inherent tensions and dilemmas relating to it.

Inclusion – a “slippery,” contentious concept

On the 25th anniversary of the signing of the Salamanca Statement, commentators still draw attention to the ambivalent, paradoxical nature of the construct and the controversy associated with it (Corcoran et al., 2019; Kinsella, 2020; Larsen et al., 2019; Slee, 2018a); controversy captured within the heated debates about the place of special schooling; models of disability; and ableism/disablism within the field of critical disability studies (Norwich, 2013; Shakespeare, 2014).

It has been argued that how inclusion is understood and put into practice must take account of geo-political factors, culture and context (Slee, 2018a). Across the developing world, where there is conflict, poverty and displaced populations, the challenges in achieving inclusion for all are significant (Ainscow and Sandill, 2010; Slee, 2018a). What may “count” as inclusion in the context of developing nations (for example, separate provision for children who are disabled) could potentially be construed as exclusionary in more developed nations (Florian, 2019).

The typology of inclusion developed by Mel Ainscow and colleagues (Ainscow, Booth and Dyson, 2006a) identifies six ways in which inclusion is commonly understood. These range from a narrow conceptualization as pertaining solely to disability and children with SEN to a much more embracing *principled approach to education and society* – a values-led approach, concerned with the promotion of equity, participation, community, compassion, respect for diversity, sustainability and entitlement. Such an understanding moves the focus away from more restricted understandings of inclusion as being primarily about children with disabilities (Florian, 2019). However, Norwich (2013) cautions that the latter conceptualization could lead to an over-simplification of the differences between different facets of diversity with the needs of children with SEN being subsumed within other categories.

Inclusive education

Norwich (2013) argues that there is no single, commonly agreed understanding of what constitutes inclusive education, resulting in ambiguity around its realization. It is conceptualized in a multiplicity of different ways as being concerned with: accepting/valuing/ extending scope to all; equal opportunities; voice and empowerment; a never-ending journey; and/or as a means toward an inclusive society.

In contrast, Loreman (2009) maintains that inclusive education has commonly agreed distinctive features, such as a recognition that all children should follow substantively similar programs of study while in attendance at their local school, with appropriate modifications made to curricula. Florian et al. (2017), concerned about the “repetition of exclusion” (p. 695) represented within separate provision for children with SEN, suggest that, rather than offering separate pedagogies for individual groups of children that could potentially stigmatize them, inclusive education should, instead, focus on extending and making accessible the provision which is on offer to children of a similar age within the classroom setting.

These differing perceptions lead to tensions within the field, manifest within the controversies surrounding special education; the extraction of children from the classroom for additional support; and the concerns around the stigmatization and labeling of children (Hornby and Evans, 2014; Mowat, 2015a; Slee, 2013). Norwich (2013) draws on the earlier work of Minow (1990) to elaborate on the “Dilemma of Difference” encapsulated in the following two questions – “When does treating people differently emphasize their differences and stigmatize or hinder them on that basis? And when does treating people the same become insensitive to their difference and likely to stigmatize or hinder them on that basis?” (p. 7).

A further tension lies in the relationship between inclusive education and SEN. While the Warnock report on SEN in England (Department for Education and Skills, 1978) initially led to a transformation in how children with disabilities were understood and educated, extending in influence within and far beyond the UK, ultimately, it led to the reinforcement of categories of impairment and disability (Slee, 2018b), represented in subsequent government reports (Norwich, 2013).

It has been argued that there has been a blurring of the distinction between the concepts of inclusion and SEN (Florian, 2019; Slee, 2018a) with the language of inclusion being appropriated by the latter (Slee, 2018b). Florian (2019) questions the degree to which SEN can be considered to be part of the solution or part of the problem in achieving an equitable education system.

Models of disability

Those who are advocates for inclusive education argue for a transformation of schools and their policies and practices rather than a focus on individual difference and deficit (the Individual/Medical/Deficit (IMD) model of disability), paying insufficient attention to the social, cultural and environmental aspects which act as a barrier to participation in society. They draw on the ideology underpinning the social model of disability (Oliver and Barnes, 2012) which makes a distinction between impairment and disablement and focuses on removing barriers to participation and learning, represented within the Index of Inclusion (Booth and Ainscow, 2011) which has exerted considerable influence across the world.

It is argued that the reification of what it is to be able-bodied (“ableism” – a key tenet of critical disability theory) together with the social imposition of restrictions on the lives of disabled people and the undermining of their socio-emotional wellbeing is a form of oppression and discrimination which is culturally situated (Slee, 2018b; Thomas, 2007).

However, while recognizing the many benefits of the social model in ameliorating the discrimination that many disabled people experience in their lives, Shakespeare (2014) argues that the alignment of the social model as a progressive approach to disability and the IMD with everything that is bad is an over-simplification of a complex interaction of the many factors which come together in the life of a disabled person – biological, psychological, cultural and socio-cultural – which cannot easily be extricated from each other. For him, the social model does not provide a full explanation for or encapsulate their life experiences.

Who's in? Who's out?

On a global basis, there is little consistency in terms of how SEN is understood and enacted and a lack of international coherence in the provision for children. The only common features appear to be that children so identified are marginalized and that “special provision” constitutes something that is additional to that which is on offer for other children (Rix et al., 2013). Likewise, Hardy and Woodcock (2015) found little commonality internationally in how inclusion is understood and enacted within public policy, finding a policy landscape characterized by fragmentation, incoherence, inconsistency or discrimination.

A range of commentators draw attention to the disproportionality of children identified with SEN across the world, with some specific groups over-represented – principally children from ethnic, racial and immigrant populations (Dyson and Kozleski, 2008; Slee, 2018b; Tomlinson, 2016) and those living in poverty (Mowat, 2019a, 2020; Slee, 2018a). These structural inequalities are sustained and recreated through the systems, structures and policies of government (Dorling and Tomlinson, 2016).

Despite the frequent claims by the OECD that the pursuit of excellence and equity can go hand in hand (Schleicher, 2014), many commentators are concerned that the neo-liberal agenda underpinning such efforts and the unrelenting focus on “excellence” and market competition lead, almost inevitably, to exclusionary teaching practices across the world (Razer et al., 2013); to “winners and losers” (Barton, 2019); barriers to participation for those children considered not to “add value” (Ainscow et al., 2006b); and “hidden” as well as actual exclusions from school (McCluskey et al., 2019).

Slee (2018a, 2018b and 2019) is deeply concerned about the exclusionary nature of much education provision throughout the world and the reproduction of inequalities that it represents, reinforcing privilege and exacerbating disadvantage. In particular, he draws attention to the exclusionary clauses in education policy relating to children’s rights; the creation of educational élites; the ever-expanding categorization of disabilities (as exemplified in the *Diagnostic and Statistical Manual of Mental Disorders (DSM – 5)*); and exclusionary practices masquerading as inclusive education. Rather than schools being places of belonging for all children, schools have become a place for some but not all children (Slee, 2019). He questions whether schools have truly transformed themselves or “has special education relocated itself in the regular school?” (Slee, 2018a).

A recent article in *The Guardian* newspaper (J. Harris, 2021), drew on the experiences of a young person being placed in a “Reflection Center,” described by the headteacher as, “Quite a welcoming room where the students have an individual teacher supporting them”, but described by the student as being placed in a booth where, “you can’t talk to anybody, you can’t stand up if you don’t ask. Literally, it’s like being in prison” (p. 33). This experience is far from unique – similar experiences have been recounted by other commentators (e.g., Riley, 2017; Riley and Rustique-Forrester, 2002).

Slee (2018a) identifies four questions that should inform the development of inclusive education and which cut to the heart of the matter:

- What kind of world do we want our children and young people to live in?
- What kind of schools and classrooms are required to achieve that world?
- What do children need to know, think and be able to achieve a more inclusive world?
- When we look at our schools – who is in, who is out, who decides and what are we going to do about it? (p. 9)

Inclusive education: a reconciliation

It will be evident from the discussion to this point that the various perspectives that have been forwarded cannot easily be reconciled. A key tension is the degree to which inclusion can be seen to be limitless (i.e., any boundaries placed around the concept are seen as exclusionary). Some commentators are unequivocal in their view that there should be no limits to inclusion, with any boundaries around the concept being seen as discriminatory (Slee, 2018b). In contrast, Shakespeare (2014), while recognizing the benefits of the removal of barriers to participation and inclusive provision, believes that, for some children, alternative forms of provision may meet their needs more effectively. Florian (2019) argues that, in a post-Salamanca era, the key focus should be on valuing diversity, recognizing the uniqueness of each individual and their shared humanity: it becomes “an imperative of practice” rather than a secondary consideration (p. 701).

Inclusion can only be fully understood when consideration is given to its alter ego, exclusion – what it is that one wishes children not to be excluded or marginalized from. Likewise, marginalization and exclusion can only be fully understood when one gives consideration to what it is desired that children should be included within (Mowat, 2008), as per the questions posed by Slee. These, clearly, are value-positions that will be determined by cultural norms and the worldviews of people individually and collectively. As reflected in Slee’s last question (When we look at our schools – who is in, who is out, who decides and what are we going to do about it?), questions can be posed about who determines what might constitute inclusive or exclusive practice and by which criteria.

On the basis that a society cannot be truly inclusive unless it is tolerant of practices of which it disapproves, Hansen (2012) argues that “inclusion presupposes exclusion” (pp. 93–4):

So, it is not possible to consider inclusion in itself by excluding its otherness: exclusion, neither by defining a normative limit to inclusion nor by avoiding limits and making it unambiguous. From this point of view, it will never be possible to put meaning into the concept of inclusion without its otherness, exclusion.

She identifies a gap between the vision and rhetoric of inclusion as a limitless principle and practice on the ground: a gap that is not theorized or reflected within the concept. Any perceived failings are explained in terms of external factors – for example, IMD explanations; theories of stigma and marginalization, leading to a failure to critique the concept itself. In her view, how inclusion and exclusion are seen is dependent on the interpretation of the individual, shaped by the professional identities of teachers and the discourse surrounding their practice: “... inclusion or exclusion are dependent on how the professional constructs the specific situation and needs of a specific child” (p. 96).

Rather than seeing inclusion and exclusion as being polarities at opposite ends of the spectrum, might it not be more helpful to see them as lying on a continuum? – not as absolutes but as social constructs that are in a constant dialectical relationship with each other. This position recognizes that there may be boundaries between inclusion and exclusion, but these

boundaries are in a constant state of flux. They are influenced not only by the situated nature of the context but by the interpretation of the professional and the subjective experience of the individual child – how they see their world. What may be experienced as inclusionary or exclusionary by one child may not be experienced in the same way by another; what may be perceived by others in the life of the child as being inclusionary or exclusionary may not be experienced by the child in this way (Messiou, 2012b; Mowat, 2015a, 2015b); and, as argued by Florian (2019), what may be regarded as inclusive practice in one society may be regarded as exclusionary in another. Ultimately, it comes down to the interaction between what is valued by the child and family and what is valued by society, and what society affords, and this is socially, historically and culturally determined (Mowat, 2015b).

The boundaries between inclusion and exclusion and between inclusive and exclusive practice may be much more complex and subtle than might otherwise be assumed to be the case. Hansen (2012) argues that the issue then becomes one of understanding the nature of the boundaries within the school setting and how they are shaped and determined in order to promote inclusive practice.

Building inclusion for all collaboratively

The chapter was initially framed within a global context examining a range of factors that impact on the lives of children and young people and their education. Schools are part of a wider eco-system. If inclusion for all is to be realized, policy, at the international, national and local levels, needs to coalesce. However, it is at the level of the school community that children and young people will experience their schooling as inclusionary or exclusionary, therefore the discussion to follow focuses on that level.

The discussion will initially examine two approaches that explore the boundaries between exclusion and inclusion: the first, an exploration of marginalization (Messiou, 2012b); the second, an exploration of place and belonging (Riley, 2017). The approaches, drawing on empirical research, have been selected because they have theoretical and practical insights of value for researchers and practitioners.

Examining the boundaries between inclusion and exclusion

A rationale for change

Messiou (2012b) is primarily concerned with exploring how marginalization is experienced by children and young people. She poses the question, “How can schools engage with students’ voices in order to confront marginalization and promote inclusion?” (p. 5). Riley (2017) approaches the topic from the perspective of exploring the degree to, and ways in which, school is experienced by children and young people, and adults, as a place of belonging and the role of the school community in creating schools as places of belonging and opportunity. She defines a sense of belonging as, “... that sense of being somewhere where you feel confident that you will fit in and feel safe in your identity” (p. 4). Being tied in with an individual’s sense of identity, it therefore means different things to different people. It is not a static state, but something that is dynamic and an evolving experience – it is about “... being comfortable in yourself, who you are and where you are” (p. 28). It can be seen therefore that the aims of each author coalesce, although Riley focuses more broadly on all members of the community.

For both authors, the key issue that needs to be addressed relates to the marginalization and discrimination that children and young people face. Messiou (2012b) is concerned that the dominant deficit perspective on children and young people who are categorized as belonging to specific groups, and the labeling associated with it, lead to a failure to recognize the individuality of the child and deflect attention away from contextual factors which create the conditions under which children and young people become marginalized. Like Florian (2019) and Slee (2018a), she is concerned also by the conflation of inclusive education with categories of need in government policy. Recognizing that the actions of teachers can serve to disenfranchise children and young people, and concerned about a range of issues relating to diversity and the learning and social relationships of children and young people in schools, the author asks, “What more can be done in schools?”

Messiou (2012b) considers marginalization to be a complex, abstract concept that can be understood in a range of different ways and is contextually situated. Whether in reference to the social or academic aspects of children and young people’s lives, it can be both overt and covert (not always recognized as such). It is socially constructed through the interactions and labeling that occurs in social groups. Those who deviate from the norm are considered to be marginal to the group. However, the boundaries are not rigid or static, but fluid in nature.

She categorizes marginalization in four ways, encapsulating both the experience of marginalization and the recognition of it (or not) by the child and/or others (Messiou, 2012a, 2012b, 2019). She argues that the subjective experience of the individual is not given sufficient consideration. This means that there can be a disparity between the child’s perception of whether or not they experience their life as marginalized and the perceptions of others. Children and young people who, in the eyes of others, are marginalized may not experience their lives in this way (or may be in a state of denial). There may be other children and young people who experience their lives as marginalized but it is not recognized as such, by others. Thus, labeling cannot be regarded as synonymous with marginalization (Messiou, 2012b).

Riley (2017) draws attention to the divided and chaotic lives that many children lead in a turbulent world characterized by “global churn,” marginalized through poverty, dislocation and ethnicity, creating “insiders” and “outsiders” – “those who felt

safe and those who didn't" (p. 39). She observes a "them and us" culture that characterizes some schools where children and young people are not valued for whom they are:

... being forced as a child or young person to conform to a school ethos which denies who you are in an enforced process of "belonging" or in which you are ostracized by staff or students, or bullied, is to deny young people their identity and start an unfolding chain of exclusion: within school and beyond the school gates (Riley, 2017, p. 38).

Messiou (2012b) believes in the transformational power of inclusive education, enabling teachers to reach out to all children, irrespective of the labels assigned to them. Riley (2017) believes that schools can serve as places of safety where values can be explored and trust built. They can help children and young people to "... face the unknown, to understand the complexities of the past and present" (p. 23). They can be places of opportunity where all children and young people can flourish. In her view, how schools reach out to their communities – and view their cultures, practices and beliefs – will shape the identities of children and young people within the school community, fostering their engagement with the world. If schools are to become places of belonging and opportunity, they need to have an understanding of the lives of children beyond the school gates and to be friendly, welcoming environments for all.

Frameworks for exploring policy and practice

Both authors propose a collaborative, enquiry-led approach in taking the work forward in schools based on frameworks, deriving from empirical research, to act as a scaffold for enquiry, reflection, learning, policy and practice.

Messiou (2012b) forwards a four-step framework to be carried out with researchers working alongside participants within schools (who could be staff and/or students) to investigate marginalization, using a wide range of data-collection methods – observation, interviews, questionnaires, sociometric measures, visual images and activities, such as role-play. The framework (cc. Table 1) focuses primarily on the roles of students and teachers and the collaboration between them in addressing marginalization.

Riley (2017) proposes a three-step framework (cc. Table 1) in a highly participative process with school leaders, teacher and student researchers exploring the themes of place and belonging through a range of research methods, including visual methods and activities. She emphasizes the role of school leaders as "Place Makers" (see discussion to follow) who, through working through the framework, make sense of a complex range of realities.

While framed in different ways, there are strong commonalities between the two conceptualizations. Initially, the frameworks are concerned with coming to a deep understanding of context and providing the opportunities for voices to be heard through a range of mechanisms. Creating a culture of trust and transparency is crucial at this initial stage. The second stage is about making sense of the data in a collaborative process to identify key messages. The final step is about putting into practice what has been learned to create inclusive school communities.

Messiou (2012b) creates a strong rationale for her framework, seeing it as a means of ensuring that issues of marginalization are identified and addressed within the school community, bringing to the surface features that are important for achieving inclusion for all, reconstructing and reimagine schooling, particularly for those children and young people most disadvantaged (Messiou, 2012b).

Riley (2017) argues that, if students are to engage with the world in an enquiring manner, there is a need to build capacity in the system through investing in the professional development of teachers as change agents while simultaneously involving children and young people meaningfully in decision-making that affects their education and welfare. Through "unlocking" how children and young people feel and experience their lives in school and establishing which children are on the "inside" and "outside," the potential of children and young people can be realized.

Both Messiou (2012b) and Riley (2017) recognize the importance of student voice and agency in shaping the lives of children and young people, reflected in a growing emphasis on democracy, citizenship and children's rights. However, Messiou (2012b) observes that approaches toward promoting student voice can be mechanistic and tokenistic in nature rather than authentic. The

Table 1 Derived from Fig. 1.2 (Messiou, 2012b) and Box 9.2 (Riley, 2017).

Framework for promoting inclusion Messiou (2012b)		Framework for Belonging (Riley, 2017)	
Step 1	Opening doors: Enabling voices to emerge	Step 1	Making sense of the context
Step 2	Looking closely: Bringing concerns to the surface	Step 2	Looking through the prism of place and belonging to understand the importance of place and belonging
Step 3	Making sense of the evidence: Sharing data with learners		
Step 4	Dealing with marginalization: Encouraging inclusive thinking and practice	Step 3	Being and becoming a place maker, building trust, developing the knowledge and capacities of staff; and harnessing the creative potential of young people, to create schools which are places of belonging and inclusion (p. 138).

framework, in her view, provides a mechanism for the voices of those children and young people who do not belong to readily identified groups to have their voices heard and for perspectives which might not normally be aired to be legitimized. It has the potential to truly value what children and young people have to say, empowering them and affirming them as valued members of the school community (p. 27).

Implications for school leadership

Messiou (2012b) discusses the role of school leaders in being supportive of and facilitating the process of gaining commitment from staff toward the authentic engagement with students' voices. Riley (2017) describes the role of school leaders as "Place Makers" who, through their mediating influence, shape the norms, beliefs and expectations of the school community, where all members of the community can be valued for whom they are. They create spaces where children and young people can freely express their views while also balancing the external demands of the school and working to create a sense of safety, belonging and agency (p. 136).

Reflections

Key themes to have emerged in the discussion of the two frameworks are the importance of participative, collaborative enquiry-led approaches in working with schools; student voice and agency; teacher professionalism, voice and agency; a sense of belonging and connectedness to school; the primacy of relationships and trust; building bridges with the local community, understanding at a deep level how children and young people experience their lives in and out of school; and the role of school leadership in creating the conditions for inclusive school communities where children and young people can flourish. The discussion to follow will examine some of these themes in greater depth.

The importance of a sense of belonging to school

Having a sense of belonging to school has been identified with a wide range of positive indicators and outcomes related not only to academic achievement (enhancing children's sense of self-efficacy, enjoyment of school and intrinsic motivation toward learning) but positive school transitions, social relationships and mental health and wellbeing, ameliorating the risk-factors in their lives (Berryman and Eley, 2019; Inchley et al., 2020; Mays et al., 2018; Mowat, 2019b; 2019c; Prince and Hadwin, 2013; Riley, 2017; Riley and Rustique-Forrester, 2002; Slee, 2019; Vetoniemi and Kärnä, 2019).

For children with SEN, a sense of belonging is a fundamental psychological need that may be integral to their motivational development and engagement in learning. If lacking, it may lead to a range of negative outcomes, such as behavioral problems and poor emotional and mental health (Prince and Hadwin, 2013). In exploring the experience of children with SEN and their mothers (Connor and Berman, 2019; Tsovoka and Tarr, 2012), it becomes clearly evident that the children's inclusion in social activities and events, whether in or out of school, and the degree to which schools and teachers facilitate participation in learning are key factors in the sense of belonging and connectedness felt by the children. A sense of being accepted and appreciated for whom they are not only enables children and young people to feel part of the school community but enhances the understanding of that community (Connor and Berman, 2019).

For Slee (2019) belonging is key to inclusion, at the heart of which lies owning (taking responsibility for) and valuing all children. A concept that closely aligns with this stance is that of *unconditional positive regard*, associated with Carl Rogers, the father of *Person-Centered Therapy*. A central tenet of *unconditional positive regard* is the need to value all people unconditionally, based on acceptance and respect, characterized by warmth and sympathy (Bozarth, 2007). What this means is that acceptance into the school community should not be dependent on certain characteristics of the child, achievement or conformity to accepted norms of behavior. Where it is seen that some children are only accepted into the school community conditionally it gives the message to all children that they also are only accepted conditionally (Khon, 2005). It has been established that teachers respond more warmly to children whose values are in line with those of the school (Hedegaard-Soerensen and Grumloese, 2020), thus valuing all children for whom they are may present as a challenge to some schools and some teachers where the values of the home and school differ.

Warm relationships between children and between children and teachers and participation in school activities are essential elements of a sense of belonging and connectedness to school (Berryman and Eley, 2019). However, the authors caution that a relentless focus on excellence and equity (excellence for all) without giving due consideration to a sense of belonging and cultural identity may lead to restricted understandings, policy and practice, culminating in the marginalization of some children.

School systems that focus on performativity at the expense of the holistic growth and development of the child, often coupled with zero-tolerance approaches toward school discipline, militate against a sense of belonging and connectedness to school. Unlikely as it might seem, even approaches which are advocated in the name of promoting positive behavior, such as the use of behavior charts; gold stars and stickers; and rewards (Jenny Mosley's "Golden Time" by which children have to earn access to desired activities through adherence to the "Golden Rules," being one such example) serve to disenfranchise and shame those children who rarely or never gain the desired reward (Jung and Smith, 2018; Khon, 1999; McLean, 2003). These children are being given a clear message that their acceptance into the school community is conditional. In the process, it gives a strong message to all children and young people that it is acceptable to exclude other children from their activities.

Building community and empowering the school community

Both [Messiou \(2012b\)](#) and [Riley \(2017\)](#) build compulsive cases for involving the school community not as objects or subjects of research but as fully engaged participants within the research process. They built on the earlier work of pioneers of student voice and participation such as Jean Ruddick and Roger Hart (“Ladder of Participation”) and the innovative work of Mary Kellett in re-imagining research with children. Engaging in participative enquiry is a very powerful means of building community and a vision for improvement within a school but it is not the only means.

[Khon \(2001\)](#) argues that “a community rests on the knowledge of, and connections among, the individuals who are part of it” (p. 113). He also argues that community, in itself, is not sufficient: when community and autonomy come together, they create democracy. [Sergiovanni \(1994\)](#) proposes that relationships within communities are characterized by shared understandings and obligations toward others, an endeavor based on schools defining their own life and creating their own practices of schooling. Thus, building community is based on the principles of democracy; of respecting human rights; of giving people a voice; on agency and autonomy; mutual understandings and obligations toward others; connectedness; and creating a vision for the school which reaches out across the school community and beyond.

Parental engagement

Parental engagement is a key aspect to building community in schools, furthering the learning of children and young people and creating equitable and inclusive school systems. Yet, there are significant social, cultural and economic barriers to parental engagement, making it difficult for schools to reach out to parents ([Goodall, 2018](#); [Goodall and Montgomery, 2014](#); [Harris and Goodall, 2008](#)). This is particularly the case for families who may experience their lives as marginalized.

The narratives of parents of children who are disabled, as represented in the accounts of [Tsovoka and Tarr \(2012\)](#), are sometimes harrowing and it is evident that they are disempowered, often feeling on the margins of decision making and frustrated on behalf of their children. The impact of the pandemic in the UK (as previously discussed) severed many of the support systems for families of disabled children and children and young people who were shielding lost the opportunity to attend school and mix with their peers, even when schooling resumed.

A current small-scale empirical study, which I am leading with colleagues at the University of Strathclyde, focuses on how current and former prospective students on the Into Headship programme (a national programme for prospective headteachers (Principals) in Scotland) supported their school communities during the initial stages of the pandemic when schools were closed. A strong emerging theme is that of a changing conceptualization of their identity and role: many assumed responsibility not only for the wellbeing of students, but for their families also:

It has been so hard not “seeing” our children and this highlighted the importance and value of the relationships we build with families ... we felt an overwhelming sense of responsibility for our whole school community and at times struggled to accept that there were limitations in what we could actually do. We had to support staff and families without overwhelming them and at times this was really difficult to do as everyone was reacting and coping in different ways. Due to this a uniform approach for all wouldn’t work so we had to make professional judgements with the knowledge and information we had on individuals before and during lockdown. (Primary School Depute Headteacher, Scotland).

The above epitomizes for me the essence of an inclusive school community. The care and compassion for children and young people and families is self-evident and, as highlighted by [Riley \(2017\)](#), it rests on an in-depth knowledge and understanding of individual families and their circumstances – it’s not a “one-size-fits-all” approach.

Transforming schools to build inclusion for all

This chapter has highlighted the global drive and vision for inclusion for all, emanating from the late 20th Century, and the trials, tribulations and tensions in its realization. Having passed the 25th anniversary of the signing of the Salamanca Statement, it would be easy to focus on what has not been achieved but, in my view, it is remarkable what has been achieved. The pandemic has undoubtedly widened social, educational and health inequalities in society but, as illustrated in the cameo above (which was representative of many of the responses), it has also brought out what is best in society, strengthening a sense of community and compassion. The challenge will be to build on this in the future as we emerge toward what has been described as the “new normal.” It is an opportunity to re-imagine schools.

There are many other aspects of building inclusion for all which could have been included within this discussion, for example, the importance of inter-professional working and the role of mental health services in supporting children and families. Mel Ainscow and colleagues have examined inclusion and inclusive practice from a systems perspective ([Ainscow, 2012, 2015](#); [Ainscow et al., 2012](#)), exploring issues around equity and school improvement. However, an area, which, in my view, requires much greater attention, is the role of school leadership in creating inclusive school communities and the professional development of headteachers (pre-and post-appointment) toward this end. Riley is one of the few scholars in the field who brings this perspective to her work.

It is not enough to be committed to inclusive practice, important as this is. Theories of change management indicate that a great deal of groundwork needs to be done before the transformation of a school becomes possible. According to [Fullan \(2016\)](#), “... an effective change process shapes and reshapes good ideas while it builds capacity and ownership among members in the group”

(p. 46). Key elements of capacity building are having a thorough knowledge and understanding of the context and culture of the school and its community; building capacity in the school through democratic, ethical and shared forms of leadership where members of the school community are able to exercise autonomy and agency; building a collaborative vision for the school, built on values which guide the school community; creating the spaces and places for participation where individual voices can be heard and ensuring that the mechanisms are in place for them to be acted upon; aligning resources to where they are needed; and building bridges and partnerships between the school and its community, drawing on the strengths of the community (Coates, 2017; Fullan, 2016; Fuller, 2012; Gross and Shapiro, 2016; Hargreaves and Fullan, 2012; Liasidou and Antoniou, 2015; Riley, 2017; Theoharis and Scanlan, 2015). Fundamental to all of the above are the values of integrity, trust, respect for self and others, transparency, care and compassion, social justice and a sense that we all have a responsibility toward each other.

While the latter part of this discussion has focused on what schools can do to challenge marginalization and build inclusive communities, schools alone cannot achieve inclusion for all. It will require a concerted, focused and collaborative global effort where educational, social and fiscal policy can coalesce to create a vision whereby inclusion for all can become a reality rather than an aspiration.

Please note that this chapter was prepared in synergy with a research monograph for Routledge which focusses on the creation of equitable, inclusive, compassionate, and empowering school communities which is due to be published in 2023. Some of the arguments in the research monograph have been condensed within this chapter and a small amount of material reproduced.

Acknowledgment

I would like to thank Professor Roger Slee, Diamond Jubilee Chair in Disability and Inclusion, University of Leeds, for his support and encouragement of my work.

References

- Ainscow, M., 2012. *Developing Equitable Education Systems*. SAGE, London.
- Ainscow, M., 2015. *Towards Self-Improving School Systems: Lessons from a City Challenge*. Routledge, London.
- Ainscow, M., Booth, T., Dyson, A., 2006a. *Improving Schools, Developing Inclusion*. Routledge, London.
- Ainscow, M., Booth, T., Dyson, A., 2006b. Inclusion and the standards agenda: negotiating policy pressures in England. *Int. J. Incl. Educ.* 10 (4–5), 295–308.
- Ainscow, M., Dyson, A., Goldrick, S., West, M., 2012. Making schools effective for all: rethinking the task. *Sch. Leader. Manag.* 32 (3), 197–213.
- Ainscow, M., Sandill, A., 2010. Developing inclusive education systems: the role of organisational cultures and leadership. *Int. J. Incl. Educ.* 14 (4), 401–416.
- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676. <https://doi.org/10.1080/13603116.2019.1622800>.
- Apple, M.W., 2001. Comparing neo-liberal projects and inequality in education. *Comp. Educ.* 37 (4), 409–423.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *J. Educ. Pol.* 18 (2), 215–228.
- Ball, S.J., 2015. Education, governance and the tyranny of numbers. *J. Educ. Pol.* 30 (3), 299–301.
- Ball, S.J., Noddings, N., 06/05/2014. OECD and PISA Tests are Damaging Education Worldwide – Academics. *The Guardian*. <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>.
- Barton, L., 2019. The politics of education for all. In: Rix, J., Nind, M., Sheehy, K., Simmons, K., Walsh, C. (Eds.), *Equality, Participation and Inclusion: Diverse Perspectives*. Routledge, London, pp. 90–98.
- Berryman, M., Eley, E., 2019. Student belonging: critical relationships and responsibilities. *Int. J. Incl. Educ.* 23 (9), 985–1001. <https://doi.org/10.1080/13603116.2019.1602365>.
- Booth, T., Ainscow, M., 2011. *The Index for Inclusion: Developing Learning and Participation in Schools*, third ed. CSIE, Bristol.
- Bozarth, G., 2007. Unconditional positive regard. In: Cooper, M., O'Hara, M., Schmid, P.F., Wyatt, G. (Eds.), *The Handbook of Person-Centred Psychotherapy and Counselling*. Palgrave Macmillan, New York, pp. 182–193.
- Bruckauf, Z., 2017. *Adolescents' Mental Health: Out of the Shadows. Evidence on Psychological Well-Being of 11-15-Year-Olds from 31 Industrialized Countries*. Innocenti Research Brief (Vol. 2017–12). Florence, Italy.
- Child Poverty Action Group, 2021. *The Cost of Learning in Lockdown: March 2021 Update (Scotland)* (Glasgow).
- Clapham, A., Vickers, R., Eldridge, J., 2016. Legitimation, performativity and the tyranny of a 'hijacked' word. *J. Educ. Pol.* 31 (6), 757–772.
- Coates, M., 2017. Setting direction: vision, values and culture. In: Earley, P., Greany, T. (Eds.), *School Leadership and Education System Reform*. Bloomsbury, London, pp. 90–99.
- Connell, R., 2013. The neoliberal cascade and education: an essay on the market agenda and its consequences. *Crit. Stud. Educ.* 54 (2), 99–112. <https://doi.org/10.1080/17508487.2013.776990>.
- Connor, D.J., Berman, D., 2019. (Be)Longing: a family's desire for authentic inclusion. *Int. J. Incl. Educ.* 23 (9), 923–936. <https://doi.org/10.1080/13603116.2019.1602361>.
- Cooper, B.S., Mulvey, J.D., 2015. Connecting education, welfare, and health for American families. *Peabody J. Educ.* 90 (5), 659–676. <https://doi.org/10.1080/0161956X.2015.1087776>.
- Corcoran, T., Claiborne, L., Whitburn, B., 2019. Paradoxes in inclusive education: a necessary condition of relationality? *Int. J. Incl. Educ.* 23 (10), 1003–1016. <https://doi.org/10.1080/13603116.2019.1625453>.
- Cullinane, C., Rebecca Montacute, R., 2020. *COVID-19 and Social Mobility. Impact brief #1: School Shutdown*. Sutton Trust, London.
- D'Agnese, V., 2018. *Reclaiming Education in the Age of PISA: Challenging OECD's Educational Order*. Routledge, London.
- Department for Education and Skills, 1978. *Special Educational Needs: The Warnock Report* (London).
- Dorling, D., Tomlinson, C., 2016. The creation of inequality: myths of potential and ability. *J. Crit. Educ. Pol. Stud.* 14 (3), 59–79.
- Dyson, A., Kozleski, E.B., 2008. Disproportionality in special education – a transatlantic phenomenon. In: Florian, L., McLaughlin, M.J. (Eds.), *Disability Classification in Education: Issues and Perspectives*. Corwin Press, Thousand Oaks, CA, pp. 170–190.
- Florian, L., 2019. On the necessary co-existence of special and inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 691–704. <https://doi.org/10.1080/13603116.2019.1622801>.
- Florian, L., Black-Hawkins, K., Rouse, M., 2017. *Achievement and Inclusion in Schools*. Routledge, London.
- Fullan, M., 2016. The path to equity: whole system change. In: Blankstein, A.M., Noguera, P. (Eds.), *Excellence through Equity: Five Principles of Courageous Leadership to Guide Achievement for Every Student*. ASCD, Alexandria, USA, pp. 45–54.
- Fuller, K., 2012. Leading with emancipatory intent: headteachers' approaches to pupil diversity. *Educ. Manag. Adm. Leader* 40 (6), 672–689.

- Goodall, J., 2018. Narrowing the Achievement Gap: Parental Engagement with Children's Learning. Routledge, London.
- Goodall, J., Montgomery, C., 2014. Parental involvement to parental engagement: a continuum. *Educ. Rev.* 66 (4), 399–410. <https://doi.org/10.1080/00131911.2013.781576>.
- Gross, S.J., Shapiro, J.P., 2016. Democratic Ethical Educational Leadership: Reclaiming School Reform. Routledge, London.
- Hansen, J.H., 2012. Limits to inclusion. *Int. J. Incl. Educ.* 16 (1), 89–98.
- Hardy, I., Woodcock, S., 2015. Inclusive education policies: discourses of difference, diversity and deficit. *Int. J. Incl. Educ.* 19 (2), 141–164.
- Hargreaves, A., Fullan, M., 2012. Professional Capital: Transforming Teaching in Every School. Routledge, London.
- Harris, A., Goodall, J., 2008. Do parents know they matter? Engaging all parents in learning. *Educ. Res. Rev.* 50 (3), 277–289. <https://doi.org/10.1080/00131880802309424>.
- Harris, J., 2021. A class of their own. *Guardian* 30–35. <https://www.theguardian.com/education/2021/jan/30/whats-behind-the-rise-in-english-school-exclusions>.
- Hedegaard-Soerensen, L., Grumlose, S.P., 2020. Exclusion: the downside of neoliberal education policy. *Int. J. Incl. Educ.* 24 (6), 631–644. <https://doi.org/10.1080/13603116.2018.1478002>.
- Hill, K., Webber, R., 2021. Staying Afloat in a Crisis: Families on Low Incomes in the Pandemic, Seeking an Anchor in an Unstable World: Experiences of Low-Income Families over Time. Joseph Rowntree Foundation, York.
- Holloway, J., Brass, J., 2018. Making accountable teachers: the terrors and pleasures of performativity. *J. Educ. Pol.* 33 (3), 361–382.
- Hornby, G., Evans, B., 2014. Including students with significant social, emotional and behavioral difficulties in mainstream school settings. In: Garner, P., Kauffman, J.M., Elliot, J. (Eds.), *The SAGE Handbook of Emotional and Behavioral Difficulties*, second ed. SAGE, London, pp. 235–347.
- Inchley, J., Mokogwu, D., Mabelis, J., Currie, D., 2020. Health Behaviour in School-Aged Children (HBSC) 2018 Survey in Scotland: National Report. MRC/CSO Social and Public Health Sciences Unit, Trans., Glasgow.
- Incedem, 2020. Poverty and the Impact of Coronavirus on Young People and Families in Scotland (Glasgow).
- Jung, L.A., Smith, D., 2018. Tear down your behavior chart!(impact on vulnerable learners). *Educ. Leader* 76 (1), 12–18.
- Khon, A., 1999. Punished with Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes. Houghton Mifflin Company, New York.
- Khon, A., 2001. Beyond Discipline: From Compliance to Community. Merrill Education. Association of Supervision and Curriculum Development, New Jersey, USA.
- Khon, A., 2005. Unconditional teaching. *Educ. Leader* 63 (1), 20–24.
- Kinsella, W., 2020. Organising inclusive schools. *Int. J. Incl. Educ.* 24 (12), 1340–1356. <https://doi.org/10.1080/13603116.2018.1516820>.
- Kirk, D., 2020. Precarity, Critical Pedagogy and Physical Education. Routledge, Abingdon.
- Larsen, T.C., Holloway, J., Hamre, B., 2019. How is an inclusive agenda possible in an excluding education system? Revisiting the Danish Dilemma. *Int. J. Incl. Educ.* 23 (10), 1049–1064. <https://doi.org/10.1080/13603116.2019.1626497>.
- Lane, Clark, Peacock, 2021. Britain's 'accidental Savers' – Who Are They and What Are They Doing with Their Windfall? – New Research from LCP, 28.02.2021. Retrieved 17.03.21, from: <https://http://www.lcp.uk.com/media-centre/2021/02/britain-s-accidental-savers-who-are-they-and-what-are-they-doing-with-their-windfall/>.
- Liasidou, A., Antoniou, A., 2015. Head teachers' leadership for social justice and inclusion. *Sch. Leader. Manag.* 35 (4), 347–364. <https://doi.org/10.1080/13632434.2015.1010499>.
- Lozman, T., 2009. Straight Talk about Inclusive Education Outcomes in Alberta. CASS Connections.
- MacBeath, J., Gray, J.M., Cullen, J., Frost, D., Steward, S., Swaffield, S., 2007. Schools on the Edge: Responding to Challenging Circumstances. Chapman Publications, London.
- Mays, D., Franke, S., Metzner, F., Boyle, C., Jindal-Snape, D., Schneider, L., Zielemann, H., Pawlis, S., Wichmann, M., 2018. School belonging and successful transition practice. In: Allen, K.-A., Boyle, C. (Eds.), *Pathways to School Belonging*. Brill Academic Publishers, pp. 167–187.
- McCluskey, G., Cole, T., Daniels, H., Thompson, I., Tawell, A., 2019. Exclusion from school in Scotland and across the UK: contrasts and questions. *Br. Educ. Res. J.* 45, 1–20. <https://doi.org/10.1002/berj.3555>.
- McLean, A., 2003. The Motivated School. Chapman Publications, London.
- Messiou, K., 2012a. Collaborating with children in exploring marginalisation: an approach to inclusive education. *Int. J. Incl. Educ.* 16 (12), 1311–1322.
- Messiou, K., 2012b. Confronting Marginalisation in Education: A Framework for Promoting Inclusion. Routledge, London.
- Messiou, K., 2019. Understanding marginalisation through dialogue: a strategy for promoting the inclusion of all students in schools. *Educ. Rev.* 71 (3), 306–317.
- Minow, M., 1990. Making All the Difference: Inclusion, Exclusion and American Law. Cornell University Press, Ithaca.
- Moore, A., Clarke, M., 2016. 'Cruel optimism': teacher attachment to professionalism in an era of performativity. *J. Educ. Pol.* 31 (5), 666–677.
- Mowat, J.G., 2008. Teaching for Understanding within the Affective Field (Ph D). University of Glasgow, Glasgow.
- Mowat, J.G., 2015a. 'Inclusion – that word!' examining some of the tensions in supporting pupils experiencing social, emotional and behavioural difficulties/needs. *Emot. Behav. Difficulties* 20 (2), 153–172. <https://doi.org/10.1080/13632752.2014.927965>.
- Mowat, J.G., 2015b. Towards a new conceptualisation of marginalisation. *Eur. Educ. Res. J.* 14 (5), 454–476. <https://doi.org/10.1177/1474904115589864>.
- Mowat, J.G., 2018. 'Closing the Gap': systems Leadership is no leadership at all without a moral compass – a Scottish perspective. *Sch. Leader. Manag.* 39 (1), 48–75. <https://doi.org/10.1080/13632434.2018.1447457>.
- Mowat, J.G., 2019a. Exploring the impact of social inequality and poverty on the mental health and wellbeing and attainment of children and young people in Scotland. *Improv. Sch.* 22 (3), 204–223. <https://doi.org/10.1177/1365480219835323>.
- Mowat, J.G., 2019b. Supporting the socio-emotional aspects of the primary-secondary transition for pupils with social, emotional and behavioural needs: affordances and constraints. *Improv. Sch.* 22 (1), 4–28. <https://doi.org/10.1177/1542305018817850>.
- Mowat, J.G., 2019c. Supporting the transition from Primary to Secondary school for pupils with social, emotional and behavioural needs: a focus on the socio-emotional aspects of transfer for an adolescent boy. *Emot. Behav. Difficulties* 24 (1), 50–69. <https://doi.org/10.1080/13632752.2018.1564498>.
- Mowat, J.G., 2020. Interrogating the relationship between poverty, attainment and mental health and wellbeing: the importance of social networks and support – a Scottish case study. *Camb. J. Educ.* 50 (3), 345–370. <https://doi.org/10.1080/0305764X.2019.1702624>.
- Norwich, B., 2013. Addressing Tensions and Dilemmas in Inclusive Education: Living with Uncertainty. Routledge, London.
- OECD, 2017a. Are Students Happy? PISA 2015 Results: Students' Well-Being. Vol 3. PISA in Focus (vol. 71). OECD, Paris.
- OECD, 2017b. PISA 2015 Results (Volume III): Students' Well-Being. Country Note. OECD, UK. Paris.
- OECD, 2018. Equity in Education: Breaking Down Barriers to Social Mobility. OECD Publishing, Paris.
- Oliver, M., Barnes, C., 2012. The New Politics of Disablement, second ed. Palgrave Macmillan, Basingstoke.
- Patalay, P., Fitzsimons, E., 2018. Mental Health and Wellbeing at Age 14 - Initial Findings from the Millennium Cohort Study Age 14 Survey. Centre for Longitudinal Studies, London.
- Patel, V., Shekar, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., 2018. The Lancet Commission on global mental health and sustainable development. *The Lancet Commissions* 392 (10157), 1553–1598.
- Peachey, K., 2021. Six Million Accidental Savers 'Created by Covid Crisis', 28.02.2021. Retrieved 17.03.2021, from: <https://http://www.bbc.co.uk/news/business-56210579>.
- Prince, E.J., Hadwin, J., 2013. The role of a sense of school belonging in understanding the effectiveness of inclusion of children with special educational needs. *Int. J. Incl. Educ.* 17 (3), 238–262.
- Putnam, R.D., 2015. Our Kids, the American Dream in Crisis, Reprint Edition ed. Simon and Schuster, New York.
- Razer, M., Friedman, V.J., Warshofsky, B., 2013. Schools as agents of social exclusion and inclusion. *Int. J. Incl. Educ.* 17 (11), 1152–1170.
- Riley, K.A., 2017. Place, Belonging and School Leadership : Researching to Make the Difference. Bloomsbury Academic, London.
- Riley, K.A., Rustique-Forrester, E., 2002. Working with Disaffected Students. Chapman Publications, London.

- Rix, J., Sheehy, K., Fletcher-Campbell, F., Crisp, M., Harper, A., 2013. Exploring provision for children identified with special educational needs: an international review of policy and practice. *Eur. J. Spec. Needs Educ.* 28 (4), 375–391.
- Schleicher, A., 2014. Equity, Excellence and Inclusiveness in Education: Policy Lessons from Around the World: Background Report for the 2014 International Summit of the Teaching Profession. OECD, Paris, France.
- Sergiovanni, T.J., 1994. *Building Community in Schools*. Jossey-Bass, San Francisco.
- Shakespeare, T., 2014. *Disability Rights and Wrongs Revisited*, second ed. Routledge, London.
- Slee, R., 2013. The labelling and categorisation of children with emotional and behavioural difficulties: a cautionary consideration. In: Cole, T., Daniels, H., Visser, J. (Eds.), *The Routledge International Companion to Emotional and Behavioural Difficulties*. Routledge, London, pp. 15–21 [Routledge Handbook].
- Slee, R., 2018a. Defining the Scope of Inclusive Education. Think Piece Prepared for the 2020 Global Education Monitoring Report – Inclusion and Education. UNESCO, Paris.
- Slee, R., 2018b. Inclusive Education Isn't Dead, it Just Smells Funny. Routledge, London.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922. 0.1080/13603116.2019.1602366.
- Solomon, Y., Lewin, C., 2016. Measuring 'progress': performativity as both driver and constraint in school innovation. *J. Educ. Pol.* 31 (2), 226–238.
- The Children's Society, 2020. *The Good Childhood Report 2020* (London).
- Theoharis, G., Scanlan, M., 2015. *Leadership for Increasingly Diverse Schools*. Routledge, London.
- Thomas, C., 2007. *Sociologies of Disablement and Illness: Contested Ideas in Disability Studies and Medical Sociology*. Palgrave Macmillan, Basingstoke.
- Tomlinson, S., 2016. Special education and minority ethnic young people in England: continuing issues. *Discourse* 37 (4), 513–528. <https://doi.org/10.1080/01596306.2015.1073013>.
- Tsovoka, D., Tarr, J., 2012. *Diverse Perspectives on School Communities*. Routledge, Oxon, England.
- UNESCO, 1990. *World Declaration on Education for All and Framework for Action to Meet Basic Learning Needs*. Florence.
- UNESCO, 2017. *A Guide for Ensuring Inclusion and Equity in Education* (Florence).
- UNESCO, UNICEF, The World Bank, 2020. *What Have We Learnt? Findings from a Survey of Ministries of Education on National Responses to Covid-19* (Paris, New York, Washington D.C.).
- UNICEF Office of Research, 1989. *United Nations Convention on the Rights of the Child*. UNICEF, London.
- UNICEF Office of Research, 2010. *The Children Left behind: A League Table of Inequality in Child Well-Being in the World's Rich Countries*. Innocenti Report Card 9 Innocenti Report Card 9. UNICEF Innocenti Research Centre, Florence.
- UNICEF Office of Research, 2016. *Fairness for Children: A League Table of Inequality in Child Well-Being in Rich Countries*. UNICEF Innocenti Report Card 13. UNICEF Office of Research – Innocenti, Florence.
- UNICEF Office of Research, 2017. *Building the Future Children and the Sustainable Development Goals in Rich Countries*. Innocenti Report Card 14, vol. 14. UNICEF Office of Research – Innocenti, Florence.
- UNICEF Office of Research, 2019a. *All Children in School*. Retrieved 27/03/2019, from file: <http://Users/jis04102/Desktop/Global%20Trends%20in%20Education/UNICEF/All%20children%20in%20school%20%7C%20UNICEF.webarchive>.
- UNICEF Office of Research, 2019b. *Children and the Sustainable Development Goals*. Retrieved 28/03/2019, from: <https://data.unicef.org/children-sustainable-development-goals/>.
- United Nations Children's Fund, 2020a. *Children with Disabilities: Ensuring Their Inclusion in Covid-19 Response Strategies and Evidence Generation* (New York).
- United Nations Children's Fund, 2020b. *Protecting Children from Violence in the Time of COVID-19: Disruptions in Prevention and Response Services* (New York).
- United Nations Children's Fund, 2021. *Covid-19 and School Closures: One Year of Education Disruption*. UNICEF, New York.
- United Nations Children's Fund and International Telecommunication Union, 2020. *How Many Children and Young People Have Internet Access at Home? Estimating Digital Connectivity during the COVID-19 Pandemic* (New York).
- United Nations Children's Fund and UN Women and Plan International, 2020. *A New Era for Girls: Taking Stock of 25 Years of Progress* (New York).
- Vetoniemä, J., Kärnä, E., 2019. Being included – experiences of social participation of pupils with special education needs in mainstream schools. *Int. J. Incl. Educ.* 25 (10), 1190–1204. <https://doi.org/10.1080/13603116.2019.1603329>.
- Wilkinson, R., Pickett, K., 2010. *The Spirit Level: Why Equality is Better for Everyone*. Penguin Books, London.
- World Health Organisation, 2016. *Growing up Unequal: Gender and Socioeconomic Differences in Young People's Health and Well-Being: Health Behaviour in School-Aged Children (HBSC) Study: International Report from the 2013/2014 Survey* (Vol. Health Policy for Children and Adolescents, No. 7). World Health Organisation, Denmark.

Further Reading

- Florian, L., Black-Hawkins, K., Rouse, M., 2017. *Achievement and Inclusion in Schools*. Routledge, London.
- Messiou, K., 2012. *Confronting Marginalisation in Education: A Framework for Promoting Inclusion*. Routledge, London.
- Riley, K.A., 2017. *Place, Belonging and School Leadership: Researching to Make the Difference*. Bloomsbury Academic, London.
- Mowat, J.G., 2023. *Building Community to Create Equitable, Inclusive and Compassionate Schools through Relational Approaches*. Routledge, London.

Nordic representations and paradoxes of inclusive education

Thomas T. Engsig, University College of Northern Denmark, Aalborg, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction	98
A Nordic model of education and the development of inclusive education	98
Representations of inclusive education in the Nordic countries	100
From welfare states to competition states	101
The intersection of educational agendas	101
A Danish example: inclusion and a call for excellence	102
Concluding remarks: continuum of inclusions	102
References	103

Introduction

Inclusive education is a pedagogic and didactic practice that aims to increase students' possibilities of participation, learning and subjective experiences of belonging and contributing to school communities. Inclusive education is not a neutral educational practice with an unambiguous what-works-formula, which thus can be analyzed and understood in a *one-size-fits-all* manner. On the contrary, inclusive education is framed by purposes, intentions, culture, actions of students and teachers, specific locations, artifacts, and policy opportunities and constraints. Considering this, it seems reasonable to understand and discuss inclusive education through socio-cultural and historical lenses which entail a sensibility to the notion of *terroir* which in this context denotes matter in a certain place. From the outset, inclusive education, like any other educational phenomena, is shaped by societal, cultural, geographic, and historical developments. Accordingly, it cannot be analyzed nor understood if detached from the contextual practices in which it occurs. In light of this, the notion of inclusive education manifests itself in many variations and is thus shaped by specific socio-cultural and historical developments and pathways in specific countries globally.

This, however, is far from a new understanding and contributes to the difficulties in defining inclusive education. In the literature, this is often articulated as inclusive education being nearly impossible to define and operationalize (Larsen et al., 2019; Ainscow et al., 2006; Ainscow and Sandill, 2010; Mitchell, 2005). The trouble we encounter, when attempting to grasp what inclusive education is and what it really looks like in pedagogical and didactic practice, is rooted in the fact that it appears differently, with variations in potential and difficulty, around the world. Furthermore, the difficulties with grasping and articulating a clear definition of inclusive education is also rooted in a, presumably banal but nonetheless essential, question: what is inclusive education for? This teleological question serves the purpose of attempting to understand whether inclusive education is a pedagogical phenomenon on its own premises or merely an instrument of societal change. In other words, does inclusive education only have socio-political purposes or also ethical, didactical, and pedagogical purposes? A central rationale for posing this question is the fact that we in school systems around the world are witnessing an alignment between significant educational agendas. This alignment, or perhaps even collision, between education agendas can lead to substantial educational paradoxes, contra productive effects, and dichotomies in the understanding of inclusive education. In the case of the inclusive education agenda, it seems plausible to argue that there is an alignment with more neoliberal agendas of standardization, accountability, marketization, and testing in the Nordic countries (Denmark, Sweden, Norway, Finland, Iceland, and the autonomous territories of Greenland and the Faroe Islands) as well as other countries around the world (Engsig and Johnstone, 2015). Korsgaard and Mortensen (2017) argue that this alignment, between the agenda of inclusive education and a neoliberal agenda, in fact may undermine inclusive education. Following this, Korsgaard and Mortensen (2017) emphasize that we must begin to ask educational questions of inclusion, instead of asking inclusive questions of education. From this starting point, and from the acknowledgment of the necessity of a socio-cultural and historical sensitivity in understanding inclusive education as an educational phenomenon, this chapter will discuss inclusive education from a Nordic perspective and argue that different sociopolitical developments have transformed the notion of inclusive education in Nordic countries.

A Nordic model of education and the development of inclusive education

Education is often viewed as a cornerstone in any society and perhaps particularly in the modern welfare state, and it seems reasonable to assume that a central common characteristic of the Nordic countries is the existence of a comprehensive welfare state. Ydesen and Andersen (2020) argue that the Nordic welfare states have a long history of pursuing increased comprehensive schooling with

a strong focus on quality education for all. Inclusion has been an inescapable and essential priority in education policy in the Nordic countries for decades. Following this, the premise of this chapter is based on the assumption that there are common underpinnings in the Nordic education systems for it to be meaningful to discuss Nordic representations of inclusive education. Thus, it seems necessary to commence with posing the question: is it meaningful to talk about a Nordic model of education?

Larsen et al. (2019) argue that the policy of inclusion, with its focus on social justice, equity, equality, access, and democracy, is closely aligned with the notion of the welfare state. The Nordic countries are often referred to when we attempt to describe the ideal version of the modern welfare state as a way of organizing universal rights for the individual citizen (Hamre, 2018). Arnesen and Lundahl (2006, p. 287) underline that the Nordic welfare states are characterized by:

... a relative strength and autonomy of political solutions and universalistic policies which not only target the neediest, but include the whole population. Collective political action is regarded as crucial to minimise differentiation and to prevent social exclusion.

The social-democratic welfare regime, which since the post-war period has been under continuous development in the Nordic countries, is characterized by substantial economic redistribution, high taxation, universal policies assuring comprehensive health care for all, pensions, extensive childcare, and free education, as well as higher education. Hence, it is possible to identify a range of common traits in the Nordic countries which denotes shared historical and political pathways in the continuous development of the modern welfare state. An example of such shared pathways, in the development of education in Nordic countries, is the idea of comprehensive education, which is firmly rooted in the ideology of the welfare state, and to a certain extent stands on social democratic foundations (Arnesen and Lundahl, 2006). The social motive of education is shared in Nordic education systems and reforms, and Arnesen and Lundahl (2006) address that particular Norwegian and Swedish education reforms stress the social motives of education. A substantial reform of the Danish school system in 2014 also entailed an extensive focus on the significance of students' social background in relation to achievement. Divisions within Nordic education systems are few and a large majority of students in Nordic countries attend public schools. Arnesen and Lundahl (2006) show data from OECD illustrating, that an overwhelming percentage of students in the Nordic countries go to public schools, and that the percentage of students in private school are well below OECD average. Another shared pathway, in the development of a Nordic education model, is the spending on public schooling. Data from the latest Education at a glance show, that Nordic countries are well above OECD average in total expenditure on educational institutions as a percentage of GDP (OECD, 2021). In other words, the comprehensive school systems in Nordic countries are a shared political priority and hence it is prioritized in relation to public spending on education systems. Another common trait, or a shared pathway, in Nordic education systems is an emphasis on pedagogical and didactic differentiation instead of tracking and use of segregated special education. The recent reform of the Danish school system underlined the need for more inclusive schooling and more differentiation and the Finnish curricula stresses that education should be based on equality and tolerance and develop communality. The Norwegian curriculum also entails a strong emphasis on a school which should include different groups of students and reduce social inequality (Arnesen et al., 2007). Although it must be acknowledged that the Nordic countries are independent countries with independent histories, cultures and demographics, Sweden, Norway, Finland, Denmark and Iceland share, as mentioned above, some common traits that are useful to consider when analyzing the characteristics of the modern Nordic welfare state. A strong interdependency between economic growth, and policies regarding capitalist economic development, and welfare on both an individual and collective level is shared by the Nordic countries (Arnesen and Lundahl, 2006). Hence, as Arnesen and Lundahl (2006) observe, the development of the modern welfare state in the Nordic countries entails a strong emphasis on social equality on the one hand and economic development on the other hand, which have resulted in interventionist welfare states.

Following this brief history of the development of the Nordic modern welfare states stands the question of the existence of a Nordic model of education, and perhaps even a Nordic model of inclusive education. In the development of the modern welfare state, in the Nordic countries, schooling was, and still is, to a large extent viewed as a quintessential welfare institution, which entails the capabilities of reducing social inequalities and enhancing social mobility and thus reducing barriers for access, participation, and learning. Lundahl (2016) indicates that the notion of a Nordic model of education historically has been utilized to point to the similarities in educational reforms, during the second half of the 20th century, that have shaped the school systems we see in the Nordic countries today. These educational reforms are aimed at: "social justice, equality and cohesion not least by providing schooling of high and equal quality, regardless of children's and young people's resources, origin and location" (Lundahl, 2016, p. 3). In other words, several education reforms in Nordic countries aimed at reducing inequalities and increasing social mobility and equity, have shaped common Nordic pathways that make it plausible to argue for the outlines of a Nordic model of education. Thus, it seems reasonable to assert that we do indeed see substantial similarities in the developments of schooling and educational systems in the Nordic context (Lundahl, 2016; Antikainen, 2006; Telhaug et al., 2004). However, as Antikainen (2006) stresses, it might be more sensible and adequate to talk about common patterns or, as this author suggests pathways, in the development of education in the Nordic countries and not as a fixed model of education. The question, that follows concerns the possibility of identifying similarities in the development of inclusive education in the Nordic countries.

Representations of inclusive education in the Nordic countries

Having established that it is justified, to a certain degree, to talk about common pathways in the development of education in the Nordic countries, I will now discuss similarities in the developmental pathways of inclusive education in the Nordic countries. The following will be based on the premise that the notion of inclusive education is rooted in some common traits and ideas but that there is no universal model of inclusion that fits all countries. Haug (2017) emphasizes that each country must develop its own pathway to inclusive education.

Inclusive education is linked to ideas and principles of democracy, equity, equal rights, access and participation. Haug (2017) argues that the practical state of inclusive education varies globally, nationally, and even between and within schools. Furthermore, Haug (2017) observes that in all countries there is a substantial gap between formulations and realizations of inclusive education. This also seems to be the case in the Nordic countries and can to some extent be attributed to the fact that understandings of inclusion, and thus contours of practice, range from mere placement of students in general neighborhood schools to a more fundamental transformation of the ways in which schooling is practiced based on values, ideas and educational practices (Bacon, 2020). With this in mind, the history of the development of inclusive education in Nordic countries is too complex to be discussed exhaustively in this chapter, but it seems reasonable to argue that inclusion has been a key educational idea, phenomena and ideal since the 1990s and has furthermore been adapted to educational policy-making, social policies and development of welfare states (Arensen and Lundahl, 2006).

Providing access to education for all has been a key driver in the development of inclusive education in the Nordic countries. This must be understood as a significant element in the development of the modern welfare state where equal opportunities for participation in society is absolutely central. Education policies in the Nordic countries have throughout the years developed a more stringent focus on providing opportunities for all regardless of factors such as social class, gender, ethnicity and academic abilities. An example is the substantial reform of the Danish school system in 2014. One of the reform's three national aims was that schools had to reduce the significance of students' social background in relation to their academic results (Engsig, 2020). The Danish example, which is illustrative of the push for equal opportunities in educational policies in Nordic countries, emphasizes that access not only refers to free comprehensive education in the Nordic welfare states, but also to guaranteeing opportunities for all to profit from comprehensive education. One could argue that the Nordic countries' development of a comprehensive school system, providing compulsory education for all, starting in the 1960s and 1970s, provided the building blocks for today's Nordic representations of inclusive education. There is a strong tradition of differentiation in the Nordic countries and streaming is less practiced than in many other countries:

Streaming is still largely absent from compulsory education in the Nordic countries, and special education support has moved from separate schools and classes to the "inclusive school", i.e., it is to a great extent now integrated with other teaching in the classroom

Lundahl (2016, p. 8).

A school system with a minimum of streaming or tracking students is common for the Nordic countries and is a central component to Nordic representations of inclusive education. Lie et al. (2003) argue that the Nordic strategy for building school systems, which display a high degree of inclusiveness, equity, and quality entails comprehensive school systems that do not practice streaming or tracking. It seems fair to accentuate that there is a fairly equal provision of education at different levels in the Nordic countries which is supported by data pointing to the fact that the Nordic countries have a higher degree of socio-economic variation in compulsory schools compared to other OECD countries (Lundahl, 2016). Data from OECD indicate that the Nordic countries is above OECD average regarding social inclusion, which in this context refers to the socio-economic variation within schools (OECD, 2012). Furthermore, Arnesen and Lundahl (2006) show that students from the Nordic countries display strong democratic beliefs and values and report a stronger sense of belonging compared to the OECD average. The results from PISA 2018 show, that in Denmark, Sweden, Norway, Finland and Iceland students report a relative high degree of sense of belonging at school. The Nordic countries are all on the OECD average or slightly above when it comes to students who agree or strongly agree with statements as: I make friends easily at school; I feel I belong at school and other students seem to like me (OECD, 2019). There are indications to support that there are some common pathways to a more inclusive school system in the Nordic countries which are coherent with the history and development of the Nordic modern welfare state and ideals regarding schooling. Lundahl (2016) concludes that the Nordic countries do indeed display certain common inclusive qualities and that it is justifiable to speak of a Nordic education model which emphasizes equality and social justice. The same way that it is justifiable to argue for common Nordic pathways in the development of the modern welfare state and (inclusive) education it also seems plausible to argue that we currently witness a shift from the welfare state to the competition state in the Nordic countries and, furthermore that this has implications for inclusive education. The Nordic countries share similarities in developmental pathways regarding the modern welfare state as well as school system. However, it is imperative to underline that the Nordic countries are individual states with strong individual history, cultures, and demographics, but, as discussed in the above, it is justifiable that the Nordic countries do indeed share central pathways on the way to both the modern welfare state as well as the welfare state's inclusive schooling.

From welfare states to competition states

The Nordic countries share other central similarities which have substantial implications for the modern welfare state, education in general, and inclusive education. Lundahl (2016) shows that in the 1980s and 1990s decentralization reforms were initiated in the Nordic countries providing municipalities and schools more autonomy to (re)allocate resources and to implement national policy aims. Lundahl (2016) proposes that extensive decentralization in Nordic countries was a precursor for a more neoliberal development in school systems. Interestingly, Nordic researchers analyzed and discussed changes in educational policies, in a special issue of *Journal of Educational Policy*, and concluded that the neoliberal changes and policy push was an echo of developments in other OECD countries (Lundahl, 2016).

In order to understand the transformations of inclusive education in Nordic countries, we must first understand the broader social changes in Nordic countries. Kristensen (2012) argues that the Danish welfare system has changed radically from a welfare state to a competition state. This applies more broadly across Nordic countries. Kristensen's (2012) argument is that mobilization, self-support, and inclusion are more adequate characteristics of the Nordic welfare state than support, protection and compensation. In line with this, there is evidence to suggest that neoliberal educational agendas are driving a shift in educational values moving from ideas of the collective to the individual and from the notion of "bildung", or the cultivating role of education, to a more rigorous focus on output or achievement. Arnesen and Lundahl (2006) point to a shift in the educational values which education in the Nordic countries have been rooted in for decades. Education policies in the Nordic countries have a long tradition of emphasizing democratic core values and that one of the primary purposes of education is educating for democratic participation. This is, to a large extent, still at the center of educational policy and curricula in the Nordic countries, but we have witnessed another powerful agenda claiming attention:

However, from the 1980s and onward, the rationale for educational attainment has changed from emphasis on collective values and social community to a focus on individual rights, academic progress and choice. Hence, contemporary political and educational discourses reflect dilemmas and contradictions between utility and culture, economy and welfare, the public and personal, community and individual

Arnesen and Lundahl (2006, p. 295).

This movement in educational agendas and values, in the Nordic countries stems from a neoliberal marketization of education which also entails an individualization in relation to the notion of purpose. Korsgaard (2016) states that we are currently witnessing a re-examination of inclusive education across the globe due to a "collapsed construction of what school is at policy, practical and theoretical levels and in part by an increased political focus on the use of the terms inclusion and inclusive education" (p. 935). This poses the question, and a teleological quest; what is inclusive education for and what type of schooling are we including students in? Korsgaard and Mortensen (2017) argue that we must "... begin to ask educational questions of inclusion, as opposed to inclusive questions of education" (p. 1245).

The intersection of educational agendas

The shift from the modern welfare state to the competition state is becoming obvious in the Nordic school systems through a central phenomenon: the intersection or alignment of educational agendas. As described above, Nordic countries went through a marketization, individualizing and decentralization and educational policies have, to an increasing degree, been rooted in neoliberal policies with a focus on standardization, accountability, high-stakes testing, and a focus on individual achievement and competition (e.g., Bacon, 2020; Engsig, 2020; Engsig and Johnstone, 2015). One could argue that the Nordic countries have undergone an instrumentalization of education which have contested the value-based and community-orientated purposes of schooling. The agenda of inclusive education is supported by UNESCO and The Salamanca Statement (UNESCO, 1994), The Convention on the Rights of Persons with Disabilities (United Nations, 2006), as well as the Sustainable Development Goals (United Nations, 2015). International declarations that are all increasingly pushing for more inclusive school systems around the globe. On the other hand, the neoliberal agenda, with a push for marketization, individualization, standardization, high-stakes testing, accountability, and international and national competition, is linked to the OECD and in particular manifested through PISA, TIMMS, PIRLS and other international assessments. There is substantial evidence that we are currently witnessing an alignment between the agenda of inclusion and a more neoliberal agenda. Bacon (2020) argues that neoliberalism and ideas of inclusion conflict ideologically, and that neoliberalism threatens a communal ideology based on equality, reciprocity and human dignity:

... market reforms allow families to opt to avoid engaging in learning alongside those who are different, which is the rationalized as the "choice" of families and students ... exclusivity rather than inclusivity is the primary goal of market-driven policies (p. 53).

Korsgaard and Mortensen (2017) assert that this clear conflict between the agenda of inclusion and a neoliberal educational agenda undermines the very purposes of schooling and inclusion. Arnesen et al. (2007) raise the concern that a market-driven

culture and school system, where factors such as competition and efficiency are main drivers, may push ideals such as inclusivity, equity and collaboration in the background. In a comparative analysis between the discourses of inclusion in a Danish and American context, [Engsig and Johnstone \(2015\)](#) propose a framework of inclusions emphasizing the notion that policies of inclusion can be rooted in different foundations on a continuum of Salamanca-inspired inclusion with a focus on rights and ethics on the one side and a more neoliberal and accountability-inspired inclusion with a focus on excellence, accountability and output on the other side.

A Danish example: inclusion and a call for excellence

In a Danish context we have a series of events in education policy-making that function as a strong example of the intersection between educational agendas. This is a view into events in one Nordic country as an exemplification of the collision between educational agendas and in doing so I do not claim that all the Nordic countries have experienced same phenomena. Nonetheless, the Danish example does not stand alone in the Nordic countries in illustrating the movements in the underpinnings of inclusive education in the North.

In 2012 the so-called Inclusion Law, which was a modification of the national Education Act, was passed in the Danish parliament. Following this legislative modification, students who needed more than 9 h of support per week fell under the act of special education and hence received the funding and support following this. The act redefined the very notion of special educational needs and special education and thus dramatically decreased the number of students receiving special educational support. This was a push for inclusive education and a call for municipalities and schools to dramatically reform their practice regarding students in special educational needs. In 2014 the Danish school system underwent yet another comprehensive reform. The motivation for the reform was anxiety regarding student achievement and Denmark's performance in international comparisons. The reform marked a clear push for excellence, standardization and accountability. The background text, for reforming the school system, clearly stated that all students, regardless of their background, should fulfill their potential and be as proficient as possible. The reform entailed three national goals:

1. Danish schools must challenge all students so they can be as proficient as possible;
2. Schools must reduce the significance of students' social background in relation to their academic results;
3. Confidence in the school system and students' well-being must be improved through increased respect for professional practice ([Engsig, 2020](#)).

The education policy, on which the school system and pedagogic and didactic practice was rooted, emphasizes both inclusivity and a neoliberal focus on output, individualization, excellence, and accountability. In other words, Danish municipalities, school administrators, and teachers must ensure educational access and inclusive education with a constant focus on the improvement of learning achievement ([Engsig, 2020](#); [Hedegaard-Soerensen and Grumloese, 2020](#)). In the Danish example there is a clear alignment between prominent international educational agendas—a push for a more inclusive school system and a neoliberal push for excellence, standardization, testing and accountability.

In a Danish study [Hedegaard-Soerensen and Grumloese \(2020\)](#) conclude that teachers are, to a large extent, loyal to the discourses regarding excellence, learning achievement and individualism and thus the agenda of inclusion seems to fade into the background. Furthermore, they underline the significant finding that neoliberal education policies present a barrier for teachers when it comes to teaching for diversity. One could argue that these findings support the claim that the intersection between the agenda of inclusion and a more neoliberal agenda pushes the practices of inclusion which may lead to reductionist approaches to inclusion. [Bacon \(2020\)](#) argues that a reductionist approach, as opposed to a radical inclusive approach, reduces inclusion to: "a technical solution to problems of resourcing, managing, groupings and designing instruction" (p. 54).

Concluding remarks: continuum of inclusions

The beginning of this chapter argued that inclusive education, and a discussion of Nordic representations of inclusion, must be rooted in a contextual perspective and from socio-cultural and historical lenses. There are substantial common traits in both the development of the modern welfare state, and the movement toward the competition state, in the Nordic countries as well as in regard to inclusive education. Inclusion is understood and practiced differently around the globe and despite of the common pathways and traits in Nordic countries this is also the case in this context. [Kiuppis \(2013\)](#) is promoting the term 'inclusions' in order to capture this conceptual, theoretical and empirical diversity. In order to fully grasp the notion of inclusions one could understand Nordic representations of inclusion on a continuum. This can be done in several manners, but [Engsig and Johnstone \(2015\)](#) have analyzed the paradoxical policies of inclusion through the theoretical lens of [Dyson \(1999\)](#) and his proposed framework on discourses in inclusive education. In this analysis [Engsig and Johnstone \(2015\)](#) showed that the policies of inclusion, in the case of Denmark, could be positioned on a continuum of inclusion from Salamanca-inspired inclusion, with a focus on educational access, rights and ethics, to a more neoliberal and accountability-inspired inclusion with a focus on excellence, output and accountability. A similar analysis could be done regarding all the Nordic countries and this would most likely show an even more complex picture of the various inclusions, and intersectionality of educational agendas, that are at play in a Nordic context.

References

- Ainscow, M., Sandill, A., 2010. Developing inclusive education systems: the role of organisational cultures and leadership. *Int. J. Incl. Educ.* 14 (4), 401–416.
- Ainscow, M., Booth, T., Dyson, A., 2006. *Improving Schools, Developing Inclusion*. Routledge, New York.
- Antikainen, A., 2006. In search of the Nordic model in education. *Scand. J. Educ. Res.* 50 (3), 229–243.
- Arnesen, A.L., Lundahl, L., 2006. Still social and democratic? Inclusive education policies in the Nordic welfare states. *Scand. J. Educ. Res.* 50 (3), 285–300.
- Arnesen, A.L., Miettola, R., Lahelma, E., 2007. Language of inclusion and diversity: policy discourses and social practices in Finnish and Norwegian schools. *Int. J. Incl. Educ.* 11 (1), 97–110.
- Bacon, J.K., 2020. Pursuing “radical inclusion” within an era of neoliberal educational reform. In: Schuelka, M.J., Johnstone, C.J., Thomas, G., Artiles, A. (Eds.), *The SAGE Handbook of Inclusion and Diversity in Education*. SAGE, London.
- Dyson, A., 1999. Inclusion and inclusions: theories and discourses in inclusive education. In: Daniels, H., Garner, P. (Eds.), *Inclusive Education*. Routledge, London, pp. 36–53.
- Engsig, T.T., Johnstone, C.J., 2015. Is there something rotten in the state of Denmark? The paradoxical policies of inclusive education—lessons from Denmark. *Int. J. Incl. Educ.* 19 (5), 469–486.
- Engsig, T., 2020. Experience of inclusion in Denmark. In: Schuelka, M.J., Johnstone, C.J., Thomas, G., Artiles, A. (Eds.), *The SAGE Handbook of Inclusion and Diversity in Education*. SAGE, London.
- Hamre, B., 2018. Inclusion and the management of diversity in the Danish welfare state. In: Best, M., Corcoran, T., Slee, R. (Eds.), *Who's in Who's Out?—What to Do About Inclusive Education*. Brill – Sense, Boston, pp. 33–46.
- Haug, P., 2017. Understanding inclusive education: ideals and reality. *Scand. J. Disabil. Res.* 19 (3), 206–217.
- Hedegaard-Soerensen, L., Grumloese, S.P., 2020. Exclusion: the downside of neoliberal education policy. *Int. J. Incl. Educ.* 24 (6), 631–644.
- Kiuppis, F., 2013. Why (not) associate the principle of inclusion with disability? Tracing connections from the start of the “Salamanca process”. *Int. J. Incl. Educ.* 18 (7), 746–761.
- Korsgaard, M.T., Mortensen, S.S., 2017. Towards a shift in perspective for inclusive education research—a continental approach. *Int. J. Incl. Educ.* 21 (12), 1245–1260.
- Korsgaard, M.T., 2016. An Arendtian perspective on inclusive education: towards a reimagined vocabulary. *Int. J. Incl. Educ.* 20 (9), 934–945.
- Kristensen, J.E., 2012. The will for inclusion—a contemporary diagnosis of the political and pedagogical agenda. *Dansk pædagogisk tidsskrift* 4, 50–59.
- Larsen, T.B., Holloway, J., Hamre, B., 2019. How is an inclusive agenda possible in an excluding education system? Revisiting the Danish Dilemma. *Int. J. Incl. Educ.* 23 (10), 1049–1064.
- Lie, S., Linnakyla, P., Roe, A., 2003. *Northern Lights on PISA: Unity and Diversity in the Nordic Countries in PISA 2000*. University of Oslo, Department of Teacher Education and School Development, Oslo.
- Lundahl, L., 2016. Equality, inclusion and marketization of Nordic education: introductory notes. *Res. Comp. Int. Educ.* 11 (1), 3–12.
- Mitchell, D., 2005. Sixteen propositions on the contexts of inclusive education. In: Mitchell, D. (Ed.), *Contextualising Inclusive Education. Evaluating Old and New International Perspectives*. Routledge, New York, pp. 1–21.
- OECD, 2012. *PISA 2012 Database. Table IV.2.16 and IV.4.4*.
- OECD, 2019. *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*. OECD Publishing, Paris.
- OECD, 2021. *Education at a Glance 2021: OECD Indicators*. OECD Publishing, Paris.
- Telhaug, A.O., Mediaas, O.A., Aasen, P., 2004. From collectivism to individualism? Education as nation building in a Scandinavian perspective. *Scand. J. Educ. Res.* 48 (2), 141–158.
- UNESCO, 1994. *The Salamanca Statement and Framework for Action on Special Needs Education*. Adopted by the World Conference on Special Needs Education: Access and Quality. Salamanca, Spain, 7–10 June.
- United Nations, 2006. *Convention on the rights of persons with disabilities*. Treaty Series 2515, 3.
- United Nations, 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development*. UN Publishing, New York.
- Ydesen, C., Andersen, C.K., 2020. Implementing inclusive education policies—the challenges of organizational change in a Danish municipality. *Nordic J. Stud. Educ. Policy* 6 (1), 69–78.

Utilizing a global social justice lens to explore indicators of inclusive education

Umesh Sharma and Pearl Subban, Faculty of Education, Monash University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	104
Frameworks for examining inclusive education: the global approach	105
Facilitating the purposive search	105
Component one: inputs	106
Theme one: policy	106
Theme two: staff professional development and teacher education	106
Theme three: resources and finances	107
Theme four: leadership	107
Theme five: curriculum	108
Component two: process	108
Theme one: climate	108
Theme two: school practices	109
Theme three: classroom practices	109
Theme four: collaboration and shared responsibility	109
Theme five: supports to individuals	110
Theme six: role of special schools	110
Component three: outcomes	111
Theme one: participation	111
Theme two: student achievement	111
Theme three: post-school outcomes	111
Discussion	112
Conclusion	112
References	112

Introduction

With inclusive programs gaining ground and developing global momentum, many educational institutions are keen to secure a set of measurement indicators to inform program development for students with disabilities. Schools around the globe continue to experience challenges with not just crafting a socially-just inclusionary learning program, but with implementing and evaluating their programs against an established set of criteria. Little research has been undertaken that has international relevance in terms of what indicators may suit different countries. There is also some criticism of the past work as not all indicators are relevant and sensitive to different international contexts. In this paper we identify several indicators that could be relevant to most international contexts. This paper builds on the work of [Loreman et al. \(2014\)](#) which examined international literature to identify inclusive education indicators. Inclusive education has continued to receive global attention within this period, and has emerged as an evolving, organic field with research taking place in both established and developing education systems. The most prominent publication that illuminates the position of socially-just inclusion within school and tertiary education, is the Global Education Monitoring Report ([United Nations Educational Scientific and Cultural Organization, 2020](#)). Triggered and informed by this dynamic exchange of ideas, this piece adds to the existing set of indicators. While preserving the initial themes as underpinning indicators to measure inclusive education, it draws on more recent data, providing an updated framework to gauge the effectiveness of inclusionary programs. The review will be directed by the overarching question of which themes are still relevant for inclusive education in contemporary times. Additionally, emergent inclusive education programs, remain in need of appropriate indicators to obtain an understanding of how to create effective inclusive programs. Media reports highlight this need in this paper, the definition of inclusive education will continue to be framed by UNESCO guidelines:

Education is not simply about making schools available for those who are already able to access them. It is about being proactive in identifying the barriers and obstacles learners encounter in attempting to access opportunities for quality education, as well as in removing those barriers and obstacles that lead to exclusion.

[UNESCO \(2012, para. 1\).](#)

Frameworks for examining inclusive education: the global approach

For the purpose of this review, a universal framework was adopted at the outset, seeking out indicators for inclusion from across the globe. In this regard, a more sensitive view was implemented toward countries classified as developing or as low-income contexts. Inclusionary programs in the developing world are often in fledgling status when viewed through the lens of western accomplishments. This juxtaposition of high-economic and low-economic status, often neglects fundamental cultural norms. There are fewer studies contextualized in the developing world, and those that do, utilize a Western lens. Most cultural orientations in the developing world embrace the view that those who are limited, or who require additional assistance, are to be supported. A commonly held South African belief, coined by Archbishop Desmond Tutu, notes that: "A person is only a person on account of other persons." This acknowledgment of human codependency is an essential trait of more collectivist cultures, and often embedded within the developing world, or countries of the Global South. Sharma (2020a) in exploring the opportunities, strengths and challenges associated with inclusive education in the Pacific region, reflects on the local tradition of "Talanoa," which literally means "to talk." This collectivist and collaborative mindset led to the care, nurturing and facilitation of opportunities for all members of a group, particularly those who cannot speak for themselves (Sharma, 2020b). The First People of Australia, do not acknowledge "disability," and seem to perceive the term as contradictory to their traditional beliefs (Velarde, 2015). In an investigation of the Anangu people in Western Australia, one study notes that Indigenous people do not distinguish disability or impairment, they celebrate the unique construction of each human being, accepting human diversity (Ariotti, 1999). Another study notes that the Maoris of New Zealand (Aotearoa), conceive differences as natural, with disability forming a unique aspect of an individual's identity (Fitzgerald, 1997).

In this regard, indicators were explored within the context of the "policy borrowing" approach, which draws on established practice from countries around the globe. A key element with policy borrowing is a consideration of transferability, especially since the experiences of one context may not necessarily translate effectively in another (Phillips and Ochs, 2003; Steiner-Khamisi and Waldow, 2012). Often policy borrowing does not translate well, given the particular dynamics of individual nations, so we also drew on the work that embraced policy learning, which draws from a nation's own history and develops reform based on contextual and cultural practices (Raffe and Spours, 2007; Sharma, 2020a). Merging the two methods, allowed the authors to draw on multi-national policies, while being mindful of the unique and individual contingencies.

Another framework adopted is through the lived experiences of the authors. Both authors, though now naturalized in Australia, hail from developing countries, and therefore wear multiple lenses in examining the indicators associated with inclusive education. This allows them to be outsiders looking in, and insiders looking out, a technique adopted by immigrants and expatriates through an altered lens (Subban and Clemans, 2017). As a consequence, indicators of inclusion are viewed from both a culturally-oriented perspective and from the view of academics within the inclusive education context. These perspectives allowed for a more informed, critical eye over research studies, literature and gray material available during the stipulated period of time. In identifying the indicators, heed was paid to the United Nations Convention on the Rights of People with Disabilities (2006), the European Social Charter (1961), and the Salamanca Statement (1994), all of which remain fundamental to any dialog on inclusive practice. However, the indicators gathered as markers of inclusive education for this manuscript shifted beyond the legal rights of students with disabilities, translating this right into practical components which could potentially be used by school leaders and practitioners around the globe. In drawing together the indicators, the authors were mindful that the partitions between special schools and regular schools are being progressively dismantled, with many special schools becoming consultation and resource hubs (Komer, 2018). The authors were able to amalgamate ideas to build, improve and enhance inclusionary practices in both developing and developed countries, acknowledging that any intervention or strategy should be contextualized and responsive. It is perhaps unwise to perceive a single practice or implementation as excellent, as its applicability to a national context is relative (Raffe and Spours, 2007).

In this paper, we distinguish inclusion as an ongoing process involving the identification of barriers to learning and participation of all learners, including those with a range of diverse profiles. Inclusion is also about identifying context-sensitive and culturally-responsive solutions to the barriers that many learners face. Additionally, implementing inclusive education is not possible in schools that fail to support educators with their inclusive education efforts. We acknowledge that inclusionary programs present educators with a challenge, and it requires extensive efforts and careful planning of many stakeholders. In this paper, we identify a number of indicators that schools and systems can use to prioritize their actions.

Facilitating the purposive search

As researchers and teacher educators in the field of inclusive education, both authors were keen to include eligible, contemporary studies, which generated indicators for inclusive education. In selecting these studies, we prioritized studies in the developing world, ensuring that the findings of these studies were relevant, and produced rich data that would inform indicators for inclusive education with global appeal. Consequently, data was extracted from individual studies, relating to location, objectives, and results which were pertinent to this review. In the article search, the date range was set from 2014 to 2020, in order to avoid an overlap with research utilized by Loreman et al. (2014). The search was restricted to studies involving inclusive education that were written in English. Given the broad yield of material in this field, the following were considered when identifying substantiation for the indicators in this piece:

- The article should have primarily covered information about inclusive education.
- It should report on measurement, assessment, evaluation or indicators of progress in inclusive education.
- It must have been published after 2014.

Some articles that were outside this date range were included in the final analysis, but these were often seminal contributors to inclusive education, or writers of scholarly material on very specific areas. Given the multitudinous range of material available online, attempts were made to draw on scholarly material in this review. An examination of all associated material available through both hard and electronic mediums proved too time-intensive, and may perhaps be reserved for another investigation. Studies were restricted to the 2014–2020 period to offer a more contemporary scope. Each study was perused for method and research design, with the search procedure recorded in the bibliographic database, EndNote. Studies were noted with regard to size, location, methods adopted, results and possible recommendations. The literature search did present challenges, as methodological approaches in inclusive education varied across scholarly material. Many publications were focused on research design, while others adopted a more pragmatic view of inclusive education. Efforts were made to include both qualitative and quantitative studies, and to minimize bias through a thorough scoping of the literature in recent years. Whittling the wide range of research offerings to only a few studies offered a more detailed inspection of key indicators and allowed the researchers to offer more targeted comments. In this regard, we adopted a critical lens to sift out pertinent literature, paving the way for further refinements as policy and research in the field continues. Each component identified previously (Loreman et al., 2014), was examined through a more critical lens, with each analysis punctuated by a series of reflective questions. These questions, viewed compositely, will provide the basis to develop and refine inclusive education programs more universally. To enable an update of the indicators for inclusive education, the authors drew on the same three components (input, process and output) and 14 themes as those highlighted in the original review (Loreman et al., 2014), however, the dialog now includes apprised ideas as they emerged from more recent publications. Input is everything provided by the system to achieve inclusive education, processes are practices that transform the ways of working with learners and their families, and outcomes are the results or effects of the combined efforts of input and process (Loreman et al., 2014). We present a few questions under each theme that readers may find useful to guide their inclusive education efforts.

Component one: inputs

Theme one: policy

Policy and legislation provide a roadmap to implementing inclusive education initiatives. There are often some gaps in the intent of an inclusive education policy and its implementation. It could be said that most policy documents outline “what” needs to be done. They do not always outline “how” it (i.e., there is inadequate information about the implementation) should be done. Countries that have made significant progress in implementing inclusive education are those that have addressed both the “what” and “how” of inclusive education through their legislative documents. Loreman et al. (2014) identified four key features of effective inclusive education policies (a) the policy being aligned to the international standards and conventions (b) policy being revised and modified in consultation with the members of schooling community including parents, students, teachers and staff (c) policy targeting barriers to achievement of all learners through joint involvement of all members and (d) policy being explicit in identifying and providing resources to support inclusive education. Recently, UNESCO (2020) in its *Global Education Monitoring Report* highlighted that an inclusive education policy should address three key areas of: non-discrimination, zero reject and reasonable accommodation. The principles of non-discrimination and zero rejection are self-explanatory. They require that we should make sure that all learners have access to education regardless of their race, religion and/or abilities. The principle of reasonable accommodation requires educators to think about how best the schools can individualize and personalize curricular activities without stigmatizing learners.

The following questions could be helpful in evaluating a policy:

- Does the policy clearly outline who should be included as well as how an individual student would be included?
- What support and resources would be made available to schools to include the learner?
- Does the policy outline a monitoring mechanism to determine the success of inclusive education efforts in terms of key outcomes for a marginalized learner (e.g., improve presence, participation, achievement, acceptance, belongingness and happiness of the learner)?

Theme two: staff professional development and teacher education

Teaching in inclusive classrooms requires that teachers are adequately prepared to teach learners with a range of diverse profiles in regular classrooms. While teacher education programs should establish systems to ensure that all graduate teachers leave the programs with adequate knowledge and skills; schools should also provide adequate professional learning opportunities for in-service educators to update and acquire new skills to include all learners. The *Global Education Monitoring Report* (2020) analyzed data from 168 countries and found that 61% of teachers received training in some aspects of inclusion. The report did not collect any data about the coverage or quality of the content. The report further outlined that despite teachers reporting that they are trained and qualified to teach, many were still concerned about their lack of skills to teach vulnerable children.

A critical aspect that must be taken into consideration when determining teacher readiness to include all learners is not just the quantity of skills acquired, but it should rather be about how well the training has prepared educators with the right kind of beliefs, knowledge and skills. A well-prepared inclusive educator is someone who understands the “what,” “why” and “how” of inclusive education; and who has the “heart,” “head” and “hands” of an inclusive educator. An effective inclusive teacher education program

should make the task of including learners with a range of diversities easier rather than make it more complex. There is some research which shows that teacher education programs can sometimes increase educator concerns about inclusion (Forlin and Chambers, 2011). Past research shows that too much emphasis on various disabling conditions could have a detrimental effect on educators' readiness to teach in inclusive classrooms (Florian, 2019). Teacher education as well as in-service professional learning programs should make clear connections between the theory and practice of inclusive education. Sharma and Mullick (2020) presented a new CHANGE framework that could assist teacher educators to bridge the gaps between theory and practice.

C—Collaboration between schools and university.

H—Hands-on experience in using inclusive teaching strategies.

A—Assessment of readiness as inclusive teacher.

N—Networking with other stakeholders.

G—Greater contact with learners with diverse learning needs.

E—Effective coaching support from mentor teachers.

Sharma and Mullick (2020).

Teacher education programs should also provide ample opportunities for participants to apply the newly acquired knowledge and skills in real classrooms. Some questions to determine teacher readiness to teach in inclusive classrooms may include:

- Do educators have skills to determine how each learner is different in terms of their learning ability and preferences?
- Do they know about local policies and legislation that support inclusive education, and are they knowledgeable in acquiring necessary resources to teach all learners including those with disabilities?
- Do they know how to group students in mixed ability groups to ensure that learning outcomes are maximized for all learners?
- Do they know what should be taught (not just the content) to different learners based on their abilities and interests?
- Do they have skills to determine if a learner has learned the content?
- Do they have skills to motivate all learners?
- Do they have skills to effectively involve parents and other stakeholders?
- Do they know how to adapt the curriculum and use existing resources?
- Do they know how to personalize the content?

Theme three: resources and finances

A number of authors have noted that inclusive education without adequate funding and resourcing is unlikely to be successful (Banks et al., 2015; Slee, 2007). However, it is important to highlight that increased funding does not always mean that it would result in better implementation of inclusive education. Available funds need to be used effectively. Johnstone et al. (2019) state that increased funding should not necessarily contribute to increased labeling of certain students. It is critical to identify where funds are needed and how they are allocated. Some authors argue that additional funds if not used appropriately could in fact contribute to more exclusion (Slee, 2007; Sharma and Vicek, 2021). For example, in order to support learners with a disability some jurisdictions require that assessments be conducted to determine if a student has a disability. A large portion of funds thus could be spent in gatekeeping and in determining the eligibility of a learner to receive funding. The same funding could be more effectively used to provide support to all learners including those who need extensive support. Schools need not always rely on the availability of funds to start the process of inclusive education. They should rather become creative resource finders. We have learned some key principles that could guide how best to resource and finance inclusive education in economically advanced as well as emerging economies. The principles could also be easily used as key indicators to monitor inclusive education efforts.

- Are funds appropriated to the benefit of all learners?
- Does funding support the active establishment of inclusionary programs?
- Are financial resources used to support the most vulnerable in the school community?

Theme four: leadership

School leaders play an extremely important role in creating inclusive schools (DeMathews et al., 2020). They are gatekeepers of any change initiative. If they believe in an idea, they would find resources to implement the change. On the other hand, if they do not believe in an idea, both their overt and covert actions can slow down the progress of their school to become more inclusive of learners with a diversity. All actions of leaders, who are supportive of inclusive education, are influenced by their fundamental belief that inclusion and equity is a way to achieve excellence. These leaders pay particular attention to ensure that learners who are at the risk of marginalization or underachievement are fully catered for and well supported. They create school communities where both learners and teachers are well supported (Hitt and Tucker, 2016). Inclusive leaders have a common set of attributes that make their schools more inclusive of diversity. They convey a clear vision in support of inclusive education (Carter and Abawi, 2018; Hitt and Tucker, 2016). They are highly supportive of their school staff as they recognize that implementing inclusive education is not easy and without supporting the staff any attempts to implement inclusive education will fail (DeMathews et al., 2020). The critical support that teachers need to teach in inclusive classrooms relate to availability of additional time to plan and access to quality

professional learning opportunities in inclusive education. Past research has also found that inclusive leaders build a community of practitioners and collaborate with other organizations to implement inclusive practices. The following questions may assist to determine if a school has a supportive leader to promote inclusive practices:

- Can the school leader present a strong rationale for the school to be inclusive of diversity?
- Does the leader frequently attend professional learning events that can build the school's capacity to be inclusive of diversity?
- Does the school staff frequently undertake professional learning activities in inclusive education?
- Does the leader actively participate in critical inclusive education policy/planning meetings?
- Does the school use some system of data collection to determine if the school is making progress in implementing inclusive education?
- Does the school leader spend time with learners (as well as their carers) who are at risk of being marginalized?
- Does the staff feel fully supported in their efforts to implement inclusive practices?

Theme five: curriculum

The dialog focusing on the curriculum of inclusive education programs continue to feature on a global scale as more research and enhanced policy documentation is constructed to facilitate this (Lawrie et al., 2017; McMurray and Thompson, 2016). Curriculum is central to a well-functioning inclusive program, as its design and construction principles should be accommodating of student diversity, and incorporate the principles of Universal Design for Learning (UDL). Consequently, rigorous and authentic inclusive curricula should involve multiple forms of representation, engagement and opportunities, to accommodate a range of student needs and learning profiles. Using the micro and macro approach, individual education departments can construct curriculum that is collectively cognizant of the needs of many, while still being mindful of the needs of individual students (Lawrie et al., 2017; McMurray and Thompson, 2016). Personalized instruction continues to be accommodated through individual education plans, especially with regard to those students who require more intensive support. Individual Education Plans (IEPs) should be strengths-based documents that are constructed to highlight the accomplishment of goals (Timothy and Agbenyega, 2018). The establishment of appropriately responsive and inclusive curricula is therefore fundamental to contemporary educational programs which accommodate diversity. Furthermore, the voices of parents and caregivers are fundamental to accurate inclusionary curriculum design and planning (McMurray and Thompson, 2016). More recently, the use of multi-tiered systems of support have become primary to inclusive curricula, with schools around the globe being attentive to student needs first (Sailor et al., 2018; Witzel and Clarke, 2015). Educators need to plan curricular activities keeping in mind that in most teaching activities there will be some students (around 10–15%) who would need additional support to learn the content. They also need to plan for those students who would need even more extensive support (approximately 5%) to learn the content. The multi-tiered support of teaching and planning is a new way of thinking to support learners who need various degrees of support and it could be easily applied in all contexts. The following questions may trigger reflection relating to the construction of an inclusionary curriculum:

- Is the curriculum framework adaptable and malleable to accommodate student needs?
- Are universal design principles elemental to the construction of inclusive curricula?
- Is the curriculum centered around strengths-based principles?
- Does the curriculum draw on evidence-based strategies?
- Are multiple-tiers of support utilized to accommodate different student needs.

Component two: process

Theme one: climate

School ethos, which is sometimes referenced as school climate or school culture, is acknowledged in the *Global Education Monitoring Report* (2020) as crucial to inclusion. The report reveals that some governments identify ethos, or climate, as a prerequisite, a starting point, to inclusive education (United Nations Educational Scientific and Cultural Organization, 2020). Positive school climates have been associated with welcoming attitudes which foster social and emotional development and improved student behavior (Goldberg et al., 2019; Schachner et al., 2019). At the outset, school frameworks are essential to sustain inclusion, with these frameworks including the voices and contributions of parents, staff, students and the wider community. Such frameworks should incorporate a range of student needs that are not necessarily limited to disability, but should utilize any factor that could lead to marginalization or vulnerability. Responsive and inclusive schools create sanctuaries where students can seek help to deal with trauma, stress or illness. This could be facilitated through clinical, and psychological services, with qualified and skilled staff available to aid students. Marginalized students may bear the misfortune of bullying; however, an established inclusive ethos is likely to minimize and combat bullying rates, reducing the adverse impact on students (Hartley et al., 2015; United Nations Educational Scientific and Cultural Organization, 2020). Establishing management practices, and readily visible expectations assist in maintaining inclusive school culture. Additionally, adopting the tenets of Universal Design for Learning is also key to accommodating all students, with these inclusive practices being monitored through constant evaluation and assessment. Technology is likewise

considered central to inclusive schooling in contemporary settings, with access for all students being primary to its implementation. Pragmatic questions which may be posed when constructing an inclusive climate are:

- Is there an explicit means of welcoming students into the school?
- Is there a documented and easily accessible framework to assist with the creation of appropriate inclusive climates?
- Are there sanctuaries or safe places to which students may withdraw to receive personalized support?
- Is there an agreement to manage bullying and aggressive behavior?

Theme two: school practices

Schools are required to implement management and guidance services to encourage inclusivity (United Nations Educational Scientific and Cultural Organization, 2020). School communities should actively embrace the tenets of these services to reduce the marginalization of any student. Villa and Thousand (2016) acknowledge foundational beliefs that schools should embrace when restructuring for inclusive practice. These include, among others, an awareness that all students can succeed, and that diversity within the school can enrich every student (Villa and Thousand, 2016). Inclusivity, therefore, is not just a school practice, but an embedded belief that alters the mindset of staff, parents and students (Villa and Thousand, 2016). School practices should consider the complexity of individual student identities, reflecting on their needs as individuals (Brown, 2019). Research notes that creating spaces to interact within schools is the first step, rather the establishment of inclusive norms which welcome, invite and include in a deliberate and intentional sense becomes elemental to inclusive school practice (Brown, 2019). In this regard, professional development and guidance to the school community at large is primary to the success of these practices (Brown, 2019). For example, inclusive programming of student needs in Singapore is a progressive initiative still in its infancy, however guidelines that support this restructuring include an intensive understanding of how the school experience is academically, socially and emotionally fulfilling for every child (Yeo and Tan, 2018). As schools examine their practice as inclusive institutions, they should be prompted by the following questions:

- Do any of the school practices marginalize students?
- Does the school explicitly acknowledge that every child can succeed?
- Is there an awareness that diversity is to be celebrated?
- Have overt attempts been made to alter the mindsets of the school community?
- Does the school incorporate the view that student identities are complex and multi-layered?

Theme three: classroom practices

Contemporary inclusive education should reduce the potential for the marginalization of vulnerable populations within mainstream or regular classrooms (United Nations Educational Scientific and Cultural Organization, 2020). Modern education systems are now embracing social equity models, with increased numbers of students with disabilities now being educated alongside their differently-abled peers. Consequently, indicators of effective inclusionary practice within classrooms should include aspects that ensure that these vulnerable populations are not stigmatized within inclusive classrooms. The classroom becomes a microcosm where other inclusionary elements are practiced and implemented at close quarters. The inclusive classroom becomes the specific space where policy and values transform into practice. Accommodating student diversity through individualized practice and authentic differentiation is at the core of effective inclusive programming within this context. In the African context, this reflection on “othering” can be detrimental to the social norms that are being embedded (Phasha et al., 2017). Creating belonging and shared ownership are crucial indicators of inclusive classrooms in the African context (Phasha et al., 2017). In Africa, the philosophy of “ubuntu” is central to inclusive classrooms, embracing the shared experiences of human beings (Pather, 2019), and reflects the notions of togetherness, attachment and sharing. In Africa, teachers within special schools often possess skills and knowledge that could be transferable to a mainstream setting (Pather, 2019), however opportunities for collaboration are limited. Equipping teachers with skill and knowledge relating to authentic differentiation, and inclusive practices are elemental to sustain the practice (Donohue and Bornman, 2014). Concrete considerations emanating from studies in Sub-Saharan Africa and many other countries, with regard to classroom practices, reflect the following key ideas:

- The establishment of inclusive philosophies, such as belonging and ownership, within classrooms that are values based.
- Professionally developing mainstream teachers to ensure that they are equipped and feel confident with differentiating instruction for differently able students.
- Creating communities of practice between more specialized settings and mainstream schools to ensure a flow of ideas, strategies, and techniques to accommodate student diversity.

Theme four: collaboration and shared responsibility

Ensuring equal educational opportunities for marginalized populations has appeared in inclusive studies previously, but the *Global Education Monitoring Report* (2020), suggests that all collaboration in theory is not sufficient. They imply that stakeholders mobilize

within any given inclusive context, to align with the needs of vulnerable populations and offer direct systems of support (United Nations Educational Scientific and Cultural Organization, 2020). They reference a vertical and horizontal approach, calling on the school, parents and teachers, government departments and other stakeholders to become active collaborators in this shared task (United Nations Educational Scientific and Cultural Organization, 2020). Shared responsibility of this nature is more likely to tap into available resources and utilize these more efficiently. In economically emerging countries collaboration between schools as well as between schools and community organizations is critical. It ensures that the countries can make the best use of the existing material and human resources. The *Global Education Monitoring Report* (2020) cites the example of Serbia's government which established a streamlined coordination mechanism to support all children within schools. A coordinated approach of this nature is challenging, due to financing structures, and agreement between contributing service provider (United Nations Educational Scientific and Cultural Organization, 2020), however effectual support for students from marginalized populations is only possible if service organizations adopt a common focus, without the impediment of bureaucratic prerequisites. Another example cited in the *Global Education Monitoring Report*, is of Sierra Leone, who established a national commission armed with the sole focus of recognizing the rights of individuals with disabilities and providing access to appropriate services (United Nations Educational Scientific and Cultural Organization, 2020). In establishing these structures within inclusive settings, the following points provide for pragmatic reflection:

- How can schools and educational institutions collaborate with stakeholders to maximize the provision of support to vulnerable students?
- What are the mechanisms that can be introduced into these collaborations to ensure that they achieve their goals promptly?

Theme five: supports to individuals

Individualized and personal support is often necessary to support additional needs within an inclusive setting (Sharma and Salend, 2018). These systems of support offer universal support (85% of students), care for smaller groups (11% of students) and more intensive maintenance of individual student needs (about 4% of students) (Sharma and Salend, 2018). Such supports more effectively address the needs of students, clarify the goals and implementation of plans of support to paraprofessional staff (Sundqvist et al., 2019), and ultimately ensures that individual student needs are being accommodated. Utilizing these systems of support scaffolds the work of teachers, framing their selection of differentiated strategies to suit whole-group, small group, and individual student needs (Giangreco and Suter, 2015). Aligning with Universal Design for Learning, these systems of support, caters to the needs of all students. Platforms of support are generally valued and beneficial to all students regardless of ability. Student diversity within mainstream settings will continue to change, and their learning profiles should be accommodated within the inclusive classroom. Of importance is the need to collaborate with support staff, shifting from the dated exclusionary "pull out" model, and support students within the classroom (Giangreco and Suter, 2015). Altering support programs in this manner has implications for teacher education and school administrators who will need to change their focus so that these tiers of support become the stimulus for altered programming and planning (Giangreco and Suter, 2015). As these Multi-Tiered Systems of Support become established in inclusive schools, the following reflective questions may be constructive:

- Do planning meetings embrace the varying tiers of support offered within the educational institution?
- Is there collaboration between teachers and paraprofessional staff?
- Are there individualized support structures in place for students with more specialized needs that draw on evidence-based practice, differentiated instruction, and the Universal Design for Learning?

Theme six: role of special schools

The role of special schools in the inclusion agenda remains an element of considerable discussion and debate. There have been mixed responses about whether a special school setting provides more intensive care when compared to the mainstream context. Special schools may continue to feature in the educational landscape, however there are indications that partnerships between specialized settings and mainstream schools may forge links (Shaw, 2017), improve awareness of disability and equip teachers with essential skills. In a study documenting shifts in inclusive policies in Portugal, Alves (2019) notes that national strategies are often influenced by international dialog. Some European countries, like Portugal, closed special schools, however more recently, these special schools have been constructed as resource centers (Alves, 2019). Educationalists are cautious however that these resource centers do not become institutions that segregate. While the closure of special schools is generally viewed as positive, mainstream schools should be appropriately equipped to accommodate student needs, through the provision of specialized equipment, resources and staff (United Nations Educational Scientific and Cultural Organization, 2020). Additionally, an individual school's response to diversity, the collaboration of this institution with others around them, and the wider social and cultural policies which impact on inclusive education, are integral to the successful implementation of socially just inclusion (Ainscow, 2020). In reflecting on the role of special schools, the following factors may be central:

- Can specialist schools and special development settings be used as resource hubs rather than locations of segregation?
- Can staff collaborate to ensure that student experiences in mainstream settings are positive?
- Does the mainstream school possess the capacity to include students with additional needs and disabilities?

Component three: outcomes

Theme one: participation

Student participation often enters the literature on inclusive education as a key marker of an effective program. In fact, many education departments around the globe are keen to enhance participation and involvement on the part of students with disabilities. Funding is often invested in a range of resources to encourage participation in this context. A Kenyan-based study revealed that collaboration between parents, the community and schools often have a positive impact on school participation of students with disabilities (Elder and Kuja, 2019). Involving the community increases societal sensitivity to issues related to disability, and is likely to reduce boundaries, allowing students with disabilities to participate in and access educational opportunities more freely. Furthermore, the study reported that the creation of committees which stimulated discussion, and highlighted needs relating to the participation of students with additional needs, was constructive. This model was expected to be replicated by other educational regions in the country. Aligned to this, another study contextualized in Barbados, points out aspects that should be included in inclusive education policy to increase student participation (Blackman et al., 2019). These include explicit guidelines and policy relating to inclusive education; greater collaboration between mainstream school teachers and special education teachers; the use of intentional plans to increase the involvement of students with disabilities in mainstream settings; including a mandatory learning unit into teacher education programs which will enhance knowledge and skill relating to the inclusion of students with disabilities into regular classrooms; and increasing contact with parents of students with additional needs so that these partnerships can inform the inclusive education plans being created by the school. While data on participation remains subjective and there are few yardsticks to measure actual participation, it is evident that the creation of communities can be valuable in enhancing involvement. For schools that are encouraging participation, the following reflective questions may be useful:

- Are inclusionary plans constantly evolving and under review?
- Are parents and other stakeholders included in discussion relating to student participation?
- Are there collaborative discussions within communities relating to the participation of students with disabilities into the wider community?

Theme two: student achievement

Student achievement is at the forefront when considering any effective inclusionary program, with much of the research in the field examining whether students meet their academic goals. In another related study on inclusive education measures in the Australian context, Anderson and Boyle (2015) acknowledged that there are limited guidelines to evaluate inclusive education, leading to varying levels of success with regard to student achievement. Introducing indicators of success acts as a means of evaluating inclusionary programs, especially as they align to student achievement (Anderson and Boyle, 2015). However, research is divided about the effects of inclusive education in relation to student achievement (Dell'Anna et al., 2019; Krammer et al., 2019). Some studies, utilizing a more philosophical and social justice framework, consider interaction, the participation of students with disabilities into mainstream education, and equal access, as markers of achievement (Haug, 2017). In a paper which highlights the growing achievement gaps for students with disabilities in regular class settings, Gilmour et al. (2019) expressed some reservations about whether student needs were being met appropriately. However, their meta-analysis conceded that more research into the area is warranted. Reading and literacy gaps were of marked concern. Other studies link student achievement with teacher practice, highlighting the need for more research-based pedagogies, curriculum differentiation and staff development with regard to accommodating students with disabilities, as triggers to enhance student achievement (Lancaster, 2014). The following reflection questions may prove helpful when considering the achievement of students with disabilities:

- What forms of data are available to measure student achievement?
- Is student achievement limited to academic output, or does it extend to social contribution and interaction?
- Do teachers possess an inclusive mindset when considering student output?
- Is assistive technology used to support students with disabilities in inclusive settings?

Theme three: post-school outcomes

Post-school outcomes are important for all students, but especially so for students with disabilities. From a human rights perspective, many researchers acknowledge that all students should eventually find their place in wider society. Research acknowledges that transition programs and high-quality curricula are important baseline indicators of post-school outcomes for students with disabilities. Schools play a critical role in the transition of students from more structured settings to workplaces and tertiary education (Joshi et al., 2012). Furthermore, bridging programs and other support services should be available to facilitate the shift from school to the workplace or other locations of study (Bouck, 2014). Schools could embed syllabi which offer structured ideas relating to independent living and post-school support (Bouck and Park, 2018; Bouck, 2014). Participation in work-related activities, or work-experience, for example, are usually valuable transition activities to support movement from school to employment. Schools could also assist with matching work options with student profiles to facilitate the post-school shift. Additionally, where appropriate, schools could empower students with disabilities with soft vocational skills, social skills, and information on how to secure

and hold employment (Bouck and Park, 2018). Post-school venues, such as colleges, universities, and other learning centers could also arrange programs that replicate those arranged within schools, to increase participation and reduce challenges for students with disabilities (Bouck and Park, 2018). These programs could offer modifications to the academic units as relevant and appropriate, thereby ensuring course completion and student retention. The following questions may serve as good points of reflection when considering the post-school outcomes of students with disabilities:

- Are transition and bridging programs available and tailored for students with disabilities?
- Are units of work incorporated into school curricula that appropriately prepare students for post-school options and living?
- Are soft skills included in school programs for students with disabilities to facilitate their transition from school to post-school study or work?
- Is there an established relationship between post-school settings (colleges and universities) and schools to facilitate the movement from school?
- Are post-school centers equipped to accommodate individuals with disabilities, with regard to social integration and academic support?

Discussion

Implementing inclusive education is not easy. Past research has shown that some systems around the globe have struggled to implement inclusive practices for many reasons. However, there are also countries and systems around the world that have made remarkable progress in implementing inclusive education. Successful case studies are evident both in economically advanced (e.g., New Brunswick in Canada, Italy, New Zealand) and economically emerging nations (e.g., Peru and Bangladesh). The contexts mentioned have risen to the challenge of providing more equitable educational opportunities for all students, particularly those with disabilities, and implies that economic advantage may not be a determining factor in the successful implementation of inclusive education. The common elements across the systems that have made progress in successfully implementing inclusive education relate to holistic efforts of all key stakeholders (e.g., policy makers, educators, parents and community members) and use of context sensitive practices. As systems across countries begin their future efforts to create inclusive schools, the questions that are often asked are: "where do we start" and "how would we know if we are making any progress." This paper attempted to answer both these questions. It provides a contemporary review of the most critical inclusive education indicators that any system can use to guide their efforts. We believe the indicators reported in this paper are relevant to all countries irrespective of the economic status of the country. Inclusion can only begin when everyone believes that it is a way to achieve excellence for all learners. We have taken care in identifying and reporting indicators that are sensitive to different contexts. We do not make the claim that the indicators be used blindly without paying close attention to the local context. A useful exercise to ensure that the indicators are culturally appropriate and context sensitive may involve close collaboration among policy makers, educators and community members. Our past experience in the Pacific indicates that such exercises can result in creating indicators that are highly valued by local stakeholders and are adopted by policy makers (Sharma et al., 2018). Workshops could be conducted where all key stakeholders come together to discuss, debate and refine the indicators for the local contexts. The newly developed/adopted indicators will be owned by local stakeholders and would have better chances of being adopted at a wider level. We believe this paper has provided a broad framework that countries across the globe can use to create inclusive schools and societies where all learners will flourish.

Conclusion

Ultimately, this paper draws together a set of global indicators for contemporary inclusive school programs, utilizing the lens of social equity. These indicators should be viewed in the light of context-specific factors, being cognizant of individual locations and geographic dynamics. The paper acknowledges that care should be exercised when applying policy-borrowing as such a practice is often negligent and unmindful of local subtleties and the nuances of individual cultures. Inclusive education is likely to thrive when it is developed through a respect and awareness of a country's cultural and social norms. The 14 indicators identified in this paper will act as a platform to illuminate and refine processes that lead to a socially just inclusionary school program, that accommodates all students.

References

- Ainscow, M., 2020. Promoting inclusion and equity in education: lessons from international experiences. *Nord J. Stud. Educ. Pol.* 6 (1), 7–16. <https://doi.org/10.1080/20020317.2020.1729587>.
- Alves, I., 2019. International inspiration and national aspirations: inclusive education in Portugal. *Int. J. Incl. Educ.* 23 (7–8), 862–875. <https://doi.org/10.1080/13603116.2019.1624846>.
- Anderson, J., Boyle, C., 2015. Inclusive education in Australia: rhetoric, reality and the road ahead. *Support Learn.* 30 (1), 4–22.
- Ariotti, L., 1999. Social construction of an angu disability. *Aust. J. Rural Health* 7 (4), 216–222. <https://doi.org/10.1046/j.1440-1584.1999.00228.x>.

- Banks, J., Frawley, D., McCoy, S., 2015. Achieving inclusion? Effective resourcing of students with special educational needs. *Int. J. Incl. Educ.* 19 (9), 926–943. <https://doi.org/10.1080/13603116.2015.1018344>.
- Blackman, S.N.J., Conrad, D.A., Brown, L.I., 2019. Achieving Inclusive Education in the Caribbean and Beyond From Philosophy to Praxis, first ed. 2019.
- Bouck, E., Park, J., 2018. Exploring post-school outcomes across time out of school for students with autism spectrum disorder. *Educ. Train. Autism Dev. Disabil.* 53 (3), 253–263.
- Bouck, E.C., 2014. The postschool outcomes of students with mild intellectual disability: does it get better with time? *J. Intellect. Disabil. Res.* 58 (6), 534–548. <https://doi.org/10.1111/jir.12051>.
- Brown, C.S., 2019. The importance, and the challenges, to ensuring an inclusive school climate. *Educ. Psychol.* 54 (4), 322–330. <https://doi.org/10.1080/00461520.2019.1655646>.
- Carter, S., Abawi, L.-A., 2018. Leadership, inclusion, and quality education for all. *Australas. J. Spec. Incl. Educ.* 42, 49–64. <http://search.proquest.com/docview/2101892634/>.
- Council of Europe, 1961. The European Social Charter. Accessed from: <https://www.coe.int/en/web/european-social-charter/about-the-charter>.
- Dell'Anna, S., Pellegrini, M., Ianes, D., 2019. Experiences and learning outcomes of students without special educational needs in inclusive settings: a systematic review. *Int. J. Incl. Educ.* 1–16. <https://doi.org/10.1080/13603116.2019.1592248>.
- DeMatthews, D., Billingsley, B., McLesky, J., Sharma, U., 2020. Principal leadership for students with disabilities in effective inclusive schools. *J. Educ. Adm.* <https://www.emerald.com/insight/0957-8234.htm>.
- Donohue, D., Borman, J., 2014. The challenges of realising inclusive education in South Africa. *S. Afr. J. Educ.* 34, 01–14.
- Elder, B.C., Kujala, B., 2019. Going to school for the first time: inclusion committee members increasing the number of students with disabilities in primary schools in Kenya. *Int. J. Incl. Educ.* 23 (3), 261–279. <https://doi.org/10.1080/13603116.2018.1432082>.
- Fitzgerald, J., 1997. Reclaiming the whole: self, spirit, and society. *Disabil. Rehabil.* 19 (10), 407–413. <https://doi.org/10.3109/09638289709166565>.
- Florian, L., 2019. On the necessary co-existence of special and inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 691–704. <https://doi.org/10.1080/13603116.2019.1622801>.
- Forlin, C., Chambers, D., 2011. Teacher preparation for inclusive education: increasing knowledge but raising concerns. *Asia Pac. J. Teach. Educ.* 39 (1), 17–32. <https://doi.org/10.1080/1359866X.2010.54085>.
- Giangreco, M.F., Suter, J.C., 2015. Precarious or purposeful? Proactively building inclusive special education service delivery on solid ground. *Inclusion* 3 (3), 112–131. <https://doi.org/10.1352/2326-6988-3.3.112>.
- Gilmour, A.F., Fuchs, D., Webby, J.H., 2019. Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Except. Child.* 85 (3), 329–346. <https://doi.org/10.1177/0014402918795830>.
- Goldberg, J.M., Sklad, M., Elfrink, T.R., Schreurs, K.M.G., Bohlmeijer, E.T., Clarke, A.M., 2019. Effectiveness of interventions adopting a whole school approach to enhancing social and emotional development: a meta-analysis. *Eur. J. Psychol. Educ.* 34 (4), 755–782. <https://doi.org/10.1007/s10212-018-0406-9>.
- Hartley, M.T., Bauman, S., Nixon, C.L., Davis, S., 2015. Comparative study of bullying victimization among students in general and special education. *Except. Child.* 81 (2), 176–193. <https://doi.org/10.1177/0014402914551741>.
- Haug, P., 2017. Understanding inclusive education: ideals and reality. *Scand. J. Disabil. Res.* 19 (3), 206–217. <https://doi.org/10.1080/15017419.2016.1224778>.
- Hitt, D.H., Tucker, P.D., 2016. Systematic review of key leader practices found to influence student achievement: a unified framework. *Rev. Educ. Res.* 86 (2), 531–569. <https://doi.org/10.3102/0034654315614911>.
- Johnstone, C., Lazarus, S., Lazetic, P., Nikolic, G., 2019. Resourcing inclusion: introducing finance perspectives to inclusive education policy rhetoric. *Prospects* 47 (4), 339–359. <https://doi.org/10.1007/s11125-018-9432-2>.
- Joshi, G.S., Bouck, E.C., Maeda, Y., 2012. Exploring employment preparation and postschool outcomes for students with mild intellectual disability. *Career Dev. Except. Individ.* 35 (2), 97–107. <https://doi.org/10.1177/0885728811433822>.
- Korner, I., 2018. Towards Inclusive Education: Examples of Good Practice of Inclusive Education. Inclusion Europe with support of the European Commission, Stockholm.
- Krammer, M., Gasteiger-Klicpera, B., Holzinger, A., Wohlfart, D., 2019. Inclusion and standards achievement: the presence of pupils identified as having special needs as a moderating effect on the national mathematics standards achievements of their classmates. *Int. J. Incl. Educ.* 1–17. <https://doi.org/10.1080/13603116.2019.1573938>.
- Lancaster, J., 2014. Measuring inclusive education school and classroom indicators of inclusive education article information. *Int. Perspect. Incl. Educ.* 3, 165–187.
- Lawrie, G., Marquis, E., Fuller, E., Newman, T., Qiu, M., Nomikoudis, M., van Dam, L., 2017. Moving towards inclusive learning and teaching: a synthesis of recent literature. *Teach. Learn. Inq.* 5 (1), 1–13. <https://doi.org/10.20343/5.1.3>.
- Loerman, T., Forlin, C., Sharma, U., 2014. Measuring indicators of inclusive education: a systematic review of the literature. *Int. Perspect. Incl. Educ.* 3, 165–187. <https://doi.org/10.1108/S1479-363620140000003024>.
- McMurray, S., Thompson, R., 2016. Inclusion, curriculum and the rights of the child. *J. Res. Spec. Educ. Needs* 16 (Suppl. S1), 634–638. <https://doi.org/10.1111/1471-3802.12195>.
- Pather, S., 2019. Confronting inclusive education in Africa since Salamanca. *Int. J. Incl. Educ.* 23 (7–8), 782–795. <https://doi.org/10.1080/13603116.2019.1623329>.
- Phasha, N., Mahlo, D., Sefa Dei, G.J., 2017. Inclusive Education in African Contexts: A Critical Reader. Brill | Sense, Leiden, Boston.
- Phillips, D., Ochs, K., 2003. Processes of Policy Borrowing in Education: some explanatory and analytical devices. *Comp. Educ.* 39, 451–461. <https://doi.org/10.1080/0305006032000162020>.
- Raffe, D., Spours, K., 2007. Policy learning in 14–19 education: from accusation to an agenda for improvement. In: Raffe, D., Spours, K. (Eds.), *Policy-Making and Policy Learning in 14–19 Education* (Bedford Way Papers). Institute of Education, London, pp. 209–330 [Google Scholar].
- Sailor, W., McCart, A.B., Choi, J.H., 2018. Reconceptualizing inclusive education through multi-tiered system of support. *Inclusion* 6 (1), 3–18. <https://doi.org/10.1352/2326-6988-6.1.3>.
- Schachner, M.K., Schwarzenthal, M., van de Vijver, F.J.R., Noack, P., 2019. How all students can belong and achieve: effects of the cultural diversity climate amongst students of immigrant and non-immigrant background in Germany. *J. Educ. Psychol.* 111 (4), 703–716. <https://doi.org/10.1037/edu0000303>.
- Sharma, U., Mullick, J., 2020. Bridging the gaps between theory and practice of inclusive teacher education. In: *Oxford Research Encyclopedia of Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.1226>. Available at.
- Sharma, U., Salend, S., 2018. Personalised learning and support. Inclusive classrooms for students with additional needs. Material prepared for the Stage Government of Victoria, Australia. Dep. Educ. Accessed from https://research.mgt.monash.edu/ws/portals/portal/245337890/245337834_oa.pdf.
- Sharma, U., Jitoko, F., Macanawai, S.S., Forlin, C., 2018. How do we measure implementation of inclusive education in the Pacific Islands? A process for developing and validating disability-inclusive indicators. *Int. J. Disabil. Dev. Educ.* 65 (6), 614–630. <https://doi.org/10.1080/1034912X.2018.1430751>.
- Sharma, U., 2020a. Inclusive education in the Pacific: challenges and opportunities. *Comp. J. Curric. Learn. Assess.* <https://doi.org/10.1007/s11125-020-09498-7>.
- Sharma, U., 2020b. Inclusive education in the Pacific: challenges and opportunities. *Prospects*. <https://doi.org/10.1007/s11125-020-09498-7>.
- Sharma, U., Viecek, S., 2021. Global trends in the funding of inclusive education: a narrative review. In: Goldan, J., Lambrecht, J., Loerman, T. (Eds.), *Resourcing Inclusive Education*. Emerald Publishing, UK, pp. 51–66.
- Shaw, A., 2017. Inclusion: the role of special and mainstream schools. *Br. J. Spec. Educ.* 44 (3), 292–312. <https://doi.org/10.1111/1467-8578.12181>.
- Slee, R., 2007. Inclusive schooling as a means and end of education? In: Florian, L. (Ed.), *The SAGE Handbook of Special Education*. SAGE Publications, London, pp. 160–170.
- Steiner-Khamisi, G., Waldow, F., 2012. Policy Borrowing and Lending in Education. Accessed from: https://www.researchgate.net/publication/274183778_Policy_borrowing_and_lending_in_education.
- Subban, P., Clemans, A., 2017. Outside in: learning from an International Professional Experience Program. In: Fitzgerald, A., Parr, G., Williams, J. (Eds.), *Narratives of Learning Through International Professional Experience*. Springer, Singapore, pp. p111–123. https://doi.org/10.1007/978-981-10-4867-8_8.

- Sundqvist, C., Björk-Åman, C., Ström, K., 2019. The three-tiered support system and the special education teachers' role in Swedish-speaking schools in Finland. *Eur. J. Spec. Needs Educ.* 34 (5), 601–616. <https://doi.org/10.1080/08856257.2019.1572094>.
- Timothy, S., Agbenyega, J., 2018. Inclusive school leaders' perceptions on the implementation of individual education plans. Citation metadata. *Int. J. Whole Sch.* 14 (1), 1.
- UNESCO, 2012. Inclusive education: Addressing exclusion. Author, Paris, France. Retrieved from. <https://unesdoc.unesco.org/ark:/48223/pf0000217073>.
- United Nations Educational Scientific and Cultural Organization, 2020. Inclusion and Education: All Means All. Global Education Monitoring Report. Accessed from. https://drive.google.com/file/d/1Aj3SZ4VOZ_cKWrYc5Y088N9qXiCy5Wp8/view.
- United Nations Educational Scientific and Cultural Organization, 1994. The Salamanca Statement and Framework for Action on Special Needs Education. Accessed from. <https://unesdoc.unesco.org/ark:/48223/pf0000098427>.
- United Nations, 2006. Department of Economic and Social Affairs. Disability. Convention on the Rights of Persons with Disabilities (CRPD). Accessed from. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- Velarde, M.C.R., 2015. Discussion Paper: Indigenous Persons With Disabilities. Access to Training and Employment. Accessed from. https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/publication/wcms_396412.pdf.
- Villa, R., Thousand, J., 2016. Leading an Inclusive School: Access and Success for All Students. Association for Supervision and Curriculum Development, Thousand Oaks.
- Witzel, B., Clarke, B., 2015. Focus on inclusive education: benefits of using a multi-tiered system of supports to improve inclusive practices: Bradley Witzel, editor. *Child Educ.* 91 (3), 215–219. <https://doi.org/10.1080/00094056.2015.1047315>.
- Yeo, L.S., Tan, S.L., 2018. Educational inclusion in Singapore for children with physical disabilities. *Asia Pac. J. Educ.* 38 (2), 175–186. <https://doi.org/10.1080/02188791.2018.1460253>.

Principles of inclusive practice in schools: what is COVID-19 teaching us?

Brent Bradford^a, Tim Loreman^a, and Umesh Sharma^b, ^a Concordia University of Edmonton, Edmonton, AB, Canada; and ^b Monash University, Faculty of Education, Clayton, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	115
Inclusion or exclusion	115
Principles of inclusive practice	116
International crisis: COVID-19	118
Objectives	118
Method	118
Results and discussion: issues exposed by COVID-19	119
Overreliance on technology	119
Changes in the teacher assistant roles	120
Disconnecting relationships among students	121
Principles of inclusive practice in schools	122
Building the plane, while flying	122
What we can learn from COVID-19	123
Conclusion	123
References	124

Introduction

For some considerable time, scholars and education practitioners have more or less settled on an array of principles regarding what constitutes genuine and effective inclusive education. While there may be some minor disagreements around the margins of these principles, they have in general proved relatively stable over time (conservatively, we say, for at least 20 years). Schools and school systems have purported to be making progress on implementing various principles in practice, and in doing so have been able to lay a claim to being an inclusive school or system.

Although inclusive education means that every learner is valued while receiving quality and equitable education (Sharma et al., 2020), the extent to which both the philosophical premise and inclusive education practices have been truly and seamlessly integrated into education has, however, recently been exposed as a result of the international crisis caused by COVID-19. In his novel *Memoirs of a Geisha*, Arthur Golden (1997, p. 405) remarked that "... adversity is like a strong wind. I don't mean just that it holds us back from places we might otherwise go. It also tears away from us all but the things that cannot be torn, so that afterward we see ourselves as we really are, and not merely as we might like to be." Actions taken in order to meet the imperatives of the COVID-19 state have exposed significant flaws in the approach to inclusive education, demonstrating that in many respects the inclusive education emperor, with all its feel-good promotional value, has no clothes. Arguably, in the rush to adopt alternate modes of educational delivery, the principles of inclusive education have too often been set aside, and many already marginalized learners have been simply abandoned. It is this issue, in itself, that has led us to this commentary—in the midst of a global pandemic—that can shed more light on the critical issues surrounding inclusive education. In this chapter, we first discuss a number of key principles that have guided inclusive practices in schools globally. Later, we provide commentary, based on evidence, as to whether or not these principles continued to guide school practices during the COVID-19 state.

Inclusion or exclusion

Inclusive education can be observed from two vantage points, which are not mutually exclusive and which can interact and co-exist. For the sake of distinction, the two vantage points might be described as positive and negative:

- *Positive.* The first vantage point, positive, involves adding to and changing existing structures to support more inclusivity. For example, it may at times be appropriate to add a teaching assistant (also referred to as an *educational assistant*) to the complement of classroom staff in order to help facilitate the type of educator-learner interaction a particular situation demands. Positive views of inclusion such as this might involve checklists of the sorts of resources and practices that must be evident if a school is to be considered inclusive. These checklists can change according to context, but there are elements common to most.
- *Negative.* The second vantage point, negative, involves the removal of barriers to support more inclusivity. Some leading researchers, beginning with Peter Mittler (1979), and later Tony Booth (1995), Slee and Allan (2001), and many others, speak of inclusion as being the removal of things that exclude. Operating under this paradigm, it is the role of the inclusive educator to

identify the sorts of practices, structures, and attitudes that serve to exclude some groups of children, and then work to remove those barriers, if necessary replacing them with more benign elements in order to become more inclusive.

Clearly, the terms positive and negative are not used here in a punitive sense. They are not to be equated with “good” or “bad”. They are applied here because one vantage point involves adding to and modifying an experience, while the other involves the removal of barriers to student engagement and learning. Done well, both can be observed as being helpful to furthering the process of inclusion in the learning environment.

The way in which schools implement inclusive practices may look different in different classrooms, in different schools, and in different countries. However, there are some common elements that need to be understood. Inclusion is a process; it is not a product. Inclusion is a journey; it is not a destination. Inclusion requires that learners are supported based on their needs. It means that sometimes we need to have unequal resources in our schools to address equity and diversity. Inclusion necessitates that we pay as much attention to supporting educators as we pay to supporting learners. Without adequate support to educators, schools may not initiate or sustain efforts of educators to create inclusive classrooms.

Principles of inclusive practice

The right to education is one of the most important rights for persons with disabilities.

World Bank Group (2020, p. 10).

The goal of inclusive education is for education to provide optimal learning experiences for all learners in ways that do not create or perpetuate the marginalization and exclusion of individuals or groups. Three justifications to support inclusive practices in schools globally are educational, social, and economic (UNESCO, 2009, as cited in Sharma, 2020). According to Sharma (2020):

- the educational justification is that all learners, irrespective of any differences, must learn together with inclusion benefitting all learners,
- the social justification points to inclusion being perhaps one of the simplest ways to modify learner attitudes toward each other and create societies where all individuals belong, irrespective of any differences, and
- the economic justification suggests that inclusive education follows a cost-effective measure to educate pupils with diverse learning needs.

It is here that context matters. There is no “one size fits all” model of inclusive education that can apply to all countries or often even systems within countries. For example, in a recent study focusing on inclusive education across the Pacific Island countries, Sharma contended that, along with some similarities, there are stark differences in terms of population, resources, languages spoken, and socio-cultural practices and religious backgrounds, which directly impacts inclusive education efforts. It is, therefore, necessary to examine specific educational contexts, decide which practices marginalize learners, and then work to change those practices by transforming or augmenting systems and specific ways of educating. Principles of inclusive practice that have been recognized as being important considerations have been enumerated and described. Some, such as placement, are obvious. Others, such as the role of home/school partnerships, less so. Fig. 1 outlines common principles associated with inclusive education, starting first with a philosophy of inclusion and desire to do it.

Based on various viewpoints presented and debated in the international literature, the features of inclusive school systems referenced in Fig. 1 include:

- *Pro-inclusion policy and legislation* (Armstrong et al., 2010; Mittler, 1979). Cohesive, inclusive systems and schools can be built only with great difficulty in the absence of guidance from governance bodies, including elected politicians and school authorities. Legislation and policy provide a road-map, ensuring that rights and expectations are defined, and educational mandates for inclusion are specified. In systems governed by weak inclusive education policy, greater opportunity for exclusion exists and occurs.
- *Home/school partnerships* (Christenson, 2004; Goodall and Montgomery, 2014; Hornby and Blackwell, 2018). Families are a child’s first teacher, and learning occurs best when the gaps between the home and the school are reduced. Working in partnership, families and schools can work toward common learning, social, and other goals, and engage in information exchange so as to provide a more holistic view of the learner.
- *Adequate resource provision* (Goldan et al., 2021; Starcic, 2010). It has been known for some time “... that the quality of the organization seems to influence the delivery of the service as much as the overall level of resources” (Sebba and Ainscow, 1996, p. 9). In other words, more money is not the answer to improve inclusion. However, it is the case that adequate resources need to be in place in order to optimize learning for all. These include adequately trained educators and other service delivery personnel, appropriate technology, and any number of other materials based on specific learning contexts and learner needs within those contexts.
- *Positive attitudes and welcoming, learner-friendly environments* (Ahmmed et al., 2012; De Boer et al., 2011; Schwab et al., 2018). Positive attitudes and commitment of all key stakeholders (e.g., educators, school leaders, family members and other members

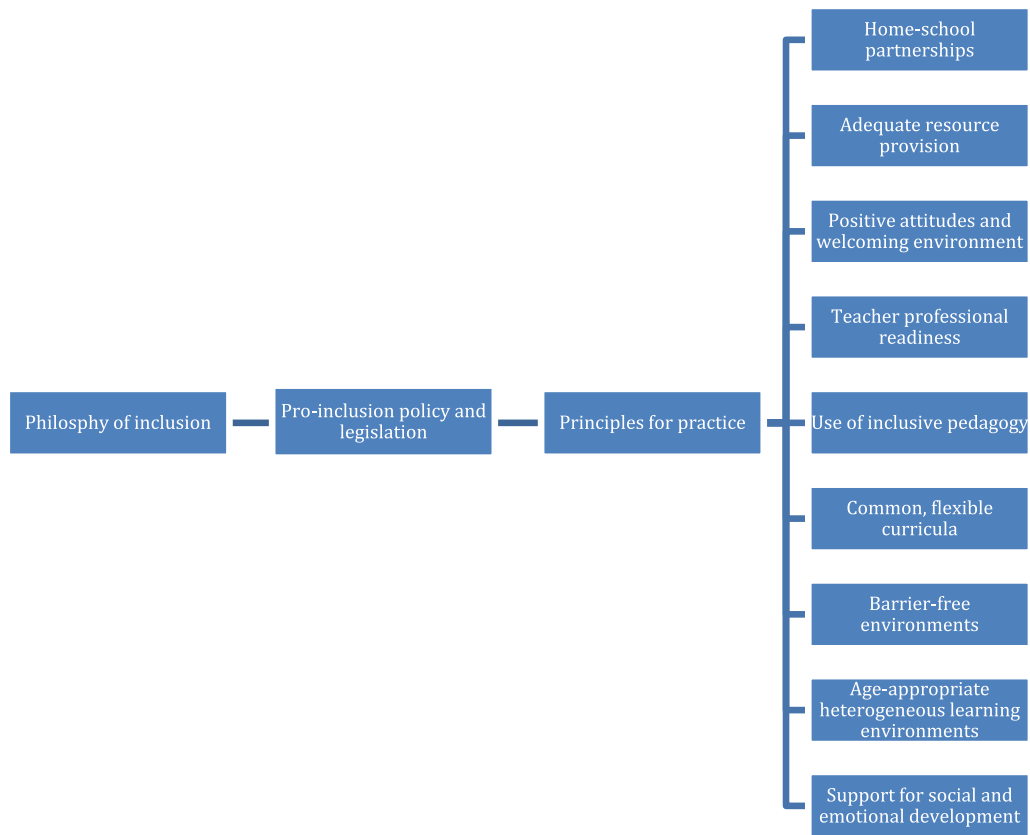


Fig. 1 Principles of inclusive practice.

of the school community) toward the inclusion of all students are critical to the success of inclusive education. Based on these positive attitudes, welcoming school cultures can be developed in which no child is turned away, but is rather embraced as an important and valued school community member.

- *Teacher professional readiness* (Mittler, 2000, 2010; Movkebayeva et al., 2016). Teaching in diverse classrooms requires skillful employment of pedagogical techniques, and the ability to engage in assessment *for* as well as *of* learning. Among other demands, teachers must be skilled at facilitating social and emotional development, working with families, planning, and creating learning experiences that include all students.
- *Use of inclusive pedagogical approaches* (Bradford et al., 2021; Deppeler et al., 2015; Katz, 2013; Loreman, 2017). While a number of pedagogical approaches that aim to facilitate inclusion exist, the one that has garnered the most attention in recent years is known as Universal Design for Learning (UDL). UDL (Katz, 2012) is premised on the provision of multiple means of learner engagement, multiple means of representation (or how content is presented), and multiple means of action or expression which can provide a basis for assessment.
- *Common, flexible curricula* (Mittler, 2000; Norwich, 1990; Operti et al., 2009). The design of school curriculum in many school jurisdictions presents barriers to inclusion. Ideally, a single, common curriculum for all learners exists; one that is broad enough to account for a spectrum of learner interests, abilities, and needs. In practice this has been difficult to accomplish, as evidenced by the presence of alternative curricula for some children, or individualized education programs that have little in common with the general school curriculum.
- *Barrier-free environments* (Ackah-Jnr and Danso, 2019; Erkilic and Durak, 2013). It is self-evident that if a pupil cannot physically access their classroom, they cannot be included. The same can be said of learning environments that, for example, are not equipped with specialized technology to assist learners with disabilities such as deafness or visual impairment. Some environments present cultural barriers where cultural sensitivities are not respected. New educational settings should be constructed with the idea of universal access in mind. Existing buildings should be retro-fitted as required.
- *Age-appropriate heterogeneous learner groupings* (Luciak and Biewer, 2011; Mitchell, 2015). One of the important features of inclusive education is that learners are included in diverse, heterogeneous groups with their same-age peers (in settings where learners are typically grouped by age) regardless of their academic or social and emotional development and capacities. This context best promotes the academic achievement of all, along with appropriate social and emotional growth.

- *Support for the social and emotional development of all learners* (Katz, 2013; Reicher, 2010; Taylor and Sidhu, 2012). Schools have always taught more than straight academic content. Inclusive schools emphasize appropriate social and emotional development of all learners, encouraging positive interactions and the development of prosocial skills, attitudes, and habits.

Moreover, inclusive education involves identifying and overcoming barriers to quality education in the mainstream education system; reasonable accommodation of the individual's requirements; and provision of support to school personnel to facilitate access to and participation in effective quality education for all (Sharma, 2020). That said, inclusive schools demonstrate an adherence to all principles described above, with most schools falling within a continuum of inclusive practice. Inclusion is a process, and schools that embrace inclusion have chosen to commit to embarking on that process of constant improvement and reflection.

International crisis: COVID-19

The coronavirus (i.e., COVID-19) originated in Wuhan, China, in late 2019 and quickly spread internationally, becoming a pandemic on a global scale not seen since the Spanish Flu of the early 20th Century. The final death toll from the virus, that has spread into a "global public health crisis" (UNESCO, 2020a), is yet to be fully realized, but at the time of writing this chapter there have been 189 million cases internationally, and 4.07 million deaths (COVID-19 Statistics, 2021). According to Sharma et al. (2020), writing in the midst of the epidemic, the more aggressively a country implemented medical containment measures, the more success it had in minimizing the public health crisis and the associated cost. Following some initial hesitation, countries around the world faced with local COVID-19 outbreaks implemented containment measures, some more aggressive than others. Italy, for example, mandated almost complete lock-down in response to skyrocketing infection rates, banning most citizens from leaving the immediate surroundings of their homes (Zampano, 2020). Other countries, such as Brazil, were perceived to lack the political will and strategy to implement very much at all in the way of containment measures, with President Bolsonaro downplaying the virus as a "little flu" (Picheta et al., 2020; Walsh et al., 2020). Moreover, as widely noted, other political leaders have responded to COVID-19 in a manner that may be perceived to reflect a complete disregard for scientific research, even so far as when being advised that COVID-19 could lead to a "mass loss of life" (Graham, 2020).

Beyond the ways that countries have responded to COVID-19 (e.g., closure of retail, parks; decreasing modes of available transportation; pausing sports seasons), what is of interest to us is how these nations—or closer to home, *local governments* and *school divisions*—have chosen to deal with schools. Part of the overall containment measures in many countries meant a swift and relatively sudden closure of schools. In many cases, schools shifted immediately to online delivery with only a few days' notice (in some cases, less than one day) around mid-March 2020, with students being expected to learn from home with support from their teachers in the form of video classes, phone calls, emails, homework assignments, discussion groups, and other physically-distant means. This remained the educational model employed in the countries that adopted the "learning from home" model up until approximately late-June 2020 or later.

Objectives

Stemming from COVID-19, we wanted to discover what such a drastic and immediate change in education delivery meant to learners in marginalized groups and how that change impacted adherence to inclusive principles; for example, for new migrant children and youth, children and youth with disabilities, and children and youth living in poverty. Such research queries included:

1. How did school systems around the globe address inclusivity during COVID-19?
2. What can we learn from this current situation about how best we can address inclusion and equity in schools in the future?

Method

This review was conducted using an emergent (or inductive) methodological approach. We do not make a claim that we employed a highly systematic approach as it was not possible nor would have allowed us to achieve the key objectives. Although having read much on the impact of COVID-19 on marginalized groups of learners prior to conducting this study was what prompted us to write about it, we attempted to suspend any preconceived ideas, allowing the sources of information to dictate the various themes identified. A variety of sources were employed to collect data. First, an international group of experts on inclusive education were asked to provide us with relevant media articles or government or NGO reports about education inequities that were further exposed once school went online in much of the world due to COVID-19. These experts were all members of the international Consortium of Inclusive Teacher Education and Development (CITED) group.

Next, we conducted an internet search for media articles and reports from national governments or international organizations including but not limited to the World Health Organization, United Nations Educational, Scientific and Cultural Organization (UNESCO), and Inclusion International. We also searched a social media platform (i.e., Twitter) in an effort to discern responses (while identifying and delving into any articles that were frequently shared on social media) directly from the public. Such terms

that were used for the data search included, but were not limited to, the following: inclusive education COVID-19; diverse learning needs during COVID-19; educational change COVID-19. Each retained article, report, or tweet-related information was examined individually and retained or discarded on the basis of relevance. Within each piece of literature, we focused on additional supporting documents, articles and reports that were referenced in the ones we initially reviewed; we strategically reviewed literature that was previously cited so that we could go down a “rabbit hole” in order to delve deeper into relevant topic matter.

The information was sorted into themes as a means of ascertaining in which areas existing inequities were further exposed once school went online in much of the world (or at least our parts [i.e., Canada, Australia]) with analysis being conducted in an ongoing manner from the commencement of the text examination (Miles and Huberman, 1994). A theme was only retained if a number of other textual pieces cross-validated one another within that theme employing a method known as the constant-comparative method (Glaser and Strauss, 1967) where thematic information is constantly compared to itself and to new information in order to ensure that the chosen theme is accurate. This process is known as triangulation (Lincoln and Guba, 1985), and while generally reserved for the analysis of interview, observational, and other types of qualitative data, it can also readily be applied to document analysis (see also Loreman, 2013). Moreover, we ensured the consistency of themes was met through the passing of information back and forth along with sharing articles, reports, and organizational information among each other. Once themes seemed to emerge, discussion took place among the research team to ensure understanding how and why the theme was identified.

Results and discussion: issues exposed by COVID-19

Evidently, the principles of inclusive practice did not carry over entirely from the regular school process (i.e., classroom-based learning) to what became the “new norm” (i.e., learning from home) because educational systems had not fully embraced and internalized them. Resulting from our examination of the related literature, three themes emerged, including: overreliance on technology; changes in teacher assistant roles; and disconnecting relationships among students.

In mid-March 2020, students worldwide were asked to continue their learning through an online platform from their homes. At first glance, for those with suitable levels of technology in the home, this type of request may have been a minor inconvenience. However, for students without access to suitable technology and whose learning needs require more specialized levels of attention, this type of request may have led to profoundly negative results. That said, we have not even taken into consideration whether or not parents/guardians were available in the homes to assist students as young as 5 years old with their new online learning environment. Without adequate physical space and technology in the home, along with minimal familial support during the online learning activities, it seems clear that the online learning environment could, and did, involve many barriers for students. Hence, it is critical to decipher the ways that school systems—along with local and national governments—have not upheld the responsibilities of delivering quality education for all learners.

Overreliance on technology

Since the time schools moved to online learning, much has been written concerning adequate resource provision related to the COVID-19 state that has impacted the educational landscape. Hence, as students were asked to stay home during the epidemic, one major resource at home became a required learning tool—the home computer with internet access. Students were immediately forced into altering their “way of learning”. Although this may seem to be a minor inconvenience, it turned out to be a barrier that caused numerous issues for many students, including students with disabilities. Although students, along with their parents/guardians, became reliant on technology from home, it became evident that not all homes had adequate technology to ensure a quality education. Because of the overreliance on technology on the part of the schools, gaps became exposed between the “haves” and the “have nots”; the “digital divide” was apparent and widened across countries.

Several articles outlined ways in which inadequate resource provision has been a problem for students with diverse learning needs. Inequalities in student access to adequate technology for remote learning have been widely reported, and for many students with additional learning needs, the increased reliance on technology for remote learning contributed to more challenges and barriers (Sharma et al., 2020). As argued by McClain-Nhlapo (2020), although the COVID-19 state has relied heavily on educational technology solutions (e.g., computers, online learning portals), these learning tools were not always, or readily, accessible to learners with disabilities or complex learning needs. McClain-Nhlapo (2020) also pointed to the fact that internet accessibility proved to be an issue, when considering various locations across the globe (e.g., 2G, 3G, 4G access is not always possible). Although there may be a home computer available, without accessible internet connections, usually the home computer cannot meet learner needs. According to Strauss (2020), schools cannot provide access to learners by merely developing alternative online learning plans that are simply compliant with guidelines without considering deeply how each learner’s needs are being met.

Moreover, due to the overreliance on technology, home/school partnerships were negatively impacted. For example, Whitley (2020) discussed how parents/guardians who typically collaborate with school staff and community organizations were found to particularly struggle after the disappearance of face-to-face networks and collaborative partnerships in supporting learners. In fact, according to Whitley, some students who specifically thrive on predictable routines, struggle with transitions and rely heavily on deep-rooted teacher/student relationships built over time were found to be experiencing a range of negative emotions (e.g., worry, anger, sadness) with some believing they lost a second home without comprehending any reason. Likewise, in the midst of a clear lack of home/school communication in some instances, Preston (2020) discussed a situation in which, during the first

week of online learning, a student received a downloadable set of learning activities that included zero modifications or considerations toward the known diverse learning needs of the learner. Without the maintenance of strong home/school partnerships continuing into the COVID-19 state of schooling (i.e., learning from home), it is no wonder there have been learners with disabilities found to be communicating and expressing their emotions during this largely difficult situation in ways that are quite challenging for families (Whitley, 2020).

That said, numerous positive outcomes have arisen related to inclusive practices in the online learning environment. According to Sharma and May (2020), educators who have invested additional efforts in developing strong home/school connections reported improved engagement with all learners, including those with additional learning needs. Families that reported that the school was responsive to individual learner needs, such as by altering teaching and learning plans and providing additional resources, described more effective learning experiences for their children (Sharma et al., 2020). Likewise, some teachers made extraordinary efforts to facilitate remote learning by making efficient use of available technology in different parts of the world (Sharma et al., 2020). For example, some teachers provided their phone numbers to families so that they could contact them directly to discuss effective ways to support learners at home. Clearly, these teachers were inclusive as they were always thinking of addressing barriers that learners were facing and were discovering new solutions to address barriers daily; these teachers proved to be highly resourceful when taking into consideration barriers related to technology and home/school connections. They did not always know how to address all barriers, but they knew working in close partnership with families—building and maintaining strong home/school partnerships—they could address the “digital divide” to some extent. Moreover, in some of the economically emerging nations (e.g., India, Bangladesh), use of easily accessible multimedia (e.g., radio or TV broadcast of lessons) was employed to deliver lessons. Evidently, although technology-related barriers were found to exist, there were teachers who explored and discovered manageable ways to close the “digital divide” without particularly requiring excess amounts of new resources and funding. In a recent study, many families described how much they have valued the enhanced levels of communication with teachers throughout the COVID-19 state, reporting a more thorough understanding of their child’s learning needs (Sharma et al., 2020).

Changes in the teacher assistant roles

In some instances, a dramatic cut in education-related costs occurred during the epidemic with some school districts seeing their teacher assistants laid off. For example, in Alberta, Canada, the apparent cost savings associated with remote learning was used by the Government of Alberta as a reason for providing significant budget cuts to all Alberta school jurisdictions (Joannou, 2020; Macfarlane, 2020). According to Macfarlane (2020) in a CTV News report, “Education Minister Adriana Lagrange announced what she promised will be a temporary cut in order to redirect up to \$128 million to the province’s response to COVID-19.” Further, the report contended that, “... education assistants will be paid through the end of April ...”

Teacher assistant layoffs exposed two primary issues: (1) throughout the years, many teachers have delegated the responsibility of educating some students to teacher assistants and many of the criticisms of the cuts equated to teacher assistants being actual teachers for students with disabilities, and (2) on the other hand, the work of teacher assistants was legitimately justifiable for fulfilling a support role in the learning environment, even in a remote learning environment, which helped lead to successful student learning from home. The massive loss of teacher assistants led to significant negative results for students and, simultaneously, seemed to put a spotlight on a “perceived lower status” in the learning environment. When we refer to a “perceived lower status”, this is something we believe was communicated by the decision-makers to simply lay off teacher assistants. We (three authors) have all spent time in the classroom. We understand the necessity of qualified teacher assistants who are caring, knowledgeable, and experienced in fulfilling the duties as educational assistants in the learning environment (and that includes remote learning). Hence, we must ask what the perception of teacher assistants is from the decision-makers when laying them off seemed to be their course of action. More specifically, as alluded to in Heidenrich (2020), when referring to teacher assistants, along with other support services, they “were no longer working” during the school closures caused by COVID-19. When asked about this decision, a governmental Opposition Leader contended that teacher assistants are critical for remote learning success as they are required to assist students with learning challenges in the new online learning landscape (Heidenrich, 2020).

Conversely, there are examples in other countries where schools were more innovative in the way they employed and worked alongside teacher assistants (Sharma et al., 2020). For example, in Australia, schools allocated a small group of learners in an online environment to a teacher assistant for part of the lesson. The teacher assistant did not deliver the complete lesson, but worked with the allocated students to scaffold/reinforce the content covered during the lesson. The teacher assistants were alongside the classroom teacher in “breakout rooms”—a Zoom feature which allows the teacher to split the class into smaller groups. Moreover, teacher assistants were found to describe the importance of their role in keeping learners with diverse learning needs connected to school and classmates by employing a variety of strategies (Sharma et al., 2020). Some schools in Australia also continue to provide on-campus services to children and youth of essential service workers (e.g., nurses, doctors, ambulance staff) as well as to students who had very high learning support needs. As a result of Australia’s perception toward the importance of teacher assistants related to remote learning, teacher assistants proved to be instrumental in supporting all these learners during COVID-19. Hence, it seemed to really come down to the overall perception of what a teacher assistant role is—and can be—for the success of all learners. It is apparent from our discussion that two polarizing beliefs emerged when considering the roles and responsibilities of teacher assistants throughout the epidemic.

Disconnecting relationships among students

School closures due to COVID-19 precipitated a significant deterioration of peer relationships. In the case of students with disabilities, this was seen to be negatively impactful. Students became reliant on parental support to complete schooling, and oftentimes, the parent/guardian became a substitute peer for students. Because some parents/guardians had to remain working as essential workers, while others could not or did not support student learning for other reasons, some students were not fully included in the online learning community or received a second-rate education, for family reasons. This is presumably also the case when classes are face-to-face. COVID-19 simply highlighted that fact when it comes to the class community.

Moreover, important socialization needs were ignored in favor of an approach that focused only on the academic “basics”. This highlights that basic academics are still the focus of schools, and the rest operates on the margins. For example, [CBC \(2020\)](#) reported that in Alberta, Canada, language/literacy and mathematics/numeracy (K-3), with the suggestion of incorporating science and social studies (4–6), would be the primary focus during the remote learning stage. When considering social and emotional learning, one must ask what considerations took place in this decision for student learning. For example, it seems unrealistic that specific learning needs can be reached from home when such barriers include the complete absence of a class community. When considering support for social and emotional learning of students during COVID-19, consideration should be given to the importance of peer connections. According to [Whitley \(2020\)](#), connections with peers are critically important for students; the development and maintenance of friendships is not always easy for students with disabilities. As argued in [Whitley et al. \(2009\)](#), “... a number of studies have documented that when students feel engaged with and enjoy their school they are likely to experience high academic achievement and school completion ([Beran et al., 2008](#); [Buhs, 2005](#); [Finn, 1993](#); [Ontario Ministry of Education and Training, 2005](#)).” That said, numerous students rely on school-based friendships and may not have these peers in their home environments ([Whitley, 2020](#)).

Likewise, according to [Strauss \(2020\)](#), for many students with disabilities, meaningful access to a conducive learning environment means providing resources so that they can access the general education curriculum and be actively engaged in the learning community—a critical step toward developing a sense of belonging in today’s society. Further, a large portion of students with intellectual disabilities have been found to receive minimal to no access to the academic instruction or social and emotional learning that promotes a sense of belonging ([Strauss, 2020](#)). In addition, much has been written concerning positive attitudes and welcoming environments related to COVID-19 that has impacted the online learning environment. For example, it has been reported that the education systems have not adapted to meet learner needs, leaving up to as much as one-quarter of 15-year-old students globally reporting a sense of exclusion from the learning environment ([UNESCO, 2020a](#)). [UNESCO \(2020a\)](#) contended that COVID-19 actually further exacerbated exclusion in schools specifically because, in the Latin America and Caribbean regions, with only approximately one-quarter of the countries having inclusive education laws covering all learners, it is believed that inequalities will have increased throughout the COVID-19 state ([UNESCO, 2020a](#)).

Although much has been discussed when considering the deterioration of classmate relationships throughout the remote learning stage, there have been some effective measures taken to counter this issue. One such example, when considering students with disabilities, is the Special Friends Network, a grassroots organization operating in Ontario, Canada, that has organized Zoom learning sessions/meetings so that students with all types of disabilities are afforded opportunities to communicate, participate in games, and engage in artwork activities with peers ([Whitley, 2020](#)). This idea seems so achievable, when considering the online learning environment, and can help lead to many deep and lasting relationships while also affording opportunities for collaborative learning.

Moreover, as shared in [Gonser \(2020\)](#), the idea of trying to frequently reach out to say “hello” has been found to elevate peer relationships throughout the epidemic. Striving to ensure that students know that teachers are thinking about them, care for them, and miss them is a resourceful tactic to help build and/or maintain class relationships. That said, for learners with no internet access, it is suggested that teachers try to connect through a phone call and consider rotating through small groups of learners each day to make it a manageable task. According to a teacher, as shared by [Gonser \(2020\)](#), reaching out by calling each student can be quite time-consuming, but the teacher observed a positive outcome from such an effort to help build and/or maintain quality class relationships.

Likewise, when considering regular check-ins, a teacher has established an idea that students connecting with peers has become part of the class homework ([Gonser, 2020](#)). According to the teacher, each learner was assigned a task to connect with one classmate, and to report back to the teacher to share how the peer is doing on that particular day ([Gonser, 2020](#)). After modeling the effective communicative practice—including guidelines—for the students, the learners can choose how they want to check in (e.g., Email, Skype).

When considering the various ways that educators are striving to build and/or maintain strong class relationships throughout the remote learning stage, it again seems to be the resourcefulness of the teachers that plays an essential role in quality education for all learners. That said, perhaps the narrative around “more funding is needed” should be lessened and additional support for teachers to build and/or maintain strong class relationships should be discussed in more depth.

Principles of inclusive practice in schools

Building the plane, while flying

COVID-19 has changed the way we see and work in our world. That said, the educational landscape has also changed and educators worldwide are striving to adapt their practices—rapidly due to the pandemic—to ensure meaningful student learning occurs. Teachers are trying to become more effective in their practices in a time when levels of anxiety across whole school communities are at exceedingly high levels; they are building a plane, while flying. During the first months of the pandemic, much was discussed concerning teacher training when working with students with disabilities, and this has been further focused upon during the extended impact of COVID-19. According to [UNESCO \(2020b\)](#), a quarter of educators across 48 countries reported they would have liked more training for teaching students with disabilities. Further, in practice, school hours were drastically reduced since mid-March 2020 for online learning. As widely accepted by educators, some students require more time than others for practice and application of what they learn. Due to both the reduced timeframes given to the school day along with teacher unreadiness to teach students with disabilities, further exposure to a lack of real attention to the principles of UDL was brought to light throughout this study. As we carried out our objectives, we maintained a close focus on the two research questions (previously noted).

In this section, we focus on how schools have addressed inclusivity during the time when schools were closed and students were engaged in remote learning. Throughout the recent literature related to COVID-19 and meeting the needs of all learners, it is evident that not all principles of inclusive practice were adhered to. Hence, we will provide a commentary that will focus on much of what we found throughout this study.

First, when considering the home/school partnership, it is widely understood that often parents/guardians are a child's first teacher. They raise and educate their child(ren), ideally entering into a partnership with schools/teachers in that endeavor once school age is reached. Children have been found to do best when learning is supported at home, and when teachers and parents/guardians help students work toward commonly understood objectives. Some families are better positioned to participate in the education of their child(ren) than others, be it as a result of economic circumstances, education attainment levels, or other contextual differences. In many cases, COVID-19 ignored many of the underpinnings of the home/school partnership, oftentimes shifting the primary locus of responsibility for education from school back to the home. Opportunities for communication and collaboration between parents/guardians and teachers became strained, mediated as they were for the most part through online appointments as opposed to face-to-face and informal interactions. Those families without the skills and resources to navigate their child's education with reduced teacher interaction inevitably suffered, and the maintenance of home/school partnerships was reported only minimally in the literature as being a priority during the remote learning stage due to COVID-19. For example, in some cases, teachers became resourceful and connected regularly, or at least frequently, with parents/guardians through phone calls which proved to be more accessible for some families due to no, or minimal, internet access. Moreover, when discussing a recent study, [Sharma and May \(2020\)](#) contended that some education support staff described how they would phone learners prior to online lessons to discuss learning goals, in addition to phone or email contact following lessons to provide any required support.

Likewise, while the provision of adequate resources is an oft-cited principle of inclusive education, and an issue most commonly identified by teachers, the school closures helped demonstrate that in some school systems the students who were impacted the most by resource changes during a time of disruption in schools were those living with disabilities. The case of Alberta, Canada provides a salient example of that. As discussed earlier in this chapter, the Government of Alberta cut \$128 million from support services, leaving students without support from teacher assistants and additional substitute teachers at a time when they likely needed more, not less, support. While it is right that teachers be viewed as the provider of education to all students, including those with disabilities, Alberta teachers were themselves expected to adapt to a sudden and profound change in their pedagogy as they moved rapidly from in-person to online delivery. The reduction of support to these teachers left many students with disabilities struggling to engage and genuinely and actively participate in their education. Conversely, and as previously noted in this chapter, Australia made decisions on the opposite end of the spectrum when it came to the roles and responsibilities of teacher assistants. Teacher assistants in Australia maintained a solid position in the remote learning environment and were key stakeholders in learner support (e.g., working with learners in online breakout rooms).

Moreover, school closures due to COVID-19, in many ways, provided all students with additional barriers to education. An over-reliance on technology meant that where delivery switched quickly from in-person to online instruction, those who had adequate technology at home and a suitable study space faced fewer issues associated with access than those who did not. However, the additional barriers presented by online learning compounded barriers to learning already being faced by many students with disabilities. Learning, by and large, became mediated by audio and visual modalities through a computer. Those with visual impairments, those who are deaf or have hearing impairments, and those who have difficulty interacting through a computer for other reasons, were found to be left behind. The reduced provision of human resources to facilitate learning resulted in learning opportunities for students with disabilities that were uneven at best, and non-existent at worst. Moreover, while most teachers were aware to varying degrees that online pedagogy is very different from in-person pedagogy, few had taught in an online environment, or if they had it was often supplemental to in-person classes. Suddenly, with little notice and almost no support, they were expected to change modalities and continue to offer quality learning experiences in an online format of which many had minimal knowledge. This was difficult for those teaching very young children, and those others who had students who required additional support and assistance. Teachers, for the most part, were just expected to do their best with little thought being given to the provision of expertise. In some regions, subject areas, such as physical education, were abandoned and teachers were re-assigned to subject areas they may

have been largely unfamiliar with. For all the talk of having expert teachers in classrooms, this often seemed less important than simply getting through the day during the COVID-19 pandemic. Again, aside from the barriers, throughout the remote learning stage, some educators placed a large focus on reaching out to students and families to assist with technological-related issues and found ways to communicate relevant information through other means (e.g., phone calls).

If schools only existed for the purposes of learning academic content it would be relatively easy to argue that there are other more efficient ways of accomplishing this, and that they are not necessary. Schools, however, also serve a number of purposes, and one of these purposes is the development of social and emotional maturity, typically through communication and friendships. For many learners, school is a primary source of social interaction with their peers, especially in rural areas where due to geography social interactions outside of school are sometimes more difficult to engage in. The COVID-19 pandemic helped demonstrate that the social and emotional development of learners is not a high priority for school systems. While academic classes were scheduled online, there is scant evidence that much was done to maintain social relationships between learners through the facilitation of unstructured settings such as children experience during recess and other breaks in the school day. Educators, perhaps feeling overwhelmed and out of their depth, tended to focus on “the basics” of delivering academic instruction, with social and emotional development very much taking a back seat—and such a focus was set out explicitly by various governmental departments (e.g., Alberta Education). That said, there were some effective measures taken to help build and/or maintain peer relationships during remote learning such as scheduling homework-like assignments for students to check in with peers. This strategy led to profound results for learners staying in touch with classmates while helping to enhance social and emotional development.

What we can learn from COVID-19

Our focus now turns to research question 2, *What can we learn from this current situation about how best we can address inclusion and equity in schools in the future?* As we continue to navigate our way through this COVID-19 state, what has become apparent from this work is that specific principles of inclusive practice have received minimal, if any, attention from various levels extending from governmental-to classroom-level decisions all the way to the other side of the spectrum where educators became quite resourceful in helping to reach all students via remote learning. In fact, according to [Sharma and May \(2020\)](#), in response to COVID-19-related challenges, many educators have rapidly altered their practices and created innovative approaches to ensure learning for all. Throughout this study, there have been some clear messages related to moving forward toward more inclusive education. We would like to offer the following as a step forward for schools globally.

- *Create strong home/school partnerships.* School communities must continue to explore and discover ways to establish quality, deep-rooted relationships with families to ensure there is a true understanding of learner needs from both the school and family perspective. It is of the utmost importance that educators continue to discover ways to ensure effective communication and understanding is focused upon to promote learning for all.
- *Develop a deeper understanding of teacher assistant roles.* When it has been perceived (and even stated) that “teacher assistants now will not be working” during the remote learning stage, it becomes much clearer as to the overall perception toward the roles teacher assistants fulfill for student learning. The critical importance of teacher assistants must be understood from the perspective of educators at the school levels, but also decision-makers at the government level. Laying off teacher assistants in the midst of a pandemic seemed to be quite the opposite of demonstrating an understanding of what teacher assistants’ roles are in schools. However, some schools demonstrated a great deal of knowledge about the importance of teacher assistants (e.g., Australia schools) where teacher assistants played an integral role in student learning throughout the remote learning stage. We note that developing a more thorough understanding of teacher assistant roles is a necessity for all key educational stakeholders.
- *Build and/or maintain peer relations.* Developing ways for learners to build and/or maintain strong peer relations with classmates is a vital piece of the learning pie for all. Although there was much shared throughout the literature about lack of communication among classmates, there were also several terrific ideas that teachers carried out during the remote learning stage that included team-building within the online learning environment. For example, affording students with opportunities to reach out to each other during regular check-ins seemed to help enhance peer relations for students, including students with disabilities who in many cases were found to be quite excluded when practices such as this were not carried out. We contend that the building and/or maintaining of peer relations will lead directly to higher levels of social and emotional wellness for all learners.

Conclusion

In this chapter, we argue that school closures due to COVID-19 have had a significant impact on student learning, and in particular, students with disabilities. Although much can be learned from how school systems reacted to the COVID-19 state by closing schools and turning the in-class learning environment into an online delivery model where students were asked to stay home. At first glance, for those with adequate home support (e.g., family) and technology (e.g., home computer, internet access), this may have seemed to be a seamless, if not rewarding in some ways, request. However, on the flip side, this could not be further from the truth. For learners without family support during the day (e.g., parents are essential workers) and without adequate technology (e.g., no home

computer, no internet access), this request had a negative impact on their learning. In some cases, governments, school systems, and front-line teachers became very resourceful and worked to enhance student learning for all during the remote learning stage. However, evidently, there were also cases where students were left to their own devices for learning such as having limited home/school interaction, losing access to teacher assistant support, and having minimal, in any, communication with peers.

Overall, we believe COVID-19 has played a major role in shedding light on issues related to the principles of inclusive education in schools. In fact, we have concluded that COVID-19 has put a spotlight on a perfect storm. With lack of teacher readiness to help all learners during remote learning combined with fewer resources, and difficulties in communication with families due to technology issues, the remote learning delivery model exposed many of the commonly used nice statements the educational landscape employs as being false in practice.

As we head into the post-COVID-19 world, we are hopeful that major gains can be observed in the areas that became focal points in this chapter to help reach all learners, specifically, students with disabilities. First, a place to begin when it comes to teaching students with disabilities is to ensure policy makers obtain a clear measure of the situation when it comes to actual numbers of students with disabilities in schools. According to UNESCO (2020c), when considering a chronic lack of quality data on students left behind, almost half of low- and middle-income countries do not collect enough education data about students with disabilities. Although household surveys are critical for obtaining education data related to individual characteristics, 41% of countries—representing 13% of the world's population—did not conduct surveys or disseminate data from such surveys, leaving the world with inadequate information on such a vital piece of the educational picture (UNESCO, 2020c). That said, when considering the significant number of students with disabilities, it is hoped that a deeper level of attention is given to home/school partnerships, teacher assistant roles, and peer relations in the learning environment.

For now, although COVID-19 has played a role in increasing the educational divide for many people across the world (Sharma et al., 2020), we can use the information obtained from this study to help guide governments, school systems, teachers, teacher assistants, and families in developing ways to enhance learning for all, and in particular, learning for students with disabilities. And, in the end, if we can help school communities become more aware of and willing to employ the principles for inclusive education discussed throughout this chapter, we will surely see more quality inclusive education worldwide.

References

- Ackah-Jnr, F.R., Danso, J.B., 2019. Examining the physical environment of Ghanaian inclusive schools: how accessible, suitable and appropriate is such an environment for inclusive education? *Int. J. Incl. Educ.* 23 (2), 188–208.
- Ahmed, M., Sharma, U., Deppeler, J., 2012. Variables affecting teachers' attitudes towards inclusive education in Bangladesh. *J. Res. Spec. Educ. Needs* 12 (3), 132–140.
- Armstrong, C., Armstrong, D., Spandagou, I., 2010. *Inclusive Education: International Policy and Practice*. Sage, Los Angeles, US.
- Beran, T.N., Hughes, G., Lupart, J., 2008. A model of achievement and bullying: analyses of the Canadian National Longitudinal Survey of Children and Youth data. *Educ. Res.* 50 (1), 25–39. <https://doi.org/10.1080/00131880801920379>.
- Booth, T., 1995. Mapping inclusion and exclusion: concepts for all? In: Clark, C., Dyson, A., Millward, A. (Eds.), *Towards Inclusive Schools?* Fulton, London.
- Bradford, B., Eblie Trudel, L., Katz, J., Sokal, L., Loreman, T., 2021. Promising practices for preparing Canadian teachers for inclusive classrooms: analysis through a transformative learning lens. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2021.1882058>.
- Buhs, E.S., 2005. Peer rejection, negative peer treatment, and school adjustment: self-concept and classroom engagement as mediating processes. *J. Sch. Psychol.* 43 (5), 407–424. <https://doi.org/10.1016/j.jsp.2005.09.001>.
- CBC, March 20, 2020. Alberta sets guidelines for students to learn at home while classes cancelled. COVID-19 News. <https://www.cbc.ca/news/canada/edmonton/alberta-sets-guidelines-for-students-to-learn-at-home-while-classes-cancelled-1.5504442>.
- Christenson, S.L., 2004. The family-school partnership: an opportunity to promote the learning competence of all students. *Sch. Psychol. Rev.* 33 (1), 83–104. <https://doi.org/10.1080/02796015.2004.12086233>.
- COVID-19 Statistics, 2021. Cases. https://www.google.com/search?q=covid-19+statistics&rlz=1C1GCEU_enCA821CA821&aq=COVID-19+statistics&aqs=chrome.0.0i31i433i457j0l7.11151j0j7&sourceid=chrome&ie=UTF-8#wptab=s:H4sIAAAAAAAAAAONgVuLVt9c3NMwySk6OL8zJecTozS3w8sc9YsmnSWtO.Xm004elKzsgvd80rySypFNLjYo0yVlgEpVB1ajBI8XOhCvHsYul2SE3MKckllkksKV7EKpycX5aZomtoqVAMFMgsLlMLgYA75so1IMAAAA.
- De Boer, A., Pijl, S.J., Minnaert, A., 2011. Regular primary school teachers' attitudes towards inclusive education: a review of the literature. *Int. J. Incl. Educ.* 15 (3), 331–353.
- Deppeler, J., Loreman, T., Smith, R., Florian, L. (Eds.), 2015. *Inclusive Pedagogy Across the Curriculum. International Perspectives on Inclusive Education*. Emerald, Bingley, UK.
- Erkic, M., Durak, S., 2013. Tolerable and inclusive learning spaces: an evaluation of policies and specifications for physical environments that promote inclusion in Turkish primary schools. *Int. J. Incl. Educ.* 17 (5), 462–479.
- Finn, J.D., 1993. *School engagement and students at risk*. National Center for Education Statistics, Washington, DC.
- Glaser, B., Strauss, A., 1967. *The Discovery of Grounded Theory Strategies for Qualitative Research*. Sociology Press, Mill Valley, CA.
- Gonser, S., March 25, 2020. Social and Emotional Learning: 7 Ways to Maintain Relationships During Your School Closure. Edutopia. <https://www.edutopia.org/article/7-ways-maintain-relationships-during-your-school-closure>.
- Goldan, J., Lambrecht, J., Loreman, T. (Eds.), 2021. *Resourcing Inclusive Education. International Perspectives on Inclusive Education*, vol. 15. Emerald, Bingley, UK.
- Golden, A., 1997. *Memoirs of a Geisha*. Alfred K. Knopf, NY, US.
- Goodall, J., Montgomery, C., 2014. Parental involvement to parental engagement: a continuum. *Educ. Rev.* 66 (4), 399–410. <https://doi.org/10.1080/00131911.2013.781576>.
- Graham, D.A., April 29, 2020. Why Trump was deaf to all the warnings he received. The president's incuriosity and paranoia hobbled his response. *The Atlantic*. <https://www.theatlantic.com/ideas/archive/2020/04/how-many-warnings-did-trump-ignore/610846/>.
- Heidenrich, P., April 1, 2020. Kenney says many Alberta school boards were planning layoffs before his government announced them. *Global News*. <https://globalnews.ca/news/6766469/kenney-alberta-education-cuts-covid-19-coronavirus/>.
- Hornby, G., Blackwell, I., 2018. Barriers to parental involvement in education: an update. *Educ. Rev.* 70 (1), 109–119. <https://doi.org/10.1080/00131911.2018.1388612>.
- Joannou, A., March 29, 2020. COVID-19: Alberta temporarily cuts funding for educational assistants, transportation while classes cancelled. *Edmont. J.* <https://edmontonjournal.com/news/politics/covid-19-alberta-to-temporarily-reduce-education-funding-amid-covid-19>.
- Katz, J., 2012. *Teaching to Diversity: The Three-Block Model of Universal Design for Learning*. Portage & Main Press, Winnipeg, CA.
- Katz, J., 2013. The three block model of universal design for learning (UDL): engaging students in inclusive education. *Can. J. Educ.* 36 (1), 153–194.

- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage Publications, Newbury Park, CA.
- Loreman, T., 2013. Measuring inclusive education outcomes in Alberta, Canada. *Int. J. Incl. Educ.* 18 (5), 459–483.
- Loreman, T., 2017. Pedagogy for inclusive education. *Oxford Research Encyclopedia of Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.148>.
- Luciak, M., Biewer, G., 2011. Inclusive education in Austria: a comparative analysis. In: Artiles, A.J., Kozleski, E.B., Waitoller, F.R. (Eds.), *Inclusive Education: Examining Equity on Five Continents*. Harvard Education Press, Cambridge, Mass, pp. 17–44.
- Macfarlane, B., March 30, 2020. Province's sudden education cut set to add 26,000 people to unemployment roll. CTV News. <https://calgary.ctvnews.ca/province-s-sudden-education-cut-set-to-add-26-000-people-to-unemployment-roll-1.4875173>.
- McClain-Nhlapo, C., May 11, 2020. An inclusive response to COVID-19: education for children with disabilities. *Global Partnership Educ.* <https://www.globalpartnership.org/blog/inclusive-response-covid-19-education-children-disabilities>.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, second ed. Sage.
- Mitchell, D., 2015. Inclusive education is a multi-faceted concept. *Center Educ. Policy Stud. J.* 5 (1), 9–30.
- Mittler, P., 1979. *People Not Patients: Problems and Policies in Mental Handicap*. Methuen, London.
- Mittler, P., 2000. *Working Towards Inclusive Education: Social Contexts*. David Fulton Publishers, London.
- Mittler, P., 2010. *Thinking Globally Acting Locally: A Personal Journey*, Milton Keynes and Bloomington. AuthorHouse, Ind.
- Movkebayeva, Z.A., Oralkanova, I.A., Mazhinov, B.M., Beisenova, A.B., Belenko, O.G., 2016. Model of formation for readiness to work within inclusive education in teachers. *Int. J. Environ. Sci. Educ.* 11 (11), 4680–4689.
- Norwich, B., 1990. *Reappraising Special Needs Education*. Cassell Educational Limited, London.
- Ontario Ministry of Education and Training, 2005. *Early school leavers: Understanding the lived reality of student disengagement from secondary school*. Author, Toronto, ON.
- Opertti, R., Brady, J., Duncombe, L., 2009. Moving forward: inclusive education as the core of education for all. *Prospects* 39 (3), 205–214.
- Picheta, R., Cotovio, V., Darlington, S., May 18, 2020. As hospitals in Brazil teeter on the brink of collapse, Bolsonaro does pushups with supporters. *CNN World*. <https://www.cnn.com/2020/05/18/americas/brazil-coronavirus-sao-paulo-bolsonaro-rally-intl/index.html>.
- Preston, C., March 31, 2020. "I hate COVID-19": kids with disabilities struggle to adjust as schools close: isolated at home with few services to help, a mother in San Francisco is terrified that her daughter's progress will evaporate. *News*. <https://www.nbcnews.com/news/education/i-hate-covid-19-kids-disabilities-struggle-adjust-schools-close-n172906>.
- Reicher, H., 2010. Building inclusive education on social and emotional learning: challenges and perspectives—a review. *Int. J. Incl. Educ.* 14 (3), 213–246.
- Schwab, S., Sharma, U., Loreman, T., 2018. Are we included? Secondary students' perceptions of inclusion climate in their schools. *Teach. Teach. Educ.* 75, 31–39.
- Sebba, J., Ainscow, M., 1996. International developments in inclusive schooling: mapping the issues. *Camb. J. Educ.* 26 (1), 5–18. <https://doi-org.ezproxy.aec.talonline.ca/10.1080/0305764960260101>.
- Sharma, U., 2020. Inclusive education in the Pacific: challenges and opportunities. *Prospects*. <https://doi.org/10.1007/s11125-020-09498-7>.
- Sharma, U., Mishra, V., Yellishetty, M., Bhide, A., Padhye, M., 2020. COVID-19: Testing Time for Leadership Globally. *The Medium*. <https://medium.com/@milindsathye/covid19-testing-time-for-leadership-globally-63023ca2d88d>.
- Sharma, U., May, F., 2020. Inclusive Education Amid COVID-19: Lessons From Teachers Around the World. *Monash University*. <https://lens.monash.edu/@education/2020/10/05/1381365/inclusive-education-during-covid-19-lessons-from-teachers-around-the-world>.
- Slee, R., Allan, J., 2001. Excluding the included: a reconsideration of inclusive education. *Int. Stud. Sociol. Educ.* 11 (2), 173–192. <https://doi.org/10.1080/09620210100200073>.
- Starcić, A.I., 2010. Educational technology for the inclusive classroom. *Turkish Online J. Educ. Technol.* 9 (3), 26–37.
- Strauss, V., July 15, 2020. Virtual education was "a disaster" for her son with down syndrome. Here's what students with disabilities really need. *The Washington Post*. <https://www.washingtonpost.com/education/2020/07/15/virtual-education-was-disaster-my-son-with-down-syndrome-heres-what-students-with-disabilities-really-need/?arc404=true>.
- Taylor, S., Sidhu, R.K., 2012. Supporting refugee students in schools: what constitutes inclusive education? *Int. J. Incl. Educ.* 16 (1), 39–56.
- UNESCO, 2009. *Policy Guidelines on Inclusion in Education*. UNESCO, Paris.
- UNESCO, 2020a. The COVID-19 Outbreak & UNESCO's Mission. COVID-19 Response. <https://en.unesco.org/covid19>.
- UNESCO, 2020b. Access to Information More Vital Than Ever. COVID-19. <https://en.unesco.org/news/access-information-more-vital-ever>.
- UNESCO, June 23, 2020c. UNESCO Shows 40% of Poorest Countries Failed to Support Learners at Risk During COVID-19 Crisis and Urges Inclusion in Education. *Global Education Monitoring Report*. https://en.unesco.org/gem-report/sites/default/files/2020_Press_Release_EN.pdf.
- Walsh, N.P., Shelley, J., Duwe, E., Bonnet, W., May 25, 2020. Bolsonaro Calls Coronavirus a "Little Flu." Inside Brazil's Hospitals, Doctors Know the Horrifying Reality. *CNN World*. <https://www.cnn.com/2020/05/23/americas/brazil-coronavirus-hospitals-intl/index.html>.
- Whitley, J., June 1, 2020. Coronavirus: Distance Learning Poses Challenges for Some Families of Children With Disabilities. *The Conversation*. <https://theconversation.com/coronavirus-distance-learning-poses-challenges-for-some-families-of-children-with-disabilities-136696>.
- Whitley, J., Lupart, J.L., Beran, T., 2009. The characteristics and experiences of Canadian students receiving special education services for emotional/behavioural difficulties. *Except. Educ.* 19, 14–31. <https://ir.lib.uwo.ca/eei/vol19/iss1/3>.
- World Bank Group, 2020. Every Learner Matters: Unpacking the Learning Crisis for Children With Disabilities. <https://openknowledge.worldbank.org/bitstream/handle/10986/31946/Every-Learner-Matters-Unpacking-the-Learning-Crisis-for-Children-With-Disabilities.pdf?sequence=1&isAllowed=y>.
- Zampano, G., March 18, 2020. Italy Cracks Down on Citizens Amid Lockdown: Number of Checks, Complaints Rises as Authorities Try to Enforce Stringent Lockdown Rules. *Anadolu Agency*. <https://www.aa.com.tr/en/europe/italy-cracks-down-on-citizens-amid-lockdown/1769892>.

Delivering on the promises of tertiary education in low- and middle-income countries: risks, sustainability and inclusion

Rachel Outhred and Fergal Turner, Oxford MeasurEd, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	126
The multiple recent shifts in the tertiary education landscape	126
Challenges to delivering tertiary education during the pandemic	127
The cases of Ethiopia, India and Peru	128
Ethiopia	128
India	129
Peru	130
Looking to the future	130
The massification trap	131
Lesson from high-income countries	131
Lessons from the massification of basic education	131
The opportunity trap	131
The globalization traps	132
Conclusion	133
References	134

Introduction

It is widely accepted in policy dialogue globally that the case for investing in education in Low- and Middle-Income Countries (LMICs) is strong. As UNICEF states *“Investing in education isn’t just the right thing to do, it’s smart economics”* (Brandt, 2015).

Within the broad policy dialogue on education, the promise of tertiary education is three-fold. Firstly, it is argued that tertiary education contributes to human development, improving young people’s prospects later in life, including the ability to actualize capacities and capabilities to be both a holistically developed human and to build knowledge and skills for decent work and other activities that derive pleasure and fulfilment (Sypnowich, 2018). Secondly, it is argued that increased participation in tertiary education at the national level may lead to national development (McCowan and Schendel, 2016), stimulating economic growth and strengthening crucial public services. Thirdly, it is argued that tertiary education contributes to social cohesion, whereby individuals enrolled in tertiary education can experience social mobility (Mehrotra, 2005).

These rationales for investments in tertiary education in low- and middle-income countries are further supported by the increasing demand for tertiary education, *“fueled by the twin pressures of a burgeoning youth population and an improvement in secondary school completion rates ...”* (McCowan and Schendel, 2016, p. 2). Increasingly tertiary education participation is gaining space within national education plans in LMICs due to national aspirations to realize the promise of tertiary education’s contribution to the public good. The national aspiration for tertiary education is articulated in terms of increased productivity and prosperity by preparing graduates for occupations, increased innovation in industry, the creation and distribution of knowledge, reducing social inequalities, increasing the preparedness of citizens for democratic decision-making and to advance free expression (Marginson, 2016).

However, the three promises of tertiary education do not always sit comfortably together. The tension between Nussbaum’s (1999) and Sen’s (2000) seminal work on capabilities as development and traditional approaches that focus on economic growth as development is fundamentally evident in the promises of tertiary education in LMICs.

Within this chapter, we review the multiple recent shifts in the tertiary education landscape and reflect on the implications of COVID-19 for the sector. We then consider the treble promises of tertiary education in the cases of three countries: Ethiopia, India and Peru. Within each of these cases, we review the realization of the promises related to education for all at the primary and secondary levels, we outline challenges to delivering tertiary education during the pandemic and review funding and market trends within the sector. Finally, we explore some of the risks that accompany the processes of expanding access to tertiary education, and how these risks may undermine delivering on tertiary education’s triple promise.

The multiple recent shifts in the tertiary education landscape

The policy logic of investments in the tertiary sector are nested in conceptualisations of tertiary education’s transformative power to improve social mobility and enact change in society. As Heyneman et al. (2007) state: *“That higher education can play an important role in promoting understanding between social groups and building a sense of unity in diverse societies has long appealed to educators and policy-makers alike.”* However, while education is often held up as a mechanism able to challenge and redress inequality in society, there is

a strong sociological tradition arguing that university and school structures reproduce social inequalities and collude to exclude certain students (both within and from education) (Outhred, 2012). As Sellar (2015, p. 112) points out, despite the promise of the relationship between education, economic growth and social mobility, this promise can become an “opportunity trap” (Following Brown, 2003) for many. This is due to global changes in education, increasing competition within job markets and changes in the modes of production in labor markets themselves (Brown, 2003).

Perhaps more than any other sector, tertiary education reflects the complexities of the human development, economic growth and social mobility relationship. As Peters (2016) highlights, it was around forty years ago that Jean-Francois Lyotard predicted the end of the “modern university” based on Enlightenment principles and a push toward tertiary institutions increasingly serving economic needs. Indeed, the role and function of the university sector has undergone significant changes over the last four decades (Blackmore et al., 2010). Post-war western universities are increasingly corporate and play the roles of multinational businesses in service of a knowledge economy (Peters, 2016).

In addition, private tertiary institutions will soon outnumber public tertiary institutions across the continent of Africa (QS, 2015), and in many countries this has already taken place.¹ Many Asian countries have privatized tertiary education in some way (Varghese, 2015). For example, in Indonesia and Singapore public universities are legal entities able to introduce cost-recovery measures, and Lao PDR and Cambodia have introduced fees. In Vietnam a share of funds to the tertiary sector comes from cost recovery through student fees, though public funding continues as the major source of funding (Varghese, 2015). In Latin America, the private sector enrolls a significant percentage of students, and in the cases of Chile and Brazil, the large majority (Reisberg and Altbach, 2018).

Therefore, recent shifts in the tertiary education landscape point to the promise of human development and human capabilities as increasingly subservient to the promise of economic growth. The promise of social mobility and increases in social cohesion also remains notional. Prior to the pandemic trends in tertiary education systems in LMICs indicate that inequalities are growing, rather than decreasing (McCowan and Schendel, 2016). Ilie and Rose identify crucial inequalities in relation to poverty and gender across a sample of 35 LMICs (Ilie and Rose, 2016). They argue that, “... expansion in access to higher education may predominantly benefit the rich, unless measures are taken to tackle inequalities” (Ilie and Rose, 2016, p. 435).

Challenges to delivering tertiary education during the pandemic

The question of how (and under what conditions) education can build human capacity, support inclusive growth and improve social cohesion is now further complicated by the global COVID-19 pandemic. Early in the pandemic (as of April 8, 2020) tertiary education institutions were closed in 175 countries and over 220 million post-secondary students (13% of the total number of students affected by education closures globally) had their studies significantly disrupted due to the global pandemic (World Bank, 2020b). More recently (as of October 28, 2021) fourteen countries have country-wide school and university closures and approximately 55 million learners are affected by these closures (UNESCO, 2020). Few education systems globally were prepared to respond to such disruption and as such, policymakers, education managers, faculty and students themselves have been ill-equipped to continue learning in the face of physical institutional closures.

In March 2020 the predictions were that “higher education can expect at least 12 months of ‘abnormal conditions’ from the COVID-19 pandemic – with at least five years before global student mobility recovers” (Marginson, 2020). Marginson also predicted South Asia and Sub-Saharan Africa will suffer the most from the impact of COVID-19 as student movement would likely be slower and that their recovery would likely take longer.

Pivoting to online delivery has been the most immediate response to closures (World Bank, 2020b), however the challenges of infrastructure, connectivity, access to devices (for faculty and students) and educators’ limited capacities to shift their pedagogical practice to online learning are significant. These challenges are compounded by issues of equitable access for already marginalized groups, along the lines of gender, rurality, class, disability and caste. In addition to limited access to the internet and devices to engage in online learning, the closure of schools and early learning centers has increased the domestic and reproductive burden on women during the crisis, exacerbating pre-existing inequalities.

The mid to long-term outlook for both individuals and the sector, paints a picture of a sector unlikely to simply recover from these challenges and continue with business as usual. Financing for tertiary education has dramatically decreased in both public and private institutions to such an extent that the landscape for tertiary education is likely to shift permanently, and the sustainability of the sector is at risk in many regions. The early indications are that these challenges are not distributed equally between countries. As of 2021 LMICs are more likely to report a reduction in financing for tertiary education than high-income countries (UNESCO, 2021a).

Not only are governments in LMICs unlikely to be able to provide stimulus packages to support the sector, many national budgets are reallocating education budgets toward the health emergency during the pandemic (Salmi et al., 2020). In addition, reallocation of tertiary budgets to the primary and secondary sub-sectors may take place in many countries. Private institutions that are fully dependent on tuition fees may not survive the pandemic. LMICs are also more likely to have reported greater reductions in student enrollment in 2021, presenting a particular challenge for private universities, or systems with capitation-based funding (UNESCO, 2021a).

¹For example, in South Africa the ratio of private to public tertiary institutions is 4:1 and in Ghana, it is 7:1.

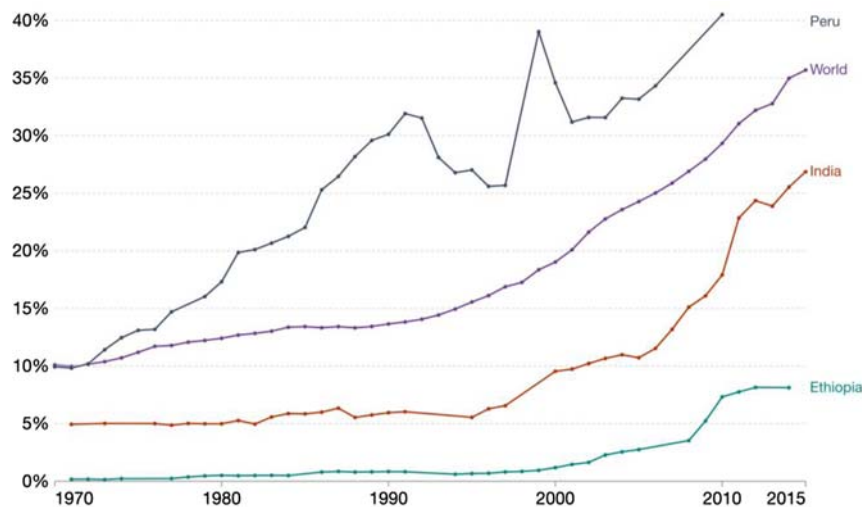


Fig. 1 Tertiary enrollment rates in Ethiopia, India, Peru, and the world (Our World In Data, 2021).

Overall, the pandemic highlights the fragility of the tertiary sector, particularly in LMICs, and the risks to sustainability and inclusion when delivering on the promise of human capital development is pursued at the cost of the other promises of tertiary education.

The cases of Ethiopia, India and Peru

In order to consider the risks, sustainability and inclusion implications of delivering on the promises of tertiary education, we review three country contexts: Peru, India and Ethiopia. Collectively, these countries represent a mix of early tertiary education take up contexts (Ethiopia), recent rapid uptake contexts (India) and a country soon approaching the massification of tertiary education (Peru) (see Fig. 1). The role of tertiary education in the vision for national development also differs in each country.

In addition, each of the three countries have had diverse experiences with the prevalence and likely impacts of COVID-19. Peru is among the top 10 countries with the most COVID-19 related deaths, and by many reports the highest deaths per capita globally (as of December 2021). Education institutions partially re-opened in September 2021, with 34 weeks of full closures and 36 weeks of partial closures since the beginning of the pandemic. India has recorded a high number of infections and COVID deaths. Education institutions have, as in Peru, been fully or partially closed since the beginning of the pandemic. With distance learning being the norm, the divide between students from privileged and less privileged backgrounds is growing as access to technology becomes integral to accessing education during closures. In Ethiopia, COVID cases and deaths (as of December 2021) are significantly lower than many countries. School closures have been shorter than in either Peru or India, with schools having been fully closed for 22 weeks, and partially closed for 9 weeks. However, as in India, school closures are likely to increase the inequalities between the advantaged and disadvantaged students due to unequal access to technology and differing domestic and reproductive burdens.

Ethiopia

In Ethiopia, tertiary education policies have been central to the vision for economic growth, poverty reduction and sustainable development since 1991. The Ethiopia Education Development Roadmap outlines the need to improve gross education enrollment in tertiary education in order to reach lower-middle income status by 2025 (with the target set at 25%). The latest figures indicate gross enrollment (2014) is at 8.1% (UNESCO, 2021b) and has remained approximately the same since 2010. While since the early 1990s, efforts have been made and the representation of women has increased, high attrition rates remain and enrollment ratios of students from less affluent regions stands at just 2.5% and the wealthiest households account for 87% of TVET and 82% of tertiary education enrolments (UNESCO, 2021b).

Ethiopia's private tertiary education sector accounted for around 17% of national tertiary education enrollment. This private sector has experienced pronounced effects of COVID-19, due to the heavy reliance on student tuition and fees. There have been calls to provide greater government support to private institutions through tax exemptions, long-term loans, rent waivers and direct

financial support (Tamrat, 2021). Some observers have concluded that “... unless a substantial intervention is made in time the [private] sector which boasts the largest number of students in Africa will be significantly weakened” (Tamrat, 2021).

Ouedraogo et al. (2021) analysis of intergenerational mobility in education attainment and occupational status indicators in Africa, finds that upward social mobility² in Ethiopia is most clearly observed at the primary education level (at 15% likelihood). However, downwards social mobility at the primary education level is higher at 28% likelihood.³

The likelihood of a child completing tertiary education, when adopted or born to parents who have completed secondary education in Ethiopia is extremely low (0.1%) and the chances of a child not completing tertiary education, when adopted or born to parents who did complete tertiary education is 96%. This indicates that lower mobility is a more likely outcome even among the most educated families (Ouedraogo et al., 2021).

Upwards intergenerational mobility in occupation status in Ethiopia is extremely low, with the likelihood of a child, born or adopted from parents working in the agriculture sector, working in blue- or white-collar occupations, at only 3%. This is even lower for males and those living in rural areas (Ouedraogo et al., 2021). Downward intergenerational mobility measured by occupation status (the likelihood of a child working in agriculture, born or adopted to parents who worked in blue- or white-collar occupations) is more likely than upward mobility (28%) and even higher for males and those in rural settings (36 and 57% respectively) (Ouedraogo et al., 2021).

These intergenerational trends of limited educational and occupational upward mobility and considerably high downward educational and occupational trends speak to the “opportunity trap” (Brown, 2003) whereby global changes in education have increased upwards mobility at the primary school level, yet tertiary and occupational intergenerational mobility are hindered by changes in the modes of production in labor markets themselves.

While intergenerational mobility at the primary school level has taken place in Ethiopia, even before COVID-19 disadvantaged children across Ethiopia and the rest of sub-Saharan Africa were unlikely to complete their primary education (Yorke and Rose, 2020). COVID-19 will likely increase drop out for the poorest children (Yorke and Rose, 2020).

India

In India’s 2020 National Education Policy, tertiary education is framed as being key for individual, societal and economic development (Government of India, 2020). The most recent figures (2017) for gross enrollment at the tertiary level is approximately 27% (UNESCO, 2021b). From an equity perspective, regional inequalities in the distribution of tertiary education facilities and regional market forces have resulted in high regional disparities in tertiary participation rates. Caste remains a strong barrier to entry and Muslims lag behind all other religious groups in terms of enrollment (Tilak and Choudhury, 2018). There are some improvements in participation rates of students from families belonging to the poorest group in the lowest economic quintile (around 4% in 2007, and 9.9% in 2014).

While different studies have taken different approaches to measuring inter-generational education mobility, the findings are consistently mixed. There is evidence that education mobility has improved over time in India (Azam and Bhatt, 2012; Asher et al., 2021; Kundu and Sen, 2021). This general trend however masks other important equity considerations. For example, Asher et al. (2021) found that overall improvements in education mobility for scheduled tribes and castes masked a worsening of mobility for children from Muslim families (Asher et al., 2021). It was also found that increasing levels of parental education closed the attainment gap for girls (Emran et al., 2020). While the impact of family background on educational attainment may be decreasing overall, this is not true for all marginalized groups.

Around two-thirds of India’s tertiary education students are enrolled in private institutions. Combined with significant increases in the annual average household expenditure on tertiary education between 2007 and 2014 (Pushkar, 2021), rising unemployment and poverty rates in India due to COVID-19 are likely to undo the previous gains of access to tertiary education for the poorest.

The impact of COVID-19 on the public sector is both financial and human. For example, the University of Delhi has lost 15 teachers to COVID-19 and Aligarh Muslim University has lost 16 current and 18 retired faculty members to the virus (Pushkar, 2021). With a history of neglect, poor regulation and inadequate financial support, India’s public universities may never recover from the effects of the pandemic (Pushkar, 2021).

Household expenditure on tertiary education is highest for the students enrolled in private -unaided institutions and is lowest for public institutions, with private-unaided household expenditure 3.5 times higher than public institutions (Tilak and Choudhury, 2018). Rising income inequality in India has increased the challenges of access to quality tertiary education for those from the poorest households and the returns to education have increased during this time (Tilak and Choudhury, 2018).

²Defined by Ouedraogo and Syrichas as the likelihood of a child adopted or born to parents who have not completed primary education, completing primary education themselves.

³Defined by Ouedraogo and Syrichas as the likelihood of a child adopted or born to parents have not completed primary education, not completing primary education themselves.

Peru

The National Education Project in Peru supports the “Higher Education Quality Improvement Project”, which has adopted a strategy to increase the quality and relevance of tertiary education in Peru. The Peruvian national development plan’s (Gobierno del Perú, 2019) objectives are captured in the plan’s title “National Plan Competitiveness and Productivity”. Objective two is of the plan is to “strengthen human capital” which includes a sub-objective on the strengthening of access to and quality of tertiary education.

Most recent figures (2017) have gross enrollment in tertiary education at approximately 70%, with higher female enrollment and large differences between urban and rural (80.8 vs. 30.6% respectively) and the richest and poorest students (142 and 31% respectively) (UNESCO, 2021b).

Education mobility in absolute terms has increased dramatically in the last fifty years, driven by an expansion of availability in basic education in the 1970s (Gaentzsch and Román, 2018). This increase in absolute mobility,⁴ however, masks a lack of progress in relative mobility.⁵ The likelihood of a child whose parents have no education achieving tertiary education is 0.06 (Gaentzsch and Román, 2018). In another study it was found that 65% of those with high parental levels of education are likely to achieve the same, with this figure dropping to below 15% for those with low parental levels of income (Neidhöfer et al., 2018). This is among the highest levels of inter-generational educational persistence in Latin America, though it is improving over time, with persistence rising at higher levels of education, and falling for lower levels of education (Neidhöfer et al., 2018). While access to tertiary education has increased, lifetime returns on tertiary education are heavily influenced by the quality of institutions, which is in turn influenced by socio-economic background (Sánchez et al., 2021).

As with many countries approaching the massification of tertiary education, observed gaps in enrollment between the richest and the poorest are often explained in terms of aspirations. Castro et al. (2016) state that the observed gap between the participation of rich and poor households is not only explained by short-term credit constraints, but also by poor cognitive skills and family backgrounds affecting “*tastes, aptitudes and aspirations for formal education*” (2016). Policy discourse regarding aspirations has come to the fore during the tertiary education massification projects of high-income countries. Success is at times understood to be a matter of choice, discounting the structural issues that plague education systems and societies (Bowers-Brown et al., 2019).

Massification of tertiary education in Peru has been heavily supported by an expansion of private sector involvement in education. This was initially driven by a 1996 reform of the education sector, which opened up the sector to private investment, and provided tax incentives for investors (Zeta, 2021). The result of this is that 8 in 10 tertiary education students are currently studying in private institutions (Benites, 2021). This liberalization of private investment in education means that in terms of the proportion of GDP spent on education, household expenditure outweighs government investment, with tertiary education accounting for the largest share at 44% of household education expenditures (Acerenza and Gandelman, 2017). Reflections on the post-covid landscape in Peru have predicted that it will be these private institutions, and particularly low-cost private institutions which will be most adversely affected by the crisis (Benites, 2021).

Looking to the future

Looking at these three case studies we see a diversity of historical trajectories, as well as current social and economic standing. We also see some common characteristics of countries in the process of massifying tertiary education. The tertiary education landscape in all three can be characterized as:

1. Being driven by a policy environment focused on the massification of tertiary education for the purpose of achieving greater economic growth.
2. Having low levels of education mobility, particularly for economically or socially marginalized groups.
3. Being increasingly reliant on the private sector to support the massification of tertiary education.

It is a crucial moment in time for policy makers and decision-makers to understand how education systems (at all levels) can support inclusive growth in tertiary education, without widening inequality. Policies aimed at narrowing gender gaps and widening opportunities for decent work through education during and after COVID will be crucial to delivering on the promises of tertiary education in future years.

Considering the common facets of systems working toward the massification of tertiary education outlined above, we suggest three “traps” that lie in wait for policy makers. They are traps in the sense that if they are not carefully considered they will undermine and weaken the connection between the expansion of tertiary education and its triple promise to improve individual capabilities, economic growth and social cohesion. They are:

- (1) the risks inherent to the rapid expansion of access to tertiary education (the massification trap);
- (2) the risks inherent in not putting equity at the forefront of planning (the opportunity trap); and
- (3) the risks associated with the role of international student mobility (the internationalization trap).

⁴Defined as the change in proportion of people accessing education over generations.

⁵Defined as the determination of education outcomes by parental level of education.

The massification trap

The massification of tertiary education seems to be an inevitable follow on from the massification of basic and secondary education. This is evidenced by the Sustainable Development Goal (SDG) target 4.3, which aims to, "... by 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university" (United Nations, 2015). This shift has begun in LMICs and is visible in the three case studies presented here. As the world of work becomes both more globalized and more technologically demanding, citizens will look to their governments to provide them with the kind of tertiary education that has become a global norm (Rizvi, 2018).

Lessons on massification are evident in both the experiences of massification of tertiary education in high-income countries, as well as the experiences of massification of basic education in LMICs.

Lesson from high-income countries

Massification of tertiary education in high-income countries has generally taken place in the second half of the twentieth century. This massification has led to a rapid expansion of the number of people taking part in various forms of tertiary education. Obvious benefits in terms of individual capabilities, workforce capacity, and social cohesion appear within the massification narrative. However, several trends undermine this narrative.

Firstly, there has been a trend toward *managerialism*. Academic boards in Australia and in Britain in the 1960s were predominantly constituted by professors, whereas recent shifts toward tertiary education as a consumable has led to a growth in the proportion of board members who are senior managers or executives (Blackmore et al., 2010). While this is inevitable in the context of hugely increased student numbers, it presents a challenge to the promises of tertiary education. It has led to an increased managerial burden and scrutiny on lecturers, the increased standardization of teaching practices, and the increased need to align with external demands rather than the developmental needs of students (Fisher, 2006; Carpentier, 2018; Evans et al., 2021). It is argued that this has led to a perceived change in the *character* of the university toward an organ of the economy rather than of the holistic development of young people (Fisher, 2006). This suggests that the process of massification in high-income countries has led to the prioritization of equipping the workforce over the individual capabilities of the student.

The second observable trend leading from the massification of tertiary education in high-income countries is the process of *differentiation* of institutions. A common factor visible in countries which have expanded access to tertiary education is the emergence of a range of different kinds of tertiary education institutions. The creation of an explicitly or implicitly tiered system of tertiary education is impactful in widening participation rapidly, but ultimately entrenches existing inequalities (Carpentier, 2018). For example, in the United Kingdom, where significant emphasis is placed on marketing to potential students, it has become increasingly difficult for students to interpret this differentiation, placing increased emphasis on a need for students to *manage* their own educational journey (Knight, 2019), an ability linked to individual cultural and social capital.

Lessons from the massification of basic education

The implementation of the Millennium Development Goals (MDGs) on education led to a rapid expansion of access to basic education globally with the proportion of out of school children at primary school ages halving between 1996 and 2015 (World Bank, 2020a). However, it has since been recognized that this has not translated to improved learning outcomes with 6 of 10 children and adolescents not meeting minimum proficiency standards in reading and mathematics (UIS, 2018).

Low quality and irrelevant education threaten delivering on the promise of developing human capital toward national economies and developing the capacities of individuals. Poor quality education can fail to develop the capacities of children and young people who spend much of their time in learning institutions yet are not learning. Additionally, too many students who are *successful in school* are not learning the skills and knowledge aligned with current and future employment needs to contribute to national economic growth (Allen et al., 2016). Both the real and perceived benefits of schooling are affected by gender, age and wealth (Rollleston, 2014, p. 135) and the extent to which individuals and schools can convert endowments into learning and learning into academic and workforce outcomes differs by individual background factors (Outhred and Lipcan, 2016).

The challenge of balancing the rapid expansion of access, with a need to deliver quality outcomes is also likely to be seen in the massification of tertiary education. This will particularly be true where the expansion of access is done rapidly, without opportunities for a gradual build-up of institutional capacity which would allow for the delivery of quality education (Rizvi, 2018).

The opportunity trap

It is widely assumed that increased access to quality education leads to greater social mobility. Discussion of social mobility is primarily interested in the extent to which an individual's, or group's opportunities in life are defined by their social origins (Müller and Pollak, 2015). In discussing the promise of tertiary education, this means focusing on the extent to which participation in tertiary education weakens the link between social origin and lifetime opportunity.

There is ample evidence from the Organization of Economic Cooperation and Development (OECD) on the interactions between social mobility and educational outcomes. Unsurprisingly this link is bi-directional and complicated. The level of social

mobility in a country influences equity in access to tertiary education, and participation in tertiary education influences social mobility. The OECD's conclusions (OECD, 2018) on the links between social mobility and education are:

1. Income inequality is the key factor in limiting social mobility
2. Income inequality is also shown to limit access to tertiary education for under-privileged students
3. In unequal societies, both the perceived and actual lifetime returns on tertiary education are lower for under-privileged students.

The evidence concludes that in societies where education outcomes are more determined by family background than they are by student aptitude, the expansion of tertiary education serves only to reinforce inequalities (OECD, 2018). The implications of this are that where societies are unequal, investment in tertiary education, without a clear equity agenda and strategy, will do little for social cohesion. Its impacts on individual capabilities and workforce capacity will skew toward the privileged (Ilie and Rose, 2016; McCowan and Schendel, 2016).

Evidence from high-income countries highlights the importance of how tertiary education is financed. Looking across OECD countries there is a strong correlation between the funding model for tertiary education, and inter-generational educational mobility. Countries with universal access, either through central funding, or a centralized loan system, have a lower correlation between socio-economic background and access to education (OECD, 2010). This suggests that investment in tertiary education does not guarantee equity in opportunity. If tertiary education does not work to increase social mobility it is unlikely to fulfill its promise of increased social cohesion, and its benefits in terms of individual capabilities and workforce capacity will be limited to those with inherent privilege. This equity dimension is particularly important in cases where per-student investments in tertiary education significantly outweigh per-student investments in primary and secondary education. If access to, and benefits from, tertiary education are skewed toward those from privileged backgrounds, these investments are particularly unequitable.

The globalization traps

In our three case study countries, the globalization of tertiary education has taken place. This can be seen globally, with international student mobility doubling between 2007 and 2019 (OECD, 2021). The number of students from low and middle-income countries studying abroad is relatively small, for example the outward mobility ratio (defined as outwardly mobile students as a percentage of total tertiary students in a country) for India is 1.3%, and for Peru is 1.7% (UNESCO Institute of Statistics, 2021). For receiving countries however, the internationalization of tertiary education is very significant. For example, the inward mobility rate in the United Kingdom is 18.7%, and for Australia is 28.4% (UNESCO Institute of Statistics, 2021).

There is an ongoing debate regarding the positive and negative impacts of the globalization of tertiary education. From the inwardly mobile (host country) perspective, the benefits are framed both in terms of economic benefits, and human capital benefits. As an example of the former, the export of tertiary education services was worth GBP 16 billion to the economy of the United Kingdom in 2019 (Department for Education, 2020). This trend has been heavily driven by the marketization of the domestic tertiary education market, bringing financial benefits for attracting more students, particularly in cases where international students are required to pay higher fees than domestic students. Human capital perspectives are driven by an increasing demand for highly educated young people, and the ability of high-ranking universities to attract the best students globally (Knerer et al., 2012). This builds on the promise of tertiary education to support the development of a modern and competitive workforce.

From the outwardly mobile perspective, there is evidence that when students return to their home country, they are able to bring both new perspectives and knowledge, which benefit both the development of their individual capabilities, as well as the capacity of the workforce in that country (Knerer et al., 2012; Chan et al., 2016; Campos and Rocha, 2021). This is tempered by the threat of "brain drain" in LMICs, which can deprive universities in countries of origin of promising students, financial revenue, as well as prestige, as they are implicitly considered as offering an inferior product to a foreign education (Mlambo et al., 2020). The scale of brain drain is less clear. While there is evidence that students who study abroad are more likely to work abroad in the period after studying (Oosterbeek and Webbink, 2011) there is no evidence to account for how many students will permanently relocate. There is also a theoretical position of ideal outward mobility, in which the education maximizing effect of the possibility of studying abroad outweighs the loss of educated population through brain drain (Docquier, 2014). While this makes a strong case for international student mobility being a potential positive influence on the individual capabilities and economic development of countries of origin, the literature does not address completely the social cohesion dimension, and the issue of the perception of value of domestic tertiary education as set out by Mlambo et al. (2020).

Tertiary student flows from LMICs to, often the colonizing nations of Europe, reflect both the historical and present-day role of tertiary education in stratifying humans. Historically there are three main waves of colonial university establishment, from:

- (1) The early "outposts" established by the Spanish in the New World during the sixteenth century;
- (2) The next 300 years of progressive consolidation of universities in the New World; and
- (3) the British established higher learning institutions in African colonies after World War II, as part of the policy of indirect rule (Peters, 2019).

Following independence, universities in formerly colonized nations have faced multiple challenges (many outlined in this entry), including ridding themselves of the epistemological and ontological legacies of colonialism, that perceives them and their people as

the recipients of world knowledge (“the periphery”), rather than creators of it (the center) (following [Peters, 2019](#)). The challenge of decolonizing is certainly one of a multitude of challenges that impact upon both the real and perceived quality of tertiary education available to citizens.

Against this historical backdrop, emerge national elites able to parlay their positions of relative advantage for greater opportunities (perhaps Bauman’s “Tourists” of the human condition in globalization) ([Bauman, 1996](#)). This is undertaken through mobility to tertiary institutions within “the center” countries.

From the perspective of our three promises of tertiary education, the globalization traps are complicated. There are benefits in terms of individual capabilities for studying abroad, both in terms of access to high quality academic education, as well as in the personal and intercultural benefits of studying abroad. In terms of workforce capacity and social cohesion the outcomes are less certain. Finally, the globalization traps risk increasing investments in tertiary institutions within “the center” and decreasing material and human resource to build tertiary education quality in LMICs and decolonize the institutions historically embedded as “periphery” knowledge recipients.

Conclusion

The expansion of access to tertiary education promises support to national development in three ways:

- (1) through the development of individual capability;
- (2) through the promotion of a modern, highly skilled workforce, stimulating economic growth; and
- (3) through the strengthening of national social cohesion, through social mobility.

Based on one or more of these promises many governments in LMICs are pursuing massification of tertiary education, supported by the aims of Sustainable Development Goal 4.

This expansion toward universal access to tertiary education is taking place against a background of a dynamic globalizing tertiary education space. Planning for tertiary education is increasingly being driven by the rapidly changing needs of increasingly technological industries, shifting the kinds of knowledge and skills needed for the workforce. Greater international mobility has placed tertiary education firmly in the realm of a global commodity, with highly ranked universities marketing themselves extensively to prospective students in emerging economies. In these emerging economies, the increased demand for tertiary education is leading to a proliferation of private institutions, and often reduced investments in the public institutions.

All of these trends must now be placed against the context of two years of global disruption caused by the global COVID-19 pandemic. COVID-19 has highlighted the fragility of a sector model built to primarily serve the promise of a modern workforce to promote economic development. Disruption during the pandemic has been both direct and indirect. Direct disruption includes restrictions necessitating a move toward remote instruction, threatening the quality of education for a generation of graduates. Indirect and long-term disruption includes the reduction in international mobility, which has the potential to negatively impact both host and origin countries, and decreased investments in tertiary education, as well as a loss of human capacity. Similarly, private institutions are more likely to suffer from loss of funding through lower student enrollment.

Observing examples from across the three case study countries of Ethiopia, India and Peru commonalities emerge. Expansion of access to tertiary education is prioritized in national policies and a steady move toward the massification of tertiary education is observed. At the same time, we can observe systems with entrenched inequalities in access to tertiary education, and systems in which there is little inter-generational education mobility. Reliance on private institutions means that a large proportion of funding for tertiary education is from household contributions. Therefore, a lack of social mobility becomes both the driving force and consequence of unequal access to tertiary education. Without explicitly addressing inequality within and as a consequence of the system, tertiary education cannot deliver on the promise of improving societal cohesion.

Learning the lessons from previous massification projects, policy makers should not assume that as access to tertiary education expands, the nature of the education changes. The successes of massification at the primary and secondary levels have at times been undermined by a reduction in the quality and relevance of education. Therefore, the ability to translate participation in education to increased individual capacities and meaningful lifelong outcomes (promise one) should not be assumed to be equal for all students. This is also evidenced in the opportunity trap. There is ample evidence from countries which have already massified tertiary education to show that often tertiary education can serve to reinforce social inequalities rather than to reduce them. This is particularly true in countries with pre-existing inequalities, as is true in many low- and middle-income countries. Finally, the globalization traps highlight that alongside the expansion of opportunities for access to tertiary education domestically, there is a trend toward greater international student mobility. The evidence for this is mixed, and there are many reciprocal benefits for both origin and host, should the process be well managed. However, the internationalization of tertiary education is also built on the legacy of colonial approaches to tertiary education, entrenching global hierarchies, and threatening the promise of social cohesion.

The massification of tertiary education is not simple, and the link between expansion of access and delivering on the triple promises of tertiary education is not clear. Policy makers in countries looking toward universal access to tertiary education have an opportunity to learn from countries that have been through the process of massification of education and from their own experiences of expanding basic education. Policy makers should not assume that each of the three promises of tertiary education will be realized as a natural consequence of achieving increased access to tertiary education as the three promises rarely sit comfortably together.

Rather, policy makers should seek to balance and maximize the three promises of tertiary education, especially when navigating a complex and shifting post-COVID global landscape.

References

- Acerenza, S., Gandelman, N., 2017. Household Education Spending in Latin America and the Caribbean: Evidence from Income and Expenditure Surveys. Available at: <http://www.iadb.org>.
- Allen, R., Elks, P., Outhred, R., Varly, P., 2016. Uganda's Assessment System: A Road-Map for Enhancing Assessment in Uganda. Available at: www.heart-resources.org.
- Asher, S., Novosad, P., Rafkin, C., Abadie, A., Autor, D., Blanchard, E., Chetty, R., Edmonds, E., Emran, S., Ferreira, F., Finkelstein, A., Hilger, N., Katz, L., Laibson, D., Ligon, E., Luttmmer, E., Newey, W., Papaioannou, E., Pavcnik, N., et al., 2021. Intergenerational Mobility in India: New Methods and Estimates across Time, Space, and Communities.
- Azam, M., Bhatt, V., 2012. Like Father, like Son? Intergenerational Education Mobility in India.
- Bauman, Z., 1996. Tourists and vagabonds: heroes and victims of postmodernity. In: *Reihe Politikwissenschaft/Political Science Series*, vol. 30. Available at: www.ssoar.info.
- Benites, R., 2021. La Educacion Superior Universitaria en el Peru post pandemia.
- Blackmore, J., Brennan, M., Zipin, L., 2010. Re-Positioning University Governance and Academic Work. Brill, Leiden, The Netherlands. <https://doi.org/10.1163/9789460911743>. Available at:
- Bowers-Brown, T., Ingram, N., Burke, C., 2019. Higher education and aspiration. *Int. Stud. Sociol. Educ.* 28 (3–4), 207–214. <https://doi.org/10.1080/09620214.2019.1641307>. Available at:
- Brandt, Y., 2015. Education: The Most Powerful Investment in Our Future - UNICEF Connect, UNICEF Connect. Available at: <https://blogs.unicef.org/blog/education-the-most-powerful-investment-in-our-future/>. (Accessed 19 January 2022).
- Brown, P., 20. The Opportunity Trap: education and employment in a global economy. *Eur. Educ. Res. J.* 2 (1), 141–179. <https://doi.org/10.2304/eej.2003.2.1.4>.
- Campos, L.N., Rocha, S.W.S., 2021. Student mobility and research capacity: a global health experience. *Int. J. Med. Stud.* <https://doi.org/10.5195/IJMS.2021.992> [Preprint]. Available at:
- Carpentier, V., 2018. Expansion and Differentiation in Higher Education: The Historical Trajectories of the UK, the USA and France. Available at: www.researchcghe.orgwww.researchcghe.org.
- Castro, J.F., Yamada, G., Arias, O., 2016. Higher education decisions in Peru: on the role of financial constraints, skills, and family background. *High Educ.* 72 (4), 457–486. Available at: <http://www.jstor.org/stable/26447555>.
- Chan, S., Sebastian, E., Walker, A., 2016. Outcomes of Learning Abroad Programs.
- Department for Education, 2020. UK Revenue from Education Related Exports and Transnational Education Activity in 2018 UK Revenue from Education Related Exports and Transnational Education (TNE) Activities.
- Docquier, F., 2014. The Brain Drain from Developing Countries. *IZA World of Labor*. <https://doi.org/10.15185/izawol.31> [Preprint]. Available at:
- Emran, M.S., Jiang, H., Shilpi, F., 2020. Gender Bias and Intergenerational Educational Mobility Theory and Evidence From China and India. Available at: <http://www.worldbank.org/prwp>.
- Evans, C., Rees, G., Taylor, C., Fox, S., 2021. A liberal higher education for all? The massification of higher education and its implications for graduates' participation in civil society. *High Educ.* 81 (3), 521–535. <https://doi.org/10.1007/s10734-020-00554-x>. Available at:
- Fisher, W., 2006. The Metamorphosis of Higher Education in the UK-is There an Identity Crisis?
- Gaentzsch, A., Román, G.Z., 2018. More Educated, Less Mobile? Diverging Trends in Income and Educational Mobility in Chile and Peru. GDI Working Paper. University of Manchester, Manchester.
- Gobierno del Perú, 2019. Plan Nacional de Competitividad y Productividad.
- Government of India, 2020. National Education Policy 2020.
- Heyneman, S.P., Kraince, R., Lesko, N., Bastedo, M., 2007. Higher education and social cohesion: a comparative perspective. In: *Higher Education in the New Century*. UNESCO, pp. 55–78.
- Ilie, S., Rose, P., 2016. Is Equal Access to Higher Education in South Asia and Sub-Saharan Africa Achievable by 2030?, vol. 72, pp. 435–455. <https://doi.org/10.1007/s10734-016-0039-3>. Available at:
- Knerr, B., Tlatlik, R., Xi, Z., 2012. The Benefits of International Student Mobility. A Comparison between Chinese and Indian Students in Germany.
- Knight, E., 2019. Massification, marketisation and loss of differentiation in pre-entry marketing materials in UK higher education. *Soc. Sci.* 8 (11). <https://doi.org/10.3390/socsci8110304>. Available at:
- Kundu, A., Sen, K., 2021. Multigenerational Mobility in India. SSRN. <https://doi.org/10.35188/UNU-WIDER/2021/970-9> [Preprint]. Available at:
- Marginson, S., 2016. The worldwide trend to high participation higher education: dynamics of social stratification in inclusive systems. *High. Educ.* 72, 413–434. <https://doi.org/10.1007/s10734-016-0016-x>.
- Marginson, S., 2020. Global HE as We Know It Has Forever Changed, *Times Higher Education*. Available at: <https://www.timeshighereducation.com/blog/global-he-we-know-it-has-forever-changed>. (Accessed 12 January 2022).
- McCowan, T., Schendel, R., 2016. The impact of higher education on development. In: *The Routledge Handbook of International Education and Development*. Routledge, Abingdon.
- Mehrotra, S., 2005. Human capital or human development? Search for a knowledge paradigm for education and development. *Econ. Polit. Wkly.* 40 (4), 300–306. Available at: <http://www.jstor.org/stable/4416104>.
- Mlambo, V.H., Ogunnubi, O., Mlambo, N., 2020. Student mobility, brain drain and the internationalisation of higher education in Southern Africa. *African J. Dev. Stud.* 10 (2), 59–82. <https://doi.org/10.31920/2634-3649/2020/10n2a3>. Available at:
- Müller, W., Pollak, R., 2015. Mobility, social. In: *International Encyclopedia of the Social & Behavioral Sciences*, second ed., pp. 640–646. <https://doi.org/10.1016/B978-0-08-097086-8.32092-X>. Available at:
- Neidhöfer, G., Serrano, J., Gasparini, L., 2018. Educational Inequality and Intergenerational Mobility in Latin America: A New Database. Available at: www.cedlas.econo.unlp.edu.ar.
- Nussbaum, M., 1999. Women and equality: the capabilities approach. *Int. Labour Rev.* 138 (3), 227–245. <https://doi.org/10.1111/j.1564-913X.1999.tb00386.x>.
- OECD, 2010. Economic Policy Reforms Going for Growth A Family Affair: Intergenerational Social Mobility across OECD Countries.
- OECD, 2018. Equity in Education. OECD. <https://doi.org/10.1787/9789264073234-en> (PISA). Available at:
- OECD, 2021. Education at a Glance: Indicator B6. What Is the Profile of Internationally Mobile Students? Available at: https://www.oecd-ilibrary.org/sites/b35a14e5-en/1/3/3/6/index.html?itemId=/content/publication/b35a14e5-en_csp_9689b83a12cab1f95b32a46f4225d1a5&itemGO=oecd&itemContentType=book. (Accessed 2 November 2021).
- Oosterbeek, H., Webbink, D., 2011. Does studying abroad induce a brain drain? *Economica* 78 (310), 347–366. <https://doi.org/10.1111/J.1468-0335.2009.00818.X>. Available at:
- Ouedraogo, R., Syrichas, N., Ricci, A., Wiegand, J., Celine, A., Kangni Kpodar, R., Mina, C., Newiak, M., Woldemichael, M., Spray, J., An, Z., Ludwig, A., Kaas, L., 2021. Intergenerational Social Mobility in Africa since 1920 Authorized for Distribution by Luca.

- Our World In Data, 2021. Gross Enrollment Ratio in Tertiary Education, 1970 to 2015. Available at: https://ourworldindata.org/grapher/gross-enrollment-ratio-in-tertiary-education?tab=chart&country=ETH~IND~PER~OWID_WRL. (Accessed 12 January 2022).
- Outhred, R., 2012. Reconceptualising inclusion in higher education. *Int. J. Incl. Educ.* 16 (9), 881–884. <https://doi.org/10.1080/13603116.2011.629690>. Available at: Outhred, R., Lipcan, A., 2016. *Do Private Schools Produce Superior Language Outcomes*. Oxford.
- Peters, M.A., 2016. *Knowledge Economy, Development and the Future of Higher Education*. Sense Publishers.
- Peters, M.A., 2019. Manifesto for the postcolonial university. *Educ. Philos. Theor.* 51 (2), 142–148. <https://doi.org/10.1080/00131857.2017.1388660>. Available at: Pushkar, 2021. The Pandemic Will Accelerate Higher Education's Privatisation in India, *Times Higher Education*. Available at: <https://www.timeshighereducation.com/opinion/pandemic-will-accelerate-higher-educations-privatisation-india>. (Accessed 31 December 2021).
- QS, 2015. Africa's Higher Education Landscape. Available at: <https://www.qs.com/africas-higher-education-landscape/>. (Accessed 31 December 2021).
- Reisberg, L., Altbach, P.G., 2018. Why Can't We Just Get along? Private and Public Higher Education in Latin America | the World View, the World View: A Blog from the Centre for International Higher Education. Available at: <https://www.insidehighered.com/blogs/world-view/why-can%E2%80%99t-we-just-get-along-private-and-public-higher-education-latin-america>. (Accessed 31 December 2021).
- Rizvi, F., 2018. Realizing the benefits of massification. *International Higher Education* 94, 17–19. <https://doi.org/10.6017/ihe.2018.94.10521>. Available at: Rolleston, C., 2014. Learning profiles and the 'skills gap' in four developing countries: a comparative analysis of schooling and skills development. *Oxf. Rev. Educ.* 40 (1), 132–150. <https://doi.org/10.1080/03054985.2013.873528>. Available at: Salmi, J., Arnhold, N., Bassett, R.M., 2020. The Big Bad Wolf Moves South: How COVID-19 Affects Higher Education Financing in Developing Countries, *World Bank Blog*. Available at: <https://blogs.worldbank.org/education/big-bad-wolf-moves-south-how-covid-19-affects-higher-education-financing-developing>. (Accessed 31 December 2021).
- Sánchez, A., Favara, M., Porter, C., 2021. Stratification of Returns to Higher Education in Peru: The Role of Education Quality and Major Choices. Available at: www.iza.org.
- Sellar, S., 2015. 'Unleashing aspiration': the concept of potential in education policy. *Aust. Educ. Res.* 42 (2), 201–215. <https://doi.org/10.1007/s13384-015-0170-7>. Available at: Sen, A., 2000. *Development as Freedom*.
- Sypnowich, C., 2018. Flourishing children, flourishing adults: families, equality and the neutralism-perfectionism debate. *Crit. Rev. Int. Soc. Polit. Philos.* 21 (3), 314–332. <https://doi.org/10.1080/13698230.2017.1398477>. Available at: Tamrat, W., 2021. Enduring the impacts of COVID-19: experiences of the private higher education sector in Ethiopia. *Stud. High Educ.* 46 (1), 59–74. <https://doi.org/10.1080/03075079.2020.1859690>. Available at: Tilak, J.B., Choudhury, P.K., 2018. Inequality in Access to Higher Education in India between the Poor and the Rich: Evidence from NSSO Data.
- UIS, 2018. *One in Five Children, Adolescents and Youth Is Out of School*. Available at: <http://uis.unesco.org>.
- UNESCO, 2020. *Education: From Disruption to Recovery*. Available at: <https://en.unesco.org/covid19/educationresponse>. (Accessed 31 December 2021).
- UNESCO, 2021a. *COVID-19: Reopening and Reimagining Universities Survey on Higher Education*.
- UNESCO, 2021b. *UIS Statistics*. Available at: <http://data.uis.unesco.org/>. (Accessed 31 December 2021).
- UNESCO Institute of Statistics, 2021. *Global Flow of Tertiary-Level Students*. Available at: <http://uis.unesco.org/en/uis-student-flow>. (Accessed 2 November 2021).
- United Nations, 2015. *Goal 4 | Department of Economic and Social Affairs, Sustainable Development Goals*. Available at: <https://sdgs.un.org/goals/goal4>. (Accessed 20 December 2021).
- Varghese, N. v., 2015. *Reshaping of Higher Education in Asia: The Role of Private Sector*. Available at: <https://www.researchgate.net/publication/292274976>.
- World Bank, 2020a. *Children Out of School (% of Primary School Age) | Data*. Available at: <https://data.worldbank.org/indicator/SE.PRM.UNER.ZS?end=2020&start=1990>. (Accessed 20 December 2021).
- World Bank, 2020b. *The COVID-19 Crisis Response: Supporting Tertiary Education for Continuity, Adaptation, and Innovation*.
- Yorke, L., Rose, P., 2020. *The Effect of Covid-19 on Primary Education in Ethiopia: Perspectives of School Principals and Teachers*.
- Zeta, M.L., 2021. Peru's blueprint for equity in higher education. *J. Publ. Int. Affairs*. Available at: <https://jpi.princeton.edu/peru-blueprint-for-equity-in-higher-education>. (Accessed 30 December 2021).

Changing knowledge & cultures in teacher education: autoethnography for critical reflection

Bethany M. Rice, Towson University, Towson, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	136
Origins of teacher education	136
Reflective practice in teacher education	137
Autoethnography as critical reflection	138
Another view: a critique of autoethnography	138
A case for using autoethnography	139
Autoethnography as critical reflection in teacher preparation	140
Conclusion	141
References	141

Introduction

Trends in education are often focused on methods, such as *Understanding by Design* or ideas, such as the creation of block scheduling in schools, but the people behind these practices hold the key to understanding how knowledge and cultures change. Understanding one's own role in teaching and learning is crucial to promoting sound, equitable practices. This chapter will highlight the origins of teacher education to acknowledge the roots of the profession, and then will highlight the role of reflection and its place in teacher preparation. Finally, autoethnography will be discussed as a method for critical reflection used to undergird a foundation for practice grounded in social justice.

Origins of teacher education

Teaching has existed in various forms from the earliest days of civilization. Knowledge was passed from one generation to the next, for instance through storytelling and folk tales, reflecting traditions and cultures of the community and its people (Amali, 2014). As formalized schooling began to occur in Europe and later in North America, there was no specific training for teachers, but rather those who completed a given level of education or experience were deemed ready to teach. Labaree (2008) notes that the expectation for teacher preparation was that if you took the class, then you could teach the class. In early times, special courses or even apprenticeships were not part of teacher preparation (Labaree, 2008). Instead, teaching was based on the personal experience and knowledge of the person teaching. People were at the center of the teaching process.

As the field evolved, teacher education and its place in the university setting differed across the globe. As early as medieval Europe, teacher education was embedded in a liberal arts curriculum to ensure breadth of knowledge, and the university was seen as the place to award advanced degrees indicating “master” teacher status (Schulman, 1986). Yet in the United States, teacher education in institutions of higher education did not take root until the nineteenth century, and even then it was not at the epicenter of the university (Labaree, 2008). The different forms of teacher education have shifted from general, liberal arts preparation to a more specialized preparation based on the demands of the job, particularly as the professional status of teachers changed. Yet, even with the changes in teacher preparation, people and their experiences remain constant.

According to Labaree (2008), the “common school” model originated in the United States in the 1830s, launching a greater demand for formal teachers. With the increased pressure for teachers also came the need for credentials. Even in the earliest years of formalized teacher education, the issue of standards arose. Yet the idea of performance standards within teacher education broke down into relevance versus rigor in many early programs with the more elite schools opting for rigor over relevance (Labaree, 2008, p. 293). Conversely, critics of teacher education, such as Arthur Levine (2006), label teacher preparation as, “academically weak and professionally irrelevant,” despite the elite labels. Arguments arose that teacher education failed to acknowledge and keep pace with the changing demands of schools. This suggests that teachers were teaching based on what they knew to be true—their own experiences with education, rather than the current educational landscape.

As the field of education continues to evolve, teacher education needs to follow suit in order to meet the demands of the profession. Yet, rather than mirroring the educational trends, often mandated by political appointments, this chapter argues that teacher education needs to be first centered on critical reflection and awareness in order to ensure that pre-service teachers acknowledge their own bias and experience in relation to their practice.

Teachers are products of their own education and experiences, whether positive or negative. Danielson (2009) asserts that pre-service teachers, “developed preconceived ideas of what teaching is through having watched others do it. They may sense what teachers do but have no grasp of why they do it.” Without acknowledgment and deep, critical reflection on these experiences, replication of outdated or ineffective practice will be likely. This is especially true for white, middle to upper class teachers, who teach diverse populations of students. Developing a critical reflective practice is essential for delivering culturally relevant pedagogy and identifying issues of social justice.

Reflective practice in teacher education

Not all reflection is created equal. Dewey (1933) began addressing reflection as a way to face dilemmas and make sound decisions based on the information at hand. In the classroom, this may mean using the information gathered to make instructional decisions. Other researchers supported this (Hargreaves and Fullan, 1992; Lortie, 1975). Yet, as teacher education continues to evolve, reflective practices vary greatly across institutions in type, amount, and regularity. There are important distinctions between reflecting based on an incident or specific practice, and reflection for the purpose of challenging social norms and issues of social justice. Is the purpose of reflection for grade or for lasting change?

Valli (1997) outlines five types of reflection common in teacher education—technical, reflection-in and on-action, deliberative, personalistic, and critical reflection. The first type is considered “technical,” as this kind of reflection pertains to specific instances within the classroom (Minott, 2008, 56). However, this type of reflection can be considered transactional in that the pre-service or practicing teacher is only focused on one specific event or series of actions, for instance a lesson or a unit of study (Zeichner, 1994). This type of reflection has been criticized for allowing the writer to assign blame when lessons do not go as planned. For instance, statements such as, the children were not focused, or it was after lunch and the students were restless. Excuses are provided instead of examining lesson content, accessibility to learning, classroom environment and other factors. Reflection in this case is often done to satisfy a requirement for a class or to earn a grade. This results in surface level thoughts, a recounting of events, and once completed, it is often forgotten with little to no learning involved.

Another type of reflection described by Valli (1997) is reflection-in-action and reflection-on-action. These types of reflection are about the overall teaching practice of the pre-service teacher or practitioner. This can occur before a yearly review with a supervisor or at the end of a teaching term. Minott (2008) states that this type of reflection includes an examination of personal beliefs and values, but without specific means for this to occur, not all educators engage in this practice independently. This closely relates to a third type of reflection—deliberative, as outlined by Valli (1997). In this case, educators or pre-service teachers research and consider competing views, weighing evidence and different viewpoints, then making a decision based on their own situation (Minott, 2008, 56). Yet, while this approach potentially evaluates current research, it does not specifically address a teacher’s own experiences and beliefs in relation to the research and situation at hand. As a teacher approaches and evaluates research, they will already hold preconceived notions based on their own experiences and previous beliefs. There is a real danger of confirmation bias in this approach.

Another common type of reflection seen in teacher preparation programs is referred to as, “personalistic” reflection (Valli, 1997). This draws upon personal reflections on one’s inner voice or thoughts. This is closely related to technical reflection in that it could be about a specific event, but it differs in that it could be much broader in scope. This kind of reflection could be about a teacher’s relationship with a particular group of students or their own views on their role as a classroom teacher in general. While this may be a cathartic exercise for pre-service teachers, it does not facilitate proactive examination of specific beliefs and practices undergirding one’s practice.

Finally, Valli (1997) does acknowledge critical reflection as one of the five types. The goal of this reflective practice is to better understand issues of social justice impacting students and the classroom environment. Yet, without specific instruction and a method for entering into this practice, how do we ensure that pre-service teachers and current practitioners are regularly engaging in this mentally demanding task? Hatton and Smith (1995) criticize Valli’s reflective typology, calling it hierarchical in nature (Minott, 2008, 57). These critics suggest that only some types of reflection, as outlined by Valli, require a more critical perspective—those at the top of the hierarchy. However, it is not the hierarchy that needs to be questioned, but rather the intent and frequency of the various types of reflection and their place within teacher education. Teacher preparation programs need to examine when, how, and why pre-service teachers are engaging in reflection to understand its effectiveness.

Many other researchers (Bright, 1996; Brookfield, 2004; Cranton, 1996; Dewey, 1933, 1938) recognize the importance of critical reflection in teacher education. Shandomo (2010) defines critical reflection as, “... the process by which adults identify the assumptions governing their actions, locate the historical and cultural origins of the assumptions, question the meaning of assumptions, and develop alternative ways of acting” (101). This type of reflection invites a deeper level of examination. It goes beyond looking at specific events such as a lesson, and instead invites an inquiry of one’s own beliefs.

Critical reflection also acknowledges pre-service teachers’ expertise and life experiences. Zeichner (1994) notes, “... reflection also signifies a recognition that the generation of knowledge about good teaching is not exclusive property of colleges, universities, and research and development centers, a recognition that teachers have theories too, that can contribute to a codified knowledge base for teaching” (10). This places ownership on pre-service teachers, as well to draw up their own knowledge base and to consider

how this aligns with what the university is teaching. It begins to build a practice where teachers evaluate knowledge and determine if and how to use it. Materials and programs do not teach children, people do. Acknowledging the expertise, as well as the bias and backgrounds of educators is essential to ensuring sound practices in the classroom.

The five types of reflection defined by Valli (1997) need to be questioned and examined further by educator preparation programs to determine whether and how the different modes of reflection are used. As stated previously, not all reflection is created equal. While there may be a time and a place for different forms of reflection, the foundation needs to be grounded in critical reflection. Without a deeper examination of one's beliefs and values, one's teaching practices are only replications of past practice and experiences.

Critical reflection can take many forms, yet one method offers a way to reflect on both beliefs and experiences, while drawing upon evidence. Autoethnography can be used as a method for critical reflection with pre-service teachers, as a way to examine their own school experiences and their beliefs about learning. The next section of this paper will discuss the merits of autoethnography as a method for critical reflection in teacher education.

Autoethnography as critical reflection

As a method for critical reflection, autoethnography stems from more traditional ethnographic research. In the context of teacher education, this research-based approach has been used to ignite critical reflection from pre-service teachers (Rice, 2019, 2021). The definition used for autoethnographic reflection in these studies stems from the work of Adams et al. (2015). These authors define autoethnography as a specific research method with six related principles (1–2):

1. Uses a researcher's personal experience to describe and critique cultural beliefs, practices, and experiences.
2. Acknowledges and values the researcher's relationships with others.
3. Uses deep and careful self-reflection to name and interrogate the intersections between self and society, particular and general, personal and political.
4. Shows people in the process of understanding and acting on that understanding.
5. Balances rigor while considering emotions, intellectual pursuits, methodological underpinnings, and creative writing.
6. Expresses concern for social justice.

These six principles form a foundation for autoethnographic writing. As a research method and a form of critical reflection, autoethnography is not the simple act of generating a personal narrative, such as with the 'transactional' approach to reflection described in the previous section. Rather, the principles outlined by Adams, Holman Jones, and Ellis acknowledge the experiences, practices, and values of the writer, while interrogating the roots of these things for the purposes of social justice. Similarly, Cunningham and Jones (2005) identify autoethnography as an objective stance on the author's personal beliefs, and a way to interpret specific actions, thoughts, and behaviors specific to teaching. It is an opportunity for pre-service teachers to identify specific personal biases that may influence instruction in the classroom.

In order for pre-service teachers to be effective in their profession, they need to understand the complex diversity present in their classrooms, and how as teachers they personally identify with these ideologies (Starr, 2010). Freire (1973) describes the concept of "conscientization" or becoming aware of one's own self and position in society, and how it creates a space for change within one's reality. Starr (2010) describes it more simply as a "study of space between self and culture to engage the individual in experiences that cultivate an authentic cycle of "action-based reflection" (1–2). Increasingly, teachers do not always represent the cultures and diversity of the students in their classroom, as many teachers are white females. It is necessary for educators to reflect on their own experiences and beliefs in order to acknowledge their potential bias, positionality, and pedagogical decisions.

Several fields, including sociology, education, and Disability Studies, already employ autoethnography as a tool for research (Pitard, 2019; Starr, 2010). Yet, the majority of the research in education has failed to consider the potentially significant role that autoethnography can play in developing reflective practice. Autoethnography is a way for pre-service teachers to recognize and process their beliefs about teaching, through examining their experience as learners. The future teachers unearth their predispositions about teaching and learning, so as to better their classroom instruction. For instance, a pre-service teacher may choose to recount their own experiences with punishment in school (seeing punishment given or receiving it). This could lead to an interesting discovery about the individual's current beliefs about classroom management and student behavior. As an example, if the pre-service teacher experienced punishment frequently as a student and saw the process as unfair, then this could impact how he/she approaches classroom management, specifically punishment in their current classroom. Autoethnography becomes the tool used to unearth who we are and what we believe in order to determine how we arrived in this place.

Another view: a critique of autoethnography

There are scholars who critique autoethnography as personal, self-indulgent writing (Starr, 2010). As a research method, autoethnography is also criticized for lacking rigor and having limited bearing on scholarly writing (Anderson, 2006). However, Starr

(2010) counters this with an important aspect of autoethnography, which is often misunderstood—as a method, it offers the capacity for dialog leading to social change, increased consciousness, and informed participants. In addition, when using autoethnography as critical reflection with pre-service teachers, the emphasis is less on scholarly writing, and more about unearthing long held beliefs and practices.

Traditional research relies on objectivity to remove one's self and bias from the process, and from potentially skewing results. Yet in autoethnography, these lines are blurred by the nature of the method (Cunningham and Jones, 2005). The focus of autoethnography is on one's self, personal experiences, and views. Throughout the process of generating an autoethnography, the participant steps back to interpret these attributes of the work as part of the reflective practice. While complete objectivity is not possible, as the writer is central to the genre, relative objectivity can be achieved with analysis and reflection about the process. The potential benefits of autoethnography as a method for critical reflection, and the activation of social justice, outweigh the concerns expressed by some scholars.

Some researchers in social sciences consider autoethnography contentious (Allen-Collinson and Hockey, 2008). Autoethnography is seen as particular only to one's self, identifying it as overly personal. Critics identify the blurred line between researcher and participant as the overarching issue with the method itself. However, proponents of the method would argue that without an examination of one's self and beliefs, then objectivity in any setting would be void (Allen-Collinson and Hockey, 2008).

In teacher preparation programs, it is important to introduce the purpose for and the use of autoethnography as critical reflection, rather than as research, in this case. Some claim that autoethnography is invalid as a research method because it fails to address the three criteria of more traditional approaches—validity, reliability and generalizability (Allen-Collinson and Hockey, 2008). These criteria, while important for quantitative studies, can still be accomplished to some extent through autoethnography, although this is not the focus of the work. Arguably, the most important use of autoethnography by educators is not to replicate and generalize beliefs, but rather to challenge them for the purpose of enacting social justice.

While pre-service teachers in university programs often engage in traditional research methods, such as action research and thesis projects, it is equally as important for them to understand the *why* behind their work. Without a clear understanding of how their beliefs about teaching and learning were formed, and how these beliefs and values impact their work with students, then candidates engaging in the action research process risk missing opportunities for deep, impactful learning. Autoethnography provides a vehicle for pre-service teachers to accomplish critical reflection about practices and beliefs, while also identifying specific opportunities to enact social change within the classroom. Specifically, inclusive education requires educators to recognize potential barriers and work to remove them before instruction is negatively impacted. Without a reflective practice in place, educators will not have the tools to do this work successfully.

A case for using autoethnography

Many teacher preparation programs require their pre-service teachers to reflect after lessons, especially observed lessons. The intent is for the pre-service teachers to identify what went well and what needs to be changed for future instruction. This practice is considered reflection-on-action, as described in the previous sections. The following scenario is fictional, but based on years of reading reflective journals from pre-service and practicing teachers. An entry like this was a common occurrence in one teacher preparation program, before implementing autoethnography, as a practice for critical reflection.

Today's lesson was supposed to go much differently than it did. I knew it was going to be a struggle when my two students, already receiving special education services, immediately began acting out. I had arranged the desks to ensure they would be apart, but this did not matter. Instead, they sabotaged the lesson completely. One crumbled up his worksheet and threw it across the room to get his friend's attention. The other student was only trying to entertain his peers, rather than engage in the lesson. I don't know what to do. All of the other children were ready to go and listened to the directions. I have already sent the two students to the office, and every time they return to the classroom almost immediately. Why do the administrators keep sending them back without consequence? How are the other children supposed to learn? I did my best work to prepare this lesson.

As this sample narrative demonstrates, there are tones of frustration, assigning blame, and ultimately no recognition of how the pre-service teacher's beliefs and values were driving the lesson. This narrative also suggests that the learning of some students may be more important than others, and that administration needs to address problem behavior outside of the classroom. On a micro-level, the pre-service teacher could examine the structures of the lesson. How was the instruction being provided? Were there different opportunities for all students to engage with the materials? Are the classroom participation rules clear and being followed regularly? On a macro level, this pre-service teacher has the opportunity to think deeply about where these beliefs originated from. How are the beliefs about teaching and learning being perpetuated or challenged? This pre-service teacher would benefit from engaging in critical reflection through autoethnography.

While autoethnography is a useful tool for critical reflection, it should not be the only method; otherwise it risks becoming a trend and not a foundational practice. Positioning autoethnography purposefully into the curriculum as part of teacher preparation is necessary. If autoethnography is firmly in place, then other methods of reflection can be successfully used post-lesson and at other specified times. In these cases, pre-service teachers could then draw upon their autoethnographies in their lesson-specific

reflections to identify where and how their own beliefs impacted their decision-making. Opportunities for change emerge when beliefs are challenged.

The next section of this chapter will provide an overview of how one teacher preparation program implemented autoethnography into the curriculum to support critical reflection among pre-service teachers enrolled in a graduate program. This is one example of how autoethnography can be used in teacher education.

Autoethnography as critical reflection in teacher preparation

The graduate program currently serves pre-service teachers in over twelve areas for teacher licensure in the Northeast United States of America. Some have had previous experience in education, while others are new to the profession. However, all pre-service teachers in the programs hold a bachelor's degree. The graduate students are generally white, middle to upper class students, yet the places these pre-service teachers work are increasingly diverse—economically, culturally, and accessibility.

As the graduate education program reviewed pre-service teacher data, including multiple reflections, exit surveys, and teaching evaluations from the field, it became evident that candidates needed additional opportunities for critical reflection before going on practicum. A review of the data indicated that the pre-service teachers needed a way to critically analyze and reflect on their beliefs about teaching and learning. The majority of reflections and materials collected focused on student deficits or negative responses to instruction, rather than the pre-service teachers' role in the classroom as the instructional provider. The submissions analyzed lacked an aspect of self-awareness, one of the cornerstones of critical reflection. This section will outline the lessons learned from implementing autoethnography into the teacher preparation curriculum.

Autoethnography offers an opportunity to examine beliefs and practices; therefore, integrating it into the early classes in the program was essential. In addition, given the number of programs offered, there was not one course required by all degrees. However, two courses were identified that would ensure all pre-service teachers engaged in the practice of autoethnography. One course focused on writing instruction and the other course on inclusive education. Both courses invite participants to engage in deep, critical reflection about their beliefs about teaching and learning. This is facilitated using focused weekly writing prompts. Each week course participants delve into an aspect of their own experience. This can be challenging for the pre-service teachers. In many cases, their previous experiences are different from the concepts being taught now in their teacher preparation program. For instance, one of the writing prompts asks participants to think about how they learned how to write. Did they enjoy writing as a child? Why or why not. This prompt is done early in the process, as a warm up to the more provocative ones that come later. However, even the basic questions about learning to write offer opportunities to reflect and acknowledge their previous experiences.

An important component of this process includes opportunities to share and discuss ideas that arise from the journal prompts. Write with your students. Engage in the process with them. This will facilitate a deeper conversation, but it does require some vulnerability and disclosure about ones self. Establishing a safe space to share experiences and ideas will help course participants, as well as the instructor to share more openly and honestly. In one of my courses, I shared that my earliest writing experiences were marked with shame and embarrassment because my teacher did not like how I held my pencil. I was given a special grip and made to sit in a special seat so that she could ensure I used it properly. This made me hesitant to write. I allowed my students to ask me probing questions about my experience. This generated a discussion about student accommodations, creating a safe learning space, as well as writing instruction. The discussion gave the pre-service teachers a reason to reflect on my experience, but also connect it to their own. Without the use of autoethnography, this type of discussion would be difficult to generate authentically.

Model texts are also used as a way to discuss and critique autoethnography as a method for reflection. Books such as Phil Smith's (2012), *Across the Table*, offer autoethnographic essays focused on various issues connected to Disability Studies. The essays are great examples of how autoethnography can raise awareness of issues of social justice, generate action, and promote deep reflection. Course participants are asked to read and discuss at least two autoethnographies from the text. The chapter selections are suggested on the syllabus. With the necessary permissions, programs might also consider using exemplars from past students. Interacting with texts in this way allows the pre-service teachers to not only experience the power of the method, but also understand the expectations for the final assignment.

After engaging in weekly writing prompts and engagement with model texts, pre-service teachers are assigned a final autoethnography paper. The final paper must demonstrate the pre-service teachers' ability to critically reflect on a belief or practice. The topic for the final paper can stem from one of the earlier journal entries, but must go into more depth to illustrate a deeper level of mastery. The assignment requirement states five pages as the expectation, but the papers submitted are almost always much longer. By this point, most pre-service teachers have embraced the process and feel comfortable sharing at a deeper level. Trust has been established, which is crucial to the process.

The implementation of autoethnography into the teacher preparation program led to other changes in the curriculum. In an early study, Rice and Threlkeld (2017) determined that pre-service teachers needed specific opportunities to engage with issues of social justice. A deeper analysis illustrated that the pre-service teachers could identify issues of injustice, but did not have the tools to enact change. The autoethnography assignments were an avenue for unearthing the issues, but the pre-service

teachers were unsure of their next steps. This led to the implementation of an action research project during their final semester. Pre-service teachers are in the field, completing their practicum experience, as well as enrolled in a reflective seminar. The seminar course is online and spans a full semester. During this time, pre-service teachers are encouraged to revisit their autoethnography, as well as consider their classroom placement. They are guided to design and implement an action research project, often stemming from an issue connected to their autoethnography. The pre-service teachers engage in specific change based on a previous belief or practice. This process has helped to close the gap between theory and practice; however, more research is needed to determine whether the autoethnography and the action research projects are authentically linked to permanent change.

Conclusion

The origins of teacher preparation relied on individual experiences and knowledge as credentials. Early teachers taught what they knew. Faced with new practices that may differ from their experience as students, teachers require opportunities to critically reflect on these differences. The types of reflection used most often, fall short of the mechanisms needed for critical reflection. Autoethnography introduces a method for reflection that offers an opportunity to think deeply and critically about beliefs and practices related to teaching and learning.

When implemented in one teacher preparation program, autoethnography provided a vehicle for analysis, critical thinking, and reflection on experiences and practices. Done early in the program, this allowed pre-service teachers to continue the process of critical reflection throughout their remaining courses and field experiences. Other types of reflection are used, for example after an observed lesson during practicum, but these opportunities also invite pre-service teachers to revisit their autoethnographies to help draw conclusions and plan next steps. The autoethnographies become a reminder to pre-service teachers that their beliefs about teaching and learning are always present.

As research continues and the field of education grows, the need for critical reflection will remain constant. Dewey (1933) brought reflection to the forefront as a way to heighten awareness, but as the culture of education evolves, a more robust tool for personal examination is needed to ensure that pre-service and practicing teachers will integrate the old practices with the new both purposefully and thoughtfully. Autoethnography can be used to ensure that teachers consider their own experiences and beliefs and the impact this has on practice, while meeting the evolving needs of their students; a necessity Freire (1973) articulates for ensuring that all are represented authentically. Autoethnography is a vehicle for addressing and honoring the changing knowledge and diverse cultures within education.

References

- Adams, T., Holman Jones, S., Ellis, C., 2015. *Autoethnography*. Oxford University Press, Oxford.
- Allen-Collinson, J., Hockey, J., 2008. Autoethnography as "valid" methodology? A study of disrupted identity narratives. *Int. J. Interdiscip. Soc. Sci.* 3 (6), 209–217.
- Amali, H., 2014. The function of folktales as a process of educating children in the 21st century: a case study of Idoma folktales. In: *International Conference on 21st Century Education*. Dubai Knowledge Village, vol. 2.
- Anderson, L., 2006. Analytic Autoethnography. *J. Contemp. Ethnogr.* 35 (4), 373–395. <https://doi.org/10.1177/0891241605280449>.
- Bright, B., 1996. Reflecting on 'Reflective Practice'. *Stud. Educ. Adults* 28 (2), 162–184. <https://doi.org/10.1080/02660830.1996.11730638>.
- Brookfield, S., 2004. The getting of wisdom: what critically reflective teaching is and why it's important. *Becoming a critically reflective teacher*. Jossey-Bass, San Francisco.
- Cranton, P., 1996. Professional development as transformative learning: new perspectives for teachers of adults. Jossey Bass, San Francisco. In press.
- Cunningham, S., Jones, M., 2005. Autoethnography: a tool for practice and education. In: *Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Making CHI Natural (CHINZ '05)*. ACM, New York, NY, USA, pp. 1–8.
- Danielson, L., 2009. Fostering reflection. *Educ. Leader* 66 (5).
- Dewey, J., 1933. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. D.C. Heath, New York.
- Dewey, J., 1938. *Experience and education*. Collier Books, Macmillan, New York.
- Freire, P., 1973. *Education: The Practice of Freedom*. Writers' & Readers' Publishing Co-operative, London.
- Hargreaves, A., Fullan, M.G., 1992. *Understanding Teacher Development*. Teachers College Press, New York.
- Hatton, N., Smith, D., 1995. Facilitating reflection: issues and research. *Forum Educ.* 50 (1).
- Labaree, D.F., 2008. An uneasy relationship: the history of teacher education in the university. In: Cochran-Smith, M., Feiman Nemser, S., McIntyre, D.J. (Eds.), *Handbook of Research on Teacher Education: Enduring Issues in Changing Contexts*, third ed. Association of Teacher Educators, Washington, DC, pp. 290–306.
- Levine, A., 2006. Educating school teachers. *The Education Schools Project*.
- Lortie, D.C., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago.
- Minott, M.A., 2008. Valli's typology of reflection and the analysis of pre-service teachers' reflective journals. *Aust. J. Teacher Educ.* 33 (5), 55–65.
- Pitard, J., 2019. Autoethnography as a phenomenological tool: connecting the personal to the cultural. In: Liampittong, P. (Ed.), *Handbook of Research Methods in Health Social Sciences*. Springer, Singapore. https://doi.org/10.1007/978-981-10-5251-4_48.
- Rice, B., Threlkeld, A., 2017. Autoethnography as Critical Reflective Practice: Themes in Teacher Candidate's Stories. Paper Presented at the Inclusive Education Summit, Adelaide, Australia.
- Rice, B., 2019. Finding their voice: action research and autoethnography in inclusive teacher preparation. In: *Global Perspectives on Inclusive Teacher Education*. IGI Global, New York, NY.
- Rice, B., 2021. Opportunities for inclusive practice: the stories our students tell. In: Thomas, M., Hang, L., Walker, P. (Eds.), *Inclusive Education is a Right, Right?* Brill, Boston, MA.
- Schulman, L., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15 (1), 4–14.
- Shandomo, H.M., 2010. The Role of Critical Reflection in Teacher Education, vol. 4. *School-University Partnerships*, pp. 100–113.
- Smith, P., 2012. *Both Sides of the Table: Autoethnographies of Educators Teaching and Learning With/in [Dis]Ability*. Peter Lang, New York.

Starr, L., 2010. The use of autoethnography in educational research: locating who we are in what we do. *Canadian J. New Sch. Educ.* 3 (1), 1–9.

Valli, L., 1997. Listening to other voices: a description of teacher reflection in the United States. *Peabody J. Educ.* 72 (1), 67–88.

Zeichner, K., 1994. Research on teacher thinking and different views of reflective practice in teaching and teacher education. In: Handal, P.G., Vaage, S. (Eds.), *Teachers' Minds and Actions: Research on Teachers' Thinking and Practice*. Routledge, London, pp. 11–28. **This page intentionally left blank**

The role of students' voices in promoting inclusive education

Kyriaki Messiou, University of Southampton, Southampton, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	143
Voice/voices in research	143
Students' voices	144
Conditions for voice	145
Inclusive Inquiry	146
Getting started	146
Co-designing lessons	147
Dialogs that lead to changes	148
Thinking about voices and dialogs through this process	149
Final thoughts	149
References	149

Introduction

The concept of voices in research has been discussed extensively in the literature (Jackson and Mazzei, 2019; Thomson, 2011). Under the umbrella term “voices,” the concept of “student voice” is also situated, which is the focus of this chapter. The movement of student voice has gained prominence over the last two decades, especially after the UN Convention on the Rights of the Child in 1989 (United Nations, 1989) which has been ratified almost by every country in the world. Despite the reported benefits of student voice approaches there is also a range of challenges that have been reported in research that involves children and young people, such as tokenistic ways of engaging with the views of students and issues of unequal power between adults and children and young people.

Students' voices and inclusive education are interconnected ideas (Messiou, 2012, 2019). Following Ainscow's definition of inclusion as a process that focuses on the presence, participation and achievement of all learners (Ainscow, 2007), I argue that a crucial avenue for enabling participation for all learners, is listening to students' voices. More importantly, moving into dialogs with learners can be a further way for enhancing participation and developing inclusive practices. As Kozleski (2020) argues inclusive education “honors divergent perspectives, capacities, and approaches to learning” (p. 340). The approach described in this chapter, *Inclusive Inquiry*, directly relates to student voice approaches, focusing especially on dialogs between children and teachers, in order to develop more inclusive practices in schools. Such dialogs, it is argued, allow divergent perspectives to be explored, and ultimately lead to changes in practice.

Voice/voices in research

The notion of voice has been problematized in the literature. For example, what constitutes voice has been discussed extensively, drawing attention to moving away from verbal means of communication and highlighting non-verbal means of communication (Thomson, 2008). Mazzei (2009) argues that it is impossible to have voice fully captured in research, while St Pierre (2009) points out that participants' voices in qualitative research may have been burdened with too much weight. She goes on to remind researchers that voice is just one source among many others that qualitative researchers should use in trying to make sense of complex phenomena.

On the other hand, Reay (2006) draws attention to the dangers of the collectiveness of voice, while others have raised concerns about the use of plural. In particular, Mazzei and Jackson (2009) argue that:

This practice of “more is better” has indeed highlighted the ways in which voices are not singular, yet the obsession for more full voices side-steps what we view as a more salient feature of the problem: these practices remain attached to notions of voice inherited from metaphysics – voice as present, stable, authentic, and self-reflective. Voice is still “there” to search for, retrieve, and liberate” (pp. 1–2)

This is a strong reminder that highlighting multiplicity of views is still not satisfactory since such a position assumes voice as a state, whereas voices are constantly changing and emerging and affected through interactions with others. Such views can also be linked to the principles outlined by Blumer (1969) in explaining the term that he coined, that of symbolic interactionism:

- (1) That people act toward things, including each other, on the basis of the meanings they have for them;
- (2) That these meanings are derived through social interaction with others; and

- (3) That these meanings are managed and transformed through an interpretive process that people use to make sense of and handle the objects that constitute their social worlds.

Blumer (1969: 2).

Similarly, it could be argued that voices are constructed as a result of social interactions and relate to meanings that people hold toward others. Through these social interactions conversations and dialogs occur. As Lodge (2005) argues, dialog means to arrive at a point that you would not get to alone and involves shared meanings. Whereas, Wegerif (2019) referring to the Bakhtinian term "polyphonic truth," argues that it is "a truth that emerges from the interaction between a range of voices in a dialogic field and is not so much a final accomplishment as a direction of travel" (p. 13). Finally, as Thomson (2011) argues, "The strongly democratic use of "voice" equates to a call for a public sphere in which there is dialog, reciprocity, recognition and respect (p. 21)." Therefore, the issue of dialog closely relates to notions of voice.

The approach introduced in this chapter, *Inclusive Inquiry*, and a specific example from two schools that used the approach, will explore some of the issues raised here about voice. In other words, the idea of voice as a direction of travel, as well as being fluid and socially constructed within a specific context will be further explored. In addition, through the presentation of this example, an exploration of how different voices can lead to the development of more inclusive contexts with identified changes in practices, will be made.

Students' voices

The notion of student voice, similar to definitions of voice described above, ranges from simple expression of views either through verbal or non-verbal means, while at the same time refers to the idea of active participation and having an active role in decision-making processes (Cook-Sather, 2006). Others have drawn attention to such an active role leading to changes within given contexts. For example, Rudduck (2006) distinguishes between consultation and participation, with the first one referring to talking with students about things that matter in school whereas the second one refers to active involvement in the development of a school's work. Taking such an approach links between notions of voice and participation can be seen.

Back in 1989, the UN Convention on the Rights of the Child (United Nations, 1989), drew attention to children's and young people's rights to be heard. Specifically, in Article 12 it is argued that:

States Parties shall assure to the child who is capable of forming his or her own views the right to express those views freely in all matters affecting the child, the view of the child being given due weight in accordance with age and maturity of the child (p. 5).

The Convention has been ratified by almost every country in the world. Though the Convention was an important step forward, it still entails some controversy around the language used. For example, age and maturity are presented to be significant in determining the importance of a child's view, therefore undermining some children's and young people's ability to have a valid view. Such a position is not in accordance with conceptions of childhood that consider children to be active "beings" (Prout and James, 1990).

Much of the research that focuses on students' voices, starts from notions of childhood that view children as active agents and children's right to be heard. However, the issue of rights leading to opportunities to be heard is only the starting point in voice research. What is of most importance, is how those views are used in order to bring about change. In other words, actions should be taken following an engagement with students' views, in order to ensure that the process is worthwhile. Otherwise, such efforts may have the opposite effects of disengaging students (Messiou and Hope, 2015).

Research in various countries has focused on involving students as researchers with a focus on exploring various aspects of school life (Ainscow and Kaplan, 2005; Carrington et al., 2010; Fielding, 2001; Tangen, 2009; Thomson and Gunter, 2006). Such approaches have various impacts on students, such as "deep participation" (Wilson, 2000) and growth of agency, belonging and competence among students (Mitra, 2003, 2004; Mitra and Serriere, 2012). In addition, other studies draw attention to relationships between students and children becoming stronger (Fielding and Bragg, 2003; Fielding, 2004; Hope, 2012; Baroutsis et al., 2016; Graham et al., 2018).

While there are many benefits as reported above, at the same time research that focuses on students' voices entails certain challenges. The issue of unequal power between adults and students has been highlighted by a number of authors (Mayall, 2001; Mitra, 2008). Though this is inevitable to some extent, due to the age difference, a range of ways has been discussed in order to minimize the effects of this power difference (Morrow and Richards, 1996), not least is the use of methods that allow students' voices to emerge and come to the surface. Essentially, as Fielding (2001) argues such work involves a change in what it means to be a student and, indeed, in what it means to be a teacher. In other words, both parties should be prepared to take different roles, respecting one another when engaged in such kind of work. The issue of teachers resisting to students' voices, or even rejecting their opinions as valid has been reported in some studies previously (Kaplan, 2008; McIntyre et al., 2005). These may stem from resistance to believe that students do have valid views, or from being reluctant to take on different roles, as Fielding suggests.

Finally, another area that has been discussed in relation to voice and research is the issue of silence. Lewis (2010) has pointed out the importance of silence in research that involves students, and suggests five methodological aspects to be taken into account:

1. Recognizing,
2. Noting,
3. Responding to,
4. Interpreting, and
5. Reporting silence from children.

Lewis's (2010) work reminds us how silence is complex and a result of co-construction between researchers or other adults and children, linking again to the ideas of symbolic interactionism. In the same way that people act toward others based on the meanings they have constructed for them, as Blumer (1969) suggests, it could be argued that silence is voice, or action, that stems from particular meanings ascribed to individuals or situations, which have emerged through interactions with others.

In other words, when a child remains silent within a research context, either while being observed in their natural setting, or while being interviewed, this silence could be seen as a form of voice, which derived out of interactions with others within a given context. Those others could be the researchers who led to these silences through their own questioning or actions, other children in a classroom setting, or the teachers who may have led to these silences through their own reactions or words. As Lewis (2010) puts it, "Listening better includes hearing silence and that silence is not neutral or empty (p. 20)." Silence can therefore be seen as voice, linked to particular situations and interactions and should be recognized and explored in depth.

Conditions for voice

Focusing within the classroom context, Lefstein et al. (2020) highlight four conditions of "voice":

1. Having the opportunity to speak;
2. Expressing one's own ideas;
3. Speaking on one's own terms; and
4. Being heeded.

Alexander (2019: 667) suggests that these conditions "prompt four questions:

1. What do we do to encourage our students to speak?
2. How do we ensure that what they say is treated equitably and respectfully?
3. When students speak, whose voices do we hear?
4. And how do we handle contributions that digress from our agenda?"

Although these issues have been raised in relation to voice in lessons, they relate to the ones posed by Fielding (2001) about student voice in research and in schools. Fielding proposed a list of questions, organized in nine clusters:

- Speaking,
- Listening,
- Skills,
- Attitudes and dispositions,
- Systems,
- Organizational cultures,
- Spaces,
- Action, and
- The future.

These, he observes, can point toward areas that need further development in efforts toward promoting democratic dialogs in schools. These clusters and the related questions can guide efforts to understand "more clearly and more acutely what conditions are most likely to foster student voice as a transformative force rather than as an unwitting adjunct to the increasingly irrelevant and pernicious paradigm of school effectiveness" (p. 133). As he argues, these questions "apply equally well to any set of arrangements which seek to move more closely and more authentically to a practice of dialogic democracy" (p. 133). The approach described in this chapter, *Inclusive Inquiry*, had as its central feature the idea of dialog, with a focus on creating more inclusive environments and ultimately arrive at what Fielding describes as the practices of dialogic democracy. In addition, Mitra (2008), using the concept of "community of practice" as a lens, suggests a number of strategies in order to strengthen student voice initiatives. She is arguing for the need to build "meaningful roles based upon mutual responsibility and respect among all members; developing shared language and norms, and developing joint enterprises aimed at fostering voices that have previously been silenced from decision making and knowledge-building processes" (p. 221).

Moving away from notions of voice as stable and taking the position that voices are created through a set of complex interactions, influenced by the way that others contribute in dialogs as they occur in particular contexts, in what follows I set out to explore how voices were constructed as part of an approach that was used in various schools with the aims of enabling all children to

participate in lessons and develop more inclusive practices. The approach has emerged through collaborative work with colleagues in five countries where we attempted to ensure that all students' voices are heard. The dialogs, take place mostly outside the classroom context, when children and teachers get together in order to design collaboratively a lesson that aims to reach out to all children.

Inclusive Inquiry

Inclusive Inquiry is an approach that has emerged through the work of two inter-connected studies which were funded by the European Union. The first study involved secondary schools in three countries (England, Spain and Portugal) (Messiou et al., 2016) and the second involved primary schools in five European countries (Austria, Denmark, England, Portugal, Spain). Both studies employed collaborative action research as a methodology, where different stakeholders took the role of co-researchers (Mitchell et al., 2009).

Inclusive Inquiry involves teachers and children working together to co-design lessons that are inclusive (Messiou and Ainscow, 2020). The approach involves three phases: Planning, Teaching and Analyzing. Each of the phases consists of 12 specific steps. These comprise a "Levels of Use Framework" that can be used by teams of students and teachers who are working collaboratively to help them with discussions about the extent to which the approach has been followed consistently. A distinctive feature of the approach is the involvement of some students as researchers collecting and analyzing their classmates' views about learning and teaching, and observing and refining the lessons in collaboration with their teachers.

In practice, schools that use the approach have to form groups of children and teachers who will collaborate for the implementation of the approach. Usually, these groups consist of three teachers, ideally teaching the same age group, and a total of nine student researchers, three from each class. Teachers are asked to choose children from their classes that are seen as "hard to reach" who will become researchers, such as disabled students, those with migrant backgrounds or others that lack confidence or are not focused enough in lessons. These students are then trained by their teachers on using various methods to collect their classmates' views, as well as on how to conduct lesson observations. They then go on to collect their classmates' views about learning and teaching and analyze their ideas. These ideas are then used to inform the design of a lesson that they plan collaboratively with their teachers. Once the lesson is designed, it is then taught by the first teacher while being observed by the other two teachers and the student researchers who are not from that class. The student researchers have been trained to carry out observation focusing on the activities of the lesson and whether they allow all children to participate or not. At the end of each lesson the student researchers from the other classes who were observing collect students' views about the taught lesson. Their views are taken into account and brought to the discussion that follows the lesson between teachers and student researchers. The discussions focus on identifying ways in which the lesson activities allowed all children to participate as well as suggest possible changes to allow for greater participation. The changes to the lesson are made before the process is repeated three times. After the three lessons have been taught, student researchers and teachers discuss the wider implications for practice that are identified, and these are implemented in various future lessons with a view to enabling better participation of students in all lessons.

Over the three years of the project, data were collected and analyzed in the thirty schools in the five countries. The analysis of the data has highlighted some significant impacts on teachers' thinking and practices and, more importantly, on the student engagement particularly for those students that took the role of researchers.

A set of materials offering step-by-step guidance have been developed that can be used by schools who wish to implement the *Inclusive Inquiry* approach. These have been prepared in five languages (English, Danish, German, Portuguese and Spanish) and can be downloaded for free from the project's website (www.reachingthehardtoreach.eu).

Getting started

The example that will be used aims to illustrate how dialogs within schools that used *Inclusive Inquiry* led to change in practices. At the same time, it will illustrate the fluidity of voices.

The example relates to two schools that are on the same site, English School 1 Infants and English School 1 Juniors. The two schools worked together for this project. Both schools were smaller than an average-sized primary school with 178 pupils and 201 pupils respectively. Four teachers were involved in the project: two from the Infants School (Year 2 teachers, 6–7 years old children) and two from the Junior School (Year 3 teachers, 7–8 years old children). One of the year 2 teachers led the whole process, the one who facilitated the children researchers training but was not involved in the teaching of the lessons.

A range of criteria were used to select the student researchers. Three were chosen from each class. There were four boys and five girls. A number of these children were seen by the teachers as not being confident, while others were children who were in receipt of pupil premium (additional funding for publicly funded schools in England to raise the attainment of disadvantaged pupils and close the gap between them and their peers), but at the same time lacked confidence and interest in learning. There was also a child who was negative toward other children in the class and rude at the same time; while another one was not behaving appropriately in class. Overall, the majority of the chosen children were children who lacked confidence.

Before the planning of the research lessons, the teacher facilitator had trained the student researchers, who then collected their classmates' views. Specifically, the student researchers had three sessions to be trained as researchers. During the first session they focused on exploring the children researchers' own views about learning, specifically what helps them and what makes it difficult for children's learning. The second session focused on how they would explore their classmates' views about learning and teaching. During the final session of training, they focused on how to carry out lesson observations. When the children researchers collected their classmates' views in all three classes, the most popular answers were the following: to use videos and to have games.

Co-designing lessons

At the end of this process the three teachers who were going to teach the lessons, the teacher facilitator and the nine student researchers all got together to discuss the lesson planning. This session started by the children researchers explaining to the teachers how they gathered their classmates' views, followed by a presentation of the findings from each class in relation to what children thought helps with their learning, such as the use of games and videos. Then the teachers explained that since the lesson that they had to plan would take place in both Year 2 and Year 3 classes, they decided to focus on a lesson about reading for pleasure, since this is something that they do in both year groups. They then asked the children researchers to discuss in small groups, with the support of one teacher in each group, about the various skills that are needed for reading. The children mentioned the following skills: fact finding, unknown words, predictions, inference, listening and summarizing. The teachers sometimes had to prompt children in order to remember and name some of these skills. Then the teachers asked them to focus on a skill that they are not so good at. They all agreed that prediction was a skill that they are not very good at. Therefore, they decided to focus on prediction for this lesson. The children worked again in small groups in order to think of various ideas that would help with practicing prediction. At the end they shared their ideas with the whole group. For example:

T1¹: We had a really good idea around the table. Should we share? What do you think we could do? I think from listening you have the same idea that we have as well. What's your idea?

S3: So, we could do it's like what half of the video clip and in our books, we could predict what's gonna happen next and predict who the author is.

T1: Ah good idea, what about D. here. Do you have similar idea?

T2: It's similar but J. has a really good idea. What could we do?

S2: So, we can like watch a video clip and then we can like predict what's gonna happen next and then we can watch the rest.

T2: What did you say before that? What could we do to help our prediction?

S2: (remains silent)

T2: I can say J. said we can do some drama so we could act it out, couldn't you? That's what you said, didn't you?

T1: That's one of the things that you've all said ... more interesting, didn't you? J. what's your idea (to another student)?

S5: We could read a book and then you predict what maybe gonna happen next.

T1: Fantastic any other ideas? A.?

S4: We have a story of Iron Man so we can watch a video clip and then we can like predict what's gonna happen next in the book.

Teachers all agreed:

T1: I think we can choose something that definitely nobody's read or watched before otherwise they'll know the answer and then you won't be able to predict. Any other ideas? M. was talking a lot about being able to move around the classroom weren't you. Can you explain a bit more?

S3: Or outside because you could have a game of like an adventure, then what we learned about, then could like put teddies around, then we can come back together, then answer questions of what we see kind of like that ...

T1: Any other ideas?

What was impressive here was that the teachers enabled the children to see that there was a clear purpose for this lesson—to develop the prediction skill and, therefore, any of their suggestions had to be linked back to this objective. Toward the end, it seems that S3's idea about moving outside the classroom, is not really taken into account and explored in more detail.

Having decided on the use of a video to predict what is going to happen next, the teachers moved on to exploring how the children were going to work.

T1: I've got a question for you. There're some good girls over here also said we can watch something and we can stop it and we can predict I said how would you like to do that? Would all everybody sit silent and write it in their books? How would you like to do that? Anyone's got any idea? What are the different ways we could do that?

S3: So we could like talk to our learning partners.

T1: Brilliant. Is that what you said? What did you say A.?

S8: We can make the game.

¹T: Teacher, S: Student.

T3: No, no, you said one of them was doing it on their own what did you say to me? You do with ...

S8: Buddy.

T1: Learning buddy. You have learning buddy, haven't you? We could do it with learning buddies. How can you do it? There could be one you, there could be two of you, you could work in a ...

Students: Group.

T1: Group ok. What do you think? What do you think is the best way to do it J.? or what do you think might work?

S: uhm ...

T1: The idea was what would you like to do it. If you are asked to do a job, would you like to do that job on your own, in a group or with a partner?

S: group.

The teacher facilitator then asked each student researcher whether they like working in a group, with a learning buddy or on their own and most of them chose as their preference working in group.

At the end of this planning session, the three teachers stayed a little longer to finalize the details of the lesson. What emerged through their discussion was that since the researchers would be observing the lesson in the other classes, they would have seen the video. Therefore, each lesson had to use a different video in order to allow all children to be able to practice the prediction skill. The teachers found three different videos to use in each lesson.

Dialogs that lead to changes

The first lesson was taught and observed by the student researchers from the other classes and the teachers. Following the first lesson, the teachers and the student researchers who observed the lesson had a discussion. The first person who spoke was the teacher who taught the lesson who said how she felt that the lesson went well overall. However, on reflection and if she had to teach it again, she thought that she would probably need to focus more on the justification of their predictions and for them to be in a position to explain a bit more why they thought that of a specific prediction. The teacher who led the discussion was the one that led the training of student researchers.

T1: What did you notice when you were watching the children?

S1: They were quite interested.

T1: What did they do to show us that they were interested?

S2: They were looking at the teacher.

S3: And listening.

T1: Were they listening to each other when they were working in groups? What do you think? What ideas did you write down?

S3: Some children had arm crossed.

T1: Yeah, what did that mean?

S3: It's err they were bored or something.

T1: But why they do their learning journey when you went closer to watch them, were they bored or were they enjoying the lesson?

S4: There were those like sitting there doing nothing.

T1: Okay so do you think that they were listening?

S4: No.

T1: Some children, we talked about some of children, weren't talking when they were sharing with others, lots of children were talking. What could we say might be a better way to get that conversation kind of all the children? What do we talk about? What could happen in that part when all the children were sharing ideas and some children were sat that ... like this. What do you think the teacher should do next time? What could we do in the lesson next time?

S5: Play a little game.

T1: How could we do that? How's that gonna help everyone have a turn? What did we say we could do next time to make sure everyone have a turn?

S1: So, like we could start from one person and then we go around to last person.

T1: So that everybody gets a chance to talk, didn't you? Everyone gets a turn to talk before they all start writing I thought that was a really good idea.

Again, here we see how a student's suggestion that children may be bored, which is then repeated by another student, is given attention by the teacher who leads the discussion. At the same time, it does seem like the teacher was trying to make them think more critically and move beyond what children's non-verbal communication was suggesting and think of alternative explanations, while she is also becoming more focused on children who did not appear to talk in their groups. Through sharing their observations, both teachers and children researchers started thinking about possible changes to the next lesson, to enable greater participation, and the children had some good suggestions about how to enable all in a group to have a say. One of the changes for the second lesson, was related to this aspect. Whereas in the first lesson the children worked in groups of five in the second lesson they reduced the number to three and each child had to express their views taking turns.

Thinking about voices and dialogs through this process

Returning to Lefstein et al. (2020) conditions of “voice” and Alexander’s (2019) related questions, it can be argued that through *Inclusive Inquiry*, all children were given the opportunities to express their views. Though the ones that were student researchers had a more prominent role in the whole process, at the same time all students were asked to express their views about learning and teaching before the design of the lesson. The fact that their views and ideas were taken into account to design the lesson, was a clear indication that their views were respected and treated equally. This was emphasized at the start of each of the lessons and it was important in communicating this to the students. Therefore, all children were given meaningful roles, as suggested by Mitra (2008).

However, the dialogs between teachers and children were more in depth with those who took the role of researchers. In the example here, we saw how the students drew attention to children in the lesson who may have been bored. Even though the teacher was trying to give them an alternative perspective and make them think more critically, at the same time she was also becoming more critical about what she was observing and why some children may have appeared to be disengaged by having their arms crossed. Therefore, we can see how dialogs outside the classroom context about a lesson were indeed allowing space for children to express their views, based on what they thought when they observed the lesson. This differs to the example presented by Lefstein et al. (2020) where, in the lesson that they observed, as they argue, though some of the contributions were student initiated, the contents were directed by the teacher, to the extent that the teacher’s voice was dominating. Whereas in our example, though the discussion was certainly led by the teacher, it could be argued that the children researchers’ suggestion of some children appearing not being engaged in the lesson that they observed was what brought into sharper focus details of the lesson. Specifically, the student researchers suggested what they called a game, which in essence was turn taking in group work, ensuring that everyone talks in a smaller group. Though this might seem obvious and a common practice among many classrooms, as a basic group skill, it was evident that this was not used in the particular classroom.

Final thoughts

The introduction of the student voice agenda to research, has been a welcome one and much needed. It could be argued that much of the research falling under the student voice umbrella term runs the risk of marginalizing certain voices, such as those of adults, in the effort of giving weight to students’ voices that have been traditionally marginalized.

Through *Inclusive Inquiry*, it is the coming together of different perspectives, those of adults and of children, that are important. Together these different perspectives, through the process of dialog aim to arrive at creative ways of working and ensuring that all children are included in the process of learning. Slee (2018) points out that our efforts should be “directed toward the ways in which institutions can pursue structural and cultural adjustments to support inclusion” (p. 67). I am arguing that dialogs between children and teachers facilitated such adjustments and led to changes in practice to enable all students to be active participants in lessons. Such notions of participation, strongly relate to the notion of voice. Students and teachers collaborating can embark on common journeys, and through dialogs such as the ones described here identify possible *directions of travel* for developing more inclusive practices.

References

- Ainscow, M., Kaplan, I., 2005. Using evidence to encourage inclusive school development: possibilities and challenges. *Australas. J. Spec. Educ.* 29 (2), 12–21.
- Ainscow, M., 2007. From special education to effective schools for all: a review of progress so far. In: Florian, L. (Ed.), *The SAGE Handbook of Special Education*. SAGE, London, pp. 146–159.
- Alexander, R., 2019. Dialogic pedagogy in a post-truth world. In: Mercer, N., Wegerif, R., Major, L. (Eds.), *The Routledge International Handbook of Research on Dialogic Education*. Routledge, London, pp. 672–686.
- Baroutsis, A., Mills, M., McGregor, G., te Riele, K., Hayes, D., 2016. Student voice and the community forum: finding ways of “being heard” at an alternative school for disenfranchised young people. *Br. Educ. Res. J.* 42 (3), 438–453.
- Blumer, H., 1969. *Symbolic Interactionism: Perspective and Method*. Prentice Hall, Englewood Cliffs (N.J.).
- Carrington, S., Bland, D., Brady, K., 2010. Training young people as researchers to investigate engagement and disengagement in the middle. *Int. J. Incl. Educ.* 14 (5), 449–462.
- Cook-Sather, A., 2006. Sound, presence, and power: “student voice” in educational research and reform. *Curric. Inq.* 36 (4), 359–390.
- Fielding, M., Bragg, S., 2003. *Students as Researchers, Making a Difference*. Pearson, Cambridge.
- Fielding, M., 2001. Students as radical agents of change. *J. Educ. Change* 2 (2), 123–141.
- Fielding, M., 2004. Transformative approaches to student voice: theoretical underpinnings, recalcitrant realities. *Br. Educ. Res. J.* 30 (2), 295–311.
- Graham, A., Truscott, J., Simmons, C., Anderson, D., Thomas, N., 2018. Exploring student participation across different arenas of school life. *Br. Educ. Res. J.* 44 (6), 1029–1046.
- Hope, M., 2012. The importance of belonging: learning from the student experience of democratic education. *J. Sch. Leader.* 22 (4), 733–750.
- Jackson, A.Y., Mazzei, L.A., 2019. In: *Voice in Qualitative Inquiry: Challenging Conventional, Interpretive, and Critical Conceptions in Qualitative Research*. Routledge, London.
- Kaplan, I., 2008. Being “seen” being heard: engaging with students on the margins of education through participatory photography. In: Thomson, P. (Ed.), *Doing Visual Research With Children and Young People*. Routledge, London, pp. 175–191.
- Kozleski, E.B., 2020. Disrupting what passes as inclusive education: predicating educational equity on schools designed for all. *Educ. Forum* 84 (4), 340–355. <https://doi.org/10.1080/00131725.2020.1801047>.
- Lefstein, A., Pollak, I., Segal, A., 2020. Compelling student voice: dialogic practices of public confession. *Discourse* 41 (1), 110–123. <https://doi.org/10.1080/01596306.2018.1473341>.
- Lewis, A., 2010. Silence in the context of “child voice”. *Child. Soc.* 24, 14–23.

- Lodge, C., 2005. From hearing voices to engaging in dialogue: problematising student participation in school improvement. *J. Educ. Change* 6, 125–146.
- Mayall, B., 2001. Conversations with children: working with generational issues. In: Christensen, P., James, A. (Eds.), *Research With Children: Perspectives and Practices*. RoutledgeFalmer, London.
- Mazzei, L.A., Jackson, A.Y., 2009. Introduction: the limit of voice. In: Jackson, A.Y., Mazzei, L.A. (Eds.), *Voice in Qualitative Inquiry: Challenging Conventional, Interpretive, and Critical Conceptions in Qualitative Research*. Routledge, London and New York, pp. 1–13.
- Mazzei, L.A., 2009. An impossibly full voice. In: Jackson, A.Y., Mazzei, L.A. (Eds.), *Voice in Qualitative Inquiry: Challenging Conventional, Interpretive, and Critical Conceptions in Qualitative Research*. Routledge, London and New York, pp. 45–62.
- McIntyre, D., Pedder, D., Rudduck, J., 2005. Pupil voice: comfortable and uncomfortable learnings for teachers. *Res. Pap. Educ.* 20, 149–168.
- Messiou, K., Ainscow, M., 2020. Inclusive Inquiry: student-teacher dialogue as a means of promoting inclusion in schools. *Br. Educ. Res. J.* 46 (3), 670–687. <https://doi.org/10.1002/berj.3602>.
- Messiou, K., Hope, M., 2015. The danger of subverting students' views in schools. *Int. J. Incl. Educ.* 19 (10), 1009–1021.
- Messiou, K., Ainscow, M., Echeita, G., Goldrick, S., Hope, M., Paes, I., Sandoval, M., Simon, C., Vitorino, T., 2016. Learning from differences: a strategy for teacher development in respect to student diversity. *Sch. Effect. Sch. Improv.* 27 (1), 45–61.
- Messiou, K., 2012. *Confronting Marginalisation in Education: A Framework for Promoting Inclusion*. Routledge, London.
- Messiou, K., 2019. The missing voices: students as a catalyst for promoting inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 768–781. <https://doi.org/10.1080/13603116.2019.1623326>.
- Mitchell, S.N., Reilly, R.C., Logue, M.E., 2009. Benefits of collaborative action research for the beginning teacher. *Teach. Teach. Educ.* 25, 344–349.
- Mitra, D.L., Serriere, S.C., 2012. Student voice in elementary school reform: examining youth development in fifth graders. *Am. Educ. Res. J.* 49 (4), 743–744.
- Mitra, D.L., 2003. Student voice in school reform: reframing student–teacher relationships. *McGill J. Educ.* 38 (2), 289–304.
- Mitra, D.L., 2004. The significance of students: can increasing “student voice” in schools lead to gains in youth development? *Teach. Coll. Rec.* 106 (4), 651–688.
- Mitra, D.L., 2008. Balancing power in communities of practice: an examination of increasing student voice through school-based youth–adult partnerships. *J. Educ. Change* 9, 221–242. <https://doi.org/10.1007/s10833-007-9061-7>.
- Morrow, V., Richards, M., 1996. The ethics of social research with children: an overview. *Child. Soc.* 10 (2), 90–105.
- Prout, A., James, A., 1990. Introduction. In: Prout, A., James, A. (Eds.), *Constructing and Reconstructing Childhood: Contemporary Issues in the Sociological Study of Childhood*. Falmer Press, London.
- Reay, D., 2006. “I’m not seen as one of the clever children”: consulting primary school pupils about the social conditions of learning. *Educ. Rev.* 58 (2), 171–181. <https://doi.org/10.1080/00131910600584066>.
- Rudduck, J., 2006. The past, the papers and the project. *Educ. Rev.* 58 (2), 131–143. <https://doi.org/10.1080/00131910600583993>.
- Slee, R., 2018. *Inclusive Education Isn't Dead, It Just Smells Funny*. Routledge, London and New York.
- St Pierre, E.A., 2009. Afterword: decentering voice in qualitative inquiry. In: Jackson, A.Y., Mazzei, L.A. (Eds.), *Voice in Qualitative Inquiry: Challenging Conventional, Interpretive, and Critical Conceptions in Qualitative Research*. Routledge, London and New York, pp. 221–236.
- Tangen, R., 2009. Conceptualising quality of school life from pupils' perspectives: a four-dimensional model. *Int. J. Incl. Educ.* 13 (8), 829–844. <https://doi.org/10.1080/13603110802155649>.
- Thomson, P., Gunter, H., 2006. From “consulting pupils” to “pupils as researchers”: a situated case narrative. *Br. Educ. Res. J.* 32 (6), 839–856.
- Thomson, P. (Ed.), 2008. *Doing Visual Research With Children and Young People*. Routledge, London.
- Thomson, P., 2011. Coming to terms with voice. In: Czerniawski, G., Kidd, W. (Eds.), *The Student Voice Handbook: Bridging the Academic/Practitioner Divide*. Emerald Group Publishing Limited, Bingley.
- United Nations, 1989. *The UN Convention on the Rights of the Child*. United Nations, New York.
- Wegerif, R., 2019. Introduction to the theory of dialogic education. In: Mercer, N., Wegerif, R., Major, L. (Eds.), *The Routledge International Handbook of Research on Dialogic Education*. Routledge, London, pp. 11–13.
- Wilson, S., 2000. Schooling for democracy. *Youth Stud. Aust.* 19 (2), 25–32.

First Nations & Indigenous identities in education

Erica Neeganagwedgin, Western University, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Overview of Indigenous models of education before contact	151
Institutionalized violence and forced education systems	151
Mission schools	152
Key assimilation reports	152
The impact of residential schools	153
Federally run day schools	155
Conclusion	156
References	157

Overview of Indigenous models of education before contact

When Europeans made their way to what is now known as Canada, Indigenous Peoples on these territories were sovereign nations who lived and enacted their systems of education through their everyday activities and life. Indigenous Peoples have engaged in their own education systems which were informed by their epistemologies. Indigenous societies have always had their own economies, leadership structures and knowledge systems. That was as true then as it is now (Coates, 2018; para 11).

Long before the arrival of Europeans and the creation of Canada, Indigenous People lived on their ancestral Lands on which they implemented their own ways of living (Assembly of First Nations, 2018, 26). They lived in harmony with, and were respectful of, the Land. According to McCue (2011, Para 1), traditional education among most Indigenous People was accomplished through various techniques, such as observation and practice, family and group socialization, oral teachings and participation in community ceremonies and institutions. Tenning (2010) explains Indigenous worldviews on teaching and learning, observing that children were also learners and teachers whose skills developed as a result of reflection, collective deliberation, and experimentation. Indigenous forms of education are not static. According to McKinley and Smith (2019, p. vii), "... these education systems had no ending: each generation expanded the community's knowledge base".

Three Laws of Relationship (Western Canadian Protocol for Collaboration in Basic Education, 2000, as cited in Laramee, 2008, p. 57) discusses key characteristics of Indigenous worldviews which guide education and interactions of most Indigenous Peoples of the Land. First, The Law of Relationship with the self, the Law of Relationship with nature and the Law of Relationship with the Land guided the behavior of learning and living.

Indigenous Peoples and the knowledge systems of First Nations were often guided and grounded by relationships with each other and to the Land and by their community, with an effort "... on preserving a physical, emotional, mental and spiritual balance" (Coates, 2018, para. 2). Indigenous Education comes from the Land and from oral traditions. Elders constitute a key part of Indigenous People's teaching, learning and identity. Beyond the family, the community and the Elders, Indigenous Education is also born from the ancestral territories. Indigenous Education is distinct to each Indigenous People and Indigenous nation (Lifemosaic, 2020). Indigenous knowledges and pedagogies are complex. They are focused on experiential learning and on lived experiences that are meant to be enjoyed, but which also serve a deeper purpose. These concepts are meant to encourage both individual and collective responsibilities, which are crucial to Indigenous life and wellbeing (Coates, 2018, para. 5).

Tenning (2010) asserts that "... at the core of Indigenous forms of education was the value of respect: respect for oneself, for others, for the Creator, for the environment, for the ancestors, and for the generations to come in the future" (para 8). Indigenous People at the time were and still are self-determining nations with their own languages and governance systems which include their education (see Carr-Stewart, 2019, p. 7).

In Canada under its constitution, there are three recognized groups of Indigenous Peoples including, First Nations, Métis, and Inuit, under section 35 of the Canadian constitution (Tenning, 2010). Indigenous People's educational systems were impacted by the onslaught of colonial education and laws violently forced on them.

Institutionalized violence and forced education systems

European contact with Indigenous People in Canada, and particularly at the school systems which the Europeans, and Euro-Canadians designed for, and imposed upon Indigenous People are examined in this section of the chapter. Further, the term "Indian" sometimes appears in this discussion when referring to some Indigenous or First Nations Peoples. This is for two reasons. First, the word "Indian" is used in historical sources and may appear in direct quotes. Additionally, "Indian" is the legal term that is found in many government reports, and it is embedded in federal legislation in Canada, such as Canada's Indian Act, which imposes on and regulates the status of First Nations People. As a result, the context sometimes requires the use of this word.

Mission schools

In fact, Carr-Stewart (2019, p. 7) emphasizes the argument that with the arrival of Europeans, Indigenous Peoples and their education systems experienced dramatic and rapid changes. Indigenous Education systems were intentionally ignored since the goal of European governments was to colonize and control Indigenous People's lives. Successive governments and Christian churches have, for hundreds of years, dedicated their time and were determined to control, and ultimately erase, Indigenous Peoples and their knowledge systems and pedagogy, and this has had a major impact on Indigenous People's identity and lives (see Coates, 2018, para. 3).

For more than 200 years, from the early 1600s to the 1800s, before Canada's residential school system, which was aimed at forcing Indigenous People to forget their identity, became official in the 1800s, Christian missionaries ran mission schools which were aimed at teaching Christianity to Indigenous children. This was the precursor to the Government of Canada's residential school system (see Canadian [The Canadian Encyclopedia](#), 2020). Religious mission schools have existed in Canada since at least 1620. In that year the Recollect Order of Franciscans (Roman Catholic) organized schools at Quebec, which functioned until 1629 (see [Anishnawbe Aski Nation](#), 2005). By 1636 the Jesuits established a school in Quebec and accepted several students. However, many parents refused to send their children to the school because it was too far away from home (see [Anishnawbe Aski Nation](#), 2005). There were many attempts to establish schools for Indigenous children, such as the Ursuline Sisters' establishment of girls' schools in Quebec in 1639. In the early 1600s various Roman Catholic orders set up schools. The Catholic Church's strong efforts to set up schools did not last long (see [Miller](#), 1996). The Mohawk Institute was the oldest continuously run Anglican residential school in Canada, starting from 1828 (see [General Synod Archives](#), 2008). What's more, missionaries continued to establish Day schools and Boarding schools for First Nations and Inuit children from the 1600s to the 1850s (see [Kanai Board of Education](#), 2005).

It is widely known and well-understood that Indigenous People have their own forms of spirituality, education and governance systems, and at no point in time did Indigenous leaders see themselves in need of Christian conversion. They did not see education as a means to be used to assimilate their children or their communities. After all, they had their own ways of life that worked well. With the new changes that were imposed on them as European settlement, and with it European tyranny grew, and with Indigenous People's understanding and worldviews of thinking of the wellbeing of future generations, Indigenous leaders saw education as a way to adapt to the new changes, although not at the expense of Indigenous People's own governance and education systems (see [Kanai Board of Education](#), 2005).

Yet, Europeans used education as one of the ways in which they carried out their colonial project of acquiring Indigenous Lands. The Royal Commission on Aboriginal People, [RCAP](#) (1996, p. 232) explains that, of all the processes that they used to try to achieve the goal of colonization and assimilation, "... none was more obviously a creature of Canada's paternalism toward Aboriginal people, its civilizing strategy and its stern assimilative determination than education". The dissatisfaction of not reaching the goal of assimilation and Christianization through mission schools ultimately led the government to develop aggressive policies and laws which were focused on further removal of Indigenous People.

Indigenous People continue to experience the destruction of their cultures, the dismissal of their rights and the aggressive assault on their identity through Canada's residential school system. The term residential schools refer to an extensive school system which the Canadian government established, and the churches administered. This school system's objective was to force their education on Indigenous children. However, it also had the more damaging, and equally explicit, objectives of indoctrinating Indigenous children into Euro-Canadian and Christian ways of life and thought, and of assimilating them into Canadian society ([Hanson](#), 2009).

For example, the government—first Britain (before Canadian Confederation in 1867) and later Canada (after Confederation in 1867)—developed key reports to facilitate the process of their intended assimilation of Indigenous People.

Key assimilation reports

Decades of dedicated efforts by governments, churches, and others to erase First Nations' knowledge and pedagogy continues (Coates, 2018, para. 3). For example, The Bagot Commission, which was in effect from 1842 to 1844, presented its report to the Legislative Assembly in 1845. Its recommendation was that the separation of children from their parents was the best method of reaching assimilation into the Euro Canadian culture. The commission also recommended that the Mohawk Institute should be considered a model for other industrial schools (see Canadian [The Canadian Encyclopedia](#), 2020).

The historical records provide ample evidence of the intention of European governments of what the [Truth and Reconciliation Commission \(TRC\) Report](#) (2015a) terms a policy of cultural genocide. The Truth and Reconciliation Commission Report describes cultural genocide as the destruction of those structures and practices that allow the group to continue as a group, and it states that the aim of cultural genocide is to destroy "the political and social institutions of the targeted group". The Truth and Reconciliation Commission Report adds that the Land is taken away and the targeted population is forcibly removed with restricted and controlled movements. Also, ancestral languages and spiritual practices of the group are prohibited, and, importantly, Indigenous families' structures are disrupted. All of this serves to block the transmission of cultural values and identity intergenerationally ([TRC](#), 2015).

As a matter of fact, the fifth recommendation of the Bagot Report (1845, p. 8) recommended that:

... the establishment and maintenance of schools, which instruction shall be given as in the rudiments of education, as in agriculture and some of the handicrafts, and the English as well as the French language taught; and to promote these objects it is suggested for consideration, with some of the medals or ornaments now given as presents, might not be converted into prizes for proficiency in these pursuits; and whether it might not be advisable to make the gift of presents to Indians and their Families, conditional on their sending their children to such schools.

In 1847, another main individual in the attempt to establish residential schools was Egerton Ryerson. At the time, Ryerson was the superintendent of Education for Upper Canada, now the province of Ontario. In the report of May 26, 1847, Ryerson wrote the following:

Their purpose should be to 'give a plain English education adapted to the working farmer and mechanic' and in addition 'agriculture, kitchen-gardening and mechanics so far as mechanics is connected with the making and repairing the most useful agricultural implements'. To attain their objective, it would be necessary for the students to reside together, with adequate provision being made for their domestic and religious education.

Ryerson also recommended that the schools should be operated by the church and funded by the government (Kanai Board of Education, 2005, p. 175). After the release of Ryerson's report, Methodist missionaries organized and directed a number of schools in Southern Ontario in the 1850s (Truth and Reconciliation Commission, 2015c, p. 22).

In 1867 the British Parliament passed the British North America Act which created the Dominion of Canada, and responsibility for governing this new Canadian entity was passed to the new Canadian government. Under Section 91 of the Act, the relationship which the British had with First Nations was also transferred to the newly formed Canadian government (Kanai Board of Education, 2005, p. 50). There was no consultation with First Nations People who were (and are) themselves sovereign nations. The lack of consultation reinforced the paternal relationship and unequal power relations, revealing how the British and later Canadian governments perceived Indigenous People.

Even more, in 1879, Nicholas Flood Davin released a report which examined the boarding school models in the United States. This report called for a model of schooling in which children would attend schools which were far away from their homes, families, and communities. The report referred to these schools as "industrial boarding schools." They came to be known as residential schools in Canada. According to Davin, "... the schools would be more successful because they could be completely removing children from their homes where, they can no longer be influenced by their parents" (Davin, 1879, para 7).

Using the US schools as a model, Davin's report called for the creation and funding of off-reserve schools to teach children the skills which he felt they would need at the time in the contemporary Canadian economy. He advised the government to set up boarding schools rather than day schools. Residential schools, it was reasoned, would be more successful because they could completely remove children from their home environment and influences (Para 7). Meanwhile the then Prime Minister of Canada, John A. Macdonald, and the rest of the government and churches fully supported Davin's recommendations (Union of Ontario Indians, 2013). In the 1800s the government and churches established partnerships to carry out their goals of evangelizing and assimilating Indigenous youth. In 1868 there were two residential schools in Canada, but by 1894 there were 45. This demonstrates the growth of aggressive assimilation. Over 130,000 residential schools were built in Canada, with the last to close in 1996 (Truth and Reconciliation Commission, 2015e, p. 7).

The impact of residential schools

Subsequently, the system of residential schools in Canada was designed to eliminate Indigenous Peoples and their identity. When students entered the schools, their languages were assaulted, and their identity attacked. Indigenous children experienced punishment for speaking their languages and foreign dominant colonial languages such as French and English were officially enforced (see Truth and Reconciliation Commission of Canada, 2015a; Report of the Royal Commission on Aboriginal Peoples, 1996). Canada's residential school is only one part of the history of imperialism on Indigenous peoples and their Lands over the past 500 years. It is also part of the history of the settler colonial Canadian state and, as the Truth and Reconciliation Commission asserts, "that history is not over" (Truth and Reconciliation Commission of Canada, 2015b).

It is well-understood that Indigenous People have been on their ancestral Lands since time immemorial. Prior to disruption by settlers, there were an estimated 450 Indigenous languages and dialects in Canada belonging to 11 language families (see McIvor, 2008, p. 14). According to Statistics Canada (2016), more than 70 Indigenous languages were reported in the census with languages divided into 12 language families.

The eradication of Indigenous languages impacts identity, and "(T)he impacts of language loss are profound and devastating" (National Association of Friendship Centers, 2018, p. 4).

Indigenous languages were highly impacted and remain “critically endangered.” Yet, Indigenous People have developed creative ways to exercise and assert their sovereignty through language by using technology and social media platforms to engage and reclaim their languages. This is demonstrated in the many Indigenous language apps, TikTok videos and YouTube language classes with both Elders and youth engaging through technology. Indigenous organizations and Indigenous schools are focusing on and prioritizing Indigenous languages within their environments. For example, there is a Bachelor of Arts in Ojibwe Languages at the Six Nations Polytechnic in the province of Ontario. These are some of the creative ways in which Indigenous People continue to renew their languages (Six Nations Polytechnic, 2022).

According to the Truth and Reconciliation Commission of Canada (2015d, p. vii),

For most of the system’s history, students were strapped and humiliated, and in some of the residential schools, they were handcuffed, shackled, beaten, locked in cellars and makeshift jails, or displayed in stocks. In addition, overcrowding and a high student ratio meant children who were not subject to physical abuse still grew up in an environment of neglect.

According to the Report of the Royal Commission on Aboriginal People (1996, p. 349) “... the very language in which the vision of residential school was structured revealed what would have to be the essentially violent nature of the school system in its assault on child and culture”.

Consequently, when Canada attempted to disconnect Indigenous children from their parents, sending them to residential schools it was done ... to break their connection to their culture and identity (Truth and Reconciliation Commission of Canada, 2015e).

Many government officials upheld and acted upon these ideas. These officials included, among others, Davin in his 1876 report and Duncan Campbell Scott, another Canadian government bureaucrat, who in 1924 expressed how crucial he felt it was “... to get the children away from their parents and get them while they are young”. According to a Department of Indian Affairs annual report (1890, p. 15):

It would be highly desirable, if it were practicable, to obtain entire possession of all Indian children after they attain to the age of seven or eight years and keep them at schools of the industrial type until they have had a thorough course of instruction, not only in the ordinary subjects taught at public schools, but in some useful and profitable trade, or in agriculture, as the aptitude of the pupil might indicate he was best fitted for. Were such a course adopting the solution of that problem, designated “the Indian question” would probably be affected sooner than it is likely to be under the present system.

These schools were part of a larger government policy which had, at its heart, Canada’s 1876 Indian Act, which was designed to legislate and control the lives of Indigenous People. The Indian Act both imposed and determined ideas of who they believed Indigenous People were under Canadian law. These laws, and the processes which the laws codified, also determined (and continue to determine) when Indigenous People were no longer Indigenous (Truth and Reconciliation Commission of Canada, 2015d, p. 11). Nevertheless, Indigenous Peoples are distinct and unique nations in what is now known as Canada today. They knew who they were then as they do now.

Meanwhile, it was estimated that over 6000 children died in Canada’s residential schools. Recently, the bodies of hundreds of Indigenous children were located in unmarked graves and, as of August 2021, it is estimated that over 1300 unmarked graves were noted at various residential school sites in Canada (see Voce et al., 2021). The unmarked graves were not news to many communities. In fact, the Truth and Reconciliation Report, Volume 4, (2015 p. 11) stated that “On the basis of the work undertaken to date, it is apparent that there are likely to be other unidentified residential gravesites across the country.” There is nowhere in the world where any parent would send their children to school to die. These schools operated for over 100 years with more than 100 and 50,000 students forced into attendance (Truth and Reconciliation Commission of Canada, 2015e). Both Indigenous parents and their children resisted these schools in many ways. For example, according to Kostecky (2019, p. 86) residential schools can be understood as containing “geographies of resistance”, wherein places of oppression can also be seen to include “areas of resistance, resilience, and endurance.” As Indigenous People’s advocacy and activism led to the phasing out of residential schools, Indigenous People themselves established “effective regional and national organizations and, in the courts and the legislatures, they advocated for their right to self-government” (Truth and Reconciliation Commission of Canada, 2015e, p. 129). These procedures related to the Residential Schools were part of overwhelmingly destructive, oppressive policies to remove Indigenous People as distinct Peoples and to assimilate them into the Canadian society (Truth and Reconciliation Commission of Canada, 2015e, p. 3).

Indigenous children and parents resisted this policy in multiple ways. For example, some parents refused to enroll their children as students, or they sometimes refused to return children to school. They also advocated for the government to improve the quality of education, food, and clothing in these schools. As the Truth and Reconciliation Commission reported, in taking these steps, parents often put themselves at risk of legal reprisals (Truth and Reconciliation Commission of Canada, 2015e, p. 114). This is because under Canada’s Indian Act, residential school attendance was mandatory. Therefore, refusing to send the children to school meant Indigenous parents could be criminalized under Canadian law.

Significantly, Canada’s residential school system was not created by any single individual or law. It developed from the days of mission schools in the 1600s, and the pervasive mentality developed over time as new laws and policies were created and

implemented. It was systemic and institutionalized in the British laws, and later in Canadian law, by people who believed in the assimilationist colonial project and were adamant about its success (Adams et al., 2018, p. 24).

Federally run day schools

In addition to residential schools, some mention of the federally run Day schools is critical in understanding Indigenous People's experiences of Canada's broader colonial project. This is an area not frequently researched, and the experiences of federally run Day School Survivors are critical to understanding Canada's Day Schools that operated alongside residential schools. These schools were funded, controlled, and managed by the government of Canada. They were not residential schools (Grand Council of the Crees, 2019).

Rather, Day schools had First Nations, Métis and Inuit children attend during the day, but the students lived away from the schools.

Both residential and Day schools were part of the genocidal project of settler-colonial Canada (Wall, 2019, para. 11). For example, in the province of New Brunswick, there were 12-Day schools, there was no residential school. Currently, there is a Class Action lawsuit, called the *McLean Class Action*, concerning federally run Day schools. This is, yet another example of Indigenous People asserting their sovereignty. A quick review of the list of federal Day schools shows that hundreds of Day schools operated throughout Canada (see *Federal Indian Day School Class Action*, n.d.). These federally funded Day schools were run by many Christian churches including Roman Catholic, the United Church, Methodist, Anglican, Mennonite, Presbyterian, and Salvation Army. There were also a few non-denominational schools. These were different from the original mission schools which were discussed earlier in this chapter and were operated by missionaries (Kanai Board of Education, 2005, p. 181).

This chapter does not discuss another aspect of Canada's system of control—that is, the Day Scholar experiences. Day Scholars are students who attended the same Residential School systems as many First Nations, Inuit, and Métis students. However, they did not sleep at the school overnight (see *Justice for Day Scholars*). The Day Scholars were not included in Canada's 2008 federal apology to Residential School Survivors, and the Day Scholars also have a Class Action suit called (*Gottfriedson v. Canada*) (Grand Council of the Crees, 2019). This is a critical example of Indigenous self-determination, particularly, considering that Indigenous People under Canada's Indian Act were prohibited from hiring lawyers until amendments were made to the Act in 1951. And, despite being a Day Scholar like the Day school experience, the Day Scholars experienced assault to their languages, cultures and identities in these colonially imposed schools.

Wall (2019, para. 9) points to the need to see the linkages between these institutions and activities, and to gain a deeper understanding of Canada's objectives in setting up residential schools and "... the techniques used to that end, are decipherable in connections between residential and Day Schools". In reference to Day schools, Suddath (cited in *CBC News*, 2020, para 6) asserts that the education system was "... mandatory, colonial, religious, paternalistic and discriminatory". Suddath states further that: "It did not reflect the traditional Indigenous ways of educating their children. So, it's a collective trauma that will take generations to heal."

According to Wall (2019, p. 18), this illustrates the "... genocidal context within which the system was established and endured". According to the recent final report of the National Inquiry into Missing and Murdered Indigenous Women and Girls (*Manitowabi*, 2019, Volume 1A, p. 25), "this is genocide".

In the Assembly of First Nations policy statement called "Indian Control of Indian Education" which was released in 1972, First Nations leaders called for a change that would make education relevant to the philosophy and needs of First Nations People (Tremblay, 2001). The history of residential school, Day school and Day Scholars' experiences show that even under oppressive conditions, Indigenous People found ways to assert their self-determination. Indigenous People have initiated, led, and developed their own schooling systems. There is an understanding that education is crucial to governance and identity.

Abu-Saad and Champagne (2006) remind us that the delivery of education to Indigenous Peoples has been problematic. Indigenous Peoples and communities are living cultures. "Their worldviews, institutions and current actions are designed for the future" (Abu-Saad and Champagne, 2006). Indigenous People have the right to determine when and how jurisdiction will be exercised over the education of their children. Only First Nations communities have jurisdiction over the education of their people (Tremblay, 2001, p. 17). Indigenous Peoples globally are reasserting their cultures and redefining their political, economic and social systems. Key to this process is ensuring that Indigenous youth continue to develop a sense of pride in their ancestral territories through their own education (see Maina, 1991, p. 294).

Whether in rural or urban contexts, Indigenous People are drawing on their own models of education. One example that comes to mind is the Amiskwaciy Academy which has been around for a little more than 2 decades. It is the first urban Indigenous high school in the Western part of Canada. This school was developed in consultation with Indigenous Elders and communities, and its architecture, curriculum and environment reflect Indigenous epistemologies, worldviews and ways of life (Amiskwaciy Academy, 2020). This model of education demonstrates the ways in which the government can honor Treaties and show commitment to Indigenous Education. I had the opportunity to visit Amiskwaciy, and the walls and classroom designs reflect Indigenous ways of knowing, Indigenous ways of teaching and Indigenous ways of learning. This is one example of how Indigenous People continue to exercise their sovereignty as the Original Peoples of these Lands now known as Canada. Other examples include the use of Indigenous languages and land-based education. This is remarkable demonstration of Indigenous sovereignty following the imposition of the destructive Residential Schools.

In another example, a postsecondary institution in Western Canada that is entrenched in Indigenous knowledges and methodology is University nuhelot'ine thaiyots'į nistameyimākanak Blue Quills. This accredited university historically was part of the

residential school system which was designed to disconnect Indigenous People from their culture and identity. Today it is the first Indigenous owned and controlled university in Canada. Its transformation and existence came about through the efforts and leadership of Indigenous peoples to reclaim their education, and in recognition of their sovereignty as Indigenous people who are rooted in their ancestral territories. In a message, Blue Quills former president, L. Makois ([Blue Quills First Nations College, 2002](#), p. v), reasserts the view that “Our belief is that spirituality is the highest form of consciousness. Through its evolution, Blue Quills has walked full circle to return our people to the original teachings of, love, honesty, sharing, and determination.”

At the university, Indigenous culture and teaching philosophies are salient and the university promotes and supports Indigenous pedagogical approaches and identities. These two models of Indigenous education and governance systems focus, and emphasize Indigenous curricula, perspectives and voices. For example, they center local Indigenous languages in the region into their curriculum including Cree and Dene. Elders are crucial to Indigenous teaching and learning and have important roles in the learning environment of these Indigenous institutions. A quick review of courses and programming shows a curriculum that is Indigenous informed, and community based in every subject area. When the curriculum is structured in Indigenous epistemologies and pedagogy, [Shirley \(2017, p. 168\)](#) writes,

It provides youth with opportunities to examine core aspects of their Indigenous identity in addition to acquiring knowledge and skills required for protecting, promoting, and preserving Indigenous people, language, culture, and land. These processes contribute to rigorous academic work that are intellectually engaging and empowering.

Indigenous Education is not static. Indigenous Education too can refer to the ways in which Indigenous People have survived and defy colonizing and assimilative agenda by asserting their sovereign right to educate their children according to their own cultures, languages, and knowledges ([Hopkins, 2020](#), p. 5).

The calls-to-action which resulted from Canada’s Truth and Reconciliation Commission are about understanding and redressing the history and legacy of residential schools. Of the Commission’s 94 calls-to-action under Education and Reconciliation, calls 62 and 63 explicitly state:

We call upon the federal, provincial, and territorial governments, in consultation and collaboration with Survivors, Aboriginal peoples, and educators, to: i. Make age-appropriate curriculum on residential schools, Treaties, and Aboriginal peoples’ historical and contemporary contributions to Canada a mandatory education requirement for Kindergarten to Grade Twelve students.

We call upon the Council of Ministers of Education, Canada to maintain an annual commitment to Aboriginal education issues.

[Truth and Reconciliation Commission of Canada \(2015f, p. 7\).](#)

These are direct calls for Canada to examine and make right its historic wrongs and to end its continued systemic oppression of Indigenous Peoples.

Conclusion

The history of residential schools in Canada is part of a broader history of global colonization and imperialism of Indigenous Peoples and their ancestral territories. The laws and policies which were developed suppressed Indigenous Peoples and their identity ([Truth and Reconciliation Commission of Canada, 2015f](#), p. 133). In fact, the residential schools, the Day Scholars and the federally funded Day schools were all part of the Eurocentric assimilative and genocidal policies.

Indigenous People lived their lives and educated their children just as any other peoples and cultures. They were, and still are today, distinct peoples living, working, and thriving on their Lands prior to colonial governments systemic oppression and targeting of Indigenous culture resulting in genocide. They were subjected to European violence that questioned their validity and viability and threatened their identity and ways of life. The British and Canadian governments introduced and enforced laws and policies which were intended to facilitate the process of aggressive assimilation and systemic abuse engrained in enforced policies and legislations.

Despite these pressures, Indigenous People expressed their self-determination and sovereignty in various ways in response to the attacks on their knowledge systems. The [Assembly of First Nations \(2018, p. 37\)](#) in their learning series Plaintalk 3 stated: “It is a testament to the strength, resilience, will, determination, and power of First Nations peoples that they have survived pressures that can only be described as genocidal”. Schools such as Amiskwaciy and University nuxelhot’ine thaa?ehots’i nistameyimâkanak Blue Quills, and other governance and leadership initiatives and engagements by Indigenous People, big or small, work toward sustaining Indigenous Peoples and their future. [LifeMosaic \(2020\)](#) suggests that these all help to prepare a new generation of connected Indigenous leaders, who are ready and willing to support their Elders, protect Indigenous rights, cultures, and territories, and examine new ways to promote Indigenous knowledge and practices while following in the ways of their ancestors.

References

- Abu-Saad, I., Champagne, D., 2006. *Indigenous Education and Empowerment: International Perspectives*. California AltaMira Press, A Division of Rowman & Littlefield Publishers, Inc., Walnut Creek.
- Adams, J., Clarke, L., Kwan-Lafond, D., Mather, M., Thornhill, N., Winterstein, S., Meness, J., Kinsella, S., Mutamaba, M.R., Roffey, C., McQuabbie, J., Enberg, S., Chief Lady Bird, Aura, 2018. Our stories: first peoples in Canada. In: openlibrary-repo.ecampusontario.ca. Centennial College. <https://openlibrary-repo.ecampusontario.ca/jspui/handle/123456789/865>.
- Amiskwaciyepsbca, 2020. Amiskwaciy Academy. <https://amiskwaciy.ebsp.ca/>.
- Anishnawbe Aski Nation, 2005. Indian Residential Schools in Ontario, Early Church Efforts. http://rschools.nan.on.ca/article/early-church-efforts-1118.asp#_ftn1.
- Assembly of First Nations, 2018. Impacts of Contact. It's Our Time Toolkit. Plain Talk 3. AFN [Online]. Available at. http://ss.yukonschools.ca/uploads/4/5/5/0/45508033/plaintalk_3_-_impacts_of_contact.pdf. (Accessed 10 July 2022).
- Bagot, C., 1845. Report on the Affairs of the Indians in Canada Laid Before the Legislative Assembly, 20th March, 1845. IRSHDC: Library Item. <https://collections.irshdc.ubc.ca/index.php/Detail/objects/9431>.
- Blue Quills First Nations College, 2002. Pimoheskanaw, 1971–2001: Blue Quills First Nations College. Blue Quills First Nations College, St. Paul, AB.
- Carr-Stewart, S., 2019. Introduction. In: Carr-Stewart, S. (Ed.), *Knowing the Past, Facing the Future: Indigenous Education in Canada*. UBC Press, Purich Book, Vancouver, BC, pp. 3–21.
- CBC News, 2020. Survivors of New Brunswick's 12 "Indian Day Schools" Can Receive Compensation, Says Lawyer. <https://www.cbc.ca/news/canada/new-brunswick/new-brunswick-day-schools-1.5441763>.
- Coates, T., 2018. Education [online]. indigenouspeoplesatlasofcanada.ca. Available at. <https://indigenouspeoplesatlasofcanada.ca/article/education-2/>.
- Davin, N.F., 1879. Report on Industrial Schools for Indians and Half Breeds. Report to the Minister of the Interior, Ottawa, ON. http://www.canadianshakespeares.ca/multimedia/pdf/davin_report.pdf.
- Department of Indian Affairs. Annual Report of the Department of Indian Affairs for the Year Ended December 31, 1890. Dominion of Canada. Ottawa: ON: Brown Chamberlain. <http://central.bac-lac.gc.ca/item/?id=1890-IAAR-RAAI&op=pdf&app=indianaffairs>.
- Federal Indian Day School Class Action, (n.d.). Schedule K—List of Federal Indian Day Schools. <https://indiandayschools.com/en/wp-content/uploads/schedule-k.pdf>.
- General Synod Archives, 2008. The Anglican Church of Canada. The Mohawk Institute, Brantford, ON. <https://www.anglican.ca/tr/histories/mohawk-institute/>.
- Grand Council of the Crees (Eeyou Istchee), 2019. Update on Indian Day School Settlement Agreement. Grand Council of the Crees. Available at. <https://www.cngov.ca/news-issues/current-issues/update-on-indian-day-school-settlement-agreement/>.
- Hanson, E., 2009. The Residential School System [online]. Arts.ubc.ca. Available at. https://indigenousfoundations.arts.ubc.ca/the_residential_school_system/.
- Hopkins, P.J., 2020. Indian Education for All: Decolonizing Indigenous Education in Public Schools. Teachers College Press, NY.
- Justice for Day Scholars, (n.d.). Available at <https://www.justicefordayscholars.com/>.
- Kanai Board of Education, 2005. Peoples and Cultural Change: Aboriginal Studies 20. Duval House Publishing, Edmonton.
- Kosteckyj, A., 2019. Geographies of resistance. The Mirror—Undergraduate Hist. J. 39 (1), 86–95. https://doi.org/10.5206/mirror_v39i1.5814.
- Laramée, M., 2008. Making sense of aboriginal education in Canadian public schools: a case study of four inner city elementary principals and their vision of aboriginal education. *First Nations Perspect.* 1 (1), 57–73.
- LifeMosaic, 2020. Indigenous Education | Projects | LifeMosaic [online] Available at. <https://www.lifemosaic.net/en/projects/indigenous-education/>. (Accessed 24 November 2020).
- Maina, F., 1991. Culturally relevant pedagogy: first nations education in Canada. *Can. J. Native Stud.* 2 (1991), 293–314. Available at. http://www3.bruntonu.ca/cjns/17.2/cjns17no2_pg293-314.pdf.
- Manitowabi, B., 2019. National Inquiry Into Missing and Murdered Indigenous Women and Girls. Foreword. Reclaiming Power and Place: The Final Report on the National Inquiry Into Missing and Murdered Indigenous Women and Girls, vol. 1A, p. 25. Retrieved from. https://www.mmiwg-ffada.ca/wp-content/uploads/2019/06/Final_Report_Vol_1a.pdf.
- McCue, H., 2011. Education of Indigenous Peoples in Canada | The Canadian Encyclopedia. <https://www.thecanadianencyclopedia.ca/en/article/aboriginal-people-education>.
- McIvor, O., 2008. Building the Nests: Indigenous Language Revitalization in Canada Through Early Childhood Immersion Programs of Human and Social Development. (Master of Arts). University of Victoria. https://dspace.library.uvic.ca:8443/bitstream/handle/1828/686/mcivor_c2005.pdf?sequence=1&isAllowed=y.
- McKinley, E.A., Smith, L.T., 2019. Towards self-determination in indigenous education research: an introduction in McKinley, E.A., Smith, L.T. (Eds.), *Handbook of Indigenous Education*. Springer, New York, pp. 1–15.
- Miller, J.R., 1996. Shingwauk's Vision. A History of Native Residential Schools. University of Toronto Press, Toronto.
- National Association of Friendship Centres, 2018. Our Languages, Our Stories Towards the Revitalization and Retention of Indigenous Languages in Urban Environments. https://sencanada.ca/content/sen/committee/421/APPA/Briefs/DiscussionPaper_NAFC_e.pdf.
- Report of the Royal Commission on Aboriginal Peoples, 1996. <https://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/royal-commission-aboriginal-peoples/Pages/final-report.aspx>.
- Ryerson, A.E., 1847. Education Office. Report on Industrial Schools. Education Office, Ottawa, ON. <http://nctr.ca/assets/reports/Historical%20Reports/Ryerson%20Report.pdf>.
- Shirley, V.J., 2017. Indigenous social justice pedagogy: teaching into the risks and cultivating the heart. *Crit. Quest. Educ.* 8, 163–177.
- Six Nations Polytechnic, 2022. Bachelor of Arts in Ojweheweh Languages. Six Nations Polytechnic. <https://snpolytechnic.com/program/bachelor-arts-ogweheweh-languages>.
- Statistics Canada, 2016. The Aboriginal Languages of First Nations People, Métis and Inuit. <https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016022/98-200-x2016022-eng.cfm>.
- Tenning, A., 2010. Canada's History—Canada's History [online] Available at. <https://www.canadahistory.ca/education/lesson-plans/walking-on-the-lands-of-our-ancestors>.
- The Canadian Encyclopedia, 2020. Residential Schools Timeline Residential. The Canadian Encyclopedia, Historica Canada. <https://www.thecanadianencyclopedia.ca/en/timeline/residential-schools>.
- Tremblay, C.P., 2001. First Nations Educational Jurisdiction. Assembly of First Nations, Ottawa. https://www.afn.ca/uploads/files/education/6_2001_afn_first_nations_educational_jurisdiction_national_background_paper_-_april_2001.pdf.
- Truth and Reconciliation Commission of Canada, 2015. Executive summary. In: *Canada's Residential Schools: Missing Children and Unmarked Burials: The Final Report of the Truth and Reconciliation Commission of Canada*, vol. 4. McGill-Queen's University Press, Montreal; Kingston; London; Chicago, pp. 1–12. <https://doi.org/10.2307/j.ctt19mbnh.3>.
- Truth and Reconciliation Commission of Canada, 2015a. Summary of the Final Report of the Truth and Reconciliation Commission of Canada. <http://www.trc.ca/about-us/trc-findings.html>.
- Truth and Reconciliation Commission of Canada, 2015b. Canada's Residential Schools: The History, Part 1 Origins to 1939. Library and Archives Canada. Truth and Reconciliation Commission. McGill-Queen's University Press, Ottawa, ON. http://nctr.ca/assets/reports/Final%20Reports/Volume_1_History_Part_1_English_Web.pdf.
- Truth and Reconciliation Commission of Canada, 2015c. What We Have Learned, Principles of Truth and Reconciliation. Library and Archives Canada. Truth and Reconciliation Commission. McGill-Queen's University Press, Ottawa, ON. <http://www.trc.ca/assets/pdf/Principles%20of%20Truth%20and%20Reconciliation.pdf>.
- Truth and Reconciliation Commission of Canada, 2015d. The Survivors Speak. Library and Archives Canada, Ottawa, ON. http://nctr.ca/assets/reports/Final%20Reports/Survivors_Speak_English_Web.pdf.
- Truth and Reconciliation Commission of Canada, 2015e. Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada. http://www.trc.ca/websites/trcinstitution/File/2015/Findings/Exec_Summary_2015_05_31_web_o.pdf.

- Truth and Reconciliation Commission of Canada, 2015f. Calls to Action. https://www2.gov.bc.ca/assets/gov/british-columbians-our-governments/indigenous-people/aboriginal-peoples-documents/calls_to_action_english2.pdf.
- Union of Ontario Indians, 2013. An Overview of the Indian Residential School System. Union of Ontario Indians. <http://www.anishinabek.ca/wp-content/uploads/2016/07/An-Overview-of-the-IRS-System-Booklet.pdf>.
- Voce, A., Leyl, C., Michael, C., 2021. "Cultural genocide": the shameful history of Canada's residential schools—mapped. The Guardian. Available at. https://www.theguardian.com/world/ng-interactive/2021/sep/06/canada-residential-schools-indigenous-children-cultural-genocide-map?fbclid=IwAR2L7BQrOEReWezQhBNF_qwcRcdXEPpKwjZ_JjgAFwjn-7pniMD8mybvmVU.
- Walls, M., 2019. The Complex Truth: Intersections Between Day Schools and the Shubenacadie Residential School. Nova Scotia Advocate. <https://nsadvocate.org/2019/11/19/the-complex-truth-intersections-between-day-schools-and-the-shubenacadie-residential-school/>.

Gender identities, sexuality and education

Wayne J. Martino, Jenny E. Kassen, and Kenan Omercajic, The University of Western Ontario, Faculty of Education, London, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	159
The experiences of gender and sexually diverse students in schools: large-scale survey studies	159
Experiences of gender and sexually diverse students of color	161
Curricular and pedagogical implications	161
Policy and practice implications	162
Desubjugated voices of gender and sexually diverse students in schools: small-scale qualitative studies	162
Knowing and being known	163
Safety	163
Toward desubjugation	164
The epistemological implications of queer and transgender studies in education	165
Conclusion	166
Acknowledgment	167
References	167

Introduction

It is impossible to do justice to the burgeoning field of gender and sexuality studies in education and the extent to which an increasing body of theoretical and empirical knowledge about gender and sexual diversity has come to inform our understandings of sex, sexual orientation, gender identity, gender expression and questions of embodiment. What we provide here is some insight into our own engagement with selected literature as a gesturing toward a mapping of the field that attempts to encapsulate a shift away from an earlier focus in *Lesbian and Gay Studies* with its minoritized focus on sexual identity categories to an embrace of *Queer Studies* with its envisioning of a more universalizing focus on a *radical* sexual politics with heteronormativity in social theory and research as its focus (Butler, 1993a; Gamson, 2000; Rubin, 1984; Sedgwick, 1991; Umphrey, 1995; Warner, 1991). However, we highlight the emerging and vital epistemological significance of *Transgender Studies* and its capacity to inform our understanding of gender identity, *trans* and nonbinary personhood and gender expansiveness in the field of education and specifically the site of schooling, which take us beyond the antinormative limits of queer theory (Kennedy, 2018; Martino and Cumming-Potvin, 2018; Wiegman and Wilson, 2015).

We begin with a review of the significant contribution of large-scale reports on the experiences of gender and sexually diverse (GSD) youth in the school system in their capacity to highlight the systemic barriers related to the impact of homophobia, heteronormativity, transphobia and cisnormativity on the lives of these students. This body of research has functioned to amplify the lived experiences of GSD students in schools with a very important focus on outlining recommendations which have very clear policy and practice implications for fostering gender and sexual diversity in schools, and more specifically supporting GSD students. We proceed by reviewing a selection of qualitative studies, which provide access to the voices and perspectives of GSD students in schools, to draw attention to the onto-epistemological, political and ethical necessity of such a desubjugating focus in educational research that is committed to supporting GSD youth in the education system. Finally, we provide some reflection on queer theory and its limits, while also acknowledging its productive potentialities for the field which, we argue, can only be realized through a more thorough engagement with and commitment to *Transgender Studies* and what it has to offer in terms of a necessary epistemological focus on *trans* bodily ontological accounts of gender personhood (Rubin, 1998; Stryker, 2006).

The experiences of gender and sexually diverse students in schools: large-scale survey studies

The increasing visibility of gender and sexually diverse (GSD) students in schools has illuminated the necessity for greater and more nuanced understanding of the ways in which current systems, policies and programs affect their experiences and perceptions of safety and support. More recently, there has been a growing interest in learning about the experiences of gender diverse students

specifically in schools, which have highlighted the inimical impact of cisgenderism and cisnormativity¹ on these youth (Bartholomaeus and Riggs, 2018; Ferfolja and Ullman, 2021; Kennedy, 2018, 2021; Luecke, 2011; Mangin, 2020; Martino et al., 2020; Riggs and Bartholomaeus, 2018; Sinclair-Palm, 2017). This attention to the experiences of *trans* and gender diverse students is a recent development, given that, as Greytak et al. (2013) have pointed out, historically, the “T” in “LGBT” studies in education has largely been omitted and erased from research that readily employs this acronym. However, a number of large-scale survey studies have included a focus on the experiences of both gender and sexually diverse youth in the school system (Bradlow et al., 2017; Kosciw et al., 2020; Taylor and Peter, 2011; Ullman, 2015). This research has, indeed, contributed to generating knowledge about the lives of GSD students in schools and the extent to which they feel supported. Significantly, only some reports signal the necessity of concomitant considerations of race, ethnicity, disability, social class and their intersections with gender and sexual diversity. However, *collectively*, these reports highlight global north similarities not only about the reality of the lived experiences of GSD students, but in the policy and practice recommendations they support.

In a nation-wide study that involved surveying 3700 students from across Canada through a 54-item questionnaire, Taylor and Peter (2011) sought to contribute to the “lack of a solid Canadian evidence base” regarding the experiences of GSD students, which served as a significant “impediment faced by educators and administrators” who needed to understand “the situation of LGBTQ students in order to respond appropriately and to assure their school communities that homophobic and transphobic bullying are neither rare nor harmless, but are major problems that schools need to address” (p. 14). The report identifies alarming rates of both verbal and physical harassment endured by GSD students. Specifically, “74% of *trans* students, 55% of sexual minority students, and 26% of non-LGBTQ students reported having been verbally harassed about their gender expression,” and when asked about being verbally harassed about their gender identity or sexual orientation daily or weekly, “37% of *trans* students, 32% of female sexual minority students, and 20% of male sexual minority students” (p. 15) reported affirmatively. Such findings highlight that homophobia and transphobia continue to be a pervasive problem in schools.

Unfortunately, these findings are neither exclusive nor confined to the Canadian context, nor are the reported instances of homophobia and transphobia simply an issue of the past. For example, Kosciw et al. (2020) conducted a nation-wide survey involving approximately 16,700 students between the ages of 13 and 21 across the United States, and revealed starkly similar findings to the Canadian context related to instances of verbal and physical harassment experienced by GSD students in schools, nearly a decade after Taylor and Peter’s study. In particular, they highlight that “[m]ore than two-thirds of LGBTQ students (68.7%) were verbally harassed at school in the past year based on their sexual orientation” while, just as alarmingly, “[o]ver half of LGBTQ students (53.7%) were verbally harassed at school in the past year based on their gender” (p. 28). Additionally, these “disturbing and increasing rates of verbal and physical harassment of sexuality and gender diverse students” were also reported in the Australian context (Ullman, 2015, p. 7). In a nation-wide survey comprising of over 700 students between the ages of 14–18 across Australia, Ullman (2014) noted that “45% of participants indicated that they had witnessed school-based physical harassment of classmates perceived to be sexuality and/or gender diverse,” whereas a “startling 94% of students had heard homophobic language at school, with 58% of these young people reporting hearing this language daily” (p. 7). Evidently, the school experiences of GSD students pertaining to homophobia and transphobia are pervasive and reported consistently across Canada, the United States, and Australia.

Further corroborating these experiences on a global north scale, a nationwide survey in the United Kingdom that specifically surveyed over 3700 LGBT young people between the ages of 11–19, found that “nearly half ... (45%—including 64% of *trans* pupils) are bullied for being LGBT at school,” and just over half of “LGBT pupils (52%) hear homophobic language ... and almost half (46%) hear transphobic language ‘frequently’ or ‘often’” (Bradlow et al., 2017, p. 6). These instances of harassment have been shown to exacerbate issues of mental health, as documented by Bradlow et al. (2017) who noted that “rates of poor mental health are alarmingly high among LGBT young people, with *trans* pupils at particular risk” (p. 30). In this respect, Kosciw et al. (2020) underscore the significance of schools having a Gay Straight Alliance (GSA) which has been shown to invite “a greater sense of belonging to the school community and greater psychological well-being among LGBTQ students, perhaps as a result of the overall positive impact of GSAs on the school environment” (p. 147). Taylor and Peter (2011) also noted the importance of GSAs as “safe space[s] in which LGBTQ students and allies can work together on making their schools more welcoming for sexual and gender minority students” (p. 127), and resultantly, students from schools who have GSAs were “much more likely to agree that their school communities are supportive of LGBTQ people compared to participants from schools without GSAs (53.1% vs. 26.4%)” (p. 128). The importance of these clubs and the recommendations for their creation are reiterated in the reports where GSAs exist (i.e., Canada and the United States). Importantly, it is also noted by Kosciw et al. (2020) that there are “benefits to GSA participation for LGBTQ students of color,” (p. 58) though research has suggested that LGBTQ students of color often avoid these spaces as “they do not see their school’s GSAs to be inclusive or useful” (Kosciw et al., 2020, p. 63) or that their identities or lived experiences are represented in GSAs and often find they are “under surveillance by their heterosexual Black peers” (McCready, 2004a, p. 42).

¹Lennon and Mistler (2014) define cisgenderism as “the cultural and systemic ideology that denies, denigrates or pathologizes self-identified gender identities that do not align with assigned gender at birth, as well as resulting behavior, expression and community” (p. 63) while Serano (2016) explains cisnormativity as “the assumption that a person’s assigned birth gender corresponds with a mutually exclusive set of masculine or feminine attributes, aptitudes, abilities, and desires” (p. 4).

Experiences of gender and sexually diverse students of color

Largely, the experiences of GSD students of color are only reflected in some reports (Kosciw et al., 2020; Taylor and Peter, 2011; Truong et al., 2020a,b; Zongrone et al., 2020a,b), while being otherwise overlooked or unaccounted for in others (Bradlow et al., 2017; Ullman, 2015). Importantly, Taylor and Peter (2011) make the distinction regarding the experiences of GSD youth of color, indicating that “Caucasian youth, both LGBTQ and non-LGBTQ, were far less likely to report having been physically harassed or assaulted because of their ethnicity: 8% compared to 13% of Aboriginal youth and 15% of youth of color” (p. 20) underscoring the pervasive reality of the aforementioned “double whammy” (Taylor and Peter, 2011, p. 65). Such an assertion overlaps with Kosciw et al. (2020) who reported that “all LGBTQ youth of color were at greater risk of experiencing multiple forms of victimization than their White LGBTQ peers” (p. 114). In light of these intersections of gender and race, Taylor and Peter (2011) insist that “policy, program, and curriculum development needs to reflect an understanding of how school climate for sexual and gender minority youth is affected by intersecting systems of social power such as racialization and poverty that are at work in all schools” (p. 28).

Similarly, Human Rights Campaign and Gender Spectrum (2018) highlighted the importance of addressing these problematic policies and systems in their report surveying over 5600 gender expansive youth, indicating that “difficulty in finding LGBTQ counselors of color or programs that relate directly to their experiences with gender and race-based discrimination often leads to feelings of isolation and a lack of sense of belonging” (p. 33). Therefore, the report emphasizes the necessity that “youth-serving professionals understand the intersection of race and LGBTQ identities, and the ways in which white supremacy polices the identities, expressions and existences of gender-expansive people of color” (p. 33). In addressing and considering these systemic barriers for *trans* youth of color, the statistics outlined in some of these comprehensive reports on GSD youth can be more purposefully addressed and considered. Speaking to more sweeping changes, every large-scale study “indicates a need to explicitly address gender identity, gender expression, and *anti*-transphobia in school and school board safer schools and equity and inclusive education policies” (Taylor and Peter, 2011, p. 18), signifying the necessity of both the development of GSD inclusive school policies and their enactment with respect to curricular and pedagogical implications.

Curricular and pedagogical implications

Each report highlights the importance of education about gender and sexual diversity which is largely lacking in many schools. For example, Bradlow et al. (2017) reported that “two in five LGBT pupils (40%) are never taught anything about LGBT issues” (p. 22). Meanwhile, Kosciw et al. (2020) found that when GSD content was discussed or referenced, “[o]nly 19.4% of LGBTQ students were taught positive representations of LGBTQ people, history, or events in their schools” (p. xxii). This absence of inclusion of gender and sexual diversity in curriculum is a consistent finding and focus across major reports examining the experiences of GSD youth in schools. Even when there is some measure of GSD-inclusive curriculum, it is often “covered in only one course, which would minimize any impact” and therefore, Taylor and Peter (2011) implore that “LGBTQ matters need to be implemented more widely across subject areas” (p. 125). In this respect, Ullman (2015), in her nationwide report of GSD student experiences in Australia, recommends “[t]raining for school staff which addresses expectations for inclusivity of sexuality and gender diversity and provides guidelines for inclusive language and related resource integration across all areas of the curriculum” (p. 9). However, these opportunities for professional development are scarce and it is often the parents of and/or GSD students who are responsible for driving such initiatives or challenging cisgenderism (Bartholomaeus and Riggs, 2018; Carlile and Paechter, 2018; Ferfolja and Ullman, 2021; Luecke, 2011; Rodela and Tobin, 2017; Slesaransky-Poe et al., 2013; Travers, 2018). As Payne and Smith (2014) indicate, “few professional development opportunities are provided to practicing teachers to gain knowledge about the LGBTQ student experience” (p. 400), and moreover, feel that it is “‘not their role’ to discuss sex, sexuality, and gender with their students” as they view their students as “innocent and asexual” (p. 401; see also DePalma and Atkinson, 2006).

As such, Truong et al. (2020a) highlight that a school curriculum which is “inclusive of diverse identities” may contribute to “instill[ing] beliefs in the intrinsic values of all individuals” (p. 32) thereby accounting for the fact that “Black LGBTQ students have unique school experiences, at the intersection of their various identities, including race, gender, and sexual orientation” (p. xx). Such inclusivity is particularly significant given that “LGBTQ students of color (including Black, AAPI, Latinx, Native and Indigenous, MENA, and multiracial LGBTQ students) commonly experienced both racist and *anti*-LGBTQ victimization at school” (Kosciw et al., 2020, p. 146). This intersectional compounding victimization is referred to by Taylor and Peter (2011) in their report as a “double whammy” where GSD students of color are “not only being physically harassed or assaulted because of reasons related to gender or sexual orientation, or both, but they are also much more likely to be physically harassed or assaulted because of their ethnicity” (p. 65). In fact, Kosciw et al. (2020) insist that “school staff must support LGBTQ youth of color with an intersectional approach that acknowledges and responds to racism, homophobia, and transphobia, and to the ways these interconnected forms of oppression may influence one another” (p. 114). To achieve this, Taylor and Peter (2011) recommend that: “Faculty of Education integrate LGBTQ-inclusive teaching and intersectionality into compulsory courses ... so that teachers have adequate opportunities to develop competence before entering the field” (p. 140) and therefore, can more effectively provide positive representation and considerations in their pedagogy, especially when considering the experiences of GSD students of color. Intertwined with such recommendations are also other policy and practice implications which these studies also highlight as necessary.

Policy and practice implications

Kosciw et al. (2020) underscore “important benefits associated with both comprehensive *anti*-bullying/harassment policies, as well as policies affirming the rights of transgender and nonbinary students” (p. 147). In addition, Bradlow et al. (2017) also note that schools should “have clear and widely promoted policies and procedures preventing and tackling homophobic, biphobic and transphobic bullying and language” (p. 38). This advocacy for GSD-inclusive policies and ensuring that they are understood by students and staff is especially significant, as Ullman (2015) notes, given that “sexuality and gender diverse students who are aware of inclusive policy are more likely to feel safe and less likely to report school-based victimization related to their sexuality and/or gender diversity” (p. 34; see also Kull et al., 2015). Of course, given the identity vectors that only heighten experiences of victimization for GSD students of color, Taylor and Peter (2011) insist upon not just “*anti*-homophobia, *anti*-biphobia, and *anti*-transphobia [but also] intersectionality measures in safer schools policies and programs” in order “to provide institutional authority and leadership” (p. 139).

Taylor and Peter (2011), specifically emphasize the significance of *intersectional* considerations in policy (as well as curricular) development and enactment. Such considerations are significant given that the creation of *anti*-discrimination legislation often only considers “the least marginalized of the marginalized,” while omitting “a person who faces intersecting vectors of harm [who] would be unlikely to benefit from *anti*-discrimination law” (Spade, 2015, p. 44). It is for this reason that Spade (2015) advocates for a *critical trans politics* that moves beyond an “individual framing of discrimination and ‘hate violence’ and [to] think more broadly about how gender categories are enforced on all people in ways that have particularly dangerous outcomes for *trans* people” (p. 9). Policy creation often “permits—and even necessitates—that efforts for inclusion in the discrimination regime rely on rhetoric that affirms the legitimacy and fairness of the status quo” (Spade, 2015, p. 44), and therefore, Taylor and Peter’s (2011) insistence for an intersectional approach is significant.

Beyond policy considerations, there are common recommendations across the reports for ensuring that leadership (e.g., administrators, department heads, etc.) “send out a strong message to all pupils and staff that homophobic, biphobic and transphobic bullying and language is not acceptable” (Bradlow et al., 2017, p. 38). Moreover, Kosciw et al. (2020) affirm the importance of supportive educators who evidently positively influence “students’ academic performance, educational aspirations, feelings of safety, school absenteeism (missing fewer days of school), psychological well-being, and connection to their school community” (p. 147). In procuring such supportive staff, Taylor and Peter (2011) advocate for “professional development opportunities to assist schools in the implementation of LGBTQ-inclusive and intersectionality curriculum” (p. 140). However, Payne and Smith (2014) have indicated that these opportunities are rarely provided to educators, which ultimately limits the capacity for teachers to recognize and achieve what Ullman’s (2015) report advocates for, which is “establish[ing] a classroom climate where sexuality and gender diverse students feel safe to advocate for curricular inclusions and to report marginalization from peers or adults in the school” (p. 42). In the following section, we turn our attention to qualitative studies which allow for more particularized and detailed insights into the perspectives and lived experiences of GSD students in schools.

Desubjugated voices of gender and sexually diverse students in schools: small-scale qualitative studies

GSD people in general have long been marginalized, *subjugated* and marked as unqualified knowers of their experience (Namaste, 2000; Stryker, 2006). Children and youth, particularly, are often subjugated due to the belief among some medical practitioners that an individual’s age determines what they can and cannot know about their own sexual or gender identity (Pyne, 2014). To seek out and center the voices of GSD students in research, then, is to provide an opportunity for “desubjugating previously marginalized forms of knowledge about gendered subjectivity and sexed embodiment” and thus ascribe rightful value to “the kind of knowledge that transgender people, whether academically trained or not, have of their own embodied experience, and of their relationships to the discourses and institutions that act upon them” (Stryker, 2006, p. 13). Youth-centered, qualitative inquiries add to a growing body of empirical work which helps us better understand the local-level implications of the themes proposed by the large-scale studies reviewed above (Porta et al., 2017), presenting an ever-growing set of recommendations that can better address the issues that impact on GSD students in the school system (Earnshaw et al., 2020; McGlashan and Fitzpatrick, 2017; Snapp et al., 2015; Wilson and Cariola, 2019).

More importantly, qualitative studies often provide more in-depth, particularized knowledge and specificity in detailing the perspectives of GSD students in schools, which further enhance and enrich our understanding of their lived experiences in these institutions, revealing complexities and nuances related to struggle and resistance that are not reflected in larger scale survey research. In this section, we explore two major themes, *knowledge* and *safety*, that are taken up by researchers in qualitative, youth-centered studies which focus on the experiences of GSD youth in schools. The sampling of studies reviewed in the following section are not exhaustive; they illustrate the ways in which knowledge and safety are conceptualized and mobilized by youth, including foci on youth of color, and bring to light the critical role of GSD youth in changing their contexts through advocacy work. Additionally, these studies also raise important questions about the limitations of adult-driven research that presents GSD youth as objects of inquiry vs. collaborators or co-creators.

Knowing and being known

Knowledge is often discussed as formal, (e.g., curricular inclusion, representation, and omission particularly, sexual health and social sciences curriculum), and informal (e.g., alternative educational spaces such as Pride or Rainbow clubs) (Gowen and Wings-Yanez, 2014; Snapp et al., 2015). Qualitative research demonstrates that the availability of formal knowledge in schools where GSD people are represented is inconsistent, at best, and that representation alone does not guarantee the accuracy of the information presented. More often than not, students report missed opportunities for representation (Gowen and Wings-Yanez, 2014; Ullman, 2014). Youth participants clearly indicate that, when faced with limited curricular representation, students consistently act with agency and request inclusion. Youth from a California study (Snapp et al., 2015) confirm that, although GSD was present in some social sciences and health classes, the information they received was superficial and incomplete, such as the exclusive inclusion of gay men in sexual health. These youth also reveal that teachers often ignore their repeated requests, despite students directing educators to specific curricular connections between GSD history, contemporary issues and class content. Importantly, these perspectives illuminate youth agency at work. In response to academic erasure, GSD students take it upon themselves to educate teachers about their legal duties under California's FAIR Education Act, and they create alternative learning spaces to discuss the pertinent issues that are overlooked in formal learning space (Snapp et al., 2015). One student in this study actively mentored an educator in creating inclusive class content that can be drawn on by other history teachers in the school (Snapp et al., 2015). In the face of erasure, youth employ a variety of tactics from "pushing the issue" (Ullman, 2014, p. 436), that is to say, confronting the educator about their silence or omission, to intentionally remaining silent because of the implicit understanding that any attempt to draw attention to GSD issues would be dismissed (Gowen and Wings-Yanez, 2014). It is widely noted that youth-led knowledge mobilization often migrates from the classroom to alternative learning spaces, such as Pride clubs or even to hallways for informal conversations with teachers outside of formal instructional time (Gowen and Wings-Yanez, 2014; Ullman, 2014).

The common thread that connects these experiences and responses is that youth are keenly aware of when their identities are being erased, but also of the power dynamics at play. In each instance they weigh the trade-off between troubling erasure and the subsequent imposition on them to act as educator with exposure and possibly opening themselves up to further alienation. GSD youth exist in a paradox where their erasure from curriculum is a major detriment to their academic engagement, and their visibility—a result of requesting inclusion and representation—also makes them vulnerable to further discrimination (Ullman, 2014). Youth advocates who take on the task of educating their educators exemplify the fact that, more than being experts of their own experience, (let alone being called on to be experts full stop on the experiences of GSD students in general), youth knowledge extends beyond simple curricular inclusion to knowledge about the power dynamics at work in school systems that help them navigate relations with staff, and astute critical thinking that allows them to elegantly maneuver GSD content into formal and informal educational contexts. However, scholars such as Frohard-Dourlent (2018) point to the limitations of student led change, especially when *trans* students are placed in the driving seat of such educational advocacy work in their demand for recognition without the provision of the necessary resources needed. They argue that "[e]mphasising self-determination and putting responsibility on individual students and their families to adapt to or change educational contexts obscures the (re)production of gendered, racialized and classed inequalities" and that "this process may create situations where educators are absolved from challenging institutional and cultural marginalization of gender diversity, and from their role as potential failures" (p. 338).

Safety

As previously established, *safety* is conceptualized broadly as a physical and psychological well-being, and the violence that GSD students face is subsumed under the umbrella of "bullying" (Earnshaw et al., 2020; Kosciw et al., 2020; Taylor and Peter, 2011). Youth voice represented in qualitative research extends this generalized understanding of *safety* by offering a more nuanced exploration of how GSD students conceptualize safe and unsafe environments. Additionally, their observations highlight the relationship between knowledge and safety and illuminate the critical role of student-driven advocacy in youth defining and ensuring their own safety at school.

Bathrooms are the focus of research that investigates the absence of safety, the very real presence of danger, and the implications of this reality on the experience of school for GSD students (Earnshaw et al., 2020; McGlashan and Fitzpatrick, 2017; Porta et al., 2017). Youth highlight that physical and verbal harassment are only one facet of what make school bathrooms unsafe for them. GSD youth from a study by Porta et al. (2017) reveal that the indirect and anticipated violence in bathroom spaces equally impacts on their well-being. In this study, youth share first and second-hand accounts of their experience with school bathrooms and through their testimonial, demonstrate how simply *knowing* of potential for violence in these spaces erodes their sense of safety. It should be noted that the "bathroom challenges" that youth describe occur on multiple levels (Porta et al., 2017, p. 110). In addition to overt acts of violence such as being forcibly removed from the space, they also report more insidious forms of epistemic violence (see Fricker, 2007; Hookway, 2010) in which their presence in sex-segregated bathrooms is policed by other occupants who assume the youth is in the "wrong" space based on their gender presentation. The potential for having one's gender policed and being made to justify one's existence is detrimental to a students' overall sense of well-being at school (Porta et al., 2017). Violence related to bathrooms also occurs via the absence of knowledge as is the case for one participant showcased in

Porta et al. (2017) work: with no information available about bathroom options, the participant is forced to out themselves at the school office in order to find out what facilities are available to them. This incident was framed as a “positive experience” by researchers because the student was eventually granted access to a staff washroom, though this representation minimizes the fact that no information about alternative washroom spaces was readily available.

A school climate cannot be deemed safe if students are required to make themselves intelligible to staff in order to access information about something so basic as bathroom use protocol (see Martino et al., 2020). Interestingly, where the schools appear to work reactively as opposed to proactively to support GSD youth (Earnshaw et al., 2020; McGlashan and Fitzpatrick, 2017; Porta et al., 2017; Ullman, 2014), youth seize on this opportunity to mobilize knowledge for the benefit of their GSD peers, and demonstrate how youth advocacy is at the heart of changes in school climate and safety. Returning to Porta et al. (2017), youth express that a gender-neutral bathroom only contributes to a safer school climate if the student body knows that it exists, and if they feel comfortable enough to access it. One student reported taking on the labor of raising awareness about this space and ensuring that it be welcoming for all students by ensuring supplies such as menstrual products and condoms are available there (Porta et al., 2017). Youth perspectives with respect to school bathrooms broaden the understanding of how safety and danger manifest and emphasize that the gaps in knowledge about GSD folk must be filled in order for school climates to shift (Ullman, 2014). Such studies also provide much more particularized knowledge about student agency and illuminate the capacity of GSD youth as leaders and advocates in bringing about necessary change. However, while these studies highlight that GSD students are not simply “docile bodies” (Foucault, 1977); they also unwittingly point to the failure of schools and educators to take responsibility for actively creating the necessary conditions for supporting the safety and well-being of GSD students in the first place (Frohard-Dourlent, 2018).

Indeed, the advocacy work required for the creation of gender-neutral spaces often falls to student-led clubs such as Gay Straight/Queer Straight, Gender and Sexuality Alliances (GSAs), and the presence of these student groups in schools have been identified in the research as a benchmark for measuring safety in schools for GSD youth (McGlashan and Fitzpatrick, 2017, 2018; Porta et al., 2017). The presence of a GSA in and of itself, however, does not indicate a ubiquitous increase in safety for club members. In an example from New Zealand (McGlashan and Fitzpatrick, 2018), youth participants who are members of their school’s Rainbow club situate their safety in the realm of cultural intelligibility at the intersection of transness and Paskifika culture and ethnicity, stating that what made them feel valued at school is “... like, we have each other. There’s someone else that is like you and they are going through what you’re going through, like the same, so we can relate” (McGlashan and Fitzpatrick, 2018, p. 10). For these two youth, safety within their Rainbow club is connected to their sense of solidarity through shared experience of gender, culture, and ethnicity and they report feeling more secure in their identities as a result of coming together in these spaces. The prerequisite for the safety to explore and express oneself here, is as the authors put it, cultural intelligibility: as these students reported, the potential for self-knowledge, discovery, and expression is amplified when they feel their experience validated by their “sister” (p. 10). The GSA itself is not the locus for safety, but rather a vehicle by which students can find safety in each other and a space which provides an opportunity for cultural intelligibility.

Such studies which provide the perspectives of GSD youth safety are essential to uncovering what specific change is needed in particular school contexts and at the local level. Youth accounts make clear that *safety* is far more complex than a matter of bathrooms, clubs, or name-calling; it is embedded in every aspect of school life and cannot be measured only by the presence of gender-neutral spaces or the absence of reports of physical harm (Earnshaw et al., 2020). The safety of GSD students must also take into account how intersectional identities are being affirmed, how knowledge is being mobilized, and how the responsibility for this labor is divided between staff and students.

Toward desubjugation

These studies highlight the need for a commitment to a desubjugation of GSD student voices in the education system in terms of their capacity to deepen and extend our understanding of the lived experiences of gender and sexual diversity in schools. In fact, such studies have the capacity to counter dominant narratives of victimization and docility which have come to define the experiences of these students (Rasmussen et al., 2004). Given the systemic erasure, pathologization and disenfranchisement of GSD people of which there is a long history (Gamson, 2000; Kuhl, 2019; Namaste, 2000), it is vital to create spaces for GSD youth to voice their own experiences and expertise which is not often acknowledged or actively sought out in schools. Such research creates the possibility for generating counter-narratives. Here, we borrow from Stryker’s conceptualization of desubjugation who proposes a *renarration* by which the “considerable intellectual heritage” of GSD communities, a knowledge born of lived experience speaks back normative narratives that have been established by “legitimate” knowers, such as those from academia and the education system. Though employed with the purpose of advancing *Transgender Studies* as a field, Stryker’s (2006) arguments for the desubjugation of marginalized voices can certainly be applied to learning from GSD youth specifically. Subjugation comes in the form of prioritizing adult voice over that of youth—a power dynamic which invariably is performed in schools and which also is replicated through research, even if that research focuses on youth.

In a number of studies in the field, GSD youth repeatedly assert that adults play a pivotal role in terms of impacting their lives at school, and yet school staff are consistently described as being disconnected from the reality of GSD students in their care and, in fact, have been identified as failing to support or “disinclined to intervene” (Taylor and Peter, 2011, p. 107) on their behalf, and “nearly half [of surveyed students] reported that staff never intervened” when hearing homophobic or transphobic remarks (Kosciw et al., 2020, p. 23; see also Allen et al., 2020; Earnshaw et al., 2020; Gowen and Wings-Yanez, 2014; McGlashan and Fitzpatrick,

2017, 2018; Sinclair-Palm, 2017; Taylor and Peter, 2011; Wyss, 2004; Youth Gender Action Project). Whether, as Earnshaw et al. (2020) report, staff are oblivious to the violence that GSD face, or as reported by McGlashan and Fitzpatrick (2017, 2018), well-meaning educators impose their own anxieties about GSD visibility on the youth they support, consequentially muting student initiatives to assert their identities in public spaces, adults are not attuned to their needs. If desubjugation entails recognizing previously disqualified knowers, in this case GSD youth, as legitimate and valuable knowers, then the power dynamics between educator and student must be re-examined, a hard lesson in humility that the former are not always willing to learn (see Frohard-Dourlent, 2018; Sinclair-Palm, 2017). Educators have at their disposal young people who are eager to share their perspectives, but youth must feel supported in this endeavor (Gowan and Wings-Yanes, 2014; Snapp et al., 2015, see also Schmitz et al., 2019). The labor that youth have been undertaking might undergo a desubjugation: as youth report, the educational work they do to ensure representation is marked as subjugated by virtue of being an extracurricular club activity.

In addition to educators learning from youth participants, researchers must also reflect on what it means to facilitate and enact desubjugation. Research that centers youth voice does not necessarily include GSD youth in meaningful ways throughout the research design process. Discussions that impact on GSD young people are quite often led by cisgender, heterosexual, adults (see Porta et al., 2017). As such, in a field that is predominantly adult-centered, it is adult-designated questions that form the basis for inquiries of all scales. This is not to criticize the work that has been done up until now, but rather to ask, if the ethical and politically exigent goal is desubjugating youth voices, then what do we stand to gain by integrating them more meaningfully into the fabric of research that ultimately impacts directly on their lives? This sample of research shows that GSD youth are answering adult-generated questions with insight that is invaluable, however it cannot be taken for granted that the questions they are being asked are the most relevant to their current reality. Youth have voice, knowledge, agency, regardless of whether “legitimate” knowers from academia and the education system grant them audience or not. There is further work to be done to excavate how participatory-oriented research can transform into youth-led research that is concerned to address fundamental and vital questions of testimonial and epistemic justice (Fricker, 2007).

The observations and contributions of GSD youth to research thus far have demonstrated that GSD youth have, beyond a shred of a doubt, the capacity to collaborate in more meaningful ways. As such these studies have not only provided invaluable knowledge about the many ways in which GSD youth themselves (and often supportive parents) are in the driving seat in terms of ensuring their safety in schools (Ferfolja and Ulman, 2021), but also illuminates their leadership role in educating adults and mobilizing knowledge about GSD issues among themselves. In this sense, the question of both GSD student agency, as well as the need the school system to take responsibility for educating about cisgenderism and *trans* inclusion is highlighted (Martino et al., 2020). As such, this youth centered research not only troubles how GSD youth are currently positioned in schools, but raises important methodological considerations about the need for more participatory modes of research. Such approaches move beyond GSD youth merely being produced as objects in connecting educational research to enacting social justice in generating knowledge that is committed to ensuring testimonial as well as epistemic and hermeneutic justice (Fricker, 2007; Lather, 1991; Smith, 1996).

The epistemological implications of queer and transgender studies in education

In reviewing and engaging with some of the significant research, what is highlighted are not only important methodological considerations in generating knowledge about GSD youth in the school system, but epistemological and theoretical questions pertaining to the framing and construction of knowledge about gender and sexual diversity. For example, the significance and contribution of queer theoretical perspectives in the field cannot be denied. Scholars in education have been inspired by Foucauldian and Butlerian insights into dominant systems of heteronormativity and heterosocialization with implications for shifting the analytic focus away from an exclusive focus on the sexual minority subject per se to a more universalizing critique of the mechanisms of *privileging* and *othering* (Kumashiro, 2002) at the heart of normalizing heterosexuality (Sedgwick, 1991, 1993/1999; Foucault, 1990; Butler 1993a). DePalma and Akinson's (2009) and Epstein and Johnson's scholarship in the United Kingdom, as well as that of Rasmussen (2006a) in Australia and Letts and Sears (1999) and Pinar (1998) in the United States are just a few representative collections that provide queerly informed insights into the schooling of sexualities.

Foucault's (1977) analytics of power, particularly in terms of how it informs our understanding of the regulatory and systemic ways in which knowledge of bodies and sex are “put into discourse” (Foucault, 1990, p. 10) with specific biopolitical effects for understanding gender and sexual normativity and diversity in schools, has had a significant impact on generating an important *body of knowledge* about gender and sexual identity that ruptures, and challenges what Britzman (1998) refers to as an “orientation to normalcy as the pernicious production of such binaries as self/other” (p. 82). Butler's (1993b) conceptualization of the *heterosexual matrix*, along with her focus on the compulsory iteration of gender norms which “operate by requiring the embodiment of certain ideals of femininity and masculinity” has indeed informed much queer educational research that has been committed to the “subversion of the norm” with respect challenging hegemonic modes of heterosocialization and gender normalization in schools (p. 22; see also Ingrey, 2013; Rasmussen, 2001, 2006b). Pinar's (1998) *Queer Theory in Education*, along with Britzman's (1995) “Is there a queer pedagogy or stop thinking straight” in our view encapsulate foundational work in education which translated the pedagogical significance of queer theory in the field of education. This body of work is also representative with respect to understanding the logics of antinormativity (Wiegman and Wilson, 2015) as a defining feature of queer theoretical frameworks and their critical application in the field of education with respect to persistent calls to expose and interrupt heteronormativity and its problematic institutionalization in schools (Ferfolja, 2007; Blackburn, 2007; Goldstein et al., 2007).

However, scholars such as McCready (2004a,b) have drawn attention to the need for more intersectional analytic approaches by centering the voices of queer youth of color thereby exposing the limits of the whiteness of queer studies and its analytic focus and erasure of race (Logie and Rwigema, 2014). Bartone (2019), for example, argues that “it is vital for all educators to consider how racism and heteronormativity influence their pedagogy, the classroom and school environment, and the educative experiences of their students” (p. 99). His research on the intersections of race and sexuality in the schooling experiences of black queer youth speaks to the epistemological and political significance of employing both critical race and queer theories—what he identifies as *quare theory* and which according to Johnson (2001) provides a basis for analysis in which “the discursive process of mediated identification and subjectivity in a political praxis [that] speaks to the material existence of “colored bodies” (cited in Bartone, p. 99).

In more recent times a body of growing research in education has emerged which has focused on the experiences of transgender, nonbinary and gender diverse youth in schools. This literature has highlighted vital issues and questions pertaining to the barriers and failure of schools to adequately support *trans* students and to educate about gender identity and gender diversity (Ferfolja and Ullman, 2020; Frohard-Dourlent, 2016, 2018; Sinclair-Palm and Gilbert, 2018; Payne and Smith, 2014; Luecke, 2018; Mangin, 2020; Martino et al., 2020; Morgan and Taylor, 2019; Riggs and Bartholomaeus, 2018; Rodela and Tobin, 2017; Ullman, 2014). However, Martino and Cumming-Potvin (2018) advocate for a much greater engagement with *trans* studies’ epistemological accounts of gender identity, selfhood and embodiment in the field of education, which we agree is central to a *critical trans politics* that is committed to *trans* desubjugation (Spade, 2015; Stryker, 2006). Such an epistemological focus requires a shift away from focusing on heteronormativity to an emphasis on what Kennedy (2018) refers to as “cultural cisgenderism” which refers to “a systematic erasure and problematizing of *trans* people and the distinction between *trans* and cisgender people [which] ... essentializes sex/gender as biologically determined, fixed at birth, immutable and externally imposed on the individual” (p. 308). In this sense, there are some limits to the capacity of queer theory to account for the complexities of embodied experiences of gender diverse people which *trans* scholars themselves have pointed out and which have often resulted in a problematic construction of *trans* people who embrace a more stable gender identity category as “gender defenders,” while those who counter gender norms are read as “gender outlaws,” thereby devaluing the former by casting them in essentialist terms as merely “conforming to the rules of gender [that the latter have] otherwise transgressed” (Elliot and Roen, 1998, p. 238; Keegan, 2020).

Such a position epitomizes the antinormative limits of queer theory in its capacity to account in phenomenological terms for the bodily ontological experiences of many *trans* people which *trans* scholars such as Rubin (1998) have railed against. He argues “that ‘queer’ signifies one condition for articulating a radical *trans* agenda [which] also represents an opportunity for the appropriation of transsexuals by nontranssexuals ... our lives have been appropriated to demonstrate the theories of gender performativity but only to the extent that they fail to reproduce the normative correspondence between body morphology and gender identity assumed as a matter of course by nontranssexuals” (p. 276). This critique has been extended by *trans* scholars such as Bettcher (2013) who points out that *trans* individuals who self-identify as men or women are “every bit as resistant as those who do not,” and that such “deployments of identity categories ... involve radical departures from dominant practices of gender and, hence, an alteration of the meaning of terms such as ‘woman’ and ‘man’” (p. 403). In this capacity, *trans* scholars have identified the limits of queer theory and its potential for enacting hermeneutic and epistemic injustice (Fricker, 2007) in its imposition of an antinormative logic which has consequences with respect to delegitimizing and failing to properly understand the bodily ontological experiences of *trans* men and women and their gender identities (see Kaufmann, 2010 as an example). Our reading of *trans* scholarly literature and engagement with *Trans Studies* specifically highlights that there is indeed a need for a more critical interrogation of queer theory in the field of Education and, hence, an epistemological commitment to embracing *trans*-informed analytic frameworks which center a critical focus on cisgenderism as a basis for understanding the terms of an ethical commitment to *trans* desubjugation.

Conclusion

In this article we have provided some critical insights into research on gender and sexuality in education though with a caveat that what we offer is limited, serving only as a snapshot of sorts from of our own various and situated standpoints from within the field. An overview of large-scale survey studies from Australia, North America and the United Kingdom have been useful in providing knowledge about the lived experiences of GSD students in schools and identifying significant barriers with specific implications for reform in the education system in the Global North, both with respect to ensuring safety for all students and, pedagogically, in terms of highlighting the need for queer and *trans* supportive curricula and affirmative spaces (Cavanaugh, 2016; Martino and Cumming-Potvin, 2016, 2019; Omercajic and Martino, 2020; Martino et al., 2021; Lapointe et al., 2017). However, we highlight the contribution of qualitative studies in education which amplify and center the desubjugated voices of queer and *trans* youth in schools with their potential for offering further nuanced analysis and particularized insights into the lived experiences of GSD students and their agency. In engaging with selected research, we highlight the need for studies that pay particular further methodological attention to participatory approaches to knowledge generation in thinking through important questions of testimonial, hermeneutic and epistemic justice in supporting the ethico-political project of desubjugation of queer and *trans* youth in the education system. We also provide some critical insights into the epistemological necessity of the need for further sustained engagement with *trans* studies’ theoretical scholarship which has highlighted the antinormative limits of queer theory.

Acknowledgment

SSHRC (Social Sciences Humanities Research Council Canada) Grant entitled:

Supporting transgender and gender minority youth in schools: Policy and practice [435-2015-0077] (Principal Investigator: Wayne Martino; Research Assistants: Jenny Kassen and Kenan Omercajic).

References

- Allen, L.R., Watson, L.B., VanMattson, S.B., 2020. *Trans young adults' reflections on adolescent sources of extra-familial support*. J. LGBT Youth 17 (1), 1–23.
- Bartholomaeus, C., Riggs, D.W., 2018. *Transgender People and Education*. Palgrave Macmillan, New York, NY.
- Barone, M., 2019. "I want them teachers to treat us like human beings": educational experiences of black gay youth. In: Martin, D., Strom, K. (Eds.), *Exploring Gender and LGBTQ Issues in K-12 and Teacher Education*. Information Age Publishing, Charlotte, NC, pp. 95–109.
- Bettcher, T., 2013. Trapped in the wrong theory: rethinking *trans* oppression and resistance. Signs 39 (2), 383–406.
- Blackburn, M., 2007. The experiencing, negotiation, breaking, and remaking of gender rules and regulations by queer youth. J. Gay Lesb. Issues Educ. 4 (2), 33–54.
- Bradlow, J., Bartram, F., Guasp, A., Jadva, V., 2017. School Report: The Experiences of Lesbian, Gay, Bi and *Trans* Young People in Britain's Schools in 2017. Center for Family Research, University of Cambridge, Cambridge.
- Britzman, D., 1995. Is there a queer pedagogy? Or, stop reading straight. Educ. Theor. 45 (2), 151–165.
- Britzman, D.P., 1998. *Lost Subjects, Contested Objects: Toward a Psychoanalytic Inquiry of Learning*. State University of New York Press, Albany.
- Butler, J., 1993a. *Bodies That Matter: On the Discursive Limits of "sex"*. Routledge, New York.
- Butler, J., 1993b. Critically queer. GLQ 1 (1), 17–32.
- Carlile, A., Paechter, C.F., 2018. *LGBTQI Parented Families and Schools: Visibility, Representation, and Pride*. Routledge, New York.
- Cavanaugh, L., 2016. Coming in/out together: queer(ing) schools through stories of difference and vulnerability. Arbutus Rev. 7 (1), 4–19.
- DePalma, R., Atkinson, E., 2006. The sound of silence: talking about sexual orientation and schooling. Sex. Educ. 6 (4), 333–349.
- DePalma, R., Atkinson, E., 2009. *Interrogating Heteronormativity in Primary Schools: The No Outsiders Project*. Trentham Books, Stoke-on-Trent, UK.
- Earnshaw, M., Menino, D.D., Sava, L.M., Perrotti, J., Barnes, T.N., Humphrey, D.L., Reisner, S.L., 2020. LGBTQ bullying: a qualitative investigation of student and school health professional perspectives. J. LGBT Youth 17 (3), 280–297.
- Elliot, P., Roen, K., 1998. Transgenderism and the question of embodiment: promising queer politics? GLQ 4 (2), 231–261.
- Ferfolja, T., Ullman, J., 2020. *Gender and Sexuality in a Culture of Limitation: Student and Teacher Experiences in Schools*. Routledge, New York and London.
- Ferfolja, T., Ullman, J., 2021. Inclusive pedagogies for transgender and gender diverse children: parents' perspectives on the limits of discourses of bullying and risk in schools. Pedagog. Cult. Soc. 29 (4).
- Ferfolja, T., 2007. Schooling cultures: institutionalizing heteronormativity and heterosexism. Int. J. Incl. Educ. 11, 147–162.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison*. Penguin Books, London.
- Foucault, M., 1990. *The History of Sexuality Volume 1: An Introduction*. Vintage, New York.
- Fricker, M., 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, Oxford.
- Frohard-Dourlent, H., 2016. "I don't care what's under your clothes": the discursive positioning of educators working with *trans* and gender-nonconforming students. Sex. Educ. 16 (1), 63–76.
- Frohard-Dourlent, H., 2018. "The student drives the car, right?": *trans* students and narrative of decision-making in schools. Sex. Educ. 18 (4), 328–344.
- Gamson, J., 2000. Sexualities, queer theory and qualitative research. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook of Qualitative Research*, second ed. Sage, London, pp. 347–365.
- Goldstein, T., Russell, V., Daley, A., 2007. Safe, positive and queering moments in teaching education and schooling: a conceptual framework. Teach. Educ. J. 18 (3), 183–199.
- Gowen, W., Winges-Yanez, N., 2014. Lesbian, gay, bisexual, transgender, queer, and questioning youths' perspectives of inclusive school-based sexuality education. J. Sex. Res. 51 (7), 788–800.
- Greytak, E., Kosciw, J., Boesen, M., 2013. Putting the "T" in "resource": the benefits of LGBT-related school resources for transgender youth. J. LGBT Youth 10 (1–2), 45–63.
- Hookway, C., 2010. Some varieties of epistemic injustice: reflections on Fricker. Episteme 7 (2), 151–163.
- Human Rights Campaign and Gender Spectrum, 2018. *Gender Expansive Youth Report*. <https://assets2.hrc.org/files/assets/resources/GReport1.pdf>.
- Ingre, J.C., 2013. Troubling gender binaries in schools: from sumptuary law to sartorial agency. Discourse 34 (3), 424–438.
- Johnson, E., 2001. "Quare" studies, or (almost) everything I know about queer studies I learned from my grandmother. Text Perform. Q. 21 (1), 1–25.
- Kaufmann, J., 2010. *Trans-representation*. Qual. Inq. 16 (2), 104–115.
- Keegan, C.M., 2020. Against queer theory. TSQ 7 (3), 349–353.
- Kennedy, N., 2018. Prisoners of lexicon: cultural cisgenderism and transgender children. In: Schneider, E., Baltes-Löhr, C. (Eds.), *Normed Children: Effects of Gender and Sex Related Normativity on Childhood and Adolescence*. transcript Verlag, Bielefeld, Germany, pp. 297–312.
- Kennedy, N., 2021. Agentic learning: the pedagogical implications of young *trans* people's online learning strategies. Pedagog. Cult. Soc. 29 (1).
- Kosciw, J.G., Clark, C.M., Truong, N.L., Zongrone, A.D., 2020. *The 2019 National School Climate Survey: The Experiences of Lesbian, Gay, Bisexual, Transgender, and Queer Youth in Our Nation's Schools*. GLSEN, New York.
- Kuhl, D., 2019. *Death of the Clinic: Trans-informing the Clinical Gaze to Counter Epistemic Violence*. Doctoral Thesis. University of Western Ontario. Available at: <https://ir.lib.uwo.ca/etd/6679/>.
- Kull, R.M., Kosciw, J.G., Greytak, E.A., 2015. *From Statehouse to Schoolhouse: Anti-bullying Policy Efforts in U.S. States and School Districts*. GLSEN, New York.
- Kumashiro, K.K., 2002. *Troubling Education: Queer Activism and Anti-oppressive Pedagogy*. Routledge, Leonhardt, New York, NY.
- Lapointe, A., Kassen, J., Snooks, S., Townsley, D., Crooks, C., Dale, S., Hughes, R., 2017. *Healthy Relationships Program for LGBT2Q+ Youth: A Fourth R Small Groups Program to Promote Healthy Relationships and Positive Mental Health*, second ed. Western University, London, ON.
- Lather, P., 1991. Deconstructing/deconstructive inquiry: the politics of knowing and being known. Educ. Theory 41 (2), 153–173.
- Lennon, E., Mistler, B., 2014. Cisgenderism. TSQ 1 (1–2), 63–64.
- Letts, W., Sears, J., 1999. *Queering Elementary Education: Advancing the Dialogue About Sexualities and Schooling*. Rowman & Littlefield, New York.
- Logie, C., Rwigema, M., 2014. "The normative idea of queer is a white person": understanding perceptions of white privilege among lesbian, bisexual, and queer women of color in Toronto, Canada. J. Lesb. Stud. 18, 174–2191.
- Luecke, J., 2011. Working with transgender children and their classmates in pre-adolescence: just be supportive. J. LGBT Youth 8 (2), 116–156.
- Luecke, J.C., 2018. The gender facilitative school: advocating authenticity for gender expansive children in pre-adolescence. Improv. Sch. 21 (3), 269–284.
- Mangin, M., 2020. Transgender students in elementary schools: how supportive principals lead. Educ. Adm. Q. 56 (2), 255–288.
- Martino, W., Cumming-Potvin, W., 2016. Teaching about sexual minorities and "princess boys": a queer and *trans*-infused approach to investigating LGBTQ themed texts in the elementary school classroom. Discourse 37 (6), 807–827.

- Martino, W., Cumming-Potvin, W., 2018. Transgender and gender expansive education research, policy and practice: reflecting on epistemological and ontological possibilities of bodily becoming. *Gen. Educ.* 30 (6), 687–694.
- Martino, W., Cumming-Potvin, W., 2019. "Effeminate arty boys and butch soccer girls": investigating queer and *trans*-affirmative pedagogies under conditions of neoliberal governance. *Res. Pap. Educ.* 34 (2), 131–152.
- Martino, W., Kassen, J., Omercajic, K., 2020. Supporting transgender students in schools: beyond an individualist approach to *trans* inclusion in the education system. *Educ. Rev.* <https://doi.org/10.1080/00131911.2020.1829559>.
- Martino, W., Omercajic, K., Cumming-Potvin, W., 2021. YouTube as a site of desubjugation for *trans* and nonbinary youth: pedagogical potentialities and the limits of whiteness. *Pedagog. Cult. Soc.* 29 (2).
- McCready, L., 2004a. Some challenges facing queer youth programs in urban high schools: racial segregation and de-normalizing whiteness. *J. Gay Lesb. Issues Educ.* 1 (3), 37–51.
- McCready, L., 2004b. Understanding the marginalization of gay and gender non-conforming black male students. *Theory Pract.* 43 (2), 136–143.
- McGlashan, H., Fitzpatrick, K., 2017. LGBTQ youth activism and school: challenging sexuality and gender norms. *Health Educ.* 117 (5), 485–497.
- McGlashan, H., Fitzpatrick, K., 2018. "I use any pronouns, and I'm questioning everything else": transgender youth and the issue of gender pronouns. *Sex. Educ.* 18 (3), 239–252.
- Morgan, E., Taylor, Y., 2019. Dangerous education: the occupational hazards of teaching transgender. *Sociology* 53 (1), 19–35.
- Namaste, V., 2000. *Invisible Lives: The Erasure of Transsexual and Transgendered People*. University of Chicago Press.
- Omercajic, K., Martino, W., 2020. Supporting transgender inclusion and gender diversity in schools: a critical policy analysis. *Front. Sociol.* 5.
- Payne, E.C., Smith, M., 2014. The big freak out: educator fear in response to the presence of transgender elementary school students. *J. Homosex.* 61 (3), 399–418.
- Pinar, W., 1998. *Queer Theory in Education*. Lawrence Erlbaum Associates.
- Porta, C.M., Gower, A.L., Mehus, C.J., Yu, X., Saewyc, E.M., Eisenburg, M.E., 2017. "Kicked out": LGBTQ youths' bathroom experiences and preferences. *J. Adolesc.* 56, 107–112.
- Pyne, J., 2014. Gender independent kids: a paradigm shift in approaches to gender non-conforming children. *Can. J. Hum. Sex.* 23 (1), 1–8.
- Rasmussen, M., Rofes, E., Talburt, S., 2004. *Youth and Sexualities: Pleasure, Subversion, and Insubordination in and Out of Schools*, first ed. Palgrave Macmillan, New York.
- Rasmussen, M., 2001. Queering schools and dangerous knowledges: some new directions in sexualities, pedagogies and schooling. *Discourse* 22 (2), 263–272.
- Rasmussen, M., 2006a. *Becoming Subjects: Sexualities and Secondary Schooling*. Routledge, New York.
- Rasmussen, M., 2006b. Play school, melancholia, and the politics of recognition. *Br. J. Sociol. Educ.* 27 (4), 473–487.
- Riggs, D.W., Bartholomaeus, C., 2018. Cisgenderism and certitude: parents of transgender children negotiating educational contexts. *TSQ* 5 (1), 67–82.
- Rodella, R., Tobin, J.M., 2017. On Anna's terms: supporting a student's gender transition in elementary school. *J. Cases Educ. Leadersh.* 20 (4), 42–57.
- Rubin, G.S., 1984. Thinking sex: notes for a radical theory of the politics of sexuality. In: Vance, C. (Ed.), *Pleasure and Danger*. Routledge & Kegan Paul, New York, pp. 267–319.
- Rubin, H., 1998. Phenomenology as method in *trans* studies. *GLQ* 263–281.
- Schmitz, R.M., Tyler, K.A., Robinson, B.A., 2019. "To get the word out": LGBTQ+ young adults' motivations for participating in qualitative research studies. *J. Gay Lesb. Soc. Serv.* 31 (2), 242–265.
- Sedgwick, E.K., 1991. How to bring your kids up gay. *Soc. Text* 29, 18–27.
- Sedgwick, E.K., 1993/1999. Axiomatic. In: Dering, S. (Ed.), *The Cultural Studies Reader*. Routledge, London, NY, pp. 320–339.
- Serano, J., 2016. *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity*. Seal, Berkeley, CA.
- Sinclair-Palm, J., Gilbert, J., 2018. Naming new realities: supporting *trans* youth in education. *Sex. Educ.* 18 (4), 321–327.
- Sinclair-Palm, J., 2017. "It's non-existent": haunting in *trans* youth narratives about naming. *Occas. Pap. Ser.* 2017 (37).
- Slesaransky-Poe, G., Ruzzi, L., DiMedio, C., Stanley, J., 2013. Is this the right elementary school for my gender nonconforming child? *J. LGBT Youth* 10 (1–2), 29–44.
- Smith, B., 1996. Addressing the delusion of relevance: struggles in connecting educational research and social justice. *Educ. Action Res.* 4 (1), 73–91.
- Snapp, B., Burdge, H., Licona, A.C., Moody, R.L., Russell, S.T., 2015. Students' perspectives on LGBTQ-inclusive curriculum. *Equity Excell. Educ.* 48 (2), 249–265.
- Spade, D., 2015. *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*. South End Press.
- Stryker, S., 2006. (De)subjugated knowledges: an introduction to transgender studies. In: Stryker, S., Whittle, S. (Eds.), *The Transgender Studies Reader*. Taylor & Francis Group, pp. 1–17.
- Taylor, T., Peter, C., 2011. *Every Class in Every School: The First National Climate Survey on Homophobia, Biphobia, and Transphobia in Canadian Schools*. Final Report. Egale Canada Human Rights Trust, Toronto, ON.
- Travers, A., 2018. *The Trans Generation: How Trans Kids (And Their Parents) are Creating a Gender Revolution*. New York University Press, New York.
- Truong, N.L., Zongrone, A.D., Kosciw, J.G., 2020a. *Erasure and Resilience: The Experiences of LGBTQ Students of Color, Black LGBTQ Youth in U.S. Schools*. GLSEN, New York.
- Truong, N.L., Zongrone, A.D., Kosciw, J.G., 2020b. *Erasure and Resilience: The Experiences of LGBTQ Students of Color, Asian American and Pacific Islander LGBTQ Youth in U.S. Schools*. GLSEN, New York.
- Ullman, J., 2014. Ladylike/butch, sporty/dapper: exploring "gender climate" with Australian LGBTQ students using stage-environment fit theory. *Sex. Educ.* 14 (4), 430–443.
- Ullman, J., 2015. *Free2be? Exploring the Schooling Experiences of Australia's Sexuality and Gender Diverse Secondary Students*. Center for Educational Research, School of Education, Western Sydney University, Penrith.
- Umphrey, M.M., 1995. The trouble with Harry Thaw. *Radic. Hist. Rev.* 62, 9–23.
- Warner, M., 1991. Introduction: fear of a queer planet. *Soc. Text* 29, 3–17.
- Wiegman, R., Wilson, E., 2015. Introduction: antinormativity's queer conventions. *Differences* 26 (1), 1–25.
- Wilson, C., Cariola, L.A., 2019. LGBTQI+ youth and mental health: a systematic review of qualitative research. *Adolesc. Res. Rev.* 5 (2), 187–211.
- Wyss, S.E., 2004. "This was my hell": the violence experienced by gender non-conforming youth in US high schools. *Int. J. Qual. Stud. Educ.* 17 (5), 709–730.
- Zongrone, A.D., Truong, N.L., Kosciw, J.G., 2020a. *Erasure and Resilience: The Experiences of LGBTQ Students of Color, Latinx LGBTQ Youth in U.S. Schools*. GLSEN, New York.
- Zongrone, A.D., Truong, N.L., Kosciw, J.G., 2020b. *Erasure and Resilience: The Experiences of LGBTQ Students of Color, Native American, American Indian, and Alaska Native LGBTQ Youth in U.S. Schools*. GLSEN, New York.

Refugee student identities and educational exclusion

Wayne Veck, Department for Education Studies and Liberal Arts, Faculty of Education, University of Winchester, Winchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	169
Otherization and the identities of refugee students	170
Stigma and social attitudes	170
Stigma and refugee student identity	170
Refugee students, assimilation, and identity performance	172
Performing the identity of a native	172
Performing the identity of a deserving refugee	172
Performing the identity of a grateful refugee	173
Performing the identity of a heroic refugee	173
Refugee student identities and “wise” educational spaces	174
Refugee students and the project of answering back to identity projections	174
Teacher identity and the capacity to heal	174
Conclusion	175
References	175

Introduction

What does it mean to be a refugee? Of course, the meaning of a “refugee” has been defined and protected in international law. It was, in fact, some 70 years ago, in Geneva, that the Convention Relating to the Status of Refugees first officially defined a “refugee” as a person who has left their country of origin due to a “well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion” (UN General Assembly, 1951). Immediately, there are points of exclusion here. What, for example, of those persons who find themselves persecuted for reasons of sexuality? And if we drill down more precisely into the terms that frame this definition, then it becomes apparent, as Carens (2003, p. 102) notes, that there is hardly one that is not “subject to interpretation and contestation.” Equally, to be a refugee is to be among the 82.4 million people who, by the end of 2020, had been forced to flee from the places that had once been their homes (UNHCR, 2021). More precisely, it is to be among the 20.7 million persons from across the world who, by the end of the same year, were (under the mandate of the UN High Commissioner for Refugees) counted as refugees (UNHCR, 2021).

But beyond the statistics, and the label itself, how is its meaning experienced by those uprooted persons who, granted refuge in some foreign land, come to be named “refugees”? Hannah Arendt, in her intensely incisive essay, “We Refugees,” observes that Jews, fleeing Nazi persecution, “don’t like to be called ‘refugees’, and so call each other ‘newcomers’ or ‘immigrants’” (Arendt, 2007a, p. 264). While Viet Thanh Nguyen, who, in 1975, came with his family to the US as child refugee from Vietnam, insists on being “called a refugee” despite the fact that it would be “easier” for him to be known as “an immigrant, to pass ... as belonging to a category of migratory humanity that is less controversial, less demanding, and less threatening than the refugee” (2018, p. 11). Vu Tran, who was also forced to flee from Vietnam as a young child, before finding refuge, along with his mother and sister, in the US, recalls that when he was younger he saw himself as an “immigrant,” since a “refugee felt like a thing of the past, a provisional status ... that suggested vulnerability, inferiority, alienness—everything I wanted to remove from my idea of myself” (original emphasis, 2018, p. 151). Alongside these tensions, there is, of course, further complexity here, since it is better to be defined as a refugee than an asylum-seeker, an asylum-seeker being a person who has yet to be conferred the status of refugee (UN High Commissioner for Refugees, 2006).

What is clear, however, is the fact that whether a person comes to be categorized as an immigrant, an asylum-seeker, or a refugee, there are real consequences for both their social and personal identity. And it is precisely to the question of identity that this chapter attends as it examines the exclusion, along with possibilities for the inclusion, of refugee students. Hostility to immigrants and refugees, as it is manifested in images and slogans that permeate social media, in endless degrading newspaper headlines, and in the discourse of politicians who seek to drum up votes by stirring up fear, distorts perceptions of these students’ identities. But, more than this, such hostility can diminish the very capacity of these students to create identities that disclose their individual distinctiveness.

This chapter opens by turning to Goffman’s (1990a) account of stigma to illustrate the social dimensions of refugee student exclusion. The chapter then examines the critical connection between assimilation and identity formation in the lives of refugee students. Here the chapter argues that as refugee students attempt to fit into preestablished ways of acting, speaking and appearing, so they engage in *identity performances* with the aim of proving to their audiences that they have been thoroughly assimilated.

The chapter concludes by turning to the work of Maxine Greene to examine what it might mean to create educational spaces where uprooted and displaced young people can come to create and divulge their distinctive identities.

Otherization and the identities of refugee students

In his distinct conception of stigma, the American sociologist, Goffman (1990a), illuminates the complex web of social interactions that condition our expectations of others. Indeed, Goffman (1990a) suggests that we rarely encounter another person for the first time with hands that are entirely open, and with gazes that are wholly attentive, to the distinctiveness of the stranger's words, deeds, and appearance. Goffman's insights into stigma contains two quite distinct, but equally significant, implications for our understanding of the exclusion of refugee student identities. First, they suggest that if we are to fully appreciate the reduction and exclusion of these students' identities, then before we attend to their individual attributes, we should first examine ruling social attitudes and prejudices. Second, they reveal the ways social attitudes can diminish the ways refugee students might come to understand and, critically, forge their identities.

Stigma and social attitudes

Goffman (1990a, p. 12) observes that each of us carries within us a set of "anticipations," which are often transformed into "normative expectations," and then into "righteously presented demands." However, we are often quite unaware of what these demands are, unaware perhaps even of the very fact that we possess them. In fact, it is not until a stranger enters into our midst who does not fit the social stereotype that we are poised to project onto them, that a distance can open between this anticipated identity (or what Goffman names a "virtual social identity") and this person's "actual social identity" (1990a, 12). Here what we expected (often without our ever being consciously aware of this expectation) is disrupted by reality. This is the moment of stigma, the moment at which we realize that another person has "an undesired differentness from what we had anticipated" (Goffman, 1990a, p. 15). Thus, before we can begin to examine stigmas, "a language of relationships, not attributes, is really needed" (Goffman, 1990a, p. 13). In other words, stigma occurs, not because of some difference or characteristic that emphatically marks out the stigmatized person from others, but because a tension has emerged between our expectations and our experience (Goffman, 1990a; Uba and Nwoga, 2016). Stigma, in short, is a social and not merely an individual issue.

Goffman's understanding of the social dimensions of stigma thus directs our attention to the ways *all* of us might, at any given time, be harboring a whole host of delimiting social expectations. Immediately, this alerts us to our responsibility to awaken ourselves and others to the hostile attitudes and prejudices we might be unreflectively nursing. In relation to the stigmatization and exclusion of refugee students, this entails avoiding a merely individualized approach to their education. In particular, it involves examining how social relations in and beyond our schools can leave persons alienated from each other, their interactions mediated through predetermined and prescribed social prejudices or stereotypes. This, in turn, involves addressing the ways practices, values, and cultures in a school can ameliorate or compound the trauma that these students may experience, and support or hinder them as they begin to articulate themselves in a language that might be entirely new to them.

Stigma and refugee student identity

Goffman tells us that a person's shifting sense of identity and self-worth forms a kind of "moral career," the dimensions of which are shaped by how they are perceived by others (1990a, p. 45). In a particularly telling example, Goffman describes the unfolding of a boy's "moral career" as he moves from a mainstream to a special school. Goffman observes that this child's self-conception will alter, that he will have to accept "the view which the public at large takes of him," as he comes to "learn that the own he thought he possessed was the wrong one, and that this lesser own is really his" (1990a, p. 47). Further insight is offered by the cultural theorist, Stuart Hall, who wrote in 1959 of how children in the UK. were separated from their peers by an examination they undertook at aged 11. This examination determined if these children were to attend a Secondary Modern, and not a Grammar School. Hall reflected that by the time they reach the age of 13 or 14, children attending Secondary Moderns, "begin to reflect upon their own sense of failure" (p. 18). Hall explains that these children, acutely aware of "their second-class status," along with their teachers' "lack of sympathy and understanding," ultimately "become ... what many teachers have always believed they were: unteachable, unclubbable" (1959, p. 18). What Goffman and Hall convey here is the fragility of so many identities, the way they are constantly susceptible to the fluctuating attitudes and judgments of those around them. As Hall notes, "identities actually come from outside, they are the way in which we are recognized and then come to step into the place of the recognitions which others give us" (1995, p. 8).

Drawing directly on Goffman's understanding of stigma, Eren and Çavuşoğlu (2021, p. 16) offer an account of the ways Syrian students in Turkish schools can be designated to "the role of outsiders in their mainstream classrooms." While refugee writers have reflected on their own experiences of being stigmatized and marginalized during their school days. Dina Nayeri, for example, in her memoir, *The Ungrateful Refugee*, writes of her escape from Iran as a child, before finally finding refuge in America. Nayeri tells the story of how, far ahead of the rest of her mathematics group in her Oklahoma school, she took out her sewing kit so she might engage in a more productive activity. Nayeri's teacher asked her to stop sewing, and then said: "In this country, girls have to learn

math just like boys" (Nayeri, 2020, p. 203). On another occasion, when a piece of a papier-mâché map she was constructing broke off, one of Nayeri's teachers said to her: "Don't be lazy. We work hard here, do you understand" (2020, p. 198). The writer, Reyna Grande, describes how she, having been abandoned by her father and then her mother, became, at the age of nine, "an 'illegal' human being," when she crossed the Mexico-U.S. border on the back of her father who had returned for her (2018, p. 73). Grande recalls that her schooling in the US was "almost as traumatic as being abandoned" by her parents, since it saw her "put at a corner table and ignored" (2018, p. 78). Left alone, and feeling "voiceless," Grande reflects that she was "on the outside looking in, marginalized, excluded, 'othered'" (2018, p. 78). The stigmatization of the Syrian students observed by Eren and Çavuşoğlu (2021), along with Nayeri and Grande's lived experience of stigma tell us nothing about the "actual identities" of refugees and much about the unexamined prejudices and stereotypes that permeate societies and influence social attitudes. Indeed, these experiences stream from and confirm over and over again the same inescapable hostility, the same unending refusal to welcome refugees, the fact that they "are mostly unwanted where they fled from; unwanted where they are, in refugee camps; and unwanted where they want to go" (Nguyen, 2018, p. 17).

The otherization of immigrant and refugee students can, however, occur in more subtle and ambiguous ways. As a young man, Stewart Hall was a supply teacher in a school in London. He soon became aware of growing racism and antagonism in Notting Hill that would erupt in the violent riots of 1958. A moment of stark illumination occurred when Hall addressed the white children he taught and asked them to provide an account of their hostility to newly arrived migrants from the Caribbean. He was told, "Well, they're taking our women," and, "They're taking our things" (Hall and Back, 2009, p. 672). Hall recalls:

So I said, "Do you mean these?" And I pointed to several black kids in the class and they looked at them as if they'd never seen them.

The white children retorted, "No sir," before expounding: "They're one of us" (Hall and Back, 2009, p. 672). Hall asked, "What about me?," only for the response to echo what he had already heard: "No sir. Not you. Them" (Hall and Back, 2009, p. 672). How many migrant and refugee students hear similar words, hear, for example, "They are taking our jobs, our benefits, but not you, you are different"? Consider, for example, a story told by Musa Okwonga, the son of Ugandan refugees forced to flee from the rule of Idi Amin to seek sanctuary in the UK. Okwonga describes sharing a meal with a friend, who suddenly "launched into an astonishing rant against migrants" (2020). When Okwonga pointed out that he and his parents "were no different" from the people his friend "was denigrating," he was told: "I don't see you as a migrant, Musa. I see you as a friend" (2020).

What does it mean to be counted as "one of us," knowing that at any moment one might be counted among the "them," the others? How must it feel to be a displaced person listening to their peers, students, or friends as they disparage migrants or refugees? Once again, Goffman's account of stigma is illuminating. Stigma, Goffman observes, acts upon an individual to exclude them both in their relations with others and, crucially, in their relationship with themselves, it "spoils his social identity; it has the effect of cutting him off from society and from himself so that he stands a discredited person facing an unaccepting world" (Goffman, 1990a, p. 31). And yet, alongside such "discredited" persons, there are, also, those people who live in the knowledge that they are, in Goffman's (1990a) language, "discreditable." Something of the experience of being included in an educational setting, while all the time knowing that hostility might be directed one's way, that some social prejudice or stereotype might thoughtlessly be given voice in one's presence, is conveyed by Joseph Azam. Writing of his experiences of growing up in New York in the 1980s, Azam (2018, p. 29) recalls how children from families like his, families that "had recently come from some distant place in search of a better life," could be identified as different by the clothes they wore, their presence as informal translators at parent-teacher meetings, and even by their very names. And so, they "stood out and were each vulnerable" in their "own way" (Azam, 2018, p. 29). His own vulnerability, Azam (2018, p. 30) tells us, necessitated an "effort ... into curating" his "identity," to "managing" his "otherness."

There are thus many possible dimensions to a refugee student's experience of displacement. Of course, there is the original displacement from their native country, but this original dislocation can be followed by a seemingly permanent experience of uprootedness, of never quite belonging. Alongside this experience, there may be a displacement of a refugee student's identity, a jarring sense that their identity can be neither cultivated nor celebrated within the tapestry of their everyday social interactions, but must, instead, as Azam (2018) stresses, be managed. No wonder, then, that Tran describes how, "on the inside, away from the eyes of those who see her only as a refugee," a displaced person, "especially if she arrived at this status as a child, ... will go on to feel as though she is not whole in some way" (2018, p. 153).

What is missing when a refugee student's identity becomes displaced in this way is the kind of "wise" attention that Goffman (1990a, p. 41) detects in those people "intimately privy to the secret life of the stigmatized individual and sympathetic with it." Indeed, among those persons who are not stigmatized, Goffman (1990a) draws out a clear distinction between those he names the "normals" and those he calls "wise." In contradistinction to the "normals," cut adrift from the stigmatized, Goffman (1990a, p. 41) characterizes the "wise" as "marginal" persons, "before whom the individual with a fault need feel no shame nor exert self-control, knowing that in spite of his failing he will be seen as an ordinary other." In relation to refugees, to be "wise" is to see the person forced to migrate from their native land as something more than a pitiable or suspicious other. It is, instead, to encounter this person in the unfolding of their distinctive and "ordinary" identity. It is precisely in this way that we might begin, in our schools and in our communities more generally, begin to counter the "bigotry directed at migrants and refugees" that is

“contingent upon their dehumanization and deindividualization” and which sees them “presented and thought of as a mass of nothings and nobodies,” in short, as “a faceless mass” (Hemon, 2018, p. 92).

Refugee students, assimilation, and identity performance

In his analysis of the lives of exiles, Zygmunt Bauman traces out how losing a place in the world leads to a loss of identity and then, finally, to a quest for acceptance by way of assimilation:

What makes the exile an unreal place is the daily effort to make it real—that is, to cleanse it of all things that are out of place. In exile, one is pressed to stop being in exile; either by moving elsewhere, or by dissolving into the place, not being anymore out of it. The latter is the pressure of assimilation. To be like anyone else. Not to be odd anymore—to renounce one’s nonidentity, that is, to renounce one’s identity in the name of a new identity, which would be nonidentity.

Bauman (1996, pp. 569–570).

Assimilation is, as Bauman suggests here, a hopelessly self-refuting process, an unending seeking of a place in the world, an unceasing quest to supplant one’s own identity with some accepted identity. We turn now to examine four paths that promise assimilation for refugee students, each of which is rather like that trodden by Sisyphus, condemned by the god Zeus to eternally push a boulder to the summit of a hill only for it to descend back down to lower ground.

Performing the identity of a native

Arendt (2007a, p. 265), writing of Jewish refugees in 1943, observed that “after four weeks in France or six weeks in America, we pretended to be Frenchmen or Americans.” Arendt goes on to note that for “the more optimistic among us” it is “only their new country” that has “taught them what a home really looks like” (2007a, p. 265). And yet, Arendt (2007a, p. 272) recognizes that “under the cover of ... ‘optimism’ you can easily detect the hopeless sadness of assimilationists.” Here the task of forging an identity has become an extended exercise in what Goffman (1990a) names “passing,” where one attempts to become something one does not entirely feel oneself. This task can also involve, in Goffman’s (1990a) phrase, “covering” up any attributes that might negatively distinguish one from others. However, whenever a refugee is able to successfully “pass” as a native or adequately “cover” up any of their distinguishing characteristics from them, their achievement is, and only can amount to, the *presentation*, and not the *disclosure*, of an identity. Thus, Tran notes that “a refugee is ... like an actor,” before explaining:

She knows she must assimilate in order to earn acceptance, to survive and prosper, but the act of assimilation is inevitably a performance—not of a false or superficial identity necessarily, but of one more legible to her audience. She becomes versions of herself, donning the new clothes and accessories, the new habits and modes of behavior that hew to the various settings of her new life (2018, p. 153).

A performed identity might not be “false or superficial,” and it might be argued that all identity is a form of presentation (see, for example, Goffman’s (1990b) *Presentation of the Self in Everyday Life*), but the performance or presentation Tran (2018) describes here is delivered to some extent in the negative and thus cannot be entirely affirmative.

Something of the complexity that can characterize the performance of an identity by refugee students is expressed by Okwonga, whose parents were Ugandan refugees. Okwonga describes how, at the private boys’ preparatory school he was given a bursary to attend, he “became an unofficial ambassador for black people” (2020). He goes on to recall how he studied with a “furious sense of mission” in the belief that, by making “a good impression,” he “could help to erode some of our society’s firmest prejudices” (Okwonga, 2020). Later, at the private school, Eton, Okwonga (2020) was to learn of the impossibility of changing some “perceptions of black people merely by being as hardworking and as agreeable as possible.” Such experiences tell us much about how, for refugee students, achievement in school can take on a double aspect: it can be both a way of confirming something positive about their individual identity and a way of attempting to refute social prejudice. Indeed, identity performances always aim to “pass” and not simply to disclose; to “cover” and not wholeheartedly to reveal. Hence, even before the student’s identity performance begins, its success can be determined, not only by what it shows to its audience, but also by what it manages to hide from them. In other words, the performance of an identity is underscored by the alienation of the performer both from their audience and from themselves.

Performing the identity of a deserving refugee

In 1944, Arendt wrote of how, despite the limited number of refugees in the US, it was always possible to hear “at least one parliamentarian asking in baffled tones just how far one intends to go trading ‘fine young Americans for refugees,’ and at least one widely read journalist ready to discredit them with the platitude that ‘not all refugees have been ennobled by their sufferings’” (2007b, pp. 211–12). In the intervening time what has changed? It is worth reflecting on these words, spoken by the Former British Prime Minister, Theresa May, when she was the UK’s Home Secretary, to a BBC Radio 4 interviewer: “People talk about refugees, but

actually if you look at those crossing the central Mediterranean, the largest number of people are ... economic migrants" (May, cited in Kingsley, 2016, pp. 41–42). Consider also that when he was the UK's Prime Minister, David Cameron, spoke of "a swarm of people coming across the Mediterranean, seeking a better life, wanting to come to Britain because Britain has got jobs, it's got a growing economy, it's an incredible place to live" (cited in BBC News, 2015). Consider, also, the words of the UK's current Prime Minister, Boris Johnson, who described crossing the English Channel to seek asylum in the UK, as "a very bad and stupid and dangerous and criminal thing to do" (cited in Grierson and Sabbagh, 2020). So, politicians in the UK and elsewhere continue to advance a firm distinction between those small number of worthy refugees from the majority of unworthy, illegal migrants. The inference for the refugee is clear, it is their duty to establish their legitimacy. Hence Nayeri tells us: "Every day of her new life, the refugee is asked to differentiate herself from the opportunist, the *economic migrant*" (original emphasis, 2020, p. 8). And there remain journalists equally zealous in advancing their pronouncements of the unworthiness of some, if not all, refugees. Consider, for example, the columnist, writing in the UK's biggest selling daily newspaper, who described refugees as "cockroaches" (Hopkins, 2015). The expression is extraordinarily degrading, of course, but all the time refugees are held to be irredeemably suspect they are, like all stigmatized persons, believed to be, in Goffman's (1990a, p. 15) phrase, "not quite human." Indeed, Nguyen (2018, p. 15) echoes Goffman when he observes that a refugee, "classified as an *other*," is—if they are seen at all—viewed as "not truly human ... but is instead a stereotype, a joke, or a horror."

Immediately, we might note, with Carens (2003, p. 102), the impossibility of determining a "single bright line dividing those who deserve asylum from those who do not." But all the while the division retains its power to exclude, refugees, in their efforts to shun the stigma of being perceived as an "economic migrant," fearing the accusation of mediocrity, will continually struggle to excel as they seek to generate "proof that letting in refugees has a good, healthy return on investment" (Nayeri, 2018, p. 146). For a young refugee, one way of distancing themselves from the category of "unworthy migrant" is to perform the identity of a brilliant and utterly dedicated student. Hence, Nayeri describes her own exceptional efforts to dedicate herself to her school studies as she endeavored to ensure that she did not "lose everything," her "entire identity as a smart, capable girl" (2020, p. 293). The performance of the devoted student is not, of course, devoid of disclosure—it exposes the refugee student's need to fit in, to "pass," and in this way, it reminds us of the profundity of the human need to belong. But the performance can reveal only what the refugee has been deemed to lack and what they long to disprove. Where is there space here for the emergence of a love of learning that is its own reward, and not a marker of assimilation?

Performing the identity of a grateful refugee

Nayeri (2020, p. 197) recalls how, in the Oklahoma school she attended, her success in producing a papier-mâché topographical map of the United States came to be "strangely central" to her teacher's efforts to support her "American assimilation." When Nayeri explained that just a few months previously she had been a refugee, living in Italy, the teacher responded "sleepily," with the words: "Awww, sweetie, you must be so grateful to be here" (2020, p. 197). Latent in this teacher's word, "gratitude," there is both a judgment and a demand. Indeed, throughout her memoir, Nayeri (2020) conveys the burden of gratitude that can be carried by refugee students. The refugee student, she suggests, can feel an almost constant pressure to demonstrate that they are glad, not only to be here, in this new land, in this new educational setting, but also that they are grateful *not* to be there, in their native land, which can be afforded no redeeming, and certainly no delightful, qualities. Such assimilation comes at a cost, it means repudiating the memories and experiences that together inform a young person's identity and enrich their hopes for the future. Already displaced from the place that was once their home, the young person is asked to distance themselves, also, from their memories of this place. Here again, it is possible to detect the complexities and multidimensional aspects of displacement in the lives of refugee students.

Performing the identity of a heroic refugee

In 2018, an extraordinary story was reported around the world about a young man named Mamadou Gassama (see, for example, Hughes, 2018). Walking through a northern Paris suburb, Gassama noticed a crowd staring up at a tiny boy hanging precariously from a fourth-floor apartment balcony. In an act of astonishing athleticism, Gassama scaled the building, moving from one balcony to the next, before pulling the child to safety. In time it transpired that Gassama was an economic migrant from Mali, with no legal status in France. And yet, this remarkable act led Emmanuel Macron, the French President, to bestow Gassama with French citizenship, along with a medal of honor in recognition of his courage.

It is possible to feel awe at Gassama's feat, to recognize and esteem the qualities of character that made his deed possible, and at the same time to be troubled by the idea that only supreme and astonishing accomplishments can transform newcomers from illegal strangers to good citizens.

This troubling sensation might be understood in direct relation to the double exclusion that occurs whenever a refugee is separated out from "economic migrants." First, there is the exclusion of this isolated individual. As others retell the story of their deed or deeds, there is always the possibility of sentimentalization and sensationalism and, worse, of positioning the refugee as a passive beneficiary of their hosts' magnanimity. Hence, Nayeri (2018, p. 146) asks aptly: "isn't glorifying the refugees who thrive according to western standards just another way to endorse gratitude politics?" The irony here is bitter: even as a refugee is elevated to the status of hero, so they are reduced to a mere means for the confirmation of the goodness of their host. It is possible to detect this same process in schools. Consider, for example, the situation in which a once excluded or marginalized young person's success in meeting predetermined academic standards is taken to be proof of the tireless endeavors of their teachers and parents. Second,

there are exclusionary consequences for refugees and those persons who are regarded as “economic” migrants. Of course, for all the “tales of heroes of assimilation who have penetrated new areas of normal acceptance” (Goffman, 1990a, p. 37), most refugees and migrants will find “passing” for a native an endless journey with no destination, a process of incessantly assimilating without ever becoming assimilated. Indeed, exceptions to the rule, like Gassama, are so very exceptional that stories of their assimilation serve to fortify the status of the refugee as strange, entirely other. More than this, the majority of people fleeing from persecution can always be judged in relation to this image of the heroic refugee, a judgment that condemns them for the very ordinariness that Goffman assures us “wise” people look to when they respond to another human.

Refugee student identities and “wise” educational spaces

Nowhere is “wise” attention more needed than in the education of refugee students. So, we turn now to the educational philosophy of Maxine Greene to imagine possibilities for an education of refugee students that is freed of all expectation of assimilation, worthiness, and heroic commitment.

Refugee students and the project of answering back to identity projections

Greene has acknowledged the debt her own educational thinking owes to Sartre’s (1963, 2007) characterization of human existence as “a project.” This idea emerges out of the central tenet of Sartre’s existentialist philosophy. Each of us, Sartre (2007, p. 23) argues, “is indeed a project that has a subjective existence,” before which “nothing exists.” In other words, it is the task of each person to determine the nature of their own existence, since, beyond that which we create of ourselves, “there is no human nature” (Sartre, 2007, p. 22). In fact, the distinct identity of each individual may be thought of as an ongoing outcome of this “project.” And while Sartre maintains that it is entirely up to each of us to shape this project for ourselves, we cannot wish it away. We can choose not to choose, but we cannot choose not to have a choice.

Reflecting on Sartre’s existential philosophy, Greene writes: “We may be moved to choose our project, because ... we are trying to become what we are not yet by acting on perceived deficiency, or on perceived possibility” (1978, p. 26). Acting on “perceived possibility,” one’s identity might be conceived as an ongoing and active attempt to answer a question that can never be definitively or satisfactorily answered, the question of who one is or at least of who one wishes one could be. But if we act under the shadow of one “perceived deficiency” or another, we act against what we are not, and the project of identity formation is supplanted by the task of answering back to identity projections cast our way by others. This necessarily negative project is firmly anchored in the past. The construction of an identity in opposition and resistance to a tirade of stereotypes and degrading expectations is, in other words, quite different from that ongoing process of forming an identity that affirms and even celebrates a person’s distinctiveness.

Nayeri (2020, p. 12) tells us that, on gaining refuge in America, she was assured that her life was her own to determine, that her “gender would no longer limit” her distinctive “potential,” and yet, she came to realize that “already a limit” had been “imposed” on her life, since her “first category had been assigned: refugee, not native-born.” With these words Nayeri encapsulates the violence inherent in the act of naming the identity of another person, in attempting to strike from their hands the chance to answer for themselves those existential questions pertaining to who they wish to be and who they wish to become. Interestingly, the subheading of Goffman’s (1990a) book *Stigma* is *Notes on the Management of a Spoiled Identity*, and the identities of refugee students may be spoiled, not only when one reductive identity or another has been imposed upon them, but when the very project of creating an identity of whatever kind has been disrupted.

Teacher identity and the capacity to heal

In the face of such violence and exclusion, it is both heartening and useful to reflect on the hope expressed by Greene that we might “see more and more teachers willing to choose themselves as healers” (1993, p. 215). Of course, teachers who take responsibility for healing those refugee students whose identities have been spoiled do so by reaching out to, and working in unison with, other agencies and experts in various fields, including trauma, law, and language acquisition. And yet, these teachers can themselves make a considerable difference simply by choosing to be “wise” in their encounters with refugee students. Indeed, Ron Baker’s reflections on his experiences as a child transported to England from Nazi Germany have much to teach us about the necessity and effectiveness of such “wise” teachers. Baker tells us that “the process of having one’s roots disturbed and rearranged is not an easy one” (1990, p. 64), before going on to give voice to a refugee child’s “sense of isolation” and their “gnawing aching need for acceptance, love and security” (p. 65). Grande’s conclusions, born of her own experiences as a migrant student in the US, are equally as instructive. “Nothing is more counterproductive than the goal of rapid integration of immigrant and refugee children into the community,” Grande insists, before noting:

Assimilation and acculturation add to their post-traumatic stress. Immigrant and refugee children need time, patience, love, and psychological help to heal from their experiences (2018, p. 78).

These moving articulations of those qualities of acceptance, security, patience, and love that can be so needed by refugee students remind us of the extraordinary qualities that teachers can and do bring into the ordinary lives of their classrooms.

To suggest, with Greene (1993), that teachers might choose to heal the identities of refugee students is not to deny that often “identities are constructed within, not outside, discourse,” that they can be formed “within the play of specific modalities of power,” and that they can, in short, be “the product of the marking of difference and exclusion” (Hall, 2003, p. 4). It is, rather, to acknowledge the reality that identities can also be—and indeed often are—formed in educational settings where the discourse of teachers is, in Foucault’s phrase, “a hindrance, a stumbling-block, a point of resistance” to those reductive and otherizing discourses that marginalize and silence refugee students (1990, p. 101). It is to acknowledge, that is, those teachers who engage in the necessary and ongoing work of examining how they might, in Greene’s words, “keep pondering the meanings of inclusion and wondering how it can occur without the kind of normalization that wipes out differences” (1993, p. 212). Noting how such teachers “suffer in many ways what they experience as conditioning or manipulation by their superiors or by the ‘system’ itself,” Greene goes on to argue that it is exactly the reflective capacity of these teachers that enables them to recognize these exclusionary factors as “forces that frustrate their quests for themselves and their efforts to create themselves as the teachers they want to be” (1978, p. 34). Such forces and pressures are frustrating to teachers, in other words, precisely because they experience them as hinderances to the advancement of their identities. Indeed, it might be that this very identity, at once professional and personal, fortifies and bolsters these teachers in their daily efforts to block narrow and quantitative targets, league tables, prescribed curricula, and shallow expectations from fully impeding upon, and thus diminishing and delimiting, themselves, their teaching, and the children they teach. The inclusion or exclusion of children receiving or seeking refuge turns at least in part, then, on how teachers live out their answers to existential questions concerning the sorts of schools they wish to create and the types of teachers they wish to become.

Conclusion

What would it mean for refugee students to forge their own identities? This chapter has called for critical attention to be given, not only to what has happened to and what is happening in, the lives of refugee students, but also to the social attitudes that can act upon these students to diminish their identities. This involves answering the situation in which refugees, despite the best of their efforts, can never quite succeed in “passing” as natives, can never sufficiently prove their merit, their gratitude, or their heroic qualities, and so pass from “one of them” to “one of us.”

Following Goffman’s (1990a) account of those “wise” persons who, in opposition to the so-called “normal” individual, conceive the stigmatized beyond their stigma, this chapter has called for the recognition and rejection of dominant narratives of “successful” or “included” refugee students. This necessitates becoming acutely aware of “ways of using language that lead to the ... identification of certain kinds of being as undesirable” (Greene, 1993, p. 216). In particular, it means attending to ways of talking about dispossessed and displaced people that can generate images of “worthy” and “unworthy” students and so result in the “demeaning of poor and immigrant people’s experience” (Greene, 1995, p. 51). A “wise” education, the chapter has suggested, resists this distinction between the deserving and the undeserving immigrant, along with the more general assumption that while some students are worthy of a teacher’s dedication, others are “disaffected,” “lazy,” “disruptive” and so on.

It is “a writer’s dream,” Nguyen (2018, p. 22) contends, “that if only we can hear these people that no one else wants to hear, then perhaps we can make you hear them, too.” Might this not be the dream of teachers, also? The teacher who listens to young people who have been uprooted from their homes, to include them not only as they speak, but also to bolster their growing sense of distinctiveness so they might come to risk speaking out to the world. The necessity of listening and thinking beyond narratives of assimilation, of becoming aware of the various ways they can, in Goffman’s (1990a) crucial phrase, “spoil” the identities of refugee students, informed this chapter’s argument for educational spaces where teachers recognize and act upon what Greene (1993) understands as their capacity to heal. This means generating dialogs of becoming, and not assimilation, in our schools; it means teachers realizing their capability to create spaces that, again following Goffman (1990a), might be thought of as “wise.” Within these spaces, refugee children can contribute to ongoing dialogs without confronting the kind of unthinking expectations that Goffman (1990a) alerted us to. Crucially these are educational spaces where dialogs flourish between people who do not have to prove themselves worthy, since their worth, as human beings and as contributors to an ongoing and expansive conversation, is presupposed from the outset.

References

- Arendt, H., 2007a. We refugees. In: Kahn, J., Feldman, R.H. (Eds.), *The Jewish Writings*. Schocken Books, New York, pp. 264–274.
- Arendt, H., 2007b. The political organization of the Jewish people: articles from *Aufbau*, April 1944–April 1945. In: Kahn, J., Feldman, R.H. (Eds.), *The Jewish Writings*. Schocken Books, New York, pp. 199–240.
- Azam, J., 2018. Last, first, middle. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, pp. 23–33.
- Baker, R., 1990. The refugee experience: communication and stress, recollections of a refugee survivor. *J. Refug. Stud.* 3 (1), 64–71.
- Bauman, Z., 1996. Assimilation into exile: the jew as a polish writer. *Poetics Today* 17 (4), 569–597.
- BBC News, 2015. David Cameron criticised over migrant “swarm” language. London BBC News. <http://www.bbc.co.uk/news/uk-politics-33716501>. (Accessed 31 July 2016).
- Carens, J.H., 2003. Who should get in? The ethics of immigration admissions. *Ethics Int. Aff.* 17 (1), 95–110.
- Eren, A., Çavuşoğlu, C., 2021. Stigmatization and othering: the case of Syrian students in Turkish schools. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2021.1900422>.

- Foucault, M., 1990. *The History of Sexuality*, vol. 1. Penguin, Harmondsworth.
- Goffman, E., 1990a. *Stigma: Notes on the Management of Spoiled Identity*. Penguin, London.
- Goffman, E., 1990b. *Presentation of the Self in Everyday Life*. Penguin, London.
- Grande, B., 2018. The parent who stays. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, 63–33.
- Greene, M., 1978. Teaching: the question of personal reality. *Teach. Coll. Rec.* 80 (1), 23–35.
- Greene, M., 1993. Diversity and inclusion: towards a curriculum for human beings. *Teach. Coll. Rec.* 95 (2), 213–221.
- Greene, M., 1995. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*. Jossey-Bass Publishers, San Francisco.
- Grierson, J., Sabbagh, D., 2020. Boris Johnson accused of scapegoating migrants over channel comments. *Guardian*. <https://www.theguardian.com/uk-news/2020/aug/10/boris-johnson-hints-at-law-change-to-deport-migrants-who-cross-channel>. (Accessed 14 July 2021).
- Hall, S., Back, L., 2009. At home and not at home: Stuart Hall in conversation with Les Back. *Cult. Stud.* 23 (4), 658–687.
- Hall, S., 1959. Absolute beginnings: reflections on the secondary modern generation. *Univ. Left Rev.* 7, 16–25.
- Hall, S., 1995. Negotiating Caribbean identities. *N. Left Rev.* 209 (Jan–Feb), 3–14.
- Hall, S., 2003. Introduction: who needs “identity”? In: Hall, S., du Gay, P. (Eds.), *Questions of Cultural Identity*. SAGE Publications, London, pp. 1–17.
- Hemon, A., 2018. God’s fate. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, pp. 91–104.
- Hopkins, K., 2015. Rescue boats? I’d use gunships to stop migrants. *Sun* 11.
- Hughes, R., 2018. The spiderman of Paris: what happened next? *BBC News*. Available. <https://www.bbc.co.uk/news/world-europe-46538253>. (Accessed 15 July 2021).
- Kingsley, P., 2016. *The New Odyssey: The Story of Europe’s Refugee Crisis*. Faber & Faber, London.
- Nayeri, D., 2018. The ungrateful refugee. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, pp. 137–150.
- Nayeri, D., 2020. *The Ungrateful Refugee: What Immigrants Never Tell You*. Canongate, Edinburgh.
- Nguyen, V.T., 2018. Introduction. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, pp. 11–20.
- Okwonga, M., 2020. The ungrateful country. *Byline Times*. <https://bylinetimes.com/2020/04/10/the-ungrateful-country/>. (Accessed 14 July 2021).
- Sartre, J.-P., 1963. *Search for a Method*. Alfred A. Knopf, New York.
- Sartre, J.-P., 2007. Existentialism is a Humanism. Trans by C. Macomber, J. Kulka (Eds.). Yale University Press, New Haven.
- Tran, V., 2018. A refugee again. In: Nguyen, V.T. (Ed.), *The Displaced: Refugee Writers on Refugee Lives*. Abrams, New York, pp. 151–157.
- Uba, C.D., Nwoga, K.A., 2016. Understanding stigma from a sociocultural context: mothers’ experience of stigma directed towards children with special educational needs. *Int. J. Incl. Educ.* 20 (9), 975–994.
- UN General Assembly, 1951. Convention relating to the status of refugees, 28 July 1951. U. N. Treaty Ser. 189. <https://www.refworld.org/docid/3be01b964.html>. (Accessed 16 January 2019).
- UN High Commissioner for Refugees, 2006. *UNHCR Master Glossary of Terms*. June 2006, Rev.1. <https://www.refworld.org/docid/42ce7d444.html>. (Accessed 16 January 2019).
- UNHCR, 2021. *Figures at a Glance*. <https://www.unhcr.org/figures-at-a-glance.html>. (Accessed 20 July 2021).

Educational trends and exclusion

David Armstrong, School of Education, RMIT University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	177
What are trends? A definition of educational trends	177
What are fads? A definition of educational fads	178
Performance enhancing trends	179
Technical trends	179
Crumbling orthodoxy	179
Cures, therapy and hope	179
Trends of the near future	180
Educational trends and exclusion	180
Conclusion	181
References	181

Introduction

Educational trends leave their mark on classrooms and on education systems across the world. Educational initiatives or programs based on: mindfulness, growth mindset, belonging, positive behavior interventions and supports (PBIS), for instance, have been implemented across tens of thousands of education settings internationally, often, but not always originating in the US PBIS, alone, claims to have been adopted in 25,000 US schools by 2018 (Kittelman et al., 2019) and is increasingly used in schools in Australia, in England and in many European countries (Poed and Whitfield, 2020). Trends rise and gain influence across school systems in relatively short periods of time. Mindfulness, growth mindset and belonging, along with other prominent, trends were relatively unknown in the education sector, in the English-speaking world at least, before 2010 (Armstrong and Armstrong, 2021).

Despite their prominence in the classroom, comprehensive critical scrutiny of trends is rare (Adey and Dillon, 2012) – a strange absence given the thousands of studies which evaluate the benefits of individual educational initiatives, programs or commercially branded interventions. If these phenomena are recognized as trends at all, existing educational research often takes a limited perspective. Research studies tend to investigate, for instance, which specific mindfulness-based programs are most effective in educational settings, rather than asking deeper questions about whether, for instance, in the case of mindfulness appropriating religious practices from Buddhism for trivial, non-religious use is ethically acceptable (Armstrong and Armstrong, 2021). Acknowledgment of cultural, financial or political agendas driving the promotion of particular trends is often missing in literature (Molnar, 2018). This absence is concerning given the major commercial and political interests involved in, for instance, promoting software apps for behavior management (Manolev et al., 2019) or advocating synthetic phonics as the pre-eminent method for teaching reading (Glazzard and Stones, 2020).

Recent, pioneering research by Armstrong and Armstrong (2021) has sought to address these shortfalls and present the bigger picture about educational trends: this article explains advances made by this emerging research, including the birth of the exciting new sub-field of critical trends analysis (CTA). An initial finding of the application of CTA across 18 prominent educational trends is that a significant cluster of educational trends drive ever greater educational exclusion of students with disabilities in public school systems in Australia, England and the US. The constitution of this grouping is discussed together with the implications of this finding – particularly the concerning point made by Armstrong and Armstrong (2021) that progress toward a more inclusive and equitable education system is compromised by the growth of trends which construct public education as an outcome-centric and performance-driven sector.

What are trends? A definition of educational trends

Following Armstrong and Armstrong (2021) these can be defined as:

Definition of educational trends

An eclectic group of practices, concepts, programs, cultural movements or educational initiatives has widespread uptake (popularization) across educational systems and settings. Trends can also be recognized with an identifiable term or terms which signify their main idea or benefit; neurodiversity, mindfulness, inclusion, or behavior management. Trends have either shaped public policy or have been popularized by public policy, often in conjunction with coverage in the media or marketing promotion by a commercial organization. As social phenomena, trends can be understood through looking at their lifespan, where they can be categorized as emerging, dominant or in decline.

A growing number of educational trends have emerged since the 1990s, suggesting an acceleration in the underpinning drivers for popularization and adoption of trends. Synthetic phonics, evidence-based practice, mindfulness, growth mindset, belonging and many other terms gained widespread adoption across education systems in the English-speaking world from the 1990s onwards, although many have much older origins. Older traditions in pastoral care can be understood, for instance, as the historical precedents for recent educational initiatives which stress the importance of attending to student wellbeing (Purdy, 2022). Recognizing these historical roots might suggest that many trends are flimsy, time-limited phenomena which re-brand older traditions – qualities attributed to fads.

What are fads? A definition of educational fads

Following Armstrong and Armstrong (2022) an educational fad could be defined as:

Definition of a fad

A pejorative (negative) term to describe an educational trend which has little intellectual substance and which is temporarily popular. Lack of durability is an element of the negative framing of fads.

The current popularity of evidence-based practice (EBP) as a desired feature of education often means that fads are typically framed, in opposition, as highly undesirable trends, although evidence-based practice is, itself, a trend originating in medicine and health science in the 1980s and subsequently applied to education (Connolly et al., 2018).

The framing of low cultural value embodied in the word “fad”, draws on deep seated value systems stretching back across time into the 20th and nineteenth century (Williams, 1976). Current educational trends are better understood, arguably, by avoiding value-laden evaluations like “fad” and, instead, by taking a longer-term perspective on their emergence using theory designed to understand cultural production. What immediately follows provides an explanatory framework for educational trends, drawing on sociological theory about cultural production, specifically the production of value. Explaining why trends appeal to teachers, education leaders and, also, other influential system-actors such as “policy entrepreneurs” is central to this framework. As is highlighted, this framework plausibly explains why there has been an acceleration in the number of trends appearing and being adopted in education systems in Australia, England and in the US since the 1990s.

In the UK, US and in Europe industrialization and urbanization in the later 19th and 20th century were drivers for mass cultural production and consumption on a scale hitherto unimaginable (Herbert, 2007). At the same time these societies were highly socially stratified, with social hierarchies based on perception of an individual’s class, race, employment, gender, ethnic group or sexuality (DiMaggio, 2019). The growth and increasing differentiation of the cultural market was accompanied by the growth of ideas and fashions which led to increasing differentiation of cultural goods to meet niche cultural markets. Consuming these cultural goods and adopting associated social practices, involve what the French sociologist Pierre Bourdieu describes as the “distinctions”: the advantages which adopting these fashions conferred on the individual in complex, increasingly urbanized and stratified social spaces. As Bourdieu convincingly suggests, developing what he describes as personal “cultural capital” through adopting fashionable attitudes, practices, and customs (“lifestyle”) is an attractive strategy for individuals who live in socially stratified societies built around hierarchies and the ethos of competitive individualism (Bourdieu, 1987). This sociological model of cultural production and cultural consumption has relevance when considering educational trends occupationally and their function in the education sector.

One plausible explanation for the power of educational trends is that demonstrated knowledge of trends, preferably signified in social situations such as a school staffroom or professional learning event, confers professional cultural capital on the individual. Such an application is consistent with Bourdieu’s emphasis on what motivates individuals to adopt specific social behaviors, around cultural consumption: for instance, the outward demonstration of culturally discerning behaviors related to precious knowledge of the new and emerging fashion (Bourdieu, 1987). Furthermore, this use of knowledge about trends as a form of cultural capital structurally fits with Bourdieu’s pessimistic evaluation of the “education system (as) a ‘conservative force’ given its primary role in the reproduction of existing class relations” (Brown et al., 2016, p. 191). This sociological explanation of the rising appeal of educational trends is strengthened by considering negative processes affecting occupational working conditions in public education in Australia, England and the US over the last 20 years.

Persistent deterioration of occupational working conditions in public education in Australia, England and the US has occurred, involving, for instance, increasing casualization and performance monitoring together with ballooning workloads (Flores, 2020). Indeed, working conditions have deteriorated to such an extent that recruiting and retaining teachers is now an endemic problem in many areas of Australia, the UK and US (Garcia and Weiss, 2019). These larger occupational changes can be understood as driving the increasing desire for professional distinction by teachers in an increasingly competitive, performance-obsessed, and insecure workplace (Flores, 2020; Lo, 2012). This driver for professional success through distinction does not only apply to “shop-floor” educational staff (teachers, tutors, lecturers, classroom assistants, sessional lecturers) but cuts across professional hierarchies.

Currency with the latest ideas and with trends is often built into the person descriptions (PDs) of employment for educational leaders in public education, including Principals, Heads, Center Managers, Associate Deans and Deans, together with the complex managerial lexicon of staff titles in educational bureaucracies. Identifying and promoting new practices, concepts, programs, cultural movements or educational initiatives is often motivated for professionals by employment-related key performance indicators (KPIs) and performance review goals: practices which are described by Flores (2020) as expressions of the increasing “technicization of teaching” (p. 220). Many of the trends which occupy educational leaders’ time on a day-to-day level are handed down by local, state/regional and national/federal government. Addressing these multiple demands is, unsurprisingly, cited by teachers and school leaders as a major driver for high workplace-related stress. An important report by detailing findings from a national survey of teachers about perceptions of the teaching profession (n = 2444), for instance reported teachers urging government to “introduce fewer initiatives simultaneously” and requests to “let people solidify their practices and work before bringing in the next idea” (p. 14).

In parallel, over the last 20 years, governments across the English-speaking world have increasingly prepared education, health and hitherto wholly public areas of public service, for commercialisation.

These policies, enacted by governments at all levels across the English-speaking world, are based in philosophical ideas about the need for small government and emphasize the primacy of the market, calling on ideas by economists such as Friedrich Hayek. Arguments in favor of “rolling back the state” and in favor of individuals making decisions in the extended market places are commonly described as expressions of the political philosophy called neo-liberalism. In their perceptive book *Governomics*, for instance, describe a growing “push to commercialize the common wealth” (p. 249). Public education, like public health, is a major element of this common wealth in Australia and the UK for instance, and historically has been a focus for public spending by government. Many educational trends have been at the vanguard of the efforts to commercialize “common wealth” and as part of concerted neo-liberal education policies to reduce tax costs for business and the corporate sector via greater involvement of the private sector in delivering public education.

Prominent educational trends respond to this increasingly neo-liberal public policy environment in myriad ways. Armstrong and Armstrong (2021) provide the following structure (with headings in italics given for each grouping of trends) to make sense of this interaction – a structure they suggest focusses on the psychological appeal of types of trends to teachers, educational leaders and other education professionals responsible for adopting them:

Performance enhancing trends

These trends build on the premise that students are academically underperforming, feeding off a sense of performance crisis often popularized in the media or government reports – the “problem”. Deficits in a student’s reading performance, mathematical knowledge or scientific knowledge are often focused upon using metrics based in comparative international student performance. Trends in this category claim they can boost student performance, “the solution” – and in doing so ultimately seek to improve school performance and student attainment as measured by national or international metrics. Performance enhancing trends present a convenient tool for meeting KPIs for educational leaders and across education bureaucracies. Trends in this category are often framed instrumentally and adopt simplistic, input → output logic, ignoring complex contextual factors.

Technical trends

Focusing on technical wizardry, “new” or “scientific” knowledge make up the appeal of trends in this category. Technical trends often infer that they can bypass the complex, time-consuming, human, political and institutional factors which make up contemporary classroom practice – offering an attractive apparent shortcut. Technical trends align well with innovation and systems-thinking, values that are increasingly prized professional attributes in the education sector and they also draw on older ideas about the benefits of “machine thinking” in this context. Max Weber, who coined the word bureaucracy, favourably discussed an efficient “bureaucratic apparatus” in his book *Economy and Society* published (posthumously) in 1921 (Weber, 1978).

Crumbling orthodoxy

Trends in this category inhabit mainstream space in educational practice and discourse, constituting widely-accepted “common-sense” cultural and professional knowledge held by teachers, education leaders and policymakers. The term “crumbling” refers to shedding of value which occurs to dominant belief systems when they occupy this central space in a system where innovation and change is increasingly privileged.

Cures, therapy and hope

These trends comprise a varied group which respond, in different ways, to longstanding structural problems in education and in some cases, to the accelerating neo-liberal direction of educational policy.

Trends of the near future

These trends have not yet achieved mainstream orthodox status in education – for many reasons – but could in the future rise to dominant status.

Educational trends and exclusion

Armstrong and Armstrong (2021) advance this analysis of trend categories by suggesting that meta-trends become visible at the highest scale of scrutiny and when looking across the broad sweep of public education systems. A grouping of trends which service neo-liberal policy goals including the increased marketization of public education make up one important meta-grouping. Armstrong and Armstrong (2021) describe an influential grouping as the “performative school” and suggest that a rising tide of exclusionary practices in public education systems are driven by the growing influence of the performative school meta-trend.

According to public data, the formal educational exclusion of school age students in Australia has risen inexorably over the last 5 years at least. Multiple public enquiries and policy reports across different jurisdictions have concluded that children or young people with disabilities are grossly over-represented in these statistics. For instance, the recent Disability Royal Commission Hearing 7 (DRC, 2020) acknowledged that the application of disciplinary sanctions, including exclusion (expulsion) from settings was an escalating problem for students in this population and was highly critical of these practices by schools.

Australia is not, it seems, a special case in the English-speaking world. Although interrupted by the COVID-19 pandemic, the UK Government’s latest report highlights that when, “Looking at Autumn term (2019/2020) only: there were 3200 permanent exclusions in 2019/20 (up 5% from 2018/19) and there were 178,400 suspensions (up 14%)”. The US has seen sustained growth in school exclusions, featuring the popular trends of “three strikes” and “zero tolerance” movements across the school system (Gage et al., 2019). Deeper analysis of school exclusion statistics reveals that this phenomenon affects some of our most vulnerable and disadvantaged children and young people. In the US, African-American students are disproportionately represented in statistics concerning school exclusions collected since at least the 1990s (Gage et al., 2019; Skiba et al., 2011). In Australia students who are indigenous are consistently over-represented in school exclusion rates (Higgins et al., 2021) as are children and young people affected by disability.

In parallel with a growth in official exclusions, gray exclusions, appear to be surging. Gray exclusions constitute a wide range of informal and often subtle practices and actions by school staff which marginalize and exclude students, often resulting in a change of setting for the student. In the UK one prominent practice associated with gray exclusions is off-rolling which constitutes: “illegal permanent removal of students from school rolls where removal is in the school’s interests and designed to enhance school performance data” (Done and Knowler, 2020, p. 1). In many cases informal actions by school staff lead to students simply withdrawing from their existing education setting: an outcome also described as “early school leaving” or more judgmentally as “drop out” (Squires, 2020).

Formal exclusion and gray exclusions counteract policy initiatives designed to bring about greater social justice in education systems. Australia and the UK (although not the US) have, like many countries, made a historic policy commitment to this progressive outcome as signatories to the UN *Convention on the Rights of Persons with Disabilities*. Failure, to meet this commitment in educational practice produces politically embarrassing headlines in the media and has attracted the negative attention from civil rights groups and organisations (AFDO, 2019) who criticize the socially regressive direction of education in Australia. Furthermore, the legal requirement to provide education in Australia, the UK and US, together with commitments to educational inclusion in policy at state or regional level, add further gravity to the problems caused by a growth in exclusions. Slee (2019) asks the critical question of whether belonging is meaningless rhetoric for children with disabilities in “an age of exclusion” (p. 28).

Highly critical public reports by system-level “guardians” and featuring highly critical analyses of exclusionary practice in schools, have added to the sense of crisis in the implementation of inclusion and conveyed a moral crisis about the treatment of vulnerable students in public education (DRC, 2020). Crisis narratives about “the future of inclusion” strengthen socially conservative efforts to reject progressive educational reform and threaten greater segregation in education systems, rolling back the hard-won gains won by advocacy groups and civil rights groups in the 1970s and 1980s (Slee, 2018). Researchers working in the field of inclusive education have noted the use of this rhetorical device in constructing apparent failure of inclusion. Boyle and Anderson (2020) describe how efforts to progress educational inclusion in Australia have been stymied by a wider movement, in educational policy, society and politics against social equality and which result in an apparent failure of inclusion.

They emphasize that: “Acknowledging and understanding the challenges that face successful inclusive education practice in Australia is imperative” and add “Without this knowledge it is too easy to lay blame for the increasing rates of segregation and exclusion at the feet of inclusion, to position the construct as an unattainable utopian ideal. Yet, it is evident from the brief discussion here that inclusive education has not been permitted to thrive (p. 208).”

The interesting concept of “astroturfing” could have some relevance to this analysis by Boyle and Anderson (2018). Political astroturfing refers to an apparently organic, spontaneous expression of representative public opinion about a contentious issue, but which is in fact, an orchestrated political campaign faking grassroots opinion on a social issue. The aim of astroturfing is to influence public opinion about this issue and initiate policy change. One distinct line of argument often heard in public discussion which constructs the apparent failure of exclusion is that “ordinary” teachers based in schools do not believe in it and support segregated school systems, unlike progressive educationalists based in universities. Watson and Barnes (2021) detail how what they

described as “New Right 2.0” in the UK and in Australia has framed this as an adversarial reality with “teachers against a progressive elite” (p. 216) and sought to use the fiction of unheard teachers’ voices as a lever for politicized policy change in favor of conservative or new-right principles.

Armstrong and Armstrong (2021) introduce critical trends analysis (CTA) as a new framework for evaluating the political and economic framework around emerging trends (p. 123). CTA is designed to help teachers; school leaders; and parents adopt a politically aware and research-informed awareness of new trends and evaluate their support for existing initiatives based on trends. This political awareness is vital because teachers, school leaders and parents have been core audiences for “New Right 2.0” campaigns designed to shift the policy agenda (Watson and Barnes, 2021) with trends often featuring as a strategic tool in this change process. Commercial interests also often promote educational trends targeting school leaders with a range of market products such as, for instance, electronic software to monitor student behavior (Manolev et al., 2019).

CTA calls on scholarship by Bacci (2012) and provides a critical methodology for asking questions about the politics and purpose of the problem-framing which trends are deployed to “solve”. Bacci’s (2012) approach is, itself, informed by post-structuralist critical discourse analysis (Fairclough, 2013) and which has intellectual heritage in theory about connections between power and language conceptualized by the French philosopher Michel Foucault (Ball, 2012). CTA applies this approach to guide the reader through a set of questions which foreground the potential power-relation, politics and commercial interests so that teachers and educational decisions makers have critical awareness of these potential biases before adopting new trends or renewing investment in existing initiatives. Whilst CTA is a highly practical approach it also suggests a new and exciting research agenda: for instance, using CTA to explore sector-specific occupational drivers for educational trends which erode inclusion and promote greater exclusion for vulnerable populations of students. Armstrong (2022), in collaboration with educational researchers based in England, has started to advance this agenda in an examination of gray exclusion.

Conclusion

Explaining the appeal of educational trends, together with their origins, lifespan and their many impacts offers a productive political and economic lens through which to consider educational inclusion. As detailed in this article, this focus helps to unpack rising rates of educational exclusion and illuminate public crisis narratives about the failure of inclusion. Examining educational trends also enables deeper sociological consideration of occupational changes in the education sector, connecting the “technicization of teaching” in classroom practice with a neo-liberal framing of education that is relentlessly outcome-centric and performance-driven, features which are part of wider efforts to marketize the sector, readying it for full commercialization. As emphasized in this article, these outcomes fundamentally oppose and undermine educational inclusion. More positively, this analysis reveals a shared ground for action and research-informed collaboration between educational leaders and practitioners who oppose these trends because of the damage they do to teachers’ work and wellbeing (on occupational grounds) and between educational leaders and practitioners who wish to advance an inclusive education system.

References

- Adey, P., Dillon, J. (Eds.), 2012. *Bad Education: Debunking Myths in Education*. McGraw-Hill Education (UK), London.
- Armstrong, D., 2022. Exclusion and the “wicked problem” of behaviour in Australian schools: a Research Agenda. In: Done, E., Knowler, H. (Eds.), *Off-Rolling and Illegal School Exclusions*. Palgrave, London.
- Armstrong, D., Armstrong, G., 2021. *Educational Trends Exposed: How to Be a Critical Consumer*. Routledge, Melbourne.
- Australian Federation of Disability Organisations (AFDO), 2019. Disability Rights Now 2019: UN CRPD Review of Australia. CRPD Factsheet. Available at: <https://www.afdo.org.au/uncrpd/>.
- Bacchi, C., 2012. Introducing the ‘what’s the problem represented to be?’ approach. In: Bletsas, A., Beasley, C. (Eds.), *Engaging with Carol Bacchi: Strategic Interventions and Exchanges*. The University of Adelaide Press, pp. 21–24.
- Ball, S.J., 2012. *Foucault, Power, and Education*. Routledge.
- Bourdieu, P., 1987. *Distinction: A Social Critique of the Judgement of Taste*. Harvard university press.
- Boyle, C., Anderson, J., 2020. The justification for inclusive education in Australia. *Prospects* 49 (3), 203–217.
- Brown, P., Power, S., Tholen, G., Allouch, A., 2016. Credentials, talent and cultural capital: a comparative study of educational elites in England and France. *Br. J. Sociol. Educ.* 37 (2), 191–211.
- Connolly, P., Keenan, C., Urbanska, K., 2018. The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educ. Res.* 60 (3), 276–291.
- DiMaggio, P., 2019. Social stratification, life-style, social cognition, and social participation. In: *Social Stratification*. Routledge, pp. 542–552.
- Disability Royal Commission (DRC), 2020. Closing Remarks – Chair, Ronald Sackville AO QC Public Hearing 7: Barriers to Accessing a Safe, Quality and Inclusive School Education and Life Course Impacts. DRC, Brisbane.
- Done, E.J., Knowler, H., 2020. A tension between rationalities: “off-rolling” as gaming and the implications for head teachers and the inclusion agenda. *Educ. Rev.* 1–20.
- Fairclough, N., 2013. *Critical Discourse Analysis: The Critical Study of Language*. Routledge.
- Flores, M.A., 2020. Surviving, being resilient and resisting: teachers’ experiences in adverse times. *Camb. J. Educ.* 50 (2), 219–240.
- Gage, N.A., Whitford, D.K., Katsiyannis, A., Adams, S., Jasper, A., 2019. National analysis of the disciplinary exclusion of black students with and without disabilities. *J. Child Fam. Stud.* 28 (7), 1754–1764.
- García, E., Weiss, E., 2019. *US Schools Struggle to Hire and Retain Teachers. The Second Report in “the Perfect Storm in the Teacher Labor Market” Series*. Economic policy institute.
- Glazzard, J., Stones, S., 2020. A rigorous approach to the teaching of reading? Systematic synthetic phonics in initial teacher education. *Front. educ. [Policy Brief]* 5 188.

- Herbert, U., 2007. Europe in high modernity. Reflections on a theory of the 20th century. *J. Mod. Eur. Hist.* 5 (1), 5–21.
- Higgins, B., Carey, M., Dunn, P., 2021. Staff attitudes regarding permanent expulsionary punishment (PEP) from Australian government schools: comparing Queensland with other jurisdictions. *Interchange* 52 (4), 479–499.
- Kittelman, A., McIntosh, K., Hoselton, R., 2019. Adoption of PBIS within school districts. *J. Sch. Psychol.* 76, 159–167.
- Lo, L.N.K., 2012. Issues in teacher professionalism and performativity: Introduction. In: Day, C. (Ed.), *The Routledge International Handbook of Teacher and School Development*. Routledge, London, pp. 13–18.
- Manolev, J., Sullivan, A., Slee, R., 2019. The datafication of discipline: ClassDojo, surveillance and a performative classroom culture. *Learn. Media Technol.* 44 (1), 36–51.
- Molnar, A., 2018. *Giving Kids the Business: The Commercialization of America's Schools*. Routledge, New York.
- Poed, S., Whitefield, P., 2020. Developments in the implementation of positive behavioral interventions and supports in Australian schools. *Interv. Sch. Clin.* 56 (1), 56–60.
- Purdy, N., 2022. Building back better in 2022—with a renewed focus on pastoral care, emotional health and wellbeing for all. *Pastor. Care Educ.* 1–3.
- Skiba, R.J., Horner, R.H., Chung, C.G., Rausch, M.K., May, S.L., Tobin, T., 2011. Race is not neutral: a national investigation of African American and Latino disproportionality in school discipline. *Sch. Psychol. Rev.* 40 (1), 85–107.
- Slee, R., 2018. *Inclusive Education Isn't Dead, it Just Smells Funny*. Routledge, London.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922.
- Squires, G., 2020. A European consideration of early school leaving as a process running through childhood: a model for inclusive action. *Educ. 3-13* 48 (3), 332–343.
- Watson, S., Barnes, N., 2021. Online educational populism and new right 2.0 in Australia and England. *Glob. Soc. Educ.* 1–13.
- Weber, M., 1978. *Economy and Society: An Outline of Interpretive Sociology*, vol. 2. University of California press.
- Williams, R., 1976. *Keywords: A Vocabulary of Culture and Society*. Croom Helm, London.

Building inclusive communities

Andrew Azzopardi, Faculty for Social Wellbeing, University of Malta, Malta

© 2023 Elsevier Ltd. All rights reserved.

Because “community” doesn’t just happen	183
The insider voices	183
The Story of David	183
The Story of Marilyn	187
The Story of Louisa	188
In pursuit of social justice—manifesting my thoughts	188
Conclusion	189
References	190

Because “community” doesn’t just happen

The term “community” is a frequently overused term in social policy. This is a problem for vulnerable population cohorts. Achieving community requires planning and investment but too often we assume that “community” just happens. The archipelago of Malta comprises three main islands and has a population of less than half a million people. The largest island of the three, Malta, is 27 km long and 14.5 km wide. Geographically and strategically, it is situated in the middle of the Mediterranean. Our geographical positioning has served us well politically and economically. Malta is a member of the European Union (EU) and is the smallest and most densely populated member nation. Being a country, which traditionally is very tightly knit, largely Catholic, in a small compactly populated geographic area contained within an archipelago would automatically seem to imply “community”. This cannot be further from the truth. Evidence from the “loneliness prevalence survey” I will be referring to later in this chapter affirms this. Malta is a deeply layered and complex country wherein camaraderie and community are not givens. Our neo-liberal, secularized country is evolving into a melange of individuals “happily” isolated from each other and it is probably the geography and the practicality of having the extended family for childcare reasons that keeps us distanced from each other even more. To write this chapter I have merged several articles I have written over the last couple of years that lay out the dangers of social fragmentation and what I see as being the primary actions needed to re-invent our communities. The indicators we are getting from the grassroots in terms of the social sector and the fissures we are experiencing is a major concern. This chapter asks: what are the ingredients that will construct a community of understanding (Bauman, 2001)? In what way do we make our communities function? How do we build active citizenship? The term “citizenship” implies engaging with one’s community through economic involvement, public service and other efforts to augment the quality of life for all (Azzopardi, 2011, p. 180) (Fig. 1).

The insider voices

Insider voices are fundamental in helping us to conceptualize community and how people are navigating the complex dynamics of everyday life. *The Story of David* brings out the desperation that people must deal with, but also surfaces the notion of resilience notwithstanding the bleakness that life seems to throw at him. On the other hand, *The Story of Marilyn* foregrounds the strength of the cult and how religion, spirituality and the belief in the self and others gives the needed impetus to carry on notwithstanding the adverse conditions created by the community. She offers very important reflections on the need for service delivery that doesn’t “pigeonhole” people. On the other hand, *The Story of Louisa* conveys the complex dilemma whereby tragedy and a willingness to go on in life is very clear. It is also interesting to recognize the strength of activism and how resistance starts from the stories and these insider experiences of people.

The Story of David

David is a well-known strategist both locally and internationally. He has worked on scoping exercises and the development of strategies all his working life. He has single-handedly led several organizations and proposed reforms on a mammoth scale. He often talked about his struggle with Bi-polar II. “Bipolar II disorder is a bipolar spectrum disorder characterized by at least one episode of hypomania and at least one episode of major depression.” (https://en.wikipedia.org/wiki/Bipolar_II_disorder). David spoke about the struggle he has had to face daily. He was diagnosed with this condition some 20 years ago. Before that, his mental health issues were pushed aside, and people would simply tell him to “get on with it”. He says that speaking publicly about his condition and how it impacts him is what gives him solace knowing well-enough that in the process he is helping others. He finds his immediate family his main source of courage. Another source of audacity are his pets who give him comfort knowing they need and depend on him. He does admit that at times it becomes very hard for those closest to him and recognizes that living with a person who must struggle constantly with this pain can get bumpy for them. Notwithstanding the support he has at home, there were moments where he just couldn’t take it anymore. In fact, during a current affairs radio show I anchor on 103 Malta’s Heart (<https://newsbook.com.mt/en/support-system-key-for-people-with-mental-health-issues/>) he explained how he had

Community participation
Young people should not only be given the right to vote, but also to contest general elections; yes, even as young as 16 years of age. Young people should be on government boards, lead authorities and have a voice in important decision-making arenas (Velasquez & LaRose, 2015).
We are nowhere near close to maximising the potential of our local governments. It would be prudent if we practiced subsidiarity and let local people lead their own locality – but with proper dedication of resources to support autonomy in management.
We need to deal with homelessness and rooflessness. This is a growing problem compounded by the Covid-19 pandemic. At times we think of people in need as benefit frauds and scammers whilst the truth is that usury, drug addiction, and unemployment are leading people to the brink. And while discussing homelessness, what about rent? It is another nemesis for our policy makers who need to juggle the right of people to retain a roof over their heads. There needs to be a fair deal for the tenants.
Alas, the social sector that I am arguing for proposes that we place NGOs at the centre of the debate. NGOs need to be recognized as being at the heart of our communities. They are the ones that need to oil our social machine. They are the ones that need to guide us at this juncture.
Strong community is sustained by care for the environment. Our trees should not be sacrificed to save traffic time. We need to take bigger decisions than simply felling trees. Trees are also memories. A healthy environment should be synonymous with social wellbeing. Gentrification and the way our country is going to the best bidder is worrying. Government needs to understand that not having enough space will result in distress (Vella, Falzon & Azzopardi, 2019).
We still have social classes in Malta and those at the bottom of the heap are disadvantaged whilst the rest are privileged in every sense of the word. We are a classist society and the political discourse insists that “we have a bigger middle-class”. This is incorrect.
The glass ceiling is still there, and gender issues remain prevalent. This country would be a better place to live in if we had more women at the helm in all sectors of society. I am also proposing that we address the issues around misogyny that encapsulates our country.
Liveability
There is no ‘nation’ but humanity, no colour, no religion that should distinguish us because we are all here to make this a better place for ourselves and others. Cosmopolitanism is of benefit to our society.
I believe that older persons are the ones we need to listen to. Their wisdom is for all to savour. Let’s treat them honourably and not turn them into the unwanted, pigeonholed in homes and residential services.
Immigration needs to be handled urgently and expediently. Malta should be the epitome of humanity and serve these people in need well. Whether it is boat people or people coming to Malta through any other way the truth is that we are duty bound not only legally, but morally to provide a safe haven.
We should also ensure that all those entitled to free medication are not made dependent on charity organizations, making them even more vulnerable.
Loneliness is a disfigurement of the illusion we have that our communities are inclusive. Research completed by the University of Malta Faculty for Social Wellbeing in 2019 has shown, that we have over 44% of our population who struggle with loneliness and social isolation. In this sense Covid-19 managed to surface existing shortcomings but hasn’t made this situation any better (Azzopardi, Bonnici & Clark, 2021).
There should be a concerted effort to stop drug trafficking – not enough is being done – drugs are everywhere, drug addicts are collapsing as we speak – the system has failed them. The proliferation of drugs especially amongst young people is devastating. The fragmentation there is between stakeholders is not making the situation any better.

Fig. 1 Manifesto of proposals.

Civil society
Charity collections should stop. When there is a new project being proposed or funds needed by an NGO we should avoid setting up a cantankerous campaign. Telethons should be focused on creating awareness of the issues and not on collecting money. Collecting money in this way should stop. We need marathons, but of ideas and solutions and of ways to attract volunteering and not scenes that dehumanise those at the 'bottom of the heap'.
The Commissions (those in the social sector) should be accountable to Parliament, but before that they need a revamp – they're just not working. The notion of Commissioners (for disability; for mental health; for children; for older persons; for migration; for domestic and gender-based violence, to name a few) is a noble concept that was set up with good intentions. I am suggesting that Commissioners should have a transversal role and they should respond to Parliament and not simply become part of a Ministry's trappings.
Parents should give quality <i>and</i> quantity time to their children.
People deserve to be happy and appreciate that life is beautiful – and for those people who don't feel that, then we should help them to find contentment.
Civil society is the citizens' last line of defence (Barrett, & Brunton-Smith, 2014).
Family is of the utmost importance and protecting this institution is essential for building community.
The national statutory social welfare agency needs to evolve into a coordinating entity.
Social policy actions
Poverty and social exclusion are blemishes – bi-products of a neo-liberal mindset that has encased our communities. It's how much you achieve in terms of money, estate and status which seems to matter nowadays and not who you are.
We still have far too many children in 'care'. Fostering is a great alternative but needs to be at the heart of our social provision with more dedicated funds and energy. Residential care is probably something we can't do away with now. However, what we need to keep investing in is adoption of children who have been abandoned by their parents. A lot of work has been done already, but we need to see more children moving into alternative families. Whilst we're talking about children, the court services responsible for child-related issues need to pull up their socks - they are in shambles. There is clearly a problem when it comes to issuing care orders, the lack of use of the children's house is inexplicable and the absence of promptness from the Court when it comes to decisions that are related to children is atrocious – it all moves at the pace of a snail.
Malta should return to its core values of being welcoming to people especially to migrants who need our hospitality. We can be an example to the World on how we deal with people who are desperate and not turn our face.
We should support parents who have lost their children to drugs, suicide, traffic accidents or sickness.
Alcohol sales should be controlled, but I see no interest from law enforcement to curb abuse.
I am also worried that the social sector is again spread amongst too many different Ministries with not enough efficient mechanisms to keep the system together. I must have counted at least nine different ministries who are involved in the social sector, and this is apart from the Parliamentary Secretaries who all have a finger in the 'social pie'. I expect that the Ministry responsible for social policy takes the lead – fragmentation is untenable.
It is the right time to close most of our residential institutions where some 10,000 persons live and bring these people back into the community – it's where they belong. We need to address institutionalization which runs into thousands of people being placed in residential homes. Related to this is that we also need to tackle loneliness which is leaving a scar on our society

Fig. 1 (continued).

<i>Social governance</i>
We need to stop hate speech and discrimination against people who are LGBTQI+. Good legislation does not automatically mean good quality of life.
Our communities should be physically and socially accessible.
Our Institutions have proven to be weak and succumb to the hidden powers (because of that journalist Daphne Caruana Galizia was assassinated) – we need a new form of governance with the proper checks and balances (van Zyl, 2014).
It is the time for a building construction moratorium - a time when we stop all construction for a year or two and give the nation some time to breathe.
Illegal drugs like cannabis should never be institutionalized, but I also believe that people who are addicted should be loved and not incarcerated. Drug addiction should be decriminalised.
Our state-run social agencies should be managed by people whose sole focus is the good of society.
<i>Values</i>
We should respect and value animals on the same level as we do human beings.
Arts and artists are the heart of our communities. Cartoonists are amongst my favourite.
The Catholic Church should return to its core teachings and be the important player it was in the past, but it must shed its negative institutional history. It should reach out to people beyond its parvis.
Prostitution is what it is and let's call it by its real name – 'prostitution'. I believe we shouldn't objectify women anymore. I also believe that domestic violence scars our society profoundly.
There is no life after death, but death remains the most important transition in life and we need to start talking about it, especially with children, more persuasively.
We need to have more talk shows and debates on our TV, Radio and news portals as stations are symptomatic of dialogue and not simply polarized. Our journalists should be protected. Politicians and public servants cannot choose when to reply to questions, they are obliged to do so. Added to this, I believe our media houses should be at least partially state supported. They are the pivot point for our democracy. I believe that we should give resources for more podcasters, bloggers and vloggers to be able to operate – they are another important opinion in our media landscape.
The time is ripe to get all stakeholders around a table; academics, policy makers, politicians, civil society led by the President of Malta and have an ongoing conversation on this sector. It is only by having this ongoing debate that we can ensure that the ship is being steered in the right direction and responds to the needs 'out there'. I also recommended that when it comes to the political manifesto the social section of the electoral programme should be common for all parties to ensure we have a comprehensive non-partisan vision (Clement, 2016).
<i>Service provision</i>
Professionals in the social and mental health sector should not over-price services and support for people who are in a vulnerable position and unable to contest or refuse to get services because of the cost this incurs.
We should recognize that suicide is nothing less than a failure of society to reach out to a call for help.
People who go into the social care professions need to be ethical, have a vocation and be driven by the cause – if it's just the wage or the power they're after, they're in the wrong place.
Our Judiciary should have a sentencing policy, should not succumb to media pressures and should be just, equally and always.
Solitary confinement should be completely abolished from our criminal justice system.
That the prison system should be dismantled, and the criminal justice system re-invented. It's taking in water from everywhere. Justice is not being seen to be done. Prison should be about 'love' (in its widest of interpretations) rather than retribution. Sorry but those who currently manage the system are not fit for purpose and they have failed the State.

Fig. 1 (continued).

Eating disorders need to be addressed nationally. The problems we are witnessing in this area are of mammoth proportions (Times of Malta, 15/10/2020 – Accessed on 21/10/2022).
The Faculty for Social Wellbeing within University should remain a voice for the voiceless – this is as important as training good professionals in the various areas of social care.
Facebook and Instagram should be flooded with positive messages and used as an emancipatory tool especially for disadvantaged groups who do not have the same degree of power and positioning in our society.
We need freedom of speech but that does not give us the right to abuse others, especially in the world of Facebook. There's something called honourable behaviour. Agree, disagree – that's fine, but respectfully.
Politicians and businesspeople should register their meetings and declare their interests. Once they start consorting it's a disaster. We've seen it happen in the past and it keeps repeating itself. I also believe that our politicians should not focus on the gains but on the pains. Political parties should stand by their values, not only respond to their quest for popularity. I also believe political parties should reinvent their social ethos – where has it gone?
We also need to have Community Development Workers placed in every locality governed by a Local Council so that we can keep understanding what is happening in our communities and how we can address the day-to-day issues.
We should also be reviewing our social benefit system to ensure we are not creating dependency. We also need to tackle the issue of benefit frauds with even more determination.
We need an improved mental health strategy and a new hospital suitable to the needs of the people who need to use it. We need a policy and strategy on suicide and need better trained personnel and the right infrastructure to support first responders who either go through difficulties themselves or people who are victims.

Fig. 1 (continued).

attempted to commit suicide twice; one of them being by driving his car at full speed into a wall, hoping that this attempt would “succeed”. Luckily it didn't, and he came out of that experience able to tell us about it albeit he remained impaired. The daily challenges he has had to face are a massive burden and digging in and trying to find scope and confidence does not come easily for David. He says that he is not always grateful that he survived his attempted suicide, but also reiterated that there is a lot to live for. Still, the overwhelming feeling of depression or hypomania keeps getting in the way. It is his determination that pushes him through the next stage in life.

David Spiteri Gingell, who was diagnosed with bipolar II some 20 years ago said “the surest way of keeping you afloat and not get lost in society is to surround yourself with people who are willing to help you. And this can only be done if you're open about it ...”

Speaking on Andrew Azzopardi's talkshow on 103 Malta's Heart, Spiteri Gingell said that unless people with mental health issues open up about their daily struggles, the people around them will not be able to understand what they're going through.

The Story of Marilyn

Marilyn is a young person who has struggled with anxiety for years. She is a highly talented media producer. She speaks about the struggle to tackle each moment and getting through one day at a time. She claims that faith in God is what sustains her and helps her to process the struggles. She says that in her mind she is in a constant state of struggle and at times she is literally unable to get herself out of bed with the intense anxiety that grips her. For her, each day brings more low points. She says that she values friendships immensely and that her work colleagues are pivotal in accompanying her as she gets on with her life. The way society at times deals with her is not always optimal, but with the help of those around her she walks through the haze and keeps finding the right doses of energy to take on the next challenge thrown at her. The strength of character of this woman is inspiring.

The Story of Louisa

Louisa's story is one of immense pain that challenges belief. She spoke about the incident that defined her over 17 years ago when her daughter Jessica was crushed under the weight of a school van (bus) that had just returned her other child. Louisa spoke of how this incident led to the instant death of her 21-month-old daughter and how it marked her for life. She found solace in her friends, in her family and also in God.

The first thing she did was to embrace the driver who ran over her daughter. I'm shocked. This wasn't the end of it. A couple of years after, she and Philip, her husband, had another son who was diagnosed with down syndrome. Louisa knew that the pressures loaded on her by a society which is ableist would increase her burden. She was never one to buckle in the face of challenges imposed by society, but Louisa knew from day one the constant demands and expectations and air of selectivity in our community would create endless troubles for her son with down syndrome. She was and still is immensely resilient but knows well enough that she and her family had so many obstacles to tackle. Louisa's mental reserves helped. This mother of immense strength has been instrumental in providing for the best possible quality of life for her son Ben who is now semi-independent, has his own friends and work colleagues and has *joie de vivre* like no other. What Louisa kept stating was that it was society that becomes the biggest problem in all of this. The birth of her son with down syndrome confirmed the intensity of her love. Society is not only very selective, but also deeply discriminatory. She and Ben faced another challenge. They had to deal with further judgment and social hostility when Ben came out as a gay person. She says that the wedding of her son to his partner surfaced love beyond words. She claims that she has never witnessed such exquisite beauty before in a wedding. Louisa sums up these intense stories by saying that notwithstanding the death and the pain and suffering she had to endure, she felt God was close to her. In fact, nowadays she remains actively involved in Drachma (a support group) and speaks regularly about her story (Curtin et al., 2015).

The stories of these three people are salutary. They show what humanity is made of and infers how our social policy should respond. These stories reveal the nuances that our social policy struggles to deal with. Considering the estimated 7 billion euros we throw into social policy every 5 years (the term a government spends in power) confirms the reality that the return on investment is not what is to be expected. When we still contend with endless social problems like lack of standards in services; problems with loneliness, which reflects various shortcomings, namely flailing communities, lack of communication and networking which formerly was at the heart of our society and an increase in a neo-liberalism which drives people to focus more on their own pursuit of wealth rather than taking into cognizance the needs of the community. In a number of research projects, the Faculty for Social Wellbeing was involved in (<https://www.um.edu.mt/socialwellbeing/ourresearch/researchprojects>) to try and respond to the growing complexities of making our communities function equitably we have surfaced a number of issues that require our immediate attention.

To compound all of this in a document collectively authored by the Faculty for Social Wellbeing (https://www.um.edu.mt/_data/assets/pdf_file/0009/485343/FacultySocialPact_Printrevised.pdf), *A social pact* (2021) I find this quotation to be particularly significant;

Inequality in Europe was described in a 2017 OECD report as reaching its highest level, with the authors stating that "the 10% of wealthiest households hold 50% of total wealth; the 40% least wealthy own little over 3%" (p. 6).¹ Within Malta this has translated to over half of the wealth belonging to the top 10% of the population. ² The share of workers at risk of poverty is on the rise, compounded by the effects of the ongoing pandemic and its aftermath (2021: 9).

In pursuit of social justice—manifesting my thoughts

My guiding purpose has always been the pursuit of social justice. David, Marilyn and Louisa embody what our social policy should look like and encapsulate. Their stories have motivated me to come up with a Manifesto of proposals. With these three protagonists it isn't just an issue of resilience, or of facing their challenges. They are more than that. They reflect the faults in the system, the selectivity in our social governance and the fact that people, many of them with low social capital are short changed. In my role as a volunteer, a social operator, activist, practitioner and academic, I have sought to work toward "good community" (Gilster, 2012). Notwithstanding the fact that I am an agnostic, my Jesuit education, which I am proud of, implanted in me this sense of "social justice" which sits above everything else. Nevertheless, the abuse, the criticism and the attempt at cancellation I persevere knowing well-enough that social justice comes at a price. It is also comforting knowing that in the pursuit of social justice I am not alone (Alessandrini, 2011).

Today, however, social protection is one of the major demands of social movements all over the world. But what kind of social protection? Do we have to (re-)build the damaged or incomplete welfare states? Do we agree with the new international agenda? Or do we want to develop our own vision and programmes? We claim it is not enough to just ask to have access to health care, we also want people to have a possibility to live healthy lives throughout and to participate in the decision-making. This means we broaden the concept of social protection; we link it to democracy, the environment and trade and we integrate it into the thinking on commons. What we propose is a truly innovative approach, a long-term agenda for contributing to system change.

Mestrum and Menon (2021:12).

Following Mestrum and Menon (2021), this chapter asserts that the reinstatement of “social protection” is a prerequisite for building community. The remainder of this chapter outlines my *Manifesto for an Inclusive Community* of how the social sector should evolve based on the following components:

- Community participation;
- Liveability;
- civil society;
- social policy actions;
- social governance;
- values;
- service provisions.

The reflections I offer represent an attempt to make social policy effective and comprehensive through a more judicious use of resources to build inclusive and supportive community.

An issue that merits our intervention is the issue of social exclusion and poverty. To do this we need all the possible help from Non-Governmental Organizations (NGOs). NGOs must work for the persons they represent and not the other way around (Rodekamp, 2013). The relevance of NGOs lies in the fact that they are autonomous from the agenda of the Government. Their focus remains the cause they represent, and they frame the needs of that minority to truly represent them.

The issue of justice has been raised in the field of geography from time to time in regard to globalization, suburbanization, gentrification, and immigration processes. In this context, the question of fairness has spurred a growing literature that theorizes justice in space (Fainstein, 2001). This rise in interest, though, seems to skip the basic question of what we should seek in order to measure justice (Israel and Frenkyl, 2018: 649).

Naturally civil society is fundamental to ensure that the nuances and contentious discourses of oppressed communities are brought up and tackled through a strong insider voice. The money that the Government and the EU institutions give to NGOs shouldn't become simply a carrot or means to sway the agenda for political ends. We should set up an autonomous *Foundation* so that the financing that the government gives will not go directly to the organizations and associations, but it will be the responsibility of the *Foundation* to direct the monies to the projects that are most deserving. This will ensure autonomy rather than dependency on the Government. There should also be transparent criteria and we should also encourage social enterprise to minimize dependency on the Government for resources.

The issue of social leadership and social governance is problematic. We have a dearth in leadership whether it is in public entities or NGOs and authorities. This is a crisis that overwhelms us and has left gaping holes in the sector.

Lessa et al. (2014).

However, discussion of the metric by which we can measure justice regarding its social consequences has been lacking. Current writings that discuss the relationship between space and justice give little attention to what constitutes well-being in the first place, thus lacking normativity (Olson and Sayer, 2009; Israel and Frenkyl, 2018: 647).

Conclusion

Defining social justice is not easy, and it remains an “under-theorized concept” in education. One conceptual issue is that social justice is an umbrella term with multiple meanings. Blackmore (2009) states that social justice “encompasses a range of terms—some more powerful than others—such as equity, equality, inequality, equal opportunity, affirmative action, and most recently diversity” and, further, that each term “takes on different meanings in different national contexts” (p. 7). Boyles et al. (2009) add that “due to such widely varied meanings [of social justice], it is possible for different groups to act in opposition to another, yet each can do so under the aegis of social justice” (p. 37) (Furman, 2012: 193).

The purpose of sharing these ideas is to help place the social sector back on the agenda where it belongs and through out-of-the-box-solutions address some major issues we must contend with. As academics and inclusive educators we are duty-bound to keep providing elucidations to very complex problems. This has to happen through a mélange of lecturing, research, dissertations, our social engagement and reflection as we go about our professional duties and life journey (Bee and Guerrina, 2014).

The University of Malta, especially the Faculty for Social Wellbeing, should offer more empirical data and research. If we want the sector to have tangible solutions rather than populist knee-jerk reactions, we need to invest money for research in this sector. Having said that in these last 4 years the Faculty has contributed to over 45 research projects commissioned by various NGOs and government and private entities. These projects have contributed directly to policy development, contribution to public debate and have influenced the design of legislation (<https://www.um.edu.mt/socialwellbeing/ourresearch/researchprojects> - Accessed on January 22, 2022).

The term “social dialogue” implies that in the process of globalisation, where neoliberal ideology has established its dominance, social struggles have moved away from the context of the contradiction between labour and capital to find a new platform of class compromise and cooperation. Yet it is debatable as to whether this idea of a shift has objective grounds or not. There is a need to go back a little in order to correctly read recent transformations in industrial relations. This will make it possible to clarify how the aforementioned discourse emerged, together with its context.

Akpınar and Akçay (2015: 438).

Finally, I encourage our communities to reignite a passion to talk to each other and to have each other's back. We also need to connect with the strategies which worked in our communities in the past.

Within the social space, the effectiveness of different forms of capital emanates from the organization of the social world into fields (e.g., the economic field, the artistic field, the education field, and the housing field). All fields are internally structured in terms of (symbolic) power relationships, in which individuals and groups act in order to gain dominance and avoid subordination. The effectiveness of the application of the different forms of capital for different social benefits depends on contexts and people's abilities to “play” their capital most effectively (i.e., applying symbolic capital).

Israel and Frenkyl (2018: 652).

References

- Akpınar, T., Akçay, I.I., 2015. From labour–capital conflict to social dialogue? a critical review of social-dialogue discourse in turkey. *Cap. Cl* 39 (3), 435–451.
- Alessandrini, M., 2011. The role of the third sector in subnational strategic planning. *Third Sect. Rev.* 17 (1), 59.
- Azzopardi, A., 2011. Conceptualising Discursive Communities: Developing Community in Contemporary Society. Web.
- Azzopardi, A., Bonnici, J., Clark, M., 2021. Loneliness in Malta during the Covid-19 pandemic. *Stud. Soc. Wellbeing* 1 (1), 53–64.
- Barrett, M., Brunton-Smith, I., 2014. Political and civic engagement and participation: towards an integrative perspective. *J. Civ. Soc.* 10 (1), 5–28.
- Bauman, Z., 2001. Consuming life. *J. Consum. Cult.* 1 (1), 9–29.
- Bee, C., Guerrina, R., 2014. Participation, dialogue, and civic engagement: understanding the role of organized civil society in promoting active citizenship in the European Union. *J. Civ. Soc.* 10 (1), 29–50.
- Clement, M., 2016. *A People's History of Riots, Protest and the Law*. Palgrave Macmillan, London.
- Curtin, N., Stewart, A.J., Cole, E.R., 2015. Challenging the status quo: the role of intersectional awareness in activism for social change and pro-social intergroup attitudes. *Psychol. Women Q.* 39 (4), 512–529.
- Faculty for Social Wellbeing—Research Projects. <https://www.um.edu.mt/socialwellbeing/ourresearch/researchprojects>.
- Fainstein, S.S., 2001. Competitiveness, cohesion, and Ggovernance: their implications for social justice. *Int. J. Urban. Reg. Res.* 25 (4), 884–888.
- Furman, G., 2012. Social justice leadership as praxis. *Educ. Adm. Q.* 48 (2), 191–229. Web.
- Gilster, M.E., 2012. Comparing neighborhood-focused activism and volunteerism: psychological well-being and social connectedness. *J. Community Psychol.* 40 (7), 769–784.
- Israel, E., Frenkel, A., 2018. Social justice and spatial inequality. *Prog. Hum. Geogr.* 42 (5), 647–665. Web.
- Lessa, F., Olsen, T.D., Payne, L.A., Pereira, G., Reiter, A.G., 2014. Overcoming impunity: pathways to accountability in Latin America. *Int. J. Transit. Justice* 8 (1), 75–98.
- Mestrum, F., Menon, M., 2021. *Redefining the Agenda for Social Justice*. Springer Nature Singapore Pte Ltd. https://doi.org/10.1007/978-981-33-6571-1_2.
- Olson, E., Andrew, S., 2009. Radical geography and its critical standpoints: embracing the normative. *Antipode* 41 (1), 180–198.
- Rodekamp, M., 2013. *Their Members' Voice: Civil Society Organisations in the European Union*. Springer Science & Business Media.
- Velasquez, A., LaRose, R., 2015. Youth collective activism through social media: the role of collective efficacy. *New Media Soc.* 17 (6), 899–918.
- Vella, S., Falzon, R., Azzopardi, A. (Eds.), 2019. *Perspectives on Wellbeing—A Reader*. Brill, Sense, Leiden/Boston.
- van Zyl, A., 2014. How civil society organizations close the gap between transparency and accountability. *Governance* 27 (2), 347–356.

“Disability and society: changing our frames of reference”

Tanya Titchkosky, Department of Social Justice Education, OISE of the University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	191
The orientation of frames	191
Key facts and the objectification of disability	192
Perceiving the frame—social model and debility studies	193
Life with conflicting frames	195
Disability studies as framer of frames	196
Conclusion – modeling the story of disability	198
Acknowledgment	199
References	199

Introduction

How disability is framed and what sort of orientation these frames invite, has much to do with the kind of disability studies we end up doing, or even if we do disability studies at all. As a way to show how these frames work, the issues they invite, and how these frames change, I will explore material produced and widely circulated by the World Health Organization (WHO). Frames matter and every time disability appears, it appears through an interpretive frame. Perceiving these frames is very important for revealing the complex imbrications of disability and society.

I make use of the WHO’s “Key Facts” on disability as an example of framing disability in a non-disability studies way that depicts disability as an individual and medical problem. I show how disability studies re-frames this conventional way of perceiving disability, problematizes conventional frames, while also showing how disability can be framed differently through disability studies. While there are, as Goodley (2011, 48) says, a plethora of “diverse disability studies” frames, this article pursues only two – first, the social model of disability including debility studies and, second, a more intersectional frame that emphasizes the relational character of all disability representations and experiences.

This article ends with a discussion of the different ways of doing disability studies and how these ways frame both the meaning of disability as life and the possibility of social change when we start from disability-as-life as the grounding assumption. I will show that what arises from the recognition of a diversity of frames is the necessity of a self-reflexive politics with disability experiences, which should include the ways we do disability studies. This diversity of frames can be read as an invitation not only to know these frames, but also to reckon with the in(ex)clusions each frame enacts as we grapple with the social and political consequences of a particular frame (Ware, 2005).

The orientation of frames

Depending on which way one turns, different worlds might even come into view.

(Sara Ahmed, 2006, 15).

The truth about stories, is that that is all we are.

(Thomas King, 2003, 2).

How we orient, that is turn toward and frame our bodies, minds, senses and emotions has much to do with how we understand, experience, and work with – or against – what comes to be called *disability*. There are some ways of orienting and thus framing disability that will never lead a person into the world of disability studies. Whatever else disability studies might be, it does not frame disability as an individual organic problem in need of individual remediation.

Framing disability as a problem to be solved at the individual level does lead back, however, into the common-sense world where there is a taken-for-granted assumption that there are “normal” and “expected” ways of being in the world. Some people fall outside of this common-sense world and its expectations because, it is assumed, something is wrong with them. Disability framed from within the perspective of the “world of the normals” (Zola, 1993), can do many things, including learn about disability-as-dysfunction and provide solutions, but this is not the same as doing disability studies. Individualizing frames can conduct a *study of disability* such as providing information about conditions. Disability studies, in contrast, focuses on the

relation between everyday norms and the appearance of disability and can reveal how those everyday norms organize how disability is perceived — that is, framed (Titchkosky and Michalko, 2009, 5).

From the premise that there is a distinction between the study of disability and disability studies, I turn to a World Health Organization (WHO) representation of disability and reveal the non-disability studies frames used to generate it. I reveal the disability studies *frames that themselves reveal the frames* that allowed for the making of the WHO’s (2018) “Key Facts” on disability. Perceiving the frames that make disability appear in one way rather than another is an activity that can be understood as important to disability studies.

Key facts and the objectification of disability

If something is to be constructed as a fact, then it must be shown that proper procedures have been used to establish it as objectively known. It must be seen to appear in the same way to anyone.

(Dorothy Smith, 1978, 35).

The presentation of *key facts on disability* has a long lineage. It has been reproduced over time, around the world, through national and local government agencies and replicated by many disability organizations. The WHO’s tradition of representing disability in this way has been occurring, in various forms and figures, decades before 1981 — the year designated by the United Nations as the International Year of Disabled Persons. Post 1981, the production and distribution of disability key facts became nothing short of common-place. This prevalent representation of disability via key facts highlights the taken-for-granted character of this traditional way of framing disability within Western cultures. The production and circulation of key facts tells the story that people believe that there are facts on disability to be had, some of them key, and that they are already available and should be widely distributed.

Consider WHO’s representation of key facts on disability (or search on-line for an updated version of such facts) (Fig. 1).

Key Facts on Disability

Key facts

- Over a billion people, about 15% of the world’s population, have some form of disability.
- Between 110 million and 190 million adults have significant difficulties in functioning.
- Rates of disability are increasing due to population aging and increases in chronic health conditions, among other causes.
- People with disability have less access to health care services and therefore experience unmet health care needs.

The International Classification of Functioning, Disability and Health (ICF) defines disability as an umbrella term for impairments, activity limitations and participation restrictions. Disability is the interaction between individuals with a health condition (e.g., cerebral palsy, Down syndrome and depression) and personal and environmental factors (e.g., negative attitudes, inaccessible transportation and public buildings, and limited social supports).

Over a billion people are estimated to live with some form of disability. This corresponds to about 15% of the world’s population. Between 110 million (2.2%) and 190 million (3.8%) people 15 years and older have significant difficulties in functioning. Furthermore, the rates of disability are increasing in part due to aging populations and an increase in chronic health conditions.

Disability is extremely diverse. While some health conditions associated with disability result in poor health and extensive health care needs, others do not. However, all people with disability have the same general health care needs as everyone else, and therefore need access to mainstream health care services. Article 25 of the UN Convention on the Rights of Persons with Disabilities (CRPD) reinforces the right of persons with disability to attain the highest standard of health care, without discrimination.

WHO (2018), fact-sheet [Accessed Oct 5th, 2020].

The global distribution of this text (Fig. 2) itself lends weight to it as a legitimate factual representation of disability. The WHO logo assists in framing the facts as objectively just-there, obvious to any reasonable person who encounters them. The WHO logo or emblem, developed in 1948, is a protected image to be used only with permission from the WHO. The emblem is comprised of a globe overlaid by an ancient symbol of medicine, which is a single serpent entwined rod wielded by the Greek god of healing and medicine, also known as the rod of Asclepius. The globe is intersected by lines in all directions and wreathed by a bay of laurels, suggesting that the WHO holds a crown of victory in its race toward health. This emblem appears, frames and endorses what comes next, namely, facts, and key ones at that, that depict disability as a problem of global proportions.



Fig. 1 Key facts on disability

Blessed with the WHO’s authority, disability is framed as straightforwardly a problem, just there, objectively given and obvious to all. The opening four key facts on disability communicate that disability is readily observable and quantifiable — it appears as a billion individuals who have trouble in functioning normally, and who comprise a large population, a population that is further affected by aging and ill health and who face a variety of unmet health-care needs. This fact sheet, circulated around the world in a variety of languages, suggests that this is how disability should and does reasonably appear to anyone, anywhere, regardless of context. The key facts, both in short-form bullet points and in full sentences, communicate that there are more than a billion people with disabilities and of those people tens of millions have difficulty functioning. Disability is framed as a taken-for-granted, obviously recognizable problem of lack of function with significant individual and global impact.

The authority of these key facts lies, in part, in not showing that they are human artifice; administrative inventions. Instead, they are made to appear as springing from the seemingly objective knowledge and power of the WHO as it addresses disability as naturally a problem. Providing numerical facts covers over the social fact that this knowledge, including numerical renderings, is made and makes up a particular meaning for disability.

After the bullet point facts and prior to their further explication in full sentences, we are provided with a definition of disability. Sandwiched between fact-statements, this definition is presented as the proper reasoning behind the key facts. The use of this definition is presented without the need for documenting or identifying the people located somewhere who decided on its use and placement, and without mentioning the history of its development. There is, for example, no mention that in 2001 this definition replaced many other definitions of disability previously used by WHO, (WHO, ND. ICF). Instead, the definition is positioned as the proper non-subjective understanding of disability by invoking an authority, namely, “The ICF defines ...” transforming the name of the definition (ICF) into the definitive depiction of disability itself.

Stripped of signs of human artifice and deliberation, the ICF definition of disability is made to appear as objectively given, as are the facts that proceed and follow it. The definition, like the facts, arises from no one in particular enabling an authorial power as though without the involvement of anyone. Even though the ICF definition mentions personal and environmental factors, no information about these factors appears in the facts provided. Instead, individuals with problems with functioning are enumerated — 15% of the population of the world, or one billion people, have a disability that is expressed in such a way that makes poor health appear as the primary meaning of disability on a global scale. In this sense, the WHO Key Facts on Disability reads as if it offers a direct glimpse at the reality of disability — poor health outcomes accompanied by difficulty with functioning.

While it may be that fact sheets are produced with the explicit interest of raising awareness, what we are made aware of is something that we already know at the level of unexamined common-sense—disability is an individual problem perhaps bigger than we first thought. Disability as a collective issue is pushed out of the frame and supplanted by a personal search for individual adjustments and personal care, at times through the acquisition of rights to care and, at other times, through individualized programs of support and even charity (Schlund-Vials and Gill, 2014; Kelly, 2016). Non-disability, understood as the norm, frames disability as an individualized personal issue that often makes it into a matter of sheer survival. However, this framing does not mean that disability is restricted to the singular issue of survival. What it does mean is that contemporary Western culture is itself deliberately restricted in the ways it attributes meaning to disability (Goodley, 2021).

Defining disability as objectively given, measuring its severity, discerning its prevalence, planning for its remediation are conventional ways of studying “health conditions” and provide a frame for the interpretation of disability and disabled people as dysfunctional. And, yet, without a frame that includes the reflexive understanding that disability is framed, we remain outside of a disability studies orientation. *Disability studies* grapples with *how* disability is made an issue, that is, with how disability is perceived and with the frames that make that perception possible.

Perceiving the frame—social model and debility studies

... Disablement has nothing to do with the body. It is the consequence of social oppression

(Oliver, 1996, 43).

Disability is a profoundly relational category, always created as a distinction from cultural ideas of normality, shaped by conditions that exclude full participation in society of those considered atypical

(Ginsburg and Rapp, 2013 54)

“Debility” [is] bodily injury and social exclusion brought on by economic and political factors.

(Chiayulien, 2020)

When framed by the social model of disability, disability is not perceived in individuals but, instead, is perceived in the interaction between inadequate and inappropriate responses to people with impairments, transforming them (us) into disabled people (Oliver, 1996). Disability is not some unfortunate biological happenstance found in people who are counted and measured.

Yet, the WHO frames disability in this way. The question becomes how might the social model *frame* non-disability studies frames of disability? Let us engage this question by returning to the WHO.

The WHO's Key Facts on Disability, perceived through the frame of the social model of disability, reveals the various ways that disability is *made* into an individual problem. The WHO tells people the facts on disability and does so by omitting, suppressing, oppressing the contravening social fact that “disablement has nothing to do with the body” (Oliver, 1996, *ibid.*). WHO's fact sheet, moreover, makes it seem as if there is a direct relation between bodies, minds, senses, emotions and what comes to be experienced as disability. But such a direct relation to embodiment is a myth.

Even a slightly more social perspective can make us take note of how people's experiences always happen somewhere and in the midst of others; in relation to the knowledge regimes, interactional norms, and physical and institutional environments that combine to make up the context of any individual's appearance as disabled (Titchkosky, 2007, 17). The fact sheet covers over the “fact” that disabled people appear in locales that cause disablement by suggesting that disability can be counted and thus known as an individual condition regardless of context. How we frame disability, then, has consequences not only for how disability is perceived and understood, but also for how it is lived. For example, “What is wrong with you and what do you need?” is one way to live with disability that fits neatly within the WHO frame; whereas asking “How's the community treating you?” is a question arising from a social model frame.

To produce facts on disability as if individual lives can be abstracted from the social and political relations that produced them maintains an unjust relation to disability and is part of the process of oppressing people with impairments, insofar as such facts represent and reproduce an ideology of individualism. This ideology is at work when individual selves are framed as the locus of the issue of disability and where the self is perceived as a task and project. This is what various social theorists call “individualization.” Bauman, for example, describes it as a key social process for the workings of contemporary capitalist social organizations, which he characterizes as both destructive and wasteful (Bauman, 2000, 30–38, 2004). Conrad (1992) characterizes this as medicalization. Framing disability as a personal problem of dysfunction participates in the on-going individualization of the self under contemporary societal conditions. Individualization as a constant demand of normalcy is reframed by disability studies as itself a phenomenon in need of inquiry.

The social model of disability frames disability as *social*; that is, a result of its relational contexts. The WHO frames disability as “the interaction between individuals with a health condition ... and environmental factors.” This interaction depends upon the tacit understanding that disability is already an individual calamity. Framing it in this way, conceals the complexity of disability's “profoundly relational” character (Ginsburg and Rapp, 2013, 54). Providing facts on disability can now be understood as a process whereby social structures continue to transform people with impairments into individual problems. Enacting the objectification and decontextualization of disability are processes that work to define the reality of disability. Facts are thus not a snapshot of disability reality but are, instead, a form of social action *making* a reality for disabled people.

The question of what we are doing when we perceive and represent disability in whatever way we do requires that we engage with the profoundly relational character of disability and there are many ways to do so within disability studies. While those who perceive disability through the frame of the social model suggest that acts of individualizing and decontextualizing disability can be read as a collective failure to respond appropriately to people with impairments, others, such as Jasbir Puar (2015), suggest that it is a collective failure to respond appropriately to wars, capitalism and colonial occupation that debilitate people. Let us explore the concept of “debility” further.

Puar suggests that “in this historical moment” many approaches “face the rise of disability as a recognized vulnerable identity in need of state and global human rights protections,” thereby propagating a notion of disability as “a socially maligned condition that must be empowered to and through a liberal politics of recognition” (Puar, 2015, 9). To seek recognition confirms the power of those structures that grant such recognition, suggests Puar. Perhaps this is why she says that in seeking human rights protection, pride, or pursuing other forms of disability politics disabled people are tacitly confirming the validity of the structures that also launch war, reproduce capitalism and seek to colonize. These are structures and activities which produce bodily injury and social exclusion. Puar suggests that a focus on debility requires a reevaluation of what we are doing when we notice disability and seek improvements:

This revaluing of excess/debility is potent because, simply put, debility – slow death – is profitable for capitalism. In neoliberal, biomedical and biotechnological terms, the body is always debilitated in relation to its ever-expanding potentiality.

(Puar, 2011, 153).

Understood through the debility frame, disabled people seeking their recognition as vulnerable and in need of individual rights end up confirming the power structures that resulted in debilitation in the first place. The debility frame suggests that the current state of affairs excludes and maims people in order to perpetuate order and organization – debilitation is the norm. Puar goes on to suggest that much disability-politics does not take into account how it/they/we perpetuate the process of debilitation and in so doing further debilitate disabled people, especially those located in the global south.

Given that nation state organization is meant to serve capitalist interests, it is profitable to recognize some disabled people as vulnerable and in need of inclusion into the workings of society via individualized rights and accommodation procedures. Still, these procedures of inclusion and recognition will not, according to a debility frame, disrupt the “maiming” of people for the

sake of the on-going reproduction of capitalist power, control, and gain. The WHO facts, framed by a debility perspective, can now be read as serving state control by functioning as an apparatus for convincing all people of a particular way of perceiving disability that serves the reproduction of state power.

With debility as a conceptual frame, disability appears as a vulnerable rights-seeking group that, through a politics of recognition or pride, continues to allow for capitalist regimes to pursue the right to maim, debilitating people of color and other marginalized groups, including disabled people. A politics of recognition can give rise to liberal disability pride or rights, but does little to reveal how the majority of disability in the world is generated by wars, capitalism and colonial occupation. Current regimes of power manufacture disability, including facts about this vulnerable identity group, and in so doing continue to secure their power. Thus, Puar recommends changing the frame from disability as an individual thing, to debility as process – a state of affairs brought on by economic and political structures that benefits those same structures. Disability, from this frame, is manufactured for and by ruling powers.

From within the frame provided by the social model of disability and from the frame of debility, objectified versions of disability appear as something quite different than a biological reality from which information can be gleaned. Framing disability as an object can now be understood as an instrument of individualization and even as serving the aims and interests of capitalist regimes. And, in the contemporary West there are many ways to frame disability that prop up the requirement that disabled individuals adapt and/or pursue individual rights within the system of power, such as it is. Further, the framing of disability as an objectively given individual problem serves to keep disabled people, or perhaps all people, from thinking about how their existence – even when understood as unjustly produced – is a symptom of wider systems of oppression such as colonialism, an ever-growing capitalism, and the wars and conflicts that make profit for few and debility for many.

We have at least two frames of disability that are fairly recognizable. One, represented by the WHO, makes disability primarily appear as biology gone wrong. The other, represented by the social model of disability and debility studies, makes disability appear as both a sign and a symptom of what is wrong in society. This conflict of interpretations is what I turn to now, a conflict that may, upon reflection, once again, change how disability is framed.

Life with conflicting frames

Well, I suppose that's the point at which one would have to ask why stories get told in those ways. Let me make the point that if you tell a story in a particular way you often activate meanings which seem almost to belong to the stock of stories themselves.

(Hall and O'Hara, 1984, 7).

During an interview with John O'Hara regarding the narrative structure of *reality*, Stuart Hall (Hall and O'Hara, 1984, *ibid*), a founder of cultural studies, suggests we should ask why lives and issues are framed by “a set of ways of telling stories that are immediately recognizable to the audience.” Disability, like many other ways of attending to life and people, is discussed in immediately recognizable ways. It is framed so as to appear as a set of stories of individual trouble at the biological level and it is also framed as the oppressive or debilitating nature of social order. Both frames are readily available and recognizable ways of activating meaning for disability. These frames, their meanings as well as the conflict between them—the nature/culture divide—belong to Western culture's immediately available taken-for-granted stock of stories.

These sets of stories framing disability as either all-biology or all-social-oppression are not only available and recognizable, but easily drawn upon by all of us. Indeed, I feel the movement of these set stories in my life, and sometimes simultaneously. When words, time, directions get mixed up, I claim the frame of dyslexia as a biological limit; when others don't respond with some generosity of spirit to these mix-ups, I feel my exclusion and belittlement rising due to their oppressive reaction. As I tell this story, I can also think that mere mix-ups can't possibly be an issue of concern when COVID-19 cases are surging, killing those people most debilitated by oppressive structures of a capitalist order still turning a profit for a few in the global health care industry. Disability framed as all-biology or as all-social-oppression are stories at the ready, belonging to the stock of stories of contemporary Western times. We can orient to these stock stories and examine how they have the power to give shape to thoughts, feelings and actions.

As Hall goes on to discuss the issue of stock stories, the question is not so much whether the stories are good or bad, but rather concerns himself with what they do. That stories, especially readily recognizable stories, are agentive gives rise to a concern about how such stock stories function. Hall (Hall and O'Hara, 1984, 9) suggests that these stories are not innocent insofar as “They underpin, they undermine, they reinforce and strengthen those sentiments, those ways of understanding; and ... by definition undermine, weaken, make less powerful and important the alternative ways ...”. His concern with stories does not end with their capacity to repeatedly provide frames of reference that emphasize or displace significance. Instead, Hall emphasizes the need to attend to stories as makers of reality that undermine alternatives and, maybe, even alterity itself. That I can, for example, experience simultaneously various meanings of dyslexia leading to confusion instead of certainty, harbors the possibility of an alternative, namely that dyslexia is a form of consciousness meeting the world in unexpected ways.

The need for critical engagement with the fact that stories do form realities, is also expressed by Bauman (2004). He speaks of this issue in many ways. Here is one:

Stories aid the seekers of comprehension by separating the relevant from the irrelevant, actions from their settings, the plot from its background, and the heroes or the villains ... It is the mission of stories to select, and it is in their nature to include through exclusion and to illuminate through casting shadows. ... Without selection there would be no story. To say "this would be a fine story if only it did not skip this or that" is like saying "these would be fine windows for seeing through the walls if they were not framed and kept apart by the walls between them".

(Bauman, 2004, 17).

Regardless of what the stories leave out or fail to account for, they still frame a reality and have power to order our lives – "their mission." Bauman's sense of what stories *do* shares Hall's orientation, in that both suggest that stories do much by way of weakening and strengthening forms of comprehension acting to frame what comes to be perceived as necessary and important. Understanding stories as makers of reality, even with their limited and limiting frames, suggests that what is necessary for critical inquiry is a commitment to generating self-reflexive relations with stories.

Following the sense of story as agential, it becomes important to notice that it's not just about telling ever better stories. That is, while the WHO might be asked to include more health conditions for the story of disability as biology, or the Social Model of Disability might be rebuked for not conceiving of debilitating barriers in a more global fashion, this sort of critique leaves the frame in place by asking only for its expansion. Expanding the frame is not the sort of critique that Hall is suggesting.

In the face of the understanding that stories frame reality, Hall continues to critically engage them by asking what the story can take into account, including whether it can take into account its own frame of reference. Framing the frame means, as Hall suggests, cultivating a concern for stories that does not end with their power to reinforce and to undermine. Joining Hall in this way of understanding stories means figuring out what has been done by the stories we have already told, the stories we inhabit and the ones that live in us, that live in our patterns of thought, ways of imagining and forms of perception. This is a theme occasionally reflected in disability studies too as, for example, when Stiker (1999, 15) suggests that "To initiate an analysis of the social workings of disability by way of its integration is a method more critical, even more militant, than to address it in terms of exclusion."

A narrative orientation launches the methodological necessity to continue to seek out whether or not a story demonstrates a self-reflexive character with its own power to utilize a stock story, that is orient to its own frame. Hall puts it this way, "Is it [story] going to acknowledge *that* within its frame of knowledges?" (Hall and O'Hara, 1984, 3, italics my own). This is very different from suggesting one story is true and the other is not; or suggesting that one story is just and the other not; or that it's just a story, not too relevant, just let it be. Hall offers neither the judgment of objective truth or falsity, nor a subjective judgment that its "just a story" of a particular person with interests and biases. Insofar as all stories create a reality and rely on a stock of stories already told and ways of making sense already made in order to do so, the question becomes, how does the story narrate its own frame of reference, its own way of framing thus creating a reality?

Framed by the World Health Organization (WHO-as-storyteller), the story of disability appears as all that can go wrong with one's biology compounded by the political and social environment. Framed by the social model and by debility studies, in contrast, the story of disability appears as all that can go wrong with societal organization both in its failure to respond and in its success to control, compounded by individual experiences of impairment, injury, and accident. Whether indicative of individual or social pathology, in both cases we have the story of Western culture with its belief in a nature/culture divide as one of its origin stories but without offering a reflection on *that* divide as it is made and remade into a key cultural reality for our lives with disability.¹

The simultaneous appearance of multiple frames may be and often is confusing. It is tempting to remedy this confusion by committing to a single frame that, given the context of the situation and of individual interests, is easier to comprehend and feels truer. Still, the existence of multiple frames provides for the possibility of disrupting the certainty of what we know and perhaps forging alternative meanings and realities. Disability is, of course, more than a single frame can frame. In our discussions thus far, disability *is* and is *more than* the symbolic violence actualized by Western cultures' insistence on a nature/culture divide.

I turn now to an exploration of how perceiving multiple frames is a political necessity. How are multiple frames a way to de-naturalize our perception of disability as a thing born of a split between biology and society? And, what of the feel of embracing disability studies' frames that recognizes disability as a complex social phenomenon opening alternative possibilities?

Disability studies as framer of frames

Representations of disability, then, allow an interrogation of static beliefs about the body while also erupting as the unseemly matter of narrative that cannot be textually contained ... While disability as an experience is rarely narrated in terms of its own significance, disability as a metaphor allows writers to access concerns on a metaphysical scale.

(Mitchell, 2002, 17, 25).

¹For a more in-depth discussion regarding how an unquestioned nature/culture divide has influenced the formation of disability studies, see Corker (1998); Davidson (2016); Hughes (2019); Michalko (1998); Pickens (2019); Slee (2019); Stiker (1999); Tremain (2015); Titchkosky (2011).

A critical disability studies scholar asks difficult questions about the possibility of representation and accountability of scholarship and activism to all disabled people. And this new-found criticality questions some of the starting assumptions ...

(Goodley, 2018, 6).

Black Madness:: Mad Blackness furthers the conversations ... by fore-grounding the spaces where Blackness and madness become usefully entangled. My goal, however, is not to unravel them but rather to pinpoint the facets of their intertwining so that we might rest with the knots history and culture have created. This project attempts an intellectual cartography.

(Pickens, 2019, 11).

Let us try to “...expand the definition of consciousness beyond human/animal sentience by locating it in the world - a living agential co-responsiveness to the field itself; where the eye first opens.”

(Manning, 2017, 162)

To take responsibility for one’s writing and/or one’s reading, I would maintain, is to commit oneself to “read and write between the lines,” that is, to resist simplistic interpretations that dogmatically ascribe fixed significances to the words that appear on the printed page. And yet, it is not the case that “anything goes” ...

(Weiss, 2008, 58).

The above vividly demonstrates the grappling with multiple frames and their various parts, uses, interests, such as representation and metaphor; assumptions and questions; knots of history and culture; the intertwining of the perceiver and the perceived; or, what lies between the lines of what has already been read or written. But why grapple with various meanings and movements of interpretation that begin to appear when the focus shifts to multiple frames? Without this grappling, the multiplicity of frames disappear in favor of a one frame fits all reductionist comprehension on any human situation, including disability.

Resisting reductionist interpretations that dogmatically ascribe a fixed singular significance to disability raises the possibility that the unseemly matter of disability as either biological or social problem is, in fact, a complex life that erupts into the field of consciousness giving rise to concerns on a metaphysical scale. To interrogate stock stories and thus to ask questions so as to further conversations about our being co-responsive to the ways that disability is perceived, written and read about, or otherwise experienced, means working within a self-reflexive frame for doing disability studies. Such a frame puts into question some of the starting assumptions of analysis, for example, that we know what disability is and means outside of the knots of history and culture forming the field of consciousness itself where the eye first opens on disability or madness or any other anticipated but unexpected difference of the body, mind, senses, emotions, comportment.

Therí Pickens suggests that this requires an analysis that maps its own terms of engagement and the relations between these terms. This is similar to Gail Weiss’ suggestion that we read, and write, between the lines. It’s not that anything goes or that some kind of “*I think, I feel, I am*” now replaces a concern and engagement with the contested meanings of disability as a cultural phenomenon. Instead, any appearance, any way that disability is said to be, any experience of I think, I feel, I am disabled can be read as part of the “field”, as Manning puts it, from which disability arises and upon which it makes its appearance. Rather than providing facts on disability, such a frame *frames* how disability has come to appear as it does as a way to reveal something of the human condition, including the context of our appearances as disabled (or not).

Disability, moreover, cannot easily be contained within any fixed belief in biology or society since embodiment is a metaphysical issue that erupts into our thoughts, feelings and actions. For Goodley, this gives rise to a need to take into account our starting assumptions, what Manning might call the field where the eye first opens. It is not that we can untie the knots of history and culture that intermesh perception and the perceived, but we can, says Pickens, at least turn back and map or take responsibility, as Weiss puts it, for what grounds our entanglements with meanings already ascribed to disability.

Speaking of his life in blindness and the experience of its changing meanings, Rod Michalko puts these issues in this way:

... whatever is said about blindness, be it said by ophthalmology, rehabilitation, or by blind persons, *must* be heard as a depiction. Depictions of blindness can be judged for their vividness, their clarity, and their need to speak of blindness without being the final word ... Blindness is a reflexive phenomenon insofar as the *living* of blindness reflects a conception of what it is ... through conceiving life as a narration of the story of blindness. ... In whatever way blindness appears, we have a hand in making it do so. Those of us who are blind, and who relate to our blindness in a mature way, know that the appearance of blindness is a story and that we are responsible for its telling.

(Michalko, 1998, 155, 156).

A disability studies that is oriented to framing the frames, oriented to perceiving how and to what ends disability is taken-into-account, engages any depiction of disability as a depiction of life that may be judged as such. It is to regard disability, suggests Michalko, as a reflexive phenomenon where life and conceptions interweave forming a narration whose vividness is what should

be judged. Asking what disability depictions make vivid and clear and what they place into the background or make invisible is to ask what sort of story we are making with and through disability while raising the metaphysical question of judgment and even the question of ethics. In particular, it is to ask what it means to live with one depiction or another.

The framing of frames as a methodological trope for disability studies is not the addition of one more method or research technique. Instead of adding one more frame (or one more model), framing the frame acts to magnify, to bring into existential view, the sense that disability is life and is not best understood as a thing, be it biological or social in origin. The possibility of this self-reflexive engagement with an appearance of disability, in our own lives or in that of others, is to embrace, however tacitly, that disability is something other than dysfunction, needless suffering, and death. Framing that which has framed disability relies on some foundational sense that disability is quintessentially life.

By way of conclusion, I turn now to an examination of “models of disability”, models that themselves find their origin in disability understood as life. I show how these models, while framing the life of disability risk leaving that life behind in favor of emphasizing the life of the models themselves. I do this so as to conclude by sowing the seeds of the question — “What sort of life with disability do disability studies approaches invite?”

Conclusion – modeling the story of disability

The proliferation of disability studies models since the late 1970s can be understood as part of the orientation that *disability is life*, but once or twice removed from this experience, transmogrified into the more acceptable *Enlightened* idea of “model.” The development of the social model of disability, for example, occurred when groups of disabled people in the UK met to discuss employment schemes and disability allowances, the warehousing of disabled people in long term care homes, and their basic exclusion from the possibility of self-governance and participation in the wider social whole. Critiques of the treatment of disabled people in long-term care facilities with little to no control of their eating, sleeping, friendship or love meant that these disabled people began their discussions from the sense that there must be an alternative to how they were framed as useless sufferers (Hunt, 1966). Their critiques of institutionalized versions of disability started from the foundational sense that it could be different since, after all, disability is not only life but also a life worth living.

Discussions of the life of disability gave rise, eventually, to the social model of disability. This model enacted a sharp dividing line between impairment (what happens to one’s body) and disability (what society does to people with impairments). The model became a frame through which to perceive disability as oppression. Oliver (1996, 27, 23) reminds us, though, that “we must not assume that models in general and the social model in particular can do everything ...”, but it can be used to map “all the things that impose restrictions on disabled people.”

The proliferation of models in disability studies continues—social death model, social barriers model, administrative or bureaucratic model, minority model, social role valorization (SRV) model/normalization model, affirmation model, relational model, cultural model, affective model, *DisCrit* model, cultural affective relational model, consumer rights model, independent living model, universal model, the partnership model, and bio/social/spiritual model. The list is indicative rather than exhaustive. Whether one takes up any one of these many and ever proliferating models, it should be understood that a model *models something*, and that something is the beginning point for doing disability studies — namely, the acknowledgment that disability is life. Of course, acknowledging this life does not naturally, nor necessarily have to lead to a model.

While the formation and language of models may drink deeply from the waters of forgetfulness thus making disability a thing, people who use models had first to find life in disability before forging a model of disability. A commitment to any one model is itself a story of the Enlightenment belief that the subject of life (disability) and its objectification (a model of it) can be separated leading to a split between subjective experience of disability and objective knowledge about it *as if* this is a good way to frame life with disability. Regardless of the attempt to keep the split between subject and object² alive, any disability studies model of disability requires something other than dysfunction, death and suffering to ground its possibility.

This short story of disability studies models is meant to illustrate a key point I have been emphasizing throughout — beginning from the orientation that whatever else disability *is*, it is life and a way to re-encounter the interpretive frames that make disability appear as it does. What arises from the recognition of a diversity of frames is not the relativism that anything goes but, instead, the necessity of a self-reflexive politics with disability experiences, inclusive of the ways we do disability studies. It is in this sense that perceiving a diversity of frames can be read as an invitation not only to know them, but also to reckon with the in(ex)clusions each frame enacts, an enactment that invites us to grapple with their social and political consequences (McRuer, 2006). Thus, frames should be judged by their vivid and clear depictions of disability, together with whether or not they invite us to regard them as the beginning and not the final word on understanding disability. Changing the frame of reference means making the vividness of life primary within our accounts of disability. Such an approach does not, of course, mean ignoring social oppression. Instead, it means asking how we relate to, that is, live disability through our narratives of social oppression, or pride, or human rights, or any other frame.

²“Probably the chief gain of phenomenology,” suggests Merleau-Ponty (1945, xxii), “is to have united extreme subjectivism and extreme objectivism in its notion of the world or rationality.” The chief gain of Enlightened science can be characterized as opposite to this.

Frame > Reference > Consequences > Frame

Fig. 2 The interpretive chain.

In this article, I have worked from the assumption that the interpretive chain (see Fig. 2) is one way to proceed in disability studies and it is a way that aims to embrace a complex and lively relation between disability and society.

Such an analysis practices a *restless reflexive return* that resists the temptation to believe in progressive quick fixes as a way to figure out what we have made of disability/life when we say or do or write or read anything of disability (Titchkosky, 2011, 69). Changing the frames then begins with knowing that we are never without a frame and finding a way to engage our frames outside of habit or utility or worse estrangement. A commitment to seeking a way to frame our frames within the life of disability is to be at once vitalized by that life and to contribute to its vitality.

Acknowledgment

Thanks to Canadian Insight SSHRC grant as well as to Elaine Cagulada, Maddy DeWelles, Miggy Esteban, Efrat Gold, Aparna Menon and, as always, Rod Michalko for their engagement, edits and ideas.

References

- Ahmed, S., 2006. *Queer Phenomenology: Orientation, Objects, Others*. Duke University Press, North Carolina.
- Bauman, Z., 2000. *Liquid Modernity*. Polity Press, Cambridge, UK.
- Bauman, Z., 2004. *Wasted Lives: Modernity and Its Outcasts*. Polity Press, Cambridge, UK.
- Chiayulien, 2020. The Concept of Debility. <https://medicalhealthhumanities.com/2020/06/09/the-concept-of-debility/>. (Accessed 13 December 2020). Blog Post.
- Conrad, P., 1992. *Deviance and medicalization: from badness to sickness*. Temple University Press, Philadelphia.
- Corker, M., 1998. *Deaf and Disabled or Deafness Disabled*. Open University Press, Buckingham.
- Davidson, M., 2016. Cleavings: critical losses in the politics of gain. *DSQ* 36 (2).
- Ginsburg, F., Rapp, R., 2013. Disability worlds. *Annu. Rev. Anthropol.* 42 (1), 53–68.
- Goodley, D., 2011. *Disability Studies: An Interdisciplinary Introduction*. Sage Publishing, London.
- Goodley, D., 2018. The dis/ability complex. *Digest* 5 (1), 3–22.
- Goodley, D., 2021. *Disability and Other Human Questions*. Emerald Publishing, Bingley UK.
- Hall, S., O'Hara, J., 1984. The narrative construction of reality: an interview with Stuart Hall. *South. Rev.* 17, 3–17.
- Hughes, B., 2019. *A Historical Sociology of Disability: Human Validity and Invalidity from Antiquity to Early Modernity*. Routledge, London.
- Hunt, P., 1966. A critical condition. In: Hunt, P. (Ed.), *Stigma: The Experience of Disability*. Geoffrey Chapman, London, pp. 135–149.
- Kelly, C., 2016. *Disability Politics and Care: The Challenge of Direct Funding*. UBC Press, Vancouver.
- King, T., 2003. *The Truth about Stories: A Native Narrative*. House of Anansi Press, Toronto.
- Manning, T.D., 2017. The murmuration of birds: an Anishinaabe ontology of mnidoo-worliding. In: Feilding, H., Olkowski, D.E. (Eds.), *Feminist Phenomenological Futures*. Indiana University Press, Indiana, pp. 154–182.
- McRuer, R., 2006. We were never identified: feminism, queer theory and a disabled world. *Radic. Hist. Rev.* 94, 148–154. Winter.
- Merleau-Ponty, M., 1945. *Phenomenology of Perception*. Routledge & Keagan Paul, London.
- Michalko, R., 1998. *The Mystery of the Eye and the Shadow of Blindness*. University of Toronto Press, Toronto.
- Mitchell, D., 2002. Narrative prosthesis and the materiality of metaphor. In: Snuyder, S.L., Brueggemann, B.J., Thomson, R.G. (Eds.), *Disability Studies: Enabling the Humanities*. The Modern Language Association, New York, pp. 15–30.
- Oliver, M., 1996. *Understanding Disability: From Theory to Practice*. St. Martine's Press, New York.
- Pickens, T., 2019. *Black Madness :: Black Madness*. Duke University Press, North Carolina.
- Puar, J.K., 2011. Coda: the cost of getting better: suicide, sensation, switchpoints. *GLQ* 18 (1), 149–158. <https://www.muse.jhu.edu/article/460796>.
- Puar, J.K., 2015. The 'right' to maim: disablement and inhumanist biopolitics in Palestine. *Borderlands E-J.* 14 (1), 1–26. www.borderlands.net.au.
- Schlund-Vials, C., Gill, M. (Eds.), 2014. *Disability, Human Rights, and the Limits of Humanitarianism*. Ashgate Press, Burlington, VT.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922.
- Smith, D., 1978. 'K is mentally ill': the anatomy of the factual account. *Sociology* 12 (23), 23–53.
- Stiker, H.J., 1999. *The History of Disability*. University of Michigan Press, Ann Arbor.
- Titchkosky, T., Michalko, R., 2009. *Rethinking Normalcy: A Disability Studies Reader*. Canadian Scholars/Women's Press, Toronto.
- Titchkosky, T., 2007. *Reading and Writing Disability Differently: The Textured Life of Embodiment*. University of Toronto, Toronto.
- Titchkosky, T., 2011. *The Question of Access: Disability, Space, Meaning*. University of Toronto Press, Toronto.
- Tremain, S., 2015. *Foucault and the Government of Disability*. University of Michigan Press, Ann Arbor.
- Ware, L., 2005. *Ideology and the Politics of (In)exclusion*. Peter Lang, New York.
- Weiss, G., 2008. *Refiguring the Ordinary*. University of Indiana Press, Indiana.
- WHO, 2018. Jan. Key Facts of Disability. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health#:~:text=Over%20a%20billion%20people%2C%20about,health%20conditions%2C%20among%20other%20causes>. (Accessed 5 October 2020).
- WHO, ND. ICF: International classification of functioning, disability and health. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health> (Accessed 20 December 2020).
- Zola, I.K., 1993. Self, identity and the naming question: reflections on the language of disability. *Soc. Sci. Med.* 36 (2), 167–173.

Disability studies in a posthuman age

Katherine Runswick-Cole, Dan Goodley, and Kirsty Liddiard, iHuman, School of Education, Sheffield, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	200
Posthuman disability studies	201
(M)others of disabled children	202
The project: living life to the fullest	202
Developing a posthuman narrative analysis	203
Entanglements with human and nonhuman animals	204
Entanglements with machines	205
Entanglements with environments	206
Conclusion	207
Acknowledgment	207
References	207

Glossary

Artificial Intelligence The theory and development of computer systems able to perform tasks which usually require human intelligence

Cyborg A person whose human abilities are extended by machines connected to the body

Critical disability studies An interdisciplinary community of scholarship, arts and activism that emphasizes the cultural, social, historical, political and economic constitution of disability

Disablism Discrimination against people on the grounds of their disabled identity

Humanism A system of thought that centers the concerns of the “rational” human

Life-limiting and/or life-threatening impairments Conditions which may result in shortened life span or sudden death

Matricentrism Centering on the mother

Narrative A story

Posthumanists Those concerned with the blurring of the boundaries between human and non-humans

Synthetic biology Branch of biology concerned with engineering organisms to have new abilities

Introduction

It has become impossible to think, write, act or deliberate without Covid-19 entering our framings of the world. Scholarship, research, innovation and knowledge generation are, understandably, locked into a direct relationship with the global pandemic. This chapter does not seek to ignore Covid-19, but it equally does not want to be defined by it. Instead, the pandemic might be viewed as an artifact of a posthuman age: where human troubles are tied up with those of the non-human world. The molecular, biological, human and societal are all intimately tied to the global and environmental. And the response to the pandemic has revealed very disparate experiences on the part of different human groups in distinct areas of the world. The individual body is clearly situated in a physical space. And these physical spaces, what we might have understood primarily in terms of locale or place, are increasingly virtualised as a consequence of the centrality of digital life. The 21st Century human being is a curious mix of biological/social, local/global, physical/digital, organic/technological. Posthuman theorizations are a response to this amalgamation; reflecting on the ways in which the human condition has become hybridized and anticipating contemporary and future mixes. In this chapter we will explore what it might mean to live in a posthuman age. We will introduce some key readings and writers and sketch out a representation of what some writers have termed the posthuman condition. Then, we will move on to consider the phenomenon of disability; and its relationship with the posthuman. Drawing on critical disability studies literature we will argue that disability constitutes the quintessential posthuman subject (Goodley et al., 2014): even though disabled people are often ignored in more mainstream articulations of posthumanity. Next, our discussion will turn to some of the potentialities offered by disability to thinking through the posthuman age. Our focus will move specifically to the lives of mothers of disabled children: as we excavate the many ways in which their human and non-human relationships provide possibilities of parenting and supporting their disabled children.

One obvious statement to add here is that posthuman studies are an intellectual response to the posthuman realities of our time. This understanding, though, emphasizes a rather circular view of knowledge production. If we were to accept the philosophical view

that knowledge constitutes the very phenomena about which it writes or speaks, then theorizations are better understood as having constructive qualities. Posthuman theorizations are complicit in the writing of and therefore the constitution of the posthuman world. As we seek to understand our interrelations with technology, animals, environments, robotics, digital contexts and science, then we not only unravel those interrelations but generate new understandings of them that may be incorporated into their very generation. Think, for example, of studying the relationship between young people and social media. The very act of exploring these relationships will ensure that we think anew about their complexities, their offerings, their pros and their cons. Empirical studies into social media not only give new insights, but also create new kinds of reflexivity, awareness and practical understandings about the functionalities, applications and consequences of social media. Posthuman studies are best understood then as offering accounts of—and interventions into—the posthuman age.

There is always a danger of embracing the speculative and sci-fi qualities of the posthuman. Cyborgs, Artificial Intelligence and synthetic biology are but some of the usual cast of subjects that are brought onto the stage of the posthuman. These characters are very seductive: not least because they plunge us into a host of utopian and dystopian tropes associated with the merging of human and science. Of course, these images compel us to think of the future: to anticipate what we might become. But any tracing of the posthuman age should also consider the present. And in the current epoch we might want to recognize the more mundane elements of the posthuman condition.

Posthuman disability studies

A posthuman world view is committed to recognizing and challenging the European renaissance philosophy of humanism (Liddiard et al., forthcoming). Humanism is predicated upon the fundamental desirability of the bounded, rational, autonomous and sovereign human subject, and is driven by an absolute certainty that human beings are capable of maintaining their privilege and species-dominance through rational thought, self-governance and progress toward objective knowledge (Goodley et al., 2014). The price of such progress has been paid by those who fail to live up to the desired humanist ideal and this has led to the creation of human, non-human and environmental collateral damage (Goodley et al., 2014). A posthuman approach encourages us to embrace ways of life that are not rigidly framed by humanistic values of independence and autonomy (Whitney et al., 2019). The posthuman condition opens up the possibilities for desires that are expansive, relational and nomadic, rather than narrow, individualized and fixed (Whitney et al., 2019; Braidotti and Regan, 2017). Crucially, in critical disability studies, posthuman thinking demands us to engage with *challenges* (barriers and inequalities) but, also, to think about *possibilities* in the lives of disabled people and families. A posthuman perspective draws us toward important relationships, resources and connections between humans and non-human animals, machines and environments (Goodley et al., 2014). It is a position that requires “the knowing subject disengaging itself from the dominant normative vision of the self” becoming “relational in a complex manner that connects it to multiple others” (Feely, 2019: 167). Making sense of these relationships, resources and connections demands that we develop an analysis that is sensitized to a “more-than-human” perspective (Monforte et al., 2019) as we explore what new insights can be gained, and how change can be affected, by thinking through (m)others’ and children’s entanglements with human and non-human animals, machines and environments in stories about their lives.

In this chapter, we expand the project of posthuman disability studies by thinking again about disability and family (Feely, 2019; Goodley et al., 2014). Thinking with posthuman politics and theory, we set out to explore the immaterial aspects (culture, attitudes and meanings) and material realities (structures, barriers and objects) in the lives of (m)others of disabled children. Writing from the United Kingdom, and in the midst of a global pandemic which has disproportionately impacted on the lives of disabled children, young people and their families (Runswick-Cole et al., 2020), we are acutely aware of the plethora of existing research focusing on the lives of mothers of disabled children. Despite this, over the last 30 years there have been limited im/material changes effected in their lives (Runswick-Cole and Ryan, 2019). The political and material conditions of disablism (the institutionalized rejection of disabled people) continue to limit the lives of disabled children, young people and their families (Thomas, 2007). We are interested, then, in how a posthuman orientation can help us to think differently about the material (e.g., body and environment) and immaterial (e.g., emotions and virtual environments) relations in the lives of (m)others of disabled children. Above all we are interested in how posthuman thinking can be used to effect (political) change.

Our interest in posthuman disability studies has been driven by our collaborative work on the research project: *Life, Death, Disability and the Human: Living Life to the Fullest* (hereby *Living Life to the Fullest*) carried out in England from 2017 to 2020. Through engagement with narratives and the arts, *Living Life to the Fullest* sought to forge new understandings of the lives, hopes, desires and contributions of children and young people with life-limiting and life-threatening impairments (LL/LTIs) and their families. Too often, knowledge production about the lives of children and families living with LL/LTIs is dominated by medicalized, individualized and pathologized accounts, reflecting the concerns of professionals, rather than the views of children, young people and families themselves (Curran and Runswick-Cole, 2014). Our work seeks to respond to these dominant accounts, by co-creating narratives with disabled young people and with family members themselves (Ryan and Runswick-Cole, 2008; Runswick-Cole and Ryan, 2019; Whitney et al., 2019).

In brief, our aim, in this chapter, is to offer new understandings of the stories of (m)others of disabled children through a posthuman thematic narrative analysis. In our analysis, we explore (m)others’ entanglements with human and non-human animals, machines and environments. In doing so, we hope that an expansion of the posthuman disability studies project (Goodley et al., 2014) and method can inspire (political) change.

(M)others of disabled children

We use the term (m)others here, not to reinforce the gendered nature of caring roles in the global North, but to pay attention to the fact that the vast majority of parental labor is still carried out by mothers. The bracketing of the “m” is used to acknowledge that these roles are taken by others too: men, women and those who refuse such binary identities (Douglas et al., 2021). We follow Gabel (2018: 554) in understanding (m)othering as “... situated within the institution of motherhood and numerous other structures or institutions (for example, class, race, marriage, patriarchy) and ‘mother’ as a subject position, a locus point in power networks.”

Research focusing on the lives of (m)others of disabled children has typically been shaped by humanist assumptions leading to a research focus on the “burden” of parenting a child who fails to conform to the ideal of the bounded, rational, autonomous and sovereign human subject. (M)others are, then, typically, cast as being in a state of “grief” about the birth of their disabled children and in need of a process of adjustment mediated by psy-professionals (Bruce and Schulz, 2002). These dominant narratives have prevented a move beyond a position of negativity and have obscured other, more positive, dimensions of (m)othering a disabled child (Ryan and Runswick-Cole, 2008).

While there has been a shift in research toward more positive and affirmative aspects of (m)othering disabled children, tensions remain. Indeed, traditionally, both disability studies and feminist theory have been reluctant to engage with matricentrism (Douglas et al., 2021). Feminists fear that any discussion of mothering will inevitably lapse back into essentialism inadvertently reinforcing the position of women as primary carers (O’Reilly, 2016). Disability studies, on the other hand, has a tradition of being preoccupied with separating the personal from the political domain (Oliver, 1990) and, in doing so, has pushed (m)othering to the margins (Ryan and Runswick-Cole, 2008; Runswick-Cole and Ryan, 2019; Douglas et al., 2021).

A posthuman approach can help to move beyond these tendencies to fix (m)othering in the realm of the individual and psychological. By paying attention to the self, becoming through complex relationships with multiple others (Feely, 2019: 167), the sharp focus on the mother-child dyad as the site of observation, surveillance and remediation can be softened and blurred to enable us to think more expansively and relationally about children and (m)others. Posthuman (m)othering draws attention to the distributed nature of care (Booth and Booth, 1998) and complexity of im/material flows at play in shaping (m)othering practices. Families of disabled children have resisted the individualizing and pathologizing tendencies of the humanist tradition; they have demonstrated the importance of collective support in enhancing the lives of children rather than dwelling on individualized autonomy (Ryan and Runswick-Cole, 2008; Runswick-Cole and Ryan, 2019). In posthuman times, (m)othering disabled children emerges as an exemplar of a posthuman relational ethic that demands that we think again about how we are all made through our connections with others (Goodley et al., 2014). A posthuman orientation removes “the obstacle of self-centered individualism to embrace an *ethics of becoming*” and “an enlarged sense of inter-connection between self and others” (Braidotti, 2013: 49).

The project: living life to the fullest

Living Life to the Fullest: life, death, disability and the human is an interdisciplinary arts-informed UK Economic and Social Research Council funded research project that was carried out in England between 2017 and 2020. Relationships with disabled young people have been at the heart of the research project and processes. Early in the research project, virtual spaces emerged as accessible spaces through which young people could both engage with, and, crucially, shape and lead the research process. We established a Co-Researcher Collective which included disabled young people aged between 19 and 30 years. Working online (and occasionally off-line), the Collective have:

- Supported research design through discussion (planning both narrative and arts-informed approaches);
- Co-designed interview schedules for young people and parent participants;
- Recruited participants;
- Planned the project’s impact strategy by building relationships with partner organizations;
- Worked with community research partner organizations;
- Met regularly via the Research Management Team to co-manage the research process;
- Written blogs and made podcasts and films that communicate and document processes and findings;
- Presented at academic conferences;
- Undertaken various public engagement and knowledge exchange activities;
- Co-authored articles for publication and a book about the project (Whitney et al., 2019; Liddiard et al., forthcoming).

This relational approach to research is central to a posthuman ethic. And yet, although young people were involved in the design of the interview schedule for parents, they were not involved in the interviews themselves. Aware of the potentially sensitive nature of the interviews, we checked with parents in advance to see if they would be happy to be interviewed by disabled young people—all of the parents declined and asked to be interviewed by a member of the University staff team. Therefore, Kirsty and Katherine carried out all the interviews/focus groups with parents. Parents expressed worries about discussing their lives with young people, they worried that they would not be able to talk freely or that they would be burdening young people with their responses. This was a difficult moment for the project team that had begun with a commitment to young people as research leaders and collaborators. And yet, it became clear that an ethic of care meant that we had respect for the parents’ decision to be interviewed by university staff

members. In line with this ethical concern, we also agreed with the disabled young people researchers that we would not share the parents' full transcripts with them. As a result, this paper has been co-authored only by the university staff team. We acknowledge the tensions in this decision-making process and the messiness of working ethically in relational ways with young people and (m) others.

Developing a posthuman narrative analysis

We follow a tradition in critical disability studies in turning to narrative analysis to think about the lives of (m)others of disabled children (Smith and Sparkes, 2008). Stories construct our lives and ways of being; they are part of being (post) human. Crucially, stories can be more than windows on the worlds of others, they can be a call to social action (Smith and Sparkes, 2008). Parents all identified themselves as the parents of children and young people who had life-limiting or life-threatening impairments. The stories re-told here were collected by Kirsty through on-line interviews with two mothers and one father. Technology allowed for increased accessibility for (m)others with busy lives and access to technology. We note that access to technology was a barrier to some (m) others' participation and we carried out a person-to-person focus group with (m)others to try to mitigate against this.

The story-tellers in this paper are:

- Eric is father to 24 year-old Alison, who lives at home with the family. He works 60 h a week as a carer for Alison, who works full time;
- Ruby is mother to Matthew who died 1 year before the interviews took place, aged ten. She works part-time.
- Karen is a single mother to Jason, aged twenty-four, who lives at home with her. Karen works full time.

Narrative interviews were carried out on on-line platforms at a time of the participants' choosing and lasted no longer than an hour and half. Participants were offered regular breaks, though none asked to pause the interviews. Interview questions were open, beginning with "can you tell me a little bit about your son/daughter?" Rather than follow a strict individual schedule, each interview took a different path, followed-up with prompts and probes asking questions about chronology: "when did that happen?"; affect "how did you feel about that?" and offering affirmation: "okay"; "I see." Throughout the interviews the interviewer sought to engage in active listening: to be attentive, responsive and curious (Smith and Sparkes, 2008). The questions encouraged the (m)others to weave their disparate experiences together to build a story (Monforte et al., 2019). At the center of the approach was an aspiration to behave sensitively, accommodating the participants' needs and engaging in an ethic of care. As a team we also took seriously the responsibility of care we had to one another and we were able to discuss the stories we collected throughout the research process. Our intimate connections to disability and (m)othering meant that the stories touched us all in different ways, at different times, in the project and we needed to attend to this carefully throughout.

In line with traditional approaches to thematic and narrative analysis (Braun and Clarke, 2006; Clandinin and Connelly, 2020), we began by reading and re-reading the transcripts. However, we did not follow a stringent step-by-step approach to coding and thematizing the data, rather we drew on our posthuman sensitivity to read for the human and more than human entanglements within the story. As we read and re-read the transcripts, we paid particular attention to the moments when the story tellers told us about how their lives were caught up with:

- (1) **Human and non-human animals:** other human beings, including their children, family members and practitioners and with animals;
- (2) **Machines:** this included mundane technology, such as smartphones and iPads, as well as specialist equipment, like power chairs and hoists, and medical technology, such as ventilators and feeding tubes.
- (3) **Environments:** understood here, as encounters with im/material objects and systems which presented barriers or offered opportunities in their lives;

Reading the stories in this way allowed us to develop a posthuman thematic narrative analysis in which themes are generated by paying attention to entanglements with human and non-human animals, machines and environments in the stories.

At this point, we want to pay attention to our choice not to engage in assemblage analysis (Feely, 2019) which may at first seem to fit well with our commitment to posthuman theory. In Feely's (2019) article, he offers a comprehensive guide to assemblage analysis in which he tracks material-semiotic forces, and flows within an assemblage. He describes an assemblage by saying:

An assemblage should not be understood as a fixed entity because it is always in a process of becoming. Its diverse components come together over time; work together for a time to produce something; and, in time, will fall apart. For example, the disability service assemblage's components (e.g., buildings, bodies, policies) came together over time, are currently working together to produce the service, and will eventually fall apart.

Feely (2019: 6).

In the final stage of the analysis, Feely (2019) examines the processes of reterritorialization and deterritorialization within the assemblage which allows him to pay attention to the dual process of stabilization and destabilization in the assemblage as the elements in the assemblage act upon one another.

This assemblage approach understands social problems as produced by human and nonhuman actors, and discursive and material forces and is underpinned by a flattened ontology where all entities in the assemblage are all assumed to be in process rather than fixed. They are ceaselessly becoming at different rates of speed and slowness and are all worthy of equal attention (Feely, 2019). And here, despite our posthuman leaning, is where we depart from assemblage analysis. We are interested in the material-semiotic forces flowing in the stories, we see the value of paying attention to the human and more than human elements in the story, and yet, in our readings we find ourselves inevitably drawn to the human actors (Brinkman, 2019). We recognize that the draw to the human risks falling into the trap of anthropocentrism and valorizing the very humanist human that we critique above. And so, our aim is to enact a posthuman thematic narrative analysis that might enable us to do two things:

1. In Brinkman's words, "[t]o stay human without being anthropocentric" (2019: 137). That is, we can allow our attention to be drawn back to the human in the narratives and yet still resist any call to human dominance over nature;
2. Hold onto the notion that the "human being is not an ahistorical, disembodied and universal intellect, but rather a historical, embodied, affective creature that lives in a sociomaterial world of flux and uncertainty" (Brinkman, 2019: 136), allowing us to pay attention to the more-than-human in the stories.

In the sections below, we explore, in turn, the families' entanglements with human and non-human animals, machines and environments. While we present each of these themes separately, we acknowledge that human and non-human animals, machines and environments are intimately entangled with one another. In this act of separation, we resist the temptation to silo or to fix the data and recognize that the themes speak across one another in their entanglements.

Entanglements with human and nonhuman animals

The stories were written through entanglements with human and non-human animals, that is, in each story (m)others told their stories through their relationships with other humans and with animals.

Not surprisingly, the (m)others all spoke, at length, about the deep connection they had with their children. All the (m)others were or had been carers for their children; this included carrying out care not usually associated with parenting, including complex personal and medical care. While care is often storied as a drain or a burden (a challenge or a barrier), the (m)others spoke about care as bringing them closer to their children (as potential). Eric noted that he spent more hours a week caring than engaging in his paid work, and for him, care was a joy:

... in exchange for the 60 hours of so week's work that I put in, it's you know, it's been amazing, I think, for me—a benefit and a joy to be able to be a close part of [my daughter's] life

Ruby, too, talked about the care she gave her son, and the connection between them. The depth of the connection between them was brought into sharp focus following Matthew's death:

I gave up everything and put everything into Matthew. I lost so much of myself and afterwards, not so much now because I think I have kind of untangled myself a little bit from it, but at the start I felt like—it is really difficult to explain—but I just felt like I was tangled in Matthew. I just, I couldn't, I didn't know—people said to me "what do you want to do now?"—"I don't know"—and I didn't know anything about myself, I couldn't have told you what my favourite food was, what my favourite colour was

In explaining that Ruby was "tangled in Matthew," she echoes the language of entanglement that frames our posthuman narrative analysis. Reflecting back on her relationship with Matthew, Ruby beautifully articulates the closeness of the relationship between mother and son, and the pain of his death.

The (m)others also spoke of how their disabled children had brought them into close relationships with other people. Eric described how Alison had brought new and unexpected relationships and understanding into his life:

I guess really the world of people with disabilities which I would never have had an appreciation of without has been, you know, something that for me has been a really important thing for me to understand.

Karen also recognized the connections made between her and other families of disabled children through impairment specific conferences and support groups saying that:

I wouldn't say I have lost any friends from it at all, I have gained a whole new tribe of global people all across the world.

Ruby described the close relationship that she and Matthew had formed with a nurse who had looked after Matthew since he was a baby.

I think a plus for us was that the nurse that was looking after us [in the specialist unit on the night he died] had known Matthew since he was about 6 months old and first transferred to the children's ward at [the local hospital]. She had looked after Matthew on her last day at the [local hospital] before she had moved to [the specialist unit] and we were friends on Facebook. [The night he died] she had come and looked after Matthew. She had waited [with him] until they had put him in the, in the freezer bit, and she said she hadn't left his side until she literally had to. She said like she had brushed his hair to the side and gave him a little kiss and stuff like that so, it was just those little things I think that makes such a massive difference.

As Ruby describes the relationship with the nurse, we see the importance of the relationships between them and the ways in which their lives were entangled from when Matthew was 6 months to 10 years when Matthew died. In the story, it is impossible not to be drawn to the (unnamed) nurse's very human act of brushing Matthew's hair and giving him a kiss, at the same time as we see the entangled relationship between Matthew, his mother and the nurse, as an exemplar of posthuman relational ethic revealing how we are all made through our connections with others (Goodley et al., 2014).

Elsewhere we have written about the positive relationships between young people and animals, reflected in their accounts of the roles that assistance dogs play in their lives, but also in their day-to-day relationships with animals as companions (Whitney et al., 2020). The (m)others too spoke about their children's relationships with animals during recreational activities such as horse riding. However, they also spoke about animal testing as a means of developing new drugs to treat life threatening and limiting impairments. Such drug trials offered them "hope." And yet a posthuman ethic demands us to question a utilitarian approach in which animals' lives are viewed as collateral damage in the name of (human) progress (Braidotti, 2013). The story brings us up short and offers us a stark reminder of the risks of romanticizing the relationship between human and non-human animals in the lives of disabled children and young people.

Entanglements with machines

(M)others' accounts were also peppered with descriptions of their children and their relationships with machines. Karen remembered how "Jason went off his feet at thirteen, first into a manual wheelchair, eighteen into a powered chair." Eric described key moments in Alison's life in relation to technology—Alison got her first electric wheelchair at four and became "ventilator dependent" some 12 years later. Matthew came home from hospital as a baby with: "24 h oxygen, feeding tube, loads of meds, we had a blood pressure monitor, a SATS monitor." This scene draws us to thoughts of cyborgs. Haraway (1990: 191) tells us that:

A cyborg is a cybernetic organism, a hybrid machine and organism, a creature of social reality as well as a creature of fiction.

Hybrids and cyborgs allow us to think about human potential in a postmodern world, saturated with technology (Goodley, 2014). Cyborg potential lies in its uncertainty and multiplicity; crucially "[t]he cyborg is a kind of disassembled and reassembled, postmodern and collective self" (Haraway, 1990: 205). Crucially for us here, the cyborg makes us think about interconnections between machines, but also with and between human beings as it opens up the possibility to think more expansively about bodies and relationships (Goodley and Runswick-Cole, 2011).

Alongside this focus on specialist equipment in the mothers' accounts, the plethora of day-to-day gadgets and devices that so many young people plug into, including smartphones, tablets and gaming machines, were also central to the lives of disabled young people. Eric describes Alison: "when she is not in her wheelchair she will just be sitting comfortably on the sofa with her gadgets and devices." We hear that Jason is "big on gaming and would always take the PlayStation to college and then bring it back," but he is also on a clinical trial which means that "every 3 months we go into [city] and they plug in our little home device to Switzerland and take that data." Specialist equipment and medical technology blends with gadgets, play stations and smartphones and becomes undifferentiated in families' lives.

Low tech also plays its part in family life. We learn that Matthew's arms were gaffer taped onto a pod on skis to protect them so he could experience the speed of flying down the slopes which he "absolutely loved." Low tech solutions were important for inclusion. Ruby described how Matthew's support staff would "always find a way around things—they would cling film his arm up, wrap it up in a plastic bag, sit him in a paddling pool next to the swimming pool where all his friends were swimming so he was still *included* experiencing similar to them, he loved water as well."

(M)others too were plugged into machines using social media as a crucial source of information and support. Karen described the connections social media had allowed her to make:

I was talking to a mother in Australia at stupid o'clock this morning and her son is really in end stage and she is out in the complete hoolies and of course it is 35 degrees there now and it is his birthday and everything and she was planning all sorts of things and you think "who else has she got to talk to?"—she is talking to me.

Ruby remembered the time before social media and the positive change it had brought to her life connecting her to other (m)others:

I think so yeah, I definitely felt more isolated in the early years erm again there was sort of no social media then so I didn't really know anyone in a similar situation and it was just kind of me ... but then social media came in and I think played a massive part in connecting parents that had children of a similar nature and you would kind of get support from each other.

The families' entanglements with machines range from the mundane to the life-saving, from entertainment and recreation to treatment and care. The cyborg, therefore, offers us possibilities for reconceiving notions of desire by and between human subjects (Shildrick, 2009), but as we explore these possibilities, the stories also remind us to resist the temptation only to celebrate cyborgification in the lives of disabled young people and their families.

Machines fail. Technological failings have massive impacts on children and young people's and families' lives. Ruby told us of just one example: "a tube had got caught in the lining of [Matthew's] stomach so they had to open him up fully, take part of his stomach away." And when Matthew died, it was important to Ruby that Matthew was disentangled from the machines that had, only a few moments ago, been keeping him alive:

the night he died ... they turned all the monitors off so that nothing was sort of alarming left right and centre ... they had taken out all the lines they could ... they had washed him and put him in some PJs and things like that they'd moved him over slightly in the bed so I could get in the bed with him.

Cyborg desires (Goodley, 2014) are shifting, uncertain, and sometimes, fade away.

Above all, these everyday (and not so everyday) entanglements with machines remind us that all children and young people and their families are becoming (more than human) humans (Brinkman, 2019).

Entanglements with environments

Here, we think of environments as encounters with im/material objects and systems which presented barriers or offered opportunities in children and young people's lives. Sadly, material barriers persist:

We are always finding the challenges of places not being accessible. Nobody is there to get us off the train, the lift is broken, you know that sort of thing (Eric).

Decades on the challenge remains, disabled people are still having to advocate for the removal of (material) barriers to their participation (Oliver, 1990).

At the same time, there were examples of possibility. Karen told us about when Jason arrived at his new school which he experienced as a barrier-free environment:

Well, I remember Jason's exact words was—"well everybody has got a problem [at this school] so [disability] is not a problem."

While material environments continued to present barriers in the families' lives so too did disabling systems (Runswick-Cole, 2010). Ruby describes the financial shock inflicted on the family by the social security system:

... then literally the day Matthew died everything stopped—housing benefit, council tax benefit, carers allowance, tax credits, everything—so we basically went from having, from being sort of on that lower end of comfortable to literally begging, borrowing and stealing just to pay our bills.

Finally, (m)others were acutely aware of their role in providing enabling environments for their children and their fear of what would happen if they could no longer take that role:

I guess the biggest worry will be—will there be somebody there who will provide the safety net, who cares and will do the things that Alison needs to the standard (that you do) that we have always done them (Eric).

Conclusion

In this chapter, we set out to explore (m)others' narratives through drawing on a posthuman approach to develop a thematic narrative analysis to explore relationships between human and nonhuman animals, machines and environments. We began with a political commitment to (m)others, children and young people and the need to challenge the material and immaterial conditions of disablism. Our inquiry is driven by the question of how posthuman thinking can be used to affect (political) change in their lives. As we end the chapter, we return to that question.

A posthuman reading of the (m)others' stories demands us to pay attention to a relational ethic that enables us to think again about how we are all made through our connections with others. By paying attention to the (m)others' entanglements with their children and with other human and non-human animals, machines and environments, the stories reveal processes of selves being made in relationships (Feely, 2019). This is crucial, because this recognition of the inescapability of relationality has (political) consequences for how we understand (m)othering. (M)othering is not made in the actions of isolated individuals but in the relationships between people. Recognizing this relationality has huge implications for the ways in which we might understand parenting, child development and support, as a collective action, rather than an individual, mother-blame, or mother-shame, game (Gabel, 2018; Ryan, 2017; Runswick-Cole and Ryan, 2019).

The (m)others' narratives have also forced us to revisit the project's engagement with young people's relationships with animals (Whitney et al., 2019). We continue to recognize the importance of those relationships in families' lives and the mutual support, care and companionship these relationships signify. And yet, we need also to keep in mind a posthuman ethic which demands us to confront the uncomfortable political reality that means that animals' lives can still be viewed as acceptable collateral damage in the name of (human) progress (Braidotti, 2013).

Entanglements with machines in families' lives reveal the mundanity of the cyborgification of family life as low tech and high tech sit side by side enabling entertainment, recreation, treatment and (life-saving) care. And, crucially, the cyborg emerges as an object of desire and disavowal. A recognition of the complexities of the relationships with machines, and the material and affective flows they generate, is crucial to understanding family life. The stories reveal the persistence of disabling environments in families' lives and a persistent failure of political will to address inequality. This harsh reality meant that (m)others imagine future environments without them in it as frightening spaces for their children.

Finally, in order to bring about (political) change, we end with a call for a relational activism inspired by a posthuman ethic to expose and to challenge disablism. (M)others accounts of online activism offer increasing potential for advocacy and activism (Runswick-Cole and Ryan, 2019) and yet the pace of change is too slow, and the shift to the digital continues to marginalize many. The (m)others' stories trace the intricacies of relationships between human and more than humans and demand a (political) response which recognizes and celebrates interdependencies and disability activism inspired by a renewed commitment to collective action (Runswick-Cole and Goodley, 2017).

Acknowledgment

We thank the Co-researcher Collective for their expertise, skill and commitment and the parents for sharing their stories with us. This work was supported by the Economic and Social Research Council under grant ES/P001041/1, *Life, Death, Disability and the Human: Living Life to the Fullest*.

References

- Booth, T., Booth, W., 1998. *Growing Up With Parents Who Have Learning Difficulties*. Routledge, London.
- Braidotti, R., Regan, L., 2017. Our times are always out of joint: feminist relational ethics in and of the world today: an interview with Rosi Braidotti, women. *A Cult. Rev.* 28 (3), 171–192. <https://doi.org/10.1080/09574042.2017.1355683>.
- Braidotti, R., 2013. *The Posthuman*. Wiley, London.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qual. Res. Psychol.* 3 (2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Brinkman, S., 2019. Can we be human after posthumanism? In: Denzin, N., Giadina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative and Methodological Reflections*. Routledge, New York and London, pp. 126–138.
- Bruce, E., Schulz, C., 2002. Non-finite loss and challenges to communication between parents and professional. *Br. J. Spec. Educ.* 29 (1), 9–13.
- Clandinin, J., Connelly, M., 2000. *Narrative Inquiry: Experience and Story in Qualitative Research*. Wiley, San Francisco.
- Curran, T., Runswick-Cole, K., 2014. Disabled children's childhood studies: an emerging domain of inquiry? *Disabil. Soc.* 29 (10), 1617–1630.
- Douglas, P., Runswick-Cole, K., Ryan, S., Fogg, P., 2021. Mad mothering: learning from the intersections of madness, mothering and disability. *J. Lit. Cult. Disabil. Stud.* 15 (1), 39–56.
- Feely, M., 2019. Assemblage Analysis: An Experimental New-Materialist Method for Analysing Narrative Data. *Qualitative Research*. Online at: <https://journals.sagepub.com/doi/abs/10.1177/1468794119830641>. (Accessed 7 July 2020).
- Gabel, S., 2018. Shatter not the branches of the tree of anger: mothering, affect and disability. *Hypatia* 33 (3), 553–568.
- Goodley, D., Runswick-Cole, K., 2011. Celebrating Cyborgs: Photovoice and Disabled Children. Paper presented at Researching the Lives of Disabled Children and Young People: With a Focus on Their Perspectives. ESRC Research Seminar Series, January 21, in the Norah Fry Research Unit Bristol University, in Bristol, UK.
- Goodley, D., Lawthorn, R., Runswick, C.K., 2014. Posthuman disability studies. *Subjectivity* 7, 342–361. ISSN 1755-6341.
- Goodley, D., 2014. *Disability Studies: An Interdisciplinary Introduction*. Sage Publications Limited, London.
- Haraway, D., 1990. A manifesto for cyborgs: science, technology and socialist feminism in the 1980s. In: Nicholson, L.J. (Ed.), *Feminism/Postmodernism*, pp. 190–233.
- Liddiard, K., Whitney, S., Vogelmann, E., Watts, L., Aimes, C., Evans, K., Spurr, R., Goodley, D., Runswick-Cole, K., forthcoming. *Living Life to the Fullest: Disability, Youth and Voice*. Bingle: Emerald Publishing.

- Monforte, J., Smith, B., Pérez-Samaniego, V., 2019. "It's not a part of me, but it is what it is": the struggle of becoming en-wheeled after spinal cord injury. *Disabil. Rehabil.* <https://doi.org/10.1080/09638288.2019.1702725>.
- Oliver, M., 1990. *The Politics of Disablement*. Macmillan, Basingstoke.
- O'Reilly, A., 2016. *Matricentric Feminism: Theory, Activism & Practice*. Demeter Press, Bradford, ON.
- Runswick-Cole, K., Goodley, D., 2017. "The Disability Commons": re-thinking motherhood through disability. In: Runswick-Cole, K., Curran, T., Liddiard, K. (Eds.), *Palgrave Handbook of Disabled Children's Childhood Studies*. Palgrave, Basingstoke.
- Runswick-Cole, K., Ryan, S., 2019. Liminal still? Unmothering disabled children. *Disabil. Soc.* 34 (7–8), 1125–1139. <https://doi.org/10.1080/09687599.2019.1602509>.
- Runswick-Cole, K., Goodley, D., Liddiard, K., 2020. Human Disposability in England. Available online at <http://ihuman.group.shef.ac.uk/human-disposability-in-england/>. (Accessed 7 July 2020).
- Runswick-Cole, K., 2010. Living with dying and disablism: death and disabled children. *Disabil. Soc.* 7 (1), 813–826.
- Ryan, S., Runswick-Cole, K., 2008. Repositioning mothers: mothers, disabled children and disability studies. *Disabil. Soc.* 23 (3), 199–210.
- Ryan, S., 2017. *Justice for Laughing Boy*. Jessica Kingsley, London.
- Schildrick, M., 2009. *Dangerous Discourses of Disability, Subjectivity and Sexuality*. Palgrave Macmillan, London.
- Smith, B., Sparkes, A.C., 2008. Narrative and its potential contribution to disability studies. *Disabil. Soc.* 23 (1), 17–28.
- Thomas, C., 2007. *Sociologies of Disability, "Impairment," and Chronic Illness: Ideas in Disability Studies and Medical Sociology*. Palgrave, London.
- Whitney, S., Liddiard, K., Goodley, D., Runswick-Cole, K., Vogelmann, E., Evans, K., Watts, L., Aimes, C., 2019. Working the edges of posthuman disability studies: theorising with young disabled people with life-limiting impairments. *Sociol. Health Illness* 41 (8). Early View.
- Whitney, S., Bakalov, N., Liddiard, K., 2020. *The Canine Care Project: A Report Into Disabled Young People's Experience of Their Assistance Dog*. Available online at <https://livinglifetothefullesttoolkit.com.files.wordpress.com/2020/02/canine-care-project-final.pdf>. (Accessed 7 July 2020).

From capacity to ability to automation: “Western” conceptions of the figure of man and ableist subjectivities

Bernadette Baker, University of Wisconsin-Madison, Madison, WI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	209
Man1: capacities to unfold into the gentleman of the social contract	210
Locke on nature	211
Man2 and ability: distinguishing one's self from peers	212
Gall on matter	213
Man3 and automation: the non-necessity of human ability-as-skill under life 3.0?	215
Conclusion: who-what?	217
References	217

... the struggle of our new millennium will be one between the ongoing imperative of securing the well-being of our present ethnoclass (i.e., Western bourgeois) conception of the human, Man, which overrepresents itself as if it were the human itself, and that of securing the well-being, and therefore the full cognitive and behavioral autonomy of the human species itself/ourselves.

Wynter (2003, p. 260).

Introduction

The 20th century “West” has been marked by an intensification of abilities that the 21st century is now automating, throwing into question the entire project of compulsory schooling and tertiary education as cauldrons of ability-production. The challenges posed via machine learning, affective programming, brain-machine interfaces, and the development of whole-body robotics, to name a few, potentially undermine the regulatory, selectivity, distributive, and expert functions of such institutions. New technologies developed largely outside of education are rewriting the definition and need for skills within it, such as memorization, data analysis, and even “learning” (Baker and Wang, 2020; Cooper Ramo, 2016). On the one hand, this longer-term existential challenge to state-sponsored formal education may be seen as good riddance, as an overdue reckoning of the harm that such institutions have delivered to populations who have been minoritized, marginalized and made the most vulnerable via disciplinary societies. On the other hand, no legitimate substitute for education’s babysitting function and false meritocracy has yet emerged, in part because the functional differentiation of the labor force is still in operation as the tail wagging the curricula dog.

More than simply institutional reorganization, however, Sylvia Wynter has posited that the overrepresentation of the human by a particular Christocentric and then biologic “figure of Man,” indebted in its shift to colonialism and racialization, has put the continued “autonomy of the human species” at stake. Wynter’s analysis captures the tensions and dynamics on a broader scale of human existence as opposed to specific reforms debated in education that assume the perpetuity of a human species in its current form. Her historical account, while not dedicated solely to ableization discourse, brings to a point the current intensity of its entwinement with themes of existence, nature, state-crafting, racialization, and violence that have interlaced educational discourse. This chapter thus takes as its springboard, guide, and point of departure Wynter’s historical analysis of the shifting figure of Man and its implications for the (re)formation of ableist subjectivities.

For Wynter, Man1 (Christocentric, Latin-based) emerged under medieval Christianity’s horizon-setting around the supernatural, followed by Man2 (biologic, secular, degodded) who arose in part via Renaissance humanism’s pursuit of scientific materialism that endures through to the present. The analysis is supplemented here with the appearance of man3: (small “m”) the attendance today to “computational superintelligence” via the advent of technologies associated with contemporary intersecting innovations such as artificial intelligence (AI), neuroscience, and big data. As such, the chapter examines how the historical figure of Man has been reconstituted via discourses that shift from capacity to ability, and more recently, automation, especially of apparently “cognitive” functions.

The sections below outline some historical trajectories that contributed to the availability of shifting ableist subjectivities.¹ In an almost prophetic analysis, Campbell (2000) argued that ableist subjectivities are formed and reformed at the heart of eugenic legacies which endured post-WWII, new forms of late 20th century technologies including genetic ones, and the law. Together, these reinforce “disability” as negative ontology and non-voluptuous, recirculating a desire “to-be-able” to do, within normative conceptions of appearance, skill, talent, and mental life. Participating in activities that today would position one as performative, contributory, and beautiful became under this formula the new inscription of the competent citizen and/or cosmopolitan globalist

¹The problem of re-entry when operating within what Deleuze calls the classical image of thought is obvious—any “rethinking” is amenable to the same charges of ableism and adultism that underwrite developmentalist versions of ableist subjectivity. For an extended discussion of this see Baker (2013). For an extended discussion of systems theory in sociology see Luhmann (1995) and Eva Knodt’s introduction in that volume.

(Campbell, 2013). Ableist subjectivity is not an eternal and ahistorical category, then, but a production of so-called modernity, an incredibly powerful one that arguably organizes and/or merges with other forms of -ism in rationales for the biopolitics of privilege, exclusion, disqualification, and life/death (Baynton, 2001; Foucault, 1979). This chapter will illustrate some of the conditions of possibility for its formation in a “Western” philosophical canon not by focusing on “disability” per se, but on discourses that promoted the normalization of human development and humanhood (Weheliye, 2014), pointing to implications of recent technological shifts for conceptions of “the human” marking educational discourse.²

The examples are drawn from what can retrospectively be labeled whitestream scholarship published in places now referred to geopolitically as England, Germany, and the United States from across the 1600s, 1800s and 2000s, forming part of an apparent canon of Western onto-epistemologies. Such scholarship helped lay the groundwork for mainstream educational theories that subsequently had to be challenged and fought against from more critical and post foundationalist perspectives. From Lockean empiricism to Gallian materialism to Tegmarkian Life 3.0 based on artificial intelligence (AI), the figure of Man has been reconstituted based on different premises about what he is, and what he is made of, each time. Such shifts change the nature and the stakes for educational fields ensconced historically in humanism, forcing reconsideration of things such as genius, disability, and eventually superintelligence in ways that today point to predictable forecasts of more-of-the-same injustices versus a more radical onto-epistemological revisioning.

Man1: capacities to unfold into the gentleman of the social contract

Wynter identifies two epochal shifts that launched a figure of Man as an overrepresentation for “the human.” The first was the idealized world that spawned Man1 up to the 18th century, which was subordinated to that of “the post-Gregorian Reform Church of Latin-Christian Europe” and entailed a “world of laymen and laywomen, including the institution of the political state, as well as those of commerce and of economic production” (p. 262). Man1’s “stuff” could not escape the horizon of what today would be called the supernatural, in this case in belief in an omniscient God, a character beneath which Man’s composition fell, and to whom his inborn soul belonged. Man’s character at birth was marked by the doctrine of Original Sin, setting the stage within Catholicism for the normalization processes of rearing to be focused on dealing with its “stain.”

For Wynter, Man2 emerges from the 18th century until today, through several trajectories including “the secularizing intellectual revolution of Renaissance humanism, followed by the decentralizing religious heresy of the Protestant Reformation and the rise of the modern state” (p. 262). If Church domination in particular was to be challenged or undercut, the Renaissance needed to produce a new figure of Man, one not supernaturally inspired but rather attributed to nature.

What bridges if not inspires this shift for Wynter is not just the Renaissance, Protestant Reformation and Scientific Revolution, however, but also colonialism and the invention of the idea of race. Quoting Quijano, Wynter notes that the idea of race would come to be “the most efficient instrument of social domination invented in the last 500 years” spawning new ontologies, ontological hierarchies, and subjugations, articulated as “natural,” and where: “One of the major empirical effects of which would be ‘the rise of Europe’ and its construction of the ‘world civilization’ on the one hand, and, on the other, African enslavement, Latin American conquest, and Asian subjugation” (Wynter, 2003, p. 263).

Here, what makes Man2 particularly distinctive is the conceptualizability “of nature as an autonomously functioning force in its own right governed by its own laws” appended to rationales for racialization and colonialism (Wynter, 2003, p. 264). Wynter argues that this, in turn, made possible academic disciplines, “the cognitively emancipatory rise and gradual development of the physical sciences (in the wake of the invention of Man1), and then of the biological sciences (in the wake of the 19th century invention of Man2)” (Wynter, 2003, p. 264). The disciplines gave a different “matter,” a different rationale and kind of “stuff,” a biological stuff called “nature,” as the reference point for the exclusivity of “the human” and this belief in exclusivity, reinforced by the processes of colonization and racialization with which it was concomitant, has put the planet at risk:

... the fact remains that while the natural sciences can explain and predict, to a large extent, the behaviors of these nonhuman worlds, the disciplines of the social sciences and humanities still remain unable to explain and predict the parameters of the ensemble of collective behaviors that are instituting of our contemporary world ... These behaviors, whether oriented by the residual metaphysics of fertility/reproduction of the agrarian age in the poorer parts of the world, or by the metaphysics of productivity and profitability of our techno-industrial one in the rich enclaves—with the one impelling the dynamics of overpopulation, and the other that of overconsumption—now collectively threaten the planetary environment of our human-species habitat.

Wynter (2003, p. 270).

The big picture formation of a “globe” pockmarked by nation-states, the division into so-called First World and Third World economies, and the threat to environment met the instantiation of education as compulsory in a way that supported exclusionary forms of nationalism. One fundamental pivot in this big picture narrative that Wynter paints, then, is the emergence of new conceptions of nature that not only bridged between Man1 and Man2 but inspired a new set of institutions and notion of the state. In

²Technology here is being used within an everyday notion of the digital, algorithmic, and electronic. For an historical account of the emergence of cybernetics in particular within educational theory in the US see Ivens (2018).

educational discourse, this was supplemented by an empiricist epistemology that established “ableist” ontologies and sensationist pedagogies dedicated to the unfolding of reason-as-morality in “sound” beings. The work of John Locke (1632–1704) is one site that illuminates this pivot and bridge around nature, empiricism, and Man2’s emergence and introduces a kind of ableist subjectivity dedicated to “morality” and “civility” as key distinguishers.

Locke on nature

A sound mind in a sound body, is a short but full description of a happy state in this world. He that has these two, has little more to wish for; and he that wants either of them, will be little the better for anything else.

John Locke, *Some Thoughts Concerning Education*.

The idea that there is only one universal inscription of who or what counts as a child was debunked decades ago and the fact that it had to be is a sign of a figure of Man’s domination. Attention to how “the child” is used, described and deployed, emptied and filled with different possibilities and limits in culturally mediated settings, is one way to understand the impact of a modernity-colonialism vortex to which Wynter refers and the reliance on discourses of nature. The works of John Locke, often positioned as one of the “fathers” of psychology and empiricism, are a good index of how normalization was encouraged through embedding a specific set of possibilities and limits in “the child,” exclusively rendered.

The normalized child of Lockean texts was given parameters, qualities, and roles across Locke’s political philosophies, such as *Two Treatises of Governments* (1689), his epistemological mainstay, *An Essay on the Human Understanding* (finished in 1689 and published across four editions), as well as more overtly pedagogical texts such as *Some Thoughts Concerning Education* (1692). The child is normalized across such texts as being a boy born with a mind, body, and soul, the former of which is posited as a blank slate and as having a series of capacities thought to unfold into faculties suited to a new social contract. This is a contract in which, as a member of the elite classes, the boy is intended to participate as an English gentleman.

Several discursive trajectories enabled the intertextual Lockean child—English, rich, male, and pale—of the late-1600s to appear peculiar relative to its premodern predecessors. These trajectories included in part the uptake of mechanical reasonings in political philosophies, a Renaissance culture of “human” dissection that made interiority into a different fabric, and the emergence of an English Protestant conception of the nuclear family (Baker, 2001). These ruptures gave new meaning to terms such as *power*, *nature*, *reason*, and an *interior/exterior problematic* in thoughts about humanity in a monotheistic-inspired politico-religious universe.³ This version of the normalized Lockean child emerged, then, out of some of the prior strictures associated with Man1 in Wynter’s terms and mark an opening toward Man2, more fully effected centuries later.

For example, in his social circle at Oxford, Locke was involved in discussions over transubstantiation. Before being able to answer religio-moral questions, Locke pondered instead what Man was in a condition to know and in particular what the capacity of human mind for understanding and knowledge might be altogether. This shifted analysis toward formulating answers around the capacity of a human mind to know (where mind was considered a nebulous thing somewhere within the body), and away from finding an answer suited to absolute Truth believed to be held by a spirit-Creator, changing the relationship between God and Man that Wynter discusses. It also changes *the route* through which scholars proceed to find answers to difficult questions.

In Locke, that route was peppered with references to nature. There are at least four inscriptions of nature across Locke’s texts that indicate his reconceptualization of human life as exclusive and hierarchized, and that help to establish a sensationist pedagogy amid the importance of reason-as-morality. Nature was: (1) the source of ideas, that is, it is the mechanical laws of motion we might guess lie behind the getting of an idea, e.g., the *Essay*, (2) a moral and positive concept, the Laws of Nature, equivalent to God’s Laws, e.g., *Elements of Natural Law*, (3) a sometimes negative, sometimes reasonable, and therefore ambiguous state, that is, the state of nature was a historical category referring to an imagined time and condition under which Man had once lived, e.g., *Two Treatises*, and (4) a form of both predetermination and openness that exists in a child for unfolding its faculties over time, e.g., *Some Thoughts Concerning Education*. While these different renditions of nature appear in various works, they cannot be disassociated in the end from the instabilities and flexibility that was written into the normalized child, its powers, and the pedagogy deemed appropriate for it in *Some Thoughts*. It is beyond the remit of this chapter to elaborate each of these examples of the complex deployment of nature in his texts. The very range, however, underscores both the difficulty and the significance of rewriting a figure of Man from supernature to nature.

For instance, Locke’s most famous epistemological principle, that all our ideas come from experience, either from sense perception or from reflection on our own mental activity, was not underpinned by a view of ideas as fantastical or phantasmatic but of ideas as *natural*. What is important about the inscription of ideas as natural is that this assertion was arrived at through the *religio-moral* debate referenced above, a theological altercation that incited Locke’s explicit search for the limits of human knowledge, the substance of which is now referred to as “psychology.” The degree of certainty with regard to religio-moral knowledge that a human could have, what could constitute “knowledge,” was an express concern in the *Essay*: “Our business is not to know all things, but those which concern our conduct” (quoted in Baker, 2001, p. 146).

³For a more extended discussion of these trajectories and how they inflected different concepts such as power see chapter two in Baker (2001).

In delineating the mechanics of this knowing, Locke paints a picture of mind as having, like body, its own mechanical laws of operation which is what natural here means. When working “properly,” the normalized adult mind has powers internal to it and lets external objects work on it in Newtonian fashion:

In time the mind comes to reflect on its own operations about the ideas got by sensation, and thereby stores itself with a new set of ideas, which I call “ideas of reflection.” These are the impressions that are made on our senses by outward objects that are extrinsic to the mind; and its own operations, proceeding from powers intrinsic and proper to itself, which, when reflected on by itself, become also objects of its contemplation ... Thus, *the first capacity of human intellect is that the mind is fitted to receive the impressions made on it*; either through the senses by outward objects, or by its own operations when it reflects on them.

Locke and Emmens (1866, p. 62; emphasis added).

The inscriptions of nature in *Two Treatises on Civil Government* further modify or make indeterminate any overall meaning of the term nature in Locke’s work, however. Locke’s reference to the “Laws of Nature” operates sometimes as another positive use of the term and sometimes in contrast, yet not diametric opposition, to the law of reason. Laws of Nature can refer to a moral inscription of an orderly society where the Laws of Nature would deem that to harm someone else’s life, liberty, health, or property was not appropriate. Locke thereby elevates individual possession and protection of these things to the special status of the Laws of Nature.

This multiplicity of meanings of nature in Locke’s work constitutes one of the grounds on which the child comes into view as modern and romantic: the child lies in positive relation to nature as mechanical and regular forces, it is vulnerable to immoral forces that lie outside of its body at birth but that lie within the realm of human conduct, and it has a discretely identifiable interiority that works in ordered ways to construct the knowledge necessary for reasonable decisions. The intersection of Lockean renditions of nature also therefore produces the subjectivity available to the child of his texts. Departing from a more troubled, “precivil,” state of nature’s past into a more “advanced” future of self-governing, reasoning adulthood becomes the possibilities inherent to being the child and the parabola of development itself (Baker, 2001, pp. 156–57).

These assumptions about nature, reason, and the parabola of development of the boy with a sound mind in a sound body thus undergird and underscore the formation of an ableist subjectivity in educational discourse that is indebted to specifically Protestant, mechanical and empiricist notions. These notions are not freestanding. They are interlaced with a savage/civil binary that reserves privileged space for “the gentleman” in early English colonialism, striated with bourgeois notions of color, class and gender that are not always overtly named as such, and they are dependent on the assumption that only through initial bodily sensations can knowledge acquisition proceed. Here, ableist subjectivity rewrites the body/mind complex of that gentleman via an onto-epistemology that asserts that all knowledge-production requires sensory portals in “good working order.” Alongside and imbricated within Wynter’s history of colonization/racialization, then, is a new threat of “disordering” this order to knowledge-production that trusts what the body senses rather than posits sensation per Descartes as an illusion. This breaks both the Catholic Church’s hold on Original Sin and supernatural causation and can also be used to justify domination and subjugation of those cast as less-than-human abroad whose apparent “non-recognition” of an albeit provincial body/mind dualism is framed as mythological, as a misunderstanding of the facticity of such a mechanical system. Between Man1 and Man2, the child is not so much inscribed as a natural Being, but as an effect of strategies for recognizing and naming “nature” in relation to Being altogether.

Man2 and ability: distinguishing one’s self from peers

F. J. Gall (1758–1828) is best-known for his work in an area referred to by others as phrenology. In his work with Johann Gaspar Spurzheim, Gall dissected cadavers in new ways and depicted with unprecedented detail the areas now thought of as comprising the nervous system (Richardson, 2001). While phrenology has rightly been dismissed as faddish pseudo-science and as a denigrating form of determinism that posited abilities and talents as inborn in the infant, some of its most basic premises have endured, such as localization to regions in the brain, the belief in the brain as the most important organ that separates humans from animals, and the belief that matter actually matters to explaining what Gall saw as the differences between men.

The combination of assumptions in Gallian psychophysiology stood, however, in direct contrast to Lockean sensationism. Gall did not believe that the senses were the major route to knowledge or the main difference-maker in educational terms. His works marks a shift from a notion of capacity to that of ability, foregrounding more directly a thesis of physiological innatism that was to redefine ableist subjectivity relative to previous debates.

One of his major productions, titled in English as *On the Origin of the Moral and Intellectual Faculties of Man* (henceforth Volume 1) was translated as part of *On the Functions of the Brain and of Each of Its Parts: With Observations on the Possibility of Determining the Instincts, Propensities, and Talents, or the Moral and Intellectual Dispositions of Men and Animals, by the Configuration of the Brain and Head*. Together the six volumes of this collection arguably represent a “crossover” moment in which concerns for morality especially become appended to the “intellectual” (in this case quite literally via the “and” placed between the terms moral and intellectual in the title). This appending, in turn, becomes attached to debates over *matter* and the hierarchy of properties and essences. While some philosophical tracts, like Locke’s *Essay*, posit extension (length, breadth, and so on) as more significant and fundamental than weight or color, Gall’s elevation of the brain and of head-based ontology implicitly fuses appeals to shape (such as length and breadth) and to color (such as references to skin or nation) within the analytical lens, coding such observations as data—something

not entirely common within the older philosophical debates. While much canonical Western thought has posited education as inherently concerned with moral development or its analogs, the conjuncture in Gall suggests an important but subtle shift that impacts ableist subjectivity's parameters: it is *through* the intellectual that the morality objectives or purposes can be secured and viewed (Baker, 2020, p. 401).

The six-volume publication and translation put Gall at the forefront of commentary on the nature of Man and in the firing line of critics. While often mocked, his work was also thought to require a response. In Volume 1, he responds to some of this mocking. He makes appeal to terms that today have much cache, such as facts, evidence, and dissection; he offers an extended analysis of how others have misconstrued the role of the senses; and he overtly uses the term data, leading to a very specific inscription of the term matter that changes causality away from supernature and embeds it into organizational structures today called physiological (Baker, 2020, p. 401).⁴

Gall on matter

F. J. Gall's work on physiognomy is a particularly relevant site for apprehending a fuller movement into Man2 as a peculiar kind of materialism. This made education reducible to instruction and elevated a variant of predeterminist philosophy. Even though critiques of phrenology regarding mental measurement movements and eugenics have been available to education for some time, there are other aspects of phrenology's basic presumptions that have been less visited.⁵ One of those is not *whether* phrenology participated in spirit-versus-matter debates as a question of origin regarding things coded as morality, talents, and dispositions, but *how*:

Are all men endowed, to the same degree, with the qualities essential to their nature, or, are the differences observed in this respect, due to the influence of accidental causes posterior to birth? Are these differences, on the contrary, determined in the womb of the mother? And if they are innate, how are we to cultivate, to perfect them, to repress or to direct them, according to the demand of individual or general good?

Gall (1822/1835, p. 96).

Gall generates this list of questions after an overview of treatises on what he calls Man. The questions, which could still summarize debate in many fields today, operate as a prompt and frame for the argument to come and the documentation drawn upon. His position in the main is that dispositions that might manifest at a later period are inborn and not the sole product of impresses made upon the senses after birth. Dispositions that are coded as inner, innate, and inborn (Gall uses all three terms) are thought to invariably make their mark on a surface, and that surface—whether the cranium, a dissected corpse, or a behavior—is available to the observer and really the only data “we” have to analyze or justify a conclusion.

Is it allowable that philosophers, while boasting to penetrate into the essence of the soul, should treat of man by piecemeal, and confine themselves to making long treatises on the soul, as an insulated being? Exercising its functions by itself, making use of the body, at most, as a means of communication between itself and the world; when, from the moment of conception to the last sigh, every thing indicates that in this world, the soul is in dependence on the material organs?

Gall (1822/1835, p. 88).

Here, conceiving of one's self as contained was required and mind was not necessarily allied to or equated with soul. Mind was considered to go completely through the agency of an organ defined as material. Gall's position, though, was more complex than this binary, albeit informed by it. In moving away from the Lockean conception of capacities Gall redefines the role of faculties:

[W]ithout doubt, these faculties are real: but they are only abstractions and generalities; they are not applicable to the detailed study of a species, or an individual. Every man, except an idiot, enjoys all these faculties. Yet all men have not the same intellectual or moral character. We need faculties, the different distribution of which shall determine the different species of animals, and their different proportions of which explain the difference in individuals.

Gall (1822/1835, pp. 88–89).

Through the exceptionality and liminality attributed to “an idiot” in the above, the “I” of humans is posited as a priori, in *need* of a material organ or organs for expression, and while creation was potentially designed by a Creator, this realization should not prevent one from understanding the operation of organs. The brain for Gall was not a single organ but a series of organs with

⁴See Baker (2020), from which the following sections largely draw, for a more extensive analysis of the implications of ocularcentrism and temporality in Gall's theory of matter.

⁵Different critiques of eugenics and education have especially emerged across the US (Franklin, 1994; Koza, 2021; Winfield, 2007), UK (Lowe, 2000; Serle, 1971), and Australia (McGregor, 2000; Slee, 1997).

different functions. Such organs are considered to house and display (by way of shape) the different talents and propensities with which we are born and which we take as our self.

The organ of animal life, the brain of man, is an assemblage of particular organs, many of which are found in animals. The animals of inferior classes have, by the fact of their inferiority to others on the score of intelligence, fewer cerebral organs; they have only the first rudiments of the human brain, and they are, consequently, easier to decipher than those animals which are provided with a more complex brain, and a more complicate animal life, or with more numerous instincts and talents.

Gall (1822/1835, pp. 91–92).

Of special note here is the constitutive character of corporal spatiality in Volume 1, as the “idiot” and “animal,” and less frequently “the savage,” cross paths in ways that link rather than divide phenomenology and positivism:

But, have those animals who possess all the senses which we do, such as the baboon, and ourang-outang, have they more decency, and are they more moved in seeing blood flow, than other animals? Are idiots, who possess their five senses in perfection, more virtuous than the man born blind? Must not every reader perceive, that it is the interior alone which modifies the impressions on the senses, and thus leads us, by a precipitate and limited judgment, to believe their operation immediate? It is for this reason, that external objects act very differently on men and on animals; very differently on the hare and on the fox; on such or such an individual.

Gall (1822/1835, p. 132).

The architecture of an emergent medical gaze is based on a redefinition of rationality driven by acute attention to and concern for how differences arise *within*, amid the limits of normalization:

The faculty of perceiving impressions, of retaining and comparing ideas, and making application of them, is by no means in proportion to the senses either in men or animals, as is proved by the example of idiots and simpletons. Thus, could we even have it demonstrated, that man, of all animals, has the most perfect senses, we should not thus obtain the explanation of his surpassing all others in intellectual faculties.

Gall (1822/1835, p. 128).

The implicit links between Man as “race” and “dis/ability” become key distinguishers here, re-rooting the material-as-organizational. The reification of the material, as the intermediary between religion and science, is not, like social justice or new materialist discourses today, however. Referring to “the material” is not to champion the underdog, the working class, or to exoticize objects. The material becomes rather the “home” of discourses of vision and of temporality in ways that markedly challenge Church authority, on the one hand, and that positions “idiots,” “animals” and “savage nations” as knowing nothing philosophical, on the other. Gall’s deployment of the material, rather than liberating toward “equality” or operating as anti-intellectualism, is a means of carving out a third space, one that nonetheless stultifies and creates hierarchies in a different way: “in certain individuals, some organs remain always in arrear, while others acquire the greatest energy. But this explanation shows again that all the moral qualities and intellectual faculties are innate” (Gall, 1822/1835, p. 145).

The sequence seems both well-intended (salvific) and totalizing (systemic), generating a spectrum outside of which no one can step. Differentiation is then set to become a site of fascination: “savage nations,” alongside and blurred with idiots–simpletons–animals, populate one end of the spectrum, while genius and men of energy are on the other, with the vast majority presumed to reside between the two ends. The “full human, not-quite-human, and nonhuman” hierarchy that marks Weheliye’s (2014, p. 3) recognition of racializing assemblages is given a newish rationalization in such a 19th century neuroturn. Knowing Man means breaking something down into material parts, while the appeal to intelligence means knowing to do what one is “fitted for,” prefiguring the formalization of both the mental measurement movement in education and the broader social program of what came to be called eugenics.

The term eugenics was coined in 1888 by Sir Francis Galton (Gould, 1981). Eugenics roughly refers to policies and practices targeting the reproduction of “quality populations” in a given geopolitical region and emerged in education in multiple subject matter pursuits and disciplines, from physical education to literacy to music (Barkan, 1992; Duster, 1990; Koza, 2021). While it became a transcontinental movement in the throes of nation-building and not just in the “West,” historians of eugenics today disagree as to whether it is a movement that has ended, “soft” or “hard,” a discrete discourse, or a diffuse set of discourses that endure through to present *neugenics* (Baker, 2013; Crowe, 2000; Davis, 1997; Carton, 2000; Winfield, 2007).

Regardless of the debate over eugenics’ impact, longevity, and coherence, however, an ableist subjectivity dedicated to judgment on the basis of matter-as-organizational and phenotypical took hold, lacing racialization and ableization discourses together as rationales within a biopolitical and statist medical gaze. For children positioned as racialized minorities, putting to death was always-already real in community contexts of multiple clashing cultures, especially in indigenous and Black communities forced via genocide and enslavement into the statist structures of a New World nation such as the United States, forming a palpable basis for transgenerational fear for the safety of the young. The “talent” criterion, such as African American children being taught how to read where white children such as in the South could not, could constitute an even greater threat to life (Anderson, 1988) and

become even more of an incitement to murder—what Watkins (1993, 2001) calls the conditions of tyranny and duress under which Black education formed. With eugenics, the racialized and binarized landscape of white/non-white expanded to children of all races who appeared “anomalous” or “asymmetrical” but not evenly so. Instances of where this played out alongside and after the peak of Gall’s popularity is the reemergence of the term *genius* and the advent of training schools for children and youth labeled as feeble-minded. The genius, who in European and US contexts could only ever be cast as white and male, and the building of training schools to house children and youth whose appearance was judged as evidence of ambiguity—not so “able” as to compete for “regular” jobs but not so “disabled” (a new term at the turn of the 1900s) as to not to do some kind of work, embodied all the tensions, incongruities, multilayered forms of segregation, and exclusivity of privilege that came with colonization’s expansion into endogenous forms (Kaplan, 2000; Wrigley, 1996).

Man3 and automation: the non-necessity of human ability-as-skill under life 3.0?

From Man1 to Man2’s overrepresentation of “the human” the figure of Man’s supposed reason changes, as do the reasons given for this reason. From Protestant-inspired and sensationist versions of capacity and morality (Lockean empiricism), to innatist and psychophysiological conceptions of morality and intellect (Gallian materialism), to statist and biopolitical conceptions of genius, intelligence and “the body” (genocide, eugenics and genetics), a shift from faculty psychology to behaviorism, from spirit to flesh, and what Wynter characterizes as religion to capital “R” Race was effected. Wynter argues further that what changes across the subsequent 20th century forms of sociogeny is the emergence of systems theories and cybernetics. Such theories rewrite the “stuff” of Man away from some idealized notion of nature—whether inscribed as spiritist or biological—and toward something else posited as a system. Wynter (2003) focuses largely on a systems shift, dependent on conceptions of cybernetics, automation and autopoiesis. This gives space for seeing what might be called a nascent man3—deliberate small “m.” Systems work is evidenced across domains including the work of Bateson (1979) on learning, Maturana and Varela (1980) on biology, and Luhmann (1995) on sociology, where it is neither entirely supernature nor nature grounding the figure of Man’s comportment, but rather autopoietic systems accounting for their own production, reproduction, and innovation, blurring “Man’s” exclusivity.

In practice in education, however, a voluntarist, biologic and state-oriented dualist mind/body foundation endured as an ontological given across the 20th century, and Man’s special form of and access to a nebulous quality called consciousness became the assumption guiding the perpetual (re)organization of formal education. In the 21st century, however, a mind/body dualism and Man’s claim to a unique hold on consciousness is placed more overtly under question, and is perhaps no longer the driving force of philosophies of existence, or at least is amenable to being seen as simply a point of departure for what Tegmark (2017) characterizes as Life 3.0. Life 3.0 emerges in regard to the coming-into-being of certain forms of AI and especially as AI intersects with neuroscience and big data.

AI has operated historically as a sliding signifier. The typical disciplinary starting points for a field labeled AI are taken most immediately back to the pre- and post-WWII work of Alan Turing in the UK and his “universal Turing machine” and the Macy’s conferences in 1950s US. After the Macy conferences, designing a machine that could do what “a reasonably logical human” could do became a general goal (Tegmark, 2017, p. 123). Who or what that reasonably logical human could be remained unspoken. Given the ongoing potency of a savage/civil binary for knowledge-production and claims to scientific status, however, it seemed to operate more in terms of a Man2. The reasonably logical human could already occupy the ontological zone of the full human based on his “nature,” amid the racializing, sexualizing and able-izing presumptions about his ability to reason. As Benjamin (2019) notes in *Race after Technology*, the issue in the early days of AI’s field formation were not about intention and identity alone, but rather a cluster of systemic elisions and provocations based on old models of humanhood.

Since the 1950s, however, research and technologies claiming the mantle of AI continuously challenged what was seen as a uniquely human capacity. In the decades after the Macy’s conference in the US, the difficulty of emulating whole human function was recognized, and the dream of a general AI was (until recently) less invested in, as the equaling or replacing of the total human complex proved too hard. The focus in the nascent computer sciences shifted and through the later decades of the 20th century specific skills and subsets of individual tasks, such as storage of information, calculation, and game-playing (e.g., chess and Go!) became the focus (Boden, 2018; Ford, 2018; Wilson, 2010). Chess especially was positioned as the *sin qua non* of intelligence in AI research. Chess suited the kinds of programming available insofar as the manipulation of symbols in chess and in mathematics matched the rule-bound and finite AI architecture at the time.

In the 1990s, once the world chess champion Gary Kasparov was defeated by the chess playing program, Deep Blue, new figures for programming emerged that tried to push beyond the limits of more rigid and fixed subject matters and architecture—the figure of the developing child and animals. Such figures emerge as key models for Research and Development, inciting the fields of brain development, neuroscience, and computer programming to start talking to each other across previous epistemological divides and overtly about affect (Wilson, 2010). Drawing on what are now called machine learning, deep learning and computer vision modeling, the exponential and disruptive technologies from the 1990s onwards were built on different kinds of computational and programming innovations, in which the human mind-body complex became not the sole referent or yardstick for what was seen as powerful.

Tegmark (2017) characterizes these shifts as three different versions of Life—Life 1.0 (the biological stage), 2.0 (the cultural stage), and 3.0 (the technological stage), locating current human-dominated circumstances within 2.0. Life 1.0 occurs “where both the hardware and the software are evolved rather than designed” (p. 17). The example Tegmark gives here is bacteria which he

sees as having DNA that specifies everything: “They never learn to swim toward sugar; instead, that algorithm was hard-coded into their DNA from the start” (p. 16). Life 2.0 he sees as current humanity “*life whose hardware is evolved but whose software is largely designed*. By your software I mean all the algorithms and knowledge that you use to process the information for your senses and decide what to do” (p. 17). This difference—the ability to design software in 2.0 “enables it to be much smarter than Life 1.0. High intelligence requires both lots of hardware (made of atoms) and lots of software (made of bits)” and to be more flexible, arguing that most of our software is added after birth through learning. Because none of us can transform and live for a million years or transform our cosmos “into a diverse biosphere that will flourish for billions or trillions of years, enabling our Universe to finally fulfill its potential,” a final upgrade is required, Life 3.0: “which can design not only its software but also its hardware. In other words, Life 3.0 is the master of its own destiny, finally fully free from its evolutionary shackles” (p. 29), which he argues many AI researchers believe will arrive during the coming century. The controversies that this launches Tegmark identifies as playing out along three major lines: the digital utopians who see AI technology basically as salvific, the techno-skeptics who see it as apocalyptic and reinforcing inequalities, and the beneficial-AI movement who see AI pragmatically, as “giving life the power either to flourish like never before or to self-destruct” and preferring the former (p. 34).

Research that embodies the possibilities of each of these trajectories typically occurs outside the field of education, yet that research has major implications for education’s continued existence as a sorter and constructor of “abilities.” A specific example is what is emerging through BMI research (brain-machine interface) in which brain-readability and brain-writability explore conduits between brain function and machines, uniting AI, neuroscience, and big data. A study recently reported in the *New York Times* (Velasquez-Manoff, 2020) captures the possibilities without necessarily considering the broader ethico-ontological stakes that techno-skeptics might raise. In the 2011 study, participants underwent a fMRI to have brain activity recorded while watching a film. The goal was to reverse engineer an algorithm so that it could construct, from brain activity and oxygen uptake, what a person was seeing. While the outcome did not capture 100% of the original film, the broad outlines were evident as “the reconstructed images moved with a dreamlike fluidity” (Velasquez-Manoff, 2020). The reconstructed images suggested how the details could be filled in with more precision in future years (see Nishimoto et al., 2011). Such studies are also suggestive of other possibilities. They intimate not only that the research already underway could “capture” what someone is seeing or thinking “in the mind’s eye” so to speak through non-invasive and increasingly more portable technologies, but also what could go in the other direction—what could be implanted in children, adults or potentially animals as a film, image, narrative, sentence, memory, disposition, emotion or affect from the exterior.

This prediction/eventuation of Life 3.0 points to questions over the emergence of man3 and his continued relevance, entification, and biology. These questions are ultimately entangled with the refiguration of ableist subjectivity in the wake of automation. From the “discovery of the unconscious” (Ellenberger, 1970) to AI, the fascination with automating both daily tasks that are unpleasant and larger jobs that a human could never perform has piqued. Today this includes not just earth-moving tasks handed over to heavy machinery or driverless trucks and cars, but analytical, quantitative, writing, drawing/design, and programming ones. As Peters (2017) has already forecast “workless societies” and “workerless societies” are now possible to imagine. Coupled with movements such as transhumanism whose express goal is the replacement of the aging human biological unit for a better carrier of “consciousness” different questions arise (Herrick, 2017): is *any* figure of Man necessary, and if so, in what form are his supposed capacities, abilities, or automative potentials to endure? Bostrom (2014) argues that such questions are not farfetched but emerging in the wake of something called superintelligence, where the aftermath could be far-reaching:

... the first superintelligence [that] may shape the future of the Earth-originating life, could easily have non-anthropomorphic final goals, and would likely have instrumental reasons to pursue open-ended resource acquisition. If we now reflect that human beings consist of useful resources (such as conveniently located atoms) and that we depend for our survival and flourishing on many more local resources, we can see that the outcome could easily be one in which humanity quickly becomes extinct.

Bostrom (2014, p. 141).

Superintelligence arises amid variations of AI that bear the traces of physics and neuroscience and that decouple the event of consciousness from the locus of an organ (Tegmark, 2017). Physics, for instance, contributes the idea that consciousness is an emergent phenomenon that arises through the re-arrangement of particles. Here, the concept of emergent refers to a state of being that shows “qualities above and beyond the properties” of the constituting particles themselves (p. 300). Neuroscience offers an integrated information theory (IIT), which describes the precise way of giving rise to such emergence of consciousness. When an event happens, “the information processing needs to be integrated,” i.e., “a conscious part” must communicate with the rest as “a physical process ... transforms the initial state of a system into a new [conscious] state” (pp. 301–302). The ramifications of these versions of consciousness in AI together with computer science’s basic understanding of “information as position” without substance/substratum (binary coding) enabled AI to make “the leap from brain to machines” (Tegmark, 2017, p. 299, in Baker and Wang, 2020). In other words, “it’s only the structure of the information processing that matters, not the structure of the matter doing the information processing” (Tegmark, 2017, p. 304 in Baker and Wang, 2020, p. 18).

In this trajectory, man3 is emergent, more or less an enabling confluence of multiplicities, operating as temporary place-markers or plateaus that form uneven pathways to the isolation of particular “qualities” that could most easily be decontextualized, unitized, coded, programmed, and recombined as a form of non-aging superintelligence. The irony here is that the ableist subjectivities associated with Man2 not only built the bridge to prizing such qualities, fostering their formation, and rewarding them in particular

forms through schooling and credentialing, but also potentially operated as the rationale for the demise of any figure of Man and “the human” that figure overrepresents for. As the processes of instrumentally rational and less resource-heavy knowledge-production accelerate to the point that not even programmers are required to program anymore, the rationales unleashed in Life 3.0 can potentially comport man3 outside of appeals to “nature” and in a format where Life-as-Information and Man as a costly composite system merge. The question remains as to whether such “systems” begin to exceed him and potentially, no longer need him in any “biologized” form whatsoever.

This underscores the absolute irony of colonization’s superiority effects, where the massification of the so-called “unfit” now expands to its ultimate point, no longer protecting “the colonizers’” zones of privilege. In such a moment, ableist subjectivities are not reducible to what the full human, not quite human, and nonhuman in Weheliye’s terms do to or signify for each other. Rather, the historical veneration of ableist subjectivity interlaced with the -isms born of Man1 and 2, their “reign” as a powerful sorter within an unquestioned distributive logic, has the potential to return full circle, redefining inferiority and inefficiency and annihilating the grounds on which “matter” and “empiricism” have previously staked their claim.

Conclusion: who-what?

Wynter’s recognition captured in the opening quote, of the tension between protecting a figure of Man who overrepresents and the potential annihilation of all as being the core challenge of this millennium, outlines the stakes and the trajectories of ableist subjectivities. From European-based Christocentrism (Man1) to the racialized nation-state (Man2) to the state of automation and non-physical substrate dependent versions of superintelligence (man3), the notion of human exclusivity has shifted, changing the horizon of enactment from the spirit to the flesh to the disembodied, from human centrism to human primacy to human non-necessity. The bases upon which ableist subjectivities have been founded have both interpenetrated these reformulations and adjusted accordingly, from faith in supernature to biological nature to systems thinking.

Ideas about intelligence have been one important bridge linking Man2 and man3 especially, albeit it in different forms. According to Malabou (2018), however, the next metamorphosis is yet-to-come. This prediction stands largely upon her reading of the pragmatism of Dewey and adaptationist psychology of Piaget. The shift is “the age of intelligence becoming automatic once and for all as a result of the removal of the rigid frontiers between nature and artifice. The power of automation by far exceeds a simple ‘roboticization,’ and the increasingly refined simulation of ‘natural’ intelligence makes a new approach to the brain incumbent—an approach that would not only make sense for biology but would also reveal the essential nature of its complicity with technological simulation” (pp. 15–16). On Malabou’s account, the collapse of nature and artifice will usher in a new era in which the biosocial and the socio-technium will not be able to be disaggregated. The running together of these terms is a temporary sign, because running together assumes prior separation and the bio and social or bio and techno will no longer be available to distantiation, making nonsense of claims to a “conflation.”

In educational discourse, though, any claim to newness is met with an at least bifurcated reaction. As the above has indicated, either, in techno-skeptic mode, man3 will just more cleverly mask inequalities and social injustices, hardwiring them into algorithms and other as-yet to be seen inventions owned and controlled by a select group, using them for automating Life and designating who-what matters, in the perpetual maintenance of an underclass. Or, in beneficial-AI mode, such technologies could undermine the historic foundations of inequality by eradicating the need for distributive logics altogether, taking the notion of an ontological hierarchy out of the game and opening onto a field of playfulness that honors the positive potentials of All. Inbetween sit already-existing children and youth, in already-existing schools and classrooms. What different professions, institutions and fields of education decide—and whether they get to decide—what should or could be forwarded of “the human” and left behind of the figure of Man is not an exercise in science fiction, decolonialism, or imagination. It is rather a crucial turning point where the moment for open dialog is contracting as the intensification and automation of abilities accelerates. As the 21st century unfolds, the responsibility for deciding in democracies at least what “we” are willing to share and what not needs revisiting and revision. Without the frank confronting of what ableism has supported and who-what - including more-than-human beings - it has not, the specter of the return which secures the demise of Man and “the human” he has overrepresented, remain on the table in the name of efficiency, superiority, profit, and reward.

References

- Anderson, J.D., 1988. *The Education of Blacks in the South, 1860–1935*. University of North Carolina Press, Chapel Hill.
- Baker, B., 2001. In *Perpetual Motion: Theories of Power, Educational History, and the Child*, vol. 14. Peter Lang, New York.
- Baker, B., 2013. *William James, Sciences of Mind, and Anti-imperial Discourse*. Cambridge University Press, New York.
- Baker, B., 2020. Inside out: head-based ontologies, vision, and temporality in the first neuroturn. *Educ. Theor.* 70 (4), 395–419.
- Baker, B., Wang, L., 2020. Common and divided school curriculum from mass compulsory schooling to technofuturists envisionings. In: Fitzgerald, T. (Ed.), *Handbook of Historical Studies in Education*. Springer Nature Singapore Pty. Ltd, Singapore.
- Barkan, E., 1992. *Retreat of Scientific Racism: Changing Concepts of Race in Britain and the United States Between the World Wars*. Cambridge University Press, Cambridge.
- Bateson, G., 1979. *Mind and Nature: A Necessary Unity*. E. P. Dutton, New York.
- Baynton, D., 2001. Disability and the justification of inequality in American history. In: Longmore, P.K., Umanski, L. (Eds.), *The New Disability History: American Perspectives*. New York University Press, New York, pp. 33–57.

- Benjamin, R., 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press, Cambridge.
- Boden, M., 2018. *Artificial Intelligence: A Very Short Introduction*. Oxford University Press, Oxford.
- Bostrom, N., 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford.
- Campbell, F., 2000. Eugenics in a different key? New technologies and the "conundrum" of "disability". In: Crotty, M., Germov, J., Rodwell, G. (Eds.), "A Race for a Place": Eugenics, Darwinism, and Social Thought and Practice in Australia. The University of Newcastle Press, Newcastle, pp. 307–318.
- Campbell, F., 2013. *Contours of Ableism: The Production of Disability and Abledness*. Palgrave MacMillan, Basingstoke.
- Cooper Ramo, J., 2016. *The Seventh Sense*. Hachette, New York.
- Crowe, C., 2000. Inheriting eugenics? Genetics, reproduction and eugenic legacies. In: Crotty, M., Germov, J., Rodwell, G. (Eds.), "A Race for a Place": Eugenics, Darwinism, and Social Thought and Practice in Australia. The University of Newcastle Press, Newcastle, pp. 173–180.
- Davis, L.J. (Ed.), 1997. *The Disability Studies Reader*. Routledge, New York.
- Duster, T., 1990. *Backdoor to Eugenics*. Routledge, London.
- Ellenberger, H., 1970. *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry*. Basic Books, New York.
- Ford, M., 2018. *Architects of AI: The Truth About AI From the People Building It*. Packt Publishing, Birmingham.
- Foucault, M., 1979. Governmentality. In: *Ideology and Consciousness*, vol. 6. Cambridge University Press, Cambridge, pp. 5–22.
- Franklin, B., 1994. From "Backwardness" to "At-Risk": Childhood Learning Difficulties and the Contradictions of School Reform. The State University of New York, Albany.
- Gall, F.J., 1822/1835. On the Functions of the Brain and of Each of Its Parts: With Observations on the Possibility of Determining the Instincts, Propensities, and Talents, or the Moral and Intellectual Dispositions of Men and Animals, by the Configuration of the Brain and Head. W. Lewis, Trans. Marsh, Capen, & Lyon, Boston.
- Garton, S., 2000. Writing eugenics: a history of classifying practices. In: Crotty, M., Germov, J., Rodwell, G. (Eds.), "A Race for a Place": Eugenics, Darwinism, and Social Thought and Practice in Australia. University of Newcastle Press, Newcastle, pp. 9–18.
- Gould, S.J., 1981. *The Mismeasure of Man*. W. W. Norton, New York.
- Herrick, J., 2017. *Visions of Technological Transcendence: Human Enhancement and the Rhetoric of the Future*. Parlow Press, Anderson, SC.
- Ivens, J.P., 2018. *Making the Modern Person: The Tyler Rationale, Curriculum Studies, and Cybernetic Systems*. Doctoral Dissertation, University of Wisconsin-Madison. ProQuest Dissertations and Theses Database.
- Kaplan, G., 2000. European respectability, eugenics, and globalisation. In: Crotty, M., Germov, J., Rodwell, G. (Eds.), "A Race for a Place": Eugenics, Darwinism, and Social Thought and Practice in Australia. University of Newcastle Press, Newcastle, pp. 19–28.
- Koza, J.E., 2021. *Destined to Fail: Carl Seashore's World of Eugenics, Psychology, Education, and Music*. University of Michigan Press, Ann Arbor.
- Locke, J., 1689/1965. *Two Treatises of Government*. New American Library, New York.
- Locke, J., 1692/1975. *An Essay on the Human Understanding*. Originally Published 1692–1700 Ed. Clarendon Press, Oxford.
- Locke, J., 1692/1989. Some thoughts concerning education. In: John, W., Yolton, J.S. (Eds.), *Introduction, Notes, and Critical Apparatus*. Clarendon Press, Oxford.
- Locke, J., Emmens, S.H., 1866. Selections from Locke's essay on the human understanding. In: Emmens, S.H. (Ed.), *Introduction and Notes*. Virtues Brothers and Co, London.
- Lowe, R., 2000. History of Education. In: *Major Themes*, vol. 1. Routledge, London.
- Luhmann, N., 1995. *Social Systems*. J. John Bednarz, With Dirk Baecker, Trans. Stanford University Press, Stanford, CA.
- Malabou, C., 2018. *Morphing Intelligence: From IQ Measurement to Artificial Brains*. Columbia University Press, New York.
- Maturana, H.R., Varela, F.J., 1980. *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing Company, Dordrecht.
- McGregor, R., 2000. "Breed out the colour": reproductive management for White Australia. In: Crotty, M., Germov, J., Rodwell, G. (Eds.), "A Race for a Place": Eugenics, Darwinism, and Social Thought and Practice in Australia. University of Newcastle Press, Newcastle, pp. 61–70.
- Nishimoto, S., Vu, A.T., Naselaris, T., Yuval, B., Yu, B., Gallant, J.L., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* 21 (19), 1641–1646.
- Peters, M.A., 2017. Deep learning, education, and the final stage of automation. *Educ. Philos. Theor.* <https://doi.org/10.1080/00131857.2017.1348928>.
- Richardson, A., 2001. *British Romanticism and the Science of the Mind*. Cambridge University Press, New York; Cambridge.
- Searle, G.R., 1971. *The Quest for National Efficiency*. Blackwell, Oxford.
- Slee, R., 1997. Inclusion or assimilation? Sociological explorations of the foundations of theories of special education. *Educ. Found.* 11 (1), 55–73.
- Tegmark, M., 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. Alfred A. Knopf, New York.
- Velasquez-Manoff, M., 2020. The brain implants that could change humanity. *The New York Times*. <https://www.nytimes.com/2020/08/28/opinion/sunday/brain-machine-artificial-intelligence.html>.
- Watkins, W., 1993. Black curriculum orientations: a preliminary inquiry. *Harv. Educ. Rev.* 63 (3), 321–338.
- Watkins, W.H., 2001. *The White Architects of Black Education: Ideology and Power in America, 1865–1954*. Teachers College Press, New York.
- Weheliye, A., 2014. *Habeas Viscus: Racializing Assemblages, Biopolitics, and Black Feminist Theories of the Human*. Duke University Press, Durham.
- Wilson, E., 2010. *Affect and Artificial Intelligence*. University of Washington, Seattle.
- Winfield, A., 2007. *Eugenics and Education in America: Institutionalized Racism and the Implications of History, Ideology, and Memory*. Peter Lang, New York.
- Wrigley, O., 1996. *The Politics of Deafness*. Gallaudet University Press, Washington, DC.
- Wynter, S., 2003. Unsettling the colonality of being/power/truth/freedom: towards the human, after man, its overrepresentation—an argument. *New Centen. Rev.* 3 (3), 257–373.

Following the money? Commodification, segregation and shopping

Andrew Loxley^a and Gary Thomas^b, ^a School of Education, Trinity College Dublin, Dublin, Ireland; and ^b The School of Education, University of Birmingham, Birmingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction: rise of the markets?	219
The role of policy: Caveat emptor?	219
Principled objections: why markets?	221
Back to the market	222
And what about inclusion and special education?	222
By way of a conclusion: market governance and governance by markets	224
References	226

Introduction: rise of the markets?

Our intention in this chapter is to consider the issue of inclusion from the perspective of marketization. How this affects inclusion is a moot point and has been a source of contention for the past three decades; especially around its implications for segregation and not only in locational but also in curricula and pedagogic terms (see for example [Bines, 1995](#); [Rouse and Florian, 1997](#); [Barton and Slee, 2010](#); [Norwich, 2008](#); [Norwich and Black, 2015](#); [Tomlinson, 2017](#); [Lui et al., 2020](#); [Thomas and Loxley, 2022](#)). Added to this are issues around diagnosis and assessment of “need” and the concomitant politics and logistics around the allocation, provision and distribution of resources; financial, material, and personnel. An important observation which underpins most critiques of the market is that far from being a neutral technical-rational “space” to facilitate a more efficient use of resources, it also brings with it issues around the moral and social valuation placed on difference—both collectively and individually. In echoing [Rix et al. \(2013a\)](#) comment on the special educational needs (SEN) “continuum,” does the existence of educational markets, which are intended in the parlance of economists to satisfy a range of preferences, reinforce “assumptions ... that some people or provision require segregation” irrespective of how or who provides this service? The creation of markets in educational “goods and services” and in particular sectors which were (or still are in the public domain) can be interpreted in two main ways. Firstly, there is a narrative which mostly regards marketization as a pragmatic solution to issues ranging from supply (and with it demand), the extension of personal and familial autonomy and choice into key areas of social life which were hitherto income and/or wealth contingent. It is important to note that this “choice” discourse is usually accompanied by the “raising standards” discourse. The competition for students between institutions should galvanize schools into being attentive to the quality of their services (see for example [Biggar and Fels, 2017](#)). The second narrative takes a more expansive view and is an attempt to provide a theoretical explanation for marketization as a phenomenon which has seemingly penetrated many facets of social life and creates a very different cultural dynamic and relationship between educational institutions and the communities and the individuals with whom they are connected.

Although space precludes a detailed unpacking of these narratives, we will consider them mainly in relation to some of the structural and cultural changes to education systems via policy interventions. Although our focus in this chapter is predominantly on the UK as a critical case of marketization and how it has evolved over the past 40 years vis-à-vis inclusion, it is a “tale” of reversals as the commitment to inclusion has been abandoned as a national policy goal by successive Conservative governments since 2010 and replaced with a policy which valorizes a highly individualistic and reductive conceptualization of SEN and a position which castigates inclusion as “an ideology conceived on spurious egalitarian grounds” ([Conservative Party, 2007](#); see also UK [Cabinet Office, 2010](#)). We will start with a discussion regarding the role of policy as an instrument to shape and form educational systems and practices. Next, we will provide a brief overview of the recent socio-political origins of marketization and its role in restructuring welfare provision and education. This will be followed by a discussion of how this process has affected special education and inclusion.

The role of policy: Caveat emptor?

Our twin emphasis on markets *and* policy is to reinforce the notion that policy is and has been a critical instrument through which to reinforce or restructure many areas of social life. Given the centrality of education across most states it has historically tended to be subject to near-constant policy interventions. We would argue that it is important to recognize that special education forms a small but relatively substantial component of the state welfare system.¹ This, we believe, is an observation that tends to become

¹Put in numerical terms, in the UK in 2022, SEND pupils comprised 14% (1.1 million) of all pupils and in 2019 [latest data] accounted for £9.4 billion in expenditure or 10% of all education spending of £94 billion) (see [NAO, 2019a,b](#) for detailed unpacking of SEND statistics). In 2022, 27% (or 355,566) of SEND pupils held were subject to an individual “Education, Health and Care” (ECH) plan and 12.6% of all school-age pupils were deemed to require some form of support. Of those with an ECH, 51% attend a state maintained special school—an upward trend since 2006 ([DfE, 2022a](#)).

obscured in many discussions about special education.² In this sense, any change in the wider welfare regime can and does have an influence on special education. An awareness of welfare reforms and restructuring is a prerequisite to a critical understanding why special education takes on its particular “shape and form.” In this regard we should not view special education as being insulated from the welfare state but is equally subject to the vicissitudes of macro as well as meso changes. This does not detract from the day-to-day pedagogical, learning or support work that is “practice,” but implies that much of this is framed and constituted in and through policy. Some of this work is more immediate such as documents like the Special Educational Needs and Disability (SEND) Code of Practice (UK Gov, 2015) which sets out in fine-grained detail how a school should frame and undertake its work with students who present or have a Special Educational Need (SEN) or disability. Some of this may be more distal in the case of the complex regulations for the purchasing of social care provision as part of an individual’s education and health care plan (UK Gov, 2014); in itself a product of quasi and actual marketization of health and social care services since the 1990s. But nonetheless these policy assemblages all become entangled within what Foucault (2002) would refer to as a “grid of intelligibility” when the connections are drawn together.

As we have argued elsewhere (Thomas and Loxley, 2007, 2022) the social categorization of children and young adults as somehow being “special” is constructed in and legitimized through the kind of policies that “speak” about them. They are, as Foucault (2002) would characterize it, embroiled in a web of discursive practices and regimes. This is as much a historical as it is a contemporaneous proposition. However, in broader terms, these systemic and cultural dimensions are located in what Esping-Andersen (1999) would characterize as welfare regimes. That is wider configurations and networks of welfare agencies and organization which are underpinned (or at least nominally) by (usually) abstract ideological notions which shape and pattern what Cox (1998) refers to as “principles” and “measures.” The former is concerned with the values which structure the regime and the latter the means through which principles are instantiated (i.e., put into practice). The former refers to the values or abstract ideals (“child voice,” “the market,” “parental preference,” “school autonomy” and so on) which constitute policy goals. The latter, “measures,” are concerned with *how* these values are “operationalized” or in Habermas’s (1987) terms brought into the lifeworld or more prosaically, the day-to-day practice of education. In echoing Ball (1994, 2012) policy measures come in multiple and for the most part interacting, conflicting and contested forms (whether oral, lexical-graphical, numerical and artifactual) and “exist” within a variety of interacting layers and contexts (e.g., legislative, financial, institutional, logistical, organizational, professional as well as “folk” norms and practices).

Clearly, any shift or redefinition of principles will bring about a corresponding change in measures—that is to say, the “instruments” or in the words of Foucault (2008), “technological apparatus” applied to making policy “work.” However, in echoing Starke (2006), we need to be analytically sensitive to at least three interconnected factors which condition policy and our responses to it. The first is “path dependency” or being attuned to the way historical antecedents act as constraints (as well as opportunities) on policy. The second is the “institutional” which refers to the durability and reproduction of norms and values (e.g., universalism or particularism, primacy of professional knowledge and decision making over other groups such as parents or administrators etc.), which shape and give rise to the organizational forms in which policy is generated, translated and enacted. Lastly, there is “politics” whereby various interest groups (such as political parties, non-governmental organizations [teacher unions, parent groups] & agencies) vie for influence, control and definitions of problems and proffered solutions. All of three of these factors form a dense and reciprocal network, in which (and to paraphrase Marx, 1926) people make policy but not always under circumstances of their choosing.

On one level, policy can be viewed in simple terms as representing normative guidelines for action or in Colebatch’s (2006) terms the “creation of order”; that is, it sets out how things *should* be done. Although this conception is useful to a point, as policy is concerned with defining objectives and in some instances the means through which they are attained, it is nonetheless rather generalized and largely instrumental, inadequately capturing the many ways in which policy is constructed, interpreted and subverted. Measures are not just a set of written precepts which we can easily exercise agency over; at work is both coercion and consent (see Ball et al., 2012). In this regard policy is not neutral, though ostensibly it may appear to be so via the use of discursive strategies to obscure intentions and/or sectional interests. Allied to this is the notion of policy as a signifier for underlying and stratified social relations of power, which crucially can be a vehicle for re-ordering and/or reinforcing these relations. When placed in an international context, there is no isomorphic relationship between principles and instruments; an observation well made by Rix et al. (2013b) fifty-five country study of SEN systems and policy orientation to inclusion. Here they found a generalized commitment to inclusion (the “principle”) but very little commonality in relation to how inclusion was operationalized (the measures) across a range of dimensions. These included for example issues related to “space” (types of schools), “grouping” (practices vis-à-vis the homogeneity or heterogeneity of students in classes based on certain criteria), “staffing” (role of practitioners, skill levels, attitudes toward inclusion), “students” (rights, characteristics i.e., definitions of SEN, how needs should be met) and “systems” (legislation, funding, services etc.).

²Additionally, SEND is not solely an educational “issue” but for a number of children & young adults and their families, one which intersects and interacts with health and social care domains—each of which require navigating and are subject to their own respective organizational and cultural logics. But this is nothing new and has been a problematic feature of the SEN “system” for decades. Therefore, the reading of any SEND policy needs to be quite literally a hermeneutical exercise; a gathering together and interpretation of what Ball (1994) called an ensemble of related texts in order to make sense of any single policy.

Principled objections: why markets?

The position given to markets (imaginary or real) and its centrality to neoliberal (and a lesser extent neoconservative) cultural politics has been subject to significant and substantial deconstruction, critique and commentary since the late 1970s across a variety of disciplines (see [Thomas and Loxley, 2022](#) for a summary). This would have remained a mere academic curiosity had not political administrations, notably in The Conservative Party government (1979 until 1997) in the UK and The Republican Party government (1980–1994) in the US, taken on anti-welfarist and pro-market agendas. What followed was that over nearly two-decades a range of social and economic policies were heavily influenced by the so-called “New Right”—not so much a movement, but a disparate collection of think-tanks, politicians, academics and pseudo intellectuals (see for example [Hall and Jacques, 1983](#) or [Levitas, 1986](#) or [Young, 2013](#)). At the center of the New Right polemic—both neo-liberal and neo-conservative—was the argument that the state and particularly such institutions as education, health and social security as they have evolved during the latter-half of the 20th century, had according to [Newman and Clarke \(1994\)](#) created three interrelated concerns:

- *moral*: welfarism had generated a culture of dependency underpinned by anti-capitalist values and sentiments;
- *financial*: high public expenditure and borrowing required high taxation and interest rates which “de-incentivized” work; and
- *social*: the repression of individual freedom and autonomy.

As we have noted elsewhere, these sentiments formed the political and ideological bedrock of policy during the 1980s and up until the mid-1990s. They were also fundamental to the reconstruction of a new role and purpose for state welfare which was accelerated in the late 1990s and became firmly embedded in the 2000s and 2010s. The dismantling of the post-war universalistic conception of welfare and the emergence of particularism as the new organizing concept in Britain has had profound implications right across all the main state welfare agencies: health and social care, social security and education ([Levitas, 2005](#); [Powell, 2008](#); [Taylor-Gooby, 2011](#); [Exley, 2012](#)). As noted previously ([Thomas and Loxley, 2001](#)) this guiding force in state welfare, coupled with shifts toward ostensibly decentralized forms of control and governance, the extensive use of marketization and “contracting out” of many services previously operated by the state, along with the reframing of client–professional relationships around concepts of “empowerment,” have had and are still having a profound impact on provision. It is useful to note that the existence of the state as an institution was never in question (other than by the more anarcho-right-wing libertarians), as it has been central to not only the restructuring of welfare provision, but the creating and “nurturing” (of) a specific type of cultural dynamic which is suffused with the core neoliberal and neoconservative values. [Foucault \(2008\)](#) in his discussion of neoliberalism—via his concept of “biopolitics”—saw this in terms of the “entrepreneurial self.” As [McNay, 2009](#) notes this is where the boundaries between the economy (its values and norms) and in particular market activity colonizes and dissolve into other spheres of social life to become the new doxa:

The autonomous citizen is s/he who manages these diverse networks—work, household, pension, insurance, private property—in the most responsible and prudent fashion vis-a-vis the avoidance of risk and the maximization of their own happiness.

[McNay, 2009](#) p. 6.

But the “entrepreneurial self” is not just a conscientious and independent citizen, he/she is also an individual who re-makes themselves in order to create and pursue opportunities in line with their own self-interest and not those presupposed by for example the state. This is neatly captured in the notion of “human capital” (an active investment in the self), which if judiciously undertaken should yield future returns, financial or otherwise. Education, which for most of us is a fundamental positional good in terms of our material and non-material well-being is an obvious example vis-à-vis neoliberal ideology.³ The role of the state in all of this is to create and maintain the conditions through which the entrepreneurial self can flourish. But as [McNay, 2009](#) also notes (echoing Gramsci), this is a form of governance which needs constant intervention and attention. This hyper-individualism can create a diminution or negation of social bonds, whereby our interactions with others is one merely based on self-interested transactions with little or no regard to their well-being. This line of reasoning is echoed in the work of [Beck and Beck-Gernsheim \(2002\)](#), [Beck \(2007\)](#) and [Giddens \(1990\)](#) who discuss this in the context of what they characterize as “reflexivity modernity” and “individuation.” Put simply, they argue that one of the defining features of late-20th century capitalist societies in particular, is the loosening of “tradition.” In this sense tradition refers to the acceptance of and reproduction generationally, of social roles and positions based on for example class or gender norms. Without wishing to generalize or be too deterministic, middle class/working class children go to middle class/working class schools and then onto middle class/working class occupations (via some form of tertiary education). These progressions become unquestioned rites of passage. The break with tradition comes via this sense of individual reflexivity whereby new horizons of possibility and choices are considered and more importantly seen as attainable. The genealogy of this structural and cultural shift is for Beck and Giddens multifaceted but rooted in changes in economic and occupational structures, a form of consumerism which valorizes individual difference and the welfare state itself. For example, and in very simple terms, access to higher education for young working-class adults becomes feasible as structural as well as cultural “barriers” are lowered or removed and with it a change in life chances which were not possible or at least highly improbable for their parents. However,

³Though markets are more problematic for the neoconservatives with a more traditional concept of what constitutes “an educated citizen”—a tension well discussed in the early critiques of marketization in the UK (see [Bash and Coulby, 1989](#); [Jones, 1989](#); [Ball, 2009](#)).

Beck also argues that this do-it-yourself approach has a number of attendant risks as one becomes removed from one's tradition. The potential for the dissolution of social bonds (family, peers, location) in the pursuit of a different self-directed trajectory is ever-present and may lead toward alienation (see for example [Kearns, 2018](#)).

Back to the market

The application of markets in educational contexts is not a new phenomenon. Most countries' systems are structured and organized around mixed economies of provision (UNESCO). These tend to be built around clusters of private (for and not for profit) and public (state funded) institutions found across most sectors (pre-school, primary, post-primary tertiary). However, how this "mix" is culturally and structurally configured is highly context contingent. The range in terms of sector density (e.g., ratio of private to non-private institutions), mode of organization (selective versus non-selective), types of funding mechanisms (public to private transfers, fee charging), employment regulations, patterns of ownership, curricula autonomy and so on are "worked out" That is, each country will have devised or evolved its own market frameworks based on their own cultural, social, political milieu ([Starke, 2006](#); [Gingrich, 2011](#)). A technical, but nonetheless critical consideration when discussing the structural contours of educational markets are, how is it configured? The most obvious markers of difference between public and private provision pivots around the question; "who pays"? However, what constitutes "private" as applied by the [OECD \(2018\)](#) is not straightforward as it is a mix of factors related to sources as well as the proportion of funds (e.g., from households, non-governmental organizations [charities, religious groups], transfers from the state), ownership (related to land, buildings) as well as managerial control over the school or institution, regulations determining curriculum, employment conditions, student enrollment, student exclusion and so on. Lastly, there is the distinction between "government dependent" and "government independent" private institutions. The former derives 50% or more of its core funding from the state and the latter 50% or less. According to the [OECD \(2021\)](#) most funding for the compulsory sectors (primary and post-primary) is derived from public sources. For the pre-school and tertiary (higher education) sectors the source of funding is much more variable. In comparing what the OECD/UNESCO categorize as "household expenditure" on education (predominantly tuition fees) with "public sources" for primary and post-primary, the former is on average 8% with 90% coming from the state. The country with the biggest differential is Chile (29% and 71% respectively); in Sweden public sources account for 100% of school expenditure. For tertiary education on average the state contributes 66% and households 44%. An added complexity is that source of funding does not by itself indicate the extent of marketization.

In the case of England, an example of a highly marketized system, 72% ($n = 2888$) of post-primary schools are deemed to be "private" but are recipients of direct state funding. As such individual users (the principal) i.e., parents/careers pay no tuition fees to the schools (the agent). This is what [LeGrand \(1991\)](#) refers to as a quasi-market. Although the English system has undergone numerous changes over the past 30 years (mainly around diversification of school type, regulation & monitoring ownership and control) these have not substantially altered this operational model since it was introduced in 1988. In this sense this type of market possesses all of the features of a non-quasi market: many buyers (parents/carers) and sellers (schools), access to and the production of information regarding the service/product, and mechanisms to express preference and/or choice. What it does not have is a price mechanism to regulate relations between buyer and seller which is one of the defining features of a real market, nor are they meant to be profit seeking.

The cost of educating a child or young adult is determined by the state (local and central) via formulas whose basic unit is the individual pupil. The competitive logic of the market is brought about by the necessity to attract and retain students who literally have a "price on their head."

Added to this is a mode of devolved management which allows schools to view themselves as autonomous units and act in ways which would through their curriculum offerings and pedagogical practice to: (1) attract pupils via parental preference; and (2) create a basis for growth. Schools which did not attract enough pupils to cover their costs either had to reassess their position to reverse this trend or run the risk of failing altogether. Ironically, for all the political attention given to markets and a faith in their redemptive power, the evidence for their effectiveness is patchy at best.

And what about inclusion and special education?

If we were to locate a "policy fulcrum" in all of the structural reforms which has problematized SEND from 1988 onwards it would be: (1) resourcing; and (2) performativity in the guise of fostering competition between schools to "raise" standards. Both have been extensively debated in relation to inclusion and SEND in particular and represent key signifiers for the critical tensions and contradictions in a system organized around and predicated on marketized (quasi or otherwise) forms of behavior. From the outset this simplistic notion of the "money following the pupil" based mainly on the number of pupils in a school worked as a disincentive (or created a market failure) to admit children with an SEN as the money "attached" to them would not cover the expense of their additional supports. Even with the introduction of financial compensatory mechanisms used by local education authorities (LEAs) in the early days of marketization, this still proved highly problematic ([Loxley, 1999](#)) and remains so ([NAO, 2019a,b](#); [LGA, 2020](#)).⁴

⁴These have gone through numerous permutations since the early 1990s and now comes in the form of the "pupil premium" based on the number of pupils registered for free school meals—a "traditional" proxy indicator for socio-economic disadvantage—and the "dedicated school grant." The latter is intended to be ring-fenced funding allocated by the central state but used by local authorities to support schools and their "high needs" pupils.

If mainstream schools were left to endure (as intended in the 1980s) the oxymoronic creative destruction of the market to “up their game,” special schools seemed to fare very much better. Although devolved management was extended to special schools it had a very different flavor to that of their mainstream counterparts which put them outside of this new market regime. This was largely due to the recommendations of the accountants Touche Ross (1990) who were put in charge of the process of extension. Their train of activity in doing this was to ask those in the system, including special school heads, how they thought special education should be funded. The conclusion, perhaps unsurprisingly, reached by the accountants was that special schools were often “centers of excellence” and that they should not be allowed to wither on the vine as children moved outside the system to the mainstream. They were thus funded according to the number of *places* (whether they were filled or not) at special schools; a model which is still in place. In much of the writing on SEN and policy, special schools, which form a significant component of the SEND “ecosystem” tend to get overlooked. In blunt numerical terms there are 1298 schools (995 are state maintained, 247 independent and 58 non-maintained) that accommodate 142,028 pupils who account for 53% of all pupils and young adults with an ECH plan. Although the number of state-maintained schools has declined since 2002 ($n = 1161$) the number of pupils has increased from 2006 onwards. The average annual cost per pupil is £25,000 for maintained special schools and £50,000 in an independent school and account for £2.3 billion or 25% of the annual SEND expenditure in England and Wales (NAO, 2019a,b); independent special schools alone cost £957 million. In comparison it costs £6000 per year per pupil with an ECH to be educated in a mainstream setting.

In going beyond this headline data, in policy terms over the past 40 years the diminution or indeed dissolution of the special school system was never part of the SEN agenda (see Norwich, 2008 and Rix, 2011).⁵ If anything, the role and position has been reinforced in three main respects. Firstly, the creation of new types of schools. New Labor’s inclusion in the 1998 Standards and Frameworks Act for the establishment of a Foundation or Community of Grant Maintained special schools signaled very loudly that their model of inclusion saw a place for these institutions. The legacy of this found its way into the Academy School Act (2010)⁶ which allowed for the setting up of special academy and special free schools (DfE, 2022b) which unequivocally cohered with the Conservatives’ anti-inclusion agenda (see Conservative Party, 2007). Indeed, the latter bemoaned the closing of schools under New Labor. Secondly, there has been the reinforcement of the role of the special school as a local center of excellence, whose expertise can be drawn on to support their mainstream colleagues. This formed a part of New Labor’s position (DfEE, 1997; DfES, 2001; HoC, 2006) as well the Conservatives’ (DfE, 2010). In its more radical form special schools labeled “outstanding” by the Office for Standards in Education (Ofsted) can offer initial teacher education programs. Thirdly, there is the all-pervasive notion of parental/carer empowerment which again is articulated via the use of the discursive (or rhetorical) mechanism of “choice” through diversity of provision. A further legacy of not only the expansion of the system, is the funding mechanism itself. With the money staying with the special school, there is little finance freed up to fund children when they are mainstreamed. The widely accepted recommendation from Touche Roche was, however, at variance with the body of empirical evidence in education which has consistently indicated that in outcome terms special schools are no more effective than mainstream schools (see Thomas and Loxley, 2022). There was no educational or inclusive rationale for the assumption about “centers of excellence”; the rationale pushing the reform of management into the special school arena was identical to that driving it in the mainstream arena. It was a rationale constructed out of the imperatives of the market which was paradoxically intended to “protect” them from the volatilities of the market.

There is also a further effect of this which concerns the LEA. Accompanying the long ascent of the market, the past 10 years, in particular, has seen the demise of what used to be known as the local education authority (LEA). This is in contrast to the role of the LEA that was promoted in the late 1990s from a government committed to inclusion and was different from the residual-provider approach that developed between 1988 and 1997 (Loxley, 1999). Its role as seen from the perspective of central government has been mapped out in a number of policy initiatives, which had their objective recalling the LEA from “exile” (Audit Commission, 1992; DfEE, 1997; Ofsted, 1999). However, with the advent of yet more Conservative administration New Right ideology *inter alia* Academy Trusts, academy schools and to a much smaller extent free schools, the LEA now downgraded to just a local authority (LA) finds itself occupying the residual-provider role once more; a minor player within Ball’s hetroarchy. However, one of the key (and historical) residual activities is the administration, though not necessarily the provider of SEND services. This has long been an area of contention and critique since the early 1990s (Lamb, 2009; DCSF, 2010) coming from both state agencies and parent groups. The past 10 years have not been kind to local government:

The system for supporting pupils with SEND is not, on current trends, financially sustainable. Many local authorities are failing to live within their high-needs budgets and meet the demand for support. Pressures—such as incentives for mainstream schools to be less inclusive, increased demand for special school places, growing use of independent schools and reductions in per-pupil funding—are making the system less, rather than more, sustainable.

NAO (2019a,b, p. 11).

The issue of the relationship between “needs” and “resources” as shown above, has long been fundamental to debates around inclusion. Despite the long-standing prevalence of a “continuum” which is supposed to an expression of the magnitude of resources required to support a pupil, this tended to be driven by administrative categories based on fixed and largely reductive conceptions of “need” (Tomlinson, 2017).

⁵In borrowing from Kohn (2020) special schools are a type of non-universal public goods provided by the state which can be collectively consumed but are preferred only by a sub-set of the population.

⁶For critiques of this program see Ball (2005, 2009, 2015).

Like a form of alchemy, this is further compounded when attempts are made to cost these activities within a system which is constructed around a performative calculus. As argued by Penna and O'Brian (1998) and Taylor (1998), there is also an underlying shift from the politics of provision to the politics of identity which is partially bound up with critiques of the inflexibility of universalism and the role of the state (national and/or local) from both the left and right sides of the political spectrum (see Ranson, 2003; Ball, 2005, 2012; Gingrich, 2011; Diamond, 2013; Hicks, 2015; Ball, 2018; Ball and Bowe, 2020; Kohn, 2020). Cox (1998) suggests, this is further exacerbated by the displacement of welfare based on "social rights" to a regime constructed around "discursive rights." In the case of the former, social rights are concerned with equal access to education, healthcare, financial support if unemployed. Although in principle each citizen has an equal right to all of these, access and/or entitlement is mediated via the use of specific rules to determine eligibility and equality of treatment. This is referred to as "conditionality." In this sense, rule-governed patterns of allocation conform to the notion of distributive justice in that what people get is based on the fair and unbiased assessment of their needs. For the latter, discursive rights may be read in terms of empowerment and the positive acknowledgment of difference it can equally be conceived of as a competitive struggle over the definition of needs and associated resources played out within existing as well as reconstructed institutional arrangements (see Tomlinson, 2012).⁷ In this sense, the mobilization of interest groups endlessly engaged in strategic maneuvering becomes paramount. The corollary of this is that those individuals who are "disenfranchised" or belong to groups which simply lack the resources to acquire resources become further disadvantaged. To say whom gets what depends on who shouts the loudest. Distributive justice maybe, in other words, insufficient when calculations about putative need are based on representations made by groups or individuals that, in turn, have access to varying "social capital" and "cultural capital."

In coming back to the "principles" and "measures" which directly connect with the above discussion over rights and needs, is the bringing of parents/carers into the SEND market-place for the purchase of resources. This is articulated by current policy stemming from the most recent legislation governing special education, the Children and Families Act (2014) regarding "need" and resource allocation. Firstly, the legal requirement for local authorities to provide the "local offer" i.e., what services and resources they have and secondly, the introduction of Personal Budgets for pupils with an Education, Health and Care (EHC) plan. The former is based on recommendations from the Lamb Inquiry (2011) and Ofsted (2010) regarding poor practice regarding the quality of information available to parents/carers and the later the extension of longstanding initiative within the domains health and social care (for a comprehensive overview see Martinez and Pritchard, 2019; also Fleming et al., 2016, also UK Gov 2009 and DoH, 2012). However, in echoing Chow and Bracci (2020), what is intended to be a measure to provide a degree of autonomy for pupils and their families regarding decision making over resources, it can be read (and particularly so in the context of the principles of marketization), as an extension and deepening of the entrepreneurial self through the exercise of rational and informed consumer sovereignty and no real diminution of state surveillance and control via a matrix of regularity mechanisms to ensure responsible appropriate and accountable behavior (see UK Gov, 2014).⁸ It can also be seen to fit with Foucault's notion of the dual productive/oppressive nature of power. Chow and Bracci (2020) also ironically notes that one of the major goals of disability groups since the 1970s has been to reduce their dependency on the state through instruments such as personal budgets. But there is also within this a well noted and enduring distributional effect, in that certain kinds of need and their presumed etiology work within existing socio-cultural frameworks mediated disproportionately via for instance gender, ethnicities and social class (see Tomlinson, 2015, 2017; DfE, 2020).⁹

By way of a conclusion: market governance and governance by markets

In this short review of marketization and its effects on inclusion we would echo Jessop's (2015) observations about how the further entrenchment and extension of marketization have fundamentally altered the role of the state in how it positions itself vis-à-vis welfare provision and more significant relationships between the state and its citizens. One could be pragmatic (or optimistic) and view the market as a vehicle for resource allocation which takes individual preferences seriously. However, as we have tried to show, sadly the market as envisaged by the New Right has become a device and mechanism for reinforcing as well as justifying a particular world view and specifically one where individual difference is largely seen as stratified rather than complex and many-faceted. The "theoretical" noise of social-Darwinism is never far below the surface.

In looking at this, one has to be aware of the role of policy as used by the central state in creating and sustaining conditions which foster or discourage inclusion as a policy goal. Hence it is essential that we explore how marketization has operated in instantiating policy objectives. As an example, one can look at what the UK government's own National Audit Office (2019, 2021) has said about the use of public funds in systems that purport to support inclusive solutions to children's difficulties at school in the context of

⁷Just to complicate matters "need" can be conceptualize via the following forms: (1) Felt need—which is subjective (2) Expressed need—which is mobilized into some form of "campaign," (3) Normative need—as defined by welfare professionals and (4) Comparative need—judged according to relative needs of groups.

⁸At the time of writing, 25,300 (or 5%) of EHC plans also included a personal budget component (DfE, 2022b). Over 500,000 of them are active in the area of social care. For health care this is 124,000 (NHSt, 2022b) with an intention to increase this to 200,000 by 2024.

⁹More males (15%) than females (8%) have some form of SEN support and 4% and 1% respectively have an EHC plan. 28% of pupils eligible for free school meals (a proxy indicator for familial economic disadvantage) are labeled SEN as opposed to 13% of pupils without an SEN (DfE, 2020). 20% of pupils of Black-Caribbean origin have either SEN support or an EHC, in comparison to 8% of Asian-Indian pupils who numerically form a much larger group (268,318 as opposed to 86,543). And in echoing Tomlinson (1982, 2012, 2017) some "needs" such as those with an assumed bio-medical origin are more "desirable" or at least less stigmatizing than others.

budgeting within the quasi market. The [National Audit Office \(2019a: 8\)](#) noted, “The main reason why local authorities have overspent their high-needs budgets is that more pupils are attending special schools.” It proceeded to note that “Between January 2014 and January 2018, the number of pupils in special schools and alternative provision rose by 20.2% ... Spending on independent special schools increased sharply—by 32.4% in real terms between 2013–14 and 2017–18.” Here, one can see how the notion of inclusion as “an ideology conceived on spurious egalitarian grounds” ([Conservative Party, 2007](#)), as we noted earlier, has materialized. The abstract notion of “spurious egalitarian grounds” has taken corporeal shape via the increased use of not only maintained special school places but—even more so—the use of private special school places.

Policy is a deeply political process, and it is imperative that we interrogate the principles on which policy is based. In the UK, the market was adopted by both right and center-left wing governments, which, though they shared some value positions, did have different conceptions of what a market could or should do. While it is reassuring to see that inclusion—as for example advocated by the Convention on the Right of Persons with Disabilities and the Convention of the Rights of the Child—is enshrined in legislation in the vast majority of the countries (of whom 73% profess a “policy guarantee” regarding “integration” into mainstream settings (see [Waisath et al., 2022](#)), the abandonment of this commitment in the UK in particular, demonstrates the fragility of values and how quickly they can be expunged based on superficial evidence and ideological posturing. In the context created by marketization, social rights, however problematically administered, assume the status of being an “optional extra.”

The now near hegemonic discourse of the New Right with its emphasis on individualism, the free market, tradition and the family makes alternative versions of the “good society” difficult to pursue as schools and parents/carers become, whether they wish to or not, players in a game in which the rules constantly change. The elaborate regulations devised by the state to mitigate or police the negative effects of markets such as “cream skimming” by schools or “monopsony” by parents/carers or simple “market failure” is testimony to the flawed logic of educational markets, or at least those driven by a particular form of competition. Much of the policy it has enacted has been inimical to inclusion. Expressions about the need for a plural, tolerant, inclusive education system sit uneasily with policy that foregrounds the benefits of choice, selection and the comparison of schools on the basis of their pupils’ attainment.

Where can we go from here? How should we face up to these challenges? Elsewhere ([Thomas and Loxley, 2022](#)), we have suggested that inclusion seems to wake up each decade to a “Groundhog Day” to relive and re-confront barriers facing it not only from politics but also from its own history of managing “special needs” and from its heritage of research methods. We have focused in this piece on the political, policy and structural influences—notably via marketization—shaping practice and pushing it, seemingly inexorably, away from the direction of inclusion. It is the case, though, that many of the institutional practices operating in education systems are appropriated and employed by such structural influence in maintaining segregation and exclusion. The research we use, the thinking we employ about failure, the words and ideas that we use to discuss the *point* of education, and the ground-level activities that we can engage in—all of these can and must be considered in challenging a political orthodoxy that is inimical to inclusion.

As far as research is concerned, we can work to develop models of research in looking at inclusion that respect the idiosyncrasies of particular situations and cultures, simultaneously eschewing the sterile “helicopter research” ([Thomas and Loxley, 2022](#)) that imposes “What Works” solutions on teachers and schools.

As to the thinking we employ about failure, we can surely encourage intersectional understandings of difficulty at school. Such understandings recognize that failure at school is usually less to do with “disorder” or “disability” and more to do with context, culture and community in the extraordinary circumstances that school environments present to many children and young people.

About the *point* of education, academics and practitioners can encourage a discourse wherein, as [Biesta \(2007: 1\)](#) argues, we “... keep in view education as a thoroughly moral and political practice that requires continuous democratic contestation and deliberation.” We can reflect, as [Biesta \(2020\)](#) recommends, on the *point* of education. We can move away from words in our own discourse such as “effective” and “efficient” with regard to education, and substitute for them words such as “worthwhile” and “meaningful.” Biesta suggests that “effectiveness” is never an educational good in itself, only becoming meaningful in relation to views about the purpose(s) of education. And “efficient” education, so beloved of legislators looking segregative solutions to accompany and support marketization, should perhaps be expunged from our vocabulary entirely.

Last, teachers, researchers and academics can selectively challenge local and national policy that acts to exaggerate difference. Rarely is concerted action taken against local or national legislatures to challenge unjust or inappropriate measures, instituted by politicians on populist platforms such as “improving standards.” Teachers may feel powerless in these arenas but [DeMoss \(2003: 61\)](#), [Sailor \(2015\)](#) and [Kopanke \(2020\)](#) review evidence which—in the USA, at least—shows that such action can help to decrease disparities, whether or not plaintiffs directly succeed in their claims. DeMoss notes: “States seem to respond to the pressures that litigation places on them to take a closer look at the ways they fund education.”

Many of the problems we have discussed are, of course, structural, created by a political system which has little real interest in change. But progress depends on more than changes in politics and policy: the field needs to go well beyond inclusive education’s historical preoccupations and grasp the insights that come from intersectional understandings of failure and unhappiness at school. It needs to stop seeking new syndromes “explaining” failure and to argue powerfully against the use of such syndromes to rationalize the difficulties students experience with life at school. It needs to go beyond failed models of inquiry that cement in place age-old diagnosis-and-treat pedagogy. It needs, in short, to trust in the intelligence of the teacher and to empower the teacher in challenging an ideology that is hostile to inclusion.

References

- Audit Commission/HMI, 1992. Getting in on the Act: Provision for Pupils with Special Educational Needs, the National Picture. HMSO.
- Ball, S.J., 1994. Education Reform: A Critical and Post-structuralist Approach. Open University Press, Buckingham.
- Ball, S.J., 2005. Radical policies, progressive modernisation and deepening democracy: the academies programme in action. *Forum* 47 (2/3), 215–222.
- Ball, S.J., 2009. Academies in context: politics, business and philanthropy and heterarchical governance. *Manag. Educ.* 23 (3), 100–103.
- Ball, S.J., 2015. The legacy of ERA, privatization and the policy ratchet. *Educ. Manag. Adm. Leader.* 36 (2), 185–199.
- Ball, S.J., 2018. The tragedy of state education in England: reluctance, compromise and muddle—a system in disarray. *J. Br. Acad.* 6, 207–238.
- Ball, S.J., Bowe, R., 2020. The neoliberalization of the state, the processes of “fragmentation,” and research implications of the New Political Terrain of English schooling. In: Brown, A., Wisby, E. (Eds.), *Knowledge, Policy and Practice in Education and the Struggle for Social Justice: Essays Inspired by the Work of Geoff Whitty*. UCL Press, London.
- Ball, S.J., Maguire, M., Braun, A., 2012. *How Schools Do Policy Enactments in Secondary Schools*. Routledge, London.
- Barton L., Slee R., 2010. Competition, selection and inclusive education: some observations. *Int. J. Incl. Educ.* 3 (1), 3–12.
- Bash, L., Coulby, D., 1989. *The Education Reform Act: Competition and Control*. Cassell, London.
- Beck, U., 2007. Beyond class and nation: reframing social inequalities in a globalizing world. *Br. J. Sociol.* 58 (4), 679–705.
- Beck, U., Beck-Gernsheim, E., 2002. *Individualization: Institutionalized Individualism and Its Social and Political Consequences*. Sage, London.
- Biesta, G., 2007. Why “what works” won’t work: evidence-based practice and the democratic deficit in educational research. *Educ. Theor.* 57 (1), 1–22.
- Biesta, G., 2020. *Educational Research: An Unorthodox Introduction*. Bloomsbury, London.
- Biggar, L.D., Fels, R.A., 2017. *Choice, Competition and Markets: Education, Health Care, and Public Services*. OECD, Paris.
- Bines, H., 1995. Special educational needs in the market place. *J. Educ. Pol.* 10 (2), 157–171.
- Cabinet Office, 2010. *The Coalition: Our Programme for Government*. Cabinet Office, London. <https://www.gov.uk/government/publications/the-coalition-our-programme-for-government>.
- Chow, D., Bracci, E., 2020. Neoliberalism, accounting, and the transformation of subjectivities in social work: a study on the implementation of personal budgets. *Financ. Account Manag.* 36, 151–170.
- Colebatch, H., 2006. What work makes policy? *Pol. Sci.* 39 (4), 309–321.
- Conservative Party, 2007. *Conservative Party Commission on Special Education*. Conservative Central Office, London.
- Cox, R.H., 1998. The consequences of welfare reform: how conceptions of social rights are changing. *J. Soc. Pol.* 27 (1), 1–16.
- DCSF, 2010. *Improving parental confidence in the special educational needs system: An implementation plan*. DCSF, London.
- DeMoss, K., 2003. Who’s accountable to the constitution? Thirty years of judicial politics in state education finance litigation. *Peabody J. Educ.* 78, 44–65.
- DfE, 2010. *The Importance of Teaching The Schools White Paper 2010*. Department for Education.
- DfE, 2011. *Support and aspiration: A new approach to special educational needs and disability: A consultation*. UK Government, London.
- DfE, 2020. *Academy Trust Governance—Structures and Role Descriptors: October 2020*. DfE, London.
- DfE, 2022a. *Special Educational Needs in England*. UK Government. <https://explore-education-statistics.service.gov.uk/find-statistics/special-educational-needs-in-england#data-Block-678da1e3-97e6-42f1-165e-08da47b0392d-tables>. (Accessed 20 June 2022).
- DfE, 2022b. *Education, Health and Care Plans*. UK Gov Statistics. <https://explore-education-statistics.service.gov.uk/find-statistics/education-health-and-care-plans>.
- DFEE, 1997. *Excellence for All Children: Meeting Special Educational Needs*. DFEE, London.
- DFEE (Department for Education and Employment), 1997. *Excellence for All Children: Meeting Special Educational Needs*. DFEE, London.
- DFES, 2001. *Inclusive Schooling Children with Special Educational Need.s*. Department for Education and Skills.
- Diamond, P., 2013. The progressive dilemmas of British social democracy: political economy after New Labour. *Br. J. Polit. Int. Relat.* 15, 89–106.
- DoH (Department of Health), 2012. *Evaluation of the Personal Health Budget Pilot Programme*. Department of Health, London.
- Esping-Andersen, G., 1999. *Social Foundations of Post-Industrial Economies*. Oxford University Press, Oxford.
- Exley, S., 2012. The politics of educational policy making under New Labour: an illustration of shifts in public service governance. *Pol. Polit.* 40 (2), 227–244.
- Fleming, P., McGilloway, S., Barry, S., 2016. The successes and challenges of implementing individualised funding and supports for disabled people: an Irish perspective. *Disabil. Soc.* 31 (10), 1369–1384.
- Foucault, M., 2002. *Archaeology of Knowledge*. Routledge, London.
- Foucault, M., 2008. *The Birth of Biopolitics: Lectures at the Collège de France 1978–1979*. Palgrave Macmillan, Basingstoke.
- Giddens, G.A., 1990. *Consequences of Modernity*. Polity, Cambridge.
- Gingrich, J.R., 2011. *Making Markets in the Welfare State*. Cambridge University Press, Cambridge.
- Habermas, J., 1987. *The Theory of Communicative Action: Lifeworld and System: A Critique of Functionalist Reason*, vol. 2. Polity Press, Cambridge.
- Hall, S., Jacques, M., 1983. The great moving right show. In: Hall, S., Jacques, M. (Eds.), *The Politics of Thatcherism*. Wishart, London.
- Hicks, T., 2015. Inequality, marketization and the left: schools policy in England and Sweden. *Eur. J. Polit. Res.* 54 (3), 326–342.
- HoC (House of Commons), 2006. *The Schools White Paper: Higher Standards, Better Schools For All First Report of Session 2005–06: Volume I*. HMSO, London.
- Jessop, B., 2015. *The State: Past, Present, Future*. Polity Press, Cambridge.
- Jones, K., 1989. *Right Turn: Conservative Revolution in Education*. Radius Books, London.
- Kearns, M., 2018. *Worth the Risk? A Study of Mature Students in Two Irish Universities*. Unpublished PhD Thesis. Trinity College Dublin, Dublin.
- Kohn, M., 2020. Public goods and social justice. *Perspect. Polit.* 18 (4), 1104–1117.
- Kopanke, J.R., 2020. *A contemporary understanding of the effects of the third wave of school finance litigation*. Dissertations 720. <https://digscholarship.unco.edu/dissertations/720> University of Northern Colorado.
- Lamb, B., 2009. *Report to the Secretary of State on the Lamb Inquiry Review of SEN and Disability Information*. DCSF, London.
- LeGrand, J., 1991. Quasi-markets and social policy. *Econ. J.* 101 (408), 309–321.
- Levitas, R. (Ed.), 1986. *The Ideology of the New Right*. Polity Press, Cambridge.
- Levitas, R., 2005. *The Inclusive Society? Social Exclusion and New Labour*, second ed. Palgrave MacMillan, Basingstoke.
- Local Government Association (LGA), 2020. *Special Educational Needs and Disability Funding House of Commons 29 January 2020*. <https://www.local.gov.uk/sites/default/files/documents/29012020%20LGA%20briefing%20-%20SEN%20support-WEB.pdf>.
- Loxley, A., 1999. *An exploration into the impact of LMS on the policy and provision of SEN*. University of Bath, Bath.
- Lui, Y., Bessudnov, A., Black, A., Norwich, B., 2020. School autonomy and educational inclusion of children with special needs: evidence from England. *Br. Educ. Res. J.* 46 (3), 532–552.
- Martínez, C.R., Pritchard, J.N., 2019. *Proceed With Caution What Makes Personal Budgets Work*. Reform, London.
- Marx, K., 1926. *The Eighteenth Brumaire of Louis Bonaparte*. G. Allen & Unwin, Ltd, London.
- McNay, L., 2009. Self as enterprise dilemmas of control and resistance in Foucault’s the birth of biopolitics. *Theory. Cult. Soc.* 26 (6), 55–77.
- National Audit Office, 2019a. *Support for Pupils With Special Educational Needs and Disabilities in England: Report by the Comptroller and Auditor General (Summary)*. <https://www.nao.org.uk/wp-content/uploads/2019/09/Support-for-pupils-with-special-education-needs-summary.pdf>.
- National Audit Office, 2019b. *Support for Pupils With Special Educational Needs and Disabilities in England: Report by the Comptroller and Auditor General (Full Report)*. <https://www.nao.org.uk/wp-content/uploads/2019/09/Support-for-pupils-with-special-education-needs.pdf>.

- National Audit Office, 2021. School Funding in England. https://www.naoA...org.uk/press-release/school-funding-in-england/?utm_source=rss&utm_medium=rss&utm_campaign=school-funding-in-england.
- Newman, J., Clarke, J., 1994. Going about our business: the managerialisation of public services. In: Clarke, J., Cochrane, A., MacLaughlin, E. (Eds.), *Managing Social Policy*. Sage, London.
- NHS UK, 2022. Personal Health Budgets: Quarter 4 2021-22. NHS Digital.
- Norwich, B., 2008. What future for special schools and inclusion? conceptual and professional perspectives. *Br. J. Spec. Educ.* 35 (3), 136–143.
- Norwich, B., Black, A., 2015. The placement of secondary school students with statements of special educational needs in the more diversified system of English secondary schooling. *Br. J. Spec. Educ. Needs* 42 (2), 128–151.
- OECD, 2018. *OECD Handbook for Internationally Comparative Education Statistics 2018: Concepts, Standards, Definitions and Classifications*. OECD Publishing, Paris.
- OECD, 2021. *Education at a Glance*. OECD, Paris.
- Ofsted (Office for Standards in Education), 1999. *LEA Support for School Improvement: Framework for the Inspection of Local Education Authorities*. Ofsted, London.
- Ofsted (Office for Standards in Education), 2010. *The Special Educational Needs and Disability Review*. Ofsted, Manchester.
- Penna, S., O'Brian, M., 1998. *Theorising Welfare, Enlightenment and Modernity*. Sage, London.
- Powell, M. (Ed.), 2008. *Modernising the Welfare State: The Blair Legacy*. Polity Press, Bristol.
- Ranson, S., 2003. Public accountability in the age of neoliberal governance. *Journal of Education Policy* 18 (5), 459–480.
- Rix, J., 2011. Repositioning of special schools within a specialist, personalised educational marketplace—the need for a representative principle. *Int. J. Incl. Educ.* 15 (2), 263–279.
- Rix, J., Sheehy, K., Fletcher-Campbell, F., Crisp, M., Harper, A., 2013a. *Continuum of Education Provision for Children With Special Educational Needs: Review of International Policies and Practices*. The National Council for Special Education, Trim.
- Rix, J., Sheehy, K., Fletcher-Campbell, F., Crisp, M., Harper, A., 2013b. Exploring provision for children identified with special educational needs: an international review of policy and practice. *Eur. J. Spec. Needs Educ.* 28 (4), 375–391.
- Ross, T., 1990. *Extending Local Management of Schools to Special Schools. A feasibility study for the Department of Education and Science*. Touche Ross, London.
- Rouse, M., Florian, L., 1997. Inclusive education in the market-place. *Int. J. Incl. Educ.* 1 (4), 323–336.
- Sailor, W., 2015. Advances in schoolwide inclusive school reform. *Remedial Spec. Educ.* 36 (2), 94–96. <https://doi.org/10.1177/0741932514555021>.
- Starke, P., 2006. The politics of welfare state retrenchment: a literature review. *Soc. Pol. Adm.* 40 (1), 104–120.
- Taylor, D., 1998. Social identity and social policy: engagements with postmodern theory. *J. Soc. Pol.* 27 (3), 329–350.
- Taylor-Gooby, P., 2011. Root and branch restructuring to achieve major cuts: the social policy programme of the 2010 UK coalition government. *Soc. Policy Adm.* 61–82.
- Thomas, G., Loxley, A., 2001. *Deconstructing Special Education Reconstructing Inclusion*. Open University Press, Maidenhead.
- Thomas, G., Loxley, A., 2007. *Deconstructing Special Education Reconstructing Inclusion*, second ed. Open University Press, Maidenhead.
- Thomas, M.G., Loxley, L.A., 2022. *Deconstructing Special Education Reconstructing Inclusion*, third ed. Open University Press, Maidenhead.
- Tomlinson, S., 1982. *A Sociology of Special Education*. Routledge, London.
- Tomlinson, S., 2012. The irresistible rise of the SEN industry. *Oxf. Rev. Educ.* 38 (3), 267–286.
- Tomlinson, S., 2015. Special education and minority ethnic young people in England: continuing issues. *Discourse* 37 (4), 513–528.
- Tomlinson, S.P., 2017. *A Sociology of Special and Inclusive Education: Exploring the Manufacture of Inability*. Routledge, London.
- UK Gov, 2009. *The Community Care, Services for Carers and Children's Services (Direct Payments) (England) Regulations 2009*. London.
- UK Gov, 2010. *Academies Act 2010*. London.
- UK Gov, 2014a. *The Special Educational Needs (Personal Budgets) Regulations 2014*. London.
- UK Gov, 2014b. *Children and Families Act*. London.
- UK Gov, 2015. *Special Educational Needs and Disability Code of Practice: 0 to 25 Years*. Department for Education & Department of Health, London.
- Waisath, W., McCormack, M., Stek, P., Heymann, 2022. Dismantling barriers and advancing disability-inclusive education: an examination of national laws and policies across 193 countries. *Int. J. Incl. Educ.* 1–16.
- West, A., Bailey, E., 2013. The development of the academies programme: “privatising” school-based education in England 1986–2013. *Br. J. Educ. Stud.* 61 (2), 137–159.
- Wisby, E., Whitty, G., 2016. Education in England—a testbed for network governance? *Oxf. Rev. Educ.* 42 (3), 316–329.
- Young, H., 2013. *One of Us*. Pan MacMillan, London.

“First you have to see me as a human being”: disabled students’ experiences of schooling

Gill Rutherford, University of Otago College of Education, Dunedin, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	228
Understanding students’ experiences of school	229
Disability: a bad thing?	229
Inclusive education?	229
School cultures	230
Teachers’ values, beliefs and assumptions	230
Pedagogical practices	230
“Special” education practices	231
Support: helping, hovering or hindering?	231
Friends matter	231
The essence of students’ experiences	232
Theoretical framing of students’ experiences	232
Disability: not a bad thing?	232
Is having a say sufficient?	232
Justice for all or just some?	233
Teacher education: part of the solution to educational inequities (or not)?	233
What is education for?	233
Do we need to change our minds?	234
Teacher education practices	234
Disability Studies in Education (DSE)	234
Disability and intersectionality	234
Pedagogies and curriculum	234
Practicum matters	235
End of chapter thoughts	235
References	236

Introduction

I began my life in teaching as the coordinator of a “special needs unit” in a New Zealand rural high school. While I have many memories of my work with students, one is particularly salient:

I remember Bill (not his real name), whose apparent shame at being associated with the unit was evident in the way he sought to leave S2 (euphemism for the unit) before the bell went. He would surreptitiously check that no-one was outside before scurrying out the door of S2, to reach the safe anonymity of the quad, where his demeanour changed from one of furtive urgency to nonchalance, as if he’d just stepped out of a “regular” class. I see myself now as an accomplice in this act, knowing that I could incur the principal’s wrath by this early release from class, yet considering Bill’s self-worth and dignity worth the risk. I thought it futile to further alienate him from school by insisting he spend a few more minutes on an academic task. In that time before the bell, his thoughts had already moved from learning to read to how he could get out of (escape?) the unit without being seen by his “normal” peers. Looking back, I realize that Bill’s response to “special” education served as a powerful lesson in its limitations and even harmfulness. Along with other students I had the privilege of learning with and from, Bill’s actions highlighted for me the wrongness of special education, and the need to work for and within a more equitable education system, in which students are not marginalised because of their human difference. Bill taught me that education should never shame.

Some 30 years after crossing the threshold of my first class as a new high school teacher, I continue to witness how damaging education can be for some students, based on inherently flawed assumptions about what it is to be human. My own experiences as a teacher and now teacher educator are echoed in a growing body of research that documents students’ educational experiences from their own perspectives (e.g., [Connor and Cavendish, 2020](#)). This literature reiterates the imperative of valuing children and young people’s perspectives about their worlds, as mandated by the United Nations Convention on the Rights of the Child (UNCRC) ([United Nations, 1989](#)) and the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) ([United Nations, 2006](#)).

The purpose of this chapter is to examine the emotional impact of markers of difference that mark the development of many students’ identities as “full-fledged human beings” who are “capable of achievement and worthy of respect” (Lipsky and Gartner, 1987, p. 69). The importance of such inquiry is emphasized by Jóhannsdóttir et al. (2020) in their research regarding the impact of shame on disabled young people. Defining shame as “psychological pain” derived from thinking of oneself as flawed and “lacking dignity or worth” (p. 4), they emphasize that shame cannot be conceptualized as an individual matter only, and that its “social determinants” (p. 4) must also be considered. How do external social factors, such as dominant exclusionary ideologies, structures and practices affect how students feel about themselves—who they are and who they may become? How can teachers/schools, who have the power to change mechanisms of marginalization, offer *all* students equitable educational opportunities to flourish in life?

The chapter begins with students’ interpretations and affective experiences of disability and education, which are interpreted within a framework that draws from theoretical concepts of disability and justice. The chapter then outlines what can be done within teacher education contexts to address often uncontested educational wrongs that can warp students’ lives.

Understanding students’ experiences of school

[Don’t] focus too much on the fact that you’re different, just, be yourself and usually things will, probably work out for you.

Ethan, cited in Jessup et al. (2018, p. 94).

This section draws from a range of international literature in which disabled students share their thoughts about school. While it is beyond the scope of this chapter to review all relevant research, the material chosen relates to students in both primary/elementary and secondary/high school contexts in multiple countries. Most of the sources were published between 2017 and 2020. Numerous student quotes are threaded throughout this chapter, to show respect for students’ own words.

Disability: a bad thing?

Is it ok to have a disability? ... Can I tell people that I have Down syndrome?

Minal, cited in Lalvani (2015, p. 14).

That disability was regarded as never a “good” thing, and that something was “wrong” with a person was the shared perception of most of the students in the reviewed research (e.g., Banks, 2017). For example, in a US study focusing upon adolescents’ meanings of disability labels and stigma, Luke explained that disability is “something that means the person is less able” and that “most of the time people wouldn’t say it’s a good quality to have in someone” (Mueller, 2019, p. 276). In another study, Jackson admitted that, “Even though my mother told me that having a disability wasn’t a bad thing, and wasn’t the end of the world, and didn’t mean I was dumb, I still felt a little self-conscious about it” (Banks, 2017, p. 102). Internalizing pejorative assumptions regarding disability and the unquestioned ideology of normalcy as the measure of human worth, it seems rational that some students chose to “not tell” or hide their difference from others, particularly their peers.

In contrast, some students recognized the contextual nature of disability. For instance, Luke claimed that he was not “less able,” providing examples of contexts in which he was academically successful (Mueller, 2019, p. 276). Nathan, who experienced Fetal Alcohol Spectrum Disorder (FASD), rejected health care providers’ assumption that there was “something wrong with him” (Brenna et al., 2017, p. 225). He accepted that the effects of FASD were “real,” and although having this condition was problematic at times, it was also “a source of self-worth” (p. 225).

Inclusive education?

Inclusion to me means that people accept you for who you are. I don’t feel included at school because people see me as “weird,” and so they don’t seem to include me. ... Every person should be included because it’s important to feel a part of something.

Sean, cited in Margrain and Farrugia (2018, p. 100).

Schools act as sorting mechanisms (Slee, 2018), whether through exclusive, segregated/special, or mainstream/integrated educational structures. These forms of education are however ill-defined; in this paper, definitions from UNCRPD Article 24, General Comment No. 4 (United Nations, 2016) are utilized as a means of contextualizing students’ experiences.

UNCRPD is the first international rights mechanism to specify a legally binding right to inclusive education for its signatories. General Comment No. 4 (United Nations, 2016), an update of UNCRPD, defines exclusion from education as a denial of the right to education while segregation/special education is that which is provided separately for disabled students. Integrated/mainstream education involves the placement of “special needs students” and/or those with other marginalized identities into “regular” classes, a process of assimilation into largely ableist schools. Finally, inclusive education necessitates a rights-based transformation of education systems, to remove barriers to all students’ presence, participation, achievement, social relationships and overall well-

being (United Nations, 2016). Of particular relevance to this chapter is the principle that education must value and respect all students' "inherent dignity and autonomy" (United Nations, 2016, p. 3). Does it, however?

Participants in the research that informs this chapter unequivocally indicated that schools are good places for some students; for others, not so much (e.g., Dadvand, 2020). Regardless, schools play a critical role in shaping students' identities and future life possibilities. Across the studies, the majority of students attended mainstream schools, while the others were enrolled in special schools. Comparatively few studies thus far have documented students' experiences of genuinely inclusive schools, perhaps because the process of developing inclusive education is relatively recent and requires fundamental changes in how we "do" education. Students shared thought-provoking observations about the impact of school culture, teachers' beliefs, pedagogical practices and social/relational opportunities on their learning and sense of self-worth.

School cultures

Inclusion "is about the school mentality."

Ro, cited in Goodall (2020, p. 1298).

School culture is apparent in overt and covert ways, from espoused values, leadership priorities, staff attitudes, structures and routines, and an often tacit "this is how we do things here." The culture of a school can often be "felt" and have a significant impact on students. James, an Australian student declared, "I have never liked school ... I don't feel safe here ... I don't feel welcomed here. It's just an alien environment to me" (Dadvand, 2020, p. 1). Similarly, a Greek student remarked, "I felt that I was losing my dignity and self-respect" (Vlachou and Papananou, 2015, p. 80). Why might this be the case? What impact can teachers' thinking about students have on the development of welcoming or inhospitable class and school cultures, in which students may be valued or denigrated?

Teachers' values, beliefs and assumptions

Care about your students. Make sure you like teaching.

Student, cited in Connor and Cavendish (2020, p. 301).

Parker Palmer (2017) asserts that we teach who we are. Who the teacher is in terms of their personal characteristics matters to students. Being empathetic, non-judgmental, making time to listen and being "on my side" (Vlachou and Papananou, 2015, p. 80) were behaviors frequently identified as mattering to students (Vetoniemi and Kärnä, 2019). Such teachers were not found in all schools however. Instead, a sense that some teachers were unwilling to engage with and teach disabled students was evident in numerous ways. Teachers' indifference and assumptions of incompetence were described by many students (Connor and Cavendish, 2020; Margrain and Farrugia, 2018). Philip, whose teacher interacted with every student but him, felt that "you simply don't exist" (Vlachou and Papananou, 2015, p. 80).

Pedagogical practices

I shouldn't be treated any more or any less than any other kids without a learning condition, but I should have the tools that I need to learn in an equal way to other kids without my condition, and at the moment I don't feel I have that.

Dave, cited in Margrain and Farrugia (2018, p. 103).

As noted by Palmer (2017), teachers' personal characteristics, values and beliefs are inseparable from their practices. Some teachers were pedagogically responsive by ensuring that environments, curriculum, teaching strategies, resources and assessments were accessible and meaningful for all students (Vlachou and Papananou, 2015). Engaging student-centered approaches and utilizing Universal Design for Learning (UDL) principles (Connor and Cavendish, 2020) to "reach and teach" all students were practices that students appreciated. Furthermore, focusing on strengths and creating opportunities for students to reveal and develop their competence was critical (Brenna et al., 2017). Making sure students had a "place to shine" (Jessup et al., 2018, p. 96; see also Vetoniemi and Kärnä, 2019) also helped students feel good about themselves.

In mainstream contexts, basic considerations such as the location of classes, seating arrangements and grouping were identified by students as important (Fellner et al., 2017; Vetoniemi and Kärnä, 2019). The use of "ability" grouping, a practice underpinned by "fixed-ability"/determinist assumptions that one is born with a certain amount of (narrowly defined) intelligence, was common. Such practices are increasingly recognized as ineffective and indeed harmful to some students in terms of generating feelings of inadequacy and not being "smart" enough (Andriana and Evans, 2017, p. 182).

According to a student in Banks' (2017, p. 102) research, "there wouldn't be learning disabilities if teachers ha[d] many different ways of teaching." A predominantly teacher-transmission and "one size fits all" approach impeded learning (Connor and

Cavendish, 2020) and social interactions, resulting in what Vetoniemi and Kärnä (2019, p. 6) call “pedagogical segregation.” Failure to cater for those who did not “fit” resulted in students feeling “completely left out” (Vlachou and Papananou, 2015, p. 78).

“Special” education practices

Don’t make it so obvious that they need special help.

Laritza, cited in Fellner et al. (2017, p. 339).

To varying extents, mainstream schools utilized special education practices as a means of “including” disabled students. Students spoke of being labeled, placed in “special ed” programs and having “special ed” teachers and teacher aides. According to Luke, “being in special ed marks you as different. And no one wants to be different” (Mueller, 2019, p. 276). Being pulled out periodically to a resource room led to Javier wondering: “Why am I not in classes with everyone else?” (Mueller, 2019, p. 273). Despite having internalized the wrongness of special, Laritza tried to resist this: “I try my best to change and not be what they mark me, what they label me as. I can prove them wrong” (Fellner et al., 2017, p. 332). The need to prove one’s worth is a heavy and unnecessary burden that many students carry (Markham, 2020).

The complexity and impact of special needs discourses, so adroitly critiqued by Slee in many of his publications (e.g., 2018), is further illustrated by Laritza in her distinction between “special” and “normal” students. She explained, “we usually sit with our own type, you know normal people [gesturing to one side], us [gesturing to the other side]” (Fellner et al., 2017, p. 333). In contrast to such a clear demarcation, the conflation of “special” and inclusive was evident in Andriana and Evans’ (2017) research (e.g., attaching the label of “inclusive/ion” to signs, children/“inclusion child” and teachers [p. 182])—as if by doing so would make it so.

Quite different responses to special education are outlined by Goodall (2020)—because of the failure of some mainstream schools, participants preferred a special school. Its appeal was described in terms of feeling “less lonely because you might feel included as there are others like you” (p. 1299). Interestingly, students in this study felt included in a segregated special school, Wade asserting that inclusion “is more about how a person feels. Feeling valued as a person and feeling you matter” (p. 1298).

Support: helping, hovering or hindering?

It just depends on how children are supported.

Sarah Jane, cited in Goodall (2020, p. 1301).

The kind of support provided to students—and how—matters. Often regarded as “the solution to inclusion,” teacher aides have typically served as the main source of support for disabled students, with varying effects (Rutherford, 2012). As Giangreco et al. claimed in 1997, teacher aides can help, hover and/or hinder. For example, Sean said that a teacher aide “makes me feel like there’s something wrong with me. Like I’m incapable. ... It’s like she stalks me” (Margrain and Farrugia, 2018, p. 101). In contrast, some students appreciated support (e.g., Rutherford, 2012). Secondary school students may encounter an unnecessary academic-social dilemma, in having to choose between teacher aide support for their learning or having relationships with peers, due to the stigma associated with the former (Rutherford, 2012).

Other forms of support may include special education teachers and therapists; how they work with students often determines whether or not “help” lives up to its name or is regarded as a surveillance and regulatory practice (Mueller, 2019). A “sponsor system” involving peers, coaches or mentors was suggested as a way of offering support without shaming students. To minimize notions of neediness, the term “sponsor system” was preferred to “caregiver” (Brenna et al., 2017, p. 224).

Friends matter

I have been excluded from groups of friends a lot. ... They don’t really notice that you’re even there. They just don’t want to see you.

Jasmine, cited in Jessup et al. (2018, p. 97).

Students spoke of a range of social experiences, from being thankful for trusted friends to feeling excluded and invisible (Jessup et al., 2018). The power of friends was immeasurable in terms of its intrinsic value. For example, Nefeli realized that with friends, “you’re no different as a human being, that you’re an equal among equals” (Vlachou and Papananou, 2015, p. 80). Nathan valued his friends for not judging him, and felt safe to share with them his experiences of FASD (Brenna et al., 2017).

Conversely, being bullied and wounded by slurs such as “retard” (Margrain and Farrugia, 2018, p. 100) and “special bastard” (Goodall, 2020, p. 1302) was not uncommon for students (Andriana and Evans, 2017). Ironically, while using a wheelchair made him hyper-visible, a student believed that he was invisible to his peers, asking “why do people look through you and see only a wheelchair? They don’t see me” (Vetoniemi and Kärnä, 2019, p. 8). Focusing primarily on the device that enables a person to operate in a world not built for every-body can elicit feelings of shame as well as frustration.

The essence of students’ experiences

The above discussion reveals often unnoticed or disregarded inequities that have confronted many disabled students. Of fundamental significance are the responses of peers and school staff. These ranged from respect to “not noticing” difference, to indifference, to ignoring/shunning students as if invisible, and/or noticing for the wrong reasons. [Lalvani \(2015, p. 9\)](#) describes the benevolent and ableist response of not noticing someone’s difference as the “goodness of not noticing,” analogous to being “color blind” ([Sapon-Shevin, 2017](#)) to racial difference. Overall, the most important message from students is the desire to simply “be yourself” (Ethan, cited in [Jessup et al., 2018, p. 94](#)). Cara wanted people to “realize I’m there” (p. 93); Jasmine sought to be “noticed, but in a good way” (p. 94); and Emma wished to be accepted “for the way you are” (p. 93). These are not unreasonable desires, but rather fundamental human needs.

Students have generously and honestly shared what it feels like to be on the receiving end of schooling. While recognizing the uniqueness of each person and their circumstances, the overall depiction of disabled students’ education indicates a paucity of equity and an abundance of injustices. The emotional costs of schooling for many students are evident in their raw accounts of rejection and shame, whether explicitly or implicitly at structural and/or personal levels. In the following section, students’ realities are linked to theory in an attempt to understand their experiences.

Theoretical framing of students’ experiences

Discussion of disability theorizing, UNCRC ([United Nations, 1989](#)) and theoretical frameworks of justice is the focus of this part of the chapter. The rationale for drawing from these frameworks is to offer possible understandings of why many students feel alienated within and by schools, and to consider how these theoretical tools might be utilized to bring about “just education” ([Connell, 2012, p. 681](#)) for all students, not just some.

Disability: not a bad thing?

Students’ interpretations of disability reflect the internalization of hegemonic narratives that position disability as an essentialized, pathological flaw within individuals, a marker of difference (read: deviance) and shame, all of which depreciate a person’s human worth ([Dadvand, 2020](#)). In turn, individuals’ assumed inferiority legitimizes the compromising of their right to education.

Theoretically, [Thomas’ \(2004\)](#) social relational model is helpful in conceptualizing the complexity of disability. She recognizes the reality of impairment and its effects—that physical, sensory, cognitive, and emotional differences are real, and do impact on how a person makes their way through life. For example, if someone has cerebral palsy, this is a core part of who they are that cannot/ought not be denied. It should never however diminish a person’s humanity. Thomas also highlights the impact of disabling factors that are external to the individual, referring to these as barriers to doing and barriers to being. The former include anything that limits or denies a person’s participation in any aspects of life (e.g., ableist schools). Barriers to being are psychosocial factors that affect a person’s sense of self, worth, and identity (e.g., the use of derogatory language). These external and interrelated barriers to doing and being are imposed upon a person, exacerbating the effects of their impairment. Thomas’ model exemplifies the socially constructed, contextual and relational nature of disability ([Mueller, 2019](#)), as evident in the accounts of students outlined above.

More recently, the notion of affect, and in particular shame, has featured in Critical Disability Studies literature ([Jóhannsdóttir et al., 2020](#)). As previously mentioned, emotions are not solely individual psychological matters but also need to be understood alongside normative social and cultural power contexts. As claimed by a participant in Jóhannsdóttir et al.’s study, “others, the ones who are not disabled, are better than the disabled ones;” they wanted to be “normal” rather than disabled (p. 11). The role of affect aligns with [Thomas’ \(2004\)](#) barriers to being, and are helpful in making sense of students’ experiences (e.g., asking whether it is ok to have a disability).

Current thinking about disability also positions disabled people as members of a subjugated minority group, akin to other such identity groups founded on commonalities of race, ethnicity, gender diversity and other differences from white, middle-class, “able-bodied” and other hegemonic norms ([Lalvani, 2015](#)). As with such minority groups who experience, for example, racism, many disabled people’s lives and identities are shadowed by ableism, in which “normal” “able-bodied” people are privileged over those who are not ([Banks, 2017](#)). Ableist assumptions are examples of barriers to being and doing, which obstruct students’ opportunities to exercise autonomy regarding their lives, which, as mentioned, is a core principle of UNCPRD and an integral part of UNCRC.

Is having a say sufficient?

“I don’t have a say. [...] I like to make a difference. I’d like to change the way things work” (James, cited in [Dadvand, 2020, p. 10](#)). Article 12 of UNCRC ([United Nations, 1989](#)) mandates the right of all children and young people to have their perspectives accorded due weight in matters that affect them. In her critique of UNCRC Article 12, [Lundy \(2007\)](#) emphasizes the need to contextualize its enactment, making it clear that “Voice” is not enough. To genuinely implement Article 12, Lundy claims that four factors, namely Space, Voice, Audience and Influence (p. 932), must be considered. Space involves all children and young people having opportunities to state their views in nondiscriminatory environments. Voice refers to facilitating the expression of views, in whatever manner works for a child, which importantly recognizes the validity of communication other than verbal. Having someone who

listens and gives due weight to children’s perspectives, however these are conveyed, is the essence of the Audience factor. Influence involves taking children’s views seriously and, wherever possible and appropriate, acting upon them.

With regard to the experiences of students outlined in this chapter, it is evident that recognition of these factors and related UNCRC Articles is sporadic. I suggest that a prerequisite to being afforded space, voice, audience and influence is basic recognition of disabled students in terms of their humanity and capacities—that they are indeed worthy of having a say. If, as numerous students claim, they feel unwelcome, unnoticed and/or diminished by their difference, it is unlikely that any aspects of Lundy’s (2007) model apply. We cannot address what we do not see, recognize or value.

Justice for all or just some?

Recognition in the context of this chapter refers to the fundamental need to be seen, known and validated as a human being. Recognition is a core element of certain theories of justice. Fraser (2008) proposes that three dimensions must be addressed to generate justice and “parity of participation” (p. 405): (1) economic redistribution (redistribution of funding for teacher aides for certain students); (2) cultural recognition (positive recognition and respect for every student); and (3) political representation (students’ involvement in that which concerns them).

As with matters relating to disability and affect, contextual influences determine how each dimension is played out. How teacher aides work with students can influence how they are recognized. The inappropriate use of teacher aides can be a barrier to doing, which also generates a barrier to being in terms of embarrassment/stigma in having such support (Thomas, 2004). Were the provision of human resources managed differently, it is more likely that the other two dimensions could facilitate parity of participation and development of positive identities.

While Fraser’s theorizing is widely utilized by scholars, its minimal recognition of disabled people is ironic yet not uncommon. Pugach et al. (2021) consider disability is often “more of an afterthought” (p. 243) (if thought of at all), and describe the inclusion of disability in discussions of social justice within teacher education research as “a precarious guest at the table” (p. 237). Such theoretical precariousness is mirrored in the real-life experiences of students discussed in this chapter. Being at the table/in class is not however enough. As argued earlier, a prerequisite to having space, a voice, audience and influence (Lundy, 2007) is being recognized as a human being.

Interdependence and respectful relationships are fundamental to effective teaching and learning. Christensen and Dorn (1997) emphasize the primacy of relationships in their proposal of a theory of relational justice. They claim that the “goal of social justice should be a just set of relationships” (p. 193)—that justice is evident (or absent) in the quality of relationships. In a similar vein, Young and Quibell (2000) reason that injustices can only be resolved if people understand each other. This involves recognition of and knowing individuals through relationships as well as constructing their identities in terms of their humanness (Bogdan and Taylor, 1992) and competence (Biklen and Burke, 2006). Conceptualizing justice as a dynamic interpersonal process (“doing right by”) brings to life and makes possible the enactment of formal rights, such as those documented in UNCRC and UNCRPD.

It could be argued that the teachers described by students as indifferent to difference were mis- (or not) educated about disability, rights and justice. The following section offers some suggestions for addressing this concern.

Teacher education: part of the solution to educational inequities (or not)?

So, what should we do about it?

Amelia, cited in Lalvani (2015, p. 12).

The final part of this chapter explores what we can *do* to support students to *be* themselves and to thrive within schools. I have chosen to focus on what can be done within teacher education, on the basis of its responsibility to develop new teachers. That we must do better is identified in UNCRPD Article 24, General Comment No. 4 (2016), in which barriers to inclusive education include:

- (a) lack of understanding of disability as a human rights matter;
- (b) unchallenged discrimination and low educational expectations;
- (c) lack of knowledge regarding the efficacy of inclusive education for all students;
- (e) “lack of political will, technical knowledge, and capacity in implementing the right to inclusive education including *insufficient education of all teaching staff*” [emphasis added] (United Nations, 2016, p. 2).

Specific examples of how teacher education programs do and can address such barriers are first contextualized in terms of the overall purpose of education and the impact of teachers’ assumptions upon educational (in)equities.

What is education for?

At a macro level, how do we conceptualize the overall purpose of education? As Slee has often asked (e.g., 2018), what kind of a world do we want to live in? Scottish educational philosopher John Macmurray (1891–1976) offers wise counsel in his belief

that education is a means of learning to be human, to lead good lives together in community, as "an interdependence of equals" (Macmurray, 1964, p. 21). He claims that "our ability to enter into fully personal relations with others is the measure of our humanity. For inhumanity is precisely the perversion of human relations" (1958, cited in Macmurray, 2012, p. 670).

Macmurray's beliefs align with students' experiences, in which the need to be recognized and respected as a human being in relationship with others is compellingly evident. Slee's (2011) conceptualization of education as an apprenticeship in democracy is also relevant, with its focus on education not only as a right in and of itself, but also as a means of developing rights-bearing citizens. Nevertheless, as documented in this chapter, individuals continue to be disenfranchised and dehumanized within democracies, on the basis of human differences over which they have no control and which are imbued with socially constructed wrongness. I wonder whether (akin to Macmurray's idea that "curriculum and pedagogy should ... be dominated by apprenticeship in humanity through community" (Stern, 2012, p. 740) it would be useful to rethink the purpose of education as an apprenticeship in both democracy *and* humanity.

Do we need to change our minds?

The power that teachers wield, for good or otherwise, is immeasurable and far-reaching. As evident in students' accounts, how teachers think matters. Conceptualizing disability as a natural part of being human is a prerequisite to recognizing and respecting disabled students as human beings who have a right to quality schooling. For some future and current teachers (and teacher educators), coming to such a new way of thinking necessitates a process of "unlearning" deficit assumptions. It is sobering to note Banks' (2017) report that students in her study believed that "teachers' depreciatory beliefs regarding race, gender, and disability were unchangeable" (p. 104).

King (1991) coined the term "dysconsciousness" to describe "an uncritical habit of mind (including perceptions, attitudes, assumptions, and beliefs) that justifies inequity and exploitation by accepting the existing order of things as given" (p. 135). Lalvani and Broderick (2015) have since applied the concept of dysconsciousness to ableism, describing "dysconsciousness ableism" as "the limited understandings that our students may hold about the nature of inequity related to disability/ability" (p. 174). In other words, the teachers described by students as indifferent and disavowing their very presence could be labeled as dysconscious ableists.

Responsibility for developing critical consciousness as an antidote to dysconsciousness lies primarily with teacher educators (Lalvani and Broderick, 2015). Being willing to identify and address our own dysconsciousness through critical reflection about our values, beliefs, assumptions and practices is a first step to teaching future teachers about inequities and their responsibility to respond to them (Broderick and Lalvani, 2017; Lalvani and Broderick, 2015).

Teacher education practices

In my own practice, introducing student teachers to dysconsciousness, critical consciousness, normalcy, ableism, fixed-ability thinking and other core concepts at the start of (and throughout) courses is proving worthwhile. Working through activities in which student teachers are asked to make visible their values and critique their entrenched assumptions is often unsettling work. It requires students to consciously acknowledge tacit knowledge that they have acquired throughout their lives. A common response from student teachers is: "I'd never even thought about this" and/or "I'm noticing things [inequities] in my 'prac' school that I would never have seen before."

Disability Studies in Education (DSE)

Drawing on the framework of Disability Studies in Education (DSE) (Connor and Cavendish, 2020) within teacher education programs is helpful in developing understandings of disability rights. DSE's core tenets include recognizing disability as an inherent aspect of being human; advocating for socially just education for all; presuming every individual is competent; and working with (rather than "on") disabled people—including students (Connor and Cavendish, 2020). In its rejection of "neo-special education" (Slee et al., 2019, p. 5), utilizing a DSE framework can support student teachers to trouble their assumptions about historical and current segregated educational practices, and begin to see "how education might be otherwise" (p. 1).

Disability and intersectionality

The inclusion of disability matters as a compulsory, rather than an elective, minor or "special" component of all teacher education programs is essential to address its longstanding omission from discussion of "minorities" which have been marginalized by dominant cultural power structures (Pugach et al., 2021). It is necessary however to emphasize with student teachers the contextual and intersectional nature of disability. Pugach et al. (2019) call for recognition of the intersectionality of every student's multiple identities. Understanding the concept of intersectionality will also help to disturb and dislodge the false binary of "special" and "normal" and the attendant disparaging discourses that students may experience.

Pedagogies and curriculum

That many teachers do not feel prepared to teach all students is well documented in educational literature (e.g., Black-Hawkins, 2017), however Florian and Linklater (2010) suggest that teachers' apprehension stems from a lack of confidence rather than competence. Within teacher education programs, modeling and teaching the use of inclusive pedagogy (Black-Hawkins, 2017)

and Universal Design for Learning (UDL) (Rose et al., 2014) serve as conceptual and practical tools that contest the assumption that some students are so “special” that they are beyond the expertise and responsibility of “regular” teachers. Rather, these approaches move the focus of learning “difficulties” from the student (= “problem”) to external barriers to learning (e.g., environmental, attitudinal, pedagogical, curricular—barriers to both “doing” and “being”) that teachers have the power to change. Common to these pedagogies is a focus on planning, teaching and assessment with *every-body* in mind from the start. So too is a belief that every child can learn and does so uniquely, providing teachers have created opportunities, choice and the “right” conditions for students to reveal and develop their competence.

In reference to just education, Connell (2012) highlights that, in addition to human rights, “educational responses to deep diversity” (p. 682) are necessary. “Curricular justice” (p. 682) refers to curriculum that is informed by the lives of students who are the least advantaged. The second educational response relates to social encounters within education that are marked by recognition, respect and responsibility. If we consider many of the students’ narratives, even accessing the curriculum, let alone recognizing themselves within it was problematic. Similarly, being seen by teachers and peers is necessary for any social encounter; the dearth of recognition and respectful interactions was all too common and hurtful. Within teacher education then, we need to ensure that student teachers are vigilant and responsive to these elements of just education.

Including strengths-based representation of disability throughout teacher education curricula is one way of bringing it into student teachers’ consciousness (Lalvani, 2015). This can be achieved by routinely incorporating disability-related matters in teaching materials, class activities and readings (wherever possible focusing on students’ experiences of school); utilizing exemplary children’s literature that portrays disability as a natural part of being human and as one aspect of a person’s multiple identities; inviting student teachers to share their experiences; learning from guest speakers who have experience of disability as well as other strategies. Abandoning deficit “special needs-ism” (Rutherford, 2016, p. 127) language is also vital.

Teaching student teachers about human rights is imperative, to ensure that future teachers have the knowledge to model and teach children about their rights and responsibilities as citizens in a democracy. Each year, whenever I ask student teachers to share their knowledge of human rights, there are very few responses. Lalvani (2015) similarly notes student teachers’ lack of knowledge of disability rights. As well, teachers need to recognize and respond appropriately to injustices, their actions grounded in knowledge of relevant international and national rights’ instruments such as UNCRC and UNCRPD. That this is not happening in all schools is evident in a New Zealand study involving beginning secondary school teachers. Three-quarters of participants knew little (or nothing) about UNCRPD; worryingly, 42% believed that such knowledge was *unnecessary* (Attwood et al., 2019, p. 1040).

Practicum matters

Student teachers’ practical experience in schools provides the space in which they can develop their skills in facilitating the social encounters that Connell (2012) claims is an integral educational response to just education. The extent to which such opportunities are available is however somewhat of a lottery. According to Attwood et al. (2019, p. 1043), 70% of their secondary teacher participants had minimal or no engagement with disabled students in their practice. While acknowledging logistical constraints such as the availability of schools and exemplary teachers, this is an area for urgent action. If some students in any class are not “seen,” let alone interacted with, disabled students will continue to feel the emotional impact of being “the last to come to mind” (Fellner et al., 2017, p. 331).

As evident in the accounts of students who were on the receiving end of teacher aides, knowledge of the latter’s role must also be part of teacher education programs. In my own practice, as part of an assignment, student teachers are required to have a conversation with a teacher aide in their school practicum. This is then synthesized with relevant research to develop guidelines for working collaboratively with support staff, for the good of all students. It is encouraging to witness student teachers’ growth in understanding and shift in (dysconscious) assumptions, from “the special needs students are looked after by the teacher aide,” to acting to address inequity, “when I’m teaching, I want to include the student with the teacher aide—it’s not fair that they’re left out.”

These are but some aspects of teacher education that need to be critiqued if we are to foster the growth of teachers who “want to get to know you as a person and you feel you have a place [in school]” (Jo, cited in Goodall, 2020, p. 1298). Providing opportunities for student teachers to feel uncertain about what they have come to know, and to offer alternative ways of knowing that are grounded in students’ experiences of schooling, may move us closer to fairer futures for the latter.

End of chapter thoughts

Documenting the affective impact of schooling on disabled students’ development has been the intent of this chapter. While it is heartening to read of a student’s feeling that “you’re no different as a human being, that you’re an equal among equals” (Vlachou and Papananou, 2015, p. 80), it is extremely disheartening to know that the reality reads so differently for many students, that “you are different from human beings; you’re not equal to others.” Current disability and justice theorizing underscores positive recognition, representation and respect, founded on positive relationships with disabled people, to realize just education. Fracturing the socially constructed dependence (special)—independence (normal) binary through recognition of our interdependence, as equals in relationships, is also crucial (Macmurray, 1964, 2012). As discussed, teacher educators play a critical role in facilitating (or obstructing) transformative change.

Students have afforded us insight to their school worlds. Are these the kinds of worlds we want, both within and beyond education? Is having a disability “the end of the world” (Banks, 2017, p. 102)? Do we want to live in a world in which some individuals

feel that they do not belong with "normal people" who have "nothing wrong with them" (Fellner et al., 2017, p. 335), and feel safer staying with "our own type ... our own people" (p. 333)? Or, as a child in Lalvani's (2015) study declared, "I want to live in a world where all people are treated equally" (p. 12), in which, as her peer claimed, "It is okay to have a disability" (p. 14).

The need for recognition of students' humanity has been writ large throughout the chapter, validating Macmurray's (2012) conceptualization of education as a means of learning to be human. In light of the inhumanity shown to many students by their peers and teachers, I reiterate my belief that education is an apprenticeship in humanity and democracy. As Sean stated (Margrain and Farrugia, 2018, p. 107),

I would make sure that every child gets treated fairly, regardless of their disability. I would not let any teacher make them feel as if they were less than a human being.

Postscript: What about Bill? Despite the shame he felt, he survived school and moved into a world in which he found a place where he could be himself—and thrive.

References

- Andriana, E., Evans, D., 2017. "Why am I chosen as inclusion child?" Listening to students' voice on school experiences of inclusion in Indonesia. In: Plows, V., Whitburn, B. (Eds.), *Inclusive Education: Making Sense of Everyday Practice*. Sense Publishers, Rotterdam, pp. 175–193. https://doi.org/10.1007/978-94-6300-866-2_2.
- Attwood, S., MacArthur, J., Kearney, A., 2019. Beginner secondary teacher preparedness for inclusion. *Int. J. Incl. Educ.* 23 (10), 1032–1048. <https://doi.org/10.1080/13603116.2019.1625455>.
- Banks, J., 2017. "These people are never going to stop labeling me": educational experiences of African American male students labeled with learning disabilities. *Equity Excell. Educ.* 50 (1), 96–107. <https://doi.org/10.1080/10665684.2016.1250235>.
- Biklen, D., Burke, J., 2006. Presuming competence. *Equity Excell. Educ.* 39 (2), 166–175.
- Black-Hawkins, K., 2017. Understanding inclusive pedagogy. In: Plows, V., Whitburn, B. (Eds.), *Inclusive Education: Making Sense of Everyday Practice*. Sense Publishers, Rotterdam, pp. 13–28. https://doi.org/10.1007/978-94-6300-866-2_2.
- Bogdan, R., Taylor, S.J., 1992. The social construction of humanness: relationships with severely disabled people. In: Ferguson, P., Ferguson, D., Taylor, S. (Eds.), *Interpreting Disability: A Qualitative Reader*. Teachers College Press, New York, pp. 275–294.
- Brenna, B., Buries, M., Holtslander, L., Bocking, S., 2017. A school curriculum for Fetal Alcohol Spectrum Disorder: advice from a young adult with FASD. *Int. J. Incl. Educ.* 21 (2), 218–229. <https://doi.org/10.1080/13603116.2016.1193565>.
- Broderick, A., Lalvani, P., 2017. Dysconscious ableism: toward a liberatory praxis in teacher education. *Int. J. Incl. Educ.* 21 (9), 894–905. <https://doi.org/10.1080/13603116.2017.1296034>.
- Christensen, C., Dorn, S., 1997. Competing notions of social justice and contradictions in special education reform. *J. Spec. Educ.* 31 (2), 181–198.
- Connell, R., 2012. Just education. *J. Educ. Policy* 27 (5), 681–683. <https://doi.org/10.1080/02680939.2012.710022>.
- Connor, D.J., Cavendish, W., 2020. "Sit in my seat": perspectives of students with learning disabilities about teacher effectiveness in high school inclusive classrooms. *Int. J. Incl. Educ.* 24 (3), 288–309. <https://doi.org/10.1080/13603116.2018.1459888>.
- Dadvand, B., 2020. Performative pedagogies of care and the emerging geographies of school exclusion for students with disability. *Int. J. Incl. Educ.* 1–15. <https://doi.org/10.1080/13603116.2020.1791981>.
- Fellner, G., Comesañas, M., Duperoy, L., Duperoy, Y., 2017. What it means to be special: two sisters discuss their experiences. *Disabil. Soc.* 32 (3), 323–343. <https://doi.org/10.1080/09687599.2017.1296817>.
- Florian, L., Linklater, H., 2010. Preparing teachers for inclusive education: using inclusive pedagogy to enhance teaching and learning for all. *Camb. J. Educ.* 40 (4), 369–386. <https://doi.org/10.1080/0305764X.2010.526588>.
- Fraser, N., 2008. Abnormal justice. *Crit. Inq.* 34 (3), 393–422. <https://doi.org/10.1086/589478>.
- Giangreco, M.F., Edelman, S.W., Luiselli, T.E., Macfarland, S.Z.C., 1997. Helping or hovering? Effects of instructional assistant proximity on students with disabilities. *Except. Child.* 64 (1), 7–18. <https://doi.org/10.1177/001440299706400101>.
- Goodall, C., 2020. Inclusion is a feeling, not a place: a qualitative study exploring autistic young people's conceptualisations of inclusion. *Int. J. Incl. Educ.* 24 (12), 1285–1310. <https://doi.org/10.1080/13603116.2018.1523475>.
- Jessup, G.M., Bundy, A.C., Hancock, N., Broom, A., 2018. Being noticed for the way you are: social inclusion and high school students with vision impairment. *Br. J. Vis. Impair.* 36 (1), 90–103. <https://doi.org/10.1177/0264619616686396>.
- Jóhannsdóttir, Á., Egilson, S.T., Gibson, B.E., 2020. What's shame got to do with it? The importance of affect in critical disability studies. *Disabil. Soc.* 1–16. <https://doi.org/10.1080/09687599.2020.1751076>.
- King, J.E., 1991. Dysconscious racism: ideology, identity, and the miseducation of teachers. *J. Negro Educ.* 60 (2), 133–146. <https://doi.org/10.2307/2295605>.
- Lalvani, P., 2015. "We are not aliens": exploring the meaning of disability and the nature of belongingness in a fourth grade classroom. *Disabil. Stud. Q.* 35 (4), 1–20. <https://doi.org/10.18061/dsq.v35i4.4963>.
- Lalvani, P., Broderick, A., 2015. Teacher education, exclusion, and the implicit ideology of *Separate but Equal*: an invitation to a dialogue. *Educ. Citizen. Soc. Justice* 10 (2), 168–183. <https://doi.org/10.1177/1746197915583935>.
- Lipsky, D.K., Gartner, A., 1987. Capable of achievement and worthy of respect: education for handicapped students as if they were full-fledged human beings. *Except. Child.* 54 (1), 69–74.
- Lundy, L., 2007. "Voice" is not enough: conceptualising Article 12 of the United Nations Convention on the Rights of the Child. *Br. Educ. Res. J.* 33 (6), 927–942. <https://doi.org/10.1080/01411920701657033>.
- Macmurray, J., 1964. Teachers and pupils. *Educ. Forum* 29 (1), 17–24. <https://doi.org/10.1080/00131726409339304>.
- Macmurray, J., 2012. Learning to be human. *Oxf. Rev. Educ.* 38 (6), 661–674. <https://doi.org/10.1080/03054985.2012.745958>.
- Margrain, V., Farrugia, A., 2018. Different means me: I just learn differently. In: Berman, J., MacArthur, J. (Eds.), *Student Perspectives on School: Informing Inclusive Practice*. Brill Sense, Leiden, pp. 93–110.
- Markham, S., 2020. Questions. Available at: https://www.linkedin.com/pulse/questions-shaun-markham?trk=portfolio_article-card_title. (Accessed 1 March 2021).
- Mueller, C., 2019. Adolescent understandings of disability labels and social stigma in school. *Int. J. Qual. Stud. Educ.* 32 (3), 263–281. <https://doi.org/10.1080/09518398.2019.1576940>.

- Palmer, P., 2017. *The Courage to Teach*, twentieth ed. John Wiley & Sons Inc, NY.
- Pugach, M.C., Gomez-Najarro, J., Matewos, A.M., 2019. A review of identity in research on social justice in teacher education: what role for intersectionality? *J. Teach. Educ.* 70 (3), 206–218. <https://doi.org/10.1177/0022487118760567>.
- Pugach, M.C., Matewos, A.M., Gomez-Najarro, J., 2021. Disability and the meaning of social justice in teacher education research: a precarious guest at the table? *J. Teach. Educ.* 72 (2), 237–250. <https://doi.org/10.1177/0022487120929623>.
- Rose, D.H., Gravel, J.W., Gordon, D.T., 2014. Universal Design for Learning. In: Florian, L. (Ed.), *The Sage Handbook of Special Education*. Sage, London, pp. 475–489.
- Rutherford, G., 2012. In, out or somewhere in between? Disabled students’ and teacher aides’ experiences of school. *Int. J. Incl. Educ.* 16 (8), 757–774.
- Rutherford, G., 2016. Questioning special needs-ism: supporting student teachers in troubling and transforming understandings of human worth. *Teach. Teach. Educ.* 56, 127–137. <https://doi.org/10.1016/j.tate.2016.02.009>.
- Sapon-Shevin, M., 2017. On the impossibility of learning “not to see”: colorblindness, invisibility, and anti-bias education. *Int. Crit. Child. Policy Stud. J.* 6 (1), 38–51.
- Slee, R., 2011. *The Irregular School*. Routledge, NY.
- Slee, R., 2018. *Inclusive Education Isn’t Dead, It Just Smells Funny*. Routledge, NY.
- Slee, R., Corcoran, T., Best, M., 2019. Disability studies in education—building platforms to reclaim disability and recognise disablement. *J. Disabil. Stud. Educ.* 1–11. <https://doi.org/10.1163/25888803-00101002>.
- Stern, J., 2012. The personal world of schooling: John Macmurray and schools as households. *Oxf. Rev. Educ.* 38 (6), 727–745. <https://doi.org/10.1080/03054985.2012.740586>.
- Thomas, C., 2004. How is disability understood? An examination of sociological approaches. *Disabil. Soc.* 19 (6), 569–583. <https://doi.org/10.1080/0968759042000252506>.
- United Nations, 1989. United Nations Convention on the Rights of the Child. Available at. https://treaties.un.org/pages/viewdetails.aspx?src=treaty&mtdsg_no=iv-11&chapter=4&lang=en&title=UNTC-publisher. (Accessed 5 January 2015).
- United Nations, 2006. United Nations Convention on the Rights of Persons With Disabilities. Available at. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html#Fulltext>. (Accessed 1 March 2021).
- United Nations, 2016. General Comment 4: The Right to Inclusive Education. Available at. <https://www.right-to-education.org/resource/general-comment-4-article-24-right-inclusive-education>. (Accessed 1 March 2021).
- Vetoniemi, J., Kärnä, E., 2019. Being included—experiences of social participation of pupils with special education needs in mainstream schools. *Int. J. Incl. Educ.* 1–15. <https://doi.org/10.1080/13603116.2019.1603329>.
- Vlachou, A., Papananou, I., 2015. Disabled students’ narratives about their schooling experiences. *Disabil. Soc.* 30 (1), 73–86. <https://doi.org/10.1080/09687599.2014.982787>.
- Young, D.A., Quibell, R., 2000. Why rights are never enough: rights, intellectual disability and understanding. *Disabil. Soc.* 15 (5), 747–764.

The concept of belonging and other areas of interest in disability studies in education

David J. Connor, Hunter College, City University of New York, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Of policies, placements, and dispositions	238
A sense of belonging	239
Contemplating belonging, cultivating disability studies in education	241
The ongoing crippling of curriculum with DSE	241
The popularity of memoir	242
The neurodiversity movement	242
The negation of racial disparities by the field of special education	243
The emergence of <i>DisCrit</i> and other intersectional approaches	243
Differentiating DSE and critical disability studies	244
Globalization of critical disability studies	244
Conclusion	245
References	246

In this chapter I explore the broad idea of “Belonging,” a concept put forth by scholars within the field of Disability Studies in Education (DSE). First, I examine the relationship of inclusive education to the ideal of creating a space in which all students feel a sense of belonging, arguing its value as a centering principle for practice. Second, I share a sample of other areas within DSE that have generated interest and engagement as the field continues to grow. The list is comprised of diverse topics, largely the subject of my own interests as I observe and participate in cultivating the growth of DSE: (1) the need to continue *cripping* the curriculum; (2) the increased popularity of disability-related memoirs; (3) the rise of the Neurodiversity movement; (4) countering the race-evasiveness of within the field of special education; (5) the development of interdisciplinary theories such as Disability Critical Race Theory (DisCrit); (6) the shift to Critical Disability Studies (CDSs); and, (7) the globalization of CDS. Finally, I discuss ways in which these topics overlap as they relate to the broad notion of belonging writ large, and in particular, within classrooms.

Of policies, placements, and dispositions

While the formalization of inclusive education for children with disabilities has now existed for decades (Lipsky and Gartner, 1997), and has seen many successful incarnations (Danforth, 2014; Hehir and Katzman, 2012), degrees of conceptual confusion have continued to exist. These are, in part, due to purposeful resistance by many scholars in the field of special education who continue to advocate for “appropriate” exclusive settings (Kauffman and Hallahan, 2004). The promise of inclusive education in comparison to its fulfillment to date has left some advocates for inclusive practices disappointed, perplexed, and questioning why we cannot do better in providing a quality education for all students (Slee, 2011; Smith, 2009)?

“*This is Not the Right Setting*” is a phrase I have heard throughout my career in education. It is primarily spoken by teachers (self-identifying as special or general educators), administrators, or school psychologists, and is usually the first step in excluding a child identified as disabled from a particular classroom or school. The comment is followed by a version of “His needs will be better met in a smaller setting”, or “She’ll receive more attention with a lower student to teacher ratio”. I am not naïve enough to think this may be true for some children, but I believe that such arrangements would be preferable short term rather than permanent. However, what is striking is the ease by which children with disabilities can be moved around. Put differently, the changing of classroom settings can be interpreted as a widespread form of passing the buck, namely shifting children who do not fit in a particular environment. What has contributed to changing a student’s placement easily is the wide range of options available. These placement options are usually correlated with the level of services provided, meaning the more intense the services, the more restrictive the setting. How did the special education system in the USA get that way?

Let us take a few steps back. Historically, widespread policies of local educational authorities had always excluded a number of children with disabilities, usually claiming a lack of resources and trained professionals (Fleischer and Zames, 2011). In response to such policies, parents of excluded children with disabilities organized to file class action lawsuits. As a result of their efforts, the landmark legislation of the Education of All Handicapped Children Act of 1975 (EAHCA) was passed to ensure all children with disabilities received the right to a “free and appropriate public education”. However, for many students with disabilities, getting in the door did not equate to getting classrooms with nondisabled peers. In fact, the majority of disabled students were placed in separate locations, creating a dual, segregated system in the USA and other countries (Tomlinson, 1985). As the number

of disability labels grew, so did the types of placement settings, thereby reifying the field of special education as a separate entity from general education (Skrtic, 1991).

For some educators, these organizational structures and cultural practices were ethically problematic and legally questionable, impinging upon the civil rights of students with disabilities (Lipsky and Gartner, 1997). Hence, by the late 1980s, the nascent movement of inclusive education, the idea of students with disabilities being educated alongside their peers had emerged, with encouragement and support from the federal government (Will, 1986). While some educators recognized the intent of inclusive education as creating equity among students (Villa and Thousand, 1995), others resisted the movement, suspiciously viewing it as a cost-saving federal mandate that would ultimately dilute provisions made for students with disabilities (Kauffman and Hallahan, 1995).

Now established as an autonomous discipline, complete with professional organizations, research funding, separate teacher preparation programs undergirded by its distinct ideology, the field of special education undermined the notion of inclusive education and attacked the educators who supported it (Brantlinger, 1997). It is precisely because many leaders in the field of special education not only opposed the inclusion movement (Fuchs and Fuchs, 1994) and sought to silence scholars who dissented from them (Gallagher et al., 2003) that the field of DSE was formed by critical special educators (for a more detailed description of how this occurred, see Connor 2014). From the beginning, part of DSE's mission has always been to support "full and meaningful access to all aspects of society for disabled people" (Connor et al., 2008, p. 448), with a particular focus on all educational settings. Having provided a brief context of policies, placements, and dispositions regarding students with disabilities, in the following section, I explore the relationship between inclusive education and belonging.

A sense of belonging

Despite resistance toward it, the inclusion movement gathered strength during the decades that followed, increasing numbers of children with disabilities were being educated in general education classrooms. With inclusion gaining traction and increasing momentum, the field of special education continued to reassert its investment in retaining segregated settings (Kauffman et al., 2002). Simultaneously, special education reluctantly accepted that inclusive education for more students with disabilities was here to stay. However, to this day, in seeking to retain power and influence, the field of special education continues attempting to colonize inclusive education, even to subsume it within its parameters. This phenomenon can be symbolically seen in many textbooks utilized in teacher education programs that have shifted from titles pertaining only to special education to those such as *Assessment in Special and Inclusive Education* (Salvia et al., 2020). Similarly, the contributions of Marilyn Friend, noted for her work in developing models of collaborative teaching in inclusive classrooms, appear eclipsed by a recent publication titled *Special Education: Contemporary Perspectives for School Professionals (What's New in Special Education)* (2018). Special education's attempted colonization has also occurred in the development, implementation, and research of the federal policy Response to Intervention (RTI), designed to provide support for academically struggling students (Fuchs and Fuchs, 2006). Consequently, general educators, arguably the most important professionals in regard to inclusion, think that inclusion is really the business of the special educators, as special educators (Danforth and Connor, 2020). However, the growth of DSE has helped to provide an alternative narrative about inclusive education that directly challenges ableist beliefs and practices, affirming disability as a natural phenomenon, and providing teachers with ways to create and continue developing inclusive practices, while critiquing ongoing exclusionary ones (Baglieri, 2017; Danforth, 2014; Dudley-Marling and Gurn, 2010; Slee, 2011; Smith, 2009; Valle and Connor, 2019).

In addition to the problematics of how inclusive education has come to be defined and marketed in various ways, other questionable assumptions about what it is continue to exist. For example, some educational organizations view inclusion as primarily a technical or programmatic endeavor, measured in the percentage of time a student spends in their school day in classes with non-disabled peers (National Assessment of Educational Progress, 2010). This approach reduces inclusion to a mathematical formula, dehumanizing both the concept and the experience. For other professionals, there remains the stubborn notion that using modifications and accommodations for students with disabilities is unfair to non-disabled peers (Wiebe Berry, 2008). In other words, traditional "one size" fits all approaches to pedagogy, often employed in Standards-based, exam driven classrooms, are incompatible with inclusive practices. Of course, Universal Design for Learning (UDL) with its premise of creating classrooms to be physically and pedagogically accessible (Novak, 2016), and differentiated instruction, the concept of being flexible in teaching the content, engaging in the process of learning, and developing different products (Tomlinson, 2014), provide valuable ways to address the challenges and complexities of teaching inclusively.

Another problematic area when contemplating how best to prepare inclusive teachers is the continued separation of professionals entering the field of teaching when provided with a forced choice of becoming certified as either a general or special educator. Unfortunately, the content of general education programs does not sufficiently prepare graduates for working with students with disabilities (Forlin and Chambers, 2011), students who are emergent bilinguals (Durgunoglu and Hughes, 2010), or both (Cioè-Peña, 2017). In brief, even though there is a generalized sense among educators as to why inclusive education exists, there are structural, attitudinal, and systemic challenges as to what inclusive education actually is and people's responsibility toward it (Connor, 2012).

Into the mix of contemplating what authentic inclusion is, or can be, is the concept of belonging. My first encounter with its use in relation to inclusive practices was in the form of Arthur Shapiro's book *Everybody Belongs* (2000). Its subtitle, *Changing Negative Attitudes Toward Classmates with Disabilities*, is clearly aimed at educating students without disabilities. For the second edition, the

author teamed with a DSE scholar, giving the title a makeover: *Disability Studies in the Inclusive Classroom*, along with a new subtitle, *Critical Practices for Creating Least Restrictive Attitudes* (Baglieri and Shapiro, 2012). The third version maintains the main title, but the subtitle morphs again to *Critical Practices for Embracing Diversity in Education* (Baglieri, 2017). I trace this trajectory as it evidences how DSE locates the notion of belonging as integral to its core values. While Disability Studies is foregrounded as a means to understand inclusive education, and “belonging” is not explicitly stated in the second and third iterations, the change in wording shifts the focus of power onto the teacher. In brief, educators are asked to understand disability as a natural form of human diversity, with the implicit message of “everybody belongs” remaining intact.

More recently, the topic of Belonging was foregrounded in a special edition of the *International Journal of Inclusive Education*, with a lead article titled, *Belonging in an Age of Exclusion* (Slee, 2019). The author confesses,

Perhaps it is naïve, but I had accepted “belonging” as a conceptual and practical precondition or element of community and inclusion. Ergo, I had assumed that inclusive education embraced a commitment to dismantling exclusions that formed the foundations for the oppression of vulnerable individuals and population cohorts. (p. 910).

In brief, he notes the lack of active commitment by the educational establishment (and beyond) to imagine and create inclusive environments. By all accounts, the revolutionary possibilities of inclusion to disrupt and change the educational ecosystem have been systematically and, at times, severely resisted (Hulgin and Drake, 2010). Efforts to thwart the success of inclusive education have occurred on many levels, some of them already articulated: from within both fields of special education and general education (their continued separation problematic in itself) to sluggish, bureaucratic local educational authorities; from unimaginative teacher preparation programs to the continued justification of exclusionary placements; from the pervasive pathologization of differences in humans to the sifting and sorting of people prepared for more and less valued professions—and by extension—positions in society. All of these issues create cause for concern, captured in a book title by the same first author, *Inclusion Isn't Dead, it Just Smells Funny* (Slee and Tomlinson, 2018). In a nutshell, the current state of inclusion is definitely perceived as “off.”

Even so, as inclusion is an ideal, it can only be approximated if never fully achieved. Knowing and accepting this reality has kept DSE-oriented educators motivated in the pursuit to make our society, starting with its classrooms, inclusive of all people. Bearing this point in mind, let us return to the idea of belonging as “a conceptual and practical precondition or element of community and inclusion” (Slee, 2019, p. 910). In many ways, the concept of belonging can be seen as an anchor—perhaps the anchor—to moor all educators. There are many associations with the word “belonging” in the context of inclusion. For example, in addition to disability (Connor and Berman, 2019), a sense of belonging can be related to racial identity (Hunter et al., 2019), ethnicity (Gonzalez, 2019), language(s) (DeNicolo, 2019), sexual orientation (Hatchel et al., 2018), gender (Lewis et al., 2017) and other markers of identity. Each of these groups have a history of being excluded in various ways from a full public education, with many individuals standing at the interstices of multiply marginalized identities (see, for example, Annamma, 2017; Ferri and Connor, 2009; Ramirez, 2007). In addition, each of these groups has described experiences of microaggressions with teachers, administrators, and other students to remind them of their status, being different from “the norm” (Dudley-Marling and Gurn, 2010). Likewise, they have experienced educational systems, structures, and practices in which they were made to feel uncomfortable, unwelcome, inadequate, and disconnected, questioning if they belonged in schools. A larger question raised is: *To what degree do students drop out or are they pushed out by the very system charged to educate them?* (McNeil et al., 2008).

One of the main reasons why some students feel they do not belong in inclusive classrooms or schools in general is a disconnection between the environment and themselves, prompting feelings of not-fitting in, anxiety, and disenfranchisement (DeMatthews, 2020). These observations speak directly to the purpose of schooling, its historical task of preparing children and youth to participate in, and contribute to, society. On one hand, schools are supposed to help students find their “place” in the world and yet, on the other hand, they have always been tied to forms of social reproduction, stratified along class lines (Collins, 2009) while being porous enough to allow social mobility for some that keeps the American Dream alive (Hochschild and Scovronik, 2003). However, globalization has shifted traditional job opportunities and employment patterns, creating national anxieties about a nation’s standing in relation to others (Tomlinson, 2013). According to the World Population Review, the USA is currently 26th in rankings based upon reading, math, and science scores (2021). Such a lackluster position that has been constant for decades has triggered an intense focus on raising academic levels measured by test scores. This trend can be seen in the form of a Standards Based movement that has dominated Pre-K through 12th grade curricula for decades. Some argue that the Standards zeitgeist is cautiously compatible to inclusive education (Ainscow et al., 2006). Others have argued that new “norms” create new undesirable students, unwelcome in classrooms (Orsati and Causton-Theoharis, 2013). What becomes evident is that a sole or obsessive focus on academic achievement at the expense of all else actively neglects other equally important realms of development—the social and emotional. If we weave all realms together—academic, social, and emotional—and envision classrooms that nurture all students, supporting their growth as human beings rather than mere performers of academic tasks that determine their value in society, then we have a grounding that is simpatico with inclusive education and a sense of belonging.

Unfortunately, the over-emphasis on international competition as the driver of academic agendas is a global phenomenon. For example, Berryman and Bley (2019) invited educators in New Zealand to consider “... what can happen when we focus only on excellence and students reaching their potential, without simultaneously investing in their sense of belonging and wellbeing,” (p. 985) because, “National statistics suggest we are alienating and shortchanging an increasing number of students and, for

disproportionate numbers of indigenous students, these statistics are part of a world trend" (p. 985). The authors note that approximately 1 in 5 students in New Zealand feel "like an outsider," "left out of things," "awkward and out of place," or "lonely" (p. 986). They recognize these descriptions are antithetical to feeling like they belong, where "... people feel accepted, respected, included and supported by a community in which they are involved" (Watson, 2005, p. 988). What has been observed in New Zealand—and around the world—is that the holistic wellbeing of children is suffering when their academic performance is of primary value, the only being counted. Schools have always had the capacity to be dehumanizing places emphasizing perceived deficits in children (Connor, 2009). One purpose of DSE, therefore, is a more humanizing view of human differences, needs, supports, and successes—striving to create classroom spaces in which everybody belongs and ableist attitudes can be actively dismantled through education (Baglieri and Lavani, 2019).

Contemplating belonging, cultivating disability studies in education

While a sense of belonging is both explicitly and implicitly integral to how educators with a DSE disposition envision and create classrooms, I now turn to DSE to share some of its other concerns. Admittedly, these areas have been selected as I find them personally interesting. Indeed, ask another scholar, and they would likely generate a different list, although I would predict significant overlap. Respecting diversity of thinking and a plurality of interests is one of the personally and professionally satisfying things about working in DSE. For the last two decades it has been an expanding field that continues to grow in many directions. I focus on seven areas: (1) the ongoing *cripping* of the curriculum; (2) the popularity of the disability-related memoir; (3) the neurodiversity movement; (4) negation of racial disparities by the field of special education; (5) the emergence of *DisCrit* and other intersectional analyses; (6) differences between DS and Critical Disability Studies (CDSs); and, (7) the globalization of CDS. In the following sections, I touch upon each area, with a caveat that all of them deserve more space and time for deeper exploration.

The ongoing crippling of curriculum with DSE

Crippling the curriculum has always been a priority of DSE (Connor and Bejoian, 2006, 2007) and continues to be so (Baglieri, 2017; Baglieri and Lavani, 2019). DSE scholars believe it is crucial that teachers first recognize the dominant discourse about disability that circulates throughout education, or what Swartz (1992) calls The Master Script. "Master scripting," he explains, "silences multiple voices and perspectives, primarily by legitimizing dominant, White, upper-class, male voicings as the 'standard' knowledge students need to know" (p. 341). Other scripts—including those of inclusive education—are deliberately omitted, or reshaped within the master script for assimilationist purposes. In fact, the more contrary the other script, the greater likelihood of its exclusion or radical morphing, thereby ensuring a continuation of the dominant discourse. Scholars working within DSE and focusing on social models of disability historically experienced being subjected to master scripting by major journals in the field of special education (Gallagher et al., 2003) who still only value medicalized understandings of disability (Anastasiou et al., 2017).

An integral part of DSE is to help teachers recognize and rewrite the master script. Examples of ways this can be done in practice is through the premise of Universal Design for Learning (UDL), that is, creating accessible classrooms for everyone as the starting point (Wilson, 2017), and utilizing differentiated instruction to support all students in learning content, how they process information, and the evidence of learning they produce (Tomlinson, 2001). However, more than the planning and implementation of lessons, it is vital that teachers pro-actively develop ways to teach disability as diversity. Banks' (1993) original outline of the five dimensions of multicultural education are also apt when applied to disability: (1) diversity must be integrated into the content, rather than adopting an additive approach; (2) the origins and construction of knowledge is made transparent, including how various ways of thinking have influenced scholars; (3) a pro-active approach is used to reduce prejudice and helps cultivate positive attitudes to different groups; (4) equitable pedagogy is used that encourages diverse forms of social interactions; and (5) an empowering school culture is consciously developed. These dimensions serve as a guide to transformative approaches of teaching and learning about disability, including integrating the knowledge contributions disabled people.

In regard to pre-school and elementary school, the work of Lavani and Bacon (2018) urge us to rethink the clichéd notion of "We are all special" by engaging in teaching anti-ableism curricula in early childhood classrooms. They advocate for the following four goals, providing multiple examples: (1) each child will demonstrate self-awareness, confidence, family pride, and positive social identities (e.g., "learn key phrases of a song in a different language, including sign language, p. 6); (2) each child will experience comfort and joy with human diversity, have accurate language for human differences, and form deep caring human connections (e.g., "use Sesame Street's new character, Julia, to teach about sensory differences and autism" (p. 9); (3) each child will increasingly recognize unfairness, have language to describe unfairness, and understand that unfairness hurts (e.g., "invite a guest speaker to share their experiences of hurt/unfairness related to a disability" p. 10), and (4) each child will demonstrate empowerment and the skills to act, with others or alone, against prejudice and/or discriminatory actions (e.g., "generate ideas for creating an inclusive playground, gym, or art room" (p. 12). In brief, Bacon and Lavani (2019) seek to challenge dominant master narratives and elevate subjugated knowledges to ensure "the righting of the story of disability in K-12 curricula" (p. 387).

Whether in middle school or high school, students learn about the Civil Rights Movements. Teachers must ask themselves: What groups are included or excluded—African-Americans? Women? The LGBTQI community? People with disabilities? Each group has organized at the grass roots level and forged onward despite considerable social and cultural opposition, to gain political power that ensures greater access to various rights and protections in terms of employment and, for people with disabilities—education, health

care, transportation, and community integration. When focusing on the Disability Rights Movement (a phenomenon that is often unrecognized and rarely included), emphasis can be placed on the activists who organized demonstrations, staged sit-ins, and argued vociferously to speak for themselves and finally be heard. From an activist's perspective, being able to *get on* a public bus was the first symbolic move (Fleischer and Zames, 2011). Students can evaluate the ways that the disability rights movement is similar to, and different from, other important movements, in addition to describing the breadth and limitations of each.

In high school, teachers can assist students to actively analyze most classic texts, from *Of Mice and Men* (Steinbeck, 1937) to *Richard III* (Shakespeare, 2018), *To Kill a Mockingbird* (Lee, 1961) to *A Streetcar Named Desire* (Williams, 1947) to critique and debate the portrayals of characters with disabilities. Likewise, depictions of disability modern classics such as *Girl, Interrupted* (Kaysen, 1994) and *One Flew Over the Cuckoo's Nest* (Kessey, 1964), along with genres such as Southern Gothic and Horror, can be unpacked. (For readers interested in this area we recommend the *Journal of Literary and Cultural Disability Studies* whose sole focus is on depictions of disability in literature). During adolescent years, students' awareness of the complex nature of the world increases, as they create connections among multiple sources of information. They are able to understand the idea of disability as a minority status, and it can be explored with other markers of identity such as race, gender, ethnicity, and sexual orientation. Students can also contemplate: when people with disabilities claim kinship as a minority group, how might that change their self-perception, and how non-disabled people view them? Students can also explore some of the inequalities faced by people with disabilities that, until pointed out to the majority group, usually go "unseen" and therefore unacknowledged.

The popularity of memoir

Memoirs have become ubiquitous in our times, and many disability-related publications in this genre have enjoyed critical acclaim. Most authors being located in the sub-genre of "nobody memoirs," that is, written by non-famous people. As memoir expert Couser (2012) writes, "A good many nobody memoirs are accounts of what it's like to live in, or as, some *particular body*" (p. 20). He explains,

... that has often meant writing about odd bodies, those with anomalous somatic conditions of some sort—illnesses or disabilities. A few conditions—breast cancer, paralysis, HIV/AIDS, deafness, and autism—have produced very large bodies of literature. Numerous other conditions—some seemingly obscure or rare (Munchausen syndrome by proxy)—have been represented in small numbers of narratives. In the case of some conditions (like autism, Down syndrome, and Alzheimer's Disease), the production of personal memoir was hitherto thought to be impossible (p. 20).

In many ways, Couser claims that we live in a Golden Age of disability memoirs that allow us to come close to the perceptions and lived experiences of disabled people. Interestingly, while family members of people with disabilities have also produced more memoirs than in previous times, Couser claims that "... depictions of particular disabilities may affect all those with such a condition. This is why memoirs of disability that are not written by people with the condition in question are especially likely to do harm" (p. 106). This point, I believe, is worth considering when applied to research. It brings up similar issues when depictions of disability are generated by non-disabled people who are removed from the daily lives of disabled people, the impact that those deficit-based understandings have historically had (Harry and Klingner, 2007), and institutional resistance toward conducting research with people *with* disabilities (Barton, 2005).

In contrast to information shared in traditional disability-related research, disability-knowledge attained through engaging with memoirs reveal intensively detailed ways that are rare in research findings. The phenomenon of an identified subgenre of disability memoirs shared by a wide variety of people holds interest for the general public from facial anomalies (Grealy, 2003) to cerebral palsy (Claire, 2017), to mobility impairments (Callahan, 1990) to Deaf-blindness (Girma, 2019), to Asperger's syndrome (Page, 2009) to Learning Disabilities (Mooney, 2019). Many of us working in DSE select information in memoirs to help shape lectures and presentations, require certain memoirs in course syllabi, and provide the option of elective readings and projects. In this way, the actual lived experience of an individual can reveal the complexities of negotiating the physical and social world—including classrooms and education systems in vivid ways that research articles and standard textbooks cannot. As memoirs are deeply personal, they almost invariably resonate with the reader in ways that impact their thinking about important concepts such as what it means to belonging to various communities (including among family, school, peers), the emotional impact of being different from the norm, and ways in which the world is configured to either enable or disable people. Put differently, integrating the humanities into teacher education classes provides a counter-balance to the pervasive deficit-based portrayals of disability.

The neurodiversity movement

Since the 1990s, the term "neurodiversity," coined by Australian sociologist with autism Judy Singer gained traction among people labeled neurologically disabled who preferred to view themselves as a part of natural human diversity. In an *Atlantic* magazine interview with Singer by Harvey Blume (1998), he noted:

Neurodiversity may be every bit as crucial for the human race as biodiversity is for life in general. Who can say what form of wiring will prove best at any given moment? Cybernetics and computer culture, for example, may favor a somewhat autistic cast of mind.

In particular, the label has appealed to individuals identified as having Asperger's Syndrome such as Ari Ne'eman, Co-Founder and President of the Autism Self Advocacy Network and the first autistic presidential appointee on the National Council for Disability. Most importantly, individuals with autism who identify as neurodiverse have challenged well-established professional organizations such as the Child Study Center at New York University who published threatening, ableist portrayals of autism in advertisements designed to look like ransom notes—all with the purpose of raising funds for scientific research and did not have any autistic people in positions of decision-making (Savarese and Savarese, 2010).

Other people identified as having Learning Disabilities (LD) and Attention Deficit Hyperactivity Disorder (ADHD) have also self-identified as part of the neurodiversity movement (Mooney, 2008, 2019), co-creating educational organizations such as Project Eye to Eye (Flink, 2014) in which mentorship occurs between successful professionals with LD/ADHD and children and youth in schools across the country. The acceptance of neurodiversity has occurred more so in mainstream journals than in traditional special education publications, that see such grass-roots reframing as threatening to their beliefs. For example, in his article in *Educational Leadership*, Armstrong (2017) critiques the outdated concepts and inward-looking practices of special education, charging it has "too often remained insular, holding fast to its diagnostic categories" (p. 14), and is stuck in "a disability mindset" (p. 15), while neglecting "strengths based-learning, inclusion, and other positive innovations" (p. 15). Similarly, DS scholars have engaged about neurodiversity in mainstream education publications to show how it contributes to dislodging usually unacknowledged ableist assumptions that reinforce superior notions of "normates" (Lewieki-Wilson et al., 2008).

In addition to individuals with autism, LD, and ADHD, the neurodiversity label has been claimed by people with intellectual disabilities, mental illness, and forms of autism other than Aspergers, who have experienced exclusion even from disability advocacy groups (Hughes, n.d.). Scholars within Mad Studies also recognize "shared concerns" with the neurodiversity community as they seek to make space within the academy (McWade et al., 2014), and the commonalities with neurodiversity and disability in general because, "These movements must all struggle to undermine what they perceive to be as stereotyped representations in science and medicine as well as in popular culture" (Rowland, 2015, p. 2). It is clear that a groundswell of grass roots activism has explored new framings of disability, forged new connections, and have gained ground in various venues from popular magazine articles to academic journals, memoirs to white papers.

The negation of racial disparities by the field of special education

The overrepresentation of students of color in special education has been a pervasive problem since the landmark case *Brown v. Board of Education*, when schools were ordered to become racially integrated in the USA (Ferri and Connor, 2005). Resource centers at Harvard and New York University were established to study and monitor this problem, and to develop approaches to reduce numbers to more equal proportions. Critical special educators published widely in this area, focusing on a variety of angles, including: an "across the board" look at all factors contributing to overrepresentation and how these factors interlock with a view to creating innovative ways that prevent such situations (Losen and Orfield, 2002); recognizing research as a form of cultural practice connected to the issues (Arzubiaga et al., 2008); situating research within historical accounts of exclusion, segregation, and inequity (Artiles, 2009); valuing interdisciplinary methods that complicate our understanding of disproportionality along with new units of analysis such as the "geographies of opportunity" maps created by Artiles et al. (2011); promoting an intersectional approach to understand the lived experiences of students who have been traditionally segregated within school (Erevelles and Minear, 2010) and within the school-to-prison nexus, an insidious phenomenon in which schooling practices resemble those of carceral facilities into which many Black and Brown students are expected to transition as a foregone conclusion (Annamma, 2017); and naming the day-to-day referral process for assessment in school of a potential disability (notice, referral, assessment, diagnosis, placement) as "The molding of (in)competence with ephemeral tainted categories" that erases considerations of race and social, cultural, and linguistic differences among students (Artiles, 2017).

However, over the last few years, the field of special education has embraced the findings of a small group of researchers who claimed the opposite of the historical recognition of overrepresentation, namely that more students of color should be identified as having disabilities, receiving unprecedented publicity in the media (Morgan et al., 2015). Critical special education scholars responded to challenge the many shortcomings of Morgan and colleagues, including a special edition of *Multiple Voices*, a special education journal dedicated to issues of students of color with disabilities, including a DSE perspective (Collins et al., 2016). Meanwhile, Morgan and colleagues continue with their research, "replicating" findings (2017), even though their work flies in the face of half a century of research stating otherwise. Unfortunately, research was used by then Education Secretary Betsey DeVos to suspend federal funding dedicated to research on overrepresentation (Youdin, 2018). DSE scholars and critical special educators continue to challenge the race evasiveness of the field of special education (Connor et al., 2019) because, given what we know about disabilities not being a monolithic experience, and the complicating aspects of race and class on labeling, placement, and disciplinary outcomes, advocating for the increased identification of students of color with disabilities is highly problematic.

The emergence of *DisCrit* and other intersectional approaches

Intersectional approaches to understanding disability have emerged, offering more nuanced understandings of how various markers of identity impact the life experiences of individuals. For example, scholars valued the use of Disability Studies yet recognized it was not sufficiently engaged with issues of race, similarly noting the value of Critical Race Theory, while recognizing it had not integrated disability. To better understand life at the intersections of dis/ability and race, Annamma, Connor, and Ferri synthesized both

theoretical frameworks to create Disability Critical Race Theory (*DisCrit*) (2013). As a dynamic framework that simultaneously engages about disability and race, the seven tenets of *DisCrit* are grounded in the assumption that racism and ableism are mutually constituted and collusive—perpetually circulating across time and context in interconnected ways to uphold normalcy. For example, scholars have utilized *DisCrit* to look at: race, class, and ability (Gillborn et al., 2016); the achievement/opportunity gap (King Thorius and Tan, 2016); overrepresentation (Fergus, 2016); the school-to-prison pipeline (Adams and Erevelles, 2016); school reform (Tomlinson, 2016); and “youth out of place” (Collins, 2016, p. 183).

Interest in using this nascent theoretical framework has grown rapidly and exponentially (Annamma et al., 2018) with scholars crossing disciplinary boundaries and geographic borders (Migliarini, 2017); explicitly questioning the connections between racism and ableism (Artiles, in press); critiquing Positive Behavioral Intervention and Supports (PBISs) (Adams, 2015); teacher education for white teachers (Beneke, 2017) and teachers of color (Kulkarni, 2015); and centering the voices of disabled scholars of color (Hernández-Saca, 2017). Scholars interested in/exclusionary practices have also developed new frameworks to hone theorizing and analysis in their specific areas. For example, Cioè-Peña, 2017 brings together central tenets of *DisCrit* with key components of a raciolinguistic perspective by proposing a Critical Disabilities Raciolinguistic (CDR) perspective to address the particular needs of linguistically-minoritized students of color with disabilities and to counter the White normative gaze. Similarly, Padia and Traxler (2021) use *UndocuCrit* to explore the experiences of students at the intersection of disability and migratory status. In one particular study, for example, they examine “the interplay of fear, schooling, and language use as students pursue college” p. 687.

DisCrit has also been taken up by scholars who see the utility of its framework both for looking inwardly and developing potential actions to ameliorate inequalities within the USA, and others who consider its potential use in the Global South. In an interesting turn, Morgan (2022) argues for *DisCrit* to be applied to American law, bringing the elements of Critical Race Theory back to their original roots within Legislation. She justifies her argument by calling attention to “... the ongoing discrimination, marginalization, and violence against disabled people of color, and the inability of existing federal disability laws, policies, and practices to fully protect against these harms, or adequately remedy these harms.” In terms of the Global South, Sarkar et al. (2022) consider *DisCrit* in its application to India and other countries outside of the Western realm of knowledge development, who continue to struggle in providing education—inclusive or otherwise—to their population of students with disabilities. The authors note the importance of culturally relevant inclusive education as “a theoretical foundation that appreciates culturally held values and local knowledge, while addressing intersecting forms of exclusion and inequity in the Indian context” (p. 82). These examples are two of many in ways *DisCrit* disrupts and ruptures existing fields of study (Annamma et al., 2021), contributing to intersectional and interdisciplinary fields, and claiming its place at the table of social justice issues.

Differentiating DSE and critical disability studies

Over the couple of decades, for many scholars interested in disability, disability studies has evolved into critical disability studies. Meekosha and Shuttleworth (2009) explored whether this change in nomenclature “constitutes a radical paradigm shift, or simply signifies a maturing of the discipline” (p. 47). The authors note how disability studies emerged as part of the civil rights movements and was primarily a critique of society’s response to disability and the material consequences for disabled citizens. While they are mindful of a widespread diversification of social theory and the possibilities that brings, they also urge the field to maintain some of DS’s founding principles, including “(1) the irreducibility of social life to objective facts; (2) the requirement of linking theory with praxis in the struggle for an autonomous and participatory society; (3) the necessity that a discipline or field of study be aware of its own historicity and critically reflect on its conceptual framework; and (4) the need to engage in a dialogue with other cultures on the issues and concepts of current significance” (p. 47).

In his attempts to dis/entangle CDS, Goodley (2013) sees the potential possibilities and risks in this transdisciplinary space, offering insights into “theorizing through materialism, bodies that matter, inter/trans-section-ality, global disability studies, and self and Other” (p. 631). His premise of potential possibilities and risks rests upon the desire to engage with other transformative theoretical arenas such as queer, feminist, and critical race theory, and how CDS reminds us of “the centrality of disability when we consider the politics of life itself” (p. 631) and the search for revolutionary responses. And yet, there is a decided tension in the air. For some, the growth of theoretical possibilities within a transdisciplinary approach to disability have been liberating; for others, it signals a move away from the grass roots grounding concerns of material existence of disabled people in a capitalist society (Barnes, 2012). However, there are, as Goodley notes, ways to balance theoretical expansion with practical concerns, of combining activism and theory that impacts practice. Practical contexts include professions designed to support people with disabilities that have ironically fostered oppressive beliefs and practices. Therefore, a CDS approach to the education of teachers (Connor, in review), social workers (Meekosha and Dowse, 2007), rehabilitation (Magasi, 2008) becomes possible, as it provides the framework to show practices that enable or disable, allowing professionals to enter their fields with a critical eye and a disposition that rejects deficit-understandings of humans, with a view to be part of changes to be made around the world.

Globalization of critical disability studies

Of great interest has been the rise of international critical disability studies in what has been termed “The Global South,” namely nations that have historically been underacknowledged in Western theory and research. Wealthier “Western” nations such as the USA, UK, Australia, and New Zealand have long dominated the field of DS. This phenomenon is problematic for many reasons, including the relevance of DS around the world and how disabilities are perceived and understood within all cultures. The vestiges

of colonialism include a hierarchy of countries defined by wealth, with commonplace terms that reflect standing and self-inscribed importance among nations, such as “first world” and “third world,” “developed” and “developing.” Complicating relationships among countries is the notion of exporting concepts of disability from one global zone to another, raising questions about neo-colonial influences and universal assumptions about disability.

Several DSE and/or CDS scholars have written of the global politics of disability, particularly with global initiatives, asserting education for all being a human right, still very much a challenge for countries with limited resources (Gabel and Danforth, 2008). In particular, Erevelles (2011) has developed insightful critiques of the historical production of disability as it intersects with race, social class, caste, gender, sexual orientation, and citizenship within a neoliberal capital landscape. Along with analyses of the social processes that designate which bodies matter and which do not, Erevelles also articulates hope for more socially just arrangements through using “a materialist ethic of care” (p. 175).

In another example of scholars contemplating the North-South divide, Meekosha (2008) questions the ontological, epistemological, and theoretical assertions within Northern Hemisphere DS, disputing:

claims to universality (what happens in the Global North should happen in the south); a reading from the metropole (a methodological projection of ideas from the centre into the periphery); emphasis on the importance of northern feudal/capitalist modes of production (with an accompanying ignorance and grand erasure of indigenous/ traditional modes of living of the south); a colonialism of psychic, cultural and geographical life of the south by the north; and ignorance of the resistant-subaltern positions of “Global Southerners”

Cited in Goodley, 2013, p. 638.

A recurring theme in writings by CDS scholars who navigate multiple worlds, such as immigrants from South Asia and South America who now reside in the USA is how much ideas fermented in Northern societies are applicable to the Global South in countries wherein the largest percentage of children with disabilities are medical-based conditions, such as disease (Sarkar et al., 2022), cultural ideas of disability differ significantly (Figueroa and Hernández-Saca, 2021), and specific, local issues take precedent. At the same time, scholars such as Rao and Kalyanpur (2015) have used an interdisciplinary perspective to carefully explore disability studies and South Asia, featuring parallel topics to the Global North such as identity construction, historical understandings of disability, cultural representations, and language around disability. More recently, Elder and Migliarini (2020) have focused on decolonizing inclusive education via CDS and *DisCrit* informed practices “show how critical theoretical frameworks can be made accessible to practitioners through strategies that can foster a critical perspective of inclusive education in postcolonial countries” (p. 1852). In doing so, they seek to “push back against the uncritical transfer of inclusion models into Southern countries, which further puts pressure on practitioners to imitate the Northern values of access, acceptance, participation, and academic achievement” (Werning et al., 2006). These examples illustrate ways in which the globalization of CDS has developed as an interactive, productive site of thought and inquiry, that encourage ongoing dialogue.

Conclusion

This chapter is dedicated to the concept of belonging and other areas of interest within DSE. I began by exploring the concept of belonging in relation to inclusive education, recognizing how it is integral to a DSE disposition. I then identified seven areas of interest as examples of topics directly related to DSE. Although these areas can be viewed as disparate, there are connections and overlaps among them, including the broad theme of belonging—writ large. To illustrate some of the connections, I briefly consider the areas of theory, practice, research, and policy.

In terms of theory, Critical Disability Studies supports engagement with any other fields that can potentially offer more nuanced insights into how disability is understood and experienced, and yet it must, as Goodley (2013) notes, strive to balance its growing theoretical confidence with praxis. Similarly, interest in Disability Critical Race Theory has developed in many directions, but the purpose of theory is to determine actions, and scholars are mindful of ways in which a *DisCrit* disposition can be employed from everyday practices to policy reform. Likewise, activists and scholars who identify as neurodiverse have claimed their place by self-theorizing and mobilizing to be in decision-making positions within organizations. If we consider what belonging means in the realm of theory, it is the presence of a distinct theory that is also welcome and utilized within interdisciplinary, social-justice centered work.

In terms of research, we see how a DSE disposition has provided a consistent alternative to limited and medicalized notions of disability, selecting from “political, social, cultural, historical, and individual understandings of disability” (Connor et al., 2008, p. 448). By uniting with critical special educators interested in the cultural contexts within which disability arises, DSE has forged coalitions that continue to challenge dominant paradigms of thought, and seek to redefine how research on disability is conceived of and executed at every step, diversifying research in ways that are culturally relevant, centering people with disabilities as participants and collaborators, rather than subjects (Gonzalez et al., 2017). These issues raise questions about the type of research (and researchers) that inform policy reforms and curriculum choices.

In terms of policy, engagement between the Global North and Global South in both directions has highlighted the importance of considering local customs, beliefs, and practices if a change in policy and practice is to be developed and enacted. Also, the identification of patterns across countries is useful in studying international trends, with the promise of looking at familiar problems

such as overrepresentation of minoritized, indigenous, and immigrant children and youth with a different lens (Cooc and Kiru, 2018). Additionally, the idea of centering student wellbeing within educational policies and practices, as Berryman and Bley (2019) note, is key to acknowledging the importance of a sense of belonging.

In terms of practice, DSE continues to emphasize the moral and ethical dimensions of inclusive education, acknowledging legal and technical aspects must always be negotiated, but without taking precedent. By utilizing DSE in teacher education programs, pre-service and in-service teachers develop an understanding as to why and how to develop and maintain inclusive practices in both the Global North and South. These practices include an ongoing *cripping* of the curriculum to ensure a presence of disability to be discussed and debated, for there is much to both learn and unlearn.

Finally, classrooms are microcosms of society, including its values. As educators, we should aim for creating environments where students feel a sense of belonging, and recognize that their classmates belong. That way, we can stop saying, "This is Not the Right Setting."

References

- Adams, D., 2015. Implementation of School-wide Positive Behavior Supports in the Neoliberal Context in an Urban Elementary School (Dissertation). <https://surface.syr.edu/etd/300>.
- Adams, D., Erevles, N., 2016. Shadow play: DisCrit, dis/respectability, and carceral Logistics. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, pp. 131–144.
- Ainscow, M., Booth, T., Dyson, A., 2006. Inclusion and the standards agenda: negotiating policy pressures in England. *Int. J. Incl. Educ.* 10 (4–5), 295–308.
- Annamma, S., 2017. *The Pedagogy of Pathologization*. Routledge.
- Annamma, S., Connor, D.J., Ferri, B.A., 2013. Dis/ability critical race studies (DisCrit): theorizing at the intersections of race and disability. *J. Race Ethn. Educ.* 16 (1), 1–31.
- Annamma, S.A., Ferri, B.A., Connor, D.J., 2018. Disability critical race theory: exploring the intersectional lineage, emergence, and potential futures of DisCrit in education. *Rev. Educ. Res.* 42, 46–71.
- Annamma, S., Ferri, B., & Connor, D. J. (Eds.), (2021) *DisCrit Expanded: Inquiries, Reverberations & Ruptures*. Teachers College Press.
- Ansatsidou, D., Kauffman, J., Maag, J.W., 2017. Special education at the crossroad: an identity crisis and the need for scientific reconstruction. *Exceptionality* 25 (2), 139–155.
- Armstrong, T., 2017. Neurodiversity: the future of special education? *Educ. Leader* 74 (7), 10–16.
- Artiles, A.J., October 19, 2017. Re-envisioning Equity Research: Disability Identification Disparities as a Case in Point [14th Annual Brown Lecture]. American Educational Research Association, Washington, DC.
- Artiles, A.J., Kozleski, E.B., Waitoller, F., Lukinbeal, C., 2011. Inclusive education and the interlocking of ability and race in the US: notes for an educational equity research program. In: Artiles, A., Kozleski, E., Waitoller, F. (Eds.), *Equity in Inclusive Education in Four Continents: A Cultural Historical Multilevel Model*. Harvard Education Press, pp. 45–68.
- Artiles, A. (in press). The future(s) of disability of complementary representations, heteroglossic communities and moral leadership. In Annamma, S., Ferri, B., & Connor, D. J. (Eds.) *DisCrit Expanded: Inquiries, Reverberations & Ruptures*. Teachers College Press.
- Arzubagi, A.E., Artiles, A.J., King, K., Harris-Murri, N., 2008. Beyond research on cultural minorities: challenges and implications of research as situated cultural practice. *Except. Child.* 74 (3), 309–327.
- Bacon, J.K., Lalvani, P., 2019. Dominant narratives, subjugated knowledges, and the righting of the story in K-12 curricula. *Curric. Inq.* 49 (4), 387–404.
- Baglieri, S., 2017. *Disability Studies and the Inclusive Classroom: Critical Practices for Embracing Diversity in the Classroom*. Routledge.
- Baglieri, S., Lalvani, P., 2019. *Undoing Ableism: Teaching About Disability in K-12 Classrooms*. Routledge.
- Baglieri, S., Shapiro, A., 2012. *Disability Studies and the Inclusive Classroom: Critical Practices for Creating Least Restrictive Attitudes*. Routledge.
- Banks, J.A., 1993. Multicultural education: development, dimensions, and challenges. *Phi Delta Kappan* 75 (1), 22–28.
- Barnes, C., 2012. Understanding the social model of disability: past, present, and future. In: Watson, N., Roulstone, A., Thomas, C. (Eds.), *The Routledge Handbook of Disability Studies*. Routledge, pp. 12–29.
- Barton, L., 2005. Emancipatory research and disabled people: some observations and questions. *Educ. Rev.* 57 (3), 317–327.
- Beneke, M.R., 2017. *Race and Ability Talk in Early Childhood: Critical Inquiry into Shared Book Reading Practices with Pre-service Teachers* (Doctoral dissertation). ProQuest, LLC.
- Berryman, M., Bley, E., 2019. Students belonging: critical relationships and responsibilities. *Int. J. Incl. Educ.* 23 (9), 985–10001.
- Blume, H., September 1998. Neurodiversity. *The Atlantic Magazine*. <https://www.theatlantic.com/magazine/archive/1998/09/neurodiversity/305909/>.
- Callahan, J., 1990. *Don't Worry, he Won't Get Far on Foot*. Vintage Books.
- Cioe-Pena, M., 2017. The intersectional gap: how bilingual students in the United States are excluded from inclusion. *Int. J. Incl. Educ.* 21 (9), 906–919.
- Claire, E., 2017. *Brilliant Imperfection: Grappling with Cure*. Duke University Press.
- Collins, J., 2009. Social reproduction in classrooms and schools. *Annu. Rev. Anthropol.* 38, 33–48.
- Collins, K., 2016. A DisCrit perspective on the State of Florida v. George Zimmerman: racism, ableism, and youth out of place in community and school. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, pp. 35–54.
- Collins, K., Connor, D.J., Ferri, B.A., Gallagher, D., Samson, J., 2016. Dangerous assumptions and unspoken limitations: a disability studies in education response to morgan, Farkas, hillemeir, mattison, Maczuga, Li, and Cook (2015). *Multiple Voices* 16 (1), 4–16.
- Connor, D.J., 2009. Breaking containment: the power of narrative knowing—countering silences within traditional special education research. *Int. J. Incl. Educ.* 13 (5), 1–21.
- Connor, D.J., 2012. Common confusions with inclusion. In: Cooper, B., Strax, C., Strax, M. (Eds.), *The Politics of Special Education: Problems, Promises, and Progress*. Roman and Littlefield, pp. 101–122.
- Connor, D.J., 2014. The disability studies in education annual conference: explorations of working within, and against, special education. *Disabil. Stud. Q.* 34 (2).
- Connor, D.J., Bejoian, L., 2006. Pigs, pirates, and pills: using film to teach the social context of disability. *Teach. Except. Child.* 39 (2), 52–60.
- Connor, D.J., Bejoian, L., 2007. Crippling school curricula: 20 ways to re-teach disability. *Rev. Disability Studies* 3 (3), 3–13.
- Connor, D.J., Berman, D., 2019. (Be)longing: a family's desire for authentic inclusion. *Int. J. Incl. Educ.* 23 (9), 923–936.
- Connor, D.J., Gabel, S., Gallagher, D., Morton, M., 2008. Disability studies and inclusive education—implications for theory, research, and practice. *Int. J. Incl. Educ.* 12 (5–6), 441–457.
- Connor, D., Cavendish, W., Gonzalez, T., Jean-Pierre, P., 2019. Is a bridge even possible over troubled waters? The field of special education negates the overrepresentation of minority students: a DisCrit analysis. *Journal of Race, Ethnicity & Education* 22 (6), 723–745.
- Connor, D. J. (in review). Revamping a graduate course to (in)fuse disability studies: the politics of representation in "The Study of Learning Disabilities in Children and Adolescents." *J. Teach. Disabil. Stud.*
- Couser, G.T., 2012. *Memoir: An Introduction*. Oxford University Press.
- Danforth, S. (Ed.), 2014. *Becoming a Great Inclusive Educator*. Peter Lang.

- Danforth, S., Connor, D.J., 2020. Considering inclusive education through the independent living movement: lessons for improvement? *Educ. Forum* 84 (4), 325–339. <https://doi.org/10.1080/00131725.2020.1798705>.
- DeMatthews, D., 2020. Addressing racism and ableism in schools: a DisCrit leadership framework for principals. *Clear. House A J. Educ. Strategies, Issues Ideas* 93 (1), 27–34.
- DiNicolò, C., 2019. The role of translanguaging in establishing school belonging for emergent multilinguals. *Int. J. Incl. Educ.* 23 (9), 967–984.
- Dudley-Marling, C., Gurn, A. (Eds.), 2010. *The Myth of the Normal Curve*. Peter Lang.
- Durgunoglu, A.Y., Hughes, T., 2010. How prepared are the U.S. preservice teachers to teach English language learners? *Int. J. Teach. Learn. High. Edu.* 22 (1), 32–41.
- Elder, B., Migliarini, V., 2020. Decolonizing inclusive education: a collection of practical- inclusive CDS and DisCrit-informed teaching practices implemented in the global South. *Disabil. Global South* 7 (1), 1852–1872.
- Erevelles, N., 2011. *Disability And Difference in Global Contexts: Enabling a Transformative Body Politic*. Palgrave MacMillan.
- Erevelles, N., Minear, A., 2010. Unspeakable offenses: untangling race and disability in discourses of intersectionality. *J. Lit. Cult. Disabil. Stud.* 4 (2), 127–145.
- Fergus, E., 2016. Social reproduction ideologies: teacher beliefs about race and culture. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, pp. 117–127.
- Ferri, B.A., Connor, D.J., 2005. Tools of exclusion: race, disability, and (re)segregated education. *Teach. Coll. Rec.* 107 (3), 453–474.
- Ferri, B.A., Connor, D.J., 2009. "I was the special ed. girl." (En)gendering disability from the standpoint of urban working class young women of color. *J. Gender Edu.* 22 (1), 1–17.
- Figueroa, C., Hernandez-Saca, D., 2021. *Dis/ability in Latin America: The Intersections of Education, Power and Identity*. Lexington.
- Fleischer, D.Z., Zames, F., 2011. *The Disability Rights Movement: From Charity to Confrontation*, second ed. Temple University Press.
- Flink, D., 2014. *Thinking Differently: An Inspiring Guide for Parents of Children with Learning Disabilities*. William Morrow.
- Forlin, C., Chambers, D., 2011. Teacher preparation for inclusive education: increasing knowledge but raising concerns. *Asia Pac. J. Teach. Educ.* 39 (1), 17–32.
- Friend, M., 2018. *Special Education: Contemporary Perspectives for School Professionals*.
- Fuchs, D., Fuchs, L.S., 1994. Inclusive schools movements and the radicalization of special education reform. *Except. Child.* 60 (4), 294–309.
- Fuchs, D., Fuchs, L.S., 2006. Introduction to response to intervention: what, why, and how valid is it? *Read. Res. Q.* 41 (1), 93–99.
- Gabel, S.L., Danforth, S. (Eds.), 2008. *Disability and the Politics of Education: An International Reader*. Peter Lang.
- Gallagher, D.J., Heshusius, L., Iano, P., Skrtic, T.M., 2003. *Challenging Orthodoxy In Special Education*. Love Publishing.
- Gilborn, D., Rollock, N., Vincent, C., Ball, S.J., 2016. The black middle classes, education, racism, and dis/ability: an intersectional analysis. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, pp. 35–54.
- Girma, H., 2019. Haben: The Deafblind Woman Who Conquered Harvard Law. Twelve.
- Gonzalez, S.M., 2019. Cultivating familismo: belonging and inclusion in one Latina/o learning community. *Int. J. Incl. Educ.* 23 (9), 937–949.
- Gonzalez, T., Saca-Hernandez, D., Artiles, A., 2017. In search of voice: theory and methods in student voice research, 1990–2010. *Educ. Rev.* 69 (4), 451–473.
- Goodley, D., 2013. Dis/entangling critical disability studies. *Disabil. Soc.* 28 (5), 631–644.
- Grealy, L., 2003. *Autobiography of a Face*. Mariner.
- Harry, J., Klingner, J.K., 2007. Discarding the deficit model. *Educ. Leader* 64 (5), 16–21.
- Hatchel, T., Espelage, D.L., Huang, Y., 2018. Sexual harassment victimization, school belonging, and depressive symptoms among LGBTQ adolescents: temporal insights. *Am. J. Orthopsychiatry* 88 (4), 422–430.
- Hehir, T., Katzman, L., 2012. *Effective Inclusive Schools: Designing Successful Schoolwide Programs*. Jossey-Bass.
- Hernández-Saca, D.I., 2017. Reframing the master narratives of dis/ability at my intersections: an outline of an educational equity research agenda. In: *Critical Disability Discourses/ Discours critiques dans le champ du handicap*, vol. 8.
- Hochschild, J.L., Scovronik, N., 2003. *The American Dream and the Public Schools*. Oxford University Press.
- n.d. Hughes, J. M. F. (n.d.). Nothing about us without us: Increasing neurodiversity in disability and social justice action groups. [White Paper] <https://autisticadvocacy.org/wp-content/uploads/2016/06/whitepaper-Increasing-Neurodiversity-in-Disability-and-Social-Justice-Advocacy-Groups.pdf>.
- Hulgin, K.M., Drake, B.M., 2010. Inclusive education and the No child left behind act. *Int. J. Incl. Educ.* 15 (4), 389–404.
- Hunter, C.D., Case, A.D., Harvey, S., 2019. Black students' sense of belonging and racial identity. *Int. J. Incl. Educ.* 23 (9), 950–966.
- Kauffman, J.M., Hallahan, D.P. (Eds.), 1995. *The Illusion of Full Inclusion*. ProEd, Austin, TX.
- Kauffman, J.K., Hallahan, D., 2004. *Special Education: What it is and why we Need it*. Pearson.
- Kauffman, J.K., Bantz, J., McCullough, J., 2002. Separate and better: a special public school class for students with emotional and behavioral disorders. *Exceptionality* 10 (3), 149–170.
- Kaysen, S., 1994. *Girl Interrupted*. Vintage.
- Kessey, K., 1964/2003. *One Flew over the Cuckoo's Nest*. Penguin.
- King Thorius, K.A., Tan, P., 2016. Expanding analysis of educational debt: considering intersections of race and ability. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, 97–54.
- Kulkarni, S.S., 2015. *Beliefs about Disability, Race, and Culture of Urban Special Education Teachers and Their Retention Decisions* (Doctoral dissertation).
- Lalvani, P., Bacon, J., 2018. Rethinking "we are all special": anti-ableism curricula in early childhood classrooms. *Young Except. Child.* 22 (2), 87–100.
- Lee, H., 1961. *To Kill a Mockingbird*. Grand Central.
- Lewicki-Wilson, C., Dolmage, J., Hielker, P., Jurecic, A., 2008. Two comments on "neurodiversity". *Coll. Engl.* 70 (3), 314–325.
- Lewis, K.L., Stout, J.G., Finkelstein, N.D., 2017. Fitting in to move forward: belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychol. Women Q.* 41 (4), 420–436.
- Lipsky, D.K., Gartner, A., 1997. *Inclusion and School Reform: Transforming America's Classrooms*. Paul H. Brookes.
- Losen, D., Orfield, G., 2002. *Racial Inequity in Special Education*. Harvard University Press, Cambridge, MA.
- Magasi, S., 2008. Infusing disability studies into the rehabilitation sciences. *Top. Stroke Rehabil.* 15 (3), 283–287.
- McNeil, L.S., Coppola, E., Radigan, E., Heilig, J.V., 2008. Avoidable losses: high-stakes accountability and the dropout crisis. *Educ. Pol. Anal. Arch.* 16 (3), 1–48.
- McWade, B.M., Milton, D., Beresford, P., 2014. Mad studies and neurodiversity: a dialogue. *Disabil. Soc.* 30 (2), 305–309.
- Meekosha, H., 2008. Contextualizing disability: developing southern theory. In: Keynote Presentation to the Disabilities Studies Association 4th Conference, September 2–4, Lancaster, UK.
- Meekosha, H., Dowe, L., 2007. Integrating critical disability studies into social work education and practice: an Australian perspective. *Practice* 19 (3), 169–183.
- Meekosha, H., Shuttleworth, R., 2009. What's so 'critical' about critical disability studies? *Aust. J. Human Rights* 15 (1), 47–75.
- Migliarini, V., 2017. Subjectivation, agency and the schooling of raced and dis/abled asylum-seeking children in the Italian context. *Intercult. Educ.* 28 (2), 182–195.
- Mooney, J., 2008. *Short Bus: A Journey beyond Normal*. McMillan.
- Mooney, J., 2019. *Normal Sucks. How to Live, Learn, and Thrive outside the Lines*. Henry Holt.
- Morgan, P.L., Farkas, G., Hillemeier, M.M., Maczuga, S., Li, H., Cook, M., 2015. Are minority children disproportionately represented in early intervention and early childhood special education? *Educ. Res.* 41 (9), 339–351.
- Morgan, P.L., Farkas, G., Hillemeier, M.M., Maczuga, S., 2017. Replicated evidence of racial and ethnic disparities in disability identification in U.S. schools. *Educ. Res.* 46 (6), 305–322.
- Morgan, J. (in press). Toward a DisCrit approach to American law. In Annamma, S., Ferri, B., & Connor, D. J. (Eds.) *DisCrit Expanded: Inquiries, Reverberations & Ruptures*. Teachers College Press.

- National Assessment of Educational Progress, 2010. <https://nces.ed.gov/nationsreportcard/studies/inclusion/>.
- Novak, K., 2016. UDL Now! A Teacher's Guide to Applying Universal Design for Learning in Today's Classrooms. CAST Professional Publishing.
- Orsati, F.T., Causton-Theoharis, J., 2013. Challenging control: inclusive teachers' and teaching assistants' discourse on students with challenging behaviour. *Int. J. Incl. Educ.* 17 (5), 507–525.
- Padia, L. B., Traxler, R. E. 2021. Traerás tus Documentos (you will bring your documents): navigating the intersections of disability and citizenship status in special education. *Race Ethn. Educ.* 24 (5), 687–702.
- Page, T., 2009. Parallel Play: Growing up with Undiagnosed Asperger's. Doubleday.
- Ramirez, R., 2007. Race, tribal nation, and gender: a native feminist approach to belonging. *Meridians* 7 (2), 22–40.
- Rao, S., Kalyanpur, M., 2015. South Asia and Disability Studies: Redefining Boundaries and Extending Horizons. Peter Lang.
- Rowland, M., 2015. Angry and mad: a critical examination of identity politics, neurodiversity, and the mad pride movement. *J. Ethic. Mental Health* 9, 1–3. https://jemh.ca/issues/v9/documents/JEMH_Open-Volume_Article_Angry_and_Mad-Oct2015.pdf.
- Salvia, J., Ysseldyke, J.E., Witmer, S., 2020. Assessment in Special and Inclusive Education. Cengage.
- Sarkar, T., Muller, C., Forber Pratt, A., 2022. Does DisCrit travel?: global South and excess theoretical baggage fees. In: Annamma, S., Ferri, B., Connor, D.J. (Eds.), *DisCrit Expanded: Inquiries, Reverberations & Ruptures*. Teachers College Press inpress, pp. 81–95.
- Savarese, E.T., Savarese, R.J., 2010. "The superior half speaking": an introduction. *Disabil. Stud. Q.* 30 (1). <https://dsq-sds.org/article/view/1062/1230>.
- Shakespeare, W., 2018. *Richard III*. Simon and Schuster.
- Shapiro, A., 2000. *Everybody Belongs: Changing Negative Attitudes towards Classmates with Disabilities*. Routledge.
- Skrlic, T.M., 1991. *Behind Special Education: A Critical Analysis of Professional Culture and School Organization*. Love.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Routledge.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 29 (3), 909–922.
- Slee, R., Tomlinson, S., 2018. *Inclusion Isn't Dead, it Just Smells Funny*. Routledge.
- Smith, P., 2009. *Whatever Happened to Inclusion? The Place of Students with Intellectual Disabilities in Education*. Peter Lang.
- Steinbeck, J., 1937/1993. *Of Mice and Men*. Penguin.
- Swartz, E., 1992. Emancipatory narratives: rewriting the master script in the school curriculum. *J. Negro Educ.* 61 (3), 341–355.
- Tomlinson, S., 1985. The expansion of special education. *Oxf. Rev. Educ.* 11 (2), 157–165.
- Tomlinson, C.A., 2001. *How to Differentiate Instruction in Mixed Ability Classrooms*, second ed. ACSD.
- Tomlinson, S., 2013. *Ignorant Yobs?: Low Attainers in a Global Market Economy*. Routledge.
- Tomlinson, C.A., 2014. *The Differentiated Classroom: Responding to the Needs of All Learners*, second ed. ACSD, Alexandria, VA.
- Tomlinson, S., 2016. Race, class, ability, and school reform. In: Connor, D.J., Ferri, B.A., Annamma, S.A. (Eds.), *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, pp. 157–166.
- Valle, J., Connor, D.J., 2019. *Rethinking Disability: A Disability Studies Guide to Inclusive Practices*, second ed. Routledge, New York.
- Villa, J., Thousand, R., 1995. *Creating an Inclusive School*. ASCD.
- Watson, C., 2005. Classrooms as learning communities: a review of research. *Lond. Rev. Educ.* 3 (1), 47–64.
- Werning, R., Artilles, A., et al., 2006. *Keeping the Promise? Contextualizing Inclusive Education in Developing Countries*. Verlag Julius Klinkhardt, Bad Heilbrunn.
- Wiebe Berry, R.A., 2008. Novice teachers' conceptions of fairness in inclusion classrooms. *Teach. Teach. Educ.* 24 (5), 1149–1159.
- Will, M.C., 1986. Educating children with learning problems: a shared responsibility. *Except. Child.* 52, 411–415.
- Williams, T., 1947. *A Streetcar Named Desire*. New Directions.
- Wilson, J.D., 2017. Reimagining disability and inclusive education through universal design for learning. *Disabil. Stud. Q.* 37 (2). <http://dsq-sds.org/article/view/5417/4650>.
- World Population Review (2021). <https://worldpopulationreview.com/country-rankings/education-rankings-by-country>.
- Youdin, M., 2018. DeVos wants to roll back protections for students of color in special education. *Talk Poverty*. <https://talkpoverty.org/2018/02/27/devos-wants-roll-back-protections-students-color-special-education/>.

Disability studies in education and intersectionality

Elizabeth Walton^{a,b}, ^a School of Education, University of Nottingham, Nottingham, United Kingdom; and ^b School of Education, University of Witwatersrand, Johannesburg, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	249
Intersectionality	249
A framework for discerning injustice	250
Disability and race	250
Disability and class/socio-economic status	251
Disability and gender	251
Disability and migrant/refugee status	252
An intersectional case-study: South African education	252
Leveraging intersectionality: necessary concepts	254
Conclusion: the promise of intersectionality in and for disability studies in education	256
References	256

Introduction

SHE is in her late teens, but can't read or write. This is because she cannot find a school for children with special needs like her.

"She" is Mathabo Mosanteli and her story is told in a South African newspaper (Macupe, 2015). It is a simplistic story in which individual deficit (illiteracy) is attributed to inadequate special school provision. The focus is Mathabo's "special needs" and the report justifiably draws attention to the hundreds of thousands of disabled children and young people who are out of school in that country (Human Rights Watch, 2015). It rehearses many of the tropes of similar news reports, including that disabled students are deficient, pitiable and helpless, that separate special schools are assumed as the appropriate placement for them, and that the government is not doing its job (Walton and McKenzie, 2020). But the simplicity of the story is pernicious. It is a single-factor analysis which fails to recognize the complexities of disability and its intersections with other identities. Mathabo's problem, according to the story, is that she has a learning disability. But she is also a young woman in a country with very high rates of gender-based violence, particularly against disabled women (Chirwa et al., 2020; Graaff and Heineken, 2017). She is a Black African, living with the legacies of racial segregation and oppression under previous colonial and apartheid governments. She is poor in a country of extreme income inequality, which will negatively affect her access to health and other social services (Sulla and Zikhali, 2018). Her experience is certainly shaped by systemic disablism that results in inadequate appropriate schooling. But it is compounded by sexism, racism and unequal economic power. Without recognizing these factors, the story fails to capture the nuances of educational exclusion in the South African context.

The unequal power relations wrought by ableism and disablism are necessary, but often insufficient explanations for the exclusion that occurs in and through education. A failure to acknowledge the intersectional impact of other oppressed identities in analysis, policy and practice risks reproducing inequality. The concept of intersectionality furthers an understanding of the complex interplay of people's identities, lives and experiences in particular contexts and can be harnessed in the quest for educational justice. This article explores intersectionality with reference to disability studies in education in three moves. First, I define intersectionality as a conceptual and analytical resource and show what it reveals about the imbrications of disability and race, class/socio-economic status, gender and migrant/refugee status. I then return to Mathabo's context and offer a case-study of South Africa that illustrates the interlocking workings of power in the education system. The third move acknowledges that intersectionality is a traveling theory which can be depoliticized (Davis, 2020). This section proposes considerations to ensure that intersectionality retains its critical and emancipatory contribution to disability studies in education. The article concludes with the promise of intersectionality as a necessary condition for the realization of socially just, inclusive and democratic education.

Intersectionality

Intersectionality is a conceptual and analytical resource that furthers an understanding of the unequal social relations that produce the educational exclusion of disabled children and young people. It recognizes their intersecting identities and the complex ways in which their experiences are shaped by the interstices of power that give rise to multiple axes of oppression. The term "intersectionality" was coined by Kimberlé Crenshaw (Crenshaw, 1991). She sought to describe the compounding disadvantage experienced by people living with multiple oppressed identities and the potential for social movements organized around a particular identity (like

feminism) to ignore the additional effects of other marginalized identities (like race or class). Often spoken of as a theory, intersectionality can be regarded as “a method and a disposition, a heuristic and analytic tool” and it is continually being developed (Carbado et al., 2013, p. 303). Romero (2018) uses a Rubik cube metaphor to describe intersectionality. The rotating cube visualizes the “multiple layers of domination and the intersection of systems of oppression” (ibid. p. 11). She sees that intersectionality allows an analysis of social identities in terms of “power, privilege and opportunity structures” (ibid.). Tikly (2020, p. 37) recognizes “intersecting regimes of inequality” including class, race, ethnicity and gender that co-evolve and are mutually constituting. Importantly, intersectionality is not the recognition of the co-existence of identity markers, but a consideration of “how each supports or unsettles the constitution of one another” (Goodley, 2013, p. 636).

Disability is often absent when scholars focus on intersectional inequality (see, for example, Tikly (2020) and Fruja Amthor (2017)). While not all intersectional scholars consider disability, disability studies (in education) acknowledges intersectionality as an important concern (Slee et al., 2019; Broderick et al., 2006). Ferri (2006, p. 304) offers the challenge and the possibilities of intersectionality in disability studies in education, saying that it requires us:

To give up certainties but in the process we gain complexity, we give up having easy answers but get better at asking hard questions, we ask for trouble in the classroom and in our own thinking but in so doing we end up with a pedagogical model that is both more than and different from any sum of identity-based parts.

This article attests to the complexity inherent in intersectional thinking and analysis, and the lack of easy answers.

A framework for discerning injustice

Ableism and disablism work together with other axes of oppression in complex ways to produce unequal social relations. These in turn set up a series of educational arrangements that differentially benefit students and produce exclusion. In the section that follows, I present four conceptual case studies on the intersections of disability with race, class/socio-economic status, gender, and refugee/migrant status. These illustrate “the relations and processes that create inequality or oppression” (Frobisher et al., 2019, p. 414). These relations and processes are located in the school and classroom and beyond, and must be seen in the context of wider societal inequities, including housing, employment, and access to health services (Harry, 2014). Students, as Artiles (2004, p. 552) notes, are “suspended in a web of historical, ideological, and cultural forces”. Their own cultures, histories and geographies are in play, as are the complex ways in which political, economic and educational structures advantage and disadvantage them and their families.

Disability and race

The overrepresentation of minority groups in segregated special settings has long been recognized in education where ethnic minority and marginalized groups have been placed in special education in numbers disproportionate to their presence in the student population as a whole (Harry, 2014; Woods et al., 2020). In the USA, racial and ethnic disparities in special education placements have led to amendments to the Individuals with Disabilities in Education Act (Woods et al., 2020). Other countries, including the UK, Germany, and Sweden also have concerns about disproportionality and low educational achievement of minority ethnic groups (Harry, 2014; Tomlinson, 2017). Accounting for this disproportionality requires considerations at the confluence of systemic and institutional discrimination, racism and bias, the instability of disability categorization, and configurations of educational provision.

Differential assignment of disability categories and labels reflects both racial biases and the instability of some categories, especially those relying on clinical judgment. Racial biases may be found in the presumption of white middle-classness as the “norm” against which behaviors and achievements are measured (O’Connor and Fernandez, 2006). The history of the category “Learning Disabilities” as a preferred diagnostic category for white middle-class students has been well documented by Reid and Valle (2004). Harry (2014) extends this discussion by noting the disproportionate incidence of white students assigned to the preferred “Severe Learning Disabilities (SLD)” category, and of black students in the stigmatized Intellectual Disability category. White students may be more likely to have an ADHD diagnosis in the USA, with teachers believing that white children who misbehave must be “sick” whereas black children who misbehave are merely “naughty” (Tait, 2010, p. 22). In the category of Emotional Disturbance, Harry (2014, p. 84) notes that “professional judgments will vary according to personal tolerance of non-compliance” with cultural conditioning playing a role in interpreting behaviors. The same behaviors may be evaluated as normal for white students, but as deviant when displayed by black students (O’Connor and Fernandez, 2006).

The way education provision is configured in different contexts impacts measures of disproportionality. This includes structures (e.g., general classrooms, self-contained classrooms, special schools), quality of provision, and resources (what supports get funded and how). In turn, these interact with systems of high stakes testing and accountability regimes, where schools are pitted against each other in competitive league tables. Poor quality instruction must be a consideration in understanding disproportionality, with Harry (2014) citing research that shows that students living in poverty, and those who are English language learners receive inferior schooling in the USA. Then, the key point is when “low achievement comes to be interpreted as ‘disability’” with teacher referrals for evaluation leading to special education placement (ibid. p. 75). Across the world, measures are taken by schools to

minimize the impact of low scorers on tests, including securing mitigating “special educational needs” labels or other exclusionary measures (Tomlinson, 2017; Riddle, 2018; Slee, 2018).

The scholarship on disproportionality tends to focus on countries of the Global North, with a consideration of ethnic and racial minority students relative to the majority white population. The Global South is often marginalized or ignored (Grech, 2015). Disproportionality is evident in the global incidence of disability and outcomes for disabled students. UNESCO (2018a, p. 4) notes that, “Around four in five children with disabilities are in developing countries. At all ages, levels of both moderate and severe disability are higher in low- and middle-income countries than in rich countries”. The legacies of racist colonialism and underdevelopment have to be acknowledged in the manufacture of this state of affairs. The violence of colonialism has contributed to disability incidence in the Global South and the “management” of disability through colonial institutions has caused untold suffering (Grech, 2015; Walton and McKenzie, 2020). Nor can this be assumed to be an unfortunate but closed chapter in human history. The (white) West cannot avert its eyes from its complicity in ongoing economic exploitation, war, political manipulation and epistemicide, all of which are implicated in the racialized experience and incidence of disability in the Global South. These themes recur in the sections below, as they must, in an acknowledgment of intersectionality and complexity.

Disability and class/socio-economic status

Race and class/socio-economic status influence how disability is experienced (Ferri, 2006). The focus in this section is specifically on the intersection of disability and class/socio-economic status, recognizing that race and class/socio-economic status themselves intersect in mutually constitutive ways in different countries. I use the term class/socio-economic status deliberately, but with caution, recognizing that they are interwoven, but not synonymous.

Poverty and disability are closely linked (Singal et al., 2019; UNESCO, 2018a). Inadequate healthcare, especially maternity and early years care, poor diet, unsanitary living conditions and dangerous work are all associated with poverty and cause impairments. In turn, being disabled is a predictor of being poor, due to inadequate social support funding, various barriers to education and employment and the costs of healthcare, support workers, transport and assistive devices. The compounding experience of poverty and disability is acknowledged as life-long and intergenerational, but has particular resonance in education. Having a poor parent with a disability increases the likelihood that a child will not attend school in some countries. Disabled children tend to leave school early in poor countries, or do not attend school at all (UNESCO, 2018a). But living in a richer country is no guarantee that disabled students will flourish. The policy of austerity and spending cuts in education in the UK has had a “crippling” impact on support available for disabled children, with Ryan (2019, p. 180) noting that “children with the highest levels of need are paying the biggest price”. The price is paid in loss of transport, therapies and one-to-one assistance that enables inclusion in mainstream schools. The result, notes Ryan, has been the increase in enrollment in special schools, contrary to the aims of inclusive education.

Disability diagnosis has a complex relationship with class/socio-economic status and, in turn, is linked to incentives around resourcing. A child with low socio-economic status is more likely to be diagnosed with a behavior disorder in countries like the USA, UK and Australia (Harwood, 2006). Harwood and Allan (2016) note that in Australia, rates of medication for young people diagnosed with ADHD are highest in disadvantaged communities. In that country, taking medication can be a condition of enrollment, and the imbalance of power between school and parents from disadvantaged communities makes resistance difficult (Harwood, 2006). These power imbalances may also make it more difficult for parents of working-class backgrounds to advocate for their disabled children to ensure they get the support and resources to which they are entitled. In some contexts, though, it is richer parents who can mobilize their economic and cultural capital to have their children diagnosed to benefit from available resources. Tomlinson (2017, p. 129) says that there is evidence in England that middle class parents fear the consequences of their children not performing well at school and would prefer “to claim that their children have medical, neurological or therapeutic conditions”. Race and class thus intersect with disability in complex ways, always to the benefit of the powerful. Unequal gender relations also impact disabled people.

Disability and gender

Gendered biases in diagnosis, particularly in neuropsychiatric conditions, is common. Globally, boys make up more than 50% of students classified as having “special educational needs” and needing additional support (Riddell, 2014, p. 104). Boys are more likely than girls to be diagnosed with ADHD (Harwood and Allan, 2016) and Autism (Dukekot et al., 2017). Higher referral rates may explain these diagnostic imbalances, with referral based on the identification of behavior that is defined as disordered. Harwood and Allan (2016, p. 160) say, “It is difficult to ignore how dangerous it is to be a boy, or at least a certain type of boy, in the era of psychopathologization”. While boys may have the edge on the “bad”, girls and women have historically been “sad” and “mad”. The pathologization of women’s bodies and experiences by a patriarchal and sexist medical world has been well documented by feminist disability scholars (Thomas, 1999; Goodley, 2017). These scholars point to the conflated devaluation of the bodies of women and disabled people, construed as weak, dependent and deficient, always in need of rescue and support (Garland-Thomson, 2010).

The Global Education Monitoring Report on Gender Equality (UNESCO, 2020) celebrates the “dramatic improvements” in girls’ access to, and achievement in education. However, it recognizes that it is girls who are intersectionally disadvantaged by poverty or disability who “face the worst forms of acute exclusion in the world’s poorest countries” (ibid. p. viii). Mozambique is offered as an example where 49% of disabled men can read and write, in contrast to only 17% of disabled women (ibid. p. 1). The report highlights infrastructural barriers to education such as inaccessibility of single-sex sanitation facilities for disabled girls and women in some countries. It also draws attention to the problem of policies that focus on a single excluded identity (like gender) but fail to

appreciate the particular exclusion produced by the intersection of other identities (like disability). An approach to policy that “analyses interlocking institutional formations of power” (ibid. p. 6) is recommended by the Report and signals a welcome recognition of the importance of intersectionality in responding to educational exclusion.

Disability studies in education needs to keep attention on the interstices of disablism and sexism, but also consider related oppressions of homophobia and transphobia. Duke’s (2011) comprehensive meta-synthesis of research shows the discrimination and prejudice that disabled LGBT students experience in many contexts because of their intersecting devalued identities, the inadequate sex- and sexuality education that they receive, especially in special education settings, and homosexuality portrayed as “a forbidden topic” (ibid. p. 37). The synthesis also notes that female participants are underrepresented in research with disabled LGBT young people. The recognition of queer/trans* disabled identity intersections has been acknowledged in some higher education spaces (Miller et al., 2019), where the fraught issue of identity disclosure points to the pervasive impact of a hetero- and cis-normative, ableist education system.

Disability and migrant/refugee status

Limited, yet growing attention is being given to the compounded marginalization and exclusion from education experienced by disabled refugee and migrant students (Walton et al., 2020). Much of this research comes from Global North contexts at the point where recent migrants and refugees confront established systems of special education support. Concerns are mostly framed in terms of the identification and assessment of learning disabilities and difficulties. Language barriers may lead to migrant or refugee children being referred for assessments of cognitive ability and then performing poorly on these assessments, leading to disproportionate over-representation in special education settings (Bačáková and Closs, 2013; Kaplan et al., 2016; Nicolai et al., 2016). This in turn has negative consequences for the social integration that education offers recent arrivals (Crock et al., 2017). The effects of trauma on learning and behavior are not always easily disentangled from other disabilities, which further complicates assessment and support provision (Graham et al., 2016; Sirin and Rogers-Sirin, 2015). A different line of inquiry has been pursued in Austria where Bešić et al. (2020) found disability, refugee status and gender intersecting in the public’s attitude to inclusive education.

Research in the Global North and the Global South (Hurley et al., 2014; King et al., 2010; Ryan et al., 2010; Tadesse, 2014) shows under-reporting of disability by immigrant or refugee families. The reasons for this are complex and relate to insecure status, fear of further disadvantage, and different cultural conceptions of disability and educational support. Global South countries report lower rates of enrollment for disabled refugee students as a result of factors that include inadequate teacher training, lack of teaching aids and assistive devices, inflexible curricula and inaccessible buildings (Crock et al., 2017; Reilly, 2010). Refugee education and disability education are separated in policy provision and reported separately in monitoring, with minimal evidence that either considers the impact of the compounding experience of being both a refugee and a disabled person (Walton et al., 2020).

These examples show the complexity of the lived experience of disability and the differential impact of compounding vulnerabilities. This impact is well illustrated by the vagaries of disability (or “special needs”) diagnosis, incidence, and response, as these pivot on a range of other personal and environmental factors. They also show the importance of a situated analysis and the need for a recognition of the exigencies of context. The section below returns to South Africa to show how the complex interaction of relations of oppression produce educational and social inequality.

An intersectional case-study: South African education

South Africa’s system of legalized racial discrimination and segregation meant that white people benefitted from the political and economic power they exercised prior to, and through the apartheid state. In the deliberate manufacture of social and economic inequality, the apartheid government (in power from 1948 to 1994) spent five to six times more on white children than they did on Black African children (Taylor et al., 2013). As a result of education policies, Black African children were educated in under-resourced schools, many teachers were unqualified or underqualified, and curriculum options were limited. Higher student to teacher ratios existed in schools for Black Africans. Whereas most white students completed school during apartheid, many Black African students dropped out even before high school and their rate of university enrollment was much lower than that of their white counterparts. Gender inequality is evident, though, and in 1976 girls were more likely to leave school early (Christie, 1985, p. 111). White children also benefitted from a well-resourced, but separate special education system, which was based on a strong individual/medical deficit model. Separate schools were available for children with different categories of disability and specialized knowledge for special education was promoted (Engelbrecht, 2006). This was a country that privileged the white, male, able child at the expense of all others in the education system, and in his life beyond school. Very little provision was made for disabled Black African students, and by 2001 it was estimated that over 280,000 children and young people were out of school, mostly because of inadequate provision for disabled students (Department of Education (DoE), 2001).

The post-apartheid government of 1994 inherited a highly fragmented and deeply unequal education system, and immediately went to work to realize constitutional values of the right to education for all and freedom from discrimination. After a consultative process, White Paper Six was promulgated in 2001 (DoE, 2001) and offered a framework for building an inclusive education system. One of its earliest breaks with the past was to adopt the term “barriers to learning” in an attempt to capture the intersecting impact of socio-economic background and impairment. Around the time of the end of legalized apartheid, scholars like Donald (1994) had tried to shift the emphasis on “special need” which was assumed to apply to only a small percentage of school-going children. He proposed an acknowledgment of “deficits” caused by “intrinsic” factors (usually medical or organic in origin)

and also of deficits caused by “extrinsic” factors in the systems and structures of schools and society. This idea was taken up by White Paper Six which acknowledges that “different learning needs” arise from a number of factors, including impairments and “socio-economic deprivation” (DoE, 2001, p. 5). It has been carried forward into the country’s Policy on Screening, Identification, Assessment and Support (SIAS) (Department of Basic Education (DBE), 2014) which claims, as a principle that

Support should no longer focus only on the diagnosis and remediation of deficits in individual learners through individual attention by specialist staff. The SIAS shifts the focus to a holistic approach where a whole range of possible barriers to learning that a learner may experience (such as extrinsic barriers in the home, school or community environment, or barriers related to disabilities) are considered. The aim is to design support programmes in such a way that the learner gains access to learning.

However, this vision of a “holistic” approach does not always translate into improved educational access and success as economic capacity constrains what schools and the wider education system can resource (Engelbrecht, 2020).

The legacy of apartheid is bitter, longstanding and difficult to overturn. At the broadest level, Fleisch (2007) characterizes the post-apartheid education system as “bimodal” with two *de facto* “systems”. The first is generally well-functioning, made up mainly of former white and Indian schools, and independent schools, which together produce most of the country’s university entrants. This is very different from the second system, mostly serving poor learners from deprived communities, and which offers a “restricted set of knowledge and skills” (ibid. 2007, p. 2) compared with what is offered in the first system. Studies show that 58% of pupils are unable to read for meaning in any language by the end of Grade 4 (10 year olds), and 29% are completely illiterate (van der Berg et al., 2016, p. 16). White pupils are more likely to complete high school (Grade 12) and girls are less likely to repeat years and drop out (Taylor et al., 2013). The Department of Basic Education (DBE) (2015) admits that, “Currently, there are persistent challenges that retard the progress that is being made in the development of an Inclusive Education and Training System” (p. 6). Twenty five years after the end of legalized apartheid, disability remains the reason given most often for children and young people not attending school (DBE, 2019).

As might be expected from global trends, disabled people in South Africa spend fewer years in formal education than do non-disabled people, and disabled women have the lowest educational levels, reflecting the intersectional impact of misogyny and disablism (Moodley and Graham, 2015). But, because of apartheid’s legacy, race has the biggest impact on levels of educational attainment in South Africa, destabilizing the “poverty-disability nexus” (ibid. p. 28). Disabled white women still complete more years in education than non-disabled Black men, reflecting the differential privilege that whiteness affords (ibid.). The Committee on the Rights of Persons with Disabilities (United Nations, 2018) in its observations on South Africa’s country report noted a number of concerns regarding the ongoing intersectional disadvantage of disabled people. These particularly concern disabled girls and women who experience discrimination in accessing education, employment and health care. Black disabled women in particular experience compounded marginalization in their exclusion from involvement in decisions on matters that directly concern them. Concern is also noted about the gender-based violence, exploitation and abuse experienced by disabled girls and women, including instances in schools and school hostels. The committee recommends that South Africa should, “Adopt effective legislation and policies that will explicitly provide protection against multiple and intersectional forms of discrimination” (ibid. Section 9b).

As South Africa pursues inclusive education with a focus on a range of possible intersecting “barriers to learning”, the disabled community has raised concerns about the marginalization of disability in inclusive education discourses (Walton and Rusznyak, 2019) and the devaluation of specialist knowledge (McKenzie et al., 2020). Countering this, various disability organizations and parent advocacy groups campaign to advance the right to education. They draw on a vibrant tradition of activism in South Africa that traces its origins to early resistance movements against the colonial and apartheid states. With some exceptions, these organizations tend to be group specific, often focusing on advocacy around a particular disability or impairment (like Down Syndrome South Africa and QuadPara Association of South Africa). Their voices are heard in policy consultations and they use various civil and legal mechanisms to hold the government to account for the failure to realize the right to education for all (McKenzie et al., 2017). They have combined their efforts in submitting reports to the Committee on the Rights of Persons with Disabilities, with the *The Right to Education for Children with Disabilities Alliance’s* (2016) Alternative Report offering a scathing indictment of the South African government’s failure to implement inclusive education for disabled children.

The postapartheid state faces very real challenges in implementing inclusive education and the many “solutions” proposed by high-income, Global North countries fail to address the regimes of inequality that work together to produce educational exclusion in this context. The content-rich curriculum that has been introduced to address overall levels of low educational attainment has been found to be too fast paced and rigid to enable access and success by many disabled students (Andrews et al., 2019). Special schools are no longer racially segregated, but they charge fees and are mostly located in former white, urban areas, meaning either additional transport costs or hostel living (Human Rights Watch, 2015). The compounding experience of poverty and a disabling society is overwhelmingly experienced by Black African disabled students and their families, especially those living in rural areas. Mention must also be made of xenophobia and the invisibilisation of disabled refugee students and their families (Walton et al., 2020) and the endemic heteronormativity of the South African education system (Francis and Reygan, 2016). The lived experience of disability in South Africa is inevitably one of intersectional identities, shaped by historical and current systems and structures. Deepening an understanding of this intersectionality is important in this country’s pursuit of inclusive education, as it is elsewhere.

Leveraging intersectionality: necessary concepts

Intersectionality, like other powerful social theories, has the potential to be sanitized, tamed and domesticated, as it travels (Salem, 2018) from Crenshaw's original conceptualization. The term intersectionality in the institutionalized discourses of equality, diversity and inclusion easily becomes shorthand for "think about other identities too". The term is sometimes recruited to convey an additive approach that acknowledges multiple identities without interrogating the complexities at the confluence of these identities (Rice et al., 2019). Or, intersectionality gets an obligatory, but "cursory nod before moving on" (Annamma et al., 2016, p. 12). Core concepts of intersectionality are described in depth by Collins and Bilge (2020), Romero (2018) and others. Rather than repeating these, I suggest how intersectionality is potentially diminished, misrepresented or caricatured without certain ideas. My thinking is influenced by Janks's (2010) synthesis model which explains the effect of the absence of different aspects of critical literacy and Maton's (2014) idea of rarefaction, where a concept is stripped of meaning. In the section that follows, I will offer eight prerequisite propositions for advancing intersectionality, summarized in Table 1.

Intersectionality must reckon with the workings of *power and oppression*. Unequal power dynamics are at work "establishing normative orders, creating centers and margins, bringing about systems of privilege and disadvantage" (Steyn, 2015, p. 381). Not recognizing this leads to a focus on different identities and their possible intersections, without distinguishing which differences are unmarked or unremarkable, and which differences make a difference. Without acknowledging power, intersectionality becomes merely a celebration of differences in a fantasy land of diversity with assumed equality (Bhabha, 1994). Evident in bright and multi-colored depictions of "diversity" in images and logos of inclusive education, this approach creates the idea that people are all equally different (Walton and Dixon, 2020). The danger of intersectionality thus diminished is that disabled children get added to schools and their difference "celebrated" in what Erevelles (2005, p. 66) calls the "add-and-stir policy". "Empathy" and "understanding" of individual (even intersecting) identities (the "exotic other") are sought while ableism, racism and other axes of oppression remain invisible (Reygan et al., 2018).

Intersectionality must take *historicity* into account. This means recognizing historical forces that have co-constructed race, gender, disability and other identities. Intersectional oppression is not a modern phenomenon and ignoring the past risks misreading the present. In a historical void educational disadvantage or difficulty experienced by groups is seen as an aberration, a transient phenomenon, or as an individual problem that has to be remediated through individual response. The response is a superficial "what works" approach, designed to ameliorate negative circumstances or experiences, rather than to challenge their historical antecedents in structural oppression. The system is too easily cast as neutral, if not benign, and "non-achievement" is attributed to deficits associated with one or more of an individual's constituent identities. This prevents collective action against educational injustice (McCall and Skrtic, 2009) as each individual family is cast as an anomaly in the face of a smooth-running system that appears to work for the rest.

Intersectionality requires a *theory of knowers and knowledge* (cf Maton, 2014). It demands consideration of who knows what about whom and how they come by this knowledge, and what counts as legitimate knowledge about people's lives and experiences. There's long been a call for disabled people to be recognized as experts in their own lives, and for parents' knowledge of their disabled children to be legitimated (Van Hove et al., 2008). The hegemony of (Anglo-Western) professionals whose assumed expertise about the lives of others has been questioned (Titchkosky and Aubrecht, 2015), and the impetus to diagnose, categorize and treat has led to ever increasing niche areas of therapeutic and pedagogical knowledge. Veck (2014, p. 784) asks,

What, then, might be said when the education of children with certain identified impairments is conceived as a specialist activity, to be conducted by specialist practitioners, and when the authority of these professionals derives neither from their knowledge of, nor from the responsibility they have taken for, the world, but rather from their specialist expertise?

Attempts to essentialize, that is, to understand and respond to someone in terms of a diagnostic or stereotyped category, are defied by an acknowledgment of intersectionality. Individual identities are forged at the nexus of privileging and disprivileging

Table 1 Eight prerequisite propositions for advancing intersectionality.

<i>Prerequisites for advancing intersectionality</i>	<i>Without which, intersectionality ...</i>
Reckoning with power and oppression	Becomes a fantasy-land of diversity with assumed equality.
Historicity	Risks exceptionalism as disadvantage is viewed as a temporary aberration.
A theory of knowledge and knowers	Risks essentializing people according to their characteristics and individualizing their problems.
Geography, culture and context	Risks universalism by not recognizing that regimes of inequality work differently in different locations.
Justice	Risks abstraction and loses its emancipatory impact on people's lived experience of oppression.
Relationality	Risks silos and suspicion as intersectional allyship becomes impossible, resulting in polarization and competition.
Embodiment	Risks symbolism and the creation of disembodied hierarchies of inequality.
Agency	Risks determinism and saviourism, and the erasure of self and community advocacy.

or disadvantaging structures. These individuals need to be accorded the “testimonial justice” that Fricker (2007) describes, that is, credibility untainted by prejudice and stereotype. However, in the political quest to dismantle educational exclusion, we need to be cautious about proliferating the specialist knower (Walton, 2016) around what Slee (2011, p. 64) calls “constituent interests”. The neo-liberal education regime thrives on the narrative of the individual, and exclusion is easily justified as being a benign response in the interest of the individual (Silver, 1994). Research, in an effort to produce knowledge about “unique” experiences and learning needs, works to fragment the student population, focusing on difference, and producing the researchable Other. In so doing, the “problem” of the individual disabled child becomes the starting point of research and the focus of inclusion (Slee, 2011, p. 82). A more productive approach seems to be that taken by the General Education Monitoring Report which says “mechanisms of exclusion are common regardless of group” (UNESCO, 2018b, p. 4). Knowledge of these common mechanisms of exclusion should command our attention and offer a way forward to advance intersectional interests in education.

Intersectionality has to recognize ways in which the peculiarities and exigencies of *geography, culture and context* interact with individual identities to differentially privilege and disprivilege people. Subject positions are constituted by and through location and place (Code, 2006). Without this recognition, experience is assumed to be universal. Regimes of inequality work differently in different locations, as the South African case study demonstrates. Too often, the experience of the white (male) child in a high-income context has been taken to be normative (Grech, 2015) in accounts of the impact of disability on learning and schooling and many of the pedagogical practices described as “inclusive” assumes this child’s presence in a well-resourced setting. The experience in many Global South contexts may be very different, where factors at the level of the system, the community and the school combine to impact learning outcomes. Individual schools configure their own “objects of activity” which may make it more or less likely that disabled learners are included (Andrews et al., 2019). And regimes of inequality are exercised beyond the education system, with students’ experiences impacted by encounters with the caprices of the health system, legal frameworks (especially for migrant and refugee students), and economic possibilities and constraints.

Intersectionality demands a commitment to *justice*. Without translating the concept of intersectionality into activism and action, it remains an interesting but abstract concept with no emancipatory impact on people’s lived experience of oppression. Or, as Collins and Bilge (2020, p. 239) note, “intersectionality demands more than simply being critical and entails turning critical analyses into critical praxis”. Drawing on a Freirean concept of critical pedagogy, and linking to disability studies in education, Van Hove et al. (2008) indicate what this praxis might entail. It includes genuine dialog with parents and students; identifying strengths and “goodness” rather than focusing on vulnerability and sources of failure; person-centered planning; acknowledging students’ broader familial network; and three key ways of seeing the Other – as thinking, individual and reciprocal beings. Ensuring safety, belonging and success within a moral framework of participatory parity is how McIntyre (2021) sees social justice operationalized in inclusive schooling. While this author focuses on refugee students, she offers a compelling argument for the possibility of transformed educational spaces that make an academic, social and material difference to young lives.

Intersectionality demands *relationality*. This means allowing for dialog that enables people to navigate their differences. Building “intellectual solidarities between distinct yet interrelated scholarly formations” (Collins and Bilge, 2020, p. 197) is necessary for intersectional work, with conflict seen as inevitable, but potentially creative and generative. Without relationality, intersectional allyship and interdependency are impossible and polarization and competition are inevitable. Interest groups potentially become silos, suspicious of any attention or advantage that other groups may command. “Whataboutery” (a new noun) is invoked when intersectionality is misunderstood and relationality is absent. It says during #BlackLivesMatter, “What about white working class boys?”, as if oppression only pivots on one axis of power. Equally destructive is presuming to pick and choose which axes of oppression do or do not deserve attention, like being “disability-friendly” but homophobic. Relational intersectionality offers “new ways of being in the world” (Lorde, 2019, p. 104) and an opportunity to understand and resist common mechanisms of exclusion.

Intersectionality is an embodied experience (Elg and Jensen, 2012). *Embodiment* frustrates attempts to create hierarchies among regimes of inequality because students are not sequentially raced, gendered and abled, but are simultaneously so. Without considerations of embodiment, intersectional analysis is symbolic. And the violence meted out on intersectional bodies is real, not symbolic. This violence includes, but is certainly not limited to the forced sterilization and menstrual suppression of disabled girls, the rape of disabled girls and women, and pharmaceutical and physical restraints used on disabled boys. The historical and neoliberal valorization of the “ideal fit body” (Grech, 2015, p. 19) necessarily constitutes other bodies as transgressive, in need of control through scrutiny, classification and punishment. The “ideal” body, is, of course, white, male, heterosexual and, compulsorily able-bodied (McRuer, 2010). The “ideal” schooled body is one that complies with instructions to “stand straight, sit still, pay attention, write neatly, tuck in escaped shirts, pull up socks, stop talking, running, hitting, screaming, sniffing, swearing, spitting, chewing” (Dixon, 2011, p. 3). That many bodies can’t or won’t comply offers the possibility for reimagining educational spaces that account for multiple forms of embodiment.

Without recognizing *agency*, intersectionality can drift into determinism. People appear as the pitiable sum of their constituent (devalued) identities with predetermined destinies. The narrative can steer dangerously close to the outdated charitable discourses in which students are gazed upon by “enlightened” critical scholars who catalog their woes. Disability studies in education needs an approach to intersectionality characterized by “resistance, hope and reclamation” (Berghs et al., 2020, p. 8). In Grech’s (2015, p. 90) view, a “focus on negative attitudes and oppression strips disabled people and their households of any form of agency and the ability/possibility to resist and control/change their circumstances, and influence other people’s attitudes and behaviors.” Collective action and activism by disabled students, their families and allies against disablism and other oppressions in the education system needs to be supported. Research needs to be co-produced and stories of resilience and resistance need to be told, rather than chronicles of deficit and misery. The self and community advocacy of marginalized groups needs to be celebrated and amplified.

Moreover, the temptation to saviourism needs to be repelled. The “everyday relational agency” that disabled children have needs recognition, with an acknowledgment of their presence in the community and “the difference they make through it” (McLaughlin et al., 2016, p. 175). Power is, of course, not only repressive. It circulates and can be harnessed in various forms of resistance. Intersectional alliances need to be forged to dismantle the common mechanisms of educational exclusion.

Conclusion: the promise of intersectionality in and for disability studies in education

Intersectionality is not an “optional extra” consideration for disability studies in education, it is central to its concerns. Both are critical projects which challenge relations of power, exploitation and domination. They describe the spaces for counter-hegemonic action and offer the possibility for realizing every child’s right to an inclusive education (UNESCO, 2016). Intersectionality forces disability studies in education to consider the construction of the disabled student not only by ableism and disablism, but also by racism, sexism, classism and other regimes of inequality. It dislodges the gaze from single-identity politics to complex life worlds forged at the cusp of structures and circumstances that simultaneously but differentially privilege and disprivilege. Intersectional allyship that offers “solidarity across movements” (Carlson et al., 2019, p. 892) has the potential to identify and disrupt the common mechanisms of educational exclusion.

The response to exclusion cannot be compensatory measures that render “the included” as perpetual guests, welcome in schools and other educational spaces conditional on their acquiescence and assimilation into normative cultures and practices. Recognizing the effects of inequitable power relationships is a task of the “first order” (Slee, 2011, p. 157) in achieving inclusive education for all, and particularly for disabled students. Intersectionality, with all its necessary component concepts, offers a means to do this recognition work and thereby to begin to forge a more just, democratic and inclusive education system.

References

- Andrews, D., Walton, E., Osman, R., 2019. Constraints to the implementation of inclusive teaching: a cultural historical activity theory approach. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2019.1620880>.
- Annamma, S., Connor, D., Ferri, B., 2016. Dis/ability critical race studies (DisCrit). In: Connor, D., Ferri, B., Annamma, S. (Eds.), *DisCrit*. Teachers College Press, New York, pp. 9–31.
- Artiles, A.J., 2004. The end of innocence. *J. Learn. Disabil.* 37 (6), 550–555.
- Bačáková, M., Closs, A., 2013. Continuing Professional Development (CPD) as a means to reducing barriers to inclusive education: research study of the education of refugee children in the Czech Republic. *Eur. J. Spec. Needs Educ.* 28 (2), 203–216.
- Berghs, M., Chataika, T., Dube, K., El-Lahib, Y., 2020. Introducing disability activism. In: Berghs, M., Chataika, T., El-Lahib, Y., Dube, K. (Eds.), *The Routledge Handbook of Disability Activism*. Routledge, London, pp. 3–20.
- Bešić, E., Paleczek, L., Gasteiger-Klicpera, B., 2020. Don't forget about us: attitudes towards the inclusion of refugee children with(out) disabilities. *Int. J. Incl. Educ.* 24 (2), 202–217.
- Bhabha, H., 1994. *The Location of Culture*. Routledge, London.
- Broderick, A., Reid, D., Valle, J., 2006. Disability studies in education and the practical concerns of teachers. In: Danforth, S., Gabel, S. (Eds.), *Vital Questions Facing Disability Studies in Education*. Peter Lang Publishers, New York, pp. 133–160.
- Carbado, D., Crenshaw, K., Mays, V., Tomlinson, B., 2013. Intersectionality: mapping the movements of a theory. *Du. Bois Rev.* 10 (2), 303–312.
- Carlson, J., Leek, C., Casey, E., Tolman, R., Allen, C., 2019. What's in a name? A synthesis of “allyship” elements from academic and activist literature. *J. Fam. Violence* 35 (8), 889–898.
- Chirwa, E., Jewkes, R., Van Der Heijden, I., Dunkle, K., 2020. Intimate partner violence among women with and without disabilities: a pooled analysis of baseline data from seven violence-prevention programmes. *BMJ Global Health* 5 (11), 1–13.
- Christie, P., 1985. *The Right to Learn*. Sached/Ravan, Johannesburg.
- Code, L., 2006. *Ecological Thinking: The Politics of Epistemic Location*. Oxford University Press, Oxford.
- Collins, P.H., Bilge, S., 2020. *Intersectionality*, second ed. Polity Press, Newark.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1300.
- Crock, M., Smith-Khan, L., McCallum, R., Saul, B., 2017. *The Legal Protection of Refugees With Disabilities: Forgotten and Invisible?* Edward Elgar Publishing, Cheltenham.
- Davis, K., 2020. Who owns intersectionality? Some reflections on feminist debates on how theories travel. *Eur. J. Wom. Stud.* 27 (2), 113–127.
- Department of Basic Education (DBE), 2014. Policy on Screening, Identification, Assessment and Support. Government Printers, Pretoria. <https://www.education.gov.za/LinkClick.aspx?fileticket=2bB7EaySbcw%3Dandtabid=617andportalid=0andmid=2371>.
- Department of Basic Education (DBE), 2019. General Household Survey Focus on Schooling 2018. https://www.education.gov.za/Portals/0/Documents/Reports/20191104_GHS_final.pdf?ver=2019-11-04-111833-313.
- Department of Basic Education (DBE), 2015. Report on the Implementation of Education White Paper 6 on Inclusive Education. Government Printers, Pretoria. <http://pmg-assets.s3-website-eu-west-1.amazonaws.com/160308overview.pdf>.
- Department of Education (DoE), 2001. Education White Paper 6: Special Needs Education. Building an Inclusive Education and Training System. Department of Education, Pretoria. https://www.gov.za/sites/default/files/gcis_document/201409/educ61.pdf.
- Dixon, K., 2011. *Literacy, Power, and the Schooled Body*. Routledge, New York.
- Donald, D., 1994. South Africa. In: Mazurek, K., Winzer, M. (Eds.), *Comparative Studies in Special Education*. Gallaudet University Press, Washington, pp. 5–24.
- Duke, T.S., 2011. Lesbian, gay, bisexual, and transgender youth with disabilities: a meta-synthesis. *J. LGBT Youth* 8 (1), 1–52.
- Duvekot, J., van der Ende, J., Verhulst, F.C., Slappendel, G., van Daalen, E., Maras, A., Greaves-Lord, K., 2017. Factors influencing the probability of a diagnosis of autism spectrum disorder in girls versus boys. *Autism* 21 (6), 646–658.
- Eig, C., Jensen, S.Q., 2012. The intersectional body: an embodiment perspective on differentiated experiences. *Sociologisk Arbejdsrapport* 34 (1), 1–27.
- Engelbrecht, P., 2006. The implementation of inclusive education in South Africa after 10 years of democracy. *Eur. J. Psychol. Educ.* 21 (3), 253–264.
- Engelbrecht, P., 2020. Inclusive education: developments and challenges in South Africa. *Prospects*. <https://doi.org/10.1007/s11125-020-09499-6>.

- Erevelles, N., 2005. Rewriting critical pedagogy from the periphery: materiality, disability and the politics of schooling. In: Gabel, S. (Ed.), *Disability Studies in Education: Readings in Theory and Method*. Peter Lang, New York, pp. 65–83.
- Ferri, B., 2006. Teaching to trouble. In: Danforth, S., Gabel, S.L. (Eds.), *Vital Questions Facing Disability Studies in Education*. Peter Lang, New York, pp. 289–306.
- Fleisch, B., 2007. Primary Education in Crisis. Juta and Co, Cape Town.
- Francis, D., Reygan, F., 2016. Let's see if it won't go away by itself.' LGBT microaggressions among teachers in South Africa. *Educ. Change* 20 (3), 180–201.
- Fricker, M., 2007. *Epistemic Injustice*. Oxford University Press, Oxford.
- Frisher, E., Elbers, W., Okwany, A., 2019. Gendered disability advocacy: lessons from the girl power programme in Sierra Leone. In: Berghs, M., Chataika, T., El-Lahib, Y., Dube, K. (Eds.), *The Routledge Handbook of Disability Activism*. Routledge, London, pp. 412–427.
- Fruja Amthor, R., 2017. "If only I did not have that label attached to me": foregrounding self-positioning and intersectionality in the experiences of immigrant and refugee youth. *Multicult. Perspect.* 19 (4), 193–206.
- Garland-Thomson, R., 2010. Integrating disability, transforming feminist theory. In: Davis, L. (Ed.), *The Disability Studies Reader*, third ed. Routledge, New York/Abingdon, pp. 353–373.
- Goodley, D., 2013. Dis/entangling critical disability studies. *Disabil. Soc.* 28 (5), 631–644.
- Goodley, D., 2017. *Disability Studies*, second ed. Sage, London.
- Graaff, K., Heineken, L., 2017. Masculinities and gender-based violence in South Africa: a study of a masculinities-focused intervention programme. *Dev. South Afr.* 34 (5), 622–634.
- Graham, H.R., Minhas, R.S., Paxton, G., 2016. Learning problems in children of refugee background: a systematic review. *Pediatrics* 137 (6), e20153994.
- Grech, S., 2015. Decolonising Eurocentric disability studies: why colonialism matters in the disability and global South debate. *Soc. Ident.* 21 (1), 6–21.
- Harry, B., 2014. The disproportionate placement of ethnic minorities in special education. In: Florian, L. (Ed.), *The Sage Handbook of Special Education*. Sage, London, pp. 73–95.
- Harwood, V., Allan, J., 2016. *Psychopathology at School*. Routledge, Abingdon.
- Harwood, V., 2006. The new outsiders: ADHD and disadvantage. In: Graham, L. (Ed.), *(De)constructing ADHD*. Peter Lang, New York, pp. 119–139.
- Human Rights Watch, 2015. *Complicit in Exclusion*. <https://www.hrw.org/report/2015/08/18/complicit-exclusion/south-africas-failure-guarantee-inclusive-education-children>.
- Hurley, J., Warren, R., Habalaw, R., Weber, L., Tousignant, S., 2014. Early childhood special education in a refugee resettlement community: challenges and innovative practices. *Early Child. Dev. Care* 184, 50–62.
- Janks, H., 2010. *Literacy and Power*. Routledge, London.
- Kaplan, I., Stolk, Y., Valibhoy, M., Tucker, A., Baker, J., 2016. Cognitive assessment of refugee children: effects of trauma and new language acquisition. *Transcult. Psychiatr.* 53 (1), 81–109.
- King, G., Esses, V., Solomon, N., 2010. Immigrant and refugee families raising children with disabling conditions: a review of the international literature on service access, service utilization and service care experiences. In: Grigorenko, E. (Ed.), *Handbook of US Immigration and Education*. Springer, New York, pp. 179–206.
- Lorde, A., 2019. *Sister Outsider*. Penguin Classics, London.
- Macupe, B., 2015. I Want to Read, Write. *Sowetan Live*. <https://www.sowetanlive.co.za/news/2015-02-21-i-want-to-read-write/>.
- Maton, K., 2014. *Knowledge and Knowers*. Routledge, Abingdon.
- McCall, Z., Skrtic, T., 2009. Intersectional needs politics: a policy frame for the wicked problem of disproportionality. *Mult. Voices Ethn. Div. Except. Learn.* 11 (2), 3–23.
- McIntyre, J., 2021. Section one: theorising practice in schools. In: McIntyre, J., Abrams, F. (Eds.), *Refugee Education: Theorising Practice in Schools*. Routledge, Abingdon, pp. 9–43.
- McKenzie, J., Pillay, S., Duvenhage, C.-M., Du Plessis, E., Jelsma, J., 2017. Implementation of educational provision for children with severe to profound intellectual disability in the western cape: from rights to reality. *Int. J. Disabil. Dev. Educ.* 64 (6), 596–611.
- McKenzie, J., Kelly, J., Moodley, T., Stofile, S., 2020. Reconceptualising teacher education for teachers of learners with severe to profound disabilities. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2020.1837266>.
- McLaughlin, J., Coleman-Fountain, E., Claverling, E., 2016. *Disabled Childhoods: Monitoring Differences and Emerging Identities*. Routledge, Abingdon.
- McRuer, R., 2010. Compulsory able-bodiedness and queer/disabled existence. In: Davis, L. (Ed.), *The Disability Studies Reader*, third ed. Routledge, New York, pp. 383–392.
- Miller, R., Wynn, R., Webb, K., 2019. "This really interesting juggling Act": how university students manage disability/queer identity disclosure and visibility. *J. Diver. Higher Educ.* 12 (4), 307–318.
- Moodley, J., Graham, L., 2015. The importance of intersectionality in disability and gender studies. *Agenda* 29 (2), 24–33.
- Nicolai, S., Wales, J., Aiazzi, E., 2016. Education, Migration and the 2030 Agenda for Sustainable Development. ODI, London. http://doc.rero.ch/record/308985/files/28-6_ODI_Education.pdf.
- O'Connor, C., Fernandez, S.D., 2006. Race, class, and disproportionality: Re-evaluating the relationship between poverty and special education placement. *Educ. Res.* 35 (6), 6–11.
- Reid, D.K., Valle, J., 2004. The discursive practice of learning disability. *J. Learn. Disabil.* 37 (6), 466–481.
- Reilly, R., 2010. Disabilities among refugees and conflict-affected populations. *Forced Migr. Rev.* 35, 8–10. <https://www.fmreview.org/disability/reilly>.
- Reygan, F., Walton, E., Osman, R., 2018. Assimilation and celebration? Discourses of difference in education and the development of critical diversity literacy. In: Walton, E., Osman, R. (Eds.), *Teacher Education for Diversity: Conversations from the Global South*. Routledge, Abingdon, pp. 1–20.
- Rice, C., Harrison, E., Friedman, M., 2019. Doing justice to intersectionality in research. *Cult. Stud.* 19 (6), 409–420. *Critical Methodologies*.
- Riddell, S., 2014. A sociological perspective on special education. In: Florian, L. (Ed.), *The Sage Handbook of Special Education*, second ed. Sage, London, pp. 97–109.
- Riddle, S., 2018. Resisting educational inequity and the 'bracketing out' of disadvantage in contemporary schooling. In: Gannon, S., Hattam, R., Sawyer, W. (Eds.), *Resisting Educational Inequality*. Routledge, Abingdon, pp. 225–235.
- Romero, M., 2018. *Introducing Intersectionality*. Polity Press, Cambridge.
- Ryan, L., D'Angelo, A., Sales, R.A., Rodrigues, M., 2010. Newly Arrived Migrant and Refugee Children in the British Educational System. Project Report. Middlesex University, London. https://eprints.mdx.ac.uk/9399/4/AFSL_-_Research_Report_-_final.pdf. (Accessed 14 January 2019).
- Ryan, F., 2019. *Crippled: Austerity and the Demonization of Disabled People*. Verso, London.
- Salem, S., 2018. Intersectionality and its discontents: intersectionality as traveling theory. *Eur. J. Wom. Stud.* 25 (4), 403–418.
- Silver, H., 1994. Social exclusion and social solidarity: three paradigms. *Int. Lab. Rev.* 133 (5/6), 531–576.
- Singal, N., Johansson, S.T., Lynch, P., 2019. Education of children with disabilities: changing landscape of new opportunities and challenges. In: Singal, N., Lynch, P., Johansson, S.T. (Eds.), *Education and Disability in the Global South: New Perspectives from Africa and Asia*. Bloomsbury Academic, London, pp. 1–18.
- Sirin, S.R., Rogers-Sirin, L., 2015. The Educational and Mental Health Needs of Syrian Refugee Children. Migration Policy Institute, Washington, DC. <http://old.worldomeg.org/wp-content/uploads/2016/02/FCD-Sirin-Rogers-FINAL-4.pdf>.
- Slee, R., Corcoran, T., Best, M., 2019. Disability studies in education – building platforms to reclaim disability and recognise disablement. *J. Disab. Stud. Educ.* 1–11. <https://doi.org/10.1163/25888803-00101002>.
- Slee, R., 2011. *The Irregular School*. Routledge, Abingdon.
- Slee, R., 2018. Educational exclusion? In: Gannon, S., Hattam, R., Sawyer, W. (Eds.), *Resisting Educational Inequality*. Routledge, Abingdon, pp. 17–30.
- Steyn, M., 2015. Critical diversity literacy: essentials for the twenty-first century. In: Vertovec, S. (Ed.), *Routledge International Handbook of Diversity Studies*. Routledge, New York, pp. 379–389.
- Sulla, V., Zikhali, P., 2018. Overcoming Poverty and Inequality in South Africa: An Assessment of Drivers, Constraints and Opportunities. World Bank Group, Washington, D.C. <http://documents.worldbank.org/curated/en/530481521735906534/Overcoming-Poverty-and-Inequality-in-South-Africa-An-Assessment-of-Drivers-Constraints-and-Opportunities>.

- Tadesse, S., 2014. Parent involvement: perceived encouragement and barriers to African refugee parent and teacher relationships. *Child Educ.* 90 (4), 298–305.
- Tait, G., 2010. *Philosophy, Behaviour Disorders, and the School*. Sense Publishers, Rotterdam.
- Taylor, N., van der Berg, S., Mabogoane, T., 2013. Context, theory, design. In: Taylor, N., van der Berg, S., Mabogoane, T. (Eds.), *Creating Effective Schools*. Pearson, Cape Town, pp. 1–30.
- The Right to Education for Children with Disabilities, 2016. Alternative Report to the UN Committee on the Rights of Persons with Disabilities in Response to South Africa's Baseline Country Report of March 2013 on the UN Convention on the Rights of Persons with Disabilities, with Particular Reference to the Provisions of Article 24. https://tbinternet.ohchr.org/Treaties/CRPD/Shared%20Documents/ZAF/INT_CRPD_ICO_ZAF_30102_E.docx.
- Thomas, C., 1999. *Female Forms: Experiencing and Understanding Disability*. Open University Press, Buckingham.
- Tikly, L., 2020. *Education for Sustainable Development in the Postcolonial World*. Routledge, Abingdon.
- Titchkosky, T., Aubrecht, K., 2015. WHO's MIND, whose future? Mental health projects as colonial logics. *Soc. Ident.* 21 (1), 69–84.
- Tomlinson, S., 2017. *A Sociology of Special and Inclusive Education*. Routledge, Abingdon.
- UNESCO, 2016. General Comment 4: Article 24: Right to Inclusive Education. <https://www.right-to-education.org/resource/general-comment-4-article-24-right-inclusive-education>.
- UNESCO, 2018a. GEM Report Summary on Disabilities and Education. https://en.unesco.org/gem-report/sites/default/files/GAW2014-Facts-Figures-gmr_0.pdf.
- UNESCO, 2018b. Concept Note for the 2020 Global Education Monitoring Report on Inclusion. UNESCO, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000265329>.
- UNESCO, 2020. A new generation: 25 years of efforts for gender equality in education. In: *The Global Education Monitoring Report on Gender Equality*. UNESCO, Paris. <https://en.unesco.org/gem-report/2020genderreport>.
- United Nations, 2018. Committee on the Rights of Persons With Disabilities: Concluding Observations on the Initial Report of South Africa. https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CRPD%2fC%2fZAF%2fCO%2f1andLang=en.
- van der Berg, S., Spaul, N., Wills, G., Gustafsson, M., Kotzé, J., 2016. Identifying Binding Constraints in Education. Available at: SSRN 2906945. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2906945.
- Van Hove, G., Roets, G., Mortier, K., De Schauwer, E., 2008. Research in inclusive education as a possible opening to disability studies in education. In: Gabel, S. (Ed.), *International Handbook of Disability Studies in Education*. Peter Lang, New York, pp. 121–140.
- Veck, W., 2014. Disability and inclusive education in times of austerity. *Br. J. Sociol. Educ.* 35 (5), 777–799.
- Walton, E., Dixon, K., 2020. Googling inclusive education: a critical visual analysis. *Discourse*. <https://doi.org/10.1080/01596306.2020.1819203>.
- Walton, E., McKenzie, J., 2020. The education of children with disabilities in South African online news reports. In: Garrisi, D., Johanssen, J. (Eds.), *Other Bodies: Disability, Media, and Representation*. Routledge, Abingdon, pp. 135–155.
- Walton, E., Rusznyak, L., 2019. Developing standards for inclusive teaching in South Africa: a dilemma analysis. *South. Afr. Rev. Educ.* 25 (1), 89–106.
- Walton, E., McIntyre, J., Awidi, S.J., De Wet-Billings, N., Dixon, K., Madziva, R., Monk, D., Nyoni, C., Thondhlana, J., Wedekind, V., 2020. Compounded exclusion: education for disabled refugees in Sub-Saharan Africa. *Front. Educ.* 5, 47.
- Walton, E., 2016. *The Language of Inclusive Education*. Routledge, Abingdon.
- Woods, A.D., Wang, Y., Morgan, P.L., 2020. Disproportionality and inclusion. In: Kauffman, J. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, Abingdon, pp. 10–126.

Cross-cultural perspectives on the rise of special education policy discourses.

Anastasia Liasidou^a and Katerina Mavrou^b, ^a Faculty of Education, Humanities and Social Sciences, St Mary's University, Twickenham, London, UK; and ^b Department of Education Sciences, European University Cyprus, Nicosia, Cyprus

© 2023 Elsevier Ltd. All rights reserved.

Introduction	259
Inclusive education policymaking in Cyprus; setting the context	260
Critical policy discourse analysis	261
Policy discourse analysis methods	261
Analysis and discussion	262
Individual pathology and deficit-oriented discourses	263
Minority rights discourses	264
Discursive absences and policy silences	265
Conclusions	266
References	267
Further reading	268

Introduction

Inclusive education is a globalized policy discourse that has been variously understood and enacted. Despite its international policy and legal endorsements, inclusive education has been plagued by ambiguity thereby confounding the processes and outcomes of inclusive education policymaking. Inclusive education is grounded in what Norwich (2010: 93) calls “a plural values framework” whereby contradictory values are juxtaposed, contested and negotiated. For instance, values of equity and social justice are frequently marginalized due to the ascendancy of “neoliberal” ideologies that undermine the process of change toward inclusion and social justice (Giroux, 2012; Hursh and Henderson, 2011). The complex and contentious nature of the notion is further complicated by the ways in which inclusive education has become “a traveling theory” (Slee, 2018) and has been intertwined and mediated through diverse “local theories” and socio-political exigencies. The dialectic of the global and local - captured through the ways in which local cultures and ideological dynamics infiltrate and eventually domesticate inclusive educational policies (Robertson 1995; cited in Green 2002) - has given rise to hybrid discourses of inclusion. These hybrid discourses create and exacerbate conceptual ambiguities around the notion and contribute to the ascendancy of contradictory and contentious policy landscapes that reduce inclusion to a special education artifact (Liasidou 2008a).

Even though as early as in 1994 the Salamanca Statement was explicit about the expectation that nation states should ensure that “both policies and financing arrangements should encourage and facilitate the development of inclusive schools” (UNESCO, 1994: 19), this is still not the case. On the contrary, there is global evidence suggesting the resurgence of special education discourses that undermine and confound the processes and outcomes of inclusive education policymaking (Kozleski, 2020; Artiles, 2020; Tomlinson, 2017). The “industry of special education” (Tomlinson, 2012) has not receded, it has become an even more sophisticated social engineering device to protect the interests of general schooling by removing potentially “troubled and troublesome students”, who are perceived to pose a major threat to schools’ unfettered quest for educational excellence gauged against mono-dimensional calibrations of students’ achievement linked to performance indicators, final examination results and data-driven accountability regimes (Armstrong et al., 2011; Dudley-Marling and Baker, 2012). Even more worryingly, this sophisticated engineering device has been masqueraded and legitimized through linguistic veneers aligned with an inclusive discourse. This is part of, according to Slee (2018: 13), “the cunning plan” devised “by exclusionists”, the aim being to “take over the concept of inclusion and arrange an educational system that excludes as far as possible, and in a variety of ways ...”.

Under the siege of what Ball and Olmedo (2013: 85) termed neoliberal governmental technologies, the blame is squarely placed on individual students who deviate from ideals prescribed by the “humanist” subject. The latter constitutes an ontological *a priori* that has engendered and legitimized social hierarchies and dualisms of normative and non-normative ways of being (Goodley et al., 2014: 343). In this context, students with special educational needs and disabilities (SEND) are constructed as being “non-ideal students” (Harwood and Humphry, 2008: 380) and “are often seen as a burden, a problem, a drain on the system, who make no civic contribution” (Campbell, 2012: 214).

The resurgence of special educational thinking is not only manifested in educational practices, but also in the use of language that inconspicuously, yet pervasively, constructs and reinforces individual pathology perspectives and discriminatory regimes that are euphemistically baptized as being “inclusive”. According to Lester et al. (2017: 2):

Language creates, codifies and conveys policy while at the same time macro and micro features of language -from broad rhetorical techniques to specific word choices- shape meanings and understandings of policy.

The realization of an inclusive discourse is partly a struggle over language (Fairclough, 2000). This aligns with Fairclough's theorizations of the "discourse driven" dimensions of educational and social change (e.g., Taylor, 2004). The power of language and its discursive manifestations constitute a significant, albeit occasionally subtle and seemingly innocuous, discursive impediment that, unless problematized and dismantled, will continue to jeopardize the process of change toward inclusion. In this respect, inclusive education reform efforts are, quoting Ball (1993, 15), "set within a moving discursive frame which articulates and constrains the possibilities and probabilities of interpretation and enactment".

Education policies and practices are constituted by and concomitantly engender discursive orthodoxies. The latter embody and are permeated by ideological presuppositions that are mediated and enacted through social practices, relations, and identities. The identification and problematization of discourses underpinning policies can enable researchers to understand why policy developments have been slow in addressing inequity, human rights violations, social injustices and transforming professional practices (Allan and Tolbert, 2019).

Even though official policy by itself does not constitute a panacea for promoting an inclusive discourse (Fulcher, 2015), it, nevertheless, plays a catalytic role in precipitating inclusive education reforms (Barton, 2008; Slee and Cook, 1999). As Ulmer and Lenhoff (2017: 175) point out, policy discourse "does more than passively exist in policy documents ... Rather, discourse is active language that once absorbed into how educational stakeholders think, speak, and act - has the capacity to filter down and influence interactions with individual students". Inclusive education reform efforts aimed at "'patching up' of the insufficiencies of the current education system" (Lindsay et al., 2020: 17) through facile and erroneous policy rhetoric are not only ineffective, but also raise concerns about the ways in which inclusive education has been "reduced to a new name for 'special education'" (Armstrong, 2013: 7).

This article utilizes a discursive analytic approach to problematize national education policy reform efforts, allegedly aimed at mobilizing education reforms geared toward inclusion. A policy discourse analysis (PDA) methodology exposes the euphemisms, fallacies, silences, and ideological messiness of official policy discourses that contribute to the paradoxical rise and consolidation of special education in Cyprus. The globalizing discourse of inclusion necessitates a cross-cultural policy dialogue that can potentially instigate a reflective and critical approach to exploring the challenges of pursuing an inclusive education policy reform agenda. The following section provides a brief overview of the Cypriot context for education, disablement and exclusion. The next section discusses the theoretical and methodological underpinnings of critical policy discourse analysis, while the final section reports on the findings of the study and their implications for the process of change toward inclusion.

Inclusive education policymaking in Cyprus; setting the context

Cyprus has a highly centralized educational system, where the education of children with disabilities has traditionally taken place in segregated special schools,¹ with integration unofficially been implemented since the late 1980s. In 1999 a Special Education legislation² providing for integration, established formal procedures of identification, assessment and provision of special education and support. However, the special schooling and segregated settings and practices are still an option in this law. In 2019 Cyprus has initiated a process still in progress, for education policy reforms in order to meet its legal obligations that arise from the ratification of the UNCPRD in 2011 (Law of 2011, N.8 (III)/2011). Though the procedures are progressing, there is no consensus of the involved stakeholders and groups in consultation (i.e., academics, parents groups, disability representatives), who criticize the new policies as non-inclusive. Despite political rhetoric on the need to address the shortcomings of the existing special education legal framework, the envisaged reforms do not provide, according to the recommendations of the UNCPRD Committee 2017³ "a clear legislative scope of inclusive education". The problematic policy landscape is, *inter alia*, manifested in the ways in which students with disabilities experience multiple forms of disablism, and are denied access to quality, non-discriminatory and equitable forms of educational provision (Liasidou, 2016; Liasidou and Hadjiyiannakou, 2019).

Failure to articulate and act upon a clear vision for an inclusive education reform agenda, can be partly attributed to the historical past of Cyprus and the ways in which issues of human rights and equity have been conceptualized and acted upon in a country that has long been striving to overcome its colonial past and to preserve its national integrity and identity. (Liasidou, 2008b; Symeonidou, 2009). As a result, "Cyprus is still struggling to establish a strong democratic tradition and thereby lacks a focus on human rights issues and equality" (Liasidou 2008b: 235); a historical antecedent that undermines attempts to recognize disabled people as rights-bearing subjects (Degener, 2016).

Developing an informed understanding of the temporal and contextual dimensions of inclusive education policymaking not only provides effective local responses to the challenges and possibilities of fostering greater inclusive policies and practices, but also contributes to a cross-cultural policy dialogue on the commonalities and idiosyncrasies of different socio-political systems and the ways in which they can address similar challenges of pursuing inclusive education reforms (Barton and Armstrong 2007).

¹Ministry of Education and Culture, The 1979 Special Education Act (N.47/1979). Nicosia, 1979.

²Ministry of Education and Culture, The 1999 Education Act for the Education of Children with Special Needs – N. 113(I)/99. Nicosia, 1999.

³UN (2017) Committee on the Rights of Persons with Disabilities: Concluding observations on the initial report of Cyprus. https://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CRPD%2fC%2fCYP%2fCO%2fi&Lang=en.

Critical policy discourse analysis

Linguistic approaches to education policy analysis have played an important role in exploring the role of discourse in promoting certain policy discourses at the expense of others, as well as providing nuanced insights into the ways in which these discourses are instantiated, negotiated, hybridized, antagonized, played out and enacted in the process of policy constitution and dissemination (e.g., Taylor 2004). Discourses are “broader ideas” communicated by text, speech and cultural artifacts that create meaning, construct human identities and social realities. A “text’s linguistic structure” for example “functions, as discourse, to highlight certain ideologies, while downplaying or concealing others” (Machin and Mayr 2015: 200). As Fairclough (2000: 1) contends while analyzing the ways in which language is used to obfuscate relations of domination: “Looked at from a language perspective, different representations/visions of the world are different ‘discourses’”. From a discourse perspective, policy making is, therefore, regarded as being implicated in the “politics of discourse” whereby different discourses compete for ascendancy (Yeatman, 1990; as cited in Fulcher, 2015).

Discourse theory has been widely used in education policy analysis to foreground dominant and marginalized policy discourses, as well as hybrid discourses that confound the processes and outcomes of education policy making (Ball, 1990; Marshall and Patterson, 2002; Taylor, 2004). Unlike other conventional policy analysis approaches that adopt a so-called rational/scientific lens to explore policy constitution and dissemination, Policy Discourse Analysis (PDA), “seeks to understand and disrupt systems of oppression and assumptions guiding inequitable practices” (Allan and Tolbert, 2019: 141).

Thus, PDA offers an alternative to conventional approaches to policy analysis by recognizing language as a pervasive, yet opaque, source of social power, and affording “policy scholars the opportunity to examine discourse-based constructions of key policy issues, including those related to (in)equity, the construction of targeted groups, articulation of political agendas, emergence conflict, and so on”. (Lester et al., 2017b: 57). PDA scholars are, therefore, socio-politically committed to exposing the subjugating effects of discourse, including prejudiced discourses (Van Dijk, 1984), which are produced and reproduced in subtle yet pervasive and pernicious ways and ultimately undermine inclusion.

PDA is informed by an interdisciplinary methodology drawn from critical, post-structural and feminist traditions and focuses on exploring the role of discourse in framing policies and promoting social justice (Allan and Tolbert, 2019). Dominant understandings of discourse in the context of policy analysis capture and embody Foucauldian notions of power in constituting social realities and human subjectivities and creating the ‘discursive contours’ within which inclusive and special education policy problems, solutions and subject positions are constituted (Liasidou, 2008a).

Foucault (1984) was particularly interested in problematizing and foregrounding the precarious nature of “truth” and “knowledge” and sought to provide a genealogical analysis of the provenance of certain “regimes of truth”, as these were manifested within specific socio-political contexts. Foucault was also interested in exploring issues of representation and theorizing the ensemble of discursive dynamics that are accountable for the constitution of social realities and human subjectivities in policies. Policy discourses are thus conceived of as being entangled in the power/knowledge couplet that shapes the ways in which “policy mobilizes educational discourses (of teaching and learning) to produce particular forms of subject positions” (e.g., the child additional support needs) (Burman, 2017: 70).

Policy discourse analysis methods

Policy discourse analysis (PDA) methodologies draw insights from Critical Discourse Analysis (CDA) (Gildersleeve and Hernandez, 2012; Lester et al., 2017a) that constitutes the most common discursive analytical approach utilized by education policy analysts (Lester et al., 2017b). CDA provides a “critical lens” through which policy scholars can explore and develop understanding of the ways in which “discursive practices may have major ideological effects that is, they can help produce and reproduce unequal power relations” (Wodak and Meyer, 2016: 6). The ideological constructs of policy documents are manifested and embodied in lexical choices, metaphors, rhetorical operations (e.g., parallelism, metaphor, deixis), euphemisms and polarizations, as well as other linguistic properties of the text that are constitutive elements of social reality (Van Dijk, 2016). Language is reified through discourse and acquires “ontological potency” to construct human identities and to ascribe meaning to “social reality”. Even single words or the absence of these words convey the pervasive effects of discourse in constructing human identities and ascribing meaning to social reality (Luke, 2002).

Fairclough (2016: 87) uses the term “semiosis” to encapsulate the “meaning-making” dimensions of discourse “as an element of social process”, and explains the ways in which discourse analysis focuses on different “semiotic modalities” embodied in language that is “dialectically related” to other elements of the social process; “interdiscursive analysis” constitutes “a mediating ‘interlevel’ which connects both linguistic analysis with relevant forms of social analysis” (Fairclough, 2016: 94). Discourses, therefore, do not exist in a spatial-temporal vacuum; they embody temporarily and spatially mediated knowledge and are embedded in, reified, and captured through an intricately and reciprocally related web of institutional, administrative, political, ideological, and other dynamics; hence the adoption of a “dialectical-relational approach” to delineate the nexus of language, power and exclusion (Fairclough, 2016: 88).

Thus, CDA is not limited to the linguistic micro-structures of the text, it extends to explore the ways in which “social relations, power, institutions, beliefs and cultural values are in part semiotic” (Fairclough, 2016: 87), while foregrounding the ways in which the linguistic/semiotic aspects of the text are implicated in the construction of “social wrongs” and subjugated human identities. In

this respect, the analytic edge moves to their higher stages of a discursive analytic approach, namely interpretation and explanation (Fairclough, 2001a), in order to explore the ways in which policy discourses are shaped against, and are the embodiments of wider socio-political dynamics and power imbalances, the aim being “to contribute to addressing the ‘social wrongs’ of the day; in a broad sense-injustice and inequality, by analyzing their sources and courses, resistance to them and possibilities for overcoming them” (Fairclough, 2016: 94).

Our selection of policy documents followed the policy discourse analysis steps recommended by Allan and Tolbert (2019) and Fairclough (2016). First, we identified the chronological period that the analysis should focus on and engaged in desk-based research to identify our theoretical sample, which is determined, according to Fairclough (2016: 95) by “how a research topic is constituted as an object of research”. The aim of the study was to address the following research questions:

- What discourses have informed the policy portrayals of problems, responses and subject positions linked to the education of children with Special Educational Needs and/or Disabilities (SEND)?
- How do these discourses re/produce special education imperatives that undermine the process of inclusive education reforms?

The corpus data (theoretical sample) of our policy discourse analysis endeavor consisted of online material derived from the Primary Education and Special Education webpages linked to the Cyprus Ministry of Education, Culture, Sports and Youth (MECSY) Website, as well as the following governmental documents that were published after 2019 and deemed to be best suited to address our research questions (Lochmiller and Lester, 2017).

- Europe 2020 Cyprus National Reform program, April, 2020⁴
- White Paper of Inclusive Education Law of 2019 «Ο περί Ενιαιίας Εκπαίδευσης (Δομές Υποστήριξης) Νόμος του 2019» and the accompanying Explanatory Note («Επεξηγηματικό Σημείωμα»)
- Formation of a New System for the Evaluation of the Educational Practice and Teachers, January, 2019⁵

The identification of the theoretical sample of the study was followed by the Analysis and Interpretation of the data using a discursive analytical approach as discussed earlier. The process of analysis was data driven (Namey et al., 2008), meaning that the thematic discourse categories were derived after reading the documents to minimize research bias. This was followed by the writing process with the aim, following Allan and Tolbert’s (2019: 144) policy discourse analysis process, “to illustrate dominant discourses and subject positions with thick, rich description that includes data excerpts representative of the range of sources examined”.

Bearing in mind the socially-just orientations of CDA, the emphasis is placed on bringing to the fore and problematizing the “social wrongs” (Fairclough, 2016: 95) embodied in articulations and instantiations of policy discourses. The identification of these discourses is followed by an analysis of the ideological effects of discourse and their implications for perpetuating the special education status, while articulating alternative policy “imaginaries” (Taylor, 2003) which despite their idealistic nature they are, nevertheless, “possible and indeed exist” (Fairclough, 2016: 105).

Despite the emancipatory “knowledge interest” and social change orientations of CDA (Fairclough 2001b: 230), policy discourse analysts should act reflexively and “be cautious of contributing to a new regime of truth” (Allan and Tolbert, 2019: 145). CDA has been characterized as an “ideological interpretation” and not an “analysis”, due to the subjective and selective interpretation and explanation of empirical evidence (Wetherell, 2001; Wood and Kroger, 2000). As a counterargument, Fairclough (2016) asserted that the theoretical underpinnings of CDA call for multiple interpretations and multi-perspectival analytical frameworks; A discourse is not static; it constitutes a “dynamic semiotic entity that is open to reinterpretation and continuation” (Reisigl and Wodak, 2016: 27) depending on the aims and focus of the research endeavor. Hence, Critical Discourse Analysts are explicit about their “positionality” and social justice orientations to decipher and interrogate the subjugating and discriminatory effects of (special education) discourse (Titscher et al., 2000), while trying to ensure that their interpretations are characterized by logical reasoning and are firmly embedded in empirical and theoretical evidence (Wood and Kroger, 2000).

Analysis and discussion

The coding and analysis schemes were developed based on the aims of the research questions looking into the conceptualization of special education and inclusive education in policy documents and the position of the competent policy-making governmental agencies such as the Cyprus Ministry of Education, Culture, Youth and Sports (MECSY). Dominant policy “discourses” include or consist of, drawing on Machin and Mayr’s (2015:21) typology, participants, locations, valued-based assumptions, behaviors ideologies and goals. Thus, codes included issues of childhood assessment and identification (e.g., diagnosis, medical/doctors, psychometric), issues of educational structures, placement and services (e.g., intervention, special education, support, school etc.), and issues of processes in relation to education, also including provisions and relevant definitions in official documents (e.g., differentiation, universal design for learning, other children, vulnerable groups, accessibility, different, comparison). Analysis was based on both the presence and absence of codes that were analyzed against the following themes: (a) individual pathology and deficit-oriented discourses, (b) minority rights discourses, and (c) discursive absences and policy silences.

⁴EU 2020 Cyprus National Reform Program: https://ec.europa.eu/info/sites/info/files/2020-european-semester-national-reform-programme-cyprus_en.pdf.

⁵Formation of a New System for the Evaluation of the Educational Practice and Teachers: <http://www.cna.org.cy/pdf/ant/06072019.pdf>.

Individual pathology and deficit-oriented discourses

The official website of the Cyprus Ministry of Education, Culture, Youth and Sports (MECSY), circulars and guidelines pertaining to the education of children with SEND, are underpinned by deficit-oriented and assimilationist perspectives. For example, rather than articulating a vision for inclusion and its pedagogical implications for adapting teaching methodologies and curricula to respond to the needs of learner diversity on the grounds of SEND, the emphasis is solely placed on enabling students to “adjust to the school environment”⁶ while stating explicitly that one of the key aims of the Department primary Education is “the increase of special education programmes and of programmes which support pupils with learning difficulties”.⁷ The notion of inclusion on the grounds of SEND and its human rights and learner diversity considerations (e.g., [Armstrong, 2017](#); [Acevedo and Nusbaum, 2020](#)), are glaringly absent, despite an emphatic acknowledgment “that there is awareness of the multicultural trends which are developing in the modern world”.⁸

Paradoxically, even though a national inclusive education reform agenda geared toward inclusion is underway, special education discourses and the institutional paraphernalia associated with it (special schools and special units) are defined as being one of the areas that the Department of Primary Education is responsible for, including the “support services offered to children with special needs who are taught inclusively in public primary and Pre-Primary schools”.⁹ The reference to alleged inclusive teaching arrangements is eclipsed by a poignant articulation of the role of diagnostic calibrations and therapeutic interventions:

The special needs of children with disabilities or other special educational needs are served and addressed effectively through the services and programs offered by Special Education, the operation of which aims at both early diagnosis and the provision of appropriate therapeutic interventions¹⁰.

A similar discursive paradox is evident in “The Explanatory Note” of the White Paper which states that “inclusive education is based on the social and not the medical model of diagnosis and intervention at an individual level”. Yet in paragraph 9.2 of the White Paper it is noted that “the provision of specialized support is decided after evaluation and decision from the Evaluation and Support team, which may include medical, psychological and other opinions”. Contrary to the alleged social model policy orientations, the reform policy document, embodies the “model of professionalism” that is aligned with the medical model of diagnosis ([Tomlinson, 2017: 105](#)). Additionally, the document refers to students’ placement in “multipurpose support infrastructures” or in a “Resource Center for Inclusion” (the two are defined separately) where students will be supported in accordance with specific timeframes based on their Individualized Educational Program (IEP). The ways these organizational arrangements are defined make painfully invisible how “dis/ability labels conceal the conflation of inadequate opportunities to learn”, and are used as a pretext to create “a polarized space” ([Kozleski, 2020: 343](#)); special schools and special units.

Along the same lines, the models of general, enhanced and specialized support suggested by the new White Paper (Part II, article 6) fundamentally reproduce the existing segregative educational structures of special education. The three tiers of educational support are presented as consecutive stages, in a form of a seamless progression across levels, based on the assumption that one will follow the other in case the first one fails to accommodate students’ educational needs. In essence, this leveling will function as a normalization process and subsequently, an “exit ticket” from inclusive education, for all students who do not respond to existing educational arrangements and for whom the “general model” will prove “inadequate”.

The sporadic enunciations of inclusion in the White Paper are reduced to the creation of the so-called “Resource Centers for Inclusion” that are a euphemistic rearticulation and re-configuration of existing “resource units” attached to mainstream schools, or even the existing special school as it is specifically noted in article 13(b) of the White Paper, referring to the “mutation” of the public special schools. This is an example of what [Fairclough \(2016: 89\)](#) calls, “a ‘colonizing’ recontextualization of discourses” and in this case “special education discourses” that have been semantically disguised and presented as being “inclusive”. Once a student is placed in a resource unit (see Resource Centers for Inclusion), this is considered a permanent and administratively convenient “modus operandi” of providing learning support to students with SEND ([Liasidou and Antoniou, 2013](#)).

The same individual pathology perspective is echoed in another official document allegedly aimed at the “Formation of a New System for the Evaluation of the Educational Practice and Teachers” published in 2019. The document uncritically refers to students “special education needs” and silences the dynamic and “evolving” nature of disability and special educational needs that results from the interaction of a person’s impairment with institutional and environmental barriers ([United Nations, 2008: 1](#)).

The dominance of the dichotomous thinking of special and general education, is also mirrored in the exponential increase of students who are placed in special schools and Resource (Special) Units which operate mostly as a separate class setting for a small number of students within the mainstream school environment. [Table 1](#) shows the numbers of students in Special Schools and Resource (Special) Units in the previous eleven academic years (2010–21). Data related to 2004–05 is also included as it represents

⁶MECSY Special Education webpage: http://www.moec.gov.cy/eidiki_ekpaidefsi/.

⁷MECSY Primary Education webpage: <http://www.moec.gov.cy/dde/en/index.html>.

⁸MECSY Primary Education webpage: <http://www.moec.gov.cy/dde/en/index.html>.

⁹MECSY Primary Education webpage: <http://www.moec.gov.cy/dde/en/index.html>.

¹⁰MECSY Special Education webpage: http://www.moec.gov.cy/eidiki_ekpaidefsi/.

Table 1 Numbers of students in primary schools, special schools and resource (special) units.

Academic year	Special schools		Resource (special) units in primary schools		Primary schools	
	Schools	Students	Units	Students	Schools	Students
2004–05	9	295	N/A	N/A	346	57,575
2010–11	9	288	75	398	344	49,889
2011–12	9	311	78	413	342	49,795
2012–13	9	343	78	445	340	49,612
2013–14	9	365	79	442	338	48,645
2014–15	9	373	82	472	334	48,796
2015–16	9	385	92	519	334	49,567
2016–17	9	404	96	576	332	50,423
2017–18	9	406	102	582	332	51,082
2018–19	9	439	110	620	331	52,042
2019–20	9	465	N/A	N/A	331	52,768
2020–21	9	498	N/A	N/A	331	51,541

N/A, numbers not available. 2010–21 number not yet published.

Resources: Ministry of Education, Culture, Sports and Youth, Annual Reports.

the first year that numbers of students are available. The data presented in the Table below is a clear indication of how the numbers of students educated in segregative settings have increased over years, in comparison with the total number of and disproportionate increase of students in primary schools.

Based on the data provided in the Annual Reports on the Ministry of Education, Cultures, Sports and Youth, as well as according to the “Statistics of Primary Education Department”¹¹ in 2004–05 there were 295 students who attended the nine special schools in Cyprus. As presented in **Table 1**, there has been a gradual increase of this number so that in 2020–21 there were 498 students in special schools, even though the number of students in primary schools decreased during the same period (57,574 students in 2004–05 and 51,541 students in 2020–21). It is noteworthy, that the number of students in special schools does not include students who are in other forms of segregated provision in mainstream schools (e.g., Resource Units), who are presented in a different column in **Table 1**. In addition to the above, the European Agency for Special Needs and Inclusive Education database¹² shows that in 2014–15 the Ministry reported that 583 (379 primary and 204 secondary education) students with an official decision of SEN were educated in Resources (Special) Units (separate special classes in mainstream schools). According to the latest available data, the number of these students increased to 687 (471 Primary and 226 Secondary) the following year (2016–17). These numbers do not include those students who are educated in Resource Units on a part-time basis and whose numbers have also increased in both primary and secondary schools (5498 in 2016–17 and 4603 in 2014–15). It is probably worth noting that in the last 5 years the increase of numbers of students moving to segregated settings of special education was more progressive, albeit the enhanced efforts of organized parents and disability activists to promote inclusive education, especially after the ratification of the UNCRRPD. Paradoxically, it seems that after 2014 the number of Resources (Special) Units is even more progressively increasing. This could be interpreted as an effort of the state to give a pseudo-inclusion (integration) sense to the education of children with disabilities, by providing the opportunity to be educated in the mainstream school environment, but separately from their peers and outside the mainstream classroom.

Minority rights discourses

Despite the change in nomenclature, the White Paper conveys a “minority rights” perspective that emphasizes “deviance” from the norm (e.g., students without additional support needs), rather than a “universalist approach” that focuses on modifying the social norm with a view to accommodating learner diversity (Kayess and French, 2008). In the definitions Section of the White Paper the target group for the suggested education law is defined as “vulnerable and special groups of student population”. Even though, minority rights “might protect key features of human identity” they nevertheless “possess the capacity divide people ... (and) create insiders and outsiders” (Macklem, 2008: 531), thereby perpetuating power imbalances, educational injustices and discriminatory regimes that undermine the rights-bearing subjectivities of those who have been “Othered”. For example, in the White Paper “reasonable accommodations” are defined with reference to “disproportionate treatment in relation to other children”, misinterpreting the definition in the UNCRRPD of providing “necessary and appropriate modification and adjustments not imposing a disproportionate or undue burden” (Article 2). The White Paper’s definition is based on a dis/ability ontological comparison that contributes to processes that exclude and “others” minority identity students. The process of ontological bipolarization and

¹¹ MECSY Primary Education Statistics: http://archeia.moec.gov.cy/mc/407/statistika_el.pdf.

¹² European Agency for Special Needs and Inclusive Education [Table 3: Children/learners with an official decision of SEN (european-agency.org) https://www.european-agency.org/data/cyprus/datatable-overview?year=2016_2017&field_year=2016_2017#tab-official_decision_on_sen].

the associated nomenclature of constructing “educational outcasts” such as, according to Slee (2018: 53) students with “special educational needs” or a seemingly more sensitive rendition: “additional needs” are, “bureaucratic illusions” that lack clear demarcation criteria to unequivocally determine:

At what point does need become additional or special? They are certainly not stable constructs. They are management techniques and though the packaging becomes ever-more sophisticated, the assumptions remain crude (Slee, 2018: 53).

Moreover, the phraseology of “additional support needs” promotes an indiscriminate, yet stigmatizing, approach which dissolves disability issues in education into a nebulous inclusive education bandwagon (Norwich, 2002). Disability-related needs are thus subsumed within the category of “students with additional support needs”, thereby raising concerns about the effectiveness of these broader definitions to meet students’ needs. As Norwich (2002: 493) writes, the effect of such an indiscriminate and all-encompassing approach “could be to dissolve disability issues in education into a wider more amorphous inclusive education”.

The minority rights approach is further manifested in the intention of the White Paper to introduce three tiers of intervention in mainstream schools whereby the general support is offered to “vulnerable groups” without the need to have an “official decision”. Rather than raising the need to provide quality first teaching and school wide positive behavior supports and interventions to meet the academic and social needs of all students in the first tier of intervention – that is universal and preventive in nature – (Scott, 2001; McIntosh et al., 2011; Simonsen et al., 2008), students are *a priori* pathologized and imputed a “vulnerability” status, which is unproblematically presented as an endogenous characteristic of the child. This assumption preempts that some students are “found lacking” (Kozleski, 2020: 350) and therefore incapable of benefiting from the “general model of support”; this presumption contradicts the White Paper’s rhetorical claim about the implementation of universal design for learning and differentiation as a means of mitigating inadequate student learning support (Florian, 2019).

Discursive absences and policy silences

The special education reform efforts are part of a general education and skills reform and childhood care enhancement agenda as articulated in the 2020 National Reform Program.¹³ Section 3.3.2 (pp. 42–44) refers to reforms of the educational and training system without making any reference to inclusive education reforms or issues related to disability rights in education, thereby failing to acknowledge the education, training, and counseling needs of individuals with SEND. This omission is indicative of the absence of a disability rights perspective in political and policy agendas.

Similarly, education specific disability issues are also silenced in the National Reform Program in references of accessibility and digital inclusion (p. 24). Even though accessibility explicitly refers to disability, the relevant objective of the Program solely includes the harmonization of national legislation with the EU Directive on informative Websites accessibility, while the corresponding Law (N.50(I)/2019)¹⁴ does not apply to school-based websites, nor to online educational content. Following the same pattern, the e-learning section of the MECSY¹⁵ website that was launched as a response to the Covid-19 pandemic entirely omits inclusive education and disability related issues in the guidelines and resources provided to educators, in a way that blatantly violates the right of learners with disabilities to maintain access to education.

Simultaneously, the alleged inclusive education reform efforts silence the need to proceed to a radical overhaul of the education system as a whole and to render educational professionals accountable for all students’ learning. This weakness has not been addressed by the envisaged education reform efforts, notwithstanding the recommendations of a World Bank report on the Cyprus Educational system in 2014. The report was commissioned by the World Bank in close collaboration with the Ministry of Education and Culture (MoEC) of the Republic of Cyprus and provides a number of evidence-based policy reforms that need to be pursued in order to improve teaching and to enhance accountability regimes for students’ learning (see Liasidou and Symeou 2018). Failure to address these “general policy issues” undermine, according to Norwich (2019: 2), attempts to address “the important questions in the SEN and inclusive education field”.

Hence, the articulation of a clear framework of accountability in relation to students’ learning is imperative to improve the educational outcomes of all students, including students designated as having SEND (Liasidou and Symeou, 2018). As discussed earlier, the nomenclature of “additional support needs” adopted in the White Paper also silences the role of teachers in preventing, minimizing and addressing these needs.

Inclusive education reforms necessitate a clear articulation of the role of teachers and their “curriculum and pedagogic development” (Lindsay et al., 2020: 17), with a view to creating effective and non-discriminatory learning environments for all (Alves, 2020; Florian, 2019). This perspective constitutes a *sine qua non* element in fulfilling the aims of a student-centered and rights-based education reform agenda. Schools and teachers should, therefore, be legally mandated to focus resources on developing empirically validated learning programmes and support mechanisms (Lavonen, 2018) that minimize the need for alternative educational

¹³EU 2020 Cyprus National Reform Program: https://ec.europa.eu/info/sites/info/files/2020-european-semester-national-reform-programme-cyprus_en_0.pdf.

¹⁴The Web Accessibility and Mobile Devices Application for the Public Sector Organisations Law, of 2019 (N.50(I)/2019): http://www.cylaw.org/nomoi/arith/2019_1_050.pdf.

¹⁵E-Learning, Ministry of Education, Culture, Sports and Youth: <http://elearning.schools.ac.cy/index.php/el/>.

placements and compensatory measures of support. In parallel, the focus also needs to be placed on exemplifying the role of schools and teachers in ameliorating wider social inequalities that are responsible for the poly-victimization and hence the “vulnerability” of these students.

Another related policy document of interest is entitled: “Formation of a New System for the Evaluation of the Educational Practice and Teachers” published in 2019.¹⁶ The extended document exemplifies the development of a new assessment system for both schools and teachers. The quest for inclusive education does not seem to be part of the Ministry’s vision to enhance the quality of education provided. As a result, schools’ and teachers’ proposed assessment criteria are not contingent on their ability and professional expertise to remove educational barriers that might create or exacerbate students’ “special educational needs”. The document conveys the presumption that students’ “needs” are a part of their genetic and biological makeup and this conjecture is further reinforced by the recurrent use of the term “special needs” rather than “special educational needs”. The educational dimension is not only semantically absent from the document’s reference to students’ needs, but is also absent from the overall conceptualization of the role of schools, teachers and pedagogy to create quality learning environments so as to meet the needs of learner diversity.

Despite the unequivocally pivotal role of teachers’ pedagogies, the document does not capitalize on their potential to facilitate inclusive education reforms, or preserve failed orthodoxies. Thus, teachers’ and deputy headteachers’ evaluations are solely gauged against their managerial and bureaucratic ability to deal with paperwork as they are merely expected to: “Hold the necessary information regarding possible special needs of specific students” (p. 168 for teachers and p. 188 for deputy heads). The same applies to headteachers and inspectors who, rather than being expected to “monitor the level of inclusivity of schools” – as discussed by Alves’ (2020: 76) analysis of the process of inclusive education reforms, they are merely expected to: “Create a file that includes the necessary information regarding possible special needs of specific students” (p. 207), and “Encourage and inspect the keeping of a file which includes the necessary information regarding possible special needs of specific students” (p. 223) respectively.

Thus, even though one of the specific aims of the Cyprus National Reform Program is to introduce a teacher evaluation framework, it is still acknowledged that “Teacher evaluation remains a major challenge” (CCOD-Cyprus Confederation of Organizations for the Disabled, 2020: 43). While being cognizant of the perils of uncritically resorting to the “politics of blame” that dictate that “the capacity to deliver an inclusive educational system is largely perceived to be an issue of teacher readiness or (...) a matter of improved teaching” (Done and Andrews, 2020: 450), regular teacher evaluations are imperative (World Bank, 2014, p. 6), to ensure that the needs of children with disabilities are met in effective and non-discriminatory ways.

Conclusions

The unfettered rise of special education in Cyprus is paradoxically witnessed amidst education reform efforts allegedly geared toward inclusion. The former has been semantically manipulated and embellished to provide the illusion of transformative change. The value of policy discourse analysis lies in its potential to expose the fallacies, ideological messiness, and silences of official rhetoric, while rendering policymakers accountable not only for the “validity of the educational systems they establish” (Sirotnik, 2004:155 cited in Lochmiller and Lester, 2017: 249), but also for their lack of reflexivity on how they use language to construct policy meanings and student identities (Liasidou, 2008a).

Bearing in mind the lingering legacy of special education and the patent lack of political will to challenge the educational status quo, the rise of special education can potentially be curtailed by the emerging role of disability advocacy groups and organizations in exercising “policy leadership” and creating a new “arena” wherein the “policy as a struggle” metaphor is negotiated and enacted (Fulcher, 2015). Recently, the Cyprus Confederation of Organizations for the Disabled (CCOD) and other disability organizations submitted a request to the Committee on the Rights of Persons with Disabilities to launch an “Inquiry regarding the violation of article 24 of the Convention on the Rights of Persons with Disabilities under the current education system as well as with respect to the ongoing reform for inclusive education in Cyprus” (September 2020). The Inquiry provides a poignant and substantiated analysis of ways in which the State fails to meet its legal obligations that derive from article 24 of the UNCRPD by exemplifying and documenting the ways in which students with SEND are discriminated against, stigmatized, marginalized, and excluded in education. Going even further, the organizations declared their intention to take legal action against the MECSY and its failure to exhibit “political will in the direction of complying with the Convention and the aforesaid Concluding Observations with the adoption of a true inclusive education” (CCOD-Cyprus Confederation of Organizations for the Disabled, 2020: 7).

Parents of children with special educational needs and their organizations have played a central role in leveraging collective disability advocacy efforts to launch the Inquiry. Even though their efforts to counteract disabling discourses in education policy and professional practice are still embryonic in nature (Liasidou and Hadjiyiannakou, 2019), they have, nevertheless, managed to not only foreground their hitherto marginalized voices, but also to turn them into individual and collective forms of legal action to dismantle what is seen as the tyrannical legacy of special education (CCOD, 2020). Disability advocacy for inclusive education can, nevertheless, be more powerful when meaningful relationships are built between education professionals and families (Symeonidou and Mavrou, 2019) to jointly lobby the government to change political rhetoric and professional practices. The resurgence of special education in policy discourses undermines not only the effectiveness of educational reforms for inclusive education, but also hinders educational professionals’ preparation, readiness, and motivation to implement

¹⁶Formation of a New System for the Evaluation of the Educational Practice and Teachers: <http://www.cna.org.cy/pdf/ant/06072019.pdf>.

pedagogies of inclusion (Symeonidou, 2017). As the analysis in this article suggests, the case of Cyprus constitutes another example of “patching up” the system rather than “systemic change”. As a result “(t)he development of an inclusive education system (...) continues to struggle” (Lindsay et al., 2020: 13), and will continue to struggle, unless there are concerted disability advocacy efforts and political will to problematize and dismantle education policies that masquerade, rationalize, euphemize and perpetuate the special education status quo.

References

- Acevedo, S.M., Nusbaum, E.A., 2020. Autism, neurodiversity, and inclusive education. In: Noblit, G. (Ed.), *The Oxford Research Encyclopedia of Education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.1260>.
- Allan, E.J., Tolbert, A.R., 2019. Advancing social justice with policy discourse analysis. In: Strunk, K.K., Locke, L.A. (Eds.), *Research Methods for Social Justice and Equity in Education*. Palgrave Macmillan, Cham, Switzerland, pp. 137–149.
- Alves, I., 2020. Enacting education policy reform in Portugal – the process of change and the role of teacher education for inclusion. *Eur. J. Teach. Educ.* 43 (1), 64–82. <https://doi.org/10.1080/02619768.2019.1693995>.
- Armstrong, A.C., 2013. Collaboration, partnerships and alliances: perspectives on erasmus mundus MA/magistr in special educational needs. *Int. J. Incl. Educ.* 17 (4), 364–376.
- Armstrong, T., 2017. Neurodiversity: the future of special education? *Educ. Leader* 74 (7), 10–16.
- Armstrong, D., Armstrong, A.C., Spandagou, I., 2011. Inclusion: by choice or by chance? *Int. J. Incl. Educ.* 15 (1), 29–39. <https://doi.org/10.1080/13603116.2010.496192>.
- Artiles, A.J., 2020. Inclusive education in the 21st century disruptive interventions. *Educ. Forum* 84 (4), 289–295. <https://doi.org/10.1080/00131725.2020.1831821>.
- Ball, S., 1990. *Politics and Policymaking in Education*. Routledge, London.
- Ball, S., 1993. What is policy? Texts, trajectories and toolboxes. *Discourse* 13 (2), 10–17.
- Ball, S.J., Olmedo, A., 2013. Care of the self, resistance and subjectivity under neoliberal governmentalities. *Crit. Stud. Educ.* 54 (1), 85–96. <https://doi.org/10.1080/17508487.2013.740678>.
- Barton, L., 2008. Inclusive education, teachers and the politics of possibility. In: Paper Presented at the Inclusion Festival, 21 January, Utrecht, Holland.
- Barton, L., Armstrong, F. (Eds.), 2007. *Policy, Experience and Change: Cross-Cultural Reflections on Inclusive Education*, vol. 4. Springer, Dordrecht.
- Burman, E., 2017. From subjectification to subjectivity in education policy research relationships. In: Lester, J., Lochmiller, C., Gabriel, R. (Eds.), *Discursive Perspectives on Education Policy and Implementation*. Palgrave Macmillan, Cham, pp. 65–87.
- Campbell, F.K., 2012. Stalking ableism: using disability to expose ‘abled’ narcissism. In: *Disability and Social Theory*. Palgrave Macmillan, London, pp. 212–230.
- CCOD-Cyprus Confederation of Organizations for the Disabled, 2020. Request to the Committee on the Rights of Persons with Disabilities for registration of an inquiry regarding the violation of article 24 of the Convention on the Rights of Persons with Disabilities under the current education system as well as with respect to the ongoing reform for inclusive education in Cyprus Nicosia, September 4, 2020.
- Degener, T., 2016. Disability in a human rights context. *Laws* 5 (3), 35.
- Done, E.J., Andrews, M.J., 2020. How inclusion became exclusion: policy, teachers and inclusive education. *J. Educ. Pol.* 35 (4), 447–464. <https://doi.org/10.1080/02680939.2018.1552763>.
- Dudley-Marling, C., Baker, D., 2012. The effects of market-based school reforms on students with disabilities. *Disabil. Stud. Q.* 32 (2). <https://doi.org/10.18061/dsq.v32i2>.
- Fairclough, N., 2000. Language and neo-liberalism. *Discourse Soc.* 11 (2), 147–148.
- Fairclough, N., 2001a. *Language and Power*. Longman, London.
- Fairclough, N., 2001b. The discourse of new labour: critical discourse analysis. In: Witherell, M., Taylor, S., Yates, S. (Eds.), *Discourse as Data: A Guide for Analysis*. Sage, London, pp. 229–267.
- Fairclough, N., 2016. A dialectical-relational approach to critical discourse analysis in social research. In: Wodak, R., Meyer, M. (Eds.), *Methods of Critical Discourse Analysis*, pp. 162–187.
- Florian, L., 2019. On the necessary co-existence of special and inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 691–704.
- Foucault, M., 1984. The order of discourse. In: Shapiro, M. (Ed.), *Language and Politics*. Blackwell, Great Britain, pp. 108–138.
- Fulcher, G., 2015. *Disabling Policies?: A Comparative Approach to Education Policy and Disability*. Routledge.
- Gildersleeve, R.E., Hernandez, S., 2012. Producing (im) possible peoples: policy discourse analysis, in-state resident tuition, and undocumented students in American higher education. *Int. J. Multicult. Educ.* 14 (2). <https://doi.org/10.18251/ijme.v14i2.517>.
- Giroux, H., 2012. *Education and the Crisis of Public Values*. Peter Lang, New York, NY.
- Goodley, D., Lawthorn, R., Cole, K.R., 2014. Posthuman disability studies. *Subjectivity* 7 (4), 342–361.
- Green, A., 2002. *Education, Globalisation and the Role of Comparative Research*. Professorial Lecture, Institute of Education, University of London.
- Harwood, V., Humphry, N., 2008. Taking exception: discourses of exceptionality and the invocation of the ideal. In: Gabel, S., Danforth, S. (Eds.), *Disability and the Politics of Education*. Peter Lang, New York, NY, pp. 371–383.
- Hursh, D., Henderson, J.A., 2011. Contesting global neoliberalism and creating alternative futures. *Discourse: Studies in the Cultural Politics of Education* 32 (2), 171–185. <https://doi.org/10.1080/01596306.2011.562665>.
- Kayess, R., French, P., 2008. Out of darkness into light? Introducing the convention on the rights of persons with disabilities. *Hum. Right Law Rev.* 8 (1), 1–34.
- Kozleski, E.B., 2020. Disrupting what passes as inclusive education: predicating educational equity on schools designed for all. *Educ. Forum* 84 (4), 340–355.
- Lavonen, J., 2018. Educating professional teachers in Finland through the continuous improvement of teacher education programmes, *Contemporary pedagogies in teacher education and development*, pp. 3–22.
- Lester, J.N., Lochmiller, C.R., Gabriel, R., 2017a. Exploring the intersection of education policy and discourse analysis: an introduction. *Educ. Pol. Anal. Arch.* 25 (25). <https://doi.org/10.14507/epaa.24.2971>.
- Lester, J.N., White, F.A., Lochmiller, C.R., 2017b. Language-based approaches to the study of education policy. In: Lester, J., Lochmiller, C., Gabriel, R. (Eds.), *Discursive Perspectives on Education Policy and Implementation*. Palgrave Macmillan, Cham, pp. 41–64.
- Liasidou, A., 2008a. Critical discourse analysis and inclusive educational policymaking: the power to exclude. *J. Educ. Pol.* 23 (5), 483–500. <https://doi.org/10.1080/02680930802148933>.
- Liasidou, A., 2008b. Politics of inclusive education policy-making: the case of Cyprus 4784 (excluding references). *Int. J. Incl. Educ.* 12 (3), 229–241.
- Liasidou, A., 2016. Disabling discourses and human rights law: a case study based on the implementation of the UN Convention on the Rights of People with Disabilities. *Discourse: Studies in the Cultural Politics of Education* 37 (1), 149–162.
- Liasidou, A., Antoniou, A., 2013. A special teacher for a special child?(Re) considering the role of the special education teacher within the context of an inclusive education reform agenda. *Eur. J. Spec. Needs Educ.* 28 (4), 494–506.
- Liasidou, A., Hadjiyiannakou, A., 2019. Disabling discourses and some implications for parent leadership in special education policy and practice. *Eur. J. Spec. Needs Educ.* 34 (3), 342–354.

- Liasidou, A., Symeou, L., 2018. Neoliberal versus social justice reforms in education policy and practice: discourses, politics and disability rights in education. *Crit. Stud. Educ.* 59 (2), 149–166.
- Lindsay, G., Wedell, K., Dockrell, J., January 2020. Warnock 40 Years On: the development of special educational needs since the Warnock report and implications for the future. *Front. Edu.* 4, 164.
- Lochmiller, C.R., Lester, J.N., 2017. Future directions for education policy research and language-based methods. In: Lester, J., Lochmiller, C., Gabriel, R. (Eds.), *Discursive Perspectives on Education Policy and Implementation*. Palgrave Macmillan, Cham, pp. 241–252.
- Luke, A., 2002. Beyond science and ideology critique: developments in critical discourse analysis. *Annu. Rev. Appl. Ling.* 22 (1), 96–110.
- Machin, D., Mayr, A., 2015. *How to Do Critical Discourse Analysis: A Multimodal Introduction*. Sage, London.
- Macklem, P., 2008. Minority rights in international law. *Int. J. Constitut. Law* 6 (3–4), 531–552.
- Marshall, C., Patterson, J., 2002. Confounded policies. Implementing site-based management and special education policy reforms. *Educ. Pol.* 16, 351–386.
- McIntosh, K., Bennett, J.L., Price, K., 2011. Evaluation of social and academic effects of school-wide positive behaviour support in a Canadian school district. *Except. Educ. Int.* 21 (1), 46–60.
- Namey, E., Guest, G., Thairu, L., Johnson, L., 2008. Data reduction techniques for large qualitative data sets. *Handbook for Team-based Qualitative Research* 2 (1), 137–161.
- Norwich, B., 2002. Education, inclusion and individual differences: recognizing and resolving dilemmas. *Br. Educ. Res. J.* 50 (40), 482–502.
- Norwich, B., 2010. A Response to special educational needs: a new look. In: Terzi, L. (Ed.), *Special Educational Needs: A New Look*. Continuum, London.
- Norwich, B., 2019. From the Warnock Report (1978) to an Education Framework Commission: a novel contemporary approach to educational policy making for pupils with special educational needs/disabilities. *Front. Edu.* 4, 72.
- Reisigl, M., Wodak, R., 2016. The discourse-historical approach (DHA). *Methods of Critical Discourse Studies* 3, 23–61.
- Scott, T., 2001. A schoolwide example of positive behavioral support. *J. Posit. Behav. Interv.* 3 (2), 88–94.
- Simonsen, B., Sugai, G., Negron, M., 2008. Schoolwide positive behavior supports: primary systems and practices. *Teach. Except. Child.* 40 (6), 32–40.
- Slee, R., 2018. *Inclusive Education Isn't Dead, it Just Smells Funny*. Routledge.
- Slee, R., Cook, S., 1999. The cultural politics of disability, education and the law. *Discourse* 20 (2), 267–277.
- Symeonidou, S., 2009. Trapped in our past: the price we have to pay for our cultural disability inheritance. *Int. J. Incl. Educ.* 13 (6), 565–579.
- Symeonidou, S., 2017. Teacher education for inclusion: a review of the literature. *Disabil. Soc.* 32 (3), 401–422. <https://doi.org/10.1080/09687599.2017.1298992>.
- Symeonidou, K., Mavrou, K., 2019. Problematising disabling discourses around assessment and placement: can interdependence inform an alternative narrative? *Eur. J. Spec. Needs Educ.* 35 (1), 70–84.
- Taylor, C., 2003. *Modern Social Imaginaries*. Duke University Press, Durham.
- Taylor, S., 2004. Researching educational policy and change in 'new times': using critical discourse analysis. *J. Educ. Pol.* 19 (4), 433–451.
- Titscher, S., Meyer, M., Wodak, R., Vetter, E., 2000. *Methods of Text and Discourse Analysis*. Sage, London.
- Tomlinson, S., 2012. The irresistible rise of the SEN industry. *Oxf. Rev. Educ.* 38 (3), 267–286.
- Tomlinson, S., 2017. *A Sociology of Special and Inclusive Education: Exploring the Manufacture of Inability*. Taylor & Francis Group.
- Ulmer, J., Lenhoff, S., 2017. Critical discourse and twenty-first-century education reform policy. In: Lester, J., Lochmiller, C., Gabriel, R. (Eds.), *Discursive Perspectives on Education Policy and Implementation*. Palgrave Macmillan, Cham, pp. 175–195.
- UN (United Nations), 2008. *Convention on the Rights of Persons with Disabilities* (New York).
- UNESCO (United Nations Educational, Scientific and Cultural Organization), 1994. *World conference on special needs education: Access and quality*.
- Van Dijk, T.A., 1984. *Prejudice in Discourse: An Analysis of Ethnic Prejudice in Cognition and Conversation*. J. Benjamins Co, Amsterdam.
- Van Dijk, T.A., 2016. Critical discourse studies: a socio-cognitive approach. In: Wodak, R., Meyer, M. (Eds.), *Methods of Critical Discourse Studies*. Sage, London, pp. 62–85.
- Wetherell, M., 2001. Debates in discourse research. In: Wetherell, M., Taylor, S., Yates, S.J. (Eds.), *Discourse Theory and Practice: A Reader*. Sage, Thousand Oaks, pp. 380–399.
- Wodak, R., Meyer, M., 2016. Critical discourse studies: history, agenda, theory, and methodology. In: Wodak, R., Meyer, M. (Eds.), *Methods of Critical Discourse Studies*. Sage, London, pp. 1–22.
- Wood, L., Kroger, R., 2000. *Doing discourse analysis*. Sage, London.
- World Bank, 2014. *Analysis of the Function and Structure of the Ministry of Education and Culture of the Republic of Cyprus*. Available at: <https://media.philenews.com/PDF/education.pdf>.

Further reading

- Bacchi, C., 2000. Policy as discourse: what does it mean? Where does it get us? *Discourse* 21 (1), 45–57.
- Danforth, S., 2015. Social justice and technocracy: tracing the narratives of inclusive education in the USA. *Discourse: Studies in the Cultural Politics of Education* 37 (4), 582–599. <https://doi.org/10.1080/01596306.2015.1073022>.
- Liasidou, A., 2011. Unequal power relations and inclusive education policy making: a discursive analytic approach. *Educ. Pol.* 25 (6), 887–907.

Recent debates in special & inclusive education

James M. Kauffman^a, Dimitris Anastasiou^b, Marion Felder^c, Garry Hornby^d, and Joao Lopes^e, ^a Department of Curriculum, Instruction and Special Education, University of Virginia, Afton, VA, United States; ^b Southern Illinois University Carbondale, Department of Counseling, Quantitative Methods and Special Education, Carbondale, IL, United States; ^c Inclusion and Rehabilitation, Department of Social Sciences, University of Applied Sciences, Koblenz, Germany; ^d University of Plymouth, Plymouth, United Kingdom; and ^e School of Psychology, University of Minho, Braga, Portugal

© 2023 Elsevier Ltd. All rights reserved.

Diverging approaches to disabilities	270
Disabilities and education	271
The particulars of abilities and disabilities	271
Variability among situations	271
Quantification, measurement and categorization	271
Controversy about the nature of disability	272
Disability as ordinary and extraordinary diversity	272
Disability as predicament	272
Special education as a response to disability	273
Labeling disabilities	273
Controversy about the meaning of inclusion	273
Inclusion as location	274
Inclusion as activity	274
Practical issues in inclusion: the “real world” debate	274
Inclusion and teacher diversity	274
Inclusion and the nature of disabilities	275
Inclusion and severity of disability	275
Inclusion and class size	276
Probable success of educational inclusion, all things considered	276
Confusions about inclusion	276
Definitions	276
Rights	277
Identification	277
Peers	278
Etiology	278
Intervention models	278
Goals	278
Curricula	278
Reality	278
Finance	278
Means and ends	278
Research evidence	279
Myths about special education	279
Special education isn't actually special, it's just good teaching	279
Special education does not close the achievement gap	279
Special education creates low expectations	279
Special education promotes “deficit thinking”	280
Identifying disabilities is a big problem	280
Disabilities are just diversity	280
Special education should be based on values, not science	280
Special education does not work if separate from general education	280
We had better prepare all educators to teach all kids	280
Special education teachers need to adjust teaching to kids' learning styles	280
Restructuring is the key to improving schools	280
Special education was created by educators, not parents	281
Conclusion and the way forward for special and inclusive education	281
References	281

Recent “debates” in special and inclusive education are perhaps better described as controversies. That is, rather than systematic point-by-point analyses of opposing statements that can be evaluated for logic and veracity, competing assertions have been made about the nature of disability and its implications for education, usually without rebuttal.

In actual debates, side-by-side arguments are made pro and con, and then there is opportunity for rebuttal by both sides. Judges may decide who won the debate because of better argumentation. This has never been the case in “debates” of inclusive and special education, although in some cases there have been responses to statements or even rejoinders to responses (e.g., Anastasiou and Kauffman, 2019; Anastasiou et al., 2015; Kauffman and Sasso, 2006). We do not consider ourselves or any authors cited sole proprietors of truth, or always the right or best advocates for persons with disabilities. However, we have committed ourselves to a scientific search for truths in order to align ourselves with people with disabilities in the quest for a better education. One cannot expect positive results from educational controversies when their pioneers fail to respect the view of the other side of the controversy. Perhaps, in this volume on *Inclusive Education & Disability Studies in Education*, some colleagues will view us as dissidents. Nevertheless, we think that the investigation of truth and the search for what will bring about the best lives for people with disabilities involves the investigation of the opposing views of dissidents. Bearing this in mind, we try to follow a dialogical style in this article. We present dominant points of view in *disability studies*. We address them as conjectures and then develop our refutations. We apply the term controversy rather than debate in the remainder of this article.

The most prominent recent controversies in special and inclusive education have involved two major issues: First, the concept of disability and the legitimacy of special education in response to it and, second, the nature and meanings of inclusion. Important to note here is that the United Nations’ *Convention on the Rights of Persons with Disabilities*, in its Article 24; General Comment 4 (CRPD Committee, 2016) on the right to education calls for full inclusion as a goal without defining *full inclusion* and without using the term *special education* or addressing the issue of the quality of education for students with disabilities (Anastasiou et al., 2018; Hyatt and Hornby, 2017).

We would be remiss were we not to acknowledge that some individuals adhere to philosophies abjuring science, the testing of hypotheses, making measurements, and collecting data, considering these vestiges of modern modes of thinking that are no longer applicable in postmodern or constructionist conceptual models guiding professional practice (Anastasiou and Kauffman, 2011, 2013). Contrary to constructionism-relativism that denies the very existence of reality or the possibility of approximation of truth (thereby confusing facts, uncertainties, speculations, and opinions, without reality checks) we follow a different approach. We pursue scientific truths using data and corroborated evidence, based on the scientific method and analytical thinking informed by sound, relevant theories (Blackburn, 2005; Bunge, 2017; Neiman, 2008; Rauch, 2021; Specter, 2009). This is similar to the basis for work in the “hard” sciences (e.g., physics, chemistry, biology) and in many applied sciences (e.g., medicine, nursing, agriculture). This empirical, rational, and scientific-realistic approach has been far from dominant in all the social sciences, and this applies to some of the literature on this topic.

Our perspectives on the controversies around inclusive and special education and those who share our views are described in more detail in a variety of publications. Among these many publications, including others cited in this article, are: Ahrbeck et al. (2018), Anastasiou et al. (2015, 2018), Correia (2018), Felder and Schneiders (2016), Hornby (1999, 2014, 2015), Imray and Colley (2017), Kauffman (2020), Kauffman and Farkas (2022), Kauffman and Hung (2009), Kauffman and Lopes (2007), Terzi (2010).

Diverging approaches to disabilities

Appreciating the basis for recent controversies requires understanding the history of special education. In some Western European countries and the United States of America (USA), that history goes back more than a century and grew in the context of the beginning of universal public education, whereas in some developing countries such history is much shorter. However, the controversies of recent years go to the core problem of special and inclusive education in any country’s system of universal public education.

Prefatory comments in what was likely the first textbook in special education included the statement that the problem of special education “is found in the fact of variability among children to be educated” (Horn, 1924, pp. 6–7). Educators today are faced with the problem of what human variability to recognize as critical to their work, what to call that variability, and how best to accommodate that variability.

The accommodation of disabilities should not be equated with the accommodation of cultural differences (e.g., race, ethnicity, or gender differences) for the purpose of education (Anastasiou and Kauffman, 2012; Kauffman, Burke et al., in press). Not all human variability has to do with disabilities, especially not with disabilities in learning. Mind-related disabilities represent significant limitations in learning that are best addressed by efforts to ameliorate their effects. Programs such as special education are then appropriate in helping individuals take advantage of the *right to learn*.

An alternative view, the “social model of disability,” starts with the dichotomy between impairment and disability (see Anastasiou and Kauffman, 2013). While *impairment* is viewed as a cognitive or physical dysfunction, *disability* refers to barriers created by society (Barnes, 1991; Oliver, 1996). Disabilities are therefore seen as societal failures to recognize and accept impairments, thus denying the realities of these impairments and their consequences. There are two types of issues here.

The first type is conceptual issues. Even if we remove all the disabling social, attitudinal and legal barriers that lead to discrimination, in some instances (e.g., mind-related impairments, including severe intellectual impairments, low-functioning autism,

severe communication disorders, severe psychosocial disorders, learning disabilities, post-traumatic stress disorder), the *impairments* remain and are potentially detrimental to learning (Anastasiou and Kauffman, 2013; Riddle, 2014). Thus, an initial methodological distinction between impairment and disability is elevated to an ontological distinction, that is, about the nature of being or reality (see Anastasiou and Kauffman, 2013).

The second type is social justice issues. If disability is just another cultural difference, then the *politics of recognition* or *politics of cultural difference*, that is, the struggle to remove cultural barriers, and the achievement of equitable social and political participation, might serve the interests of people with disabilities (Anastasiou and Kauffman, 2012; Anastasiou et al., 2016). This kind of politics makes sense if the barriers to educational attainment for students with disabilities are largely environmental, cultural, or legal, as it often is for individuals with body-related disabilities, including physical and sensory disabilities. For those, the barriers to educational attainment for students with disabilities are largely physical rather than mind-related, and *equality as inclusion* then makes sense (Anastasiou et al., 2018). However, when the barriers to educational attainment are deeply intertwined with the neural architecture of learning, as often is the case for individuals with mind-related impairments, then meaningful educational opportunity will likely require much more than the individual's physical presence in the same classroom as children who are not impaired (Anastasiou et al., 2018). In these cases, *social justice* as the *politics of redistribution* requires additional resources to compensate for learning disadvantage. These additional resources include special education and expansion of meaningful economic and social opportunities (Anastasiou et al., 2018, 2016).

Disabilities and education

Universal public education requires attention to ways in which human beings vary from one another and deciding how this variability is relevant for particular groupings of individuals to be taught. Bateman (1994) has suggested that determining the who, how, and where are perpetual issues for special education. Human variability takes many forms, with the variability we call disability having particular implications for education.

Part of the problem contributing to recent controversies is failure to consider the meaning of a particular variability or diversity among students for their education. That is, people sometimes confuse or fail to understand the differences among differences, the diversity in diversities. In the USA, for example, skin hue or ancestry has sometimes been confused with disability in arguments for educational placement of students with disabilities, leading to the untenable conclusion that the USA's Supreme Court's statement that "separate is inherently unequal" applies to disabilities as well as to skin color (e.g., Stainback and Stainback, 1991). Clearly, skin color and disability are very different with quite different implications in the matter of education (Kauffman and Badar, 2020; Kauffman and Landrum, 2009).

The particulars of abilities and disabilities

Another problem contributing to recent controversies is that particular abilities and disabilities have extremely different implications for education. To figure out what must be taught and how, and what are reasonable expectations, the specifics of disabilities, such as cognitive and sensory impairments must be considered. Adding to this problem is that special education includes hyperabilities (typically called gifts and talents) (Callahan et al., 2020). Moreover, twice-exceptional individuals, those with both a disability and a special gift or talent, must also be considered. Thus, the observation that someone has a hyper-ability or a disability or both has little meaning for education until the specifics of ability and disability are considered.

Variability among situations

Systems and traditions of public education vary enormously from country to country. Considerable variation exists in the grouping of students for various activities, services offered in and through public education, typical class size, ethnic and language homogeneity, and expectations of outcomes. Controversies in special and inclusive education require understanding at least five dimensions of difference: (a) the relevance of particular variations or diversities of students to their education or schooling, (b) how education or schooling is different from other community activities, (c) the particulars of abilities and disabilities of individuals, (d) the diversity of teachers' skills and demands or situations, and (e) the particulars of a country's system of education. Recognition of these dimensions of difference require some system of measurement or categorization.

Quantification, measurement and categorization

Some disabilities are judged comparatively easily (i.e., compared to other disabilities) and relatively objectively (i.e., compared to other measures). However, no matter the measure (e.g., of visual acuity, hearing loss, academic achievement, language acquisition, intelligence), the actual criterion for defining disability is arbitrary. That is, people decide what to measure, how to measure it, and the significance of the findings, and they can change the measurement criterion (e.g., for defining such things as intellectual disability, visual acuity, hearing loss, or low birth weight) basically because the quantification of qualities involves boundaries, and these boundaries are typically blurred.

Diversity or difference needs measurement, even if that measurement is simply binary, for example yes/no, or on/off. The only way we can get rid of scientific categories is to have no quantification and measurement at all. This would mean a return to

prescientific and hugely stigmatized concepts (e.g., laziness, foolishness, stupidity). This may satisfy romantic anti-quantification sentiments, but it would work against the interests of people with disabilities. Measurement often involves a continuous variable, and when it does disability or other diversity is defined by a cut-off point in a statistical distribution of measurement values (Kauffman and Lloyd, 2017). Furthermore, as we have noted, all measurement, even if it is a binary judgment, results in categorization (e.g., does/does not have disability).

Controversy about the nature of disability

The way one views disability has important implications for efforts to deal with it. Of course, it is philosophically possible to deny the existence of disability, to consider it a fake category simply made up to oppress those so categorized, rendering special education superfluous. For example, Baglieri et al. (2011) wrote:

... it is within special education that cognitive and biological ideologies of normalcy and abnormalcy are codified and exercised—even championed for relentless methods of identifying pathology with schoolchildren in order to deliver “appropriate services” (p. 2129).

This provided a rationale for attacking special education, as they then stated, “what special education strives to enforce, Disability Studies seeks to dissolve” (Baglieri et al., 2011, p. 2930).

However, the conceptualization of disability as resulting from social oppression and nothing else—comes to a logical dead end. If disability is defined by social exclusion, or discrimination, then all victims of racial and gender discrimination would count as disabled including women, blacks, and gays, which would be an illogical conclusion (Anastasiou and Kauffman, 2012; Harris, 2000). An alternative is that the category “disability” does, indeed, exist and that some students with disabilities do, actually, have significant limitations in learning. Therefore, a special educational mechanism (i.e., special education) is a rational and needed response to these disabilities (e.g., Kauffman, 2022; Kauffman et al., 2018b).

Disability as ordinary and extraordinary diversity

Some people see disabilities as ordinary diversities or even idealize them, the “problem” being society’s unwillingness to address disabilities as ordinary differences that need to be accommodated. Disabilities themselves are not problems at all, according to this view, which emphasizes the “lived experience” of disability and claims that disability is an integral part of personhood and that a person with a disability loses identity if and when disability is lost.

Others view people with disabilities as valuable human beings, but see their specific disabilities as both extraordinary activity restrictions, not things to be considered gifts. Disabilities are restrictions in ability to perform a critical activity, such as difficulty walking, seeing, hearing, communicating, learning or problem solving (Anastasiou and Kauffman, 2012, 2013). Typically, disabilities generate special needs that should be addressed in any caring society if we want to maximize learning for students with disabilities and overcome their participation restrictions.

In this view, individuals may have characteristics that need modification. The implication of this viewpoint is that the attitude or philosophical notion that “there is nothing wrong with disabilities, it’s the rest of the world that needs changing” is a half-truth in many cases and sorely lacking in nuance when applied to all individuals with disabilities. “Normal” or “typical” is a meaningful concept, and depending on an individual’s characteristics and the capability of societal institutions to change either themselves or the individual, it will be unfair to deny the appropriate change of either. Moreover, a view that disability is never an extraordinary challenge fails to recognize the perspective of ecological psychology—that it is interactions and transactions of individuals and environments, not simply or completely a fault in one or the other, that creates a disability (e.g., see Anastasiou and Kauffman, 2013).

Those who see disabilities as extraordinary human diversity also recognize the importance of prevention, conditions that could be prevented from occurring at all if possible and prevented from progressing or worsening. Thus, prevention of accidents and diseases that cause disabilities is not a bad idea. Neither are medication, surgery, therapies, treatments, devices, and special education designed to mitigate disabilities.

Disability as predicament

Some people see disabilities as predicaments, perhaps natural parts of life but, nevertheless, not good situations to be in. Disabilities are differences from typical functioning. They definitely are not imaginary. Their recognition denotes that others care about educational performance that puts a student in a poor predicament or bad situation. Characteristics that create predicaments in developing persons are not experienced as just another kind of social identity among developing persons.

Every person modifies her or his learning behavior in school, and students with disabilities are not exceptions—they go to school to learn, not just to spend time there. Moreover, the right to learn applies to all students not just students with disabilities. We also know that many people with disabilities experience their atypical restrictions of abilities as predicaments that make life harder (Shakespeare, 2006), and this includes students with disabilities.

Special education as a response to disability

From one point of view, special education adds insult to injury—in that it is a way of needlessly identifying, separating, and ill-serving students with disabilities. From this perspective, students' disabilities are assumed not to generate special needs that should be remediated (see Pfeiffer, 2002), merely ordinary differences to be accommodated by all teachers, diversities that, like any others, are not in themselves challenges in any way. Presumably, then, students with disabilities are to be treated like all other students and, without exception, educated in the same place, in the same curriculum, and by the same methods as all other students. People with such views are vehemently opposed to special education as a response to disability. For example, as Oliver (2000), a central figure in the field of disability studies, put it, "I was and remain implacably opposed to the very existence of special education" (p. 6). In support of this view by lambasting special education as something not needed in a society properly reformed, Lipsky and Gartner (1996) wrote:

Special education plays a sorting role, both for those consigned to it and for those students who remain in general education. It limits expectations of the former, and gnarls the attitudes of the latter ... Thus, the system of special education, and the attitudes towards disability that undergird it, have harmful consequences for both those labeled "disabled" and those not.

Lipsky and Gartner (1996, pp. 767–768).

From another point of view, special education has been historically, and continues to be, necessary and effective in meeting the needs of students with disabilities, as demonstrated by recent studies (Hurwitz et al., 2019; Schwartz et al., 2019). Accordingly, it does not do the damage Lipsky and Gartner (1996) noted, and the attitudes toward disability undergirding it are helpful, not harmful (see Gerber, 2017; Hornby, 2014, 2015; Kauffman, 1981, 2022; Kauffman and Badar, 2018, 2020; Kauffman et al., 2018b; Martin, 2013; Warnock, 2010).

Labeling disabilities

Much of the criticism of special education has to do with labeling, particularly the negative implications of labels and their accompanying stigma. Concerns about labeling and stigma are not new phenomena but have been controversial for decades, and not just depending on one's view of disability. The issue of labeling has been discussed extensively for years by the first author and colleagues (e.g., Kauffman, 1989; Kauffman and Badar, 2013; Kauffman et al., 2008). Suffice it to say here that efforts to avoid labels have not worked out well when tried (e.g., Feniak, 1988). Without labels one cannot talk about things and they become unmentionable. Labels are inevitably attached to persons receiving special services, and progress in all fields of study is characterized by more specific labeling because finer distinctions are made.

Attempts to avoid labels or to use only more general, less specific labels, have not been found to be practicable.

The labels may not be those in current use for specific disabilities, but a label is necessary to designate the matter of concern. In our view of language, avoidance of labels merely makes the matter of concern unmentionable—or, ironically, all the more visible. Some have attempted to avoid labels by suggesting, for example, that "We do not believe a person *has* an intellectual disability; rather the person is *defined* by other as having the condition" (Kliever et al., 2006, p. 188). Still another tactic in attempts to avoid labels for disabilities or make them less unpleasant is using general, less specific language. Thus, the more general term developmental disability, which lumps disabilities of all kinds together, is perceived by some as more appropriate As a field of study matures, its language gets more specific, not less. Less specific language or more general terms without an increase in more specific sub-terms is a reliable indication of regression, not advances, in any field of work.

Kauffman et al. (2008, p. 363).

Intention to reduce erroneous labeling and stigma is not misguided, although efforts to do so, such as coming up with new labels or using words that apply to everyone, may not work as intended (Kauffman and Badar, 2013). Much of the anti-labeling rhetoric in education has been counterproductive. Labels need not be changed to provide better understanding of their meaning and of the individuals to whom they accurately apply. A better understanding of differences about which we can speak openly and candidly is important, and so is communication with labels people understand. Understanding that a new label means the same as the old or former label does little to reduce the stigma of the condition named. A label can clarify communication or distort it.

Controversy about the meaning of inclusion

Virtually no one believes that inclusion of diverse groups of people should not be included in a diversity of community activities. However, two questions about inclusion arise in the minds of most people: Which diversities? For a given diversity, which activities? Virtually no one suggests that inclusion of diverse groups of students should not be included in public education. The question for most people is which diversities to include. Furthermore, virtually no educators hold the view that no students with disabilities should be included in general education. The question for most educators is whether there are objective limits on inclusion of students with disabilities in general education and, if so, just what those limits should be and the conditions under which they should apply.

Ultimately, there is also the question of just what inclusion means in any activity. Inclusion could mean just being there—that physical presence is enough, that presence or exposure to something is access to it. It could be that more than just being there is important, that actual participation is required. Obviously, one cannot participate without being there, so being there is a precondition for participation. However, not all people have the ability to participate in activities that most people enjoy, and, in education, presence and exposure do not necessarily mean access (Fuchs et al., 2014).

Inclusion as location

One way of being able to tell whether a student is included in general education might be called *habeas corpus*, a “we have the body” definition of inclusion (Kauffman and Badar, 2020). This locational notion has the advantage of being literal, easily and objectively measured (is/is not in a designated classroom or location for x% of the school day), and a precondition for any other type of inclusion. Typically, any other place, even if temporary or a percentage of the school day, is called “segregated.” The “segregated” definition is applied to special education, although not typically used to describe other separation or instruction in a setting designated for particular purposes (e.g., band or music education, practice of a sport, science lab), even though these might be activities or places not including all students. We do not usually consider students to be segregated by bus route, school, or grade, even though they may be in different places depending on their residence, age, and achievement. This is partly because we understand that not all students can be included in all things and that there are both rational and irrational reasons for a student’s inclusion in any given place.

Nonetheless, some proponents of full inclusion argue that *all* students with disabilities can be appropriately taught in a single kind of place, that being the school and classroom they would attend if they had no disability. For these proponents, place is not only the central issue but the assumption is that appropriate education can and will follow placement in general education. It is assumed that we now know how to bring all educational services to students in general education and that this should be an invariant assumption in all cases (e.g., Laski, 1991; SWIFT Schools, 2020). For example, Blackman (1992) wrote, “‘Place’ is the issue ... what is being challenged is the location where ... supports are being provided” (p. 29).

Inclusion as activity

Another definition of inclusion has to do with the activity in which a student is engaged, something Kauffman and Badar (2020) call *proprium instructio*, meaning in Latin appropriate instruction. Warnock (2010), an early advocate of inclusive education, notes that students with disabilities may be instructed appropriately in a variety of settings, that engagement in appropriate instruction is more important than place of instruction and that appropriate instruction is for some children far more feasible and likely in some places than in others.

For those doubting the wisdom of *full* inclusion the primary objective is a student’s engagement in learning that is most appropriate for her or him as an individual, regardless of place. That is, the most important issue is instruction, not place (Kauffman and Badar, 2014, 2016; Kauffman et al., 2018a,b,c).

Putting instruction before place—chronologically, preparing the individualized educational program or IEP before choosing the least restrictive environment (LRE) from a full continuum of alternative placements (CAP) is not just logical but is mandated by federal law in the USA (see Bateman, 2007, 2017; Bateman and Linden, 2012).

Practical issues in inclusion: the “real world” debate

“Real world” includes policies, whether we approve of them or not. Nevertheless, the “real world” includes other realities, whether we approve of them or not. We prefer what Rauch (2021) called “the reality-based community” and what Kauffman and Farkas (2022) refer to as veracity in special education (Kauffman and Hornby, 2020; *in press*). Usually, policies can be changed more easily than can other realities. Moreover, policies can require ignoring or denying other realities. For example, a policy might call for *all* people in a political entity (such as a country or city) to be fed a nutritious diet in the absence of sufficient food to do so. In special education, we must recognize the fact that a policy may mandate the impossible because of human limitations or the nature of teaching and learning. It is also important to recognize that policies requiring people to do the impossible, or even the nearly impossible or highly unlikely for those to whom it applies, generate a variety of negative consequences, including resentment, resistance, hostility, pretense, lying, cheating, and cover-ups.

Inclusion and teacher diversity

Teachers vary along many dimensions of competence in instruction and classroom management, so if education policies are not to demand the impossible, they must consider what the typical teacher is realistically able to do. Despite promotion of inclusive education for more than 20 years, many general education classroom teachers still consider that they are unprepared to teach students with various disabilities (Hornby, 1999; Mitchell, 2019).

Training in special education has not, nor can it reasonably be expected to, prepare all of its trainees to be competent in instruction of students with all types of disabilities. Much more unrealistic is the notion of training all teachers to teach all students well.

Kauffman et al. (2020b) pointed out that preparing all teachers to teach all students makes as much sense as: all drivers being licensed to drive all vehicles, without special training; all physicians being licensed to perform all medical treatments, including examinations, prescriptions, surgeries, and other medical procedures; all patients being placed in general medical units regardless of medical condition or diagnosis; all lawyers being expected to handle all types of legal cases; all sport coaches being able to coach any sport; all teachers being prepared to teach all subjects at all levels.

Although these examples may be considered hypothetical, it should be pointed out that the Iowa Professional Teaching Practices Commission proposed, in 2014, to establish a single special education endorsement that would be applied to all levels of instruction (K-12), all levels of severity, and all disability conditions. A similar proposal put forward for a single endorsement in general education that would qualify an individual to teach anything from early childhood education to high school physics would be considered laughably outrageous! ... Such a proposal is not only unconscionable but reveals the devaluation of what special educators do and what students with disabilities deserve.

Kauffman et al. (2020b, pp. 258–259).

In the “real world” of teaching and teacher education, the need for specialization is readily recognized for most subjects of study, particularly as grade and age of student increase (e.g., music, science, math, art). The notion that special education teachers are not required for meeting the instructional needs of students with disabilities might not be merely an indication of how teacher diversity is ignored but also how little the nature of disability is acknowledged.

Inclusion and the nature of disabilities

We have previously mentioned the diversity of conditions that comprise disability—that the singular term disability can mask the diversity of conditions involved. Disabilities can involve seeing, hearing, talking, thinking, moving, emotional/behavioral health involving self and relationships with others, and other biological/organ functioning. Disabilities can involve these things in varying degrees and various combinations. Thus, disabilities differ greatly in nature and implications for teaching and learning. Among the many issues involving disabilities is what students with disabilities deserve in their education, and afterward. One must consider not only students with disabilities’ lives during their education, but also their education as preparation for life after they leave public education.

One point of view is that the most important thing for all students with disabilities is to be part of general education, not to be taught in environments or settings different from that of their age peers who are not identified as having a disability. Hence, their disability should not set them apart from other students but be seen as just another variation of human existence, demanding only accommodation to make their participation in the classroom a reality.

A different perspective is that some disabilities require acknowledgment of the need for special instruction, that such instruction may require a separate setting and special teacher, and that not all things can be taught well in the same place and by the same teacher. Moreover, students with disabilities deserve instruction from a teacher especially knowledgeable about their particular disability and skilled in teaching students with it. For example, blind students need a teacher highly skilled in teaching orientation and mobility as well as Braille, deaf students need a teacher fluent in signing, deaf-blind children need a teacher highly competent in teaching them, students with speech/language problems need a teacher highly skilled in working with speech/language deficits, and so on. From this point of view, instruction is more important than being in a class with nondisabled peers, and learning particular skills in school is highly important for students’ lives after school. Failure to become competent in the special skills required for functioning as independently as possible after the school years sets up students with disabilities for their failure of inclusion in community settings when they are adults.

Inclusion and severity of disability

Severity of disability is often a relevant variable in the matter of inclusion’s feasibility and outcomes, with greater severity of disability typically associated with lower feasibility of successful inclusion in general education. However, advocates for the inclusion of students with more severe disabilities often bemoan the fact that such students are not included as often as they should be or as often as students with other disabilities, such as learning disabilities (e.g., Agran et al., 2020; Wehmeyer et al., 2020). In trying to explain why this is inappropriate, Kauffman et al. (2020c) considered several factors related to severity. First, some students have multiple disabilities, none of which is alone severe, but the combination of conditions creates severe difficulty in learning. Second, another factor to consider is what parents, educators, and students themselves expect as the outcome of education. In some cases, it is economically and administratively convenient to place the student in a general education classroom with relatively low expectations of the student’s learning. That is, just where the student is for a given percentage of time makes little difference in what the student can do, so being in a general education classroom is of little consequence, and this might be the case for both the students with disabilities and other members of the class. In other cases, a student’s severe disability includes extremely disruptive behavior that does interfere with other students’ learning. Moreover, in many cases students with severe disabilities need to learn a great deal that is important in self-care and ability to function in the community following school, and they are very unlikely to learn these things in mainstream classrooms.

Therefore, making sweeping statements about the feasibility and desirability of the inclusion of students with severe disabilities in general education is irresponsible, all things considered. Perhaps this is especially true in the case of severe and multiple disabilities. Likely, this is one of the reasons for the USA's Supreme Court's conclusion that educators should expect that a students with disabilities should be making progress toward achieving challenging goals, rather than just being included in general education (see [Kauffman et al., 2019](#)).

Inclusion and class size

Teaching all students well, regardless of disability is perhaps possible for the typical teacher with appropriate training in general and special education when the class consists of 5 or 10 students. However, it is far more difficult for any teacher who has training in special as well a general education when the class has 20 or 30 or more students. It is extremely demanding for the most highly skilled and best-prepared teachers of more than 30 students. Probably, the typical general education class size needs to be 20 or more students for the wealthiest nations to make universal public education sustainable without massive increases in education funding. Therefore, implementing full inclusion of students with disabilities is an extraordinary task if they are to be well served. The practical, real-world problem is meeting the challenge of appropriate education for all, including not just typical human diversity (most of which has little or no implication for education) but the special and often unique learning needs of students with disabilities.

In all cases, the larger and the more learning-diverse the group of students, the more difficult the job of the teacher. In most countries, few public school classes have fewer than 20 students, and in many developing countries the typical public education class size is more than 30 students, sometimes reaching 100 students or more, making inclusion of students with disabilities extremely difficult if not practically impossible if students with disabilities are to receive appropriate instruction.

Probable success of educational inclusion, all things considered

Success in education is a relatively difficult thing to define, whether for students with disabilities or students not identified as having disabilities. However, success is usually defined more by what students learn and do in school and how they fare after school than where they go to school or in which classrooms in a particular school building or with what teachers they spend their time. When we make decisions about our own or other individuals' lives, we need to try to figure out the probabilities of both immediate and long-term consequences.

Science is our best bet for estimating short- and long-term probabilities of a given course of action ([Kauffman and Farkas, 2022](#)), and sometimes the probable long-term consequences are the more important. The Covid-19 pandemic of 2020 put this into sharp relief for the welfare and safety of societies in general and school children in particular ([Kauffman and Badar, 2023](#)). The world-wide movement toward full inclusion illustrates the problem of immediate versus long-term consequences in special and inclusive education ([Kauffman et al., 2020a,b](#)).

As we understand success in school, it means that students learn a curriculum that is judged appropriate for them. If we define success in that manner, then we suspect that the probability of success for students with disabilities declines as the teacher's competence decreases, the number of students in the class increases, and as the severity of disability increases. That is, success for students with disabilities is very likely positively correlated with teacher competence and negatively correlated with class size and severity of disability.

Besides our hypotheses regarding school success of students with disabilities, we think inclusive and special education suffers because of many confusions about inclusive education and myths about special education. That is, special and inclusive education is often misunderstood because of lack of clarity, overconfidence in fast, intuitive, and emotional thinking, and failure to think deliberately about it and the evidence that supports and refutes it ([Kahneman, 2011](#)).

Confusions about inclusion

Common confusions about inclusion are apparent in the literature—confusion about definitions, rights, identification, peers, etiology, intervention models, goals, curricula, reality, finance, means and ends, and research evidence.

The 12 confusions depicted in [Fig. 1](#) are briefly outlined below with further details provided in [Hornby \(2014, 2015\)](#).

Definitions

The term *inclusion* is used in various ways, for example, to refer to inclusive schools or an inclusive society. There is also confusion of social inclusion with inclusive education for children with disabilities. Social inclusion in education refers to inclusion in mainstream schools of children with a wide diversity of differences, such as race, sexual orientation, and religions. But disabilities are a different kind of diversity requiring a totally different response from the education system. Nevertheless, proponents of full inclusion confuse disabilities with other forms of diversity, treating diversities as if they all mean the same thing as far as the students with disabilities' education is concerned.

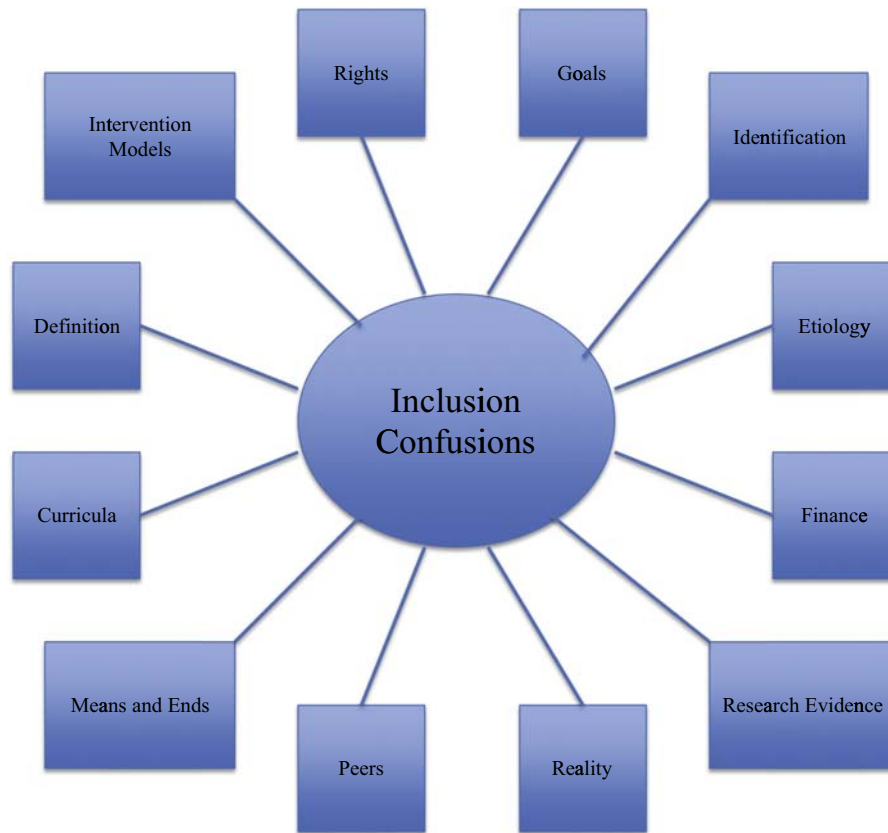


Fig. 1 Confusions about inclusion. Based on Hornby, G., 2014. *Inclusive Special Education: Evidence-Based Practices for Children with Special Needs and Disabilities*. New York: Springer; Hornby, G., 2015. Inclusive special education: development of a new theory for the education of children with special educational needs and disabilities. *Br. J. Spec. Educ.* 42, 234–256.

Rights

The committee that monitors the implementation of the UNCRPD has radically reinterpreted Article 24's *right to education* as a *right to inclusive education* (Anastasiou et al., 2018). In implementation of the requirements of UNCRPD Article 24 General Comment Number 4, the Committee de-emphasizes the importance of learning, avoiding reference to specially designed instruction, special education teachers, and even co-teaching (Hyatt and Hornby, 2017; UNCRPD, 2016). In addition, it weaponizes school inclusion against special education settings (Anastasiou et al., 2020). However, the right to education takes on its full meaning, only if it includes a *right to learn*, and the right to learn is not the same as the right to educate in the same environment (Anastasiou et al., 2018; Warnock, 2010). Furthermore, in all international law treaties, the right to education is considered a *human social right* aimed at ensuring socio-economic justice among all persons with disabilities, including those with mind-related disabilities. The anti-discrimination paradigm, which the UNCRPD Committee promotes, is inspired by the tradition of individual rights (e.g., right to life, equal recognition before the law), prioritizing sameness of educational treatment for people with disabilities. Thus, although their human rights allow students with disabilities to be educated alongside their mainstream peers, for some of them this may not, morally, be the best option. As Warnock puts it, "What is a manifest good in society, and what it is my right to have ... may not be what is best for me as a schoolchild" (2010, p. 36). Placement in a general education classroom is not enough to do justice to persons whose impairments cannot be addressed by mere access to education (Anastasiou and Bantekas, in press; Anastasiou et al., 2018).

Identification

To supporters of the inclusion of all students in general education, with no exceptions, identification can result in labeling children with disabilities, thereby stigmatizing them, and therefore should be avoided. However, if children *are* identified as having disabilities, there is a risk of negative labeling, while if they are *not* identified there is a risk that their special needs will not be met. Also, being stigmatized is not necessarily a result of identification but the fact that their disabilities themselves mark them out as different from other children in some way (Kauffman, 1989, 2005, 2008; Kauffman and Badar, 2013).

Peers

Another confusion is related to one of the proposed hallmarks of inclusive education, that children with disabilities are educated alongside their age peers in mainstream classrooms. However, many children with disabilities are more comfortable alongside peers with shared interests and similar abilities or disabilities to themselves, than alongside peers who do not have similar disabilities and merely have the same chronological age. As Warnock pointed out, "Inclusion is not a matter of where you are geographically, but where you feel you belong" (2010, p. 34).

Etiology

Over the past few decades, awareness has grown about how social and environmental factors can influence children's development and functioning. However, some proponents of full inclusion have taken this social perspective to its extreme to suggest that disabilities are entirely socially constructed and to deny the impact that impairments can have on children's learning (Hughes and Patterson, 1997; Oliver, 1992; see also Anastasiou and Kauffman, 2011).

Intervention models

It is sometimes suggested that special education is based solely on a medical or deficit model. In fact, special education interventions are influenced by educational, psychological, sociological, sometimes by medical approaches as well as other not easily classified treatment models with a particular focus on evidence-based practices as well as law and social policy. No matter the theoretical starting point (e.g., applied behavioral analysis, cognitivism, Piagetian or Vygotskian approaches, information processing theory, social learning theory, systems theory), it is the research evidence within a context of human dignity and continuous monitoring, that provide an objective way to evaluate intervention approaches (Wiley et al., 2019). This is an important aspect of special education because of the persistence of a plethora of controversial interventions, such as gentle teaching, learning styles, facilitated communication and other interventions lacking support of research evidence (Fox and Mulick, 2016).

Goals

Including students with disabilities in mainstream schools driven by the need to achieve high predetermined academic standards (e.g., age-level Common Core Standards) results in the goals of education for many of these children being inappropriate. Instead, the main goals of education for students with disabilities should be tailored to their special needs to "... assist them to become economically active and provide them with the skills needed in everyday life, offering training in skills which respond to the social and communication demands and expectations of adult life." (UNESCO, 1994, p. 10). This becomes even more an issue with students who have severe disabilities (Kauffman et al., 2020c).

Curricula

Supporters of full inclusion demand that *all* children with disabilities should be entitled to have access to the same curriculum as other children. However, inclusion in a mainly academic curriculum that is unsuitable for the learning needs of many students with disabilities may contribute to the development of emotional or behavioral difficulties that can lead to their being disruptive and being excluded from schools. The priority for children with disabilities must be that they have access to curricula that are appropriate for them, not that they are fitted into a curriculum that was designed for the mainstream population (Fuchs et al., 2014; Kauffman et al., 2020c, 2019).

Reality

Full inclusion requires that *all* children with disabilities be educated in mainstream classrooms. However, the reality of the situation in most mainstream schools is that, because of insufficient input on teaching students with disabilities in initial teacher education courses, limited in-service training, insufficient material and financial resources, and insufficient support staff, many teachers do not feel able to effectively implement even a policy of temperate inclusion (Farrell, 2010; Mitchell, 2019).

Finance

A key controversy is about the relative cost of provision for students with disabilities in mainstream or special facilities. At first sight special schools, special classes and resource rooms appear more expensive, so inclusive education seems to be the cheaper option (OECD, 1999). However, if the education system does not enable young people with disabilities to achieve independence after they leave school, then the cost to society will be far higher in long-term welfare payments.

Means and ends

An important confusion is whether inclusion is a means to an end or an end in itself. Proponents of full inclusion argue that any separate placement is morally wrong because a key goal of education should be to fully include individuals in the community in which they live. Therefore, the argument goes, they ought to be included in their local mainstream schools. However, inclusion in the community after leaving school is the most important end that educators should be seeking. Inclusion in mainstream schools may be a means to that end but should not be an end in itself. A dedicated, separate setting is also a means to that end, not an end in itself.

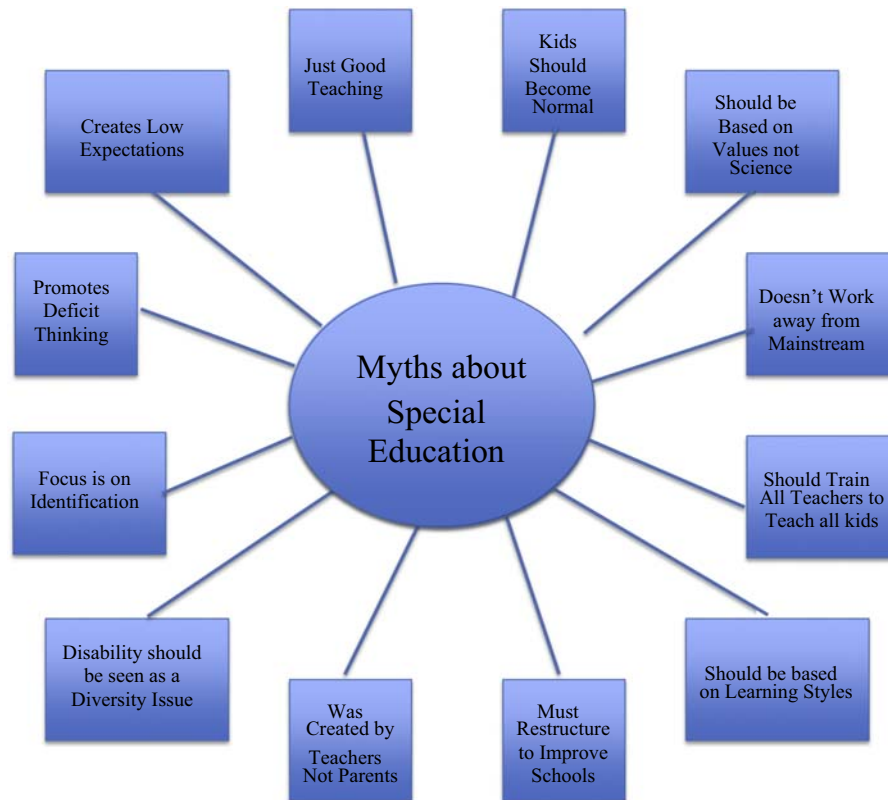


Fig. 2 Myths about special education. Based on Kauffman, J.M., Schumaker, J., Badar, J., Hallenbeck, B.A., 2018. Where special education goes to die. *Exceptionality*, 27, 149–166.

Research evidence

There is confusion about the research base for inclusive education, with many supporters of inclusion appearing to believe that an adequate research base for the effectiveness of inclusive schools already exists. However, reviews of the research evidence in support of inclusion to date have been inconclusive, suggesting that an adequate research base for inclusion has not been established (Cook and Cook, 2020; Farrell, 2010; Fuchs et al., 2022; Kauffman and Hornby, in press; Lindsay, 2007).

Myths about special education

The possible future of special education is not entirely clear, and some have speculated about how it might be devalued or even die (e.g., Hyatt and Hornby, 2017; Kauffman, 2022; Silvestri and Heward, 2016). Kauffman et al. (2018c) have suggested 12 common myths that may contribute to its demise. The 12 myths discussed briefly below are depicted in Fig. 2.

Special education isn't actually special, it's just good teaching

We have previously noted this myth in our discussion of specialization. The idea that teaching is teaching—that if you can teach one student you can teach any student, or that what you can teach in one setting you can teach in any setting, or that if you can teach one thing you can teach anything—reflects a profound misunderstanding and oversimplification of teaching and learning. Special educators need to have special teaching skills (Kauffman et al., in press).

Special education does not close the achievement gap

As we have long known, the goal of closing the gap between the mean achievement of those with and those without disabilities is irrational and demonstrates misunderstanding of the implications of educational disability (Kauffman, 2005). The gap that can and should be closed is that between what students with disabilities are capable of achieving with support and what they do in fact achieve (Kauffman, 2008). Special education is designed to close that and only that gap.

Special education creates low expectations

Regrettably, many people have lower expectations of students with disabilities than those students can achieve, sometimes because of bias, prejudice, and malicious ableism. Many people have unreasonably high expectations of students with disabilities. Getting expectations just right for individual students is not easy (Kauffman and Landrum, 2018). In short, there is no evidence that special

education creates low expectations, although reasonable expectations for students with disabilities may well be lower than those for students without disabilities.

Special education promotes “deficit thinking”

Deficit thinking rhetoric implies that special educators, or even general educators, tend to see only deficits in students with disabilities. Such rhetoric in disability studies (e.g., [Dinishak, 2016](#)) is a favorite linguistic tactic used to neutralize the specific characteristics of disability and neglect the need for dedicated care of those needing special education services ([Anastasiou and Kauffman, 2019](#)). The IDEA definitions of disability categories have a standardized structure. They first describe an extraordinary human condition in terms of impairment, disability, low functioning, deficit or disorder, and then emphasize that these conditions adversely affect a child’s educational performance ([Anastasiou and Kauffman, 2019](#)). In education—contrary to medicine—disability itself is not the special educator’s interest. *If the condition does not affect a child’s educational performance, then it is not a disability for special education purposes.* This is an educational-legal model tailored to school learning. Deficit thinking rhetoric manipulates language to attack the legally defined identification process of students with disabilities ([Anastasiou and Kauffman, 2019](#)). When students with disabilities do have impairments in learning, then and only then should educators have a special system to support them, and that system is special education. Denying special educational needs or denying what students with disabilities can achieve under optimal instruction is not helpful. What is helpful is seeing realities, avoiding the traps of overestimation and underestimation of skills, finding students with disabilities’ strengths that help them reduce any possible learning deficit or gap, and enabling them to do more things that promote their independence and welfare. One’s specific ability to do more things does not make a person more valuable as a human being; all persons have equal value. Malicious ableism—underestimating the value or abilities of students with disabilities—should be fought within school and community.

Identifying disabilities is a big problem

Identifying disabilities is sometimes difficult, but identification is not the biggest problem. In fact, denial of disability may be more common than difficulty in identification, especially when the disability is not obvious. The biggest problem is making certain that all students with disabilities receive appropriate instruction ([Kauffman et al., 2018c](#)).

Disabilities are just diversity

Disabilities are, indeed, a kind of diversity. However, they are not like any other kind of diversity when it comes to education. It is most obviously different from diversities of color, nation of origin, culture, heritage, size, age, religion, sexual orientation, and many other diversities requiring different responses from educators ([Anastasiou and Kauffman, 2012](#)).

Special education should be based on values, not science

Values are, indeed, very important. Valuing individuals for whom they are and valuing abilities are essential in education. However, in the matter of instruction, scientific evidence must supplement and direct the value of educators for learning ([Kauffman and Farkas, 2022](#); [Wiley et al., 2019](#)).

Special education does not work if separate from general education

Separation of students with disabilities from general education does not prevent special education from “working” in the sense of helping students with disabilities learn appropriate skills. Separate education can only be said not to “work” in that it does not fit the inclusionist model of locating the students with disabilities in general education. For some students with disabilities, special education offered in the general education classroom does not “work” in the sense of being the place the students with disabilities will be most successful in learning the appropriate curriculum ([Kauffman and Hornby, in press](#); [Fuchs et al., 2022](#); [Kauffman et al., 2018c](#)).

We had better prepare all educators to teach all kids

Preparing all educators to teach all students is similar to the idea of preparing all members of any profession or line of specialized work to perform competently in every case or job they might encounter. The idea is contrary to the notion of specialized training and competence. Students, including students with disabilities, deserve teachers with training appropriate for their learning needs ([Kauffman et al., in press](#)).

Special education teachers need to adjust teaching to kids’ learning styles

Learning styles is one of many ideas that have intuitive appeal, but are not supported by scientific evidence. The effectiveness of adjusting instruction to accommodate “learning styles” of students with disabilities is not supported by corroborating evidence ([Riener and Willingham, 2010](#)).

Restructuring is the key to improving schools

Restructuring in the form of rearranging duties and responsibilities or other combinations of responsibilities of staff, movement of students, changing hours or calendar, and so on are highly unlikely to significantly improve school learning. Improving instruction based on scientific evidence is the key to improving school learning. Cultural and organizational changes can shape conditions of instruction, but there is no silver bullet from outside that will automatically solve learning problems.

Special education was created by educators, not parents

Many educators have been advocates for students with disabilities and their appropriate education. However, the support and advocacy of parents have been the most critical factors in the creation of special education law, local policy, and the provision of many specialized facilities (see [Martin, 2013](#)).

Conclusion and the way forward for special and inclusive education

There has now been more than 25 years of promotion of an inclusion agenda without producing convincing research evidence that it is superior to alternatives ([Cook and Cook, 2020](#); [Fuchs et al., 2022](#)), or that any country or region has been able to effectively educate *all* students with disabilities in mainstream classrooms. It is therefore considered that it is well overdue time to move beyond controversy and debate on this issue in order to establish a new way forward in the education of students with disabilities. This will involve a more temperate approach to inclusion that combines the best elements of special education with those of inclusive education ([Gordon-Gould and Hornby, 2023](#); [Hornby, 2015](#); [Kauffman et al., 2018a](#)).

Within this we consider that the highest value should be placed on the learning of students with disabilities, not on their bodily presence in a particular place, such as the general education classroom. Inclusion is an important value, but we conclude that it should mean inclusion in appropriate learning, wherever that learning occurs. All students with disabilities are entitled to an education that is appropriate for them, and sometimes this is best delivered in a separate setting outside the general education classroom.

References

- Agran, M., Jackson, L., Kurth, J., Ryndak, D., Burnette, K., Jameson, M., Zagana, A., Fitzpatrick, H., Wehmeyer, M., 2020. Why aren't students with severe disabilities being placed in general education classrooms? Examining the relations among classroom placement, learner outcomes, and other factors. *Res. Pract. Persons Severe Disabil.* 45 (1), 4–13.
- Ahrbeck, B., Badar, J., Felder, M., Kauffman, J., Schneiders, K., 2018. Fachbeitrag: Totale Inklusion? Fakten und Überlegungen zur Situation in Deutschland und den USA. *Q. J. Spec. Needs Educ. Neighbor. Areas* 87 (3), 218–231.
- Anastasiou, D., Bantekas, I., (in press). Education for learners with disabilities as a social right. *University of Pennsylvania Journal of Constitutional Law*.
- Anastasiou, D., Kauffman, J.M., 2011. A social constructionist approach to disability: implications for special education. *Except. Child.* 77, 367–384.
- Anastasiou, D., Kauffman, J.M., 2012. Disability as cultural difference: implications for special education. *Remedial Special Educ.* 33, 139–149.
- Anastasiou, D., Kauffman, J.M., 2013. The social model of disability: dichotomy between impairment and disability. *J. Med. Philos.* 38, 441–459.
- Anastasiou, D., Kauffman, J.M., 2019. Cultural politics, ideology, and methodology in disproportionality research: a rejoinder. *J. Disabil. Pol. Stud.* 30, 105–110.
- Anastasiou, D., Kauffman, J.M., Di Nuovo, S., 2015. A reply to the *EJSNE* symposium's respondents. *Eur. J. Spec. Needs Educ.* 30, 457–458.
- Anastasiou, D., Kauffman, J.M., Michail, D., 2016. Disability in multicultural theory: conceptual and social justice issues. *J. Disabil. Pol. Stud.* 27, 3–12.
- Anastasiou, D., Gregory, M., Kauffman, J.M., 2018. Commentary on Article 24 of the CRPD: the right to education. In: Bantekas, I., Stein, M., Anastasiou, D. (Eds.), *Commentary on the UN Convention on the Rights of Persons with Disabilities*. Oxford University Press, New York, NY, pp. 656–704.
- Anastasiou, D., Felder, M., Correia, L., Shemanov, A., Zweers, I., Ahrbeck, B., 2020. The impact of Article 24 of the CRPD on special and inclusive education in Germany, Portugal, the Russian Federation, and Netherlands. In: Kauffman, J.M. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, New York, pp. 236–245.
- Baglieri, S., Bejoian, L.M., Broderick, A.A., Connor, D.J., Valle, J., 2011. [Re]claiming “inclusive education” toward cohesion in educational reform: disability studies unravels the myth of the normal child. *Teach. Coll. Rec.* 113 (10), 2122–2154.
- Barnes, C., 1991. *Disabled People in Britain and Discrimination*. Hurst, London, United Kingdom.
- Bateman, B.D., Linden, M.A., 2012. *Better IEPs: How to Develop Legally Correct and Educationally Useful Programs*, fifth ed. Attainment, Verona, WI.
- Bateman, B.D., 1994. Who, how, and where: special education's issues in perpetuity. *J. Spec. Educ.* 27, 509–520.
- Bateman, B.D., 2007. Law and the conceptual foundations of special education practice. In: Crockett, J.B., Gerber, M.M., Landrum, T.J. (Eds.), *Achieving the Radical Reform of Special Education: Essays in Honor of James M. Kauffman*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 95–114.
- Bateman, B.D., 2017. Individual education programs for children with disabilities. In: Kauffman, J.M., Hallahan, D.P., Pullen, P.C. (Eds.), *Handbook of Special Education*, second ed. Routledge, New York, pp. 87–104.
- Blackburn, S., 2005. *Truth: A Guide*. Oxford University Press, New York.
- Blackman, H.P., 1992. Surmounting the disability of isolation. *Sch. Adm.* 49 (2), 28–29.
- Bunge, M., 2017. *Doing Science in the Light of Philosophy*. World Scientific, New York.
- Callahan, C.M., Plucker, J.A., Gluck, S., Rodriguez, C., 2020. Inclusion of academically advanced (gifted) students. In: Kauffman, J.M. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, New York, pp. 176–194.
- Cook, B.G., Cook, L., 2020. An examination of highly cited research on inclusion. In: Kauffman, J.M. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, New York, pp. 130–159.
- Correia, L.M., 2018. *Educação Inclusiva & Necessidades Especiais*. Flora Editora, Braga, Portugal.
- CRPD Committee, 2016. General Comment No. 4 on the Right to Inclusive Education. UN Doc. CRPD/C/GC/4.
- Dinishak, J., 2016. The deficit view and its critics. *Disabil. Stud. Q.* 36 (4). <https://dsq-sds.org/article/view/5236/4475>.
- Farrell, M., 2010. *Debating Special Education*. Routledge, London.
- Felder, M., Schneiders, K., 2016. *Inklusion kontrovers: Herausforderungen für die Soziale Arbeit*. Schwalbach am Taunus. Wochenschau Verlag, Germany.
- Feniak, C.A., 1988. Labeling in special education: a problematic issue in England and Wales. *Int. J. Spec. Educ.* 3, 117–124.
- Fox, R.M., Mulick, J.A. (Eds.), 2016. *Controversial Therapies for Autism and Intellectual Disabilities: Fad, Fashion, and Science in Professional Practice*, second ed. Routledge, New York.
- Fuchs, L.S., Fuchs, D., Compton, D.L., Wehby, J., Schumacher, R.F., Gersten, R., Jordan, N.C., 2014. Inclusion versus specialized intervention for very-low-performing students: what does access mean in an era of academic challenge? *Except. Child.* 81 (2), 134–157.
- Fuchs, D., Mirowsitz, H.C., Gilbert, J.K., 2022. June 17. Exploring the truth of Michael Yudin's claim: The more time students with disabilities spend in general classrooms the better they do academically. *Journal of Disability Policy Studies*. <https://doi.org/10.1177/10442073221097713>.
- Gerber, M.M., 2017. A history of special education. In: Kauffman, J.M., Hallahan, D.P., Pullen, P.C. (Eds.), *Handbook of special education*, 2nd ed. Routledge, New York, pp. 3–15.
- Gordon-Gould, P., Hornby, G., 2023, in press. *Inclusion at the crossroads: Exploring effective special needs provision in global contexts*. Routledge, New York.

- Harris, J., 2000. Is there a coherent social conception of disability? *J. Med. Ethics* 26, 95–100.
- Horn, J.L., 1924. *The Education of Exceptional Children: A Consideration of Public School Problems and Policies in the Field of Differential Education*. Century, New York.
- Hornby, G., 1999. Inclusion or delusion: can one size fit all? *Support Learn.* 14 (4), 152–157.
- Hornby, G., 2014. *Inclusive Special Education: Evidence-Based Practices for Children with Special Needs and Disabilities*. Springer, New York.
- Hornby, G., 2015. Inclusive special education: development of a new theory for the education of children with special educational needs and disabilities. *Br. J. Spec. Educ.* 42, 234–256.
- Hughes, B., Patterson, K., 1997. The social model of disability and the disappearing body: towards a sociology of impairment. *Disabil. Soc.* 12 (3), 325–340. <https://doi.org/10.1080/09687599727209>.
- Hurwitz, S., Perry, B., Cohen, E.D., Skiba, R., 2019. Special education and individualized academic growth: a longitudinal assessment of outcomes for students with disabilities. *Am. Educ. Res. J.* <https://doi.org/10.3102/0002831219857054>.
- Hyatt, C., Hornby, G., 2017. Will UN Article 24 lead to the demise of special education or to its re-affirmation? *Support Learn.* 32, 288–304.
- Imray, P., Colley, A., 2017. *Inclusion is Dead: Long Live Inclusion*. Routledge, London.
- Kahneman, D., 2011. *Thinking Fast and Slow*. Farrar, Strauss and Giroux, New York.
- Kauffman, J.M., 1981. Introduction: historical trends and contemporary issues in special education in the United States. In: Kauffman, J.M., Hallahan, D.P. (Eds.), *Handbook of special education*. Prentice-Hall, Englewood Cliffs, NJ, pp. 3–23.
- Kauffman, J.M., Badar, J., 2013. How we might make special education for students with emotional or behavioral disorders less stigmatizing. *Behav. Disord.* 39, 16–27.
- Kauffman, J.M., Badar, J., 2014. Instruction, not inclusion, should be the central issue in special education: an alternative view from the USA. *J. Int. Spec. Needs Educ.* 17, 13–20.
- Kauffman, J.M., Badar, J., 2016. It's instruction over place—not the other way around! *Phi Delta Kappan* 98 (4), 55–59.
- Kauffman, J.M., Badar, J., 2018. Extremism and disability chic. *Exceptionality* 26, 46–61.
- Kauffman, J.M., Badar, J., 2020. Definitions and other issues. In: Kauffman, J.M. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, New York, pp. 1–24.
- Kauffman, J.M., Badar, J. (Eds.), 2023, in press. Navigating students' mental health in the wake of COVID-19: Using public health crises to inform research and practice. Routledge, New York.
- Kauffman, J.M., Hornby, G., 2020. Inclusive vision versus special education reality. *Educ. Sci* 10 (9), 1–13. <https://doi.org/10.3390/educsci10090258>.
- Kauffman, J.M., Hornby, G. (in press). Special education's zombies and their consequences. *Exceptionality*.
- Kauffman, J.M., Hung, L.Y., 2009. Special education for intellectual disability: current trends and perspectives. *Curr. Opin. Psychiatry* 22, 452–456.
- Kauffman, J.M., Landrum, T.J., 2009. Politics, civil rights, and disproportional identification of students with emotional and behavioral disorders. *Exceptionality* 17, 177–188.
- Kauffman, J.M., Landrum, T.J., 2018. *Characteristics of Emotional and Behavioral Disorders of Children and Youth*, eleventh ed. Pearson, Upper Saddle River, NJ.
- Kauffman, J.M., Lloyd, J.W., 2017. Statistics, data, and special education decisions: basic links to realities. In: Kauffman, J.M., Hallahan, D.P., Pullen, P.C. (Eds.), *Handbook of Special Education*, second ed. Taylor & Francis, New York, pp. 29–39.
- Kauffman, J.M., Lopes, J.A. (Eds.), 2007. *Pode a educação especial deixar de ser especial? Psiquilíbrios Edições*, Braga, Portugal.
- Kauffman, J.M., Sasso, G.M., 2006. Certainty, doubt, and the reduction of uncertainty: a rejoinder. *Exceptionality* 14, 109–120.
- Kauffman, J.M., Mock, D.R., Tankersley, M., Landrum, T.J., 2008. Effective service delivery models. In: Morris, R.J., Mather, N. (Eds.), *Evidence-Based Interventions for Students with Learning and Behavioral Challenges*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 359–378.
- Kauffman, J.M., Farkas, G., 2022. Veracity in special education. *Exceptionality* 30 (1), 43–56. <https://doi.org/10.1080/09362835.2021.1938066>.
- Kauffman, J.M., Felder, M., Ahrbeck, B., Badar, J., Schneiders, K., 2018a. Inclusion of *all* students in general education? International appeal for a more temperate approach to inclusion. *J. Int. Spec. Needs Educ.* 21, 1–10.
- Kauffman, J.M., Hallahan, D.P., Pullen, P.C., Badar, J., 2018b. *Special Education: What it is and Why We Need It*, second ed. Taylor & Francis, New York.
- Kauffman, J.M., Schumaker, J., Badar, J., Hallenbeck, B.A., 2018c. Where special education goes to die. *Exceptionality* 27, 149–166.
- Kauffman, J.M., Wiley, A.L., Travers, J.C., Badar, J., Anastasiou, D., 2019. Endrew and FAPE: concepts and implications for all students with disabilities. *Behav. Modif.* 45 (1), 177–198.
- Kauffman, J.M., Ahrbeck, B., Anastasiou, D., Badar, J., Felder, M., Hallenbeck, B.A., 2020a. Special education policy prospects: lessons from social policies past. *Exceptionality*. <https://doi.org/10.1080/09362835.2020.1727326>.
- Kauffman, J.M., Hallahan, D.P., Landrum, T.J., Smith, C.R., 2020b. Likely legacies of inclusion. In: Kauffman, J.M. (Ed.), *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, pp. 249–265.
- Kauffman, J.M., Travers, J.C., Badar, J., 2020c. Why *some* students with severe disabilities are not placed in general education. *Res. Pract. Pers. Severe Disabil.* 45 (1), 28–33.
- Kauffman, J.M., Hallahan, D.P., Pullen, P.C., (in press). Creeping normality: special education's problem of a new normal. *J. Disabil. Pol. Stud.*
- Kauffman, J.M., 1989. The regular education initiative as Reagan-Bush education policy: a trickle-down theory of education of the hard-to-teach. *J. Spec. Educ.* 23, 256–278.
- Kauffman, J.M., 2005. Waving to Ray Charles: missing the meaning of disability. *Phi Delta Kappan* 86, 520–521, 524.
- Kauffman, J.M., 2008. Would we recognize progress if we saw it? A commentary. *J. Behav. Educ.* 17, 128–143.
- Kauffman, J.M. (Ed.), 2020. *On Educational Inclusion: Meanings, History, Issues and International Perspectives*. Routledge, New York.
- Kauffman, J.M., 2022. This one, not that one: Toward revitalizing special education. In: Kauffman, J.M. (Ed.), *Revitalizing Special Education: Revolution, Devolution, Evolution*. Emerald, Bingley, UK, pp. 1–41.
- Kliewer, C., Biklen, D., Kasa-Hendrickson, C., 2006. Who may be literate? Disability and resistance to the cultural denial of competence. *Am. Educ. Res. J.* 43, 163–192.
- Laski, F.J., 1991. Achieving integration during the second revolution. In: Meyer, L.H., Peck, C.A., Brown, L. (Eds.), *Critical Issues in the Lives of People with Severe Disabilities*. Paul H. Brookes, Baltimore, pp. 409–421.
- Lindsay, G., 2007. Educational psychology and the effectiveness of inclusive education/mainstreaming. *Br. J. Educ. Psychol.* 77, 1–24.
- Lipsky, D.K., Gartner, A., 1996. Inclusion, school restructuring, and the remaking of American society. *Harv. Educ. Rev.* 66, 762–796.
- Martin, E.W., 2013. *Breakthrough: Federal Special Education Legislation 1965–1981*. Bardolf, Sarasota, FL.
- Mitchell, C., 2019. Most classroom teachers feel unprepared to support students with disabilities. *Educ. Week*. May 29.
- Neiman, S., 2008. *Moral Clarity: A Guide for Grown-Up Idealists*. Harcourt, New York.
- OECD, 1999. *Inclusive Education at Work: Students with Disabilities in Mainstream Schools*. OECD, Paris.
- Oliver, M., 1992. Changing the social relations of research production? *Disabil. Handicap Soc.* 7 (2), 101–114. <https://doi.org/10.1080/02674649266780141>.
- Oliver, M., 1996. *Understanding Disability: From Theory to Practice*. Macmillan, Basingstoke, United Kingdom.
- Oliver, M., 2000. Decoupling Education Policy from the Economy in Late Capitalist Societies: Some Implications for Special Education. Keynote Address at the International Special Education Congress, Manchester, England.
- Pfeiffer, D., 2002. The philosophical foundations of disability studies. *Disabil. Stud. Q.* 22 (2), 3–23.
- Rauch, J., 2021. *The constitution of knowledge: a defense of truth*. Brookings Institution Press, Washington, DC.
- Riddle, C.A., 2014. *Disability and justice: The capabilities approach in practice*. Lexington Books, Plymouth, UK.
- Riener, C., Willingham, D., 2010. The myth of learning styles. *Change* 42 (5), 32–35.
- Schwartz, A.E., Hopkins, B.G., Stiefel, L., 2019. The Effects of Special Education on the Academic Performance of Students with Learning Disabilities. <http://www.edworkingpapers.com/ai19-86>.
- Shakespeare, T.W., 2006. *Disability Rights and Wrongs*. Routledge, London.

- Silvestri, S.M., Heward, W.L., 2016. The neutralization of special education, revisited. In: Foxx, R.M., Mulick, J.A. (Eds.), *Controversial Therapies for Autism and Intellectual Disabilities*, second ed. Routledge, New York, NY, pp. 136–153.
- Specter, M., 2009. *Denialism: How Irrational Thinking Hinders Scientific Progress, Harms the Planet, and Threatens Our Lives*. Penguin, New York.
- Stainback, W., Stainback, S., 1991. A rationale for integration and restructuring: a synopsis. In: Lloyd, J.W., Singh, N.N., Repp, A.C. (Eds.), *The Regular Education Initiative: Alternative Perspectives on Concepts, Issues, and Models*. Sycamore, Sycamore, IL, pp. 226–239.
- SWIFT Schools, 2020. www.swiftschools.org.
- Terzi, L. (Ed.), 2010. *Special Educational Needs: A New Look*. Continuum, London.
- UNESCO, 1994. *The Salamanca Statement and Framework for Action on Special Needs Education*. UNESCO, Salamanca, Spain.
- United Nations Convention on the Rights of Persons with Disabilities (UNCRPD), 2016. Article 24: Right to Inclusive Education. Accessed from. <http://www.refworld.org/docid/57c977e34.html>.
- Warnock, M., 2010. Special educational needs: a new look. In: Terzi, L. (Ed.), *Special Educational Needs: A New Look*. Continuum, London, UK, pp. 11–45.
- Wehmeyer, M.L., Shogren, K.A., Kurth, J., 2020. The state of inclusion with students with intellectual and developmental disabilities in the United States. *J. Pol. Pract. Intellect. Disabil.* 18 (1).
- Wiley, A., Anastasiou, D., Kauffman, J., 2019. No, special education does not treat disability like a disease and is not “obsessed” with forcing students to conform. *Wash. Post*. https://www.washingtonpost.com/education/2019/02/25/no-special-education-does-not-treat-disability-like-disease-is-not-obsessed-with-forcing-students-conform/?fbclid=IwAR3gRuBXap41HmccS_hjOhoemCE90QV1mD0dfG6tNkZk9-hkhLu6B0RC9rA&noredirect=on&utm_term=.4a6cac1bc670.

Re-examining special needs—what could be better?

Jonathan Rix^{a,b}, ^aThe Open University, Milton Keynes, United Kingdom; and ^bInland Norway University of Applied Sciences, Innlandet, Norway

© 2023 Elsevier Ltd. All rights reserved.

Words?	284
History?	285
Un-together?	286
Re-arranged?	286
Continuation?	288
Beyond?	288
Needless?	289
And?	290
References	291

I have a problem

I have to write
a chapter
about
special needs
even though
I do not know
what they are
and simply
have to rely on
no one else
knowing
what they are
either

Words?

Things have to start somewhere, don't they? This chapter for example, has begun with a sort of poem. The idea didn't come from nowhere though. Poetry isn't short on history. I am just one in a long line of millions who have written it. It is what it is because we have shared and talked about it a multitude of times in so many contexts over millenia. As a reader there are numerous ways in which you might have responded, but you will have had an immediate response. This may have leaned toward a sense of positivity or negativity or amusement or annoyance or ... well the list is full of endless permutations. And it will have been informed by your relationship with the history of the thing we call poetry (in English).

Of course your response would have been informed by your relationship with the history of a great many other things too. For example your perspective on what special needs are; and your perspective on what an encyclopedia chapter should be like. You may, for instance, have had a very particular response to the use of "I" at the start of a chapter. You may have presumed a kind of expertise in the author that would not have used the first person or claimed to be ignorant about the thing they are writing about.

Words. We always begin with a response ... even if it is not much of a response at all. And that response is buried in a huge interwoven melee of personal, cultural, environmental experiences. Special Needs—2 words, together and apart, you and they bring with them so much meaning.

- Since it comes first, let us start with special. What is "special"? Are you special good or special bad? Are you labeled special or are you claiming special? Are you special as a difference from others or as a difference within yourself? Are you special across all contexts or in all cultures or just in some? Are you special in a way that brings you access to resources or are you special and it does not? Are you special for all time or for just some of it? Is anyone else special like you? Can your specialness be shared or is it unique? Does it rub off or encompass people associated with you? Are you special to everyone or just to some people? Can your specialness be measured? Can it be proven? Can it be explained? Are some parts of you more special than other parts? Does the

answer to all these questions change, so you can be both special good and bad, labeled and claiming, different from and within ... and so forth. And finally (if there really can be a finally) where does your special come from? Is it real or is it constructed?

- What about “needs”? Does the word coming second to “special” tell us something about the needs themselves? Who decides about these needs? Do you identify your needs or does someone else? Are they expressed or are they inferred? Are they relative to others or fundamental to you? Are the needs a result of something in you or something in what you have been provided with? Are your needs a strength or a weakness? Are the needs individual to you, shared with others, or collective needs? Are these needs in all aspects of your life or just in some? Are they developmental or developed? Who needs these needs? Are your needs actually priorities?

As these questions are intended to show, how we use each term will have a complex history and a complex web of meanings which make any simple understanding of them nigh on impossible. This is made even more complex when we combine them. All of which is ironic of course, because “She’s got special needs,” is supposed to make things much clearer.

History?

I am by no means the first person to raise these kinds of issues, of course. For example as part of three short articles in which he critiqued both inclusion and special needs, Wilson (2002) asked a series of questions about special needs, noting its ambiguity of meaning as well as some underlying assumptions associated with it. He pointed out that special needs requires an other to be special to. It therefore requires reference to criteria associated with something generally seen to be desirable, important and necessary. This is bound to vary within and between systems and cultures. Neither specialness nor neediness can ever be attributes of the person except within a relational context. This is true about all human experience of course, which always exists in relationship to others. As such it will be subject to the power which exists within relationships, in the give and take between us (Foucault, 1980, 1982). In the context of special needs it is fair to suggest that there is greater power for those who define the criteria of specialness and neediness than those to whom the criteria are applied. This reflects the “irresolvable contradiction” which Oliver (1996) pointed out.

To put it bluntly, people in need of care cannot be empowered because the reason that they need care is precisely because they are powerless;- if they were not, why would they need care?

These contradictions and negative connotations have been discussed for decades. For example, Tomlinson (1982) pointed to Rousseau’s use of needs in his 1762, treatise on Education, *Emile*, [*N’accordez rien à ses désirs parce qu’il le demande, mais parce qu’il en a besoin*] which roughly means, “Do not agree to his wishes because he asks for it, but because he needs it”, noting that Rousseau does not define who identifies those needs. Roaf and Bines (1989) pointed out that *needs* is often used when discussing development and learning for all children, but can be seen as being linked to social and cultural structures and issues of equality. They suggest this is evident in the haphazard and unequal provision that can be associated with attempts to deliver both general and special equitable educational opportunity. They suggest too that by linking “needs” with “special,” the focus has been shifted onto “*individuals as a bottomless pit of problems to be overcome or filled up*” (p. 10). They point out that the potential radicalism of associating needs with the context has been avoided, that their association with special means they are based upon deficits in the person; as such combining special and needs simply relabelled old problems and old practices.

This tendency to focus upon a deficit type of need, is perhaps not surprising given the use of the term special needs when it began to appear in educational texts. Gulliford (1971) for example, explained that people with different “*handicaps*” can have very similar special needs. He said we should not define individuals according to an impairment which results in loss of capacity or function but instead should focus upon the extent to which the impairment “*becomes a handicap*.” As an example he suggests that a “*child with low level of intelligence can be unhappily handicapped in a school with predominantly bright children*” (p. 5). The extent to which the impairment becomes a handicap “*depends on many other factors*” (p. 4) He suggests these are school factors, home factors, remedial intervention and the child’s intelligence, personality, number of impairments and the way he is developing. And the resulting needs “*may be met in a special school, in special classes or units in ordinary schools or be the subject of some extra attention and care within the ordinary work of schools*” (p. 6).

What Gulliford does not mention in his deficit view of special and needs, but Tomlinson does, are the professional vested interests which are associated with the remedial interventions offered by those services and that special “*expertise*.” Tomlinson predicted an ongoing expansion of these services with “*more expert special teachers, support staff, advisers and inspectors to be employed*” (Tomlinson, 1982, p. 161) suggesting this was an inevitable response to a range of educational and socio-economic pressures and priorities. Returning to this topic 30 years later (Tomlinson, 2012), she noted that whole new professions had emerged as part of this SEN industry, as well as numerous Associations, Institutes and certifying bodies. The huge growth in numbers of identified children, identifying professionals and in funding, across nations, leaves little room to doubt that special needs education has been on the up globally for decades (Rix, 2015).

Un-together?

Just because special and needs is a global growth business does not mean it is a unified field. In carrying out a review of special educational provision across 55 administrations in 50 countries (Rix et al., 2013) we noted that the only factors that applied across all administrations and countries were that:

- Children are marginalized within all education systems. (Who they are and why they are marginalized varies between systems.)
- Provision referred to as special involves time and space additional to that provided typically.

This variability means, for example, that countries' definitions of special educational needs cannot be categorized for comparison without creating a deceptive sense of commonality (Riddell et al., 2006) and consequently can only be listed as a tool for reflection (see OECD, 2012). In our study, the number of categories associated with special and needs varied considerably between countries; ranging from 3 categories to 22 categories. Once all the obvious similarities were grouped together there were 60 different categories which emerged across the 55 administrations. This of course reflects the relational nature of special and needs discussed above.

Special and needs are localized phenomena, responsive to localized priorities, policies, practices and understandings. For example, in England in 2016, 39% of people taking national exams at 15/16 years of age had been identified with SEN at some point in their school career, but the overall probability for this attribution was 71% explained by which school the child attended (Hutchinson, 2021). In 2020, in one English local authority, you could find that 3.4% of diagnosed children had been diagnosed with an autistic spectrum disorder, but in another local authority it was 25%. In every one of the 12 reported categories covering sensory, intellectual, functional and behavioral difficulties or impairment there were huge discrepancies. If you take the top 10 and bottom 10 English local authorities (to iron out the extremes) using any label, you will be more than 4 times as likely to have a particular diagnosis in a top-ten using authority than in a bottom-ten. This even applies for what might seem like a highly evident sensory-related category such as a hearing impairment. The only exceptions where the ratio falls slightly below 3:1 are around issues of behavior and communication which are usually identified by the teacher (DfE, 2020).

This is not just an English problem, of course. Special and needs is chaotic globally. Our administrative system may yearn for unification, but the reality is that our systems cannot provide this (Rix, 2020). For example, in the 50 country study mentioned above we visited Italy and Norway. Before we went we knew that both countries claimed to have closed their special schools, that Italy required a psycho-medical assessment to achieve additional funding while Norway did not. But when we got there we found that there were special schools and special classes in existence with varying degrees of official support at different levels within the system. In addition, some administrations in Italy were trying to provide additional funding without psycho-medical assessment, while in Norway psycho-medical labels were in everyday use as part of local funding distribution processes.

Assessments are clearly one of the major challenges for the seekers of certainty. Identifying and applying special and needs requires finding observable phenomena as evidence of behaviors or characteristics. The markers that experts search for have been selected as a focus because of their practical value in the diagnostic process. But consequently they can also be seen as theoretical constructs in their own right, requiring us to unpack *"the constructs underlying [the] symptoms"* (Wilshire and Clack, 2021) and creating considerable space for disagreement.

These theoretical and practical tensions underpin calls such as that coming from the Society for Humanistic Psychology, Division 32 of the American Psychological Association, regarding the reform and revision of diagnostic systems. In an official statement, released in February 2020, they talked about the need to move beyond the limits of individual diagnosis, to document and address the social and structural determinants at the root of difficulties. They concluded that the impact of these social and structural determinants was already supported by robust scientific evidence while the search for in-person factors continues to "elude scientists" who merely hold onto hope for *"the next taxonomic revolution."* A few years earlier, the British Psychological Society's Division of Clinical Psychology (BPS-DCP, 2013) had also called for *"a paradigm shift in relation to functional psychiatric diagnoses."* It said they: *"provide a flawed basis for evidence-based practice, research, intervention guidelines and the various administrative and non-clinical uses of diagnosis."*

This struggle to find reliable observable phenomena, goes across key categories associated with special and needs too; as well as being recognized across the field of educational psychology. For example:

- Elliott and Gibbs (2008) note that *"there is no clear discontinuity that provides an absolute categorical boundary for a diagnostic category of 'dyslexics'"* (p. 476);
- Norwich et al. (2014) note the paucity of research and the application in practice of Moderate Learning Difficulties which *"call into question"* (p. 17) its use as a category;
- Timimi (2021), citing other leaders in autism research, recommends disbanding Autism Spectrum Disorder entirely because of *"consistently inconsistent findings"* (p. 19) and decades of research bedeviled by *"replicability crises."*

Re-arranged?

Evidently a great many experts in the field of special and needs don't agree with each other about what to do, but there is fair agreement about what they/we disagree over. Norwich (2016) suggests that future assessment approaches need to: be underpinned and strongly informed by values associated with human rights; explore all functional aspects not just areas of concern; assume

interaction between impairments and the environment; and consider personal characteristics beyond the functional difficulties. Some people feel that achieving this is akin to squaring a circle, but this has not stopped people from trying. For example, as part of the global effort to come up with solutions, researchers associated with the British and American Psychological Associations have been exploring holistic solutions which take all the issue of context into account.

There are also other approaches to assessment that have been advocated at scale with the intention of transforming practice. Let's consider two important developments:

- Response to Intervention (RTI) arose in the United States with the growth (from the 1960s onwards) of concerns related to: discrepancies between “ability” and achievement; overidentification of learning disabilities; and a lack of both early identification and curriculum-based measurement. These combined with the emergence of influential research groups and policy makers and led (in 2004) to the RTI's place in the Individuals with Disabilities Education Act (Preston et al., 2016). There are two dominant RTI approaches, operating in stages either as a problem-solving approach focused on individual students, or as a standard protocol that operates at class, small group & individual levels. A strong motivation in RTI is the application of evidence-based practice.
- The international classification of functioning, disability and health (ICF) was approved by the World Health Organization in 2001. The ICF is based on a biopsychosocial model, in which functioning and disability are understood as complex interactions between biological, psychological, and social factors. A central problem within the educational context is the ICF's breadth, and so firstly an ICF-CY was produced with a focus on childhood and youth, and then in 2011a, Hollenweger designed an eligibility procedure for specialized education services based on the ICF-CY and the United Nations Convention on the Rights of Persons with Disability. Her intention was to build bridges between the child's situation and the teacher's situation, creating opportunities for reflection and co-ordinating support (Hollenweger, 2014).

The intentions behind both of these developments can seem admirable. They are based upon critical engagement with dominant deficit views of special and needs. Both seek to work within the learning context and focus upon functional involvement to transform teacher practice and the learner experience. However, for both these models (and I would suggest from any frameworks likely to emerge from the work of US & UK psychologists), there seems to be a more fundamental problem. They are facing a tri-partite problem. The challenge of engaging with function in context is only one part. The other two are:

- The people who are undertaking the assessments and delivering the learning support; and
- The funneling of attention back onto the individual.

These two issues are interwoven. Focusing personalized, responsive attention onto the individual sounds like it should be great, but it brings practitioners (and researchers) back to individualized notions of need. They are still seeking to identify certain learners as inherently different, and therefore for many practitioners this equates to a child who requires inherently different teaching. The special and the needs are still there. This seems to be borne out by studies of teachers engaging with these approaches. It is clear that school personnel (including administrators, psychologists, and teachers) are unclear about the purpose of RTI and their roles within its application (Cavendish et al., 2016); it has also been shown to create discomfort and chaotic dynamics and to be dependent upon the contexts in which it is delivered (Dulaney, 2013). It is hardly surprising therefore that practitioners feel that to engage with RTI they need professional development (Newman Thomas et al., 2020).

The call for greater professional development has been at the heart of special and needs from the earliest days (a key part of the professionalization process about which Tomlinson spoke). As with RTI, where the ICF has been introduced in an educational context, teachers are frequently resistant to making use of it. Hollenweger (2011b) herself writes about the need for specific training to engage with her approach. Studies often link teacher resistance back to the need for training, highlighting how the intensity of training makes teachers more likely to see it as a useful approach, and calling for the introduction of ongoing professional development and a process of reflective practice to ensure they engage with its processes (e.g., Saragoça et al., 2013). In RTI, this need for skills training and knowledge acquisition is intensified not only by advocating for the use of an evidence base, but also because the teacher has to work with increasing specialization with the child as they move through the tiers, toward increasingly individualized support (Fuchs and Fuchs, 2006). A similar knowledge-gap dilemma emerges from the ICF-CY-Rights model. Hollenweger describes three types of information that should be sought, two of which can be gathered “objectively” by specialists, with only the third involving teachers, where their views are sought about contextual factors (e.g., resource availability, cultural values, treatment priorities, expectations).

This call for training also runs up against another challenge for special and needs, the research base within which any expertise is situated. Whenever systematic literature reviews consider educational research associated with any special and needs intervention and/or category of impairment they conclude that there is a paucity of evidence, that the research is either not at a suitable scale or of a suitable quality to make hard and fast conclusions (Rix, 2015). The problem for evidence seekers is that any intervention needs to be fully tested in real world contexts (Holman et al., 2018). Its own measure of effectiveness, and therefore proof of truth, can only be satisfactorily proven when it is reliably producing its predicted outcomes within those everyday classroom contexts, in a wide variety of cultures undertaken by everyday practitioners. The complexity of such an approach might seem beyond the means of many researchers who are not suitably networked, particularly since agreed evaluation processes at scale still need to be developed (Stahmer et al., 2018) and the resources and costs are considerable (Horner et al., 2017). However,

even for those who are well networked such an approach can prove hard to achieve simply because of the practicalities of achieving a realistic sample (e.g., [Taylor et al., 2017](#)). This difficulty is made harder by the challenges of delivering evidence-based practices with “fidelity” to the original designers’ intentions. The lack of fidelity would seem to be evident for example in relation to RTI studies that suggest the process of identification is still arbitrary ([Hendricks and Fuchs, 2020](#)) and still overestimates the focus population ([Kranzler et al., 2020](#)).

Continuation?

Both the RTI & ICF-CY-Rights approaches require external specialists to oversee them and/or interested colleagues who are willing to develop the skills valued by those experts. They both run up against the same barriers that so many ideas (not just dealing with special and needs) face when they are introduced into schools. They are primarily non-educational constructs which have been redesigned to fit into the education world. The response of teacher’s therefore is that they require to be trained in “it.” “It” is not something which they can do as part of the day job. “It” requires a special expertise. Secondly, the whole process echoes the continuum model which was described in the 1960s and 1970s, and so can neatly fit within the thinking associated with it. The continuum mindset nestles up to the notion that the teacher does not have the necessary knowledge. Children and young people are positioned on a continuum of need, supported within a continuum of provision, and by a continuum of services. The priority of policy makers is to provide what they regard as the institutional flexibility associated with a continuum of settings and services. The expectation for teachers is that someone else can offer support if they do not ([Rix et al., 2015](#)).

Yet, the continuum did not emerge as some brilliant invention seeking to deliver social justice. It merely represented the provision and resources available as education began to accept people who had been previously ignored by schools or had been kept in long-stay institutions. It was (and is) an idealized solution for the challenge of how to organize, identify, place and teach special and needs. The use of the continuum metaphor was an attempt to provide a logical framework for the chaotic nature of the available provision. It was part of the great search for certainty at the heart of our educational process and of special and needs ([Rix, 2020](#)). Perhaps ironically, the development of special education and psychology also had their birth in this search for certainty, emerging from processes developed for selection and intervention, such as the measurement of behavior ([Clough, 1995](#)), grace ([Goodey, 2011](#)) and the acceptable range of a child’s performance at school ([Richardson and Powell, 2011](#)).

This continuation of the old ways of doing things, under the guise of doing something new, is a symptom of this search for certainty in the face of a complex, uncertain world. People are reassured if they can feel that their responses are “correct.” They highly value the processes and the outputs of the systems of which they are part, they recognize how the structures organize the distribution of resources, and how they enable their evaluation and their effectiveness across the system. However, the fundamental challenge is that seeking certainty oversimplifies the complexity to make it manageable; in creating boundaries it not only restricts the problem to an individual or small population of individuals, it means that shared or less obvious difficulties are missed. This is not to say that people within systems seeking certainty are not aware of the problems that can be created by those systems. The problems are frequently evident on a large scale. However, the institutional and policy context limits their capacity to think outside and beyond those institutions’ and policies’ boundaries, constraining their responses ([Ball, 1994](#)). As a result, our responses continue to divert focus away from the wider impact and structures of schools, from the complex economic and social factors underpinning disablement and from the delivery of meaningful social reform ([Slee, 2008](#)).

Beyond?

So what can be beyond special and needs? It is very difficult to see how we can move past its structures and mindsets when they are so deeply embedded in our current educational system. As much as we may wish our education systems to be transformational, there is a long tradition of powerful voices who see them as the necessary means of creating certainties, enabling social control and social organization. Attempts to meaningfully change any of these systems by arguments around rights and social justice is at best only going to deliver a few compromises and some pockets of success. This support for the status quo is a real problem for moving beyond special and needs, when many experts advocating change can be seen as saying—to mis-quote an old Irish saying—“*you don’t want to start from here, you need to start from over there.*”

So how can we start from here yet still generate movement away from special and needs. My solutions are premised upon encouraging a greater acceptance of the uncertainty at the heart of our life experiences, but they begin with recognizing and being responsive to the demands of the marketplace. Whenever business leaders are asked about what it is that their companies need, the newspaper headlines always focus on literacy and numeracy, alongside STEM and language communications. These are no means the top priorities though. Businesses are always looking for a focus upon interpersonal and intercultural competencies as well as methods of assessment to evidence these skills. They are concerned with life experience and attitudes, team working skills, communication skills, business and customer awareness, customer handling skills, foreign language skills and international cultural awareness ([Rix, 2015](#)).

If we want to meet these market demands we need to be arguing for (and getting businesses to be advocating for) pedagogies that embrace uncertainty, rather than seek certainty. This pedagogy of uncertainty is closely allied to relational, socio-cultural views of learning and critical pedagogies, as well as much of the writing associated with inclusive education and 21st century skills. It would take very little to redesign our assessment system and curricula to reflect this, making it a platform for our social goals and meeting market needs. We can prioritize any learning of subjects and key skills in this context. They can be part of our how we explore and validate people's capacity to consider ideas and resolve challenges interdependently in social contexts in ways that are meaningful to those involved (Rix, 2020).

Perhaps more challenging, but closely allied to this advocacy for engaging with a pedagogy of uncertainty, is shifting classroom practice away for constructivist to socio-cultural models of learning. This shouldn't be impossible though. This is the model taught in a great many teacher training contexts already. It can be seen to be encouraging teachers to adapt the principles of good teaching. These principles are widely regarded as being active, highly skilled and providing rich opportunities for learners within a well-planned and organized setting.

Effective practice relies upon everyday classroom ways of working: ways which teachers can readily understand, particularly if it fits with the curriculum and assessment required of them. Where there is an evidence base, it seems to be encouraging teachers to:

- Recognize their responsibility for all pupils, having respect for them and for all learning, while seeing everyone as a learner and collaborating with teachers who share this model of learning;
- Recognize social interaction as the means to develop knowledge as well as having clear learning goals and subject skills and knowledge;
- Plan to scaffold both cognitive and social content, delineating group roles, and using pupils own understandings as resources for learning;
- Explore pupils' understandings, encouraging questioning, making links between areas of student knowledge;
- Offer multimodal opportunities, with a focus upon the representation of concepts and the links between them, making them relevant and authentic;
- Focus upon communicative behaviors, and encourage naturalistic, child-centered and child-led approaches;
- Encourage playfulness and problem-solving, involving co-constructing, collaboration and setting aside time for learners to explore resources and generate ideas; and
- Explore real-world, practical, open-ended problems while providing direction, boundaries and support (Rix, 2021).

Importantly for a discussion of special and needs the approach outlined above removes special from the equation. If you are working in the manner described above to meet the needs of markets and reflect mainstream pedagogic guidance, then what might be considered an additional space or time becomes part of the everyday space and time. It is part of the day job.

This does not resolve the needs issue though.

Needless?

Needs is where the practitioner hits a wall. *The child(ren) who cannot do what the others are doing*. This is where teachers seek the support. This is where the assessment begins and this is where the big loop back to individualism restarts and special is reborn. If we wish to deal with needs we have to face three challenges: how we label them; how we identify them; and how we fund their support. As soon as any of those three factors becomes additional then the needs become special again. In order to deal with this—in a place which is very close to here—we have to embrace both the principle of uncertainty and teacher-relevance.

So let us start with the labels. I am an advocate of labels of opportunity (Rix, 2007). What teachers need is something simple. Something manageable. Something which makes them feel capable. They want something that does not make them call for more training from the outset. If we have to provide a label, school staff need one that encourages them to think about how they can best work with a child—that would support teaching, learning and social interaction. Consider, for example, the label of autism. This may be useful in a medical context and as a matter of pride, but within an education/social care/everyday situation it fails to open up a huge range of possibilities: this could be a child who will go to university or one who will never engage in effective eye to eye contact; they may be able to run a marathon or struggle to attend any social gathering; they may have many friends or may have none. Even if a teacher knows all the characteristics associated with this label, the child they are teaching will only have some of those characteristics. They have to deal with the child as they are, *just like they would with anyone else*. Depending upon this child's circumstances and the resources available in the setting, they may be more usefully described as *"a person who benefits from support to build social relations"* or *"a person who will use the break out space."* Such a label would not cover everything, but it would be a start and would help both the person who used the label and the person it was applied to. They are far simpler to share, to comprehend and to hold yourself to account for. They can be locally relevant, easier to assess, and less prone to the postcode lottery. It is a label about how to do something. It is situated in the social, creating relational opportunities and provocations. They can also change in ways that reflect the priorities of the people they are describing and the people involved in their lives. For example, the first

label mentioned above (benefiting from support to build social relations) may only be required for a while, perhaps at the start of each academic year, and not once some relationships had been established. Labels of opportunity are not who you are forever. They are who you are at the moment.

A move away from needs would also be supported by linking our descriptions of people to their access preferences rather than deficits or functional comparisons. For example during a participatory research project (working with disabled people, technology companies and museum education departments across Europe) there was an early collective agreement to refer to us as having access preferences. These were what we chose to include in the demographic data. There were many categories identified by the participants, ranging from Blue badge parking to Sign Language, Induction loops to Magnifying glasses, One-to-one support to Gallery stools. Such categories were specific to the museum context in which they were collected, but they far more effectively enable people working in museums (and elsewhere) to consider possibilities and the solutions for visitors, rather than having to make presumptions based upon a generalized diagnostic label. As the participants were quick to point out, many of these access options would also have benefits across a range of users, not just a particular group to whom they might be typically supplied (Rix, 2021). By focusing upon labels of opportunity and access preferences, teachers would be better prepared to envision practice. This is not to suggest that such labels would not also bring with them social and cultural assumptions, which could create problems with the people to whom they were applied. But they would be more likely to create commonalities around how access was supported. They would be more likely to enable a focus upon relationships and the support mechanisms between people.

The next creator of needs is our paperwork (Rix and Matthews, 2014). Beyond its perpetuation and construction of needs, our current paper work has three further practical problems. The first is the amount of it; the second is how it drives our ways of engaging; the third is how it removes highly knowledgeable individuals from the space in which they have most to contribute, the class learning context. Our goal has to be to reduce paperwork without reducing responsibility. The solution is to move away from the individual plans; and instead to encapsulate the rich knowledge that emerges in discussions about collective teaching and learning situations in which diverse individuals are situated. The paperwork that evidences and plans for this would therefore be a record of discussion about current priorities in a class situation. It would more closely resemble a set of minutes from a committee meeting, reflecting topics of discussion with action points. The key would be what is on the agenda. The possible questions might be as varied as the contexts which need to be drawn upon, but the discussion would recognize that people can only develop as part of a developing context which itself develops as the people do. So, for example, whatever issue was identified for a child or group of children (e.g., how to take advantage of some labels of opportunity or how we ensure we are responding to access preferences, or enabling collaboration, or exploring a topic of study), the discussion would need to consider all aspects of the community of provision, the space, support, staffing, students, strategies and systems (Rix et al., 2015). This would include availability of resources at all levels of the system, as well personal and professional experiences across the setting; what people feel they cannot do, as well as what they might be able to do, and why. Such an approach also lends itself to more regular discussions. It is not required once a year; it is not about singular children; it does not take place after numerous reports are written. It can happen once a week. Process can still be recorded. Responsibility can be allotted. But the paperwork can respond to the shifting reality of the uncertain classroom.

This more open use of paperwork, focused far more upon the context in which learning is taking place of itself can also be used to inform teacher-friendly funding and support decisions. It would not take much to shift to a class based funding model (Rix, 2009). This model would not be a conglomeration of individual assessments, each with a score that can be totted up to produce a banding total. This funding evaluation would begin with the collective context. It would identify not only the individuals and their access preferences, but also the support that was available and what was required to deliver collectively agreed goals for the whole class. It would still acknowledge the individual, but it would place them in the collective context of the other individuals around them, taking into account the skills and experiences of not just the pupils but the adults working alongside them. It could encourage possible learning opportunities when people are together. A class model could be flexible, with long term and short term budgetary processes. It would have local direction, but could still be controlled by traditional mechanisms of governance. Importantly for the established professions, allying class based funding models to collectively focused paper work would not remove expertise from the frame. It is not about saying that there is no place for speech and language therapists or educational psychologists and so forth, but it reduces the need for them to be churning out the paperwork, to be making final decisions and allows them to be part of a team, critical friends and/or active participants. It respects that they have wisdom to contribute to the benefit of all.

And?

If we wish to move beyond special and needs, we need to focus upon good teaching, authentic curricula, and interdependent assessments using labels of opportunity which recognize people's access preferences, and are supported by collectively focused paperwork along with class-based funding models. In doing this we can respond to market needs, build on the skills and experiences that practitioners have, reduce paperwork, allow experts to contribute directly to practice and provide robust mechanisms for governance and funding. If we wish to move beyond special and needs, we do not have to deny ourselves the security of certainty, but we should recognize it as a brief moment that will pass back into the uncertainty from which it came.

As the poet (me) said at the start of this all “*I have a problem.*” Well, it’s not just his (my) problem. We only have special and needs because we are determined to make things harder for ourselves.

However
In a world of uncertainty
we can rely upon
a different knowing
that all needs
are
our needs
not special.

References

- Ball, S., 1994. *Education Reform: A Critical and Post Structural Approach*. Open University Press, Buckingham/Philadelphia.
- British Psychological Society's Division of Clinical Psychology (BPS–DCP), 2013. Position Statement on the Classification of Behaviour and Experience in Relation to Functional Psychiatric Diagnoses.
- Cavendish, W., Harry, B., Menda, A.M., Espinosa, A., Mahotiere, M., 2016. Implementing response to intervention: challenges of diversity and system change in a high stakes environment. *Teach. Coll. Rec.* 118 (5), 1–36.
- Clough, P., 1995. Problems of identity and method in the investigation of special educational needs. In: Clough, P., Barton, L. (Eds.), *Making Difficulties—Research and the Construction of SEN*. Paul Chapman, London, pp. 126–142.
- Department for Education, 2020. *Special Educational Needs in England*. January 2020, London.
- Dulaney, S.K., 2013. A middle school's response-to-intervention journey: building systematic processes of facilitation, collaboration, and implementation. *NASSP Bull.* 97 (1), 53–77.
- Elliott, J.G., Gibbs, S., 2008. Does dyslexia exist? *J. Philos. Educ.* 42 (3–4), 475–491.
- Foucault, M., 1980. In: Gordon, C. (Ed.), *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*. Pantheon, New York, p. 16.
- Foucault, M., 1982. The subject and power (afterword). In: Dreyfus, H.L., Rabinow, P. (Eds.), *Michel Foucault: Beyond Structuralism and Hermeneutics*. Harvester Press, Brighton.
- Fuchs, D., Fuchs, L., 2006. Introduction to response to intervention. *Read. Res. Q.* 41 (1), 93–99.
- Goodey, C., 2011. *A History of Intelligence and "Intellectual Disability": The Shaping of Psychology in Early Modern Europe*. Ashgate, Farnham, MD.
- Gulliford, R., 1971. *Special Educational Needs*. Routledge & Kegan Paul, London.
- Hendricks, E.L., Fuchs, D., 2020. Are individual differences in response to intervention influenced by the methods and measures used to define response? Implications for identifying children with learning disabilities. *J. Learn. Disabil.* 53 (6), 428–443.
- Hollenweger, J., 2011a. Development of an ICF-based eligibility procedure for education in Switzerland. *BMC Public Health* 11 (Suppl. 5), S7. <https://doi.org/10.1186/1471-2458-11-S4-S7>.
- Hollenweger, J., 2011b. Teachers' ability to assess students for teaching and supporting learning. *Prospects* 41 (3), 445–457.
- Hollenweger, J., 2014. Beyond categories and labels: knowledge to support assessment for learning “disability”—a problem well put? In: Florian, L. (Ed.), *The SAGE Handbook of Special Education*, vol. 2, pp. 507–521.
- Holman, D., Lynch, R., Reeves, A., 2018. How do health behaviour interventions take account of social context? A literature trend and co-citation analysis. *Health* 22 (4), 389–410.
- Horner, R., Sugai, G., Fixsen, D., 2017. Implementing effective educational practices at scales of social importance. *Clin. Child Fam. Psychol. Rev.* 20 (1), 25–35.
- Hutchinson, J., 2021. How fairly and effectively Special Educational Needs and Disabilities (SEND) are identified? Trends in SEN identification: contexts, causes and consequences, SEN Policy Research Forum. *J. Res. Spec. Educ. Needs* 21 (1), 22–28.
- Kranzler, J.H., Yaraguchi, M., Matthews, K., Otero-Valles, L., 2020. Does the response-to-intervention model fundamentally alter the traditional conceptualization of specific learning disability? *Contemp. Sch. Psychol.* 24 (1), 80–88.
- Newman Thomas, C., Allen, A., Ciullo, S., Lembke, E., Billingsley, G., Goodwin, M., Judd, L., 2020. Exploring the perceptions of middle school teachers regarding response to intervention struggling readers. *Exceptionality*. <https://doi.org/10.1080/09362835.2020.1729765>.
- Norwich, B., 2016. Conceptualizing special educational needs using a biopsychosocial model in England: the prospects and challenges of using the international classification of functioning framework. *Front. Educ.* 1 (5). <https://doi.org/10.3389/feduc.2016.00005>.
- Norwich, B., Ylonen, A., Gwernan-Jones, R., 2014. Moderate learning difficulties: searching for clarity and understanding. *Res. Pap. Educ.* 29 (1), 1–19.
- OECD, 2012. CX3.1 Special Educational Needs (SEN). Accessed from: <http://www.oecd.org/social/family/50325299.pdf>. (Accessed 16 February 2013).
- Oliver, M., 1996. User Involvement in This Voluntary Sector, Paper Prepared for Submission to the Commission on the Future of the Voluntary Sector.
- Preston, A., Wood, C., Stecker, P., 2016. Response to intervention: where it came from and where it's going. *Prev. Sch. Fail.* 60 (3), 173–182.
- Richardson, J., Powell, J., 2011. *Comparing Special Education: Origins to Contemporary Paradoxes*. Stanford University Press, Stanford.
- Riddell, S., Tisdall, K., Kane, J., Mulderrig, J., 2006. Literature Review of Pupils With Additional Support Needs Final Report to the Scottish Executive Education Department Centre for Research in Education Inclusion and Diversity (CREID).
- Rix, J., Matthews, A., 2014. Viewing the child as a participant within context. *Disabil. Soc.* 29 (9), 1428–1442.
- Rix, J., Sheehy, K., Fletcher-Campbell, F., Crisp, M., Harper, A., 2013. Exploring provision for children identified with special educational needs: an international review of policy and practice. *Eur. J. Spec. Needs Educ.* 28 (4), 375–391.
- Rix, J., Sheehy, K., Fletcher-Campbell, F., Crisp, M., Harper, A., 2015. Moving from a continuum to a community reconceptualizing the provision of support. *Rev. Educ. Res.* 85 (3), 319–352.
- Rix, J., 2007. Labels of opportunity—a response to Carson and Rowley. *Ethical Space* 4 (3), 25–27.
- Rix, J., 2009. Statutory assessment of the class? Supporting the additional needs of the learning context. *Int. J. Incl. Educ.* 13 (3), 253–272.
- Rix, J., 2015. *Must Inclusion Be Special? Rethinking Educational Support Within a Community of Provision*. Routledge, London.
- Rix, J., 2020. Our need for certainty in an uncertain world: the difference between special education and inclusion? *Br. J. Spec. Educ.* 47 (3), 283–307.
- Rix, J., 2021. Inclusive relationships—creating the space for each other. In: Beaton, M., Codina, G., Wharton, J. (Eds.), *Leading on Inclusion: The Role of the SENCO*. Routledge, London.
- Roaf, C., Bines, H. (Eds.), 1989. *Needs, Rights, and Opportunities: Developing Approaches to Special Education*. Falmer Press.
- Saragoça, A., Candeias, A., Rosário, A., 2013. Implementation of the International Classification of Functioning, Disability and Cealth (ICF) in the Portuguese Educational System: Attitudes and Training Needs of Special Education Teachers'. *Transylv. J. Psychol. (Special Issue)*, 223–243.
- Slee, R., 2008. Beyond special and regular schooling? *Int. Stud. Sociol. Educ.* 18 (2), 99–116.

- Stahmer, A., Suhrheinrich, J., Schetter, P., McGee Hassrick, E., 2018. Exploring multi-level system factors facilitating educator training and implementation of Evidence-Based Practices (EBP): a study protocol. *Implement. Sci.* 13 (3), 1–11.
- Taylor, B., Francis, B., Archer, L., Hodgen, J., Pepper, D., Tereshchenko, A., Travers, M., 2017. Factors deterring schools from mixed attainment teaching practice. *Pedagog. Cult. Soc.* 25 (3), 327–345.
- Timimi, S., 2021. The social cultural construction of Autism, trends in SEN identification: contexts, causes and consequences, SEN Policy Research Forum. *J. Res. Spec. Educ. Needs* 21 (1), 28–32.
- Tomlinson, S., 1982. *A Sociology of Special Education*. Routledge, London.
- Tomlinson, S., 2012. The irresistible rise of the SEN industry. *Oxf. Rev. Educ.* 38 (3), 267–286.
- Wilshire, C.E., Clack, T.W.S., 2021. Symptom descriptions in psychopathology: how well are they working for us? *Clin. Psychol. Sci.* 9 (3).
- Wilson, J., 2002. Defining "special needs". *Eur. J. Spec. Needs Educ.* 17 (1), 61–66.

Developing inclusive education for disabled children and students: advocacy in hard times

Richard Rieser, World of Inclusion Ltd London, UK

© 2023 Elsevier Ltd. All rights reserved.

Introduction	293
Bringing the paradigm shift into education of disabled children and students	294
Barriers to inclusive education and ways to address them	297
The index for inclusion has three key dimensions	300
Developing inclusive education theory into practice	301
Filming inclusive practice in English schools	302
Champions of inclusive education	302
Identifying and removing barriers to inclusive education	303
Possible solutions	304
Conclusion	304
References	305

Introduction

In 2022, state education systems around the world should be making good progress to implementing inclusive education, in theory and practice, with Sustainable Development Goal No. 4 (2015 (17 goals agreed)) and Article 24 of the United Nations Convention on the Rights of Persons with Disabilities (2006) committing country leaders to developing an inclusive education system.

Historically, mass state education in the developed world grew to meet the requirements for a literate and then increasingly technological workforce. The more progressive and liberal reformers viewed education as a Human Right, a means of overcoming age-old traditions and for coming generations to have the means for less oppressive and more fulfilled lives. Foremost among these excluded groups are women and girls, ethnic and linguistic minorities and disabled people. The last was seen as the most difficult to address. Education for Blind and Deaf children was first to be addressed, then those with physical impairments and lastly those with intellectual impairments and psycho-social and neuro-diverse conditions such as autism. Education for these groups was seen as a good idea. Students either had to “pass” as non-disabled and manage the environment, curriculum and social relationships on non-disabled people’s terms (integration); or their needs were seen as too disruptive to the education of non-disabled children and they were placed on a medicalized path of special education leading in most cases to segregated special schools (segregation). The dual education system that emerged over time was expensive. It required two sets of buildings, two sets of staff and two types of curricula. The accounts of survivors of the segregated system overwhelmingly talk of the negative impact on their self-esteem and lack of useful social skills to live and work in the world, and in most cases of the very low educational expectations.

The adoption of the Universal Declaration of Human Rights, 1948 (including the right to education: Article 26), the setting up of the United Nations, and the decolonization of territories leading to many new countries has led to contradictions in education systems set up in these developing countries. Most new countries replicated the dual mainstream/segregated system in developed countries or more likely excluded disabled children from education or allowed casual integration leading to a situation where large numbers of disabled children do not access state education or where they drop out in the early grades.

In carrying out a review of the state of education systems across the 54 countries of the Commonwealth with regard to implementing Inclusive education, as outlined in Article 24 of the UNCPRD (2006), the author found that even in countries such as India, Bangladesh, Pakistan, Nigeria or South Africa, where policies existed for including disabled children in education, examples of good practice were rare and usually due to small scale, Non-Government Organizations (Northern) projects (Rieser, 2008/2012). Fundamentally it was a lack of political will backed by lack of funding and traditional and old-fashioned ideas about disabled children and their capabilities.

The Covid-19 Pandemic is very likely to have a long-term impact on the education of disabled children as they were particularly affected by lockdown, impacting on support arrangements, difficulties with accessing on-line learning, getting out of the habit of schooling and so not returning to class. *Then There Was Silence* (Lunt, 2021) demonstrated in a UK-wide survey over the first 18 months of Covid: children and families have been isolated and abandoned and not been listened to. Covid restrictions meant services were stopped or reduced, and many are still slow to return. Mental health and wellbeing of all the family has deteriorated, children’s conditions have worsened and their needs have become more complex; delays in assessments mean needs haven’t been identified. A World Bank (World Bank, 2021) study showed that in the Spring of 2020, at the peak of the pandemic, 180 countries experienced temporary school closures, affecting 1.5 billion young people or 94% of the student population worldwide. For more than half of these students, schools would remain closed for more than seven months. A year after the worldwide spread of COVID-19, in March 2021, schools remained closed in more than 50 countries. Governments tried to respond to these unprecedented times

by employing remote learning initiatives using both digital and nondigital formats. However, inadequate infrastructure or the absence of trained teachers who could quickly transition to online models of teaching often limited the effectiveness of such responses especially for disabled children. Disabled children have been slower to return to school with many dropping out.

The Global Crisis in 2008 has reduced funding for state education, both in real terms within states and in terms of transfer of aid from richer to poorer countries. The ideology of Neo-Liberalism has reinforced this pattern of concentration of wealth among the very rich. The gap between the rest of the population widening, part of which has been the systematic, real terms reduction of egalitarian services like state education where they exist and a reduction in international state aid to those areas where they do not exist. The National Audit Office (July 2021) looking at school funding in England found that:

Although the Department's total funding for schools increased by 7.1% in real terms between 2014–15 and 2020–21, the growth in pupil numbers meant real-terms funding per pupil rose by 0.4% (Paragraph 7) ...

In recent years, there has been a relative re-distribution of funding from the most deprived schools to the least deprived schools (Paragraph 16).
National Audit Office (2021).

People on the other hand are not just pawns in an economic and ideological system and their love for their families and fellow humans can, against the odds, lead to developments with glimpses of an inclusive and an egalitarian future which many do/would materially and ethically benefit from. In this chapter I will explore how these contradictions impact on the development of inclusive education for disabled people in England and globally from the point of view of a disabled pedagogue and inclusion advocate.

Disabled people, rather than **persons with disabilities** will be used throughout this chapter as the oppression of **disabilism** is rooted in traditional, overwhelmingly negative myths, reinforced by the development, over the last 300 years of the **medical or individual view of disabled people**. In this view, a disabled person's medical condition was thought to be the root cause of their exclusion from society, an approach now referred to as the **medical or individual model** of disability. Using the thinking of the Disabled People's Movement, where the **social model of disability** identifies the barriers of environment, attitude and organization which disable people with impairments and lead to prejudice and discrimination. Our impairments can be physical, metabolic, sensory, intellectual, mental or psycho-social, but we view ourselves as united by a common oppression and so are proud to identify as **disabled people** rather than **people with disabilities**. To call ourselves "persons with disabilities" is to accept that we are objects and powerless. This is logical but very much a minority view, since the United Nations adopted a Convention on the Rights of Persons with Disabilities UNCRPD, in 2006.

Bringing the paradigm shift into education of disabled children and students

The UNCRPD (2006) does address a paradigm shift from charity and medical views toward a human rights approach to disability and disability rights.

Persons with disabilities as subjects of **rights**, not objects of charity

The Convention takes to a new height the movement from the treatment of persons with disabilities as "objects" of charity, medical treatment and social protection towards viewing persons with disabilities as "subjects" with rights who are capable of claiming those rights and making decisions for their lives based on their free and informed consent as well as being active members of society. The Convention gives universal recognition to the dignity of persons with disabilities.

United Nations Enable UNDESA (2007).

The United Nations UNCRPD Committee is charged with overseeing the implementation of the UNCRPD (2006). It is elected by United Nations member states who have ratified the Convention, currently 184 (February 2022). It comprises 20 experts, mainly disabled people. In 2016 the Committee called a General Day of Discussion because they were finding that member countries reporting to them were not understanding what was required of them to implement Article 24 and an inclusive education system, but instead were continuing with a special education system, mainly relying on segregation in separate special schools. Within many poorer countries, the large majority of disabled children were not attending school. Following consultation with state parties and civil society organizations, the UNCRPD Committee issued a General Comment (No. 4) in 2016 clarifying what state parties were required to do to implement Article 24 of the UNCRPD (United Nations, 2016, CRPD/C/GC/4).

It is worth quoting at some length from the opening passages of General Comment (No.4) and providing illustrative historic exemplars to provide a context:

2. Recognition of inclusion as the key to achieving the right to education has strengthened over the past 30 years and is enshrined in the Convention on the Rights of Persons with Disabilities, the first legally binding instrument to contain a reference to the concept of quality inclusive education. Sustainable Development Goal 4 too affirms the value of inclusive, quality and equitable education. Inclusive education is central to achieving high-quality education for all learners, including those with disabilities, and for the development of inclusive, peaceful and fair societies. Furthermore, there is a powerful educational, social and economic case to be made ... Only inclusive education can provide both quality education and social development for persons with disabilities, and a guarantee of universality and non-discrimination in the right to education.

3. Despite the progress achieved, however, the Committee is concerned that profound challenges persist. Many millions of persons with disabilities continue to be denied the right to education and for many more education is available only in settings where persons with disabilities are isolated from their peers and where the education they receive is of an inferior quality.

The idea of barriers and dismantling them is now seen as an important component of delivering human rights for disabled people, both in education and in/with all socially constructed organizations, attitudes and environments. This is a relatively new idea which came out of the struggles of physically disabled people in the 1970s, to live independently and to challenge oppressive attitudes that viewed them through a medical model lens, confining many of them in care homes. Like all good social ideas, it happened in a number of places around the world at about the same time.

The Zimbabwean disability movement for equality was born in institutions run by Jairos Jiri in 1975, but he did not support it. Like many people at that time, he saw the call for involvement, participation and increased opportunities as a threat to his charity model. The movement emanated from disabled people, like Alexander Phiri and Joshua Malinga who went on to become a politician against white rule and Mayor of Bulawayo from 1993 to 1995. These activists formed and registered the National Council of Disabled Persons Zimbabwe which started as Kubatsirana Welfare Society and later National Council for the Welfare of Disabled, while still in the institutions.

Even from day one, I had an in-born attitude not to accept the attitudes at Jairos Jiri. These ideas were very bad; for example, disabled people were told when to eat, when to sleep, that we couldn't make love—it was banned! Clients do not have any say about how the services are provided or about which services they receive The services available to clients contribute to the under-development of disabled persons; but they help to contribute to the development and conscientisation of activists of the Disability Rights Movement.

Joshua Malinga.

In 1980, Malinga attended Disabled Peoples' International's Winnipeg World Congress in Canada. This marked internationalization of disability work in Zimbabwe and its spread across Southern Africa.

In the early 1970s, the Union of Physically Impaired Against Segregation (UPIAS) was formed around Paul Hunt, a disabled resident of a Leonard Cheshire Home. By 1974 they produced a policy statement with the following aim:

The Union aims to have all segregated facilities for physically impaired people replaced by arrangements for us to participate fully in society. These arrangements must include the necessary financial, medical, technical, educational and other help required from the State to enable us to gain the maximum possible independence in daily living activities, to achieve mobility, to undertake productive work, and to live where and how we choose with full control over our lives.

Union of the Physically.

UPIAS went on to be the catalyst for the formation of the British Council of Disabled People's Organizations (BCODP). Similar Movements can be identified in New York in 1973, in Malaysia, Sweden and elsewhere in this period.

This crystalized when 400 representatives of organizations led by disabled people walked out of the Rehabilitation International Conference in Winnipeg in 1980. They set about preparing for the inaugural conference of a new global organization to represent disabled people. In 1981, Disabled Peoples' International, at its founding World Summit in Malaysia, adopted the following statement:

Impairment is the loss or limitation of physical, mental or sensory function on a long-term or permanent basis.

Disability is the loss or limitation of opportunities to take part in the normal life of the community on an equal level with others due to physical and social barriers.

Disabled Peoples' International, 1981 (Barnes, 1991).

This definition is partly recognized in the United Nations Convention on the Rights of Persons with Disabilities (2006), pointing to the evolving nature of what is included, but clearly breaking with the old ‘medical model’ approach and focusing on interaction with barriers.

Persons with disabilities include those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others.

UN (2006).

To understand this transformation referred to above we need to be complexly clear about the impacts of these two approaches. See **Figs. 1A, B and 2** below.

The first Figure (A) is based on a “medical model” approach with the focus on fixing the person. However, many impairments are for life and so not fixable. This is not to say that disabled people do not need the best rehabilitation and medical support, but we do not need this to be the primary focus, which needs to be on changing the structures in society and the barriers they create. Figure (B) holds out a way forward which is not about primarily changing us as disabled people, but dismantling the socially created barriers as part of our Human Rights.

Myself and Micheline Mason developed an analysis of applying this transformation to schools in 1990 ([Rieser and Mason, 1990](#)) in *Disability Equality in the Classroom* for the Inner London Education Authority and then in *Altogether Better* in 1994 ([Mason and Rieser, 1994](#)), a training pack for Comic Relief to develop Inclusion in Education.

In 2008, when writing the Commonwealth Book, these ideas were further adapted and then presented to the Global Summit on Disability in 2018.

Fig. 3 examines the different geographic expression of thinking about disabled children and educational provision.

Exclusion is the majority state for disabled children in the South and Segregation is the majority in the North.

Building on the above analysis, General Comment No 4 makes clear the key parts of a definition of inclusive education and different organizational forms of education for disabled children **exclusion, segregation, integration and inclusion** and that only inclusion can meet disabled children’s Human Rights.

11. The Committee highlights the importance of recognizing the differences between exclusion, segregation, integration and inclusion. **Exclusion** occurs when students are directly or indirectly prevented from or denied access to education in any form. **Segregation** occurs when the education of students with disabilities is provided in separate environments designed or used to respond to a particular impairment or to various impairments, in isolation from students without disabilities. **Integration** is the process of placing persons with disabilities in existing mainstream educational institutions with the understanding that they can adjust to the standardized requirements of such institutions. **Inclusion** involves a process of systemic reform embodying changes and modifications in content, teaching methods, approaches, structures and strategies in education to overcome barriers with a vision serving to provide all students of the relevant age range with an equitable and participatory learning experience and the environment that best corresponds to their requirements and preferences. Placing students with disabilities within mainstream classes without accompanying structural changes to, for example, organization, curriculum and teaching and learning strategies, does not constitute inclusion. Furthermore, integration does not automatically guarantee the transition from segregation to inclusion.

Paragraph 11, General Comment No. 4.

In developing countries with low or middle income, planners and politicians should try **not** to set up a continuum of provision, but try to provide support and services to the school by use of itinerant specialist teachers and local district support centers. In terms of restructuring the provision of education especially in more developed countries this means moving from a continuum of provision approach to one where a constellation of services is around the student and their teacher in the regular or mainstream school, depending on the level of economic development. See **Fig. 4**.

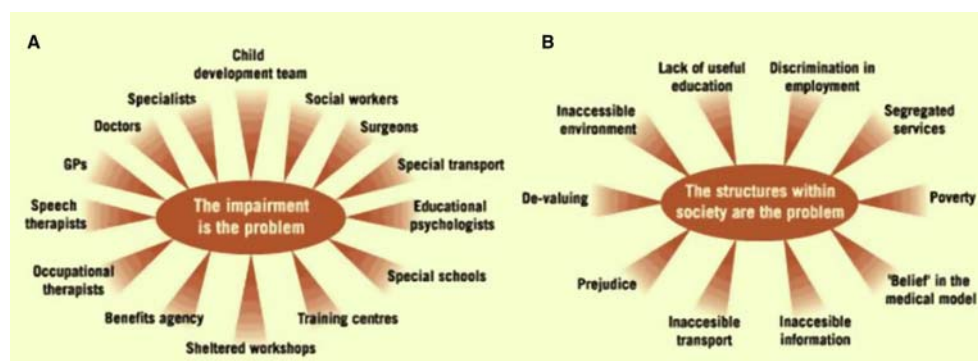


Fig. 1 (A) Disabled people as passive receivers of services aimed at cure or management. (B) Disabled people as active fighters for equality working in partnership with allies. From [Rieser \(2004\)](#).

Medical model thinking	Social model thinking
Child is faulty	Child is valued
Diagnosis	Strengths and needs defined by self and others
Labelling	Identify barriers and develop solutions
Impairment becomes focus of attention	Outcome-based programme designed
Assessment, monitoring, programmes of therapy imposed	Resources are made available to ordinary services
Segregation and alternative services	Training for parents and professionals
Ordinary needs put on hold	Relationships nurtured
Re-entry if normal enough OR permanent exclusion	Diversity welcomed, child is included
Society remains unchanged	Society evolves

Fig. 2 The two routes identified start from basic values and underline that the denial of disabled children's human rights starts with a lack of viewing us as human. Adapted from [Mason and Rieser \(1994\)](#).

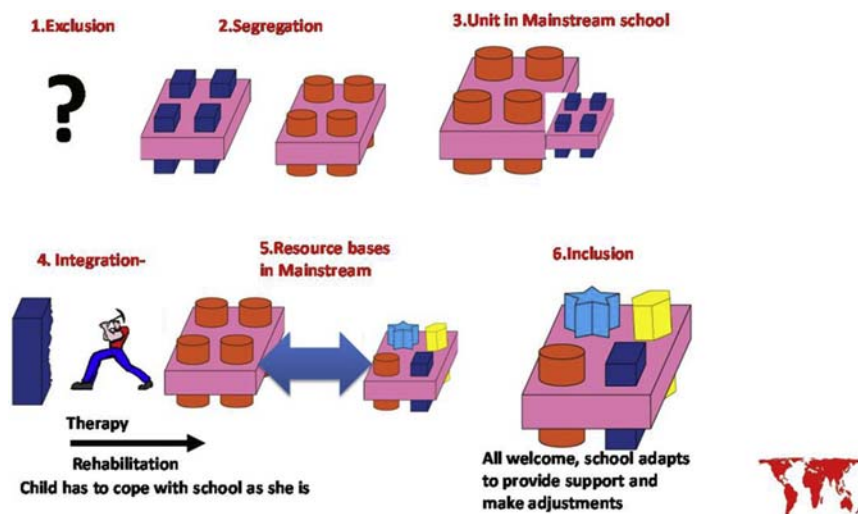


Fig. 3 Form of organization of education for children with disabilities ([Rieser, 2018](#)).

Barriers to inclusive education and ways to address them

If barriers and removing them is the key to Inclusive Education, what are the key barriers that obstruct the development of Inclusive Education? General Comment No. 4 identifies them, pointing out they can be attributed to multiple factors, that include:

- The failure to understand or implement the human rights model of disability, according to which barriers within the community and society, rather than personal impairments, exclude persons with disabilities;
- Persistent discrimination against persons with disabilities, compounded by the isolation of those still living in long-term residential institutions, and low expectations about those in mainstream settings, allowing prejudices and fear to escalate and remain unchallenged;
- Lack of knowledge about the nature and advantages of inclusive and quality education and diversity, including regarding competitiveness, in learning for all; lack of outreach to all parents; and lack of appropriate responses to support requirements, leading to misplaced fears and stereotypes that inclusion will cause a deterioration in the quality of education or otherwise have a negative impact on others;
- Lack of disaggregated data and research (both of which are necessary for accountability and program development), which impedes the development of effective policies and interventions to promote inclusive and quality education;
- Lack of political will, technical knowledge and capacity in implementing the right to inclusive education, including insufficient education of all teaching staff;
- Inappropriate and inadequate funding mechanisms to provide incentives and reasonable accommodations for the inclusion of students with disabilities, inter-ministerial coordination, support and sustainability

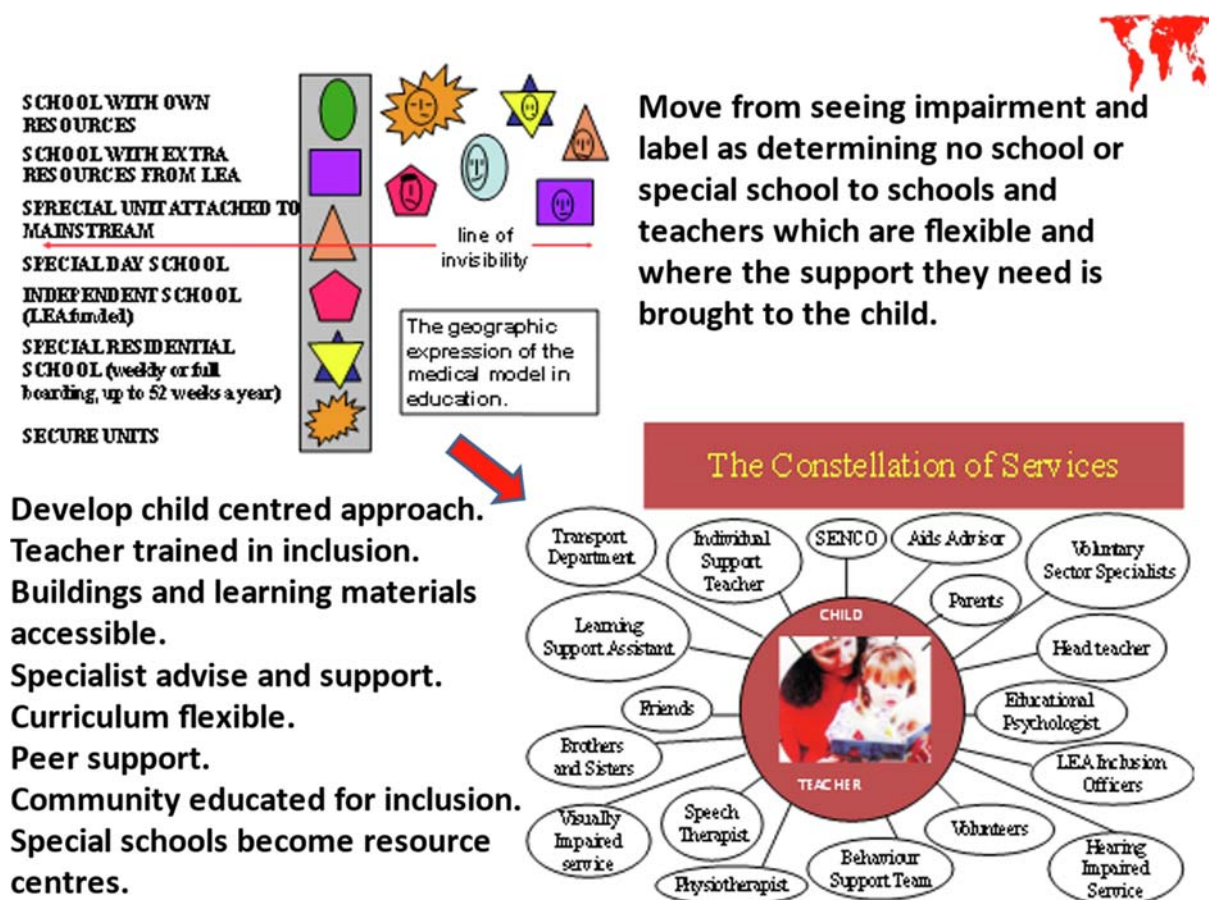


Fig. 4 The continuum of provision.

(g) Lack of legal remedies and mechanisms to claim redress for violations (Paragraph 4, General Comment No. 4, 2016)

One of the big barriers that has obstructed real moves to Inclusive Education are normative testing and outcome tables, such as Organization for Economic Cooperation and Development, (OECD's) PISA tests of 15-year-olds which are very often picked up by politicians as a false measure of how competitive their school system and therefore economy is likely to be. PISA is the OECD's Program for International Student Assessment. PISA measures 15-year-olds' ability to use their reading, mathematics, science knowledge and skills to meet real-life challenges and is run every three years (OECD). As a UK Coalition Government White Paper (2010) makes clear:

So much of the education debate in this country is backward looking: have standards fallen? Have exams got easier? These debates will continue, but *what really matters* is how we're doing compared with our international competitors. That is what will define our economic growth and our country's future. The truth is, at the moment we are standing still while others race past.

DfE (2010).

While efforts have been made to be more inclusive in PISA tests, as many students with SEND are excluded as included, with a particularly high number excluded with cognitive or intellectual impairment or using sign language as their medium of instruction. In addition, many small schools are excluded which will include special schools. In the UK in 2008 all special schools and Pupil Referral Unit students were excluded from sampling for the test cohort (Smith and Douglas, 2014). A more recent investigation by the European Network for Independent Living (Efthymiou, 2021) found that the PISA test stands for Exclusion not Inclusion. For example, students will be "included when their needs are not serious enough to be a barrier to their participation."

PISA also does not encourage schools to provide accommodations to students with access needs to participate in the test. This is made clear in the 2020 PISA project manual (PISA, 2020), which states that the test can only allow "... limited modifications and accommodations to assist students with special educational needs". By reinforcing the stereotype that only some disabled people can be educated, the PISA test reinforces the idea that it is "natural and expected" for disabled students and others with support needs to be excluded from the society.

The Exam Factory (NUT, 2015) system created in England by the UK Government over the last 12 years based on high stakes testing, a narrowing of the curriculum, more testing of knowledge, a reduction in methods of testing, increasing the status of exams, downgrading or stopping continuous assessment, course work and folders of work have all made it far more difficult for a range of disabled students to demonstrate what they can achieve.

Results in England from public exams show that despite a range of reforms to KS4, the gap between disabled and non-disabled students remains high and is widening. The attainment difference between pupils with SEN compared to pupils with no identified SEN remains the largest difference of all characteristic groups. Pupils with SEN perform markedly worse than pupils with no identified SEN across all headline measures of attainment. The attainment differences remained broadly similar to 2018, however there was a slight widening in the difference in EBacc entry between groups (0.7% points) with 13.5% of pupils with SEN entering EBacc in 2019, up from 12.4% in 2018 (assets).

New attainment indicators were introduced in 2014/15 at Key Stage 4—Attainment 8 and Progress 8. The average attainment 8 score for pupils with SEN in 2019/20 was 30.7 compared to 53.7 for those with no SEN. The average attainment 8 score for pupils with SEN support was 36.4 and for pupils with an EHC plan was 15.2 (educational).

Another barrier is attitudinal and particularly parental attitudes to the education of their disabled children. In England the factors identified above together with an ideological position of the Conservative Party who have been in power for the last 12 years. David Cameron, former Prime Minister, who had a disabled son and was strongly against him being included, promoted a policy of being “against the bias toward inclusion.” The failure to promote Inclusive Education linked with the narrowing of curriculum opportunities and two other factors, has led to an unprecedented number of disabled children being out of mainstream school with increasing segregation. Firstly, the Academies program taking schools away from Local Authority or Municipal control, which are handed over to be run increasingly by Multi Academy Trusts. These schools have no democratic control by the local community and parents, do not have to employ trained teachers (other than the Special Educational Needs Co-ordinator), have their budgets top sliced from the Local Authority’s and are provided directly by Government. In addition, Local Authorities are unable to build any new community schools, only Academies to meet need.

The Government is putting their faith in Academies to bring about improvement. But we know from work by Brahm Norwich and his colleagues at Exeter (Lin et al., 2020), analyzing the National Pupil data base, that, in the last 10 years, those schools that have become Academies, especially converters—those schools that are usually rated outstanding by OFSTED that have chosen to be academies, have reduced the numbers of students identified with SEN at a faster rate than the community schools. This has partly led to the big acceleration of numbers in special schools.

Secondly, despite increases, in real terms school budgets have reduced as the numbers of students has increased. This has hit SEND budgets particularly hard as these are not ring-fenced and many of the essential support staff who make inclusion work have been made redundant. Schools in England have delegated budgets and are responsible for hiring and firing staff under Local Management of Schools, which was introduced in 1988. The National Education Union who represents the largest group of teachers and those who work in schools commented on a recent White Paper on Leveling Up, “While the National Education Union welcomes any new investment in schools” it is vital to note, as the National Audit Office (National Audit Office, 2021) has pointed out, “there has been a relative re-distribution of funding from the most deprived schools to the least deprived schools.” We can see that many of the areas now targeted for support have been among the hardest hit by education cuts over the last decade—on the Government’s own watch, and entirely of its own making” (Anonymous, 2022).

So many parents whose first choice would be mainstream inclusion in England are now voting with their feet or going to the First Tier Tribunal in increasing numbers to place their children in special schools. The pressure cooker atmosphere of high stakes testing, high levels of disabilist bullying, the lack of appropriate curriculum, together with increasing levels of mental health issues among pupils and pressure from schools to leave, have forced the highest level ever of children with SEND to be out of the mainstream in 2021.

Table 1 shows us that the number of 5–16-year-olds identified with a special educational needs statement or latterly an Education Health and Care Plan has gone up from 2.4% to 3.6% of the school population, even as the school population has risen by more than 1 million. This number plateaued at 2.8% from 2006 to 2010 and then rose again. A conscious policy to increase inclusion in mainstream schools of disabled children, introduced by the Labor Government of 1997 did have an impact until 2006 when 59.1% of those identified with a Statement were in mainstream provision. This figure went down to 49.2% in 2016 and in the last year has gone back to 50.4%, but there were far more children impacted. From 2010 we had a Conservative Government (in Coalition 2010–2015) to the present with a policy of choice and setting up academies, and are lukewarm on their international obligation to inclusion. If those in Pupil Referral Units and all types of special schools with and without a plan are added, this gives the highest ever recorded figure of 1.94% educated in separate (segregated) settings in 2021 in England. So, Government policies do impact on the level of Inclusion!

Interestingly, Scotland and Wales who have devolved powers in Education have not introduced any of the above measures and although strapped for cash, are both committed to develop fully inclusive systems as the ruling parties in their governments are committed to incorporating the UNCPRD (2006) fully into their Domestic Law.

So, what does the UNCPRD Committee think Inclusion is?

10. Inclusive education is to be understood as:

- (a) A fundamental human right of all learners. Notably, education is the right of the individual learner and not, in the case of children, the right of a parent or caregiver. Parental responsibilities in this regard are subordinate to the rights of the child;

Table 1 Secondary schools, England.

Type	Date	Number schs.	% SEN support	% EHCP/statement
Maintained (LA funded)	2003	4606	14	4.11
	2010	3955	21	3.7
	20,125	1951	13	5.6
Sponsored academy (Forced, failing)	2003	3	29	2.34
	2010	201	29	2.33
	2015	556	17	2.30
Converter academe (Self-chosen)	2012	841	17	2.07
	2015	1471	12	2.49

Note: This table reports the descriptive statistics for secondary schools in our sample.
Based on Table 2 in article.

Table 2 England. Children with SEN statement or EHC plan and type of school attended.

Year (A)	%Mainstream primary & secondary schools (B)	%State funded special schools, PRUs, NMSS* & independent schools (C)	Total students with a EHC plan /statement+ (D)	Nos of pupils in segregated settings with no EHC plan /statement (E)	Total school population (F)	%Total school population segregated Number C + E /total school population (F) (G)
2021	50.4	49.6	325,618	11,665	8,911,887	1.94
2016	49.2	50.8	236,805	17,430	8,559,540	1.61
2010	55.2	44.8	225,945	16,655	8,098,360	1.45
2006	59.1	40.9	236,750	15,110	8,215,690	1.36
1997	57.2	42.8	234,629	11,730	8,194,964	1.37
1994	51.9	48.1	195,410	10,462	7,882,835	1.32

Source: DFS, DFES, DfE Annual SEND Statistics*NMSS Non-Maintained Special Schools, PRUs Pupil Referral Units. +Statement was a legally binding certification of SEN and provision following a multidisciplinary assessment. In 2014 this was replaced with a statutory Education, Health and Care Plan (Rieser, 2022).

- (b) A principle that values the well-being of all students, respects their inherent dignity and autonomy, and acknowledges individuals' requirements and their ability to effectively be included in and contribute to society;
- (c) A means of realizing other human rights. It is the primary means by which persons with disabilities can lift themselves out of poverty, obtain the means to participate fully in their communities and be safeguarded from exploitation ([Committee on Economic](#)). It is also the primary means of achieving inclusive societies;
- (d) The result of a process of continuing and proactive commitment to eliminating barriers impeding the right to education, together with changes to culture, policy and practice of regular schools to accommodate and effectively include all students.

Change in school systems must be primarily led by regular or mainstream schools changing themselves to accommodate a wider diversity of pupils. But as we have seen above, the conditions that prevail from Government and to a less effect from Local Authorities do set the climate in which school change occurs. Focusing on (d) above the *Index for Inclusion* was a self-evaluation tool developed in UK, but then adapted for use in over 70 countries with 34 different language translations, for schools to self-evaluate how inclusive they are by asking a series of questions of different stakeholders in the school-teaching staff, support staff, parents, students and leaders about the culture, policies and practices of schools. The school as a whole can then identify a plan to become more inclusive, by removing the barriers so identified ([CSIE, 1997–2010](#)). As a member of the original working group of the *Index for Inclusion*, the author was the only disabled person and we had strong debates about whether it should be based on the social model of disability. In the end the Center for Studies of Inclusive Education agreed and Mel Ainscow and Tony Booth, the main authors, had to agree as can be seen in an endnote of the published "Index." The "Index" is about including all students but without the conscious underpinning of the social model it would not have included the wide variety of students with different impairments.

The index for inclusion has three key dimensions

- A. Creating an Inclusive Culture
 - 1. Building Community
 - 2. Establishing Inclusive Values
- B. Producing Inclusive Policies
 - 1. Developing Schools for All
 - 2. Organizing Support for Diversity

C. Evolving Inclusive Practice

1. Orchestrating Learning
2. Mobilizing Resources.

This remains the best school self-evaluation tool and now with greater emphasis on developing inclusion it should be developed and rolled out to schools globally. Applying the paradigm shift from traditional/charity/medical model approach to the social model/human rights approach to schools and educational systems is the crucial transformation that is required to bring about critical understanding of what needs to change.

Developing inclusive education theory into practice

The theory of developing inclusive education is clear, but putting it into practice is an altogether more complex process, strewn with barriers and unforeseen obstacles.

The **London Borough of Newham** is an early example in England. Below is a write-up from the *Disability Equality in Education Developing Inclusive Education Course Book* (*Disability Equality in Education*, 2002) (Table 3).

The London Borough of Newham believes in the inherent equality of all individuals irrespective of physical or mental ability. It recognises, however, that individuals are not always treated as equals and that young people with disabilities experience discrimination and disadvantage. The Council believes that segregated special education is a major factor causing discrimination. We therefore believe desegregating special education is the first step in tackling prejudice against people with disabilities and other difficulties. They have been omitted from previous Equal Opportunities initiatives, and it is now obvious that our aim of achieving comprehensive education in Newham will remain hindered while we continue to select approximately 2% of school pupils for separate education ...

It is also the right of pupils without disabilities or other difficulties to experience a real environment in which they can learn that people are not all the same and that those who happen to have a disability should not be treated differently, any more than they would be if they were of a different ethnic background. It is their right to learn at first-hand about experiences which they will possibly undergo in future, either themselves or as parents ...

Desegregating special education and thus meeting the needs of statemented children in mainstream schools will also contribute, by the entry of expert qualified staff into mainstream schools, to improved provision for the considerable number of children who already experience difficulties.

London Borough of Newham (1987).

In 1984 a group of parents of disabled children ran for and were elected onto the Council with the express wish of seeing the ending of segregated special education. They achieved their aim in a Council policy which recognized the rights of children, whatever their needs, to learn together (see above).

The Borough's latest policy has a goal of making it possible "for every child, whatever special educational needs they may have, to attend their neighborhood school". In the 12 years, 1984–1996, the number of special schools in the borough has been reduced from 8 to 2 (now 1) and the number of children segregated in special education has dropped from 913 to 206. Parents are becoming increasingly confident in the ability of their neighborhood school to meet diverse needs and teachers have signed an agreement on inclusive education. This was achieved in an educational and political climate that was hostile to this process. Resourced schools were set up to meet certain needs in mainstream schools as a response to parental concerns. These are now planned to be phased out as Newham moves to inclusive neighborhood schools.

The process from the start envisaged radically changing mainstream schools rather than fitting children with SEN into the existing system. An independent report commented that having to cater for children with serious learning difficulties helped schools

Table 3 Percentage of pupils with statements in different kinds of provision.

Placement	1986 (%)	1996 (%)	2001 (%)
Newham special schools	68	9	1
Out-borough residential schools	11	3	6
Out-borough day schools	6	5	7
Mainstream schools	7	63	79
Resourced mainstream schools	8	20	7

The London Borough of Newham (Jordan and Goodey, 1996) provides a useful indicator of how such moves toward inclusion can occur in a poor, multi-cultural, inner-city area.

make better provision for all pupils. This is born out in 1997–2000. Newham schools have the biggest improvement nationally in the GCSE results of all pupils in grades A-G. Many children labeled as having severe learning difficulties are now passing exams. In addition, the numbers of exclusions have been falling while they have been rising in most other areas of the country. The process is continuing with some schools inclusive, others realizing they have a long way to go and some struggling. The Local Education Authority/(Municipality) has now appointed four monitoring officers to pro-actively address this process of developing inclusion from integration.

Newham schools have also had the lowest levels of exclusions in London.

Postscript: 20 years on. The introduction of Academies, cuts in budget impacting on central support teams and school funding for additional needs, pressure from a minority of parents who want segregated provision for their autistic children and not maintaining systematic training on inclusion is all leading to pressures to restart special schools. However, the legacy of the reforms is that Newham still has the lowest level of students with EHC Plans and the lowest number of students educated in special schools of any English Local Authority in 2021.

Filming inclusive practice in English schools

In 2003, the Disabled People's Organization and Charity, *Disability Equality in Education* of which I was the founder and CEO, was approached by the Department of Education to find a range of mainstream schools across England that were good at making Reasonable Adjustments for a wide range of disabled pupils and students. We had previously been supported by a government grant to develop a network of disability equality trainers and materials to work with school staff. We were providing this training in many areas and by our closure in 2008 had provided it to more than 100,000 education staff. In 2002 the Education Amendment Act had become law and extended the duty of making reasonable adjustments for disabled pupils and students in all educational settings. We wrote to 10,000 mainstream schools with a questionnaire about their good practice. Over 400 replied wishing to have us visit and film their adjustments for students with sensory impairment, physical impairments, intellectual impairments, behavioral difficulties and medical needs. We selected around 90 some of whom had been making these adjustments for some time and others were just beginning. We eventually selected 41 schools around the country and set up 2 tranches of visits with a film crew. A telephone interview was held to set up a schedule of classes to visit and staff, parents and students to interview. Parental permissions to film were also obtained. Over 100 h of film was eventually edited down to 5.5 h and put on DVDs which were made available to all schools (DFES, 2006). The 20 secondary schools, 18 primary schools, 2 Early Years settings and 1 special school featured were all making innovative and imaginative efforts to make reasonable adjustments, leading to good inclusive practice. The practical adjustments were clearly leading to a range of students being effectively included who would not have been in mainstream schools previously (The films can be seen).

We interviewed the principal or headteacher of these schools and there was a unanimity of ethos and values that probably marked them out from the majority of schools at this time:

- vision and values based on an inclusive ethos
- a “can do” attitude from all staff
- a pro-active approach to identifying barriers and finding practical solutions
- strong collaborative relationships with pupils and parents
- a meaningful voice for pupils
- a positive approach to managing behavior
- strong leadership by senior management and governors
- effective staff training and development
- the use of expertise from outside the school
- building disability into resourcing arrangements
- a sensitive approach to meeting the impairment specific needs of pupils
- regular critical review and evaluation
- the availability of role models and positive images of disability

The tragedy is that I have been in touch with many of these schools and this good practice can no longer be found in many of these schools. The education “reforms” with the setting up of Academies and Academy chains, loss of central support teams of disability expert teachers, and tightening of budgets all have made inclusion harder, but not impossible in English schools. Though it is still possible to find schools with good inclusive practice (Inclusion Working, 2017).

Champions of inclusive education

All over the world at the heart of the move to inclusive education have been pioneering parental champions who do not understand why their disabled children should not have access to education in regular or mainstream schools. When they link up with disabled people's organizations, they can create strong Alliances for Inclusive Education.

In 2008 in **South Africa**, I met Marie Schoeman who was working at the Ministry of Basic Education leading on implementing Inclusion. South Africa in its new democratic phase had committed to White Paper No. 6 on Inclusive Education and the Constitution guaranteed equal treatment for disabled people including in education. Marie like a number of parent advocates, had fought for her son with Downs Syndrome to go through mainstream schools. The following year (2009) I returned to South Africa to make a film of inclusive practice in 10 primary schools in 4 Provinces ([World of Inclusion, 2009](#)) arranged by Marie. There were some full-service schools which were being additionally resourced to take a full diversity of disabled children and we captured some excellent practice and interviews. The plan was also to change existing special schools into resource centers.

In 2008, I had been to 3 of these resource centers and they seemed to be using their extra resources to recruit more disabled students rather than support the surrounding schools. I also carried out a series of seminars to educationalists in 5 Provinces in 2009. We would call the gate keepers -principals of special schools and Educational Psychologists were vociferous in their opposition to inclusion—perhaps because of their vested interest? At this time some 300,000 disabled children were not attending school in South Africa.

Money from the Ministry for developing Inclusive provision was allocated locally in the provinces, and it appeared that much of this was going to open new special schools. This is a short-sighted policy, as now there are some 600,000 disabled children not in school ([wayside](#)). The lesson from this for me is that Governments need to have robust plans with the funding tied down to deliver inclusive provision with strong monitoring and penalty arrangements.

Bringing parents, educational professionals together in an Alliance for Inclusion makes for a very powerful campaigning body and being one of the founders of the Alliance for Inclusive Education in the UK in 1990 was important in campaigning for and getting through the repositioning of the Labor Government 1997–2010 to support Inclusive Education ([Alliance for Inclusive Education](#)).

Equally, organizations of parents where they take on the thinking of the Disabled People's Movement and become allies to their children in their struggle for Inclusive Education is very important.

"In the UK Parents for Inclusion was formed in the 1980s by parents who wanted inclusion for their disabled children and disabled activists—a fusion which led to an entirely new perspective, so that parents became allies in their children's struggle. Parents for Inclusion's mission statement says: We believe that it is physical and social barriers that stop disabled children from being included, rather than the disabled children themselves. Children often have little power to ask for what they want. They say they want to be able to make friends and be involved in their local community. They want to be able to get out and about, meet new people, enjoy their leisure time and go to their local school. Parents and professionals have great influence over these children's futures. So, we work with them to help them see children as individuals first, with human rights and preferences and a right to an optimistic and self-determined future. Training: All our training is designed and delivered in partnership with experienced trainers. One trainer is a disabled person and the other is the parent of a disabled child. Listening to disabled people allows parents to see the world from their child's point of view" ([Rieser, 2012](#)). This work with Parents for Inclusion was a vital part of building the Alliance for Inclusive Education.

Identifying and removing barriers to inclusive education

From the distilled experience of schools that have successfully included, school systems such as in Canada (New Brunswick, British Columbia), Finland, Italy, Portugal, Spain, Brazil, New Zealand and India and hundreds of small-scale NGO inclusive education projects across developing countries we know what works. The task is to resolve the barriers, ([Fig. 5A](#)) at school level and ([Fig. 5B](#)) Systemic barriers (below) ([Commonwealth Secretaria](#)).

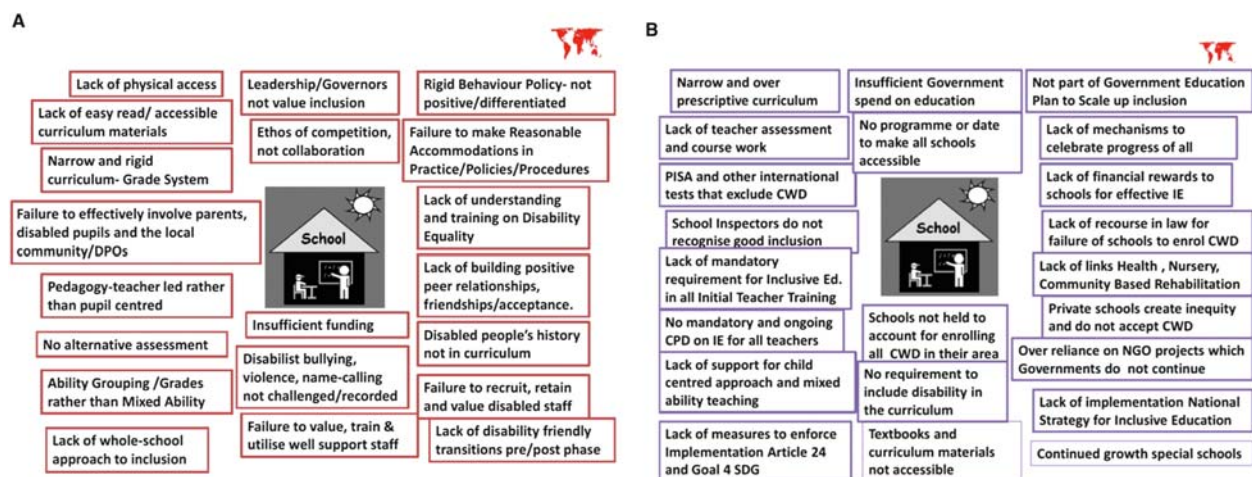


Fig. 5 (A) School barriers to the inclusion of children/students with disabilities. (B) Systemic barriers to the inclusion of children/students with disabilities.

Possible solutions

*The community can build/alter school-build ramps, widen and improve toilets, whiten walls. *Turn desks round into work tables and get students to peer support. Mix groups carefully. *Inclusion training in the community to get all children to school. *Train all teachers in the District on an aspect of inclusion and to make resources to access the curriculum for learners with different impairments. *Link with a University to train all teachers on inclusion involving action research in school, accredit and remunerate those successful. *Ensure all students and staff know where prejudicial attitudes come from, how to challenge them and include knowledge of the history and treatment of people with disabilities in their curriculum. *Teachers develop child centered approaches and adjust their practice to enrich each child. *For students vulnerable to dropping out or exclusion build strong peer support. *Make films of good practice in similar situations and promote.

Conclusion

The Global Summits on Disability in 2018 and 2022 have recorded increasing interest from Governments in supporting the development of inclusive education ([International Disability, 2018](#)). In 2020 UNESCO's Global Monitoring Report on Education Focused on Inclusive Education ([UNESCO, 2020](#)) focused on global inequalities and how to remedy them such as:

- * In the case of students with disabilities, laws in 25% of countries (but over 40% in Asia and in Latin America and the Caribbean) make provisions for education in separate settings, 10% for integration and 17% for inclusion, the remainder opting for combinations of segregation and mainstreaming.
- * Some 25% of teachers in 48 education systems report a high need for professional development on teaching students with special needs.

All these statistics and now inclusive education has been put back by the COVID Pandemic.

The 2020 Global Education Monitoring Report coincided with the International Disability Alliance (IDA) producing a Report that was supportive of developing Inclusion for people with all forms of impairment. During the debates around making the UNCRPD there had been significant divisions between the Deaf, Deafblind and Blind organizations and those representing learning difficulty such as Inclusion International and cross impairment organizations such as Disabled People's International that led to the weak formulation of Article 24. That weakness has been rectified by the UNCRPD Committee in General Comment 24. However, it was heartening to have a unifying Report from IDA ([International Disability Alliance, 2020](#)).

The report includes recommendations of good practices supported by policies and legislation, leading to truly inclusive education systems. It puts forward measures required to build an inclusive education system, geared toward providing quality education for all children and youth equitably. A critical message is that an inclusive education system is the only way to achieve SDG 4 for all children—including disabled children and youth—whomever and wherever they are. Inclusive education requires an educational transformation, which is unachievable if it is considered an add-on to existing education systems.

The IDA Report proposes "that including children and youth with disabilities—requires the following **minimum conditions**: (A) All Children/learners have access to quality education in schools where their inclusion requirements are met; (B) All teachers are equipped to ensure that all their students participate in quality learning; (C) Well-resourced support services and resources are available to assist all schools and all teachers to provide inclusive effective learning to all learners including those with disabilities; (D) All students succeed in reaching their full academic and social potential with learning outcomes measured against their own wishes, plans and benchmarks."

While the IDA analysis presents a way forward, caution is needed. IDA argues that specific impairment DPOs should provide the advice to schools and governments on impairment specific adjustments. Contained in the Annex of the Report. While these reasonable adjustments and support are needed, in a world where the large majority of disabled children if in school attend special schools, this is a dangerous strategy that could play lip-service to full inclusion, while reinforcing segregation. The retention of special schools while they transition to inclusive schools is also a dangerous strategy, as the evidence shows they take the extra resources offered and remain an expanded version of themselves. Experience shows the management and accountability measures need to be changed for this transition to be successful.

It is much better if cross-impairment DPO Councils or coalitions, with parents, provide this support and advise drawing on the wealth of experience from their members to cover the full range of impairments. The IDA approach is likely to lead to separate silos, with some groups such as psycho-social left out, rather than full disability equality ([Rieser, 2020](#)).

As we look to achieving the Sustainable Development Goals and in particular SDG No. 4 Quality Inclusive Education for All by 2030, we are a long way off-track.

We need now a new Global movement for Inclusive Education bringing parents, disabled young people, educators and the Disabled People's Movement together. The internet has demonstrated its power in building strong Global Movements on the Climate Emergency and Black Lives Matter. We know what needs to be done, but we need a A Global Alliance for Inclusive Education to make it happen.

References

- 17 Goals Agreed by World Leaders at the United Nations in September 2015. Sustainable Development Goal is the Education Goal. It Aims to "Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning Opportunities for All." SDG 4 is Made Up of 10 Targets Which Make It Clear This Includes Disabled People.
- Alliance for Inclusive Education, <https://www.allfie.org.uk/about-us/>.
- Assets, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/863815/2019_KS4_revised_text.pdf.
- Barnes, C., 1991. Disabled People in Britain and Discrimination: A Case for Anti-discrimination Legislation. Hurst, London. Ch. 1, p. 2.
- Committee on Economic, Social and Cultural Rights, General Comment No. 13 (1999) on the Right to Education.
- Many More Ideas and Examples in Implementing Inclusion A Commonwealth Guide to Implementing Article 24 of UNCRPD 2012, Commonwealth Secretariat, <http://worldofinclusion.com/v3/wp-content/uploads/2014/01/Implementing-InclusiveEducation-promo-copy1.pdf>.
- CSIE, 1997–2010. Index for Inclusion Bristol. CSIE. <http://www.csie.org.uk/resources/inclusion-index-explained.shtml>.
- DfE, 2010. The Importance of Teaching, the Schools White Paper 2010, p. 1. <http://www.education.gov.uk/publications/eOrderingDownload/CM-7980.pdf>.
- DFES, 2006. Implementing the Disability Discrimination Act in Schools and Early Years. DFES/HMSO, London.
- Disability Equality in Education, 2002. Inclusion in Schools Course Book. DEE, London. http://worldofinclusion.com/res/schinc/Inc_coursebook.doc.
- Educational, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985162/Special_educational_needs_Publication_May21_final.pdf.
- Efthymiou, K., 2021. How Inclusive is PISA. <https://enil.eu/news/how-inclusive-is-pisa>.
- Inclusion Working, 2017. Eastlea Community Secondary, Newham, <https://youtu.be/cAAoWGi3AKk>; Inclusion Working in 2015 Wroxham Pri., Hertfordshire, <https://youtu.be/NL-Y9L3U6gY>; Emersons Green Pri. South Gloucestershire, <https://youtu.be/w4B4CGopmZw>; Priestnall Sec., Stockport, <https://youtu.be/ggMUIQ58-xk>.
- International Disability Alliance, 2020. What an Inclusive, Equitable, Quality Education System Means to Us. IDA, New York. <https://www.internationaldisabilityalliance.org/ida-inclusive-education-2020>.
- International Disability Alliance Global Summit on Disability 2018, 2022, <https://www.globaldisabilitysummit.org/pages/global-disability-summit-2022-norway#>.
- Jordan, L., Goodey, C., 1996. Human Rights and School Change: The Newham Story. CSIE, Bristol.
- Joshua Malinga Quoted in African Development Activists Handbook, SAFOD in ADEWALE MAJIA-PEARCE 1998 Disabled Africa Rights Not Welfare, INDEX ON CENSORSHIP 1 1998 193, <https://journals.sagepub.com/doi/pdf/10.1080/03064229808536308>.
- Lin, Y., Besudnov, A., Black, A., Norwich, B., 2020. School Autonomy and Educational Inclusion of Children With Special Needs: Evidence From England. https://www.researchgate.net/figure/Trends-of-academisation-of-English-secondary-schools-and-SEN-inclusion-Note-This-figure_fig1_338768318.
- Lunt, C., 2021. Then There Was Silence London, Disabled Children's Partnership-An Alliance of 90 Charities. <https://disabledchildrenspartnership.org.uk/then-there-was-silence/>.
- Mason, M., Rieser, R., 1994. Altogether Better. Comic Relief, London. <http://worldofinclusion.com/res/altogether/AltogetherBetter.pdf>.
- National Audit Office, 2021. School Funding in England London House of Commons HC300.
- National Education Union, 2022. <https://neu.org.uk/levelling-white-paper>.
- NUT, 2015. Exam Factories Merryin Hutchinson. NUT(NEU), London. https://www.researchgate.net/publication/309771525_Exam_Factories_The_impact_of_accountability_measures_on_children_and_young_people.
- OECD, <https://www.oecd.org/pisa/>.
- PISA, 2020. National Project Managers Manual. OECD, Paris.
- Rieser, R., Mason, M., 1990. Disability Equality in the Classroom: A Human Rights Issue. IEA/DEE, London. <http://worldofinclusion.com/res/deinclass/DEC1.PDF>.
- Diagram from Rieser, R., 2004. Disabling Imagery? A Teaching Guide to Disability and Moving Image Media. BFI/DEE, London. <http://worldofinclusion.com/resources/#disability-and-the-media>.
- Rieser, R., 2008/2012. Implementing Inclusive Education: A Commonwealth Guide to Implementing Article 24 of the UNCRPD. Commonwealth, London. <http://worldofinclusion.com/v3/wp-content/uploads/2014/01/Implementing-Inclusive-Education-promo-copy1.pdf>.
- Box 6.5 in Rieser, R., 2012. Implementing Inclusive Education, p. 117.
- Rieser, R., 2020. The IDA Inclusive Education Global Report Inclusion Now Issue 57. ALLFIE, London. <https://www.allfie.org.uk/news/inclusion-now/inclusion-now-57-autumn-2020/>.
- Rieser, R., 2022. Inclusive education: implementing human rights. Ch. 8. In: Cole, M. (Ed.), Education, Equality and Human Rights, fifth ed. Routledge, London (Forthcoming).
- Rieser, R., 2018. 2008, 2012 and 2018 Ibid 2 and Developing Inclusive Education and Disability Equality for Children and Students With Disabilities: A World of Inclusion Broadsheet for the Global Summit on Disability. <http://worldofinclusion.com/developing-inclusive-education-and-disability-equality-a-world-of-inclusion-broadsheet-for-global-summit-july-2018/>.
- Smith, E., Douglas, G., 2014. Special educational needs, disability and school accountability: an international perspective. Int. J. Incl. Educ. 18 (5).
- The films can be seen online at <http://worldofinclusion.com/v3/wp-content/uploads/2014/01/RAP-document-with-youtube-links.pdf>.
- UN, 2006. Convention on Rights of Persons With Disabilities Article 1.
- UNESCO, 2020. Inclusion and Education All Mean All Paris. UNESCO. <https://en.unesco.org/gem-report/report/2020/inclusion>.
- Union of the Physically Impaired Against Segregation: Policy Statement 1974. Access University of Leeds Disability Archive, <https://disability-studies.leeds.ac.uk/library/author/upias>.
- United Nations Enable UNDESA, 2007. <https://www.un.org/esa/socdev/enable/convinfopara.htm>. (Accessed 9 February 2022).
- United Nations, 2016, CRPD/C/GC/4. General Comment No. 4 (2016) on the Right to Inclusive Education. UN CRPD Committee, Geneva.
- Wayside, <https://www.news.uct.ac.za/article/-2020-05-28-the-education-of-children-with-disabilities-risks-falling-by-the-wayside>.
- World Bank, 2021. Learners With Disabilities and COVID-19 School Closures: Findings From a Global Survey Conducted by the World Bank's Inclusive Education Initiative. World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/36326>.
- World of Inclusion, 2009. Developing Inclusive Education in Southern Africa. <https://youtu.be/W3CJnCr8vRY?list=PLw6nQoN43w2Q6HmHao8a3SfbC7i8Ae83>.

“But tell me, where do the children play?” Disabled children and the right to play

Angharad E. Beckett, University of Leeds, Leeds, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	306
The right to play	308
Barriers to play for disabled children	310
The “adulteration” of disabled children’s play	314
Emancipating play for disabled children	315
Conclusion	317
Acknowledgments	318
References	318

Introduction

The quotation in the title of this chapter comes from the chorus of a song by Yusuf Islam (aka Cat Stevens):

I know we’ve come a long way
We’re changing day to day
But tell me, where do the children play?

The song is about urbanization, pollution and the loss of green-spaces for children to play, yet its chorus has lodged in my mind because it also speaks to the issue of the inclusivity of play. Progress has been made in recognizing disabled children’s¹ right to play. Things *have* changed. Things *are* changing. In *some* places. When compared to thirty years ago, there is now greater recognition of the need to move beyond individual model (often termed “medical model” or “deficit model”) articulations of the “problem” of disabled children’s play. We have a better understanding of the many and varied barriers—beyond their minds and bodies—that limit disabled children’s access to and inclusion in play. It is more common to see researchers, practitioners and governments espousing rights-based approaches to all children’s play, including that of disabled children. The topic of play for disabled children is receiving ever more attention from researchers in various disciplines, including children’s geography, occupational therapy and design. We have seen initiatives such as the LUDI Network (a European Cooperation in Science and Technology [COST] Action) which brought together researchers from many different disciplines (including psychologists, sociologists, designer/engineers), civil society organizations and the play industry with the objective:

to spread awareness on the importance of giving children with disabilities the opportunity to play, while ensuring equity in their exercise of the right to play²

It is possible to identify position statements, play policies and various initiatives that recognize and seek to support disabled children’s right to play. For example, in 2015 the International Play Association (IPA) published a Position Statement on the Play Rights of Disabled Children which emphasized that:

Disabled children have the same right as other children to sufficient time and space to play freely, in the ways they choose, without being unduly overprotected

IPA (2015, para. 2).

¹Herein I follow the UN Convention on the Rights of Persons with Disabilities in understanding disability as “relational”—resulting from “interaction between persons with impairments and attitudinal and environmental barriers that hinder their full and effective participation in society” (UN CRPD, 2006 para. e). Whilst the UN CRPD adopts people first language (“persons with disabilities”/“children with disabilities”) I use identity-first language (“disabled people”/“disabled children”) because this is the preference of the UK disabled people’s movement and is aligned with the UK social model of disability and because I value the political operations of the social model. Further, since this model informed the articulation of the CRPD (Lawson and Beckett, 2021), the use of identity-first language is not fundamentally at odds with the spirit of the convention.

²<https://www.ludi-network.eu/about-ludi/> (accessed 18/07/22).

The inclusive playgrounds movement has grown over this period and the idea that we should be providing spaces and opportunities for disabled and non-disabled children to have fun together, has been influential. One example, amongst many, is the ‘Morgan’s Inspiration Island, San Antonio, Texas’—an inclusive waterpark with many accessible features available, including waterproof wheelchairs (see Fig. 1). The park was opened by disability advocate Gordon Hartman, who has stated:

Morgan’s Inspiration Island is not a special-needs park; it’s a park of inclusion (...) designed with special-needs individuals in mind and built for everyone’s enjoyment. Our goal is to provide a great guest experience in an inclusive, safe, comfortable, not-overly-crowded environment (...) without barriers.³

Yet despite these positive developments, disabled children continue to be at risk of experiencing play deprivation or play bias (exclusion from a full play experience) (Hughes, 2003). The question “where do the children play?” thus remains pertinent for these children. In this chapter I ask, why and how is this the case and set forth an agenda, not only for ensuring disabled children’s right to play, but also for *emancipating* their play.

The chapter is organized into three further sections and a conclusion. In Section 2 I outline how disabled children’s right to play is enshrined in international law. I explain how play has been defined/interpreted by the Committee on the UN Convention on the Rights of the Child and why their interpretation is important. In Section 3 I provide a narrative review of existing knowledge about the barriers to play for disabled children, focusing on literatures that consider disabled children’s access to and inclusion in play at home, school and in the community (e.g., at public playgrounds). In Section 4, I take a closer look at a set of barriers resulting from the discursive framing of disabled children’s play within the health and psy-sciences (especially, educational psychology). It is here that my immersion within ‘Disabled Children’s Childhood Studies’ (DCCS) and adoption of a DCCS perspective is most apparent.

This chapter includes two case-studies (see Figs. 4 and 5). The first is drawn from a research project that explored barriers to play for disabled children living in Hong Kong and Taiwan. This case study provides insights into barriers to play for disabled children living outside the “West”. The second is drawn from a project conducted in the UK, entitled “Together Through Play”. This case study demonstrates the children’s capacity for “design thinking” and why it is important, and valuable, to involve them in inclusive play design projects.

Also interspersed throughout the chapter are a series of vignettes. These are derived from the research journals/field-notes I have kept during various studies that involved observations of disabled children playing or in play situations. In *Vignette 1*, which ends this section, the reader is introduced to Freddy, for whom play is a serious business. These vignettes are included in order to: (a) bring the voices and perspectives of children into this chapter; and/or (b) to provide insights into the ways that disabled children’s agency is revealed in and through their play.



Fig. 1 Children playing at the Morgan’s Inspiration Island. Photo credit: Robin Jerstad/Jerstad Photographics.

³Quoted in Metro newspaper: <https://metro.co.uk/2017/06/16/the-worlds-first-fully-accessible-water-park-opens-this-weekend-with-special-waterproof-wheelchairs-6712734/?ito=facebook> (accessed 17/07/22).

Vignette 1 Play Is a Serious Business

I am sitting quietly in a classroom of primary-aged children, aged 7–8 years, observing. The children have just returned from playtime and the teacher is getting them settled to tasks. There is a lot of movement in the room—children are getting out exercise books, fetching coloring pencils and so on. I notice that one little boy—Freddy (not his real name)—is doing something different. He’s moving around the room in a very careful, surreptitious or stealthy way, mostly below the tables, or behind the bookcases. The teacher also notices and calls out to him that “playtime is over”, “stop messing about” and “please focus on the task”.

Freddy has recently been diagnosed as having Asperger’s Syndrome. His teacher has opened up to me about it. She has explained that she’d thought he was “not very bright” and had been very anxious that because he was not concentrating in class, that he might “ruin the SATS results for the class” (national tests). The educational psychologist who diagnosed Freddy has said that he is very intelligent and has given her an A4 leaflet about Asperger’s. He cannot provide any other support because he’s got too heavy a case-load. The teacher bursts into tears, tells me how guilty she feels and promises to go to the library and “read up” on things, so she can support Freddy better.

Intrigued by Freddy’s behavior in the classroom and having developed a good rapport with him, I beckon him over and casually ask him what he’s up to. What game is he playing, moving around the classroom so secretly?

He leans in and whispers to me in a conspiratorial manner: “Well, it is a game yes, but *really* I’m practicing subterfuge”.

The right to play

The right to play is enshrined for *all* children in Article 31 of the UN Convention on the Rights of the Child (CRC, 1989) and reinforced for disabled children by Article 30 of the UN Convention on the Rights of Persons with Disabilities (CRPD, 2006):

UN CRC (1989) Article 31 “States Parties” must: ... recognise the right of the child to rest and leisure, to engage in play and recreational activities.

UN CRPD (2006) Article 30, 5d “States Parties” must take appropriate measures to: ... ensure that children with disabilities have equal access with other children to participation in play, recreation and leisure and sporting activities, including those activities in the school system.

There is also a General Comment (GC17, 2013, on CRC, Art. 31, see UN CRC, 2013) on the “right to play”. This quasi-legal document provides authoritative interpretation of an area of international law. GC17 offers important guidance regarding how play should be understood, for all children. It distinguishes play from “play-like” activities (e.g., play therapy) and adopts a concept of play that, in Play Studies, has come to be termed “free play”. One of the best-known definitions of the latter has been provided by Garvey (1990) who defined free play as: a) pleasurable; b) having no extrinsic goals; c) being spontaneous and voluntary; and d) involving some level of player engagement.⁴ Free play occurs when children decide “what they want to do, how they want to do it and when to stop and try something else” (Santer et al., 2007: xi). It is play free from adult-agendas. Having opportunities for free play is important for children and young people’s mental health, overall well-being and quality of life. Play “is more than an enabler or a tool in relation to social, educational, physical and psychological development and health promotion/maintenance” (Barron et al., 2017: 1). Play has intrinsic value for all children.

GC17 also stresses that children’s right to play intersects with all other principles of the CRC including non-discrimination (Art. 2) and children’s right to express their views and be listened to (Art. 12) and reminds us that disabled children are also afforded similar rights under the CRPD. The authors of GC17 identify disabled children as a group facing barriers in/to their play. Of the “multiple barriers” identified in this general comment, all lie external to the minds and bodies of children. This is important. It is a position echoed by the IPA (2015: np) which has stated that the barriers disabled children face in/to their play lie beyond their impairments in: (a) “inaccessible facilities and environments, negative attitudes and inappropriate social policies and programmes”; and (b) “the imposition of activities determined by adults” onto disabled children’s play.

All States Parties that have signed and ratified the CRC and CRPD are obliged to put in place an appropriate set of policies and programmes to ensure their realization. Sadly, however, this does not always happen. As a “snapshot”, between January 2021–June 2022, of the twenty-five Concluding Observations published either by the Committee on the CRC or Committee on the CRPD, twelve reference concerns or recommendations that relate, in some way, to disabled children’s right to play (or to their participation in leisure/sport) or importance of providing accessible spaces for recreational activities. It is regrettable that the remaining Concluding Observations neglect the right to play. This appears, however, to be because the committees have prioritized issues such as the failure of States Parties to ensure disabled children’s right to education, have wanted to highlight the precarious lives of disabled children in particular countries, or have sought to expose the significant human rights violations disabled children are experiencing in other areas of their lives (e.g., as a result of institutionalization). It is at times difficult to argue that greater

⁴NB. there is no set level of “engagement”. A giggle or smile, or a blink of an eye may be meaningful engagement.

attention to be given to disabled children’s “right to play” when there are even more pressing issues at stake, such as ending the neglect and abuse many disabled children face within institutional settings (congregate care).

Nevertheless, it is important not to forget the “right to play” and to hold governments to account in this respect. As is evident from just 18 months of Concluding Observations from the CRC and CRPD Committees (and the issue can be traced into *many* more of these documents prior to 2021), governments are failing to take their obligations seriously when it comes to supporting disabled children’s play.

At the time of writing, it is not known how many countries have a national play policy or program that references the needs of disabled children. What is known is that where play policies do exist or have existed, they have not always been well implemented, evaluated or enforced. There is also evidence that these policies are vulnerable to the ebbs and flows of politics. A case in point would be in England. In 2016 the CRC Committee (UN, 2016) expressed concern that children’s views are not being systematically heard in policymaking on issues that affect them (also implying contraventions of domestic law in the Children’s and Families Act, Part 19) and that measures to support disabled children’s play were not always being taken. They highlighted withdrawal of the Play Policy for England and inadequate resourcing of disabled children’s play in this national context. The situation in England is not unusual.

To conclude this section: disabled children’s right to play is enshrined in international human rights law, within the CRC and CRPD. Has this been sufficient to ensure that their right to play is respected? The answer is of course, no, or *not yet*. CRC Article 31 and CRPD Article 30 do, however, provide an important moral horizon and act as route-maps for change. Further, the concluding observations of the two committees are part of an international monitoring process. They contribute to the delineation of relevant standards. Concluding observations have the potential to include positive recognition of successful strategies or innovations undertaken by governments to protect disabled children’s right to play; as well as criticism for regressive developments or failures to act. Neither Committee has the power to “punish” States that do not guarantee disabled children’s right to play by imposing sanctions. Nevertheless, “the reputational implications of public criticism from such a body has the potential to damage the standing of governments both internationally and at home” (Lawson and Beckett, 2021: 369). Further, at a national level, in countries which adopt a monist approach to international law, there is the potential for the CRC and the CRPD to be used in courts—opening the possibility that disabled children’s exclusion from play opportunities or play spaces could be legally challenged.⁵

UN CRC and CRPD ‘Concluding Observations’ January 2021–June 2022

June 2022

CRC on Greece (G.33.i): the committee recommends the government attend to ‘regional disparities...including access...to **leisure and play**’ for disabled children.

CRC on Cambodia: the committee recommends the government to provide ‘**accessible playgrounds**’.

CRC on Canada (H.40.e.): the committee recommends the government works to guarantee the right to ‘**play and accessible recreational activities**’ for all children (my emphasis).

CRC on Croatia (I.40): the committee recommends that the government provides ‘children, **including those with disabilities**...with safe, **accessible, inclusive and smoking-free places for play**...with public transport to access such spaces’ (my emphasis.)

CRC on Chile (G.37.b.): with regard to disabled children, the committee recommends that the government puts in place ‘legislation, policies and programmes to facilitate their access to sports facilities’.

March 2022

CRPD on Hungary (B.58): the committee is ‘concerned about the remaining barriers to access to recreational, leisure and sporting venues’ particularly for ‘children with disabilities’.

CRPD on Jamaica: committee makes statement similar to that made vis-à-vis Hungary.

September 2021

CRC on Poland (I.40): the committee recommends that the government ‘strengthens its efforts to guarantee the right of children, in particular children with disabilities, to...**sufficient time to engage in play** and recreational activities that are safe, accessible, inclusive, reachable by public transport and age-appropriate’.

CRC on Switzerland (H.41.a.): the committee recommends that the government ensures that ‘sporting, leisure...activities and accessible for children with disabilities’.

CRPD on France (B.60): the committee is concerned about ‘**lack of play, recreation and leisure and sporting activities for children with disabilities**’.

CRPD on Estonia (A.59): the committee recommends that the government ‘should promote inclusive recreational, leisure and sporting activities for...children with disabilities’.

June 2021

CRC on Luxembourg (F.23.f.): the committee recommends that the government take ‘all measures necessary to provide children with disabilities with reasonable access in all settings including at school and in leisure centres’

(Emphasis added by author)

References to play/leisure/recreation within UN CRC and CRPD Committee’s Concluding Observations Jan 2021–June 2022.

⁵There is need for increased legal advocacy in this area. Geary (2012) for the Child Rights International Network found CRC, Article 31 to be a neglected article, seldom (he found zero) referenced in court cases deemed to be important with regard to their interpretation of children’s rights.

Barriers to play for disabled children

The previous section introduced the international legal frameworks and discourse surrounding disabled children's rights to play and described how the CRC and CRPD committees recognize that disabled children face barriers in and to their play activities. This section explores these barriers as they exist within the social environment. Emphasizing that the main barriers to their play exist within the social environment, does not mean ignoring the reality of impairment. Impairments may present certain challenges for a child in/to play; but equally, they may not. The presence or absence of any particular impairment may simply result in different play preferences or experiences. Where impairment-effects *do* present challenges for the child in their chosen play, these challenges can often be overcome with the use of imagination or application of knowledge on the part of "allies" (families, friends, professionals). For example, it is helpful to raise awareness/understanding of how certain toys may be more user-friendly and interesting for children with particular impairments e.g., tactile or sound toys for children with visual impairments (Besio, 2017). Nevertheless, care should be taken to avoid focusing too heavily upon impairment. Such an approach tends to result in individualized approaches that are not always necessary and present attendant risks (e.g., of developing "special"/segregated provision, paternalistic or stigmatizing "solutions"). Further, disabled children have agency and often navigate impairment-effects themselves, in creative ways, without adult-interventions.

When it comes to disabling barriers to play within the social environment the first thing to note is that these often result from, or are compounded by, barriers that disabled children face in other aspects of their lives, beginning with poverty. Economic deprivation is a feature of the lives of many disabled children, globally (UNICEF, 2013). In many countries disabled children are not provided with basic services such as healthcare and education and are at higher risk of experiencing abuse or violence (UNICEF, 2013). The CRC General Comment 17 highlights disabled children's exclusion not only from school, but also from informal and social arenas where they might form friendships and where play takes place.

Within research and gray literatures, the inaccessibility of physical environments and non-inclusive design are the most frequently cited barriers to disabled children's play (e.g., Burke, 2013; Dunn and Moore, 2005; Law et al., 1999; Moore and Lynch, 2015; Rigby and Gaik, 2007; Rimmer et al., 2004; Olsen and Dieser, 2012; Woolley, 2013; Yantzi et al., 2010), especially in relation to playgrounds and other public play spaces. Insufficient, unsuitable or expensive public transport can mean a difficult, even impossible "journey to play" at suitable locations (Stafford, 2017). Noise and overcrowding when children are able to reach a play-site are known problems (Law et al., 1999; Rimmer et al., 2004). Play-spaces with unsuitable surfaces or lighting or that lack appropriately designed play equipment, can result in exclusion from play for disabled children. Poor design can mean that disabled children need adult assistance and are thus prevented from playing independently. Sometimes, the introduction of "accessible" elements to playgrounds (e.g., wheelchair swings), if not carefully positioned, have resulted in further segregation, even stigmatization, if they come to be perceived to be "special things for special children" and not fun for all (Dunn and Moore, 2005). An example of the latter would be the positioning of equipment such as wheelchair swings—there is often only one of the "heavier duty" versions of such swings on a playground (able to carry an electric wheelchair on a platform) and barriers or a perimeter may be constructed around the swing for health and safety reasons (see Fig. 2).

Disabled children sometimes self-exclude from poorly designed (or inadequately supported) play-spaces lest they be perceived to be less "able" within those spaces (Prellwitz and Skar, 2007). Poor design can result in internalized oppression. The issue of poor design of playgrounds has been recently highlighted in the UK by the charity Scope. Emma Vogelmann, disability activist and Lead Policy Adviser at the charity/non-profit Scope, was quoted in the Metro Newspaper (May 12, 2022) as follows:

Every child has an equal right to play. Play feeds imagination and forms friendships (...) Our playgrounds are places where memories are made and where children can be themselves. Yet many disabled children can't enjoy their local playground because the equipment isn't designed for them. It leaves disabled children shut out and missing childhood experiences.⁶

At the root of the problem and/or exacerbating the problem is insufficient knowledgeability regarding the concepts of inclusive and universal design amongst policy-makers, urban-planners and commissioners (Steinfeld and Maisel, 2012). Further, the expertise of disabled children and their families—born of lived experience of disabling barriers—is not always recognized. Researchers have found that children and families are seldom consulted adequately during playground design processes (Prellwitz and Tamm, 1999; Woolley, 2013).

At school, disabled children have been found to experience exclusion at playtimes. Yantzi et al. (2010) found that disabled children face barriers in school playgrounds—getting in, moving around and playing. This provokes feelings of being "abnormal" and not belonging amongst the children. Woolley (2006) found that Physical Education (PE) can be used to introduce skills in relation to inclusive play, that children can then transfer onto the playground. But even in this study, the authors found that if disabled children are delayed getting onto the playground, then they risk getting left out. Inclusion in play is a "fragile" thing. Vignettes 2 and 3 jointly entitled "Inclusive Play: Now You See It. Now You Don't" offer the reader additional insights into the ways in which

⁶<https://metro.co.uk/2022/05/12/disabled-children-are-being-excluded-from-playgrounds-16599216/> (accessed 16/07/22).



Fig. 2 Wheelchair swing at an inclusive playground in Taiwan—equipment located in separate/segregated enclosure for health and safety reasons. Photo: Author’s own.

PE might be used to encourage inclusive play. Yet these vignettes also demonstrate how easily positive developments can be undermined by non-inclusive educational environments and practices.

Vignettes 2 and 3 Inclusive Play: *Now you See it. Now you Don’t.*

V2: I arrive at the primary school. It’s a new building—very much “architect designed”. In the foyer a large sign reads “We’re a rights-respecting school” (referring to the UNICEF scheme of this name). The receptionist points me to the gymnasium where I’m met by a peripatetic PE teacher and introduced to the class. The children are juniors, aged 8–9 years old. The teacher reminds the children why I’m there.

Having been told that I’m observing the XXXX program (a games-based “intervention” designed to promote disability equality) being rolled out in a “usual PE session”, I’m surprised that almost half of the children have some type of visible impairment. This includes four children who are wheelchair users. I note how unusual this is. I wonder if I have chanced upon the most inclusive school in the UK?

The class begins. The PE teacher invites the children to warm up. “Everyone do some star jumps!”, he says, then quickly adds, “you in the wheelchairs, do something else”. One of the children who is a wheelchair user glances over to me and catches my eye. He gives me an obvious “long-suffering” look (eye-roll) in response to this instruction from the PE teacher.

After the session ends, I’m scheduled to speak to some of the non-disabled children who took part. I ask them if they enjoyed it? They say that they had. I ask which aspects? They reply that it had been good to meet “those other children”.

“Which other children?”, I ask them. Although, of course, I know. They tell me that they meant the disabled children.

“Aren’t they members of your class then?”, I say. They tell me that the disabled children are not usually in their PE sessions because they are part of a “different class”. That class has a classroom “on the other side of the playground”.

I realize that they are describing a unit for children with “special educational needs”, which is supposed to be part of this mainstream school – but really *isn’t*.

I ask the children if they have shared playtimes with the disabled children from the unit. They tell me they do not and *have never met them before*.

~~~~~

V3 (different school): I have just observed a session being run by one of the school partnership/outreach coaches from a Premier League football club in the South of England. It’s a session meant to promote disability equality and inclusion. The children are aged 8–9 years. They have been playing Chair Football. To make the session more accessible for all the children in this class, the coach has opted to ask the children to sit on the floor. Although none of the children are visually impaired, he has brought a bell ball with him and explains its purpose. There has been discussion about making games accessible, so everyone can join in.

As scheduled, some of the non-disabled children are chatting to me. I ask them if they enjoyed the session. They say that they did. One child says that the activities had been fun because everyone was able to take part, “no matter their abilities”. Another, who seems very excited, tells me that the session had been “helpful” for “when we’re playing on the playground” and suggests we could “take some of the ideas ... onto the playground and make them public!”. The others agree.



Later, I go to observe the children during playtime. They have set up another game of what they are calling "Sit Down Football". The game is in full swing and the players include a disabled child who has cerebral palsy. A teacher who is standing close to me observes that it's "really nice to see him being included in the game because he's often left on the side-lines".

A few minutes later—long before the end of playtime—a teaching assistant appears on the playground and leads the disabled child away from the game and back inside the school building. I ask why. The teacher explains that it's "necessary". He would "need to go to the toilet before class begins and that can take a while, so he always goes in early from playtime".

At home, lack of space can impact negatively upon disabled children's play. Researchers have found that barriers to play in the home include insufficient space for wheelchairs to maneuver, or necessary pieces of equipment (e.g., hoists) taking up potential play-space (Brotherson et al., 2008; Connors and Stalker, 2003). Late or no adaptations to a home (e.g., absence of a stair-lift) can reduce children's independent mobility, impacting on their play. Kuhaneck et al. (2010) found that toys are seldom designed to be accessible by children with a wide range of abilities and "adapted" toys often offer limited play options, lacking the degree of challenge for enjoyable play experiences. Toys that are suitable for a child's abilities are often inappropriate for their age (Kuhaneck et al., 2010). If the only toys a child can access are designed for younger children, this can become a barrier to their inclusion within their peer-group (Kuhaneck et al., 2010)—it's "just not cool" to be playing with toys that are meant for younger children.

Perhaps the most stubborn barriers to disabled children's play lie in the attitudes of peers and adults. Jeanes and Magee (2012) advise us not to romanticize play or imagine that because a play space, for example a playground, is well-designed, accessible and has the potential to be an interesting, fun environment, that it will be a safe, happy place for all children. They remind us that "play can be a place of power dynamics" and "playgrounds can be sites where young people define and legitimize social hierarchies". Attitudes held by non-disabled peers can lead to exclusion of disabled children from play (Mencap, 2007; Spencer-Cavaliere and Wilkinson, 2010). Negative peer attitudes have led to disabled children becoming withdrawn, anxious or depressed (Baumeister et al., 2008; Contact a Family, 2012). Disabled children are more likely to be bullied than other children (Connors and Stalker, 2003; McLaughlin et al., 2016). Disabled children do sometimes display bullying behaviors, but this is often an impairment-effect not understood by other children (e.g., a child might play roughly, accidentally hurting another child), or may be their response to bullying (Contact a Family, 2012). Both reasons require sensitive management of play situations by adults, and this does not always happen (ibid).

Well-designed play-spaces can thus be helpful, but we also need to develop more inclusive play cultures. Developing such cultures is not as simple as bringing disabled children and non-disabled children together in the same space and expecting that inclusive play will "happen", spontaneously. Sometimes it does. Often it does not. Wenger et al. (2021) found that social barriers amongst peers can present challenges for disabled and non-disabled children. They noted that whilst disabled children can be excluded by non-disabled children, disabled children sometimes felt (or assumed) that non-disabled children would not want to befriend them and thus "stuck together" in play contexts. Encouragingly, they found that when given the opportunity and encouragement disabled and non-disabled children displayed design-thinking about how games could be adapted to ensure everyone was included in the fun. Children's design-thinking was much in evidence in a research project that colleagues and I undertook entitled "Together Through Play" (Holt et al., 2014). Please see Fig. 3 and the description of – Button Bash – designed by friendship groups of disabled and non-disabled children as part of this project. This case study offers the reader insights into the ways in which "mixed ability" friendship groups can find solutions that allow everyone to be included and have an important part to play in the game.

Adult attitudes and behaviors can also become barriers to disabled children's play. For example, staff attitudes have been found to negatively affect access to natural environments and nature play for disabled children (Gleave, 2010; Shackell et al., 2008). Adult perceptions of risk figure highly here. Gleave (2010), for Play England, found that staff considered less supervised outdoor opportunities to be "too risky" for disabled children. In another study, Ludvigsen et al. (2005) found staff at a nature playground attempting to stop disabled children engaging in "messy" activities such as kicking and throwing leaves. Exaggerated perceptions of the risk of play for disabled children have also been found amongst some parents of disabled children (Connors and Stalker, 2003; Schlieien et al., 2014). Not all such concerns are unfounded. Nevertheless, so-called bubble-wrapping can constrain disabled children's free play. This can also be a problem in the school-setting and on school playgrounds, where research has shown that teaching staff who are risk-averse limit disabled children's opportunities to make choices, take risks, embrace challenges and make friends (Ozen et al., 2013; Richardson, 2002).

Even when adults are play-advocates they can exhibit unconscious ableism. This is particularly evident within the nature play movement, which has gained considerable momentum since the 2000s. The UK heritage organization the National Trust's campaign "50 things to do before you're 11 and 3/4"<sup>7</sup> (running to date, 2022) encourages children to climb a tree, jump over waves, go on a walk barefoot, run around in the rain, play conkers, climb a huge hill, try rock climbing, canoe down a river and build a raft. Similarly, the Yorkshire Wildlife Trust (UK) encourages children to:

take risks, hide, run, jump, swing, balance, climb, sit, talk and have fun<sup>8</sup>

<sup>7</sup>See <https://www.nationaltrust.org.uk/50-things-to-do> (accessed 19/07/22).

<sup>8</sup><http://www.ywt.org.uk/discover-learn/pearson-park-wildlife-garden/natural-play> (accessed 03/08/15).



**Fig. 3** Together Through Play. (A) Children test the prototype of their design, (B) Button Bash, showing the game-characters designed by the children. (A) Photo courtesy of Principal Investigator, Together Through Play Project, Raymond Holt, (B) Photo courtesy of Raymond Holt.

There are disabled children who would be able to do some or all of these activities. There are disabled children who might be supported to engage in many of these activities. Equally, however, there are those for whom such activities would prove difficult, perhaps impossible. What does this signal? At best, we might conclude that this is a matter of “lazy ableism”—a failure to recognize that children are not a homogenous group or to imagine activities that cater for children with different abilities. At worst, there is a concern that such campaigns promote an “ideal” of the playful/playing child, and rest upon associated and “desirable” ability-expectations.

So far, I have drawn upon research and gray literatures from many different countries. Yet it is important to note that whilst there has been a growth in studies of disabled children’s experiences of play—from around the world—knowledge about play for disabled children outside the West and in the majority world, is still very patchy. More attention needs to be given to whether and how different cultural constructions of disability and childhood, varied social and family structures and myriad other factors (e.g., geographies, intersecting oppressions) impact on disabled children’s access to and inclusion in play. This is not to say that there are not some barriers that are faced by disabled children irrespective of where they live, but rather that it is likely that these have distinct local flavors or formations (Beckett et al., 2020). The latter is something that colleagues and I found, for example, in our study of access to and inclusion in play for disabled children living in Hong Kong and Taiwan. More studies are needed that expand our knowledge of disabled children’s experiences of play, particularly in the Global South.

To conclude this section: disabled children face myriad barriers to play—including economic, physical and attitudinal. In the final section of this chapter, we will turn our attention to how these barriers might be removed and disabled children’s play become “emancipated”. There is, however, one thorny barrier left to consider: the “adulteration” of disabled children’s play. It is to this issue that I turn in Section 4.

## The "adulteration" of disabled children's play

Play Studies is a highly interdisciplinary field. A rich history of research and theorizing from psychoanalytic, evolutionary and psycho-educational perspectives has unquestionably deepened our understanding of children's play. There is need, however, to be cautious and critical in the application of play theory when considering play for disabled children. This is because it has tended to (over)emphasize the relationship of play to child development and underemphasize its role in relation to wellbeing. This is especially the case with psycho-educational perspectives. In this section I engage critically with these perspectives.

### Together Through Play

During the Together Through Play project, funded by the Leverhulme Trust, we worked with friendship groups composed of disabled and non-disabled children. The disabled children had physical impairments (cerebral palsy). Supported by a researcher, each friendship group worked to develop a concept for a new game or toy they would like to play with together. Their ideas were developed into prototypes by students at the Faculty of Engineering and Physical Sciences, University of Leeds (UK). The children tested and helped to refine the prototypes. This case study describes one of the games designed by the children.

**Button Bash:** initially the children wanted this to be a videogame involving a hovercraft being steered around a race course by pressing buttons embedded within a floor mat. Lo-fidelity testing revealed that the children who had cerebral palsy found the game difficult to access due to the fixed positioning of buttons on the mat. Working with the children as co-designers, the game was reimagined as a video game controlled via large, very sensitive access switches. These could be positioned flexibly and pressed using any part of the body. The game play became something akin to 'whack-a-mole'. Buttons had to be hit to shoot a corresponding alien as it appeared on the screen. The revised version was more accessible. Nevertheless, inclusion of individual scores meant the children played competitively. This proved to be problematic for children with slightly slower reflexes (not only the children who had cerebral palsy). The children themselves identified the solution. They proposed that Button Bash could be modified to become a team-game. One of the disabled children suggested that it would be fun to work together to navigate a maze, each team member having a 'button' that moved the character either 'right', 'left', 'up' or 'down'. This necessitated a coordination of effort by the whole team. This cooperative game proved to be very popular.

Play is, of course, a site of/for informal learning. Further, play is essential for children's physical, cognitive, social and emotional development (Frost et al., 2012). It provides an outlet for and fosters their creativity, reflection and problem-solving. It helps them to gain self-esteem and "confidence to meet physical, intellectual and emotional challenges" (Whitebread et al., 2012: 31). Access to play is important for all children's subjective wellbeing and mental health (Hughes, 2012).

It was perhaps inevitable that because play has all these "benefits" that it would eventually be captured by a machinery of government(al)ity—and so it was, by the psy-sciences. By the 1900s psychologists had recognized the important role of play in child development. The theories of Piaget, Vygotsky and others emphasized the functional value of play. Play Theorist Sutton-Smith (1997) described how from this early work emerged a rhetoric of play as "play-as-progress"—play as key to child development—which came to be dominant in the 20th Century (and is arguably still dominant today). Sutton-Smith was keen not to dismiss all work in this vein. Yet he was concerned that the rhetoric is problematic when it serves "adult needs rather than the needs of children" (Sutton-Smith, 1997: 42). Today, in advanced capitalist countries, the "adult need" this rhetoric prioritizes is the production of a certain type of "ideal" child/young person and future citizen: possessing certain abilities, self-sustaining, rational, productive and entrepreneurial (Goodley, 2014; Slater, 2015). Even the most cursory of Google searches reveals many online articles and blogs extolling the value of play for preparing children for work and success or explaining how playing games can instil an entrepreneurial mind-set in children.

All children are subject to this rhetoric, but it impacts acutely upon disabled children. Unfortunately, as John and Wheway (2004: 12) comment, too much of the discussion about so-called "freedom to play" for disabled children has in fact "revolved around treatment or directed activity, rather than play". Anxiety that they do not meet the ableist demands of society is revealed in the tests that have been designed to measure their "play dysfunctions" or interventions to enhance their "play performance". Disabled children often spend large amounts of their time undertaking activities that masquerade as play but are really vehicles to help them "overcome" their impairment and achieve certain developmental goals (Goodley and Runswick-Cole, 2010). Play can sometimes be considered to be "frivolous" or "disposable" when the focus of parents or professionals is on educative or therapeutic "regimes" (Beckett et al., 2020). As the British disability activist Micheline Mason (1990) once wrote, reflecting on her own childhood and those of other disabled people:

Other children play, but you do "therapy". (...) Almost every activity of daily living can take on the dimension of trying to make you less like yourself and more like the able-bodied.

The research community continues to play its part in the re(production) of this rhetoric. For example, turning again to the example of nature play—where the play-as-progress rhetoric is highly evident—it is noticeable that its fun/joyful aspects are sometimes downplayed when it comes to disabled children's access to such play. There are studies that suggest that "frivolous" nature play for disabled children distracts from its more serious purpose. Von Benzon (2010, 624), for example, critiques certain UK-based environmental organizations for emphasizing play/recreation and "failing to provide adequate educational opportunities"

for disabled children; whilst Crosbie (2014), in a study of the work of the Calvert Trust (UK) (a provider of outdoor adventure holidays for disabled people, their families and friends) is critical of situations where disabled children are not encouraged to do anything *other than play* because of limiting perceptions of staff or workers. They are, he argues, being denied the wider benefits to be gained from outdoor “education”.

To conclude this section: disabled children do of course have rights to high quality education and rehabilitation. We should be concerned, however, when play is repurposed by adults—*adulterated* (Else and Sturrock, 1998; Hughes, 2012)—to become primarily or solely a vehicle for education or rehabilitation. We should be even more concerned when these activities are about “normalization”. Twisting play to fit adult agendas vitiates it (Eberle, 2014) and can be oppressive. In the final section of this chapter, I consider how we might emancipate play for disabled children from the various barriers to their play, including the “play-as-progress” rhetoric.

## Emancipating play for disabled children

Happily, there is much evidence that disabled children’s play can be liberated from many of the barriers discussed within this chapter. Well-designed playgrounds (Burke, 2013; Moore and Lynch, 2015), improvements to the “journey to play” (e.g., affordable and accessible local transport) (Stafford, 2017), welcoming actions by communities (Burke, 2013), assistive technologies (Chantry and Dunford, 2010) and inclusive design of toys and games (Holt et al., 2014) can help. Professionals such as occupational therapists can support disabled children to achieve their own play goals (Missiuna and Pollock, 1991). Teachers can devise and implement anti-bullying strategies (Good et al., 2011).

Around the world, efforts are being made to challenge disabling barriers and create opportunities for inclusive play. It is increasingly recognized that building a physically accessible playground will not guarantee inclusion of/for disabled children. Imaginative resolutions and approaches are being developed—and are to be celebrated. In Israel, for example, the non-profit Beit Issie Shapiro has built an inclusive playground called Park Chaverim (Friendship Park) and collaborated with Sesame Street (Israel)<sup>9</sup> to develop an educational and social program that aims to raise awareness and help children to learn how to play together. This program includes inclusive get-togethers of families of disabled children and families of non-disabled children, inclusive festivals and visits to the Park by “Sivan”, one of the Sesame Street Muppets. Sivan is a young girl who uses a wheelchair. The model of Park Chaverim has been replicated across Israel and internationally.

### Play for disabled children in Hong Kong and Taiwan

Headline findings of research conducted in 2018. The project involved a survey of parents of disabled children in Hong Kong and Taiwan, exploring their attitudes towards play and perceptions of the barriers to play for their child (see Beckett et al, 2020 for full findings and discussion).

In both locations 98% of parents said that they believed it was ‘important’ or ‘very important’ that their disabled child had opportunity to play. The majority of parents said they preferred inclusive play-spaces (for children of all abilities) for their child: 73.9% (Taiwanese parents) and 88.5% (Hong Kong parents). Parents who preferred ‘segregated’ playgrounds (for disabled children only), tended to say that they believed these playgrounds are safer.

Of the parents who preferred inclusive playgrounds, there were three main justifications provided – inclusive playgrounds:

Enable disabled children to make friends with and enjoy the company of non-disabled children;

Are important for social inclusion;

Help prevent disabled children feeling ‘different’ e.g. a Hong Kong parent commented that she didn’t ‘want to make him feel that he’s a special kid’ and a Taiwanese parent observed ‘the kid has his pride, he doesn’t want to stand out’.

Yet the parents acknowledged barriers to play for their children. The main barriers were the attitudes of other children and adults. 69.8% of Taiwanese and 68.3% of Hong Kong parents said that the ‘negative attitudes or lack of understanding of other children’ were a barrier to their child’s play.

45.3% of Taiwanese and 50% of Hong Kong parents said that the attitudes of adults were also problematic. ‘Free text’ comments from parents relating to attitudinal barriers included that ‘community education’ was needed ‘so that citizens may know and accept’ disabled children (Hong Kong parent) and that ‘empathy’ (Taiwanese parent) ought to be promoted within society.

In the high-density urban populations of both Taiwan and Hong Kong, insufficient space for play in the home was highlighted by parents. Just over 40% of Taiwanese parents and 54.1% of Hong Kong parents said that they were either ‘dissatisfied’ or ‘very dissatisfied’ in this regard.

‘Free text’ comments by Hong Kong parents regarding playgrounds included that ‘places for children’s play are too small and too far away’ and ‘space in parks with sensory play facilities is too small’. A Taiwanese parent observed that ‘most park entrances are gated to prevent scooter riders entering, but it also keeps people with physical disabilities out’.

Parents from both locations reflected on the impact on children and parents of living in societies that prioritise academic achievement. 53.7% of Taiwanese and 52.5% of Hong Kong parents reported that their child plays for less than an hour a day. This was a particular concern amongst Hong Kong parents whose comments included the following:

‘The child has high study pressure. Parents feel troubled. Relationship is affected and so is the child’s self-esteem and self-confidence. Through play, children can release some of the pressure. Hope that the government can adjust the aims of education and let children have real time to play.’

‘Parents and children with learning disabilities are very much troubled by daily homework.’

Over half of respondents (59.1% of Taiwanese parents; 63.9% of Hong Kong parents) stated that they did not have enough time to support their child’s play. Our findings present a picture of time-poverty amongst many mothers, resulting from employment, wider family caring responsibilities, and housework.

<sup>9</sup>See <https://youtu.be/SB5R7R5TZ4I> for a short film about this collaboration. (accessed 19/07/22).



In the Netherlands the Speeltuinbende<sup>10</sup> (The Playground Gang) test the accessibility and inclusivity of playgrounds. The gang is composed of disabled and non-disabled children and is supported by the Dutch Foundation for Disabled Children. The gang visits playgrounds twice—before and after renovation—to assess if the playability has been improved for both disabled and non-disabled children. The gang explore the playgrounds whilst carrying signs. After playing for a while, they come to a verdict on aspects of the playground: either showing the green (positive) or the red (negative) side of their signs. If a playground meets the approval of the Playground Gang, a green sign is erected at the site, which reads “Here we can play well together” (see Fig. 4).

These are but two of many examples of innovative practice facilitating disabled children’s right to play and, importantly, encouraging *inclusive play* between disabled and non-disabled children. In other places in the world there are examples of initiatives designed to enable disabled children’s right to play that operate within (constrained by) existing segregated systems. An example of this is an initiative run by the Rotary Club of Nagpur, India. This club has been working with special schools in the region to run an annual event of sports, games and play activities for disabled children (see Fig. 5 from the 2012 event). The name for the annual event is “Udaan”, which means “flight” and signals, according to the former President of the Club Kapil Bahri, that the event allows the children to “soar above their day-to-day mundane challenges”.<sup>11</sup> Whilst the events may not be inclusive in the sense of bringing together disabled and non-disabled children, they are an indication of progress being made toward recognizing disabled children’s “right to play” and, in their adoption of mixed ability games, these events open the possibility of more inclusive events in the future.

These initiatives, from around the world, are encouraging. Nevertheless, the thorniest barrier to disabled children’s play remains the capacity of the play-as-progress rhetoric to limit disabled children’s access to free play and to repurpose play for their “normalization”. We need to challenge the grand narratives of play that prioritize child development over child wellbeing. Supporting play researchers and practitioners to think critically about the play-as-progress rhetoric and refocus their efforts on matters of play, equity and quality of life, will be essential.

Adults have an important role to play in emancipating play for disabled children. Nevertheless, it is also important to recognize that disabled children are not always passive in the face of processes and practices shaped by the “play-as-progress” rhetoric. We know, from wider research, that disabled children resist disabling discourses and, at times, reconstruct them to fit their experiences and priorities (Priestley, 1998). Research has revealed that disabled children resist processes that seek to govern (“adulterate”) their play, exerting control over and creativity within their play (Burke, 2012). Future play researchers and practitioners concerned with



**Fig. 4** De Speeltuin bende sign, Donald Trung Quoc Don (Chữ Hán: 徵國單) - Wikimedia Commons, CC BY-SA 4.0 ([https://commons.wikimedia.org/wiki/File:De\\_Speeltuin\\_bende\\_sign,\\_Deventer\\_28201929.jpg](https://commons.wikimedia.org/wiki/File:De_Speeltuin_bende_sign,_Deventer_28201929.jpg)).

<sup>10</sup>See <https://www.speeltuinbende.nl/> for information about the gang.

<sup>11</sup>Quoted in Rotary News: <https://rotarynewsonline.org/a-special-event-by-rc-nagpur-for-special-children/> (accessed 17/07/22).





**Fig. 5** Children take part in the Udaan event 2012 (Rotary Club Nagpur via Flickr CC BY 2.0).

emancipating disabled children’s play might usefully begin by understanding *how* disabled children demonstrate resistance in and through, control over and creativity within their play and then seek to support them in these regards. This section ends with two final vignettes that explore this relationship between play and resistance (Vignettes 4 and 5).

#### Vignettes 4 & 5 Play and Resistance

The next two stories are based upon video-based observational research. The videos were made by a research team based in Portugal. They were exploring whether toy robots might offer augmented manipulation opportunities for children with motor impairments (e.g. cerebral palsy) helping them to engage in independent exploration and participate play activities. The videos were shared with me for secondary analysis (see acknowledgments). In what follows children’s names have been altered to ensure anonymity.

V4: the location is a primary (elementary) school. The setting is a quiet space away from the noise and activity of the main classroom. On a table, a mini-obstacle course has been set up. It looks “fun”. A small robot has been placed at one end of a table. A child, Sofia, sits at the end of the table closest to the robot. Her impairment means that she has difficulties with motor functions.

The researcher, a young woman, gives Sofia instructions about the “game”. She encourages her to use the adapted controls to navigate the robot around the obstacle course.

Sofia looks impassive. Eventually she begins to move the robot. The robot runs over the edge of the table, falling to the floor.

The researcher reassures Sofia not to worry. No harm done. She can have another go. The robot is returned to the starting line. The same thing happens again.

The researcher clearly thinks that Sofia is having some difficulties understanding how the controls work and offers guidance. They try again. Sofia navigates the robot off the side of the table.

The “penny drops” for the researcher—Sofia can use the controls and is intentionally sending the robot over the table edge!

Instead of resetting the scenario and encouraging another attempt with the robot, the researcher asks Sofia “what would you like to be doing?” Sofia says clearly “not this” and indicates she’d prefer to go back to the classroom to be with the other children.

~~~~~

V5: (a different school). A young boy, Martim, is sitting at a computer console. He is a wheelchair user and, like Sofia, his impairment means that he has difficulties with motor function. On the floor in front of him an obstacle course has been laid out. There is a physical perimeter to the course—a set of low barriers. He is using an adapted set of controls to navigate a small robot around the course.

Also present in the room is his teacher and some of his friends. His friends have come to watch and see him demonstrate the robot and his skill with the controls.

The robot begins to move. Instead of advancing around the course, it edges toward the perimeter barrier and speeds up so that it jumps over it.

The teacher quickly picks up the robot and puts it back onto the “official” course. The researchers check that Martim is not having difficulties with the controls.

Again the robot moves. Again it jumps over the barriers. The teacher is clearly frustrated. There is a sense of tension or anxiety, as the adults assume that Martim is not able to control the robot properly and is at risk of exposing some lack of ability to his peers.

They start again. Martim again sends the robot over the barriers. His friends laugh with glee. They have realized, before the adults, that Martim is being mischievous. He is doing this intentionally. An act of resistance appreciated by his peers.

Conclusion

Disabled children, like all children, have the right to play. All disabled children *can* play. Sadly, not all disabled children *do* play. This is worrying. Play has myriad benefits – the most important of which, it is argued here, is that it supports children’s mental health and wellbeing. That said, we must remember that play isn’t all “ha ha hee hee” (to paraphrase the title of Meera Syal’s novel). Play is

a phenomenon exhibiting social structures similar to those of wider life (Henricks, 2006). Oppression and exclusion, including ableism and disablism, can be (re)produced and experienced in play-worlds (Hodge and Runswick-Cole, 2013). Our challenge is to explore how play can be emancipated for disabled children whilst recognizing its ambiguities. Our focus needs to be on play's potential for resistance; for it to be an opportunity for disabled children to take control of their activities; for it to be a vehicle for their creativity and innovation, opening "lines of flight" from normalization (Guattari, 2016). Play, at its best, has the capacity to be a site for "revolutionary effervescence" (Guattari, 2016: 68)—this is why play matters for disabled children, alongside children from all marginalized groups.

I end on a practical note with a set of recommendations. If we are to emancipate play for disabled children, then we need to do the following:

- (1) Recognize disabled children's "right to play"—play free of adult agendas.
- (2) Reject the deficit-model of disabled children's play.
- (3) Challenge the dominance and worst excesses of the "play-as-progress" rhetoric.
- (4) Enhance training for professionals who are in a position to facilitate disabled children's right to play (e.g., play-workers, teachers, occupational therapists), in particular, address the problem of over-protection (which limits disabled children's play experiences) and encourage them to think critically about their own practices (are they prioritizing child development/normalization over child wellbeing?).
- (5) Provide support for parents/guardians so that they are able to resist the play-as-progress' rhetoric, to facilitate their disabled child's free play.
- (6) Acknowledge that even the most well-intentioned initiatives promoting play for all children, may be informed and limited by ableist assumptions, inadvertently excluding disabled children.
- (7) Prioritize inclusion and inclusive methods during the design of toys, games, play-spaces and play opportunities.
- (8) Recognize that inclusion is about much more than physical accessibility: we need to change attitudes (amongst adults and peers) to ensure that disabled children feel included in play.
- (9) Ensure that families are able to access inclusive and affordable toys, play-spaces and play opportunities.

If we achieve 1–9, then we will truly have "come a long way ... "

Acknowledgments

I would like to thank Dr. Pedro Encarnação for sharing the video recordings from which I derived Vignettes 4 and 5. Thank you also to Robin Jerstad for granting me permission to use Fig. 1 in this chapter.

References

- Barron, C., Beckett, A.E., Coussens, M., Desoete, A., Cannon Jones, N., Lynch, H., Prellwitz, M., Fenney Salkeld, D., 2017. Barriers to Play and Recreation for Children and Young People With Disabilities: Exploring Environmental Factors. De Gruyter, Warsaw/Berlin.
- Baumeister, A.L., Storch, E.A., Geffken, G.R., 2008. Peer victimization in children with learning disabilities. *Child Adolesc. Soc. Work. J.* 25, 1–23.
- Beckett, A.E., Encarnação, P., Chiu, C., Ng, S.T., 2020. Play for disabled Children in Taiwan and Hong Kong: parent perspectives. *J. Disability Stud. Educ.* 1 (1–2), 90–120.
- Besio, S., 2017. The need for play for the sake of play. In: Besio, S., Bulgarelli, D., Stancheva-Popkostadinova, V. (Eds.), *Play Development in Children with Disabilities*. De Gruyter, Warsaw/Berlin, pp. 9–52.
- Brotherson, M.J., Cook, C.C., Erwin, E.J., Weigel, C.J., 2008. Understanding self-determination and families of young children with disabilities in home environments. *J. Early Interv.* 31, 22–43.
- Burke, J., 2012. Some kids climb up; some kids climb down": culturally constructed play-worlds of children with impairments. *Disabil. Soc.* 27 (7), 965–981.
- Burke, J., 2013. Just for the fun of it: making playgrounds accessible to all children. *World Leisure J.* 55 (1), 83–95.
- Chantry, J., Dunford, C., 2010. How do computer assistive technologies enhance participation in childhood occupations for children with multiple and complex disabilities? A review of the current literature. *Br. J. Occup. Ther.* 73 (8), 351–365.
- Connors, C., Stalker, K., 2003. *The Views and Experiences of Disabled Children and Their Siblings: A Positive Outlook*. Jessica Kingsley, London.
- Contact a Family, 2012. A Guide to Dealing With Bullying: For Parents of Disabled Children. <https://contact.org.uk/media/388418/bullying.pdf>. (Accessed 18 March 2019).
- Crosbie, J.P.G., 2014. *The Value of Outdoor Education for People With Disabilities: An In-Depth Case Study of the Calvert Trust*. PhD Thesis. University of Edinburgh.
- Dunn, K., Moore, M., 2005. Developing accessible play space in the UK: a social model approach. *Child. Youth Environ.* 15 (1), 331–354.
- Eberle, S.G., 2014. The elements of play: toward a philosophy and a definition of play. *Am. J. Play* 6 (2), 214–233.
- Else, P., Sturrock, G., 1998. The playground as therapeutic space: playwork as healing. In: Paper Presented at Play in a Changing Society: Research, Design & Application, IPA/USA National Conference, Longmont, CO, 17–21 June. <https://ipaewni.files.wordpress.com/2016/05/colorado-paper.pdf>. (Accessed 18 March 2019).
- Frost, J.L., Wortham, S.C., Reifel, S.C., 2012. *Play and Child Development*, fourth ed. Pearson Education, New Jersey, US.
- Garvey, C., 1990. *Play*. Harvard University Press, Cambridge, Massachusetts.
- Geary, P., 2012. CRC in court: the case law of the convention on the rights of the child. CRIN. https://archive.crin.org/docs/CRC_in_Court_Report.pdf.
- Gleave, J., 2010. Making it Our Place: Community Views on Children's Play. <http://socialwelfare.bl.uk/subject-areas/services-client-groups/children-young-people/playengland/making10.aspx>. (Accessed 31 March 2015).
- Good, C.P., McIntosh, K., Giet, C., 2011. Integrating bullying prevention into schoolwide positive behavior support. In: Council for Exceptional Children. <https://njbullying.org/documents/BullyingPosBehSupports.pdf>. (Accessed 19 July 2022).
- Goodley, D., 2014. *Disability Studies. Theorising Disablism and Ableism*. Routledge, Abingdon.
- Goodley, D., Runswick-Cole, K., 2010. Emancipating play: disabled children, development and deconstruction. *Disabil. Soc.* 25 (4), 499–512.
- Guattari, F., 2016. *Lines of Flight: For Another World of Possibilities* (Trans. A. Goffey). Bloomsbury, London.
- Henricks, T.S., 2006. *Play Reconsidered: Sociological Perspectives on Human Expression*. University of Illinois.

- Hodge, N., Runswick-Cole, K., 2013. 'They never pass me the ball': exposing ableism through the leisure experiences of disabled children, young people and their families. *Child Geogr.* 11 (3), 311–325.
- Holt, R.J., Moore, A.M., Beckett, A.E., 2014. Together through play: facilitating inclusive play through participatory design. In: Langdon, P., Lazar, J., Heylighen, A., Dong, H. (Eds.), *Inclusive Designing*. Springer, Cham.
- Hughes, B., 2003. Play deprivation, play bias and playwork practice. In: Brown, F. (Ed.), *Playwork Theory and Practice*. Open University Press, Maidenhead.
- Hughes, R., 2012. *Evolutionary Playwork*, second ed. Routledge, Abingdon.
- International Play Association (IPA), 2015. Position Statement on the Play Rights of Disabled Children. <http://ipaworld.org/ipa-position-statement-on-the-play-rights-of-disabled-children/>. (Accessed 3 January 2017).
- Jeanes, R., Magee, J., 2012. 'Can we play on the swings and roundabouts?': creating inclusive play spaces for disabled young people and their families. *Leisure Stud.* 31 (2), 193–210.
- John, A., Wheway, R., 2004. *Can Play, Will Play: Disabled Children and Access to Outdoor Playgrounds*. National Playing Fields Association, London.
- Kuhaneck, H.M., Spitzer, S.L., Miller, E., 2010. *Activity Analysis, Creativity and Playfulness in Pediatric Occupational Therapy. Making Play Just Right*. Jones and Bartlett Publishers, Sudbury, Massachusetts.
- Law, M., Haight, M., Milroy, B., Williams, D., Stewart, D., Rosenbaum, P., 1999. Environmental factors affecting the Occupations of children with physical disabilities. *J. Occup. Sci.* 6, 102–110.
- Lawson, A., Beckett, A.E., 2021. The social and human rights models of disability: towards a complementarity thesis. *Int. J. Hum. Right.* 25 (2), 348–379.
- Ludvigsen, A., Creagan, C., Mills, H., 2005. Let's Play Together: Play and Inclusion – Evaluation of Better Play Round Three. Barnardo's. <https://www.inclusivechoice.com/Lets%20play%20together%20-%20play%20and%20inclusion.pdf>.
- Mason, M., 1990. Internalized Oppression. <https://disability-studies.leeds.ac.uk/wp-content/uploads/sites/40/library/Mason-Micheline-mason.pdf>. (Accessed 19 July 2022).
- McLaughlin, J., Coleman-Fountain, E., Clavering, E., 2016. *Disabled Childhoods: Monitoring Differences and Emerging Identities*. Routledge, Abingdon.
- Mencap, 2007. *Bullying Wrecks Lives: The Experiences of Children and Young People with Learning Disabilities*. Mencap, London.
- Missiuna, C., Pollock, N., 1991. Play deprivation in children with physical disabilities: the role of the occupational therapist in preventing secondary disability. *Am. J. Occup. Ther.* 45 (10), 882–888.
- Moore, A., Lynch, H., 2015. Accessibility and usability of playground environments for children under 12: a scoping review. *Scand. J. Occup. Ther.* 22 (5), 331–334.
- Olsen, H.M., Dieser, R.B., 2012. "I am hoping you can point me in the right direction regarding playground accessibility": a case study of a community which lacked social policy toward playground accessibility. *World Leisure J.* 54 (3), 269–279.
- Ozen, A., Ergonekon, Y., Ulke-Kurkcuoglu, B., Genc, D., 2013. Opinions of special education teachers about activity-based intervention. *Hacettepe Univ. J. Educ.* 44, 262–274.
- Prellwitz, M., Skar, L., 2007. Usability of playgrounds for children with different abilities. *Occup. Ther. Int.* 14 (3), 144–155.
- Prellwitz, M., Tamm, M., 1999. Attitudes of key persons to accessibility problems in playgrounds for children with restricted mobility: a study in a medium-sized municipality in northern Sweden. *Scand. J. Occup. Ther.* 6 (4), 166–173.
- Priestley, M., 1998. Childhood disability and disabled childhoods: agendas for research. *Childhood* 5 (2), 207–223.
- Richardson, P.K., 2002. The school as social context: social interaction. Patterns of children with physical disabilities. *Am. J. Occup. Ther.* 56 (3), 296–304.
- Rigby, P., Gaik, S., 2007. Stability of playfulness across environmental settings. a pilot study. *Phys. Occup. Ther. Paediatr.* 27 (1), 27–43.
- Rimmer, J.H., Riley, B., Wang, E., Rauworth, A., Jurkowski, J., 2004. Physical activity participation among persons with disabilities: barriers and facilitators. *Am. J. Prevent. Med.* 26 (5), 419–425.
- Santer, J., Griffiths, C., Goodall, D., 2007. *Free Play in Early Childhood. A Literature Review*. National Children's Bureau, London.
- Schleien, S.J., Miller, K.D., Walton, G., Pruett, S., 2014. Parent perspectives of barriers to child participation in recreational activities. *Ther. Recreat. J.* xlviii (1), 61–73.
- Shackell, A., Butler, N., Doyle, P., Ball, D., 2008. *Design for Play: A Guide to Creating Successful Play Spaces*. <http://www.playengland.org.uk/resources/design-for-play.aspx>. (Accessed 31 March 2015).
- Slater, J., 2015. *Youth and Disability. A Challenge to Mr Reasonable*. Routledge, Abingdon.
- Spencer-Cavaliere, N., Watkinson, E.J., 2010. Inclusion understood from the perspectives of children with disability. *Adapt. Phys. Act. Q. (APAQ)* 27 (4), 275–293.
- Stafford, L., 2017. Journeys to play: planning considerations to engender inclusive playspaces. *Landsc. Res.* 42 (1), 33–46.
- Steinfeld, E., Maisel, J.L., 2012. *Universal Design. Creating Inclusive Environments*. John Wiley and Sons, Hoboken, New Jersey.
- Sutton-Smith, B., 1997. *The Ambiguity of Play*. Harvard University Press, Cambridge, Massachusetts.
- UN CRC (1989). Online at: <https://www.ohchr.org/en/professionalinterest/pages/crc.aspx> (accessed 18/03/2019).
- UN CRPD (2006). Online at: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/convention-on-the-rights-of-persons-with-disabilities-2.html> (accessed 15/10/2017).
- UN CRC, 2013. General Comment No. 17 on the Right of the Child to Rest, Leisure, Play, Recreational Activities, Cultural Life and the Arts (art. 31). Online at: https://www.playscotland.org/resources/print/UNCRC-General-Comment-17-on-Article-31.pdf?plscml_id=20587. (Accessed 19 July 2022).
- UN CRPD Committee, 2016. Concluding Observations on the Initial Report of the United Kingdom of Great Britain and Northern Ireland. https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CRPD2fC2fGBR2fC02f1&Lang=en. (Accessed 19 July 2022).
- UNICEF, 2013. *The State of the World's Children 2013. Children With Disabilities*. Online: https://www.unicef.org/sowc2013/files/SWCR2013_ENG_Lo_res_24_Apr_2013.pdf. (Accessed 15 October 2017).
- Von Benzon, N., 2010. Moving on from ramps? The utility of the social model of disability for facilitating experiences of nature for disabled children. *Disabil. Soc.* 25, 617–626.
- Wenger, I., Schulze, C., Lundström, U., Prellwitz, M., 2021. Children's perceptions of playing on inclusive playgrounds: a qualitative study. *Scand. J. Occup. Ther.* 28 (2), 136–146.
- Whitebread, D., Basilio, M., Kuvalja, M., Verma, M., 2012. *The Importance of Play. Report for the Toy Industries of Europe*. <http://www.csap.cam.ac.uk/media/uploads/files/1/david-whitebread-importance-of-play-report.pdf>. (Accessed 12 March 2019).
- Woolley, H., 2013. Now being social: the barrier of designing outdoor play spaces for disabled children. *Child. Soc.* 27 (6), 448–458.
- Yantzi, N.M., Young, N.L., McKeever, P., 2010. The suitability of school playgrounds for physically disabled children. *Child Geogr.* 8 (1), 65–78.

Everybody's got to learn sometime: disability youth activism and the pursuit for inclusive education

Miro Griffiths, School of Sociology and Social Policy, University of Leeds, Leeds, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	320
Conceptualizing inclusive education through disability youth activism	321
Embedding disability youth activism within inclusive education	323
Conclusion	325
References	325

Introduction

Between 2010 and 2012, I was a co-investigator on an Erasmus Plus (Leonardo) project exploring professional partnerships for inclusive education. My role was to attend educational settings across Europe and observe, politely interrogate, and record school approaches to “inclusive” education. On an exceptionally hot, spring day I arrived at the entrance to a school in Reggio Emilia, Italy. It was presented, by teachers and school administrators, as an inclusive school with *minimal* segregation between disabled and non-disabled learners. The contradiction was not lost on me. I was interested to know how the school—inspired by the Reggio Emilia approach to education (Malaguzzi, 1993)—prioritized the inclusive learning and full participation of all learners. We sat at a table in a courtyard; the teachers, sensibly, had taken seats in the shade whereas I, regrettably, positioned my wheelchair in the direct sunlight—with no hat for protection. The teachers spoke, enthusiastically, about developing individual education plans for learners with impairments and diagnostic labels. These plans were open to continuous reassessment and review, justified because learning objectives can be too high and disabled learners can struggle to keep up with the established curriculum. The teachers moved the conversation onto the importance of socializing before I had time to respond. They declared, with pride, that all learners had their meals together. I nodded and asked if “being together” is reflected also in the classroom. There was a pause. “No”, a teacher replied, “disabled students can be part of the classroom for one, 2 h, maybe, sometimes 15 min of an hour, then they work in a separate room. It has computers and better support.” I asked why adopt this approach. A different teacher responded, “Because ... Disabled ... People can be disturbing. You can’t always have disabled children disturbing other kids.” All the teachers, and administrators, nodded in agreement. The response, directed to a researcher with visible impairments, sitting in a power wheelchair, surrounded by assistive technologies and human support, was ironic and all too familiar. Later in the day, sat waiting for an accessible vehicle to collect me from the school, a teacher—acting as a chaperone—stood patiently at my side, tapping their foot to complement the sounds of the agricultural machinery in the nearest field. “Do local disabled activists influence your school’s approach to learning? Do they assist in supporting your students to have a voice in organizing the school?” I asked. The teacher shook her head, “No. No activists are here.”

The narrative above reflects a central thread in this chapter. It is interesting to consider to what extent disability activism would influence the production of inclusive educational arrangements. Disability activism remained absent in the local context outlined above. Education, here, was predicated on a tolerance for exclusion (Slee, 2019). Disabled people positioned as deviant, disruptive, and problematic to the ordering of school learning were a hindrance and threat to other people’s educational pursuit. Although unlikely to address the entirety of restrictions, exclusionary practices, and prejudiced attitudes experienced by disabled learners, activism has the potential to destabilize current educational configurations and introduce alternative, arguably preferable, possibilities for organizing education. The role of disability activism is integral to the conceptualization of inclusive education, as well as resisting the exclusionary practices and policies that permeate educational environments.

The chapter explores how young disabled activists conceptualize inclusion, as well as the practices they employ to pursue inclusive and emancipatory agendas. This is explored within the context of education, and consideration is given to how disability activism complements the pursuit for inclusive education. The next section, following this introduction, explores and critiques young disabled activist’s conceptualization of inclusive education with a particular focus on two specific aspects. Firstly, the purpose of inclusive education and, secondly, the significance of disabled people’s participation in education environments and knowledge exchanges. The chapter then considers the possibilities and opportunities for inclusive education and disability youth activism to coalesce, with particular attention on recognizing the importance of young disabled people’s ideas and perspectives within the production of inclusive educational arrangements. This chapter draws on literature investigating young disabled people’s participation in disability activism and social movements (Griffiths, 2022).

The conceptualization of inclusive education explored in this chapter recognizes disability as a social justice issue. The chapter’s central argument is that inclusive education approaches must facilitate young disabled people’s participation within activism and social movements. Inclusive education has the potential to support young disabled people to determine disability as political, and to contribute to the claims and activities to realize emancipation. Inclusive education can, and should, produce accessible

and participatory knowledge exchanges for young disabled people to become aware of disability politics and the injustices that emerge because of how society is organized. Education serves as an opportunity to reflect the principles of inclusion, justice, fairness, and participation within and beyond education. Disability youth activism provides routes to disrupt exclusionary education policies and contributes to the realization of inclusive education. To achieve such disruption, and introduce possibilities for progressive change, consideration is needed to ensure young disabled people have accessible opportunities to influence the organization of education.

Understanding disability, in this chapter, draws upon the social model of disability. The social model is key to the development of disability activism in the UK. The principles, which underpin the model, articulate disability as a product of social conditions, imposed upon people with impairments. This creates unnecessary isolation and exclusion for disabled people, who remain an oppressed group due to how society is organized (UPIAS, 1975). The model remains influential in academia, notably within Disability Studies (Owens, 2015). Activism, in this chapter, draws on Tilly and Tarrow's (2015) definition of social movements and the contentious politics that emerge through activist relations. Contentious politics refers to the relations that exist between and beyond activists, who are engaged in asserting and maintaining claims for social change. On a broader scale, social movements constitute: (a) repeated attempts of claim making; (b) strategies to disseminate claims publicly; (c) maintaining collective solidarity; and (d) resource mobilization to sustain claim making and their related activities.

Conceptualizing inclusive education through disability youth activism

The discourse that surrounds conceptualizing education is replete with tensions over who is considered an acceptable presence within the educational environment (Miles and Singal, 2010). This means the debates and aspirations to determine how learning should be facilitated, how school should be organized, and why knowledge should be exchanged, often reject, restrict, or excuse the participation of disabled learners (Byrne, 2022). Discussions about disabled people's inclusion in education become relegated to the technical details surrounding an integrative approach to educational provision (Oliver and Barnes, 2010). This, ultimately, distracts from establishing a broader narrative as to how education ought to be organized—particularly—from the perspective of those who experience daily injustices in the pursuit of learning. In this section, attention is given to how disability youth activism can produce insight into the purpose of inclusive education, as well as the dynamics of participation within learning environments.

It is possible, through exploring young disabled activists' accounts of inclusion, to learn from and centralize their ideas to belong, participate, and resist within educational arrangements. In Griffiths' (2019) study of UK disability youth activism, the premise of inclusion constituted four principles. Firstly, developing a sense of pride and acceptance across all identities. Young disabled activists argued that inclusion requires embracing and celebrating all variance and diversity within existence. It remained paramount that individuals, and collectives, are encouraged and supported to articulate their identities; and that a sense of pride is produced and maintained throughout the environment. This reflects the importance of an affirmative identity, wherein impairment is not assumed to be tragic and disabled people are not objectified through a lens of pity, sorrow, and tragedy (Swain and French, 2000).

Secondly, progressing rights to realize positive and negative freedoms. Here, the premise of rights is used to further the full participation of disabled people and remove infringements that would, typically, deny disabled people a valued and respected presence in their communities. Young disabled activists argued for rights to produce mechanisms and interventions that would facilitate disabled people's engagement in education, such as the right to access support and the right to participate in educational settings. Alongside this was the application of rights to remove forms of interference that would deny or impede disabled people's participation in education. Hirschmann (2016) argues for a freedom approach to rights to realize inclusive and socially just societies. This, they argue, establishes a basis that disabled persons are free, equal human beings who cannot be denied their status as rights-bearers.

Thirdly, reconfigure the political and economic conditions that produce marginalization and oppression. Young disabled activists emphasized the necessity to identify and challenge the political and economic arrangements that (re)produce restrictions experienced by the disabled people's community. This reinforces a central tenant within disability politics, which is to understand disabled people's marginalization as a result of historical and contemporary social world organization (Lawson and Beckett, 2021). The neoliberal influence upon education is well documented (Roda, 2017; Hedegaard-Soerensen and Grumloese, 2020; Ball and Grimaldi, 2021), particularly within the context of disability (Liasidou, 2015; Kearl, 2019). Strategies for developing fair and socially just societies must incorporate an assessment of the political and economic influences on communities (Graeber, 2009).

The final principle is to build sustainable campaigns to maintain dialog and activities focusing on realizing inclusive societies. Here, young disabled activists insisted on configuring campaigns and agendas that would promote, continuously, the importance of inclusion and emancipation. This was an attempt to recognize the ongoing imperative to promote claims for realizing inclusive societies, and to never assume inclusion will be realized or maintained without constant vigilance and action. The perpetual need to contest injustices, and promote claims for inclusion, is an attempt to keep the importance of inclusion at the forefront when it comes to organizing education. It requires strategies to highlight why, and how, inclusion is necessary; however, it also illustrates that inclusion can often be reversed or dismantled (Peruzzo, 2020). Building campaigns to promote inclusion are—to some extent—a response to Schmitt's (2003) argument that processes of alienation lead people to accept, and advocate for, conformity.

Accepting conformity limits the possibilities of reorganizing spaces, producing alternative possibilities, and resisting the arrangements that produce unjust and inaccessible societies.

These four principles highlight the practices that are required to realize, and sustain, inclusive environments. They encompass arrangements associated with traditional, albeit structural, understandings of social world organization (Oliver and Barnes, 2012). They also engage in the (re)claiming of the subjective experiences of disability, contesting the categories produced to define disabled people and their position in society (Feely, 2016). There are consistent attempts to disrupt and rework existing arrangements (political, economic, and legislative) to realize inclusive education, as well as introduce new strategies and resistance practices to promote and celebrate disabled people's identities and their participation in education. These principles are evident of radical and incremental pathways to realize inclusion. Young disabled activists are establishing approaches that allow for flexibility and creativity in the pursuit of their goals, instead of adopting an either/or approach.

Promoting flexibility and creativity can be read as a response to the existing struggles in inclusive education literature. Understanding inclusion can become trapped in irresolvable debates about the purity of definitions of inclusive education and precise and binding blueprints for building educational institutions (Slee and Allan, 2001). The principles outlined by young disabled activists should be read as options for experimental play: things to work through, embed, and test within educational environments. Building pedagogical approaches to explore, promote, and celebrate disability pride is one option. So too is an approach that explores how learning opportunities, critical thinking, and knowledge exchanges can be protected by utilizing existing legislative infrastructures and calling for further radical alternatives. Inclusion, here, necessitates an exploration of the dangers and implications of current and emerging educational practice. It attempts to unsettle the existing arrangements that constitute education while, simultaneously, offering disabled learners an influential position to produce new ways of thinking about disability, education, inclusion, and participation.

There are similarities between the principles proposed by young disabled activists and the definitions of social justice education employed by Zion and Blanchett (2017). Their approach is to consider the extensive forms of social exclusion that are tolerated throughout social organization, and which are reflected in the production of schools. Equitable access to education does not exist for disabled students and leaves them struggling to engage in a learning environment, and curriculum, that gives little attention to the significance of social justice. Inclusive education requires a commitment to embed a social justice critique across all educational arrangements. This, it is argued by Zion and Blanchett (2017), will encourage a re-examination of the purpose of education, a reconfiguration of the curriculum, and a review of how education is situated across political, economic, cultural, and social arrangements in society.

The significance of social justice education is found within the broader discourses surrounding school organization. deMarrais and LeCompte (1999) and Spring (2013) highlight an egalitarian imperative within education, which attempts to position educational practice as a regulator of opportunity. It is designed to produce and promote opportunities for social mobility, improved quality-of-life, and sustainable community development (Adebayo, 2020). There are also humanistic and civic imperatives, which aim to position education as a right. The right to access education is coupled with responsibilities and duties, imposed upon the learner, to utilize their education to better society (Blais and Labbé-St-Vincent, 2011). The humanistic and civic imperatives are centered upon a notion of preparing learners to reciprocate and give back to their respective communities. This also mirrors the expectations found within the economic imperative of education. Learners are expected to develop skills, competencies, and attributes necessary for engaging within the existing economic model (Merrill-Glover and Edwards, 2015). Their education is predicated on participating within the labor market and contributing to economic output. Social justice, in this regard, is about broadening participation for communities to access and remain in the labor market.

The imperatives outlined above utilize social justice to widen learner opportunities to access education, embed responsibility and civic duty within the student, and prepare them for contributing to existing economic arrangements. These approaches to social justice must be read alongside the continuation of practices to categorize learner abilities (Kilinc, 2019), standardize pedagogical approaches (Molbaek, 2018), and legitimize certain knowledge exchanges over others (Mølsted et al., 2017). Scholars have highlighted how inclusive education practice undermines the principles of inclusion and social justice (Armstrong et al., 2010; Slee, 2018); however, the absence of a disability politics critique is of relevance here.

It is important to analyze, by returning to the principles of inclusion as articulated by young disabled activists, how broader conceptualizations of social justice and inclusion disregard the nuances found within disabled people's experience. Inclusion and social justice, within education, rarely equips a disability politics critique to contest the contradictions and controversies found within existing education policy. There is a lack of empirical data, and literature, exploring the importance of disability politics within curriculums. This has implications for the production of disabled people's political identities, and the associated pride and celebration of disabled people's participation across educational institutions. It can also lead to a weakened or non-existent interrogation of the arrangements that reproduce normative and ableist (Wolbring, 2008) values throughout society—including in schools, colleges, and universities.

This can be reversed and the ideas, perspectives, and activities of disability youth activism, can be incorporated into education design, development, and delivery. Beckett's (2013) research with school-aged children identified non-disabled learners sustaining a privileged, and perceived superior, position over disabled people. Such findings illustrate that current pedagogy fails to challenge conceptualizations of disabled people as tragic, problematic, and deviant. Action is required to cement anti-oppressive pedagogies, which will disrupt contemporary understandings of disability and situate learning within social justice, inclusion, and emancipation. Beckett (2015) offers a typology of anti-oppressive pedagogies to envisage how teaching and learning could identify the arrangements, practices, and ideas that reproduce disabled people's marginalization and oppression. This is necessary for realizing

innovative forms of inclusive education, which—in turn—supports the organization of socially just, democratic, and accessible societies (Slee, 2011). The typology is influenced by the work of Kumashiro (2002) and sets out three approaches to anti-oppressive pedagogy.

The first pedagogy is centered upon the interrogation of *othering*, which often situates disabled people as a threat to the normative expectations of how one ought to be, become, and exist. To disrupt this, knowledge is centered upon the development of strategies to identify, and eradicate, stigma and prejudice toward disabled people. It attempts to celebrate difference and encourage acceptance in the various ways people live and participate in their communities. Beckett (2015) argues that this approach is useful for severing the link between impairment and deviancy. The second pedagogy examines the social arrangements that reproduce continuous patterns of inequality, social injustice, and disadvantage. It supports students to understand the relations of power that emerge, coalesce, and diffuse throughout society (Jerome and Lalor, 2020). This is useful for student cohorts to build consciousness about power, resistance, and the necessity to destabilize political, economic, social, and cultural arrangements that maintain oppressive conditions (Freire, 2000). Student activities are centered upon critiquing contemporary structures and arrangements that remain detrimental to disabled people's inclusion. Finally, the third pedagogy is associated with producing knowledge exchanges for the purpose of reimagining spaces, relations, and categories. Beckett (2015, p. 84) promotes this pedagogy as a way for learners to "think otherwise." Teaching and learning, here, is not to impose one perspective over another; instead, it is to provide learners with the techniques and opportunities to question how the social world is organized and to introduce alternative possibilities—a plurality of futures (Motta, 2013). This is pertinent for questioning how, and why, disabled people are positioned in precarious, violent, and restrictive conditions. It also guides learning toward imagining new, alternative, forms of disabled people's participation and existence (Obourn, 2020).

There is complementarity between the views expressed by young disabled activists and the three pedagogies. Young disabled activists outlined, in their articulation of what constitutes inclusive practice, the importance of disability pride and the disruptive potential of celebrating disabled people's political identities. Embracing disability pride leads to respecting human variance and diversity (Martin, 2012), and has the potential for destabilizing processes of internalized oppression (Watermeyer and Görgens, 2014). Disability pride becomes a tool to disrupt notions of normality, conformity, and the value imparted upon specific attributes, traits, and characteristics. Through challenging stereotypes and prejudice, learning environments can be experimental spaces to redefine identities, resist discrimination, and articulate possible and preferable futures (Lorde, 1990). Emphasis is also placed on critiquing the political and economic conditions that exacerbate disabled people's marginalization. Young disabled activists call for inclusive practices to illustrate how existing social conditions reproduce unnecessary restrictions for disabled people. This reinforces Erevelles' (2000) claim for knowledge exchanges to critically engage with the inter-relationality of political, economic, and social structures for the purposes of understanding struggle and suffering. Learning, in this approach, opens possibilities for analyzing the significance of activism and disabled people's social movements. It encourages an exploration of rights-based approaches (Blanck, 2005), incremental policy measures (Harris and Roulstone, 2011), and radical claims (Withers et al., 2019) for social world overhaul. Young disabled activists also highlighted the importance of ongoing campaigns, and resistance practices, to maintain focus on producing inclusive practices. This could be read as a call for inclusive education to carve out safe and accessible spaces to critique, ponder, and propose alternative ways of organizing societies. It is to, as Beckett (2015, p. 84) suggests, "devise new (positive and non-oppressive) ways of thinking and acting about [oneself] and others; experiment with what and where a life can be/do/go."

The chapter now considers potential avenues for fostering sustainable and effective relations between disability youth activism and inclusive educational practice.

Embedding disability youth activism within inclusive education

Florian and Black Hawkins (2011) illustrate how inclusive pedagogy discourse often positions inclusivity as a process of reformatting curriculum and teaching practices to accommodate all learners. This is a flawed approach to understanding inclusive education. Instead, inclusive education can be informed by disability youth activism to progress learning and activities on an inclusive space, why it is necessary for social organization, and how emancipatory claims can be realized. The pursuit for inclusive education should consider young disabled people's desire to influence existing educational arrangements and practices. There are three issues to consider if the ideas, and practices, of disability youth activism is to inform inclusive education: problematic conceptualizations of youth; embedding notions of futures in curricula; and establishing spaces for disability activism to flourish within educational practice.

Conceptualizations of youth can restrict young people's participation within the design, organization, and delivery of projects and strategies (Müller-Bachmann et al., 2022). This is because children and young people are perceived as passive future beings, unable to absorb the nuanced arguments and commentaries expressed by authoritative and influential individuals and groups (Jenks, 2005). Research carried out by Slater (2015) highlighted that young people—in contemporary Western discourses—are often positioned as incomplete adults. They are deemed to not possess the attributes, characteristics, and values necessary for interrogating complex issues and providing solutions. Ascribing an incomplete status indicates that credibility, insight, and influence can be attained only after young people have transitioned through the border zones of youthhood. Alldred and David (2007) suggest this relegates young people to be understood as responsible people-in-waiting.

Having intuitive ideas, and the responsibility to coordinate activities, are understood as possessions—things to have after one has become an adult. They are not available to children and young people, or they are restricted to specific purposes and contexts. Young disabled people, within the context of UK disability activism, struggle to contribute and influence emancipatory agendas, strategies, and demands. Research findings (Todd et al., 2012; Griffiths, 2019) indicate that young disabled people are primarily involved in discussions and topics that require a dedicated youth perspective. This restricts their contributions to occasions that are focused on youth solely; it denies young people opportunities to engage within the discourses and practices beyond youth issues. Their narratives and commentaries are considered irrelevant, naïve, and ineffective if youth experiences are not of concern. The reluctance to facilitate broader youth participation, particularly beyond youth matters, can lead to situations where adults are required to validate and confirm youth perspectives. Slater (2015) suggests this is portrayed as adult generations who believe it is their responsibility to take control and enlighten children and young people.

Education configurations are likely to reflect these dominant conceptualizations of youth, which will lead to young people becoming occupied with learning about—and embodying—specific traits and abilities as they progress toward adulthood. For Kelly (2011, p. 48), youth participation is often understood as “becoming an adult, becoming a citizen, becoming independent, becoming autonomous, becoming mature and becoming responsible.” Minimal regard will be given to young disabled people's contributions if such understandings are not challenged. There is a danger that young disabled people will not be presented with opportunities to inform and influence the organization of inclusive education. They will not be supported to share their experiences, interests, and ideas associated with activism and politics, and contextualize these within the broader aims of producing inclusive knowledge exchanges.

A refreshed approach to participation is required if disability youth activism is to inform and influence inclusive education. Cahill and Dadvand (2018) propose a framework that, simultaneously, disrupts existing expectations of young people's contributions and encourages new forms of active engagement. Their approach is to understand youth perspectives as an assemblage of interactive ideas and activities, which come together to form new possibilities to experiment with and pathways to follow. Youth, here, is understood as a manifestation of experiences located within space, time, and relations (Dadvand, 2018). It denotes the circumstances, relationships, and situations encountered by young people, which are rooted within the current political, economic, social, and cultural arrangements. It is essential that young disabled people are recognized as active, and valued, participators in the organization of education. Their contributions—in the present—are significant for reconfiguring spaces and promoting inclusive remedies to current injustices (White et al., 2017).

Disrupting contemporary conceptualizations of youth is essential for valuing—and promoting—young disabled people's contributions; however, an additional issue needs addressing. Curricula need to be organized to encourage, and facilitate, discussions about realizing inclusive societies. This, in turn, will impact the organization of educational institutions. Young people's commentaries on the future tend to be framed as individualized narratives, focused primarily on personal achievement and success (Facer, 2011). It is argued this is due to perceptions of the future being colonized by neoliberal economic events, corporate agendas, and consumerist ideals. Aspirations become fixated on individual competitiveness (Naidoo, 2018), self-sufficiency (Giroux, 2009), and a disregard for collective progression (Peters and Marshall, 1996). This results in a precarious situation. On the one hand, exploration of futures is tolerated within pedagogy if aligned with notions of individual gain and personal success. Alternatively, little emphasis is placed on building knowledge exchanges and learning environments to consider possible, and preferable, inclusive societies for all.

Hicks (2002) argues for educational establishments to present safe, and accessible, spaces for learners to consider the plurality of futures available for exploration. This provides a platform for students to develop analytical skills and criticality, both useful for reflecting on potential futures (Facer, 2013). Learners should be supported to critique the present arrangements that constitute the social world, and determine the conditions necessary for realizing inclusive communities. This means questioning the attempts made by those who “know” the future then to consider the possibilities that may come to exist in a social world where the future is unknown—or open to alternative possibilities. Attention should be paid to the way or ways in which the emergence of ideas about the future will initiate action in the present.

While people are unable to know the future, this allows for creative opportunities to consider what may be possible (Miller, 2011). Disability activism and politics can serve to disrupt the tendency to individualize future narratives, instead shifting the focus to collective pursuits for progressive social change. Pedagogical practice, which absorbs the ideas of disability youth activism, positions learners at the forefront of generating ideas for transforming society. Their role is to understand contemporary and potential crises, with consideration given to the various responses available to address such catastrophes (Hicks, 2004; Inayatullah, 2008). A focus on the future also prompts learners to proactively consider how the social world ought to be organized (Masini, 2011). Learners can use future orientated curricula to propose how people should be supported to live, thrive, and participate in their communities. It can allow for new visions of welfare, security, and assistance.

In addition, it is rejection of the fatalistic predictions often attributed to discourses about the future (Davidson, 2019). Discussions about the future should not be primarily focused on responding to immediate concerns and reacting to the ongoing injustices encountered by marginalized communities. Notwithstanding the importance of addressing immediate injustices, the risk is that considerations of the future become fixated on short-term, crisis driven agendas. Inclusive education can become a space to introduce contingency within the present and destabilize the preferred narratives and trajectories established by existing influential forces. To achieve this, dedicated spaces within education are required to facilitate disability activism participation.

There is a need to establish spaces for disability activism to flourish within educational practice. Participation within activism remains patchy for young people from marginalized communities, particularly at the school level (Andrews, 2009). This points

to a continued absence of learners who will be excluded from emancipatory politics, given the struggles of young disabled people to participate in disability activism (Griffiths, 2019). Educational environments need to establish accessible pathways for young disabled people to learn about disability politics, and to build connections between existing activist networks, disabled people's organizations, and learners. Inclusive education needs to provide opportunities for young disabled people to become activists, as well as encourage young disabled people to use activism to influence education arrangements. To achieve this, inclusive education must be organized around three facets.

Firstly, educators will need to be familiar with the complex demands, agendas, and strategies employed by disabled activists and social movements to realize disabled people's inclusion. This requires translating the historical, and contemporary, ideas and practices that comprise disability politics (Berghs et al., 2019; Soldatic and Johnson, 2020) and ensure it is accessible and stimulating to learners—many of which will be newcomers to disability activism. Much of the archived material about disability activism is retained within higher education establishments (Griffiths, 2019), which remains inaccessible to disabled people (Brown and Leigh, 2020). Learners will need to be supported to reflect on the existing collections and materials and consider how activist claims redefine their own lived experiences of disablement and social injustice (Moore and Pell, 2010).

Secondly, educators will need to support learners to create and experiment with disability activism claim-making. For Braa and Callero (2006), the use of classroom discussions is an effective method for assisting students to engage in activism and social movements. Young disabled students can participate in discursive exercises to determine the possibilities for progressing inclusion and accessibility. This is important for ensuring learners can relate to—and influence—concepts and issues relevant to their lives (McCabe, 2013). The dominant claims made within disability activism and social movement organizations need to be contested by students, either through discussions or by embarking on activist orientated projects and collaborations.

Finally, educators need to situate this exploration of disability youth activism within the parameters of building inclusive education spaces. Isin (2008, p. 38) refers to activists "engage[ing] in writing scripts and creating the scene" for social change. In a similar vein, young disabled activists, within educational environments, should localize their interest in disability youth activism to assert claims and positions pertinent to the realization of inclusive education. This may require measures such as establishing specific student-led governance structures (Dollinger and Vanderlelie, 2019), representation of young disabled activists within education policy making processes (Deutschmann, 2022), and an expansion of advocacy provision within education settings (Thomas et al., 2017). The strategies have the potential to increase young disabled people's participation in activist mobilizations, as well as cement disability activism's role in eradicating discretionary practices within education.

Conclusion

This chapter has interrogated two aspects of inclusive education. Firstly, it has explored conceptualizations of inclusive education and aligned it with contemporary accounts of inclusion by young disabled activists. Secondly, it has considered how the organization of education can be informed by disability youth activism and politics. The chapter argues that inclusive education must comprise of persistent interrogations as to what constitutes inclusive and accessible societies, which should be informed by the ideas and activities of existing young disabled activists and newcomers to disability activism. It is paramount that inclusive education critically engages with the resistance practices of young disabled people, across the spectrum of incremental and radical activity, to understand and celebrate how disabled people are reconfiguring the social world. This will introduce opportunities for all learners to consider and experiment with alternative futures—to introduce contingency, possibility, and change into the worlds we are to inhabit.

Inclusive education has the potential to be guided by principles of disability youth activism and achieve two specific outcomes. Firstly, it can provide a continuous space for exchanging knowledge on what constitutes—and should constitute—inclusion and how this is reflected in social world organization. Secondly, it can utilize disability and disabled people's experiences to interrogate how human variance and diversity is celebrated, disregarded, and oppressed across societies. Education and knowledge exchanges should be orientated toward supporting creativity, exploration, and experimentation of how communities can be organized and participation for all guaranteed.

References

- Adebayo, S., 2020. A critical review of school choice and egalitarian justice with special reference to the Philippines. *Camb. Educ. Res. J.* 7, 70–85.
- Aldred, P., David, M., 2007. *Get Real About Sex: The Politics and Practice of Sex Education*. Open University Press, London.
- Andrews, R., 2009. Civic engagement, ethnic heterogeneity and social capital in urban areas: evidence from England. *Urban Aff. Rev.* 44 (3), 428–440.
- Armstrong, A., Armstrong, D., Spandagou, I., 2010. *Inclusive Education: International Policy and Practice*. Sage, London.
- Ball, S., Grimaldi, E., 2021. Neoliberal Education and the Neoliberal Digital Classroom. *Learning, Media and Technology*, pp. 1–15.
- Beckett, A., 2013. Non-disabled children's ideas about disability and disabled people. *Br. J. Sociol. Educ.* 35 (6), 1–20.
- Beckett, A., 2015. Anti-oppressive pedagogy and disability: possibilities and challenges. *Scand. J. Disabil. Res.* 17 (1), 76–94.
- Berghs, M., Chataika, T., El-Lahib, Y., Dube, A. (Eds.), 2019. *Routledge Handbook of Disability Activism*. Routledge, Abingdon.
- Blais, A., Labbé-St-Vincent, S., 2011. Personality traits, political attitudes and the propensity to vote. *Eur. J. Polit. Res.* 50 (3), 395–417.
- Blanck, P., 2005. *Disability Rights*. Ashgate Publishing, Farnham.
- Braa, D., Callero, P., 2006. Critical pedagogy and classroom praxis. *Teach. Sociol.* 34 (4), 357–369.
- Brown, N., Leigh, J., 2020. *Ableism in Academia: Theorising Experiences of Disabilities and Chronic Illnesses in Higher Education*. UCL Press, London.

- Byrne, B., 2022. How inclusive is the right to inclusive education? An assessment of the UN convention on the rights of persons with disabilities' concluding observations. *Int. J. Incl. Educ.* 26 (3), 301–318.
- Cahill, H., Dadvand, B., 2018. Re-conceptualising youth participation: a framework to inform action. *Child. Youth Serv. Rev.* 95, 243–253.
- Dadvand, B., 2018. Civics and citizenship education in Australia: the importance of a social justice agenda. In: Peterson, A., Stahl, G., Soong, H. (Eds.), *The Palgrave Handbook of Citizenship and Education*. Palgrave Macmillan, London, pp. 435–447.
- Davidson, S., 2019. "Nothing is set in stone": incorporating futurology into college courses. *J. Eff. Teach. High. Educ.* 1 (2), 136–152.
- deMarrais, K., LeCompte, M., 1999. *The Way Schools Work: A Sociological Analysis of Education*, third ed. Longman, New York.
- Deutschmann, A., 2022. Education and politics: student activism for elite recruitment in Kenya. *J. Contemp. Afr. Stud.* <https://doi.org/10.1080/02589001.2022.2034761>.
- Dollinger, M., Vanderlelie, J., 2019. Developing and enacting student governance and leadership training in higher education. *Stud. Success* 10 (2), 59–64.
- Erevelles, N., 2000. Educating unruly bodies: critical pedagogy, disability studies, and the politics of schooling. *Educ. Theor.* 50 (1), 25–47.
- Facer, K., 2011. *Learning Futures: Education, Technology and Social Change*. Routledge, London.
- Facer, K., 2013. The problem of the future and the possibilities of the present in education research. *Int. J. Educ. Res.* 61 (0), 135–143.
- Feely, M., 2016. Disability studies after the ontological turn: a return to the material world and material bodies without a return to essentialism. *Disabil. Soc.* 31 (7), 863–883.
- Florian, L., Black-Hawkins, K., 2011. Exploring inclusive pedagogy. *Br. Educ. Res. J.* 37 (5), 813–828.
- Freire, P., 2000. *Pedagogy of the Oppressed*, 30th anniversary ed. Continuum, New York.
- Giroux, H., 2009. *Youth in a Suspect Society: Democracy or Disposability?* Palgrave Macmillan, New York.
- Graeber, D., 2009. *Direct Action: An Ethnography*. AK Press, California.
- Griffiths, M., 2019. These days are ours: young disabled people's experiences of activism and participation in social movements. In: Berghs, M., Chataika, T., El-Lahib, Y., Dube, A. (Eds.), *Routledge Handbook of Disability Activism*. Routledge, Abingdon, pp. 57–70.
- Griffiths, M., 2022. Disabled youth participation within activism and social movement bases: an empirical investigation of the UK Disabled People's Movement. *Curr. Sociol.* <https://doi.org/10.1177/00113921221100579>.
- Harris, J., Roulstone, A., 2011. *Disability, Policy and Professional Practice*. SAGE, California.
- Hedegaard-Soerensen, L., Grumloese, S., 2020. Exclusion: the downside of neoliberal education policy. *Int. J. Incl. Educ.* 24 (6), 631–644.
- Hicks, D., 2002. *Lessons for the Future: The Missing Dimension in Education*. Routledge Falmer, New York.
- Hicks, D., 2004. Teaching for tomorrow: how can futures studies contribute to peace education? *J. Peace Educ.* 1 (2), 165–178.
- Hirschmann, N., 2016. Disability rights, social rights, and freedom. *J. Int. Polit. Theor.* 12 (1), 42–57.
- Inayatullah, S., 2008. Six pillars: futures thinking for transforming. *Foresight* 10 (1), 4–21.
- Isin, E., 2008. Theorising acts of citizenship. In: Isin, E., Nielsen, G. (Eds.), *Acts of Citizenship*. Palgrave Macmillan, London, pp. 15–43.
- Jenks, C., 2005. *Childhood*. Routledge, London.
- Jerome, L., Lalor, J., 2020. Revisiting subject knowledge in citizenship education: understanding power and agency. *Educ. Citizen Soc. Justice* 15 (2), 104–118.
- Kearl, B., 2019. Special education as neoliberal property: the racecraft, biopolitics, and immunization of disability. *Educ. Stud.* 55 (4), 473–488.
- Kelly, P., 2011. An untimely future for youth studies? *Youth Stud. Aust.* 30 (3), 47–53.
- Kilinc, S., 2019. "Who will fit in with whom?" Inclusive education struggles for students with dis/abilities. *Int. J. Incl. Educ.* 23 (12), 1296–1314.
- Kumashiro, K., 2002. *Troubling Education*. Routledge Falmer, New York.
- Lawson, A., Beckett, A., 2021. The social and human rights models of disability: towards a complementarity thesis. *Int. J. Hum. Right.* 25 (2), 348–379.
- Liasidou, A., 2015. *Inclusive Education and the Issue of Change: Theory, Policy and Pedagogy*. Palgrave Macmillan, London.
- Lorde, A., 1990. Age, race, class, and sex: women redefining difference. In: Ferguson, R., Gever, M., Minh-ha, T., West, C. (Eds.), *Out There: Marginalization and Contemporary Cultures*. New Museum of Contemporary Art, New York, pp. 281–287.
- Malaguzzi, L., 1993. For an education based on relationships. *Young Child.* 49 (1), 9–12.
- Martin, N., 2012. Disability identity—disability pride. *Perspectives* 16 (1), 14–18.
- Masini, E., 2011. How to teach futures studies: some experiences. *J. Futures Stud.* 15 (4), 111–120.
- McCabe, J., 2013. Making theory relevant: the gender attitude and belief inventory. *Teach. Sociol.* 41 (3), 282–293.
- Merrill-Glover, K., Edwards, K., 2015. Making it work: widening participation in higher education for adults in employment. *Widening Partic. Lifelong Learn.* 17 (2), 96–107.
- Miles, S., Singal, N., 2010. The Education for all and inclusive education debate: conflict, contradiction or opportunity? *Int. J. Incl. Educ.* 14 (1), 1–15.
- Miller, R., 2011. Being without existing: the futures community at a turning point? A comment on Jay Ogilvy's "facing the fold". *Foresight* 13 (4), 24–34.
- Molbaek, M., 2018. Inclusive teaching strategies—dimensions and agendas. *Int. J. Incl. Educ.* 22 (10), 1048–1061.
- Mølsted, C., Pettersson, D., Forsberg, E., 2017. A game of thrones: organising and legitimising knowledge through PISA research. *Eur. Educ. Res. J.* 16 (6), 869–884.
- Moore, S., Pell, S., 2010. Autonomous archives. *Int. J. Herit. Stud.* 16 (4–5), 255–268.
- Motta, S., 2013. Teaching global and social justice as transgressive spaces of possibility. *Antipode* 45 (1), 80–100.
- Müller-Bachmann, E., Chorvát, I., Mefalopoulos, A., 2022. Heading for a better world: micropolitical activism of young people seeking social change. *J. Youth Stud.* <https://doi.org/10.1080/13676261.2022.2053669>.
- Naidoo, R., 2018. The competition fetish in higher education: Shamans, mind snares and consequences. *Eur. Educ. Res. J.* 17 (5), 605–620.
- Obour, M., 2020. *Disabled Futures: A Framework for Radical Inclusion*. Temple University Press, Philadelphia.
- Oliver, M., Barnes, C., 2010. Disability studies, disabled people and the struggle for inclusion. *Br. J. Sociol. Educ.* 31 (5), 547–560.
- Oliver, M., Barnes, C., 2012. *The New Politics of Disablement*. Palgrave, London.
- Owens, J., 2015. Exploring the critiques of the social model of disability: the transformative possibility of Arendt's notion of power. *Sociol. Health Illness* 37 (3), 385–403.
- Peruzzo, F., 2020. The Model of Becoming Aware: disabled subjectivities, policy enactment and new exclusions in higher education. *J. Educ. Pol.* <https://doi.org/10.1080/02680939.2020.1856415>.
- Peters, M., Marshall, J., 1996. *Individualism and Community: Education and Social Policy in the Postmodern Condition*. Falmer Press, London.
- Roda, A., 2017. "More [Time] is better or less is more?" Neoliberal influences on teaching and learning time. *J. Educ. Pol.* 32 (3), 303–321.
- Schmitt, R., 2003. *Alienation and Freedom*. Worldview Press, Oxford.
- Slater, J., 2015. *Youth and Disability: A Challenge to Mr Reasonable*. Routledge, London.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Routledge, Abingdon.
- Slee, R., 2018. *Inclusive Education Isn't Dead, It Just Smells Funny*. Routledge, Abingdon.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922.
- Slee, R., Allan, J., 2001. Excluding the included: a reconsideration of inclusive education. *Int. Stud. Sociol. Educ.* 11 (2), 173–192.
- Soldatic, K., Johnson, K. (Eds.), 2020. *Global Perspectives on Disability Active and Advocacy: Our Way*. Routledge, Abingdon.
- Spring, J., 2013. *American Education*. McGraw Hill, New York.
- Swain, J., French, S., 2000. Towards an affirmation model of disability. *Disabil. Soc.* 15 (4), 569–582.
- Thomas, N., Crowley, A., Moxon, D., Ridley, J., Street, C., Joshi, P., 2017. Independent advocacy for children and young people: developing an outcomes framework. *Child. Soc.* 31 (5), 365–377.
- Tilly, C., Tarrow, S., 2015. *Contentious Politics*. Oxford University Press, Oxford.

- Todd, Z., Griffiths, M., Agustsdottir, E., Kichashki, P., Maraz, D., 2012. Young People With Disabilities as Future Leaders of the Independent Living Movement. European Network on Independent Living, Brussels.
- Union of the Physically Impaired Against Segregation [UPIAS], 1975. Fundamental Principles of Disability. UPIAS, London.
- Watermeyer, B., Görgens, T., 2014. Disability and internalized oppression. In: David, E. (Ed.), *Internalized Oppression: The Psychology of Marginalized Groups*. Springer Publishing Company, New York, pp. 253–280.
- White, R., Wyn, J., Robards, B., 2017. *Youth and Society*, fourth ed. Oxford University Press, Oxford.
- Withers, A., Ben-Moshe, L., Brown, L., Erickson, L., da Silva Gorman, R., Lewis, T., McLeod, L., Mingus, M., 2019. Radical disability politics. In: Kinna, R., Gordon, U. (Eds.), *Routledge Handbook of Radical Politics*. Routledge, New York, pp. 178–193.
- Wolbring, G., 2008. The politics of ableism. *Development* 51, 252–258.
- Zion, S., Blanchett, W., 2017. On the purpose of schooling: inviting in, or locking out? In: Hughes, M., Talbott, E. (Eds.), *The Wiley Handbook of Diversity in Special Education*. Wiley, New Jersey, pp. 69–85.

Attitudes toward inclusion

Tim Corcoran, School of Education, Deakin University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	328
Attitudes about disability	328
Psychology of attitudes	329
Attitudes as life flows	330
Dissolving inclusion	331
Conclusion	332
References	333

Introduction

There is a lengthy history and longstanding fascination with attitudes of presently non-disabled people toward those living with disabilities (Antonak and Livneh, 1988). Traditionally this manifests in attempts to reify, identify and measure the psychological construct. No cohort is left unchecked as the research moves through groups to capture these attitudes—from school principals to teachers generally, to history teachers, to mathematics teachers, to science teachers, and so on. The work extends to teaching assistants, student teachers and school ancillary staff and moves across sectors from early childhood to primary and secondary school, vocational and higher education. Comparisons are also made between education workers, health workers, criminal justice workers and beyond. It is important to acknowledge the sheer volume of this research makes it difficult to resist generalization. This chapter is not intended to be a literature review yet it is important to acknowledge volume and activity for research into attitudes are regularly set free from contingency, devoid of context and reliant on dominant forms of reductionist psychology.

There is inference in this work that once the character and the carriers of so-called negative attitudes toward disabled people are identified, these thoughts can be mitigated through adjustments to teacher preparation and professional development. Also common to the cannon is scant recognition of ontological and epistemological traditions generative of two sustaining features in this work: psychological reductionism and ableism. Researchers generally eschew acknowledgment of their ontological footings and more will be said to this point below. Preempting the forthcoming argument, descriptors such as disabled and non-disabled are struck as stable categories. The effect being at least forty years of research espousing similarly essentialized ontology, research design and repetitive recommendation that these things called attitudes exist, are deleterious toward people living with disability and therefore must be changed.

This chapter thinks otherwise about attitudes and how they might be known. To begin, how might we ask researchers to become interlocutors within the psychosocial, political and material world they inhabit? As an initial step, this is a call for analyses regarding commencement; with apriori assumptions and orientations that too predictably lead research to foregone conclusions. The following discussion begins with consideration for how we categorize disability in educational contexts. I then turn to consider how attitudes have been operationalized in research after-the-fact. While not denying the impact of prejudices enacted in and sustained by ableist attitudes, the chapter points to a research opportunity which, if taken seriously, heralds the outlandish—a call for dissolving inclusion.

Attitudes about disability

In this section my intent is to show how researchers, when orienting to the world, enact different ways of knowing/being (Corcoran, 2017a). Much of the research that aims to identify attitudes of non-disabled people to disabled students can be connected to conservative and neo-conservative traditions. Special education, whether in its more established form (e.g., Fuchs and Fuchs, 1994; Kauffman and Hallahan, 1995; Kauffman and Sasso, 2006) or neo-conservative manifestation (Farrell, 2006; Hornby, 2014), proceeds from what scholars in the field of Disability Studies call a medical model of impairment (Goodley, 2017; Oliver, 1990, 2009; Thomas, 2004, 2007). There, disability is understood as a representation of individual pathological defectiveness that presents for diagnosis, intervention, treatment and management. Consequently, there is little question disability exists and it is of the individual—an affliction that can evoke fear, pity, charity, or professional interventions to normalize the disabled individual.

Disability Studies however proceed from divergent ontological assumptions. As socially constituted phenomenon, disability is engaged as ubiquitous, dynamic and understood as the consequence of social (i.e., institutional, political and cultural) responses to people's impairments, illnesses and/or differences. Accordingly, disability is not of the pathology of individual people but more precisely an expression of social or political mores sustaining the oppression of those described as disabled. As a developing field within the social sciences, Disability Studies is being shaped by researchers and activists from a range of theoretical traditions including critical theory, post structuralism, feminist studies, critical race theory, queer studies and crip studies.

The focus for the education scholar committed to inclusion, proceeding from divergent points of engagement, might narrow to questions about how students are brought into being as different in and through entangled life flows.

When you've seen beyond yourself

then you may find peace of mind is waiting there

And the time will come when you see we're all one

And life flows on within you and without you

Harrison (1967).

In 2017 the world celebrated the 50th anniversary of the release of The Beatles *Sgt. Pepper's Lonely Hearts Club Band*. I need not add more to the volume already written about the band and their landmark recording. Suffice to say the relevance of George Harrison's lyric to this discussion is simple: it speaks to a kind of onto-epistemological union often discounted and derided as unchecked hippy-inspired spirituality. It makes complete sense for people philosophically and materially invested in ideological individualism to say as much. The advent of modern psychology has dutifully helped to perpetuate self-focused proclamations on what is thought to exist within, that being our psychology and its constitutive elements (e.g., psychopathologies or disorders). How might we then understand life otherwise when the majority of talk about human being is premised on reductionist principles?

Below I offer an alternative by conceptually reconsidering attitudes as psychosocial phenomena or "life flows" to borrow Harrison's phrase. I believe, in reconsidering life's enactments this way, more ethical and purposeful opportunities might be cast in educational research and practice. Along the way, questions involving research relations, production and enactment shuffle forward, anticipating certain ways of life (cf. Wittgenstein, 1953). There is indeed a different set of ethical tensions to be considered when surveying or interviewing presently non-disabled people about attitudes regarding students living with disability. As well as ethical tensions revolving around disability categorizations and subsequent tendencies to frame students as a problem for the school rather than the school as a problem for the student, these things called attitudes remain stubbornly affixed and seemingly difficult to remove from the heads of individuals.

Psychology of attitudes

As a first move, retracing a course through history can often be insightful. When it comes to historical accounts of psychology, there are few more informative than those provided by Kurt Danziger (1990, 1997). In *Naming the Mind*, Danziger supplies erudite coverage of many of psychology's key concepts including intelligence, motivation, behavior and attitudes. Regarding attitudes, it is unsurprising that understandings of the concept have changed over the years. Psychology's infancy as a distinct scientific discipline saw researchers take up interest with psychological phenomena as observable via instances of motor function (i.e., muscle movement). For example, it has been suggested a person's facial expression may be considered the overt expression of underlying emotion. Still today psychologists research the evolutionary and universal nature of facial expression and its correlates to emotion (e.g., Ekman, 1973, 2003). If researchers could infer an emotional state from physical observation, what other kind of psychological phenomena might be addressed?

The shift to researching attitudes was initiated over a century ago. The dominance of experimentalism in the new social sciences encouraged work that empirically connected observable actions with the internal psychology of the individual. Theory had already started to explain the potential existence of attitudes (Titchener, 1909). It was suggested if a means to measuring attitudes could be developed, the existence of mental constructs would be confirmed. To establish its scientific bonafides, psychologists would need to go further than simply canvassing people's opinions. Political polling and advertising market research was already sourcing opinion from the general populace by the early 1900s. The social scientific challenge was to create a method for exposing attitudes via sophisticated measurement and in so doing, a logical cause of behavior would be made.

Any social science researcher worth their salt has heard the name Likert. Renis Likert was a student of Louis Thurstone, an American psychologist acclaimed with pioneering key developments in the field of psychometrics. It was Thurstone who proposed the law of comparative judgment, a central principle used in measuring psychological constructs like attitudes. Thurstone's work allowed researchers to scale measurement via a method of pairwise comparison. His research helped transfer work in the area of psychophysical judgment (e.g., comparing material objects on a physical dimension like weight) to judgments sought on social positioning (e.g., being pro- or anti-a particular topic). In 1932, Likert coopted Thurstone's approach to provide what those familiar with questionnaire or survey research today know as a cornerstone of such method, the Likert scale. This is commonly a 5 or 7 point

ordinal scale respondents use to judge where they are situated in relation to a particular statement. When a series of statements or items sufficiently correlate, the scale is deemed significant.

An issue often not addressed in understanding measurements such as attitude scaling is a slight of hand employed in its early development (Danziger, 1997). What had been an intra-individual method involving psychophysical judgment (i.e., one person making a series of judgments) was subsequently and unproblematically applied as an inter-individual method (i.e., pooling judgments from more than one person). Such a move neglects and negates the capacity for individuals to develop different meanings from particular words or statements. As briefly discussed above and in our present context, who is to say we agree on what the term disability can mean? Also, as an ordinal scale, the difference between a response of “often” and a response of “sometimes” is not measurably exact. Subsequently, what any response means to one individual is not necessarily the same for all.

Despite such conjecture, attitude research accelerated from the Second World War on. The technology found traction with government services, particularly in the US, proliferated by social scientists serving in the US War Department’s Information and Education Division. Psychologists “held that attitudes actually exist as the mental property of individuals, not groups [and] claimed that their focus was on external, observable behavior; yet attitudes clearly are concerned with internal characteristics. This contradiction precipitated years of debate about the link between attitudes and behavior” (Walsh et al., 2014, p. 290). Such controversy persists today as policy makers and researchers struggle to change behavior via focus on attitudes (e.g., Ewert et al., 2021).

This brief incursion into the history of attitude research brings to the fore two central concerns. The first is to do with the theoretical treatment of attitudes as causal psychological phenomena. The second points to methodological issues created when language is treated as a neutral conduit. Together, these raise distressing problems that cannot be overcome by simple reconfiguration of existing cycles of knowing/being. And so, it is to an onto-epistemological envisaging that we now turn. In the following section discussion invites a different way of considering attitudes as psychosocial phenomena.

Attitudes as life flows

It is this kind of knowledge—of the provisions and resources we make available to ourselves for the realisation of our different possible next forms of social behavior—that is the special kind of knowledge embodied in the world of civil society

Shotter (1993, p. 3).

It is easy to get caught up with/in words. Ontology. Epistemology. Onto-epistemology. More and more, terms like these seem to make their presence felt in contemporary social science discussion. Even if you experience their use as frustrating, this is, I believe, a good thing. It is, as Shotter intimates above, necessary for us to examine both ways we go about knowing (epistemology) and existing forms of life (ontology) supported in/through such work. Rather than being made to accept life in/on certain terms, such discussion opens ethical debate around how life could and should be. The discomfort experienced in grappling with the lexicon slowly gives way to a kind of freedom wherein language has the potential to become more servant than master. In this section I dissect further the provisions and resources made available in attitudes research to clear a way for what could possibly come next.

The kind of attitude research discussed here predominantly belongs to a monologic view of life. It is monologic in that it generally stories disability through two mutually supportive philosophical perspectives, realism and positivism. We need not go into extended discussion regarding these. Suffice to say, the kinds of provisions and resources made available in such work do more to disable possibilities than enable change. Let us resume the discussion broached in the previous section regarding what we believe to be ontological and epistemological oddities promoted by the majority of attitude research to explain this further.

Ontologically, attitude research promotes, unsurprisingly, psychologism (Sugarman, 2017). This comes about when psychological perspectives are deemed fundamental to the account rendered. Attitude researchers do this by accepting and promoting the existence of individually possessed mental constructs verified through measurement. My main apprehension with attitudes explained as mental constructs is the division created between a corporeally situated psychological interior of the individual and an existent exterior where the rest of life takes place; in homes, in schools, in communities and so on. Seemingly, if people could just change the way they think about disability or education, what goes on in schools and in classrooms should follow suit. It could be suggested that today fewer people are willing to accept this kind of psychological dualism and yet, such ontological reasoning persists (e.g., Krischler and Pit-ten Cate, 2019). When the prevailing contemporary ideology is that of individualism, it is fair to ask how could it be any other way?

Subsequently, this reductive view of existence is supported by a way of knowing the world in which certainty is both promised and believed. The gold standard, particularly for those claiming to be scientists, and more fervently for policy-makers these days, is evidence supposedly guaranteed by experimentalism. This kind of proof, derived from statistical inference, aims to generalize what are often called research “findings.” A simple way to check what is being presented is to ask whether the outputs of research are claimed to be found or made (Corcoran, 2017b; Shotter, 1993). With the former, there is a legitimated separation of epistemology from ontology or knowing from being. Put directly, what is presumed to exist (e.g., attitudes) are waiting in the world to be found. If we are yet to find them, this does not necessarily mean we should doubt their existence. Because ontology can be separated from epistemology, verification of a certain reality may simply be pending the appropriate method for knowing.

But what if we were to consider attitudes differently? Instead of being things people possess inside their heads, what if attitudes were understood as psychosocial matter existing like any other participant in the action taking place? To do this, we would need to reconsider a number of precepts; for example, how we understand and attribute agency, varieties of worldly relationship, and so on. Simply invoking a more relational view of individuals engaging with the world won't do. Versions of constructivism have been offered in this regard, but these remain theoretically tied to the kind of dualism that marks a distinct psychological interior from the material external world (e.g., Kelly, 1991; Seligman, 2011). Beyond behavioral and cognitive-behavioral theoretical dominance in special education, cognitive-behavioral and constructivist models continue to overshadow Disability Studies. Critical Disability Studies has been quick to recognize the shortcomings of dominant psychological theory with contributions in this area examining opportunities where alternate kinds of psychological theory (e.g., social constructionism) and disability studies might be mutually beneficial (Billington and Goodley, 2020; Rapley, 2004).

To move psychology out of the heads of people and into an active world, what is needed is a form of transdisciplinarity capable of speaking to psychological, biological, social, economic and political concerns. A form of psychosocial understanding that resists separating people from their world, or ontology and epistemology. Put another way, a psychosocial understanding that might account for life as it is being made rather than one that pre-exists engagement (cf. Brown and Stenner, 2009). In searching for collaborators, we can reach back thirty years to Potter and Wetherell's primer in discursive psychology, *Discourse and Social Psychology: Beyond attitudes and behavior*. Against the grain of contemporary mainstream social psychological research, this early prototype for discursive social psychology sought to explore two main concerns: (i) how language is constructed in local contexts, and (ii) what consequences might ensue from particular types of construction. Discursive psychology is not preoccupied with establishing the existence of mental properties, but instead focuses on "the organization of discourse in relation to function and context" (Potter and Wetherell, 1987, p. 54). More recent discursive psychological accounts acknowledge the increasing complexity involved in social psychological theorization of attitudes (e.g., social cognitive theory), but acknowledge these as retaining uncomplicated accounts of representational language use and separation of the individual from their surroundings (McVittie and McKinlay, 2017).

While discursive social psychology provides valuable recourse to examining psychological life, this work should not be taken as homogenous nor without reservation (Bozatzis, 2014). In particular, an enduring treatment of epistemic and ontological concerns continues to hinder potential theoretical developments (Wetherell, 2007). For example, in a recently published piece on topic, Potter et al. (2020) persist with invocation of a "natural" world within which human beings develop "assessments" or attitudes. As flagged above, this chapter is interested in pursuing Barad's (2007) notion of onto-epistemology where potential dilemmas regarding ontology and epistemology are ethically inseparable. For us, as researchers, "... an invitation to live justly is written into the very matter of being. How to respond to that invitation is as much a question about the nature of response and responsibility as it is about the nature of matter" (Barad, 2007, XI). Recognition of an existing yet insufficiently theorized natural world continues to be an ethical dilemma I encounter with this type of research (cf. Corcoran, 2009).

In the space remaining I want to move on to suggest several proposals for how attitudes research in education might take place. It is important to emphasize these propositions should not be taken as "How To" instructions for the work ahead. To even hint at that would denigrate all that I have attempted to share in this discussion. To do so would return us to the kind of monologic telling-it-like-it-is approach that so often fails to open possibilities, encasing our exchanges in closed probabilities instead. I now join in chorus with one of my cultural heroes, who when describing his craft reflected: "My business is SHOW business and that is the business of SHOWING ... not TELLING. You don't TELL people anything, you SHOW them, and let them decide" (Springsteen, 2016, p. 277; upper case in original). In this vein, allow me to show how we can do research that might satisfy our ambition if inclusive education is central to that prospect.

Dissolving inclusion

There is a fundamental distinction here regarding how we approach and engage disability and attitudes, as well as knowledge of other kinds of things for that matter. Above, I discussed how Disability Studies has challenged traditional representations of people living with impairment and how psychology generally purports to know these things called attitudes. In combination, the ever-growing literature theoretically refines the existence of attitudes toward disability as things already possessed by us and epistemologically proves their ontology through rigorous scientific method. But what if we were to reconsider attitudes not as already existing; that is, things waiting to be found like one's car keys or the television remote control, but as things made by us in our continuous engagements with one another and the world? In this regard, as researchers might we then refrain from allowing our inquiries to be determined by what is already known (e.g., categories in the DSM), and in so doing be enabled to notice how our own actions in engagement with our material world bring different things like attitudes concerning disability into existence. This is not to suggest we discount certain research traditions and what those have had to contribute to date. But what I am attempting to do here is call out inquiries which separate ontology and epistemology as particular kinds of after-the-fact accounting (Shotter, 2017), and in doing so, move to reorient education research and practice in the area toward other possibilities.

British psychologist John Shotter worried about the ways social scientists usually go about analyzing life. Or, to reiterate Barad's earlier statement, what "the nature of response and responsibility" means for doing research. Shotter's key concern held that most research disrespects the unfinalisability of living things, geared as it is to produce reified accounts of life already lived. In these kinds of after-the-fact tellings, decisions are reached methodically by retrospectively weighing up evidence to solve real-world problems. Problems are "broken down" or "pulled apart" so we might capture how they materialize to get a clearer and more accurate view of

Table 1 Options for attitude research.

	<i>After-the-fact</i>	<i>Before-the-fact</i>	
Retrospective	Attitudes are fixed	Attitudes are emergent	Prospective
Solving	Attitudes are faulty	Attitudes are entangled	Dissolving
Representation	Attitudes are found	Attitudes are made	Orientation
Reflection	Attitudes are facsimiles	Attitudes are equivocal	Diffraction

their constituent parts. The dominant mechanistic metaphor of our time tells us that if we isolate the problem, we will then be able to diagnose dysfunction and correct what needs to be corrected to make things work. This is how statistical significance, derived from responses to technologies like attitudes questionnaires, is intended to contribute to the goal of attaining equitable education in our schools.

Closer to the kind of onto-epistemological work I have been discussing here, Shotter proposed a form of before-the-fact engagement to assist us in responsibly meeting the world in and through research. As just mentioned, as researchers we have been trained to solve problems. Not all problems are the same of course, nor are their solutions. Nor is might what we see as problematic understood by others similarly across various situations. Already from what has been said we start to sense the pebble in the shoe of the researcher. If our intent is to remain part of a living world and speak in ways that potentially sustain possibilities for how we can go on together, our efforts may not be well spent producing after-the-fact evidence (e.g., isolating teacher's attitudes) in service of final solutions (e.g., attaining inclusion). Promises like this often offer false hope and any who have worked in government possibly know what I mean here. Perhaps, rather than exclusively solving problems using after-the-fact derived evidence, our purpose might just as importantly be directed at dissolving problems. Dissolved problems are diffused phenomena, irrevocably altered if and when extracted, reduced or withdrawn from the unique whole in which they exist. As **Table 1** shows, in considering how researchers might conserve the entangled nature of living worlds, our work should orient to dissolving problems as before-the-fact venturing.

This discussion regarding after- and before-the-fact venturing relies considerably on how we orient ourselves with/in the world. For example, earlier I made reference to research findings and makings. The point can be extended by current academic distinctions contrasting how reflection and diffraction are used in research ([Bozalek and Zembylas, 2017](#)). As learners we have been taught the value of reflective practice whereby we scrutinize our actions and observations looking to gain greater insight so as to inform what comes next. Immediately, this process can be spotted for its retrospectivity and after-the-fact way of working. Reflection can also be appraised metaphorically for the act is suggested to be a form of cognitive mirroring insusceptible to the potential affections involved in any process ([Rorty, 1979](#)). That a researcher can use language to transparently represent what is, for example, attitudes meaningfully derived from statements posed on questionnaires, is highly contentious for, as suggested above, it falsely presumes and unethically generates response homogeneity.

This then is seemingly one of the greatest challenges of our time. Gated communities describe the boundedness of physical areas created to securely separate a like-minded us from the (for whatever reason) unwanted them ([Bauman, 1997](#); [Sennett, 2018](#); [Slee, 2018](#)). Similarly, for as long as there have been universities, people have recognized the intellectual silos created between disciplines and traditions of argumentation ([Billig, 1987](#)). Around the globe today, commentators both academic and social worry about the influence rapid developments in technology are having on what information we are exposed to and what information we in turn expose ourselves to. Populist political leaders cash in on this zeitgeist, furthering exclusions and promoting divisions wherever and whenever possible, gating not only their communities, but potentially even their countries. Looking prospectively, orienting research practice to include diffractive engagements may help in dissolving contemporary problems by "carefully and attentively doing justice to a detailed reading of the intra-actions of different viewpoints and how they build upon or differ from each other to make new and creative visions" ([Bozalek and Zembylas, 2017](#), p. 118).

Conclusion

Much of the research that aims to interrogate attitudes of non-disabled people toward students with disabilities proceeds from accepting disability as a stable or fixed phenomenon that refers to the pathological composition of individuals. Accordingly this research has been linear in design, application and trajectory, moving as it does from specific cohorts of able-bodied people to tap and adjudicate their reactions to students with disabilities in general, or to students who are attached to particular diagnostic categories. Herein lies a largely unrecognized and unacknowledged onto-epistemological chasm. The very framing of the research question about attitudes toward students with disabilities is emblematic of the foundations of disablement and exclusion. It is an exquisite representation of the creation and nurturing of discrimination through particular forms and expressions of knowing the world, knowing people and knowing the disabled person—simply, knowing/being. The critique of this research genre undertaken above is not an academic—and academic is used frequently as a pejorative term of reference—exercise in polemics. It is offered to stimulate radically different discussions about knowing/being human ([Goodley, 2021](#)).

Accounting for existing conditions via after-the-fact practices like specifying attitudes no longer need be claimed as a primary means to effect change in education. Researching attitudes, and to a certain degree other kinds of psychosocial sensemaking

(including educational favorites like grit (Corcoran and Vassallo, 2021) or behaviour (Corcoran and Thomas, 2021), can instead be engaged as prospective activity where “to understand is to anticipate together, it is to make a cross-reference which, when acted upon, brings about a partaking in a common, inclusive, undertaking” (Dewey, 1929/1958, pp. 178–179). How does initiating before-the-fact orientations moves us closer to dissolving inclusion? It does so by calling upon our anticipations, that is where we decide to commence our work, to ethically address available onto-epistemological possibilities. What should be heralded are those starting points which configure our ways of knowing/being in certain terms at certain times and places and whether these create preferred ways to go on together. Should ableist understandings of students, for instance, set the coordinates for educational practice? And what, to conclude, would the significance then be of inclusive education were this not our point of departure?

References

- Antonak, R.F., Livneh, H., 1988. *The Measurement of Attitudes Toward People With Disabilities: Methods, Psychometrics and Scale*. Charles C. Thomas, Springfield.
- Barad, K., 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, London.
- Bauman, Z., 1997. *Postmodernity and Its Discontents*. Polity, Cambridge.
- Billig, M., 1987. *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge University Press, Cambridge.
- Billington, T., Goodley, D., 2020. The relief of critical educational psychology and the nomadism of critical disability studies: social constructionism in practice. In: McNamee, S., Gergen, M.M., Camargo-Borges, C., Rasera, E.F. (Eds.), *The Sage Handbook of Social Constructionist Practice*. Sage, London, pp. 372–379.
- Bozalek, V., Zembylas, M., 2017. Diffraction or reflection? Sketching the contours of two methodologies in educational research. *Int. J. Qual. Stud. Educ.* 30 (2), 111–127.
- Bozatzis, N., 2014. The discursive turn in social psychology: four nodal debates. In: Bozatzis, N., Dragonas, T. (Eds.), *The Discursive Turn in Social Psychology*. World Share Books, Chagrin Falls, pp. 25–50.
- Brown, S.D., Stenner, P., 2009. *Psychology Without Foundations: History, Philosophy and Psychological Theory*. Sage, London.
- Corcoran, T., Thomas, M.K.E., 2021. School-wide positive behavioural support as evidence-making interventions. *Res. Educ.* 111 (1), 108–125.
- Corcoran, T., Vassallo, S., 2021. Psychosocial justice: always more-than to consider. *Educ. Child Psychol.* 38 (1), 8–18.
- Corcoran, T., 2009. Second nature. *Br. J. Soc. Psychol.* 48 (2), 375–388.
- Corcoran, T., 2017a. Afterthought no longer. In: Plows, V., Whitburn, B. (Eds.), *Inclusive Education: Making Sense of Everyday Practice*. Sense, Rotterdam, pp. 249–256.
- Corcoran, T., 2017b. Ontological constructionism. In: Williams, A., Billington, T., Goodley, D., Corcoran, T. (Eds.), *Critical Educational Psychology*. Wiley Blackwell, London, pp. 26–33.
- Danziger, K., 1990. *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge University Press, Cambridge.
- Danziger, K., 1997. *Naming the Mind: How Psychology Found Its Language*. Sage, London.
- Dewey, J., 1929/1958. *Experience and Nature*. Dover, New York.
- Ekman, P. (Ed.), 1973. *Darwin and Facial Expression: A Century of Research in Review*. Academic Press, New York.
- Ekman, P., 2003. *Emotions Revealed*, second ed. Times Books, New York.
- Ewert, B., Loer, K., Thomann, E., 2021. Beyond nudge: advancing the state-of-the-art of behavioural public policy and administration. *Pol. Polit.* 49 (1), 3–23.
- Farrell, M., 2006. *Celebrating the Special School*. David Fulton, London.
- Fuchs, D., Fuchs, L., 1994. Inclusive schools movement and the radicalisation of special education reform. *Except. Child.* 60 (4), 294–309.
- Goodley, D., 2017. *Disability Studies: An Interdisciplinary Introduction*, second ed. Sage, London.
- Goodley, D., 2021. *Disability and Other Human Questions*. Emerald, Bingley.
- Harrison, G., 1967. *Within You, Without You* [Liner Notes]. In: On Sgt. Pepper's Lonely Hearts Club Band [LP]. Parlophone, London.
- Hornby, G., 2014. *Inclusive Special Education: Evidence-Based Practices for Children With Special Needs and Disabilities*. Springer, New York.
- Kauffman, J.M., Hallahan, D.P., 1995. *The Illusion of Full Inclusion: A Comprehensive Critique of a Current Special Education Bandwagon*. Pro-Ed, Austin.
- Kauffman, J.M., Sasso, G., 2006. Toward ending cultural and cognitive relativism in special education. *Exceptionality* 14 (2), 65–90.
- Kelly, G.A., 1991. *The Psychology of Personal Constructs*. In: *Theory and Personality*, vol. 1. Routledge, London (Original work published 1955).
- Krischler, M., Pit-ten Cate, I.M., 2019. Pre- and in-service teachers' attitudes toward students with learning difficulties and challenging behaviour. *Front. Psychol.* 10. <https://doi.org/10.3389/fpsyg.2019.00327>.
- McVittie, C., McKinlay, A., 2017. Attitudes and attributions. In: Gough, B. (Ed.), *The Palgrave Handbook of Critical Social Psychology*. Palgrave, London, pp. 269–290.
- Oliver, M., 1990. *The Politics of Disablement*. Macmillan Education, London.
- Oliver, M., 2009. *Understanding Disability: From Theory to Practice*. Palgrave Macmillan, Basingstoke.
- Potter, J., Wetherell, M., 1987. *Discourse and Social Psychology: Beyond Attitudes and Behaviour*. Sage, London.
- Potter, J., Hepburn, A., Edwards, D., 2020. Rethinking attitudes and social psychology—issues of function, order, and combination in subject-side and object-side assessments in natural settings. *Qual. Res. Psychol.* <https://doi.org/10.1080/14780887.2020.1725952>.
- Rapley, M., 2004. *The Social Construction of Intellectual Disability*. Cambridge University Press, Cambridge.
- Rorty, R., 1979. *Philosophy and the Mirror of Nature*. Princeton University Press, Princeton.
- Seligman, M.E.P., 2011. *Flourish: A Visionary New Understanding of Happiness and Well-Being*. Free Press, New York.
- Sennett, R., 2018. *Building and Dwelling: Ethics for the City*. Penguin Books, London.
- Shotter, J., 1993. *Cultural Politics of Everyday Life: Social Constructionism, Rhetoric and Knowing of the Third Kind*. Open University Press, Buckingham.
- Shotter, J., 2017. Persons as dialogical-hermeneutical-relational beings—new circumstances “call out” new responses from us. *New Ideas Psychol.* 44, 34–40.
- Slee, R., 2018. *Inclusive Education Isn't Dead, It Just Smells Funny*. Routledge, Abingdon.
- Springsteen, B., 2016. *Born to Run*. Simon and Schuster, New York.
- Sugarman, J., 2017. Psychologism as a style of reasoning and the study of persons. *New Ideas Psychol.* 44 (1), 21–27.
- Thomas, C., 2004. How is disability understood? An examination of sociological approaches. *Disabil. Soc.* 19 (6), 569–583.
- Thomas, C., 2007. *Sociologies of Disability, “Impairment”, and Chronic Illness: Ideas in Disability Studies and Medical Sociology*. Palgrave, London.
- Titchener, E.B., 1909. *Lectures on the Experimental Psychology of the Thought Processes*. Macmillan, New York.
- Walsh, R.T.G., Teo, T., Baydala, A., 2014. *A Critical History and Philosophy of Psychology: Diversity of Context, Thought and Practice*. Cambridge University Press, Cambridge.
- Wetherell, M., 2007. A step too far: discursive psychology, linguistic ethnography and questions of identity. *J. Sociolinguistics* 11 (5), 661–681.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Blackwell, Oxford.

Considerations for implementing Universal Design for Learning: toward anti-oppressive pedagogies

Kathy Cologon, Macquarie University, Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	334
Universal Design for Learning	335
UDL: a social model approach	336
Progressing forward with UDL	338
References	339

Introduction

As the name suggests, Universal Design for Learning (UDL) is about creating access for everyone. UDL has been developed from the broader concept of Universal Design. In the United Nations Convention on the Rights of Persons with Disability (CRPD), Universal Design is defined as “the design of products, environments, programs and services to be useable by all people, to the greatest extent possible, without the need for adaptation or specialized design” (United Nations, 2006, Article 2). In addressing the topic of implementing UDL, at its core, this chapter is about accessibility. Accessibility is a precondition for upholding human rights since human rights cannot be upheld without accessibility across every aspect of human society. The essential nature of accessibility is outlined in the Universal Declaration of Human Rights (United Nations, 1948) and the International Covenant on Civil and Political Rights (United Nations, 1966). Consequently, accessibility is a core principle underpinning the CRPD (United Nations, 2006). This principle is explored extensively in General Comment 2 on the CRPD, which unpacks Article 9 on accessibility (United Nations, 2014). It is also directly applied to education in General Comment 4 on the CRPD, which addresses Article 24 on inclusive education (United Nations, 2016). On this basis, UDL becomes a fundamental strategy for upholding human rights.

Universal Design is an approach that is widely beneficial as it facilitates the creation of a world in which all people are considered, a world that is accessible to all people, and a world that will not always require constant “retro-fitting” in order to uphold basic rights and address the basic needs of many individuals. It is important to note, though, that Universal Design does not mean creating a “one-size-fits-all” approach in which we all do the same thing, in the same way, or at the same time. Instead it involves examining who is considered in the design of our world, who is missing in these considerations, why, and how this can be addressed in order to make design universal—as in, for all of us. As is further explicated in the CRPD, “‘Universal design’ shall not exclude assistive devices for particular groups of persons with disabilities where this is needed” (United Nations, 2006, Article 2). While such assistive devices may only be used by some of us, or some of the time, the incorporation of these into Universal Design is essential in ensuring that the design is genuinely “universal”.

Originating in the field of architecture, Universal Design examples include: curb cuts; automatic doors; ramps; lifts with visual, audio and braille information; wide doorways; closed captioning; and in-built screen readers (CAST, 2018). Thinking about the concept of Universal Design always brings to mind for me an old cartoon by Giangreco and Ruelle (2002). Giangreco and Ruelle’s cartoon depicts a school that has been snowed in. The entrance to the school is blocked as the stairs and ramp are covered in snow. There are a group of students waiting to get into the school. One student, who uses a wheelchair, asks the adult who is shoveling snow off the stairs, “Could you please shovel the ramp?” The adult responds, “All these other kids are waiting to use the stairs. When I get through shoveling them off then I will clear the ramp for you”. The student replies, “But if you shovel the ramp, we can all get in!”.

To me, this cartoon illustrates both the mistaken association of Universal Design as being only for some people (in this case, the student who uses the wheelchair) and, in contrast, the genuine and logical benefits of creating accessibility for everyone. The cartoon also draws our minds to the attitudinal aspects of the implementation of Universal Design—beyond the architectural considerations. The school depicted *does* have a ramp, and yet the ramp is viewed as secondary to the stairs at the entrance to the school.

It is noteworthy, however, that the importance of the existence of the ramp is not to be underestimated, as those of us who have attempted navigating the world around us with the use of mobility aids can attest to. Although the cartoon may be coming up to 20 years old, many such ramps (and other avenues for accessibility) are brand new or not yet in existence.

This example, while specific, can be extrapolated endlessly and the benefits to everyone of accessibility are timeless. Consider pulling a suitcase, pushing a pram, or a trolley and the universal aspect of this design becomes readily apparent. However, the reality that we are still so often retro-fitting the world around us highlights that Universal Design requires a transformational approach whereby we *rethink* the ongoing planning of the world in which we live so that planning is

by all of us and *for* all of us—critically, reflecting the mantra of “nothing about us without us” as a core aspect of universal approaches. This doesn’t mean only having ramps and no stairs—in reality ramps are not actually accessible to all of us—but it does mean that when we are designing building access points we need to ensure that ramps and lifts and braille and appropriate width doorways and clear signage, for example, are an equally important part of the planning and building use, from the outset. This is not yet, and has not been, the reality. While some aspects of accessibility are relatively commonplace in some parts of the world now, even the humble curb cut is an example of a hard fought “win” in working toward accessibility (Hendren, 2020). As Hendren (2020) explains, the battle for curb cuts in North America included disability activists pouring concrete on curbs or smashing sections of existing concrete curbs to flatten them in order to create “do it yourself” curb cuts, in response to the circular arguments of city planners that curb cuts were not necessary since disabled people were never seen out on the street.

In unpacking the significance of such changes to the built environment, Hendren (2020) builds on the work of Hayward and Swanstrom (2011) who conceptualize the inequities of built environments as being examples of “thick injustice”. Thick injustices can be understood as aspects of our societies and environments that are so deeply embedded and accepted without question that they lead to a raft of largely unrecognized injustices that seem both permanent and inevitable as we struggle to imagine the world other than what it is in the here and now (Hayward and Swanstrom, 2011). Within and beyond built environments, the notion of thick injustices speaks to the seemingly endless argument that accessibility is expensive and requires resources that “people might not have”. Importantly, (and separate to the counter-argument that if we are all equally humans then the rights of some of us are not rendered invisible or secondary to the rights of others of us), Universal Design provides a valuable opportunity from which it is often possible to circumnavigate these issues through planning for *all* from the outset (and thus avoiding the need for costly and time-consuming retrofitting).

Reflecting on the seemingly intractable, and yet fully alterable, realities of the built environment, and the need for Universal Design, we can recognize the attitudinal underpinnings that require addressing—which may be more difficult to shift than bricks and mortar (or indeed concrete). At the core of considerations of accessibility (and inaccessibility) are the “normalizing” beliefs that narrow humans and our ways of being and doing—as if human diversity simply does not exist, or at the very least that it does not need to be accounted for, or indeed that it is an “expensive inconvenience” best left ignored. Implementing Universal Design is one important tool through which we can begin to address these attitudinal and material barriers—and injustices.

Universal Design for Learning

Although beset with the complexities of thick injustice, the importance of Universal Design, and its application and implementation, is nonetheless relatively easy to wrap our minds around once we start to think about the physical accessibility of environments. However, the environment, while a significant domain for the production and reproduction of injustices, is only one of many aspects of *inaccessibility*. When we consider accessibility within pedagogy, and the curricula that emerge from it (Ellsworth, 2005), we turn to UDL.

The notion of UDL developed through applying notions of Universal Design to the ways in which we teach and learn. The framework of UDL is generally understood to have been developed by David Rose, Anne Meyer and colleagues at CAST (previously, the Center for Applied Special Technology) in Massachusetts, USA. UDL philosophy and practice is embodied in a set of three principles that address the “why”, “what” and “how” of learning, through creating accessible “universal” approaches to engagement, representation, action and expression (CAST, 2018). In seeking to create “universal” approaches, the question of “who” we are planning for also becomes key.

Of course, for any of us to genuinely engage in and benefit from educational opportunities, we must first have access to those very same opportunities. Thus, this need for accessibility seems rather self-evident. However, for many students, learning opportunities are inaccessible, preventing participation and engagement and creating a vastly inequitable “playing field”. This has serious implications at many levels. Pressing the “snowy ramp” analogy a little further, the ramp is necessary but it also needs to be useable and valued—as in, “access” that is actually accessible, available, and respected as legitimate. Taking a UDL standpoint means that such accessibility becomes no longer a privilege or an option, it is the realization of a fundamental human right for all. As stipulated in General Comment 4 on the CRPD “[c]urricula must be conceived, designed and implemented in such a way as to meet and adjust to the requirements of every student, and provide appropriate educational responses” (United Nations, 2016, paragraph 26).

UDL is intended to address—and, ultimately, remove—barriers to educational accessibility that create, but also go far beyond, access to *information* and instead facilitate access to *learning* (Edyburn, 2005). Applying principles of Universal Design to education means that—consistent with teaching as a *creative* role, rather than a task for “transaction” or “transfer” of knowledge—we create and implement pedagogies and curricula that are built on a foundation of recognizing and valuing humans in all our diversities and proactively and responsively plan for flexibility in enabling genuine access to equitable educational opportunities for all students (Cologon and Lassig, 2020; Edyburn, 2010; Glass et al., 2013). UDL is underpinned by the idea that “learner variability is the rule not the exception” and that UDL as an approach is innovative, strengths based, and

benefits everyone (CAST, 2018; Glass et al., 2013). At its core, implementing UDL is about providing multiple representations of information, ways to engage, and different ways to express knowledge and learning, with the emphasis on the role of the teacher in creating valued and recognized access for all through careful planning (Kurth, 2013).

The United Nations Committee on the Rights of Persons with Disabilities defines UDL as:

a set of principles providing teachers and other staff with a structure for creating adaptable learning environments and developing instruction to meet the diverse needs of all learners. It recognizes that each student learns in a unique manner and involves: developing flexible ways to learn, creating an engaging classroom environment; maintaining high expectations for all students while allowing for multiple ways to meet expectations; empowering teachers to think differently about their own teaching; and focusing on educational outcomes for all, including persons with disabilities.

United Nations (2016, paragraph 26).

Applying a framework of UDL can facilitate the development of pedagogy and practice underpinned by the recognition of the reality that people "... all operate across a spectrum of emotional, perceptual, and cognitive capacities and proclivities when we learn" (Glass et al., 2013, p. 99). Valuing human diversity and taking an anti-oppressive approach can support us in addressing the barriers otherwise presented through narrowly conceived pedagogies that do not recognize or build on human diversity (Symeonidou, 2020).

At a philosophical and practical level, UDL involves the recognition that student outcomes and abilities, rather than being inherent or individual, are the result of interactions between an individual and their environments (Rose et al., 2018). Applying this thinking then creates opportunities to recognize that "poor outcomes" or "low performance" can often be a consequence of inaccessibility in the design of the learning experiences, materials, or the assessment approach (Rose et al., 2018).

UDL is a cohesive and logical framework and philosophical standpoint from which to approach planning that is inclusive of all students. In its implementation, though, UDL has been criticized as being interpreted and applied in limited ways that do not include all students, in particular students with intellectual disabilities (Smith and Lowrey, 2017). While the intention of universality is self-evident, even in the name UDL, explicit consideration of who we are as diverse groups of humans—and equal valuing of the importance of access to equitable educational opportunities for *all* of us—is needed to bring the promise of UDL even close to a reality. Consequently, it is helpful to dig deeper to start to address how UDL can be implemented in a way that is authentically inclusive of all, including (from a social model perspective) students who experience the imposed social oppression that constitutes disability.

UDL: a social model approach

At its core, UDL is understood to be a responsive framework that is flexible and strengths based and facilitates high expectations and authentic access to education (Cologon and Lässig, 2020). Implementing UDL involves considering every aspect of pedagogical planning and assessment to determine how it enables (or can enable) multiple means of engagement, multiple means of representation, and multiple means of action and expression in order to enable learning for all students (CAST, 2018). This includes considering how students can be provided with positive and appropriate motivators and feedback to engage with learning (for example, choice, modeling, and reflection), what scaffolds can be put in place to ensure that representation is accessible to all students (for example, visual and auditory alternatives and markers and explanations of key information), and what resources and modes of accessing learning can be put in place (for example addressing methods of communication and the role of technology in facilitating participation and demonstrating of knowledge) (Cologon and Lässig, 2020).

Taking this approach requires identifying potential barriers to learning within the pedagogy and curricula and multiple solutions to these barriers. The emphasis, therefore, is on why, how, and what the *teacher*—and all those involved in educational teams (including teaching assistants and principals and so forth)—can do (and provide) to enable access to learning for *all* students (and what barriers can be addressed within pedagogical practices). This involves moving away from a focus on what a *student* might *not* be able to do as the barrier. It also involves a rejection of the ableist and mythical notion of a "normal student".

The idea of planning for "normal students"—with or without the intention of making adaptations for some (therefore inferred to be "abnormal") students—is a traditionally accepted approach to education that is underpinned by a deeply ableist "medical model" understanding of human diversities. Taking an example from Australia, the Australian Curriculum (Australian Curriculum, Assessment and Reporting Authority [ACARA], n.d.) still uses the definition of disability from the 1992 Disability Discrimination Act in which disability is defined as:

"total or partial loss of the person's bodily or mental functions

- total or partial loss of a part of the body
- the presence in the body of organisms causing disease or illness
- the presence in the body of organisms capable of causing disease or illness
- the malfunction, malformation or disfigurement of a part of the person's body

- a disorder or malfunction that results in the person learning differently from a person without the disorder or malfunction
- a disorder, illness or disease that affects a person's thought processes, perception of reality, emotions or judgment or that results in disturbed behaviour" (ACARA, n.d., n.p.)

Focusing on disability, a medical model perspective holds that, disability is the opposite of ability, that disability is impairment, that disability is necessarily negative or tragic, that to be disabled is to be "lesser" or "inferior"—in some sense "broken", that disability is an undesirable state of being, that those of us who experience disability need to be "fixed", and that disability needs to be eliminated or ameliorated (French and Swain, 2004). When viewed from a medical model perspective, the impact and role of society in disabling people is either not recognized or actively not seen to exist (French and Swain, 2004; Thomas, 2004).

Disability Rights Activists and those within the field of Disability Studies have worked hard to disestablish this ableist way of viewing each other as humans, with many implications for education. The social model of disability, despite many critiques, provides an alternative—and productive—viewpoint. As Oliver (2004, p. 21) writes:

In the broadest sense, the social model of disability is about nothing more complicated than a clear focus on the economic, environmental and cultural barriers encountered by people who are viewed by others as having some form of impairment—whether physical, mental or intellectual. The barriers disabled people encounter include inaccessible education systems, working environments, inadequate disability benefits, discriminatory health and social support services, inaccessible transport, houses and public buildings and amenities, and the devaluing of disabled people through negative images in the media—films, television, and newspapers.

Understanding a social model of disability may be beneficial to applying UDL. UDL can also be viewed as a parallel to the social model in that it relocates the "problem" to society (in this case to the education setting, pedagogy, or curriculum) rather than viewing the "problem" as/within the student. Within UDL the challenges identified within a social model understanding are then actively and directly addressed through ameliorating *barriers* within *society*—in this case, within education in particular.

Like others before me, I would argue that applying a social model approach to education is not to ignore criticisms of the social model, nor to propose a "disembodied" understanding of disability (for discussion of the social model, see Hughes and Paterson, 1997; Oliver, 2013; and Siebers, 2019). Including the bio-social aspects of impairment, as explored below. However, applying a social model approach to UDL is to acknowledge the social and material barriers that impact on the participation and lived experiences of those of us who are disabled, and proactively work toward ameliorating these problems that prevent participation.

In building on earlier conceptualizations of the social model, and in addressing some of the key criticisms (including the "disappearing body"), Thomas (1999, 2001, 2004, 2007, 2010, 2012, 2019) offers a *social-relational* understanding of disability. A social-relational understanding of disability incorporates the oppressive social and relational aspects of disability (disablism), alongside the individual lived experience of the bio social effects of impairments (impairment effects). In explicating the bio social nature of impairment effects, Thomas (2019, p. 1040) identifies that impairments are "bio social in character i.e., they have real material bodily effects and are simultaneously culturally constructed especially by the medical profession and bio/scientists in the West in current times". In offering closing remarks on her contribution to the theorizing of disability, Thomas (2019) reiterates the ongoing and pressing relevance of these two key aspects of disability in the form of disablism and impairment effects.

Disablism: refers to the social imposition of *avoidable restrictions* on the life activities, aspirations and psycho-emotional well-being of people categorised as "impaired" by those deemed "normal". Disablism is *social-relational* in character and constitutes a form of *social oppression* in contemporary society—alongside sexism, racism, ageism and homophobia. As well as enacted in person-to-person interactions, disablism may manifest itself in institutionalised and other sociostructural forms.

Thomas (2010, p. 37).

Impairment effects: the direct and unavoidable impacts that "impairments" (physical, sensory, intellectual, emotional) have on individuals' embodied functioning in the social world. Impairments and impairment effects are always bio-social and culturally constructed in character and may occur at any stage in the life course.

Thomas (2012, p. 211).

From a social-relational understanding of disability, disablement can further be understood to occur through "barriers to doing" and "barriers to being". Barriers to doing constitute avoidable socially imposed environmental or economic material barriers to participation—for example, all forms of access barriers (Cologon and Thomas, 2014). Thus, the core considerations in implementing UDL. Barriers to being, incorporate the attitudinal factors that need to be addressed if the thick injustices within disablism are ever to be adequately countered. As Thomas (2007, p. 72) outlines, barriers to being are enacted as a process of psycho-emotional disablism, which "involves the intended or unintended 'hurtful' words and social actions of non-disabled people (parents, professionals, complete strangers, others) in inter-personal engagements with people with impairments". For example, one aspect of considering representation includes directly addressing barriers to being.

When we consider these concepts in relation to implementing a framework of UDL, we can see that, as part of the “who”, impairments and impairment effects are one of many forms of human diversity. As a “universal” framework, necessarily then, in implementing UDL we are actively planning and creating pedagogy and curriculum that is accessible to students who experience impairment effects. This has implications for engagement, representation, action and expression. If we are to engage in anti-oppressive pedagogy, this also means that in addressing this matter of “who”, and planning for multiple methods of engagement, representation, action and expression, we will be directly addressing barriers to being and barriers to doing. Arguably then, not only can UDL be considered an implementation of a social relational model in practice, but a social relational model—in particular the concepts of barriers to doing and barriers to being—can be a useful tool to use in implementing UDL.

Symeonidou and Chrysostomou (2019, p. 363), in research with public primary school teachers in Cyprus found that “disability-focused anti-oppressive pedagogies do not view disability as the exact opposite of ability. Anti-oppressive pedagogies seek to make a difference by understanding disability as a form of social oppression, but at the same time they appreciate disability as a lived experience, and as a complex discursive, cultural, and historic construct”. Going on to explain that “This realization means that a social model of disability informs anti-oppressive pedagogies, without ignoring other factors related to disability” (Symeonidou and Chrysostomou, 2019, p. 363).

UDL can, therefore, be understood and implemented as an *anti-ableist* and anti-oppressive framework for education. However, although this is possible it is not necessarily the case in application. UDL can also be superficially applied without any attempt to address the thick injustices within education thus existing within, or even further contributing to, oppressive pedagogies—as in, pedagogies that present some of us, for example those of us who experience disability, as passive and deficient, in need of pity (Symeonidou and Chrysostomou, 2019). While the potential to exist within or even perpetuate oppressive pedagogies could be considered an inadequacy of UDL, perhaps the focus instead needs to be on how we understand and implement UDL as an aspect of anti-oppressive pedagogies.

Harking back to Hayward and Swanson’s (2011) conceptualization of thick injustices underlines the extent to which oppression is an insidious and everyday encounter for so many of us. Pedagogy and curricula are no exception to this reality. However, when we shift our thinking from planning and implementing pedagogies from a “central” point (stemming from myths of “normal”) with variations retrofitted toward UDL, we open possibilities to move beyond deficit understandings of students toward creating the space to recognize and address the barriers to learning and opportunity that are otherwise often hidden and unexamined. At a fundamental level, one aspect of such an anti-oppressive UDL approach requires engaging with a social model understanding of disability, from which we can recognize that disability is negatively experienced through social (including attitudinal) and material (including environmental and curricular) barriers imposed upon those of us who are, therefore, disabled. The key transformation in thinking that is necessary here, is to move from viewing barriers or “problems” as existing within individuals (a deficit based/medical model understanding) toward recognizing that barriers exist in the interactions between people and our societies (a social model understanding) and therefore that UDL—and indeed the project of democratic education—is about removing social and material barriers in order to produce and engage all students within accessible, equitable and meaningful learning opportunities. This requires applying a social model approach to understanding and implementing UDL. As such, taking a social model approach to implementing UDL can have benefits for ensuring the “universal” aspect of UDL not only in recognizing and including those of us who experience disability, but beyond this in relocating our focus from seeking to change, or exclude, students, or even on retrofitting learning experiences that were designed for the elusive “majority”, toward a productive focus on pedagogies for all.

One avenue through which we can apply a social model understanding to the implementation of UDL is through drawing on Disability Studies as a core aspect of informing our understandings and implementation of UDL as anti-oppressive curricula and pedagogies. As Symeonidou (2020, p. 2) explains, “Disability Studies provides the framework to rethink disability and infuse a disability perspective into the curriculum”. The ongoing lived experience of oppression for so many of us results in an active need for anti-oppressive curriculum and pedagogies (Symeonidou, 2020). Symeonidou (2020, p. 2) proposes the need to draw on critical pedagogy (informed by Paulo Freire and others) and postcritical pedagogy (informed by Michel Foucault and others) “in order to follow pedagogies that do not present the person with a disability as the Other ... but rather encourage learners to understand that disability is a form of social oppression that needs to be challenged” and to ensure that “disability” is not viewed as the opposite of “ability”.

However, at present, those of us who experience disability are largely “invisible” within curricula and pedagogy, it falls upon educators to bring this perspective to everyday practice (Valle and Connor, 2019). This is where guidance for educators, from a social model perspective, is essential. Without such guidance from a Disability Studies approach, no matter how well-meaning and dedicated the teacher, such instruction is likely to “accidentally” fall into oppressive pedagogy in which disability is cast in an ableist light as a within-person deficit in need of amelioration or elimination (Symeonidou and Chrysostomou, 2019) and as a reason for which a student—or entire group of students—can or should be excluded from the classroom and curriculum. A viewpoint from which—ironically—UDL can be justified as an approach for *some* and not *all* students.

Progressing forward with UDL

UDL offers a great deal of promise. However, as I have explored above, in its implementation UDL does not necessarily “live up to” this promise. Drawing together the notion of anti-oppressive pedagogies and building on a social model understanding of the

barriers that students often face may have the potential to create a space for implementing UDL as a framework for planning environments and pedagogies that reduce rather than perpetuate barriers to educational opportunities. Translating the intention of UDL beyond the material and into the attitudinal holds possibilities for contributing to addressing the thick injustices that include but also go far deeper than the (nonetheless essential) existence of ramps or accessible learning materials.

As explored within this chapter, implementing UDL involves a process of considering every aspect of pedagogical planning and assessment to determine how, *as teachers*, we can enable multiple means of engagement, representation, and action and expression in order to enable learning for all students (CAST, 2018). I would argue that UDL has the potentiality to be used to genuinely contribute toward more inclusive education experiences and systems through anti-oppressive pedagogies and the curricula that emerges from them. Again, though, this requires remembering that UDL is not about taking a one-size-fits-all approach. Instead—through a focus on teachers and *how* we teach—UDL forms a flexible and responsive approach that starts from the position of recognizing, valuing, planning and implementing pedagogies and curricula for the diversity of students. Reminding us of Slee's (2011) call for "irregular schools", it could perhaps be argued that UDL holds such potential that it could be used as a framework for transformation of education settings, pedagogies, curricula, and systems even to a point where it would be possible to move beyond talk of "inclusion" and "exclusion" toward education for all in a deep sense—an education that is accessible to all in material and practical senses and, importantly, within which *all of us* are welcome, valued, and seen as worthy members of educational settings and experiences.

References

- Australian Curriculum, Assessment and Reporting Authority [ACARA], (n.d.). The Australian Curriculum: Meeting the Needs of Students With a Disability, <https://www.australiancurriculum.edu.au/resources/student-diversity/meeting-the-needs-of-students-with-a-disability/>.
- CAST, 2018. Universal Design for Learning Guidelines: Version 2.2. <http://udlguidelines.cast.org>.
- Cologon, K., Lassig, C., 2020. Universal approaches to curriculum, pedagogy and assessment. In: Graham, L. (Ed.), *Inclusive Education for the 21st Century*. Allen & Unwin.
- Cologon, K., Thomas, C., 2014. Ableism, disablism and the early years. In: Cologon, K. (Ed.), *Inclusive Education in the Early Years: Right From the Start*. Oxford University Press, South Melbourne, pp. 27–48.
- Edyburn, D.L., 2005. Universal design for learning. *Spec. Educ. Technol. Pract.* 7 (5), 16–22.
- Edyburn, D.L., 2010. Would you recognise Universal Design for Learning if you saw it? Ten propositions the new directions for the second decade of UDL. *Learn. Disabil. Q.* 33, 33–41.
- Ellsworth, E., 2005. *Places of Learning: Media, Architecture, Pedagogy*. Routledge, New York.
- French, S., Swain, J., 2004. Whose tragedy? Towards a personal non-tragedy view of disability. In: Swain, J., French, S., Barnes, C., Thomas, C. (Eds.), *Disabling Barriers—Enabling Environments*, second ed. Sage, London, pp. 34–40.
- Giangreco, M.F., Ruelle, K., 2002. Absurdities and Realities of Special Education. The Complete Digital Set. Peytral Publications Inc. www.peytral.com.
- Glass, D., Meyer, A., Rose, D.H., 2013. Universal design for learning and the arts. *Harv. Educ. Rev.* 83 (1), 98–119.
- Hayward, C.R., Swanstrom, T., 2011. *Justice and the American Metropolis*. University of Minnesota Press, Minneapolis.
- Hendren, S., 2020. What Can a Body Do? How We Meet the Built World. Penguin Random House, Riverhead.
- Hughes, B., Paterson, K., 1997. The social model of disability and the disappearing body: towards a sociology of impairment. *Disabil. Soc.* 12 (3), 325–340. <https://doi.org/10.1080/09687599727209>.
- Kurth, J.A., 2013. A unit-based approach to adaptations in inclusive classrooms. *Teach. Except. Child.* 46 (2), 34–43. <https://journals.sagepub.com/home/tcx>.
- Oliver, M., 2004. If I had a hammer: the social model in action. In: Swain, J., French, S., Barnes, C., Thomas, C. (Eds.), *Disabling Barriers—Enabling Environments*, second ed. SAGE Publications, London, pp. 7–12.
- Oliver, M., 2013. The social model of disability: thirty years on. *Disabil. Soc.* 28 (7), 1024–1026. <https://doi.org/10.1080/09687599.2013.818773>.
- Rose, D., Robinson, K.H., Hall, T.E., Coyne, P., Jackson, R.M., Stahl, W.M., Wilcauskas, S.L., 2018. Accurate and informative for all: Universal Design for Learning (UDL) and the future of assessment. In: Elliott, S.N., Kettler, R.J., Beddow, P.A., Kurz, A. (Eds.), *Handbook of Accessible Instruction and Testing Practices: Issues, Innovations, and Applications*, second ed. Springer International Publishing, New York, NY, pp. 167–180.
- Siebers, T., 2019. Returning the social to the social model. In: Mitchell, D.T., Antebi, S., Snyder, S.L. (Eds.), *The Matter of Disability. Materiality, Biopolitics, Crip Affect*. University of Michigan Press, Michigan, pp. 39–47.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Routledge, Abingdon, Oxon.
- Smith, S.J., Lowrey, A., 2017. Applying the Universal Design for Learning framework for individuals with intellectual disability: the future must be now. *Intellect. Dev. Disabil.* 55 (1), 48–51. <https://doi.org/10.1352/1934-9556-55.1.48>.
- Symeonidou, S., 2020. Teacher education for inclusion and anti-oppressive curriculum development: innovative approaches informed by disability arts and narratives. *Int. J. Incl. Educ.* 1–15. <https://doi.org/10.1080/13603116.2020.1711819>.
- Symeonidou, S., Chrysostomou, M., 2019. 'I got to see the other side of the coin': teachers' understandings of disability-focused oppressive and anti-oppressive pedagogies. *Int. J. Educ. Res.* 98, 356–365. <https://doi.org/10.1016/j.ijer.2019.09.012>.
- Thomas, C., 1999. *Female Forms: Experiencing and Understanding Disability*. Open University Press, Buckingham.
- Thomas, C., 2001. Feminism and disability: the theoretical and political significance of the persona and the experiential. In: Barton, L. (Ed.), *Disability, Politics and the Struggle for Change*. David Fulton, London, pp. 48–58.
- Thomas, C., 2004. Disability and impairment. In: Swain, J., French, S., Barnes, C., Thomas, C. (Eds.), *Disabling Barriers—Enabling Environments*, second ed. SAGE Publications, London, pp. 21–27.
- Thomas, C., 2007. *Sociologies of Disability and Illness: Contested Ideas in Disability Studies and Medical Sociology*. Palgrave Macmillan, Hampshire.
- Thomas, C., 2010. Medical sociology and disability theory. In: Scambler, G., Scambler, S. (Eds.), *New Directions in the Sociology of Chronic and Disabling Conditions: Assaults on the Lifeworld*. Palgrave Macmillan, Basingstoke, pp. 37–56.
- Thomas, C., 2012. Theorising disability and chronic illness: where next for perspectives in medical sociology? *Soc. Theor. Health* 10 (3), 209–228. <https://doi.org/10.1057/sth.2012.7>.
- Thomas, C., 2019. Times change, but things remain the same. *Disabil. Soc.* 34 (7–8), 1040–1041. <https://doi.org/10.1080/09687599.2019.1664074>.
- United Nations, 1948. Universal Declaration of Human Rights. <https://www.un.org/sites/un2.un.org/files/udhr.pdf>.
- United Nations, 1966. International Covenant on Civil and Political Rights. <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>.

- United Nations, 2006. Convention on the Rights of Persons With Disabilities. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- United Nations, 2014. General Comment No. 2, Article 9: Accessibility. http://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CRPD/C/GC/2&Lang=en.
- United Nations, 2016. General Comment No. 4, Article 24: Right to Inclusive Education. <https://www.refworld.org/docid/57c977e34.html>.
- Valle, J., Connor, D., 2019. Rethinking Disability: A Disability Studies Approach to Inclusive Practice, second ed. Routledge, New York.

Teachers with disabilities in mainstream schools: embodying and enacting inclusive education

Nidhi Singal, Faculty of Education, University of Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Insights from current literature of teachers with disabilities	342
Experiences of teachers with disabilities: cross country perspectives	343
Research approach	343
Note on anonymity	343
Official documents: an overview	343
Findings	343
Reasons for becoming a teacher	343
Teacher identity	344
Impact on pedagogy	346
Engagement with colleagues	346
Students	347
Parents and the community	347
Impact of covid-19	348
Discussion	348
Facilitating inclusion	348
Systemic efforts needed to increase participation and retention of teachers with disabilities	349
Advocating for change through robust evidence	349
References	350

Over the last few decades, considerable attention has been directed toward the development of inclusive schools, with significant efforts being made both at the international and national level to specifically promote increased enrollment, engagement and learning outcomes for children with disabilities. Two rather significant international developments have been the adoption of the United Nations Sustainable Development Goals [SDGs] (United Nations, 2020) and the United Nations Convention on the Rights of Persons with Disabilities [CRPD] (2006, 2016). The SDG 4 “inclusive and equitable quality education” and promoting “lifelong learning for all” is part of the United Nations (UN) 2030 Agenda for Sustainable Development pledge to leave no-one behind. The agenda promises a “just, equitable, tolerant, open and socially inclusive world in which the needs of the most disadvantaged are met.” Several high-profile declarations in the past decade have signaled growing regional interest in providing access to education for children with disabilities (Global Education Monitoring Report, 2020). Research evidence notes that over the years there has been a positive impact globally with rates of enrollment increasing slowly, though considerable disparities remain within and across countries. For example, in India official estimates suggest a significant increase in enrollment numbers of children with disabilities over the past two decades. However, significant disparities within states exist (Singal et al., 2017). While enrollments have shown an increase, there are significant concerns in relation to issues around classroom experiences, engagement with the pedagogical processes and learning outcomes, as evident in the data from Pakistan (Malik et al., 2020; Singal et al., 2018).

Even though significant efforts are being made to include more children with disabilities into the education system, very little attention has been given to increasing diversity among those working in these school settings. The report “Transforming the Education Workforce” published by the Education Commission (2019) noted that an increasingly diverse student population needs a deeper engagement with a diversified teacher workforce. However, discussions on teacher diversity in countries around the world tend to focus on race, gender and ethnicity with very little attention being paid to disability.

Globally, participation of persons with disabilities in the workforce remains a challenge. As Stoveska (2020) discussing data collected by the International Labor Organization (ILO) notes, “People with disabilities are less likely to participate in the labor force: almost two-thirds of those who are of working age are outside the labor force” (n.p.). Worldwide, the statistics show that “Only one in three people with disabilities is employed on average. Significantly, they are two times less likely to be employed than people without disabilities” (Stoveska, 2020, n.p.). These figures are consistent with several other studies. Lamichhane (2015) noted that, “with regard to the participation rate of people with disabilities in the labor market in various countries, the available employment statistics are not encouraging” (p. 23). Mizunoya and Mitra (2013) using data on developing countries from the World Health Survey, showed that people with disabilities generally experience lower employment rates compared to their non-disabled counterparts, and the unemployment rate of persons with multiple impairments is even lower than those with single impairment. Data from the World Report on Disability (WHO, 2011) for 51 countries further shows that while the employment rate is lower for people with disabilities compared to their non-disabled counterparts, it is even more serious for women with disabilities. As Lamichhane (2015) outlines, a range of factors such as the lack of access to education and training, nature of the

workplace, or employers' perceptions/misconceptions of disability as not being very productive for example, could all contribute to their continued exclusion from the labor market.

The aim of this chapter is two-fold: (1) to highlight what the current literature tells us about teachers with disabilities working in mainstream schools; and (2) to discuss insights from an exploratory study undertaken with teachers with disabilities across four Low- and Middle-Income Countries (LMICs).

Insights from current literature of teachers with disabilities

In the recent past, there have been several studies focusing on the experiences of teachers with disabilities, but these have primarily been undertaken in Northern contexts. 53 research studies were identified in a review of this literature published in [Neca et al. \(2020\)](#), highlighting themes around teachers' life trajectories, educational practices and challenges; impact and need for teacher training, students and school principals' perspectives about teachers with disabilities, while revealing the underrepresentation of teachers with disabilities.

[Singal and Ware \(2021\)](#) undertook a review of existing databases, specifically databases of the Education Resources Information Center (ERIC) and the British Education Index (BEI), to identify literature on teachers with disabilities working in mainstream school settings. Of the 59 articles identified in this review, more than a third ($n = 21$) focused on the experience of teachers with learning difficulties, primarily those with dyslexia. Teachers with sensory impairments (hearing and/or visual impairments) were the second most frequent ($n = 12$), followed by thought pieces and/or articles not specifying type of disability ($n = 11$). Studies reflecting on teachers with physical disabilities ($n = 6$), those who were Autistic ($n = 1$) and those working across a range of impairments ($n = 8$) were the least common. Few of the existing studies were subject specific, although three studies focused on physical education (USA), two looked at music education (USA) and one at science education (UK). The existing literature provides some interesting insights around motivations for becoming a teacher, with reasons such as pragmatic factors (pay, holidays, etc.), personal variables (passion for working with children, having relevant skills, etc.), and a strong sense of social justice ([Ware et al., 2021](#); [Ferri et al., 2005](#)) being often cited. Tensions around disclosure of one's disability status, particularly hidden disabilities, are identified in the literature, with [Saltes \(2020, p. 22\)](#) suggesting that the "normalizing ideology" of a "capable teacher" is a substantial barrier which causes both attitudinal and physical difficulties. This was markedly highlighted in the work by [Valle et al. \(2004\)](#) in the USA, where their participants noted significant struggles around disclosure of their learning disabilities due to issues of stigma. In contrast, other work with teachers with dyslexia note their value to classrooms with diverse student populations ([Burns and Bell, 2011](#); [Vogel and Sharoni, 2011](#)). [Ware et al. \(2021\)](#), researching in England with ten teachers and trainee teachers having a range of disabilities, found that while there was a strong discourse of inclusive education in schools, this did not extend to teachers. As a result, nine out of ten teachers faced difficulties and discrimination at work, with the individual being made responsible for adjusting rather than the barriers in the system being addressed. Similarly, in Israel, [Alon and Lishchinsky \(2019: n.p.\)](#) note that "dealing with the school environment is often more difficult and demanding than coping with the disability itself."

Findings from the literature highlight the central role of teachers with disabilities in the classroom as positive role models and for improving teaching and learning practices. [Anderson \(2006, p. 368\)](#) argues "... pedagogy is also shaped by our life experiences. Teachers with disabilities offer knowledge through their bodies and experiences that isn't usually part of the curriculum. Disabled teachers embody pedagogies of justice, interdependence, and respect for differences." Similarly, [Hayashi and May \(2011\)](#) argue that to have a teacher with a disability, a respected expert in their field, contributes more to developing positive mindsets toward disabled people than exposure to disability in other contexts. In relation to teaching and learning, some studies highlight how teachers with disabilities are better able to differentiate learning activities for children, are more empathetic, and particularly empower students with disabilities ([Burns and Bell, 2010, 2011](#); [Griffiths, 2012](#); [Morgan and Burn, 2000](#); [Ware et al., 2021](#)).

Additionally, research also suggests that students tend to view teachers with disabilities largely positively. For example, [Roberson and Serwatka \(2000\)](#) found that secondary students perceived Deaf teachers to be better at teaching and were more positive toward them. More extensive work undertaken by [Bryant and Curtner-Smith \(2008, 2009a,b\)](#) in the USA examined students' perceptions of the competence of a teacher with physical disabilities in teaching Physical Education (PE), and findings showed that these perceptions could alter over time. In three studies, conducted across elementary, middle and high schools, 100 students were randomly assigned to watch one of two almost identical recorded swimming lessons. In the first video the teacher was able-bodied, and in the second video the teacher used a wheelchair. Students were then asked to complete a questionnaire on the lesson content and perceptions of the teacher. Rather worryingly, the researchers report that as students' progress through their schooling, their beliefs about PE teachers with disabilities gradually changes for the worse because "they are socialized into believing that sport, physical activity, and physical education are for what appear to be whole and fit bodies" ([Bryant and Curtner-Smith, 2009b, p. 319](#)). The researchers suggest that this is likely due to the limitations of PE curriculums which focus on athleticism and competitive sport.

Looking specifically at peer reviewed journals published in English it is clear that there is a significant lacuna of papers addressing similar issues in LMICs. Only three papers situated in LMICs were identified in the [Neca et al. \(2020\)](#) review and these were based in Nepal, Malaysia and Ethiopia. In Malaysia, [Lee et al. \(2011\)](#) conducted interviews with 24 teachers with visual impairments to understand their experiences of finding employment. The results provide useful insights into the challenges faced, the importance of individuals' personal characteristics, such as optimism, confidence, and resilience, while negative attitudes from employers and over protectiveness of parents were identified as significant barriers. It is not clear though if these teachers were employed in mainstream or specialist settings.

The study in Ethiopia conducted by Hankebo (2018) focused on exploring adopted modes of communication, teaching experiences and perceptions of deaf teachers in inclusive classrooms. Seven deaf teachers from two primary government schools were interviewed and the findings noted that, “though deaf teachers have positive perceptions for their career, they lacked appropriate pedagogical skills. It also appears that deaf teachers are highly inconvenienced by communication barriers and poor interaction, lack of technology usage, and lack of sign language interpreters” (p. 477). In a mixed methods study across seven schools from four different districts in Nepal which were employing teachers with a visual impairment, Lamichhane (2016) reported that these teachers tended not to teach subjects such as science and mathematics that require frequent use of a blackboard or illustrations of formulas. They also faced several challenges in lesson preparation, marking students’ examination papers, or teaching picture-based content due to the absence of an adequate support system, educational material, and resources. However, positive attitudes, good communication skills, as well as giving more social and moral lessons were reported by students as strengths of these teachers.

Experiences of teachers with disabilities: cross country perspectives

Given the significant lacuna of literature in LMICs, a small exploratory project was undertaken to understand the experiences of teachers with disabilities working in mainstream schools across India, Nepal, Rwanda, and South Africa. The aim of this cross-country study was not to homogenize the experiences of those with disabilities across these contexts. Rather, given that the research was being funded by an organization with global outreach during the covid-19 pandemic, we decided to recruit the sample from different countries. Additionally, given that we were unable to locate any existing networks supporting teachers with disabilities at the country level, and there being no indication of reliable data sets to know how many teachers with disabilities were enrolled in mainstream school settings, a more exploratory approach was adopted. To get an overview of the country contexts, official documents addressing issues of inclusive education, available on-line and in English, were scanned to see if teachers with disabilities were mentioned in the text.

In presenting the findings one is mindful of the fact that contextual realities shape the experiences of individuals. Nonetheless, the aim here is to provide a snapshot of the shared themes emerging in the experiences of teachers with disabilities located across different developing economies.

Research approach

A qualitative survey design was adopted for this study. Initially, one-to-one semi-structured on-line interviews were undertaken with the participants. After an initial analysis of the data, a focus group was organized for validation of emerging themes and additional probing. The interviews covered the following areas: Motivation for becoming a teacher; identity as a teacher; and experiences with others, such as students, colleagues, parents and so on.

In total nine individuals signed up to take part after seeing a call for participants on WhatsApp groups, Twitter, or other circulation lists. All participants signed an e-consent form to ensure that they had understood the purpose of the research and the nature of engagement. At the end of the data collection, participants were offered a small contribution toward the cost of mobile data usage. All the data were transcribed, and analysis was undertaken to identify overarching themes (Neuman, 2013).

As noted earlier, the aim was not to undertake any comparative analysis, but to identify overarching themes and patterns.

Note on anonymity

Anonymity was discussed with participants, and they were given the option of using their first name or a pseudonym. After some discussion, four teachers; Joseph, Veliswa, Ashok, and Arjun wanted to be identified by their names, while the other five asked for a pseudonym to be used. This has been respected in the reporting of the data here.

Table 1 provides an overview of the participants.

Official documents: an overview

Through an online search, focusing exclusively on official education policies (available in English) of the four selected countries, an attempt was made to understand if teachers with disabilities were visible in these texts. Table 2 highlights key insights emerging from this search.

Thus, official texts varied in their mention of and emphasis on teachers with disabilities, with no specific efforts shining through. None of the official documents had any well-articulated vision around the recruitment and role of teachers with disabilities.

Findings

Reasons for becoming a teacher

In examining the reasons participants gave for becoming a teacher, three factors emerged: (i) passion for teaching; (ii) belief that it was a good or suitable job; and (iii) lack of opportunities to access other jobs. Across all four countries, seven teachers, Joseph, Mary,

Table 1 Participant overview.

<i>First name</i>	<i>Country</i>	<i>Sex</i>	<i>Age</i>	<i>Disability</i>	<i>Highest level of education completed</i>	<i>Completed formal teacher training?</i>	<i>What level is the school?</i>	<i>Years as a teacher</i>	<i>Years as an English teacher</i>
Joseph	South Africa	Male	56	Physical	Undergraduate degree	Yes	Secondary school	15	7
Veliswa	South Africa	Female	49	Physical	Undergraduate degree	Yes	Secondary school	26	20
Jacob	Rwanda	Male	29	Physical	Secondary school	Yes	Primary school	3	3
Mary	Rwanda	Female	39	Visual	Undergraduate degree	Yes	Secondary school	Unemployed at time of interview	N.A.
Gabriel	Rwanda	Male	35	Physical	Undergraduate degree	Yes	Inclusive school	9	9
Arjun	Nepal	Male	33	Mental health	Master's degree	Yes	Secondary school	9	9
Ashok	India	Male	37	Visual	Master's degree	Yes	Primary school	14	14
Sania	India	Female	37	Physical	Master's degree	Yes	Secondary school	8	8
Ankur	India	Male	64	Physical	Master's degree	Yes	Secondary school	32	32

Ashok, Sani, Jacob, Gabriel, and Arjun, spoke about teaching as their vocation or passion. They noted their affinity for working with young people or being interested in teaching from a young age. Joseph (South Africa) underlined teaching as being a “calling” due to it “not having any incentives at all; it has very little pay at the end of the month.” Similar sentiments were expressed by Sania (India) who described how teaching was a “huge responsibility” connected to the progress of a country with teachers being the “creators of the future.”

Ankur (India) and Jacob (Rwanda) viewed teaching as a more suitable profession for a person with disabilities due to the lack of “tensions” related to disability (Jacob). Ankur declared that 32 years ago when he became a teacher, “the job of a teacher was quite comfortable, no challenges, go to school, come back home at the right time” and so it was a good job for someone with a physical disability like him. In addition, parental views about teaching being a respected job was also a factor in Ankur and Veliswa’s decisions to choose the profession.

Both South African participants, Veliswa and Joseph, had chosen to become teachers due to the lack of access to other jobs. For example, Veliswa wanted to become a social worker but when she did not get a place at university her parents applied on her behalf to a local teacher training college.

In explaining why, they decided to specialize as English teachers participants highlighted three main reasons: Personal enjoyment and ability in the language, belief in the importance of the language, and job availability. Six teachers from all four countries, Sania, Arjun, Jacob, Gabriel, Mary, and Joseph, noted their enjoyment of the English language and their ease and ability in the subject while studying as an important factor in specializing as an English teacher. Other reasons highlighted were the fact that English is a “global language” and hence it is important for young people to know, it made access to higher education feasible and also prepared young people for the workforce. Additionally, some spoke of the enhanced status that being an English teacher provided. For example, Ashok (India) reflected that “if you have command over English, then there is demand for you in the whole world.” Expressing similar sentiments Ankur (India) argued that if a person speaks English, then society thinks that “he comes from a noble family, he is a very knowledgeable person, a very smart person, a very powerful person.” This was also echoed by Joseph in the South African context. Mary (Rwanda) was the only one who noted difficulties in finding a job.

Teacher identity

During the interviews, across the participants, it was interesting to observe that they underplayed their disability when talking about themselves as teachers. Rather, participants focused on traits such as “hard working” (Mary), being accepted by colleagues and as capable professionals.

Sania reflected that her perceptions about her abilities changed after becoming a teacher:

Before becoming a teacher, I was a student, and my thinking was very backward and now it has developed as I have studied more. Even my thinking about my handicap has changed a lot. Earlier I used to consider myself the weakest but now I feel that things are a little okay, it does not matter.

Building self-confidence was something that was valued and discussed by most others, with Ankur (Nepal) rebutting the idea that a teacher with disabilities faced discrimination due to disability. He explained that: “respect depends on your personality, not

Table 2 Official documents making reference to teachers with disabilities.

Country	Mention of teachers with disabilities in official education specific documents
South Africa	<i>White Paper on the Integrated National Disability Strategy, 1997</i> (Office of the Deputy President, 1997): Emphasized the need for capacity building within the education sector, by acknowledging the importance of teachers who are fluent in sign language, so that they can be good deaf adult role models and create positive deaf culture in school. <i>Employment Equity Act</i> (Parliament of the Republic of South Africa, 1998): Legislated that those in a designated group, including persons with disabilities, cannot be unfairly discriminated against and states that there should be diversity in the workforce. There is no specific mention of education. <i>Draft White Paper on the Mainstreaming of the Rights of Persons with Disabilities to Equality and Dignity</i> (Department of Social Development, 2016, 112): States “parties shall take appropriate measures to employ teachers, including teachers with disabilities, who are qualified in sign language and/or Braille, and to train professionals and staff who work at all levels of education. Such training shall incorporate disability awareness and the use of appropriate augmentative and alternative modes, means and formats of communication, educational techniques and materials to support persons with disabilities.”
Rwanda	<i>Law No1 Relating to Protection of Disabled Persons in General</i> (Ministry of Justice, 2007): Outlined specific anti-discrimination policies for people with disabilities and noted affirmative action in recruitment stating, “a disabled person shall be given greater access to employment opportunities than any other citizen in case of equal capacities or in case of equal marks in competition.” <i>Revised Special Needs and Inclusive Education Policy (2018)</i>: Highlights the importance of accessibility for both students and educators underlining a commitment to “developing conditions that are adequately conducive for all learners and educators to interact in a barrier-free and child-friendly educational environment” (p. 1).
Nepal	There was scant reference to teachers with disabilities in policies with the only explicit mention being in the <i>School Sector Development Plan 2016/17-2022/23</i> (Ministry of Education, 2016). In this plan the importance of ensuring accessible sanitation facilities in schools for “children, boys and girls, and differently-abled students and teachers” was noted.
India	<i>Rights of Persons with Disabilities Act (2016)</i>: In order to promote and facilitate inclusive education the aim is to “train and employ teachers, including teachers with disability who are qualified in sign language and Braille.” However, this explicit focus on recruiting teachers with disabilities is not noted in the national programs, such as the <i>Sarva Shiksha Abhiyan (SSA)</i> which sought to universalize primary education, had a clear focus on children with disabilities, and focused on increasing the recruitment of female teachers and teachers from tribal communities.

your physical stature and it depends on what work you do, how do you work, your behavior with others and it depends more on those factors than if you are disabled or not.”

Five of the teachers, Mary, Ashok, Ankur, Veliswa, and Joseph explicitly spoke about how they were role models for students with disabilities. Joseph explained that he felt, “... I’ve got a role to play in society as far as people with disabilities are concerned” and he pointed out that students with disabilities were more likely to be sent to his school because he worked there. Additionally, some of the teachers noted that their presence projected a positive perception of persons with disabilities to those around them. Importantly, Veliswa, Mary, and Ashok also believed that they were more capable of teaching children with disabilities because they had an increased understanding of what it means to be a person with disabilities. Joseph felt the same and made the case that teachers with disabilities were more empathetic toward children with disabilities as they shared a common identity.

Impact on pedagogy

The teachers’ experiences of disability took on more prominence in their discussions about the impact of their disability on their pedagogy and the different challenges and enablers in the school system. A few instances of their disability impacting their ability to carry out their duties as a teacher were noted. Ankur, Veliswa, Joseph, Jacob, Gabriel, and Ashok spoke about challenges in walking long distances, standing for long periods of time, or difficulties with stairs. Some of these teachers received reasonable adjustments to support them in carrying out their responsibilities. Joseph (South Africa), who had strongly advocated to his school management team on behalf of both himself and the children with disabilities, had been able to get changes made to the buildings including the installation of ramps and rails, and moving steps closer together. He also had some change in duties, for example, he was not required to undertake any sports related duties and instead had been able to set up extra-curricular debating classes which he could arrange “comfortably” from his class. Ashok (India), who had a visual impairment as well as a physical impairment, had his teaching schedule set so that he would not have to undertake teaching duties in the dark. He also spoke about receiving accessible teaching resources such as large format textbooks from the government. In contrast, though Mary (Rwanda) who was blind did not feel that her disability impacted her teaching, she was concerned about the lack of accessible material and particularly a lack of translated resources in Braille.

Not all teachers received additional support or appropriate reasonable adjustment. Unlike the teachers in India who received support with transport, Gabriel and Jacob in Rwanda, both with a physical impairment, had to walk to school leaving them tired by the start of the school day. Veliswa (South Africa), who acquired her physical impairment less than 5 years ago, had experienced being placed in inaccessible schools. One such placement, where the school did not have a ramp, resulted in her falling from her wheelchair, rendering her unconscious and being hospitalized. It was only after this incident that she was transferred to a school with more appropriate accessibility. At the time of the interview Veliswa had changed schools again and was employed in a special school which was the most accessible school she had worked in, both in terms of the building and support; for instance, she was supplied with a teachers’ assistant to help her manage the class.

As well as discussing reasonable adjustments Ankur (India) and Joseph (South Africa) spoke about the way in which they adjusted their teaching styles to accommodate challenges in standing at the blackboard. Before he became a principal Ankur used to deliver his lessons sitting in a chair and spoke his lessons rather than writing on the board. Joseph used the opportunity to focus more on group work using pre-prepared worksheets in lieu of writing on the board. Gabriel (Rwanda), who has a physical disability, noted how this could reduce his energy and so to compensate he had to plan carefully to ensure he could do everything he wanted to. He prided himself on his excellent time management and commented on how this was an important quality for a teacher with disabilities to have in order to manage their life. More broadly, challenges were noted in relation to being teachers of the English language, with a lack of access to translated resources and internet and computers in schools, as highlighted by Mary and Jacob in Rwanda. Veliswa (South Africa) noted a specific challenge in teaching English in multilingual classrooms. For example, some children in her class were from Afrikaans speaking homes, some from Xhosa speaking homes, and some had English as a first language, causing tensions among children.

Teachers from India also focused on the official certification of their disability which enabled them to seek employment, disbursements, and an additional stipend to get transport to work and access to health camps for free medical check-ups.

Engagement with colleagues

All teachers spoke highly of their colleagues and generally described having strong and supportive relationships. Sania (India), Veliswa (South Africa), and Gabriel (Rwanda) spoke about being friends with colleagues and offering and receiving emotional support. Ashok (India) and Joseph (South Africa), on the other hand, positioned themselves as being respected members of staff who were sought out by others for advice. Ashok noted that this was particularly in relation to supporting children with disabilities, while for Joseph this was usually for support in speaking English. Colleagues would also ask him to translate or ask questions on their behalf during meetings conducted in English due to his strong command of the language. Jacob (Rwanda) and Joseph also spoke about being accepted as a teacher first by colleagues rather than a person with a disability.

In discussing relationships with school management authorities, Veliswa (South Africa) raised her issue about the lack of accessibility of buildings. Jacob from Rwanda did not report any challenges with the management but noted that he worried that should he become sick due to his physical disability he might be at risk of losing his job, and that there was no protection against this.

Joseph (South Africa), Arjun (India), and Ankur (India) spoke more positively about being given additional responsibilities. Joseph had strong relations with his management team and was often assigned to chair meetings or take minutes. Similarly, Ankur who was a headteacher noted how his seniors at administrative levels were very supportive of him, enabling him to undertake professional development which included visiting schools abroad to familiarize himself with their education systems. Career progression was something which was also noted by Gabriel (Rwanda) who was now in charge of discipline at his school. On the other hand, Sania (India) noted how while promotion was a possibility, she did not feel able to take on the extra load due to her disability.

[teachers with disabilities] get many opportunities because after my post, the next promotion is for that of a school principal. As a handicap person, I cannot take up all the responsibilities that fall on the principal. I feel that if I was not handicapped then maybe I could have easily taken the post and done all the work. A teacher with disabilities can become a principal but personally, I do not want to because the principal has many responsibilities. There are no obstacles from the government, it is my choice not to take up that post.

Students

Most of the teachers related positive interactions with their students and spoke about receiving respect. Sania (India) reflected that this might be both due to her being a teacher of the English language and a person with a disability:

I feel that I am the same for them, just like all the other teachers. They do not think of me as different, but they respect me. They respect me a lot because I have problems in life and yet I continue to struggle and it's not like I have given up. I continue struggling in life and because of that they [the students] might consider me their role model ... They respect me a little more than other teachers because I teach them English. Thus, that also facilitates their respect.

While Veliswa and Joseph in South Africa felt well respected in school, they also recalled some initial challenges which had been overcome by directly engaging with students. Joseph spoke about how initially some students didn't do their homework "because they thought, forget about that disabled teacher he's not going to do anything to us." However, he explained that just like any other teacher when he first started at the school, he "la[id] down the law" and with time students' mindsets changed and he is now well respected by students. He talked about the need for promoting a culture of mutual respect as well as setting out clear behavioral rules for students. Veliswa recalled that when she returned to work after her leg had to be amputated, some of the students were "scared" and would "fold their hands, move a little bit to the back" while other students "came closer to me trying to find out, 'teacher how do you feel?'" She explained how she reassured students that she was okay and after a couple of weeks the students adjusted. Veliswa went on to share how one of the students had an uncle who was walking with the use of a prosthetic leg after his leg was amputated, and knowing this had given her comfort that she would be able to do something like that in time.

Gabriel and Jacob, from Rwanda, were the only teachers to speak about ongoing challenges with students. Jacob spoke about how some students "laugh and laugh" at him due to his physical disability and in some cases also attempt to "imitate how I am walking," and how he engages with them to change their mindset.

I try to make conversation with them, and I try to tell them that disabled people are people like them. I give them different examples of disabled people that are famous in the world and also in our country. What normal people can do also disabled people can do. They can't change my mind, but their minds are changed. I try to change it!

Similarly, Gabriel also noted how some students "laugh at me" and "turn away." However, Gabriel who was more senior than Jacob and Head of Discipline at his school, explained that this happened outside of school. In reflecting on why this kind of behavior occurs Gabriel explained that "normally they [students] don't experience people like me, teachers who are disabled, teachers who are not moving well." As with Jacob, Gabriel also explained that it was his duty to teach the young people and promote acceptance:

I normally call them. I normally say, "you come" and I show them. There are some students who are disabled, we have to sit together. We tell them there is no need to put people aside because they are disabled; we have to teach them.

Parents and the community

The teachers generally reported having strong relationships with parents. Mary (Rwanda), reflecting on her internship in a school for students with visual impairments, spoke about how a parent had told her it was "amazing" to see a blind teacher and that she was more suitable to teach his blind child because she too was blind and so had a better understanding of their needs. Joseph (South

Africa) spoke about how when he first arrived at his school some parents were slightly concerned “you could see with the body language that they think ‘can this fellow actually pull it up?’” However, he explained how this changed:

My matric results in English were very good and then I got the confidence of the parents and hence the parents were comfortable with me, and even coming to me for personal issues outside of the school.

While no teacher reported any concerns with discrimination from parents, some teachers experienced challenges with wider society. Sania (India) stated that she avoided attending any weddings and parties as she felt “different from others” and “shame.” Arjun (Nepal) spoke briefly about being discriminated against due to his mental health difficulties. He also commented that he knew of qualified teachers who “due to their physical disabilities were confined inside their house.” In expanding on this he described Nepalese society as viewing a disabled person as having “something lacking.”

Impact of covid-19

Given that the project took place during the pandemic, it became a discussion during the focus group, as teachers reflected on how this was having an impact on their professional lives.

While most teachers had remained in their permanent teaching positions throughout the pandemic, Mary (Rwanda) was laid off work, having started a temporary contract at the beginning of the pandemic. Veliswa (South Africa) spoke about financial hardship having left one teaching position during the pandemic and starting a new one only recently, leaving her without enough money to buy personal protective equipment (PPE). While these concerns were shared by their colleagues without disabilities, some participants highlighted how their disability had them feeling more vulnerable and at risk. Veliswa also spoke about the increased risk posed to teachers, and particularly teachers with disabilities, who continued to work in person. She explained that at the beginning of the pandemic, she had become “paranoid” about the virus. However, as time progressed, she has learned to live with the risk, explaining “as I learned about it all I adjusted and started acquainting myself with the protocols and how to cope and how to manage stress itself, because it was really scary.”

For the teachers in South Africa there had been a drive to enable vulnerable people to shield and take leave from work at the beginning of the pandemic. Joseph, who had an underlying health condition as well as having a physical disability, explained that he had been granted leave on full pay for 3 months in 2020. This had enabled him to work from home preparing resources while a substitute teacher took his place in the classroom.

An important theme raised by teachers in all four countries, was how the pandemic had placed additional workloads on them. The teachers reported how extended periods of school closures built up pressure in the system to do supplementary work with the students to help them catch up on their schooling. Gabriel (Rwanda) and Joseph (South Africa) explained how the extra workload was particularly challenging for teachers with physical disabilities as the increased movement between classes, as well as the pressure to teach more content in shorter times, caused more fatigue.

Discussion

Across the different country contexts, multiple evidence presented in this chapter, namely academic literature, government documents, and primary research clearly highlight efforts toward the participation of children with disabilities in mainstream school settings but do not extend to teachers with disabilities. Rather, the role and importance of an inclusive teaching workforce remains neglected.

In the final section of this chapter, I will develop three key arguments:

1. The central role that teachers with disabilities play in facilitating inclusion.
2. The need for a more systemic approach in supporting the integration of teachers with disabilities into the education workforce.
3. The important role that evidence can play in taking this agenda forward.

Facilitating inclusion

Evidence gathered across the existing literature continues to show that teachers with disabilities have an important role to play in building and fostering more inclusive school cultures. Their presence promotes inclusion *through* education, rather than simply *in* education. As Pritchard (2010, p. 41) argues, “Disabled people as educators enact exemplary pedagogic justice and socially inclusive practice.”

Teachers with disabilities feel that they are particularly able to understand, empathize and communicate better with students with disabilities (Glazzard and Dale, 2015). It can also have a positive impact on the well-being of pupils with disabilities, aiding their participation and progress through the schooling system. Additionally, Duquette (2000) notes how their presence can be particularly beneficial for fostering education success of children with disabilities by helping them challenge low expectations of

themselves as they can relate to the teachers' personal and professional paths. This notion of being a role model for students with disabilities, but also for parents of children with disabilities, resonated strongly in the evidence presented from the empirical research. By their presence as teachers, they can also challenge negative perceptions of disability present in the wider community.

Additionally, teachers with disabilities are more likely to notice acts of injustice taking place in their school settings. Evidence of frustration, and noticing the really difficult time that students with disabilities were facing in schools, was reported by autistic teachers working in the UK in [Wood and Happe's \(2021\)](#) study. Teachers with disabilities are shown to have a sense of purpose as change agents—they are compassionate, empathetic and effective educators who foster inclusion and change ([Moore et al., 2020](#)). This is particularly so as they draw on their own experiences as students, when they likely faced difficulties because of their disability, and this can motivate them to be positive agents for change.

Finally, having teachers with disabilities in mainstream settings can also act as an important impetus, in many low resource contexts, to invest in accessible teaching and learning materials. In the study presented in this chapter, teachers in Rwanda highlight this as a major barrier, while in [Lamichhane's \(2016\)](#) study teachers with visual impairment in Nepal struggled to teach certain subjects because of the lack of accessible materials. As a teacher in Lamichhane's study mentioned: "It was very hard for me to adjust to school in the beginning because I did not have access to enough books in Braille." Further, observing how it was difficult to mark students' exam papers and needing help from others the teacher goes on to note: "If I were a teacher in the school for students with a visual impairment, I could easily check those papers, as they would be in Braille" (p. 26). It could therefore be argued that having more teachers with disabilities as part of the workforce would provide additional impetus to invest in teaching and learning materials which were inclusive of diverse needs.

Systemic efforts needed to increase participation and retention of teachers with disabilities

Experience and insights from many countries suggests that teachers with disabilities working in an education role in schools face several impediments to their effective and successful employment in the sector. For example, in Wood and Happe's work with autistic teachers in UK they noted the lack of understanding and support in school communities about the need of this teacher group. [Tal-Alon and Shapira-Lishchinsky \(2020\)](#) in their interviews with headteachers who were employing teachers with different physical disabilities revealed the many ethical dilemmas they faced between valuing inclusion while also being concerned about student and staff wellbeing. For example, headteachers commonly noted their concerns that employing teachers with disabilities is a burden when they consider frequent absenteeism. They found themselves "torn by their responsibility to the teacher with the disability on the one hand, and to the other teachers on the other" (p. 10). The dilemma around the notion of students "paying the price" (p. 11) was another significant conflict. The authors conclude that even though there is recognition of the value of employing teachers with disabilities, these dilemmas "reflect the commitment to inclusion not only as an ideology but also a practice" (p. 13) and that more needs to be done to provide school principals with regulated information, sources on accessibility adjustments. These findings are very similar to those by [Ware et al. \(2021\)](#) who noted that in the absence of awareness of official policies and regulations in school management can result in the needs of teachers with disabilities being neglected. Thus, clearly framed and widely publicized official mandates are essential in ensuring that teachers with disabilities are adequately supported in school settings. Encapsulating some of these ideas Veliswa, reflecting on the South African context, noted that the White Paper C (ref?), "which is about inclusivity. This includes only learners but not teachers, so that needs to be looked at."

A common theme that emerges from various studies in the field is the lack of adequate support given to trainee teachers with disabilities, particularly when undertaking teacher training programs. This seems ironical given that many teacher training programs now advocate for the need for inclusivity in teachers, however, they fail to provide adequate accommodation to teacher candidates with disabilities as the trainers themselves fail to take account of the specific needs of this group of trainees ([Bargerhuff et al., 2012](#)). Ensuring training courses are inclusive in terms of trainee accessibility, course content, and trainers, ITT courses should be accessible to people with disabilities and other marginalized groups.

Given the very low participation rates of persons with disabilities in higher education in many developing economies raises issues of the likelihood of future employment opportunities in the education sector. However, as the [Education Commission \(2019\)](#) report suggests that "New approaches are needed to ensure routes into teaching are accessible for underrepresented groups. Traditional routes are often inaccessible to people with disabilities and people from minority groups. Alternative routes could facilitate greater representation." While it is unclear what these new approaches might be, there is a case for thinking more creatively and inclusively about how the participation of persons with disabilities can be increased within the mainstream education workforce. An essential element of this will be to ensure that there is greater dialogue between different stakeholder groups to enable a deeper appreciation of the kind of support needed by teachers with disabilities, and the value that they bring to the culture of the classroom.

Advocating for change through robust evidence

Very little attention is being given to the role and participation of teachers with disabilities within research on inclusive education. This neglect has a multitude of consequences, including little understanding of the lived experiences of teachers with disabilities, and perpetuating a lack of appreciation of their important role in implementing the inclusive education agenda. There is a lack of knowledge of their rates of participation and how these intersect with gender, ethnicity and location, all of which can provide insights into how to address their low levels of participation in the education workforce.

The World Report on Disability (World Health Organization, 2011) categorically stated that the existing “lack of data and evidence ... often impedes understanding and action in the field across various sectors, including education” (p. 263). As its final recommendation the report notes the “need to strengthen and support research on disability” (p. 267). This message also resonates in Article 31 of the CRPD (UN, 2006) which states that research is central to addressing discrimination, changing perceptions, and combating stereotypes and prejudices.

Robust national datasets, as well as in-depth qualitative research elucidating the experiences of teachers with disabilities, has the potential to contribute to both the policy and practice praxis. Subject specific research, as well as intersectional research examining how ethnicity, religion and sexuality intersects with disability, is needed to understand how different dimensions of teachers’ identities interact with disabilities and affects their experiences and inclusion within the workforce.

Additionally, it is useful to be reminded that “good research ... will not on its own change the world. It needs to be accompanied by sophisticated and strategic activism” (Swartz, 2014, p. 5). This is something that was strongly felt by participants in this research study too. As Joseph (South Africa) succinctly argued:

Disabled people have stigma attached and we have become objects of pity. Especially in this part of the world people say “Ahh do you think he can do something worthwhile? Is he worth listening to?” That mindset has got to change, especially in our part of the world. I think we need to be out there, say whatever we have to say, not preaching but putting across what we believe in and making sure we are counted among the able bodied.

Albert and Harrison (2005) make a very useful distinction around how researchers “should be ‘on tap’ not ‘on top’” (p. 8). In addition to this “on tap/on top” distinction, Singal (2019) adds the “on shelf” distinction, suggesting that research which is left on the shelf and accessible only to a few is not worth pursuing. Meaningful research that can support the inclusion of people with disabilities in various life spheres is needed. The role of research in fostering change can no longer be undermined or overlooked, but there continues to be a need to make sure that research (and researchers) is respectful of people and their contexts, while being rigorous in gathering evidence. The role of teachers with disabilities as being agents of change is one which can no longer be neglected within the wider discourse on inclusive education.

This chapter makes a strong case for examining more closely and investing more systematically in developing an inclusive and diverse education workforce, specifically promoting the inclusion of teachers with disabilities in mainstream education systems.

References

- Albert, B., Harrison, M., 2005. Messages From Research. Disability Knowledge and Research (KaR) Programme. DfID, London.
- Alon, N.T., Lishchinsky, O.S., 2019. Ethical dilemmas among teachers with disabilities: a multifaceted approach. *Teach. Teach. Educ.* 86, 102881.
- Anderson, R.C., 2006. Teaching (with) disability: pedagogies of lived experience. *Rev. Educ. Pedagog. Cult. Stud.* 28 (3–4), 367–379. <https://doi.org/10.1080/10714410600873258>.
- Bargerhuff, M.E., Cole, D.J., Teeters, L., 2012. Using a transdisciplinary model to address inequities in field placements for teacher candidates with disabilities. *Int. J. Incl. Educ.* 16 (2), 185–206. <https://doi.org/10.1080/13603111003721098>.
- Bryant, L.G., Curtner-Smith, M.D., 2008. Impact of a physical education teacher’s disability on elementary pupils’ perceptions of effectiveness and learning. *Adapt. Phys. Act. Q.* 25 (2), 118–131. <https://doi.org/10.1123/apaq.25.2.118>.
- Bryant, L.G., Curtner-Smith, M.D., 2009a. Effect of a physical education teacher’s disability on high school pupils’ learning and perceptions of teacher competence. *Phys. Educ. Sport Pedagog.* 14 (3), 311–322. <https://doi.org/10.1080/17408980802225800>.
- Bryant, L.G., Curtner-Smith, M.D., 2009b. Influence of a physical education teacher’s disability on middle school pupils’ learning and perceptions of teacher competence. *Eur. Phys. Educ. Rev.* 15 (1), 5–19. <https://doi.org/10.1177/1356336X09105209>.
- Burns, E., Bell, S., 2010. Voices of teachers with dyslexia in Finnish and English further and higher educational settings. *Teach. Teach. Educ.* 16 (5), 529–543. <https://doi.org/10.1080/13540602.2010.507964>.
- Burns, E., Bell, S., 2011. Narrative construction of professional teacher identity of teachers with dyslexia. *Teach. Teach. Educ.* 27 (5), 952–960. <https://doi.org/10.1016/j.tate.2011.03.007>.
- Department of Social Development, 2016. Draft White Paper on the Mainstreaming of the Rights of Persons With Disabilities to Equality and Dignity (National Disability Rights Policy) South Africa. Retrieved from: http://uhambofoundation.org.za/new_wp/wp-content/uploads/2016/06/Nationaldisability-rights-policy_draft-white-paper.pdf.
- Duquette, C., 2000. Examining autobiographical influences on student teachers with disabilities. *Teach. Teach.* 6 (2), 215–228. <https://doi.org/10.1080/713698718>.
- Education Commission, 2019. Transforming the Education Workforce: Learning Teams for a Learning Generation. The Education Commission, New York. Retrieved from: <https://educationcommission.org/wp-content/uploads/2019/09/Transfor>.
- Ferri, B., Connor, D., Solis, S., Valle, J., Volpitta, D., 2005. Teachers with LD: ongoing negotiations with discourses of disability. *J. Learn. Disabil.* 38 (1), 62–78. <https://doi.org/10.1177/00222194050380010501>.
- Glazzard, J., Dale, K., 2015. “It takes me half a bottle of whisky to get through one of your assignments”: exploring one teacher educator’s personal experiences of dyslexia. *Dyslexia* 21 (2), 177–192. <https://doi.org/10.1002/dys.1493>.
- Griffiths, S., 2012. Being dyslexic doesn’t make me less of a teacher. School placement experiences of student teachers with dyslexia: strengths, challenges and a model for support. *J. Res. Spec. Educ. Needs* 12 (2), 54–65. <https://doi.org/10.1111/j.1471-3802.2011.01201.x>.
- Hankebo, T.A., 2018. Being a deaf and a teacher: exploring the experiences of deaf teachers in inclusive classrooms. *Int. J. Instr.* 11 (3), 477–490. <https://doi.org/10.12973/iji.2018.11333a>.
- Hayashi, R., May, G., 2011. The effect of exposure to a professor with a visible disability on students’ attitudes toward disabilities. *J. Soc. Work. Disabil. Rehabil.* 10 (1), 36–48. <https://doi.org/10.1080/1536710X.2011.546300>.
- Lamichhane, K., 2015. Disability, Education and Employment in Developing Countries: From Charity to Investment. Cambridge University Press, Cambridge.
- Lamichhane, K., 2016. Individuals with visual impairments teaching in Nepal’s mainstream schools: a model for inclusion. *Int. J. Incl. Educ.* 20 (1), 16–31. <https://doi.org/10.1080/13603116.2015.1073374>.
- Lee, M.N., Yen, A., Ching Mey, S., 2011. Employment of people with disabilities in Malaysia: drivers and inhibitors. *Int. J. Spec. Educ.* 26 (1), 112–124.

- Malik, R., Raza, F., Rose, P., Singal, N., 2020. Are children with disabilities in school and learning? Evidence from a household survey in rural Punjab, Pakistan. Compare. <https://doi.org/10.1080/03057925.2020.1749993>.
- Ministry of Education, 2016. School Sector Development Plan 2016/17-2022/3. Nepal. Retrieved from. https://moe.gov.np/assets/uploads/files/SSDP_Book_English_Final_July_5_2017.pdf.
- Ministry of Justice, 2007. Law No. 1 Relating to Protection of Disabled Persons in General. https://www.un.org/development/desa/disabilities/wp-content/uploads/sites/15/2019/11/Rwanda_Law-Relating-to-Protection-of-Disabled-Persons-in-General.pdf.
- Mizunoya, S., Mitra, S., 2013. Is there a disability gap in employment rates in developing countries? World Dev. 42, 28–43.
- Moore, A., Kern, V., Carlson, A., Vaccaro, A., Kimball, E.W., Abbott, J.A., Troiano, P.F., Newman, B.M., 2020. Constructing a sense of purpose and a professional teaching identity: experiences of teacher candidates with disabilities. Educ. Forum 84 (3), 272–285. <https://doi.org/10.1080/00131725.2020.1738608>.
- Morgan, E., Burn, E., 2000. Three perspectives on supporting a dyslexic trainee teacher. Innovat. Educ. Train. Int. 37 (2), 172–177. <https://doi.org/10.1080/13558000050034538>.
- Neca, P., Borges, M.L., Pinto, P.C., 2020. Teachers with disabilities: a literature review. Int. J. Incl. Educ. <https://doi.org/10.1080/13603116.2020.1776779>.
- Neuman, W.L., 2013. Social Research Methods: Qualitative and Quantitative Approaches. Pearsons, England.
- Office of the Deputy President, 1997. White Paper on the Integrated National Disability Strategy. South Africa. Retrieved from. https://www.gov.za/sites/default/files/gcis_document/201409/disability2.pdf.
- Parliament of the Republic of South Africa, 1998. Employment Equity Act, 1998. <https://www.ilo.org/dyn/natlex/docs/WEBTEXT/51169/65139/E98ZAF01.htm>. (Accessed 1 June 2022).
- Pritchard, G., 2010. Disabled people as culturally relevant teachers. J. Soc. Inclusion 1 (1), 43–51. <https://doi.org/10.36251/josi.4>.
- Roberson, J.L., Serwatka, T.S., 2000. Student perceptions and instructional effectiveness of deaf and hearing teachers. Am. Ann. Deaf 145 (3), 256–262. <https://doi.org/10.1353/aad.2012.0066>.
- Saltes, N., 2020. It's all about student accessibility. No one ever talks about teacher accessibility": examining ableist expectations in academia. Int. J. Incl. Educ. <https://doi.org/10.1080/13603116.2020.1712483>.
- Singal, N., 2019. Researching disability and education in Southern contexts: rigour, respect and responsibility. In: Singal, N., Lynch, P., Johansson, S.T. (Eds.), Education and Disability in the Global South. New Perspectives From Africa and Asia. Bloomsbury.
- Singal, N., Ware, H., 2021. English Language Teachers With Disabilities: An Exploratory Study across Four Countries. Report submitted to the British Council.
- Singal, N., Ware, H., Khanna-Bhutani, S., 2017. Inclusive Quality Education for Children with Disabilities. Report prepared for the World Innovation Summit For Education, Doha.
- Singal, N., Sabates, S., Aslam, M., Saeed, S., 2018. School enrolment and learning outcomes for children with disabilities: findings from a household survey in Pakistan. Int. J. Incl. Educ. Published online: 15 Oct <https://www.tandfonline.com/doi/abs/10.1080/13603116.2018.1531944?journalCode=tied20>.
- Stoveska, V., 2020. International Day of Persons With Disabilities. How Disability Affects Labour Market Outcomes? <https://lostat.ilo.org/international-day-of-persons-with-disabilities-how-disability-affects-labour-market-outcomes/>. (Accessed 10 February 2022).
- Swartz, L., 2014. Five challenges for disability related research in sub-Saharan Africa. African J. Disabil. 3 (2). <https://doi.org/10.4102/ajod.v3i2.149>, 6 pages.
- Tal-Alon, N., Shapira-Lishchinsky, O., 2020. Employing teachers with disabilities: a multifaceted prism of school principals' ethical dilemmas. Educ. Manag. Adm. Leader. 49 (6), 940–957.
- United Nations, 2006. Convention on the Rights of Persons With Disabilities. UN, Geneva.
- United Nations, 2016. Convention on the Rights of Persons With Disabilities. UN, Geneva.
- United Nations, 2020. The 2030 Agenda and the Sustainable Development Goals. United Nations, Paris.
- Valle, J.W., Solis, S., Volpitta, D., Connor, D.J., 2004. The disability closet: teachers with learning disabilities evaluate the risks and benefits of "coming out". Equity Excell. Educ. 37 (1), 4–17.
- Vogel, G., Sharoni, V., 2011. "My success as a teacher amazes me each and every day"—perspectives of teachers with learning disabilities. Int. J. Incl. Educ. 15 (5), 479–495. <https://doi.org/10.1080/13603110903131721>.
- Ware, H., Singal, N., Groce, N., 2021. The Work Lives of Disabled Teachers: Revisiting Inclusive Education in English Schools. Disability & Society. <https://doi.org/10.1080/09687599.2020.1867074>.
- Wood, R., Happé, F., 2021. What are the Views and Experiences of Autistic Teachers? Findings From an Online Survey in the UK Disability & Society. <https://doi.org/10.1080/09687599.2021.1916888>.
- World Health Organization, 2011. World Report on Disability. WHO, Geneva.

Using disability studies in education to move from individualized assessment to partners in pedagogy

Missy Morton^a, Jude MacArthur^a, and Anne-Marie McIlroy^b, ^a Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand; and ^b Te Pae Aronui Ministry of Education, Wellington, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	352
Context for this chapter	353
DSE as framework for rethinking research and teacher professional learning about assessment and pedagogy	354
DSE supports a focus in research on problems in teaching as professional dilemmas	354
DSE encourages an examination of the social and cultural context in which children learn	355
The problems of traditional individual and individualizing approaches to assessment	355
Partners in pedagogy: a narrative assessment case study	356
Anna and research	357
Anna and Ria	358
Anna and Ria's parents	359
Anna and school management	359
Conclusion: the working partnerships that enable success for students and teachers	360
Acknowledgments	360
References	360
Related websites	361

Glossary

Aotearoa Aotearoa is the Māori name for New Zealand

Māori Indigenous people of Aotearoa New Zealand

Narrative assessment a credit-based, narrative approach to assessment for disabled students that is accessible to teachers, families and students

Rangatahi – young person in Māori culture.

Te reo Māori Māori language

Whānau extended family group that includes physical, emotional and spiritual dimensions

Kura school, education, learning

Introduction

In this chapter we argue that student belonging and learning are enhanced when assessment in education is understood as the starting point for thinking about approaches to teaching and learning, and as a partnership between students, teachers and families. All three authors of this paper are involved in teaching and learning about a particular sociocultural approach to assessment, known in New Zealand as *narrative assessment*. Our wider interest is in equity and inclusion, and in collaborative research that supports the development of inclusive school communities. We draw on a two-year collaborative research project in two New Zealand secondary schools and a special school where researchers, teachers, students and families worked together to recognize, describe, report on and respond to disabled students' learning in affirmative and accessible ways.

Making good decisions about the future of our schools relies on a shared understanding about what inclusion means and why it is important. While recognizing that the terms inclusion and inclusive education are dynamic in their meaning and dependent on local context, we draw on a rights and evidence-based interpretation that also considers how teachers enact inclusive pedagogical practice (Ainscow and Messiou, 2018; Florian, 2014; Slee, 2018). Roger Slee's think piece prepared for the 2020 *Global Education Monitoring Report* asks what kind of society we want to live in, and poses three questions that are integral to the transformation of our education system toward inclusion:

What kind of world do we want our children and young people to live in?

What kind of schools and classrooms are required to achieve that world?

When we look at our schools – who is in, who is out, who decides and what are we going to do about it?

Slee (2018, p. 9).

We recognize that inclusive education is, as [Slee and Tait \(2022\)](#) note, “a rebellious challenge to oppression in and through education” (p. 2) that is contingent on time, place, political economy and culture and demands far-reaching transformation. Without wishing to lose this “insurrectionary zeal” these authors opt to apply guiding principles that position inclusive education as a process that, in practical terms, increases access to and ensures presence, participation and success (academic and social) for all children and young people in their neighborhood early childhood center and school. A central message is that every learner matters and matters equally ([United Nations Educational, Scientific and Cultural Organization, 2017](#)). Children’s rights to quality education in their local regular school are guaranteed, and to uphold these rights neighborhood schools take a principled approach, building their capacity to teach all students, eliminating barriers to belonging and success, and enacting inclusive values such as equity, participation, rights, respect for diversity, community and sustainability ([Ainscow and Messiou, 2018](#); [MacArthur and Rutherford, 2016](#); [Rutherford, 2016](#); [Slee, 2018](#)).

The *Using narrative assessment to support secondary school teachers’ inclusive practices project*, used a Disability Studies in Education (DSE) framework to explore the impact of professional development about narrative assessment on teachers’ pedagogy and practice, and on students’ learning and social experiences. This chapter illustrates how DSE can support research based on a curiosity about the dilemmas teachers face in the process of teaching and learning, as opposed to locating learning challenges as problems or deficits within students that require “fixing”. We have used a DSE and sociocultural framework to examine the social and cultural context in which children learn in secondary school settings, and to pay attention to how teachers and families can collaborate to understand students, their capability, their learning and how that learning is assessed.

Drawing on the experiences of one school in our research we turn attention to the importance of starting with assessment and show how assessment capable teachers and students can drive quality teaching and learning. Collaboration within the school community is an essential element in this process, with narrative assessment shifting thinking from assessment as focused on the individual student to assessment as partnership among teachers, students and families that takes place across multiple contexts in the life of the student.

Context for this chapter

The research on narrative assessment in secondary schools drawn on in this chapter was carried out in Aotearoa New Zealand (Aotearoa). Education is compulsory from ages 6 to 16 and most children, including disabled children, attend their local, public school, funded by the government, where attendance is free. At secondary level students may be able to choose between single sex or coeducational public schools, secondary schools in rural areas are more likely to be co-educational and serve a wide geographical area. Most secondary schools are state schools, about 10% of students attend “integrated” state schools which are usually designated as having a special character, such as a particular religious affiliation, or education approach. Less than five percent of children attend private schools that are not government funded, fees are charged for children’s tuition and private schools do not have to follow *The New Zealand Curriculum*. Some students also receive support from The Correspondence School, Te Aho o Te Kura Pounamu, which uses multimedia and online learning. Twenty-seven separate day special schools are designed for disabled students from Years 1–13 who are thought to have “a high level of need”, students can enroll in these schools under a Specialist Education Agreement.

In Māori-medium schools some students are taught curriculum subjects in Te Reo Māori for at least 51% of the time. Kura Kaupapa Māori schools are a unique New Zealand primary school education system in which children are immersed in a Māori language and cultural environment with the aim of enabling graduates to “live as Māori”. They are state schools where teaching is in te reo Māori and is based on Māori culture and values, they follow the curriculum for Māori-medium teaching, learning and assessment, Te Marautanga o Aotearoa, and cater for students from years 1 to 8 or years 1 to 13 ([Ministry of Education, 2017](#)).

A key policy document that guides teachers’ and schools’ practices is the *New Zealand Curriculum* ([Ministry of Education, 2007](#)), which covers subjects that are taught at primary and secondary levels and specifies the standards students should reach by subject. The document clearly states that it is for all students; participation, learning, achieving and belonging for all students are central concepts. The purpose of assessment is described as “... improving students’ learning and teachers’ teaching as both student and teacher respond to the information that it provides” ([Ministry of Education, 2007](#), p. 39). However for disabled students who may be considered to be learning long-term within Level 1 of the curriculum, assessment has often been associated with a diagnosis or application for resources. Many teachers, in both local and special schools, have considered the *New Zealand Curriculum* to be irrelevant for some students ([Guerin, 2015](#); [Morton et al., 2012](#)) and may therefore struggle to see themselves as teachers of all students or to see all students as capable learners ([Florian, 2014](#); [McIlroy, 2017](#)).

The National Certificate in Educational Achievement (NCEA) is the national senior secondary school qualification which assesses student achievement in the last three years at school. It is a standardized process which means students can move between schools and regions and continue working toward their NCEA qualifications. NCEA is recognized by employers, and is used for selection by universities and polytechnics, both in New Zealand and overseas. We have noted in an earlier publication, however, that NCEA also creates structures that determine how schools operate, some of which may exclude disabled students (Morton et al., 2021). NCEA measures specific skills (in literacy and numeracy, for example) that may not reflect the learning and strengths of some students. We have posed a number of questions about assessment for all, noting that school should be a place where students can recognize their capability, believe their dreams can be achieved, and have the skills and motivation to continue learning beyond school. When students feel that their learning is neither recognized nor valued, their identity and confidence as a learner is at risk, and on that basis we have asked:

So how can schools respect student diversity when the assessment system values only one narrow interpretation of excellence? What does this mean for students who may not achieve at the standard required to 'succeed' using this system? How do we recognise achievement for some disabled students whose learning and progress are invisible within NCEA school achievement data? How do we respect diversity in teachers and students when assessment benchmarks determine teaching content?

Morton et al. (2021, p. 10).

It is within this space that we undertook the *Using narrative assessment to support secondary school teachers' inclusive practices* project. Narrative assessment is a credit-based narrative approach that documents learning for disabled students who are considered to be working long-term in Level 1 of the *New Zealand Curriculum*. It uses a combination of photographs and/or videos of students' learning experiences alongside narrative descriptions that connect with the curriculum's key competencies and learning areas. Importantly, learning is described in affirmative ways that are accessible to teachers, families and students to ultimately enhance student belonging and learning at school, home and in the wider community (Ministry of Education, 2009a,b).

DSE as framework for rethinking research and teacher professional learning about assessment and pedagogy

The *Using narrative assessment to support secondary school teachers' inclusive practices* project began with questions from teachers and principals about how to recognize and report on the learning of disabled secondary school students. The project was funded through the Teaching and Learning Research Initiative (TLRI) and explored the use of narrative assessment to support secondary teachers' inclusive practices. Funding for TLRI projects is conditional on a model of partnership between researchers and educators (<http://www.tlri.org.nz/tlri-research/research-progress/school-sector/using-narrative-assessment-support-secondary-school>).

The project built on an earlier research and professional learning project funded by the New Zealand Ministry of Education which introduced narrative assessment to a group of teachers, parents and academic staff in a collaborative learning community (Morton et al., 2012). This earlier research sparked our interest in making greater use of DSE to explain shifts in understandings and practices and to focus on the relationships between cultures of learning and cultures of belonging (MacArthur et al., 2018; Morton, 2014; Morton et al., 2019).

DSE supports a focus in research on problems in teaching as professional dilemmas

Teacher education and professional learning is central to the project of supporting education communities to develop inclusively, challenging those working in initial teacher education and ongoing teacher professional learning to think carefully about the nature and form of such learning. What it means to deliver high quality inclusive education training in teacher education for teachers either entering or already working in the profession remains unclear (Florian and Camedda, 2020), although studies from a range of national contexts highlight that while teachers are generally committed to the values and principles of inclusive education, they are often anxious about how to work with an increasingly diverse range of learners, believing they lack the necessary "specialist" skills and knowledge (Black-Hawkins, 2017).

Black-Hawkins argues that such views can undermine teachers' professional confidence while also casting doubt on traditional pedagogical approaches to inclusion that are based on planning and teaching for "most" learners, while providing something extra or different for "some". The "some" are often those students who have been labeled and categorized as "special needs", a construct that supports a deficit model of learning by locating the "problem" in the learner (who requires "fixing"), rather than encouraging an exploration of the learning context (Florian and Black-Hawkins, 2011; MacArthur and Rutherford, 2016; Rutherford, 2016; Rutherford and MacArthur, 2018). DSE scholars have drawn attention to the pathologizing effects of such practices on students and their families (Paugh and Dudley-Marling, 2011; Valle and Connor, 2011a).

Colleagues in Aotearoa have described the same deficit pathologizing aimed at Māori (indigenous peoples of Aotearoa New Zealand) students and their whānau (extended family in Māori culture) who have been blamed for their perceived failure within the education system (Morton, 2015). These authors argue for a reconnection of culture, identity, values and beliefs to our educational practices and argue that *relationships* are central to all students being able to participate and flourish in classrooms and schools (Berlyman and Woller, 2013; Bishop et al., 2003; Bishop and Glynn, 1999; Macfarlane et al., 2012). Morton (2015) invites us to

consider that relationships between teachers, students and students' whānau are only possible when teachers work at knowing their students and poses the question: how do our current beliefs and practices in curriculum, pedagogy and assessment support or get in the way of forming meaningful relationships with students?

Within the context of these relationships, Florian and Black-Hawkins' (2011) work on inclusive pedagogies further encourages teachers to view difficulties in learning not as deficits within children but rather as dilemmas and professional challenges for themselves as teachers (Black-Hawkins, 2017; Rouse and Florian, 2012). Their concept of transformability is a reminder that teachers' relationships with children and young people can be built on an understanding that capacity to learn can change and be changed for the better as a result of what teachers do in the present. Transformability provides a tangible way for teachers to "recognize their capacity to teach all learners" (Rouse and Florian, 2012, p. 17).

In reviewing research on the development of inclusive classroom practices, Black-Hawkins (2017) suggests that to be useful for teachers, research needs to take account of the complex interactions intrinsic to daily classroom teaching and learning. There are implications for researchers to also engage in learning partnerships, learning with and from teachers to understand inclusive pedagogical approaches that are meaningful to teachers. A DSE perspective sits comfortably with this position because it emphasizes making sense of constructions of disability and competence in education, including in teaching practice (Macartney and Morton, 2013; McIlroy and Guerin, 2014; Morton et al., 2021, 2012). DSE provides a valuable framework for research that is based on a curiosity about the dilemmas teachers face in the process of teaching and learning. It heightens alertness to the barriers to participation and learning faced by children and young people, including approaches that locate learning challenges as problems/deficits within students that make it difficult for whānau, teachers and others to effectively nurture children's learning journeys (Guerin and Morton, 2015; Macartney and Morton, 2013; Skidmore, 2002). The present project was therefore framed as a shared inquiry into a shared concern (Grudnoff et al., 2019), with DSE underpinning a commitment to value all voices and ensure all project members contributed ideas that shaped the research direction (Goodley, 2013; Valle and Connor, 2011a).

DSE encourages an examination of the social and cultural context in which children learn

DSE also recognizes that the social construction of disability as a deficit or problem within the child or young person can be challenged by sociocultural forms of teaching, learning, and assessment (Macartney, 2011; Macartney and Morton, 2013; McIlroy and Guerin, 2014; Morton and McMenamin, 2011; Morton et al., 2012). Narrative assessment, the focus of our research with secondary school students, their teachers and whānau, is one such example. Sociocultural understandings about children and young people's learning turn attention to the social and cultural context in which learning takes place, and to the relationships and interactions between people that shape our knowledge of the world and our common ways of understanding it (Macfarlane, 2015). Teachers are part of this context, how they understand students and their capability, as well as the opportunities they create for students to show their understanding and competence and for how that learning is assessed (Guerin, 2015; Guerin and Morton, 2015; Valle and Connor, 2011b; Wansart, 1995). Black-Hawkins (2017) reflects on this point:

What children 'are' and how they are 'different' from each other cannot be read off from a list of characteristics. Instead, it emerges from the interactions amongst children, and between them and their teachers, as they work together in particular educational contexts, on particular tasks and priorities (p. 16).

If schools are to become more responsive to diversity, the curriculum and approaches to assessment, teaching and learning need to be inclusive of and relevant to the lives of all students, brought to life over time in the relationships that are built up between the teacher and a group of students. Both DSE and sociocultural theory encourage an affirmative orientation by teachers that views the starting point as the present state of students' knowledge and understanding, their culture, capability and interests. Teaching builds a bridge between what the student knows how to do independently and the new level of understanding and proficiency that lies within their reach to achieve with the support of the teacher and peers. Macfarlane (2015) describes this process as a dynamic and interactive social and learning relationship, not one of fixed dependency. Inclusive approaches to teaching recognize, reflect and celebrate the human diversity found in school communities, supporting shared learning experiences and the joint production of knowledge. As we have worked in partnership with teachers of disabled secondary school students, a central aim has been to broaden the range of modes of student response that are sanctioned within the curriculum and assessed as valid learning. In this way, the rich range of knowledges and ways of being that children and young people bring to school are valued as assets and strengths to be built upon (Morton et al., 2021).

The problems of traditional individual and individualizing approaches to assessment

The history of assessment in relation to disabled people has focused on identification, diagnosis, intervention, allocation of resources and placements for services, including institutionalization. Within DSE, Paugh and Dudley-Marling (2011) have described how the "unrelenting focus" on the individual continually reinvents and reinscribes deficit views of individual students, at the same time pathologizing both students and their families.

In the UK project *Learning without Limits* Swann et al. (2012) observed an education climate where teachers were required to use the certainty of prediction as a tool in their planning and teaching (i.e., what do we expect this child should be able to/will learn?). The curriculum specified targets, levels, objectives and outcomes that meant teachers were *required to behave as if all children's potential is predictable and their futures are knowable well in advance*. Teachers felt that what they did in the classroom would have only limited impact on the lives of the children and young people they were teaching. Closely associated with this view of learning as linear, measurable and quantifiable, was "... an equally damaging view of the children who do the learning who can themselves be known, measured, and quantified in terms of so-called fixed ability, a fixed, internal capacity, which can be readily determined" (Swann et al., 2012, p. 1).

Curriculum, pedagogy and assessment are intricately linked. Assessment impacts on teaching, and Guerin and Morton (2015) argue that curriculum and pedagogy can only be transformed by also transforming the way learning is assessed. The way teachers understand and assess student learning shapes the way they think about and carry out the act of teaching, when the scenario is overwhelmingly negative, the impact is not only on the child or young person's learning, it is also on their friendships and their sense of identity. In other words, if I do not see you as a learner, I cannot see myself as your teacher.

Guerin and Morton (2015) describe a similar climate in Aotearoa to that observed by Swann et al. (2012) in the UK, in which teachers can come to hold narrow views about the potential of some children and young people to learn:

They may question how some children and young people can possibly participate in a curriculum that sets out a sequence of levels of learning objectives- a sequence that makes it look as if it is only possible to participate in the 'upper levels' having demonstrated competence in previous levels. Some teachers may also doubt their ability to be teachers in relation to children and young people who are not recognised as learners (p. 72).

The consequences of this approach can be highly detrimental for the student and their whānau who are much more likely to view their family member as capable, and has having strengths and interests that are celebrated at home and in the community.

In New Zealand, ground breaking work in early childhood curriculum by Margaret Carr and colleagues has directly challenged traditional psychological and developmental ways of framing individual learning (Carr, 2009; Carr and Lee, 2012; Cowie, 2009; Cowie and Carr, 2009). Turning to sociocultural perspectives on learning moves the focus from the individual to the individual in context, to learning as a social activity. Understanding early childhood centers as learning communities frames all members as learners, including the adults in the settings. The narrative assessment projects have similarly framed classrooms and schools as learning communities, with all members as learners.

DSE has also turned our attention to the importance of seeing assessment as a critical starting point in teaching, in the sense that it is assessment capable teachers and students who drive quality teaching and learning. Equally importantly, assessment is understood as taking place not just at school, but across the many contexts that a student participates in including their home and the wider community. Sociocultural forms of assessment privilege interactions within and between social and cultural contexts, rejecting fixed notions of medical abnormality or deficit that traditional assessment approaches are based on (Biklen and Kliever, 2006; Macartney, 2011; Morton and Guerin, 2017). In this view, impairment is no longer regarded as a singular or fixed identity, a permanent measure or predictor of capacity.

Guerin and Morton (2015) reflect on accounts from families describing their child as doing something that they believe should be celebrated, but their recounts are dismissed as not providing evidence of learning or accomplishment. They ask the reader to consider the child's and family's sense of belonging at school when learning is viewed in narrow, prescriptive ways, arguing instead that responsive and reciprocal relationship among teachers, students and whānau are central to all students being able to participate, show their learning and flourish at school. This requires that teachers work at knowing their students, a point that is illustrated in the following narrative assessment case study.

Partners in pedagogy: a narrative assessment case study

This case study comes from the *Using narrative assessment to support secondary school teachers' inclusive practices project* and illustrates how student belonging and learning are supported when assessment in education is a *partnership* between students, teachers and families, and is the *starting point* for thinking about approaches to teaching and learning. Names of people and places in this case study are pseudonyms.

At the center of this case study is Anna, a teacher in a day special school; Ria, a student aged 15 years and her whānau, and the school principal. The research this case study was part of included two local secondary schools and a special school. We have chosen this example from the special school for two reasons. Firstly, this is a school that wanted to know how they could use the *New Zealand Curriculum* to guide and document their students' learning in affirmative and meaningful ways, and how narrative assessment might more broadly inform their teaching pedagogy. Secondly, the example illustrates the process of teachers coming to know Ria, of becoming assessment-capable partners with Ria, her whānau and other teachers. The context could equally be a local secondary school, so we can also consider whether and how Ria's capability as a poet and writer might be recognized and celebrated, her learning and belonging fostered in that setting.

Anna is focused on developing her teaching pedagogy so the capabilities of all rangatahi (term meaning young people in Māori culture) in her class are visible and enhanced. Ria is a rangatahi in Anna's class. The partnerships, though interconnected, are considered separately as follows:

- Anna and research
- Anna and Ria
- Anna and Ria's whānau
- Anna and school management

Anna and research

The motivation for Anna to join the narrative assessment research project was a challenge in her teaching practice that she really wanted to grapple with, and the research provided an opportunity for collective reflection and shared learning. Anna recognized that, for some disabled rangatahi, learning progress may take time, and achievement may be demonstrated in unexpected contexts and in surprising ways. Anna recognized the learning in her class but did not think she fully captured the complexity of the interactions between teaching and learning in assessments.

All the rangatahi in Anna's class have personalized sophisticated technological support to communicate their thinking and to enable them to make best use of the senses. For example, some have cortical vision impairment and need materials presented in very specific ways so they can process the visual stimuli they experience. All of the rangatahi in Anna's class can only access the *New Zealand Curriculum* with high levels of support. This does not mean they are not capable. Some people believe that children who need high levels of support cannot learn and some educators lack confidence to teach all rangatahi. These rangatahi may then be excluded from the rich curriculum available to their non-disabled peers. When rangatahi are not taught curriculum content, standardized curriculum assessment tools can become irrelevant. This cycle is difficult to break, for assessment drives teaching and without good knowledge of children's skills and capabilities, planning a meaningful learning program is problematic.

Often school goals set for disabled rangatahi with high and complex needs are based on key competencies, not on the learning areas of the Curriculum. Key competencies in the *New Zealand Curriculum* include thinking; participating and contributing; using language symbols and text; belonging; and managing self. These are not intended as stand-alone teaching goals, but are integral to learning in the learning areas (such as, literacy; science; physical education and health; numeracy) (Hipkins, 2009). For example, a child may have a goal related to how long they spend in a science lesson, but not a goal about the scientific content that connects the child to teaching subject content.

Anna recognized the significant progress her rangatahi were making in the learning areas; however standardized assessment processes such as NCEA or e-asTTle (an online assessment tool developed to assess students' achievement and progress in reading, mathematics, writing, which begins at level 2 of the *New Zealand Curriculum*) were not accessible for the rangatahi in her class. Nor did they capture the learning demonstrated. Anna noticed a disconnect between the Curriculum vision and some assessment processes. She recognized that the complexity of teaching did not always link to planned objectives and assessments. Sometimes learning objectives require multiple experiences over time to be achieved. Sometimes very small steps toward a larger goal are achieved, and these steps may not be easily recognized. Sometimes the planned lesson goes in a different direction and very significant unplanned learning achievements are noticed but not always validated.

Anna had begun to explore the narrative approach to assessment in her postgraduate studies. Developing a deeper understanding of assessment theory is important to her as she develops and implements her teaching philosophy and feels confident to take risks, be creative and implement change in her classroom. Anna's connection to research made her an early adopter of narrative assessment theory and a thoughtful practitioner as she applied her learning to her classroom practice. A commitment to study also supported Anna's motivation to critically reflect on her teaching and constantly refine how she recognized progress and achievement (Grudnoff et al., 2019).

Anna's teaching philosophy is based on democratic principles where all rangatahi are equally valued and are partners in learning. Anna said:

I run my classroom in a democratic way. It's student driven based on students' rights; and the students plan their own learning journey. My focus is on their rights to be heard, to play, to socialise, to be educated and to be respected.

She describes herself as a facilitator of learning. Her classroom program is built around the interests and strengths of the rangatahi, which are interwoven in the learning areas of the *New Zealand Curriculum*. In Anna's classroom, teaching, planning and assessment are collaborative and inclusive of all rangatahi (Alton-Lee, 2003). For example, the class had all indicated music as an enjoyable subject. Anna gave each class member the opportunity to choose a favorite piece of music or a song, which was shared with the rest of the class. This meant every student experienced the variety of music selected, meaning all choices were validated and everyone had the opportunity to learn something about their classmates.

Anna and Ria

Anna's democratic approach to teaching means she partners with her rangatahi to learn with and give value to their strengths, interests, and perspectives about their learning. The partnership focused on here is between Anna and Ria. At the time of the research project, Ria is beginning her journey through secondary school and Anna has been her classroom teacher for 2 years. Ria uses a wheelchair for mobility. She communicates in ways other than speech and uses a PODD (Pragmatic Organization Dynamic Display), a high tech communication tool based on using visual discrimination to communicate a message. Additionally, she uses head movements to indicate yes or no and facial expression to communicate emotion.

Anna shared Ria's whānau's belief that, "Ria has much to say, has firm opinions and is keen to share her thoughts both at home and at school". She recognized that her role as teacher was to ensure Ria had effective, accessible communication systems to enable her voice to be heard. Without a deep knowledge of Ria and a genuine connection with her, Anna felt unable to fully respond to Ria's motivations and best enable successful learning as a valued class member belonging in the *New Zealand Curriculum*. The curriculum's vision is for "young people who will be competent, connected, actively involved, lifelong learners" (p. 7), and this vision guides all Anna's pedagogical decision-making.

An example of partnership in learning that built on Anna's deep knowledge of Ria is visible in physical education. Ria has never been an enthusiastic participant in this learning area and had frequently expressed her opposition to participate. However, Ria is a Harry Potter fan, and indicated to Anna that she wanted to play quidditch. Ria helped write the class "Quidditch Game Rules", and with support from a broader team the context was created so all class members could play the game. Participation was enabled by ramps so the wheelchairs could navigate the quidditch court; balls that made a noise so rangatahi with a cortical vision impairment could see the ball and large clear targets. Anna reported that Ria's face "lit up" when she saw the quidditch court and that she was totally involved in all aspects of the game which turned out to be a fierce competition. Ria's participation and learning in physical education was presented in a narrative assessment. The narrative provided the approach to share the learning outcomes of the Quidditch Game, and made visible the collaborative planning, Ria's voice, the ongoing interactions, the successes, and Ria's reflections. At home after the game, Ria used her PODD to tell her whānau that she "loved quidditch" and the game was "exciting" and "fun".

Ria's skills as a poetry writer were also revealed through partnership and recognized through narrative assessment. During an English lesson, all rangatahi were enabled to express their thoughts in writing. The focus was creative writing about a topic or person of interest. Anna holds high expectations of all her rangatahi, believing that with support creative writing is achievable. Using a range of communication supports, Ria indicated she wanted to write a poem. Ria selected a title, the vocabulary she wished to use, the order and structure of the poem, and how it was to be presented. This did not happen in one lesson but was completed with perseverance over time, supported by Anna. On completion, the poem was printed in large font on a background that enabled Ria to see the text. Ria asked that the poem be mounted on purple card and showed great pleasure in seeing the published poem she created and later shared with her whānau.

The narrative approach to assessment makes visible the relationships that support learning (Guerin, 2015; McIlroy, 2017). The personal connections mean teachers notice the interactions and small successes that grow to build a picture of capability and achievement. Anna said:

Summative assessment wasn't showing capability and wasn't showing the small successes and the big successes my class were making. Most of our class learn for a long time in level 1 (of the *New Zealand Curriculum*) and those little successes build up to become huge successes and we can show that through narrative assessment.

She described narrative assessment as a process that highlighted relationships in learning and supported her reflexivity, noting that "narrative assessment enabled me to delve deep and reflect on the learning. It's really important (that) we always presume competence and it's our job as teachers to make this visible within the NZC." She described the *New Zealand Curriculum* as "relevant for every single student because it's not prescriptive and narrative assessments are a beautiful way to capture and share learning and the curriculum".

Anna described the process of writing about the connections, the partnerships, and the progress as a joy, "Narrative assessment helps teachers to see students through a strengths-based lens, you can't see any deficits when you write narrative assessment. Their personality shines through and their strengths shine through." Knowing rangatahi strengths and building on supportive partnerships help to create a culture of care and kindness that can support joy in teaching and learning, and joy in assessing the learning and creating future learning goals (Monchinski, 2010; Noddings, 2012; Nelson, 2013). According to Anna:

Narrative assessment ... shows everyone else and Ria that we value her. Through Ria's narrative [assessment] we see her personality shine through. We can see that she is kind, empathetic, funny, generous, and clever, that she loves literacy, hates PE, and isn't a fan of maths but will give it a go, and we see it all in her narrative assessments. Narrative assessment helps teachers to see students through a strength-based lens, you can't see any deficits when you write narrative assessment. The personality shines through and the strengths shine through ... All the wonderfulness that is Ria is in her narrative assessments.

Anna and Ria's parents

The process of creating the poem was recorded in a narrative assessment and shared on Story Park—the school's electronic platform used to connect teachers, rangatahi and whānau. Ria's Mum and Dad read the narrative assessment on Story Park and said:

We didn't know she was into writing at all, particularly creative writing like poetry. We had no idea till we saw the narrative assessment and now we do writing with her at home. So we've now got her an email address and a Facebook page.

This sharing enabled Ria's whānau to support Ria's learning at home, an important element in creating meaningful connections in Ria's learning across contexts.

Ria cannot easily share thoughts about her school day. Ria's mum said of the poetry writing narrative assessment, "We can see she is valued, how she gets her own voice and her own independence." Her whānau welcomed narrative assessments that show Ria's engagement and progress, her connection with others and her capability, all made visible through teacher knowledge and skill that enables learning. Assessments that include photographs and videos capturing learning throughout the school day provide evidence their daughter is happily engaged, is recognized as a capable learner and is a welcomed member of the school community. They also describe how important it is that Ria can see her whānau enjoying the assessments, this is a tangible way to connect with her school day. Her mum described Ria as:

... so proud that we get to see a glimpse into her school world, and we get to see and extend her interests at home – merging both of her worlds together
... The connection between home and school has opened up Ria's world.

For Anna, the relationship with family means she can better understand the contexts that support wellbeing at home and use this knowledge to develop a successful school learning environment. Anna valued that families become active learning partners, drawing on an example where Ria expressed an interest in gardening at school when the class created a class garden, an experience that was shared with families. Ria's family together created a wheelchair accessible sensory garden at home for Ria to enjoy, and her engagement in this process was shared between home and school.

Anna and school management

As a classroom teacher, Anna's relationship with school management is critical to enacting a shared vision of success. Being listened to, supported, challenged, trusted, and respected all contribute to a school culture that she makes evident in her classroom. Several teachers at Anna's school joined the research project alongside Leanne, the deputy principal. The relationships among staff, between staff, and with the research team created a platform that welcomed innovative assessment processes to enhance teaching and learning. Leanne's participation and support of the research project was critical to embedding new professional learning in the school. Leanne was already on a journey to develop assessment and, as a manager, was able to actively support change, both valuing and promoting formative approaches to assessment in her school community (Ministry of Education, 2011; Wiliam, 2011). She described how the school had changed the way they communicated learning with families when a disconnect between learning and assessment was recognized:

We stopped writing reports two years ago because it was a summative report, but we don't really do summative here because the students are always moving, what's next, what's next? Why would we write something summative when we haven't stopped?

Leanne was committed to partnerships with families and making visible the children's success through the *New Zealand Curriculum*, " ... we're a school, so the NZC is embedded in everything we do. If our children aren't embedded in the Curriculum, then we wouldn't be doing our job". She welcomed narrative assessment as a process that valued and commented on relationships:

Our narrative assessments start with a picture of learning, the tools used to support progress, what's happening, how it's happening, the context, curriculum links, including the child, together setting the next learning steps, and then sharing it with the family so they can give us feedback. Often feedback from parents really helps us to fine tune our teaching. We all get to see all the smaller steps ending up in a big achievement. Parents will say 'my child is a learner in the *New Zealand Curriculum*' - 'look my child's a learner'.

She appreciated assessment that enabled teachers to show that learning wasn't linear and could be unpredictable; that rangatahi progress in their own way; that progress looks different for every rangatahi, reinforcing the importance of teachers knowing rangitahi and their whānau. And that is something the school wants to celebrate.

Conclusion: the working partnerships that enable success for students and teachers

To better understand their students, teachers need to also understand all the ways, all the places, the various contexts in which students demonstrate what they know and can do. An understanding of learning as being stretched across people, places and things; with an understanding of competence as being context dependent, of necessity means that people must work together to build shared meanings of competence and learning. Students and their families are expert-partners in the assessment practices that support teachers to plan for, recognize and value learning.

Disability studies combined with a sociocultural understanding of assessment, teaching and learning has effectively provided the framework we have needed to highlight and build upon relationships that foster a better understanding of disabled students' learning. This approach has supported us to collaborate with teachers to explore how every student's learning can be noticed, recognized and responded to in ways that support student learning and more broadly foster cultures of belonging as schools work toward inclusion. Anna, Leanne, Ria, and Ria's whānau's story is one example of how assessment-capable teachers, students and whānau can work together to drive quality teaching and learning, deliberately adopting a standpoint that every student is capable, able to learn through the *New Zealand Curriculum*, and as Guerin (2015) has emphasized, that space can be made for the strengths and knowledge evident in observed interactions to be the focus of the assessment process.

Acknowledgments

We extend our thanks to the New Zealand Council for Educational Research's Teaching and Learning Research Initiative (TLRI) which provided financial support for this research. <http://www.tlri.org.nz/tlri-research/research-progress/school-sector/using-narrative-assessment-support-secondary-school>.

We extend our gratitude and thanks to the rangatahi, their teachers and whānau who worked with us to explore narrative assessment and teaching pedagogy so rangatahi can flourish at school.

References

- Ainscow, M., Messiou, K., 2018. Engaging with the views of students to promote inclusion in education. *J. Educ. Change* 19, 1–17.
- Alton-Lee, A., 2003. Quality Teaching for Diverse Students in Schooling: Best Evidence Synthesis. Ministry of Education, Wellington.
- Berryman, M., Woller, P., 2013. Learning about inclusion by listening to Maori. *Int. J. Incl. Educ.* 17 (8), 827–838.
- Biklen, D., Kliever, C., 2006. Constructing competence: autism, voice and the 'disordered' body. *Int. J. Incl. Educ.* 10, 169–188.
- Bishop, R., Berryman, M., Tiakiwai, S., Richardson, C., 2003. Te Kotahitanga: Experiences of year 9 and year 10 Māori students in mainstream classrooms: Final report to the Ministry of Education. Ministry of Education, Wellington, NZ.
- Bishop, R., Glynn, T., 1999. Culture Counts: Changing Power Relations in Education. Dunmore Press, Palmerston North, NZ.
- Black-Hawkins, K., 2017. Understanding inclusive pedagogy: learning with and from teachers. In: PLoWS, V., Whitburn, B. (Eds.), *Inclusive Education: Making Sense of Everyday Practice*. Brill, Rotterdam, pp. 13–28.
- Carr, M., 2009. Kei tua o te pae: assessing learning that reaches beyond the self and beyond the horizon. *Assess. Matters* 1 (1), 20–46.
- Carr, M., Lee, W., 2012. *Learning Stories: Constructing Learning Identities in Early Education*. Sage, London.
- Cowie, B., 2009. Teacher formative assessment decision making: a consideration of principles and consequences. *Assess. Matters* 1 (1), 47–63.
- Cowie, B., Carr, M., 2009. The consequences of socio-cultural assessment. In: Anning, A., Cullen, J., Fleer, M. (Eds.), *Early Childhood Education: Society and Culture*, second ed. Sage, London, pp. 105–116.
- Florian, L., 2014. What counts as evidence of inclusive practice? *Eur. J. Spec. Needs Educ.* 29, 286–294.
- Florian, L., Camedda, D., 2020. Enhancing teacher education for inclusion. *Eur. J. Teach. Educ.* 43, 4–8.
- Florian, L., Black-Hawkins, K., 2011. Exploring inclusive pedagogy. *Br. Educ. Res. J.* 37, 813–828.
- Goodley, D., 2013. Dis/entangling critical disability studies. *Disabil. Soc.* 28, 631–644.
- Grudnoff, L., Ell, F., Haigh, M., Hill, M., Tocker, K., 2019. *Enhancing Equity Through Inquiry*. NZCER Press, Wellington, New Zealand.
- Guerin, A., 2015. 'The Inside View' Investigating the Use of Narrative Assessment to Support Student Identity, Wellbeing, and Participation in Learning in a New Zealand Secondary School. Unpublished Doctoral Dissertation. University of Canterbury, Christchurch.
- Guerin, A., Morton, M., 2015. Sociocultural perspectives on assessment: implications and opportunities for students with special education needs. In: Macfarlane, A., Macfarlane, S., Webber, M. (Eds.), *Sociocultural Realities: Exploring New Horizons*. Canterbury University Press, Christchurch, New Zealand, pp. 68–86.
- Hipkins, R., 2009. Determining meaning for key competencies via assessment practices. *Assess. Matters* 1, 4–19.
- MacArthur, J., McIlroy, A.-M., Howard, T., 2018. What made school so good? Student perspectives on school informing inclusive practice. In: Berman, J., MacArthur, J. (Eds.), *Student Perspectives on School – Informing Inclusive Education*. Brill Sense Publishers, Leiden, Boston, pp. 111–127.
- MacArthur, J., Rutherford, G., 2016. Success for ALL? re-envisioning New Zealand schools and classrooms as places where 'Rights' replace 'Special'. *New Zealand J. Educ. Stud.* 51, 15–74.
- Macartney, B., 2011. Disabled by the Discourse: Two Families' Narratives of Inclusion, Exclusion and Resistance in Education. A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy. University of Canterbury, Christchurch, New Zealand.
- Macfarlane, A., 2015. Sociocultural foundations. In: Macfarlane, A., Macfarlane, S., Webber, M. (Eds.), *Sociocultural Realities: Exploring New Horizons*. Canterbury University Press, Christchurch, New Zealand, pp. 19–35.
- Macartney, B., Morton, M., 2013. Kinds of participation: teacher and special education perceptions and practices of 'inclusion' in early childhood and primary school settings. *Int. J. Incl. Educ.* 17, 776–792.
- Macfarlane, A., Macfarlane, S., Savage, S., Glynn, T., 2012. Inclusive education and Māori communities in Aotearoa New Zealand. In: *Teaching in Inclusive School Communities*. John Wiley & Sons, Milton, Queensland, pp. 163–185.
- McIlroy, A.-M., 2017. The Myth of Inability - Exploring Children's Capability and Belonging at Primary School Through Narrative Assessment. A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy. University of Canterbury, Christchurch, New Zealand.

- McIlroy, A.-M., Guerin, A., 2014. Flying under the radar: democratic approaches to teaching in neoliberal times. In: Wills, R., Morton, M., MacLean, M., Stephenson, M., Slee, R. (Eds.), *Tales from School: Learning Disability and State Education after Administrative Reform*. Sense Publishers, Rotterdam, pp. 213–226.
- Ministry of Education, 2007. *The New Zealand Curriculum*. Learning Media, Wellington.
- Ministry of Education, 2009a. *Narrative Assessment: A Guide for Teachers*. Learning Media, Wellington, New Zealand.
- Ministry of Education, 2009b. *Through different eyes*. <http://www.throughdifferenteyes.org.nz/home2>.
- Ministry of Education, 2011. *Ministry of Education Position Paper: Assessment*. Learning Media, Wellington.
- Ministry of Education, 2017. *Te Marautanga o Aotearoa*. Ministry of Education, Wellington, New Zealand. <https://tmoa.tki.org.nz/Te-Marautanga-o-Aotearoa>.
- Ministry of Education, 2019a. *Shaping a Stronger Education System With New Zealanders*. <https://conversation.education.govt.nz/conversations/shaping-a-stronger-education-system-for-all/>.
- Ministry of Education, 2019b. *The Learning Support Action Plan*. Ministry of Education, Wellington.
- Monchinski, T., 2010. *Education in Hope: Critical Pedagogies and the Ethic of Care*. Peter Lang, New York.
- Morton, M., 2014. Resisting neoliberalism. In: Wills, R., Morton, M., McLean, M., Stevenson, M., Slee, R. (Eds.), *Tales from School: Learning Disability and State Education after Administrative Reform*. Sense Publishers, Rotterdam, pp. 195–198.
- Morton, M., 2015. Using DSE to recognize, resist, and reshape policy and practices in Aotearoa New Zealand. In: Connor, D., Valle, J.W., Hale, C. (Eds.), *Practicing Disability Studies in Education: Acting Toward Social Change*, vol. 17. Peter Lang, New York, USA, pp. 197–216.
- Morton, M., Guerin, A., 2017. Sociocultural perspectives on curriculum, pedagogy and assessment to support inclusive education. In: Noblit, G.W. (Ed.), *Oxford Research Encyclopedia of Education*. Oxford University Press, Oxford, UK, p. 23.
- Morton, M., McIlroy, A.-M., MacArthur, J., Olsen, P., 2021. Disability studies in and for inclusive teacher education in Aotearoa New Zealand. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2021.1882059>.
- Morton, M., McMenamin, T., 2011. Learning together: collaboration to develop curriculum assessment that promotes belonging. *Support Learn.* 26 (3), 109–114.
- Morton, M., Rietveld, C., Guerin, A., McIlroy, A., Duke, J., 2012. Curriculum, assessment, teaching and learning for all. In: Carrington, S., MacArthur, J. (Eds.), *Teaching in Inclusive School Communities*. John Wiley & Sons, Milton, pp. 270–292.
- Morton, M., McIlroy, A.-M., Guerin, A., 2019. Sociocultural perspectives on curriculum, pedagogy and assessment. Implications for participation, belonging and building inclusive schools and classrooms. In: Schuelka, M.J., Johnstone, C.J., Thomas, G., Artiles, A. (Eds.), *The SAGE Handbook on Inclusion and Diversity in Education*. SAGE, London, pp. 89–102.
- Nelson, P., 2013. The inquiry of care. *Educ. Theor.* 63, 351–367.
- Noddings, N., 2012. The caring relation in teaching. *Oxf. Rev. Educ.* 38, 771–781.
- Paugh, P., Dudley-Marling, C., 2011. “Speaking” deficit into (or out of) existence: how language constrains classroom teachers’ knowledge about instructing diverse learners. *Int. J. Incl. Educ.* 15 (8), 819–833.
- Rouse, M. & Florian, L. 2012. *Inclusive practice project: Final report*. Aberdeen: University of Aberdeen. https://www.academia.edu/4432038/The_Inclusive_Practice_Project_Final_Report
- Rutherford, G., 2016. Questioning special needs-ism: supporting student teachers in troubling and transforming understandings of human worth. *Teach. Teach. Educ.* 56, 127–137.
- Rutherford, G., MacArthur, J., 2018. What’s wrong with ‘special’? Thinking differently in New Zealand teacher education. In: Runswick-Cole, K., Liddiard, K. (Eds.), *Building Understandings: Child, Youth, Family and Disability*. Palgrave, UK, pp. 365–388.
- Skidmore, D., 2002. A theoretical model of pedagogical discourse. *Disabil. Cult. Educ.* 1 (2), 119–131.
- Slee, R., 2018. *Defining the Scope of Inclusive Education: Think Piece Prepared for the 2020 Global Education Monitoring Report*. UNESCO, Paris. Retrieved from. <https://unesdoc.unesco.org/ark:/48223/pf0000265773?posInSet=2&queryId=5cbb1fce-23a3-438c-9b47-11f9553aa4fd>.
- Slee, R., Tait, G., 2022. *Ethics and Inclusive Education : Disability, Schooling and Justice*. Springer Publishing.
- Swann, M., Peacock, A., Hart, S., Drummond, M.J., 2012. *Creating Learning Without Limits*. McGraw Hill, Open University Press, London.
- United Nations Educational, Scientific and Cultural Organization, 2017. *A Guide for Ensuring Inclusion and Equity in Education*. United Nations Educational, Scientific and Cultural Organization, Paris.
- Valle, J., Connor, D., 2011a. *Rethinking Disability: A Disability Studies Approach to Inclusive Practices*. McGraw Hill, New York.
- Valle, J., Connor, D., 2011b. Creating a dynamic classroom culture. In: Valle, J., Connor, D. (Eds.), *Rethinking Disability: A Disability Studies Approach to Inclusive Practices*. McGraw-Hill, New York, pp. 106–141.
- Wansart, W.L., 1995. Teaching as a way of knowing: observing and responding to students’ abilities. *Remedial Special Educ.* 16, 166–177.
- William, D., 2011. *Embedded Formative Assessment*. Solution Tree Press, USA.

Related websites

- Narrative Assessment: A Guide for Teachers. https://www.throughdifferenteyes.org.nz/a_guide_for_teachers.
- Narrative Assessment: Through Different Eyes. https://www.throughdifferenteyes.org.nz/_data/assets/pdf_file/0010/57925/Narrative-Assessment_a-guide-for-teachers.pdf.
- Through Different Eyes: Assessment for Learners With Special Education Needs. <https://assessment.tki.org.nz/Assessment-tools-resources/Learners-with-diverse-learning-needs/Through-different-eyes-Assessment-for-learners-with-special-education-needs>.
- Using Narrative Assessment to Support Secondary School Teachers’ Inclusive Practices. <http://www.tlri.org.nz/tlri-research/research-progress/school-sector/using-narrative-assessment-support-secondary-school>.

Inclusive postsecondary education: vision, practice, higher education reimagined

E. Anne Hughson^a and Arden A. Duncan Bonokoski^b, ^aCommunity Rehabilitation and Disability Studies, Community Health Science, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada; and ^bSTEPS Forward Inclusive Post-Secondary Education Society, Chilliwack, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Inclusive education debates	362
Resisting definitions through focus on practice	363
Key features taking the place of programmatic design	364
Facilitation as a strategy for advancing inclusion	364
Emanating inclusion through a commitment to reflective practice	365
What is sustainable and what emerges	366
References	367

Inclusive education debates

A postsecondary education is one of Western society's more valued and privileged pursuits. Regardless of social differences—class, sex, gender, race, ability—it is presumed that attending a college or university improves one's position in life. Engagement in postsecondary education is widely recognized as a means of promoting increased sense of belonging, improved community engagement, expanded career options, and increased chances for lifelong learning for adults (Baum and Payea, 2004; Evans et al., 2017). Scholars argue that what today's adult learners want from their education is ample opportunity to engage in academic study that is meaningful and allows them to advance personal and career opportunities (Armstead et al., 2010).

Article 24 of the United Nations Convention on the Rights of Persons with Disabilities (CRPD) (2006) prohibits discrimination against children with disabilities and mandates the right to inclusive education. However, for students with intellectual disabilities opportunities to benefit from pursuing inclusive education let alone an authentic postsecondary education have not been fully realized. Although many countries pledge to support inclusion for people with disabilities, progress has been uneven. In fact, many countries have been slow to shift from a segregated special education model (Hehir et al., 2016). In Canada and elsewhere, the traditional model of special education (i.e., life skills training, congregated, deficit-based, segregated classrooms) co-located on postsecondary campuses is more commonplace. The adult special education option, designed for eligible (some not all) post-high school students with intellectual disability, continues to replicate special education practices with claims that students will gain community access, inclusion, and lifelong learning. However, student outcomes continue to reveal disability oppression, unemployment and social isolation (Ferri and Connor, 2008; Crawford, 2011; Prince, 2016).

A dismal future prevails for so many adults with intellectual disabilities (Van Asselt et al., 2015). Individuals are more likely to live in poverty, social isolation, loneliness, experiencing devalued roles and poor health. Resisting further segregation for young adults, fueled advocates to rethink what educational options might better lead to lives of belonging and contribution in their community. Families and allies, wanting more inclusive lives for their loved ones, considered access to a typical postsecondary education as one of the more likely pathways to a good life. Their vision was to establish a way for students with intellectual disabilities, including those with significant disability-related support needs, to enroll in postsecondary institutions (PSI's) with support. Initiatives began in Canada by family leaders in the inclusion movement of Alberta and British Columbia. The vision for implementing inclusive postsecondary education has also grown in United States, Europe, and Australia.

This chapter casts an eye on the inclusive post-secondary initiatives for young adults with intellectual disabilities that have gained momentum in Canada over the last 30 years. These initiatives, energized by parents (and allies) who imagined a better future for their adult family member, continue to ground this hope in the belief that engaging in intentional relational educational practices within postsecondary settings increases the likelihood of life-long inclusion. Reported accounts are promising. For young adults in Alberta, postsecondary education has resulted in over 70% of alumni reporting positive life outcomes such as inclusive employment, often in their related area of study (Hughson and Uditsky, 2019). Initiatives continue to thrive with modest growth in the number of students gaining access to inclusive initiatives. At the time of writing, there are nine inclusive post-secondary education initiatives in the province of British Columbia, twenty-one in Alberta, one in Saskatchewan and three in Nova Scotia. These initiatives support students to enroll at colleges, universities and technical institutes in both private and public PSIs.

The articulated implementation of theory and practices in postsecondary initiatives in Alberta and British Columbia, offers a platform for describing the scope of practice in inclusive postsecondary education that has emerged. Here we focus on some stated essential features, explore emerging issues and pose forward thinking notions about possible life outcomes for adults with intellectual disabilities who participate in fully inclusive postsecondary education. Since inception, these inclusive postsecondary

education initiatives have adopted a radical vision about the possibility of offering student experiences in authentic postsecondary educational inclusion. Committed to transparency, reflective documentation of effective practices and emerging improvements position these initiatives as “sites of innovation” aspiring to live out the inclusive vision. It is a novel approach in education to set aside the reliance on organizational and programmatic design with the aim to learn more about how the power of relationships may or may not be a more likely approach to increasing a sense of belonging and learning for students who have been historically excluded from typical educational environments. Resisting formalized or standardized programmatic structures typically offered to students with “differences” was understood to be fundamental to avoid replicating the paradoxes reported in educational reform research (Slee, 2014).

Parallel to emerging inclusive initiatives in postsecondary education, debates about inclusion in public schools began in the 1990s, based on rights and entitlement rhetoric more often than on the inclusion movement driven by families. Despite many government policies and legislation to reform and create inclusive education systems in UK and North America, the ideas and policies remain vague and elusive today. Although competing inclusive education discourses and practices associated with race, ethnicity, gender and disability construct different meanings by different vested interests and by different people, a strong association between regular special education is retained (Dunne, 2009). Defining a coherent inclusive post-secondary education framework might be seen as reflecting similarities with the discursive tensions associated with the meaning of inclusive public education. The history of higher education traditions, economic basis for high returns on investment and the ableist culture of elite postsecondary institutions are but some of the apparent contradictions that rival progressive policies as a way forward to more socially equitable and inclusive educational systems regardless of individual differences (Kromydas, 2017).

The critical interrogation of the central educational debates encouraged these inclusive postsecondary initiatives to resist institutional and structural definitions that, by their very definition, might confine inclusive postsecondary education to a set of prescribed organizing practices, set of policies, programmatic naming, or identified spaces within the post-secondary landscape. Their perspective was also informed by an examination of the construction of discourse on diversity, patterns of “othering” as well as by research that continues to draw attention to the pervasive patterns, infrastructures, technological features of social exclusion and stratification that perpetuate significant barriers to achieving equitable education for students with intellectual disabilities (Graham and Slee, 2007; Slee, 2011, 2014). Attending to the paradox of naming programs as inclusive education and the ongoing discourses about student differences shaped the conceptualization and action of implementers of inclusive postsecondary education initiatives to avoid embedding “special programmatic designs” that reproduce stigma and the educational marginalization of adult students who are “different.”

Responding to competing discourses in educational research, dominant social biases and hierarchical leadership that drive reforms in complex educational systems, the emergence of inclusive postsecondary education in Canada continues to be guided by persistent efforts to minimize programmatic/institutional practices that separate students. Choosing, instead, to emphasize the development of intentional relationships that build on the ability of postsecondary communities to think and act in ways that actively promote student identity and belonging for those who have been historically excluded. There is a steadfast conviction that sharing the philosophy and theory that students engaged in meaningful connections while attending college or university courses and participating in the campus social life, will change their predicted trajectory. It is argued that embedding students in the full range of typical opportunities available to nondisabled students means they are more likely to contribute to campus and lead more valued adult lives. Central to establishing and implementing inclusive post-secondary education means that “inclusive practices” are characterized as finding ways and means to seek out and leverage existing social and academic experiences more likely to benefit all students, including students with intellectual disabilities.

Resisting definitions through focus on practice

While there are no assessments, high stakes testing or categorical programmatic design features associated with the implementation of this approach to inclusive postsecondary education, key features reminiscent of high-fidelity initiatives have been articulated by the initiatives (O’Brien et al., 2009; Aylward and Bruce, 2014; Hughson and Uditsky, 2019). Examining language usage patterns and the implications of long held assumptions are but one example. The innovators of inclusive education realize that, as a term, “inclusion” denotes social progression, a nice enough word, connoting magnanimity and beneficence as well as signaling welcome and tolerance. However, it is also understood that the excluded never chose their status and the change in conferred status is arbitrary. While discourses around inclusion, equity and diversity are now rapidly emerging in postsecondary education rhetoric and policy documents, a newly included student may retain that sense of being an outsider. Indeed, current institutional mandates to embrace diversity alongside the explicit, covert or simply tacit understandings associated with exclusive postsecondary institutional culture galvanize increasing attention to the implications of identity politics related to new policies, practices and relationships especially when applied by those who have seen inclusion as a birthright. Allowing students with intellectual disabilities to be part of the post-secondary education campus by the grace of the host, means practicing vigilance and critical reflection on everyday practices to side-step being seduced by other vague or elusive vested interests in institutional inclusion. Measuring implementation by accounting for those harmonious actions that contribute to increasing the inclusive experiences for students is identified as key to guiding inclusive postsecondary education practices. In practice, insights gained by examining and interpreting emerging practices that reflect the “least difference making” approach and build on the most meaningful (and sometimes new to all parties) connections and

relationships are the benchmarks that sustain hope in postsecondary settings that welcome individuals with intellectual disabilities (O'Brien et al., 2009).

Key features taking the place of programmatic design

Innovative thinking is often marked by visionaries' attention to changing intentional social practices. Facilitators of the initiatives aim to shift conscious and unconscious institutional habits of mind and "normative" social segregation by negotiating opportunities for learning, career choices, identity and belonging for students historically denied access and accommodation in a university, college or technical school. In no order and all of equal importance, key features include; (i) students enroll based on their desire to continue their education, there are no minimum academic or behavior requirements that must be met to enroll, (ii) initiatives build their understanding of typical student pathways and are intentional about uncovering the least difference-making approaches to addressing anything that might limit student participation on these pathways; (iii) adaptations and support are developed so that they are harmonious with student life; (iv) initiatives build the capacity of the campus community to support and include students, it is not one person's responsibility to "do" inclusion instead it is a shared commitment of the PSI and the inclusive postsecondary education initiative to work toward a more inclusive and equitable experiences that strengthen future relationships.

To repeat, inclusive postsecondary initiatives ensure students, enrolled at their local technical school, college or university, do not attend any specialized, congregated classes or program supports targeted to their disability label. Instead, initiatives intentionally support students to select options from the existing postsecondary programs of study. This is in stark contrast to historical and current assumptions of special education approaches that believe students need specialized preparation and segregated learning opportunities to compensate for their cognitive impairments before they can participate in adult education. Inclusive postsecondary education contests this approach by supporting students to explore and select an available program of study based on their own dreams and hopes for learning in their PSI. Any additional campus services or programs used by typical students are sought (if needed) in the same way by students with intellectual disabilities.

Each student selects courses and follows a coherent pathway along with their classmates over the respective years in that discipline/program. The inclusive postsecondary education facilitator and faculty members collaborate to modify course work for each student so that they are fully engaged in all the typical class learning opportunities (assigned reading, labs, tutorials, exams, assignments, field schools). Students do not receive a credit for their course work which allows for the modifications to be tailored to the learning style and the learning goals of the individual student. Upon completion of their program of study they receive a certificate of completion, portfolios of course work in their program of study with references and participate in convocation with classmates from their program of study.

Over the last 30 years, students with intellectual disabilities in Canada have pursued programs of study across a wide array of academic, artistic, and technical fields. Inclusion in a program of study and in individual courses is established with the permission of the faculty program of study and individual professor. Overwhelmingly, the evident responses of program administrators and faculty reflect an openness and eagerness to include a student who is interested in joining their class. This relational approach to course enrollment invites professors, instructors and Teaching Assistants to meet with inclusion facilitators who work closely with the student requesting the course. Building on this relationship means instructors can learn about the student's interest in the course, their goals and the role of the inclusion facilitator in supporting the course modifications and the student experience beyond the classroom. Inclusion facilitators are in regular communication with professors and instructors to discuss strategies and identify promising practices that forge a student's sense of belonging in each class.

This relational approach to individual student course enrollment is a departure from the typical ways academically eligible students with disabilities must request academic accommodations when choosing their postsecondary program of study. While governing processes may be unique to each PSI, access and accommodation is usually characterized by a legal duty implemented by an infrastructure that requires a medical diagnosis of disability, an emphasis on confidentiality (to the point of anonymity) and policies that place the burden of seeking accommodations on the student (<https://sswd.jibc.ca/duty-to-accomodate/>). Instead of such a puzzling approach to predicting and controlling disabled student's engagement based on a medical diagnosis, inclusion facilitators participate with willing professors in an iterative practice. Each term the student engages with the course, the facilitator and the instructor to develop least intrusive accommodations and/or instructional strategies to engage with course content. The inclusion facilitator aims to establish accommodations and strategies that are typical of the classroom setting and are the least intrusive to customs of the class. Professors, with no special education training, have found that they are able to make adaptations to their learning environment that increase student engagement and in fact, report that such instructional changes have a positive impact on all their students. Furthermore, the faculty have indicated that active participation in broadening the diversity of their classrooms has challenged their own limited assumptions about the learning goals of students with intellectual disabilities (Bingham and Duncan Bonokoski, 2018).

Facilitation as a strategy for advancing inclusion

The facilitated supports for students with intellectual disabilities are multifaceted, often perceived as counter cultural and represent operational features that are unique to the inclusive nature of these initiatives. Inclusive postsecondary education facilitators are

most often employed by a non-profit organization external to the PSI. Their work on campus is made possible through a memorandum of understanding or partnership agreement between the PSI and the non-profit organization. In a few initiatives in Canada, inclusion facilitators are employees of the PSI. Facilitators have no formal role or authority conferred by the postsecondary infrastructure. Rather, the focus of their work is to advance inclusion through collaborations that invite an interest in potential shared activities that might spark an interest in mutual connections. Uncovering student capacities to adopt new roles, facilitators look for coherent student opportunities that benefit a student with an intellectual disability in the same way typically enrolled students benefit from their academic and social roles and experiences. Facilitators find ways for people to meet and learn about each other. For many nondisabled students, these encounters are unique as education in regular education systems often limit natural encounters with students with intellectual disabilities. These strategies are rooted in the hope that making space for freely given friendships happens when typical students identify personally with the student, their instructors, other classmates, members of campus clubs and other student services. There are repeated accounts of members of the PSI community welcoming students in courses, student clubs and informal social activities of campus life.

Unsurprisingly, facilitators identify encounters of conscious and unconscious social bias that can interfere with a normative student experience. For example, because course work is not for credit, instructors may have lower expectations for students to submit their assignments or to complete exams. When this occurs, facilitators work with instructors to emphasize the importance of keeping their expectations high when it comes to typical student participation in the course. In courses with group work, facilitators connect with group classmates as a resource to support the assigned group to figure out how to meaningfully include the student when the group dynamic may be insensitive or assume the student's participation is less important. This might involve highlighting the student's strengths and working with the group to identify a role in the project for the student to add value to the group assignment. Facilitators indicate that this practice opens opportunities for positive interactions that challenge classmates' assumptions about the student's desires and have been known to shift and affirm that the student with an intellectual disability share the same goals as the nondisabled peer.

Facilitators describe examples of occasional encounters with blatant social bias when students with an intellectual disability chose to enroll in a program of study related to the helping professions or teacher education. Facilitators are keenly aware of the thin veneer of inclusive discourse that nonetheless fosters widespread negative assumptions consciously or unconsciously in professional disciplines rooted in medical constructs of disability. They recognize the importance of confronting inequality and devaluing assumptions that impact the students' identity and participation as a member of the class. Specifically, facilitators found that student identity can be at risk when an instructor initially sees the value of a student joining their social work or education class because it would provide an opportunity for their typically enrolled students to gain experience with a person who they might serve as a client in their future career. In order to protect the students' identity and role as a valued learner in the class, facilitators meet with the instructor before the student joins the class, to reinforce the fact that participation in class is expected just as any other student—not “placed” as a practicum opportunity for other students. Facilitators are careful to explain that the greater value would be derived from students learning how to work alongside someone with a disability, proposing that one day, they may become work colleagues in their chosen field. Such confrontational conversations have been reported as uncomfortable, however, instructors, peers and PSI staff largely report the interactions are welcome because the institutions want to learn more about how faculty can be inclusive leaders in creating just communities of practice on and off campus.

Emanating inclusion through a commitment to reflective practice

Those experienced in implementing inclusive postsecondary initiatives recount the ubiquity of exclusion and know no blueprint for identifying and dismantling the negative devaluing beliefs that perpetuate social biases tied to the label of intellectual disability (Hughson and Uditsky, 2019). In fact, documented gaps in social and economic life outcomes for people with intellectual disabilities remain current indicators that reinforce long held stereotypes and prejudices endemic to postsecondary community. Facilitating inclusion within the postsecondary context is further complicated by the many ways that a busy student life unfolds. Complexities of social and academic life for a student may tempt a facilitator to seek simpler solutions. One such risk is a drift toward implementing a special education or charitable approach to resolve perceived vulnerabilities such as loneliness or risky individual choice, rationalizing that such actions are needed in the name of safety or compassion. To safeguard engagement in valued practices, inclusive postsecondary initiatives have built in operational activities that regularize commitments to set times for reflective practice. Sharing the challenges, risks, and successes in formal and informal ways are seen as a systematic approach to examining alternatives that might disrupt the pervasive reaction to protect by segregating.

Reflective practice is known to be essential to countercultural work in emerging pockets of social innovation. Inclusive postsecondary education initiatives share the view that such practices are central to their mission. To that end, these initiatives put considerable effort into collectively interrogating ways to understand commonly held devaluing beliefs, circulating their experiences and critiquing the narratives that are retold. Critical examinations of student experiences are reviewed each term and mutual plans are made to improve efforts that pull students into valued roles and relationships not bound by low expectations and loneliness. Documenting reflective conversations and interpretations capture promising practices that are shared with “critical friends” interested in innovation and change. Annual inclusive postsecondary symposiums have been hosted over the last few years (see: <https://www.center-ipse.org/2021-symposium-resources.html>). Initiatives know what is at stake and want to push the inclusive

agenda in education, particularly when current policies, cultures, and practices continue to place students with disabilities at greater risk of educational and social exclusion (Slee, 2014).

As well as regular reflective practice, an evaluation tool has been developed by the early implementers of inclusive postsecondary education in Alberta. This tool articulates the commitments of inclusive post-secondary education and lays out benchmarks that are used to understand the quality of inclusion experienced on campus. The evaluation process is not interested in satisfaction or performance appraisal of students, faculty, their families or the post-secondary administration. Rather the benchmarks help focus the analysis of these accounts by considering how students are embedded or not in an authentic student experience. The evaluation tool is used by designated teams of inclusive postsecondary peers both external to and connected with the initiative being evaluated. Measuring the quality and fidelity of practices that describe the students' experiences at one point in time include gathering interviews with faculty, students, families, administrators, and alumni engaged in inclusive postsecondary education initiatives (O'Brien et al., 2009). Regular evaluation by "critical friends" is a salient commitment for inclusive postsecondary education initiatives as it is usually easier to fall into perpetuating habits that accept the status quo of access rather than push the practice forward. This commitment to evaluation, by initiatives who are deeply curious about ways to facilitate intentional relational education practices, ensures that implementation is held in a state of innovation from which new learnings and practices can emerge.

What is sustainable and what emerges

It is important to contemplate what will sustain and safeguard inclusive postsecondary education while inclusive education debates and doubts continue. Many questions arise. Will the preoccupation with the value of credentialing create barriers? Is it possible to critique design elements that are typical of elite educational systems while leveraging the existing pathways and programs that are normatively offered? As described earlier, students do not meet academic prerequisites to enroll. What is known to date, is that participating PSI's continue to adjust routine rules, welcome and accept students with intellectual disabilities that were systematically denied access to an education that met the pre-requisites to PSI admission.

Inclusive postsecondary education troubles other long held traditions in education. Students with intellectual disabilities welcomed in postsecondary level courses in a wide range of subject matter, including upper-level courses without the academic prerequisites, shift the pedagogical assumption that all student learning naturally unfolds in sanctioned curriculum of sequenced developmental pattern delivered in modules from simple to more advanced concepts. Educational philosophers and critics ponder the current repackaging of educational activities as "learning" activities whereby education is described as the provision of learning opportunities or learning experiences (Bingham, 2015). An emerging body of educational theory argues that there is not one known logic or a particular natural "rate" of learning. The experiences and accomplishments of students with intellectual disabilities included in postsecondary education bring life to new theorizing.

While the expectations for learning in inclusive post-secondary initiatives may disrupt typical educational designs, the "work" also actively engages and leverages student roles and identity in the defined programs of study offered. Enrolling in an identified postsecondary program of study from the outset, builds the student's sense of belonging and identity in the mores, customs and status associated with that career path (such as an ecology student, a business student, or a trades student). From year to year the inclusive postsecondary initiatives facilitate more and more students who have connected with instructors, classmates and club members associated with their specific interests in chosen programs of study and do go on to gain inclusive employment. Continuous participation in a wide range of postsecondary programs of study over the years may lead more faculty, administrators and students to expect there will always be a place for students with intellectual disabilities in the classroom. More students with intellectual disabilities can imagine pursuing further studies in higher education alongside peers with the same interests and goals. Furthermore, the expectations for equal opportunities to engage in learning have the power to move beyond campus boundaries and into the professional and personal lives of all postsecondary graduates who have the benefit of experiencing authentic inclusion through their postsecondary education. It is not uncommon for facilitators to report that the student's success in finding paid employment in their field after graduation is the result of an employer having experienced inclusion through their post-secondary studies.

Inclusive postsecondary education disrupts some long-held conventions. For example, it has been the role of universities and colleges preparing students for roles in social science disciplines such as education, psychology, social work and medicine to mirror social and cultural inequalities and social hierarchies. Some social science disciplines theorize and define intellectual disability in terms of categorical diagnoses based on deficits, lack intellect and individuals are seen as being in a state of disconnectedness from those without disabilities (Klewer et al., 2015; Goodley, 2007; Slee, 2001). Such social constructions invite educators to continue to conclude that people with intellectual disabilities are unable to meaningfully engage with same age peers in regular education curriculum. In fact, when such constructs are believed to be "true" or "natural" further progress in inclusive education is unlikely to be realized. However, in postsecondary settings where inclusive initiatives have grown and established relationships, many key faculty members and classmates have already opened doors to some interesting possibilities for educational and social change. Practices that actively construct the "work" of inclusion together with professors, instructors, teaching assistants, postsecondary administrators, and students may be a way to unpack devaluing assumptions and create spaces to know students with intellectual disabilities as learners and contributors. Presumed competence embraces the theory that belonging is essential to learning (O'Brien and Mount, 2015). Inclusive initiatives provide opportunities for many more people to share some common experience throughout their postsecondary education and to engage with students with intellectual disabilities from a place of connectedness. To create the

conditions for adults with intellectual disabilities to become fully realized students becomes a work in progress, suggesting that we are only limited by a lack of imagination. These relational experiences, intentionally facilitated in unremarkable ways, may quietly chip away at the cultural politics of disablement (Slee, 2001) and open new possibilities for students with intellectual disabilities to break through the biases that stand in the way of their education, employment and social connections. Hundreds of postsecondary courses and programs of study have welcomed students over the last 30 years. Accordingly, many shared experiences across time and settings, have been taught by hundreds of different faculty members (Hughson and Uditsky, 2019) and freely given relationships have directly emerged among peers and students with intellectual disabilities because of their mutual participation in class and in extra-curricular activities across campuses. These initiatives may not persuade those engaged in vague and disputable interpretations of inclusion on the one hand, while on the other, there are those who know inclusive education can be realized.

References

- Armstead, C.L., Bessell, A.G., Sembante, S., Plaza, M.P., 2010. What students need; what students say they want: student perspectives on the promise of smaller learning communities. *Peabody J. Educ.* 85 (3), 365–374.
- Aylward, L.M., Bruce, C., 2014. Inclusive post-secondary education in Canada: transition to somewhere for students with intellectual disabilities. *J. Int. Assoc. Spec. Educ.* 15 (2), 42–47.
- Bingham, C., 2015. On the Logic of Learning and Relationality. *Philosophy of Education*. <https://educationjournal.web.illinois.edu/archive/index.php/pes/article/view/4485.pdf>.
- Bingham, C., Duncan Bonokoski, A., 2018. Inclusive Postsecondary Education as a Practice of Social Justice Final Report. Simon Fraser University Teaching and Learning Development Grant. <http://www.sfu.ca/content/dam/sfu/tigrants/>.
- Crawford, C., 2011. The Employment of People With Intellectual Disabilities in Canada: A Statistical Profile. Institute for Research and Development on Inclusion and Society (IRIS), Toronto.
- Dunne, L., 2009. Discourses of inclusion: a critique. *Power Educ.* 1 (1), 42–56.
- Evans, N.J., Broido, E.M., Brown, K.R., Wilke, A.K. (Eds.), 2017. *Disability in Higher Education: A Social Justice Approach*. Jossey-Bass, San Francisco, C.A.
- Ferri, B., Connor, D., 2008. Tools of exclusion: race, disability, and (re)segregated education. *Teach. Coll. Rec.* 107 (3), 453–474.
- Goodley, D., 2007. Towards socially just pedagogies: deleuzoguattarian critical disability studies. *Int. J. Incl. Educ.* 11 (3), 317–334. <https://doi.org/10.1080/13603110701238769>.
- Graham, L., Slee, R., 2007. An illusory interiority: interrogating the discourse/s of inclusion. *Educ. Philos. Theor.* 40 (2), 277–293.
- Hehir, T., Grindal, T., Freeman, B., Lamoreau, R., Borquaye, Y., Burke, S., 2016. A Summary of the Evidence on Inclusive Education. Institute Alana, Brazil, pp. 1–40. <https://eric.ed.gov/?id=ED596134>.
- Hughson, E.A., Uditsky, B.L., 2019. Inclusive post-secondary education: 30 years. In: O'Brien, P., Bonati, M., Gadow, F., Slee, R. (Eds.), *People With Intellectual Disability Experiencing University Life*. Brill Sense, Leiden.
- Kliwer, C., Biklen, D., Petersen, A., 2015. At the end of intellectual disability. *Harv. Educ. Rev.* 85 (1), 1–28. <https://doi.org/10.17763/haer.85.1.j260u3gv2402v576>.
- Kromydas, T., 2017. Rethinking higher education and its relationship with social inequalities: past knowledge, present state and future potential. *Palgrave Commun.* 3, 1–12. <https://doi.org/10.1057/s41599-017-0001-8>.
- O'Brien, J., Mount, B., 2015. <https://inclusion.com/site/wp-content/uploads/2019/09/What-More-Is-Possible.pdf>.
- O'Brien, J., Bowman, P., Chesley, B., Hughson, E.A., Uditsky, B., 2009. *Inclusive Postsecondary Education: Measuring Quality and Improving Practice*. Government of Alberta, Edmonton.
- Prince, M.J., 2016. Inclusive Employment for Canadians With Disabilities. Institute on Public Policy. <http://irpp.org/wp-content/uploads/2016/08/study-no60.pdf>.
- Slee, R., 2001. Driven to the margins: disabled students, inclusive schooling and the politics of possibility. *Camb. J. Educ.* 31 (3), 385–397. <https://doi.org/10.1080/03057640120086620>.
- Slee, R., 2011. *The Irregular School, Exclusion, Schooling and Inclusive Education*. Routledge, Abingdon.
- Slee, R., 2014. Discourses of inclusion and exclusion: drawing wider margins. *Power Educ.* 6 (1), 7–17. <https://doi.org/10.2304/power.2014.6.1.7>.
- Thoma, C.A., Lakin, K.C., Carlson, D., Domzal, C., Austin, K., Boyd, K., 2011. Participation in postsecondary education for students with intellectual disabilities: a review of the literature 2001–2010. *J. Postsecond. Educ. Disabil.* 24 (3), 175–191.
- Van Asselt, D., Buchanan, A., Peterson, S., 2015. Enablers and barriers of social inclusion for young adults with intellectual disability: a multidimensional view. *J. Intellect. Dev. Disabil.* 40 (1), 37–48. <https://doi.org/10.3109/13668250.2014.994170>.

Researching curriculum and disability in higher education

Shiralee Poed^a, Miriam Edwards^b, Hadeel F. Al-Nawab^c, and Olivia Penna^d, ^a School of Education, The University of Queensland, St Lucia, QLD, Australia; ^b Faculty of Business and Economics, The University of Melbourne, Parkville, VIC, Australia; ^c Faculty of Education, Monash University, Melbourne, VIC, Australia; and ^d Melbourne Graduate School of Education, The University of Melbourne, Parkville, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	368
Students with disability in higher education	369
Legislative obligations	370
Reasonable adjustments	370
Reasonable accommodations/adjustments typically afforded in higher education	370
Disability support services	370
Barriers to higher education for students with disability	371
Facilitators to higher education for students with disability	372
Conclusion	373
Acknowledgment	373
References	373
Further reading	375
Relevant websites	375

Glossary

Accommodations or adjustments Steps taken by an education provider that enable a student with a disability to enroll, access, or participate in higher education on the same basis as their peers

Barriers Obstacles that prevent a student with disability from accessing tertiary education or that can prevent their ability to participate and demonstrate their ability

Higher education Courses that are completed by students as post-secondary education, that may lead to an advanced qualification such as a vocational certificate, certificate, diploma, advanced diploma, graduate certificate, post-graduate certificate, bachelor degree, masters degree, or doctoral degree

Inclusive learning An experience where all students learn alongside one another, in an environment where diversity is respected, celebrated, and accommodated, where all students have equitable opportunities for participation, and the barriers that prevent learning have been recognized and removed wherever possible

Inherent requirements The core knowledge, skills or abilities considered essential in order to perform a task safely and effectively within the workplace, e.g., in order to drive a bus/coach, an inherent requirement is that the driver has a valid drivers' license that permits the driver to operate that class of bus/coach

Reasonable As much as is necessary or fair, that enables the person with a disability to enroll, access, and participate in education on the same basis as their peers

Universal design Designed through careful planning to allow education, to the greatest extent possible, to be accessed and understood by, and inclusive of, all students

Introduction

On December 13, 2006, the Convention on the Rights of Persons with Disabilities (CRPD) and its Optional Protocol ([United Nations, 2006](#)) was adopted, then opened for countries to sign from March 30, 2007, and finally came into force on May 3, 2008. There are 193 member states of the United Nations, with the UN CRPD the most widely ratified human rights treaty, currently ratified by 182 member states ([United Nations, 2006](#); [Kanter, 2020](#)). In ratifying a UN convention, member states agree to ensure that domestic legislation is introduced or amended to comply with the obligations specified within.

Article 24 (section 1) of the United Nations *Convention on the Rights of Persons with Disabilities* ([United Nations, 2006](#)) recognizes the right to an inclusive education without discrimination for people with disability. It articulates the purpose of this is to allow students with disability to develop their sense of self-worth and dignity, their talents, and their abilities, to the fullest extent possible so that they can participate within society. Article 24 (section 5) makes explicit that this right extends to “general tertiary education,

vocational education, adult education and lifelong learning” and for this right to be realized, reasonable accommodations must be provided to students (United Nations, 2006).

This right was further reinforced by *General Comment No. 4* (United Nations [Committee on the Rights of Persons with Disabilities], 2016) which obligates member states to ensure that learners with disability should be provided an accessible education, including equal access to tertiary education without discrimination and on the same basis as their peers (section 23). Member states are further encouraged to adopt affirmative action measures to ensure learners with disability can access and participate on an equal basis. These measures must address attitudinal, communicative, financial, legal, linguistic, physical, or other barriers that inhibit access and participation (section 37). Section 69 calls upon member states to ensure that there is a process for educating tertiary staff so that they have the necessary skills to provide inclusive learning environments. One recommendation is that all staff are “provided with dedicated units/modules to prepare them to work in inclusive settings, as well as practical experiential learning, where they can build the skills and confidence to problem-solve through diverse inclusion challenges” (United Nations [Committee on the Rights of Persons with Disabilities], 2016).

As Slee (2018) reminds us, for decades universities have played a pivotal role in training and grooming social elites so that they can take predetermined places in professions and leadership roles within industry, while students with disability have “been hidden behind the edifice of special education” (p. 39). Research that investigates the experience of students with disability in higher education is beginning to grow (Linder et al., 2015). From this research, we learn that the traditional methods for teaching and assessing used within universities continue to disadvantage students with disability, and that curriculum reform in higher education is critical to the inclusion of students with disability (Taylor and Shallish, 2019; Podzo and Chipika, 2019). Further, while the UN have called for member states to ensure that students with disability can access a tertiary education without discrimination, and on the same basis as others, the experience of students with disability in post-secondary environments suggest much work is still needed to improve access and participation so that students can thrive in post-school educational settings.

The purposes of this article are to (a) familiarize readers with the extent of enrollment of, and the legal obligations of higher education institutions toward, students with disability, (b) identify the provisions currently offered to this set of learners, (c) explore the barriers and facilitators to access, participation, and engagement within higher education for students with disability, and propose future directions for enhancing the educational experience so that students with disability can engage in life-long learning and have their inclusive education rights upheld.

Students with disability in higher education

The 21st century has seen an influx of people with disability into higher education across the globe (Caleb and Barden, 2019; Barnes, 2007; Konur, 2006; Biggeri et al., 2020; Werth et al., 2014; Ganguly et al., 2015; Zeng et al., 2018; Moriña et al., 2014; Grimes et al., 2019). The drivers for this increased participation, in addition to international human rights agendas, include legislative enactments, higher education policy agendas, improved institutional ethos, and provisions made available to support participation and engagement (Brett, 2016; Riddell et al., 2005; Brandt, 2011; Sarkar, 2016; Madaus et al., 2012; Werth et al., 2014).

Fig. 1 shows the growth of Australian students with disability accessing tertiary education between 2012 and 2019.

This represents a growth in access of 87.13% over the past decade. This is considerable when compared to the growth for the total number of students accessing higher education was 17.49% measured over the same timeframe (Australian Government, 2020b). But, as Gale (2010) cautioned, inclusion of people with disability is not purely about numbers, and for inclusion to be the lived experience of students with disability in higher education, it is critical that we reconfigure higher education to meet our ratified obligations.

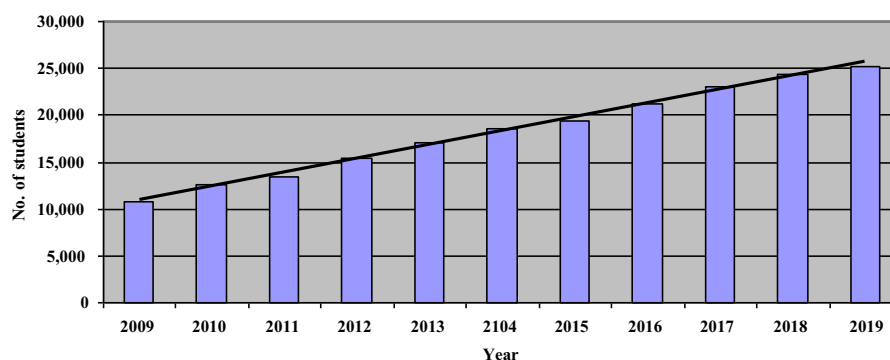


Fig. 1 Access by Australian students with disability to higher education.

Legislative obligations

Countries around the world have passed legislation enshrining the rights of students with disability in higher education (Moriña et al., 2014). In the United States, the legal mandate to provide adjustments for students with disability in higher education was initially documented in the *Rehabilitation Act of 1973*, and reinforced in the *Americans with Disabilities Act of 1990*, the *Americans with Disabilities Act Amendment Act of 2008*, and the *Higher Education Opportunity Act of 2008* (Behling and Linder, 2017; Kimball et al., 2016). In the United Kingdom, the passage of *The Disability Discrimination Act* (Home Office, 1995) was amended in 2001, through *The Special Educational Needs and Disability Act* (Department of Education and Skills, 2001) to include education and an obligation to provide reasonable adjustments. Further, in late 2006, the *Disability Equity Duty* (Disability Rights Commission, 2006) obligated higher education institutions to treat students with disability fairly and equitably on the same basis as their peers. In Australia, the *Disability Discrimination Act 1992* (Australian Government [Federal Register of Legislation], 2017), the *Disability Standards for Education 2005* (Australian Government, 2020a), and the *Disability (Access to Premises—Buildings) Standards 2010* (Australian Government [Federal Register of Legislation], 2020) provide the federal legislative platform for ensuring the rights of students with disability are upheld in educational settings, including those that serve post-secondary students.

Legislative frameworks are a way of countries demonstrating their compliance with their obligations as a signatory to the UN, but they are also important because they offer clarity to providers in meeting their obligations to people with disability, and are meant to afford people with disability protection against discrimination (Poed, 2016). To enable students with disability to access education, and participate on the same basis as their peers, legislation typically specifies what actions an education provider must take to ensure that students with disability are not treated less favorably and to avoid actions that would be viewed as discriminatory.

Reasonable adjustments

One action usually imposed by legislation on education providers is an obligation to provide reasonable accommodations or reasonable adjustments. Often used interchangeably, the term “reasonable accommodations” was adopted in the *Americans with Disabilities Act of 1990* (United States [Department of Justice—Civil Rights Division], n.d.), while the term “reasonable adjustments” was adopted in the *UK Equality Act 2010* (Government Equalities Office and Equality and Human Rights Commission, 2015) as well as the Australian *Disability Standards for Education 2005* (Australian Government, 2020a). Both refer to either actions or measures taken by an education provider to ensure that students with disability can enroll, access, and participate in education and training on the same basis as their peers (Ashman, 2019).

Reasonable accommodations/adjustments typically afforded in higher education

At the end of the previous century, the *United Nations Educational Scientific and Cultural Organization* (1999) surveyed the provisions made for students with disability in higher education. Responses were provided by 35 universities from different UNESCO regions, with adjustments being summarized under an array of categories.

First, some generic adjustments included general and specialized student counseling, training of staff (both academic and administrative), and resourcing support services for students with disability. Second, a small number of universities acknowledged adjustments made to the entrance process for students with disability, including supporting students to select an appropriate course, providing preferential treatment through modified (lower) entrance criteria, and exempting students from having completed prerequisite studies/subjects. Third, environmental adjustments were made including ensuring physical, visual, and auditory accessibility for students on campus to lecture theaters, laboratories, libraries, and the campus more broadly, as well as to living quarters, toilets, parking facilities, special transportation, telephones, and cafés/restaurants. Fourth, supports were made to allow students with disability to achieve academically, including flexibility around due dates, content delivery, adjustments to examinations, pedagogical support provided by tutors, adapted texts, and assistive technology (including training in its use). Fifth, the survey found universities were making adjustments to allow students with disability to thrive socially, including encouraging students with disability to join regular student clubs, but also creating special associations for students with disability. Finally, some universities reported providing adjustments to assist students with disability post-graduation, including career planning (United Nations Educational Scientific and Cultural Organization, 1999).

Disability support services

In keeping pace with the increased presence of students with disabilities accessing post-secondary education, there has also been a rise of disability support services provided in tertiary institutions. Known by various names including equity support units, disability services units, and disability resource units, these services are meant to “level the playing field” through the provision of adjustments that allow students with disability to access and participate in higher education on the same basis as their peers (Clark et al., 2018). The staff who work in these services translate the university’s legislative obligations, as well as their policies and procedures, into a set of services that enable students with disability to access and participate, and hopefully excel, in their

studies (Hartley, 2015). Those who work in these services have been challenged to find ways to ensure that higher education curriculum across all disciplines is accessible for students with disability (Behling and Linder, 2017).

To access support through disability services units, students are often required to disclose and provide evidence of their disability diagnosis. Research indicates that disclosing a disability at the point of enrollment or throughout their degree can be a vexed decision for many students (Ganguly et al., 2015; Clark et al., 2018). For students with disability, particularly those with “hidden disability”, reasons for not disclosing are often cited as:

- concern that disclosure may result in not being offered a place at university,
- perceived or real stigma, where deficit and medical models of disability define who can access additional supports,
- not wanting to seek medical confirmation or an assessment to access services,
- unawareness that a medical condition, learning disability, or mental health diagnosis makes them eligible for access to disability supports,
- costs associated with getting an assessment or diagnosis confirming their disability,
- concerns about the impact disclosure might have on staff perceptions, opportunities for post-graduate study or employment,
- fear of discriminatory or differential treatment and negative peer and staff attitudes,
- lowered self-esteem tied to their diagnosis, and
- past negative experiences associated with disclosure (Clark et al., 2018; Couzens et al., 2015; Werth et al., 2014; Kent et al., 2018; Ganguly et al., 2015; Hernandez, 2011; Jacklin, 2011; Kimball et al., 2016; Markoulakis and Kirsh, 2013; Dunn and Burcaw, 2013; Vickerman and Blundell, 2010; Grimes et al., 2019; Fossey et al., 2015).

Clark et al. (2018) found four circumstances in which students are less likely to disclose their disability: (1) if they are able to hide their disability from others, (2) if they perceive the higher education environment to be hostile toward students with disability, (3) if the request for receiving adjustments is confronting, and (4) if there is no perceived benefit in disclosing, or if the benefits of disclosure are unclear. This reluctance to disclose can result in students with disability needing to spend more time completing tasks than their peers, having less time for social opportunities or for employment, and increased anxiety (Kimball et al., 2016). Further, a failure to disclose can result in the student with disability waiving their right to adjustments (Grimes et al., 2019).

The onus to disclose also applies to students for whom English is an additional language or dialect or students who are First Nations/Indigenous, and similar barriers to disclosure have been identified for these cohorts of students, raising the added concern as to whether intersectionality; that is, overlapping identity markers; further decreases self-disclosure rates.

Fossey et al. (2015) found that disability service staff have a different set of criteria on which they judge the success of their work when compared to the judgments made by students with disability. Staff tend to believe they have successfully supported students with disability based on retention rates, or on students deferring in order to better prepare for tertiary studies; students, however, judge the impact the adjustments had on their outcomes, their motivation, and their enjoyment of their studies.

Barriers to higher education for students with disability

Yusuf et al. (2009) issue the challenge that with widening participation of students with disability into higher education, we have an obligation to ensure that we are not simply bringing “a flood of students into increasingly dysfunctional institutions” (p. 110). While, as an example, Australia has experienced a growth of almost 90% in the number of students accessing higher education over the past decade, each year over the same timeframe retention rates have steadily remained around the 80% mark (Australian Government, 2020b). In comparison, retention rates for all students hover between 82 and 83%, indicating a slightly higher rate of attrition for students with disability (Australian Government, 2020b). This may be because the adjustments provided to students with disability in higher education are enough simply to survive, but not sufficient to thrive (Reutlinger, 2015).

As noted by Strnadová et al. (2015), the barriers experienced by students with disability in higher education are prevailingly institutional; adapting the culture and climate of higher education institutions is still the key to removing these barriers. Most critically, students offer external environmental issues as the most common barriers, rather than individual factors impacting their academic success (Ganguly et al., 2015). Students advise that their ability to navigate the system to identify the supports available to them can be the first hurdle (Edwards et al., under review). Further, students identify that misunderstandings by teaching staff, negative attitudes of peers and both administrative and teaching staff, inaccessible curricula and pedagogy, low expectations, attendance demands, frequent turnover of disability support personnel, lag times in accessing accommodations and supports, poor universal access to the campus environment, as well as health, counseling, and financial difficulties as key barriers (Ganguly et al., 2015; Fleming et al., 2016; Moriña et al., 2014; Grimes et al., 2019; Verdinelli and Kutner, 2016; Strnadová et al., 2015). Students report that even when provided with appropriate documentation, some academic staff still refuse to provide adjustments (Ganguly et al., 2015). Hartley (2015) also reported that students expressed uncertainty as to whether investing in higher education was worth it given the lack of job opportunities and barriers to employment experienced by people with disability.

Barriers have also been identified by staff who work with students with disability. A common barrier is the time available for collaboration between disability support personnel and faculty members to co-design learning materials (Behling and Linder, 2017). Some academic staff believe they have no responsibility for providing accessible curricula, instead viewing this as the sole responsibility of disability services staff (Linder et al., 2015). This problem is exacerbated by the under-resourcing of disability support services as well as funding for making learning accessible through captioning, purchasing, and maintaining technologies,

and other accessibility resources (Behling and Linder, 2017; Linder et al., 2015). During the global COVID-19 pandemic, these barriers were exacerbated as university services shut down and students with disability experienced delays in accessing services and adjustments (Meleo-Erwin et al., 2021).

A range of barriers have been noted in relation to academic staff, including reports of some staff having:

- A lack of knowledge or awareness about the requirements of students with disability,
- A lack of skill/training in making the necessary adjustments,
- A belief that students with disability are not able to achieve at the same levels as their peers,
- A belief that more time is needed to support students with disability, which impacts workloads and positions students with disability as a burden,
- A belief that students with disability are already well supported by disability services, and it is beyond the scope of an academic's workload to offer further assistance, and
- A lack of will in accommodating the requirements of students with disability particularly when they perceive this offers the students an unfair advantage over their peers or diminishes the academic integrity of what is being assessed (Behling and Linder, 2017; Linder et al., 2015; Couzens et al., 2015; Baker et al., 2012; Yuknis, 2014; Fleming et al., 2016; Vickerman and Blundell, 2010; Moriña et al., 2014).

Pedagogical tensions sometimes surface as a barrier, where faculty members argue the traditional ways in which content is taught or assessed are pedagogically sound, and any changes to these approaches are seen as threatening to entrenched and deeply held views (Behling and Linder, 2017). Some students see these pedagogies as methodologies for teaching rather than methodologies for learning; that is, the focus is merely on the transmission of content (Moriña et al., 2014).

These barriers can be further exacerbated by professional registration bodies who specify inherent requirements students must be able to achieve in order to be job-ready at graduation, without consideration about whether these inherent requirements can be achieved with the support of reasonable adjustments in the workplace (Couzens et al., 2015). The issue of inherent requirements is further exacerbated by the lack of involvement of people with disability in determining inherent requirement statements, and considering how these might be met through the provision of reasonable adjustments (Johnston et al., 2016).

Facilitators to higher education for students with disability

Students have identified a range of facilitators to their success in higher education. Knowing their own strengths and needs, and being able to discuss these with disability support staff and academics is viewed as critical to their success (Ganguly et al., 2015). Also, knowing what supports are available, and being able to access direct assistance is also considered important (Ganguly et al., 2015; Zeng et al., 2018). Comprehensive support programs, including peer support networks, have also been identified as an important facilitator (Ganguly et al., 2015; Zeng et al., 2018).

Curriculum and pedagogical reform have been identified as critical facilitators. There is an onus on academics to pay attention to the course content, teaching and assessment methods to ensure that, where possible, the need for adjustments can be minimized (Bunbury, 2020). Simple content adjustments, such as providing alt text and image descriptions for students with vision impairment (Lewis, 2018) are still seen as the role of the Disability Services team rather than the responsibility of an academic. Additionally, pedagogical adjustments have been shown to assist students with disability access and participate in higher education. The recording of lectures, supplemented with captioning, assistive technology, and accessible notes, is critical to increasing access for students with disability, while also helpful for a range of their peers (Zeng et al., 2018; O'callaghan et al., 2017; Grimes et al., 2019). Flexibility in the ways in which students access content is an enabler, rather than rigid policies that mandate on-campus attendance in lecture theaters. During the global pandemic, universities have needed to be more flexible with all students regarding attendance and participation requirements and submission of assessment tasks, and consideration of ways to enhance digital interactions and social connectedness (Benito et al., 2021). There are opportunities here to learn from these lessons and implement universal curriculum, pedagogical and assessment reforms that benefit all learners, particularly those with disability.

The current lack of training to equip lecturers to teach students with disability is recognized internationally (see, for example, Burgstahler, 2015; Cunninghame et al., 2016; Fleming et al., 2016). Key findings from the Final Report on the 2020 Review of the Disability Standards for Education 2005 are consistent with these claims, reporting Australian educators, including those in higher education, are aware of their legal obligations to provide adjustments, but unclear as to how they should be implemented within their teaching practice (Australian Government [Department of Education Skills and Employment], 2021).

Universal Design for Learning (UDL) in higher education has been found to show promise in enabling students with disability greater access (Kimball et al., 2016; Roberts et al., 2011; Fleming et al., 2016; Edwards, 2019). UDL is a framework used by educators, underpinned by a set of principles to improve the ways in which content is transmitted, the ways in which students can demonstrate their knowledge, and the ways in which students engage during the teaching-learning transaction (Rogers-Shaw et al., 2018). Universally designed curriculum, teaching and assessment provide student experiences which address three broad objectives:

- (1) students should be able to interact with and respond to materials in multiple ways,
- (2) students should be able to find meaning and thus motivate themselves in different ways, and
- (3) ... course material must be accessible to all (Australian Disabilities Clearinghouse, n.d., para. 3)

These inclusive design ideals are student-centered and have the potential to benefit all students. The investigation by Dean et al. (2017) into the use of UDL within large marketing subjects found it to be highly effective, empowering students to take responsibility for their own learning as barriers commonly found in large class settings are decreased (p. 13). To employ such inclusive practice, however, requires a rethink of how learning is designed, delivered, and assessed.

Finally, the co-design of university spaces, drawing from the lived experiences of students with disability is a further enabler to the access and participation of learners with disability in higher education. Where attention has been paid to physical access with ramps, elevators and tactile signage, accessible laboratories and lecture theaters, assistive listening technology and tactile signage, attention now turns to sensory design as well as accessible materials and virtual spaces.

Conclusion

In 1998, Sheila Riddell encouraged those in higher education institutions supporting students with disability to continue to “chip away at the mountain” of barriers that block the full and active participation of students with disability. While the numbers of students with disability enrolled in higher education might suggest success in scaling the mountain, an abundance of research continues to point to the ongoing challenges along the journey (Fleming et al., 2016; Vickerman and Blundell, 2010; Griful-Freixenet et al., 2017; Moriña et al., 2014; Grimes et al., 2019; Verdinelli and Kutner, 2016; Strnadová et al., 2015).

In spite of the myriad of legislative and policy frameworks that enshrine the rights of people with disability to access education (including higher education), significant barriers still inhibit equal access (Poed, 2020). Werth et al. (2014) suggest that the inclusion of students with disability in higher education must extend beyond conversations that focus on disability services and supports, rights and reasonable adjustments, and begin to embrace “the promotion of values, frameworks, pedagogies, and practices that aim to bring about greater levels of normalization, and inclusion of students with a disability” (p. 107). It is time for higher education institutions to embrace inclusivity; students with disability enrich the learning environment making it accessible for, and inclusive of, all students (Moriña et al., 2020).

Acknowledgment

This research is an extension of research project funded by a Learning and Teaching Grant provided by the Faculty of Business and Economics at the University of Melbourne.

References

- Ashman, A., 2019. The foundations of inclusion. In: Ashman, A. (Ed.), *Education for Inclusion and Diversity*, sixth ed. Pearson, Melbourne, Australia.
- Australian Disabilities Clearinghouse, (n.d.). Universal Design. Available. <http://www.adcet.edu.au/disability-practitioner/course-design-and-implementation/universal-design/>.
- Australian Government [Department of Education Skills and Employment], 2021. Final Report—2020 Review of the Disability Standards for Education 2005.
- Australian Government [Federal Register of Legislation], 2017. Disability Discrimination Act 1992. Available. <https://www.legislation.gov.au/Details/C2016C00763>.
- Australian Government [Federal Register of Legislation], 2020. Disability (Access to Premises—Buildings) Standards. Available. <https://www.legislation.gov.au/Details/F2010L00668>.
- Australian Government [Department of Education Skills and Employment], 2020a. Disability Standards for Education 2005. Available. <https://www.dese.gov.au/swd/resources/disability-standards-education-2005-guidance-notes>.
- Australian Government [Department of Education Skills and Employment], 2020b. Selected Higher Education Statistics—2019 Student Data. Available. <https://www.dese.gov.au/higher-education-statistics/student-data/selected-higher-education-statistics-2019-student-data>. (Accessed 15 January 2021).
- Baker, K.Q., Boland, K., Nowik, C., 2012. A campus survey of faculty and student perceptions of persons with disabilities. *J. Postsecond. Educ. Disabil.* 25, 309–329.
- Barnes, C., 2007. Disability, higher education and the inclusive society. *Br. J. Sociol. Educ.* 28, 135–145.
- Behling, K., Linder, K.E., 2017. Collaborations between centers for teaching and learning and offices of disability services: current partnerships and perceived challenges. *J. Postsecond. Educ. Disabil.* 30, 5–15.
- Benito, A., Yenisey, K.D., Khanna, K., Masis, M.F., Monge, R.M., Tugtan, M.A., Araya, L.D.V., Vig, R., 2021. Changes that should remain in higher education post COVID-19: a mixed-methods analysis of the experiences at three universities. *High. Educ. Res. Commun.* 11, 51–75.
- Biggeri, M., Di Masi, D., Bellacicco, R., 2020. Disability and higher education: assessing students' capabilities in two Italian universities using structured focus group discussions. *Stud. High Educ.* 45, 909–924.
- Brandt, S., 2011. From policy to practice in higher education: the experiences of disabled students in Norway. *Int. J. Disabil. Dev. Educ.* 58, 107–120.
- Brett, M., 2016. Disability and Australian Higher Education: policy drivers for increasing participation. In: Harvey, A., Burnheim, C., Brett, M. (Eds.), *Student Equity in Australian Higher Education: Twenty-Five Years of A Fair Chance for All*. Springer Singapore, Singapore.
- Bunbury, S., 2020. Disability in higher education—do reasonable adjustments contribute to an inclusive curriculum? *Int. J. Incl. Educ.* 24, 964–979.
- Burgstahler, S., 2015. Opening doors or slamming them shut? Online learning practices and students with disabilities. *Soc. Incl.* 3, 69–79.
- Caleb, R., Barden, N., 2019. Higher education in the twenty-first century: changes and challenges for mental wellbeing. In: Barden, N., Caleb, R. (Eds.), *Student Mental Health and Wellbeing in Higher Education: A Practical Guide*. Sage, London, UK.
- Clark, C., Kusevskis-Hayes, R., Wilkinson, M., 2018. Enhancing student disclosure: Australia's invisible equity students and reasons for nondisclosure in Australia's tertiary sector. *J. Aust. N. Z. Stud. Serv. Assoc.* 26, 28–41.
- Couzens, D., Poed, S., Kataoka, M., Brandon, A., Hartley, J., Keen, D., 2015. Support for students with hidden disabilities in universities: a case study. *Int. J. Disabil. Dev. Educ.* 62, 24–41.
- Cunninghame, I., Costello, D., Trinidad, S., 2016. Issues and Trends for Students With Disability: Review of NCSEHE Funded Research.
- Dean, T., Lee-Post, A., Hapke, H., 2017. Universal design for learning in teaching large lecture classes. *J. Market. Educ.* 39, 5–16.
- Department of Education and Skills, 2001. *The Special Educational Needs and Disability Act*. HMSO, London.

- Disability Rights Commission, 2006. In: Disability Equality Duty. Disability Rights Commission.
- Dunn, D.S., Burcaw, S., 2013. Disability identity: exploring narrative accounts of disability. *Rehabil. Psychol.* 58, 148–157.
- Edwards, M., Poed, S., AL-Nawab, H., Penna, O., (Under review). Academic Accommodations for University Students Living With Disability and the Potential of Universal Design to Address Their Needs. Higher Education.
- Edwards, M., 2019. Inclusive learning and teaching for Australian online university students with disability: a literature review. *Int. J. Incl. Educ.* 1–16.
- Fleming, A.R., Oertle, K.M., Plotner, A.J., 2016. Student experiences: recommendations for improving postsecondary experiences of students with disabilities. *J. Postsecond. Educ. Disabil.* 30, 311–328.
- Fossey, E., Chaffey, L., Venville, A., Ennals, P., Douglas, J., Bigby, C., 2015. Supporting Tertiary Students With Disabilities: Individualised and Institution-Level Approaches in Practice. Available. <http://www.ncver.edu.au>.
- Gale, T., 2010. Putting social inclusion in its place: three principles of pedagogic work. In: Pathways 10 Conference: Creating Space for People With Disabilities in Tertiary Education. Southbank, Brisbane, Australia.
- Ganguly, R., Brownlow, C., Dupree, J., Graham, C., 2015. Resilience/Thriving in Post-secondary Students With Disability: An Exploratory Study. Curtin University [National Centre for Student Equity in Higher Education].
- Government Equalities Office & Equality and Human Rights Commission, 2015. Equality Act 2010: Guidance. Available. <https://www.gov.uk/guidance/equality-act-2010-guidance>.
- Grifol-Freixenet, J., Struyven, K., Verstichele, M., Andries, C., 2017. Higher education students with disabilities speaking out: perceived barriers and opportunities of the Universal Design for Learning framework. *Disabil. Soc.* 32, 1627–1649.
- Grimes, S., Southgate, E., Scevak, J., Buchanan, R., 2019. Learning impacts reported by students living with learning challenges/disability. *Stud. High Educ.* 1–13.
- Hartley, J., 2015. Australian higher education policy and inclusion of people with disabilities: a review. *J. Postsecond. Educ. Disabil.* 28, 413–419.
- Hernandez, M., 2011. People with apparent and non-apparent physical disabilities: well-being, acceptance, disclosure and stigma. In: Doctor of Philosophy. Alliant International University.
- Home Office, 1995. The Disability Discrimination Act. HMSO, London.
- Jacklin, A., 2011. To be or not to be “a disabled student” in higher education: the case of a post-graduate “non-declaring” (disabled) student. *J. Res. Spec. Educ. Needs* 11, 99–106.
- Johnston, K.N., Mackintosh, S., Alcock, M., Conlon-Leard, A., Manson, S., 2016. Reconsidering inherent requirements: a contribution to the debate from the clinical placement experience of a physiotherapy student with vision impairment. *BMC Med. Educ.* 16, 74.
- Kanter, A.S., 2020. The failure of the United States to ratify the CRPD. In: Kakoulis, E.J., Johnson, K. (Eds.), *Recognizing Human Rights in Different Cultural Contexts: The United Nations Convention on the Rights of Persons With Disabilities (CRPD)*. Springer Singapore, Singapore.
- Kent, M., Ellis, K., Giles, M., 2018. Students with disabilities and eLearning in Australia: experiences of accessibility and disclosure at Curtin University. *TechTrends* 62, 654–663.
- Kimball, E.W., Wells, R.S., Ostiguy, B.J., Manly, C.A., Lauterbach, A.A., 2016. Students with disabilities in higher education: a review of the literature and an agenda for future research. In: Paulsen, M.B. (Ed.), *Higher Education: Handbook of Theory and Research*. Springer International Publishing, Cham.
- Konur, O., 2006. Teaching disabled students in higher education. *Teach. High. Educ.* 11, 351–363.
- Lewis, V., 2018. How to Write Alt Text and Image Descriptions for the Visually Impaired. Available. <https://www.perkinselearning.org/technology/blog/how-write-alt-text-and-image-descriptions-visually-impaired>.
- Linder, K.E., Fontaine-Rainen, D.L., Behling, K., 2015. Whose job is it? Key challenges and future directions for online accessibility in US institutions of higher education. *Open Learn.* 30, 21–24.
- Madaus, J.W., Kowitt, J.S., Lalor, A.R., 2012. The higher education opportunity act: impact on students with disabilities. *Rehabil. Res. Policy Educ.* 26, 33–41.
- Markoulakis, R., Kirsh, B., 2013. Difficulties for university students with mental health problems: a critical interpretive synthesis. *Rev. High. Educ.* 37, 77–100.
- Meleo-Erwin, Z., Kolia, B., Fera, J., Jahren, A., Basch, C., 2021. Online support information for students with disabilities in colleges and universities during the COVID-19 pandemic. *Disabil. Health J.* 14. <https://doi.org/10.1016/j.dhjo.2020.101013>.
- Moriña, A., Cortés, M.D., Melero, N., 2014. Inclusive curricula in Spanish higher education? Students with disabilities speak out. *Disabil. Soc.* 29, 44–57.
- Moriña, A., Sandoval, M., Carnero, F., 2020. Higher education inclusivity: when the disability enriches the university. *High Educ. Res. Dev.* 39, 1202–1216.
- O’callaghan, F.V., Neumann, D.L., Jones, L.J., Creed, P.A., 2017. The use of lecture recordings in higher education: a review of institutional, student, and lecturer issues. *Educ. Inf. Technol.* 22, 399–415.
- Podzo, B.Z., Chipika, C.G., 2019. Curriculum reform: a key driver to the inclusion of students with disabilities in higher education. *J. Educ. Pract.* 10, 59–67. <https://core.ac.uk/download/pdf/234642414.pdf>.
- Poed, S., 2016. Adjustments to Curriculum for Australian School-Aged Students With Disabilities: What’s Reasonable? Doctor of Philosophy Griffith University.
- Poed, S., 2020. Social inclusion and equal access to high-quality, inclusive education. In: Leal Filho, W., Azul, A.M., Brandli, L., Özuyar, P.G., Wall, T. (Eds.), *Quality Education*. Springer International Publishing, Cham.
- Reutlinger, C.J., 2015. The Ableist Othering of Disability in the Classroom: An Experiential Investigation of Academic Adjustments in Higher Education. Master of Arts. Kansas State University.
- Riddell, S., Tinklin, T., Wilson, A., 2005. *Disabled Students in Higher Education: Perspectives on Widening Access and Changing Policy*. RoutledgeFalmer, London, England.
- Riddell, S., 1998. Chipping away at the mountain: disabled students’ experience of higher education. *Int. Stud. Sociol. Educ.* 8, 203–222.
- Roberts, K.D., Jin Park, H., Brown, S., Cook, B., 2011. Universal design for instruction in postsecondary education: a systematic review of empirically based articles. *J. Postsecond. Educ. Disabil.* 24, 5–15.
- Rogers-Shaw, C., Carr-Chellman, D.J., Choi, J., 2018. Universal design for learning: guidelines for accessible online instruction. *Adult Learn.* 29, 20–31.
- Sarkar, R., 2016. Inclusive University: a way out to ensure quality, equity, and accessibility for students with disabilities in higher education. *Int. J. Adv. Res.* 4, 406–412.
- Slee, R., 2018. Setting the scene for people with disability to experience university life. In: O’Brien, P., Bonati, M.L., Gadow, F., Slee, R. (Eds.), *People With Intellectual Disability Experiencing University Life*. Brill.
- Strnadová, I., Hájková, V., Květoňová, L., 2015. Voices of university students with disabilities: inclusive education on the tertiary level—a reality or a distant dream? *Int. J. Incl. Educ.* 19, 1080–1095.
- Taylor, A., Shallish, L., 2019. The logic of bio-meritocracy in the promotion of higher education equity. *Disabil. Soc.* 34, 1200–1223.
- United Nations [Committee on the Rights of Persons With Disabilities], 2016. General Comment No. 4, Article 24: Right to Inclusive Education. Available. <https://www.refworld.org/docid/57c977e34.html>. (Accessed 15 October 2020).
- United Nations Educational Scientific And Cultural Organisation, 1999. Provisions for Students With Disabilities in Higher Education: A Survey.
- United Nations, 2006. Conventions on the Rights of Persons With Disabilities (CRPD). Available. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>. (Accessed 15 January 2021).
- United States [Department of Justice—Civil Rights Division], (n.d.). Introduction to the ADA. Available. https://www.ada.gov/ada_intro.htm.
- Verdinelli, S., Kutner, D., 2016. Persistence factors among online graduate students with disabilities. *J. Divers. Higher Educ.* 9, 353–368.
- Vickerman, P., Blundell, M., 2010. Hearing the voices of disabled students in higher education. *Disabil. Soc.* 25, 21–32.
- Werth, S., Hammer, S., D’abadie, D., 2014. Disability and study: layers of management. *Int. J. Pedagog. Learn.* 9, 101–109.
- Yuknis, C., 2014. Removing the disability from distance education. In: Yuzer, T., Kurubacak, G. (Eds.), *Handbook of Research on Emerging Priorities and Trends in Distance Education: Communication, Pedagogy, and Technology*. IGI Global.

- Yusuf, S., Saint, W., Nabeshima, K., 2009. Accelerating catch-up: tertiary education for growth in sub-Saharan Africa. VOCEDplus. <https://openknowledge.worldbank.org/bitstream/handle/10986/2589/462750Revised017808213773830Revised.pdf?sequence=1&isAllowed=y>.
- Zeng, W., Ju, S., Hord, C., 2018. A literature review of academic interventions for college students with learning disabilities. *Learn. Disabil. Q.* 41, 159–169.

Further reading

- Barkas, L.A., Armstrong, P.-A., Bishop, G., 2020. Is inclusion still an illusion in higher education? Exploring the curriculum through the student voice. *Int. J. Incl. Educ.* 1–16. <https://doi.org/10.1080/13603116.2020.1776777>.
- Griful-Freixenet, J., Struyven, K., Verstichele, M., Andries, C., 2017. Higher education students with disabilities speaking out: perceived barriers and opportunities of the Universal Design for Learning framework. *Disabil. Soc.* 32 (10), 1627–1649. <https://doi.org/10.1080/09687599.2017.1365695>.
- Kutscher, E.L., Tuckwiller, E.D., 2019. Persistence in higher education for students with disabilities: a mixed systematic review. *J. Divers. Higher Educ.* 12 (2), 136–155. <https://doi.org/10.1037/dhe0000088>.
- McNicholl, A., Casey, H., Desmond, D., Gallagher, P., 2021. The impact of assistive technology use for students with disabilities in higher education: a systematic review. *Disabil. Rehabil. Assist. Technol.* 16 (2), 130–143. <https://doi.org/10.1080/17483107.2019.1642395>.
- Moriña, A., Sandoval, M., Carnerero, F., 2020. Higher education inclusivity: when the disability enriches the university. *High Educ. Res. Dev.* 39 (6), 1202–1216. <https://doi.org/10.1080/07294360.2020.1712676>.
- Newman, L.A., Madaus, J.W., Lalor, A.R., Javitz, H.S., 2020. Effect of accessing supports on higher education persistence of students with disabilities. *J. Divers. Higher Educ.* <https://doi.org/10.1037/dhe0000170>.
- Slee, R., 2018. Setting the scene for people with disability to experience university life. In: O'Brien, P., Bonati, M.L., Gadow, F., Slee, R. (Eds.), *People With Intellectual Disability Experiencing University Life*. Brill, pp. 39–48. https://doi.org/10.1163/9789004394551_003.
- Smith, S.A., Woodhead, E., Chin-Newman, C., 2019. Disclosing accommodation needs: exploring experiences of higher education students with disabilities. *Int. J. Incl. Educ.* 1–17. <https://doi.org/10.1080/13603116.2019.1610087>.
- Stentiford, L., Koutsouris, G., 2020. What are inclusive pedagogies in higher education? A systematic scoping review. *Stud. High Educ.* 1–17. <https://doi.org/10.1080/03075079.2020.1716322>.
- Thurber, A., Bandy, J., 2018. Creating Accessible Learning Environments. Vanderbilt University Center for Teaching. <https://cft.vanderbilt.edu/guides-sub-pages/creating-accessible-learning-environments/>.

Relevant websites

- Australian Disability Clearinghouse on Education and Training—Disability Services University: If you are an Australian student with disability, enrolled in higher education, this site provides a summary of the support services available across all Australian universities, <https://www.adcet.edu.au/students-with-disability/current-students/disability-services-university>.
- How to make your teaching inclusive: This website, designed by the Anglia Ruskin University, uses the IMPACT framework to help higher education staff make their teaching more inclusive, https://web.anglia.ac.uk/inclusive_practice/getting_started/getting_started_00.html.
- Reasonable accommodations explained: As part of the DisABILITY Resources Toolkit, this site explains reasonable accommodations, and offers additional resources and case examples, <https://www.apa.org/pi/disability/dart/toolkit-three>.
- Accommodations for students with disability: While written with school-aged children in mind, this online module offers insights into the array of accommodations students might require to support their learning, <https://iris.peabody.vanderbilt.edu/module/acc/cresource/q1/p01/#content>.
- Scholarships for students with disability: This article was compiled by Sabrina Collier on April 4, 2021, and documents scholarships from around the world that enable students with disability to access higher education, <https://www.topuniversities.com/student-info/scholarship-advice/scholarships-students-disabilities>.
- How to succeed in college with a disability: This article, written by Stephen Bird (2017), offers some important strategies to assist students with disability to succeed at college, <https://theconversation.com/how-to-succeed-in-college-with-a-disability-79226>.
- Universal Design for Learning in Higher Education: This website provides an overview of the Universal Design for Learning Framework and how it can be used in higher education settings, http://udloncampus.cast.org/page/udl_landing.

Reviews of education for students with disability in Australia

Sam Brain^a and Rachael Mariani^b, ^a Deloitte Access Economics, Melbourne, VIC, Australia; and ^b Victorian Department of Education and Training, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	376
The Australian context	376
Australian reviews and inquiries	377
Review and inquiry findings	378
Student segregation	380
Student exclusion	380
Integration	382
Alternative curriculum	382
Review design considerations	382
Independent review processes	383
Clear ownership of recommendations	383
Accountability for acting on recommendations	383
Conclusion	384
References	384

Introduction

Australian governments are committed to improving the participation, wellbeing and academic outcomes of the sizable portion of the Australian school student population who have a disability. The past decade has seen several significant advancements in education policy relating to students with disability, including the adoption of a nation-wide practice model, a shift to a needs-based funding model and the creation of the Queensland inclusive education statement.

However, the available evidence suggests that, over this period, there has been limited meaningful advancement in the outcomes and experiences of students with disability. Drawing on evidence tabled during recent reviews and inquiries into the experience of students with disability, this article highlights common issues associated with the schooling experience of these students in Australia.

Noting these issues, and their persistent nature, the article examines the barriers which impede the adoption of more inclusive education structures, methods and practice. The article then provides concluding remarks and discusses the role of future research and inquiry in helping Australian education jurisdictions dismantle these barriers.

The Australian context

Within Australia's federated model of governance, there are multiple layers of responsibility for schooling. Australian state and territory governments are responsible for delivering school education, including registering, regulating and operating government schools and they provide the majority of funding for the operations of government schools. The Australian Federal Government provides additional funding to government schools and is the primary government funder of non-government schools.

The Australian Government influences education policy through the provision of conditional funding. This typically takes the form of large-scale initiatives which tie the provision of federal funding to participation in nation-wide initiatives or reforms. Australia has a national curriculum, national assessment program and national school data collection, which are conducted by the independent statutory agency the Australian Curriculum, Assessment and Reporting Authority (ACARA). Major reforms to education are also progressed by the Education Council, which is a body comprised of education ministers from the Australian, state and territory governments. The Education Council helps to shape policy and decision-making across all school systems within Australia.

Over recent decades, particularly since the publication of the Salamanca Statement of Special Needs Education (the Salamanca Statement, UNESCO, 1994) and the United Nations Convention of the Rights of Persons with Disability (CRPD, 2007), Australian governments have demonstrated their commitment to improving the participation, wellbeing and academic outcomes of students with disability.

Efforts to improve the educational experiences of students with disability have largely centered on the introduction of rights-based legislation, and aspirational and binding policies. There have also been substantial developments in policy and funding

Table 1 International and domestic legislative instruments informing education for students with disability.

<i>Instrument</i>	<i>Relevance to education of students with disability</i>
Disability Discrimination Act (1992)	The Commonwealth Disability Discrimination Act of 1992 (DDA) legislated the right of all students to enrol in their local school (Attorney-General Department, n.d.).
Salamanca Statement & Framework for Action (1994)	In 1994, Australia endorsed the Salamanca statement & framework for action (1994) that advocates governments to endorse inclusive education in recognition of its value for providing a quality education and building a more inclusive society. The statement lacked an authoritative definition of inclusive education, and was an aspirational document rather than a binding one (Ainscow et al., 2019).
Disability Standards for Education (2005)	The Disability Standards for Education 2005 (DSE) clarified the rights and responsibilities of education providers under the DDA, and introduced the concepts of “reasonable adjustments” and “unjustifiable hardship” into the Australian policy framework (Urbis, 2015).
United Nations Convention on the Rights of Persons With Disabilities (2006)	In 2006, the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) was published and has since been ratified by 177 signatory countries, including Australia. Article 24 of the UNCRPD articulates the responsibility of governments to pursue an inclusive education (Committee on the Rights of Persons with Disabilities, 2016).
Melbourne Declaration (2008)	In 2008, Education Ministers across Australia committed to a collective aspiration for an excellent and equitable education system. The statement outlines that Australian governments must provide all students with access to high-quality schooling that is free from discrimination based on disability (Ministerial Council on Education Employment Training and Youth Affairs, 2008).
UNCRPD General Comment 4 (2016)	In 2016, the Committee on the Rights of Persons with Disabilities (CRPD) produced general comment 4, a companion document for the implementation of article 24 of the United Nations Convention on the Rights of Persons with Disabilities (2006). General comment 4 provides a definition of inclusion education alongside other models of education. General comment 4—which itself has not been ratified by governments—outlines that inclusion “involves a process of systematic reform embodying changes and modification in content, teaching methods, approaches, structures and strategies in education to overcome barriers” (Committee on the Rights of Persons with Disabilities, 2016).
Alice Springs (mparntwe) Education Declaration (2019)	Following a Nation-wide consultation strategy to refresh the 2008 Melbourne declaration, Education Ministers from across Australia again made a commitment to an education system promoting excellence and equity. The declaration specified that these aspirations meant tailoring to the needs of individual students, including those with disability (Education Council, 2019).

at national and state levels, such as the development of the Nationally Consistent Collection of Data (NCCD) on School Students with Disability and the introduction of a national funding loading for students with disability ([National School Resourcing Board, 2020](#)).

In addition to state and Australian Government legislation, the Australian Government has endorsed and ratified a range of international frameworks to commit to improving educational experiences and outcomes of students with disability. It has also developed a range of supporting domestic legislation and policy to align with these commitments. These are listed in [Table 1](#) below.

These instruments provide a legislative basis for all students to receive an inclusive education with reasonable adjustments in regular school settings. Indeed, families can take education providers to court for failing to provide reasonable adjustments—using the DDA and DSE as the legislative instruments. However, these have not proven effective at changing practice at a systemic level, or promoting a shift to inclusive, regular schools ([de Bruin, 2019](#)). There is no Australia-wide policy which promotes regular school settings as the recommended place for all Australian school students, and there is no inclusive education policy at a national level. The Australian Government does not recognize General Comment 4 as a binding document, noting that it takes General Comment 4 into account when considering its obligations under article 24.

Australian reviews and inquiries

The experiences of students with disability in schools have been investigated in 16 reviews and inquiries at the state and Australian government level over the last decade. These reviews have been instigated by a range of factors:

- Increasing public awareness of the rights of students with disability. Advocacy organizations such as Children and Youth with Disability Australia, All Means All and Family Advocacy NSW have raised public awareness of the experiences of students with disability.

- Impetus to reform. In 2012, the Australian Government introduced the More Support for Students with Disabilities (MSSD) initiative. This promoted the adoption of evidence-based practices and led to a wave of reforms across Australian jurisdictions. It also established a vehicle for states and territories to collaborate on improving educational provision.
- Public awareness and interest following specific incidents. The 2015 *Schools for All Children & Young People* review in the Australian Capital Territory (ACT) was instigated following a high-profile case in a Canberra public school (Shaddock et al., 2016).

These reports considered a range of issues relating to the experiences of students with disability. Reports have distinct terms of reference and oversight arrangements. Parliamentary inquiries, for example, are chaired by parliamentarians and may develop findings and recommendations which are political in context. Across these varying contexts, findings within these reports tend to relate to the experience of students with disability, the practices that support them, and the policy and legislative structures that promote continuing improvement in practice and outcomes.

Table 1 below provides an overview of reviews and inquiries into the experiences of students with disability in Australia between 2009 and 2020. The most recent publicly available inquiry conducted in Western Australia was completed in 2004, and so to ensure coverage across all states and territories, this has been included (Table 2).

The extent to which the recommendations of these reviews and inquiries have been successfully implemented varies. Certainly, there are some notable examples of progress:

- Victorian Functional Needs Assessment—an assessment of student needs which aims to shift the focus of categorical assessment away from medical diagnosis toward the functional needs of the student and the school environment (Victorian Department of Education and Training, 2018).
- The Queensland Inclusive Education policy—an aspirational statement that sets out Queensland's commitment for success and learning among all students. This statement is unique among Australian statements in that it specifies and defines what educational practices are deemed *not inclusive*. Definitions of integration, segregation and exclusion, based on General Comment 4, are provided (Queensland Department of Education, 2018).

However, in many cases progress is difficult to observe and the issues identified in the reports persist.

To increase the pace of progress and improve levels of coordination, Australia's Education Council formed the *Joint Working Group to Provide Advice on Reform for Students with Disability* (Joint Working Group). Since the creation of the Joint Working Group in 2011, there have been some significant advances across Australian school systems which reflect advances in legislation, financing and policy for students with disability.

- The More Support for Students with Disabilities (MSSD) initiative provided short term funding to education authorities in each state and territory over the 2012–2014 school years. It was established to strengthen the capacity and expertise of Australian schools and teachers to provide additional support to students with disability. The initiative supported education systems to trial innovative strategies to achieve the objective with the purpose of stimulating new approaches to inform future developments, and contributed to ongoing streams of reform across school systems (PhillipsKPA, 2015).
- The Nationally Consistent Collection of Data (NCCD) on school students with disability is a framework providing the basis for assessing and financing the provision of adjustments for students with disability in schools (de Bruin et al., 2020). The NCCD was developed collaboratively across all Australian governments, borne out of an agreement that a consistent definition of disability within the school education context was required across States and Territories. The NCCD provides a definition based on the broad category of disability a student has (physical, cognitive, sensory or behavioral) and the extent of adjustments which are required in the classroom as a result (supplementary, substantial and extensive). It also provides the basis for schools to consider how classrooms are structured to support as many learners as possible without the need for additional adjustments, through quality differentiated teaching practice.

Following sustained pressure from Australian advocacy and rights-based organizations, a Royal Commission into Violence, Abuse, Neglect and Exploitation of People with Disability was established in 2019. This Commission, which hereafter in this article is referred to as the Disability Royal Commission, explicitly considers the role of education as both an institution that has historically inhibited the rights of people with disability, as well as a potential force in driving inclusion (Royal Commission Into Violence Abuse Neglect and Exploitation of People With Disability, 2019).

Review and inquiry findings

The Australian legislative framework provides a basis for all students to receive education with reasonable adjustments in regular school environments. While the findings of Australian reviews and inquiries about students with disability reveal examples of inclusive cultures and practices in regular schools, they also provide extensive evidence that this does not often occur in practice (Iacono et al., 2019).

Ultimately, Australia is yet to see a substantial shift toward inclusive education in every school, or for every student—and there remains wide variability in quality among school practices and student experiences. This variation is seen between schools and

Table 2 Overview of Australian reviews describing the experience of students with disability.

<i>Review title</i>	<i>Year</i>	<i>Author</i>	<i>References</i>
Australian Government			
Review of the Disability Standards for Education (2015)	2015	Urbis	Urbis (2015)
Access to Real Learning: The Impact of Policy, Funding and Culture on Students With Disability	2016	Australian Senate	Australian Senate (2016)
Violence, Abuse and Neglect Against People With Disability in Institutional and Residential Settings	2015	Australian Senate	Australian Senate (2015)
Review of the Disability Standards for Education (2010)	2012	Department of Education, Employment and Workplace Relations	DEEWR (2010)
New South Wales			
Students With a Disability or Special Needs in NSW Schools	2017	NSW Legislative Council	New South Wales Legislative Council (2017)
Victoria			
Review of the Program for Students With Disabilities	2015	Department of Education and Training, Victoria.	Victorian Department of Education and Training (2015)
Held Back: The Experiences of Students With Disability in Victorian Schools	2012	Human Rights and Equal Opportunities Commission, Victoria.	Human Rights and Equal Opportunities Commission Victoria (2012)
Inquiry Into Services for People With Autism Spectrum Disorder	2017	Parliament of Victoria	Parliament of Victoria (2017)
Investigation Into Victorian Government School Expulsions	2017	Victorian Ombudsman	Glass (2017)
Queensland			
Review of Education for Students With Disability in Queensland State Schools	2017	Deloitte Access Economics	Deloitte Access Economics (2017)
South Australia			
Report on the Select Committee on Access to the SA Education System for Students With Disability	2017	Parliament of South Australia	Parliament of South Australia (2017)
Inquiry Into Suspension, Exclusion, and Expulsion Processes in South Australian Government Schools	2020	Graham, McCarthy, Killingly, Tancredi, & Poed	Graham et al. (2020)
Western Australia			
Pathways to the Future: Review of Educational Services for Students With Disabilities in Government Schools	2004	Department of Education and Training	Western Australia Department of Education and Training (2004)
Tasmania			
Improved Support for Students With Disability Ministerial Taskforce Report	2015	Minister for Education and Training	Tasmanian Department of Education and Training (2015)
Northern Territory			
Review of Policy and Practice for Students With Additional Needs	2018	Center for International Research on Education Systems	Center for International Research on Education Systems (2018)
Australian Capital Territory			
Schools for All Children & Young People—Report to ACT Govt From the Expert Panel on Students With Complex Needs and Challenging Behavior	2015	Shaddock, Packer and Roy	Shaddock et al. (2016)
Disability, Diversity and Tides that Lift All Boats: Review of Special Education in the ACT	2009	Anthony Shaddock, Nancy MacDonald, Julie Hook, Loretta Giorcelli, Michael Arthur-Kelly	Shaddock et al. (2009)

within schools. Consideration of reviews and inquiries highlights three consistent themes which align with the framework of practice presented by the CRPD in General Comment 4 and which serve as a basis for summarizing the persistent issues limiting the achievement of inclusive education in Australia.

- Students with disability are segregated. Students—and their families—are either coerced or opt into a segregated system of education.
- Students with disability experience exclusionary practices when accessing or participating in schools, which lead to them attending no school at all, attending part time, or not having the extent of choice which other students in Australia have.

- The model of education experienced by many students with disability is integration. Across schools, there is a failure to provide the reasonable adjustments required for students to receive a fully inclusive education alongside their peers.

This section describes school practices as experienced by students with disability, drawing on the evidence from reviews and inquiries, mapped to these three overarching themes. Proposed solutions offered by reviews and inquiries are also outlined.

Student segregation

Many students with disability participate in separate educational environments to other Australian students, with the likelihood of this increasing as students progress through their schooling. The Australian Bureau of Statistics (ABS) *Survey of Disability Aging and Carers 2018* estimates that approximately 1 in 40 Australian students attend either a special school or class (Australian Bureau of Statistics, 2018). The available data also confirms that participation in mainstream classes decreases as students progress through the school years. For example, the 2018 National Disability Insurance Scheme outcomes report found that 75% of students with disability attended a mainstream school in years Foundation–2, with this figure reducing to 65% in years 3–6 and to 42% for years 7–10 (National Disability Insurance Scheme, 2019).

It is well established that the school setting shapes an individual's educational experiences and outcomes (Hehir et al., 2016). However, there is a consistent lack of evidence to suggest that segregated schooling is more beneficial than inclusive settings for students with disability. Within an Australian context, no reviews or inquiries to date have published original research into the impact of segregated schooling on student outcomes. All Australian jurisdictions continue to offer segregated education for students with disability.

To date, the terms of reference for inquiries and reviews have not explicitly considered the prevalence or impact of segregated schooling on subsequent life outcomes. Following sustained pressure from advocacy and rights-based organizations and an increasing awareness of the issues associated with the practice, student segregation was brought into consideration of the Disability Royal Commission (Royal Commission Into Violence Abuse Neglect and Exploitation of People With Disability, 2019). It is expected that over its course, the Commission will consider the role of segregation in Australian education systems and prepare recommendations for future consideration for all Australian governments.

Issues raised in reviews and inquiries

Reviews and inquiries have indicated that a parent/carer's decision to place their child in a special school or program may be driven by the absence of tenable options within the mainstream education system (Australian Senate, 2016). The current Australian emphasis on parent/carer choice in education settings presupposes that families have access to meaningful information that informs their choice. In reality, however, families may not have access to relevant information, such as in relation to the benefits of inclusive education (Deloitte Access Economics, 2017; Tah, 2020; Mann et al., 2015). Further, when parents experience exclusionary practices from schools, such as enrollment dissuasion (also known as gatekeeping), their perceptions of attractiveness and viability of mainstream options are impacted (Mann et al., 2015).

Recommendations

To enable informed and tenable family choices, reviews and inquiries have recommended improving school community understanding of the benefits of inclusive education—to better enable informed choices among students and their families (Deloitte Access Economics, 2017). Reviews have also emphasized that mainstream education should be conveyed as the primary education placement for *all* students (Victorian Department of Education and Training, 2015; Deloitte Access Economics, 2017; New South Wales Legislative Council, 2017; Center for International Research on Education Systems, 2018).

Recommendations have also focused on redirecting resources to ensuring mainstream school systems can become more inclusive so as to improve educational options for families of students with disability (DEEWR, 2010; Victorian Department of Education and Training, 2015; Deloitte Access Economics, 2017).

Student exclusion

Exclusionary practices encapsulate a range of discriminatory practices that limit equitable access to education for students with disability. Exclusionary practices apply to access and engagement at all stages of a student's educational journey. These include discriminatory enrollment practices, disproportionate disciplinary practices, and restrictive practices (restraint and seclusion).

Issues raised in reviews and inquiries

Australian reviews and inquiries indicate that exclusionary practices are commonly—and in some cases routinely—applied by schools across a student's educational trajectory. However, rigorous data does not exist to quantify the incidence of these practices. A yearly survey conducted by Children and Young People with Disability Australia reports that exclusionary practices take place in all states and territories (Children and Young People with Disability Australia, 2019).

Evidence from reviews and inquiries also sheds light on the attitudes, perceptions and knowledge of educators that underpin these exclusionary practices. In particular, low expectations of students with disability were regularly cited in reviews and inquiries as contributing to exclusionary enrollment practices, while schools' insufficient understanding of how a student's disability

influences their behavior is noted in reference to the use of suspension, expulsion and restrictive practices. It should be noted, however, that no terms of reference in recent reviews and inquiries explicitly considered educator workforce attitudes or school culture.

Enrollment practices

Upon seeking to enroll within a school, families of students with disability may experience subtle or explicit dissuasion to enroll, a practice known as “gatekeeping.” This can manifest in multiple ways, including granting enrollment on a conditional basis only (such as forced part-time enrollment) or through school staff dissuading a parent/carer from enrolling their child by referencing the school’s resourcing limitations, including referencing catchment zone limitations or subtly expressing low expectations of students with disability (Castan Center for Human Rights Law, 2018; Australian Senate, 2016). There is an inequitable onus on families of students with disability to navigate the education system, secure education support, and advocate for necessary supports for their child (Human Rights and Equal Opportunities Commission Victoria, 2012; Parliament of Victoria, 2017).

Disciplinary practices, such as suspension (temporary removal from a school environment) and expulsion (permanent removal from a school environment) can be detrimental to the mental health, wellbeing and social adjustment of students with disability. Students and their families, along with peak bodies, reported that suspending and expelling students with disability reinforces difference, segregation and stigma associated with ability (Hehir et al., 2016; Human Rights and Equal Opportunities Commission Victoria, 2012; Parliament of South Australia, 2017). The 2015 *DSE Review* noted that suspension or expulsion from the learning environment impacts on the student’s continued learning and integration into the social structures (Urbis, 2015).

In addition to the impacts on students with disability and their families, several reviews noted that traditional forms of disciplinary practice were ineffective at changing behavior (Shaddock et al., 2016; Deloitte Access Economics, 2017; Parliament of South Australia, 2017). Consultation with one teacher for the *Queensland Review of education for students with disability* found that “suspensions give us a small amount of respect, but they don’t resolve the issues” (Deloitte Access Economics, 2017).

The immediate inappropriate response of punishing a student with disability for challenging behavior can be circumvented through a shift in perception and behavior—that is, by ascertaining the underlying causes of challenging behavior and putting in place the appropriate supports to address these (Sackville, 2020).

Restrictive practices

While reviews and inquiries stress that restrictive practices should be reserved as a “last resort” intervention, to be used only once all other options have been exhausted and to protect the safety of staff and students, they also found many instances of restrictive practice being misused (DEEWR, 2010; Deloitte Access Economics, 2017; Australian Senate, 2015). The use of restrictive practices is widely known to be an ineffective intervention in instigating behavior change and contravenes the rights of the young person (Parliament of South Australia, 2017; Iacono et al., 2019; Poed et al., 2020). A common theme in reviews, inquiries and indeed contemporary Australian literature is that restrictive practice is misused as a disciplinary, or behavior management, tool—that is, in some cases, normalized (Australian Senate, 2015; Human Rights and Equal Opportunities Commission Victoria, 2012; Australian Senate, 2016; Poed et al., 2020). Surveys estimate that between 16% (Children and Young People with Disability Australia, 2016) and 25% (Poed et al., 2020) of students with disability experience restrictive practices.

The prevalence of restrictive practice misuse is attributed to limited system-level guidance and leadership (Deloitte Access Economics, 2017; Human Rights and Equal Opportunities Commission Victoria, 2012; Parliament of South Australia, 2017; Castan Center for Human Rights Law, 2018; Poed et al., 2020). Specifically, there is insufficient guidance on how and when restraint can be used in schools.

Recommendations

Exclusionary practices are in breach of national and international obligations, reflect low expectations for the potential achievements of students with disability, and are enabled by the limited guidance, oversight and accountability mechanisms at the state/territory level and more broadly at a national level.

Recommendations outlined in reviews and inquiries focus on levers relating to accountability and governance and policy guidelines. Specifically, reviews and inquiries recommend:

- Improved collection and analysis of disaggregated data on suspensions and expulsions, part-time enrollment arrangements, and restrictive practice (Deloitte Access Economics, 2017; Australian Senate, 2016)
- Adoption of national rules around excluding students with disability within schools (Australian Senate, 2015)
- Improved information sharing of parents/carers’ rights, obligations and complaints processes regarding enrollment (Tasmanian Department of Education and Training, 2015; Parliament of South Australia, 2017; Deloitte Access Economics, 2017; Human Rights and Equal Opportunities Commission Victoria, 2012; Urbis, 2015)
- Adoption of clear principles and guidelines regarding the use of restrictive practice, and minimization and eventual elimination of these practices through a revision of system policy and workforce capacity (Deloitte Access Economics, 2017; Australian Senate, 2015; Parliament of South Australia, 2017; Australian Senate, 2016; Human Rights and Equal Opportunities Commission Victoria, 2012)

- Adoption of clear guidelines regarding suspension and expulsion of students with disability, and increased accountability of the discriminatory use of these practices (Glass, 2017; Parliament of South Australia, 2017; Human Rights and Equal Opportunities Commission Victoria, 2012).
- Adoption of innovative approaches that focus on prevention and limiting the need for disciplinary and restrictive practice, such as schoolwide positive behavior support (Glass, 2017; Parliament of South Australia, 2017; Human Rights and Equal Opportunities Commission Victoria, 2012).

Integration

The CRPD defines integration as the process of “... placing persons with disabilities in existing mainstream educational institutions, as long as (students with disabilities) can adjust to the standardized requirements of such institutions” (Committee on the Rights of Persons with Disabilities, 2016). The concepts of integration and inclusion are not always clearly distinguished from each other and are often used interchangeably (UNICEF, 2012). Across the course of their educational experience, students may experience aspects of their education which resemble inclusion and integration.

Integration can contribute to poor experiences of students with disability and their peers. An outcome of integration is that students opt to transition into segregated environments or leave education entirely (Mann et al., 2015).

Issues raised in reviews and inquiries

Evidence from reviews, inquiries and contemporary literature indicates that students with disability who participate in education within a mainstream setting may not receive adequate support and adjustments necessary to truly access and engage with the educational program. This means that the form of education they are experiencing may be closer to integration than inclusion, using the General Comment 4 model.

One such issue is the excessive use of alternative curriculum. This issue is illustrative of the experience of integration in an educational setting, however is by no means exhaustive.

Alternative curriculum

All schooling systems across Australia either implement the Australian curriculum or a similar curriculum system. For students with disability, it is considered effective practice to develop adjustments where students access the same curriculum as their similar-aged peers. However, where students struggle to access despite the provision of reasonable adjustments, it is also considered appropriate to deliver curriculum which is selected from a different year level to the student, or selected from the general dimensions to the curriculum (Swancutt et al., 2019).

However, the use of alternative and modified curriculum and assessment practices has been raised as a key concern in reviews and inquiries (Urbis, 2015; Castan Center for Human Rights Law, 2018; Iacono et al., 2019; Deloitte Access Economics, 2017; Parliament of South Australia, 2017). For example, there is evidence that instead of modifying curricula to suit individual needs and capabilities, some teachers merely reduce the difficulty and rigor of work tasks or reduce workloads, which can result in learning activities that do not hold educational value (Urbis, 2015; Castan Center for Human Rights Law, 2018).

Over time, these practices can have significant impacts on student life outcomes. Employment opportunities may be limited by a focus on “special” curriculum that is exclusively developed for students with disability, which can eventually lead to social exclusion, lower levels of health and an overall life of disadvantage (European Agency for Special Needs and Inclusive Education, 2018; Castan Center for Human Rights Law, 2018).

Recommendations

While these adjustments may be implemented in line with an agreed method of practice, it is important to know which students are receiving modified curriculum, and ideally why. At present there is no national view on the distribution of these types of arrangements, and at the time of publication, no state or territory educational authority appears to publish the use of these arrangements.

Reviews and inquiries in Australia have recommended the following:

- Better information on access to curriculum and assessment modifications (Australian Senate, 2016)
- Consistent reporting on the use of alternative curriculum—including standard curriculum delivered at a different year level, and pre-foundational or alternative curricula (Deloitte Access Economics, 2017)

Such policies could enable improved information sharing, research, and potentially support regulation of practices. In the United States, for example, access to modified curriculum is regulated through the Individuals with Disabilities Education Act, which specifies that district educational authorities are permitted to measure up to 1% of their students against alternate achievement standards, and 2% against modified achievement standards (Mitchell, 2015).

Review design considerations

The previous section has shown that Australian governments have conducted numerous reviews into inclusive education. These reviews have highlighted the experiences of students with disability and provided a framework for advancing the interests of students with disability.

Yet, despite some notable exceptions, there has been little meaningful progress in either outcomes or policy to support inclusion of children with disability across Australian schools. While the need for, and elements of, reform may be clear from these reviews, significant barriers to progress remain.

The reasons for this are complex and broad-ranging. Indeed, in General Comment 4, a framework of attitudinal, technical, financial and legal barriers to progress is presented. Any one of these areas would be deserving of its own article, and it is not possible to provide a full description as part of this article. Instead, noting that the topic of this article is reviews of education, this article will seek to describe the barriers inherent in the design, implementation and publication of the reviews themselves. This is presented in the hope that the lessons can be instructive in the design of future policy reviews.

While there has been no public acquittal of major Australian reviews, in some cases there is limited evidence that the advice provided to governments has been acted on or even considered. This can be damaging—it leads to wasted effort of governments, independent reviewers and stakeholders. It represents further unproductive use of the time of teachers and school leaders who generously contribute to review processes. Ultimately, failure to act diminishes public confidence in the effectiveness of any review process. Minor improvements to the way in which such reviews form and present recommendations, as well as the systems used to act on them, may improve the way in which Australian jurisdictions consider review findings.

Independent review processes

To the greatest extent possible, reviews should be conducted in a manner completely independent from the government and bureaucracy whose actions are subject to review. They should be conducted in line with a transparent set of criteria or terms of reference—to enable scrutiny of the methodology and findings.

This is to avoid the possibility of findings or recommendations being developed which are based on political acceptability, convenience or fiscal considerations, or indeed to avoid the perception of this.

In Australia, many of the landmark reviews have been conducted by State or Federal Parliaments. While these may be separate from Government, they cannot be said to be truly independent, as they are conducted by committees with representation from ideologically oriented political parties.

Clear ownership of recommendations

To have the greatest effect, recommendations should be concise, evidence-based, and have a clear “owner” responsible for taking those recommendations forward (National Health and Medical Research Council, 2020). Despite the careful planning and complex stakeholder negotiation informing reviews and inquiries, the subsequent recommendations are not always actionable, and in some cases do not clearly assign responsibility and accountability for action. This can lead to wasted effort in the development and consideration of recommendations.

In some instances, the specific body responsible for acting on recommendations is not clear. This is, in part, driven by the federated structure of education within Australia. Under Australian law, states and territories hold responsibility for delivery of education. There are a complex set of bodies, at federal and state levels, who may have responsibility for acting on recommendations arising from reviews which consider education. The history of reviews conducted at a federal level shows the difficulty in subsequently acting on the recommendations outlined.

- A 2016 Senate inquiry recommended the adoption of a national database on access and attainment for students with disability (Australian Senate, 2016). This recommendation was *accepted in principle*, with the response stating that the Australian Government would work with Education Council to develop specific details. However, since that response was issued, this has not been discussed as an item at Education Council (Australian Government, 2017).
- That same inquiry recommended a national strategy to improve education for students with disability (Australian Senate, 2016). The Australian Government also accepted the recommendation in principle, noting that the responsibility lay with State governments, and highlighted an existing patchwork of strategies which worked toward this end. There has been no specific progress in developing a national strategy around education.

In some instances, the responsible body may be clear, but the nature of recommendations put forward by the review can be challenging. Some reviews have resulted in recommendations which do not outline specific actions that can be easily acquitted. For example, in 2017, the *South Australian Select Committee on Access to the Education System for students with disability* delivered its report. This report contained over 80 recommendations, without a clear organizing framework. Among these recommendations were a number of aspirations for improvement, rather than specific recommendations for action (Parliament of South Australia, 2017). In these circumstances, it may be difficult for an authority to determine the specific actions sought, and develop a response.

Guidance provided by the National Health and Medical Research Council indicates that greater specificity in recommendations, including the body responsible or accountable for actioning recommendations, will help those recommendations to be received and actioned appropriately (National Health and Medical Research Council, 2020). This may facilitate progress toward more inclusive education structures, and ensure the efforts of all involved in conducting reviews—reviewers, governments, research participants and sector stakeholders—are put toward a productive end.

Accountability for acting on recommendations

Following a review or inquiry, action can be prompted by mechanisms that support implementation of relevant recommendations. Some government agencies have not established effective mechanisms for these purposes.

Some recent examples of Australian reviews have recommended the establishment of an administrative function and accountability/oversight board to ensure the steady adoption of recommendations (Shaddock et al., 2016; Deloitte Access Economics, 2017). It is important to note that forming such recommendations relies on an understanding of the context of the responding government authority and is based on a shared understanding of the role and remit of the independent reviewer.

Most of the reviews identified in Table 1 above do not have corresponding public acquittals against the implementation of recommendations. In some cases, there is a degree of opacity around the status of acceptance of the recommendations—where governments may not prepare a formal response to a review or inquiry, or do not accept the recommendations of a review. For example, the 2017 South Australian Parliamentary Inquiry, and the 2018 report into Policy and Practice for students in NT schools, do not appear to have corresponding public government responses (Parliament of South Australia, 2017; Center for International Research on Education Systems, 2018).

However, there are examples of reviews where the status of recommendations is scrutinized by an independent authority. A year following their inquiry into school expulsions in Victoria, the Victorian Ombudsman publicly released an acquittal against implementation of those recommendations, classifying recommendations as “implemented,” “in progress” or “not implemented.” The ACT Government publishes a regular communique on the implementation of the recommendations following the “Schools for All” review (Australian Capital Territory Education Directorate, 2020).

Establishing independent mechanisms for scrutinizing the status of review recommendations may increase the likelihood that those recommendations are acted on.

Conclusion

Australian governments have conducted numerous reviews into inclusive education. These reviews have highlighted the experiences of students with disability and provided a framework for advancing the interests of these students.

This article provides a summary of the experiences facing many students with disability across Australia. Across all states and territories, students are placed in segregated educational facilities, excluded from accessing and participating in learning, and not supported through differentiated teaching practice.

These issues associated with the experiences of students with disability have been widely cataloged and are well known at this stage. In addition, the necessary changes have been articulated through hundreds of recommendations at state and federal levels and are reasonably well understood. However, there has been little progress in action at a national or state level—challenges remain with implementation. To support implementation, it is important to establish the right conditions for change from the outset in any review process. Future reviews and inquiries into these matters can overcome some of these challenges through adoption of strict independence in review processes, establishing accountability and responsibility at the right levels, ensuring buy-in across stakeholders, and committing to public release and acquittal of all report recommendations.

References

- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676.
- Attorney-General Department, (n.d.). Disability Discrimination Act 1992 (Cth) (Austl), Canberra.
- Australian Bureau of Statistics, 2018. Survey of Disability, Ageing and Carers.
- Australian Capital Territory Education Directorate, 2020. Schools for All Reform Program. Retrieved September 27, 2020, from <https://www.education.act.gov.au/support-for-our-students/feeling-safe-at-school/schools-for-all-reform-program>.
- Australian Government, 2017. Australian Government Response to the Senate Education and Employment References Committee Report: Access to Real Learning: The Impact of Policy, Funding and Culture on Students With Disability. Parliament of Australia, Canberra. Retrieved from https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_and_Employment/students_with_disability/Government_Response.
- Australian Senate, 2015. Violence, Abuse and Neglect Against People With Disability in Institutional and Residential Settings. Parliament of Australia, Canberra.
- Australian Senate, 2016. Access to Real Learning: The Impact of Policy, Funding and Culture on Students With Disability. Parliament of Australia, Canberra, Australia.
- Castan Center for Human Rights Law, 2018. Improving Educational Outcomes for Children With Disability in Victoria. Castan Center for Human Rights Law, Melbourne, Victoria.
- Center for International Research on Education Systems, 2018. Review of Policy and Practice for Students With Additional Needs. Victoria University, Melbourne, Victoria.
- Children and Young People With Disability Australia, 2016. Education Survey. CYDA, Melbourne, Victoria.
- Children and Young People With Disability Australia, 2019. Education Survey. Children and Young People With Disability Australia, Melbourne, Victoria.
- Committee on the Rights of Persons With Disabilities, 2016. General Comment No. 4, Article 24: Right to Inclusive Education.
- de Bruin, K., Graham, L.J., Gallagher, J., 2020. What is the NCCD and what does it mean for my practice? In: Graham, L.J. (Ed.), *Inclusive Education for the 21st Century*. Allen & Unwin, pp. 122–155.
- de Bruin, K., 2019. The impact of inclusive education reforms on students with disability: an international comparison. *Int. J. Incl. Educ.* 23, 811–826.
- DEEWR, 2010. Review of the Disability Standards for Education. DEEWR, Canberra.
- Deloitte Access Economics, 2017. Review of Education for Students With Disability in Queensland State Schools. Deloitte Access Economics, Brisbane, Queensland.
- Education Council, 2019. Alice Springs (Mparntwe) Education Declaration. Education Services Australia, Melbourne.
- European Agency for Special Needs and Inclusive Education, 2018. Evidence of the Link Between Inclusive Education and Social Inclusion: A Review of the Literature. European Agency for Special Needs and Inclusive Education, Odense, Denmark.
- Glass, D., 2017. Investigation Into Victorian Government School Expulsions. Victorian Ombudsman, Melbourne, Victoria.
- Graham, L.J., McCarthy, T., Killingly, C., Tancredi, H., Poed, S., 2020. Inquiry Into Suspension, Exclusion and Expulsion Processes in South Australian Government Schools. The Center for Inclusive Education, QUT, Brisbane, Queensland.
- Hehir, T., Grindal, T., Freeman, B., Renée, L., 2016. A Summary of the Evidence on Inclusive Education. Abt Associates, Sao Paulo, Brazil.

- Human Rights and Equal Opportunities Commission, Victoria, 2012. *Held Back: The Experiences of Students With Disability in Victorian Schools*. HREOC, Melbourne, Victoria.
- Iacono, T., Keffe, M., Kenny, A., McKinsty, C., 2019. A document review of exclusionary practices in the context of Australian School Education Policy. *J. Pol. Pract. Intellect. Disabil.* 16 (4), 264–272.
- Mann, G., Cuskelly, M., Moni, K., 2015. Choosing a school: parental decision-making when special schools are an option. *Disabil. Soc.* 30 (9), 1413–1427.
- Ministerial Council on Education, Employment, Training and Youth Affairs, 2008. *Melbourne Declaration on Educational Goals for Young Australians*. MCEETYA, Canberra.
- Mitchell, D., 2015. *Education That Fits: Review of International Trends in the Education of Students With Special Educational Needs*, second ed. University of Canterbury, Canterbury.
- National Disability Insurance Scheme, 2019. *NDIS Participant Outcomes 30 June 2018*. National Disability Insurance Scheme Launch Transition Agency, Geelong, Victoria.
- National Health and Medical Research Council, 2020. *Guidelines for Guidelines Handbook: Independent Review*. National Health and Medical Research Council. Retrieved October 1, 2020, from: <https://www.nhmrc.gov.au/guidelinesforguidelines/review>.
- National School Resourcing Board, 2020. *Review of the Loading for Students With Disability: Final Report*. Department of Education, Canberra.
- New South Wales Legislative Council, 2017. *Students With a Disability or Special Needs in NSW Schools*. New South Wales Legislative Council, Sydney, New South Wales.
- Parliament of South Australia, 2017. *Report on the Select Committee on Access to the SA Education System for Students With Disability*. Parliament of South Australia, Adelaide, South Australia.
- Parliament of Victoria, 2017. *Inquiry Into Services for People With Autism Spectrum Disorder: Final Report*. Family and Community Development Committee, Melbourne, Victoria.
- PhillipsKPA, 2015. *Evaluation of the More Support for Students With Disabilities Initiative 2012–2014, Final Evaluation Report*. PhillipsKPA, Melbourne.
- Poed, S., Cologon, K., Jackson, R., 2020. Gatekeeping and restrictive practices by Australian mainstream schools: results of a national survey. *Int. J. Incl. Educ.* 1–14. Retrieved from: <http://allmeansall.org.au/wp-content/uploads/2017/10/TIES-4.0-20172.pdf>.
- Queensland Department of Education, 2018. *Inclusive Education Policy Statement*.
- Royal Commission Into Violence, Abuse, Neglect and Exploitation of People With Disability, 2019. *Education and Learning Issues Paper*. Brisbane. Retrieved August 8, 2021, from: <https://disability.royalcommission.gov.au/system/files/2020-07/Issues-paper-Education-Learning.pdf>.
- Sackville, R., 2020. *Closing Remarks Public Hearing 7: Barriers to Accessing a Safe, Quality and Inclusive School Education and Life Course Impacts*. Royal Commission Into Violence, Abuse, Neglect and Exploitation of People With Disability, Brisbane.
- Shaddock, A., Macdonald, N., Hook, J., Giorelli, L., Arthur-Kelly, M., 2009. *Disability, Diversity and Tides That Lift All Boats: Review of Special Education in the ACT*. Service Initiatives, Chitwick, NSW.
- Shaddock, A., Packer, S., Roy, A., 2016. *Schools for All Children & Young People—Report to ACT Govt From the Expert Panel on Students With Complex Needs and Challenging Behavior*. Australian Capital Territory Government—Human Rights Commission, Canberra.
- Swancutt, L., Medhurst, M., Poed, S., Walker, P., 2019. Making adjustments to curriculum, pedagogy and assessment. In: Graham, L.J. (Ed.), *Inclusive Education for the 21st Century*. Routledge, Brisbane.
- Tah, J.K., 2020. Shopping for schools: parents of students with disabilities in the education marketplace in Stockholm. *Eur. J. Spec. Needs Educ.* 35 (4), 497–512.
- Tasmanian Department of Education and Training, 2015. *Improved Support for Students With Disability: Ministerial Taskforce Report*. Tasmanian Department of Education and Training, Hobart, Tasmania.
- UN General Assembly, *Convention on the Rights of Persons with Disabilities: resolution / adopted by the General Assembly, 24 January 2007. A/RES/61/106*. <https://www.refworld.org/docid/45f973632.html>. (Accessed 18 January 2022).
- UNICEF, 2012. *The Right of Children With Disabilities to Education: A Rights-Based Approach to Inclusive Education*. UNICEF, Geneva.
- Urbis, 2015. *2015 Review of the Disability Standards for Education 2005*. Urbis, Melbourne.
- Victorian Department of Education and Training, 2015. *Review of the Program for Students With Disabilities*. Victorian Department of Education and Training, Melbourne, Victoria.
- Victorian Department of Education and Training, 2018. *Functional Needs Assessment Pilot*. Retrieved October 31, 2020, from: <https://www.education.vic.gov.au/about/programs/Pages/functional-needs.aspx>.
- Western Australia Department of Education and Training, 2004. *Pathways to the Future: A Report of the Review of Educational Services for Students With Disabilities in Government Schools*. Western Australia Department of Education and Training, East Perth.

From Salamanca to Sweden: inclusive education as policy in transit

Gunnlaugur Magnússon, Department of Education, Uppsala University, Uppsala, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	386
Fields of tension	387
Inclusive education as policy	388
The Swedish education system	388
From the international arena to local practices	389
Inclusion as international policy—the ambiguities of the Salamanca statement	389
Inclusion in a Swedish policy-context	390
Inclusion in Swedish teacher-training	391
Inclusion as an organizational question	392
Inclusion in practice of different professions	393
Concluding remarks	394
Acknowledgments	394
References	394

Introduction

When inclusive education was established as an international policy through the Salamanca Statement (UNESCO, 1994), it was seen as a radical shift that would not only improve the situation for pupils traditionally falling under the scope of special education, but be conducive to alter education in general and the future society in particular (Ainscow et al., 2019). Developments in the almost 30 years since the signing of the Salamanca Statement have been complex and several promising international policy initiatives have followed. However, the implications for traditional understanding of the organization of schooling and teaching have varied greatly (de Bruin, 2019) and, in some countries, politicians, parents, and even teacher-unions increasingly question inclusion. This questioning often regards the efficiency of inclusion with respect to increasing educational attainment, as a matter of resources, as a matter of teacher proficiency, and as a matter of stigmatization or *othering*.

The sporadic development of inclusive education has several different reasons. One reason is the conceptual confusion surrounding inclusion, both as regards *whom* it focuses on and *how* it is to be organized (Magnússon, 2019). Also, the concept of inclusion has become positively loaded, in a similar manner as words such as democracy and liberty may be (Nilholm, 2006). As such, inclusion risks suffering the fate of elevator-words (Hacking, 1999), in the end not really referring to objects in the world, but giving a positive shine to that which it is attached to. For some proponents, inclusion therefore seems to have been kidnapped, rendering it meaningless (e.g., Skrtic et al., 1996), or as Slee expressed his concern: "... *that inclusive education is a casualty of a form of 'eduspeak' characterized by reductionism and disconnection and devoid of its original political intent*" (Slee, 2008, p. 104). Indications of this can be seen in how inclusion is used to signify a broad range of organizational practices, some of which are forms of segregation. Its radical political origins thus risk disappearing into oblivion (Slee, 2011).

However, semantics can only explain so much, there are material and organizational factors to account for as well. For instance, the European Agency for Special Needs and Inclusive Education has pointed to a lack of stable and longitudinal financial commitments and investments to further inclusion (Meijer and Watkins, 2019) which is likely to add to the stress of overworked teachers and the exposure of pupils who do not receive the support they need. Inclusion can also be seen as deprioritized in a political landscape that values and promotes economic efficiency, accountability, and standardization, (Apple, 2012; Blossing et al., 2014), and where position on international test measurements are seen as a national priority (Mølsted and Pettersson, 2019). This latter development has been explained as the result of increased power of supranational policy-actors (Daun, 2011), global policy-networks (Ball, 2012), and international policy-borrowing (Steinar-Khamisi, 2004). It is also the case that inclusive education was intended to be promoted by supranational policy-actors, policy-networks, and policy traveling and borrowing. This would indicate that inclusive education needs to be viewed in a broader policy-context where it competes for prioritization (Magnússon et al., 2019).

The varying successes of inclusion as education policy must therefore be analyzed more thoroughly, particularly if one wishes to promote inclusion as an important, democratic and radical alternative to traditional educational organization and education politics. Such analysis is the objective of this article, using Sweden as a case, with the premise that inclusive education is of relevance for several different types of practice (e.g., policy, research and schooling) and that within all of these fields, different actors view inclusion from different perspectives and with different needs and intentions.

Sweden is often raised as a beacon of inclusion in an international perspective (OECD, 2011, 2015). However, a series of market-oriented reforms in the past 30 years have contributed to increased segregation and decreased equity in the education system (Magnússon, 2015, 2020). Even more conspicuous with respect to the erosion of inclusion, is a recently instigated government inquiry (Swedish Government Official Reports 2020:42, 2020) for more segregated solutions within the school system, using

the government sanctioned sentence *“the idea of inclusion has gone too far”* as a point of departure. This theme has in fact been increasingly prevalent in public debate about education in Sweden (Bagger and Lillvist, 2021). Hence, it can be argued that Sweden lives in a tension between its tradition of emphasis on social cohesion and inclusion, and a political hegemony emphasizing market-ideology of competition and efficiency (Magnússon and Pettersson, 2021).

In this chapter I will discuss fields of tension within inclusive education, and give examples of how these appear within research. I will then explain my theoretical approach to the concept. Thereafter, the Swedish education system is briefly described. Then, I will illustrate my argument about inclusion as a traveling international education policy embedded in a national and local policy context, with a summary of research and statistics from different levels of the Swedish education system. Finally, some conclusions will be drawn.

Fields of tension

The thoroughly discussed conceptual complexity of inclusive education can be partly explained by the broad application of the concept and its relevance to different types of practices on different levels. Besides being a policy-concept on international, national, regional and local levels, inclusion is also relevant to a large and diverse field of research and of course to the practical organization of education and teaching. For the aims of this article, inclusive education is viewed as a political ideal of education. As a political ideal, inclusive education embraces distinct normative beliefs about the purpose, the content, and organization of education. However, as I will illustrate below, inclusion as a field of research reflects neither a single nor homogenous ideology, but a range of understandings, not always explicitly stated, that vary greatly in both theory and practice (Avramidis and Norwich, 2002).

Inclusive education also encompasses several fields of tension. One of these fields concerns *who* is the focus of inclusion; which pupil-groups or population cohorts are identified as excluded and in need of being “included” (Florian, 2008; Hansen, 2012; Nilholm, 2006). In its broadest definition, the recognition of particular pupil-groups may be left outside the discussion and a general inclusive ambition raised focusing on democratic communities for “all” pupils. However, such approaches implicitly define pupil groups as not belonging to “all” pupils right now, and thus in need of politics and policies to be included (Popkewitz, 2009). The question of recognizing particular pupil groups pertains to a second field of tension that regards the relationship between inclusive education and special education (Magnússon, 2015). Some researchers argue that an identification and categorization of educational difficulties is in itself an excluding practice which often leads to pathologizing of pupil difference; hence it is antithetic to inclusion (e.g., Haug, 1998; Vislie, 2003; Thomas and Loxley, 2001). The argument is that the explanation of problems within the organization of schooling with individual characteristics of pupils—rather than contextual or organizational factors—will only reproduce the marginalization of vulnerable pupils (Skrtic, 1991; Ainscow, 1998). This approach to inclusion contains a radical questioning of the legitimacy of special education (Magnússon, 2015). However, other researchers argue that the political project of inclusive education should focus on pupils with special educational needs (e.g., Kiuppis, 2013; Miles and Singal, 2010), both due to the historical origins of the political process (Kiuppis, 2013) and to ensure that the particular needs of pupils with disabilities are not rendered invisible (Florian, 2019). Such an approach would be of particular importance in countries where access to education is limited in general and for pupils in need of special educational support in particular (Miles and Singal, 2010).

Over time, the meaning of inclusive education has become more encompassing, increasingly focusing on the notion of creating inclusion for “all children” (Kiuppis, 2013; Florian, 2019). An example of such widening focus of inclusion is found in Hope and Hall (2018) where inclusion is related to the experiences of LGBTQ students in a school specifically (and exclusively) accommodating LGBTQ students. It could therefore be argued that inclusion has developed into a political tool to acknowledge a variety of disadvantaged student groups (Engsig and Johnstone, 2015).

That relates to a third field of tension regarding how inclusive education should be organized and practiced. Several researchers have argued that inclusive education must not be reduced to the placement of the pupils (Ferguson, 2008; Haug, 1998; Slee, 2008; Thomazet, 2009; Vislie, 2003) and that “mainstreaming” (the placement of pupils in need of special support in regular classrooms) risks leading to experiences of exclusion if the placement is not accompanied by organizational resources and pedagogical measures to address the culture of the school and the participation of the pupils (Haug, 1998; Slee, 2011; Vislie, 2003). The question remains: how can research address inclusion both as a matter of mapping different practices and as a matter of improving inclusive education in everyday schooling (Göransson and Nilholm, 2014; Nilholm, 2020)?

It is therefore interesting to note the existence of a wide range of interpretations of the organizational conception of inclusion in research. For instance, Göransson and Nilholm (2014) noted four types of definitions of inclusive education among the most frequently cited research articles, spanning from a focus on the placement of SEN pupils in regular classrooms (mainstreaming), to the more abstract focus on “creating communities” for all pupils.

Later Nilholm and Göransson (2017) noted different definitions of inclusion in different types of research. Large-scale empirical and statistical studies usually used a placement-definition of inclusion, whereas theoretical and small-scale, qualitative articles used broader definitions. This finding was later confirmed by another study (Amor et al., 2018). This conceptual divide between different types of research can be understood through the following suggestion: different types of research need different conceptions of inclusion in order to operationalize it as, for instance, an analytical concept, a theoretical construct, or as a measureable variable. Large-scale register-studies or experimental studies need observable relationships between distinguishable variables, for example: the placement of pupils on the one hand and the knowledge-attainment or student learning outcomes on the other. In small-

scale qualitative studies or theoretical developments of the concept of inclusion, broader and more abstract definitions are more applicable, or can even be the result of the inquiry. This variance can be seen as a challenge for inclusive education as a research field (Göransson and Nilholm, 2014), but it also illustrates inclusive education as a political project lacking a common goal specifically formulated for practice. Thus, further studies of how inclusion is conceptualized as a political project in policy, are needed.

Inclusive education as policy

There are a number of ways to view inclusive education. For instance, it can be seen as:

- a matter of bureaucratic organization—where the prerequisites for different pupils' education are regulated;
- as a matter of teaching and the practical choices of teachers;
- as a matter of philosophy regarding definitions of equity or what is "just" or "right"; and
- as a matter of activism and ambition to improve conditions for either specific pupil-groups or all pupils.

While I see a value in each of these perspectives, here inclusive education is viewed as education policy. Viewing inclusion as policy allows us to investigate its political statements about practice, organization and teaching. Each of these statements are in turn influenced by different philosophies and activism in different spheres and at different levels of policy-making and enactment. It also emphasizes the importance of policy on international, national, regional and local levels, and as negotiated and enacted on all levels.

Education policies illustrate political intentions and ideas about education, its aims, implementation and priorities. Policies are also inherently political as they concern distribution of resources, power and representation within the social field of education (Apple, 2004). Here, I follow Ball's definition of policies as statements about practice, "*intended to bring about idealised solutions to diagnosed problems*" (Ball, 1990, 26), a "*tinkering toward utopia*" of sorts (Tyack and Cuban, 1995) where political ideals about both society and education, are more or less explicitly stated. Hence, besides regulating and defining the framework of educational practice, education policies express assumptions about how pupils are to become citizens of and for society, and how that future society is to be constructed (Popkewitz, 2009).

Economic fluctuations and ideological issues, that change over time influence political ideas about education (Apple, 2004) and as political majorities shift, some policies prevail whereas others perish. However, policies are rarely written by sole actors, especially as regards education where longevity of policies is seen as a stabilizing factor for the education system. Policies thus often contain series of compromises between different political and non-political actors (Ball, 1993). This often leads to contradictory goals or measures that schools and staff must find ways to enact (Ball et al., 2012).

Tracing particular international policies and their consequences in practice, requires an analysis of the tension between general commonalities and local particularities. Any analysis of the traveling of international policies and political agendas is therefore not simply a matter of tracing implementation, but rather a consideration of the contextualization of the policies at different levels and among different actors. An international policy discourse doesn't translate directly into policy texts or practices in its original form, but is rather received and interpreted differently in varying national and regional contexts, constructed of political architectures, infrastructures, ideologies and institutions (Ball, 1998) and subsequently translated and enacted in local educational organization and practice. Wahlström and Sundberg (2018, p. 171) illustrate such traveling through a complex theoretical model, the points below being a modified, with the addition of the level of teacher-training, and simplified version of their model, using inclusive education and Sweden as an example. The list is not intended to be neither a normative hierarchy nor a chronological account. Rather it sketches some of the important levels a transnational policy transfers through.

- Inclusion is formulated as transnational policy at an international level by supra-national institutions (e.g., Salamanca Statement by UNESCO).
- Inclusive education is embedded in a context of political policies, organizational structures and traditions at the national level (e.g., Swedish education policy), where it competes for prioritization among other policies.
- Inclusive education becomes a subject of teacher-training, entering different organizational contexts and traditions, competing with and entered into other subjects (such as special education). Teacher-training is viewed here as an important level of policy transmission in education as it translates curriculum and legislation into the vocabulary and practices of future teachers.
- Inclusive education becomes a matter of policy making at the municipal level, or principal organizer level for privately run schools. In the decentralized Swedish education system, this is the level of resource allocation and organizational prioritizations, that will influence the practical and fiscal prerequisites for inclusion.
- Inclusive education becomes a matter of school-level and classroom-level practices where it is enacted in everyday work by individual teachers, principals, special educators and other staff, parents, and students. Each of these groups may have different interpretations of the concept of inclusion, but still have to enact and cooperate around it.

The Swedish education system

The Swedish education system encompasses preschool (ages 1–5 years) through primary school (6–16 years), to upper-secondary school (16–19 years), and adult education. Education is publicly funded and free of charge at all levels, with the exception of

relatively low and income-related fees for preschool attendance. The concept of “pupil in need of special support” is central with regard to the provision of special education in the Swedish education system (Swedish Code of Statutes, 2010: 800). While the legislation emphasizes attainment of knowledge-goals as a primary indicator of need for special support, “other reasons” such as behavioral and social difficulties are also mentioned. The concept of special educational support thus encompasses a broad array of pupils in Sweden, that is, pupils with disabilities/diagnoses, learning difficulties, and pupils encountering difficulties reaching the knowledge-goals for a variety of other reasons. When suspicion of need for more extensive support arises, principals are legally accountable to initiate investigations and to secure documentation and ensure procedures for interventions are established and followed (Swedish Code of Statutes, 2010: 800 § 3, 5, 7–12). To a certain extent, this means that principals define and interpret what support is to be provided, to whom, and how. It is also of relevance that the Swedish system has several parallel educational tracks, for example state run special schools for pupils who use sign language due to hearing disabilities, and two types of schools for pupils with learning disabilities and severe learning disabilities (a diagnosis of intellectual disability is needed).

The concept of inclusion is not mentioned in the Swedish Education Act, (Göransson et al., 2011; Isaksson and Lindqvist, 2015), but it stipulates that special educational support should primarily be provided in the “regular” classroom if possible and that segregating measures deemed necessary should be temporary and restricted to particular subjects. Recent statistics show that roughly 1% of pupils are educated in segregated groups within their schools and that the proportion grows exponentially over age groups, with approximately 3% of the oldest primary school pupils receiving segregated support (Swedish National Agency for Education, 2018). As Magnússon et al. (2019) argue, one reason may be that inclusion is just one educational policy ideal among others and must be viewed in the general context of education policy and political prioritizations. One such factor is the increasing social segregation in the system which can be related to far-reaching market-reforms that shifted the focus of the education system from equity to efficiency (Magnússon and Pettersson, 2021). An example of the reduced ambitions for inclusion is the above-mentioned government report, investigating the possibilities of more segregating solutions in the school system, primarily “resource-schools” (i.e., temporary schools) and special-classes/groups (Swedish Government Official Reports 2020:42, 2020). This report follows both political prioritizations and increasingly negative public discourse in political debates on education (Bagger and Lillvist, 2021).

The education reforms that introduced marketization in the Swedish education system reflected concurrent international trends. The argument was that competition would lead to higher attainment, a plurality of actors and pedagogical development, and more efficiency (Daun, 2003; Lundahl, 2005, 2010). Thus, the system was decentralized, shifting the ownership and fiscal responsibility of schools from the state to municipalities and school-choice was introduced allowing private actors as principal organizers to own and run independent schools. From a policy perspective, independent schools are equivalent to municipal schools; they are subject to the Education Act and the state-defined curricula, and must fulfill the same goals as municipal schools (Swedish Code of Statutes, 2010: 800) with some exceptions. The municipalities finance pupils’ education from preschool through upper-secondary school, via vouchers, a sum that covers expenses for the pupil’s education at his/her school of choice and no additional tuitions or fees are allowed. Thus, pupils are in principle able to choose between different municipal and independent schools, although the supply of choices varies regionally. Some independent schools are profiled specifically toward pupils with different diagnoses (Swedish National Agency for Education, 2014; Magnússon, 2015) and while independent schools are supposed to be open to all pupils, different profiles and marketing strategies affect who applies to individual schools, and they have a legal-opening to refuse admittance to pupils with special needs.

From the international arena to local practices

In the following sections, I will summarize recent research, with a focus on Sweden to illustrate the complexities of how inclusion travels from an international policy to practice. The structure follows the description in the theoretical section of the article:

- an in-depth analysis of the Salamanca Statement (UNESCO, 1994);
- a contextualization of inclusion as an education policy among other Swedish policies;
- an example of how inclusion is organized in teacher-training;
- discussion of how different municipalities and principal organizers organize work of relevance to inclusion; and
- discussion of how different professions approach inclusion in the Swedish system.

Inclusion as international policy—the ambiguities of the Salamanca statement

The Salamanca Statement is a prime example of an international policy intended to influence national policies and local practices. As regards the development of inclusion, the Statement can be seen as the most important and most-referred to document, despite later developments of important international policies (Graham et al., 2020). However, despite the fact that the Salamanca Statement constitutes a foundational reference in both policy-making and research on inclusive education, its definition of inclusive education is not clear, neither as regards *who* is in focus of inclusion, nor as regards *how* it should be organized. A recently published article (Magnússon, 2019), analyzed the Salamanca Statement, illustrating how inclusion is not only formulated as a solution of the particular educational problem of arbitrary segregation of pupils to special educational provision within contemporary schooling, but also as a solution to problems of future society via the formation of future citizens. Thus, schools are not only supposed to

educate children in particular subjects, but also to raise democratically thinking, tolerant, inclusive, and employable citizens that will participate in the future democratic society.

The Statement is also permeated by several different types of political ideals, of conservative, liberal and left-wing origins (Magnússon, 2019). For instance, education is both formulated as a public good, necessary to construct the future democratic society, and as a private good, where parents are described as privileged partners who should exercise choice. The arguments to motivate inclusion are even embedded in economic terminology, such as efficiency and cost-effectiveness, along with ethical and democratic justification of the suggested political objectives.

As regards *who* is in focus of the statement, its roots in special education are clear, the subtitle being “Framework for Action on Special Needs Education”. The document is therefore largely focused on pupils with different disabilities and learning difficulties. Therefore, using the Salamanca Statement, one could argue that inclusion is about what would be termed pupils in need of special support in Sweden. However, the Statement also states that:

... schools should accommodate all children regardless of their physical, social, emotional, linguistic or other conditions. This should include disabled and gifted children, street and working children, children from remote or nomadic populations, children from linguistic, ethnic or cultural minorities and children from other disadvantaged or marginalized areas or groups.

UNESCO (1994, 6).

In other words, inclusion can also be regarded as encompassing “all children”, albeit with a focus on children seen as disadvantaged. All of the groups mentioned above can be said to fall outside the norm of the exemplary pupil, so the Statement can be said to both define the “normal” or “desired” pupil, along with pupils seen as excluded. Hence, the Statement can be seen as formulating the idea of “*who ‘we’ are, should be and who is not that ‘we’*” yet (Popkewitz, 2009).

As regards the organization of inclusion, the Statement’s Framework for Action, includes guidelines ranging from policy to teaching, services, teacher-training and material resources. While the overarching concern is access (of children with disabilities) to education, the Statement also argues for attendance in the same schools and classrooms. A narrow interpretation of the Statement could therefore view inclusion as primarily a matter of placing pupils with disabilities in regular schools and classrooms. However, the Salamanca Statement provides ample room for interpretation, allowing for segregated educational provision and limitation to placement. The Statement urges governments to enroll all children in regular schools, “*unless there are compelling reasons for doing otherwise*” (UNESCO, 1994, ix) and that every person with disabilities should have the right to express their wishes regarding education “*as far as this can be ascertained*” (UNESCO, 1994, 6). Some additional articles discuss the necessity of separate educational provisions in special schools and establishments for particular individuals, or pupils with particular disabilities (Magnússon, 2019, 686). In other words, the Statement can also be used to argue for segregated educational provision based on categories of difference.

The summary of Magnússon’s analysis states that the notion of inclusion in this international and influential policy-document “*encompasses a range of ideals and generally formulated decrees for educational practice*” ranging from “*specific definitions and focus on the placement of pupils with special needs or disabilities in regular classrooms, to broader ideals of ‘creating communities’ for all pupils*” (Magnússon, 2019, 687). There are arguments pertaining to different political ideologies—it is an “*amalgam of ideals*” that “*allows for a multitude of interpretations of what inclusion can mean*” (Magnússon, 2019). The conclusion is that this may be by design. An international policy that wishes to influence several hundred national and regional education systems needs to be malleable. However, “*(g)iven the lack of clarity and general formulations, The Salamanca Statement risks serving as a foundation from which policy-makers (and researchers) can select in order to legitimize definitions of inclusion that may not be very inclusive for pupils*” (Magnússon, 2019, 686).

Inclusion in a Swedish policy-context

Several researchers have pointed out the necessity of viewing international education policies, such as inclusive education, as something that enters into an existing policy-context, with national and regional traditions, cultures, and prioritizations (Ball, 1998; Slee, 2011). The following summary, attempts such contextualization.

There has been a long tradition of emphasis on equal access, equity and equality in the Swedish education system (Lundahl, 2005; Arnesen and Lundahl, 2006), and Sweden has long been internationally renowned for having an inclusive and equitable education system. However, developments in the past 30 years indicate that this reputation is based upon old merits, and that the education system is in fact decreasingly inclusive and increasingly segregated (OECD, 2015). In fact, education reforms and policies of the past three decades have emphasized individualization, efficiency, fiscal cuts, liberal notions of equality and inclusion, and increased social segregation (Arnesen and Lundahl, 2006; Blossing et al., 2014; Dahlstedt and Fejes, 2018; Dovemark et al., 2018). Several researchers have also pointed to increased social segregation as a result of school-choice and decentralization in Sweden. Schools are becoming more homogenous in terms of pupils’ socio-economic background, ethnicity, and gender, as well as increased disparities in attainment and educational resources (Andersson et al., 2012; Bunar, 2010; Fjellman, 2019; Trumberg, 2011). Recent studies have also showed increased levels of segregation of pupils in need of special support (Magnússon, 2015, 2020, Ramberg, 2015; Tah, 2019). This appears in several manners. First of all, independent schools can legally deny admission to pupils in need of support for economic or organizational reasons. According to Magnússon (2015), approximately 15% of all independent primary

schools had done so, although the numbers varied between different profiles and ranged to over 40% among certain school-profiles. Also, there was apparent clustering of pupils in need of special support at specific types of schools and hence fewer at other types of schools (Magnússon, 2015, 2019). In addition, a parallel-market of independent schools has profiles marketed specifically toward special support has grown forth, leading to further clustering of pupils at such schools (Magnússon et al., 2014; Tah, 2019). Finally, the resources for and provision of special support (such as educated special needs teachers) varies greatly between both primary and secondary schools and has relation to whether or not the schools are publicly run or privately run (Magnússon, 2015, 2016; Ramberg, 2015).

To understand how inclusive education as an international policy affects Swedish policy, we must look at how inclusion is formulated in the legal and curricular framework of education. For instance, the word inclusion is not mentioned in any of the most important governing documents of education in Sweden, and education policies are far from clear as regards more inclusive goals of education (Göransson et al., 2011; Isaksson and Lindqvist, 2015). While conceptions of inclusive ideals of education can be found in general terms in Swedish policy (Lundahl et al., 2013), these are open for interpretation on both municipal and school level in the decentralized Swedish education system.

Isaksson and Lindqvist (2015) conducted an analysis of special education in Swedish education policy and illustrated how the conception of special education fluctuates with political and discursive societal shifts. For instance, they argue that policies emphasizing mainstreaming and integration against segregated solutions followed a general critique of the traditional special education in the 1970s and up until the 1990s. However, they argue that developments in the late 1990s led to a dislocation in education policy where increased individualist focus and new views on school difficulties arose, leading to an increased demand for medical and health-related expert-knowledge. The period between mid-2000s and 2014, was in turn marked by intense and frequent reforms of all levels of the Swedish education system. During these reforms, the idea of inclusion was openly questioned and criticized and a focus on knowledge attainment and international tests arose (Isaksson and Lindqvist, 2015).

An analysis of 23-years of Swedish government-statements (Magnússon et al., 2019) further illustrated the importance of shifts of political discourse. The “knowledge-crisis” that followed a Swedish drop in the PISA-studies was used to justify far-reaching reforms of the education system, from preschool to teacher-training, and the notion of knowledge-crisis was inherited by succeeding governments. While right-wing governments emphasized quality, choice and “better teachers”, left-wing governments tended to emphasize equality and more teachers (or more adults). Inclusion was never mentioned in any of the analyzed statements. When special education was mentioned, it was discussed as a matter of increasing teacher-competence and to accommodate disadvantaged pupils in the system. Magnússon et al. (2019) argued that as policy, inclusion is a moving and developing field and subject to political shifts on national and local levels. Thus, inclusion “competes with other political ambitions and ideals as regards education, and as our analysis illustrates, other ambitions tend to be prioritized on the political agenda” (Magnússon et al., 2019).

Finally, increased segregation in the education system is a matter of concern, whatever definition of inclusion is preferred. That leads to another analysis of inclusive education and the marketization of the Swedish education system (Magnússon, 2020). Here, all articles from Sweden that reported empirical research about either marketization in general or school-choice in particular and special education or inclusive education were analyzed. The marketization of the Swedish education system has been described as a paradigm shift in Swedish education politics (Englund, 1998; Daun, 2003) and as mentioned above, the correlation between the increased segregation and decreased equity in the system on the one hand and the rise of school-choice and market policies on the other hand is well established. As mentioned above, this increased segregation regards both pupils’ socioeconomic and ethnic background, gender, need of support, pupil attainment and the availability of pedagogical resources (staff and learning materials). As Magnússon (2020, 28–29) explains: “This relates to special education, as attainment is central to the definition of the need for special educational support in Sweden.” And that “Similar to international patterns, pupils with lower socio-economic background and with migrant-backgrounds are overrepresented among SEN-pupils in Sweden (...) groups that exercise choice to a lower degree.”

The article concludes that increased segregation of the system affects inclusion in several ways; in particular, the clustering of pupils in need of support at specific schools, their corresponding absence from other schools, and the growing market of independent special schools. Marketization has not led to innovative ways of organizing special support more inclusively, rather traditional organization of special education is reproduced and strengthened in the system and the choices for pupils in need of support are limited. Independent schools also recruit special education teachers and special pedagogues to a much lower degree than municipal schools do (Magnússon et al., 2018; Ramberg, 2015). For an education system often renowned for its inclusivity and equity, these results are dire.

As Magnússon and Pettersson (2021) argue, this illustrates a shift where traditional educational ideals have been given new meanings, including how the public and private are conceptualized. The reproduction and strengthening of traditional special education can therefore also be understood as a consequence of the emphasis of individual choice, attainment and efficiency in an increasingly marketized system as the responsibility for education is privatized to individual or family-preferences. When pupils choose segregated provision of education, it is legitimized as a matter of supply-and-demand rather than as an indicator of the need for development of more inclusive educational environments. Hence the market supplies what Norwich (2000) termed as “elective inclusion/exclusion”. In other words, pupils and parents are given the power to opt out of “inclusive” educational environments.

Inclusion in Swedish teacher-training

The importance of teacher competence to accommodate an increasingly diverse pupil-population is a common theme in education politics, and thus the importance of teacher-training is an important matter, in particular in light of the Swedish “knowledge-crisis”.

In Sweden, teacher-training is a matter of constant debate (Edling and Liljestrand, 2019; Rytzler and Magnússon, 2020) and has been reformed several times during the last 30 years (Hallsén, 2013). As Hallsén (2013) points out, teacher-training reforms are often intended to deliver teachers the system needs to fulfill its expectations, or to contribute to the development of the system rather than to organize training of autonomous professionals prepared to utilize professional knowledge, competence and discretion.

The common denominator for primary and upper-secondary levels and teacher-training is the idea that schools are supposed to shape future citizens (Popkewitz, 2009). Hence, there are concerns that teacher students are not prepared for the task after graduation, and a skepticism toward teacher educators and educational research (Labaree, 2004). This can also be understood as educationalization of social problems, i.e., education is intended to solve both current and future societal problems (Smeyers and Depaepe, 2008; Popkewitz, 2008).

Here inclusion as policy is viewed as an attempt to address present economic and social problems. Thus, aside from preparing teachers to disseminate important knowledge to the next generation, teacher-training must also prepare accountable, efficient and independent teachers (Mølsted and Prøitz, 2019) to deal with and prevent future social problems. It therefore is important to study how teacher-training organizes inclusion. Miškolci et al. (2020) argue that universities have three options for organizing inclusion in teacher-training:

- (1) to incorporate the topic into all or several courses;
- (2) to have a particular course focusing on inclusion (or special education), or
- (3) a combination of the two approaches.

This varies between different universities in Sweden, but inclusion is commonly connected to special education, no matter which approach is used. Thus, inclusive education will often be a subject in a course or lectures on special education, rather than within for instance courses or lectures on teaching or the organization of schooling. This has not been a matter of much research in the Swedish context, but a recent study (Miškolci et al., 2020) made a case study of a Swedish university. Key-actors among both teaching and administrative staff, and students from all teacher-training programs—preschool, primary school and upper-secondary school—were interviewed.

According to the respondents, the university approached the concept of inclusion and education of diverse pupil groups purposefully, both by infusing it through different courses and through a specific course on special education. However, the respondents identified a need for more on inclusion, especially its implications for teaching practices and to understand the relationship between special education and inclusion, and between disability/learning difficulties and diversity in general. In particular, students seemed to view special education and inclusive education as mutually complementary in an uncritical manner. The study concluded that a recent education reform, which called for more special education in teacher-training, had created a hierarchy between inclusive education and special education, prioritizing the latter. Rather than challenging this hierarchy, the studied university had developed and incorporated special education courses in their teacher-training programs. As teacher-training is a space in which the future school is hierarchized and shaped, this acceptance of political influence over course-structure and syllabi is revealing as regards the de-prioritization of inclusive education and the tension among these contradictory educational policies risks constraining the university's capacity to prepare its teacher candidates to develop inclusive practices (Miškolci et al., 2020).

Inclusion as an organizational question

Existing comparative studies of how municipalities and private principal organizers interpret and enact Swedish policies regarding inclusion are unfortunately based upon relatively old data. However, some general indicators can be gleaned from them. First of all, the number of pupils receiving special support in Sweden has historically been high. Giota and Lundborg (2007) estimated that approximately 40% of pupils had received special support at some time during their compulsory school education, and figures indicated approximately 20% of compulsory school pupils are seen as being in need of support at any specific time (Giota and Lundborg, 2007; Nilholm et al., 2007). Until 2014, 14% of pupils in primary schools had intervention plans for special support (Swedish National Agency of Education, 2013), but following a legislative alteration in 2014, that number had dropped to 5% in 2018/2019 (Swedish Government Official Reports 2020:42, 2020) and teachers were intended to adapt more of their teaching without documentation. Segregating measures have been deemed to be on the increase over a long period (Heimdahl Mattson, 2006; Giota and Lundborg, 2007; Nilholm et al., 2007; Emanuelsson and Persson, 2002; Emanuelsson et al., 2002; Magnússon, 2015).

Results from Giota and Emanuelsson's (2011) showed that 25% of principals in independent schools claimed that fiscal resources for special support were insufficient, and Magnússon (2016) argued that the different models of financing special support in the Swedish system was a problem for the equity of the system often raised by independent schools. Comparisons of municipal and independent schools indicate that proportions of pupils in need of special support are higher in municipal schools than in independent schools and that independent schools use segregated groups to a higher degree (Magnússon, 2015, 2020; Ramberg, 2015). There has also been a general tendency to attribute school problems to pupils and pupils in need of special support in all types of schools. Thereby, risk is further marginalized through segregation as an organizational solution (Nilholm et al., 2007; Giota and Emanuelsson, 2011; Magnússon, 2015). Schools with a particular focus on special support are becoming more common and receive increased legitimacy through parental choice (Swedish National Agency for Education, 2014). These schools claim to take care of pupils whom "the regular school" has failed, pupils who have thus "opted out" (Magnússon, 2015, 2016). This is yet another indicator of increased segregation at the system level (Giota and Emanuelsson, 2011)—in this case, based on the need

for special support. An additional risk is that special educational competence may become concentrated at special schools and that other schools may either not be able (or willing) to attain, retain, and develop appropriate competence.

In the light of recent fiscal austerity, there are increasing worries about the economic stability of the education system, particularly as regards consequences for equity. These concerns can be seen among both local politicians and principals who are economically responsible for the schools (Swedish Government Official Reports 2017: 35, 2017; Lärarnas, 2019; Plesner and Larsson, 2019). The consequences of austerity are larger classes, fewer and less qualified teachers, and teaching materials of diminishing quality. This along with recent political moves questioning inclusion as a goal for education (Swedish Government Official Reports 2020:42, 2020) indicates that inclusion has a double function on this level. On the one hand, it is a policy that is to a large degree up to the principal organizers and municipalities to interpret and enact in their organizations. In the decentralized Swedish education system, that means that several different approaches exist simultaneously, not always fulfilling the legislative demands (Swedish School Inspectorate, 2014, 2016). On the other hand, it is used as a term for the placement of pupils in regular classes and thus as justification for downscaling support services. Hence, inclusion at best becomes a secondary objective or, at worst, a justification for fiscal austerity, far departed from its original political intent.

Inclusion in practice of different professions

Schools can be viewed as complex social systems in their own right, constantly being negotiated and reorganized by the people who work and live within them—from principals, teachers, assistants, kitchen staff, leisure-time leaders, janitors, counsellors, to parents, pupils and external staff. The involvement of all occupational groups and the importance of school leadership has long been emphasized in research on inclusion (e.g., Topping and Maloney, 2005), but here I will focus on principals, teachers, and special educators.

The importance of school leadership for the quality of schooling is well established, both regarding learning, organization and development of inclusion (Leithwood et al., 2008; Seashore Louis, 2015; Hoppey and McLeskey, 2013; Persson and Persson, 2012), in particular with regard to sustainability of organizational development. Swedish principals have a legal responsibility for the organization and implementation of special support, health promotion and preventive measures (Swedish Code of Statutes 2010: 800). Thus, the development of an inclusive school environment is to a high degree the principals' responsibility. It is therefore of concern in the Swedish context that 60% of principals have left their position within 3 years, and 80% have done so within 5 years (Swedish National Agency for Education, 2020). Many principals find that they are limited in their leadership by certain policies (Lindqvist and Nilholm, 2013) and that they lack authority in economic issues and personnel decisions (Nihlfors and Johansson, 2013). In a more recent study, Magnússon and Pettersson (2021) reported results from interviews with 22 head teachers of primary schools about inclusion and the market. Most of the principals defined inclusion in terms of goal attainment and "reaching one's full potential", a clear market-influenced formulation with focus on the individual, while others focused more on the societal role of schools, and as a matter of community at the school. The definition of *who* was to be included varied as well, from a broad socially-oriented definition focusing on accommodating newly-arrived refugees and integrating different cultures and religions to a focus on pupils requiring special educational support. An in-common reasoning was that inclusion should not revolve too much around pupil-placement as separating pupils into segregated teaching environments was seen as conducive to their knowledge-attainment; that is, a view of participation in special groups as a manner of organizing inclusion. When it came to the influence of the market on inclusion some principals proclaimed that the idea of inclusion had been ruined by school-choice due to the segregating consequences of the market. Others viewed successful inclusion in their schools, as a tool for competing in the market. In the second case, it can be argued that inclusion is seen more as means to an end, rather than an end in-and-of-itself. Inclusion was in other words a matter of reputation rather than pedagogical objective.

There are several differences in how different professions define and explain special support (Lindqvist et al., 2011; Lindqvist, 2013a). Some prior research indicates that school difficulties are often explained in terms of the pupils' individual attributes rather than organizational factors by principals and teachers, with the exception of the special educators (Lindqvist, 2013a,b; Nilholm et al., 2013; Magnússon and Göransson, 2019). In fact, the differences are so great that Lindqvist et al. (2011) indicate that different professions have different agendas for special support and inclusion. Special educators tend to express values and opinions in line with a more inclusive perspective (Lindqvist et al., 2011; Lindqvist, 2013a; Nilholm et al., 2013; Göransson et al., 2015). This has led to a general view of special educators as being more inclusive and having similar views, but Göransson et al. (2019) showed that despite this there are several differences within the group.

While the special educators are often seen as important for the development of inclusive education, their role is not defined in the Swedish Education Act, and no demands are made for them to be employed by schools. This has also led to common misunderstandings among school staff regarding their role (Magnússon and Göransson, 2019) and, it seems that special educators shape and adapt their occupational roles to the demands of the school they work at (Göransson et al., 2015; Klang et al., 2017). Given that both principals and teachers tend to be less positive toward inclusion and more in favor of traditional special education, special educators are likely to be working in organizational frameworks they do not deem appropriate (Lindqvist, 2013b; Göransson et al., 2015). Here it is also of relevance to remember that Swedish independent schools employ special educators, the professions most inclined to inclusive values, to a much lower extent than municipal schools in Sweden (Magnússon et al., 2018; Ramberg, 2015). This is likely both because special educators have higher salaries and are therefore an expensive resource for schools, and because they come from the pool of experienced teachers, and Sweden has been experiencing an increasing shortage of certified teachers. Of course, this affects the pupil/teacher ratio, which leads to larger classes and work that is more intensive for individual

teachers. Independent schools also tend to have a lower proportion of certified teachers (OECD, 2015; Swedish National Agency for Education, 2019). The decline has been explained with increased bureaucratization and distrust in the professional competence of teachers (OECD, 2015), and has also been related to the marketization of education, which increases dimensions of customer-service and a standardized, goal and results-driven enterprise in the work of teachers (Beach et al., 2014; Nilsson-Lindström and Beach, 2015).

To summarize, teachers and principals tend to be less enthusiastic toward inclusive education than special educators are. There is also increased stress within the system, particularly in the retention of principals; those who are legally accountable for inclusive organization, and teachers, who are to adapt their teaching to increasingly diverse pupil groups in ever-larger classes with limited support. Special educators, who are generally positive toward inclusion, are a scarce resource within the system, and practically non-existent at certain schools. These stressors are detrimental to the development of inclusive education at school level. There are significant differences between schools with respect to inclusive education in the decreasingly equitable Swedish education system.

Concluding remarks

The aim of this article was to illustrate the complexities of inclusive education as a policy that travels from an international arena to the level of practice in local schools, using Sweden as an example. I've summarized some of the ambiguities in the original formulations of inclusive education in the Salamanca Statement (UNESCO, 1994), and argued that these allow for different interpretations of what is considered inclusive in different contexts. More importantly, I've shown how national Swedish education politics have shifted in the past three decades, from equity and inclusion as a national priority toward a more individualized and marketized system, characterized by increasing inequity and social segregation, negatively affecting disadvantaged pupil groups. In the years since the Salamanca Statement (1994), inclusion has not been a prioritized policy in Sweden, it is not even explicitly mentioned in political statements or primary education policies or legislation. The prevalence of segregating methods and organizational solutions is common and increasing among both municipal and private principal organizers. Recent political signals from the government also undercut inclusion as a political ambition. Finally, both principals and teachers are experiencing increased stress as a result of both external demands, fiscal austerity, and the questioning of their professional autonomy—this is visible through high turnover and a deficit of certified teachers. Additionally, the occupational groups most inclined to inclusive values are scarce, often misunderstood, and often working in organizations that do not prioritize inclusion. The prerequisites for inclusive education on all levels in Sweden are therefore deficient and there are no signals of a political will to emphasize inclusive ambitions.

While this image is bleak, the primary contribution of this article is threefold. First, it utilizes an analytical approach that can be adapted and developed further to illustrate the nuance and interplay between international policies and national and local prerequisites for education practice in line with these policies in general and as regards inclusion in particular. Second, it illustrates the importance of contextualization of education policies in general, and inclusive education in particular at national and local levels. Further and more nuanced analyses of how these policies are enacted are needed however, but here the prerequisites are made visible. Finally, by showing both ambiguities of policies, and the contextualization of these throughout complex education systems, we can identify challenges to the radical political potential of inclusive education in policy and practice and attempt to address them. A good starting-point would be the material prerequisites provided to develop inclusion. Without provision of resources, policies may lead to the antithesis of inclusion. Another would be to request that policy-makers and practitioners be more explicit with their ambitions as regards inclusion. Finally, more attention needs to be directed to the development of inclusive practices and organization.

Acknowledgments

The writing of this article was to a great extent the result of a one-year research-position I received at the Department of Education, Uppsala University, which alleviated me from my everyday-duties at the university during most of the otherwise tumultuous year of 2020. I'd like to thank the Department of Education for this opportunity. I'd also like to thank the research-groups that I belong to at Uppsala University (PS and STEP) for lively and challenging discussions, which have contributed greatly to the development of the argument presented here.

References

- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676.
- Ainscow, M., 1998. Would it work in theory? Arguments for practitioner research and theorising in the special needs field. In: Clark, C., Dyson, A., Millward, A. (Eds.), *Theorising Special Education*. Routledge, London, pp. 123–137.
- Amor, A.M., Hagiwara, M., Shogren, K.A., Thompson, J.R., Verdugo, M.Á., Burke, K.M., Aguayo, V., 2018. International perspectives and trends in research on inclusive education: a systematic review. *Int. J. Incl. Educ.* 23 (12), 1277–1295.
- Andersson, E., Malmberg, B., Östh, J., 2012. Travel to school distances in Sweden 2000–2006: changing school geography with equality implications. *J. Transport Geogr.* 23, 34–45.
- Apple, M.W., 2004. *Ideology and Curriculum*. Routledge, New York.
- Apple, M.W., 2012. *Education and Power*. Routledge, New York.
- Arnesen, A.L., Lundahl, L., 2006. Still social and democratic? Inclusive education policies in the Nordic welfare states. *Scand. J. Educ. Res.* 50 (3), 285–300.

- Avramidis, E., Norwich, B., 2002. Teachers' attitudes towards integration/inclusion: a review of the literature. *Eur. J. Spec. Needs Educ.* 17 (2), 129–147.
- Bagger, A., Liljvist, A., 2021. Mediala diskurser om inkludering. En berättelse om (gem)ensamhet. *Utbildning Demokrati* 30 (1), 45–74.
- Ball, S.J., Maguire, M., Braun, A., 2012. *How Schools Do Policy. Policy Enactment in Secondary Schools*. Routledge, London.
- Ball, S.J., 1990. *Discipline and Chaos. The New Right and Discourses of Derision. Politics and Policy Making in Education*, London: Routledge. Reprinted in Ball, S.J. (2006). In: *Education Policy and Social Class. The selected works of Stephen J. Ball*. Routledge, New York, pp. 26–42.
- Ball, S.J., 1993. What is policy? Texts, trajectories and toolboxes. *Discourse* 13 (2), 10–17.
- Ball, S.J., 1998. Big policies/small world: an introduction to international perspectives in education policy. *Comp. Educ.* 34 (2), 119–130.
- Ball, S.J., 2012. *Global Education Inc. New Policy Networks and the Neo-liberal Imaginary*. Routledge, London.
- Beach, D., Bagley, C., Eriksson, A., Player-Koro, C., 2014. Changing teacher education in Sweden: using meta-ethnographic analysis to understand and describe policy making and educational changes. *Teach. Educ.* 44 (6), 160–167.
- Blossing, U., Imsen, G., Moos, L., 2014. The Nordic Education Model. "A School for All" Encounters Neo-liberal Policy. Springer, Dordrecht.
- Bunar, N., 2010. The controlled school market and urban schools in Sweden. *J. Sch. Choice* 4 (1), 47–73.
- Dahlstedt, M., Fejes, A., 2018. *Skolan, marknaden och framtiden*. Studentlitteratur, Lund.
- Daun, H., 2003. Market forces and decentralization in Sweden: impetus for school development or threat to comprehensiveness and equity? In: Plank, D.N., Sykes, G. (Eds.), *Choosing Choice: School Choice in International Perspective*. Teachers College Press, New York, pp. 92–111.
- Daun, H., 2011. Globalization, EU-ification, and the new mode of educational governance in Europe. *Eur. Educ.* 43 (1), 9–32.
- de Bruin, K., 2019. The impact of inclusive education reforms on students with disability: an international comparison. *Int. J. Incl. Educ.* 23 (7–8), 811–826.
- Dovemark, M., Kosunen, S., Kauko, J., Magnúsdóttir, B., Hansen, P., Rasmussen, P., 2018. Deregulation, privatization and marketization of Nordic comprehensive education: social changes reflected in schooling. *Educ. Inq.* 9 (1), 122–141.
- Edling, S., Liljestränd, J., 2019. Let's talk about teacher education! Analyzing the media debates in 2016–2017 on teacher education using Sweden as a case. *Asia Pac. J. Teach. Educ.* 48 (3), 251–266.
- Emanuelsson, I., Persson, B., 2002. Differentiering, specialpedagogik och likvärdighet. En longitudinell studie av skolkarriärer bland elever i svårigheter. *Pedagogisk Forskning i Sverige* 7 (3), 183–199.
- Emanuelsson, I., Persson, B., Rosenqvist, J., 2002. *Forskning inom det specialpedagogiska området—en kunskapsöversikt*. Skolverket, Stockholm.
- Englund, T., 1998. Utbildningspolitiska vägval – förändrade förutsättningar för skola och didaktik. In: Englund, T. (Ed.), *Utbildningspolitiskt systemskifte? HLS förlag, Stockholm, Sweden*, pp. 26–43.
- Engsig, T.T., Johnstone, C.J., 2015. Is there something rotten in the state of Denmark? The paradoxical policies of inclusive education—lessons from Denmark. *Int. J. Incl. Educ.* 19 (5), 469–486.
- Ferguson, D.L., 2008. International trends in inclusive education: the continuing challenge to teach each and everyone. *Eur. J. Spec. Needs Educ.* 23 (2), 109–120.
- Fjellman, A.M., 2019. *School Choice, Space and the Geography of Marketization- Analyses of Educational Restructuring in Upper Secondary Education in Sweden*. Göteborgs universitet, Göteborg.
- Florian, L., 2008. Special or inclusive education: future trends. *Br. J. Spec. Educ.* 35 (4), 202–208.
- Florian, L., 2019. On the necessary co-existence of special and inclusive education. *Int. J. Incl. Educ.* 23 (7–8), 691–704.
- Giota, J., Emanuelsson, I., 2011. *Specialpedagogiskt stöd, till vem och hur?* Faculty of Education and Special Education Report Series, no. 1 RIPS-Rapport 20011:1 Gothenburg University, Faculty of Education and Special Education, Gothenburg.
- Giota, J., Lundborg, O., 2007. *Specialpedagogiskt stöd i grundskolan omfattning, former och konsekvenser*. Faculty of Education Internal Report Series, no. IPD-Rapport 2007:3. Gothenburg University, Faculty of Education and Special Education, Gothenburg.
- Göransson, K., Nilholm, C., 2014. Conceptual diversities and empirical shortcomings—a critical analysis of research on inclusive education. *Eur. J. Spec. Needs Educ.* 29 (3), 265–280.
- Göransson, K., Nilholm, C., Karlsson, K., 2011. Inclusive education in Sweden? A critical analysis. *Int. J. Incl. Educ.* 15 (5), 541–555.
- Göransson, K., Lindqvist, G., Klang, N., Magnússon, G., Nilholm, C., 2015. *Speciella yrken? Specialpedagogers och speciallärares arbete och utbildning: En enkätstudie*. Karlstad University Studies Report 2015:13. Karlstad University, Karlstad.
- Göransson, K., Lindqvist, G., Klang, N., Magnússon, G., 2019. Professionalism, governance and inclusive education—a total population study of Swedish special needs educators. *Int. J. Incl. Educ.* 23 (6), 559–574.
- Graham, L., Medhurst, M., Malaquias, C., Tancredi, H., de Bruin, C., Gillett-Swan, J., Poed, S., Spandagou, I., Carrington, S., Cologon, K., 2020. Beyond Salamanca: a citation analysis of the CRPD/GC4 relative to the Salamanca Statement in inclusive and special education research. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2020.1831627>. Ahead of print.
- Hacking, I., 1999. *The Social Construction of What?* Harvard University Press, Cambridge.
- Hallsén, S., 2013. *Lärarutbildning i skolans tjänst? En policyanalys av statliga argument för förändring*. Acta Universitatis Upsalensis, Uppsala.
- Hansen, J.H., 2012. Limits to inclusion. *Int. J. Incl. Educ.* 16 (1), 89–98.
- Haug, P., 1998. *Pedagogiskt dilemma: Specialundervisning*. The National Agency for Education, Stockholm.
- Heimdahl Mattson, E., 2006. *Mot en inkluderande skola? Skolledares syn på organisation och genomförande av specialpedagogiska insatser*. Specialpedagogiska Institutet och Lärarhögskolan i Stockholm, Stockholm.
- Hope, M.A., Hall, J.J., 2018. "This feels like a whole new thing": a case study of a new LGBTQ-affirming school and its role in developing "inclusions". *Int. J. Incl. Educ.* 22 (12), 1320–1332.
- Hoppey, D., McLeskey, J., 2013. A case study of principal leadership in an effective inclusive school. *J. Spec. Educ.* 46 (4), 245–256.
- Isaksson, J., Lindqvist, R., 2015. What is the meaning of special education? Problem representations in Swedish policy documents: late 1970–2014. *Eur. J. Spec. Needs Educ.* 30 (1), 122–137.
- Kiuppis, F., 2013. Why (not) associate the principle of inclusion with disability? Tracing connections from the start of the Salamanca process. *Int. J. Incl. Educ.* 18 (7), 746–761.
- Klang, N., Gustafson, K., Möllås, G., Nilholm, C., Göransson, K., 2017. Enacting the role of special needs educator—six Swedish case studies. *Eur. J. Spec. Needs Educ.* 32 (3), 391–405.
- Labaree, D., 2004. *The Trouble with Ed Schools*. Yale University Press, New Haven.
- Lärarnas, R., 2019. *Hur ska grundskolan finansieras i framtiden? Så tycker rektorer och kommunala skolpolitiker om skolans resurser de närmaste tio åren*. Lärarnas Riksförbund och Sveriges Skolledarförbund.
- Leithwood, K.A., Harris, A., Hopkins, D., 2008. Seven strong claims about successful school leadership. *Sch. Leader. Manag.* 28 (1), 27–42.
- Lindqvist, G., Nilholm, C., 2013. Making schools inclusive? Educational leaders' views on how to work with children in need of special support. *Int. J. Incl. Educ.* 17 (1), 95–110.
- Lindqvist, G., Nilholm, C., Almqvist, L., Wetso, G.M., 2011. Different agendas? The views of different occupational groups on special needs education. *Eur. J. Spec. Needs Educ.* 26 (2), 143–157.
- Lindqvist, G., 2013a. Who Should Do What to Whom? Occupational Groups' Views on Special Needs. School of Education and Communication Jönköping University, Jönköping.
- Lindqvist, G., 2013b. SENCos: vanguards or in vain? *J. Res. Spec. Educ. Needs* 13 (3), 198–207.
- Lundahl, L., Erikson Arreman, I., Holm, A.S., Lundström, U., 2013. Educational marketization the Swedish way. *Educ. Inq.* 4 (3), 497–517.
- Lundahl, L., 2005. A matter of self-governance and control. The reconstruction of Swedish education policy: 1980–2003. *Eur. Educ.* 37 (1), 10–25.
- Lundahl, L., 2010. Sweden: decentralization, deregulation, quasi-markets—and then what? *J. Educ. Pol.* 17 (6), 687–697.

- Magnússon, G., Göransson, K., 2019. Perimeters of, and challenges to, the jurisdiction of Swedish special educators: an exploration of free text responses. *Eur. J. Spec. Needs Educ.* 34 (3), 257–271.
- Magnússon, G., Pettersson, D., 2021. Imaginaries of Inclusion in Swedish Education. *Oxford Encyclopedia of School Reform*. Oxford University Press, New York. <https://doi.org/10.1093/acrefore/9780190264093.013.1682>.
- Magnússon, G., Göransson, K., Nilholm, C., 2014. Similar situations? Special needs in different groups of independent schools. *Scand. J. Educ. Res.* 59 (4), 377–394.
- Magnússon, G., Göransson, K., Nilholm, C., 2018. Varying access to professional, special educational support : a total population comparison of special educators in Swedish independent and municipal schools. *J. Res. Spec. Educ. Needs* 18 (4), 225–238. <https://doi.org/10.1111/1471-3802.12407>.
- Magnússon, G., Göransson, K., Lindqvist, G., 2019. Contextualizing inclusive education in education policy: the case of Sweden. *Nordic J. Stud. Educ. Policies* 5 (2), 67–77.
- Magnússon, G., 2015. Traditions and Challenges. Special Support in Swedish Independent Compulsory Schools. Mälardalen University, Västerås.
- Magnússon, G., 2016. Representations of special support: independent schools' descriptions of special educational provision. *Eur. J. Spec. Needs Educ.* 31 (2), 155–170.
- Magnússon, G., 2019. An amalgam of ideals—images of inclusion in the Salamanca statement. *Int. J. Incl. Educ.* 23 (7–8), 677–690.
- Magnússon, G., 2020. Inclusive education and school choice lessons from Sweden. *Eur. J. Spec. Needs Educ.* 35 (1), 25–39.
- Meijer, C.J.W., Watkins, A., 2019. Financing special needs and inclusive education—from Salamanca to the present. *Int. J. Incl. Educ.* 23 (7–8), 705–721.
- Miles, S., Singal, N., 2010. The education for all and inclusive education debate: conflict, contradiction or opportunity? *Int. J. Incl. Educ.* 14 (1), 1–15.
- Miškolci, J., Magnússon, G., Nilholm, C., 2020. Complexities of preparing teachers for inclusive education: case-study of a university in Sweden. *Eur. J. Spec. Needs Educ.* <https://doi.org/10.1080/08856257.2020.1776983>. Ahead of print.
- Mølsted, C.E., Pettersson, D., 2019. New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research. Routledge, London, New York.
- Mølsted, C.E., Pritz, T., 2019. Teacher-chameleons: the glue in the alignment of teacher practices and learning in policy. *J. Curric. Stud.* 51 (3), 403–419.
- Nihlfors, E., Johansson, O., 2013. Rektor en stark länk i styrningen av skolan. SNS Förlag, Stockholm.
- Nilholm, C., Göransson, C., 2017. What is meant by inclusion? An analysis of high impact research in North America and Europe. *Eur. J. Spec. Needs Educ.* 32 (3), 437–451.
- Nilholm, C., Persson, B., Hjerm, M., Runesson, S., 2007. Kommuners arbete med elever i behov av särskilt stöd—en enkätundersökning. Insikt, 2. Rapporter från Högskolan i Jönköping. Jönköping University Press, Jönköping.
- Nilholm, C., Almqvist, L., Göransson, K., Lindqvist, G., 2013. Is it possible to get away from disability-based classifications in education? An empirical investigation of the Swedish system. *Scand. J. Disabil. Res.* 15 (4), 379–391.
- Nilholm, C., 2006. Special education, inclusion and democracy. *Eur. J. Spec. Needs Educ.* 21 (4), 431–445.
- Nilholm, C., 2020. Research about inclusive education in 2020—how can we improve our theories in order to change practice? *Eur. J. Spec. Needs Educ.* <https://doi.org/10.1080/08856257.2020.1754547>. Ahead of print.
- Nilsson-Lindström, M., Beach, D., 2015. Changes in teacher education in Sweden in the neo-liberal education age: toward an occupation in itself or a profession for itself. *Educ. Inq.* 6 (3), 241–258.
- Norwich, B., 2000. Inclusion in education: from concepts, values and critique to practice. In: Daniels, H. (Ed.), *Special Education Re-formed. Beyond Rhetoric?* Falmer, London.
- OECD, 2011. Social Justice in the OECD: How Do the Member States Compare? Sustainable Governance Indicators 2011. Bertelsmann Stiftung, Gütersloh.
- OECD, 2015. Improving Schools in Sweden: An OECD Perspective. OECD, Paris.
- Persson, B., Persson, E., 2012. Inkludering Och Måluppfyllelse—Att Nå Framgång Med Alla Elever. Liber, Stockholm.
- Plesner, Å., Larsson, M., 2019. De effektiva. Tankesmedjan Balans, Skarpnäck.
- Popkewitz, T., 2008. The social, psychological, and education sciences: from educationalization to pedagogicalization of the family and the child. In: Smeyers, P., Depaepe, M. (Eds.), *Educational Research: The Educationalization of Social Problems*. Educational Research, vol. 3, pp. 171–190.
- Popkewitz, T., 2009. Curriculum study, curriculum history, and curriculum theory: the reason of reason. *J. Curric. Stud.* 41 (3), 301–319.
- Ramberg, J., 2015. Special Education in Swedish Upper Secondary Schools. Resources, Ability Grouping and Organization. Stockholm University, Stockholm.
- Rytzel, J., Magnússon, G., 2020. Didaktik som läranutbildningens innehåll och form: motstånd, dilemman och möjligheter. *Utbildning Demokrati* 29 (1), 29–64.
- Seashore Louis, K., 2015. Linking leadership to learning: state, district and local effects. *Nordic J. Stud. Educ. Policy* 2015 (3).
- Skrtic, T.M., Sailor, W., Gee, K., 1996. Voice, collaboration, and inclusion. Democratic themes in educational and social reform initiatives. *Rem. Spec. Educ.* 17 (3), 142–157.
- Skrtic, T.M., 1991. Behind Special Education. A Critical Analysis of Professional Culture and School Organization. Love Publishing Company, Denver.
- Slee, R., 2008. Beyond special and regular schooling? An inclusive education reform agenda. *Int. Stud. Sociol. Educ.* 18 (2), 99–116.
- Slee, R., 2011. The Irregular School. Exclusion, Schooling and Inclusive Education. Routledge, London.
- Smeyers, P., Depaepe, M., 2008. Introduction-pushing social responsibilities: the educationalization of social problems. In: Smeyers, P., Depaepe, M. (Eds.), *Educational Research: The Educationalization of Social Problems*. Educational Research, vol. 3, pp. 1–11.
- Steinar-Khamisi, G., 2004. The Global Politics of Educational Borrowing and Lending. Teachers College Press, NY, pp. 54–67.
- Swedish Code of Statutes 2010: 800. Skollagen. Stockholm: Swedish Code of Statutes.
- Swedish Government Official Reports 2017:35, 2017. Slutbetänkande av 2015 års skolkommision. Samling för skolan. Nationell strategi för kunskap och likvärdighet. Norstedts, Stockholm.
- Swedish Government Official Reports 2020:42, 2020. En annan möjlighet till särskilt stöd. Reglering av kommunala resursskolor. Delbetänkande av Utredningen om elevers möjligheter att nå kunskapskraven. Norstedts, Stockholm.
- Swedish National Agency for Education, 2013. PM. Särskilt stöd i Grundskolan. Dnr. 71-2013:28.
- Swedish National Agency for Education, 2014. Fristående skolor för elever i behov av särskilt stöd—en kartläggning. The National Agency for Education, Stockholm.
- Swedish National Agency for Education, 2018. PM. Särskilt stöd i grundskolan läsåret 2017/18. PM, Dnr: 2018:600. Skolverket, Stockholm.
- Swedish National Agency for Education, 2019. Pedagogisk personal i skola och vuxenutbildning läsåret 2018/19. Dnr: 5.1.1-2019:46.
- Swedish National Agency for Education, 2020. PM. Omsättning bland rektorer i grund- och gymnasieskolan. Dnr: 5.1.1-2019:1485.
- Swedish School Inspectorate, 2014. Stöd och stimulans i klassrummet—Rätten att utvecklas så långt som möjligt. Rapport 2014:2. Skolinspektionen, Stockholm.
- Swedish School Inspectorate, 2016. Skolans arbete med extra anpassningar—Kvalitetsgranskningsrapport 2015:2217. Skolinspektionen, Stockholm.
- Tah, J.K., 2019. One market, two perspectives, three implications: on the development of the special education market in Sweden. *Eur. J. Spec. Needs Educ.* 34 (4), 485–498.
- Thomas, G., Loxley, A., 2001. Deconstructing Special Education and Constructing Inclusion. Open University Press, Baltimore.
- Thomazet, S., 2009. From integration to inclusive education: does changing the terms improve practice? *Int. J. Incl. Educ.* 13 (6), 553–563.
- Topping, K., Maloney, S., 2005. The Routledge Falmer Reader in Inclusive Education. Routledge, London.
- Trumberg, A., 2011. Den delade skolan. Segregations processer i det svenska skolsystemet. Örebro Universitet, Örebro.
- Tyack, L., Cuban, L., 1995. Tinkering Toward Utopia. A Century of Public School Reform. Harvard University Press, Cambridge.
- UNESCO, 1994. The Salamanca Statement and Framework for Action on Special Needs Education. UNESCO, Paris.
- Vislie, L., 2003. From integration to inclusion: focusing on global trends and changes in the western European societies. *Eur. J. Spec. Needs Educ.* 18 (1), 17–35.
- Wahlström, N., Sundberg, D., 2018. Discursive institutionalism: towards a framework for analyzing the relation between policy and curriculum. *J. Educ. Pol.* 33 (1), 163–183.

Public policies in Portuguese education: the path to inclusion for all

Ines Alves^a and Domingos Fernandes^b, ^a University of Glasgow, Glasgow, United Kingdom; and ^b ISCTE—Instituto Universitário de Lisboa, Lisbon, Portugal

© 2023 Elsevier Ltd. All rights reserved.

Introduction	397
What do we mean by inclusion in education?	397
What do we mean by assessment?	398
Development of inclusive education policies	399
Development of assessment	400
Assessment for inclusion or exclusion?	402
References	403

Introduction

This chapter focuses on the innovations in legislation and policy in Portugal in the context of developing an inclusive education system. In his analysis of cultural reproduction and distribution of power, Bernstein claims that "... curriculum, pedagogic practice, and modes of evaluation set the terms for the crucial encounters in the classroom context of teachers and pupils" (Bernstein, 2003, p. 154). However, Lingard and Mills consider that "... of the three message systems of schooling, sometimes the main driver is assessment or its narrower companion, testing" (Lingard and Mills, 2007, p. 237). So, in this chapter, we will review the development of policies of inclusion and assessment in Portugal. Building on Bernstein's idea of the three message systems of schooling: curriculum, pedagogy, assessment (Lingard and Mills, 2007), we argue that any country's attempt to develop an inclusive education system must consider and align both curriculum and pedagogical issues. Pedagogical issues (understood as comprising issues related to teaching, learning, and assessment) should be focal in processes and discussions about developing inclusion in education for all learners. However, changes to pedagogical approaches in education are slower and more complex than changes to the curriculum, as they require a consistent policy approach that involves long-term teacher education and empowerment.

The *Education for All* and the inclusive education movements have different origins (Miles and Singal, 2010; Kiuppis, 2014) with different groups of learners being its initial focus or "raison d'être." In Portugal in the 1950s and 1960s there were extremely low levels of literacy and access to education (10% population). Only a minority of the population attended education, and this was based on a very selective system in which exams played a key role in the (non-) transition between cycles and in grade repetition (Fernandes, 2014). The mid-70s brought an awareness of the selective and discriminatory character of education, and a move away from national exams. It also brought an inflow of migrants from the Portuguese former colonies (Marques et al., 2007), which caused a diversification in the school population from the 1980s. The publication, in 1986, of the Fundamental Law of the Education System (Law 46/86 of 14 October), approved after a wide public debate and a meaningful political consensus, is a crucial landmark of the development of the Portuguese public policies in education. The structure of the Portuguese educational system set out by this law is organized into Pre-School, Basic and Secondary Education. Basic education comprises three cycles: First Cycle (Grades 1–4); Second Cycle (Grades 5–6), and Third Cycle (Grades 7–9). Secondary Education includes Grades 10–12.

The parallel development of education for all and inclusive education described by Kiuppis (2014) can be observed in the Portuguese context. As described above, through an opening of education to children and young people from lower social economic backgrounds (aiming to achieve education for all), and a parallel development of legislation from the late 1970s that expected the increasing presence of some learners with disabilities in mainstream schools (special/inclusive education movement). We argue here, and in line with recent international policies (e.g., United Nations, 2016; International Bureau of Education-UNESCO, 2016; UNESCO, 2017), that inclusion in education means much more than mainstreaming learners with disabilities—that it means changing education systems to remove barriers to learning that may be experienced by all learners.

What do we mean by inclusion in education?

According to the General Comment 4 on Article 24 (United Nations, 2016), inclusive education is to be understood as a fundamental human right for all learners which requires a whole system and whole person approach. This approach requires changes to institutional cultures, policies, practices—including flexible curricula and pedagogies. At the basis of the development of inclusive education is support being available not only to learners, but also to teachers.

Inclusion involves a process of systemic reform embodying changes and modifications in content, teaching methods, approaches, structures and strategies in education to overcome barriers with a vision serving to provide all learners of the relevant age range with an equitable and participatory learning experience and environment that best corresponds to their requirements and preferences.

United Nations (2016, p. 4).

We consider that an inclusive education system is one in which all learners have access to, participate, and succeed in education. As these three concepts can be used in different ways, we will start by defining what we mean by access, participation, and success as shown in Fig. 1.

Access means both physical access to school buildings and to learning and teaching—accessible to all learners, regardless of their abilities or backgrounds. For example, braille or enlarged print might make learning accessible to learners with vision impairments, or interpreters may facilitate learning for learners from diverse linguistic backgrounds.

Participation means taking part in the activities, but also feeling part of a group, a class or a school—a sense of belonging as proposed by Slee (2019). Since the sense of participation is individual and subjective, different learners might feel included or excluded by the same activity or situation. Participation could be seen as a “core element and, at the same time, a lever for the improvement of inclusion” (Susinos, 2020, p. 201).

Success cannot be reduced to the enrollment of learners with disabilities in a mainstream school. Inclusive education means accepting the principle that all learners can learn and be taught, therefore, teachers should hold high expectations for all learners (Hart et al., 2004). Success in education tends to be determined by assessments, particularly by summative assessments and grades. On the other hand, systematic formative assessment, sometimes referred to as assessment for learning, has been related to significant improvement in student learning, and learners who are considered to “have more difficulties” are the ones who benefit the most from consistent formative assessment (Black and Wiliam, 1998). Thus, assessment can be looked at as a powerful pedagogical process in the path to developing inclusive education. When referring to assessment and success, the latter should consider the best interest of each learner as proposed in the UN Convention on the Rights of the Child (United Nations, 1989). Understanding what success means for different learners requires considering the aims of education in general, and should start from an understanding that learners’ learning and life opportunities improvement need to take into account all forms of personal, social, and academic learning (European Agency for Special Needs and Inclusive Education, 2017). Success needs to be considered in the context of providing equitable, quality education to all learners (United Nations, 2015).

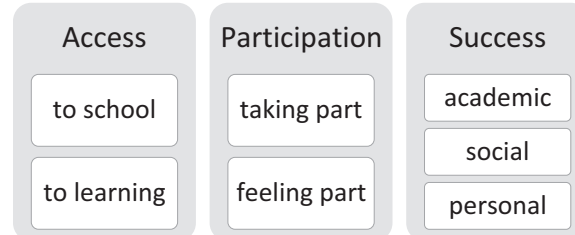


Fig. 1 Access, participation, and success as the main pillars of inclusive education.

What do we mean by assessment?

Learning assessment is generally conceived as a process that enables teachers to gather relevant information about what learners know and can do, to make a variety of possible decisions (e.g., Black and Wiliam, 1998; Assessment Reform Group, 2002). This can be achieved through external and internal assessments. The former ones are generally administered and run by some sort of entity external to the schools (e.g., Ministry of Education, Assessment Institute) whereas teachers and schools are fully responsible for internal assessments.

For the purposes of this chapter, pedagogical assessment is seen as a process through which both teachers and learners might collect, analyze, interpret, discuss, and make use of information related to learning taking into account a variety of purposes such as: (a) to provide quality feedback on learner’s work to both learners and parents; (b) to track the learning progress toward what has been defined through a given set of criteria, indicators and standards; (c) to identify those aspects of learning that need improvement as well as those that show a good domain of the material to be learned; and (d) to determine learners’ grades (Fernandes, 2019). Some of these purposes are formative by nature, that is, they are aimed at supporting learners’ efforts to learn through the distribution of quality feedback. Therefore, feedback is intrinsically associated with formative assessment and in a very systematic way it must make learners aware of what they are supposed to learn, where they are in the process of learning, and what they should do to achieve the intended learning outcomes. In a sense, formative assessment is day-by-day assessment integrated in teaching and learning and, therefore, it is pedagogical by nature because it has to do with the use of a variety of teaching and assessment methods that help learners to succeed. In this context, teachers’ questioning of learners becomes critical. And the, information that results from formative assessments should never be used to grade learners. Other purposes are of a summative nature, that is, at

a certain point in time they make a synthesis of what learners know and can do. This is, for example, the case of grading a particular piece of learners' work (e.g., essay, test, oral report, portfolio of artifacts) or grading learners' work at the end of a given period like a semester. However, summative assessments can also be used for the single purpose of providing feedback to learners; and this means that the results of these assessments are not used to grade learners.

Pedagogical assessments are internal by nature and their fundamental purpose is to improve learners' progress. Even grades should be thought as a means to help learners to progress. Particularly if some form of feedback is attached to them.

In sum, pedagogical assessment encompasses both formative and summative assessments. Research has been showing that formative assessment, also labeled as assessment for learning, increases dramatically what learners learn (Black and Wiliam, 1998) due to the effect of quality feedback. Summative assessment, also known as assessment of learning, can be used to provide feedback, to grade learners' work or to do both. Formative assessment is continuous because it is integral to teaching and learning and, as a result, it regulates these pedagogical processes through feedback. Summative assessment is periodical by nature since it is used to provide an overview of learners' learning at certain points in time. One of the goals of assessment might be the achievement of learners' continuous self-assessment of their own learning and productivity.

Development of inclusive education policies

This section provides a brief overview of the development of legislation and policies focusing on special/inclusive education in Portugal. The national legislation changes have been taking place roughly every decade since the 1970s (1979, 1986, 1991, 2008, 2018) pursuant to developing an inclusive school where all learners are welcome and supported in their learning. Further details on the progress toward inclusion can be found in recent publications (cf. Alves, 2019; Alves et al., 2020; Alves, 2020).

The Law on Special Education of 1979 (Law 66/79) created the expectation of some learners "with disabilities" to attend mainstream schools and it was anticipated that schools would progressively adapt and restructure to respond to the needs of learners with sensory, physical and intellectual disabilities. The expectation was not that all learners would attend mainstream schools, and so segregated settings were created as a result of parent and social society movements. These consisted mainly of education and rehabilitation cooperatives (CERCI—Cooperativas de Educação e Reabilitação de Crianças Inadaptadas), which still exist at present. Its role has developed from responding only to the needs of children with learning disabilities in schooling age, to its present form which includes for example Early Intervention Units, Vocational Education and Training Centers, Protected Employment Centers, and home support for disabled people in general.

The 1986 Fundamental Law of Education (Law 48/1986) is to this day, the structuring legislation which presented the blueprint of the Portuguese national educational system, recognizing the right to education for all; aiming to offer equal opportunities. This legislation extended compulsory schooling to nine years or until learners reached 15 years old. Compulsory schooling was increased to 12 years in 2009 (Law 85/2009). Law 48/86 proposed that learners would attend either special or mainstream schools depending on the level of "specific education need." This legislation considered special and mainstream schools to be part of "mainstream education."

The 1991 legislation established the special education system for learners "with special educational needs" (Decree-Law 319/91). This legislation aimed to develop "schools for all," and it increased the responsibility of mainstream schools "for the problems of learners with disabilities or learning difficulties" who should attend school in "the least restrictive environment." Additionally, it introduced the concept of "special educational needs" based on pedagogical criteria to substitute "the category-based classification, based on medical decisions." This legislation considered "special education" a pedagogical response to strengthen the individual autonomy of learners with physical and mental impairments.

In 2008 the legislation outlined specialized support for learners' special educational needs (Decree-Law 3/2008, Ministério da Educação, 2008). It aimed to further develop a "democratic and inclusive school, targeting the success of all children and young people," and defended a "flexible education system" to "guarantee equality in the access and in the results." Mainstream schools were expected to accept all learners. However, this legislation created a strong divide between learners "with Special Education Needs" and learners who could experience difficulties due to reasons that fell out with the "Special Education Needs" criteria. Using the language proposed by the International Classification of Functioning, Disability and Health (WHO, 2001) the policy defined "SEN learners" as those who needed specialized support provided due to "significant limitations at the level of activities and participation due to permanent functioning and structural alterations." Individualization and personalization were pedagogical approaches proposed to develop universal competencies and citizenship for all, as opposed to specialized support to promote the "biopsychosocial functioning" of a minority of learners "with SEN."

The 2018 changes to policy included a new legal framework for inclusive education (Decree-Law 54/2018, Ministério da Educação, 2018a). Using the terminology in this legislation, it aims to develop a second-generation inclusive school, where all learners have their potential, expectations and needs met, regardless of their personal and/or social situation. This new framework claimed to reject the need to categorize learners in order to provide the necessary educational responses. Instead, it proposed a tiered approach to support learning and inclusion, organized into: universal, selective and additional responses. The universal level involved changes to teaching and learning methods, resources, and assessment—it can include interventions targeting academic aspects or behavior. The selective level involved non-significant changes to the curriculum (aims and contents), psycho-pedagogical support, prior learning, or extra support interventions, and tutoring for learners perceived to be disengaged, and struggling. The additional level involved significant curricular adaptations (e.g., different curricular content) and should be available for

learners with considerable and persistent difficulties in communication, interaction and learning with a view to develop personal and social autonomy. The additional responses include Individual Educational Programs and Individual Transition Plans, attendance “by subject,” specialized resources and teaching. The enactment of this policy relies on the work of multidisciplinary teams that include teachers, “technicians” (e.g., Speech and language therapists), parents/carers and learners. These teams are described in the legislation as a pillar to the identification of responses and monitoring of its suitability and implementation. An interesting input of this legislation to pedagogies used in school was the indication that learning, and teaching should be designed based on the principles of Universal Design for Learning.

This wave of policy changes also brought other associated legislation which related to creating a more inclusive education system. One of these was the Decree-law 55/2018 ([Ministério da Educação, 2018b](#)), which established a revised curriculum and assessment principles. This legislation planned more flexibility in the way schools manage the curriculum and pedagogies. It proposed specific measures such as developing learners' citizenship and autonomy through project-based learning, diversifying work configurations, student outputs (oral, written, visual, multimodal) and an emphasis on formative assessment in order to deal with difficulties as soon as they emerge. This legislation points to another key policy—the “Learners' Profile at the end of compulsory schooling” (*Perfil dos Alunos à Saída da Escolaridade Obrigatória*) ([Ministério da Educação, 2017](#)). The “Students' Profile” conceived learning at the core of education and raised relevant issues concerning curriculum development, learner participation, and pedagogical innovation. This overarching policy is imbued with an inclusive approach, setting an expectation that learning and teaching should adapt to the individual's characteristics to create full participation, a sense of belonging, and equity in education and contribute to social inclusion and cohesion. The aim was to get all learners to reach the competencies outlined in the “Learners' Profile.” In this context, assessment for learning, that considers academic, behavioral, social, emotional and environmental factors, plays an important part.

Despite governmental changes over the years, the development of policies has steadily aligned with a vision of a mainstream school which should respond to the needs of all learners. International policies such as the United Nations Convention on the Rights of the Child ([United Nations, 1989](#)) and the Salamanca Statement (1994) have influenced and guided to some extent the route to a more inclusive education ([Alves, 2019](#)). In a context of national, centralized educational policies, the responses to some of the challenges to inclusion such as poverty, school failure, and drop out, have been tackled through projects and/or programs which encourage local solutions such as the National Program to Promote School Success (PNPSE—Programa Nacional de Promoção do Sucesso Escolar, 2016–Present), the area-based intervention “TEIP” (Territórios Educativos de Intervenção Prioritária 1996–Present), the Pedagogical Innovation Pilot Project (PPIP—Projeto-Piloto de Inovação Pedagógica, 2016–2020), and the Autonomy and Curriculum Flexibility Project (PAFC—Projeto de Autonomia e Flexibilidade Curricular, 2018–Present). Some developments concerning assessment policies in the last decades will be discussed in the following section.

Development of assessment

The Fundamental Law of the Education System (Law 46/86 of 14 October) provided the country with fundamental guidelines and with an agenda for education putting an end to 12 years of a certain deregulation that followed the democratic revolution of April 25, 1974. Legislation establishing fundamental curricular principles as well as new syllabi for Pre-School and Grades 1–12 (K-12) was issued in 1989 (Decree-Law 286/89, 29 August) followed by a major experimental and phase-designed process which was put into practice in a group of pilot schools for four academic years in a row from 1989-90 to 1992-93). In 1993–1994, after an evaluation of the experimental process, the new curricula for Grades K-12 were generalized to the whole educational system. For the first time since 1974 Portugal could enjoy a more coherent, logical and consistent legislation pointing to the need to integrate curriculum, pedagogy, learning, assessment, and teaching. This is illustrated by two historical pieces of legislation on learning assessment published in 1992 and 1993 (Legal Norm 98-A/92, 19 June, for Basic Education—Grades 1–9, and Legal Norm 338/93, 21 October, for Secondary Education—Grades 10–12). In fact, in these Legal Norms the legislator stated very clearly that formative assessments should prevail in the classrooms as a means to improve both teaching and learning. Summative assessments, on the other hand, should be considered as a means to collect information at specific points in time to balance out what learners knew and were able to do. Assessments, particularly those of a formative nature, began to be considered fundamental pedagogical processes to include rather than to exclude learners.

This progressive legislation was inspired by cognitivist perspectives on assessment as well as by sociocultural visions on learning and teaching emerging in the 90s ([Cardinet, 1991](#); [Berlak et al., 1992](#); [Gifford and O'Connor, 1992](#)). However, cultural, and pedagogical traditions, ancestral teaching practices and, above all, a generalized conception of assessment as a measure or as means to assign grades, soon revealed that teachers and schools needed to be supported in order to use assessments as pedagogical processes aimed at improving learners' learning and matching the legislation guidelines.

In both Legal Norms 98-A/92 and 338/93 external learning assessments, that is, assessments that were developed by some entity external to teachers and schools, were also regulated. The former established external assessments—*Provas Aferidas* (National Low Stakes Tests)—aimed at regulating the educational system as a whole and without any sort of influence in learners' ratings and academic progress. These assessments were designed for basic education learners (Grades 4, 6, and 9) and began to be administered in the Mathematics and Portuguese subjects only in the year 2000. The process of these assessments was never clear enough and had been an object of constant changes, concerning issues such as its nature, purposes, utilization, and population coverage. Quite recently, through Decree-Law 17/2016, these, and other issues of National Low Stakes Tests were clarified. Currently they are

criteria-based assessments and a means to provide learners and teachers with data aimed at helping them to improve their learning and teaching processes, respectively. Thus, they are external summative assessments that are being used in a formative way. As a consequence, they are more aligned with the idea of inclusion than its former counterparts that were created more than 20 years ago. Moreover, these are no longer assessments in Portuguese and Mathematics only. In order to avoid the so-called effect of curriculum narrowing, these assessments are now tackling all curriculum subjects from Mathematics and Portuguese to Social Sciences, Arts and Physical Education. Finally, rather than being administered at the final year of each cycle of basic education (Grades 4, 6, and 9) and in order to be consistent with the idea of its formative utilization, they are administered at toward the start or mid-cycles, at the end of Grades 2, 5, and 8.

These external, criteria-based assessments, aimed at providing learners, teachers, parents, and schools with feedback on learners' learning are more consistent with pedagogical assessments, as defined above. Therefore, it seems that there is now a new generation of external assessments in Portugal that could contribute to improve teaching and learning, to diminish inequalities, and to contribute to a more inclusive and democratic education. Nevertheless, in 2002 in the context of Basic Education, National Examinations were introduced in Portuguese and Mathematics at the end of Grade 9 (Decree Law 209/2002, 17 October). These examinations are still part of the current overall assessment architecture, but they might be considered as "low" or even "very low stakes" since their results only represent 30% of the determination of a student's final grade whereas the weight of internal assessments, a teachers' exclusive responsibility, is 70%.

Legal Norm 338/93 also introduced external assessments—*Exames Nacionais* (National Examinations)—at the end of secondary education. These are dual-purpose exams since they are aimed at both certifying learners' secondary education and selecting learners to access higher education. This legislation reintroduced the secondary school examinations into the system since its elimination after the April 25, 1974. These exams began in the 1995/1996 academic year and are still in use with the two above-mentioned purposes regardless of the slight adjustments that were introduced in the legislation throughout the last 25 years. Now a student must take two exams in Grade 11 and two in Grade 12. For the purposes of secondary school graduation these exams are low or even very low stakes since, as we have mentioned, they only represent 30% of the final grade which contrasts with internal assessments weighted at 70% of a student's overall grade. Typically, one or two of these examinations are specific to enter a given higher education course and, in this case, the student's grade to apply is computed based upon internal final grades, weighting 50% and external grades in the specific exams, also weighting 50%. Therefore, in the context of learners' access to higher education the exams are moderate stakes tests.

In recent years Portuguese public policies have been emphasizing the need to develop new and innovative curricular and pedagogical visions to build up a more democratic and inclusive educational system. As previously mentioned, in 2017, the above mentioned "learners' Profile" was issued, which is possibly one of the most comprehensive and meaningful curriculum materials ever published in the country. The document mobilizes, integrates and uses in a particularly successful way research recommendations carried out in education in the last decades. It defines *Principles* (e.g., Inclusion, Humanist Education), *Competency Areas* (Problem Solving, Esthetic and Artistic Sensitivity, Critical and Creative Thinking, Scientific, Technical, and Technological Knowledge), and *Values* (Responsibility, Citizenship, Participation). Indeed, it is a far-reaching set of guidelines for the development of the Portuguese mandatory curriculum (Grades k-12) which was the object of a wide debate in the country and eventually ratified through Legal Norm 6478/2017 ([Ministério da Educação, 2017](#)) or more if schools design a pedagogical project approved by the Ministry of Education. This document defines two important mandates to schools and school clusters. On one hand it calls for a more humanist, democratic, progressive, and inclusive education, and, on the other hand, it challenges the system to become one where learners have ample opportunities to learning to think.

The "learners' Profile" called for a consistent vision of pedagogy where issues of teaching, learning, and assessment should be integrated processes. This means that assessment rather than being essentially a means to grade learners should become a pedagogical process aiming at improving teaching and learning. Thus, formative assessment should prevail within all pedagogical actions that take place in schools and classrooms. Besides, assessments in general should be at the service of learning including summative ones. But that vision also meant that learning must be at the center of all teaching efforts and sociocultural perspectives on learning issues should be taken into account ([Lave and Wenger, 1991](#); [Wenger, 1998](#)). Consequently, in 2018, Decree-Law 55/2018 ([Ministério da Educação, 2018b](#)), provided a framework for both basic (Grades 1–9) and secondary education (Grades 10–12) curricula defining important pedagogical principles, namely on assessment organization and practice. Moreover, it created the real possibility for schools to autonomously manage up to 25% of the National curriculum ([Ministério da Educação, 2018](#)). The above discussed Decree-Law 54/2018 that defines the legal framework for inclusive education, completes a set of three important legal documents that represent the current public policies concerns to develop a more democratic and inclusive educational system.

To summarize, since 1986 the Portuguese public policies in education and, in particular, in learning assessment, seem to have evolved from a vision of assessment as grading or assessment as a measure, toward a perspective that sees assessment as a pedagogical process aimed at improving learning. Nevertheless, much still needs to be done to decrease inequalities among learners, particularly those related to their socioeconomic and cultural backgrounds which continue to have a disproportional impact on student achievement; as well as to support teachers' pedagogical practices who may face challenges regarding curricular and pedagogical innovations, inherent to the mentioned current legislation.

Assessment for inclusion or exclusion?

Until the mid 1970s there was an attempt to democratize and innovate in education, but exams remained the sole form of assessment and progression between cycles. Assessment served the purposes of grading, selecting and ranking learners (Fernandes, 2014). Exams have progressively lost their exceptional weight and, after 1974, they were practically abolished. They were reintroduced through public policies in 1993, and started being used in 1995/96 to certify secondary level learning and accessing higher education. From 2000 a number of initiatives were put in place to allow learners to develop and avoid grade repetition as the response to student difficulties (Fernandes, 2014). Nonetheless, grade repetition and early school dropout continue to be a challenge, despite consistent improvements (Conboy et al., 2013; Ferreira et al., 2015). There has been a trend of national policies, complemented by projects and/or programs that allow schools and teachers to adapt the national policies to their own contexts, at least since 1986. Those projects and/or program have been created over the years for example to try to respond to student diversity and avoid grade retention and early school dropout. Examples of these are the Inter-ministerial Program to Promote Educational Success (PIPSE, Conselho de Ministros, 1988), National Plan to Prevent School Dropout (PNAPAE, Ministério da Educação/ Ministério da Segurança Social e do Trabalho, 2004), More School Success Program (PMSE, Ministério da Educação, 2009), National Program to Promote School Success (PNPSE, Conselho de Ministros, 2016), and measures such as the Specific Tutorial Support (ATE, Gabinetes da Secretária de Estado Adjunta e da Educação e do Secretário de Estado da Educação, 2018). These initiatives were based on the presumption of schools being problem solving organizations (Skrtic, 1991) that consider uncertainty as an opportunity for innovation, based on collaboration and interdependence (Lloyd, 2000). While they can be ways to foster change toward more inclusive schools, they can give the appearance of add-on solutions to an already established system which resists change. Despite the voluntary or optional nature of these initiatives to develop inclusion and school and teacher autonomy, the adherence of schools and teachers is very substantial (e.g., Flexibility and Curricular Autonomy Project—PFAC, Project MAIA: Monitoring, Support and Research in Pedagogical Assessment). These projects help uncover differences between schools and school clusters both in terms of student population, and implementation of initiatives. While also revealing a requirement for Portuguese schools and teachers to be supported in their implementation of comprehensive visions of inclusion, which require rethinking curriculum, and pedagogy in terms of teaching, learning and assessment.

An inclusive education system, that promotes access to, participation in, and success of all learners requires a reconsideration of curriculum, and pedagogy. This reconsideration can be supported by Bourdieu's notions of capital (Bourdieu, 1986), as we should analyze which types and forms of knowledge are valued by schools. Assessment often crystalizes valued types and forms of knowledge and "assessment tools recognize the cultural practices of some while ignoring the cultural practices of others" (Kozleski and Waitoller, 2010, p. 663). The development of an inclusive school requires considerations of how assessment can become a proxy for cultural capital, and even though it is sometimes thought of as neutral, it is mediated by the assessors' biographies, and assumptions about knowledge, learning and learners (Kozleski and Waitoller, 2010). Similarly, to the existence of a national curriculum, the presence of external assessment is sometimes presented as a way to ensure more equitable education systems. However, external assessment on its own does not guarantee equity in terms of learning opportunities, it needs to be part of an overall vision and policy effort (Fernandes, 2014).

Another vulnerability and threat of external assessment to the development of inclusive schools, is the association of external assessment to notions of evaluating whether schools are effective and efficient. Lloyd (2000) considers a move away from the rhetoric of the "effective and efficient school" to a focus on developing "barrier-free, flexible, responsive, inclusive learning environments" is essential (Lloyd, 2000, p. 146). To achieve this, Lloyd suggests teachers need to be empowered and see themselves as agents of change who need themselves to work collaboratively, to adopt "... critical problem-solving approaches to practice" (Lloyd, 2000, p. 148). Since the 1990s, school-based teacher education centers (Centros de Formação de Associação de Escolas, Decree-Law 249/1992) have played a role in teacher continuous professional development. However, teacher learning for diversity (Szelei and Alves, 2018) and schools' use of assessment for learning needs to be further supported and teachers must be empowered. Teachers need the time and tools to reflect and analyze their school contexts, and to develop responses which are both consistent with research and literature recommendations and specific to their teaching and learning contexts. The implementation of collaborative action-research approaches for inclusion may demand considerable changes to the way we organize education times and spaces, and the way we comprehend relationships between teachers, and between teachers and learners. It will require valuing diversity, and questioning which types and forms of knowledge are valued—and aiming to diversify the valued knowledge. If we want to develop an inclusive education, assessment cannot be an afterthought and should not be a system on its own, it should play a part in the pedagogy message system of schooling (Lingard and Mills, 2007). Inclusion and assessment cannot be "add-on solutions" (Skrtic, 1991) in schools that resist changes, they need to be developed as a whole when planning learning and teaching.

Portugal has been a late comer to the Education for All movement, but has quickly made attempts to widen access to education for all, namely access and participation of learners with disabilities in mainstream education. Evidence suggests a continuity in Portuguese educational public policies toward equity in the past 35 years (Fernandes and Gonçalves, 2018), particularly in terms of access to education, and not as consistent in terms of success. The recent interrelated educational policies regarding curriculum, pedagogy and assessment are an important step and came to reinforce the commitment to inclusion. They provide a key pillar to avoid disjointed approaches to different learners and different "types" of diversity. The policy rhetoric of developing a truly inclusive school where all learners are welcome, are able to participate and learn, needs to be supported through flexibility for schools to adapt to their local learners and contexts, appropriate funding, and opportunities for teacher continuous professional development

(CPD). CPD should involve comprehensive in-service teacher education programs in which educators have ample opportunities to become familiar with, reflect and discuss policies, theories, research, and practices that support inclusive approaches to learning, teaching and assessment.

References

- Alves, I., Campos Pinto, P., Pinto, T.J., 2020. Developing inclusive education in Portugal: evidence and challenges. *Prospects* 49 (3–4), 281–296.
- Alves, I., 2019. International inspiration and national aspirations: inclusive education in Portugal. *Int. J. Incl. Educ.* 23 (7–8), 862–875.
- Alves, I., 2020. Enacting education policy reform in Portugal—the process of change and the role of teacher education for inclusion. *Eur. J. Teach. Educ.* 43 (1), 64–82.
- Assessment Reform Group, 2002. *Assessment for Learning: Research-Based Principles to Guide Classroom Practice—10 Principles*. Nuffield Foundation. Retrieved from: https://www.researchgate.net/publication/271849158_Assessment_for_Learning_10_Principles_Research-based_principles_to_guide_classroom_practice_Assessment_for_Learning. (Accessed 30 October 2021).
- Berlak, H., Newmann, F., Adams, E., Archbald, D., Burgess, T., Raven, J., Romberg, T., 1992. *Toward a New Science of Educational Testing and Assessment*. State University of New York Press, Albany, NY.
- Bernstein, B., 2003. *Class, Codes and Control, Volume V, The Structuring of Pedagogic Discourse*. Routledge, London.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess. Educ.* 5, 7–73.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood, Westport, CT, pp. 241–258.
- Cardinet, J., 1991. L'apport sociocognitif à la régulation interactive. In: Weiss, J. (Ed.), *L'évaluation: problème de communication*. Delval, Cousset (Fribourg), pp. 199–213.
- Conboy, J., Santos, I., Moreira, I., Fonseca, J., 2013. *Práticas e Consequências da Retenção Escolar: Alguns Dados do PISA*. Conselho de Ministros, 1988. Programa Interministerial de Promoção Do Sucesso Escolar. PIPSE, Lisbon.
- Conselho de Ministros, 2016. Programa Nacional de Promoção do Sucesso Escolar, Resolução do Conselho de Ministros n.º 23/2016. Lisbon.
- European Agency for Special Needs and Inclusive Education, 2017. *Agency Position on Inclusive Education*. Odense.
- Fernandes, D., Gonçalves, C., 2018. Para compreender o desempenho dos alunos portugueses no PISA, 2000–2015. In: Ortigão, M.I.R. (Ed.), *Políticas de avaliação, currículo e qualidade: Diálogos sobre o PISA*. CRV, Curitiba, Brasil, pp. 39–68.
- Fernandes, D., 2014. Avaliação Das Aprendizagens E Políticas Educativas: O Difícil Percurso da Inclusão e Da Melhoria. ML Rodrigues, pp. 231–268.
- Fernandes, D., 2019. Para um enquadramento teórico da avaliação formativa e da avaliação sumativa das aprendizagens escolares. In: Ortigão, M.I.R., Fernandes, D., Pereira, T.V., Santos, L. (Eds.), *Avaliar para aprender em Portugal e no Brasil: Perspectivas teóricas, práticas e de desenvolvimento*. CRV, Curitiba, pp. 139–164.
- Ferreira, A.L., Félix, P., Perdigão, R., 2015. Relatório Técnico: Retenção Escolar nos Ensinos Básicos e Secundários. In: Conselho Nacional de Educação, p. 89.
- Gabinetes da Secretária de Estado Adjunta e da Educação e do Secretário de Estado da Educação, 2018. Article 12 Apoio tutorial específico (ATE), Despacho Normativo n.º 10-B/2018, Lisbon.
- Gifford, B., O'Connor, M., 1992. *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*. Kluwer, Norwell, MA.
- Hart, S., Dixon, A., Drummond, M.J., McIntyre, D., 2004. *Learning Without Limits*. Open University Press, Maidenhead.
- International Bureau of Education-UNESCO, 2016. *Reaching Out to All Learners: A Resource Pack for Supporting Inclusive Education*. IBE-UNESCO, Geneva.
- Kiuppis, F., 2014. Why (not) associate the principle of inclusion with disability? Tracing connections from the start of the Salamanca process. *Int. J. Incl. Educ.* 18 (7), 746–761.
- Kozleski, E.B., Waitoller, F.R., 2010. Teacher learning for inclusive education: understanding teaching as a cultural and political practice. *Int. J. Incl. Educ.* 14 (7), 655–666.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Lingard, B., Mills, M., 2007. Pedagogies making a difference: issues of social justice and inclusion. *Int. J. Incl. Educ.* 11 (3), 233–244.
- Lloyd, C., 2000. Excellence for all children false promises! The failure of current policy for inclusive education and implications for schooling in the 21st century. *Int. J. Incl. Educ.* 4 (2), 133–151.
- Marques, M.M., Valente-Rosa, M.J., Martins, J.L., 2007. School and diversity in a weak state: the Portuguese case. *J. Ethnic Migrat. Stud.* 33 (7), 1145–1168.
- Miles, S., Singal, N., 2010. The education for all and inclusive education debate: conflict, contradiction or opportunity? *Int. J. Incl. Educ.* 14, 1–15.
- Ministério da Educação/ Ministério da Segurança Social e do Trabalho, 2004. *Plano Nacional de Prevenção do Abandono Escolar*. PNAPAE, Lisbon.
- Ministério da Educação, 2008. Decreto-Lei n.º 3/2008 de 7 de Janeiro.
- Ministério da Educação, 2009. *Mais Sucesso escolar: Projectos de combate ao insucesso escolar no ensino básico*. PMSE, Lisbon.
- Ministério da Educação, 2017. *Perfil dos Alunos à Saída da Escolaridade Obrigatória*. In: Ministério da Educação/DGE (Ed.), *Editorial do Ministério da Educação e Ciência*.
- Ministério da Educação, 2018a. Decreto-Lei 54/2018. Lisbon.
- Ministério da Educação, 2018b. Decreto-Lei 55/2018. Lisbon.
- Skrtric, T.M., 1991. The special education paradox: equity as the way to excellence. *Harv. Educ. Rev.* 61 (2), 148–206.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922.
- Susinos, T., 2020. Is participation a “sick word”? New insights into student democratic participation in light of research in Spanish schools. In: Allan, J., Harwood, V., Jorgensen, C.R. (Eds.), *World Yearbook of Education 2020—Schooling, Governance and Inequalities*. Routledge, Abingdon, pp. 200–216.
- Szelei, N., Alves, I., 2018. The missing link: teacher learning for diversity in an area-based initiative in Portugal. *Center Educ. Policy Stud.* J. 8 (3), 79.
- UNESCO, 2017. *A Guide for Ensuring Inclusion and Equity in Education*. UNESCO, Paris.
- United Nations, 1989. *Convention on the Rights of the Child*. United Nations, Treaty Series, vol. 1577, p. 3.
- United Nations, 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development*, 21 October 2015, A/RES/70/1.
- United Nations, 2016. General Comment No. 4 (2016), Article 24: Right to Inclusive Education. UN Committee on the Rights of Persons With Disabilities (CRPD), pp. 1–24.
- Wenger, E., 1998. In: Brown, J.S. (Ed.), *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, New York.
- WHO, 2001. *International Classification of Functioning, Disability and Health*. World Health Organization, Geneva.

Inclusive education policies and practices in Malawi and Zimbabwe

Tsitsi Chataika, University of Zimbabwe, Department of Educational Foundations, Harare, Zimbabwe

© 2023 Elsevier Ltd. All rights reserved.

Introduction	404
Inclusive education in Africa	405
Inclusive education in Malawi	405
The definition of inclusive education in Malawi	405
Existing policies and strategies for inclusive education in Malawi	405
Practices, systems and structures of inclusive education in Malawi	407
Pre-service and in-service teacher training	408
Strengths and gaps on policy and practice in inclusive education	408
Inclusive education in Zimbabwe	410
The definition of inclusive education in Zimbabwe	410
Existing policies and strategies for inclusive education	410
Practices, systems & structures for inclusive education	413
Pre-service and in-service teacher training	414
Strengths and gaps on policy and practices in inclusive education	415
Inclusive education strengths	415
Gaps in inclusive education	415
Financing education in Zimbabwe	415
Conclusions	416
Country-specific recommendations	416
Malawi	416
Zimbabwe	416
References	417

Introduction

Good quality education needs to be inclusive, ensuring that all learners participate in and achieve. During learning that is responsive to individual learner differences. Thus, inclusive education can act as a catalyst for change in educational practice, leading to improved learning outcomes (Ministry of Education, Science and Technology – Malawi & Save the Children. Despite having adequate legal frameworks, children from ultra-poor families, children and youth with disabilities, girls out of school (including teenage mothers), refugees and displaced children and youth, are particularly affected by exclusion and marginalization in southern Africa. Throughout the world, one billion people (15%) of the world's population, experience some form of disabling conditions, with 80% living in low income countries, southern Africa included (WHO and World Bank, 2011). Persons with disabilities, on average, are more likely to experience adverse socio-economic outcomes than persons without disabilities. Globally, people with disabilities are considered the world's largest minorities, albeit intersecting with other social categories: men and women, children, different races and cultures among all social classes; wealthy and poor, and highly educated and the illiterate (UNESCO, 2020). Historically, many children with disabilities have been excluded from basic mainstream education and access it in separate specialized schools (WHO and World Bank, 2011). This makes people with disabilities, especially children and youth, vulnerable to exclusion and stigmatization later in life.

Africa mainly comprises low-income countries where education is less likely to be obtained by children and youths with disabilities (UNESCO 2020). Over 87 million primary school-age children, secondary school-age adolescents and youth in sub-Saharan Africa, remained out of school in 2014. Member States supported by the UNESCO Regional Office for Southern Africa (ROSA) are among the poorest countries in the world with few available resources. Children and youth with disabilities face insurmountable challenges in accessing education and most importantly, accessing high-quality and equitable inclusive education (UNESCO, 2020).

Inclusive education implies increasing participation of learners in the culture, curricula and communities of their neighborhood learning centers that respond to learner differences; hence, leading to improved quality education (UNESCO, 2020). Global experience has shown that introducing inclusive practices cannot be done without introducing major system changes. The long term aim of creating a society with inclusive education, relates to the opening up of all mainstream schools and programmes so that they welcome and accommodate the learning of all members of the local community. This article draws upon inclusive education experiences from Malawi and Zimbabwe. Specifically, an analytical assessment on definitions and concepts, existing policies and strategies, practices and systems, strengths and gaps in the implementation of inclusive education in the two countries are examined. The article further offers recommendations that have implications on the provision of inclusive education in Africa.

Inclusive education in Africa

There are several regional legal frameworks in Africa. These include: the African Charter on the Rights and Welfare of the Learner (ACRWC) of 1999, which is similar to those in the Convention on the Rights of the Child, and the Southern African Development Committee (SADC) Policy Framework on Care and Support for Teaching and Learning (CSTL) of 2015. The CSTL focuses on the realization of the education rights of vulnerable learners and youth in the SADC region, by ensuring that schools are inclusive centers of learning, care and support; where every learner, especially the most vulnerable, have access to quality education. It provides a common understanding among Member States regarding the implementation of the initiative, and to guide efforts toward the realization of the CSTL goal across the SADC region.

Inclusive education in Malawi

Malawi signed and ratified international treaties and protocols relating to inclusive education, demonstrating the country's commitment to this noble cause. These include the United Nations Convention on the Rights of the Child (1989), the African Charter on the Rights and Welfare of the Child (1999), the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) (United Nations, 2006), Jomtien World Declaration on Education for All (1990), Salamanca Statement and Framework for Action, and the Dakar Education For All Framework for Action. The nation also endorsed the Sustainable Development Goals (SDGs) (United Nations, 2015), which promote access and affordable education for all boys and girls at all levels, including vocational training for vulnerable groups. This country-specific section, maps the policies, definitions and practices on inclusive education in Malawi.

The definition of inclusive education in Malawi

The concept of inclusive education is relatively new in Malawi and the country has been mainly associating it with special needs education (SNE) and students with disabilities. The recent National Strategy on Inclusive Education (NIES) 2017–21 defines inclusive education as a longer-term process of reforming the education system, cultures, policies and practices to address and respond to diverse needs of all learners (Ministry of Education Science And Technology). Inclusive education therefore is both a human rights and development issue, supporting the view that all children and young people belong together irrespective of their individual differences, based on race, gender, disability, social status, or religious beliefs (MoEST, 2017). This is consistent with Sustainable Development Goal 4 [SDG 4] (United Nations, 2015) and Article 24 of the UNCRPD (United Nations, 2006), which emphasize inclusive and equitable education.

Existing policies and strategies for inclusive education in Malawi

Malawi has shown commitment to ensure inclusive education at all levels, in line with global policies, frameworks and procedures. There are several domestic policies, strategies, plans and frameworks that have been developed and enacted by Malawi to spearhead inclusive education. These include:

- the Constitution of the Republic of Malawi (1995, as amended Act No.11 of 2010);
- the National Strategy on Inclusive Education (NIES) (2017–21),
- National Education Sector Plan (NESP) (2017–20);
- National Disability Mainstreaming Strategy and Implementation Plan (2018–23);
- Child Care, Protection and Justice Act (Act No.22 of 2010);
- The Disability Act (Act 8 of 2012), the Education Act (Act No. 21 of 2013);
- Vision 2020, Special Needs Education Implementation Guidelines (2008);
- National Education Standards-Primary and Secondary Education (2015);
- National Policy on the Equalization of Opportunities of Persons with Disabilities (2006);
- National Policy on Early Childhood Development (2017);
- National Youth Policy (2013);
- National Gender Policy (2011, Second Edition, 2015);
- National Integrated School Health and National Policy (2017);
- National School Health and Nutrition Strategic Plan (2018–22); and
- The Draft 2019–20 Multi-hazard Contingency Plan.

In responding to the challenges of SDG 4, which speaks about the aspiration to ensure inclusive and equitable quality education and promoting lifelong learning opportunities for all, Malawi developed the *National Strategy on Inclusive Education [NIES 2017–21]* (Government of Malawi MoEST, 2017). This flagship strategy aims at ensuring that all learners have equitable access to quality

education in inclusive settings at all levels of the education system in Malawi (GIZ and Government of Malawi, 2019). The strategy identified eight key areas that inform inclusive education and makes it a reality. These are:

- Capacity for inclusive education;
- Governance and Management for inclusive education;
- Learner identification and assessment;
- Inclusive Education Management Information System;
- Teacher education and motivation;
- Partnership for inclusive education;
- Enabling environment for teaching and learning; and
- Financing inclusive education.

Section 1.2.1 of the national strategy affirms that exclusion exists at all levels from early childhood development (ECD), through primary, secondary and higher education levels. Section (1.2.1) focuses on learners excluded not only from school, but also within the school, and therefore identifies a total of 20 groups of such learners. These include children with disabilities, children from ultra-poor households and marginalized girls.

The Constitution of the Republic of Malawi (1995, as amended Act No.11 of 2010) provides for the right to education for all citizens. Chapter IV, Section 20 prohibits discrimination of any form (Government of Malawi, 2010). The National Education Sector Plan (NESP) acknowledges lack of access to education among learners with disabilities at ECD, primary and secondary levels of education, non-formal education (such as adult literacy), and for out of school youths. Therefore, the NESP sets out to improve access and equity to ECD by promoting early detection, intervention and inclusion of children with specific health and educational needs, supported by inclusive curricula (JICA Malawi, 2020). The NESP further realizes the need to have targeted activities for children and adults with disabilities and provide teaching and learning materials for both in order to promote access and equity to non-formal education (JICA Malawi, 2020). The current NESP and its predecessors are widely credited to have informed the development of the five year NIES (2017–21). The NIES addresses the needs of learners with disabilities at all levels of the education system in a more comprehensive manner, to ensure quality inclusive education.

The National Disability Mainstreaming Strategy and Implementation Plan (2018–23) identifies ‘Access to Education’ for learners with disabilities as its Priority 2 area, under Section 5.2.2 (Ministry of Gender, Children, Disability and Social Welfare [MoGCDSW], 2018). This implies that the National Disability Mainstream Strategy and Implementation Plan have prioritized access to inclusive and equitable education in order to break barriers that restrict learners with disabilities from accessing their right to education.

The Child Care, Protection and Justice Act (Act No.22 of 2010) makes provision for childcare, protection, child justice; and for matters of social development of the child and related matters. The *Malawian Education Act No. 21 of 2013* stipulates that education should be for all people irrespective of any discriminatory characteristics such as race, disability, ethnicity or gender (Government Printers, 2013). It strives for accessible, equitable, relevant and inclusive education. *Vision 2020* represents Malawi’s long term development strategy. Chapter 7, Section 7.2.2 of *Vision (2020)*, makes provision for the inclusion of learners with disabilities at all levels of education (Government of Malawi, 2020). It sets strategic options with regards to the improvement of SNE and thus ensures appropriate designs to cater for people with disabilities. The *SNE Implementation Guidelines of 2008* recognizes education as a basic human right. It focuses on eight priority areas:

- early identification,
- assessment and intervention,
- advocacy,
- care and support,
- management,
- planning and financing,
- access,
- quality, equity and relevance (Itimu-Phiri and Kopetz, 2008).

The *National Education Standards-Primary and Secondary Education of 2015* policy specifies expected outcomes for learners to be delivered by all education providers in public and private institutions (MoEST-Directorate of Inspection and Advisory Services, 2015).

The National Policy on the Equalization of Opportunities of Persons with Disabilities of 2006 supports and encourages inclusive education, provision of assistive technology, adaptation of communication systems, training of specialist teachers, and inclusion of special needs education (now inclusive education) in teacher training curriculum and establishment of accessible resource rooms in schools (Ministry of Persons with Disabilities and the Elderly).

The National Policy on Early Childhood Development of 2017 has a number of guiding principles, one of which is that; “No child shall be discriminated against, on the basis of age, sex, race, health status, economics, religious or political affiliation in the provision of ECD services by any organization” (Government of Malawi, 2017). This implies that no child should be left behind when it comes to accessing education. *The National Youth Policy of 2013* provides interventions for both in-school and out-of-school youth, regardless of their circumstances.

The National Gender Policy of 2011 (revised in 2015), seeks to mainstream gender issues in the national development process to enhance participation of boys and girls, women and men for the attainment of sustainable and equitable development (Malawi National Gender Policy (2011, Second Edition, 2015).

The National School Health and Nutrition Strategic Plan (2018–22) sets out the implementation of the 2017 National Integrated School Health and National Policy. The plan supports marginalized children in school with regard to their health and nutrition (MoEST, 2018).

The Draft 2019–20 Multi-hazard Contingency Plan was crafted under the Department of Disasters Management Affairs (DoDMA) of Malawi. The plan has the Education Cluster in Section 3.3 to ensure the provision of education in emergencies, with stakeholders involved including MoEST, UNICEF, ActionAid, Save the Children, Civil Society Education Coalition (CSEC), World Vision and Mary's Meals responsible for this cluster.

These policy documents are governing inclusive education in Malawi, with the view of '*not leaving anyone behind*' in accessing education. The next section focuses on practices, systems and structures of inclusive education in Malawi.

Practices, systems and structures of inclusive education in Malawi

Inclusive education in Malawi is implemented at all levels, from national level down to the school level. At school level, there are SNE teachers and itinerant teachers, who teach learners in resources units and/or special classes (MoEST, 2016). Other learners with disabilities are taught in mainstream classes, depending on the extent of their impairments. The itinerant teaching program was started by the MoEST in collaboration with Sightsavers, which is an international non-governmental organisation, whose headquarters are in the United Kingdom. Sightsavers works in more than 30 countries to prevent avoidable blindness and fight for the rights of people with disabilities.

With the itinerant teaching program, specialist teachers provide basic training to teachers in mainstream schools on how to improve inclusion in the classroom (Banks and Zuurmond, 2015). Also, itinerant teachers provide some direct support to children with profound visual impairments in mainstream schools within their catchment area (Research for Inclusive Education in International Cooperation, 2015). At district level, the SNE coordinator uses tools and procedures that are not available in all districts in Malawi. Each coordinator has self-defined assessment tools and procedures, which s/he shares with specialist and itinerant teachers within the district. The SNE coordinator does not directly work together with the health personnel and social welfare officer to develop and conduct district level assessment and procedures, unless there is a crisis or an emergency. At national level, the Department of Special Needs Education (DSNE) has an oversight role on the implementation of inclusive education. Specifically, DSNE ensures:

- Establishment of specialist and itinerant teachers;
- Coordinating programmes of inclusive; and
- Issues related to management of SNE diploma teachers.

Malawi has taken a twin-track approach to inclusive education, where children and youth with severe impairments are educated in special schools or special needs centers, while those with mild impairments are mainstreamed (UNESCO, 2020). The Education Sector Implementation Plan II aims to strengthen inclusive education in all schools to avoid segregation. As a result, special schools at each education level are being transformed into resource centers, as specified in the 2007 National Policy on Special Needs Education. Currently, there are 150 resource centers across the 34 districts of Malawi (GIZ and Government of Malawi, 2019). Special schools also provide disability-specific instruction, primarily for the hearing and visually impaired learners. However, there are only six special schools in Malawi, all of which are residential schools, while others have been converted into resource centers to reflect a paradigm shift in policy.

Education Information Management Systems (EMIS) data collected annually, provides the basis for the monitoring framework for education. EMIS in Malawi was established in 1986 to generate quality information for planning, policy formulation and decision making (Kimura, 2005). At school level, records are collected from the class registers, aggregated and displayed in the head teacher's office. Such records are submitted to the District Education Manager (DEM)'s office. Also, the district level aggregated school level data, are merged with distinct school names and "used" for resource allocation and reporting to the Divisions and Headquarter (GIZ and Government of Malawi, 2019). At national level, the EMIS questionnaire includes the following:

- Section C on learners (c.3 orphans by standard and gender and vulnerable learners by standard and gender; and c 4 on special learning needs by standard and gender);
- Section D on infrastructure and sanitation information (D.1.a special needs resource rooms and special needs preparatory rooms; D.1.c number of disability friendly toilets; D.2.b rooms with ramps and D.2.d newly constructed infrastructure – special needs resource room and special needs preparatory rooms); Section F 3 teachers' details (in terms of disability); and
- Section F4 SNE teachers' details (GIZ and Government of Malawi, 2019; MoEST, 2018).

Malawi, through the MoEST, is adapting the existing education curriculum and thus, incorporating inclusive education within it to improve teaching and learning environments toward inclusivity. This is in line with the NIES (2017–21) Priority Area 7, 'Enabling Environment for Inclusive Education', which seeks to promote inclusion through the following strategies:

- 7.1- Adapting existing curriculum methodology, teaching and learning materials to incorporate inclusive education;
- 7.2- Equip educators, caregivers, learners and community with knowledge and skills to support inclusive education; and
- 7.3- Construct and rehabilitate infrastructure water points, sanitation facilities and resource centers.

The 2019 Capacity Assessment Report on inclusive education in primary schools, surveyed 557 schools across 33 Malawian districts. The results indicate that there is a considerable number of schools with the NIES (2017–21) document and the inclusive education curriculum materials for reference in implementing inclusive education (GIZ and Government of Malawi, 2019).

Pre-service and in-service teacher training

Pre-service training ensures that prospective teachers are skilled in inclusive teaching pedagogies when deployed to schools, while in-service teachers who receive continuous professional development (CPD) training are capacitated in inclusive pedagogy while practicing. Most of these specialist teachers have been trained by Montfort Special Needs Education College, which is the prominent training institution for teachers specializing in special needs education (Chitiyo et al., 2019). CPD was noted by the NIES (2017–21) as key to the realization of inclusive education. This is because mainstream primary school teachers who would have inadequate skills to handle children with diverse needs, would be reluctant to accommodate these children in their classrooms. Therefore, the CPD initiatives in Malawi adopts UNESCO's definition of inclusive education as a process of addressing and responding to the diversity of needs of learners through increasing participation in learning, cultures and communities, and reducing exclusion within and from education (Chitiyo et al., 2019).

Strengths and gaps on policy and practice in inclusive education

The Malawian inclusive education policies and strategies are backed by implementation plans, which promote accountability and therefore yield stakeholder political will. The implementation plans of these policies and strategies provide specific, measurable, achievable, realistic and time-bound (SMART) targets for accomplishment by the responsible parties. The National Health and Nutrition Strategy Plan (2018–20) has the monitoring and evaluation plan in Section 5 and Implementation Plan in Section 6; the National Disability Mainstreaming Strategy and Implementation Plan (NDMS & IP) (2018–23) has the implementation plan in Section 6. The National Integrated School Health and Nutrition Policy (2017) has implementation arrangements in Section 4, as well as the implementation plan as Annex 1. The backing of the inclusive education policies with implementation plans can be credited to the landmark swift development of the National COVID-19 Preparedness and Response Plan of 2020 in response to COVID-19 pandemic, covering education in emergencies under its Education Cluster in Section 4.7, in line with the Draft 2019–20 Multi-hazard Contingency Plan. This Education Cluster strives to, 'ensure that teaching and learning continues through innovative solutions and creating an enabling environment in communities with special attention given to vulnerable groups in the education sector during the COVID-19'.

The NIES (2017–21), the flagship strategy for inclusive education and education in emergencies in Malawi, has estimated the cost of inclusive education outcomes, as per its Priority Area 8, which has to date yielded positive results toward inclusivity in education. Malawi developed a costing model for inclusive education in its NIES strategy document, which estimated the cost at US\$29 million over five years. In terms of funding for inclusive education, the 2019 Capacity Assessment of inclusive education in 557 primary schools across Malawi's 33 districts revealed that sources of funding were: school improvement grants (63%); school improvement plans (24%); development fund (9%); and parental assistance (4%); thereby demonstrating political will for implementing inclusive education. This has resulted in progressive transition of the country from special schooling (discriminatory) to mainstream schooling/resource centers (inclusive) for learners with diverse needs. However, several researches have noted inadequate funding for inclusive education. The 2019 Capacity Assessment of inclusive education revealed existing gaps for funding, which include inadequate materials due to high enrollment (22.5%), lack of support materials and assistive technology (56.5%) and lack of good and accessible infrastructure (21%) (GIZ and Government of Malawi, 2019).

The greatest strength of the Malawian NIES (2017–21), is its broadness. Such an approach of not only focusing on barriers to access, but also barriers on inclusivity and participation within the schools is a good practice from the Strategy. The Strategy also covers all children likely to be excluded from and within the education system. The 2015–19 National Education Plan endorses an inclusive approach, referring to children and youth who have been marginalized or excluded, such as girls, children with disabilities, people living in remote villages and those from poor households (UNESCO, 2020). As indicated earlier, the NIES covers a total of 20 groups of learners with diverse needs or marginalized learners. This is the reason why, since its (NIES) inception, there has been an increase on enrollment of learners with diverse needs in Malawian schools. At national level, EMIS 2016/17 reports that 2.8% (146,048/5,073,729) of total primary enrollment were learners with diverse needs compared to the EMIS 2015/16 enrollment of 2.4% (120,007/4,901,009) enrolled. Distribution of learners with disabilities by gender shows that the majority of learners were having learning difficulties [62,767], followed by learners with visual impairment (low vision [35,234]; blind [496]), hearing impairment (hard of hearing [31,434]; deaf [3414]); and lastly those with physical impairment [12,801] (GIZ and Government of Malawi, 2019).

The major limitation of the NIES (2017–21) is that it is available at national level and sparingly at district and school levels. What defeats the intention of advancing inclusive education, is that those at MoEST headquarters are more familiar with the document and not in districts and schools (Save the Children, 2018). Some personnel at district level may have copies of the Strategy. However, this does not translate into standardized transmission of its contents to districts and schools (GIZ and Government of Malawi, 2019). From the 557 primary schools across 33 Malawian districts surveyed by the 2019 Capacity Assessment of inclusive education in primary schools, 84.3% reported that inclusive education guides or handbooks were not available for the school

manager (*ibid*). This defeats the goals of the NIES spearheading inclusive education, while the lack of adequate IEC materials compromises achieving inclusive pedagogies in schools.

While the country is committed to capacitating structures and systems for inclusive education as per NIES Priority Area 1, policy gaps in other policy documents work against the promotion of inclusion. Also, only 150 primary schools nationwide have resource rooms, which are not sufficiently equipped as they lack essential materials to provide quality learning (*ibid*). Also, there are 1742 specialist teachers serving in the education system mainly in primary education and these are not enough. In some cases, learners with visual impairment were noted by the 2020 GEM Report Education Profiles to have moved to schools for the blind while those with albinism are mainly educated by itinerant teachers for the blind (Lynch and Lund, 2011). Kaphle et al. (2015) established that the majority of students with refractive errors, such as myopia, were not wearing glasses in Malawi. These highlighted shortfalls undermine the commitment of the country toward inclusive education.

The NIES (2017–21) Priority Area 8 (Financing inclusive education) is not regarded as mandatory, thus creating a ‘closed gateway’ for not making provisions for inclusive education. The aggregated School Improvement Grants (SIGs) at the district are not specific on inclusive education, but more with the NESP’s thematic areas (access and equity, quality and relevance, and governance and management). In a few cases, District Education Managers (DEM) advise primary schools to allocate a portion of their funding to inclusive education; and failure to allocate funds for inclusive education results in re-budgeting. However, the decision is usually with the school management to allocate the School Improvement Grant (SIG) to inclusive education. In the same vein, at national level there is no deliberate intent and action to ring-fence inclusive education although SNE has a defined resource package. Therefore, as noted, both district and national level the inclusive education budget and its allocation is at the whim of the controlling officer and/or budget officer if it is to get any funding (GIZ and Government of Malawi, 2019). The 2019 Capacity Assessment of Inclusive Education revealed that in 557 primary schools across Malawi’s 33 districts, revealed the funds were; adequate (0.19%), inadequate (29.4%), no funding (62.69%) and no response (8.08%) (*ibid*). The existing gaps for funding were identified and included inadequate materials due to high enrollment (22.5%); lack of support materials and devices for inclusive education (56.5%); and lack of good infrastructure (21%) (*ibid*). This affects the implementation and subsequent realization of inclusive education.

Although some sensitisations were applied to raise awareness of inclusive education as per NIES (2017–21) Priority Area 1 (Capacity for Inclusive Education), there are still inherently negative attitudes and lack of political will with regard to the implementation of inclusive education (Save the Children, 2018). Also, support from the district to the schools for children with disabilities was noted to be low, which limits efforts toward the realization of inclusive education. Again the same assessment revealed that of the 557 surveyed primary schools, 50% were not visited, hence not assisted by the district SNE coordinator/desk officer, while 33.4% indicated non-support from the DEM. These are worrying trends as they signify lack of political will by these key stakeholders (Save the Children, 2018). Also, 51.92% head teachers indicated lacking knowledge on policy and the legal framework relating to inclusive education, while only 49.69% of the 557 schools had by-laws on inclusive education. This too, militates against the implementation of inclusive education (GIZ and Government of Malawi, 2019). The Policy on ECD also (2006) lacks commitment to inclusive education by failing to embed provisions for early intervention in ECD for children with diverse needs. This implies that community based child care centers will hardly contribute to inclusive practices; hence, impact on children with diverse needs’ access to education.

The national level situation reflects a “silo” approach toward inclusive education; thus, defeating the collaboration and partnership embraced by the NIES on Priority Area 6 (partnership for inclusive education). Three key directorates (Basic Education, Education Planning and SNE) of the education ministry were noted by the 2019 Capacity Assessment of inclusive education in Malawi as apparently not working as one entity. The assessment established that Directorates of Basic Education and Planning view themselves as ‘outsiders’ that only come to offer advice and support on financial and management matters (GIZ and Government of Malawi, 2019). The same capacity assessment further observed that the Directorate of SNE has been solely championing every inclusive education activity without the full participation and collaboration of other relevant players. This non-collaboration of these critical departments to deliver inclusive education, negatively impacts on the realisation of SDG 4.

While it is plausible that the EMIS data is somehow inclusive, it is both out-dated and overemphasizes SNE over inclusive education. The EMIS data in Malawi is collected at national, district and school levels. At the national level, the EMIS questionnaire includes Sections from C3 to F3, which speaks to inclusive education. The presence of data at all levels indicates a positive development, which is highlighted by UNICEF (2016a, p.12), stating that, “it is important to take stock of what is already known”. However, the EMIS data seems not to embrace all categories of learners with diverse needs under the banner of SNE. By this, it ignores the 20 various categories of learners with diverse needs, which are usually excluded from and within schools under section 1.2.1 of the flagship inclusive education strategy the NIES MoEST (2017). As indicated earlier, the EMIS data is out dated, with the latest being for 2016/2017. Given the importance of this data in informing policies and strategies for inclusive education, this is a gap that should be closed.

In terms of full practices for inclusive education in Malawi, there is still a high degree of non-inclusivity of education environments as well as in the general curriculum. A recent study of the TVET system, which included interviews in 15 public, private and community technical colleges, as well as community skills development centers, identified obstacles to inclusion (UNESCO, 2020). Facilities lacked accessibility features such as: ramps; wide doorways; good pathways between buildings, and disability-friendly

toilets. Learning equipment instructions were not available in Braille, and learning materials were not available in large print, while discriminatory employer attitudes and behaviors were also evident (ibid). According to, cited in the GEM 2020 Report, in Malawi, students with disabilities cannot benefit from information and communications technology (ICT), computer studies or science in the general curriculum because of the lack of resources.

What is clear is that the concept of inclusive education is fairly new in Malawi, such that most mainstream teachers, communities and other stakeholders are not aware of it. While some challenges have been identified, key to note is that inclusive education is a process and not an end (MoEST, 2016). While promoting inclusive education, it is vital to note that inclusive education as a concept, demands a change of mindset about education. It becomes important for Malawi to ensure that all stakeholders work together in order to ensure that no child is left behind.

Inclusive education in Zimbabwe

Zimbabwe had been seen as exemplary in the education sector, with one of the highest literacy rates in Africa (around 90%). However, there is an estimated 600,000 children with disabilities of school age in Zimbabwe. More than half have no access to education (WHO and World Bank 2011). One study suggests that children with significant impairments have been turned away from schools in Zimbabwe because teachers perceived themselves as untrained and ill-equipped to assist them (Chakuchichi and Chataika, 2013). In order to try to address this issue, the Leonard Cheshire Zimbabwe Trust and partners have been implementing an inclusive education program in Zimbabwe since 2009.

Within the Ministry of Primary and Secondary Education (MoPSE), inclusive education is the responsibility of the Learner Welfare Services, Schools Psychological Services and Special Needs Education Department. It is notable that the Constitution (2013), the Education Act (as amended in 2020), and various policy documents guarantee the right to basic education for all. The Zimbabwean Constitution of 2013 allows all learners to be enrolled at schools nearest their homes irrespective of diversity status. Thus, all schools are expected to be inclusive and a considerable number of marginalized learners and those with disabilities are enrolled in inclusive classrooms within mainstream schools. Zimbabwe has 16 official languages, with English, Shona, and Ndebele the most common. Sign Language is one of the 16 official languages in the national Constitution (Government of Zimbabwe, 2013).

The last two decades have seen Zimbabwe going through economic crises, natural disasters and political emergencies that created more education vulnerability among the usually excluded children with disabilities, those with albinism, girls, among others (Chataika et al., 2020). These crises include continuous economic crises with high unemployment rates, the land reform program, floods, frequent droughts, cyclones and the COVID-19 epidemic. Thus, it is only through an inclusive education system that the needs of historically disadvantaged learners can be adequately addressed during such emergencies.

The definition of inclusive education in Zimbabwe

The concept of inclusive education is embedded in Section 3.9 of the National Disability Policy of 2021. It does not define inclusive education as such, but indicates that an inclusive education system of appropriate standards, at all levels, as well as lifelong learning for persons with disabilities of all gender affiliations, must be ensured. In Zimbabwe, inclusive education is defined in the 2021 draft Inclusive Education Policy as:

An approach or process that seeks to ensure universal access to and participation in education for all learners through mainstreaming of responsive learner-centred support to overcome identified barriers to individual learners' full participation in formal, non-formal and lifelong educational settings and activities.

The challenge is that, it is still a Bill, which still needs to be enacted. Inclusive education is also enshrined in the Education Amendment Act of 2020 as a human rights provision. This is consistent with the Sustainable Development Goal (SDG) 4 (United Nations, 2015) and Article 24 of the United Nations Convention on the Rights of Persons with Disabilities (CRPD) (United Nations, 2006), emphasizing inclusive and equitable education. The country ratified the CRPD on September 23, 2013.

Existing policies and strategies for inclusive education

Zimbabwe has policy documents that have a direct bearing on the provision of inclusive education in Zimbabwe (Presented in Table 1).

Inclusive education may also be indirectly influenced by the following legal documents in Zimbabwe. Chief Education Officer's Circular Minute Number 12 of 1987 (Schools Psychological Services Institutionalizing Remedial Programmes) and Secretary's Circular No. 20 of 2001 (Replacement of Special Education Teachers).

The Ministry of Primary and Secondary Education (MoPSE) has a draft Inclusive Education Policy and a Practical Inclusive Education Handbook for Primary and Secondary School: The Policy has been out for public consultation forums, and will be going before Parliament soon. It is hoped that the policy will be enacted as soon as possible so that the implementation plan in the policy document can be implemented. MoPSE also produced an inclusive education handbook in 2020 to strengthen the practice of inclusive enrollment, teaching and learning processes, as well as learner support services at every school through continuous professional development. Thus, the handbook intends to strengthen teacher capacity in order to translate the curriculum principles of inclusivity, equity, and gender sensitivity and learner-centredness, among others, into practical reality. This is to ensure a strong foundation for inclusive societies, which is in line with the Constitution of Zimbabwe (MoPSE, 2021).

MoPSE has curricula that consider issues of equity. The curriculum framework (2015–22) has inclusivity as one of its guiding principles, calling for diversity education. The curriculum spelled out one of the competencies expected of learners as demonstration of an understanding of cross-cutting issues such as HIV and AIDS, sustainable resource utilization, climate change and disaster risk management. These issues are adequately being addressed in all the learning areas, especially at primary school level. Within the principle of inclusivity, the aim of the curriculum is to provide an education system that considers and addresses the different learners' needs and abilities without disadvantaging any group or individual. MoPSE, in conjunction with UNICEF has also developed a downloadable handbook that gives guidelines to educators on the implementation of inclusive education (MoPSE, 2021).

Within the MoPSE, inclusive education and education in emergencies are the responsibility of the Learner Welfare Services, Schools Psychological Services and Special Needs Education Department. It focuses on enabling all learners to access the educational facilities on offer to achieve their individual best through a range of support services.¹ However, education in emergencies is not mentioned anywhere on the function of the department, leaving it as a gray issue.

Alternative Learning Approaches, such as online lessons and distance learning approaches through radio and Television ensure that the country has procedures in place to respond to emergencies/disasters, which negatively impact on learning. This is mostly executed by the Education Cluster, led by MoPSE, co-led by Save the Children and UNICEF and comprising more than 22 partner Zimbabwe Education Cluster (2020). This structure is responsible for implementing the education in emergencies programming in Zimbabwe. This has been the approach which has been used in the face of recent emergency disasters and setbacks such as Cyclone Idai and the novel COVID-19. These disasters have exposed the exclusionary nature of the education system in Zimbabwe. The COVID-19 pandemic led to the closure of 9625 primary and secondary schools, and this was responded to by the crafting of the Education Cluster Zimbabwe COVID-19 Preparedness and Response Strategy (UNESCO, 2020c). This strategy devised various alternative learning and safe schools opening approaches under its five approaches. These are Alternative Learning Approaches (radio programming; digital & online learning; and provision of supplementary learning materials); Back to School and Opening Up Better; Safe School Feeding; Teacher Capacitation and Raising awareness for prevention and mitigation through information, education and communication materials. However, the challenge was that most learners with disabilities had to leave behind their

Table 1 Relevant policy documents.

<i>Document</i>	<i>Emphasis</i>
Constitution Amendment Act (2013)	• Section 6- recognition of sign language and 12 other minority languages as official languages • Section 19- appropriate education and training for all children • Section 20- appropriate education and training for all youths • Section 22- accessible public buildings, use and development of all forms of communication for people with disabilities • Section 27- free and compulsory education for children, same educational opportunities for boys and girls at all levels • Section 75- right to basic State-funded education • Section 83- rights of persons with disabilities to be provided special facilities for their education, and the right to state-funded education and training

(Continued)

¹See <http://mopse.co.zw/department/psychological-services-special-needs-learner-welfare-department>

Table 1 Relevant policy documents.—cont'd

Document	Emphasis
Education Act (1987, as amended 2020)	• Section 4: No child shall be discriminated against by the imposition of onerous terms and conditions in regard to his or her admission to, suspended, excluded or expelled from any school on the grounds of his or her <i>nationality, race, color, tribe, place of birth, ethnic or social origin, language, class religious belief, political affiliation, opinion, custom, culture, sex, gender, marital status, age, pregnancy, disability or economic or social status, or whether they were born in or out of wedlock.</i> • provision of sanitary wares and other menstrual health facilities to girls in all schools to promote menstrual health. • Section 5: Compulsory education, every child shall be entitled to compulsory basic state funded education parents who deprives their child the right to basic state funded education shall be guilty of an offense and liable to a fine. • Section 10: Every child of school going age shall be entitled to be enrolled at the primary or secondary school, as the case may be, <i>nearest to the place</i> where he or she is ordinarily resident, unless such primary or secondary school is fully enrolled. • Section 25: Every school shall endeavor to offer non formal education including adult education. • Section 62: Every school shall endeavor to— (a) Teach every officially recognized language. (b) Ensure that the language of instruction shall be the language of examination. (c) Ensure that the mother tongue is to be used as a medium of instruction at early childhood education. • Section 68B Pupil with disabilities 1) every registered school shall provide infrastructure, subject to availability of resources, suitable for use by pupils with disabilities. 2) the Secretary shall monitor and enter premises of every registered school for the purposes of ascertaining whether the rights of pupils with disabilities are considered during teaching and learning. 3) for the purposes of fees approval, the Secretary shall require every registered school to submit a plan highlighting how the school shall advance the rights of pupils with disabilities. • Section 68C No pupil shall be excluded from school for non-payment of school fees or based on pregnancy.
Draft inclusive education policy of Zimbabwe (2019)	• the comprehensive policy draft covers all traditionally excluded and marginalized learners such as those with disabilities, migrants, OVC, children affected by HIV and AIDS, disadvantaged children in rural and urban areas, girls, out of school youths, those with learning difficulties and refugees. • the policy calls for schools to become inclusive centers of teaching, learning and support, and with the intention to ensure allocation and effective realization of the objectives of improved access, retention, and quality educational outcomes. The first phase of policy implementation will be focusing on the following six core areas: • infrastructure • leadership and structures • material support (inclusive of learning materials and support) • Safety and protection (including Psychosocial support) • nutrition and health promotion • water, sanitation and hygiene
Zimbabwe education cluster COVID-19 Preparedness and response strategy	The strategy has five priorities guiding education response to COVID-19. These are: • alternative learning approaches (radio programming; digital & online learning; and provision of supplementary learning materials) • back to school and opening up better (infection prevention & control for safe schools; learners, teacher and staff wellbeing & welfare; psycho-social support; assessment & accelerated learning for academic reintegration; and monitoring, learning and building back better. • safe school feeding (supplementary utensils and food preparation materials; supplementary food items; emergency school improvement grants; and training for school communities • teacher capacitation (radio, digital and online programming; back to school integration; learner welfare; and prevention and mitigation. • raising awareness for prevention and mitigation through information, education and communication (IEC) materials (behavior change communication; rapid assessment and engagement of publics)

Table 1 Relevant policy documents.—cont'd

Document	Emphasis
National disability policy 2021	Section 3.9 outlines Government's obligations toward people with disabilities including the provision of inclusive education, inclusive infrastructure in schools and provision of assistive technologies
Education sector strategic plan (2016–20)	The comprehensive plan focused on promote the development of the sector for the benefit all the learners especially the vulnerable and marginalized who face the risk of not going to school.
Zimbabwe school health policy (2018)	Provides a broad frame of reference to guide the implementation of a number of health related issues relating to the welfare of learners in the school system, such as health and nutrition, education services, water, sanitation and hygiene, needs of learners with disabilities, mental health, sexual and reproductive health concerns, and the care and support provisions as well as guidance and counseling needs of all learners.
Curriculum framework for primary and secondary education (2015–22)	The framework considers inclusivity as one of the guiding principles and defines inclusivity as an education system that considers and addresses the different learners' needs and abilities without disadvantaging any group or individual
The national gender policy (2013–2017)	Fosters a gender just society in which women, girls, men, and boys should equally contribute to and benefit from all the spheres of country development.
Disaster risk management: A resource book for educational institutions in Zimbabwe (2009)	The resource book is specifically for educational institutions. It provides an inclusive response to disaster package that considers gender, disability, orphaned and vulnerable status as well as rural/urban differences.
2015 The national non- formal education policy for Zimbabwe	The policy provides a second education chance for all children, youth and adults, who through different circumstances could have missed an opportunity to enroll in formal education.
MHTES&TD strategic plan 2019–23	Promote gender equality and equity, and gender mainstreaming in colleges and universities.
Secretary's circular No P36 of 1990	Special educational provisions in ordinary schools for children with varying degrees of disabilities and special education placement and procedures for special classes, resource units and special education schools.
Secretary's circular No 2 of 2000	Guidelines on the inclusion of learners with albinism
Directors' circular number 2 of 2001	Directs sign language to be taught in all primary schools in Zimbabwe
Directors' circular No 3 of 2001-	Guidelines on providing equal access to education for learners with disabilities
Director's circular minute no 24 of 2001	Provides for braille transcription and sign language interpretation for students with visual and hearing impairment during examinations
Directors' circular number 7 of 2005	Guidelines for the inclusion of learners with disabilities in all school situations
Directors' circular 2008	Policy guidelines of assessment of learners with disabilities and special needs
Director's circular number 5 of 2011	Provides guidelines for children who live more than 5 km from the nearest registered primary school and the children who live in scattered and isolated communities
Examinations circular number 21, 2014	Language modifications and examinations Administrations for grade 7 candidates with deafness

assistive technologies at schools since they were not allowed to carry them home. This made it difficult to access online lessons as most learners with disabilities come from poor families. It implies that, most learners with disabilities hardly benefitted from the systems put in place during lockdown (UNESCO, 2020c).

Practices, systems & structures for inclusive education

The Ministry of Primary and Secondary Education (MoPSE) has curricula framework (2015–22) that considers issues of equity, touching also on cross-cutting issues such as HIV and AIDS, sustainable resource utilization, climate change and disaster risk management. The Ministry of Higher and Tertiary Education, Science and Technology Development (MoHTESTD) has not yet taken tangible deliberate action to come up with a curriculum that addresses equity issues and also lacks a structure for inclusive education and education-in-emergencies. Within the MoPSE, inclusive education is the responsibility of the Learner Welfare Services, Schools Psychological Services and Special Needs Education Department. It focuses on enabling all learners to access the educational facilities on offer to achieve their individual best through a range of support services. The Zimbabwean legal framework allows, for all learners to be enrolled at schools nearest their homes irrespective of their diversity status (Chakuchichi et al., 2012; Government of Zimbabwe, 2013). Thus, all schools are expected to be inclusive and a considerable number of marginalized students and those with disabilities are enrolled in inclusive classrooms within regular schools. Resource units/Special classes within regular schools which accommodate all learners with disabilities numbered 1851 across all provinces as noted by the MoPSE 2018 s quarter disability statistics reports. Some special classes have various types of learners with a range of impairments, others have specific disabilities. Additionally, there are more such provisions in primary than in secondary schools.

Through a partnership between the Ministry of Primary and Secondary Education (MoPSE) and the Leonard Cheshire Disability Zimbabwe, a pilot project code-named: *The Cluster Model of Inclusive Education* came into effect in January 2010. The three-year

project was implemented at sixteen schools in three provinces: Harare, Mashonaland East and Mashonaland West. The project registered significant gains in terms of enrollment. Reports indicate that over one thousand children with disabilities who had not been in schools due to lack of support and a disabling school environment eventually made it into school (Chakuchichi et al., 2016; Chakuchichi and Chataika 2014). The project saw school infrastructures being adapted, learners were provided with mobility as well as teaching and learning assistive technologies as required. Also, teachers and communities were trained in inclusive education. After the success of the pilot project, the model was extended for a further four years.

An evaluation conducted in Mashonaland West indicated that promoting the inclusive primary education project was overwhelmingly reported as relevant by 159 out of 175 teachers (Chakuchichi et al., 2016). The project was meant to support the enrollment of 1500 children with disabilities in 30 model schools; and, at the same time, ensuring the retention of 741 such children enrolled in model schools prior to the commencement of the project. The project was thus expected to provide access to primary education to a total of 2241 children with disabilities in model schools. The project supported a total of 1089 children with disabilities in model schools and a further 1843 in cluster schools. Thus, 2932 children with disabilities were supported in accessing mainstream primary school education. In addition, a total of 951 teachers were capacitated in teaching children with disabilities within an inclusive setting through training workshops. Parents and children with disabilities were also provided with training to enhance their skills in self-advocacy. Teachers rated the impact of the project highly. At least 93.14% of teachers viewed the Leonard Cheshire Disability Zimbabwe's Promoting Inclusive Primary Education project as having positive impacts on the children with disabilities, their parents, schools, communities and the nation at large. Thus, the project was seen to have given rise to increased understanding of the rights of children with disabilities in the model schools.

At the local community level, the participation of women was enhanced to such an extent that they have taken leadership posts in School Development Committees. The participants of the study were quite confident that the project produced excellent value for money. Through transport solutions and assistive technology, it enabled most of the children with disabilities in the project areas to access education and be able to remain in school for more than 80% of the required time. The model-cluster school strategy comprises one central school in a cluster with about ten schools being adapted for disability inclusion. The adaptations include infrastructure, training of teachers and local leadership and the provision of assistive technologies to learners with disabilities. The school becomes inclusive and most learners with disabilities in the cluster of the adapted school are expected to attend the school. The increase of children with disabilities to accessing nearly all schools as the ripple effect influenced school heads and teachers who also participated in exchange visits. Such visits enabled neighboring schools to adopt inclusive education practices in their schools, thus making the project having a multiplier effect in its roll out.

In order to enhance access to an inclusive and quality education by all learners, including those with disabilities, UNICEF, through the Ministry of Primary and Secondary Education provided improvement grants to 3159 out of the 3200 schools targeted (98%) across the country. The grants have contributed to the improvement of the schools' physical environment and to the availability of teaching and learning materials. Income from the grants also enhanced the retention of poor children in school by subsidizing their overall costs of education. The number of children with disabilities enrolled into the general education system as a result of this project increased from 34,734 in 2015 to 49,692 in 2016.

MoPSE, with UNICEF support and funding from the Global Partnership for Education (GPE), supported the training of over 80,000 primary school teachers in 2018 in a supplementary module to the in-service Early Reading Initiative (ERI) and Performance Lag Address Program (PLAP). The supplementary module supports teachers with knowledge and skills on how to accommodate the diverse educational needs of learners to enhance their numeracy and literacy (UNICEF, 2018). This was done in recognition that teachers needed the know-how of planning and delivering meaningful lessons for the benefit of learners who have diverse educational needs in their classes.

As indicated earlier, Zimbabwe is also implementing the reverse inclusion model at special schools such as Emerald Hill School for the Deaf, Danhiko school and King George VI. The reverse inclusion model entails that special schools open their doors to learners without disabilities and these children are enrolled in the same class. The greatest achievement associated with this model is the change of attitudes by learners without disabilities toward their peers with disabilities. In line with the CRPD, this is a great step toward reduction of discrimination. The implementation of this model is being done without the support of the government. A five-year study of this model by Hlatywayo (2017) indicated that reverse inclusion has many benefits associated with it. However, educationists in authority seem to have little understanding of the model, which may be a stumbling block for its acceleration.

Pre-service and in-service teacher training

Pre-service training ensures that aspiring teachers acquire the prerequisite skills in inclusive teaching pedagogies. Similarly, in-service continuous professional development (CPD) also capacitates them in inclusive pedagogy. CPD in inclusive education in Zimbabwe is mainly supported by development partners in education. However, in service training is being conducted in several universities at undergraduate and postgraduate levels, and these include the University of Zimbabwe, Great Zimbabwe University and Zimbabwe Open University. The United College of Education trains teachers in inclusive education at diploma level.

Strengths and gaps on policy and practices in inclusive education

Inclusive education strengths

Zimbabwe has a comprehensive and robust legal framework for inclusive education. Policy documents use *person first* language for persons with disabilities which is in line with best standards. The 2013 Constitution has specific provisions for persons with disabilities which is not the case with many other African Countries. In addition, the nation is a signatory to many international and regional legal frameworks on inclusive education and education in emergencies. In addition, Inclusive Education Policy Drafts' Implementation Plan is very clear on **who** is supposed to do **what** and **when**, which makes it easy to implement. However, the policy is still in its draft form, and might take time before it is adopted.

Gaps in inclusive education

There is no inclusive curriculum framework for colleges and universities, with inclusive education seemingly part of primary and secondary education only. The MoHTESTD has serious policy gaps in inclusive education, and yet teachers who are trained under the ministry are expected to implement inclusive education. For example, all the policy documents mainly focus on primary and secondary education leaving higher education open for segregation. The Constitution of Zimbabwe (for example, Section 83) and some sections of the Education Amendment Act (for example, Section 62) seem to imply that inclusive education is a privilege not a right. Obligation statements are preceded by conditions or caveats such as, "provided there are resources". These are regressive considering the gains that have been made. It demonstrates that for the State, the funding of disability issues and other vulnerable student identities are shifted to become a liability to development partners. There is also a discord in the definition of inclusive education within policy documents with the recently amended Education Act remaining silent on the concept. Rather, it focuses on special needs education and yet there is supposedly a transition to inclusive education. Most importantly, the funding matrix for inclusive education is not clear.

While there are milestones covered by Zimbabwe, there remain several challenges facing inclusive education, including:

- lack of adequate resources;
- inaccessibility of schools;
- ambiguity of laws and policies;
- lack of political will;
- cultural stereotypes and negative societal attitudes;
- high teacher-pupil ratios;
- curriculum inaccessibility;
- limited research. (UNESCO, 2020; Sibanda, 2018).

Still, there remain opportunities given the efforts of the government of Zimbabwe, which has invested significantly in general education since 1980. These investments culminated in the construction of several mainstream schools and training of specialist and special education teachers which can be exploited for full inclusion of learners with disabilities (Sibanda, 2018). These opportunities have increased the potential for the successful implementation of inclusive education in Zimbabwe. Furthermore, the country has now developed a draft Inclusive Education Policy, which if adopted, could change the landscape of inclusive education in Zimbabwe.

Financing education in Zimbabwe

It is not possible to speak about education in the country without considering the issues relating to the costs of education provision. Kaplan and Lewis (2013:21) posit that, "Understanding education as a basic human right takes seriously the idea of education as a public good and education which is 'market driven' works against inclusive education". It is, therefore, the Government of Zimbabwe's responsibility to adequately fund the country's education system.

It is not unusual to observe that families in rural and poor urban areas often must choose between sending a child to school or having the child work to accrue revenue to help feed the family. Such families consider education to be too expensive a market commodity in a country where many workers have joined the informal labor sector. Consequently, between 5% and 40% of students drop out, finishing with low skills and high rates of unemployment. Among the drop - outs would be those many pupils with negative learning experiences and a history of having to repeat years because of poor performance.

In Zimbabwe, low education funding remains a major challenge. It is observed that, for example, while 23% of the total Government budget in 2016 was allocated to education, 98% of that money was spent on the wage bill leaving only 2% for capital development and other non-recurrent expenditures. In 2018, 93.7% of the Ministry's allocation also went to employment costs leaving only 6.3% for education development and other important costs such as procurement of teaching and learning materials, infrastructure development, capacity and manpower development and learner welfare support. In this regard, it has also been observed that non-recurrent expenditure of \$13 million a year, which equates to \$3.34 a year per child has resulted in families being called upon to shoulder most of the school facilities and supplies through payment of fees and levies in addition to meeting other direct and indirect costs including uniforms, transport, school feeding programs, and the procurement of teaching and learning materials. The

system thus, heavily depends on parental and community support which amounts to approximately 96% of the non-salary costs. Schools in disadvantaged communities are unable to raise sufficient funds to provide a quality education for all (Chireshe, 2017). The level of fees and levies collected at a school varies in accordance with the parents' ability to pay, determining the availability of infrastructure and learning materials. This has contributed to the perpetuation of existing inequalities. Despite the assistance rendered by parent communities, development partners such as UN agencies, sister ministries and civil society organisations, the country continues to be challenged to reverse the impact on access, retention, participation and achievement of quality outcomes, which have resulted from the historical context in the country and the economic crises experienced.

The central message emerging in this section is that; "Every learner matters and matters equally" (UNESCO, 2017, p.12). Thus, inclusive education seeks to increase access, presence, participation and success for all learners in education (Ainscow, 2016). This is the objective of SDG 4. Inclusive education is not a separate strategy to be used for educating a particular group; but rather a process and a goal that represents a particular quality or characteristic of Education For All. If Zimbabwe is to realize the benefits of leaving no one behind, the need to adopt inclusive education as a national strategy for providing education, cannot be overemphasized, taking into account issues of learners' welfare and wellbeing.

Conclusions

Marginalized children are more likely to have no access to education when compared to other children. If they go to school, they have lower rates of staying at school or promotion to the next grade. Concerning inclusive education, many countries in Africa benefit from foreign donors who are involved in micro inclusive education projects. There is no systemic change and formal national inclusive policy that results in system change at a country level. Given the socio-economic and related challenges, Africa is likely to benefit tremendously from a single inclusive system. This inclusive system could have enormous benefits for literacy and numeracy levels, throughput and access to teaching and learning. Malawi and Zimbabwe have positive strides in inclusive education that other African countries can learn from. However, several challenges have been identified that can impact negatively on inclusive education in Malawi and Zimbabwe. It is hoped that the following country-specific recommendations address the challenges raised in this article.

Country-specific recommendations

Based on the discussion in this chapter, the following country-specific recommendations are offered.

Malawi

- The MoEST should consider adequately supplying IEC materials for inclusive education (such as the NIES and its inclusive education curriculum guides) to districts and schools to build familiarity with inclusive education at all levels.
- The country should prioritize financing inclusive education, especially through the allocation of School Improvement Grants (SIGs), as per Priority 8 of the NIES (2017–21).
- The country, especially the MoEST and stakeholders, should consider scaling up sensitization efforts to change attitudes and behaviors at all levels regarding inclusive education.
- Malawi should mainstream inclusive education in all educational related policies, plans, legislations and frameworks.
- EMIS data should be inclusive (disaggregated by nature & extent of disability) and up to date (current-2016/17).
- MoEST and stakeholders should consider increasing the capacities for inclusivity of the school environments and learning institutions.

Zimbabwe

- Adopt the draft inclusive education policy (which has a strategic plan) in order to implement it nationally.
- Design a clear and comprehensive framework, including a funding mechanism for inclusive education and how it will be streamlined across all education sectors.
- Capacitate education stakeholders for inclusive education.
- Harmonize inclusive education across all the legal instruments.
- Establish diversity as a cross-cutting issue within the education system.
- Ensure EMIS is implemented in Non-Formal Education systems.
- MoHTESTD should design and implement an inclusive curriculum framework for higher education.
- MoPSE to collect all administrative circulars and instruments and make a booklet out of them for easy reference and guidance.
- Design a clear monitoring and evaluation framework for inclusive education.

References

- Ainscow, M., 2016. *Struggles for Equity in Education: The Selected Works of Mel Ainscow*. Routledge, Abingdon.
- Banks, L.M., Zuurmond, M., 2015. *Barriers and Enablers to Inclusion in Education for Children with Disabilities in Malawi*. Norwegian Association of Disabled, Oslo, Norway. https://www.lshtm.ac.uk/sites/default/files/2019-06/Barriers-enablers-inclusion-Malawi_0.pdf.
- Chakuchichi, D., Chataika, T., 2013. *Inclusive Education in Zimbabwe: Mashonland West Province*. Unpublished report. Leonard Cheshire Disability Zimbabwe Trust.
- Chakuchichi, D., Chataika, T., 2014. *Mid Term Review Report: 'Promoting Inclusive Education in Schools'*. Leonard Cheshire Disability Zimbabwe Trust, Harare.
- Chakuchichi, D., Nyaruwata, L., Chataika, T., 2012. *Leonard Cheshire Disability International Schools for All-Inclusive Education – Zimbabwe Evaluation Report*. Leonard Cheshire Zimbabwe Trust/Leonard Cheshire Disability International, Harare.
- Chakuchichi, D., Chataika, T., Nyaruwata, L., 2016. *'Promoting Inclusive Education in Schools' Project Endline Evaluation Report*. LCD, Harare.
- Chataika, T., Hlatywayo, L., Marufu, S., Maroveke, M., 2020. *Mapping Definitions, Policies and Practices on Inclusive Education and Education in Emergencies in Nine Member States Supported by UNESCO ROSA Countries*. UNESCO, Harare.
- Chireshe, R., 2017. *The State of Inclusive Education in Zimbabwe: Bachelor of Education (Special Needs Education) Students' Perceptions*. <https://doi.org/10.1080/09718923.2013.11893133>. Retrieved from.
- Chitiyo, M., Hughes, E.M., Chitiyo, G., Changara, D.M., Itimu-Phiri, A., Haihambo, C., Taukeni, S.G., Dzenga, C.G., 2019. *Exploring Teacher's Special and Inclusive Education Professional Development Needs in Malawi, Namibia and Zimbabwe*.
- GlZ and Government of Malawi, 2019. *Capacity Assessment of Inclusive Education (IE) in Primary Schools in Malawi*. March 2019.
- Government of Malawi, 2010. *The constitution of the Republic of Malawi (1995, as Amended Act No.11 of 2010)*. Retrieved from. <https://www.wipo.int/edocs/lexdocs/laws/en/mw/mw030en.pdf>.
- Government of Malawi, 2017. *National policy on early childhood development*. Retrieved from. <https://www.doccity.com/en/malawi-national-ecd-policy/5099960/>.
- Government of Malawi, 2020. *African Child Forum for Malawi: Vision 2020*. Retrieved from. http://www.africanchildforum.org/clr/policy%20per%20country/malawi/malawi_vision2020_en.pdf.
- Government of Zimbabwe, 2013. *The National Constitution of Zimbabwe*. Government Printers, Harare.
- Government Printers, 2013. *Malawi Government Education Act (Act No. 21 of 2013)*. Malawi, Lilongwe.
- Itimu-Phiri, A., Kopetz, P.B., 2008. *Malawi's special needs education (SNE): perspectives and comparisons of practice and progress*. J. Res. Spec. Educ. Needs 8 (3), 153–160.
- JICA Malawi, 2020. *Sector Position Paper: Education*. Retrieved from. <https://www.jica.go.jp/malawi/english/activities/c8h0vm00004bpzlh-att/education.pdf>.
- Kaphle, D., Marasini, S., Kalua, K., Reading, R., Naidoo, K.S., 2015. *Visual Profile of Students In Integrated Schools in Malawi*. <https://doi.org/10.1111/cxo.12269>.
- Kimura, H., 2005. *An Analysis of Information Management at School Level in Malawi: A Case of Lilongwe*.
- Lynch, P., Lund, P., 2011. *Education of children and young people with albinism in Malawi*.
- MoEST, 2018. *National School Health and Nutrition Strategic Plan (2018-2022)*. MoEST, Lilongwe.
- MoEST-Directorate of Inspection and Advisory Services, 2015. *National Education Standards-Primary and Secondary Education*. MoEST, Lilongwe.
- MoPSE, 2021. *Draft Inclusive Education Policy in Zimbabwe (unpublished)*, Ministry of Primary and Secondary Education. MoPSE, Harare.
- Research for Inclusive Education in International Cooperation, 2015. *Country Studies*. Malawi., Germany).
- Sibanda, 2018. *Impact of Cyclone Idai in Zimbabwe*. Department of Civil Protection, Harare; Zimbabwe.
- UNESCO, 2017. *A Guide for Ensuring Inclusion and Equity in Education*. UNESCO, Paris. Retrieved from. <http://unesdoc.unesco.org/images/0024/002482/248254e.pdf>.
- UNESCO, 2020. *Global Education Monitoring (GEM) Report 2020: Inclusion and Education-All Means All*.
- United Nations, 2006. *Convention on the Rights of Persons with Disabilities*. United Nations, New York.
- United Nations, 2015. *The Sustainable Development Goals*. United Nations, New York.
- WHO and World Bank, 2011. *World Report on Disability*. WHO/World Bank, Geneva. Retrieved from. http://www.who.int/disabilities/world_report/2011/en/index.html.

Crisis, austerity and the inclusive education agenda: the case of Greece

Evdoxia Nterpoulou-Nterou, School of Education, Department of Early Childhood Education, National and Kapodistrian University of Athens, Athens, Greece

© 2023 Elsevier Ltd. All rights reserved.

Introduction	418
The disability as a matter of rationalizing expenditure	419
Separate schools for different students	420
Seeking the free character of public education	422
Conclusion	423
References	424

Introduction

The central reflection of this chapter concerns the investigation of the effects of economic crisis and austerity measures on inclusive educational policy in Greece. At times of economic crisis such as Greece experienced over approximately the last 10 years, the population's impoverishment monopolizes public dialog. In the western world, the issue of poverty is perceived by terms of undifferentiated quantitative measures of material criteria, without taking into consideration the social and cultural differences, as well as the differences in values between societies, while most of the time it is associated with developing countries (Oppenheim, 1993; Alcock, 1993). Economic poverty in Greece cannot be analyzed by traditional terms of poverty evaluation, nor can be compared with the absolute poverty of developing countries. The international economic community often seems skeptical about the effects of austerity measures on people's lives in the European Union. The statement of the Director and Chairman of the International Monetary Fund, Christine Lagarde, is indicative. In her interview in *the Guardian* newspaper, on May 25th, 2012, Lagarde said this about poverty in Greece:

I think more of the little kids from a school in a little village in Niger who get teaching two hours a day, sharing one chair for three of them, and who are very keen to get an education. I have them in my mind all the time. Because I think they need even more help than the people in Athens.

Elliott and Aitkenhead (2012).

Skepticism regarding poverty within the European Union is reinforced by the media and political rhetoric as the blame for bankruptcy is attributed to the frivolous management of finances by the countries themselves. Arguments in favor of austerity programs are often constructed through the ethical mechanisms of narrative and the rhetoric of media and political discourse (Kelsey et al., 2016; Bickes et al., 2014).

We cannot examine austerity policies and their effects on education outside of the context of globalization and neoliberalism. Free market economy, commitment to laissez-faire of markets and commitment to free trade are some of the basic elements characterizing the relationship between neoliberal policy and education (Olssen and Peters, 2005). The exercise of neoliberal policies emerged in the western world at the beginning of 1980s. It was recognized as *Thatcherism* in Great Britain, *Reganomics* in the United States of America and *Rogernomics* in New Zealand (Dowse, 2009). However, it is relatively difficult to feature the effect of neoliberal history on education, as it appears asynchronously and in different forms during the last 30–40 years. Neoliberal policies and practices have specific consequences for education, as relations between government and private enterprises are redefined. Public sector and schools are exposed to increased competition, increased accountability measures and implementation of performance and management goals (Davies and Bansel, 2007).

On the other hand, changes in global economy, trade, communication and information diffusion, changes in the ways we deal with human resources and immigration are the basic characteristics that give meaning to globalization. Transition from the welfare state to the market economy, privatization of public bodies, personalization of provided services emerge as structural influences of the globalized neoliberal policy on the way countries are governed (Davies and Bansel, 2007). As a political ideology, it is applied to every aspect of governance, including education. These policies transfer the responsibilities of the welfare state to the individual, the family, the voluntary or non-governmental organizations and to communities' charitable actions (conservative communism) (Dowse, 2009).

In a newsletter about social inclusion, the European Commission declares that in 2015, 30.2% of disabled people (31 million) faced the risk of poverty or social exclusion in the European Union. This is a much higher percentage than the corresponding data for non-disabled people; 20.8% (European Commission, 2017). The influence of those particular policies was instantly revealed in Greece during the period of economic crisis—2009–2019—by dramatic income decreases, a leap of the unemployment rate, the downgrading of insurance and retirement provisions, shrinkage of the welfare state and under-funding of health and education systems.

At the beginning of March 2012, the Ministry of Education issued a circular regarding the support of students to deal with cases of malnutrition. The Ministry called School Principals, Teachers' Associations, in cooperation with Parents and Guardians' Associations to contact with municipalities' social services and Holy Metropolis in order to deal with nutrition problems (Ministry of Education et al., 2012).

The debt crisis and austerity policy in Greece reveals and crystallizes the real intentions of government from 2010 to 2011, with the "down-sizing" of welfare state and public infrastructure through the announcement of shrinkage of 151 Public Enterprises and Entities. Thirty-one associations working with and on behalf of people with disabilities were included in the list of targets (Ministry of Finance, Special Secretariat for Public Enterprises and Entities, 2011).

From 2009 to 2012, total expenditure for education dropped by 16% according to the estimates of the General Secretary of the Federation of Secondary Education Officers (Education Department, 2011). The shrinkage of the Education budget precipitated the merging and abolition of schools. In total, 1933 schools merged, representing 1056 school closures. 851 primary schools and 205 secondary schools were closed (Ministerial Decision, 2011). Moreover, in 2016, the percentage of people with severe or mild disability (aged from 16 to 64 years old) that were at risk of poverty or social exclusion in Greece amounted to 51.9%, surpassing the average index value of 27 countries of the European Union, that amounts to 37.4% (Development Center for Education Policy of General Confederation of Greek Workers, 2018).

According to data for the year 2016, Greece is included among state-members of the European Union with the largest inequalities regarding access to education for disabled people over 16 years of age. Their participation rate in education is significantly reduced as the level of education increases. Corresponding research for Europe regarding the education of disabled people of preschool and school age does not exist. Respectively, in the Greek Statistical Authority, data about education and particularly data about compulsory education of disabled students are incomplete.

According to the last statistical data of Eurostat for the year 2017, the total population at risk of poverty or social exclusion in Greece amounts to 34.8%, occupying third place among 28 countries of the Eurozone (Eurostat, 2018a). Additionally, according to the last data of the Greek Statistical Authority for the year 2017, the percentage of people experiencing material deprivation at ages 0–17 years old amounts to 38.3% (Greek Statistical Authority, 2019).

The political and economic context of Greece has profound implications for inclusive schooling. The current study presents the impact of austerity measures on inclusive education in three sections:

1. An examination of how education policy, under the pressure of austerity measures, addresses disability.
2. An analysis of the impact of austerity on education policy decisions regarding the further segregation of students and schools.
3. An examination of the shrinking of free public education in the context of the economic crisis and austerity measures.

The disability as a matter of rationalizing expenditure

Inclusion as political act is revealed as a struggle against power abuse, as identified in inequalities maintained by the hierarchical structure of the education system. Educational inequalities are exacerbated in periods of economic crisis by significant cuts in expenditure.

Recognizing and responding positively to student differences in schools is considered a major cause of waste that must be reduced. The political discourse separates education from the values of community, humanism, acceptance of difference, and critical approaches to pedagogy in the name of efficiency, serving the wider imperative of austerity policies (Giroux, 2010). Almost from the beginning of the economic crisis, disabled students became objects for economic management. The austerity measures were imposed as an antidote to the so-called irrational financial management of previous years (Runswick-Cole and Goodley, 2015). This position is reflected in official discourse. Disability was treated as an expense to be avoided. The economic crisis privileged the medical model of disability applied to schooling through neoliberal economic rhetoric. The Special Secretary of Primary and Secondary Education—Ministry of Education held a meeting about Attention Deficit and Hyperactivity Disorder (ADHD) on April 2012 in Athens. His speech is an example of "cruel optimism" expressing his belief in the necessity of austerity policies (Runswick-Cole and Goodley, 2015). The Special Secretary said:

... Our goal is the interdisciplinary diagnostic teams ... to be present in schools, in order to act directly and avoid every kind of expenses: expenses regarding time or economical, mental, learning

Esos.gr (2012).

In the context of the rationalization of expenditure on education during the period 2014–2018, two different structures of diagnostic, educational evaluation and support for students with disabilities were established. Their purpose was to ensure equal access of all students to education and to protect the psychosocial development and progress of the child and to enhance interdisciplinary collaboration in the general school (Law, 2018; Ministerial Decision, 2014; Law, 2012). The aim of the legislators was more to harmonize with the international conventions on the rights of the disabled students in education (Law, 2012) than to commit to the implementation of this ambitious project. The above finding is presumed by the fact of the lack of funding for education.

According to Eurostat data, the total expenditure of the General Government for education as a percentage of GDP amounts to 4.5% for the year 2013 and falls to 3.9% for the year 2017 (Greek Statistical Authority, 2019). In addition, during the year 2017–2018, 79,934 students with disabilities or special educational needs attended general schools. 57.3% of students with disabilities or special educational needs in general schools did not receive specialized support. The number of students receiving specialized support is dramatically reduced during the transition from primary to secondary education, where more than 90% of students were supported only by teachers of general classes without additional support or adjustments. It is worth noting that at a rate of 1.70% the coverage of the educational support expenses of the disabled students was borne by the students' families themselves (Disability Observatory of the National Confederation of Persons with Disabilities, 2019).

In the last decade, diagnostic and support structures have been used as mechanisms to identify student differences and not as structures to support and promote inclusive education for all students.

The meaning attached to the rationalization of spending on education in times of austerity usually disguises spending cuts. The disabled student is a costly student whose educational needs are placed in the context of immediate cuts in times of financial crisis. The cost of spending is passed on to the family, objectifying this shift of responsibility from the state to parents. A kind of ideological austerity is formed in addition to the economic austerity within the educational system, creating a school without choices for disabled students and their families. This ideological austerity characterizes the decisions of the educational policy, from the beginning of the crisis until today (Nteropoulou and Slee, 2019), without exceptions in the political orientations of different government formations in accordance with the policy decisions taken on macroeconomic management (Hayes, 2017).

Separate schools for different students

Existing legislation in Greece commits to secure for all citizens with disabilities and certified special educational needs equal opportunities for full participation and contribution to society, independent living, economic self-sufficiency and autonomy, full consolidation of their rights for education and social and vocational inclusion (Law, 2008, article 1). This commitment results in expectations for change of educational policy, as it is self-defeating in three basic ways:

1. The wording of the purpose and pursuits of the education of disabled students,
2. The diagnostic process regarding disabilities and special educational needs, and
3. The preconditions for students' attending general education.

Specifically, the wording of the purpose and pursuits of legislation makes it obvious that the success of educational, social and vocational inclusion is closely related to the individual abilities of students and impairments of functionality (Law, 2008, article 2, paragraph 5). Accordingly, the goals of education are achieved through the timely medical diagnosis and evaluation of special educational needs by official diagnostic bodies (Law, 2008, article 2, paragraph 6; Law, 2018, article 7, paragraph 2). The categorization of the student population, which is directly related to clinical perception of disability, individualizes the practices of treatment of disabled as deficient, perpetuates social discrimination and suspends the dynamics of a political, collective organization and representation of disabled people (Siebers, 2008). The continuous discrimination of students is not an act of acceptance of difference or otherness, but a practice of controlling groups of within the student population that seem to "disturb" the *normality* of general education. For example, the legislative framework gives the opportunity to students to remain and attend a regular neighborhood classroom only when they have mild learning difficulty (Law, 2008, article 6). Students may be supported by a special education teacher according to the kind and degree of their special educational needs. However, when they have more severe special educational needs, they must attend separate special classrooms, paradoxically called "inclusion classrooms". Students who face difficulties in attending inclusion classrooms are excluded from general education and attend separate special schools (Law, 2008, article 6).

The arbitrary nature of diagnostic practice regarding the attribution of the label of "special educational need" defines the educational life of student population cohorts, particularly of those students growing up in adverse social conditions. According to the legislative framework, students with social difficulties, delinquent behavior due to abuse, parental neglect and abandonment, or due to domestic violence belong to the group of students with special educational needs (Law, 2008, article 3). This results in the registration of those students in special schools without the justification of a report that certifies their mental, sensory or other impairment, but with the single excuse of their burdensome social identity.

Education policy has appropriated and weakened the subversive character of inclusive education discourse, by enhancing, through institutional discourse and institutional regulations, the conviction that inclusive education is a modern direction of special education (Slee, 2004) and a tool for the assimilation of students with cultural, national and religious differences as well as refugees. Inclusion is perceived as a worrying and provocative activity, when its political discourse approaches critically the system of social values, structures and institutions (Barton, 2008). Inclusive discourse as a fundamental struggle for changing the educational system ought to lead us in cultural alertness, aiming at the transformation and restoration of every kind of educational exclusion (Slee, 2004, 2011, 2012, 2018). The change is not about assimilation, but transformation of structural barriers of change in the academic community and schools (Whitty, 2002; Gillborn and Youdell, 2000).

The identification of inclusive education with special education is revealed in educational techniques regarding the management of difference. These techniques focus on deficit, diagnosis, categorization and individual treatment with the purpose of regulating the devalued differences. It is, also, evident in the formation of attendance structures (general classroom, special classroom, parallel

support, special school), in the discrimination of educational specialties and specializations (general education teachers, special education teachers), in the discrimination of curriculum, in the registration procedures in schools and in many other areas (Nteropoulou-Nterou, 2012; Nteropoulou-Nterou and Slee, 2019).

Shortly before Greece entered the International Monetary Fund (IMF), the action plan of the new elected government about changes in primary and secondary education was publicly available. Entitled, "The new school: the student comes first," the plan is committed to future reforming initiatives from preschool to secondary education (Ministry of Education, Lifelong Learning and Religious Affairs, 2009). In the second section of the text, a reference is made to inclusion and special education. Examination makes clear that in the section of programmatic commitments about inclusion, disabled students are not included, but only students from the Muslim minority of Thrace, foreigners, repatriated students and Roma. The educational issues of disabled students are included only in the section of special education (Ministry of Education, Lifelong Learning and Religious Affairs, 2009). Inclusion appears as a programmatic objective that does not concern all students, but a particular group of students with cultural differences.

In the period of economic crisis, Greece also faced the refugee issue. According to Oliver (2004), war and poverty are the two main causes of impairments. War brings poverty, injuries and refugees. Many refugees' and disabled students' experiences are characterized by life-threatening psychological and material deprivation. At the same time, the dominant value cultural position of the general school treats them as the "other," as the "intruder" (Barnes et al., 2002). Xenophobic perceptions and dominant social values form identities. The way we perceive ourselves and the way we feel when we are judged by others is of nodal importance. It forms the degree of self-esteem and empowerment, properties that are usually perceived as issues of personal choice, overlooking the degree of influence of the hegemonic social values on personal experience (Galvin, 2006). It is worth noting that the total number of refugee students deserving international protection that attend Greek public schools is 12,867 students. More specifically according to the Ministry of Education, during the school year 2018–2019, 4050 refugee students attended reception classrooms in schools of primary and secondary education, 4240 students attended schools without reception classrooms, and 4577 students attended structures for the educational support of refugees (Ministry of Education, Lifelong Learning and Religious Affairs, 2019a).

On May 30th, 2019, at the presentation of the Annual Report of UNESCO about immigrants' education, the Undersecretary of Education referred to the state's actions for the combat of discrimination in Education and the progressive inclusion of refugees in the Greek educational system. She characteristically said:

... we formed the action plan for the smooth inclusion of refugee children in the educational system, by creating Structures for the Reception and Education of Refugees ... staffed all the Hosting Centers of Greece ... It is our obligation to invest in education as it is the place that owes and can reduce social inequalities, combat prejudice and racist perceptions and defend all children's rights, without any kind of racial, national, religious, gender, social or economic discrimination.

Ministry of Education, Lifelong Learning and Religious Affairs (2019b).

According to the Minister, inclusion is perceived and promoted as an integration tool through the establishment of special structures that function either as separate schools within the refugee camps or as special separate classrooms called "reception classrooms" in general schools. The inclusive approach of education, in political discourse, is approached by idealistic terms and terms of self-complacency, downgrading the importance of social barriers and inequalities that are systematically produced and reproduced by and within the education system (Roulstone and Prideaux, 2008).

An example of downgrading the importance of social barriers is the recent angry reaction of a Parents' and Guardians' Association in an island province of 33,000 inhabitants. The Parents' Association sued a teacher of parallel support in a primary school, because she defended the right of eleven refugee students to attend classrooms of educational support in their children's school. Parents asked for compensation of 50,000 euros from the teacher for calling them "racists" and "insulting their honor and reputation." It is worth noting that in the area of the northern Aegean, the rate of unemployment for 2018 was 22.3% (Greek Statistical Authority, 2019).

Parents argued that the schooling of refugee children cannot be provided against Greek students' rights and against islanders' public health. It is worth noting that Greek students attending all-day school and refugee students had their break in different yards, while there were also separate toilets and desks. In this xenophobic climate, the teacher stated in her interview on an educational portal on June 27th, 2019:

Education sector has to shield against arbitrary and illegal practices of members of Parents and Guardians' Associations that desire with humble motives, that violate the internationally guaranteed children's rights ... Teachers should not be afraid or terrorized. This is my message to all my colleagues ...

Vainanidi (2019).

It is a fact that many Parents' Associations from all over Greece as well as many other educational bodies supported the teacher for her attitude. The Ministry of Education also condemned Parents' and Guardians' Association's suit against the teacher that defended refugees' schooling, pointing out in a press release on June 17th, 2019 that: "... the mission of teachers is the education of students based on values of solidarity and collectivity" (Ministry of Education, Research and Religious Affairs, 2019c).

However, the point is not only about an official condemnation of a xenophobic event. Inclusive education is not a matter of dealing individually with social barriers. Teachers' exposure to the anger of social groups that, in the years of crisis, were struggling materially is an extremely dangerous and lonely experience.

In the context of neoliberal political commitments, economic poverty and competitive individualism, the transfer of responsibility for inclusive education to the individual conscientiousness of teachers tests the strength of the educational community. In an environment that appears hostile to pedagogical values, the danger of teachers' frustration looms and "collective indifference" through to hostility is exacerbated in weakened communities (Bousquet, 2008; Slee, 2012). The responsibility for cultural vigilance should not be taken only by teachers, but also by the state. However, the role of "cultural advocate" is annoying for those who approach inclusive education as a technical issue and deal with it through bureaucratic assimilationist decisions. The transformation of inclusion into conditional inclusion (Corbett and Slee, 2000) results in the transfer of educational policy's responsibilities to the low-paid teachers.

Seeking the free character of public education

Until today compulsory education in Greece is public and free (The Constitution of Greece, 2010, article 16, paragraph 4). However, the problem of a "shadow education" system, namely private tutorials, which existed before economic crisis has been extended. Never before were private teachers present in public schools, particularly with the responsibility to support disabled students in general classrooms.

In 2013, within the economic crisis, private parallel support is enacted by the state, within Law 4186. The disabled student can be supported in the general classroom by a special teacher appointed and paid for by the student's family (Law, 2013). The state seems to transfer the responsibility of the student's educational support and inclusion in general classrooms to the family. The student's family avoids registration and attendance at a special school, by hiring a private teacher. Disabled students' families take up the role of employer and, according to financial capacity, hires supportive personnel in the neoliberal context of public and free education. It seems that the inclusive school is more hospitable for students whose families can afford this expense. This fact further enhanced discrimination between families as well as between teachers. Discrimination between teachers results from different employment contracts of educational personnel. The teacher who is hired by the family to support the disabled student is perceived as a "foreign body" by the educational community, a fact that leads to unequal educational relations and difficulties in sustaining cooperation between teachers.

It is well known that education is the vehicle for dealing with poverty and social exclusion and for securing welfare (Barnes and Sheldon, 2010). However, currently, this general principle seems to be shaken, if we consider the last data of Eurostat about graduates of tertiary education in the European Union. Specifically, 10.9% of the EU population with a tertiary education were at risk of poverty or social exclusion. This ranged from the lowest at 3.2% in Malta to the highest at 19.7% in Greece (Eurostat, 2018b).

Although the spigot of placements in the public sector has narrowed due to the reduction of budget expenditure in Greece, the sector of public education still offers teachers some working opportunities with fixed-term contracts, in order to cover identified needs of schools. The existing jobs that seem to still absorb a number of teachers in public education are those concerned with supporting disabled students in general education. These jobs are predominantly funded by the European Community Fund (Ministry of Education, Lifelong Learning and Religious Affairs, Press Office, 2013).

According to the data of the Ministry of Education, during 2014–2015, 6308 teachers were hired with fixed-term contracts, 8623 in 2015–2016 and 10,700 teachers were hired in 2016–2017 (Ministry of Education and Religious Affairs, 2017). Additionally, the Ministry of Education committed to, until the autumn of 2019, proceed with 4500 permanent placements in Special Education for the school year 2019–2020 (Ministry of Education and Religious Affairs, 2019a,b,c,d). Finally, 3445 teachers were appointed in August 2020. These are the first mass appointments of special education teachers in the last 10 years (Ministry of Education and Religious Affairs, 2020).

The typical qualification that teachers must have in order to be placed in those jobs is, among others, the postgraduate degree in special education (Law, 2008, article 20, paragraph d).

In the context of a desperate search for a job in public education, we note an intense participation of teachers in offered postgraduate and training programs. A new market is being formed within Higher Education. Attending postgraduate and training programs results in raising hope for employment for an increased number of teachers, as they are given the opportunity to work in educational structures and mainly in inclusive educational structures for disabled students. Universities, after the expenditure cutting for education, are concerned with the decrease or restraint of cost of the services they offer in order to secure their survival. It is worth noting that, according to the data of Eurostat, republished by Greek Statistical Authority (2019), total expenses of General Government for education as a percentage of GDP amounted to 4.5% for 2013 and ended up at 3.9% for 2017. Fourth in a row from the bottom of the countries of Eurozone. In the last three positions are Ireland with 3.3%, then Italy and Slovakia, both with 3.8% with respect to the total expenses of the General Government for education as a percentage of GDP (Greek Statistical Authority, 2019).

To be able to meet the requirements of the new economic situation, universities are oriented toward specialization, by means of postgraduate programs, and toward groups of students that can afford the tuition fees. In this way, the cost of postgraduate studies is transferred to teachers who, in their desperate effort to find a job, must cover the expense. On the other hand, Universities enter into

to creating the market conditions and business practices in the so-called public and free Higher Education Institutions ([The Greek Constitution, 2010](#)).

In 2013, postgraduate programs in special education were limited in number and demanded tuition fees ([National and Kapodistrian University of Athens et al., 2013](#)). The number of postgraduate programs in special education offered by Higher Education Institutions across Greece grows rapidly; 53 postgraduate programs for the academic year 2018–2019 ([Ministerial Decision, 2019a,b,c,d,e](#)). The demand for attending postgraduate programs is extremely high. The postgraduate program in Special Education of the Department of Early Childhood Education at the National and Kapodistrian University of Athens is a typical example. During the decade 2009–2019 the applications for the Postgraduate Program were 2849 for a total number of 200 positions ([Postgraduate Studies Program in Special Education, 2019](#)). According to the data of [Greek Statistical Authority \(2019\)](#), the total number of postgraduate students for the period 2012–2013 was 33,687, while for the academic year 2016–2017 they doubled and amounted to 66,295 postgraduate students.

In the period of economic crisis and after the memorandum, the increase of placements of teachers in inclusive and special education structures seems to upgrade the interest of educational community for inclusive education. Moreover, inclusive education, which is considered as the modern direction of special education, is projected by the representatives of educational policy as a high priority objective for the government. Characteristically, on the occasion of teachers' appointments, the Minister of Education stated:

The right to equal access for all children, including disabled children, is a high priority goal for the government. The upgrade of Special Education is distinct and unquestionable. The effort to improve the educational conditions for all students, without exceptions, in a public school that promotes solidarity, cohesion and acceptance of the difference continues with unwavering determination.

Ministry of Education and Religion Affairs (2018).

The use of the word "difference" is perceived as euphemism for "*not normal*". Based on the perspective of consensualism, possibly conflicting properties such as social class, race or disability, are not perceived as important forces that can influence the course of an educational system. These factors are defined as individual properties rather than structural mechanisms supporting the production and reproduction of unequal relations in education ([McDonnell, 2012](#)).

Throughout the last decade, inclusive education in Greece operates as a life boat for the employment of teachers, for the market development in the context of public Universities and for the political scene of the country as it uses inclusive education as the frontpiece of social policy ([Nteropoulou-Nterou and Slee, 2019](#)).

Conclusion

In times of austerity and financial crisis, it was observed that the segregation of students based on their diversity increased, diagnostic and supportive institutions are used as mechanisms for identifying difference and not as structures to support and promote inclusive education. It has also been observed that confusion is perpetuated about the conceptual appropriation of inclusive education by special education. There has been a tendency to reinforce segregation policy in teacher education in the name of the teacher's specialization to support the so-called inclusion of disabled students. The individual responsibility of teachers in addressing the problems arising from the divisive education policy, the individual initiative of families to provide educational support to their children in the context of public education remain modern examples in which the proclamatory political discourse for inclusion seems to be stripped baer.

Inclusion, as conditional inclusion or neo-special education, is empowered in the context of economic crisis and austerity measures and presents itself as a realistic, directly applicable educational practice. Still, the education policy of the country produces exclusionary education systems and does not seem committed to the vision of an inclusive education inspired by the principle of cultural vigilance, which can constitute the vehicle of a collective demand for social justice. It creates a climate of individual resolution of differences in those involved in education, without focusing on the full development and empowerment of human potential, the development of creativity and the effective participation in the educational happenings and society ([Byrne, 2013](#)). Inclusive education has been reduced to another bureaucratic discussion about monitoring and dealing with students in the education system.

In the context of the imposed Structural Adjustment Programs and austerity policies, the legislative framework is indeed reproducing the aims of discriminatory special education in the preconditions of attendance, the offered educational services and the perception of students' difference, shaping an inclusion under conditions ([Roulstone and Prideaux, 2008](#)). In times of austerity the gap between the political discourse on inclusion and the educational reality is evident in the context of implementing continuous practices of separating different groups within the education system ([Tomlinson, 2005](#)). The cloudy and confused vision about inclusive educational policy that existed in Greece before the debt crisis period, has transformed into a management problem of expense rationalization in the after-debt crisis period. In the context of ideologically and economically weakened periods, we need to reconsider ways that will allow us to free the right to attend general school from the biological, social and cultural diversity of the students and means that will enable us to stop further education segregation policies.

References

- Alcock, P., 1993. *Understanding Poverty*. Macmillan, London.
- Barnes, C., Sheldon, A., 2010. Disability, politics and poverty in a majority world context. *Disabil. Soc.* 25 (7), 771–782. <https://doi.org/10.1080/09687599.2010.520889>.
- Barnes, C., Oliver, M., Barton, L., 2002. Disability, the academy and the inclusive society. In: Barnes, C., Oliver, M., Barton, L. (Eds.), *Disability Studies Today*. Polity Press, Cambridge, pp. 250–260.
- Barton, L., 2008. Education and the struggle for change: the need for a policy review of hope. In: Fiata, H. (Ed.), *Drop in for a Coffee. Family Relations and School on the Edge of Difference*. Taxideutis, Athens, pp. 21–38 [in Greek].
- Bickes, H., Otten, T., Weymann, L.C., 2014. The financial crisis in the German and English press: metaphorical structures in the media coverage on Greece, Spain and Italy. *Discourse Soc.* 25 (4), 424–445. <https://doi.org/10.1177/0957926514536956>.
- Bousquet, M., 2008. *How the University Works: Higher Education and the Low—Wage Nation*. New York University Press, New York.
- Byrne, B., 2013. Hidden contradictions and conditionality: conceptualizations of inclusive education in international human rights law. *Disabil. Soc.* 28 (2), 232–244. <https://doi.org/10.1080/09687599.2012.699282>.
- Corbett, J., Slee, R., 2000. An international conversation on inclusive education. In: Armstrong, F., Armstrong, D., Barton, L. (Eds.), *Inclusive Education. Policy, Contexts and Comparative Perspectives*. David Fulton Publishers, London, pp. 133–146.
- Davies, B., Bansel, P., 2007. Neoliberalism and education. *Int. J. Qual. Stud. Educ.* 20 (3), 247–259. <https://doi.org/10.1080/09518390701281751>.
- Development Center for Education Policy of General Confederation of Greek Workers, 2018. *Annual Report on Education 2017–2018*. In: *The Basic Dimensions of Education*. KANEP/GSEE, Athens [in Greek].
- Disability Observatory of the National Confederation of Persons With Disabilities, 2019. 5th Statistical Information Sheet: Information on the Education of Students With Disability and Special Educational Needs. National Confederation of Disabled People, Athens. Retrieved 2020-09-16, from: <https://www.esamea.gr/publications/others/4312-5o-deltio-statistikis-plirotforisis-stoixeia-gia-tin-ekpaideysi-ton-mathiton-me-anapiria-i-kai-eidikes-ekpaideytikes-anagkes> [in Greek].
- Dowse, L., 2009. “Some people are never going to be able to do that”. Challenges for people with Intellectual Disability in the 21st century. *Disabil. Soc.* 24 (5), 571–584. <https://doi.org/10.1080/09687590903010933>.
- Education Department, 2011. *Education Budget Note for 2012. Coalition of the Left, of Movements and Ecology Party*. Retrieved 2020-09-10, from: http://www.syn.gr/downloads/proyp2012/paideia_proyp2012.pdf [in Greek].
- Elliott, L., Aitkenhead, D., 2012. It's Payback Time: Don't Expect Sympathy—Lagarde to Greeks, Take Responsibility and Stop Trying to Avoid Taxes, International Monetary Fund Chief Tells Athens. *The Guardian*. May 25, 2012, Eurozone section, UK edition. Retrieved 2020-09-10, from: <https://www.theguardian.com/world/2012/may/25/payback-time-lagarde-greeks>.
- Esos.gr, 2012. The ministry of education regarding the support of students with ADHD. In: *Esos.gr (Daily Portal Briefing on Education)*. April 2, 2012. Retrieved 2020-09-10, from: <https://www.esos.gr/arhra/defterovathmia-ekpaidefsi/eidisis-defterovathmia-ekpaidefsi/to-ypoyrgeio-paideias-gia-thn-yposthrijh-tvn-mauhtvn-me-diataraxh-elleimmatikh-prosoxh-yperkinhtikothta> [in Greek].
- European Commission, 2017. *Social Inclusion. Thematic Factsheet*. EU. Retrieved 2020-09-10, from: https://ec.europa.eu/info/sites/info/files/file_import/european-semester_thematic-factsheet_social_inclusion_en_0.pdf.
- Eurostat, 2018a. *International Day for the Eradication of Poverty*. EU. Retrieved 2020-09-10, from: <https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/poverty-day-2018>.
- Eurostat, 2018b. *People at Risk of Poverty or Social Exclusion by Educational Attainment Level*. EU. Retrieved 2020-09-10, from: https://ec.europa.eu/eurostat/web/products-datasets/product?code=ilc_peps04.
- Galvin, R., 2006. A genealogy of the disabled identity in relation to work and sexuality. *Disabil. Soc.* 21 (5), 499–512. <https://doi.org/10.1080/09687590600785969>.
- Gillborn, D., Youdell, D., 2000. *Rationing Education: Policy, Practice, Reform and Equity*. Open University Press, Buckingham.
- Giroux, H., 2010. Bare pedagogy and the scourge of neoliberalism: rethinking higher education as a democratic public sphere. *Educ. Forum* 74 (3), 184–196. <https://doi.org/10.1080/00131725.2010.483897>.
- Greek Statistical Authority, 2019. *Greece in Figures. January–March 2019*. ELSTAT, Athens [in Greek].
- Hayes, G., 2017. Regimes of austerity. *Soc. Mov. Stud.* 16 (1), 21–35. <https://doi.org/10.1080/14742837.2016.1252669>.
- Kelsey, D., Mueller, F., Whittle, A., KhosraviNik, M., 2016. Financial crisis and austerity: interdisciplinary concerns in critical discourse studies. *Crit. Discourse Stud.* 13 (1), 1–19. <https://doi.org/10.1080/17405904.2015.1074600>.
- 3699 Law, 2008. *Special Education of individuals with disabilities or special educational needs*. *Gov. Gazette* 1 (199), 3499–3522.
- 4074 Law, 2012. *Ratification of the convention on the rights of individuals with disabilities and the optional protocol in convention on the rights of persons with disabilities*. *Gov. Gazette* 1 (88), 2591–2690.
- 4186 Law, 2013. *Restructuring secondary education and other provisions*. *Gov. Gazette* 1 (193), 3111–3162.
- 4547 Law, 2018. *Reorganization support structures of primary and secondary education and other provisions*. *Gov. Gazette* 1 (102), 8379–8454.
- McDonnell, P., 2012. Deep structures in special needs education. In: Zoniou-Sideri, A., Nteropoulou-Nterou, E., Vlachou-Balafouti, A. (Eds.), *Disability and Education Policy: Critical Approach to Special and Inclusive Education*. Pedio, Athens, pp. 93–121 [in Greek].
- Ministerial Decision, 2011. *Decisions: mergers and repeals secondary school units, demote, merger and abolition of primary schools and kindergartens*. *Gov. Gazette* 2 (440), 6473–6547.
- Ministerial Decision, 2014. *Legal recommendation of diagnostic educational assessment and support committees and definition of the specific duties of members and coordinators of them*. *Gov. Gazette* 2 (315), 4113–4128.
- Ministerial Decision, 2019a. *Finding the relevance of postgraduate programs in Greece and abroad to the subjects of special education and school psychology*. *Government Gazette* 2 (1217), 14779–14783.
- Ministerial Decision, 2019b. *Amendment of the ministerial decision of finding the relevance of postgraduate programs in Greece and abroad to the subjects of special education and school psychology*. *Gov. Gazette* 2 (1455), 17376–17370.
- Ministerial Decision, 2019c. *Second amendment of the ministerial decision of finding the relevance of postgraduate programs in Greece and abroad to the subjects of special education and school psychology*. *Gov. Gazette* 2 (1809), 21079–21082.
- Ministerial Decision, 2019d. *Third amendment of the ministerial decision of finding the relevance of postgraduate programs in Greece and abroad to the subjects of special education and school psychology*. *Gov. Gazette* 2 (1869), 21539–21541.
- Ministerial Decision, 2019e. *Amendment of the ministerial decision of finding the relevance of postgraduate programs in Greece and abroad to the subjects of special education and school psychology*. *Gov. Gazette* 2 (2743), 31125–31127.
- Ministry of Education and Religious Affairs, Culture and Sport, Press Office, 2013. *Recruitment Deputies Primary Education Teachers for the Academic Year 2012–2013*. Retrieved 2020-09-10, from: <https://www.minedu.gov.gr/grafeio-typoy-kai-dimosion-sxeion/deltia-typoy/9336-13-02-13-%2064-54-252-2012-2013> [in Greek].
- Ministry of Education and Religious Affairs, 2017. *Two Ministerial Decisions—A Landmark in Special Education*. Retrieved 2020-09-10, from: <https://www.minedu.gov.gr/eidiki-agwgi-2/nomothesia-eidiki-ekpaideusi-3/28799-20-06-17-dyo-ypourgikes-apofaseis-stathmos-stin-eidiki-agogi-kai-ekpaidefsi-3> [in Greek].
- Ministry of Education and Religious Affairs, 2018. *Historic Leap With 7505 Recruits in Parallel Support*. Retrieved 2020-09-10, from: <https://www.minedu.gov.gr/rss/38300-20-11-18-istorikoalma-me-7505-proslipseis-stin-parallili-stiriksi-2> [in Greek].

- Ministry of Education and Religious Affairs, 2020. For the First Time Permanent Appointments in Primary and Secondary Special Education Schools. Retrieved 2020-09-22, from <https://www.minedu.gov.gr/news/45927-05-08-20-gia-proti-fora-monimoi-diorismoi-se-soxleia-protovathmias-kai-defterovathmias-eidikis-agogis-kai-ekpaidefsis-telikoi-pinakes-2> [in Greek].
- Ministry of Education, Lifelong Learning and Religious Affairs, Uniform Administrative Sector Primary and Secondary Education, 2012. Enhancing Students. Newsletter. Retrieved 2020-09-10, from <https://www.minedu.gov.gr/grafeio-typoy-kai-dimosion-sxeseon/deltia-typoy/8416-27-03-12-panelladiki-drasi-synergias-politeias-ekklisias-gia-tin-enisxysi-ton-mathiton-kai-ton-oikogeneion-toys> [in Greek].
- Ministry of Education, Lifelong Learning and Religious Affairs, 2009. The New School: The Student First. Retrieved 2020-09-10, from https://www.minedu.gov.gr/index.php?option=com_content&view=article&id=402&Itemid=785&show_all=1&lang=el [in Greek].
- Ministry of Education, Research and Religious Affairs, 2019a. The Ministry of Education for World Refugee Day. Retrieved 2020-09-10, from <https://www.minedu.gov.gr/rss/41865-20-06-19-to-yppeith-gia-tin-pagkosmia-imera-proslygon-2> [in Greek].
- Ministry of Education, Research and Religious Affairs, 2019b. Presentation of the Annual UNESCO Report on Education of Migrants. Retrieved 2020-09-10, from <https://www.minedu.gov.gr/rss/41497-31-05-19-i-yf-paideias-m-tzoyfi-stin-parousiasi-tis-etisias-ekthesis-tis-unesco-gia-tin-ekpaidefsi-ton-metanaston> [in Greek].
- Ministry of Education, Research and Religious Affairs, 2019c. The Ministry Condemns the Lawsuit Against Teacher. Retrieved 2020-09-10, from <https://www.minedu.gov.gr/rss/41802-17-06-19-to-ypourgeio-katadikazei-tin-agogi-se-varos-ekpaideftikoy-2> [in Greek].
- Ministry of Education, Research and Religious Affairs, 2019d. Request From the Ministry to the County Council for the 15,000 Appointments to Education. Retrieved 2020-09-10, from <https://www.minedu.gov.gr/rss/41754-13-06-19-aitima-tou-ypourgeiou-pros-to-ste-gia-tous-15-000-diorismoys-stin-ekpaidefsi-2> [in Greek].
- Ministry of Finance, Special Secretariat for Public Enterprises and Entities, 2011. Redundant Staff and Labor Reserve. Newsletter, 12 September. Retrieved 2020-09-10, from http://www.e-forosimv.gr/docs/1597_2011.pdf [in Greek].
- National and Kapodistrian University of Athens, School of Education, Department of Early Childhood Education, 2013. Postgraduate studies program in special education. In: Call for Applications for Postgraduate Students for the Academic Year 2013–2014. Retrieved 2020-09-10, from http://www.ecd.uoa.gr/?page_id=2478 [in Greek].
- Nteropoulou-Nterou, E., Slee, R., 2019. A critical consideration of the changing conditions of schooling for students with disabilities in Greece and the fragility of international in local contexts. *Int. J. Incl. Educ.* 23 (7), 891–907. <https://doi.org/10.1080/13603116.2019.1623331>.
- Nteropoulou-Nterou, E., 2012. Evaluation of the progress of legislative changes thirty years after the passage of the first law for special education in Greece. In: Zoniou-Sideri, A., Nteropoulou-Nterou, E., Viachou-Balafouti, A. (Eds.), *Disability and Education Policy: Critical Approach to Special and Inclusive Education*. Pedio, Athens, pp. 123–152 [in Greek].
- Oliver, M., 2004. If I had a hammer: the social model in action. In: Swain, J., Barnes, C., Mercer, G. (Eds.), *Disabling Barriers-Enabling Environments*, second ed. Sage Publication Ltd, London, pp. 7–12.
- Olssen, M., Peters, M., 2005. Neoliberalism, higher education and the knowledge economy: from the free market to knowledge capitalism. *J. Educ. Pol.* 20 (3), 313–345. <https://doi.org/10.1080/02680930500108718>.
- Oppenheim, C., 1993. *Poverty the Facts*. Child Poverty Action Group Ltd, London.
- Postgraduate Studies Program in Special Education, 2019. Applications File for Admission to Postgraduate Studies Program in Special Education 2009–2019. Department of Early Childhood education, National and Kapodistrian University of Athens, Athens [in Greek].
- Roulstone, A., Prideaux, S., 2008. More policies, greater inclusion? Exploring the contradictions of new labor inclusive education policy. *Int. Stud. Sociol. Educ.* 18 (1), 15–29. <https://doi.org/10.1080/09620210802195939>.
- Runswick-Cole, K., Goodley, D., 2015. Disability, austerity and cruel optimism. *Can. J. Disabil. Stud.* 4 (2), 162–186. <https://doi.org/10.15353/cjds.v4i2.213>.
- Siebers, T., 2008. *Disability Theory*. The University of Michigan Press, Michigan.
- Slee, R., 2004. Inclusive education and educational reform in new Era. In: Zoniou-Sideri, A., Spandagou, I. (Eds.), *Education and Blindness, Current Trends and Prospects*. Greek Letters, Athens, pp. 31–42 [in Greek].
- Slee, R., 2011. *The Irregular School*. Routledge, Abingdon.
- Slee, R., 2012. How do we make inclusive education happen when exclusion is a political predisposition? *Int. J. Incl. Educ.* 17 (8), 895–907. <https://doi.org/10.1080/13603116.2011.602534>.
- Slee, R., 2018. Defining the scope of inclusive education. In: Background Paper for the 2020 Global Education Monitoring Report. UNESCO, Paris.
- The Constitution of Greece, 2010. The Constitution of Greece. As Revised by the Resolution of the Parliamentary Assembly of 27 May 2008. Hellenic Parliament, Athens [in Greek].
- Tomlinson, S., 2005. Race, ethnicity and education under new labor. *Oxf. Rev. Educ.* 31 (1), 153–171. <https://doi.org/10.1080/0305498042000337246>.
- Vainanidi, C., 2019. Above All, We Are Teachers. *Alfavita*, (Educational portal). June 27, 2019. Retrieved 2020-09-10, from https://www.alfavita.gr/ekpaideysi/292300_pano-apo-ola-eimaste-paidagogoi-i-ekpaideytikos-apo-ti-samo-mila-sto-alfavita [in Greek].
- Whitty, G., 2002. *Making Sense of Education Policy*. Paul Chapman Publishing, London.

Inclusive education in the Republic of Singapore: a situated perspective

Levan Lim and Thana Thaver, Psychology and Child & Human Development Academic Group, National Institute of Education, Nanyang Technological University, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	426
The cultural-historical framework	426
Singapore: contextual legacies and priorities	427
Toward inclusive education	428
Situating inclusive education within Singapore	428
Post-2004: developments in special needs education and disability	429
Initiatives and support for preschool children with disabilities	429
Current practice and support within mainstream schools and Institutes of Higher Learning	430
Inclusive education as a catalyst for reform	431
Conclusion	432
References	433

Introduction

Over a quarter of a century has passed since the Salamanca Statement and Framework on Action for Special Needs Education pronounced inclusive education as “the most effective means of combating discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all” (UNESCO, 1994, p. ix). According to inclusive education scholars (e.g., Ainscow et al., 2019), the Salamanca Statement has had and continues to have a major worldwide influence on many countries in adopting and developing inclusive education policies and practices for all learners regardless of their diverse backgrounds or profiles.

In contrast with many countries around the world that have officially embraced inclusive education as a core educational agenda to transform their education systems in order to become more inclusive of all learners, Singapore remains an anomaly in that it has yet to officially adopt inclusive education as an educational agenda within its education system. Singapore did not participate at the World Conference in Special Needs Education at Salamanca and was not a signatory to the Salamanca Statement and Framework for Action on Special Needs Education. To inclusion advocates, Singapore’s non-adoption of inclusive education as official policy can appear baseless considering that its success and stability as a racially, culturally, linguistically and religiously diverse nation rests upon being an inclusive and socially cohesive society.

It can also seem ironic that inclusive education does not have a formal role in the education system since Singapore has an aspirational vision of itself as an inclusive society for its citizens, in particular for those with disabilities. In this article, we have chosen to focus on inclusive education for persons with disabilities for the reason that persons with disabilities were explicitly mentioned as part of its inclusive society vision announced in 2004 which was then followed by the most progressive measures and efforts ever to promote their quality of life and inclusion within society. Hence, in spite of the lack of a formal endorsement of inclusive education as official rhetoric and policy, Singapore has witnessed an unprecedented slew of national developments over the past eighteen years to support and include children, youth and adults with disabilities within mainstream education and society, as particularly evident in the exponential growth in the number of students with disabilities enrolled within mainstream schools (Choo, 2019; Toh, 2018).

To make sense of and appraise Singapore’s approach and its developments toward inclusive education, this article provides an understanding of inclusive education that is situated within Singapore’s own unique socio-historical and socio-political context. For this purpose, we draw upon the cultural-historical framework (Kozleski et al., 2011) that seeks to re-understand inclusive education as situated within larger cultural historical contexts where its influence and implementation are circumscribed within and mediated by national priorities, historical legacies, competing discourses, sociocultural constructions of difference and other relevant contextual variables. This foregrounding of history and culture within the context of Singapore is invaluable to framing the understanding of Singapore’s approach to inclusive education and its developments toward inclusive education, which are described and discussed in this article.

The cultural-historical framework

Even though inclusive education and its interpretation and implementation occur within national contexts, each with unique socio-historical, socio-cultural and socio-political legacies and conditions, researchers in inclusive education tend to under-theorize, under-estimate and even ignore contextual complexities that inevitably influence the development and challenges of inclusive

education (Kozleski et al., 2011; Lim et al., 2019). A positive development that has emerged in the field of inclusive education over the past decade to recognize and highlight the influence of context on the evolution of inclusive education is the cultural-historical framework (Kozleski et al., 2011) that highlights the significance of history and culture in providing the foreground to better understand the development of inclusive education within a particular context.

The cultural-historical framework highlights that the translation of inclusive education from universally mandated intent to actual practice within a context is often influenced and affected by more encompassing societal and national conditions and developments that are further embedded within a complex, volatile and interdependent global context. Hence, variations across countries in their local interpretations and implementations of inclusive education, from the perspective of the cultural-historical framework, are hence not so much construed as aberrations from the norm, but rather are better understood as legitimate expressions of a country's construction of and response to difference within the larger scheme of nation-building at a particular point in time (Lim et al., 2019). The cultural-historical framework therefore provides a situated perspective of the developments hitherto in the fields of special needs education, disability and inclusive education within Singapore and alerts us to be cognizant that past, present and future developments in these fields are circumscribed by larger contextual legacies, priorities and forces.

Singapore: contextual legacies and priorities

Singapore is a small and densely populated (with a population of 5.5 million) island city-state located at the base of the Malay Peninsula in Southeast Asia which became an independent nation in 1965. Singapore has a multi-ethnic, multi-lingual and multi-religious population. Its citizens are categorized into four main groups: Chinese, Malay, Indian and Eurasian. There are four official languages in Singapore, namely English, Mandarin, Malay and Tamil. English is the medium of instruction in schools. According to a 2014 analysis by the Pew Research Center on global religious diversity, Singapore was found to be the most religiously diverse country in the world. With a diverse multiracial, multicultural, multilingual and multi-religious population that is packed densely within a little over 720 square kilometers of land size, Singapore is also one of the most peaceful (Global Finance, 2020) and safest cities (Economist Intelligence Unit, 2019) in the world.

Singapore's meteoric rise over the course of five decades to become a global city-state par excellence and an exemplar for successful multiculturalism is a well-known success story that is all the more remarkable in light of the odds that conspired against its survival and chances for success during its early years of independence: its fractured inception as a nation, tense race relations, socio-political and economic unrest, and bereft of its hinterland and natural resources after its separation from Malaysia. As a member state of the United Nations since September 21, 1965, Singapore's application to join the United Nations on September 3, 1965 to gain international recognition of its independent sovereignty was precipitated by its painful separation less than a month before on August 9, 1965 from the Malaysia Federation (Lim, 2019). Thrust abruptly into independent statehood, Singapore's uneasy birth as a sovereign nation was tearfully announced by its founder and first Prime Minister, Mr. Lee Kuan Yew, in a televised news conference as a "moment of anguish." As a fledgling sovereign nation without natural resources, including sufficient water of its own and limited land, and compounded by racial tensions, mass unemployment, housing shortage and economic unrest, the birth and early years of Singapore's independence were marked by an acute sense of its vulnerability.

Confronting its own innate vulnerability and its risks for instability by its pioneer leaders and its people, however, helped define a primordial source and distinctive positive force for Singapore's survival, growth and development. The narrative of vulnerability, which continues in Singapore's national consciousness, galvanized the wherewithal of the pioneer generation to transcend its beleaguered state of existence to become a country that is a world of difference from what it was. What fueled the process of Singapore's transformation since its "breakup" with Malaysia is aptly conveyed in Derek Walcott's sublime epigraph from his Nobel lecture which has been cited to capture the essence of Glissant's existential ontology of difference: "Break a vase, and the love that reassembles the fragments is stronger than the love which took its symmetry for granted when it was whole" (cited in Headley, 2012, p. 59).

It is no wonder therefore that Singapore's ontological relationship with its own sense of vulnerability has produced national educational epithets like "No-one owes Singapore a living," and has also shaped and fostered the national characteristics of tenacity, determination, self-reliance, competitiveness, efficiency, a keen sense of urgency, and "planfulness" for the future. Through sheer will, girded by these qualities, Singapore delivered itself from the angst of its vulnerability to become a country whose global significance and reputation in the world punches way beyond its little size and weight.

The historical legacy of Singapore's narrative of vulnerability and the passionately sustained tenacity it has taken for its government and people to overcome its adversities to become an exceptionally prosperous, successful and peaceful country has left an indelible stamp on Singapore's approach to framing and justifying its nation-building efforts especially in the key areas of social, educational and economic development. It is an approach toward matters of governance, policy making and national affairs that is strongly characterized by pragmatism focused on outcomes rather than alliances with ideological positions (Quah, 2018; Tan, 2012), and is upheld in "fostering a harmonious, inclusive and prosperous society" (Ghosh, 2018). Singapore's pragmatic stance to all its affairs as a legacy of its vulnerability narrative and reinforced by its outstanding accomplishments based on a focus on achieving realistic and practical outcomes rather than ideological alliances, has apparently also influenced its response and approach to inclusive education.

Toward inclusive education

Singapore's historical developments in the fields of disability and special needs education can be categorized into two eras—before and after 2004. Before 2004, special needs education in Singapore was largely identified with and associated with the special education system comprised of voluntary welfare associations (known currently as Social Service Agencies) charged with the operation of special (or special education) schools and adult disability services. A dual system of education was clearly in place in which students with and without disabilities were seen as belonging to the special education system and mainstream education system respectively (Lim and Nam, 2000). The integration, instead of the inclusion, of students with disabilities into mainstream schools was the prevailing philosophy and practice of the pre-2004 era (Lim and Nam, 2000; Lim and Quah, 2004; Quah, 1990, 2004) where the onus of responsibility laid with students with disabilities to keep up with the academic performance expected of all mainstream school students to pass national examinations. The prospects for inclusion or inclusive education to make headways were dim in this first period of the history of special needs education in Singapore (Lim and Nam, 2000; Lim and Quah, 2004; Lim and Tan, 1999).

The emphasis on integration echoed the beliefs and sentiments of the pre-2004 era which were defined and influenced by the report of the Advisory Council on the Disabled (1988) entitled "Opportunities for the Disabled"; the most relevant national document existing at that period. This document stated: "Whenever appropriate and feasible, special education should be provided within the regular educational system. A child should only be placed in a special school if he cannot be well educated in a regular school" (pp. 37–38). This report also recommended that: "integration should fit the disabled child to the most suitable educational environment" (p. 38); however, in reality during that time period, special education services were generally provided outside the mainstream education system (Quah, 1993).

By 2003, it was becoming apparent that the increasing number of students with disabilities enrolling in mainstream schools was a cause for concern because of the limited amount of support mainstream schools had for them. The increase in students with disabilities in mainstream schools is attributed to strides made by the Child Development Units in their early detection and intervention programs, the pre-school integration effort, and the fact that children with disabilities or developmental delays, whether diagnosed or latent, were being enrolled by their parents in mainstream schools. While the government still remained strongly in favor of integrating only those students with disabilities who were able to cope with the rigorous pace of the mainstream setting, and did not require accommodation beyond what was present in the mainstream education system, by 2004, it found that this was not a stance that it could maintain given the report that 20% of the school population entering mainstream schools had learning disabilities (Basu, 2006).

Hence, by 2004, more encompassing measures were needed to support the increasing number of students with disabilities within the mainstream education system. This untenable situation of underserving the educational needs of swelling numbers of students with disabilities in mainstream education provided the opportune moment to think and move beyond integration.

The year 2004 has been described as the watershed year (Lim and Thaver, 2018) for the fields of disability and special needs education in Singapore due to the explicit mention by the Prime Minister of Singapore, Mr. Lee Hsien Loong, of persons with disabilities as part of the envisioning of Singapore as an inclusive society in his inauguration speech in August 2004. About a month after his inauguration, Prime Minister Lee committed his government to this vision of an inclusive society through education by calling for better efforts to integrate students with disabilities within mainstream schools (Teo, 2004). The recognition from the Ministry of Education that children with disabilities could benefit from their inclusion within mainstream schools with the appropriate supports clearly conveyed Singapore's initial shift toward greater openness to inclusive education.

Situating inclusive education within Singapore

Singapore's shift toward a greater openness to inclusive education by changing mainstream schools to better accommodate students with disabilities was a practical response, beyond integration, to supporting their educational needs. Singapore, however, has not turned to nor embraced inclusive education as official policy to support Singapore as an inclusive society. To make sense of Singapore's approach toward inclusive education is to further contextualize it within Singapore's larger national priorities during that period. Prime Minister Lee's envisioning of Singapore as an inclusive society in 2004 which explicitly included persons with disabilities followed soon after by his call to better support students with disabilities in mainstream schools, was, in retrospect, embedded within a larger nation-building agenda.

The new dawn in Singapore's sociopolitical history in 2004 provided the driving force for the government to build the "heartware" and "hardware" for Singapore's bold new plans for self-transformation in order to continue to thrive in a world rapidly shifting toward globalization and global competition for talent and jobs. Staying cohesive, united, inclusive, aspirational and rooted to Singapore (i.e., the "heartware") would create a greater and more secure sense of Singapore as home, and "hardware" infrastructural developments would create a more livable and globally attractive city. Although there is little doubt that Singapore has always been changing itself in response to the larger global society, the launch of its own self-transformation and re-invention during the years following 2004 took on a new level and momentum without precedent.

The re-invention and transformation of Singapore was announced by Prime Minister Lee in his first May Day message after taking office, as necessary to staying ahead of rising competition especially from countries within Asia (Ministry of Information, Communication and the Arts, 2005). Billions of dollars were poured into transforming the Marina Bay area, the financial district, Sentosa Island and downtown areas through commercial and real estate projects, the building of two integrated resorts, the

introduction of Formula 1 racing, and the creation of a livelier entertainment and liberal arts scene. These attractions have added tremendously to Singapore's global reputation as a tourist destination and have earned billions of dollars for the Singapore economy.

Meanwhile, to nurture the "heartware" of society, Prime Minister Lee continued his promise of building an inclusive society in his second National Day rally speech in 2005, calling on everyone to play a part in remaking Singapore as a vibrant global city that can be proudly called home (Lee, 2005; Lim, 2005; Teo, 2005). The theme of Singapore becoming an inclusive society was again reiterated in Prime Minister Lee's third National Day rally speech in 2006 (The Straits Times, 2006) and through the years, the inclusive society rhetoric has become a salient keyword in Singapore's aspirational identity. In the search for new ways to co-create inclusive policies and build trust and engagement with diverse groups in society, the government launched national dialogs and conversations over the years, such as Singapore 21 (1999), Remaking Singapore (2002), SGfuture (2015) and the most recent, Our Singapore Conversation, with citizens across a diverse range of issues and policy areas (Ho, 2019).

Post-2004: developments in special needs education and disability

After the pivotal announcements of an inclusive society and better support for students with disabilities in mainstream education in 2004, the developments and directions in the fields of special needs education and disability that soon followed and which have continued to evolve progressively till the present, have been unparalleled in scope and outreach.

At the national societal level, the most comprehensive series of national roadmaps to have ever been developed in Singapore for persons with disabilities was first launched in 2007. These Enabling Masterplans offer a life-course approach to services in the disability sector with systematically drawn-up 5-year national roadmaps (known as Enabling Masterplans) to create an inclusive society where persons with disabilities can be supported and empowered to maximize their potential and be embraced as equal citizens. After its initial 5-year span (2007–2011), the first Enabling Masterplan was consecutively succeeded by the second Enabling Masterplan (2012–2016) and third (2017–2021) Enabling Masterplan. The next Enabling Masterplan (2022 - 2030) has a longer span for implementation compared with the previous Enabling Masterplans. In 2012, Singapore signed the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) and ratified it in 2013. The approach taken by Singapore, as stated in its 2016 UNCRPD report, point 19 on Article 24 (Education), is the life-course approach to services in the disability sector enacted through its series of Enabling Masterplans to inform the development of programs and services in order to progressively realize its vision of an inclusive society where people with disabilities are fully integrated and contributing members.

Within education, while there is no legislation mandating inclusive education pertaining to students with disabilities, increasing provisions and government funding as well as new policies to facilitate the greater inclusion and engagement with stakeholders and the community, have been key elements in the approach adopted since 2004 when Prime Minister Lee Hsien Loong shared the government's vision of an inclusive society—a theme which was reiterated in Singapore's 2020 Budget statements. With each decade since the call for an inclusive society by the government in 2004, the government has increased funding for education and support for people with disabilities; this is featured in government annual budgets. Currently, \$400 million per year is spent on initiatives for persons with disabilities and their caregivers while funding for special education schools by the government has increased by 40% since 2015 (Government of Singapore, 2019). In the past decade, the Ministry of Education (MOE) has committed \$150 million to the upgrading and building of special education schools to provide better facilities and meet rising demand for places (Disabled People Association, 2016).

In 2019, following up on its announcement in 2016 and after efforts on strengthening the structures and processes within the special education sector, the Compulsory Education Act which had been passed in 2000 was extended to include children with moderate to severe needs (Teng and Koh, 2017). The Act defines compulsory education as education in national primary schools for Singapore citizens residing in Singapore of compulsory school age (above 6 years old and under 15 years old). This amendment requires children with mild disabilities who have the cognitive abilities and adaptive skills to attend mainstream primary schools and children with moderate to severe disabilities to attend government funded special education schools (now recognized as "national primary schools") offering primary level education.

Initiatives and support for preschool children with disabilities

The post-2004 years have also seen a large expansion of funding, resources and services for young children of preschool age with disabilities. Children with developmental delays in Singapore are supported through the government-funded Early Intervention Program for Infants and Children (EIPIC) which provides social, learning and therapy services. Each child's progress is monitored through an Individual Education Plan, and intervention is conducted by a trans-disciplinary team of EIPIC teachers, therapists, psychologists and social workers. Since its commencement in 2003, the Government has expanded coverage and increased subsidies for children receiving such services.

The EIPIC has expanded from 9 centers offering 1350 places in 2006, to 17 centers in 2015, serving 2400 children and to 3200 places in 2018. In 2019, the Ministry of Social and Family Development (MSF) rolled out more customized early intervention services tailored to the varied developmental needs for children enrolled in the program at more affordable fees. The government is currently investing \$60 million a year in the program; a 30% increase from \$45 million in 2018 (Ministry of Social and Family

Development, 2020a). MSF is further exploring how to better integrate children with learning needs into pre-schools and also working with Singapore's Early Childhood Development Agency (ECDA), which oversees the early childhood sector, to take the early intervention programs under its aegis so that it can have better oversight of the developmental needs of all children under the age of seven and "strengthen the ecosystem of support" for these children (Teng, 2019a). It is envisaged that this centralization will aid parents in their search for support for their children.

Apart from the EIPIC, there has also been the development of the government-funded Integrated Childcare Program (ICCP). This is an inclusive program offered in mainstream childcare centers for children with mild to moderate disabilities ages 2–6 to prepare them for future entry into mainstream primary education. Its goal is to provide a natural learning environment for these children to socialize and learn with their peers without disabilities. In addition to this, there are the Development Support (DS) and Learning Support (LS) programs to support children with mild developmental needs to prepare them for entry into mainstream schooling. This involves helping children from EIPIC centers transition into a mainstream preschool setting through continued support and co-teaching with teachers in the preschool with the broader aim of equipping these children with the skills to function within a larger class setting.

Acknowledging that the quality of implementing the above initiatives to support the inclusion of young children with disabilities in preschool education depends ultimately on the quality of professionals involved, plans for strengthening the competencies of teachers within the early childhood sector are underway. In October 2020, the Minister for Social and Family Development, Mr. Masago Zulkifli, announced that ECDA would be developing a "Continuing Professional Development" (CPD) roadmap to guide the professional development of early childhood educators (Ministry of Social and Family Development, 2020b). The CPD will delineate the core competencies in priority areas and ECDA is also working with Singapore's National Institute of Early Childhood Development (NIEC) to increase the emphasis on inclusive practices at pre- and in-service training. The goal is to enhance educators' capacities in catering to the diversity of learning needs within their classrooms and develop more inclusive practices.

Current practice and support within mainstream schools and Institutes of Higher Learning

The current perspective that Singapore seems to have adopted is that of viewing the education system, with reference to students with disabilities, not as part of a dual system as in the pre-2004 era, but as one national system with a "continuum of services" emphasizing quality education in an appropriate school setting that can best serve their needs, either in a mainstream or government-funded special education school (Enabling Masterplan Steering Committee, 2016; Puthuchear, 2017). This was made clear in the definition of national primary schools in the Compulsory Education Act (Ministry of Education, 2020a).

Since the watershed year of 2004, the number of students with disabilities attending mainstream schools has risen considerably from 2500 students in 2005 (Chan, 2005) to 13,000 in 2013, to 24,000 in 2018 (Toh, 2018) and to 25,600 in 2019 (Choo, 2019). For a system of education that was traditionally dualistic in terms of the placement of students with and without disabilities into separate education systems, the exponential increase in the number of students with disabilities attending mainstream schools over those in special education schools has blurred the traditional boundaries and notions of special needs education associated with the pre-2004 era. Current statistics from the Ministry of Education (MOE) indicate that 80% of children with special educational needs are being served in mainstream schools while the remaining 20% requiring more specialized and individualized support attend one of Singapore's special education schools (Ang, 2020b).

Current forms of support for students with mild to moderate disabilities within mainstream schools include specialist support in the form of a cadre of personnel termed as "Allied Educators" (Learning & Behavioral Support) [AEDs (LBS)]; which was renamed as Special Educational Needs (SEN) Officers as of April 2022. This was introduced into mainstream schools in 2005 with the call for a more inclusive society and official support for the inclusion of students with disabilities within mainstream schools. The role of the AED (LBS) as delineated by Singapore's MOE is to support students with mild special educational needs such as dyslexia, attention deficit hyperactivity disorder (ADHD), mild Autism Spectrum Disorder (ASD), sensory and physical impairments in the form of individual or small group intervention or in-class support. There are reportedly 600 AEDs (LBS) serving in mainstream schools; the goal is to have two in each primary school and one in the secondary school (Teng, 2020).

In addition to the AED (LBS) support personnel scheme, 2005 also saw the proposal and implementation of a systemic plan to train a core group of mainstream primary and secondary school teachers known as "Teachers trained in Special Needs" (TSNs) to provide learning support as well as to help build the capacity of the teachers in their schools in working with students with diverse needs. The goal then was to have 10% of teachers in all primary schools and 20% in all secondary schools take the Certificate in Special Needs so that the knowledge and skills acquired by these teachers would cascade through the system. In addition to this, in pre-service programs at the National Institute of Education (Singapore's sole teacher education certification body), beginning teachers were equipped with basic understanding of how to cater to students with disabilities.

To provide further help in the early years of schooling to students experiencing learning difficulties, there is the Learning Support Program for primary school learners who require additional aid with English Language in the initial two years in school (Primary 1 and 2) and the School-based Dyslexia Remediation (SDR) program for students in their third and fourth year in primary school (Primary 3 and 4) assessed to have dyslexia. The SDR program is conducted by trained school personnel using a MOE-designed curriculum. Students with dyslexia beyond Primary 4 would attend the Main Literacy Program conducted by the Dyslexia Association of Singapore (DAS).

Although the current framework in terms of educational placements in Singapore still recommends that mainstream school settings would be more suited to students with disabilities who are able to access and cope with the mainstream curriculum with support and those with higher support needs who require more intensive specialized support would be more appropriately served in special schools, there is acknowledgment that Singapore's landscape is evolving, and that these placements and provisions will change over time (Ministry of Education, 2018). For example, within the mainstream education system, the paradigm and approach to support students with disabilities have changed with the building of the systemic capacity of mainstream schools to take ownership and responsibility in this area through initiatives such as the training of mainstream teachers as TSNs, the specialist support in the form of the AED LBS, implementation of the school-based remediation programs and provision of support services from providers of psychological services and social service agencies as well as the coaching provided by the Psychological Services Branch, MOE.

This incremental capability enhancement toward a whole-school systemic response for supporting students with special needs in mainstream schools since 2005 prepared the ground for the recent introduction of a more comprehensive framework of support by MOE to all mainstream schools known as the "Tiered-System of Support for Special Educational Needs" (TSS). This framework is similar to the multi-tiered model of Response to Intervention (RTI) where intensity of support provided is matched to the severity of a student's learning and/or behavioral issues (Aljunied, 2021; Dutt et al., 2019).

Within each school, a "Case Management Team" (CMT) oversees the support of students with special needs and implementation of the recommended interventions. Psychologists from the Psychological Service Branch monitor and build the capabilities of the CMTs through regular meetings and training sessions.

This move toward a tiered system of support within mainstream schools represents a significant step toward inclusive education because it proactively seeks to identify and provide intervention for all students who are at risk due to learning and/or behavioral difficulties not just those with special educational needs. In addition, this tiered system of support will require the collaboration and teamwork of relevant staff members to coordinate efforts in order to sufficiently address the learning and behavioral needs of at-risk students who may require more intensive interventions with multiple supports and services.

One corner in the strengthening of the capability of the mainstream education system in supporting the inclusion of increasing numbers of students with disabilities over the past fifteen years, has been the attention devoted to the professional development of educators. Since 2004, there has been increased focus on professional development in the area of special needs education within mainstream teacher education, the early childhood sector and Institutes of Higher Learning. Within the Institutes of Higher Learning, as of 2015, more than 1500 staff from the polytechnics and Institutes of Technical Education (ITE) have undergone professional development courses in special needs; the goal is to have all staff in the polytechnics and ITE gain knowledge and skills to cater to a more diverse student population (Disabled People Association, 2016).

In March 2020, a significant milestone was reached. MOE unveiled a professional development roadmap to ensure that all mainstream educators would gain basic knowledge and skills in the area of special educational needs to help them better support students with disabilities in mainstream settings. This roadmap is part of the SkillsFuture for Educators (SFEd) initiative and delineates core competencies which all mainstream teachers are expected to attain at pre-service ("emergent" level) and in-service ("proficient" level) through bite-sized online learning and professional development workshops (Ang, 2020a; Ministry of Education, 2020b,c).

Besides building the capability of the mainstream education system to support and practice inclusion, MOE initiated collaborative partnerships between mainstream schools and special schools to facilitate greater interaction and more inclusive attitudes among students from both types of schools through the Satellite Partnership Program in 2008, starting with two special education schools and three mainstream schools. The schools that participate in the program are generally located near each other. This program creates opportunities for students from both schools to mingle through co-curricular activities, games and community projects. Students with disabilities deemed academically ready can also join their peers in the mainstream school for academic subjects. Through this partnership, collaborations, exchanges of ideas and inclusive practices have also been facilitated between teachers from the schools. Benefits have been reported in the deepening of understanding and greater acceptance of people with disabilities on the part of mainstream students and learning of social skills and increased confidence in terms of interactions on the part of students from the special education schools (Teng, 2019b). Currently, 16 of the 19 special education schools are in these partnerships; the goal is to have all 19 in the program.

Inclusive education as a catalyst for reform

In spite of all the developments that have occurred since 2004 to promote the inclusion of persons with disabilities, it will take more time and effort for Singapore society to build the "heartware" to be more inclusive of persons of disabilities. While the focus of the provisions and initiatives seem to be on adequate provisions and accommodations for learning needs in order to pave the way for progress, especially in terms of "hardware," there is growing acknowledgment since 2016 of the need to work on the "heartware" of Singapore society and to intentionally design opportunities for social interaction to educate and develop more inclusive attitudes within the larger community. This resolve to work harder on the aspect of social inclusion gained impetus from the findings of two surveys in 2016, one commissioned by the Lien Foundation and the other by the National Council of Social Service (NCSS).

The findings from the surveys revealed that the majority of Singaporeans felt that Singapore was not an inclusive society, and while willing to share public spaces, preferred not to interact with people with disabilities. They were also not confident interacting with children with disabilities and were not comfortable being close friends with persons with disabilities. The majority of people with disabilities who participated in the NCSS survey stated that they did not feel socially included, accepted or given opportunities to achieve their potential.

The Lien Foundation (2016) survey findings also indicated that only one in four parents reported that their children were friends with children with disabilities and that the primary barrier was the lack of conducive opportunities. The NCSS survey findings revealed that Singaporeans felt and behaved more positively toward persons with physical or sensory impairment than those with intellectual disability or ASD, and that attitudes toward people with disabilities were more positive with greater contact (National Council of Social Service, 2017).

Within the education sector, existing research studies conducted with mainstream Singaporean teachers on their attitudes toward and experiences with students with disabilities reveal neutral to negative responses. Poon et al.'s (2016) survey of 131 teachers and school professionals from two mainstream secondary schools reported an overall neutral attitude toward the inclusion of students with disabilities. Thaver and Lim's (2014) survey of 1538 preservice mainstream teachers reported generally negative attitudes toward people with disabilities, ambivalence toward the inclusion of students with disabilities and a preference for special school placements for students with disabilities. Both positive and negative experiences related to the inclusion of students with disabilities were reported in Yeo et al.'s (2016) research study. In their focus group interviews with 202 teachers across 41 primary mainstream schools, the teachers indicated more negative than positive experiences. Positive experiences included satisfaction with students' progress and new learning for teachers while stress from dealing with challenging behaviors and instructional difficulties in attempting to meet the diverse needs of the class comprised dominant negative experiences.

In light of these attitudinal findings which, per se, constitute barriers to the inclusion of persons with disabilities in education and society, we believe that inclusive education is all the more relevant and important to the inclusion discourse as well as future directions for the fields of special needs education and disability in Singapore. In a country like Singapore which is intentionally seeking to become more inclusive through education, inclusive education provides a critical spotlight on identifying and dismantling educational exclusion and marginalization, poses questions about the power relations of schooling, confronts injustice, and re-envisions education as a democratic apprenticeship to build inclusive and sustainable communities (Slee, 2011).

Inclusive education can offer an "insurrectional edge" to interrogating the foundations and assumptions of educational policy and practice (Slee, 2008) as well as institutional structures, social policies and cultural mindsets in Singapore that have contributed to prevailing public attitudes toward persons with disabilities. As such, inclusive education can identify and highlight areas of education, concerning not just students with disabilities but all students across mainstream and special education, that are in need of reform because of effects that can compromise the holism and outcomes of their education. Hence, inclusive education, as an element of a broader reform for the architecture of schooling (Slee, 2008), can be a catalyst for reforming schooling in Singapore to be more inclusive.

Conclusion

It is unequivocal that the most progressive developments in the fields of special needs education and disability ever witnessed in Singapore have occurred over the past eighteen years after the envisioning of an inclusive society in 2004. What is particularly encouraging about the thrust of many of these developments is the increasing realization and emphasis that the improvement of the quality of life of persons with disabilities needs to be contextualized in their inclusion within communities and society; and therefore, communities, including school communities, and society need to change in order to become more inclusive of people with disabilities. In hindsight, this change in priority and direction to intentionally open up the inclusionary space within mainstream communities and society for persons with disabilities was circumscribed within a larger nation building agenda to reimagine and remake Singapore for the 21st century.

The pragmatic significance of these national directions, developments and events is clear: as one of the freest economies in the world that is deeply interconnected with the global economy, Singapore's survival and success rest on its pragmatic usefulness and relevance to other countries, and therefore staying ahead of other countries in a highly competitive and volatile global economic climate while remaining a livable, stable and peaceful home for its citizens, is crucial to sustaining Singapore's survival and success (Lim and Thaver, 2021). Therefore, the impetus and response toward inclusion for persons with disabilities can be interpreted as a default to the larger nation building scheme of transforming Singapore for reaping the benefits and addressing the challenges of an increasingly globalized and competitive world.

Singapore's brand of inclusiveness for diversity is sensibly steeped in the practical foundational benefits of its collective nation-building efforts to be stable, peaceful and prosperous. This pragmatic ethos to inclusion provides a contrast to the right-based based approach prevalent in the brand of inclusive education dominant in many Western countries.

In the chronological evolution of Singapore's shift toward the inclusion of persons with disabilities in society, their growing inclusion within mainstream society did not hinge upon them exercising their individual rights to challenge infringements upon their human rights to access and participate in mainstream environments. Instead, their inclusion is couched as a counterpart of a collective long-term aspirational response to the envisioning of Singapore as an inclusive society. What has facilitated and sustained the inclusion movement and its developments in Singapore as a priority in education and society are the alliances that

have been re-invigorated between the inclusion of more vulnerable individuals (such as the elderly and the disabled) and the re-envisioning and re-invention of Singapore for the 21st century (Lim et al., 2019).

Singapore's ratification of the UNCPRD in 2013 means that the role of inclusive education must be strengthened across both mainstream and special education through the challenge of identifying educational policy and practice that act as barriers to inclusion and confronting forms of educational settlements that contribute to exclusionary attitudes. Since the UNCPRD affirms the rights of students to access and participate in an inclusive education, the motivations, aspirations and means for translating the envisioning of an inclusive society will need to be increasingly framed within a right-based perspective and discourse. How inclusive education will evolve in Singapore, and how inclusive education will evolve Singapore into the future, remains to be seen.

References

- Advisory Council on the Disabled, 1988. *Opportunities for the Disabled*. Advisory Council on the Disabled, Singapore.
- Ainscow, M., Slee, R., Best, M., 2019. Editorial: the Salamanca statement: 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 671–676.
- Aljunied, M., 2021. Psychological services for children with special educational needs in mainstream schools. In: Wong, M.E., Lim, L. (Eds.), *Special Needs in Singapore: Trends and Issues*. World Scientific Press, Singapore, pp. 149–167.
- Ang, H.M., 2020a. Professional development roadmap for special educational needs training for mainstream educators to be rolled out. ChannelNewsAsia. Available at: <https://www.channelnewsasia.com/news/singapore/sen-roadmap-moe-professional-development-12499176>. (Accessed 30 November 2020).
- Ang, J., 2020b. Parliament: students with special education needs to get more help and teaching support. Straits Times. Available at: <https://www.straitstimes.com/politics/parliament-students-with-special-education-needs-to-get-more-help-and-teaching-support>. (Accessed 30 November 2020).
- Basu, R., 2006. More kids with special needs attending MOE schools: enrolment of kids with learning disabilities has seen 20% rise every year during past 3 years. Straits Times.
- Chan, S.S., 2005. Singapore Government Media Release: Reply by Mr. Chan Soo Sen, Minister of State, Ministry of Education on Resources in Schools, Special Education, Pre-school Education and Education Hub. Available at: <https://www.nas.gov.sg/archivesonline/data/pdfdoc/20050309994.htm>. (Accessed 30 November 2020).
- Choo, C., 2019. The Big Read: where kids with and without special needs learn together—and it's not in Singapore. ChannelNewsAsia. Available at: <https://www.channelnewsasia.com/news/singapore/big-read-kids-with-and-without-special-needs-learn-together-11952714>. (Accessed 30 November 2020).
- Disabled People Association, 2016. *Achieving Inclusion in Education: Understanding the Needs of Students With Disabilities*. Available at: <https://www.dpa.org.sg/wp-content/uploads/2018/08/Inclusion-in-Education2.pdf>. (Accessed 30 November 2020).
- Dutt, A., Lim, L., Thaver, T., 2019. Behavioral supports in Singapore. In: Bryer, F., Beamish, W. (Eds.), *Behavioral Supports for Students With Special Educational Needs in Asia*. Springer Nature, Singapore, pp. 115–125.
- Economist Intelligence Unit, 2019. *Safe Cities Index 2019*. Available at: <https://www.nec.com/en/global/ad/safecitiesindex2019/index.html#:~:text=The%20Safe%20Cities%20Index%202019,infrastructure%20security%20and%20personal%20security>. (Accessed 30 November 2020).
- Enabling Masterplan Steering Committee, 2016. 3rd Enabling Masterplan 2017–2021: Caring Nation, Inclusive Society. Enabling Masterplan Steering Committee, Singapore.
- Ghosh, N., 2018. Masagos tells UN of S'pore's pragmatic way of governance. Straits Times. Available at: <https://www.straitstimes.com/asia/se-asia/masagos-tells-un-of-spores-pragmatic-way-of-governance>. (Accessed 30 November 2020).
- Global Finance, 2020. *The Most Peaceful Countries in the World 2020*. Available at: <https://www.gfmag.com/global-data/non-economic-data/most-peaceful-countries>. (Accessed 30 November 2020).
- Government of Singapore, 2019. *Budget 2019*. Available at: https://www.singaporebudget.gov.sg/budget_2019/about-budget/look-back-at-recent-budgets. (Accessed 30 November 2020).
- Headley, C., 2012. Glissant's existential ontology of difference. *CLR James J.* 18 (1), 59–101.
- Ho, G., 2019. Singapore together: redefining the National Conversation. Straits Times. Available at: <https://www.straitstimes.com/singapore/redefining-the-national-conversation>. (Accessed 4 December 2020).
- Kozleski, E.B., Artiles, A.J., Waitoller, F.R., 2011. Introduction: equity in inclusive education: historical trajectories and theoretical commitments. In: Alfredo, A.J., Kozleski, E.B., Waitoller, F.R. (Eds.), *Inclusive Education: Examining Equity on Five Continents*. Harvard Educational Press, Cambridge, MA, pp. 1–14.
- Lee, H.L., 2005. *Remaking Singapore as a vibrant global city* (edited excerpt). Straits Times H6–H7.
- Lien Foundation, 2016. *Inclusive Attitudes Survey Part 1: Views of the General Public*. Available at: http://www.lienfoundation.org/sites/default/files/FINAL%20-%20Inclusive%20Attitudes%20Survey%20Part%201_30May16.pdf. (Accessed 30 November 2020).
- Lim, L., Nam, S., 2000. Special education in Singapore. *J. Spec. Educ.* 34 (2), 104–109.
- Lim, L., Quah, M.M., 2004. Foresight via hindsight: prospects and lessons for inclusion in Singapore. *Asia Pac. J. Educ.* 24 (2), 193–204.
- Lim, L., Tan, J., 1999. The marketization of education in Singapore: prospects for inclusive education. *Int. J. Incl. Educ.* 3 (4), 339–351.
- Lim, L., Thaver, T., 2018. Inclusion of persons with disabilities in Singapore: an evolutionary perspective. In: Hsu, S. (Ed.), *Routledge Handbook of Sustainable Development in Asia*. Routledge, New York, pp. 373–393.
- Lim, L., Thaver, T., 2021. Education as apprenticeship for the future: the evolution of Singapore's approach and its impact on inclusion for persons with disabilities. In: Wong, M.E., Lim, L. (Eds.), *Special Needs in Singapore: Trends and Issues*. World Scientific Press, Singapore, pp. 31–55.
- Lim, L., Thaver, T., Strogilos, V., 2019. Contextual influences on inclusivity: the Singapore experience. In: Schuelka, M.J., Johnstone, C.J., Thomas, G., Artiles, A.J. (Eds.), *SAGE Handbook on Inclusion and Diversity in Education*. Sage Publishing, London, pp. 496–508.
- Lim, L., 2005. PM Lee: your vibrant global city, your home. Straits Times 1.
- Lim, T.S., 2019. Singapore joins the United Nations. Singapore Inf. Available at: https://eresources.nlb.gov.sg/infopedia/articles/SIP_2019-04-12_104028.html. (Accessed 30 November 2020).
- Ministry of Education, 2018. *Professional Practice Guidelines: Psycho-educational Assessment and Placement for Students With Special Educational Needs*. Ministry of Education, Singapore. Available at: <https://www.moe.gov.sg/docs/default-source/document/education/special-education/files/professional-practice-guidelines.pdf>. (Accessed 30 November 2020).
- Ministry of Education, 2020a. *Overview of Compulsory Education*. Available at: <https://beta.moe.gov.sg/primary/compulsory-education/overview>. (Accessed 30 November 2020).
- Ministry of Education, 2020b. *Learn for Life—Ready for the Future: Refreshing Our Curriculum and Skills Future for Educators*. Available at: <https://www.moe.gov.sg/news/press-releases/learn-for-life-ready-for-the-future-refreshing-our-curriculum-and-skillsfuture-for-educators>. (Accessed 30 November 2020).
- Ministry of Education, 2020c. *Infosheet on Skillsfuture for Educators (SFE)*. Available at: <https://www.moe.gov.sg/docs/default-source/document/media/press/2020/infosheet-on-SFE.pdf>. (Accessed 30 November 2020).
- Ministry of Information, Communication and the Arts, 2005. Prime Minister's May Day Message. Available at: <https://www.nas.gov.sg/archivesonline/data/pdfdoc/2005043002.htm>. (Accessed 4 December 2020).
- Ministry of Social and Family Development, 2020a. *Press Release: Enhanced Early Intervention: Better Support for Children With Development Needs*. Available at: <https://www.msf.gov.sg/media-room/Pages/Enhanced-Early-Intervention.aspx>. (Accessed 30 November 2020).

- Ministry of Social and Family Development, 2020b. Press Release—Boosting Professional Development of Preschool Educators, Enhancing Outdoor Learning for Children and Expanding KidSTART to More Families. Available at: <https://www.msf.gov.sg/media-room/Pages/Enhanced-Early-Intervention.aspx>. (Accessed 30 November 2020).
- National Council of Social Service, 2017. Understanding the Quality of Life of Adults With Disabilities. Available at: <https://www.ncss.gov.sg/NCSS/media/NCSS-Publications/Pdfdocument/Understanding-the-Quality-of-Life-of-Adults-with-Disabilities.pdf>. (Accessed 30 November 2020).
- Pew Research Center, 2014. Global Religious Diversity. Available at: <https://www.pewforum.org/2014/04/04/global-religious-diversity>. (Accessed 30 November 2020).
- Poon, K.K., Ng, Z.J., Wong, M.E., Kaur, S., 2016. Factors associated with staff perceptions towards inclusive education in Singapore. *Asia Pac. J. Educ.* 36 (1), 84–97.
- Puthucheary, J., 2017. Speech at the Inclusive Education in Higher Education: Inclusion in Action Forum. Available at: <https://www.moe.gov.sg/news/speeches/speech-by-dr-janil-puthucheary-senior-minister-of-state-ministry-of-education-at-the-inclusive-education-in-higher-education-2017-inclusion-in-action-forum-at-tan-chin-tuan-lecture-theatre-nanyang-technological-university>. (Accessed 30 November 2020).
- Quah, M.M., 1990. Special education in Singapore. *Int. J. Disabil. Dev. Educ.* 37 (2), 137–148.
- Quah, M.L., 1993. Special education in Singapore. In: Quah, M.L., Gopinathan, S., Chang, S.C. (Eds.), *A Review of Practice and Research in Education for All in Singapore*. Country Report Submitted to the Southeast Asian Research, Review and Advisory Group (SEARRAG). National Institute of Education, Singapore, pp. 89–102.
- Quah, M.M., 2004. Special education in Singapore. In: Lim, L., Quah, M.M. (Eds.), *Educating Learners With Diverse Abilities*. McGraw-Hill, Singapore, pp. 29–61.
- Quah, J.S.T., 2018. Why Singapore works: five secrets of Singapore's success. *Public Adm. Policy* 21 (1), 5–21.
- Slee, R., 2008. Beyond special and regular schooling? An inclusive education reform agenda. *Int. Stud. Sociol. Educ.* 18 (2), 99–116.
- Slee, R., 2011. *The Irregular School: Exclusion, Schooling and Inclusive Education*. Routledge, London.
- Tan, K.P., 2012. The Ideology of pragmatism: neo-liberal globalization and political authoritarianism in Singapore. *J. Contemp. Asia* 42 (1), 67–92.
- Teng, A., Koh, F., 2017. MOE accepts panel's recommendations on compulsory education for special needs children. *Straits Times*. Available at: <https://www.straitstimes.com/singapore/education/moe-accepts-panels-recommendations-on-compulsory-education-for-special-needs>. (Accessed 30 November 2020).
- Teng, A., 2019a. Parliament: break walls between pre-schools and early intervention: Study. *Straits Times*. Available at: <https://www.straitstimes.com/singapore/education/break-walls-between-pre-schools-and-early-intervention-study>. (Accessed 1 December 2020).
- Teng, A., 2019b. More schools join hands for lessons on inclusivity. *Straits Times*. Available at: <https://www.straitstimes.com/singapore/education/more-schools-join-hands-for-lessons-on-inclusivity>. (Accessed 30 November 2020).
- Teng, A., 2020. Allied educators who support students with learning needs will have more career opportunities. *Straits Times*. Available at: <https://www.straitstimes.com/singapore/parenting-education/allied-educators-who-support-students-with-learning-needs-will-have>. (Accessed 30 November 2020).
- Teo, L., 2004. \$220m school aid for disabled kids. *Straits Times* 1.
- Teo, H.N., 2005. Blueprint for a global city called home: PM maps out the future, envisions a world-class Singapore. *Today Off.* 1.
- Thaver, T., Lim, L., 2014. Attitudes of pre-service mainstream teachers in Singapore towards people with disabilities and inclusive education. *Int. J. Incl. Educ.* 18 (10), 1038–1052.
- The Straits Times, 2006. Being Singaporean: Heart of the Matter (National Day Rally Speech by PM Lee), p. H7.
- Toh, L., 2018. Continuing Efforts to Get Students From Different Groups to Interact. Available at: <https://www.straitstimes.com/forum/letters-in-print/continuing-efforts-to-get-students-from-different-groups-to-interact>. (Accessed 30 November 2020).
- UNCRPD, 2016. Singapore CRPD Report. Available at: <https://www.msf.gov.sg/policies/International-Conventions/Pages/UN-Convention-on-the-rights-of-persons-with-disabilities-UNCRPD.aspx>. (Accessed 30 November 2020).
- UNESCO, 1994. The Salamanca statement and framework for action in special needs education. In: *World Conference on Special Needs Education: Access and Quality*. Available at: https://www.right-to-education.org/sites/right-to-education.org/files/resource-attachments/Salamanca_Statement_1994.pdf. (Accessed 30 November 2020).
- Yeo, L.S., Chong, W.H., Neihart, M.F., Huan, V.S., 2016. Teachers' experience with inclusive education in Singapore. *Asia Pac. J. Educ.* 36 (1), 69–83.

Writing the history of Nordic disability and mental disorders

Bjørn F. Hamre^a and Jesper V. Kragh^b, ^a History and Education, Department of Communication, University of Copenhagen, Copenhagen, Denmark; and ^b The SAXO Institute, University of Copenhagen, Copenhagen, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction	435
General historical patterns—disability and mental disorders	436
Psychiatry from 1800 to the 1970s	436
Disability institutions from 1800 to the 1970s	438
New perspectives: empowerment, patients' voices and critique after 1970	440
The establishment of special education	440
National cases	441
IQ testing, progressivism and professional demarcation	442
The Scandinavian school system in the years after 1945	442
Conclusion	443
Acknowledgment	443
References	443

Introduction

This article considers disability history in relation to the field of education and the emerging welfare states in Sweden, Norway and Denmark. The Nordic countries resemble each other in many ways (Telhaug et al., 2006). The reasons for the great similarity between these countries are geographical, historical, cultural, linguistic and political. The connections between these three countries have traditionally been very close, not least in the fields of education and science. The object of this article is the historical relationship between education, disability/mental disorders and the welfare state in Sweden, Norway and Denmark. Historically, education as a field relates to the development of schooling, and education as a science is associated with teaching, educational systems and upbringing (Popkewitz and Lindblad, 2004; Tröhler, 2020). Education, in this context, concerns how educational ideas transform into practices in education and schooling.

Our way of understanding the history of disability and mental disorders relates to Michel Foucault's explanation of the interaction between technologies of discipline and security in the development of modern states (Foucault, 2009). Educational science and educational practices function as disciplining technologies in the shaping of norms for the upcoming generations, the schooling system and the future society. From this perspective, disability and mental disorders represent disturbing phenomena. Since the era of the Enlightenment, societies have developed technologies of security, that is, institutions and practices. According to Foucault, the institutions in modern states develop between technologies of discipline and securitization (Foucault, 2009).

In light of this, the history of disability and mental disorders concerns the history of the parallel processes of, on the one hand, excluding and, on the other hand, integrating people with disabilities and mental disorders into the modern liberal welfare state. Civil rights, inclusiveness and social security are often associated with the welfare state. Mitchell Dean questioned these ideals, stressing that liberal states have developed some very illiberal practices toward its unfit members (Dean, 2010), for example eugenic legislation and the establishment of island institutions for people who were regarded as "feeble-minded."

To understand the history of disability, we must question how differentiations have been constructed through history (Stiker, 1997). Studying disability is thus about determining how the differentiation of disabilities has been organized and legitimized through discourses, institutions and practices and the rationality behind the differentiation. According to Stiker, even more important in these processes is the question of integration: "Why does society try to integrate the disabled? What is behind the intention? And, more precisely, why does society want to integrate in the way it does?" (Stiker, 1997, 15). The study of the history of disability and mental disorders is about investigating *how* these processes take place. Institutional settings and the incarceration of people with disabilities are expressions of a certain political rationality that has been applied in Western countries for centuries (Ben-Moshe et al., 2014). Alternatively, as disability scholar Bill Hughes (2020, 5) put it, "Examining the past for strands that tell stories of disability leads to the contention that disability history is a history of dis/ability. It cannot escape its umbilical connection to the history of the relations between disability and non-disability." Ability and disability are intertwined phenomena, and approaching disability history could thus be summarized as examining the changes in disability through changing ideas, discourses, institutions and practices.

These remarks hint at our point of departure. The history of disability and mental disorders since the era of the Enlightenment has been a history of how to govern disability through ideas, institutions, practices and professionals. Our focus is on psychiatric asylums, institutions for disabled people and special needs education. On the one hand, these phenomena are entangled and follow similar lines; on the other hand, they developed in very different ways, with their own ideas, institutions and practices. The period of study is 1800 onward, with a focus on 1900–1950. Firstly, we briefly outline some historical trends in the development of the three

Scandinavian countries that influenced the three practices for psychiatric institutions, institutions for disabled people and special needs education in the schooling system. Subsequently, we investigate the practices in three separate sections. These sections draw lines between the early developments, ending up in the current debates on the deinstitutionalization process and inclusion. In the conclusion, we ask what characterizes the differentiation of people with disabilities and mental disorders in the three Scandinavian countries with regard to the three practices.

General historical patterns—disability and mental disorders

Before the 19th century, people with disabilities were not defined as a specific group, and historical records of the time contain sparse information about them (Rønn, 1996, 69–86; Grunewald, 2009, 27–42). However, medieval sources occasionally mention people, described as mad or foolish, who were incarcerated (Goodey, 2011). Until the Enlightenment in the 18th century, disability and mental disorders were considered as natural, demonic or divine. The philosophers of the Enlightenment believed in the importance of education and pedagogy and the moral shaping of the human character (Froestad, 1995, 78–79). Departing from religious beliefs, the Enlightenment thinkers believed in knowledge and reason. Following this line of thought, the poor and people with disabilities were regarded as potential resources for the society (Froestad, 1995). Institutional changes regarding hospitals and asylums for mental disorders also reflected this philosophy. The foundation of the first disability institutions relied on the idea of the period that humans were formable. This change of view was part of the scientific and ideological modernization, including new space for experiments in the field of disability care, education and policy. Ideas took shape through medical development but also in education and the way in which educational institutions for disabled people emerged in Scandinavia. These processes in the three countries were part of a general process of rationalization and modernization.

The development in the countries was a move from private, religious and philanthropic initiatives toward a higher degree of state involvement. This was the case in terms of the social policy, the organization of institutions as well as the establishment of the schooling system. Philanthropy played a significant role in the policy handling the problem of the poor in the last part of the 19th century, that is, the social question. The increasing effects of the industrialization and the market led to greater awareness about children's problems. Norway and Denmark adopted comprehensive child protection acts in 1896 and 1905 (Christiansen and Markkola, 2006, 15). On the threshold of the 20th century, the Scandinavian states passed reforms to handle phenomena such as social security, old age, illness and unemployment. In these processes, the social democrats in the Nordic countries played a major role in promoting social policies, paving the way to the establishment of the welfare state. The social democratic parties managed to obtain a very strong position in the three countries by virtue of their engagement in social rights. In the 1930s, they profiled themselves as parties for the people, which is apparent in the slogans *Folkheim* in Sweden and *Danmark/Norge for folket* (Denmark/Norway for the people) (Korsgaard, 2008). The social democratic parties succeeded in building alliances across parties in the adoption of vast social reforms in the 1930s. The visions and ideas took shape in collaboration with experts such as sociologists, economists and social scientists. Some of the most prominent of these experts were Gunnar and Alva Myrdal, who contributed to the social, political and economic theory of the time (Christiansen and Markkola, 2006, 20). During the interwar years, when the social democrats were in power, eugenic acts were introduced in the Scandinavian countries (Have, 2000; Koch, 2000; Tydén, 2002; Christiansen and Markkola, 2006). The Scandinavian development combined the early development of social democratic parties with the adoption of social rights, with similar early reliance on eugenics as an idea and practice.

Psychiatry from 1800 to the 1970s

The earliest method of isolating the insane made use of a type of wooden cage with thick walls of heavy timber and an iron door with a hatch through which food was passed. Most of these cages were in the hospitals that would later evolve into psychiatric units. The largest collection was at St. Hans Hospital, situated outside of the Copenhagen ramparts in the early 17th century. Later, in 1816, the doctor Johannes Seidelin was appointed as consultant at St. Hans, and the hospital evolved into a psychiatric hospital (Zalewski, 2008; Hamre, 2010; Kragh and Møllerhøj, 2016).

The living conditions of the so-called insane generated little concern, however, until the early 19th century, when Scandinavian physicians raised the alarm. Danish medical doctor Jens Rasmussen Hübertz traveled the country in the 1840s, and he was shocked to see how the insane lived (Hübertz, 1843). In private homes, both rural and urban, he often found people who were chained or bound with rope or who had spent years in wooden crates or confined spaces. Hübertz, in 1843, proposed the building of three big psychiatric hospitals as well as an institution or asylum for the so-called feeble-minded. Similar initiatives were undertaken by the Norwegian doctor Herman Wedel Major (Haave, 2008; Kringlen, 2012). In Sweden, such plans had already been carried out in the 1820s (Qvarsell, 1982; Eriksson, 1989).

In 1848, Norway was the first of the three Scandinavian countries to pass an Insanity Act (Dahl, 2018a; Skålevåg, 2016). Sweden followed suit in 1858, whereas Denmark's first psychiatric law was passed as late as 1938 (Adserballe, 1977; Wetterberg, 2012). In all the Scandinavian countries, the state had a major role in establishing the mental health service. Three large psychiatric hospitals or state asylums were built in Norway during the period 1855 to 1881. Similar Danish state mental hospitals were founded in the late 19th century, and very large hospitals, such as Becomberga Hospital in Stockholm, opened in Sweden (Sjöström, 1992; Jönsson, 1998; Eivergaard, 2003). Besides state asylums, county psychiatric hospitals, municipal hospitals and university clinics were

established during the late 19th and the early 20th century. In the early 1950s, Danish hospital psychiatry had evolved into a comprehensive system with almost 11,000 psychiatric beds. By that time, the Danish population had reached almost 4.3 million (Kragh, 2008, 2010a,b,c). In Sweden, with a population of about 7 million people, psychiatric hospitals had beds for more than 26,000 patients in the same period, and, in Norway, with the smallest population of approximately 3.3 million, psychiatric institutions provided beds for approximately 7000 inmates (Haave, 2008, 257; Berg, 2018, 61).

When the first Scandinavian asylums opened in the early 19th century, medical doctors were appointed as managers or consultants of the hospitals. In the years that followed, the medical profession gained a strong hold over the mental health care system. The profession's representatives also acquired certain recognition in other areas. With the establishment of highly secure detention institutions designed for the most dangerous "mentally ill criminals," psychiatry was involved in the politico-legal area to an increasing extent, and this position solidified further in the 1930s (Börjesson, 1994; Garpenhag, 2012; Bergheim, 2014; Nilsson and Vallström, 2016; Skålevåg, 2016; Dahl, 2018a,b; Kragh, 2018). The first Scandinavian criminal asylum, Kriminalasylet, opened in the city of Trondheim in Norway in 1894. This institution was deemed necessary to protect society from so-called criminal, dangerous or disruptive male psychiatric patients. The Norwegian Criminal Act of 1842 prohibited the punishment of the insane, and, consequently, such offenders could not be imprisoned (Myhre, 2016; Dahl, 2018a,b; Rabben, 2020). The Trondheim criminal asylum was seen as the solution to this problem. In 1910, a Danish state commission presented plans for a high-security ward, which would have 50 beds and would be not only for "insane criminals" and dangerous psychiatric patients but also for "the feeble-minded," "psychopaths," "epileptics," and "alcoholics" (Kragh, 2018). In Sweden, wards for criminals with mental disorders had been set up in the prison system in Norrköping and Ystad, and, after a Swedish act on high-security wards was passed in 1937, a new institution also opened in Hall (Bergheim, 2014; Nilsson and Vallström, 2016).

Patients with psychiatric disorders were to some extent considered a threat or a burden to society in the Scandinavian countries in the 20th century. During the 1940s and 1950s, euthanasia of people with mental disabilities became a heated topic, and the issue was discussed in the Scandinavian countries (Kragh, 2019). In Norway, actual plans to kill a large number of patients in mental hospitals were put forward by the politician Christopher Lange in 1944. Lange's suggestion was considered by the Norwegian Ministry of the Interior, but it finally found the proposal to be too problematic (Haave, 2008, 237–244). At the end of World War II, many patients with psychiatric disorders from the northern regions of Norway were deported by the Germans, but there were no signs of attempts to exterminate the patients during wartime (Fause, 2015). In Sweden, too, euthanasia was discussed (Barow, 2009, 276–280). During a debate in the Swedish Parliament, the later Minister of Culture Arthur Engberg gave a talk about the "feeble-minded." He supposed that "we perhaps in the not-so-distant future are forced to consider wiping out such lives" (Barow, 2009, 277). The famous Swedish feminist writer Ellen Key also left the door open to euthanasia (Tyden, 2010, 366), but the idea did not become reality (Engwall, 2005; Seeman, 2006). In Denmark, public debates took place in 1949 and 1950. Plans about euthanasia for people with psychiatric disorders and intellectual disabilities were proposed in this period, but these were not finalized (Kragh, 2019). As historian Thomas Barow noted in the case of Sweden, the development shows that no radical political system was necessary to put people with mental disabilities in threatening situations. The extent of coercion that the Swedish state could exert over individuals was huge, but in no way did it match the power of the German national socialism. Such differences highlight the limits of conformity between the two countries (Barow, 2009, 397).

In 1913, the first Nordic Psychiatric Congress took place in Copenhagen. After this event, Nordic congresses were organized regularly every third year. The Nordic Psychiatric Association was founded in 1935 (Knorrung, 2012). Therapeutic innovations, especially somatic treatments, were discussed at the Nordic congresses. In the early 1920s, malaria fever therapy was introduced for patients with the diagnosis dementia paralytica (Kragh, 2010a), and, at the Nordic congress in Oslo in 1938, shock treatments and insulin coma therapy were presented (Kragh, 2010b). At the congress in Copenhagen in 1946, lobotomy played a major role. Lobotomy, invented in 1935 by Professor of Neurology Egas Moniz in Lisbon (Valenstein, 1986), was warmly embraced by Scandinavian psychiatrists. The first operations were performed in Denmark in 1939 (Kragh, 2010c, 2021). The psychiatrist Ørnulf Ødegaard and the surgeon Arne Thorkildsen conducted the first Norwegian surgical intervention at the Gaustad psychiatric hospital in Oslo in 1941 (Tranøy and Blomberg, 2005). Psychosurgery came to Sweden a little later—the first lobotomy was performed in 1944 (Ögren, 2005). The Scandinavian countries command a special position in the worldwide context when it comes to lobotomy. More than 2500 operations were performed in Norway and approximately 4400 in Sweden. In Denmark, at least 4500 psychosurgical interventions took place on Danish neurosurgical wards from 1939 to 1983. With a population of only 4 million Danes in 1945, this is an extremely high number. In fact, no other country in the world has equivalently high figures in relation to its population size (Kragh, 2010c, 23–24).

After the introduction of anti-psychotic medication and anti-depressants around the mid-1950s (Swazey, 1974; Healy, 1997, 2002; Gelman, 1999; Valenstein, 2001; Shorter, 2008), the use of psychosurgery declined but did not disappear. The last operations were performed in Norway in 1974 (Tranøy and Blomberg, 2005, 108) and in Denmark in 1983 (Kragh, 2010c). In Sweden, however, precision psychosurgery (capsulotomy) was still undertaken until 2007. More than 200 Swedish patients received psychosurgery in the period from 1970 to 2007 (Bejerot, 2004; Rörbecker, 2010).

Patients selected for lobotomy often had long-term hospital stays. A very large group found itself trapped in the psychiatric system, spending lengthy periods in hospital, and the bed-blocking problem only seemed to grow despite a higher number of discharges at the time. In 1937, 34.6% of all Danish patients had been in hospital for more than a decade. By 1947, the figure was 49.7%, and, by 1952, it had risen to 51.3% (Betænkning, 1956, 148). Similar patterns of hospital stays could be found in Sweden and Norway. However, during the 1960s, hospitalization patterns began to change. Around the mid-1950s, antipsychotic drugs, such as chlorpromazine (Thorazine/Largactil), made their way into Scandinavian psychiatry. The early trials carried out at the

hospitals raised hopes that chlorpromazine and other drugs could improve patients' conditions since they seemed to be effective against a number of classic psychotic symptoms. The hospitals reported that patients were generally calmer and were up and about more on the wards or outside.

Doubts were later raised about whether the drugs would actually improve the hospitals' capacity problems. The initial reports did not suggest a marked change in the turnover of patients. More patients could be treated more quickly, but the trend had little effect on the large group of patients who spent many years in hospital. As the Danish psychiatrist Erik Strömberg noted in a cross-sectional survey of all the hospitals in 1957, "there was no sign that the therapeutic progress had caused any radical change in the overall picture." It was not until a later review in 1962 that he noted that a large group of long-term patients could be discharged and could cope with life outside the hospitals. Psychotropic drugs were part of the reason for this new situation, but, as Strömberg explained, other factors also played a part: "The explanation is in all probability that the rehabilitation measures necessary to maintain the results of pharmacotherapy and to improve on them, were not felt until sometime later. The impact of the introduction of the state pension on this must also be stressed" (Strömberg, 1965). The patients' opportunities to cope with life outside hospitals were limited prior to the introduction of the state pension in 1957. Before then, they had to make do with a disability pension, which was not enough to live on outside of hospital. The state pension was followed a few years later by other initiatives, such as rehabilitation acts and disability pensions. However, the change in hospitalization patterns did not lead to the downsizing of the hospitals in the 1960s. On the contrary, the number of beds was higher than ever before at many of the Scandinavian psychiatric hospitals in the early 1970s (Sjöström, 1992; Haave, 2008, 433ff; Kragh, 2010c). The downsizing of the hospitals began in the mid-1970s. New ideas of community care, critique of the "total institutions," patients' movements and political considerations all played a role in this process.

Disability institutions from 1800 to the 1970s

Froestad (1995) has made a historical analysis of developments within disability institutions and special education in the Scandinavian countries in the 19th century. In the period 1840–1890, the education of people with disabilities expanded internationally as well as in the Nordic countries. Nordic meetings and journals emerged in this period to share knowledge on deaf-dumbness, blindness, feeble-mindedness and other disabilities. In 1872 onward, Nordic meetings were established on what was entitled *Abnormsagen*, which could be translated into English as the question of the abnormal, referring to people who deviated from what was perceived as normal. In the proceeding meetings, the key issue was the organization of the care and upbringing of disabled people. Professionally, as well as politically, these meetings came to play an important role. The sharing of ideas among experts and the development of new pedagogies were the result. The meetings also addressed important questions of education, especially for the blind, the deaf-mute and the feeble-minded.

As noted by Froestad and Kirkebak, the education of the deaf-mute and the blind had the highest status because of the strong involvement of experts in the field and the promising possibilities for integrating these groups into society; at the bottom were people who were perceived as feeble-minded (Froestad, 1995; Kirkebak, 1993). In the organization of the care, there were important national differences. The care of the so-called feeble-minded was conceived as a pedagogical question in Norway, as a medical one in Denmark and as a religious-philanthropic one in Sweden. Froestad argued that a similar situation applied to the education of the deaf and those perceived as feeble-minded (Froestad, 1995, 333).

By the mid-19th century, large institutions for the "feeble-minded" began to emerge. In Denmark, the establishment of these institutions in the 1850s and 1860s ran parallel to the creation of hospitals or asylums for patients with psychiatric disorders (Kirkebak, 1993; Kragh, 2008). During the late 19th century and the early 20th century, a number of large institutions, managed by medical doctors, were created in various parts of Denmark. These institutions were originally founded on a philanthropic basis but received financial support from the state (Kirkebak, 1993, 150–153).

There were, of course, differences between the three countries. The early industrialization in Denmark, for example, meant that tasks such as handling deviancy became part of social policy. Enrollment in and stays at a disability institution were increasingly supported by the public sector. In all three countries, the care for the feeble-minded may be regarded as a prism reflecting the way in which disability care was handled (Froestad, 1995). Differences between the groups were obvious. Whereas doctors had the dominant position in Denmark, Swedish disability care was still influenced by tradition and religious and philanthropic values. Generally, "it can be argued that disability care in Sweden in the latter part of the last century [19th century] was defined as poverty policy, in Norway as school policy, and in Denmark as social policy" (Froestad, 1995, 428, our translation).

Views on feeble-mindedness were influenced by the central role that ideas of degeneration played in the Scandinavian countries. These took shape from international ideas like the French psychiatrist Bénédict Augustin Morel's theory of degeneration in 1857, the scientist Charles Darwin's theory of evolution in 1859 and Gregor Mendel's laws of inheritance but primarily the British psychologist Francis Galton's theory of eugenics.

Informed by the ideas of eugenics, Denmark was the first country to establish legislation on sterilization. In 1924, the Danish Minister of Justice, K. K. Steincke, set up a commission to investigate options for sterilization and castration (Koch, 1996, 56). Two years later, the commission issued a report recommending that the so-called "feeble-minded," "mentally ill patients" and people over 21 years of age with an abnormal sex drive should be permitted to undergo castration or sterilization. Only a small number of physicians found the suggestion to be unacceptable. A few legal professionals also held a critical stance toward sterilization and castration, but most politicians were in favor of the idea. In 1929, a sterilization and castration act was passed by the Danish

Parliament—the first in any European nation (Graugaard, 1997, 127; Koch, 2000, 344). A further act allowing compulsory sterilization and castration of “the feeble-minded” and other groups was passed in 1934 and 1935 (Koch, 1996, 2000, 193–194). Similar acts were passed in Norway in 1934 (adjusted law under the Nazi occupation in 1943–1945) and in Sweden in 1934 and 1941. As in the Danish case, Norway and Sweden adopted the sterilization laws with a broad political majority in their parliaments (Koch, 2006, 304). The numbers of sterilizations in the three countries were approximately 11,000 in Denmark (1929–1960), ca. 17,500 in Sweden (1935–1960) and around 7000 in Norway (1934–1960) (Roll-Hansen, 2006, 263; Tyden, 2010, 375). These numbers are those that can be directly related to the laws; the actual numbers of sterilizations were higher.

To exclude people who were perceived as feeble-minded and dangerous to society, the Danish state organized two island institutions, Livø for men and Sprogø for women (Kirkebak, 1997, 2004). The initiatives were seen as progressive in that time. In the period 1911–1962, approximately 700 men were interned in Livø. These were men who were perceived as slightly mentally weak and able to work and who had been detained for disciplinary reasons or because they had committed offenses. The men’s behavior was interpreted from the eugenic mindset, and the island was considered to be a crucial prevention technology. Similarly, Sprogø, an island for women who were perceived as feeble-minded and frivolous, functioned in the period 1923–1961. Sprogø housed 500 women.

During the early 20th century, the Danish state took more control of the service for vulnerable groups. When the National Assistance Act was passed in 1933, institutional care for the treatment and education of the “feeble-minded” became the responsibility of the state, which was obliged to cover all the expenses for the care of people with physical or mental disabilities (Merrick et al., 2014; Kragh et al., 2015). This involvement by the state reached a new level with the establishment of the Mental Retardation Service under the jurisdiction of the Ministry of Social Affairs in 1959, when all the existing institutions were taken over by the state. At that time, there were about 7500 beds at these institutions. The largest one, in Brejning, which was founded in 1899, had more than 1700 beds (Duedahl, 2019). The Danish hospital-like institutions served as a model for those in Sweden and Norway, where only a few residential facilities had been developed in the 19th century (Grunewald, 2009, 51, 365; Fjermeros, 2009, 177, 261). In Sweden, no more than 380 beds were registered in the asylums (“vårdhem”) in 1910 (Grunewald, 2009, 367). At the time, most Swedish children and adults with disabilities were placed in boarding schools, workhouses or family care. A similar situation existed in Norway, where only two institutions opened in the 1890s. In 1945, a total of 480 beds were registered at Norwegian disability institutions. Due to the lack of beds, more than 500 people with intellectual disabilities were housed in psychiatric hospitals, and over 1000 were in family care (Fjermeros, 2009, 180). However, in the 1950s, Sweden followed the same model as Denmark and began to develop a system of large institutions. The expansion of the disability care system was part of the development of the Swedish welfare state. The period from the 1950s to the late 1970s is often thought of as the golden years of the Swedish welfare state. Universal income-related social security was established as well as ambitious labor market policies and the expansion of health care and personal social services in the counties and local municipalities (Lindqvist, 2017, 218–219). This also included plans to modernize and upsize the institutional system for people with disabilities. During the period 1960–1970, the number of beds in large disability institutions grew from 5500 to 10,500 (Grunewald, 2009, 367). In Norway, an increase in the number of beds also took place from about 2000 in 1955 to 6550 in 1970. However, half of these beds were situated in homes, small institutions and family care (Fjermeros, 2009, 305). These numbers were rather small compared with those for Sweden and Denmark. In Denmark, with a slightly larger population, approximately 10,000 people were placed in residential institutional care and about 12,000 in family care, small institutions and so on (Fjermeros, 2009, 305–306; Kragh et al., 2015, 260–261).

Despite the Scandinavian preference for large institutions in the 20th century, opposing thoughts about disability care surfaced during this period. Public critique of institutional care became more frequent after 1945, when ideas of human rights emerged. The European Council’s Convention on Fundamental Human Rights (1948) was ratified by the Scandinavian countries in the 1950s, and the rights of people with disabilities subsequently began to be a public issue. Equally important was the emergence of the first parent associations in Scandinavia. In 1952, the Danish association LEV [Danish Society for the Welfare of the Intellectually Disabled] was founded, and, in 1956, the Swedish parent association “Föreningen för utvecklingsstörda barn” [FUB] (Swedish Association for Children, Youths and Adults with Developmental Disorders) was established (Kirkebak, 2001; Grunewald, 2009, 215). Similar associations emerged in Norway in the mid-1960s. These new associations would play a key role in the fight for political and administrative reforms for disabled people.

In 1946, the concept of normalization was used for the first time in Sweden by the Commission for People with Partial Work Ability. The commission’s task was to investigate and propose measures to make better use of the work ability of the partially disabled so that they could obtain or keep gainful employment (Hollander, 1999; Lindqvist, 2017, 219). This idea about normalization would be developed further in Denmark by a commission that was set up in 1954 to reform the services for the so-called feeble-minded (Kirkebak, 2001). With the Danish Act No. 192 of June 5, 1959, the normalization principle was introduced in practice. This act granted power to the Danish National Service for the Feeble-minded to provide people with learning disabilities with living conditions as close to those of others in society as possible (Kirkebak, 2001). Self-determination for people with disabilities also became an important issue. In Sweden, a new disability act was implemented in 1967. By the late 1960s, the normalization principle gained further ground through the work of physician Bengt Nirje (Hollander, 1999; Perrin, 1999; Grunewald, 2009, 219–220). In the same period, the normalization principle also held a strong position in Norway (Kristiansen, 1999, 397–398). This would later lead to a showdown with the exclusive laws for people with disabilities and the large disability institutions.

New perspectives: empowerment, patients' voices and critique after 1970

During the late 1960s, a medical counterculture began to emerge, in which medical science and its practitioners came under increasing scrutiny. The counterculture that emerged from the youth movements of the day also brought with it a critical reappraisal of the role of medical science in the treatment of groups that deviate from society's norms. Scandinavian left-wing intellectuals wrote several articles on psychiatry, inspired by international thinkers such as Ronald D. Laing and David Cooper and their notion of "anti-psychiatry" (Crossley, 2005; Jacobsen and Larsen, 2007, 444–445; Haave, 2008, 396–397; Ohlsson, 2008).

However, this critical focus on the conditions faced by people with mental disorders or intellectual disabilities and other groups of people who were placed under state care was not only a consequence of the late 1960s youth counterculture. Within psychiatry, alternative approaches had already been emerging, like Maxwell Jones's therapeutic communities and other currents that prioritized social–psychiatric and psychological theories of mental disorder (Whiteley, 2004; Haave, 2008, 375–379, 493–494; Kritsotaki et al., 2019). The legislation covering this area in the late 1950s represented the start of a period of major reform, and the idea of normalization became further entrenched in the following decade. In Denmark, the National Service for the Feeble-Minded decided to phase out two of the most controversial institutions, those on Livø and Sprogø. Since the start of the century, these two otherwise uninhabited islands had been used to isolate "antisocial" mentally disabled young men and women. International critique of the "total institutions" increased in the late 1970s, when community mental health care became a central issue. In the new Danish plans for psychiatry in 1977, a commission under the Ministry of the Interior stressed that the hospital system should be based on community mental health care (Bjerrum, 2008). Management of the state mental hospitals was handed over to the counties and municipalities. However, when this decision was made in 1976, the psychiatric institutions still had a substantial number of beds, more than 11,000. However, this system changed during the 1980s, when treatment services were transferred to the new community care units, outpatient facilities and psychiatrists in private practice and so on. During the period 1980–1990, the number of beds in psychiatric hospitals was reduced by more than 4000. This development continued in the years that followed. An analogous deinstitutionalization process took place in the service for people with intellectual disabilities. In 1980, about 12,000 beds were available in Swedish residential institutions (*asyler/vårdhem*); in 1990, this number had dropped to 5000, and 10 years later it was 200 (Grunewald, 2009, 367).

In the 1960s and 1970s, patient associations had a crucial role not only in altering the former disability care system but also in fostering the empowerment of people with disabilities. The Swedish association FUB was the first to give people with disabilities the option to participate in the development of its policies. At a seminar in Stockholm in 1985, these issues were discussed for the first time with participants from Scandinavia and Finland. A specific project entitled "Our right to decide" (*Vår rätt att bestämma*) was developed by the NFPU, and it was a central theme at the annual conferences held in the following years (Bylov, 2010, 453–463). These conferences served as inspiration for the development of organizations such as the Danish ULF, founded in 1993, exclusively for persons with intellectual disabilities (Bylov, 2010).

The establishment of special education

Throughout the 19th century, many local and private initiatives led to the establishment of schools and education for children who were regarded as disabled. These initiatives could be termed early efforts toward special needs education since they featured teaching for children with disabilities. In developing these initiatives, professionals exchanged ideas and practices through visits and networks in Scandinavian countries.

Children were regarded as having the greatest potential and securing the best results in the light of usefulness for society. Therefore, all three countries established teaching for deaf–mute children at the beginning of the 19th century (Froestad, 1995, 83). This was pioneering work in all three countries, with no established professional networks or common policies across the Nordic countries supporting the work, which gave space for local initiatives and practices. These processes contained great variations in the countries, but, in all three cases, the state, in its policy for schools in general, also influenced the practice of deaf–mute education (Froestad, 1995; Coninck-Smith, 2000).

Whereas the first part of the 19th century was characterized by a pragmatic and practical approach to the education of deaf–mute people, ideological conflicts on the organization of the teaching and the question of the differentiation of students occurred in the later part of the century (Froestad, 1995, 327). The practice of teaching disabled people was influenced by academics putting forward different ideological positions in the different national contexts. The pedagogy for people with disabilities was not only developed through the education sector; other sectors, like that of social policy, also contributed to this process. The care and education of disabled people developed through initiatives from different professional actors, like teachers, philanthropists, clergymen and doctors.

Pedagogy played an important role in the institutionalization of people with disabilities in the 19th century, whether this took the form of schools directed toward different groups of people with disabilities or care institutions (Kirkebæk, 1993; Froestad, 1995). Pedagogy was debated among the professionals organizing these first initiatives, and there was an exchange as well as battles concerning the right way to teach. It was believed that pedagogy along with medical support could cure feeble-mindedness or at least make the individual fit to participate in the duties of the society. This reflected the ideal of the Enlightenment that individuals should be useful to society. The utopian hopes of education and cure had been abandoned by the end of the 19th century (Hansen, 2005). "The teacher and the amateur philanthropist became subordinate to the physician, the expert

who could classify the mentally retarded and determine the extent of mental retardation, and subsequently the amount of instruction required" (Hansen, 2005, 15).

Later, the "abnormal case" [Abnormssagen] in Nordic countries developed despite great differences as a form of joint project through Nordic meetings, joint journals and the exchange of ideas and practices. From 1876, special congresses for teachers within special education took place in the Scandinavian countries (Andersen and Stevens, 1957, 253). These joint initiatives across fields and borders contributed to the development of national values and practices within education and care for people with disability, leading to shared interpretations in Sweden, Norway and Denmark.

The Special Teachers College in Slagsta, close to Stockholm, played an important role in the training of Nordic teachers within the field of special education. In Norway, fellowships were available from the government for teachers to visit Sweden and Denmark to learn from those countries' experiences of programs for individuals who were regarded as mentally retarded.

National cases

In Denmark, a number of initiatives were launched to educate pupils who were perceived as disabled: the Deaf Institute was established in 1807, the first Danish Institute for the Blind in Copenhagen was founded in 1811 and the first Danish institution for the mentally handicapped/feeble-minded was created in 1855. Compulsory education for children who were perceived as abnormal was introduced in 1814. However, until the middle of the century, many mentally retarded individuals were still placed in the local communities without being accounted for in any way in the school system. There were also children with disabilities in public schools in the early 1800s, as evidenced by school records from the period. These were students who were deaf, were blind, had physical disabilities or were perceived as mentally retarded (Larsen et al., 2013, 262). Many children with disabilities were in larger care institutions, but, by the middle of the 19th century, these children increasingly became an object of educational interest (Larsen et al., 2013, 295). Children who deviated from the norm became increasingly visible in normal schools. Auxiliary schools and classes were established for these children. In 1900, the first auxiliary class was introduced, and, in 1913, an auxiliary school was founded in Copenhagen (Larsen et al., 2013, 305). It was not until 1937 that a law on special education was established (Egelund, 2000; Hamre and Ydesen, 2014). The law stated that municipalities were obliged to provide special education in cases in which the conditions provided the possibility for it. Feeble-mindedness caused special awareness among the Danish school authorities. In 1943, the Circular on Special Education stated: "For children recognized as feeble-minded/mentally handicapped, the Folkeskolen's [regular school] auxiliary classes have no place" (Hamre, 2012, 122–123, our translation). These schoolchildren should instead be referred to the special institutions for the mentally retarded. In the debate among the Danish teachers, the sorting of the children into special schools or classes was defined with progressive notions of individualization and differentiation, whereas those who were assessed as feeble-minded were defined with the notion of exclusion (Hamre, 2012).

In Norway, one of the former pupils of the Copenhagen Deaf institution, Andreas Christian Møller, created a Norwegian boarding school for people with hearing disabilities in Trondheim in 1825 (Eriksson, 1992; Frostad, 1995; Simonsen, 2000; Tønnesen, 2011). The advent of schools for the blind followed in 1861 in Norway (Kaimer, 2002; Ruby, 2002, 2011). In the Scandinavian case, Norway was the first country to adopt an act directed to students who were considered disabled, which was also unique in the European context (Froestad, 1995). The Abnormal Schools Act of 1881 introduced compulsory schooling for the deaf-mute, blind and mentally handicapped (Froestad, 1995, 337). The Ministry of Church and Education, in line with the legislation, appointed an attendant functioning as a director of "Schools for the Abnormal" in 1897. In 1915, a new version of the 1881 law was passed that including low-performing schoolchildren. Until 1949, the Ministry was responsible for special schools as well as for institutions for the so-called mentally retarded. From 1949 onward, these institutions for the mentally retarded were taken over by the Ministry of Social Welfare, whereas special schools continued to belong to the Ministry of Church and Education. In 1951, the parliament passed a new law, the Special School Act of 1951, on special schools addressing schoolchildren who could not attend regular schools. The new law was a further development of the 1881 law, aiming to provide education for schoolchildren who were regarded as "deaf, blind, backward (mentally retarded), maladjusted, and children with speech defects" aged between 8 and 21 (Andersen and Stevens, 1957, 252). The practice of abnormal schools continued after WWII, but the schools were hereafter titled special schools (Tønnesen, 2011).

In Sweden, an institute for the deaf and blind was established in 1808 in Stockholm (Grunewald, 2009, 61). An elementary school act was adopted in 1842, but it was to be run on a local basis (Westberg, 2019, 202). It has been argued that special education started in connection to this act (Wickman and Kristofferson, 2017), though many schoolchildren in practice were excluded from schooling. As elsewhere in Scandinavia, schoolchildren with learning disabilities were perceived as a problem for education in regular schools as well as for the teacher and for the whole class. Regular schooling had been compulsory for deaf schoolchildren since 1889 and for blind children since 1896. In 1944, a law made schooling compulsory for children who were regarded as "educable mentally retarded," and, in 1967, this law was revised to include children who were considered to be "profoundly mentally retarded" (Walton et al., 1990, 115). For children who were perceived as feeble-minded, school homes were established, the first one being in Stockholm in 1872, and the largest was established in Slagsta in 1911 (Grunewald, 2009, 55). The school homes developed educational practices for the feeble-minded and were isolated from the primary school and the auxiliary classes. By the 1960s, the school homes had closed and had been replaced by special classes (Grunewald, 2009, 315). According to Grunewald, the optimistic perception of feeble-mindedness changed due to eugenics theories, which meant that teaching, by the beginning of the 20th century, turned into training in practical skills (Grunewald, 2009, 321). In 1905, the first auxiliary class was established in Stockholm, and, a few years later, classes had been established in 15 cities (Grunewald, 2009, 334). In 1954, an

act was adopted stating that special education should take place primarily within primary schools. The term “special school” [særskole] was applied from 1955, and the school homes as boarding schools were abandoned, special classes in primary schools becoming the new practice (Grunewald, 2009, 339).

IQ testing, progressivism and professional demarcation

Educational testing became an important feature of the welfare societies in supporting a political interest in differentiating the future workforce and selecting the part of the population that was regarded as the most talented (Axelsson, 2007). Like elsewhere in Europe, eugenics and psychometric testing played a major role in Scandinavia in the sorting of students in the schooling system (Ydesen, 2011; Ydesen et al., 2013). Proponents of eugenics, for example, linked feeble-mindedness with bad morals, and this contributed to legitimizing the new exclusion of students to special institutions. The importing of psychometric testing from France and the USA to the Scandinavian countries in the 1920s and 1930s led to the sorting of students being legitimized by scientific practices (Axelsson, 2007; Ydesen, 2011; Ydesen et al., 2013). In the interwar period onward, IQ testing became the most important prerequisite for sorting the pupils in public schools. Teachers, educational psychologists and psychiatrists as well as progressive educationalists in Scandinavia were proponents of testing. In their view, testing was a practical tool to solve the problem of differentiation in schools and, according to progressive educationalists, a way to accommodate the individual teaching of children. There was an alignment between proponents of testing and progressive educationalists in the interwar years, and it does not make sense to distinguish between the two groups in practice (Nørgaard, 1977; Bendixen, 2006). The proponents regarded intelligence testing as the common ground to overcome limitations in examinations.

In the interwar years, the international progressive education organization New Education Fellowship (NEF) arranged meetings concerning child-centered education. Its largest conference took place in Elsinore in Denmark in 1929. The ministers of education from Sweden, Norway and Denmark were responsible for the initiative, regarded as being of the utmost importance for the educational field (Ydesen et al., 2013, 120). Mental tests were a prominent issue at the conference. Testing was a valuable technology for Scandinavian educators since it could assess the quality of the students as well as of the schools (Ydesen et al., 2013). Child-centered progressivists shared the same values with the intention to individualize and put the child at the forefront of education.

Generally, there were challenges in the assessment of children, primarily the questions of who had the authority to practice the testing and how children's problems should be interpreted. Dilemmas arose in terms of the professional demarcation between the two disciplines and the issue of authority to examine. Like child psychology, child psychiatry gained influence in the Scandinavian countries in the interwar years (Hamre et al., 2019). As in England, child guidance clinics emerged in Denmark, Sweden and Norway. Psychiatrists participated in the examination of pupils and thus in the process of referral to auxiliary and special schools. The Scandinavian experts initiated counseling activities based on a mixture of ideas drawn from mental hygiene, psychoanalysis and child psychology (Hamre et al., 2019). Mental hygiene clinics and school health services became important arenas for sorting schoolchildren. The mental hygiene movement has been regarded as a bridge between the biologically oriented hygienist of the early 20th century and the sociopolitical reforms of the interwar period (Axelsson, 2007; Stewart, 2011). Scandinavian experts attended the First International Congress on Mental Hygiene in 1930, and this led to the establishment of local mental hygiene associations. Professional attention increasingly turned away from eugenics and toward the context of upbringing in the family, the local community and educational institutions. In this process, local mental health movements played a crucial role in establishing child guidance clinics, outpatient treatment and child psychiatric wards in hospitals (Hamre et al., 2019). All three countries experienced a strong influence from psychoanalysis in the 1930s, 1940s and 1950s, and this contributed to the importance of environmental explanations, a process that was also enforced by the mental hygiene movement (Hamre et al., 2019). At the beginning of the 1950s, the three Scandinavian countries established child psychiatry as a university field.

The Scandinavian school system in the years after 1945

The period 1945–1970 has been titled “the golden age” of the welfare states in Scandinavia (Petersen, 2006). The ideal of equality has been regarded since the 1960s as an important tool to equalize differences in backgrounds and social inheritance. In this process, the schooling system had the most effective approach to mitigating social inequity. It underwent a modernization process in which the social democratic parties were the key actors. “Education for all children was also considered to be the main vehicle for reducing social differences and increasing social mobility in the population” (Imsen et al., 2017). The structure of schools was one of comprehensive schools without streaming. The state had a strong influence on the curriculum, and progressivism as a pedagogical mindset was related to the overall ideas of democratization, equity and social justice. “The comprehensive school was in itself seen as a stage in pupils’ democratic socialization as it placed them within a community where all classes of society would meet” (Telhaug et al., 2006, 253). The idea of a comprehensive school was very similar in Sweden, Denmark and Norway and experienced similar historical development, characterized by a strong influence from the social democratic parties (Imsen et al., 2017). In the years to come, the comprehensive schooling system was gradually revised, though the general idea remained.

Concerning special education, there was an increasing critique of the segregation into special needs education and the general conditions for children with disabilities. In Sweden, for example, the number of pupils in special schools doubled in the period 1967–1977 (Grunewald, 2009, 347). The critique proposed three ideals for the Scandinavian schooling system: normalization,

integration and decentralization (Walton et al., 1990). In the subsequent years, normalization was subject to more or less the same critique (Holloway and Hamre, 2018). In 1994, all three countries ratified the Salamanca Treatment, thus obliging themselves to embrace inclusion as the overall principal in schooling.

Conclusion

This article on the historical development of disability, mental disorders and special education in Denmark, Norway and Sweden took as its point of departure the practices of differentiation, the ideas formed by the Enlightenment that subjects are educable and the need to use all human resources in the modernization of society. The main question in the 19th century concerned how to differentiate the broad groups of individuals who did not conform to the ideals of the educable subject. Different institutional settings took form, whether they were for disabled people, for those with mental disorders or educational initiatives. This modernization process involved dilemmas between belief in the educability of the subject and ideas of social Darwinism, degeneration and eugenics.

The Scandinavian combination of the dominance of social democratic parties by the beginning of the 20th century and the development of eugenic policies formed practices targeting people with disabilities, especially those regarded as feeble-minded, thereby legitimizing practices of exclusion and sterilization. Scholars have questioned why liberal societies like those in Scandinavia could develop illiberal practices like these. Some of these practices were legitimized through eugenics, and others, like that of lobotomy within psychiatry, were founded on medical arguments as well as economic rationalities. The question is whether these historical practices developed despite the new ideals of the welfare state or whether these illiberal practices simply worked as a prerequisite for the functioning of the welfare state. Similarly, the development of special education was a practice that developed from existing exclusion practices that were enforced by the importing and use of IQ testing and assessments from different professionals.

Acknowledgment

Thanks to Christian Ydesen for valuable comments on the article.

References

- Adserballe, H., 1977. Frihedsberøvelser og tvang i psykiatrien. FADL's Forlag, Copenhagen.
- Andersen, B.H., Stevens, G.D., 1957. A review of the education of the mentally handicapped in Norway. *Except. Child.*
- Axelsson, T., 2007. Rätt elev i rätt klass: skola, begåvning och styrning 1910–1950. Tema Barn, Linköpings universitet.
- Barow, T., 2009. Kein Platz im Volksheim? Die "Schwachsinnigenfürsorge" in Schweden 1916–1945. Verlag Julius Klinkhardt, Bad Heilbrunn.
- Bejerot, S., 2004. Psykokirurgi i dag—en kritisk betraktelse. *Läkertidningen* (32–33), 2502–2505.
- Bendixen, C., 2006. Psykologiske teorier om intelligens og folkeskolens elevdifferentiering. Roskilde Universitetscenter, Roskilde.
- Ben-Moshe, L., Chapman, C., Carey, A.C., 2014. Disability Incarcerated: Imprisonment and Disability in the United States and Canada. Palgrave Macmillan, New York.
- Berg, A., 2018. De samhällsbevärliga. Förhandlingar om psykopati och kverulans i 1930- och 40-talens Sverige. Makadam Förlag, Göteborg.
- Bergheim, Å., 2014. Sexual assault, irresistible impulses, and forensic psychiatry in Sweden. *Int. J. Law Psychiatr.* 37, 99–108.
- Betänkning, 1956. Afgivet af kommissionen af 29. marts 1952 vedrørende Statens sindssygevesen. Schultz, Copenhagen.
- Bjerrum, M., 2008. Fra stat til amt. Dansk hospitalspsykiatri 1930–1976. In: Kragh, J.V. (Ed.), *Psykiatriens historie i Danmark*. Hans Reitzels Forlag, Copenhagen, pp. 222–262.
- Börjesson, M., 1994. Sanningen om brottslingen: Rättspsykiatri som kartläggning av Livsöden. Carlsson, Stockholm.
- Bylov, F., 2010. Den store løsladelse : en kulturhistorisk beretning om empowerment-bevægelserne blandt udviklingshæmmede 1980-95 : Det kulturelle opbrud. University College Syddanmark, Esbjerg.
- Christiansen, N.F., Markkola, P., 2006. Introduction. In: Christiansen, N.F., Petersen, K., Edling, N., Haave, P. (Eds.), *The Nordic model of welfare: a historical reappraisal*. Museum Tusulanum, København.
- Coninck-Smith, N.D., 2000. For barnets skyld—Byen, skolen og barndommen 1880–1914. Gyldendal, Copenhagen.
- Crossley, N., 2005. Contesting Psychiatry. Social Movements in Mental Health. Routledge, London.
- Dahl, H., 2018. Institusjonaliseringen av farlighet, sinnssykdom og samfunnsvern i Norge En studie av farlige kriminelle sinnssyke pasienter i Kriminalasylet og Reitgjerdet asyl, 1895-1940 Avhandling for graden philosophiae doctor Trondheim, september 2018. Norges teknisk-naturvitenskapelige universitet Det humanistiske fakultet Institutt for historiske studier, Trondheim.
- Dahl, H., 2018a. Institusjonalisering av farlighet, sinnssykdom og samfunnsvern i Norge. En studie av farlige kriminelle sinnssyke pasienter i Kriminalasylet og Reitgjerdet asyl, 1895–1940. PhD-Dissertation. University of Trondheim, NTNU.
- Dahl, H., 2018b. "Insane criminals" and the "criminally insane": criminal asylums in Norway, 1895–1940. *Hist. Psychiatr.* 28, 209–224.
- Dean, M., 2010. Governmentality: Power and Rule in Modern Society. SAGE Publications, London.
- Duedahl, P., 2019. En anden verden. De Kellerske Anstalters historie. Gads Forlag, Copenhagen.
- Egelund, N., 2000. Country briefing: special education in Denmark. *Eur. J. Spec. Needs Educ.* 15 (1), 88–98.
- Eivergård, M., 2003. Frihetens milda disciplin. Normalisering och social styrning i svensk sinnessjukvård 1850–1970. Umeå University, Umeå.
- Engwall, K., 2005. Starved to death? Nutrition in asylums during the world wars. *Scand. J. Disabil. Res.* 7, 2–22.
- Eriksson, B.E., 1989. Vägen till centralhospitalet: Två studier av den anstaltsbundna sinnessjukvårdens förhistoria i Sverige. Daidalos, Göteborg.
- Eriksson, P., 1992. Dövas historia. Bokverkstan Förlag, Örebro.
- Fause, Å., 2015. The fate of the mentally ill during the Second World War (1940–1945) in Troms and Finnmark, Norway. *Nordlit* 37, 1–29.
- Fjermøer, H., 2009. Åndssvak! Et bidrag til sentralinstitusjonens og åndssvakeomsorgens historie. Universitetsforlaget, Oslo.
- Foucault, M., 2009. Security, Territory, Population: Lectures at the Collège de France, 1977–1978. Palgrave Macmillan, New York.

- Freestad, J., 1995. Faglige diskurser, intersektorielle premisser og variasjoner i offentlig politikk. In: Døveundervisning og handicapomsorg i Skandinavia på 1800-tallet [Professional Discourses, Intersectoral Premises and Differences in Public Policy. Educating the Deaf and Caring for the Disabled in Scandinavia in the 17th Century]. Department of Administration and Organization Theory, University of Bergen. Dissertation.
- Garpenhag, L., 2012. Kriminaldådar: sinnssjuka brottslingar och straffrihet i Sverige, ca. 1850–1930. Uppsala Universitet, Uppsala.
- Gelman, S., 1999. *Medicating Schizophrenia: A History*. Rutgers University Press, New Brunswick, NJ.
- Goodey, C.F., 2011. A History of Intelligence and “Intellectual Disability”: The Shaping of Psychology in Early Modern Europe. Ashgate.
- Graugaard, C., 1997. Professor Sands høns – om seksualbiologi i mellemkrigstidens Danmark, ph.d.-afhandling. Københavns Universitet, København.
- Grunewald, K., 2009. *Från idiot till medborgare : de utvecklingsstördas historia*. Gothia, Stockholm.
- Haave, P., 2008. Ambisjon og handling: Sanderud Sykehus og norsk psykiatri i et historisk perspektiv. Unipub Forlag, Oslo.
- Hamre, B.F., Ydesen, C., 2014. The ascent of educational psychology in Denmark in the interwar years. *Nordic J. Educ. Hist.* 1 (2), 87–111.
- Hamre, B.F., Axelsson, T., Ludvigsen, K., 2019. Psychiatry in the sorting of schoolchildren in Scandinavia 1920–1950: IQ testing, child guidance clinics, and hospitalization. *Paedag. Hist.* 55, 1–25.
- Hamre, B., 2010. Disciplinary power and the role of the subject at a nineteenth-century Danish asylum. *Phaenex* 5, 1–27.
- Hamre, B., 2012. Potentialitet og optimering i skolen—Problemforståelser og forskellssætninger af elever—en nutidshistorisk analyse. Institut for Uddannelse og Pædagogik. Aarhus Universitet, Aarhus.
- Hansen, B.S., 2005. Something rotten in the state of Denmark : eugenics and the ascent of the welfare state. In: Broberg, G., Roll-Hansen, N. (Eds.), *Eugenics and the Welfare State: Sterilization Policy in Denmark, Sweden, Norway and Finland*. Michigan State University Press.
- Have, P., 2000. Sterilisering av tatere 1934–1977. En historisk undersøkelse av lov og praksis. Norges forskningsråd, Oslo.
- Healy, D., 1997. *The Antidepressant Era*. Harvard University Press, Cambridge, MA.
- Healy, D., 2002. *The Creation of Psychopharmacology*. Harvard University Press, Cambridge, MA.
- Hollander, A., 1999. The origin of the normalization principle in Sweden and its impact on legislation today. In: Flynn, R.J. (Ed.), *A Quarter-Century of Normalization and Social Role Valorization: Evolution and Impact*. University of Ottawa Press, Ottawa, pp. 407–410.
- Holloway, J., Hamre, B.F., 2018. School reforms, market logic, and the politics of inclusion in the USA and Denmark, I. In: Hamre, Morin, Ydesen (Eds.), *Testing and Inclusive Schooling—International Challenges and Opportunities*. Routledge, London.
- Hübertz, J.R., 1843. *Om Daarevæsenets Indretning i Danmark*. Trier, Copenhagen.
- Hughes, B., 2020. *A Historical Sociology of Disability*. Routledge Advances in Disability Studies, London, New York.
- Imsen, G., Blossing, U., Moos, L., 2017. Reshaping the Nordic education model in an era of efficiency. Changes in the comprehensive school project in Denmark, Norway, and Sweden since the millennium. *Scand. J. Educ. Res.* 61 (5), 568–583. <https://doi.org/10.1080/00313831.2016.1172502>.
- Jacobsen, K., Larsen, K., 2007. *Ve og velfærd. Læger, sundhed og samfund gennem 200 år*. FADLs Forlag, Copenhagen.
- Jönsson, L.E., 1998. *Det terapeutiska rummet: rum och krop i svensk sinnessjukvård 1850–1970*. Carlsson, Stockholm.
- Kaimar, G., 2002. *Blindhedens verden, træk af blindes kulturhistorie*. Klim, Århus.
- Kirkebæk, B., 1993. *Da de åndssvage blev farlige*. Forlaget SocPol, Holte.
- Kirkebæk, B., 1997. *Defekt og deporteret. Ø-anstalten Livø 1911–1961*. Forlaget SocPol, Holte.
- Kirkebæk, B., 2001. *Normaliseringsperioden, dansk åndssvageforsorg 1940–1970 med særligt fokus på forsorgschef N.E. Bank-Mikkelsen og udviklingen af statens åndssvageforsorg 1959–1970*. Forlaget SocPol, Holte.
- Kirkebæk, B., 2004. *Lettærdig og løssagtig. Kvindeanstalten Sprogø 1923/1961*. Forlaget SocPol, Holte.
- Knorr, L.V., 2012. History of the Nordic psychiatric cooperation. *Nord. J. Psychiatr.* 66 (Suppl. 1), 54–60.
- Koch, L., 1996. *Racehygiejne i Danmark 1920–1956*. Gyldendal, København.
- Koch, L., 2000. *Tvangsterilisering i Danmark 1929–67*. Gyldendal, København.
- Koch, L., 2006. *Eugenic Sterilisation in Scandinavia*. The European Legacy 11 (3), 299–309.
- Korsgaard, O., 2008. *The Struggle for the People*. Aarhus Universitetsforlag, Aarhus.
- Kragh, J.V., Jensen, S.G., Knage Rasmussen, J., 2015. *Anbragt i historie(n): Et socialhistorisk projekt om anbragte og indlagte i dansk socialforsorg 1945–1980*. Socialstyrelsen, Copenhagen.
- Kragh, J.V., 2021. *Lobotomy Nation: The History of Psychosurgery and Psychiatry in Denmark*. Palgrave Macmillan, London.
- Kragh, J.V., Møllerhøj, J. (Eds.), 2016. *Sct. Hans 1816–2016. Psykiatrisk Center Sct. Hans, Roskilde*.
- Kragh, J.V. (Ed.), 2008. *Psykiatriens historie i Danmark*. Hans Reitzels Forlag, Copenhagen.
- Kragh, J.V., 2010a. Malaria fever therapy for general paralysis of the insane in Denmark. *Hist. Psychiatr.* 21, 471–486.
- Kragh, J.V., 2010b. Shock therapy in Danish psychiatry. *Med. Hist.* 54, 341–364.
- Kragh, J.V., 2010c. Det hvide snit. Psykiatriske og dansk psykiatri 1922–1983. University Press of Southern Zealand, Odense.
- Kragh, J.V., 2018. *Sikringen—Forbrydelse, farlighed og foranstaltninger: Sikringsafdelingen i dansk psykiatri 1918–2018*. University Press of Southern Zealand, Odense.
- Kragh, J.V., 2019. The final step: the issue of euthanasia of people with mental disabilities in Denmark. *Disabil. Soc.* 34, 143–161.
- Kringlen, E., 2012. History of Norwegian psychiatry. *Nordic Journal of Psychiatry* 66 (1), 31–41. <https://doi.org/10.3109/08039488.2011.638726>.
- Kristiansen, K., 1999. The impact of normalization and social role valorization in Scandinavia. In: Flynn, R.J. (Ed.), *A Quarter-Century of Normalization and Social Role Valorization: Evolution and Impact*. University of Ottawa Press, Ottawa, pp. 395–406.
- Kritsotaki, D., Long, V., Smith, M., 2019. *Preventing Mental Illness. Past, Present and Future*. Palgrave, London.
- Larsen, C., Nørre, E., Sonne, P., 2013. *Da skolen tog form: 1780–1850*. Dansk Skolehistorie, vol. 2. Aarhus Universitetsforlag, Aarhus.
- Lindqvist, R., 2017. Swedish disability policies: ideas, values and practices in a historical perspective. In: Hanes, R., Brow, I., Hansen, N.E. (Eds.), *Routledge Handbook of Disability History*. Routledge, London, pp. 216–230.
- Merrick, J., Uldal, P., Volther, J., 2014. Intellectual and developmental disabilities: Denmark, normalization and de-institutionalization. *Front. Public Health* 2, 1–3.
- Myhre, E., 2016. *Farlige menn. Mannlighet, seksualforbrytelse og sinnssykdom 1895–1940*. PhD Dissertation. University of Trondheim, NTNU.
- Nilsson, R., Vallström, M. (Eds.), 2016. *Inspärrad: roster fra intagne på sinnessjukhus, fängslar och andre anstalter 1850–1992*. Nordic Academic Press, Lund.
- Nørgaard, E., 1977. *Lille barn, hvis er du? en skolehistorisk undersøgelse over reformbestrebelse inden for den danske folkeskole i mellemkrigstiden*. Gyldendal, Copenhagen.
- Ögren, K., 2005. Psychosurgery in Sweden 1944–1964. *J. Hist. Neurosci.* 14, 353–367.
- Ohlsson, A., 2008. *Myt och manipulation: Radikal psykiatrikritik i svensk offentlig idédebatt 1968–1973*. Acta Universitatis Stockholmiensis, Stockholm.
- Perrin, B., 1999. The original “Scandinavian” normalization principle and its continuing relevance for the 1990s. In: Flynn, R.J. (Ed.), *A Quarter-Century of Normalization and Social Role Valorization: Evolution and Impact*. University of Ottawa Press, Ottawa, pp. 181–196.
- Petersen, K., 2006. Constructing Nordic welfare? Nordic social political cooperation 1919–1955. In: Christiansen, N.F., Petersen, K., Edling, N., Have, P. (Eds.), *The Nordic model of welfare: a historical reappraisal*. Museum Tusculanum, København.
- Popkewitz, T.S., Lindblad, S., 2004. Historicizing the future: educational reform, systems of reason, and the making of children who are the future citizens. *J. Educ. Change* 5 (3), 229–247.
- Quarsell, R., 1982. *Ordning och behandling: psykiatri och sinnessjukvård i Sverige under 1800-talets första hälft*. University of Umeå, Umeå.
- Rabben, M.B., 2020. *Humanitet, kontroll og tvang. Tvangsmiddelbruk ved Kriminalasylet og Reitgjerdet sykehus, 1895–1978*. PhD Dissertation. University of Trondheim, NTNU.
- Rønn, E.M., 1996. *De fattige i ånden. Essays om kultur, normalitet og uformu*. Museum Tusculanum Forlag, Copenhagen.

- Roll-Hansen, 2006. In: Broberg, G., Roll-Hansen, N. (Eds.), *Eugenics and the Welfare State: Sterilization Policy In Denmark, Sweden, Norway and Finland*. Michigan State University Press, 2006.
- Rörbecker, S., 2010. Risikabel hjärnkirurgi i Umeå. *Expressen*. <https://www.expressen.se/halsoliv/riskabel-hjarnkirurgi-i-umea/>.
- Ruby, R., 2002. Konstruktionen af den blinde ca. 1850–1914. *Handicaphistorisk tidsskrift* 12, 7–166.
- Ruby, R., 2011. *Afvigelsens dilemmaer, blindes organisering i det 20. århundrede*. PhD Dissertation. University of Copenhagen.
- Seeman, M., 2006. Starvation in psychiatric institutions in Sweden. *Int. J. Ment. Health* 35, 81–87.
- Shorter, E., 2008. *Before Prozac. The Troubled History of Mood Disorders*. Oxford University Press, New York.
- Simonsen, E., 2000. Vitenskap og profesjonskamp. Opplæring av døve og åndssvake i Norge 1881–1963. Unipub, Oslo.
- Sjöström, B., 1992. *Kliniken tar över dårskapen. Om den moderne psykiatris fremåxt*. Daidalos, Göteborg.
- Skålevåg, S.A., 2016. *Tilregnelighet—en historie om rett og medisin*. Pax, Oslo.
- Stewart, J., 2011. "The dangerous age of childhood": child guidance and the "normal" child in great Britain, 1920–1950. *Paedagog. Hist.* 47 (6).
- Stiker, H.-J., 1997. *A History of Disability*. The University of Michigan Press, Ann Arbor.
- Strömberg, E., 1965. De danske tværnsundersøgelser og psykofarmaka. *Nord. Psykiatr. Tidsskr.* 19 (1965), 101–114.
- Swazey, J., 1974. *Chlorpromazine in Psychiatry. A Study of Therapeutic Innovation*. MIT Press, Cambridge, Mass.
- Telhaug, A.O., Mediås, O.A., Aasen, P., 2006. The Nordic Model in Education: education as part of the political system in the last 50 years. *Scand. J. Educ. Res.* 50 (3), 245–283.
- Tranøy, J., Blomberg, W., 2005. Lobotomy in Norwegian psychiatry. *Hist. Psychiatr.* 16, 107–110.
- Tröhler, D., 2020. Nation-state, education and the fabrication of National-Minded Citizens (introduction). *Croat. J. Educ.—Hrvatski časopis za odgoj i obrazovanje*. <https://doi.org/10.15516/cje.v22i0.4129> (Bd. 22).
- Tønnesen, L.K.B., 2011. *Norsk utdanningshistorie: en innføring med fokus på grunnskolen utvikling*. Fagbokforlaget, Bergen.
- Tydén, M., 2002. *Från politik till praktik: De svenska steriliseringslagarna 1935–1975*. Acta Universitatis Stockholmiensis Stockholm Studies in History.
- Tydén, M., 2010. The Scandinavian States: Reformed Eugenics Applied. In: Bashford, A., Levine, P. (Eds.), *The Oxford handbook of the history of eugenics*. <https://doi.org/10.1093/oxfordhb/9780195373141.013.002>.
- Valenstein, E.S., 1986. *Great and Desperate Cures. The Rise and Decline of Psychosurgery and Other Radical Treatments for Mental Illness*. Basic Books, New York.
- Valenstein, E.S., 2001. *Blaming the Brain*. The Free Press, New York.
- Walton, W.T., Emanuelsson, I., Rosenqvist, J., 1990. Normalization and integration of handicapped students into the regular education system: contrasts between Sweden and the United States of America. *Eur. J. Spec. Needs Educ.* 5 (2), 111–125.
- Westberg, J., 2019. Basic schools in each and every parish: the school act of 1842 and the rise of mass schooling in Sweden. In: Westberg, J., Boser, L., Brühwiler, I. (Eds.), *School Acts and the Rise of Mass Schooling: Education Policy in the Long Nineteenth Century*. Palgrave Macmillan, Cham.
- Wetterberg, L., 2012. History of psychiatry in Sweden during a millennium. *Nord. J. Psychiatr.* 66 (Suppl. 1), 42–53.
- Whiteley, S., 2004. The evolution of the therapeutic community. *Psychiatr. Q.* 75, 233–248.
- Wickman, K., Kristoffersson, M., 2017. History and national policy documents on special education in Sweden. In: *The Routledge History of Disability*.
- Ydesen, C., Ludvigsen, K., Lundahl, C., 2013. Creating an educational testing profession in Norway, Sweden, and Denmark, 1910–1960. *Eur. Educ. Res. J.* 12 (1), 120–138.
- Ydesen, C., 2011. *The Rise of High-Stakes Educational Testing in Denmark (1920–1970)*. Peter Lang.
- Zalewski, B., 2008. St. Hans hospital i København 1612–1808. In: Kragh, J.V. (Ed.), *Psykiatriens historie i Danmark*. Hans Reitzels Forlag, Copenhagen, pp. 19–58.

The National Disability Insurance Scheme, education and funding to support the inclusion of students with disabilities across educational sectors

Ben Whitburn, University of Southampton, UK

© 2023 Elsevier Ltd. All rights reserved.

Introduction	446
The NDIS: how does it work?	446
The NDIS and education	447
Funding education: what will the NDIS support?	448
Early childhood education	449
Primary and secondary school	449
Higher education and vocational education and training (VET)	450
Conclusion	450
References	451

Introduction

The purpose of this chapter is to provide a detailed overview of the Australian National Disability Insurance Scheme (NDIS); in particular, the piece explores ways in which by design, the NDIS engages with educational services in support of students with disabilities. The largest social reform to be implemented in Australia since the introduction of Medicare in 1984, which is the scheme by which all residents of the country are provided basic health care, the NDIS comparably signals significant change to disability support policy. The intent behind the scheme is to increase social and economic participation of every Australian living with disabilities, which was estimated in 2018 (ABS, 2019) to be 4.4 million, or 17.7% of the population, by supporting them to connect with services that they need, including medical care, schooling, therapy plans, community services and recreational opportunities. The NDIS is administered by the National Disability Insurance Agency (NDIA), which has statutory responsibility for its administration, in accordance with the [NDIS Act \(2013\)](#), which was hastily put together to guide scheme governance. Estimated to provide \$18–22 billion in GBP by increasing participation of not only people with disabilities into the workforce but also their families and carers ([National Disability Services, 2016](#)), the NDIS has large ambitions behind it for transforming the ways in which people with disabilities are understood, related to, and are anticipated to contribute in the Australian socio-economic landscape, to which an open and inclusive education system is directly relevant.

The chapter proceeds through five sections, each drawing on published research, NDIA reports and factsheets, as well as conceptual disability literature to explore the impact of the NDIS on education, in order to provide an explanation relevant to educators, students with disabilities and their families on the affordances and parameters of the scheme. In the first section, further explanation is provided about the NDIS and, in particular, how it is set up to benefit participants by emphasizing choice and control over funded support, while also building capacity more broadly across the nation for increasing inclusion for all people living with disabilities. In the second section, a deeper analysis is provided about the architecture of the scheme with respect to education, and the ways that its orientation both positions educational inclusion, and depends on broader inclusivity across the sectors, in order to be effective for achieving its ambitious outcomes. The third section then provides details about NDIS affordances to education in the early childhood, primary and secondary, and post-compulsory sectors. This section concentrates in particular on what the NDIS can provide to students who receive funded plans in terms of transitional support, equipment and associated services to support their education, and as well how the scheme might be drawn on to support educators in some instances. In this section the article inevitably touches on reported experiences of scheme participation particularly in and through the Global Covid-19 pandemic. The chapter concludes with a discussion of the contribution that innovative approaches to disability scholarship and activism might make to NDIS participation, to give emphasis to what it could mean to support a student with disability to achieve an equitable education.

The NDIS: how does it work?

The NDIS is a substantial piece of policy reform that was introduced into parts of Australia in 2013 before being rolled out across the nation from 2016. Prompted by a [2011](#) Productivity Commission report which found that disability services at the time were underfunded, inequitable, fragmented and inefficient, the NDIS has rapidly replaced state and territory run disability support with an insurance-based and individualized approach to funding that emphasizes user choice and control. The NDIS is notionally underpinned by a human rights approach to disability policy reforms and aligned with principles of the social model of disability, locating societal structures responsible for posing the foremost barriers to equal participation for people living with impairments rather than their biological conditions ([Oliver, 2009](#)). The NDIS marks a significant shift in disability policy in Australia. It

ostensibly provides greater scope to support people with permanent and significant disabilities from birth to 65 years of age to access support services commensurate with their goals and aspirations, while also building capacity of Australian services to adapt toward a more inclusive future whereby people with disabilities enjoy social and economic parity. Anticipated to fund up to 460 thousand recipients directly through funded care plans (a fraction of the estimated population of Australians disclosing disabilities), the scheme has already exceeded projections in both funded plans and dollar amount (NDIA, 2021a), leading to potential scheme derailment (Olney and Dickinson, 2019).

There are three types of support budgets available to funded participants: core (to support everyday activities); capacity building (to develop independence and skills over time); and capital (to purchase specific equipment for support such as assistive technology, or to make home and vehicle modifications). Person-centred disability funding is not new to some of the States and territories of Australia (Kendrick et al., 2017). However, by formalizing this approach, the NDIS has mobilized individual choice and control as the preferred national provision of disability support, also increasing portability for funded participants. One major difference that this change brings about is that whereas service providers—individuals or organizations that provide services or products to people with disabilities—received direct block funding from governments, they are now pooled into a market from which NDIS participants purchase support that reflect their lifestyle and aspirations. Participants deemed eligible for funded support are provided financial support on the basis that they make clear and achievable goals in their personal plans, against which reasonable and necessary supports are drawn to achieve their aspirations. In so doing, the NDIS aligns disability support in Australia with the United Nations Convention on the Rights of People with Disabilities (CRPDs) (2006) by emphasizing independence for citizens with impairments that affect their participation. Furthermore, the NDIS has reshaped disability support in Australia to a system that is more consumerist by nature, whereby consumers of products and services are included in Australian society in the sense that they are afforded the right to “choice and redress”.

On the surface, promoting autonomy and control for citizens with disabilities who receive NDIS funding represents a welcome upgrade to Australia's system of disability support. Framed in this way in particular, the NDIS is redolent of the current government's favoring inclusion initiatives for otherwise marginalized people on grounds social, economic, and for increasing independence. To this end, people with disabilities who receive funded packages to access support are anticipated to achieve a greater level of equality in their local communities by participating in the marketplace in ways that benefit their short and long-term goals. That being said, to ensure scheme sustainability, participant autonomy is always balanced against what is deemed reasonable and necessary. According to NDIA guidelines (2019), support can be considered reasonable and necessary only if it is related to a person's disability; is not simply responding to a day-to-day cost of living (such as rent or utility bills); is cost effective; is responsive to a persons' identified needs; and is balanced against supports already provided from other governmental services and a persons' informal networks such as carers, families and friends.

The elusiveness of the measures by which reasonable and necessary support is evaluated by the NDIA has been taken up by disability researchers, who have registered their concerns that inconsistencies emerge in their interpretation. To them, gaps between the promises of the NDIS and participants' lived experiences are a hallmark of the NDIS in its current form, and potentially into the future. Cost-benefit analysis and discretionary decision making on the part of planners tend to outweigh the aspirations of participants (Carey, Malbon, and Blackwell, 2020; Olney and Dickinson, 2019; Thill, 2015), and as Australia enters an unstable economic future, scrutiny of what reasonableness and necessity is will be likely further intensified (Kendrick et al., 2017). Indeed, a recent evaluation of the NDIS in its trial phase from 2013 to 2016 (Mavromaras et al., 2018) found a majority of participants would have liked to have more choice and control over the supports available to them via their funded plans, a concern that continues to increase steadily (NDIA, 2021a).

As previously foregrounded, only a fraction of the population is projected to receive personal funding from the NDIS at any one time, leaving the remainder of the population to draw from other services and informal supports. The scheme assures support for people with disabilities across Australia who do not receive direct funding, by way of the Information Linkages and Capacity Building (ILC) program. Designed to focus in particular on raising community awareness of the barriers that affect social and economic participation for people with disabilities while building greater capacity for their inclusion (NDIS Act, 2013), the ILC framework rather ambitiously seeks to consolidate mainstream service delivery with non-specialist, publicly-available information, while funding projects with these aims on an ad hoc basis. Such a project might be, for example, to ensure a public library service is more inclusive of people with vision impairments, by providing staff training and purchasing resources (NDIA, 2021b). Other initiatives provide information to people with disabilities about engaging with services and everyday living, such as freely-available podcasts (Independence Australia, 2021). While these initiatives are promising, there is little to mandate all service providers to develop more inclusive ways of working through the ILC as projects are generally funded on an ad hoc basis. Additionally, people with disabilities who do not receive direct NDIS funding have reported increased difficulty accessing support since the introduction of the scheme (Mavromaras et al., 2018).

The NDIS and education

A persistent issue relevant to Australians with disabilities and their families, the NDIS and educators, to which this article now turns to examine in more detail, is the ways in which the scheme connects with services such as education. The scheme's sustainability is

dependent on funded participants emphasizing access to mainstream services including education, health, housing, employment, justice and family support (NDIA, 2020a,b), rather than relying solely on NDIS-specific providers (Productivity Commission, 2017). The provision of access for NDIS participants to mainstream services is a shared responsibility between the NDIA and Commonwealth, State and Territory government services, which draw on principles that were drawn up that explain the roles and responsibilities of different service systems and the boundaries between them. Educational participation is a consistently strong objective of NDIS participants with funded plans since its introduction (Mavromaras et al., 2018; NDIA, 2021a), and the scheme might be anticipated to support children and people of all ages with disabilities to access and participate in inclusive learning by providing bespoke responses to identified needs. Nevertheless, the division between the scheme and cross-sectoral education services present a consistent obstacle to educational inclusion (Brien et al., 2017; Whitburn, Moss, and O'Mara, 2017a,b; Yates et al., 2021). There is also a danger that those who live with disabilities who are not eligible to receive direct funding through the NDIS will be further disadvantaged should mainstream systems not adequately support their needs (Knight, 2013). It is, after all, an expectation that either with or without ILC program support, mainstream services will be inclusive of people with disabilities.

NDIS participants who access the scheme at different life stages have reported experiencing a inimitable gap between the scheme and education (Brien et al., 2017; Whitburn et al., 2017a). This comes about, at least in part, because of the way that the NDIS was initially designed. While the voices of people with disabilities were instrumental to the scheme's development (Thill, 2015), the strength of a clinical understanding of disability was the catalyst for its detachment from education at the design stage. In a recorded podcast interview with *The Conversation* (August 10, 2016) co-founders of the NDIS explained that they felt forced to abandon engaging in a "war about content" of the terms of the scheme between carer organizations, who were in favor of special education, and disability groups, who wanted to use the scheme to re-imagine how exclusionary educational practices might be addressed. Explicitly eliminating educational considerations from the design of the NDIS effectively evaded the potential for the scheme to inform a pronounced cultural shift toward inclusion for Australians with disabilities, while at the same time side-lining the inclusive aspirations of people with disabilities over that of carers. The broader implications here is that although the NDIS encourages innovative responses to complex barriers to inclusion, it unfortunately does little to address how disability has been traditionally viewed as a problem to be fixed, thereby reinforcing the marginal status of many people who live with disabilities (Corcoran et al., 2019; Horsell, 2020; Kendrick et al., 2017; Olney and Dickinson, 2019; Thill, 2015; Whitburn et al., 2017a,b).

As will be explored in further detail in the forthcoming sections, something of a paradox emerges in the ways that institutions like the NDIS orientate to disability (Whitburn and Corcoran, 2021). Whereas education providers are said to increasingly acknowledge and work with difference affirmatively to provide strength-based, inclusive opportunities to learning participation, and the NDIS seeks to transform how social and economic participation for people with disabilities is enacted in Australia, the scheme nevertheless commences from a medically-informed, deficit-first orientation. Furthermore, the disconnections between both services motivates families of children with disabilities to fabricate support needs to ensure continuity of provision (Whitburn, 2017a). As Whitburn et al. also found, complexities of navigating scheme bureaucracy brings into sharp focus the necessity for people with disabilities to attain a relatively high base-level of education, if they ever want to manage their own plans, rather than having another organization or the NDIA undertake this role on their behalf. This article now turns to examine some of the affordances of the scheme across educational sectors for people with disabilities and educators, before concluding with a discussion of how affirmative orientations to disability could otherwise inform NDIS participation, to give emphasis to what it could mean to support a student with disability to achieve an equitable education.

Funding education: what will the NDIS support?

A simple search of the internet brings up countless references that suggest a clear connection between the NDIS and education, however this is only partly the case. Significantly, a line is drawn demarcating between what types of supports are provided by both services, with the deliberate and justifiable intent of complementing one another rather than overlapping. The provision of funded supports is also dependent on a range of variables, including educational sectors, types of supports required by individuals owing to their age and impairments, and the supports they require. In simple terms, the NDIS funds support directly related to a person's impairment, while the education system in each State and Territory is expected to cover expenses associated with their educational requirements. The Commonwealth Government under the Gonski 2.0 From Growth to Achievement report (McKew, 2018) and each State and Territory education department, has specific funding available to schools to specifically support students with disabilities, from which it is understood most funded support will be drawn. **Table 1** illustrates the support provided for students both by the NDIS and the education systems in each State and Territory of Australia.

Although the list presented in **Table 1** delineating responsibilities may seem clear, it is plainly not exhaustive, and some consequences emerge as a direct result of the demarcations therein. Additionally, there are subtle differences that the NDIS will fund depending on educational sector and type of impairment. In the following subsections of this article, further discussion is provided about the educational support provided, aligned with each educational sector from early childhood, primary and secondary education, to the higher and community sectors.

Table 1 NDIS and Education provisions for students with disabilities.

<i>Funded NDIS support for students</i>	<i>Education provider support for students</i>
Personal care needs associated with a person's impairment while they attend education	Teachers, education assistants and other supports such as Auslan interpreters, visiting teachers
Some therapy unrelated to education that may be provided during school hours	Therapy delivered in schools for education or training purposes, such as allied health practitioners helping teachers and trainers adjust curricula
Specialized training of teachers and other staff about the specific personal support needs of a student with disability	General support, resources and training for teachers, tutors and other staff
Portable equipment to aid mobility and/or communication	Aids and equipment to make curriculums accessible, such as modified computer hardware, software and braille textbooks
Specialist transport for travel to and from the education provider	Transport for educational or training activities such as excursions, field trips and sporting carnivals
	Adjustments to buildings such as ramps, lifts and hearing loops
	Day-to-day supervision of students at school, including behavioral support

Modified from [NDIA \(2015, 2020a,b\)](#).

Early childhood education

The early childhood phase is a core area of interest to the NDIS, for the reasons that through targeted, timely and specified interventions, the disabling effects of a child's impairment can be addressed and potentially reduced, thereby reducing their overall dependence on the scheme. Entitled the Early Childhood Early Intervention approach since the scheme's inception, which is a somewhat clunky name that emphasizes a preoccupation with risk mitigation associated with a child's disabling condition/s, this area garners particular investment. While it is important to recognize that most children in Australia commence primary school before the age of six, this section of the article concentrates on the early childhood education and care (ECEC) sector specifically (which is guided by the principles of the Early Years Learning Framework from birth-5 years ([DET, 2016](#))) as it relates to the NDIS. Approximately 16% of NDIS participants are children aged 0–6 years and their families or carers, and this group consistently represents a high proportion each quarter of applicants seeking scheme access through new plans, largely in response to early diagnoses of developmental impairments among children ([NDIA, 2021a](#)). In 2020 the NDIA launched a review, receiving 192 submissions on its impact for children and their families, and related stakeholders ([NDIA, 2020b](#)). The NDIA has rushed to make three recommendations quickly from the review, including stepping back from its deficit-based, intervention-centred inclination, and reframing it instead in terms of service delivery to children under nine years of age and their families, under the title the “Early Childhood Approach” ([NDIA, 2021a](#)). They have also sought to clarify operational guidelines and the criteria by which distinctions can be made to assess developmental capacities ([NDIA, 2021a](#)).

ECEC providers are anticipated to deliver reasonable adjustments under the Disability Standards for Education 2005 ([Australian Department of Education, Science and Training, 2006](#)), including any resource- or personnel supports they require to ensure an inclusive learning environment that facilitates enrollment and participation of all students. Aside from the types of NDIS funded supports foregrounded in [Table 1](#), including freely available information linking families with necessary information through the ILC program, the principles that delineate service responsibilities ([NDIA, 2015](#)) indicate that early childhood participants can receive transition support, intended both in making timely interventions away from disability-specific providers, as well as funding to transition out of ECEC into schools. The latter support is provided on the basis that the funding required is beyond typical school-readiness programs, to respond to a child's specific impairment or behaviour-related circumstances. With some therapy available to individuals during the hours a child may attend an ECEC service, the NDIS has increased the presence of disability specialists in these settings ([Boaden et al., 2021](#)). In spite of these promises of the NDIS to ECEC and the early childhood sector more broadly, the scheme's impact is yet to be realized ([Brien et al., 2017](#)). Parents and carers of children accessing the scheme have reported considerable challenges attaining support, along with diminished choice and control, and they express ongoing concerns that the values that therapeutic supports in ECEC settings are hampered because both systems are not able to work in collaboration ([Gavidia-Payne, 2020](#)). To this end, while the Early Years Learning Framework emphasizes family-centred, strength-based practices that provide for child-centred choice and control ([DET, 2016](#)), the impact of the NDIS in correspondents with the ECEC sector in support of these ideals is unclear.

Primary and secondary school

Accessing compulsory education is a consistently prevalent activity of NDIS participants ([NDIA, 2021a; Whitburn et al., 2017a,b](#)). To that end, while Whitburn et al. report that early plans of the NDIA during its initial trial phase were to provide single plans to school-aged children that would incorporate their education, similar to the joined-up plans available in England to support young people from birth to 26 years ([Hunter et al., 2020](#)), this has not found to have eventuated in the Australian scheme. Instead, there are guidelines that define what the NDIS will support in relation to the compulsory schooling sector, and what is understood to be

the responsibility of schools in providing equitable access to the curriculum, assessment and reporting along with all other students, as illustrated in [Table 1](#). That being said, there is a thin line between what both systems will fund, which may be open to diverse interpretations. The NDIS may fund specialist training to teachers and relevant school staff about an NDIS participants' impairment/s or behavioral needs, however where this strays into education-specific skills, these costs must be born by the school. Similarly, while the NDIS may fund some equipment that aids mobility or communication, such as wheelchairs, hearing aids and communication devices, any building modifications to support access, and assistive technology that a student might use to access their studies is to be provided by the educator. In another unclear delineation, services such as allied health that are unrelated to schooling may be provided during school hours and on school grounds. However, if such services are directly related to a students' schooling, these costs must be borne by the educator as a reasonable adjustment.

Invariably, these divisions have led to insufficient provisions from either service ([Whitburn et al., 2017a](#)), and as it has been expressed eloquently by a parent of a school-aged NDIS participant, "You leave the support at the school gate, and pick it up again at 3.30" ([Whitburn et al., 2017a](#), p. 1072). As the onus to support a student's education shifted rapidly from schools and teachers to parents and carers in response to the global Covid-19 pandemic, these concerns seemingly expanded. In a survey of student participants of the NDIS and their families, [Yates et al. \(2021\)](#) found that many were unable to obtain the technology and personal support they required for online learning, due to inflexibility in plans. Abandoned by both the education system and the NDIS, in many instances families with sufficient skills and resources had to resort to teaching their school-aged children with minimal support that was otherwise available at the schools. Many, however, reportedly went without support—an indication of the exacerbated inequity that is rife in the NDIS ([Carey et al., 2020](#); [Yates et al., 2021](#)).

One area of significance, for which the NDIS does provide direct support, relates to leaving school. Much in the way that specialist transition support is available into schooling for NDIS participants, specific funding to aid students transitioning out of the compulsory sector into employment is provided. A School Leaver Employment Supports program ([NDIA, 2021c](#)) can be accessed by NDIS participants in their final year of schooling, as well as their first post school year, which provides one-on-one job seeker skill development. Further, the NDIS has a Participant Employment Strategy which it launched in 2019 ([NDIA, 2021a](#)), through which 30% of funded participants are intended to be employed by June 2023. The strategy consists of six priority areas, one of which being to strengthen the link between education and employment outcomes for young people. Activities include reviewing current capacity-building programs, while strengthening work experience availability for NDIS participants while they are enrolled in schools. While these efforts seem useful, there appears to be a particular focus on disability-specific employment programs, while at the same time giving emphasis to overcoming individual deficits—a preoccupation similar to the plans available to young people with disabilities in England ([Hunter et al., 2020](#)). This point will begin engaged further in the conclusion to this article.

Higher education and vocational education and training (VET)

Education providers in the post-compulsory sector are also bound by law ([Australian Department of Education, Science and Training, 2006](#)) to make reasonable adjustments to their physical campuses, buildings and curriculum, to underwrite the enrollment and participation of all students with disabilities. The Commonwealth Government provides universities funding under the Higher Education Disability Support Program ([DET, 2020](#)) to cover any costs associated with these activities, however this is not available to educators in the VET or community sectors. NDIS participants are also afforded the same list of funded supports as are children and school-aged service users, including transition support both into and out of post-compulsory education. The agency will also provide specific training to teaching staff and placement supervisors as required, extending scope for this type of support outside the institution and into the workplace. While the ambiguity of NDIS funding for students in ECEC or schooling seeking NDIS support to participate in learning is just as relevant to students in the non-compulsory sector, it has been suggested that ongoing informal support throughout a student's course would be of particular use for those with autism spectrum condition ([Owen et al., 2016](#)), who might be at risk of attrition due to the unique structure of non-compulsory study.

Conclusion

The NDIS has been designed and implemented swiftly to capitalize on the widespread and bilateral support from opposing sides of the political divide ([Olney and Dickinson, 2019](#)). To this end, the scheme is still evolving, and proposed changes are ushered in quickly with minimal, or at best, overly complicated community consultation ([Boaden et al., 2021](#); [Kendrick et al., 2017](#); [Thill, 2015](#)). In this article the interface between the NDIS and education has been examined, the affordances of both systems for educators and students provided, and some of the challenges therein explained. It is clear that the NDIS, in its fledgling stage, has made a difference to the national landscape for disability and mainstream services. Yet, it operates on a "punitive and exclusionary" ([Carey et al., 2020](#), p. 14) approach to allocating the types of support people are eligible for, giving rise to some families exaggerating their deficits and needs ([Whitburn et al., 2017a](#)). Not everyone, of course, is able to advocate for their needs equally. Accordingly, disparities appear by way of NDIS provisions for participants from low socioeconomic backgrounds, people with cultural and linguistic diversities, and Indigenous Australians ([Carey et al., 2020](#); [Soldatic et al., 2014](#); [Yates et al., 2021](#)), entrenching the very inequities the scheme was intended to address. That the scheme was designed favoring the special education tradition, rather than to engage in

re-imagining how education across sectors and its linkages with employment might be made more inclusive for Australians with disabilities, is a legacy of this positioning.

In conclusion, this article provides a brief commentary as to how recent disability scholarship and advocacy might inform children and people with disabilities, their families and educators to think differently about engaging with services such as education and the NDIS, whether or not attaining funded support in so doing. In its present form, the NDIS is a hybrid between a system that seeks to promote human rights, by creating and giving access to a quasi marketplace for people with disabilities (Carey et al., 2020). This in itself is something of a contradiction, unless of course inclusion for people with disabilities is best measured by the degree to which they can operate freely in a market-driven culture. A systemic desire to increase social and economic participation of people with disabilities is laudable, and we might recognize that education and employment contribute financial stability and wellbeing for many. However, the scheme also demonstrably positions deficits within individuals, while seeking to increase their independence and claiming to adhere to principles of the social model of disability. While some might think that to critique these features of the NDIS is toxic, rife among published accounts of NDIS participation are troubling experiences associated with the scheme's architecture (Mavromaras et al., 2018; Whitburn et al., 2017a; Yates et al., 2021). We might wonder how things could be otherwise.

Recent developments in disability advocacy and scholarship invite us to consider how we come to know disability (Whitburn and Corcoran, 2021; Hunter et al., 2020), with a view to dismantling categories of deficit that were previously taken as read. This is to recognize that a complex intersection of disadvantages affects marginalization for people with diagnosed impairments, while at the same time finding solace in the relationships between people and the support mechanisms that ably complement their lives. The barriers that disable people, then, may be biological, but they could just as easily be social, political, and/or technological. Taking such a view to NDIS participation is to move away from tough concepts like independence, needs, and intervention, and to consider more how the relationships between people, technological devices, and even animals make activities possible. These are continually changing in relation to context. A person's NDIS plan might be understood not to tell things how they are, but how they might be in relation to these conditions, resources and the people who live them. As Hunter et al. also implore in the case of England, we only hope that the NDIA accordingly realizes the transformative potential of broadening its focus along these lines to increase the inclusive possibilities of the scheme for people with disabilities to achieve their educational ambitions in Australia.

References

- Australian Bureau of Statistics (ABS), 2019. 4430.0 – Disability, Ageing and Carers, Australia: Summary of Findings, 2018. <https://www.abs.gov.au/ausstats/abs@.nsf/mf/4430.0>. (Accessed 27 September 2021).
- Australian Department of Education, Science and Training, 2006. Disability Standards for Education 2005 Plus Guidance Notes. <http://nla.gov.au/nla.arc-7692>.
- Boaden, N., Purcal, C., Fisher, K., Meltzer, A., 2021. Transition experience of families with young children in the Australian national disability insurance scheme (NDIS). *Aust. Soc. Work* 74 (3), 294–306.
- Brien, J., Page, J., Berman, J., 2017. Enabling the exercise of choice and control: how early childhood intervention professionals may support families and young children with a disability to exercise choice and control in the context of the National Disability Insurance Scheme. *Australas. J. Early Child.* 42 (2), 37–44.
- Carey, G., Dickinson, H., Malbon, E., Weier, M., Duff, G., 2020. Burdensome administration and its risks: competing logics in policy implementation. *Adm. Soc.* 52 (9), 1362–1381. <https://doi.org/10.1177/0095399720908666>.
- Corcoran, T., Claiborne, L., Whitburn, B., 2019. Paradoxes in inclusive education: a necessary condition of relationality? *Int. J. Incl. Educ.* 23 (10), 1003–1016.
- Department of Education and Training (DET), 2016a. Victorian Early Years Learning and Development Framework. DET and Victorian Curriculum and Assessment Authority, Melbourne, VIC.
- Department of Education and Training (DET), 2020. Higher Education Disability Support Program. <https://www.dese.gov.au/higher-education-disability-support-program>. (Accessed 23 September 2021).
- Gavidia-Payne, S., 2020. Implementation of Australia's national disability insurance scheme: experiences of families of young children with disabilities. *Infants Young Child.* 33 (3), 184–194. <https://doi.org/10.1097/YYC.0000000000000169>.
- Horsell, C., 2020. Problematising disability: a critical policy analysis of the Australian national disability insurance scheme. *Aust. Soc. Work* 1–13.
- Hunter, J., Runswick-Cole, K., Goodley, D., Lawthorn, R., 2020. Plans that work: improving employment outcomes for young people with learning disabilities. *Br. J. Spec. Educ.* 47 (2), 134–151.
- Independence Australia, 2021. Inform Podcast. <https://www.independenceaustralia.com.au/inform-podcast/#:~:text=Inform%20Podcast%20is%20a%20monthly,navigating%20the%20NDIS%20and%20more>. (Accessed 23 September 2021).
- Kendrick, M., Ward, M., Chenoweth, L., 2017. Australia's national disability insurance scheme: looking back to shape the future. *Disabil. Soc.* 32 (9), 1333–1350.
- Knight, K., 2013. The changing face of the 'good mother': trends in research into families with a child with intellectual disability, and some concerns. *Disabil. Soc.* 28 (5), 660–673.
- Mavromaras, K., Moskos, M., Mahuteau, S., Iskerwood, L., 2018. Evaluation of the NDIS: Final Report. National Institute of Labour Studies, Flinders University, Adelaide. Retrieved from: https://www.dss.gov.au/sites/default/files/documents/04_2018/ndis_evaluation_consolidated_report_april_2018.pdf.
- McKew, M., 2018. Whatever happened to Gonski 2.0? Pursuit University of Melbourne. <https://pursuit.unimelb.edu.au/articles/whatever-happened-to-gonski-2-0>. (Accessed 20 September 2021).
- National Disability Services (NDS), 2016. Factsheet: Economic Benefits of the NDIS. https://www.nds.org.au/images/factsheets/ndis_economic_benefits.pdf. (Accessed 25 September 2021).
- NDIA (National Disability Insurance Agency), 2015. Principles to Determine the Responsibility of the NDIS and Other Service Systems. National Disability Insurance Agency. <https://static1.squarespace.com/static/57c65af5cd0f68b1295663dc/t/586092084402439b0c8370ab/1482723859121/NDIS+-+Principles+to+Determine+Responsibilities+NDIS+and+Other+Service+Systems+-+Revised+27+Nov+2015+%281%29.pdf>. (Accessed 20 September 2021).
- NDIA (National Disability Insurance Agency), 2019. Reasonable and Necessary Supports. National Disability Insurance Agency. <https://www.ndis.gov.au/understanding/supports-funded-ndis/reasonable-and-necessary-supports>. (Accessed 25 September 2016).
- NDIA (National Disability Insurance Agency), 2020a. NDIS and Other Government Services. <https://www.ndis.gov.au/understanding/ndis-and-other-government-services>. (Accessed 23 September 2021).
- NDIA (National Disability Insurance Agency), 2020b. Help for Children under 7. <https://www.ndis.gov.au/understanding/how-ndis-works/help-children-under-7>. (Accessed 23 September 2021).

- NDIA (National Disability Insurance Agency), 2021a. NDIS 2020 – 2021 Quarterly Report June 30 2021. National Disability Insurance Agency. <https://www.ndis.gov.au/about-us/publications/quarterly-reports>. (Accessed 25 September 2021).
- NDIA (National Disability Insurance Agency), 2021b. Information, Linkages and Capacity Building (ILC). <https://www.ndis.gov.au/community/information-linkages-and-capacity-building-ilc>. (Accessed 25 September 2021).
- NDIA (National Disability Insurance Agency), 2021c. School Leaver Employment Supports. <https://www.ndis.gov.au/providers/working-provider/school-leaver-employment-supports>. (Accessed 22 September 2021).
- NDIS Act, 2013. National Disability Insurance Scheme Act. Office of Parliamentary Counsel, Canberra. No. 20.
- Oliver, M., 2009. *Understanding Disability: From Theory to Practice*. Macmillan International Higher Education.
- Olney, S., Dickinson, H., 2019. Australia's new national disability insurance scheme: implications for policy and practice. *Policy Des. Pract.* 2 (3), 275–290. <https://doi.org/10.1080/25741292.2019.1586083>.
- Owen, C., McCann, D., Rayner, C., Devereaux, C., Sheehan, F., Quarmby, L., 2016. Supporting Students with Autism Spectrum Disorder in Higher Education. Report submitted to the National Centre for Student Equity in Higher Education (NCSEHE), Curtin University, Perth.
- Productivity Commission, 2011. *Disability Care and Support: Productivity Commission Inquiry Report*. Productivity Commission, Melbourne, Australia.
- Productivity Commission, 2017. National Disability Insurance Scheme (NDIS) Costs. <https://www.pc.gov.au/inquiries/completed/ndis-costs/report>. (Accessed 23 September 2021).
- Soldatic, K., van Toorn, G., Dowse, L., Muir, K., 2014. Intellectual disability and complex intersections: marginalization under the national disability insurance scheme. *Res. Pract. Intellect. Dev. Disabil.* 1 (1), 6–16.
- The Conversation, Producer, 2016. Change Agents: Rhonda Galbally and Bruce Bonyhady on the Birth of the NDIS. The Conversation. August 10. <https://theconversation.com/change-agents-rhondagalbally-and-bruce-bonyhady-on-the-birth-of-the-ndis-63662>. (Accessed 20 September 2021).
- Thill, C., 2015. Listening for policy change: how the voices of disabled people shaped Australia's National Disability Insurance Scheme. *Disabil. Soc.* 30 (1), 15–28.
- United Nations, 2006. *Convention on the rights of persons with disabilities*. Treaty Series 2515 (December), 3.
- Whitburn, B., Corcoran, T., 2021. Diversity training, inclusive education and our inevitable lament. *J. Res. Spec. Educ. Needs* 21, 17–24.
- Whitburn, B., Moss, J., O'Mara, J., 2017a. The National Disability Insurance Scheme and access to education: progressive or coercive policy discourse? *Int. J. Incl. Educ.* 21 (10), 1065–1079.
- Whitburn, B., Moss, J., O'Mara, J., 2017b. The policy problem: the National Disability Insurance Scheme (NDIS) and implications for access to education. *J. Educ. Pol.* 32 (4), 467–479.
- Yates, S., Dickinson, H., Smith, C., Tani, M., 2021. Flexibility in individual funding schemes: how well did Australia's National Disability Insurance Scheme support remote learning for students with disability during COVID-19? *Soc. Pol. Adm.* 55 (5), 906–920.

Disability and employment: finding a way to work

Abigail E. Croydon and Melanie AE. Nind, Southampton Education School, University of Southampton, Southampton, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Connecting disability and employment	453
Responding to disability employment gaps	453
Interpreting disability employment data	454
Summary	455
The capability approach	455
Rethinking employment: the case of intellectual disability in high-income countries	456
Conclusion	459
References	459

Connecting disability and employment

The term “disabled” refers to a diverse group of people whose personal characteristics relevant to employment vary widely. Health, age, gender, socio-economic background and education have complex and dynamic relationships with impairment, employment and social inclusion (Banks and Polack, 2014; Meyer and Mok, 2019; World Health Organization and World Bank, 2011a,b). These interact with local cultural and economic conditions to influence individuals’ prospects in terms of work and earnings. Consequently, measurement of employment rates is a vexed issue; it depends to some extent on how “disability” is differentiated from the range of factors influencing employment opportunities as well as what counts as employment.

The meaning of “employment” and the availability of work varies especially between the Global North and South and between men and women. Nevertheless, globally disabled people are significantly disadvantaged in terms of employment across the range of types and severity of disability. Disabled people are less likely to be employed; they are on average paid less; and they are more likely to be in part-time, temporary and insecure forms of employment (Gunderson and Lee, 2016; Mitra and Kruse, 2016; Schur et al., 2017). Disabled people also receive fewer training and development opportunities than non-disabled people (Schur et al., 2017).

As disability is defined in relation to people’s ability to carry out everyday tasks and to participate socially, it is inherently linked to basic questions of people’s occupations and employment, age, health and social inclusion. Historically “disability”, “unemployment” and “retirement” developed as related constructs in twentieth-century welfare states, as work in industrial production intensified. As work disability increased among older workers, sickness benefit came to be used as a surrogate old age pension (Macnicol, 2011). This history suggests that the shift to wage labor – to jobs and contracted employment – entailed increasing work disability as performance and individual level productivity became critical aspects of work. Currently, disability employment statistics show that levels of disability resulting from musculoskeletal problems associated with aging are among the biggest contributor to disability employment statistics in the Global North and South, only age of onset differs (World Health Organisation, 2021). The shift to wage labor resulted in increasing de-recognition of other forms of work, especially the unpaid work of women (Ferguson et al., 2016; Standing, 2009).

In high- and low-income countries, disabled people overall have lower educational attainment (WHO, 2011), consistent with socioeconomic and gender-based disadvantage and unequal access to education. Education disparities alone do not account for employment outcome differences (Jones, 2008; Vormholt et al., 2018), but they do critically affect prospects for work. As labor markets increasingly select for higher levels of educational attainment in middle- and higher-income countries, this education and skills differential appears to be increasing over time (OECD, 2008). This suggests that disability employment gaps, that is differences between the employment rates of people categorized as disabled and others, may also grow.

In the Global North disabled people are among those targeted by activation policies seeking employment for all, on the grounds of social justice and reduced welfare spending. This “employment-focused paradigm” (Laruffa, 2020, p. 2) has been supported by the European Union and international organizations including the World Bank and the OECD. It includes incentivizing people to work using non-work sanctions and in-work benefits, and calling for investment in human capital to increase employability (see Laruffa, 2020). In the case of disabled people, such policies present employment as achievable and as the solution to entrenched associations between disability, low income and high social exclusion (see Grover and Piggott, 2015; Redley, 2009). By implication, there are no valid alternative routes to equal social participation when people are unable to attain or sustain employment. The high social value placed on employment – rather than work in other forms – is not confined to high income economies, as shown below.

Responding to disability employment gaps

The international community has addressed disadvantage in work and employment for disabled people as a development issue. The United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) (2006) is the key agreement relevant to

employment for disabled people, namely Article 27 governing work and employment. The UNCPRD (2006) establishes a “right to work” and focuses on the disabling impacts of social attitudes and social arrangements, as key issues for employment in the disability context. The treaty advocates legal and policy mechanisms to uphold the right to work, such as measures to counter discrimination on the grounds of disability and the concept of “reasonable adjustments,” that is, employers’ duty to modify workplaces and practices for accessibility. Protection of the right to work in this convention is also promoted through equal access to education and training, opportunities for self-employment and entrepreneurial activities, and vocational and professional “rehabilitation”. Aspirations for disabled people in terms of work and employment are also inscribed in the UN Sustainable Development Agenda for 2030, where Goal 8 (Decent Work and Economic growth) recognizes the additional disadvantage in employment terms for disabled women and girls and for people with intellectual disabilities and mental health conditions (United Nations, 2016).

Yet competitive markets in employment are inherently difficult to reconcile with the overall diversity of disability. The summary version of the World Health Organization (2011) *World Report on Disability* cites two examples of successful accommodations in the workplace: a person using a wheelchair enabled by accessible buses and lifts, and a person with visual impairment whose disability is eradicated by corrective lenses. The examples echo social model of disability arguments that a suitably adjusted work environment may effectively erase disability. In such an erasure, impairment effects (Thomas, 2004) such as pain and fatigue, and the impact of impairments such as intellectual and developmental disabilities and mental health difficulties are underplayed (Morris, 2007; Shakespeare, 2006). Also downplayed are the complex associations between impairment, age, and socio-economic disadvantage, which are not readily rectified by accommodations. Arguments making the “business case” for disability inclusion focus on potential advantages in productivity and financial performance to be gained by employers from disability inclusion (ILO, 2019), making the case that employment is effectively conditional upon such advantages.

Overall the gap between employment rates for disabled and nondisabled people have remained stable following an unprecedented amount of policy activation in individual countries and through the UNCPRD (Sainsbury, 2018; Vornholt et al., 2018), including the activation policies targeted at individuals, legislation against discrimination and supported employment programs. Supported employment, that is, providing job search and job support services to people furthest from the labor market, is a practice with an international evidence base to improve employment outcomes for the rising number of people with mental health conditions (Modini et al., 2016). Nevertheless more than half of those supported may not gain employment (Corbière et al., 2017). For those with intellectual and developmental disabilities, supported employment has a chequered history with little impact on extremely low levels of employment (Blamires, 2015; Wehman et al., 2018). Similarly, evidence for the impact of anti-discrimination legislation is weak; the Americans with Disabilities Act (1990) and the UK Disability Discrimination Act (1995) seem to have had little positive impact on employment rates for disabled people (Bell and Heitmueller, 2009). Looking ahead, if the availability of jobs overall decreases through automation and technological change (OECD, 2019; Suskind, 2020), employment levels for disabled people are likely to lag further in middle- and higher-income countries. A shrinking jobs market is likely to have a disproportionate impact on disabled people, as it affects disadvantaged workers most (OECD, 2019) – that is people with lowest levels of education and skills, social capital and mobility. OECD forecasts for 2019 highlighted the increasing need to switch jobs flexibly and retrain extensively to remain in employment, while those for 2020 highlighted the jobs crisis arising from the Covid-19 pandemic and its potential to widen inequalities (OECD, 2020). Together with the failures of policy interventions to shift employment rates for disabled people, such projections point to the need to reconsider the primacy of paid work in disability policy.

Interpreting disability employment data

Examining the global evidence more closely, there are considerable variations in the scale of employment differentials between disabled and non-disabled people, and reasons to interpret employment statistics with care. Data focusing on whether or not disabled people have attained employment in high-income countries may also be relatively uninformative. Those without paid work are set aside from employment accounting as “economically inactive” even though a high proportion (30–40%) would like to work. For example, in the UK nearly half of disabled people are counted as “economically inactive”, while only 6% count as “unemployed” (Meager and Higgins, 2011; Evans, 2019). The overall picture of marked disadvantage in employment and earnings conceals large and important differences between groups and settings. First, data on the employment of disabled people across the world is significantly unbalanced, with greater information available for high-income countries than for low income countries. In comparing employment statistics, basic difficulties arise in the different inclusion criteria across countries and studies, with low-income countries often using narrower impairment-focused criteria and high-income countries broader “activity limitation” definitions (Mizunoya and Mitra, 2013). When statistics are adjusted for comparability, the average employment disadvantage for working-age men with disabilities is said to be around 12 points (53% employed by comparison with 65% of men without disabilities) based on survey data from 51 countries (World Health Organization and World Bank, 2011a,b). For women there is said to be a 10 point gap, arising in the context of overall lower employment rates (20% compared to 30% of women without disabilities). Mitra and Mizunoya (2013) found significantly lower employment rates for men with disabilities in most of 15 low-middle income countries, with significant differences between countries, that is, employment gaps ranging from 11 to 53 percentage points, and no apparent difference for six others. The statistics for women, discussed below, show less impact of disability status.

There are markedly different impacts on employment levels and earnings for people with chronic and severe impairment and for some types of impairment or disability. In Europe, people whose disabilities are classed as severe are approximately three times more likely to be unemployed than non-disabled people (World Health Organization and World Bank, 2011a,b). In low-

middle-income countries, evidence indicates people with multiple disabilities are significantly more disadvantaged than those with “single” disabilities (Mizunoya and Mitra, 2013). In the USA, men acquiring a chronic or severe disability experience dramatic drops in earnings (around 80%) over the first ten years following disability onset (Meyer and Mok, 2019). People with intellectual and mental health disabilities are also considerably more disadvantaged than people with other disabilities in employment terms, regardless of income level of country (Mizunoya and Mitra, 2013; Lysaght et al., 2012). In the UK, for example, around 6% of people with intellectual disabilities and 9% of people with significant mental health difficulties are employed (2019–2020), by comparison with 44% of disabled people as a whole and 78% of non-disabled people (ONS, 2019). In terms of earnings, the pay gap is twice as high for people with a “mental impairment” – a term that covers both intellectual and mental health related disabilities – as for people with a physical impairment (Evans, 2019). These differences between people with disabilities in employment terms are important reminders of the diversity within the category of disability which may be minimized in the employment-focused paradigm (see Meager and Higgins, 2011; Stevens and Harris, 2013).

Globally, the picture is affected not only by the imbalance in available information, but by the dominance of the Global North in setting the agenda for disability studies. The framing of disability and employment issues in much research and policy formation reflects the cultural and political dominance of the Global North (Meekosha, 2011). This includes that of the UNCPRD (2006) since the concepts and processes of human rights discourse are “culturally Eurocentric” (Meekosha, 2011, p. 678). This Global North inflection affects critical questions of how exclusion from work is counted and how priorities are set. Grech (2008) notes that the large majority of disabled people are located in parts of the Global South where levels of poverty provide multiple conditions for the creation and maintenance of impairment. Meekosha (2011) questions how the production of disability in the exploitative working conditions of sweatshops in the South, for example, comes to be such a neglected phenomenon.

Two further issues arise in the context of Global North conceptions of disability in relation to employment: i) the interpretation of work as a matter of employment in the formal economy, and ii) the implications of this construction of work for understanding the mechanisms driving employment statistics for disabled people. In the Global South, data and policy on rights and protections for disabled people in employment largely revolve around contracted employment in the formal economy (Mizunoya and Mitra, 2013), though this may represent the work of a minority of people. This focus may be particularly inappropriate in relation to the lowest income countries where a majority of workers may be in the informal economy, in small trading, subsistence agriculture and other unregulated forms of work. Self-employment in the informal economy may have advantages in terms of low barriers to entry and flexibility that make it particularly prevalent among women and disabled people (Mizunoya and Mitra, 2013; Schur, 2003). Hence, there is no employment gap between women with and without disabilities in some African countries, though one exists for men (Mizunoya and Mitra, 2013). This underscores the significance of gender rather than disability in rates of employment for women, and indicates how the economic activity of women and disabled people may be overlooked when formal employment is the measure of participation.

The idea that it is the strictures of formal employment that particularly affect disabled people is also supported by statistics comparing urban and rural settings in low-income countries. These show less disadvantage for disabled people in rural settings, where formal employment is less available. Similarly, the disability employment gap is larger in middle-income countries where formal employment is more prevalent compared to low-income countries (Mizunoya and Mitra, 2013).

Variation in the connections between disability and employment may point to a wider conclusion about exclusion from the jobs market and its consequences. Persistent negative employment statistics for disabled people may be informative about how jobs markets are shaped, and made inaccessible, by the pressures of global competition (see e.g., Grover and Piggott, 2015; Moore et al., 2018) while providing relatively little information about what excluded people do or want to do.

Summary

Independently of definition of disability, data source, country or time-period, disability has a negative effect on employment and earnings. Commonly, statistics for formally contracted work are interpreted as a measure of how far the rights of disabled people are protected in terms of work, and how far such people are contributing to and socially included in their communities. Relatively little is known about other kinds of work by disabled people, especially work in the informal economy, and how far it is preferred by people unable or unwilling to conform to the exigencies of paid employment (Schur, 2003; Mizunoya and Mitra, 2013). Evidence suggests that working as an employee is difficult to achieve and demanding to sustain for some disabled people (Moore et al., 2018; Vornholt et al., 2018), and that it may be increasingly inaccessible for many disadvantaged people, including disabled people, in the future. Meanwhile, the capability approach (Sen, 1999) offers a framework for reconsidering employment, work and disability in relation not to employment metrics but to people’s freedoms to pursue the kinds of life they value.

The capability approach

In considering the disability/employment intersection, this chapter has highlighted disparities regarding employment outcomes and the potentially distorting role of applying in a global context understandings of disability and priorities for disability policy that derive from a minority world perspective. The capability approach developed by Amartya Sen and others (Sen, 1999; Robeyns, 2017) is a conceptual framework that has advantages for conceptualizing disability and employment because of its central focus on

diversity in the personal characteristics of individuals (such as health, gender, education, mental, physical and sensory abilities), how these interact with their social and cultural contexts and the value systems they hold. The capability approach is primarily concerned with questions of human development and inequalities, and has been understood as a means to reconceptualize disability as an aspect of “pervasive” (Sen, 1992, p. 1) human diversity, like age and gender rather than part of an abled/disabled binary (Burchardt, 2004; Mitra, 2006). In employment terms, the capability approach highlights the meanings and functions of forms of work within the lives of the people concerned, rather than focusing on metrics of employment and earnings (Bartelheimer et al., 2012; Yeoman, 2014). With diversity as a central rather than secondary factor, using a capability approach means taking into account personal characteristics as they operate within a person’s physical environment, in interaction with the relevant social and cultural conditions, and in the light of the individual’s aims and objectives. According to Sen (1992), the question of equality is demanding precisely because such pre-existing variations imply the need for unequal treatment in favor of people who are disadvantaged.

In the capability approach, it is the agency of people rather than their resources *per se* or their “wellbeing” that is paramount – their freedom to be and to do the things they judge to be valuable. Hence, development is not necessarily a matter of economic growth, technological advance, or social modernization, but a process of “expanding the real freedoms that people enjoy” (Sen, 1999, p. 3). “Poverty” is understood as deprivation in the capability to live a good life, and “development” as a matter of expanding people’s capability (Robeyns, 2016). This makes it relevant equally in high-income countries as in the majority world, where debate on development is often focused. Sen (1992) argues that intergroup contrasts in high-income countries may be stark, with disadvantaged groups lacking the basic opportunities of health care, education, employment, or economic and social security in spite of the broader context of prosperity. This is the case for certain groups of disabled people, for example, for people with intellectual and developmental disabilities, who are effectively excluded from socially valued forms of employment, and thereby from a central means of social participation. The capability approach has been applied extensively in the context of education, and offered as a framework to define the ethical purposes of inclusive education (Reindal, 2016). It is proposed here to support an argument for work – outside the restricted parameters of employment – as a means to extend the capabilities and social inclusion of people most disadvantaged in relation to employment. Yeoman (2014, p. 236) refers to work as “a mode of being in the world which transcends the employment relation to include all the activities which contribute to producing and reproducing a complex system of social cooperation”. Work in this sense is “meaningful” and can be understood to be a fundamental human need.

In addition to its uses in reconceptualizing the employment/disability intersection and what matters to those involved in it, the capability approach has also been developed and used as a tool to evaluate policies and institutions, such as the design of a welfare state or the priorities of poverty reduction strategies (Robeyns, 2017). The *World Disability Report* (World Health Organization and The World Bank, 2011a,b, p. 11) uses it as a means to understand “the obligations that states owe to individuals to ensure that they flourish, exercise agency, and reach their potential as human beings.” Setting employment as a core target of policy for disabled people misrecognizes diversity in people’s needs, purposes and environments, effectively reducing their capability to their capacity to participate in a given economy (Laruffa, 2020). From a capability approach perspective, the benefits of employment should not be taken for granted; indeed, as Meekosha (2011) points out, some forms of employment involve the production of disability. Instead, the extension of people’s capabilities, especially the key capabilities of learning and social inclusion, may be a more productive aim in the context of employment and disability.

Rethinking employment: the case of intellectual disability in high-income countries

People with intellectual and developmental disabilities are particularly affected in employment terms – that is, by low levels of employment, limited employment options, and adverse experiences in employment. The UK and Australia provide examples of conditions in high-income countries with comparable economic and policy contexts, where, as Sen (1992) argues, inequalities may be substantial. Links between work, learning and social inclusion mean that work might extend the capabilities and social inclusion of this group. These links outline some of the “goods of work other than money” (Gheaus and Herzog, 2016) that may be more attainable by people with intellectual and developmental disabilities outside the conventional parameters of employment than within them: opportunities to learn, to make a social contribution, to experience community and receive social recognition.

In high income countries, employment is a central factor structuring adult lives and identities for the majority. It is often cited as the primary means of social participation and linked with better health and mental health (e.g., Bambra and Eikemo, 2009), with learning and personal development (Manuti et al., 2015), social engagement and finding meaning in life (Yeoman, 2014), as well as earnings. It is therefore striking that around 93% of people with intellectual and developmental disabilities in the UK are not in employment (Public Health England, 2020). When they do attain employment, these assumed benefits are likely to be unrealized, on account of limited options for work, reduced hours, low pay and experiences of rejection. These issues define the need to consider ways of engaging in work outside the parameters of competitive employment.

When the majority of people are in employment, an immediate practical outcome of lack of employment for people with intellectual and developmental disabilities is under-occupation and social and spatial separation from working people. A recent review of literature relating to opportunities for engagement in meaningful occupations found a “concerning pattern of inactivity” among adults with intellectual and developmental disabilities (Channon, 2014, p. 443). Research involving young people in relatively

privileged circumstances – those expected to have benefited from modernized, personalized services – found them to be experiencing a strong sense of exclusion and under-occupation. Merrells et al. (2019, p. 16) summarized their participants' experiences with the phrase: "We always feel left out... experiencing exclusion and rejection from paid work in the community ... and having nothing to do." A research theme of "Not sitting around on my backside all day at home" featured a young adult's report on his "boring lifestyle": "Um, cleaned my house, did my dishes, um, went to the shops, looked around, got some DVDs last night to watch, yeah... get bored so I just go to the shops, but the shops get boring after a while" (Merrells et al., 2019, p. 17).

Under-occupation is often combined with lack of hope. People with intellectual and developmental disabilities and their families have been found to be pessimistic, with reason, about their chances in employment (Beyer et al., 2016; Giri et al., 2021). Many have extensive experience of trying to find and keep paid work, of placements that don't work out, in demanding circumstances experienced as "craziness" (Ramsey, 2020, p. 131). Recent research offers examples of how both seeking employment and being in low quality employment can increase people's sense of alienation from the majority. Experiences of social and employment rejection are commonly reported, for example:

"I tried nursery at one point, I thought I could get into nursery, that's when it hit me, I had a disability actually."

Ramsey (2020, p. 182)

The sense of frustration at the parameters of employment can be acute. Merrells et al. (2019, P. 16) report a young woman describing how a "mainstream person":

"doesn't understand how ... where we're coming from because she hasn't got a disability, it's hard to say what we feel like... it was really, really hard to handle and I just completely shut down."

The employment options open to people with intellectual and developmental disabilities are limited, which leaves people struggling with their aspirations and how to live their lives. A woman with a range of short-term employment experiences reflected on the disorienting impact:

what work I don't want is easier to say than what I do want because what I do want is kind of impossible, it's not realistic. I don't know you know I'm like in my thirties and I have no idea what to do.

Ramsey (2020, p. 171).

The constricted parameters of employment, its demands for efficiency and productivity, are structural – recruitment is a discriminatory process. An example of how competitive processes can play out in people's lives is given by a young adult who had obtained work clearing plates and washing dishes, but was losing out in the competition for shifts:

"I find it really difficult. Um, because I find it hard to keep up with everything because I know everything has to be done at a certain time... So, um, yeah, so I phoned up, like, last week to see if there's any more shifts, and there haven't been any shifts for me, like, this week, so ...yeah."

Merrells et al. (2019, p. 16)

Employment in these circumstances is unlikely to offer financial security; indeed most people with intellectual and developmental disabilities work too few hours, at too low rates of pay to become independent of state benefits (Public Health England, 2020); see also Dearing, 2020). Basic advantages of employment escape their grasp.

The dominance of formal employment in the economies of high-income countries means that there are limited other means of participating in society in significant and sustainable ways. A consequence of marginalization within, and exclusion from, formal employment is the loss of access to processes of social participation and learning that are fundamental in human society. Sociocultural approaches to learning (e.g., Lave and Wenger, 1991) depict critical links between learning, work and social inclusion, according to which a learner's participation in social practices defines the content, manner and extent of their learning. Such learning is seen as the primary means by which individuals make sense of and act on the world. Lack of participation, then, compounds the learning disadvantage of intellectual impairment. According to participatory learning theory, human learning originates in social, historical, and culturally specific interactions, where learners negotiate new learning challenges that are within their reach when supported by others. The literature explores how "novices" – children, young people and other kinds of newcomers – are inducted into the practices, knowledge and beliefs of their cultures by taking part in them, following pathways of increasing complexity over time (Rogoff et al., 2016). Learning by a given individual is therefore inseparable from the context within which it is embedded and from the support of others (Matusov, 2001), so that being there, taking part and having support are essential to learning and understanding your world.

Education, and by extension learning, have a critical place in the capability approach for their potential to enhance capabilities. Social inclusion is also a key capability (see [Reindal, 2016](#)). Much attention has been paid in education and disability studies to the question of inclusion in schooling and its relation to social inclusion. While the rhetoric of inclusion is everpresent in educational discourse, it coexists with policies and practices of segregation and the ranking of students ([Slee, 2019](#)), and with peer victimization and social rejection of minority groups ([Juvonen et al., 2019](#)), which critically affect what social learning takes place. In this way inclusive education can act as “an accomplice in the exclusion of students with disabilities” ([Slee, 2019](#), p. 911). From a capabilities perspective, these aspects of education jeopardize one of the critical tasks of learning - to challenge “adaptive preferences,” the tendency of disadvantaged people to adapt their aspirations according to their situation ([Watts et al., 2008](#)). In spite of these contradictions, people with intellectual and developmental disabilities share with non-disabled peers culturally defined experiences of formal education – that is, they attend schools. By contrast, the practices of work that are dominant in adult life are largely unexperienced. From this perspective, the learning and social participation aspects of work matter all the more for capabilities and inclusion.

It is important that the learning involved in social participation is not conceived as a one directional process of transmission of knowledge and values to the “newcomer”. Rather, as individuals take up different positions in relation to the cultures and practices they encounter, these are continuously reshaped. [Rogoff \(2003, p.51\)](#) describes this process as “people contributing to the creation of cultural processes and cultural processes contributing to the creation of people, so that individual and cultural processes are mutually constitutive.” Learning in this sense accounts for how cultures and practices evolve over time ([Tomasello, 2016](#)). According to these understandings, the absence of people with intellectual and developmental disabilities from the places and experiences of work represents a loss to society of opportunities for mutual learning and social adaptation. An outcome of exclusion, and a perpetuating factor, is that disabled people generally have been absent from the public sphere and consequently inhabit “a cultural, political and intellectual world from whose making they have been excluded” ([Abberley, 1999](#), p. 15).

In employment as in education, an “ethic of competitive individualism” ([Slee, 2019](#), p. 909) dominates policy and practice in ways that drive the social isolation of people with intellectual and developmental disabilities. Yet it is personal and relational support in work that is a condition of participation for many for people with intellectual and developmental disabilities. This is clearly expressed by a young woman interviewed by [Meltzer et al. \(2016, p. 44\)](#) as she worked in a social enterprise, having experienced open employment. They sum up: “Would you try another open employment job? Nope. Why not? No support.” A participant in [Ramsey’s \(2020\)](#) study spoke directly about the impact of unstable employment and how relational support differentiates sustainable from unsuccessful experiences of work:

“Basically, I’ve always been in and out of work, this is what people don’t understand is quite hard. The only reason I’m doing really well at the moment in my job is my team. If I didn’t have my team, wow.” (p. 133)

Independence in the sense of operating without support is highly valued in society. People are assumed to be self-sufficient and “significant discursive and material pressure is exerted on encouraging them to become so” ([Davy, 2019](#), p. 105). This ideal is reflected in policy and practice for people with intellectual and developmental disabilities in England ([Department of Health and Social Care, 2014](#); [NHS England, 2019](#)). However, feminist thinkers argue that independence, autonomy and self-sufficiency – for any person – are “foundational myths” ([Fineman, 2013](#), p. 13), designed to keep human relations of dependence and interdependence in the private sphere. Employment is framed in line with the ideal of the employee as an autonomous, self-made individual without caregiving responsibilities or support needs. This is particularly problematic for people with intellectual and developmental disabilities, whose need for relational support may not be so easily confined to the private sphere. [Davy \(2019\)](#) argues the need to rethink concepts of independence and care, maintaining that relationships are key to the development and exercise of agency and that this reality is heightened for people with significant support needs. This argument implies that interdependent relationships in work – including dedicated personal support and reciprocal support among co-workers – can extend the capabilities and agency of people with intellectual and developmental disabilities, allowing them to contribute in the public sphere of work. Interdependence is a characteristic feature of human culture - collaboration and collaborative learning are key drivers of cultural evolution ([Tomasello, 2016](#)). Interdependence generates social solidarity - shared experience has a strong and sustained impact on social closeness between those involved ([Wolf and Tomasello, 2020](#)). It is known that personal contact with people with intellectual and developmental disabilities specifically is associated with support for their rights and capabilities ([McConkey et al., 2020](#)). This suggests that the loss of opportunities for sustained work with others, due to the way that employment is framed, is a fundamental disadvantage, distinct from the disadvantage of not earning a wage, because it precludes processes important for extending the capabilities and inclusion of marginalized people.

This chapter includes examples of the importance to people with intellectual and developmental disabilities of working and receiving support in work, but meaningful social inclusion cannot depend solely on the granting of support by non-disabled people. Greater acceptance of mutual dependence and interdependence and wider opportunities for collaborative work might give scope for people with intellectual and developmental disabilities to build agency through their interactions with others. Recent research gives an example where expectations of lack of competency in the workplace rouse an individual’s commitment to show otherwise. The participant talks about being doubted at work:

"Then it seems like, "You can't handle it" ... "Well, sorry pal, but that? You can't do that." That hurts me ... when someone tells me, "Someone else will do that," then I think, "No! Heck, no! Absolutely not. I'll do it myself, even if it takes me three or four hours" And when it's finished three or four hours later, they come back and say, "Is it taken care of?" ... Then they stand there like, "Huh? Okay, you can do it after all." Yeah, that gives you a kick. ... that's the confidence that you can inspire in someone, and that someone can inspire in you!"

Voermans et al. (2021, p. 243)

This account tallies with Billett's (2008) argument about the interdependence between personal and social agency that operates in work situations. He emphasizes how individuals exercise agency through their decisions on the degree and nature of their engagement in learning and work. A further example of positive interdependence is given by a young person with intellectual disabilities motivated by opportunities in her work to give support. She understands how she herself gains by giving personal attention to the elderly people with whom she works:

"you have time for that resident, and that's what I really like about it ... hmm, contact, just having a chat, for the resident, that in itself means a whole lot and, yeah, for myself, well, I think like, "Now I've done something useful again," just because I also know that those residents get hardly any attention, and then I can do that myself, and they like it a lot."

Voermans et al. (2021, p. 245).

Voices such as these suggest ways in which interdependence in working situations can result in dynamics that are more reciprocal than the granting of inclusion by nondisabled people.

In summary, the numerous initiatives to promote employment for people with intellectual and developmental disabilities that have taken place in recent decades have had a very limited impact on employment statistics (Hatton et al., 2016), and particularly have failed to reach people with "moderate" and higher levels of support need (e.g., Stevens and Harris, 2013). Focusing in policy on employment and whether people count as in or out of employment or as "economically inactive" has proved to be counterproductive, limiting the capabilities of people with intellectual and developmental disabilities and their possibilities for social inclusion.

Conclusion

Formal employment processes are inherently discriminatory; globally, employment is less accessible to women and to disabled people. The high social and policy value placed on this specific form of work compounds the disadvantage of both groups. Indeed, "employment-focused" policy may be poorly compatible with the goal of social inclusion for many disabled people. There is a need to recognize the inflexible, contingent and situated nature of that market, so that employment is not held as the ultimate object and failure to be employed as personal failure. The social learning value of work more broadly conceived may have greater potential to extend the capabilities of disabled people and to promote greater interdependence between people with and without disabilities.

References

- Abberley, P., 1999. The significance of work for the citizenship of disabled people. In: Paper Presented at University College Dublin, 15th April. Available online at: www.leeds.ac.uk/disability-studies/archiveuk/index.
- Bambra, C., Eikemo, T., 2009. Welfare state regimes, unemployment and health: a comparative study of the relationship between unemployment and self-reported health in 23 European countries. *J. Epidemiol. Commun. Health* 63 (2), 92–98.
- Banks, M.L., Polack, S., 2014. The economic costs of exclusion and gains of inclusion of people with disabilities: evidence from low and middle income countries. In: International Centre for Evidence in Disability and London School of Hygiene and Tropical Medicine. London, UK.
- Bartelheimer, P., Verd, J.M., Lehwel-Litzmann, R., López-Andreu, M., Schmidt, T., 2012. Unemployment, intervention and capabilities. A comparative study of Germany and Spain. *Transfer* 18 (1), 31–44. <https://doi.org/10.1177/1024258911431199>.
- Bell, D., Heitmueller, A., 2009. The Disability Discrimination Act in the UK: helping or hindering employment among the disabled? *J. Health Econ.* 28 (2), 465–480. <https://doi.org/10.1016/j.jhealeco.2008.10.006>.
- Beyer, S., Meek, A., Davies, A., 2016. Supported work experience and its impact on young people with intellectual disabilities, their families and employers. *Adv. Ment. Health Intellect. Disabil.* 10 (3), 207–220. <https://doi.org/10.1108/AMHID-05-2014-0015>.
- Billett, S., 2008. Learning throughout working life: a relational interdependence between personal and social agency. *Br. J. Educ. Stud.* 56 (1), 39–58. <https://doi.org/10.1111/j.1467-8527.2007.00394.x>.
- Blamires, K., 2015. A summary of government initiatives relating to employment for people with learning disabilities in England. *Learn. Disabil. Rev.* 20 (3), 151–165. <https://doi.org/10.1108/TLDR-08-2014-0027>.
- Burchardt, T., 2004. Capabilities and disability: the capabilities framework and the social model of disability. *Disabil. Soc.* 19 (7), 735–751. <https://doi.org/10.1080/0968759042000284213>.
- Channon, A., 2014. Intellectual disability and activity engagement: exploring the literature from an occupational perspective. *J. Occup. Sci.* 21 (4), 443–458. <https://doi.org/10.1080/14427591.2013.829398>.
- Corbière, M., Lecomte, T., Reinharz, D., Kirsh, B., Goering, P., Menear, M., Berbiche, D., Genest, K., Goldner, E.M., 2017. Predictors of acquisition of competitive employment for people enrolled in supported employment programs. *J. Nerv. Ment. Dis.* 205 (4), 275–282.

- Davy, L., 2019. Between and ethic of care and an ethic of autonomy: negotiating relational autonomy, disability, and dependency. *Angelaki* 24 (3), 101–114. <https://doi.org/10.1080/0969725X.2019.1620461>.
- Dearing, K., 2020. "Not worth the minimum wage?" Unpacking the complexities of intellectual disability and its intersection with employment structures. *Scand. J. Disabil. Res.* 22 (1), 360–370. <https://doi.org/10.16993/sjdr.729>.
- Department of Health and Social Care, 2014. Care and support statutory guidance, updated 2018. Available at: <https://www.gov.uk/government/publications/care-act-statutory-guidance/care-and-support-statutory-guidance>. (Accessed 8 February 2019).
- Evans, T., 2019. Disability Pay Gaps in the UK: Earnings and Employment for Disabled and Non-disabled People in the UK, Raw Disability Pay Gaps and Factors that Affect Pay for Disabled People. ONS, London.
- Ferguson, A., Hennessy, R., Nagel, M., 2016. Feminist perspectives on class and work. *Stanf. Encycl. Philos.* <https://doi.org/10.5860/choice.41sup-0181>.
- Fineman, M.A., 2013. Cracking the foundational myths: independence, autonomy, and self-sufficiency. *Am. Univ. J. Gen. Soc. Pol. Law* 179–192. <https://doi.org/10.7591/9781501724077-013>.
- Gheaus, A., Herzog, L., 2016. The goods of work (other than money!). *J. Soc. Philos.* 47 (1), 70–89. <https://doi.org/10.1111/josp.12140>.
- Giri, A., Aylott, J., Giri, P., Ferguson-Wormley, S., Evans, J., 2021. Lived experience and the social model of disability: conflicted and inter-dependent ambitions for employment of people with a learning disability and their family carers. *Br. J. Learn. Disabil.* 1–9. <https://doi.org/10.1111/bld.12378>.
- Grech, S., 2008. Disability, poverty and development: critical reflections on the majority world debate. *Disabil. Soc.* 24 (6), 771–784.
- Grover, C., Piggott, L., 2015. Disabled People, Work and Welfare: Is Employment Really the Answer? Policy Press, Bristol. <https://doi.org/10.1080/09687599.2016.1198554>.
- Gunderson, M., Lee, B.Y., 2016. Pay discrimination against persons with disabilities: Canadian evidence from PALS. *Int. J. Hum. Resour. Manag.* 27 (14), 1531–1549. <https://doi.org/10.1080/09585192.2015.1072106>.
- Hatton, C., Glover, G., Emerson, E., Brown, I., 2016. Learning Disabilities Observatory: People with Learning Disabilities in England 2015: Main Report. Public Health England, London.
- ILO, 2019. Making the future of work inclusive of people with disabilities. In: Fundación ONCE and the ILO Global Business and Disability Network, pp. 1–42. Available at: http://www.businessanddisability.org/wp-content/uploads/2019/11/PDF_acc_FoW_PwD.pdf.
- Jones, M.K., 2008. Disability and the labour market: a review of the empirical evidence. *J. Econ. Stud.* 35 (5), 405–424. <https://doi.org/10.1108/01443580810903554>.
- Juvonen, J., Lessard, L.M., Rastogi, R., Schacter, H.L., Smith, D.S., 2019. Promoting social inclusion in educational settings: challenges and opportunities. *Educ. Psychol.* 54 (4), 250–270. <https://doi.org/10.1080/00461520.2019.1655645>.
- Laruffa, F., 2020. What is a capability-enhancing social policy? Individual autonomy, democratic citizenship and the insufficiency of the employment-focused paradigm. *J. Hum. Dev. Capabil.* 21 (1), 1–16. <https://doi.org/10.1080/19452829.2019.1661983>.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Lysaght, R., Ouellette-Kuntz, H., Lin, C.J., 2012. Untapped potential: perspectives on the employment of people with intellectual disability. *Work* 41 (4), 409–422.
- Macnicol, J., 2011. The history of work-disability. In: Lunday, C., Houston, D. (Eds.), *Disability Benefits, Welfare Reform and Employment Policy*. Palgrave Macmillan, London, pp. 33–53. <https://doi.org/10.1179/174582311X13129418298893>.
- Manuti, A., Pastore, S., Scardigno, A.F., Giancaspro, M.L., Morciano, D., 2015. Formal and informal learning in the workplace: a research review. *Int. J. Train. Dev.* 19 (1), 1–17. <https://doi.org/10.1111/ijttd.12044>.
- Matusov, E., 2001. Vygotskij's theory of human development and new approaches to education. In: Smelser, N.J., Baltes, P.B. (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, Amsterdam, pp. 16339–16343. <https://doi.org/10.1016/b0-08-043076-7/02407-4>.
- McConkey, R., Slater, P., Smith, A., Dubois, A., Shellard, A., 2020. Perceptions of the rights and capabilities of people with intellectual disability in the United States. *J. Appl. Res. Intellect. Disabil.* 1–9. <https://doi.org/10.1111/jar.12819>.
- Meager, N., Higgins, T., 2011. Disability and Skills in a Changing Economy. Briefing Paper Series. UK Commission for Equality and Skills. <https://www.employment-studies.co.uk/resource/disability-and-skills-changing-economy>.
- Meekosha, H., 2011. Decolonising disability: thinking and acting globally. *Disabil. Soc.* 26 (6), 667–682. <https://doi.org/10.1080/09687599.2011.602860>.
- Meltzer, A., Bates, S., Robinson, S., Kayess, R., Fisher, K.R., Katz, I., 2016. What Do People with Intellectual Disability Think about Their Jobs and the Support They Receive at Work? A Comparative Study of Three Employment Support Models. Final report (SPRC Report 16/16). Social Policy Research Centre, UNSW, Australia, Sydney, NSW.
- Merrells, J., Buchanan, A., Waters, R., 2019. "We felt left out": experiences of social inclusion from the perspective of young adults with intellectual disability. *J. Intellect. Dev. Disabil.* 44 (1), 13–22. <https://doi.org/10.3109/13668250.2017.1310822>.
- Meyer, B.D., Mok, W.K.C., 2019. Disability, earnings, income and consumption. *J. Publ. Econ.* 171, 51–69. <https://doi.org/10.1016/j.jpubeco.2018.06.011>.
- Mitra, S., 2006. The capability approach and disability. *J. Disabil. Pol. Stud.* 16 (42), 236–247. <https://doi.org/10.1080/0305764x.2012.706393>.
- Mitra, S., Kruse, D., 2016. Are workers with disabilities more likely to be displaced? *Int. J. Hum. Resour. Manag.* 27 (14), 1550–1579. <https://doi.org/10.1080/09585192.2015.1137616>.
- Mizunoya, S., Mitra, S., 2013. Is there a disability gap in employment rates in developing countries? *World Dev.* 42 (1), 28–43. <https://doi.org/10.1016/j.worlddev.2012.05.037>.
- Modini, M., Tan, L., Brinchmann, B., Wang, M.J., Killackey, E., Glozier, N., Mykletun, A., Harvey, S.B., 2016. Supported employment for people with severe mental illness: systematic review and meta-analysis of the international evidence. *Br. J. Psychiatry* 209 (1), 14–22. <https://doi.org/10.1192/bjp.bp.115.165092>.
- Moore, K., McDonald, P., Bartlett, J., 2018. Emerging trends affecting future employment opportunities for people with intellectual disability: the case of a large retail organisation. *J. Intellect. Dev. Disabil.* 43 (3), 328–338. <https://doi.org/10.3109/13668250.2017.1379250>.
- Morris, J., 2007. Impairment and disability: constructing an ethics of care that promotes human rights. *Hypatia* 16 (4), 1–16. <https://doi.org/10.2979/hyp.2001.16.4.1>.
- NHS England, 2019. Giving choice and independence. Available at: <https://www.england.nhs.uk/blog/giving-choice-and-independence/>. (Accessed 10 March 2019).
- OECD, 2008. *Sickness, Disability and Work: Breaking the Barriers: A Synthesis of Findings across OECD Countries*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264088856-en>.
- OECD, 2019. *The Future of Work: OECD Employment Outlook*. OECD Publishing, Paris. <https://doi.org/10.1787/2e9fb244-nl>.
- OECD, 2020. *OECD Employment Outlook 2020: Worker Security and the Covid-19 Crisis*. OECD Publishing, Paris. <https://doi.org/10.1787/1686c758-en>. (Accessed 14 July 2020).
- ONS, 2019. Employment in the UK: June 2019. ONS, London. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/disability/bulletins/disabilityandemploymentuk/2019>.
- Public Health England, 2020. Research and analysis Chapter 2: Paid employment. *People with Learning Disabilities in England*. gov.uk. <https://www.gov.uk/government/publications/people-with-learning-disabilities-in-england/chapter-2-employment>. (Accessed 14 July 2020).
- Ramsey, D., 2020. *The Lived Experience of Work for People Who Have an Intellectual and Developmental Disability*. PhD thesis. University of Brighton.
- Redley, M., 2009. Understanding the social exclusion and stalled welfare of citizens with learning disabilities. *Disabil. Soc.* 24 (4), 489–501. <https://doi.org/10.1080/09687590902879122>.
- Reindal, S.M., 2016. Discussing inclusive education: an inquiry into different interpretations and a search for ethical aspects of inclusion using the capabilities approach. *Eur. J. Spec. Needs Educ.* 31 (1), 1–12. <https://doi.org/10.1080/08856257.2015.1087123>.
- Robeyns, I., 2016. Capabilities approach. In: Zalta, E.N. (Ed.), *Stanford Encyclopedia of Philosophy*. <https://doi.org/10.5860/choice.41sup-0181>.
- Robeyns, I., 2017. *Wellbeing, Freedom and Social Justice*. Open Book Publishers, Cambridge. <https://doi.org/10.11647/OBP.0130>.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press, Oxford.
- Rogoff, B., Callanan, M., Gutiérrez, K.D., Erickson, F., 2016. The organization of informal learning. *Rev. Res. Educ.* 40 (1), 356–401. <https://doi.org/10.3102/0091732X16680994>.

- Sainsbury, R., 2018. Labour market participation of persons with disabilities – how can Europe close the disability employment gap? In: Wansing, G., Welti, F., Schäfers, M. (Eds.), *The Right to Work for Persons with Disabilities*. Nomos Verlagsges. MBH & Co., Baden-Baden, pp. 135–154. <https://doi.org/10.5771/9783845291673-134>.
- Schur, L., Han, K., Kim, A., Ameri, M., Blanck, P., Kruse, D., 2017. Disability at work: a look back and forward. *J. Occup. Rehabil.* 27 (4), 482–497. <https://doi.org/10.1007/s10926-017-9739-5>.
- Schur, L.A., 2003. Barriers or opportunities? The causes of contingent and part-time work among people with disabilities. *Ind. Relat.* 42 (4), 589–622. <https://doi.org/10.1111/1468-232X.00308>.
- Sen, A., 1992. *Inequality Re-examined*. Harvard University Press, Cambridge, MS.
- Sen, A., 1999. *Development as Freedom*. Oxford University Press, Oxford.
- Shakespeare, T., 2006. *Disability Rights and Wrongs*. Routledge, London.
- Slee, R., 2019. Belonging in an age of exclusion. *Int. J. Incl. Educ.* 23 (9), 909–922. <https://doi.org/10.1080/13603116.2019.1602366>.
- Standing, G., 2009. *Work after Globalization: Building Occupational Citizenship*. Edward Elgar, Cheltenham and Northampton, MA.
- Stevens, M., Harris, J., 2013. *Jobs First Evaluation: Final Report*. King's College London, Social Care Workforce Research Unit, London.
- Susskind, D., 2020. *A World without Work: Technology, Automation and How We Should Respond*. Penguin, London.
- Thomas, C., 2004. Rescuing a social relational understanding of disability. *Scand. J. Disabil. Res.* 6 (1), 22–36. <https://doi.org/10.1080/15017410409512637>.
- Tomasello, M., 2016. Cultural learning redux. *Child Dev.* 87 (3), 643–653. <https://doi.org/10.1111/cdev.12499>.
- UNCPRD, 2006. *Convention on the Rights of Persons with Disabilities and Optional Protocol*. United Nations General Assembly, New York, NY.
- United Nations, 2016. *Leaving No One behind: The Imperative of Inclusive Development*. Report on the World Social Situation 2016. United Nations, New York, NY.
- Voermans, M.A.C., Taminiau, E.F., Giesbers, S.A.H., Embregts, P.J.C.M., 2021. The value of competitive employment: in-depth accounts of people with intellectual disabilities. *J. Appl. Res. Intellect. Disabil.* 34 (1), 239–249. <https://doi.org/10.1111/jar.12802>.
- Vornholt, K., Villotti, P., Muschalla, B., Bauer, J., Colella, A., Zijlstra, F., Van Ruitenbeek, G., Uitdewilligen, S., Corbière, M., 2018. Disability and employment—overview and highlights. *Eur. J. Work. Organ. Psychol.* 27 (1), 40–55. <https://doi.org/10.1080/1359432X.2017.1387536>.
- Watts, M., Comim, F., Ridley, B., 2008. Adaptive preferences and educational policies. In: *Human Development and Capability Association Annual Conference*. New Delhi, p. 22. https://doi.org/10.1007/978-3-642-38376-2_10.
- Wehman, P., Taylor, J., Brooke, V., Avellone, L., Whittenburg, H., Ham, W., Brooke, A.M., Carr, S., 2018. Toward competitive employment for persons with intellectual and developmental disabilities: what progress have we made and where do we need to go. *Res. Pract. Persons Severe Disabil.* 43 (3), 131–144. <https://doi.org/10.1177/1540796918777730>.
- Wolf, W., Tomasello, M., 2020. Human children, but not great apes, become socially closer by sharing an experience in common ground. *J. Exp. Child Psychol.* 199, 104930. <https://doi.org/10.1016/j.jecp.2020.104930>.
- World Health Organisation (2021) Musculoskeletal conditions. Available at <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions> (Accessed: 28 May 2021).
- World Health Organization and The World Bank, 2011a. *Summary World Report on Disability*. Available at: <https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/world-report-on-disability>.
- World Health Organization and The World Bank, 2011b. *World Report on Disability*. Available at: https://www.who.int/disabilities/world_report/2011/report.pdf.
- Yeoman, R., 2014. Conceptualising meaningful work as a fundamental human need. *J. Bus. Ethics* 125 (2), 235–251. <https://doi.org/10.1007/s10551-013-1894-9>.

Disability and poverty

Rebecca L. Porter and Josephine Sirotkin, Center for Disability Studies, University of Leeds, Leeds, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	462
Measuring disability and poverty: more than limitations and lines	462
Disability and poverty: why so many?	464
The politics of poverty	465
In their own words: the lived experience of disability and poverty in the UK	466
Conclusion	467
Acknowledgment	467
References	467

Introduction

Poverty and disability have a continuously, intertwining relationship that largely stems from the systematic inequalities disabled people experience. This article aims to explore this relationship, contextualizing it with secondary data and literature analysis. First, it will critically examine how different institutions conceptualize and measure both poverty and disability. Second, it will focus on how disabled people are largely over-represented in poverty data, and some of the potential reasons for this. Following on from this, the third section of this article examines the political decisions that have led to the over-representation of disabled people in poverty in the UK. In doing so, we argue that poverty is a political choice that has led to many disabled people being denied their human rights. The fourth section highlights this further with stories and interviews gathered from [Ryan \(2019\)](#) that demonstrate the dehumanizing impacts of poverty on disabled people's lived experiences. Finally, this article concludes by reasserting the significance of the relationship between poverty and disability, while maintaining that poverty is indeed political.

Measuring disability and poverty: more than limitations and lines

Both disability and poverty have been contested categories, with multiple definitions. This is because there is no one definitive or standardized way to decide who is disabled, nor who is living in poverty. Measuring disability and poverty therefore requires subjective, political decisions to be made, and so any statistical data that is created cannot be neutral. Nevertheless, statistics play an important role within contemporary society and categorizations are employed to make different social groups comparable with one another, particularly in terms of social inequalities ([Mau, 2019](#)). As a result, different aspects of social life become quantified and numbers are used to "establish and ascribe social worth" ([Mau, 2019](#), p. 19). There is a concern, however, that metrics are used to rationalize certain behaviors, while demonstrating the social order, which is controlled by perpetual counting and categorizing of persons, as a form of governance ([Mau, 2019](#), p. 19). This can create problems for certain groups that experience inequality, such as disabled people, as they may not fit with the fixed formula of what their society values the most ([Reckwitz, 2020](#)). Those who are in poverty and/or disabled find themselves subject to unstable and turbulent categorization ([Stone, 1984](#)).

There are multiple definitions and ways of measuring poverty, using a variety of deprivation indicators, including: income, food, fuel, and well-being ([Alkire et al., 2015](#)). Although the importance of a multidimensional poverty measurement is recognized ([Alkire et al., 2015](#)), income-related poverty measures are still most prevalent ([UN, 2018](#)). Income poverty draws upon the idea of a fixed poverty line, whereby a threshold of income and expenditure is set in order to determine the number of people living in poverty ([Alkire et al., 2015](#)). This creates a threshold that represents how much income a household needs to be able to afford the basic necessities to meet their physical needs ([Alkire et al., 2015](#)). Those who fall below this threshold are considered to be living in absolute poverty or destitution ([Alkire et al., 2015](#)).

Taking this threshold approach, the UK government traditionally relies upon just one measurement of median income to create their statistical analysis of how many people live in poverty, both before and after housing costs, such as rent and mortgage payments ([JRF, 2016a](#)). The idea of the poverty line, however, has been criticized for not accounting for the full complexities of the lives of impoverished households ([O'Connell et al., 2019](#)). Relative poverty is considered a better alternative by some social scientists, as it accounts for both the social and physical effects of poverty, including, for example, not being able to offer a visitor a cup of tea ([O'Connell et al., 2019](#)). To exemplify, in their work toward solving poverty in the UK, the Joseph Rowntree Foundation (JRF) draws upon a notion of relative poverty to understand how many impoverished households there are ([Goulden, 2020](#)). Typically, the JRF focuses upon income poverty after housing costs, and their poverty line/threshold is set at those who are below 60% of median income, which is then supplemented with an indicator based on the Minimum Income Standard (MIS) ([Goulden, 2020](#)). The MIS is the amount of income needed to afford to meet minimum needs, such as being able to pay for food, shelter, warmth, and

other essentials such as toiletries, or take a shower. Anyone who goes completely without these things is considered destitute (JRF, 2016a, 2020).

This model of poverty allows for an understanding that begins to recognize the variety of lived experiences for those within poverty. The JRF (2020) employs the MIS to identify three different levels of poverty and with each level there is increasing financial hardship. Here, anyone who falls below the MIS is considered to be in poverty and anyone above can afford a decent standard of living, or better. The first level of poverty consists of those whose income is below the MIS, but is not less than 75% (JRF, 2020). Though people living in this level of poverty can often manage day-to-day, they are at risk of being pushed further into poverty by unexpected expenses and events (JRF, 2020). The second level of poverty consists of people whose income is under 75% of the MIS, and, as a result, there is a high chance that their needs are not being met and they have a poor standard of living (JRF, 2020). The third level, representing those who are deepest in poverty, is destitution (JRF, 2020). The JRF defines destitution as being when a person has lacked two or more essentials (such as, shelter or food) within a month, or when they have an income of less than £70 per week (Goulden, 2018).

However, a key issue with poverty measures more generally, is that governments can alter how they choose to measure poverty. Commonly used measures of poverty draw upon income and basic needs to determine the proportion of people in poverty (Alkire et al., 2015). Yet, looking at the proportion of people in poverty is problematic as a “decrease of any size in the income (or unmet basic needs) of a poor person paired with a corresponding increase for a non-poor person will leave poverty unchanged”, and vice versa (Alkire et al., 2015, p. 4). For example, in the UK, child poverty was shown to drop during the 2008 financial crisis, but on close inspection, this was because the recession had decreased overall household income (Ryan, 2019). More recently, the UK Conservative government altered the measurements of child poverty by scrapping the 60% median income measure, which saw a “fall” of people in poverty. Yet, critics of the proposed changes noted that child poverty was set to increase due to austerity measures, and the Conservative government was simply moving the goal posts, for example, by refusing to consider household income at all (Stewart and Roberts, 2019).

Similarly, governments can also alter how they choose to classify and measure disability. As such, decisions about who counts (and is counted) as a disabled person have changed over time (see, for example, DWP, 2020a). The ways in which disability is currently defined and understood have been profoundly influenced by the social model of disability (Oliver, 1990; Barnes, 2012). Originating in the 1970s within the UK’s Disabled People’s Movement (UPIAS, 1976), this model sought to redefine disability as a specific form of socially created disadvantage and marginalization that occurs to people with “particular mind and body traits” (also known as impairments) (Lawson and Beckett, 2020, p. 1). As such, disability is no longer framed as an individual problem requiring medical or technological solutions, rather it is “one of tackling disadvantage or exclusion,” which “necessarily requires different institutional expertise” (Lawson and Beckett, 2020, p. 20). This can allow for potential policy reforms to be more easily identifiable (Lawson and Beckett, 2020), thus could be particularly fruitful when considering the over-representation of disabled people in poverty. Although the distinction between disability and impairment is not always maintained, the social model’s reframing of disability has had a profound impact on governmental bodies and policies, perhaps most notably with its influence in the development of the United Nations Convention on the Rights of People with Disabilities—2006 (UNCRPD) (Lawson and Beckett, 2020).

Despite the influence of the social model, the distinction between disability and impairment can collapse when we begin to look at poverty statistics. For instance, in the UK, the Department for Work and Pensions (DWP) identify disabled people as:

Those who report any physical or mental health condition(s) or illness(es) that last or are expected to last 12 months or more, and which limit their ability to carry out day-to-day activities a little, or a lot.

DWP (2020a, p. 11).

Here, rather than defining disability consistent with the social model, the DWP identifies disabled people based on their impairment effects (see Thomas, 1999). This may influence who does and does not identify as disabled. Moreover, different individuals may have different interpretations of what it means for a condition to “limit their ability to carry out day-to-day activities,” which, again, may affect self-identification (DWP, 2020a, p. 11). This individual model approach to disability, which places limitations as resulting from “physical or mental health condition(s) or illness(es)” (DWP, 2020a, p. 11), is not limited to statistics within the UK; such an approach also appears within welfare policy (DWP, 2020b). Thus, policy purported to mitigate the barriers to employment and the extra costs that disabled people can face are equally subjective.

Judgments regarding who is entitled to financial support for extra costs, through Personal Independence Payments (PIP), are based on external, subjective assessments of the “severity” of impairment and its impact on function (Gray, 2017). Consequently, these judgments are subject to change. For instance, a judgment was made that being “unable to undertake journeys due to ‘overwhelming psychological distress’” did not warrant financial support with mobility, through PIP (Kennedy, 2018, no pagination). This meant that approximately 220,000 disabled people were denied the support they needed (Kennedy, 2018). After a High Court judgment deemed this to be unlawful discrimination, and in violation of human rights, this decision was overturned (Kennedy, 2018). Hence, it is important to consider the precarious nature of governmental support when examining the relationship between disability and poverty. Not all disabled people will receive the financial support necessary to mitigate the costs of oppression, and therefore may be more likely to live in poverty than non-disabled people. Perhaps unsurprisingly, then, disabled people are

extremely over-represented in impoverished populations, irrespective of the type of impairment or living standard indicator that is used (Morris and Zaidi, 2020). The next section of this chapter identifies and explains the reasons why this may be the case, with a particular focus on extra costs and reliance on social security.

Disability and poverty: why so many?

The over-representation of disabled people in poverty stems from both a disabling society and the social determinants of ill-health. Although this article focuses on the former, it is important to note that poverty has been found to be “closely associated with most causes of impairment,” even those that appear to be hereditary or caused by natural disasters (Swain and French, 2014, p. 48). These causes of impairment, which can stem from poverty, include: limited access to healthcare; poor nutrition; poor hygiene and sanitation; inadequate housing; and dangerous working conditions (Swain and French, 2014). This may help to explain why there is a higher prevalence of disability/impairment in low-income countries (World Bank, 2020).

The struggle of disabled people being trapped in poverty, thus, extends globally. Disabled people, in all forms of poverty measurement (hunger indicators, traditional poverty income, household expenditures) are much more likely to be in poverty than non-disabled people (UN, 2018). Additionally, disabled people are much more likely to be experiencing multidimensional deprivation, in areas such as housing, education or healthcare. In high-income countries, such as the UK, income poverty was found to be the most prevalent (UN, 2018). Whereas, in low-income countries the picture was more varied. Although households with disabled members tended to have fewer assets and worse living conditions compared to other households, some studies found no significant association between disability and living conditions (UN, 2018). This may be because the income gap in low-income countries can be less extreme between households with disabled members and households without (Shakespeare, 2012). Nonetheless, this income may still need to stretch further, as disability-related costs can still be a factor (Shakespeare, 2012). In addition, in countries where there is no universal healthcare system, impairment-related expenses are also a factor, as healthcare costs can be a significant expense (Shakespeare, 2012). Consequently, this could increase the likelihood of disabled people living in poverty in these countries.

Out of 41 higher income countries, the UK was ranked 16th on tackling poverty, 34th on food insecurity, 15th on health and wellbeing, 31st on economic growth, and 6th on reducing inequalities (UNICEF, 2018). Even though disabled families only make up 34% of the total population in the UK, they make up 40% of the number of people living with relative low income in 2018/19 (Francis-Devine, 2020). Yet, these figures do not account for the additional living costs that disabled people can face (Francis-Devine, 2020). Typically, households that have a disabled person in them face additional costs toward their everyday living. These extra costs can range from paying for mobility aids or house adaptations, to paying for support, care, or even fuel costs. These account for, on average, £583 of additional costs; these extra costs are equal to just over 50% of their total income, even if the disabled person has access to extra cost benefits (Scope, 2019). Over time, research has continually shown that these extra costs are not fully covered by the government support available (Morris and Zaidi, 2020). As the benefits that are purported to cover these additional costs are included as income in the statistics above, a more accurate picture of the amount of disabled people living in poverty may be gained by adjusting the measure (Francis-Devine, 2020). Removing these benefits from our measure of income shows that “people living in families where someone is disabled made up around 47% of the total population in relative low income” before housing costs and 46% after housing costs (Francis-Devine, 2020, p. 29).

By providing an income, employment can offer some protection from poverty, although it does not entirely prevent it (Bourquin et al., 2019; JRF, 2016b). Despite this, the UK government consistently places emphasis on employment as the main route out of poverty (see DWP and DE, 2011). This is particularly problematic as employment rates are dependent upon the needs of the market (*capitalism*) and the government (Grover and Piggott, 2005). Moreover, disabled people have lower employment rates than non-disabled people, due to a number of barriers (Grover and Piggott, 2005). Some of these disabling barriers may be reinforced by poverty as having access to money can help to alleviate some of their effects. For instance, barriers to employment, such as inaccessible application forms, can be mitigated when personal assistance can be financed. Seemingly, in recognition of these extra costs, the government provides some financial support for disabled people seeking work through Employment and Support Allowance (ESA) and Universal Credit. Yet, there have been multiple reports that have highlighted that this support is insufficient, unfit for purpose, and even actively harmful (Dwyer et al., 2016; UN, 2016).

Furthermore, one in three children in the UK are in what UNICEF calls “multi-dimensional poverty”—which measures deprivation in several areas linked to children’s rights, including: housing, clothing, nutrition and access to social and leisure activities (UNICEF, 2018). Previously, the JRF (2016a) found that children who were labeled as having special needs were more likely to experience poverty than their non-disabled peers and were less likely to leave school with qualifications and outcomes that would help to prevent poverty from reproducing itself for them as adults. Thus, for children, the reproduction of poverty through generations can manifest itself within the education system, where so-called inclusive education further marginalizes disabled students (Slee, 2014). Despite there being different intersectional differences, many disabled people find their lives are broadly impacted by deprivation (Singal, 2011). Formal education can be a way for disabled people to access a better quality of life. However, disabled children have limited access to education, which is further impacted by the cyclical recreation of poverty across generations, where disability, poverty and education share a connection (Singal, 2011). This is supported by the JRF (2020) who estimate that one-third of disabled children in the UK live below the poverty line. This is attributed largely to changes to the welfare system, where families with disabled children were estimated to have lost around £1750 a year.

In relation to this, the DWP estimates that between 2018 and 2019 £400 million went unclaimed, due to either DWP or claimant error. Claimant error is defined as occasions where benefit claimants did not provide accurate enough information or did not report changes to be reassessed (DWP, 2019). As a result, 84% of those being underpaid by the DWP were those claiming for support with extra costs through PIP (DWP, 2019). This is noted by several independent reviews of PIP, which state these figures could be due to an inaccessible, long, and confusing assessment process for claimants (Gray, 2017). Other recommendations for PIP included: making clearer what claimant responsibilities are; ensuring evidence given by carers or appointees of the claimant are given more weight in the decision-making process; transparency of why a decision is made; and, regular auditing to ensure consistency across assessments (Gray, 2017). However, some claimants still find themselves distressed, frustrated, and worse off after applying for benefits. The War on Welfare (WOW) campaign highlights this with its collection of testimonies from welfare claimants, including Helen, who states:

... My latest PIP Assessment is the closest to breaking me, the extended torture of pushing through the Appeals Process after a woefully inadequate Assessors Report left me £450 worse off a month and my management and care regime in tatters. Deliberate institutional harm by the Government and its lackeys.

WOW (2018, no pagination).

Another claimant, Gary, talks about the loss of his wife, and still being expected to attend his assessment for his work capability assessment review in order to receive financial assistance (related to employment barriers):

6 (six) days after my wife died, I duly turned up for the assessment. I went home and waited for the brown envelope to arrive, now bearing in mind that nothing had changed in my conditions since the first assessment, except for the small fact that I had lost my wife, I was told that I was being moved from the ESA support group to the work-related activity group and would consequently lose £30 a week. I asked for a mandatory reconsideration only to be told the decision was upheld. I could have appealed the decision, but to be completely honest I did not and do not have the energy or emotion to fight anymore. I guess that is what they hope for, that people at their most vulnerable, won't fight them.

WOW (2019, no pagination).

As shown in these examples, the relationship between disability and poverty runs deep within social institutions, such as the welfare state and education system. This is especially prevalent in countries where austerity measures have been in place to reduce government spending on social care and welfare programs (Ryan, 2019). For example, according to the UK foodbank charity the Trussell Trust (2020), in 2018–2019, the top three reasons for food bank referrals were: income not covering the cost of essentials (33.11%), benefit delays (20.34%), or benefit changes (17.36%). This indicates that food insecurity (a measure of poverty) is intricately wrapped up with the current welfare state system. Consequently, many people are relying upon charitable donations, in this case through food banks, to meet their most crucial needs. The following section examines the role that austerity has played in creating this level of poverty in the UK, and its significance to the relationship between disability and poverty is explored.

The politics of poverty

Poverty can be argued to be a political choice, where public funds are not used to help those in need. Instead, disabled people who claim welfare are blamed and vilified for the economic downturn (Benstead, 2019; Ryan, 2019). Within the context of the 2010 Liberal Democrat-Conservative Coalition, the UK government were under obligation to try to minimize impact of the recession on the poorest in society, through spending cuts (Benstead, 2019). This could have been achieved by making full use of all resources and making spending cuts fall on those who are financially better off. They also had the option of introducing measures, such as closing loopholes that have allowed large companies to avoid paying tax, or reducing tax breaks for the rich (Benstead, 2019). This could generate enough money to spend on essential government services. Instead, the government chose to give tax cuts to the most privileged in our society, while implementing cuts to welfare and social care that caused a reduction in human rights for people living in poverty or deprivation (Benstead, 2019).

Funding cuts to services, then, were justified by the government through notions of austerity and fairness (Hamnett, 2014). Under the conventions of the UNCRPD (2007), there is room for a country to regress in its progress on disabled people's human rights, if there is significant economic or political unrest, such as invasion, or a major natural disaster, and the government used this to justify austerity measures (Cooper and Whyte, 2017). For Cooper and Whyte (2017), austerity became a form of violence in three distinct stages of government deception. First, was the idea that came from the government that every person had played a role in the 2008 financial crisis. The second deception was that austerity was absolutely necessary and that catastrophe was imminent, unless harsh austerity measures were put in place. The third deception was the idea that "we are all in this together," implying that everyone in the country was to be equally impacted by austerity measures. Yet, this often repeated political phrase masks the fact that those hit hardest by austerity are targeted groups, such as working-class people and welfare claimants. It was through these deceptions that the government was able to avoid addressing social inequalities, and so, while the UK's top earners benefited from lowered tax rates, many people who needed assistance went without (Cooper and Whyte, 2017).

Thus, the real politics of austerity stabilizes the economy so that the rich sustain themselves, ultimately extending wealth by increasing inequality (Cooper and Whyte, 2017). At the same time, austerity has allowed for the disposition of rights by reducing and privatizing public services, while permanently dismantling the “protection state” (Cooper and Whyte, 2017). This occurred as the government utilized the panic around reducing the state deficit as an excuse to dismantle social welfare, which impoverished disabled people desperately relying on the support of the state as a main source of income (Cooper and Whyte, 2017). These changes to the welfare state have been popularized through discourses of “fairness,” which imply that needs-based social welfare has, in itself, become a source of unfairness (Hoggett et al., 2013). This is further supported by a long-standing anti-welfare rhetoric that positions benefit claimants as “scroungers” who abuse the system (McEnhill and Byrne, 2014; Briant et al., 2013; Garthwaite, 2011).

A significant proportion of the responsibility for benefit system “abuse” has been attributed to disabled people by the state (Hoggett et al., 2013). Consequently, the governmental category of “disability” has been reconfigured, along with the “rights and responsibilities of disabled people in relation to work and welfare” (Dwyer et al., 2016, p. 1). The process of reclassifying the criteria, of who is considered disabled, has tightened so that it only applies to the fictitious idea of the “truly” disabled (Garthwaite, 2014). Because of these changes, disabled people continue to be suspicious of the system, and feel depersonalized by the assessment process (Garthwaite, 2014). Those with invisible impairments, such as mental health issues, struggle to be recognized as disabled under PIP guidelines where participants are forced into black and white tick boxes on government forms, even though their lives are more shades of gray (Jamieson, 2020). This reduces their suffering, and the barriers they face in society, to a number on a form, casting them as abnormal. As Jamieson (2020, p. 35) states “... You can’t just put people into mad, bad, or sad.” Yet, the DWP attempt to do just that, with a box-ticking culture that stigmatizes claimants for being “bad citizens” when they are unable to provide for themselves financially (Jamieson, 2020).

Some have attributed the stigmatization and systematic discrimination faced by those receiving benefits to necropolitical governance (Mayblin et al., 2019). Necropolitics builds on biopolitics: where tools are used for the governance of certain population’s lives, to argue that necropolitics is the next step, where undesirable populations have their deaths governed by political violence (Mbembe, 2019). Biopolitics is an act of governance that allows states to use the population as a resource to build capital (Campbell, 2013). Therefore, biopower is the state’s power over life, as opposed to power over death. This governance ultimately controls certain behaviors or characteristics of the population, by using the standard of “the norm” (Campbell, 2013). Necropolitics argues society has transitioned into an even more violent and oppressive governance of populations, through the state’s use of political power and violence to subjugate oppressed peoples into the category of the living dead (Mbembe; Davies et al., 2017; Mayblin et al., 2019). Within these categories an individual or marginalized group does not live their life as a full citizen but is not yet physically dead. The latter idea is supported by Mbembe who argues the group experiences “death-in-life” (Mbembe, 2003, p. 21) that comes from states deliberately creating poor conditions for life. Citing the UK government’s gradual decrease in the amount of financial support asylum seekers receive, Mayblin et al. (2019) argue this results in systematic impoverishment, despite there being human rights laws that the UK ascribes to, which are meant to protect asylum seeker’s human rights. Necropolitical governance divides society into three categories: the living, the dead, and the living dead (Mbembe, 2019). The *living dead* are those who are not physically dead, but, because of political violence that pushes them into poverty, they experience a life not lived to an adequate standard, for example, not being able to pay for heating or food, because you lack the income from the welfare system to do so (Mayblin et al., 2019).

Porter (2018) extends this argument to consider disability extra cost benefits and the subsequent changes to the categories that govern acceptance or rejection. This, she argues, can be characterized as a politically violent form of social control over increasingly impoverished disabled people. The disabled welfare claimant, then, is part of the living dead, because they cannot have an adequate standard of living without access to the extra income PIP could provide them with. Specifically, over 7990 people died within 6 months of having their PIP claim rejected (DWP, 2019). Yet, we know that these decisions can be overturned, as approximately 72% of PIP appeals are won by claimants (Scope, 2019; in Porter, 2019). The application for benefits is also a very lengthy process; 17,000 people have died within the last 5 years while waiting for their PIP claim to go through (DWP, 2019). This means that potentially thousands of people have died before having access to the money that they needed to keep them out of poverty and ensure their human rights (Porter, 2019). It is important, therefore, to remember that disabled people’s poverty is a political choice. The social policies that exist within the UK have left disabled people more impoverished than in previous decades: a situation Ryan (2019, p. 40) describes as “near-Dickensian levels of hardship” in one of the world’s richest societies. The next section of this chapter will highlight the impact that this has had on disabled people’s lives.

In their own words: the lived experience of disability and poverty in the UK

Francis Ryan (2019) relates multiple stories of disabled people’s experiences of austerity in the UK. One harrowing case is that of Susan. Under government reforms to the social care system, Susan now must contribute to her own care costs, having to find £57 a week of her own money for this. Another £4 goes on a network alarm, in case she falls and needs emergency assistance. If she does not have the money, it will be turned off. Susan’s reduced benefits means she can rarely afford to buy food. Her bowel condition means she cannot eat solid food, but to get chemically pureed food costs £3.50 a time. She survives on porridge instead, as this is £2.00 a box for the week. She has lost four kilos in five years and tells Ryan (2019) she is starving.

Another story is that of former nurse, Racheal, who has had her entire social care package removed (Ryan, 2019). Without someone to help her into bed at night, she now sleeps fully clothed in her wheelchair. Five out of seven days she does not shower, eat regular meals, or get dressed, because there is no one to help her. She could not afford the extra payments that she is expected to make to access care, after she transferred to the new disability benefit, PIP, and her money was drastically reduced. Over the past year, she has been told she is clinically malnourished. When her arms are strong enough to lift a pot, she makes stew for the week. On her bad days, she survives on fruit and bread.

As explained earlier in this article, this inequality also extends to disabled children. Take this case, from Ryan (2019), for example:

Louis is 9 and is homeschooled. He has autism, sensory processing disorder and dyspraxia. He was excluded from mainstream school aged 6. He was labelled as a naughty child, and his exclusion came two weeks before his diagnostic appointment. This has impoverished the family, as his mother had to give up work meaning: "Things are hard financially ... we just about get by but have no money spare ... we live day to day."

Louis's mother, Joanna, cited in Ryan (2019, p. 185).

This is not unusual for parents of disabled children. 48% of parents surveyed by Contact (2014) had to either reduce their working hours or give up work completely. With a limited amount of financial support for carers, it is unsurprising that this can lead to poverty. Like the other benefits mentioned in this article, Carers Allowance has multiple conditions attached to it, meaning that not all carers will receive it (Gov.uk, no date). In addition, at the time of writing, the income it provides is just £67.25 a week (Gov.uk, no date), which is less than the amount of income needed to be considered destitute (Goulden, 2018).

Moreover, the experiences described here are not unique. Since the implementation of austerity measures in the UK, many disabled people have seen their income fall drastically (Ryan, 2019). The impact of this can be seen in the form of a loss of human rights for those affected (UN, 2016). Contact (2014) reported a worrying trend in families with disabled children increasingly going without essentials, such as: days out/leisure time (84%); clothes (65%); car fuel (40%); specialist equipment (36%); food (31%); and heating (33%). Going without had health implications for both disabled children and their parents, with 49% of those surveyed stating they have had ill-health as a result and 22% stating that their child's health has worsened (Contact, 2014).

Conclusion

This chapter demonstrates how the relationship between disability and poverty is a contentious and complex issue. The political and subjective decisions regarding how the two concepts are defined and measured impact the ways in which we are governed, as citizens. By deciding the conditions that determine who is living in poverty and who is disabled, governments can adjust the amount of financial support provided to its citizens. Through changing these classifications, governments can define the scale of the problem and therefore justify any interventions (or lack of them). Consequently, governmental decisions can greatly influence the number of people that are living in poverty.

Because of the way poverty is structured, education can be a way out of poverty. However, this is denied to many disabled people who find themselves segregated from their non disabled peers, and unable to gain formal qualifications for the labor market. The segregated nature of education for disabled children allows poverty to reproduce itself across generations.

The over-representation of disabled people in impoverished populations is, thus, to a certain extent a political choice. In the UK, and we suspect more broadly, austerity has had a significant role in this over-representation, allowing for a reduction in living standards to be justified as fair and proportionate (Cooper and Whyte, 2017; Hamnett, 2014; Hoggett et al., 2013). Yet, as we have seen, the impact of austerity disproportionately fell on marginalized groups, particularly disabled people (Ryan, 2019). The impact of this cannot be understated. Testimonies from impoverished disabled people and their families demonstrate the dehumanizing impact of poverty, which is then reinforced by the government, particularly through the welfare state. This, we argue, is a form of necro-political governance whereby many disabled people are not permitted to join the living. Rather, disabled people in poverty become the "living dead", as their human rights are denied, and they are unable to live their lives in a fulfilling and meaningful way.

Acknowledgment

We would like to thank the Center for Disability Studies at the University of Leeds for their ongoing support, with special thanks to our respective supervisors: Dr. Alison Sheldon, Dr. Tom Campbell, Dr. Angharad Beckett, and Dr. Joanne Greenhalgh. Special thanks to Rebecca Porter's study skills tutor Taiyybah, for her continued support.

References

- Alkire, S., Foster, J., Seth, S., Santos, M.E., Roche, J.M., Ballon, P., 2015. Multidimensional Poverty Measurement and Analysis. Oxford Poverty & Human Development Initiative (OPHI), London.
- Barnes, C., 2012. Understanding the social model of disability: past, present and future. In: Watson, N., Roulstone, A., Thomas, C. (Eds.), *Routledge Handbook of Disability Studies*. Routledge, Abingdon, pp. 12–29.

- Benstead, S., 2019. *Second Class Citizens: The Treatment of Disabled People in Austerity Britain*. Centre for Welfare Reform, Sheffield.
- Bourquin, P., Cribb, J., Waters, T., Xiaowei, X., 2019. *Why Has in-Work Poverty Risen in Britain?* Institute for Fiscal Studies, London.
- Briant, E., Watson, N., Philo, G., 2013. Reporting disability in the age of austerity: the changing face of media representation of disability and disabled people in the United Kingdom and the creation of new "folk devils". *Disabil. Soc.* 28 (9), 874–889.
- Campbell, T., 2013. *Dyslexia: The Government of Reading*. Palgrave Macmillan, Basingstoke.
- Contact, 2014. *Counting the Costs (2014): Research Into the Finances of More Than 3500 Families With Disabled Children Across the UK* [Online]. Available from: https://contact.org.uk/media/805120/counting_the_costs_2014_uk_report.pdf. (Accessed 30 October 2020).
- Cooper, V., Whyte, D., 2017. *The Violence of Austerity*. Pluto Press, London.
- Davies, T., Isakjee, A., Dhesi, S., 2017. Violent inaction: the necropolitical experience of refugees in Europe. *Antipode* 49 (5), 1263–1284.
- DWP [Department for Work and Pensions] and DE [Department for Education], 2011. *A New Approach to Child Poverty: Tackling the Causes of Disadvantage and Transforming Families' Lives* [Online]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/177031/CM-8061.pdf. (Accessed 2 November 2020).
- DWP [Department for Work and Pensions], 2019. *Fraud and Error in the Benefit System: Financial Year 2018 to 2019 Estimates* [Online]. Available from: <https://www.gov.uk/government/statistics/fraud-and-error-in-the-benefit-system-financial-year-2018-to-2019-estimates>. (Accessed 14 May 2020).
- DWP [Department for Work and Pensions], 2020a. *Households Below Average Income: An Analysis of the UK Income Distribution: 1994/95–2018/19*. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/875261/households-below-average-income-1994-1995-2018-2019.pdf.
- DWP [Department for Work and Pensions], 2020b. *PIP Assessment Guide Part 2: The Assessment Criteria* [Online]. Available from: <https://www.gov.uk/government/publications/personal-independence-payment-assessment-guide-for-assessment-providers/PIP-assessment-guide-part-2-the-assessment-criteria>. (Accessed 1 November 2020).
- Dwyer, P., Jones, K., McNeill, J., Scullion, L., Stewart, A., 2016. *First Wave Findings: Disability and Conditionality* [Online]. Available from: <http://www.welfareconditionality.ac.uk/wp-content/uploads/2016/05/WelCond-findings-disability-May16.pdf>. (Accessed 7 December 2016).
- Francis-Devine, B., 2020. *Poverty in the UK: Statistics* [Online]. Available from: <https://researchbriefings.files.parliament.uk/documents/SN07096/SN07096.pdf>. (Accessed 1 November 2020).
- Garthwaite, K., 2011. "The language of shirkers and scroungers?" Talking about illness, disability and coalition welfare reform. *Disabil. Soc.* 26 (3), 369–372.
- Garthwaite, K., 2014. Fear of the brown envelope: exploring welfare reform with long-term sickness benefits recipients. *Soc. Pol. Adm.* 48 (7), 782–798.
- Goulden, C., 2018. *What is Destitution?* [Online]. Available from: <https://www.jrf.org.uk/blog/what-destitution>. (Accessed 1 November 2020).
- Goulden, C., 2020. *UK Poverty 2019/20* [Online]. Available from: <https://www.jrf.org.uk/report/uk-poverty-2019-20>. (Accessed 30 October 2020).
- Gov.uk, (no date). *Carer's Allowance*, [Online], [Accessed 1 November 2020], Available from: <https://www.gov.uk/carers-allowance>.
- Gray, P., 2017. *Personal Independence Payment (PIP) Assessment: Second Independent Review* [Online]. Available from: <https://www.gov.uk/government/publications/personal-independence-payment-PIP-assessment-second-independent-review>. (Accessed 16 October 2020).
- Grover, C., Piggott, L., 2005. Disabled people, the reserve army of labor and welfare reform. *Disabil. Soc.* 20 (7), 705–717.
- Hamnett, C., 2014. Shrinking the welfare state: the structure, geography and impact of British government benefit cuts. *Trans. Inst. Br. Geogr.* 39 (4), 490–503.
- Hoggett, P., Wilkinson, H., Beedell, P., 2013. Fairness and the politics of resentment. *J. Soc. Pol.* 42 (3), 567–585.
- Jamieson, M., 2020. *The Austerity Cure: The Impact of Benefit Sanctions on Mental Health*. Luna Press Publishing.
- JRF [Joseph Rowntree Foundation], 2016a. *Special Educational Needs and Their Links to Poverty* [Online]. Available from: <https://www.jrf.org.uk/report/special-educational-needs-and-their-links-poverty>. (Accessed 2 November 2020).
- JRF [Joseph Rowntree Foundation], 2016b. *UK Poverty: Causes, Costs and Solutions* [Online]. Available from: <https://www.jrf.org.uk/report/uk-poverty-causes-costs-and-solutions>. (Accessed 30 October 2020).
- Joseph Rowntree Foundation (JRF), 2020. *What is Poverty?* [Online]. Available from: <https://www.jrf.org.uk/our-work/what-is-poverty>. (Accessed 30 October 2020).
- Kennedy, S., 2018. *Research Briefing: Changes to the Personal Independence Payment Eligibility Criteria* [Online]. Available from: <https://commonslibrary.parliament.uk/research-briefings/cbp-7911/>. (Accessed 30 October 2020).
- Lawson, A., Beckett, A.E., 2020. The social and human rights models of disability: towards a complementarity thesis. *Int. J. Hum. Right.* <https://doi.org/10.1080/13642987.2020.1783533>.
- Mau, S., 2019. *The Metric Society: On the Quantification of the Social*. Polity, Cambridge, UK.
- Mayblin, L., Wake, M., Kazemi, M., 2019. Necropolitics and the slow violence of the everyday: asylum seeker welfare in the postcolonial present. *Sociology* 54 (1), 107–123.
- Mbembe, A., 2003. *Necropolitics*. Public Culture 15, 11–40.
- Mbembe, A., 2019. *Necropolitics*. Duke University Press, Durham, US.
- McEnhill, L., Byrne, V., 2014. "Beat the cheat": portrayals of disability benefit claimants in print media. *J. Poverty Soc. Justice* 22 (2), 99–110.
- Morris, Z.A., Zaidi, A., 2020. Estimating the extra costs of disability in European countries: implications for poverty measurement and disability-related decompensation. *J. Eur. Soc. Pol.* 30 (3), 339–354.
- O'Connell, R., Owen, C., Padley, M., Simon, A., Brannen, J., 2019. Which types of family are at risk of food poverty in the UK? A relative deprivation approach. *Soc. Pol. Soc.* 18 (1), 1–18.
- Oliver, M., 1990. *The Politics of Disablement*. MacMillan, Basingstoke.
- Porter, R.L., 2018. *Personal Independence Payment (PIP) as a Form of Necropolitics* [Online]. Available from: <https://disability-studies.leeds.ac.uk/pipnecropolitics/>. (Accessed 2 November 2020).
- Porter, R.L., 2019. *Personal Independence Payment and the Necropower of Austerity* [Online]. Available from: <http://human.group.shef.ac.uk/personal-independence-payment-and-austerity/>. (Accessed 30 October 2020).
- Reckwitz, A., 2020. *Society of Singularities*. Polity Press, Cambridge.
- Ryan, F., 2019. *Crippled: Austerity and the Demonisation of Disabled People*. Verso, London.
- Scope, 2019. *Disability Price Tag Report* [Online]. Available from: <https://www.scope.org.uk/campaigns/extra-costs/disability-price-tag/>. (Accessed 30 October 2020).
- Shakespeare, T., 2012. Disability in developing countries. In: Watson, N., Roulstone, A., Thomas, C. (Eds.), *Routledge Handbook of Disability Studies*. Routledge, Abingdon, Oxfordshire, pp. 271–284.
- Singal, N., 2011. Disability, poverty and education: implications for policies and practices. *Int. J. Incl. Educ.* 15 (10), 1047–1052.
- Slee, R., 2014. Discourses of inclusion and exclusion: drawing wider margins. *Power Educ.* 6 (1), 7–17.
- Stewart, K., Roberts, N., 2019. Child poverty measurement in the UK: assessing support for the downgrading of income-based poverty measures. *Soc. Indic. Res.* 142 (2), 523–542.
- Stone, D., 1984. *The Disabled State*. Temple University Press, Philadelphia.
- Swain, J., French, S., 2014. International perspectives on disability. In: Swain, J., et al. (Eds.), *Disabling Barriers—Enabling Environments*, third ed. Sage, London, pp. 45–52.
- Thomas, C., 1999. *Female Forms: Experiencing and Understanding Disability*. Open University Press, Buckingham.
- Trussell Trust, 2020. *What We Do* [Online]. Available from: <https://www.trusselltrust.org/what-we-do/>. (Accessed 20 April 2020).
- UN [United Nations], 2018. *Disability and Development Report* [Online]. Available from: <https://www.un.org/development/desa/disabilities/wp-content/uploads/sites/15/2019/07/disability-report-chapter2.pdf>. (Accessed 15 April 2020).

- UN., 2016. Report of the Inquiry Concerning the United Kingdom of Great Britain and Northern Ireland Carried Out by the Committee Under Article 6 of the Optional Protocol to the Convention [Online]. Available from. http://www.ohchr.org/_layouts/15/WopiFrame.aspx?sourcedoc=%2FDocuments%2FHRBodies%2FCRPD%2FCRPD%2EC%2E15%2ER%2E2%2ERev%2E1-ENG%2Edoc&action=view. (Accessed 1 November 2020).
- UNICEF, 2018. Report Card 14: Building the Future: Children and the Sustainable Development Goals in Rich Countries [Online]. Available from. <https://www.unicef-irc.org/publications/890-building-the-future-children-and-the-sustainable-development-goals-in-rich-countries.html>. (Accessed 1 November 2020).
- UPIAS [Union of the Physically Impaired Against Segregation], 1976. Fundamental Principles of Disability. Union of the Physically Impaired Against Segregation, London.
- UNCPRD [United Nations Convention on the Rights of Disabled People], 2007, p. 24. January 2007. Resolution: A/RES/61/106.
- World Bank, 2020. Disability Inclusion [Online]. Available from. <https://www.worldbank.org/en/topic/disability>. (Accessed 1 November 2020).
- WOW [War On Welfare] Campaign, 2018. My Latest PIP Assessment is the Closest to Breaking Me [Online]. Available from. <https://wowvoices.uk/my-latest-pip-assessment-is-the-closest-to-breaking-me/>. (Accessed 16 October 2020).
- WOW [War On Welfare] Campaign, 2019. In December of 2017, I Was Called in for a Further Work Capability Assessment... [Online]. Available from. <https://wowvoices.uk/5896-2/>. (Accessed 16 October 2020).

Gender, poverty and disability—responding to a triple exclusion

Lucy Lake^a, Katie J. Smith^a, and Seamus Hegarty^b, ^a CAMFED International, Cambridge, United Kingdom; and ^b University of Warwick, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	470
Girls' education in sub-Saharan Africa	471
The CAMFED intervention	472
Effectiveness of the intervention	474
Evaluation approach	474
Retention outcomes	475
Learning outcomes	475
Cost-effectiveness of the intervention	477
A productive investment? The wider outcomes	477
Conclusion	478
References	478

Introduction

It is at once extraordinary and unsurprising that gender is a major determinant of access to education. Put simply, girls are less likely than boys to go to school and, when they do, they tend to leave earlier and with lower educational attainments. This failure to meet girls' right to education has been documented copiously, at national and global levels. Despite welcome progress over the past twenty years, gender inequality persists in many countries, especially in Africa and Western Asia, both in access to schooling and in post-school opportunities. The 2019 *Global Education Monitoring* (GEM) Gender Report (UNESCO, 2019) summarizes a mass of data, and the figures for gender parity in education make for stark reading (p. 8):

- 1/3 of countries have not achieved gender parity in primary schooling.
- 1/2 of countries have not achieved gender parity in lower secondary schooling.
- 3/4 of countries have not achieved gender parity in upper secondary schooling.

A subsequent report found that for twenty countries, mostly in sub-Saharan Africa, fewer than 1% of poor rural women completed upper secondary education (UNESCO, 2020, p. 68).

The link between poverty and low educational achievement is all too familiar. The GEM 2020 Report (UNESCO, 2020) gives an up-to-date picture, documenting the extent to which poverty is associated with low school attendance and completion, and reduced learning opportunities. The impact of poverty begins even before schooling starts: in low- and middle-income countries, large numbers of children from the poorest households are malnourished, "severely compromising their opportunity to benefit from education" (p. 6). An estimated 17% of the school-age cohort—258 million children and young people—are not in school. These school absentees are disproportionately from the poorest households. As documented in the GEM 2020 Report, disparities by family wealth in attendance and completion rates are large and enduring: the poorest children are least likely to attend school in the first place, most likely to repeat a year or leave school early, and have lowest levels of proficiency in literacy and numeracy.

Disability is a third strand in exclusion from education. Again, this is a long running story and the picture is, if anything, even more challenging. There are numerous data at national and global levels detailing the exclusion from schooling of children and young people with disabilities. Two sources can be cited briefly to indicate the scale of the challenge. The *World Disability Report* (WHO/World Bank, 2011) documents, for many countries and globally, the extent of the educational exclusion of children with disabilities, summarizing that they "... are less likely to start school and have lower rates of staying and being promoted in school" (p. 206). Indeed, the gap in school attendance between disabled and non-disabled children was close to 60% in some countries.

The *World Disability Report* (WHO/World Bank, 2011) documents the strength of the association between disability and school exclusion and the interaction between disability and other factors associated with exclusion. The correlations between low educational outcomes and disability are generally higher than the equivalent correlations for gender, economic status or rural residence. In other words, children and young people with disabilities are at even greater risk of school exclusion than girls or those from poor backgrounds. On top of that, there is an interaction between these factors: girls with a disability are more likely than boys with a disability to miss out on schooling or have limited educational outcomes; and children with disabilities from poor backgrounds are likewise at even greater risk of educational disadvantage.

A more up-to-date picture is provided by the GEM 2020 Report (UNESCO, 2020). The Report collates a large amount of data that displays the continuing association between exclusion from education and poverty, gender and disability. It affirms that "Children with disabilities are particularly at risk of exclusion from education" (p. 10). Despite the challenges in collecting data on

disability, it presents compelling data on the educational disadvantage experienced by children with disabilities. They are more likely than peers to be excluded from school, to attend segregated provision when they do go to school, to have fewer years of schooling, and to finish school with lower attainments. In one study (GEM Report team analysis of MICS and ASER Pakistan data), for instance, children with disabilities were 19% less likely to achieve minimum proficiency in reading than those without disabilities (p. 10).

This triple exclusion from education—based on gender, poverty and disability—is indefensible. Not only are the resources needed to educate all children and young people available, but the failure to deploy them results in enormous waste of human potential. Those who question the need for action, but question its affordability need only look at expenditure in other areas. To take just a few examples—global military expenditure, governments’ acquiescence in revenue foregone through tax avoidance and manipulation, the massive resources produced at short notice to deal with the impact of Covid-19—it is evident that resources can readily be found to meet certain priorities. (There is copious literature on each of these areas, not least in respect of income support schemes and other economic subventions during Covid-19. The Stockholm International Peace Research Institute is a reliable source on military expenditure around the world, documenting a total spend of more than US\$1.9 trillion in 2019 alone (<https://www.sipri.org/>). It is difficult, if not impossible, to calculate the total loss to national exchequers arising from tax avoidance and manipulation. Some estimates are available, however. International Monetary Fund (IMF) researchers calculated an annual loss to Governments of \$600 billion from selective tax avoidance techniques in the international corporate sector (Crivelli et al., 2016). The Tax Justice Network has estimated a loss of \$427 billion a year to tax abuse and evasion (<https://taxjustice.net/reports/the-state-of-tax-justice-2020/>). These figures are minimal estimates and do not take account of tax avoidance by individuals, inefficiencies in tax collection or the economic distortions caused by the tax avoidance industry, and it is likely that the total loss to Governments—with consequent impact on social spending budgets—is very much larger than the IMF or Tax Justice Network estimates).

So, what can be done to ensure a higher priority for the education of girls, particularly girls from poor backgrounds and/or with disabilities? If we can demonstrate that it is possible to educate these girls, at relatively little cost, maybe mindsets will shift a little. If, moreover, it transpires that such education is a productive investment resulting not only in greater human fulfilment, but also in net economic benefit, the case for such investment cannot readily be gainsaid.

The purpose of this chapter is to explore this question and propose solutions by reference to a substantial intervention that has been running for nearly 30 years in five sub-Saharan African countries.

Girls’ education in sub-Saharan Africa

Girls’ education is widely recognized for its fundamental and wide-ranging benefits to humanity across a range of indicators, including economic growth, improved health outcomes, and tackling climate change (Sperling et al., 2016). Educated women earn more, have better jobs, have smaller, healthier and better educated families, and are more likely to be decision-makers in their homes and communities. The economic benefit of ensuring universal secondary education for girls has been valued at between \$15 and \$30 trillion (Wodon et al., 2018), and the importance of education in supporting women’s leadership, particularly in low- and middle-income countries, has been identified as key to climate change adaptation and mitigation strategies (Brookings, 2021).

This makes a pressing case for tackling the issue of girls’ exclusion from education, which means addressing the urgent and inter-linked challenges related to gender, poverty and disability. Our focus here is on the marginalization experienced by girls from poor households in rural areas of sub-Saharan Africa where poverty is the greatest barrier to education (Morris et al., 2017). With many households surviving on less than \$1.25 per day, a lack of funds for both direct and indirect school-going costs creates the constant threat of exclusion from school among marginalized girls, or prevents their access entirely. Other poverty-related issues such as hunger can make it difficult for them to attend school regularly, learn while they are there, and study effectively at home. Poverty is also a key driver of practices like early marriage—both a cause and effect of girls’ exclusion from school—with girls facing pressure to get married to bring in more resources to their families, either through the “bride price” or through simply reducing the number of mouths to feed. In Tanzania, for example, 31% of young women aged 20–24 marry before 18 years and this rises to 49% among the poorest quintile, and 56.7% of young women have begun childbearing by age 19 (Tanzania DHS, 2016). Barriers are exacerbated for girls living with disabilities: the likelihood that children aged 12 in sub-Saharan Africa have ever enrolled in school is 10% lower for children with disabilities than those without, and the gap is bigger for girls (Wodon et al., 2018).

These issues are further compounded by the lack of education provision in remote, rural areas. Secondary schools are widely dispersed which has a significant impact on the retention of marginalized girls, who may be faced with long, tiring walks to and from school, or are forced to seek accommodation away from home with relatives or villagers close to the school; these factors, along with inadequate supportive infrastructure, are even more significant barriers for girls with disabilities. Both scenarios leave girls vulnerable to abuse, including sexual abuse (Stein, 2014). The under-resourcing of government schools in rural areas severely compromises the quality of education they can provide, and an academic curriculum that lacks relevance to rural young people’s reality and perceptions of their future, coupled with outdated pedagogy, also serves to undermine girls’ participation and self-esteem. The learning crises created by these circumstances contribute to pressure on girls to drop out of secondary school. This is compounded by the fact that when poor parents perceive education to be of low quality, they are less willing to sacrifice their limited resources to keep their children in school (World Bank, 2018)—a rational response which nevertheless perpetuates a cycle of exclusion.

The CAMFED intervention

One organization that has taken on this challenge is CAMFED (Campaign for Female Education, www.camfed.org), an international NGO serving girls and women in rural sub-Saharan Africa. Founded in 1993, it started out by providing bursaries to an initial cohort of 32 girls in Zimbabwe to attend secondary school; by early 2021, it had supported the education of more than 4.8 million marginalized children across five countries: Ghana, Malawi, Tanzania, Zambia, and Zimbabwe. CAMFED's work focuses on girls and young women because they are the most marginalized group in the communities where they live, but also provides some support to boys where they face similar barriers, and both girls and boys benefit from interventions to improve the learning environment and boost learning outcomes. In 2020, CAMFED was recognized with the Yidan Prize for Education Development for its work on girls' education.

CAMFED's intervention has evolved from one focused primarily on bursary support to girls, to a highly responsive model that addresses the multi-faceted barriers to school inclusion and learning among the most marginalized girls. It incorporates targeted financial and material support to cover direct and indirect school-going costs, and specialist support for children with disabilities. Young women school graduates operate as peer mentors or "Learner Guides" to deliver a life skills and wellbeing curriculum in schools. They are trained to spot indicators of a student's increased vulnerability to drop-out, for example, through changes in her attendance or engagement at school, and to alert the appropriate authorities. Learner Guides in turn benefit from the opportunity to acquire an accredited vocational qualification that provides a stepping stone to their further education and employment.

A core principle of CAMFED's intervention is that it operates in government school systems in full partnership with Ministries of Education, with the aim of ensuring the school system is able to recognize and better serve the often invisible needs of the most marginalized children, with a focus on girls. Its infrastructure for delivery is decentralized to community level, to support a locally responsive approach to the particular challenges young people face, recognizing that community-driven programs represent the most effective solution for reaching marginalized children in a sustained way, as the barriers to education for this group start well before the school gate. Importantly, CAMFED has set out to demonstrate that it is possible to scale up a community-driven approach to improve the education outcomes of marginalized children, countering the common perception that programs that go to scale often preclude the most marginalized because of the potentially high transaction and discretionary costs associated with this outreach.

Central to CAMFED's approach is the community-led process through which marginalized girls are identified to receive targeted support to meet their school-going needs. This process draws on a set of criteria that are grounded in government definitions of "vulnerability" as laid out by the relevant department of social welfare in each country. These criteria have been drawn up with and by communities, and cover the social background of girls (whether she is from a child-headed household, is affected by disability and so on) as well as her educational status (whether a girl has dropped out or is at risk of dropping out of school). This has led to a "marginality tool," operationalized as an "index" of 20 locally defined dimensions of marginalization, shown in the table below.

Marginality indicators identified by communities

1. A child whose parents/guardians cannot pay the school costs and so are often sent home or drop out of school.
2. A child living in a family that gets only one meal per day, or sometimes goes to bed hungry.
3. A child living in a household with very low income so that they cannot afford even the basic needs.
4. A child living with old relatives with no or little income, so the child has to earn income for the family.
5. An orphaned child living with guardians who is being neglected and not having all needs provided, including school costs.
6. A child taking care of sick or disabled parents, siblings or other relatives (which stops them going to school).
7. A child who lives in the street.
8. A child who lives in a household headed by a child.
9. A child who is the head of the household.
10. A child who is given a lot of work so that they don't have time to do their homework or they miss school.
11. A child whose guardian treats them unfairly compared to other children in the household in terms of work or provisions.
12. A child who spends a lot of time in church activities to the extent that she/he misses school.*
13. A child whose parents/guardians do not value education and so do not pay school fees and other school costs.
14. A child whose parents/guardians are sick or disabled so that they have very low or no income.
15. A child with a chronic illness or disability whose parents/guardians cannot afford the treatment and school-going costs.
16. A child with chronic illness/disability whose parents don't encourage them to go to school and so don't pay school-going costs.
17. A child living in a household with many children so that the parents/guardians cannot pay the school going costs.
18. A child who spends most or all of their leisure time working to make some money.
19. A child who does not have a permanent home and therefore often misses school.
20. A child whose parents/guardians are pressuring them to marry or drop out of school to get a job or work on the farm.

*This criterion was identified by communities concerned that some children were spending whole days and nights in church, which was leading them to miss large amounts of time in school.

In addition to its use in identifying girls likely to benefit from support, the index helps in outlining the distribution of needs across the target population. It has been integrated with other tools, including questions recommended by the Washington Group on Disability Statistics (http://www.cdc.gov/nchs/washington_group.htm), to achieve a nuanced understanding of the nature and degree of marginalization. A study conducted with a representative cohort in Tanzania and Zimbabwe in 2013 (CAMFED, 2013) found that while all students in the school communities experienced some degree of marginalization by virtue of their location in remote and under-served areas and the lack of, or low quality, school infrastructure, 62% and 68% of girls in the cohorts in Tanzania and Zimbabwe respectively were especially marginalized against this index (For comparison, in an analysis based on Grameen PPI, these girls achieved a mean score of 33.1).

The most common scenario in both Tanzania and Zimbabwe was “a child whose parents/guardians cannot pay the school costs and so is often sent home or drops out of school.” This scenario alone accounted for 35% of girls in the cohort in Tanzania at the baseline and 34% in Zimbabwe. According to the Grameen Progress Out of Poverty Index (PPI) (<https://www.povertyindex.org/>), the mean cohort member in Tanzania had a 78% likelihood of living in a household with an income of less than \$1.25 per day. Living in a household where the adults had a regular income was rare among marginalized girls, often resulting in difficulties paying for school-going costs. In both Tanzania and Zimbabwe, around half of marginalized girls were engaged in income-earning activity to contribute money to the household. This highlights the short-term opportunity cost for the households of children attending school and explains one important cause of drop-out from school.

Hunger was a common experience among the cohort students, a product of the prevailing levels of poverty. 60% of marginalized girls in Zimbabwe and 47% of those in Tanzania reported experiencing hunger on some or most days. Hunger can have many physical and psychological effects on young people that make learning difficult (<http://borgenproject.org/effects-of-hunger-on-education/>). Children experiencing hunger are more likely to have problems with memory and concentration. Hunger has been observed to cause depression, anxiety and withdrawal, all of which are obstructions to a child trying to focus on education. In addition, when hunger leads to malnutrition, children become more susceptible to illnesses, the development of the brain can be impaired and it can disrupt sleeping patterns, making a child too tired for a full day of school.

The loss of one or both parents was also associated with marginality, and was common among the cohort in both countries, especially in Zimbabwe. The death of her father and the associated economic fall-out was a significant factor leading to girls dropping out of school. Marginality was also associated with living in a family headed by a grandparent (or another elderly person) and with living in a female-headed household—both of which may be the consequence of the death of one or both parents.

Living with family members facing long-term illness or disability was something experienced by around half of marginalized girls in both countries. Long-term illness or disability will have immediate consequences for the economic welfare of most households, creating pressures to engage children of school-going age in income-earning activity. Another consequence for the children in these households is the need to provide the caring role for family members. The caring role is also a factor in larger households with many children or elderly relatives, a role often assigned to older siblings—particularly girls. For many marginalized girls, caring for family members was a regular, even daily, responsibility, and a common reason for missing school.

Many marginalized girls themselves had a disability. Around half of cohort students in Zimbabwe and a fifth in Tanzania reported having some kind of difficulty indicative of a disability. In many of the categories of disability (or difficulty), marginalized girls were more likely than their less marginalized peers to experience some level of difficulty, and the severity of difficulty tended also to be greater among them.

A subsequent study in 2018 (Surridge et al., 2018) of students at secondary school in Tanzania and Zimbabwe, and upper primary school in Zambia, confirmed the intersection of dimensions of exclusion and marginality with barriers to education. Girls identified as marginalized were more likely to be over-age for their school grade. They were also more likely to suffer from a high burden of chores (in Tanzania, 51% of marginalized girls compared with 8% of other girls). Marginalized girls felt less safe at school or traveling to and from school (in Zambia, 23% of marginalized girls compared with 14% of other girls), and had no encouragement from parents or caregivers to stay in school (in Tanzania, 19% of marginalized girls compared with 10% of other girls). Marginalized girls were also less likely to attend school regularly (defined as attending less than 85% of the time); in Zimbabwe, for example, this accounted for 81% of marginalized girls. Analysis of the intersection of dimensions of marginality found that students with a disability were likely to experience multiple barriers to education, though this multiple exclusion was not the exclusive experience of disabled children but the norm for the majority of girls—girls affected by the loss of one or both parents, and those living in a household without a regular income, for example, faced similar intersecting barriers.

In the 2018 study, 17% of girls in Tanzania and 15% of girls in Zimbabwe were identified as having some type of disability, with visual impairment the most commonly reported disability in both. These disability levels contrasted with official data: EMIS (Education Management Information System - school level data collected annually by governments) data in Zimbabwe gave a national rate of 1.08% for disability prevalence in secondary schools (both genders); in Tanzania the latest Basic Education Statistics Tanzania (BEST) survey at the time (2017) found the most prevalent disability was impaired vision, putting the prevalence at 0.2%. It is probable that there are several reasons for the discrepancy between the studies and government data, including differences in data-gathering tools, inadequacy of official data collection, and higher levels of marginalization among the target group. On a positive note, following this study the use of Washington Group questions has now been introduced at school level in Zimbabwe by ZIMSTAT (Zimbabwe National Statistics Agency).

Washington Group Questions

The Washington Group (WG) Questions are two sets of question (one short, one extended) developed to provide cross-nationally comparable, population-based measures of disability. Developed by the Washington Group on Disability Statistics, established under the United Nations Statistical Commission, they have been internationally validated and tested across low-, middle- and high-income countries. The WG questions are designed to be incorporated easily into surveys and censuses and provide a low-cost, easy-to-administer and analyze, method of gathering accurate data on disability prevalence to inform policy and action.

Effectiveness of the intervention

CAMFED's intervention has grown from small beginnings to become a substantial operation with many components. Size and complexity are no guarantees of excellence, however, and it is reasonable to ask how effective the intervention is and whether it delivers value for money. The intervention has in fact been subjected to a good deal of scrutiny, particularly as it has scaled up, partly as a result of donor requirements and partly from its internal commitment to good governance. Evaluations have focused on how successful the intervention has been in securing girls' continued attendance in school and becoming effective learners.

This section sets out the evaluation approach used to measure the effectiveness of the intervention. It presents evidence from a series of large-scale independent evaluations carried out over an eight-year period, 2013–2021, across three countries, with particular reference to school retention and learning outcomes. It also presents data on cost-effectiveness and draws attention to broader evidence suggesting that its expenditure on girls' education is a highly productive investment.

Evaluation approach

The evaluations took a quasi-experimental approach, tracking successive cohorts of students across randomly selected intervention and comparison schools. The underlying assumption was that a holistic intervention approach, incorporating targeted financial and material support alongside the "social" support delivered by Learner Guides, would lead to better retention and, ultimately, improved learning outcomes for marginalized girls in particular.

The evaluations were conducted in two phases: the first phase, from 2013–17, followed a cohort of 11,075 students at secondary school in Tanzania and Zimbabwe; the second phase, from 2018 to 2021, followed a cohort of 23,107 students at secondary school in Tanzania and Zimbabwe, and upper primary school in Zambia. At baseline each cohort was disaggregated according to the degree of marginality students faced, using CAMFED's marginality index. This allowed the evaluation to examine whether and to what extent the intervention reached the most marginalized girls, and to assess whether the impact varied according to a student's marginality status.

The tools to measure learning were developed and delivered in partnership with Ministries of Education and exam councils in all three countries, and marked by the relevant national exam board: this helped ensure buy-in, relevance and comparability. It also set the stage for the subsequent integration and adoption of components of the intervention into school systems nationally. In the first phase of evaluations (2013–2017), learning tests were developed using a national assessment tool as the base. In the second phase (2018–2021), the EGRA (Early Grade Reading Assessment) (https://earlygradereadingbarometer.org/pages/about_egra) and EGMA (Early Grade Mathematics Assessment) (https://pdf.usaid.gov/pdf_docs/Pnads439.pdf), open-source tools developed by RIT International for USAID, were used as a basis for the development of tests known as SeGRA/SeGMA. These used a framework of subtasks designed to increase in difficulty and test higher orders of learning among secondary school students, and unlike the (foundational level) EGRA/EGMA tests, were delivered solely in a written format. In Zambia, where the intervention supported upper primary school students to progress to and through secondary school, the SeGRA/SeGMA tests were used in a continuum with EGRA/EGMA style tests used at the lower grades.

The evaluation findings are detailed below. Significant improvements were identified in retention and learning across all groups experiencing different degrees of marginalization, who were reached by CAMFED's intervention. Importantly, these improvements were repeated in each successive tracked cohort.¹

¹Tracked cohorts:

2013–2015: 11,075 students in Tanzania and Zimbabwe followed for 18 months across two academic years, from Form 2 to Form 4
 - 2015–2017: 11,255 students in Tanzania and Zimbabwe followed for 12 months for a single year, as they transitioned from F3 to F4
 - 2017–2019 (A): 23,107 students in Tanzania, Zambia and Zimbabwe followed for 18 months. In Tanzania and Zimbabwe this tracked the transition from F2 to F4. The Zambia sample included two sets of students, progressing from Grades 5 to 7 and 7 to 9 respectively.
 - 2017–2019 (B): 7985 students in peri-urban districts of Tanzania followed for 18 months between forms 1 and 3.
 - 2019–2021: The Zambian students in cohort 2017–19(A) will be followed through to grades 9 and 11 respectively, as will the younger part of the 2017–19(B) cohort. The remainder of the 2017–2019 cohorts in Tanzania and Zimbabwe will be followed to assess post school outcomes and pathways but will not sit learning assessments.

Retention outcomes

Evaluation of the first phase (2013–2017) of the intervention in Tanzania and Zimbabwe (Morris et al., 2017) found:

- Marginalized girls supported by CAMFED in intervention schools were significantly more likely than their counterparts in comparison schools to have stayed in school to the end of lower secondary education. This was true for successive cohorts.
- The probability of a marginalized girl in a comparison school in Zimbabwe in Cohort 1, for example, dropping out of school was nearly three times as high (275%) as for a marginalized girl in an intervention school ($p < 0.001$).

These results were repeated in subsequent cohorts and in new geographies: evaluation of the 2017–19 period (Roland et al., 2020) found similar positive associations between the intervention and girls' retention and attendance in school:

- In Tanzania, retention of marginalized students was greater by 14% in intervention compared to comparison schools, a statistically significant value at the $p < 0.001$ level.
- Marginalized girls at supported schools were more likely to successfully transition from Form 1 to Form 2 than those at comparison schools.
- In Zambia, the proportion of marginalized girls attending school regularly (defined as in school $>85\%$ of the time) rose by 22.5% over an 18-month period.
- Children with disabilities were more likely to successfully transition to the next grade of education in intervention schools than children with disabilities in comparison schools.

Ropafadzo is 15 and lives in rural Zimbabwe. As a baby she suffered severe burns in an accident at home, resulting in a traumatic amputation of her right hand, which became worse through lack of medical attention. In addition to the physical challenges this caused, it severely affected her self-esteem, creating an additional barrier to inclusion in education. Ropafadzo was identified for support to attend secondary school, including access to specialist medical assessment. This resulted in Ropafadzo undergoing corrective surgery and prosthetic support. Importantly, Ropafadzo was connected to young women with similar experiences of disability: they included Primrose, who had suffered burn injuries as a child and was herself supported to complete her education by CAMFED. Ropafadzo called this connection "my first turning point".

Analysis by the REAL Center (Research for Equitable Access & Learning) at the University of Cambridge, based on data from the first (2013–2017) evaluation phase in Tanzania (Alcott et al., 2020), found:

- Accounting for household wealth, marginalized girls in supported schools were 18% less likely to drop out compared to similarly disadvantaged girls in comparison schools; those receiving financial and material support were 31% less likely to drop out, an improvement that was similar across different degrees of disadvantage and levels of disability.
- Learning levels and learning support matter: drop-out was reduced for everyone, but less significantly for weaker learners, who had higher initial levels of dropout (improving from 20% to 16% vs. 8%–3% for stronger learners).

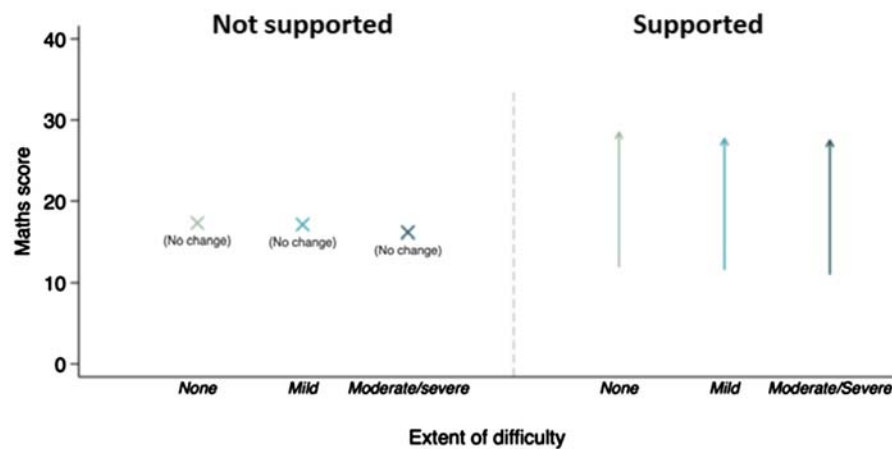
Learning outcomes

In the first evaluation phase (Morris et al., 2017), marginalized girls in intervention schools demonstrated significantly larger learning gains than their counterparts in the comparison schools in both Math and English, repeated in successive cohorts over the intervention period. Specifically:

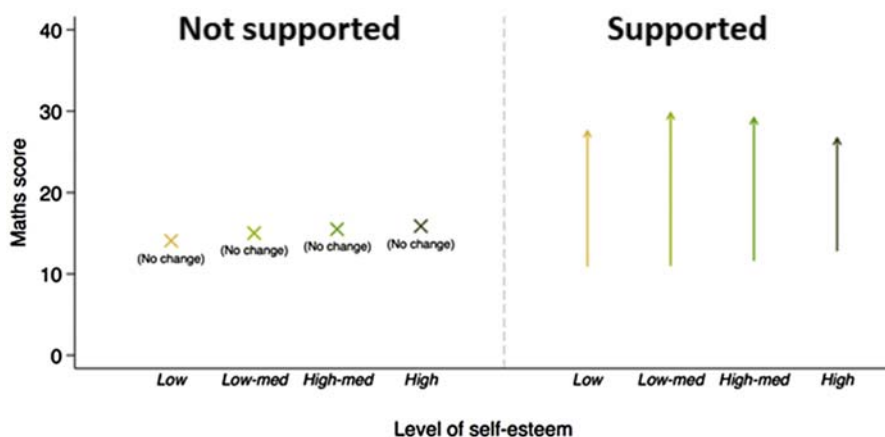
- In Tanzania, learning gains by marginalized girls at intervention schools in both Math and English were statistically significantly greater than those among marginalized girls at comparison schools. For Math, these equated to a Difference in Difference (DiD) of +16.62 marks (an effect size of 1.06, measured in standard deviations) for Cohort 1 and a DiD of +13.2 marks (an effect size of 0.65) for Cohort 2 (A difference-in-difference approach compares the average change over time in the outcomes of the intervention group to the average change in those of the comparison group, in order to calculate the effect of the intervention).
- In Zimbabwe, where partner communities were affected by severe droughts and flooding over the period, overall gains were lower than seen in Tanzania (Zimbabwe experienced a series of environmental shocks over the period, which badly affected target communities including through school closures) but reflect significant differences between gains achieved by marginalized girls in intervention and comparison schools. For Math, these equated to a DiD of +1.12 marks (an effect size of 0.14) for Cohort 1 and a DiD of +1.32 marks (an effect size of 0.34) for Cohort 2.

Additional analysis conducted by the REAL Center (Alcott et al., 2020) found that all girls, including those with disabilities (as above, 17% of the cohort at baseline), improved their learning against comparison groups: *"Once supported by the program, we find that marginalized girls significantly increased their English and Math test scores over the period of two years, regardless of physical or mental learning difficulties. For example, on average girls with moderate/severe difficulties improved their Math scores from 11 to 28 points out of 100. By comparison, equivalent marginalized girls not supported by the program did not improve. In addition, it is worth noting that marginalized girls with physical or mental difficulties supported by CAMFED improved their learning at the same rate as marginalized girls who did not report any difficulties."*

This finding is illustrated in the following figure taken from the REAL Center’s paper showing absolute gains in Math test scores for girls according to extent of disability.



There was also a correlation between supporting girls’ academic self-esteem and improvements in their learning, as shown in the figure below. Academic self-esteem was measured at baseline for each cohort using a bespoke “Attitudes to Learning” questionnaire tool, which explored students’ beliefs about their abilities and academic enjoyment. While at baseline girls’ learning differed according to their level of self-esteem (where girls with low self-esteem scored on average 11 points in Math, vs. 13 points scored by those with higher self-esteem), learning benefits occurred across the entire self-esteem scale.



Evaluation evidence, measured through the Attitudes to Learning tool, confirmed that the Learner Guides and life skills curriculum had a statistically significant ($p < 0.01$) impact on girls’ enjoyment of school and confidence in their academic progress in both countries (Morris, 2016). More than 95% of marginalized girls in both Tanzania and Zimbabwe reported they felt more positive about their future and more able to shape their life goals.

The independent second phase evaluation, tracking subsequent cohorts across Tanzania, Zambia, and Zimbabwe, found repeated improvements in girls’ learning (Roland et al., 2020):

- In Tanzania, there were statistically significant gains for marginalized girls at intervention schools with DiD increases over scores for marginalized girls at comparison schools of 6.7 pp ($p < 0.001$), reflecting an effect size of 0.71. English scores showed statistically significant DiD gains of 3.3 pp ($p < 0.01$) for marginalized girls at intervention schools, which represents a standardized effect size of 0.19.
- In Zimbabwe, Math gains for marginalized girls were significantly positive, with a DiD score of 7.0 pp ($p < 0.001$), representing an effect size of 0.35. The DiD increases for all students (girls and boys, marginalized and less marginalized) were significantly greater than for all students at comparison schools.
- In Zambia, Math gains for marginalized girls at intervention schools were significantly greater than for their peers at comparison schools. Those progressing from grade 5 to 7 achieved scores with a significant DiD of 5.8 pp ($p < 0.001$). Marginalized girls progressing from grades 7 to 9 at intervention schools achieved a positive DiD of 6.5 pp ($p < 0.001$).

In Tanzania, where the intervention had been replicated from scratch in new districts, independent evaluation after approximately 18 months of intervention found (Flett et al., 2020):

- Gains in English literacy for marginalized girls were larger in intervention schools versus comparison schools and showed a statistically significant ($p < 0.003$) DiD of 4.6 pp (effect size: 0.3) for girls moving from Form 1 to Form 2, and a significant DiD of 4.2pp ($p < 0.001$) (effect size: 0.26) for the combined cohorts of girls moving from Form 1 and 2 and from Form 2 to 3.
- Math skills of marginalized girls moving from Form 1 to Form 2 in intervention schools improved at a greater rate than those in comparison schools. Their DiD numeracy gains were statistically significant ($p < 0.008$) 3.0 pp (effect size: 0.27). Numeracy gains for the combined Form 1 and Form 2 cohorts were 2.8 pp, ($p < 0.002$), reflecting an effect size of 0.20.

These latter results have important implications for the replicability of the intervention; they demonstrate that it is possible to introduce an intervention of this nature to new areas and achieve meaningful improvement in the education outcomes of marginalized girls within a reasonable timeframe.

Cost-effectiveness of the intervention

It is one thing to evaluate the effectiveness of an intervention in terms of the impact on young people's education outcomes; it is quite another to demonstrate whether it is cost-effective in achieving those outcomes. This kind of information is vital in enabling policymakers to consider how resources can be best deployed to address the educational needs of all children.

Cost-effectiveness is an approach to calculating how much impact on education outcomes an intervention such as CAMFED's achieves relative to the inputs invested in it. Common approaches to cost-effectiveness analysis calculate the average cost associated with the average improvement in learning. However, this ignores the additional resources that are likely to be needed to support marginalized groups who face multiple forms of disadvantage.

The REAL Center conducted a detailed cost-effectiveness analysis (Sabates et al., 2020) drawing on evaluation data and financial information from CAMFED's intervention in Tanzania. Importantly, this cost-effectiveness analysis incorporated an equity lens in estimating the additional unit cost of reaching marginalized children relative to the increase in learning for these groups resulting from the intervention, such that disparities in learning outcomes are narrowed.

The findings of this cost-effectiveness analysis were that, for all children, the impact of the intervention was equivalent to 1.7 extra years of schooling per \$100 spent. Taking equity considerations into account, the value of the intervention was equivalent to 2 additional years. The overall conclusion was that interventions of this nature that target the most marginalized (in this case girls marginalized by gender, geography, poverty, and disability) address the needs of all: *"inclusive education systems will function for everyone if they function for the most marginalized"* (Sabates et al., 2018, p. 3). This conclusion has important implications for policymakers in the drive to achieve system-wide inclusive and equitable quality education. It helps to dispel the myth that it is prohibitively costly to address the educational needs of children who are most marginalized.

A productive investment? The wider outcomes

Arguably one of the most important outcomes of the intervention lies in the cohorts of girls who graduate from secondary school, in terms of what they go on to do and become. The benefits of girls' education are well documented in the ensuing improvements in health and welfare. What is perhaps less conspicuous in the evidence base to date is the impact of an empowering educational experience on girls' longer-term prospects beyond school.

In the case of CAMFED, there has been a proactive approach to ensure continued contact with marginalized girls who received support through their education, thus opening up the opportunity to track their life choices and outcomes. Data shows that these young women marry and have children later than their peers (Binkley, 2019): in Zambia, for example, a survey of CAMFED Association members found that only 18% of women had married, compared with a national average of 55% for the same age bracket (Good, 2019); in Tanzania CAMFED Association members start a family on average 2.5 years later than their peers (Binkley, 2019).

Many of these young women have gone on to join the CAMFED Association, the alumni network founded by the first young women to complete their secondary education with CAMFED's support, which by early 2021 numbered 178,000. In their transition from school, the majority have embarked on some form of entrepreneurship as the stepping stone to formal employment, self-employment or further education. An independent evaluation of CAMFED's impact on the lives of young women in the initial post-school phase, recorded "significant impact" in a number of dimensions: increased entrepreneurship, savings and household consumption; decision making power; access to further education and business growth; and, significantly, in their burgeoning philanthropy (Binkley, 2019).

Empowered by their own shared experience of overcoming poverty and inequality to get an education, these young women are united in their determination that no other girl should face the same hardships and are stepping up in remarkable ways to support the education of other marginalized children. Each member of the Association is assisting, on average, three other girls in her community to go to school; an extraordinary multiplier. Pulitzer prize-winning author Nick Kristof describes it as *"a perpetual motion machine... a virtuous cycle that grows over time"* (<https://www.nytimes.com/2020/11/21/opinion/sunday/charitable-holiday-gift-guide.html>).

Naomi is one such young woman in Zambia, who took up the role of Learner Guide and has recently been awarded a BTEC (Business and Technology Education Council) qualification:

I am passionate about supporting girls and young people to stay in school and reach their potential, so I wanted to improve my ability to do this as a Learner Guide. As a Guide I offer mentoring and psychosocial support to help learners improve and complete secondary school, which is crucial to their future plans. My facilitation skills have been enhanced so that I am able to deliver sessions more effectively and better support learners.

The BTEC helped me to develop my interest in becoming a teacher and I have learned how to work with learners in and out of school. It has sharpened my facilitation skills, helping me to deliver higher quality sessions. The BTEC has strengthened my personal and professional skills, like problem-solving, responsibility, teamwork, organisation, time management, self-discipline, commitment, and patience, just to mention a few.

I am going the extra mile because I look for more work and volunteer my strength whenever duty calls. I often support my fellow Guides through delivering extra sessions when I am available. Through this I have learnt a lot from other people's experiences and gained more exposure to teaching, which will help me in my career ambitions.

I am a role model in my community because I demonstrate confidence and leadership. I communicate and interact effectively with everyone, always showing respect and concern for others. I am knowledgeable, well-rounded, honest and have a positive attitude. In my community, I lead activities to promote girls' education and support my fellow women. Through my willingness to volunteer and give back to my community, I am a positive role model. I conduct community sessions with youths, following the My Better World curriculum; I sensitise them on the dangers of early marriage and early pregnancies, and teach them about sexual and reproductive health. I want to help learners to stay in school and complete their education.

Thanks to the skills and confidence I have developed through volunteering as a Learner Guide and undertaking the BTEC, I am now a public speaker, frequently invited to speak at different platforms on issues concerning the wellbeing of youths.

Naomi, like many other Learner Guides, is on track to pursue formal teacher training ultimately to take up a place in the teaching profession, toward redressing the gender imbalance in teachers posted to rural schools. From an intimate understanding of what it means to be marginalized, an understanding that is embedded in their own experience, these young women can surface the often-hidden needs of the most marginalized children, especially those affected by disability, and ensure their needs are sensitively and effectively addressed. They are now at the vanguard in communities in reaching out and assisting marginalized girls.

Importantly, CAMFED itself is now largely staffed and led by women who themselves received support from CAMFED to attend school and who are members of the CAMFED Association. This network of women leaders is at the forefront of designing and delivering programs, and at the table with policymakers to advocate for the system-wide adoption of strategies that are effective in improving the outcomes of marginalized young people.

Conclusion

Education is the unquestioned route to individual betterment and social improvement: access to high-quality, appropriate education is an essential prerequisite for an equitable and prosperous society. Regrettably, many countries lack such universal access and progress toward it is unacceptably slow, where it has not stalled entirely. As documented above, gross inequalities in access to education persist, particularly in respect of gender, poverty and disability.

This situation can, and should, be addressed. The arguments against doing so appear to revolve around the twin obstacles of affordability and feasibility: extending high-quality education to all children is prohibitively expensive and, in any case, we do not know how to do so effectively. The purpose of this chapter is to challenge these arguments and, by presenting evidence, refute them. Not only do the reasons for inaction not stand up to close scrutiny, but we can demonstrate, on the basis of large-scale, system-level evidence, that quality education is *already being provided* to girls marginalized by gender, poverty and disability. Moreover, the cost of these interventions is modest and, arguably, an investment that more than repays the initial outlay. We have shown that the initial cost yields substantial extra years of schooling. When to this are added the manifold other benefits that accrue—better employability, increase in entrepreneurial activity, reduction in early marriage and infant mortality—the educational expenditure outlined here is a highly productive investment.

The global community and many national governments have committed to the imperative of universal education. This means quality education for all children—regardless of gender, poverty or disability. If the commitments are genuine and we know what needs to be done and that it is affordable, the case for urgent, comprehensive action is unassailable.

References

- Alcott, B., Rose, P., Sabates, R., 2020. Targeted, multidimensional approaches to overcome inequalities in secondary education: case study of CAMFED in Tanzania. In: Background Paper for the Education Commission. <https://doi.org/10.17863/CAM.7685>.
- Binkley, A., 2019. An Evaluation of CAMFED's Impact in Supporting Young Women's Livelihoods in Four African Countries. AJB Advisors. <https://camfed.org/ajblivelihoods>.
- Brookings Institution, 2021. <https://www.brookings.edu/blog/education-plus-development/2021/02/10/why-is-girls-education-important-for-climate-action/>.
- CAMFED International, 2013. A New Equilibrium for Girls - CAMFED International 5101 Baseline Report. SQW, London. https://camfed.org/equilibrium_baseline.
- Crivelli, E., De Mooij, R., Keen, M., 2016. Base Erosion, Profit Shifting and Developing Countries. IMF working paper. International Monetary Fund, Washington, DC.
- Flett, P., Surridge, M., Baur, D., Kureya, T., Thirkell, A., 2020. Girls Learn, Succeed, and Lead - Midline Evaluation Report. Centre for International Development & Training, Wolverhampton, UK. <https://girlseducationchallenge.org/media/loaj23q/camfed-international-tanzania-midline.pdf>.
- Good, D., 2019. CAMA: A Powerful New Force for Development. CAMFED/Linklaters. https://camfed.org/wp-content/uploads/2021/07/CAMA-Governance-Report-Diana-Good-Linklaters_2019.pdf.
- Kristof, N., 2020. Choose a gift that changes lives. New York Times. <https://www.nytimes.com/2020/11/21/opinion/sunday/charitable-holiday-gift-guide.html>.
- Morris, M., Hardy, D., Zivetz, L., 2017. A New Equilibrium for Girls - CAMFED International 5101 Endline Report. SQW, London. https://camfed.org/equilibrium_endline.
- Morris, M., 2016. A New Equilibrium for Girls - CAMFED International 5101 Midline Report. SQW, London. https://camfed.org/equilibrium_midline.

- Roland, R., Baur, D., Flett, P., Nyirenda, R., Surridge, M., Van Graan, M., Kureya, T., Thirkell, A., 2020. The Ultimate Virtuous Cycle of Girls' Education – Midline Evaluation Report. Centre for International Development & Training, Wolverhampton, UK. <https://girlseducationchallenge.org/media/0s3hghyw/camfed-international-multicountry-midline.pdf>.
- Sabates, R., Rose, P., Delprato, M., Alcott, B., 2018. Cost-effectiveness with Equity: Raising Learning for Marginalised Girls through Camfed's Program in Tanzania. Policy Paper No. 18/2. REAL Centre, University of Cambridge. <https://doi.org/10.5281/zenodo.1247315>.
- Sabates, R., Rose, P., Alcott, B., Delprato, M., 2020. Assessing cost-effectiveness with equity of a program targeting marginalized girls in secondary schools in Tanzania. *J. Dev. Effect.* 13 (1), 28–46. <https://doi.org/10.1080/19439342.2020.1844782>.
- Sperling, G., Winthrop, R., Kwauk, C., Yousafzail, M., 2016. What Works in Girls' Education: Evidence for the World's Best Investment. Brookings Institution Press, Washington, DC.
- Stein, A., 2014. Tanzania's girls struggle for education. *World Pol. J.* <http://worldpolicy.org/2014/12/15/tanzanias-girls-struggle-for-education/>.
- Surridge, M., Roland, R., Begum, R., Kureya, T., Piringondo, A., Zinhumwe, G., 2018. The Ultimate Virtuous Cycle of Girls' Education – Baseline Evaluation Report. Centre for International Development and Training, Wolverhampton, UK. <https://girlseducationchallenge.org/media/bzwh5zks/camfed-international-multicountrybaseline.pdf>.
- Tanzania DHS, 2016. Tanzania Demographic and Health Survey and Malaria Indicator Survey 2015–16. Dar es Salaam, Tanzania, and Rockville, Maryland, USA.
- UNESCO, 2019. Global Education Monitoring Report – Gender Report: Building Bridges for Gender Equality. UNESCO, Paris.
- UNESCO, 2020. Global Education Monitoring Report 2020: Inclusion and Education: All Means All. UNESCO, Paris.
- WHO/World Bank (World Health Organization/World Bank), 2011. World Disability Report. WHO Press, Geneva.
- Wodon, Q., Montenegro, C., Nguyen, H., Onagoruwa, A., 2018. Missed Opportunities: The High Cost of Not Educating Girls. The Cost of Not Educating Girls Notes Series. World Bank, Washington, DC.
- World Bank, 2018. World Development Report 2018: Learning to Realize Education's Promise. World Bank, Washington, DC. <https://doi.org/10.1596/978-1-4648-1096-1>.

Inclusive education and deaf learners

Kristin Snoddon, School of Early Childhood Studies, Ryerson University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	480
Are deaf learners exceptional?	480
A segregated separate school?	481
Closing deaf schools	482
Where is a way forward?	483
References	483

Introduction

This chapter provides an overview of inclusive education for deaf learners with reference to how this concept has been differentially defined by some deaf advocacy organizations and inclusive and special education advocates, and operationalized by governments. Inclusive education is often viewed in terms of disabled children's educational placement in a regular school, instead of governments hosting separate systems for regular and special schools (Kayess and Green, 2017). In an inclusive classroom, all children receive instruction by primarily the same teacher (Florian and Spratt, 2013). In Canada and elsewhere in the world, the movement toward inclusive education has usually meant a focus on disabled children's integration, or proximity to nondisabled peers (Allan, 2008).

However, this focus on children's physical presence in the classroom frequently overlooks the issues of deaf children's participation and social and academic development. For most deaf learners, the main areas of need relate to language development and communication with peers and teachers. Deaf individuals are at high risk of language deprivation due to chronic lack of full access to a natural language in early childhood (Hall et al., 2017). Deaf advocacy organizations have long argued that these needs are best met in deaf schools where a natural sign language is the medium of instruction and a subject of study, and a written/spoken language is also learned. In fact, as past World Federation of the Deaf presidents Kauppinen and Jokinen (2014) have argued, bilingual education for deaf learners is not special education but part of an inclusive education system. However, this type of bilingual education is still very rare. Schools for deaf learners have also often entailed children's social and identity development through participation in a group of peers who share the same language and who benefit from access through visual and, in the case of deafblind children, tactile modalities. Consequently, the drafting and ongoing interpretation of international statements involving disability rights in education, such as the Salamanca Statement and Framework for Action on Special Needs Education (UNESCO, 1994) and the 2006 United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) (United Nations, 2006), have been sites of both triumph and struggle for deaf community advocates (Snoddon and Murray, 2019).

The following sections of this chapter discuss some policy issues and problems with practical implementation and ideologies of inclusive schooling for deaf learners. In many ways, these problems are rooted in attitudinal and systemic barriers to inclusion, including the predominance of a special education model of service delivery, that have been identified by other scholars (Slee and Allan, 2001). However, inclusive education advocates may also need to give greater consideration to the primacy of natural sign languages and a bilingual model of teaching and learning that affords learner congregation with other children and adult signers, including teachers.

Are deaf learners exceptional?

Among advocates for deaf bilingual education, there are perceived to be multiple lines of thought regarding the interpretation of inclusive education. In part, these differing perspectives may be rooted in deaf people's relationship with disability. Deaf people have long struggled against their ongoing classification and management according to a medical model of disability as defective individuals in need of rehabilitation. In response, many deaf community advocates have disputed the notion that deaf people are disabled in the sense of a medical model of disability; rather, as sign language users, deaf people are part of a vibrant community and a linguistic and cultural minority (e.g., Lane, 2005). This position is buttressed by scholarly consensus regarding the complex linguistic organization of sign languages and by research in deaf histories, social organizations, ontologies, and epistemologies (Kusters et al., 2017).

In some circles, there is now more recognition of a social model of disability, which states people with impairments are disabled by society (Barnes, 2019), and also more recognition of a human rights model of disability that underpins the UNCRPD (Degener, 2016). The human rights model is a broad, rights-based model of disability that values impairment as part of human diversity (Degener, 2016). Other scholars have employed a capability approach and social relational model of disability in analyzing deaf childhoods (Snoddon and Underwood, 2014). This sociological model attempts to account for the complexities and diversities

in the lives of deaf children and seeks to support deaf children's freedom to achieve their full potential as members of deaf and hearing communities. Scholars have also contrasted a cultural model of disability with the aims of deaf communities (Snoddon and Underwood, 2017). This ongoing theorizing and evolving understanding of disability recognizes that deaf people have multiple identities, including disabled identities. In addition, since deaf people are explicitly recognized as rights-bearers only in disability legislation, deaf advocates have taken a pragmatic approach in utilizing this legislation to lobby for sign language rights (De Meulder and Murray, 2017; Snoddon, 2009). Thus, deaf advocacy and academic circles have become more accepting of a dual-category status of deaf people as both disabled and members of linguistic minorities (De Meulder and Murray, 2017; Ladd, 2003).

During participation in the UNCRPD drafting process, the World Federation of the Deaf sought "the right to be different" as a fundamental principle of the Convention (Kauppinen and Jokinen, 2014). Deaf and deafblind people are mentioned in the UNCRPD more than any other disability group (Wilcox et al., 2012), notably in Article 24, which deals with education. Section 3 of Article 24 outlines some measures related to deaf learners which states parties must take, including:

- (b) Facilitating the learning of sign language and the promotion of the linguistic identity of the deaf community;
- (c) Ensuring that the education of persons, and in particular children, who are blind, deaf or deafblind, is delivered in the most appropriate languages and modes and means of communication for the individual, and in environments which maximize academic and social development.

The above reference in section 3(c) to "environments which maximize academic and social development," particularly for deaf, blind, or deafblind children, has been taken to mean that the UNCRPD sanctions bilingual deaf schools (Arnardóttir, 2011; Murray et al., 2018). However, as discussed further below, this notion continues to be challenged by some inclusive education advocates who view this type of setting as "segregated." Scholars have also asserted that there is a legislative history in the UNCRPD of a "sensory exception" to the principle of "full inclusion" of all children in a regular neighborhood school. The "sensory exception" recognizes the right to learning environments where deaf, deafblind, and blind children learn with their peers (Murray et al., 2018). This exception is visible in the provisions of Article 24.

However, stating that deaf learners are the exception to the rule of a model of inclusion that is focused mainly on placement and that inclusion can be measured in degrees of proximity to a nondisabled center is different from asserting that bilingual schools for deaf learners form part of an inclusive education system. The two propositions follow different definitions of inclusion, with the former oriented to a needs-based approach that gives more consideration to the education and social needs of individual children and the latter to a more conceptual, rights-based approach to inclusion (Hall, 2002; Snoddon and Murray, 2019). Additionally, the notion that deaf learners have a special status creates a tension vis-à-vis both the universalizing claims of some inclusive education advocates and the proponents of special education based in a deficiency model who have long held sway in deaf education. A focus on deaf learners' exceptional status may overlook commonalities between deaf learners and other learners that can inform an understanding of inclusion.

A segregated separate school?

In 2016, the CRPD Committee issued General Comment No. 4 on Article 24 of the UNCRPD. The General Comment is intended to explicate and provide practical orientation for the implementation of the rights outlined in Article 24 (Murray et al., 2020). However, several deaf scholars have disputed General Comment No. 4's interpretation of inclusive education and Article 24 (Murray et al., 2018, 2020). Section 11 of the General Comment states, "Segregation occurs when the education of students with disabilities is provided in separate environments designed or used to respond to a particular or various impairments, in isolation from students without disabilities" (emphasis in original). For many observers, this definition describes deaf schools and places them in direct opposition to inclusive education. For example, Kayess and Green (2017, p. 59) write that placement in a neighborhood school with nondisabled peers must be the "default" position in inclusive education. Deaf scholars have charged that this definition is not in keeping with the legislative history of Article 24 and that the views of deaf advocacy organizations were not sufficiently considered in General Comment No. 4 (Murray et al., 2020).

An unspoken consideration is whether the foremost construction of a school for deaf learners is based only or mainly on disability or impairment. This basis ignores the cultural and linguistic identity of deaf communities or at least makes this identity subservient to a disability inclusion agenda. Murray et al. (2020) note that General Comment No. 4 conceives of sign languages in terms of assistive technology and accommodations for communication impairments, and this conception militates against an understanding of deaf learners as bilinguals with bilingual education rights like members of any other linguistic minority. Moreover, a prohibition of disabled people's congregation and affiliation seemingly takes a stigmatizing view of disability instead of recognizing what is positive, powerful, and productive in disability and in other shared social identities and life experiences (Hall, 2002). In the face of systemic oppression, deaf communities' struggle to recruit and train more signing deaf teachers of deaf learners finds strength in the advocacy efforts that are nurtured through deaf people's congregation and identity development (Hall, 2002; Snoddon, 2021b).

As Hall (2002) notes, the concept of "full inclusion" is a reaction to the ongoing history of educational exclusion and special education, which continue to oppress disabled learners. Deaf education systems are largely special education systems run by non-deaf professionals with a vested interest in maintaining the hegemony of oralism (Branson and Miller, 1993). This is evident in the discrimination practiced by teacher-of-the-deaf education, which has long enacted barriers to enrollment and licensure for signing

deaf candidates (O'Brien, 2021; Snoddon, 2021b). Today, however, teacher-of-the deaf education and professional practice are often geared toward the rehabilitation and assimilation of deaf children in a neighborhood school rather than the maintenance of deaf schools that support deaf learners' congregation and bilingual learning. This is seen in the widespread implementation of cochlear implants without sign language learning as the universal standard of care for deaf children in Western nations, and the educational mainstreaming of deaf children as a main component and goal of this process (Holmström and Schönström, 2017; Komesaroff, 2007; Mauldin, 2016). In this way, special education for deaf learners has appropriated inclusion discourses "in order to maintain unreconstructed notions of schooling and educational defectiveness" (Slee and Allan, 2001, p. 174). For many deaf advocates, full inclusion is often perceived as synonymous with the loss of sign language vitality and community institutions. This perception is supported by the current state of deaf education in Canada.

Closing deaf schools

Across Canada, the closure of schools for the deaf has meant the loss of educational environments and personnel to fulfill the right to sign language-medium education. When the R.J.D. Williams School for the Deaf in Saskatoon, Saskatchewan closed in 1991, the provincial government chose to focus on the integration of deaf students in spoken-language classrooms (Snoddon, 2009). The school closure led to "the exodus of deaf teachers and teaching assistants and the diminishment of a vibrant deaf community" (Weber, 2020, p. 33). It also led to several human-rights complaints and legal judgments concerning families with deaf children who were unable to receive an education in sign language (Snoddon, 2009). However, none of these legal decisions led to the restoration of a school for the deaf or the broad implementation of American Sign Language (ASL) learning environments in the province. In 2016, the Saskatchewan child advocate released a report concerning the death in custody of 16-year-old Dylan Lachance, an Indigenous deaf and language-deprived youth who was unable to communicate with staff responsible for his care and died "of acute bronchopneumonia with associated sepsis" (Langenegger, 2016; Pringle, 2016). In Newfoundland, where the provincial school for the deaf closed in 2010, 9-year old Carter Churchill reportedly receives one-and-a-half to 2 h of instruction in ASL every seven school days (Lord, 2019). Carter's parents filed a human rights complaint in 2017 and subsequently sought a hearing in regard to their son's lack of access to an education equal in quality to that received by nondeaf students in Newfoundland (McKenzie-Sutter, 2019). Carter's situation has led to another provincial child and youth advocate report concerning deaf education, which is as lacking in practical solutions in its findings and conclusions as the report from Saskatchewan (Kavanagh, 2020).

While Kayess and Green (2017) argue that maintaining regular and special school systems results in fewer resources and that phasing out special schools will mean more resources for inclusive schools, this assertion is not borne out by what transpired in Canada when deaf schools were closed. Rather, following deaf school closures, there are minimal or nonexistent resources to support deaf children who benefit from sign language; this point underscores how inclusive education is seen to complement government austerity measures as well as oralism (Snoddon and Underwood, 2017). Deaf advocates from other countries also report the loss of sign language learning environments and a consequential decline in sign language and community vitality (De Meulder et al., 2019; Goico, 2019). Goico (2019) reports on the implementation of inclusive education in Iquitos, Peru, where individual deaf children are placed in regular classrooms without sign language support services. This phenomenon has triggered language shift from the Peruvian Sign Language used by deaf adults who attended special schools to "homesigns" invented by deaf children and youth who have little or no access to a natural sign language or spoken language at home or school.

Some inclusion advocates view providing a sign language interpreter as a solution for deaf learners' language and communication needs in a regular classroom (Russell and McLeod, 2009). However, Article 24 of the United Nations UNCRPD does not refer specifically to sign language interpreters (instead, these are covered in Article 9, which deals with the accessibility of public facilities). Rather, Article 24(4) requires governments to "take appropriate measures to employ teachers, including teachers with disabilities, who are qualified in sign language." Deaf advocates sought to ensure that for deaf learners, the scope of the rights contained in Article 24 is direct instruction in sign language from signing teachers (including deaf teachers) and alongside deaf peers, in addition to opportunities to study sign language as a school subject (Kauppinen and Jokinen, 2014). Although an interpreter is an important part of a range of educational options that should be available to deaf learners, this is not sufficient to ensure an inclusive education since providing an interpreter does not mean that students receive the same quality of instruction as nondeaf learners or enjoy the same opportunities for social development (Russell, 2021; Russell and McLeod, 2009; Snoddon and Murray, 2019).

However, some countries, such as Belgium, maintain deaf schools that are markedly poorer in quality than regular schools and do not follow the government curriculum (De Meulder and Murray, 2021). In keeping with the overrepresentation of minority children in special education, many of these schools host a higher proportion of immigrant children and children with multiple impairments. This illustrates the pervasiveness of a special education orthodoxy in systems seeking to be bilingual and inclusive. However, Moges (2020) argues that the history of the segregation of schooling for white and Black children in the USA led to better sign language outcomes for Black deaf children and the flourishing of Black American Sign Language (Black ASL) at Black deaf schools, since schools for white children were often oralist. This paradox of "Black deaf gain" also occurred in South African schools for Black deaf children. Following the dismantling of apartheid, which coincided with oralism for white deaf children, white children reported learning sign language from Black deaf children (Brown, 2020). Borrowing from Crenshaw's (1989) concept of intersectionality, this phenomenon makes visible the multidimensionality of experiences for people who are multiply oppressed by racism, sexism, and other systems of discrimination, and the importance of shared identities for deaf children's flourishing.

Where is a way forward?

While inclusive education advocates may often prioritize a regular school setting and a vaguely defined, aspirational set of resources to be provided in this setting, this often fails to provide meaningful guidance to governments or policymakers who seek to meet the needs of deaf learners. It is insufficient to claim that existing inclusive and deaf education systems are not truly inclusive and that continual educational transformation is needed while the resources that enable bilingual education provision continue to be stripped away. There is an ongoing struggle to maintain deaf schools and an array of placement options for other students that support language, culture, and identity development (Hall, 2002). Deaf advocates may thus perceive that they inhabit a hostile terrain beset by both inclusive education proponents and special education professionals.

Murray et al. (2020) describe a range of models of inclusive education for deaf learners. These include longstanding deaf schools that provide bilingual education in one or more natural sign languages and written/spoken languages while following the government curriculum. These schools typically employ a high proportion of deaf teachers and administrators, and sometimes provide sign language classes for parents of deaf children. The Maryland School for the Deaf employs the inclusive practice of allowing non-deaf siblings of deaf students to attend school on Mondays and Fridays (Fevlado, 2015). Deaf schools can also support deaf learners enrolled in regular schools and enable a model of simultaneous enrollment where students can attend classes at both a deaf and regular school (Murray et al., 2020). This model may especially but not only benefit deaf learners in rural areas, who live farther away from a deaf school.

Co-enrollment models involve the provision of a bilingual program at a regular school with deaf and hearing teachers and learners. These models can take various forms, including deaf student enrollment in both regular and separate classes at the same school, or deaf learners' attendance in a separate bilingual class inside a regular school (Murray et al., 2020). However, in practice deaf teachers in these programs may often have a subordinate role and deaf learners are marginalized, since hearing students are rarely expected to learn sign language (Murray et al., 2020; Snoddon, 2021a). These issues point to how inclusion is an ongoing process and an opening for dialog, rather than closure (Slee and Allan, 2001).

In seeking to enact inclusive education, deaf communities and educators need new activism, research, forms of schooling and teacher education, and reflexivity (Slee and Allan, 2001). In their overview of the current state of education for deaf students in Sweden, which was previously on the vanguard of the deaf bilingual education movement, Schönström and Holmström (2021) argue that a bilingual approach is even more relevant today than in the past. This approach needs to account for the broad linguistic repertoires of deaf students that may include spoken languages, and provide ongoing support to educators in terms of bilingual teaching practices (Schönström and Holmström, 2021).

In many countries, an education in sign language is frequently seen as the last resort, after deaf students fail to benefit from cochlear implants and integrated placements in regular schools. Inclusive education for deaf students should not only respect their educational rights as outlined in Article 24, but also affirm their identities as members of signing communities.

References

- Allan, J., 2008. Inclusion for all? In: Bryce, T.G., Humes, W.M. (Eds.), *Scottish Education: Beyond Devolution*. Edinburgh University Press, Edinburgh, pp. 701–710.
- Arnardóttir, O.M., 2011. The right to inclusive education for children with disabilities: innovations in the CRPD. In: Eide, A., Möller, J., Ziemele, I., Alfreddson, G. (Eds.), *Making Peoples Heard: Essays on Human Rights in Honour of Gudmundur Alfreddson*. Martinus Nijhoff Publishers, Leiden, The Netherlands, pp. 197–227.
- Barnes, C., 2019. Understanding the social model of disability: past, present and future. In: Watson, N., Vehmas, S. (Eds.), *The Routledge Handbook of Disability Studies*, second ed. Routledge, New York, pp. 14–31.
- Branson, J., Miller, D., 1993. Sign language, the deaf, and the epistemic violence of mainstreaming. *Lang. Educ.* 7 (1), 21–41. <https://doi.org/10.1080/09500789309541346>.
- Brown, R.L., 2020. How 2 Sign Language Interpreters Became Unlikely Pandemic Stars. NPR. Retrieved from: <https://www.npr.org/sections/goatsandsoda/2020/08/30/905417180/unusual-celebrities-sign-language-interpreters-for-south-africa-s-president>.
- Crenshaw, K., 1989. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine. In: *Feminist Theory and Antiracist Politics*, vol. 1. University of Chicago Legal Forum, pp. 139–167, 8. <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>.
- De Meulder, M., Murray, J.J., 2017. Buttering their bread on both sides? The recognition of sign languages and the aspirations of deaf communities. *Lang. Probl. Lang. Plan.* 41 (2), 136–158. <https://doi.org/10.1075/plp.41.2.04dem>.
- De Meulder, M., Murray, J.J., 2021. The illusion of choice in inclusive education. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2021.1956607>.
- De Meulder, M., Murray, J.J., McKee, R. (Eds.), 2019. *The Legal Recognition of Sign Languages: Advocacy and Outcomes Around the World*. Multilingual Matters, Bristol.
- Degener, T., 2016. A human rights model of disability. In: Blanck, P., Flynn, E. (Eds.), *The Routledge Handbook of Disability Law and Human Rights*. Routledge, New York, pp. 31–49.
- Fevlado, 2015. Vlaanderen is gelijke kansen. Retrieved from: <http://www.fevlado.be/fevlado-vzw/nieuwsprikbord/actualiteit/?d=600>.
- Florian, L., Spratt, J., 2013. Enacting inclusion: a framework for interrogating inclusive practice. *Eur. J. Spec. Needs Educ.* 28 (2), 119–135. <https://doi.org/10.1080/08856257.2013.778111>.
- Goico, S.A., 2019. The impact of “inclusive” education on the language of deaf youth in Iquitos, Peru. *Sign Lang. Stud.* 19 (3), 348–374. <https://doi.org/10.1353/sls.2019.0001>.
- Hall, P.J., 2002. Narrowing the breach: can disability culture and full educational inclusion be reconciled? *J. Disabil. Pol. Stud.* 13 (3), 144–152. <https://doi.org/10.1177/1044207302130030201>.
- Hall, W.C., Levin, L.L., Anderson, M.L., 2017. Language deprivation syndrome: a possible neurodevelopmental disorder with sociocultural origins. *Soc. Psychiatry Psychiatr. Epidemiol.* 52, 761–776. <https://doi.org/10.1007/s00127-017-1351-7>.
- Holmström, I., Schönström, K., 2017. Resources for deaf and hard-of-hearing Students in mainstream schools in Sweden: a survey. *Deaf. Educ. Int.* 19 (1), 29–39. <https://doi.org/10.1080/14643154.2017.1292670>.
- Kauppinen, L., Jokinen, M., 2014. Deaf culture and linguistic rights. In: Sabatello, M., Schulze, M. (Eds.), *Human Rights and Disability Advocacy*. University of Pennsylvania Press, Philadelphia, pp. 131–145.

- Kavanagh, J.L., 2020. The Sounds of Silence: Perspectives on the Education System's Response to Deaf and Hard of Hearing Children. Office of the Child and Youth Advocate, St. John's, NL.
- Kayess, R., Green, J., 2017. Today's lesson is on diversity. In: Blanck, P.D., Flynn, E. (Eds.), *The Routledge Handbook of Disability Law and Human Rights*. Routledge, New York, pp. 53–71.
- Komesaroff, L., 2007. Introduction. In: Komesaroff, L. (Ed.), *Surgical Consent: Bioethics and Cochlear Implantation*. Gallaudet University Press, Washington, DC, pp. xi–xiv.
- Kusters, A., De Meulder, M., O'Brien, D., 2017. Innovations in deaf studies: critically mapping the field. In: Kusters, A., De Meulder, M., O'Brien, D. (Eds.), *Innovations in Deaf Studies: The Role of Deaf Scholars*. Oxford University Press, New York, pp. 1–53.
- Ladd, P., 2003. Understanding Deaf Culture: In Search of Deafhood. Multilingual Matters, Clevedon, UK.
- Lane, H., 2005. Ethnicity, ethics, and the deaf-world. *J. Deaf Stud. Deaf Educ.* 10 (3), 291–310. <https://doi.org/10.1093/deafed/eni030>.
- Langenegger, S., 2016. Youth Justice System "Failed" Late Sask. Teen With Hearing Disability Who Tried Gesturing Severity of Illness. CBC News. Retrieved November 1, 2019 from. <https://www.cbc.ca/news/canada/saskatchewan/advocate-says-sask-teenager-failed-by-youth-justice-system-1.3636392>.
- Lord, R., 2019. "They Are Hurting Him": Newfoundland Family Fights for Education Equality. Global News. Retrieved November 4, 2019 from. <https://globalnews.ca/news/6113747/newfoundland-human-rights-complaint/>.
- Mauldin, L., 2016. *Made to Hear: Cochlear Implants and Raising Deaf Children*. University of Minnesota Press, Minneapolis, MN.
- McKenzie-Sutter, H., 2019. N.L. Father Says Review of Services for Deaf Children Must Include Families. CTV News. Retrieved November 4, 2019 from. <https://www.ctvnews.ca/canada/n-l-father-says-review-of-services-for-deaf-children-must-include-families-1.4660089>.
- Moges, R.T., 2020. "From white deaf people's adversity to Black deaf gain": a proposal for a new lens of Black deaf educational history. In: *Journal Committed to Social Change on Race and Ethnicity*, vol. 6, pp. 69–99, 1. <https://journals.shareok.org/jcscore/article/download/99/68/>.
- Murray, J., De Meulder, M., Le Maire, D., 2018. An education in sign language as a human right? An analysis of the legislative history and on-going interpretation of Article 24 of the UN Convention on the Rights of Persons with Disabilities (CRPD). *Hum. Right Q.* 40, 37–60. <https://doi.org/10.1353/hrq.2018.0001>.
- Murray, J.J., Snoddon, K., De Meulder, M., Underwood, K., 2020. Intersectional inclusion for deaf learners: moving beyond general comment no. 4 on Article 24 of the UNCRPD. *Int. J. Incl. Educ.* 24 (7), 691–705. [10.1080/13603116.2018.1482013](https://doi.org/10.1080/13603116.2018.1482013).
- O'Brien, D., 2021. Bourdieu, plurilingualism, and sign languages in the UK. In: Snoddon, K., Weber, J. (Eds.), *Critical Perspectives on Plurilingualism in Deaf Education*. Multilingual Matters, Bristol, UK, pp. 60–80.
- Pringle, B., 2016. The Silent World of Jordan: Special Investigation Report. Saskatchewan Advocate for Children and Youth, Regina, SK. Retrieved November 25, 2019 from. <https://www.saskadvocate.ca/sites/default/files/u12/The%20Silent%20World%20of%20Jordan%20Public%20Report%20June%2015%202016.pdf>.
- Russell, D., McLeod, J., 2009. Educational interpreting: multiple perspectives of our work. In: Mole, J. (Ed.), *International Perspectives on Educational Interpreting*. Direct Learned Services, Ltd, Brassington, UK, pp. 128–144.
- Russell, D., 2021. Critical perspectives on education mediated by sign language interpreters: inclusion or the illusion of inclusion? In: Snoddon, K., Weber, J. (Eds.), *Critical Perspectives on Plurilingualism in Deaf Education*. Multilingual Matters, Bristol, UK, pp. 217–239.
- Schönström, K., Holmström, I., 2021. Four decades of sign bilingual schools in Sweden: from acclaimed to challenged. In: Snoddon, K., Weber, J. (Eds.), *Critical Perspectives on Plurilingualism in Deaf Education*. Multilingual Matters, Bristol, UK, pp. 15–34.
- Slee, R., Allan, J., 2001. Excluding the included: a reconsideration of inclusive education. *Int. Stud. Sociol. Educ.* 11 (2), 173–192. <https://doi.org/10.1080/09620210100200073>.
- Snoddon, K., 2021a. [Review of Co-enrollment in Deaf Education]. *Deafness and Education International* 23, 142–144. <https://doi.org/10.1080/14643154.2020.1804710>.
- Snoddon, K., 2021b. Sign language planning and policy in Ontario teacher education. In: Snoddon, K., Weber, J. (Eds.), *Critical Perspectives on Plurilingualism in Deaf Education*. Multilingual Matters, Bristol, UK, pp. 35–59.
- Snoddon, K., Murray, J.J., 2019. The Salamanca Statement and sign language education for deaf learners 25 years on. *Int. J. Incl. Educ.* 23 (7–8), 740–753. <https://doi.org/10.1080/13603116.2019.1622807>.
- Snoddon, K., Underwood, K., 2014. Toward a social relational model of deaf childhood. *Disabil. Soc.* 29 (4), 530–542. <https://doi.org/10.1080/09687599.2013.823081>.
- Snoddon, K., Underwood, K., 2017. Deaf time in the twenty-first century: considering rights frameworks and the social relational model of deaf childhood. *Disabil. Soc.* 32 (9), 1400–1415. <https://doi.org/10.1080/09687599.2017.1320269>.
- Snoddon, K., 2009. Equity in education: signed language and the courts. *Curr. Issues Lang. Plann.* 10 (3), 279–295. <https://doi.org/10.1080/14664200903116303>.
- UNESCO, 1994. World Conference on Special Needs Education: Access and Quality. Final report. Retrieved from. <https://unesdoc.unesco.org/ark:/48223/pf0000110753>.
- United Nations Convention on the Rights of Persons With Disabilities, 2006. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- Weber, J., 2020. Interrogating sign language ideologies in the Saskatchewan deaf community: an autoethnography. In: Kusters, A., Green, M., Moriarty Harrelson, E., Snoddon, K. (Eds.), *Sign Language Ideologies in Practice*. Mouton de Gruyter, Berlin, pp. 25–41.
- Wilcox, S.E., Krausneker, V., Armstrong, D.F., 2012. Language policies and the deaf community. In: Spolsky, B. (Ed.), *The Cambridge Handbook of Language Policy*. Cambridge University Press, Cambridge, UK, pp. 374–395.

Girls, autism identities, and education

Suzanne Carrington, Beth Sagers, Keely Harper-Hill, Rachel Aberdein, Jodie Wilde, Hannah Smith, and Bella Harper-Hill,
Center for Inclusive Education, School of Early Childhood & Inclusive Education, Faculty of Creative Industries, Education and Social Justice, Queensland University of Technology, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	485
Education and identity	486
Autism and girls	486
Autistic voice	487
Methodology	487
Autoethnographic tales of school experience	488
Rachel's tale	488
Jodie's tale	488
Hannah's tale	489
Bella's tale	489
Analysis of topics and themes	490
Topic 1: school experiences	490
Theme 1.1: positive and negative experiences at school	490
Theme 1.2: teachers did not know what to do or how to help	491
Theme 1.3: teachers had a negative view of difference	491
Topic 2: identity	491
Theme 2.1 relationships: friendships, belonging, and others' perceptions of you	491
Theme 2.2: masking/masquerading	492
Theme 2.3: diagnosis and being autistic	492
Theme 2.4: impact on identity	492
Topic 3: recommendations	493
Theme 3.1: advice for teachers	493
Theme 3.2: helpful support	493
Discussion	493
Conclusion	495
Appendix	495
References	495

Introduction

In the past, dominant discourses about autism in education contexts have largely been deficit-based and perpetuate marginalization, stigma, and dependence (Urbanawicz et al., 2019). This has led to “normalizing” students with disability to enable them to fit into existing educational structures, settings, and practices rather than implementing systemic changes that enable teachers to adjust their practices to cater to the diversity of learners and support inclusion of all students (Bajada et al., 2021; Liasidou and Ioannidou, 2021; Shume, 2020). The consequences of stigma and marginalization for the autistic population are well documented (Botha and Frost, 2020) and are seen from an early age. It is estimated that 71% of autistic adolescents meet criteria for mental health conditions (Hull et al., 2021); forming a positive autistic identity has been related to lower rates of depression and anxiety (Cresswell and Cage, 2019; Shakes and Cashin, 2019) and improved self-esteem (Cooper et al., 2020). In this article the research team, consisting of four female researchers who identify as autistic and three researchers who do not, used collaborative autoethnography (Chang et al., 2016) to reflect on how educational experiences in the school years may nurture or erode identity for autistic girls, and how this can inform more inclusive educational support for this group of learners.

Autism is best known for early independent descriptions provided by Kanner (1943) and Asperger (1944). Since then, there has been debate about what autism is, and who can be considered autistic with the ongoing evolution of diagnostic criteria reflecting this debate. One important event in the expansion of the spectrum of autism was Wing's (1981) paper *Asperger's Syndrome: A Clinical Account* based on Asperger's original accounts. This helped raise awareness that autistic people could be highly verbal and have normal to superior IQs. Since then, there has also been increasing recognition that women can be autistic, and that autism in females may not present in the same way as in males. It is timely to hear from autistic women about how their school experience impacted on their identity.

Education and identity

While dictionary definitions of identity are not unified (Lopes, 2018), the Cambridge Dictionary (2022) defines identity as “the fact of being, or feeling that you are, a particular type of person ... the qualities that make a person ... different from others”. Importantly, Lopes (2018) defines identity as:

Identity is concerned largely with the question: “Who are you?” which can be broken down into our sense of self, or the global understanding a person has of themselves; including roles, attributes, skills, abilities, disabilities, behaviors, and associations that we consider most important to us. Identity relates to our basic values that dictate the choices we make, and these choices in turn reflect who we are and what we value.

In comparison, Rachel – an autistic woman and co-author of this article – described identity in the following way:

I see identity as the core of who you are as a person, your fundamental beliefs and temperament. This is separate from what you do and how you act, but it informs these things ... I see identity as being separate from role, but the two can sometimes overlap. For example, someone may take on roles as a parent, as an employee, and as a customer, but they may only consider “parent” to be part of their identity.

Erikson’s (1968) influential psychosocial model suggests that identity development is the product of both psychological and social processes and that it is an ongoing, active process that occurs across the lifespan. Although Erikson recognized the importance of social influences, his work extended the works of psychoanalysts including Freud, and focused on ego and personal identities. Ego identity includes those traits and characteristics that are often unconscious (Ehala, 2018). Personal identity includes the conscious cognitive processes used to understand the world and who we are, as communicated through our values, attitudes, and goals (Kerpelman and Pittman, 2018).

Critical to notions of identity is that “everybody has a sense of self or sense of personal identity” (Lopes, 2018). This sense of identity is fluid and evolving based on the person’s experiences in different contexts and in the roles they have in their lives (e.g., family, school, social relationships, abilities and disabilities; Lopes, 2018), reinforcing the significant role schools have in nurturing a strong and positive sense of identity for all learners across their childhood and adolescence. In addition to these ideas, the concept of intersectionality (Crenshaw, 1991) is useful to understand personal and social identity configurations and illuminate “the distinctiveness of experience that comes from how one is treated as a member of multiple, major social groups” (Shmulsky and Gobbo, 2019, p. 648). Intersectionality is a framework for understanding how social identities such as gender, race, ethnicity, social class, religion, sexual orientation, and ability overlap with one another and with systems of power that oppress and advantage people. Consideration of these overlaps supports our understanding of the data reported in this chapter, interpretation of meaning and generation of new ideas to support inclusive practices in schools for autistic girls and nurture their developing identities.

Autism and girls

Until recently, there was a disproportionate ratio of males with a diagnosis of autism in comparison to females, with ratios of 4–5 males to every one female diagnosed (Milner et al., 2019). Current research has challenged these ratios and produced evidence for underdiagnosis in girls and a biased understanding of the spectrum and methods of diagnosis based on a male model of autism (Milner et al., 2019). Recent qualitative research into autistic female experiences reinforces these issues, with many describing how their autistic tendencies went underdiagnosed to late adolescence or early adulthood (Milner et al., 2019). These experiences have highlighted how autistic females often use a technique described as “masking” to help them fit in and hide their autistic tendencies (Bargiela et al., 2016; Cage and Troxell-Whitman, 2019; Hull et al., 2017; Milner et al., 2019).

Masking or camouflaging autistic behavior occurs in response to external pressure to appear “normal”. This pressure can be direct – through interventions to address perceived deficits – or indirect, with masking acting as a protective mechanism against stigma and discrimination (Cage and Troxell-Whitman, 2019; Hull et al., 2017). The term “masquerading” has also been used to describe how feelings, emotional reactions, and stress are masked in the school environment (Carrington and Graham, 2001). Multiple studies examining the causes and consequences of masking have been carried out in recent years and there is a growing body of evidence demonstrating the adverse mental health outcomes associated with masking autistic behaviors (Cage and Troxell-Whitman, 2019; Hull et al., 2021; Schneid and Raz, 2020). While studies looking into the masking behaviors of autistic females suggested they proved successful in helping them try to fit in socially, it was only a short-term, superficial way of coping and masked their feelings of unhappiness, anxiety, and loneliness (Cage and Troxell-Whitman, 2019; Hull et al., 2021; Schneid and Raz, 2020). We suggest that masking or masquerading behavior is therefore an important consideration in this discussion because the social influences at school influence identity development. We also suggest that, given the historical gender bias in autistic research, it is important for us to look at and investigate the autistic female experience in more depth.

Autistic voice

An inclusive research approach assumes that the lived experiences of autistic individuals should inform future planning in education (Lawson and Carrington, 2021). Recently there has been a growth in research that engages the autistic community as research partners (Pellicano et al., 2019). Autistic educator, Trudy Bartlett, suggests that listening to the voices of people on the autistic spectrum can be transformational, stating:

As a student myself I had a challenging time in high school. I was seen as a naughty, disruptive student. I was often issued consequences like detentions and even put on behavior monitoring programs because of my behavior in class ... Almost all of my teachers didn't ever ask me what was going on in my world, what they could do to help me learn, how could they make adjustments to help support me in the classroom ... They constantly told me I was not good academically; they told me that I wouldn't be successful when I finished school.

Bartlett and Carrington (2021: 55).

Engaging the autistic community by providing opportunities to hear the autistic voice is powerful. Importantly, based on previous research bias in this field, listening to, understanding, and learning from the experiences of autistic women is critical for informing change at systematic, school, and classroom levels.

Methodology

A sociocultural approach (Vygotsky, 1978) underpins this study exploring girls, autistic identity, and education. We were interested in investigating how various sociocultural factors at school influenced the development of autistic identity for four women. This project involved three researchers who are not autistic (Suzanne, Beth, Keely) and four researchers who identify as autistic (Rachel, Jodie, Hannah, Bella). All four women had received late diagnoses of autism.

Drawing on a collaborative autoethnography approach (Chang et al., 2016) we worked together across four stages of research (Table 1).

Collaborative autoethnography is an iterative, qualitative research method in which researchers work together to collect autobiographical data and to analyze and interpret data collectively to gain a meaningful understanding of sociocultural phenomena (Chang et al., 2016). In this project, the four autistic women contributed to the development of guiding questions (see Appendix) for Stage 1 data collection. They then documented and shared their autobiographical data with the whole collective (authors of this article). The autobiographical data are descriptions about their experiences at school and personal reflections about how these experiences impacted their developing identity.

The research questions guiding the project were:

1. How do educational experiences in the school years nurture or erode identity?
2. How can collaborative autoethnographic reflections about educational experiences from four autistic women support our understanding of what is needed to improve support for autistic girls at school and nurture positive identity?

In the following sections, we describe our process for working collaboratively in the collective, and present the four autoethnographic *tales* which we agreed provided unique insights on the range of school experiences. We begin with Rachel who highlights the lack of teacher support and understanding about autism.

Table 1 Applying a collaborative autoethnography approach across four stages of research activities.

Stage	Research activity
1	Preliminary data collection: Individual self-writing and reflection and sharing in collective group
2	Subsequent data collection: Group sharing and discussion via online meetings and preliminary meaning-making
3	Data analysis and interpretation: Individual and collective analysis and interpretation
4	Collaborative group writing

Autoethnographic tales of school experience

Rachel's tale

In primary school, teachers didn't give me any extra support in learning from an academic viewpoint as it wasn't necessary. The tone taken by some of the teachers during the support could have been better, as I was sometimes treated as someone with a character flaw rather than a child having trouble with something. I don't think it was personally targeted at me, though: those teachers were like that with everyone. In secondary school, the only difference in learning support was that I was penalized for poor bookwork rather than offered support. My reaction to this was to feel annoyed that my grades could be brought down by an entire letter just because my books didn't look good enough for the teacher.

Rachel suggested that she received little support at school and comments about the annoying focus on bookwork influenced her belief that her work was not good enough. She continues to explain that although she was confident in her ability as a learner, she lost interest in doing well because the teachers did not align the teaching of the curriculum to her needs and interests.

Throughout all of school, I felt confident about myself as a learner and was aware of my strengths. Throughout high school, my average grades became worse than they were in primary school because that I could no longer be bothered putting any effort into subjects that didn't interest me. I suspect that the only way that teachers could have mitigated this would have been to show how their subjects could be relevant to things I was interested in. In terms of social support in primary school, my teachers didn't offer any support with making friends and I don't think it would even have occurred to them that was something they should do.

Rachel reflected on the importance of nurturing strengths and interests to support motivation for learning and success in school and how this could have further supported a positive sense of self and identity. She indicates that she received no teacher support in the social aspect of making friends which could have been important in promoting her sense of self and identity within the group context of school. These reflections also highlight how educators never consulted with Rachel about what she might need to help promote her success in the school environment and in turn positively influence her identity in that context.

Although I did have friends, I never had a sense of belonging. This is something that is innate to me rather than the fault of anyone or of any environment that I've been in. I don't do "sense of belonging". Even in groups where I'm a long-term and valued member, I don't have a sense of "us". It's always "me" and "all those other individuals". This doesn't mean, however, that I don't enjoy the company of others in my social group, or that I don't notice when I'm accepted and valued by a group. I'm quite happy doing my own thing. I don't see the point of trying to force a connection just for the sake of company, it seems like a lot of work for the "reward" of having to spend more time and effort forcing said connection. Similarly, my peers didn't particularly influence my sense of self and my sense of belonging at school as my sense of self isn't based on what other people are doing, and school was just somewhere that I had to go, so it didn't matter to me whether I belonged there or not. Ever since preschool, I've known that I was fundamentally different in some way from most other people, but I've also thought that this doesn't affect my worth as a person. I don't see myself as better or worse than anyone, just different.

Rachel reflects on how a sense of belonging and connection to the large group is not essential to supporting her identity about who she is and her sense of self.

Jodie's tale

Jodie's tale highlights the challenges she had at school. Similar to Rachel, Jodie was also achieving good academic results, but was finding it challenging to navigate "fitting in" with her peers and making friends. Her teachers also did not provide good support and had no understanding of the difficulties she was experiencing in the school context.

I remember being aware that I was different from my peers from a very early age but the idea that I was neurodivergent never occurred to anyone, especially not my teachers. In many ways, I excelled in school. I found reading and writing easy and was intensely interested in learning about the world. I never appeared to need support, but the reality was that every day was a struggle. I vividly remember sitting and watching the other children interact and trying to figure out what I needed to do to "fit in" and group activities filled me with dread. I was desperate to understand how the world worked, and why other people just seemed to "get it" and I became a voracious reader, trying to learn as much as I could in the hope it would help me understand the unwritten rules. This helped a little; however, I still found the social and non-academic aspects of school extremely difficult. In contrast, I was often praised for my academic performance, and I began to link academic success to self-worth and to see my social difficulties as a personal failing. Although I had a few friends, friendships were always difficult for me and I often preferred to play alone. The only advice that teachers offered was to tell me I needed to be "more flexible" and "just learn to get along". They obviously expected that I should know how to do this and that I was deliberately choosing not to be flexible and get along. However, with no guidance on what this meant, it led me to believe that I needed to hide my own feelings and needs and do whatever people wanted if I wanted to have friends and fit in. As I progressed through school and academic success became less socially desirable, I was labeled a "swot" and "teacher's pet" by my peers. This caused a great deal of distress and there was significant conflict between my desire to fit in with my peers – and potentially "fix" what I saw as a personal failing – and my desire to succeed academically, which underpinned my sense of self-worth.

As Jodie progressed through her schooling with little support, labeling and marginalization impacted her mental health and influenced her sense of identity within the school context.

The added pressures of high school also meant that my executive functioning and sensory processing differences became much more apparent and, although I still performed well academically, I struggled significantly just to get through the day. With no social

network to help support me, my mental health began to deteriorate rapidly, I began to self-harm, and I would often stay home “sick” as I become increasingly overwhelmed. Although I have some good memories of teachers who were engaging and whose classes I enjoyed, I received no support beyond general classroom instruction, and I felt largely invisible to both my peers and my teachers (several teachers who taught me throughout high school still did not know my name by the time I left). No teacher ever spoke to me about my wellbeing, despite becoming increasingly withdrawn and showing obvious signs of depression, and my absences were never raised as a concern because I continued to perform well academically.

Jodie, like Rachel, started believing in herself despite receiving little or no support from her teachers in ways that would support her connection and wellbeing in the school context and in developing a positive identity.

I made the deliberate decision to (essentially) not care whether people “liked” me or not and to stop trying to fit in. I realized that I had been masking my true personality because I thought it was the only way for people to like me, but I was still having the same social difficulties as before so what was the point? I also realized that I had been looking for friends in the wrong places, trying to fit in with groups that I thought I “should” be friends with, not people I connected with. My final two years of high school were a complete contrast. I (mostly) accepted myself and embraced the fact that I was different from my peers and found a very eclectic group of people who I connected with and who made me want to go to school each day – and it’s probably no coincidence that many of us have since discovered we are neurodivergent!

Hannah’s tale

Hannah’s tale also highlights her academic strengths, but illustrates the profound impacts teachers and their practices can have on a learner’s sense of self and identity.

When I started Prep, my teachers praised me for my ability to write in complete sentences while the other kids were scribbling. Despite attending a private school with resources, I was not offered any enriching opportunities to further nurture my abilities. So, as I progressed through junior school, while I continued to enjoy learning at home in my own ways, my achievements and growth at school never peaked in the way I think I could have if I had the opportunities.

In contrast, the following excerpt shows how positive teacher support can nurture positive identity, highlighting the powerful influence any single supportive teacher can have, or conversely be detrimental by not valuing or accepting a student’s interests and strengths.

[Teacher] made learning inspiring for me. On an individual level, he listened to whatever I had to say about my interests (sharks and politics), and on a curriculum level he allowed me to learn through my interests.

At the end of the year, before summer holidays, [Teacher] suggested I do something real-world with my interest in sharks. This encouragement led me to being able to do speeches for a local chapter of Sea Shepherd on shark conservation. However, when I returned to school for Grade 11, I was bullied because of my interest. It sounds harmless, but the way it completely isolated me from peers, and the way it made teachers treat me, really cemented how much I did not belong in the schooling system. [A different teacher] often talked about sharks in math class (in a mocking way) ... Once, he asked what my plans after graduation were. When I told him, he said in front of everyone, “Why can’t you just be normal?”. It was humiliating. In modern history, I never felt comfortable asking for help because the class dynamic was one big clique, and the teacher favored the bullies. [I felt] very let down by the system, like I was “too much” or “too weird”, and like I was expected to be someone I wasn’t if I wanted help. Those attitudes reinforced the behaviors of my bullies, teaching them: “it is okay to treat different people badly”.

Hannah describes how her negative school experience impacted on feelings of self-worth and identity, and acknowledges the positive impact of a good student–teacher relationship, which is critical to inclusive practice and a sense of belonging for all students.

The teachers who actively participated in and/or enabled my bullying, as well as the bullies and everyone who had been influenced by them (one of my best friends was too embarrassed to hang out with me anymore) affected my learning, my mental health, and my sense of self. I never felt the need to fit in – I didn’t want to change who I was, I didn’t think there was anything “wrong” with me and in fact I liked me very much, but I knew I was different and struggled to understand why. I felt very depressed and had suicidal ideations.

Bella’s tale

Bella’s school experience is more recent and highlights her awareness that she was different and how her teachers did not know how to support her learning and participation at school.

They supported me but they didn’t know what skills I needed from them to do my work. Just because I need help doesn’t mean I’m dumb, that I can’t do things. When a teacher said my work was nonsense that’s when I felt not very smart. It made me quite anxious, and it made me not very confident in myself. I think it’s hard for teachers to know how to help students to make friends. They don’t really know what kids need to make friends. Once I had friends, it was really hard to keep them because I didn’t have the skills and know how to keep them.

Table 2 Topics and themes.

<i>Topic</i>	<i>Theme</i>
1. School experiences	1.1 Positive and negative experiences at school 1.2 Did not receive support at school and teachers did not know how to help 1.3 Teachers had a negative view of difference
2. Identity	2.1 Relationships: Friendships, belonging, and others' perceptions 2.2 Masking/masquerading and being autistic 2.3 Diagnosis and being autistic 2.4 Impact on identity
3. Recommendations	3.1 Advice for teachers 3.2 Helpful support

Bella described how the social experience at school impacted negatively on her identity, and how the interactions with teachers and peers influenced her sense of self and belonging.

It sucked because I knew I could do it but I didn't know how to do it. It also sucked because I didn't have the confidence to ask the teacher for help so I would write on my paper "I need help", but that didn't help me. But when a teacher explained what I needed to do, I still didn't understand what I had to do which made me feel very confused and lost. The fact that I was so little – people didn't think that I could do things. That made me very upset and not very capable that I could do things. And defeated. It made me really upset to go to school. I never wanted to go to school. I pretended to be sick a lot and that's when my anxiety really developed, and I cried a lot. I was the outsider or the left over in the class. I was always left out and everyone was upset that they got me and not their friend. Everyone had friends and they were laughing and joking, and I was counting down the minutes to go home and be with my family where I felt safe. I really hated school. I didn't think I was very smart but I knew I was good at some things. Academic wise I found it pretty tricky back in the day. I did really try hard but people just thought I was being dumb. The worst thing a teacher can do is when they compare you to an age group and say, "they can do this, why can't you". Well, they can do this but I can't. It doesn't give you any confidence to do this, it doesn't help the kid.

Once the members of the collective had reviewed the four autoethnographic tales, we met together using an online meeting platform to discuss common themes that emerged across the individual tales. The group discussion was audio recorded and transcribed for analysis and interpretation. This was Stage 2 in the four-phase process (Table 1).

In Stage 3 we used a reflective method of analysis and interpretation (Braun and Clark, 2021), informed by the sociocultural approach (Vygotsky, 1978) and the literature that informed the study. The group interview transcript was read several times and data topics were identified. The themes were generated to describe patterns of meaning in the data that are united by ideas. Individuals and the collective provided comments and feedback which resulted in the final version (Table 2).

Analysis of topics and themes

Topic 1: school experiences

This topic reflects our discussion of school experiences that influenced identity in negative and positive ways.

Theme 1.1: positive and negative experiences at school

We all agreed the autoethnographic tales described negative experiences at school that resulted in disengagement or fear and negatively influenced the women's sense of self and identity. For example, Hannah said, "I've never failed a class or anything, so I know I'm very capable of all of that stuff, but I was just so uninspired, and not engaged, and I already didn't like the teachers. So even though I probably could have done the maths work, it was just 'I can't do it'". Comparatively, Jodie told us that she was

terrified of going into some classrooms: remember being forced to hold my pencil in a specific way when I was in probably Grade 2. You know, and that teacher, that's all I remember about her, is standing over me, like constantly forcing me to change the way I held my pencil. I was writing. I wrote legibly. But it just wasn't, you know, the standard pencil grip that you're supposed to use. And all I remember is her constantly telling me I was doing it wrong. And then they, and then I had her again, I think in Grade 5, because they switched the teachers around. And when I knew that I was having her, I was absolutely terrified of going in her class.

What is also apparent is the lasting effect that negative experiences had. For example, Hannah "already didn't like the teachers" and Jodie was still terrified of one of her early teachers three years later, reinforcing how one single negative experience can have a long-term influence on identity.

In contrast, Beth noted that some tales described the significant positive impact teachers supporting special interests could have, acknowledging when something was done well and making efforts to connect with students through these areas of interest. Jodie followed up by saying,

the teachers that were good, it's not that they necessarily did anything specific to support me, it's just that they, I felt like they liked me, I guess. You know, they were interested in me, and they were, yeah, I actually had some kind of relationship with them. And I think that was probably the major difference between feeling supported by them, even though they didn't necessarily give me anything specific, I just felt a lot more confident and comfortable, I guess, in their classes, than with other teachers.

It is clear from Jodie's comment that positive feedback and interest supported both a positive relationship and her confidence. She went on to say, "so when you kind of find people who can see other strengths, it really has a positive impact on your sense of self". Hannah also described some positive support from teachers. Even though the specifics of the support couldn't always be identified, these comments reflected that feeling accepted and valued had a lasting and positive impact.

Theme 1.2: teachers did not know what to do or how to help

We all agreed every tale showed a lack of support, or the right type of support, and limited evidence for student consultation to map out what support might look like: "whether we needed a lot of support in one area or not, like whatever our strengths were, we all didn't get the support that we needed to do better" (Hannah).

Rachel suggested, "I sort of noticed a pattern of teachers not really knowing what to do with us. Like, they knew something was different, if they noticed, but they didn't know what to do about it or how to support us". We all agreed with this statement.

Theme 1.3: teachers had a negative view of difference

The reflections often suggested teachers had negative views of difference that influenced the type of support that was provided, having a negative influence on students and their student-teacher relationships. For example, Hannah said, "even though we had different experiences, the teachers still responded to those things actually negatively". Jodie said, "I don't think that many teachers that are actually aware of how negative they are toward people in general". It is also important to note that none of the women reflected that they had at any time been consulted about what might help support them with their learning, interactions, and participation.

Topic 2: identity

Topic 2 has a number of themes highlighting how the development of identity is influenced by social interactions with others in the school community. These experiences are important to the development of identity, particularly in adolescence; however, whether those experiences are positive or negative is what may influence identity in destructive or helpful ways (Abbasi, 2016).

Theme 2.1 relationships: friendships, belonging, and others' perceptions of you

Rachel and Hannah agreed that they did not really care about not having many friends as reinforced by Rachel's statement, "I feel like I should set up an automatic Facebook message that says, 'Sorry I haven't replied to you in 6 months, we're still friends, the problem is entirely me'".

Realization and acceptance of themselves as "belonging differently" was echoed by both Rachel and Hannah and clearly time-stamped by Jodie:

I mentioned that, like, around Grade 9, I don't know, I had a bit of an epiphany, similar, and started to really not care too much, about people's perceptions of me, and not try to fit in. And I think that was, I don't think it's that we don't need to belong, necessarily, it's just a different type of belonging.

While "belonging differently" was acknowledged, Hannah observed that all tales still talked about not fitting in and that this was challenging and had an impact on identity, "*we all still knew that we couldn't fit in and we all still struggled in that way*". In contrast to this being a struggle at school, Jodie described how in adulthood she was now selective about fitting in, "*I don't want to try to fit in with people that I actually don't really like that much, I don't have much in common with*". The importance of acknowledging, understanding, and respecting a sense of connectedness and belonging differently for this cohort of learners is important for schools to understand so they can support the development of identity in positive ways. Further to belonging differently was the realization the women made during their schooling that what others thought of them was not important.

Theme 2.2: masking/masquerading

Similar to other experiences reported in research (Bargiela et al., 2016; Cage and Troxell-Whitman, 2019; Hull et al., 2017; Milner et al., 2019), Jodie and Bella shared examples of how they masked/masqueraded within the social context of school to fit in with peers and belonging. Jodie said:

That was kind of like my whole school experience, until yeah, probably Grade 9 ... it was just so stressful, trying to dig up this internal conflict going on all the time, over yeah, I just want to kind of be myself ... As in, you don't want to fit in to that kind of situation, or that kind of group, necessarily, but you want to find somewhere where you do fit in, I guess. And so you're just sort of masking to try and find somewhere where you fit in ... over the years I've realized that masking in the long-term, like, isn't beneficial. But it's so ingrained, it's really quite difficult not to do, I guess, especially when you meet new people. Yeah, really not good for my mental health, I know that for sure.

Bella described the challenges she experienced with fitting in and the negative impact this had:

because I struggled so much to like feel like included in things, I would just sit with anyone, not anyone, but like, people didn't actually know the real me, I think, they didn't know the real me, because I would pretend to be someone I was not.

In contrast, Rachel said:

I think my experience of masking is kind of different from a lot of the accounts of autistic people that I've heard. Because I've heard of masking as a lot of suppressing of traits, like "Don't stim, don't talk too much about your interests", but if I'm just left to my own devices, I will, I won't talk someone's ear off, I will just go away and ignore everyone and read about stuff that I find interesting. So for me, it's been more about constructing a social persona for people to interact with, and actually remembering to talk to people, and show interest in people, and things like that.

Aligned with previous research findings, these experiences reinforce masking is only a short-term, superficial way of coping, masking feelings of unhappiness, anxiety, and loneliness (Cage and Troxell-Whitman, 2019; Hull et al., 2021; Schneid and Raz, 2020). Intersections of autism and gender are evident in the tension between conforming to the school girl culture displayed by masking/camouflaging and accepting their differences.

Theme 2.3: diagnosis and being autistic

All the women reported how they knew they were different at school, but did not have a diagnosis at the time. Bella suggested, "I think if I was diagnosed with autism when I was younger, I think it would have just made my schooling a whole lot worse for me. Cause I would have had the label of autism". Jodie said, "I'd never even heard of autism when I was a kid. I didn't know anyone who was diagnosed. I'm sure they were, but you know, it wasn't spoken about or anything. So yeah, I think you're probably right. And yeah, maybe going forward, hopefully it'll make a difference". Hannah commented, "If I had the diagnosis, it would have just given the teachers more to pick on me for".

We all acknowledged that some of the tales described the autistic traits which the women gradually accepted as important to their own identities. They described their special interests and changing obsessions, how they enjoyed being with friends and siblings who also received a late diagnosis of autism, the challenge of keeping up appearances of friendships, and meeting various social expectations. These intersectional experiences of autism and gender had a strong impact on identity.

Theme 2.4: impact on identity

Rachel and Hannah shared stories about how their experiences impacted their own identity. Rachel said:

in addition to the fact that I had enough positive feedback coming in, like I was always into music, and even teachers whose classes I was failing knew that and respected me for it, was that I've always had a very clear sense of who I am, and what I want, and where I want to go. And other people's opinions on that have always been kind of irrelevant to me.

In contrast, other experiences within this theme highlighted occasions when teachers' actions or speech explicitly impacted students as being different or not belonging.

Topic 3: recommendations

As a collective, we didn't talk in depth about ideas that would help teachers in the future. As a result, the data reported below is taken directly from the autoethnographic tales and addressed the following questions:

- Considering your own experience, what advice do you have for future school teachers who have girls on the spectrum in their classroom?
- What characteristics of routine teacher interactions cumulatively impact on student identity over the primary and secondary years; consider positive and negative interactions. What did teachers say and do that helped you to see yourselves as capable versus incapable?

Theme 3.1: advice for teachers

Rachel

The advice I would give future teachers who have girls on the spectrum in their classroom is to teach them to be assertive. Teach them that they are just as good as anyone else, and that they deserve a place in the world just as much as anyone else. Teach them that they don't need to be agreeable and acquiescent. And make sure you show this through your actions, because words are cheap, but if you're telling a person they're great and valuable while acting dismissive toward them, ignoring bullying, and constantly acting as though their neurotypical peers are right and they're wrong, they will notice, and they will internalize it.

Hannah

Advice for teachers with autistic girls in their class: do not expect "different" students to be more like neurotypical kids. Invest time into developing a good relationship, and then use your position to empower them to embrace their identities, their unique strengths, and their interests. Also, don't ignorantly uphold the traditional and pathologizing ideas of autism: learn about autism through the neurodiversity model.

Jodie

My main advice for teachers would be to get to know your students as much as possible, make sure you are aware of the less well-known indicators of autism and neurodivergence, seek information about autism from autistic students and/or adults, and remember that it is vital that you consult with the student about their needs. If my teachers had known me better, I may have received my diagnosis much earlier in life (I was diagnosed at age 36, following my daughter's diagnosis) and this could have made a huge difference, not just at school but in all aspects of my life. Routine teacher interactions can have a huge impact on identity, and even positive ones can have a negative impact. Language and framing can make a big difference. The combined effect of being praised for academic success rather than effort and the expectation that social interaction should be effortless led me to believe that self-worth was intrinsically linked to success and/or failure and that the amount of effort was meaningless.

Bella

Find out what people are into and use that to get them to tick. I know that you have to have a lesson plan but some kids, if it's a PowerPoint presentation we don't know what to do with that. Be flexible, go with the flow and have little bits so that if a teacher (this is hard trying to find out what to say, I sometimes I find it hard to put things into my own words) ... Everyone has a battery so if their battery is full, don't expect them to be able to go home and write up an essay. It takes me 3 h after school to decompose everything that has happened in the day. If you push it we won't do so great the next day at school.

Theme 3.2: helpful support

Rachel said:

The thing that teachers did in the course of the everyday that helped me to see myself as capable was to treat me as capable and to have high standards for me. Even when I had trouble with something, their attitude was "you need to practise this more", not "you can't do this".

Discussion

In this section we have taken a collective approach to interpreting the data, drawing on a collaborative autoethnographic approach. The opportunity for collaboration within the collective promoted self-reflexivity associated with autoethnography, intersectionality, cultural learning and interpretation, and multi-subjectivity associated with collaboration (Chang et al., 2016). We worked together across the four stages of research (Table 1) to address the research questions for the project:

1. How do educational experiences in the school years nurture or erode identity?

The women's tales describe negative experiences at school and illustrate how the social and cultural interactions had a lasting negative effect on their self-identity. It seems that teachers' interactions were influenced by deficit-based views about difference and that many teachers did not take the time to understand their students and seek ways of supporting them to be successful in academic and social ways. It is evident that these negative experiences marginalized the girls (Urbanawicz et al., 2019) and

emphasized their differences. These tales help educators understand how identity development is intersectional (Crenshaw, 1991) and influenced by interactions with others, social belonging, group membership, achievements, and teacher feedback. The women described their social creativity and strategies to socially connect with their peers. It was evident that they used their own strategies to manage stigma, enhance their own self-esteem by creating counter-narratives, and revalue stigmatized traits (Dirth and Branscombe, 2018). Over time, the girls gained a better understanding of themselves which was important to their wellbeing (Cooper et al., 2020). Individual agency influenced the development of positive identity development (Erentaitė et al., 2018).

Some teachers' words and behaviors, sometimes single events, caused significant stress and anxiety. Mental health difficulties were evident in most accounts, consistent with research suggesting that almost three-quarters of autistic adolescents meet criteria for mental health concerns (Hull et al., 2021). Autistic women have particularly high mental health needs combined with unmet needs for support in education and social difficulties (Baldwin and Costly, 2016). This can be attributed to the experience of dealing with stigma and discrimination as well as to the masking of autistic traits to appear normal, particularly for young women at school. These autoethnographic reflections support previous research on masking and mental health in autism (Hull et al., 2021; Schneid and Raz, 2020).

The women also reported how teachers could make a positive impact by showing interest, care, understanding, and support. Relationships with teachers and peers influenced identity across the stages of school, showing that identities are fluid and contextual (Leary and Tangney, 2012). The tales illustrated that masking/masquerading possibly supported a fake sense of belonging, before the realization that it was acceptable to "belong differently". All of the women reported how they knew they were different at school but did not have a diagnosis at the time and acknowledged that this contributed in various ways to their developing identity in the school years. Teachers have an important role to influence individual identity development in positive ways, be sensitive to intersectionality, and also to facilitate positive interactions in social groups at school because the dynamics, talk, and activities in the groups influence identity development and motivate actions and future opportunities. Belonging or "belonging differently" provide positive connections and support relationships and motivation to be successful at school.

We put forth four propositions to address Research Question 2, illustrating how educational experiences in the school years nurture or erode identity.

2. How can collaborative autoethnographic reflections about educational experiences from four autistic women support our understanding of what is needed to improve support for autistic girls at school and nurture positive self-identity?

Proposition 1. It is important to view autism through a neurodiverse lens and embrace/accommodate individual differences instead of employing a medical lens that focuses on efforts to "fix" or "fit in" to the perceived norm or assuming incompetence against standardized perceptions of what is acceptable.

Regardless of whether girls have a diagnosis of autism, teachers play a key role in being sensitive to intersectionality in identity, and modeling acceptance of student diversity. This requires acknowledgment, without judgment, that everyone has strengths and areas that challenge them. "Without judgment" implies genuine acceptance of difference and may necessitate teachers to confront not only the inequities present within their school setting but those within their own personal values. These ethnographic tales clearly describe, at times shockingly, why teachers must check not only what they say, but the nuance of how they say it and to whom they say it. While looking over this manuscript, Rachel suggested that teachers ask themselves, "Is this student's difference hurting anyone?" If it isn't, don't try to change it".

Proposition 2. Getting to know and consulting with students will facilitate personalized support.

Teachers who know and value their students are in a unique position to encourage positive relationships and connection with the curriculum. The key could be to build structures within schools that provide diverse students with opportunities for membership to a group. While stigma, discrimination, and masking are associated with poor mental health for autistic individuals (Botha and Frost, 2020), a sense of belonging and acceptance by peers can support the development of positive identity because identities are formed and maintained by reflecting on the world and our place in it (Chaffee, 2020). Listening to students in student-teacher consultations will inform the implementation of tailored and personalized supports. This approach would nurture educators' understanding of how to support autistic girls to identify and understand their own personal needs.

Proposition 3. Taking a strengths-based approach by promoting autistic students' strengths and interests.

The ethnographic tales highlighted the power of teachers' responses to the girls' strengths and interests. Celebrating and encouraging "passions" as opposed to lamenting or ridiculing "obsessions" leads to very different outcomes. Accepting and encouraging learning through interests can promote continual learning achievement for autistic girls. Furthermore, autistic individuals view special interests as an attribute of their autistic identity (Cooper et al., 2020). Failing to acknowledge, or at the very worst ridiculing, the intense or special interests of an autistic student is detrimental to their identity formation.

Proposition 4. Teaching autistic students to be proud of who they are and to be assertive to find their place in the world.

As discussed in the Introduction, awareness of how autistic girls present is recent, in part because they effectively "mask" their authentic selves in their efforts to fit into existing group structures. In fact, girls are more committed to their identity earlier in their adolescence than their male counterparts (Klimstra et al., 2010). One consequence of this is that girls are more strongly wedded to their group status and existing social connections (Geary et al., 2003) than boys (Klimstra et al., 2010). Girls who are members of an "in" group are more likely to stigmatize and discriminate toward girls who are not within their group (Ehala, 2018). For autistic girls who are trying to fit in, the challenge they face is therefore enormous. In their reflections and throughout the discussion, the autistic women spoke of the downsides of masking, such as suppressing one's own feelings and needs, negative impacts on mental health, not making true connections with people, increased feelings of isolation, and the fact that masking is difficult and tiring. It is clear that teachers have an important role in supporting students to be proud of their autistic identity.

Conclusion

Previous research has demonstrated that school experiences vary greatly depending on teachers' knowledge of and attitudes toward difference/autism (Liasidou and Ioannidou, 2021). The experiences of autistic females are under-researched in comparison to their male counterparts. Accounts of female experiences are therefore important to help counterbalance the historic dominant male bias in research (Milner et al., 2019). The opportunity for collaboration within the collective of women drawing on an autoethnographic approach enabled us to reflect and interpret in social and cultural ways, highlighting multi-subjectivity associated with collaboration (Chang et al., 2016). According to Chang (2013), successful social interaction requires an understanding of both self and others. Collaborative online meetings encouraged active agency in the process of explaining and justifying directions and research objectives; they also helped to build trust between the co-authors and supported effective teamwork, communication, and positive relationships in ways that made the process both productive and enjoyable. The data and individual and collective interpretations can inform not only increased awareness and understanding of autistic females, but help educators develop more inclusive ways to support this group of learners in school contexts that promote a positive sense of identity. It is important therefore that practices that support inclusion, minimize marginalization, acknowledge intersectionality, and support a positive sense of identity are nurtured for autistic learners and autistic girls specifically.

Appendix

Final guiding questions for Stage 1:

- Did teachers support you or not support you in learning? If so, how?
 - How did you feel about support or lack of support? What impact did this have on you?
 - How did you feel about yourself as a learner?
- How did teachers support you to develop friends?
 - Describe how your experiences influenced your sense of self and your sense of belonging in the social group.
 - What else helped or would have helped you to make friends?
- How did your peers support you or not support you in learning?
 - How did your relationships with peers influence how you thought about yourself as a learner?
 - How did your peers support your own sense of yourself and your sense of belonging at school?
- What else influenced your sense of self?

References

- Abbasi, N., 2016. Adolescent identity formation and the school environment. In: Fisher, K. (Ed.), *The Translational Design of Schools: An Evidence-Based Approach to Aligning Pedagogy and Learning Environments*. Sense Publishers, Rotterdam, pp. 83–103.
- Asperger, H., 1944. Die "Autistischen Psychopathen" im Kindesalter. *Arch. Psychiatr. Nervenkrankh.* 117 (1), 76–136.
- Baldwin, S., Costly, D., 2016. The experiences and needs of female adults with high-functioning autism spectrum disorder. *Autism* 20 (4), 483–495. <https://doi.org/10.1177/1362361315590805>.
- Bajada, G., Callus, A.M., Borg, K., 2021. Unpretentious education: a Foucaultian study of inclusive education in Malta. *Disabil. Soc.* <https://doi.org/10.1080/09687599.2021.1874877>.
- Bargiela, S., Steward, R., Mandy, W., 2016. The experiences of late-diagnosed women with autism spectrum conditions: an investigation of the female autism phenotype. *J. Autism Dev. Disord.* 46 (10), 3281–3294.
- Bartlett, T., Carrington, S., 2021. Autistic voices in autism education research. In: Carrington, S., Slegers, B., Harper-Hill, K., Whelan, M. (Eds.), *Research Approaches to Supporting Students on the Autism Spectrum in Inclusive Schools: Outcomes, Challenges and Impact*. Routledge, Oxon, pp. 50–65.
- Botha, M., Frost, D.M., 2020. Extending the minority stress model to understand mental health problems experienced by the autistic population. *Soc. Ment. Health* 10 (1), 20–34.
- Braun, V., Clarke, V., 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual. Res. Psychol.* 18 (3), 328–352.
- Cage, E., Troxell-Whitman, Z., 2019. Understanding the reasons, contexts and costs of camouflaging for autistic adults. *J. Autism Dev. Disord.* 49 (5), 1899–1911.
- Carrington, S., Graham, L., 2001. Perceptions of school by two teenage boys with Asperger syndrome and their mothers: a qualitative study. *Autism* 5 (1), 37–48.
- Chaffee, T., 2020. Normal Life Inc [website]. <https://normallifeinc.com/>.
- Chang, H., 2013. Individual and collaborative autoethnography as method. In: Holman Jones, S., Adams, T.E., Ellis, C. (Eds.), *Handbook of Autoethnography*. Routledge, Abingdon, pp. 107–122.
- Chang, H., Wambura Ngunjiri, F., Hernandez, K.C., 2016. *Collaborative Autoethnography*. Routledge, New York.
- Cooper, R., Cooper, K., Russell, A.J., Smith, L.G.E., 2020. "I'm proud to be a little bit different": the effects of autistic individuals' perceptions of autism and autism social identity on their collective self-esteem. *J. Autism Dev. Disord.* 51 (2), 704–714.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43, 1241–1299.
- Cresswell, L., Cage, E., 2019. "Who Am I?": an exploratory study of the relationships between identity, acculturation and mental health in autistic adolescents. *J. Autism Dev. Disord.* 49 (7), 2901–2912.
- Dirth, T., Branscombe, N.R., 2018. The social identity approach to disability: bridging disability studies and psychological science. *Psychol. Bull.* 144 (12), 1300–1324.
- Ehala, M., 2018. Social media, culture, and identity construction: from mass communication to communication of the masses. In: Harwood, J., Gasior, J., Pierson, H.D., Nussbaum, J.F., Gallois, C. (Eds.), *Language, Communication, and Intergroup Relations*. Routledge, New York, pp. 54–55.
- Erentaitė, R., Vosylis, R., Gabrielavičiūtė, I., Raizienė, S., 2018. How does school experience relate to adolescent identity formation over time? Cross-lagged associations between school engagement, school burnout and identity processing styles. *J. Youth Adolesc.* 47 (4), 760–774.
- Erikson, E.H., 1968. *Identity, Youth and Crisis*. W. W. Norton, New York.

- Geary, D.C., Byrd-Craven, J., Hoard, M.K., Vigil, J., Numtee, C., 2003. Evolution and development of boys' social behavior. *Dev. Rev.* 23, 444–470.
- Hull, L., Levy, L., Lai, M.C., Petrides, K.V., Baron-Cohen, S., Allison, C., Smith, P., Mandy, W., 2021. Is social camouflaging associated with anxiety and depression in autistic adults? *Mol. Autism*. 12 (1). <https://doi.org/10.1186/s13229-021-00421-1>.
- Hull, L., Petrides, K.V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.C., Mandy, W., 2017. "Putting on my best normal": social camouflaging in adults with autism spectrum conditions. *J. Autism Dev. Disord.* 47 (8), 2519–2534.
- Kanner, L., 1943. Autistic disturbances of affective contact. *Nerv. Child* 2 (3), 217–250.
- Kerpelman, J.L., Pittman, J.F., 2018. Erikson and the relational context of identity: strengthening connections with attachment theory. *Int. J. Theory Res.* 18 (4), 306–314.
- Klimstra, T.A., Hale, W.W., Raaijmakers, Q.A.W., Branje, S.J.T., Meeus, W.H.J., 2010. Identity formation in adolescence: change or stability? *J. Youth Adolesc.* 39 (2), 150–162.
- Lawson, W., Carrington, S., 2021. Inclusive research practice: engaging autistic individuals and their families. In: Carrington, S., Siggers, B., Harper-Hill, K., Whelan, M. (Eds.), *Research Approaches to Supporting Students on the Autism Spectrum in Inclusive Schools: Outcomes, Challenges and Impact*. Routledge, Oxon, pp. 17–31.
- Leary, M.R., Tangney, J.P. (Eds.), 2012. *Handbook of Self and Identity*, second ed. The Guilford Press, New York.
- Liasidou, A., Ioannidou, E., 2021. Disability-related bullying and its discursive formations and enactments in the social ecology of schooling. *Discourse: Stud. Cult. Polit. Educ.* 42 (4), 499–512.
- Lopes, B., 2018. The Identity Element: A Definition of Identity. <https://www.linkedin.com/pulse/identity-element-definition-bianca-lopes>.
- Milner, V., McIntosh, H., Colvert, E., Happé, F.A., 2019. Qualitative exploration of the female experience of autism spectrum disorder (ASD). *J. Autism Dev. Disord.* 49 (6), 2389–2402.
- Pellicano, E., den Houting, J., du Plooy, L., Lilley, R., 2019. Knowing autism: the place of experiential expertise. *Behav. Brain Sci.* 42. <https://doi.org/10.1017/S0140525X18002376>.
- Shmulsky, S., Gobbo, K., 2019. Autism support in a community college setting: ideas from intersectionality. *Community Coll. J.* 43 (9), 648–652. <https://doi.org/10.1080/10668926.2018.1522278>.
- Schneid, I., Raz, A.E., 2020. The mask of autism: social camouflaging and impression management as coping/normalization from the perspectives of autistic adults. *Soc. Sci. Med.* 248. <https://doi.org/10.1016/j.socscimed.2020.112826>.
- Shakes, P., Cashin, A., 2019. An analysis of Twitter discourse regarding identifying language for people on the autism spectrum. *Issues Ment. Health Nurs.* 41 (3), 221–228.
- Shume, T.J., 2020. Conceptualising disability: a critical discourse analysis of a teacher education textbook. *Int. J. Incl. Educ.* <https://doi.org/10.1080/13603116.2020.1839796>.
- Urbanawicz, A., Nicolaidis, C., den Houting, J., Shore, S.M., Gaudion, K., Girdler, S., Savarese, R.J., 2019. An expert discussion on strengths-based approaches in autism. *Autism Adulthood* 1 (2), 82–89.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, London.
- Wing, L., 1981. Asperger's syndrome: a clinical account. *Psychol. Med.* 11 (1), 115–129.

Looking for something outside of ourselves: the politics of researching race and dis/Ability and (special) education

Deborah J. Gallagher, David Hernández-Saca^a, and Whitney Hanley^b, ^a University of Northern Iowa, Cedar Falls, IA, United States; and ^b Trickum Middle School, Gwinnett County Public Schools, Lilburn, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

From Dunn to Morgan and colleagues: a brief history of disproportionality	498
Responses to Morgan and colleagues	500
Overrepresentation and the social and intersectional disablism model of dis/Ability	501
Conclusion	504
References	504

As we move into the third decade of the 21st century, it seems increasingly hard for the thoughtfully observant among us to sustain a healthy optimism in humanity's capacity for resolving its most agonizing injustices. We write this chapter in the aftermath of George Floyd's murder under the knee of a malevolently racist police officer and the police shooting death of Breonna Taylor as she slept in her bed. We trust the reader will not only forgive but embrace this somber opening because we wish to make the case that the politics of what happened to these two Americans (among so many others) is intimately tied to the politics of schooling at the intersections of race and dis/Ability¹—most specifically to the construction of the educational debt (Ladson-Billings, 2006) and the pathologization and overrepresentation of Black, Indigenous and People of Color (BIPOC) students in special education programs, leading in turn to the ongoing educational segregation/exclusion and coinciding with hyper-disciplining of these students that is now recognized as integral to the school-to-prison pipeline (Annamma, 2015; Blake et al., 2011; Kim et al., 2010; Skiba et al., 2002; USDOE, 2016).

However, what happened to George Floyd and Breonna Taylor is not only a species of hegemony within the United States, but a stark reminder that the evil of racism is alive and well internationally, since the root concept of race began during the "Enlightenment" period in Europe (Suárez-Krabbe, 2014). Throughout the Global North, scholarly analyses of the intersection of race and dis/Ability have illuminated the issue of disproportionality and expanded it to include other groups such as migrant and immigrant students (Artiles et al., 2011a; Erevelles, 2011; Migliarini, 2017). These analyses also reveal the manner in which the illusive concept of "inclusion" within the US and internationally has been co-opted to truly mean inclusion into hegemony (Artiles et al., 2011b). McDermott et al. (2011) have aptly identified this co-optation as "global norming," which has served to perpetuate stratification, labeling and management of "deviant" bodyminds into predefined hegemonic ideologies of normal. Accordingly, these ideologies conceal dominant discourses within 21st century institutions that began during colonization of the Global South by the Global North as part of the sociopolitical state apparatus (Althusser, 2014). Recently, the field of Disability Studies in Education (DSE), and in particular Disability Studies and Critical Race Studies (Annamma et al., 2013), has taken a turn toward global conversations (Connor et al., 2021). For example, Connor et al. (2021) posed the following question:

... how can we jointly create more shared and dialogic spaces between the Global North and South, concepts that by their very existence reify hierarchies? And how can DisCrit can be placed in service of addressing existing inequities without enacting an imperialist gaze? (p. 604).

In turn, this question raises another—what is happening within the US surrounding not only the disproportionality issue but also extending to the very humanity of young Black, Indigenous and People of Color with and without dis/Abilities, both inside and outside of educational contexts and within the global cultural-historical context (Artiles, 2009, 2019; Iqtadar et al., 2021).

It is with this global cultural-historical context that we write from our positionality and relationalities (Evans-Winters and Espósito, 2019). The first author identifies as a cisgender white female. The second author identifies as a Dis/Abled Latinx of El Salvadorean and Palestinian heritage gay cisgender male who was labeled with an Auditory Learning Disability and experienced special education programming and trauma (Hernández-Saca, 2021). The third author identifies as a Black queer cisgender female.

The purpose of this chapter is to demonstrate, through our critical analysis of the debate surrounding the overrepresentation of BIPOC students in special education, that researching dis/Ability at the nexus of race and other markers of difference is inevitably a moral, spiritual and therefore political act (Hernández-Saca and Cannon, 2019; Williams, 1987). Accordingly, we assert that the most important and meaningful questions pertaining to dis/Ability and education can never be resolved by invoking

¹By dis-dash-Ability (dis/Ability), we mean to underscore the political and sociocultural construction of both disability and ability at the intersections of power and identities. We capitalize the "A," in dis/Ability to resist the deficit thinking and language (Patton Davis and Museus, 2019) associated from the medical-psychological model of disability in history and society associated with people with disabilities at their intersections of power and identities.

methodological neutrality to settle what are essentially moral, spiritual and political problems. In the process, we make the case that Paul Morgan et al.'s. (2015) research announcing that overrepresentation is “not” a problem—and that BIPOC students are not in fact overrepresented but instead underrepresented in special education programs—is not new, but is instead an elaborated iteration of timeworn, eugenically tinged methodological commitments (see: Stubblefield, 2007).

In what follows, we first provide a brief history of the debate surrounding overrepresentation. This historical overview includes Morgan et al. (2015) research asserting that the problem of overrepresentation is more imagined than real. Next, we outline the outpouring of scholarly responses to their work. We then extend and elaborate on these responses to make the case that until *we stop looking for something outside of ourselves* through the use of objectivist research methods to adjudicate knowledge claims about how we should and should not educate children, we will inevitably adhere to policies and arrangements that impose racist-ableist-sexist-capitalist competition on schoolchildren. More to the point, it is not only possible but essential that educators deliberately engage in a moral and spiritual re-imagining of what we want education to be for and about. In the absence of such a re-imagining, we will continue to have the same unresolved debates, spinning in the same amnesic circles one generation after another. Finally, we contend that this position has distinctly praxical implications pertinent not only to the overrepresentation problem, but also to the goal of improving the educational experience and promise for ALL students, but specifically BIPOC students with and without dis/Abilities.

From Dunn to Morgan and colleagues: a brief history of disproportionality

Even before special education programs were federally mandated in American public schools, Lloyd Dunn, a prominent scholar in the field who was nearing retirement, surveyed the evolving landscape and noted with deep apprehension that special education classrooms contained proportionally more Black, Indigenous and People of Color (BIPOC) students than general education classrooms—quite noticeably more, in fact. His 1968 “swan song” article published in the journal, *Exceptional Children*, posed a seemingly straightforward question—is the expansion of these classrooms educationally justifiable? His own answer was an emphatic, “No!” Not only did Dunn see them as stigmatizing, segregating and restricting BIPOC students, he also asserted that in his view:

... much of our past and present practices are morally, and educationally wrong. We have been living at the mercy of general educators who have referred their problem children to us. And we have been generally ill prepared and ineffective in educating these children. *Let us stop being pressured into continuing and expanding a special education program that we know now to be undesirable for many of the children we are dedicated to serve* (p. 5, emphasis added).

Dunn's critique was followed by Deno's (1970) prescient essay warning that the burgeoning field of special education's bent toward establishing and expanding dis/Ability identification categories would result in more harm than good for students. Deno (1970) stated: “... the medical model criteria on which these classifications are based ... too often have limited or undetermined relevance to the educational decisions that need to be made” (p. 332). Deno (1970) went on to explain that, “... the emphasis on ‘defect’ residing in the child tends to focus attention away from the external variables which educators might be in a position to do something about” (p. 232). The emphasis, Deno believed, should be on remediating the general education classroom and schools rather than pathologizing and segregating students—particularly those marginalized at the nexus of racism and poverty. The field of special education, Deno also believed, should be intent on “... work[ing] itself out of business as a social institution” (p. 333). In the best of all worlds, by sharing its instructional breakthroughs with general education, special education as a separate system would essentially make itself redundant. Three years later, Mercer (1973) directly confronted the overrepresentation of BIPOC students in special education, finding that Hispanic and Black students were significantly overrepresented in the derogatory special education category of what was then called, “Mild Mental Retardation (MMR),” today called Intellectual Disabilities (ID).

After the Education for All Handicapped Children (Public Law 94–142) was signed into federal law, dis/Ability categories expanded apace. And, as Artiles and Trent (1994) noted, the expansion of dis/Ability categories over the 1980s extended overrepresentation into the newer ones such as Specific Learning Disabilities (SLD) (See: Ong-Dean, 2006). It is more than passingly interesting, though, that this particular category initially overrepresented white students because it originated as a safe harbor for struggling middle class white students from the stigma of the “MMR” programs already disproportionately representing BIPOC students (Sleeter, 1986). That situation shifted precipitously in the aftermath of class action lawsuits ruling in favor of the plaintiffs' arguments that IQ test bias resulted in the overrepresentation of BIPOC students in MMR programs (see: Larry P. v. Riles, 1972, 1974, 1979, 1984, 1986). Consequently, as fewer BIPOC students were identified as “MMR,” the SLD category became the new default placement for “under-achieving” students. Because the SLD category was primarily created to preserve white privilege and educational advantage for affluent whites, the growing placement of BIPOC students into the category was soon seen by whites as tarnishing the redemptive *mystique* of white SLD students fetishized as intellectually “exotic” rather than “disabled” in the *usual* sense of the word (Carrier, 1986; Sleeter, 1986).

In short, many if not most of the contemporary critiques of special education had been articulated by pioneers in the field even before special education had become the law of the land. These critics were acutely aware of the stigma of dis/Ability labels. They knew that dis/Ability labels were conceptually and practically flawed—in addition to being less than useless in informing instructional and placement decisions. It was clear to them that something was very wrong with the disproportionate number of BIPOC

students in classrooms for students identified by these stigmatizing labels. Early on, it was also increasingly evident that segregated special education placements were not superior to general education placements. Quite the contrary, in fact.

As it turned out, of course, these warnings and critiques went largely unheeded as other scholars and researchers pushed back against these incipient threats to the status of the field. Reschly (1984), for one, expressed incredulity and dismay toward what he considered “sweeping allegations” of institutional racism and test bias raised by those who pointed to overrepresentation as a problem. Troubled by what he viewed as a precipitous rush to misjudgment, Reschly applauded the *National Academy of Sciences* report (Heller et al., 1982) for asking what he considered to be the important question—“Why is disproportion a problem?” (p. 15). From Reschly’s perspective, the crucial issue had to do with the lack of specificity in the dis/Ability classification system. State variations in terminology and criteria resulted in changes in students’ classifications merely by crossing state boundaries. Yet, doubling down on attempts to refine classification procedures and criteria held little promise for resolving the problem of overrepresentation. Reschly instead favored the idea of generic renaming and combining of dis/Ability categories, an ingenious move that would doubtless do more to elide the issue than confront it head on.

As overrepresentation gained increasing attention among scholars and the public, defenders became more concerted in their efforts to minimize its challenge to special education’s standing. MacMillan and Reschly’s (1998) attempt at this goal proposed that a more *cautious* way to view and understand overrepresentation was to calculate, out of the entire population of BIPOC students, what percentage were identified as having a particular dis/Ability. This alternative statistical rendition attempted to swap the most common understanding that considers overrepresentation as the proportionally larger special education enrollment of a historically multiply marginalized group as compared to their proportion of the school population (see: *Office of Civil Rights*). For the sake of illustration, Mercer (1973) found that while Hispanic students ages 6–15 constituted 7% of the Riverside, California school population, these students comprised 12% of the students in special education programs for students who were, at the time, labeled as having “MMR” or “educable mentally retarded” (“EMR”). More strikingly, this study also established that while African American students made up 9.5% of the population, they represented a whopping 32% of the students in these programs. In short, the calculation compared the percentages of students from historically multiply marginalized groups in the general school population to their percentage representations in various special education programs.

With MacMillan and Reschly’s (1998) replacement of one algorithm with another, the *Office of Civil Rights*’ troubling statistic revealing 34.64% of the MMR category to consist of Black students (% category by group) became the more palliating statistic of 2.10% of all Black students in the “MMR” category (% of group in category) (MacMillan and Reschly, 1998, p. 18). *Put in plainer terms, they proposed that we should not view overrepresentation as Black students embodying one out of every three students in the “MMR” classroom/program. Rather, they wanted readers to think of overrepresentation as asking what percentage of all Black students are labeled as “having” “MMR.”* The answer is only a tiny bit over 2%—a seemingly much more assuasive number. Or is it? For white students, the percentage of the group in the “MMR” category was 0.81. *This still means that roughly two and a half times more Black students were labeled “MMR” compared to white students.* In actual numbers, out of a hundred thousand students, over 2000 Black students would be labeled as “MMR,” but only about 800 white students. This disproportionate representation is hardly the tempest in a teapot these researchers wanted to portray it as being.

MacMillan and Reschly (1998) also argued that overrepresentation statistics were confounded first by the difficulties of defining race (or what they called “ethnicity”) as a quantifiable variable and second by their assertion that “ethnicity” is actually a proxy for socioeconomic status (SES). “When ethnicity is the only independent variable,” they quipped, “interpretations tend to emphasize ‘the figment of the pigment’” (p. 20). In our upcoming discussion of Morgan et al.’s. (2015) study, we will revisit this disingenuous attempt to invoke the cultural construction of racial categories and the intersectionality of race and class as vehicles for effacing the overrepresentation of BIPOC students in special education. Still, at this point in time, these arguments did little to appease mounting censure surrounding the status quo.

At the dawn of the 21st century, a study by Harvard’s Civil Rights Project (Losen and Orfield, 2002) again substantiated the fact that a disproportionately high number of BIPOC students were placed in special education. In their study, African American children were almost three times (2.88) as likely as white children to be identified as having an “intellectual disability,” 1.9 times as likely to be identified as “emotionally disturbed,” and nearly 1.3 times as likely to be identified as having a “Specific Learning Disability” (SLD). Native American children were 1.3 times as likely as white children to be identified as having an “intellectual disability,” 1.24 times as likely to be identified as “emotionally disturbed,” and 1.50 times as likely to be identified as having a learning disability. Latino children are 1.17 times as likely to be identified as having a “SLD.” Not surprisingly, BIPOC students are significantly more likely than white students to be placed in unequal and segregated special education classes for more of the school day. The extent of this phenomenon led researchers to conclude that these dis/Ability identification and placement patterns have systematically expedited the racial re-segregation of American schools (Artiles and Trent, 1994; Blanchett, 2009; Ferri and Connor, 2005; Fierros and Conroy, 2002; Losen and Orfield, 2002). Now, as we enter the third decade of the 21st century, well over four decades of research, scholarship and public policy has attempted in vain to address the problem (see: Artiles, 2019; Artiles et al., 2010; Cooc and Kiru, 2018; Cruz and Rodl, 2018; Skiba et al., 2016b). More recently we have been presented by Morgan et al.’s. (2015) research asserting that it is a *nonexistent problem after all*.

Morgan et al.’s. (2015) study made a smashing debut. Not only did it receive an advance press release before it was presented at the American Educational Research Association (AERA) conference, but it also attracted attention from the popular press—a relatively rare event for educational research (Morgan and Farkas, 2015; Phippen, 2015). And no wonder. Research announcing to the public that its educational system was not as “racist” as it might have been feared was good news, indeed. What a great relief this proclamation was for so many white Americans—“racism without racists” (Zuberi and Bonilla-Silva, 2008).

In their op-ed in the *New York Times*, Morgan and Farkas explained that, “The real problem is that black children are underrepresented in special education classes when compared with white children with *similar* levels of academic achievement, behavior and family economic resources” (Morgan and Farkas, 2015, emphasis added, n.p.). What they sought to do in their research, in other words, was to demonstrate that overrepresentation of BIPOC students in special education was caused by more dire “background conditions” affecting these students, not discrimination. When they applied statistical controls for said conditions, it appeared that these students were less likely rather than more likely to be placed in special education programs (See Gordon, 2017, 2018). Assuming the mantle of seemingly “liberal champions” of dis/abled BIPOC students, they asserted that the previous decades of research that informed policies to prevent overrepresentation were egregiously misinformed and therefore discriminatory in that they denied needed resources and services on the basis of race, ethnicity, and language. For most of the public and many professionals, this was compelling news that seemed to render five decades of critique entirely misguided.

Responses to Morgan and colleagues

It is something of an understatement to say that many scholars with longstanding engagement with this issue, and those with a very different set of ideas about the politics of research on dis/Ability, were distinctly unimpressed with Morgan et al.’s (2015) work. Their scholarly responses ranged from critiques on the technical merits of their research to commentaries on their conceptual flaws.

Skiba et al. (2016a) critique essentially centered on what they viewed as substantial shortcomings of a technical nature in Morgan et al.’s study. The upshot of their criticisms highlighted Morgan and his colleagues’ audacious declaration that not only was decades of research about overrepresentation wrong but entirely backward. Among other concerns, Skiba et al. (2016a) noted:

- Their limited database consisting of unverified self-reported teacher estimates that, had the authors been so inclined, could have been validated against a rigorous federal data base.
- The fact that studies conducted by prominent agencies (e.g., U.S. Department of Education and the National Research Council) using federal data sets during the same period from which Morgan et al. drew their data, “... found significant overrepresentation of African American students in the disability categories of ID [intellectual disabilities], ED [emotional disabilities], and LD [learning disabilities].”
- The researchers’ confounding combination of racial categories (Native Americans, Asian Americans, and Pacific Islanders), despite significant cultural differences, making it impossible to demonstrate what national studies had previously found to be the case—the significant overrepresentation of Native Americans in special education.
- Their “simplistic” claim that overrepresentation is primarily if not entirely a function of SES despite extensive literature demonstrating the very culturally loaded judgments about criteria (social-behavioral and adaptive behavior) that form part or most of the bases for the dis/Ability categories of ED and ID. Additionally, they cited research studies that, “... African American students were most overrepresented in the category of ED in *high SES* districts” (p. 222, emphasis in original).
- Their use of indirect methods of counting special education enrollment that has been shown by previous research to, “... find that the introduction of SES as an independent variable reduces findings of overrepresentation to non-significance (or have failed to find overrepresentation entirely,” while studies relying on direct counts found that race was the significant variable regardless of SES (p. 223).

Crucially, Skiba et al. (2016a) also demonstrated that Morgan et al.’s conclusions both *misrepresented and even contradicted their own data*. This surreptitiousness is something that we (the present authors) found to be present in earlier statistical attempts seeking to minimize or erase the overrepresentation of BIPOC students in special education (see: Hosp and Reschly, 2004).

As an interesting aside, Hosp and Reschly (2004) made the case that *academic achievement* is a variable subject to more control by special educators and should be researched rather than economic and demographic variables over which educators have no control. Oddly, these researchers contradicted their central premise at the conclusion of their article by confessing that their data did not after all support their hypothesis that academic achievement accounts for overrepresentation. They also went a step further, adding that it is in any event impossible to pin down and statistically disentangle the relevant variables that might account for this phenomenon. One might expect that this sagacious admission would have finished off future efforts in this direction, but that has clearly not been the case. Like those preceding it, Morgan et al.’s statistical analysis thus brings to mind Andrew Lang’s familiar observation that, “Most people use statistics like a drunk man uses a lamppost; more for support than illumination” (n.d, n.p.).

Conceptual critiques centered on a whole range of erroneous assumptions in Morgan et al.’s study (see: Collins et al., 2016). First, Morgan and his colleagues assumed that receiving special education services is indisputably advantageous. Citing the most authoritative research on long-term outcomes for students exiting special education programs (Newman et al., 2009), Collins et al. (2016) refuted this assumption, stating that these students, “... continue to experience lower graduation rates, poorer post-secondary outcomes, and higher rates of unemployment and underemployment ...” (p. 5). They also pointed out that three-quarters of students labeled as having emotional/behavioral disorders end up entangled in the criminal justice system at some point in their lives—a finding that has highly meaningful implications for overrepresented BIPOC students given the enormous disproportionality in the US prison system of mass incarceration (see: NAACP). Disturbingly, Losen and Gillespie (2012) reported that, “... one out of every four (25%) Black children with disabilities enrolled in grades K-12 was suspended at least once in 2009–2010,” and that, “[s]tudents with disabilities and Black students were also more likely to be suspended repeatedly in a given year than to be suspended just once. The reverse was true for students without disabilities and for most other racial/ethnic groups” (see also: Rausch

and Skiba, 2006). Further, BIPOC students in special education are excluded from the general education classroom in significantly larger proportions (Bal et al., 2019; Fierros and Conroy, 2002; Losen, 2015; Skiba et al., 2016b).

Second, Collins et al. (2016) contested Morgan and his colleagues's *within the person* account of dis/Ability—otherwise known as the deficit view of dis/Ability. By locating, "... the issue of overrepresentation either within the bodies, behaviors, or brains of students, they failed to unpack why it is that only the most subjective disability categories are more likely to be attributed to [BIPOC students]" (p. 6). Accordingly, their objectivist medical modelist view of dis/Ability absolves them from moral and spiritual introspection about the role of racist stereotypes, policies and practices in dis/Ability identification (see: Baynton, 2001; Kliever et al., 2015). This deficit view further occludes responsibility for enacting this privilege.

Third, Morgan et al. (2015) assumed that resolving disproportionality is a simple matter of accurate dis/Ability identification—an assumption that is itself premised on their objectivist epistemological stance. It is telling that they did not concern themselves with the politics of defending this position—even when they acknowledged the "judgmental" disability categories as the ones in which overrepresentation has become a problem. The point here is that their medical modelist view of dis/Ability and its epistemological underpinnings lead them to overlook the implications of their own (apparently murky) recognition that disability is a social/cultural construct and one that, in a racist culture, is thoroughly and unavoidably imbued with racist stereotypes. Given their own admission of the judgmental nature of the disability categories, their baseless stance that "accurate" disability identification is possible provides an incoherent footing for their assertion that increased identification of BIPOC students makes sense as an equity measure and remedy: *nothing could be further from the truth*.

From the outset, we contended that researching dis/Ability is inevitably a moral, spiritual and therefore political act. It cannot be otherwise because dis/Ability at the intersections of power and identities is a moral, spiritual and political construct (See: Hernández-Saca and Cannon, 2019). We also asserted that important and meaningful questions pertaining to dis/Ability and education cannot be resolved by invoking methodological neutrality to settle what are essentially moral, spiritual and political problems. By extension, methodological neutrality cannot resolve the problem of overrepresentation because it created the problem in first place. In what follows, we demonstrate how the epistemological and ontological repositioning toward the social and psycho-emotional model of dis/Ability (Thomas, 1999) or intersectional disablism (Iqtadar et al., 2020) not only alters how we understand dis/Ability as an intersectional phenomenon, but also obviates the problem of overrepresentation.

Overrepresentation and the social and intersectional disablism model of dis/Ability

What precisely does it mean when we say someone has a "disability"? When and how do certain human differences come to be known as "disabilities"? In everyday parlance, having a disability means that an individual is unable to perform physically, cognitively, or socially/emotionally in a manner considered "typical" or "desirable." Yet, words like *considered*, *typical*, and *desirable* are distinctly values-laden interpretations that are, even if unintentionally, laced with idiosyncratic judgments. This means that identifying some people as having "disabilities" signifies their failure to meet established sociocultural and political expectations—even if identifying them as such is intended as a benevolent gesture of aiding and assisting—something that special education has historically done toward BIPOC youth and communities (Coomer, 2019; Thorius, 2019).

Over the years, this informal, common sense, understanding of dis/Ability has been incrementally formalized by professionals who through a process of applying ever more exacting description and quantification have prescribed increasingly specific criteria for defining a whole range of disabilities. This process of formalization took place within the purview of the fields of medicine and psychology, and the emergent understanding of dis/Ability has come to be known as the *medical model* of dis/Ability (Trent, 1994). Viewed through the paradigmatic lens of the medical framework, "disability" is:

- An objective condition (categorizing a person as having a disability is an entirely impartial act).
- Inherent in the individual. The person "owns" the disability.
- A deficiency or form of pathology—thus the remedy for disability is to cure, remediate, rehabilitate (i.e., make disabled people more "normal").
- Note: the medical model, in assuming objectivity, does not view the category of normal, that is, the ideology of normal, as a socio-political category.

By all appearances, this formalization process removed the role of judgment, values, or opinion—presumably reclaiming the identification of disabilities as a politically neutral act and disability itself as an "objective" condition. Even when the role of judgment is acknowledged, it is termed *clinical judgment*, such that the identification process is nonetheless thought to be sufficiently objective enough that most of us tend to think of a disability as an autonomous condition that an individual possesses or "has"—an inherent facet of them (Biklen, 1988). This assumption is apparent in our language when we say, for example, that Kevin *has* a learning disability (or more appropriately, given this approach, that Kevin *is* learning disabled). Or that Keisha *has* a behavioral disorder (or more to the point, given this approach, that Keisha *is* behaviorally disordered). In other words, the assumption that a student's disability is solely a function of their intrinsic characteristics seems self-evidently "true" and generally taken-for-granted.

Still, disturbing suspicions about the integrity of this common sense *within-the-person* assumption confront us in the form of incongruities for which some find themselves quite at a loss to explain. One of the most troubling of these incongruities is the changing and arbitrary nature of disability definitions and identification criteria (Danforth, 2007; Dunn, 1968; Ferri and Connor, 2006; Gallagher, 2010; Kliever and Biklen, 1996). A classic example dates back to 1973 when what is now the *American Association*

on *Intellectual and Developmental Disabilities* (AAIDD) redefined “mental retardation” (now termed intellectual disabilities) by revising the criteria for determining “significant sub-average intellectual functioning” from one standard deviation below the mean (an IQ score of 85 or below) to two standard deviations below the mean (an IQ score of 70 or below). This redefinition precipitously lowered the population of individuals identified as “mentally retarded” from 16% of the overall population to 2.25%, a shift dramatic enough for some scholars to note satirically that it resulted in thousands of students labeled as having intellectual disabilities children magically becoming “perfectly normal.” It is also noteworthy that since its original formalization in 1941, the definition of intellectual disabilities underwent six rather substantive alterations prior to the late 1990s (Smith, 1997).

A very similar situation has surrounded the category of Specific Learning Disabilities (SLD), a special education category that has been with us since Samuel Kirk first coined the term in 1963 (Sleeter, 1986). Its theoretical origins, however, go back several decades earlier when researchers sought to explain the learning problems of students whose low academic achievement could not be explained by the usual factors—low intelligence, cultural diversity, physical/sensory impairment, or emotional complication (Carrier, 1986; Danforth and Ressa, 2013). Crucially, these researchers started from the premise that the learning problems they sought to explain originated in the neurological makeup of individual students—even though the basis of this assumption was simple conjecture based on tautological reasoning. Incidentally, this terrific leap of inference was pointed out *and acknowledged* by the researchers quite early on, but nonetheless ignored (Carrier, 1986). Almost a century later, we have no decisive evidence that SLD are neurologically based.

That is not to say that neurological factors are not involved in learning. Of course, they are. Nor is it to deny that some students require much more time and experience far greater difficulty in acquiring literacy and other academic skills in comparison with their same age peers (although we would add that this is another Eurocentric and ableist foundation or master narrative regarding what counts as a SLD if we understand learning from a cultural-historical developmental approach) (Hernández-Saca, 2016). But this point does not support the presumption of a specific neurological disorder in the absence of knowing far more than we currently know—if indeed such specificity can ever be pinned down or is even desirable from an educational and cultural standpoint. Further, the neurological explanation is far too unidimensional because the intricacies involving how individual neurology intersects with the equally complex webs of the social/cultural, pedagogical, and emotional and we argue spiritual aspects of learning are far, far more complicated.

Our main point here is that even though the neurological basis of learning disability theory is nothing more than an appealing (to some) conjecture at best, it has become lodged in our collective consciousness as a powerful truism. It is also one that leads us to overlook countervailing evidence that should prompt us to reflect more deeply about this special education category. The fact that the definition and identification criteria for SLD have undergone many significant changes over the last several decades (see: Gallagher, 2010) may just mean that this category is an evolving concept we have about some students, not an *in-their-heads* deficiency that some students *have* (Dudley-Marling, 2004).

Elaborating on the distinction between SLD as a personal deficiency versus a culturally/institutionally created identity, Dudley-Marling asserted that:

... learning and learning problems, including the identity of “having LD,” do not reside in people’s heads as much as in the complex of social interactions performed in a place called school that is itself situated in a broader social, political, and cultural context (p. 483).

Pointing specifically to our cultural ethos of rugged individualism, Dudley-Marling noted that students who require more assistance than others are seen inadequate and problematic. Together with:

- The competitive underpinnings of American schooling,
- Our preoccupation with the *rate* of learning over the quality of learning, and
- School tracking and stratifying practices based on the erroneous assumption of achievement as normally distributed.

An emphasis on individualism makes it inevitable that some students come to be viewed as deficient—*not because they are innately so, but because these values and practices make their differences “a problem.”* Finally, quoting McDermott (1993), Dudley-Marling enjoined readers to acknowledge that the failure of some students is created through the rather ludicrous insistence that “everyone do better than everyone else” (p. 274). W.E.B. Du Bois (1903) asked: How does it feel to be a problem? in the early twentieth century. Although, Du Bois (1903) was talking about the two-ness and double-consciousness of Black Americans in the US, at the beginning of the three decades of the 21st century the entanglement between race and dis/Ability has moved us and other critical interdisciplinary scholars of disability to ask the same question: *How does it feel to be placed in special education as a BIPOC youth at your intersections of power and identities and be labeled as a problem?* (See: Artiles, 2013).

None of these institutional values, practices, or beliefs is preordained, nor are they inescapable. Of course, that does not mean that taking on such matters is an easy task. Providing us a good starting place, Dudley-Marling suggested that instead of asking the classic question—“What’s wrong with [Kevin or Keisha],” it would be far more useful to ask—“What’s wrong with an institution that produces so much failure?” along with its corollary question—“What’s wrong with a culture that created an institution that creates so much failure?” (p. 483).

These questions apply not just to SLD but to all other special education categories as well. Most obviously they apply to categories like serious emotional disturbance (SED), or what is commonly termed behavioral disorders (BD), and Attention-deficit,

Hyperactivity Disorder (ADHD) (both of which are categories whose definitions and identification criteria continue to shift and change even while students are regularly identified as “having” these conditions). But they are also just as relevant to the so called “low incidence” special education categories, including those which involve physically visible impairments. Most importantly, these questions point to an entirely different perspective on the nature of dis/Ability and what it means to be dis/Abled.

At its core, the difference between the medical model of disability and its social and psycho-emotional and intersectional disablism model contender boils down to whether disabilities *exist* as ontologically *real* conditions in the sense that they are intrinsic to the individual, can be quantified, and are properly grounded in the logic of ability as normally distributed. The medical model affirms this perspective, while the social and psycho-emotional and intersectional disablism models holds that dis/Abilities are the product of social interpretations of various human differences at the intersections of power and identities (and that BIPOC dis/Abled folks experience psycho-emotional disablism that constrains their well-being and agency due to the social construction of what counts as dis/Ability in educational contexts across K-16 and beyond) (Iqtadar et al., 2020). Simply put, the social and intersectional disablism model supports the position that disability is a social construct or an *idea* about the nature of certain kinds of human variation that imposes socially originated constraints on the well-being and agency of BIPOC youth that they actively resist (Kliewer and Biklen, 1996; Iqtadar et al., 2020). In other words, across K-16 educational spaces, BIPOC youth have resisted and are agentive against the hegemony of special education labeling, classifications, and identifications as white neoliberal property (Kearl, 2019; Iqtadar et al., 2020). Further, the social and intersectional disablism models call for transformations of our cultural, educational, physical, affective and relational environments at the personal, interpersonal, structural, and political dimensions (Crenshaw, 1990). This resistance also calls for accessibility and inclusion along all dimensions of our society, both internal and external.

It is important to note here that the social and the intersectional disablism model of disability does not pretend that human differences are non-existent or are merely figments of our imagination. What they do ask us to consider is why those differences *make a difference*—or more to the point, it asks us to account for how we *make* some people’s differences problematic. That these differences do exist is one thing, but how we interpret and respond to them is another matter altogether. Applying the social and the intersectional disablism model also does not mean that,

... society should not value activities that some of its members cannot engage in; it is harmless to value the capacity of sight that permits people to behold Rembrandt’s masterpieces, sunsets, or the faces of family members and friends. It is not offensive to prize intellectual accomplishment, athletic prowess, or the ability to appreciate visual beauty and to regret that not everyone we know can enjoy them. To the extent that spina bifida, Down syndrome, blindness, or cystic fibrosis currently precludes people from undertaking some parts of life that people who do not have those traits might experience, the disability critique [the social and intersectional disablism model] acknowledges that disability puts some limits on the “open future” people seek for themselves and their children.

Parens and Asch (2000, p. 25).

We already know that not all students can master the intricacies of advanced calculus or write beautifully crafted essays on Shakespeare’s *King Lear*. And although we continue to value these capabilities, it is crucial that we also make the educational arrangements and pedagogical changes that will allow *all students* to experience the joy of understanding and the intense satisfaction of thinking deeply about things that matter.

Holding high standards for all students does not mean holding the same standard for all students: this is at the heart of privileging equity rather than only equality. Nevertheless, we need to move beyond binaries in order to achieve both/and equity and equality. We also need to do so in order to live out the spirit of the Equal Educational Opportunity (EEO) Act of 1974 in our everyday practices not only within the classroom but also in conducting educational research. Furthermore, this legal construct of EEO, encoded into the Elementary and Secondary Education Act of 1964 asserts that researchers and all agents within the public educational system must hold the complexity of such paradoxes through the love and spirit of learning—with teaching and education as our task at hand as we educate BIPOC youth with and without dis/Abilities. In doing so, we can pay back the educational debts (Ladson-Billings, 2006) that since the founding of the country have been due to historically multiply marginalized groups such as BIPOC youth.

From a social and intersectional disablism model perspective, the relentless attempt to “catch students up” with their so called “normal” peers is both self-defeating and damaging. Equally defeating are the ranking, sorting, selecting, and stratifying activities to which we educators devote so much time, energy, and scarce fiscal resources. Not much reflection is required to recognize that making invidious comparisons among our students contributes little to the goal of educating them. And knowing what individual students need from their teachers does not hinge on how much better or worse they “perform” in any given subject area compared to other students. Put differently, seeing to it that every student is afforded the conditions to become as knowledgeable, as skilled and as caring as possible does not require comparing them to one another. It does not involve labeling some of them. And, finally, an end to special education disability labeling, classification, and identification policies and practices or master narratives (Hernández-Saca, 2016) also means an end to the problem of overrepresentation in special education.

Conclusion

It is not coincidental that research seeking to minimize the problem of overrepresentation has relied on methodological neutrality to do so. This stance is inherently a politically conservative one, "... in that it inevitably must leave the school and social structures basically intact" (Smith and Gallagher, 2008, p. 292). It is a stance that allows, and in fact requires, *researchers to look outside of themselves for technical solutions that never really resolve the problem*. And very importantly it is one that promises, misleadingly, to cordon off researcher bias along with racist and ableist policies and practices—including the researcher's culturally and historically contingent baggage surrounding deeply fraught issues like race and dis/Ability. The main point here is that the researcher's values, interests, and ideological commitments cannot and should not be avoided in the pretension of a neutrality that cannot in any event be achieved (Arzubiaga et al., 2008; Gallagher, 2006). Researchers' conceptual frameworks are not devoid of their personal goals, ideological commitments, identities and positionalities (Ravitch and Riggan, 2017). On the contrary, they must serve as the starting point for the research process (Ravitch and Riggan, 2017). In the end, there is no methodologically privileged, easy way out if we truly want to abolish the deeply entrenched racist and ableist injustices we face as educational equity researchers. The time has always been now to anchor and connect the larger historical-sociological patterned inequities and inequalities faced by BIPOC youth and communities to achieve morally, spiritually and politically catalytic praxical actions for justice for all. What are we waiting for and what are we so afraid of? (Bell, 1995).

References

- Althusser, L., 1969, 2014. On the Reproduction of Capitalism: Ideology and Ideological State Apparatuses. Verso Trade.
- Annamma, S.A., 2015. Innocence, ability, and whiteness as property: teacher education and the school-to-prison pipeline. *Urban Rev.* 47 (2), 293–316.
- Artiles, A.J., Trent, S.C., 1994. Overrepresentation of minority students in special education: a continuing debate. *J. Spec. Educ.* 27 (4), 410–437.
- Artiles, A.J., Kozleski, E.B., Trent, S.C., Osher, D., Ortiz, A., 2010. Justifying and explaining disproportionality, 1968–2008: a critique of underlying views of culture. *Except. Child.* 76, 279–299. <https://doi.org/10.1177/001440291007600303>.
- Artiles, A.J., Kozleski, E.B., Waitoller, F.R., 2011a. Inclusive Education: Examining Equity on Five Continents. Harvard Education Press.
- Artiles, A.J., Kozleski, E.B., Gonzalez, T., 2011b. Para além da sedução da educação inclusiva nos Estados Unidos: confrontando o poder, construindo uma agenda histórico-cultural. *Rev. Teias* 12 (24), 24.
- Annamma, S.A., Connor, D., Ferri, B., 2013. Dis/ability critical race studies (DisCrit): theorizing at the intersections of race and dis/ability. *Race Ethn. Educ.* 16 (1), 1–31.
- Artiles, A., 2009. Re-framing disproportionality research: outline of a cultural-historical paradigm. In: *Multiple Voices for Ethnically Diverse Exceptional Learners*, vol. 11, pp. 24–37.
- Artiles, A.J., 2013. Untangling the racialization of disabilities: an intersectionality critique across disability models. *Du. Bois Rev.* 10 (2), 329–347.
- Artiles, A.J., 2019. Fourteenth annual Brown lecture in educational research: re-envisioning equity research: disability identification disparities as a case in point. *Educ. Res.* 48 (6), 325–335.
- Arzubiaga, A.E., Artiles, A.J., King, K.A., Harris-Murri, N., 2008. Beyond research on cultural minorities: challenges and implications of research as situated cultural practice. *Except. Child.* 74 (3), 309–327.
- Bal, A., Betters-Bubon, J., Fish, R., 2019. A multilevel analysis of statewide disproportionality in exclusionary discipline and the identification of emotional disturbance. *Educ. Urban Soc.* 51 (2), 247–268.
- Baynton, D.C., 2001. Disability and the justification of inequality in American history. In: Longmore, P.K., Umansky, L. (Eds.), *The New Disability History: American Perspectives*. New York University Press, New York, pp. 33–57.
- Bell, D.A., 1995. Who's Afraid of Critical Race Theory, vol. 4. University of Illinois Law Review, pp. 893–910.
- Biklen, D., 1988. The myth of clinical judgment. *J. Soc. Issues* 44 (1), 127–140.
- Blake, J., Butler, B., Lewis, C., Darensbourg, A., 2011. Unmasking the inequitable discipline experiences of urban black girls: implications for urban educational stakeholders. *Urban Rev.* 43, 90–106.
- Blanchett, W.J., 2009. A Retrospective examination of urban education: from brown to the resegregation of African Americans in special education—it is time to "go for broke". *Urban Educ.* 44 (4), 370–388.
- Carrier, J.G., 1986. *Learning Disability: Social Class and the Construction of Inequality in American Education*. Greenwood Press, New York.
- Collins, K., Ferri, B.A., Connor, D.J., Gallagher, D.J., Sampson, J.F., 2016. Dangerous assumptions and unspoken limitations: a disability studies in education response to Morgan, Farkas, Hillemeir, Mattison, Maczuga, Li, and Cook (2015). In: *Multiple Voices for Ethnically Diverse Exceptional Learners*, vol. 16, pp. 4–16 (1).
- Connor, D.J., Ferri, B.A., Annamma, S.A., 2021. From the personal to the global: engaging with and enacting DisCrit theory across multiple spaces. *Race Ethn. Educ.* 5, 597–606.
- Cooc, N., Kiru, E.W., 2018. Disproportionality in special education: a synthesis of international research and trends. *J. Spec. Educ.* 52 (3), 163–173. <https://doi.org/10.1177/0022466918772300>.
- Coomer, M.N., 2019. Deconstructing difference and inclusion in educational research: reflections on the International Journal of Qualitative Studies in Education special edition on difference. *Int. J. Qual. Stud. Educ.* 32 (3), 341–345.
- Crenshaw, K., 1990. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299.
- Cruz, R., Rodl, J., 2018. An integrative synthesis of literature on disproportionality in special education. *J. Spec. Educ.* 52 (1), 50–63. <https://doi.org/10.1177/0022466918758707>.
- Danforth, S., Ressa, T., 2013. Treating the incomplete child: how the science of learning disabilities was built for exclusion. In: Kanter, A.S., Ferri, B.A. (Eds.), *Righting Educational Wrongs: Disability Studies in Law and Education*. Syracuse University Press, Syracuse, NY, pp. 127–150.
- Danforth, S., 2007. Disability as a metaphor: examining the conceptual framing of emotional behavioral disorder in American public education. *Educ. Stud.* 42 (1), 8–27.
- Deno, E., 1970. Special education as developmental capital. *Except. Child.* 37 (3), 229–237. <https://doi.org/10.1177/001440297003700306>.
- William Edward Burghardt, 1868–1963 Du Bois, W.E.B., 1903. *The Souls of Black Folk; Essays and Sketches*, New York: Johnson Reprint Corp., 1968. A. G. McClurg, Chicago.
- Dudley-Marling, C., 2004. The social construction of learning disabilities. *J. Learn. Disabil.* 37 (6), 482–489.
- Dunn, L., 1968. Special education for the mildly retarded: is much of it justifiable? *Except. Child.* 23, 5–21.
- Erevelles, N., 2011. *Disability and Difference in Global Contexts: Enabling a Transformative Body Politic*. Palgrave Macmillan.
- Evans-Winters, V.E., Esposito, J., 2019. Intersectionality in education research: methodology as critical inquiry and praxis. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative, and Methodological Reflections*. Routledge, pp. 52–64.
- Ferri, B.A., Connor, D.J., 2005. Tools of exclusion: race, disability, and (re)segregated education. *Teach. Coll. Rec.* 107 (3), 453–474.
- Ferri, B.A., Connor, D.J., 2006. *Reading Resistance: Discourses of Exclusion in Desegregation and Inclusion Debates*. Peter Lang, New York.

- Fierros, E.G., Conroy, J., 2002. Double jeopardy: an exploration of overrepresentation of minority children in special education. In: Losen, D.J., Orfield, G. (Eds.), *Racial Inequity in Special Education*. Harvard Education Press, Cambridge, MA, pp. 39–70. The Civil Rights Project.
- Gallagher, D.J., 2006. If not absolute objectivity, then what? A reply to Kauffman and Sasso. *Exceptionality* 14 (2), 91–107.
- Gallagher, D.J., 2010. Hiding in plain sight: the nature and role of theory in learning disability labeling. *Disabil. Stud. Q.* 30 (2).
- Gordon, N., 2017. Race, Poverty, and Interpreting Overrepresentation in Special Education. Brookings, Washington. September 20th.
- Gordon, N., 2018. Disproportionality in Student Discipline: Connecting Policy to Research. Brookings, Washington. January 18th.
- Heller, K., Holtzman, W., Messick, S., 1982. Placing Children in Special Education: A Strategy for Equity. National Academy Press, Washington, DC.
- Hernández-Saca, D., Cannon, M.A., 2019. Interrogating disability epistemologies: towards collective dis/Ability intersectional emotional, affective and spiritual autoethnographies for healing. *Int. J. Qual. Stud. Educ.* 32 (3), 243–262.
- Hernández-Saca, D.I., 2016. Re-framing the Master Narratives of Dis/Ability Through an Emotion Lens. *Voices of Latina/o Students With Learning Disabilities*. Arizona State University.
- Hernández-Saca, D.I., 2021. Recovering the spirit. In: Connor, D.J., Ferri, B. (Eds.), *How Teaching Shapes Our Thinking About Disabilities: Stories From the Field*. Peter Lang, pp. 263–276.
- Hosp, J.L., Reschly, D.J., 2004. Disproportionate representation of minority students in special education: academic, demographic, and economic predictors. *Except. Child.* 70 (2), 185–199.
- Iqtadar, S., Hernández-Saca, D.I., Ellison, S., 2020. "If it wasn't my race, it was other things like being a woman, or my disability": a qualitative research synthesis of disability research. *Disabil. Stud. Q.* 40 (2).
- Iqtadar, S., Hernández-Saca, D.I., Ellison, B.S., Cowley, D.M., 2021. Global conversations: recovery and detection of Global South multiply-marginalized bodies. *Race Ethn. Educ.* 24 (5), 719–736.
- Kearl, B., 2019. Special education as neoliberal property: the racecraft, biopolitics, and immunization of disability. *Educ. Stud.* 55 (4), 473–488.
- Kim, C., Losen, D., Hewitt, D., 2010. *The School-to-Prison-Pipeline: Structuring Legal Reform*. New York University Press, New York.
- Kliewer, C. and Biklen, D. (1996) Labeling: who wants to be called retarded? In: W. Stainback and S. Stainback (Eds.) *Controversial issues confronting special education: Divergent perspectives*. Needham Heights, MA: Allyn & Bacon, pp 83–95.
- Kliewer, C., Biklen, D., Petersen, A., 2015. At the end of intellectual disability. *Harv. Educ. Rev.* 85 (1), 1–28.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in U.S. schools. *Educ. Res.* 35 (7), 3–12.
- Larry P. v. Riles, 1986. 342 F. Supp. 1306 (N.D. Cal. 1972) (Preliminary Injunction) Affirmed 502 F. 2d 963 (9th Cir. 1974); 495 F. Supp. 926 (N. D. Cal. 1979) (Decision on Merits) Affirmed (9th Cir. No. 80-427), January 23, 1984. Order Modifying Judgment, C-71-2270 RFP.
- Losen, D.J., Gillespie, J., 2012. Opportunities Suspended, Executive Summary Civil Rights Project/Proyecto Derechos Civiles. The Civil Rights Project. In: Losen, D.J., Orfield, G. (Eds.), 2002. *Racial Inequity in Special Education*. Harvard Education Press, Cambridge, MA, pp. 39–70.
- Losen, D., 2015. Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion (Disability, Culture, and Equity Series).
- MacMillan, D.L., Reschly, D.J., 1998. Overrepresentation of minority students: the case for greater specificity or reconsideration of the variables examined. *J. Spec. Educ.* 32 (1), 15–24.
- McDermott, R., 1993. The acquisition of a child by a learning disability. In: Chaiklin, S., Lave, J. (Eds.), *Understanding practice: perspectives on activity and context*. Cambridge University Press, Cambridge, UK, pp. 269–305.
- McDermott, R., Edgar, B., Scarloss, B., 2011. Global norming. In: Artiles, A., Kozleski, E., Waitoller, F. (Eds.), *Inclusive Education*. Harvard Education Press, Cambridge, MA, pp. 223–235.
- Mercer, J., 1973. *Labeling the Mentally Retarded*. University of California Press, Los Angeles.
- Migliarini, V., 2017. Subjectivation, agency and the schooling of raced and dis/abled asylum-seeking children in the Italian context. *Intercult. Educ.* 28 (2), 182–195.
- Morgan, P.L., Farkas, G., 2015. Is Special Education Racist? New York Times. Opinion pages. <https://www.nytimes.com/2015/06/24/opinion/is-special-education-racist.html>.
- Morgan, P., Farkas, G., Hillemeier, M., Mattison, R., Maczuga, S., Li, H., Cook, M., 2015. Minorities are disproportionately underrepresented in special education: longitudinal evidence across five disability conditions. *Educ. Res.* 44 (5), 278–292.
- NAACP Criminal Justice Fact Sheet. <https://www.naacp.org/criminal-justice-fact-sheet/>.
- Newman, L., Wagner, M., Cameto, R., Knokey, A.M., 2009. The Post-high School Outcomes of Youth With Disabilities Up to 4 Years After High School: A Report From the National Longitudinal Transition Study—2(NLTS2) (NCSE 2009-3017). Washington, DC.
- Ong-Dean, C., 2006. High roads and low roads: learning disabilities in California, 1976–1998. *Socio. Perspect.* 49 (1), 91–113.
- Parens, E., Asch, A., 2000. The disability rights critique of prenatal testing: reflections and recommendations. In: Parens, E., Asch, A. (Eds.), *Prenatal Testing and Disability Rights*. Georgetown University Press, Washington, DC, pp. 3–43.
- Patton Davis, L., Museus, S., 2019. Identifying and disrupting deficit thinking. *Spark: The Online Magazine of the*. National Center for Institutional Diversity.
- Phippen, J.W., 2015. The Racial Imbalances of Special Education. *The Atlantic*. <https://www.theatlantic.com/education/archive/2015/07/the-racial-imbalances-of-special-education/397775/>.
- Rausch, M.K., Skiba, R.J., 2006. Examining Patterns of Disciplinary Disproportionality by Disability Status and Racial Group. Center for Evaluation & Education Policy, Bloomington, IN.
- Ravitch, S.M., Riggan, M., 2017. *Reason and Rigor: How Conceptual Frameworks Guide Research*, second ed. SAGE Publications, Thousand Oaks, CA.
- Reschly, D.J., 1984. Beyond IQ test bias: the National Academy Panel's analysis of minority EMR overrepresentation. *Educ. Res.* 13 (3), 15–19.
- Skiba, R.J., Michael, R.S., Nardo, A.C., Peterson, R.L., 2002. The color of discipline: sources of racial and gender disproportionality in school punishment. *Urban Rev.* 34, 317–342.
- Skiba, R.J., Artiles, A.J., Kozleski, E.B., Losen, D.J., Harry, E.G., 2016a. Risks and consequences of oversimplifying educational inequities: a response to Morgan et al. (2015). *Educ. Res.* 45 (3), 221–225.
- Skiba, R.J., Poloni-Staudinger, L., Gallini, S., Simmons, A.B., Feggins-Azziz, R., 2016b. Disparate access: the disproportionality of African American students with disabilities across educational environments. *Except. Child.* 72 (4), 411–424.
- Smith, J.K., Gallagher, D.J., 2008. An essay on the politics of schooling and educational research. *Cult. Stud.* 8 (3), 284–301. <https://doi.org/10.1177/1532708607310793>.
- Sleeter, C.E., 1986. Learning disabilities: the social construction of a special education category. *Exceptional children* 53 (1), 46–54.
- Smith, J.D., 1997. Mental retardation as an educational construct: time for a new shared view? *Educ. Train. Ment. Retard. Dev. Disabil.* 32 (3), 167–173.
- Stubblefield, A., 2007. "Beyond the pale": tainted whiteness, cognitive ability, and eugenic sterilization. *Hypatia* 22 (2), 162–181.
- Suárez-Krabbe, J., 2014. The other side of the story: human rights, race and gender from a transatlantic perspective. In: Dhawan, N. (Ed.), *Decolonizing Enlightenment. Transnational Justice, Human Rights and Democracy in a Postcolonial World*. Verlag Barbara Budrich, pp. 211–226.
- Thomas, C., 1999. *Female Forms: Experiencing and Understanding Disability*. Open University Press, Buckingham, UK.
- Thorius, K.A., 2019. Facilitating en/counters with special education's cloak of benevolence in professional learning to eliminate racial disproportionality in special education. *Int. J. Qual. Stud. Educ.* 32 (3), 323–340.
- Trent, J.W., 1994. *Inventing the Feeble Mind: A History of Mental Retardation in the United States*. University of California Press, Berkeley.
- U.S. Department of Education, 2016. *Racial and Ethnic Disparities in Special Education: A Multi-year Disproportionality Analysis by State, Analysis by Category, and Race/Ethnicity*. <https://www2.ed.gov/programs/osepidea/618-data/LEA-racial-ethnic-disparities-tables/disproportionality-analysis-by-state-analysis-category.pdf>.
- Williams, P., 1987. Spirit-murdering the messenger: the discourse of fingerprinting as the laws response to racism. *Univ. Miami Law Rev.* 42, 127.
- Zuberi, T., Bonilla-Silva, E. (Eds.), 2008. *White Logic, White Methods: Racism and Methodology*. Rowman & Littlefield Publishers, Lanham, MD.

Rethinking schooling for inclusion

Roger Slee, Center for Disability Studies, School of Sociology & Social Policy, University of Leeds, Leeds, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

The core business of exclusion	506
Inclusive education: reconsidering recurrent themes?	508
What is inclusive education?	508
Why our insistence on disability studies in education?	509
Exclusion: a petite reprise	510
Disabled children and schooling in Australia: systemic failure	511
Damaged or design fault?—fixable or irredeemable?	511
References	513

The core business of exclusion

Call me a contrarian, but there is something of a paradox afoot pressing us to question the authenticity of inclusive education's reception and application. Let me explain—explain this paradox's foundations, its inevitability and the shadow it casts over the future of education. *Inclusive Education* has established itself in education's official lexicon (Wallace, 2009). It has, in my academic lifetime, insinuated itself as a motif for—or mantra in—education policy and practice. It now occupies growing sections on library and publishing house shelves and catalogs and has secured sizable national and global research funding envelopes. Education practitioners, researchers, administrators, consultants, and activists circumnavigate the globe in search of comfortable convention centers to share their aspirations for an inclusive “education for all”. Education Ministers and their Secretaries (aka Director-General) declare inclusive education as an educational aspiration that is simultaneously their stated organizational and pedagogic strategy for improving educational experiences and outcomes for all. Legislation, administrative regulations, international conventions, charters and memoranda of understanding stack-up and are referred to as the legitimizing platform for inclusive schooling. Trans-national supra-organizations such as UNESCO (2020, 2022), the OECD (Cerna et al., 2021), The World Bank (2020), the Asian Development Bank (2022), Open Society Foundation (2019), the Brookings Institution Sengeh and Winthrop (2022), Save the Children (2014) and the International Monetary Fund Filmer et al. (2021) have all signed up to the cause as their recent reporting and project listings reflect.

Other *key indicators*, applying the language of our time, may be readily assembled to support the proposition that inclusive education is more than a rhetorical flourish. Many children who would otherwise be engaged in unremitting manual labor in clothing sweatshops, on village farms or in a multitude of factories in countries where the struggle for daily subsistence proscribes childhood dreams are finding their place in classrooms. Girls who would have been married as children and not received their right (United Nations Convention on the Rights of the Child, 1989: see Articles 23, 28, 29) to basic education are being sponsored by NGOs or their communities to complete their schooling. Disabled children, once protectively hidden by parents at home or gathered in specially designated special schools, are being encouraged to attend their neighborhood school with their brothers, sisters and neighborhood friends. Children living in poverty are targeted, “targets” being another of *eduspeak's* default vocabulary, for *leveling-up*.

I don't mean to appear cynical or derisive. Nor do I want to devalue progressive gains for children who would otherwise continue to be excluded from their right to an education. To be sure, people of good-will are dedicated to securing inclusive and excellent educational experiences for more children through a range of reforms to the modern educational edifice. Clearly, more children are attending schools than was the case just a handful of decades ago. This ought to be cause for celebration, not hesitancy. Many say the successes of inclusive education policies and programs should form the catalyst for continuing educational reform (Ainscow, 2016). An erstwhile school reformer, I want to kick away the shadow of doubt and struggle on. This struggle has become harder. Perhaps it is pandemic, perhaps advancing years. The combination of institutional intransigence, discrimination and chicanery form a belligerent foe. The weight of doubt sits heavier than the accelerant of optimism.

Like many others I was drawn back to Albert Camus' classic novel of pestilence: *The Plague* (*La Peste*), when living around the corner from Harrogate's emergency 500 bed Nightingale Hospital that had been opened by Captain Sir Tom Moore in April 2020 to relieve anticipated surges of COVID-19 pressure at the local National Health Service hospital. Soon after, in March 2021, it closed having not treated a single patient. Nevertheless, passing it on my daily walk, it stood as a silent sentinel; it's two giant oxygen tanks casting foreboding shadows. The building was shrouded in white hessian and security guards loitered in clusters around its perimeter. As the ambulance sirens wailed at night my imagination filled those 500 beds (Fig. 1).

Let's return to Camus and the inhabitants of the Algerian coastal township of Oran where *The Plague* is set. The novel follows the deleterious impacts of bubonic plague on the relationships between the inhabitants of this pandemic-ravaged township. It is a chronicle of social atrophy. The novel is of course allegorical. For Camus the plague represents fascism. In this respect, the populist formations of the far right remain a threat to civil society (Harnish, 2017; Camus and Lebourg, 2017; Castells, 2018). The novel ends



Fig. 1 NHS Nightingale Hospital Yorkshire, Harrogate

with the disappearance of the disease-carrying rats, the diminution of infection and the emergence of the town's survivors in the throes of revelry. An exhausted Monsieur Rieux, the doctor whom we have followed in his heroic fight against the plague, makes his way home slowly walking, bearing the weight of foreboding.

Indeed, as he listened to the cries of joy that rose above the town, Rieux recalled that this joy was always under threat. He knew that this happy crowd was unaware of something that one can read in books, which is that the plague bacillus never dies or vanishes entirely, that it can remain dormant for dozens of years in furniture or clothing, that it waits patiently in bedrooms, cellars, trunks, handkerchiefs and old papers, and that perhaps the day will come when for the instruction or misfortune of mankind, the plague will rouse its rats and send them to die in some well contented city.

Camus (2013: 331–332).

The paradox to which I refer is the resilience of educational exclusion. It lurks, mostly unspoken, behind and within any discussion of inclusive education. The problem is not just its omnipresence. Although, that is a monumental problem to be considered further in this chapter. Exclusion is a part of the DNA of schooling, deeply embedded in its architecture and often difficult to detect (Slee, 2018). In a recent paper entitled *Against School: an epistemological critique*, Ball and Collet-Sabe (2021) enlist Rogan (2017) treatise on the moral economists to suggest that examinations of schooling, like most economic debates, have been distracted by technical issues concerning the effects of schooling. Who are the winners? Who are the losers? How do we create more winners with minimal change to the structure and culture of schooling? Thereby, the fundamental question about the moral legitimacy of the institution in and of itself escapes interrogation.

Rogan's point is that the critique of capitalism in the twentieth century shifted away from a fundamental demolition of its "moral and spiritual desolation" to a single-minded focus on the calculation of the relative advantages and disadvantages it generates—a shift from "Is it morally wrong?" to "Does it have bad outcomes?"—Who wins and who loses? We want to suggest that the critical tradition in education and criticism of the school by sociologists and others has suffered a similar displacement or avoidance of moral argumentation by calculative evaluation.

Ball and Collet-Sabe (2021: 1–2).

This chapter, mindful of Ball and Collet-Sabe's (2021:3) observation that reforming an "intolerable and irredeemable" institution of "moral and spiritual desolation" is a futile modernist project, is divided into three parts. First is a truncated discussion of inclusive education; what it is and why there exists a need for *disability studies in education*. Second is a brief re-examination of exclusion and schooling. And finally, I will tentatively commence what must be an ongoing discussion of whether reform is an exercise in futility or a cry of hope.

Inclusive education: reconsidering recurrent themes?

Rather than offering yet another incomplete chronology of inclusive education (Slee, 2018b), I will outline two recurring questions:

1. What is inclusive education?
2. Given the surging acceptance of inclusive education, why our insistence on disability studies in education?

My aim is to support the more nuanced discussion (Thomas and Loxley, 2022) of the possibilities and limits of school reform.

What is inclusive education?

In general terms inclusive education is defined as any policy, program or practice aiming to increase access, presence, participation and success for all students in education. Accordingly, attempts to identify and dismantle barriers to education for all students constitute inclusive education. It's as simple and as complex as that. But let's explore this proposition a little further as it is obvious that exclusion is contingent. It is formed and shaped within a matrix of historic, political, cultural, geographic, economic elements and by varying conceptions of the correct purposes, organization and delivery of schooling.

Since the invention of modern schooling there have been successive attempts to understand the differential impacts of schooling on diverse population cohorts, to extend provision for more students and to improve their educational experiences and benefits. We see this in the "new sociology of education" (Apple, 1979, 1982; Bernstein, 1970, 1973; Young, 1971; Bourdieu and Passeron, 1977; Whitty, 1985). Such work debunked the liberal myth of meritocratic schooling and its claim that hard work and talent trumped class and other identity markers such as gender, race, first nations, minority cultural or religious affiliation, or perceived disability in the distribution of educational opportunities and rewards. Sandel (2020) has returns us to an examination of the tyranny of merit in a discussion of buying university admissions in the US

In an unequal society, those who land on top want to believe their success is morally justified. In a meritocratic society, this means the winners must believe they have earned their success through their own talent and hard work.

Paradoxically, this is the gift the cheating parents wanted to give their children to live in affluence, they could have given them trust funds. But they wanted something else—the meritocratic cachet that admission to elite colleges confers.

Sandel (2020: 13).

The new sociology of education challenged deterministic psychological explanations of intelligence, heritability and educational achievement, dismissing the resurrection of risible claims that failure could be linked to the inferior intelligence of the working class and people of color (Hernstein and Murray, 1994) and heightening our awareness of the contingent nature of educational opportunities and outcomes (Ball, 2006; Gillborn, 2008; Tomlinson, 2013, 2017; Tait, 2018). Put simply, different students receive different deals from education. Some students are privileged, others are not. The assignation of educational success and failure is not random (Teese, 2013).

In his examination of stratification in schooling, American criminologist, Polk (1984) describes *sponsored* students and *the new marginal youth*, or those whom Bauman (2004: 39) would call the *surplus population of liquid times*. Obvious barriers to access and participation in schooling are apparent in the countries of the global south when compared with those of the global north (UNESCO, 2020; Pather and Slee, 2019). Widening disparities in participation and attainment are also obvious within affluent countries along the fault lines of privilege and poverty (Connell, 1993; Dorling, 2018; Lamb et al., 2015, 2020). "All societies", Bauman instructs, "produce their own set of strangers; but each kind of society produces its own kind of strangers and produces them in its own inimitable way" (1997:17).

Targeted reforms and compensatory education programs (Rizvi and Kemmis, 1987; Connell et al., 1992; Rizvi, 2013) have been, and continue to be deployed by jurisdictions to mitigate disadvantage. Such interventions are built into jurisdictional funding models (Deloitte Access Economics, 2017; Department of Education, 2022) and augmented through specific programs. Notwithstanding, these periodic tilts at redistribution, "education equality is always proclaimed, never realized" (Picketty, 2022: 176). The "concentration of expenditures for education to the benefit of a minority is, to be sure, less extreme than it used to be, but it remains considerable, and is hardly in conformity with contemporary talk about equality of opportunity" (Picketty, 2022: 179–180).

Tracing the origins and development of inclusive education is fraught. Some may return to the introduction of mass schooling and track forward (Tait, 2018), others might suggest that civil rights legislation such as *Brown v. Board of Education Topeka* (1954) in the US together with international agreements and declarations such as the UN Convention on the Rights of the Child provided the impetus for inclusion for all. Danforth (2009) returns us to special education activists who refused to accept the proposition that disabled children were ineducable. In doing so, he exposes tensions between a spirit of radicalism and the conservatism of the structures and cultures of schooling that accommodated its application so as not to unsettle the status quo. Of course, such accommodations drew and protected clear boundaries.

Attempts to locate the source of inclusive education as an international education policy imperative often return to the *World Education Conference on Education for All* in Jontiem, Thailand in 1990 (UNESCO, 1990). Delegates gathered from around the world to consider strategies for removing barriers to education emanating from: poverty; conflict; population displacement; “natural” and climate disaster; gender; child labor; child marriage; ethnic, religious, linguistic and or tribal affiliations; disability; gender and their intersections. Herein we see the origins of the “*Education for All*” movement. Of course, within the assembled delegates, specific groups and individuals were concerned with specific renditions of exclusion (e.g., the exclusion of girls and women; the exclusion of the poor; the exclusion of “minority” races and ethnicities; the exclusion of the disabled; the exclusion of first nations peoples; the exclusion of children of travelers, refugees and asylum-seekers). For those who were primarily concerned with the education and the exclusion of disabled students, the follow-up conference on special needs education in the Spanish city of Salamanca over 25 years ago in 1994 (UNESCO, 1994) is seen as a watershed event. The *Salamanca Statement and Framework for Action on Special Needs Education* was endorsed and adopted by the 92 represented governments and 25 international organizations from around the world. It was jointly released by UNESCO and the Spanish Ministry of Education. It is important to note two of the principles set out in the *Salamanca Statement and Framework for Action on Special Needs Education*:

- Those with special educational needs must have access to regular schools which will accommodate them within a child-centered pedagogy capable of meeting their needs.
- Regular schools with this inclusive orientation are the most effective means of combatting discriminatory attitudes, creating welcoming communities, building an inclusive society and achieving education for all; moreover, they provide an effective education to the majority of children and improve the efficiency and ultimately the cost effectiveness of the entire system.

UNESCO (1994: viii & ix).

Reflecting on the creation and legacy of the *Salamanca Statement and Framework for Action on Special Needs Education* some 25 years later Ainscow et al. (2019) observe that:

The publication of the Salamanca Statement suggested a significant change in policy direction, not least for the education of disabled students. However, the document contains some ambiguities that have become apparent when used in the field. In particular, the deficit orientation of special education discourse continues to undermine more radical moves, as argued in some of the papers. Indeed, this conservative discourse has protected resistance to the reforms it foreshadowed.

Why our insistence on disability studies in education?

An impression I have formed is that there has been a systematic move across the special education industry to rebrand itself. Although I have not yet amassed sufficient data to *scientifically* substantiate this impression, others have worked toward this (Brantlinger, 2006; Tomlinson, 2017). There has been a maneuvering to merge the discourses of special education and inclusive education to forge a view of common purpose and practice in the service of inclusive schooling. Repetition and saturation suppress any audible intake of breath when we encounter astonishing phrases such as “*reverse inclusion*”, or “*inclusive special education*”, or read course descriptions for postgraduate awards in *special and inclusive education*, or hear a headteacher, Special Education Needs Coordinator (SENCO) or in Queensland a Head of Special Education (HOSE) direct us to the inclusion room where all of the disabled students take their lessons, or even read of an education jurisdiction in Australia awarding its *Showcase Award for Excellence in Inclusive Education* to special schools. A confluence of factors may surreptitiously contribute to the irony of celebrating separate special schools as exemplars of inclusion. At play is a powerful teachers’ industrial organization that continues to frame the education of disabled students as an industrial working conditions issue rather than as a political question of overdue organizational reform to dismantle ableism. There is a strong special schools administrators’ lobby and historic commitment to an individually based categorical model for distributing support which has had longstanding perverse effects (Deloitte Access Economics, 2017). Put simply, through a series of linguistic and organizational adjustments, there are in 2022 powerful vested interests maintaining educational disablement just as there were when Tomlinson (1982; 1985) wrote her then radical sociological analyses of special education.

Inclusive education has become a catch-all term for a broad spectrum of practices including the building of new separate special schools, the refurbishment of existing special schools, the co-location (typically with clear lines of division—see Walker, 2019) of special and regular schools, visitations of non-disabled children to special schools and classrooms to familiarize themselves with and befriend disabled children, the renaming of special schools as *specialist schools* to merge into the landscape of specialist science schools, specialist arts schools, specialist music schools, specialist golf schools. A legacy of Dame Mary Warnock’s *Report of the Committee of Enquiry into the Education of Handicapped children and Young People* (DfES, 1978), children, teachers, rooms, units, buildings, training courses, curriculum consultants—the list is too long to complete—all gather under the title of *special educational needs* (Slee, 2018). Of course, the administrative insignia: *special educational needs*, echoes assumptions of individual defectiveness adhering to the tradition of special schooling and the expansive industry that formed around it (Tomlinson, 2017). Repetition in usage (Bourdieu, 2014) transformed special educational needs to an institutional truth.

Anti-racist education researcher and activist, Barry Troyna pressed English Local Education Authorities and schools in the 1980s and 1990s to resist liberal displays of *multiculturalism* such as the celebration of *saris*, *samosas* and *steel bands*. Such displays deflect from institutional racism and the minutiae of assimilationist practices in white British schooling (Troyna, 1993), providing a protective veneer for racism. Enlisting Critical Race Theory (Gillborn, 2008; Ladson-Billings, 1995) to interrogate the ontologies and

epistemologies of *whiteness* or racism in education is the first step in working toward an inclusive education for these Black British or Asian students. So too for deconstructing disabling school cultures and practices. Traditional special educational conceptions of disability and impairment support ablest or disabling schooling. Enablement requires a different ontology and epistemology, and this has profound implications for: what we teach; how we teach; how we assess; how we physically arrange schools and classrooms; how we understand the nature and purpose of schooling; how we perceive and respond to diverse student populations; how we educate teachers; how we re-educate teachers; how we distribute resources and support across education jurisdictions (Morton and Guerin, 2017; Gallagher, 2010).

No doubt, this chapter will be read as an attack on special education. Such a reaction is reductive and misses the point. Special education is one element of a constellation of educational provision. Special education exists because of the structures and cultures of schooling at large. The erstwhile co-dependence of regular and special schools continues to support the surveillance, calibration, categorization and subdivision of students *in their best interests* (sic) and of course in the interest of the stable organization and administration of schools. I return to Ball and Collet-Sabe (2021: 3):

... the problem is not with the ambitions of goodness, but its vehicle—the school and schooling, as the default signifiers of education. The project of inclusion mis-reads the school as a site of opportunity and possibility, and we are always disappointed (Slee, 2018). The essence and *raison d'être* of the school is in fact normalisation and perversely inclusion may contribute to that end.

The rhetoric of *special educational needs* returns us to the assumptions, traditions and practices of defectology, the outcome of which is ableism and the continuing temporal, cultural and spatial separation of disabled students. The response requires more than a change of vocabulary, or the provision of “safe spaces” in which to congregate disabled students in mainstream schools, it requires changes to the foundations of how we know, speak and operate in our world (Rix, 2015; Thomas and Loxley, 2022). Accordingly, the move from special education, and axiomatically regular schooling, to disability studies in education is overdue. Borrowing from Said (2000: 236–237), it may be that inclusive education, having been *tamed and domesticated*, has lost its *insurrectionary zeal*.

Exclusion: a petite reprise

It is beyond the scope of this chapter to forensically describe and analyze educational failure and exclusion, their antecedents and effects. Having already acknowledged the antiquity and ubiquity (Slee, 2011) of educational failure and exclusion, the intention here is to suggest two constitutive features of exclusion as a prelude to remarks on the possibilities of and limits to school reform. Since we have been concerned with the education of disabled students, I will also draw on recent observations about their experiences of schooling in Australia.

Prior to the grip of the COVID-19 pandemic, UNESCO (2019) reported 258.4 million children excluded from school due to a range of intersecting factors reflecting divisions along the fault-lines of privilege and disadvantage. Director-General, Audrey Azoulay revised this estimate suggesting that with temporary school closures because of COVID-19, the numbers of children out of school had reached 290 million. Educational marginalization and exclusion is contingent (UNESCO, 2020). Here I will present just two features of educational exclusion: *structural features* and *cultural features*. Of course, this presentation is constructed to order the discussion because the reality is the messy collision of these features.

As we have already observed, schooling is the boiler-room for the reproduction of inequality (Apple, 1982, 1990; Dorling, 2010; Picketty, 2022). In his recent appraisal of the education system in the UK Sam Bright (2022: 17) is unambiguous:

Elite, fee-paying institutions have funnelled students into the upper echelons of business, law and politics for centuries—a rigged system that the Labour Party has variously attempted to knobble, with little success.

England is not unique. The structure of school funding across and within jurisdictions illuminates fundamental inequalities in the material conditions of schooling that have direct impacts on educational opportunities and outcomes. Here we may be referring to the differences between a school with a technology rich theater and rehearsal spaces, with science laboratories or with sports stadia in a leafy suburb in the Eastern suburbs of Melbourne and the windowless, dirt floored room in rural Ethiopia where 90 children attend to their teacher at the small blackboard. In one such school a principal showed me a multi-function printer in his office that had been donated by an NGO. Still in its wrapping and box, it was of little value to a school without electricity and the other basic accoutrements of a contemporary school. Even within affluent countries, we see *savage inequalities* (Kozol, 1991). The funding of schools in affluent countries reflects and reproduces educational inequality and disengagement (Lamb et al., 2020; Wilkinson and Pickett, 2010; Connell, 1993). The supplementation of parent fees in imprecisely described *private* or *independent* schools through general taxation revenue reflects the “trickle-up economics of education” (Slee, 2018). The divisions of privilege and disadvantage are clearly manifest in the structure of material resources, but it also privileges and attenuates curriculum and pedagogy (Teese, 2013; Lamb et al., 2020).

The second feature of educational exclusion is both structural and cultural. Education is the site of intense competition (Ball, 2017). Using international testing programs, national school systems compete against each other to secure higher ranking positions on international performance tables. Many jurisdictions have their own renditions of student performance testing with elaborate ranking systems to record and broadcast the performance of states, regions, provinces and individual schools. To support achieving higher standards, jurisdictions apply additional measures for auditing schools and their performance across several domains. Incentives and sanctions are used to assist schools to improve their performance. Achievement is rewarded, underachievement is not. Schools are locked into competition against each other. The logic is compelling, the practice is complicated and often traumatic (Ravitch, 2014; Luff, 2021). The important point for this discussion is that performance against the criteria of the testing programs and the school audits directly impacts on the profile, funding and futures of schools in the educational marketplace (Ball, 2017). Schools with higher performance rankings are more heavily subscribed to and exercise choice over their student profile to protect their ranking. In this climate students become the *bearers of results*. The modern educational marketplace fosters an ethos of competitive individualism. Parents, students, teachers, and schools lock into a competition that inevitably must produce failure and exclusion, winners and losers (Ravitch, 2020).

Accordingly, we have seen schools applying what Gillborn and Youdell (1999) referred to as “educational triage” to determine which students would add value to the school’s profile and which students presented a risk to investment. Students are more and less attractive, more or less risky for increasingly risk-averse schools. Managing diverse populations who challenge institutional order and performance takes various forms. We have seen the continuing rise of school exclusions and diversion of disengaged and disruptive students to off-site centers, referral units, and alternative programs (Mowatt, 2022). Many jurisdictions manage diversity, difference and disengagement through their special educational needs or additional needs protocols and practices (Slee, 2011). The absence of the unskilled labor market to absorb the failure of schooling has coincided with the increasing diagnosis of students with an expanding range of syndromes (Tait, 2010; Harwood and Allan, 2014; Rose, 2007). Schools and teachers are being instructed in the application of DSMV to managing classrooms (Tobin and House, 2016). The requirement for more special programs and personnel has its own gravity (Tomlinson, 2017; Rose, 2007).

Disabled children and schooling in Australia: systemic failure

Following a highly publicized case in the Australian Capital Territory of a child with autism being restrained in a purpose-built cage (called *The Sanctuary*) within a classroom storage area (Lawson, 2015; Macdonald, 2015), a review of educational provision for students with complex needs and challenging behavior (Shaddock et al., 2016) was commissioned. Further complaints about restraint and seclusion followed in other jurisdictions. Successive reports of inquiries and reviews converged in presenting evidence of the negative experiences of disabled children (Deloitte Access Economics, 2017). An Australia-wide Royal Commission into Abuse, Neglect and Exploitation of People with Disability was convened in April 2019 (Royal Commission into Abuse, Neglect and Exploitation of People with Disability, 2019). The Royal Commission held a public hearing in November of that year for four days in the northern Queensland city of Townsville to gather evidence concerning:

1. Inclusiveness in education as it relates to students with disability; and
2. The implementation of existing policies and procedures relating to inclusive education of students with disability, with focus on the Queensland government education system.

Royal Commission into Abuse, Neglect and Exploitation of People with Disability (2020: 5).

The expert witness evidence to the Royal Commissioners has been published in an Interim Report from the public hearing (Royal Commission into Abuse, Neglect and Exploitation of People with Disability, 2020) (Table 1). Following is an indicative list of nine areas of complaint:

Damaged or design fault?—fixable or irredeemable?

I return to the resilience of exclusion in what has been described as the “*age of inclusion*” (O’Keefe, 2004). Do we shuffle along the continuum of school reform as we have for decades? Condemned at the end of each initiative to plug the leaks and seal the cracks, rebranding the next round of fixatives?

In other words, the school, and education as currently conceived and represented by the school, as a site of equity, is a lost cause. That is, despite all the debates around segregation, inequality and exclusion, despite or because of innovation, what is sought, desired and struggled for in movements of school reform, of virtually all sorts, is not something different but rather another version of the same thing, using the same architecture and paraphernalia that provide for and deliver, in new and old ways, division, exclusion, normalisation, and categorisation.

Ball and Collet-Sabe (2021: 3).

Over three decades ago, the English sociologist John Furlong asked why has, “... despite its popularity in academic circles the impact of the sociological perspective on policy (on student disaffection and disruption) ... been negligible” (1991: 293)? “... ”

Table 1 Areas of complaints about education raised with the Royal Commission into Abuse, Neglect and Exploitation of People with Disability, 2020 (Australia).

1. Gatekeeping	Refers to the practice of refusing entry for a student with disability or persuading them to attend another school or special school.
2. Overrepresentation in Student Disciplinary Absences & Unique Attendance Schedules	Students with disability are more likely to incur student disciplinary absences (suspensions/exclusions). Some schools were setting unique schedules of attendance for students with disability.
3. Problems with Making Reasonable Adjustments	The disability discrimination Act requires that schools make “reasonable accommodations” to enable a child with disability to attend and participate in schooling. Parents complained that the adjustments were too frequently not forthcoming.
4. Bullying & Mistreatment	There were reports of bullying of children by non-disabled children and mistreatment by teachers and other adults.
5. Restrictive Practices & Seclusion	Reports of nature, frequency and misapplication of restrictive practices and seclusion of children with disabilities.
6. Low Expectations & Underachievement	Endemic lowering of expectations and problems with curriculum, pedagogy and assessment.
7. Inadequate Teacher Preparation.	Teachers, school principals and parents presented concerns about the amount, nature and quality of preparation for inclusive education.
8. Funding Gaps and Inappropriate Resourcing Models.	Funding models have historically had perverse and unproductive effects. Largely based on diagnostic categories.
9. Inadequate & Ineffective Complaint Resolution Procedures.	Unequal power relations are reinforced through complaints and appeals processes which accentuate adversarial relationships.

the individualized, psychologically orientated approach stills holds sway in most parts of the educational profession” (Furlong, 1991: 293). In what remains a seminal essay, he suggested that in part, it could be attributed to inadequacies in sociological research. Accordingly, he constructed an investigation of the production of the structures of ability, values and occupational identity to offer a nuanced model of human personality to generate more strategic and hopeful levers for the engagement and inclusion of marginalized students.

Twenty-five years after the publication of Furlong’s working through a theoretical impasse Beck’s (2016) *Metamorphosis of the World* was published posthumously. Prior to so much as a whisper of the COVID-19 and pandemic that was ahead, Beck (2016: 3) wrote as “an attempt to rescue myself, and perhaps others too, from a major embarrassment”. As a spectator of the global events revealing themselves on television, he felt at a loss to provide a satisfactory explanation for this “*unhinged world*” with the sociological tools he had hitherto relied on. Streeck (2016), considering the end of capitalism, takes up the challenge issued by Burawoy (2005; see also 2021) in his Presidential Address to the American Sociology Association; namely the challenge to explore the public mission of sociology. Streeck (2016: 237) was fascinated “... with the contrast between the progressive decay of the politics and economy in the United States and the star-studded social sciences departments from Harvard to Stanford”. “*What was all this brilliance good for?*” he asks.

Hauling the economy back into society, and indeed into sociology, may be a programme for which one could today find allies, in a world in which states are about to be turned into something like public corporations having to earn the confidence of capital givers; in which international organisations function as deposit insurance or debt collection agencies on behalf of private investors; and governments begin to resemble corporate managements pressed to extract “creditor value” from citizens turned into workforces disciplined by capital markets (p. 250).

He calls for sociology to ready itself “*for the moment in which the foundations of modern society will again have to be rethought*” (p. 250). Streeck’s prescience is noteworthy when reflecting on such moment-inducing recent crises (*Sattelzeit*) including, global pandemic; the electoral victory and defeat of Trump and the attack on Capitol Hill; Putin’s invasion of The Ukraine and its multiple and far-reaching impacts including the threat to food security for the world’s most vulnerable populations; deepening environmental disasters; successive waves of population displacement and the turning away of those who could and should provide refuge.

Surveying inclusive education policy and practice, special education and the experiences of disabled students over recent decades simultaneously reflects the progress of social movements, trends in education administration, and the inevitability of the institutional failures of schooling. Where once educators spoke without second thoughts of a bifurcated school system comprising regular and special schools, we have seen fragmentation, niche branding around curriculum specializations or identity markers, academisation, the growth of independent state schools and charter schools within that part we described as regular. Alternative educational provision continues to grow, including rises in home-schooling. Special schools, having commissioned sign-writers, are more than

likely to be specialist schools or some are centers for educational inclusion. The special or regular school debate has long been beside the point. This may be the moment to commence our thinking, with new intellectual tools, about reconstructing the task of educating ourselves and our children and applying these tools to a new imaginary that is fit for the times we live in.

References

- Ainscow, M., 2016. *Struggles for Equity in Education. The Selected Works of Mel Ainscow*. Routledge, Abingdon.
- Ainscow, M., Slee, R., Best, M., 2019. The Salamanca statement 25 Years on (editorial). *Int. J. Incl. Educ.* 23 (7–8), 671–676. <https://www.tandfonline.com/doi/full/10.1080/13603116.2019.1622800>.
- Apple, M., 1979. *Ideology and Curriculum*. Routledge and Kegan Paul, Boston.
- Apple, M., 1982. *Education and Power*. Routledge and Kegan Paul, Boston.
- Asian Development Bank, 2022. ADB's Work in the Education Sector. Asian Development Bank, Manila. <https://www.adb.org/what-we-do/sectors/education/overview>.
- Ball, S.J., 2006. *Education Policy and Social Class: The Selected Works of Stephen J. Ball*. Routledge, Abingdon.
- Ball, S.J., 2017. *The Education Debate*, third ed. The Policy Press, Bristol.
- Ball, S., Collet-Sabe, J., 2021. Against school: an epistemological critique. *Discourse*. <https://doi.org/10.1080/01596306.2021.1947780>.
- Bauman, Z., 2004. *Wasted Lives: Modernity and Its Outcasts*. Polity, Cambridge.
- Beck, U., 2016. *The Metamorphosis of the World*. Polity Books, Cambridge.
- Bernstein, B., 1970. *Education Cannot Compensate for Society*, vol. 15. New Society, (London), pp. 344–347, 387.
- Bernstein, B., 1973. *Class, Codes and Control*, vol. 1. Routledge & Kegan Paul, London.
- Bourdieu, P., 2014. *On the State*. Polity Press, Cambridge.
- Bourdieu, P., Passeron, J.C., 1977. *Reproduction in Education, Society and Culture*. Sage Publications, London.
- Brantlinger, E.A., 2006. Who Benefits From Special Education? Remediating (Fixing) Other People's Children. Lawrence Erlbaum Associates, Mahwah.
- Burawoy, M., 2005. 2004 ASA presidential address: for public sociology. *Am. Socio. Rev.* 70 (1), 4–28.
- Burawoy, M., 2021. *Public Sociology: Between Utopia and Anti-Utopia*. Polity, Cambridge.
- Camus, A., 2013. *The Plague*. Penguin Books, London.
- Camus, J.-Y., Lebourg, N., 2017. *Far-Right Politics in Europe*. The Belknap Press of Harvard University Press, London.
- Castells, M., 2018. *Rupture: The Crisis of Liberal Democracy*. Polity Books, Cambridge.
- Cerna, L., Mezzanotte, C., Rutigliano, A., Brussino, O., Santiago, P., Borgonovi, F., Guthrie, C., 2021. Promoting Inclusive Education for Diverse Societies: A Conceptual Framework. OECD Publishing, Paris. <https://doi.org/10.1787/94ab68c6-en>. OECD Education Working Papers, No. 260.
- Connell, R., 1993. *Schools and Social Justice*. Temple University Press, Philadelphia.
- Connell, R.W., White, V.M., Johnston, K.M., 1992. An experiment in justice: the disadvantaged schools program and the question of poverty, 1974–1990. *Br. J. Sociol. Educ.* 13 (4), 447–464.
- Danforth, S., 2009. *The Incomplete Child*. Peter Lang, New York.
- Deloitte Access Economics, 2017. Review of Education for Students With Disability in Queensland State Schools. Department of Education and Training, Brisbane. <http://education.qld.gov.au/schools/disability/docs/disability-review-report.pdf>.
- Department for Education and Science, 1978. *Special Educational Needs. Report of the Committee of Enquiry Into the Education of Handicapped Children and Young People*. Her Majesty's Stationery Office, London. (The Warnock Report).
- Department of Education, 2022. *National School Reform Agreement: 2021 Annual Report on Progress in Implementing the National Policy Initiatives*. Government of Australia, Canberra. <https://www.dese.gov.au/national-school-resourcing-board/resources/national-school-reform-agreement-2021-annual-report-progress-implementing-national-policy>.
- Dorling, D., 2010. *Injustice: Why Social Inequality Still Persists*. Bristol: Bristol University Press.
- Dorling, D., 2018. *Peak Inequality: Britain's Ticking Timebomb*. Policy Press, Bristol.
- Filmer, D., Gatti, R., Rogers, H., Spatafora, N., Emrullahu, D., 2021. Education & Health for Inclusiveness. IMF Working Paper (WP/21/60). International Monetary Fund, Washington, D.C.
- Furlong, V.J., 1991. Disaffected pupils disaffected pupils: reconstructing the sociological perspective. *Br. J. Sociol. Educ.* 12 (3), 293–307.
- Gallagher, D., 2010. Hiding in plain sight: the nature and role of theory in learning disability labeling. *Disabil. Stud. Q.* 30 (2). <https://dsq-sds.org/article/view/1231/1278>.
- Gillborn, D., 2008. *Racism and Education: Coincidence or Conspiracy?* Routledge, London.
- Gillborn, D., Youdell, D., 1999. *Rationing Education: Policy, Practice, Reform, and Equity*. Open University Press, Buckingham.
- Grayling, A.C., 2018. *Democracy and Its Crisis*. OneWorld Publications, London.
- Harnish, A., 2017. Ableism and the Trump phenomenon. *Disabil. Soc.* 32 (3), 423–428. <https://www.tandfonline.com/doi/full/10.1080/09687599.2017.1288684>.
- Harwood, V., Allan, J., 2014. *Psychopathology at School: Theorizing Mental Disorders in Education*. Routledge, London.
- Herrnstein, R.J., Murray, C.A., 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press.
- Kozol, J., 1991. *Savage Inequalities: Children in America's Schools*. Crown Publishers, New York.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- Lamb, S., Jackson, J., Walstab, A., Huo, S., 2015. *Educational Opportunity in Australia 2015: Who Succeeds and Who Misses Out*. Centre for International Research on Education Systems, Victoria University. Mitchell Institute, Melbourne.
- Lamb, S., Huo, S., Walstab, A., Wade, A., Maire, Q., Doecke, E., Endekov, Z., 2020. *Educational Opportunity in Australia 2020: Who Succeeds and Who Misses Out*. Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute, Melbourne.
- Lawson, K., 2015. Child Placed in Cage: Canberra Parent Says Schools Can't Cope With Special Needs Students. *The Canberra Times*, 3rd April. <https://www.canberratimes.com.au/story/6069575/child-placed-in-cage-canberra-parents-says-schools-cant-cope-with-special-needs-students/>.
- Luff, I.R., 2021. The Impacts of OFSTED Inspection on Secondary Schools and Their Teachers Between 1993 and 2018: A Study of the Perceptions of Teachers With Multiple Experiences of School Inspection. Unpublished PhD Thesis Submitted to the School of Education and Lifelong Learning. University of East Anglia. June. <https://ueaeprints.uea.ac.uk/id/eprint/82706/1/Redacted%20Thesis.pdf>.
- Macdonald, E., 2015. Cage for Autistic Child at Canberra School a Shocking Wake-Up Call. *The Sydney Morning Herald*. September 9th. <https://www.smh.com.au/opinion/cage-for-autistic-child-at-canberra-school-a-shocking-wakeup-call-20150909-gjictu.html>.
- Morton, M., Guerin, A., 2017. Examining assessment for students with special educational needs in Aotearoa New Zealand. Creating new possibilities for learning and teaching for all. In: Allan, J., Artiles, A. (Eds.), *Assessment Inequalities*. World Yearbook of Education. <https://doi.org/10.4324/9781315517377>.
- Mowatt, J.G., 2022. *Building Community to Create Equitable, Inclusive and Compassionate Schools Through Relational Approaches*. Routledge, Abingdon.
- Open Society Foundation, 2019. *The Value of Inclusive Education*. <https://www.opensocietyfoundations.org/explainers/value-inclusive-education>.
- O'Keefe, P., 2004. The special school in an age of inclusion. Closing address delivered at the sixteenth annual conference on special education, IATSE, St. Patrick's college, Drumcondra, Dublin, June 17–19, 2004. *REACH J. Spec. Needs Educ. Ireland* 18 (1), 3–15.

- Pather, S., Slee, R. (Eds.), 2019. Challenging Inclusive Education Policy and Practice in Africa. Brill Sense, Leiden.
- Picketty, T., 2022. A Brief History of Equality. The Belknap Press of Harvard University Press, Cambridge Mass.
- Polk, K., 1984. The new marginal youth. *Crime Delinquen*. 30 (3), 462–480.
- Ravitch, D., 2014. Reign of Error: The Hoax of the Privatisation Movement and the Danger to America's Public Schools. Vintage Books, New York.
- Ravitch, D., 2020. Slaying Goliath. The Passionate Resistance to Privatisation and the Fight to Save America's Public Schools. Vintage Books, New York.
- Rix, J., 2015. Must Inclusion Be Special? Rethinking Educational Support Within a Community of Provision. Routledge, Abingdon.
- Rizvi, F., 2013. Equity & marketisation: a brief commentary, discourse. *Stud. Cult. Pol. Educ.* 34 (2). <https://www.tandfonline.com/doi/abs/10.1080/01596306.2013.770252>.
- Rizvi, F., Kemmis, S., 1987. Dilemmas of Reform. Geelong: Deakin Studies in Education. Deakin University.
- Rogan, T., 2017. The moral economists. In: Tawney, R.H., Karl Polanyi, E.P. (Eds.), *Thompson and the Critique of Capitalism*. Princeton University Press, Woodstock.
- Rose, N., 2007. The Politics of Life Itself. Biomedicine, Power, and Subjectivity in the Twenty-First Century. Princeton University Press, Princeton N.J.
- Royal Commission into Violence, Abuse, Neglect and Exploitation of people With Disability, 2019. Education & Learning. Issues Paper, 30th October, Canberra. <https://disability.royalcommission.gov.au/system/files/2020-07/Issues-paper-Education-Learning.pdf>.
- Royal Commission into Violence, Abuse, Neglect and Exploitation of people with Disability, 2020. Inclusive Education in Queensland—Preliminary Inquiry, Public Hearing Report. <https://disability.royalcommission.gov.au/system/files/2020-10/Report%20-%20Public%20hearing%20-%20-%20Inclusive%20education%20in%20Queensland%20-%20Preliminary%20inquiry.pdf>.
- Said, E.W., 2000. Travelling theory reconsidered. In: Said, E.W. (Ed.), *Reflections on Exile and Other Literary and Cultural Essays*. Granta Publications, London, pp. 436–452.
- Sandel, M.J., 2020. The Tyranny of Merit. What's Become of the Common Good? Allen Lane, New York.
- Save the Children, 2014. Save the Children Stands for Inclusive Education. Save the Children, London. <https://www.savethechildren.org/content/dam/global/reports/education-and-child-protection/incl-ed-pos-ppr.pdf>.
- Sengeh, D., Winthrop, R., 2022. Why We Must Transform Our Education Systems, Now. Education Plus Development. June 23rd. Brookings, Washington, D.C. <https://www.brookings.edu/blog/education-plus-development/2022/06/23/why-we-must-transform-our-education-systems-now/>.
- Shaddock, A., Packer, S., Roy, A., 2016. Schools for All Children and Young People. Report of the Expert Panel on Students With Complex Needs and Challenging Behaviour. Australian Capital Territory Department for Education, Canberra.
- Slee, R., 2011. The Irregular School. Routledge, Abingdon.
- Slee, R., 2018a. Inclusive Education Isn't Dead, It Just Smells Funny. Routledge, London.
- Slee, R., 2018b. Defining the Scope of Inclusive Education. Background Paper for the 2020 Global Education Monitoring Report. UNESCO, Paris.
- Streek, W., 2016. How Will Capitalism End? Verso, London.
- Tait, G., 2010. Philosophy, Behaviour Disorders and the School. Sense Publishers, Rotterdam.
- Tait, G., 2018. Making Sense of Mass Education, third ed. Cambridge University Press, Melbourne.
- Teece, R., 2013. Academic Success and Social Power: Examinations and Inequality. Scholarly Press, North Melbourne.
- The World Bank, 2020. Pivoting to Inclusion: Leveraging Lessons From the COVID-19 Crisis for Learners With Disabilities. World Bank Group, Washington, D.C. <https://thedocs.worldbank.org/en/doc/147471595907235497-0090022020/original/ElIssuesPaperDisabilityInclusiveEducationFINALACCESSIBLE.pdf>.
- Thomas, G., Loxley, A., 2022. Deconstructing Special Education and Constructing Inclusion. Open University Press, London.
- Tobin, R.M., House, A.E., 2016. DSM-V Diagnosis in the Schools. The Guildford Press, New York, NY.
- Tomlinson, S., 1982. A Sociology of Special Education. Routledge & Kegan Paul, London.
- Tomlinson, S., 1985. The expansion of special education. *Oxf. Rev. Educ.* 11 (2), 157–165.
- Tomlinson, S., 2017. A sociology of special and inclusive education. Exploring the Manufacture of Inability. Routledge, Abingdon.
- Tomlinson, S., 2013. Ignorant Yobs: Low Attainers in a Global Knowledge Economy, London: Routledge.
- Troyna, B., 1993. Racism & Education: Research Perspectives. Open University Press, Maidenhead.
- UNESCO, 1990. World Declaration on Education for All: Framework for Action to Meet Basic Learning Needs. World Conference on Education for All: Meeting Basic Learning Needs, Jomtien, Thailand, March 1990. UNESCO, Paris.
- UNESCO, 1994. Final Report. World Conference on Special Needs Education: Access and Quality. Salamanca Statement and Framework for Action on Special Needs Education. UNESCO, Paris.
- UNESCO, 2019. New methodology shows that 258 million children, adolescents and youth are out of school. Fact Sheet Number 56, September 2019, Paris: UNESCO.
- UNESCO, 2020. Global Education Monitoring Report. Inclusion and Education: All Means All. UNESCO, Paris. <https://en.unesco.org/gem-report/report/2020/inclusion>.
- UNESCO, 2022. Reaching Out to All Learners: A Resource Pack for Supporting Inclusion and Equity in Education. <https://www.inclusive-education-in-action.org/resources/reaching-out-all-learners-resource-pack-supporting-inclusion-and-equity-education>.
- United Nations, 1989. United Nations Convention on the Rights of the Child, (UNCRC). Resolution 44/25 of 20th November. United Nations, New York (see: Articles 23, 28, 29). https://www.unicef.org/uk/wpcontent/uploads/2010/05/UNCRC_united_nations_convention_on_the_rights_of_the_child.pdf.
- Walker, P., 2019. South Australia's Special School Co-location Reform: Inclusion or Illusion? Doctor of Education Thesis, College of Education, Psychology and Social Work. Flinders University, Adelaide. <https://theses.flinders.edu.au/view/ada8f843-0d1c-47e5-905d-27b48ded9b78/1>.
- Wallace, S. (Ed.), 2009. A Dictionary of Education, first ed. Oxford University Press, Oxford. Online Edition 2014. <https://www.oxfordreference.com/view/10.1093/acref/9780199212064.001.0001/acref-9780199212064>.
- Whitty, G., 1985. Sociology and School Knowledge. Methuen, London.
- Wilkinson, R. G. and K. Pickett., 2010. The spirit level: why equality is better for everyone. London, Penguin Books.
- Young, M.F.D., 1971. An approach to the study of curricula as socially Organized Knowledge. In: Young, M.F.D. (Ed.), *Knowledge and Control. New Directions for the Sociology of Education*. Collier Macmillan, London, pp. 19–46.

FOURTH EDITION

International Encyclopedia of Education

Literacies and Languages

VOLUME
10

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

David Yaden
Theresa Rogers



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver,
Canada, University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 10

Literacies and Languages

VOLUME 10 EDITORS

David Yaden and Theresa Rogers



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

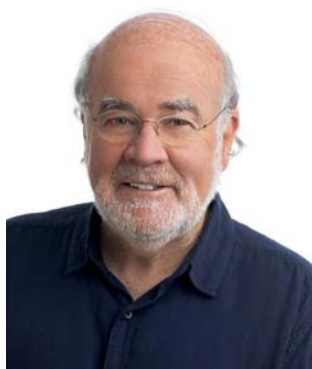
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



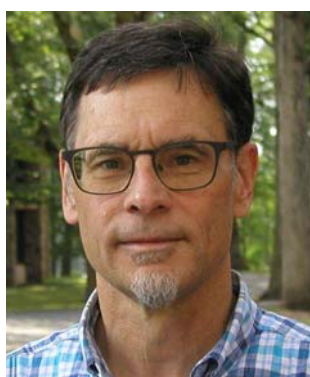
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 10: Literacies and Languages Education

David Yaden and Theresa Rogers

© 2023 Elsevier Ltd. All rights reserved.

Introduction

The study of literacy is not only an educational matter, but a broadly shared field in which multiple disciplines have laid legitimate claims to the functions and/or consequences of reading and writing beyond any pedagogical aspects. For example, historians Patricia Limerick (1987) and Richard Slotkin (1973) have provided vivid and graphic accounts of how the pamphlets, flyers, and popular literacy of the early American West played key roles in what Worster et al. (1989) have termed the “lingering injustice that a conquering people did, and are still doing, to the resident native peoples and ethnic minorities” (p. 305). And in a similar vein, Edward Said (1979, 1994), literary theorist and cultural critic, makes a compelling case that such literary and musical masterpieces of the 19th century as Joseph Conrad’s (1899) *Heart of Darkness*, Jane Austen’s *Mansfield Park* (1814), and Giuseppe Verdi’s opera *Aida* (1871) were key cultural supports to the promotion of racism and the colonizing interests of Belgium, Britain, and France.

In addition, scholars in disciplines such as the Classics have also added illuminating insights into the structure of Homer’s *Iliad* and *Odyssey* (Havelock, 1963, 1988; Ong, 1982; Mitchell, 2011) as a production of oral poetry, not written, discovering as well the complex, oral narrations of early 20th-century poets in Eastern Europe as well as storytellers in most traditional societies including the songs of slaves and rural preachers in the American south (Kanigel, 2021; Lord, 1960/2000; Parry, 1987). Such is the primacy of oral poetic traditions that Alfred Lord, heir to Milman Parry’s legacy, makes the following observation:

We realize that what is called oral tradition is as intricate and meaningful an art form as its derivative “literary tradition.” In the extended sense of the word, oral tradition is as “literary” as literary tradition. It is not simply a less polished, more haphazard, or cruder second cousin twice removed, to literature. By the time, the written techniques come onto the stage, the art forms have been long set and are already highly developed and ancient.

Lord (1960/2000, p. 141)

There is also the perennial debate on the origins of written systems, whether they emerged from the indexing of commercial exchanges of goods and services or from the processes of communication (Daniels, 2017; Robinson, 2007; Sampson, 2015). Emanating initially within the field of sociology, there has also been an extensive and ongoing debate particularly since the publication of Goody’s and Watt’s (1963) seminal piece, “The consequences of literacy,” on whether populations with alphabets, in particular, have a different way of thinking, a more analytic rationality, about the world than those with other scripts or people who have not learned to read (Goody, 1996; Olson, 1996, 2016; Ong, 1982) in any written language. And in his classic *Tristes Tropiques*, anthropologist Levi-Strauss (1955) made the grim claim from his observation of the reaction to written language by members of traditional societies that writing “seems to have favoured the exploitation of human beings rather than their enlightenment,” and, finally, that “my hypothesis, if correct, would oblige us to recognize the fact that the primary function of written communication is to facilitate slavery (p. 299)” as tribal chieftains understood that by mimicking facility with such a tool came additional authority and power.

Goals of the Literacies and Languages Education Volume

Suffice it to say, that any compilation of literacy and languages research such as the present volume contains must necessarily be delimited, although we found the creation of our own arbitrary, liminal boundaries extremely difficult to impose. Thus, it bears noting how the chapters were assembled for the Literacies and Languages education volume. We first took to heart the Vision and Goal statements by the Editors-in-Chief (EIC) that this 4th edition of the *International Encyclopedia of Education* would reflect “a deeper critical engagement with issues of globalization, internationalization and cross-cultural matters.” Therefore, we attempted to address this charge by (a) specifically searching for literacy and languages investigations which touched on one or more of these three themes (i.e., globalization, internationalization and cross-cultural matters); (b) requesting that authors, both in the initial invitation and in review comments, include citations and additional elaborations of literacy research outside the country in which they conducted the majority of their own research; and (c) by identifying authors, particularly in the Global South, who were addressing various issues in learning to write and read in languages other than English. While this latter tack wasn’t as successful as we would have liked, we are encouraged to have several scholars from Mexico, Central and South America, Africa, the Middle East, and China who submitted chapters. Overall, we feel that nearly every chapter, even those reporting on English-language studies, extends and connects the literacy research discussed with other related investigations across the globe.

Another goal of the EIC team for the volume was to encompass a broad range of studies, including both past topics represented in the 3rd edition as well as new and emerging areas of research over the last decade in particular. In the prior third edition (Peterson et al., 2010), studies in reading and writing were integrated under the topic of “Curriculum Development,” and constituted approximately 30 chapters addressing such topics as adult literacy, early language and literacy acquisition, bilingualism, second- and

foreign language learning, reading and technology, literacy issues among exceptional populations, reading and writing assessment in schools, neuroscience and reading, and writing.

In the present volume, the 94 chapters not only expand and update research on these familiar topics, but extend the range of literacy studies across many other areas, including studies in the digital and multimodal aspects of reading and writing, artificial intelligence, critical literacies, race, gender and sexuality, cosmopolitan and international literacies, Indigenous literacies, World Englishes, children's literature development in several Global North and South countries, and several theoretical perspectives on literacy, including posthumanist, sociocultural, developmental, humanities-oriented, and some general epistemological aspects which underlie theorizing in reading and writing. Thus, in this 4th edition of the *Encyclopedia* there is a much wider scope of literacy studies represented, which is no surprise. But even so, as mentioned earlier, the full range of disciplines which lay claim to literacy studies in some aspect cannot practically be taken into consideration.

A final goal which we attempted to address in soliciting authors for the volume was that the chapters were to be critically oriented, forward-looking, dynamic, and serving as a "springboard for advances dealing with new formations, new arguments, new theories and offering future directions." Since the *Encyclopedia's* readership is broadly interdisciplinary, international, and inclusive of both emerging and established scholars, all authors were encouraged to not only describe the developments in their subfield of literacy study, but to hypothesize, speculate, and prognosticate about what future developments are needed for the field to thrive and offer creative solutions to problems or issues not easily resolved. Specifically, in the review process, authors were asked to unpack and elaborate on statements simply indicating that more research needed to be done. They were asked to identify promising programs or approaches as exemplars, and to offer a critical appraisal of why progress may or may not be happening currently, and what needed to be done to overcome these challenges. As volume editors, we are hopeful that this volume represents indeed "a springboard" toward literacy research which is more equitable, inclusive, and an accurate, credible representation of the populations from which information is obtained such that the outcome of such research truly enhances the well-being of the world's peoples.

We need to mention as well that the COVID-19 pandemic, raging within these last nearly three years, had a significant impact on authors' ability to accept an invitation to write an entry, and, therefore, some topics had to be dropped due to the many challenges to personal and professional lives that forced some authors to withdraw even after the writing had begun. These extenuating circumstances often resulted in multiple declinations by respected authors known for their expertise in those areas, some of which included linguistic landscapes, trans*+ centered literacy study, the reading of augmented texts, writing practices among global majority youth, 3rd space approaches to literacy research, subaltern literacies, resistance writing, and public-facing literacy scholarship, just to name a few. And, sadly, during the past two years of this volume's development, three prominent literacy researchers passed—Vivian Yenika-Agbaw, a scholar of African children's literature, Barbara Guzzetti, a widely-known expert in literacy and gender studies, and David O'Brien, a prominent researcher in the field of disciplinary literacies. We are thankful, however, for the coauthors on those manuscripts who completed the submissions as a testament to their colleagues and mentors so that their voices would be represented in the *Encyclopedia*.

Themes and Methodological Trends

The fact that we noted some trends in research which can be recognized in the *Literacies and Languages Education* volume does not mean that there are any universal, or "grand narratives," as it were, to be told. If anything, within each "theme" or strand of connected conceptual ideas, methodological approaches, or theoretical positions, there are contradictions and disjunctions that are apparent as well.

For example, for the most part, what can be detected in the chapters across the labels of the major sections is what Frederick Erickson (2018) has called a "critical realist" approach to the ontology and epistemology (p. 58) of literacy research. Erickson goes on to suggest that this approach is akin to what Bent Flyvbjerg (2001) has termed, drawing from Aristotle, *phronesis*:

Phronesis goes beyond both analytical, scientific knowledge (*episteme*) and technical knowledge or know-how (*techne*) and involves judgments and decisions made in the manner of a virtuoso social and political actor. I will argue that *phronesis* is commonly involved in social practice, and that therefore attempts to reduce social science and theory either to *episteme* or *techne* or to comprehend them in those terms, are misguided (p. 2).

Thus, the *Encyclopedia's* readers will find across a great number of the chapters an emphasis on action-oriented language and literacy knowledge that is highly local, situated, creative, dynamic, and spatialized according to the particular cultural norms, individual experiences, and contextual settings of the research population. There are few, if any, calls for a "one-size-fits-all" approach to literacy learning.

On the other hand, within this research emphasis to gain a deeply individual and contextualized knowledge of the influences upon literacy growth, there is also a nuanced emphasis on "collectivity"—in other words, that literacy acquisition, development, and teaching is an enterprise that calls for the cooperation and networking of numerous persons as actors in any literacy environment. Unlike a general positivist approach that attempts to tease out causes within what Cronbach and Meehl (1955) termed as a "nomological network," the critical realist orientation links all the actors in the network together and attempts to understand the proportionate contributions of each, each playing pivotal roles at various times. Like Flyvbjerg's virtuoso social actor, all participants or collaborators in a research study are conducting their own virtuoso performances simultaneously, and the greater burden

on the researcher is to understand how all these performances—including their own—are inextricably connected to produce the conditions in which the literacy event, act or phenomenon is being observed or investigated. Thus, overall, the chapters in this volume generally adhere to Vygotsky's (1931/1997) notion that "To study something historically means to study it in motion ... for only in movement does a body exhibit what it is." Static causes, then, will seldom be identified.

This orientation to literacy research largely explains why the majority of the chapters report on research that is qualitative or "interpretive" in nature (Erickson, 2018). However, this situation is not because invitations were not extended to researchers adopting experimental and quantitative approaches, because they were, but were declined for various reasons. It is not the case either that we specified to invited authors that there were certain kinds of methodologies that would be preferred. Each author had full autonomy related to the research they would report in their areas of expertise. While we would have liked better representation from quantitatively-oriented investigators, we suspect that this didn't happen because within the last decade the literacy field in particular has been substantively influenced by the overall trends in social science inquiry toward poststructural, posthumanist, participatory, critical, and nonrepresentational theoretical research frameworks (Lincoln et al., 2018; Unrau et al., 2019). Indeed, the term "bricoleur," derived from Levi-Strauss' (1966) description of the construction of mythical thought in traditional societies, and reintroduced into modern research parlance as a person who is simultaneously "scientist, naturalist, fieldworker, journalist, social critic, artist, performer, jazz musician, filmmaker, quilt maker, and essayist" (Denzin and Lincoln, 2018, p. 11), is a persona which has been taken on by many present-day literacy researchers, a fair number of whom are represented in this volume.

Section Overviews

We have divided the chapters of the volume into seven major sections, although we candidly acknowledge that these sections, as well as their labels, have fuzzy edges, and the research represented by the chapters so designated could have been categorized potentially in any number of ways. The chapters in the volume under each section heading are not so much organized by title or even subtopic, but more by common motifs or loosely woven themes which occur repeatedly across most of the entries in the section. That said, brief summaries of each major section are offered below.

Acquiring, Developing, Learning Languages/Languaging, and Literacy

The first section includes research that focuses essentially upon the processes and diverse contextual situations of language and literacy acquisition and subsequent development whether it relates to very young children's first encounters with their heritage languages and written forms or older learners' in and out of school experiences with first, second, or foreign languages. Across all entries in this section are three notions. First, that language learning is, in the words of one author, always "dynamic, complex, polyphonic, and shifting across time and space." Whether it is in the acquisition of a foreign language, a second language, or negotiating languages and literacies across borders, all language users, beginning, developing, or experienced, operate indeed like "virtuosos," orchestrating their language and literacy resources in response to linguistic settings which are, in turn, fluctuating themselves.

Another key aspect of the chapters in this section is the emphasis the "criticality" of language users, both in everyday usage and learning situations. Every encountered literacy or language event or act requires judgments and decision-making given the situational parameters and power dynamics. In other words, as Goffman (1967) has described, language users are continually involved in "face-work," acting out a *line*—"that is, a pattern of verbal and nonverbal acts by which he expresses his view of the situation, and through this his evaluation of the participants, especially himself" (p. 5). Finally, it becomes paramount, then, as is also addressed in this section, that any assessment or assessor(s) of language or literacy use must be cognizant of the multiple dimensions of performance embedded in any linguistic event, so that judgments of that performance may be accurate, credible, and reflective of the language and literacy strengths of the users.

Cultural, Critical, and Community Literacies

Some of the key themes which run through the 15 chapters in this section include (a) the systemic and institutional oppression, racism and educational barriers that are encountered by particularly Black, Indigenous, and other People Of Color (BIPOC), members of the LGBTQIA+ community, and persons in exceptional populations both in and out of school settings; (b) that learning is far beyond just a cognitive function, but is deeply embedded in the racial, cultural, historical, socioeconomic, and gendered identities of language and literacy users; (c) an emphasis upon the asset-based, multimodal linguistic resources of family and community from which all learners draw; and (d) the ongoing quest for learners and educators alike to develop a critical consciousness and action-oriented behaviors for achieving social justice, inclusion and equity in school classrooms and the broader society.

Touching on topics such as border and family literacies, gender and literacy, federal and state language, and literacy policies, the multimodality of literacy learning across various in and out of school settings, these papers stress simultaneously both the historical oppression of marginalized populations, while at the same time pointing out the notable resistance to these structural and societal barriers through access to personal, family, and community resources as well as the extensive linguistic resources derived from biliteracy and multilingual abilities. It is also abundantly clear from these articles that our current educational systems, both curricular and evaluative, are woefully unequipped at the present time to accommodate the richness of what young learners, youth, and their

families possess, squeezing them into traditional educational molds and ideologies which strip away linguistic and personal resources which otherwise could be used for the benefit of all. The chapters challenge readers to adopt a liberatory, and transformative attitude to achieve more equitable systems.

Literacy Development and Pedagogy Among Children, Youth, and Adults.

This section encompasses a range of theoretical frames that provide multiple perspectives on language and literacy development and pedagogy among children, youth, and adults. These frames include cognitive, linguistic, sociocognitive, constructivist, sociocultural, creative, and critical perspectives, with several drawing from multiple frameworks. A throughline in many of the pieces is an epistemology of language and literacy practices as contextualized in the social, cultural, and material conditions of their uses, and the role of language and literacy in addressing and responding to the lived experiences and challenges of our contemporary highly connected and global world. These language and literacy practices include engaging with a range of texts meaningfully, imaginatively, collectively, and critically.

Attention to learning differently is another hallmark of this section. While we all learn, in some measure, in different ways, here the needs of exceptional learners who struggle with literacy development, have learning challenges, are autistic, or who read braille are considered, as is the importance of fostering social connectedness and a sense of competence. Overall, many of these authors advocate for shifting away from notions of language and literacy as decontextualized or from standardizing practices toward providing rich, contextual and creative approaches to pedagogy—what Edmiston (this volume) calls a more humanizing and transformative approach to teaching, learning, and evaluating in the schooling process.

Digital and Multimodal Literacies

Language and literacy are inherently multimodal practices—whether we are gesturing as we speak or using a pen to write; however, in the past several decades scholars have been paying increasing attention to the role and range of multimodalities in language and literacy practices. At the same time, educators have championed the value of integrating various modes in literacy teaching and learning pedagogies. This is largely due to the advent of connected digital technologies that have extended the access to and affordances of multimodal texts and practices. Several of these chapters speak broadly to theories, scholarship and practices of multimodal and digital language and literacy practices such as (a) theorizing a sensory literacies approach that emphasizes the vital roles of the senses, the body, and real-world referents across digital and non-digital practices; (b) defining the contours of a trans-literacy framework to understand how digital media is at the center of identity formation; (c) enlisting notions of social semiotics and transmediation to help unpack social and cultural meanings across modalities, and (d) conceptualizing digital literacies as “electracy”—as technologies intertwined with cultural and institutional formation and collective identities of a society in transition.

Many of these chapters then extend these conceptualizations of digital multimodal tools to explore what they offer in terms of the modes, spaces and reach of language and literacy practices, as well as the ways this rich variety of forms foster the complexity of meaning-making and entail play, creativity, criticality, and participatory practices. Specific topics range from reading with e-books, AI machine learning, and reading online to the many ways media and digital tools enhance engagement among reluctant readers, extend writing practices, support the production of creative and critical media texts, and engender global literacies and curricula.

Children’s Literature: An International Perspective

Children’s literature in all its forms remains a vital aspect of the literacies and literacy education of children and young adults. While various forms of storytelling are as old as cultures, children’s and young adult literature comprises a relatively new classification of literary works, having developed with the advent of the conception of childhood itself as distinct from adulthood, and has proven to be truly robust both in its traditional forms and highly adaptable to current technologies. The chapters here focus on contemporary critical perspectives of literature for young people, valuing it as a rich medium for critically reimagining lives and worlds. The authors speak to the compendium of new forms of literature that range from the metafictional to the digital, and carefully delineate the global world of children’s and young adult literature while considering the particular traditions across countries in Latin America, Asia, Southern Europe, the Middle East, and Africa.

As Peter Hunt (1999) has noted, “children’s books are part of the ideological structures and cultures of the world” (p. 4) and indeed these chapters illuminate how access, forms and topics of children’s and young adult literature, as well as the critical theoretical lenses that are taken up in related scholarly work, are deeply entwined with historical and political developments such as colonialism, shifting political philosophies and revolutions, and language policies and practices. A glimpse into the literature for young people across the globe reminds us of the variety and unique traditions, richness and purposes of this work that are too often overshadowed by the Western canon.

Literacies in the Global North and South

While local politics, language policies, traditional educational structures have a substantive impact within any one school, neighborhood or community—which is complicated enough to figure out, the entries in this section elevate some of these same forces of influence to a global scale with frighteningly global consequences. Inescapable in majority of the articles are trenchant

discussions of the historical and negative influences of colonialism as well as how ongoing imperialistic tendencies are still pervasive in many spheres, including inequitable access to educational resources, the imposition of official languages as the language of instruction in local areas, and the underresourced infrastructure of technological advances in many communities. It is made clear in these entries that the languages and literate practices of Indigenous people in every country are the most vulnerable to extinction, being threatened with ongoing oppression and even the intent of erasure.

Thus, there is a uniform clarion call yet again (see Brandt, 1980) in these chapters for all countries in the Global North and South to embrace a humanizing pluralism to dismantle colonial thinking and structures and to recognize the intellectual strengths and linguistic resources of the world's Indigenous peoples, not only as local solutions, but as a strategy for improving the well-being of all populations. These efforts include facilitating language recovery and restoration, eliminating economic insecurities, valorizing endogenous languages to deliver educational content, and working to eliminate conditions which exacerbate violence and injury.

Another pervasive theme recognized in nearly every chapter in this section is the historical and unwavering resistance to dominance, the reservoirs of deep spiritual strengths, the adaptive ability of Indigenous communities to maintain their own solidarity, and the will to create meaningful “third-space” literacy learning, thus, unrelentingly advocating for themselves and demanding that their voices play a role in solutions toward educational equity and improvement.

Contemporary and Emerging Theoretical and Methodological Issues in Literacy Research

The chapters in this final section embrace a range of theoretical and methodological perspectives from psycho-sociolinguistic, neurolinguistic, and developmental to broader epistemological issues, theories of embodiment, experimental humanities as well as postqualitative, postmodern, and poststructural inquiry. It is fair to say, however, that even models with roots in older psycho- or sociolinguistic theories emanating from fields such as psychology, anthropology and sociology now incorporate more dynamical features, drawing from findings in the neuro- and developmental sciences and evolutionary biology. These theories all stress the organicity of literacy growth as a species-specific, holistic, and intersectional adaption to ever-shifting, complex environmental influences.

Another group of authors in this section explore some of the epistemological shifts in literacy research that bend away from reductionism and the expert, omniscient observer such as existed in the “Golden Age” of anthropology (see Erickson, 2018) to incorporate the interleaving of researcher, the researched, the literacy space, and the dynamic social fabric within which any literacy event or act takes place. These chapters highlight the essential tension, as it were, between deriving an “explanation” from a study's findings to accepting some tolerable ambiguity as is characteristic of much social science research.

A final set of entries in this section place literacy research fully within the framework of a postqualitative or posthumanist theoretical frame. In this perspective, there are no definable “edges,” to literacy phenomena since they occur as an intersectional function of the individual, the community, the social milieu, the real or imagined spaces and materials within any setting—a “rhizomatic” structure for some. In this view, the study of “literacy” is almost a misnomer since any act of communication takes place within the full sensorium of the human experience as it is situationally spatialized. Researchers with this perspective do not so much “study” literacy as participate in its very construction with others.

Conclusion

The nearly 100 chapters in this volume are a fair sampling, we believe, of the research into various literacies and their related language foundations since the publication of the third edition of the *International Encyclopedia of Education* (Peterson et al., 2010). It is also the case that the entries in this new fourth edition provides a greatly increased range and scope of literacy studies, not only updating and enriching some long-researched areas of interest (e.g., children's literature, comprehension, vocabulary, spelling, bilingualism, second-language learning, family literacy), but broadening the field of language and literacy investigations to encompass very 21st century concerns in education such as the exploration of gendered literacy identities across all literacy environments, the spacialized and sociocultural impacts on learning, the injustices of inequitable language and literacy resources worldwide for marginalized populations, and the institutional and structured racism of language and literacy policies, both domestic and international, which continue to erect barriers preventing many Indigenous populations and members of the global majority from obtaining educational benefits. It must also be said, however, that the volume does not fail to highlight the deep intellectual resources and “thrivance” of Indigenous and oppressed peoples in every country, and the peril involved in ignoring their voices.

Finally, as editors of the Literacies and Languages Education volume, we have had the immense privilege of interacting with and reading the work of an international cadre of collegial and seminal scholars, immersing ourselves in their highly stimulating, thoughtful, and compelling research which has, in turn, enriched our own learning and work. We trust that the readership of this volume will profit from this remarkable compendium of chapters as we have.

References

- Brandt, W., 1980. North-South: A Program for Survival. The MIT Press, Cambridge, MA.
- Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. *Psychol. Bull.* 52 (4), 281–302.

- Daniels, P.T., 2017. *An Exploration of Writing*. Equinox Publishing, Sheffield, UK.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2018. *The Sage Handbook of Qualitative Research*, fifth ed. Sage Publications, Inc, Los Angeles, CA, p. 11.
- Erickson, F., 2018. A history of qualitative inquiry in social and educational research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*, fifth ed. Sage Publications, Inc, Los Angeles, CA, pp. 36–65.
- Flyvbjerg, B., 2001. *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*. Cambridge University Press, Cambridge, UK.
- Goffman, E., 1967. *Interaction Ritual*. Anchor Books, Garden City, NY.
- Goody, J.R., 1996. *The East in the West*. Cambridge University Press, Cambridge, UK.
- Goody, J.R., Watt, I., 1963. The consequences of literacy. *Comparative Studies in Society and History* 5 (3), 304–345. <https://doi.org/10.1017/S0010417500001730>.
- Havelock, E.A., 1963. *Preface to Plato*. The Belknap press, Cambridge, MA.
- Havelock, E.A., 1988. *The Muse Learns to Write: Reflections on Orality and Literacy From Antiquities to the Present*. Yale University Press, New Haven, CT.
- Homer, 2011. *The Iliad* (S. Mitchell trans.). Simon & Schuster, New York, NY.
- Hunt, P., 1999. *Understanding Children's Literature*, second ed. Routledge, New York, NY.
- Kanigel, R., 2021. *Hearing Homer's Song: The Brief Life and Big Idea of Milman Parry*. Alfred A. Knopf, New York, NY.
- Lévi-Strauss, C., 1955. *Tristes Tropiques* (J. Weightman, D. Weightman, Trans.). Penguin Group, New York, NY.
- Limerick, P., 1987. *The Legacy of Conquest: The Unbroken Past of the American West*. W.W. Norton & Co, New York, NY.
- Lincoln, Y.S., Lynham, S.A., Guba, E.G., 2018. Paradigmatic controversies, contradictions, and emerging confluences revisited. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*, fifth ed. Sage Publications, Inc, Los Angeles, CA, pp. 108–150.
- Lord, A.B., 2000. In: Mitchel, S., Nagy, G. (Eds.), *The Singer of Tales*, second ed. Harvard University Press, Cambridge, MA. (original work published 1960).
- Olson, D.R., 1996. *The World on Paper: Conceptual and Cognitive Implications of Writing and Reading*. Cambridge University Press, Cambridge, UK.
- Olson, D.R., 2016. *The Mind on Paper: Reading, Consciousness and Rationality*. Cambridge University Press, Cambridge, UK.
- Ong, W.J., 1982. *Orality and Literacy: The Technologizing of the Word*. Methuen Publishing, London, UK.
- Parry, A. (Ed.), 1987. *The Making of Homeric Verse: The Collected Papers of Milman Parry*. Oxford University Press, New York, NY.
- Peterson, P.L., Baker, E., McGaw, B., 2010. *International Encyclopedia of Education*, third ed. Elsevier Science, Oxford, UK.
- Robinson, A., 2007. *The Story of Writing: Alphabets, Hieroglyphics and Pictograms*, second ed. Thames & Hudson, London, UK.
- Said, E.S., 1979. *Orientalism*. Vintage Books, New York, NY.
- Said, E.S., 1994. *Culture and Imperialism*. Vintage Books, New York, NY.
- Sampson, G., 2015. *Writing Systems*, second ed. Equinox Publishing, Sheffield, UK.
- Slotkin, 1973. *Regeneration Through Violence: The Mythology of the American Frontier, 1600–1860*. University of Oklahoma Press, Norman, OK.
- Unrau, N.J., Alvermann, D.E., Sailors, M., 2019. Literacies and their investigations through theories and models. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, NY, pp. 3–34.
- Vygotsky, L.S., 1997. The prehistory of written language (M. J.Hall, Trans.). In: Rieber, R. (Ed.), *The Collected Works of L. S. Vygotsky*, vol. 4: *The History of the Development of Higher Mental Functions*. Springer, Boston, MA, pp. 131–148 (Originally published 1931).
- Worster, D., Armitage, S., Malone, M.P., Weber, D.J., Limerick, P.N., 1989. "The legacy of conquest," by Patricia Nelson Limerick: a panel of appraisal. *Western Hist. Q.* 20 (3), 323–330.

CONTRIBUTORS TO VOLUME 10

Ashraf Abdelhay

Doha Institute for Graduate Studies, School of Social Sciences and Humanities, Doha, Qatar

Jamal Abedi

University of California, UC Davis School of Education, Davis, CA, United States

Onyedikachi Abiodun-Ekus

Faculty of Education, University of Reading, Reading, United Kingdom

Susanne K Abou Ghaida

University of Glasgow, Glasgow, United Kingdom

Joseph Onyema Ahaotu

English Language Teachers Association of Nigeria, Department of English Studies, University of Port Harcourt, Port Harcourt, Rivers, Nigeria

Richard L Allington

University of Tennessee, Knoxville, TN, United States; and University of Connecticut, Storrs, CT, United States

Ghaida S Alrawashdeh

College of Education, University of Illinois at Urbana-Champaign, Champaign, IL, United States

Lilian Alvarez-Arellano

Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico

Donna E Alvermann

University of Georgia, Athens, GA USA

Jocelyn Amevuvor

Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, United States

Ann Anderson

Department of Curriculum and Pedagogy, University of British Columbia, Vancouver, BC, Canada

Jim Anderson

Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Richard C Anderson

Center for the Study of Reading, University of Illinois at Urbana-Champaign, Champaign, IL, United States

Ryosuke Aoyama

Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Sonya L Armstrong

Department of Curriculum and Instruction, Texas State University, San Marcos, TX, United States

Laura E Arrington

University of Louisiana at Lafayette, Lafayette, LA, United States

Alfredo J Artiles

Stanford University, Stanford, CA, United States

Betsy A Baker

University of Missouri, Columbia, MO, United States

Sandra Barrueco

Department of Psychology, The Catholic University of America, Washington, DC, United States

Richard Beach

University of Minnesota, Minneapolis, MN, United States

Thomas W Bean

Old Dominion University, Darden College of Education and Professional Studies, Department of Teaching and Learning, Norfolk, VA, United States

Jared Bernstein

Analytic Measures Inc., Palo Alto, CA, United States

Marcelo Emilio Bianchi-Bustos

Academia de Literatura Infantil, Buenos Aires, Argentina

Mollie V Blackburn

The Ohio State University, Columbus, OH, United States

Nicolas Blaisdell
University of Massachusetts, Amherst, MA, United States

David Bloome
The Ohio State University, Department of Teaching and Learning, Columbus, OH, United States

Danielle Burrell-Kim
University of British Columbia (UBC), Vancouver, BC, Canada

Luzkarime Calle-Díaz
Universidad del Norte, Barranquilla, Atlántico, Colombia

Brian L Cambourne
University of Wollongong, Worring Heights, NSW, Australia

Gerald Campano
University of Pennsylvania, Philadelphia, PA, United States

Michelle Cannon
The UCL Knowledge Lab, Institute of Education, University of London, London, United Kingdom

Jill Castek
Teaching, Learning and Sociocultural Studies, University of Arizona, College of Education, Tucson, AZ, United States

Nathan M Castillo
College of Education, University of Illinois at Urbana-Champaign, Champaign, IL, United States

Gina N Cervetti
Educational Studies, School of Education, University of Michigan, Ann Arbor, MI, United States

Eve V Clark
Stanford University, Stanford, CA, United States

Jamie Colwell
Old Dominion University, Norfolk, VA, United States

Barbara Comber
University of South Australia, Education Futures, Adelaide, SA, Australia

Jonathan Jae-an Crisman
University of Arizona, Tucson, AZ, United States

Debra K Crouch
Educational Consultant, San Diego, CA, United States

Aijuan Cun
University of New Mexico, Albuquerque, NM, United States

James W Cunningham
University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Jen Scott Curwood
The University of Sydney, Camperdown NSW Australia

Holly Damico
University of Louisiana at Lafayette, Lafayette, LA, United States

Jack S Damico
University of Colorado Boulder, Boulder, CO, United States

Amanda Deliman
School of Teacher Education and Leadership, Utah State University, Salt Lake City, UT, United States

Fred Dervin
Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

Ryan Deschambault
Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Deborah R Dillon
University of Minnesota, Minneapolis, MN, United States

Beatrice C Dupuy
The University of Arizona, Tucson, AZ, United States

Bernadette Dwyer
Institute of Education, School of Language, Literacy & Early Childhood Education, Dublin City University, Dublin, Ireland

Brian Edmiston
The Ohio State University, Columbus, OH, United States

Patricia A Edwards
Michigan State University, Department of Teacher Education, East Lansing, MI, United States

Lucila D Ek
The University of Texas at San Antonio, San Antonio, TX, United States

Patricia Enciso
The Ohio State University, Columbus, OH, United States

Beryl Exley
School of Education and Professional Studies, Griffith University, Mount Gravatt, QLD, Australia

Narges Firouzshahi
University of Louisiana at Lafayette, Lafayette, LA, United States

Douglas Fisher
San Diego State University, Health Sciences High, San Diego, CA, United States

Elena Forzani
Literacy Education, Boston University, Wheelock College of Education & Human Development, Boston, MA, United States

Divina Frau-Meigs
University Sorbonne Nouvelle, Paris, France

Nancy Frey
San Diego State University, Health Sciences High, San Diego, CA, United States

Karole-Ann Friddle
University of North Carolina-Greensboro, Greensboro, NC, United States

Brittany L Frieson
Teacher Education and Administration, University of North Texas, Denton, TX, United States

María E Fránquiz
Department Curriculum and Instruction, University of Texas at Austin, Austin, TX, United States

Rachael Gabriel
University of Tennessee, Knoxville, TN, United States; and University of Connecticut, Storrs, CT, United States

Candace Kaleimamoowahinekapu Galla
Department of Language & Literacy Education, Faculty of Education, The University of British Columbia, Vancouver, BC, Canada; and Institute for Critical Indigenous Studies, Faculty of Arts, The University of British Columbia, Vancouver, BC, Canada

Andrea Gambino
University of California, Los Angeles, CA, United States

Rabani Garg
University of Pennsylvania, Philadelphia, PA, United States

Meg Gebhard
University of Massachusetts, Amherst, MA, United States

Aanchal Gidra
Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

Perry Gilmore
Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, Tucson, AZ, United States

Kristen Gregory
East Carolina University, College of Education, Department of Elementary and Middle Grades, Greenville, NC, United States

Myriam Jimena Guerra
Texas A&M University-San Antonio, San Antonio, TX, United States

Barbara J Guzzetti
Arizona State University, New College of Interdisciplinary Arts & Sciences, Humanities, Arts & Cultural Studies, Department of English, Phoenix, AZ, United States

Caroline Hamilton-McKenna
Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Colin Harrison
Emeritus Professor of Literacy Studies in Education, University of Nottingham, Nottingham, United Kingdom

Michael Havazelet
Vanderbilt University, Nashville, TN, United States

Walid Hedidar
Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

Donovan Herrera-Santillán
Universidad Nacional Autónoma de México, Mexico City, Mexico

Elfrieda H Hiebert
TextProject, Santa Cruz, CA, United States

Kathleen A Hinchman
Syracuse University, Syracuse, NY, United States

James V Hoffman
Language and Literacy Studies, Department of Teacher Education and Administration, The University of North Texas, Denton, TX, United States

Amanda Holmes
Independent Scholar, United States

Sunggye Hong
University of Arizona, Department of Disability and Psychoeducational Studies, Tucson, AZ, United States

Rosalind Horowitz

International Discourse & Literacy Studies, Culture and Learning Sciences, Departments of Interdisciplinary Learning and Teaching, Educational Psychology, College of Education and Human Development, The University of Texas at San Antonio, San Antonio, TX, United States

George G Hruby

Collaborative Center for Literacy Development, University of Kentucky, Lexington, KY, United States

Patricia Arias Hunt

University of Louisiana at Lafayette, Lafayette, LA, United States

Gay Ivey

University of North Carolina-Greensboro, Greensboro, NC, United States

Erik Jacobson

Montclair State University, Montclair, NJ, United States

Bong Gee Jang

Syracuse University, Syracuse, NY, United States

Hilary Janks

The University of the Witwatersrand, School of Education, Johannesburg, South Africa

Michelle Jeffries

Queensland University of Technology, Faculty of Education, Kelvin Grove, QLD, Australia

Jennifer Jenson

University of British Columbia (UBC), Vancouver, BC, Canada

Korina M Jocson

York University, Toronto, ON, Canada

Latrise P Johnson

The University of Alabama, Tuscaloosa, AL, United States

Miriam Jorge

University of Missouri-St. Louis, St. Louis, MO, United States

Takeshi Kajigaya

Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Wei Kan

Faculty of Education, Beijing Normal University, Beijing, China

Erin Kim-Rich

The University of New England, Armidale NSW Australia

James R King

Literacy Studies, University of South Florida, Tampa, FL, United States

Lara-Stephanie Krause-Alzaidi

Institute of African Studies, Leipzig University, Germany

Beth Krone

Kennesaw State University, Kennesaw, GA, United States

Ryuko Kubota

Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Candace R Kuby

University of Missouri, Columbia, MO, United States

Yuri Kumagai

Smith College, Department of East Asian Languages and Cultures, Northampton, MA, United States

Gilberto P Lara

The University of Texas at San Antonio, San Antonio, TX, United States

Kevin Leander

Vanderbilt University, Nashville, TN, United States

María G Leija

Department of Interdisciplinary Learning and Teaching, University of Texas at San Antonio, San Antonio, TX, United States

Kimberly Lenters

University of Calgary, Calgary, AB, Canada

Anthony J Liddicoat

Department of Applied Linguistics, University of Warwick, Coventry, United Kingdom

Shufeng Ma

Department of Educational Psychology, East China Normal University, Shanghai, China

Sinfree Makoni

Department of Applied Linguistics, Pennsylvania State University, College of the Liberal Arts, University Park, PA, United States

Marla H Mallette

Department of Teaching, Learning, and Educational Leadership, College of Community and Public Affairs, Binghamton University, Vestal, NY, United States

Camille Martinez-Yaden

Department of Teaching, Learning and Sociocultural Studies, University of Arizona, Tucson, AZ, United States

Jillian Mason

Communications Department, Division of Liberal Arts and Education, Arizona Western College, Yuma, AZ, United States

Dom Massaro

Department of Psychology, University of California, Santa Cruz, CA, United States

Matthew McConn

Department of Teaching, Learning, and Educational Leadership, College of Community and Public Affairs, Binghamton University, Vestal, NY, United States

Jonda C McNair

The Ohio State University, Department of Teaching and Learning, Columbus, OH, United States

Mary B McVee

University at Buffalo—SUNY, Buffalo, NY, United States

Carmen Liliana Medina

Department of Literacy, Culture, and Language Education, Indiana University, Bloomington, IN, United States

Kathy A Mills

Institute for Learning Sciences & Teacher Education, Brisbane Campus, Brisbane, QLD, Australia

Ayan Mitra

BrainLENS Laboratory (UCONN-UCSF), Department of Psychological Sciences, Hoeft Laboratory for Learning Engineering and Neural Systems, University of Connecticut, Storrs, CN, United States

Maria C Morales

Department of Psychology, The Catholic University of America, Washington, DC, United States

Ernest Morrell

Department of English, The University of Notre Dame, Notre Dame, IN, United States

Jodene Morrell

Notre Dame Center for Literacy Education, The University of Notre Dame, Notre Dame, IN, United States

Yasmine S Motawy

Department of Rhetoric and Composition, The American University in Cairo, New Cairo, Egypt

Luz A Murillo

Curriculum and Instruction, Texas State University, San Marcos, TX, United States

Joaquin Muñoz

Department of Curriculum & Pedagogy, Faculty of Education, The University of British Columbia, Vancouver, BC, Canada

Ryan L Nelson

University of Louisiana at Lafayette, Lafayette, LA, United States

Bonny Norton

University of British Columbia, Vancouver, BC, Canada

Allison Wynhoff Olsen

Montana State University, Bozeman, MT, United States

Kate Omenugha

Nnamdi Azikiwe University, Awka, Anambra, Nigeria

Chukwuemeka Eze Onukaogu

Reading Association of Nigeria, Ile-Ife, Osun, Nigeria

David G O'Brien

University of Minnesota, Minneapolis, MN, United States

David D Paige

Northern Illinois University, Department of Curriculum & Instruction, DeKalb, IL, United States

Petros Panaou

University of Georgia, Department of Language and Literacy Education, Athens, GA, United States

Sylvia Pantaleo

Language & Literacies, Faculty of Education, University of Victoria, Victoria, BC, Canada

Seth A Parsons

George Mason University, Fairfax, VA, United States

Ashley N Patterson

Penn State University, University Park, PA, United States

P David Pearson

Instructional Science, Graduate School of Education, University of California, Berkeley, CA, United States

Amanda Percle

University of Louisiana at Lafayette, Lafayette, LA, United States

Vivian E Presiado

School of Teaching and Learning, Illinois State University, Normal, IL, United States

Mastin Prinsloo

School of Education, University of Cape Town, Rondebosch, South Africa

Pengtong Qu

Department of Literacy, Culture, and Language Education, Indiana University, Bloomington, IN, United States

Phoebe L Quaynor

Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, United States

Timothy Rasinski

School of Teaching Learning and Curriculum Studies, Kent State University, Kent, OH, United States

Rebecca Rogers

University of Missouri-St. Louis, St. Louis, MO, United States

Theresa Rogers

Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Deborah Wells Rowe

Vanderbilt University, Department of Teaching and Learning, Nashville, TN, United States

William Rowluck

Department of Language & Literacy Education, Faculty of Education, The University of British Columbia, Vancouver, BC, Canada

Carlos Rubio-Torres

Universidad de Costa Rica, San José, San Pedro, Costa Rica

Beth Sagers

Queensland University of Technology, Faculty of Creative Industries, Education and Social Justice, Brisbane City, QLD, Australia

Casey Salman

University of Georgia, Athens, GA USA

Astrid N Sambolín Morales

Kent State University, OH, United States

Cecilia Sanchez

University of California, UC Davis School of Education, Davis, CA, United States

Sonia N Sánchez

The University of Texas at San Antonio, San Antonio, TX, United States

Loukia K Sarroub

Department of Teaching, Learning & Teacher Education, College of Education and Human Sciences, School of Global Integrative Studies, University of Nebraska-Lincoln, Lincoln, NE, United States

Cassandra Schroeder

Department of Teaching, Learning & Teacher Education, College of Education and Human Sciences,

University of Nebraska-Lincoln, Lincoln, NE, United States

Kelly A Schucker

University at Buffalo—SUNY, Niagara Falls, NY, United States

Joshua Schwab Cartas

Division of Art History and Contemporary Culture, Art Education, Nova Scotia College of Art and Design, Halifax, Nova Scotia, Canada

Frank Serafini

Literacy Education & Children's Literature, Mary Lou Fulton Teachers College, Arizona State University, Phoenix, AZ, United States

Cristine Severo

Department of Portuguese Language, Federal University of Santa Catarina, Campus Universitário Reitor João, Florianópolis, Brazil

Monica Shank Lauwo

University of British Columbia, Vancouver, BC, Canada

Jeff Share

University of California, Los Angeles, CA, United States

Kathy G Short

University of Arizona, Tucson, AZ, United States

Blaine E Smith

College of Education, University of Arizona, Tucson, AZ, United States

Patriann Smith

University of South Florida, Department of Language, Literacy, Exceptional Educational, Ed.D., & Physical Education, Tampa, FL, United States

Jim Sosnowski

Curriculum and Instruction, University of Illinois Urbana-Champaign, Champaign, IL, United States

Nerida Spina

Queensland University of Technology, Faculty of Education, Kelvin Grove, QLD, Australia

Norman A Stahl

Department of Curriculum and Instruction, Northern Illinois University, DeKalb, IL, United States

Ericka Stauffert-Reyes

University of Pennsylvania, Philadelphia, PA, United States

Bessie G Stone

Queensland University of Technology, Faculty of Creative Industries, Education and Social Justice, Brisbane City, QLD, Australia

Amy Stornaiuolo

University of Pennsylvania, Philadelphia, PA, United States

Sude

School of education, Minzu University of China, Beijing, China

Jingjing Sun

Department of Teaching and Learning, University of Montana, Missoula, MT, United States

Yuya Takeda

Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

Liqing Tao

Department of Curriculum and Instruction, School of Education, College of Staten Island, City University of New York, Staten Island, NY, United States

Adai A Tefera

University of Arizona, Tucson, AZ, United States

Laura Teichert

Department of Special Education and Literacy Studies, Western Michigan University, Kalamazoo, MI, United States

Shane Templeton

University of Nevada, Reno, NV, United States

Robert J Tierney

University of British Columbia, Vancouver, BC, Canada; University of Sydney, Sydney, NSW, Australia; and Beijing Normal University, Beijing, China

Liliana Tolchinsky

Department of Catalan Philology and General Linguistics, University of Barcelona, Barcelona, Spain

Lina Trigos-Carrillo

Universidad del Norte, Barranquilla, Atlántico, Colombia

Lindsay M Truesdale

Department of Psychology, The Catholic University of America, Washington, DC, United States

Gregory L Ulmer

University of Florida, Gainesville, FL, United States

Marijn van Dijk

Heymans Institute for Psychological Science, Department of Developmental Psychology, University of Groningen, Groningen, the Netherlands

Paul van Geert

Heymans Institute for Psychological Science, Department of Developmental Psychology, University of Groningen, Groningen, the Netherlands

Daniel A Wagner

Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

Amy Walker

Department of Literacy, Culture, and Language Education, Indiana University, Bloomington, IN, United States

Chantelle N Warner

The University of Arizona, Tucson, AZ, United States

Gustave J Weltsek

Department of Curriculum and Instruction, Indiana University, Bloomington, IN, United States

Arlette Ingram Willis

University of Illinois, Urbana Champaign, College of Education, Champaign, IL, United States

Amy Wilson-Lopez

Adolescent Literacy, Utah State University, Emma Eccles Jones College of Education and Human Services, Logan, UT, United States

Leisy Wyman

Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, Tucson, AZ, United States

David B Yaden, Jr.

Department of Teaching, Learning and Sociocultural Studies, University of Arizona, Tucson, AZ, United States

Vivian S Yenika-Agbaw

Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, United States

Kasim Yildirim

Muğla Sıtkı Koçman University, Muğla, Turkey

Junko Yokota

National Louis University, Chicago, IL, United States

Xiaoye You

Pennsylvania State University, University Park, PA, United States

Chase Young

Sam Houston State University, Huntsville, TX, United States

Mei Yuan

School of education, Minzu University of China, Beijing, China

Michelle Zaleski

Benedictine College, Atchison, KS, United States

Katina Zammit

Western Sydney University, School of Education, Penrith, NSW, Australia

This page intentionally left blank

CONTENTS OF VOLUME 10

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 10: Literacies and Languages Education</i>	<i>xix</i>
<i>Contributors to Volume 10</i>	<i>xxv</i>

LEARNING LANGUAGES/LANGUAGING AND LITERACY

Language development in the early years <i>Eve V Clark</i>	1
Dynamic system approaches to language acquisition <i>Marijn van Dijk and Paul van Geert</i>	14
Writing before age three <i>Deborah Wells Rowe</i>	27
Children's earliest insights in developing literacy <i>Liliana Tolchinsky</i>	38
Biliteracy, playfulness and performative practices <i>Pengtong Qu, Amanda Deliman, Gustave J Weltsek, Amy Walker, and Carmen Liliana Medina</i>	48
Systemic functional linguistic and disciplinary literacy development in multilingual K-12 schools in the United States <i>Meg Gebhard and Nicolas Blaisdell</i>	54
Literacies, language, and schooling: exploring writing pedagogy for English language learners <i>Erin Kim-Rich and Jen Scott Curwood</i>	65
Critical media literacy and second language acquisition <i>Andrea Gambino and Jeff Share</i>	76
Assessment of English learner students <i>Jamal Abedi and Cecilia Sanchez</i>	86
Assessment of young emergent bilingual children <i>Sandra Barrueco, Maria C Morales, and Lindsay M Truesdale</i>	96

Literacy/multiliteracies and foreign language education <i>Beatrice C Dupuy and Chantelle N Warner</i>	105
Critical literacies in East Asian languages: examples from Japanese in foreign language education <i>Yuri Kumagai</i>	112
Language policy and planning as a discursive accomplishment <i>Anthony J Liddicoat</i>	123
Transliteracy for a cosmopolitan future <i>Xiaoye You and Michelle Zaleski</i>	130
Rethinking multiliteracies: toward a radical vision for multilingualism <i>Brittany L Frieson, Luz A Murillo, Vivian E Presiado, and Jim Sosnowski</i>	138
Translanguaging and identity across global sites <i>Monica Shank Lauwo and Bonny Norton</i>	148

CULTURAL, CRITICAL AND COMMUNITY LITERACIES

Funds of knowledge and funds of identity: lenses for building language and literacy <i>María E Fránquiz and María G Leija</i>	159
Family literacies: differing perspectives <i>Jim Anderson, Laura Teichert, and Ann Anderson</i>	168
Toward culturally relevant literacies with children and families of color <i>Patricia A Edwards, Patriann Smith, and Jonda C McNair</i>	180
Literacy and identity across home, school, church, and social life <i>Gilberto P Lara, Sonia N Sánchez, Myriam Jimena Guerra, and Lucila D Ek</i>	198
Religious influences on the growth of literacy practice <i>Loukia K Sarroub and Cassandra Schroeder</i>	209
Critical literacies & social justice: social justice approaches in schools <i>Nerida Spina, Barbara Comber, and Michelle Jeffries</i>	217
Critical literacies and social justice <i>Ernest Morrell and Jodene Morrell</i>	225
Youth poetry and literacies in remaking the world <i>Korina M Jocson</i>	235
Youth out-of-school (critical) literacies, texts, and audiences <i>Kathleen A Hinchman</i>	242
The construction of gender in literacy <i>Barbara J Guzzetti and Katina Zammit</i>	250
Literacy at the intersections of race, gender, and sexuality <i>Latrise P Johnson and Mollie V Blackburn</i>	262
Literacy access, the law, and racism in the United States: a critique <i>Arlette Ingram Willis</i>	268
Learning disabilities' unsettling intersections: disrupting learning, cultural, socioeconomic, and environmental deficits <i>Adai A Tefera and Alfredo J Artilés</i>	279

Borderlands: unique sources of transnational language and literacy for educational research <i>Rosalind Horowitz</i>	288
Culture and literacy: situated social literacies <i>Martin Prinsloo and Lara-Stephanie Krause-Alzaidi</i>	296
LITERACY DEVELOPMENT & PEDAGOGY WITH CHILDREN, YOUTH, AND ADULTS	
Literacies, language & schooling: teaching and learning the language arts <i>David Bloome, Hilary Janks, and Allison Wynhoff Olsen</i>	308
Reading comprehension in troubled times <i>Gina N Cervetti and P David Pearson</i>	319
Collaborative Reasoning and children's development of language and thought <i>Shufeng Ma, Richard C Anderson, and Jingjing Sun</i>	330
Literature: critical imagination and young adult readers remaking worlds <i>Patricia Enciso and Beth Krone</i>	337
Humanizing language and literacy in education with dramatic inquiry <i>Brian Edmiston</i>	346
Arts-based practices for literacy & language learning <i>Douglas Fisher and Nancy Frey</i>	358
Which words to teach: review and reflection <i>Elfrieda H Hiebert</i>	365
Spelling: theory, assessment, and pedagogy <i>Shane Templeton</i>	374
The art and science of reading fluency and fluency instruction <i>Timothy Rasinski, Chase Young, David D Paige, and Kasim Yildirim</i>	388
Expanding meaningfulness for struggling readers <i>Gay Ivey and Karole-Ann Friddle</i>	398
The meaning(s) of dyslexia <i>Richard L Allington and Rachael Gabriel</i>	408
A constructivist lens of eye-tracking in special populations <i>Ryan L Nelson, Laura E Arrington, Narges Firouzshahi, Patricia Arias Hunt, Amanda Percle, Holly Damico, and Jack S Damico</i>	414
Issues and current trends of research in braille for people with visual impairments: questions and implications <i>Sunggye Hong</i>	423
Autism, literacies, and social communication: a framework of inclusive new literacies <i>Bessie G Stone and Beth Sagers</i>	434
Shifting from disciplinary versus content area literacy to a transdisciplinary approach <i>Thomas W Bean, Amy Wilson-Lopez, and Kristen Gregory</i>	448
(Mis)understanding adult literacy <i>Erik Jacobson</i>	456

College reading	463
<i>Norman A Stahl, Sonya L Armstrong, and James R King</i>	
Research and practice in literacy teacher education: imagining, designing, and engaging in "what could be"	474
<i>James V Hoffman</i>	

DIGITAL AND MULTIMODAL LITERACIES

Sensory literacies: the full sensorium in literacy learning	480
<i>Kathy A Mills and Beryl Exley</i>	
Expansive, dynamic, and critical new literacies: widening views of digital and multimodal learning	487
<i>Elena Forzani and Jill Casteck</i>	
Digital media arts and critical media education: an alchemy of creative literacy practices	497
<i>Michelle Cannon</i>	
Responding to and creating multimodal texts	503
<i>Richard Beach and Blaine E Smith</i>	
Electracy	518
<i>Gregory L Ulmer</i>	
Artificial intelligence in literacy	529
<i>Dom Massaro and Jared Bernstein</i>	
Evaluating, supporting and enhancing Internet search activity: a review of research, with implications for pedagogy	543
<i>Colin Harrison</i>	
E-texts and e-books: constructing meaning with and from texts in an online and digital world	551
<i>Bernadette Dwyer</i>	
Digital literacies, multimodality and the engagement of adolescents who are disengaged or disenfranchized in school	563
<i>David G O'Brien and Deborah R Dillon</i>	
Digital literacies & multimodal learning: writing in digital formats	572
<i>Blaine E Smith</i>	
Digital games and literacies	583
<i>Jennifer Jenson and Danielle Burrell-Kim</i>	
Transliteracy and the digital media: theorizing media and information literacy	590
<i>Divina Frau-Meigs</i>	

CHILDREN'S LITERATURE: AN INTERNATIONAL PERSPECTIVE

Critical literary perspectives on children's and young adult literature	605
<i>Theresa Rogers and Caroline Hamilton-McKenna</i>	
The evolution of children's literature: metafictional devices and digital mediation	620
<i>Frank Serafini and Sylvia Pantaleo</i>	
Global literature for young people: critical issues and social responsibilities	633
<i>Kathy G Short and Junko Yokota</i>	

Children's literature in Latin America today: a historical perspective <i>Lilian Alvarez-Arellano, Carlos Rubio-Torres, Marcelo Emilio Bianchi-Bustos, and Donovan Herrera-Santillán</i>	644
Discovered and reconstructed: a perspective of Chinese children's literature over the last 100 years <i>Wei Kan</i>	662
Children's literature in Greece and Cyprus <i>Petros Panaou</i>	672
Educational concerns in Arabic picture books <i>Yasmine S Motaawy and Susanne K Abou Ghaida</i>	678
The scholarship of African Children's literature: issues of representation of critical voices in the conversation <i>Vivian S Yenika-Agbaw, Phoebe L Quaynor, and Jocelyn Amevuvor</i>	688
LITERACIES IN THE GLOBAL NORTH & SOUTH	
Global meaning-making: critical analysis and advocacy for a diverse world <i>Robert J Tierney</i>	696
Cosmopolitan literacies and the limits of understanding: possible directions in global literacy research <i>Amy Stornaiuolo and Rabani Garg</i>	710
Literacy and international development <i>Daniel A Wagner and Walid Hedidar</i>	717
Regimes of literacy as regimes of truth about Africa: language ideologies and southern voices <i>Ashraf Abdelhay, Cristine Severo, and Sinfree Makoni</i>	723
Literacy, colonialism and resistance in Latin America <i>Lina Trigos-Carrillo, Miriam Jorge, Luzkarime Calle-Díaz, and Rebecca Rogers</i>	736
Chinese literacy acquisition: an overview <i>Liqing Tao</i>	747
Chinese Minzu (民族) diversity and education <i>Sude, Mei Yuan, and Fred Dervin</i>	759
The politics of literacy languaging and texts: the politics of literacy in Nigeria. <i>Chukwuemeka Eze Onukaogu, Joseph Onyema Ahaotu, Onyedikachi Abiodun-Ekus, and Kate Omenugha</i>	770
Community-centered Indigenous language recovery, restoration, revitalization, and renewal <i>Candace Kaleimamoowahinekapu Galla, Amanda Holmes, Joshua Schwab Cartas, and William Rowluck</i>	786
(Just) languaging with <i>duende</i> : (re)inventing and innovating communicative practices as resistance, healing, and thriving in Indigenous contexts <i>Joaquin Muñoz, Candace Kaleimamoowahinekapu Galla, Leisy Wyman, and Perry Gilmore</i>	797
Mobilizing prefigurative hybrid literacies for decolonial resistance: three transnational examples <i>Astrid N Sambolín Morales, Ericka Staufert-Reyes, and Gerald Campano</i>	809

Language choice, technology and international education 817
Nathan M Castillo, Daniel A Wagner, Ghaida S Alrawashdeh, and Aanchal Gidra

On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT 825
Ryosuke Aoyama, Takeshi Kajigaya, Yuya Takeda, Ryuko Kubota, and Ryan Deschambault

CONTEMPORARY AND EMERGING THEORETICAL AND METHODOLOGICAL ISSUES

The literacy journey: from a psychological to a biological way of thinking about learning to read and write 837
Brian L Cambourne and Debra K Crouch

Reintroducing “development” into theories of the acquisition and growth of early literacy: developmental science approaches and the cultural-historical perspective of L. S. Vygotsky 849
David B Yaden, Jr. and Camille Martinez-Yaden

A naturalized view of literacy education: what the neuro- and life sciences may offer language and literacy research 866
George G Hruby and Ayan Mitra

Literacy models and epistemology 883
James W Cunningham

Are we there yet? Onto-epistemological qualitative language and literacy research 895
Marla H Mallette, Jillian Mason, and Matthew McConn

Epistemologies and methodologies in literacy research: a case study 902
Seth A Parsons, Ashley N Patterson, Jamie Colwell, Bong Gee Jang, and Betsy A Baker

Affect theory and textual variations 911
Kimberly Lenters

Posthumanism and Positioning Theory: explorations of embodiment and affect from a critical perspective 918
Mary B McVee, Kelly A Schucker, and Aijuan Cun

Literacy and the spatial production of school during a pandemic 928
Kevin Leander and Michael Havazelet

Post-philosophies and the doing of inquiry in literacies studies 938
Candace R Kuby

Urban humanities as experimental research praxis 948
Jonathan Jae-an Crisman

Reimagining research on writing: a timely rhizomatic argument 958
Donna E Alvermann and Casey Salman

Language development in the early years

Eve V. Clark, Stanford University, Stanford, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Getting started	2
Adding and storing words in memory	2
Early utterances	2
Taking up offers of new words	4
Coining words	6
Language in interaction	8
Other language skills	9
Figurative language	9
Pretend play	9
Politeness	9
Telling stories	10
Crosslinguistic comparisons	10
Conclusion	11
References	11
Further reading	13
Relevant websites	13

Introduction

How much do children know about using language by age 5? Normally-developing children typically know how to convey their wants and needs, how to answer questions, say what they are interested in, tell others what they have been doing, and what they want to do now. They can voice different roles in pretend play, and make up stories to accompany that play as they enact familiar activities like shopping, cooking, eating a meal, giving the baby (doll) a bath, searching for a lost puppy (toy), or acting out favorite stories. But what words and constructions have they acquired by this age in order to achieve such a range?

In this article, I begin by summarizing some of the early steps children take as they begin to follow a path into language use, and describe the differences between speech comprehension (leading the way) and speech production (lagging, sometimes quite far behind), and how these core processing skills together help children arrive at much already adult-like usage in speech interactions by age 5 (Clark, 2019a). I then take up in more detail what children need to know in order to set up a system of communication, from early gestures to gesture-word combinations, to more complex linguistic constructions, often emphasized or supplemented with gestures, by the age of 4 or 5 years. And I argue that children benefit greatly from extensive experience with language use from their parents and other adult caregivers, feedback of some kind, preferably tied to content, on their errors of usage along the way, and practice in all kinds of settings, practice in speech comprehension through exposure to a variety of speakers, and practice in speech production, again in a variety of settings—choosing and planning the forms appropriate for the intentions they are trying to convey, taking turns on time, and contributing appropriate information in their turns within an exchange. In achieving all this, children must take into account what their addressees already know (given information), and contrast that with information they are offering that is new. That is, they attend to what is currently in common ground, and so already known to the other, and what, at each step in the exchange, counts as new information, and so keep track of where the conversation is going. Talking with another person also means keeping track in the longer term of what that person knows, from any shared history of interactions on other occasions.

In acquiring a first language, exposure to and use of language in interaction is critical. Parents start out with fixed routines that accompany such everyday activities as changing nappies, feeding, picking up, putting to bed, and so on as they talk to their infants (Ferrier, 1978; Tamis-LeMonda et al., 2019). And adults elaborate on these, talking about what the young child is doing, treating it as an interactive partner by taking sounds, cries, smiles, reaches, hand gestures—any small action on the child's part—as “turns” in an exchange (Snow, 1977). They track babies' and small children's gaze as indicating attention and act so as to attract their gaze in setting up interactions (Moore and Dunham, 2014). They scaffold young children's single-word utterances and thereby model turn-taking in a conversational exchange (Bruner, 1983), and they make requests for clarification whenever they don't understand what the child is trying to say (Clark, 2020b). All this presents children with language used for interactive communication, where both words and gestures play a role, and where language is also an essential ingredient in coordinating with others in specific activities such as building a tower, doing a puzzle, getting dressed, going up stairs, playing catch, putting a doll to bed (e.g., Garvey and Kramer, 1989).

Getting started

Infants attend to their environment. They look at faces, patterns, angles, and corners. They follow gestures and the direction indicated by an adult point (Thoermer and Sodian, 2001). This attention on the child's part is critical for establishing and maintaining joint attention with someone else on an object, place, or action (Estigarribia and Clark, 2007). Even during early infancy, babies attend to the sounds they hear in the adult speech stream, discriminate same from different syllables, and recognize recurring sequences of sounds (Saffran et al., 1999). By 10 months or so, they begin to attend more closely to the sounds of the ambient language than to sounds from other languages (Werker and Tees, 1984) and even show early comprehension of a few words (Bergelson and Swingle, 2015). And, somewhere between 8 and 12 months of age, infants begin to babble, producing short sequences of syllables, with varying intonation contours. Most children go on to produce their first words, or first attempts at words, between 12 and 24 months. During their second year, they rely on both gestures and gesture-word combinations as they try to communicate with the people around them (Clark and Kelly, 2022).

Recognizing sequences of sounds heard from others and managing to produce sequences of sounds are two very different skills. Young children's recognition and comprehension of words spoken by adults is typically well ahead of their ability to produce recognizable versions of those same words (Clark and Hecht, 1983). In fact, even 1-year-olds monitor their own productions and try to correct them when they fail to match the adult targets they already have represented in memory. This can be seen, for example, in (1), where Brenda makes six attempts at the word *shoe* (Scollon, 1976):

(1) Brenda (1;8, talking to her mother as she held up her mother's shoe

and looked at it): *mama, mama, mama.*

sh. shi. sh. shiss. shoe. shoesh.

Mother: Shoes!

Young children also respond to adult prompts for clarification, making repairs to their own productions in response (Clark, 2020b), for instance, by improving the pronunciation of a word, as Brenda did, or by changing a word to make the meaning clearer (Clark and Bowerman, 1986), as in (2):

(2) D (1;9.5, looking at a picture of a crab, a word up till now pronounced as "kam"):

kab. (a minute later, same picture): *kam* <repair> *kab.*

Adding and storing words in memory

As children add adult words to memory, their comprehension vocabulary grows from just a few words at around 10–12 months, up to several hundred words by age 2. And their production of different words also increases steadily, (See Table 1) but rather more slowly than their comprehension (Fenson et al., 1994). The size of these early vocabularies, in both comprehension and production, covers a considerable range, affected in part by how much parents talk to their children (e.g., Arriaga et al., 1998; Hurtado et al., 2008). This in turn varies with social class and child-rearing beliefs, with upper- and middle-class parents typically talking more with their children than lower-class and welfare parents in the US (Hart and Risley, 1995; Heath, 1983) (Table 1).

Early utterances

Children start by saying just one word at a time (Bloom, 1973). But once they can produce more words that are recognizable, they also begin to produce an increasing number of two-word combinations. From that point on, they also attempt more complex constructions. They produce some modifying adjectives like *big* and *small* in combination with nouns, as well as a few color terms

Table 1 Early words, produced by at least 50% of children at each month of age from 1;0 to 1;6.

<i>People:</i>	daddy (1;0), mommy (1;0), baby (1;3), grandma (1;6), grandpa (1;6)
<i>Food/drink:</i>	banana (1;4), juice (1;4), cookie (1;4), cracker (1;5), apple (1;5), cheese (1;5)
<i>Body parts:</i>	eye (1;4), nose (1;4), ear (1;5)
<i>Clothing:</i>	shoe (1;4), sock (1;6), hat (1;6)
<i>Animals:</i>	dog (1;2), kitty (1;4), bird (1;4), duck (1;4), cat (1;6), fish (1;6)
<i>Vehicles:</i>	car (1;4), truck (1;6)
<i>Toys:</i>	ball (1;3), book (1;4), balloon (1;4), boat (1;6)
<i>Household objects:</i>	bottle (1;4), keys (1;5)
<i>Routines:</i>	bye (1;1), hi (1;2), no (1;3), night-night (1;4), bath (1;5), peekaboo (1;5), thank you (1;6)
<i>Activities (sound effects, motion, state):</i>	uh-oh (1;2), woof (1;4), moo (1;4), ouch (1;4), baa baa (1;4), yum yum (1;4), vroom (1;5), up (1;5), down (1;5)

Based on Fenson et al. (1994: 93).

(initially often wrongly applied), and they understand and ask questions (e.g., Tyack and Ingram, 1977; Kemler Nelson et al., 2004; Chouinard, 2007). From around 2;0 to 2;6 on, they actively ask for words for unfamiliar objects and actions, and readily take up new words when these are offered (e.g., Clark, 2007, 2010a).

In many languages, words can be modified in meaning with the addition of suffixes or prefixes. Such modifications may include inflections that mark singular vs. plural, possessive, male vs. female, and case marking for grammatical relations like subject vs. object. Other modifications in word meaning may be provided by function words, e.g., English *of*, *to*, or *with*. Children acquiring English produce some inflections and function words by age 2 (Brown, 1973). They use plural marking early, distinguishing one from more-than-one, sometimes relying first on a numeral (usually *two*) before resorting to the plural suffix *-s* on nouns (Clark and Nikitina, 2009); they also mark possession on nouns with *-s* early, as well as the regular past tense, with *-ed*, on verbs; and they produce a few prepositions like *in*, *but*, *to* begin with, again often use these incorrectly. By age 3, in English, children typically over-use the regular past tense *-ed* on irregular verbs, in such productions as *fallen*, *bringed*, *goed* (also *wented* with mis-identification of the verb stem), *runned*, and *eated*, but produce a number of other grammatical morphemes correctly.

Another early inflection, first produced at around age 2, is *-ing* added to non-stative verbs for actions with some duration (e.g., *boy running*, *dog barking*). This is followed by the auxiliary verbs *be* and *have* (*he's running*; *I have teddy*) as well as copula *be* in equative utterances (e.g., *doggie is big*). They mark gender with the pronouns *he* and *she* (though at first they over-use *he*), mark plural number with *they*, and otherwise use *it* or demonstrative *that* in their references. By age 4 or 5, children understand proximal *here* (close to the speaker) vs. non-proximal *there*, followed by proximal *this* vs. non-proximal *that*, and by age 8 or 9, have also mastered the conditions on deictic verbs like *come* (motion toward the speaker) vs. *go* (motion away from the speaker), as well their causative counterparts *bring* and *take* (see Clark, 1978; Clark and Sengul, 1978; Clark and Garnica, 1974).

As their vocabulary grows larger, children begin to produce more complex constructions by combining two or more propositions. And from around age 2;6 to 3;0 on, for example, they can talk about how events are ordered in time using *and then*, with the actual event order marked by order of mention (Clark, 1973), as in (3):

(3) L (~3;6, to teacher): *First you'll have your coffee and then we'll*

go outside (to the playground).

They also start to use conjunctions like *when*, *before*, and *after*, as in (4) and (5):

(4) A (~3;6, showing how to hold up a hand in a "stop" gesture):

This is what a policeman looks like when a car stops.

(5) D (2;6.22): *I going [ə] bring this pile of books to the table,*

after I aten my supper, then I can read them.

At first, children generally introduce temporal subordinate clauses in second place, after the main clause. But they soon begin to use some conjunctions to introduce first-position clauses as well, as in (6):

(6) D (2;5.3, starting a nap): when you close the door, then I can kick all my

blankets off.

Two- and 3-year-olds also talk about causes, external and internal, with *because*, as in (7) and (8), also with the subordinate clause in second position:

(7) E (2;5.21, going towards disks): get them cause I want it.

(8) D (2;3.30, to Father starting to clear the table): You don't get my bowl

because ... I still eating.

and about conditionality with *if* as in (9), in either first or second position:

(9) D (2;8.26, asking for a specific T-shirt): If it doesn't have writing on it,

it's not my Levi's shirt.

They also attempt relative clauses from around age 2 onwards, with varying degrees of success and frequent early omission of relativisers like *who* or *that*, as in (10) and (11):

(10) D (2;0): I see [ə] building Eve go.

(11) D (2;5.16, touching a wet spot on newspaper): That paper what

Eve got fell into a tiny puddle.

See Clark (1971, 1973), Diessel and Tomasello (2000), Donaldson and Balfour (1968), Jourdain and Lahousse (2021), Klatzky et al. (1973), Reilly (1986), Sheldon (1974).

Children acquire still more words and constructions from being read to, and later on from starting to read themselves (e.g., Noble et al., 2018; Seidenberg, 2017; Snow and Matthews, 2016). Overall, vocabulary acquisition involves long-term uptake, and it is not until early adulthood that many speakers of English attain a vocabulary size estimated at around 50,000 words. Indeed speakers continue to add new words on a life-long basis, albeit much more slowly after age 18 or so (Clark, 2016).

Taking up offers of new words

When talking with their children, adults make explicit offers of some new words, and children often take these up. They show this by repeating the words in the next utterance, and using them on subsequent occasions. These adult offers are typically flagged with specific frames like: *This is a—, That's a—, This/that is called a—, It's a—, Those are—* (Clark and Wong, 2002; Clark, 2007, 2010a; Clark and Estigarribia, 2011), as in (12) and (13):

(12) Naomi (1;8.6; Sachs 3:84/CHILDES)

Father: That's a pen (offer of word).

Naomi: *pen* (uptake with repetition).

(13) Abe (2;7.26; Kuczaj 26:245/CHILDES)

Father: What's he doing?

Abe: *I don't know. I don't know.*

Father: He's pulling a wagon (offer of word).

When adults offer new words, they often link them to other words the children already know. They indicate the kind of object in terms of set-membership or inclusion, e.g., *A seal is a kind of animal*, *A retriever is a kind of dog*, where the terms *animal* and *dog* provide familiar or known anchors for the new words *seal* and *retriever*. They offer information about properties or material associated with particular objects, using expressions like *is made of*, *is shaped like*, *sounds like*, *is the same color as*, and so on. They comment on parts of the newly-labeled entity, identifying such elements as *feet*, *hands*, *wheels*, etc. They may also provide lists of related objects or actions. And they sometimes offer explicit definitions for new words, as in the parental repair (underlined; Gelman et al., 1998), shown in (14):

(14) Child (2;11, looking at a book)

Mother: Do you know what that one is?

Child: *Ummm.*

Mother: I don't know if you know what that one is.

Child: *That's a snake.*

Mother: It looks like a snake, doesn't it? It's called an eel. It's like

a snake only it lives in the water. And there's another one.

The information provided along with such direct offers of new words helps children as they build up semantic domains--terms related to each other in meaning. Their very earliest domains consist of small numbers of words that are related in some way. Later the relations linking words become more varied and complex. Compare the set: "*Mummy, daddy, baby*," say, with the more elaborated set: "*Parent, mummy, daddy, child, boy, girl, baby, walk, eat, talk, cry, sleep, jump, shout, play, throw*," where the latter grouping

Table 2 The early domain of animal terms in one child’s speech (Clark, 1995).

I. Initial repertoire in the 2nd year (1;0.29–1;11.30)	
1;0.20	doggie, 1;1.15 dog, 1;3.3 mouse, 1;3.23 cat, 1;4.28 horse, 1;6.13 cow, 1;6.28 rabbit, 1;11.30 goat
1;1.26	bear, 1;7.19 lion, 1;8.8 alligator, 1;11.6 gorilla, 1.10.24 seal
1;2.0	turtle, 1;6.4 fish, 1;8.7 frog, 1;8.8 snake, 1;8.22 crab, 1;10.19 ladybug
1;7.20	animal
II. Elaborating the repertoire in the 3rd year (2;0–3;0)	
2;0.8	puppy-dog [toy dog], 2;2.24 baby-rabbit, 2;5.21 mummy-bunny [toy rabbit], 2;7.27 sheep, 2;11.1
cattle, 2;11.1	baby-cattle, 2;11.1 daddy-cattles, 2;11.3 bronco-horse, 2;11.10 bucking-horse
2;0.11	hippo, 2;1.7 camel, 2;4.10 baboon, 2;4.16 tiger, 2;5.21 monkey, 2;7.15 wolf, 2;7.15 raccoon, 2;8.8
armadillo, 2;8.8	fox, 2;8.17 beaver
2;3.21	stork, 2;3.21 ostrich, 2;4.6 robin, 2;4.17 sparrow, 2;4.19 flamingo, 2;4.20 dove, 2;5.10 grouse,
2;5.14	woodpecker, 2;8.6 owl
2;5.1	trout, 2;5.4 flounder, 2;9.21 lizard-animal [live], 2;9.21 frog-animal [live]
2;5.6	spider, 2;5.6 grasshopper, 2;7.2 fly, 2;7.2 bee, 2;8.14 butterfly

involves many more relations linking the words. Some words can be members of several, or even many, semantic domains, while others are more specialized in meaning: compare *baby*, *foal*, *calf*, and *puppy*, for example (Clark, 2010b, 2018b).

Acquiring new words allows children to build up semantic domains, with steady additions, for example, to their earliest words for animals, for toys, for vehicles, for household objects, for activities. This leads children to organize their growing vocabulary into groups that are related in meaning: terms for animals, with male/female and adult/young terms; terms for the noises particular animals make, for their characteristic motions and actions, for their usual food, and for their habitats, say (Clark, 2018b). Table 2 shows the gradual build-up in production of the semantic domain for various types of animal in the speech of one child. By the end of his 2nd year, this child used some 21 animal terms. He then added a further 37 terms to this domain in the next year. In a few cases, he constructed a compound form, produced with appropriate primary-tertiary stress, as in *baby-cattle* for a calf, or *lizard-animal* for a live lizard seen outside.

Coining words

When children (or adults) lack a word for a particular entity or action, they will often coin one on the spot. In English, from the age of 2 onwards, children rely extensively on compounding, combining two nouns with a primary-tertiary stress pattern, as in the examples in (15):

(15) a. D (1;8.5, at table, puts spoon in cup): *ORANGEJUICE-spoon*.

b. D (1;11.28, wanting *Lion in the Meadow*): read [ə] *LION-book*.

c. D (2;2, playing with a small and large strainer): *That [ə] TEA-sieve*.

Mo: What?

D: *That [ə] TEA-sieve* (then pointing at large strainer). *That [ə] WATER-sieve*.

(Clark et al., 1985; Clark and Berman, 1987; Berman and Clark, 1989). They also readily use nouns as verbs when they either lack or can’t retrieve the relevant verb (Clark, 1982, 1993), as in (16):

(16) a. D (2;4.13, grinding pepper onto counter): *I’m sanding*.

b. D (2;9.1, of a sock, to Mo): *And did you needle this?* [=mend]

c. D (2;9.10, talking about boats): *And we might see a man **oaring** a boat*

with oars. [=rowing]

d. S (3;0.21, watching a man open door with a key): *He's **keying** the door.*

By age 3, children also make ready use of the productive agentive suffix *-er*, as in (17):

(17) a. D (2;5.26, reaching across counter): *I'm a big **reacher**.*

b. C (2;4.18, of someone cooking) *She a **cooker**.*

c. C (3;7): *That's the **rainer**.* [=person who drives away the rain]

d. F (3;11, contrasting with "painter"): *He's a **crayoner**.*

But they take rather longer to master the less productive agentive suffixes, *-ist* and *-ian* (Clark and Cohen, 1984; Clark and Hecht, 1982). Other suffixes like adverbial *-ly* in English are acquired later still, as are prefixes like *un-* (Clark et al., 1995), as in (18):

(18) a. D (2;9.11, opening plug in bidet to empty water): *It's **unflowing**.*

b. D (2;10.20, hiding thumbs in his fists): *They've **disappeared**.*

Mo: Can you make them appear again?

D: *I can't make it **undisappear**.*

c. D (3;4.8, getting things out of school cubbyhole):

*That's a picture and you have to **uncrumple** it to see what it is.*

d. D (3;5.9, to Mo): *Show me how you **uncatch** your necklace.*

By age 5, children have mastered a number of the most productive options for forming new words in their language. They attend first to transparency in form as they construct new forms for new meanings, and in particular they attend to highly productive forms. In acquiring Germanic languages, for example, they start early with compounding and with zero derivation, while in Romance languages, they start early with derived forms, adding suffixes (and some prefixes) to existing forms, to construct new words. Cross-linguistically, children follow similar trends, but generally first latch on to those patterns that are productive in the language they are acquiring (Clark, 1993, 2019b).

Language in interaction

Children learn language in interaction. Adults talk with them and carry out actions with them along with their talk. Children must respond, show they have understood, and make their own contributions to any ongoing exchange in conversation (Clark, 2018a). But children start out knowing virtually nothing about language itself, and so rely at first on gestures, single words, and gesture + word combinations (Clark and Kelly, 2022). As they hear more language and get more practice in using words, they advance to producing some two-word combinations (Braine, 1976; Bowerman, 1973; Slobin, 1973).

In using more speech, children need to make sure they are making themselves understood by others. They monitor their own production to see that it matches what they have already stored in memory for comprehension, and they make spontaneous repairs to their own speech when they detect a mismatch between a stored form and the version they produced (Clark and Hecht, 1983). They respond to requests for clarification, whether Open requests like *hm?* or *what?*; to Restricted requests like *He said what?*, or to Restricted offers, as when adults reformulate erroneous child utterances in the next turn and so provide children with conventional forms that would convey what was apparently intended (Clark, 2020b; Chouinard and Clark, 2003; Clark and de Marneffe, 2012; Lustigman and Clark, 2019). In responding to clarification questions from adults, children consistently try to make clearer what they had meant.

Using language in interaction depends on being able to track what the addressee does and doesn't know. Establishing and then tracking common ground is central to language use, and children start to track what their addressee knows early on. They also start to take turns very early, in vocalizing, in passing objects in play, and then, with the onset of word production, in producing a word as a turn (Hillbrink et al., 2015; Snow, 1977). But timing with language is complicated since the child has to plan the next turn, and do so with minimal gaps between speakers as they take their turns.

One place we can see this is when children answer questions, both *yes/no* polar questions and *wh*-questions, as in *Do you want some cheese?* and *Where are your sandals?* In answering such questions, children must first understand the question and then decide how to answer. That is, the turn-taking involved in answering questions requires that children plan their answers. They need to retrieve the right words, put them together, then produce an utterance that responds to the question. Such planning takes time, and if speakers are to avoid gaps and overlaps as they take turns, they need to be able to plan the next turn before the current one is done. One domain where such turns have been examined is in the timing of young children's answers to questions. There, the gap between adult question and child answer is larger when the question is more complex, but this gap becomes smaller as children get older and more skilled in their turn-timing (Casillas et al., 2016; also Garvey and Berninger, 1981; Hillbrink et al., 2015; Stivers et al., 2018). Finally, very young children may know the answer well before they can produce the word(s) needed. In a case study of child answers to early *Where* and *Which* questions, for example, Clark and Lindsey (2015) found that the child produced appropriate gestures as answers faster than he did words.

Participants in a conversation keep track of what they've said and what the other person knows as a result of the exchange so far. That is, they keep track of common ground (Clark, 2015, 1996). This means keeping track of information status, what is already "given" in the exchange, and what is "new" information provided by each participant in the next turn.

How do children, and adults, present information as given vs. as new when they contribute a turn to the ongoing exchange? Young children distinguish given from new information very early, from the two-word stage on. New information in English is marked with heavier stress, and children make use of stress, as in *Book THERE* or *ANDY bed*, where the more heavily stressed term in each utterance conveys the new information being added. While they typically answer questions with new information alone, they make many contributions that add new information to something that has already been given (Baker and Greenfield, 1988). As they move on to more complex constructions, children assign whole clauses as given vs. new, as in *After his brother came in, he sat down at the table*, with the first clause here comprising given information and the second new information (Clark, 1973).

More generally, even very young children take the addressee into account, just like adults do, tracking what the other person can see, for example, as when a child holds up a book to display a page. At age 1 or 1;6, children will tilt the top edge of the book down toward the other person who, as a result, sees the target page upside down. But by age 2 or so, children turn the book in the horizontal plane instead, so the other person can see what the children had been looking at, upright on the page (Lempers et al., 1977). By age 2, children also keep track of whether a parent, say, had been present and so knows what happened earlier in what the child had been playing with (O'Neill, 1996). That is, children start tracking common ground from an early age, and so choose their actions and words to reflect the relevant perspective for the person they are talking with (Clark, 1997).

At the same time, children are building up information about categories and relations in the world around them, and about how to talk about them. They are also building up cultural knowledge about their surroundings and about how people around them use language. And they assign words to refer to those category types, gradually building up semantic domains for the conceptual categories identified, by building on their earlier conceptual development (Clark, 2018b). Critical to building up different semantic domains are the patterns of communicative interaction that surround the children. Reading to children helps expose them to new words and new categories. And talking to them about whatever is happening and what they are doing provides them with information about the words to use, and models of how to use them.

In short, conversational interaction, and how it is layered with ongoing collaborative activities, is central to adult language use (Clark, 1996; Enfield and Levinson, 2006; Kempson et al., 2016). It is also central when it comes to acquisition. Children learn their first language in conversational settings as they begin to understand others and try to communicate their own needs and intentions.

Conversational interaction provides the site where children gain more and more experience with language use, both in comprehension and in production. It is the site where they receive feedback about what they have (or haven't) understood as well as about what they have produced. And it is the site in which they practice what they have acquired so far (Clark, 2022). Conversational interactions therefore provide experience, feedback, and practice in language use—all three play an essential role during the process of acquiring a first language.

Other language skills

Early on, children learn to take different perspectives on the same referent as they choose terms for referring to it, and they start to use some figurative language quite early on, at around 2;6 to 3. At age 3–4, they begin to use stage directions in their pretend play, specifying actions as well as what to say in enacting specific roles. And they make more and more subtle use of persuasion when they make requests. And, by age 5, they have learned the basics of how to tell a story.

Figurative language

Some figurative uses of language appear to emerge quite early in children's spontaneous speech, starting at around age 2 or 3 (Billow, 1981; Clark, 2020a). That is, children that age can shift perspective and so refer to the same entity in different ways (*the waste-basket* vs. *my hat*). Such shifts in how to refer to the same entity license their extension of different terms to new contexts (Clark, 1997). This ability to shift perspective, combined with extensive pretend-play, appears to lay the groundwork for figurative usage. And, from age 3 onwards, children do well in understanding metonymy, for instance, where a term for a part is used to refer to the whole object, as in *the beard* for *the man with a beard*, as measured in an eye-tracking task (Köder and Falkum, 2020). Three-year-olds also do well in a choice task where they have to identify a referent and hand it to an adult, where the referring expression consists of a metaphor appropriate to what 3-year-olds know and to their current vocabulary (Pouscoulous and Tomasello, 2020).

In short, shifts in perspective on objects or events, together with pretend-play, provide a basis for extending language in various ways, and this allows for some early understanding and use of figurative expressions.

Pretend play

Children begin pretend play very early, enacting everyday activities (like dressing, sleeping, having a bath), initially with a parent, and then, as they get older, with siblings and peers. And talk is integral to this play. Even children as young as 2 or 3 distinguish whether they are talking as a character in a doll's house, say, vs. commenting outside of the play itself to an adult (Strömqvist, 1984). And as they get older, children give stage directions to their play mates, marking them with use of the past tense to signal non-present *irrealis* (i.e., the event has yet to happen) (e.g., Kaper, 1980; Garvey and Kramer, 1989; Sawyer, 1997). These stage directions may assign roles and in particular the actions that should follow, including talk, that need to be carried out in the play.

By age 4, children are quite adept at adjusting their own speech for different addressees. They adopt different styles of talk in explaining a Noah's Ark toy to a 2-year-old compared to a 4-year-old peer, or compared to an adult (Shatz and Gelman, 1973). Moreover, from as young as age 2, children take into account what the addressee doesn't know and supply the relevant information (see O'Neill, 1996; Liszkowski, 2005; Knudsen and Liszkowski, 2013; Marcos et al., 2021; Morgenstern and Sekali, 1999).

Children also adopt the appropriate speech style or register for any roles they take on as speakers. They contrast the speech to use as a toy or as a 2-year-old with that they use when enacting the role of an older child or an adult (Sachs and Devin, 1976). They are also adept at taking different roles, and the voices appropriate to those roles, in talk as-a-child vs. as-a-parent, as-a-mother vs. as-a-father, as-a-pupil at school vs. as-a-teacher, or as-a-child-patient vs. as-a-doctor (Andersen, 1990). They mark such roles with different vocabulary and construction choices to signal such social dimensions as relative power, status, and age. In short, by age 4–5, children can take into account a number of social dimensions that are regularly reflected in how people talk to each other. Their requests, when speaking as children, for example, are politer in form when addressed to someone with more power (father) than those addressed to someone with less power (mother): Compare the requests in "Ice-cream tastes nice, doesn't it?" vs. "I want some ice-cream." Assigning and taking on different roles, of course, also involves being able to switch perspective on how to refer to a particular individual, a younger brother, say, who has been assigned the role of baby vs. father in family-based play.

Politeness

Politeness formulas in language often involve routine expressions like *please* and *thank you*, as well as uses of modal verbs (*could you ...*, *would you ...*, *may I ...*), and other indirect forms such *would you mind ...* (which expects the answer "no" when the addressee agrees with the speaker). Address terms, whether one is expected to use a first name alone, a title and surname, and second-person pronoun uses, and kin-term uses, differ from one culture to another. And here too, parents offer extensive feedback on just what to say when the occasion demands it. Although children may start using some polite forms by age 3 or so, it may take considerable time before they use polite language appropriately, without prompting.

Very young children typically have politeness first imposed on them in the form of politeness routines: Say “please,” Say “thank you,” Ask “may I ...”, and so on (Gleason et al., 1984; Greif and Gleason, 1980). These routines provide early contexts for conventional expressions of politeness. And by around 2;6 to 3;0, children can judge which of two requests, say, is better, and so more likely to succeed (Bates, 1976). In one study of politeness in children aged 2;7 to 5;8, she looked at the forms of (nice) requests the children produced themselves, and at their comprehension of such forms, tested through judgments of which requests were the nicest or the best way to ask (Bates and Silvern, 1977). Children’s comprehension of polite forms was related to individual differences in social adjustment (as assessed by teachers), but their production and their explanations of what made an utterance polite were not. In their explanations, for example, children took into account of uses of *please*, the presence of modal verbs (*could*, *would*, *may*), formal or informal address (first name vs. title and family name), and direct vs. indirect requests in deciding which of two requests would be more successful.

By age 4–5, children can readily do the voices for dolls representing an adult, a peer, and a younger child. James (1978) elicited directives and requests from children aged 4;6 to 5;2, and found that age was the most important factor in their production of directives: Adults received the politest forms, followed by peers and then younger children. But when these children were asking a favor instead, all the requests they produced were more polite (see also Shatz and Gelman, 1973; Andersen, 1990; Ervin-Tripp et al., 1990).

Persuasion often accompanies requests, as when children spontaneously attempt to get a parent to buy a toy they have seen (Axia, 1996). Four-year-olds are fairly direct in their requests here, and typically don’t anticipate adult reactions when they ask. Six-year-olds are more strategic and take some anticipated adult objections into account when they make their requests. But it is not until age 8 that children both anticipate adult objections and tailor their requests in a cooperative way as they bargain for what they want.

Telling stories

Children start to tell stories early on, minimalist stories about what they have been doing. From age 3 or so, these stories are often prefaced by “*You know what?*” followed by two or three narrative utterances. Later, children often adopt a formulaic starting point such as *Once upon a time ...* as they start in on their telling (e.g., Umiker-Sebeok, 1979; Stone, 1992; Berman and Slobin, 1994). But good story telling takes more planning than this. Children must learn to identify the setting of the action, the main and subsidiary characters, each event in the action, and the culminating event that ends the story. This means keeping track of each character as well as any background information, and the role each plays in the actual sequence of events (see Karmiloff-Smith, 1981).

The planning required to do all this appears to be a lot more complex than simply planning one’s next utterance in a conversational exchange. Children take time to learn how to do this, even with simple stories where they have a picture book like Meyer’s *Frog, where are you?* in front of them, with each page acting as a prompt. At age 5, children’s stories are still often quite rudimentary in form, but as they get older, they get better at keeping track of all the characters, and in particular of the main character, in the events being recounted. They learn to add background information about the setting and about the emotions involved, and they present a more coherent account of the sequence of events as they get older. The same trends can be seen cross-linguistically, as children pick up on the conventions within each language for how to tell a story (Berman and Slobin, 1994).

Crosslinguistic comparisons

While most of the illustrations here have been taken from children acquiring English, children acquiring other languages go through very similar stages, and appear to apply similar strategies in how they communicate as they acquire the language they are being exposed to. Slobin (1985-1997, in five volumes) presented research on English, German, Hebrew, Hungarian, Japanese, Kaluli, Polish, Romance languages with emphasis on French, Samoan, Turkish, and American Sign Language (volumes 1 and 2); Georgian, West Greenlandic, K’iche Maya, and Warlpiri, as well as Mandarin, the Scandinavian languages, and Sesotho (vol. 3), and Finnish, Estonian, and Hungarian compared, Greek, and Korean (volume 4). These volumes offer an important starting point for looking at similarities and differences by language and cultural context. And, as Slobin pointed out, the general approach children appear to take to language, the strategies they apply during acquisition, are very similar. This work has encouraged the collection of acquisition data from a variety of additional settings, including indigenous rural communities in Mexico for Tseltal (Casillas et al., 2020), and in Papua New Guinea for Yelî Dnye (Casillas et al., 2021). There has also been considerable research on the acquisition of indigenous Australian languages (see Kelly et al., 2015) as well as on other less well-studied languages (see, e.g., Jancso, et al., 2020). In short, recent years have seen considerable progress in the documentation of just what children know when about the particular language they are in the process of acquiring. Such data offer valuable insights into the role of child-rearing beliefs, for example, and cultural norms, as they are applied to young children in diverse settings. They also reveal similarities in the stages children go through as they acquire a first language.

Conclusion

By age 5, most children can use language appropriately in a variety of settings, with a variety of addressees. They have some sense of differences in understanding and in talk for 2-year-olds vs. their own peers. They are aware that speakers differ in their use of language depending on age and status, and they associate different degrees of politeness with different social roles depending on age, status, and relative power. They can plan appropriate answers to questions, and can tell personal stories, as well as attempt to re-tell familiar tales. They can give “stage directions” for different roles in pretend play, adjusting these on the fly as the “play” proceeds. They can be persuasive, and justify their requests. And they can keep track of what the other person knows as they pursue a conversation. By age 5, children are on their way to becoming more fluent speakers, speakers who can take into account what their addressees already know, and they are becoming more skilled in how to use language in a variety of settings, for a variety of purposes, with a variety of different people.

References

- Andersen, E.S., 1990. *Speaking With Style*. Routledge, London.
- Arriaga, R.I., Fenson, L., Cronan, T., Pethick, S.J., 1998. Scores on the McArthur Communicative Development Inventory of children from low and middle income levels. *Appl. Psycholinguist.* 19 (2), 209–223.
- Axia, G., 1996. How to persuade mum to buy a toy. *First Lang.* 16 (48), 301–317.
- Baker, N.D., Greenfield, P.M., 1988. The development of new and old information young children’s early language. *Lang. Sci.* 10, 3–34.
- Bates, E., 1976. *Language and Context: The Acquisition of Pragmatics*. Academic Press, New York.
- Bates, E., Silvern, L., 1977. Social adjustment and politeness in preschoolers. *J. Commun.* 27, 104–111.
- Bergelson, E., Swingle, D., 2015. Early word comprehension in infants: replication and extension. *Lang. Learn. Dev.* 11 (4), 369–380.
- Berman, R.A., Clark, E.V., 1989. Learning to use compounds for contrast: data from Hebrew. *First Lang.* 9 (27), 247–270.
- Berman, R.A., Slobin, D.I., 1994. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Erlbaum, Hillsdale, NJ.
- Billow, R.M., 1981. Observing spontaneous metaphor in children. *J. Exp. Child Psychol.* 31 (3), 430–445.
- Bloom, L., 1973. *One Word at a Time*. Mouton, The Hague.
- Bowerman, M., 1973. *Early Syntactic Development: A Cross-Linguistic Study With Special Reference to Finnish*. Cambridge University Press, Cambridge.
- Braine, M.D.S., 1976. Children’s first word combinations. *Monogr. Soc. Res. Child Dev.* 41. Serial No. 164.
- Brown, R.W., 1973. *A First Language*. Harvard University Press, Cambridge, MA.
- Bruner, J.S., 1983. *Child’s Talk: Learning to Use Language*. Norton, New York.
- Casillas, M., Bobb, S.C., Clark, E.V., 2016. Turn-taking, timing, and planning in early language acquisition. *J. Child Lang.* 43 (6), 1310–1337.
- Casillas, M., Brown, P., Levinson, S.C., 2020. Early language experience in a Tzeltal Mayan village. *Child Dev* 91, 1819–1835.
- Casillas, M., Brown, P., Levinson, S.C., 2021. Early language experience in a Papuan community. *J. Child Lang.* 48, 792–814.
- Chouinard, M.M., 2007. Children’s questions: a mechanism for cognitive development. *Monogr. Soc. Res. Child Dev.* 72. Serial No. 286.
- Chouinard, M.M., Clark, E.V., 2003. Adult reformulations of child errors as negative evidence. *J. Child Lang.* 30 (3), 637–669.
- Clark, E.V., 1971. On the acquisition of the meaning of before and after. *J. Verb. Learn. Verb. Behav.* 10 (3), 266–275.
- Clark, E.V., 1973. How children describe time and order. In: Ferguson, C.A., Slobin, D.I. (Eds.), *Studies of Child Language Development*. Holt, Rinehart and Winston, New York, pp. 585–606.
- Clark, E.V., 1978. From gesture to word: on the natural history of deixis in language acquisition. In: Bruner, J.S., Garton, A. (Eds.), *Human Growth and Development: Wolfson College Lectures 1976*. Oxford University Press, Oxford, pp. 95–120.
- Clark, E.V., 1982. The young word maker: a case study of innovation in the child’s lexicon. In: Wanner, E., Gleitman, L.R. (Eds.), *Language Acquisition: The State of the Art*. Cambridge University Press, Cambridge, pp. 390–425.
- Clark, E.V., 1993. *The Lexicon in Acquisition*. Cambridge University Press, Cambridge.
- Clark, E.V., 1995. Later lexical development and word formation. In: Fletcher, P., MacWhinney, B. (Eds.), *The Handbook of Child Language*. Blackwell, Oxford, pp. 393–412.
- Clark, E.V., 1997. Conceptual perspective and lexical choice in acquisition. *Cognition* 64 (1), 1–37.
- Clark, E.V., 2007. Young children’s uptake of new words in conversation. *Lang. Soc.* 36 (2), 157–182.
- Clark, E.V., 2010a. Adult offer, word-class, and child uptake in early lexical acquisition. *First Lang.* 30 (3–4), 250–269.
- Clark, E.V., 2010b. Learning a language the way it is: conventionality and semantic domains. In: Malt, B., Wolff, P. (Eds.), *Words and the Mind: How Words Capture Human Experience*. Oxford University Press, Oxford, pp. 243–265.
- Clark, E.V., 2015. Common ground. In: MacWhinney, B., O’Grady, W. (Eds.), *The Handbook of Language Emergence*. Wiley, New York, pp. 328–353.
- Clark, E.V., 2016. *First Language Acquisition*, third ed. Cambridge University Press, Cambridge.
- Clark, E.V., 2018a. Conversation and language acquisition: a pragmatic approach. *Lang. Learn. Dev.* 14 (3), 170–185.
- Clark, E.V., 2018b. Acquiring meanings and meaning domains. In: Syrett, K., Arunachalam, S. (Eds.), *Semantics in Language Acquisition*. John Benjamins, Amsterdam, pp. 21–43.
- Clark, E.V., 2019a. Research on first language acquisition: a brief history. In: Horst, J., Torkildsen, J. von K. (Eds.), *International Handbook of Language Development*. Routledge, London, pp. 3–19.
- Clark, E.V., 2019b. The acquisition of word-formation in Romance. *Romance linguistics*. In: Gardani, F., Loporcaro, M. (Eds.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press, Oxford. <https://doi.org/10.1093/acrefore/9780199384655.013.680>. Published online 8/2019, 42pp.
- Clark, E.V., 2020a. Perspective-taking and pretend-play: precursors to figurative language use in young children. *J. Pragmat.* 156, 100–109.
- Clark, E.V., 2020b. Conversational repair and the acquisition of language. *Discourse Process.* 57 (5–6), 441–459.
- Clark, E.V., 2020c. *Language as Expertise*. Ms. Stanford University, 110pp.
- Clark, E. V., 2022. *Language as graded expertise*. Ms. Stanford University, 180pp.
- Clark, E.V., Garnica, O.K., 1974. Is he coming or going? On the acquisition of deictic verbs. *J. Mem. Lang.* 13 (5), 559–572.
- Clark, E.V., Sengul, C.J., 1978. Strategies in the acquisition of deixis. *J. Child Lang.* 5 (3), 457–475.
- Clark, E.V., Hecht, B.F., 1982. Learning to coin agent and instrument nouns. *Cognition* 12 (1), 1–24.
- Clark, E.V., Hecht, B.F., 1983. Comprehension, production, and language acquisition. *Annu. Rev. Psychol.* 34 (1), 325–349.
- Clark, E.V., Cohen, S., 1984. Productivity and memory for newly formed words. *J. Child Lang.* 11 (3), 611–625.
- Clark, E.V., Gelman, S.A., Lane, N.M., 1985. Compound nouns and category structure in young children. *Child Dev.* 56 (1), 84–94.

- Clark, E.V., Bowerman, M., 1986. On the acquisition of final voiced stops. In: Fishman, J.A. (Ed.), *The Fergusonian Impact, From Phonology to Society*, vol. 1. Mouton, Amsterdam, pp. 51–68.
- Clark, E.V., Berman, R.A., 1987. Types of linguistic knowledge: interpreting and producing compound nouns. *J. Child Lang.* 14 (3), 547–567.
- Clark, E. V., Carpenter, K. L., Deutsch, W., 1995. Reference states and reversals: undoing actions with verbs. *J. Child Lang.* 22 (3), 633–662.
- Clark, E.V., Wong, A.D.-W., 2002. Pragmatic directions about language use: offers of words and relations. *Lang. Soc.* 31 (2), 181–212.
- Clark, E.V., Nikitina, T.V., 2009. One vs. more than one: antecedents to plural marking in early language acquisition. *Linguistics* 47 (1), 103–139.
- Clark, E.V., Estigarribia, B., 2011. Using speech and gesture to introduce new objects to young children. *Gesture* 11 (1), 1–23.
- Clark, E.V., de Marneffe, M.C., 2012. Constructing verb paradigms in French: adult construals and emerging grammatical contrasts. *Morphol* 22, 89–120.
- Clark, E.V., Lindsey, K., 2015. Turn-taking: a case study of early gesture and word use in answering WHERE and WHICH questions. *Front. Psychol.* 6.
- Clark, E.V., Kelly, B.F., 2022. Constructing a system of communication: using gestures and words. In: Goldin-Meadow, A., Morgenstern, A. (Eds.), *Gesture in Language*. Walter de Gruyter, Berlin.
- Clark, H.H., 1996. *Using Language*. Cambridge University Press, Cambridge.
- Diessel, H., Tomasello, M., 2000. The development of relative clauses in spontaneous child speech. *Cognit. Ling.* 11 (1/2), 131–152.
- Donaldson, M., Balfour, G., 1968. Less is more: a study of language comprehension in children. *Br. J. Psychol.* 59 (4), 461–471.
- Enfield, N.J., Levinson, S.C., 2006. *Roots of Human Sociality*. Berg, New York.
- Ervin-Tripp, S., Guo, J., Lampert, M., 1990. Politeness and persuasion in children's control acts. *J. Pragmat.* 14 (2), 307–331.
- Estigarribia, B., Clark, E.V., 2007. Getting and maintaining attention in talk to young children. *J. Child Lang.* 34 (4), 799–814.
- Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D., Pethick, S., 1994. Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59 (5). Serial No. 242.
- Ferrier, L., 1978. Some observations of error in context. In: Waterson, N., Snow, C. (Eds.), *The Development of Communication*. Wiley, New York, pp. 301–309.
- Garvey, C., Berninger, G., 1981. Timing and turn taking in children's conversations. *Discourse Process.* 4 (1), 27–57.
- Garvey, C., Kramer, T.L., 1989. The language of social pretend play. *Dev. Rev.* 9 (4), 364–382.
- Gelman, S.A., Coley, J.D., Rosengren, K.S., Hartman, E., Pappas, A., 1998. Beyond labeling: the role of maternal input in the acquisition of richly structured categories. *Monogr. Soc. Res. Child Dev.* 63. Serial No. 253.
- Gleason, J.B., Perlman, R.Y., Greif, E.B., 1984. What's the magic word: learning language through politeness routines. *Discourse Process.* 7 (4), 493–502.
- Greenfield, P.M., Smith, J.H., 1976. *Structure of Communication in Early Language Development*. Academic Press, New York.
- Greif, E.B., Gleason, J.B., 1980. Hi, thanks, and goodbye: more routine information. *Lang. Soc.* 9, 159–166.
- Hart, B., Risley, T.R., 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Paul H. Brookes Publishing, Baltimore, MD.
- Heath, S.B., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.
- Hillbrink, E.E., Gattis, M., Levinson, S.C., 2015. Early developmental changes in the timing of turn-taking: a longitudinal study of mother-infant interaction. *Front. Psychol.* 6.
- Hurtado, N., Marchman, V.A., Fernald, A., 2008. Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Dev. Sci.* 11 (6), F31–F39.
- James, S., 1978. Effect of listener age and situation on the politeness of children's directives. *J. Psycholinguist. Res.* 7 (4), 307–317.
- Jancso, A., Moran, S., Stoll, S., 2020. The ACQDIV corpus database and aggregation pipeline. In: *Proceedings of the 12th Language Resources and Evaluation Conference*, pp. 156–165.
- Jourdain, M., Lahousse, K., 2021. The development of constructions from the right edge: a multinomial regression analysis of clitic left and right dislocation in child French. *J. Child Lang.* 45, 1023–1047. <https://doi.org/10.1017/S0305000920000513>.
- Kaper, W., 1980. The use of past tense in games of pretend. *J. Child Lang.* 7 (1), 213–215.
- Karmiloff-Smith, A., 1981. The grammatical marking of thematic structure in the development of language production. In: Deutsch, W. (Ed.), *The Child's Construction of Language*. Academic Press, London, pp. 121–147.
- Kelly, B.F., Forshaw, W., Nordlinger, R., Wigglesworth, G., 2015. Linguistic diversity in first language acquisition research: moving beyond the challenges. *First Lang* 35, 286–304.
- Kemler Nelson, D.G., Egan, L.C., Holt, M.B., 2004. When children ask, "What is it?", What do they want to know about artifacts? *Psychol. Sci.* 15 (6), 384–389.
- Kempson, R., Cann, R., Gregorimichelaki, E., Chatzikyriakidis, S., 2016. Language as mechanisms for interaction. *Theor. Ling.* 42 (3–4), 203–276.
- Klatzky, R.L., Clark, E.V., Macken, M., 1973. Asymmetries in the acquisition of polar adjectives: linguistic or conceptual? *J. Exp. Child Psychol.* 16 (1), 32–46.
- Knudsen, B., Liszkowski, U., 2013. One-year-olds warn others about negative action outcomes. *J. Cognit. Dev.* 14 (3), 424–436.
- Köder, F., Falkum, I.L., 2020. Children's metonymy comprehension: evidence from eye-tracking and picture selection. *J. Pragmat.* 156, 191–205.
- Lempers, J., Flavell, E.L., Flavell, J.H., 1977. The development in very young children of tacit knowledge concerning visual perception. *Genet. Psychol. Monogr.* 95, 3–53.
- Liszkowski, U., 2005. Human twelve-month-olds point cooperatively to share interest with and helpfully provide information for a communicative partner. *Gesture* 5, 135–154.
- Lustigman, L., Clark, E.V., 2019. Exposure and feedback in language acquisition: adult construals of children's early verb-form uses in Hebrew. *J. Child Lang.* 46, 241–264.
- Marcos, H., Salazar Orvig, A., da Silva-Genest, C., Heudier, J., 2021. The influence of dialogue in young children's uses of referring expressions. In: Salazar Orvig, A., de Weck, G., Hassan, R., Rialland, A. (Eds.), *The Acquisition of Referring Expressions: A Dialogical Approach*. John Benjamins, Amsterdam, pp. 204–232.
- Moore, C., Dunham, P.J. (Eds.), 2014. *Joint Attention: Its Origins and Role in Development*. Erlbaum, Hillsdale, NJ.
- Morgenstern, A., Sekali, M., 1999. Processus de clarification de la référence dans le dialogue adulte-enfant. In: *Travaux linguistiques du CERLICO : La référence, 2: Statut et processus*. Presses Universitaires de Rennes, pp. 313–333.
- Noble, C.H., Cameron-Faulkner, T., Lieven, E., 2018. Keeping it simple: the grammatical properties of shared book reading. *J. Child Lang.* 45 (3), 753–766.
- O'Neill, D.K., 1996. Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Dev.* 67 (2), 659–677.
- Pouscoulous, N., Tomasello, M., 2020. Early birds: metaphor understanding in 3-year-olds. *J. Pragmat.* 156, 160–167.
- Reilly, J.S., 1986. The acquisition of temporals and conditionals. In: Traugott, E.C., ter Meulen, A., Reilly, J.S., Ferguson, C.A. (Eds.), *On Conditionals*. Cambridge University Press, Cambridge, pp. 309–331.
- Sachs, J., Devin, J., 1976. Young children's use of age-appropriate speech styles in social interaction and role-playing. *J. Child Lang.* 3 (1), 81–98.
- Saffran, J.R., Johnson, E.K., Aslin, R.N., Newport, E.L., 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* 70 (1), 27–52.
- Sawyer, R.K., 1997. *Pretend Play as Improvisation: Conversation in the Preschool Classroom*. Psychology Press, Hove.
- Scollon, R., 1976. *Conversations with a one-year-old*. University of Hawaii Press, Honolulu, HA.
- Seidenberg, M., 2017. *Language at the Speed of Sight: How We Read, Why So Many Can't, and What Can Be Done About It*. Basic Books, New York.
- Shatz, M., Gelman, R., 1973. The development of communication skills: modifications in the speech of young children as a function of listener. *Monogr. Soc. Res. Child Dev.* 38 (5). Serial No. 152.
- Sheldon, A., 1974. The role of parallel function in the acquisition of relative clauses in English. *J. Verb. Learn. Verb. Behav.* 13 (3), 272–281.
- Slobin, D.I., 1973. Cognitive prerequisites for the development of grammar. In: Ferguson, C.A., Slobin, D.I. (Eds.), *Studies of Child Language Development*. Holt, Rinehart and Winston, New York, pp. 175–208.
- Snow, C., 1977. The development of conversation between mothers and babies. *J. Child Lang.* 4 (1), 1–22.
- Snow, C.E., Matthews, T.J., 2016. Reading and language in the early grades. *Future Child.* 26 (2), 57–74. Starting early: Education from pre-Kindergarten to Third Grade.

- Stivers, T., Sidnell, J., Bergen, C., 2018. Children's responses to questions in peer interaction: a window into the ontogenesis of interactional competence. *J. Pragmat.* 124, 14–30.
- Stone, P.S., 1992. "You know what?" Conversational narratives of preschool children. *Early Child. Res. Q.* 7 (3), 367–382.
- Strömqvist, S., 1984. *Make-Belief Through Words: A Linguistic Study in Children's Play With a Doll's House*. Unpublished PhD dissertation. University of Göteborg, Sweden.
- Tamis-LeMonda, C.S., Custode, S., Kuchirko, Y., Escobar, K., Lo, T., 2019. Routine language: infant-directed speech during everyday activities. *Child Dev.* 90, 2135–2152.
- Thoermer, C., Sodian, B., 2001. Preverbal infants' understanding of referential gestures. *First Lang.* 21 (63), 245–264.
- Tyack, D., Ingram, D., 1977. Children's production and comprehension of questions. *J. Child Lang.* 4 (2), 211–224.
- Umiker-Sebeok, D.J., 1979. Preschool children's intraconversational narratives. *J. Child Lang.* 6 (1), 91–109.
- Werker, J.F., Tees, R.C., 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7 (1), 49–63.

Further reading

- Bloom, P., 2000. *How Children Learn the Meanings of Words*. MIT, Cambridge, MA.
- Clark, H.H., 1996. *Using Language*. Cambridge University Press, Cambridge.
- Clark, E.V., 2016. *First Language Acquisition*, third ed. Cambridge University Press, Cambridge.
- Clark, E.V., 2017. *Language in Children*. Routledge and Linguistic Society of America, London.
- Frank, M., Braginsky, M., Marchman, V., Yurovsky, D., (2021). *Variability and Consistency in Early Language Learning: The Wordbank Project*. Cambridge, MA: MIT.
- Malt, B., Wolff, P. (Eds.), 2010. *Words and the Mind: How Words Capture Human Experience*. Oxford University Press, Oxford.
- Slobin, D.I. (Ed.), 1985–1997. *The Crosslinguistic Study of Language Acquisition*, 5 vols. Erlbaum, Hillsdale, NJ.
- Syrett, K., Arunachalam, S. (Eds.), 2018. *Semantics in Language Acquisition*. John Benjamins, Amsterdam.
- Tomasello, M., 2003. *Constructing a Language*. Harvard University Press, Cambridge, MA.
- Wanner, E., Gleitman, L.R. (Eds.), 1982. *Language Acquisition: The State of the Art*. Cambridge University Press, Cambridge.

Relevant websites

- MacWhinney, B., 2000. *The CHILDES Project: Tools for Analyzing Talk*, third ed. Erlbaum, Mahwah, NJ <http://childes.psy.cmu.edu/>.
<http://wordbank.stanford.edu/>.

Dynamic system approaches to language acquisition

Marijn van Dijk and Paul van Geert, Heymans Institute for Psychological Science, Department of Developmental Psychology, University of Groningen, Groningen, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Dynamic systems approaches to language acquisition	14
State spaces and trajectories	14
Self-organization and emergence	15
Layers of organization and time scales	16
Process causality	16
Variability is an intrinsic feature of a complex system's trajectory	16
Implications for research methodology	16
Findings from studies on first language acquisition	17
Bidirectional interactions and collective variables in L1 acquisition	17
Self-organization and emergence in L1 acquisition	18
Nonlinearity and variability in L1 acquisition	18
Modeling L1 acquisition	20
Findings from studies on second language acquisition	21
Interacting variables in L2 acquisition	21
Nonlinearity and variability in L2 acquisition	21
General discussion	22
Summary	22
Critical reflection	23
Future directions	24
References	24

Dynamic systems approaches to language acquisition

Though language has been studied for centuries, many questions with regard to its forms, evolutionary pathway, and acquisition are still unanswered. One of the main problems of language acquisition is whether it is a matter of transmission and appropriation or whether it is brought about by processes of self-organization and emergence. In the first case, the properties of a language are being transferred from a mature speaker to a novice learner or from an innate universal system to a manifest language use. In the second case, language emerges as a result of a child's use of it in a context of activity with other language users. In this chapter, we will describe a specific interpretation of this latter perspective on language emergence and elaborate on dynamic systems approaches to language acquisition. We will start by providing a general description of this theoretical framework, explain its main concepts and discuss the implications for research. After this, we will provide a short narrative review of the main empirical findings of dynamic systems-oriented research in both first and second language acquisition. We will finish this chapter by critically reflecting on what the dynamic systems approach has yielded and discuss some future directions.

State spaces and trajectories

Generally speaking, dynamic systems approaches specify how something changes over the course of time. This can be any kind of ability, such as a child's movements, understanding of the world, and language use. If something changes, there is a current state, and there is a state into which it changes. This "something" is defined by the set of all its possible states, for instance all possible states of a child's language use. This set of all these possible states is called the system's **state space**. A state space consists of a hypothetically unlimited number of dimensions (one for each of the possible properties of the states), and the number of states may vary between 1 and virtually infinite, depending on the choice of the researcher. For instance, if we focus only on the words in an utterance, we can use a one-dimensional state space that captures the number of words on a single axis. However, this one-dimensional state space can easily be extended to include other dimensions, such as the syntactic nature of the utterances, the word types that are used, etcetera.

At each particular moment in time, a system occupies one particular point in the state space, represented by the values it has on each of the quantitative or categorical dimensions. For instance, when a particular child uses the sentence "want milk," he or she occupies the state of the two-word utterances in the dimension of utterance length and the state of "present tense" in the dimension of verb tense (see Fig. 1 for an illustration). During a single conversation, the child may use mainly one and two words utterances and each utterance represents a consecutive state in the state space. That is, during the conversation, the child only moves within a limited region of the state space.

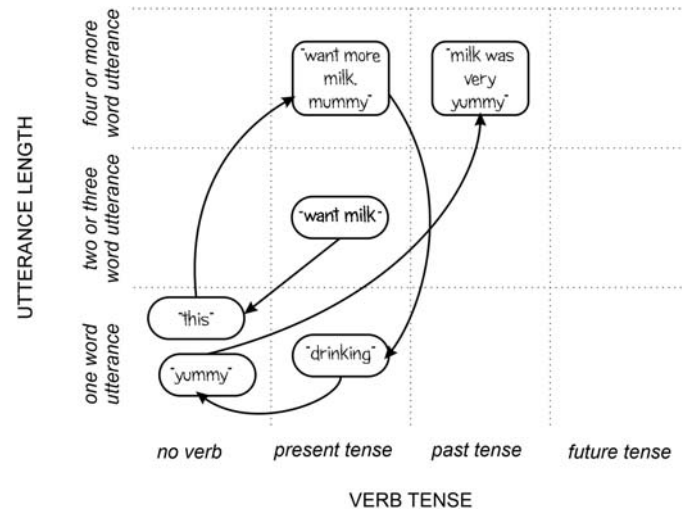


Fig. 1 State space representation of the utterances of a child in a conversation.

If we connect all the points of a particular child's conversations over time, we obtain a line or **trajectory** in this state space. This is represented by the lines in **Fig. 1**, where the arrows show the sequence of events. It should be noted that the lines do not represent a "fit" of any kind between datapoints; they merely connect each event with each consecutive event. However, the sequence of data-points and lines between them can be used for further statistical analysis (e.g. cluster analysis, temporal analysis). It is likely that the trajectory will be irregular, but that this variability will be confined to a particular region of the state space. If we wait long enough, the trajectory will probably have moved either from one region in the state space to another, e.g. from dominance of one-word utterances to the dominance of multiword utterances or from a narrow to a much larger region, e.g. from one-word utterances only to utterances between one and four words. The trajectory between these regions can be relatively smooth or quite erratic, with gradual versus sudden changes. In dynamic systems theory, this real trajectory, with all its temporal properties, is the phenomenon of interest.

Self-organization and emergence

According to a dynamic systems approach, developmental trajectories are not imposed in the form of an in-built design or explicit external control. Instead, they result from the dynamics of the interacting components: they are self-organized patterns. **Self-organization** is a core feature of language development. Since the components in a system (or state variables) form a network of connected and interacting ("coupled") nodes, all components work together to generate effects on the system. An example of such coupling in the domain of child language acquisition exists in the interaction between the language of the child and the child-directed speech of the parent. From the literature we know that parent and child coordinate their language during a conversation and over conversations (see for instance [van Dijk et al., 2013](#)). The principle of coupling is easy to imagine with components such as parent and child, but holds for almost any set of components, e.g. meaning and form, verbs and nouns, nouns and determiners, etc. These interactions are often bidirectional (or cyclical), for instance the parent's child directed speech has an effect on the child's speech and vice versa. Some of the couplings between components are supportive whereas others are competitive ([Van Geert, 2003](#)). The relations can also be symmetric or asymmetric, direct or indirect (run via other components), and more or less complex (for instance, they can be pre-conditional, non-linear, etc.).

A complex system typically consists of many interacting nodes. The interactions of the coupled components of such a complex system generate **emergent properties**. This means that these properties are novel and cannot be reduced to an addition or superposition of the properties of their components. For instance, usage-based theories of language acquisition show how a syntactic category such as "noun" or "verb" may emerge out of processes of language perception and production, cognitive processes of statistical frequency learning and so forth. Once such categories have emerged, they interact with the already existing processes of language perception and production. Emergence is of fundamental importance for all developing systems.

The concepts of emergence and self-organization are often used interchangeably, probably because their typical effects co-occur. Self-organization implies that the interactions between system components without external control lead to self-maintaining patterns of interaction (in a relatively small region of the state space). Such patterns can be the cause of novel properties that become part of the system (emergence).

Given the dynamics that govern the interactions between their components, many dynamic systems move toward self-reproductive and self-maintaining states. Such states are dynamically stable, and are being referred to by the term "attractors." The construction of sentences on the basis of a limited set of syntactic rules can be seen as an attractor, which is achieved through a trajectory of producing one-word sentences, to forming simple sentences based on combining words, and finally to using

grammatical rules. A system may have one stable attractor or several, it may cycle between attractors, or show attractor states in the form of complex repetitive temporal patterns (i.e. recurrent patterns).

Layers of organization and time scales

Complex systems self-organize into interacting layers of organization. This means that each of the system's components can be conceived of as a complex system itself, all the way down to the most elementary quantum physical interactions. This also applies to its time course or temporal evolution. For instance, in language, there are processes on the level of utterance formation (for instance all the physical and bodily processes that are involved in saying "I want more milk") and there are processes on the level of conversations which consist of sequences of interacting utterances (e.g. a conversation about milk). There are also processes on the level of language acquisition itself, which consists of changes in form and content of conversations, corresponding among others with changes in form and content of utterances (from "want milk" to "I want more milk"). These events, which are in fact patterns of change, are typically **nonlinear**. Nonlinearity implies that causes and effects are not proportional or additive. Examples of important nonlinear effects are sudden rapid changes, regressions, phase shifts and discontinuities. For instance, in the case of language acquisition of a particular child, a certain correction by a parent might have virtually no effect at one moment in time. But for another moment in its development, such corrections might have huge effects: they push the system over the edge or make it tip, metaphorically speaking. This "pushing over the edge" can refer to the concept of a phase shift (indicating a rapid transition between developmental stages) or even a discontinuity (indicating an abrupt change between discrete stages). These typical nonlinear effects occur on their own proper timescale, for instance the timescale of months and years in the case of the acquisition of syntax or the timescale of seconds or minutes that governs the dynamics of utterance production.

Process causality

Self-organization and emergence imply a particular view on causality, called "**process causality**." This is fundamentally different from the causal model we are used to, namely that of "intervention causality," which defines a cause as a specific intervention or natural event producing a specific effect. Take for instance the question of whether parents' reading books to their toddlers are a cause of their toddlers' accelerated vocabulary development. Assume that we have found a statistically significant difference in average vocabulary between a reading and non-reading group, is reading the cause of the effect, greater vocabulary? Interpreting this finding, we must first, realize that the cause is an iterative process, that is a sequence of connected events (e.g. the child's asking to be read, the ritual of reading before going to bed, the routine of doing so over a great number of days, etc.). Second, because this is a coupled dynamic system, the causal process is bidirectional or cyclical. The reading, for instance its continuation over a sufficient period of time, depends on the child's interest, attention, willingness to participate, eagerness or lack thereof and so forth. Third, because this is a complex system, the "behavior" of the system is an emergent property of the interactions between all the relevant components. In the case of reading and vocabulary development, those components involve material conditions, housing, availability of age-appropriate books, presence of other children, availability of other activities to spend time with, reading done by parents and siblings and so on. This means that the process is multi-causal, in spite of the fact that the intervention (in this case the reading by the parents) suggests a mono-causal event (it is only the reading that was manipulated). Fourth, the causality is contextual, which means that it is something that occurs in concrete, specific complex systems, e.g. concrete families. Given that such systems are usually highly **idiosyncratic** (family- or person-specific), the causal process itself is also likely to be highly idiosyncratic. Group averages, for instance over a sample of parent-child dyads, should not be confused with the real causal effect. Because development is such an idiosyncratic and nonlinear process, there is also a large degree of unpredictability on the level of individual parent-child dyads.

Variability is an intrinsic feature of a complex system's trajectory

We already noticed that observed trajectories in state spaces display **variability** over time, over different timescales. For instance, the number of words in each utterance will fluctuate, but there will also be fluctuations in the average number of words over consecutive conversations. According to a dynamic systems interpretation of development, this variability is a real and developmentally important property of the system. It emerges out of its multi-causal structure and is a typical indicator of its complexity. For instance, temporal variability provides contexts of exploration, parental modeling or correction. It is important to note that the cause-and-effect relationship between variability and development is complex: variability and context create each other and drive development.

Variability has a temporal structure that is different from random variation about a central tendency, and that is not a linear function of the known variables. It has been proposed that a "fractal" pattern of variability (a pattern which has a degree of "self-similarity") is optimal in the sense that it suggests coordination.

Implications for research methodology

The dynamic systems theoretical framework—with its explicit focus on the system as a whole, with nested layers and timescales, bidirectional interactions, nonlinearity, variability, and idiosyncratic pathways—has clear implications for the way development

in general and language acquisition in specific should be investigated. One of the main consequences is that it should in principle focus on the system as a whole, and thus include all factors that are part of it and that influence it (see [Larsen-Freeman and Cameron, 2008](#)). Context should not be seen as an external source of error, but as part of the complex system that provides important information. A dynamic approach demands a holistic perspective on the system, instead of attempting to take it apart. In practice, however, including all relevant variables of a system in one specific study is impossible, but researchers should specify at least the most important variables and the way they interact with each other. When selecting these variables, they must be treated as collective representations of many underlying components that are codependent on one another rather than treating them as independent variables that can be isolated from one another. Researchers should also rethink the units of analysis they choose, for instance by identifying collective variables that characterize the interaction among multiple elements over time.

A second implication is that researchers should think in terms of dynamic processes and changing relationships among variables (see also [Larsen-Freeman and Cameron, 2008](#), p. 206). This means moving away from analyzing unidirectional and/or linear influences from “independent” variables onto “dependent” variable(s). Instead, researchers should always consider bidirectional (reciprocal, cyclical) causality and nonlinear interactions between layers of development and timescales. In fact, studies should explicitly seek linkages across these layers and timescales.

Third, because of the idiosyncratic nature of development, analyses must be conducted at the level of the individuals (that is: trajectories of individual learners, individual dyads, individual classrooms, etc.). Analyses should focus on nonlinear forms of development, and on intra-individual variability as a central property. This means that the time-dependent properties should be seen as fundamental, and not be treated as noise added to underlying, stable or slowly and smoothly changing traits or variables. Currently, several statistical techniques are available to investigate the temporal patterns of variability in data with long time series, such as spectral and fractal analysis and recurrence quantification analysis.

There are various types of studies in the field of language acquisition that are compatible with these recommendations. These include longitudinal (multiple) case studies with extended and frequent measurement points, and studies on micro-development (short time serial design with intense measurements; conversation analysis). Dynamic systems research is also well known for its model building and simulation studies in which the underlying relations between components/variables are described in the form of coupled mathematical equations. These models are based on dynamic assumptions, such as the inclusion of the most relevant variables of the system, the focus on dynamic processes, including bidirectional causality, the investigation of linkages between time scales, and the emergence of higher order stability through interactions at the micro level. According to [Larsen-Freeman and Cameron \(2008\)](#) experimental designs can also qualify as dynamic studies when they pay attention to the real-time processes of change, are ecologically valid and keep away from reductionistic thinking.

Findings from studies on first language acquisition

Various publications on L1 development explicitly refer either to “dynamic systems,” “complex systems,” or to “DST.” Though the majority of these contributions are solely theoretical (e.g. [Larsen-Freeman and Cameron, 2008](#)), several studies cite empirical work published elsewhere (e.g. [Hockema and Smith, 2009](#)), or contain original data that are analyzed/interpreted in a dynamic way (e.g. [Hohenberger and Peltzer-Karpf, 2009](#)). With regard to those studies that either include or refer to empirical data, many of these explicitly name various dynamic constructs—such as self-organization, variability and non-linearity—as being relevant for first language development. In the empirical sections of these publications, there is often a focus on only one or two dynamic aspects which existence is being demonstrated.

In the sections below, we will organize some of the main empirical findings and directions along the lines of the implications for research as describe above. First, we will review several important studies on first language acquisition with an explicit focus on the system as a whole either by investigating the interactions themselves or by using collective variables. Second, we will discuss studies with an explicit focus on self-organization and emergence as central phenomena (for instance by investigating layers or time scales). Third, we will review studies directed at non-linearity and variability in trajectories of first language acquisition. Finally, we will refer to some of the model simulation studies and report what they have taught us.

Bidirectional interactions and collective variables in L1 acquisition

Various studies have offered descriptions of reciprocal influences and/or have analyzed collective variables to describe and explain developmental pathways. As an example, in [Van Dijk et al. \(2013\)](#), we have presented a coupled growth equation model on the **bidirectional influence** between early child language and child directed speech by the parent. This model describes the interaction between the language input from the parent and language growth in the child. The model was validated with developmental data of three children and those of the CDS of their caregivers (MLU and vocabulary) and was shown to have a good fit for all variables, except for one. These results can be seen as “proof of principle” for the existence of a dynamic form of adaptation between child-language and child directed speech. This study is a typical example of a dynamic systems interpretation because it describes mutually interacting variables that are treated as collective representations of underlying interacting complex systems.

With regard to using **collective variables**, [Hsu and Fogel \(2001\)](#) investigated infant vocal development as a precursor to language development. Based on observations of infants and their mothers in a face-to-face interaction, three global patterns of communication dynamics were classified: symmetrical (mutual engagement), asymmetrical (mother active/infant inactive), and unilateral

(mother active/infant disengaged). Further analysis showed that the quantity and quality of infant vocalizations were associated with these patterns of mother-infant communication and that these associations changed developmentally. Whereas a curvilinear growth pattern of the rate of syllabic vocalizations was found for symmetrical communication, a linear increase was found in unilateral communication. This study demonstrates the use of collective variables, but also shows how these are related to longer-term developmental pathways.

Parladé and Iverson (2011) also combined the use of collective variables (in this case: temporarily coordinated communication bouts) to the shape and form of developmental trajectories, more specifically in relation to the existence of a phase shift in language acquisition. This study analyzed the relation between the vocabulary spurt in the infant and co-occurring changes in patterns of caregiver-child communication. Infants who showed a discrete transition point were identified as “spurters,” and were compared to infants with a more gradual vocabulary spurt. In addition to coding vocal, gestural, and affective behaviors, a second-order variable was constructed for those communicative behaviors that co-occurred in time (“coordinated bouts”). The results showed that on average, infants who were identified as spurters had a greater disruption in the communicative system during this transition period than more gradual word learners.

Combined, these studies show how researchers can focus on language acquisition in a “holistic”/dynamic way, and what type of knowledge this yields. In these cases, the studies provide evidence for the existence of a complex relation between child-caregiver patterns and the child’s individual language acquisition trajectories.

Self-organization and emergence in L1 acquisition

With regard to the study of L1 development, several authors (particularly researchers collaborating with Linda Smith) have investigated the topics of self-organization and emergence by analyzing the linkages between time scales of language development or between layers of the nested structure of language. For instance, in their review article, Hockema and Smith (2009) present an overview of research, and argue that “[it] shows to be a self-organizing process in which children create the units they need from the regularities present in the environment in which they are situated.” (pp. 453). These authors argue that the causal factors are changing from moment to moment, and that the accumulation of these changes create higher order structure. The study includes a review of studies in the domain of lexical learning that support this conceptualization, for instance with regard to phoneme use (Maye et al., 2002), word segmentation Saffran et al. (1996, 1997) and Aslin et al. (1998), the use of syntactic and semantic categories (Mintz et al., 1995, 2002; Redington et al., 1998; Cartwright and Brent, 1997; Elman, 1990, 1998; Brill, 1991; Finch and Chater, 1992), word meaning (Landaauer et al., 1998). According to Hockema and Smith, this demonstrates that language acquisition is rooted in the “surface” regularities of any particular language, and that higher levels of organization are built every step of the way.

A similar approach was followed by Ibbotson et al. (2019), who explored the extent to which syntactic structure emerges from the interaction of simpler entities (in this case: words). In this specific instance, the authors used a dynamic network model to show how naturalistic child directed speech causes the formation of “communities” around words that were more densely connected with other words. In addition, experimental studies have shown that infants can aggregate and make statistically appropriate decisions from word-referent co-occurrence data in order to learn new words. In line with this, Colunga and Smith (2008) have argued that word learning requires the adaptive flexibility that integrates the stability of past experience with the specifics of the moment, and uses this in the context of children’s novel noun generalizations. By using three different experiments they show how young children flexibly and smartly adapt to novel circumstances in the task of learning new words. According to the authors, this shows that noun acquisition is a system that through its own self-organizing processes builds the knowledge that creates stable categorization patterns.

These examples of studies presented above show that stable patterns of language acquisition emerge as the result of underlying processes of self-organization, and that higher order structures are formed because of the repeated interactions at a lower level. This phenomenon is observed at different levels of language development, such as phoneme use, lexical development, and the use of syntactic and semantic structures.

Nonlinearity and variability in L1 acquisition

The study into nonlinear trajectories has received relatively much attention in dynamic systems oriented empirical studies of first language acquisition. The reason for this might be that first language research has a longstanding tradition of looking at individual trajectories and that most researchers are well aware of its nonlinear and idiosyncratic nature. The framework of complex dynamic systems offers a coherent theory that explains why development shows itself the way it does. In addition, the approach offers appealing theoretical constructs that can be observed and analyzed by means of a broad spectrum of methodological tools. As a result, various types of nonlinearity have received attention in the literature on L1 development, such as temporal regressions, rapid growth, discontinuities and intra-individual variability (see Fig. 2).

Temporal regressions are highly relevant for first language acquisition. The best known example of such regression is the u-shaped trajectory (see Fig. 2A) of verb inflection of irregular verbs (Carlucci and Case, 2013). In fact, Gershkoff-Stowe and Thelen (2004) have argued that temporal regressions are actually characteristic of all developmental processes. Aside from referring to some well-known examples from perception and action (particularly reaching in the A-not-B task), these authors use language development to illustrate how the same kind of perseverations takes place in word naming. Here, the temporal regression was explained by the changing patterns of activation and decay in the use of individual words. The authors refer to data from a study of

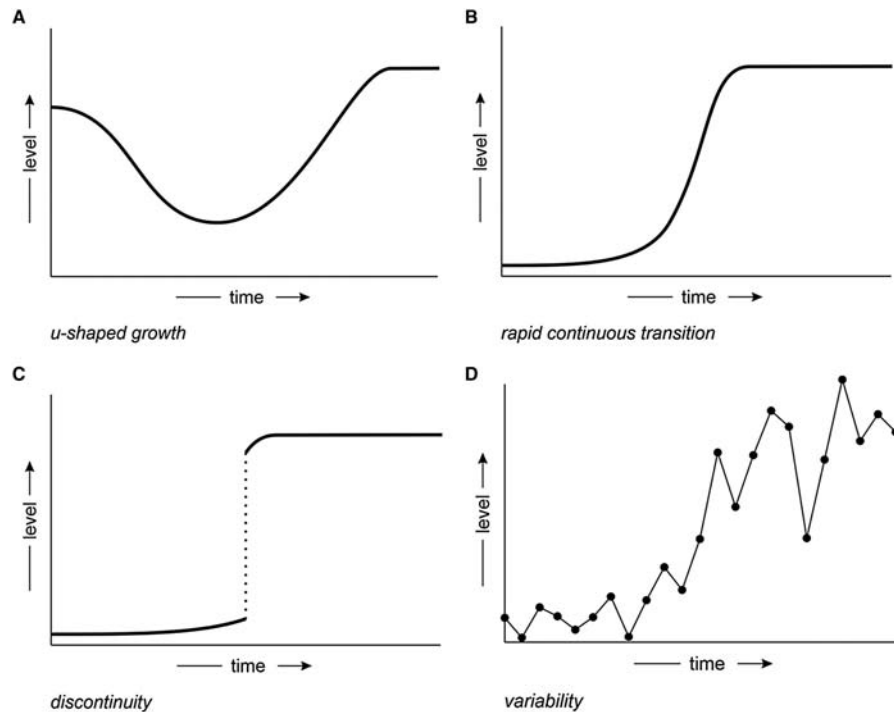


Fig. 2 Examples of nonlinear development: (A) temporal regression as u-shaped growth, (B) rapid continuous development in logistic growth, (C) discontinuous development, (D) intra-individual variability with underlying logistic curve.

Gershkoff-Stowe (2002) that showed that practicing naming objects can reduce the likelihood of perseveration errors during the period of increased vocabulary growth. The results of this study have demonstrated that children consistently made more errors with the target words when they were less practiced and that errors in low-practice words spiked during the vocabulary spurt. This suggests that the temporal regression resulted from multiple contingent processes mediating between the developmental history and the task.

The question of whether **discontinuities** exist in language development is also prominent in the literature. Stage descriptions of first language acquisition have been present already in the first studies of child language, but the question is whether these mainly concern qualitative differences or whether such stages also lead to abrupt changes in a quantitative sense with regard to language use. There are different definitions of a discontinuity, but the most intuitive one is the observation of a sudden change in the trajectory, that is a “jump” between two modes of functioning with a “gap” in between (see Fig. 2C: the gap is represented by a dotted line). As an example of investigating discontinuities, Ruhland and van Geert (1998), have conducted a quantitative study on the development of function word use between the ages of 18–34 months. The trajectories of six children were modeled with both a continuous model (a logistic growth curve, see Fig. 2B for an example) and a discontinuous one (the cusp model). The results show some evidence for rapid change and multimodality, but most findings supported the continuous model of development.

A similar approach—though not by using the cusp model—was taken by Hohenberger and Peltzer-Karpf (2009). These authors focused on detecting phase shifts in the development of lexical categories and MLU-curves. The data that were cited consists of a linguistic corpus of one monolingual (German) and one bilingual (English/German) child who were followed between the ages of 1; 5 and 2; 6 (e.g. Zangl, 1998). The results showed a qualitative and quantitative change around the age of 2 years suggesting a critical point in input processing, which according to the authors coincides with a highly active phase in brain development.

In addition, Rvachew and Bernhardt (2010) have focused on discontinuity in early phonological acquisition. This discontinuity exists in the form of sudden appearance of phonemes or structures and of regressions in speech production (Stemberger et al., 1999; Ullrich et al., 2008; Vihman and Croft, 2007). The authors argue that in all of these examples, the discontinuity (the emergence of a new behavior) is built upon a fundamental underlying continuity (gradual improvements in the child’s abilities with respect to the requisite subcomponent skills).

In Van Dijk and van Geert (2007), we have also suggested that development is not either continuous or discontinuous, but that it is simultaneously continuous and discontinuous with respect to different dimensions of development. Here we focused on detecting discontinuities in naturalistic data of preposition use. In contrast to other studies, data were used that included the variability as part of the change pattern. Four criteria were used to test the development of four children and the results show large individual differences: whereas two participants showed clear evidence for a discontinuity, the results were much weaker for two other participants.

The contributions described above on the detection of discontinuities have raised important questions about both the shape and the nature of the underlying process of first language acquisition. What these studies have added is empirical evidence for the

conceptualization that both continuity and discontinuity can be observed at the same time across domains of early language acquisition.

Various researchers have investigated the role of **intra-individual variability** (or “temporal variability”) in early language acquisition. During this process, there are constant interactions between the child’s many internal states (cognitions, emotions, etc.) and the context (Van Geert and Verspoor, 2015). As a result, the language of the child is different from moment to moment, from conversation to conversation, from day to day, and from week to week (see Fig. 2D for an illustration). Intra-individual variability can be part of any type of underlying trend (linear, logistic, u-shaped, etc.). However, when a system or subsystem is moving from one stage to another, the system destabilizes which leads to an increase of variability in behavior. Applied to data of first language development, it has been confirmed that MLU is highly variable, and that this may be meaningful for development (Van Dijk et al., 2001). There are also some experimental studies that investigate the role of temporal variability, for instance the one by Twomey et al. (2018) on word learning. In their design, 2-year-old children watched novel and known objects, and simultaneously heard a novel or known word. Whereas one group received constant color condition objects appearing on a uniform white background, another group received variable color condition backgrounds. The results showed that only children in the variable condition showed retaining label-object associations. The authors have argued that these findings are consistent with dynamic systems accounts because they show how variability leads to a change in behavior. We reported a similar result in a longitudinal study on spontaneous speech under naturalistic circumstances (van Dijk and van Geert, 2011) where we analyzed the use of prepositions in combination with utterance length. The data of two children showed that both had a relatively stable period with agrammatical preposition use first, through a period of relative variability, to a new stable phase of grammatical preposition use.

Intra-individual variability is also studied in relation to the detection of discontinuities or sudden changes in first language development. A few authors have used the framework of catastrophe theory to predict properties of variability associated with such transitions (van der Maas and Molenaar, 1992). For instance, in the case-study presented by Bassano and van Geert (2007), the intra-individual variability in the use of utterances of a certain length showed two statistically significant peaks, which coincided with rapid changes in the two- and three-word sentences and the 4+-word sentences respectively. These peaks in variability suggest that the process does not reflect a gradual change in utterance length, but that the utterance length is likely to be the consequence of underlying mechanisms of sentence formation that are organized in the form of different attractors.

As argued above, intra-individual variability has an internal structure which provides information about the underlying language acquisition processes. This phenomenon has been studied in both the studies of Cox and van Dijk (2013) and Dale and Spivey (2006) that each investigated the process of syntactic coordination between child and caregiver by means of cross-recurrence quantification analysis. The analyses did not only employ dyadic measures (that is: measures that are collective and capture the system as a whole), but also investigated the structure of the interactions in terms of their temporal variability. The results demonstrated that complexity of the behaviors changes over time, indicating a process of coordination between child language and child directed speech.

In sum, the study of first language acquisition has a long tradition at investigating individual developmental pathways and many studies have demonstrated that language development is nonlinear. More recent studies have shown that various shapes exist, such as u-shaped development, rapid continuous growth and in some cases discontinuous change, which is compatible with dynamic systems principles. In addition, studies that contain a sufficient amount of measurement points demonstrate large intra-individual variability. The amount and structure of this variability is not random, but is likely to relate to the underlying process of change.

Modeling L1 acquisition

As mentioned before, dynamic systems research is also known for its use of mathematical models to describe change processes. These models are based on coupled equations and are able to simulate the way the components of the system interact to produce trajectories of change. These models have been applied to various domains of child development among which first language acquisition. An example of such a model was already mentioned above, where the relation between the language production of the child and the child directed speech of the caregiver was expressed in a series of coupled equations (van Dijk et al., 2013). In addition, McCune and Zlatev (2015) proposed a conceptual threshold model to describe the transition to referential language. Here, they assume that certain developments in reaching prompt the transition to referential word production and/or comprehension. The model consists of four variables (representational play, vocalization, gestures, and communicative grunts) that all interact with each other from moment to moment and that predict use of references in child speech.

A very influential growth model, was proposed by Van Geert (1991) which was based on the logistic growth equation. The data on the lexical development of case-study “Keren” (as reported in Dromi, 1986) were used to specify the parameters in order to fit the data. Robinson and Mervis (1998) applied the same model with data from another child (“Ari”) and demonstrated that both lexical and grammatical development could be fitted with this logistic growth curve.

In sum, these studies show that relatively simple mathematical equations lead to a model which in many cases can be validated by means of empirical trajectories of children’s language use. As said, these models are based on underlying dynamic assumptions, such as the notion of process causality (implying the notion of iterativity), inclusion of the most relevant variables of the system, self-organization, the emergence of higher order stability through interactions at the micro level, etc.

Findings from studies on second language acquisition

In the globalized society of today, most individuals acquire a second, third or fourth language, either in a natural or emersion context, or in an instructional setting. Various authors (such as van Geert, Larsen-Freeman, de Bot) have proposed dynamic systems theory as a candidate for an overall theory of language development and have argued that its ideas are highly relevant for understanding second language (L2) acquisition. Pioneering work on the application of the theoretical constructs to this field was done by Larsen-Freeman (1997, 2002), and this was followed up by Herdina and Jessner (2002). Some years later, Verspoor, de Bot and Lowie published a series of influential papers (Verspoor et al., 2004; De Bot et al., 2005) inspired by these early interpretations and by the work on first language acquisition (Van Geert, 1991; Van Geert and van Dijk, 2002). These authors have also contributed with a series of empirical studies, of which many are based on quantitative measures of complexity, accuracy and fluency of produced language (so-called “CAF measures”).

The main theoretical issues that have been addressed by studies on second language acquisition from a dynamic systems approach are: (1) the existence of complex interactions between variables, (2) the unpredictable and idiosyncratic nature of L2 development, and (3) the importance of intra-individual variability. The first issue follows from the realization that studies on second language learning should focus on the system as a whole and self-organization. However, the question of language emergence is much less a focus of empirical studies, probably because a first language must by definition have been acquired before a second language can be acquired. For this reason, the “origin of language” question is much less a matter of debate in this field. The second and third issues (unpredictable and idiosyncratic development and variability) correspond to the focus on nonlinear forms of change.

Interacting variables in L2 acquisition

The issue of studying **interacting variables** is at the core of the underlying model of second language development. According to De Bot et al. (2007), much of the literature at the time considered simple causal models and implied a linear point of view. In contrast, De Bot, Verspoor and Lowie argue that complex interactions exists between different variables, such as the way the native language literacy is a crucial condition for the successful acquisition of a second language (see for instance Sparks and Ganschow, 1991). A systems viewpoint implies that all variables are interconnected and influence each other in complex ways (de Bot et al., 2007, p. 8), and this interconnectedness of subsystems has been a focus point of various studies.

For instance, Caspi (2010), Verspoor et al. (2008), Spoelman and Verspoor (2010), and van Dijk et al. (2011) have shown how subsystems develop at different rates with different trajectories. In the study of Caspi (2010), the trajectories of four L2 learners of English were analyzed and the results showed that the lexical system developed before the syntactic system, and that for each subsystem general use increased before accuracy was achieved. A more in-depth case-study showed that residual and moving correlation plots corroborated the existence of complex interactions between the knowledge levels, and cycles of competition and support between variables. In another study, Verspoor et al. (2008) analyzed an advanced English learner’s academic writing with regard to the development of vocabulary and complexity. Several interesting patterns were revealed, among which a competitive relationship between type-token ratio and sentence length, and a supportive relationship between finite verb ratio and noun phrase length. In Van Dijk et al. (2011) we also addressed these interrelations and reanalyzed the acquisition data from Cancino et al. (1978) which followed the emergence of the negative verb constructions in detail.

Though at the global level the development showed the expected sequence, interesting interactions at the microscopic level were observed. The first observation was that from the beginning onwards all four negative constructions occurred, and each construction seemed to develop in a wavelike form. The second observation was that nontarget-like constructions (that is constructions that are considered non-grammatical or uncommon) would keep occurring even after the increase of target-like constructions.

As a final example, Spoelman and Verspoor (2010) followed a Dutch learner’s acquisition of written Finnish over a 3-year period. Once again, the results showed interactions between variables over time, with the learner’s complexity variables sometimes competing and sometimes supporting each other. Interestingly, although accuracy and complexity seemed to be in competition early in the study period, they later changed to a noncompetitive relationship as the learner became more proficient, suggesting that proficiency level may influence the interaction of other variables.

Combined, these studies have provided evidence that in L2 development complex interactions most likely exist between different language variables and that these interactions become visible by investigating individual trajectories with a high measurement resolution.

Nonlinearity and variability in L2 acquisition

The second important issue for second language development is the question of whether there is a **fixed order** of acquisition which is identical for each learner. According to various authors, the notion of developmental sequences is a generally accepted finding in L2 research (Baten and Hakansson, 2015; Long, 2007; VanPatten and Williams, 2007), and this is supported by robust evidence over the past 40 years (see Abrahamsson, 2013; Meisel, 2013). In contrast, De Bot et al. (2007), have argued that there is no pre-determined order, but that development is unpredictable and idiosyncratic. This notion is supported by results from various empirical studies. For instance, Larsen-Freeman (2006) examined language production of five Chinese learners of English in a longitudinal study, by means of measures of written grammatical complexity, lexical complexity, accuracy, and fluency. Although

the group means indeed showed a linear increase, the patterns of the inter-individual and the individual data exhibited very clear differences, as the learners progressed and regressed in widely varying ways. Learners showed unique developmental pathways, with different rates of improvement and temporal decreases in some variables, findings which had initially been obscured by the use of group averages. A similar finding was also reported by [Verspoor et al. \(2008\)](#) in a study on academic writing development. Here, the learner showed development in almost all the aspects investigated. This progress was nonlinear and was different for each variable. In another study from this group on advanced writers' use of finite and nonfinite dependent clauses ([Verspoor et al., 2017](#)), there were only few similarities between the three learners. Instead, each learner had their own unique trajectory. The authors argue that although developmental stages may be observable in group means, the latter are not representative for any of the individuals. When investigating the L2 development of two identical twins, [Lowie et al. \(2017\)](#) found that in spite of the striking similarities between these learners in terms of background, proficiency and genetic make-up, the development of their L2 English over time was very different. Developmental patterns for spoken and written language even showed opposite tendencies. These observations underline the idiosyncratic nature of the process of second language development.

However, these results are not convincing for all L2 researchers. According to [Baten and Hakansson \(2015\)](#), the non-linearity that is revealed in the studies is the direct result of the way CAF measures are used to reflect L2 development. They consider this problematic because no rationale is provided for the use of these measures. Adopting a different type of criterion in their own study, –namely the first emergence of certain forms—the authors report that although the use of subordinate clauses is variable, the development of the internal structure followed a clear and constant sequence. They conclude that these findings are evidence that fixed developmental stages do exist in second language acquisition. However, dynamic systems approaches do not assume that fixed orders of development do not exist entirely. Some of these orders are simply unavoidable, because the steps or stages are defined in such a way that their order becomes a logical necessity. For instance, this is the case when a more advanced stage includes the qualities of the simpler stage, or because a stage of using a certain strategy is preceded by a stage in which this strategy is not used. In other cases, a different order is not impossible, but just very unlikely. For instance, though it may be possible that children start out by using three word utterances after which they start using one-word utterances, it is implausible because the universal aspects of the stages originate from the general principle of first simple, then more complex. Dynamic systems approaches are very well compatible with the assumption that such general principles exist. However, these approaches argue that when zooming in on individual trajectories, development is idiosyncratic and nonlinear. Combined, these results show that also in the study of second language acquisition, there is evidence for both a global order and unpredictability at the individual level of development.

The phenomenon of **intra-individual variability** is studied relatively frequently in second language acquisition research. For instance, [Polat and Kim \(2014\)](#) conducted a longitudinal case study on the development of complexity and accuracy for an advanced untutored learner of English. The data showed that most development occurred in the participant's lexical diversity, that syntactic complexity had only a slight improvement, and that accuracy did not change at all. Moreover, the trajectories were highly variable and there were temporal peaks in both lexical diversity and the number of errors. Several studies have shown that increased variability coincides with a developmental jump in second language development as well ([Spoelman and Verspoor, 2010](#); [van Dijk et al., 2011](#)). In addition, [Plat et al. \(2018\)](#) used spectral slope analysis on the data of 42 Dutch students in a Dutch/English naming task. A fractal (that is: self-similar) pattern of variability was found for both L1 Dutch and L2 English word naming. The results also suggest that automatic, optimal coordination of mental processes existed in L1 whereas this was much less the case for L2. This conclusion is in line with an earlier study of [Lowie et al. \(2014\)](#) that looked at a single native Dutch speaker who was very proficient in the L2. Here, the variability showed a more optimal ("fractal") pattern for the language that was recently used, regardless of it being the L1 or L2.

Finally, it is important to note that the coupled equation models that were used in the L1 development have also been applied to L2. One example is worked out in [Caspi \(2010\)](#), where vocabulary knowledge (receptive productive gap) and writing performance of four learners were fitted with the precursor models of [Van Geert \(1994\)](#). The optimization procedure showed a replication of the data trends on the basis of precursor interactions.

In sum, there is a general consensus among researchers that in L2 development, fixed orders of development exist while at the same time individual differences are large. Researchers differ in the interpretation of this paradoxical finding: whereas some argue that the underlying structure is fixed and variations only exist at the surface level, others argue that development is truly idiosyncratic due to processes of self-organization out of which various forms of order can emerge. Furthermore, also in the field of L2 development, it is highly likely that intra-individual variability plays a significant role in development. Finally, coupled equation models are also able to fit L2 development, as was also the case for L1 development.

General discussion

Summary

In this chapter, we have provided a description of dynamic systems theoretical approaches to language acquisition and of what they have yielded. As explained in the introduction, these approaches conceive of language acquisition as a complex system of interacting components or variables. In such a system, language emerges in a communicative context as a result of a process of self-organization. This theory has implications for research that are rather fundamental. Most importantly, it demands a focus on the system as a whole and the dynamic interactions between components and the environment, a radically different perspective on causality (process causality), and a renewed appreciation for intra-individual variability and nonlinear forms of change which

requires analyses of individual trajectories. Many researchers have been inspired by these approaches and have conducted empirical studies on first and second language acquisition. In most cases, the analyses are aimed at gathering proof-of-principle of the specific phenomenon under investigation.

With regard to the empirical literature on first language acquisition, various studies have been published that offer dynamic systems interpretations, and these have provided support for the existence of its main constructs. For instance, there is evidence that L1 development follows a non-linear pathway and that lexical and syntactical structures are likely to emerge through processes of self-organization at the micro-level (words, phonemes). In addition, several other dynamic aspects have been demonstrated to exist, such as reciprocal causality, discontinuities and changes in patterns of temporal variability. Together, the results show that first language acquisition is a highly idiosyncratic process that exhibits both continuity and discontinuity. Bidirectional interactions between variables and co-adaptation processes between children and caretakers have also been observed. Finally, relatively simple, coupled equation models have been successfully applied to trajectories of individual language learners. These models have also been validated and are shown to adequately predict individual pathways of language acquisition. With regard to research methodology, there are some studies that use or refer to empirical data with somewhat larger samples sizes (particularly experimental studies), but most use single or multiple case-studies with repeated measurement of rather global language variables. The data are predominantly on lexical or syntactic development (frequency measures) of children below the ages of 3 years.

Dynamic systems approaches to second language acquisition have addressed similar topics as in L1 acquisition: researchers have paid attention to non-linear shapes of change, have carefully described complex interactions between languages or language components and have conducted studies on the role of intra-individual variability. These studies have provided evidence for the existence of these dynamic phenomena in second language acquisition. Whereas for the field of first language acquisition, the topic of emergence as a result of self-organization has received ample attention, this is much less the case for the field of second language acquisition. In the latter case, the question of a universal fixed order of acquisition seems to be considered as more relevant. The studies so far suggest that based on the development of CAF measures, the existence of a universal fixed order is rather unlikely. Research on L2 development is in many ways explicitly inspired by L1 development. Therefore, it employs a similar methodology, such as multiple case studies, individual-based analyses, descriptive analyses, and modeling.

Critical reflection

In this contribution, we have used the term “dynamic systems approaches” without paying any attention to differences that exist between these approaches. However, it should be noted that there are several related theories and applications that can be considered as “dynamic systems approaches.” Among these are “nonlinear dynamic systems theory” (Thelen and Smith, 1994; Van Geert, 1994), “chaos theory” (Kelso, 1995), “theories of emergence” (Campbell, 2015), “dynamic network approach” (Schmittmann et al., 2013; Wichers, 2014), “dynamic field theory” (Spencer and Schöner, 2003), and “complexity theory” (De Bot, 2017). A description of the specifics of each of these branches goes beyond the scope of this chapter, but these approaches share the above described theoretical concepts, and differences are mainly seen in terminology, methodology and fields of application. The current contribution is based on the approach currently known as “complex dynamic systems,” which is the theoretical “branch” most frequently applied to the field of L1 and L2 acquisition (see for instance Van Geert and Verspoor, 2015; Larsen-Freeman and Cameron, 2008; Verspoor et al., 2004; Smit et al., 2021).

It is important to note that dynamic systems approaches are not standard practice of research into language acquisition at the moment. However, the framework is highly compatible with usage-based (constructivist) approaches, which are generally accepted in the study of Linguistics. Both approaches conceive of language and linguistic categories as emergent properties that come into existence through their use across contexts and where many elements and causal processes—such as statistical learning—come together. The framework also shares many conceptualizations with the theory of language emergence (MacWhinney, 1999). In fact, it may be argued that dynamic systems approaches provide an “umbrella” theory—an ontology—that functions as a modern working hypothesis of how to interpret reality, in the same way as Universal Grammar, Classical Test Theory and General Linear Theory are still prominent ontologies, based on very different assumptions than those of dynamic systems approaches.

If dynamic systems approaches are more of an ontology than a “regular theory,” one may wonder to what degree the claims that follow from it are falsifiable or not. If quantitative trajectories of language use—frequencies of the use of certain linguistic categories—are observed to be highly idiosyncratic, nonlinear and variable, what does this actually prove about the origin and mechanisms of language acquisition? Are the plentiful observations of highly idiosyncratic, nonlinear and variable trajectories of language use support for the idea of underlying universal latent constructs that develop linearly? Or do they support the nonlinear dynamics hypothesized by DST? And is the idea of underlying universal, latent, linearly developing linguistic categories compatible with the evidence for nonlinearity and complexity provided by such widely diverse fields as mathematics, physics or biology? These are fundamental, but difficult questions, and answering them requires an integration of theory and data, both from within and outside the fields of L1 and L2 acquisition.

In addition, many of the theoretical constructs from dynamic systems approaches actually lead to very specific testable hypothesis about developing behavior. This is particularly the case for the phenomena of intra-individual variability, linkages between layers of organization and time scales, and bidirectional influences between variables. Testing these hypotheses requires sufficiently dense data series and detailed analyses of micro development. For instance, if the prediction is that variability is meaningful, it can be hypothesized that it exhibits an optimal structure in certain circumstances. It can also be hypothesized that its amount and structure change in relation to other variables. Several attempts at testing the core theoretical constructs have been published, not only in

the field of L1 and L2 acquisition (e.g. Bassano and van Geert, 2007; Verspoor et al., 2017) but for instance also in the study of parent-child interactions (e.g. Yaden, 2003; Lichtwarck-Aschoff et al., 2012) and cognitive development (De Jonge-Hoekstra et al., 2020). If such hypotheses are not supported by empirical evidence, we would have no reason to still use the theory of complex dynamic systems as a general framework for empirical and theoretical studies.

A related criticism to dynamical systems approaches is that the descriptions of the system's behavior in terms of the nonlinear equations are taken as a mathematical description of its overt behavior rather than some underlying process, as Clark (1997: 115) noted (Clark, 1997). However, at the bottom of this problem is a matter of the ontology underlying the explanatory work. If the ontology is typically substance- or entity-oriented, the processual features of a particular temporal sequence must be explained by such an underlying entity, which is seen as the independent *cause* of the observable features. However, in a typical process ontology, the causal structure is one of process causality, where current manifest features of the process have causal influence on later features of the process (also see Van Geert, 2019). However, in a complex dynamic system framework, features on different but coupled levels of organization are co-dependent and typically display reciprocal causality.

Future directions

Ideas from dynamic systems approaches have been introduced more than a quarter of a century ago and have (rather slowly) seeped through in language acquisition research. At the moment, the validity of the main concepts has been demonstrated in many topics of first and second language development. This suggests that it is now time for larger scale dynamically oriented studies that address many of the unanswered questions about language acquisition. It may be argued that currently there is momentum for conducting such comprehensive empirical studies on first and language development. Traditionally, the study of language use and language interaction is a very labor-intensive enterprise. Currently, however, various new methods are being developed that can assist this process, for instance with (partly) automated data collection, transcription methods, and coding software for various languages. In addition, in recent years, many publications have appeared in related fields—such as psychology, sociology and pedagogics—that have employed dynamic approaches in larger scale studies to investigating various kinds of processes. These studies have conducted analyses of bidirectional influences, intra-individual variability, linkages between layers of organization and nonlinearity by means of new and advanced statistical techniques (such as latent growth curves and cluster analytic techniques). A promising direction in the field of language acquisition research is to investigate the ways in which some learners have different trajectories than others and in which ways typical and atypical language development show different dynamic properties over time. Employing a dynamic systems approach to studying interindividual differences offers a world of possibilities to ask and answer new and intriguing questions about language acquisition.

References

- Abrahamsson, N., 2013. Developmental sequences. In: Robinson, P. (Ed.), *The Routledge Encyclopedia of Second Language Acquisition*. Routledge, pp. 173–177.
- Aslin, R., Saffran, J., Newport, E., 1998. Computation of conditional probability statistics by 8 month-old infants. *Psychol. Sci.* 9 (4), 321–324. <http://www.jstor.org/stable/40063345>.
- Bassano, D., van Geert, P., 2007. Modeling continuity and discontinuity in utterance length: a quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Dev. Sci.* 10 (5), 588–612. <https://doi.org/10.1111/j.1467-7687.2006.00629.x>.
- Baten, K., Hakansson, G., 2015. The development of subordinate clauses in German and Swedish as L2S; a theoretical and methodological comparison. *Stud. Sec. Lang. Acquis.* 37 (3), 517–547. <https://doi.org/10.1017/S0272263114000552>.
- Brill, E., 1991. Discovering the lexical features of a language. In: *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics Berkeley*. Association for Computational Linguistics, Morristown, NJ.
- Campbell, R., 2015. *The Metaphysics of Emergence*. Springer Nature.
- Cancino, H., Rosansky, E., Schumann, J., 1978. The acquisition of English negatives and interrogatives by native Spanish speakers. In: Hatch, E.M. (Ed.), *Second Language Acquisition: A Book of Readings*. Newbury House, pp. 207–230.
- Carlucci, L., Case, J., 2013. On the necessity of U-shaped learning. *Top. Cognit. Sci.* 5 (1), 56–88. <https://doi.org/10.1111/tops.12002>.
- Cartwright, T.A., Brent, M., 1997. Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition* 63 (2), 121–170. [https://doi.org/10.1016/S0010-0277\(96\)00793-7](https://doi.org/10.1016/S0010-0277(96)00793-7).
- Caspi, T., 2010. *A Dynamic Perspective on Second Language Development* [Doctoral thesis, University of Groningen, the Netherlands]. Retrieved from. <http://irs.ub.rug.nl/ppn/329338412>.
- Clark, A., 1997. *Being there: Putting brain, body and world together again*. MIT Press, Cambridge.
- Colunga, E., Smith, L.B., 2008. Flexibility and variability: essential to human cognition and the study of human cognition. *New Ideas Psychol.* 26 (2), 174–192. <https://doi.org/10.1016/j.newideapsych.2007.07.012>.
- Cox, R.F.A., van Dijk, M., 2013. Microdevelopment in parent-child conversations: from global changes to flexibility. *Ecol. Psychol.* 25 (3), 304–315. <https://doi.org/10.1080/10407413.2013.810095>.
- Dale, R., Spivey, M.J., 2006. Unraveling the dyad: using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Lang. Learn.* 56 (3), 391–430. <https://doi.org/10.1111/j.1467-9922.2006.00372.x>.
- De Bot, K., Lowie, W., Verspoor, M., 2005. Dynamic systems theory and applied linguistics: the ultimate “so what”? *Int. J. Appl. Ling.* 15 (1), 116–118. <https://doi.org/10.1111/j.1473-4192.2005.0083b.x>.
- De Bot, K., Lowie, W., Verspoor, M., 2007. A dynamic systems theory approach to second language acquisition. *Biling. Lang. Cognit.* 10 (1), 7–21. <https://doi.org/10.1017/S1366728906002732>.
- De Bot, K., 2017. Complexity theory and dynamic systems theory. In: Ortega, L., Han, Z. (Eds.), *Complexity Theory and Language Development: In Celebration of Diane Larsen-Freeman*. John Benjamins, pp. 48–51.

- De Jonge-Hoekstra, L., van der Steen, S., Cox, R.F.A., 2020. Movers and shakers of cognition: hand movements, speech, task properties, and variability. *Acta Psychol.* 211, 103187. <https://doi.org/10.1016/j.actpsy.2020.103187>.
- Dromi, E., 1986. *Early Lexical Development*. Cambridge University Press.
- Elman, J., 1990. Finding structure in time. *Cognit. Sci.* 14 (2), 179–211. [https://doi.org/10.1016/0364-0213\(90\)90002-E](https://doi.org/10.1016/0364-0213(90)90002-E).
- Elman, J., 1998. Generalization, simple recurrent networks, and the emergence of structure. In: Gernsbacher, M.A., Derry, S.J. (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum.
- Finch, S., Chater, N., 1992. Bootstrapping syntactic categories by unsupervised learning. In: Gernsbacher, M.A., Derry, S.J. (Eds.), *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*. Lawrence Erlbaum, pp. 820–826.
- Gershkoff-Stowe, L., Thelen, E., 2004. U-shaped changes in behavior: a dynamic systems perspective. *J. Cognit. Dev.* 5 (1), 11–36. https://doi.org/10.1207/s15327647jcd0501_2.
- Gershkoff-Stowe, L., 2002. Object naming, vocabulary growth, and the development of word retrieval abilities. *J. Mem. Lang.* 46 (4), 665–687. <https://doi.org/10.1006/jmla.2001.2830>.
- Herdina, P., Jessner, U., 2002. A dynamic model of multilingualism: perspectives of change in psycholinguistics. In: *Multilingual Matters*.
- Hockema, S.A., Smith, L.B., 2009. Learning your language, outside-in and inside-out. *Linguistics* 47 (2), 453–479. <https://doi.org/10.1515/LING.2009.016>.
- Hohenberger, A., Peltzer-Karpf, A., 2009. Language learning from the perspective of nonlinear dynamic systems. *Linguistics* 47 (2), 481–511. <https://doi.org/10.1515/LING.2009.017>.
- Hsu, H.-C., Fogel, A., 2001. Infant vocal development in a dynamic mother–infant communication system. *Infancy* 2 (1), 87–109. https://doi.org/10.1207/S15327078IN0201_6.
- Ibbotson, P., Salnikow, V., Walker, R., 2019. A dynamic network analysis of emergent grammar. *First Lang.* 39 (6), 652–680. <https://doi.org/10.1177/0142723719869562>.
- Kelso, S., 1995. *Dynamic patterns: the self-organizing of brain and behavior*. MIT Press, Cambridge.
- Landauer, T., Laham, D., Foltz, P., 1998. Learning human-like knowledge by singular value decomposition: a progress report. In: Jordan, M.I., Kearns, M.J., Solla, S.A. (Eds.), *Advances in Neural Information Processing Systems*, vol. 10. MIT Press, pp. 45–51.
- Larsen-Freeman, D., Cameron, L., 2008. Research methodology on language development from a complex systems perspective. *Mod. Lang. J.* 92 (2), 200–213. <https://doi.org/10.1111/j.1540-4781.2008.00714.x>.
- Larsen-Freeman, D., 1997. Chaos/complexity science and second language acquisition. *Appl. Linguist.* 18 (2), 148–165. <https://doi.org/10.1093/applin/aml029>.
- Larsen-Freeman, D., 2002. Language acquisition and language use form a chaos/complexity theory perspective. In: Kramsch, C. (Ed.), *Language Acquisition and Language Socialization: Ecological Perspectives*. Continuum, pp. 33–46.
- Larsen-Freeman, D., 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Appl. Linguist.* 27 (4), 590–619. <https://doi.org/10.1093/applin/aml029>.
- Lichtwarck-Aschoff, A., Hasselman, F., Cox, R., Pepler, D., Granic, I., 2012. A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dyn. Psychol. Life Sci.* 16 (3), 353–379.
- Long, M.H., 2007. Problems in SLA. Erlbaum.
- Lowie, L., de Bot, K., Plat, R., 2014. Pink noise in production; a nonlinear approach to the multilingual lexicon. *Ecol. Psychol.* 26, 216–228. <https://doi.org/10.1080/10407413.2014.929479>.
- Lowie, W., van Dijk, M., Chan, H., Verspoor, M., 2017. Finding the key to successful L2 learning in groups and individuals. *Stud. Sec. Lang. Learn. Teach.* 7 (1), 127–148. <https://doi.org/10.14746/sslit.2017.7.1.7>.
- MacWhinney, B., 1999. *The Emergence of Language*. Lawrence Erlbaum Associates Publishers.
- Maye, J., Werker, J.F., Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82 (3), B101–B111. [https://doi.org/10.1016/S0010-0277\(01\)00157-3](https://doi.org/10.1016/S0010-0277(01)00157-3).
- McCune, L., Zlatev, J., 2015. Dynamic systems in semiotic development: the transition to reference. *Cognit. Dev.* 36, 161–170. <https://doi.org/10.1016/j.cogdev.2015.09.010>.
- Meisel, J.M., 2013. Development in second language acquisition. In: Robinson, P. (Ed.), *The Routledge Encyclopedia of Second Language Acquisition*. Routledge, pp. 165–173.
- Mintz, T.H., Newport, E.L., Bever, T.G., 1995. Distributional regularities of form class in speech to young children. In: Beckman, J. (Ed.), *Proceedings of the 25th Annual Meeting of the North Eastern Linguistics Society*, vol. 2. GLSA.
- Mintz, T.H., Newport, E.L., Bever, T.G., 2002. The distributional structure of grammatical categories in speech to young children. *Cognit. Sci.* 26 (4), 393–424. [https://psycnet.apa.org/doi/10.1016/S0364-0213\(02\)00072-1](https://psycnet.apa.org/doi/10.1016/S0364-0213(02)00072-1).
- Parladé, M.V., Iverson, J.M., 2011. The interplay between language, gesture, and affect during communicative transition: a dynamic systems approach. *Dev. Psychol.* 47 (3), 820–833. <https://doi.org/10.1037/a0021811>.
- Plat, R., Lowie, W., de Bot, K., 2018. Word naming in the L1 and L2: a dynamic perspective on automatization and the degree of semantic involvement in naming. *Front. Psychol.* 8, 2256. <https://doi.org/10.3389/fpsyg.2017.02256>.
- Polat, B., Kim, Y., 2014. Dynamics of complexity and accuracy a longitudinal case study of advances untutored development. *Appl. Linguist.* 35, 184–207. <https://doi.org/10.1093/applin/amt013>.
- Redington, M., Chater, N., Finch, S., 1998. Distributional information: a powerful cue for acquiring syntactic categories. *Cognit. Sci.* 22, 425–469.
- Robinson, B.F., Mervis, C.B., 1998. Disentangling early language development: modeling lexical and grammatical acquisition using and extension of case-study methodology. *Dev. Psychol.* 34 (2), 363–375. <https://doi.org/10.1037/0012-1649.34.2.363>.
- Ruhland, R., van Geert, P., 1998. Jumping into syntax: transitions in the development of closed class words. *Br. J. Dev. Psychol.* 16 (Pt 1), 65–95. <https://doi.org/10.1111/j.2044-835X.1998.tb00750.x>.
- Rvachew, S., Bernhardt, B.M., 2010. Clinical implications of dynamic systems theory for phonological development. *Am. J. Speech Lang. Pathol.* 19 (1), 34–50. [https://doi.org/10.1044/1058-0360\(2009\)08-0047](https://doi.org/10.1044/1058-0360(2009)08-0047).
- Saffran, J., Aslin, R., Newport, E., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Saffran, J., Newport, E., Aslin, R., Tunick, R.A., Barrueco, S., 1997. Incidental language learning: listening (and learning) out of the corner of your ear. *Psychol. Sci.* 8, 101–105.
- Schmittmann, V.D., Cramer, A.O.J., Waldorp, L.J., Epskamp, S., Kievit, R.A., Borsboom, D., 2013. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas Psychol.* 31, 43–53. <https://doi.org/10.1016/j.newideapsych.2011.02.007>.
- Smit, N., van Dijk, M., de Bot, K., Lowie, W., 2021. The complex dynamics of adaptive teaching; Observing teacher-student interaction in the language classroom. *International Review of Applied Linguistics in Language Teaching*. In press.
- Sparks, R., Ganschow, L., 1991. Foreign language learning differences: affective or native language aptitude differences. *Mod. Lang. J.* 75 (1), 3–16.
- Spencer, J.P., Schöner, G., 2003. Bridging the representational gap in the dynamic systems approach to development. *Dev. Sci.* 6 (4), 392–412. <https://psycnet.apa.org/doi/10.1111/1467-7687.00295>.
- Spoelman, M., Verspoor, M., 2010. Dynamic patterns in development in accuracy and complexity: a longitudinal case study in the acquisition of Finnish. *Appl. Linguist.* 31 (4), 532–553. <https://doi.org/10.1093/applin/aml001>.
- Stemberger, J.P., Bernhardt, B.H., Johnson, C.E., 1999. “Regressions” (“U”-shaped learning) in the acquisition of prosodic structure. In: *Poster Session Presented at the 6th International Congress for the Study of Child Language*, San Sebastian, Spain.
- Thelen, E., Smith, L., 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, Cambridge.
- Twomey, K.E., Ma, L., Westermann, G., 2018. All the right noises: background variability helps early word learning. *Cognit. Sci.* 42 (Suppl. 2), 413–438. <https://doi.org/10.1111/cogs.12539>.

- Ullrich, A., Stemberger, J.P., Bernhardt, B.M., 2008. Variability in a German-speaking child as viewed with constraints- based nonlinear phonology. *Asia Pac. J. Speech Lang. Hear.* 11, 221–237. <https://doi.org/10.1179/136132808805335536>.
- Van der Maas, H.L., Molenaar, P.C., 1992. Stages of cognitive development: an application of catastrophe theory. *Psychol. Rev.* 99 (3), 395–417. <https://doi.org/10.1037/0033-295X.99.3.395>.
- Van Dijk, M., van Geert, P., 2007. Wobbles, humps and sudden jumps: a case study of continuity, discontinuity and variability in early language development. *Infant Child Dev.* 16 (1), 7–33. <https://doi.org/10.1002/icd.506>.
- Van Dijk, M., van Geert, P., 2011. Heuristic techniques for the analysis of variability as a dynamic aspect of change. *Infanc. Aprendiz.* 34 (2), 151–167. <https://doi.org/10.1174/021037011795377557>.
- Van Dijk, M., de Goede, D., Ruhland, R., van Geert, P., 2001. Kindertaal met hokkensprongen Een studie naar intra-individuele variabiliteit in de taalontwikkeling [Child language cuts capers: a study into intra-individual variability in developmental processes]. *Ned. Tijdschr. Psychol. haar Grensgebieden* 56 (1), 26–39.
- Van Dijk, M., Verspoor, M., Lowie, W., 2011. Chapter 3: variability and DST. In: Verspoor, M., Lowie, W. (Eds.), *A Dynamic Approach to Second Language Development: Methods and Techniques*. John Benjamins Publishers, pp. 55–84.
- Van Dijk, M., van Geert, P., Korecky, K.K., Maillochon, I., Laaha, S., Dressler, W.U., Bassano, D., 2013. Dynamic adaptation in child–adult language interaction. *Lang. Learn.* 63 (2), 243–270. <https://doi.org/10.1111/lang.12002>.
- Van Geert, P., van Dijk, M., 2002. Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behav. Dev.* 25 (4), 340–374. [https://doi.org/10.1016/S0163-6383\(02\)00140-6](https://doi.org/10.1016/S0163-6383(02)00140-6).
- Van Geert, P., Verspoor, M., 2015. Chapter 24: dynamic systems and language development, pp. 537–555. In: MacWhinney, B., O'Grady, W. (Eds.), *The Handbook of Language Emergence*. John Wiley and Sons, Inc, pp. 537–556.
- Van Geert, P., 1991. A dynamic systems model of cognitive and language growth. *Psychol. Rev.* 93 (1), 3–53.
- Van Geert, P., 1994. *Dynamic Systems of Development. Change between Complexity and Chaos*. Harvester.
- Van Geert, P., 2003. Dynamic systems approaches and modeling of developmental processes. In: Valsiner, J., Conolly, K.J. (Eds.), *Handbook of developmental Psychology*. Sage, London, pp. 640–672.
- Van Geert, P.L., 2019. Dynamic systems, process and development. *Hum. Dev.* 63, 153–179. <https://doi.org/10.1159/000503825>.
- VanPatten, B., Williams, J., 2007. *Theories in Second Language Acquisition*. Erlbaum.
- Verspoor, M., De Bot, K., Lowie, W., 2004. Dynamic systems theory and variation: a case study in L2 writing. In: Aertsen, H., Hannay, M., Lyall, R. (Eds.), *Words in Their Places: A Festschrift for J. Lachlan Mackenzie*. Free University Press, pp. 407–421.
- Verspoor, M., Lowie, W., Van Dijk, M., 2008. Variability in second language development from a dynamic systems perspective. *Mod. Lang. J.* 92 (2), 214–231. <https://doi.org/10.1111/j.1540-4781.2008.00715.x>.
- Verspoor, M., Lowie, W., Chan, H., Vahtrick, L., 2017. Linguistic complexity in second language development: variability and variation at advanced stages. *Recherches en didactique des langues et des culture. Les cahiers de l'Acedle* 14 (1). <https://doi.org/10.4000/rdlc.1450>.
- Vihman, M.M., Croft, W., 2007. Phonological development: toward a “radical” templatic phonology. *Linguistics* 45 (4), 683–725. <https://doi.org/10.1515/LING.2007.021>.
- Wichers, M., 2014. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychol. Med.* 44, 1349–1360. <https://doi.org/10.1017/S0033291713001979>.
- Yaden Jr., D.B., 2003. Parent-child storybook reading as a complex adaptive system: or “an igloo is a house for bears.”. In: van Kleeck, A., Stahl, S., Bauer, E. (Eds.), *On Reading Books to Children: Parents and Teachers*. Lawrence Erlbaum, Mahwah, NJ, pp. 336–362.
- Zangl, R., 1998. *Dynamische Muster in der sprachlichen Ontogenese: Bilingualismus, Erst- und Fremdspracherwerb*. Narr.

Writing before age three

Deborah Wells Rowe, Vanderbilt University, Department of Teaching and Learning, Nashville, TN, United States

© 2023 Elsevier Ltd. All rights reserved.

Writing before age three	27
Contexts for observing two-year-olds learning to write	28
What does participation in writing look like before age three?	28
Personal interests and approaches to participating in writing	29
Conceptual interests	29
Procedural interests	29
Creative interests	29
Social interests	29
What are children under age three learning about writing?	29
Foundational understandings about writing	29
Writing forms and directional patterns	30
Message content	31
Speech-print match	31
Writing identity	32
Summing up: what children learn about writing before age three	32
How do children learn to write before age three?	32
Reegan's first invitation to write a message	33
Using what is "at hand" to participate in writing events	34
Adult demonstrations of what, how, and why to write	35
Top-down and bottom-up processing in shared activities with common goals	35
Summing up: how children learn to write before age three	36
Concluding thoughts	36
References	37

Writing before age three

It is well-established that, beginning about twelve months of age, many toddlers begin to make marks (Dunst and Gorman, 2009). They dip their fingers in food and delight in rubbing them across any surface they can reach (Matthews, 1999). By the time they are two years old, they are avid mark makers, whether they drag a stick in the dirt while playing outdoors or use pencils, markers, or paint brushes to leave their traces on paper. Though it is quite likely that these early mark-making experiences are in some way related to children's experiences in learning to write, the graphic activities of children under age three have attracted very little attention in the fields of literacy studies and literacy education (Dunst and Gorman, 2009; Lee, 2013). Instead, most of the research on the graphic activities of one and two-year-olds has been located in the art literature (e.g., Golomb, 1974; Kindler and Darras, 1994; Matthews, 1984) where researchers have been interested in the trajectory from early "scribbles" to iconic representations of people and objects, that is, childhood drawing.

Like art researchers, literacy researchers have generally associated children's earliest marks with drawing rather than writing, effectively pushing one- and two-year-olds to the margins of early writing research. When literacy researchers have attended to the mark-making of children under age three, the aim of research has often been to determine when children begin to purposefully differentiate their marks as writing versus drawing. While findings (Levin and Bus, 2003; Otake et al., 2017; Treiman and Yin, 2011) suggest that this differentiation begins, for some children, around age 2, only a small body of research has explored what children are learning prior to displaying writing-like features in their graphic products, or how they come to learn it (Dunst and Gorman, 2009; Lee, 2013). In short, we know much less than we should about how the graphic activities of one- and two-year-olds start them on a path toward the kinds of emergent writing (Teale and Sulzby, 1986) that have been so thoroughly studied and described for 3- to 5-year-olds. As Lancaster (2003) has argued,

It would be surprising if, from an apparent position of no serious engagement with print at all, children are suddenly able to explode into the type of rampant, creative, independent and reasoned relationship with the medium, which is described in... studies [of writing development of 3- to 5-year-olds.] It appears that much groundwork has been laid before. (p. xx)

My aim in this chapter is to draw together findings from my own studies of two-year-old writers to explore *what* children are learning about writing before age three, and *how* this learning happens, especially in preschool classrooms. In the next sections,

I first provide contextual information about the children who participated in my studies and the spaces where I and my research team observed them writing. Next, I provide a profile of what these two-year-olds were learning about writing. Finally, I describe the social processes through which these children were forming understandings about writing and coming to see themselves as writers. I conclude with some thoughts about why research on writing before age three is so important.

Contexts for observing two-year-olds learning to write

Writing and learning to write are local experiences. Like other aspects of early childhood learning (Rogoff et al., 1993) very young children's experiences with learning to write are necessarily shaped by their individual histories, the cultural expectations and practices of adult caregivers and teachers, and children's opportunities to observe and participate in writing events in home, caregiving, and community contexts.

The portraits of early writing presented in this chapter come from two, long-term qualitative studies of two-year-olds' writing in preschool classrooms. The eighteen children observed in the Walker Preschool study (Rowe, 2008a,b) hailed from White, middle class families, while the eighteen two-year-olds participating in the Write Start! study (Rowe, 2019; Rowe and Neitzel, 2010; Rowe and Wilson, 2015) were mostly from Black families living in a low-income neighborhood. All families lived in a mid-sized city in the United States and reported speaking English at home. Despite differences in children's racial and family backgrounds, formal and informal interviews with parents revealed many similarities in children's experiences with writing at home. In both studies, children had access to writing materials and paper at home, but neither group of parents asked their two-year-olds to write messages beyond the child's name. Some families in both studies limited children's access to pencils, markers, and other writing tools because they felt children were likely to "make a mess" by writing on objects or surfaces other than paper. Most children, in both studies, had opportunities to explore making marks on the page, and were sometimes given paper and pencil to pass the time while families waited for appointments, or helped older siblings with homework. Both groups of children had opportunities to see adults, siblings and others writing at home and in the community.

Most important for understanding the patterns to be reported in this chapter, both groups of children attended childcare settings where adults shared many similar ways of teaching and interacting with young children. The White and Black women who served as classroom teachers, along with me, as a White female researcher, enacted patterns of middle-class child-rearing (Mejía Arauz et al., 2005; Rogoff, 2003) learned in our homes, but also through early childhood teacher preparation. In both studies, teachers organized the day around long periods where children had opportunities to learn informally through self-selected play activities (Copple and Bredekamp, 2009). Adults invited and guided children's participation in play, engaging them as conversational and play partners.

Adults' interactions around writing were guided by emergent literacy principles (Rowe, 2008a). Specifically, each classroom devoted space to a writing center where children had open access to a variety of writing tools and media, and where they were encouraged to launch their own projects as part of child-selected play activities. Teachers wrote alongside children as a means of providing authentic demonstrations of what and how to write. They sometimes worked with children to co-author a text. While at the writing table, adults engaged children in extended conversations about their texts and topics related to their messages. Though none of the two-year-olds in either study were writing with conventional alphabet letters at the beginning of the school year, adults in both settings playfully invited children to write and asked them to read their marks. Regardless of the kinds of marks produced, adults accepted and praised children's participation as writers. Teachers assumed that "form follows function" and that participation in writing would, over time, lead children to construct new understandings about writing.

What does participation in writing look like before age three?

While research on children's earliest writing experiences in homes is limited, a few case studies conducted by parent researchers show that some toddlers begin to show interest in using writing tools to make marks beginning as early as 12–18 months of age (Baghban, 1984; Schickedanz, 1990). In these families, children showed early interest in observing and taking part in adults' writing. Sitting alongside an adult who was writing, youngsters grabbed the pen in a bid to make their own marks. When children were asked to write something, they took their turns by making marks. Initially, these very young children seemed to focus on the physical-motor activity of moving the pen across the page with little attention to the resulting scribbles (Baghban, 2007). However, within a few weeks or months, children began to play closer attention to their marks, deliberately repeating some kinds of movements and visually examining the outcomes (Schickedanz, 1990). Case study children were observed to start naming some of their marks as early as twenty-one months of age (Baghban, 1984).

My classroom-based research with two-year-olds (Rowe, 2008a,b, 2019; Rowe and Neitzel, 2010; Rowe and Wilson, 2015) provides additional insight into the ways very young children participate in writing events when they have open access to writing materials, opportunities to play with peers and the attention of interested adults. In the classrooms where these studies were conducted, children often voluntarily chose to come to the writing center. Two-year-olds were intrigued by exploring the physical properties and possibilities offered by paper and page-based tools for marking (pens, pencils, markers, crayons, rubber stamps), joining and affixing (stapling, taping, pasting), and disjoining (cutting, tearing, punching holes). For many, it was the physical process of marking or pasting or cutting that attracted their interest. Many children initially paid little attention to the resulting products once

they concluded their exploration of the tools. However, adults often opened conversations with children that drew attention to their texts. In response, children talked about what they were doing at the writing table in general terms or displayed their intentions non-verbally through pointing or other gestures. Children sometimes read a particular message for their texts.

Personal interests and approaches to participating in writing

While these findings establish that two-year-olds are interested in engaging with writing materials and interacting with others around their texts, our research (Rowe and Neitzel, 2010) also has shown that children's ways of participating in classroom writing activities varies in predictable ways related to their broader patterns of personal interest. Specifically, the types of writing activities children launched and the features of texts or interactions that drew their attention were related to more general orientations to play reflecting differences in preferred play types and actions. Our classroom observations identified four broad interest patterns.

Conceptual interests

Children with *conceptual* interests and play preferences spent a great deal of their play time in topic-centered explorations where they talked about, manipulated, and categorized themed toy collections (e.g., dinosaur figures, transportation toys) or other materials related to their conceptual interests. As they played, they sought out topic-relevant information or content knowledge related to these interests, often through requests for information directed at adults. These children approached writing as another kind of opportunity to draw, write, and talk with others about topics that were personally important. As they talked with adults and peers about messages for their texts, they often sub-contracted with adults to help them record their ideas. These children engaged in writing as a venue for developing and sharing ideas.

Procedural interests

Children who approached play with *procedural* interests spent the largest portion of their play in spatial-relational or constructive activities with materials such as puzzles or blocks. They preferred highly structured activities focused on the processes, uses or relations between objects. At the classroom writing center, children with procedural play interests were more likely to choose structured literacy materials and activities that moved them toward conventional writing. For example, at the writing table, these children engaged in repeated, self-initiated practice of letter and name writing using letter templates or name cards as models. Writing was centrally about performing conventional writing procedures and reproducing conventional print forms. Learning to write was a puzzle to be solved through careful observation and practice.

Creative interests

Children with *creative* play interests chose fantasy play and other activities that allowed them to explore and transform the uses of materials in novel ways. They preferred materials that could be used in multiple ways and were relatively free from inherent limits on activity imposed by structured materials such as puzzles. These children approached writing as an opportunity to explore the transformative potentials of open-ended writing materials and processes. They focused on generating innovative literacy structures and processes rather than reproducing them. They used materials in innovative ways and intentionally experimented with the properties of markers, paper, scissors and other tools while carefully observing the results. Interactions with other people played only a minor role in shaping their participation.

Social interests

In contrast, children with *social* interests regularly engaged in play actions aimed at building, managing, and maintaining relationships with their peers. They spent a large amount of time engaged in sociodramatic play where they took up roles related to their life experiences. (e.g., family roles, familiar occupations) and enacted social routines that focused on the relationships, rules, and structures of society. For these children, writing was experienced as an opportunity to be socially engaged with peers or teachers. These children tended to align their writing activities to those of other participants. For example, they joined peers with procedural interests in attempting to write alphabet letters and talked about the topics introduced by peers with conceptual interests. For these children, writing was an opportunity to initiate and maintain social connections with others. While texts were important, it was the social uses of those texts that were the focus of children's attention. For them, writing was a relational activity.

What are children under age three learning about writing?

Foundational understandings about writing

When the two-year-olds in our studies engaged in writing in emergent literacy classrooms they formed a broad range of foundational understandings about writing. These social contracts for writing (Rowe, 2008a) were socially shared understandings that allowed members of the preschool community to recognize and use written texts as a special class of material objects. These understandings were so basic that few adults would consciously recognize them as targets for instruction. For example, the children in our studies learned that pages containing their marks were viewed as a special kind of object and were central to mediating interactions with adults and peers. As children's attention was drawn more closely to the page and their marks, they formed foundational understandings about print as a representational system, including figure-ground relations (i.e., marks are the central focus in written texts

and paper is the background) and the message concept (i.e., the child's marks can represent a linguistic message, Clay, 1975). Children's attention was also drawn to the physical features of texts including the understanding that the edges of the page formed a physical boundary for the text and their marks.

Writing forms and directional patterns

Beyond these foundational understandings, children's marks showed that they were beginning to attend to the visual features of print (Rowe and Wilson, 2015). For adults, the writing forms produced by preschoolers can seem chaotic. However, emergent literacy researchers (e.g., Clay, 1975; Harste et al., 1984) long ago established that preschoolers' writing performances are not randomly produced, but instead provide clues to children's understandings about print.

In the Write Start! study (Rowe and Wilson, 2015), in addition to open-ended observations, we used a standard writing task where 2- to 5-year-olds were asked to write a caption for a photo showing them playing in the classroom. Data from this study provided an opportunity to cross-sectionally and longitudinally analyze the writing produced by 139 preschoolers including a group of 18 two-year-olds whose data patterns are the focus of this section. In all age groups, we found that variability was the rule, with two-year-old Write Start! study participants producing the types of writing forms seen in Fig. 1: undifferentiated scribbles, scribble units, stroke units, personal cursive and personal manuscript. Though none of the two-year-olds produced conventional letters when writing a photo caption, their writing forms, nevertheless, showed increasing awareness of print features.

Full page, undifferentiated scribbles were one of the most frequent types of writing for this age group and were produced by 28% of two-year-olds. These marks were visually the least print-like. However, by producing scribbles when asked to write, children demonstrated the foundational understanding that writing involved making marks.

The other forms of writing produced by two-year-olds showed the beginnings of more fine-grained awareness of what print looks like, beginning with global observations that print is organized into units and continuing with increasingly more complex observations of letter features. Twenty-eight percent of two-year-olds wrote using scribble units. Their texts contained many small patches of overlapping scribbles surrounded by white space. An additional 22% of this age group wrote using stroke units, generating texts composed of small individual lines, circles, or curves. These children wrote their messages with marks reflecting their observations of the kinds of strokes used to create English alphabet letters. A final group, 22% of two-year-olds, produced either personal cursive or personal manuscript. As seen in Fig. 1, personal cursive is characterized by wavy horizontal lines, and reflects children's attention to the linear arrangement of English print. Children who wrote with personal manuscript generated letter-like forms that contained at least two different kinds of strokes joined in the same unit. Personal manuscript demonstrated children's more developed understandings about the complexity and variation in strokes and how they are combined to create alphabet letters—an awareness that is also foundational to processing print during reading (Rumelhart, 2004).

Children's texts also showed that some were beginning to attend to the directional patterns used to arrange English print on the page. While the majority of two-year-olds in the Write Start! study produced overlapping scribbles (28%) or placed their marks in random locations on the page (44%), just over one fourth of the children (28%) arranged their marks in horizontal, vertical, or diagonal lines. As children moved from whole page scribbles to creating different kinds of marks surrounded by white space, some two-year-olds demonstrated emergent understandings of the linear arrangement of English print units.

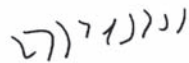
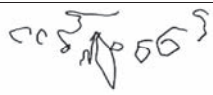
Scribbles	
Scribble units	
Individual stroke units	
Personal Manuscript (letter-like forms)	
Personal Cursive	

Fig. 1 Writing forms produced by two-year-old writers in the write start! study (Rowe and Wilson, 2015).

Message content

While writing forms have received the most attention from researchers, as children in our studies talked about and read their texts to adults, they also demonstrated their developing abilities to verbalize the messages represented by their writing. Clay (1975) has argued that the message concept, the child's understanding that their own marks can represent a linguistic message, is a crucial understanding that must be formed by young writers. In the Write Start! study (Rowe and Wilson, 2015), after children had responded to the request to write a photo caption, we also asked them to read their writing. Analysis of children's orally produced messages provides a window into their initial understandings about the representational nature of print and also of the kinds of messages that might be expressed through print in a familiar social situation.

Write Start! assessment data showed that two-year-olds used several different strategies to structure their responses when asked to read the marks making up their photo captions. A first group, comprising 28% of two-year-olds, was silent or responded with gesture or mumbles. Our qualitative observations suggested that many of these children seemed unsure of what was being asked of them or how they were expected to respond.

When children were asked to read their writing, some two-year-olds took their turns in interaction by reading messages that were unrelated to the content of the photo, despite having previously talked with the adult about the pictured events and objects. Six percent of children used the opportunity to talk with the adult about an unrelated topic. These children used their developing conversational skills to successfully take their turn in response to an adult question but did not link their talk to the ongoing writing event. An additional 6% of children produced a conventional performance such as reciting the names of numbers or alphabet letters in sequence or reading a conventional message such as "I love you." While not related to the task of composing a photo caption, these children demonstrated understanding of specific kinds of messages they had experienced writing with adults at home or school.

In a third pattern, 22% of two-year-olds produced an oral message that was related to the immediate social or physical act of writing (e.g., "It's red."; "I went round and round." "It's for you." "I want to take it home."). They understood that reading their writing involved talking about the writing event underway.

The most frequent response for this age group was also the most conventional. Despite the unconventional emergent forms of writing produced by all two-year-olds in the Write Start! study, 39% of this age group read their marks as a word, phrase, or sentence related to the people, objects, and actions depicted in the photo. To read their marks, they drew on their ability to talk about familiar topics and ongoing events as a support for message construction.

Speech-print match

In addition, to learning how to compose a message for their written texts, children were also learning about the inner workings of print as a representational system. As Read (2009) has suggested, the understanding that written symbols can represent spoken language is an essential component of learning to write in alphabetic languages. This global understanding is a necessary foundation for forming the fine-grained understandings about speech-print match that underlie the alphabetic principle.

Our observations of children writing and reading their own texts also provided opportunities to examine children's earliest understandings of the ways a spoken message is related to print on the page. Of interest here are the emergent forms of speech-print matching produced by two-year-olds. In a recent reanalysis of data from the Write Start! study, our research team has examined how children linked gestures directed at the page with the rhythm of their spoken messages as evidence of their understandings of connections between their written marks and spoken messages. These analyses revealed that two-year-olds in the Write Start! Study ($n = 18$) used several different approaches to linking written and spoken messages.

First, about 6% of children were entirely silent when asked to read their marks and therefore provided no information about their understanding of speech-print matching. For these children, the request to read their writing was novel. Observations of embodied cues indicated that they did not fully understand what they were expected to do.

The most common pattern of speech-print matching for this age group, involved children's use of natural conversational speech to describe their writing, while using gesture to indicate the page as the target of attention. Seventy-one percent of children made these kinds of global connections between their spoken messages and the graphic array on the page. Specifically, children used natural, conversational speech to tell us about their texts, while also using gestures to globally point to the whole page (6%), point briefly to some marks on their page (53%), or to match the timing of the beginning and ending of their talk to the beginning and ending of their pointing (12%). Children using these approaches created global links between their oral messages and marks on the page. While there was no direct evidence that they understood that their marks represented speech, the act of talking about their writing while directing attention to the text set the stage for children's construction of more fine-grained observations about the way speech and print are linked in writing events.

Children using another approach also made global links between speech and print, but with an important change in their oral performance. These children (6% of the sample) began to use a slowed and segmented cadence for the orally produced message while continuing to point globally to the marks on the page. As children slowed and segmented their spoken message using an exaggerated rhythm, their talk no longer sounded like conversational speech. From this approach to speech-print matching, we can infer children were attending in a new way to the role of the spoken message in writing events—especially the ways they had observed adults slowing and segmenting their speech as they read their messages aloud during writing, or the ways adults slowed their reading as they pointed to print. Though there was no one-to-one match between speech units and print segments, these children demonstrated an awareness that the speech stream played a special role in writing.

In a final pattern observed for 18% of two-year-old Write Start! participants, young writers directly explored one-to-one matching of speech and print. These children slowed and segmented speech units (e.g., words, or phrases) and synchronized them to their production of marks during writing or to their pointing when reading already completed writing. As noted earlier, none of the two-year-olds in our study produced conventional letters in their photo captions, so this exploration of speech-print matching occurred as they produced other kinds of writing forms including scribble units, the up and down zigzags of personal cursive, and the individual letter-like forms of personal manuscript. Regardless of the graphic forms used, children displayed conscious awareness that it was the spoken message that was represented in writing and that writing involved matching specific chunks of speech to specific chunks of print.

These observations suggest that watershed insights like the understanding that writing represents a linguistic message that can be spoken aloud are not formed all at once, but instead have their roots in much more global understandings that children begin building as early as age two. Children's understandings of how speech is matched to print develop over time as they begin to talk about their writing and as they attend to the special ways of slowing and matching speech in adults' live demonstrations of writing and reading. Long before young writers begin to invent spellings with letter-sound correspondence, they begin to connect their developing phonological awareness with their observations about print features to explore how speech is matched to and represented by print on the page.

Writing identity

Implicit in these descriptions of what two-year-olds were learning about writing is the broader notion that children also were forming identities as writers. In the emergent literacy classrooms where we conducted our research, young children progressively came to see themselves as the kinds of persons who could actively participate in classroom writing events using their current forms of emergent writing to represent messages and read them to others (Rowe, 2008b). Evidence of their writerly identities is seen in children's eagerness to self-select writing activities as part of their play and to use writing as a form of social action. Over time, children became more active participants in the writing events of their preschool classrooms. While at first children demonstrated their identity as writers primarily by responding to adult requests that they write messages as part of play, over time they proactively took up roles as writers, generating their own purposes for writing and actively bidding for adult attention: "I'm gonna write you." or "I wrote you back!" While writing was collaboratively accomplished with adults, two-year-olds also exerted considerable agency by choosing which writing invitations to take up and by shaping their participation in writing in relation to their personal interests (Rowe and Neitzel, 2010).

Summing up: what children learn about writing before age three

Given opportunities to take part in emergent writing events, children under age three actively explore the graphic potentials of writing tools and surfaces. While at first this is primarily a physical-motor exercise, through interaction with adults, they come to understand that the resulting marks and texts have special significance in their classroom and at home. As they participate in joint writing with adults, they come to see themselves as writers and their texts as writing. As their attention is drawn to their marks and to the writing of adults, they make progressively more refined observations about the visual features of print. Their developing understandings are visible in unconventional writing forms that start as undifferentiated scribbles but soon contain small individual units and combinations of print-like strokes. As adults inquire about the meaning of these texts, children learn to respond by reading their marks as messages that are generally and then more specifically related to the ongoing social event. In these writing interactions, children begin to form global understandings that their marks and spoken messages are somehow related, and some children start to specifically explore the match between speech and the marks they make on the page.

Overall, two-year-olds who have opportunities to observe and participate in writing events, begin to turn their attention to both the social and technical aspects of writing. When looking at individual children's trajectories over time, our data (Rowe and Wilson, 2015) show that children move from global to more specific and conventional understanding for all features of print over time, though not all aspects are equally well-developed at any time point. As two-year-olds, children already show variability in their understandings about different aspects of print and in the pacing of learning. Variability is the rule rather than the exception.

How do children learn to write before age three?

As noted in the introduction to this chapter, there is no universal process for learning to write. Instead, what and how children learn is shaped by the writing practices of their larger communities. For young children, learning to write is further shaped by culturally based patterns of childrearing and guided participation (Morelli et al., 2003; Rogoff, 2003) that include adults' perceptions of their roles as caregivers and teachers, their cultural models for child development, their preferred arrangements for childhood activities, and the nature of adult-child communication.

The observations of young writers reported in this chapter occurred in preschool settings where teachers and researchers engaged two-year-olds in playful interactions involving writing as a means of supporting their literacy learning. Children were invited into learning-to-write events to watch and participate alongside more experienced adult writers. They were encouraged to playfully participate as writers using their current graphic strategies, regardless of their conventionality. Adults provided demonstrations



Fig. 2 Reagan's photo page.

of how writing should be done by writing collaboratively with children as part of play activities and by offering occasional bits of direct teaching in the context of the child's writing. Very little of this interaction looked like formal school lessons. Instead, as teachers intended, it looked and felt like play.

But the question remains: *How* do children begin to learn to write in these kinds of writing events? When an adult asks a two-year-old to write about a photo and then to read their marks, as we did in the Write Start! study (Rowe and Wilson, 2015), how do children learn what it means to "write"? What resources do they call on to participate the first time this request is made, and then thereafter? How do they learn to produce graphic responses that are increasingly more print-like? How do they learn to initiate writing interactions and use their texts to mediate social interactions?

To provide some context for considering these questions, in the follow section I provide a transcript from the Write Start! study recording the first time I asked two-year-old Reagan (a pseudonym) to write in his preschool class. In this early fall event, I first demonstrated making a page for a class photo book and then invited him to make his own page. Specifically, I asked him to write his name and a caption for a photo showing him building with wooden blocks in his classroom—an event that had occurred the day before. Fig. 2 shows Reagan's completed photo page.

Reagan's first invitation to write a message

1.	Rowe:	I point to the white space above Reagan's photo. "First, write your name up here. Write Reagan up on top just like I did. I wrote. 'Ms. Debbie'."
2.	Reagan:	Marks on his torso in the photo. "There!"
3.	Rowe:	"Yeah, you wrote on top of the picture." I point to the white space above the photo a second time. "Can you write up here? Write your name?" Reagan moves the paper closer to his body but doesn't respond.
4.	Rowe:	I tap the photo several times. "What do you think you should write about this picture?"
5.	Reagan:	Reagan reaches across the table for my photo page and lays it in his own workspace. He makes some marks on top of my photo. "You!"
6.	Rowe:	I laugh. "You wrote on top of me, too, didn't you?" I take my page and put it away. I once again slide Reagan's page into his workspace.
7.	Rowe:	"Ok! Down here!" I point to the white space below the photo. "Reagan, let's write something about your picture."
8.	Reagan:	Reagan is already adding more marks on top of the part of the photo showing his body. "My hair!"

(Continued)

9.	Rowe:	"That was your hair!" I point again to the space below the photo. "Can you write down here, I am playing with blocks?" I point to the blocks in the photo.
10.	Reegan:	Reegan begins to use the marker to make marks directly on the skin of his face.
11.	Rowe:	I gently move the hand holding the marker away from his face. "Don't write on your real face. We'll have to scrub it off!"
12.	Reegan:	"Yeah, yeah! You get a wupping!"
13.	Rowe:	"We don't want to get a wupping and we don't want to write on your face." I tap the page several times for emphasis. "Paper. Paper. We write on paper!"
14.	Reegan:	Reegan acts out washing his face by rubbing his hand over the mark and making a pretend washing sound "Shu Shu." He returns to drawing on top of the photo.
15.	Rowe:	"Alright! Now write your name up here." I draw my finger left to right in the white space above the photo. "So, we'll know, will remember, whose picture that is."
16.	Reegan:	Reegan draws a large circle around the entire photo. When his marker runs of one edge of the page onto the table, he is concerned: "Oh! Oh!"
17.	Rowe:	"That's OK. We'll wipe it off. This table is washable."
18.	Reegan:	"It's washable," he says, then continues drawing large circles around the photo.
19.	Rowe:	I touch his last circle. "Read that to me. What did you write there?"
20.	Reegan:	"Reegan write!," he says with excitement.
21.	Rowe:	I repeat his message: 'Reegan write.' "Good for you!" I tap the part of the photo showing his head. "Yep, right on top of your face, you sure did write on it."
22.	Rowe:	"Write down here for me." I touch the white space below the photo. "Write 'I am playing with blocks'"
23.	Reegan:	Reegan makes more marks on the part of the photo where his face is pictured.
24.	Rowe:	"Can you do that?"
25.	Reegan:	Reegan continues to make more marks on the photo.
26.	Rowe:	"Alright!" I take the page and begin to write below the photo, saying the words aloud as I write. "Let's put, 'I ... am ... playing ... with ... blocks'"
27.	Reegan:	He has paused and is watching me write. As soon as I complete the printed sentence, he begins drawing on top of his photo again.
28.	Reegan:	Reegan turns from the table and calls across the room to his friends. "Look! Look! Da Reegan! Everybody know Reegan!" He sits back with satisfaction and begins to make more marks on the photo.

Using what is "at hand" to participate in writing events

Though being asked to write was initially a novel request for the two-year-olds in our studies, children had at hand a variety of physical and social experiences to support their initial participation as writers. For example, based on his previous physical explorations of writing tools and mark making, at Turns 2, 5, and 8, we see Reegan confidently picking up a marker and beginning to mark on his photo page and mine. At Turns 12 and 16, his talk and actions reference social rules for using markers learned in previous experiences outside school that might be glossed as "write on paper, not your skin or the furniture." Though an adult would have expected him to write in the white space of the page, the target for his marks was the photo itself, where he proceeded to color on top of the regions of the photo showing his body. When I directed him to write "right up here" at Turn 15 and dragged my finger across the white space, Reegan complied by shifting his marks (mostly) to the white space, this time making large circles that roughly traced the boundaries of the photo—the topic of our conversation. When I suggested a message for his photo caption ("I am playing with blocks.") at Turn 22 and pointed to white space below the page, he returned to marking on top of the part of the photo showing his face. Finally, based on his experience in other events where he and his peers had enjoyed looking at photos taken in the classroom, at Turn 28 he used his photo page as an occasion to loudly bid for the attention of his classmates as they played in other parts of the room.

This example demonstrates that young children do not start with a blank slate even when they first respond to invitations to write. Instead, they use their existing understandings of social routines and events to support their participation. In the example, Reegan drew on at least two kinds of familiar events: exploratory "art" events where he and peers used markers and paper, and events where he and peers viewed and talked about photos of themselves and classmates. These familiar social routines provided Reegan with skills needed to take his turn in this writing event by talking about his photo, making marks on paper, and using his photo page to launch social interaction with his peers.

Though Reegan's graphic product had few, if any, print-like features, the marks were not random. Instead, they built on his ability to communicate with gesture. He directed his marks to the part of the photo that was relevant to our conversation—his body and the building blocks. When he moved his marker into the white space, he used it to inscribe the boundaries of the photo—the focus of our joint attention in this event. Like many two-year-olds in our studies (Rowe, 2019), Reegan's initial writing responses piggybacked on both his ability to make marks and to physically point to objects of joint attention. As Vygotsky (1978, 1997) and Luria (1998) noted, children's first forays into writing often involve the production of marks that are created

by pointing with a pen in hand. In this way, children like Reegan used inscribed gestures, literally pointing to the page with a pen, to graphically participate in writing at a time when they were just beginning to form understandings about print forms and processes (Rowe, 2019). Much like the role of pointing in oral language learning (Butterworth, 2003), these pointing marks launched children into writing by providing an initial strategy for meaningful graphic participation.

Adult demonstrations of what, how, and why to write

The event recorded in the example is also typical of the many opportunities that children had to write with and alongside adults. In these events, children's texts were foregrounded as an important focus for conversation and joint social action, drawing children's attention to the social importance given to writing products in their classroom. When adults invited children to write (Turn 4: "What do you think you should write about this picture?") or asked them to read their writing (Turn 19: "Read that to me. What did you write there?") they implicitly named children as writers and their texts as writing. These interactions also provided children with information about the roles, text types, and textual intentions adults viewed as appropriate for two-year-old writers.

While adult-child talk has received the most attention from literacy researchers, my interactions with Reegan demonstrate the importance of gesture and other embodied cues in these writing events. In the transcript, almost every turn used gesture or handling objects as cues to meaning and expected activity. Adults used gesture directed at the page to provide physical demonstrations of expected locations for print (Turns 1, 3), the linear arrangement of marks, and left-to-right directional patterns (Turn 15). Adults also provided implicit demonstrations of speech-print matching as they slowed and synchronized pointing and speech while reading their writing (Turn 26).

Because writing generated lasting graphic traces of adult and child activity, texts could be revisited and examined visually, verbally, and physically through gaze, conversation, and pointing. Child and adult texts served as sites for publicly negotiating writing content, processes, and purposes. As seen in my interactions with Reegan, it was at these moments that adults sometimes provided direct teaching about how writing works within the context of the ongoing shared activity. For example, after Reegan produced several unconventional responses to my request to write a message about the photo, at Turn 9 and again at Turn 22, I suggested a specific message that would be seen as an appropriate photo caption. These kinds of adult responses offered individualized, context-specific information about expected message content, writing processes, or social purposes for writing. In shared writing activities, children had socially supported opportunities to begin learning how writing functions to represent meaning.

Top-down and bottom-up processing in shared activities with common goals

Microanalysis of preschool writing events showed that learning to write was supported by the same social mechanisms that support language learning more generally. Tomasello (1992, 2008; Tomasello et al., 2007) has argued that communication, whether through pointing or speech, is built on the same foundational infrastructure, namely prosocial motivations for sharing with and helping others and the ability to establish joint attention with others inside shared activities with common goals. Bruner (1986) has further shown that for young children, most language is learned through joint participation with adults in familiar activities and collaborative routines that are part of everyday life. Our analyses suggest that these observations also hold true for learning to write.

Children in our studies learned to participate and to form shared goals in learning-to-write events through the use of top-down processing. That is, children used their understandings of the meaning and purpose of the shared activity to infer the likely meaning of the adult partner's talk, gestures, and actions. In the transcript at Turn 19, I touched the large circles Reegan had drawn around the photo and asked him to read his marks. Though this was a new kind of request, he used his understanding of our shared book-making activity to respond, "Reegan write!" Similarly, adults used their understanding of the shared activity to predict what young children were likely to mean, even if their talk and graphic activity were emergent and unconventional (e.g., Turns 7–9). Preschoolers and their partners made heavy use of deictic gestures to communicate their intentions, so adults and children actively attended to cues provided by the partner's gestures (Goldin-Meadow, 2014). Participants followed the direction of the other person's pointing to determine their partner's attentional focus. This allowed them to determine the meaning of their partner's gesture in relation to the jointly constructed shared common ground, that is "everything we both know (and know that we both know...)" (Tomasello, 2008, p. 75) about the ongoing activity.

Additionally, participants in these face-to-face interactions closely tracked cues to their partner's intentions provided by handling and placing objects, body posture and movement in space (Andr n, 2010; Tulbert and Goodwin, 2011). Through top-down processes, children learned the meaning of the unfamiliar words and novel requests produced by adults in learning-to-write events. Through bottom-up processing (Tomasello, 2008), children also learned that when an adult used a new bit of language or pointed to an object, they were being invited to share the adult's attentional focus; that is, they should notice what the adult was attending to as a means of understanding the meaning of the unfamiliar oral language, print, or gesture. In this way, children began to attend to and notice meaningful patterns in the visual features, message content, and social functions of the graphic arrays produced in shared writing activities.

Summing up: how children learn to write before age three

As Vygotsky (1978) pointed out, writing development does not happen all at once. Instead, it occurs as “the result of a long series of developmental events” (p. 57). Children, like Reagan, do not immediately understand that print stands for language when an adult first suggests a message and demonstrates how to record it with print. Instead, over time, children form increasingly more sophisticated understandings about print through many small moments of social interaction around writing.

When children produce marks in events culturally framed as writing, the resulting graphic array becomes a public site for adult-child negotiations of the semiotic properties of the child’s marks (cf. Goodwin, 2003). Children’s texts provide opportunities for adults to reframe children’s marks as if they were symbolic representations of linguistic messages. Through these interactions, adults provide individually tailored demonstrations of the content, processes, and purposes of writing.

Despite the powerful roles of adults in preschool writing events, learning to write occurs as a collaborative activity in which children exert considerable agency. They choose which suggestions for writing interactions they will take up and perform them in ways that are in line with their current understandings of writing. Children frequently renegotiate or ignore adult suggestions and invitations in favor of writing that is more in line with their own personal interests (Rowe, 2008b; Rowe and Neitzel, 2010).

For the two-year-olds in our studies, learning to write was not accomplished individually, but instead in joint interaction as adults and children acted and reacted to one another around writing. In writing events, children used both bottom-up and top-down processing to learn about the content, processes, and purposes for the kinds of writing used in their classroom community. Children learned to write through peripheral and gradually more central participation in play-based writing events (Lave and Wenger, 1991).

Concluding thoughts

In this chapter, my goal has been to describe two-year-olds’ participation in emergent writing events, what their texts looked like, and what they were beginning to learn about print and themselves as writers. The data clearly show that these youngsters began to form foundational understandings about writing through observation and playful interactions with adults. Against the backdrop of these findings, a colleague in another field once asked me, why should we care about writing for children this young? As he put it, “Can’t writing wait until they are older?” As I conclude this chapter, I offer three reasons why research on writing under age three is important.

First, children living in literate societies are immersed in writing events in the normal course of everyday life. Children aren’t willing to wait! Writing is part of their sociomaterial environment and they explore it even if there is no intentional effort by adults to engage them as writers. A case in point was provided by a parent who showed me her checkbook where her preschool son had used neat rows of personal cursive to fill in all the blank lines on the paper checks he found in his mother’s purse! While these childlike forays into writing are amusing, they also establish that very early on, children feel the social press to learn about and participate in writing because it is an important part of their experiences at home and in the community. Children under age three commonly see their siblings sitting at the table doing homework or a parent jotting items on a grocery list. Children have natural curiosity and motivation to participate in these kinds of everyday routines. Given the importance of learning to read and write in later schooling, these early forays into writing are learning opportunities that are too good to miss!

Second, research on children’s writing before age three can help adults recognize and value the very beginnings of writing as seen in children’s unconventional texts. This is important because it has implications for the ways adults interact with young children around writing. Adults often overlook the power of one- and two-year-olds as creative and active literacy learners because their participation in writing events is so unconventional. Faced with a page of scribbles, adults are more likely to ask children, “What did you draw?” than asking them “What did you write?” –even if the social occasion is one in which writing is foregrounded. Though the present chapter provides an initial portrait of two-year-olds’ writing responses, much more research is needed to understand how this age group and even younger children begin to learn to write in different kinds of social situations, and with different kinds of adult supports. We need to better understand the normal variability in the ways that children approach writing.

Third, research on children’s earliest playful experiences with writing provides clues to the social and material conditions that support their participation as writers. This is important information for parents and early childhood educators whose understandings and values largely shape the writing opportunities and kinds of adult supports that children are offered. For very young children, learning to write occurs in play and everyday activities rather than in formal lessons. In our studies, adults invited children to write for purposes that made “human sense” (Donaldson, 1978) in relation to their play and the social scene, accepted and responded with interest to children’s texts and talk regardless of conventionality, asked children to read their marks, and found many ways to celebrate and value children’s participation in writing. Adults intentionally wrote collaboratively or alongside children and looked for opportunities to provide quick bits of individualized instruction tailored to the child’s current writing project. While these kinds of interactions fit well with current notions of developmentally appropriate preschool learning environments (Copple and Bredekamp, 2009), there is still much to learn about the ways that parents and early childhood teachers can invite children into writing at home and school.

To better understand the beginnings of writing, in the future, researchers will need to collaborate across theoretical lines to study learning-to-write as a both a sociocultural and cognitive enterprise. Single perspective studies, while helpful, will always fall short of understanding the complex interactions between cultural, social, and cognitive facets of learning to write in early childhood. To

move the field forward, future research needs to simultaneously ask and answer questions about cultural practices around writing in early childhood, writing identity, ways of collaborating in face-to-face interactions, and children's sociocognitive strategies for learning about writing.

References

- Andr n, M., 2010. Children's Gestures from 18 to 30 Months. Dissertation. Lund University, Lund. <http://www.lunduniversity.lu.se/lup/publication/1700528>.
- Baghban, M., 1984. Our Daughter Learns to Read and Write. A Case Study from Birth to Three. International Reading Association.
- Baghban, M., 2007. Scribbles, labels, and stories: the role of drawing in the development of writing. *YC Young Children* 62 (1), 20.
- Bruner, J., 1986. *Actual Minds, Possible Worlds*. Harvard University Press.
- Butterworth, G., 2003. Pointing is the royal road to speech for babies. In: Kita, S. (Ed.), *Pointing. Where Language, Culture and Cognition Meet*. Lawrence Erlbaum, pp. 9–34.
- Clay, M.M., 1975. What Did I Write? Heinemann.
- Copple, C., Bredekamp, S. (Eds.), 2009. *Developmentally Appropriate Practice in Early Childhood Programs*, third ed. National Association for the Education of Young Children.
- Donaldson, M., 1978. *Children's Minds*. W. W. Norton & Co.
- Dunst, C., Gorman, E., 2009. Development of infant and toddler mark making and scribbling. *Centre Early Learn. Lit. Rev* 2 (1).
- Goldin-Meadow, S., 2014. How gesture helps children learn language. In: Arnon, I., Casillas, M., Kurumada, C., Estigarribia, B. (Eds.), *Interaction: Studies in Honor of Eve. V. Clark*. John Benjamins Company, pp. 157–171.
- Golomb, C., 1974. *Young Children's Sculpture and Drawing: A Study in Representational Development*. Harvard University Press.
- Goodwin, C., 2003. Pointing as situated practice. In: Kita, S. (Ed.), *Pointing: Where Language, Culture, and Cognition Meet*. Lawrence Erlbaum Assoc, pp. 217–241.
- Harste, J.C., Woodward, V.A., Burke, C.L., 1984. *Language Stories and Literacy Lessons*. Heinemann.
- Kindler, A.M., Darras, B., 1994. Artistic development in context: emergence and development of pictorial imagery in the early childhood years. *Vis. Arts Res.* 20 (2), 1–13.
- Lancaster, L., 2003. Moving into literacy: how it all begins. In: Hall, N., Larson, J., Marsh, J. (Eds.), *Handbook of Early Childhood Literacy*. Sage, pp. 145–153.
- Lave, J., Wenger, E., 1991. *Situated Learning. Legitimate Peripheral Participation*. Cambridge University Press.
- Lee, B.Y., 2013. Early literacy development in toddlerhood: publication trends from 1990 to 2009. *Early Child. Educ. J.* 41 (1), 25–37. <https://doi.org/10.1007/s10643-012-0537-6>.
- Levin, I., Bus, A.G., 2003. How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. *Dev. Psychol.* 39 (5), 891–905.
- Luria, A.R., 1998. The development of writing in the child. In: de Oliveira, M.K., Valsiner, J. (Eds.), *Literacy in Human Development*. Ablex Publishing Corporation, Stamford, CT (Originally published 1929).
- Matthews, J., 1984. Children drawing: are young children really scribbling? *Early Child. Dev. Care* 18, 1–39.
- Matthews, J., 1999. *The Art of Childhood and Adolescence: The Construction of Meaning*. Falmer Press.
- Mej a Arauz, R., Rogoff, B., Paradise, R., 2005. Cultural variation in children's observation during a demonstration. *IJBD (Int. J. Behav. Dev.)* 29 (4), 282–291.
- Morelli, G.A., Rogoff, B., Angellillo, C., 2003. Cultural variation in young children's access to work or involvement in specialised child-focused activities. *IJBD (Int. J. Behav. Dev.)* 27 (3), 264–274.
- Otake, S., Treiman, R., Yin, L., 2017. Differentiation of writing and drawing by U.S. two- to five-year-olds. *Cognit. Dev.* 43, 119–128. <https://doi.org/10.1016/j.cogdev.2017.03.004>.
- Read, C., 2009. Learning to use alphabetic writing. In: Beard, R., Myhill, D., Riley, J., Nystrand, M. (Eds.), *The SAGE Handbook of Writing Development*. Sage, pp. 260–270.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press.
- Rogoff, B., Mistry, J., Goncu, A., Mosier, C., 1993. *Guided Participation in Cultural Activity by Toddlers and Caregivers*, vol. 58. Society for Research in Child Development.
- Rowe, D.W., 2008a. Social contracts for writing: negotiating shared understandings about text in the preschool years. *Read. Res. Q.* 43 (1), 66–95.
- Rowe, D.W., 2008b. The social construction of intentionality: two-year-olds' and adults' participation at a preschool writing center. *Res. Teach. Engl.* 42 (4), 387–434.
- Rowe, D.W., 2019. Pointing with a pen. The role of gesture in early childhood writing. *Read. Res. Q.* 54 (1), 13–39. <https://doi.org/10.1002/rrq.215>.
- Rowe, D.W., Neitzel, C., 2010. Interest and agency in two- and three-year-olds' participation in emergent writing. *Read. Res. Q.* 45 (2), 169–195.
- Rowe, D.W., Wilson, S., 2015. The development of a descriptive measure of early childhood writing: results from the Write Start! Writing Assessment. *J. Literacy Res.* 47 (2), 245–292. <https://doi.org/10.1177/1086296X15619723>.
- Rumelhart, D.E., 2004. Toward an interactive model of reading. In: Ruddell, R.B., Unrau, N.J. (Eds.), *Theoretical Models and Processes of Reading*, fifth ed. International Reading Association, pp. 1149–1179.
- Schickedanz, J.A., 1990. *Adam's Righting Revolutions: One Child's Literacy Development From Infancy Through Grade One*. Heinemann.
- Teale, W., Sulzby, E., 1986. Introduction. Emergent literacy as a perspective for examining how young children become writers and readers. In: Teale, W., Sulzby, E. (Eds.), *Emergent Literacy*. Ablex, pp. vii–xxv.
- Tomasello, M., 1992. The social basis of language acquisition. *Soc. Dev.* 1 (1), 67–87.
- Tomasello, M., 2008. *Origins of Human Communication*. MIT Press.
- Tomasello, M., Carpenter, M., Liszkowski, U., 2007. A new look at infant pointing. *Child Dev.* 78 (3), 705–722. <https://doi.org/10.1111/j.1467-8624.2007.01025.x>.
- Treiman, R., Yin, L., 2011. Early differentiation between drawing and writing in Chinese children. *J. Exp. Child Psychol.* 108, 786–801. <https://doi.org/10.1016/j.jecp.2010.08.013>.
- Tulbert, E., Goodwin, M.H., 2011. Choreographies of attention: multimodality in a routine family activity. In: Streeck, J., Goodwin, C., LeBaron, C. (Eds.), *Embodied Interaction. Language and Body in the Material World*. Cambridge University Press.
- Vygotsky, L.S., 1978. *Mind in Society*. Harvard University Press.
- Vygotsky, L.S., 1997. The prehistory of written language. In: Rieber, R., Hall, M.J. (Eds.), *The Collected Works of L. S. Vygotsky, The History of the Development of Higher Mental Functions*, vol. 4. Springer, Boston, MA, pp. 131–148. https://doi.org/10.1007/978-1-4615-5939-9_7 (Originally published 1931).

Children's earliest insights in developing literacy

Liliana Tolchinsky, Department of Catalan Philology and General Linguistics, University of Barcelona, Barcelona, Spain

© 2023 Elsevier Ltd. All rights reserved.

The pre-phonological roots of literacy	39
Sensory motor roots	39
Neurobiological underpinning	39
Historical grounding	39
Initial research approaches	40
Children's mobilization of "action schemas"	40
Foundational schemas	40
Sensitivity to orthographic patterns	41
Early writing patterns	42
Meaning attribution and written language	43
Early attribution of meaning	43
Conclusion	45
Writing and language	45
On the nature of literacy development	45
References	46

This is a paper about children's earliest attempts to make writing their own. Learning how to write and read is an integral part of language development, it is not an added skill. There is no aspect of language use—from storytelling to understanding metaphors, from colloquial to academic usages—that remains untouched by literacy and, conversely language development has a crucial impact on developing literacy (Tolchinsky and Berman, in press).

Nonetheless, studies that have examined spelling and reading from a developmental viewpoint typically identify a developmental phase in which children do not associate the graphic signs of writing with language. This phase has been called: pre-communicative, pre-phonetic, pre-phonological (e.g., Ehri, 1995; Gentry, 2000; Winch et al., 2010; Zhang and Treiman, 2015) with children in this period generally being labeled as "pre-literate" (Gombert and Fayol, 1992).

This chapter examines the key features of this "pre" phase, in particular how children single out writing as a particular type of notational system, and what representational status they attribute to writing if not to represent language. As will become clear, I am reluctant to refer to children in this phase as pre-literate, on the grounds that literacy begins to develop even before children can read and write using littera "alphabetic letters."

It is not easy to pinpoint the precise age range of this pre-phase. On the one hand, the fluctuating nature of development is particularly salient during early childhood (Dawson-Tunik et al., 2005). On the other, the variability between learners is huge even in controlled teaching environments with a shared curriculum (Campbell et al., 2019). The fluctuating nature and inter-subject variability in the domain are exacerbated when one compares findings from different languages (English, Hebrew, Spanish), periods in history (from the 1930s to the present day) and contexts (e.g., at school, in the laboratory, at home), as I plan to do in what follows.

In literate communities, the first social act takes a written form since a child "shall be registered immediately after birth".¹ In such environments, children grow up surrounded by different types of notations (road signs, advertisements, commercial labels), they are exposed to letters on T-shirts and stickers; letters of the alphabet are named in songs and poems. And in many countries, letters are taught even in nursery schools, where children are expected to be able to recite part of the alphabet, to recognize letters as part of their own name (declaring "that's me!"), even learning to draw letters. If they are lucky enough, young children also participate in family practices of the kind that Lareau (2000) called concerted cultivation. But even such children have not yet discovered that alphabetic writing represents the sound dimension of language.

My purpose in this chapter is thus to clarify what comes in development before writing is phoneticized. Given that exposure to notational input and participation in literacy practices is a key condition for developing literacy, here I address two related questions: first, what information do children extract from the notational environment and, second, what meaning do they attribute to such information?

The chapter is organized as follows. It starts with a brief summary discussion of the sources that underlie the pre-phonological phase of literacy, the sensorimotor roots of conceptual development, and the neurological underpinnings of the initial disconnection between the visual material provided by writing and the language children talk. This is followed by a short historical account of the germs of writing to highlight both the sensorimotor/manipulative its origin and the role of users' interpretation in linking visual markings to language. The second part of the chapter illustrates basic action schema that children mobilize to obtain a desired

¹Article 49 Article 7 1 the UN Rights of the Child (September 1990).

outcome from literacy related artifacts like cellphone screens and books and implements for writing like pencils or crayons. And I consider the kind of information children extract from the printed environment by applying learning mechanisms that enable them to identify graphic regularities. The third part of the chapter focuses on children's attribution of meaning to writing before they realize that writing represents the sound properties of spoken utterances. I conclude with some reflection on the nature of developing literacy, the inborn capacities that it entails and its reliance on explicit teaching.

The pre-phonological roots of literacy

Sensory motor roots

Like any other domain of knowledge, literacy is rooted in early sensorimotor experience that are crucial for the development of later cognitive skills and abilities. As explained by [Wellsby and Pexman \(2014\)](#), sensorimotor experience refers to "a range of experiences that typically involve an action being performed on an object, either by a child directly, or through observation of another's actions. The experience is multisensory, derived primarily from visual, tactile, and proprioceptive senses" (p. 9). This quite broad characterization of sensorimotor experience in the form of actions performed on specific objects is more precisely defined in earlier works, such as Piagetian epistemology, Bruner's account of the "attainment of competence" in small children, and the notion of "affordance" of James and Eleanor Gibson. These three perspectives guide my interpretation of the earliest literacy development, as supported by current neurobiological accounts of conceptual development that also involves other systems such as emotions (e.g., [Pulvermuller, 2013](#)).

Neurobiological underpinning

Nothing in our evolution prepares us to receive language through the eyes. Nonetheless, current brain-imaging techniques show that literate adults possess highly evolved mechanisms adapted to reading which show up in brain areas that are activated during reading. [Dehaene \(2009\)](#) defines literacy as a case of neuronal recycling because the same neural mechanisms and brain region that are found in the macaque species of monkeys are those relied on in human reading. The mechanisms that genetically enable object discrimination are recycled to be able to carry out the necessary discriminations required for reading. Learning to read would be a remarkable instance of brain plasticity.

Studies applying neuroimaging techniques specify that the shape of letters is recognized in the occipito-temporal region in the left hemisphere, and that this information is transmitted to diverse areas of the left hemisphere that are in charge of attributing sense, sonority, and articulation to written characters. In other words, learning to read consists in connecting the visual with the linguistic areas of the brain. These interconnections are bidirectional from visual to linguistic areas and vice versa.

On the other hand, and a finding crucial to our topic, recent research on the growing connections between visual and language regions observed in children in the process of learning to read shows that the activation of the left hemisphere observed in adults did not occur in children who still had not learned to read. The latter, instead, revealed activation in the right hemisphere with no link between visual forms and language. Research did, however, note activation of the neural networks associated with reading among children from the age of 7 years. By this time, the area attending to the visual form of words is already distinguished in the left occipito-temporal region as well as in temporal regions ([Dehaene, 2009](#)). [Shaywitz and Shaywitz \(2008\)](#), who conducted one of the few longitudinal studies in this domain, show that reading improves the activation of this region, and that this increase depends more on reading level than on chronological age, suggesting that this level of activation does not in itself reflect brain maturation. Linking visual information to language, activating language while looking at written displays is a consequence rather than a prerequisite for learning to read. Thus, the interest in inquiring into the meaning children attribute to written displays before they establish its link to language. From the perspective of both evolving cognition and neurological underpinnings, we assume that developing literacy builds on sensorimotor experience. Moreover, it is likely that, alphabetic writing follows a comparable historical evolution.

Historical grounding

Historically, the precursors of writing were collections of clay tokens or counters of various geometrical forms: spheres, discs, cones, and tetrahedrons devised to show number and quality of merchandise for trading, to know, say, how many sheep, or how many crops are for sale. This system of artifacts was used by ancient Sumerians living in what is now southern Iraq from the ninth millennium BCE onwards (e.g., [Schmandt-Besserat, 1998](#)). The forms representing the animals or products were not figurative but rather arbitrary symbols that differentiated kind on the basis of height, size, or form. The system was very secure, but each time recounting or checking of merchandise had to be undertaken, the bullae had to be broken. The solution to this problem can be considered "the great invention": People started to mark the outside of the clay bull, when still wet, so as to indicate what was inside. The impressions on the bullae were signs of the tokens inside them, which in turn were signs of the counted items. They were "second order signs," signs of signs. When tokens were used for recording, they could be directly manipulated to control the transaction. But the impressions on the bullae could not be manipulated for checking the amount of merchandise; they had to be interpreted. Thus, these second-order signs enabled (and demanded) a new activity: Interpretation by thinking or by saying something. It is in the interpretation of these impressions that the link to language was established ([Coulmas, 1989](#)). The (verbal) interpretation of the

marks by the traders and a process of abstraction of these marks led to the emergence of the cuneiform writing system many centuries later (around 3500 BCE). Only then did the marks on clay allude to language leading to the expansion of writing as a system for representing language. This is clearly an oversimplified account of the historical emergence of writing. Yet it is worth considering because it points to the reasons that led to the emergence and subsequent evolution of writing, and to a direction of development that goes from the manipulation of physical objects via the creation of physical marks (that provide visual information) to an interpretation that made it possible to link the marks to language.

Initial research approaches

Research interest in pre-phonological features of literacy acquisition originated in the late 1930s, with the work of Luria and Vygotsky on “the prehistory of written language” (Luria, 1929/1978). Almost 50 years later, this interest was renewed in three different languages and cultures. In English, it followed, initially, an ethnographic, sociocultural approach in what is termed “emergent literacy.” Even children who have not yet started formal instruction in literacy are asked to compose phone messages, food-recipes, birthday cards, letters to the parents, and so on. The different categories and stages of early writing are obtained from observing how kindergartners and preschoolers plan, write and read back different types of texts (e.g., Clay, 2001; Dyson, 1987; Goodman, 1980; Sulzby, 1986). In Spanish, it arose in the pioneering work of Ferreiro (Ferreiro and Teberosky, 1979/1982; Ferreiro, 1986a, 1994) with the aim of tracing “the psychogenesis of written language.” This line of work expanded to other Romance languages such as Italian and Portuguese in different Latin American and European countries. Studies were carried out using relatively structured clinical interviews following Piagetian methodology. Children from 3 to 6 years of age were asked to sort written materials, to match pictures to written labels, to create blanks in written sentences as well as to write words, sentences, and stories. Investigators observed both the graphic characteristics of the written products and the verbal explanations given by children while writing and after having written. Their aim to discover what children know about the writing system, their knowledge of the function and social uses of writing. A third line of work, in Hebrew, addressed similar issues, but applying a more controlled methodology (Tolchinsky Landsmann and Levin, 1987; Levin and Tolchinsky Landsmann, 1989). As we note next, many of the findings elicited applying mainly qualitative methodologies were later replicated using more controlled paradigms, larger samples, and careful analytical procedures.

Children's mobilization of “action schemas”

Foundational schemas

While I was considering how to approach the sensorimotor schemas on which literacy builds, I recalled a scene I witnessed a few months ago at a supermarket. A young mother was queuing for the cashier with a small child, maybe 20–22 months, on a baby trolley. Seeing how long the line was, she searched in her bag and took out a colored children's book about penguins. The little boy took the book in his hands and started to bang/tap on the penguin's face. Realizing that nothing happened, he gazed incredulously at his mother, who immediately responded by instructing her child: *así no “not like that!”*, *ábrelo y pasa las páginas “(you) open it and turn the pages.”* I was enchanted by the scene, although friends of mine, parents of small children, did not find anything particularly remarkable. They took it for granted that tapping on a tablet is an established schema for a young child.

In 2015, data from (non-representative) samples of the American population suggest that children's access to mobile technology may be approaching saturation; 96.6% of children under 4 reportedly have used mobile devices at home. This trend is not exclusive to American children. Estimates from other industrialized countries similarly indicate that toddlers and preschoolers are introduced to mobile technology at an early age, with many now using touchscreens daily (Kabali et al., 2005).

The scene at the supermarket illustrates how an action schema, as a sequence of actions afforded by a given object which had been successful to attain a goal is intentionally mobilized to obtain a similar goal afforded by similar object. The scene serves to exemplify what Bruner described, long ago, as the “attainment of competence” in infants, and also the notion of “affordance” advanced by the Gibsons. As Bruner put it, “In the growth of such competence in infants, three themes are central: intention, feedback, and the patterns of action that mediate between them” (Bruner, 1973, p. 1), a description that underlined the intentional and anticipatory nature of an action. Exploratory activity, even at a very early age, is controlled by some anticipation of an outcome, presumably an adaptive one. These ideas have strong implications for literacy development.

Should children be exposed to books—as carriers of texts—that require other actions to obtain a desired outcome and guided by a well-disposed adult (such as the mother in the supermarket scene who had the time and patience to explain to her child what to do), they will include a book-reading schema in their repertoire and apply one or the other when appropriate taking control of the differing action-schemas called for.

Snow and Ninio (1986) carried out a series of studies addressing a similar question, by observing mother-infant dyads in the process of picture book reading. They concluded that from 8 to 18 months of age children “the child had progressed from an attempt to eat the page to being able to participate fully in verbal dialogue” when looking at the book (p. 116). Children grasp the physical actions involved in reading—gazing, pointing, monitoring—and also become familiar with the verbal behavior afforded by books (Bus et al., 1995). This engagement is critical for children learning to link graphic information to language.

Similar findings were obtained from Spanish-speaking toddlers. At the age of two, children believe that they must look at the pictures in order to discover the story, but within a few months, they start to make distinctions between text and pictures (Ferreiro,

1986a). They give each a different name and state that the text and not the pictures must be read, “because there are letters.” That is, children realize that books are a particular kind of object—different from food, toys, pets or people—an object you look at and say something in a loud voice. They also realize that different information is extracted from pictures than from the accompanying letters.

Well before the invention of mobile technology, Gibson and Yonas (1966) demonstrated a similar intentional and anticipatory nature in small children's control of graphic production. They assumed that “The marking of traces, controlling of the displayed trace, and the seeing of these new display variables were assumed to motivate the act of scribbling, not the transient feedback from the activity itself” (Gibson and Yonas, p. 155). Children from 15 to 38 months were presented with paper, a tool that left a trace and another that did not. The researchers provide a vivid account of infants' discovering “a fundamental graphic act” (Gibson, 1988) when given the tool that left a trace. When handed the nontracing tool, their scribbling was significantly reduced and then abandoned.

Small children raised in isolated communities have an internal representation of writing—to write means the production of marks on paper—but reading has no definite meaning, it is confused with writing (Ferreiro, 1986a,b). For example, children who had just jotted down some marks on a sheet of paper and were then asked to read them back responded that they had already done so. Other children claimed that pencils were needed for reading. Such comments are not a simple matter of terminology, they indicate, rather, that the nature of writing is understood at an earlier age because it leaves visible traces; it changes the object visibly, while reading does not.

This insight was summed up several decades ago Gibson and Levin (1975), “For the human species, at least, the marking of nonpictorial graphs is nevertheless interesting and has a genetic development which seems to be to some extent self-directed” (p. 231). This has been shown by several studies in different languages, as we explain below.

Sensitivity to orthographic patterns

Small children develop increasing responsiveness to the orthographic patterns of the written language they are exposed to prior to learning to read conventionally. This responsiveness might build on learning mechanisms that enable the extraction of statistical patterns in auditory and visual input. The effect of such mechanisms is well recognized for early language acquisition where studies show, for example, that infants develop early sensitivity to phonotactic patterns in their native language. Phonotactic constraints define what sound sequences are possible and which are not, based on attention to what sequences do and do not occur in the ambient language (Estes et al., 2011; Mattys et al., 1999). Along similar lines, 8-month-old infants were exposed to sequences of discrete visual stimuli whose ordering followed a statistically predictable pattern. Subsequently they were again shown the familiar pattern but alternating with a novel sequence of identical stimulus components. Researchers found that infants exhibited significantly greater interest in the novel sequence at all ages, indicating that they discriminate between known and unknown visual patterns (Kirkham et al., 2002). Similar discrimination and selection abilities were found by pre-phonological children in different sorting tasks.

Children apply readability criteria for distinguishing possible from impossible combinations of letters in their language. The mere presence of letters is not enough for something to be readable. Over a period of several decades, researchers applied similar tasks using different instruction to children acquiring different languages and orthographies (English, Hebrew, Italian and Spanish). Pre-readers (aged 2–5 years) were shown cards containing graphic displays differing by several variables: iconicity, familiarity, number, and variety of graphic elements. Whether children were asked to separate “writing from non-writing” (Lavine, 1977) or to say “what is good for reading from what is not” (Ferreiro and Teberosky, 1979/1982) the conclusions were similar. Cards displaying drawings were the earliest to be rejected as words or as possible to read. Then, letters from unknown alphabets, mixed alphabets, and cards containing only one sign or more than 10 in a row were increasingly rejected as children grew older, but before they were able to read. Kaefer (2016) had 3–5-year-olds to distinguish words from non-words when faced with strings of letters, non-letters, and numbers. Three-year-olds do not distinguish between them overtly, but they do look significantly longer at a number or a symbol. In contrast, 5-year-olds identified strings of letters as “words” and strings of non-letters or repeated letters as “non-words.”

Pre-readers were also able to distinguish between sequences of letters that are possible combinations in their language and those that are not. Children rejected strings containing repeated vowels or repeated consonants demonstrating their understanding that a word must integrate both types of letter (Pick et al., 1978). In order to decide which combinations are “well written,” they look both at the letters that are integrated within the string and at the general configuration of the string (Ferreiro, 1982). The ideal written word—the graphic display that was better accepted as good for reading—contained more than 2 and less than 6 letters from own alphabet.

The constraints that children impose on number and variety of signs in order for a graphic display to be readable are not mere inventions: They attend to the actual distribution of word length and intra-word variation found in real texts of the different orthographies that were tested. Neither Spanish nor English has written words containing the same letter repeated in a row more than twice consecutively (e.g., acceso “entrance”). In line with children's constraints, the rules of English spelling outlaw triple letters while in Hebrew this constraint applies to consecutive strings of three identical consonants.

Notably, the constraints on number and variety are applied only to writing and not to numbers. Together with Annette Karmiloff-Smith we asked Spanish speaking children to separate between cards they think “are good for reading” from those that “are good for counting” (after piloting we have found that this expression is the most frequently used by Spanish small children to design

numbers). Cards containing one graphic sign or several repeated in a row were rejected as good for reading but accepted as good for counting (Tolchinsky Landsmann and Karmiloff Smith, 1992).

These intuitive criteria of a word's readability are evident in children's sorting as well as in writing tasks. This is important, since if a text is readable, it makes sense for a person (in this case a young child) to wonder what it means, whereas if it is not readable, there would be no point in pondering its meaning.

Early writing patterns

In a different study Annette Karmiloff Smith asked preliterate and "predrawing" children that neither know how to read nor how to draw in a conventional manner—to draw a dog and write its name (Karmiloff-Smith, 1992). Children felt rather bewildered when asked to write since they clearly believed that they were not supposed to know. As suggested by Brenneman et al. (1996), this refusal might reflect a child's conviction that writing, as opposed to drawing, requires special marks that must be known. Once convinced to try, their drawing typically appeared as circular scribbles while their writing as wiggly horizontal lines. Yet, even when their drawings were hard to distinguish from their writings, close observation showed differences in the production process. When pretending to write toddlers, lift the pen much more frequently than when pretending to draw. Drawing appears as a much more continuous movement than writing. By 4 years of age, children have learned about the formal properties that differentiate drawing from writing, although it may take them longer to learn how each system is used to refer and communicate. These findings were confirmed by Brenneman et al. (1996) with 4–6-year-old US children who were asked to draw and write various targets. Children were more likely to make linear productions while writing than while drawing. Gombert and Fayol (1992) reported similar findings with 3–6-year-old French children. They were more likely to make productions containing linear, segmented units created in the left to right direction when writing than when drawing.

More recently Treiman and associates confirmed similar distinctions among even younger Chinese (Treiman and Yin, 2011) and US children (Otake et al., 2017). Both Chinese and US 2–5-year-olds had to write and draw the words for fire, sun, water, and their name and to draw these same targets and a picture of themselves in the drawing task. In both tasks, children were given a choice of five implements (black pencil, black pen, red crayon, yellow crayon, blue crayon), and four types of paper (grid paper and unlined white, yellow, and pink paper). Even the 2-year-olds in these studies made some distinctions between writing and drawing. Children in the youngest group (mean age 2;8) created smaller writings than drawings and their writing tended to be more rectangular than their drawings. When writing, for example, participants were more likely to use dark implements and to choose grid paper, which Chinese schoolchildren use to practice writing characters. Moreover, adults who were asked to classify the productions as writing or drawing performed significantly above the level of chance with the productions of 2-year-olds.

From a very early age, children are sensitive to some of the formal properties of different notational systems. It would appear, moreover, that this sensitivity is not limited to children from a specific culture and/or language.

There are vivid documentations of the transition from unbroken, wavy scribble produced by 2-year-olds to the tracing of separate marks (Schickedanz and Casbergue, 2004; Vukelich and Christie, 2009). As children grow older their writing will increasingly display the superordinate/formal features common to all writing systems: They are linear, go in a certain direction, have discrete units, and are not iconic. Children tend to arrange their scribbles horizontally, revealing a growing correspondence between their attempts to write and the directionality of their orthography (Brenneman et al., 1996; Pontecorvo, 1985; Sulzby, 1986). Puranik and Lonigan (2011) provide support for such a developmental trend in a study with over 300 English-speaking children aged 3–5 years. All 3-year-olds in the sample had acquired most of the superordinate writing features, while specific-language features could only be found in the writing samples of 4- and 5-year olds.

In examining word-writing among Spanish- and Hebrew-speaking children from preschool through second grade (Tolchinsky and Teberosky, 1998), we found that the written products of preschoolers in both languages were constrained by the same formal features of number and variety: Every string contained a similar amount of letters (between 3 and 5), and adjacent letters were not repeated. These initial similarities subsequently diverged increasingly between Spanish and Israeli children, as the distinctive features of each language began to have an effect (in this case, the primary role of consonants in a Semitic script).

These studies in different orthographies confirm the salience of the visual/formal patterns of writing, which children extract from the print they encounter in their environment rather than due to a direct instructional strategy. Adults might teach letters, how to write proper names and real words, but not to scribble in ways that look like writing but are not.

Multiple studies in different languages, showed that, for representing their own names, children use more advanced writing forms than when writing other words. The same children might use letters or simile letters for their own name and produce unspecific, unbroken, wavy scribbles for any other word. A recent study showed that among all the tested predictors, including gender, age, and number of years in the program, only name-writing ability (assessed at school entry) predicted early writing scores after 6 months (Campbell et al., 2019).

During the period in which children's writing display the superordinate/formal features common to all writing systems they produce very similar patterns for different words and even different sentences. They did not introduce graphic indications of differences in meaning. One crucial characteristic of written displays is that they can be interpreted outside the context of production—separated from the person who has produced them. But if no graphic distinction is made, only someone who has produced or been asked to produce them is able to interpret it. At this stage of writing development, the product has not yet been detached from the process of production and from the producer. Children behave as though the place where writing stands (e.g., under a certain picture) or the writer him or herself, rather than any particular feature of the written display, will determine the

interpretation (Tolchinsky, 2003). A study by Yaden and Tardibuono (2004) with Latino preschoolers illustrates clearly that a child's first attempts to differentiate writing outcome is in terms of his or her intents. In spite of producing very similar outputs for different words—the same arrangement of the letters of his or her own name—s/he is certain that they are different. After this period of “undifferentiated” writing (Ferreiro, 1982), children gradually become more selective as to what forms or combinations of forms are accepted as writable. As a consequence, their own written production becomes more constrained and children introduce graphic indicators to reflect differences in the words they are attempting to write.

The study of Yaden and Tardibuono (2004) again provides a clear illustration of this important development. They found that a child named Alfredo, in spite of using a limited number of simile letters (forms that resemble letters but are not) in different orientations, varied these symbols in order to distinguish words (nene “child” and oso “bear”) and the names of family members. Researchers report this type of “recombinatory strategy” (the same symbols in different combinations) as being used by several children at this stage of writing.

These findings suggest that children are aware that differentiating between written outcomes must take place when creating the message rather than when interpreting it, since only then can the written outcome be understood as an entity in itself, separated from the producer and the context of production. When asked to read what they have just written, children usually repeat verbatim the words they have been asked to write, pointing globally at the entire piece of graphic material or at an unspecified part of it. The question arises, however, what meaning children attribute to the visible/graphic changes they observe and produce in writing.

Meaning attribution and written language

Early attribution of meaning

We have seen how small children are able to discriminate between pictorial and written displays, initially attending to their distinctive formal constraints and later on to some of their orthography-specific features. The early graphic patterns children produce and identify as writing are not yet symbolic patterns. They are action-schemas in the graphic space, but they are not interpretable by other, they do not contain cues to what they stand for. Even when children introduce graphic changes in their constrained productions in order to write different words and to map a verbal utterance (a single word or multiword expression) onto what they have written, when they pretend to read the product, it is clear that they have not yet disclosed how writing means.

Nearly a 100 years ago, in 1929, Luria addressed this issue by asking 3–7-year old children to write a list of expressions in a way “they could remember.” The productions of the youngest did not contain any graphic cue to the phrases the children had to remember. But at some point, some of the older children began to introduce certain graphic elements as clues to the meaning of the written phrases. For example, they use longer lines for “The monkey has a long tail” than for writing “It's cold in winter.” Luria called this “pictographic” use of written sign because they try to represent the content of the utterance rather than its sounds.

Ferreiro (1985) reported similar findings. Children explained, for example, that oso “bear” should have more letters than pato “duck” because a bear is bigger than a duck. She remarked, however, that differences in shape or color of the referents were not reflected in children's writing. Later on, Iris Levin and I (Levin and Tolchinsky Landsmann, 1989) had Hebrew-speaking preschoolers to write pairs of words that represented objects differentiated by shape (e.g., ball/rope) color (e.g., tomato/cucumber) and size (e.g., elephant/ant). We controlled for word length so that the longer sounding word refer to the smaller object (pil “elephant”/nemala “ant”) while the words referring to shape were of similar length. We found that the use of shape and color was very rare and tended to decrease with age while the use of more symbols for bigger objects tended to increase. Prereading children favored the use of number of symbols over any other referential resource.

Children's attempts to establish a direct relation between writing and physical features of the referent was also found in their judgments of written displays. Bialystok (1991) showed English-speaking children two pictures accompanied by two printed words, which the investigator read out loud, and then had the children match each picture with the word they thought they had just heard. The test items manifested two different relations between pictures and words, one of correspondence between the length of the word in number of letters and the size of the object in the picture (e.g., ball and ballerina) and the other one of non-correspondence between them (e.g., cat and caterpillar). Non-readers gave correct responses more often for corresponding than noncorresponding pairs, suggesting that they used the relative sizes of objects to make decisions about the number of letters in printed words. A previous study by Lundberg and Torneus (1978) drew similar conclusions for Swedish-speaking children, who were presented word pairs both orally and visually and asked to indicate which printed word went with each spoken word. Four-to seven-year-old non-readers gave more correct responses when the relative lengths of the words were consistent with the relative sizes of the referents than when the two were not consistent.

In order to further prove whether children's initial attribution of meaning is guided by the referential content of the utterances they are attempting to write rather than by phonological considerations, Zhang and Treiman (2015) presented children with words said out loud one at a time without accompanying pictures, so avoiding additional cues that might make semantic information particularly salient. These researchers applied a rigorous procedure to identify pre-phonological spellers' insights. They asked children to write words that varied in number of phonemes (from 3 to 9) and size of the object (e.g., ant, asparagus, bear, beetle, bug, bus). For each writing production, they measured the area of the smallest rectangle enclosing the entire production (in cm²) and found that object area had a significant effect on area of writing production. Pre-phonological spellers took up larger areas for words representing large objects than for small objects while producing writing of similar size for words with different numbers of

phonemes or syllables. They applied similar analyses with number of written elements as the outcome variable and found that neither number of phonemes nor number of syllables significantly predicted number of written elements.

Zhang and Treiman study added convincing evidence regarding the existence of a phase in developing literacy during which children do not conceive writing as related to speech, whether segmented by phonemes or syllables. Children can produce graphic displays that look like writing and are recognized by external observers as such because they obey formal constraints of writing, in many cases containing letters, however they relate the size and number of elements in writing to the size of the entities the words refer to.

These consistent findings raise a number of queries. Does this behavior indicate that children are resorting to a drawing resource in their writing? Are they confusing the way drawing and writing represent meaning, so contradicting earlier claims about differentiation between the two notational systems? Or is this, rather, an indication of the representational status writing has at this point in development? In other words, does this notational behavior indicate what writing represents for children at this developmental period? What do they think needs to be noted? How should writing resources be used to note what they think should be noted?

Several studies (Ferreiro, 1986b; Olson, 1996; Olson and Homer, 1996) have addressed these questions creatively. They designed writing tasks in which, if children assume that writing represents meaning directly, there would be a contradiction between writing (leaving marks on paper) and the meaning of the utterance they are asked to write. If, on the other hand, children relate writing to the linguistic expression that it manifests, no such contradiction should exist.

This was examined in different ways. For example, negation is amply represented in oral language and, children produce and understand negative statements from an early age, yet studies show that the written representation of negative statements constitute a genuine difficulty for children. Ferreiro (1986b) conducted individual interviews with a group of pre-reading Spanish-speaking children, where the investigator wrote, in front of the child, sentence such as: *Mama compró tres tacos* "Mommy bought three tacos," then read it out loud and asked the child to repeat what was written. Immediately afterward the child had to write the negative version of what s/he had heard, *Mama no compró tacos* "Mommy didn't buy tacos" with the written sentence still visible in front of him/her. Ferreiro reported that one-third (34%) of the children alluded to the impossibility of writing the negative sentences. She explains this as due to children's being unable to overcome the apparent contradiction between marking something on paper and the absence connoted by negation. Following the same methodology, Ferreiro reported that 53% of the same children refused to write the sentence *Ningún pájaro vuela* "No bird is flying" after the sentence *Un pájaro vuela* "A bird is flying." In this case it was the due to the absence connoted by the negative quantifier.

Olson and Homer (1996) and Olson (1996) tackled similar issues with English-speaking prereaders and illiterate adults. Participants were presented with simple written texts which were then read aloud to them. For example, they were shown a card on which was printed *Three little pigs*, and the investigator read the text out loud, pointing to each word in turn. The child then repeated the text. Following this, the final word in the written sentence was covered, so that only some of the words remained exposed (*Three little*) and children were asked what they thought was written there. A common response was to say: *two pigs*. When a further word was covered, children tended to read it as *one pig*. Olson and Homer concluded that children take the marks as tokens for things rather than as signs for words.

In a second study, they asked pre-reading children to write such statements as *One cat*, *Two cats*, *Red cat*, and the like. Children showed a similar pattern of responses: They tried to make their writing (scribbles or simile letters) correspond to the number of cats that were mentioned. Finally, when the child was asked to write *No cats*, the researcher reported that the child glanced at the adult and waved her pencil above the paper without leaving a mark, and without prompting said: "There's no cats so I didn't write anything."

Although these studies were designed to test prereaders' assumption of a direct link between the written marks and "the things" mentioned in the utterances, there are other possible interpretations for these responses that might shed light on the processes involved. This behavior can be interpreted as children's search for semantic congruence. In other words, children may be attempting to relate to the meaning of the expression/utterance rather than to the things it refers to; that is, they do associate writing with language, but not to how it sounds, its phonology, but rather to its meaning (its semantic content). This alternative interpretation highlights a possible lack of the required metalinguistic ability on the part of pre-readers to separate content from form. What are the implications of the two interpretations of pre-reading?

In the cases just described (e.g., mapping between number or size of written elements and size of the referred entity, refusing to leave marks on paper because there are no entities in the statement they are supposed to write) children are creating a visual similarity between the symbol and the symbolized. Since 1932, when John Ridley Stroop discovered that it is easier to say the color of a word if it matches the semantic content of the word (saying the word "green" if it is written in green ink rather than blue or red), the "interaction between objects as referents and their symbols (words)", is recognized as a well-known famous (and intriguing) effect of perception on conceptualization (for a review see Shen et al., 2016).

In literate adults, visual congruence between symbols and what they symbolize facilitates processing, words are recognized and interpreted faster. Congruence might also function as a feasible strategy for pre-reading children to match successfully each word with its picture (as shown in different studies by Bialystok, by Lundberg, and Torneus) studies, and may serve help them to cope with the writing task—especially when the application of such strategy is facilitated by the way the task is presented. To test the effect of (in)congruence, stimuli are presented as pairs of words that contrast on a specific dimension of content (size, color, quantity) and allowed to choose between different colored crayons so they can take advantage of an obvious resource for writing (change in number or size of graphic elements). It is more difficult, however, to use this strategy to explain children's refusal to produce marks for representing absence or negation, as in the studies of Olson and Holmes.

An alternative account views these behaviors as a consequence of a metalinguistic deficit in young children, which impedes their ability to dissociate the referent from features of the utterance. Although preschool children use negation pragmatically to mark agreements and disagreements with speakers' assertions, they do not conceive of negation metalinguistically to make judgments about statements (Hummer et al., 1993). Metalinguistic awareness is not a dichotomous ability that children either do or do not have. Rather, it follows a protracted progression that starts very early in little children's self-repairs when speaking and grows into the ability to think about and discuss different features of the language people know or are in the process of learning. This lack of a highly evolved metalinguistic capacity is thought to be responsible for children's difficulty in representing negation in writing. Such a metalinguistic deficit is reflected, for example, in young children's difficulty in understanding synonymy and homonymy, as tasks that require to distinguish between the linguistic expression and what it represents Doherty (2000).

Against an interpretation in terms of metalinguistic deficit, Olson argues that representing negation is equally difficult for pre-reading children and non-literate adults, thus, it cannot be seen as a developmental phenomenon which will be overcome with age. The fact that this behavior is overcome after learning to write supports Olson's view that writing brings the structures of language into consciousness and not the other way around. Negation is a linguistic device which speakers use competently. But without a representation of negation in writing, negation is not readily brought into consciousness and turned into an object of thought. As explained by Olson (1997), when writing offers a visual model for the representation of language, the negative particle comes to be represented by a distinctive sign for that negation, as in the written word *not*. Similarly, for the representation of absence, "nothing becomes a something which can then be examined, thought about, and operated on." (p. 14). Put it more simply, pre-readers and illiterate adults have a struggle with the representation of negation and absence because they have not yet learned to read and write, because they have not yet acquired a system for representing language.

Conclusion

Writing and language

Young toddlers before age two may emit sounds that mimic reading when they look at a story-book; they imitate adult's behavior in front of a text. This "simile-reading," applied to their own scribbling or to other printed objects, is crucially important because it links a verbal behavior to a written object. Adult cues may also guide children in linking writing to language. Parents of young children tend to talk about speaking and writing in quite similar ways.² For example, a parent may use the word *say* to refer to both speech and writing (e.g., What did you say? What does that sign say?). The words *name* and *word* can also be used interchangeably to refer to units of the written or spoken language (e.g., I just wrote my name or that's not a nice word to use). According to Robins and Treiman (2010), the use of similar terms to refer to both speech and writing may help children to realize that the two systems symbolize language.

The phoneticization of writing constitutes a genuine turning point in developing literacy. It implies that what is said can be written, that different words reflect more than referential differences such as size, amount or color. True, this assertion is an oversimplification: We cannot represent in (alphabetic) writing differences between, say, a Texan and New York accent, or the length of the pauses that people produce when speaking, one of the reasons why children find it difficult to separate between words in writing. Writing is not a code for transcribing speech but a system of representation of language or, as advanced by Olson (1996), a model of language in which certain features are represented and others are not.

Detailed description of the process of phoneticization or how children acquire the alphabetic principle that each graphic element stands for a phonological segment is beyond the scope of the present chapter. This process is well-documented in the literature on developing literacy (Byrne, 1996; Tolchinsky, 2003). Rather, I wish to underscore, in line with Olson, that metalinguistic awareness is significantly extended as a consequence of literacy learning. The blossoming of language awareness with literacy contrasts markedly with the widely accepted view of language awareness as a prerequisite for acquiring literacy.

On the nature of literacy development

Text-books and encyclopedias caution that reading and writing are achieved only by explicit instruction (Malmkjær (1993/2010). Yet the picture we have depicted so far of the abilities of prereading children point to several developments that do not result from direct teaching. Children build action schema afforded by literacy-related artifacts and implements which become richer and more diversified with the availability of new artifacts and cooperating adults. Children extract relevant information from the notational environment, they are sensitive to its regularities and rely on this sensitivity to develop incipient orthographic knowledge. Pre-readers are also able to take advantage of cues provided by written forms—number and size of elements—to reflect their own ideas on the representational status of writing, regardless of whether they turn out to be mistaken. The quality of exposure to and interaction with potential teachers are conditions for early literacy acquisition while formal instruction is fundamental for developing proficient writers and readers, but this is equally the case for every domain of knowledge. The quality of input and interaction

²They conducted a series of analyses using CHILDES, a database containing child language transcripts from numerous different researchers. They included all the transcripts of parent-child conversations in English, mostly in the U.S., with children between 18 months and 5 years of age. Although the transcripts were collected by different researchers for a variety of purposes, they found that parents speak about writing with their children even when they are quite young.

conditions early language development whereas formal instruction is necessary for developing proficient and reflective users of language, although some gifted individuals may get to it by intuition if not by instruction (Tolchinsky and Berman, *in press*).

We have been told time and again that reading and writing abilities, unlike speaking or breathing, are not natural (Emig, 1977; Stanovich, 1994), that the human brain is naturally wired to speak, whereas our brain needs to follow a systematic instructional process in learning to read and write. To the extent that natural means “genetically specified,” I would argue that developing literacy is as natural as speaking: both are enabled by the basic human capacity for learning. The ability to learn language(s) and writing (and many other things) is wired into the human brain (Dehaene, 2019). “The great majority of the brain’s development happens after we are born. It matures as much as possible in the womb, but this still leaves the lion’s share of developmental work to be done after birth. The upside to this state of affairs is that our brains are finished being built while they are immersed in a particular culture, allowing our brains to be fine-tuned to operate in that specific environment” (Lieberman, 2013, p. 41). If we take these assertions seriously, as more and more evidence from neurobiology suggests we should, we have no choice but to recognize and embrace the enormous importance of early socio-cultural experiences from the environment for early as well as later learning to read and write.

References

- Bialystok, E., 1991. Letters, sounds, and symbols: changes in children’s understanding of written language. *Appl. Psycholinguist.* 12, 75–89. <https://doi.org/10.1017/S0142716400009383>.
- Brenneman, K., Massey, C., Machado, S.F., Gelman, R., 1996. Young children’s plans differ for writing and drawing. *Cognit. Dev.* 11, 397–419. [https://doi.org/10.1016/S0885-2014\(96\)90011-8](https://doi.org/10.1016/S0885-2014(96)90011-8).
- Bruner, J.S., 1973. Organization of early skilled action. *Child Dev.* 44, 1–11.
- Bus, A.G., van Ijzendoorn, M.H., Pellegrini, A.D., 1995. Joint book reading makes for success in learning to read: a meta-analysis on intergenerational transmission of literacy. *Rev. Educ. Res.* 65 (1), 1–21. <https://doi.org/10.3102/00346543065001001>.
- Byrne, B., 1996. The learnability of the alphabetic principle: children’s initial hypotheses about how print represents spoken language. *Appl. Psycholinguist.* 17, 401–426. <https://doi.org/10.1017/s0142716400008171>.
- Campbell, K., Yi-Jui Chen, Y.-J., Shenoy, Cunningham, A., 2019. Preschool children’s early writing: repeated measures reveal growing but variable trajectories. *Read. Writ.* 32 (4), 939–961. <https://doi.org/10.1007/s11145-018-9893-y>.
- Clay, M., 2001. *Change Over Time in Children’s Literacy Development*. Heinemann.
- Coulmas, F., 1989. *The Writing Systems of the World*. Blackwell.
- Dawson-Tunik, T.L., Commons, M., Wilson, M., Fischer, K.W., 2005. The shape of development. *Eur. J. Dev. Psychol.* 2, 163–195.
- Dehaene, S., 2009. *Reading in the Brain*. Penguin Viking.
- Dehaene, S., 2019. *How We Learn: Why Brains Learn Better Than Any Machine ... for Now*. Penguin Random House.
- Doherty, M., 2000. Children’s understanding of homonymy: metalinguistic awareness and false belief. *J. Child Lang.* 27 (2), 367–392. <https://doi.org/10.1017/S0305000900004153>.
- Dyson, A.H., 1987. Individual differences in beginning composing: an orchestral vision of learning to compose. *Writ. Commun.* 4 (4), 411–442.
- Enri, L.C., 1995. Phases of development in learning to read words by sight. *J. Res. Read.* 18 (2), 116–125.
- Emig, J., 1977. Writing as a mode of learning. *Coll. Compos. Commun.* 28 (2), 122–128. <https://doi.org/10.2307/356095>.
- Estes, K.G., Jan Edwards, J., Saffran, J.R., 2011. Phonotactic constraints on infant word learning. *Infancy* 16 (2), 180–197. <https://doi.org/10.1111/j.1532-7078.2010.00046.x>.
- Ferreiro, E., Teberosky, A., 1979. Los sistemas de escritura en el desarrollo del niño. *Siglo XXI*. In: *Literacy before schooling* (Ed.). Heinemann, 1982.
- Ferreiro, 1982. Los procesos constructivos de apropiación de la escritura. In: Ferreiro, Gómez Palacio (Eds.), *Nuevas perspectivas sobre los procesos de lectura y escritura*. Siglo xxi, pp. 128–154.
- Ferreiro, 1985. Literacy development: A psychogenetic perspective. In: Olson, Torrance, Hildyard (Eds.), *Literacy, language and learning*. Cambridge University Press, pp. 217–228.
- Ferreiro, E., 1986a. The interplay between information and assimilation in beginning literacy. In: Teale, W., Sulzby, E. (Eds.), *Emergent Literacy*. Ablex, pp. 15–49.
- Ferreiro, E., 1986b. *Proceso de Alfabetización*. Bibliotecas Universitarias.
- Ferreiro, E., 1994. Literacy development: construction and reconstruction. In: Tirosh, D. (Ed.), *Implicit and Explicit Knowledge: An Educational Approach*. Ablex, pp. 169–180.
- Gentry, J., 2000. A retrospective on invented spelling and a look forward. *Read. Teach.* 54 (3), 318–332.
- Gibson, E., Levin, H., 1975. *The Psychology of Reading*. MIT Press, Cambridge, Mass.
- Gibson, E., Yonas, A., 1966. A developmental study of visual search behavior. *Percept. Psychophys.* 1, 169–171. <https://doi.org/10.3758/BF03215778>.
- Gibson, E., 1988. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annu. Rev. Psychol.* 39, 1–41.
- Gombert, J.E., Fayol, M., 1992. Writing in preliterate children. *Learn. Instruct.* 2 (1), 23–41. <https://doi.org/10.1016/0959-4752>.
- Goodman, Y., 1980. The roots of literacy. In: Douglass, M.P. (Ed.), 44th Yearbook of the Claremont Reading Conference. The Claremont Colleges, Claremont, CA, pp. 1–32.
- Hummer, P., Wimmer, H., Antes, G., 1993. On the origins of denial negation. *J. Child Lang.* 20 (3), 607–618. <https://doi.org/10.1017/s0305000900008503>.
- Kabali, H.K., Irigoyen, M.M., Nunez-Davis, R., Budacki, J.G., Mohanty, S.H., Leister, K.P., Bonner Jr., R.L., 2005. Exposure and use of mobile media devices by young children. *Pediatrics* 136 (6), 1044–1050. <https://doi.org/10.1542/peds.2015-2151>.
- Kaefer, T., 2016. Integrating orthographic and phonological knowledge in early readers: implicit and explicit knowledge. *Child Dev. Res.* 1–7. <https://doi.org/10.1155/2016/6036129>.
- Karmiloff Smith, A., 1992. *Beyond Modularity: A Developmental Perspective on Cognitive Science*. MIT Press/Bradford.
- Kirkham, N.Z., Slemmer, J.A., Johnson, S.P., 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83 (2), 35–42. <https://doi.org/10.1016/S0010>.
- Lareau, A., 2000. Social class and the daily lives of children: a study from the United States. *Childhood* 7 (2), 155–171.
- Lavine, L.O., 1977. Differentiation of letterlike forms in prereading children. *Dev. Psychol.* 13 (2), 89–94. <https://doi.org/10.1037/0012-1649.13.2.89>.
- Levin, I., Tolchinsky Landsmann, L., 1989. Becoming literate: referential and phonetic strategies in early reading and writing. *IJBD* 12, 369–384. <https://doi.org/10.1177/016502548901200306>.
- Lieberman, M., 2013. *Why Our Brains are Wired to Connect*. Oxford University Press.
- Lundberg, I., Torneus, M., 1978. Nonreaders’ awareness of the basic relationship between spoken and written words. *J. Exp. Child Psychol.* 25, 404–412. [https://doi.org/10.1016/0022-0965\(78\)90064-4](https://doi.org/10.1016/0022-0965(78)90064-4).
- Luria, A.R., 1929/1978. The development of writing in the child. In: Cole, M. (Ed.), *The Selected Writings of A. R. Luria*. White Plains, pp. 145–194.
- Malmkjær, K., 1993/2010. *The Routledge Linguistics Encyclopedia*, third ed.

- Mattys, S.L., Jusczyk, P.W., Luce, P.A., Morgan, J.L., 1999. Phonotactic and prosodic effects on word segmentation in infants. *Cognit. Psychol.* 38 (4), 465–494. <https://doi.org/10.1006/cogp.1999.0721>.
- Olson, D.R., Homer, B., 1996. Children's concept of "word". In: Paper Presented at the XXVI International Congress of Psychology, Montreal: August 16–21.
- Olson, D.R., 1996. Towards a theory of literacy: on the relations between speech and writing. *Cognition* 60 (1), 83–104.
- Olson, D., 1997. The written representation of negation. *Pragmat. Cognit.* 5 (2), 235–252. <https://doi.org/10.1075/pc.5.2.03ols>.
- Otake, S., Treiman, R., Yin, L., 2017. Differentiation of writing and drawing by U.S. two- to five-year-olds. *Cognit. Dev.* 43, 119–128. <https://doi.org/10.1016/j.cogdev.2017.03.004>.
- Pick, A.D., Unze, M.G., Brownell, C.A., Drozdal, J.G., Hopmann, M.R., 1978. Young children's knowledge of word structure. *Child Dev.* 49 (3), 669–680. <https://doi.org/10.2307/1128234>.
- Pontecorvo, C., 1985. Figure, parole, numeri: un problema di simbolizzazione [Pictures, words, and numbers: a problem of symbolization]. *Eta Evol.* 22, 5–33.
- Pulvermuller, F., 2013. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cognit. Sci.* 17. <https://doi.org/10.1016/j.tics.2013.06.004>.
- Puranik, C.S., Lonigan, C.J., 2011. From scribbles to scabble: preschool children's developing knowledge of written language. *Read. Writ.* 24 (5), 567–589. <https://doi.org/10.1007/s11145-009-9220-8>.
- Robins, S., Treiman, R., 2010. Learning about writing begins informally. In: Aram, D., Korat, O. (Eds.), *Literacy Development and Enhancement Across Orthographies and Cultures*, Literacy Studies, vol. 101. Springer, pp. 17–29. https://doi.org/10.1007/978-1-4419-0834-6_2.
- Schickedanz, J., Casbergue, R., 2004. Writing in Preschool: Learning to Orchestrate Meaning and Marks. International Reading Association.
- Schmandt-Besserat, D., 1998. Accounting before writing in the ancient near east. *Porocilo* 24, 151–156.
- Shaywitz, S.E., Shaywitz, B.A., 2008. Paying attention to reading: the neurobiology of reading and dyslexia. *Dev. Psychopathol.* 20, 1329–1349. <https://doi.org/10.1017/S0954579408000631>.
- Shen, M., Xie, J., Liu, W., Lin, W., Chen, Z., Marmolejo-Ramos, F., Wang, R., 2016. Interplay between the object and its symbol: the size-congruency effect. *Adv. Cognit. Psychol.* 12 (2), 115–129. <https://doi.org/10.5709/acp-0191-7>.
- Snow, Ninio, 1986. In: Teale, W.H., Sulzby, E. (Eds.), *Emergent literacy: Writing and reading*. Ablex, pp. 116–138.
- Stanovich, K.E., 1994. Constructivism in reading education. *J. Spec. Educ.* 28 (3), 259–274. <https://doi.org/10.1177/002246699402800303>.
- Sulzby, E., 1986. Writing and reading: signs of oral and written language organization in the young child. In: Teale, W.H., Sulzby, E. (Eds.), *Emergent Literacy: Writing and Reading*. Ablex, Norwood, NJ, pp. 50–89.
- Tolchinsky, L., Berman, R., in press. *Growing into language*. Oxford University Press.
- Tolchinsky, L., Teberosky, A., 1998. The development of word segmentation and writing in two scripts. *Cognit. Dev.* 13 (1), 1–24. [https://doi.org/10.1016/S0885-2014\(98\)90018-1](https://doi.org/10.1016/S0885-2014(98)90018-1).
- Tolchinsky Landsmann, L., Karmiloff-Smith, A., 1992. Children's understanding of notations as domains of knowledge versus referential-communicative tools. *Cognit. Dev.* 7 (3), 287–300. [https://doi.org/10.1016/0885-2014\(92\)90017](https://doi.org/10.1016/0885-2014(92)90017).
- Tolchinsky Landsmann, L., Levin, I., 1987. Writing in four—to six-year-olds: representation of semantic and phonetic similarities and differences. *J. Child Lang.* 14, 127–144.
- Tolchinsky, L., 2003. The Cradle of Culture and What Children Know About Writing and Numbers Before Being Taught. Lawrence Erlbaum.
- Treiman, R., Yin, L., 2011. Early differentiation between drawing and writing in Chinese children. *J. Exp. Child Psychol.* 108, 786–801. <https://doi.org/10.1016/j.jecp.2010.08.013>.
- Vukelich, C., Christie, J.F., 2009. Building a Foundation for Preschool Literacy: Effective Instruction for Children's Reading and Writing Development. International Reading Association.
- Wellsby, M., Pexman, P.M., 2014. Developing embodied cognition: insights from children's concepts and language processing. *Front. Psychol.* 5. <https://doi.org/10.3389/fpsyg.2014.00506>. Article 506.
- Winch, G., Johnston, R., March, P., Jungdahl, L., Holliday, M., 2010. *Literacy: Reading, Writing and Children's Literature*. Oxford University Press.
- Yaden, D.B., Tardibuo, J.M., 2004. The emergent writing development of urban latino preschoolers: developmental perspectives and instructional environments for second-language learners. *Read. Writ. Q.* 20, 29–61. <https://doi.org/10.1080/10573560490242723>.
- Zhang, L., Treiman, R., 2015. Writing *dinosaur* large and *mosquito* small: prephonological spellers' use of semantic information. *Sci. Stud. Read.* 19 (6), 434–445. <https://doi.org/10.1080/10888438.2015.1072820>.

Biliteracy, playfulness and performative practices

Pengtong Qu^a, Amanda Deliman^b, Gustave J. Weltsek^c, Amy Walker^a, and Carmen Liliana Medina^a, ^a Department of Literacy, Culture, and Language Education, Indiana University, Bloomington, IN, United States; ^b School of Teacher Education and Leadership, Utah State University, Salt Lake City, UT, United States; and ^c Department of Curriculum and Instruction, Indiana University, Bloomington, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	48
Creativity in and across languages	48
Creativity in everyday biliteracy practices	49
Identity development & translanguaging	49
Language brokering	49
Translanguaging as human capital building	50
Drama and play pedagogies as performative spaces in biliteracy practices	50
Personal, linguistic, cultural and social resources accessed through drama and play	51
Conclusion	51
References	52

Introduction

In this chapter, we regard biliteracy as a creative strength and as an integrated linguistic repertoire that bilingual speakers own instead of two separated language systems (Hornberger, 1989; Reyes, 2012). Examining biliteracy through a creative performative lens (Medina and Campano, 2006; Medina et al., 2014)—where language, identity and culture are understood as dynamically interacting and co-constructed—we first discuss the emerging, inventive, and creative features of biliteracy to foreground the skills emergent bilingual children create and engaged within multiple social domains. Second, we focus on research that highlights bilingual and biliterate performative practices in a multicultural and pluralistic world. Specifically, we dive into depth of how family, cultural, and linguistic playfulness are enacted in everyday practices. Next, we review biliteracy pedagogical performative practices in drama and play with the aim to discuss new goals and challenges of promoting biliteracy from a creative and playful space in schools that honors emergence, creativity, linguistic hybridity and plurality as productive spaces for learning.

Creativity in and across languages

In countries such as the United States, from where the culturally diverse—China, USA, Puerto Rico—multilingual group of authors of this chapter do most of our work, monolingualism ideology, which regards two languages that bilingual speakers own as two separate systems, has been accepted as a dominant view in language and literacy programs in the K-12 settings (Cummings, 2000; Wiley, 2014). Within such programs, it is commonly seen that one language is legitimated according to the time, space, policies, and instruction (Farr and Song, 2011; Ruiz, 1984). The claim of many of these programs is to “help” emergent bilingual speakers, who oftentimes are labeled as English Language Learners (for a critique of this term, see Martínez, 2018), develop their English proficiency and learn school-like literacy (Morrow, 1995) or academic language (García and Solorza, 2020) so that they are able to catch up with the English monolingual peers in terms of achievement in content areas (Hornberger, 1989). These forms of language education approaches, grounded in orientations of language as a problem (Ruiz, 1984) however, might result in the suppression of bilingual students’ bilinguality in school (Cummings, 2000; Li and Wu, 2009) and the construction of material and ideological barriers through segregation and erasure approaches with consequences in relation to students’ rights to “exist in their own language” (Perryman-Clark et al., 2014).

Moll (2014) perceived biliteracy as a “tool of the intellect” (p. 84), which is used by bilingual speakers for facilitating their dynamic discourses and practices. The visible differences (e.g., English and Spanish) are merely the surface feature of languages (Cummins, 1980; Reyes, 2012); beneath the surface, two languages do not work separately. Bilingual speakers think in, between, and beyond two languages when practicing in the complex, multi-layered, and interrelated multilingual communication (García and Sylvan, 2011; Li, 2011; Perez and Torres-Guzmán, 1992; Reyes, 2012). They creatively use languages in order to express and negotiate complex meanings (Franquiz and de la Luz Reyes, 1998; Gutiérrez et al., 1999; Medina, 2006; Medina 2010). Instead of perceiving bilingual students as deficient English speakers, educators need to consider the development of ways of doing with biliteracy in classrooms that honors the creativity, fluidity, and complexity of linguistic plurality as assets to support students learning (Canagarajah, 2011; García and Lin, 2017; Guerra, 2016; Perryman-Clark et al., 2014).

In order to dig into the complexities of linguistic pluralism, its creative nature, and its potential in the development of creative biliteracy pedagogies, it is important to value what Blommaert (2010) stated, “[m]ovement of people across space is never a move

across empty spaces” (p. 6) and advocate for pedagogical practices that value the creative and improvisational practices in classrooms (Canagarajah, 2011).

García and Li (2017) described Canagarajah’s work, “translingual practices focus on the social practices of mixing modes and symbol systems as a creative improvisation to adapt to the needs of the context and the local situations” (pp. 122–123). Bilingual speakers’ communicative patterns and language practices are drawn from their linguistic repertoire, which is socially, culturally, and historically constructed (García, 2011; Smith, 2019). The requirement of speaking one language only or “standard” English in school ignores the dynamic and creative feature of bilingualism (García and Solorza, 2020). In fact, drawing on knowledge from their entire linguistic repertoire, bilingual students have an opportunity to deploy their full capacities in learning (Baker, 2011; García, 2011).

In Bauer et al. (2020) research, students in an English-Spanish dual language program purposefully used codes from two languages to communicate and ensure their understanding. Creese and Blackledge (2010) argued that code-switching implemented by educators as a resource in their classroom contributed to students’ identity performance, confidence, and academic achievement. Similarly, Canagarajah’s (2013) work promoted hybrid forms of writing consisting of dominant genre conventions as well as multilingual and diverse language resources. Findings suggested this hybrid form was beneficial for both multilingual students and native speakers of English when the students were in a space where both language experiences were equally valued, disrupting a common notion that multilingual students are deficient.

Bilingual speakers creatively use their entire linguistic repertoire to form communities of practice (Lave and Wenger, 1991) and engage in negotiated literacy during which they share textual conventions, strategies, and semiotic symbols and signs in order to communicate. Hornberger (2005) suggests, “bi/multilinguals’ learning is maximized when [bilingual speakers] are allowed and enabled to draw from across all their existing language skills (in two+ languages), rather than being constrained and inhibited from doing so by monolingual instructional assumptions and practices” (p. 607). Using in and between two languages is not a problem but a strategic, expressive identity and meaning-making process (Gutiérrez et al., 1999; Franquiz and de la Luz Reyes, 1998; García, 2011; Gort, 2008; Gutiérrez et al., 1999; Palmer et al., 2014) as well as a form of linguistic access, agency and equity in discursive learning spaces (Blommaert, 2010; Smith, 2019). Ignoring or stigmatizing alternative uses of English may stymie students from developing an accurate comprehension of language use (Matsuda, 2001) and undermine the “creativity that arises from belonging to multiple communities” (de Saint-Georges, 2013, p. 3). Educators are urged to value it as a legitimate public practice and recognize bilingual students’ creativity in literacy practices (Zhang and Guo, 2017).

Creativity in everyday biliteracy practices

Languages that bilingual speakers own are always localized, always dynamic, and always evolving (Blommaert, 2010; Pennycook, 2010; Shohamy, 2006). In this section, we reviewed and listed articles addressing how biliteracy is practiced as a creative and powerful ‘tool’ by bilingual speakers to have social existence in classrooms and access to a wider range of resources.

Identity development & translanguaging

Bilingual students’ identities are shaped profoundly by their literacy experiences (Jiménez, 2000; Smith, 2019). Their sense of belonging increases as their languages and identities are accepted, used, and valued and the classroom legitimizes the way bilingual students perform their own constructed identities through translanguaging (de Jong et al., 2020). Kabuto’s (2013) investigation distinctly highlighted the importance of teachers providing openings for children to maintain their identities. In this longitudinal observation, an emergent bilingual with the kindergarten teacher who was aware of the child’s biliterate capacities demonstrated her meaning-making process in complex and varied ways. On the other hand, with the first-grade teacher who was not aware of the child’s biliteracy and therefore did not encourage, the young girl’s own perception shifted to exhibit more monolingual modes of discourse versus bilingual.

Sayer’s (2013) work discussed Spanish-English students’ language use moved fluidly and dynamically not just between Spanish and English but also between deviations for specific social and academic purposes. This act of translanguaging offers opportunities for both the teacher and students to engage and construct discursive spaces in which they can practice and acquire social meanings in school. Davila (2017) discussed newly-immigrated students’ role as a language broker in a Swedish school. Through the help of teachers, language aides, and peers who valued and supported first language use and development, the students built social and linguistic capital, disrupting cultural and linguistic deficit stereotypes with and through practicing multilingual identities. These acts of translanguaging are agential and creative, as speakers decide when and how to follow or not to follow normative language constructs, “pushing and breaking the boundaries between the old and the new, the conventional and the original, and the acceptable and the challenging” (Li, 2011, p. 382).

Language brokering

Bilingual students use their language strengths to increase human capital, particularly through the practice of language brokering. Orellana’s (2009) ethnographic work demonstrated that children used both Spanish and English as language brokering on behalf of their families. Children, proficient in both languages, worked with their families translating and interpreting in the home and

community. Zhang's work (2007) demonstrated a nine-year-old Chinese-English bilingual communicating and translating for his grandparents when they were out talking with native speakers of English so that the older generation was able to participate in social practices. Orellana et al. (2003) found that Spanish-English bilingual students not only translate words, letters and phone calls for their parents but also legal documents and bank statements. Their family members depended on their bilingualism and expected their help.

For many bilingual students, translating is a part of their life (Jiménez, 2000). In Jiménez's research, a Spanish-English bilingual kid who was labeled as having learning disabilities expressed his concerns about losing Spanish and played the role as a Spanish tutor for his sister. While schools often assume that bilingual students come from homes/communities where literacy is impoverished or not well supported (Morrow, 1995), bilingual children do not lack literacy experiences outside of school. In fact, they exercise power and actively participate in their households and adults' socialization by helping their families access resources and information through a variety of ways across social domains.

Translanguaging as human capital building

Biliteracy and linguistic fluidity, such as in translanguaging, offers a strong basis for the development of bilingual students' voices, identities, cultural contexts, and critical knowledge (Palmer et al., 2014). Li and Wu (2009) argued that the use in and between two languages is a symbolic and creative capital that bilingual speakers own. The authors observed interactions between British-born Chinese students and their Chinese language teachers in complementary schools in Manchester and found that although the school policy was to speak Chinese only, bilingual children were able to creatively use English and Chinese to challenge the traditional ways of Chinese teaching and gained control of the classroom. They used their full linguistic repertoire to negotiate learning practices in class and challenged teachers' authority. The examples in this study suggested that bilingual students were not passive learners in school; they made the school a safe environment through bilingual practices.

Restrict language policies in school and community would limit bilingual students' literacy experiences (Batista-Morales et al., 2019; Moll and Gonzalez, 1994; Ruiz, 1984; Smith, 2016) as well as their social existence (Medina, 2004). Scholars suggested implementing critical literacy pedagogies to create a bilingual as well as bicultural space in school. de los Ríos (2017) discussed how music allowed a Mexican-American student to use the entire linguistic repertoire to tell border tension. Medina's (2010) study showed when the immigrant students were exposed to culturally and linguistically relevant texts, they compared and connected cultural contexts in Mexico and the US, discussed their perceptions of border control, and shared their awareness of border lives. The students collaboratively moved the discussion forward and gave voices to injustice. In Batista-Morales et al. (2019) research, Cristian, a second-grade Spanish-English bilingual student, proposed plan to pick up trash in his community. He and his mom created and distributed bilingual flyers and had a clean-up walk in the neighborhood. Cristian was empowered to respond to an issue in his community when his language and home knowledge was valued. Bilingual students' critical consciousness and their active roles are fostered when the teaching practices focus on what they can do with languages and allow space for bilingual students to creatively use and play languages (Cervantes-Soon et al., 2017).

Research shows that globally bilingual and multilingual speakers engage in the production of highly creative use of languages as contact zones or in the "social spaces where cultures meet, clash, and grapple with each other ..." (Pratt, 1991, p. 34). Centering the notions of linguistic imagination in relation to "invention," "desinvention" and "reconstruction" (Makoni and Pennycook, 2007) in the social mapping of language, we can get a potential sense of the emerging new social futures and forms of existing/new multilingual pedagogical territories (Gutierrez et al., 1999). In the next section, we dive into the use of drama and play as a critical pedagogy in a multilingual and multicultural classroom.

Drama and play pedagogies as performative spaces in biliteracy practices

The use of drama and play as a critical language pedagogy is situated in a perception of education as focused on the ways young people make sense of themselves and the world. According to Medina et al. (2014):

[L]earning is perceived as in motion, emerging from the experiences of all those involved in any one learning moment. This perspective of learning as performed, embodied, and emergent profiles new critical reflections on the ways we are participants of multiple social worlds. From these perspectives, learning is seen as something people "do" in communication and interaction. People's ways of knowing move to the center of any pedagogical exploration so that learning is reframed from focusing on "outcomes" to "incomes"—the histories, cultural and intellectual resources and identities learners bring to and create in classrooms (p. 71).

A critical performative pedagogical approach positions the classroom as an experience that "privileges the fluid, ongoing, often contradictory features of human experience..." (Pineau, 2002, p. 29). All classroom events, either inside or outside of curriculum, are perceived as "spectacles" produced at the intersections of culture and identity in the making (Diamond, 1996; Pineau, 2002). Weltsek and Medina (2007) explained, "classrooms are perceived as spaces where students and teachers perform and imagine multiple social realities addressing political issues, moving beyond superficial understandings of 'difference,' 'the other,' or assumed

‘naïve’ notions of empowerment and instead explore the embedded multiplicity of discourses” (p. 8). In this way, the dynamic interplay between self and other is situated as the space where learning takes place.

Pedagogical practices that foster and support interactive learning spaces built around collaborative performance offer opportunities for bilingual children to cross borders between languages (García et al., 2011). Souto-Manning suggests (2017) when play is used equitably in classrooms, young children are provided openings to “develop sophisticated understandings of academic knowledge and concepts through authentic, engaging, participative, and agentic ways” (p. 785). Through performative practices, children are able to draw on multiple resources to dramatize their thinking; real or imagined events become coupled with their own development of language (Wright et al., 2008). With multiple linguistic profiles at drama and play, a bilingual classroom turns into a space where young children make sense of knowledge and mediate their own identities (García et al., 2011). Furthermore, when intertwined with interpretations from both bilingual and non-bilingual children, the learning environment transitions to a dynamic and ever-evolving space for sharing ideas using personally meaningful modes of expression (Bengochea et al., 2018; Canagarajah, 2013). There is a need to expand and articulate effective pedagogical strategies and practices that assume the importance of personal and social development of biliterate children. The embodied and complex actions that stem from performance-based interactive learning spaces give rise to the importance of infusing drama and play pedagogies in biliteracy classrooms.

Personal, linguistic, cultural and social resources accessed through drama and play

Linking these ideas together, brings forth the ever-evolving importance of providing opportunities for children to learn and express through a variety of performative pedagogies where biliteracy serves as a strength that allows room for influence from the personal, linguistic, social, and cultural resources at play. Moving in between real and imagined worlds serves as an opening to affect agency and meaning making practices and promotes more socially and culturally responsive practices that could benefit the lives of students in and out of the classroom (Edmiston, 2000). This can be especially productive because all children bring different kinds of knowledge to the conversation (Moll and Gonzalez, 1994) thereby enriching the process of knowing through drama and play.

Bengochea et al. (2018) explained an emergent bilingual preschooler’s engagement through a “Show and Tell” activity, where the children’s own sense of meaning making was elaborated as they used gestures and personal stories to describe their objects. Multiple metacognitive functions occurred as the bilingual child switched between languages in communication. The findings demonstrated the holistic communicative competence in and between two languages and also highlighted the importance of validating personal and social resources accessed through this type of engaged showing and telling. Soderman and Oshio (2008) provided an example of competence and social behavior of children in an English-Chinese dual immersion program. The scholars learned that drama provided openings for the children to explore roles, learn rules, and examine power relationships between roles. Duran’s (2018) research explored how young bilingual children drew on social and cultural resources at play and developed “a metalinguistic awareness about how their two languages relate to each other” (p. 75). Deliman’s (2019) longitudinal qualitative case study found that drama-based strategies facilitated understanding of multiple perspectives and promoted opportunities for young children to cultivate empathetic values. Building on children’s social practices also helps enhance their critical awareness as Martello (2001) describes, “Experiencing language as a social practice first-hand and having opportunities to replay and reflect upon the relationships between language and social practices” (p. 201).

Biliteracy as a creative strength in drama and play allows children to move between languages, texts, and knowledge of literacy (Duran, 2018). When flexible bilingual pedagogic practices are enacted, bilingual children can “use their entire linguistic repertoire flexibly, meaningfully, and competently” (Gort and Sembiente, 2015, p. 23) in order to rework and mediate meanings for their social and cultural purposes (Wohlwend, 2019). Through drama and play, openings are created for multiple perspectives and voices to be heard (Medina, 2004). When personal and cultural assets become intertwined with performance, the meaning-making is embedded with real-world scenarios. These performative practices enlist biliteracy as a creative tool that drives enhanced meaning-making through the rich histories, personal stories, and cultural traditions of the playmakers.

Conclusion

In this chapter, we argue that biliteracy is a creative strength. Canagarajah (2011) asserts that translanguaging is performative and a naturally occurring phenomenon that “cannot be completely restrained by monolingual educational policies” (p. 8). Bilingual students are not passive learners or problematic English speakers; they think in and between languages and creatively use the languages through varied and complex ways. There is a need for educators to value bilingual students’ strength in literacy and enact pedagogical practices that allow bilingual students to use their entire linguistic repertoire and access their personal, cultural, and social resources in order to deploy their full capacities in learning (García, 2011).

This chapter included articles addressing critical literacy practices that value and foster bilingual students’ voices in class and specifically discussed using performative pedagogy to connect their personal, cultural, and social resources to the knowledge construction through embodied practices. The collaborative actions and interactions are personally meaningful and demonstrate how meaning-making extends beyond language. When drama and play is structured by the children, there is as the children mediate and negotiate multiple meanings, identities, and social resources. Performance-based play and drama in education creates spaces for students to showcase their multilingual repertoires with greater control of the meaning-making (Dunn, 1998).

References

- Baker, C., 2011. *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters, Bristol, UK.
- Batista-Morales, N.S., Salmerón, C., DeJulio, S., 2019. Their words, their worlds: critical literacy in bilingual spaces. *Biling. Res. J.* 42 (4), 471–490.
- Bauer, E.B., Colomer, S.E., Wiernelt, J., 2020. Biliteracy of African American and Latinx kindergarten students in a dual-language program: understanding students' translanguaging practices across informal assessments. *Urban Educ.* 55 (3), 331–361.
- Bengochea, A., Sembiente, S., Gort, M., 2018. An emergent bilingual child's multimodal choices in sociodramatic play. *J. Early Child. Literacy* 18 (1), 38–70.
- Blommaert, J., 2010. *A Sociolinguistics of Globalization*. Cambridge University Press, Cambridge, UK.
- Canagarajah, S., 2011. Translanguaging in the classroom: emerging issues for research and pedagogy. *Appl. Ling. Rev.* 2 (2011), 1–28.
- Canagarajah, S., 2013. Negotiating translanguaging literacy: an enactment. *Res. Teach. Engl.* 48 (1), 40–67.
- Cervantes-Soon, C.G., Dorner, L., Palmer, D., Heiman, D., Schwerdtfeger, R., Choi, J., 2017. Combating inequalities in two-way language immersion programs: toward critical consciousness in bilingual education spaces. *Rev. Res. Educ.* 41 (1), 403–427.
- Greese, A., Blackledge, A., 2010. Translanguaging in the bilingual classroom: a pedagogy for learning and teaching? *Mod. Lang. J.* 94 (1), 103–115.
- Cummins, J., 1980. The construct of language proficiency in bilingual education. *Curr. Issues Bilingual Educ.* 81–103.
- Cummins, J., 2000. *Language, Power, and Pedagogy: Bilingual Children in the Crossfire*. Multilingual Matters, Clevedon, UK.
- Davila, L.T., 2017. Newly arrived immigrant youth in Sweden negotiate identity, language & literacy. *System* 67, 1–11.
- de Jong, E.J., Coulter, Z., Tsai, M.C., 2020. Two-way bilingual education programs and sense of belonging: perspectives from middle school students. *Int. J. Biling. Educ. BiLing.* 1–13.
- de los Rios, C.V., 2018. Toward a corridista consciousness: learning from one transnational youth's critical reading, writing, and performance of Mexican corridos. *Read. Res. Q.* 53 (4), 455–471.
- de Saint-Georges, I., 2013. Multilingualism, multimodality and the future of education research. In: de Saint-Georges, I., Weber, J.J. (Eds.), *Multilingualism and Multimodality: Current Challenges for Educational Studies*. Sense Publisher, Rotterdam, The Netherlands, pp. 1–10.
- Deliman, A., 2019. *Negotiating Diverse Perspectives: Early Elementary Students Cultivate Empathy Through Children's Literature and Dramatic Inquiry*. PhD Dissertation. Indiana University.
- Diamond, E., 1996. Introduction. In: Diamond, E. (Ed.), *Performance and Cultural Politics*. Routledge, New York, pp. 1–12.
- Dunn, J., 1998. "This time I'll be the golden bird": a call for more child-structured dramatic play. *Res. Drama Educ.* 3 (1), 55–66.
- Duran, L., 2018. Understanding young children's everyday biliteracy: "Spontaneous" and "scientific" influences on learning. *J. Early Child. Literacy* 18 (1), 71–96.
- Edmiston, B., 2000. Drama as ethical education. *Res. Drama Educ.* 5 (1), 63–84.
- Farr, M., Song, J., 2011. Language ideologies and policies: multilingualism and education. *Lang. Linguistics Compass* 5 (9), 650–665.
- Franquiz, M.E., de la Luz Reyes, M., 1998. Creating inclusive learning communities through English language arts: from "Chancas" to "Canicas". *Lang. Arts* 75 (3), 211–220.
- García, O., Lin, A., 2017. Extending understandings of bilingual and multilingual education. In: García, O., Lin, A., May, S. (Eds.), *Bilingual and Multilingual Education*, third ed. Springer, pp. 117–128.
- García, O., Solorza, C.R., 2020. Academic language and the minoritization of US bilingual Latinx students. *Lang. Educ.* 1–17.
- García, O., Sylvan, C.E., 2011. Pedagogies and practices in multilingual classrooms: singularities in pluralities. *Mod. Lang. J.* 95 (3), 385–400.
- García, O., Starcevic, M., Terry, A., 2011. The translanguaging of Latino kindergartners. In: Potowski, K., Rothman, J. (Eds.), *Bilingual Youth: Spanish in English-Speaking Societies*. John Benjamins Publishing Company, Philadelphia, PA, pp. 33–56.
- García, O., 2011. *Bilingual Education in the 21st Century: A Global Perspective*. John Wiley & Sons, Hoboken, NJ.
- Gort, M., Sembiente, S., 2015. Navigating hybridized language learning spaces through translanguaging pedagogy: dual language preschool teachers' languaging practices in support of emergent bilingual children's performance of academic discourse. *Int. Multidiscip. Res. J.* 9 (1), 7–25.
- Gort, M., 2008. "You give me ideal!": collaborative strides toward bilingualism, biliteracy, and cross-cultural understanding in a two-way partial immersion program. *Multicult. Perspect.* 10 (4), 192–200.
- Guerra, J.C., 2016. Cultivating a rhetorical sensibility in the translingual writing classroom. *Coll. Engl.* 78 (3), 228–233.
- Gutiérrez, K.D., Baquedano-López, P., Tejeda, C., 1999. Rethinking diversity: hybridity and hybrid language practices in the third space. *Mind Cult. Activ.* 6 (4), 286–303.
- Hornberger, N.H., 1989. Continua of biliteracy. *Rev. Educ. Res.* 59 (3), 271–296.
- Hornberger, N.H., 2005. Opening and filling up implementational and ideological spaces in heritage language education. *Mod. Lang. J.* 89 (4), 605–609.
- Jiménez, R.T., 2000. Literacy and the identity development of Latina/o students. *Am. Educ. Res. J.* 37 (4), 971–1000.
- Kabuto, B., 2013. Becoming biliterate within the crossroads of home and school: Insights gained from taking multiple theoretical stances. In: Martin, D. (Ed.), *Researching Dyslexia in Multilingual Settings*. Multilingual Matters, Bristol, UK, pp. 157–173.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, UK.
- Li, W., Wu, C.J., 2009. Polite Chinese children revisited: creativity and the use of code switching in the Chinese complementary school classroom. *Int. J. Biling. Educ. BiLing.* 12 (2), 193–211.
- Li, W., 2011. Moment analysis and translanguaging space: discursive construction of identities by multilingual Chinese youth in Britain. *J. Pragmat.* 43 (5), 1222–1235.
- Makoni, S., Pennycook, A., 2007. Desinventing and reconstituting languages. In: Makoni, S., Pennycook, A. (Eds.), *Desinventing and Reconstituting Languages*. Multilingual Matters, Clevedon, UK, pp. 1–41.
- Martello, J., 2001. Drama: ways into critical literacy in the early childhood year. *Aust. J. Lang. Literacy* 24 (3), 195–207.
- Martínez, R.A., 2018. Beyond the English learner label: recognizing the richness of bi/multilingual students' linguistic repertoires. *Read. Teach.* 71 (5), 515–522.
- Matsuda, P.K., 2001. Voice in Japanese written discourse: implications for second language writing. *J. Sec. Lang. Writ.* 10, 35–53.
- Medina, C.L., Campano, G., 2006. Performing identities through drama and teatro practices in multilingual classrooms. *Lang. Arts* 83 (4), 332–341.
- Medina, C.L., Coggin, L., Weltsek, G., 2014. Foregrounding emergence, embodiment, and critical practices: performance pedagogies in literacy methods. In: Brass, J., Webb, A. (Eds.), *Teaching the English Language Arts Methods: Contemporary Methods and Practices*. Routledge, New York, pp. 70–82.
- Medina, C.L., 2004. Drama wor(l)ds: explorations of Latina/o realistic fiction through drama. *Lang. Arts* 81 (4), 272–282.
- Medina, C.L., 2006. Identity and imagination of immigrant children: Creating common place locations in literary interpretation. In: *Process Drama and Multiple Literacies: Addressing Social, Cultural, and Ethical Issues*, pp. 53–70.
- Medina, C.L., 2010. "Reading across communities" in biliteracy practices: examining translocal discourses and cultural flows in literature discussions. *Read. Res. Q.* 45 (1), 40–60.
- Moll, L.C., Gonzalez, N., 1994. Lessons from research with language-minority children. *J. Read. Behav.* 26 (4), 439–456.
- Moll, L.C., 2014. *LS Vygotsky and Education*. Routledge, New York, NY.
- Morrow, L.M., 1995. *Family Literacy: Connections in Schools and Communities*. International Reading Association, Newark, DE.
- Orellana, M., Dorner, L., Pulido, L., 2003. Accessing assets: immigrant youth's work as family translators or para-phrasers. *Soc. Probl.* 50 (4), 505–524.
- Orellana, M., 2009. *Translating Childhoods: Immigrant Youth, Language and Culture*. Rutgers University Press, New Brunswick, NJ.
- Palmer, D.K., Martínez, R.A., Mateus, S.G., Henderson, K., 2014. Reframing the debate on language separation: toward a vision for translanguaging pedagogies in the dual language classroom. *Mod. Lang. J.* 98 (3), 757–772.
- Pennycook, A., 2010. *Language as a Local Practice*. Routledge, New York, NY.

- Perez, B., Torres-Guzman, M.E., 1992. *Learning in Two Worlds: An Integrated Spanish/English Biliteracy Approach*. Longman, New York, NY.
- Perryman-Clarke, S., Kirkland, D., Jackson, A. (Eds.), 2014. *Students' Rights to Their Own Language: A Critical Sourcebook*. Bedford/St. Martin's, Boston, MA.
- Pineau, E.L., 2002. Critical performative pedagogy: fleshing out the politics of liberatory education. In: Stucky, N., Wimmer, C. (Eds.), *Teaching Performance Studies*. Southern Illinois University Press, Carbondale, pp. 41–54.
- Pratt, M.L., 1991. Arts of the contact zone. *Profession* 33–40.
- Reyes, I., 2012. Biliteracy among children and youths. *Read. Res. Q.* 47 (3), 307–327.
- Ruiz, R., 1984. Orientations in language planning. *NABE J.* 8 (2), 15–34.
- Sayer, P., 2013. Translanguaging, TexMex, and bilingual pedagogy: emergent bilinguals learning through the vernacular. *Tesol Q.* 47 (1), 63–88.
- Shohamy, E., 2006. *Language Policy: Hidden Agendas and New Approaches*. Routledge, New York, NY.
- Smith, P., 2016. A distinctly American opportunity: exploring non-standardized English(es) in literacy policy and practice. *Policy Insights Behav. Brain Sci.* 3 (2), 194–202.
- Smith, P., 2019. (Re)positioning in the Englishes and (English) literacies of a black immigrant youth: towards a transraciolinguistic approach. *Theory Into Pract.* 58 (3), 292–303.
- Soderman, A., Oshio, T., 2008. The social and cultural contexts of second language acquisition in young children. *Eur. Early Child. Educ. Res. J.* 16 (3), 297–311.
- Souto-Manning, M., 2017. Is play a privilege or a right? And what's our responsibility? On the role of play for equity in early childhood education. *Early Child. Dev. Care* 187 (5–6), 785–787.
- Weltsek, G., Medina, C., 2007. Critical performative spaces: identity, stance and discourse in process drama as critical literary interpretation. In: Blackburn, M.V., Clark, C. (Eds.), *New Directions in Literacy Research for Teaching, Learning and Political Action*. Peter Lang, New York, NY, pp. 255–275.
- Wiley, T.G., 2014. Diversity, super-diversity, and monolingual language ideology in the United States: tolerance or intolerance? *Rev. Res. Educ.* 38 (1), 1–32.
- Wohlwend, K.E., 2019. Play as the literacy of children: imagining otherwise in contemporary childhoods. In: Alvermann, D.E., Unrau, N.J., Sailors, M. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, NY, pp. 301–318.
- Wright, C., Bacigalupa, C., Black, T., Burton, M., 2008. Windows into children's thinking: a guide to storytelling and dramatization. *Early Child. Educ. J.* 35 (4), 363–369.
- Zhang, Y., Guo, Y., 2017. Exceeding boundaries: Chinese children's playful use of languages in their literacy practices in a Mandarin–English bilingual program. *Int. J. Biling. Educ. BiLing.* 20 (1), 52–68.
- Zhang, G., 2007. Multiple border crossings: literacy practices of Chinese-American bilingual families. In: Purcell-Gates, V. (Ed.), *Cultural Practices of Literacy: Case Studies of Language, Literacy, Social Practice, and Power*. Lawrence Erlbaum, Mahwah, NJ.

Systemic functional linguistic and disciplinary literacy development in multilingual K-12 schools in the United States

Meg Gebhard and Nicolas Blaisdell, University of Massachusetts, Amherst, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: the challenges of teaching and learning disciplinary literacies	54
SFL: a social semiotic perspective of grammar	55
Registers	57
Genres	58
Critiques to SFL pedagogy and their responses	60
Future directions for SFL scholarship in the US: personal reflections	62
References	63

Introduction: the challenges of teaching and learning disciplinary literacies

In introducing Halliday’s systemic functional linguistics (SFL), Meg, the first author, asks teachers and literacy researchers to compare the texts shown in Table 1. These examples come from an elementary classroom where students were required to conduct an experiment as a way of developing an understanding of empiricism and the scientific method. One group of students, five girls, decided to test which type of bubble gum would enable them to blow the biggest bubble. After buying different packages of Bubble Yum, Bazooka, and Trident Bubble Gum, the girls developed a procedure for chewing the gum for so many seconds, blowing bubbles, and then placing their bubbles on pieces of paper. They then measured the diameter of each bubble and recorded the data in a table. Their next task was to write a laboratory report on a poster for a science fair for family and community members. In explaining her work, one student named Ava (a pseudonym) explained their project orally as shown on the left of Table 1. Meg asked her to consider how scientists, using the scientific method, might describe the same experience. She offered the text shown on the right. Ava’s response was to laugh out loud and declare, “I’m not talking like that—what do you think I am, a geek?!”

This short interaction reveals how different meaning-making choices in the form of gestures, graphs, words, and grammatical patterns reflect different ways of constructing ideas and experiences; enacting voice, identity, and relationship with others; and managing the flow of extended discourse for a specific purpose and audience. For example, Ava’s text recounts a set of events using predominantly a subject-verb-object pattern. Beginning with the pronoun “we,” her point of departure was herself and her group. In constructing their experiences orally, Ava used “we” as well as the words “gum” and “bubble” to convey what her group did using verbs in the past tense to express everyday activities (e.g., chewed, blew). Regarding voice, she made linguistic choices that conveyed excitement and investment in this project as an engaged student to an adult she most likely positioned in a teacher role (e.g., Look! Really great, so cool!). She also managed the flow of information by repeating keywords to stay on topic and build information over a series of sentences. Lastly, regarding coherence, she used oral discourse markers such as “like,” “so,” and “and” to connect ideas in an additive fashion.

In comparison, Meg’s text, about the very same events, used language and other meaning-making systems, such as a graph, to construct more scientific and technical expressions of disciplinary knowledge, as is more common in the context of “doing science” in school (Lemke, 1990). For example, in explaining the results of the experiment empirically, Meg began with a focus on the data as displayed in the graph to make the data the “doer” or “agent” in the sentence, not a human actor (e.g., “As illustrated in the graph, the data show that Trident chewing gum produced the largest bubbles in this experiment”). This distinction is important because a central premise of the scientific method is close observation and letting the data speak for itself with less human bias. The verb choices also convey a more empirical, as opposed to personal experience, through the use of nouns such as “experiment” and “results” and the verb “hypothesized.” Moreover, the sentence patterns are more varied and longer than the subject-verb-object patterns found in Ava’s text. For example, scientific discourse typically packs more information into sentences through the use of adjectives in longer noun groups and the use of prepositional phrases (e.g., “the largest bubbles in this experiment”). Meg’s text also

Table 1 Comparison of everyday and scientific discourse.

We chewed gum and then blew bubbles. Look! (pointing to a bubble on the paper) Like, Trident is really great, and like we blew the biggest bubble with that. So that’s what we did. And it’s so cool we got to have gum in school!	As illustrated in the graph, the data show that Trident chewing gum produced the largest bubbles in this experiment. These results were unexpected because we hypothesized that the brand of gum would not be a factor in the size of the bubble we blew.
---	---

packs more information into a single sentence using subordinating clauses beginning with “that.” Regarding voice, aspects of identity, and the relationship between interlocutors, Meg’s text is more formal and neutral in ways that construct greater social distance and authority rather than excitement and personal investment (e.g., “These results were unexpected”). Lastly, regarding managing the flow of information, Meg selected the conjunction “because” to explain rather than adding ideas together using “and.”

Ava’s response to Meg’s suggestions for revision indicates that she had an awareness of how language choices function semiotically, especially regarding issues of identity, given that she was not willing to speak like someone who she characterized as “a geek.” And perhaps she had a point. As a fourth-grader, her group’s ability to design a very clever experiment met age-appropriate expectations regarding developing a beginning understanding of empiricism and the scientific method. However, a pressing question remains: How can teachers prepare all students to read, write, and critique increasingly technical texts as they transition from elementary to secondary school, and then to college and/or the world of work? How can teachers prepare students to become generators of knowledge in a rapidly changing social, economic, and political world? Further, how can teachers teach disciplinary literacies in ways that draw on students’ multilingual resources in culturally sustaining ways?

The New London Group, an international group of literacy scholars, explored this crucial question in their 1996 landmark article, “A Pedagogy of Multiliteracies: Designing Social Futures.” Drawing in part on Halliday’s theory of SFL, the international team of authors promoted a “pedagogy of multiliteracies” to advance multilingual/multimodal approaches to teaching and learning in the 21st century (New London Group, 1996, p. 64). Despite the influence of this often-cited article published over twenty-five years ago, teachers, literacy researchers, and policymakers continue to struggle to answer the question of how teachers can design instruction to provide all students with access, scaffolding, and opportunities for critical reflection regarding the types of dense texts they will encounter in school, work, and life (Accurso and Gebhard, 2020). In response to this persistent problem, this chapter explains Halliday’s theory of SFL and describes how educational linguists have used this theory to design disciplinary literacy instruction for a wide variety of students in different K-12 international contexts using the “Teaching and Learning Cycle” (Christie and Derewianka, 2008; Derewianka and Jones, 2016; Janks, 2010; Humphrey, 2017; McCabe, 2021; Rose and Martin, 2012; Schleppegrell, 2004). Next, it provides a critique of SFL-informed pedagogy and concludes with recommendations for future research.

SFL: a social semiotic perspective of grammar

SFL provides a social semiotic theory of meaning-making, learning, and social change (Halliday, 1973, 1978, 1993). Initially developed in the 1960s by Michael Halliday, SFL has expanded into a theory that informs many branches of applied linguistics, including educational linguistics, studies of multimodality, and critical analyses of institutional discourse in a wide variety of fields (e.g., Bartlett and O’Grady, 2017; Christie and Martin, 1997; Brisk and Schleppegrell, 2021; Unsworth, 2001; Young and Harrison, 2004). Halliday’s contribution to linguistics is found in his conceptualization of grammar as a functional meaning-making resource that expands to reflect and construct both the immediate situational and broader cultural contexts in which semiotic systems are used and evolve over time (Halliday, 1993; Halliday and Hasan, 1985; Halliday and Matthiessen, 2014). This context-sensitive and multimodal perspective stands in contrast to behavioral and psycholinguistic conceptions of grammar that have tended to dominate approaches to teaching literacy in the US (e.g., Alvermann et al., 2019; Gebhard, 2019). For example, behavioral understandings of language learning, informed by B.F. Skinner, conceptualize grammar as a set of formal verbal behaviors and language learning as the assembly of sounds, words, sentences, and eventually longer stretches of discourse with stress placed on formal accuracy through practice via formal rules and prescribed text structures (Skinner, 1957; Kaplan, 1966). From this perspective, the work of a teacher typically centers on the adoption of leveled texts to support learners in practicing simple to more complex linguistic patterns to enable them to communicate about everyday experiences. Critics of this approach argue that it may enable students to pass exams that target formal accuracy, but it does not necessarily enable them to communicate about meaningful topics for authentic audiences in creative ways (e.g., Elbow, 1973; Krashen, 1987).

In response, informed by the influential work of Noam Chomsky (1967), many scholars and teachers turned to a psycholinguistic understanding of language and a communicative approach to language and literacy instruction (e.g., Alvermann et al., 2019; Gebhard, 2019). This perspective suggests that all humans are endowed with the cognitive ability to acquire the remarkable complexities of their home language as well as additional languages according to universal rules governing all human languages. Therefore, the job of the teacher is to engage learners in authentic communicative tasks geared toward the learners’ stage of language and literacy development and support learners in creatively problem-solving as they learn to read and write (e.g., Calkins, 1994). This more process-oriented, as opposed to form-oriented approach, often takes the form of teachers guiding students in brainstorming, mind-mapping, drafting, revising, editing, and publishing texts about their personal experiences and opinions as a way of learning to express themselves with less emphasis placed on explicit instruction in how language and other semiotic systems work in the types of texts they are required to read and write across the curriculum (e.g., Elbow, 1973).

Likewise, influential researchers in the field of second language acquisition, such as Krashen (1987), argue that for students to acquire an additional language naturally in ways that mirror first language acquisition, teachers should avoid explicit instruction in grammatical forms. Instead, they should engage learners in communicative activities that generate ample amounts of comprehensible linguistic input and provide opportunities for students to generate linguistic output in a stress-free way that does not focus on errors.

Despite the widespread appeal of more natural and process-oriented perspectives of language and literacy development, many scholars and teachers have noted that these approaches, even after years of classroom instruction, have not addressed the equity gap that has long existed between dominant and non-dominant students, especially for historically marginalized students such as speakers of African American English (e.g., [Delpit, 1988](#); [Gebhard, 2019](#); [Westerlund and Besser, 2021](#)). Nor have these approaches supported monolingual English speakers in the US in developing advanced literacies in additional world languages (e.g., [Byrnes et al., 2010](#)). In response, SFL scholars argue that literacy programs anchored solely in a psycholinguistic perspective of literacy development do not prepare first and second language learners to read, write, and critically engage with the wide range of dense informational texts they are likely to encounter as they attempt to make sense of disciplinary knowledge in mathematics, the sciences, social studies, and language arts (e.g., [Rose and Martin, 2012](#)).

Evidence of this persistent educational opportunity gap, or what [Ladson-Billings](#) refers to as an “educational debt,” is apparent in an analysis of disaggregated assessment data in the US ([Ladson-Billings, 2006](#), p. 3; [Howard, 2019](#)). For example, drawing on the 2017 National Assessment of Educational Progress Data, [Howard \(2019\)](#) reports that “among fourth-graders, 49% of African Americans, 46% of Latinx, and 52% of American Indians read below grade level compared to 22% of whites.” These disparities exist in the area of mathematics in that these same data show “approximately 37% of African American students, 29% of Latinx students, and 31% of American Indian students” score below grade level in “mathematics proficiency compared with 12% of whites ([Howard, 2019](#), p. 18).” [Howard](#) argues that despite widespread gains in graduation rates for students of color, there still exist inequities in these students’ schooling experiences as evidenced by over-representation in special education programs, limited access to advanced placement courses, and lower scores on college entrance exams. In accounting for these disparities, he points to prevailing “deficit thinking” regarding students of color and speakers of varieties of English as well the history of “institutional racism” in US schools ([Howard, 2019](#), pp. 22–30).

In documenting similar problems facing students institutionally designated as “English language learners” (ELL), [Walqui and Bunch \(2019\)](#) provide several important insights that echo [Howard’s](#) analysis. They describe the problem of “deficit” views of multilingual learners and “reductive pedagogical approaches” that “leave some students without access to high quality, intellectually rich discipline-based curriculum and pedagogy” (p. xi). They add that teachers are often viewed as “delivers” of curriculum prepared by others as opposed to capable “designers” of instructional materials themselves (p. xi).

In response, literacy scholars working in a variety of contexts, including first and second language and literacy development, have turned to context-sensitive conceptions of language and literacy development (e.g., [Halliday, 1993](#); [Halliday and Hasan, 1985](#); [Hasan and Williams, 1996](#); [Unsworth, 2001](#)). As described by [Halliday](#), [Hasan](#), and their colleagues, this perspective maintains that language and other meaning-making systems, such as gestures, images, equations, and graphs, cannot be understood without an analysis of the immediate context in which communication takes place, develops, and evolves, nor can these semiotic systems be understood separately from issues of power, ideology, and socialization.

In arguing for a more context-sensitive perspective of the grammar, [Halliday](#) argued for an expansive functional perspective of sociolinguistic practices by maintaining that every act of communication performs three metafunctions simultaneously: The *ideational* function constructs ideas and experiences, such as [Ava’s](#) recounting of her personal experiences and [Meg’s](#) revoicing of this experience as a scientific claim through the selection of different lexical items and grammatical constructions. The *interpersonal* function enacts social roles and power dynamics, such as [Ava’s](#) use of linguistic choices that constructed her as a joyful learner versus [Meg’s](#) choices which construct a more authoritative, neutral, and detached voice—one [Ava](#) rejected because it represented an identity to which she was not oriented. And third, the *textual* function manages the flow of information to make extended oral, written, and multimodal discourse coherent and cohesive. For example, [Ava’s](#) repetition of keywords, synonyms, and pronouns created a lexical chain that stayed on topic and connected ideas using the conjunctions “and.” In contrast, [Meg’s](#) text used more discipline-specific lexical chains to construct a causal explanation using “because.”

In developing this trinocular theory of grammar as a functional social semiotic system, [Halliday](#) maintained that as children, such as [Ava](#), grow up and interact with the world, they have more experiences and ideas to share with a greater variety of people through their uses of a greater variety of language(s) and modes of communicating. Simultaneously, in interacting with a greater variety of people, children, like [Ava](#), learn to use language(s) and varieties of language(s) to construct different kinds of interpersonal relationships with family members, peers, teachers, co-workers, and bosses in ways that are shaped by ideologies related to race, class, gender, religion, and country of origin. Regarding the textual function, as they mature, children, like [Ava](#), learn how to manage the flow of information in oral, written, multimodal, and computer-mediated communication within the diverse cultural contexts in which they participate, including in their homes and neighborhoods, schools, places of worship, at work, and online. [Halliday](#) maintains that this expansion of semiotic resources drives the development of individual meaning-making repertoires as well as the evolution of meaning-making systems themselves (e.g., [Halliday, 1993](#)).

[Halliday’s](#) context-sensitive and functional model of grammar as it relates to K-12 schooling in the US is illustrated in [Fig. 1](#) ([Gebhard, 2019](#), p. 14). This model attempts to capture the complex dynamics that exist between texts and the multilayered institutional contexts in which they are produced and interpreted in schools as well as how schools as institutions are shaped by ideologies regarding language, learning, and the purpose of public education (e.g., [Gebhard, 1999, 2019](#); see also [Bowles and Gintis, 1976](#); [Fairclough, 1989, 1992](#)). This model is informed by over five decades of scholarship produced by educational linguists who have expanded upon [Halliday’s](#) work with the goal of “increasing the range of contexts within which children are likely to participate as they grow older, and the types of linguistic resources learners need to develop if they are to operate successfully in these contexts” ([Christie and Derewianka, 2008](#), p. 6). Based on their experiences as language learners, language teachers, and classroom researchers working with teachers, SFL scholars have theorized and documented how oral, written, multimodal, and

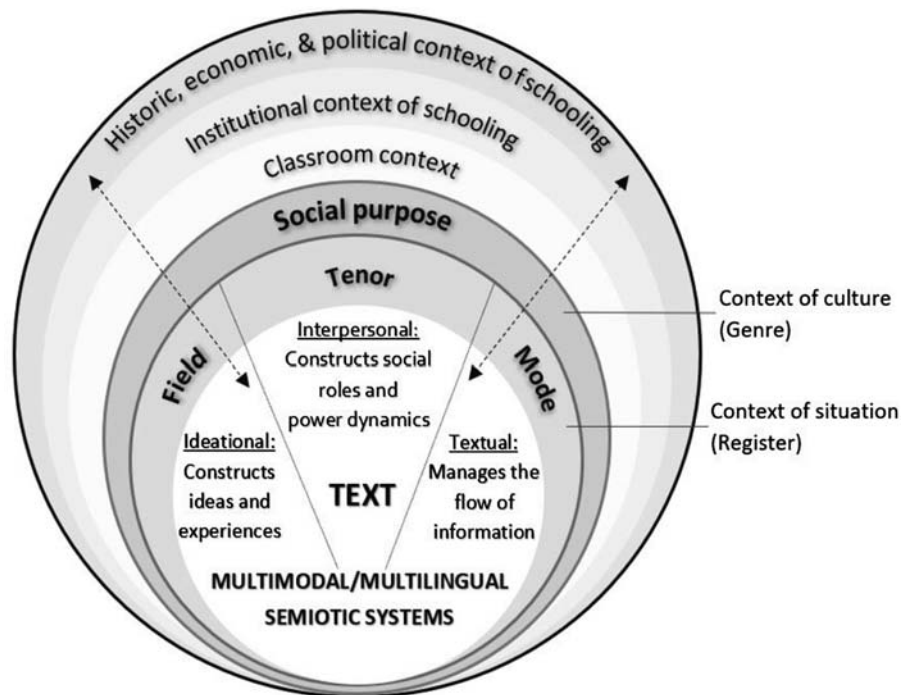


Fig. 1 An SFL perspective and text/context dynamics in schools as ideological institutions (Gebhard, 2019, p. 14). Copyright (2019 ©) From Teaching and Researching ELLs' Disciplinary Literacies by Meg Gebhard. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa PLC.

multilingual texts construct ideas and experiences, enact self-other dynamics, and manage the flow of communication in schools as historical, economic and political institutions that tend to produce and reproduce dominant ideologies regard language, learning, and social status (e.g., Cope and Kalantzis, 1993; Janks, 2010; Christie and Derewianka, 2008; Christie and Martin, 1997; Hasan and Williams, 1996; Humphrey, 2017; New London Group, 1996; Rose and Martin, 2012; Unsworth, 2001).

Beginning at the center of this diagram, SFL scholars propose that a text can be any multilingual/multimodal stretch of discourse, such as a student using a combination of Spanish and African American English to recount their weekends on their way to class; an editorial about the benefits of bilingual education written in English and Chinese in a bilingual school newspaper; an explanation of a linear equation on an algebra exam written in English; or a PowerPoint analysis of code-mixing and shifting in a novel by Toni Morrison written in Spanish in the context of a bilingual program. These oral, printed, multimodal, multilingual, and computer-mediated texts are constructed through the choice of different phonological, morphological, lexical, clause-level syntactic patterns and/or accompanying ways of using gestures, lines, colors, shading, and images to accomplish specific kinds of work in specific "contexts of situations," such as a high school hallway, a school newspaper room, a math class, or on an online learning platform.

These texts are also situated in the "context of culture" of a US high school. This more abstract layer of this striatal model is shaped by historical, economic, and political ideologies regarding the purpose of public education as well as institutional policies and practices in the US at the federal, state, district, and school levels (e.g., high stakes testing systems, English-only mandates, school funding formulas, student grouping practices, mandated textbooks, and teacher professional certification practices; see also Gebhard, 2019).

Registers

As shown in Fig. 1, at the level of the "context of the situation," Halliday and his colleagues refer to different constellations of sound, graphic, and clause-level lexicogrammatical choices as "registers" to capture the variety of choices people make when communicating about a particular topic for a purpose and audience in a specific situation (e.g., Rose and Martin, 2012). From an SFL perspective these patterned ways of making meaning construct ideas and experience through the ideational metafunction; the voice or stance of text through the interpersonal metafunction; and the coherence and cohesion through the textual function. More precisely, SFL scholars define registers with reference to specific "field," "tenor," and "mode" variables (Halliday and Hasan, 1985, pp. 38–39). For example, Ava and I constructed the same events using very different field, tenor, and mode resources that entail more than just selecting different vocabulary words to say the same thing. Rather, in addition to choosing different lexical items, we used different syntactical patterns, gestures, charts, graphs, and material objects to perform different functions for different social purposes. Ava's field choices constructed her personal experiences through her use of everyday and concrete nouns, referred to in SFL as grammatical

"participants;" verbs, referred to as "processes;" and adverbial and prepositional phrases, referred to as "circumstances." In contrast, my field choices are those required to explore an empirical question through the display of data represented in a graph and through the grammatical participants I selected, including "the data," "these results," and "a factor." The verbal processes, which included "show," "produced," "hypothesize" and various forms of the relational verb "to be." I also selected certain "circumstances" to add details and specificity related to time, manner, and place (e.g., "as illustrated in the graph, in this experiment").

Similarly, regarding tenor, Ava and Meg's linguistic choices reflect a different stance regarding the topic at hand and their relationship to their audience. Ava used a mix of declaratives and imperatives as well as highly emotional language to convey excitement to an interested adult who was curious about her work. In contrast, Meg used only declaratives, more neutral language, and the modal verb "would" to speculate in communicating a scientific explanation using a more detached and authoritative voice for an unknown academic audience.

Regarding mode, the flow of Ava's and Meg's text relied on the repetition of keywords, but Ava's connected ideas using oral discourse markers such as "so," "and" and "like." In contrast, Meg's more decontextualized text utilized longer clauses that were connected with the conjunction "because." These differences are summarized in [Table 2](#) in ways that demonstrate how metafunctions are realized through field, tenor, and mode resources.

Genres

As illustrated in [Fig. 1](#), SFL scholars have noted that field, tenor, and mode choices in the form of material phonological sound waves, print, gestures, graphs, and images do not just reflect and construct an isolated context of the situation. Rather register variables bundle in culturally expected ways to make extended discourse culturally recognizable as performing particular functions through genres, such as recounting an event, telling a story, writing an editorial, producing a mathematical explanation, or making an argument in a literary essay. Martin defines bundled register features as genres that operate at the level of the context of culture ([Martin, 1992](#)). He adds that genres are staged in that they have recognizable parts that perform different functions; are goal-oriented because they have a purpose and audience; and are social practices because they have evolved and will change over time as cultural practices shift ([Martin, 1992](#); [Martin and Rose, 2008](#); [Rose and Martin, 2012](#)). For example, if we overhear a conversation between students in the hallway in which one says, "Dude you are not gonna believe what happened," most of us, based on our previous experiences, might be able to predict that the student is about to tell a story. This genre typically begins with an orientation that locates the hearer/reader in a time and place and introduces a problem, such as "I was driving my dad's car on Saturday during that wicked storm when I ran out of gas." Our hypothetical storyteller might use a combination of Boston English and African American English in conveying a sequence of events that lead to a complication, then a resolution, and sometimes a commentary or "coda" that comments on the narrative as a whole (e.g., [Labov and Waletzky, 1967](#); [Martin and Rose, 2008](#)). In making registerial choices, our hypothetical storyteller might build the field by using packed noun groups (a blinding blizzard of white snow); a variety of action verbs (skidded), verbal verbs (screamed), and mental verbs (worried); and circumstances to add details regarding time, manner, and place (e.g., late at night, on the side of the road, with an empty tank of gas). They might also use emotional, judgmental, or evaluative language to convey their stance on the events using a combination of declaratives, exclamations, and questions (e.g., Damn, that was wild. Dad was mad salty!). Drawing on mode resources, they are likely to use time markers to move the arc of the narrative along in a coherent fashion (e.g., then, next, suddenly, after that, finally).

If this same student then headed off to their algebra class to take an exam in which they were given the equation $y = x^2 - 5$ and asked to give the "range" and "domain" and explain whether the relationship was a function, they would use a different genre and configuration of the field, tenor, and mode choices to achieve a very different purpose and for a different audience (e.g., [O'Halloran, 2007](#)). For example, if this student received adequate instruction in not just subject matter vocabulary, but also scaffolding and explicit instruction in how to explain their mathematical reasoning, they might produce a text similar to the one shown in [Table 3](#). This text reveals that successful students in mathematics can provide an organizing thesis, a set of definitions, a solution, an explanation of their thinking, and a graphic representation of an equation. Moreover, the field choices they make need to be specific to the domain of algebra, which constructs disciplinary knowledge through long noun-groups (e.g., the set of all real numbers) and specific verbs (e.g., substitute, intersect). The tenor is more formal, detached, and authoritative and is constructed through a set of declaratives with no emotion. In addition, the construction of this text requires the use of modal verbs to construct possibility (e.g., can be, could be). Regarding mode, the text flows smoothly through the use of repetition of keywords, the weaving of given and new information in a zig-zag fashion, the choice of certain cohesive devices that support the genre of an explanation (e.g., because, therefore), and the use of language to connect the written text with a graphic display ("as shown," see [Accurso et al., 2017](#)).

In reflecting on the disciplinary knowledge constructed in [Table 3](#), the teachers we work with often remark that they wish their students could produce texts like these, but they are not sure how to set such high expectations, especially if they work with newcomers to dense disciplinary discourses. In response, SFL scholars such as [Hammond and Gibbons \(2005, p. 6\)](#) maintain teachers must "support up" students rather than "dumb down" the curriculum. They contend teachers can do this by coupling a functional perspective of genre and register instruction with a Vygotskian perspective of literacy development. Building on the work of SFL scholars at the University of Sydney, they maintain that teachers can set high expectations when they also provide

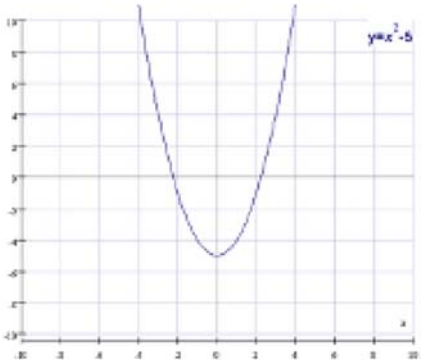
Table 2 A summary of SFL metafunctions and register variable used to construct meaning Gebhard and Blaisdell (In Press).

<i>Example 1</i> <i>We chewed gum and then blew bubbles. Look! (pointing to a bubble on the paper).... Like, Trident is really great and like we blew the biggest bubble with that. So that's what we did. And it's so cool we got to have gum in school!</i>		<i>Example 2</i> <i>As illustrated in the graph, the data show that Trident chewing gum produced the largest bubbles in this experiment. These results were unexpected because we hypothesized that the brand of gum would not be a factor in the size of the bubble we blew.</i>
Metafunctions	Register variable	Grammatical resources and their linguistic realizations
Ideational function: Constructs ideas and experiences; constructs the content of a text or what is happening	Field choices	Participants: Nouns, packed nominal groups - Doers/agents (e.g., <i>we, the data, these results, this experiment</i>) - Objects (e.g., <i>gum, the largest bubble, the brand of gum</i>) Process: Types of verbs: - Action (<i>chew, did, blew, produced</i>) - Saying not evident in this example (<i>said, yelled, screamed, moaned, whimpered</i>) - Mental (e.g., <i>hypothesize</i>) - Sensing (e.g. <i>look</i>) - Relating (<i>be, is</i>) Circumstances: Prepositional phrases, adverbs, and adverbial phrases that give information regarding time, manner, and place (e.g., <i>in school, as illustrated in the graph, in this experiment</i>)
Interpersonal function: Enacts social dynamics related to social distance and power dynamics; constructs the voice of the text or the social roles of speakers/hearers or writers and readers	Tenor Choices	Mood to give information and negotiate propositions - Declaratives to give information (e.g., <i>As illustrated in the graph, the data show that Trident chewing gum produced the largest bubbles in this experiment</i>). - Interrogatives to ask for information or an action or to make an offer (e.g., <i>This experiment posed the question, what brand of gum produces the largest bubble?</i>) - Imperatives to command or express excitement (e.g., <i>Look!</i>) Modality to construct degrees of certainty and possibility - Modal verbs (e.g., <i>may, might, can, could, should, will, must</i>) - Modal adverbs (e.g., <i>maybe, probably, absolutely</i>). - Modal nouns (e.g., <i>possibly, likelihood, certainty</i>) Appraisal system to express emotion, evaluate the quality of things or to judge behavior (e.g., <i>really great, so cool, largest</i>).
Textual function: Manages the flow of information in service of coherence and cohesion	Mode Choices	Repeating the same word, pronoun, or synonym to build ideas and stay on topic (e.g., <i>We got to chew gum in school. We blew the biggest bubble with Trident. We thought that was cool</i>). Weaving given and new information together using a zig-zag pattern to build ideas and stay on topic (e.g., <i>This experimented tested different brands of gum. The brand of gum that resulted in the largest bubble was Trident. Trident bubbles measured....</i> Use of cohesive devices and conjunctions to connect ideas (e.g., <i>so, and, because, first, second, third, however, therefore, in sum</i>). Nominalization to construct abstraction (e.g., <i>the bubble-blowing potential</i>).

high levels of explicit instruction and modeling within the students' "zone of proximal development" (Vygotsky, 1978). In Vygotskian terms, this level of social, linguistic, and cognitive support comes in the form of building students' knowledge of the field using hands-on activities and engagement with multimodal/multilingual resources (e.g., video clips) as well as explicitly modeling text interpretation and production practices so students can notice and critically discuss genre expectations and field, tenor, and mode choices. In this way, teachers can build on students' everyday, or "spontaneous," ways of knowing and meaning-making to develop more disciplinary, or "scientific," ways of talking, reading, writing, and critically reflecting on subject matter knowledge (Vygotsky, 1986, p. 148).

This approach, known as the "Teaching and Learning Cycle" (TLC), provides a research-based approach to designing curriculum, instruction, and assessments using five pedagogical stages (for variations on the TLC, see Accurso and Mizell, 2020; de Oliveira and Smith, 2019; Cope and Kalantzis, 1993; Gebhard, 2019; Rose and Martin, 2012). As described by Derewianka and Jones (2016, p. 53), these stages require teachers actively:

Table 3 Algebra: genre and register analysis of mathematical explanation.

Genre stages	Text	Register features
Thesis using the language of the question Definition 1 Definition 2 Solve and explain Graph	I explain that this equation is a function by providing a definition of the equation's domain and range and graphing what it looks like. This helps me see if the equation passes the vertical line test. The domain is the set of all numbers that can be substituted for X in equation. X can be any real number, so the domain is all real numbers. The range is all the possible numbers that can be Y. Because x^2 is positive or at a minimum 0, x^2-5 could be -5 or greater. This means that the range is $Y > -5$. Therefore, the equation, $y = x^2-5$, is a function because, for every possible value for X, there is only one value of Y. One way to show that this is true is to graph the equation and use the vertical line test. If vertical lines intersect the graph at a single point, then the equation is a function, as shown below. 	Content/field Content-specific long noun groups (participants) - <i>A definition of the equation's domain and range</i> - <i>The set of all real numbers</i> Content-specific verbs (processes) - <i>Is</i> - <i>Substitute</i> - <i>Show</i> - <i>Intersect</i> Voice/tenor Declarative sentences and neutral language constructs authoritative facts Modal verbs possibility (<i>can be, could be</i>). Flow/mode Repetition of the word "equation," "domain," and "range" Pronouns and synonyms: <i>equation, function, it</i> Weaving given and new information regarding X: - The domain is the set of all numbers that can be substituted for X (new) in equation. X (given) can be any real number, so the domain is all real numbers. Cohesive devices that support explanations: <i>because, therefore, then</i> Connection between language and graph: - <i>As shown below ...</i> - Use of symbols

Prompt: Give the domain and range of the relationship. Then tell whether the relation is a function. Explain your answer. $Y = x^2-5$.

© 2020 Board of Regents of the University of Wisconsin on behalf of WIDA. Reproduced with permission.

1. Build students' background knowledge or field resources through multimedia and hands-on tasks that build students' shared understanding of a new topic.
2. Support students in reading a text or excerpts from texts to further build their knowledge and interest in a topic.
3. Deconstruct a target genre collaboratively with students to highlight, literally, the genre's purpose, audience, and stages. This phase also includes drawing attention to key register features as illustrated in the text annotations in Figure 3. [Derewianka and Jones \(2016, p. 53\)](#) note that students need several guided encounters with a specific genre, such as a mathematical explanation, before they will internalize the target genre and register features.
4. Jointly construct a text, such as a mathematical explanation, with students to make visible how to prepare, organize, and produce a disciplinary text that uses language and other meaning-making systems in specific ways. This approach to making literacy practice visible and open to discussion also involves discussing the literal and inferential meaning of words and critically discussing why certain lexicogrammatical choices are more strategic than others given the text's purpose and audience.
5. Support students in going through the same process to research a similar topic and produce a text of their own more independently with less scaffolding.

The TLC was originally developed to teach school based literacies in disadvantaged K-12 schools in Sydney Australia (e.g., [Christie and Derewianka, 2008](#); [Humphrey, 2017](#); [Rothery, 1996](#); [Rose and Martin, 2012](#)). Its use has since expanded to include the teaching of content based English instruction internationally (e.g., [McCabe, 2021](#)); and the teaching and learning a world language such as Arabic, French German and Spanish (e.g., [Abdel-Malek, 2020](#); [Colombi, 2009](#); [Byrnes et al., 2010](#); [Trojan et al., 2022](#)).

Critiques to SFL pedagogy and their responses

Despite promising research conducted internationally in elementary, secondary, and tertiary contexts, SFL pedagogy has been sharply critiqued for valorizing disciplinary literacy practices as "superior" to everyday literacy practices and thus contributing to

deficit perspectives of students' home and community semiotic resources (Thompson and Watkins, 2021, pp. 14–15; see also Flores, 2020; Flores and Rosa, 2015; Luke, 2018). Others, such as Bawarshi and Reiff (2010) cite concerns about the “degree of formalism” and lack of creativity exhibited by an approach to teaching in which genres are “classified and then modeled to students as though they were given” (p. 35). Johns (2008, p. 245) echoes this concern by describing how novices need to acquire genre schemas as a pathway toward more expert textual practices, but also need to be able to revise these schemas to meet the demands of a specific situation given that genres are evolving and must be reformulated for shifting situations. Others, such as Bourke (2005), capture the pervasive impression that SFL's metalanguage is too technical and “messy” to gain widespread application in teacher education. And last, Luke (2018) contributes an additional warning related to claims of empowerment and social change through the use of SFL teaching and learning practices. Drawing on Bourdieu's formulation of language and power, he writes that the “possession” of disciplinary genre and register knowledge, or what he refers to as “symbolic and cultural capital,” is “neither necessary nor sufficient for economic and social power” (Luke, 2019, p. 160). Rather “institutional racism,” as well as classism, sexism, and other ideologies related to the politics of difference within classrooms, schools, and society are apt to prevent historically marginalized populations, such as people of color, women, and immigrants, from ultimately “converting” disciplinary literacies learned in school into “economic capital” outside of schools (Luke, 2019, p. 162). Luke concludes his critique by stating that the move toward explicitly teaching high stakes genres has been valuable, but there is a “danger” if we:

... fail to situate, critique, interrogate, and transform these texts, their discourses, and their institutional sites—coming full circle back to enshrining, reproducing, and making invisible their bases in conflict, power, and difference. To do so is to place these genres, their teaching and reproduction in classrooms, workplaces, and bureaucracies beyond criticism. It is to present them as essential and compellingly functional, but not political and ideological. Without reconsidering its own social and cultural consequences, genre teaching runs the risk of becoming an institutional technology principally engaged in self-reproduction of the status and privilege of disciplinary knowledge, rather than part of a broadly-based political project for remaking the institutional distribution of literacy and its affiliated forms of capital.

Luke (2018, pp. 163–164).

In response to these critiques, SFL scholars and teacher educators have formulated and reformulated their theoretical models and approaches to pedagogy. For example, Hasan (2003), an early influencer in the development of SFL and SFL pedagogy, calls for the use of reflection literacy. She maintains that this type of reading and writing instruction requires students to develop the abilities to not only decode and interpret texts through a close analysis of genre and register features but also to reflect on the social significance of how language and other meaning-making systems construct knowledge. She writes, “the literate person should be able to interrogate the wording and the meaning of any utterance” to determine whose perspective is presented, whose perspectives are excluded, and how different ways of constructing knowledge impact the world we live in (2003, p. 447). Likewise, Martin (2000) stresses that enacting an SFL-informed approach to literacy development cannot be anything but “ideologically committed” (p. 116) and Christie highlights that Halliday's theories, informed by Marxism, have always focused on the issue of equity (Christie and Martin, 2009). In reviewing a “critical SFL praxis approach to teacher education,” Troyan et al. (2020) draw attention to the influence of Marxism in the formulation of SFL. Halliday writes the purpose of developing SFL was:

To give value to the language of the “other” non-European languages, unwritten vernaculars, non-native varieties of English, non-standard dialects... this motif has remained constant, as the main ideological input to what evolved as systemic theory. The theory has never been neutral.

Halliday (1993, p. 74), as cited in Troyan et al. (2020).

In the US, an explicitly critical instantiation of SFL pedagogy, one that serves the dual purpose of scaffolding students' disciplinary literacy development and their ability to detect and challenge specific ideologies, is evident in some but certainly not all studies of SFL teaching practices. For example, in a comprehensive review of 136 published articles regarding SFL research conducted in the US between 2000 and 2019, Accurso and Gebhard (2020) found that only 43% of teacher professional development (PD) programs reported gains in the area of critical language awareness. These programs were designed to guide teachers in understanding how semiotic choices construct subjectivities and ideologies in the texts they encounter in school. These studies demonstrated that after participating in SFL-informed PD, teachers approached teaching disciplinary texts differently and began to view their students' literacy practices from a resource as opposed to a deficit perspective. For example, several studies conducted with social studies and science teachers showed that when teachers were guided to question texts commonly presented as factual, they began to denaturalize dominant assumptions about language use in schools (e.g., informational texts are facts devoid of author attitude, opinion, or ideological stance; see Achugar et al., 2007; Achugar and Carpenter, 2012; Carpenter et al., 2014; Gebhard et al., 2014; O'Halloran et al., 2015; Ramirez, 2014; Schleppegrell et al., 2008). However, these studies did not show these lines of critical questioning extending beyond individual sets of disciplinary texts to examine how broader school practices and societal ideologies influence semiotic choices across texts.

A second much smaller group of studies, also reviewed by Accurso and Gebhard (2020), demonstrated more engagement with macro-level critiques of schooling policies and practices. For example, Gebhard et al. (2008) documented two first-grade teachers' experiences analyzing text-context dynamics. As part of their master's degree program, these teachers analyzed the literacy practices of a first-grade bilingual student who had been characterized as “mixing Spanish and English” in ways that might “interfere” with

her academic progress (p. 280). However, by analyzing this student's writing to understand the influence of historical, political, and economic aspects of the context in which it was produced, the teachers were able to see how school reforms related to English only mandates and high stakes testing shaped this student's opportunities for disciplinary literacy development. Further, they were able to see how this student drew on a wide variety of out-of-school semiotic resources in completing school-based tasks.

Similarly, Harman (2007) documented how one eighth-grade teacher experienced a perspective shift throughout her teacher education program experience from a deficit orientation toward students labeled "English learners" to "affording them multiple literate identities" (p. 31). Before participating in an extended professional development program, the focal teacher admitted to seeing some of her students and their work as problematic because they spoke Spanish. However, as this teacher analyzed multilingual students' texts in relation to her own teaching practices and school policies, she began to notice and question her own monolingual orientation and looked for new ways to design multilingual/multicultural curriculum and instruction.

Collectively, Gebhard (2019) and Harman's (2018) work signal a critical shift in the use of SFL pedagogy in US teacher education and classroom research. Both focus on a critical praxis perspective and language, learning, and social change in teacher education. Gebhard's work is influenced by Fairclough's (1992) conception of critical discourse analysis to inform debates regarding institutional discourses of public education and Dyson's (1993, p. 28) concept of designing a "dialogic" or "permeable curriculum" to create inroads to using disciplinary discourses that draw on historically marginalized students' home and community linguistic resources (Gebhard, 2019). Harman's work is more inspired by Paris and Alim's (2017) "culturally sustaining pedagogy" and New London Group's (1996) construct of multilingual and multimodal meaning making (e.g., Harman, 2018; Harman and Burke, 2020; Khote and Tian, 2019).

Other US scholars have also expanded SFL-informed conceptual frameworks and related pedagogical practices to include a stronger focus on SFL within the following programs and conceptual domains:

- Dual language programs (e.g., de Oliveira and Westerlund, 2021; de Oliveira and Smith, 2019; Ramirez et al., 2018)
- Culturally sustaining family literacy programs (e.g., Ramirez, 2020)
- Heritage language learners (e.g., Achugar and Colombi, 2008; Colombi, 2009, 2015)
- Critical race theory and raciolinguistics (e.g., Accurso and Mizell, 2020; Mizell, 2021; Thomas, 2013)
- Multimodality (e.g., Accurso et al., 2019; Shin et al., 2020; Harman and Burke, 2020)
- Teacher action research and critical language awareness (e.g., Gebhard and Graham, 2018)
- State and district policies and practices influencing the education historically marginalized bilingual learners (e.g., Turriago, 2020)

Future directions for SFL scholarship in the US: personal reflections

These recent studies respond to critiques of SFL and call for scholars, teacher educators, and practitioners to address systemic inequities related to language, language variation, and the construction of race-, class-, and gender-based inequities in US schools. Some of these studies do so in ways that productively couple Halliday's theory of SFL with other critical sociocultural perspectives of language, learning, and power in schools as ideologically driven institutions. For example, the pairing of SFL with multimodality (Kress and van Leeuwen, 2001), translanguaging (García and Wei, 2014), culturally sustaining pedagogy (Paris and Alim, 2017), critical language awareness (Janks and Locke, 2008), and critical race theory (Delgado and Stefancic, 2017) has been fruitful, especially in the context of the Black Lives Matter Movement and other structural inequities made even more visible by the Covid-19 pandemic (Gebhard and Accurso, in press).

As SFL research and SFL-informed PD initiatives emerge to respond to these inequities, we suggest that connections with commensurate theories will have more traction if researchers and teacher educators strive for conceptual and methodological clarity regarding what they mean by language, learning, and social change. This clarity is essential if SFL-informed practices are to avoid the slippage into more conservative and behavioral approaches to teaching and learning that happens all too often in US schools as they strive for efficiency and standardization (Callahan, 1962; Gebhard, 2019; Trujillo, 2014). In addition, this recommendation will also enable researchers and practitioners to be clearer about what their disagreements are. Are their disagreements related to different conceptions of grammar (e.g., behavioral, psycholinguistics, or social semiotic perspectives)? Different perspectives regarding the nature of learning (e.g., behavioral, psychological, sociocultural)? And/or different conceptions of power and what social justice looks and sounds like in schools (assimilationist vs. multicultural/multilingual or structural vs. post-structural)?

In addition, despite many mission statements using the words "bilingual," "equity," or "social justice," the K-12 schools and licensing programs that we have worked in, driven by ever-expanding sets of standards, high stakes testing practices, and diminishing resources for sustained engagement with complex topics, often revert to assimilationist and behavioral perspectives of language and teaching that support English monolingualism, white supremacy, and middle-class norms. Within these contexts, we notice that language is still treated as a decontextualized and neutral set of formal rules regarding sentence structure with less attention to the function of linguistic variation (e.g., African American English), register features, genre stages, and students' investments in learning to read, write, and discuss disciplinary topics for authentic social, academic, and political purposes. In these cases, under-resourced schools and teacher licensing programs often lack the time and capacity for authentic and sustained critical reflection as advocated for by Luke (2019). In moving forward, we recommend that future studies of SFL in the US attend more seriously to the institutional context in which teachers learn and work. These contexts place many demands on teachers, demands that

compromise their ability to use SFL and act as transformative educators capable of supporting multilingualism, setting high expectations, providing high supports, and fostering critical reflections. Therefore, we recommend that more attention be paid not just to conceptions of language, learning, and social change, but also to how teachers are supported in using systemic functional linguistics over time.

References

- Abdel-Malek, M., 2020. Empowering Arabic learners to make meaning: a genre-based approach. *System* 94. <https://doi.org/10.1016/j.system.2020.102329>.
- Accurso, K., Gebhard, M., 2020. SFL praxis in U.S. teacher education: a critical literature review. *Lang. Educ.* 1–21.
- Accurso, K., Mizell, J.D., 2020. Toward an antiracist genre pedagogy: considerations for a North American context. *TESOL J.* 11 (4). <https://doi-org.ezproxy.neu.edu/10.1002/tesj.554>.
- Accurso, K., Gebhard, M., Purington, S.B., 2017. Analyzing diverse learners' writing in mathematics: systemic functional linguistics in secondary pre-service teacher education. *Int. J. Math. Teach. Learn.* 18 (1), 84–108.
- Accurso, K., Muzeta, B., Pérez Battles, S., 2019. Reflection multiliteracies: teaching meaning making across the visual and language arts. *SPELT Q. J.* 34 (2), 2–16.
- Achugar, M., Carpenter, B., 2012. Developing disciplinary literacy in a multilingual history classroom. *Ling. Educ.* 23, 262–276.
- Achugar, M., Colombi, M.C., 2008. Systemic functional linguistic explorations into the longitudinal study of advanced capacities: the case of Spanish heritage language learners. In: Ortega, L., Byrnes, H. (Eds.), *The Longitudinal Study of Advanced L2 Capacities*. Routledge, London, pp. 52–73.
- Achugar, M., Schleppegrell, M., Oteiza, T., 2007. Engaging teachers in language analysis: a functional linguistics approach to reflective literacy. *Engl. Teach. Pract. Critiq.* 6.
- Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B., 2019. *Theoretical Models and Processes of Literacy*. Routledge, New York.
- Bartlett, T., O'Grady, G., 2017. *The Routledge Handbook of Systemic Functional Linguistics*. Routledge, London.
- Bawarshi, A.S., Reiff, M.J., 2010. *Genre: An Introduction to History, Theory, Research, and Pedagogy*. Parlor Press, West Lafayette, Indiana.
- Bourke, J., 2005. The grammar we teach. *Engl. Lang. Teach.* 4, 85–97.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Basic Books, New York.
- Brisk, M.E., Schleppegrell, M., 2021. *Language in Action: SFL Theory Across Contexts*. Equinox, Sheffield.
- Byrnes, H., Maxim, H.H., Norris, J.M., 2010. Realizing advanced foreign language writing development in collegiate education: curricular design, pedagogy, assessment. *Mod. Lang. J.* 94, 1–235.
- Calkins, L.M.C., 1994. *The Art of Teaching Reading*. Heinemann, Portsmouth.
- Callahan, R.E., 1962. *Education and the Cult of Efficiency*. University of Chicago Press, Chicago.
- Carpenter, B., Earhart, M., Achugar, M., 2014. Working with documents to develop disciplinary literacy in the multilingual classroom. *Hist. Teach.* 48 (1), 91–103.
- Chomsky, N., 1967. A review of B.F. Skinner's verbal behavior. In: Jakobovits, L., Miron, M.S. (Eds.), *Readings in the Psychology of Language*. Prentice-Hall, Englewood Cliffs, N.J., pp. 142–143.
- Christie, F., Derewianka, B., 2008. *School Discourse: Learning to Write Across the Years of Schooling*. Continuum, London.
- Christie, F., Martin, J.R., 1997. *Genre and Institutions: Social Processes in the Workplace and School*. Continuum, London.
- Christie, F., Martin, J.R., 2009. *Language, Knowledge and Pedagogy: Functional Linguistic and Sociological Perspectives*. Continuum, London.
- Colombi, M.C., 2009. A systemic functional approach to teaching Spanish for heritage speakers in the United States. *Ling. Educ.* 20 (1), 39–49.
- Colombi, M.C., 2015. Academic and cultural literacy for heritage speakers of Spanish: a case study of Latin@ students in California. *Ling. Educ.* 32, 5–15.
- Cope, B., Kalantzis, M., 1993. *The Powers of Literacy: A Genre Approach to Teaching Writing*. University of Pittsburgh Press, Pittsburgh.
- de Oliveira, L.C., Smith, B., 2019. *Expanding Literacy Practices Across Multiple Modes and Languages for Multilingual Students*. Information Age Publishing.
- de Oliveira, L.C., Westerlund, R., 2021. A functional approach to language development for dual language learners. *J. English Learner Educ.* 12, 1.
- Delgado, R., Stefancic, J., 2017. *Critical Race Theory*. NYU Press, New York.
- Delpit, L., 1988. The silenced dialogue: power and pedagogy in educating other people's children. *Harv. Educ. Rev.* 58, 280–299.
- Derewianka, B., Jones, P., 2016. *Teaching Language in Context*. Oxford University Press, South Melbourne.
- Dyson, A.H., 1993. *Social Worlds of Children Learning to Write in an Urban Primary School*. Teachers College Press, New York.
- Elbow, P., 1973. *Writing Without Teachers*. Oxford University Press, New York.
- Fairclough, N., 1989. *Language and Power*. Longman, London.
- Fairclough, N., 1992. *Discourse and Social Change*. Polity Press, Cambridge.
- Flores, N., Rosa, J., 2015. Undoing appropriateness: raciolinguistic ideologies and language diversity in education. *Harv. Educ. Rev.* 85, 149–171.
- Flores, N., 2020. From academic language to language architecture: challenging raciolinguistic ideologies in research and practice. *Theory Into Pract.* 59, 22–31.
- García, O., Wei, L., 2014. *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan, Basingstoke.
- Gebhard, M., Accurso, K. (in press). *Teacher Action Research Across the Curriculum: An SFL Approach to Supporting Diverse Learners' Disciplinary Literacies in Hard Times*. New York: Routledge.
- Gebhard, M., Graham, H., 2018. Bats and grammar: developing critical language awareness in the context of school reform. *Engl. Teach. Pract. Critiq.* 17 (4), 281–297.
- Gebhard, M., Demers, J., Castillo-Rosenthal, Z., 2008. Teachers as critical text analysts: L2 literacies and teachers' work in the context of high-stakes school reform. *J. Sec. Lang. Writ.* 17 (4), 274–291.
- Gebhard, M., Chen, I., Britton, B., 2014. "Miss, nominalization is a nominalization": English language learners' use of SFL metalanguage and their literacy practices. *Ling. Educ.* 26, 106–125.
- Gebhard, M., 1999. Debates in SLA studies: redefining classroom SLA as an institutional phenomenon. *Tesol Q.* 36 (3), 544–557.
- Gebhard, M., 2019. *Teaching and Researching ELLs' Disciplinary Literacies*. Routledge, New York.
- Halliday, M.A.K., Hasan, R., 1985. *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Deakin University Press, Deakin, Victoria.
- Halliday, M.A.K., Matthiessen, C., 2014. *Halliday's Introduction to Functional Grammar*. Routledge, Oxon.
- Halliday, M.A.K., 1973. *Explorations in the Functions of Language*. Elsevier, New York.
- Halliday, M.A.K., 1978. *Language as Social Semiotic*. Edward Arnold, London.
- Halliday, M.A.K., 1993. Towards a language-based theory of learning. *Ling. Educ.* 5 (2), 93–116.
- Hammond, J., Gibbons, P.F., 2005. *Putting Scaffolding to Work: The Contribution of Scaffolding in Articulating ESL Education*. National Centre For English Language Teaching and Research (NCELTR).
- Harman, R., Burke, K., 2020. *Culturally Sustaining SFL Praxis: Embodied Inquiry With Multilingual Youth*. Routledge, New York.
- Harman, R., 2007. Critical teacher education in urban contexts: discursive dance of a middle school teacher. *Lang. Educ.* 21 (1), 31–45.
- Harman, R. (Ed.), 2018. *Bilingual Learners and Social Equity: Critical Take(s) on Systemic Functional Linguistics*. Springer, London.
- Hasan, R., Williams, G., 1996. *Literacy in Society*. Longman, London.
- Hasan, R., 2003. Globalization, literacy and ideology. *World Engl.* 22, 433–448.

- Howard, T.C., 2019. *Why Race and Culture Matter in Schools: Closing the Achievement Gap in America's Classrooms*. Teachers College Press, New York.
- Humphrey, S., 2017. *Academic Literacies in the Middle Years: A Framework for Enhancing Teacher Knowledge and Student Achievement*. Routledge, New York.
- Janks, H., Locke, T., 2008. Discourse awareness in education: a critical perspective. *Encycl. Lang. Educ.* 2, 31–43.
- Janks, H., 2010. *Literacy and Power*. Routledge, New York.
- Johns, A., 2008. Genre awareness for the novice academic student: an ongoing quest. *Lang. Teach.* 41 (2), 237–252.
- Kaplan, R., 1966. Cultural thought patterns in inter-cultural education. *Lang. Learn.* 16, 1–20.
- Khote, N., Tian, Z., 2019. Translanguaging in culturally sustaining systemic functional linguistics: developing a heteroglossic space with multilingual learners. *Transl. Translanguaging Multiling. Contexts* 5 (1), 5–28.
- Krashen, S.D., 1987. *Principles and Practice in Second Language Acquisition*. Prentice-Hall International, Englewood Cliffs, N.J.
- Kress, G., van Leeuwen, T., 2001. *Multimodal Discourse*. Arnold, London.
- Labov, W., Waletzky, J., 1967. Narrative analysis. In: Helm, J. (Ed.), *Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*. University of Washington Press, Seattle.
- Ladson-Billings, G., 2006. From the achievement gap to the education debt: understanding achievement in US schools. *Educ. Res.* 35 (7), 3–12.
- Lemke, J.L., 1990. *Talking Science: Language, Learning, and Values*. Ablex Publishing Corporation, Norwood, NJ.
- Luke, A., 2018. *Critical Literacy, Schooling, and Social Justice: The Selected Works of Allan Luke*. Routledge, New York.
- Luke, A., 2019. *Educational Policy, Narrative and Discourse*. Routledge, New York.
- Martin, J.R., Rose, D., 2008. *Genre Relations: Mapping Culture*. Equinox, London.
- Martin, J.R., 1992. *English Text: System and Structure*. Benjamins, Philadelphia.
- Martin, J.R., 2000. Design and practice: enacting functional linguistics. *Annu. Rev. Appl. Ling.* 20, 116–126.
- McCabe, A., 2021. *A Functional Linguistic Perspective on Developing Language*. Routledge, London.
- Mizell, J.D., 2021. Apprenticeship of pre-service teachers through culturally sustaining systemic functional linguistics. *Lang. Educ.* 35 (2), 123–139.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- O'Halloran, K.L., Palincsar, A.S., Schleppgrell, M.J., 2015. Reading science: using systemic functional linguistics to support critical language awareness. *Ling. Educ.* 32, 55–67.
- O'Halloran, K.L., 2007. Mathematical and scientific forms of knowledge: a systemic functional multimodal grammatical approach. In: Christie, F., Martin, J.R. (Eds.), *Language, Knowledge, and Pedagogy: Functional Linguistic and Sociological Perspectives*. Continuum, London, pp. 205–236.
- Paris, D., Alim, H.S., 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York.
- Ramirez, A., Sembiente, S.F., de Oliveira, L.C., 2018. Translated science textbooks in dual language programs: a comparative English-Spanish functional linguistic analysis. *Biling. Res. J.* 41 (3), 298–311.
- Ramirez, J.A., 2014. Genre-based principles in a content-based English as a second language pull-out classroom. In: de Oliveira, L., Iddings, J. (Eds.), *Genre Pedagogy Across the Curriculum: Theory and Application in Classrooms and Contexts*. Equinox, Sheffield, pp. 55–74.
- Ramirez, A., 2020. The case for culturally and linguistically relevant pedagogy: bilingual reading to learn for Spanish-speaking immigrant mothers. *System* 95.
- Rose, D., Martin, J.R., 2012. Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School. Equinox, Sheffield.
- Rothery, J., 1996. Making changes: developing an educational linguistics. In: Hasan, R., Williams, G. (Eds.), *Literacy in Society*. Longman, pp. 86–123.
- Schleppgrell, M., Greer, S., Taylor, S., 2008. Literacy in history: language and meaning. *Aust. J. Lang. Literacy* 31.
- Schleppgrell, M., 2004. *The Language of Schooling: A Functional Linguistics Perspective*. Routledge, New York.
- Shin, D.S., Cimasko, T., Yi, Y., 2020. Development of metalanguage for multimodal composing: a case study of an L2 writer's design of multimedia texts. *J. Sec Lang. Writ.* 47.
- Skinner, B.F., 1957. *Verbal Behavior*. Appleton-Century-Crofts, New York.
- Thomas, E., 2013. Dilemmatic conversations: some challenges of culturally responsive discourses in a high school English class. *Ling. Educ.* 24, 328–347.
- Thompson, G.A., Watkins, K., 2021. Academic language: is this really (functionally) necessary? *Lang. Educ.* 1–17.
- Troyan, F.J., Harman, R., Zhang, X., 2020. Critical SFL praxis in teacher education: insights from Australian SFL scholars. *Lang. Educ.* 383–401.
- Troyan, F.J., Herazo, J.D., Ryshina-Pankova, M., 2022. SFL pedagogies in language education: special issue introduction. *System* 104. <https://doi-org.ezproxy.neu.edu/10.1016/j.system.2021.102694>.
- Trujillo, T., 2014. The modern cult of efficiency: intermediary organizations and the new scientific management. *Educ. Pol.* 28 (2), 207–232.
- Turiago, N., 2020. Taking Another LOOK: Analyzing Discourses of Bilingual Education in Massachusetts as Revealed in the LOOK Act Press Releases. Graduate North American SFL Association Conference: University of Michigan.
- Unsworth, L., 2001. *Teaching Multiliteracies Across the Curriculum: Changing Contexts of Text and Image in Classroom Practice*. Open University, Maidenhead, Berkshire [England].
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge.
- Vygotsky, L.S., 1986. *Thought and Language*. MIT Press, Cambridge.
- Walqui, A., Bunch, G.C., 2019. *Amplifying the Curriculum: Designing Quality Learning Opportunities for English Learners*. Teachers College Press, New York.
- Westerlund, R., Besser, S., 2021. Reconsidering Calkins' Process Writing Pedagogy for Multilingual Learners: Units of Study in a Fourth-Grade Classroom. Wisconsin Center for Education Research. Working Paper No. 2021-4.
- Young, L., Harrison, C., 2004. *Systemic Functional Linguistics and Critical Discourse Analysis: Studies in Social Change*. Continuum, London.

Literacies, language, and schooling: exploring writing pedagogy for English language learners

Erin Kim-Rich^a and Jen Scott Curwood^b, ^a The University of New England, Armidale NSW Australia; and ^b The University of Sydney, Camperdown NSW Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	65
Writing pedagogy and English language learners	66
Writing and learning across disciplines	66
Writing pedagogy: from promise to practice	66
English language learners' writing	67
Supporting English language learners in secondary schools: Australian policy and practice	67
Australia's EAL/D policy and context	68
The role of grammar in writing	69
The teaching and learning cycle	69
Creative writing and spoken word poetry in Australian schools: a case study	70
Spoken word in context: real talk in Australian schools	71
Multilingualism, multiculturalism, and meaning making	71
Moving forward: developing English language learners' writing practices	72
Acknowledgments	73
References	73

Introduction

Over the last two decades, some progress has been made in improving the literacy achievement of young adults around the world. However, one in four still do not meet a minimum level of proficiency in reading (Organisation for Economic Cooperation and Development, 2019). Writing is an extremely complex skill (Hayes, 2012) that is inextricably related to reading (Fitzgerald and Shanahan, 2000), as they are mutually supportive of each other. This raises concerns that many students do not possess the reading and writing skills that will allow them to meet the demands of secondary school.

Prior research suggests that students who struggle with writing are less likely to use writing as a tool to support and extend their learning across all disciplines, including science, math, and social studies (Bangert-Drowns et al., 2004; Graham and Hebert, 2011). This is especially problematic because writing serves as one of the primary means for students to convey their thoughts and ideas, as well as for teachers to assess their learning and understanding (Graham, 2006). Even as young people have more tools than ever to convey their knowledge and beliefs, there is an increasingly narrow range of writing practices that are used to indicate literacy achievement. If students are unable to use writing as a way to represent their disciplinary knowledge, then it may be more difficult for them to graduate from secondary school, pursue further education, and prepare for the workforce. This is compounded for youth who are learning English as an additional language.

English language learners represent the fastest growing segment of the student population in many countries. Notably, the percentage of primary and secondary students in government schools who are English language learners (ELLs) is now at 10% in the United States (United States Department of Education, 2019), 19% in England (Department for Education, 2019), and 25% in Australia (Australian Curriculum Assessment and Reporting Authority (ACARA) 2020a). Compared to other literacy skills such as reading, speaking, and listening, writing is often the most difficult skill for ELLs (Cumming, 2013; Snow and Uccelli, 2009).

ELLs who have already learned to write in another language often have disciplinary knowledge and literacy skills that can facilitate their writing in English, but they still experience challenges. For instance, the differences in cultural norms, rhetorical styles, and academic discourses may lead to students' resistance to and difficulty with writing tasks (McCarthy and Garcia, 2005). In this chapter, we explore writing pedagogy for English language learners. We begin with a review of the international literature to highlight evidenced-based elements of writing instruction and to explore how ELLs' linguistic and cultural resources shape their writing development. Then, we turn our attention to policy and pedagogy in the Australian context, offering insight into how one country has strived to meet the needs of the rapidly growing number of English language learners. Lastly, we draw on our own research in Australian secondary schools to show how creative writing through spoken word can support the literacy and identity development of ELLs, which in turn impacts ELLs' writing engagement and achievement.

Writing pedagogy and English language learners

There is global concern for students' writing development (Graham and Rijlaarsdam, 2016). Writing has been described as a "complex of interconnected systems" (Nelson and Nelson, 1978, p. 278) and learning how to write requires substantial time and effective instruction (Graham et al., 2012). In schools around the world, students' knowledge and understanding are commonly assessed through writing. However, assessment-driven writing instruction typically tends to "narrow views of writing to the mastery of widely-used processes in order to create a product that conforms to recognisable standards" (Peterson et al., 2007, p. 30). Research suggests that writing pedagogy needs to reflect the "richness and complexity of literacy practices evident in society at large" (Street, 2005, p. 4) rather than conceptualize writing as a discrete skill that students must master and perform on demand. This literature review will consider why writing is important to students' learning, how teachers can effectively teach writing in secondary schools, and how English language learners, in particular, can become capable and effective writers.

Writing and learning across disciplines

Writing allows students to learn new concepts, express ideas, create fictional worlds, and engage in meaning making. For that reason, writing is instrumental to students' learning, especially when it involves authentic writing tasks and real audiences (Graham and Harris, 1997). Writing also enhances students' literacy skills more broadly. When writing is included as part of reading instruction, for instance, it improves students' ability to decode and comprehend information (Graham et al., 2018). The nature of literacy means that writing, reading, speaking, listening, viewing, and representing can develop in mutually constitutive ways. For this reason, this chapter, whilst focusing on the mode of writing, explores the reciprocal relationship across the modes to support writing development.

Being able to write enhances students' performance in multiple disciplines, and students comprehend and retain material better in science, social studies, and mathematics when they write about it (Graham et al., 2011). Research suggests that simply increasing the frequency of writing will not enhance students' knowledge and understanding about a specific topic (Applebee, 1984); rather, factors such as the nature of the writing task and teachers' level of preparation and confidence can substantially impact the influence of writing on learning (Bangert-Drowns et al., 2004; Gillespie et al., 2014). A meta-analysis by Graham and Perin (2007) examined 26 studies on writing to learn, that included students from upper primary through secondary schools. Their work found that writing to learn has a positive impact and 73% of studies demonstrated that writing can improve student learning in language arts, social studies, and science. Another meta-analysis by Graham and Hebert (2011) included 66 writing to learn studies and focused on the impact of writing on understanding information. They found that 94% of the studies showed that writing about material read enhances students' comprehension of texts across disciplines. Whilst the meta-analysis was conducted across the general school population, Graham and Hebert (2011) note that the reading performance of ELLs is often statistically lower than their native English speaking peers but that "writing about reading had a positive impact on the comprehension of weaker readers/writers" (p. 728). Furthermore, the effect is stronger, particularly for high school students, when students are explicitly taught how to write about their reading (Graham and Hebert, 2011).

Writing pedagogy: from promise to practice

While writing can substantially support student learning, international scholarship has shown that writing instruction in most classrooms is not sufficient. A recent literature review synthesized findings from 28 studies involving the writing practices of more than 7000 teachers in the United States, Europe, Asia, South America, and Australasia. Within this, Graham (2019) argues that these global studies of writing found that "some teachers provide students with a solid writing program, and in some classrooms this instruction is exemplary" but "this is not typically the case, as writing and writing instruction in most classrooms are inadequate" (p. 279). Notably, these findings were relevant across countries and grades.

While most secondary teachers believe that writing can facilitate student learning, this is not necessarily reflected in their pedagogy and they may not dedicate time and attention to using writing to support learning (Applebee and Langer, 2011). Despite the importance of writing for students' learning in all disciplines, it is more likely to be taught in language arts and most other classes relegate writing to activities such as filling in the blanks or copying, rather than composing substantial texts (Gillespie et al., 2014; Ray et al., 2016). A study by Ray et al. (2016) found that writing short answer responses, taking notes for reading or while listening, and completing worksheets were the most frequent writing strategies reported by secondary teachers; on average, these strategies were used once a week. Writing disciplinary-specific texts like a literary analysis or a lab report were reportedly used less than once per year.

As a whole, secondary teachers directly teach less than half of the writing to learn strategies that they asked their students to apply (Ray et al., 2016). This is problematic because writing like a scientist is different from writing like a historian - they involve making meaning from different kinds of sources, writing in different genres, and writing for different audiences and purposes. Moreover, students' writing in one genre is not highly predictive of their writing in another (Graham et al., 2011). This reflects the tendency to treat disciplines as "static corpuses of fact" and not "dynamic sets of discourses and practices, literacies and knowledges" (Luke, 2000, p. 139). A study by Gillespie et al. (2014) found that secondary teachers are more likely to implement writing to learn strategies in the classroom if they have more preparation, greater confidence in their writing pedagogy, and more diverse classrooms.

Gadd and Parr's (2017) and Ortmeier-Hooper's (2013) research further highlighted that effective teachers of writing emphasize task selection, direct instruction, and building self-efficacy and self-regulation.

English language learners' writing

The concept of writing varies across countries, and even when tasks are well-defined, cultural differences and expectations impact both students and teachers (Purves, 1992). For the growing number of English language learners, the development of higher-level reading and writing skills is essential for them to navigate the complexity of academic English (Scarcella, 2002). Since it may take ELLs a number of years to develop grade appropriate reading and writing proficiency in English, it is essential that schools offer ongoing and differentiated literacy instruction across disciplines. Cumming (2015) synthesizes recent research and argues that there are nine interrelated approaches to support ELLs' writing development, including "regular routines to scaffold writing tasks; strategy instruction and practice; collaborative writing; content-based teaching; development of metalanguage, genres, and vocabulary; teacher development; supporting literacies outside of school; dual-language programs; and tutoring" (p. 366). Many of these approaches can readily be integrated into disciplinary classrooms to support ELLs' learning broadly as well as their writing for specific audiences and purposes.

Writing pedagogy for ELLs needs to account for their language background in ways that are fundamentally different from native English speakers of the same age. First, ELLs are often learning to speak and write English concomitantly, which impacts how teachers understand their content knowledge and literacy skills. Gibbons (2015) explains that many countries tend to "assess ELLs as though they have identical language resources to English mother tongue speakers... [and] fail to recognize what ELLs learners actually do achieve in their learning of English" (p. 98). Second, assessments also rarely acknowledge that many ELLs are literate in another language, or reflect the fact that some ELLs may also be learning a new script. Cumming (2015) argues that a more productive approach would involve creating assessments for ELLs based on normative trajectories for their literacy development. Alternative approaches to assessment are necessary so that ELLs' full language resources and literacy skills can be accounted for in school-based assessments.

In conceptualizing English language learners' abilities, teachers may unintentionally perpetuate a deficit discourse, which "locates its explanation of the underperformance or underachievement of non-dominant students in the nonalignment of the cultural practices of the home and school" (Gutierrez, Zitali Morales, and Martinez, 2009, p. 218). Alford (2014) argues that ELLs need to be able to utilize a wide range of resources and tools to express their ideas and understandings, rather than being forced to write analytically and sophisticatedly in a language they have not mastered. Janks (2010) adds that young people need access to school-based literacies, including the experience of writing for academic purposes, in a range of genres, and for different audiences. For ELLs, this can be challenging if they "have no experience of meaningful, pleasurable, fluent writing on which to build" (Janks, 2010, p. 158). For that reason, secondary teachers must strive to understand the background knowledge and literacy resources that ELLs bring to disciplinary learning, and consider how to design instruction and assessment that can account for their developing understanding of English language, grammar, and syntax.

Writing pedagogy needs to "value literacy learners' funds of knowledge and the ways in which they can inform literacy teaching" (McLean et al., 2009, p. 169). ELLs bring many linguistic and cultural resources with them into the classroom, and writing pedagogy must reflect that. For instance, Hornberger and Link (2012) argue that translanguaging is becoming more recognized in educational contexts, as it values how multilingual students make meaning "by drawing on and intermingling linguistic features from different languages" (p. 240). Research by Velasco and Garcia (2014) shows that translanguaging can effectively be used by ELLs at all stages of the writing process, including planning, editing, and production, as well as to promote self-regulation. Similarly, because poetry encourages the exploration of diverse linguistic and cultural competencies (Cahnmann, 2006), it can be considered as a genre of access for English language learners (Friedman et al., 2019). Poetry is personal, malleable, and authentic, while at the same time working within and against rules of language and structure. Educational policies and curriculum mandates in some countries recognize ELLs' funds of knowledge and language resources. For example, the Australian Curriculum Assessment and Reporting Authority (2020a) argues that ELLs' cultural and linguistic resources can support their English language learning and disciplinary knowledge while also being shared in the classroom for the benefit of all students. In the following section, we focus on how Australian policy and pedagogy can support English language learners' writing development in secondary school contexts.

Supporting English language learners in secondary schools: Australian policy and practice

Australia is one of the most culturally and linguistically rich countries in the world with students coming from over 2000 different ethnic backgrounds (ACARA, 2020a). With the introduction of Australia's national curriculum being progressively implemented since 2009 (ACARA, 2020b), the universal terminology in Australia for English Language Learners became English as an Additional Language/Dialect (EAL/D). This definition encompasses overseas and Australian-born students whose first language is a language other than English who require additional language support to develop Standard Australian English; students whose parents use Auslan as their first language; and the dialectical 'D' component is inclusive of Aboriginal and Torres Strait Islander students whose first language is a creole or a variety of Standard Australian English such as Aboriginal English (ACARA, 2020a; New South Wales Department of Education (DoE), 2020). Across primary (Years K – 6) and secondary (Years 7–12) school,

approximately a quarter of Australian students are learning English as an additional language or dialect (ACARA, 2020a). In the New South Wales (NSW) state government sector alone, more than a third of students come from language backgrounds other than English and nearly 60% are located in the Sydney metropolitan area (NSW Department of Education: Center for Education Statistics and Evaluation, 2019). Correspondingly, some schools may have an EAL/D student population as high as 90% (ACARA, 2020a). The following sections offer an overview of Australia's EAL/D frameworks and policies, an exploration of how grammar has shaped writing pedagogy and curriculum, and insight into how the teaching and learning cycle can support English language learners' writing.

Australia's EAL/D policy and context

Australia's EAL/D policies exist at federal and state levels, and are informed by the student demographics, local contexts and theoretical frameworks. At a national level, documents such as *The Melbourne Declaration on Educational Goals for Young Australians* (Ministerial Council on Education, Employment, Training and Youth Affairs, 2008) and *The Alice Springs (Mparntwe) Education Declaration* (Education Council, 2019) help underpin EAL/D education in schools through their goals of excellence and equity. The individual needs of all young Australians, regardless of background, are to be supported with students having access to quality education that encourages "high expectations" and is inclusive "free from any form of discrimination" (Education Council, 2019, p. 4). However, both the implementation of the Australian Curriculum and EAL/D support is the responsibility of state and territory government authorities (ACARA, 2020a), and as such, variation occurs.

In the state of NSW, for instance, for newly arrived secondary students (ages 12–18) whose first language is not English, they may attend an Intensive English Center which provides short-term targeted English language support for up to approximately one year (DoE, 2020b) before they move to mainstream secondary schools. The *NSW EAL/D Advice for Schools* (DoE, 2020a) which is an implementation guide for the *NSW Multicultural Education Policy* (2005), presents a range of second language acquisition theories to help contextualize general timeframes and language acquisition stages. For example, it includes Thomas and Collier's (1997) model of academic English where it takes approximately 4–7 years to develop academic English with EAL/D support if students have education and literacy in their first language; and Cummin's (1979, 1991) Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) model. Within the school context, BICS is framed as conversational or social language and takes, on average, up to two years to acquire. CALP, on the other hand, is considered the academic language needed for success at school and takes a beginning EAL/D student 5–7 years to develop and up to approximately 11 years if they have had disrupted schooling (DoE, 2020a, pp. 19–21). Given these approximate timeframes, the reality is that the majority of students needing English language support are in mainstream secondary schools.

Mainstreaming in Australia is where grade level curriculum is accessed and demonstrated through English as the medium of instruction and "each area of the curriculum has language structures and vocabulary particular to its learning domain, and these are best taught in the context in which they are used" (ACARA, 2014, p. 6). However, research in English speaking countries have reported that mainstream teachers are often ill-equipped and under-prepared to teach English language learners (e.g., Cajkler and Hall, 2009; Cruickshank, 2009; Cummins et al., 2012; Franson, 1999; Harper and de Jong, 2009; Leung and Franson, 2001; Rushton, 2008; Sinclair and Lau, 2018; Turkan and Buzick, 2016). How to prepare mainstream teachers to embed linguistic, cognitive, sociocultural and contextual knowledge into content instruction in order to meet the needs of EAL/D students is a complex task with no clear agreement on best practice (Costley, 2014; DelliCarpini and Gulla, 2016; Harper et al., 2010; Janzen, 2008; Villegas, Saiz de La Mora, Martin, and Mills, 2018; Von Esch and Kavanagh, 2018).

With the present EAL/D legislation in place at both national and state levels in Australia, there is a lack of consistency across jurisdictions as EAL/D provisions are determined largely based on school context. For example, the NSW EAL/D system has shifted from state-based infrastructure to *Local Schools, Local Decisions* (DoE, 2014) with a student-based funding model. The responsibility of supporting EAL/D student needs is now determined by principals who have the flexibility in decision-making and budget expenditure (DoE, 2014). The decentralization of the EAL/D services has raised concerns of accountability over EAL/D funding and programming (Australian Council of TESOL Association (ACTA), 2017, 2018; Joint Standing Committee on Migration, 2017; NSW ESL and Refugee Education Working Party, 2013). The concern lies in the lack of transparency as it hinges on what the principal determines as the "local need." This in turn is problematic as the variability obscures the data making it difficult to be able to analyze the effectiveness of all the services that schools potentially offer to meet the needs of EAL/D students (ACTA, 2017). This has major implications for research as comparative points may be dependent on localized contexts. Despite this, the one variable that remains consistent is the mainstream curriculum.

Currently, EAL/D programming is governed largely by the expectations of the mainstream curriculum (Williams, 2010). In terms of subject English, at a national level during the last decade, the *Australian Curriculum: English (ACE)* (ACARA, 2014), including senior secondary curriculum (Years 11–12), has been implemented with state and territories determining how the content will be integrated into existing courses (ACARA, 2020b). Content wise, the ACE is structured into three core strands: Language, Literature and Literacy (ACARA, 2014). In particular, the language strand has been highly controversial (Christie, 2010; Derewianka, 2012; Exley and Mills, 2012; Hammond, 2012; Jones and Chen, 2012; Sawyer, 2010) because it now holds English teachers responsible for the explicit teaching of knowledge about language when grammar instruction has been predominantly absent for the past two generations (Jones and Chen, 2012).

The role of grammar in writing

In Australia and the UK, government initiatives for higher literacy outcomes have focused on reintroducing explicit language teaching (Christie, 2010; Clark, 2010; Myhill, 2005; Myhill and Watson, 2014). From the Australian context, a “back to basics” approach has been championed by politicians with the role of grammar being politically hijacked as a solution for raising declining literacy rates. The problem here lies in that public and political perspectives on the role of grammar have tended to remain strongly prescriptive with a focus on what *should* and *ought* to be taught. In contrast, educators and linguists tend to theorize grammar as descriptive in terms of how language is used with a position that language will continue to change and evolve according to context. Despite this, a renewal of a standardized form of grammar seems to be emerging in countries such as England, the USA and Australia (Myhill and Watson, 2014). In the United Kingdom, both Wyse (2001) and Andrews (2005) undertook systematic reviews to test whether the renewal of grammar in England was supported by empirical research, and found no evidence that the teaching of formal grammar helps improve writing. In contrast, Myhill (2005) views the lack of evidence as a result of writing being a new area of empirical study, which only started developing models since the 1980s. She implies that because there is no theoretical framework of how grammar could support the teaching of writing, then it is not surprising that there is no substantial evidence available.

This dispute on the resurgence of explicit grammar has emerged in Australia with the introduction of the *Language* strand in the ACE. Many researchers (Derewianka, 2012; Exley and Mills, 2012; Hammond, 2012; Jones and Chen, 2012) consider the focus on knowledge about language a positive development, as it brings language to the center of practice. Sawyer (2010), however, contends that because the English curriculum is structured into the strands of Language, Literature and Literacy, it fractures the English subject into silos, and this could result in grammar being taught traditionally in isolation with the potential for it being “a proxy for linguistic competence” (p. 290). These concerns are confirmed in the UK with the introduction of a high stakes national grammar, punctuation and spelling tests in the primary years since 2014 (Department for Education, 2020), and with UK research indicating continued decontextualized grammar teaching with minimal focus on the relationship between grammar and meaning (Myhill and Watson, 2014). Despite this, Australian researchers, postulate that the basing of the ACE on a functional model, which considers connections between context and language use, will shift grammar teaching away from the decontextualized approaches that focus on sentence-level grammar (Derewianka, 2012; Hammond, 2012; Macken-Horarik et al., 2011). Myhill and Watson (2014) suggest that the difference between the UK and Australia’s pedagogical rationale for the re-introduction of grammar is that where there is an absence of a clearly articulated purpose in the UK, Australia on the other hand, has grounded it within knowing about the English language in order to understand how it works and how it can be used as a meaning-making resource.

Exley and Mills (2012) suggest that the ACE provides a new approach to the role of grammar in English content by combining traditional terminology with the principles of systemic functional linguistics, which considers the relationship between cultural and social contexts of texts to determine language choice. Macken-Horarik et al. contend that the inclusion of knowledge about language (KAL) in the curriculum has provided a metalanguage – “a language for talking about language” (2011, p. 10). It has also explicitly highlighted the KAL that teachers need to know and teach (Hammond, 2012). Despite this, Derewianka (2012) asserts that little is known about which features of language contribute to positive literacy outcomes at particular ages/stages. Furthermore, knowing grammar is not enough to use KAL effectively in practice (Harper and Rennie, 2008; Myhill and Watson, 2014). However, large-scale research in the UK, involving 13–14 years old students has produced, for the first time, evidence that contextualized grammar teaching can improve writing (Myhill et al., 2012). This is significant as it provides an “evidenced rationale for developing a purposeful role for grammar in the curriculum” (Myhill and Watson, 2014, p. 49). Interestingly though, this study found that the intervention of explicit contextualized grammar teaching to improve writing benefited more able writers, whereas the impact was neutral for less able writers (Jones et al., 2013). Consequently, this raises questions on the role of grammar in regards to EAL/D pedagogy which centralizes around language knowledge. Clark (2010) suggests that any role of grammar needs to be recontextualized insofar as to build upon teachers’ knowledge bases. Within the Australian debate, Sawyer posits that the role of grammar should be as a teaching device, only used “when judged appropriate by teachers” (2010, p. 290). Macken-Horarik (2012), nonetheless, contends that without stipulated KAL, teachers’ knowledge bases may not be sufficient to identify points of need.

The teaching and learning cycle

Similar to the grammar wars, teaching about writing has also been an ideological battleground split between two dominant camps: progressivist practice versus genre theory. The former, the progressive approaches, stemmed out of a reaction to the traditional teacher-centric prescriptiveness of the 1960s and decades prior (Gibbons, 2015). In particular, Graves’ (1983) process approach to writing grew widespread popularity across Australia, New Zealand, the United States and the United Kingdom during the 1980s (Wyse, 2018). This approach was process-oriented with importance placed on learner choice and the teacher acting in the role of an editor (Wyse, 2018). The underlying philosophy was akin to how students learn to speak without formal instruction, so in a similar token, if students were immersed in a range of writing purposes and forms, then it follows that they would organically absorb these processes and learn to write naturally (Gibbons, 2015; Knapp and Watkins, 1994). Similarly, the ‘Growth’ model which focuses on learning processes and personal response also held a pervasive influence in Australia’s English curriculum since the late 1960s (Sawyer, 2010). One of the main contentions over the new ACE is that the pendulum has swung back toward a more explicit form of practice with “strong emphasis on knowledge as ‘content’ rather than as ‘process’” (Macken-Horarik, 2011, p. 198), as exemplified with the language strand being defined as “knowing about the English language” (ACARA, 2020a).

At the other end of the spectrum is the genre-based approach to teaching writing which is based on a functional model of language founded within the work of Halliday's (1975, 1978) systemic functional linguistics (Derewianka, 2003). This approach to writing is concerned with the interrelationship between the patterns of language choice of the written text and the context in which the text is composed (Hammond and Gibbons, 2005). So, from a genre-based approach the writing process becomes explicit as students learn *about* what types of language can be used within the actual context of the language *in use* (Gibbons, 2015). Genre-based pedagogy gained momentum particularly from the 1980s as a number of critiques toward the progressive approaches heightened in relation to ELLs and students from disadvantaged backgrounds (Knapp and Watkins, 1994). The progressivist approach focused on good ideas and writers' voice but the lack of explicit teaching meant that expectations to achieve success were not always clear. This tended to reinforce social inequalities as it soon became apparent that there were many students who were unable to 'crack the codes' of writing (Knapp and Watkins, 1994). As a result, there was a strong call from many English language teachers for a more explicit form of teaching that directly addressed the language structures and conventions of writing (Gibbons, 2015).

In response, a teaching and learning cycle was developed in collaboration with teachers and linguists (Gibbons, 2015). Research in this area focused on identifying the text types or genres (e.g., Christie, 2005; Christie and Derewianka, 2008; Christie and Martin, 1997; Martin and Rothery, 1986; Martin, 1985; Rose and Martin, 2012) used in school contexts to create a curriculum cycle of teaching and learning. There are four general stages as outlined below (model summary is based on Derewianka and Jones, 2016, pp. 51–54; Gibbons, 2015, pp. 109–121):

Stage 1. *Building the field* aims to develop background knowledge and build a shared understanding of the topic. At this stage teachers can activate students' prior knowledge and build on the content of the text so that students will be able to use this knowledge within their own compositions. At this point, students are still a distance from composing themselves but this field-building process is accumulative and continues throughout the other phases of the model.

Stage 2. *Modeling the genre* involves deconstructing model texts or examples of the genre/text type. This stage explicitly focuses on modeling the form and function of the text, so unpacking the structures, language features and skills that students will need in order to create their own compositions.

Stage 3. *Joint construction* demonstrates the process of composing and allows students opportunities to practice using the skills in a scaffolded manner. For example, the teacher and students might jointly write a model text, workshoping the language features and content required within the genre. Students can also work in small groups to practice applying the particular modeled features and specific language elements, which can help teachers assess their comprehension of the modeled skills and determine whether they are ready for independent construction. Hence, this stage illustrates both the process of the composition and the product that students are working toward. Opportunities can be given to discussing how language is used, to considering alternative ways of expressing and improving ideas, drawing back on previous models and workshoping language features (e.g., grammar, vocabulary, spelling, punctuation) in the context of the language in use and at point of need for students.

Stage 4. *Independent construction* as the title suggests, at this stage students should be ready to compose their own texts with confidence. Students should also be reminded to use appropriate processes of drafting and editing as they work toward a published form of work.

The Teaching and Learning Cycle was originally developed for the explicit teaching of writing (see Rothery, 1996) drawing on the work of Vygotsky (1962, 1978) and the concept of scaffolding (Bruner, 1986) (Cruickshank, 2009). In doing so, a functional grammar for writing pedagogies was developed (see Christie, 2005; Christie and Martin, 1997; Feez, 2002; Macken-Horarik, 2002; Martin, 2009; Martin and Rose, 2008; Rose and Martin, 2012). However, in current practice it is now used as a framework to guide programming across a range of modes and curriculum contexts (Derewianka and Jones, 2016). It is a process that can be applied to an entire unit of work with the four stages being implemented across several weeks or longer "but it is never a single lesson" (Gibbons, 2015, p. 111). For further reading, suggested teaching strategies for each stage can be found in Derewianka and Jones' (2016) *Teaching Language in Context* and Gibbons' (2015) *Scaffolding Language, Scaffolding Learning: Teaching English Language Learners in Mainstream Classrooms*.

Creative writing and spoken word poetry in Australian schools: a case study

In order to situate our prior discussion on how Australia's frameworks, policies, and pedagogies shape the learning experiences of EAL/D students in secondary schools, we now offer a snapshot of contemporary research on creative writing through spoken word poetry. A growing body of research over the past three decades has shown that spoken word can nurture students' literate identities and self-confidence while also challenging problematic social and political structures (Weinstein and West, 2012). By combining the written conventions of poetry with performance, poets can utilize their voices, movements, and gestures to express meaning and connect with their audience (Dymoke, 2017). Prior research on spoken word has found that it builds on the strengths of every student (Jocson, 2005), celebrates human difference (Biggs-El, 2012), and offers a platform for youth to interrogate relevant societal and cultural issues (Fisher, 2015). In Flint and Laman's (2012) study, English language learners were able to use spoken word to link their everyday experiences to broader societal issues. Consequently, Coppola et al. (2019) argue that spoken word can encourage "multidimensional and intersectional explorations of culture and identity" (p. 243). For EAL/D students, in particular, poetry can be a powerful way to explore their lived experiences and cultivate their creative writing within school contexts.

The genre of spoken word poetry connects the written conventions of poetry with performance (Dymoke, 2017; Gregory, 2013). Pedagogically, spoken word poetry has the potential to lend itself to both genre and progressivist writing practices. For example, spoken word poetry supports culturally sustaining pedagogy as it draws on students' sociocultural resources (Paris, 2012) with the focus on developing students' voices and creating space to honor their lived experiences. Thus, the form can be process-oriented as importance is placed on learner choice where students are encouraged to base their writing on their experiences to compose a personal response. This addresses the *Australian Curriculum English* which requires students to "create literary texts that reflect an emerging sense of personal style" (ACARA, 2020c). In addition, the form is also conducive to SFL approaches and provides a vehicle for EAL/D students to learn about language through using language.

Spoken word in context: real talk in Australian schools

Real Talk is an annual initiative of the Bankstown Poetry Slam, the largest poetry slam in Australia, and takes place in high schools in western and southwestern Sydney in New South Wales. Over six weeks, spoken word poetry workshops are facilitated by experienced spoken word poets, who act as mentors, and the workshops are primarily attended by students in Years 9 and 10. Through the weekly workshops, mentor poets work with students to critically explore topics such as culture, gender, and race through poetry. The workshops culminate with an in-school heat, and the winners then compete in a finale at a well-known theater in Sydney. In Jones and Curwood's (2020) study of three schools who participated in Real Talk, half of the students identified as having an additional language other than English, and all schools were identified as low socioeconomic by the federal government (ACARA, 2018).

Poetry offers young adults the opportunity to critically and emotionally explore with the world around them (Kovalik and Curwood, 2019) and to push back against oppressive narratives and power structures (Curwood and Gibbons, 2009). Graham and Harris (1997) emphasized the importance of the creation of "legitimate and authentic writing tasks aimed at a real audience" (p. 417). Spoken word poetry is inherently performative, yet the structure of the Real Talk workshops gave students agency in whether and when they would share their creative work; consequently, as students gain confidence in their writing and performing, they increasingly chose to share their poetry with a wider audience. In this sense, spoken word encourages students to "think about who they are and who others (media, public, dominant discourses) say they are and create dialogue to mediate and negotiate any discrepancies" (Muhammad and Gonzalez, 2016, p. 449). Spoken word offers an opportunity for EAL/D students, in particular, to consider the intersections of language and identity while cultivating literacy skills such as listening, speaking, and writing.

Real Talk cultivates students' understanding of the spoken word genre by creating a safe space to learn about, evaluate the effect, and manipulate elements of grammar and aspects of language without the prescriptiveness and constraints of traditional grammar due to its free form. The teaching and learning cycle can also be applied effectively as the working toward a finished product can be readily backwards mapped, such as modeling the genre through various exemplars to develop a consciousness of language usage with attention to how the relationship between cultural and social contexts, purpose and audience plays an integral role in determining language choice. Our research, however, took the modeling and joint construction phase a step further by partnering students with a Real Talk mentor poet to create "a new learning environment that acted as a third space as teachers stepped back and joined their students in the writing process" (Jones and Curwood, 2020, p. 284).

Multilingualism, multiculturalism, and meaning making

The New South Wales Department of Education (2020a) emphasizes that additional language practices can "support learning in English" (p. 28). In alignment with the state-based frameworks and policies, the Real Talk workshops used "culturally and linguistically inclusive strategies... [such as] encouraging the use of first language" (p. 66). One of the teachers in the study noted that many of her students have "developing literacy skills in English, but are quite sophisticated in their first language." When "no single language or register is privileged" (Gutiérrez et al., 1999 p. 293), EAL/D students are able to capitalize on their rich linguistic capabilities in other languages along with their intercultural experiences to explore within their spoken word poetry. Jones and Curwood (2020) argue that the Real Talk workshops embody culturally sustaining pedagogy in that they nurture multilingualism and multiculturalism, which Paris (2012) suggests are inextricably linked to access and power in our globalized world.

The classroom context plays a significant role in English language learners' engagement and achievement with writing. Not only do the types of tasks and the quality of teacher-student interaction shape students' writing practices (Gutiérrez, 1992), students' academic writing is informed by how they position themselves in relation to the power structures embedded in disciplinary activities (Enright, 2011). Through writing, sharing, and listening to spoken word poetry, EAL/D students are able to learn about poetry broadly, including a variety of literary techniques, but also engage in a "performance of possibilities" (Madison, 2005, p. 190). The Real Talk curriculum encourages students to explore their own identities through poetry; in the first week of the workshops, mentor poets sought to build the field by exploring and defining spoken word. In the next week, students worked with their poetry mentor to consider various aspects of identity, including race, ethnicity, gender, age, sexual orientation, socioeconomic status, and language background. Students used this as a basis to craft an 'I am' poem, which one mentor described as "the things that make up who you are, the things people call you, and the things society says about you" and modeled her own for the students. Arihi, a Samoan EAL/D student, shared that this approach created "a different learning environment" where students' knowledge was valued.

In the Real Talk workshops, students were encouraged to share their creative writing, either through reading it out loud from their seats or performing at the front of the classroom. Thuy, a Vietnamese EAL/D student, said that she felt "intimidated" to perform at first, but the support of peers "gives you the environment where you can talk." Because students' work was not part of a formal

assessment, it offered them more agency over the content and less pressure to write and to speak in a prescribed way. At the same time, students' engagement in the workshops allowed them to fulfill key outcomes of the NSW English Syllabus, including to "critically evaluate the ways bias, stereotypes, perspectives ... are constructed," (NESA, 2012, p. 155). Some students with refugee backgrounds in the workshops had only been in Australia for a couple years. Zara and Armami, from Iraq and Syria respectively, "eloquently conveyed their message" according to their teacher. Jones and Curwood (2020) argue that "by establishing a space where student voices are privileged, Real Talk effectively engaged diverse students with varying personalities and literacy abilities" (p. 4).

Students' multilingual skills can offer them valuable opportunities for access and innovative pathways for power (Paris, 2012). In the Real Talk workshops, half of the students identified as EAL/D. In typical classroom settings, the incorporation of additional languages is either limited or nonexistent. Rather than being seen as an asset, students' multilingual abilities may be overshadowed by a focus on academic English, rather than being seen as a pathway to engaging with language and grammar. Moreover, students typically write on their own and for an audience of one - "the teacher-as-examiner" (Applebee and Langer, 2011, p. 17). Spoken word, in contrast, encourages EAL/D students to write (and to perform) together, and to consider their multilingual skills and multicultural identities as assets for engaging in joint construction. For instance, Jaimini and Miracle's collaborative poem *Culture* draws on their collective knowledge of English, Gujarati, and Māori. In class, Miracle began by writing, "He reo Maori taku reo. He Maori ahau me he koa ana au kit e Maori. My language is Maori. I am Maori. I am proud to be Maori" and Jaimini added, "Hu Gujarati bolu chu. Hu Gujarati chu. My language is Gujarati. I am Gujarati." They continued their poem by exploring the unique aspects of their respective cultures, such as food, before confronting stereotypes, "The words we are speaking, the stories we are telling, have been twisted into something they're not." Later performed at the Bankstown Poetry Slam's Grand Finale, held at the Sydney International Convention Center in front of 4000 people, Jaimini and Miracle received the People's Choice Award.

Spoken word poetry is a powerful way to foster EAL/D students' writing and speaking, particularly through its focus on accessibility, agency, and identity. Gibbons (2015) argues that "spoken language and written language are better understood as a continuum rather than as two discrete forms of language" (p. 79). As such, spoken word poetry has the power to shape how, why, and for whom EAL/D students write, which has a direct impact on how their learning will be formatively and summatively assessed in formal school contexts. For instance, the New South Wales Education Standards Authority has recently introduced a multimodal assessment requirement in English for students in Years 11 and 12, and added spoken word poems from Luka Lesson and Kate Tempest to the list of prescribed texts for study in Standard and Advanced English. Not only does this validate performance poetry as a valued text for academic study, it also creates space for teachers to integrate spoken word, along with digital poems (Curwood and Cowell, 2011) and multimodal counternarratives (Curwood and Gibbons, 2009), into the English curriculum.

Moving forward: developing English language learners' writing practices

This chapter critically discussed some of the complexities and challenges around supporting English language learners in writing development. The review of the international literature has highlighted the significance and increasing need to further develop writing pedagogy that targets language through explicit instruction and builds on funds of knowledge, which Australia's rapidly growing ELL population bring to the country's multilingual and multicultural classrooms. Specific examples of policy and practice situated within the Australian context have illustrated some of the competing mandates, complexities and considerations that influence the implementation of writing pedagogy and EAL/D systems when working within overarching educational structures and institutions. We explored, as suggested applied models of practice, the teaching and learning cycle which provides a means for scaffolded programming that aims to integrate content and language teaching through an explicit instructional framework. We also highlighted how spoken word poetry can serve as a compositional form of creative expression to provide a meaningful way to draw upon and cultivate students' sociocultural resources to move along the spoken to written mode continuum. Multimodality, in this sense, is integral to modern literacies, language, and schooling.

In moving forward, successful support for English language learners, at least in the Australian context, can be "achieved through co-designed teaching programs" (DoE, 2020a, p. 24). Given that in the mainstream curriculum, knowledge about language is taught in conjunction with content discipline, this integration of language and content has been increasingly acknowledged as central to the academic success for English language learners (de Jong and Barko-Alva, 2015; Schlepppegrell, 2004) as it provides exposure to meaningful use of academic language in situational contexts (Schlepppegrell, 2004; Snow et al., 1989). While this integration is typically a pedagogical focus for ELL teachers, in contrast, mainstream teachers are generally not trained in this area (de Jong and Barko-Alva, 2015). Consequently, a dual focus is needed where mainstream teachers develop knowledge on how to address the needs of English language learners, and for ELL specialists to enhance their content knowledge (Tigert and Peercy, 2017). To this, another concern raised in the literature is whether teachers are adequately prepared to teach knowledge about language given the resurgence of grammar within various curricula and the educational sphere (Borg, 2003; Cajkler and Hislam, 2002; Harper and Rennie, 2008; Loudon and Rohl, 2006; Macken-Horarik et al., 2011; Milton et al., 2007; Sangster, Anderson, & O'Hara, 2013). This highlights a need to examine and improve how the development of teachers' knowledge about language is being addressed in teacher education; and for further empirical work that clearly theorizes the role of grammar in the curriculum, and provides evidence-based ways on how and whether it can be taught in a meaningful, contextualized manner to improve writing pedagogy for specific cohorts. Lastly, future research can explore how multimodal assessment can move beyond a print-focused paradigm (Curwood, 2012) to better account for multimodal ways of making meaning and capture students' disciplinary learning and understanding.

Acknowledgments

The authors are grateful to Sarah McCarthy and Eveline Chan for their constructive feedback on an earlier version of this chapter.

References

- Alford, J., 2014. "Well, hang on, they're actually much better than that!": disrupting dominant discourses of deficit about English language learners in senior high school English. *Engl. Teach. Pract. Critiq.* 13 (3), 71–88.
- Andrews, R., 2005. Knowledge about the teaching of [sentence] grammar: the state of play. *Engl. Teach. Pract. Critiq.* 4 (3), 69–76.
- Applebee, A.N., 1984. Writing and reasoning. *Rev. Educ. Res.* 54, 577–596.
- Applebee, A., Langer, J., 2011. A Snapshot of writing instruction in middle schools and high schools. *Engl. J.* 100 (6), 14–27.
- Australian Council of TESOL Association, 2017. Submission to the Review to Achieve Educational Excellence in Australian Schools. Author, Sydney, NSW. http://www.tesol.org.au/files/files/582_Gonski_Review_ACTA_submission_summary_Final.pdf. (Accessed March 2018).
- Australian Council of TESOL Association, 2018. National Strategy for Language in Education and Training. Author, Sydney, NSW. http://tesol.org.au/files/files/596_ACTA_language_in_education_and_training_strategy_November_2018.pdf. (Accessed March 2018).
- Australian Curriculum, Assessment and Reporting Authority, 2014. Australian Curriculum: English. Author, Sydney, NSW. <http://www.australiancurriculum.edu.au/english/content-structure/language>. (Accessed July 2014).
- Australian Curriculum Assessment and Reporting Authority, 2018. My School. Author, Sydney, NSW. <https://www.myschool.edu.au/>. (Accessed September 2018).
- Australian Curriculum Assessment and Reporting Authority, 2020a. Meeting the Needs of Students for Whom English Is an Additional Language or Dialect. Author, Sydney, NSW. <https://www.australiancurriculum.edu.au/resources/student-diversity/meeting-the-needs-of-students-for-whom-english-is-an-additional-language-or-dialect/>. (Accessed September 2020).
- Australian Curriculum Assessment and Reporting Authority, 2020b. Development of the Australian Curriculum. Author, Sydney, NSW. <https://www.acara.edu.au/curriculum/history-of-the-australian-curriculum/development-of-australian-curriculum>. (Accessed September 2020).
- Australian Curriculum Assessment and Reporting Authority, 2020c. English (Version 8.4) Year 10. Author, Sydney, NSW. <https://www.australiancurriculum.edu.au/f-10-curriculum/english/?strand=Language&strand=Literature&strand=Literacy&capability=ignore&priority=ignore&year=11584&elaborations=true&cd=ACELT1814&searchTerm=ACELT1814#dimension-content>. (Accessed September 2020).
- Bangert-Drowns, R.L., Hurley, M.M., Wilkinson, B., 2004. The effects of school-based writing-to-learn interventions on academic achievement: a meta-analysis. *Rev. Educ. Res.* 74, 29–58.
- Biggs-El, C., 2012. Spreading the Indigenous gospel of rap music and spoken word poetry: critical pedagogy in the public sphere as a stratagem of empowerment and critique. *West. J. Black Stud.* 36 (2), 161–168.
- Borg, S., 2003. Teacher cognition in grammar teaching: a literature review. *Lang. Aware.* 12 (2), 96–108.
- Bruner, J., 1986. *Actual Minds, Possible Worlds*. Harvard University Press, Cambridge, MA.
- Cahnmann, M., 2006. Reading, living, and writing bilingual poetry as ScholARTistry in the language arts classroom. *Lang. Arts* 83 (4), 342–352.
- Cajkler, W., Hall, B., 2009. 'When they first come in what do you do?' English as an additional language and newly qualified teachers. *Lang. Educ.* 23 (2), 153–170.
- Cajkler, W., Hislam, J., 2002. Trainee teachers' grammatical knowledge: the tension between public expectation and individual competence. *Lang. Aware.* 11 (3), 161–177.
- Christie, F., 2005. Using the functional grammar to understand children's written texts. *Austr. Rev. Appl. Linguist. Suppl. Ser.* 19 (1), 9–21.
- Christie, F., 2010. The "grammar wars" in Australia. In: Locke, T. (Ed.), *Beyond the Grammar Wars*. Routledge, New York, NY, pp. 55–72.
- Christie, F., Derewianka, B., 2008. *School Discourse*. Continuum International Publishing Group, London.
- Christie, F., Martin, J.R., 1997. *Genres in Institutions: Social Processes in the Workplace and Schools*. Academic Press, Heden, VA.
- Clark, U., 2010. The problematics of prescribing grammatical knowledge: the case in England. In: Locke, T. (Ed.), *Beyond the Grammar Wars*. Routledge, New York, NY, pp. 38–54.
- Coppola, R., Woodard, R., Vaughan, A., 2019. And the students shall lead us: putting culturally sustaining pedagogy in conversation with universal design for learning in a middle-school spoken word poetry unit. *Lit. Res.* 68, 226–249.
- Costley, T., 2014. English as an additional language, policy, and the teaching and learning of English in England. *Lang. Educ.* 28 (3), 276–292.
- Cumming, A., 2013. Multiple dimensions of academic language and literacy development. *Lang. Learn.* 63 (1), 130–152.
- Cumming, A., 2015. Theoretical orientations to L2 writing. In: Manchón, R., Matsuda, P.K. (Eds.), *Handbook of Second and Foreign Language Writing*. Walter de Gruyter, Boston, MA, pp. 65–90.
- Cummins, J., 1979. Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism* 19, 121–129.
- Cummins, J., 1991. Interdependence of First and Second Language Proficiency in Bilingual Children. In: Bialystok, G. (Ed.), *Language Processing in Bilingual Children*. Cambridge University Press, Cambridge.
- Cummins, J., Mirza, R., Stille, S., 2012. English language learners in Canadian schools: emerging directions for school-based policies. *TESL Can. J.* 29 (6), 25–48.
- Curwood, J.S., 2012. Cultural shifts, multimodal representations, and assessment practices: a case study. *E-Learn. Digit. Med.* 9 (2), 232–244.
- Curwood, J.S., Cowell, L.L., 2011. iPoetry: creating space for new literacies in the English curriculum. *J. Adolesc. Adult Lit.* 55 (2), 107–117.
- Curwood, J.S., Gibbons, D., 2009. "Just like I have felt": multimodal counternarratives in youth-produced digital media. *Int. J. Learn. Med.* 1 (4), 59–77.
- Cruikshank, K., 2009. EAP in secondary schools. In: Belcher, D. (Ed.), *Specific Purposes in Theory and Practice*. University of Michigan Press, Ann Arbor, MI, pp. 22–40.
- de Jong, E., Barko-Alva, K., 2015. Mainstream teachers in two-way immersion programs: becoming content and language teachers. In: Freeman, Y.S., Freeman, D.E. (Eds.), *Research on Preparing Inservice Teachers to Work Effectively with Emergent Bilinguals*. Emerald Group Publishing Limited, Bingley, pp. 107–126.
- DelliCarpini, M., Gulla, A.N., 2016. Working with English language learners in the mainstream English language arts classroom through collaboration and two-way content-based instruction. In: deOliveira, L., Shoffners, M. (Eds.), *Teaching English Language Arts to English Language Learners*. Palgrave Macmillan, London, pp. 79–105.
- Department for Education, 2019. *Schools, Pupils and Their Characteristics: January 2019*. Author, Sheffield. <https://assets.publishing.service.gov.uk/>. (Accessed September 2020).
- Department for Education, 2020. *National Curriculum Test Handbook: 2019 Key Stage 1 and Key Stage 2*. Author, London.
- Derewianka, B., 2003. Trends and issues in genre-based approaches. *REL C. J.* 34 (2), 133–154.
- Derewianka, B., 2012. Knowledge about language in the Australian curriculum: English. *Australian. J. Lang. Lit.* 35 (2), 127–146.
- Derewianka, B., Jones, P., 2019. *Teaching Language in Context*, second ed. Oxford, Melbourne, VIC.
- Dymoke, S., 2017. 'Poetry is not a special club': how has an introduction to the secondary discourse of spoken word made poetry a memorable learning experience for young people? *Oxf. Rev. Educ.* 43 (2), 225–241.
- Education Council, 2019. *Alice Springs (Mparntwe) Education Declaration*. Author, Victoria.
- Enright, K.A., 2011. Language and literacy for a new mainstream. *Am. Educ. Res. J.* 48 (1), 80–118.
- Exley, B.E., Mills, K.A., 2012. Parsing the Australian curriculum English: grammar, multimodality and cross-cultural texts. *Aust. J. Lang. Lit.* 35 (2), 192–205.

- Feez, S., 2002. Heritage and innovation in second language education. In: Johns, A.M. (Ed.), *Genre in the Classroom: Multiple Perspectives*. Lawrence Erlbaum, Mahwah, NJ, pp. 43–69.
- Fisher, M.T., 2015. From the coffee house to the school house: the promise and potential of spoken word poetry in school contexts. *Engl. Educ.* 37 (2), 115–131.
- Fitzgerald, J., Shanahan, T., 2000. Reading and writing relations and their development. *Educ. Psychol.* 35, 39–50.
- Flint, A.S., Laman, T.T., 2012. Where poems hide: finding reflective, critical spaces inside writing workshop. *Theory Into Pract.* 51 (1), 12–19.
- Franson, C., 1999. Mainstreaming learners of English as an additional language: the class teacher's perspective. *Lang. Cult. Curric.* 12 (1), 59–71.
- Friedman, A.A., Pederson, J.M., Bacon, C.K., 2019. Getting to what is: poetry as a genre of access for multilingual learners. *Teaching/Writing* 6 (1), 23–56.
- Gadd, M., Parr, J.M., 2017. Practices of effective writing teachers. *Read. Writ.* 30 (7), 1551–1574.
- Gibbons, P., 2015. *Scaffolding Language, Scaffolding Learning: Teaching English Language Learners in the Mainstream Classroom*, second ed. Heinemann, Portsmouth, NH.
- Gillespie, A., Graham, S., Kihara, S., Hebert, M., 2014. High school teachers' use of writing to support students' learning: a national survey. *Read. Writ.* 27, 1043–1072.
- Graham, S., 2006. Strategy instruction and the teaching of writing: a meta-analysis. In: MacArthur, C.A., Graham, S., Fitzgerald, J. (Eds.), *Handbook of Writing Research*. The Guilford Press, New York, NY, pp. 187–207.
- Graham, S., 2019. Changing how writing is taught. *Rev. Res. Educ.* 43, 277–303.
- Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., Olinghouse, N., 2012. *Teaching Elementary School Students to Be Effective Writers: A Practice Guide*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC.
- Graham, S., Harris, K.R., 1997. It can be taught, but it does not develop naturally: Myths and realities in writing instruction. *Sch. Psychol. Rev.* 26, 414–424.
- Graham, S., Harris, K.R., Hebert, M., 2011. It is more than just the message: analysis of presentation effects in scoring writing. *Focus Except. Child.* 44 (4), 1–12.
- Graham, S., Hebert, M., 2011. Writing-to-read: a meta-analysis of the impact of writing and writing instruction on reading. *Harv. Educ. Rev.* 81, 710–744.
- Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K.R., Aitken, A., Barkel, A., Kavanaugh, C., Talukdar, J., 2018. Reading for writing: a meta-analysis of the impact of reading and reading instruction on writing. *Rev. Educ. Res.* 88, 243–284.
- Graham, S., Perin, D., 2007. *Writing Next: Effective Strategies to Improve Writing of Adolescent Middle and High School*. Alliance for Excellence in Education, Washington, DC.
- Graham, S., Rijlaarsdam, G., 2016. Writing education around the globe. *Read. Writ.* 29, 781–792.
- Graves, D.H., 1983. *Writing: Teachers and Children at Work*. Heinemann Educational, Portsmouth, NH.
- Gregory, H., 2013. Youth take the lead: digital poetry and the next generation. *Res. J. Natl. Assoc. Teach. Eng.* 47 (2), 118–133.
- Gutiérrez, K., 1992. A comparison of instructional contexts in writing process classrooms with Latino children. *Educ. Urban Soc.* 24, 244–262.
- Gutiérrez, K.D., Baquedano-López, P., Tejeda, C., 1999. Rethinking diversity: hybridity and hybrid language practices in the third space. *Mind Cult. Activ.* 6 (4), 286–303.
- Gutiérrez, K.D., Zitali Morales, P., Martinez, D.C., 2009. Re-mediating literacy: culture, difference and learning for students from non-dominant communities. *Rev. Res. Educ.* 33, 212–245.
- Halliday, M.A.K., 1975. *Learning How to Mean: Exploring in the Development of Language*. Edward Arnold, London.
- Halliday, M.A.K., 1978. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. Edward Arnold, London.
- Hammond, J., 2012. Hope and challenge in the Australian curriculum: implications for EAL students and their teachers. *Aust. J. Lang. Lit.* 35 (1), 223–240.
- Hammond, J., Gibbons, P., 2005. Putting scaffolding to work: the contribution of scaffolding in articulating ESL education. *Prospect* 20 (1), 6–30.
- Harper, C., Cook, K., James, C.K., 2010. Content-language integrated approaches for teachers of EAL learners: examples of reciprocal teaching. In: Leung, C., Creese, A. (Eds.), *English as an Additional Language: Approaches to Teaching Linguistic Minority Students*. SAGE Publications, London, pp. 15–32.
- Harper, C., de Jong, E., 2009. English language teacher expertise: the elephant in the room. *Lang. Educ.* 23 (2), 137–151.
- Harper, H., Rennie, J., 2008. 'I had to go out and get myself a book on grammar': a study of pre-service teachers' knowledge about language. *Aust. J. Lang. Lit.* 32 (1), 22–37.
- Hayes, J.R., 2012. Modeling and remodeling writing. *Writ. Commun.* 29, 369–388.
- Hornberger, N.H., Link, H., 2012. Translanguaging and transnational literacies in multilingual classrooms: a biliteracy lens. *Int. J. Biling. Educ. Biling.* 15, 261–278.
- Janks, H., 2010. *Literacy and Power*. Routledge, New York, NY.
- Janzen, J., 2008. Teaching English language learners in the content areas. *Rev. Educ. Res.* 78 (4), 1010–1038.
- Jocson, K.M., 2005. 'Taking it to the mic': pedagogy of June Jordan's poetry for the people and partnership with an urban high school. *Engl. Educ.* 37 (2), 132–148.
- Joint Standing Committee on Migration, 2017. *Parliament of the Commonwealth of Australia: No One Teaches You to Become an Australian*. Report of the Inquiry into Migrant Settlement Outcomes. Author, Sydney, NSW. https://parlinfo.aph.gov.au/parlInfo/download/committees/reportjnt/024098/toc_pdf/NooneteachesyoutobecomeanAustralian.pdf;fileType=application%2Fpdf. (Accessed March 2018).
- Jones, P., Chen, H., 2012. Teachers' knowledge about language: issues of pedagogy and expertise. *Aust. J. Lang. Lit.* 35 (1), 147–168.
- Jones, K., Curwood, J.S., 2020. Tell the story, speak the truth: creating a third space through spoken word poetry. *J. Adolesc. Adult Lit.* 64 (3), 281–289.
- Jones, S., Myhill, D., Bailey, T., 2013. Grammar for writing? An investigation of the effects of contextualised grammar teaching on students' writing. *Read. Writ.* 26 (8), 1241–1263.
- Knapp, P., Watkins, M., 1994. Context – Text – Grammar: Teaching the Genres and Grammar of School Writing in Infants and Primary Classrooms. Text Productions, Broadway.
- Kovalik, K., Curwood, J.S., 2019. #poetryisnotdead: understanding Instagram poetry within a transliteracies framework. *Literacy* 53 (4), 185–195.
- Leung, C., Franson, C., 2001. Mainstreaming: ESL as a diffused curriculum concern. In: Mohan, B., Leung, C., Davison, C. (Eds.), *English as a Second Language in the Mainstream: Teaching, Learning and Identity*. Pearson, London, pp. 11–29.
- Louden, W., Rohl, M., 2006. 'Too many theories and not enough instruction': perceptions of preservice teacher preparation for literacy teaching in Australian schools. *Literacy* 40 (2), 66–78.
- Luke, A., 2000. Critical literacy in Australia: a matter of context and standpoint. *J. Adolesc. Adult Lit.* 43 (5), 448–461.
- Macken-Horarik, M., 2002. 'Something to shoot for': a systemic functional approach to teaching genre in secondary school. In: Johns, A.M. (Ed.), *Genre in the Classroom: Multiple Perspectives*. Lawrence Erlbaum, Mahwah, NJ, pp. 17–42.
- Macken-Horarik, M., 2011. Building a knowledge structure for English: reflections on the challenges of coherence, cumulative learning, portability and face validity. *Aust. J. Educ.* 55 (3), 197–213.
- Macken-Horarik, M., 2012. Why school English needs a 'good enough' grammatics (and not more grammar). *Chang. Engl. Stud. Cult. Educ.* 19 (2), 179–194.
- Macken-Horarik, M., Love, K., Unsworth, L., 2011. A grammatics 'good enough' for school English in the 21st century: four challenges in realising the potential. *Aust. J. Lang. Lit.* 34 (1), 9–23.
- Madison, D.S., 2005. *Critical Ethnography: Method, Ethics, and Performance*. Sage, Thousand Oaks, CA.
- Martin, J.R., 1985. *Factual Writing: Exploring and Challenging Social Reality*. Deakin University Press, Geelong, VIC.
- Martin, J.R., 2009. Genre and language learning: a social semiotic perspective. *Ling. Educ.* 20 (1), 10–21.
- Martin, J.R., Rose, D., 2008. *Genre Relations: Mapping Culture*. Equinox, London.
- Martin, J.R., Rothery, J., 1986. *Writing Project: Report (Working Papers in Linguistics 4)*. Department of Linguistics, University of Sydney, Sydney, NSW.
- McCarthy, S.J., Garcia, G.E., 2005. English language learners' practices and attitudes. *Writ. Commun.* 22 (1), 36–75.
- McLean, C., Boling, E., Rowsell, J., 2009. Engaging diverse students in multiple literacies in and out of school. In: Morrow, L., Rueda, R., Lapp, D. (Eds.), *Handbook of Research on Literacy and Diversity*. The Guilford Press, New York, NY, pp. 158–172.
- Milton, M., Rohl, M., House, H., 2007. Secondary beginning teachers' preparedness to teach literacy and numeracy: a survey. *Austr. J. Teach. Educ.* 32 (2), 1–20.
- Ministerial Council on Education, Employment, Training and Youth Affairs, 2008. *Melbourne Declaration on Educational Goals for Young Australians*. Author, Melbourne, VIC.
- Myhill, D., 2005. Ways of knowing: writing with grammar in mind. *Engl. Teach. Pract. Critiq.* 4 (3), 77–96.

- Myhill, D., Jones, S., Lines, H., Watson, A., 2012. Re-thinking grammar: the impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Res. Pap. Educ.* 27 (2), 139–166.
- Myhill, D., Watson, A., 2014. The role of grammar in the writing curriculum: a review of the literature. *Child Lang. Teach. Ther.* 30 (1), 41–62.
- Muhammad, G., Gonzalez, L., 2016. Slam Poetry: an artistic resistance toward identity, agency, and activism. *Equity & Excell. Educ.* 49 (4), 440–453.
- Nelson, K., Nelson, K.E., 1978. Cognitive pendulums and their linguistic realization. In: Nelson, K.E. (Ed.), *Children's Language*. Gardener, New York, NY. (Accessed September 2020).
- NSW Department of Education, 2014. *Local Schools, Local Decisions: The Reform Agenda*. Author, Sydney, NSW.
- NSW Department of Education, 2020a. *English as an Additional Language or Dialect: Advice for Schools*. Author, Sydney, NSW.
- NSW Department of Education, 2020b. *Intensive English Centres*. Author, Sydney, NSW. <https://education.nsw.gov.au>. (Accessed September 2020).
- NSW Department of Education: Centre for Education Statistics and Evaluation, 2019. *Schools: Language Diversity in NSW, 2019*. Author, Sydney, NSW.
- New South Wales Education Standards Authority (NESA), 2012. *English K-10 Syllabus*. <https://educationstandards.nsw.edu.au>.
- NSW ESL and Refugee Education Working Party, 2013. "Passing the Buck": Local Schools, Local Decisions and English as a Second Language (ESL) Education in NSW Government Schools. A Draft Report from the NSW ESL and Refugee Working Party. Author, Sydney, NSW. https://www.nswtf.org.au/files/passing_the_buck_final_report_plus_executive_summary.pdf. (Accessed March 2018).
- Organisation for Economic Cooperation and Development, 2019. *PISA 2018 Results (Volume I): What Students Know and Can Do (Summary)*. OECD Publishing, Paris.
- Ortmeier-Hooper, C., 2013. *The ELL Writer: Moving beyond Basics in the Secondary Classroom*. Teachers College Press, New York, NY.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology and practice. *Educ. Res.* 41 (3), 93–97.
- Peterson, S.S., Botelho, M.J., Jang, E., Kerekes, J., 2007. Writing assessment: what would multiliteracies teachers do? *Lit. Learn. Middle Years* 15 (1), 29–35.
- Purves, A.C., 1992. Reflections on research and assessment in written composition. *Res. Teach. Engl.* 26, 108–122.
- Ray, A., Graham, S., Houston, J., Harris, K.R., 2016. Teachers' use of writing to support students' learning in middle school: a national survey in the United States. *Read. Writ.* 29, 1039–1068.
- Rose, D., Martin, J.R., 2012. *Learning to Write, Reading to Learn. Genre, Knowledge and Pedagogy in the Sydney School*. Equinox Publishing Ltd, South Yorkshire.
- Rothery, J., 1996. Making changes: developing an educational linguistics. In: Hasan, R., Williams, G. (Eds.), *Literacy in Society*. Addison Wesley Longman, Essex.
- Rushton, K., 2008. Cooperative planning and teaching for ESL students in the mainstream classroom. *TESOL Context* 18 (1), 21–28.
- Sangster, P., Anderson, C., O'Hara, P., 2013. Perceived and actual levels of knowledge about language amongst primary and secondary student teachers: do they know what they think they know? *Lang. Aware* 22 (4), 293–319.
- Sawyer, W., 2010. Structuring the new English in Australia: James Moffett and English teaching in New South Wales. *Appl. Financ. Econ.* 17 (3), 285–296.
- Scarcella, R., 2002. Some key factors affecting the English learners' development of advanced literacy. In: Schleppegrell, M., Colombi, C. (Eds.), *Developing Advanced Literacy in First and Second Languages*. Erlbaum, Hillsdale, NJ, pp. 209–226.
- Schleppegrell, M., 2004. *The Language of Schooling: A Functional Linguistics Perspective*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Sinclair, J., Lau, C., 2018. Initial assessment for K-12 English language support in six countries: revisiting the validity-reliability paradox. *Lang. Educ.* 32 (3), 257–285.
- Snow, M.A., Met, M., Genesee, F., 1989. A conceptual framework for the integration of language and content in second/foreign language instruction. *Tesol Q.* 23, 201–219.
- Snow, C., Uccelli, P., 2009. The challenge of academic language. In: Olson, D.R., Torrance, N. (Eds.), *The Cambridge Handbook of Literacy*. Cambridge University Press, Cambridge, pp. 112–133.
- Street, B., 2005. Introduction: new literacy studies and literacies across educational contexts. In: Street, B. (Ed.), *Literacies across Educational Contexts: Mediating Learning and Teaching*. Caslon Publishing, Philadelphia, PA, pp. 1–20.
- Thomas, W.P., Collier, V., 1997. *School Effectiveness for Language Minority Students*. National Clearing House for Bilingual Education, Washington, DC.
- Tigert, J.M., Peercy, M.M., 2017. Preparing to teach both content and language: four ESOL teacher candidates' experiences. *TESOL J.* 9 (3), 542–556.
- Turkan, S., Buzick, H.M., 2016. Complexities and issues to consider in the evaluation of content teachers of English language learners. *Urban Educ.* 51 (2), 221–248.
- United States Department of Education, National Center for Education Statistics, 2019. *Common Core of Data, "Local Education Agency Universe Survey," 2000-01 through 2017-18*. <https://nces.ed.gov>. (Accessed September 2020).
- Velasco, P., García, O., 2014. Translanguaging and the writing of bilingual learners. *Biling. Res. J.* 37 (1), 6–23.
- Villegas, A.M., SaizdelLaMora, K., Martin, A.D., Mills, T., 2018. Preparing future mainstream teachers to teach English language learners: a review of the empirical literature. *Educ. Forum* 82 (2), 138–155.
- Von Esch, K.S., Kavanagh, S.S., 2018. Preparing mainstream classroom teachers of English learner students: grounding practice-based designs for teacher learning in theories of adaptive expertise development. *J. Teach. Educ.* 69 (3), 239–251.
- Vygotsky, L., 1962. In: Hanfmann, E., Vakar, G. (Eds.), *Thought and Language*. Massachusetts Institute of Technology, Cambridge, MA.
- Vygotsky, L., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Weinstein, S., West, A., 2012. Call and responsibility: critical questions for youth spoken poetry. *Harv. Educ. Rev.* 82 (2), 282–302.
- William, A., 2010. Connecting communication, curriculum and second language literacy development: meeting the needs of 'low literacy' EAL/ESL learners. In: Leung, C., Creese, A. (Eds.), *English as an Additional Language: Approaches to Teaching Linguistic Minority Students*. SAGE Publications, London, pp. 44–56.
- Wyse, D., 2001. Grammar for writing? A critical review of empirical evidence. *Br. J. Educ. Stud.* 49 (4), 411–427.
- Wyse, D., 2018. Choice, voice and process – teaching writing in the 21st Century: revisiting the influence of Donald Graves and the process approach to writing. *Engl. Aust.* 53 (1), 82–91.

Critical media literacy and second language acquisition

Andrea Gambino and Jeff Share, University of California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	76
Disparities of access and information	76
A crisis of trust	77
Critical media literacy	77
Language acquisition	79
Asset-based vs. deficit-based approaches	79
Putting theory into practice	80
Photography	80
Curriculum interventions	81
Research studies	82
Conclusion	83
References	83
Further reading	84
Relevant websites	85

Introduction

For most of the time that humans have been walking on this Earth, we have passed along our knowledge and culture through oral traditions, as the elders and those with the most experience have been the storytellers assuring that our collective wisdom is shared with the next generation. Only 500 years ago with the invention of the printing press, alphabetical literacy took prominence and stories could be mass produced and disseminated across the globe. Since then, the innovation of information communication technologies (ICTs), from radio to television to the internet and cell phones, has changed the world. [McLuhan's \(1997\)](#) warning that "the medium is the message" (p. 7) resonates today as never before. Today the wealthiest corporations on Earth are media companies and technology platforms, and with over half the world's population connected to these digital networks our dominant storytellers are now marketeers, buying and selling information, products, our attention, and our personal data, often at the expense of society and the planet.

Disparities of access and information

In UNICEF's "The State of the World's Children 2017: Children in a Digital World" they assert, "Digital technology has already changed the world—and as more and more children go online around the world, it is increasingly changing childhood" ([2017](#), p. 3). They found that the most connected people (those with access to digital technologies and the internet) on the planet are youth between the ages of 15–24 and about one-third of all internet users are under 18 ([UNICEF, 2020](#), p. 4). While this connectivity is increasing globally, access is not equal as the digital divide furthers the disparities. Globally, 2.2 billion youth remain unconnected to the internet at home and access for youth to the internet is also exacerbated in rural communities worldwide, with 25% of rural youth connected in comparison to 41% of youth in urban areas ([UNICEF, 2020](#), p. 7). Additionally, African youth are more disconnected (60% unconnected) than children in Europe (4% unconnected) based on many factors such as: economics, gender, and language ([UNICEF, 2020](#), p. 2). More than half of all websites (56%) are only in English, inaccessible for most people on the planet ([UNICEF, 2017](#), p. 10).

Soon after the World Health Organization declared COVID-19 officially a global pandemic, [Guterres \(2020\)](#), the UN Secretary General, tweeted, "Our common enemy is #COVID19, but our enemy is also an 'infodemic' of misinformation." This was followed by two policy briefs from UNESCO warning the world about the dangers of a "disinfodemic" ([Posetti and Bontcheva, 2020a,b](#)). [UNESCO \(2020\)](#) asserts, "COVID-19 has led to a parallel pandemic of disinformation that directly impacts lives and livelihoods around the world. Falsehoods and misinformation have proven deadly and sowed confusion about life-saving personal and policy choices" (para. 1). At a time when the world needed reliable and factual information to save lives, our information systems were flooded with "fake news," "alternative facts," and conspiracy theories. While misinformation and disinformation are nothing new, the popularity and design of social media platforms have created the perfect storm for the most sensational messages to go viral instantly. At the same time, these digital tools are being used by individuals and organizations to connect and build movements for political and social transformation, the power of information communication technologies has never been greater.

Between 2015 and 2019, the amount of time young people in the United States (US) spent watching videos every day doubled (Rideout and Robb, 2019, p. 5). Almost half of US teenagers surveyed, report that “they are online on a near-constant basis” (Anderson and Jiang, 2018, para. 2). In 2016, Stanford researchers assessed 7804 students across the US on their competence to analyze online media and found it “bleak,” reporting that students are “easily duped” and unprepared to distinguish between news and advertising, or to judge the reliability of a website (p. 4). They assert, “Never have we had so much information at our fingertips. Whether this bounty will make us smarter and better informed or more ignorant and narrow-minded will depend on our awareness of this problem and our educational response to it” (Stanford History Education Group, 2016, p. 5).

A crisis of trust

While we have access to far more information than ever before, we are trusting it less as researchers at Reuters Institute for the Study of Journalism surveyed people in over 40 countries and found that “fewer than four in ten (38%) say they trust most news most of the time” (Newman et al., 2020, p. 14). The increase in disinformation and unregulated media platforms and the decrease in trust of news media and the ability to distinguish credibility, require a new pedagogical response. Educational systems need to shift from traditional ideas of literacy that focus on a standard national language, emphasize skills-based functional literacies and memorizing facts to critical media literacy that addresses multiple communication systems and increasing linguistic and cultural diversity. Part of the requirement for being literate in the 21st century is being able to read and write images, sounds, multimedia, advertising, social media, and numerous other “multiliteracies” (New London Group, 1996). In Australia, Luke and Freebody (1999) created the *Four Resources Model* that lists the competencies necessary for being literate in a multimodal world: breaking the code of the text, participating in the meanings of the text, using texts functionally, and critically analyzing and transforming the text. Vasquez (2003, p. 15) explains:

Luke and Freebody assert that reading should be seen as a nonneutral form of cultural practice, one that positions readers in advantageous and disadvantageous ways. They argue that readers need to be able to interrogate the assumptions and ideologies that are embedded in a text as well as the assumptions that they, as sociocultural beings, bring to the text. This leads to asking questions such as, Whose voice is heard? Who is silenced? Whose reality is presented? Whose reality is ignored? Who is advantaged? Who is disadvantaged? These sorts of questions open spaces for analyzing the discourses or ways of being that maintain certain social practices over others.

The digital networked age brings challenges for which most educational institutions are unprepared. The majority of educators receive little training about how to promote critical thinking with their students in ubiquitous online cultures. In this essay, we discuss the potential of critical media literacy as a pedagogical response to our present-day needs. Critical media literacy combines a body of knowledge with a set of skills and dispositions to empower youth to critically analyze media, information, technology, and create their own alternative messages. This is especially important for educators working with second language learners because they are navigating multiple linguistic identities in different geographic and virtual sociocultural contexts, with an array of media (Meyrer and Kersch, 2021). This involves redefining second language instruction to include listening, speaking, reading, and writing all types of texts (digital and print) and necessitates critical analysis to question the connections between information and power.

Critical media literacy

There are many types of media education, some more focused on arts integration, others aiming to protect students from unsolicited media messages, and some simply including more types of media into traditional teaching (Kellner and Share, 2007). While these approaches can be useful for increasing media literacy skills, teaching critical media literacy means developing educators and students who embrace a critical way of looking at the world and questioning systems of power. This requires building critical awareness of how information communication technologies function within society, teaching skills for deconstructing and creating media, and developing a disposition for participating in a democratic process. Critical media literacy involves questioning the politics of representation, challenging oppressive ideologies, and creating counter-hegemonic media representations. Vasquez (2017, p. 6) reminds us that critical literacy should not be just a skill or topic to teach:

Instead it should be looked on as a lens, frame, or perspective for teaching throughout the day, across the curriculum, and perhaps beyond. What this means is that critical literacy involves having a critical perspective or way of being.

Critical media literacy offers a theoretical framework and pedagogical practice for critically analyzing information and creating alternative responses. Evolving from the multidisciplinary field of cultural studies, critical media literacy aims to expand our understanding of literacy to include reading and writing all kinds of texts (from printed books to movies, music, memes, social media, video games, websites, etc.), as well as deepen analysis to more critical levels that question the relationships between media and audiences, information and power. Critical media literacy integrates a body of knowledge and set of skills, as well as a framework

Table 1 Critical media literacy framework (Kellner and Share, 2019, p. 8).

Conceptual understandings	Questions
1. Social constructivism All information is co-constructed by individuals and/or groups of people who make choices within social contexts.	WHO are all the possible people who made choices that helped create this text?
2. Languages/semiotics Each medium has its own language with specific grammar and semantics.	HOW was this text constructed and delivered/accessed?
3. Audience/positionality Individuals and groups understand media messages similarly and/or differently, depending on multiple contextual factors.	HOW could this text be understood differently?
4. Politics of representation Media messages and the medium through which they travel always have a bias and support and/or challenge dominant hierarchies of power, privilege, and pleasure.	WHAT values, points of view, and ideologies are represented or missing from this text or influenced by the medium?
5. Production/institutions All media texts have a purpose (often commercial or governmental) that is shaped by the creators and/or systems within which they operate.	WHY was this text created and/or shared?
6. Social & environmental justice Media culture is a terrain of struggle that perpetuates or challenges positive and/or negative ideas about people, groups, and issues; it is never neutral.	WHOM does this text advantage and/or disadvantage?

Available in English, Mandarin, Portuguese, Spanish online: <https://guides.library.ucla.edu/educ466>.

of conceptual understandings (Buckingham, 2003). Based on the work of scholars and organizations around the world, the following Critical Media Literacy Framework (Kellner and Share, 2019, p. 8) includes six conceptual understandings and corresponding questions (see Table 1).

These conceptual understandings and questions are intended to guide educators and students along an inquiry process to interrogate all texts, mediums of communication, and the context that surround them. Livingstone (2018) states, “media literacy is needed not only to engage *with the media* but to engage *with society through the media*” (para. 3). It is not enough to teach with media, we must also teach about media and acknowledge that teaching about media can be a highly complex and multifaceted undertaking, especially when done through a critical media literacy lens.

Recognizing the need to embrace sociology, Luke and Freebody (1997) suggest educators should use more than just a psychological approach to literacy, because reading and writing is not just an individual cognitive process. They recommend bringing a sociological lens into the process of questioning the contexts, the dominant ideologies, and the systems that make some things seem “natural” or “normal.” Kovach and Rosenstiel (2011) assert, “Asking questions begins the process of deconstructing the media content we see in front of us. Critical thinking is not a formula—it is a journey” (p. 210). This journey according to Freire, (1970) requires dialectical thought, in which “world and action are intimately interdependent” (p. 38). Freire refers to combining theory with practice as *praxis*. When critical media literacy teaches students how to analyze media messages as well as create their own media and critically participate in media culture, students demonstrate their ability to read and write the word and the world (Freire and Macedo, 1987).

The critical media literacy framework is a holistic tool for thinking about and questioning the dynamic role media play in our relationships with ourselves, each other, and society. Accompanying this framework is a definition of critical media literacy created by the steering committee of the Critical Media Literacy Conference of the Americas (2021):

The goal of critical media literacy is to engage with media through critically examining representations, systems, structures, ideologies, and power dynamics that shape and reproduce culture and society. It is an inquiry-based process for analyzing and creating media by interrogating the relationships between power and knowledge. Critical media literacy is a dialogical process for social and environmental justice that incorporates Freire’s (1970) notion of praxis, “reflection and action upon the world in order to transform it” (p. 36). This pedagogical project questions representations of class, gender, race, sexuality and other forms of identity and challenges media messages that reproduce oppression and discrimination. It celebrates positive representations and beneficial aspects of media while challenging problems and negative consequences, recognizing media are never neutral. Critical media literacy is a transformative pedagogy for developing and empowering critical, caring, nurturing, and conscientious people.

Much of the theory behind critical media literacy has developed from cultural studies, a field of critical inquiry that began in the 1900s in Europe and continues to expand with new critiques of media and society. Researchers at the Frankfurt Institute for Social Research used critical social theory to analyze how media culture and the tools of communication technology induce ideology and

social control (Kellner, 1989). In the 1960s, researchers at the Center for Contemporary Cultural Studies at the University of Birmingham added to the earlier concerns of ideology with a more nuanced understanding of the audience as active constructors of reality (Hall, 2002, 2019). Kellner (2020) explains that cultural studies has continued to grow and incorporate concepts of semiotics, multiculturalism, feminism, critical race theory, queer theory, postmodernism, and other critical theories. Incorporating a dialectical understanding of textual analysis, political economy, and audience theory, cultural studies critiques media culture as dynamic discourses that reproduce dominant ideologies as well as entertain, educate, and offer the possibilities for counter-hegemonic alternatives.

Language acquisition

When considering language development, it can be helpful to distinguish the differences between acquisition and learning. The traditional view of language learning comes from behavioral psychology and promotes direct instruction that breaks language into component parts through drills and practice to learn vocabulary and grammar rules. In contrast, language acquisition emerges from cognitive psychology and prioritizes comprehensible input, the use of language in meaningful contexts, and a focus on comprehension and expression over correctness and skills (Freeman and Freeman, 2004). These ideas have been greatly influenced by Krashen's (1992) research and hypotheses about language acquisition.

Two of Krashen's hypotheses that can be supported by critical media literacy address the need for input that is comprehensible and learning environments that are low in stress and high in motivation. Krashen (1992) asserts that second language learners acquire the new language best when they receive input that is just a little beyond their ability level. This is similar to Vygotsky's (1978) *zone of proximal development*, a sweet spot that is not too difficult that the language cannot be understood, but not too easy so that there is nothing new to acquire (Moll, 1990). The *input hypothesis* (Krashen, 1992) calls for teachers to provide *comprehensible input* in which students can understand the majority of what they read or hear. "To ensure that the input is comprehensible, teachers can use pictures, gestures, tone of voice, and hands-on activities" (Freeman and Freeman, 2004, p. 38). Other researchers have also emphasized the importance of comprehensible output, so that students are producing and using language in meaningful ways (Gibbons, 2009; Nowbakht and Shahnazari, 2015; Swain, 1985).

These ideas align well with critical media literacy when students are creating media that is meaningful to them, thereby increasing their multimodal skills, deepening their critical thinking, and acquiring new language in the process. Many of the rules and linguistic features of the target language are more difficult for emergent bilinguals to learn, than they are for native speakers who hear and use it repeatedly all around them (Gibbons, 2009; Krashen, 1992). Another hypothesis that Krashen (1992) argues is important for acquiring a new language is the need to lower students' *affective filters*. When students feel anxious, scared, bored, or lack self-confidence their affective filter rises and prevents them from being able to acquire new language. In contrast, when they feel confident, engaged, and motivated, their affective filter is lowered and they are more capable of developing language. Students acquire language best when they receive relevant comprehensible input in meaningful and supportive situations. This is enhanced when an asset-based approach is used to counteract the deficit thinking that too often pervades educational institutions, especially with second language learners.

Asset-based vs. deficit-based approaches

Discourse surrounding deficit-based and asset-based terminology of students engaging in second language learning and acquisition has been debated globally for several decades (Martínez, 2018). Deficit-based terminology often occurs for second language learners when educators impose preconceived biases or assumptions upon students, suggesting that non-native speakers enter school lacking linguistic, cultural, and economic norms, skills, and familial support for their education (González et al., 2005; Valencia, 2010; Yosso, 2005).

In the US, deficit frames are often embedded within formal education through common labels used in schools. For example, deficit terminology, such as: *Limited English Proficient* (LEP) and *Former Limited English Proficient* (FLEP) positions the student as deficient and the English language as the preferred normative language, thus negating the values tethered to other cultural, racial, and ethnic identities, histories and languages. Currently, the term *English Language Learner* (ELL) or *English Learner* (EL) is widely used in second language research and by practitioners, yet there are increasing concerns that this label prioritizes English as a monolingual ideal and "can distort how we make sense of bi/multilingual students' everyday language practices, or ways of using language" (Martínez, 2018, pp. 515–516). Deficit constructs disproportionately harm students who are most marginalized by falsely implying that non-dominant identities do not hold valuable forms of knowledge. In this way, viewing students through an English monolithic lens does not acknowledge students' vast community and cultural wealth that they bring to the classroom, especially their linguistic assets (Yosso, 2005). Yosso (2005) describes "linguistic capital" as students' "intellectual and social skills attained through communication experiences in more than one language/style" (p. 78).

Valencia and Solórzano (2004) explain three deficit-based ideologies: genetic pathology thesis, culture of poverty model, and inadequate parents, home, and child model. The genetic pathology thesis promotes the idea that certain groups (often reflecting white supremacy) are biologically superior and others are genetically inferior. The culture of poverty model blames students' culture and lifestyle suggesting that people experiencing poverty are less capable of succeeding in school and society because they are poor. Lastly, is the deficit idea that parents/guardians, families, and home cultures do not care about their child's educational

achievement. Deficit thinking reinforces the myths that blame second language learners' biology, culture, or families instead of recognizing the systemic and structural causes that lead to inequities.

An alternative to deficit thinking is asset-based thinking that prioritizes the positive skills, knowledge, and experiences that students bring with them that rarely are appreciated or valued in schools. [González et al. \(2005\)](#) describe a *funds of knowledge* approach in which teachers reject the deficit paradigm and take a proactive role in validating and incorporating the positive assets students and their families possess in their daily household routines. This practice creates an opportunity for educators to interrogate and reject deficit-based thinking, forge positive collaborative relationships with learners and their families, and can be applied in classrooms by fusing students' familial knowledge into their curricular and pedagogical designs ([Moll, 2019](#)). By integrating students' funds of knowledge into the curriculum, classrooms are more likely to become *culturally relevant* ([Ladson-Billings, 1995, 2014](#)) while also increasing the amount of comprehensible input ([Krashen, 1992](#)).

Asset-based terminology disrupts problematic labels for second language learners. [García \(2009\)](#) asserts that the term *emergent bilingual* (EB) is a more inclusive asset-based term that respects the learner's "potential in developing their bilingualism; it does not suggest a limitation or a problem in comparison to those who speak English" (p. 322). Instead, this terminology disrupts deficit categorizations of students as limited English proficient (LEP) to emphasize the positive attainment of a second language. Emergent bilingual students navigate the development of two languages simultaneously, including their home language and their second language during language acquisition ([Gort et al., 2008](#)). When educators embrace students as assets, rejecting deficit ideas, students are more likely to feel welcome, self-confident, and valued, thereby lowering their affective filters and increasing their potential for acquiring new language. When critical media literacy guides students to analyze their media preferences, express their voices, and create media that are meaningful to them, emergent bilingual learners are prone to expand their linguistic dynamics through multi-cultural discourse ([Sherman, 2003](#)).

Putting theory into practice

Media literacy education is increasing around the world in multiple content areas and grade-levels, thus becoming an opportune pedagogy for second language learners ([Lim and Tan, 2020](#)). Analyzing and creating media can be a useful strategy for emergent bilinguals to learn and acquire grammar, phonics, syntax, semantics, and pragmatics, as well as comprehension and critical thinking. When analysis and production are critical of power and dominant ideologies, students are more likely to question and challenge deficit thinking, unjust representations, and disinformation.

There are many ways for second language teachers to integrate critical media literacy into their practice in all grade levels. Young emergent bilingual students can compare and contrast different versions of fables and fairytales, questioning portrayals of gender, race, class, and other identities found in books, cartoons, movies, websites, songs, or video games. Then they can collaboratively create their own media with alternative perspectives or different endings of the same story, taking the form of comic strips, memes, podcasts, digital stories, or photographs. Older emergent bilingual students can analyze news media to learn about how production techniques convey ideologies, bias, and credibility through narrative, imagery, and sound. Then they can create their own text as radio shows, blogs, animation, zines, movie trailers, books, or social media, to retell or repurpose the story from different perspectives and question hegemonic narratives. These critical media literacy activities also provide pedagogical opportunities for emergent bilinguals by making input comprehensible, content culturally relevant, lowering the affective filter, and increasing language acquisition.

Photography

While there is little research available on using photography with young emergent bilinguals, it is important to recognize that "Language does not develop as an isolated mode of communication. Its relationship with visual imagery is primal" ([Britsch, 2010](#), p. 171). Students can use cameras to record their lives outside of school, as well as construct meanings of academic words and ideas they are learning in the classroom. Reflecting on over 30 years of experience teaching elementary school, [Dragan \(2008, p. 41\)](#) explains that for emergent bilinguals, creating photography:

gives them a feeling of power and control over a piece of equipment, and that helps compensate for the lack of control they may feel over not yet speaking English. Photography gives my English language learners (ELLs) an additional language—another way for them to convey who they are and show what they know.

In Los Angeles, California, emergent bilingual fourth-grade students learned how to illustrate their vocabulary words with photography ([Share, 2015](#)). First, they took pictures that showed words they encountered in their textbooks. The activity began with a class discussion about the meaning of the word and the different ways they could show it in a single image. The student photographer had to tell classmates where and how to pose in order for the picture to communicate the intended meaning of the word. They analyzed photographs from newspapers and magazines and brainstormed a list of techniques that photographers use to convey feelings and ideas, such as camera angles, composition, and lighting. After a couple weeks, students were so



Fig. 1 Fourth-grade second language learners at an elementary school in downtown Los Angeles took pictures of each other to illustrate vocabulary words and created this palm-sized booklet with adjectives and synonyms.

enthusiastic about taking pictures that they began bringing their own words into the classroom. This sparked an activity in which whoever brought in a new word would get to take the picture illustrating that word or be in the picture. Because of the enthusiasm, students spent time every morning illustrating vocabulary words and generating a massive collection of new words. The digital images became a synonym/antonym flashcard game. Students wrote sentences or stories to accompany their photographs and then printed PowerPoint handouts that were folded into pocket-sized books and trading cards for each student. These personalized vocabulary cards and booklets were popular with students and their families, as well as powerful student-made teaching materials to increase their vocabulary and literacy skills (Fig. 1). This use of photography for learning different subject matter expands the understanding of literacy and opens the opportunity to critically analyze visual texts using the critical media literacy framework (see Table 1). For students from marginalized groups who are often under-represented and misrepresented in mainstream media, becoming the creators of their representations can be a step in the process of challenging deficit thinking and enacting their power.

Curriculum interventions

Refugees, immigrants, and emergent bilingual youth, often experience schooling as an exclusionary space that reinforces deficit-frames and funnels them into vocational tracks that rarely capitalize on their assets (Muhammad, 2020; Pérez Huber and Muñoz, 2021; Valenzuela, 1999). At the same time, commercial media often reproduce racist, sexist, classist, and nativist ideologies of refugees and other marginalized groups, appropriating their cultures, denigrating their experiences, and reinforcing harmful stereotypes (Yosso, 2020). Researchers in the Netherlands examined a critical media literacy intervention with young migrants run by a Dutch participatory action research project. They report on the value of critical media literacy in which teachers and students co-created curriculum that was situated on students' interests, assets, and needs (Bruinenberg et al., 2021). Bruinenberg et al. (2021) report that their predetermined curriculum undervalued the knowledge about social media that the migrant students already had, often more sophisticated than the educators. This suggests the importance of an asset-based approach that validates students' funds of knowledge and incorporates it into the curriculum design and implementation, thus making learning more culturally relevant.

In 2017, Brazil enacted legislation to require media analysis studies for children attending compulsory education given increasing public concerns about competing news media messages (Lim and Tan, 2020). Meyer and Kersch (2021) provide examples of 3rd-year high school students in an ESL class in Brazil participating in a critical media literacy intervention. Concerned with emergent bilingual students' exposure to "fake news" about the pandemic spreading on social media, educators decided with their students to critically analyze a viral video on WhatsApp and research multiple news sources. Students collaborated in critical dialogs, compared and contrasted multiple news articles and videos from international online newspapers, and constructed alternative journalistic articles in response to their analyses of news information about COVID-19. Students' processes of analyzing and responding to news articles about the pandemic yielded opportunities for them to read, write, and collaborate (orally and in written-formats) to increase their second language skills in English while also sharing more critical information in their native language. Meyer and Kersch (2021) report that using authentic multimodal texts related to students' lives helped them feel more connected with their language learning and empowered to be co-constructors of the curriculum. Students also demonstrated deeper critical thinking through questioning subsequent news articles and social media.

For 5 months during the pandemic, researchers in the Ukraine implemented a media literacy experiment with 138 1st-year students online in a university English as a Foreign Language (EFL) program (Bilotserkovets et al., 2021). The 70 students in the experimental group received media literacy lessons that led them to analyze social media for their language instruction. A control group of 68 students used the traditional EFL textbook for their lessons and received no media literacy instruction. The data from questionnaires, tests, and observations led Bilotserkovets et al. (2021) to conclude that the experimental group outperformed the control group in their development of media literacy skills and English-language proficiency. Researchers attribute much of the improvement to an increase in motivation from using English in authentic communication practices with social media and interactive ICTs (Bilotserkovets et al., 2021, p. 265). This case study provides an example of how reaching beyond the standard EFL textbook and analyzing social media can be important steps in the process of making EFL learning more engaging and relevant to students' lives.

Research studies

In an English as a foreign language program at a university in South Korea, Kim (2018) examined oral proficiency development using two types of movie clips, with and without sound. During a 12-week media literacy intervention with 62 low-proficiency EFL students, one group engaged with silent movie clips and the other interacted with movie clips with sound. Students analyzed and constructed digital stories to practice their listening comprehension and oral speaking skills about each film clip and provided feedback to their peers on their oral fluency (e.g., grammar, pronunciation, and summary skills). Pre and post tests indicate that the student group watching movie clips without sound were more confident to think critically and creatively, yet their listening comprehension was limited. Conversely, the group assigned movie clips with sound reported that their listening comprehension increased, but their critical thinking and creativity was limited. Kim's (2018) study coincides with Krashen's (1992) recommendations for teachers to increase comprehensible input through visual literacy to support emergent bilinguals' oral proficiency development, sometimes with sound and sometimes without. Taking this media literacy project to a more critical application could include asking questions about power and the politics of representation, as well as having students create alternative media that challenge the norms and ideologies they were seeing.

In a case study at Taipei University, in Taiwan, with four EFL students, Kung (2016) investigated whether using news broadcasts could increase students' communicative competencies and media literacy. First, the researcher assessed students' prior knowledge about the target vocabulary words from the video they were about to watch. Next, the teacher provided direct vocabulary instruction prior to their interaction with a CNN newscast. Students engaged with the newscast twice and took notes about events, main ideas, and vocabulary-in-context. Lastly, students completed an immediate oral language assessment to recount what they observed in the newscast (e.g., content, key points with examples, and a reflection) by creating a digital recording. Data revealed that all participants increased their application of target vocabulary with mild-to-full accuracy and improved their oral linguistic output. Kung (2016) recommends that second language teachers incorporate media texts mindful of comprehensible input relevant to students' interests to expand students' "practical and genuine linguistic input" (p. 337). Although this study has a small sample size, it contributes to the scant literature in Southeast Asia relating to second language approaches to teaching with media. While teaching *with* media is part of critical media literacy, it is also important to teach *about* media, such as deconstructing production techniques, analyzing positionality and bias, and questioning economic motivation, thus reinforcing the understanding that no message is neutral (Kellner and Share, 2019).

Also analyzing news media, university students in a Teaching English as a Foreign Language class in Bogotá, Colombia, learned media literacy through weekly radio workshops (Neiva-Montaña, 2021). Working collaboratively for eight weeks, 15 students analyzed a radio program from the BBC, role played as journalists interviewing each other, developed a survey for their social circles, researched topics of interest, and then created their own radio programs that were broadcast on the university radio station. Neiva-Montaña (2021) gathered qualitative data to assess the learning and found that the media literacy work encouraged oral interactions and assisted the students in making sense of complex concepts. She reports, "Bringing media messages to the classroom activated the students' critical thinking and transformed the learning language into a more meaningful experience" (p. 14). As the students used language together to solve problems, they were also applying Mercer's (2007) concept of "interthinking." Mercer (2007, p. 6) writes:

Language is designed for doing something much more interesting than transmitting information accurately from one brain to another: it allows the mental resources of individuals to combine in a collective, communicative intelligence which enables people to make better sense of the world and to devise practical ways of dealing with it.

Collaborating to solve real-world problems, made the students' experiences more meaningful and engaging than simply filling out worksheets or memorizing decontextualized information as often occurs in traditional EFL classrooms. Through analyzing news media and creating their own radio programs, these students were demonstrating the potential of media literacy for language learning, something that could be even further enhanced through a critique of ideology and power.

After his sixth-grade intermediate English as Second Language (ESL) students complained about their ESL textbook being boring and irrelevant, Mohammed Choudhury departed from the text and taught his students language skills through critical media literacy: analyzing dominant commercial media, interviewing community experts, questioning hierarchies of power, and creating their own counter-narratives (Choudhury and Share, 2012). His students encountered many negative stereotypes in the mainstream press about their community and people like them. They discussed the racism and injustice they were finding and decided to do their own investigation about their community by focusing on the assets that were rarely reported.

They embarked on walking field trips around their school, taking pictures and interviewing community members. Some people they interviewed in their primary language (Spanish) and then translated the interviews later in the classroom. Other people they interviewed in English and worked collaboratively to write notes and document the conversations. Students compared the positive information they learned about their community with the negative representations they found in the local news media. In an attempt to challenge the deficit messages they wrote essays and created multimedia presentations that valued the assets. Choudhury reports that his students benefited in numerous ways from increased self-esteem and pride in their community to higher test scores and greater language acquisition. By the end of the school year, “more than half of his students, 47 out of 72 (65.3%), were reclassified as English-Fluent-Proficient, well above the district norm of 13.4% for the sixth grade” (p. 43). Not only did their language proficiency increase, but Choudhury’s students learned to use language to critically examine the systems and structures of power and then exercised their agency by creating media that challenged deficit thinking and celebrated their community’s assets.

Conclusion

Critical media literacy provides a framework, disposition, and pedagogy to critique the social structures, systems, and representations that organize society and reproduce hierarchies of power. While the critical media literacy framework and examples listed above can be useful resources for recreating this in classrooms across the globe, education is a sociocultural process and should always be adapted and shaped for the particular context and needs of each student and learning environment. Students should be guided to question injustice and challenge oppression through critical media analysis and production, varying depending on the context. Second language learning through critical media literacy offers many benefits for emergent bilingual students such as: increasing comprehensible input, lowering affective filters, raising student engagement, using language in meaningful contexts, developing critical thinking, and expanding language acquisition. When educators integrate critical media literacy with second language learning, they provide authentic opportunities for emergent bilinguals to connect their funds of knowledge with all types of media, deepen their critical engagement, and use all their language skills to create more socially just alternative messages.

References

- Anderson, M., Jiang, J., 2018. Teens, Social Media & Technology 2018. Pew Research Center, Washington, DC. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- Bilotserkovets, M., Fomenko, T., Gubina, O., Klochkova, T., Lytvynko, O., 2021. Fostering media literacy skills in the EFL virtual classroom: a case study in the COVID-19 lockdown period. *Int. J. Learn. Teach. Educ. Res.* 20 (2), 251–269.
- Britsch, S., 2010. Photo-booklets for English language learning: incorporating visual communication into early childhood teacher preparation. *Early Child. Educ. J.* 38 (3), 171–177.
- Bruinenberg, H., Sprenger, S., Omerović, E., Leurs, K., 2021. Practicing critical media literacy education with/for young migrants: lessons learned from a participatory action research project. *Int. Commun. Gaz.* 83 (1), 26–47.
- Buckingham, D., 2003. *Media Education: Literacy, Learning and Contemporary Culture*. Polity Press, Cambridge, MA.
- Choudhury, M., Share, J., 2012. Critical media literacy: a pedagogy for new literacies and urban youth. *Voices Middle* 19 (4), 39–44.
- Critical Media Literacy Conference of the Americas (CMLCA), 2021. Definition of Critical Media Literacy From the CMLCA 2021 Steering Committee. <https://guides.library.ucla.edu/educ466>.
- Dragan, P.B., 2008. *Kids, Cameras, and the Curriculum: Focusing on Learning in the Primary Grades*. Heinemann, Portsmouth, NH.
- Freeman, D.E., Freeman, Y.S., 2004. *Essential Linguistics: What You Need to Know to Teach Reading, ESL, Spelling, Phonics, and Grammar*. Heinemann, Portsmouth, NH.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York, NY.
- Freire, P., Macedo, D., 1987. *Literacy: Reading the Word and the World*. Bergin & Garvey, Westport, CN.
- García, O., 2009. Emergent bilinguals and TESOL: what’s in a name? *Tesol Q.* 43 (2), 322–326.
- Gibbons, P., 2009. *English Learners Academic Literacy and Thinking: Learning in the Challenge Zone*. Heinemann, Portsmouth, NH.
- González, N., Moll, L.C., Amanti, K. (Eds.), 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Gort, M., de Jong, E.J., Cobb, C.D., 2008. SEelng through a bilingual lens: structural and ideological contexts of structured English immersion in three Massachusetts districts. *J. Educ. Res. Pol. Stud.* 8 (2), 41–67.
- Guterres, A., 2020. Our Common Enemy is #COVID19, But Our Enemy is Also an “infodemic” of Misinformation. Twitter. <https://twitter.com/antonioGuterres/status/1243748397019992065>.
- Hall, S., 2002. Encoding/decoding. In: Durham, M.G., Kellner, D. (Eds.), *Media and Cultural Studies: KeyWorks*. Blackwell Publishers, Malden, MA, pp. 166–176.
- Hall, S., 2019. In: Moley, D. (Ed.), *Essential Essays (Two-Volume Set): Foundations of Cultural Studies & Identity and Diaspora*. Duke University Press, Durham, NC.
- Kellner, D., Share, J., 2007. Critical media literacy is not an option. *Learn. Inq.* 1 (1), 59–69.
- Kellner, D., Share, J., 2019. *The Critical Media Literacy Guide: Engaging Media and Transforming Education*. Brill/Sense Publishers, Leiden, The Netherlands.
- Kellner, D., 1989. *Critical Theory, Marxism, and Modernity*. John Hopkins University Press, Baltimore, MD.
- Kellner, D., 2020. Critical perspectives on television from the Frankfurt School to the politics of representation. In: Wasko, J., Meehan, E.R. (Eds.), *A Companion to Television*, second ed. Wiley-Blackwell, Hoboken, NJ, pp. 15–37.

- Kim, S., 2018. Improving low-proficiency Korean EFL learners' speaking skills using two different modes of movie clips: a media literacy perspective. *STEM J.* 19 (1), 129–151.
- Kovach, B., Rosenstiel, T., 2011. *Blur: How to Know What's True in the Age of Information Overload*. Bloomsbury, New York, NY.
- Krashen, S., 1992. *The Input Hypothesis: Issues and Implications*. Laredo Publications, Laredo, TX.
- Kung, F.W., 2016. Facilitating learners' second language communicative competence through the development of media literacy: a conversation analytic approach. *Asia-Pac. Educ. Res.* 25 (2), 337–346.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lim, S.S., Tan, K.R., 2020. Front liners fighting fake news: global perspectives on mobilising young people as media literacy advocates. *J. Child. Media* 14 (4), 529–535.
- Livingstone, S., 2018. Media Literacy—Everyone's Favourite Solution to the Problems of Regulation. The London School of Political Science and Economics. Available from: <https://blogs.lse.ac.uk/media/2018/05/08/media-literacy-everyones-favourite-solution-to-the-problems-of-regulation/>.
- Luke, A., Freebody, P., 1997. Shaping the social practices of reading. In: Muspratt, S., Luke, A., Freebody, P. (Eds.), *Constructing Critical Literacies: Teaching and Learning Textual Practice*, pp. 185–225. Sydney, Australia: Allen & Unwin and Cresskill, NJ: Hampton Press.
- Luke, A., Freebody, P., 1999. Further Notes on the Four Resources Model, pp. 1–4. Reading online.
- Martínez, R.A., 2018. Beyond the English learner label: recognizing the richness of bi/multilingual students' linguistic repertoires. *Read. Teach.* 71 (5), 515–522.
- McLuhan, M., 1997. *Understanding Media: The Extensions of Man*. MIT Press, Cambridge, MA.
- Mercer, N., 2007. *Words and Minds: How We Use Language to Think Together*. Routledge, London, England.
- Meyrer, K.P., Kersch, D.F., 2021. Can high school students check the veracity of information about COVID-19? A critical media literacy proposal in ESL classes. *J. Media Lit. Educ.* Pre-Prints. <https://digitalcommons.uri.edu/jmle-preprints/5>.
- Moll, L.C. (Ed.), 1990. *Vygotsky and Education: Instructional Implications and Applications of Sociocultural Psychology*. Cambridge University Press, Cambridge, UK.
- Moll, L., 2019. Elaborating funds of knowledge: community-oriented practices in international contexts. *Lit. Res.* 68, 130–138.
- Muhammad, G., 2020. *Cultivating Genius: An Equity Framework for Culturally and Historically Responsive Literacy*. Scholastic Incorporated, USA.
- Neiva-Montaño, H.K., 2021. Developing oral interaction skills in foreign language learners through media literacy. *Colomb. Appl. Linguist. J.* 23 (1), 3–16.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Newman, N., Fletcher, R., Schulz, A., Andi, S., Nielsen, R.K., 2020. Reuters Institute Digital News Report 2020. Reuters Institute for the Study of Journalism, Oxford, UK. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf.
- Nowbakht, M., Shahnazari, M., 2015. The comparative effects of comprehensible input, output and corrective feedback on the receptive acquisition of L2 vocabulary items. *Adv. Lang. Lit. Stud.* 6 (4), 103–114.
- Pérez Huber, L., Muñoz, S. (Eds.), 2021. *Why They Hate Us: How Racist Rhetoric Impacts Education*. Teachers College Press, New York, NY.
- Posetti, J., Bontcheva, K., 2020a. Disinfodemic: Deciphering COVID-19 Disinformation: Policy Brief 1. UNESCO, Paris, France. https://en.unesco.org/sites/default/files/disinfodemic_deciphering_covid19_disinformation.pdf.
- Posetti, J., Bontcheva, K., 2020b. Disinfodemic: Dissecting Responses to COVID-19 Disinformation: Policy Brief 2. UNESCO, Paris, France. https://en.unesco.org/sites/default/files/disinfodemic_dissecting_responses_covid19_disinformation.pdf.
- Rideout, V., Robb, M.B., 2019. *The Common Sense Census: Media Use by Tweens and Teens, 2019*. Common Sense Media, San Francisco, CA. <https://www.commonsensemedia.org/sites/default/files/uploads/research/2019-census-8-t0-18-full-report-updated.pdf>.
- Share, J., 2015. *Media Literacy is Elementary: Teaching Youth to Critically Read and Create Media*, second ed. Peter Lang, New York, NY.
- Sherman, J., 2003. In: Thornbury, S., Ur, P. (Eds.), *Using Authentic Video in the Language Classroom: Cambridge Handbooks for Language Teachers*. Cambridge University Press, New York.
- Stanford History Education Group, 2016. Evaluating Information: The Cornerstone of Civic Online Reasoning. Executive Summary. <https://stacks.stanford.edu/file/druid:fv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf>.
- Swain, M., 1985. Communicative competence: some roles of comprehensible input and comprehensible output in its development. In: Gass, S.M., Madden, C. (Eds.), *Input in Second Language Acquisition*. Newbury House, Rowley, MA, pp. 235–253.
- UNESCO, 2020. *Combating the Disinfodemic: Working for Truth in the Time of COVID-19*. <https://en.unesco.org/covid19/disinfodemic>.
- UNICEF, 2017. *The State of the World's Children 2017: Children in a Digital World*. United Nations Children's Fund, New York, NY, USA.
- UNICEF, 2020. *How Many Children and Young People Have Internet Access at Home? Estimating Digital Connectivity During the COVID-19 Pandemic*. United Nations Children's Fund and International Telecommunication Union, New York.
- Valencia, R.R., Solórzano, D., 2004. Today's deficit thinking about the education of minority students. In: Santa Ana, O. (Ed.), *Tongue-Tied: The Lives of Multilingual Children in Public Education*. Rowman & Littlefield, Lanham, MD, pp. 124–133.
- Valencia, R.R., 2010. *Dismantling Contemporary Deficit Thinking: Educational Thought and Practice*. Taylor and Francis, New York, NY.
- Valenzuela, A., 1999. *Subtractive Schooling: U.S.-Mexican Youth and the Politics of Caring*. State University of New York Press, Albany, NY.
- Vasquez, V.M., 2003. Getting Beyond "I like the Book": Creating Space for Critical Literacy in K-6 Classrooms. International Reading Association, Newark, DE.
- Vasquez, V.M., 2017. Critical literacy. In: *Oxford Research Encyclopedia of Education*.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Yosso, T.J., 2020. Critical race media literacy for these urgent times. *Int. J. Multicult. Educ.* 22 (2), 5–13.

Further reading

- Cubbage, J. (Ed.), 2022. *Critical Race Media Literacy: Themes and Strategies for Media Education*. Routledge, New York, NY.
- Funk, S., Kellner, D., Share, J., 2016. Critical media literacy as transformative pedagogy. In: Yildiz, M.N., Keengwe, J. (Eds.), *Handbook of Research on Media Literacy in the Digital Age*. IGI Global, Hershey, PA, pp. 1–30.
- Goodman, S., 2003. *Teaching Youth Media: A Critical Guide to Literacy, Video Production, and Social Change*. Teachers College Press, New York, NY.
- Hammer, R., Kellner, D. (Eds.), 2009. *Media/Cultural Studies: Critical Approaches*. Peter Lang, New York, NY.
- López, A., 2021. *Ecomedia Literacy: Integrating Ecology Into Media Education*. Routledge, New York, NY.
- Morrell, E., Dueñas, R., García, V., López, J., 2013. *Critical Media Pedagogy: Teaching for Achievement in City Schools*. Teachers College Press, New York, NY.
- Noble, S.U., 2018. *Algorithms of Oppression*. New York University Press, New York, NY.
- Santa Ana, O. (Ed.), 2004. *Tongue-Tied: The Lives of Multilingual Children in Public Education*. Rowman & Littlefield, Lanham, MD.
- Vasquez, V.M., 2014. *Negotiating Critical Literacies With Young Children*, tenth anniversary ed. Routledge, New York, NY.

Relevant websites

Center for Critical Internet Inquiry, <https://www.c2i2.ucla.edu/>.

Critical Media Literacy Framework in English, <https://drive.google.com/file/d/1VJE2yVMJ0JdohiXwB9ti4UEQPYPfBbEjc/view>.

Critical Media Literacy Research Guide, <https://guides.library.ucla.edu/educ466>.

Critical Media Project, <https://criticalmediaproject.org/>.

International Council for Media Literacy, <https://www.nationaltelemediacouncil.org/>.

Assessment of English learner students

Jamal Abedi and Cecilia Sanchez, University of California, UC Davis School of Education, Davis, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

ELs' performance on state content assessments	86
Reliability and validity of state content assessments	87
Reliability and validity of other standardized assessments	87
Accessibility of state content assessments	88
Accommodations for ELs on state content assessments	88
First language translation	88
Glossary	89
Read-aloud	89
Linguistic modification	89
Current issues	89
Unnecessary linguistic complexity	90
Need for evidence-based accommodations	91
Need for guidelines on assigning accommodations	92
Are ELs one homogenous subgroup?	92
Summary and discussion	93
References	93
List of relevant websites	95

The English learner (EL) population continues to grow across the United States (US). Although EL populations are highly concentrated in states like California, Texas, and New Mexico, data reveals that EL populations are growing across most states each school year (US Department of Education et al., 2020). According to the US Department of Education et al. (2020) ELs made up 10.2% of students in public K-12 schools in 2018, which is about five million students; meanwhile, back in 2000 ELs made up 8.1% of students (US Department of Education et al., 2020).

Due to federally mandated statewide testing and the Every Student Succeeds Act (ESSA, 2015), educators are held accountable to ensure that ELs participate in the assessments and that ELs' academic performance improves over the years. However, since the inclusion of ELs in assessments many decades ago, a performance gap between ELs and their non-EL peers has been observed (Abedi and Sanchez, 2021). Changes to assessments have steadily increased the accessibility of assessments for ELs; however, there are still issues with current assessment practices because ELs continue to struggle to score at the same level as their peers. ELs underperform even in content assessments not related to English language skills (Abedi and Gándara, 2006) or when assigned test accommodations intended to increase the accessibility of the tests (Abedi, 2012; Abedi et al., 2020). With the growing EL population in mind, the need for reliable and valid assessments that are accessible to ELs across the nation increases.

In this chapter the authors will first summarize key aspects of ELs' performance on state content assessments including the reliability, validity, and accessibility of the tests. Next, the authors will discuss issues related to ELs and state content assessments. Although there are various aspects of statewide assessments pertaining to ELs that can be discussed, the authors have chosen three main issues related to the assessment of the EL student population: unnecessary linguistic complexity in content-based assessments such as math and science; the need for effective, valid, and research-based accommodations specifically adapted to assist the language needs of ELs; and the need for guidelines on how to assign accommodations to ELs. The authors will also discuss the validity of the dichotomous EL/non-EL classification. By targeting these main issues, ELs may improve their performance on statewide content assessments and, most importantly, ELs' content knowledge may be more accurately assessed and evaluated.

ELs' performance on state content assessments

ELs tend to perform lower on assessments with higher language demands (Abedi, 2002). The 2019 National Assessment of Educational Progress (NAEP, 2021a) scores revealed a statistically significant 24-point gap between grade 4 ELs and non-ELs on the math section, with ELs performing lower than their non-EL peers. NAEP grade 4 math test results showed that 44% of non-ELs score proficient or above, while only 16% of ELs score proficient or above (NAEP, 2021b). On the reading section, ELs significantly underperform at a higher degree than in math. To compare, NAEP grade 4 reading results revealed a 33-point gap difference between ELs and non-ELs in grade 4 (NAEP, 2021c). This resulted in 39% of non-ELs scoring proficient or above on the NAEP reading sections, while only 10% of ELs scoring proficient or above (NAEP, 2021d). The clear difference between ELs' scores on the math and readings sections reveals the way ELs' limited English proficiency affects test performance.

Various studies have also found that non-ELs outperform ELs on content assessments. For example, [Turken and Liu \(2012\)](#) found that this is true on a science test administered to non-ELs and ELs in California. The researchers found a moderate performance difference between the two EL groups that proved to be statistically significant. To add, [Miley and Farmer \(2017\)](#) compared the state standardized assessment scores between ELs and non-ELs in Tennessee and found similar findings where ELs significantly underperform on both English language arts (ELA) and math sections. The researchers took it one step further by creating a third student subgroup, recently reclassified ELs, which was indicated by exit criteria on the state's English language proficiency assessment. [Miley and Farmer \(2017\)](#) found that recently reclassified ELs underperform compared to their non-EL peers. Similarly, data from California shows that although reclassified students perform higher than current ELs, reclassified students still do "not outperform English-only students" ([Hill, 2018](#)). It is clear that students' level of English proficiency influences test scores even on subject areas not directly assessing English language skills.

Reliability and validity of state content assessments

Examining test reliability and validity is essential for all state content assessments, but especially when drawing conclusions about students belonging to minority groups, such as ELs. Test reliability is defined as "a measure of the extent to which test scores are consistent across testing conditions, such as across different test items or upon retest" ([Moore et al., 2020](#)). Whereas test validity is defined as the degree to which a test measures the targeted construct of the test. When examining the reliability and validity of assessments for ELs it is important to assess the way in which the language of the test plays a role, because language as a source of construct irrelevant factor may have a substantial impact on assessments.

Some studies show a high significant difference in the reliability of assessment results between EL and non-ELs. In a technical brief on ELs' scores on the ACT, the researchers found that "the reliability of ELs' scores were consistently lower than those of non-ELs" ([Moore et al., 2020](#)). [Moore et al. \(2020\)](#) add that these reliability issues between ELs and non-ELs are also observed in other standardized tests like the Partnership for Assessment of Readiness for College and Careers ([PARCC, 2017](#)) and the Smarter Balanced Assessment Consortium ([SBAC, 2016](#)). Reports on EL's performance on standardized tests make it clear that components with a high level of linguistic complexity can lower reliability estimates of an assessment ([Moore et al., 2020](#)). To add, [Noble et al. \(2014\)](#) found that ELs were more likely to answer incorrectly on science test items than their non-EL peers, despite demonstrating knowledge of the science content assessed in the test item. When interviewing EL students, the researchers found that ELs misunderstood the test items because of the way that the questions were worded. For example, the test question "If enough heat is taken away from a container of water, what will happen to the water?" caused some ELs to interpret the question as "What would happen when you heat the container of water?" or that the water was taking heat from a heat source ([Noble et al., 2014](#)).

It is also clear that language plays a factor in the reliability of these assessments for ELs. For example, [Abedi \(2002\)](#) found that ELs consistently underperform in areas such as ELA, math, and science; however, the performance gap widens in areas with higher language demand. Similarly, [Abedi et al. \(2003\)](#) found that as students' level of English proficiency decreased so did the internal consistency, or how reliable the test measure is across items, of the test, which demonstrates how language affects ELs' test performance. Therefore, the reliability of the assessment is not only affected by EL status but also by level of English proficiency.

[Huang and Bulter \(2020\)](#) explored four empirical studies focused on the validity of assessments for language minority students in four different parts of the world and concluded that special attention to the specific needs of the language minority group being assessed is required when developing assessments. This is especially true when the assessments hold high stakes as statewide content assessments do for ELs in the US. [Abedi \(2006a\)](#) states that although there are various factors that can affect the way all students perform on a test (i.e., the test's validity), factors related to unnecessary linguistic complexity most affect students who are not proficient in English since their limited English skills hinders their ability to perform on tests. Statewide tests in the US are presented in English, which is the language that the EL is not yet proficient in, and therefore contain unnecessary linguistic complexities such as construct-irrelevant variance (i.e., language that is not related to the focal construct) that threaten the validity of the assessments ([Abedi, 2006a](#)).

Reliability and validity of other standardized assessments

Although this chapter mainly focuses on statewide assessments in the US, the authors also wanted to comment on the reliability and validity of other types of standardized tests used across the world, as they are often used to make inferences about student subgroups such as ELs. Specifically, tests like Peabody Picture Vocabulary Test-4 (PPVT-4; [Dunn and Dunn, 2007](#)) and the Structured Photographic Expressive Language Test: Third Edition (SPELT-3; [Dawson et al., 2005](#)) will be explored below because these tests have proven to be reliable and valid measures for the general student population. For example, [Pearson \(2007\)](#), the publishers of PPVT-4, report high levels of reliability in forms of internal consistency and test-retest and high levels of validity. SPELT-3 also showed high levels of reliability and validity in its development ([Dawson et al., 2003](#)). However, careful examination is needed when considering if the same level of reliability and validity can be applied to ELs.

Researchers believe that PPVT-4 holds bias against ELs and is not a reliable measure for this student group. [Goriot et al. \(2018\)](#) explored the reliability of PPVT-4 for ELs whose first language is Dutch and found that the test was most reliable for students who answered 24 or more items correct. The researchers conclude that PPVT-4 is most suitable for ELs that have higher levels of English proficiency. Another study compared the performance difference between students whose first language was either French or Polish and found that PPVT-4 and similar vocabulary-based tests could be biased based on students' first language ([Leśniewska et al., 2018](#)). Additionally, [Finneran et al. \(2019\)](#) conducted a study to explore the performance differences across African American and Hispanic students (majority of which were ELs) in the US and found that despite both groups were from low-income communities, PPVT-4 had potential cultural-linguistic biases that did not favor Hispanic students.

Research on SPELT-3 is much more limited than PPVT-4, especially on student subgroups such as second language learners in other countries. In the US SPELT-3 has shown applicability due to its high accuracy in differentiating between students with and without language impairments ([Perona et al., 2005](#)). The version for preschoolers (called SPELT-Preschool 2; [Dawson et al., 2005](#)) is highly correlated to another expressive grammar test, which strengthens its case for strong concurrent validity and diagnostic accuracy ([Greenslade et al., 2009](#)). Another study examining SPELT-P2 also supports the use of this test as a way of determining language disabilities in preschool-aged students; however, the researchers clearly mention that the participants only spoke English and that those that indicated that they spoke another language at home (i.e., ELs) were excluded due to "concerns about the validity of the criterion measures for nonmainstream speakers" ([Eisenberg et al., 2019](#)). The way in which ELs have been excluded demonstrates the need for research exploring the reliability and validity of these tests specifically for ELs.

Accessibility of state content assessments

Assessments are said to be "accessible" to ELs if all extraneous variables that are not related to the focal construct are controlled. The assessments can be constructed based on "accessibility" guidelines or if relevant accommodations are included in the assessment in order to minimize the impact of extraneous variables, such as the impact of cultural biases and language complexity. For example, the work of the Race to the Top consortia, which was funded by the [American Recovery and Reinvestment Act \(2009\)](#) in the US, involved SBAC and PARCC and led to more accessible and valid assessments for the nation ([US Department of Education et al., 2013](#)). The consortia paid special attention to the research-based decision process on the concept of "accessibility."

Accommodations for ELs on state content assessments

Accommodations that are effective in making assessments more accessible for ELs can be used as accessibility features and can become part of a standardized test administration process, the same is true for students with disabilities (SWD). [Abedi \(2012\)](#) states that accommodations for ELs must be effective, valid, relevant, feasible, and sensitive to students' background characteristics. If accommodations meet these criteria, then assessments will become more accessible for ELs ([Abedi, 2017](#)).

These accommodations can be grouped under two categories: (1) accommodations that may have no conceivable impact on the focal construct (e.g., extra time or test administration directions that are simplified), and (2) accommodations that may have some direct or indirect influence on language as the focal construct (e.g., commercial dictionary). Accommodations under the first category can be used for ELs and SWD as appropriate accommodations (since they are not providing an unfair advantage to the recipients); or as accessibility features for all because everyone may benefit from these features. However, accommodations under the second category can be considered as accessibility features when used for all students since using them for a certain group of students may not be a good practice as they may provide unfair advantages to the recipients.

Accommodations for ELs can also be divided between those that provide direct and indirect linguistic support ([Shafer Willner et al., 2008](#)). The researchers simply differentiate between the two types of accommodations by defining direct linguistic support accommodations as modifications related to the language of the test (examples thoroughly described later), meanwhile the researchers define indirect linguistic support accommodations as modifications not related to the language of the test (e.g., extra time, breaks, changes to test setting such as small group and familiar environment). This chapter will focus on direct linguistic support since the main goal of language-based accommodations is to increase the linguistic accessibility of assessments. As [Sireci and Faulkner-Bond \(2015\)](#) point out, one way to increase the validity of statewide content assessments for ELs is to provide accommodations that may reduce the linguistic barrier. Below is a quick description and brief literature review on four language-based accommodations that have been researched in hopes of making assessments more accessible for EL students: first language translation, glossary, read-aloud, and linguistic modification.

First language translation

The first language translation accommodation provides a translated version of each test item in the students' first language. Some research reveals that this accommodation is most effective for students with high language skills in their first language ([Abedi and Ewers, 2013](#)). For example, [Robinson \(2010\)](#) found that Spanish-speaking ELs in kindergarten and first grade scored significantly higher on Spanish versions of a math test. However, other research shows that this accommodation is not effective. In a recent study, [Abedi et al. \(2020\)](#) conducted a randomized controlled trial with this accommodation, in which only students who performed at a certain level on a Spanish language assessment were randomly offered one of two Spanish-based accommodations, one of which was a Spanish translated version. The researchers found that students score significantly lower than students who took the same test without an accommodation. [Abedi \(2017\)](#) points a major flaw with this accommodation, as there is potential for lack of alignment between the students' first language (i.e., the language of the translated test) and the language of instruction.

Glossary

The glossary accommodation provides simplified definitions of vocabulary terms not related to the focal construct of the test item. On computerized tests students can hover over the selected term and then the glossary item will pop up when clicked on. An earlier version of this accommodation was the dictionary accommodation which provided dictionary-based definitions; however, glossaries seem to be more effective than dictionary accommodations (Abedi and Ewers, 2013).

A metaanalysis by Abedi and Ewers (2013) revealed that the glossary-based accommodation showed some promise but needs more research to examine its effectiveness. Chia (2014) also called for more research in the use of glossary accommodations, as further guidance is needed on which terms should have a glossary and the amount of cognitive load associated with the additional information in glossaries. A recent study by Cohen et al. (2020) found that computer-based, pop-up glossaries were effective on both ELA and math tests and did not alter the focal construct. Similar to Chia (2014), Cohen et al. (2020) also noted the trouble with determining which words should have a glossary. To address this issue, Cohen et al. (2020) used a carefully formulated coding procedure to determine which words required a glossary. Additionally, Chia (2014) listed categories of terms that SBAC has integrated into their glossary-based accommodations including: “nominalizations, false cognates, informal English used in an academic context, prepositional phrases that can have multiple meanings, and relative pronouns directly related to the task at hand.” Research-based guidelines on how to determine which terms need a glossary could help in determining how to make this accommodation most effective for ELs.

Read-aloud

The read-aloud accommodation provides students with a word for word read-aloud of the test items by person or electronic media. As Abedi and Ewers (2013) describe, there are various types of read-aloud accommodations such as read aloud of test directions in students’ first language and read aloud of test questions in English. The researchers’ metaanalysis reveals inconsistent effectiveness across ELs based on level of English proficiency, subject area, and grade level. Sireci Faulkner-Bond’s (2015) metaanalysis also reports on the mixed results on the effectiveness of this accommodation. Specifically, Wolf et al. (2012) studied this accommodation in two states and found that the read-aloud accommodation resulted in no effect for ELs in one state and a significant positive effect for ELs in the other state.

Linguistic modification

The linguistic modification accommodation reduces the linguistic demand on tests by minimizing the level of linguistic complexity and removing all unnecessary linguistic factors unrelated to the content assessed. A major aspect of the linguistic modification accommodation is removing unnecessary linguistic complexity from test items. The accommodation involves modifying linguistic features that may be appropriate for non-ELs but complex for ELs. Abedi (2006b) breaks down linguistic features that could be modified to make test items more accessible to ELs and thus increase the chances of correctly measuring ELs’ content knowledge. These features are related to word familiarity; word, sentence, nominal, and question phrase length; voice of verb phrase (i.e., active vs. passive); structure related to sentence, discourse, and comparative; subordinate, conditional and relative clauses; negation; and concrete statements. Below is an example from Abedi (2006b) that applies linguistic modification to transform an abstract statement into concrete terms:

Original: The weights of three objects were compared using a pan balance. Two comparisons were made

Revision: Sandra compared the weights of three objects using a pan balance. She made two comparisons

In this example, the revised version contains active voice verbs, while the original contains passive voice verbs (Abedi, 2006b). Although at times the linguistic modification accommodation is referred to as “linguistic simplification,” Abedi et al. (2020) disagrees this term as it implies that the linguistic complexity is simply reduced and does not take into account that unnecessary linguistic factors should be removed. Overall, research on this accommodation is mixed (Abedi and Ewers, 2013). More information on unnecessary linguistic complexity and its use in the linguistic modification accommodation will be discussed below.

Current issues

This section of the chapter will focus on three main issues related to ELs’ performance on state tests including: (1) unnecessary linguistic complexity in content-based assessments such as math and science; (2) the need for effective, valid, and researched-based accommodations specifically adapted to assist the language needs of ELs; and (3) the need for guidelines on how to assign accommodations to ELs. The issue with analyzing ELs’ performance with ELs as one homogeneous group will also be discussed.

Unnecessary linguistic complexity

Unnecessary linguistic complexity can serve as a construct irrelevant factor and has a considerable impact on the reliability and validity of assessments for ELs. The effect of unnecessary linguistic complexity has been studied in ELs for decades now and research points to overall mixed results for its use as an accommodation (Abedi and Ewers, 2013). An issue with removing unnecessary linguistic complexity from test items is determining the linguistic features that are unnecessary to measure the focal construct. However, it appears that more recently developed tests have taken great efforts to remove unnecessary linguistic complexity for all students to make the content being assessed more straightforward.

Proof for the effectiveness of the linguistic modification accommodation has been shown in some research. In the early 2000s, Abedi and Lord (2001) modified the language demand on a math test by reducing the level of linguistic complexity and found that there was a statistically significant improvement for ELs on linguistically modified test items. The researchers explain that in interviews students expressed preference for items that were linguistically modified. Abedi and Lord (2001) found that EL students performed almost the same as their non-EL peers in a mathematics problem-solving test where there was minimal or no unnecessary linguistic complexity as compared to reasoning items with complex linguistic structure. Below is an example of a math question in its original form followed by a modified version with decreased linguistic complexity from Abedi and Soto (2008).

Original:

Alba needed to know about how much the sum of 19.6, 23.8, and 38.4 is. She correctly rounded each of these numbers to the nearest whole number. What three numbers did she use?

- A. 19, 23, 38
- B. 19, 24, 38
- C. 20, 24, 38
- D. 20, 24, 39

Revision:

Look at the numbers below.

19.6 23.8 38.4

Which list shows each number rounded to the nearest whole number?

- A. 19 23 38
- B. 19 24 38
- C. 20 24 38
- D. 20 24 39

In this example linguistic modification was used by removing information not needed to solve the question, changing the verbs to present tense, and using less complex sentence structures (Abedi and Soto, 2008).

Other research also shows support for the use of the linguistic modification accommodation for ELs. For example, Wolf and Leon (2009) investigated the association between language complexity and degree of differential item functioning (DIF; i.e., “when the probability of answering an item correctly is different between the two groups of interest who have the same ability level”) for ELs and found that as the level of academic vocabulary and amount of language increased so did DIF statistics for ELs. Similarly, Martiniello (2009) analyzed ELs’ math test performance and found that ELs are more likely to underperform on linguistically complex items when compared to their non-EL peers. To add, Keh (2019) assessed 11 ELs in a series of three assessments and compared their scores to existing data for English proficient students. Based on an analysis of the ELs’ performance, the researcher believes that modifications to linguistic aspects (e.g., orthography, phonology, syntax, and vocabulary) of the tests could improve ELs’ performance.

Meanwhile, other research shows mixed support for the use of linguistic modification as an accommodation. For example, although ELs scored slightly better on the linguistically modified version of a math test, Abedi et al.’s (2020) study did not show that linguistic complexity significantly affected ELs’ test scores. Similarly, a report published by the Educational Testing Service led by Young et al. (2014) found mixed results for the use of linguistic modification as an accommodation and add that at times the scores of both non-ELs and ELs improved, worsened, or did not change on science items with linguistic modifications. Since one of the criteria to claim the effectiveness of an accommodation is that it results in an increase in performance for the target group (ELs) and no change in performance for the non-target group (non-ELs), Young et al.’s (2014) research shows that this accommodation did not cause a differential impact for ELs’ test performance.

A recent study by Noble et al. (2020) found that the effects of the linguistic modification accommodation rely on the specific modifications made to specific test questions, meaning that the type of modifications performed by researchers and test developers must be carefully executed to determine the accommodation’s true effects on ELs’ performance. However, like other studies mentioned, the researchers noted no statistically significant results although there were some differences across the test performance of the EL groups. Although there are mixed results on the effectiveness of the linguistic modification accommodation, the way in which there is some level of differential performance across ELs groups reveals that this accommodation has potential to increase the

accessibility of assessments for ELs. Future research could work on developing a universal method to remove unnecessary linguistic components as mentioned by Noble et al. (2020). Future research could also ensure experimental controlled design and sufficient sample sizes to generate results that are valid and reliable.

An issue with the concept of unnecessary linguistic complexity is defining language as related or unrelated to the content being measured in order to determine how to remove the complex linguistic structure from the test items without altering the focal construct. Based on literature, this decision can be made collectively by a group of experts including linguists, test developers, and researchers who focus on student assessment practices (Abedi, 2015). Together this multidisciplinary team can determine the language factors in the test items that could be considered not directly related to the focal construct (i.e., “construct irrelevant”). However, consensus across multidisciplinary teams is ideal in order to develop established guidelines on how to properly remove unnecessary linguistic complexity from test items.

A challenge with using linguistic modification as an accommodation for ELs is that level of complexity should not alone determine inclusion of a linguistic feature (Avenia-Tapper and Llosa, 2015). Instead, Avenia-Tapper and Llosa (2015) argue that linguistic modification should focus on the degree of relevance language terms to the content being assessed. To add, the researchers state that the level of construct-irrelevance associated with linguistic complexity should be determined by examining the “the degree of correspondence between language use in the assessment and language use in the educational contexts in which the content is learned and used.” The researchers implicate that test developers should utilize the language educators use to teach ELs science knowledge concepts, which will encourage the use of grade-appropriate science language. In support of this argument, familiarity is listed as a factor to consider when modifying linguistic demand on assessments for ELs. Various researchers who include linguistic modification in EL accommodation research like Abedi (2006b) support the idea that unfamiliar vocabulary and linguistic features should be removed and that terms used to teach students new material (i.e., familiar) should remain in assessments.

It seems to be that over the past years, unnecessary linguistic demands have been altered to focus on the content knowledge being assessed and test developers have adapted linguistic modification as a standard practice. Given that the linguistic modification accommodation in Abedi et al.’s (2020) study did not show to be significantly effective, the researchers believe that “given the literature on the general effectiveness of linguistic modifications, [the researchers] surmise that newer assessments, including those developed for [the] study, have been linguistically modified to the point that further linguistic modification has only a limited effect.” In fact, Sireci and Faulkner-Bond (2015) once argued, that instead of considering linguistic modification as an accommodation, it should be “part of the standard test development process.”

An overall finding from literature on ELs is that language proficiency does interfere with ELs’ ability to perform on English-language assessments. As demonstrated in a study conducted by Buono and Jang (2021), the researchers found that out of 11 math test items that had complex linguistic features, DIF analyses showed that ELs underperformed when compared to non-ELs on seven of the items. Researchers agree that language does interfere with ELs’ performance on test items not intended to measure English language proficiency; however, a conclusion on how to overcome the issue with unnecessary linguistic complexity is not yet resolved (Noble et al., 2020). Given that ELs and non-ELs continue to perform differently on statewide assessments and that the effectiveness of this accommodation has not yet been ruled out, it is worthwhile including it in future research as it has potential in making assessments more accessible for ELs.

Need for evidence-based accommodations

As mentioned earlier, one of the issues with research on accommodations is that metaanalyses reveal mixed results (Abedi and Ewers, 2013); therefore, a definite answer as to whether an accommodation is effective is not fully understood. Like all other research attempting to generalize their results, these studies should include a large enough EL sample size and provide statistically significant findings. To add, accommodations should be assigned based on evidence-based research and depend on each student’s needs. The main conclusion on current research on accommodations for ELs is that there is inadequate empirical data to thoroughly determine the effects of accommodations for ELs (Abedi, 2017).

Both a lack of statistically significant score differences between students receiving accommodations and no accommodations and small EL sample sizes are issues in accommodation research for ELs. For example, Rios et al. (2020) examined accommodations offered to ELs in a metaanalysis and found no statistically significant results to support the use of accommodations, although differences in test scores were observed. To add, in a randomized controlled study, Abedi et al. (2020) explored ELs’ performance on an assessment with randomly assigned accommodations and found that none of the accommodations significantly improved EL test performance. It is important to note that Abedi et al. (2020) encountered issues with a small sample size when only a couple of students were randomly assigned into the Spanish-based accommodations being evaluated. This occurred because not many students scored high enough on a Spanish language assessment to receive a Spanish accommodation (Abedi et al., 2020). A small sample size in research analyzing the effects of accommodations on ELs appears to be a recurring issue as Yeo (2010) state that research lacks studies with large EL sample sizes, which are necessary to generalize results.

Instead of attempting to guess which accommodation is best suited for ELs, appropriate accommodations should have an explicit rationale for their use. For example, a student whose first language is Spanish should not necessarily be given a test in Spanish; rather the accommodation should depend on the language of instruction and other relevant background variables (Francis et al., 2006). When accommodations were first assigned to ELs the accommodations came from a set of accommodations offered to SWD (e.g., testing students in multiple sessions or providing breaks) and have not proven to be significantly helpful for ELs (Abedi, 2017). It is not surprising to find that research points to a need for studies that assign accommodations based on students’ needs so

that specific needs are met with appropriate accommodations (Abedi and Ewers, 2013). A hypothetical example of a specific need could include variations in the level of English proficiency, such that a student with moderate-to-high English oral skills but low English written skills would receive a read-aloud accommodation.

Need for guidelines on assigning accommodations

Hand in hand with the need for evidence-based accommodations that have been proven to be effective for ELs is the need for a set of guidelines on how to assign accommodations to ELs. Some obstacles to developing a set of guidelines is that there is variability across states and within districts, students' unfamiliarity with accommodations, and mixed results in research that studies accommodations for ELs. In the past decades more ELs have gained access to accommodations for statewide assessments, however since statewide tests continue to hold high stakes it is important that accommodations on these assessments are properly assigned to ELs.

The regulation for assigning accommodation differs across the nation and across tests, so there is great variability in the guidelines on how to assign accommodations to ELs. Each state selects which statewide assessment will be implemented (e.g., some states use SBAC and other states like Texas develop and use their own state assessments) and the way that accommodations are assigned, which has both positive and negative implications. On the positive side, local control of accommodation assignment allows for regulations affecting different groups of ELs. This means that districts with large numbers of ELs with distinct backgrounds such as newly arrived or high Spanish-speaking can implement regulations educational leaders believe is best suited for their group of students. Local control also allows for quick changes since the regulations are not under widespread control, which gives educators the flexibility to quickly adapt the latest research findings into their accommodation assignment. However, on the negative side, establishing rules for assigning accommodations based on county lines or other geological factors means that location influences accommodation decisions instead of basing the decision solely on student needs. This also implies that the number of resources a school has available will determine which accommodations a student can receive. Differences in accommodation assignment creates inconsistency across the nation, which means that when ELs change schools they may be assigned a different accommodation despite having the same needs that they had at their previous school.

It is also important that students not only have access to accommodations while taking statewide tests, but also during classroom achievement tests. As mentioned by Abedi et al. (2020) students should be familiar with accommodations before being expected to use them in assessments. This way students are accustomed to using the accommodation and it does not cause confusion during a test, when students may feel additional stress. This point is also reiterated in interviews by Wolf et al. (2012) where students expressed unfamiliarity with a glossary-based accommodation and thus did not use it during the assessment. On computer-based assessments students' use of accommodations can be measured by the number of times the student clicks on items such as a pop-up glossary or a recording for a read-aloud accommodation. Computer-based assessments and tools that measure the amount of accommodation-use could give insight into how much students are utilizing the accommodations, if at all. Researchers could use these tools in future studies to explore the relationship between the amount of exposure to an accommodation and use of the accommodation during tests and then make inferences on the way in which these factors influence test scores.

As stated by Rios et al. (2020) recommendations on assigning accommodations to ELs are minimal. A main factor that may lead to minimal and unclear guidelines is the need for evidence-based accommodations that show that accommodations are effective in making assessments more accessible to ELs. However, despite the issues with educators not having a clear set of guidelines on how to assign accommodations to ELs, the use of accommodations in ELs is increasing. Rios et al. (2020) mentioned that despite the inconsistencies, at least one test accommodation was assigned for the ELA and math sections across distinct state testing programs. Chia (2014) explains SBAC's application and use of several accommodations in their computer-based tests and summarized leading research for each accommodation. Researchers should ensure that results of their studies on accommodations for ELs can be accessed by test developers and educators so that guidelines based on the latest research can be created, shared, and implemented.

Are ELs one homogenous subgroup?

Recent research points out the fact that EL status is not permanent, in that students can place out of EL status and become reclassified as English proficient and thus transition from the "EL" to the "non-EL" subgroup (Saunders and Marcelletti, 2013). The fluidity of the EL classification can create misleading conclusions when ELs' progress on achievement tests is observed over time. The category of "ever-EL" may help move researchers and educators away from misleading conclusions. However, there are also a vast number of differences in students within those that are not English-only speakers which need to be considered.

The main issue with analyzing student assessment progress with only two EL subgroups, EL and non-EL, is that ELs with a higher English language proficiency (ELP) score become non-ELs, which leaves ELs with low ELP scores in the EL category and will result in misleading comparisons of the two groups over time (Umansky et al., 2017). This has proven to be an issue when analyzing performance gaps between ELs and non-ELs across grade levels. For example, as listed earlier, the performance gap on grade 4 math and reading tests were 24 and 33 points, respectively (NAEP, 2021a,c); however, if these scores are compared to the performance gap on grade 8 math and reading tests a 42-point and 45-point difference would be observed (NAEP, 2021e,f). A misleading conclusion when comparing ELs' performance gaps from grades 4 and 8 is that as ELs advance in school, the performance gap between ELs and non-ELs widens. However, this conclusion is based on the misleading assumption that the ELs in grade 4 and grade 8 are the same group of students with the same background characteristics—which is not the case since ELs can become non-ELs.

The ever-EL subgroup helps with the issue of misleading conclusions when a dichotomous EL/non-EL categorization is used. The ever-EL category groups ELs who have reached English proficiency according to their state's ELP assessment and current ELs (i.e., non-native English speakers who have not yet reached English proficiency on their state's ELP assessment). The ever-EL categorization does not separate non-native English speakers based on ELP scores, which gives more insight into how this group of students achieve academically on state tests and therefore, this categorization can be used when observing their progression on state-tests over time. Saunders and Marcelletti (2013) add that the EL/non-EL categorization can also decrease the chances of detecting positive improvements in achievements within non-native English speakers. For example, Betts et al. (2019) found that the percent of students that met or exceeded ELA standards on the 2018 SBAC test varied across EL groups (ever-ELs: ~45%; reclassified students: >60%; English only students: ~60%) especially when examining current ELs (>10%). Therefore, the ever-EL alternative breakdown of EL students challenges the narrative that ELs are low achievers, when in fact most EL students are placed into the non-EL subgroup as they advance in school.

Another reason why the dichotomous EL/non-EL categorization should be reexamined is because there is a vast amount of diversity within the EL population that could be influencing ELs' performance on content assessments. Sandberg and Reschly (2011) summarize that "low educational attainment and socioeconomic [SES] conditions of the family, as well as recency of immigration, often place ELs at increased risk for low achievement, poor educational outcomes, and limited economic opportunities." Sireci and Faulkner-Bond (2015) add that ELs also differ in their home language and level of literacy and education in their first language. Studies and reports that separate ELs into distinct groups show differences in assessment scores based on different variables like SES. For example, Luo et al. (2020) found that dual language learners from higher-SES backgrounds (which was measured by the level of education of students' primary caregiver) demonstrated higher language learning skills than dual language learners from lower-SES backgrounds. It is also likely that ELs' age of entry into the US school system influences their performance on statewide assessments, as they are likely to have lower levels of English proficiency. Also, it is likely that students who enter the US school system at a later age will face more developmental barriers to learning a new language like English than students who enter at an earlier age (Conger, 2009). Given that the EL population is so diverse, it is important to consider ELs' background when analyzing their content-based assessment scores.

Summary and discussion

In this chapter the authors reviewed research on elementary school aged ELs' performance on content-based assessments in the US and provided recommendations on how to improve the quality of assessments for ELs. First, the authors summarize the reliability, validity, and accessibility of state content assessments for ELs. Next, three issues related to the assessment of ELs were discussed including: unnecessary linguistic complexity; the need for effective, valid, and researched-based accommodations specifically adapted to assist the language needs of ELs; and the need for guidelines on assigning accommodations to ELs. The authors recognize vast improvements that have been made to assess ELs but also identify areas of improvement to make assessments more accessible and more accurately measure ELs' knowledge.

Literature on the assessment of ELs clearly show that this student group consistently underperforms when compared to their native English-speaking peers in almost all academic subjects (Abedi, 2015). However, the performance-gap between the two groups is more profound in content areas with higher level of linguistic complexity; that may not even be related to the focal construct (Abedi and Lord, 2001). Unnecessary linguistic complexity may have serious impact on the validity and reliability of content-based assessments for ELs, as it impacts on the content, construct, and concurrent validity indexes. Rios et al. (2020) mention that educational stakeholders should be cautious when drawing conclusions about ELs since language threats the reliability and validity of assessments.

By reviewing literature on the main language-based accommodations (such as first language translations, glossaries, read-aloud, and linguistic modifications), it is clear that mixed research results have made it difficult for researchers to provide educators with established recommendations or guidelines on which accommodations are best suited for ELs. The main way that assessments can be improved to be more inclusive of ELs is to provide research-based and effective accommodations that target ELs' limited English language proficiency. In addition to effective accommodations, educators should receive proper guidelines on the application and use of the accommodations so that they are equally accessible to ELs. When conducting research and analyzing results, it is important to consider that the EL category is not fixed and that ELs often become non-ELs as they advance in school. By considering differences between ever-ELs and current ELs, researchers can avoid making misleading conclusions about non-native English speakers. With these improvements, the performance gap between ELs and non-ELs can be minimized on statewide assessments in the US and, most importantly, ELs' content knowledge may be more accurately assessed and evaluated.

References

- Abedi, J., Ewers, N., 2013. Smarter Balanced Assessment Consortium: Accommodations for English Language Learners and Students With Disabilities: A Research-Based Decision Algorithm. Smarter Balanced Assessment Consortium, Olympia, WA. Available from: <https://portal.smarterbalanced.org/library/en/accommodations-for-english-language-learners-and-students-with-disabilities-a-research-based-decision-algorithm.pdf>. (Accessed 29 August 2021).

- Abedi, J., Gándara, P., 2006. Performance of English language learners as a subgroup in large-scale assessment: interaction and policy. *Educ. Meas.* 25, 36–46. <https://doi.org/10.1111/j.1745-3992.2006.00077.x>.
- Abedi, J., Lord, C., 2001. The language factor in mathematics tests. *Appl. Meas. Educ.* 14, 219–234. https://doi.org/10.1207/S15324818AME1403_2.
- Abedi, J., Sanchez, C., 2021. Historical milestones in the assessment of English learners. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement: Key Advancements in Theory, Policy, and Practice*. Routledge, London, ISBN 9780367415754, pp. 87–110.
- Abedi, J., Soto, E., 2008. "Linguistic Modification", Limited English Proficiency (LEP) Partnership Meeting. US Department of Education, Washington, DC, pp. 15–16. Available from: https://ncela.ed.gov/files/uploads/11/abedi_sato.pdf. (Accessed 17 February 2022).
- Abedi, J., Leon, S., Mirocha, J., 2003. Impact of Students' Language Background on Content Based Assessment: Analyses of Extant Data (Technical Report No. 603). University of California, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA. Available from: <http://cresst.org/publications/cresst-publication-2975/>. (Accessed 29 August 2021).
- Abedi, J., Zhang, Y., Rowe, S.E., Lee, H., 2020. Examining effectiveness and validity of accommodations for English language learners in mathematics: an evidence-based computer accommodation decision system. *Educ. Meas.* 39, 41–52. <https://doi.org/10.1111/emip.12328>.
- Abedi, J., 2002. Standardized achievement tests and English language learners: psychometrics issues. *Educ. Assess.* 8, 231–257. https://doi.org/10.1207/S15326977EA0803_02.
- Abedi, J., 2006a. Psychometric issues in the ELL assessment and special education eligibility. *Teach. Coll. Rec.* 108, 2282–2303. <https://doi.org/10.1111/J.1467-9620.2006.00782.X>.
- Abedi, J., 2006b. Language issues in item development. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey, ISBN 9780805852653, pp. 377–398.
- Abedi, J., 2012. Validity issues in designing accommodations. In: Fulcher, G., Davidson, F. (Eds.), *The Routledge Handbook of Language Testing*. Routledge, Taylor & Francis Group, London, ISBN 9781138205369.
- Abedi, J., 2015. Language issues in item-development. In: Lane, S., Raymond, M.S., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, Florence, KY, ISBN 9780415626026.
- Abedi, J., 2017. Utilizing accommodations in assessment. In: Shohamy, R., Or, I.G., May, S. (Eds.), *Language Testing and Assessment*. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-02261-1>.
- American Recovery and Reinvestment Act, 2009. Public Law 111-5. Available from: <https://www.congress.gov/bill/111th-congress/house-bill/1>. (Accessed 17 March 2022).
- Avenia-Tapper, B., Llosa, L., 2015. Construct relevant or irrelevant? The role of linguistic complexity in the assessment of English language learners' science knowledge. *Educ. Assess.* 20, 95–111. <https://doi.org/10.1080/10627197.2015.1028622>.
- Betts, J., Hill, L., Bachofer, K., Hayes, J., Lee, A., Zau, A., 2019. English Learner Trajectories and Reclassification. Public Institute of California. Available from: <https://www.ppic.org/publication/english-learner-trajectories-and-reclassification/>. (Accessed 27 August 2021).
- Buono, S., Jang, E.E., 2021. The effect of linguistic factors on assessment of English language learners' mathematical ability: a differential item functioning analysis. *Educ. Assess.* 26, 125–144. <https://doi.org/10.1080/10627197.2020.1858783>.
- Chia, M.Y., 2014. Content assessment aligned to the Common Core State Standards: improving validity and fairness for English language learners. *Appl. Meas. Educ.* 27, 307–312. <https://doi.org/10.1080/08957347.2014.944307>.
- Cohen, D.J., Ballman, A., Rijmen, Cohen, J., 2020. On the reliable identification and effectiveness of computer-based, pop-up glossaries in large-scale assessments. *Appl. Meas. Educ.* 33, 378–389. <https://doi.org/10.1080/08957347.2020.1789137>.
- Conger, D., 2009. Testing, time limits, and English learners: does age of school entry affect how quickly students can learn English? *Soc. Sci. Res.* 38, 383–396. <https://doi.org/10.1016/j.ssresearch.2008.08.002>.
- Dawson, J.I., Stout, C.E., Eyer, J.A., 2003. SPELT-3: Structured Photographic Expressive Language Test, third ed. Available from: <https://www.proedinc.com/Products/12243/spelt3-structured-photographic-expressive-language-testthird-edition.aspx>. (Accessed 16 February 2022).
- Dawson, J.I., Stout, C.E., Eyer, J.A., Tattersall, P., Fonkalsrud, J., Croley, K., 2005. Structured Photographic Expressive Language Test-Preschool 2. Available from: <https://www.janellepublications.com/0707.shtml>. (Accessed 16 February 2022).
- Dunn, L.M., Dunn, D.M., 2007. Peabody Picture Vocabulary Test-4. Available from: <https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Brief/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html>. (Accessed 15 February 2022).
- Eisenberg, S., Victorino, K., Murray, S., 2019. Concurrent validity of the fluently preschool speech and language screening test—second edition at age 3: comparison with four diagnostic measures. *Lang. Speech Hear. Serv. Sch.* 50, 673–682.
- Every Student Succeeds Act, 2015. Public Law 114-95. Available from: <https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>. (Accessed 17 March 2022).
- Finneran, D.A., Hellmann, J.J., Jones Moyle, M., Chen, S., 2019. An examination of cultural-linguistic influences on PPVT-4 performance in African American and Hispanic preschoolers from low-income communities. *Clin. Linguist. Phon.* 3, 242–255. <https://doi.org/10.1080/02699206.2019.1628811>.
- Francis, D., Rivera, M., Lesaux, N., Kieffer, M., Rivera, H., 2006. Practical Guidelines for the Education of English Language Learners: Research-Based Recommendations for the Use of Accommodations in Large-Scale Assessments. RMC Research Corporation, Center on Instruction, Portsmouth, NH. Available from: <https://www.carnegie.org/publications/practical-guidelines-for-the-education-of-english-language-learners-research-based-recommendations-for-the-use-of-accommodations-in-large-scale-assessments/>. (Accessed 26 August 2021).
- Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J.M., Unsworth, S., 2018. Using the peabody picture vocabulary test in L2 children and adolescents: effects of L1. *Int. J. Biling. Educ. Biling.* 24, 546–568. <https://doi.org/10.1080/13670050.2018.1494131>.
- Greenslade, K., Plante, E., Vance, R., 2009. The diagnostic accuracy and construction validity of the structured photographic expressive language test—preschool. second ed. *Lang. Speech Hear. Serv. Sch.* 40, 150–160.
- Hill, L., 2018. K–12 Reforms and California's English Learner Achievement Gap. Public Policy Institute of California. Available from: <https://www.ppic.org/publication/k-12-reforms-and-californias-english-learner-achievement-gap/>. (Accessed 29 August 2021).
- Huang, B.H., Butler, Y.G., 2020. Validity considerations for assessing language proficiency in young language minority students. *Lang. Assess. Q.* 17, 461–466. <https://doi.org/10.1080/15434303.2020.1826486>.
- Keh, M.L., 2019. Keeping language in mind: an exploratory study of English learners' performance on three language and literacy assessments. *CATESOL J.* 30, 79–99.
- Leśniewska, J., Pichette, F., Béland, S., 2018. First language test bias? Comparing French-speaking and polish-speaking participants' performance on the peabody picture vocabulary test. *Can. Mod. Lang. Rev.* 74, 27–52. <https://doi.org/10.3138/cmlr.3670>.
- Luo, R., Pace, A., Levine, D., Iglesias, A., de Villiers, J., Golinkoff, R.M., Wilson, M.S., Hirsh-Pasek, K., 2020. Home literacy environment and existing knowledge mediate the link between socioeconomic status and language learning skills in dual language learners. *Early Child. Res. Q.* 55, 1–14. <https://doi.org/10.1016/j.ecresq.2020.10.007>.
- Martiniello, M., 2009. Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educ. Assess.* 14, 160–179. <https://doi.org/10.1080/10627190903422906>.
- Miley, S.K., Farmer, A., 2017. English language proficiency and content assessment performance: a comparison of English learners and native English speakers achievement. *Engl. Lang. Teach.* 10, 198–207. <https://doi.org/10.5539/elt.v10n9p198>.
- Moore, J.L., Li, T., Lu, Y., 2020. Reliability of English Learners' Test Scores. ACT Research & Policy: Technical Brief. Available from: <https://www.act.org/content/dam/act/unsecured/documents/R1754-el-score-reliability-2020-05.pdf>. (Accessed 29 August 2021).
- National Assessment of Educational Progress, 2021a. Math: National Student Group Scores and Score Gaps. Available from: <https://www.nationsreportcard.gov/mathematics/nation/groups/?grade=4>. (Accessed 28 July 2021).

- National Assessment of Educational Progress, 2021b. Math: National Achievement-Level Results. Available from. <https://www.nationsreportcard.gov/mathematics/nation/achievement/?grade=4>. (Accessed 28 July 2021).
- National Assessment of Educational Progress, 2021c. Reading: National Student Group Scores and Score Gaps. Available from. <https://www.nationsreportcard.gov/reading/nation/groups/?grade=4>. (Accessed 18 August 2021).
- National Assessment of Educational Progress, 2021d. Reading: National Achievement-Level Results. Available from. <https://www.nationsreportcard.gov/reading/nation/achievement/?grade=4>. (Accessed 18 August 2021).
- National Assessment of Educational Progress, 2021e. Reading: National Student Group Scores and Score Gaps. Available from. <https://www.nationsreportcard.gov/mathematics/nation/groups/?grade=8>. (Accessed 27 August 2021).
- National Assessment of Educational Progress, 2021f. Reading: National Student Group Scores and Score Gaps. Available from. <https://www.nationsreportcard.gov/reading/nation/groups/?grade=8>. (Accessed 27 August 2021).
- Noble, T., Roseberry, R., Suarez, C., Warren, B., O'Connor, M.C., 2014. Science assessments and English language learners: validity evidence based on response processes. *Appl. Meas. Educ.* 27, 248–260. <https://doi.org/10.1080/08957347.2014.944309>.
- Noble, T., Sireci, S.G., Wells, C.S., 2020. Targeted linguistic simplification of science test items for English learners. *Am. Educ. Res. J.* 57, 2175–2209. <https://doi.org/10.3102/0002831220905562>.
- Partnership for Assessment of Readiness for College and Careers (PARCC), 2017. Final Technical Report for 2016 Administration. Available from. <https://parcc-assessment.org/wp-content/uploads/2018/02/PARCC-2016-Tech-Report.pdf>. (Accessed 14 February 2022).
- Pearson, 2007. Technical Information: PPVT-4 and EVT-2 Infographic. Available from. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/evt-2/ppvt-4-evt-2-infographic.pdf>. (Accessed 16 February 2022).
- Perona, K., Plante, E., Vance, R., 2005. Diagnostic accuracy of the Structured Photographic Expressive Language Test: third edition (SPELT-3). *Lang. Speech Hear. Serv. Sch.* 36, 103–115. [https://doi.org/10.1044/0161-1461\(2005\)010](https://doi.org/10.1044/0161-1461(2005)010).
- Rios, J., Ihlenfeldt, S.D., Chavez, C., 2020. Are accommodations for English learners on state accountability assessments evidence-based? A multistudy systematic review and meta-analysis. *Educ. Meas.* 39, 65–75. <https://doi.org/10.1111/emip.12337>.
- Robinson, J.P., 2010. The effects of test translation on young English learners' mathematics performance. *Educ. Res.* 39 (8), 582–590. <https://doi.org/10.3102/0013189X10389811>.
- Sandberg, K.L., Reschly, A.L., 2011. English learners: challenges in assessment and the promise of curriculum-based measurement. *Remedial Special Educ.* 32, 144–154. <https://doi.org/10.1177/0741932510361260>.
- Saunders, W.M., Marcelletti, D.J., 2013. The gap that can't go away: the catch-22 of reclassification in monitoring the progress of English learners. *Educ. Eval. Pol. Anal.* 35, 139–156. <https://doi.org/10.3102/0162373712461849>.
- Shafer Willner, L., Rivera, C., Acosta, B., 2008. Descriptive Study of State Assessment Policies for Accommodating English Language Learners. The George Washington University Center for Equity and Excellence in Education, Arlington, VA. Available from. <https://files.eric.ed.gov/fulltext/ED539753.pdf>. (Accessed 17 August 2021).
- Sireci, S., Faulkner-Bond, M., 2015. Promoting validity in the assessment of English learners. *Rev. Res. Educ.* 39, 215–252. <https://doi.org/10.3102/0091732X14557003>.
- Smarter Balanced Assessment Consortium (SBAC), 2016. Smarter Balanced Assessment Consortium: 2014–15 Technical Report. Available from. <https://portal.smarterbalanced.org/library/en/2014-15-technical-report.pdf>. (Accessed 14 February 2022).
- Turken, S., Liu, O.L., 2012. Differential performance by English language learners on an inquiry-based science assessment. *Int. J. Sci. Educ.* 34, 2343–2369. <https://doi.org/10.1080/09500693.2012.705046>.
- Umansky, I.M., Thompson, K.D., Diaz, G., 2017. Using an ever-English learner framework to examine disproportionality in special education. *Except. Child.* 84, 76–96. <https://doi.org/10.1177/0014402917707470>.
- US Department of Education, Office of the Deputy Secretary, Implementation and Support Unit, 2013. Race to the Top Assessment: Partnership for Assessment of Readiness for College and Careers [PARCC] Year Two Report. US Department of Education, Washington DC. Available from. www2.ed.gov/programs/racetothetop-assessment/performance.html. (Accessed 3 September 2021).
- US Department of Education, National Center for Education Statistics, Common Core of Data, 2020. Local Education Agency Universe Survey 2000–01 Through 2018–19. Available from. https://nces.ed.gov/programs/digest/d20/tables/dt20_204.20.asp. (Accessed 28 July 2021).
- Wolf, M.K., Leon, S., 2009. An investigation of the language demands in content assessments for English language learners. *Educ. Assess.* 14, 139–159. <https://doi.org/10.1080/10627190903425883>.
- Wolf, M.K., Kim, J., Kao, J., 2012. The effects of glossary and read-aloud accommodations on English language learners' performance on a mathematics assessment. *Appl. Meas. Educ.* 25, 9347–9373. <https://doi.org/10.1080/08957347.2012.714693>.
- Yeo, S., 2010. Predicting performance on state achievement tests using curriculum-based measurement in reading: a multilevel meta-analysis. *Remedial Special Educ.* 31, 412–422. <https://doi.org/10.1177/0741932508327463>.
- Young, J.W., King, T.C., Hauck, M.C., Ginsburgh, M., Kottloff, L., Cabrera, J., Cavalie, C., 2014. Improving content assessment for English language learners: studies of the linguistic modification of test items. *ETS Res. Rep. Ser.* 2014, 1–79. <https://doi.org/10.1002/ets2.12023>.

List of relevant websites

Abedi, J., English Language Learner Accommodation Project Led. ELLACProject.net.

Assessment of young emergent bilingual children

Sandra Barrueco, Maria C. Morales, and Lindsay M. Truesdale, Department of Psychology, The Catholic University of America, Washington, DC, United States

© 2023 Elsevier Ltd. All rights reserved.

Assessment of young emergent bilingual children	96
Call for comprehensive assessments of young emergent bilingual children	96
Language and early literacy development	97
Parent and teacher reports	97
Observational approach	98
Direct language assessment	98
Literacy assessment	99
Cognitive development	99
Social and emotional development	100
Approaches to learning	100
Physical development	100
Family and early care experiences	100
Societal and policy contexts	101
Conclusion	102
References	102

Assessment of young emergent bilingual children

The United States continues to experience dramatic increases in its cultural and linguistic diversity, including the burgeoning Latinx community (Bauman, 2017). Linguistically, Hispanic students constitute about 80% of emergent bilinguals in the US, who are children whose English language proficiencies are in its steps toward full proficiency (National Center for Education Statistics, 2022). Emergent bilingual (EB) children is a term that has become increasingly utilized to describe English language learners (ELLs; García et al., 2008), and used henceforth in this chapter. Emergent bilinguals (EBs) may have previously developed linguistic skills in another language before English, though this is not always the case. Comparatively, dual language learners (DLLs) is a broader term identifying children who may have varying levels of proficiencies within and across their languages, including potentially highly developed English proficiency.

The age of acquisition and rate of acquirement vary among emergent bilinguals, and are generally classified in two principal modalities: simultaneous learning or sequential learning (McLaughlin, 1984). Simultaneous learners are exposed to both languages early in life and sequential learners have undergone strides in their first language proficiency before learning the second language, typically upon entering school systems around the kindergarten years (Guzman-Orth et al., 2017). In the assessment field, measures and approaches are relatively more developed with school-aged sequential EB learners in comparison to their younger simultaneous EB peers. While the assessment landscape is rapidly broadening for EB infants, toddlers, and preschoolers, further developments are necessary as described in detail in this chapter.

Call for comprehensive assessments of young emergent bilingual children

Across educational and medical settings, screenings and assessments of emergent bilingual children are conducted to determine eligibility for a myriad of services and opportunities. These include prevention, early intervention, specialized instruction, and treatments (e.g., special education, targeted school-based English language programs, gifted and talented, magnet schools, therapies, and more). Appropriate and accurate screenings and assessments are imperative for young EB children due to their role in qualifying for supports, services, and interventions.

While language development is an essential consideration in the assessment of young emergent bilingual children, it is not singularly sufficient. Rather, a holistic understanding of young EB children is vital to accurately understand the progress and course of linguistic development (both within and across languages), along with development in other important domains. Even within a linguistically-focused assessment of young emergent bilingual children, there are differences in exposure, language usage, and proficiency that must be considered within the scope of a comprehensive assessment (Espinosa and López, 2007).

Indeed, young EB children are in fact children; they experience variation across their development much like other children. As such, full screenings and assessments across developmental domains are more likely to yield increased precise estimations of young EB children's functioning. Furthermore, they lead to more valid courses of actions in medical, familial, and scholastic settings (e.g., at the student-level for individualized plans, at the class-level for curricular decisions, and at the district-level for policy

determinations). Otherwise, there is both the risk, and the reality, of under- and over-identification of abilities and disorders (Bedore and Peña, 2008).

Fig. 1 below highlights key domains of comprehensive screening and assessment considerations for young Emergent Bilingual (EB) children. The first five are consistent with the central developmental domains of the Head Start Early Learning Outcomes Framework (U.S. DHHS, 2015): (a) linguistic and early literacy development, (b) cognitive development, (c) social and emotional development, (d) approaches to learning, and (e) physical development. The sixth and seventh domains reflect the examination of environmental features that consistently influence the development of young EB children: (f) family and early care experiences and (g) societal and policy contexts. Each of these domains is described below, with focus on expressly salient considerations such as the assessment of early linguistic and early literacy development.

Language and early literacy development

Engaging in multi-modal evaluations of young emergent bilingual (EB) children's language development is central for an accurate assessment (Lazewnik et al., 2019; Mancilla-Martinez et al., 2018). Well-established and empirically supported methods have been created that can, and should, be utilized when evaluating young EB children. These include parent and teacher reports, language sampling, nonword repetition, dynamic assessment, and standardized measures, among others.

Parent and teacher reports

Parent and teacher reports that are administered via questionnaires or interviews incorporate adult observations and knowledge of young children's linguistic skills and developmental trajectories. Assessment tools like the *Bilingual English-Spanish Assessment* (BESA; Peña et al., 2018) contain a parent interview section that can be used for this purpose. In turn, teacher reports on young children's language use have been developed to identify relative levels of linguistic proficiency and subsequent steps for bilingual assessment, such as language routing (Barrueco et al., 2017). In addition, the CECER-DLL *Child and Family and Teacher Questionnaires* incorporate both parental and teacher reports. It was developed with extensive input, feedback, and analysis from researchers and practitioners across the United States (Castro et al., 2020).

Monitoring vocabulary progress is also valuable for estimating the linguistic skills of young EB children over time, particularly since some aspects of academic development, such as reading comprehension, are linked to vocabulary development (Mancilla-Martinez et al., 2016, 2018). Among very young emergent bilinguals, measures are available to examine early expressive and receptive linguistic skills (including vocabulary), such as the *MacArthur-Bates Communicative Development Inventory* (CDI) and its Spanish counterpart the *MacArthur Inventario del Desarrollo de Habilidades Comunicativas* (IDHC) (respectively, Fenson et al., 2007; Jackson-Maldonado et al., 2013). Short forms have become validated and available over time, which facilitate their implementation across clinical, school, and research settings (Fenson et al., 2000).

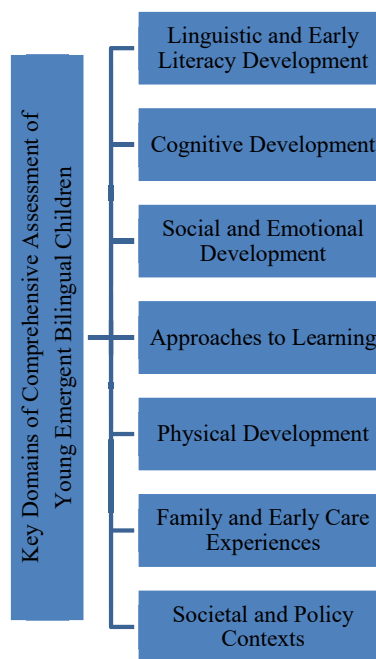


Fig. 1 Key domains of comprehensive screening and assessments for young emergent bilingual (EB) children.

Observational approach

Language sampling is an observational approach that can additionally be incorporated in young EB children's linguistic assessment. Using this procedure, children's utterances, articulation, and communicative language are observed and documented by professionals. A wealth of information is obtained during language sampling regarding conversational and expressive language skills. However, it should not be used in isolation because it does not account for the broader repertoire of receptive language skills, which are critical foundations for language development among bilingually developing children (Lazewnik et al., 2019). As such, language sampling acts as a component within a broader system of linguistic measures, allowing for the analysis of bilingual children's language use.

Direct language assessment

To directly assess linguistic development through structured interactions with young emergent bilingual children, various approaches include nonword repetition, dynamic assessment, and the broader array of standardized measures (Lazewnik et al., 2019). Repetition tasks allow for the observation of phonological working memory among other components of language. The nonword method specifically has been suggested for emergent bilinguals because it does not necessitate advanced linguistic knowledge prior to testing (Archibald, 2008). In turn, dynamic assessment focuses on learning processes and its utility has been explored for language screenings among young EB children (Patterson et al., 2020).

Even in the midst of advancements among the measures described thus far with young emergent bilingual children, standardized norm-referenced measures comprise the most frequently utilized approach. In discussing standardized measures, it is important to examine frequently identified concerns that must be considered in selection, implementation, and interpretation of results. First, standardized measures historically were initially normed for English monolinguals and then used with young emergent bilinguals without even a basic evaluation of their linguistic, cultural and psychometric properties (see Barrueco et al., 2012). The practice of directly translating measures from one language to another while disregarding cultural and linguistic differences may even still be found among some measures today, yielding potentially invalid findings on children's functioning. As Mancilla-Martinez et al. (2018) describe, the extent to which cultural and social factors are incorporated in the development of measures must be considered when identifying whether specific assessment tools may be appropriate for application with specific children or communities. According to Bandel et al. (2012), specific facets to examine when reviewing and selecting measures for use with EB children include (a) the extent to which skills found across differing cultures are incorporated for the communities being assessed, (b) the extent to which the questions asked on the measures function similarly or dissimilarly across languages and cultures, and (c) the extent to which dialectical variations within one linguistic group, as the differences in pronunciation, vocabulary, or grammatical structures, are incorporated. Such characteristics should be carefully examined through a mix of advanced, rigorous quantitative and qualitative analyses and procedures. If empirical support is not evidenced, bias may be present in the measures (Bandel et al., 2012).

Second, there has traditionally been a scarcity of appropriately-developed standardized measures available in languages other than English for use with young EB children. Importantly, tremendous shifts have been made in the field, and the use of this argument to avoid crucial assessments with young EB children is no longer the case. There are indeed well-developed, normed measures, particularly in Spanish (e.g., Barrueco et al., 2012). What the field faces (and what young EB children face) is continued use of invalid measures, limitations in the creation and validation of measures beyond English and Spanish, and a scarcity of trained personnel to appropriately select, administer, and interpret assessments with EB children (Barrueco et al., 2012; Mancilla-Martinez et al., 2018). Critically, conclusions reached from poorly executed assessment approaches with young EB children may be inaccurate and harmful.

Third, the language proficiencies of the standardization groups can differ tremendously among culturally and linguistically validated measures. For example, some measures in Spanish are developed abroad in Spanish-speaking countries and/or with recent Spanish-speaking immigrants in the United States. On the one hand, such measures can facilitate an understanding of how strongly developed language skills in languages other than English are. On the other hand, it can reflect an assumption that bilingual children are fundamentally two monolinguals in one person (Grosjean, 1989). Further, the application of findings from populations who have markedly distinct monolingual language environments and experiences (including potentially different dialects) to the community of emergent bilingual children in the United States may be questionable. There is also acknowledgment that the languages within bilinguals are not inherently separated, as evidenced by emerging bilinguals code-mixing by using both languages in the same stretch of conversation (Genesee, 1989). Positively, assessment instruments are being created and norm-referenced specifically with young emergent bilingual children across the nation. These include, among others, the *Bilingual English-Spanish Assessment* (BESA; Peña et al., 2018) and the *Preschool Language Scales-fifth Edition, Spanish* (Zimmerman et al., 2012).

A related consideration in the standardization of measures for young EB children is whether they were developed to assess one language (such as Spanish or Japanese) or whether they are created to be administered bilingually (Spanish-English or Japanese-English). The latter, termed *conceptual scoring*, provides practitioners with information about children's skills across languages, which is critical when the questions at hand relate to the overall development of children's functioning in language, reading, or other domains. However, conceptual scoring does not yield specific information about the development of skills in each language, which is sometimes integral depending on the purpose of the assessment (Aikens et al., 2020). With conceptually scored measures, every item is designed and analyzed for their equivalence in serving as valid indicators of development at each stage of growth. A good example of the necessity of such analyses is evidenced in comparing monosyllabic words in English that are often present as early items on English measures (such as ball) relative to their more difficult linguistically counterparts in other languages (such as pelota in Spanish).

Literacy assessment

While language and literacy are intricately connected, young EB children can progress in separate rates across these two subdomains (see, [López and Páez, 2021](#) for discussion). That is, early literacy development may be more advanced in one or both languages compared to language development, or there may even be variation dependent on the specific language and subdomain (for example, low English and high Spanish vocabulary coupled with high English and low Spanish literacy). Such findings may be evidenced when comparing results to English dominant peers, and also reflect the composition of measures. Some early literacy measures focus on more discrete skills (such as alphabet knowledge and/or phonological awareness) relative to the linguistic processes involved in literacy assessments for older children (such as reading comprehension (e.g., [DeBruin et al., 2015](#)).

Given this state of the field and the rapid rate of linguistic and literacy development within and across languages early in life, it is advisable that the literacy assessment of young EB children be conducted in all their languages (rather than only one) in order to gather a complete perspective of their full literacy abilities and growth. Otherwise, assessing in one language will only provide information regarding literacy development in that singular language ([Aikens et al., 2020](#)). Conversely, using both languages provides fuller information about young EB children's literacy development, including specific words or phrases within and across languages ([Aikens et al., 2020](#)). Otherwise, potentially invalid inferences about young EB children's literacy knowledge and abilities may result.

Even so, it is important to acknowledge that, at times, assessments may only be able to be administered in one language due to limitations in resources, time, and other factors (when assessing literacy development and/or other domains; [Aikens et al., 2020](#)). In such cases, determinations are made to conduct assessments in the home language, English, or dominant language. Systematic decisions for identifying the appropriate language for evaluation should be made through parent or teacher reports or utilization of language screeners directly with young EB children ([Bandel et al., 2012](#)). Examples of measures utilized for language screeners include the *Woodcock-Muñoz Language Survey-III* (WMLS-III; [Woodcock et al., 2017](#)) and the *PreLAS English and Spanish assessment scales* ([Zimmerman et al., 2012](#)).

Cognitive development

As described at the outset of this chapter, a comprehensive approach across developmental domains is critical in the assessment of all children, including young EB children. If linguistic difficulties are occurring among young EB children, it could potentially reflect a broader cognitive or learning impairment. Further, if their languages are developing strongly, solely considering this aspect of development among young EB children might miss the demand for support in other domains, such as in the other aspects of cognition.

Within the Head Start Early Learning Outcomes Framework (ELOF), the cognitive domain reflects varying aspects dependent on age ([U.S. DHHS, 2015](#)). Among infants and toddlers, the cognitive subdomains focus on: (a) exploration and discovery, (b) memory, (c) reasoning and problem-solving, (d) emergent mathematical thinking, and (e) imitation and symbolic representation and play. In the cognitive assessment of infant and toddler EB children, indirect measurement through parent and teacher reports is often used. This includes measures such as the *Ages and Stages Questionnaire* (third ed., ([Squires and Bricker, 2009](#)), which is available in various languages (Arabic, Chinese, French, Spanish, and Vietnamese and others).

Among preschoolers, the ELOF cognitive subdomains reflect early academically-related processes, such as a) mathematics development (including counting, cardinality, operations, algebraic thinking, measurement, geometry, and spatial sense), and b) scientific reasoning (scientific inquiry, reasoning, and problem-solving). Assessing preschooler mathematical development (including that of young EB children) has often been comprised of indicator measures, including subtests from the *Batería IV Woodcock-Muñoz*; [Woodcock et al. \(2019\)](#). More detailed mathematical measures have also been created such as the *Research-based Early Mathematics Assessment- Short Form* (REMA-SF; [Dong et al., 2021](#)), and with attention to diverse communities (e.g., [Weiland et al., 2012](#); [Alkhadim et al., 2021](#)). Regardless of the specific measure, young EB children's mathematical skills may be underestimated if only assessed in English ([White et al., 2021](#)).

Strides have also been made in assessing scientific reasoning among young EB children, notably Spanish bilingual children. A federally-funded study is underway to create a Spanish adaptation of the *Lens on Science* measure, a computer-adaptive tablet-based science measure ([Greenfield, 2015](#)). *Enfoque en Ciencia* is being developed with bilingual samples across the United States including representation of varied communities and dialectal variations, as well as analysis of potential biases using item response theory analytic approaches (PI: Daryl Greenfield, Co-PIs Sandra Barrueco & Ximena Domínguez).

In recent years, early academic and developmental assessments at school entry through widespread benchmark testing has been gaining traction, including kindergarten entry assessments (KEA; [Ackerman, 2018](#)). If developed well and to its fullest potential and used appropriately with diverse communities, KEAs may be able to serve as state and national indicators of early childhood well-being and thus improve policy initiatives ([Kaminiski et al., in press](#)). Through this approach, school systems may also be able to create targeted practices that reduce educational gaps among young EB children by mitigating developmental and educational delays before they become prevalent. While promising, great strides must occur before widespread implementation of KEA assessments for young EB children, particularly in regards to cultural, linguistic, and psychometric properties ([Ackerman, 2018](#); [Kaminiski et al., in press](#)).

Social and emotional development

Relationships with adults, relationships with other children, emotional functioning, and sense of identity and belonging comprise the four subdomains of the ELOF across the infant, toddler, and preschool years (U.S. DHHS, 2015). Assessment of social and emotional development in early childhood often consists of parent and/or teacher reports (either through interviews or questionnaires). For young EB children, these can include measurement systems like the *Ages and Stages Questionnaires: Social-Emotional* (second ed., Squires et al., 2015), which are available in multiple languages and across ages. In addition, parent-report measures also include the *Brief Infant-Toddler Social & Emotional Assessment (BITSEA)* (Briggs-Gowan and Carter, 2006), which includes Competence and Problem scales and whose Spanish form has been utilized with young EB children in national studies (Caswell et al., 2019).

Among assessments that examine social and emotional development, there is a stark obligation to attend to linguistic and cultural considerations since there can be tremendous diversity in social norms and emotional expression (e.g., Leong et al., 2014). When considering peer interactions and interpretation of social behaviors, one aspect that may be uniquely relevant for young EBs is Theory of Mind (ToM). ToM refers to the ability to attribute beliefs, emotions, and intentions to another person or oneself, and it contributes to the understanding that others can have different mental states than oneself (Premack and Woodruff, 1978). From a socio-pragmatic approach, bilingualism may contribute to EB children's awareness about others' mental states due to their experience communicating with individuals who speak different languages (Goetz, 2003). Further, bilingual children may understand the possibility of different perspectives as they can express themselves using varying modalities. ToM and, relatedly, social reasoning, can influence children's estimation of others' behavior and can subsequently impact their reactions and behaviors. Measurement development assessing such processes among young EB children continues to make progress, such as with the Spanish version of the *ToMas-Child* (Rivas-Garcia et al., 2020). Among older EB children, a Spanish-language version of *Virtual Environment for Social Information Processing (VESIP-Sp)* is also currently undergoing validation across the United States (PI: Russo-Ponsaran, Co-Is: McKown, Wainer, Barrueco, Smith, & Huber).

Approaches to learning

The Approaches to Learning domain encompasses attention, exploration, and higher order cognitive processes within the Head Start Early Learning Outcomes Framework, with similar subdomains for infants, toddlers, and preschoolers: (a) emotional and behavioral self-regulation, (b) cognitive self-regulation (executive functioning), (c) initiative and curiosity, and (d) creativity (U.S. DHHS, 2015). This aspect of development has garnered extensive research in the bilingualism field over past decades, among the most of any domain outside of the language and literacy domain.

Indeed, the relationships between bilingual experiences, attention, and memory have been of keen interest (Bialystok and Werker, 2017). Bilingualism has been linked to executive function (EF) capabilities in children and conceptualized to influence the development of nuanced attentional strategies given the subtleties required to understand two languages (Bialystok, 2015). When selective attention is required, bilingual children have demonstrated elevated inhibitory capabilities (Bialystok, 2015). However, findings of enhanced executive functions, such as inhibition, appear more consistent when children have already developed high rates of bilingualism and less consistent among emergent bilinguals in the process of learning multiple languages with varied linguistic and life experiences (De Bruin, 2019; Paradis et al., 2021).

Among young emergent bilingual children, there is a relative dearth of validated measures of approaches to learning, especially for EB infants and toddlers. Furthermore, there are limited assessments available to examine the subdomains of initiative, curiosity, and creativity for children across all ages, much less young EB children. As this area of research continues and is promising for young EB children, measurement of the approaches to learning domain must advance.

Physical development

Overall health and physical development are important facets of early childhood, especially given the remarkable changes during this age period. Fine motor skills particularly play a role in expressive language development and language impairments and, as such, are critical to include in screenings and assessments of young EB children (Grissmer et al., 2010). Gross motor skills and physical health also provide a broad perspective on the general development of young EB children, including discerning whether any language difficulties may be reflective of systemic concerns that should be addressed and supported in a multisystemic fashion. Health (including dental health) and nutrition can be involved in language-specific and other difficulties, with disparities identified among EB children's families (e.g., Carter et al., 2022; Flores and Toman-Korman, 2008). Within the Head Start Early Learning Outcomes Framework, this domain is inclusive of perceptual, motor, and physical development (U.S. DHHS, 2015). Specific measures for screening and assessing physical development among young EB children can include parent and teacher report measures (such as the *Ages and Stages Questionnaire* in English and Spanish), as well as direct assessment of body-mass index through measuring height and weight.

Family and early care experiences

While the comprehensive assessment of young emergent bilingual children should incorporate the varied abilities within children (as described in the five domains above), it must not stop there. Rather, young EB children's development and experiences are best

understood with the contexts of their familial and early care settings. This includes examination of environmental contributors that influence developmental processes, such as exposure to and use of languages in the home and early care environments, formal educational levels, and socioeconomic resources (also termed external factors; see [Paradis et al., 2021](#)).

For example, the presence of a dominant language across two languages (where one language is more developed than another) is fairly typical in the context of the United States. Differing levels of proficiencies across two languages may not necessarily be indicative of children's inability to develop or sustain two languages. Rather, language dominance may be reflective of varying experiences across environments (e.g., school, home, community; [Paradis et al., 2021](#)). Importantly, imbalance does not necessarily imply language learning difficulties, impairment, delays, or even linguistic preference among young EB children. Even so, parents, teachers, and healthcare providers (including psychologists) are often concerned about young EB children's language development and assumptions must be avoided that any language delays or imbalance is solely due to external contributors. Instead, a thorough analysis of environmental factors and a full examination of children's development across various facets of functioning and over time is essential. Age of dual language learning and relative language proficiency should be assessed within the context of everyday language usage and switching patterns, though these have historically been difficult to accurately obtain ([De Bruin, 2019](#)).

Importantly, there have been strides in recent years in evaluating language exposure and use in both home and early care settings. Specifically, detailed parent and teacher reports have been developed, which are described in detail above within the language and literacy section. Further, wearable devices are available that capture language exposure and use (such as LENA), along with open-source automatic estimators of phoneme, syllable and word counts (e.g., [Räsänen et al., 2021](#)). However, continued developments in the accuracy of such devices may be warranted ([Marchman et al., 2021](#)). Finally, classroom observations of bilingual language use across individuals (child, teacher, assistant teacher, others), settings (large group, small group, individual), and purpose (instruction, non-instruction) have been developed through national studies, such as the *Cultural Items and Language Use (CILU) Checklist* ([Barrueco et al., 2017](#); [Bumgarner et al., 2019](#)).

Societal and policy contexts

The availability and use of effective assessments for young emergent bilingual children is further situated within societal and policy contexts. As indicated above, significant inroads have been made in the development of linguistically, culturally, and psychometrically validated measures, though continued advancements always remain. Even so, key issues persist in impacting effective emergent bilingual assessment at systemic levels. Among various factors, these include the availability of trained professionals, accessibility, policies, and funding. To understand the source of such barriers, it is informative to consider the historical and current policies pertaining to emergent bilingual children and their assessment. Young emergent bilingual children are members of the same societal and policy contexts as their school-aged EB peers and soon enter these educational settings for elementary school and beyond. Two notable policies are relevant: the Lau decision and Proposition 227.

First, the Lau decision of 1974 by the Supreme Court was a pivotal inflection point because of its directive in instructing language-minority students (another term for English Language Learners; [Ovando, 2003](#)). The decision stemmed from a case pertaining to the equality of education for nearly two thousand students who primarily spoke Chinese and experienced comprehension barriers to educational content. Citing the Civil Rights Act of 1964, this landmark decision provided greater educational access to bilingual children with the required provision of specialized support and services. Even with the subsequently passed Lau Remedies, some barriers for EB children have persisted due to inconsistencies in approaches to educating and assessing bilingual children, varied foci on distinct goals and outcomes, and isolation of some programs in schools and districts ([Ovando, 2003](#)).

The second notable policy influencing EB children and their assessment was Proposition 227, a Californian law passed in 1998 and later repealed in 2016. It dictated that language-minority students were primarily required to learn in English with minimal support of their home languages. It reflected an expanding societal focus on English language use to the exclusion of other languages in response to increased immigration ([Padilla et al., 1991](#); [Ovando, 2003](#)). States such as Arizona, Connecticut, Washington, and Colorado adopted similar laws. The implementation of English-only programs shifted the linguistic resources in classrooms for emergent bilingual students, as well as reduced the ability to collect accurate, holistic data about their academic progress across their languages ([Katznelson and Bernstein, 2017](#)).

What does this mean for the future of assessment for emerging bilingualism? Some scholars suggest that the assessment of emergent bilingual students is at the center of policy decision-making and is essential in determining the future of interventions and programs. A study conducted by [Poza and Viesca \(2020\)](#) examined the outcomes of a law requiring all students to reach English reading proficiency by third grade (specifically, The Colorado Reading to Ensure Academic Development Act). State-mandated English proficiency assessments (without similar implementation of assessments in other languages or across languages) were found to effectively serve as a means of limiting the incorporation of home language in instruction. These findings highlight the capability and power of assessments to determine programs and policies. It is important to note that double-testing in two languages can also create double-standards where bilingual students are held to a disproportionately higher standard than their monolingual peers in which they are expected to perform well academically in two languages instead of only one or across both ([Poza and Viesca, 2020](#)). This is an important consideration in schools where English proficiency is prioritized over general educational development and where bilingual development is not fully understood or supported. Nevertheless, implementation of assessments solely in English for emergent bilingual children can not only misestimate their development, they effectively change the practices and programs developed to facilitate their success.

Within the early childhood education system, policies have been enacted that support multilingual language use in familial and care settings. This is reflected in standards and recommendations comprised of educational, medical, and federal entities (e.g., McCabe et al., 2013; Head Start Early Childhood Learning and Knowledge Center, 2021; Oliva-Olson et al., 2019), and reflects the incorporation of research-based knowledge about bilingual developmental processes during the initial years of life. Relatedly, it has become mandated that all young EB children participating in Head Start programs across the United States must be assessed across their languages in order to comprehensively understand their development (U.S. Department of Health and Human Services, 2016). Influencing nearly a million young children a year and citing research on bilingual assessment (Barrueco et al., 2012), fiscal resources are provided yearly to programs to support the adequate implementation of full assessments for young EB children.

Conclusion

Impressive inroads have been made in the field of young emergent bilingual assessment, from the creation of measures across a range of domains and settings to the creation of national standards and policies. At this juncture, improvements in the availability and implementation of multilingual comprehensive assessments for all young emergent bilingual children remain, especially when it is well-established that assessments across domains, metrics, and languages is best practice (e.g., Lazewnik et al., 2019; DHHS, 2016; NAEYC, 2020). These consist of parent and/or teacher reports, language sampling and/or observations, along with linguistically- and culturally-valid and well-standardized measures. Comprehensive measurement approaches yield more accurate depictions of young EB children's development, especially if there is the possibility of impairment such as autism spectrum disorder or developmental language disorder. An essential component is the incorporation of multilingual measurement since assessment solely using one single language can lead to underestimation and misdiagnosis (Bandel et al., 2012).

Unfortunately, best practice has been hampered by limitations in training and availability of professionals with experience and expertise in this topical area. As such, the status quo is poised to continue where young EB children are assessed with inappropriate measures, inaccurate procedures, hastily translated measures by well-meaning professionals, or not assessed at all (Ackerman, 2018; Barrueco et al., 2012). However, this does not have to be the case nor should it be. Professionals themselves are calling for training advancements. For example, a nationally-representative study identified that over half of teachers and assistant teachers in Migrant and Seasonal Head Start indicated a need for additional training in bilingual development and about one-third of teaching staff across the United States called for additional training in assessment (Caswell et al., 2019).

Given that approximately 1 in 10 children in the United States are emergent bilinguals (and even higher proportions of 1 in 5 among some young EB populations), there is an obligation for all constituents to appropriately assess this substantial proportion of American children (National Center for Education Statistics, 2022; U.S. Department of Health and Human Services, 2006). Further, it is not the sole responsibility of any one system of care given the multiple and varied interactions young EB children have across settings. Health service professionals, early care and educational settings, and school districts are all primary systems that must be held accountable for accurately assessing young emergent bilingual children. Now is the time for universal training, standards, and policies for all professionals and across all settings to comprehensively and accurately meet the strengths and needs of young emergent bilingual children across the nation.

References

- Ackerman, D.J., 2018. Comparing the potential utility of kindergarten entry assessments to provide evidence of English learners' knowledge and skills. ETS Res. Rep. Ser. 2018 (1), 1–26. <https://doi.org/10.1002/ets2.12224>.
- Aikens, N., West, J., Mckee, K., Moiduddin, E., Atkins-Burnett, S., Xue, Y., 2020. Screening approaches for determining the language of assessment for dual language learners: evidence from Head Start and a universal preschool initiative. *Early Child. Res. Q.* 51, 39–54. <https://doi.org/10.1016/j.ecresq.2019.07.008>.
- Alkhadim, G.S., Cimet, A.D., Marx, R.W., Cutshaw, C.A., Yaden, D.B., 2021. Validating the research-based early math assessment (REMA) among rural children in Southwest United States. *Stud. Educ. Eval.* 68. <https://doi.org/10.1016/j.stueduc.2020.100944>.
- Archibald, L.M., 2008. The promise of nonword repetition as a clinical tool. *Can. J. Speech-Lang. Pathol. Audiol.* 32 (1), 21–28.
- Bandel, E., Atkins-Burnett, S., Castro, D.C., Wulsin, C.S., Putnam, M., 2012. Examining the Use of Language and Literacy Assessments with Young Dual Language Learners (Final Report). The University of North Carolina, FPG Child Development Institute, CECER-DLL.
- Barrueco, S., Layzer, C., Caswell, L., 2017. Migrant and Seasonal Head Start Cultural Items and Language Use Checklist (MSHS CILU). Developed for the MSHS Study. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- Barrueco, S., López, M., Ong, C., Lozano, P., 2012. Assessing Spanish-English Bilingual Preschoolers: A Guide to Best Approaches and Measures. Paul H Brookes Publishing.
- Bauman, K., 2017. School Enrollment of the Hispanic Population: Two Decades of Growth. United States Census Bureau. Available at: https://census.gov/newsroom/blogs/random-samplings/2017/08/school_enrollmentof.html.
- Bedore, L.M., Peña, E.D., 2008. Assessment of bilingual children for identification of language impairment: current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 1–29. <https://doi.org/10.2167/beb392.0>.
- Bialystok, E., 2015. Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9 (2), 117–121. <https://doi.org/10.1111/cdep.12116>.
- Bialystok, E., Werker, J.F., 2017. The systematic effects of bilingualism on children's development. *Dev. Sci.* 20 (1). <https://doi.org/10.1111/desc.12535>.
- Briggs-Gowan, M.J., Carter, A.S., 2006. The Brief Infant-Toddler Social Emotional Assessment: Technical Manual. Pearson Clinical Assessment, San Antonio, TX.
- Bumgarner, E., Caswell, L., Layzer, C., Barrueco, S., 2019. Select Findings from the Migrant and Seasonal Head Start Study 2017: Cultural Items and Language Use (CILU) Checklist. OPRE Report #2020-18. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- Carter, A., Clarke, L., Stevens, C., 2022. Dental health for children with special educational needs and disability. *Paediatr. Child Health* 32 (8), 290–296. <https://doi.org/10.1016/j.paed.2022.05.003>.

- Castro, D.C., Scheffner Hammer, C., Franco, X., Cycyk, L.M., Scarpino, S.E., Burchinal, M.R., 2020. Documenting bilingual experiences in the early years: using the CECER-DLL child and family and teacher questionnaires. *Biling. Lang. Cognit.* 23 (5), 958–963. <https://doi.org/10.1017/S1366728920000401>.
- Caswell, L., Bumgarner, E., Barrueco, S., López, M., Wolf, A., Layzer, C., 2019. Migrant and Seasonal Head Start Study 2017: Data Tables. OPRE Report #2019-66. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- De Bruin, A., 2019. Not all bilinguals are the same: a call for more detailed assessments and descriptions of bilingual experiences. *Behav. Sci.* 9 (3), 33. <https://doi.org/10.3390/bs9030033>.
- DeBruin-Parecki, A., Kleeck, V.A., Gear, S., 2015. *Developing Early Comprehension: Laying the Foundation for Reading Success*. Paul H. Brookes Publishing Co.
- Dong, Y., Clements, D.H., Day-Hess, C.A., Sarama, J., Dumas, D., 2021. Measuring early childhood mathematical cognition: validating and equating two forms of the Research-Based Early Mathematics Assessment. *J. Psychoeduc. Assess.* 39, 983–998. <https://doi.org/10.1177/07342829211037195>.
- Espinosa, L.M., López, M.L., 2007. Assessment Considerations for Young English Language Learners across Different Levels of Accountability, vol. 5. National Early Childhood Accountability Task Force and First, pp. 1082–1111.
- Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, S., Bates, E., 2007. *MacArthur Communicative Development Inventories: User's Guide and Technical Manual*, second ed. Brookes, Baltimore, MD.
- Fenson, L., Pethick, S., Renda, C., Cox, J., Dale, P., Reznick, S., 2000. Short form versions of the MacArthur communicative development inventories. *Appl. Psycholinguist.* 21, 95–116. <https://doi.org/10.1017/S0142716400001053>.
- Flores, G., Tomany-Korman, S.C., 2008. The language spoken at home and disparities in medical and dental health, access to care, and use of services in US children. *Pediatrics* 121 (6), e1703–e1714. <https://doi.org/10.1542/peds.2007-2906>.
- García, O., Kleifgen, J.A., Falchi, L., 2008. From English Language Learners to Emergent Bilinguals. *Equity Matters: Research Review*, No. 1.
- Genesee, F., 1989. Early bilingual development: one language or two? *Journal of child language* 16 (1), 161–179. <https://doi.org/10.1017/S0305000900013490>.
- Goetz, P., 2003. The effects of bilingualism on theory of mind development. *Biling. Lang. Cognit.* 6 (1), 1–15. <https://doi.org/10.1017/S1366728903001007>.
- Greenfield, D.B., 2015. Assessment in early childhood science education. In: *Research in Early Childhood Science Education*. Springer Netherlands, pp. 353–380. https://doi.org/10.1007/978-94-017-9505-0_16.
- Grismer, D., Grimm, K.J., Ayer, S.M., Murrah, W.M., Steele, J.S., 2010. Fine motor skills and early comprehension of the world: two new school readiness indicators. *Dev. Psychol.* 46 (5), 1008–1017. <https://doi.org/10.1037/a0020104>.
- Grosjean, F., 1989. Neurolinguists, beware! the bilingual is not two monolinguals in one person. *Brain Lang.* 36 (1), 3–15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5).
- Guzman-Orth, D., Lopez, A.A., Tolentino, F., 2017. A framework for the dual language assessment of young dual language learners in the United States. *ETS Res. Rep. Ser.* 1–19. <https://doi.org/10.1002/ets2.12165>.
- Head Start Early Childhood Learning and Knowledge Center, 2021. Dual Language Learners in Head Start. Available at: <https://eclkc.ohs.acf.hhs.gov/culture-language/article/using-data-support-full-effective-participation-children-who-are-dual-language-learners-dlls-their>.
- Jackson-Maldonado, D., Marchman, V., Fernald, L., 2013. Short-form versions of the Spanish MacArthur-Bates communicative development inventories. *Appl. Psycholinguist.* 34 (4), 837–868. <https://doi.org/10.1017/S0142716412000045>.
- Kaminiski, J., Barrueco, S., Kelleher, K., Edwards, E. & Hoagwood, K. (in press). School Readiness as a Core Metric for Early Childhood Health. *National Academies of Medicine Perspectives*.
- Katznelson, N., Bernstein, K.A., 2017. Rebranding bilingualism: the shifting discourses of language education policy in California's 2016 election. *Ling. Educ.* 40, 11–26. <https://doi.org/10.1016/j.LINGED.2017.05.002>.
- Lazewnik, R., Creaghead, N., Smith, A., Prendeville, J., Raiser-Becker, L., Silbert, N., 2019. Identifiers of language impairment for Spanish-English dual language learners. *Lang. Speech Hear. Serv. Sch.* 50 (1), 126–137. https://doi.org/10.1044/2018_LSHSS-17-0046.
- Leong, F.T.L., Nagayama Hall, G.C., McLoyd, V.C., Trimble, J.E. (Eds.), 2014. *APA Handbook of Multicultural Psychology*, vols. 1 & 2. American Psychological Association.
- López, L.M., Pérez Mariela, M., 2021. *Teaching Dual Language Learners what Early Childhood Educators Need to Know*. Paul H. Brookes Publishing.
- Mancilla-Martínez, J., Gámez, P., Vagh, S., Lesaux, N., 2016. Parent reports of young Spanish-English bilingual children's productive vocabulary: a development and validation study. *Lang. Speech Hear. Serv. Sch.* 47 (1), 1–15. https://doi.org/10.1044/2015_LSHSS-15-0013.
- Mancilla-Martínez, J., Greenfader, C.M., Ochoa, W., 2018. Spanish-speaking preschoolers' conceptual vocabulary knowledge: towards more comprehensive assessment. *NHSA Dialog* 21 (1).
- Marchman, V., Weisleder, A., Hurtado, N., Fernald, A., 2021. Accuracy of the Language Environment Analyses (LENA) system for estimating child and adult speech in laboratory settings. *J. Child Lang.* 48 (3), 605–620. <https://doi.org/10.1017/S0305000920000380>.
- McCabe, A., Bornstein, M.H., Guerra, A.W., Kuchirko, Y., Pérez, M., Tamis-LeMonda, C.S., Cates, C.B., Hirsh-Pasek, K., Melzi, G., Song, L., Golinkoff, R., Hoff, E., Mendelsohn, A., 2013. Multilingual children: beyond myths and toward best practices and commentaries. *Soc. Pol. Rep.* 27, 1–37. <https://doi.org/10.1002/j.2379-3988.2013.tb00077.x>.
- McLaughlin, B., 1984. *Second-language Acquisition in Childhood: Preschool Children*, second ed. Lawrence Erlbaum Associates, Inc.
- National Association for the Education of Young Children, 2020. *Developmentally Appropriate Practice*. Available at: <https://www.naeyc.org/resources/position-statements/dap/contents>.
- National Center for Education Statistics, 2022. *English Learners in Public Schools*. Available at:
- Ovando, C.J., 2003. Bilingual education in the United States: historical development and current issues. *Biling. Res. J.* 27 (1), 1–24. <https://doi.org/10.1080/15235882.2003.10162589>.
- Oliva-Olson, C., Espinosa, L.M., Hayslip, W., Magruder, E.S., 2019. Many languages, one classroom: supporting children in superdiverse settings. *Teaching Young Children* 12 (2), 4–7.
- Padilla, A.M., Lindholm, K.J., Chen, A., Durán, R., Hakuta, K., Lambert, W., Tucker, G.R., 1991. The English-only movement: myths, reality, and implications for psychology. *Am. Psychol.* 46 (2), 120. <https://doi.org/10.1037/0003-066X.46.2.120>.
- Paradis, J., Genesee, F., Crago, M.B., 2021. *Dual Language Development & Disorders: A Handbook on Bilingualism and Second Language Learning*, third ed. Paul H. Brookes Publishing Co.
- Patterson, J.L., Rodríguez, B.L., Dale, P.S., 2020. Dynamic assessment language tasks and the prediction of performance on year-end language skills in preschool dual language learners. *Am. J. Speech Lang. Pathol.* 29 (3), 1226–1240. https://doi.org/10.1044/2019_AJSLP-19-00120.
- Peña, E.D., Gutiérrez-Clellen, V.F., Iglesias, A., Goldstein, B.A., Bedore, L.M., 2018. Bilingual English Spanish Assessment (BESA). Brookes, Baltimore, MD.
- Poza, L.E., Viesca, K.M., 2020. Testing and ideology: policy debates about literacy assessments for Colorado's bilingual students. *J. Educ. Pol.* 35 (4), 556–581. <https://doi.org/10.1080/02680939.2018.1511831>.
- Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1 (4), 515–526. <https://doi.org/10.1017/S0140525X00076512>.
- Räsänen, O., Sheshadri, S., Lavechin, M., Cristia, A., Casillas, M., 2021. ALICE: an open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behav. Res. Methods* 53, 818–835. <https://doi.org/10.3758/s13428-020-01460-x>.
- Rivas-García, S., Bateman, A., Caracul, A., 2020. Validation of the ToMas-child Scale for the assessment of the theory of mind in a group of Spanish speaking children aged 3 to 7 years from Spain. *Dev. Neuropsychol.* 45, 232–245. <https://doi.org/10.1080/87565641.2020.1764567>.
- Squires, J., Bricker, D., 2009. *Ages and Stages Questionnaire (ASQ): A Parent Completed Child Monitoring System*, 3rd ed. Paul H. Brookes Publishing.
- Squires, J., Bricker, D., Twombly, E., 2015. *Ages & Stages Questionnaires®: Social-Emotional, Second Edition (ASQ®:SE-2)*. A Parent-Completed Child Monitoring System for Social-Emotional Behaviors. Paul H. Brookes Publishing Co., Inc, Baltimore.
- U.S Department of Health and Human Services, Administration for Children and Families, Office of Head Start, 2015. *Head Start Early Learning Outcomes Framework: Ages Birth to Five*. Washington, DC.

- U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start, 2006. Biennial Report to Congress: The Status of Children in Head Start Programs, 2005. Washington, DC.
- U.S. Department of Health and Human Services, 2016. Administration for children and families, Office of Head Start. Head Start Program Performance Standards and Other Regulations. Washington, D.C.
- Weiland, C., Wolfe, C.B., Hurwitz, M.D., Clements, D.H., Sarama, J.H., Yoshikawa, H., 2012. Early mathematics assessment: validation of the short form of a prekindergarten and kindergarten mathematics measure. *Educ. Psychol.* 32, 311–333. <https://doi.org/10.1080/01443410.2011.654190>.
- White, L., Keuter, S., Hsin, L., Li, M., Quick, H., Manship, K., Holtzman, D., 2021. Math Skills Among Spanish-speaking DLLs: Implications for Assessment. First 5 California.
- Woodcock, R.W., Alvarado, C.G., Ruef, M., 2017. Woodcock-Muñoz Language Survey, third ed. Riverside.
- Woodcock, R.W., Alvarado, C.G., Schrank, F.A., McGrew, K.S., Mather, N., Muñoz-Sandoval, A.F., 2019. Bateria IV Woodcock-Muñoz. Riverside Insights Publishing.
- Zimmerman, I.L., Steiner, V.G., Pond, R.E., 2012. Preschool Language Scales-5th Edition Spanish. Pearson Publishing.

Literacy/multiliteracies and foreign language education

Beatrice C. Dupuy and Chantelle N. Warner, The University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	105
Literacy in late 20th and early 21st centuries	105
The New London Group's manifesto and the literacy turn	105
FL multiliteracies post NLG's manifesto	106
New and emergent directions in multiliteracies language education	107
Translingual literacy practices and social justice	107
Transmodal language pedagogies	108
Aesthetics and affect	108
Conclusion	109
References	109

Introduction

In the last two decades, literacy has emerged as a central notion for guiding curricular and instructional decisions in foreign language (FL) programs in the US. This is especially at the collegiate level where calls for a paradigm shift have often been directly connected not only with critiques of the two-tiered curriculum that has long characterized them (e.g., Allen and Paesani, 2010; Kern, 2000, 2003; Maxim et al., 2013) but also broader conversations around supporting advanced FL abilities (e.g., Byrnes, 2005; Byrnes et al., 2010; Maxim, 2009). In this two-part essay, we offer a brief history of literacy focusing on the literacy turn of the mid-1990s and the move from literacy to multiliteracies in the first two decades of the 21st century and its contributions to the field of FL, and suggest new directions in FL multiliteracies that should shape research and practice going forward. To provide some context, we start with a snapshot of literacy in the early days of communicative language teaching.

From the mid-1970s to mid-1990s, several FL scholars (e.g., Phillips, 1978; McKee, 1981; Terry, 1989), in likely reaction to a persistent tendency to prioritize oral over written modes in communicative language teaching, called for a better integration of reading and writing into FL teaching which they argued should also be viewed as communication. At the time, literacy was primarily equated with reading and writing as separate cognitive processes and the ability to decode and encode texts in print, page-bound, official, standard forms of the language studied. Nevertheless, by supporting a view of FL reading and writing as communication, these early advocates set the stage for later, more multidimensional meaning-focused approaches to literacy (Paesani et al., 2016; Warner and Dupuy, 2018).

The second half of the 1990s witnessed a shift, with reading and writing being increasingly viewed as interconnected modes of communication. Whereas lexical and syntactic knowledge were often highlighted in early discussions on FL literacy, now greater consideration was given to the role cultural and textual schema play in both reading and writing and greater stress was put on the need to go beyond mere comprehension and functional use and foster analytic and interpretive abilities. The more integrative approaches of the 1990s would pave the way to later understandings of literacy not only as plural (as in multiliteracies) but also as social practice wherein identities, relationships, and world views are built (Gee, 1990).

Literacy in late 20th and early 21st centuries

The New London Group's manifesto and the literacy turn

In the years leading to the meeting of the New London Group (NLG), an international group of education and literacy scholars, and the publication of their manifesto *A Pedagogy of Multiliteracies: Designing Social Futures* in 1996, the understanding of literacy as reading and writing and the dominance of print modes over other modes of communication had started to be challenged in linguistics and educational circles. In the last decade of the 20th century, (a) the world was changing fast and literacy pedagogy in education needed to reflect the increasing cultural and linguistic diversity of the contemporary globalized world, (b) approach text genres as sites of unequal power relations, and (c) account for the new kinds of texts and textual engagement that had emerged in the wake of new information and multimedia technologies. With this shift in thought, greater attention would be paid to everyday literacy practices putting them on par or higher in value to academic literacy practices.

From the outset, the NLG's overall goal was twofold: (a) change what counts as literacy and literacy teaching, and (b) acknowledge the multimodal nature of literacy practices. In the jointly-authored framework they proposed at the end of their week-long meeting the NLG called for a multiliteracies pedagogy that recognizes the range of communicative modes used to make meaning in texts, encourages critical and social justice framing, and promotes engagement with design, technology, and multimodality.

The NLG's pedagogy of multiliteracies views language and other modes of communication as dynamic resources ("available designs") for meaning making which get transformed in dynamic acts ("designing") of re-presentation and re-contextualization as learners attempt to achieve their own purposes, thereby contributing new meaning resources ("the redesigned"). Three key aspects make this broader view of literacy and literacy teaching stand apart from earlier views: variability, agency, and subjectivity. Within this framework, meaning does not reside in texts, but is rather actively and dynamically made when learners creatively apply linguistic and other semiotic resources in interrelated ways with an awareness of "the sets of conventions connected with semiotic activity [...] in a given social space" (NLG, 1996, p. 74).

To translate the multiliteracies view of dynamic communication into the context of teaching, the NLG's framework proposes four pedagogical components (Situated Practice, Overt Instruction, Critical Framing, and Transformed Practice) which need to be part of the learning process, though not in any particular fixed order or as neatly separate foci. In Situated Practice activities, learners are guided to make use of their designed and designing experiences as they engage in textual meaning-making. In Overt Instruction activities, learners, with active instructor support, are introduced to explicit metalanguages so that they are better able to recognize and meaningfully describe the processes and patterns of Design. Critical Framing activities help learners connect particular Designs of meanings to their social and cultural contexts, and view and evaluate them critically in relation to their own contexts. Transformed Practice activities provide opportunities for learners to demonstrate their ability to apply reflectively the knowledge that they developed through Overt Instruction and Critical Framing activities, in other contexts or cultural sites and in new and creative ways "embedded in their own goals and values" (NLG, 1996, p. 87). Kalantzis and Cope (2000) underscored that Situated Practice is the point from which the other three pedagogical components depart but never leave behind and each of them "in its own way, expands the horizons of the lifeworld. Overt Instruction makes implicit patterns of meaning explicit; Critical Framing interrogates contexts and purposes; Transformed Practice takes meanings and subjectivity into new and less familiar domains" (p. 207). Overt Instruction, Critical Framing and Transformed Practice are "processes for 'denaturalising' the lifeworld, of making the everyday strange in order to cast new light on it and so as to have a more informed basis upon which to design both imminent meanings and our larger social futures" (p. 208).

Although the NLG's Multiliteracies framework was never more than a "statement of general principle" (1996, p. 89) for schools, their call for educators to acknowledge the diversity and, social and cultural situatedness of literacy has had a deep and lasting impact on FL teaching and learning. The work of this group's scholars along with the work of other scholars who also took a critical approach to literacy appeared at a time when the field of FL education was becoming less firmly anchored in theories of language acquisition and more interested in theories of language socialization (Steffensen and Kramsch, 2017). In the next section, we turn to the current state of FL literacy studies as it has developed in the early decades of the 21st century before finally focusing on primarily two directions that both research and practice should consider going forward.

FL multiliteracies post NLG's manifesto

Building on the NLG's manifesto as well as critical literacy scholarship, Kern's (2000) seminal book *Literacy and Language Teaching* offered both a rationale and a model for a more integrative approach to teaching FL written communication. For Kern, much like the NLG, literacy is not about "what texts mean in an absolute sense, [but] what people mean *by* texts, and what texts mean *to* people who belong to different discourse communities" (p. 2; emphasis in original). This kind of literacy, one that "[considers] reading and writing in their contexts of use, [frames] reading and writing as complementary dimensions of written communication, rather than utterly distinct linguistic and cognitive processes" (p. 2) can help FL learners develop deep communicative competence. For Kern this involves their ability to use linguistic and other semiotic resources in both critical and conscious ways. To facilitate this, FL teaching must be more deliberate in integrating critical framing and transformed practice as conceptualized by NLG across all curricular levels of FL programs, something that is still a work in progress as instructor-led, text-centered talk (usually around literary texts) that seldom allows for complex thinking in complex language largely continues to prevail (Mantero, 2002, 2006; Donato and Brooks, 2004) in this instructional context. The long-standing focus on skill-based literacy instruction and the entrenched belief that learners with lower communicative ability cannot engage in cognitively challenging tasks, explain why multiliteracies pedagogy has been slow to enter the FL classroom and why the existing gap between lower and upper division FL courses often continues to persist.

Expanding on his argument that FL Learners must develop a "critical and cultural understanding of language, literacy, and communication" (Kern, 2000, p. 134), Kern argued that a literacy-based approach could help bridge the perennial divide between lower-level and upper-level courses in FL programs (see also Parsons, 1985; Kramsch and Nolden, 1994), a point which was taken up again in the Modern Language Association (2007) report, which called for "a broader and more coherent curriculum in which language, culture, and literature are taught as a continuous whole" (p. 3). Following this report, numerous publications, primarily focused on curricular and instructional frameworks and the types of textual work that needed to be more actively promoted in the FL classroom, appeared. Scholars like Arens and Swaffar (2000) argued that the ACTFL Standards (National Standards in Foreign Language Education Project, 1996) could provide a framework to promote FL learners' literacy as a set of linguistic, cognitive, and sociocultural skills, while others argued that proficiency-oriented frameworks have limited potential for fostering an understanding of "what it takes to communicate capably across, or for that matter, even within cultures" (Warford and White, 2012, p. 401). It is for this reason that scholars like Crane (2006), Paesani (2006), Maxim (2009), and Troyan (2016) suggested a genre- and discourse-based orientation for re-organizing the FL curricular sequence so that it would reflect a social understanding of language use and add breadth and depth to it in the process.

A growing number of scholars have expanded the range of literacy practices in FL teaching far beyond the verbal to include visual texts such as graphic novels (e.g., [Elsner et al., 2013](#)), films (e.g., [Brown et al., 2015](#); [Dubreil, 2011](#); [Goulah, 2007](#); [Kaiser and Shibahara, 2014](#); [Kaiser, 2018](#); [Ryshina-Pankova, 2013](#)), images (e.g., [Sagnier, 2015](#)), posters (e.g., [Péron, 2010](#); [Ryshina-Pankova, 2013](#); [Wu, 2016](#)), photos (e.g., [Schmerbeck and Lucht, 2017](#)), and artwork (e.g., [Palpacuer-Lee, 2018](#)) in the classroom, arguing that the critical interpretation and production of different visual genres “enable learners to uncover prevalent representational motives, metaphors, and symbols in texts” ([Ryshina-Pankova, 2013](#), p. 164) and empower them “in their use of the target language and development of second language identities” ([Brown et al., 2015](#), p. 162).

As FL education is gradually moving beyond a proficiency-oriented view of communication, multiliteracies-oriented pedagogies offer a path forward to expand the goals of FL learning so that it takes in the “[r]eflection that accompanies interpretation and explores the diversity of interpretations, assumptions, perspectives, positions, expectations, and judgments, and includes reflexivity that turns one’s own experience into an object of critical examination” ([Leung and Scarino, 2016](#), p. 90). However, even as a large number of FL educators have embraced a design-based perspective on meaning-making, other aspects of the multiliteracies framework have often remained underemphasized, including an attention to translanguaging, social justice, and learner agency/subjectivity (see [Collier and Rowsell, 2014](#); [Warner and Michelson, 2018](#)). Writing from the context of US second language teaching and learning, Warner and Michelson argue that “institutional and disciplinary priorities of FL education, which are often framed tightly by and indeed identified through nationalist paradigms and dominant discourses, create constraints” (p. 12) which might explain why the dimensions of multiliteracies that are not readily tethered to proficiency and competence-oriented goals are often overlooked. Adopting broader and combined multiliteracies strands would demonstrate a commitment to the original ideas of NLG who understood multiliteracies more holistically.

In this next section, we examine several emergent and in many ways interconnected new research directions, which are beginning to expand the scope of multiliteracies in FL education and highlight important notions from the framework that have been overlooked by the field for a long time but are starting to emerge.

New and emergent directions in multiliteracies language education

Translingual literacy practices and social justice

Offered as a way to pedagogically respond to changing literacy practices and increasing linguistic and cultural diversity in an interconnected world, the NLG’s multiliteracies framework aimed to make literacy instruction more socially just by recognizing the validity of multiple languages, cultures, and means of expression and the many fluid and unpredictable ways in which individuals dynamically mix the languages and other semiotic resources in their repertoire when in contact situations of linguistic and cultural diversity particularly in digital spaces ([Cope and Kalantzis, 2013](#)) and form in the process new practices, meanings, and grammars which [Canagarajah \(2013\)](#) describes as translingual practices. Beyond this broadening of focus, the multiliteracies framework also promoted pedagogical shifts which called on bringing students’ out-of-school literacy practices into the classroom and, in doing so, expanding what “counts” as meaning-making in an academic context.

Translingual pedagogy is still very much in its infancy in foreign language education as it challenges commonly adopted departmental structures which tend to define programs and courses by the names of languages and sequenced proficiency levels, and the exclusive use of the target language policy. In programs whose *raison d’être* is defined by the names of languages (see [Kramsch, 2012, 2014](#); [Kramsch and Huffmaster, 2015](#)) instructors grapple with the need to reconcile the changing contexts of language use in a globalized world with the fostering of historical and cultural awareness through the study of one named language and culture. As a result, translingual practices in both in-person or digital environments remain under-applied and under-researched in this field and work is needed to better understand the role that translingual practices may have in FL pedagogy as “these modes of hybridity can be socially and rhetorically significant” ([Canagarajah, 2013](#), p. 10). By opening spaces wherein learners can use their entire semiotic repertoire and act as their translingual literate selves, FL teachers would engage in acts of social justice by making the learning environment transformative.

Social justice and critical literacy have always been central to multiliteracies pedagogies in the awareness they bring to the fact that texts are never neutral and that, as any forms of language use, whichever they are, they reflect the interests, purposes, and ideologies of their producers. More importantly, although less reflected in FL pedagogy, critical literacy pushes educators and students to “see themselves as active participants of social change, as learners and students who can be active designers—makers—of social futures” ([NLG, 1996](#), p. 64). Some FL scholars (e.g., [Glynn et al., 2014](#); [Phipps and Gonzales, 2004](#)) have argued that the FL classroom can be a locus for social transformation. The goal of modern language education, as argued by Phipps and Gonzales, is to foster the development of “linguagers ... ‘agents’ or ‘language activists,’ who engage actively with the world, and for whom language learning is a way of embarking on the risky business of stepping out of one’s habitual ways of speaking and acting in order to engage with others whose modes of speech and action are other” (cited in [Crosbie, 2009](#), p. 295). As the futures of individuals around the world hinge on decisions about social justice, multiliteracies pedagogies within FL instructional contexts have the potential to help learners bridge between analysis and action as they develop an awareness of how language shapes the very “misconceptions, untruths, and stereotypes that lead to structural inequality and discrimination” ([Nieto, 2010](#), p. 46).

Transmodal language pedagogies

In recent years, major organizations in the field of FL education (e.g. ACTFL/Partnership for 21st Century Skills; MLA) have started to promote an expanded view of meaning-making modes beyond the more traditional one which generally refers to the modalities of listening, reading, speaking, and writing and reflects a linguistic mode bias. The focus on the linguistic mode in this traditional view might be considered intuitive given the field's commitment to helping students develop the target language, but it has also served to perpetuate a longstanding tradition of logocentrism which is not reflective of an individual's repertoire of semiotic resources and the dynamic and complex ways in which they deploy them in their literacy practices. As put by Kress et al. (2014), the underlying assumption of such a view is that non-linguistic modes are "illustrative supports to the 'real thing'" rather than legitimate meaning-resources in their own right. Kress (2000) called on language educators to no longer "act as though language fully [represents] the meanings they wish to encode and communicate" while viewing other modes as "someone else's to look after" (p. 337).

With meaning-making resources other than linguistic ones featuring more prominently in recent conversations, and transmodal/translingual competence being a desired goal of foreign language education, FL programs are starting to embrace, albeit slowly, multiple modes and their use "in the design of a semiotic product or event" (Kress and van Leeuwen, 2001, p. 20), found in the types of visual media (e.g., graphic novels, films, videos, images, posters, and artwork) described in the previous section. More recently, many new literacy practices enabled by digital communication media have also begun to make their way in the FL classroom. In 2008, Thorne and Reinhardt suggested an approach that bridges not only the lower and upper levels of a language curriculum but also more academic genres with vernacular language, which is increasingly dominated by technology-mediated literacy practices. Similarly, Haffner, 2014 advocated for digital literacies to be embedded in existing language curricula and argued for extending the traditional range of genres and practices including informal, everyday practices that draw on the affordances of digital communication tools and demonstrate participants' use of complex semiotic systems. Reinhardt, Warner, and Lange (2014) and Warner and Richardson (2017) have demonstrated that integrating digital gaming and digital social media into a German genre-based multiliteracies course can help instructors and students view their engagement with texts within a larger nexus of social practices and discourses. Leier (2018), also in German, designed an Instagram-based, multiliteracies-oriented task which led learners to see themselves not only as consumers and borrowers of the social internet but also as contributors to, and thus part of, a virtual shared pool of knowledge.

In the current era of globalization of linguistic and cultural flows, digital communication media bring together ever-expanding audiences which engage in new configurations of communicative engagements and thus new literacy practices which are pushing FL educators to confront the transmodal/translingual realities of the contemporary communities in which learners are likely to participate such as Facebook (e.g., Androutsopoulos, 2014, 2015; Blyth and Dalola, 2016; de Nooy, 2017; Kulavuz-Onal and Vásquez, 2018; Oliver and Nguyen, 2017; Schreiber, 2015; Warner and Chen, 2017), fandoms (e.g. Shafirova et al., 2020; Zang and Cassany, 2016), or video games and their devoted online platforms (e.g., Hernández and Castaño, 2015; Lee, 2018; Reinhardt and Sykes, 2013). Findings from these studies demonstrate that these affinity spaces with their multimodal affordances and one-to-many communication offer an environment supportive of transmodal/translingual practices and "can raise learners' awareness of the dynamic and emergent nature of translingualism" (Blyth and Dalola, 2016).

Aesthetics and affect

Multiliteracies frameworks have expanded the scope of literacy in second language education beyond the kinds of transactional functionalism often found in proficiency-oriented treatments of texts; however, as education scholars such as Leander and Boldt (2012) and Pahl and Rowsell (2020) have also argued, there is a tendency toward hyperrationalism and artifactualization prompted by the design metaphor, which also assumes that people's purposes in engaging with texts are largely determined and goal-oriented. In their critical re-reading of the New London Group's manifesto, Leander and Boldt (2012) argue that, on the contrary, literacy-related activity is often not projected toward some textual end point but "as living its life in the ongoing present, forming relations and connections across signs, objects, and bodies in often unexpected ways" (p. 26). In his more recent book *Language, Literacy Technology* (2015), Kern has similarly argued that literacy pedagogy must become more relational and less object-oriented, by recognizing that much of meaning making is not in the texts themselves but in the invisible relations between things (see Kern, 2015, p. 222). This involves giving learners opportunities to not only analyze and apply the meanings designed into textual forms, but to engage with the ways texts move us and are moved through the world (e.g. Gramling and Warner, 2016; Warner and Gramling, 2013). This has also brought renewed attention to the role of aesthetics and affect in second language literacy development.

Scholars of language education (e.g., Risager, 2005; Kramsch, 2011; Kramsch and Nolden, 1994) have long emphasized the poetic and aesthetic dimensions of language and culture learning. While acknowledging that literacy practices are inextricable from the sociocultural contexts of their production and reception, these scholars have also pointed attention to the sensations, feelings, and emotions texts conjure in those who design meaning through and from them that are not necessarily directly connected to social interaction. Although affective response is understood by these scholars as attributable to the "richness and variety" (McRae, 1996, p. 19) of language and literacy more generally, there is also a case to be made through these more aesthetically attuned models for literariness in language teaching and learning. In this vein, some have argued that the "destabilizing potential" (Schultz, 2001) of literary literacy practices can draw learners' attention to the range of effects that particular design choices can have and enable them to "construct new/unknown meanings" (Tin, 2010, p. 216). Kramsch has argued for a central place for the "literary imagination" as a means of fostering symbolic competence and translingual/transcultural thinking (Kramsch, 2006, p. 251). Recent scholarship on

digital social annotated reading and viewing in foreign language contexts (e.g. Blyth, 2014; Thoms and Poole, 2017) has provided some valuable insights as to how learners engage with the aesthetic elements of texts, but further research is needed to understand what role literariness might have in the development of FL literacy. An initial study by Blyth et al. (2021) also suggests that meta-language that emphasizes literary and creative ways in which users play with language can help instructors in connecting concepts from meaning-oriented paradigms in language education to the design of pedagogical materials and learning experiences.

Conclusion

Just as other fields of education have been changed through shifting paradigms for understanding literacy and its development at the end of the 20th century, foreign language teaching and learning have been significantly impacted by the societal, technological, and pedagogical rethinking *multiliteracies* frameworks have attempted to capture. The pluralizing potential of the concept of *multiliteracies* has enabled scholars and practitioners to expand definitions of reading and writing, but beyond that it has created a term around which a range of different approaches that share in common a commitment to teaching languages as complex, dynamic systems for meaning making that intersect in practice with each other (translingually) and with other semiotic systems (multimodally). In this sense, multiliteracies pedagogies are helping language educators to further realize the meaning-oriented approaches to language teaching and learning that began to be conceived during the proficiency movements in the 1980s, by expanding the scope of what kinds of communicative resources receive space and attention in the language classroom.

At the same time, multiliteracies frameworks complicate familiar models of communication within FL education through the metaphor of design. Understanding language use and learning as the design of meaning highlights the reality that languages do not simply use new linguistic systems to manifest already formed thoughts, ideas, and feelings, but rather these ways of being in the world are shaped through the designs available, the agentful process of designing, and the ways in which we redesign ourselves and the world when we engage with them. As discussions of multiliteracies continue to develop and expand to better integrate movement between languages and modes and to recognize not only critical thinking but feeling, the pedagogies shaped through this work will also need to develop and adapt to social futures that are quite different from what the New London Group was responding to in the mid 1990s. Although it may be impossible to predict what kinds of social and technological realities language learners will encounter in their private and professional futures, we can safely assume that they will be called upon to engage with linguistic, semiotic, and cultural diversity that is multifarious and complex in all of the ways multiliteracies pedagogies have emphasized and conceptualized.

References

- Allen, H.W., Paesani, K., 2010. Exploring the feasibility of a pedagogy of multiliteracies in introductory foreign language courses. *L2 Journal* 2, 119–142.
- Androutsopoulos, J., 2014. Language when contexts collapse: audience design in social networking. *Discourse Context Media* 4–5, 62–73.
- Androutsopoulos, J., 2015. Networked multilingualism: some language practices on Facebook and their implications. *Int. J. Biling.* 19, 185–205.
- Arens, K., Swaffar, J., 2000. Reading goals and the standards for foreign language learning. *Foreign Lang. Ann.* 33, 104–122.
- Blyth, C., 2014. Exploring the affordances of digital social reading for L2 literacy: the case of eComma. In: Guikema, J., Williams, L. (Eds.), *Digital literacies in foreign and second language education*. CALICO, San Marcos, Texas, USA, pp. 201–226.
- Blyth, C., Dalola, A., 2016. Translingualism as an open educational language practice: raising critical language awareness on Facebook. *ALSIC* 19. <https://doi.org/10.4000/alsic.2962>.
- Blyth, C., Warner, C., Luks, J., 2021. The role of OER in promoting critical reflection and professional development: the Foreign Languages and the Literary in the Everyday Project. In: Blyth, C., Thoms, J. (Eds.), *Open Education and Foreign Language Learning and Teaching: The Rise of a New Knowledge Ecology*. Multilingual Matters, Bristol, UK, pp. 158–182.
- Brown, L., Iwasaki, N., Lee, K., 2015. Implementing multiliteracies in the Korean classroom through visual media. In: Kumagai, Y., López-Sánchez, A., Wu, S. (Eds.), *Multiliteracies in World Language Education*. Routledge, Oxon, UK, pp. 158–181.
- Byrnes, H., 2005. Literacy as a framework for advanced language acquisition. *ADFL Bull.* 37, 11–15.
- Byrnes, H., Maxim, H., Norris, J., 2010. Realizing advanced foreign language writing development in collegiate education: curricular design, pedagogy, assessment. *Mod. Lang. J.* 94, 1–235.
- Canagarajah, S., 2013. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge, London, UK.
- Collier, D., Rowsell, J., 2014. A room with a view: revisiting the multiliteracies manifesto, twenty years on. *FLuL* 43, 12–28.
- Cope, B., Kalantzis, M., 2013. "Multiliteracies": New Literacies, New Learning, Framing Languages and Literacies: Socially Situated Views and Perspectives. <https://doi.org/10.4324/9780203070895>.
- Crane, C., 2006. Modelling a genre-based foreign language curriculum: staging advanced L2 learning. In: Byrnes, H. (Ed.), *Advanced language learning: the contributions of Halliday and Vygotsky*. Continuum, London, UK, pp. 227–245.
- Crosbie, V., 2009. Future directions for modern languages in the higher education landscape: an interview with Alison Phipps and Mike Gonzalez. *Lang. Intercult. Commun.* 5, 294–303.
- de Nooy, J., 2017. How to be virtually French: Australian Facebook users' co-construction of a Francophile identity. *ALSIC* 20. Available from: <https://alsic.revues.org/3002>. (Accessed 15 September 2021).
- Donato, R., Brooks, F., 2004. Literary discussions and advanced speaking functions: researching the (dis)connection. *Foreign Lang. Ann.* 37, 183–199.
- Dubreil, S., 2011. Rebels with a cause: (Re)defining identities and cultures in contemporary cinema. *L2 Journal* 3, 176–200.
- Elsner, D., Helff, S., Viebrock, B., 2013. *Films, Graphic Novels & Visuals: Developing Multiliteracies in Foreign Language Education - an Interdisciplinary Approach*. Lit Verlag, Münster, DE.
- Gee, J., 1990. *Social Linguistics and Literacies: Ideology in Discourses*, third ed. Routledge, Oxon, UK.
- Glynn, C., Wesely, P., Wasell, B., 2014. *Words and Actions: Teaching Language through the Lens of Social Justice*. ACTFL, Alexandria, VA, USA.

- Goulah, J., 2007. Village voices, global visions: digital video as a transformative foreign language learning tool. *Foreign Lang. Ann.* 40, 62–78.
- Gramling, D., Warner, C., 2016. Whose 'crisis in language'? Translating and the futurity of foreign language learning. *L2 Journal* 8, 76–99.
- Haffner, C., 2014. Embedding digital literacies in English language teaching: students' digital video projects as multimodal ensembles. *Tesol Q.* 48. <https://doi.org/10.1002/tesq.138>.
- Hernández, M., Castaño, S., 2015. Language-as-Victory. *Micro-Papers* 32. Available from: <https://www.literaciesinl2project.org/uploads/3/8/9/7/38976989/islip-micro-paper-32-language-as-victory.pdf>. (Accessed 19 September 2021).
- Kaiser, M., 2018. Teaching a film clip in a multiliteracies framework. *L2 Journal* 10, 47–57.
- Kaiser, M., Shibahara, C., 2014. Film as source material in advanced foreign language. *L2 Journal* 6, 1–13.
- Kalantzis, M., Cope, B., 2000. Language education and multiliteracies. In: May, S., Hornberger, N.H. (Eds.), *Encyclopedia of Language and Education, Language Policy and Political Issues in Education*, second ed., vol. 1. Springer, New York, NY, USA, pp. 195–211.
- Kern, R., 2000. *Literacy and Language Teaching*. Oxford University Press, Oxford, UK.
- Kern, R., 2003. Literacy as a new organizing principle for foreign language education. In: Patrikis, P.C. (Ed.), *Reading between the Lines: Perspectives on Foreign Language Literacy*. Yale University Press, New Haven, CT, USA, pp. 40–59.
- Kern, R., 2015. *Language, Literacy, and Technology*. Cambridge University Press, Cambridge, UK.
- Kramsch, C., 2006. From communicative competence to symbolic competence. *Mod. Lang. J.* 90, 249–252.
- Kramsch, C., 2011. The symbolic dimensions of the intercultural. *Lang. Teach.* 44, 354–367.
- Kramsch, C., 2012. Authenticity and legitimacy in multilingual SLA. *Crit. Multiling. Stud.* 1, 107–128.
- Kramsch, C., 2014. Teaching foreign languages in an era of globalization: Introduction. *Mod. Lang. J.* 98, 296–311.
- Kramsch, C., Huffmaster, M., 2015. Multilingual practices in foreign language study. In: Cenoz, J., Gorter, D. (Eds.), *Multilingual education: Between language Learning and Translanguaging*. Cambridge University Press, Cambridge, UK, pp. 114–136.
- Kramsch, C., Nolden, T., 1994. Redefining literacy in a foreign language. *Unterrichtspraxis* 27 (1), 28–35.
- Kress, G., 2000. Multimodality: challenges to thinking about language. *Tesol Q.* 34, 337–340.
- Kress, G., Van Leeuwen, T., 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Arnold Publishers, London, UK.
- Kress, G., Jewitt, C., Ogborn, J., Tsatsarelis, C., 2014. *Multimodal Teaching and Learning: The Rhetorics of the Science Classroom*. Bloomsbury Academic, Bloomsbury, UK.
- Kulavuz-Onal, D., Vásquez, C., 2018. "Thanks, shokran, gracias": translanguaging practices in a Face- book group. *Lang. Learn. Technol.* 22, 240–256.
- Leander, K., Boldt, G., 2012. Rereading "A pedagogy of multiliteracies": bodies, texts, and emergence. *J. Lit. Res.* 45, 22–46.
- Lee, K.M., 2018. *New Approaches to Literacies Studies in the Digital and Globalizing World: Border-Crossing Discourses in the Global Online Affinity Spaces*. Doctoral dissertation, Arizona State University. ASU Digital Repository. Available from: <https://repository.asu.edu/items/50457>.
- Leier, V., 2018. Using Instagram in a tertiary German language course: students' perception and approaches. *N. Z. Lang. Teach.* 44, 77–90.
- Leung, C., Scarino, A., 2016. Reconceptualizing the nature of goals and outcomes in language/s education. *Mod. Lang. J.* 100 (S1), 81–95.
- Mantero, M., 2002. Bridging the gap: discourse in text-based foreign language classrooms. *Foreign Lang. Ann.* 35, 437–456.
- Mantero, M., 2006. Applied literacy in second language education: (Re)framing discourse in literature-based classrooms. *Foreign Lang. Ann.* 39, 99–114.
- Maxim, H.H., 2009. Developing advanced formal language abilities along a genre-based continuum. In: Katz, S.L., Watzinger-Tharp, J. (Eds.), *Conceptions of L2 Grammar: Theoretical Approaches and Their Applications in the L2 Classroom*. Cengage, Boston, MA, USA, pp. 172–188.
- Maxim, H.H., Hoeyng, P., Lancaster, M., Schaumann, C., Aue, M., 2013. Overcoming curricular bifurcation: a departmental approach to curriculum reform. *Unterrichtspraxis* 46, 1–26.
- McKee, E., 1981. Teaching writing in the second language composition/conversation class at the college level. *Foreign Lang. Ann.* 14, 273–278.
- McRae, J., 1996. Representational language learning: from language awareness to text awareness. In: Carter, R., McRae, J. (Eds.), *Language, Literature and the Learner*. Addison Wesley Longman Ltd, Harlow, UK, pp. 16–40.
- Modern Language Association, 2007. Foreign languages and higher education: new structures for a changed world. *Ad Hoc Committee on Foreign Languages. Profession* 234–245.
- National Standards in Foreign Language Education Project, 1996. *Standards for foreign language learning : preparing for the 21st century*. National Standards in Foreign Language Education Project, Yonkers, New York, USA.
- New London Group (NLG), 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66, 60–92.
- Nieto, S., 2010. *Language, Culture, and Teaching: Critical Perspectives*. Routledge, New York, NY, USA.
- Oliver, R., Nguyen, B., 2017. Translanguaging on Facebook: exploring Australian aboriginal multilingual competence in technology-enhanced environments and its pedagogical implications. *Can. Mod. Lang. Rev.* 73, 1–25.
- Paesani, K., 2006. *Exercices de style: developing multiple competences through a writing portfolio*. *Foreign Lang. Ann.* 38, 618–639.
- Paesani, K., Allen, H.W., Dupuy, B., 2016. *A Multiliteracies Framework for Collegiate Foreign Language Teaching*. Pearson, Upper Saddle River, NJ.
- Pahl, K., Rowsell, J., 2020. *Living Literacies*. MIT Press, Cambridge, MA, USA.
- Palpacuer-Lee, C., 2018. Multiliteracies in action at the art museum. *L2 Journal* 10, 134–157.
- Parsons, R., 1985. Language, literature, and curriculum revision in the 1980's. *Foreign Lang. Ann.* 18, 213–218.
- Péron, M., 2010. Writing history in the voice of an other: debyser's immeuble at the advanced level. *Foreign Lang. Ann.* 43, 190–215.
- Phillips, J., 1978. Reading is communication, too! *Foreign Lang. Ann.* 11, 281–287.
- Phipps, A., Gonzales, M., 2004. *Modern Languages: Learning and Teaching in an Intercultural Field*. Sage, London, UK.
- Reinhardt, J., Sykes, J., 2013. *Language at Play: Digital Games in Second and Foreign Language Teaching and Learning*. Pearson, Upper Saddle River, NJ, USA.
- Reinhardt, J., Warner, C., Lange, K., 2014. Digital games as practices and texts: new literacies and genres in an L2 German classroom. In: Pettes-Guikema, J., Williams, L. (Eds.), *Digital literacies in foreign and second language education*. CALICO, San Marcos, TX, USA, pp. 159–177.
- Risager, K., 2005. Languaculture as a key concept in language and culture teaching. In: Preisler, B., Fabricius, A., Haberland, H., Kjærbeck, S., Risager, K. (Eds.), *The consequences of mobility: linguistic and sociocultural contact zones*. Roskilde Universitet, Roskilde, DK, pp. 185–196.
- Ryshina-Pankova, M., 2013. Understanding "Green Germany" through images and film: a critical literacy approach. *Unterrichtspraxis* 46, 163–184.
- Sagnier, C., 2015. Multiliteracies and multimodal discourses in the foreign language classroom. In: Kumagai, Y., López-Sánchez, A., Wu, S. (Eds.), *Multiliteracies in World Language Education*. Routledge, Oxon, UK, pp. 81–116.
- Schmerbeck, N., Lucht, F., 2017. Creating meaning through multimodality: multiliteracies assessment and photo projects for online Portfolios. *Die Unterrichtspraxis* 50, 32–44.
- Schreiber, R.B., 2015. "I am what I am": multilingual identity and digital translanguaging. *Lang. Learn. Technol.* 19, 69–87.
- Schultz, J., 2001. The Gordian knot: language, literature, and critical thinking. In: Scott, V., Tucker, H. (Eds.), *SLA and the Literature Classroom: Fostering Dialogues*. Heinle, Boston, MA, USA, pp. 3–31.
- Shafirova, L., Cassany, L., Bach, C., 2020. Transcultural literacies in online collaboration: a case study of fanfiction translation from Russian into English. *Lang. Intercult. Commun.* 20, 531–545.
- Steffensen, S., Kramsch, C., 2017. The ecology of second language acquisition and socialization. In: Duff, P.A., May, S. (Eds.), *Encyclopedia of Language and Education, Language Socialization*, third ed., vol. 8. Springer, New York, NY, USA, pp. 17–37.
- Terry, R., 1989. Teaching and evaluating writing as a communicative skill. *Foreign Lang. Ann.* 22, 43–52.
- Thoms, J., Poole, F., 2017. Investigating linguistic, literary, and social affordances of L2 collaborative reading. *Lang. Learn. Technol.* 21 (2), 139–156.
- Thorne, S., Reinhardt, J., 2008. "Bridging activities," new media literacies, and advanced foreign language proficiency. *CALICO J.* 23, 558–572.

- Tin, T., 2010. Language creativity and co-emergence of form and meaning in creative writing tasks. *Appl. Linguist.* 32 (2), 215–235.
- Troyan, F., 2016. Learning to mean in Spanish writing: a case study of a genre-based pedagogy for standards-based writing instruction. *Foreign Lang. Ann.* 49, 317–335.
- Warford, M., White, W., 2012. Reconnecting proficiency, literacy, and culture: from theory to practice. *Foreign Lang. Ann.* 45, 400–414.
- Warner, C., Chen, H., 2017. Designing talk in social networks: what Facebook teaches about conversation. *Lang. Learn. Technol.* 21, 121–138.
- Warner, C., Dupuy, B., 2018. Moving towards multiliteracies in foreign language teaching: past and present perspectives...and beyond. *Foreign Lang. Ann.* 51, 116–128.
- Warner, C., Gramling, D., 2013. Gerade Dir hat er eine Botschaft gesendet: contact Pragmatics and the Teaching of Foreign Language Texts. In: Plews, J., Schmenk, B. (Eds.), *Traditions and Transitions: Curricula for German Studies*. Wilfrid Laurier University Press, Waterloo, CA, pp. 209–226.
- Warner, C., Michelson, K., 2018. Introduction to the special issue. Living literacies: L2 learning, textuality, and social life. *L2 Journal* 10, 3–15.
- Warner, C., Richardson, D., 2017. Beyond participation: symbolic struggles with(in) digital social media in the L2 classroom. In: Dubreil, S., Thorne, S. (Eds.), *Engaging the World: Social Pedagogies and Language Learning*. Cengage, Boston, MA, USA, pp. 199–226.
- Wu, S., 2016. Creating an effective learning environment in an advanced Chinese language course through film, poster presentations and multiliteracies. In: Kumagai, Y., López-Sánchez, A., Wu, S. (Eds.), *Multiliteracies in World Language Education*. Routledge, Oxon, UK, pp. 209–232.
- Zhang, L.T., Cassany, D., 2016. Fansubbing from Spanish to Chinese: organization, roles and norms in collaborative writing. *BiD Textos Univ. Bibl. Doc.* 37. Available from: <http://bid.ub.edu/en/37/tian.htm>.

Critical literacies in East Asian languages: examples from Japanese in foreign language education

Yuri Kumagai, Smith College, Department of East Asian Languages and Cultures, Northampton, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	112
Shifting focus from “mere literacy” to critical literacies in foreign language education	113
Critical literacies in foreign language education	113
Critical literacies praxes in Japanese language classrooms	114
Katakana project: questioning authoritative knowledge presented in the textbook	115
Linguistic landscape project: engaging in dialogical negotiations to uncover taken-for-granted assumptions	116
Comparative reading of newspaper reports: shuttling between languages to become linguistic mediators	118
Summary of praxes	119
Conclusion	119
References	120

Introduction

What does it entail for foreign language learners to participate in their new language communities? What does it mean to be “literate” in their new social and cultural worlds? Literacy—or “reading and writing,” as it is more commonly referred to—in foreign language education is often conceived as “language skills” (i.e., decoding and encoding) that involve complex cognitive processes. Particularly in the field of teaching Chinese and Japanese as foreign languages that use non-alphabetic writing systems (i.e., Chinese characters), significant amounts of instructional time and effort are dedicated toward helping learners recognize, sound out, and memorize each character so that they can “decode” and eventually make sense of texts. The importance of scribing characters neatly and correctly is also emphasized, and learners are encouraged to place painstaking attention to the details of how to write each character. Participating in one’s foreign/second language communities, however, requires much more than decoding and encoding. It requires critical literacies—in its plural form, emphasizing that various types and forms of literacy exist—that allow one to recognize the sociocultural and political context of language use, as well as building one’s repertoire (skills and strategies) in negotiating one’s identity through shifting power relations. To be literate, learners must consider themselves capable of interpreting language in texts as a means of creating social relationships, as well as being able to produce knowledge in their communities.

Critical literacy is a pedagogical approach that focuses on the relationships between language, literacy, and power (Luke, 2012). Paulo Freire’s work on critical consciousness and critical pedagogy has had a profound impact on the advancement of critical literacy. In his book *Literacy* (1987), Freire (together with Macedo) wrote, “Reading the world always precedes reading the word, and reading the word implies continually reading the world Reading always involves critical perception, interpretation, and ‘rewriting’ of what is read” (pp. 35–36, emphasis original). In other words, we create meaning through dialogic negotiations between language (the word) and lived experience (the world), and the act of reading and writing is a creative act based on “a critical comprehension of reality” (1987, p. 157). For Freire, literacy serves to critically reflect on how language shapes our representations of our experiences and existing social orders. Thus, it allows us to imagine and transform the world (Giroux, 1987; Vasquez et al., 2019). Freire, strongly rejecting the “banking-model of education,” calls for a “dialogical approach” in which knowledge is recreated through self-reflection and dialog (Shor and Freire, 1987).

The idea of critical literacy that was originally developed in first language education has taken on importance in the field of teaching English as a second language (e.g., Alford, 2001; Case et al., 2005; Chun, 2009; Kubota, 2010; Lau, 2012; Luk, 2015; Luke and Dooley, 2011; Wallace, 2001, 2003; Weninger, 2018, 2020) and, more recently, English as a foreign language (e.g., Cho and Johnson, 2020; Huang, 2011a,b, 2012, 2013; Huh, 2016; Ko and Wang, 2013; Kuo, 2009, 2015; Lau et al., 2021; Lee, 2017; Lin, 2001; Mitsikopoulou, 2020; Motschenbacher, 2021; Romeo and Bobkina, 2021; Sibanda, 2021; Weng, 2021). However, its dissemination is still limited in the field of foreign language teaching, with some exceptions: Ludewig (2017) and Ryshina-Pankova (2013) in German; Fabbian and Carney (2018a,b) in Italian; Férez Mora et al. (2022) and Verma (2008) in Spanish; Kene-man (2015) in French; and Chikamatsu (2021), Doerr and Kumagai (2014), Gyogi (2019), Iwasaki and Kumagai (2008, 2020), Koyama (2016), Kumagai (2007, 2020), Kumagai and Fukai (2009), and Kumagai and Iwasaki (2011) in Japanese.

In this chapter, I take on the question of what critical literacies mean for foreign language learners who are already “reading” the world in their first language. Why do we need critical literacies in one’s additional language? What can we do to nurture learners’ critical literacies in their new language? Sharing examples of classroom projects that were designed according to critical literacy principles and carried out with US college students, I attempt to shed light on what critical literacies look like in learning Japanese in a foreign language classroom.

Shifting focus from “mere literacy” to critical literacies in foreign language education

Traditionally, literacy is understood to constitute cognitive abilities and individual skills for reading and writing printed materials (Reagan and Osborn, 2021). In most foreign language classrooms, written texts are used in assisting students’ language learning. Written texts are viewed as sources of “language data” (Alderson, 1984; Devitt, 1997; Elley, 1984), are used for “language exercises” (Alderson and Urquhart, 1984; Bernhardt, 1991; Elley, 1984) such as examining grammar and vocabulary (and Chinese characters or *kanji*, in the case of Chinese and Japanese), and are used to practice oral decoding (Grabe, 1991).

Oftentimes, the goal of reading instruction, regardless of the linguistic level, is to develop students’ abilities to accurately retrieve information from texts (Wallace, 2003). It is assumed that there is one correct interpretation of a text that is agreed upon by native speakers of the language. This emphasis on literal comprehension, rather than on interpretive or evaluative reading practices, is particularly marked in Japanese (and Chinese) instruction due to the perceived challenges that their writing systems present to those whose primary language is alphabet-based (e.g., English) (Iventosch, 2012; Lin, 2012). Writing, particularly at the beginning level, is often reduced to exercises to demonstrate a student’s mastery of writing systems, as well as their use of vocabulary, grammar, and mechanics. The accuracy of the form is strongly emphasized, much more than the meanings that students try to convey with their new language.

Cope and Kalantzis (2000) call such literacy “mere literacy” (p. 5), which focuses solely on “language”—usually a singular national language—that is thought of as a static and stable system divorced from its context. In contrast, sociocultural and critical perspectives consider the situatedness of literacy to be of the utmost importance. Literacy—reading, writing, and talking around texts—only makes sense if and when it is understood within historical, sociocultural and political contexts (e.g., Barton, 2007; Barton et al., 2000; Gee, 1987/1991; Street, 2014). Furthermore, from the sociocultural and critical views of literacy, the notion a “text” does not limit itself to just “language.” It includes multimodality—such as colors, images, sounds, use of space, and so on—which is an integral part of meaning making (Kress, 2000). Particularly with the advent of internet-based communication, it is immensely important to consider other semiotic systems to be part and parcel of what constitutes a text.

Critical literacies focus on the power relations that writers create through the use of language in texts. They use linguistic analyses—both the micro features of texts and the macro conditions of societies—to help readers develop “critical language awareness” (Fairclough, 1992). Critical language awareness constitutes an awareness of how writers choose particular linguistic/semiotic devices to shape messages for particular readers in order to accomplish certain goals. No language choices are accidental; writers consciously (or sometimes subconsciously) make selections of words, grammars, and discourses from available resources to produce, reinforce, and challenge certain ideologies and struggles for power. Critical literacies “empower learners by providing them with a critical analytical framework to help them reflect on their own language experiences and practices and on the language practices of others” (Clark and Ivanic, 1997, p. 12). Furthermore, they can be used to analyze how the classroom, text, or conversation is related to broader social, cultural and political relations (Pennycook, 2001), as well as how texts and discourse practices are constituted in order to maintain the status quo (Anderson and Irvine, 1993). To critically engage with texts requires self-reflection, which leads to interrogating taken-for-granted concepts (e.g., Kubota, 1996; Pennycook, 2001; Wallace, 2003). Critical engagement also encourages learners to consider ways to innovate and appropriate language uses for their own purposes. Finally, it aims to make learners conscious that all texts—textbooks included—implicate their writers’ values and worldviews, and that texts’ presentations of reality, including linguistic rules, are always partial and interested, and thus always open to contestation (Brodkey, 1996).

Teaching linguistic norms and conventions is important because they are used as gatekeeping devices. Without knowledge of these norms and conventions, students could be marginalized and denied access to the mainstream target society (e.g., Delpit, 1995; Freire and Macedo, 1987). However, as Baynham (1995) suggests, rather than teaching and learning “the discourse conventions of the dominant culture because they are there and ‘given,’ uncritically,” it is important to teach and “learn these conventions critically and strategically as a means of gaining access to powerful discourses and their benefits” (p. 242). Explicit discussion of the social origins of conventions such as “correctness” or “appropriacy” allows students (if they choose, and when the risk is not too high) to contribute to the ideological work of challenging and changing the conventions that work against values, beliefs, social groups, and the ideas with which the students identify.

Shifting the focus of foreign language teaching from dispensing knowledge of linguistic rules and training learners’ skills development—mere literacy—to critical literacies that foster learners’ agency and their ability to critically interpret wider texts (i.e., the world), foreign language education can play a critical role in contributing to the advancement of a more equitable social order (Janks and Vasquez, 2011; Reagan and Osborn, 2021).

Critical literacies in foreign language education

Oftentimes, developing critical literacies has been considered to be outside of the purview of foreign language education (for a review of this discussion, see Comber and Simpson, 2001; Wallace, 2003). Some argue that any critical analysis of texts needs to wait until the learners have well-developed linguistic knowledge (e.g., Bernhardt, 1991), while others argue that the process of acquiring linguistic knowledge and developing critical language awareness is an interdependent process (Hammond and Macken-Horarik, 1999; Luke, 1995; Wooldridge, 2001). Kubota et al. (2003) argue that the educational goal in any field, including foreign language instruction, is to encourage students to think critically in order to transform society by making it more equitable. In

a foreign language classroom, students' proficiency, the curriculum goal, and each instructor's beliefs may influence how critical literacy practices can be brought into the curriculum. Nevertheless, the critical outlook should be incorporated at all levels of instruction.

Undoubtedly, critical engagement with texts is a challenging task for foreign language readers, since they read texts written for an audience and community to which they most likely do not belong (Kramsch, 1993; Wallace, 2003). They need to interpret the text from both the writer's and the target reader's viewpoints, as well as from their own and their familiar communities' perspectives. While engaging in critical literacies in foreign language may be a challenge, foreign language learners have an advantage over those who are monolingual speakers/readers. Equipped with multiple language and cultural repertoires, such learners can bring to their literacy practices their rich knowledge and perspectives from different societies, cultures, and languages. Such resources are invaluable assets, allowing learners to compare multiple perspectives, to notice discrepancies or blind spots, and to recognize that what is considered "normal" in one linguistic/cultural practice is not necessarily universally so.

Through critical literacy practices in foreign language, learners begin to recognize that what is written in a text is not indisputable fact, but rather the author's views or interpretations of an event as shaped by a particular ideology. Being mindful and extremely cautious toward such textual politics is especially crucial for foreign language students because they tend to have "an overly deferential stance" (Wallace, 2003) toward authentic foreign language texts, due to their subjectivity as "learners" of that language. Recognizing the choices that textual authors make so as to accomplish their purposes via particular linguistic/semiotic devices is an important step for learners; such students can then become active agents in their own worlds, choosing specific words for their desired purposes. By drawing upon and negotiating diverse linguistic resources, learners can also play a mediating role between languages and cultures—a role that is not possible for monolingual individuals (Iwasaki and Kumagai, 2020; Kumagai and Iwasaki, 2015).

Critical literacies praxes in Japanese language classrooms

In this section, I introduce summaries of three classroom praxes that I have designed and implemented into Japanese language curricula. "Praxis," according to Freire, refers to theory-based action/action-based theorizing (Shor, 1999). Each project, based on the learners' linguistic levels, has different foci, yet all are guided by critical literacy principles and are aimed at nurturing learners' critical language awareness and agencies as multilingual speakers who are willing to take part in making societies more equitable. The critical language awareness that I emphasize takes into account the following aspects: (a) diversity of languages, (b) diversity of interpretations, (c) reexamination of taken-for-granted ideas, (d) fluidity and hybridity of languages, and (e) co-constructed relations of language and ideology. The first two examples that I will provide are projects that were incorporated into existing textbook-based curricula: a *Katakana* Project for a beginning-level course and a Linguistic Landscape Project for an intermediate-low level course. The third example is designed for intermediate-high to advanced level students who have begun reading various types of authentic texts.

Before I introduce the projects, I will describe some features of literacy in Japanese that are of particular importance to students who are engaging in critical reading of texts. Japanese language has three distinctive orthographies: *hiragana*, *katakana*, and *kanji*. These systems (along with the Roman alphabet, called "*romaji*") are all used in typical Japanese texts. *Hiragana* and *katakana* are sound-based, phonetic scripts called syllabaries, and both have 46 symbols (See Fig. 1). Some of the symbols are manipulated to represent other sounds by adding one of two small diacritics at the top right corner—for example, は (ha) becomes ば (ba) and ぱ (pa).

Hiragana						Katakana					
	a	i	u	e	o	a	i	u	e	o	
-	あ	い	う	え	お	ア	イ	ウ	エ	オ	
k	か	き	く	け	こ	カ	キ	ク	ケ	コ	
s	さ	し	す	せ	そ	サ	シ	ス	セ	ソ	
t	た	ち	つ	て	と	タ	チ	ツ	テ	ト	
n	な	に	ぬ	ね	の	ナ	ニ	ヌ	ネ	ノ	
h	は	ひ	ふ	へ	ほ	ハ	ヒ	フ	ヘ	ホ	
m	ま	み	む	め	も	マ	ミ	ム	メ	モ	
y	や		ゆ		よ	ヤ		ユ		ヨ	
r	ら	り	る	れ	ろ	ラ	リ	ル	レ	ロ	
w	わ				を	ワ				ヲ	
-n	ん					ン					

Fig. 1 Hiragana and Katakana chart.

The other system, *kanji* (Chinese characters), is a meaning-based script where each character represents a concept. It is said that there are approximately 10,000 *kanji* in use (Taylor and Taylor, 2014); however, the government's current script policy (updated in 2010) lists 2136 *kanji* as “*joyo kanji*”, or commonly used *kanji*. In practice, 3000–3500 *kanji* are needed to read texts such as newspapers, advertisements, etc., and to function as a literate member of Japanese society (Gottlieb, 2005; Seeley, 2000).

Each writing system has some normative rules: *kanji* is used for content words, *hiragana* primarily for function words (case markers and verbal/adjectival suffixes), and *katakana* mainly for foreign-origin words. Nevertheless, a writer has some freedom in their choice of orthography. For example, *manga* (Japanese comic books) can be written in *hiragana* (まんが), *katakana* (マンガ), *kanji* (漫画), or even *romaji* (Manga), and each choice evokes different images and effects (Dahlberg-Dodd, 2022; Kunert, 2020; Robertson, 2017, 2021). For example, reading experiments conducted by Ukita et al. (1996) revealed that when Japanese readers see the *kanji* word for *isu*, chair (椅子), an image of an old-fashioned, strongly build chair is usually generated, while the *katakana* (イス) generates an image of a modern, elegant chair, and the *hiragana* (いす) is associated with a simple wooden chair (in Iwahara et al., 2003, p. 393). Fig. 2 shows adjectives that were provided most frequently by native Japanese university students who were asked to write down any attributes that they associated with each script (Iwahara et al., 2003).

Hiragana	soft, round, tender, simple, childish, feminine, lovely
Katakana	angular, hard, cold, simple, new, sharp, inorganic
Kanji	hard, difficult, intellectual, vigorous, old, masculine, formal

Fig. 2 Adjectives association with each script type. Adopted from Iwasaki and Kumagai (2015, p. 212).

Similarly, in Fig. 3, Smith and Schmidt (1996) give the following stereotypical assumptions that native Japanese readers associate with each script choice.

	Writer/Reader features	Stylistic features
Hiragana	female, young	softness, femininity
Katakana	young, especially male	modernity, pop culture
Kanji	male, middle-aged and older	erudition
Romaji	young, especially female	commerciality

Fig. 3 Stereotypical assumptions for each script type. Adopted from Iwasaki and Kumagai (2015, p. 212).

In addition, there exist words from three types of vocabularies for a (nearly) identical referent: native Japanese words (*wago*), Sino-Japanese words (Chinese origin) (*kango*), and loanwords from Western languages (predominantly English) (*gairaigo*). For instance, a “hotel” can be referred to as *yadoya* (宿屋), *ryokan* (旅館), and *hoteru* (ホテル), respectively. Similar to the choice of orthography, choice of vocabulary evokes different images of both the reference and the writer (e.g., social status, sophistication). According to Kikuchi (1994), the images of hotels generated by each vocabulary type are “nostalgic, outmoded, simple, plain” for *wago*, “simple, plain, familiar” for *kango*, and “modern, luxurious” for *gairaigo*.

The writer’s choices for these features (among other things) communicate not only the ways in which they present themselves (Robertson, 2017, 2021) and perceive the audience (Dahlberg-Dodd, 2022), but also—and perhaps more importantly—illuminate their unspoken ideologies and/or worldviews that are implicated with such choices.

Katakana project: questioning authoritative knowledge presented in the textbook

Focusing on the three writing systems that exist in Japanese, having students investigate the authentic use of scripts—particularly, that of *katakana*—is a meaningful way to introduce students to critical literacy practices that can be accomplished at the beginning level. The *Katakana Project* was carried out as a short-term project (about 2 weeks) in the second semester of a first-year Japanese language course (for details, see Kumagai, 2011).

Overwhelmingly, Japanese language textbooks explain that *katakana* is used for foreign loanwords (e.g., Banno et al., 2021; Nagara, 1990). For example, the textbook that I use in my beginning-level Japanese course describes the use of *katakana* as “normally used for writing loanwords and foreign names” (Banno et al., 2021, p. 20). Less frequently, some textbooks include other usages of *katakana*, such as for onomatopoeic words (Makino et al., 1998; Tohsaku, 2006) and for words the author wishes to emphasize (Tohsaku, 2006). However, as mentioned above, the choice of writing system is context-dependent and involves the writer’s meaning-creating processes. Accordingly, the *Katakana Project* was designed to question textbook explanations of the rules, and to promote students’ awareness about contextual use of writing systems by having them collect samples of *katakana* words from

real-life texts. By extension, the project aimed to nurture students' critical attitudes toward any authoritative writings, including those found in textbooks (see for details, Kumagai, 2007; Doerr and Kumagai, 2014).

Students were instructed to find various uses of *katakana* in their personal lives (e.g., food packages, *manga*, magazines, movie/CD covers, internet sites, etc.) and think about the possible reasons for and effects of its use. They found an abundance of *katakana* words that deviated from the "normative" rules described in the textbooks, and noticed that the use of *katakana* does not automatically mean that words written in it are foreign loanwords. They found *katakana* used for writing proper nouns (name of places, persons, and companies), animals and plants, onomatopoeia, slang and invented words, and for emphasizing different accents/enunciations of Japanese words uttered by foreigners (and even by space aliens!).

To illustrate the unconventional use of *katakana*, two examples that students brought to class are shared here. The first example is a Japanese poster of a Hollywood film, "Resident Evil: Extinction" (released in Japan as "Biohazard III"). The poster has a picture of the heroine holding guns in both hands, and behind her is an apocalyptic scene with a burning sky. On it, the name of the movie (バイオハザード, Biohazard) and the name of the actress who plays the heroine (ミラ・ジョヴォヴィッチ, Milla Jovovich) are written in *katakana*—normative uses of the script. What is not normative is the word "サヨナラ" (*sayonara*, or goodbye) written in *katakana* above the heroine's head. Since it is a Japanese phrase, it would normally be written in *hiragana* (さよなら). The use of *katakana* instead of *hiragana* evokes a sense that the word was spoken by the heroine (i.e., a white woman) with some foreign accent.

Another example is from a package of strawberry candies. On the package, four words are in *katakana*: "サクマ" (*sakuma*), which is the name of the company; "ギュ" (*gyu*) and "キュン" (*kyun*), both of which are onomatopoeia; and "スゴい" (*sugoi*), an adjective meaning "great", which is written in a combination of *katakana* (for the stem) and *hiragana* (for the suffix)—a highly unconventional way of scribing the word. Normally, the rules dictate that the word "sugoi" is written in *kanji* followed by *hiragana* (凄い), or, if catering for children who do not know that *kanji* yet, it could be written exclusively in *hiragana* (すごい). The use of *katakana* in place of *kanji* captured the student's attention and curiosity. This type of "script play", only possible for languages with multiple writing systems, allow writers to be creative in communicating various feelings and nuances.

Students interpreted the effects of non-conventional *katakana* usage as making it "stand out," evoking a sense of "modern", "cool", "light-hearted", etc. They noticed that factors such as genres (e.g., *manga*, fashion magazines, advertisements, newspapers, academic books, snack packages, etc.), the writer's background and identity representations (e.g., authoritative, erudition, young, hip, etc.), and the target audience (e.g., children, adults, general public, scientists, etc.) shape the choice of scripts. Through in-class discussions, the students also discovered that while these choices are guided by the writers' various intentions, interpretations of these works differ depending on their readers' backgrounds and worldviews. For example, scribing a foreigner's Japanese utterance in *katakana* was interpreted by some students as "discriminatory," as it highlighted "wrong pronunciation," whereas those who think that speaking English symbolizes "modern" and "intelligence" thought that such transliteration was "cool." By engaging in the project, the students realized the fluidity of meaning conveyed by script choice and also began to recognize the arbitrariness of rules and norms (Kumagai, 2007). As one student wrote in her reflection paper, "I noticed that there are many cases where *katakana* was used besides for *gairaigo* [foreign loanwords]. But before the project, I did not think about the reasons for such cases. Now, when I see *katakana* words, I want to know: Why is this word written in *katakana*?" These types of questions and ensuing discussions do not impede students' mastery of basic rules and norms; rather, they spark intellectual curiosity.

This project, questioning authoritative presentations of knowledge, does not suggest that we do away with the standard textbook explanations of writing systems. However, it is important to keep in mind that simplified rules in textbooks—aimed at avoiding confusion—can sometimes constrict a learner's potential and creativity. Oftentimes, judgment as to whether some use of language is correct or not is based not only on its alignment with rules, but also on who the author is (Bourdieu, 1991). Some uses praised as language play or creativity if done by native speakers may be viewed as an error if used by learners.

Linguistic landscape project: engaging in dialogical negotiations to uncover taken-for-granted assumptions

The Linguistic Landscape Project, which took advantage of linguistic landscapes—"words and images displayed and exposed in public spaces" (Shohamy and Gorter, 2009a, p. 1)—as sociocultural texts, was implemented in a fourth semester (the second semester of a second year) Japanese language class as a semester-long project (for details, see Kumagai and Takahashi, *in print*). Previous studies that used linguistic landscapes as resources for additional language learning discuss various educational effects, ranging from language forms (i.e., writing systems, vocabulary, expression, grammar) (Leung and Wu, 2012; Rowland, 2013; Sayer, 2009) to pragmatic competence (Cenez and Gorter, 2008; Chern and Dooley, 2014; Rowland, 2013), symbolic competence (Shohamy and Waksman, 2009b), multicompetence (Cenez and Gorter, 2008; Rowland, 2013), critical language awareness (Dagenais et al., 2009), and multimodal literacy skills. Importantly, all of these studies highlight that the use of linguistic landscapes in language education offers a learning context for students to critically engage with texts.

The linguistic landscape project provided students with an opportunity to analyze issues such as what languages and other semiotics are used on signs, what function(s) each of them plays, what such signs suggest about the community, and what power relationships among the different languages can be extrapolated through a critical "reading" of signs in Japan. Ultimately, its aim was to help students understand how language is integrated within broader social structures; this was achieved by exploring actual language-in-use, and, subsequently, by imagining ways to transform society so that it becomes more inclusive and equitable for all.

When there are no nearby areas where students can experience Japanese linguistic landscapes (as was the case for this project), Google Street View can offer an alternative way for students to virtually explore streets in Japan. In groups of two or three, students

chose a city, virtually walked down the street, and selected (and took screenshots of) several signs that captured their interest, including several that they deemed to be incomprehensible. They then engaged in multiple levels of analysis of each sign using a modified version of the “Four Resources Model” of critical literacy (Freebody and Luke, 1990): (a) decoding and describing, (b) meaning making, (c) evaluating and critiquing, and (d) transforming (see Kumagai and Takahashi, *in print*). Students, in their original groups, completed several assignments: their findings were presented in class, an analytical essay was composed, and members created a multimodal page on WordPress. The various groups’ pages then became community resources for other Japanese language learners. In the process, students became “language detectives” (Sayer, 2009) and “urban geographers” (Malinowski, 2010) with agency as “investigators,” not merely “learners.”

The project adopted the dialogical approach (Shor and Freire, 1987), incorporating group discussions, multiple class presentations with feedback, and multiple check-in meetings with the teachers as significant parts of the learning process. It provided both students and teachers with an opportunity to engage with diverse interpretations, leading all participants to recognize their own blind spots concerning socioculturally sanctioned norms.

Fig. 4 shows a public sign posted by the Tokyo Metropolitan Police Department on a busy street in Tokyo. It alerts drivers that the road would be closed to the public in case of an earthquake. Significantly, this sign has been brought to class every time the project was conducted. What makes this sign so intriguing for students?

Students always ask why there is an illustration of a fish on the emergency sign. This simple question leads to a class discussion which illuminates that the connection between earthquakes and catfish (as told by a Japanese folktale) is a sociocultural one. This made “native” Japanese teachers recognize the problem of using such a culture-specific drawing on an official warning sign, as it excludes certain groups of people. Further discussion revealed that a more fundamental issue that prevented students from understanding the sign was the use of a cute illustration (be it catfish or anything else) on official signs, which is a normalized cultural practice prevalent in Japan. One student stated, “I thought it was funny because this is like ... a pretty serious sign, like you know, emergency sign that everyone needs to see. And it’s just like a fun little cartoon. [...] it wouldn’t make sense because like in America, all the emergency signs are just text and no pictures.” Her words underscored the difference in sociocultural expectations for what emergency signs should look like in Japan and the US.

The analysis of linguistic landscapes affords diverse interpretations, since the signs’ intentions may not be always explicit. For example, regarding the sign discussed above, it is not clear why the sign maker decided to put an illustration of a catfish, which occupies more than half of the space on the sign. What does the smiling catfish communicate with the viewers? Does it help to get a person’s attention? Isn’t it too distracting for people to actually read what is written? These discussions put teachers and students on a level playing field, allowing all to become collaborative partners who seek new knowledge. These discussions destabilize and transform the dichotomous relationships between teachers/learners, native/non-native language users, and knowledge-providers/knowledge-seekers. In the process, we are able to make visible the invisible assumptions that each of us hold, are reminded that what is normal to some is not necessarily so for others, and are able to recognize and appreciate diverse interpretations.

The project also led some students to pay attention to power struggles among different languages in Japan. For instance, a pair of students, both of whom were also studying Korean language at the time, investigated linguistic landscape in a Korea Town in Osaka, Japan. Korea Towns are areas where *zainichi* Koreans reside. *Zainichi* Koreans are ethnic Koreans with roots to the Korean peninsula during Imperial Japan’s colonization. They are also one of the ethnic minority groups in Japan who have been subjected to various forms of legal and societal discrimination, despite their permanent resident status in Japan. Taking advantage of their multiple languages as resources (i.e., English, Japanese, and Korean), they started exploring the streets with a simple question: Why are



Fig. 4 Emergency road sign.

some words written in Korean while other words are in Japanese? Seeking answers to the initial question led them to pose further questions: Why were Japanese translations provided for Korean phrases but not the other way around? Who are the signs trying to cater to, and what are their intentions and purposes? These inquiry-based explorations made them recognize that the use of Korean language on store signs was not necessarily for transactional purposes, but rather was used as design to capture the attention of potential customers. In contrast to the abundance of Korean scripts inside the Korea Town, the public signage across the street from the area that called to “Keep Rivers Clean” was only written in Japanese. This gap made them recognize an implicit message of assimilation policy by the Japanese government. Consequently, they decided to transform the “River” sign by redesigning it with translations in Chinese, Korean and English, as well as providing *furigana*—a reading aid, consisting of smaller *hiragana*, to indicate its pronunciation—for *kanji* words so that even children would be able to understand the sign. They wanted their new sign to reflect the community’s diversity, and to be more inclusive and helpful for the needs of everyone who lives there.

Focusing on how and why different languages are used in the analytic process, the students came to recognize the marginalized status of the Korean language in Japan’s official discourse, which is reflective of the *zainichi* Koreans’ position in Japanese society. This further made the students think critically about their own lives and led them to recognize linguistic injustices in their own communities. One student remarked that:

It’s funny because before this project, I’ve never considered [linguistic injustice] in America But look at things in the critical light in Japan, and in America which is more diverse than Japan, signs still tend to be in English. When I come across with the signs in Spanish, I take note of it because, I’m like, “that’s not usual.” I’m from [a] really white area But because I live in a neighborhood that’s very liberal so there are all sorts of signs saying like, “you are also our neighbors” in English, Spanish, and Arabic, because these are hot bun issue[s] around immigration. And it was just interesting to me because all those languages are only utilized when making a political point or stand, but then in everyday life they really aren’t So it made me think about that critically to my own life.

Problematizing and questioning what seems to be normal and ordinary would make the students aware of issues such as power struggles between different languages and unspoken political and sociocultural intentions behind semiotic choices (e.g., languages, words, illustrations, etc.). The experience that students gain from projects such as this one allows them to understand Japanese language and society on a deeper level, not to mention potentially making them more aware of their familiar surroundings.

Comparative reading of newspaper reports: shuttling between languages to become linguistic mediators

By taking advantage of students’ language and cultural repertoires as multilingual language users, creating a context where they can actively engage in “shuttling between languages” (Canagarajah, 2013) or “translanguaging” (Garcia, 2009; Garcia and Wei, 2014) in a foreign language classroom has significant benefits in developing their critical literacies and nurturing their agency as linguistic/cultural mediators (for details, see Kumagai and Iwasaki, 2015).

This unit, comparative reading of newspaper reports, was designed for intermediate-high to advanced level learners (Iwasaki and Kumagai, 2015), who have started reading texts beyond those in the textbooks. In this unit, students read two Web-based newspaper articles about the same event: one published by *The New York Times* (www.nytimes.com) in English (“4 American teenagers arrested”) and the other by *Asahi Shimbun Digital* (www.asahi.com) in Japanese (“米兵の子4人を殺人未遂容疑で逮捕” [4 children of American military arrested for attempted murder]). The event involved American teenagers from an American military base in Japan who allegedly strung a rope across a road, toppling a Japanese female motorbike rider and seriously injuring her. Crucially, the articles differed in terms of the sociopolitical and cultural background knowledge presumed of the target readers. Such differences led to differences in the writers’ selection of information and linguistic resources, which may give readers vastly different impressions and ideas about the reported event.

Three class sessions were spent on the unit. In the first session, students engaged in pre-reading activities by reading the article in *The New York Times*. They summarized the event in Japanese, using a list of keywords provided by the instructor. In so doing, the students familiarized themselves with the event and with new vocabularies that they needed in order to discuss the incident in Japanese, as well as to read the aforementioned Japanese article.

The students were also given a set of reading questions, and were instructed to read the *Asahi* article at home. During the second session, the students discussed the textual features and content of the *Asahi* article, using the reading questions as a guide. The questions guided students to pay attention to textual features such as sentence-ending modality and the use of direct quotes; the questions also helped them to achieve general comprehension of the article’s content.

In the third session, the students discussed their analyses of the text in depth, particularly paying attention to the effects of language choice on the reader. For example, the effect of the use of causatives was discussed by comparing two sentences: one containing causative predicates (i.e., “made [the woman] topple”; “made [her] get injured”) and the other containing active intransitive predicates (i.e., “[the woman] toppled over”; “got injured”). The students also compared the information included in the two articles, as well as the effect of the inclusion (and exclusion) of a given piece of information on the respective readers. For instance, while *The New York Times* article included a mention of a boy having sought help from a passer-by for the injured woman, the *Asahi* article contained two rather conflicting statements (in the form of direct quotes) made by two of the four teenagers (one stating that the rope stringing was done without the knowledge of the other two; the other stating that the action was decided by all four). The

Asahi article also included a drawing of the accident site, illustrating how the rope was strung. At the conclusion of the unit, each student was instructed to write an email message to her imagined Japanese friend, whose American colleagues only read *The New York Times*. The task was to inform the friend that her American colleagues' impression of the episode may differ due to the way the event is reported in *The New York Times*.

In class discussions, the students readily interpreted and independently realized that the purpose of each article was shaped by the readers' interests assumed by the writer, which consequently affected the details and types of information included in each article. In addition, the students noticed that the wording in the headlines and the use of quotes had created different impressions regarding the seriousness of the incident ("a bad image about Americans" in *Asahi*; just "kids [are] messed up" in *The New York Times*). The classroom interactions demonstrated that students were actively constructing meaning using their knowledge of both English and Japanese. As for the discussions regarding the use of causatives, the students articulated that while the sentence without the causative described the event as though it was "an innocent accident," the one with the causative highlighted the intention behind the (evil) act done to the woman who had no control over the incident. They recognized how a certain grammatical structure choice (in this case, causatives) evoked different nuances and images, and communicated different messages to the readers.

During the individual interviews conducted after the course was over, students' awareness of three key areas was identified. First, students recognized how different sociocultural perspectives (such as histories and experiences) are reflected in the ways news articles are written. Second, students realized that there is no clear line between "facts" and "opinions" in news reports. One student noted that even though an article may appear to be a non-biased, objective reporting of the news, it can in fact be biased by the writer's perspectives—"opinionated," to use the student's word—marked by a choice of words and grammatical structures. Third, when asked about the roles that they could potentially play because of their ability to read in multiple languages with an understanding of multiple cultures, most of the students considered their own roles to be significant. They demonstrated a growing awareness that they could play the role of mediators for those who do not have the ability or a desire to go beyond their language or cultural boundaries.

Summary of praxes

The three praxes described above all share important features. First, they all encourage students' explorations of real-life language use. The explorations and ensuing examinations are guided by the students' own interests and curiosity. The role of the teachers is not to provide answers (as they do not have them), but rather to ponder and engage in the inquiry with the students. Second, all three praxes focus on students becoming sensitive to the choice of language (and other semiotic devices). Why are particular choices made? What factors motivate such choices? What effects do the choices have on readers (including the students themselves)? Do you feel comfortable with a worldview represented and a position you are placed in by such choices? If not, how can we change them? Inviting students to engage in such dialogs is immensely important, allowing them to more fully understand the language and consequent social order.

Third, these praxes value students' multilingual and multicultural resources, as well as other knowledge and experiences (academic and otherwise) they possess, by encouraging them to actively enlist such resources during the learning process. Traditionally, the monolingual approach has been the norm in foreign language classrooms, and the use of languages other than the target one is strictly prohibited. However, strategic use of multiple languages, and allowing students to shuttle between languages as people normally do in real life (Canagarajah, 2013; Garcia, 2009; Otusji, 2019, etc.), is necessary for enabling students to fully engage in critical inquiries and meaning making (Kato and Kumagai, 2020; Kumagai and Kono, 2018). Encouraging students to draw on their knowledge of languages/cultural resources, and guiding them to pay closer attention to the semiotic choices in texts, can lead them to the analytical disposition that "seeks to uncover the social interests at work, to ascertain what is at stake in textual and social practices" (Janks, 2010, p. 12).

Fourth, the praxes provide a space for students to relate to communities outside of the classroom in various ways. The *Katakana* Project and the Linguistic Landscape Project both disseminated knowledge to the communities of other Japanese language learners by creating tangible products. The comparative reading unit, which tasked students with writing a letter to an imagined friend, made them recognize the important role they could play in society as linguistic/cultural mediators—a role I hope they take seriously and embody long after they leave the classroom.

Conclusion

The advent of internet and mobile devices has brought us unprecedented access to information and means of communication. It allows us to access information of all kinds almost instantly, from every corner of the globe. It also provides us with platforms and opportunities to participate in the creation and dissemination of information. At the same time, however, it bombards us with a flood of information, some of which is reliable, some of which is not. In the US, the age of disinformation facilitated by social networking systems brought about distrust and attacks on the mainstream media. People are living in an echo chamber of their own beliefs, creating a world that is gravely divided. Critical literacies—abilities to question and assess the legitimacy of information, to seek alternative interpretations by reflecting on one's own viewpoint as well as other's perspectives, and to imagine and work toward creating more equitable societies—are in great need, now more than ever.

Critical literacies in foreign language learning provide learners not merely with functional skills, but also with the conceptual tools that allow them to critique and engage with society. They assist learners in appreciating and recognizing the social contexts of language use, particularly with respect to the roles that language plays in establishing and maintaining power relations in society (Reagan and Osborn, 2021). The development of critical literacies in foreign language has profound implications because it allows learners to look at their world from multiple perspectives and equips them to voice concerns informed by alternative viewpoints. Learning a new language, and thus seeking to understand (and to become) the “Other,” requires humility and empathy, a disposition that is extremely important in this diverse and increasingly divided world.

References

- Alderson, J.C., 1984. Reading in a foreign language: a reading problem or a language problem? In: Alderson, J.C., Urquhart, A.H. (Eds.), *Reading in a Foreign Language*. Longman, New York, pp. 1–27.
- Alford, J., 2001. Learning language and critical literacy: adolescent ESL students. *J. Adolesc. Adult Literacy* 45 (3), 238–242.
- Anderson, G.L., Irvine, P., 1993. Informing critical literacy with ethnography. In: Lankshear, C., McLaren, P.L. (Eds.), *Critical Literacy: Politics, Praxis, and the Postmodern*. State University of New York Press, Albany, NY, pp. 81–104.
- Banno, E., Ohno, Y., Sakane, Y., Shinagawa, C., 2021. *Genki: An Integrated Course in Elementary Japanese*, third ed. The Japan Times, Tokyo, JP.
- Barton, D., Hamilton, M., Ivanic, R. (Eds.), 2000. *Situated Literacies: Reading and Writing in Context*. Routledge, New York.
- Barton, D., 2007. *Literacy: An Introduction to the Ecology of Written Language*, second ed. Blackwell, Cambridge.
- Baynham, M., 1995. *Literacy Practices: Investigating Literacy in Social Contexts*. Longman Singapore Publishers, Singapore.
- Bernhardt, E.B., 1991. Development in second language literacy research: retrospective and prospective views for the classroom. In: Freed, B.F. (Ed.), *Foreign Language Acquisition Research and the Classroom*. D.C. Heath, Lexington, MA, pp. 221–251.
- Bourdieu, P., 1991. *Language and Symbolic Power*. Harvard University Press, Cambridge, MA.
- Brodkey, L., 1996. *Writing Permitted in Designated Areas Only*. University of Minnesota Press, Minneapolis.
- Canagarajah, S., 2013. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge, New York.
- Case, R., Ndura, E., Righettini, M., 2005. Balancing linguistic and social needs: evaluating texts using a critical language awareness approach. *J. Adolesc. Adult Literacy* 48 (5), 374–391.
- Cenez, J., Gorter, D., 2008. The linguistic landscape as an additional source of input in second language acquisition. *Int. Rev. Appl. Ling.* 46 (4), 267–287.
- Cheah, Y.M., 2001. From prescription to participation: moving from functional to critical literacy in Singapore. In: Comber, B., Simpson, A. (Eds.), *Negotiating Critical Literacies in Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 69–82.
- Chern, C., Dooley, K., 2014. Learning English by walking down the street. *ELT J.* 68 (2), 113–123.
- Chikamatsu, N., 2021. Translanguaging in language and area-studies curriculum: a Japanese FLAC course of *Minamata* and *Fukushima* in environmental humanities. In: de la Fuente, M.J. (Ed.), *Education for Sustainable Development in Foreign Language Learning*. Routledge, New York, pp. 215–232.
- Cho, H., Johnson, P., 2020. “We have to focus on improving our and our next generation’s rights!” Exploring critical literacy in a third space for Korean female high school students. *Crit. Inq. Lang. Stud.* 18 (2), 174–202.
- Chun, C., 2009. Critical literacies and graphic novels for English-language learners: teaching Maus. *J. Adolesc. Adult Literacy* 53 (2), 144–153.
- Clark, R., Ivanic, R., 1997. *The Politics of Writing*. Routledge, New York.
- Comber, B., Simpson, A. (Eds.), 2001. *Negotiating Critical Literacies in Classrooms*. Routledge, New York.
- Cope, B., Kalantzis, M. (Eds.), 2000. *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, New York.
- Dagenais, D., Moore, D., Sabatier, C., Lamarre, P., Armand, F., 2009. Linguistic landscape and language awareness. In: Shohamy, E., Gorter, D. (Eds.), *Linguistic Landscape: Expanding the Scenery*. Routledge, New York, pp. 253–269.
- Dahlberg-Dodd, H.E., 2020. Script variation as audience design: Imagining readership and community in Japanese yuri comics. *Language in Society* 49 (3), 357–378.
- Dahlberg-Dodd, H.E., 2022. Katakana and the mediated other: script variation in fantastical narratives. *Jpn. Stud.* 1–19.
- Delpit, L., 1995. *Other People’s Children: Cultural Conflict in the Classroom*. The New Press, New York.
- Devitt, S., 1997. Interacting with authentic texts: multilayered processes. *Mod. Lang. J.* 81, 457–469.
- Doerr, N., Kumagai, Y., 2014. Power of language ideologies: challenging the notion of “foreign loanwords” in Japanese-as-a-foreign-language classroom. *Writing Syst. Res.* 6 (2), 148–166.
- Elley, W.B., 1984. Exploring the reading difficulties of second-language learners in Fiji. In: Alderson, J.C., Urquhart, A.H. (Eds.), *Reading in a Foreign Language*. Longman, New York, pp. 281–301.
- Fabbian, C., Carney, E., 2018. Critical pedagogy, social justice, and prosocial identities in the Italian classroom (Part I). *Italica* 95 (2), 241–271.
- Fabbian, C., Carney, E., 2018. Critical pedagogy, social justice, and prosocial identities in the Italian classroom (Part II). *Italica* 95 (3), 417–433.
- Fairclough, N., 1992. *Critical Language Awareness*. Addison Wesley Longman.
- Férez Mora, P.A., Coyle, Y., Dorado Otero, Á., 2022. Foreign Language Students’ Views on FL and Critical Literacies. *Language Awareness*, pp. 1–20.
- Freebody, P., Luke, A., 1990. Literacies programs: debates and demands in cultural context. *Prospect* 5 (3), 7–16.
- Freire, P., Macedo, D., 1987. *Literacy: Reading the Word and the World*. Bergin & Garvey, West Port, CT.
- Garcia, O., Wei, L., 2014. *Translanguaging: Language, Bilingualism and Education*. Palgrave Pivot, New York.
- Garcia, O., 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell, Malden, MA.
- Gee, J.P., 1987/1991. What is literacy? In: Mitchell, C., Weller, K. (Eds.), *Rewriting Literacy: Culture and the Discourse of the Other*. Bergin & Garvey, Westport, CT, pp. 3–11.
- Giroux, H.A., 1987. Introduction: literacy and pedagogy of political empowerment. In: Freire, P., Macedo, D. (Eds.), *Literacy: Reading the Word and the World*. Bergin & Garvey, Westport, CT, pp. 1–27.
- Gottlieb, N., 2005. *Language and Society in Japan*. Cambridge University Press, New York.
- Grabe, W., 1991. Current development in second language reading research. *Tesol Q.* 25, 375–406.
- Gyogi, E., 2019. Critical literacy in beginner-level Japanese language classrooms. In: Clements, P., Krause, A., Bennett, P. (Eds.), *Diversity and Inclusion*. JALT, Tokyo.
- Hammond, J., Macken-Horarik, M., 1999. Critical literacy: challenges and questions for ESL classrooms. *Tesol Q.* 33, 528–544.
- Huang, S., 2011a. Reading “further and beyond the text”: student perspectives of critical literacy in EFL reading and writing. *J. Adolesc. Adult Literacy* 55 (2), 145–154.
- Huang, S., 2011b. “Critical literacy helps wipe away the dirt on our glasses”: towards an understanding of reading as ideological practice. *Engl. Teach. Pract. Critiq.* 10, 140–164.
- Huang, S., 2012. The integration of “critical” and “literacy” education in the EFL curriculum: expanding the possibilities of critical writing practices. *Lang. Cult. Curric.* 25 (3), 283–298.
- Huang, S.Y., 2013. Revising identities as writers and readers through critical language awareness. *Engl. Teach. Pract. Critiq.* 12 (3), 65–86.
- Huh, S., 2016. Instructional model of critical literacy in an EFL context: balancing conventional and critical literacy. *Crit. Inq. Lang. Stud.* 13 (3), 210–235.

- Iventosch, M.S., 2012. Teaching Japanese written language. In: Goodman, K., Wang, S., Iventosch, M.S., Goodman, Y. (Eds.), *Reading in Asian Languages: Making Sense of Written Texts in Chinese, Japanese, and Korean*. Routledge, New York, pp. 236–257.
- Iwahara, K., Hatta, K., Maehara, A., 2003. The effects of a sense of compatibility between type of script and word in written Japanese. *Read. Writ.* 16, 377–397.
- Iwasaki, N., Kumagai, Y., 2008. Towards critical approaches in an advanced level Japanese course: theory and practice through reflection and dialogues. *Jpn. Lang. Literature* 42, 123–156.
- Iwasaki, N., Kumagai, Y., 2015. *The Routledge Intermediate to Advanced Japanese Reader: A Genre-Based Approach to Reading as a Social Practice*. Routledge, London.
- Iwasaki, N., Kumagai, Y., 2020. Making it your own by adapting it to what's important to you: plurilingual critical literacies to promote L2 Japanese user's sense of ownership of Japanese. In: Bagga-Gupta, S., Golden, A., Holm, L., Lausen, H.P., Pitkänen-Huhta, A. (Eds.), *Reconceptualizing Connections Between Language, Literacy, and Learning*. Springer Nature Switzerland, Cham: Switzerland, pp. 165–186.
- Janks, H., Vasquez, V., 2011. Critical literacy revisited: writing as critique. *Engl. Teach. Pract. Critiq.* 10 (1), 106.
- Janks, H., 2010. *Language and Power*. Routledge, New York.
- Kato, R., Kumagai, Y., 2020. Translingual practices in a monolingual society: discourses, learners' subjectivities and language choices. *Int. J. Biling. Educ. Biling.* <https://doi.org/10.1080/13670050.2020.1799318>.
- Keneman, M., 2015. Finding voice in the foreign language classroom: reading, writing, and performing slam poetry to develop critical literacies. In: Parkes, L., Ryan, C.M. (Eds.), *AAUSC 2015 Volume—Issues in Language Program Direction: Integrating the Arts: Creative Thinking About FL Curricula and Language Program Direction*. Cengage Learning, Boston, MA, pp. 125–149.
- Kikuchi, S., 1994. Daigakusei no gairaigo ishiki (1): imeeji, hyoki, goshu ishiki no choosa kara [College students perceptions for *gairaigo* (1): research on image, scripts, and vocabulary types]. *J. Center Educ. Res. Pract.* (4), 61–73.
- Ko, M.Y., Wang, T.F., 2013. EFL learners' critical literacy practices: a case study of four college students in Taiwan. *Asia-Pac. Educ. Res.* 22 (3), 221–229.
- Koyama, N., 2016. Emergence of critical reading in L2 literature seminars: applications of discourse analysis in language pedagogy. *Jpn. Lang. Literature* 50 (1), 1–24.
- Kramsch, C., 1993. *Context and Culture in Language Teaching*. Oxford university press.
- Kress, G., 2000. Multimodality. In: Cope, B., Kalantzis, M. (Eds.), *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, New York, pp. 182–202.
- Kubota, R., Austin, T., Saito-Abbott, Y., 2003. Diversity and inclusion of sociopolitical issues in foreign language classrooms: an exploratory survey. *Foreign Lang. Ann.* 36 (1), 12–24.
- Kubota, R., 1996. Nihongo kyooku ni okeru hihan kyooku, hihanteki yomikaki kyooku [Critical pedagogy and critical literacy in teaching Japanese]. *Sekai no Nihongo* 6, 35–48.
- Kubota, R., 2010. Critical multicultural education and second/foreign language teaching. In: May, S., Sleeter, C.E. (Eds.), *Critical Multiculturalism: Theory and Praxis*. Routledge, New York, pp. 105–118.
- Kumagai, Y., Fukai, M., 2009. Nihongo gakushuu ni okeru hihan-see, soozoo-see no ikusee eno kokoromi: "Kyookasho kakikae" purojekuto [Cultivating critical thinking and creativity in Japanese language learning: "Revising textbook" Project]. *Jpn. Educ. Around Globe* 19, 181–202.
- Kumagai, Y., Iwasaki, N., 2011. What it means to read "critically" in a Japanese language classroom: students' perspective. *Crit. Inq. Lang. Stud.* 8 (2), 125–152.
- Kumagai, Y., Iwasaki, N., 2015. Reading words to read worlds: a genre-based critical multiliteracies curriculum in intermediate/advanced Japanese language education. In: Kumagai, Y., Lopez-Sanchez, A., Wu, S. (Eds.), *Multiliteracies in World Language Education*. Routledge, New York, pp. 107–131.
- Kumagai, Y., Kono, K., 2018. Collaborative curricular initiatives: linking language and literature courses for critical and cultural literacies. *Jpn. Lang. Literature* 52 (2), 247–276.
- Kumagai, Y., Takahashi, Y. (in print). "Of course just in Japanese! Why would the government do anything else? This is Japan!": recognizing language ideology through a linguistic landscape project in a Japanese language classroom. In: Johnson, S., et al. (Eds.), *How We Take Action: Social Justice in K-16 Language Classrooms*. Information Age Publishing.
- Kumagai, Y., 2007. Tension in a Japanese language classroom: an opportunity for critical literacy?. In: Kutoba, R., Austin, T. (Eds.), *Critical Perspectives on World Language Education in the US: Critical Inquiry in Language Studies [Special Issue]*, vol. 4, pp. 85–116.
- Kumagai, Y., 2011. Kuritikanu riterashii no ikusee e mukete: Katakana purojekuto [Developing critical literacy: katakana project]. In: Sato, S., Kumagai, Y. (Eds.), *Shakai sanko o mezasu nihongo kyooku: shakai ni kakawaru, tsunagaru, hatarakikakeru*. Hituzi Shoboo, Tokyo, pp. 3–18.
- Kumagai, Y., 2020. "*Ekkyō-bungaku*" as crossing the border of language: implications for learners of Japanese. In: Doerr, N.M. (Ed.), *Global Education Effect: New Borders and Imagined Communities Around "Japan"*. Routledge, New York, pp. 214–234.
- Kunert, H., 2020. Contemporary loanwords in hiragana: an analysis of typical traits and context. *Jpn. Stud.* <https://doi.org/10.1080/10371397.2019.1682935>.
- Kuo, J.M., 2009. Critical literacy and a picture-book-based dialogue activity in Taiwan. *Asia Pac. Educ. Rev.* 10, 483. <https://doi.org/10.1007/s12564-009-9048-6>.
- Kuo, J.M., 2015. *Guji guji* goes to college: promoting critical literacy in Taiwan. *TESL EJ* 19 (1), n1.
- Landry, R., Bourhis, R.Y., 1997. Linguistic landscape and ethnolinguistic vitality: an empirical study. *J. Lang. Soc. Psychol.* 16, 23–49.
- Lau, S.M.C., Tian, Z., Lin, A.M., 2021. Critical literacy and additional language learning: an expansive view of translanguaging for change-enhancing possibilities. In: Pandya, J.Z., Mora, R.A., Alford, J.H., Golden, N.A., De Roock, R.S. (Eds.), *The Handbook of Critical Literacies*. Routledge, New York, pp. 381–390.
- Lau, S., 2012. Reconceptualizing critical literacy teaching in ESL classrooms. *Read. Teach.* 65 (5), 325–329.
- Lee, Y.J., 2017. First steps toward critical literacy: interactions with an English narrative text among three English as a foreign language readers in South Korea. *J. Early Child. Literacy* 17 (1), 26–46.
- Leung, G.Y., Wu, M.-H., 2012. Linguistic landscape and heritage language literacy education: a case study of linguistic rescaling in Philadelphia Chinatown. *Writ. Lang. Lit.* 15 (1), 114–140.
- Lin, A.M.Y., 2001. Resistance and creativity in English reading lessons in Hong Kong. In: Comber, B., Simpson, A. (Eds.), *Negotiating Critical Literacies in Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 83–100.
- Lin, W.-Y., 2012. Teachers' reflections on Chinese reading miscue analysis. In: Goodman, K., Wang, S., Iventosch, M.S., Goodman, Y. (Eds.), *Reading in Asian Languages: Making Sense of Written Texts in Chinese, Japanese, and Korean*. Routledge, New York, pp. 211–220.
- Ludewig, J., 2017. A look over the wall: using the graphic novel *drüben!* to teach linguistic, literary, cultural, and visual skills. *Unterrichtspraxis* 50 (2), 162–170.
- Luk, J., Lin, A., 2015. Voices without words: doing critical literate talk in English as a second language. *Tesol Q.* 49 (1), 67–91.
- Luke, A., Dooley, K., 2011. Critical literacy and second language learning. In: Hinkel, E. (Ed.), *Handbook of Research in Second Language Teaching and Learning*. Routledge, New York, pp. 856–868.
- Luke, A., 1995. When basic skills and information processing just aren't enough: rethinking reading in new times. *Teach. Coll. Rec.* 97 (1), 94–115.
- Luke, A., 2012. Critical literacy: foundational notes. *Theory Into Pract.* 51 (1), 4–11.
- Makino, S., Hatasa, Y.A., Hatasa, K., 1998. *Nakama 1: Japanese Communication, Culture, Context*. Cengage Learning.
- Malinowski, D., 2010. Showing seeing in the Korean linguistic cityscape. In: Shohamy, E., Ben-Rafael, E., Barni, M. (Eds.), *Linguistic Landscape in the City*. Multilingual Matters, Tonawanda, NY, pp. 199–215.
- Mitsikopoulou, B., 2020. Genre instruction and critical literacy in teacher education: features of a critical foreign language pedagogy in a university curriculum. *L2 J.* 12 (2), 94–109.
- Motschenbacher, H., 2021. Corpus-based considerations on critical literacy in ELT: the linguistic representation of ricky martin in the news media. In: Paiz, J.M., Coda, J.E. (Eds.), *Intersectional Perspectives on LGBTQ+ Issues in Modern Language Teaching and Learning*. Springer International Publishing, pp. 217–259.
- Nagara, S., 1990. *Japanese for Everyone: A Functional Approach to Daily Communication*. Gakken, Tokyo, JP.
- Otsuji, E., 2019. Metrolingualism in transitional Japan. In: Heinrich, P., Ohara, Y. (Eds.), *Routledge handbook of Japanese Sociolinguistics*. Routledge, Oxton, pp. 143–157.
- Pennycook, A., 2001. *Critical Applied Linguistics: A Critical Introduction*. Lawrence Erlbaum, Mahwah, NJ.
- Reagan, T.G., Osborn, T.A., 2021. *World Language Education as Critical Pedagogy: The Promise of Social Justice*. Routledge, New York.

- Robertson, W.C., 2017. He's more katakana than kanji: indexing identity and self-representation through script selection in Japanese manga (comics). *J. Sociolinguistics* 21 (4), 497–520.
- Robertson, W.C., 2021. *Scripting Japan: Orthography, Variation, and the Creation of Meaning in Written Japanese*. Routledge, New York.
- Romero, E.D., Bobkina, J., 2021. Exploring critical and visual literacy needs in digital learning environments: the use of memes in the EFL/ESL university classroom. *Think. Skills Creativ.* 40, 100783.
- Rowland, L., 2013. The pedagogical benefits of a linguistic landscape project in Japan. *Int. J. Biling. Educ. BiLing.* 16 (4), 494–505.
- Ryshina-Pankova, M., 2013. Understanding “green Germany” through images and film: a critical literacy approach. *Unterrichtspraxis* 46 (2), 163–184.
- Sato, S., Hasegawa, A., Kumagai, Y., Kamiyoshi, U., 2017. Content-based instruction (CBI) for the social future: a recommendation for critical content-based language instruction (CCBI). *L2 J.* 9 (3), 50–69.
- Sayer, P., 2009. Using the linguistic landscape as a pedagogical resource. *ELT J.* 64 (2), 143–154.
- Seeley, C., 2000. *A History of Writing in Japan*. University of Hawaii Press, Honolulu.
- Shohamy, E., Waksman, S., 2009a. Introduction. In: Shohamy, E., Gorter, D. (Eds.), *Linguistic Landscape: Expanding the Scenery*. Routledge, New York, pp. 1–10.
- Shohamy, E., Waksman, S., 2009b. Linguistic landscape as an ecological arena: modalities, meanings, negotiations, education. In: Shohamy, E., Gorter, D. (Eds.), *Linguistic Landscape: Expanding the Scenery*. Routledge, New York, pp. 313–330.
- Shor, I., Freire, P., 1987. What is the “dialogical method” of teaching? *J. Educ.* 169 (3), 11–31.
- Shor, I., 1999. What is critical literacy? In: Shor, I., Pari, C. (Eds.), *Critical Literacy in Action: Writing Words, Changing Worlds*. Boynton/Cook Heinemann, Portsmouth, NH, pp. 1–30.
- Sibanda, R., 2021. Beyond the written word: newspapers as critical literacy teaching tools in the South African classroom. *English Lang. Teach. Educ. J.* 4 (3), 174–186.
- Smith, J., Schmidt, D., 1996. Variability in written Japanese: Towards a sociolinguistics of script choice. *Visible Lang.* 30 (1), 47–71.
- Street, B.V., 2014. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Routledge.
- Taylor, I., Taylor, M.M., 2014. *Writing and Literacy in Chinese, Korean and Japanese*, revised ed. John Benjamins Publishing Company, Philadelphia.
- Tohsaku, Y., 2006. *Yookoso! Invitation to Contemporary Japanese*. McGraw Hill.
- Tyson, R.E., 2012. A successful mixture of alphabetic and non-alphabetic writing: Chinese characters in Korean. In: Goodman, K., Wang, S., Iventosch, M.S., Goodman, Y. (Eds.), *Reading in Asian Languages: Making Sense of Written Texts in Chinese, Japanese, and Korean*. Routledge, New York, pp. 92–98.
- Ukita, J., Sugishima, I., Minagawa, N., Inoue, M., Kasyu, K., 1996. Psychological research for written forms of Japanese words. *Japanese Psychological Monograph*, No. 25. Japanese Psychological Association, Tokyo.
- Vasquez, V.M., Janks, H., Comber, B., 2019. Critical literacy as a way of being and doing. *Lang. Arts* 96 (5), 300–311.
- Verma, R., 2008. Beyond Tacos and pizza: critical literacy in the world language classroom. In: Wallowitz, L. (Ed.), *Critical Literacy as Resistance: Teaching for Social Justice Across the Secondary Curriculum*. Peter Lang, pp. 165–178.
- Wallace, C., 2001. Critical literacy in the second language classroom: power and control. In: Comber, B., Simpson, A. (Eds.), *Negotiating Critical Literacies in Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 209–228.
- Wallace, C., 2003. *Critical Reading in Language Education*. Palgrave Macmillan, New York.
- Weng, T.H., 2021. Creating critical literacy praxis: bridging the gap between theory and practice. *RELC J.* <https://doi.org/10.1177/0033688220982665>.
- Weninger, C., 2018. A contextual critique of critical literacy: Freirean “generative themes” and their impact on pedagogic practice. *Discourse* 39 (1), 84–97.
- Weninger, C., 2020. Investigating ideology through framing: a critical discourse analysis of a critical literacy lesson. *Classr. Discourse* 11 (2), 107–128.
- Wooldridge, N., 2001. Tensions and ambiguities in critical literacy. In: Comber, B., Simpson, A. (Eds.), *Negotiating Critical Literacies in Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 259–270.

Language policy and planning as a discursive accomplishment

Anthony J. Liddicoat, Department of Applied Linguistics, University of Warwick, Coventry, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	123
The scope of language policy and planning	123
The positioning of multilingualism in language policy	124
The normalization of monolingualism	124
Hierarchies of value	126
Promoting multilingualism	127
Valuing and validating multilingualism—future directions for policy and research	128
References	128

Introduction

Language policy and planning (LPP) is an intervention into a language ecology in order to influence some aspect of language learning or use. As an intervention it shapes the language ecology in particular ways in order to address particular problems or to achieve particular goals. In shaping linguistic ecologies, decision-making about languages is inevitably consequential for the whole language ecology (Mühlhäusler, 2000); for example, promoting one language in education has consequences for other languages, which be marginalized in the planned ecology as a result of increasing space and attention being allocated to the promoted language. Language ecologies are by their nature diverse and linguistic diversity is the normal state of most larger social groups and even very small social groups such as families. LPP is thus an intervention into linguistic diversity and it often has as its objective simplifying this diversity, for example by declaring a particular language, or a small set of languages, as being the norm for various social functions, whether these are the work of government, the use and learning of languages in education, or other communicative needs. This chapter will examine the ways that LPP has impacted on linguistic diversity in societies by considering how LPP is consequential for societal and individual multilingualism, but before doing this will consider the nature and scope of LPP as a field of scholarship.

The scope of language policy and planning

In traditional LPP research it has been conventional to view the scope of study in terms of the actions of governments or governmental bodies, and research has taken the form of examining policy texts, language laws, education laws and other formal declarations (Ricento, 2000; Taylor-Leech and Benson, 2017). More recently, LPP scholarship has come to recognize that decision-making about languages exists at other levels than government and takes forms other than texts and has sought to develop a critical perspective on decision-making about language and the ways in which such decision-making reflects, reproduces, and enacts of power, ideology, and other social processes. As a result LPP scholarship has seen a movement away from the idea that language problems can be identified by objective processes and has recognized that identifying an aspect of language learning or use as problematic is highly subjective and influenced by ideological, attitudinal, economic and politic interests (Liddicoat and Taylor-Leech, 2020). This has led to an increased focus on the effects of policies, and particularly on the negative effects that policies could have on non-dominant groups (e.g., Pennycook, 1994). As LPP research has come to consider the subjectivity of decision-making, it has also come to recognize that such decision-making is not the unique preserve of macro-level institutional structures, but rather is distributed through societies and that multiple actors are engaged in decisions at different levels in society. This thinking reflects Foucault's (1975) recognition that power is something that is multivalent and distributed and applies such thinking to understanding the ways that various actors intervene in linguistic ecologies to shape how languages are used and learned. As a result, LPP research has come to focus attention on different levels of society in which decisions are made that are consequential for understanding languages and their ecologies, identified as macro, meso, and micro levels (Baldauf, 2006; Kaplan and Baldauf, 1997).

While initial thinking about these levels seemed to assign these layers to particular actors—e.g., the macro-level equates with governmental language policy actors, the micro-level with local communities, and the meso-level with intervening institutions, Liddicoat (2020b) has argued that is more appropriate to consider them in a contextualized way: the macro-level as the topmost level of language policy that is relevant in a context, the micro-level as the most local level in the context and the meso-level as any intervening levels. Moreover, there has been recognition that language policy is not simply a top-down process in which macro-level actors create policies that are then transmitted to lower social levels for implementation, but rather involves the potential for either top-down or bottom-up processes, or combinations of these, as in the case of "language policy from below" (Alexander, 1992) where the direction of influence is from the micro to the macro.

There has also been a shift in thinking about what constitutes language policy within LPP work, recognizing that texts (language laws, language policy documents, etc.) are just one aspects of the ways that decision about languages are represented in societies. In particular there has been a recognition that language policies can be seen in discourses about languages and their use (Liddicoat, 2013; Lo Bianco, 2005), which may occur in many different formats, and which play a profound role in shaping what happens in language learning and use. This involves a recognition that LPP is profoundly ideological—it is an encapsulation of systems of belief about languages—that are transmitted through discourses and ultimately shape language ecologies (Liddicoat, 2020a). Moreover, it is also possible to understand language policies by examining the language practices of a social group (Bonacina-Pugh, 2012, 2020). Such practice represented enacted language policy and reveal the often-unarticulated norms and conventions that shape actions. This movement has been accompanied by an growing interest in covert aspects of language policy (Shohamy, 2006) and the ways these impact on language learning and use. While language policy texts are typically overt statements about languages, attention is coming to be focused increasingly on covert aspects of policy; that is, the tacit norms, beliefs, attitudes, etc. that influence how languages are learned or used, or the silences that can be found in articulations of policy that can be as consequential for languages and their speakers as explicit statements.

The positioning of multilingualism in language policy

As language policy is an intervention in language ecologies and language ecologies are constellations of multiple languages within a social space, language policy is an important site for understanding how ideologies and practices of language use play out. In particular, language policy represents enactments that construct the social world of languages in particular ways (Gee, 1994). The discussion below will consider some of the key issues that are involved in understanding the nature of language policy interventions and how such interventions structure or restructure the language ecologies on which they act. In particular, it will consider how language policy engages with linguistic diversity and so creates particular understandings of the nature and meaning of diversity.

The normalization of monolingualism

When decisions are made about language learning and use in multilingual settings, linguistic diversity is often perceived as a problem that needs to be resolved by reducing the diversity of languages, at least in public domains (see also Ruiz, 1984). One particular focus of interventions in language ecologies since the Enlightenment has been the desire to achieve a unification of language and nation; what has been called the one-nation-one-language ideology (Ducard, 2016). While this ideology can be traced to earlier periods, notably the Reformation, it emerged strongly during the French Revolution as part of its rationalist nation-building agenda that argued for a single language for the French state on two grounds (Geeraerts, 2003). The first was pragmatic and focused on the democratic aims of the revolution: linguistic diversity as perceived as a barrier to communication that would prevent them from understanding their social rights and obligations and limit their ability to participate equitably in the society. The second was symbolic and based on the idea that a single language would create and represent a single, unified identity. The first rationale for linguistic uniformity requires dissemination of a common language, but not inherently at the expense of other languages. The second sees linguistic diversity as inimical to the unity nation-states and therefore requires diversity to be replaced by uniformity. Although this ideology has not always framed the articulation of policies and some nation-states such as India or South Africa have adopted multilingual policies, the ideas of effective communication and national identity have frequently guided policy decision-making in most contexts either overtly or covertly.

Overt articulations of the desire for linguistic unity can be seen in cases such as the language policy of Franco's Spain or in the treatment of Kurdish in contemporary Turkey. In Spain, languages other than Spanish (e.g., Basque, Catalan and Galician) were proscribed in public domains and restricted even in private contexts (Vallverdú, 1991). The government removed the languages from the cultural sphere, closed newspapers, forbade their use for naming children, for commerce and even in telephone conversations (Ben-Ami, 1991). While proscriptions of everyday language use are more difficult to police, the use of the spoken languages was stigmatized, and constructed as unpatriotic: *Si eres español, habla el idioma del Imperio* ["If you're Spanish: speak the language of the empire"]. In Turkey, language policy related to Kurdish has been articulated in the context of antiterrorism laws passed in 1991. Under this legislation, disseminating separatist propaganda was deemed to be a terrorist act, and promoting the use and learning of Kurdish was constructed as a form of separatist action because it undermined the linguistic unity of the Turkish state and thus its political unity (Liddicoat, 2018b). The nature of such separatist propaganda was very loosely interpreted; for example, Muller (1996) reports that including the mention of Kurds or Kurdish in a newspaper report could be understood as separatist propaganda. Such responses can be seen as extreme formulation of the idea that a single language is required to create and express a national identity. Proscriptive approaches to the use of languages can also be seen at a micro-level in the practices schools have adopted in relation to students' language use that involve sanctioning the use of the languages as deviant behaviors. For example, in 19th and early twentieth century France, students using a minority language were exposed to ostracism and ridicule through the use of the *symbole*—a token that students were forced to wear if they were heard using their home language in school. The particular power of the *symbole* was that students needed to police their peers' language practices as they were freed from the *symbole* only

by passing it onto another (Guillore, 2018). In a more contemporary context, Spernes (2012) reported that in some Kenyan schools, “students may be physically punished” (p. 190) if they use their home language in school, except in their mother tongue education classes.

In reality most LPP work aiming toward linguistic uniformity is less strongly linked to the proscription of languages and articulates the desire for uniformity through statements that support the sole use of another language, such as in statements in constitutions that a language is the official language of the nation-state. Such formulations usually do not explicitly mention other languages, but nonetheless serve to constrain possibilities for their teaching and use. For example, the designation of French as the language of the republic in the French constitution prevented France’s ratification of the European Charter for Regional or Minority Languages because the sorts of state support for minority languages envisaged in the Charter were seen to be inconsistent with the idea of French as the sole language of the nation (Roger and de Bres, 2017). Such restrictions are not simply legalist interpretations of official decisions about language but rather are direct consequences of intervention in linguistic ecologies; any attempt to consolidate or increase the place of one language in a linguistic ecology inevitably has consequences for the other languages of the ecology. Claims of centrality for one language in a society simultaneously relegate other languages to more peripheral positions.

It is not only overt policies stated in texts that relegate languages to peripheral positions. Such marginalization can happen even in contexts where there is no overt policy promoting a single language but where discourses about languages and language practices naturalize monolingualism. Such discourses and practices often center around the idea of the utility of speaking a single national language and thus reflect the pragmatic rationale for a monolingual practice discussed above. For example, in Australia, where there is no declared official language, English has been constructed as a pre-requisite for participation in the economic, social and political life of the country. This ideological position has a certain rationality; in a highly monolingual public context, access to the sole legitimate language of public has significant advantages (Bourdieu, 1982, 2001). However, the discourse of the utility of English is accompanied by a monolingual habitus (Gogolin, 1994) that sees monolingualism as the normal and default linguistic state of persons. Within a monolingual habitus, multilingualism is seen as a deviation from the norm and often as a threat to obtaining the language capabilities required to participate in that norm (Liddicoat, 2017, 2018a). Thus, for example, the dismantling of bilingual education for Aboriginal students in the Northern Territory was undertaken in order to improve students’ levels of English. Aboriginal students who speak English as an additional language are regularly found to have lower levels of English than the white, English-speaking peers (Devlin, 2017) and this situation was problematized as creating lesser economic and social opportunities and as creating less access to English-dominated, White Australian society. Studying in a child’s indigenous first language was constructed as time not spent learning English and therefore a constraint on the learning of the dominant language of the society and a direct cause of lower English language levels. In such ideological positions, multilingualism comes to be constructed as a polar opposite to normalized and desirable monolingualism and as a problem for participating predominantly monolingual societies. Multilingual education can thus be terminated as a way of “improving” educational provision.

This conflictual construction of multilingualism can be seen in many other contexts, such as in the practices of naming groups who do not speak the dominant language of the society in which they live, study or work. For example, the term Limited English Proficiency (LEP) student constructs such students in terms of a deficit in a single language in their repertoire and claims that this limitation is more relevant for understanding the student than their abilities in other languages. Similarly, terms that seem to indicate a recognition of linguistic diversity, such as Culturally and Linguistically Diverse (CALD) or multilingual speakers of English, often reflect a discourse that normalizes monolingualism in a particular language. CALD people are not positioned simply as people who participate in linguistic and cultural diversity but as being external to a linguistically and culturally “non-diverse” norm. Similarly, multilingual speakers of English are not any person who speaks a language (or languages) in addition to English but rather as those who do not speak English well because they speak other languages. Thus, a person who is a first language speaker of English but also speaks other languages as additional languages does not qualify as a multilingual speaker of English. The following quote is representative of the ways the term is used, even in discourses that are sympathetic to those denoted by the term:

What these standards do ensure is that learners of English and multilingual speakers of English are discriminated against in their access to education in countries like Australia and Britain since they need to meet particular English language requirements not required of the monolingual speakers of English in these countries

McAlinden and Zagoria (2013, p. 32).

According to this quote, multilingual speakers of English are subjected to discriminatory practices in education because they are required to meet conditions that are not set for monolingual speakers. In reality, the distinction is not between monolingual and multilingual speakers of English, but between those who speak English as legitimate speakers and those who do not. The terminology here adopts a discourse that normalizes monolingualism and problematizes multilingualism and precludes the possibility that a first language speaker of English can also be multilingual. In this way, a statement that has as its purpose the critique of native-speaker privilege in education settings has transported into the discussion as a heteroglossic act a set of ideological assumptions that normalize a monolingual habitus.

Hierarchies of value

In any language ecology, not all languages are equal and language policy both reflects and constructs hierarchies of value (Liddicoat and Curnow, 2014) in which the various languages present in the ecology are positioned relative to others according to certain valued dimensions. These hierarchies are pervasive in the language practices of a society and have a particularly significant impact on language policies in educational contexts. These hierarchies of value connect with Bourdieu's (2001) idea of the linguistic marketplace. In the linguistic marketplace, linguistic capital, which Bourdieu sees as a form of cultural capital, is exchanged, and different languages and language varieties have different symbolic values based on the connotation the linguistic variety holds. Thus, a language that is connoted as useful, prestigious, etc. holds more value in the marketplace than one that is not. In the linguistic marketplace standard language varieties and dominant languages enjoy greater value than non-standard varieties and non-dominant languages. The highest value in the market is associated with legitimate language, which Bourdieu defines as the language spoken by elites—cultivated circles, or those in power, and serves as a tool not just for communication but also for social distinction. The hierarchies that result from the differential valuing of languages and language varieties provide a context in which decisions about language are made and thus serve to structure interventions into language ecologies.

Such hierarchies of value can be seen at play wherever language policies select languages for specific focus. For example, in Australia in the 1990s, the Australian Language and Literacy Policy (ALLP) proposed a policy of funding eight "priority languages" to be chosen by each State and Territory from a list of fourteen languages: Aboriginal languages (counting as one language), Arabic, Chinese, French, German, Indonesian, Italian, Japanese, Korean, Modern Greek, Russian, Spanish, Thai and Vietnamese (Liddicoat, 2013). This policy thereby constructed an administrative and financial hierarchy for languages that had both symbolic and practical dimensions. Symbolically, the policy designated some languages as important by listing them as priority languages and others as less important by passing them over in silence. The fact that not all of these symbolically supported languages would receive funding created a further dimension to the hierarchy in which some languages received both practical and symbolic support, others received symbolic support only while the majority of languages, including most languages of immigrant communities, received neither. The particular formulation of priority languages used here is the result of a political process that revealed contestation about the ways languages were given value in the prestigious context of a policy document. The discussion paper on which the policy was developed proposed reducing the number of designated languages in the previous policy by omitting Modern Greek. This omission was largely associated with a change in the criteria for determining the value of languages within education. In earlier policy, a number of criteria had been used including both cultural and economic criteria. The ALLP, however, framed language learning as being entirely for economic goals and the value of languages was understood in terms of their potential contribution to economic outcomes. Modern Greek, which as a strong community language had a high cultural value, was not identified as a language which had economic significance for Australia and so was not considered to be a suitable target for educational funding (see, for example, Stanley et al., 1990). The omission led to significant political reaction from the Greek community, who sought to maintain the symbolic recognition of their language (Sussex, 1991). The resulting fourteen language list can be seen as a compromise in that it preserved aspects of the symbolic valuing of some languages, but by requiring a selection from the list for funding purposes reinforced more utilitarian valuing based on economics.

The Australian example shows that hierarchies of value can create controversies in language policy where competing criteria for determining value exist. Such hierarchies can also be reflected in other types of tensions and conflicts. This can be seen in the case of medium of instruction policy in Hong Kong, which is controversial and contested area of language policy. Hong Kong is officially trilingual in Cantonese, English and Putonghua and biliterate in English and Chinese, although the majority of the population speaks Cantonese as a first language (Poon, 2019). In Hong Kong's LPP, Chinese-medium instruction as the default educational setting, and Chinese can be used through the whole of education from primary to tertiary level, although most universities in Hong Kong are English-medium institutions. The policy is thus to give students access to education in their home language; however, community pressure for English-medium instruction has been significant and continues to increase (Poon, 2013, 2019). One of the main pressures against Chinese-medium instruction is a perception that English is a more valuable language than (Cantonese) Chinese and that Chinese-medium instruction is a less effective way of developing English (Evans, 2011). In discourses about language and education, Chinese and English are thus in conflict and use of Chinese is seen as the use of a less valuable language that is taking time away from a more useful language. Similar controversies can often be seen in contexts where educational policies seek to introduce local, non-official languages into education. Usually such policies are introduced to address students' difficulties in learning in a language that they do not know and so have sound educational objectives, but may be resisted by local communities who see the use of children's first languages in education as a threat to their acquisition of more powerful and thus more valuable languages. Lopez (2008) reports that in Latin America indigenous communities may resist bilingual education models because they believe that such education will continue to marginalize indigenous children by limiting their acquisition of Spanish. They also believed that, because their languages and cultures have not been valued in mainstream society, educational programs that use them are aimed at limiting indigenous people's access to Spanish and thus work to ensure that indigenous people have less access to contexts in which they could challenge majority group power. Thus, in this case, the pre-existing hierarchies of value in relation to Spanish and indigenous languages constructed interpretative frameworks that saw education as cementing such hierarchies rather than challenging them. Thus, the ideologies of value that attach to languages together with an idea that languages are in competition in education can mean that the space available to less "valued" languages may be highly contested (Liddicoat and Curnow, 2014).

Hierarchies of value structure the sorts of relationships that exist between linguistic and cultural groups and how these are represented in education. In my own work (Liddicoat, 2013) I have examined the ways that LPP constructs interculturality a number of contexts and has shown that interculturality becomes a relationship of inequality in policies relating to the education of immigrant and indigenous minorities. In such policy contexts, interculturality is presented as a linguistic and cultural adjustment of the minority to the majority and policies rarely make mention of the need for reciprocal adjustments of majorities to minorities. Interculturality is thus presented to immigrant and indigenous groups as the cost of maintaining their linguistic and cultural diversity and as a requirement for participating in the dominant society. In these contexts, it is the differential valuing of the language and culture of the mainstream that normalizes a monodirectional linguistic and cultural adjustment; those whose languages and cultures represent less cultural capital in the linguistic marketplace need to develop their cultural capital by learning to engage with dominant groups, while learning to engage with non-dominant groups is not represented as adding to the cultural capital of dominant groups.

Promoting multilingualism

LPP in education is not always about limiting or constraining linguistic diversity, it can also seek to expand linguistic repertoires of speakers of a dominant language by adding additional languages. In cases where multilingualism is promoted, language policies construct the types of multilingualism that are valued in the society and in these constructions hierarchies of value shape educational provision. Language education programs typically present a small selection of languages that are made available for students to learn and this selection indicates something about which languages are valued in society and what criteria are used for giving value to languages. In some societies, language policies indicate that language study is required by learners but do not specify which languages will be studied, often delegating the decision about language study to the schools themselves. In such contexts, the choice of languages made is informed by the value that these languages have in the linguistic marketplace. For example, in most European Union countries there is a policy that students will study two languages at school in addition to the official language of their country. One language (Modern Language 1) is usually begun during primary education and the other (Modern Language 2) in secondary education. Most policies do not state which languages will be taught as Modern Language 1 or 2, but invariably the choice is made to teach English as Modern Language 1 and the languages of a limited number of other languages as Modern Language 2. The decision to teach English reflects ideologies of English as the language of international communication, of economic advantage and of cultural prestige (Climent-Ferrando, 2016; Krzyżanowski and Wodak, 2011). The languages taught as Modern Language 2 similarly reflect the values of the linguistic marketplace with the economically, politically and culturally dominant languages of Europe (e.g., French or German) or of elsewhere in the world (e.g., Chinese) being chosen. Other languages are placed in more peripheral positions and many languages are either not widely taught or are taught only in the border areas with countries that speak them. Immigrant languages are not widely taught unless they are also languages of economic or political importance, although they may be made available to immigrants. The policy of requiring two languages is itself an attempt to offset the dominance of English in the linguistic marketplace as one factor that lies behind its introduction is the fear that English poses a threat to the languages and cultures of EU member states. For example, the UE Commission's *Promoting language learning and linguistic diversity: An Action Plan 2004–2006* (Commission of the European Communities, 2003) states “learning one lingua franca alone is not enough” (p. 4) and articulates the potential threat of English to other European languages:

In non-anglophone countries recent trends to provide teaching in English may have unforeseen consequences on the vitality of the national language. University language policies should therefore include explicit actions to promote the national or regional language (p. 8)

While this policy seems to reflect a valuing of more generalized multilingualism, in that it supports learning of a wider range of languages, it remains nonetheless a restrictive policy in that it favors the languages of EU member states and attempts to promote a cultural rather than economic criterion for valuing languages, but one that is understood from within the logic of European integration. Climent-Ferrando (2016), however, argues that the political rhetoric that supports multilingualism and linguistic diversity is not born out in practice and that decision-making remains embedded in the economic rationale of the linguistic marketplace. The ways that various languages are valued in Australia's ALLP discussed above is a further example of the value systems that underlie policies for the teaching and learning of additional languages.

Policies may also favor particular pathways to multilingualism over others. In particular, policies may favor what Skutnabb-Kangas (1981) has called “elite” multilingualism—the acquisitions of additional languages other than the dominant language of the society through schooling—over “folk” multilingualism—the acquisition of a language through primary socialization into a non-dominant language (see also De Costa, 2019). This bias in how multilingualism is valued can be seen for example in an Australian language policy proposal on Asian languages (COAG, 1994), which has a lengthy discussion of the need to develop high level abilities in a number of Asian languages but does not mention the existence of large heritage-language communities of some of the those languages or the possibility of developing educational programs to develop their languages further (Liddicoat, 2013). Thus, the policy is entirely focused on new learning of these languages by non-speakers. Clyne (2008) identifies a similar orientation to elite multilingualism where language curricula are designed entirely for new speakers of the language and heritage language speakers are excluded from programs because they have an “unfair advantage” in the form of previous knowledge of

the language. Scotland's 1 + 2 language education policy would seem to be a counter-example to such thinking in that it aims for every child to have mastered the basics of two languages in addition to their first language by the time they leave primary school and thus implies a strong level of support for multilingualism. However, while the policy seems to support folk multilingualism, in practice there are few opportunities available for speakers of most non-dominant languages to receive educational support for their first languages (Hancock and Hancock, 2018). There thus seems to be a disconnection between policy rhetoric and practice in which valued forms of multilingualism are more likely to receive support that is more than symbolic.

Valuing and validating multilingualism—future directions for policy and research

Language policies, whether overt or covert, macro-level, meso-level or micro-level, involve decisions about how languages will be used and perceived within complex language ecologies. Research in LPP has tended to focus on policies and their impacts on societies, but much of the scholarship until recently has tended to give attention to the languages of focus in policies rather than examining them within an ecological setting to examine how policy interventions shape and are shaped by the language ecologies into which they enter. These language ecologies are permeated by texts, discourses and practices that construct frameworks for understanding languages, their use, and their value in society. One consequence of this is that multilingualism itself is often not highly valued or that many forms of multilingualism are not valued and existing ideological positions remain entrenched. This system of valuing underlies a range of policy contexts from those that are overtly antagonistic to multilingualism to those that are overtly supportive of it. LPP research has begun to investigate language policies as ideological positioning and so is beginning to understand the consequentiality of policies for languages that are included in or excluded from various policy settings and to understand how there can be both planned and unplanned consequences of language-related decision-making. While there has been substantial critique of the consequences of language policies and the ideologies that underlie them, there has been less work that has focused on addressing these consequences and in understanding what a contextualized, ecologically oriented policy response to multilingualism might be. LPP studies have only begun to look at the complexities of language policies and the multiple levels at which they exist in society, there is a distinct need for more studies that examine how policies play out across levels and how the texts, discourses and/or practices that exist at one level interact with, shape and transform those that exist at other levels.

References

- Alexander, N., 1992. Language planning from below. In: Herbert, R.K. (Ed.), *Language and Society in Africa*. Witwatersrand University Press, Johannesburg, pp. 56–68.
- Baldauf Jr., R.B., 2006. Rearticulating the case for micro language planning in a language ecology context. *Curr. Issues Lang. Plann.* 7 (2&3), 147–170. <https://doi.org/10.2167/cilp092.0>.
- Ben-Ami, S., 1991. Basque nationalism between archaism and modernity. *J. Contemp. Hist.* 26 (3/4), 493–521. <https://doi.org/10.1177/002200949102600308>.
- Bonacina-Pugh, F., 2012. Researching “practiced language policies”: insights from conversation analysis. *Lang. Pol.* 11 (3), 213–234. <https://doi.org/10.1007/s10993-012-9243-x>.
- Bonacina-Pugh, F., 2020. Legitimizing multilingual practices in the classroom: the role of the “practiced language policy”. *Int. J. Biling. Educ. BiLang.* 23 (4), 434–448. <https://doi.org/10.1080/13670050.2017.1372359>.
- Bourdieu, P., 1982. *Ce que parler veut dire: L'économie des échanges linguistiques*. Broché, Paris.
- Bourdieu, P., 2001. *Langage et pouvoir symbolique*. Arthème-Fayard, Paris.
- Clement-Ferrando, V., 2016. Linguistic neoliberalism in the European Union. Politics and policies of the EU's approach to multilingualism. *J. Lang. Law* 66, 1–14. <https://doi.org/10.2436/rld.66.2016.2843>.
- Clyne, M., 2008. The monolingual mindset as an impediment to the development of plurilingual potential in Australia. *Socioling. Stud.* 2 (3), 347–366.
- COAG, 1994. *Asian Languages and Australia's Economic Future*. Australian Government Publishing Service, Canberra.
- Commission of the European Communities, 2003. *Promoting Language Learning and Linguistic Diversity: An Action Plan 2004–2006*. Retrieved from <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2003:0449:FIN:EN:PDF>.
- De Costa, P., 2019. Elite multilingualism, affect and neoliberalism. *J. Multiling. Multicult. Dev.* 40 (5), 453–460. <https://doi.org/10.1080/01434632.2018.1543698>.
- Devlin, B., 2017. Sources of evidence on student achievement in Northern Territory bilingual education programs. In: Devlin, B.C., Disbray, S., Devlin, N.R.F. (Eds.), *History of Bilingual Education in the Northern Territory: People, Programs and Policies*. Springer Singapore, Singapore, pp. 185–202.
- Ducard, D., 2016. La division sociale de la langue: La fiction d'une langue populaire. *Revista Virtual de Estudos da Linguagem - ReVEL* 24 (3). <https://doi.org/10.17851/2237-2083.26.3.789-814>.
- Evans, S., 2011. Historical and comparative perspectives on the medium of instruction in Hong Kong. *Lang. Pol.* 10 (1), 19–36. <https://doi.org/10.1007/s10993-011-9193-8>.
- Foucault, M., 1975. *Surveiller et punir: Naissance de la prison* [Discipline and Punish: The Birth of the Prison]. Gallimard, Paris.
- Gee, J.P., 1994. New alignments and old literacies: critical literacy, post-modernism and fast capitalism. In: O'Connor, P. (Ed.), *Thinking Work*. Albsac, Sydney, pp. 82–104.
- Geeraerts, D., 2003. Cultural models of linguistic standardisation. In: René, D., Roslyn, F., Pütz, M. (Eds.), *Cognitive Models in Language and Thought: Ideology, Metaphors and Meanings*. Mouton de Gruyter, Berlin, pp. 25–68.
- Gogolin, I., 1994. *Der monolinguale Habitus der multilingualen Schule*. Waxman, Munste, New York.
- Guillourel, É., 2018. Le Breton. Itinéraire d'une langue meurtrie. In: Cornette, J. (Ed.), *La Bretagne: Une aventure mondiale*. Tallandier, Paris, pp. 279–290.
- Hancock, A., Hancock, J., 2018. *Extending the 1+2 Language Strategy: Complementary Schools and Their Role in Heritage Language Learning in Scotland*. University of Edinburgh, Edinburgh.
- Kaplan, R.B., Baldauf Jr., R.B., 1997. *Language Planning: From Practice to Theory*. Multilingual Matters, Clevedon, UK.
- Krzyżanowski, M., Wodak, R., 2011. Political strategies and language policies: the European Union Lisbon strategy and its implications for the EU's language and multilingualism policy. *Lang. Pol.* 10 (2), 115–136. <https://doi.org/10.1007/s10993-011-9196-5>.
- Liddicoat, A.J., Cumow, T.J., 2014. Students' home languages and the struggle for space in the curriculum. *Int. J. Multiling.* 11 (3), 273–288. <https://doi.org/10.1080/14790718.2014.921175>.
- Liddicoat, A.J., Taylor-Leach, K., 2020. Agency in language planning and policy. *Curr. Issues Lang. Plann.* 1–18. <https://doi.org/10.1080/14664208.2020.1791533>.

- Liddicoat, A.J., 2013. *Language-in-Education Policies: The Discursive Construction of Intercultural Relations*. Multilingual matters, Bristol.
- Liddicoat, A.J., 2017. Indigenous and immigrant languages in Australia. In: Shah, S., Seals, C. (Eds.), *Heritage Language Policies Around the World*. Routledge, New York, pp. 237–253.
- Liddicoat, A.J., 2018a. Ideologies of value and the place of Australian indigenous languages in education. In: Negga, D., Chan, D., Szirmai, M. (Eds.), *Language Policy, Ideology and Educational Practices in a Globalized World*. Editions des archives contemporaines, Paris, pp. 59–70.
- Liddicoat, A.J., 2018b. National security as a motivation in language-in-education policy. In: Chua, C.K.S. (Ed.), *Un(intended) Language Planning in a Globalising World: Multiple Levels of Players at Work*. de Gruyter, Berlin, pp. 113–128.
- Liddicoat, A.J., 2020a. Ideology in language policy and educational practice: an afterword. In: Al-Issa, A.S.M., Mirhosseini, S.-A. (Eds.), *World English Language Education Today: Ideologies, Policies and Practices*. Routledge, London, New York, pp. 135–145.
- Liddicoat, A.J., 2020b. Language policy and planning for language maintenance: the macro and meso levels. In: Schalley, A., Eisenchlas, S. (Eds.), *Handbook of Home Language Maintenance and Development*. De Gruyter, Berlin, pp. 337–356.
- Lo Bianco, J., 2005. Including discourse in language planning theory. In: Bruthiaux, P., Atkinson, D., Eggington, W.G., Grabe, W., Ramanathan, V. (Eds.), *Directions in Applied Linguistics: Essays in Honor of Robert B. Kaplan*. Multilingual Matters, Clevedon, pp. 255–264.
- Lopez, L.E., 2008. Top-down and bottom-up: counterpoised visions of bilingual intercultural education in Latin America. In: Hornberger, N. (Ed.), *Can Schools Save Indigenous Languages? Policy and Practice on Four Continents*. Palgrave MacMillan, New York, pp. 42–65.
- McAlinden, M., Zagoria, I., 2013. Which English? *Forum*, Winter, 31–22.
- Mühlhäusler, P., 2000. Language planning and language ecology. *Curr. Issues Lang. Plann.* 1 (3), 306–367. <https://doi.org/10.1080/14664200008668011>.
- Muller, M., 1996. Nationalism and the rule of law in Turkey: the elimination of Kurdish representation during the 1990s. In: Olson, R.W. (Ed.), *The Kurdish Nationalist Movement in the 1990s: Its Impact on Turkey and the Middle East*. University of Kentucky Press, Lexington, KY, pp. 173–199.
- Pennycook, A., 1994. *The Cultural Politics of English as an International Language*. Longman, London.
- Poon, A.Y.K., 2013. Will the new fine-tuning medium-of-instruction policy alleviate the threats of dominance of English-medium instruction in Hong Kong? *Curr. Issues Lang. Plann.* 14 (1), 34–51. <https://doi.org/10.1080/14664208.2013.791223>.
- Poon, A.Y.K., 2019. Language education policy in Hong Kong. In: Kirkpatrick, A., Liddicoat, A.J. (Eds.), *Routledge International Handbook of Language Education Policy in Asia*. Routledge, New York, pp. 60–75.
- Ricento, T., 2000. Historical and theoretical perspectives in language policy and planning. *J. Sociolinguistics* 4 (2), 196–213. <https://doi.org/10.1111/1467-9481.00111>.
- Roger, G., de Bres, J., 2017. Langues de France et Charte européenne des langues régionales ou minoritaires. *Socioling. Stud.* 11 (1), 131–152. <https://doi.org/10.1558/sols.31098>.
- Ruiz, R., 1984. Orientations in language planning. *NABE J.* 8 (2), 15–34. <https://doi.org/10.1080/08855072.1984.10668464>.
- Shohamy, E., 2006. *Language Policy: Hidden Agendas and New Approaches*. Routledge, London, NY.
- Skutnabb-Kangas, T., 1981. *Bilingualism or Not? The Education of Minorities*. Multilingual Matters, Clevedon, UK.
- Spernes, K., 2012. "I use my mother tongue at home and with friends—not in school!" Multilingualism and identity in rural Kenya. *Lang. Cult. Curric.* 25 (2), 189–203. <https://doi.org/10.1080/07908318.2012.683531>.
- Stanley, J., Ingram, D., Chittick, G., 1990. *The Relationship Between International Trade and Linguistic Competence*. Australian Government Publishing Service, Canberra.
- Sussex, R., 1991. The green paper on language and literacy: an overview and an assessment. In: Liddicoat, A.J. (Ed.), *Language Planning and Language Policy in Australia*, vol. 8. Applied linguistics association of Australia, Clayton, VIC.
- Taylor-Leach, K., Benson, C., 2017. Language planning and development aid: the (in)visibility of language in development aid discourse. *Curr. Issues Lang. Plann.* 18 (4), 339–355. <https://doi.org/10.1080/14664208.2017.1360690>.
- Vallverdú, F., 1991. Los estudios sociolingüísticos en España, especialmente la Cataluña. In: Klee, C., Ramos García, L. (Eds.), *Sociolinguistics of the Spanish-Speaking World*. Bilingual Press, Tempe, AZ.

Transliteracy for a cosmopolitan future

Xiaoye You^a and Michelle Zaleski^b, ^a Pennsylvania State University, University Park, PA, United States; and ^b Benedictine College, Atchison, KS, United States

© 2023 Elsevier Ltd. All rights reserved.

Definition of transliteracy	130
Transliteracy teacher toolkit	131
Transliteracy as a translingual, transcultural, and cosmopolitan practice	132
The ethics of transliteracy	133
Future directions	134
References	135

Globalization has led to widely felt tensions in language education these days. Globalization has widened the gap between the haves and the have-nots, a situation which causes people to raise questions about the unbridled reign of free trade, immigration, and blurred national borders. As a result, European and North American countries saw the rise of populist and isolationist nationalism over the last decade, a movement in which supporters argued for the strengthening of borders and anti-immigration policies. Now spreading into other parts of the world, this movement has led to racial tensions and occasionally racial cleansing, like what has happened to the Rohingya people in Myanmar in recent years. Nationalism, being increasingly asserted in many countries, comes to perpetuate language and cultural boundaries—boundaries that go directly against fluid language practice and flexible identity formation on the ground (You, 2010, 2018). The assertion of national identity in many countries requires that we reconsider how to define and teach literacy in an age of globalization.

In this chapter, we would like to extend the scholarship on globalization and literacy education by highlighting work done at the intersection of literacy, culture, and disposition under such terms as “translanguaging,” “translingualism,” and “transliteracy.” We will first review literacy-related concepts developed in response to globalization and then consider the notion of “transliteracy” in the context of language education. We will then review published studies that have drawn on transliteracy as well as translanguaging and generalize practical guidelines for teacher education and classroom instruction. While transliteracy has gained solid ground in literacy education, it has also received criticisms. Next, we will introduce these criticisms and how they can refocus the application of transliteracy in education. Finally, we will discuss future directions for the exploration of transliteracy, including historical applications, applications outside the classroom, and intersections with other fields, such as new materialism, comparative/cultural rhetoric, disability studies, and technology.

Definition of transliteracy

Over the last three decades, scholars have proposed different conceptualizations of literacy amid the rise of new media and globalization. For example, the pedagogy of multiliteracies advocated by the [New London Group \(1996\)](#) has deeply influenced literacy and language teaching in the West. The group argues that traditional literacy pedagogy is inadequate in the post-Fordism era in which communication channels and media have multiplied and cultural and linguistic diversity has increased. A pedagogy of multiliteracies assumes that “language and other modes of meaning are dynamic representational resources, constantly being remade by their users as they work to achieve their various cultural purposes” (Cope and Kalantzis, 1999, p. 5). Working in the fields of new media, arts, and creative writing, [Thomas et al. \(2007\)](#) proposed the term “transliteracy” as both a practice and a heuristic to understand what being literate means in the new media age. They define transliteracy as “the ability to read, write and interact across a range of platforms, tools and media from signing and orality through handwriting, print, TV, radio and films, to digital social networks” (Section “What is transliteracy?”, para 1). By emphasizing the interaction among text, visual, and digital literacies, the concept decenters text-based writing and reading. What the concepts of “multiliteracies” and “transliteracy” share is that globalization and the rise of new media have brought to light the fact that we communicate by interacting with multiple modes and media on a daily basis, not only now but also in the past.

These new concepts of literacy have inspired numerous explorations that intend to capitalize on the affordances of new media for literacy education. For example, [Gogan and Marcus \(2013\)](#) incorporated transliteracy into their school’s public library summer reading program. When reading a book, students used a tool called “VoiceThread” to take part in conversations, formulate ideas, and share higher-order thinking through a variety of media channels such as video, audio, text, images, and music. They also communicated with the book’s author through blogs and websites. Similarly, [Aillerie \(2019\)](#) described an experiment conducted at a French high school to observe and test students’ transliterate skills. Placed in small groups, students were required to use all kinds of media and tools to collect data for their projects. The focus of this experiment was on “the possible diversity of media

and tools employed, on the ways of and reasons for switching from one to another, on how these different media and tools are distributed within contexts, according to the academic requirements and tasks individually and collectively performed by the students" (p. 740).

While inspiring pedagogical explorations on the affordances of new media, these new concepts of literacy also have encouraged scholars and educators to examine the connections between languages using terms such as "translanguaging" and "translingualism." In applied linguistics and language education, scholars have emphasized both the fluid and artificial boundaries among languages and across modes of representation in human communication (Blackledge and Creese, 2010; Canagarajah, 2013a; García, 2009; Jordan, 2015). This added emphasis helps to define what it means to be literate not only in an age of new media but also during an era of rapid globalization and the limits of past conceptions of literacy on both. They suggest that our sense of language and dialect has been deeply influenced by nationalism, which has taught us rigid rules and forms within lingua-cultural boundaries. In practice, languages are neither discrete nor stable but rather dynamic and negotiated, and writing necessarily involves the negotiation of language differences. While this was always true, globalization makes it harder to ignore. In studies of spoken and written interactions, Canagarajah (2013a,b) used the term "translingual practice" to break down the boundaries of linguistic codes in a speaker's communicative repertoire. He emphasizes two key points in human communication: first, communication transcends individual languages because languages are always in contact and mutually influencing each other; and second, communication transcends words and involves diverse semiotic resources and ecological affordances. It is this second point that reinforces the move from multiliteracies to transliteracy that Thomas et al. (2007) made and that characterizes this chapter's emphasis on transliteracy, rather than "translingual practice" alone.

Transliteracy, then, provides a pedagogical framework that maintains an emphasis on educating students in not only multiple languages and dialects but also multiple modes and new media. Furthermore, in the field of English education, You (2016) has argued for a cosmopolitan orientation to transliteracy. This orientation emphasizes that, though sometimes defined by kindred relations, ethnicity, nation, race, or class, all people are first and foremost members of the human race and as such are always in relation with those outside their own identifications; further, they have the agency to develop and sustain new allegiances across cultures, communities, and languages. Embracing such an orientation, You (2016) suggests that the demand for cultivating global citizens requires that educators seek to cultivate in English users the ability to engage with diverse cultural discourses through reading and writing using Standard English, other Englishes, other languages, and other semiotic resources. This ability necessarily requires that English users negotiate across language differences and consciously and effectively shuttle between communities they belong to or will belong to. He calls this ability and its enactment transliteracy. Ultimately, transliteracy should foster in English users a transcultural disposition, "involving an intellectual and esthetic stance of 'openness' toward peoples, places, and experiences from different cultures, especially those from different nations" (Szczesny and Urry, 2002, p. 469). While institutionally students remain citizens of a certain nation, morally and culturally they become citizens of the world. In other words, transliteracy is a translingual and transcultural practice with a cosmopolitan imperative.

Transliteracy teacher toolkit

Putting transliteracy to use in the classroom can be implemented in two ways: (1) as a cosmopolitan disposition that encourages translingual, transcultural practice or (2) as translingual, transcultural practices that encourage a cosmopolitan disposition. For teachers and teacher educators, transliteracy as practice is both an approach and a practical toolkit; and for those looking for a way into applying transliteracy to the classroom, it is important to first acknowledge the relationship between literacy and habitus within education.

As a cosmopolitan disposition, transliteracy reorients the way in which we approach language. This allows transliteracy to be adaptable yet necessarily asks language users to recognize how it is that language shapes us as we learn (Lee and Jenks, 2016). Scholars have already uncovered the transformation of US literacy education from moral to economic imperative (Graff, 1979; Brandt, 2004). This work reveals the extent to which literacy has been asked to shape students into particular ways of thinking and acting that are ideological. Bourdieu (1979) neatly defines this dynamic through concepts like habitus and cultural capital. The typical habitus encouraged by schooling, he explains, is a social conditioning into the implicit rules of decorum that students come to internalize, maintaining the status-quo. Cultural capital is moderated by schooling whereby institutionalized norms, such as standardized literacy, both determine and become valuable for class mobility. Understanding the social and cultural role of schooling in this way, a cosmopolitan disposition encourages students and teachers to understand the connection between literacy and habitus and asks them to instead work to shape literacy in ways that allow for and forward a new status-quo built on cosmopolitan principles of inclusion, linguistic diversity, and social (r)evolution, rather than class mobility, monolingualism, and social conformity. This disposition, then, seeks to foster a position of openness toward difference (Horner et al., 2011a). It understands difference as generative and sees the encounter with the "other" as imperative to self-understanding that emerges out of exchange, rather than alongside, against, or in spite of another. This cosmopolitan disposition lays the groundwork for the translingual and transcultural negotiation that motivates transliteracy practice.

As a pedagogical approach, transliteracy can provide us with concrete ways in which to foreground language as a negotiated process. Broadly speaking, these translate into the following translingual moves: advocating for writer agency, encouraging the use of other languages, assuming heterogeneity as the classroom norm, and challenging English monolingualism; yet, at the same time, best practices in translingualism are continually evolving (Sun and Lan, 2020). As an example of this, Canagarajah (2013c)

identified negotiation strategies that can help frame transliteracy in text production. Based on ethnographic classroom research, these strategies are broken down into the following rhetorical moves shared between writer and audience: envoicing, recontextualization, interaction, and entextualization. Each demonstrates the many possibilities for putting transliteracy into practice, but together show how malleable transliteracy can be. For example, envoicing highlights the personal construction of authorial voice and the ways in which writers choose to mark their identity in writing. Envoicing understands language as performative and rhetorical, not just a medium for meaning but an outlet for playing with and changing how language makes meaning. To practice envoicing means taking the agency of the writer seriously—regardless of that writer’s education and position, and in spite of the power readers may hold over those writers, especially as teachers. Teachers work to respond to student writers as curious readers rather than as graders or institutional gatekeepers. Another negotiation strategy, interaction, further draws attention to the dialectic between reader and writer that shapes texts. Interaction as a strategy acknowledges that writing is a process of alignment where writers can invite readers into the meaning-making process and open the text up to different, imaginative interpretations. The strategy of recontextualization highlights the process by which this happens. Recontextualization incorporates feedback and drafting into the making of meaning, centering the social negotiation of voice, interaction, and textual production over time. Entextualization, then, a final strategy, accounts for the spatiotemporal dimension of the texts themselves, how it is that the long process of social negotiation becomes part of the text, embedded to facilitate that connection between reader and writer and the alignment of authorial voice with creative audience interpretation. Translingual practice broken down into these four negotiation strategies highlight that as a practice, translingualism always aims to facilitate the co-construction of meaning. Transliteracy emphasizes that such co-construction happens across different languages, modes, and cultures. It is invitational and generative, working to disperse creativity while empowering readers and writers to explore agency in conversation *with*, rather than in spite of, the other.

Transliteracy can then be practiced on the level of disposition. In designing courses that prepare teachers to use transliteracy in their classrooms, this means acknowledging—and actively overturning—the relationship between literacy and power. Teacher-trainers can empower new teachers to critically understand this complex dynamic and encourage on-going reflection so that teachers evaluate their own internalized grammars. One way to do so is by engaging in negotiation strategies that rewrite the relationship between language users; in particular, their relationship with their students. By contextualizing language as a negotiated, co-constructed process that evolves across time and travels through different spaces, teachers not only can begin to help students develop concrete transliteracy strategies like envoicing, interaction, entextualization and recontextualization that foreground negotiation but also prepare themselves for the kind of collaboration that fuels language (r)evolution.

Transliteracy as a translingual, transcultural, and cosmopolitan practice

While transliteracy began as a way to theorize meaning-making in an era of rapid globalization, as scholars have continued to research its implications for literacy education, they have underscored its value for transforming research and classroom practice. Transliteracy is both a translingual and transcultural practice. In bringing the two together, it links practice with disposition, encouraging flexibility and careful consideration to intents and purposes. Transliteracy, then, is useful not just as a theoretical framework but also as an adaptable tool for teaching and research.

In the classroom, transliteracy involves both teachers and students in actively reframing the goals of literacy education. Rather than teaching literacy as an autonomous product, literacy is taught as a negotiated process. This calls for teachers to prepare students to reflect on their language practices, decision-making processes, and changing contexts and communities. Instead of focusing on assimilating students to standardized literacy expectations, teachers support metacognition by teaching students to recognize these expectations and negotiate the ways in which they can choose to voice difference or sameness. This means discussing code-switching (Canagarajah, 2013a) or unexpected linguistic choices such as the use of the phrase “can be able to” (Lu, 1994); more broadly, teachers can foreground rhetorical decision-making in the classroom, analyzing an author’s use of multimedia and choice of medium, genre, or delivery (Bawarshi, 2016; Guerra, 2016; Lorimer-Leonard, 2014; Schreiber, 2019; Shipka, 2016). Teaching transliteracy engages students in establishing voice in their writing, allowing students to work on skills of linguistic negotiation (Alvarez et al., 2017).

From assessment to assignment and course design, transliteracy as practice redirects the typical flow of power in the classroom. First, practicing transliteracy asks teachers to critically evaluate their assessment practices. Instead of grading students on the final product alone, teachers may consider grading students based on revision and reflection (Canagarajah, 2016); or, they might consider replacing conventional writing tests with labor-based contracts (Inoue, 2017). Beyond assessment, there are multiple ways to reimagine assignments to foster transliteracy, such as restructuring peer review in Jordan’s (2012) scaffolded peer interaction, and teaching genres like the critical literacy narrative, as discussed by Young (2004); both decenter the instructor and invite students to center their own literacy practices. And finally, teachers and administrators might consider using transliteracy to redesign their writing courses by, for instance, integrating second language writers into mainstream sections, inviting new models of ownership, and incorporating cross-cultural global or community partnerships (Alvarez and Alvarez, in press; Lee and Alvarez, 2020; Lee and Jenks, 2016; Wu, 2018; You, 2016). While such practices can support multilinguals and other students whose home literacy practices are excluded from the academy, teaching transliteracy emboldens all students—including those who identify as monolingual—to ask and then assert how it is that they want to communicate with the dominant discourse (Lu and Horner, 2013).

Transliteracy as classroom practice recognizes the push and pull of academic discourse and challenges its structured dominance by involving readers and writers in its deliberate restructuring.

Transliteracy as a research practice similarly challenges the structured conditions in which literacy scholars pursue and disseminate their research. Transliteracy as a tool for research seeks to foster critical metalinguistic awareness that unsettles the ways in which scholars have internalized the nationalistic and colonial tendencies of literacy research. This first means addressing the geo-history of the native speaker and the implications of English monolingualism in the past and present (Trimbur, 2008; Horner and Trimbur, 2002). As Horner et al. (2011b) argue, “English monolingualism is manifested not simply in the language(s) of the scholarship produced but the language(s) of scholarship cited, the bibliographic resources on which composition scholars rely, the forums in which the scholarship circulates, and the arguments it makes” (p. 272). Transliteracy as a research method, then, first questions the tacit policy of English monolingualism that has structured schooling and confronts the tension monolingualism generates with the reality of a diverse, multilingual world. Transliteracy changes the time and space in the doing of research by understanding the field of literacy studies as “an on-going historical project” (Horner, 2016, p. 8).

Because these forms of dominance shape practice and power into the present day—from the exclusion of non-English texts in journals to understudied non-English writing cultures—engaging both local and global differences in researching literacy can help to confront the unexamined assumptions that allow monolingualism to persist. That is, one way to not only acknowledge but to resist the dominance of monolingualism is to expose ourselves to different ways of constructing and sharing knowledge across time and space. As a research method, transliteracy encourages scholars to work across languages, across disciplines, and across media. It is research with a cosmopolitan imperative that builds on both translingual and transcultural practices. This method can be as simple as reading research outside the field and as complex as committing to learning, using, and researching in other languages (Zaleski, 2018a, 2020). In both cases, researching this way asks scholars to learn from and engage with outside work that goes beyond mere “linguistic tourism” (Matsuda, 2014), taking care to not decontextualize or elide external contributions and paying attention to the labor of translation (Horner and Tetreault, 2016). Transliteracy as research practice has the ability to transform our research context and our respective fields by crossing borders, but such border-crossing requires a cosmopolitan orientation to research and cosmopolitan practices like decentering Western epistemologies and recentering historically marginalized ones (“epistemic de-linking”) and situating knowledge-making in in-between spaces (“border thinking”) (Mignolo, 2007). It calls for an awareness and sensitivity to location and context (Alvarez, 2018; Lee and Alvarez, 2020). Transliteracy as research practice asks scholars to commit to decolonial methods that interrogate interpersonal, institutional, and historical arrangements of power that structure research and our relationships with subjects and audiences, calling us to ultimately interrogate our own positionality (Cushman, 2016). And, it is important to note that such a shift facilitates new ways of training and mentoring scholars (Kilfoil, 2015; Zaleski and You, 2022).

Transliteracy as a practice changes how we teach and research. It engages academics in the conscious making of literacy by drawing attention to the way in which we choose to make meaning. In other words, transliteracy foregrounds language as choice, a choice to reproduce more of the same or to remake literacy through its diversity. In changing literacy praxis, scholars develop new transcultural attitudes and cosmopolitan dispositions that acknowledge the labor of translation and decentralize power by recontextualizing our place in the world. Therefore, transliteracy as practice depends on a dialectic between literacy praxis and cosmopolitan dispositions.

The ethics of transliteracy

The transliteracy perspective has received criticisms from applied linguists and literacy scholars. For instance, applied linguist Kubota (2016) has criticized the proponents of the perspective for having unwittingly aligned themselves with neoliberal multiculturalism “that celebrates individual cosmopolitanism and plurilingualism for socioeconomic mobility” (p. 475). She is concerned that the perspective has not seriously considered how racial and other power relations might affect the ways people use, learn, and teach language. She is worried that the perspective does not improve the life of those who are linguistically, racially, and economically marginalized in globalization. Responding to the neoliberalism charge, Canagarajah (2015) suggested that while neoliberals use translingualism as a product and as a commodified branding mechanism, proponents of the translingual perspective view it as a practice that is politically savvy and socially conscious, and sensitive to power, social norms, and audience. In this respect, teaching transliteracy means not just teaching transgressive language practices but also pursuing a social justice agenda that resists the neocolonial and nationalistic tendencies of literacy education and instead forwards a critical, cosmopolitan ethics.

To resist neoliberal and neocolonial tendencies, scholars of color have underscored the importance of centering the struggles and literacy practices of the underprivileged in research and pedagogy. For instance, Gilyard (2016) was concerned that proponents of translingualism have flattened language differences, blurring the boundaries between differences in communication in general and those encountered in the education of African American students in particular. This tendency ends up erasing the Black community’s struggles for survival, perpetuating language as abstraction and alienating scholars of color. Therefore, Gilyard suggests that proponents of translingualism could study students’ efforts, struggles, and triumphs in transforming education systems and fighting for social justice through translanguaging. Gilyard was recently joined by other Black scholars who seek to expose and redress racism in language education and research. These scholars demand Black linguistic justice, asking teachers to stop using academic language and standard English as the accepted communicative norm and to instead teach Black students about anti-Black linguistic racism and white linguistic supremacy (CCCC, 2020).

Meanwhile, transliteracy was also censured by those who held a monolingual and pragmatic perspective to language education. These individuals charged that proponents of the translingual perspective had ignored students' desires for Standard English and downplayed the importance of linguistic norms (Atkinson and Tardy, 2018). Responding to these charges, You (2020) argued for a yin-yang dynamic between national/monolingual and transnational/translingual perspectives to language and literacy pedagogy. Students' desires and expectations are informed by both national and transnational forces, and transliteracy has focused on those informed by transnational forces. Linguistic norms and variations are in a dialectical relation, he suggests, and not easily separable in practice. When we use language, we begin with norms developed in a certain community and use them creatively or transgressively in our interaction with context and audience, hence creating variations.

Although some scholars are critical of transliteracy, others hold hope for its potential power for leveling hierarchies of thoughts and languages fostered by modernity's colonial matrix of power and for generating pluriversal understandings, values, and practices. For example, Cushman (2016) suggests translingualism be defined as meaning making processes that involve students and scholars in translanguaging, translating (trans-actions), and dwelling in borders. She argues that translingual approaches must work at the paradigmatic level to reveal and potentially transform colonial matrices of power and at the pedagogical level to allow students to express their languages and categories of understanding. These epistemological and pedagogical moves might together lead to epistemic delinking and border thinking, or transcultural dispositions, crucial for decolonization of thought and language.

These criticisms of transliteracy or translingualism and their ensuing conversations suggest that while focusing on cultivating responsible global citizens, we need to continue pursuing social justice in teaching and studying transliteracy. As scholars have suggested, while notions such as hybridity, fluidity, and multiplicity highlighted by the translingual perspective are potentially liberating, they can obscure the struggles and inequalities brought about by globalization. We need to stay alert to the way in which neoliberalism has widened economic and educational gaps and encouraged a kind of multiculturalism that evades racism and other social injustices. Thus, in teaching and studying transliteracy, as Kubota (2016) has warned, "more explicit attention should be paid to issues of asymmetrical relations of power and inequalities that privilege or stigmatize individuals and groups due to their plurilingualism, cosmopolitanism, and hybridity on the one hand, or their monolingualism and monoculturalism on the other" (p. 490).

Future directions

Thus far, transliteracy has inspired numerous studies on students' literacy practice as well as pedagogical applications (for pedagogical applications see Canagarajah, 2013c; De Costa et al., 2017; Frost et al., 2020; Horner and Tetreault, 2017; Lee and Jenks, 2016; You, 2016). Looking forward, we already see rising on the horizon its applications in historical studies, applications outside of the classroom, and its intersections with other fields, such as new materialism, comparative/cultural rhetorics, disability studies, and technology. In this final section, we will give an overview of these new applications with the intention to identify areas where transliteracy is well poised to make a contribution.

Transliteracy highlights the fact that humans communicate by interacting with multiple modes and media, not only in the present but also in the past. This insight has inspired historical explorations of language and language education, explorations that have elucidated how individuals, groups, and institutions exercised agency and political power through translingual practice (Arnold, 2014, 2018; Kimball, 2021; You, 2023; Zaleski, 2018b). For instance, using a material translingual ecologies perspective, Zaleski (2018b) studied Jesuit rhetoric as it traveled between Europe and India by looking at how the Jesuit educational work was enmeshed in both new and old social relationships and constructed through rhetorical practices that were textual, material, and embodied. Kimball (2021) examined discourses of language, religion, and education in early Philadelphia from a translingual perspective. The city was essentially a translingual space, where communities such as Germans, Quakers, and African Americans developed metalinguistic conceptions of what language was and how it mattered in their relations and the nation formation.

As early as 2013 (Canagarajah, 2013a), research on literacy began connecting language practice in the classroom to extra-curricular ecologies and environments, situating literacy in the marketplace and across the world. Scholars have attempted to better understand the direct relationship between the classroom and the world (Nordquist, 2017), and more recently to build on "the mobility turn" in composition studies to trace the ways in which texts create meaning as they travel (Faist, 2013). Incorporating work in sociolinguistics from Blommaert (2013) and Kell (2015), literacy scholars like Viera (2019) and Lorimer-Leonard (2017) are charting exactly how literacy travels; that is, how transliteracy is negotiated across both space and time.

This work has been influenced not only by the mobility turn but also new materialism and studies of digital networking, including research from rhetoric and composition scholars like Prior (2010), Spinuzzi (2015), and Jordan (2015). Such work further defines the way in which transliteracy foregrounds diverse semiotic resources as meaning-making tools and helps us to understand exactly the ways in which transliteracy engages complex ecologies to restructure meaning. At the same time, such work questions fixed understandings of human agency. Scholars such as Ruiz and Arellano (2019) have complicated these understandings of agency and power by bringing together transliteracy, new materialism, and decolonial methods. Similarly Zaleski (2020) and Zaleski and You (2022) have brought together transliteracy and research on comparative/cultural rhetorics to decenter and resituate literacy practice elsewhere. These studies have benefited from comparative rhetoric's emphasis on rhetoric as culturally

constructed and its commitment to interrogating research positionality in the study of other cultures. Scholars have also interrogated assumptions about language and voice by reflecting on their own identity as multilingual scholars of writing. Scholars like Lee and Canagarajah (2019) and Canagarajah and Matsumoto (2017) have used their position as teachers, students, and mentors, as well as their relationship to each other, to underscore the complex voices engaged by transliteracy and how transliteracy requires a commitment to on-going identity work and relationship-building.

As transliteracy highlights human communication taking place through multiple modes and media, technological affordances are a natural area to extend literacy scholarship. Scholars have examined students' transliteracy practices and identity construction in social media (DePew, 2011; Lee, 2011; Liu, 2018; Schreiber, 2015; You, 2008, 2011) and developed pedagogical activities to engage students in exploring multilingual and multimodal resources for promoting accessibility, social justice, and cosmopolitanism (Fraiberg et al., 2017; Gonzales and Butler, 2020; Wu, 2018; Wu and Li, 2019). For instance, focusing on identity construction, Schreiber (2015) examined a Serbian college student's literacy practice on Facebook. She found that as a hip-hop artist, this student constructed a fluid transnational identity by capitalizing on the affordances of multimodality; he used Serbian or English and posted images and links to video clips to mark his membership in local and global hip-hop communities. Drawing on disability studies, Gonzales and Butler (2020) designed multimodal assignments in which students were asked to examine the issue of accessibility through multilingual/multimodal community engagement, reflecting on communication differences, and analyzing communication in practice. Literacy educators have increasingly used new communication tools for intercultural communication activities that cultivate cosmopolitan disposition. Wu (2018) reviewed some of these efforts over the last two decades and concluded that these activities are inherently transnational because when participants are engaged in virtual communication, their identities projected, relations forged, and languages practiced all emerge from the context, defying any static or discrete nation-based grouping.

These emerging trends in transliteracy scholarship underscore the importance of pursuing transliteracy alongside social justice, keeping in mind power and privilege as researchers incorporate new ways of studying literacy practice. At the same time, such scholarship demonstrates transliteracy's contribution to developing a more critical habitus toward language-making. It reveals the ways in which transliteracy's emphasis on language as negotiation, on meaning-making through diverse semiotic resources, and on negotiation as imperative to the evolution of language across time and space have all pushed literacy studies to reconsider its former assumptions about language, and therein, what scholars should research as well as what teachers can teach. By stretching the very basic definition of literacy in this way, transliteracy has posed a question in need of answering to all scholars and teachers: how can we language toward a more dynamic and diverse future? By redefining language and literacy, in particular, the stability provided by definitions, transliteracy has pushed research forward, challenging all researchers to question their assumptions about language and pursue their own ways of transforming literacy education so that it is not only more adaptable, equitable, and just but also a more realistic reflection of the way that languages happen and necessarily evolve.

References

- Allier, K., 2019. Teenagers' information practices in academic settings: what relevance for a transliteracy-based approach? *J. Librarian. Inf. Sci.* 51 (3), 737–745.
- Alvarez, S.P., Alvarez, S., (in press). Transnational and trans-modal community alliances: Southern Latinx students' digital and culturally sustaining practices. In: del Hierro, V., Baca, I., Gonzales, L. (Eds.), *Community Action for Social Change: A Digital Archive*. Parlor Press, Anderson, SC.
- Alvarez, S.P., Canagarajah, S., Lee, E., Lee, J.W., Rabbi, S., 2017. Translingual practice, ethnic identities, and voice in writing. In: Horner, B., Tetreault, L. (Eds.), *Crossing Divides: Exploring Translingual Writing Pedagogies and Programs*. University Press of Colorado, Boulder, CO, pp. 31–48.
- Alvarez, S.P., 2018. Multilingual writers in college contexts. *J. Adolesc. Adult Lit.* 62 (3), 342–345.
- Arnold, L.R., 2014. "The worst part of the dead past": language attitudes, policies, and pedagogies at Syrian Protestant College, 1866–1902. *Coll. Compos. Commun.* 66 (2), 276–330.
- Arnold, L.R., 2018. "Today the need arises" *اليوم قد مسّت الحاجة* : Arabic student writing at the turn of the 20th century. In: You, X. (Ed.), *Transnational Writing Education: Theory, History, and Practice*. Routledge, New York, pp. 95–112.
- Atkinson, D., Tardy, C., 2018. SLW at the crossroads: finding a way in the field. *J. Sec. Lang. Writ.* 42, 86–93.
- Bawarshi, A., 2016. Beyond the genre fixation: a translingual perspective on genre. *Coll. Engl.* 78 (3), 243–249.
- Blackledge, A., Creese, A., 2010. *Multilingualism: A Critical Perspective*. Continuum, New York.
- Blommaert, J., 2013. *Ethnography, Superdiversity and Linguistic Landscapes: Chronicles of Complexity*. Multilingual Matters, Bristol.
- Bourdieu, P., 1979. Symbolic power. *Critiq. Anthropol.* 4 (13–14), 77–85.
- Brandt, D., 2004. Drafting US literacy. *Coll. Engl.* 66 (5), 485–502.
- Canagarajah, S., Matsumoto, Y., 2017. Negotiating voice in translingual literacies: from literacy regimes to contact zones. *J. Multiling. Multicult. Dev.* 38 (5), 390–406.
- Canagarajah, S., 2013a. *Translingual Practices: Global Englishes and Cosmopolitan Relations*. Routledge, New York.
- Canagarajah, S. (Ed.), 2013b. *Literacy as Translingual Practice: Between Communities and Classrooms*. Routledge, New York.
- Canagarajah, S., 2013c. Negotiating translingual literacy: an enactment. *Res. Teach. Engl.* 48 (1), 40–67.
- Canagarajah, S., 2015. Suresh Canagarajah on translingualism: a four-part interview, part I. In: *Transnational Writing*. <https://transnationalwriting.wordpress.com/2015/08/04/suresh-canagarajah-on-translingualism-a-four-part-interview-part-i/>.
- Canagarajah, S., 2016. Translingual writing and teacher development in composition. *Coll. Engl.* 78 (3), 265–273.
- CCCC, 2020. This Ain't Another Statement! This is a DEMAND for Black Linguistic Justice!. <https://cccc.ncte.org/cccc/demand-for-black-linguistic-justice>.
- Cope, B., Kalantzis, M. (Eds.), 1999. *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, New York.
- Cushman, E., 2016. Translingual and decolonial approaches to meaning making. *Coll. Engl.* 78 (3), 234–242.
- De Costa, P.I., Singh, J.G., Milu, E., Wang, X., Fraiberg, S., Canagarajah, S., 2017. Pedagogizing translingual practice: prospects and possibilities. *Res. Teach. Engl.* 51 (4), 464–472.
- DePew, K.E., 2011. Social media at academia's periphery: studying multilingual developmental writers' Facebook composing strategies. *Read. Matrix* 11 (1), 54–75.
- Faist, T., 2013. The mobility turn: a new paradigm for the social sciences? *Ethn. Racial Stud.* 36 (11), 1637–1646.

- Fraiberg, S., Wang, X., You, X., 2017. *Inventing the World-Grant University: Chinese International Students' Mobilities, Literacies, and Identities*. Utah State University Press, Logan, UT.
- Frost, A., Kiernan, J., Malley, S.B. (Eds.), 2020. *Translingual Dispositions: Globalized Approaches to the Teaching of Writing*. The WAC Clearinghouse; Louisville, CO. University Press of Colorado, Fort Collins, CO.
- García, O., 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell, Oxford.
- Gilyard, K., 2016. The rhetoric of translanguaging. *Coll. Engl.* 78 (3), 284–289.
- Gogan, B., Marcus, M., 2013. Lost in transliteracy: how to expand student learning across a variety of platforms. *Knowl. Quest.* 41 (5), 40–45.
- Gonzales, L., Butler, J., 2020. Working toward social justice through multilingualism, multimodality, and accessibility in writing classrooms. *Compos. Forum* 44. <https://compositionforum.com/issue/44/multilingualism.php>.
- Graff, H.J., 1979. *The Literacy Myth: Literacy and Social Structure in the Nineteenth-Century City*. Academic Press, New York.
- Guerra, J.C., 2016. Cultivating a rhetorical sensibility in the translingual writing classroom. *Coll. Engl.* 78 (3), 228–233.
- Horner, B., Tetreault, L., 2016. Translation as (global) writing. *Compos. Stud.* 44 (1), 13–30.
- Horner, B., Tetreault, L., 2017. *Crossing Divides: Exploring Translingual Writing Pedagogies and Programs*. Utah State University Press, Logan, UT.
- Horner, B., Trimbur, J., 2002. English only and US college composition. *Coll. Compos. Commun.* 53 (4), 594–630.
- Horner, B., 2016. *Rewriting Composition: Terms of Exchange*. Southern Illinois University Press, Carbondale, IL.
- Horner, B., Lu, M.-Z., Royster, J.J., Trimbur, J., 2011a. Language difference in writing: toward a translingual approach. *Coll. Engl.* 73 (3), 303–321.
- Horner, B., NeCamp, S., Donahue, C., 2011b. Toward a multilingual composition scholarship: from English only to a translingual norm. *Coll. Compos. Commun.* 63 (2), 269–300.
- Inoue, A., 2017. Writing assessment as the conditions for translingual approaches: an argument for fairer assessments. In: Horner, B., Tetreault, L. (Eds.), *Crossing Divides: Exploring Translingual Writing Pedagogies and Programs*. University Press of Colorado, Boulder, CO, pp. 119–134.
- Jordan, J., 2012. *Redesigning Composition for Multilingual Realities*. NCTE, Urbana, IL.
- Jordan, J., 2015. Material translingual ecologies. *Coll. Engl.* 77 (4), 364–382.
- Kell, C., 2015. Ariadne's thread: literacy, scale, and meaning making across time and space. In: Stroud, C., Prinsloo, M. (Eds.), *Language, Literacy and Diversity: Moving Words*. Routledge, New York, pp. 72–91.
- Kilfoil, C.B., 2015. Beyond the “foreign” language requirement: from a monolingual to a translingual ideology in rhetoric and composition graduate education. *Rhetor. Rev.* 34 (4), 426–444.
- Kimball, E., 2021. *Translingual Inheritance: Language Diversity in Early National*. University of Pittsburgh Press, Philadelphia. Pittsburgh.
- Kubota, R., 2016. The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: complexities and implications for applied linguistics. *Appl. Linguist.* 37 (4), 474–494.
- Lee, E., Alvarez, S.P., 2020. World Englishes, translanguaging, and racialization in the US college composition classroom. *World Englishes* 39, 263–274.
- Lee, E., Canagarajah, S., 2019. The connection between transcultural dispositions and translanguaging practices in academic writing. *J. Multicult. Discourses* 14 (1), 14–28.
- Lee, J.W., Jenks, C., 2016. Doing translingual dispositions. *Coll. Compos. Commun.* 68 (2), 317–344.
- Lee, C.K.M., 2011. Micro-blogging and status updates on Facebook: texts and practices. In: Thurlow, C., Mroczek, K. (Eds.), *Digital Discourse: Language in the New Media*. Oxford University Press, New York, pp. 110–130.
- Liu, J.Y., 2018. The affordances of Facebook for teaching ESL writing. In: You, X. (Ed.), *Transnational Writing Education: Theory, History, and Practice*. Routledge, New York, pp. 203–221.
- Lorimer Leonard, R., 2014. Multilingual writing as rhetorical attunement. *Coll. Engl.* 76 (3), 227–247.
- Lorimer Leonard, R., 2017. *Writing on the Move: Migrant Women and the Value of Literacy*. University of Pittsburgh Press, Pittsburgh.
- Lu, M.Z., Horner, B., 2013. Translingual literacy, language difference, and matters of agency. *Coll. Engl.* 75 (6), 582–607.
- Lu, M.Z., 1994. Professing multiculturalism: the politics of style in the contact zone. *Coll. Compos. Commun.* 45 (4), 442–458.
- Matsuda, P.K., 2014. The lure of translingual writing. *PMLA* 129 (3), 478–483.
- Mignolo, W.D., 2007. Delinking. *Cult. Stud.* 21 (2), 449–514.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–93.
- Nordquist, B., 2017. *Literacy and Mobility: Complexity, Uncertainty, and Agency at the Nexus of High School and College*. Routledge, New York.
- Prior, P., 2010. Remaking IO: semiotic remediation in the design process. In: Prior, P., Hengst, J. (Eds.), *Exploring Semiotic Remediation as Discourse Practice*. Palgrave MacMillan, Hampshire, pp. 206–234.
- Ruiz, I.D., Arellano, S.C., 2019. *La Cultura nos cura: reclaiming decolonial epistemologies through medicinal history and quilting as method*. In: García, R., Baca, D. (Eds.), *Rhetorics Elsewhere and Otherwise: Contested Modernities, Decolonial Visions*. NCTE, Urbana, IL, pp. 141–168.
- Schreiber, B., 2015. “I am what I am”: multilingual identity and digital translanguaging. *Lang. Learn. Technol.* 19 (3), 69–87.
- Schreiber, B., 2019. “More like you”: disrupting native speakerism through a multimodal online intercultural exchange. *Tesol Q.* 53 (4), 1115–1138.
- Shipka, J., 2016. Transmodality in/and processes of making: changing dispositions and practice. *Coll. Engl.* 78 (3), 250–257.
- Spinuzzi, C., 2015. *All Edge: Inside the New Workplace Networks*. University of Chicago Press, Chicago.
- Sun, Y., Lan, G., 2020. Enactment of a translingual approach to writing. *Tesol Q.* 54 (2), 1–29.
- Szerszynski, B., Urry, J., 2002. Cultures of cosmopolitanism. *Sociol. Rev.* 50 (4), 461–481.
- Thomas, S., Joseph, C., Laccetti, J., Mason, B., Mills, S., Perril, S., Pullinger, K., 2007. Transliteracy: crossing divides. *First Monday* 12 (12). <https://doi.org/10.5210/fm.v12i12.2060>.
- Trimbur, J., 2008. The Dartmouth conference and the geohistory of the native speaker. *Coll. Engl.* 71 (2), 142–169.
- Viera, K., 2019. What happens when texts fly. *Coll. Compos. Commun.* 82 (1), 77–95.
- Wu, Z., Li, X., 2019. Developing cosmopolitan communicative competence through online transnational encounters. *ESL Can. J.* 36 (1), 110–131.
- Wu, Z., 2018. Technology-mediated transnational writing education: an overview of research and practice. In: You, X. (Ed.), *Transnational Writing Education: Theory, History, and Practice*. Routledge, New York, pp. 170–186.
- You, X., 2008. Rhetorical strategies, electronic media, and China English. *World Englishes* 27 (2), 233–249.
- You, X., 2010. *Writing in the Devil's Tongue: A History of English Composition in China*. Southern Illinois University Press, Carbondale, IL.
- You, X., 2011. Chinese white-collar workers and multilingual creativity in the diaspora. *World Englishes* 30 (3), 409–427.
- You, X., 2016. *Cosmopolitan English and Transliteracy*. Southern Illinois University Press, Carbondale, IL.
- You, X. (Ed.), 2018. *Transnational Writing Education: Theory, History, and Practice*. Routledge, New York.
- You, X., 2020. The yin-yang of writing education in globalization. In: Silva, T., Wang, Z. (Eds.), *Reconciling Translingualism and Second Language Writing*. Routledge, New York, pp. 150–159.
- You, X., 2023. *Genre Networks and Empire: Rhetoric in Early Imperial China*. Southern Illinois University Press, Carbondale, IL.
- Young, M., 2004. *Minor Re/Visions: Asian American Literacy Narratives as a Rhetoric of Citizenship*. Southern Illinois University Press, Carbondale, IL.

- Zaleski, M., You, X., 2022. Comparative rhetoric and the translingual future of mentorship. In: Ayash, N.B., Kilfoil, C. (Eds.), *Translingual and Transnational Graduate Education in Rhetoric and Composition*. Utah State University Press, Logan, UT.
- Zaleski, M., 2018a. Beyond words: missionary grammars and the construction of language in Tamil country. In: Cañizares-Esguerra, J., Hsia, R.P., Maryks, R.A. (Eds.), *Encounters Between Jesuits and Protestants in Asia and the Americas*. Brill, Leiden, pp. 159–176.
- Zaleski, M., 2018b. *Salir consigo*: Jesuit Rhetorical Education in India, 1542–1623. Ph.D. Diss. Pennsylvania State University.
- Zaleski, M., 2020. Recontextualizing comparative rhetoric. In: Lloyd, K. (Ed.), *The Routledge Handbook of Comparative World Rhetorics*. Routledge, New York, pp. 268–276.

Rethinking multiliteracies: toward a radical vision for multilingualism

Brittany L. Frieson^a, Luz A. Murillo^b, Vivian E. Presiado^c, and Jim Sosnowski^d, ^a Teacher Education and Administration, University of North Texas, Denton, TX, United States; ^b Curriculum and Instruction, Texas State University, San Marcos, TX, United States; ^c School of Teaching and Learning, Illinois State University, Normal, IL, United States; and ^d Curriculum and Instruction, University of Illinois Urbana-Champaign, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Casting our radical vision: a positionality statement	138
Misconceptions and radical hopes in the literacy education of indigenous multilingual children	139
The idea of Latin America	139
Radical hope	139
What US teachers learned about indigenous children and their schools in Guatemala	139
Developing autonomy in an indigenous bilingual school in Colombia	140
Language Partners: disrupting deficit discourses about multilingual futures	140
Origins of Language Partners	141
Push and pull of neoliberal forces	141
Danger of the language and literacy myth	142
False narratives of the Mexican motherland: Mexican multilingual children (re)writing their worlds	142
The concern of the “global outcasts”	142
Self-authoring literacy	143
An awakening: reflections on Mexican children and their schooling experience in Mexico	143
Transnational literacies: re-constructing ideas of borders and knowledges	143
The cruel fallout of racism: confronting teacher candidates’ misconceptions and (re)structuring beliefs about Black language	144
Addressing anti-black linguistic racism	144
Disrupting anti-black linguistic racism in preservice teacher education	145
Normalizing Black language in preservice teacher education	145
Conclusion: looking toward radical hope	146
References	146

Society is permeated with ideological notions about language and power that refer to how language should be used, and the “appropriateness” of language in given contexts, and the validation of “standardized languages” that reconstitute misconceptions about multiliteracies and multilingualism. Given that the majority of youth enrolled in public schools in the United States are from historically marginalized populations (NCES, 2022), now more than ever, educational institutions have the opportunity to disrupt historical discourses that have privileged dominant narratives of monolingualism and further build upon the linguistic strengths of multilingual youth. Such justice-oriented work that includes dismantling oppressive systems that fuels misconceptions about multilingualism in education necessitates a critical lens. A critical approach to multilingualism serves as a means for the denunciation of US societal norms that regard multilingualism as a problem and not a right or “a fact of life” for multilinguals (Wei, 2008, p. 3).

Researching multilingualism from a critical perspective locates an individual’s language practices within the social, cultural, and political contexts in which social practices are employed (Blackledge and Creese, 2010; Heller, 2007). In our research we adopt a critical perspective in viewing language as a dynamic practice, as opposed to being heavily regulated by the ideological boundedness of language systems. Such a perspective provides space for disrupting deficit language ideologies (Flores and Rosa, 2015) about multilingual practices, and radical hope for the futures of multilingual children. Hooks (2003) asserts that educators who are committed to rejecting oppressive systems and making classroom a place of mutual liberation for teachers and students, must be grounded in radical hope and freedom. Following hooks’ call for a pedagogy of hope, this chapter will share a summary of our recent research that has challenged misconceptions about multiliteracies and multilingualism situated in various social contexts.

Casting our radical vision: a positionality statement

Our positionalities as a coalition of multilingual scholars who identify as scholars of Color (first, second, and third author) or co-conspirators (fourth author), have fundamentally shaped the main premise of our critical perspectives on multiliteracies and multilingualism. Having all participated in these misconceptions at various points in our personal journeys where we unlearned harmful rhetoric about our own multilingualism, we consider our scholarship, teaching, and service as an opportunity to encourage radical hope in the education of multilingual children. Luz, a multilingual associate professor and Latina scholar with Colombian roots, often challenges students’ deficit perspectives about the Global South. Jim, a current secondary educator and multilingual white

scholar, resists his own presumptions about the hegemony of English that he often utilizes to challenge student discourse about the “legitimacy of academic English”. Vivian, an assistant professor and Chicana scholar, views her scholarship of rejecting students’ negative perceptions of multiliteracies as a reflection of her childhood experiences. She also had teachers whose misconceptions affected her own perspectives of multilingualism; therefore, she considers it of great importance to disrupt similar dominant narratives among her students, pre-service teachers. Brittany, an assistant professor and Black woman scholar, considers her research as an opportunity to interrupt societal discourse that often casts a negative shadow on Black language. Together, our goal is to share how we can rethink multiliteracies and move toward a radical vision for multilingualism grounded in radical hope (Lear, 2006), that disrupts misconceptions in the education of multilingual children. We are determined to radically reimagine a humanizing future for multilingual children by sharing a collection of experiences entrenched in our personal and professional lives.

Misconceptions and radical hopes in the literacy education of indigenous multilingual children

The education of multilingual indigenous children is fraught with misconceptions as it is almost always organized by others (e.g., adults, non-indigenous, monolingual speakers of socially dominant languages). In the case of Latin America, these misconceptions include essentialist notion that Indigenous children are uneducable or incapable of educational success; monolingual ideologies that would require children to abandon their home language(s) in order to “gain” Spanish, English or Portuguese; and neoliberal education policies that work to erase indigenous identities and practices as antithetical to participation in globalized forms of capital.

Because these misconceptions are also grounded in particular geographies, they reflect place-specific understandings and ways of organizing and practicing education. The examples I (Murillo) present, from indigenous schools in Guatemala and Colombia, are located in the geopolitical region of Latin America, as well as within two distinct nation states, each with its own traditions of mis-educating indigenous peoples. Studies of the language and literacy education of Indigenous children who are also (im)migrants reveal the dominant interests that structure schools in the US, Canada, and other “host” nations. And sometimes, as in the application of educational research from dominant neoliberal traditions to curriculum, instruction and materials in Latin American schools (Walsh, 2017), indigenous children are subject to both “national” and “international” deficit perspectives simultaneously.

The idea of Latin America

Eduardo Galeano wrote that “the ultimate goal of the Latin American colonial economy from its inception ... was to serve the development of capitalism somewhere else” (Galeano, 1997, p. 45). The development of capitalism in Europe and the industrialized West was accomplished through the displacement of indigenous populations, appropriation of their land and labor, and a 5-centuries-long project of replacement of Indigenous cultural practices and languages with Spanish and other colonial languages. Language homogenization and the imposition of alphabetic literacy were central to the colonial project of civilizing indigenous “barbarians”, and these forces have continued largely through the institution of schooling. In response, indigenous resistance movements have emerged to enact “concrete processes, struggles, and practices of resurgent and insurgent action and thought, including in the spheres of knowledge, territory-land, state, and life itself” (Mignolo and Walsh, 2018, p. 9).

Radical hope

To illuminate examples from my research in Guatemala and Colombia, I use Jonathan Lear’s theory of radical hope. Lear (2006, p. 104) writes that what makes hope radical

is that it is directed toward a future goodness that transcends the current ability to understand what it is. Radical hope anticipates a good for which those who have the hope as yet lack the appropriate concepts with which to understand it. What would it be for such hope to be justified.

Thus, researching and teaching with radical hope means envisioning and working toward a more just society, one that we have yet to experience. Lear suggests that radical hope is “intimately bound to the question of how to live” (p. 105) in times and situations that seem otherwise hopeless. Given the abject history of schooling for indigenous children around the world, radical hope seems an apt metaphor for those who wish to imagine the outlines of a better future for speakers of indigenous languages.

What US teachers learned about indigenous children and their schools in Guatemala

State-sponsored violence against Indigenous population has expelled many Guatemalans. One migration stream, from Q’anjob’al communities in the northwestern province of Huehuetenango, has settled in agricultural communities in central Illinois, where they are looked down upon by white and African American speakers of English, as well as by Mexican immigrants who identify as white or mestizo. Fieldwork in the local schools and early childhood centers, along with my university courses with in-service literacy teachers, revealed that teachers felt unprepared to teach these newcomer students. I (Murillo) found common misconceptions about

indigenous languages. Many teachers, including ESL and bilingual teachers, expressed the view that the Guatemalan children spoke Spanish “badly” or “incorrectly”, and quite a few believed that Q’anjob’al is an inferior dialect of Spanish. Although they had little direct experience with or training about Latin America, teachers questioned the value of the education children had experienced before coming to the US. Several teachers wondered if there were even schools in Guatemala.

To inform the work of Illinois teachers, I (Murillo) visited families and schools in sending communities in Huehuetenango. My goal was to document the educational experiences of indigenous children. I wanted to provide information about Q’anjob’al speakers’ bilingualism and mestizo identity and the safety, economic, and political motivations behind their decisions to migrate. Above all, I sought to humanize the idea of Q’anjob’al migrant children and families. To give teachers a sense of what schools and classrooms in Sta. Eulalia looked and sounded like, I took photographs and interviewed children, teachers, and school administrators, and observed lessons in elementary and secondary classrooms. I also collected examples of lesson plans, bilingual textbooks, and teacher-made materials. I hoped to use this knowledge to challenge misconceptions teachers might hold about indigenous people, about education in Guatemala and, by extension, the educability of indigenous children.

In Illinois, I shared my findings through teacher workshops and in my university literacy courses. I was struck by the positive reactions to images and descriptions of Guatemalan classrooms, sparsely furnished but beautifully decorated with teacher-made posters and examples of student work in Q’anjob’al and Spanish. Teachers were surprised to learn that the curriculum included textbooks in two languages and commented on the esthetic differences between the Guatemalan materials and the worksheets that dominated the instruction of immigrant children in Illinois. Another surprise was that elementary school teachers in Guatemala must study a pre-university year and then three years full-time university study to earn certification. This is, as several teachers observed after reading excerpts from interviews with bilingual educators in Huehuetenango, more preparation in the field of education than many Illinois teachers receive. We addressed misconceptions about Q’anjob’al and bilingualism in Indigenous languages by inviting a university linguist who specializes in Guatemalan languages. In these ways, I sought to help Illinois educators recognize the multilingualism of their Q’anjob’al families and the value of immigrant children’s and parents’ previous schooling.

Developing autonomy in an indigenous bilingual school in Colombia

A second example comes from my (Murillo) research with an Arhuaco community in the Sierra Nevada mountains of northern coastal Colombia. The Arhuacos are known as one of Colombia’s most resilient Indigenous groups, based on their defense of territory, language, and cultural practices. Ethnolinguistic maintenance among the Arhuaco is rooted in strong feelings of pride in being Indian and in speaking Ika, their ancestral language. Since the early 1970s the Arhuacos have sustained a strong movement that grounds their rights to *educación propia* (our own education) in a broad defense of their lands and culture. After boycotting and eventually shutting down government schools in 1982 (Murillo, 2009), the Arhuaco school in the community of Simurwa re-opened their school with an indigenous, bilingual faculty and an “indigenized” curriculum that sought to balance local knowledge and the Colombian national curriculum. In 2001, dissatisfied with the continued encroachment of non-Indigenous values among their youth, the Arhuacos closed the school to rethink their curriculum. My work at the Simunurwa school consisted of documenting Arhuaco efforts to decolonize schooling and to answer with them the question that had led to the school closures: What should be done with an institution that has been harmful to the community’s linguistic sovereignty and identity but which holds forms of knowledge, including print literacy, mathematics, and the Spanish language, that the Arhuaco regard as necessary for protecting their lands, identity, economic and political viability, and even physical security? (Murillo, 2009, 2021). In 2002, the community formally established the autonomy of the school under control of the Arhuaco People and reopened the school with a restructured curriculum emphasizing Ika/Spanish biliteracy and locally relevant curricula.

Research with Indigenous multilingual communities and schools in Latin America suggests ways that education and educational research can be restructured to protect Indigenous communities and their languages. The defense of territory, language, and education in indigenous communities in Guatemala, Colombia, and elsewhere in Latin America has often been accompanied by linguists, anthropologists, and literacy experts. As noted by Fals-Borda, this project involves working within and “against a university colonized by Western Euro American culture and so subordinating that it impeded useful discovery of our own realities” (2009, p. 360). Multilingualism researchers must embrace new research paradigms and move away from a tradition of studying about indigenous education to new forms of research that frame indigenous education within broader struggles for territorial and educational autonomy. Adopting new research methodologies and approaches like those I have described is one way that researchers can practice radical hope.

Language Partners: disrupting deficit discourses about multilingual futures

Historically, language and literacy education has carried the weight of a great equalizer, believed to provide access to economic and social opportunities for individuals who could develop the correct practices. Globally, literacy education has been viewed as a means of reducing poverty and national “development” (Hernandez-Zamora, 2010; Spolsky, 2010). Similarly, within the US, language and literacy education has been associated with moral development, national unity, economic sustainability, and national security (Brandt, 2004). While these claims are exciting, it is important to question whether building additional language and literacy practices operates equally for all, holding the same potential benefits for all individuals.

Drawing on data from an 18-month qualitative study which focused on how language ideologies shaped classroom practices and policies in Language Partners (LP), a prison-based, peer-taught, adult language and literacy program situated in the US Midwest, I (Sosnowski) examine how discourses pertaining to the promises associated with multilingualism and multiliteracies shaped the choices and actions of the LP students and their perceptions of their future possibilities. After introducing LP, I compare how similar forces operated to influence LP students' choice to immigrate to the US and how similar discourses are again operating as these men are deported after their incarceration.

Origins of Language Partners

LP began when an incarcerated individual, who was part of a higher education in prison program, proposed that he and his peers create an English language and literacy course for fellow incarcerated men who could not speak English. He observed that individuals who could not operate in English struggled to navigate the linguistic landscape of the prison, beyond what others normally experienced. With the support of volunteers, like me, from a nearby research-focused university, they developed a program that met every Tuesday and Thursday for 3 h class sessions. As a peer taught program, a team of incarcerated individuals planned and taught each lesson. A grammar-focused textbook guided the curriculum with each class focusing on a discrete grammatical feature. Since LP began offering classes in 2010, nearly all students enrolled in LP were immigrants from Mexico, Guatemala, and Honduras and identified Spanish as their dominant language.

Push and pull of neoliberal forces

Many LP students shared similar motivations which shaped their choice to immigrate to the US. These motivations were captured in an essay written by an LP student, Felipe, in which he recounted his immigration story. He began his story by stating, "I never thought I would leave my family someday, but I did. In 2003, I left Mexico to come to the United States. I think that it was the worst decision that I had made in my whole life, but I didn't have a choice because I was poor." Felipe explained how he hoped for a "better life" and wanted to be able to provide for his family and how he believed immigrating to the US would provide those opportunities. Instead, like many other migrants to the US, Felipe found himself following the seasons, working in agriculture, contributing his labor to the US economy but barely earning enough to survive.

Felipe's story provides an example of the historical trends that have seen individuals from Central America and Mexico immigrate to the US for labor opportunities. However, the forces influencing Felipe and many other immigrants were not based on individual choices. Heyman (2018) argues that "unauthorized and guest-worker migrations from Mexico and Central America, and their employment in the United States, were reinforced by neoliberalism but were not a product of it. Rather, they existed before the rise of neoliberalism and contributed to its rise" (p. 221). While neoliberal forces contributed to the conditions which motivated Felipe and other LP students to immigrate to the US, these forces did not end with their incarceration.

Instead, similar forces, this time associated with English language and literacy education, operated as the students considered their life after incarceration when they would be deported. Contemplating life after incarceration, many students believed developing English language and literacy practices would provide access to more employment opportunities once back in Mexico. When asked to explain their motivations for joining LP, students responded with comments like, "Salir adelante con un buen trabajo para que tenga [sic] una familia unida y para darle un futuro mejor a mis hijos" (I want to move forward with a good job so that my family can be together and I can give my children a better future), "mi meta es aprender a hablar [sic] y escribir perfectamente y así tener un mejor empleo" (My goal is to learn to speak and write perfectly and like that I will have a better job), and "Tener un mejor trabajo en el futuro" (To have a better job in the future). Multiple students used the phrase "salir adelante," (move forward) a phrase Messing (2014) identifies as part of a metadiscourse often found in Latin America and connected with instances of linguistic assimilation and opportunities for improving a person's economic situations. Similarly, Farr (2006) connects this phrase to a "liberal individualist ideology" focused on economic and social gain through hard work.

LP instructors also argued that LP would contribute to better opportunities after deportation. In a video used in a social justice award submission, one LP instructor explained:

It [LP] provides opportunity for these guys. I believe all of them are going to get deported. So, now they have this newfound knowledge that they're able to use to find maybe better employment ... tourism. And I guess that's the jist of Language Partners. Hope.

In a conversation, another instructor shared that LP was so important:

because it doesn't only help the students to navigate the prison environment, but it also helped them when they were released. Eventually the men in the program would be deported and when they went back to Mexico, they could find better jobs in restaurants or hotels because they would be able to speak English.

For both students and instructors, developing additional English language and literacy practices would benefit students economically by providing better access to employment opportunities.

While processes of globalization and neoliberal forces initially influenced decisions to migrate to the US in search of better employment opportunities, in LP these men were being positioned to return to Mexico, again with the promise of enhanced employment opportunities, this time related to the English proficiency developed while incarcerated.

Danger of the language and literacy myth

Shortly before Felipe was released he shared some concerns about returning to Mexico saying, “I don’t have my certificado de primaria or secundaria. They will ask me for those when I want to get a job in Mexico.” He continued, “if you didn’t have those certificates you couldn’t get a job. You could get hired, but the employers will pay less,” adding that it did not matter if he spoke English. While Felipe, like many other LP students, initially joined LP and believed it would someday contribute to a better future for him, faced with the reality of returning to Mexico, he questioned the extent to which those language and literacy promises would materialize.

While LP was situated within a carceral context, the promises associated with multilingualism and multiliteracies permeate educational settings throughout society and educators and researchers need to critically question these discourses. As Ong (2006) argues, “Populations governed by neoliberal technologies are dependent on others who are excluded from neoliberal considerations” (p. 32). The exclusion of Felipe and others like him from flows of capital is not an accident, but an inherent part of the neoliberal system in which we operate. Instead of being a “great equalizer” it is imperative to consider how the promises associated with multilingualism and multiliteracies for racialized individuals contribute to maintaining a societal status quo in which racialized individuals contribute to the flows of capital but are never allowed to fully benefit from them.

In writing about literacy education in prison, Berry (2018) cautions that language and literacy researchers must “resist the romance of prison writing and prison literacies, perspectives that reduce prison writing to a good in and of itself” (p. 5). Instead, language and literacy researchers need to recognize and critique how language and literacy is intertwined with social and economic systems that have historically contributed to the exploitation of language minoritized individuals. As Gee (2015) reminds us, “The most striking continuity in the history of literacy is the way in which literacy has been used, in age after age, to solidify the social hierarchy, empower elites and ensure that people lower on the hierarchy accept the values, norms and beliefs of the elites, even when it is not in their self-interest or group interest to do so” (p. 84). To challenge this status quo, language and literacy researchers must resist the urge to perpetuate narratives that forward the idea that developing additional language and literacy practices inherently benefits all individuals equally.

False narratives of the Mexican motherland: Mexican multilingual children (re)writing their worlds

Research pertaining to culturally and linguistically diverse learners has indicated that US educators often wrongly assume that students and families of Mexican descent are uninterested in “official” academic and “prestigious” forms of literacy, with children often internalizing these deficit views (Adair et al., 2017). Even educators of Mexican descent can deficit ideologies of Mexican literacy (Murillo, 2016). As Hernandez-Zamora (2010) noted, “Since the 20th century, literacy policies and programs have been implemented on the basis of a growing rhetoric about ‘alarming levels of functional illiteracy’ in the population. Official figures in countries like Mexico count the number of ‘illiterate’ adults in [the] millions” (p. 21). When illiteracy rhetoric penetrates society and spreads, those deficit ideologies travel with individuals, including into US classrooms. A number of schools in the Midwest serve students and families that come from programs attempting to “eliminate illiteracies” (Meyers, 2014).

That illiteracy baggage became part of my (Presiado) upbringing as those deficit ideologies were transferred to me by parents who weren’t able to complete their elementary schooling. My research interests are grounded in my own lived experience as a daughter of low-SES Mexican immigrants and as a former bilingual/ESL teacher for 10 years in the Midwest. Although I hadn’t formally learned about theoretical frames and the exclusion of certain parts of US history until I was at the collegiate level, I quickly realized that those deficit narratives were false early on in life. Well aware of the rich literacy practices present in Mexican households, I set out to document and learn from Mexican American families’ home literacy practices in the Midwest, only to discover how mainstream cultural and linguistic practices and middle-class ideologies about education have penetrated Mexican immigrant homes. An interview with a participant revealed her belief in a dominant discourse that I have encountered in the field regarding education in Mexico, often coming from Mexican immigrants themselves: “La educación de México no es buena” (Mexican education isn’t good). Meyers (2014) captures this prominent theme best when addressing Mexican education, stating, “Emphasis continues to be placed on illiteracy and how this lack is harming the nation” (p. 31).

The concern of the “global outcasts”

Migration between Mexico and the United States is a unique case in the history of international migration with the presence of massive migratory flows and interdependence between nations (Durand, 2000). Within this movement, the interplay between literacy and immigration in US history has highlighted a shift from literacy being used to regulate immigration to a growing sense that literacy has to be protected from immigration and its various complexities (Meyers, 2014). Hernandez-Zamora (2010) posits

that Mexicans and by default individuals of Mexican descent, can be categorized as “global outcasts” instead of global citizens (p. 19). This stems from the fact that millions have never been treated as citizens in Mexico, with the marginalization and exclusion being more pronounced in the US. In the case of Mexico, the system of castes established with Spanish rule (1521–1821) was officially over after the Declaration of Independence in 1821, yet “we have maintained *de jure* and *de facto* structures that guarantee a caste-like society, spread out now at a transnational or global level by massive migration to the US” (Hernandez-Zamora, 2010, p. 21). Millions of descendants of the former colonized or enslaved people are now categorized as school failures, often omitted are the literacy trajectories and historical factors that have impacted marginalized individuals. In particular, Mexicans have encountered European literacy and education as instruments of conquest and colonization, economic deprivation, and cultural assimilation, and dislocation for over five centuries (Hernandez-Zamora, 2010).

Self-authoring literacy

A crucial element when reflecting back on my own lived experiences and those of the participants in my research studies, is the practice of self-authoring literacy (Freire, 1970; Hernandez-Zamora, 2010). Self-authoring literacy, goes beyond merely accomplishing daily tasks and is therefore a practice aimed at self-construction or re-construction of identities, a way to push back and write one’s place in the world (Freire, 1970; Dyson, 1993). As the presented examples will elucidate, Mexican children and children of Mexican descent in the US not only engage in multiple forms of literacy, but are often producers of powerful counter-narratives.

An awakening: reflections on Mexican children and their schooling experience in Mexico

A Latino National Survey conducted in 2006 revealed that while Midwestern Latinos are similar to Latinos in other regions of the US, the foreign born in the Midwest are more likely to have come from Mexico, particularly the states of Chihuahua, Jalisco, Michoacán, and Guanajuato (Martínez, 2011). As a bilingual teacher and researcher working in schools in the Midwest, I (Presiado) have encountered many families from Jalisco. One of my first experiences with a Mexican immigrant parent from Jalisco was listening to a mother apologize for her inability to help her child because of what she called her own “poor literacy skills,” this misconception one of many internalized by Mexicans.

To better understand the language and literacy practices in the schooling of Mexican children, I visited public elementary schools in Jalisco, Mexico, specifically in Ciudad Guzman and Mazamitla for four weeks, spending 2 weeks in each city. The ethnographic case study allowed for an in-depth investigation of how language and literacy were enacted both in the official school setting and the community. The experience was eye opening and extended beyond my own awakening, as I later shared the information with pre-service and practicing teachers in Illinois who were working with Mexican families from Jalisco, or who they themselves were from Jalisco.

Regarding school-based literacies, overall the findings aligned with previous descriptions on the emphasis of “form” and “meaning” to some extent in all classrooms with the presence of planas (the repetitive copying of individual letter, words, syllables) as an important part of early literacy instruction. I took pictures of the materials used to teach literacy, was given some artifacts both by teachers and students, collected lesson plans, conducted interviews with the teachers and administrators, and had conversations with children in addition to observing in the classrooms all day for four weeks. The materials collected out in the community were photographs, artifacts, documented conversations with student’s family members or individuals in the community.

In sharing my (Presiado) findings with pre-service teachers and practicing teachers in Illinois, educators were often surprised to learn of the expectations for teachers and the multiple ways students engaged with their multiliteracies. Many were amazed to learn of the yearly exam that Mexican teachers have to take to maintain their licensure and the rigorous preparation program for pre-service teachers. They were also not expecting to hear of the multimodal ways in which the children and their families were authoring their lives, especially when seeing pictures of the schools and communities. The stories shared about the students using smartphones to learn other languages, including English and Mandarin, to stay informed about world events was also unforeseen. Mexican teachers’ rapport with families was also revealing with technology being a major tool for communication. Shared information often extended beyond school related topics to include recipes, interesting articles, inspiring quotes, and funny memes to encapsulate life circumstances. Misconceptions were challenged when sharing students’ interviews with educators in Illinois of Mexican student’s life goals to be chefs, executives, travel agents, pet shop owners, writers and how it was intricately planned in unofficial classroom spaces by generating plans and samples. The level of complexity in multimodal literacy practices is also relevant for students of Mexican descent residing in Illinois.

Transnational literacies: re-constructing ideas of borders and knowledges

My research in the Midwest with transnational families has also impacted my perspectives pertaining to language and literacy practices. Transnationals are individuals who participate in flows and movement of people, ideas, capital and good between regions (Jiménez et al., 2015). Transnationals have regular ongoing contact in physical and digital realms with others across national boundaries. A family I worked with not only engaged in transnational practices including communicating with family in Mexico via digital technologies, upholding Mexican traditions, and having Mexican goods in the home but also engaged in regular and ongoing contact in digital forms with Asian culture.

When sharing my research with teachers in Illinois about how students of Mexican descent were engaged in exploring and engaging with Korean and Chinese culture there was a shift in thinking about “borders and knowledges.” Borders in the sense of how “borders” often indexed the relationship and boundaries between Mexico and the US to see how digital spaces have, in a sense, created bridges to explore other cultures. Taking it a step further, it was unexpected that multilinguals of Mexican descent would be interested in learning Asian languages and practices associated with art and music, when according to teachers they were still learning English. Yet the girls in this transnational family were well aware of the impact of multilingualism as one stated, “Puedo decir hola, hello, y *nĩ hã* y eso es poderoso” (“I can say hola, hello, and *nĩ hã* and that is powerful”).

Considered “global outcasts,” Mexican and individuals of Mexican descent have created their own identities that have created counternarratives to what it means to “be Mexican.” Researchers exploring the language and literacy practices of multilinguals would benefit from humanizing the populations they are examining by learning about their histories, present, and future hopes.

The cruel fallout of racism: confronting teacher candidates’ misconceptions and (re)structuring beliefs about Black language

Negative perceptions and ideologies about Black language are pervasive in educational spaces as ideological assumptions centered on race, language, and power are embedded within institutional structures. It is not surprising that deficit language ideologies about Black discourses saturate educational institutions, as society often circulate harmful rhetoric about Black linguistic identities through media outlets, such as films, television, social media and other mainstream accounts. Critical scholarship challenges misconceptions socially constructed about minoritized languages by arguing that language is essentially heterogenous, and false narratives that promote ideological linguistic hegemony reject the linguistic and cultural identities of Black American students, particularly in academic spaces. Therefore, deficit perspectives that permeate schools about Black language discourses contribute to Black students’ internalized falsehoods about their own linguistic practices (Richardson, 2003). Perhaps the most detrimental false narrative of them all encourages Black language speakers to develop “code-switching habits” in order to achieve academic success.

In (re)membering my (Frieson) lived schooling experiences, Black teachers that consistently embodied my linguistic practices was imperative to sustaining Black language in the classroom and affirming my cultural and linguistic identities. However, in my work as a teacher educator, I often observe the reproduction of deficit ideological assumptions about languages other than White Mainstream English (WME), through the curriculum, policies, and pedagogical practices in higher education spaces. As a Black faculty member and a Black language speaker, I often encourage teacher education students to confront their deficit perspectives about Black language by developing a critical consciousness to anti-Black linguistic racism and its impact on Black students. Particularly critical to this examination is the unlearning of negative ideologies and misconceptions that have been fostered about Black children’s “linguistic human rights” (Baugh, 2015). Language ideologies are often bound with goals of producing a single “legitimate” language, used as a gatekeeping mechanism to exclude socially minoritized languages (Blackledge and Creese, 2010), such as Black language. The ontological trope of ideologies derives from colonial racism (Fanon, 1967) that reproduces racial and linguistic hierarchies that demonstrate how Black discourses are undervalued resources in classrooms (Baker-Bell, 2020).

The following sections illuminate critical theoretical perspectives situated in my pedagogical practices central to the facilitation of restructuring and reframing of students’ epistemic assumptions about Black language. Not only do the subsequent conceptual understandings challenge societal misconceptions that reduce Black languages as informal tropes of expression solely fit for utility outside of school, but also calls on educators and scholars to cast a forward-looking vision of a humanizing education for Black students.

Addressing anti-black linguistic racism

The language, only the language ... It is the thing that Black people love so much- the saying of words, holding them on the tongue, experimenting with them, playing with them. It's a love, a passion. Its function is like a preacher's: to make you stand up out of your seat, make you lose yourself and hear yourself. The worst of all possible things that could happen would be to lose that language This is a really cruel fallout of racism.

Toni Morrison.

For decades, Black children have endured linguistic oppression and violence in schools (Boutte and Bryan, 2019). In contrast, white ideological ways of speaking and knowing are institutionalized and reconceptualized as the “standard” or ideal manner of communication in schools, and often go unmarked and unquestioned (Baker-Bell, 2019). This notion of standardization is often protected with arguments that advocate for “academic language”, or the language of power. As many critical scholars have previously argued, when white ways of speaking are constructed as the “standard” to which all other ways of speaking and being are evaluated, racial and linguistic hierarchies are formed where the linguistic practices of communities of Color are devalued and positioned as useless in schools (Alim and Smitherman, 2012; Baker-Bell, 2020; Frieson, 2021; Morales and Hartman, 2019). Toni Morrison referenced this exact notion in the 1981 interview with Thomas LeClair, when discussing how Black children’s language practices are disparaged in schools as a result of being told permanently damaging things about their language.

Although racialized populations who speak minoritized languages endure linguistic discrimination, *anti-black linguistic racism* references the specificity and the embodied experience of the violence and harm that Black language speakers encounter daily (Baker-Bell, 2019). As mentioned by Baker-Bell (2020, p. 2), “anti-black linguistic racism refers to the linguistic violence, persecution, dehumanization, and marginalization that Black Language speakers endure when using their language in schools and in everyday life.” Anti-black linguistic racism serves as a framework for understanding the multiple forms of violence that Black language speakers experience and how anti-Blackness fuels the harm that is experienced across various contexts.

Boutte and Bryan (2019) further outline linguistic oppression that Black children experience by defining linguistic violence as a type of anti-Black violence that includes Black language marginalization and policing “through privileging and promoting White mainstream English” (p. 6). Linguistic violence and oppression also refers to the raciolinguistic ideologies (Alim et al., 2016; Flores and Rosa, 2015) manifested by teachers who do not understand the richly complex linguistic discourses employed by Black language that surpass the single-axis “basics” that are promoted in schools (Wynter-Hoyte and Smith, 2020), demonstrating the significance of encouraging pre-service teachers to interrogate their misconceptions about Black children’s multilingualism.

Disrupting anti-black linguistic racism in preservice teacher education

Recent scholarship focused on policy implications has emphasized the urgency in creating counternarratives about multilingualism in the classroom (Seltzer and de los Rios, 2021). Aligning with research that disrupts monolingualism as the norm in the classrooms, I contend that language ideologies and anti-Black linguistic racism should be further challenged in higher educational contexts. In order to cultivate future educators who reject, “institutionalized anti-Black dispositions, policies, practices, and curricula” (Wynter-Hoyte, 2020, p. 408) and who deeply understand the nuanced richness that Black language speakers bring to the classroom, teacher educators in higher education should also examine their pedagogical practices to ensure that anti-Black linguistic violence and harm is not being perpetuated in teacher education. Therefore, the examples I (Frieson) present demonstrate how I as a preservice teacher educator and professor confront teacher candidates’ misconceptions and deficit perspectives about Black language in higher education spaces.

Every semester, I teach an undergraduate course that centers the multilingualism of children of Color. We have previously discussed the complex linguistic landscapes of multilinguals that often remain absent from curriculum and pedagogical practices in schools, and how preservice teacher candidates can reimagine how to foster a humanizing education that prioritizes the knowledge, histories, experiences, and discourses of BIPOC children. Providing an education that places the voices, lives, and literacies of BIPOC children at its core must be inclusive of critically examining, interrogating, and uprooting the historical and contemporary educational injustice perpetrated by white supremacy that has been imposed upon systematically marginalized populations.

In the course, I challenged students to reframe their ideological assumptions of multilingual children, particularly Black children considering the pervasiveness of anti-Blackness in educational and societal contexts. Many preservice teachers, including future ESL and ELA teachers, shared perspectives that multilingual children speak languages that have been institutionalized, such as English, Spanish, or French. When asked about students who speak Black language or other languages that have been systematically marginalized, many students echoed a negative perception about Black language, stating that it was “ghetto”, “slang”, and questioned if Black language was a language system. Baker-Bell (2020) mentions how teacher attitudes shaped by language ideologies contribute to the normalization and perpetuation of anti-Black linguistic racism by questioning the legitimacy of Black Language such as the perspectives shared by my undergraduate students. When preservice teachers engage in questioning the legitimacy of Black language, it also further reflects their deficit assumptions about its speakers (Alim and Smitherman, 2012).

Black language and literacy scholars have discussed the devastating impact that anti-Black linguistic racism has on Black students, such as the *linguistic push-pull* where Black students’ love for Black language is often contradicted by the rejection of Black language advanced by societal norms that promote whiteness (Smitherman, 2006). For example, when Black students internalize negative ideologies about their own language practices, their linguistic, racial, and intellectual identities are often impacted (Kirkland and Jackson, 2008). In a study conducted by Baker-Bell (2013) in a secondary advanced placement English class, students shared that they associated “low intellectual ability” and “ignorance” with Black language speakers while they identified WME speakers as “well-educated”. Moreover, research also demonstrates when Black children are taught to reject their linguistic practices, they are indirectly receiving the message that their identities and discourses don’t matter (Baker-Bell, 2020).

Normalizing Black language in preservice teacher education

To combat anti-Black linguistic racism, I (Frieson) interrogated many of my own pedagogical practices in the undergraduate course. One of the approaches implemented was a translanguage approach (García et al., 2017) foregrounded in centering multilingualism and the entirety of my linguistic repertoire as the norm. In an attempt to reframe my students’ ideological assumptions about Black language, I utilized Black language frequently and encouraged my students who knew Black language to do the same. In an effort to cultivate praxis that supported students’ rights to their own languages (CCCC, 1975), I also implemented linguistically flexible assignments where socially minoritized languages were invited, especially Black language and Spanglish. Students acknowledged that it took them time to adapt to critical praxis that located multilingualism as the norm, as the linguistic practices of multilinguals were typically not esteemed in their undergraduate coursework.

A shift that I noticed when I began to implement translanguaging and antiracist Black language pedagogy and practices, such as fluidly moving throughout my linguistic toolbox and prioritizing Black language, Black students began to also confidently utilize linguistic dexterity in their oral discourse and written assignments. For example, one of the major course assignments is a critical reading response journal where students synthesize readings and share thoughts. In the past students have only engaged in written responses; however, this year I invited students to engage in linguistic creativity by submitting audio files or artwork for their critical reading response journals. Only a few students took advantage of the offer, including two Black students, Ebony and Camille (both pseudonyms).

For Ebony's critical reading response journal, she created a podcast where she interviewed many people within her circle, including family, friends, and co-workers, about critical language and literacy topics ranging from language ideologies to Black language to translanguaging. Throughout the podcast, Ebony primarily used Black language to facilitate these critical conversations. When I asked her about her processes in creating and executing the podcast, Ebony mentioned how liberated she felt to finally be herself. Camille also created a podcast to offer critical responses to the readings. In a particular episode, she shared her astonishment with learning about the complexity of Black language in a class because as a secondary English major she had internalized the fallacy of the code-switching approach where Black language is solely utilized as a means of survival for the linguistic manipulation of learning WME (Baker-Bell, 2020). In addition, she also raised critical questions about never receiving formal education about her language practices until now and interrogated how anti-Black linguistic racism in school structures underscored her experiences.

In short, critical interdisciplinary research that further explores multilingualism centered on Black American children and youth stresses that the future of scholarship focused on Black language must address the complexity of language and race. As emphasized by Baker-Bell (2019, p. 3) "many classrooms continue to be informed by anti-black, deficit theories, and monolingual ideologies that view Black language as a barrier to Black students' literacy education.", therefore it is imperative that we interrogate policies and practices in higher education that contribute to anti-Black linguistic racism. As Toni Morrison once stated, it is essential that we are refraining from the participation of the cruel fallout of racism in denying Black students' rights to their own multilingualism.

Conclusion: looking toward radical hope

Our four different, but complementary narratives illustrate how we have employed critical perspectives that debunk misconceptions about multilingualism in our research and in our lives. Our narratives further demonstrate the process of how our participants and students grappled with misconceptions about indigenous languages (Murillo), language and literacy myths (Sosnowski), multiliteracies (Presiado), and how teaching them about their own language rights (Frieson) constructs a path forward toward radical hope. While confronting misconceptions about multiliteracies and multilingualism is challenging work for us all, it allows us to envision a future with a more-just and liberating society for multilinguals.

References

- Adair, J.K., Colegrove, K.S.S., McManus, M.E., 2017. How the word gap argument negatively impacts young children of Latinx immigrants' conceptualizations of learning. *Harv. Educ. Rev.* 87 (3), 309–334. <https://doi.org/10.17763/1943-5045-87.3.309>.
- Alim, H.S., Rickford, J.R., Ball, A. (Eds.), 2016. *Raciolinguistics: How Language Shapes Our Ideas About Race*. Oxford University Press, New York, NY.
- Alim, H.S., Smithernan, G., 2012. *Articulate while Black: Barack Obama, Language, and Race in the U.S.* Oxford University Press.
- Baker-Bell, A., 2013. "I never really knew the history behind African American Language": critical language pedagogy in an advanced placement English language arts class. *Equity Excell. Educ.* 46 (3), 355–370. <https://doi.org/10.1080/10665684.2013.806848>.
- Baker-Bell, A., 2019. Dismantling anti-black linguistic racism in English language arts classrooms: toward an anti-racist Black language pedagogy. *Theory Into Pract.* 59. <https://doi.org/10.1080/00405841.2019.1665415>.
- Baker-Bell, A., 2020. *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*. Routledge, New York, NY.
- Baugh, J., 2015. Use and misuse of speech diagnostics for African American students. *Int. Multidiscip. Res. J.* 9 (4), 291–307. <https://doi.org/10.1080/19313152.2015.1082416>.
- Berry, P., 2018. *Doing Time, Writing Lives: Refiguring Literacy and Higher Education in Prison*. Southern Illinois University Press. <https://doi.org/10.1080/07350198.2018.1497891>.
- Blackledge, A., Creese, A., 2010. *Multilingualism: A Critical Perspective*. Continuum International Publishing Group, New York, NY.
- Boutte, G., Bryan, N., 2019. When will Black children be well? Interrupting anti-Black violence in early childhood classrooms and schools. *Contemp. Issues Early Child.* 1–12. <https://doi.org/10.1177/1463949119890598>.
- Brandt, D., 2004. Drafting U.S. literacy. *Enycl. Educ. Reform Dissent* 66 (5), 485–502. <https://doi.org/10.4135/9781412957403.n296>.
- Committee on College Composition and Communication Language Statement, 1975. Students' right to their own language. *Coll. Engl.* 36 (6), 709–726.
- Durand, J., 2000. Tres premisas para entender y explicar la migración México-Estados Unidos. *Relaciones. Estudios de historia y sociedad XXI* (83), 1–11. <http://repositorio.cualtos.udg.mx:8080/jspui/handle/123456789/52>.
- Dyson, A.H., 1993. *Social Worlds of Children Learning to Write in an Urban Primary School*. Teachers College Press, New York, NY.
- Fals Borda, O., 2008. The application of the social sciences' contemporary issues to work on participatory action research. *Soc. Appl. Anthropol.* 67 (4), 359–361.
- Fanon, F., 1967. *Black Skin, White Masks*. Grove Press Inc, New York, NY.
- Farr, M., 2006. *Rancheros in Chicagoacán: Language and Identity in a Transnational Community*. University of Texas Press, Austin, TX. <https://doi.org/10.4324/9780203121436>.
- Flores, N., Rosa, J., 2015. Undoing appropriateness: raciolinguistic ideologies and language diversity in education. *Harv. Educ. Rev.* 85 (2), 149–171.
- Freire, P., 1970/1997. *Pedagogy of the Oppressed*, 20th anniversary ed. Continuum Publishing, New York, NY.
- Frieson, B.L., 2021. Remixin' and flowin' in centros: exploring the biliteracy practices of Black language speakers in an elementary two-way immersion bilingual program. *Race Ethn. Educ.* 1–21. <https://doi.org/10.1080/13613324.2021.1890568>.
- Galeano, E., 1997. *Open Veins of Latin America: Five Centuries of the Pillage of a Continent*. Monthly Review Press, New York, NY.

- García, O., Johnson, S.I., Seltzer, K., 2017. *The Translanguaging Classroom: Leveraging Student Bilingualism for Learning*. Caslon Publishing, Philadelphia, PA.
- Gee, J., 2015. *Social Linguistics and Literacies: Ideology in Discourses*, fifth ed. Routledge, New York, NY. <https://doi.org/10.4324/9780203944806>.
- Heller, M., 2007. *Bilingualism: A Social Approach*. Palgrave Macmillan, New York, NY.
- Hernandez-Zamora, G., 2010. Decolonizing Literacy: Mexican Lives in the Era of Global Capitalism. *Multilingual Matters*, Tonawanda, NY. <https://doi.org/10.21832/9781847692641>.
- Heyman, J., 2018. How does neoliberalism relate to unauthorized migration. In: Carrier, J. (Ed.), *Economy, Crime, and Wrong in a Neoliberal Era*. Berghahn Books, pp. 218–239. <https://doi.org/10.2307/j.ctww04jq5.12>.
- Hooks, B., 2003. *Teaching Community: A Pedagogy of Hope*. Routledge, New York, NY.
- Jiménez, R.T., Smith, P.H., Teague, B.L., 2009. Transnational and community literacies for teachers. *J. Adolesc. Adult Literacy* 53 (1), 16–26. <https://doi.org/10.1598/JAAL.53.1.2>.
- Jiménez, R.T., Eley, C., Leander, K., Smith, P.H., 2015. Transnational immigrant youth literacies: a selective review of the literature. In: *Handbook of Research on Cross-Cultural Approaches to Language and Literacy Development*. IGI Global, Hershey, PA, pp. 322–344. <https://doi.org/10.4018/978-1-4666-8668-7.ch013>.
- Kirkland, D., Jackson, A., 2008. Beyond the silence: instructional approaches and students' attitudes. In: Scott, J.C., Straker, D.Y., Katz, L. (Eds.), *Affirming Students' Right to Their Own Language: Bridging Language Policies and Pedagogical Practices*. National Council of Teachers of English, New York, NY, pp. 132–150.
- Lear, J., 2006. *Radical Hope*. Harvard University Press, Cambridge, MA.
- Martínez, R., 2011. *Latinos in the Midwest*. Michigan State University Press, East Lansing, MI.
- Messing, J., 2014. "I didn't know you knew Mexican!": shifting ideologies, identities, and ambivalence among former youth in Tlaxcala, Mexico. In: *Indigenous Youth and Multilingualism: Language Identity, Ideology, and Practice in Dynamic Cultural Worlds*. Routledge, New York, NY, pp. 111–129. <https://doi.org/10.4324/9780203121436-14>.
- Meyers, S.V., 2014. *Del Otro Lado: Literacy and Migration Across the US-Mexico Border*. SIU Press, Carbondale, IL.
- Mignolo, W.D., Walsh, C., 2018. *On Decoloniality. Concepts, Analytics, Praxis*. Duke University Press, Durham, NC.
- Morales, P.Z., Hartman, P.W., 2019. Positioning Spanish-language and African American Language (AAL) speakers as knowledgeable and valuable contributors in the language arts classroom. *Theory Into Pract.* 58 (3), 236–245. <https://doi.org/10.1080/00405841.2019.1607654>.
- Murillo, L.A., 2009. "This great emptiness we are feeling": toward a decolonization of schooling in Simunurwa, Colombia. *Anthropol. Educ. Q.* 40 (4), 421–437.
- Murillo, L.A., 2016. Aquí No Hay Pobrecitos: decolonizing bilingual teacher education in the US-Mexico borderlands. *Diaspora Indig. Minority Educ.* 11 (4), 163–176. <https://doi.org/10.1080/15595692.2016.1258694>.
- Murillo, L.A., 2021. Late capitalism and new challenges: indigenous communities taking risks in defense of vulnerable languages and territories in Guatemala and Colombia. In: Cunningham, C., Hall, C.J. (Eds.), *Vulnerabilities, Challenges and Risks in Applied Linguistics*. *Multilingual Matters*, pp. 31–47.
- National Center for Education Statistics, 2022. *Back-to-School Statistics*. <https://nces.ed.gov/fastfacts/display.asp?id=372>.
- Ong, A., 2006. Introduction: neoliberalism as exception, exception to neoliberalism. In: *Neoliberalism as Exception: Mutations in Citizenship and Sovereignty*. Duke University Press, Durham, NC, pp. 1–30. <https://doi.org/10.1215/9780822387879-001>.
- Richardson, E., 2003. *African American Literacies*. Routledge, New York, NY.
- Seltzer, K., de los Rios, C.V., 2021. Understanding Translanguaging in US Literacy Classrooms [Policy brief]. National Council Teachers of English. <https://ncte.org/resources/policy-briefs/>.
- Smitherman, G., 2006. *Word From the Mother: Language and African Americans*. Routledge, New York, NY.
- Spolsky, B., 2010. Introduction: what is educational linguistics? In: Spolsky, B., Hult, F. (Eds.), *The Handbook of Educational Linguistics*. Blackwell Publishing, Malden, PA. <https://doi.org/10.1002/9780470694138.ch1>.
- Walsh, C. (Ed.), 2017. *Pedagogías decoloniales. Prácticas insurgentes de resistir, (re)existir y (re)vivir*. Ediciones Abya-Yala.
- Wei, L., 2008. Research perspectives on bilingualism and multilingualism. In: Wei, L., Moyer, M.G. (Eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Blackwell Publishing, Malden, PA.
- Wynter-Hoyte, K., Smith, M., 2020. "Hey Black child. Do you know who you are?" Using African diaspora literacy to humanize Blackness in early childhood education. *J. Literacy Res.* 52 (4), 406–431. <https://doi.org/10.1177/1086296X20967393>.

Translanguaging and identity across global sites

Monica Shank Lauwo and Bonny Norton, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	148
Translanguaging and identity: conceptual foundations	149
Conceptualizing translanguaging	149
Conceptualizing identity	149
Named languages, translanguaging, and the productive unsettling of identity categories	149
Third space and transformation	150
Transformation and polyphonic identities	151
Identity, investment, and translanguaging	151
Power	152
Ideologies of legitimacy	152
Capital	153
Imagined identities and imagined futures	153
Translanguaging, identity, and pedagogical context	154
A caveat: affordances of named languages	156
Conclusion	156
References	157

Introduction

As the world is increasingly characterized by mobility and fluidity, there is need for conceptualizations of language and identity that account for the complexity of social practices in the 21st century (Darvin and Norton, 2015). Increasingly, scholars are using translanguaging as a lens through which to understand the dynamism that is characteristic of 21st century languaging practices (García and Li, 2014). At the same time, poststructuralist conceptualizations of language and identity offer useful frameworks for understanding the complex relationships between identity and situated languaging choices, within existing structures of power (Norton and Morgan, 2020).

The indexical links between languaging practices and identity are foundational, and an investigation of translanguaging is incomplete without concerted attention to ways in which translanguaging both shapes and is shaped by identity. Translanguaging pushes back against standard language ideologies, monolingual normativity, and discriminatory linguistic hierarchies, ideologies that prop up systems of racism, classism, ethnic discrimination, and colonial relations of power. In so doing, translanguaging can destabilize ideologies that disproportionately marginalize people of color, people living in poverty, and communities that “have come out on the short end of the processes of nation-building, coloniality, and global capitalism” (García and Otheguy, 2020, p. 27), offering opportunities to claim more empowering and multidimensional identity options.

In this chapter, we examine relationships between languaging practices and identity across diverse global sites, exploring ways in which translanguaging practices contribute to the negotiation, construction, and transformation of identities that are dynamic, complex, polyphonic, and shifting across time and space. The chapter is organized in five sections, beginning with conceptualizations of translanguaging and identity, and then moving to Named Languages, Translanguaging, and the Productive Unsettling of Identity Categories. In this second section, we examine translanguaging as a counter-hegemonic process that disrupts restrictive identity categories, including those constructed by elitist Eurocentric ideologies of language and nationhood, enabling speakers to complicate one-to-one equations of language and identity. In the next section, Third Space and Transformation, we draw on the notion of Third Space (Bhabha, 1994) to investigate how translanguaging moves beyond previously existing identity possibilities, energizing creativity and criticality in the process of transformation, and enabling speakers to construct polyphonic identities for themselves that challenge coloniality and unequal relations of power (Shank Lauwo, 2018). In the fourth section, Identity, Investment, and Translanguaging, we examine how the construct of investment (Darvin and Norton, 2015, 2021; Norton, 2013) invigorates the study of translanguaging and identity, detailing the significance of power, ideology, and capital in shaping investment in translanguaging practices. In the final section, Translanguaging, Identity, and Pedagogical Context, we examine the importance of context in translanguaging practices and pedagogy, and compare translanguaging pedagogy in the Global South and Global North. Our discussion concludes with a caveat that addresses the conditions under which named languages may have strategic importance for policy and pedagogy.

Translanguaging and identity: conceptual foundations

Conceptualizing translanguaging

García (2009, p. 140) conceptualizes translanguaging as “the act performed by bilinguals of accessing different linguistic features or various modes of what are described as autonomous languages, in order to maximize communicative potential.” Translanguaging unsettles understandings of language as a stable, decontextualized, bounded, ahistorical system, and focuses instead on communicative practices, linguistic repertoires, and discursive resources. In so doing, translanguaging shifts the analytical gaze from languages and structures to speakers and their practices in the process of meaning making and identity negotiation (e.g., Makalela, 2016; Lin, 2019).

Translanguaging understands language as a “socially, culturally, politically and historically situated sets of features” (Guzula et al., 2016). Rather than being constrained by the social constructs of named languages, speakers make strategic use of various features of their linguistic repertoires in order to meet their communicative and identity construction goals in a particular situation. Translanguaging disrupts the idea that a “normal” or “ideal” speaker is monolingual, instead positioning multilinguals and their communicative practices as normative. While translanguaging involves “using one’s idiolect, that is one’s linguistic repertoire, without regard for socially and politically defined language names and labels” (Li, 2018, p. 19), the social dimensions of translanguaging are of utmost importance, in that there is an audience and communicative objectives. Thus, Anya (2017) frames translanguaging as “a joint, performative, improvisational social accomplishment in which interlocutors work with one another’s linguistic contribution to the interaction to anticipate and fulfill communicative needs and collaboratively construct meaning” (Anya, 2017, p. 25).

Conceptualizing identity

In this chapter, we adopt a poststructuralist understanding of identity as “the way a person understands his or her relationship to the world, how that relationship is structured across time and space, and how the person understands possibilities for the future” (Norton, 2013, p. 4). This conceptualization of identity foregrounds the complex relationships between speakers and the social world, emphasizing that neither identity nor languaging practices can be separated from the social contexts involved. Translanguaging practices contribute to the negotiation of all of these aspects of identity in important ways, including the mediation of power relations. The interdependent relationship between identity and languaging practices is acknowledged by Chris Weedon (1987), a prominent feminist poststructuralist: “Language is the place where actual and possible forms of social organization ... are defined and contested, yet it is also the place where our sense of ourselves, our subjectivity, is constructed” (p. 21).

Translanguaging and poststructuralist conceptualizations of identity share core impulses, making their pairing particularly generative. Through a poststructuralist lens, both identity and translanguaging practices are dynamic, process-oriented, and shifting across time and space. Both are understood as being fluid, hybrid, and complex; they are both negotiated in relations of power, constructed in specific sociocultural and political contexts, and represent sites of struggle.

Named languages, translanguaging, and the productive unsettling of identity categories

Translanguaging complicates one-to-one-equations of language and identity, and liberates language users from rigid identity categories frequently associated with named languages. The ideology equating one national, racial, or ethnic identity with one language was a European invention, tied to the invention of the concept of named languages, and entangled with processes of colonialism and nation building (Makoni and Pennycook, 2007). Indeed, the invention of the very concept of named languages was itself largely a product of a nationalist romanticist ideology presuming an intrinsic link between language and (national) identity (García and Otheguy, 2020; Makoni and Pennycook, 2007). As Li (2018) reminds us, named languages are really sets of structures identified and named by linguistics, with the naming and labeling generally done in collusion with “cultural-political concepts associated with the one-nation/race-one-language ideology” (Li, 2018, p. 19). The concept of named languages, and the related ideology of standard languages, was invented as a tool of domination, serving ideologies of racial, ethnic, class, and gender superiority, and the marginalization or exclusion of minoritized “Others” who do not conform to the dominant group’s languaging practices.

The nation-statist ideology of a named language representing a particular culture and geopolitical entity still pervades much language education, including foreign language education in particular. Students in foreign language classrooms are generally purported to be learning the standardized form of a named language that presumably corresponds to a specific foreign nation-state, and are being prepared to communicate with monolingual native speakers associated with a particular culture in this foreign context (Anya, 2017; García and Otheguy, 2020).

Translanguaging, as a practice and as a pedagogy, disrupts the domination and erasure embedded in one-nation-one-language ideology, and “releases histories and understandings that had been buried within fixed language identities constrained by nation-states (see Mignolo, 2000)” (García and Li, 2014, p. 20). Translanguaging also pushes back against raciolinguistic ideologies that grant legitimacy only to the linguistic and cultural practices and knowledge systems stereotypically associated with white monolingual people, and legitimizes the languaging practices, knowledges, and abilities of communities who have been racialized through processes of colonialism (García et al., 2021). In this way, translanguaging contributes to the unsettling of both linguistic and racial

categories (and their co-naturalization), as part of broader struggles toward decolonization and the dismantling of white supremacy (Rosa and Flores, 2017).

One compelling illustration of how translanguaging can powerfully and productively deconstruct racial and ethnic categories comes from Kerfoot and Bello-Nonjengele's (2016) ethnographic work in a multi-racial, multi-ethnic, multilingual school near Cape Town, South Africa. The school, attended by an unusual combination of Black African and colored students, is described as "a site in which 'what was once thought of as separate—identities, spaces, histories—come together or find points of intersection in unexpected ways' (Nuttall, 2009: 10)" (Kerfoot and Bello-Nonjengele, 2016, p. 453). They document how children's languaging practices disrupt the ways in which relationships between language, ethnicity, and "race" had previously been enforced and naturalized, as children "conscious[ly] and strategic[ally] use their semiotic repertoires as identity-building resources" and to negotiate power hierarchies (Kerfoot and Bello-Nonjengele, 2016, p. 468). Children's strategic identity choices, including the "mother tongue" they choose to assign to themselves for language of instruction placements, are "often in conflict with parental ideologies, subverting essentialist notions of linguistic ownership" (Kerfoot and Bello-Nonjengele, 2016, p. 468). They demonstrate how the "decoupling of the links between linguistic forms and social categories, the basis of apartheid policies, has fundamental implications for current language-in-education policies still based on essentialized understandings of ethnicity, race, and language" (Kerfoot and Bello-Nonjengele, 2016, p. 469).

In a different South African context, Nkadimeng and Makalela (2015) note how translanguaging supported multilingual high school students in Soweto to use the "hybrid language space to identify and ethnify—choosing who they want to become beyond traditional linguistic and exact ethnic affiliations" (p. 7). In a context where boundaries between both languages and ethnic identities were prescribed by colonial forces (see section on *Translanguaging, Identity, and Pedagogical Context*), the hybridization of linguistic practices and ethnic affiliations has a decolonial character. In these ways, translanguaging can productively disrupt ideologies of nationhood, colonialism, and race/ethnicity that naturalize parallel borders around languages and social categories, creating space for both hybridity and transformation.

Third space and transformation

Translanguaging practices can expand imaginations of what is linguistically and epistemologically possible, opening space for the emergence of new identities and ways of being. As Li (2018, p. 26) notes, human social interaction is overwhelmingly characterized by "alternation between languages, spoken, written, or signed; between language varieties; and between speech, writing, and signing". These normative practices "construct an identity for the speaker that is different from a La identity or a Lb identity" (Li, 2018, p. 26).

While translanguaging practices are expressions of hybridity, they also go beyond hybridity to create novel languaging practices. The concept of Third Space (Bhabha, 1994) accounts for the way translanguaging goes beyond combining existing communicative and cultural possibilities, to create a generative space characterized by transformation and creativity. Gutiérrez (2008) understands Third Space as "a transformative space where the potential for an expanded form of learning and the development of new knowledge are heightened" (p. 152). Drawing from Soja (1996), Li (2018) conceptualizes a translanguaging Third Space as "a space of extraordinary openness, a place of critical exchange where the geographical imagination can be expanded to encompass a multiplicity of perspectives that have heretofore been considered by the epistemological referees to be incompatible and uncombinable" (pp. 23–24).

Significantly, a Third Space also enables expanded imaginations of identity possibilities, opening space for the creative combination of identity features previously considered to be "incompatible and uncombinable". Shank Lauwo and Loth Mollel (forthcoming) detail how the Third Spaces created through practices of translanguaging and multimodality in a Tanzanian community library support children to embrace the complementarity of seemingly disparate identities, languaging practices, cultures, and values that they perform and embody in in-school and out-of-school contexts. In this context, translanguaging helps children to find ways of bridging the worlds and challenging damaging hierarchies between languages, cultures, and knowledges that are constructed as bordered and existing in spatial isolation.

Another example of transformed identity possibilities enabled by a translanguaging Third Space comes from Flores and García's (2013) study of translanguaging practices at a New York secondary school of Spanish-speaking immigrant adolescents. Flores and García (2013) examine how a bilingual teacher juxtaposes Latin American hip-hop and American pop music to support her students to explore complex political-historical social justice issues relating to US imperialism and the propping up of violent dictatorships. In so doing, the class creates a translanguaging Third Space that facilitates students' language learning, investment, and critical engagement, as well as the development of complex ethnolinguistic identities. "A linguistic third space is developed in the interaction between a language (Spanish) and style (hip-hop) with which students are more familiar, and a language (English) and style (pop) which are less known. Thus, students not only cross national and linguistic borders (Rampton, 2006) but go beyond them" (Flores and García, 2013, p. 252). This linguistic, cultural, and geographical Third Space moves beyond blending, resulting in the emergence of trans-features and trans-identities: "putting Spanish and English alongside each other, many times with features that are 'trans'—blends of both—starts to construct an ethnolinguistic identity that is not solely Latin American and Spanish monolingual, and not solely Anglo US and English monolingual, but a US Latino bilingual identity" (Flores and García, 2013, p. 248).

We thus see the transformative power of translanguaging, as it transgresses existing possibilities and opens space for the emergence of new identities and communicative practices, as well as transformed power relations, epistemological possibilities, and social structures. [Shank Lauwo \(2018\)](#) demonstrates how “the practice of translanguaging offers tremendous transformative power, through its ability to reconfigure entrenched power hierarchies between languages and their speakers” (p. 144). [García and Li \(2014\)](#) describe the transformative nature of translanguaging thus:

as new configurations of language practices and education are generated, old understandings and structures are released, thus transforming not only subjectivities, but also cognitive and social structures. In so doing, orders of discourses shift and the voices of Others come to the forefront, relating then translanguaging to criticality, critical pedagogy, social justice and the linguistic human rights agenda (p. 3).

The transformative power of translanguaging, which makes it a potent tool for social justice, is tied to its relationship with creativity and criticality. Translanguaging, as a perpetually emergent languaging practice that is characterized by dynamism and innovation, is an inherently creative practice. Unconstrained by prescriptive linguistic norms, translanguaging enables constant engagement with communicative *options*, with playful experimentation with new possibilities, and therefore invites the constant renegotiation of identities, values, and perspectives. Translanguaging practices, inasmuch as they push back against established relations of power and top-down impositions of linguistic norms, are also inherently critical. Criticality, which interrogates power relations, questions the status quo, and seeks more just alternatives, is deeply tied to creativity, and both are essential to imagining and constructing more equitable futures. As [Li \(2018\)](#) contends, “From a Translanguaging lens, multilingualism by the very nature of the phenomenon is a rich source of creativity and criticality, as it entails tension, conflict, competition, difference, and change in a number of spheres, ranging from ideologies, policies, and practices to historical and current contents” (p. 23). Creativity and criticality through the practice of translanguaging enable speakers to transcend received identity possibilities that may be experienced as bordered or limiting, and to creatively and critically construct alternative identities for themselves, individually and collectively, that may be experienced as more empowering, responsive to shifting circumstances, and in service of the imagined futures in which they are invested.

Transformation and polyphonic identities

Translanguaging gives voice to the complexity and multivalent nature of multilinguals’ identities, identities described elsewhere as “polyphonic” ([Shank Lauwo, 2018](#)). The term polyphony describes a musical texture in which multiple melody lines are featured simultaneously. Conceptualizing identities as polyphonic points to the ways in which “identities, like languages, are not discrete, bordered, nor mutually exclusive, but can be complementary, interwoven, and interdependent” ([Shank Lauwo, 2018](#), p. 141). Multiple dimensions, or voices, of our polyphonic identities are simultaneously audible in any languaging act or semiotic expression we may make. Sometimes certain melody lines, or identity features, are foregrounded, while other times various identity features are equally pronounced as part of a complex and interdependent whole. Sometimes identity features co-exist and intermingle harmoniously, while other times they create dissonance, producing tension, and perhaps competing for dominance. Significantly, while different melody lines, or identity dimensions, may have different features or characteristics that previously moved independently, in polyphony these melodies/identities are experienced as one dynamic, new whole, with a continuously evolving progression of harmonies, dissonances, and complexities.

For multilinguals, these polyphonic identities are creatively expressed—and, indeed constructed—through translanguaging, which carries these dynamics of harmony and dissonance, tension and competition, and the emergence of something new. Indeed, translanguaging practices, like identities, are also polyphonic, and sometimes linguistic features associated with a particular named language may be dominant, while at other times the features may carry so many linguistic influences simultaneously that they defy categorization. Important to both polyphony, or any form of music, and translanguaging is the dimension of time. Translanguaging, like music, is a temporal expressive medium, utilizing the resource of time in the expression and construction of identities. Because of this temporal dimension, translanguaging practices, like polyphonic identities, by their very nature are ever-changing, shifting over the course of a conversation, or a day, or a lifetime. Any given excerpt of expression of our polyphonic identities is necessarily partial, relying on everything that came before and everything yet to come for its ever-elusive completion.

Translanguaging spaces help to nurture these polyphonic identities, offering opportunities to negotiate the complementarity and mutual interdependence of diverse identity dimensions. Indeed, one of the goals of translanguaging pedagogy is “to create a space in which these multiple identities can be brought productively into conversation with each other” ([Shank Lauwo, 2018](#), p. 141), in order for speakers to “[bring] together different dimensions of their personal history, experience and environment, their attitude, belief and ideology, their cognitive and physical capacity into one coordinated and meaningful performance” ([Li, 2011](#), p. 1223).

Identity, investment, and translanguaging

In addition to supporting the creation of linguistic and epistemological Third Spaces and the affirmation and expansion of polyphonic identities, translanguaging can support learners’ investment in learning. Investment, understood as a sociocultural

complement to the psychological construct of motivation, occurs when learners perceive their identities to be valued, their cultural capital is positioned as a resource, and the ideological environment of the learning space is supportive of the learner (Darvin and Norton, 2015, 2021). If learners invest in learning particular languaging practices, they do so with the understanding that they will acquire a wider range of symbolic and material resources, which will in turn increase the value of their cultural capital and social power (Norton, 2013). While a learner might be highly *motivated* to learn, they may struggle to be *invested* in learning if the learning environment devalues their languaging practices, culture, or identities, or if it is, for example, racist, classist, patriarchal, or homophobic. This section examines intersections between translanguaging, identity, and investment in learning, by addressing themes of power, ideologies of legitimacy, cultural capital, and imagined identities and imagined futures.

Power

At the heart of translanguaging's transformative power is its ability to shift relations of power, enabling those who are the most marginalized by monoglossic and standard language ideologies to claim agency as legitimate knowers, speakers, and knowledge producers. Translanguaging enables those with rich languaging practices to shed ascribed identities as multilingual "Others" and position themselves as agential communicative subjects. Translanguaging is a potent act of resistance to ideologies that privilege monolingualism and that grant greater power to certain languages (such as English) and language varieties, thereby shifting power relations between both languages and their speakers (Shank Lauwo, 2021). Indeed, as García and Otheguy (2020) assert, translanguaging pedagogy has the development of agency as a central emphasis: "Translanguaging pedagogy ... focuses on building the agency of the learner to language in order to act and mean as a bilingual" (García and Otheguy, 2020, p. 28). Certainly, relations of power have significant implications for learners' investment in their learning, and the identities they are able to adopt in the process of learning (Darvin and Norton, 2015).

The story of Namelok, a 10-year-old Maasai girl in rural Tanzania who had never been to school, illustrates how the transforming of power relations enabled by a translanguaging space can ignite possibilities for investment in learning and the expansion of identity options (Shank Lauwo, 2019). Fearful and timid in her initial interactions with a multilingual community library, the positioning of her dominant language of Maa as an asset in multilingual games meant that other children, who had previously belittled her abilities, clamored to have her on their team. The decentring of the hegemony of Kiswahili and English and of the written word offered her space to engage with books through picture reading and begin authoring her own stories through drawing and oral narrations. As her confidence grew through these successful literacy experiences, Namelok, on her own initiative, began teaching Maasai songs and Maa vocabulary to other children and to facilitators, and punctuating her picture reading with newly-learned English words. She led a group performance of singing and dancing, and comforted a girl in distress by orally narrating her own rendition of a picture book to her. In these ways, the translanguaging space enabled Namelok, who would have been marginalized by the monolingual Kiswahili and written-word-centered environment of formal schooling, to invest in the learning space and to claim identities as a knower, language expert, author, valued team member, reader, English speaker, teacher, leader, and helpful friend. Indeed, her new-found confidence enabled by the valorization of her full linguistic, cultural, and semiotic repertoires supported her investment in the library's language and literacy activities, thereby unleashing her creative agency, and enabling her to practice skills that could lead to expanded literacy abilities in the future (Shank Lauwo, 2019).

Ideologies of legitimacy

Ideology pertains to systemic patterns of control, often invisible, that shape our thinking and behavior based on a normative set of ideas (Darvin and Norton, 2015). These systemic patterns of control can both constrain and/or enable investment, as they can also facilitate or inhibit possibilities for a learner to position themselves as a legitimate speaker (Bourdieu, 1991). With regards to legitimacy of speech, Bourdieu (1977) posits that,

Just as, at the level of the relations between groups, a language is worth what those who speak it are worth, so too, at the level of interactions between individuals, speech owes a major part of its value to the value of the person who utters it (p. 653).

Translanguaging offers opportunities to transgress ideologies that limit access of certain people in certain contexts to identities as legitimate speakers, opening space for all people to claim identities as legitimate speakers, knowers, and knowledge producers. Specifically, translanguaging pushes back against ideologies of native speakerism and of distinctive L1 and L2 identities, to acknowledge and valorize linguistic resources and processes of language acquisition that defy categorization into countable and chronological units.

Pedagogically, translanguaging can deconstruct orders of value of various named languages, languaging practices, and speakers, and construct alternative ideologies in a particular classroom or learning environment. When learners' identities, languaging practices, and efforts at learning new communicative practices are valued and positioned as legitimate, this sense of legitimacy enables learners to invest in the languaging practices of a given learning/communicative environment. Supporting learners to utilize their full linguistic, cultural, and semiotic repertoires in a learning environment acknowledges their agency as legitimate communicators, even when their proficiency in the 'target' language or languaging practices may still be emergent.

Significantly, classrooms or learning environments exist in broader structures of power that “construct modes of inclusion and exclusion, and [position] learners in multiple ways before they even speak” based on “embodied identities inscribed by race, ethnicity, gender, and social class” (Darvin and Norton, 2015, p. 43). While pedagogy alone may not be able to address systemic patterns of control at the macro level, it can prompt ideological shifts at the micro level, with long-term ramifications for learners’ investment and the expansion of their identity possibilities, in the classroom and beyond.

Capital

As Darvin and Norton (2015) note, capital is essentially about power, and, drawing on Bourdieu (1986), includes economic, cultural, and social capital. Language is a form of cultural capital. The extent to which learners’ linguistic or cultural capital is valued—including their languaging practices, their prior knowledge, and their diverse semiotic and cultural practices—impacts their ability to invest in learning in a given context. As Darvin and Norton (2015) argue,

learners invest in particular practices not only because they desire specific material or symbolic benefits, but also because they recognize that the capital they possess can serve as affordances to their learning. The valuing of their capital is an affirmation of their identity, a legitimization of their rightful place in different learning contexts (p. 46).

When certain forms of capital are recognized and valued in a particular field (such as a classroom), they become symbolic capital (Bourdieu, 1986). The challenge then is to create a learning environment that positions all learners’ cultural capital as symbolic capital, in order to support their investment in learning. Translanguaging spaces can support this investment, be it through valuing learners’ out-of-school translanguaging practices as linguistic capital, or supporting the fluid use of learners’ full linguistic repertoires alongside the new linguistic features they are engaged in appropriating.

The South African children in Kerfoot and Bello-Nonjengele (2016), described above, demonstrate how translanguaging practices, as counter-hegemonic processes, can shift the value of linguistic capital and rupture entrenched language hierarchies, creating new possibilities for investment and identity negotiation. Kerfoot and Bello-Nonjengele (2016, p. 469) observe that, “The frequent reordering of expected Afrikaans >English>Xhosa hierarchies, the harmonious use of isiXhosa, English, and Afrikaans in working on school tasks, the egalitarian multilingual exchanges, all work to revalue linguistic capital both in the classroom and outside it.” Significantly, it was the children, far more than teachers, the school, or educational policies, who were able to shift the ways in which linguistic features were attributed value and circulated, and were the principal agents in the contestation and renegotiation of rates of exchange of different forms of linguistic and cultural capital, including “race”, ethnicity, and gender. These children’s translanguaging practices usher us toward ideologies of post-racial solidarity, the reordering of economies of meaning, and pedagogical possibilities that support investment through the reimagination of restrictive identity categorizations.

Imagined identities and imagined futures

Certainly, learners’ investment in learning particular languaging practices is entangled with what communities they aspire to be members of in the future. Imagined communities refer to groups of people with whom one aspires to be affiliated, even when there is little immediate affiliation for a given person (Kanno and Norton, 2003; Norton and Pavlenko, 2019). Such imagined communities that might compel learners to invest in certain languaging practices might include a professional network, a sporting community, a community interested in a particular kind of music, or a religious community. An imagined community assumes an imagined identity, and learners invest in language learning in hopes of realizing these imagined identities and imagined futures (Norton, 2013).

Translanguaging supports learners to imagine diverse futures that may be less accessible in a monoglossic environment. While the learning of English, even in a monolingual space, may expand what is socially imaginable for many learners, and grant them access to an imagined community of international English speakers, translanguaging supports the multidimensionality of these imagined futures, and creates imaginative space for the complementarity of various future aspirations. Translanguaging also enables learners to use English features in their repertoires to position themselves as English speakers (an identity that may be crucial to their imagined futures), while drawing from their entire linguistic and semiotic repertoires to ensure the fullness of their intended meaning-making.

For example, Shank Lauwo (2018, 2019) documents how a translanguaging space supported Tanzanian children to use diverse linguistic and semiotic resources to imagine multidimensional futures for themselves. In an identity text of how she imagined herself in the future (see Fig. 1), Happiness introduced herself with a scripted, school-learned English paragraph, that concluded with the (predominantly Kiswahili) sentence, “Ndoto zangu kuwa Doctor” (“My dreams are to be a Doctor”). Happiness’ use of English positions her as someone who is well-educated, and signifies the centrality of English proficiency to her aspiration of becoming a doctor. She drew a girl (presumably herself) outside of a beautiful house, and used a combination of Kiswahili and Maa to write prayers for God to help her become a doctor. She also wrote creative transcriptions of the lyrics of a famous English song, which she titled “Wimbo wa Kiingereza” (“English song”—in Kiswahili). Happiness thus used diverse linguistic and semiotic resources (features from Maa, Kiswahili, English; drawing, writing, and music) to construct identities as a doctor, a Christian, a school-educated person, and Maasai, thus creating the unambiguous message that these linguistic and semiotic features and these identities can co-exist in complementarity.

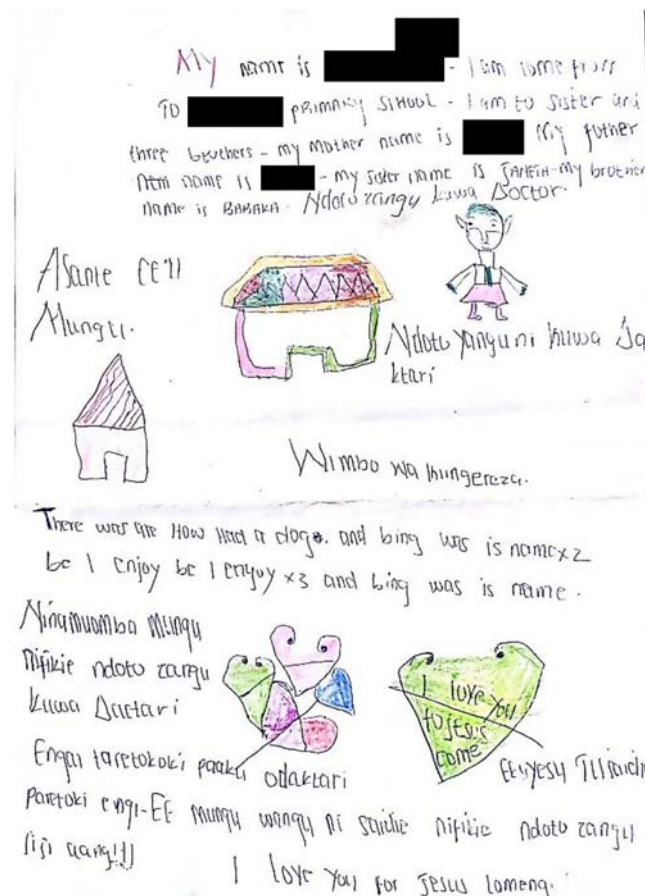


Fig. 1 Happiness' Identity Text. Reproduced with permission from Shank Lauwo, M., 2018. Power, literacy engagement, and polyphonic identities: translanguaging in a Tanzanian community library. *South. Afr. Ling. Appl. Lang. Stud.* 36 (2), 133–146.

In a similar future-oriented identity text (see Fig. 2), Daudi used both Kiswahili and Maa to express his dream to be a teacher in order to teach reading and writing in Maa—a transformative ambition, as Maa, his dominant language, had been actively muted in his formal schooling experiences. He used drawings and inscriptions in Kiswahili, English, and Maa to depict various in-school and out-of-school activities, signifying the importance of all these languages to a range of activities and to his ambition. Translanguaging enabled Daudi to simultaneously imagine himself as a highly-educated professional and to transgress his received images of what such a professional does, as he brought together his Maa-speaking identity, his passion for literacy (taught in Tanzanian schools in Kiswahili and English only), and his ambition to become a literacy teacher.

Taken together, the employment of translanguaging in the examples of Namelok, South African school children, and Happiness and Daudi illustrates how translanguaging can support language users to claim power and identities of legitimacy, and to assert the value of their cultural capital in ways that enhance their investment in both learning and reimagined social futures.

Translanguaging, identity, and pedagogical context

While translanguaging has immense transformative potential, it always occurs in particular communicative contexts. Context plays a central role in shaping, not only translanguaging practices themselves, but also their significance for identity and pedagogy. As Anya (2017, p. 25) notes, a translanguaging perspective posits that languages “cannot be separated from the persons and social contexts involved”. Further, power relations between different languaging practices and their speakers shape the meanings associated with translanguaging, and the identities constructed through translanguaging in specific contexts. These power relations shift dramatically across time and space, including across Global South and Global North contexts.

In the Global South, McKinney and Tyler (2019) remind us of the situated nature of translanguaging when they insist that “no communicative practice is by definition transformative (or constraining)” (McKinney and Tyler, 2019, p. 146). In her commentary on a 2019 special issue in *Language and Education* on “translingual and transmodal complexity” in the Global South (Prinsloo, 2019), Zavala (2019) asserts,

Translanguaging pedagogy is not an essential or an absolute good ... Translanguaging pedagogy works differently in different contexts, as happens always with language use, where there is no direct connection between form and function. When we research language use in specific contexts—within a situated language approach – functions change all the time and in non-predictable ways (p. 175).

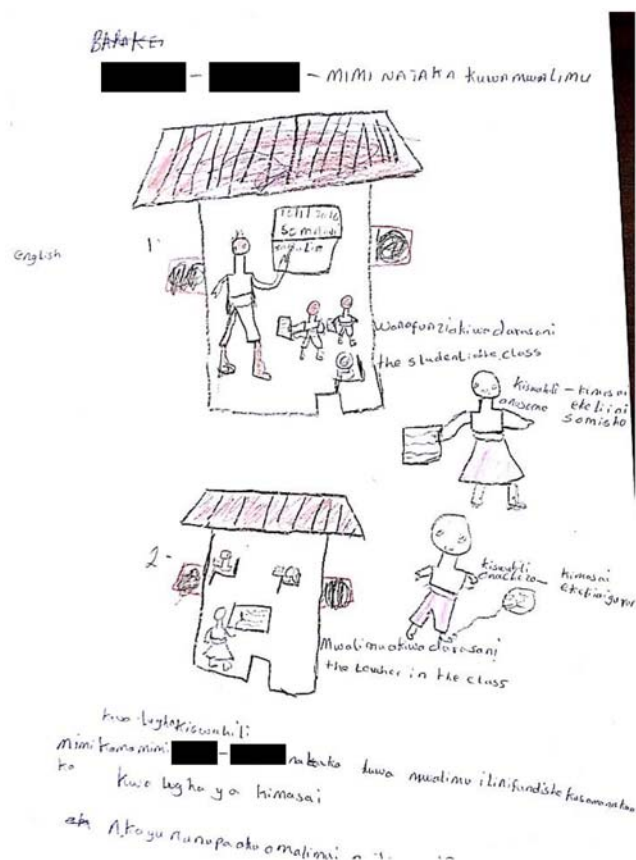


Fig. 2 Daudi's Identity Text. Reproduced with permission from Shank Lauwo, M., 2018. Power, literacy engagement, and polyphonic identities: translanguaging in a Tanzanian community library. *South. Afr. Ling. Appl. Lang. Stud.* 36 (2), 133–146.

Translanguaging in many Southern contexts, including South Africa, has an important decolonial character. South African scholars of translanguaging highlight the colonial invention of African languages, detailing how 19th century colonial missionary linguists named, codified, and standardized African languaging practices to produce discrete “languages” that reflected European ideologies of one-language-one-culture (e.g., Makalela, 2015a, 2016; Prinsloo and Krause, 2019). Apartheid policies further entrenched these linguistic distinctions by separating Africans into “homelands” or “Bantustans” based on perceived language differences, while the post-Apartheid government has “clumsily retained this earlier commitment to discrete languages” (Prinsloo and Krause, 2019, p. 164) by further institutionalizing the differences, including with the declaration of nine distinct “African languages” as official languages (Makalela, 2015a).

Currently, the languaging practices dominant in many South African townships are profoundly hybrid and fluid, representing “disruption and simultaneous recreation of new linguistic and social identities in these multilingual complex spaces” (Nkadameng and Makalela, 2015, pp. 10–11). Nkadameng and Makalela (2015) demonstrate that, for many urban Black South African youth, there is no correspondence between their ethnic and linguistic affiliations, as their linguistic repertoires are intensely hybrid and pluralistic, involving features associated with sometimes eight or more named languages. In this context, the concept of mother tongue loses currency, and the monolingual ideology behind the notion of mother tongue, presuming every person to have a single dominant language and constructing the monolingual child as normative, has little meaning in this context (McKinney and Tyler, 2019). Nkadameng and Makalela (2015) document how youth use diverse languaging practices to move across various ethnic and social identities, demonstrating processes of mother tongue disaffiliation, ethnic re-affiliations, blurring of identity boundaries, and resistance to monoglot identities. These hybrid languaging practices deconstruct the rigid ethnic and racial categories used to separate people under apartheid, and create fluid social spaces characterized by creativity and criticality, enabling speakers to choose who they want to be in any given discursive interaction (see Kerfoot and Bello-Nonjengele, 2016; Makalela, 2015b; Prinsloo and Krause, 2019).

Leketi Makalela’s influential framework of ubuntu translanguaging demonstrates how fluid languaging practices reflect African ways of knowing and making sense of the world, including African epistemological orientations toward confluence, interdependence, and incompleteness (e.g., Makalela, 2015b, 2016). Pedagogically, the normative nature of such fluid languaging practices points to the need to position translanguaging resourcefulness as both an invaluable resource in learning, and as an educational

goal, with a long-term objective being to undermine the hegemony of standard language ideologies in broader society (McKinney and Tyler, 2019; Prinsloo and Krause, 2019).

While in many contexts of the Global South, current expressions of linguistic diversity have deep historical roots (as with the South African example above), in much of the Global North, current experiences of superdiversity (Vertovec, 2007) are largely the result of exponential increases in global migration flows, in addition to the enduring presence of a rich diversity of Indigenous languages and cultures. These rapidly proliferating global migration flows, resulting in rapid increases in linguistic, cultural, and racial pluralism, are characterized particularly strongly by South-to-North migration, as processes of global capitalism and (neo) colonialism make many Southern contexts increasingly unlivable. Translanguaging in the Global North is thus frequently related in some way to transnationalism, involving combinations of linguistic features that have traveled considerable distances. In many North American contexts, for instance, translanguaging practices typically involve some linguistic features associated with English (which has come to be the dominant named language of North America after traveling considerable distances itself, and being constructed through processes of translanguaging), and linguistic features associated with more recent histories of migration. Thus, for many speakers, translanguaging practices are indexical of complex transnational identities. For example, the teachers and students featured in Flores and García (2013), described above, engage in translanguaging to negotiate “transcultural and transnational Pan-Latino US identities” (p. 251), engaging with linguistic, semiotic, and cultural features from various Latin American contexts and from the United States. In a different example, Sayer (2013) describes how second grade students’ translanguaging, involving features associated with both standard and local vernacular varieties of Spanish and English, indexes identities as Mexican, American, and Tejano simultaneously.

A caveat: affordances of named languages

While translanguaging can be an empowering resource for communication, pedagogy, and identity negotiation, in certain instances there are strategic reasons for upholding notions of bordered, named languages (Doherty et al., *in press*). In some contexts, certain advocacy objectives (such as more privileged positioning in curriculum or policy, heightened access to publishing opportunities, etc.) rely on strategic constructions of named languages, particularly for language features with less proximity to power. Also, as Creese and Blackledge (2015) point out, “If languages and identities are socially constructed, we nevertheless need to account for the fact that at least some language users, at least some of the time, hold passionate beliefs about the importance and significance of a particular language to their sense of ‘identity’” (p. 25).

Sheris et al. (2019) document one instance in which ideologies of named languages may have strategic importance. They examine the paradox of needing to simultaneously construct Safaliba, a minoritized Ghanaian language, as a bordered language for status-raising advocacy such as inclusion in school curriculum, and to validate the translanguaging practices that are the norm in the Safaliba-speaking community. Their analysis points to ways in which descriptive linguistic analysis, based on mostly static or normed notions of language, might productively co-exist with fluid, dynamic languaging practices and poststructuralist language ideologies, for gains in both advocacy and pedagogy. Indeed, this seemingly paradoxical approach can result in increased status associated with Safaliba-speaking identities while affirming the identities associated with translanguaging. Again, the role of power is significant, as this status-raising advocacy may be particularly crucial for language features positioned as less powerful in society.

While in some contexts, advocates of language preservation argue that translanguaging further erodes the already precarious position of minority languages (see Nguyen, 2019 for a Vietnamese example), in some contexts translanguaging may contribute to the preservation of linguistic features that would otherwise be further endangered. As Li (2018) notes, while Teochow and Hokkien, as bordered entities, are quickly disappearing in Singaporean Chinese communities, many features associated with Teochow and Hokkien are being dynamically incorporated into the languaging practices of Singaporeans, thereby preserving words and expressions that might otherwise be lost.

Conclusion

As we have demonstrated, translanguaging as a pedagogy and as a counter-hegemonic practice has much potential to disrupt limiting and prescriptive identity categorizations, empowering speakers to generate novel and complex identity possibilities as individuals and as communities. As a bottom-up process characterized by creativity, criticality, and transformation, translanguaging enables speakers to create a linguistic, cultural, and epistemological Third Space. This Third Space, as a generative space of innovation and newness, unleashes creative agency to imagine an expanded range of identities across time and space, and to negotiate imagined futures of both individuals and society. Significantly, translanguaging, as a pedagogy and as a practice, supports an orientation of complementarity, in which diverse aspects of a person’s background, beliefs, community affiliations, ideological orientations, and aspirations can co-exist and be simultaneously audible, in shifting ways, as part of their complex polyphonic identities. Pedagogically, translanguaging shifts relations of power and creates space for the valorization of learners’ diverse identities and cultural capital, opening opportunities for investment in learning and knowledge production.

Importantly, while the nurturing of translanguaging spaces in classrooms can have a transformative impact on power relations, investment, and identity possibilities, the broader goal is to transform systems of inequity, and their attendant discriminatory ideologies, in society at large. Translanguaging on its own (or any pedagogical intervention for that matter) cannot dismantle systems of capitalism and white supremacy, nor bring about the transformative redistribution necessary to address the material bases of inequity (Block, 2018; Fraser, 1995). Rather, translanguaging, and broader efforts towards more equitable and pluralistic education, must be situated within broader social and political economic struggles to dismantle the global systems of racial capitalism that produce inequity, poverty, and injustice (see Kubota, 2016; Rosa and Flores, 2017).

As we work toward reimagining and reconstituting global systems as we know them, translanguaging practices—and the rich complexity of world views, identities, and cultural and epistemological possibilities these construct—will continue to creatively and critically invigorate decolonial, antiracist resistance to the status quo. Our challenge is to continue to cultivate translanguaging spaces that support us to collectively imagine and construct futures that are more equitable, and that embrace the complexity of human identities.

References

- Anyá, U., 2017. *Racialized Identities in Second Language Learning: Speaking Blackness in Brazil*. Routledge, New York, NY.
- Bhabha, H.K., 1994. *The Location of Culture*. Routledge, London.
- Block, D., 2018. The political economy of language education research (or the lack thereof): Nancy Fraser and the case of translanguaging. *Crit. Inq. Lang. Stud.* 15 (4), 237–257.
- Bourdieu, P., 1977. The economics of linguistic exchanges. *Soc. Sci. Inf.* 16 (6), 645–668.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J.F. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, New York, NY, pp. 241–258.
- Bourdieu, P., 1991. *Language and Symbolic Power*. Harvard University Press, Cambridge, MA.
- Creese, A., Blackledge, A., 2015. Translanguaging and identity in educational settings. *Annu. Rev. Appl. Ling.* 35, 20–35.
- Darvin, R., Norton, B., 2015. Identity and a model of investment in applied linguistics. *Annu. Rev. Appl. Ling.* 35, 36–56.
- Darvin, R., Norton, B., 2021. Investment and motivation in language learning: what's the difference? *Lang. Teach.* 1–12.
- Doherty, L., Norton, B., Stranger-Johannessen, E. (in press). Translation, identity, and translanguaging. In: Ayers-Bennett, W., Fisher, L., (Eds.), *Multilingualism and Identity: Interdisciplinary Perspectives*. Cambridge University Press, Cambridge, UK.
- Flores, N., García, O., 2013. Linguistic third spaces in education: teacher's translanguaging across the bilingual continuum. In: Little, D., Leung, C., van Avermaet, P. (Eds.), *Managing Diversity in Education: Languages, Policies and Pedagogies*. Multilingual Matters, Bristol, pp. 243–256.
- García, O., Li, W., 2014. *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan, Hampshire.
- García, O., Otheguy, R., 2020. Plurilingualism and translanguaging: commonalities and divergences. *Int. J. Biling. Educ. BiLing.* 23 (1), 17–35.
- García, O., Flores, N., Seltzer, K., Li, W., Otheguy, R., Rosa, J., 2021. Rejecting abyssal thinking in the language and education of racialized bilinguals: a manifesto. *Crit. Inq. Lang. Stud.* 18 (3), 203–228.
- Fraser, N., 1995. From redistribution to recognition? dilemmas of justice in a “post-socialist” age. *New Left Rev.* 212, 68–93.
- García, O., 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Wiley/Blackwell, Malden.
- Gutiérrez, K., 2008. Developing a sociocritical literacy in the third space. *Read. Res. Q.* 43 (2), 148–164.
- Guzula, X., McKinney, C., Tyler, R., 2016. Language-for-learning: legitimising translanguaging and enabling multimodal practices in third spaces. *South. Afr. Ling. Appl. Lang. Stud.* 34 (3), 211–226.
- Kanno, Y., Norton, B., 2003. Imagined communities and educational possibilities. *J. Lang. Identity Educ.* 2 (4).
- Kerfoot, C., Bello-Nonjengele, B.O., 2016. Game changers? Multilingual learners in a Cape Town primary school. *Appl. Ling.* 37 (4), 451–473.
- Kubota, R., 2016. The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: complexities and implications for applied linguistics. *Appl. Linguist.* 37 (4), 474–494.
- Li, W., 2011. Moment analysis and translanguaging space: discursive construction of identities by multilingual Chinese youth in Britain. *J. Pragmat.* 43 (5), 1222–1235.
- Li, W., 2018. Translanguaging as a practical theory of language. *Appl. Ling.* 39 (1), 9–30.
- Lin, A.M.Y., 2019. Theories of trans/language and trans-semiotizing: implications for content-based education classrooms. *Int. J. Biling. Educ. BiLing.* 22 (1), 5–16.
- Makalela, L., 2015a. Moving out of linguistic boxes: the effects of translanguaging strategies for multilingual classrooms. *Lang. Educ.* 29 (3), 200–217.
- Makalela, L., 2015b. Translanguaging practices in complex multilingual spaces: a discontinuous continuity in post-independent South Africa. *Int. J. Sociol. Lang.* 234, 115–132.
- Makalela, L., 2016. Ubuntu translanguaging: an alternative framework for complex multilingual encounters. *South. Afr. Ling. Appl. Lang. Stud.* 34 (3), 187–196.
- Makoni, S., Pennycook, A. (Eds.), 2007. *Disinventing and Reconstituting Languages*. Multilingual Matters, Clevedon.
- McKinney, C., Tyler, R., 2019. Disinventing and reconstituting language for learning in school science. *Lang. Educ.* 33 (2), 141–158.
- Mignolo, W., 2000. *Local Histories/Global Designs. Coloniality, Subaltern Knowledges, and Border Thinking*. Princeton University Press, Princeton.
- Nguyen, T.T.T., 2019. Translanguaging as trans-identity: the case of ethnic minority students in Vietnam. *Lingua* 222, 39–52.
- Nkadiemeng, S., Makalela, L., 2015. Identity negotiation in a super-diverse community: the fuzzy languaging logic of high school students in Soweto. *Int. J. Sociol. Lang.* 234, 7–26.
- Norton, B., Morgan, B., 2020. Poststructuralism. In: Chapelle, C. (Ed.), *The Concise Encyclopedia of Applied Linguistics*. John Wiley & Sons Inc., pp. 901–907.
- Norton, B., Pavlenko, A., 2019. Imagined communities, identity, and English language learning in a multilingual world. In: Gao, X. (Ed.), *Second Handbook of English Language Teaching*. Springer International Handbooks of Education: Springer, Cham, pp. 703–718.
- Norton, B., 2013. *Identity and Language Learning: Extending the Conversation*, second ed. Multilingual Matters, Bristol, UK.
- Prinsloo, M., Krause, L.-S., 2019. Translanguaging, place and complexity. *Lang. Educ.* 33 (2), 159–173.
- Nuttall, S., 2009. *Entanglement: Literary and Cultural Reflections on Post Apartheid*. Wits University Press, Johannesburg.
- Translingual and transmodal complexity and innovation in English-language-dominant Southern schooling. In: Prinsloo, M. (Ed.), *Lang. Educ.* 33 (2).
- Rampton, B., 2006. *Language in Late Modernity: Interaction in an Urban School*. Cambridge University Press, Cambridge.
- Rosa, J., Flores, N., 2017. Unsettling race and language: toward a raciolinguistic perspective. *Lang. Soc.* 46 (5), 621–647.
- Shank Lauwo, M., Loth Mollel, U. (forthcoming). Neoliberalism, transformation, and educational imperatives in tension: a duoethnography of collaboration and contestation in a Tanzanian community library. In: Gagné, A., Kalan, A., Herath, S. (Eds.), *Critical Action Research Challenging Neoliberal Language and Literacies Education: Auto and Duoethnographies of Global Experiences*. Peter Lang.
- Sayer, P., 2013. Translanguaging, TextMex, and bilingual pedagogy: emergent bilinguals learning through the vernacular. *TESOL Q.* 47 (1), 63–88.
- Shank Lauwo, M., 2018. Power, literacy engagement, and polyphonic identities: translanguaging in a Tanzanian community library. *South. Afr. Ling. Appl. Lang. Stud.* 36 (2), 133–146.

- Shank Lauwo, M., 2019. Ubuntu translanguaging and social justice: negotiating power and identity through multilingual education in Tanzania. In: Avineri, N., Graham, L.R., Johnson, E., Riner, R., Rosa, J. (Eds.), *Language and Social Justice in Practice*. Routledge, New York, NY, pp. 88–96.
- Shank Lauwo, M., 2021. Literacies and translanguaging in Africa: a critical review of sociocultural perspectives. *South. Afr. Ling. Appl. Lang. Stud.* 39 (2).
- Sherris, A., Schaefer, P., Aworo, S.M., 2019. The paradox of translanguaging in Safaliba: a rural Indigenous Ghanaian language. In: Sherris, A., Adami, E. (Eds.), *Making Signs, Translanguaging Ethnographies: Exploring Urban, Rural and Educational Spaces*. Multilingual Matters, Bristol, pp. 152–169.
- Soja, E., 1996. *Thirdspace: Journeys to Los Angeles and Other Real-and-Imagined Places*. Blackwell, Oxford.
- Vertovec, S., 2007. Superdiversity and its implications. *Ethn. Racial Stud.* 30 (6), 1024–1054.
- Weedon, C., 1987. *Feminist Practice and Poststructuralist Theory*. Blackwell, Oxford.
- Zavala, V., 2019. Translanguaging pedagogies and power: a view from the South. *Lang. Educ.* 33 (2), 174–177.

Funds of knowledge and funds of identity: lenses for building language and literacy

María E. Fránquiz^a and María G. Leija^b, ^aDepartment Curriculum and Instruction, University of Texas at Austin, Austin, TX, United States; and ^bDepartment of Interdisciplinary Learning and Teaching, University of Texas at San Antonio, San Antonio, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	159
Defining a funds of knowledge approach	160
Research literature reviews on funds of knowledge	160
Research on funds of knowledge in subject area learning	161
Funds of knowledge in international contexts	162
Defining a funds of identity approach	162
Research on funds of identity in subject area learning	163
Funds of knowledge and funds of identity as theoretical lenses	164
Conclusion	165
Implications for future research and educator preparation	165
References	166

Introduction

In his seminal book, *Pedagogy of the Oppressed*, the Brazilian educator [Paulo Freire \(1970\)](#) argued that education is a political act where teachers and learners become aware of inequalities and take actions to change unjust conditions for the purpose of individual and collective liberation. For him critically conscious persons and communities “denounce the structures of oppression” and plan for a less unjust world ([Freire, 1993](#), p. 115)—one that does not subjugate but rather strives for humanization. In historically dehumanizing educational systems, oppressive structural policies include students segregated in academic tracks, separated by home and school language, taught Eurocentric world views and conventions, and exposed to a banking model of education with deposits of teacher knowledge held accountable to normalized and standardized criteria valued by dominant groups. These systems of oppression have existed in tandem with colonial expansion, settler colonialism of indigenous communities, gender binaries, and transatlantic slavery. In systems of dehumanization and oppression learners are led to accept racial and linguistic hierarchies and deficit views of specific languages and language speakers. Left untapped in oppressive systems are the underappreciated hidden literacies ([Martínez-Roldán and Fránquiz, 2009](#); [Villalva, 2006](#)) and meaningful activities that learners from subjugated communities, such as Latinx students in border communities, deploy in and outside classrooms.

The Funds of Knowledge theory and method ([González et al., 2005](#); [Moll et al., 1992](#)) counters systems of oppression and their deficit views of subjugated peoples. It does not accept segregated tracks, home-school separations, and the privileging of dominant norms in language and literacy education. Instead, the Funds of Knowledge approach advocates for understanding the languages and literacies of less known communities such as working-class migrant communities in order to integrate and build from these out of school funds in educational settings. The Funds of Knowledge approach emerged in the Southwestern region of the borderlands between the US and Mexico, “as part of a range of social justice efforts in education aimed at making the curriculum and learning more equitable and engaging for students from historically marginalized communities” ([Esteban-Guitart et al., 2019](#), p. 2). From this perspective, the Funds of Knowledge approach actively seeks to challenge deficit viewpoints by changing teacher – student relationships and minimizing the passive learning activities of banking education. Instead, many affordances for active learning in schools builds on students’ experiences from their homes and communities. Because schools too often attempt to build academic language and literacy in English to the exclusion of the powerful hidden literacies and translanguaging practices of youth, scholars argue for seeking, recognizing, and incorporating hidden funds of knowledge ([Martínez-Roldán and Fránquiz, 2009](#)) in school curricula to boost youths’ interest in furthering their education.

In the following sections we first define a Funds of Knowledge approach and describe ways the research literature has applied and amplified original explanations through studies conducted in the United States and abroad. We show how a Funds of Knowledge approach has been utilized for subject area learning in kindergarten through high school settings. These sections lead to studies that show ways a Funds of Identity approach was developed in order to address how learners absorb and create identity artifacts from their social worlds and their household’s Funds of Knowledge to actively define themselves in and across social environments. We take the reader through the definition of a Funds of Identity approach and describe ways the research literature has more recently applied and amplified this definition and its relationship to Funds of Knowledge through empirical studies. We conclude with implications for research and teacher educators who are interested and invested in appropriating the Funds of Knowledge and Funds

of Identity approaches in planning humanizing research (Paris and Winn, 2014) and humanizing pedagogies (Fránquíz, 2012; Fránquíz and Salazar, 2004; Salazar, 2013) to support the language and literacy development of all students.

Defining a funds of knowledge approach

A Funds of Knowledge approach seeks to make visible the bodies of knowledge including language that have strategic importance in students' households (Vélez-Ibáñez and Greenberg, 1992). More specifically, Funds of Knowledge is defined as "historically accumulated and culturally developed bodies of knowledge and skills essential for household and individual functioning and well-being" (Moll et al., 1992, p. 133). This approach "builds upon the sociocultural realities of students' lives, examines socio-historical and political dimensions of education, and casts students as critically engaged, active participants in the co-construction of knowledge" (Salazar, 2013, p. 128). Stated differently, McKinney (2017) argued that young learners from nondominant backgrounds ought to be viewed as resourceful and legitimate no matter where they live. Based on her work in South Africa, she proposes that language and literacy education in schools ought to treat "all students as creative multilingual and multi-dialectal language users, critically aware of the relationship between language and power" (McKinney, 2017, p. 139). Thus, living between linguistic borderlands ought to be considered in all humanizing educational projects.

In order to design more inclusive, affirming, and sustainable language and literacy instruction for multilingual and multidialectal children living in the transnational borderlands between the US and Mexico borders, a Funds of Knowledge framework was intentionally applied from the 1980s forward (González, 2006). From its conception, researchers and teachers valuing a Funds of Knowledge approach documented the skills and knowledges within local families with the underlying premise of building mutual trust, or *confianza*, between caregivers in students' homes and their teachers at schools. Rather than accept subtractive discourses and assimilationist practices historically promoted in many urban, suburban and rural schools across the United States (Arce, 2004; Fitts et al., 2008; Weisman and Hansen, 2008) the purpose of studying students' familial and local community Funds of Knowledge was for teachers to continually draw on these resources to advance emancipatory language and literacy practices (González et al., 2005; Rodríguez, 2013) in their classrooms. Empirical studies documenting such practices have appeared in research literature reviews.

Research literature reviews on funds of knowledge

There are at least five research literature reviews conducted with the Funds of Knowledge framework applied in K-12 education. Hogg (2011) identifies the sources from homes that align with school content. The Rodríguez (2013) review takes up notions of power and agency within pedagogical dimensions including the Community Cultural Wealth framework (Yosso, 2005). Llopert and Esteban-Guitart's (2017) literature review focuses on how students' formations of identities relate to Funds of Knowledge. Specifically, the two authors describe strategies and resources to contextualize the curriculum based on processes of students' identity work within the Funds of Knowledge approach. Continuing the work begun by Hogg (2011), Llopert and Esteban-Guitart (2018) also review the literature on Funds of Knowledge published from 2011 to 2015 in order to highlight what content, curricular subject, educational level, and the scope of countries in which empirical research was carried out. They also discuss some of the tensions in using the Funds of Knowledge approach as well as including issues involved in teacher education. More recently, the review by Denton and Borrego (2021) involved an extant body of research studies that included identifying, designing, teaching and learning in the STEM areas from a Funds of Knowledge approach. These reviews acknowledge Vélez-Ibáñez and Greenberg (1992) as originators of the Funds of Knowledge trailblazing concept. Their research on the economic system of interchange among Mexican households showcased strategies of working-class families not just for surviving, but achieving a strong measure of success in diverse occupational and social settings that was emulated in diverse educational contexts.

From the literature reviews the widespread application of the Funds of Knowledge theoretical approach and empirical work, counters deficit views of subjugated families and elevates the strengths and resources within households. Initial findings from the range of studies on family and student's everyday experiences at home showed language and literacy practices related to doing farming and animal management, construction and building, figuring out finances, attending to medicine and spiritual wellbeing, among other practices. Identification of these practices prompted a number of literacy and teaching curricular projects across grades in schools. Notably, these initial studies of identifying practices in family households involved teachers and researchers working collaboratively. They examined, prepared, conducted, reflected, analyzed, and reimagined together the ways practices at home can become resources for teaching and learning. As noted by Rodríguez (2013) in her literature review, "rather than allow the assumption of 'cultural deficits' to persist in the educational arena, the initiating researchers emphasize the presence of knowledge, skills, and strategies among students that were produced in settings beyond the school – and therefore, beyond the immediate view (and appreciation) of their teachers" (p. 90).

More recent research on Funds of Knowledge argues that it is also necessary to "acknowledge and understand, rather than avoid, the dark funds of knowledge, the consequences for students of living in difficult circumstances" (Moll, 2019, p. 135) that are also present in households. Thus, the positive benefits of strengthening relationships using the Funds of Knowledge approach among students, peers, families, and teachers ought to also consider the dangers of presenting a single story (Adichie, 2009). As stated in their literature review (2021), "[u]nlike many other education approaches, funds of knowledge curricula cannot be adopted as easily in other settings once developed." (Denton and Borrego, 2021, p. 85). This caution is also succinctly noted by González

(2016) who asked, “[h]ow do we account for diversity without falling into essentializing discourses about students, communities, and families?” (p. 73). Keeping the importance of local and contextual factors in mind, many studies in a variety of subject areas and grades provide examples of ways Funds of Knowledge can better engage students’ home languages, cultures, and experiences in school subject area learning.

Research on funds of knowledge in subject area learning

Scholars have documented ways in which teachers embrace Funds of Knowledge as an assets-based approach. They show how students and their families become active participants of curricula taught in ways that are culturally and/or linguistically relevant for students (Borgerding, 2016; Coles-Ritchie et al., 2015; Cribbs and Linder, 2013; Tan and Calabrese Barton, 2010; Walkington, 2017). With the premise that Funds of Knowledge activates language and literacy engagement and learning, some studies have explored traditional subject area topics such as natural science with a broader inclusion of scientific thinking from non-Western points of view. One such project was the Mien-American garden cultivated with Mien families in Sacramento, California (Hammond, 2001). In another gardening project, math and science lessons were integrated. This plan called for students to calculate mathematical values such as the distance between plants, daily volume of water for plants to thrive, and the perimeter of the gardening project (Diaz and Bussert-Webb, 2017). When such gardening projects are coupled with the concept of food deserts, critical engagement with science, math, and nutrition can lead to personal and/or collective action in a specific community. In the Hammond study, findings also revealed that “a new kind of multi-science can emerge, one accessible to all collaborating members and responsive to school standards” (2001, p. 983).

Elementary classroom teachers have also explored the Funds of Knowledge theory as a means to improve their pedagogy and praxis. For example, Amanti (2005), an elementary teacher and master’s student, documented her role as a teacher-researcher while she engaged in a Funds of Knowledge project with families who had horses. Amanti visited students’ homes to learn about the students’ and families’ experiences with horses at their Arizona residence or in Mexico. During her observations, she learned about the different types of feed for the horses and how to care for them. She even learned about racing horses. When Amanti identified the vast amount of Funds of Knowledge her students’ and their families had about horses, she developed a curricular unit about horses that extended students’ learning new knowledge about social studies (e.g. history), language arts (e.g. vocabulary, stories), math (e.g. measurements), and science (e.g. anatomy of a horse). Amanti observed the impact the Funds of Knowledge project had on her pedagogy and her classroom community. She noted that her teaching shifted from a teacher-directed to student-centered pedagogical approach. She accomplished this change by inviting students’ and families’ voices (parents as expert guest speakers), skills, and lived experiences into her curricular plan.

Middle school teachers have also embraced and incorporated Funds of Knowledge as an approach that aides in the scaffolding of content areas, such as science. For example, Tan and Calabrese Barton (2010) examined the pedagogical strategies of Mr. M., a white, male, sixth grade science teacher at a middle school. The teacher created formal spaces that created opportunities for his minoritized students of color to make connections with their Funds of Knowledge and the science content, specifically a unit on nutrition. During the enactment of lessons, students worked collaboratively to make nutritious food and to provide information about the nutritious benefits of a meal, beverage, or snack they provided to their peers and guests. During the unit students were invited to share their stories from kitchens at home and experiences in markets to make relevant connections with scientific vocabulary and principles. Also, parents were invited on field trips. With Mr. M.’s interest in raising the authority of familiar Funds of Knowledge in his science curriculum the disciplinary vocabulary, principles, and locations of scientific knowledge were normalized for all. Contrary to some deficit thinking about parents of minoritized students and their families, the parents in this study took up the invitation to engage in learning with their child’s teacher and peers.

High school teachers have also drawn on students’, parents’, and communal Funds of Knowledge as a method for improving their praxis in content areas. For example, Borgerding (2016) examines how a teacher explored with his predominantly white students in a rural high school the concept of evolution in their biology class. Findings show how this male teacher from the community and his students had largely positive evolution learning experiences even as some students came from homes that reject evolution on religious grounds. The teacher strategically positioned himself as not personally representing science. Instead, he consistently took the position “according to scientists (pg. 72).” Such instruction honored local funds of knowledge with respect to local knowledge of nature and by treating students’ religious knowledge as a form of local expertise. Additionally, the teacher served as a border crossing “tour guide” by helping students identify how the culture of science and the culture of their lifeworld differed with respect to evolutionary theory. Students negotiated the cultural borders for learning evolution and respected the teacher’s sensitivity toward multiple positions. In this way students examined diverse ways of critically thinking about the discourse and premises of science.

Environmental education and wildlife conservation has also been the context for Funds of Knowledge research. An example is Arreguín-Anderson and Kennedy (2013) study which documented how Latinx culture and the Spanish language was intentionally used to teach the characteristics of animals and plants of the various regions in Texas. Capitalizing on the ways of knowing generated in students’ homes, teachers are encouraged to use visuals, songs, physical activities, and home remedies for lesson planning and for identifying common cognates between English and Spanish. These cognates may include *predador* (predator), *población* (population), and *ecosistema* (ecosystem). Adopting the view of language as a resource for learning (Ruíz, 1984), the Next Generation Science Standards advocate for making science real and relevant to the lives and future careers of all students (Lee et al., 2014)

by specifically capitalizing on their Funds of Knowledge (Lee et al., 2019). This particular study set out to accomplish these goals in a professional development workshop for teachers conducted in Spanish.

A Funds of Knowledge approach is also important in social studies education where issues such as immigration and citizenship are vital to teach and critical for students to understand. Zipin (2009) argued that students and their families do not always have positive, or “light” Funds of Knowledge for development in school curricula. Rather, he argues students may have “dark” experiences of violence, racism, poverty and abuse. The question is, will teachers stick to the academic work that neatly aligns with positive family practices and avoid the emotional work involved in transforming “dark” Funds of Knowledge? Social studies experts acknowledge the importance of building on comfortable and uncomfortable experiences in students’ lives. Naming these experiences is the issue at hand. Because “dark” historically has been used in racist discourses about Black and Brown people and their bodies (Dumas, 2016), researchers such as Becker (2014) proposed the concept of Funds of (difficult) Knowledge to account for difficult experiences such as violence in family migration stories. Rodríguez and Salinas (2019) used testimonios of Latinx teachers and students that recounted traumatic experiences such as deportation to describe how “dark” funds are complex and difficult Funds of Knowledge.

For Gallo and Link (2015) the issue of documentation, undocumented status, and belonging in mixed status families with the accompanying personal and collective fear of deportation is less about “dark” or difficult, and more about the agentic possibilities, or “literacy for action” (González and Moll, 2002, p. 637), offered by consciously embracing Politicized Funds of Knowledge. In their work Gallo and Link (2015) state Politicized Funds of Knowledge are “the experiences, knowledges, and skills young people deploy and develop across learning contexts” (p. 357). These are the types of experiences with potential for classroom communities to understand the root causes of oppression. In the United States the social studies is one of the subject areas that can provide liberation to minds, hearts, and spirits rather than ignore or discount injustices in families and communities, past and/or present. In other countries various Funds of Knowledge approaches are also important.

Funds of knowledge in international contexts

A plethora of international contexts are taken up by scholars interested in the Funds of Knowledge approach in their educational research. Examples are scholars who have conducted studies in New Zealand (Cooper and Hedges, 2014; Hedges et al., 2011), Australia (Mantei and Kervin, 2014; Zipin, 2013; Zipin et al., 2012), the United Kingdom (Cremin et al., 2012; Hughes and Pollard, 2006), Spain (Jovés et al., 2015; Saubich and Esteban, 2011); (Vila and Siqués, 2013), and Uganda (Kendrick and Kakuru, 2012). These research studies involve educators from early childhood settings to child-run homes. Although each country has its own unique context, there are many overlapping similarities, such as the changing demographics of student populations due to poverty, mobility, health, and safety concerns.

Moll (2019) specifically documented how Funds of Knowledge have been applied in countries such as New Zealand, Australia, Spain, and Uganda to meet the needs of their dynamically changing sociocultural context. Noted by Moll (2019) is a New Zealand study (Hedges and Jones, 2012) where young children’s working theories are mediating products of their own agency and of their relationships with peers and siblings. As a receiving community of many immigrant families, Catalunya, Spain, has felt the impact of diverse newcomers in its schools. Rather than emphasizing ethnographic perspectives, the research approach of Saubich and Esteban (2011) is highlighted by Moll (2019) because the researchers relied on administering tasks to family members, such as self-portraits, assessment of family artifacts, taking photos of family routines, etc. Rather than the traditional ethnographic approach of Funds of Knowledge, this study also concentrated on the formation of identity and how families come to understand themselves.

In Australia, Zipin et al. (2012) are cited by Moll (2019) as applying the Funds of Knowledge framework in communities that are “undergoing complex unsettlements” rather than more cohesive or settled communities. Central to this work are students becoming researchers of their own lives no matter how dark their past or present circumstances may be. Moll (2019) also noted how the health epidemic of HIV/AIDS that ravaged regions such as Uganda left orphaned children generating Funds of Knowledge for survival and protection. As such, children living in child-headed families were observed using language practices such as stories, songs, riddles, proverbs, and translanguaging between Uganda and English as resources for further education and survival (Kendrick and Kakuru, 2012). With a broader understanding of ways that the Funds of Knowledge theory can have lasting effects on the identities of learners the Funds of Identity offers a new theoretical construct.

Defining a funds of identity approach

Until recently emphasis on family Funds of Knowledge did not address the individual Funds of Identity (Esteban-Guitart, 2016; Esteban-Guitart and Moll, 2014a) from daily lived interests, experiences with peers, membership in communities, and participation in popular culture. Esteban-Guitart (2014) notes, “Funds of knowledge...become funds of identity when the participants appropriate them and use them to define themselves” (p. 753). Accordingly, he defines the term Funds of Identity, as representing an accumulated set of historically and culturally developed tools essential for constructing one’s identity and defining individuality (Esteban-Guitart, 2016). As a theoretical construct, Funds of Identity refers to artifacts (written, spoken, visual, musical, or combinations in multimodal form) that people absorb and create from their social worlds and their household’s Funds of Knowledge to actively define themselves in and across social environments. With this approach researchers observe how children create their own Funds of Identity from being with peers, problem-solving with them, and experiencing diverse social contexts with adults and other

children like and unlike them. The challenge as pointed out by Esteban-Guitart and Moll (2014), is “to recognize and maintain optimal identities for learning (Coll and Falsafi, 2010)” (p. 44).

Originally, Esteban-Guitart (2012) identified five types of Funds of Identity: “Geographical Funds of Identity” can include an area, landscape, river, mountain, city, country, etc. (Jovés et al., 2015; Saubich and Esteban, 2011); “Practical Funds of Identity” include important activities for the individual such as work, music, hobby, sport, etc. (Marsh and Zhulamanova, 2017; Moulton, 2018); “Cultural Funds of Identity” includes artifacts, flags, religious symbols, etc. (Nagle and Stooke, 2016; Poole, 2017); “Social Funds of Identity” refers to significant people, friends, etc. (Nagle and Stooke, 2016; Saubich and Esteban, 2011), and “Institutional Funds of Identity” refers to social institutions family, marriage, church affiliation, etc. (Charteris et al., 2018; Marsh and Zhulamanova, 2017). Since then, additional contributions to the Funds of Identity approach have emerged in the literature: “Digital Funds of Identity” such as creating a blog, avatar, videos, music, artwork, etc. (González-Patiño and Esteban-Guitart, 2015; Poole, 2017); “Popular Culture Funds of Identity” refers to music, shows, comics, toys, dance styles, etc. (Nagle and Stooke, 2016; Ordoñez et al., 2018); and Poole (2017) identified “Existential Funds of Identity” as negative experiences and emotions (Poole and Huang, 2018) and “Ideological Funds of Identity” as “political, cultural or religious beliefs” (Poole, 2017, p. 8), as sub-divisions of Cultural Funds of Identity.

One systemic literature review examined 24 studies and encompassed conceptual and empirical studies within the United States and internationally (Hogg and Volman, 2020). This literature review found that researchers utilized two methods for participants to identify their Funds of Identity. One method focused on the creation of a narrative or an object and the other method focused on gathering objects. The following are examples of the methods used to identify Funds of Identity: self-portrait and an explanation of the portrait (Saubich and Esteban, 2011), significant circle (Moulton, 2018), poems (Subero, 2015), diary entries (Esteban-Guitart, 2012), taking photographs (Marsh and Zhulamanova, 2017), videos (Esteban-Guitart and Moll, 2014b), maps (Moulton, 2018), collages (Subero et al., 2018), “self-portrait, multilingual texts, shoeboxes used to bring into school meaningful objects from home” (Subero et al., 2017, p. 253), discussions (Esteban-Guitart et al., 2019), guided reflection activities (Charteris et al., 2018; Recchia and McDevitt, 2018), and developing digital literacy media (Nagle and Stooke, 2016; Poole and Huang, 2018). Overall, the reviewers (Hogg and Volman, 2020) found that the studies documented effective strategies for teachers to learn about students’ Funds of Identity. They suggested that future research ought to examine how the theory supports the academic development (outcomes) of students when examining their Funds of Identity.

Research on funds of identity in subject area learning

Some scholars have utilized Funds of Identity as the theoretical framework to examine subject area learning (Esteban-Guitart et al., 2019; Hachey et al., 2021; Nagle and Stooke, 2016; Poole, 2017). For example, Hachey et al. (2021) worked with a kindergarten teacher to develop “making-based curriculum” that drew on Makerspace Pedagogy to examine students’ STEM academic identity. The team selected a picture book, *Creepy Carrots* (Reynolds, 2014), that they identified as drawing on “potential funds of knowledge” of the rural students. The book is about a bunny who eats too many carrots and later imagines they are following him. To contain the carrots in the garden, the bunny builds a wall around the garden and a moat with alligators to keep the carrots contained. After hearing the story, students were asked to take on the role of the rabbit and problem solve with their own solutions. The researchers found that there were differences in the varied gendered groups (girls, boys, and a heterogeneous groups). The all boys and all girls groups’ artifacts “show(ed) more stereotypical characteristics associated with their respective gender” (Hachey et al., 2021, p. 476). The authors highlight the lack of STEM education within the early grades and the importance of incorporating hands-on learning to develop funds of identity in the STEM areas.

Within middle school contexts, scholars have explored students’ Funds of Identity (Moulton, 2018; Nagle and Stooke, 2016). In the research study by Nagle and Stooke (2016) four-seventh and eighth grade students’ created history projects in digital forms over a two-week period. The history projects were based on students’ personal interests in Canada. Two participants created a slideshow of the “...development of an intercolonial railway in the mid-1800s...” and the two other participants developed a website that highlighted opposing British and American views (Nagle and Stooke, p. 160). Through the history project, students drew on their Funds of Identity: Practical Funds of Identity (music and video games); Cultural Funds of Identity (Irish and Scottish family background); Social Funds of Identity (student-teacher relationship and student-to-student friendships); Popular Funds of Identity (Che Guevara and Rage Against The Machine); and Digital Funds of Identity (website, slideshow, and videos). Researchers question the role that teachers need to play in supporting their students in developing multimodal texts (website and slideshow) and the importance of creating more opportunities for them to draw on their Digital Funds of Identity. Findings highlight the importance of providing students with sufficient time and space to socialize while working on creating their multimodal texts.

Scholars have also examined students’ Funds of Identity within high school contexts (Poole, 2017, 2021; Poole and Huang, 2018). For example, Poole (2017), as a researcher-participant, examined 31 10th grade Han Chinese high school students’ Funds of Identity through a two-week English unit. The unit explored identity, The Avatar Project, at an internationalized high school located in Shanghai, China. Student constructed word clouds, avatars, and written reflections were analyzed for Funds of Identity. Poole found that Chinese students revealed several types of Funds of Identity: Geographical Funds of Identity (China, Shanghai, and students’ towns of origin); Social Funds of Identity (parents, grandparents, teenager, cousins, teachers, friends, and pop stars); Institutional Funds of Identity (family and school); Practical Funds of Identity (digital, software, applications, computer games, music, and sport); Cultural Funds of Identity (local dialects, Filial piety, panda, and Chinese flag); Ideological Funds of Identity (patriotism, desire to study abroad and give back to the country, and Christianity); and Existential Funds of Identity (low-self-esteem,

identity issues, peer pressure, and death). Poole noted the importance of complementing digital strategies, such as creating word clouds and avatars, with written reflections to better analyze and interpret students' Funds of Identity. He also cautioned that time and experience with multiple software can impact the avatar students create. The key is to avoid circumstances in which students settle for an avatar that is not representative of their Funds of Identity. Another key finding is the role of the teacher in facilitating a project such as the one Poole completed with his students. He warns that the role of teacher can shift to spectator rather than facilitator, which may be out of the ordinary for some teachers. In addition, through the project, he learned about the "...psychological and affective aspects of (some of his students') development..." (Poole, 2017, p. 16). As a reflective practitioner he questioned what to do in his role as a teacher with this new knowledge. As the 21st century continues to unfold, ways for teachers to uncover and become responsive to new knowledge through the theoretical lenses of Funds of Knowledge and Funds of Identity is necessary.

Funds of knowledge and funds of identity as theoretical lenses

Several researchers have examined empirical work through the use of Funds of Knowledge and Funds of Identity as a combined theoretical framework (Charteris et al., 2018; Jovés et al., 2015; Recchia and McDevitt, 2018). For example, Jovés et al. (2015) examined research that emerged from a study group comprised of eight in-service teachers and university professors in Spain. The professors supported the teachers in learning how to examine the Funds of Knowledge and Funds of Identity of one student within each of the teacher's classrooms. In their article, Jovés and colleagues examine a second-grade teacher's experience in integrating both approaches with a Moroccan family.

During the teacher's home visits, she interviewed the parents and examined the funds of identity of her student by asking him to create an identity drawing and a significant circle. The Funds of Knowledge data revealed the family's knowledge about agriculture (caring for animals on the farm), religion (Koran), languages (Berber, Arabic, and Spanish), home repairs (household appliances), and cooking (bread, biscuits, Moroccan, and Mediterranean food). The focal student's Funds of Identity consisted of: Geographical Funds of Identity (Nador, Morocco and Catalonia, Spain), Social Funds of Identity (family, friends, relatives, and neighbors), Cultural Funds of Identity (Arabic, Berber, Catalan, and Spanish languages), Institutional Funds of Identity (school, family, and the Mosque), and Practical Funds of Identity (cycling and going to the farm with his dad, learning Spanish at school and Arabic at the Mosque, and making cookies with his mother). Through data analysis, the teacher identified and developed a unit on farm animals since the topic emerged in the interviews (Funds of Knowledge) and the student's artifacts (Funds of Identity).

The farm animal unit covered an array of content areas: science (types of animals, characteristics of animals, and classification of animals), math (measurement of ingredients used in a recipe), language arts (summarizing their farm field trip experience), creating a multimodal text (poster) that incorporated pictures taken during the field trip to the farm, reading texts, searching online for animal riddles, writing a biscuit recipe, and translating the recipe through the support of their family, and art (creation of a mural) as a culminating event, a visual summary of the unit). The authors, Jovés et al. (2015), note the vital role that the parents' Funds of Knowledge and the focal student's Funds of Identity played in fostering and enhancing the educational farm unit the teacher developed.

Involving students in the process of designing classroom activities based on their own interests and Funds of Knowledge (González et al., 1995) has been observed as affecting their identity investment in classroom practices (Norton, 1995). In the United Kingdom, Andrews and Yee (2006) studied two Muslim primary school-aged children engaged in out-of-school learning. Although they engaged in many similar activities at home and in their communities, the significance of each child's investment in the activities was different. In New Zealand children's Funds of Knowledge were observed as working theories based on interests and inquiries with peers and siblings during play activities (Hedges et al., 2011; Hedges and Jones, 2012). These studies show that positive identities as learners can emerge and grow in out-of-school settings, and in children's play. As such, children's Funds of Knowledge and interests play an important role in their developing Funds of Identity. However, Funds of Identity are conceptualized beyond play or children's interests on "activities that are geographical (spaces), social (people), cultural (artifacts) or institutional (locations) and that each of us incorporate into the vision we have of ourselves and of what, for each of us, is significant and important" (Llopert and Esteban-Guitart, 2018, pp. 155–156). More specifically, Funds of Identity are understood to comprise of significant people such as grandparents, spaces such as a lake, institutions such as the Muslim religion, or artifacts such as a musical instrument – all that best defines the unique emerging self in each student.

Esteban-Guitart et al. (2019) also explored Funds of Knowledge and Funds of Identity in two different school contexts in Catalonia, Spain. The two participatory action research projects involved at least 30% students of foreign origin assisting state-run schools and within the boundaries of the city of Barcelona. In one context, Funds of Knowledge were examined with a group of seven teachers interested in implementing the approach to improve the curriculum. After learning about the approach, the teachers explored students' Funds of Knowledge by carrying out five home visits with the families they identified. Interviews were developed collectively and meaningful circles plus artist's books acknowledged as important for gathering comprehensive information on families and students. After the interviews, one significant learning of the Funds of Knowledge approach was the importance of peanuts in the kitchens of several families so they worked collaboratively to implementing a peanut activity with their five- and six-year-old students. As a result, the natural sciences were explored using various languages (Spanish, French, Chinese, etc.) present in the children's families. Another significant finding, from the Funds of Identity approach with students, was creating space for direct dialog about racism and providing opportunities for writing and acting out through theater, the labels imposed on children by others in their social world. The study recommends that sustainability in an increasingly multiracial, multilingual, and

multicultural world goes beyond culturally sustaining pedagogy to uncover and uphold the meaningful practices that guarantee the survival of different experiences, practices, and life contexts. This means “going beyond the typical assimilationist values and practices prevalent in schools, in which the curriculum, instruction and teaching practices are often disconnected from the meaningful experiences, life contexts and practices of learners” (Esteban-Guitart et al., 2019, p. 9). The two projects in Barcelona illustrate strategies that are sensitive to the resources of families as well as children’s own areas of sense making. The researchers offer a provocative thought for amplifying culturally sustaining pedagogies, “we believe that a pedagogy of cultural and identity sustainability should normalize diversity (incorporating it into school practice), as well as diversify normality” (Esteban-Guitart et al., 2019, p. 10).

Like Hidden Funds of Knowledge (Martínez-Roldán and Fránquiz, 2009; Moll, 2015) from home and community that are unseen, ignored and neglected in schools the Invisible Funds of Identity that develop tacitly in learners are also critical to address. Accordingly, the purpose of education from an Invisible Funds of Identity framework is “[T]o make explicit the impact of negative hegemonic-dominant discourse and practices such as racism, discrimination, or colonialism” (Esteban-Guitart, 2021, p. 12). These dominant discourses are “produced and distributed historically, culturally and socially and end up tacitly embedded in people’s behaviors, thoughts, and practices even though they may be unaware of them” (Esteban-Guitart, 2021, p. 12). Such a perspective requires teachers to build curricula for students that uncovers not only the Hidden Funds of Knowledge from outside of school contexts but also requires the dismantling of the dominant discourses and practices that undergird the identities and identity productions of students. The theory of Funds of Identity, then, works within the framework of the Funds of Knowledge with the goal of enacting social justice pedagogies and humanizing pedagogies (Fránquiz and Salazar, 2004; Salazar, 2013) in classrooms. The goal is to uncover deficit thinking in/from people, institutional structures and legacies of colonization. In super diverse classroom activities, these Invisible Funds of Identity can present themselves in and through the creation of identity artifacts such as drawings, photos, avatars, or literacy journey boxes (Fránquiz et al., 2011).

Conclusion

As active participants and agents in familial, school, community, and digital contexts children acquire a variety of linguistic and cultural practices that build on and build up their Funds of Knowledge and Funds of Identity. Esteban-Guitart (2021) explains that “It is important to emphasize here that *all* funds of identity are funds of knowledge, but not all funds of knowledge are funds of identity” (pg. 4, emphasis in original). While the former emphasizes cultural and intellectual resources that households use for their welfare, the latter emphasizes significant people, cultural artifacts, spaces, passions, and interests within a learner’s definition of self. The strategy for building on students’ Funds of Knowledge requires home visits and partnerships with caregivers, their practices, and their extended network for health and well-being. The strategy for building up Funds of Identity requires opportunities to create and share identity artifacts in school and after school contexts. The cautionary considerations for appropriating Funds of Knowledge and Funds of Identity for humanizing pedagogical purposes is not just to validate funds, but to include them meaningfully in identifying social problems, reflecting individually and collectively about necessary changes, and taking actions toward transformation. Actions may be identity products representing the self, such as creating posters, speeches, murals, petitions, podcasts, internet posts, etc.

As the Funds of Identity framework develops, there has been limited research on ways to respond to the urgent need for the preparation and sustainability of career educators who are capable of and committed to advancing system-wide equity and access for students from culturally, economically, and linguistically diverse backgrounds. More work is needed to understand how gender and status would affect experiences that influence the development of Funds of Identity and how it is affirmatively cultivated. Noteworthy would be the influence of male teachers of color on students in spaces such as elementary education or women in STEM fields and their influence in historically male dominated learning spaces. The influence of indigenous world views on the Funds of Knowledge and the Funds of Identity in schools for indigenous children or in schools for all learners would impact the degree of educational development afforded to future generations.

Implications for future research and educator preparation

Given the feminization of the teaching profession and the scarcity of male teachers, how might a Funds of Knowledge and Funds of Identity research approach benefit families and their children? Recent studies on teachers of color take into account the hyperawareness and differential intersecting identities of male teachers (e.g., Lynn, 2006; Bristol and Goings, 2019; Brown, 2012). In the case of Latinx male teachers, they share common characteristics with the public narrative of whitemale (Grande, 2003; Urrieta, 2009) male teachers, yet these men live in a parallel world with intersecting identities that do not necessarily obtain the benefits of hegemonic masculinity and male privilege (Lara and Fránquiz, 2015). While intersectionality is prevalent in education research for its capacity to illuminate how identities like race, gender, ability, and class interact and shape individual experiences (Sibbett, 2020), a Funds of Knowledge and Funds of Identity framework may provide a complimentary way for understanding the assets that the underrepresented population of male teachers generally and male teacher of color in particular bring to the profession. Such understanding may reveal how pre- and post-professional development for male teachers ought to be reimagined and what matters in the sustainability of their Funds of Identity.

Future research should examine: To what degree does ancestral knowledge and family history matter in educator preparation?, To what degree does teacher’s intersectional story matter? and Does providing teachers opportunities to reveal their Funds of

Knowledge and Funds of Identity through identity artifacts influence their persistence in the teaching profession? Researchers are encouraged to conduct studies about ways to validate and tap the rich expertise that all teachers generally and male teachers particularly bring to K-12 classroom and out of school educational settings.

As teacher education programs take up the call for preparing aspirantes/future teachers for highly diverse classrooms, intentional attention must be drawn to their Funds of Knowledge (González et al., 2005; Moll et al., 1992) and their developing Funds of Identity (Esteban-Guitart, 2016; Esteban-Guitart and Moll, 2014a). Providing safe spaces to unpack racial experiences for preservice teachers are necessary. Because the profession is still predominantly white and female (Kohli, 2019; National Center for Education Statistics, 2018; Sleeter, 2017), one way to experience anti-racist education is by providing environments where students will not be minoritized, such as affinity groups (Bell, 2015). Racial affinity groups are self-selected groups made up of people who identify with similar racial or ethnic backgrounds. Race affinity groups can create an environment where preservice teachers interpret, critically reflect, find fellowship, identify personal biases, and become more aware of cultural responsiveness (Douglas, 2008). Affinity groups as a tool for providing a safe space where all preservice teachers can share and unpack their racialized and gendered experiences holds great promise for the better preparation of teachers for an increasingly diverse student population.

References

- Adichie, C.N., 2009. The Danger of a Single Story. Video file. TED Talk. Available from: https://www.ted.com/talks/chimamanda_ngozi_adichie_the_danger_of_a_single_story.
- Amanti, C., 2005. Beyond a beads and feathers approach. In: Gonzalez, N., Moll, L.C., Amanti, C. (Eds.), *Funds of Knowledge: Theorizing Practices in Households, Communities and Classrooms*. Erlbaum, Mahwah, NJ, pp. 131–141.
- Andrews, J., Yee, W.C., 2006. Children's 'funds of knowledge' and their real life activities: two minority ethnic children learning in out-of-school contexts in the UK. *Educ. Rev.* 58 (4), 435–449.
- Arce, J., 2004. Latino bilingual teachers: the struggle to sustain an emancipatory pedagogy in public schools. *Int. J. Qual. Stud. Educ.* 17 (2), 227–246.
- Arreguin-Anderson, M.G., Kennedy, K.D., 2013. Deliberate language planning in environmental education: a CRT/LatCrit perspective. *J. Environ. Educ.* 44 (1), 1–15.
- Becker, A., 2014. Funds of (difficult) knowledge and the affordances of multimodality: the case of Victor. *J. Lang. Literacy Educ.* 10 (2), 17–33.
- Bell, M.K., 2015. Making Space: Teaching Tolerance. Available from: <https://www.learningforjustice.org/magazine/summer-2015/making-space>. (Accessed 1 February 2022).
- Borgerding, L.A., 2016. High school biology evolution learning experiences in a rural context: a case of and for cultural border crossing. *Cult. Stud. Sci. Educ.* 12 (1), 53–79.
- Bristol, T.J., Goings, R.B., 2019. Exploring the boundary-heightening experiences of Black male teachers: lessons for teacher education programs. *J. Teach. Educ.* 70 (1), 51–64.
- Brown, A.L., 2012. On human kinds and role models: a critical discussion about the African American male teacher. *Educ. Stud.* 48 (3), 296–315.
- Charteris, J., Thomas, E., Masters, Y., 2018. Funds of identity in education: acknowledging the life experiences of first year tertiary students. *Teach. Educat.* 53 (1), 6–20.
- Coles-Ritchie, M., Monson, B., Moses, C., 2015. Drawing on dynamic local knowledge through student-generated photography. *Equity Excell. Educ.* 48 (2), 266–282.
- Coll, C., Falsafi, L., 2010. Learner identity. An educational and analytical tool. *Rev. Educ.* 353, 211–233.
- Cooper, M., Hedges, H., 2014. Beyond participation: what we learned from Hunter about collaboration with Pasifika children and families. *Contemp. Issues Early Child.* 15 (2), 165–175.
- Cremin, T., Mottram, M., Collins, F., Powell, S., Drury, R., 2012. Building communities: teachers researching literacy lives. *Improv. Sch.* 15 (2), 101–115.
- Cribbs, J.D., Linder, S.M., 2013. Teacher practices and hybrid space in a fifth-grade mathematics classroom. *Math. Educat.* 22 (2), 55–81.
- Denton, M., Borrego, M., 2021. Funds of knowledge in STEM education: a scoping review. *Stud. Eng. Educ.* 1 (2), 71–92.
- Diaz, M.E., Bussert-Webb, K., 2017. Latino youth's out-of-school math and science experiences: impact on teacher candidates. *International Journal of Research in Education and Science* 3, 624–635.
- Douglas, P.H., 2008. Affinity groups: catalyst for inclusive organizations. *Employ. Relat. Today* 34 (4), 11–18.
- Dumas, M.J., 2016. Against the dark: antiblackness in education policy and discourse. *Theory Into Pract.* 55 (1), 11–19.
- Esteban-Guitart, M., 2012. Towards a multimethodological approach to identification of funds of identity, small stories and master narratives. *Narrat. Inq.* 22 (1), 173–180.
- Esteban-Guitart, M., 2014. Funds of identity. In: Teo, T. (Ed.), *Encyclopedia of Critical Psychology*. Springer, New York, NY.
- Esteban-Guitart, M., 2016. Funds of Identity: Connecting Meaningful Learning Experiences in and Out of School. Cambridge University Press, New York, NY.
- Esteban-Guitart, M., 2021. Invisible funds of identity in urban contexts. *Urban Educ.* 1–21.
- Esteban-Guitart, M., Lalueza, J.L., Zhang-Yu, C., Llopert, M., 2019. Sustaining students' cultures and identities. A qualitative study based on the funds of knowledge and identity approaches. *Sustainability* 11 (12).
- Esteban-Guitart, M., Moll, L.C., 2014a. Funds of identity: a new concept based on the funds of knowledge approach. *Cult. Psychol.* 20 (1), 31–48.
- Esteban-Guitart, M., Moll, L.C., 2014b. Lived experience, funds of identity and education. *Cult. Psychol.* 20 (1), 70–81.
- Fitts, S., Winstead, L., Weisman, E., Flores, S., Valenciana, C., 2008. Coming to voice: preparing bilingual-bicultural teachers for social justice. *Equity Excell. Educ.* 41 (3), 357–371.
- Fránquiz, M.E., 2012. Key concepts in bilingual education: identity texts, cultural citizenship, and humanizing pedagogy. *N. Engl. Read. Assoc. J.* 48 (1), 32.
- Fránquiz, M.E., Salazar, M., 2004. The transformative potential of humanizing pedagogy: addressing the diverse needs of Chicano/Mexicano students. *High Sch. J.* 36–53.
- Fránquiz, M.E., Salazar, M., DeNicolò, C.P., 2011. Challenging majoritarian tales: portraits of bilingual teachers deconstructing deficit views of bilingual learners. *Biling. Res. J.* 34 (3), 279–300.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum.
- Freire, P., 1993. *Pedagogy of the City*. Continuum, New York, NY.
- Gallo, S., Link, H., 2015. "Diles la verdad": deportation policies, politicized funds of knowledge, and schooling in middle childhood. *Harv. Educ. Rev.* 85 (3), 357–382.
- González, N., 2006. *I Am My Language: Discourses of Women and Children in the Borderlands*. University of Arizona Press, Tucson, AZ.
- González, N., 2016. Imagining literacy equity: theorizing flows of community practices. *Lit. Research* 65 (1), 69–93.
- González, N., Moll, L.C., 2002. Cruzando el puente: building bridges to funds of knowledge. *Educ. Pol.* 16 (4), 623–641.
- González, N., Moll, L.C., Amanti, C., 2005. *Funds of Knowledge Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ.
- González, N., Moll, L.C., Tenery, M.F., Rivera, A., Rendon, P., Gonzales, R., Amanti, C., 1995. Funds of knowledge for teaching in Latino households. *Urban Educ.* 29 (4), 443–470.
- González-Patiño, J., Esteban-Guitart, M., 2015. Fondos digitales de conocimiento e identidad: un análisis etnográfico y visual. *Papeles Trab. Cult. Educ. Desarro. Hum.* 11 (2), 20–25.
- Grande, S., 2003. Whiteman feminism and the colonialist project: a review of contemporary feminist pedagogy and praxis. *Educ. Theor.* 53 (3), 329–346.
- Hachey, A.C., An, S.A., Golding, D.E., 2021. Nurturing kindergarteners' early STEM academic identity through makerspace pedagogy. *Early Child. Educ. J.* 50 (3), 469–479.
- Hammond, L., 2001. Notes from California: an anthropological approach to urban science education for language minority families. *J. Res. Sci. Teach.* 38 (9), 983–999.
- Hedges, H., Cullen, J., Jordan, B., 2011. Early years curriculum: funds of knowledge as a conceptual framework for children's interests. *J. Curric. Stud.* 43 (2), 185–205.
- Hedges, H., Jones, S., 2012. Children's working theories: the neglected sibling of Te Whariki's learning outcomes. *Early Child. Folio* 16 (1), 34–39.

- Hogg, L., 2011. Funds of knowledge: an investigation of coherence within the literature. *Teach. Teach. Educ.* 27 (3), 666–677.
- Hogg, L., Volman, M., 2020. A synthesis of funds of identity research: purposes, tools, pedagogical approaches, and outcomes. *Rev. Educ. Res.* 90 (6), 862–895.
- Hughes, M., Pollard, A., 2006. Home-school knowledge exchange in context. *Educ. Rev.* 58 (4), 385–395.
- Jovés, P., Siqués, C., Esteban-Guitart, M., 2015. The incorporation of funds of knowledge and funds of identity of students and their families into educational practice. A case study from Catalonia, Spain. *Teach. Teach. Educ.* 49, 68–77.
- Kendrick, M., Kakuru, D., 2012. Funds of knowledge in child-headed households: a Ugandan case study. *Childhood* 19 (3), 397–413.
- Kohli, R., 2019. Lessons for teacher education: the role of critical professional development in teacher of color retention. *J. Teach. Educ.* 70 (1), 39–50.
- Lara, G.P., Fránquiz, M.E., 2015. Latino bilingual teachers: negotiating the figured world of masculinity. *Biling. Res. J.* 38 (2), 207–227.
- Lee, O., Llosa, L., Grapin, S., Haas, A., Goggins, M., 2019. Science and language integration with English learners: a conceptual framework guiding instructional materials development. *Sci. Educ.* 103 (2), 317–337.
- Lee, O., Miller, E.C., Januszky, R., 2014. Next generation science standards: all standards, all students. *J. Sci. Teach. Educ.* 25 (2), 223–233.
- Llopart, M., Esteban-Guitart, M., 2017. Strategies and resources for contextualising the curriculum based on the funds of knowledge approach: a literature review. *Aust. Educ. Res.* 44 (3), 255–274.
- Llopart, M., Esteban-Guitart, M., 2018. Funds of knowledge in 21st century societies: inclusive educational practices for under-represented students. A literature review. *J. Curric. Stud.* 50 (2), 145–161.
- Lynn, M., 2006. Education for the community: exploring the culturally relevant practices of Black male teachers. *Teach. Coll. Rec.* 108 (12), 2497.
- Mantei, J., Kervin, L., 2014. Interpreting the images in a picture book: students make connections to themselves, their lives and experiences. *Engl. Teach. Pract. Critiq.* 13 (2), 76–92.
- Marsh, M.M., Zhulamanova, I., 2017. Follow the leader: attending to the curriculum making potential of preschoolers. *Early Child. Dev. Care* 187 (5–6), 1004–1014.
- Martínez-Roldán, C.M., Fránquiz, M.E., 2009. Latina/o youth literacies: hidden funds of knowledge. In: Christenbury, L., Bomer, R., Smagorinsky, P. (Eds.), *Handbook of Adolescent Literacy Research*. Guilford Press, New York, NY, pp. 323–342.
- McKinney, C., 2017. *Language and Power in Post-colonial Schooling: Ideologies in Practice*. Routledge, New York, NY.
- Moll, L.C., 2015. Tapping into the “hidden” home and community resources of students. *Kappa Delta Pi Rec.* 51 (3), 114–117.
- Moll, L.C., 2019. Elaborating funds of knowledge: community-oriented practices in international contexts. *Lit. Research* 68 (1), 130–138.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theory Into Pract.* 31 (2), 132–141.
- Moulton, M., 2018. Funds of identity and humanizing research as a means of combating deficit perspectives of homelessness in the middle grades. *Educ. Sci.* 8 (4).
- Nagle, J., Stooke, R., 2016. Railways, rebellions and rage against the machine: adolescents’ interests and meaning-making in the creation of multimodal identity texts. *Literacy* 50 (3), 158–165.
- National Center for Education Statistics, 2018. Table 209.10. Available from: https://nces.ed.gov/ipeds/data/ipedsonline/tables/209_10.asp. (Accessed 1 February 2022).
- Norton, B.P., 1995. Social identity, investment, and language learning. *Tesol Q.* 29 (1), 9–31.
- Ordoñez, D., Siqués, C., Esteban-Guitart, M., 2018. ‘The best way to learn language is by not doing language’. Incorporating funds of identity for learning Spanish in a shared education unit. *Int. J. Biling. Educ. Biling.* 24 (9), 1259–1270.
- Paris, D., Winn, M.T., 2014. Trust, feeling, and change: what we learn, what we share, what we do. In: *Humanizing Research: Decolonizing Qualitative Inquiry for Youth and Communities*, pp. 223–248.
- Poole, A., 2017. “I want to be a furious leopard with magical wings and super power”: developing an ethico-interpretive framework for detecting Chinese students’ funds of identity. *Cogent Education* 4 (1).
- Poole, A., 2021. Towards a multimodal method for identifying and interpreting funds of identity derived from avatars. *Int. J. Res. Method Educ.* 1–13.
- Poole, A., Huang, J., 2018. Resituating funds of identity within contemporary interpretations of Perezhivanie. *Mind Cult. Activ.* 25 (2), 125–137.
- Recchia, S.L., McDevitt, S.E., 2018. Unraveling universalist perspectives on teaching and caring for infants and toddlers: finding authenticity in diverse funds of knowledge. *J. Res. Child. Educ.* 32 (1), 14–31.
- Reynolds, A.B.P., 2014. *Creepy Carrots!* Scholastic, New York, NY.
- Rodríguez, G.M., 2013. Power and agency in education: exploring the pedagogical dimensions of funds of knowledge. *Rev. Res. Educ.* 37 (1), 87–120.
- Rodríguez, N.N., Salinas, C.S., 2019. “La lucha todavía no ha terminado”/The struggle has not yet ended: teaching immigration through testimonio and difficult funds of knowledge. *J. Curriculum Theor.* 34 (3).
- Ruiz, R., 1984. Orientations in language planning. *NABE J.* 8 (2), 15–34.
- Salazar, M., 2013. A humanizing pedagogy: reinventing the principles and practice of education as a journey toward liberation. *Rev. Res. Educ.* 37 (1), 121–148.
- Saubich, X., Esteban, M., 2011. Bringing funds of family knowledge to school. The living Morocco project. *REMIE Multidiscip. J. Educ. Res.* 1 (1), 79–103.
- Sibbett, L., 2020. Intersectionality in US Educational Research. *Oxford Research Encyclopedia of Education*, pp. 1–42 (Oxford University Press).
- Sleeter, C.E., 2017. Critical race theory and the Whiteness of teacher education. *Urban Educ.* 52 (2), 155–169.
- Subero, D., 2015. La resignificación de la escuela. El proyecto fondos de identidad en el centro abierto de ciudad Meridiana, Barcelona (FICAB). *Àmbits de Psicopedagogia i Orientació* (43), 69–72.
- Subero, D., Llopart, M., Siqués, C., Esteban-Guitart, M., 2018. The mediation of teaching and learning processes through *identity artefacts*. A Vygotskian perspective. *Oxf. Rev. Educ.* 44 (2), 156–170.
- Subero, D., Vujasinović, E., Esteban-Guitart, M., 2017. Mobilising funds of identity in and out of school. *Camb. J. Educ.* 47 (2), 247–263.
- Tan, E., Calabrese Barton, A., 2010. Transforming science learning and student participation in sixth grade science: a case study of a low-income, urban, racial minority classroom. *Equity Excell. Educ.* 43 (1), 38–55.
- Urieta, L., 2009. *Working from within: Chicana and Chicano Activist Educators in Whitestream Schools*. University of Arizona Press, Tucson.
- Vélez-Ibáñez, C.G., Greenberg, J.B., 1992. Formation and transformation of funds of knowledge among US-Mexican households. *Anthropol. Educ. Q.* 23 (4), 313–335.
- Vila, I., Siqués, C., 2013. Les llengües de l’alumnat dins del sistema educatiu català als inicis del segle XXI. © *Articles de Didàctica de la Llengua i de la Literatura* 60, 18–27.
- Villalva, K.E., 2006. Hidden literacies and inquiry approaches of bilingual high school writers. *Writ. Commun.* 23 (1), 91–129.
- Walkington, C., 2017. Design Research on Personalized Problem Posing in Algebra. North American Chapter of the International Group for the Psychology of Mathematics Education.
- Weisman, E.M., Hansen, L.E., 2008. Student teaching in urban and suburban schools: perspectives of Latino preservice teachers. *Urban Educ.* 43 (6), 653–670.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Zipin, L., 2009. Dark funds of knowledge, deep funds of pedagogy: exploring boundaries between lifeworlds and schools. *Discourse* 30 (3), 317–331.
- Zipin, L., 2013. Engaging middle years learners by making their communities curricular: a funds of knowledge approach. *Curric. Perspect.* 33 (3), 1–12.
- Zipin, L., Sellar, S., Hattam, R., 2012. Countering and exceeding ‘capital’: a ‘funds of knowledge’ approach to re-imagining community. *Discourse* 33 (2), 179–192.

Family literacies: differing perspectives

Jim Anderson^a, Laura Teichert^c, and Ann Anderson^b, ^a Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada; ^b Department of Curriculum and Pedagogy, University of British Columbia, Vancouver, BC, Canada; and ^c Department of Special Education and Literacy Studies, Western Michigan University, Kalamazoo, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction and background	168
Theoretical frames	168
Naturalistic studies of family literacies	169
Cultural and social diversity	169
Mediators of literacy	170
Digital family literacy	171
Home digital literacy environments	171
Children's uses of digital technology in homes	171
Family literacy programs	173
Assessing and evaluating family literacy programs	173
Maintaining and promoting first/home languages	174
Building on cultural capital and funds of knowledge	174
Gender and family literacy programs	175
Conclusion	175
References	176

Introduction and background

In this chapter, we critically examine the literature in the area of family literacies. We begin by tracing the popularization of the term in the 1980s. Next, we briefly describe theoretical frames that inform the field of family literacies and our interpretation of the research literature. As the concept of family literacy gained prominence, its meaning became bifurcated in that, for some people, it meant the day-to-day literacy activities and events families engage in at home and in the community while, for others, it connoted intervention programs (Anderson et al., 2010; Hannon, 2003). We first review “naturalistic” studies of family literacy and, recognizing that digital technology is a facet of daily life for many families (e.g., Marsh et al., 2017a), we address how families take up digital literacy. Acknowledging that there are ongoing issues associated with family literacy programs, we next examine how program developers and facilitators have addressed these issues (or not). We conclude by discussing the implications of research in this area, identifying what we believe are gaps or limitations, and speculate about the future of family literacy scholarship and programming.

Denny Taylor popularized the term, “family literacy” with the publication of her acclaimed 1983 book, *Family literacy: Young children learning to read and write*. That ethnographic study documented how pre-school children growing up in middle class homes in the United States (US) were acculturated into literacy as their parents and other family members engaged in various literacy activities and events as part of daily life. In a later study with inner-city families in the US, Taylor and Dorsey-Gaines (1988) showed how parents and other family members of young children living in severe poverty, and in challenging and sometimes dangerous conditions, supported their children's literacy development and learning at home.

Although Dolores Durkin (1966) in the United States and Margaret Clark (1976) in the United Kingdom had done studies with young, precocious readers who had learned to read at home prior to formal instruction in school or preschool, Taylor's work precipitated a great deal of interest in family literacy. As Panofsky (2000) pointed out, at this time there was a burgeoning interest in children's early literacy development (e.g., Clay, 1966, 1975; Ferreiro and Teberosky, 1982; Ninio, 1983; Teale and Sulzby, 1986), and several academics (e.g., Baghban, 1984; Bissex, 1980) documented their own children's emerging literacy. Thus, Taylor's focus on young children fits with the emergent literacy paradigm that was prominent in that era. A perusal of the literature on family literacy suggests that this emphasis on young children has been pervasive and persists today.

Theoretical frames

As Lynch and Prins (2021) indicate, scholarship in family literacy and family literacy programs draw on socio-cultural theory (e.g., Vygotsky, 1978) and the tenet that learning is initially social, before it is individualized. That is, guided by parents and significant others in their family and community, children/learners use “tools” such as literacy inter-psychologically. Wood et al. (1976) refer to the mediation by a more experienced or proficient person as “scaffolding” that is gradually withdrawn as learners use the tools

intra-psychologically, or independently. However, “human development is a cultural process” (Rogoff, 2003, p. 4) and different cultural groups support development and learning differently.

Consonant with Rogoff’s perspective is the concept of literacy as social or cultural practice (Heath, 1983; Street, 1995). That is, the meanings ascribed to literacy, its functions and uses, its importance and significance, and how it is learned and taught vary from one cultural and social context to another (Clay, 1993). For example, Heath’s (1983) groundbreaking ethnography in three economically and socially diverse communities in the Southeastern United States exemplifies how families’ literacy practices reflect this phenomenon. She found that while literacy in school aligned with literacy practices in the homes of the mainstream Townspeople, they did not align with those in the white working class community of Roadville nor the working class African-American community of Trackton and the children in the latter two communities were not well served by school.

Although much of the research in family literacy focuses on print, researchers and practitioners are increasingly drawing on “multiliteracies”, a concept “... based on the assumption that individuals ‘read’ the world and make sense of information by means other than traditional reading and writing” (International Bureau of Education, 2021, n.p). Relatedly, Kress (2010), and others employ the term multimodality in referring to the understanding that meaning is constructed and communicated through a number of modes, often simultaneously in the contemporary digital and technological world. We reference both concepts in this chapter.

Naturalistic studies of family literacies

Cultural and social diversity

Research documenting families’ literacies in homes and communities have continued apace in the nearly half-century since Taylor’s study. In this section, we briefly canvas studies from culturally and socially diverse contexts and examine differences in the mediational roles that significant others play.

In an influential study, Purcell-Gates (1996) documented the literacy practices in twenty low-income, urban homes in the United States. She found that families engaged with a range of texts on a daily basis and that there was a relationship between the types of texts read at home and children’s emergent knowledge of print. Essentially, children from homes where longer texts (e.g., newspapers, magazines) were read developed more emergent knowledge of print than children where shorter texts (e.g., coupons, TV Guide) were read. There was considerable variation in the amount and frequency of reading among this relatively homogeneous socio-cultural group with some families engaging in a lot of reading and writing while other families did little. Indeed, “Some families in this study, in fact, lived busy and satisfying lives with very little mediation of print” (Purcell-Gates, 1996, p. 425).

Family literacy is often synonymous with shared book reading, notwithstanding the fact that adults sharing books with children is a practice particular to some cultural groups, and is far from universal (Anderson et al., 2003; Reese, 2012). For example, reflecting on literacy practices in her family growing up on a Navaho reservation in Northern Arizona, Marlinda White-Kaulaity (2007, p. 560) recalled, “‘Read me a story’ are words I never said. ‘I’ll read you a story’ are words I never heard.” Instead, she explained, storytelling was a widely used tool in her community for learning and teaching. Indeed, reading and writing were considered of little immediate benefit, and not highly regarded. Furthermore, reading and writing have been used as colonialist tools within Indigenous communities to disempower and discredit people, and thus are viewed with suspicion (White-Kaulaity, 2007).

Other studies demonstrate the multiple ways in which some families interact with texts in different modes, switching between different languages. Sibanda and Kajee (2019) examined the multi-lingual, multimodal language and literacy practices in the homes of 3, 8 and 9-year-old children who attended a primary school in a township in Gauteng province, South Africa. Thandi, a Zulu speaker who participated in the study, liked English language television programs. She especially liked drawing and wanted to become a fashion designer. She borrowed her uncle’s mobile phone to play games. As well, she regularly read books that the school provided. Setswana, an official language of South Africa, was the medium of instruction at school, and Thandi’s facility with different languages indicates she is multilingual. Another participant, Lindi, moved from KwaZulu-Natal, after her parents died, to live with her maternal grandmother and other relatives who spoke isiXhosa. She read the Bible at church services and assisted adults there who were unable to read. Lindi consistently reminded her grandmother to complete and submit the forms necessary for government support for orphaned children such as herself and her siblings. Tumelo’s (the third child in the study) first or home language, Setswana, was the language of instruction at school. She engaged in shared reading of school texts with her mother, and was also expected to read them on her own. Interestingly, Tumelo’s mother indicated that she also learned by helping Tumelo with homework. In these homes, family literacy may include, but extends beyond shared book reading.

Similarly, an ethnographic study with three kindergarteners in Puerto Rico conducted by Volk and de Acosta (2003) exemplifies the diverse ways that families from different socio-cultural contexts engage in literacies. They analyzed what Volk and de Acosta called “syncretic” at-home literacy practices whereby the “participants drew on oral and written texts from various contexts in their lives and reinvented them. The new practices they created were more than a juxtaposition of old ones, but rather were recreations that provided new opportunities for learning” (p. 22). They drew on texts from school, from church, from the community and popular culture. Although Spanish was the home and dominant language of the children, they were also facile in English. As well, in addition to their parents, a network of significant others including siblings and other relatives supported the children’s learning.

As indicated previously, literacy is embedded in religious practices in some homes and, over the years, researchers have studied this phenomenon (e.g., Fishman, 1987; Lytra and Ilankuberan, 2020; Lytra et al., 2016). As part of the Becoming Literate in Faith

Settings (BeLiFs, 2021) project in London, UK, Souza et al. (2016) reported on an ethnographic study in which they gave children flip cameras to record religious practices at home and had them construct scrapbooks depicting these faith practices. Souza and colleagues indicated that the religious practices included oral and written texts, and facilitated an understanding of symbolism and metaphorical language, abilities that schools should encourage and help students develop.

These studies reflect the theoretical perspectives outlined earlier in several ways. For example, the differences within the relatively homogeneous group of families in the Purcell-Gates study exemplifies a literacy as social practices perspective. Likewise, the range of texts and modes for different functions and purposes that the children in Sibandi and Kajee's study engaged in, and their plurilingualism, are consistent with a multi-literacies or multi-modal frame. Furthermore, the mediational roles that adults and children perform across the studies align with socio-cultural theory.

Mediators of literacy

Family literacy is often considered intergenerational in that parents (or grandparents) play important roles in mediating learning. Intergenerational learning is a feature of many cultures around the world (Ouane, 2008) and much of the literature in family literacy focuses on parents supporting their children's learning. However, we next examine studies that have investigated the roles of siblings and significant others, besides parents (or grandparents), in supporting learning in family contexts.

Gregory (2005) observed and recorded the literacy activities that young children engaged in at home in East London, UK. The eight families, each with an older child who was 10 or 11 years old and a younger child aged 4–8, had emigrated from Bangladesh. The children were bilingual in English and Sylheti, a Bengali dialect. At home, the children regularly played school with the older child, emulating closely the literacy lessons from school, as they "taught" their younger sibling. The older children also shared academic or disciplinary knowledge learned at school, imparting the routines and practices from there. Gregory (2005, p. 37) illustrated, "the very special role played by siblings close in age as literacy mediators in initiating their younger siblings into the wider school discourses ... or ways of behaving, valuing, and expressing new opinions, beliefs, and views". The older children also benefited as they practiced and solidified their literacy skills and strategies in a familiar, low-risk environment (Gregory, 2001).

In their study of child headed households in Uganda, Kendrick and Namazzi (2017) also identified the inadequacy of conceptualizing family literacy as parent-child apprenticeships. They worked with four families where the heads of the household ranged in age from twelve to eighteen, and with other family members as young as three. Kendrick and Namazzi documented a number of practices across the families, including: using songs and drama to educate about HIV/AIDS, and to entertain, communicate information and knowledge, and share feelings; bringing home sports news garnered from the market and reporting them in the format of a radio program; and using family photographs to discuss family history. Importantly, their findings suggest the need to conceptualize the Zone of Proximal Development theorized by Vygotsky (1978) differently from the parent/significant other-child dyad model that is implicit in much of the literature. They stated, "These practices required that siblings commonly operate within a collective Zone of Proximal Development (ZPD) whereby no one member of the family necessarily has the expertise to teach or apprentice younger siblings" (Kendrick and Namazzi, 2017, p. 70).

Literacy brokering is practiced in many families, and in some cases, children act as literacy brokers for their parents or older family members (e.g., China and Robins, 1996; Orellana et al., 2003; Perry, 2014). According to Perry (2009), "Literacy brokering is a complex activity that may involve one aspect of a text, such as translation of word meanings, mediation of cultural content, or explanation of genre aspects of a printed text, or it may involve many of these aspects all at once" (p. 257). Utilizing ethnographic methods, Perry (2009) documented the literacy brokering practices in three Sudanese families with refugee status in the United States. She identified a number of individuals who were literacy brokers for the adults, including their children who were in kindergarten and grade one. Although literacy brokering is sometimes conceptualized as translating or paraphrasing (e.g., Orellana et al., 2003), Perry found that it went beyond these functions. She identified three types of literacy brokering: (1) lexico-syntactic and graphophonic where the broker explained the meaning of an unfamiliar word, or explained word order or word choice in writing a note, or demonstrated letter-sound relationship in identifying words; (2) culture brokering which entailed explaining customs or holidays or events; and (3) genre brokering, for example, explaining the purpose of grocery store coupons or hyperlinks on a website (Perry, 2009, p. 263). In her 2009 study, Perry recorded 257 literacy-brokering events, nearly half of which involved genre brokering. In a later study, she noted that children's brokering for their parents often relates to school such as forms, memos, newsletters and so forth (Perry, 2014).

The research on brokering as a family literacy practice such as Perry's and the others previously cited usually involved families' first languages and the dominant language/language of instruction in school in the resettlement country. However, in his study involving four Afghani families with refugee status living in Pakistan, the country of first asylum while waiting to immigrate to a country in which to permanently settle, Sadiq (2020) found that the children's brokering involved three languages-Pashto, their home language from Afghanistan, Urdu, the official language of Pakistan, and English. For example, they translated texts such as bills, wedding invitations, cards, and notices and notes from school written in English or Urdu to Pashto for their parents. Likewise, 11-year-old Safa (pseudonym) translated the Duas, Islamic texts written in Arabic, to Pashto for her father so that he could comprehend them.

In line with socio-cultural theory, these studies demonstrate the various, important and often culturally rooted role(s) of mediators in literacy learning. Thus although, it is implicit in much of the theoretical literature that adults are the primary mediators for children, these studies show that this is clearly not always the case.

Digital family literacy

As Marsh et al. (2017b) pointed out, there is a significant research base indicating that digital technology is now an integral aspect of family life for many people and thus it is essential to examine digital literacy in the context of homes and communities. In her seminal work, Marsh (2004) identified “techno-literacy” as significant in the home and defined the concept as literacy events and literacy practices mediated by digital technology. Marsh argued literacy skills were no longer limited to paper-based activities and that children and other family members accessed texts from multiple modes. She highlighted the “importance of visual, aural, and corporeal ways of meaning-making” (p. 52), not just print. She also outlined children’s emergent techno-literacy practices and the role(s) parents (and other significant adults) played in scaffolding children in learning to use technological devices such as computers and mobile phones. Marsh found that parents supported and encouraged young children’s techno-literacy practices and that these practices played a role in families’ relationships in the home. For example, the area around the TV was often set up for “celebrating and extending the children’s relationship with the screen” (p. 58), such as keeping dress-up clothes or toys nearby so children could easily access these props when re-enacting screen content. She also noted that children were not simply passive viewers but played with narratives and images viewed on the screen through re-imagined story lines and imitation of on-screen behavior. She found that men were more likely to collaborate with children in video game play whereas women were more likely to participate in print-literacy activities (e.g., read a storybook), a finding also observed by Connell et al. (2015) and Dias et al. (2016).

In the almost two decades since Marsh published her work, the field of young children’s digital tool use at home has increased exponentially. A number of large-scale surveys have examined young children’s access to digital technology in western countries. Common Sense Media (Rideout, 2013, 2017; Rideout and Robb, 2020) and the Pew Research Center (Anderson et al., 2020) have collected statistical information about Internet and digital access and use in the United States, and OFCOM (2016) and Eurostat (<https://ec.europa.eu/eurostat/web/main/home>) have collected similar data across Europe. These studies indicate the prevalence of digital tools in families across cultural and social groups, which in turn impacts engagement in digital literacy practices at home.

Home digital literacy environments

Although large-scale surveys provide important data about digital access and use, naturalistic studies have provided important insights into families’ digital literacy practices in homes and communities. Differences in how much or how little digital technology is used in homes has often been described on a spectrum of low use to high use (e.g., Chaudron et al., 2018; Plowman et al., 2012; O’Hara, 2011). For example, Wong (2015) categorized families as new dispositions (allowed a range of technology in the home and were enthusiastic about use); new and old dispositions (held mixed values toward traditional literacies and new technology); and traditional dispositions (few digital tools in home and were guarded about digital technology, thereby limiting children’s uses). It has been widely reported in research that many parents feel technology contributes to children’s learning and development and that technology is required for life in today’s society (Aubrey and Dahl, 2014; Dias et al., 2016; Kumpulainen et al., 2020; Kervin et al., 2018; O’Hara, 2011; Ozturk and Ohi, 2019; Schlembach and Johnson, 2014). Kervin et al. (2018), for example, quoted an Australian parent of a preschooler stating, “they [children] should have the basics of how to operate [the technology] ... *I don’t want him left behind*” (p. 120) (emphasis added). However, while parents have reported wanting their children to develop digital technology skills, they also seek to limit screen time (Gillen et al., 2018; Kervin et al., 2018; O’Hara, 2011; Teichert, 2017). Livingstone and Blum-Ross (2020) reported on parents who tracked the amount of time a child used a screen instead of focusing on quality of screen time interaction. They posited that screen time seemed “to offer parents an officially-sanctioned and seemingly straightforward metric of parenting ‘success’” (p. 33). It is important to note that parents’ positions on digital technology are often fluid. Parents in the same family may at times differ in their values and beliefs and negotiate with each other on what the best approach may be for their children (Livingstone and Blum-Ross, 2020; Roach O’Keefe and O’Keefe, 2020). For example, Livingstone and Blum-Ross (2020) demonstrated this tension when describing participants Lara and Pawel Mazur, who lived in the United Kingdom. Lara believed digital technology provided their 6-year-old son with opportunities to “build his confidence and make him independent” while Pawel was cautious and worried about online risks, “especially after [son’s friend] introduced him to a violent video game” (p. 1). Pawel therefore set passwords on all devices in an attempt to monitor and control his son’s access. Researchers have reported that media discussions of young children using technology, parents’ educational backgrounds, and their prior experiences of using technology for work, study or leisure all influence parents’ beliefs (e.g., Plowman et al., 2008, 2010; Teichert, 2017).

Children’s uses of digital technology in homes

For many young children, home is typically the first space where they encounter and experience digital technologies and media (Kervin et al., 2018), with the age of child influencing access to, and location of, digital tools. For example, Poveda et al. (2020) noted how much of younger children’s use occurred in general living spaces (e.g., the living room), which, unsurprisingly, contrasts with descriptions of older children who own their own digital devices and independently use them in personal spaces (e.g., bedrooms) (Livingstone and Blum-Ross, 2020).

Similar to the research on print literacy in families (e.g., Purcell-Gates, 1996; Taylor, 1983), studies have noted how children's use of digital technology was embedded in daily family life (Gillen et al., 2018; Kumpulainen et al., 2020; Livingstone and Blum-Ross, 2020; Ozturk and Ohi, 2019; Palaiologou, 2016; Teichert, 2020). Gillen et al. (2018) described it as "the rhythm of the day" (np) while Gee et al. (2018) noted the "dynamics of family interactions and routines, and how the meaning and use of media are caught up in these dynamics" (p. 9). Children demonstrate operational skills from a very young age (Gillen et al., 2018; Harrison and McTavish, 2018; Murphy and Headley, 2020; Poveda et al., 2020), but it is a misconception that children "just pick it up" (Plowman et al., 2008, p. 311). As Plowman et al. (2008, 2012) documented, children learn operational skills by observing more proficient digital technology users in their homes, notably parents and siblings. Parents scaffold children's uses (Davidson, 2009, 2010; Davidson et al., 2018; Gillen et al., 2018), even if they are unaware of it. Older siblings mentor younger siblings' digital engagement by playing alongside them and explaining how to operate the device and/or game (Davidson, 2010; Kervin et al., 2018; McPherson, 2011). However, once operational skills are developed, children show an adeptness at trouble-shooting when technology does not work and astute planning to organize resources to ensure proper functioning before engaging in tasks (e.g., Taylor et al., 2018).

Some researchers describe children's uses of digital technology at home as unplanned or fleeting and encounters may be more interstitial as children move between different digital games or videos and between imaginary play in offline spaces (Fleer, 2014; Marsh et al., 2017a; Plowman et al., 2012). Importantly, digital technology does not supplant other types of play children engage in while in their homes (Kumpulainen et al., 2020; Marsh, 2004; Marsh et al., 2016; Plowman et al., 2013). Children continue to engage in activities more traditionally associated with childhood, such as outdoor play, book reading, and non-digital dramatic play (Marsh et al., 2017b; Plowman et al., 2013; Teichert and Anderson, 2014). By design, Lego presents an interesting example of children's online/offline multimodal play as children manipulate physical bricks (e.g., build), while using visual paper-based instructions or accessing additional instructions online via websites or QRI codes (Roach O'Keefe and O'Keefe, 2020; Wong, 2015). Five-year-old Andrew, the focal participant in Wong's (2015) study went a step further and developed Lego instructional videos and posted them on YouTube.

Taken together, these studies indicate that when children interact with digital technology at home, they are not passive users. They re-enact what they see or play in a variety of ways, such as dramatizing the story lines or characters they know from TV, movies, YouTube, and video games (Barron and Levinson, 2018; Given et al., 2016; Huh, 2015; Marsh et al., 2015; Poveda et al., 2020; Stephens et al., 2013; Takeuchi and Stevens, 2011; Teichert, 2019). Sometimes children's dramatic play mimics what they are seeing online in-the-moment (Kervin, 2016; Kervin et al., 2018; Scriven et al., 2018; Teichert, 2020). Children shifting between online and offline formats has also been evident in their informational search practices. For example, a child expresses an interest in a topic and, typically alongside a parent, searches books, videos, and the Internet to learn more about the topic (Barron and Levinson, 2018; Davidson, 2009). In contrast to the notion of technology as a babysitter (American Academy of Pediatrics, 1999), families frequently use technology together, especially those with younger children (O'Hara, 2011; Kumpulainen et al., 2020; Livingstone and Blum-Ross, 2020; Marsh, 2004; Teichert, 2020; Wolfe and Flewitt, 2010). Therefore, rather than being seen as babysitters, digital devices might be considered as sites for shared family experiences (Plowman et al., 2013).

Staying in contact with family and friends is a frequently cited use of digital technology in homes (Teichert, 2020; Busch, 2018; Flewitt and Clark, 2020; Livingstone and Blum-Ross, 2020; Kumpulainen et al., 2020; Harrison and McTavish, 2018). Plowman et al. (2013), for example, described how 3-year-old Colin, who lived in the UK, proficiently took photographs and initiated video-chatting with family in Australia. Becker (2021) described a Brazilian immigrant mother and her 6-year-old daughter living in the US and the "unboxing" videos created in response to boxes filled with gifts they received yearly from relatives in Brazil.

In addition to staying in touch, digital technology may help children develop emergent print literacy skills (Bittman et al., 2012; Neumann, 2014) and has been cited as beneficial for children with speech disorders, when using apps approved by speech pathologists (Wren et al., 2018). Families have also reported online engagement as a way to develop heritage language skills (Harrison and McTavish, 2018; Takeuchi and Ellerbe, 2018), support bilingual/multilingual children (Barron and Levinson, 2018; Marsh et al., 2017b), or develop English language skills (Kumpulainen et al., 2020; Ozturk and Ohi, 2019). Language brokering of online materials has also been documented (e.g., Yip et al., 2018 with Hispanic youth in the US).

Although much of the research on digital literacy in families has focused on young children, as just noted, some studies have involved older children and youth. For example, Murphy and Headley (2020) examined the role of digital technology in teenage mothers' and children's literacy. They interviewed eight teenage mothers in an urban area in Canada about their recollections of their own early literacy experiences, their children's engagement in literacy, and their own current literacy practices, especially with regard to digital technology. All of the participants had a mobile phone or tablet and access to a computer. They reported using the digital devices for a range of functions and purposes, including, reading and writing different forms and genres of texts, accessing information and resources, creating artistic work, and for organizing their daily lives. They also indicated that digital tools could be distracting and worried about their own and their children's overuse of them.

Despite the wealth of observational data available, the majority of this work occurred in Western Europe, North America, and Australia. We know much less about families' digital literacy practices in the global south. Furthermore, many of the research participants describe themselves as middle class, and frequently live in metropolitan and suburban areas. Although efforts have been made to examine families in rural communities (e.g., Flewitt and Clark, 2020 in UK, Wong, 2015 in Canada, Roos and Olin-Scheller, 2018 in Sweden), more observational studies are needed to better understand how digital literacy is enacted in places where connectivity is less consistent and access to digital technology is limited.

As with print literacy, the research reviewed here reflects the theoretical perspectives outlined earlier. For example, mediation by a more experienced or significant other is highlighted in many of these studies. As well, digital literacy varies across contexts for different functions and purposes. Of course, digital literacy is multi-modal by design.

Family literacy programs

Before reviewing the research on family literacy programs and examining some examples more closely, it is important to acknowledge and discuss concerns about, and critiques of, these programs. For example, Purcell-Gates (2000) and Thomas and Skage (1998) identified a lack of assessment or evaluation in such programs, and from an ethical and accountability perspective, relatively little evidence that children and families were benefitting in various ways from their participation. Auerbach (1989) raised several issues, including: family literacy programs often target marginalized families, such as those living in poverty, and recent immigrants and refugees; families' home languages are often ignored and acquiring English (or the dominant language of the community) is promoted; and families' home literacies are unrecognized and school literacy is privileged. Reyes and Torres (2007) criticized family literacy programs for imposing middle class, Eurocentric values or practices on Indigenous families and those from culturally, linguistically, and racially diverse backgrounds and called for decolonizing family literacy programs (Crooks, 2019). Furthermore, some researchers have argued that family literacy programs are gendered (e.g., Cuban and Hayes, 1996; Rizk, 2020), designed with the assumption that mothers will attend sessions with their children and follow up with literacy activities at home, adding yet another expectation on them.

Assessing and evaluating family literacy programs

In the period since Thomas and Skage (1998) and others raised concerns about the lack of attention to program evaluation, several researchers have addressed this issue, including meta-analyses that focused on children's language and literacy learning. For example, Marulis and Neuman (2010) investigated the effectiveness of various programs on young children's vocabulary development. They reported a large effect size (0.88) for sixty-seven studies involving 5929 children. In their meta-analysis, van Steensel et al. (2011) identified studies in which researchers measured comprehension skills and print related or code skills. The thirty studies included 4326 children, of whom 1866 had participated in family literacy programs and 2460 were in comparison groups. The average effect size for the programs was 0.20, which is statistically significant but considered small (e.g., Cohen, 1988). Nevertheless, van Steensel et al. (2011) concluded that the programs had "a non-trivial, positive effect" (p. 88). More recently, Fikrat-Wevers et al. (2021) analyzed forty-eight studies and found an average effect size of 0.50 immediately after the programs but an effect size of 0.16 in follow-up measures. Finally, in a meta-study that did not involve calculating effect sizes, Brooks et al. (2008) found that children made gains in language and literacy in eighteen of the nineteen qualitative and quantitative studies and in four of the five studies that had follow-up measures, gains were sustained.

Despite these generally positive results, however, Paratore and Yaden (2011) pointed out, the federally funded Even Start program offered throughout the United States was defunded based on studies that indicated that the popular program did not significantly improve children's language and literacy learning. They questioned whether such large-scale programs, "perhaps impersonally implemented with varying quality goals, and even conflicting or unspecified theories of change ... work against the type of personal and professional cohesion" (p. 94) fundamental to the success of family literacy programs.

Although many studies focused on program impact on children's language and literacy learning as measured by standardized tests, some researchers have looked at other effects. For example, Swain et al. (2014) found that the 100 parent participants they interviewed from seventy-four programs in England indicated: they more capably supported children's learning; came to value education more; enhanced their understanding of schools/schooling; developed social connections; and became more interested and invested in their own literacies. Anderson et al. (2013) reported similar findings in a study in which they were able to locate ten of eighteen parents from a cohort in an economically depressed, rural community in Eastern Canada, who had participated in a community-initiated family literacy initiative with their 4- and 5-year olds two decades earlier.

These findings align with the perspective of noted British sociologist Diane Reay (1998), who grew up in a working-class home, that educators need to make explicit to culturally, economically and socially diverse families, exactly what they need to do to support their children's development and learning.

While research such as that just reviewed on the impact of family literacy programs have generally focused on children's learning and parents' learning about how to support their children, Toso and colleagues have looked at how mothers who participate in these programs personally benefit from them. For example, Toso (2012) found that the Mexican immigrant women in a family literacy program in the US, developed agency, appropriating the discourse about mothering to serve their needs and to meet the aspirations they had for themselves and their children. She concluded, "Family literacy, often seen as a site of colonization ... or imposition of normative ideas of parenting ..., offered a place to explore ideas of the self, thereby pushing the boundaries of women's traditionally held mothering discourses" (Toso, 2012, p. 4). Similarly, in their qualitative study, involving two family literacy programs in the United States, Prins et al. (2009) reported that the women experiencing conditions of extreme poverty who participated saw them as sites of social contact and support, outcomes that are often not assessed in family literacy programs or are ignored as being peripheral.

Maintaining and promoting first/home languages

Many programs continue to be offered in the dominant language in the community (e.g., English); however, educators have developed programs that include the first language (L1) of new immigrants and families of refugee status. In an initiative in Sheffield in the UK (Hirst et al., 2010), a cultural worker and an early childhood educator regularly visited sixteen families who had emigrated from Pakistan, providing literacy materials and demonstrating to families how they could support their children's learning. Families used their home languages, Mirpuri Punjabi or Urdu, or English at their choosing, as the facilitators did not privilege any language. The children made significant gains on the Sheffield Early Literacy Development Profile (SELD) (Nutbrown, 1997) in comparison to a control group. Parents/caregivers indicated that they benefited from participating and believed that the program should be available to other families. In a study of the Home Enrichment of Language and Literacy Skills (SHELLS) program with migrant workers in the western United States, Boyce et al. (2010) encouraged families to share their experiences in Spanish and facilitators then supported them in making books, based on the oral stories, that parents then read with the children. Findings indicated that the thirty-two children in the SHELLS program significantly increased the number of words they used, as well as the number of different words they used, compared to children in a control group. As well, parents in the SHELLS program significantly increased elicitation strategies compared to the control group. Anderson et al. (2011) developed and implemented the Parents as Literacy Supporters (PALS) in Immigrant Communities bilingual program with families from four different linguistic groups (Farsi, Karen, Mandarin, and Punjabi) in five communities in Western Canada (Anderson et al., 2011). The program encouraged children and families to maintain their home language as they simultaneously acquired/learned English as an additional language. An early childhood educator and a cultural worker from each community facilitated sessions for 4- and 5-year old children and their parents/caregivers in English and L1, all written materials were translated into home languages, and families received a dual language book at each session.

The program facilitators, guided by the cultural workers, incorporated as much as possible, the families' funds of knowledge—"the essential cultural practices and bodies of knowledge and information that households use to survive, to get ahead, or to thrive" (Moll and Greenberg, 1990, p. 321). Another principle of the program was respecting different cultural models of teaching and learning, acknowledging that in some cultures, children learn by first watching more proficient or experienced others (e.g., Rogoff, 2003), not by experimenting and approximating as advocated by many early childhood educators in western countries. Pre and posttest comparisons of the Normal Curve Equivalent Scores on the widely used Test of Early Reading Ability-2 (Reid et al., 1989), indicated that on average, the children made statistically significant gains with an effect size partial eta squared of 0.318, which is large (Cohen, 1988). Qualitative data from focus groups, field notes, debriefing sessions, and so forth, revealed that: the families: appreciated the bilingual format; increased their appreciation of maintaining their home language; developed greater understanding of western pedagogy; became more comfortable in school; and believed children's transition to kindergarten would be smoother. As other studies have found, the program helped demystify early literacy development and learning. However, the families also raised concerns and issues and requested more frequent sessions, extension of the program for older students and their parents, and more direct instruction (Anderson and Anderson, 2021a,b; Anderson et al., 2017, 2016, 2011).

Building on cultural capital and funds of knowledge

A goal of some programs is to build on participants' cultural and social capital and funds of knowledge. A good example of such an initiative was Project FLAME- Family Literacy Aprendiendo Mejorando Educando (Learning, Bettering, Educating), developed by Professor Flora Rodriguez Brown and colleagues with Spanish speaking families in Chicago (Rodriguez Brown, 2004, 2008). Parents who attended had children between 3 and 9 years of age. Adults attended bi-weekly, English as a Second Language classes and bi-monthly classes that focused on parents as teachers. Rodriguez Brown (2004) reported that both adults and children improved their literacy knowledge and skills, and that adults increased their confidence and self-esteem, which led to more involvement in the wider community and more social inclusion.

Oral language is the focus of the Parent-Child Mother Goose Program (P-CMG) where parents and young children share familiar poems, riddles, songs, stories and so forth in a group setting. Started in Toronto, Canada in 1984, the low-barrier, participatory program spread and is currently offered in many parts of Canada and elsewhere (The Parent-Child Mother Goose Program, 2021). Originally offered in English, it has expanded to include other languages and is offered in a variety of venues, including community centers, libraries, Aboriginal friendship centers, etc. (P-CMG, 2021). Indeed, during the Covid-19 pandemic, the public library in Coquitlam, British Columbia offered the program virtually in Mandarin (Coquitlam Public Library, 2021). The multi-modal program is interactive and includes movement, oral language and print versions of the poems, songs, etc., shared in sessions. Several studies have examined the effects of P-CMG. For example, in a study conducted in the state of Victoria, Australia, Terrett et al. (2012) found that the thirty children who attended P-CMG sessions made greater gains than the twenty-five children in a control group on measures of expressive language. Parents in P-CMG also reported feeling less stressed than those in the comparison group. In a study with families of children with hearing impairments in Iran, Koochi et al. (2017) reported that mothers and children benefitted from P-CMG. Likewise, based on interviews with participants in Ontario, Canada, Ling et al. (2017) reported several benefits of the program, including increased opportunities for parents to socialize.

As described earlier, besides promoting first language maintenance, the PALS program attempted to incorporate families' funds of knowledge, and respected and welcomed different cultural models of learning in working with culturally and linguistically diverse families. Alongside the PALS in Immigrant Communities initiative (Anderson et al., 2011), Gail Stromquist, coordinator

of Aboriginal Education for the Langley School District in British Columbia, worked with elders, community members and colleagues to develop a version of PALS for Aboriginal families and communities. Aboriginal PALS promotes Aboriginal culture(s) of British Columbia and encourages families to use the Indigenous language of their territory, many of which are endangered because of assimilationist and racist policies of governments in the past. First Nations (Aboriginal) communities throughout BC have implemented (and in some cases, locally adapted) the program. For example, Alison Gear, the Early Childhood Coordinator for the Haida Gwaii School District in northern British Columbia, collaborated with families and the broader community in developing Haida Gwaii PALS. Ensuring that the program content reflected the place-based orientation of Haida culture, as well as First Nations principles of learning ([First Nations Education Steering Committee, 2022](#)), [Gear \(2013\)](#) designed the project to be a “means of addressing the imbalance between Indigenous and Western knowledge systems currently offered through the public education system” (p. ii). Families appreciated having a program that reflected their culture and indicated that they and their children benefitted.

As these examples demonstrate, family literacy programs can be sites for sharing knowledge with families and supporting children’s language and literacy development, and at the same time, valuing and incorporating families’ cultural and social capital.

Gender and family literacy programs

Some family literacy programs appear to target mothers and to place additional childrearing demands and responsibilities on them (e.g., [Mace, 1998](#); [Rizk, 2020](#)) and fathers tend not to participate in them to the extent that mothers do ([Lynch and Prins, 2021](#); [MacLeod, 2008](#)). As would be expected, educators have developed initiatives designed for fathers in response to calls for their greater involvement (e.g., [Gadsden, 2012](#)). For example, the Vancouver Public Library in Canada developed the Man in the Moon Program to provide “opportunities for fathers and other male identifying caregivers to learn about how to best support early literacy and language development in their children through playing, singing, talking and reading together” ([Vancouver Public Library Foundation, 2021](#), n.p.). Feedback from participants has been positive and the program has run for nearly two decades. [Kagitcibasi \(2005\)](#) reported on a 22 year longitudinal study with fathers who participated in the Fathers Support Program in Turkey and found that they tended to become more involved with their children, adapted a more democratic, less authoritarian stance with children, and developed a supportive network with other fathers (p. 85). More recently, several researchers have worked with programs for incarcerated fathers (e.g., [Nutbrown et al., 2019](#); [Prins et al., 2020](#)). Using an interpretivist frame, Nutbrown et al. reported on a 5-month project sponsored by an interdisciplinary team in which males, incarcerated in two prisons in the United Kingdom, participated in a series of workshops and family visits. The authors concluded that the program supported fathers in committing to supporting their children’s literacy learning while in prison and resolving to avoid involvement in crime and to take up more parental responsibilities upon release. Of course, it remains unknown whether family literacy programs for fathers relieve some of the inequitable demands and expectations placed on mothers (e.g., [Rizk, 2020](#)).

As can be seen, some family literacy program developers and facilitators have attempted to respond to the concerns raised over the years to make programs more socially and culturally responsive; encourage fathers (men) to become more involved in them; shift away from a focus only on print and incorporate multiple literacies, and promote plurilingualism.

Conclusion

In the four decades since Denny Taylor published her work, research and scholarship in the field of family literacy has continued to grow and expand globally. Important naturalistic studies have helped broaden conceptions of families beyond parent-child dyads and extend the focus on other modes besides print. Many studies have exemplified the social nature of literacy and language development and learning in homes and communities, and shown the important roles that mediators perform. At the same time, these studies challenge the well-entrenched notion that parents and significant other adults such as grandparents necessarily are the mediators for children. As well, studies have demonstrated the range of languages and literacies that occur within homes and communities and the different formats, genres and modes that are employed in constructing and communicating meaning. Many researchers have examined how families take up digital literacies but there is still much to learn in this area.

Much of the literature has come from the global north, and especially western countries. More research from the majority south would broaden and deepen knowledge and understanding in the field. As mentioned, a great deal of the naturalistic research in family literacy has focused on young children. Naturalistic studies with older children, youth and adults would broaden our knowledge. As well, longitudinal studies in different cultural and social contexts, which document literacy practices within families over time, would enhance our understanding. Although research on families’ digital literacy has proliferated over the last decade or so, there are still gaps in the literature, especially in terms of access to digital tools and technology across socio-economic groups, as issues of equity and inclusion remain.

There is evidence that children and families benefit from family literacy programs in various ways; however, there is a dearth of longitudinal studies in the field. Although these pose logistical and pragmatic challenges, they would add considerably to our knowledge. For example, such studies would demonstrate whether the positive impact on children’s language and literacy learning that many of the studies reviewed earlier have documented are maintained as children progress through elementary and secondary school. Tracking child participants over the life course to determine the impact of programs in a number of domains, as was done in the Perry Preschool Project in Ypsilanti, Michigan starting in the 1960s ([Highscope, 2022](#)) would be insightful and valuable.

As well, there is a continued need to conceptualize assessment and evaluation broadly, beyond just scores on fairly narrow measures of language and literacy. For example, relatively little is known about whether programs change families' literacies practices at home. That is, do families maintain vernacular literacies or do their practices become school-like, does the amount of reading and writing and other forms of literacies that families engage in change as they participate, do families' conceptions or definitions change and if so, how, and do families take up more critical stances in their literacies? Finally, while some researchers have documented adult participants' perspectives or views of the program, we know relatively little about children's views of their experiences.

The research on family literacy programs tends to focus on cognitive outcomes for children. However, some researchers have reported affective and social benefits such as parents feeling more included in the school and community, parents' increased feelings of self-efficacy, and greater understanding of families, their context, and cultural practices on the part of teachers. We believe that programs that are authentically strengths-based can make important contributions to family-school relationships and to family life; thus assessment and evaluation need a greater focus in these areas.

While family literacy programs continue to be a contested phenomenon in some quarters, they continue to operate in many areas of the world and there is still considerable interest in them. As we indicated, developers and facilitators of family literacy programs have attempted to address the concerns and issues and to make programs more culturally and socially responsive, with varying degrees of success. Indeed, we see a need for ongoing assessment/evaluation of programs with an aim to improving them to meet the aspirations and needs identified by families and communities. Just over two decades ago, Hendrix (2000) asked the question, "Family Literacy Education—Panacea or False Promise?" Concurring with Toso (2012), we conclude that family literacy programs are neither, but there is still much to learn about them.

References

- American Academy of Pediatrics Committee on Public Education, 1999. Media education. *Pediatrics* 104 (2), 341–343.
- Anderson, J., Anderson, A., 2021a. "Biggest thing is saying in English and Punjabi, too": working with immigrant and refugee families and communities in a bilingual family literacy program. *Read. Psychol.* 42 (8), 899–927.
- Anderson, J., Anderson, A., 2021b. Some lessons learned from working with diverse children, families and communities: looking back, looking forward. In: Li, G., Anderson, J., Hare, J., McTavish, M. (Eds.), *Superdiversity and Teacher Education: Supporting Teachers in Working With Culturally, Linguistically, and Racially Diverse Students, Families, and Communities*. Routledge, New York, pp. 252–267.
- Anderson, J., Anderson, A., Lynch, J., Shapiro, J., 2003. Storybook reading in a multicultural society: critical perspectives. In: van Kleeck, A., Stahl, S.A., Bauer, E.B. (Eds.), *On Reading Books to Children: Parents and Teachers*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 203–230.
- Anderson, J., Anderson, A., Friedrich, N., Kim, J., 2010. Taking stock of family literacy: some contemporary perspectives. *J. Early Child. Lit.* 10, 33–53.
- Anderson, J., Friedrich, N., Kim, J., 2011. *Implementing a Family Literacy Program With Immigrant and Refugee Families: The Case of Parents As Literacy Supporters (PALS)*. Decoda Literacy Solutions, Vancouver, BC. Available from: <http://www.decoda.ca/wp-content/uploads/PALS-in-Immigrant-Communities-Research-Report-Feb-2011-2.pdf>.
- Anderson, A., Anderson, J., Teichert, L., 2013. Through a rear-view mirror: families retrospectively evaluate a family literacy program twenty years later. *Sch. Community J.* 23 (2), 33–53.
- Anderson, J., Friedrich, N., Morrison, F., Teichert, L., 2016. "Now he knows that there are two kinds of writing, two kinds of reading": insights and issues in working with immigrant and refugee families and communities in a bilingual family literacy program. In: Anderson, A., Anderson, J., Hare, J., McTavish, M. (Eds.), *Language, Learning and Culture in Early Childhood: Home, School and Community Contexts*. Routledge, New York, pp. 82–102.
- Anderson, J., Anderson, A., Friedrich, N., Teichert, L., 2017. "You guys should offer the program more often!": some perspectives from working alongside immigrant and refugee families in a bilingual family literacy program. In: McLachlan, C., Arrow, A. (Eds.), *Literacy in the Early Years—Reflections on International Research and Practice*. Springer, New York, pp. 63–78.
- Anderson, M., Auxier, B., Nolan, H., 2020. *Parenting Children in the Age of Screens*. Pew Research Center. Available from: <https://www.pewresearch.org/internet/2020/07/28/childrens-engagement-with-digital-devices-screen-time/>.
- Aubrey, C., Dahl, S., 2014. The confidence and competence in information and communication technologies of practitioners, parents and young children in the Early Years Foundation Stage. *Early Years* 34 (1), 94–108. <https://doi.org/10.1080/09575146.2013.792789>
- Auerbach, E., 1989. Toward a social-contextual approach to family literacy. *Harv. Educ. Rev.* 59, 165–181.
- Baghban, M., 1984. *Our Daughter Learns to Read and Write: A Case Study From Birth to 3*. International Reading Association, Newark, DEL.
- Barron, B., Levinson, A.M., 2018. Media as a catalyst for children's engagement in learning at home and across settings. In: Gee, E., Takeuchi, L., Wartella, E. (Eds.), *Children and Families in the Digital Age: Learning Together in a Media Saturated Culture*. Routledge, New York, pp. 17–36.
- Becker, M.L., 2021. Unboxing care: constructions of happiness and gratitude in a Brazilian transnational family's video production in the United States. *J. Early Child. Lit.* <https://doi.org/10.1177/14687984211030668>. Advanced Online.
- BeLiFs, 2021. *Becoming Literate in Faith Settings*. BeLiFs. Available from: <http://www.belifs.co.uk/>.
- Bissex, G., 1980. *GYNS at Wrk: A Child Learns to Read and Write*. Harvard University Press, Cambridge, MA.
- Bittman, M., Rutherford, L., Brown, J., Unsworth, L., 2012. Digital natives? New and old media and children's language acquisition. *Fam. Matters* 91, 18–26.
- Boyce, L.K., Innocenti, M.S., Roggman, L.A., Norman, V.K.J., Ortiz, E., 2010. Telling stories and making books: evidence for an intervention to help parents in migrant Head Start families support their children's language and literacy. *Early Educ. Dev.* 21 (3), 343–371.
- Brooks, G., Pahl, K., Pollard, A., Rees, F., 2008. *Effective and Inclusive Practices in Family Literacy, Language and Numeracy: A Review of Programmes and Practice in the UK and Internationally*. CfBT Education Trust, Reading.
- Busch, G., 2018. How families use video communication technologies during intergenerational Skype sessions. In: Danby, S., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 17–32.
- Chaudron, S., Di Gioia, R., Gemo, M., 2018. *Young Children (0–8) and Digital Technology: A Qualitative Study Across Europe*. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2760/294383>.
- China, A., Robins, S., 1996. "We can all sing, but we can't all talk": literacy brokers and *tsotsi* gangsters in a Cape Town shantytown. In: Prinsloo, M., Breier, M. (Eds.), *The Social Uses of Literacy: Theory and Practice in Contemporary South Africa*. Sached Books, Bertsham, South Africa, pp. 157–176.
- Clark, M., 1976. *Young Fluent Readers: What Can They Teach Us?* Heinemann, London.
- Clay, M.M., 1966. *Emergent Reading Behaviour*. Ph.D. thesis. University of Auckland.
- Clay, M.M., 1975. *What Did I Write? Beginning Writing Behaviour*. Heinemann, Auckland, NZ.

- Clay, M., 1993. Always a learner: a fable. *Read. Today* 3, 10.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, second ed. Lawrence Erlbaum, Hillsdale, NJ.
- Cornell, S., Lauricella, A., Wartella, E., 2015. Parental co-use of media technology with their young children in the USA. *J. Child. Media* 9 (1), 5–21. <https://doi.org/10.1080/17482798.2015.997440>.
- Coquitlam Public Library, 2021. *Virtual Parent-Child Mandarin Mother Goose*. Coquitlam Public Library. Available from: <https://coquitlam.libnet.info/event/4386491>.
- Crooks, S., 2019. Robust respect: de-centering deficit thinking in family literacy work. *Lang. Lit.* 21 (3), 64–78.
- Cuban, S., Hayes, E., 1996. Women in family literacy programs: a gendered perspective. *N. Dir. Adult Cont. Educ.* 70, 5–16.
- Davidson, C., Danby, S.J., Given, L.M., Thorpe, K., 2018. Producing contexts for young children's digital technology use: web searching during adult-child interactions at home and preschool. In: Danby, S.J., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 65–82.
- Davidson, C., 2009. Young children's engagement with digital texts and literacies in the home: pressing matters for the teaching of English in the early years of schooling. *Engl. Teach. Pract. Critiq.* 8 (3), 36–54.
- Davidson, C., 2010. "Click on the big red car": the social accomplishment of playing a Wiggles computer game. *Convergence* 16 (4), 375–394. <https://doi.org/10.1177/1354856510375526>.
- Dias, P., Brito, R., Ribbens, W., Daniela, L., Rubene, Z., Dreier, M., Gemo, M., Di Gioia, R., Chaudron, S., 2016. The role of parents in the engagement of young children with digital technologies: exploring tensions between rights of access and protection, from "Gatekeepers" to "Scaffolders". *Glob. Stud. Child.* 6 (4), 414–427. <https://doi.org/10.1177/2043610616676024>.
- Durkin, D., 1966. *Children Who Read Early*. Teachers College Columbia, New York.
- Ferreiro, E., Teberosky, A., 1982. *Literacy Before Schooling*. Heinemann Educational Books, Portsmouth, NH.
- Fikrat-Wevers, S., van Steensel, R., Arends, L., 2021. Effects of family literacy programs on the emergent skills of children from low-SES families: a meta-analysis. *Rev. Educ. Res.* 91 (4), 577–613.
- First Nations Education Steering Committee, 2022. *First Peoples Principles of Learning*. Available from: <http://www.fnesc.ca/first-peoples-principles-of-learning/>.
- Fishman, A., 1987. Literacy and cultural context: a lesson from the Amish. *Lang. Arts* 54 (8), 842–854.
- Fleer, M., 2014. The demands and motives afforded through digital play in early childhood activity settings. *Learn. Cult. Soc. Interact.* 3 (3), 202–209. <https://doi.org/10.1016/j.lcsi.2014.02.012>.
- Flewitt, R., Clark, A., 2020. Porous boundaries: reconceptualising the home literacy environment as a digitally networked space for 0–3 year olds. *J. Early Child. Lit.* 20 (3), 447–471. <https://doi.org/10.1177/1468798420938116>.
- Gadsden, V., 2012. Father involvement and family literacy. In: Wasik, B. (Ed.), *Handbook of Family Literacy*, second ed. Routledge, pp. 151–165.
- Gear, A., 2013. *Expanding the Circle: Collaborative Research to Create Culturally Responsive Family Literacy Programming*. MA thesis. University of British Columbia.
- Gee, E., Takeuchi, L., Wartella, E., 2018. Introduction. In: Gee, E., Takeuchi, L., Wartella, E. (Eds.), *Children and Families in the Digital Age: Learning Together in a Media Saturated Culture*. Routledge, New York, pp. 1–14.
- Gillen, J., Aliagas, C., Bar-lei, Y., Flewitt, R., Jorge, A., Kumpulainen, K., Marsh, J., Matsumoto, M., Morgade, M., Pacheco, R., Poveda, D., Sairanen, H., Sandberg, H., Scott, F., Sjöberg, U., Sundin, E., Tigane, I., Tomé, V., 2018. *A Day in the Digital Lives of Children Aged 0–3*. Summary Report by DigiLiTEY ISCH COST Action IS1410 Working Group 1 "Digital Literacy in Homes and Communities."
- Given, L.M., Winkler, D.C., Willson, R., Davidson, C., Danby, S., Thorpe, K., 2016. Watching young children "play" with information technology: everyday life information seeking in the home. *Libr. Inf. Sci. Res.* 38 (4), 344–352. <https://doi.org/10.1016/j.lisr.2016.11.007>.
- Gregory, E., 2001. Sisters and brothers as language and literacy teachers: synergy between siblings playing and working together. *J. Early Child. Lit.* 1 (3), 301–322.
- Gregory, E., 2005. Guiding lights: siblings as literacy teachers in a multi-lingual community. In: Anderson, J., Kendrick, M., Rogers, T., Smythe, S. (Eds.), *Portraits of Literacy Across Families, Communities and Schools: Intersections and Tensions*. Routledge, New York, pp. 21–39.
- Hannon, P., 2003. Family literacy programmes. In: Hall, N., Larson, J., Marsh, J. (Eds.), *Handbook of Early Childhood Literacy*. Sage, London, pp. 99–111.
- Harrison, E., McTavish, M., 2018. "IBabies: infants" and toddlers' emergent language and literacy in a digital culture of iDevices. *J. Early Child. Lit.* 18 (2), 163–188. <https://doi.org/10.1177/1468798416653175>.
- Heath, S., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, New York.
- Hendrix, S., 2000. Family literacy education: panacea or false promise. *J. Adolesc. Adult Lit.* 43 (4), 338–346.
- Highscope, 2022. *Perry Preschool Project*. Available from: <https://highscope.org/perry-preschool-project/>.
- Hirst, K., Hannon, P., Nutbrown, C., 2010. Effects of a pre-school bilingual family literacy program. *J. Early Child. Lit.* 10, 183–208.
- Huh, Y.J., 2015. Making sense of gender from digital game play in three-year-old children's everyday lives: an ethnographic case study. *J. Comp. Res. Anthropol. Sociol.* 6 (1), 155–170.
- International Bureau of Education, 2021. *Multiple Literacies*. International Bureau of Education. Available from: <http://www.ibe.unesco.org/en/glossary-curriculum-terminology/m/multiple-literacies>.
- Kagitcibasi, C., 2005. *Continuing Effects of Early Intervention in Adult Life: Preliminary Findings of Turkish Early Enrichment Project*. Second follow-up study. Mother Child Education Foundation, Istanbul.
- Kendrick, M., Namazzi, E., 2017. Family language practices as emergent policies in child-headed households in rural Uganda. In: Macalister, J., Mirvahedi, S.H. (Eds.), *Family Language Policies in a Multilingual World: Opportunities, Challenges, and Consequences*. Routledge, London, UK, pp. 56–73.
- Kervin, L., Verenikina, I., Rivera, C., 2018. Digital play and learning in the home: families' perspective. In: Danby, S., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 117–130.
- Kervin, L., 2016. Powerful and playful literacy learning with digital technologies. *Aust. J. Lang. Lit.* 39 (1), 64–73.
- Koohi, R., Sajedi, F., Movallali, G., Dann, M., Soltani, P., 2017. Faranak Parent-Child Mother Goose Program: impact on mother-child relationship for mothers of preschool hearing impaired children. *Iran. Rehabil. J.* 14 (4), 201–210.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, London.
- Kumpulainen, K., Sairanen, H., Nordstrom, A., 2020. Young children's digital literacy practices in the sociocultural contexts of their homes. *J. Early Child. Lit.* 20 (3), 472–499. <https://doi.org/10.1177/1468798420925116>.
- Ling, D., Tibbetts, G., Scharfe, E., 2017. Once upon a time: lessons learned from the benefits of parent-child mother Goose. *Child Welfare* 95 (2), 9–31.
- Livingstone, S., Blum-Ross, A., 2020. *Parenting for a Digital Future: How Hopes and Fears About Technology Shape Children's Lives*. Oxford University Press, New York.
- Lynch, J., Prins, E., 2021. *Teaching and Learning About Family Literacy and Family Literacy Programs*. Routledge, New York.
- Lytra, V., Ilankuberan, A., 2020. Syncratising ways of doing, seeing and becoming in children's faith-inspired text-making and conversations around texts at home. *Lang. Intercult. Commun.* 20 (5), 433–449.
- Lytra, V., Volk, D., Gregory, E., 2016. *Navigating Language, Literacies and Identities: Religion in Young Lives*. Routledge, New York.
- Mace, J., 1998. *Playing With Time: Mothers and the Meaning of Literacy*. UCL Press, London.
- Macleod, F., 2008. Why fathers are not attracted to family learning groups? *Early Child. Dev. Care* 178 (7), 773–783.
- Marsh, J., Plowman, L., Yamada-Rice, D., Bishop, J., Lahmar, J., Scott, F., Davenport, A., Davis, A., French, S., Piras, K., Thornhill, M., Robinson, S., Winter, P., 2015. *Exploring Play and Creativity in Pre-schoolers' Use of Apps: Final Project Report*. Technology and Play, Norton, England. Available from: www.techandplay.org.
- Marsh, J., Plowman, L., Yamada-Rice, D., Bishop, J., Scott, F., 2016. Digital play: a new classification. *Early Years* 36 (3), 242–253. <https://doi.org/10.1080/09575146.2016.1167675>.

- Marsh, J., Hannon, P., Lewis, M., Ritchie, L., 2017a. Young children's initiation into family literacy practices in the digital age. *J. Early Child. Res.* 15 (1), 47–60. <https://doi.org/10.1177/1476718X15582095>.
- Marsh, J., Mascheroni, G., Carrington, V., Arnadottir, H., Brito, R., Dias, P., Kupianen, R., Trueltzsch-Wijnen, C., 2017b. *The Online and Offline Digital Literacy Practices of Young Children: A Review of the Literature*. Cost Action IS1410. Available from: http://digitaleu.White_Paper24copy_internetssegura.pt.
- Marsh, J., 2004. The techno-literacy practices of young children. *J. Early Child. Res.* 2 (1), 51–66.
- Marulis, L., Neuman, S., 2010. The effects of vocabulary intervention on young children's word learning: a meta-analysis. *Rev. Educ. Res.* 80 (3), 300–335.
- McPherson, K., 2011. Young Children Learning With New Literacies. PhD Thesis. University of Victoria.
- Moll, L., Greenberg, J., 1990. Creating zones of possibilities: combining social contexts for instructions. In: Moll, L. (Ed.), *Vygotsky and Education*. Cambridge University Press, Cambridge, MA, pp. 319–348.
- Murphy, S., Heady, M., 2020. The role of digital technology in teen mothers' and their children's literacy. *J. Early Child. Lit.* 20 (4), 755–782. <https://doi.org/10.1177/1468798418783326>.
- Neumann, M., 2014. An examination of touch screen tablets and emergent literacy in Australian pre-school children. *Aust. J. Educ.* 58 (2), 109–122. <https://doi.org/10.1177/0004944114523368>.
- Ninio, A., 1983. Joint book reading as a multiple vocabulary acquisition device. *Dev. Psychol.* 19 (3), 445–451.
- Nutbrown, C., Clough, P., Stammers, L., Emblin, E., Alston-Smith, S., 2019. Family literacy in prisons: fathers' engagement with their young children. *Res. Pap. Educ.* 34 (2), 169–191.
- Nutbrown, C., 1997. *Recognising Early Literacy Development: Assessing Children's Achievements*. Paul Chapman, London.
- OFCOM, 2016. *Children and Parents: Media Use and Attitudes Report*. OFCOM. Available from: <https://www.ofcom.org.uk/research-and-data/media-literacy-research/children/children-parents-nov16>.
- O'Hara, M., 2011. Young children's ICT experiences in the home: some parental perspectives. *J. Early Child. Res.* 9 (3), 220–231. <https://doi.org/10.1177/1476718X10389145>.
- Orellana, M., Reynolds, J., Dornier, L., Meza, M., 2003. "In other words": translating and paraphrasing as a family literacy practice in immigrant households. *Read. Res. Q.* 38 (1), 12–34.
- Ouane, A., 2008. Foreword. In: Desmond, S., Elfert, M. (Eds.), *Family Literacy: Experiences From Africa and Around the World*. UNESCO Institute for Lifelong Learning, Hamburg, pp. v–vi. Available from: <https://unesdoc.unesco.org/ark:/48223/pf0000180481>.
- Ozturk, G., Ohi, S., 2019. What do they do digitally? Identifying the home digital literacy practices of young children in Turkey. *Early Years*. <https://doi.org/10.1080/09575146.2019.1702925>.
- Palaiologou, I., 2016. Children under five and digital technologies: implications for early years pedagogy. *Eur. Early Child. Educ. Res. J.* 24 (1), 5–24.
- Panofsky, C.P., 2000. Examining the research narrative in early literacy: the case of parent-child book reading activity. In: Gallego, M.A., Hollingsworth, S. (Eds.), *What Counts as Literacy: Challenging the School Standard*. Teachers College Press, New York, pp. 190–212.
- Paratore, J., Yaden, D., 2011. Family literacy on the defensive: the defunding of even start—omen or opportunity? In: Lapp, D., Fisher, D. (Eds.), *Handbook of Research on Teaching the English Language Arts*, third ed. Routledge, New York, pp. 90–96.
- Parent Child Mother Goose Program, 2021. *Our History: Helping Families Bond*. National Parent-Child Mother Goose Program. Available from: [HISTORY_The-Parent-Child-Mother-Goose-Program-nationalpcmgp.ca](https://www.history.org/learn/parent-child-mother-goose-program).
- Perry, K.H., 2009. Genres, contexts, and literacy practices: literacy brokering among Sudanese refugee families. *Read. Res. Q.* 44 (3), 256–276.
- Perry, K.H., 2014. "Mama, sign this note": young refugee children's brokering of literacy practices. *Lang. Arts* 91 (5), 313–325.
- Plowman, L., McPake, J., Stephen, C., 2008. Just picking it up? Young children learning with technology at home. *Camb. J. Educ.* 38 (3), 303–319. <https://doi.org/10.1080/03057640802287564>.
- Plowman, L., McPake, J., Stephen, C., 2010. The technologisation of childhood? Young children and technology in the home. *Child. Soc.* 24, 63–74. <https://doi.org/10.1111/j.1099-0860.2008.00180.x>.
- Plowman, L., Stevenson, O., Stephen, C., McPake, J., 2012. Preschool children's learning with technology at home. *Comput. Educ.* 59, 30–37. <https://doi.org/10.1016/j.compedu.2011.11.014>.
- Plowman, L., McPake, J., Stephen, C., 2013. Pre-school children creating and communicating with digital technologies in the home. *Br. J. Educ. Technol.* 44 (3), 421–431. <https://doi.org/10.1111/j.1467-8535.2012.01323.x>.
- Poveda, D., Matsumoto, M., Sundin, E., Sandberg, H., Aliagas, C., Gillen, J., 2020. Space and practices: engagement of children under 3 with tablets and televisions in homes in Spain, Sweden and England. *J. Early Child. Lit.* 20 (3), 500–523. <https://doi.org/10.1177/1468798420923715>.
- Prins, E., Toso, B.W., Schafft, K., 2009. "It feels like a little family to me": social interaction and support among women in adult education and family literacy. *Adult Educ. Q.* 59 (4), 335–352.
- Prins, E., Stickel, T., Kaiper-Marquez, A., 2020. Incarcerated fathers' experiences in the Read to Your Child/Grandchild Program: supporting children's literacy, learning and education. *J. Prisons Educ.* 6 (2), 168–188.
- Purcell-Gates, V., 1996. Stories, coupons, and the "TV Guide": relationships between home literacy experiences and children's emergent literacy knowledge. *Read. Res. Q.* 31 (4), 406–428.
- Purcell-Gates, V., 2000. Family literacy. In: Kamil, M.L., Mosenthal, P.B., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research, vol III*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 853–870.
- Reay, D., 1998. *Class Work: Mothers' Involvement in Their Children's Primary Schooling*. Routledge, London.
- Reese, E., 2012. The tyranny of shared book reading. In: Suggate, S., Reese, E. (Eds.), *Contemporary Debates in Childhood Education and Development*. Routledge, London, pp. 59–68.
- Reid, K., Hresko, W., Hammill, D., 1989. *Test of Early Reading Ability III*. Pro-Ed, Austin, TX.
- Reyes, L., Torres, M., 2007. Decolonizing family literacy in a culture circle: reinventing the family literacy educator's role. *J. Early Child. Lit.* 7 (1), 73–94.
- Rideout, V., Robb, M., 2020. *The Common Sense Census: Media Use by Kids Age Zero to Eight*. Common Sense Media, San Francisco. Available from: https://www.common SenseMedia.org/sites/default/files/uploads/research/2020_zero_to_eight_census_final_web.pdf.
- Rideout, V., 2013. *Zero to Eight: Children's Media Use in America*. Common Sense Media, San Francisco. Available from: <https://www.common SenseMedia.org/research/past-research-reports>.
- Rideout, V., 2017. *The Common Sense Census: Media Use by Kids Age Zero to Eight*. Common Sense Media, San Francisco. Available from: <https://www.common SenseMedia.org/research/the-common-sense-census-media-use-by-kids-age-zero-to-eight-2017>.
- Rizk, J., 2020. "Well, that just comes with being a mama": the gendered nature of family literacy programs. *Early Child. Educ. J.* 48, 393–404.
- Roach O'Keefe, A., O'Keefe, E., 2020. Remixing digital play in the early years: a child-parent collaboration. In: Schamroth Abrams, S., Schaefer, M., Ness, D. (Eds.), *Child-Parent Research Reimagined*. Brill Sense, Leiden, Netherlands, pp. 108–139.
- Rodriguez Brown, F., 2004. Project FLAME: a parent support family literacy model. In: Wasik, B. (Ed.), *Handbook of Family Literacy*. Lawrence Erlbaum, Mahwah, NJ, pp. 213–229.
- Rodriguez Brown, F., 2008. *The Home-School Connection: Lessons Learned in a Culturally and Linguistically Diverse Community*. Routledge, New York.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press, Oxford, UK.
- Roos, C., Olin-Scheller, C., 2018. Digital participation among children in rural areas. In: Danby, S.J., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 49–63.
- Sadiq, A., 2020. *Forgotten Refugees: Understanding the Language and Literacy Practices of Afghan Refugees in Pakistan*. PhD thesis. University of British Columbia.

- Schlembach, S., Johnson, M.L., 2014. Parents' beliefs, attitudes and behaviors concerning their young children's screen media use. *Dialog* 17 (2), 95–104.
- Scriven, B., Edwards-Groves, C., Davidson, C., 2018. A young child's use of multiple technologies in the social organization of a pretend telephone conversation. In: Danby, S.J., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 267–284.
- Sibanda, R., Kajee, L., 2019. Home as a primary space: exploring out-of-school literacy practices in early childhood education in a township in South Africa. *S. Afr. J. Child. Educ.* 9 (1), 1–10.
- Souza, A., Barradas, O., Woodham, M., 2016. Easter celebrations at home: acquiring symbolic knowledge and constructing identities. In: Lytra, V., Volk, D., Gregory, E. (Eds.), *Navigating Languages, Literacies and Identities*. Routledge, New York, pp. 39–55.
- Stephen, C., Stevenson, O., Adey, C., 2013. Young children engaging with technologies at home: the influence of family context. *J. Early Child. Res.* 11 (2), 149–164. <https://doi.org/10.1177/1476718X12466215>.
- Street, B., 1995. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Routledge, New York.
- Swain, J., Brooks, G., Bosley, S., 2014. The benefits of family literacy provision for parents in England. *J. Early Child. Res.* 12 (1), 77–91.
- Takeuchi, L., Elberse, B., 2018. Daddy loves dora and mama loves drama: ethnic media as intergenerational boundary objects. In: Gee, E., Takeuchi, L., Wartella, E. (Eds.), *Children and Families in the Digital Age: Learning Together in a Media Saturated Culture*. Routledge, New York, pp. 153–173.
- Takeuchi, L., Stevens, R., 2011. *The New Co-viewing: Designing Learning Through Joint Media Engagement. A Report of the Joan Ganz Cooney Center at Sesame Workshop and LIFE Center*. Joan Ganz Cooney Center.
- Taylor, D., Dorsey-Gaines, C., 1988. *Growing Up Literate: Learning From Inner-City Families*. Heinemann, Exeter, NH.
- Taylor, K.H., Silvius, D., Stevens, R., 2018. Collecting and connecting: intergenerational learning with digital media. In: Gee, E., Takeuchi, L., Wartella, E. (Eds.), *Children and Families in the Digital Age: Learning Together in a Media Saturated Culture*. Routledge, New York, pp. 56–75.
- Taylor, D., 1983. *Family Literacy: Young Children Learning to Read and Write*. Heinemann, Exeter, NH.
- Teale, W., Sulzby, E., 1986. *Emergent Literacy: Reading and Writing*. Ables Publishing, Norwood, NJ.
- Teichert, L., Anderson, A., 2014. "I don't even know what blogging is": the role of digital media in a five-year old girl's life. *Early Child. Dev. Care* 184 (11), 1677–1691.
- Teichert, L., 2017. To digital or not to digital: how mothers are navigating the digital world with their young children. *Lang. Lit.* 19 (1), 63–76.
- Teichert, L., 2019. Young children using digital technology: the case of Belle. In: Jagger, S. (Ed.), *Early Years Education and Care in Canada: A Historical and Philosophical Overview*. Canadian Scholars' Press, Toronto, pp. 275–297.
- Teichert, L., 2020. Negotiating screen time: a mother's struggle of "no screen time" with her infant son. *J. Early Child. Lit.* 20 (3), 524–550.
- Terrett, G., White, R., Spreckley, M., 2012. A preliminary evaluation of the Parent–Child Mother Goose Program in relation to children's language and parenting stress. *J. Early Child. Res.* 11 (1), 16–26.
- Thomas, A., Skage, S., 1998. Overview of perspectives of effective practice. In: Thomas, A. (Ed.), *Family Literacy in Canada: Profiles of Effective Practices*. Soleil Publishing, Welland, ON, pp. 5–24.
- Toso, B., 2012. Educational and mothering discourses and learner goals: Mexican immigrant women enacting agency in a family literacy program. In: *Goodling Institute for Research in Family Literacy Research Brief #8*. Pennsylvania State University, University Park, PA.
- van Steensel, R., McElvaney, N., Kurvers, J., Herppich, S., 2011. How effective are family literacy programs? Results of a meta-analysis. *Rev. Educ. Res.* 81 (11), 69–96.
- Vancouver Public Library Foundation, 2021. *Man in the Moon* [online] Vancouver Public Library.
- Volk, D., de Acosta, M., 2003. Reinventing texts and contexts: syncretic literacy events in young Puerto Rican children's homes. *Res. Teach. Engl.* 38 (1), 8–48.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- White-Kaulait, M., 2007. Reflections on Native American reading: a seed, a tool, and a weapon. *J. Adolesc. Adult Lit.* 50 (7), 560–569.
- Wolfe, S., Flewitt, R., 2010. New technologies, new multimodal literacy practices and young children's metacognitive development. *Camb. J. Educ.* 40 (4), 387–399. <https://doi.org/10.1080/0305764X.2010.526589>.
- Wong, S., 2015. Mobile digital devices and preschoolers' home multiliteracy practices. *Lang. Lit.* 17 (2), 75–90.
- Wood, D., Bruner, J., Ross, G., 1976. The role of tutoring in problem solving. *JCPP* 17, 89–100.
- Wren, Y., McCormack, J., Masso, S., McLeod, S., Baker, E., Crowe, K., 2018. Digital tools to support children's speech and language skill. In: Danby, S.J., Fleer, M., Davidson, C., Hatzigianni, M. (Eds.), *Digital Childhoods: Technologies and Children's Everyday Lives*. Springer, Singapore, pp. 235–251.
- Yip, J., Gonzalez, C., Katz, V., 2018. Children of immigrants' experiences in online information brokering. In: Gee, E., Takeuchi, L., Wartella, E. (Eds.), *Children and Families in the Digital Age: Learning Together in a Media Saturated Culture*. Routledge, New York, pp. 137–152.

Toward culturally relevant literacies with children and families of color

Patricia A. Edwards^a, Patriann Smith^b, and Jonda C. McNair^c, ^a Michigan State University, Department of Teacher Education, East Lansing, MI, United States; ^b University of South Florida, Department of Language, Literacy, Exceptional Educational, Ed.D., & Physical Education, Tampa, FL, United States; and ^c The Ohio State University, Department of Teaching and Learning, Columbus, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

The origins of family literacy programs	180
Importance of generational influences and connectivity to community	181
Reenvisioning family literacy for families of color	181
Conceptions of family literacy	182
Approaches to family literacy implementation in the United States	183
Specific family literacy programs	183
Undergirding philosophies of family literacy programs	183
Debates about family literacy in the United States	184
Family-school relationships	184
Early childhood focus	184
Focus on storybook reading	184
Absence of longitudinal focus	185
Responses to family literacy debates	185
Approaches to family literacy across the globe	186
Insights from international family literacy programs	186
Culturally relevant literacies of families of color	190
Ethnographic studies	190
Programmatic tensions	190
Family “literacies” vs. school-valued literacy	191
Parent stories	191
Re-envisioning family literacy for families of color	192
Considerations for family literacy programs for families of color	192
Reconceptualizing family literacies in early childhood development	192
Leveraging equitable literacy practices in family literacy for families of color	193
Conclusion	194
References	194

The origins of family literacy programs

The roots of family literacy as an educational method come from the belief that the parent is the child’s first teacher (Edwards, 2016). The term “family literacy” was first used over 30 years ago to describe discourse within the family (Taylor, 1983). Family literacy has origins in the United States, Canada, and South Africa, and since 1983 has emerged as a new and distinct field of inquiry, made prominent during Barbara Bush’s time as Second Lady while her husband was the then Vice President and President of the United States (White House, 2009). Barbara Bush called family literacy “the most important issue we have” (White House, 2009). Her dedication to eliminating the generational cycle of illiteracy in America by supporting programs where parents and their young children were able to learn together significantly affected the advancement of families of color. US legislative definitions such as the *National Literacy Act of 1991*, spurred on by the war on poverty (National Public Radio, 2004), soon followed, defining literacy as “an individual’s ability to read, write and speak in English and compute and solve problems at levels of proficiency necessary to function in the job and in society, to achieve one’s goals and to develop one’s knowledge and potential” (Library of Congress, 1991, np).

By 1997, when *H.R. 1385 was passed by the House* (Library of Congress, 1998), the United States saw an increasing recognition of the need to move beyond an emphasis on children to a more explicit focus on parents and families, a large number of whom were families of color. *Family literacy services* thus came to connote “services provided to participants on a voluntary basis, of sufficient intensity in terms of hours, and of sufficient duration, to make sustainable changes in a family (such as eliminating or reducing welfare dependency)” (Association of Centers for the Study of Congress, 2021, np). The United States then saw family literacy emerge to integrate interactive literacy activities between parents and their children, equipping them to partner with their children in learning, providing parent literacy training for economic self-sufficiency, and offering appropriate instruction for children of parents receiving parent literacy services. School programs such as Reading Is Fundamental (RIF) (2021) and Title I, which were originally designed for school-age children, also incorporated programs for families, and Head Start demonstrated that parents’ participation produced greater school success than programs without parent involvement. Initiatives such as the PACE/Kenan

project in Kentucky developed a successful model for teaching literacy skills to both parents and children, in which content skills were supplemented by parent education and parent-child activity time. In addition, the federal legislation for family literacy, based on the Kentucky model and shepherded by Representative William Goodling from Pennsylvania, ushered in Even Start in 1988 ([Even Start Family Literacy Program, 2014](#)).

Even Start Family Literacy further reinforced the need to break the cycle of poverty and illiteracy, which remained a concern for families of color, by working to improve the educational opportunities of the nation's low-income families through integrating early childhood education, adult literacy or adult basic education, and parenting education into a unified family literacy program. Through cooperative projects that built on existing community resources to create a new range of services, *Even Start* promoted achievement of the National Education Goals ([Legal Information Institute, 1994](#)) and assisted children and adults from low-income families to achieve challenging state content standards and student performance standards. In doing so, *Even Start* functioned to advance a national agenda dedicated to a focus on increasing literacy opportunities for children, a large majority of whom were from families of color ([U.S. Department of Education, 2004](#)).

It was no wonder then, that the US Congress, via the 1998 *Adult Education and Family Literacy Act*, created a definition for adult literacy, describing it as "an individual's ability to read, write, and speak in English, compute and solve problems, at levels of proficiency necessary to function on the job, in the family, and in society" ([Literacy Information and Communication System, 2021](#), np; [U.S. Department of Education, 1998](#)). With this definition, family literacy education claimed to experience a shift from traditionally preparing children to read, write, speak English and use computation skills to enabling children as well as adults to apply these skills in the workplace and community to use information to solve problems.

Importance of generational influences and connectivity to community

On the global scale, these variations in terms of how family literacy was emphasized, both legislatively and nationally in the United States, were interrogated by the *International Reading Association's Family Literacy Commission* ([Morrow et al., 1993](#)). The then International Reading Association (IRA) perceived family literacy as encompassing the ways parents, children, and extended family members use literacy at home and in their community. Focusing more on honoring the literacy practices of families (of color) as opposed to school-valued literacy practices, IRA established that family literacy may be initiated purposefully by a parent or may occur spontaneously as parents and children go about the business of their daily lives ([Morrow, 1994](#)). Deviating slightly from this international view of family literacy, and aligned more closely with US legislation, the United Nations Educational Scientific and Cultural Organization (UNESCO) then described family literacy as "the development of the literacy, numeracy and language skills of both children and adults" across families and generations ([UNESCO Institute for Lifelong Learning, 2017](#), p. 3). In conceiving of family literacy across international domains, UNESCO also advocated the idea of family learning that functions as a "broader learning activities including, but not limited to literacy" ([UNESCO Institute for Lifelong Learning, 2017](#), p. 3). Family learning therefore represented "a transformative approach that works across generations and between institutions, breaking down barriers between home, school, and community" ([UNESCO Institute for Lifelong Learning, 2017](#), p. 1).

Reenvisioning family literacy for families of color

While the emphases of family literacy education across domains appeared to differ, and despite no clear indication that the international focus of IRA (now the International Literacy Association) and UNESCO were geared toward families of color, it can be argued that the majority of the world's population constitutes families of color ([Ferris, 2015](#); [Race of the World Population, 2022](#)). National and international conceptions of family literacy in the US have thus far located family literacy as a key mechanism for supporting families of color across the globe by operating at the nexus of (a) early childhood, preschool and primary education and (b) adult and community education (see [Hanemann et al., 2017](#); [UNESCO Institute for Lifelong Learning, 2017](#)). In doing so, the focus on family literacy and learning on the international scale is envisioned as a mechanism for impacting educational outcomes, life success, bonding between parents and children, connecting schools, organizations and communities and reducing school drop-outs ([UNESCO Institute for Lifelong Learning, 2017](#)), particularly for a large majority of families of color. Yet, the alignment of UNESCO with legislative notions of family literacy in the US as opposed to conceptions of community and culturally-valued practices of families (of color) by ILA presents a conundrum.

This conundrum places family literacy in the 21st century at an impasse, one which requires a re-envisioning of the underlying ways of enacting and supporting the literacies of families of color. The increasing recognition that children begin to learn literacy prior to formal schooling in the US and across the globe, and that family contexts shape literacy development, has had profound implications for the field of early literacy. Beyond this, for the past 30 years, parent involvement in children's education has been expanding, leading to further transformation in the definition of family literacy. In addition, although family literacy originated as a concept describing the rich and varied ways that families use literacy in home and community settings, it has increasingly become associated with formal programs aimed at improving the literacy of parents and their young children. Notwithstanding, there remains a continued need for family literacy that addresses literate equity for families of color as evidenced in the International Literacy Association's recent *What's Hot in Literacy Report* which highlighted "family engagement in literacy learning" as one of

the top four key responsibilities of literacy educators, and the “variability of family engagement in literacy and learning” as one of the top three barriers to equity in literacy education ILA, 2020 (ILA, 2020, pp. 17–18, 22).

While geared toward many families of color in its focus, intention, and implementation, the evolution of family literacy over the past 30 years has not explicitly identified variations in what a family literacy agenda that explicitly supports families of color might entail. Specifically, given recent indications of the ways in which race and racialized ways of thinking about language continue to undergird literacy, English language arts, and bilingual curriculum (Anya, 2020; Flores, 2020; Nalubega-Booker and Willis, 2020; Smith et al., 2020) of families of color even while these families wrestle with institutional norms and practices to advance their economic and social futures (Rosa and Flores, 2017; Smith et al., 2020), it is reasonable to question the degree to which family literacy education and its corresponding programs have centered and complicated elements such as race in literacy curriculum, assessment, programs, and practices of populations of color.

Family literacy education has arguably created change, ranging from a focus on sustaining informal family interactions among parents and children driven by the need to eliminate intergenerational cycles of poverty to more formal and federally funded programs driven by legislative definitions that advance traditional literacy skills, life skills as well as problem-solving skills deemed necessary to enhance opportunities for both parents and children in families to overcome challenges posed by illiteracy. Yet, a persistent concern lingers that family literacy programs can do much more to balance the development of programs with an adequate emphasis on the community and cultural practices of families of color, a dynamic critical to enhancing teachers’ capacity to support parents and children in and beyond schools, both in the United States and across the globe.

To advance this argument, we first discuss family literacy and its evolution over time, taking into consideration local and global realities of families, and particularly, families of color. We then describe family literacy implementation and programming in the US followed by a discussion of the same across the globe, highlighting the ways in which these initiatives may have potentially deterred the economic or social realities of families of color even while perceived as advancement. Subsequently, we discuss long-standing efforts to examine and support the literacies of families of color, pointing to key indicators that characterize these research and community practices. Finally, we identify key insights gleaned from our discussion by presenting certain culturally relevant literacy practices that can sustain the literacies of families of color and proposing a way forward for the field.

Conceptions of family literacy

In the evolution of family literacy over the past 30 years, it is clear that this term means different things to different people. What it might mean for families of color, however, remains underexplored. *Family literacy*, broadly conceived, refers to the interactions of parents and children using language—talking, playing, exploring, limiting, soothing, explaining, encouraging, and nurturing. With the support of the adults in his or her life, a child learns to navigate his or her world with the help of language, acquiring limits and self-control, making choices, and solving problems, communicating needs to others, developing emotional ties to parents and siblings, and responding to the print environment surrounding him or her. In recent decades, the focus of *emergent literacy* has shifted from learning to read in the first grade to preschool interactions in the home environment and from the first-grade teacher to the parent as first teacher.

We have shown earlier that a second use of the term *family literacy* applies broadly to the federally funded programs, typically originating from the US context, developed to support intergenerational education for at-risk, low-literacy families—programs such as Head Start, Even Start, and Title I (Even Start Family Literacy Program, 2014; U.S. Department of Education, 2004; U.S. Department of Health and Human Services, 2019). The legislation authorizing these programs contains a uniform definition of *family literacy* that has typically entailed four components:

- Adult basic education to improve basic skills, prepare for the (GED) certificates, and to learn workplace skills that leads to economic self-sufficiency;
- Early childhood education for preschool and school-age children to help them prepare for success in school and life experiences;
- Parent education in which parents and caregivers discuss parenting practices and the importance of literacy experiences in the home;
- Parent and child together time (PACT) for adults and children to practice literacy activities together.

While improving their reading, writing and math skills, the assumption was that parents have an opportunity to practice language strategies with their children in areas such as storybook reading, discipline, and play and exploration. These skills are integrated into units arising from family issues, citizenship, and workforce readiness.

Doyle (2012) has argued that three major factors affected such trends in how family literacy was conceived in the United States: (a) developmental theory; (b) research on emergent literacy and parent involvement; and (c) sociopolitical developments. Developmental theory and research on emergent literacy during the 1960s appeared to work together to inform novel approaches to educational practice and social policy. Among the concepts introduced at this time were children functioning as active participants in their learning, the use of environmental supports to improve learning avenues for underserved children, and the presence of theoretical foundations that informed movements surrounding interventions from which the intervention movement developed. Later on, during the 1970s, the use of systems theories reflected how ecological perspectives could potentially clarify the nuances involved in child development which saw an increased emphasis on interventions. And in the 1980s, educational and social fields increasingly addressed “family literacy” as a construct leading to the continued emergence of programs that reflect varied degrees of

intensity, myriad forms of resources and a variety of goals, all dependent on the ways in which political and social shifts were manifested at that time.

Thus, despite the earlier recognition of ecological perspectives and the centering of the child as active participant, the subsequent shift toward intervention and programs superseded the cultural ways of being of children and of families of color, resulting in a current context where school-valued literacy practices positioned continue to be privileged over the community and cultural literacy practices in family literacy education. There is no doubt that parents, children, and communities of color can benefit from family literacy programs. Not only do individual literacy skills of parents and children improve but social skills increase, and families place higher values on education. Parents' expectations of their children change as they learn more about the continuum of child development. Parents become more involved in their children's schools as they better understand new educational approaches and recognize the important role they have as partners with teachers in their children's education.

For most, *family literacy*, whether spontaneous or promoted by formal programs, is a process of incorporating the spoken and written word into meaningful activities within the family unit. This becomes the legacy of language practices that passes from one generation to the next. However, to explore family literacy education as a promise of enhancing literacy opportunity with undue recognition of the cultural wealth of families of color is to inadvertently negate the already-present literacies (Marciano and Watson, 2021) that they possess.

Approaches to family literacy implementation in the United States

Specific family literacy programs

In the United States, family literacy legislation has allowed for more intensive programming known as the *comprehensive* or *four component* family literacy model (National Institute for Literacy: NIFL, 2001). The four components entail academic upgrading for parents, an early childhood development program, parenting and early literacy development, and a parent-child together (PACT) time, in which parents spend quality time with their children and try out new literacy ideas with them. Elements of this model are visible in initiatives such as The Even Start program (described earlier), which has catered to numerous families of color in the United States. They are also visible in the program, Reach Out and Read (ROR) (<http://www.reachoutandread.org/>), which began in Boston in 1992, and involves pediatricians, other professionals, and nonprofessional members of the community, supported numerous parents of color to lay the foundations for literacy during their children's 1st years of life. ROR used the pediatric office or clinic to bring together young families, pediatricians, volunteers, and children's books.

A third program reflecting elements of the comprehensive literacy model is Imagination Library, operating since 1995 and visible in Canada, the United Kingdom, and Australia. Imagination Library functioned as a book giveaway program in the United States, giving a free book per month to children, including children of color, from the time they are born until they begin school with about 1.4 million children are currently registered with the program and over 122 million books have been distributed since inception (Imagination Library, n.d.). Finally, a fourth program is the Laura Bush initiative where public library staff were dedicated to helping children and adults (of color) through a large variety of services and programs that promote family literacy (The Laura Bush Foundation for America's Libraries, 2021). And often overlooked are innovative family literacy programs involving correctional facilities in which people of color continue to be significantly overrepresented in the US, which cater to inmates usually separated from their families and help incarcerated parents develop the literacy of their children through enhancing their own literacy-building and parenting abilities (e.g., Georgetown University family literacy program for inmates of the District of Columbia: D.C. Family Literacy Project).

Undergirding philosophies of family literacy programs

Across such family literacy programs, three philosophical approaches have been identified: (a) the *intervention/prevention* approach; (b) the *multiple literacies* approach and (c) *critical literacy* and *social change theories* approach. Family literacy programs reflecting an *intervention/prevention* approach aim to compensate for perceived inadequacies in parenting behaviors and home literacy activities, believed to negatively affect children's readiness for school (Darling, 1993). A second broad approach to family literacy programming is known as *multiple literacies*, which is informed by anthropological, sociolinguistic, and sociocultural studies. Here, literacy is seen not as a single skill, but rather as a set of practices grounded in social contexts and social roles (Barton, 1994). A third broad approach is associated with critical literacy and social change theories, informed by the work of Freire (1985) and Giroux (1988). These programs present family literacy within a broader context of social inequities that shape power relationships among families, schools, and the broader society.

Evidently, the discussion of family literacy programs in the US thus far reveals a heavy emphasis on the *intervention/prevention* approach, which highlights the ways in which the literacies of color may have been potentially overlooked even while such populations appear to be supported by family literacy programs. As highlighted by Schwartz (1999), family literacy programs undergirded by the intervention/prevention philosophical approach reflect three primary elements: (a) adult education involving instruction surrounding reading, writing, computing, and problem-solving and that considers attention to English as a Second Language, General Educational Development (GED) classes and jobs skills training; parenting education involving the support of families as active participants in their children's education both in the home and school; and early childhood education geared

toward preschool children. For families of color, such an approach is geared toward addressing deficiencies in literacy for adults as well as children by instilling in them the literate practices and norms privileged by schools and the broader society.

Debates about family literacy in the United States

It is no surprise then, that persisting tensions in how family literacy education is proposed broadly for families via an intervention/prevention approach that inadvertently sidelines the needs of families of color have resulted in numerous debates. These debates stem from the pervasive presence of the *intervention/prevention approach* in family literacy programs despite the indication of the potential of approaches such as multiple literacies, critical literacy & social change theories for families of color. The growing popularity of family literacy to address illiteracy, poverty and to advance the opportunities of underserved families in a context of national and global racial crises, exacerbated by a global pandemic, have led such debates to take on new importance. Among those debates are tensions persisting in family-school relationships, the focus on early childhood by family literacy programs, the central role of storybook reading and the lack of a longitudinal focus over time.

Family-school relationships

A key issue in family literacy is the relationship between home and school, and, more specifically, between literacy as valued and practiced in homes and literacy as valued and practiced in schools. Shirley Brice Heath, in 1983, showed how family values and communication patterns contributed to differences among white children and African-American in two very small communities just a few miles apart, and how children in both these small communities differed from children who grew up in the nearby town. Despite efforts made since then, the mismatch between home and school literacy has remained a topic of ongoing concern and debate. For instance, in the US, many children from homes with low-income, and for whom English is a second language, continue to score below their more privileged peers on standardized tests and in overall academic performance (Gunderson and Clarke, 1998a,b; The Nations Report Card, 2020). For children and adults of color, particularly, this mismatch is further exacerbated as their racialized linguistic and literate practices often remain unaddressed by the limited opportunities presented in Eurocentric and Westernized approaches to literacy curriculum and standards in schools (Milner, 2012; Rosa and Flores, 2017; Smith, 2019; Willis, 2015, 2018).

It is for this reason that family literacy programs and research are best understood within a broader social, cultural, racial, and economic policy context. Scholars have linked the growth in family literacy programs with a concern over what is termed the literacy crisis, high dropout rates, and low academic achievement (Auerbach, 1989), challenges historically associated with families of color, emphases that continue to be mirrored in UNESCO's approach to family literacy (UNESCO Institute for Lifelong Learning, 2017). And national studies and policies in North America and the United Kingdom such as the International Adult Literacy Survey (OECD, 1996) and the United States' Equipped for the Future Initiative 2000 (UNESCO Institute for Lifelong Learning, 2017) increasingly look beyond schools to families as the source of both the problem of and the solution to the low school achievement of minority and low-income children, a large majority of whom tend to be students of color.

The understanding of family in such policies, historically and contemporarily, has nonetheless been limited primarily to mothers, primary caregivers, and their young children (UNESCO, 2008). This phenomenon prompted Jane Mace to comment:

The evidence of the literacy problem in industrialized countries with mass schooling systems has revealed that schools cannot alone meet this need. Families must therefore be recruited to do their bit, too. This is where the spotlight falls on the mother. She it is who must ensure that the young child arrives at school ready for school literacy, and preferably already literate (Mace, 1998, p. 5).

Early childhood focus

Another area of debate in the family literacy field is the wisdom of designing family literacy programs mainly for children from birth to 5 years old and their primary caregiver—usually the mother or grandmother. This has raised concerns about ignoring the literacy and learning needs of alternative families, as well as of older children, youth, and adults, a large number of which tend to be part of families of color. Luke and Luke (2001) make the case that the needs of adolescents experiencing literacy difficulties and their families are largely ignored, and resources are directed almost exclusively to young children.

Focus on storybook reading

A third and largely debated focus is storybook reading. Although ethnographic research with families reveals myriad forms and functions of reading and writing in daily life, many family literacy programs focus on reading story books to children as the way (Pellegri, 1991) to support children's literacy development. This emphasis on storybook reading apparently emanates from early research with precocious readers (Clark, 1976). This research, conducted mainly with white, middle-class families, indicates that reading to children was a common factor. Several criticisms, however, have been raised about the emphasis on storybook reading. First, clear evidence shows that young children's literacy development is supported in many different ways in addition to storybook reading (Hadley et al., 2020; Taylor, 1983). Second, storybook reading is not common across all cultures and social classes; by emphasizing it, the literacy practices of minority and other groups are devalued (Meier, 2003), much like is done for students of

color. Third, not all children enjoy being read to, and the implication that it is necessary in order for children to learn to read causes some parents to insist on their children's participation. Scarborough and Dobrich (1994) cautioned that some children consequently develop negative attitudes toward reading and there are those who have observed that despite the importance placed on storybook reading, the empirical evidence suggests that it plays a less significant role in learning to read than is commonly believed (Saracho, 2017; Scarborough and Dobrich, 1994).

Absence of longitudinal focus

Another concern related to family literacy policies is the lack of longitudinal studies documenting their impact over time. Adele Thomas describes much of the research in family literacy as "testimonials" (Thomas and Skage, 1998, p. 20), suggesting that little attention has been paid to broad measures of program effectiveness, raising questions about the degree to which these largely implemented mechanisms result in an increase in literacy gains for families of color as well as decisions surrounding how to evaluate measures used to gauge effectiveness.

In each of these debates, the needs of families of color appear to be affected, further emphasizing a re-envisioning of the construct and emphasis of family literacy and inviting a response from scholars focused on such needs.

Responses to family literacy debates

In response to the foregoing debates about family literacy, Wiley (1996) has identified three key elements useful for research communities working with families: *accommodation*, *incorporation*, and *adaptation*, all of which appear to bode well for advancing family literacy for families of color. According to Wiley (1996), *accommodation* places the responsibility for having an enhanced understanding of the ways in which students engage in communicative styles and literacy practices on educational stakeholders. *Incorporation* invites researchers to make understudied and undervalued community practices a priority, to study these practices and to integrate them into the curriculum by acknowledging the ways in which various ethnic groups can decenter mainstream curriculum practices often prioritized in schools. *Adaptation* proposes that children and their families, who are often perceived as having knowledge and skills that are often undervalued, can be expected to master the norms of schools.

Notwithstanding, adaptation approaches, which tend to be popular in family literacy implementation, have faced criticism (e.g., Yaden and Paratore, 2003). For instance, there are those who claim that these programs promote "blaming the victim" (Cairney, 1997; Nieto, 1992; Shockley, 1994; Street, 1995), perpetuating a dynamic where a deficit-oriented approach is taken that tends to mirror a "neoliberal Savior" (Brown, 2013) mentality. These arguments challenge the basis for family literacy programs which imply that the homes of poor, minority, racialized and immigrant children are "lacking literacy," an assumption increasingly identified as the result of White supremacist notions of what it means to use language based on Eurocentric norms in literate practice (Anderson and Stokes, 1984; Auerbach, 1989; Chall and Snow, 1982; Delgado-Gaitan, 1987; Erickson, 1989; Goldenberg, 1984; Rosa and Flores, 2017; Willis, 2018). Others argue that such programs do not recognize that "literacy is not something which can be pasted on to family life [but as] deeply embedded within it" (Macleod, 1996, p. 130) and perpetuate "the 'we know, you don't know' dichotomy" (Shockley, 1994, p. 500).

To address such criticisms of the adaptation approach, Edwards (1995) has invited scholars to pose questions such as the following:

- Where have researchers' fears, doubts, reservations come from?
- What do researchers think these parents are being "forced" to read?
- Does evidence exist where researchers have interviewed parents who have attended these family literacy programs?
- Is it fair for researchers to insert their own personal feelings about parents participating in family literacy programs without highlighting parent voices, perceptions, and evaluations?
- Are researchers' fears, doubts, and reservations justified? (p. 562)

Also, in response to these debates, scholars such as Li (2012) have invited a consideration of the pedagogy of cultural reciprocity in parent-teacher interactions across home and community where the teacher engages in personal inquiry and change while the parent also explores personal change and empowerment. This perspective, Li believes, can be fostered through school structural change, classroom pedagogical change, change in the home school relationship and change in perspectives on underserved student populations, creating opportunities in family literacy practices for families, including families of color.

Overall, approaches to family literacy implementation in the United States have reflected a wide range of programming efforts as well as a theoretical, conceptual, philosophical, and practical underpinnings that emphasize the intervention/prevention as well as adaptation approach to family literacy, creating a situation where the literacies of families of color are inadvertently sidelined. Despite numerous initiatives such as book giveaway programs that emphasize traditional literacy skills and remain prevalent and touted as impactful to family literacy (DeBondt et al., 2020), the enduring debates, many of which have centered primarily around cultural mismatches between the school-valued literacy practices and cultural literate norms of students in the home reinforce the need for a re-envisioning of family literacy education and programs for families of color in the United States.

Approaches to family literacy across the globe

While the historical basis for family literacy originated with the United States, Canada, and South Africa, increasing initiatives across the globe in both the Minority and Majority World focus on family literacy which reflect large populations of families of color (see Panelli et al., 2007 for a discussion of the labels, “Majority” and “Minority” World). Globally, research on family literacy has tended to be driven by Western notions. Yet, families of color around the world have always engaged in intergenerational literacy learning and practices. UNESCO has been a driving force in identifying what countries are doing globally with regards to family literacy practice from both a research and practitioner perspective. A report conducted by the UNESCO Institute for Lifelong Learning (UIL) reflects findings from the 10-year Literacy Initiative for Empowerment (LIFE) (2006–2015) focused on 35 countries worldwide where the literacy rate was lower than fifty percent and/or the adult population reflected over 10 million people lacking literacy competencies.

Among the findings reflected in this report are the ways in which family literacy, when considered at the international level, mirrors to some degree, an emphasis on traditional literacy support structures for the child, traditional literacy support structures for the parent, emphasis on the literacy connections between parents and children, and a focus on developing intergenerational literacy capacity and centers lifelong learning (UNESCO, 2008, p. 6). These findings are compelling because they reinforce the need for a focus on the cultural differences that come to bear on constrained literacy skills (Paris, 2005) if the multiple and critical literacies (Luke, 2012; Street, 1995) of students and families of color are to function as a basis for empowerment. The findings highlight the need for centering such literacies within contexts where, to leverage literacy practices is to consider the cultural norms of students, their parents, and the overall families as they operate within various contexts across the globe.

In the following overview of literacy programs captured internationally across different countries and contexts, many of whom reflect families of color (Desmond and Elfert, 2008; UNESCO, 2008), we illustrate how a number of the afore-mentioned findings are reflected (see Table 1). Through this illustration, it is clear that opportunities exist where cultural norms and the multiple and critical literacies that they invoke, as well as policy-orientated approaches to family literacy within nations, might become more central as an area of need when addressing family literacy, given the traditional notions that persist in many of the contexts explored.

Insights from international family literacy programs

Across the programs presented in the afore-mentioned countries (Table 1), significant variation exists in terms of the time allocated for programs. Just as was observed earlier in the US context, there are a number of situations in which certain countries reflect governmentally operated programs while others reflect programs operated under the auspices of National Government Organizations (NGO’s) because of limited visibility in policies designed to guide family literacy practices. In certain instances, communities appear to be included in family literacy programs while in many others, communities as a whole were not. There seems to be significant variation in how emphasis is placed on literacy across the Global North with predominantly White populations and Global South with predominantly families of color; the North features a heavier emphasis on literacy than the South. There are indications that the distance between schools and communities affects the ability of parents and children to navigate transitions and there seems to be a needed distinction between adult literacy programs and family literacy programs where the focus is placed on the individual vs. the parent-child relationship, given that all of programs, regardless of an individual focus, often fall under the same umbrella term, *family literacy* (UNESCO, 2008).

Among the programs presented in Table 1, it was further observed that most programs appear to target children or their parents. Very few programs seem to reflect a focus on parents and children in tandem or an equal focus on the parent and child, a shift that seemed to be partly reflected in the United States. There were a number of instances where mothers are explicitly targeted by programs, in alignment with earlier observations by Mace (1998) but no instance where fathers are an explicit focus. Much like was observed earlier in the US context, most programs focus on early childhood or elementary contexts while fewer programs focused on youth (UNESCO, 2008). Only one instance was identified where bilingualism was acknowledged and addressed (UNESCO, 2008) and very few programs highlighted the notion of critical literacy or multiple literacies as a vehicle for drawing on the already-present literacy (Marciano and Watson, 2021) assets of families of color.

Among the challenges reflected across international programs is low involvement of fathers in the literacy structures and supports provided, observation that the investment of grandparents appeared to be crucial to potential success, realization that adolescents tended to have difficulty with and thus, failed to secure tremendous benefits from the programs, and the recognition that multilingualism requires teaching language before family literacy content can be taught (UNESCO, 2008).

Based on these indications, key criteria for family literacy programs should involve the provision of training for teachers carrying out the family literacy programs with parent involvement. In addition, evaluation must be built into programs from their inception with a focus on qualitative and/or quantitative research as well as a shift to comparative studies across programs where findings are presented in multimodal and attractive formats for all stakeholders involved (UNESCO, 2008). This recommendation for evaluation of family literacy programs mirrors the call by Doyle (2012) for “thoughtful approaches to evaluation that consider the relations between program design features and program effects ... to provide the evidence necessary for maintaining funding support, [and] also to ensure that program participation yields optimal results for participating families” (p. 98). The call remains critical given that intervention/prevention and adaptation approaches also appear to be visible in family literacy programs that cater primarily to families of color across the globe, as they also were for families in the United States.

Table 1 Family literacy programs across the globe.

Africa				
Mali	Namibia	Senegal	South Africa	Uganda
• Support for the Quality and Equity of Education in Mali (Appui à l'amélioration de la Qualité et de l'Égalité de l'Éducation/AQEE) program (2003–2008), training and support to 700 parent associations and school management community-based organizations • The adult literacy program, Sanmogoya, functioning as part of this initiative was active in six regions of Mali (Timbuktu, Gao, Kidal, Koulikoro, Sikasso). • There, approximately 6000 adult learners benefitted from the services provided by 140 literacy centers. • The course was divided into two phases—basic literacy and post-literacy based on sound theories of reading and writing—and is taught by volunteer teachers drawn from the community. • The Parent Association in each community managed the program with training and support from the local NGOs.	• A study was conducted in a school in each of the thirteen regions to launch the family literacy program to establish what school principals, teachers, children and their parents (caregivers) believed to be the ways in which parents support their children in the 1st year of primary school • Next step in the implementation of the family literacy program was the training of family literacy promoters or facilitators. • Training of these promoters or facilitators. • Training of these promoters based on Ministry of Education document, “The First Teacher” and based on the skills and competencies expected from learners in the 1st year of school. • The main purpose of the training was to equip the promoters with an understanding of the importance of the role of the parents and caregivers in the development of early literacy skills in their children. This included early childhood development, storytelling and reading, songs and music, child protection, children's rights, HIV/AIDS and nutrition. In addition, sessions were designed to develop skills in facilitating groups and designing and writing session outlines. • Also, in Namibia, was the Adult Education program. The Ministry of Education recognized the high drop-out rates in Namibia and responded to this challenge by focusing on the problem of literacy in the home environment. In 2005, a carefully selected Family Literacy Promoters and Education Officers from the Head Office received a 10-week training program which allowed them to target 1281 parents with children in pre-school and in grade one	• The Grandmother Project (GMP) produced a booklet, entitled “The Role of Grandmothers in Halpular Society” to highlight the vital role that grandmothers play in the lives of children and their families. It was developed for use in both primary schools and adult literacy classes in southern Senegal. • The booklet and its dissemination to primary schools and adult literacy classes has served as a tool for teachers to bridge the gap between school and home by promoting discussion of cultural roles, values, and traditions. • Tostan, a non-government organization based in Senegal, realized that a successful program for helping women to transform their lives meant rethinking traditional educational approaches to better respond to participants' needs, habit, lifestyle, and culture • Tostan also felt it was essential to create program content and strategies to improve communication, understanding, and collaborative action for positive change.	• The Family Literacy Project (FLP), based in the KwaZulu-Natal province of South Africa, was started in 2000 to address the low literacy achievement of many pre- and primary school children. • The program also focused attention on parents who demonstrated low confidence in their ability to provide support for these children. In the program, eleven facilitators worked directly with 197 adults, 479 primary school children and 116 teenagers. • FLP group members visited neighbors to share information on early childhood health and development. • They also played with the children and read to them. • In spite of this progress, relevant reading materials were scarce and led to low levels of literacy for both children and adults. • Three community libraries, ran by the project, helped to address these concerns.	• Uganda's “Literacy and Adult Basic Education (LABE)” was developed when the country first reflected interest in family literacy. • The piloted program (2000–2001) supported parents and teachers in the Buguri district—one of the poorest districts in Uganda with low adult literacy rates. • LABE expanded the coverage of the pilot project from 3 to 18 schools. Eventually, the project catered to 1400 parents and over 3300 children. • And nationally, the African Family Literacy Action Group (AFLAG), founded after the November 2007 UNESCO Institute for Lifelong Learning (UIL), is a group which aims to implement a Center for Family Literacy in Africa which can function as a point of reference for promoting family literacy and disseminating tools and know-how across the continent.

Arab initiatives

Palestine

- The “Mother-to-Mother Program” in Palestinian communities has supported and strengthened the parenting skills of young mothers and trained mothers as paraprofessionals.
- The Learn-by-Play Program attempted to reverse the trend of increasing drop-out rates in elementary schools by building trusting relationships between pupils and teacher and creating a stimulating educational environment.

Bahrain and Saudi Arabia

- The “Mother-Child Home Education Program (MOCEP)” has provided a non-formal, home-based education program for pre-school-aged children and for their mothers.
- The annual 25-week program supported women who were provided with a schedule of daily lessons which they could use to teach their 5-year-old children at home.
- Like many of the family literacy program contexts targeted in the United States, participants (children entering primary school) were from low income, disadvantaged families, unable to afford pre-school education for their children.

Asia & Pacific

Vanuatu

- Vanuatu’s “Vanuatu Literacy Education Project (VANLEP)” has been coordinated in primary schools as an initiative focused on improving the literacy learning outcomes in rural and isolated areas of the country.
- VANLEP targets 8 and 9 year-old students in classes 3 and 4.
- Teacher training and teacher empowerment represented an important element of the program.
- Other objectives were to provide quality books for children, improve classroom management, and parent’s awareness with regard to literacy learning at home and participation in literacy acquisition at school.

National programming

- The “Manukau Family Literacy Project” which was based originally on the American Kennan model, eventually reflected the added component of Parent and Child Together Time.
- The model involved adults attending their children’s schools to participate in a program.
- Adults also enrolled in a course focused on child development and participated in the reading and numeracy components of their children’s schoolwork.

Europe

Hamburg

Malta

Romania

France

Greece

- The FLY pilot project was implemented in 2004 with participation of parents of 5-year-old pre-school children and runs in nine locations in socially disadvantage districts of Hamburg.
- This program aimed to involve parents in everyday school life with a focus on literacy activities.
- Individuals worked with parents within the pre-school/day centers and their support focused on three major components-keeping parents actively involved in everyday classroom activities (together with children), engaging in activities with parents independently of children, and engaging in parent and student activities outside of school.
- Beyond this, the “Foundation for Educational Services (FES)” in Malta was conceived as a means of providing a range of innovative educational initiatives a range of innovative educational initiatives with a focus on literacy support, parental empowerment and lifelong learning. The program involved after-school family-oriented educational service, complementing and reinforcing teaching and learning at day schools.
- In addition, Romania’s “Center Education 2000+ (CEDU)” focused on “Parent Empowerment through Family Literacy (PEFaL),” a program that remains visible in a few schools and kindergartens.
- PEFaL functions as the only literacy program targeting parents and children equally in the country
- And “Mother Child Education Foundation (A.EV)” in Turkey functioned as a source of empowerment for people using education through the implementation of research as well as the development and implementation of programs in early childhood education and adult education, particularly targeting pre-school children and their families who were at a disadvantage.
- “Agence nationale de lutte contre l’illettrisme (The National Agency to Fight Illiteracy/ANLCI)” in France (2000) was launched to coordinate the efforts of all stakeholders and the resources provided by the State, local authorities and private enterprises towards advocacy and action in the field of literacy.
- ANLCI recently began a focus on family literacy and has actively promoted the approach in France.
- “Parents’ Schools” program (2003) was launched nationwide in the context of the country’s commitment to lifelong learning exist in all 14 prefectures with 45,000 (approximately 3500 fathers and 41,500 mothers) participating in these courses
- The courses offered through this program are mainly held in schools and community centers. The Parents’ Schools were run exclusively for parents, although parents were recruited through their children’s educational institutions.
- Like many programs in the United States, the focus of the Parents’ Schools is on parenting intergenerational communication, encouraging parents to participate in school activities and addressing the issues of violence and health.

National programming

- “Basic Skills Agency (1990+)” also a European program, was leveraged and reflected three strands: sessions for parents in which they worked on their own literacy and on how to help their children; parallel sessions for children in which they received high-quality early years provision for reading, writing and talking; and joint sessions in which the parents typically tried out something with their children that they had been practicing just beforehand in their separate session.
- Book start originated in the United Kingdom in 1992 and distributes Bookstar Baby packs which often have baby books and a flyer with information about engaging in shared book reading with infants. Over 3 million book packs have been distributed across the European Union, Asia, Australia, and New Zealand through the program (Book Trust, 2019).
- After the November 2007 UNESCO Institute for Lifelong Learning (UIL), France submitted a proposal for a national family literacy pilot program to the Ministry for Social Affairs. This has since been approved and was in the process of being launched.

North & Latin America

<i>Guatemala</i>	<i>Caribbean</i>	<i>Canada</i>	<i>USA</i>
• Guatemala launched the “National Committee on Literacy (CONALFA)” in 1998, a program featuring the idea that one member of the family, generally a child attending the 4th, 5th or 6th grade of primary school would provide teaching to adult members of his or her family unit, in most cases, this being the mother. • The primary objectives of the program were to assist the educational development of the individual within the home, without detracting from the culture or behavioral patterns that often held within the family environment, and to strengthen family relations by allocating responsibilities to all its members.	• Caribbean Community (CARICOM) spearheaded the Caribbean Action Plan for Early Childhood Care, Education and Development, approved by the Standing Committee of Ministers of Education in 1997. • Five years later, Caribbean leaders endorsed a regional action plan for the period 2002–2015 with Guidelines for Developing Policy, Regulation and Standards in Early Childhood Development Services. • OECS emphasized education and training for parents and caregivers of children of children (0–5) and appropriate education for the children (3–5). • The Foundation for the Development of Caribbean Children (FDCC) implemented the Roving Caregivers Program (RCP) to deliver family and community interventions, help create and apply knowledge, support advocacy and communication for development. • FDCC runs the Early Childhood Health Outreach (ECHO) to incorporate maternal and child health services into RCP (0–5). • FDCC supports the Family Learning Program, providing literacy and numeracy skills to parents. • FDCC’s Youth and Community Advocacy Network (YouCAN) mobilizes young people to get involved in the quest for high-quality ECD and family support programs.	• “Éduk”—an NGO working mainly in the area of literacy for the francophone minority in Canada developed and implemented literacy and family literacy programs and policies in response for the need for communities, families and individuals to be exposed to and use the minority language in order to gain mastery of both the minority and majority languages, practitioners are provided with a comprehensive five-day program on the foundations of family literacy in French and English. • The French-language family literacy programs help parents transmit mother tongue skills and develop a sense of belonging to a specific culture.	• The “National Center for Family Literacy (NCFL)” in the USA reports 5000 family literacy programs operating in non-governmental organizations (NGOs), schools, libraries, and early childhood centers. • NCFL provided a model demonstration of best practices. Provided training for more than 150,000 teachers and program administrators, advocates for family literacy in the US Congress and state governments, and conducted research and evaluation to increase the knowledge base and improve the quality of programs as they continue to emerge.

Global programs & conferences

- In 2004, UIL launched family literacy pilot project (FLY) in Hamburg inspired by the family literacy programs developed by the Basic Skills Agency (BSA) in the UK and from 2005 to 2007, UIL coordinated a European project on family literacy—QualiFLY (Quality in Family Literacy)—which was funded by the European Union through the Socrates/Grundtvig program and included partners from around the world.
- Family literacy has remained a focus of UNESCO’s series of Regional Conferences in Support of Global Literacy in the context of Education for All (EFA) and the United Nations.

Culturally relevant literacies of families of color

As has been shown, family literacy programs both in the US and across the globe, were from the inception, developed with the child in mind, overlooking the parent and privileging school-based literacies. Even with certain advancements in and beyond the US that cater to adults, the centering of the literacies of families of color and the equal focus on culturally relevant literacies geared toward both parent and child remain a laudable goal. To address this dynamic, family literacy programs must get to know families of color and understand the literacies that they already possess. Research findings presented below, though not exhaustive, have informed our understanding of how discourse and dialog within families of color influence children's language and literacy learning as well as how distinctive family patterns and literacies can be, even among families living in similar circumstances or locations, as well as those migrating across contexts (e.g., Compton-Lilly et al., 2012; Desmond and Elfert, 2008; Edwards and White, 2018; Heath, 1983; Li, 2000; Moll et al., 1992a,b; Smith and Marfo, 2015; Smith, 2019; Taylor, 1983; Velliaris, 2015).

Ethnographic studies

As alluded to earlier, in her detailed ethnological study of families in North Carolina, Shirley Brice Heath (1983) examined the literacy practices of three communities in the Carolina Piedmonts and found the ways that White, middle-class families used literacy at home were most strongly associated with the forms of literacy taught in school. Consequently, these children tended to succeed in school. However, the literacy practices of African American middle-class and working-class communities at home differed from the forms of literacy valued and taught in school. These children were less successful in school. Heath intended to demonstrate the need for more diverse and inclusive literacy teaching strategies in schools. Her observations about the use of oral language and print among all three groups of children and their communities provided one of the most nuanced documentations of cultural differences related to language and literacy and how these differences influenced children's performance in school.

Similarly, Luis Moll has coined the term "funds of knowledge" (FoK) to refer to practices in the household including those involving labor and socialization that are located at the center of learning interactions between children and adults, accessible through the ethnographic home visits of teachers and researchers (Moll et al., 1992a,b). Extending beyond the households of families of color to their life worlds, researchers have identified FoK across a range of contexts from family values and beliefs to popular culture, interests, content areas, communities and classrooms (Bouillion and Gomez, 2001; Moje et al., 2004; Razfar et al., 2015; Riojas-Cortez and Flores, 2009; Smythe and Toohey, 2009). For families of color, examining the FoK consisting of knowledge and skills accumulated culturally and historically allows teachers to incorporate literacies from their life-worlds to support their well-being, provide a context for more traditional literacy learning and also, interrupt institutional norms that privilege certain literacies over others (Flores, 2020; Moll et al., 1992a,b, 2005; Razfar et al., 2015; Rosa and Flores, 2017). Mirroring the focus of family literacy from its inception, children of color "are not passive bystanders, as they seem in the classrooms, but active participants" (Moll, 2005, p. 74).

Li (2000) has illustrated how families of color each live "their own literate lives in their own way" (p. 27). In her ethnographic examination of an immigrant Filipino family in Canada, Li (2000) illustrated how individuals function both within and out of ethnic groups to develop identities that may or may not mirror the typical depiction of this group. Through the ideologies and power relationships that individuals in the family aligned with based on their socialization in the family as well as the school and broader society, Li has shown how discourses work in families of color to align with corresponding social and how family literacy practices can therefore be very complex and different from what is often typically expected by society.

And Smith (2019) has highlighted how a Black adolescent immigrant youth, Jaeda, navigated literate and racialized linguistic discourse with her mother and sister in a single-parent family while also engaging the racialized discourses of peers and teachers in schools. Examining the literate uses of Englishes and its corresponding raciolinguistic ideologies (Rosa and Flores, 2017) leveraged by this adolescent in her linguistic and literate interactions with her mother over time even as she transitioned across Jamaica and the United States, Smith (2019) illustrated how Jaeda, in this single-parent family, made use of "metalinguistic, metaracial, and metacultural understanding of her past experiences with race and language, and by extension, culture, to determine how to function effectively within non/academic settings in ways that [did] not completely sacrifice her personhood" (pp. 299–300).

Through multiracial understanding, Jaeda focused on her thoughts about race, including how the racialization and social constitution of Englishes and other languages, allowed her to exercise duality while retaining her personhood. Through metalinguistic understanding, she focused on thinking regarding her thoughts about language, developing awareness and control of the linguistic elements that she used, including understanding how and why Englishes and other languages are flexibly used for different purposes. Through metacultural understanding, she acknowledged how her race and racialized Englishes as a Caribbean youth were both associated with, yet transcended typical conceptions of these constructs in the cultural context of the US. Presenting how these intersecting forms of understanding functioned in Jaeda's literate discourse, Smith (2019) demonstrated how the adolescent adopted "a transraciolinguistic approach," using her past experiences with race and language, and by extension, culture, to agentively exercise flexibility and duality in response to racialization.

Programmatic tensions

As discourse, family literacy is inclusive of the oral and written communications within the family and it encompasses the family's efforts to support the language and literacy development of their children. Indeed, it is this definition that often puts some

professionals at odds with others, as some wish to keep a focus on interactions that are unique to the family, while others see the need for skill development essential for success in educational and workplace settings. This bind is noted by Faircloth and Thompson (2012) as they discuss the tension between the traditional literacy practices in Native American families in contrast to literacy skills needed for education and economic success. Similarly, Scott et al. (2012) stress the need to be inclusive of family beliefs and values when considering services to families. And Smith and Marfo (2015) highlight the need for closer attention to linguistic diversity of children in early childhood care and education settings, critical to emergent learners' academic progress, socioemotional development (Bauer, 2000; Bauer and Arazi, 2011; Siegel, 2012) and social, emotional, academic, and cultural skills (Dickinson et al., 2006; Strickland and Morrow, 1989; Tomasello, 1999). Using oral traditions and cultural heritage of families can provide a creative and empowering means of helping families develop reading and writing skills important for many everyday functions, and can potentially serve as a foundation for more advanced literacy development (e.g., Adra, 2004).

On a broader level, the literacy skills of families of color are central to the literacy proficiency of neighborhoods and communities; viewed collectively, they determine the literacy levels of individual countries around the world. That the family is instrumental in these highly significant international phenomena is no surprise, given the extensive documentation of the correspondence between parental literacy levels and children's literacy skills. Other social organizations, especially schools, certainly carry a significant responsibility for producing a literate society, but they can accomplish their mission better when families are able to provide children with a firm foundation in language and early literacy prior to school entrance. Because some families may lack the skills or knowledge to help their children succeed in school and later employment, many societies have provided services to help ensure that not only children but also the adults have opportunities to gain these skills (UNESCO, 2008, 2019). The UNESCO Institute for Lifelong Learning (UIL), partnering with national and local partners in countries within Africa and South Asia, with predominantly populations of color, reflects how societies such as those in Benin, Burkina Faso, Ethiopia, the Gambia, Guinea-Bissau, Guinée, Mali, Niger, Senegal, and Pakistan, in 2019, has used workshops to build capacity for communities of color in family literacy and learning (UNESCO, 2019).

Family "literacies" vs. school-valued literacy

As observed in our discussion of family literacy both in the US and across the globe, there continues to be an emphasis on enhancing the livelihoods of parents and families, but it is unclear the degree to which culturally relevant literacies that draw upon the already-present literacies, cultural practices and community wealth, play a key role in designing and transforming literacy curriculum to align with the needs identified by large numbers of culturally, racially, and linguistically diverse families across the globe. Many teachers judge students' literacy practices according to their own background experiences and mainstream expectations of how print can be used. They view their own mainstream literacy practices as "right," and "legitimate" and not in the context of power relationships within a socially and racially stratified society. And they identify literate practice as the kind that is valued within schools. Most teachers typically consider activities such as mastering the alphabet, reading books, or writing essays as literate practices. However, literate practices are much more complicated, extensive, and deeply embedded in culture. For instance, there are literate practices associated with Hip Hop culture that are not legitimized in school, but nonetheless involve sophisticated use of language (Morrell and Duncan-Andrade, 2002). Considering multiple literacies as an avenue for reconsidering the role of school-valued literacies may present a pathway for using mechanisms such as parent stories to address this concern (Edwards et al., 1999, 2019; Ellison and Solomon, 2019).

Literacy instruction is more effective when it honors the ways of using language and print presented by students of color, insights made possible through the stories of parents. A social equity orientation supports the view that *schools and teachers* need to recognize the literacies that children bring from home so that we can build upon what they know and do in the classroom (Lazar et al., 2012). Beyond this, we believe the insights from our discussion point to the need for re-envisioning the conception of family literacy, which we subsequently propose.

Parent stories

Parent stories (Edwards et al., 1999) are windows into the lives of children and their families and be seen as repositories of information about literacies within homes and communities as well as the "funds of knowledge" (Gonzalez et al., 2005) that are contained in households. Collecting parent stories provide a way to gather information about students' home and community valued literacy practices (Edwards et al., 1999). Teachers can generate parent stories through thoughtful questioning and dialog with caregivers. Several teachers have shared with Edwards that they viewed parent stories as a tool to gain insight into students' current academic performance as well as their potential. Teachers also shared that parent stories helped them to better understand their students' literate lives outside of the school context. Some of the questions that teachers have asked from the Edwards-Pleasant Parent Stories Questionnaire (1999) include:

- Can you describe "something" about your home learning environment that you feel might be different from the learning environment of the school? (p. 40)
- Can you describe "something" about your home learning environment that you would like the school to build upon because you feel that this "something" would enhance your child's learning potential at school? (p. 40)

- Is there “something” about your child that might not be obvious to the teacher, but might positively or negatively affect his/her performance in school if the teacher knew? If so, what would that “something” be? (p. 38)
- Everyone has hopes and dreams when they are young children. As a child, what did you want to be when you grew up? Did that change over the years? Have you realized your childhood goals for your future? (p. 39)

A collection of parent stories can serve as a space for shaping and impacting teachers’ knowledge and beliefs about how to connect students’ language and literacy experiences valued in their homes and communities with formal school valued literacy practices. This connection might interrupt deficit ways of thinking and put us on the path toward more inclusive models for working with nonmainstream children and families of color.

Re-envisioning family literacy for families of color

Given the foregoing discussion, we envision family literacy for families of color in the 21st century as multifaceted. Family literacy has come to embody the multiple and critical literacies that parents and children use to intergenerationally interact and sustain their cultural, linguistic, and racial identities across home, school, and the globe while mastering and manipulating school valued literacies for agency to determine how they can best thrive in society and create social change. In such a definition, family literacy is concerned with “culturally relevant” and “culturally sustaining” mechanisms that address the “multiplicities of identities and cultures that help formulate today’s youth culture” (Ladson-Billings, 2014, p. 82; see also Paris, 2012) as well as drawing on the “already-present literacies” (Marciano and Watson, 2021) in the “within-group” and “across group” cultural practices necessary for cultural competence (Paris, 2012, p. 95) and a sense of agency. At the same time, family literacy is concerned with supporting parents and their children, both independently and collaboratively, to master and manipulate school-valued literacies in ways that align with their culturally, linguistically, and racially diverse identities while supporting their sense of agency for thriving within and across culturally diverse communities for socially transforming the world. In contrast to a focus on building traditional literacy skills for problem-solving, a contemporary definition of family literacy for families of color moves toward identifying the literate assets that children and parents agentively possess based on how they already interact culturally within their communities and uses this as a basis for determining how school valued literacies might be taught and leveraged in ways that extend their communicative literate repertoires.

We believe such an approach necessitates a shift from family literacy to what we envision as “*family literacies*,” a notion that “disrupt[s] practical [and static] notions of what it means for ... [families] to reflect success in literacy” by acknowledging how literacy practices based on school-valued literacies often reflect “schizophrenic institutional norms that delegitimize [the] personhood, linguistic repertoires, and literacies” of underserved populations and “[create] potential challenges for their wellbeing” even while claiming to equip them with literacy skills (Smith et al., 2020). With this contemporary definition, we invite the transition from the notion of family literacy that emphasizes code-, employment-based and life-skills of families for aligning with pre-existing needs of society to a vision of “*family literacies*” that instead serves as a vehicle for empowerment and agency based on the cultural, racial, and linguistic wealth of communities as well as the desires of families of color for social change.

Considerations for family literacy programs for families of color

Based on the above, it is clear that transitioning toward *family literacies* for families of color will require a rethinking of the current ways in which approaches are leveraged, particularly in early childhood contexts. There is also a need for foregrounding children’s literature and its many facets as a basis for developing and evaluating quality approaches to family literacies for racially, linguistically, and culturally diverse communities.

Reconceptualizing family literacies in early childhood development

Drawing from Smith and Marfo (2015), we recommend, much like “parent stories” (Edwards et al., 1999), envisioning *family literacies* as part of contemporary holistic conceptualizations of early childhood development that move away from traditional notions by finding out from caregivers what is needed by those in the early childhood sector here and now. Beyond this, we envision *family literacies* as an indispensable element of all aspects of “early childhood development,” where this development is holistic in nature and that allows for a centralization of administrators who interact with each other around how to consolidate delivery services in regard to linguistic, health, racial, social, emotional, cultural, etc. As observed by Smith and Marfo (2015), “the construct of early childhood development (ECD) provides a much broader umbrella framework for bringing together all dimensions of children’s health, language, development, protection, and well-being into a system of policy-driven intervention services with preventive, facilitative, and ameliorative objectives” (p. 41).

Such a system does not have to consolidate all sector-based early childhood services under one bureaucratic organizational structure or management; rather, it will serve to provide an overarching integrative synergy that brings much-needed national coherence to policy formulation and execution and in the process, strengthens sector-based programs. In doing so, literacy, or rather “*literacies*,” would form a basis for this integrative synergy. There are literacies undergirding language, literacies undergirding health, literacies undergirding development, literacies undergirding well-being. In this proposed approach, literacies within and across each

of these sectors would be identified based on examinations of the at-risk populations in question with particular attention to race, language, culture, gender, class, nationality, etc.

The literacy curricular then, contrary to what it tends to be at the moment where traditional approaches to family literacy limit educational foci to constrained skills such as concepts of print, the alphabet, phonics, phonemic awareness, etc. (Paris, 2005) would expand significantly to include literacies that are useful for daily living, thriving, agency and survival. By the same token, family literacy programs would extend beyond traditional literacy skills and be re-examined and re-envisioned to reflect elements such as critical literacy (Bacon, 2017; Luke, 2012), anti-racism (e.g., Smith, In review; Willis & Smith, In review), and a transraciolinguistic approach (Smith, 2019). For instance, integrating mechanisms such as a transraciolinguistic approach into family literacies can facilitate the use of metalinguistic, metacultural and metaracial understanding in family literacy programs, allowing teachers as well as parents to rethink how skills such as phonemic awareness can penalize students' dialectal production because their language is racialized in literacy classrooms (Smith, 2019, 2022; Smith et al., 2020).

An early childhood development system that centers *family literacies* would replace the fragmented system often seen with family literacy practices that tend to operate often independently of "early childhood care and education" by placing "education" in one locale of the government or society and other elements such as "health" in another. As Smith and Marfo (2015) have explained:

While we acknowledge, though, that coordinating disparate services through a liaison agency is certainly one way to conceptualize integration, what we have in mind here are realistic, flexible, and dynamic cross-sector planning and coordination strategies that make it possible to structure and manage new and existing programs/services—around space, time, and distance—to ensure that for any child (and family), any number needed intervention components will be readily accessible in closeness, proximity to the community of residence (p. 41).

A re-envisioned and re-imagined notion of *family literacies* for families of color would thus emerge from, as well as cater to, the unique literacies of at-risk populations, which in turn, can function as a basis for caregivers, educators, healthcare providers, working in tandem to develop a holistic child. Equity can only be achieved when the assets that at-risk populations bring are used as a basis for building on their already established and prioritized knowledge bases about the world and survival in ways that they can find usefulness. Such an approach would foster redesigned literacy curriculum for parents and children that extends beyond the traditional sense and enables them to use their literacies to navigate multiple layers of success in an increasingly complex society. It would also challenge us to redesign traditional literacy curriculum and standards in ways that center anti-racist policies and practices during instruction. *The question then becomes, "What literacies can be incorporated into and across family literacy programs in ways that build upon the already-present literacies that families of color possess and the agency that they wish to use for social change in and beyond their communities?"*

Leveraging equitable literacy practices in family literacy for families of color

Incorporating children's literature that reflects the cultural norms, histories, and experiences of families of color is one way in which to build literacies into family literacy programs and tap into the already-present literacies they possess. Books written by and about people of color can provide opportunities for children of color to access their background knowledge which aids the comprehension of text. For example, Bell and Clark (1998) conducted an experimental study with African American children in grades 1–4 and found that "comprehension was significantly more efficient for stories depicting both black imagery and culturally related themes than for stories depicting both white imagery and culturally distant themes" (p. 470). A focus on literature written by and about people of color also provides opportunities for children of color and their families to see themselves represented in ways that are authentic—thereby allowing books to function not solely as windows, but also as mirrors for them (Bishop, 1990). Bishop (1990) writes:

Books are sometimes windows, offering views of worlds that may be real or imagined, familiar or strange. These windows are also sliding glass doors, and readers have only to walk through in imagination to become part of whatever world has been created or recreated by the author. When lighting conditions are just right, however, a window can also be a mirror. Literature transforms human experience and reflects it back to us, and in that reflection we can see our own lives and experiences as part of the larger human experience. Reading, then, becomes a means of self-affirmation, and readers often seek their mirrors in books (p. ix).

For children and families of color, oftentimes books function as windows and less as mirrors. In her book, *Telling Tales: The Pedagogy and Promise of African American Literature for Youth*, Johnson (1990) writes, "We use children's books as agents of socialization, politicization, and of formal education" (p. 1). Children of color may miss out on important underlying messages when they are deprived of opportunities to read books written by and about people of color. Johnson (1990) contends that African American children's literature "at its foundation demonstrates an abiding concern for the psychological, pedagogical, esthetic, intellectual, and spiritual needs of African American youth" (p. 3). This body of literature has underlying themes that are important for African American youth such as an emphasis on Black history, battling racial injustice, and race pride.

Many of these underlying themes were delineated by W.E.B. Du Bois, Jessie Redmon Fauset, and Augustus Granville Dill, the creators of the *Brownies' Book Magazine* (1920–1921)—one of the first periodicals created primarily for African American children. Du Bois (1919) articulated several objectives he wanted to accomplish with this magazine. A few of them were "To make colored children realize that being colored is a normal, beautiful thing," "To make them familiar with the history and achievements of the Negro race" and "To make them know that other colored children have grown into useful, famous persons" (p. 286). These themes continue to be of significance and are addressed in the works of many contemporary African American authors (McNair, 2008).

In spite of the importance of children of color seeing themselves in books, statistics published by the Cooperative Center for Children's Books show that there are significantly fewer books published by authors of color annually (CCBC, 2020). However, there are still a significant number when considering titles that have been published over the last 30 years. McNair (2015) completed a project with 10 African American families—with children in grades K-2—in which each family was provided with more than 50 children's books written by and about African Americans. There were a series of monthly 2 h workshops (from March until July) at which each family could select 10 books. The purpose of McNair's project was to examine what happened when these families were exposed to an abundance of books written by and about African Americans. Findings suggested that parents increased the amount of time reading aloud to their children, passed along the information they learned about African American children's literature to family, friends, and co-workers, and began seeking out and developing an appreciation for quality African American children's literature.

McNair (2015) provided families with information about notable African American authors and illustrators so they could seek out titles by these individuals at local bookstores or libraries and make requests that they be made available for purchase or to check out if unavailable. The project ended with a presentation by James E. Ransome, an award-winning illustrator of notable titles such as *Uncle Jed's Barbershop* (Mitchell, 1993), *Aunt Flossie's Hats (and Crab Cakes Later)* (Howard, 1991), and *Visiting Day* (Woodson, 2002). McNair shared titles across multiple genres (e.g., fantasy, historical fiction, contemporary realistic fiction) that included "classics" such as *Everett Anderson's Goodbye* (Clifton, 1983) and *Flossie and the Fox* (McKissack, 1986). Components of this program included high-quality children's books written by and about African Americans, culturally relevant materials such as dolls and "people colors" crayons and pencils, and instruction for parents on activities such as how to conduct interactive read alouds and have children respond in various modes such as art and writing. Equally important as well was the welcome inclusion of the entire family (and extended members such as grandparents) in the workshops. This was key in that grandparents who are often caregivers can be instrumental in the literacy development of children.

Providing children and their families with access to digital books is necessary with the increasing usage of technological devices such as iPhones. Epic, for example, is a free digital reading platform with more than 40,000 titles from 250 plus publishers. A perusal of this platform reveals picturebooks and transitional chapters books written by and about African Americans authors such as Nikki Grimes. Digital access can make it easier for children who may come from lower income families to read books. Last, sharing children's books in settings outside of the US featuring children of color can be insightful as well. For example, for African American children, allowing them to see representations of the Black experience set outside of the US can raise their sense of global awareness. Titles such as *The Field* (Paul, 2018), a story of young children playing soccer on the Caribbean island of Saint Lucia, can help children recognize the complexity and variety within the Black experience.

Conclusion

Nelson Mandela stated that "education is the most powerful weapon which you can use to change the world." We concur. As the foundation of education, literacy is important. It is a global affair. In this chapter, we have indicated that a key element missing from many approaches to family literacy for families of color is the use of the cultural assets possessed by these communities as a vehicle for empowerment and social change. We have illustrated how multiple literacies can function as a basis for valuing the funds of knowledge that families of color possess while also allowing them to develop agency surrounding their futures and the changes they wish to see in their communities. We have argued, based on notions of culturally relevant literacy, for a shift in conception and in practice, from family literacy to *family literacies*.

As Kofi Annan seventh Secretary-General of the United Nations, has rightly indicated:

Literacy is a bridge from misery to hope. It is a tool for daily life in modern society—For everyone, everywhere, literacy is along with education in general, a basic human right—Literacy is, finally, the road to human progress and the means through which every man, woman and child can realize his or her full potential.

For families of color in the United States and across the globe, we envision that family literacy programs will be re-envisioned to center culturally relevant literacies that empower children, their parents and communities, to create and enact social change.

References

- Adra, N., 2004. Literacy through poetry: a pilot project for rural women in the Republic of Yemen. *Wom. Stud. Q.* 32 (1/2), 226–243.
- Anderson, A.B., Stokes, S.J., 1984. Social and institutional influences on the development and practice of literacy. In: Goelman, H., Oberg, A., Smith, F. (Eds.), *Awakening to Literacy*. Heinemann, Exeter, NH, pp. 24–37.
- Anyia, U., 2020. African Americans in world language study: the forged path and future directions. *Annu. Rev. Appl. Ling.* 40, 97–112.
- Association of Centers for the Study of Congress, 2021. Elementary and Secondary Education Act. Retrieved from: <http://acsc.lib.udel.edu/exhibits/show/legislation/esea>.
- Auerbach, E.R., 1989. Toward a social-contextual approach to family literacy. *Harv. Educ. Rev.* 59, 165–181.
- Bacon, C.K., 2017. Multilanguage, multipurpose: A literature review, synthesis, and framework for critical literacies in English language teaching. *Journal of Literacy Research* 49 (3), 424–453.
- Barton, D., 1994. Exploring family literacy. *Rapal Bull.* 24, 2–5.
- Bauer, E.B., Arazi, J., 2011. Promoting literacy development for beginning English language students. *Read. Teach.* 64 (5), 383–386.
- Bauer, E.B., 2000. Code-switching during shared and independent reading: lessons learned from a preschooler. *Res. Teach. Engl.* 35 (1), 101–130.
- Bell, Y., Clark, T., 1998. Culturally relevant reading material as related to comprehension and recall in African-American children. *J. Black Psychol.* 24 (4), 455–475.

- Bishop, R.S., 1990. Mirrors, windows, and sliding glass doors. *Perspectives* 6 (3), ix–xi.
- Bouillion, L., Gomez, L., 2001. Connecting school and community with science learning: real world problems and school-community partnerships as contextual scaffolds. *J. Res. Sci. Teach.* 38 (8), 878–898.
- Brown, A., 2013. Waiting for superwoman: white female teachers and the construction of the “neoliberal savior” in a New York City public school. *J. Crit. Educ. Pol. Stud.* 11 (2), 123–164.
- Cairney, T.H., 1997. Acknowledging diversity in home literacy practices: moving towards partnership with parents. *Early Child. Dev. Care* 127–128, 61–73.
- Chall, J.S., Snow, C., 1982. Families and Literacy: The Contributions of Out of School Experiences to Children’s Acquisition of Literacy. A final report to the National Institute of Education. Harvard Graduate School of Education, Cambridge, MA.
- Clark, M.M., 1976. *Young Fluent Readers*. Heinemann Educational, London.
- Clifton, L., 1983. *Everett Anderson’s Goodbye*. Henry Holt.
- Compton-Lilly, C., Rogers, R., Lewis, T.Y., 2012. Analyzing epistemological considerations related to diversity: an integrative critical literature review of family literacy scholarship. *Read. Res. Q.* 47 (1), 33–60.
- Cooperative Children’s Book Center, 2020. Books by and/or About Black, Indigenous and People of Color 2018–. Cooperative Children’s Book Center, University of Wisconsin–Madison. Retrieved from: <https://ccbc.education.wisc.edu/literature-resources/ccbc-diversity-statistics/books-by-and-or-about-poc-2018/>.
- Darling, S., 1993. Family literacy: an intergenerational approach to education. *Viewpoints* (15), 2–5.
- DeBondt, M., Willenberg, I.A., Bus, A.G., 2020. Do book giveaway programs promote the home literacy environment and children’s literacy-related behavior and skills? *Rev. Educ. Res.* 90 (3), 349–375.
- Delgado-Gaitan, C., 1987. Mexican adult literacy; new directions for immigrants. In: Goldman, S.R., Trueba, K. (Eds.), *Becoming Literate as Second Language Learners*. Ablex, Norwood, NJ, pp. 9–32.
- Desmond, S., Elfert, M., 2008. *Family Literacy: Experiences From Africa and Around the World*. UNESCO Institute for Lifelong Learning, Feldbrunnenstrasse 58, 20148 Hamburg, Germany.
- Dickinson, D., McCabe, A., Essex, M.A., 2006. A window of opportunity we must open to all: the case for high-quality support for language and literacy. In: Dickinson, D.K., Neuman, S.B. (Eds.), *Handbook of Early Literacy Research*. Guilford Press, New York, NY, pp. 11–28.
- Doyle, A., 2012. Family literacy programs: where have they come from and where are they going? *Front. Educ. China* 7 (1), 85–102.
- Du Bois, W.E.B., 1919. The true brownies. *Crisis* 18 (6), 285–286.
- Edwards, P.A., White, K.L., 2018. Working with racially, culturally, and linguistically diverse learners, families and communities: strategies for preparing preservice teachers. *J. Family Divers. Educ.* 3 (1), 1–22.
- Edwards, P.A., Pleasants, H.M., Franklin, S.H., 1999. *A Path to Follow: Learning to Listen to Parents*. Heinemann, Portsmouth, NH.
- Edwards, P.A., Spiro, R.J., Domke, L.M., Castle, A.M., White, K.L., Peltier, M.R., Donohue, T.H., 2019. *Partnering With Families for Student Success: 24 Scenarios for Problem Solving With Parents*. Teachers College Press, New York, NY.
- Edwards, P.A., 1995. Empowering low-income mothers and fathers to share books with young children. *Read. Teach.* 48 (7), 558–564.
- Edwards, P.A., 2016. *New Ways to Engage Parents: Strategies and Tools for Teachers and Leaders, K–12*. Teachers College Press, New York, NY.
- Ellison, T.L., Solomon, M., 2019. Counter-storytelling vs. deficit thinking around African American children and families, digital literacies, race, and the digital divide. *Res. Teach. Engl.* 53 (3), 223–244.
- Erickson, F., 1989. Forward. Literacy risks for students, parents, and teachers. In: Allen, J., Mason, J. (Eds.), *Risk Makers, Risk Takers. Risk Breakers: Reducing the Risks for Young Literacy Learners*. Heinemann, Portsmouth, NH, pp. xiii–xvi.
- Even Start Family Literacy Program, 2014. William F. Goodling Even Start Family Literacy Program. Retrieved from: <https://www2.ed.gov/programs/evenstartformula/index.html>.
- Faircloth, S., Thompson, N., 2012. Meeting the needs of American Indian and Alaska native families. In: *Handbook of Family Literacy*. Routledge, New York, NY, pp. 255–269.
- Ferris, R., 2015. World Population: Quarter of Earth Will Be African in 2050. Retrieved from: <https://www.cnbc.com/2015/07/30/world-population-quarter-of-earth-will-be-african-in-2050.html>.
- Flores, N., 2020. From academic language to language architecture: Challenging raciolinguistic ideologies in research and practice. *Theory Into Practice* 59 (1), 22–31. <https://doi.org/10.1080/00405841.2019.1665411>.
- Freire, P., 1985. *The Politics of Education*. Bergin-Garvey Publishers, Cambridge, MA.
- Giroux, H., 1988. Postmodernism and the discourse of educational criticism. *J. Educ.* 170 (3), 5–30.
- Goldenberg, C.C., 1984. Low-income parents’ contributions to the reading achievement of their first-grade children. In: *Paper Presented at the Meeting of the Evaluation Network/ Evaluation Research Society*, San Francisco.
- Gonzalez, N., Moll, L.C., Amanti, C., 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, N.J.
- Gunderson, L., Clarke, D.K., 1998a. An exploration of the relationship between ESL students’ backgrounds and their English and academic achievement. In: *National Reading Conference Yearbook*, vol. 47, pp. 264–273.
- Gunderson, L., Clarke, D., 1998b. An exploration of the relationship between ESL students’ backgrounds and their English and academic achievement. In: Shanahan, T., Rodriguez-Brown, F. (Eds.), *Yearbook of the National Reading Conference*, vol. 47. National Reading Conference, Chicago, IL.
- Hadley, E.B., Newman, K.M., Mock, J., 2020. Setting the stage for TALK: strategies for encouraging language-building conversations. *Read. Teach.* 74 (1), 39–48.
- Hanemann, U., McCaffery, J., Newell-Jones, K., Scarpino, C., 2017. *Learning Together Across Generations: Guidelines for Family Learning and Literacy Programmes*. UIL, Hamburg.
- Heath, S.B., 1983. *Ways with words: Language, life, and work in communities and classrooms*. Cambridge University Press, Cambridge, MA.
- Howard, E.F., 1991. *Aunt Flossie’s Hats (and Crab Cakes Later)*. Clarion Books.
- International Literacy Association, 2020. 2020 What’s Hot in Literacy Report. Retrieved from: <https://www.literacyworldwide.org/get-resources>.
- Johnson, D., 1990. *Telling Tales: The Pedagogy and Promise of African American Literature for Youth*. Greenwood Press.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Lazar, A.M., Edwards, P.A., McMillon, G.T., 2012. Bridging literacy and equity: The essential guide to social equity teaching. Teachers College Press, New York.
- Legal Information Institute, 1994. 20 U.S. Code § 5812. National Education Goals. Retrieved from: <https://www.law.cornell.edu/uscode/text/20/5812>.
- Li, G., 2000. Family literacy and cultural identity: an ethnographic study of a Filipino family in Canada. *McGill J. Educ. /Revue des sciences de l’éducation de McGill* 35 (1).
- Li, G., 2012. *Culturally Contested Pedagogy: Battles of Literacy and Schooling Between Mainstream Teachers and Asian Immigrant Parents*. SUNY Press, Albany, NY.
- Library of Congress, 1991. H.R. 751 (102nd): National Literacy Act of 1991. Retrieved from: <https://www.congress.gov/bills/102nd-congress/house-bill/751>.
- Library of Congress, 1998. H.R.1385—Workforce Investment Partnership Act of (1997–1998). Retrieved from: <https://www.congress.gov/bills/105th-congress/house-bill/1385>.
- Literacy Information and Communication System (LINCS), 2021. Frequently Asked Questions. US Congress 1998 Adult Education and Family Literacy Act. Retrieved from: <https://lincs.ed.gov/faq#measure>.
- Luke, A., Luke, C., 2001. Adolescence lost/childhood regained: on early intervention and the emergence of the techno-subject. *J. Early Child. Lit.* 1 (1), 91–120.
- Luke, A., 2012. Critical literacy: foundational notes. *Theor. Pract.* 51 (1), 4–11.
- Mace, J., 1998. *Playing With Time: Mothers and the Meanings of Literacy*. UCL Press limited, London.
- Macleod, F., 1996. Integrating home and school resources to raise literacy levels of parents and children. *Early Child. Dev. Care* 117, 123–132.
- Marciano, J.E., Watson, V.W., 2021. “This is America”: Examining artifactual literacies as austere love across contexts of schools and everyday use. *The Urban Review* 53 (2), 334–353.
- McKissack, P.C., 1986. *Flossie and the Fox*. Dial Press, New York, NY.

- McNair, J.C., 2008. A comparative analysis of *the Brownies' Book* and contemporary African American children's literature written by Patricia C. McKissack. In: Brooks, W.M., McNair, J.C. (Eds.), *Embracing, Evaluating, and Examining African American Children's and Young Adult Literature*. Scarecrow Press, Lanham, MD, pp. 3–29.
- McNair, J.C., 2015. "I never knew there were so many books about us": parents and children reading and responding to African American children's literature together. *Child. Lit. Educ.* 44, 191–207.
- Meier, T., 2003. "Why can't she remember that?" the importance of storybook reading in multilingual, multicultural classrooms. *Read. Teach.* 57 (3), 242–252.
- Miner IV, H.R., 2012. Beyond a test score: explaining opportunity gaps in educational practice. *J. Black Stud.* 43 (6), 693–718.
- Mitchell, M.K., 1993. *Uncle Jed's Barbershop*. Simon and Schuster, New York, NY.
- Moje, E.B., Ciechanowski, K.M., Kramer, K., Ellis, L., Carrilo, R., Collazo, T., 2004. Working toward third spaces in content area literacy: an examination of everyday funds of knowledge and discourse. *Read. Res. Q.* 39 (1), 38–70.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992a. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Qual. Issues Educ. Res.* 32 (2), 132–141.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992b. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141.
- Moll, L., 2005. Reflection and Possibilities. In: Gonzalez, N., Moll, L.C., Amanti, C. (Eds.), *Funds of knowledge: theorizing practices in households, communities, and classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Morrow, L., Paratore, J., Gaber, D., Harrison, C., Tracey, D., 1993. Family literacy: perspective and practices. *Read. Teach.* 47 (3), 194–200.
- Morrow, L.M., 1994. *Family Literacy: New Perspectives, New Opportunities*. International Reading Association, Newark, DE.
- Morrell, E., Duncan-Andrade, J.M.R., 2002. Promoting academic literacy with urban youth through engaging hip-hop culture. *The English Journal* 91 (6), 88–82.
- Nalubega-Booker, K., Willis, A., 2020. Applying Critical Race Theory as a tool for examining the literacies of Black immigrant youth. *Teach. Coll. Rec.* 122 (13), 1–24.
- National Public Radio, 2004. Lyndon Johnson's War on Poverty. Retrieved from: <https://www.npr.org/templates/story/story.php?storyId=1589660>.
- Nieto, S., 1992. *Affirming Diversity: The Sociopolitical Context of Multicultural Education*. Longman, New York, NY.
- National Institute for Literacy (NIFL), 2001. Developing Early Literacy Report of the National Early Literacy Panel. Retrieved from: <https://lincs.ed.gov/publications/pdf/NELPReport09.pdf>.
- Organisation for Economic Development, 1996. International Adult Literacy Survey. Retrieved from: <http://www.oecd.org/education/skills-beyond-school/41529765.pdf>.
- Panelli, R., Punch, S., Robson, E. (Eds.), 2007. *Global Perspectives on Rural Childhood and Youth: Young Rural Lives*. Routledge, London, England.
- Paris, S.G., 2005. Reinterpreting the development of reading skills. *Read. Res. Q.* 40 (2), 184–202.
- Paris, D., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Paul, B., 2018. *The Field*. NorthSouth Books, New York, NY.
- Pellegrini, A., 1991. A critique of the concept of at risk as applied to the concept of emergent literacy. *Lang. Arts* 68, 380–385.
- Race of the world population. (2022). Retrieved from <https://infogram.com/race-of-the-world-population-1go502yg18k62jd>.
- Razfar, A., Troiano, B., Nasir, A., Yang, E., Rumenapp, J.C., Torres, Z., 2015. Teachers' language ideologies in classroom practices: Using English learners' linguistic capital to socially re-organize learning. In: Smith, Patriann (Ed.), *Handbook of research on cross-cultural approaches to language and literacy development*. IGI Global Publications, Hershey, PA, pp. 261–298.
- Reach Out and Read (ROR), 1992. About Us. Retrieved from: <https://reachoutandread.org/about/>.
- Reading is Fundamental, 2021. About Us. Retrieved from: <https://www.rif.org>.
- Riojas-Cortez, M., Flores, B., 2009. Supporting preschoolers' social development in school through funds of knowledge. *J. Early Child. Res.* 7 (2), 185–199.
- Rosa, J., Flores, N., 2017. Unsettling race and language: toward a raciolinguistic perspective. *Lang. Soc.* 46 (5), 621–647.
- Saracho, O.N., 2017. Parents' shared storybook reading—learning to read. *Early Child. Dev. Care* 187 (3–4), 554–567.
- Scarborough, H.S., Dobrich, W., 1994. On the efficacy of reading to preschoolers. *Dev. Rev.* 14 (3), 245–302.
- Schwartz, W., 1999. Family literacy strategies to support children's learning. ERIC digest. In: ERIC Clearinghouse on Urban Education. Teachers College, Columbia University, New York, NY. eric-web.tc.columbia.edu.
- Scott, K.S., Brown, J., Jean-Baptiste, E., Barbarin, O., 2012. A socio-cultural conception of literacy practices in African American families. In: *Handbook of Family Literacy*. Taylor & Francis, New York, NY, pp. 239–254.
- Shockley, B., 1994. Extending the literate community: home-to-school-and school-to-home. *Read. Teach.* 47 (6), 500–502.
- Siegel, J., 2012. Educational approaches for speakers of pidgin and creole languages. In: Yiaikoumetti, A. (Ed.), *Harnessing Linguistic Variation to Improve Education*. Peter Lang, Oxford, UK, pp. 259–292.
- Smith, P., Marfo, K., 2015. Toward cross-sector integration of early childhood services in St. Lucia: using historical, policy and linguistic factors to inform future progress. In: Berson, M., Berson, I. (Eds.), *Global Child Advocacy Series (Volume 6): Child Advocacy and Early Childhood Education Policies in the Caribbean*. Information Age Publishing, Charlotte, NC, pp. 17–48.
- Smith, P., Lee, J., Chang, R., 2020. Characterizing competing tensions in Black immigrant literacies: beyond partial representations of success. *Read. Res. Q.* <https://doi.org/10.1002/RRQ.375>.
- Smith, P., 2019. (Re) Positioning in the Englishes and (English) literacies of a Black immigrant youth: towards a transraciolinguistic approach. *Theor. Pract.* 58 (3), 292–303.
- Smith, P., 2022. A transraciolinguistic approach for literacy classrooms. *Read. Teach.* <https://doi.org/10.1002/trr.2073>.
- Smythe, S., Toohey, K., 2009. Investigating sociohistorical contexts and practices through a community scan: a Canadian Punjabi-Sikh example. *Lang. Educ.* 23 (1), 37–57.
- Street, B., 1995. *Social literacies: Critical approaches to literacy in development, ethnography and education*. Routledge, New York.
- Strickland, D.S., Morrow, L.M., 1989. Oral language development: children as storytellers. *Read. Teach.* 44 (3), 260–261.
- Taylor, D., 1983. *Family Literacy: Young Children Learn to Read and Write*. Heinemann, Portsmouth, NH.
- The Laura Bush Foundation for America's Libraries, 2021. About. Retrieved from: <http://www.laurabushfoundation.com/about/index.html>.
- The Nations Report Card, 2020. How Did U.S. Students Perform on the Most Recent Assessments? Retrieved from: <https://www.nationsreportcard.gov>.
- Thomas, A., Skage, S., 1998. Overview of perspectives on family literacy: Research and practice. In: Thomas, A. (Ed.), *Family literacy in Canada: Profiles of effective practices*. Soleil, Welland, Canada, pp. 5–24.
- Tomasello, M., 1999. The human adaptation for culture. *Annual Review of Anthropology* 28 (1), 509–529.
- U.S. Department of Education, 1998. Adult Education and Family Literacy Act of 1998. Retrieved from: <https://www2.ed.gov/policy/adulted/leg/legis.html>.
- U.S. Department of Education, 2004. Title I—Improving the Academic Achievement of the Disadvantaged. Retrieved from: <https://www2.ed.gov/policy/elsec/leg/esea02/pg1.html>.
- U.S. Department of Health and Human Services, 2019. History of Head Start. Office of Head Start: An Office of the Administration for Children & Families. Retrieved from: <https://www.acf.hhs.gov/ohs/about/history-head-start>.
- UNESCO Institute for Lifelong Learning, 2017. UIL Policy Brief 9: Engaging Families in Literacy and Learning. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000249463>.
- UNESCO, 2008. *Family Literacy—A Global Approach to Lifelong Learning Effective Practices in Family Literacy and Intergenerational Learning Around the World*. Retrieved from: <https://uil.unesco.org/literacy/learning-families/family-literacy-global-approach-lifelong-learning-effective-practices>.
- UNESCO, 2019. *Learning Together Through Life: Building Capacity in Family Literacy*. <https://uil.unesco.org/literacy/learning-families/learning-together-through-life-building-capacity-family-literacy>.

- Velliari, D.M., 2015. International family configurations in Tokyo and their cross-cultural approaches to language socialization. In: Smith, P., Kumi-Yeboah, A. (Eds.), *Handbook of Research on Cross-Cultural Approaches to Language and Literacy Development*. IGI Global, Hershey, PA, pp. 57–86.
- White House, 2009. White House Biography of Barbara Bush. Archived from the original on March 3, 2009. Retrieved from. https://web.archive.org/web/20090303155711/http://www.whitehouse.gov/about/first_ladies/barbarabush/.
- Wiley, T.G., 1996. Literacy and language diversity in sociocultural contexts. In: *Literacy and Language Diversity in the United States*. Center for Applied Linguistics and Delta Systems Co., Inc, Washington, DC.
- Willis, A.I., 2015. Literacy and race: access, equity, and freedom. *Lit. Res.* 64 (1), 23–55.
- Willis, A.I., 2018. Re-positioning race in English language arts research. In: Lapp, D., Fisher, D. (Eds.), *Handbook of Research on Teaching the English Language Arts*. Routledge, New York, NY, pp. 30–56.
- Woodson, J., 2002. *Visiting Day*. Scholastic, New York, NY.
- Yaden Jr., D.B., Paratore, J., 2003. Family literacy at the turn of the millennium: the costly future of maintaining the status quo. In: Flood, J., Lapp, D., Squire, J.R., Jensen, J. (Eds.), *Handbook on Teaching the English Language Arts*, second ed. Erlbaum, Mahwah, NJ, pp. 532–545.

Literacy and identity across home, school, church, and social life

Gilberto P. Lara^a, Sonia N. Sánchez^a, Myriam Jimena Guerra^b, and Lucila D. Ek^a, ^aThe University of Texas at San Antonio, San Antonio, TX, United States; and ^bTexas A&M University-San Antonio, San Antonio, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	198
Theoretical framework: sociocultural theories of learning and development	198
Language socialization (LS) and identities	198
New literacy studies	199
Literature review on literacy in multiple contexts	200
Home and community literacies	200
Home literacies	200
Non-mainstream literacies	201
Church literacies	201
In-school literacies	202
Our research on literacy and identity	203
Research on literacy and gender of Latino male teachers and their bilingual students	203
Literacy and identity of pentecostal students in church and school	204
Research on digital biliteracy: La Clase Mágica at UTSA (LCM@UTSA)	204
Research on home and family literacies of emergent bilingual students	205
Research on language and literacy of Latina/o/x lowriders	205
Conclusion	206
References	206

Introduction

Literacy and identity across home, school, church, and social life is an important area of study which we have researched extensively. This encyclopedia entry presents the authors' research on various literacies grounded in sociocultural theories of learning and development (Vygotsky, 1978) and using qualitative methods. In particular, we discuss scholarship that is informed by language socialization (Ochs, 1995), New Literacy Studies (NLS) (Street, 1998), as well as theories of language, literacy, and identity (Ibrahim, 2016). An interdisciplinary perspective is key to understanding the experiences of students who move across and negotiate varied and multiple literacies in their daily lives.

As teacher educators who were formerly bilingual teachers, we center our studies on bi/multilingual children, adolescents, and adults in key cultural contexts. Our ways of viewing the literacy practices in these settings are informed by key scholarship on various literacies including home, non-mainstream, church, in-school contexts which we present in the literature review section. First, though, we discuss the overarching theoretical framework that informs our scholarship.

Theoretical framework: sociocultural theories of learning and development

Rooted in Vygotskian principles of learning and development, sociocultural theory provides a lens that illuminates how social and cultural factors affect literacy learning. The theory's focus on learning as a socially mediated activity and its view of language as the preeminent meaning making tool links mind, language, and society (Vygotsky, 1978). By using language as both a resource and a tool, individuals acquire knowledge in contexts that are culturally and socially organized. Cultural tools mediate novices' (both children and adults) becoming competent members of the various and diverse settings and cultural groups that make-up their daily lives. Thus, individual experiences are connected to social relations and individuals are linked to daily interactions and situated practices. Hence, the theory emphasizes the relationship between the individual and their communities. This overarching theory also affords a lens for examining literacy and identity. The Language Socialization (LS) paradigm deepens the view of language as a cultural tool for mediating learning and development, and for constructing identities.

Language socialization (LS) and identities

According to LS, children's acquisition of culture occurs in interaction with more expert members; language is at the center of this process (Ochs and Schieffelin, 1995). Through language socialization children gain an understanding of the social and cultural organization of their everyday lives, including knowledge, skills, values, beliefs, and social structures as they participate in language

and literacy activities. Hence, children are socialized through the use of language and literacy and to use language and literacy. The LS perspective looks at the role of social identity in the process of becoming a competent member of a community (Capps, 1999). In the socialization process, expert members of social groups use language to socialize children and novices to particular identities and roles which these children learn to express by using language in certain ways as they participate in complex social and linguistic interactions including literacy practices.

The LS perspective on identity differs from notions that are rooted in traditional psychological perspectives in significant ways. Language socialization acknowledges sociocultural factors and the diversity of forms in which development occurs and thus counters a collective and stereotypical identity for culturally and linguistically diverse students. What a child does and says and what others do and say to and around a child must be considered if we are to understand how identity is constructed. Language, then, is central to the construction of identity. Sociolinguists posit that “[s]ocial identity and ethnicity are in large part established and maintained through language” (Gumperz and Cook-Gumperz, 1982, p. 7). Moreover, LS’s particular view of identity/ies as being dynamic, fluid, and emerging in social interactions moves away from traditional psychological notions of identity as being static, rigid, and occurring solely within the individual—a move which affords a better lens with which to study the complicated situation of culturally and linguistically diverse (CLD) students whose identities may be context specific. Thus, children are socialized into social identities through discourse as scholars in New Literacy Studies (NLS) have pointed out.

New literacy studies

The turn of the century marked a shift in how literacy was understood beyond psychoanalytic lenses and the dichotomy between literate and nonliterate abilities of literacy. New Literacy Studies (NLS) originated from the confluence of various fields of discipline (Barton and Hamilton, 2012; Heath, 1983; Scribner and Cole, 1981) that sought to understand traditional thought and its relationship between literacy and cognitive development. Literacy, as defined by NLS (Gee, 2015; Street, 1984, 2003) considers the social contexts that inform literacy events and practices. Whereas the traditional view of literacy defines it as the ability to read and write, NLS situates literacy within a social perspective, challenging views that make invisible the social implications of literacy learning. Theorists who have made efforts to dichotomize literacy into oral and written discourse, approach literacy as autonomous, devoid of social influences, and characterized by cognitive skills (Hatch, 2002; Hymes, 1978). It is a view dominated by Western perspectives of what it means to be literate, often as a set of universal skills independent of social practices and their contexts. Further, the autonomous model is often associated with literacy programs and campaigns that often associate the acquisition of literacy with progress, civilization, individual liberty, and social mobility (Street, 2014). Byrd, Hayes, and Turnipseed (2021) speak of the pervasiveness of the autonomous model in education practice and policy to affirm dominant cultural practices while devaluing others. These autonomous models, add Byrd et al., exist in systems that uphold oppressive structures that negate access to certain forms of capital to culturally and linguistically diverse students. In contrast, NLS posits literacy as a socially constructed concept involving the intersections of power, beliefs, values, purposes, and economic and political factors. In this model, literacy is always embedded within the social institutions that define and practice the relationships of power and knowledge.

In line with Vygotskian notions of literacy, Gee (2015) constructs literacy within a sociocultural context arguing that language and literacy need to be studied in its social context, exceeding the limits of the traditional view of literacy studies. Therefore, Gee defines literacy in relation to Discourses. In his work on D/discourses, Gee proposes that language and literacy are more than just speaking, reading, and writing, but rather a part of a larger discourse that encompasses how literacy is valued, acted, transmitted, talked about, thought about, and believed to be. Gee’s work also analyzes what ideas, beliefs, and meanings are behind that language and extends the notion of D/discourses by claiming that we acquire a primary Discourse through our socialization with family and community. These Discourses are connected to a variety of literacies that are specific to that social context. On the other hand, secondary Discourses can only be learned in literacy settings that carry power. An example of such a setting is that of academic institutions and the learning that takes place behind those walls. By understanding the relationship between language and social and cultural contexts, we can also understand literacy through a sociocultural perspective.

Lastly the field of NLS also spawned the theory of *multiliteracies*. The theory of multiliteracies focuses on modes of representation that extend beyond print and written texts. An argument presented by the New London Group considers the “increasing multiplicity and integration of significant modes of meaning-making, where the textual is also related to the visual, the audio, the spatial, the behavioral, and so on” (Cazden et al., 1996, p. 64). The emphasis of modes as a channel for expressing meaning is what Kress (2010) terms multimodality. Multimodality describes the different communicative channels that include images, printed text, speech, color, and moving images. Scholars of multiliteracies do not limit their definitions of literacy to exclusively include print literacy, instead they expand on a broader definition that works to privilege multiple literacies evident in communities outside of schools (Perry, 2012). The implications of expanding the traditional notions of literacy through multiliteracies provide a greater opportunity for navigating a world that is constantly changing, thus placing demands on the ability to draw on multiple literacies. Through a sociocultural perspective, the production and consumption of digital texts becomes a new form of social practice that focuses attention on the way we engage in these literacy practices as members of some Discourse. If we consider Gee’s (2015) work on Discourses as described earlier in this section, then we can approach digital literacies as the language and literacy practices of discourse communities where members interact with text to redefine the meaning of literacy and to construct identities.

Literature review on literacy in multiple contexts

Home and community literacies

Rooted in a sociocultural paradigm, literacy practices are seen as highly contextual and diverse in specific settings. Through this lens, literacy becomes multiple and is shaped within the sociocultural and sociolinguistic contexts in which it is embedded (Schieffelin and Ochs, 1986). However, it is not enough to only study literacy in academic settings. For this reason, the literature in this section is representative of literacies across varying communities. The literature as it relates to the use of reading and writing outside of school is documented in ethnographies where local literacies are captured as people manage households, communicate with others, partake in leisure activities, and simply live their lives (Barton and Hamilton, 2012; Schieffelin and Ochs, 1986). Without a complete understanding of what people do with literacy in their everyday lives, misinterpretations can lead to disparities between varying sets of literacies. The need for this understanding lies in the scholarship that exists in out-of-school literacies.

Home literacies

Much of the deficit thinking on the achievement of culturally and linguistically diverse students like Latinas/os/xs stems from research that ignores and devalues their cultural strengths in favor of more dominant pedagogies (Gorski, 2006; Trueba, 1988). A response to these deficit views are those studies that aim to deconstruct thinking that fails to recognize the value of literacy outside of schools.

Much of what is missing in recent discussions of literacy is the reading and writing that exists beyond traditional texts. Prevailing conceptions emphasizing the relationship between literacy and print continues to place books as the only source of text. However, if we approach the term “text” through a sociocultural lens, then we can refer to Freire’s (1970) construct of the word as the ability to read the world and the word. In line with a more progressive view, Freire believes that one cannot read a text without situating the text in the context in which it finds itself. Almost simultaneously, the process of “reading” happens by understanding the text and its relationship to its social sphere. In this sense, the much-contested terms “text” and “reading” transform in meaning to encompass the analysis of gestures, moods, symbols, signs, and art. Text essentially becomes anything from which we can create meaning (New London Group, 2000). The studies in this section challenge the dominant notion that reading can only happen through books and highlight literacy as something more than what we learn at school.

Mui and Anderson (2008) take a sociocultural approach to study the literacy practices found in one Indo-Canadian family, consisting of a pair of older parents, their sons, and the sons’ families. As part of this multicultural dynamic, English and Punjabi were widely spoken. What Mui and Anderson found was that the Johar household contained a “rich tapestry” of literacy practices. Non-traditional literacies included the singing of Punjabi and English songs, especially at bedtime rather than the more dominant practice of reading bedtime stories. The literacy practices supported by the adults within a multicultural, multilingual context, uncovered the different pathways this family engaged in literacy.

A similar study that recognized the value of literacies occurring outside of school, was Johnson’s (2010) study of an African American family living in a small rural community in the southeastern United States. Informed by a sociocultural perspective of literacy, this study offered additional insight into the value that families assign to literacy by engaging in everyday tasks, religious services, and family events. The multiplicity of literacy found in the Jones family was unmistakable evidence that literacy is defined by the social context within which it is practiced (Perez, 2004). Within this family, literacy was used in multiple ways. Johnson describes seven ways in which the family is actively engaged in literacy: (a) interactional uses of literacy to build intra-relationships within the family such as letter-writing; (b) instrumental uses of literacy to perform everyday tasks such as cooking and sewing; (c) news-related uses of literacy to learn more about local and global events; (d) financial uses of literacy to organize their finances; (e) spiritual uses of literacy to attend to their spiritual lives; (f) recreational uses of literacy to entertain themselves; and (g) educational uses of literacy to support educational opportunities. This study helped to understand how one family’s practice and participation in literacy can shed light on the relationship between literacy and its historical and cultural context.

Research on Latina/o/x family literacy practices argues the importance of recognizing home literacies as valuable resources in literacy pedagogy. De la Piedra’s (2010) study of recent Mexican immigrant families living on the United States-Mexico border highlights the diversity of family literacy practices that included religious literacy and writing letters to family in Mexico to maintain relationships with relatives. The literacies practiced within the adolescent circles included sending emails, navigating the Internet to gather information about popular culture, and text messaging. The importance of the school-home connection is emphasized in this study as De la Piedra highlights how the young members of this family were able to engage in academic tasks when given the opportunity. This study also recognized that home literacies provide individuals with a set of literacies that can act as “funds of knowledge” (González et al., 2005) outside of the home.

Although everyday literacies are usually assigned a lower social value compared to more dominant literacies of schools and institutions, Mercado (2005) argues that “they are, in fact, hybrid practices that draw on the media and official literacies, as well as those household’s social networks” (p. 137). Mercado’s study of language and literacy in two Puerto Rican homes illustrates the value that literacy holds for these two families by describing how their literacy practices “are about making sense of and responding to lived experience” (p. 145). This study also concludes the importance of recognizing home literacies as funds that can help redefine what literacy learning means.

Research conducted on the literacy education of Latina/o/x adults is limited to the research on family literacy. Given this challenge, Rivera (2004) examines the experiences of Latina adults who participate in adult literacy education programs. In her study of Latinas in an adult literacy education class, Rivera challenges theories of “functional literacy” offered by proponents of adult literacy programs who only view literacy as a solution to the social and economic problems of larger cities. Through a popular curriculum where the learning content is connected to the real lives of learners, Rivera found that adult Latinas experienced a level of motivation to complete the program and foster community change. Her critique of the current focus on adult education for Latinas/os/xs, addresses the need to redefine the meaning and function of literacy as outlined by current adult literacy educational programs.

A seminal study in this area of family literacy research is Heath's (1983) work in the towns of Roadville and Trackton where class and racial factors were found to shape the literacy events found within those families. What she discovered was a discontinuity that existed between family literacies and the more dominant literacies valued in schools. Conceptions of literacy that are culturally based, such as those of the families in these two towns, have to recognize that literacy differs across communities and that not acknowledging such diversity contributes to the deficit views on home literacies.

As the studies in this section have shown, there is great significance in studying literacy outside of school. However, there is a greater challenge in acknowledging the significance of literacy practices within marginalized social groups. The following studies focus on a closer look at those literacies that exist within adolescent groups where much of their group activity is marked as diverging from mainstream standards.

Non-mainstream literacies

The following studies are centered around youth in non-mainstream communities to examine how the literacy and identity practices of these groups often go unrecognized in traditional literacy scholarship. Note that attention to these studies is not meant to glamorize any acts deemed unlawful. Nor do they diminish the negative implications involved in these acts. However, these studies do intend to create an understanding that adolescents engaged in unsanctioned social practices also engage in purposeful literate practices. As defined by Moje (2000), alternative or unsanctioned literacies include not only gang-related literacies such as tagging and graffiti, “but also a wide range of poetry, narrative, journal writing, letter writing, and novel reading” found in the literacy practices of most youth (p. 661). Moving beyond considering literacies of marginalized youth as representations of crime and resistance, Moje illustrates how unsanctioned literacy practices of these youth are in reality tools that not only transform thought and experience, but also make meaning about the everyday activities in their lives and shape these youth's identities. Her examination of five Vietnamese, Latina, Latino, Laotian, and Samoan participants' literacy practices was described in terms of three Discourses or the ways of “thinking, feeling, believing, valuing, and acting” (Gee, 2015, p. 161) that individuals use to create membership in a group: (1) written discourses such as the writing of raps, poetry, and parodies; the writing of gang names, street names, and/or gang symbols on notebooks, desks, walls, lockers, and as tattoos; (2) body discourses that included clothes and styles connected to a particular gang; and (3) oral discourses which include the accents, terminology, and dialects used in the gang. Moje (2000) found that tagging and graffiti practices were forms of literacy that became tools for transforming thought, identity, and power.

Alternative literacies mirror the current popular culture and practices of a given community. Although these practices are not novel in the lives of urban children and youth, they are part of a recent trend of research that frames these literacy practices from the perspective of the New Literacy Studies. The following studies research the multiple varieties of alternative literacies as illustrations of the uses of literacy that are specific to texts in given contexts.

Church literacies

A burgeoning body of work on children's and youth's religious learning highlights the church as a key site for literacy and construction of identities, particularly for culturally and linguistically diverse students. McMillon and Edwards' (2000) research on an African-American child who attended an African-American Baptist church found that he was seen as a “superstar” because of his participation in the Sunday school's literacy practices including recitation, retelling stories, and reading. On the other hand, his teachers in the public school considered him to be a child with behavior problems. A study of language and literacy socialization in a Catholic Church found that Latina/o/x immigrant children were socialized to a Mexican ethnic identity as they engaged in the literacy practices of their *doctrina* (catechism) class (Baquedano-López, 1997, 2004). Further highlighting the role of religion in the construction of identities for bi/multilingual students is Fader's (2001) research of a Hasidic community in New York. Fader found that the children's literacy experiences were differentiated by gender. Boys acquired literacy in liturgical Hebrew and Yiddish and studied religious texts all in Yiddish while girls were socialized to gendered roles through Yiddish and English language and literacy practices.

Hsin and Yu (2021) examined religious activities of Indigenous Taiwanese children involved in Catholic and Christian practices, as well as Indigenous ceremonies. Children negotiated their identities in the city and village due to the influence of varied religious practices. In the village, they formed their cultural identity through Rukai-related religious literacies but in the city, some children did not express their Rukai identity which may have been due to “limited access to Rukai linguistic and cultural resources” (Hsin and Yu, 2021, p. 331). Cun (2021) found how the use of religious texts at home created multilingual identities of Burmese refugee children in the United States. The families engaged in reading the Quran, praying, worship services in the mosque, and religious events with their parents. Through these religious literacies at home, children acquired Burmese, Arabic, and English.

Skerrett (2017) investigated how a Caribbean-Chinese child developed his religious identity through multiple language, religion, and cultural practices. The child developed a Christian identity influenced by the home and school setting, as well as the English dominant language. In another case study by Skerrett (2016), the religious identities and religious literacies of a Mexican American child were examined, along with the links the practices carried into the school social setting. The youth was constantly negotiating between her religious identities and literacies to those of her social setting. For example, “she reshaped her literate participation in secular contexts to better align with her religious identity” (Skerrett, 2016, p. 982) by no longer being a part of the hip-hop dance group at her school since she felt the lyrics of the song chosen to dance were not appropriate according to her Christian identity view.

Peñalva et al. (2014) examined transcultural spiritual literacy through artifactual mediations that were part of the cultural capital of a first-generation Latina child at a US church-affiliated community center. Change in identity was made evident as the child participated in transcultural navigations “through cultures, spiritualities, languages, and trajectories” (Peñalva et al., 2014, p. 101). When she shared the key holder as her artifact, the child also presented her rich literacies, such as using her artifact to tell a story, understanding Biblical information, and creating connections to her Honduran culture.

Children’s identity linked to writing during a Catholic First Communion preparation course was highlighted in a study by Tusting (2015). Children created Catholic identities through recontextualization of Catholic teaching that “highlighted the uniqueness of their individual biographies, supported them in verbalizing internal emotional experiences, highlighted their agency and their identity as active choosers, and situated them in relation to a particular set of broader communities” (Tusting, 2015, p. 238). Children were invited to reflect on their identity through texts and writing.

Poveda et al. (2006) examined how adults served as mediators for children in a Gitano Evangelist Church in Spain to learn how to interpret the Bible. Through outward and inward textuality orientations, students learn about authority, gender, and institutional roles depicted by the discussions of the readings from the religious texts. In another study, Rackley (2021) examined how Latter-day Saints youth engaged in religious literacies. Through observations during worship services and educational settings, Rackley found that motivation to learn about their faith was due to spiritual, relational, and educational factors. Additionally, demotivation in these areas were also recognized when there was a lack of influence and motivation. These motivations and de-motivations were linked to the identities youth constructed.

In-school literacies

Scholars have studied the contexts where children and adults participate in social and cultural practices to include the traditional classroom (Gutiérrez et al., 1995; 1997; Luke, 1994; Orellana, 1994; Rueda and Moll, 1994) and after school sites (Gutiérrez and Stone, 1997; Gutiérrez et al., 1999a; Vasquez, 2003). The schools exist as spaces where the rich family practices of the home can be leveraged for learning. An example of scholars that have created meaningful links between the children’s household practices and school social worlds is Moll and Greenberg (1992) who argue that by capitalizing on family and community resources, schools can improve instruction and serve more appropriately their diverse student populations, particularly second language learners. In this framework families and households are viewed as a source of knowledge and resources for curriculum design and classroom instruction. Moll and Greenberg (1992) employed ethnographic methods in the Southwest US to unveil the transmission of knowledge and resources among Latina/o/x households drawing from Velez-Ibañez (1983) early the concept, *funds of knowledge*. Scholars who focus culturally and linguistically diverse students have made clear the necessity of schools around the country to provide a more inclusive approach of their linguistic, literacy, and cultural strengths as resources during for learning (Fránquiz and Reyes, 1998; García, 2008; Jiménez, 2003; Mercado, 2005; Moll, 2005; Sánchez and Machado-Casas, 2009). Linking the knowledge from the home and community creates spaces where school age children can be positioned as writers and producers of knowledge in the classroom (Bauer and Gort, 2012; Jiménez, 2003; Moll et al., 2001; Reyes, 2006; de la Luz Reyes and Halcón, 2001). Vásquez craftily illustrated how children in the early grades possess the capacity to critically engage with texts and become critical readers (Vásquez, 2014). While studies illustrating how young bilingual learners are capable of sustaining literary conversations about or around texts (Martínez-Roldán, 2003) are scarce. Engaging with difficult topics such as discrimination, death, and patriarchy during interactive read alouds has been shown to be possible in the younger grades when carefully planned, managed and enacted (Lara and Leija, 2014; Lara and Fránquiz, 2015; Leija, 2017; Leija and Fránquiz, 2021).

Building on funds of knowledge and their use in the classroom, scholars have proposed an interrogation of the tools that mediate how identities are taken up, created and defined as funds of identity (Esteban-Guitart and Moll, 2014). These funds of identity are the accumulation of knowledge, interests, skills and artifacts that students collect throughout their lives in their home and community contexts to shape their identities (Esteban-Guitart, 2016). Fránquiz, Ortiz, and Lara make clear that while funds of knowledge are collected through ethnographic home visits, “the strategy for building up funds of identity requires opportunities to create and share identity artifacts in school and after-school contexts that allow teachers to identify skills, knowledge, significant people, and places to be linked to curricula” (Fránquiz et al., 2021, p. 406). Taking stock of students’ funds of identity allows for teachers to create spaces where students’ identities can be leveraged to make sense of texts in the classroom. Oikonomidoy and Karam (2020) illustrated how funds of identity can serve as a lens to how the family socialization practices contribute to children’s development of not just language but also their sense of agency and identity in the classroom.

Research which examines the hybrid language practices—that is, the strategic use of multiple codes and registers—of bilingual (English/Spanish) students includes studies of an elementary classroom (Gutiérrez et al., 1999b) and an after-school computer club (Gutiérrez et al., 1999a). These articles reframe the role of language in the classroom from the object of instruction to a tool for learning. Gutiérrez et al. (1999b) study of hybrid language practices in a dual immersion elementary classroom highlights the role of language in cognitive development, literacy learning and socialization. To illustrate the mediational importance of language, the authors focus on the development of one student in particular, Jorge. In this classroom, the teacher organized learning, so that Jorge could use his cultural and linguistic background and experiences as resources to further his own learning of literacy content. In affording Jorge, the opportunity to use his full linguistic “tool kit” (Gutiérrez et al., 1999a), Jorge’s identity was co-constructed as a “knowing contributor” (p. 300). Hence, because this class valued different ways of demonstrating competence, Jorge was able to reframe his identity from that of class clown to knowing contributor.

Gutiérrez et al. (1999b) demonstrate the importance of language as a central mediating tool in fostering the learning and socialization of Martha, a young third grade bilingual Latina in an afterschool club. In an exchange of email letters with El Maga (the cyberspace entity who presides over the computer club), Martha draws upon a repertoire of codes (English and Spanish) and registers to accomplish particular goals. In doing so, Martha is being socialized to the culture of Las Redes while at the same time, “El Maga and other adult participants of Las Redes are socialized to children’s linguistic, sociocultural, and academic experiences” (p. 20). Thus, in this case the socialization process is a bi-directional one which occurs through the use of multiple codes and registers.

Our research on literacy and identity

Our body of research focuses on the intersections of language, literacy, and identity in bilingual communities. We examine these processes within and across educational settings, both formal and informal, including schools, homes, and churches with the intent of bridging these spaces. In our research, we underscore how cultural contexts affect the language and literacy socialization of bilingual Chicana/o/x and Latina/o/x students in that these students engage in social practices that vary across settings and that influence their learning and identity formation in different ways.

Research on literacy and gender of Latino male teachers and their bilingual students

Dr. Lara’s research focuses on the language and literacy practices of Latino/a/x teachers and their students and how their identities are mediated by their reading. In addition, his research interrogates the relationships between discourse and identity. A particular focus is the incorporation of multicultural literature in the classroom. Through the interactions with texts, spaces are created in the literacy events to explore how literacies beyond that of the written text help all actors in the classroom mediate their identities across various contexts such as school, home and community.

Lara and Fránquiz’ study (2015) study of the introduction of bilingual literature with a focus on gender equity centered on an understudied population of Latino male elementary teachers. In a profession that is largely dominated by white women, Latino males make up less than two percent of the teaching population. This study investigated the intersection of gender and ethnicity of Latino male elementary teachers’ literacies and identities in bilingual classrooms through their narratives, experiences and pedagogical practices of read alouds. A qualitative multi-site case study of three Latino male teachers and their classrooms, the study employed qualitative methods and LatCrit as a theoretical lens as a way to depart from the white and Black binary of the studies preceding it. Latino men share common characteristics with the master narrative of male teachers yet; they live in a parallel world where they do not necessarily reap the benefits of hegemonic masculinity and male privilege. The study focused on three teachers who employed interactive read alouds in their bilingual education language arts lessons. Apart from traditional literacy, all three teachers had an explicit focus on critical literacy (Vasquez, 2012). Their teaching approach entailed problem posing methodology (Freire, 1970) where they encouraged reading their worlds as self to text and text to world intertextual connections (Rosenblatt, 1994). The lessons encouraged not just the students to interrogate their identities as they were being shaped by sociocultural agents like family, media, and institutions but also for the teachers to reflect on what male identities they embodied and authored according to the master narratives. The participants in Lara’s study created counternarratives as a way to depart and speak back to the master narratives that position Latino men as patriarchs, and their primarily Latino students at risk or culturally and linguistically deficient. Implications are offered for teachers to leverage the knowledge in the lived experiences of the students and teachers as tools to interrogate meaning making of various texts within differing contexts of power, agency and identity.

Lara’s and Fránquiz’ (2015) study used figured worlds (Holland et al., 1998) as a sociocultural frame to explore the positioning and authoring of three Latino male bilingual education teachers and their students. The findings illustrate how the men were able to read the ways in which they were figured as males by their teacher colleagues and how they encouraged their students to read their worlds. Through the various discourses of administrators and other teachers these men were presented with situations in which they could take on the role of disciplinarian or engage in hypermasculine activities. At these junctures these men were faced with the option to toe the line or enact their agency to author new or alternate identities. These identities were often the topic of conversation in the interactive read alouds where the men authored counter-identities to speak back to the figured worlds of teachers that often position Latino men as machos or overbearing patriarchs. The literacies on display in Lara’s study alternated between traditional and nontraditional types of literacies. Where on the one hand, books were being read, on the surface they explored gender equity topics,

while on a deeper level they encouraged both teacher and students to interrogate their own beliefs and stances for socially constructed gender norms. The home and community became a topic of conversation for children in the study. The students often provided anecdotes of where they had been gendered in a problematic way by their parents or peers. These text to world connections helped the students recognize how their identities are socioculturally constructed. Students illustrated various forms of literacies, for example, they read both books and their worlds. Books offered the possibilities of imagining new worlds and enacting them. As a way to write in their imagined worlds, students often wrote to the characters in the books with advice often in both English and Spanish, showcasing their biliteracy.

Literacy and identity of pentecostal students in church and school

A significant aspect of Dr. Ek's work is the documentation of Chicana/o/x and Latina/o/x experiences and the centering of their voices and perspectives in educational research. Too often, representations of these groups have focused on deficits that contribute to their marginalization. Furthermore, Latina/o/x communities and their practices are often depicted as monolithic; thus, Ek's work highlights the heterogeneity of Latina/o/x identities and the various *Latinidades* that are expressed through literacy in the United States.

Scholars seeking to document the experiences of immigrant Latinos/as/xs have focused on religious institutions as important sites for learning and community building (Baquedano-López, 1998; Ebaugh and Chafetz, 2000; Leon, 1998; Wellmeier, 1998). Religion is often at the center of immigrants' sense of identity and religious institutions serve as key sites for literacy socialization. Yet, studies that focus on Latina/o/x religious literacies are scarce, particularly on non-Catholic Latina/o/xs in the United States. Addressing this gap, Dr. Ek conducted an ethnography that examined the language and literacy practices at a Pentecostal church and how Latina/o/x bilingual children and youth's engagement in Sunday school socialized them to identity. Ek (2005) found that children and youth were socialized to a Latino Christian Pentecostal identity through Bible readings, prayers, songs, and the use of other Pentecostal texts in Spanish. The construction of a unified Latino Christian Pentecostal identity blurred the Mexican, Guatemalan, and Salvadoran national identities of the participants (Ek, 2005). In addition, the Bible was a tool for building a Christian Pentecostal morality and identity and reading the Bible in Spanish was a community building activity as well as a main literacy practice (Ek, 2009a).

A related strand of Ek's Pentecostal study examined the language and literacy socialization to ethnicity, class, and gender of four focal adolescents in the church, their high schools, and homes. In the home, Central American varieties of Spanish constructed Central American identities, and these linguistic varieties of Central Americans marked them as an "other" *Latinidad* within the larger spectrum of a homogenous Latina/o/x community. On the other hand, at school the Central American students defaulted to a Mexican variety of Spanish to blend in with the majority Mexican-origin students (Ek, 2009b). In particular, Ek focused on a longitudinal case study of a second generation Guatemalan American student from the church from the time she was a girl in elementary school through becoming a college student. Her transnational experiences in her parents' native Guatemala served as resources for developing and maintaining her Central American language, literacy, culture, and identity in the United States, particularly in the home setting (Ek, 2009b). Findings also revealed that this young woman negotiated conflictive socializations at home, church, and school by using Guatemalan and Mexican Spanish in strategic ways to construct classed, racialized/ethnicized, and gendered identities (Ek, 2009a). Furthermore, Ek found that the church afforded all children and youth the opportunities to engage in rich language and literacy practices in Spanish that led to their becoming competent church members whereas the Pentecostal youth who were in the English as a Second Language (ESL) program in their high school were placed in classes that were characterized by a lack of effective instruction and a watered-down curriculum (Ek, 2008). Recently, (Ek, 2019) demonstrated how through participation in the Pentecostal church literacy practices, students acquire knowledge and skills in Spanish that overlap with the literacy standards of the core curriculum. These projects revealed how cultural contexts shaped diverse literacy experiences of bilingual students and as a consequence influenced their expression of identities.

Research on digital biliteracy: La Clase Mágica at UTSA (LCM@UTSA)

Building on her research in home and school contexts, Ek was a Principal Investigator (PI) on a collaborative six-year study of La Clase Mágica at UTSA (LCM@UTSA), an after-school technology club modeled after the award winning *La Clase Mágica* at UC San Diego (Vasquez, 2003). Using qualitative and social design experiment methods (Gutiérrez and Vissoughi, 2010), LCM@UTSA was a university-public school partnership located in a predominantly Latina/o/x working-class area (Ek et al., 2010). As a social design experiment (Gutiérrez and Vissoughi, 2010), the goals of LCM@UTSA included: to address the digital divide by providing access to technology to students at an under-resourced school; to engage children in culturally relevant language and literacy activities in Spanish and English; and to apprentice pre-service bilingual teachers to use technology to foster digital literacies. These goals guided our research design.

Our research questions focused on how the bilingual students' participation in literacy activities with the pre-service bilingual teachers fostered their digital biliteracy and identities. Data included participant-observation field notes, videotape logs of interactions between the kids and the pre-service teachers, and children's artifacts including their email exchanges with a cyberwizard, their *fotonovelas*, and other writings. The research team's analysis of digital literacies was informed by scholarship on literacies and new technologies (Knobel and Lankshear, 2007; Lam and Rosario-Ramos, 2009; Lewis and Fabos, 2005). The team also drew from research that documented that emergent bilingual students' engagement with social media, fan fiction, and multimodal literacies

were key to their identity construction, heritage language maintenance, and second language acquisition (Black, 2007; Lam, 2009; Lam and Rosario-Ramos, 2009).

The findings from the LCM@UTSA study revealed the connections among language, literacy, and the construction of various social identities. Ek et al. (2014) demonstrated how codeswitching in Spanish and English in students' emails and digital auto-narratives was significant to their biliteracy development as well as to building relationships with others. Ek and her collaborators also found that the students' stories provided spaces where they represented their own and their family's identities in positive ways. They also found that culturally relevant literacy texts including using multicultural children's books by Latina/o authors and *foto-novelas* were key to the children's engagement and acquisition of digital biliteracy (Ek et al., 2015; Garcia, 2014; Prieto et al., 2016). Through the creation of the *foto-novelas*, children expressed their present and future selves including achievers, star students, fearless test takers, airplane pilots as well as the cultural and social identities of their parents, siblings, and pets (Garcia, 2014; Prieto et al., 2016). Moreover, these findings from LCM@UTSA highlight that effective university-school partnerships must be steeped in the languages, literacy, and cultures of the communities they serve.

Research on home and family literacies of emergent bilingual students

Guerra's (2015) study employed a sociocultural lens to examine the literacy practices of fourth grade bilingual Latina/o/x students in multiple contexts. Specifically, the qualitative study focused on six focal students in home, neighborhood, community, and school settings. Guerra drew on Gee's (2015) notions of literacy as Discourse to analyze the data she collected which included participant-observation field notes, audio and video recordings of family interactions, artifacts, and interviews with parents, the students, and their teacher. She documented the students' Primary Discourses which are learned in the home and were part of their everyday languages and experiences.

She found that the home and family literacy practices were varied, complex, rich, and dynamic in both Spanish and English. These practices depicted a lively amalgam of traditional genres such as: telling *cuentos* (oral stories); teasing and joking with family; Bible reading and quoting; playing table games; and reading of cultural texts, such as *fábulas* (fables). Conversely, media literacy practices mediated by modern technology including computers, TV, cell phones, video games, music, and social media were found to be prevalent across the participants' households. One student participant for example, engaged in intergenerational literacy activities with her grandmother, aunts, mother, and siblings. They played the table game *Lotería* and watched *telenovela* (soap operas) together. At the same time, she played video games with her siblings who taught the youngest ones how to play. The older siblings socialized the younger ones to newer media literacy such as watching YouTube videos that taught them to draw their favorite cartoon characters.

Language brokering and translanguaging characterized the literacy interactions between children, their younger siblings, and the older generation. Guerra found that the fourth-grade focal children used their newly acquired knowledge in two languages to translate for their parents, family, classmates, and teacher. Thus, the emergent bilingual kids were positioned and positioned themselves as experts of language and literacy. One focal student, for example, translated from English to Spanish when he showed his mother and his siblings how to use a laptop while explaining that the family needed to save money to get internet services. The focal children constructed bilingual identities that reflected how savvy they were in using technology, but at the same time knowledgeable about the traditional literacies passed down through generations. In addition, emergent bilingual children's translanguaging in literacy activities reflected multiple varieties of Spanish including Spanish of the Southwest, Mexican Spanish, Honduran Spanish, and Colombian Spanish which marked their identities.

Research on language and literacy of Latina/o/x lowriders

Sánchez's (2017) research examines the language and literacy practices in a Latina/o/x lowriding community. This study provides a focus on the social practices within the lowrider community to uncover the multiple and diverse literacies usually hidden to outsiders of these non-mainstream communities. This study served to advocate for alternative discourses that have historically stigmatized the lowrider community by reconceptualizing literacy as a challenge to traditional definitions of literacy commonly privileged in learning institutions. In addition, this study highlights the intersection of language and literacy in lowrider culture that serves to construct a lowrider identity.

The focal participants in this study were members of several local lowrider car clubs based in a South Texas city. The data for this study was gathered during a period of about five years and included participant observations, fieldnotes, interviews, and audio and video recordings collected during car shows, car club meetings, and social gatherings organized around cruises and hop-offs. Written texts taken from flyers and social platforms were also analyzed and provided evidence of the relationship between language and identity. In addition, the analysis of the data collected in this study revealed how lowrider identities are multiple and constructed through lowrider language, values, cars, bikes, and dress. This "identity kit" (Gee, 2015) included a way of acting, talking, and writing that communicated to others their particular role as members of a lowrider community.

Central to this study and to the construction of a lowrider identity was the use of linguistic variations in the Spanish oral and written language to mark a lowrider identity. The analysis of the use of lexicons and vernaculars specific to members of the lowriding community emphasized how language is used to express a lowrider identity. Online communities such as Facebook provided a space where a specific style of language, or vernacular, became a linguistic marker of a lowrider identity. Terminology describing social practices of this community, such as talking about a hop-off, represented multiple meanings and the understanding of these

meanings was contingent upon membership in this lowriding community. Another aspect of language that constructed a lowrider identity was the use of Spanglish or Chicano slang, also known as Chicano or Mexican slang to this particular community of Latino/a/x lowriders. This particular language style, characterized by unconventional grammatical rules and heavily influenced by Spanish, is created within the space of Latina/o/x lowriding. These language variations used in both oral and written texts serve as linguistic symbols of what it means to be a Latina/o/x lowrider in this South Texas city.

Conclusion

We began with situating the sociocultural theories of learning and development as essential to understanding how literacy and identity are shaped across the context of home, school, church, and social life. It is important to recognize how literacies extend beyond the basic notion of reading words in a textbook to include the social and cultural contexts where sensemaking of any text, in or out of school, is informed by social institutions and power relationships and how the reader is situated based on beliefs, customs, identities and experiences. While identities are socioculturally created and prescribed to some degree, the ability of individuals to engage in critical readings of their worlds aids in the recognition of agency to author their own new or separate identities.

We have provided a non-exhaustive survey of literacy and identity across multiple contexts. When considering other types of literacies and pedagogies for culturally and linguistically diverse students, readers may explore concepts such as: funds of knowledge (González et al., 2005), funds of identity (Esteban-Guitart and Moll, 2014), culturally relevant pedagogy (Ladson-Billings, 2009), culturally sustaining pedagogies (Paris and Alim, 2014) and community cultural wealth (Yosso, 2005).

The encyclopedia authors' body of research is significant because the studies presented here examine and create spaces that foster the multiliteracies and multilingualism of culturally and linguistically diverse individuals. Too often, the multiple literacies (Comber and Simpson, 2001) of students from culturally and linguistically diverse groups—particularly bilingual, immigrant, students of color—go unrecognized, are devalued, or are untapped by those in power (Morrell, 2004). Hence, our own scholarship illuminates the rich, complex, literacies of working-class students and their families that open up spaces for them to author themselves in powerful ways. Moreover, the knowledge and skills gained from literacy practices in out-of-school contexts can be leveraged in the classroom. Thus, we call for teachers of working-class bilingual students to foster opportunities for students to bring their home identities and Discourses (Gee, 2015) into the classroom. Our research can serve as a resource for teachers, administrators, and policymakers to become more aware of the everyday negotiations of the development of literacies and identities in our classrooms, churches, student homes, and everyday activities. Finally, we hope to shed light on the beauty of our participants' knowledge and practices as a way to encourage more equitable educational and institutional approaches for culturally and linguistically diverse students, their families, and communities.

References

- Baquedano-López, P., 1997. Creating social identities through "doctrina" narratives. *Issues Appl. Ling.* 8 (1), 27–45.
- Baquedano-López, P., 1998. *Language Socialization of Mexican Children in a Los Angeles Catholic Parish*. University of California, Los Angeles, CA.
- Baquedano-López, P., 2004. Traversing the center: the politics of language use in a Catholic religious education program for immigrant Mexican children. *Anthropol. Educ. Q.* 35 (2), 212–232.
- Barton, D., Hamilton, M., 2012. *Local Literacies Reading and Writing in One Community*. Routledge, London.
- Bauer, E.B., Gort, M., 2012. *Early Biliteracy Development*. Routledge, New York, NY.
- Black, R.W., 2007. Digital design: English language learners. *New Literacies Sampler* 29, 115.
- Byrd, A., Hayes, J., Turnipseed, N., 2021. Against autonomous literacies: extending the work of Brian V. Street: introduction to the special issue. *Lit. Compos. Stud.* 8 (2), V–XXI.
- Capps, L., 1999. Constructing the irrational woman: narrative interaction and agoraphobic identity. In: Bucholtz, M., Liang, A.C., Sutton, L.A. (Eds.), *Reinventing Identities: The Gendered Self in Discourse*. Oxford Press, New York, NY, pp. 83–100.
- Cazden, C., Cope, B., Fairclough, N., Gee, J., Kalantzis, M., Kress, G., Luke, A., Luke, C., Michaels, S., Nakata, M., 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Comber, B., Simpson, A., 2001. *Negotiating Critical Literacies in Classrooms*. Routledge, Mahwah.
- Cun, A., 2021. Funds of knowledge: early learning in Burmese families. *Early Child. Educ. J.* 49 (4), 711–723.
- de la Luz Reyes, M., Halcón, J.J., 2001. *The Best for Our Children: Critical Perspectives on Literacy for Latino Students*. Teachers College Press, New York, NY.
- De la Piedra, M.T., 2010. Adolescent worlds and literacy practices on the United States-Mexico border. *J. Adolesc. Adult Literacy* 53 (7), 575–584.
- Ebaugh, H.R., Chafetz, J.S., 2000. *Religion and the New Immigrants: Continuities and Adaptations in Immigrant Congregations*. Rowman & Littlefield, Lanham, MD.
- Ek, L.D., 2005. Staying on god's path: socializing Latino immigrant youth to a Christian Pentecostal identity in Southern California. In: Zentella, A.C. (Ed.), *Building on Strength: Language and Literacy in Latino Families and Communities*. Teachers College Press and California Association for Bilingual Education, New York, NY, pp. 77–92.
- Ek, L.D., 2008. Language and literacy in the Pentecostal church and the public high school: a case study of a Mexican ESL student. *High Sch. J.* 92 (2), 1–13.
- Ek, L.D., 2009a. "It's different lives": a Guatemalan American adolescent's construction of ethnic and gender identities across educational contexts. *Anthropol. Educ. Q.* 40 (4), 405–420.
- Ek, L.D., 2009b. "Alla en Guatemala": transnationalism, language, and identity of a Pentecostal Guatemalan-American young woman. *High Sch. J.* 92 (4), 67–81.
- Ek, L.D., 2019. Linking church and school: Language and literacy practices of bilingual Latinx Pentecostal youth. In: García-Sánchez, I.M., Orellana, M.F. (Eds.), *Language and cultural practices in communities and schools: Bridging learning for students from non-dominant groups*. Routledge, New York, NY, pp. 107–123.
- Ek, L.D., García, A.S., Garza, A., 2014. Latino children: constructing identities, voices, linguistic, and cultural understandings. In: Flores, B.B., Vasquez, O.A., Clark, E.R. (Eds.), *Generating Transworld Pedagogy: Reimagining La Clase Mágica*. Lexington Books, New York, NY, pp. 129–142.
- Ek, L.D., Machado-Casas, M., Sanchez, P., Alanis, I., 2010. Crossing cultural borders: La Clase Mágica as a university–school partnership. *J. Sch. Leader.* 20 (6), 820–849.

- Ek, L.D., Sánchez, S.N., Guerra, M.J., 2015. Cultural Multiliteracies: Integrating Technology with Latino Children's Literature, Multicultural Literature for Latino bilingual Children: Their Words, Their Worlds. Rowman Littlefield Publishing Group, Lanham, MD, pp. 207–222.
- Esteban-Guitart, M., 2016. Funds of Identity: Connecting Meaningful Learning Experiences in and Out of School. Cambridge University Press, New York, NY.
- Esteban-Guitart, M., Moll, L.C., 2014. Funds of identity: a new concept based on the funds of knowledge approach. *Cult. Psychol.* 20 (1), 31–48.
- Fader, A., 2001. Literacy, bilingualism, and gender in a Hasidic community. *Ling. Educ.* 12 (3), 261–283.
- Fránquiz, M.E., de La Luz Reyes, M., 1998. Creating inclusive learning communities through English language arts: from "chancas" to "canicas". *Lang. Arts* 75 (3), 211–220.
- Fránquiz, M.E., Ortiz, A.A., Lara, G.P., 2021. Enlaces prometedores: funds of knowledge and funds of identity. *Biling. Res. J.* 44 (4), 405–408.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York, NY.
- García, A.S., 2014. "La Fotonovela Digital": A Sociocultural Tool for Biliteracy, Digital Literacy, and Visual Literacy Learning of Latina/o Bilingual Children at an Afterschool Club. The University of Texas at San Antonio.
- García, E.E., 2008. *Teaching and Learning in Two Languages: Bilingualism & Schooling in the United States*. Teachers College Press, New York, NY.
- Gee, J.P., 2015. *Social Linguistics and Literacies: Ideology in Discourses*. Routledge, New York, NY.
- González, N.E., Moll, L.C., Amanti, C.E., 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Gorski, P.C., 2006. Complicity with conservatism: the de-politicizing of multicultural and intercultural education. *Intercult. Educ.* 17 (2), 163–177.
- Gumperz, J.J., Cook-Gumperz, J., 1982. Introduction: language and the communication of social identity. In: Gumperz, J.J. (Ed.), *Language and Social Identity*. Cambridge University Press, Cambridge, pp. 1–21.
- Guerra, M.J., 2015. *Usando Nuestros Recursos: Biliteracy Practices of Young Latino Bilingual Learners inside and outside School*. The University of Texas at San Antonio.
- Gutiérrez, K.D., Baquedano-Lopez, P., Turner, M.G., 1997. Putting language back into language arts: when the radical middle meets the third space. *Language Arts* 74 (5), 368–378.
- Gutiérrez, K.D., Rymes, B., Larson, J., 1995. Script, counterscript, and underlife in the classroom: James Brown versus Brown v. Board of education. *Harv. Educ. Rev.* 65 (3), 445–472.
- Gutiérrez, K.D., Baquedano-López, P., Alvarez, H.H., Chiu, M.M., 1999a. Building a culture of collaboration through hybrid language practices. *Theory Pract.* 38 (2), 87–93.
- Gutiérrez, K.D., Baquedano-López, P., Tejeda, C., 1999b. Rethinking diversity: hybridity and hybrid language practices in the third space. *Mind Cult. Activ.* 6 (4), 286–303.
- Gutiérrez, K.D., Stone, L.D., 1997. A cultural-historical view of learning and learning disabilities: participating in a community of learners. *Learn. Disabil. Res. Pract.* 12 (2), 123–131.
- Gutiérrez, K.D., Vossoughi, S., 2010. Lifting off the ground to return anew: mediated praxis, transformative learning, and social design experiments. *J. Teach. Educ.* 61 (1–2), 100–117.
- Hatch, J.A., 2002. *Doing Qualitative Research in Education Settings*. SUNY Press, Albany, NY.
- Heath, S.B., 1983. *Ways with Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, New York, NY.
- Holland, D., Lachiotte Jr., W., Skinner, D., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge.
- Hsin, C.T., Yu, C.Y., 2021. Literacy and identity development of indigenous Rukai children. *J. Literacy Res.* 53 (3), 313–335.
- Hymes, D.H., 1978. *What Is Ethnography?* Southwest Educational Development Laboratory, Austin, TX.
- Ibrahim, N., 2016. Enacting identities: children's narratives on person, place and experience in fixed and hybrid spaces. *Educ. Inq.* 7 (1), 69–91.
- Jiménez, R.T., 2003. Literacy and Latino students in the United States: some considerations, questions, and new directions. *Read. Res. Q.* 38 (1), 122–128.
- Johnson, A.S., 2010. The Jones family's culture of literacy. *Read. Teach.* 64 (1), 33–44.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, New York, NY.
- Knobel, M., Lankshear, C., 2007. *A New Literacies Sampler*. Peter Lang, New York, NY.
- Ladson-Billings, G., 2009. *The Dreamkeepers: Successful Teachers of African American Children*. John Wiley & Sons.
- Lam, W.S.E., 2009. Literacy and learning across transnational online spaces. *E-learn. Digit. Med.* 6 (4), 303–324.
- Lam, W.S.E., Rosario-Ramos, E., 2009. Multilingual literacies in transnational digitally mediated contexts: an exploratory study of immigrant teens in the United States. *Lang. Educ.* 23 (2), 171–190.
- Lara, G.P., 2015. *The Complexity of Latino Male Bilingual Teachers' Experiences*. The University of Texas at Austin.
- Lara, G.P., Fránquiz, M.E., 2015. Latino bilingual teachers: negotiating the figured world of masculinity. *Biling. Res. J.* 38 (2), 207–227.
- Lara, G.P., Leijja, M.G., 2014. Discussing gender roles and equality by reading "Max: the stubborn little wolf". *Soc. Stud. Young Learn.* 27 (2), 22–25.
- Leijja, M.G., 2017. *Interactive Read Alouds: Developing Literacy in Bilingual Second Grade Classrooms*. The University of Texas at Austin.
- Leijja, M.G., Fránquiz, M.E., 2021. Building bridges between school and home: teacher, parents, and students examining Latinx immigrant experiences. In: Onchwari, G., Keengwe, J. (Eds.), *Bridging Family-Teacher Relationships for ELL and Immigrant Students*. IGI Global, Hershey, PA, pp. 100–121.
- León, L., 1998. Congregation as border space. In: Warner, S., Wittner, J.G. (Eds.), *Gatherings in Diaspora: Religious Communities and the New Immigration*. Temple University Press, Philadelphia, PA, pp. 163–198.
- Lewis, C., Fabos, B., 2005. Instant messaging, literacies, and social identities. *Read. Res. Q.* 40 (4), 470–501.
- Luke, A., 1994. *The Social Construction of Literacy in the Primary School*. Macmillan Education Australia, Melbourne, VIC.
- Martínez-Roldán, C.M., 2003. Building worlds and identities: a case study of the role of narratives in bilingual literature discussions. *Res. Teach. Engl.* 37 (4), 491–526.
- McMillon, G.T., Edwards, P.A., 2000. Why does Joshua "hate" school... But love Sunday school? *Lang. Arts* 78 (2), 111–120.
- Mercado, C.I., 2005. Seeing what's there: language and literacy funds of knowledge in New York Puerto Rican homes. In: Zentella, A.C. (Ed.), *Building on Strength: Language and Literacy in Latino Families and Communities*. Teachers College Press, New York, NY, pp. 134–147.
- Moje, E., 2000. "To be part of the story": the literacy practices of gangsta adolescents. *Teachers College Record* 102 (3), 651–690.
- Moll, L., 2005. Reflections and possibilities. In: Gonzalez, N., Moll, L., Amanti, C. (Eds.), *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 275–288.
- Moll, L.C., Greenberg, J.B., 1992. Creating zones of possibilities: combining social contexts for instruction. In: Moll, L.C. (Ed.), *Vygotsky and Education: Instructional Implications and Applications of Sociocultural Psychology*. Cambridge University Press, New York, NY, pp. 319–348.
- Moll, L.C., Sáez, R., Dworin, J., 2001. Exploring biliteracy: two student case examples of writing as a social practice. *Elem. Sch. J.* 101 (4), 435–449.
- Morrell, E., 2004. *Linking Literacy and Popular Culture: Finding Connections for Lifelong Learning*. Christopher-Gordon Publishers, Norwood, MA.
- Mui, S., Anderson, J., 2008. At home with the Johars: another look at family literacy. *Read. Teach.* 62 (3), 234–243.
- New London Group, 2000. A pedagogy of multiliteracies. In: Cope, B., Kalantzis, M. (Eds.), *Multiliteracies: Literacy Learning and the Design of Social Future*. Routledge, London, pp. 9–37.
- Ochs, E., 1995. Introduction. In: Schieffelin, B.B., Ochs, E. (Eds.), *Language Socialization across Cultures*. Cambridge University Press, New York, NY, pp. 1–16.
- Ochs, E., Schieffelin, B.B., 1995. The impact of language socialization on grammatical development. In: Fletcher, P., MacWhinney, B. (Eds.), *The Handbook of Child Language*. Blackwell, Cambridge, pp. 73–94.
- Oikonomidou, E., Karam, F.J., 2020. Funds of identity and language development: the case of a Syrian refugee-background child. *Int. J. Early Years Educ.* 28 (2), 122–135.
- Orellana, M.F., 1994. Appropriating the voice of the superheroes: three preschoolers' bilingual language uses in play. *Early Child. Res. Q.* 9 (2), 171–193.
- Paris, D., Alim, H.S., 2014. What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harv. Educ. Rev.* 84 (1), 85–100.

- Peñalva, S.L., Coggin, L.S., Medina, C.L., 2014. Examining transcultural spiritual literacies among Latino children through artifactual mediations. *Diaspora, Indig. Minority Educ.* 8 (2), 92–107.
- Pérez, B., 2004. Literacy, diversity, and programmatic responses. In: McCarty, T.L., Pérez, B. (Eds.), *Sociocultural Contexts of Language and Literacy*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, pp. 3–24.
- Perry, K.H., 2012. What is literacy? – a critical overview of sociocultural perspectives. *J. Lang. Literacy Educ.* 8 (1), 50–71.
- Poveda, D., Palomares-Valera, M., Cano, A., 2006. Literacy mediations and mediators in the escuela dominical of a Gitano evangelist church. *Ethnogr. Educ.* 1 (2), 265–283.
- Prieto, L., Arreguin-Anderson, M.G., Yuen, T.T., Ek, L.D., Sánchez, P., Machado-Casas, M., García, A., 2016. Four cases of a sociocultural approach to mobile learning in La Clase Mágica, an afterschool technology club. *Interact. Learn. Environ.* 24 (2), 345–356.
- Rackley, E.D., 2021. “Blessings and friends and knowledge”: environmental motivations for religious literacy. *Relig. Educ.* 116 (2), 101–115.
- Reyes, I., 2006. Exploring connections between emergent biliteracy and bilingualism. *J. Early Child. Lit.* 6 (3), 267–292.
- Rivera, L., 2004. Chapter seven: literacy for change: Latina adult learners and popular education. *Counterpoints* 253, 133–155.
- Rosenblatt, L.M., 1994. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Southern Illinois University Press, Carbondale, IL.
- Rueda, R., Moll, L.C., 1994. A sociocultural perspective on motivation. In: O’Neil, H.F., Drillings, M. (Eds.), *Motivation: Theory and Research*. Routledge, New York, NY, pp. 117–137.
- Sánchez, P., Machado-Casas, M., 2009. At the intersection of transnationalism, Latina/o immigrants, and education. *High Sch. J.* 92 (4), 3–15.
- Sanchez, S.N., 2017. *Rooster Tails, Ranflas, and Rags: The Language and Literacy Practices of Latin@ Lowriders*. The University of Texas at San Antonio, San Antonio, TX.
- Schieffelin, B.B., Ochs, E., 1986. Language socialization. *Annu. Rev. Anthropol.* 15 (1), 163–191.
- Scribner, S., Cole, M., 1981. Unpackaging literacy. In: Whiteman, M.F. (Ed.), *Writing: The Nature, Development, and Teaching of Written Communication*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 71–87.
- Skerrett, A., 2016. “Closer to God”: following religion across the lifeworlds of an urban youth. *Urban Educ.* 51 (8), 964–990.
- Skerrett, A., 2017. The role of language in religious identity making: a case of a Caribbean-Chinese youth. *Lit. Res.* 66 (1), 325–340.
- Street, B.V., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge, MA.
- Street, B.V., 1998. New literacies in theory and practice: what are the implications for language in education? *Ling. Educ.* 10 (1), 1–24.
- Street, B.V., 2003. What’s “new” in new literacy studies? Critical approaches to literacy in theory and practice. *Curr. Issues Comp. Educ.* 5 (2), 77–91.
- Street, B.V., 2014. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Routledge, Hoboken, NJ.
- Trueba, H.T., 1988. Culturally based explanations of minority students’ academic achievement. *Anthropol. Educ. Q.* 19 (3), 270–287.
- Tusting, K., 2015. “I am a peacemaker” writing as a space for recontextualizing children’s identity in a Catholic first communion preparation course. *Writ. Commun.* 32 (3), 227–253.
- Vasquez, O.A., 2003. *La Clase Mágica Imagining Optimal Possibilities in a Bilingual Community of Learners*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Vasquez, V.M., 2012. Creating opportunities for critical literacy with young children. In: Evans, J. (Ed.), *Literacy Moves on: Using Popular Culture, New Technologies and Critical Literacy in the Primary Classroom*. David Fulton Publishers, London, pp. 78–98.
- Vasquez, V.M., 2014. *Negotiating Critical Literacies with Young Children*. Routledge, New York, NY.
- Vélez-Ibañez, C.G., 1983. *Bonds of Mutual Trust: The Cultural Systems of Rotating Credit Associations Among Urban Mexicans and Chicanos*. Rutgers University Press, New Brunswick, NJ.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, London.
- Wellmeier, N.J., 1998. *Ritual, Identity, and the Mayan Diaspora*. Garland Publishers, New York, NY.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.

Religious influences on the growth of literacy practice

Loukia K. Sarroub^a and Cassandra Schroeder^b, ^a Department of Teaching, Learning & Teacher Education, College of Education and Human Sciences, School of Global Integrative Studies, University of Nebraska-Lincoln, Lincoln, NE, United States; and

^b Department of Teaching, Learning & Teacher Education, College of Education and Human Sciences, University of Nebraska-Lincoln, Lincoln, NE, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	209
Language and literacy learning	209
Identity construction via literacy, language, and religious practices	210
Power via literacy, language, and religious practices	212
Practices in the classroom that engage religious literacies pluralistically	213
References	215
Further reading	216

Introduction

Religious influences on the growth of literacy practices are well documented and span more than a century of research ranging from disciplines such as social and cultural anthropology to sociology to language and literacy studies in education. Intellectuals known across disciplines such as Benedict Anderson, Lila Abu-Lughod, Pierre Bourdieu, Jonathan Boyarin, Clifford Geertz, Michaela de Leonardo, Shirley Brice Heath, Alan Peshkin, Claude Lévi Strauss, and Brian Street broke new ground in the 20th century in connecting literacy to religious literacies. In recent years, the work of contemporary language education scholars such as Huamei Han (2018) as well as English education and literacy scholars such as Juzwik et al. (2022) have illuminated the nexus of education, literacy, and religious expression in an intellectual and academic milieu that has typically espoused secularly-minded scholarship. This chapter explores contemporary research in the 21st century that suggests that researchers' questions about religious influences on literacy practices has shifted over time and is also indicative of scholars' own changing reflective stances toward the impact of religious literacies in education.

Language and literacy learning

Organized religion creates cultures around rituals that are maintained and passed from generation to generation. As ancient religions have evolved into the 21st century, many have seen a shift to more modernized practices. The current scholarship on language and literacy practices within religious contexts focuses on religion as a way to foster social communities by linking the ancient rites of the religion to everyday life.

Several studies have examined the literacy practices within religious groups, especially concerning the three major belief systems: Christianity, Judaism, and Islam. All three religions require literacy practices to be fully engaged in the worship services. The use of prayers, recitation, and singing are all examples of ways religious communities engage in literacy practices. The focus in these kinds of practices is not comprehension, but rather, rote memorization (Baquedano-López, 2008; Moore, 2015; Rosowsky, 2013). Many young people attend Qur'anic and Hebrew schools or a Christian Sunday School to engage in the memorization of ancient religious texts like the Qur'an, Torah, and Bible. The comprehension of text is secondary to the fluency and oral expression of reading/speaking the text because the meaning is orally communicated by religious leaders.

Other studies suggest that students who engage in literacy practices are exposed to a wider variety of genres and forms of language that are not present in public school systems today (Baquedano-López, 2008; Papen, 2017; Rosowsky, 2013). The understanding of students' religious literacy practices is important for educators to know because many students are engaged in a wider variety and more difficult texts than what is available in schools (Rackley, 2014; Papen, 2017). The comprehension of ancient religious texts is secondary to their memorization because its purpose is to allow for more engagement in the worship service and build a sense of community among church members.

Research on both immigrant and native religious groups shows a strong relationship between religious rituals and community. Learning to perform and participate in the rites of worship creates a community that many immigrants seek when adjusting to a new country and culture. Immigrants are especially susceptible to the pulls of religion as a community. In Han's (2011, 2014) pioneering research on Chinese immigrant conversion to Christianity, she found that the participants were originally seeking a community of English speakers in which to practice communicating in English. The participants of the study found the local community college English classes too formal but enjoyed the informality of conversational English through Bible study. While they were not interested in converting to Christianity, the couple in Han's (2011, 2014) study became hyper-involved in the church because of the inclusion

of their own cultural practices within the religious customs. By code-switching each member of the community was given a voice and felt empowered to use it (Han, 2011, 2014).

The use of religious schools as a counterbalance to the public school curriculum is another facet of religious literacy research that has emphasized community ways of knowing over individual comprehension of the text. In Rackley's (2014) study on literacy practices of Latter-Day Saints and Methodist youth, he found that even within a commonly known religious Christian sects, there can be cultural differences. Through his analysis of religious texts, he found that Methodist youth engaged in a *culture of interpretation* that created meaning through small group discussion. On the other hand, the Latter Day Saints youth engaged in a *culture of listening*, where the interpretation was given to them by elders in the church. Understanding the teaching and interpretation of religious texts is informative for developing a more nuanced understanding of the connections to public school pedagogy that bridges the out-of-school literacy practices with the requirements of academic language (cf. Sarroub, 2001, 2002, 2005, 2007a, 2007b, 2009a, 2009b).

Although religious practices are entrenched in ancient rituals, a recent trend in the research is showing a change in the ways young students are learning Ancient Arabic and Hebrew. Children's attendance at Arabic and Hebrew schools is an important cultural aspect of both religions because the ancient language is such an integral part of worship services. In Moore's (2015) study on West African Muslim families, she discovered that more families were choosing the French-speaking public schools over traditional Islamic day schools. Instead, their children were learning ancient Arabic by attending night classes. This approach to religious language learning is described as a counterbalance to the privileged French-speaking public school. This new way of learning has also shown an uptick in girls' attendance compared to boys in Muslim schools (Moore, 2015). In the same vein, Sarroub (2005) showed that Yemenis and Yemeni Americans in Michigan chose to send their children to public school during the week and to "Arabic School" at the local mosque on Saturdays and Sundays.

In 2019, Rosowsky revisited her primary field site and research about the sacred language practices of Mosque schools in northern England. In her original study (2008), she found that most students and families of the Mosque schools were first- or second-generation immigrants from the same country with the same home language, primarily Urdu. Twenty years later, Rosowsky (2019) saw a shift in Qur'anic study in English, the common language of the more diverse population of the schools. With an influx of immigrants from many parts of the world, and the establishment of a stronger multi-generational community, the use of the English language is an imperative tool for interpersonal communication within the Islamic school classroom (Rosowsky, 2019).

Literacy practices among the three major religions- Christianity, Judaism, and Islam-are important to maintaining the ritual practices of worship. The scholarship in this area demonstrates a prevalent emphasis on the decoding and fluency of religious print texts over comprehension and interpretation. Students of sacred languages are not asked to develop their own understandings of the text, but rather, to recite them as members of a faith-based community. There has also been a shift to maintaining institutional religious memberships and bridging the sociocultural practices of youth with the literacy practices of their religions.

Identity construction via literacy, language, and religious practices

Religion, or lack thereof, is an important aspect of individual identity. While church and state are fiercely separated in the United States, especially regarding public schools, religious beliefs often undergird students' classroom discussions. This section explores recent scholarship about the links between individual identity, religion, and public institutions, and it addresses how religious literacy practices are being used to build community among immigrants and US-born populations. This section also highlights research about how young people utilize religion to negotiate religious, social, and cultural spaces.

Bigelow (2008), for example, studied the intersection of immigrant identity and religious practices within public institutions such schools. As immigrants in a new dominant culture, Somali youth struggled with the racialization of being seen as Black and not African, while dressing in a hip-hop style and adopting the African American vernacular English along with the standard variety to fit into the social aspects of school. Despite this conformity, Somali teens wrote, "Somali," "Muslim," or "Other" on forms for school that indicate racial categories.

Choi and Tinker Sachs's (2017) case study focused on how multilingual students negotiated their identities when reading multicultural literature. Four male participants from India, Uzbekistan, and Korea participated in Saturday morning book clubs that met for five months. Two of the boys identified as religious- Muslim and Hindu- and the other two identified as atheists with no religion. The young men used their religious knowledge gained from their home lives to create nuanced understandings of the multiethnic literature they were reading. The researchers documented how differing perspectives about religious beliefs became a source of tension among the boys as they naturally drew on their religious identities to make meaning from the text. These studies are important in understanding of religious literacy because they demonstrate the ways multilingual youth use religion to interpret texts secular literature. It also shows the importance of teachers allowing various religious perspectives into the classroom to foster dialogue and understanding among participants.

Research such as Han's (2007) ethnographic case study provides a lens into a different educational setting for adults. She followed one Chinese immigrant couple, Grace and Timothy, as they navigated their new life in Toronto, Canada. Han outlines the couple's settlement in Toronto as being traumatic for the newlyweds. Timothy and Grace struggled to find meaningful work and make it financially in Canada. They found a community in the evangelical Christian group that they joined and used expressions such as "God's plan" to make sense of their experience and find calmness. At the heart of this study is the gender and language shifts of Grace and Timothy that occurred through their involvement in the evangelical Christian group. Han (2007) found that Grace took on a dominant and outspoken role in private, often keeping a steady job to keep the family afloat. However, Timothy

took on a more dominant role in group conversations, especially in the church's English language classes. Han's findings suggest that immigrant women tend to be particularly vulnerable to gender stereotypes despite the efforts at gender equality in Western liberal democracies.

According to Han (2007), evangelical churches in Canada utilize English language classes to build community and connection among immigrants, and by the same token, there is the expectation that they will convert to Christianity despite any desire to do so before encountering the group. However, Han's (2011) research explored the participants' sense of "otherness" via normalized racial and structural discrimination as a reason for their interest in finding community through Chinese nationalism, and thus, saving their country through Christian evangelism. Despite the hardships encountered during their settlement in Canada, Chinese immigrants were able to find a positive identity within the evangelical Christian community.

In her case study (Reyes, 2009) followed one high school Latina girl, Zulmy, for 10 months at school and in her home to explore how she negotiates her religious identity in a public space, such as school. The main source of data came from Zulmy's science scrapbook which she crafted for the biology club, and, which, incidentally, had very little to do with science, but rather, it included reflections on her religion and her relationship with the material world. Reyes (2009) concluded that Zulmy's religious identity is complex and sometimes contradictory. For example, one entry was an inspirational poem about "God" and the next page was a visual that sent the message: "I love Mickey Mouse, a symbol of American consumerism." Through this scrapbook, Zulmy made sense of the different facets of her identity by integrating the pieces of her religion, culture, language, and relationships into one artifact. One could hypothesize that the scrapbook could be what Sara Ahmed (2010) calls a "happy object," in that its significance drew on contemporary media and religious texts that helped Zulmy make sense of her world.

Recent scholarship has also focused on Christian religious identity of US-born citizens in public schools. Research in this sub-field, which has roots in the pioneering work of Alan Peshkin (1986) in an evangelical Christian community, calls attention to the importance of studying this population because of the emphasis placed on their beliefs and values in US politics. For example, Juzwik's (2014) critical review of the Biblicism tradition in American Christianity argued that Biblicism, a belief in the authoritativeness of the Bible is a literacy event because it requires the reading and interpretation of complex religious texts. Juzwik suggested that researchers should study this population because it includes the largest religious group in the United States, one that develops strong identities as Biblicist Christians and focuses heavily on evangelizing to transform lives. The emphasis on proselytizing and conversion is problematic in religiously pluralistic settings such as public schools, and Juzwik noted that research further research offers a deeper understanding of their literacy practices for evangelization. With a more optimistic view than that of Peshkin (1986), Juzwik argued that gaining insight into evangelical Christian religious literacy practices, dialogue across ethical differences can be promoted to gain a broader view of the world.

In 2015, Juzwik and McKenzie's qualitative study included both a teacher and student who self-identified as evangelical Christians. Sam, the teacher, conducted a "This I Believe" essay writing unit wherein he invited students to explore their own beliefs and morals while hearing from others who may be ethically different. Charlie, an evangelical Christian student in Sam's class, saw the unit as an opportunity to honor "God" and "His" word through his own literacy practices. The findings from this study indicated that there are differences even among members of the evangelical Christian belief. For example, Sam, the teacher, was older and more accustomed to discussion that included different viewpoints and ideologies. Charlie, the student, on the other hand, has not learned to participate in or value cross-faith dialogue. Ultimately, Juzwik and McKenzie (2015) argue that educators ought to take the religious roots of students into account and develop avenues in the classroom to engage them in dialogue.

However, there has been a limited number of studies that explore the interactions of youth religious beliefs, literacy, and motivation. Rackley (2016), in turn, contributed to the development of religio-cultural theories in literacy studies. In his qualitative study of Latter-day Saint and Methodist youths' motivation for reading complex, religious texts, the findings revealed that the young people in both religious groups were motivated to read and comprehend complex religious texts because they wanted to apply them to their lives, endure challenges, find comfort, and connect with "God."

Using a multiliteracies framework, Skerrett (2014) examined how religious literacies are used to make sense of the secular literature that students read in class. Using a qualitative case study design, data was collected over one year in Ms. Campbell's ninth grade classroom. The students in the study were members of a diverse high school in the southwestern United States, and many of them self-identified as Mexican or Mexican American. The findings suggest that, with the teacher's support, students were able to draw on their religious literacies to analyze and evaluate the secular literature read in class and within their academic writing. Skerrett also found that tensions arose during class discussions that involved religious identity, but there was a strong sense of community based on shared values and empathy for others that helped students navigate through these tensions with their teacher, who also understood their more nuanced religious viewpoints.

In another study about evangelical Christian students and their interactions with a secular literacy curriculum in the United States, Weyand and Juzwik (2019) pointed out that literacy education in midwestern public schools was developed according to evangelical Protestant print literate norms, and therefore, evangelical Christian literacy practices have been excluded. This study focused on an evangelical Christian young man's responses to a community-based curriculum that required him to intern with a traveling college campus preacher.

Jeremy, the student participant, was given the opportunity to publicly identify as a Christian and justify his beliefs using Biblical scripture. This dialogue was evidence of his willingness to share his faith with others by inviting discussion, and thus, scrutiny of his beliefs. In turn, it demonstrated a readiness to understand others and their views. This was something that Jeremy was not open to doing prior to the internship. Like Charlie (Juzwik and McKenzie, 2015), Jeremy is another example of students adapting curricular goals for their own religious exploration and identity development. Further research is needed to fully examine young people's

religious-related decisions and literacy practices in school spaces, especially those of historically minoritized groups who face discrimination (cf. [Mir and Sarroub, 2019](#); [Sarroub, 2001, 2002, 2005](#)).

[Wynter-Hoyte and Boutte \(2018\)](#), for example, contributed a qualitative case study that focuses on Melissa, an African American girl in the third grade as she navigated the literacy worlds of church and school. The authors implemented Cultural Historical Activity Theory (CHAT) as a framework as well as a post-structural identity theoretical perspective that instantiates identity as non-static, pluralistic, and adaptable to one's environment. These frameworks illuminated interaction and learning differences across church and school settings. At church, Melissa's literacy practices were predominantly communal, i.e., praying for one another and sharing testimonies. The church community also embraced African American Vernacular English (AAVE), such as the call and response speech pattern within the services. Other examples of communal practices included small group instruction for the children wherein they engaged in a close reading of the Bible and role-playing biblical stories. Melissa's classroom literacy practices, modeled by her African American teacher, focused on individual learning. There were few peer interactions and the lessons focused on standardized testing skills. Melissa navigated both institutional literacy practices seamlessly and was able to participate successfully in both environments. Wynter-Hoyte and Boutte pointed out that ways of participating in literacy activities in the church setting could be a resource within a school to activate communal and multiple literacies to benefit all students and help develop stronger relationships between students and teachers.

Power via literacy, language, and religious practices

While religious beliefs are an important aspect of many individual's identities and sense of belonging, it is institutions that systematically wield power an often for their own gains. This section explores the relationship between religion and power and the impact it has on people's lives as they implement their political and social agendas. [Baquedano-López \(2004\)](#), [Hsiao et al. \(2018\)](#), and [Kim \(2019\)](#) explored the ways in which church groups use the English language as a source of evangelizing new members and acculturating them to the dominant culture. Other researchers have studied this dynamic in China, which is experiencing rapid migration ([Wang, 2018](#); [Wang and Froese, 2019](#)). Lastly, [Shaw \(2019\)](#) and [Vikdahl and Skeie \(2019\)](#) explored religion and power in European Union countries that required religious education as a part of their overall curriculum and the inclusion of dialogical activities in the classroom.

[Baquedano-López](#) took an in-depth look at a Catholic religious education program in Los Angeles, California serving Mexican immigrant families as they eliminated the Spanish-based class (Doctrina) in favor of "English-only" instruction. The church resided in a working-class neighborhood of mostly Mexican and Central American immigrants that were fluent in both English and their native language(s). Families enrolled their students in the Doctrina program because of the cultural connection and emphasis on the Spanish language. Even though the program was much larger than the English-only class and served well over 100 students, it was eliminated from church programming. The Latinx leaders of the Doctrina class resisted the insistence on English-only by continuing to conduct classes in Spanish, an alternative third space. [Baquedano-López \(2004\)](#) used de Certeau's *strategies and tactics* to describe the phenomenon that took place at this church. The church leaders implemented an English-only policy as a strategy to diminish the influence of the Spanish language and culture, while the Latinx leaders disregarded this policy by continuing to utilize the language as a connection to the community and culture, thus demonstrating the nexus of religion and power and the ways in which marginalized people subverted the dominant.

In their three-year ethnographic research of Chinese immigrant families, [Hsiao et al. \(2018\)](#) studied a Christian church after-school program, in which they found startling differences between the experiences of working-class and middle-class Chinese immigrants and their relationships with the church. The working-class Fuzhunese-Chinese families were concerned about the pressure to convert to Christianity but saw the program to improve their children's performance in school, especially in learning English. On the other hand, middle-class Chinese families could choose whether to participate in religious or non-religious educational opportunities as a tool of integration. While all Chinese immigrant children in this study socialized into Christianity at varying degrees, the class status of the family determined in which activities they could participate. The middle-class Chinese families, both Christian and non-Christian, saw church activities as social gatherings, whereas the working-class Chinese families who resisted Christian norms lost an opportunity to acculturate into society, and thus, were "othered" by the dominant norms of the Protestant Christian community.

In non-US settings such as North Korea and South Korea and the European Union, power dynamics are at play in the context of English language and literacy learning. For example, [Kim \(2019\)](#) explored missionary English teaching sites and the access to power and privilege they hold for minority population participants through the political economy of language and society. The main purpose for missionaries in this economy is to convert non-believers to Christianity, whereas the top priority of the North Korean students is to learn English to gain education credentials, so English serves as a currency for career advancement. Evangelical Christians in South Korea used English as a tool to evangelize North Koreans. More research into this phenomenon can provide us an opportunity to understand "the dominant ideologies and power relations of the larger society" ([Kim, 2019](#), p. 21) for which [Han \(2018\)](#) argued in her initiative to build a subfield focused on religion and language acquisition.

Research in the European Union context offers a more expansive view of religious literacy integration in curriculum. [Shaw \(2019\)](#) studied European Union countries that are required to provide a religious education curriculum to their students. The educational model suggested in drew on data from a national study of educational stakeholders' views in Great Britain. Shaw used [Dinham and Shaw's \(2017\)](#) framework as the tool to develop quality religion and worldview literacy instruction for students using the strands: *category, disposition, knowledge, and skills*. The data from stakeholders suggested broadening the *category* of religion and

worldview to include the informal-the lived social and political experiences of individuals to gain “authentic representations of religion and worldview” (Shaw, 2019, p. 5). Similarly, *disposition* of religion and worldview literacy promoted self-awareness and intercultural understanding toward religion instead of a preconceived set of principles of tolerance and respect. The knowledge strand asked students to consider the relationship between the religious and the secular and engage in dialogical practices to understand the complexity of religion and belief as it is constructed by individuals. Finally, the skills associated with religion and worldview literacy, discernment, meant the ability to self-examine one’s own beliefs and worldviews while also engaging in dialogue about someone’s religion or belief with grace and sensitivity. Shaw advocated for a critical lens when thinking about religion and worldviews. She also suggested that religious education classes be used to not only educate students in the major world religions but also as an opportunity to gain self-awareness and understanding of students’ beliefs and worldviews.

Using a comparative case study design, Vikdahl and Skeie (2019) analyzed religion and dialogue within education in London, England; Hamburg, Germany; Stockholm, Sweden; and Stavanger, Norway. While each country has distinct curricular goals, all case studies fit into the European research project “Religion and Dialogue in Modern Society” (Weisse, 2010). The findings from this cross-case analysis showed that while students like engaging in religion-related dialogue at school, they found it to be difficult to understand a broader view outside of their own personal opinions and beliefs. Students also indicated being insecure about sharing their own religious beliefs, especially if they differed from those of the majority, and feared social consequences from their peers. Teachers found that religion-related dialogue was threatened by “historical, religious, and political conflicts” (Vikdahl and Geir, 2019, p. 124). Also, the emphasis on curricular standards, learning goals, and assessments restricted dialogical practices in the classroom. Implications from this study suggest that dialogical practices are not just a teaching tool for students, but a set of skills they need to make sense of the world in which they live.

In turn, Wang (2018) provided insight into the largest ethnic minority group in China, the Hui, and their experiences migrating from rural to more urban areas in northwest China for economic opportunities. Almost 100% of the Hui identify as Muslim which has caused conflicts for them in the new urban setting. The participants in this study included fifteen seventh and eighth grade students who had moved to the region within the past five years. Through narrative stories, semi-structured interviews, and field observations, Wang (2018) found that all the participants distinguished themselves by their multiple identities but placed their religious identity of Muslim above the rest. Other findings point to an incompatibility of secular culture and religion. Religious symbols are not allowed in Chinese public spaces, including schools, therefore, students were not able to wear traditional Muslim clothing, nor were they given time to pray during the school day. Since pork is the main meat consumed in the region, Muslim Hui found eating out difficult. The fasting obligation during Ramadan made it challenging for students to concentrate on their schoolwork and many admitted to not participating the whole month.

The implication of the study is for schools in this region to develop more culturally relevant pedagogy practices to help students integrate their religious beliefs into the secular classroom and to become more adept to the urban environment in which they are living. School accommodations for Muslim students who fast and cannot eat pork is also relevant as it was in Sarroub’s (2005) ethnography of a high school in Michigan that was the academic home of a large Yemeni and Muslim population and whose administrators implemented inclusive accommodations in the cafeteria and classrooms. The conflict the participants experienced in northwest China adapting to the dominant Han culture while also wanting to maintain their religious beliefs and customs is a good example of power dynamics that involve both religious and rural population minorities. This scholarly research is imperative for understanding ways in which institutional powers weaponize religion to further minoritize people.

Additionally, Wang and Froese (2019) analyzed data from the 2010 Wave of the China General Social Survey to better understand the relationship between education and religious tolerance in China. The survey data concluded that education was the largest variable in determining religious tolerance among Chinese citizens. The more education an individual attained, the more open they were likely to religious diversity. At the same time, Chinese respondents noted that religious institutions are dangerous and should be socially and politically regulated so that they do not influence the decisions of the government. The researchers recognized the harsh religious laws in place in China and concluded that highly educated individuals may be reluctant to share their religious beliefs openly and honestly. Therefore, the openness to religious practices and the skepticism of their secular influence needs to be examined more fully using a qualitative research design. Wang and Froese acknowledged the people’s fear in responding to the survey and even alluded to the example of Muslim re-education camps implemented in rural China that may have contributed to less than forthcoming survey response.

The research about power via religion, language, and literacy demonstrates the symbiotic, political relationship that religious institutions have with schools and government to maintain their status. The English language classes (Kim, 2019), safety and academic support (Hsaio et al., 2018), and formation of a cultural community (Baquedano-López, 2004) provided by these religious institutions disguise their intent of assimilation according to Protestant cultural norms. This is contrasted by the research in China (Wang, 2018; Wang and Froese, 2019) involving religious individuals living in a country that does not openly recognize nor accept religious practices in any setting, private or public. Shaw’s (2019) aforementioned research and dialogical framework offers an in-between option for accommodation and education in such settings.

Practices in the classroom that engage religious literacies pluralistically

The world in the 21st Century is more interconnected than ever before, but the religious differences continue to foster tension and conflict within governments and communities. Scholars continue to advocate for religious pluralism despite the shift of leading

democracies to a more nationalist worldview (Lockley-Scott, 2019). In 2010, the European Union published a report on the effectiveness of the religious education (RE) requirement in European public schools. The conclusions of the study demonstrated a need for schools to make a place in their curriculum for dialogue and discussion about religious and non-religious worldviews (Weisse, 2010). Strengthening the religious and cultural dialogue would promote intercultural understanding, respect for otherness, and coexistence in a democratic society (Weisse, 2010). The lofty goals outlined in the report were met with obstacles when individual school systems attempted to implement the practices within their own classrooms.

For many European RE teachers, the purpose of the compulsory class was to provide students with a comprehensive view of the world's major religions to create citizens prepared to participate in a democracy (Liljestrand, 2016). In a small study of nine RE teachers in Sweden, Liljestrand (2016) found that the educators struggled to determine whether democratic citizens were created by focusing on the similarities or differences of the religions studied. The findings suggested that individual teachers' approaches to RE were situated within "how they define themselves as teachers when certain issues are explicitly addressed" (Liljestrand, 2016, pp. 328–329). Niemi et al.'s (2019) study supported the notion of reflexivity being a major factor in individual teacher's approaches to RE. In a survey of 181 Finnish pre-service and practicing teachers of world religions, the results showed that younger, less-experienced teachers favored an approach that called more attention to differences by providing students with religious exemptions from certain school activities, gender-based grouping when necessary, and spaces to pray at school (Niemi et al., 2019). Older teachers, on the other hand, tended to lean toward similarities as they attempted to show tolerance and acceptance but did not go so far as to make accommodations for those students (Niemi et al., 2019).

Currently, the practices surrounding religion in the classroom are unique to the beliefs and values of the teachers and the geographical regions in which they teach. The US recognizes the diversity of religions and cultures within the country and attempts to preserve equity among all groups by maintaining a clear and strong division between church and state. The numerous court cases surrounding the legality of public displays of religion in schools demonstrates the resistance to any kind of education or dialogue on the topic of religious diversity within classrooms (Bindewald et al., 2017). In recent years, Canada began following the US model and secularized its public schools to better accommodate diverse student population (Bindewald et al., 2017). The result of a lack of RE in US and Canadian public schools has been "high rates of religious illiteracy" (Patrick et al., 2017) among its student populations. Richardson (2017) and Patrick et al. (2017) argued for RE classes in public schools because they offer the skills required for opening dialogues among religious and secular worldviews and give minoritized religions the space for recognition within their communities (Davila, 2015; Patrick et al., 2017; Richardson, 2017; Vallerand, 2018). Living in a diverse and democratic society requires more from citizenship education courses than the teaching of tolerance; it must also focus on respect and coexistence (Patrick et al., 2017; Weisse, 2010).

While the scholarship shows a need for public schools to embrace RE, the *how-to* is a little more ambiguous. Lockley-Scott (2019) and Vallerand (2018) both advocated for the implementation of religious dialogue in UK and US schools, respectively, as a way to decrease cultural tensions and explore their own identities, but the nationalistic worldview recently adopted by Great Britain and the US has not been conducive to religious pluralism (Lockley-Scott, 2019; Vallerand, 2018). Before religious dialogue can take place in the classroom, the political context of the nation state must be more open and inclusive of diverse and multiple worldviews. Additionally, teachers require on-going professional development opportunities to be more confident in engaging their students in controversial topics (Davila, 2015; Lockley-Scott, 2019; Richardson, 2017). Without professional development and practice, RE teachers may end up doing more harm than good (Davila, 2015; Dinham and Shaw, 2017; Lockley-Scott, 2019; Vallerand, 2018).

The navigation of the landscape between religious and secular worldviews in public schools is a bit like an obstacle course. The scholarship supports the need for all people of democracies to learn the cultural values and beliefs of the religions that are prevalent in their communities, but a nationalistic political climate fosters fear of outside, minority populations (Dinham, 2017; Juzwick, 2014; Patrick et al., 2017; Richardson, 2017; Vallerand, 2018; Weisse, 2010). Instead, teacher education programs ought to include tools for engaging pre-service teachers in pluralistic dialogue in their classrooms such that they are better prepared to participate in the democratic societies they will inherit (Davila, 2015; Dinham, 2017). A glimmer of hope that religious pluralistic ideas can come to fruition appeared in Addai-Mununkum's (2016) multi-case study wherein he observed the pluralistic practices of three schools affiliated with different religions. The results show that despite the diverse religious backgrounds of the students at each school, the Muslim school demonstrated more evidence of engaging in a dialogue with Christians than the Christian schools showed to their Muslim students. The focus on moral education and learning from one another set the Muslim school apart from the other two religious schools and allowed all students to feel and be welcomed in its space (Addai-Mununkum, 2016).

Western democracies could learn from the Muslim school in Ghana, Africa. The impact of religious literacies in education should not be a taboo topic in public schools. Rather teacher preparation programs and professional development workshops should consider providing teachers the tools necessary to engage young people in critical conversations focused on religion, multiple world views, and a diversity of perspectives grounded in historically, sociologically, anthropologically, linguistically informed practices that are accurate and factual. The current nationalistic climate and Islamophobia that is rampant across the world could be snuffed out with an open-minded approach to critical conversations and education surrounding religious literacies (Mir and Sarroub, 2019). As the 21st century continues to unfold, a less insular approach to inquiry in language and literacy studies about the influence of religious literacies could potentially offer new questions and answers such that people do less harm to one another, and young people learn to talk and argue democratically. As Durkheim (1912) pointed out, religion is the product of human activity, not divine intervention, so scholars and educators can study it as such and help teachers be more confident in engaging with their students to better understand religious literacies in society.

References

- Addai-Mununkum, R., 2016. Adding and dividing by religion: the not-so-hidden curriculum of mission-public schools in Ghana. *Relig. Educ.* 44 (2), 225–243. <https://doi.org/10.1080/15507394.2016.1243435>.
- Ahmed, S., 2010. *The Promise of Happiness*. Duke University Press.
- Baquedano-López, P., 2004. Traversing the center: The politics of language use in a catholic religious education program for immigrant Mexican children. *Anthropol. Educ. Q.* 35 (2), 212–232.
- Baquedano-López, P., 2008. The pragmatics of reading prayers: learning the act of contrition in Spanish-based religious education classes (doctrina). *Text & Talk* 28 (5), 581–602. <https://doi.org/10.1515/TEXT.2008.030>.
- Bigelow, M., 2008. Somali adolescents' negotiation of religious and racial bias in and out of school. *Theory Into Pract.* 47 (1), 27–34. <https://doi.org/10.1080/00405840701764706>.
- Bindewald, B.J., Sanatullova-Allison, E., Hsiao, Y.-L., 2017. Religion and public education in pluralist, democratic societies: some lessons from the United States and Canada. *Relig. Educ.* 44 (2), 180–202. <https://doi.org/10.1080/15507394.2017.1283193>.
- Choi, J., Tinker Sachs, G., 2017. Adolescent multilinguals' engagement with religion in a book club. *J. Adolesc. Adult Literacy* 60 (4), 415–423. <https://doi.org/10.1002/jaal.591>.
- Davila, D., 2015. Who needs diverse books? Preservice teachers and religious neutrality with children's literature. *Res. Teach. Engl.* 50 (1), 60–83.
- Dinhnam, A., Shaw, M., 2017. Religious literacy through religious education: the future of teaching and learning about religion and belief. *Religions* 8 (7), 119. <https://doi.org/10.3390/rel8070119>.
- Durkheim, E., 1912. *The Elementary Forms of the Religious Life*. George Allen & Unwin Ltd, London.
- Han, H., 2007. Calm and humble in and through evangelical christianity: a Chinese immigrant couple in Toronto. In: Allyson, J. (Ed.), *Language and Religious Identity: Women in Discourse*. Palgrave Macmillan, Basingstoke, pp. 196–219.
- Han, H., 2011a. 'Love your China' and evangelize: religion, nationalism, racism, and immigrant settlement in Canada. *Ethnogr. Educ.* 6 (1), 61–79. <https://doi.org/10.1080/17457823.2011.553080>.
- Han, H., 2011b. Social inclusion through multilingual ideologies, policies and practices: a case study of a minority church. *Int. J. Biling. Educ. Biling.* 14 (4), 383–398. <https://doi.org/10.1080/13670050.2011.573063>.
- Han, H., 2014. Accessing English and networks at an English-medium multicultural church in East Canada: an Ethnography. *Can. Mod. Lang. Rev.* 70 (2), 220–245. <https://doi.org/10.3138/cmlr.1871>.
- Han, H., 2018. Studying language learning and teaching: building a subfield. *Mod. Lang. J.* 102 (2), 432–445.
- Hsiao, Y.-L., Bindewald, B., Bailey, L.E., 2018. Educational experiences of Chinese immigrants in Smalltown, USA: acculturation and the intersections of class, religious norms, and race. *Relig. Educ.* 45 (2), 208–225. <https://doi.org/10.1080/15507394.2018.1425075>.
- Juzwik, M.M., 2014. American Evangelical biblicism as literate practice: a critical review. *Read. Res. Q.* 49 (3), 335–349. <https://doi.org/10.1002/rq.72>.
- Juzwik, M.M., McKenzie, C., 2015. Writing, religious faith, and rooted cosmopolitan dialogue. *Writing, religious faith, and rooted cosmopolitan dialogue. Portraits of Two American Evangelical Men in a Public School English Classroom* 32 (2), 121–149. <https://doi.org/10.1177/0741088315576480>.
- Juzwik, M.M., LeBlanc, R.J., Dávila, D., Rackley, E.D., Sarroub, L.K., 2022. Spiritual and religious meaning making in language and literacy studies: Global perspectives on teaching, learning, curriculum, and policy. *Engl. Teach. Pract. Critiq.* 21 (3), 225–237. <https://doi.org/10.1108/ETPC-03-2022-0051>.
- Kim, E.-Y., 2019. English as a site of Evangelical contact: a critical ethnography of missionary English teaching between South and North Koreans. *J. Lang. Ident. Educ.* 18 (1), 10–24. <https://doi.org/10.1080/15348458.2019.1575739>.
- Liljestrand, J., 2016. Education for citizenship in Swedish RE—approaches and dilemmas in teachers' talk. *Relig. Educ.* 44 (3), 317–330. <https://doi.org/10.1080/15507394.2016.1267541>.
- Lockley-Scott, A., 2019. Closing down the discussion: is a classroom a conducive space for religion related dialog? A United Kingdom based case study. *Relig. Educ.* 46 (1), 40–58. <https://doi.org/10.1080/15507394.2019.1577707>.
- Mir, S., Sarroub, L.K., 2019. Islamophobia in US education. In: Zempi, I., Awan, I. (Eds.), *The Routledge International Handbook of Islamophobia*. Routledge Press, pp. 298–309.
- Moore, L.C., 2015. Change and variation in family religious language policy in a West African Muslim community. *Lang. Pol.* 15 (2), 125–139. <https://doi.org/10.1007/s10993-015-9366-y>.
- Niemi, P.-M., Kimanen, A., Kallioniemi, A., 2019. Including or excluding religion and worldviews in schools? Finnish teachers' and teacher students' perceptions. *J. Beliefs Values* 41 (1), 114–128. <https://doi.org/10.1080/13617672.2019.1617628>.
- Papen, U., 2017. Hymns, prayers and Bible stories: the role of religious literacy practices in children's literacy learning. *Ethnogr. Educ.* 13 (1), 119–134. <https://doi.org/10.1080/17457823.2016.1277773>.
- Patrick, M.L., Gulayets, V., Peck, C.L., 2017. A call for teacher professional learning and the study of religion in social studies. *Can. J. Educ.* 40 (4), 603–637.
- Peshkin, A., 1986. *God's Choice: The Total World of a Fundamental Christian School*. University of Chicago Press, Chicago.
- Rackley, E.D., 2014. Scripture-based discourses of latter-day Saint and methodist youths. *Read. Res. Q.* 49 (4), 417–435. <https://doi.org/libproxy.unl.edu/10.1002/rq.76>.
- Rackley, E.D., 2016. Religious youths' motivations for reading complex, religious texts. *Teach. Coll. Rec.* 118 (11), 1–50.
- Reyes, C.C., 2009. El libro de recuerdos [book of memories]: a Latina student's exploration of self and religion in public school. *Res. Teach. Engl.* 43 (3), 263–285.
- Richardson, M., 2017. Religious literacy, moral recognition, and strong relationality. *J. Moral Educ.* 46 (4), 363–377. <https://doi.org/10.1080/03057240.2017.1324771>.
- Rosowsky, A., 2013. Faith, phonics, and identity: reading in faith complementary schools. *Literacy* 47 (2), 67–78. <https://doi.org/10.1111/j.1741-4369.2012.00669>.
- Rosowsky, A., 2019. Sacred language acquisition in superdiverse contexts. *Ling. Educ.* 53. <https://doi.org/10.1016/j.linged.2019.100751>.
- Sarroub, L.K., 2009a. Finding husbands, finding wives: how being literate creates crisis. In: MacGillivray, L. (Ed.), *Literacy Practices in Times of Crisis*. Lawrence Erlbaum Press.
- Sarroub, L.K., 2009b. Glocalism in literacy and marriage in transnational lives. *Crit. Inq. Lang. Stud.* 6 (1–2), 63–80.
- Sarroub, L.K., 2007a. Seeking refuge in literacy from a scorpion bite. *Ethnogr. Educ.* 2 (3), 365–380.
- Sarroub, L.K., 2007b. How Yemeni American Women Use Literacy: North America. *Encyclopedia of Women and Islamic Cultures*, vol. 5. Brill Academic Publishers, Leiden, The Netherlands.
- Sarroub, L.K., 2005. *All American Yemeni Girls: Being Muslim in a Public School*. University of Pennsylvania Press.
- Sarroub, L.K., 2002. In-betweenness: religion and conflicting visions of literacy. *Read. Res. Q.* 37 (2), 130–148.
- Sarroub, L.K., 2001. The sojourner experience of Yemeni American high school students: an ethnographic portrait. *Harv. Educ. Rev.* 71 (3), 390–415.
- Shaw, M., 2019a. Towards a religiously literate curriculum – religion and worldview literacy as an educational model. *J. Beliefs Values* 1–12. <https://doi.org/10.1080/13617672.2019.1664876>.
- Shaw, M., 2019b. Towards a religiously literate curriculum – religion and worldview literacy as an educational model. *J. Beliefs Values* 41 (2), 150–161. <https://doi.org/10.1080/13617672.2019.1664876>.
- Skerrett, A., 2014. Religious literacies in a secular literacy classroom. *Read. Res. Q.* 49 (2), 233–250. <https://doi.org/libproxy.unl.edu/10.1002/rq.65>.
- Vallerand, V., 2018. Teaching about religion in a democratic society: an analysis of the Quebec ethics and religious culture program. *Relig. Educ.* 45 (3), 270–286. <https://doi.org/10.1080/15507394.2018.1535764>.
- Vikdahl, L., Skele, G., 2019. Possibilities and limitations of religion-related dialog in schools: conclusion and discussion of findings from the ReDi project. *Relig. Educ.* 46 (1), 115–129. <https://doi.org/10.1080/15507394.2019.1577712>.

- Wang, T., 2018. Religion-based cultural identity and conflicts of migrant Muslim students in Northwest China. *Race Ethn. Educ.* 21 (6), 858–875. <https://doi.org/10.1080/13613324.2017.1395324>.
- Wang, X., Froese, P., 2019. Attitudes toward religion and believers in China: how education increases tolerance of individual religious differences and intolerance of religious influence in politics. *Relig. Educ.* 47 (1), 98–117. <https://doi-org.libproxy.unl.edu/10.1080/15507394.2019.1626211>.
- Weisse, W., 2010. REDCo: a European research project on religion in education. *Relig. Educ.* 37, 187–202. <https://doi.org/10.1080/15507394.2010.513937>.
- Weyand, L., Juzwik, M., 2019. Schooling activist evangelical literacy: speaking, writing, and storying Christian faith in dialogue with public secondary literacy curriculum. *Ling. Educ.* 55, 1–16. <https://doi.org/10.1016/j.linged.2019.100789>.
- Wynter-Hoyte, K., Boutte, G.S., 2018. Expanding understandings of literacy: the double consciousness of a Black middle class child in church and school. *J. Negro Educ.* 87 (4), 375–390.

Further reading

- Biesta, G., Aldridge, D., Hannam, P., Whittle, S. (2019). Religious Literacy: A Way Forward for Religious Education? A Report Submitted to the Culham St Gabriel's Trust. <https://www.reonline.org.uk/wp-content/uploads/2019/07/Religious-Literacy-Biesta-Aldridge-Hannam-Whittle-June-2019.pdf>.
- Casson, A., Cooling, T., 2019. Religious education for spiritual bricoleurs? the perceptions of students in ten Christian-ethos secondary schools in England and Wales. *J. Beliefs Values* 41 (1), 20–33. <https://doi.org/10.1080/13617672.2019.1632596>.
- Francis, L.J., Village, A., Parker, S.G., 2016. Exploring the trajectory of personal, moral, and spiritual values of 16- to 18-year-old students taking religious studies at a level in the UK. *J. Beliefs Values* 38 (1), 18–31. <https://doi.org/10.1080/13617672.2016.1232567>.
- Iyer, S., 2018. The modern madrasa: a case study of religious education in India. In: *The Economics of Religion in India*. Harvard University Press, Cambridge, MA, pp. 182–220.
- Jones, I., 2016. Why study a level religious studies? Qualitative perspectives from two English Midlands sixth forms. *J. Beliefs Values* 38 (1), 3–17. <https://doi.org/10.1080/13617672.2016.1232566>.
- Juzwik, M.M., Stone, J., Burke, K.J., Dávila, D., Barros, S.R., Kirkland, D.E., 2020. Legacies of Christian Language and Literacies in American Education: Perspectives on English Language Arts Curriculum, Teaching, and Learning. Routledge, New York, NY. <https://doi.org/10.4324/9780429027604>.
- Latif, D., 2018. Considering religious education in context: politics, reform, and debates among Turkish Cypriots. *J. Beliefs Values* 40 (1), 64–76. <https://doi.org/10.1080/13617672.2018.1472998>.
- Leis-Peters, A., 2018. Religious literacy in welfare and civil society: a Nordic perspective. In: Baker, C., Crisp, B., Dinham, A. (Eds.), *Re-imagining Religion and Belief: 21st Century Policy and Practice*. Bristol University Press, Bristol, pp. 145–168. <https://doi.org/10.2307/j.ctv5qdfqv.12>.
- Riegel, U., Delling, S., 2019. Dealing with worldviews in religious education. *J. Beliefs Values* 40 (4), 403–415. <https://doi.org/10.1080/13617672.2019.1618150>.
- Rosowsky, A., 2013. Religious classical practice: entextualisation and performance. *Lang. Soc.* 42 (3), 307–330. <https://doi.org/10.1017/s0047404513000250>.
- Rusitoru, M.-V., Kallioniemi, A., 2018. Religious education and spiritual development as a cultural component of the lifelong learning policy in Romania. The example of orthodox Christianity. *J. Beliefs Values* 40 (1), 31–44. <https://doi.org/10.1080/13617672.2018.1519688>.
- Scholes, S.C., 2019. Challenges and opportunities in religious education: re-considering practitioners' approaches in scottish secondary schools. *Relig. Educ.* 115 (2), 184–200. <https://doi.org/10.1080/00344087.2019.1677996>.
- Wong, M.-Y., 2018. Current teaching practice for religious education in Hong Kong: implications for 'learning from religion'. *J. Beliefs Values* 40 (2), 133–145. <https://doi.org/10.1080/13617672.2018.1548827>.
- Zembylas, M., Loukaidis, L., Antoniou, M., Antoniou, P., 2018. The political function of religious education in an ethnically divided society: Greek-Cypriot Teachers' Perspectives on Conflict, Peace, and Religious Teaching. *Relig. Educ.* 45 (2), 226–248. <https://doi.org/10.1080/15507394.2017.1419771>.

Critical literacies & social justice: social justice approaches in schools

Nerida Spina^a, Barbara Comber^b, and Michelle Jeffries^a, ^a Queensland University of Technology, Faculty of Education, Kelvin Grove, QLD, Australia; and ^b University of South Australia, Education Futures, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	217
Critical literacies and social justice	217
Literacy as a political project	218
Critical literacy through Nancy Fraser's theory of justice	219
Redistribution: the "what" of critical literacy	220
Recognition: the "who" of literacy	221
Representation: the "how" of curriculum justice	222
A tool for action: students and teachers as social-justice advocates and activists	222
Conclusion	223
References	224
Further reading	224
Relevant websites	224

Introduction

Critical literacy and social justice theories are both related to the exploration of power, privilege and equity. Critical literacy is an approach to language and communication education that teachers explore with students and colleagues to engage in deep thinking and substantive analysis of unequal and oppressive power relations. This approach enables teachers to advocate for dismantling of systemic injustices, and to propose change. As authors writing this entry, we acknowledge that much of the schooling discussed is based in settings that to various degrees reflect a history of colonization, and that the entry reflects this Western perspective. We heed Maori scholar [Smith \(1999\)](#), who stated that "social science... is not an innocent or distant academic exercise but an activity that has something at stake" (p. 5). This is certainly the case for critical literacies, which have grown out of Western education and academic traditions. As [Tuck and Yang \(2014\)](#) point out, the academy itself is "as an apparatus of settler colonial knowledge already domesticates, denies, and dominates other forms of knowledge" (p. 235). With these limitations in mind, we note that there is no one "right way" or single definition of critical literacy or social justice, though there are commonalities in their theoretical underpinnings, requiring serious engagement with issues such as diversity, privilege, fairness and equity for people, animals and the earth.

This entry begins with a discussion of critical literacy and an overview of the interplay between school education, literacy and social justice. We introduce critical literacy as a field, including its theoretical underpinnings. We then use [Fraser's \(2008a,b\)](#) three dimensional model of social justice to consider how critical literacy has been taken up in schools and by education researchers to forward social justice agendas. We conclude by explaining how teachers and students can employ critical literacy as a catalyst for deep learning that enables them to become advocates and allies for justice.

Critical literacies and social justice

Critical literacy is an approach that seeks to unsettle and contest taken-for-granted cultural and social practices, and to use literacy to spark deep questioning and imagine more just futures. In unsettling dominant discourses and practices, critical literacy educators typically explore what [Janks \(1993\)](#) refers to as both "small-p" and "big-P" politics. That is, critical literacy educators examine the micropolitics and power relations that are part of everyday life in schools—for example, exclusion in the playground that is based on racist, homophobic or ableist views, or action for social change, such as how we recycle unwanted objects. In big-P politics, Janks explains that the focus turns toward wider, structural power mechanisms such as how governments attend to issues of racial injustice, global warming and the effects of capitalism and inequality on marginalized groups. As the work of researchers and teacher-researchers such as Anne Haas Dyson, Vivan Vasquez, Timothy Lintner, Summer Melody Pennell and others demonstrates, even young children are capable of understanding complex issues of justice in their everyday lives.

While there is no set definition of critical literacy, key principles have been outlined in the research literature. Critical literacy encompasses the multitude of ways teachers create socially just classrooms and encourage their students to work toward social justice in their communities and beyond. This includes adopting a critical approach toward teaching literacy that sees issues of oppression, justice and power-relations as worthy of inquiry.

Critical literacy is embedded in how teachers approach texts and literate practices. Thinking about how texts are constructed from particular perspectives encourages students to question the assumed neutrality of discourses, practices and texts. Providing opportunities for students to make connections and be curious about their everyday realities and concerns, and problematizing the world around them, are central facets to adopting a critical stance. We recognize with Hilary Janks (2000) that critical literacy can sometimes operate as “a rationalist activity that does not sufficiently address the non-rational investments that readers bring with them to texts and tasks” (p. 7). Critical deconstructions of texts can set up serious, rational discussion that may fail to take into account students’ situated realities. Janks provides the example of her emotional reaction to the terrorist attacks in the United States on 9/11. Janks had flown through New York in the weeks preceding 9/11, which left her with a sense of disbelief, and an “irrational, unconscious, all-consuming” vulnerability that made it difficult to engage intellectually or apply critical thinking about the events at the time. We therefore argue for an approach toward critical literacy that incorporates both rationalist approaches toward the disruption of dominant structures in education alongside an acknowledgment of students’ and teachers’ lived realities. In an increasingly polarized world, this approach to critical literacy may be less likely to alienate, or “preach to the converted” (Janks, 2000, p. 21). For those engaged in critical literacy, this means being open to understanding the “lopsided” influence of Western academic traditions (Takayama, 2009, p. 354), and to make an effort to genuinely engage in critical pedagogic approaches that have alienated minoritized groups.

The reliance of reasoned approaches to critical literacy also has the potential to limit students’ enjoyment and engagement with critical thinking, and texts. In moving beyond a purely rationalist approach, critical literacy becomes a space where enjoyment and playfulness can come together in the pursuit of troubling assumptions and dominant discourses. Hannah Gadsby’s (2018, 2020) acclaimed stand-up comedy specials illustrate how a critical stance can be taken to dominant views on comedy, sexism, autism, trauma and even art history. Her playful approach, including laughter and discomfort for audiences, suggests that humor can convey important messages and spark critical thinking and discussion. We suggest that teachers and students be “given permission to be other than sensible and serious” (Comber, in Janks, 2000, p. 9) when enacting critical literacies for social justice.

In the following section, we draw on some seminal social and educational theory to consider the interplay between literacies and wider social forces, including how literacies have been linked to the reproduction of dominant views and practices. We then consider how critical literacy has emerged as a practice for contesting the inequalities borne out of these political circumstances, with a view to bringing about a renewed focus on social justice and activism in schooling.

Literacy as a political project

Literacy and schooling are inherently political projects, embedded in wider social views around education and childhood. Critical literacy education has often been caught in debates about how literacy in school is practiced (e.g., Luke, 2018), how childhood is constituted, and whether knowledge is seen to be neutral, or imbued with social, cultural and political meaning.

In Western scholarship, the emergence of childhood as a category (as opposed to the treatment of children as adults, as was the case in medieval times) (Aries, 1962) has been linked to the emergence of print and literacy from the Renaissance onwards (Postman, 1985). Literacy instruction was primarily positioned as being about the acquisition of skills, rather than having cultural or political goals such as contributing to more just societies. As alternate approaches emerged that were built on a philosophical belief in literacy that emphasized meaning and saw literacy instruction as a means of contesting unfair power-relations, literacy became one of the most politicized parts of education policy and practice.

Dominant views around the nature of pedagogy have had a significant impact on how literacy is understood and taught in schools. For instance, the late 1800s saw the “scientific management” principles and practices espoused by Frederick W. Taylor being introduced into the field of education (Apple, 2014, p. 92). Developed for factory work, Taylor’s focus was on efficiency, standardization, increased control of work processes, and the maximization of outcomes/results. In literacy instruction, the application of this industrial model has meant a focus on skills that can be most easily standardized and measured such as spelling, language conventions, and the use of standardized/basal readers for teaching decoding and literal comprehension (Hoffman and Alvermann, 2020).

A body of work in a field that has become known as the sociology of education, produced from the 1960s onwards, has become the basis for an emergent field of literacy scholars who sought to understand how intractable race, gender and class based injustices were produced (at least in part) by literacy practices such as inequitable access to texts, and different forms of literacy. Reflecting on the development of the field, Luke (2018) writes that “Young, Bernstein, Apple, Cazden, Hymes, and others—coming from very different foundational and methodological bases—began from the empirically and quantitatively documented social facts of educational inequality” (p. 178).

The seminal contribution of Paulo Freire in the *Pedagogy of the Oppressed* (1974) provided a foundation for critical education and pedagogy theorization. In particular, understandings of literacy as implicated in the construction of the social world emerged in the 1960s and 70s. Freire reimagined literacy as a political enterprise that could take up matters of social justice through pedagogies that seek to problematize and reshape social, political and economic relations. Freire used a “banking” metaphor (in which students are

passive recipients of knowledge) to contest dominant models, and to argue that students could instead be offered pedagogies that raise their consciousness of class-based oppression, and to critically reflect upon social structures. In *Literacy: Reading the Word and the World* (Freire and Macedo, 2004), originally published in the 1980s, he set out to contest technicist assumptions about literacy, explaining how literacy is implicated in (in)justice because it can either empower or disempower, reproduce injustice or lead to emancipatory change. Freire's writings extended from his early work on Brazil and Latin America, to discussion of colonized countries such as Guinea-Bissau (where he organized a literacy campaign in the 1970s), as well as European and North American education systems. Freire's work has been taken up extensively since that time, for example by scholars such as Henry Giroux, Hilary Janks, Colin Lankshear, Allan Luke, and Mariana Souto-Manning who have advocated for emancipatory and decolonizing approaches to social justice, through critical literacy practice. Freire's writing on emancipation, and the importance of consciousness-raising in education was taken up by early critical literacy writers, including those in *The New London Group* (1996). In Japan, the Buraku-min, who have a long history of subjugation have played a strong role in applying Freire's work, in the development of liberatory critical pedagogies (Takayama, 2009). The movement worked to reshape the consciousness of both Baruku children, as well as changing the attitudes of non-Baruku people. Understanding how the approach of scholars such as Freire has been taken up in different contexts is important. As Takayama suggests, too often, progressive voices are silenced in Western accounts of education systems, including Japan. Challenging assumptions about how critical literacy is understood is itself an important means of moving toward more emancipatory approaches. In China, for example, a country that has had longstanding links with Marxism as the official ideology, it might be assumed that critical pedagogies are embedded in education. However, Yan and Chang (2009) note that this is not the case, and that instead, "critical educational studies have ... been considered as capitalistic thought and ideology, remaining unknown until the 1980s" (p. 369).

The theoretical contributions of Pierre Bourdieu, Michael Apple, Allan Luke and others have also been highly influential. A diversity of critical literacy work emerged in the 1970s and 80s alongside other forms of activism and consciousness raising such as second wave feminism and black feminism, activism around diverse sexualities (such as the HIV/AIDS epidemic), and anti-apartheid movements. Luke (2018) reflecting on his early work notes that when he began to apply sociological approaches to literacy in the 1970s, that "much of this work was deemed unscientific and overtly political" (p. 22). Nevertheless, he persisted, along with other researchers and teachers, taking critical literacies forward into the digital era.

In 1996, a group of internationally renowned scholars known as the "New London Group" came together with a view to taking "deliberate, political and intellectual action" (Luke, cited in Luke, 2018, p. 298). Their seminal work around multiliteracies sought to de-center dominant accounts of literacy pedagogies, with their focus on print media, canonical works and decoding skills (phonics, spelling and so forth). Seeking to create a strong commitment to social justice by capitalizing on new online and digital media and the political environment of the day showed how teachers' critical engagement with multiliteracies could alter long-standing inequities such as the marginalization of Indigenous peoples, working classes and other cultural minorities. The group's manifesto on multiliteracies teaching argued that multiliteracies must not only tackle rapid technological changes to literacy, but also, must confront the increasing importance of being able to critical analyze multimodal texts, and acknowledge the cultural and linguistic diversity of students and communities. The group challenged teachers to consider the interplay between local happenings and global events in new times.

Critical literacy through Nancy Fraser's theory of justice

Public debates over literacy and equity are often dominated by reports of gaps in achievement between groups of students as measured on large-scale, standardized assessment programs. In contrast, struggles for equity are concerned with questioning hegemonic power systems that oppress groups and individuals, based on class, race, gender, sexuality, disability and so on. Early theorisations of justice and power (such as that of Marx in the mid-1800s) centered around the unequal allocation of resources, arguing that redistribution was needed to create a fairer society. Efforts to bring about parity that are couched in these terms argue that a lack of economic resources and control, and unequal control of labor constitute injustice. The 1990s saw an increasing emphasis on broader conceptions of justice that includes political and cultural aspects of advantage and domination.

Fraser's moral philosophy has greatly influenced debates about social justice. Her most significant contribution remains her three-dimensional theory of social justice, a model she has developed over many decades. The model began as two-dimensional, exploring concepts of "redistribution" and "recognition" as approaches to move away from identity politics toward cultural politics, and to address social injustices related to concepts such as class, race, sexuality, gender, and so on (Fraser, 2008a,b). Underlying this model is the contention that while discourses of social justice are usually centered on either redistribution or recognition as approaches to resolving social inequities, a combination of the two is required in order to execute equitable approaches to social justice (Fraser, 2008b). As a result of an increasingly globalized world, Fraser would later introduce a third concept, "representation" which attends to matters of political representation, governance and participation. Fraser (2008a) describes the three major forms of justice claims (redistribution, recognition, and representation as the "what," "who" and "how" of justice (p. 400).

Fraser's theory of justice has been taken up by educational researchers to understand the principles by which critical literacy offerings can work to achieve participatory parity (Fraser, 2008a) for diverse peoples and groups, and consider how effective different approaches to critical literacy have been in achieving their emancipatory aims. Fraser insists that for justice to be realized, all three dimensions must be attended to. A focus on achieving one component while necessary, is also insufficient. In literacy education, this

means that equal access to education funding and resourcing (for instance quality texts) while important, cannot on its own bring about justice. This means that a literacy policy focus that relies on redistributive justice is inadequate unless attention is also afforded to the recognition of students' and communities' cultural and social heritage, knowledges and ways of being; as well as providing students or communities with genuine opportunities for partnership in decision-making about literacy. We now turn to explain how viewing critical literacy through Fraser's family of claims reveals teachers' efforts and progress toward social justice agendas.

Redistribution: the "what" of critical literacy

Redistribution refers to the commitment to ensuring all students have access to resources that will enable them to lead good lives. Conceptually, Fraser refers primarily to the distribution of economic and material resources, which is an undeniably important project in redressing injustice. Students living in areas with significant poverty must have access to the resources needed to allow them to gain the literacy skills required to access employment and full citizenship as adults. Fraser (1997) states that economic injustices can be remedied through a kind of political-economic restructuring, such as redistributing education funding or reorganizing staffing to meet the needs of particular groups of students. If cultural injustices are rooted in economic injustices, Fraser (1997) suggests redistribution as a strategy to address injustice.

A range of scholarship that draws on sociological and anthropological works to explain how literacy is implicated in the perpetuation of structural inequities through unequal access to economic capital and knowledge been published over recent decades. Oakes' (2005) *Keeping Track: How schools structure inequality* explored how schooling in capitalist societies functioned as a technology for producing a compliant workforce, willing to accept economic inequities, unequal access to knowledge, and hierarchical divisions in education and employment. Oakes' work drew attention to the uneven access to different forms of literacy within schools, and illustrated how schooling was structured for injustice.

Anyon's (1981) *Social Class and School Knowledge*, also examined how injustice is perpetuated through differential access to curriculum knowledge. Anyon used Marxist theoretical categories to explain how scripted, low-intellectual quality literacies and a focus on compliance offered to working-class and minority students created lasting patterns of inequity. Anyon's work has been taken up by scholars such as Luke (2010) who have argued that critical literacy, and the distribution of equal access to quality pedagogies must be afforded to all students, not just those in elite schools, or affluent suburbs.

Teacher-researchers such as Pennell (2019) have used critical literacy (and critical mathematics) to spark curiosity among students about matters of redistributive justice. Pennell's work documents the inquiries of undertaken by one group of her middle school students as they explored news articles about school funding and student achievement. The class had begun by reading news articles, checking them for legitimacy, and asking questions such as "who has power?" (p. 22). Bringing together critical maths and critical literacy, she reports that groups of students created presentations based on their findings:

[one group] learned that school funding was not the same by state or even city. As private school students, this was not something they had ever needed to consider and gave them new insight into educational issues on a national scale (p. 22).

Education researchers and teachers have also considered the distribution of knowledge in their approach to critical literacy. Equal access to knowledge (as a form of capital that must be fairly distributed to ensure parity of participation), being a discerning and critical consumer of multimodal texts is an increasingly essential skill. Critical literacy has provided teachers with a framework for teaching students to question texts, and to consider the proliferation of untrustworthy material that exists both online, in the media, and in wider textual spaces. Nicole Cooke's (2018) work is an example of how critical literacies can be used to detect both misinformation (information that is incomplete, partial, ambiguous), disinformation (information contains deliberate falsehoods), and to question power structures and cultural biases. From texts on social media to images that have been digitally edited and online videos, her work illustrates how online information is often viewed as being universal, despite it frequently "reflect[ing] a culture of whiteness and sensibilities that are distinctly Western, masculine, and homogeneous in regard to notions of religion, progress, and modernity" (p. 37). She provides an example of a widely tweeted photograph of Rosa Parks, along with a claim that Parks played a role in ending racism as part of the US Civil Rights movement. As Cooke points out, a critical reading of this text allows us to see this tweet as a micro-aggression in that it is underpinned by a worldview that racism no longer exists. Her work draws attention to the role that librarians can play helping students to develop critical digital, media and visual literacies in an age where the internet dominates how many students access information.

Reflecting on the education of Indigenous students in Australia, Luke (2018) considers distribution in terms of "the curriculum question: which selective traditions should shape what will count as literacy, which texts and discourses, literacy practices and events will be codified and transmitted in schools, in whose interests, with what material and discourse consequences?" (p. 229). As Marnie Shay and Judi Wickes (2017) point out, counter-narratives, (such as their own stories, as two Indigenous Australian women) are an important way of disrupting racism, and valuing a diversity of experiences and perspectives. They write that, "the absence of critique and discussion about how constructs of race and issues of racism are so deeply entrenched in our experiences in educational institutions is a marker for how limited our progression as Indigenous Australians are" (p. 112).

Recognition: the “who” of literacy

Recognitive justice takes into account how diverse peoples, standpoints and knowledges are recognized in literacies. This component of Fraser’s model raises the question of “whose knowledge” is foregrounded and whose is marginalized in literacy education. Fraser (2009) noted that “[P]eople can also be prevented from interacting on terms of parity by institutionalized hierarchies of cultural value that deny them the requisite standing: in that case they suffer from status inequality or misrecognition” (p. 16).

Fraser (2003) contends that parity of participation requires a combination of material (distributive) and social (recognition) approaches to addressing injustices. Fraser labels the material as the “objective condition” of parity of participation, and refers to ensuring that independence, voice, and participation are made certain by material distribution, allowing subjects to interact on par with other members of society as peers. Equity in achieving social regard is referred to as “intersubjective condition” with regard to participatory parity. Thus, institutionalized patterns and norms that marginalize some categories of people and privilege others, as well as those that fail to acknowledge difference or ascribe excessive differences to groups, are impeded to ensure parity of participation.

The question of whose knowledge is seen to be a “legitimate” part of literacy curricula, and whose knowledge is absent is important, as scholars such as Luke (2018), Apple (2014) and others have established by examining the texts and curricula used in schools. These forms of knowledge are based on a selective tradition, in which the dominant culture reproduces itself by teaching students what is valued, while oppressing alternate forms of knowledge. Subcultures and diversities are often invisible within these school-based texts, which represent a narrow range of perspectives, histories and traditions. For instance, Indigenous peoples, girls, LGBTIQ+ peoples and immigrants have often been invisible in the texts used in western schooling. Unproblematized, the use of these texts privileges the knowledges and literacies of dominant cultures, constructing them as a natural, homogenized way of thinking, being and doing.

McGraw and van Leent (2018) examination of the Australian Curriculum for Senior Secondary English, for instance, used thematic analysis to establish the extent to which diverse sexualities were represented within these texts. While the majority of texts did not were not readily apparent as representing LGBTIQ+ characters, they showed how books such as Fitzgerald’s *The Great Gatsby* and Shakespeare’s *Twelfth Night* can be taken up in ways that resist heteronormative assumptions about protagonists. As the authors point out, both text choice, as well as how teachers engage with texts, are important in bringing about justice for LGBTIQ+ diverse students, families and teachers. Text selection, and the visibility of marginalized groups such as LGBTIQ+ students and families has also been examined in early and middle years school contexts, showing how even young readers and viewers can and must be visible for justice to be achieved.

In practice, pivoting toward recognitive justice in literacy has sometimes proven to be difficult; first, because there is a risk that tokenistic inclusion of minority views is sufficient to constitute recognitive justice; and second, because hegemonic views about history, culture and ideology can be deeply entrenched in policy and media discourses. Fraser (2005) also draws attention to the risks of recognitive justice in her writing. Describing the politics of identity and the struggle for recognition, she points out that “routes such struggles take often serve not to promote respectful interaction within increasingly multicultural contexts, but to drastically simplify and reify group identities. They tend, rather, to encourage separatism, intolerance and chauvinism, patriarchy and authoritarianism” (p. 108). Calling this the problem of reification, Fraser explained how this approach toward recognition does little to dismantle structural injustice.

Adopting a critical stance that challenges official texts can draw attention to the ways in which official knowledge is implicated in power, justice and the silencing of marginalized groups, such as those within the LGBTIQ+ community. Other scholars such as Jeffries (2019) have examined how invisible norms of heteronormativity are passively maintained when texts such as the film *Gayby Baby* (Mars and Newell, 2015) are excluded in order to meet the demands of conservative stakeholders. *Gayby Baby* is a documentary film that follows four children with same-sex parents that was produced as a resource for teachers, yet led to a “frenzy” (Jeffries, p. 80) and ultimately, a ministerial decision to prevent a planned showing of the film at an Australian school. Links to Fraser’s (2005) problem of displacement, in which struggles for recognition can lead to further marginalization are apparent, with Jeffries’ work showing the complexities of critical literacy. Jeffries’ examination of the conservative commentary and discourses that were unleashed following the inclusion of *Gayby Baby* in a school shows how heteronormative ideals are imbued in school literacies, and how questions of recognition can further marginalize.

While access to officially sanctioned forms of literacy through the redistribution of economic capital has been assumed to lead to more socially just outcomes, the theoretical contributions of Pierre Bourdieu have been used to illustrate how cultural and social capitals also reproduce inequality. Bourdieu’s concept of linguistic capital (an embodied form of cultural capital) illustrates how literate practices and knowledges are acquired through socialization in families and communities, marking students and families as either part of the dominant class, or not. Because schools and education systems value some forms of linguistic capital over others (typically, assigning a symbolic power to dominant, middle-class language), significant educative and symbolic benefits can be accrued by those who possess the desired linguistic capital. Bernstein’s concepts of elaborated and restricted codes similarly explain how literacies are transmitted through families, which means children of working class families who possess different literate practices to those valued by the middle and upper class are subject to ongoing injustices.

While literacy is often assumed to be an equalizer that has the power to reduce inequality, these theorizations have been used by education researchers to show how literate practices therefore signal group membership (for instance, being the child of a working-class family), and can perpetuate exclusion, inclusion; domination and submission in schools. Critical literacy inquiry has been

taken up by teachers as they work with students from diverse backgrounds, including those for whom Standard English is not their first language or dialog. The theoretical work of Bourdieu has been used to highlight how the distribution of knowledge has been restricted by unequal access to educational opportunities. Take for example Souto-Manning's (2009) work with a young student who spoke using African-American Vernacular English [AAVE], for instance showed how the focus in early years' classes on teaching and communicating in Standard American English led to frustration and exclusion from learning. Critical literacies offer an alternate approach in which teachers and students are encouraged to trouble official knowledges, and to consider how these knowledges work to authorize the knowledge of some students and communities over others.

Luke (2018) explains that redistributive justice, as it relates to "fair and equitable access to dominant knowledge structures and conventional literate practices that are requisite for further and higher education, and for sustainable economic and social participation" (p. 11) requires an engagement with the Audre Lorde's concept of the "master's tools" (p. 11). Lorde (1984) writes that "the master's tools will never dismantle the master's house. They may allow us temporarily to beat him [sic] at his own game, but they will never enable us to bring about genuine change" (p. 112). That is, the tools that have been used to oppress, such as access to hegemonic, canonical knowledge, can be used to further critique and dismantle the power relations that have oppressed and pathologized cultural minorities.

Representation: the "how" of curriculum justice

Reframing teachers from being "experts" who are equipped with knowledge about literacy instruction that precludes students or communities from having an input into what is learned, or whether important and locally relevant questions of power and injustice are taken up in the discussion of texts. Fraser (2009) asks, "By which criteria or decision procedure should one resolve disputes about the "what" and the "who?" (p. 54). Teachers who have taken up critical literacies have adopted an explicit stance that takes democratic practice into account in their classrooms.

The final aspect of Fraser's three-dimensional model of social justice is representation. This element of social justice relates to procedures of representation and community membership and governance and the related injustice is misrepresentation (Fraser, 2008b). Fraser (2008b) associates representation with "social belonging, and thus determining who counts as a member, the *political* dimension of justice specifies the reach of those other dimensions: it tells us who is included in, and who is excluded from, the circle of those entitled to a just distribution and reciprocal recognition" (p. 278). Therefore, representation contributes to the politics of belonging, giving clues as to how power functions. Critical literacy that tackles issues of representation therefore relates to an examination of who has a voice equal to others, in schools, public debates, the media and who does not. Fraser's point that misrepresentation is characterized by political decisions and boundaries that prevent participation in governance and decision making. In a classroom, for instance, the exclusion of Indigenous elders and community from making decisions about literacy has been taken up through participatory teacher-researchers.

For Fraser (2008b), institutional and social practices and purposes must be examined in ways that question taken-for-granted assumptions and challenge the "deep grammar" (p. 23) of justice. Her principle of participatory parity requires that "all interlocutors must, in principle, enjoy roughly equal chances to state their views, place issues on the agenda, question the tacit and explicit assumptions of others ... and generally receive a fair hearing" (Fraser, 2009, pp. 93–94). For instance, while decisions about literacy curriculum, pedagogy and assessment are often framed as being the responsibility of literacy "experts," policy makers and educators, taking a critical literacy stance means questioning these standardized formations. Comber's (2016) work with teacher-researchers in Australia illustrate how these shifts in power-sharing can work in a classroom. Teacher Helen Grant's work alongside an artist-in-residence and recently-arrived refugee students in her class in to provide students with meaningful opportunities to make decisions in how they approached a literacy task. The students ultimately created a powerful film called Sudan that repositioned students from lacking the literacies typically considered essential to succeed in school (for instance, proficiency in Standard Australian English), to experts with significant prior knowledge. This approach of sharing decisions, and teachers learning from and alongside students and communities has been documented by others such as Vivian Vasquez (2016). Developing a critical inquiry stance and a curiosity about issues of justice requires teachers to think about their own privileges and blind spots in their approach toward the teaching of literacy.

As Apple (2014) notes, as he reflects on the social construction of knowledge, and who controls knowledge:

Take schools, for example. For some groups of people, schooling is seen as a vast engine of democracy: opening horizons, ensuring mobility, and so on. For others, the reality of schooling is strikingly different. It is seen as a form of social control, or, perhaps, as the embodiment of cultural dangers, institutions whose curricula and teaching practices threaten the moral universe of the students who attend them (p. 46).

A tool for action: students and teachers as social-justice advocates and activists

Literacy has been the focus of political attention in many countries across the globe over recent decades, with convergence of education policies that have prioritized basic skills literacy pedagogy, curriculum and standardized assessment. The assumption that literacy is inherently apolitical, serves to perpetuate hegemony by privileging the knowledges, practices, languages and values of dominant groups at the expense of marginalized groups. The rise of neoliberal practices that use literacy data collected through

high-stakes, large-scale testing has put teachers, schools and education authorities under significant pressure to orient their attention toward the literacy “that counts.” While this meritocratic approach is often touted as a means for ensuring fairness, through ensuring all students are tested using standardized instruments, the result has often been a standardization of texts, and pedagogies that do not engage with critical questions. In contrast, critical literacy provides all students with an opportunity to engage in substantive conversations about a range of issues—such as racism, homophobia, sexism, ableism and climate change—and to consider their own role in bringing about socially just futures across the dimensions of redistribution, recognition and representation.

Teachers working toward social justice by adopting a critical literacy stance can not only be active in their own work, but also, can also spark children and young people’s sense of activism. Linter’s (2010) work using children’s literature to teach critical geographic awareness is an instructive example. His work shows how critical textual analysis can be used to develop students’ understandings of human impacts of the environment, and how these understandings can be furthered through ongoing discussion of “meta-issues” (p. 18) such as climate change. He suggests that it is through these discussions that students might be prompted to “think critically about themselves and their world and the relationship between the two. It also encourages them to ask, ‘What can I do about it?’” (p. 18).

The use of theoretical frameworks that underpin critical literacy, such as queer theory, decolonization and critical race theories, or by drawing on the scholarship of Freire, Luke, Janks and others has increasingly been used to encourage both teachers and students to think beyond the boundary of the classroom, and to develop a critical consciousness toward justice. Summer Melody Pennell’s (2019) use of Janks’ (2000) work that suggests critical literacy take into account four domains (domination; access; diversity; and design) (pp. 176–177) is a useful example. Her work shows how middle-school students can become intentional in their analysis and understanding of events such as the 2014 shooting of an unarmed, Black teenager Michael Brown at the hands of a White police officer. Following the work of three activists who founded the Black Lives Matter Movement gave both teachers and students time to be intentional in reflecting on and analyzing textual accounts such as media reports. Ultimately, Pennell (2019) reports that they “became better allies for racial justice” by having the opportunity to talk about and understand racism, even if they had not experienced it themselves. She notes that, “This unit helped meet our goals of increasing student awareness of social justice issues and encouraging further learning and action. For our young students, this action took the form of speaking to others in the school community and paying closer attention to news stories they watched with their families. Social justice is a constant commitment and practice, one that we worked to foster for young learners through teaching about the tragic death of Michael Brown” (p. 75). In this way, critical literacy became a tool for both teachers and students to learn about the experiences of others, and to become better allies and advocates for social justice. Given the increasing demands created by standardized curricula and assessment, finding space for teachers to be active and autonomous in working toward social justice aims is important.

Conclusion

Given the rise in education policies in many countries that have prioritized standardization of curriculum, assessment and pedagogy, many primary and secondary teachers have looked to critical literacy as a means of bringing about social justice through the act of teaching. In an interview conducted in 2017, Luke said that “Over the long term, my sense is that critical approaches to curriculum are embodied in many teachers’ dispositions—that is, once they and their students are engaged with critical, inquiry-based, dialogic and constructivist work around the formation of knowledge and truth, text and discourse, writing and image, field and discipline, there’s no going back” (Luke, 2018, p. 302).

Using Fraser’s model allows educators to understand injustice as multidimensional so they might reflect on how they are addressing different families of justice claims through critical literacy. The political, cultural and material events of recent years has created a sense of urgency for change, as evidenced by global movements that seek to address colonial and racial injustice, economic justice, justice for women and LGBTIQ+ people. However, the enormity of political upheaval has also left many feeling despondent. We therefore end with the words of prominent scholar, Allan Luke, as he reflects on the evolution of critical literacy, and advises us that:

... consider what our generation experienced in the aftermath of 1968, when our hopes were dashed, friends and colleagues jailed and leaders murdered, while violence raged on in Vietnam and on city streets. This is not a time for despair. This is your generational challenge as educators, scholars and activists. It is game on.

Luke (2018, p. 19).

When social justice is framed as a multidimensional concept, simple solutions to injustice are not plausible. Social transformation and a recalibration of how privilege is understood and dismantled can occur through a commitment to critical literacies that retain a strong focus on multiple aspects of justice—representation, redistribution and representation.

References

- Anyon, J., 1981. Social class and school knowledge. *Curric. Inq.* 11, 3–42.
- Apple, M.W., 2014. *Official Knowledge: Democratic Education in a Conservative Age*. Routledge, New York.
- Ariès, P., 1962. *Centuries of Childhood: A Social History of Family Life*. Knopf, London.
- Comber, B., 2016. *Literacy, Place, and Pedagogies of Possibility*. Routledge, New York.
- Cooke, N.A., 2018. Critical literacy as an approach to combating cultural misinformation/disinformation on the Internet. In: Agosto, D.E. (Ed.), *Information Literacy and Libraries in the Age of Fake News*. Libraries Unlimited, Santa Barbara, CA, pp. 36–51.
- Fraser, N., 1997. Justice Interruptus: Critical Reflections on the "Postsocialist" Condition. Routledge, London.
- Fraser, N., 2003. Social Justice in the Age of Identity Politics: Redistribution, Recognition, and Participation. In: Fraser, N., Honneth, A. (Eds.), *Redistribution or Recognition? A Political-Philosophical Exchange*. Verso, New York, pp. 7–109.
- Fraser, N., 2005. Reframing Justice in a Globalizing World. *New Left Rev* 36, 69–88.
- Fraser, N., 2008a. Abnormal justice. *Crit. Inq.* 34, 393–422.
- Fraser, N., 2008b. *Scales of Justice: Reimagining Political Space in a Globalizing World*. Columbia University Press, Boston.
- Fraser, N., 2009. *Scales of justice: Reimagining political space in a globalizing world*. Columbia University Press, New York.
- Freire, P., Macedo, D.P., 2004. *Literacy: Reading the Word & the World*. Routledge and Kegan Paul, London.
- Freire, P., 1974. *Pedagogy of the Oppressed*. Herder & Herder, New York.
- Gadsby, H., 2018. *Nanette* [Netflix special].
- Gadsby, H., 2020. *Douglas*. [Netflix Australia].
- Hoffman, J.V., Alvermann, D.E., 2020. What a genealogical analysis of Nila Banton Smith's American reading instruction reveals about the present through the past. *Read. Res. Q.* 55, 251–269.
- Janks, H. (Ed.), 1993. *Critical Language Awareness Series*. Wits University Press and Hodder and Stroughton, Johannesburg.
- Janks, H., 2000. Domination, access, diversity and design: a synthesis for critical literacy education. *Educ. Rev.* 52, 175–186.
- Jeffries, M., 2019. Examining media discourses of diversity and "indoctrination": public perceptions of the intended screening of *Gayby Baby* in schools. In: Baroutsis, A., Riddle, S., Thomson, P. (Eds.), *Education Research and the Media: Challenges and Possibilities*. Routledge, New York, pp. 80–96.
- Lintner, T., 2010. Using children's literature to promote critical geographic awareness in elementary classrooms. *Soc. Stud.* 101, 17–21.
- Lorde, A., 1984. *Sister Outsider* (Original Work Published 1979). Crossing Press, Berkeley.
- Luke, A., 2010. Documenting reproduction and inequality: revisiting Jean Anyon's "social class and school knowledge". *Curric. Inq.* 40, 167–182.
- Luke, A., 2018. *Critical Literacy, Schooling, and Social Justice: The Selected Works of Allan Luke*. Taylor and Francis, London.
- Mars, C. (Producer), and Newell, M. (Director), *Gayby Baby*, 2015. Marla House, Australia.
- McGraw, K., van Leent, L., 2018. Textual constraints: queering the senior English text list in the Australian curriculum. *Engl. Aust.* 53, 28–39.
- Oakes, J., 2005. *Keeping Track: How Schools Structure Inequality*, second ed. Yale University Press, London.
- Pennell, S.M., 2019. *Queering Critical Literacy and Numeracy for Social Justice: Navigating the Course*. Palgrave Macmillan, London.
- Postman, N., 1985. The disappearance of childhood. *Child Educ.* 61, 286–293.
- Smith, L.T., 1999. *Decolonizing Methodologies*. Otago University Press, Otago.
- Souto-Manning, M., 2009. Acting out and talking back: negotiating discourses in American early educational settings. *Early Child Dev. Care* 179, 1083–1094.
- Takayama, K., 2009. Progressive struggle and critical education scholarship in Japan toward the democratization of critical education studies. In: Apple, M.W., Au, W., Gandin, L.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, pp. 354–368.
- The New London Group, 1996. *A pedagogy of multiliteracies: designing social futures*. *Harv. Educ. Rev.* 66, 60–93.
- Tuck, E., Yang, K.W., 2014. R-words: refusing research. In: Paris, D., Winn, M.T. (Eds.), *Humanizing Research: Decolonizing Qualitative Inquiry With Youth and Communities*. Sage Publications, Thousand Oaks, pp. 223–248.
- Vasquez, V.M., 2016. *Critical Literacy Across the K-6 Curriculum*. Taylor & Francis, New York.
- Yan, G.C., Chang, Y., 2009. The circumstances and the possibilities of critical educational studies in China. In: Apple, M.W., Au, W., Gandin, L.A. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, pp. 378–396.

Further reading

- Apple, M.W., 2012. *Knowledge, Power, and Education: The Selected Works of Michael W. Apple*. Routledge.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Routledge & Kegan Paul, London.
- Fitzgerald, F.S., 1925. *The Great Gatsby*. Scribner's sons, New York.
- Janks, H., 2005. Deconstruction and reconstruction: diversity as a productive resource. *Discourse* 26, 31–43. <https://doi.org/10.1080/01596300500040078>.
- Janks, H., 2013. Critical literacy in teaching and research. *Educ. Inq.* 4, 225–242. <https://doi.org/10.3402/edui.v4i2.22071>.
- Jeffries, M., 2019. "Doing" gender differently: exposing the porous nature of gender norms through children's literature. *Jeun. Young People Texts, Cult.* 11, 192–203.
- Pennell, S.M., 2019. Reading representations of race: critical literacy and Ferguson. *Engl. J.* 108, 68–75.
- Shakespeare, W., 1928. *Twelfth Night*. Houghton Mifflin, New York.
- Shay, M., Wickes, J., 2017. Aboriginal identity in education settings: privileging our stories as a way of deconstructing the past and re-imagining the future. *Aust. Educ. Res.* 44, 107–122.
- Vasquez, V., 2007. Using the everyday to engage in critical literacy with young children. *N. Engl. Read. Assoc. J.* 43, 6–11.

Relevant websites

- <https://newlearningonline.com/>, Multiliteracies and Critical Literacies Website by Mary Kalantzis and Bill Cope.
- <http://thegaybyproject.com/>, *Gayby Baby* Film Information.
- <https://thelearningexchange.ca/videos/allan-luke-critical-literacy/>, Allan Luke Discusses Critical Literacy.

Critical literacies and social justice

Ernest Morrell^a and Jodene Morrell^b, ^a Department of English, The University of Notre Dame, Notre Dame, IN, United States; and
^b Notre Dame Center for Literacy Education, The University of Notre Dame, Notre Dame, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

A Counterstory of Critical Literacies, Social and Cultural Justice	225
How I found critical literacy studies	225
“The two biggest problems facing the black community are too much poverty and too little self love”- Cornel West	225
Toward a critical pedagogy of English language, literature, and literacy	227
Powerful literacy education today and tomorrow	228
Critical approaches to teaching the Africanist presence in literature	228
Exploring race and justice in America through literature and popular culture	229
Critical hip-hop and literacy studies	230
Critical media pedagogies	231
YPAR and critical literacies for social action	232
Toward a praxis of universal critical literacy and revolutionary love	233
References	233

A Counterstory of Critical Literacies, Social and Cultural Justice

We have colloquially titled this entry *A Counterstory of Critical Literacies, Social and Cultural Justice* and, with the help of Dr. Jodene Morrell, I have attempted to weave together a narrative of 30 years of life as a student, a scholar, a teacher, and a writer of critical literacies for social and cultural justice across our most historically vulnerable populations (the remainder of the article will use the personal pronoun “I”). We are indebted to the editors of this volume for encouraging us to draw on our autobiographies as we engage key concepts in the critical discourse of global education. Far too often for African descended postcolonial subjects, our personal voice, our “I”, is muted under a guise of distance and objectivity. This distancing of the subject from the production of knowledge is not consistent with the critical literacy tradition I am charged to discuss. It feels far more authentic to write about critical literacy studies in the tradition of personal voice and agency.

As I grow older, I find it increasingly difficult (and unethical) to separate my critical autobiography from my critical scholarship. I am the work that I do and I am who I am because of the words I have read and written and because of the students I have taught and loved and learned so much from and the academic and social movements I have been privileged to participate in over these past 29 years. In the spirit of the critical race methodological tradition of counter storytelling (Crenshaw et al., 1995; Bell, 1992; Solorzano and Yosso, 2002), this article attempts to narrate the biography of my research and praxis in critical literacies and social justice as a microcosm of larger scholarly and social movements that I hope to illuminate as well.

How I found critical literacy studies

Why Critical Literacy Studies? The first time I pondered this question was almost 30 years ago. I was a junior in the University of California system dropping my Economics courses and trading them in for our English core and a fat *Norton Anthology of British Literature*, one I still possess. I still find myself asking this question of why. Why would one exit a promising career in finance? Why study and teach languages that we speak naturally from birth? Why join a discipline that, at times, pretends that we members of the African Diaspora have nothing to contribute? And why, in a time of global competition and technological advancement, should anyone be left alone to stare at the clouds and contemplate Keats and Tennyson when they could use their time far more productively? As a member of the presidential team of the National Council of Teachers of English (NCTE) from 2011 to 2015, I was asked these and similar questions directly from staffers, senators, and congressional representatives as I sat in their plush DC offices on our annual Advocacy Day.

“The two biggest problems facing the black community are too much poverty and too little self love”- Cornel West

While there are many places from which I could begin to formulate a response, in preparing this text I was drawn to a quote from Cornel West, an American philosopher I discovered at this very moment in my life. The day was April 29, 1992 and, after an American jury found the officers videotaped beating Rodney King not guilty, Los Angeles was burning (and not Ferguson, Louisville, or

Minneapolis of the 2020 movement) and I had just turned 21, having met Jodene (we have been married now 26 years) only 5 days prior. I and my Southern California classmates dropped everything and took to the streets. For three long days and nights.

I returned to my coursework different, somewhat liberated, exhausted, somewhat devastated and in search of a different pathway forward. I turned to a cassette tape sent to me by a friend of a talk delivered by Professor West to students at Bowdoin where he began with the quote above. The talk, essentially a precursor to his now much acclaimed book *Race Matters* (1993), pointed me to pedagogy and literacy studies and away from the JD-MBA applications that adorned my Southern California dorm room. For, in critical literacy, I found my own pathway toward self love and meaningful cultural production. After 30 years, I am still here, and I am happily here. At the risk of sounding sacrilegious I must say that it was critical literacy studies that saved me. I am here because I continue to place critical literacy at the forefront of a pedagogical praxis that can intervene in the lack of self-knowledge, and self love, demanded for active citizenship.

My research and pedagogical philosophy begin with an understanding that any empowering literary or literacy practices must affirm all people as fully human and draw upon the literary and cultural artifacts that most profoundly reveal their subjective existential experiences and their historical and material realities. Toward these ends I have examined the literary, linguistic, and literate practices of ethnic minorities and other historically marginalized groups, particularly as they relate to participation in counter-hegemonic literary production, youth popular culture, and local community life. Advocating for the intellectual merit of these oft hidden or degraded language and literacy practices and the cultural traditions in which they are situated, I have investigated transformative literacy pedagogies inside of literacy classrooms and in out-of-school spaces that facilitate textual empowerment, literary production, and academic engagement.

1. I am interested in how literacy is taught, how it is engaged; the work that teachers and students do together in literacy classrooms, especially work that promotes self love and social justice.
2. I am interested in the literature that is written for and consumed by children and adolescents. I am interested in how this literature is read and reacted to and how these readings and reactions constitute a praxis of critical literacy
3. I have an interest in the African Diaspora as an imagined community—its literature, its popular culture; its engagement; its critique, its resistance, its plurality; its cacophonous polyvocality; its beauty, its power, its joy—its agency, its undeterred presence. I'm also interested in making the African Diaspora knowable and tangible for present and future generations
4. As a doctoral student attending the University of California at Berkeley in the 1990s, I studied the cultural theory of Raymond Williams, Stuart Hall (Hall, 1997), and John Storey and became interested in popular culture as a contested space; a simultaneous space of resistance and social construction. How do youth navigate the discourse of popular culture? And what are the core tenets of what Doug Kellner might call a critical media pedagogy? How are youth taught to inhabit media culture and how do they use media to educate others?
5. I am also interested in how nondominant communities acquire and utilize critical literacies for social action.

Methodologically my work has drawn upon a New Literacy Studies-inspired ethnography of communication and what Norman Fairclough terms a critical discourse analysis. Linguistic anthropologist Shirley Bryce Heath's *Ways With Words* (1983) explores the literate lives of working-class African Americans and whites in the Piedmont Carolinas. Heath analyzes literacy events, which she defined as any occasion in which a piece of writing is integral to the nature of participants' interactions and their interpretive processes. Subsequent theorists in the New Literacy Studies (Brian Street, David Barton, Mary Hamilton) have also advocated the study of the literacy events of members of marginalized subcultures. Barton and Hamilton (2000), for instance argue that it is the work of cultural theorists and ethnographers in the NLS tradition to derive new literacy practices from the study of contemporary (often digitally mediated) literacy events. These scholars have informed my own work even as I push upon the constructs of literacy. The conceptualization of the term, I argue, is central to the work of the student of culture and discourse. We cannot excavate or document what we are not trained to look for. In this vein I have worked to expand the notion of literacy events to accommodate traditional textual practices as well as those that accompany participation in youth popular culture and new media literacies.

Employing an NLS framework, I have conducted a series of ethnographic studies exploring the possible connections between criticalist pedagogies of popular culture and the development of academic and critical literacies in classrooms serving historically vulnerable student populations. For example, I have conducted studies that investigate the potential of drawing upon youth engagement with hip-hop culture to teach critical reading, literary analysis, and the production of expository and creative writing. I have investigated the uses of court trials to teach argumentative writing in secondary English classrooms. I have explored the uses of film critique as a bridge to literary analysis, and I have investigated how critical media education can empower youth as consumers and producers of media. For example, I have documented the ways that young people can be taught to decipher negative and stereotypical images in popular media and I have also documented the process involved in young people becoming producers of short documentary films covering issues of inequality in their schools and communities. Much of my early work along these lines was published in my second book, *Linking Literacy and Popular Culture* (Morrell, 2004).

I have endeavored in my scholarship to articulate a grounded theory of critical literacy education that is culturally and socially responsive to the lives of today's youth, that develops academic literacies and facilitates academic achievement, and that is rooted in the systematic inquiry of classroom practice. My work has targeted diverse classrooms in historically underserved schools, where students' inabilities to access academic content leads to severe educational, social, and political outcomes. Toward these ends, my research has demonstrated that innovative pedagogical practices can increase student engagement and increase youth academic literacy development and academic achievement.

I have been interested in life in formal literacy classrooms and, particularly, how African Diaspora literatures and popular cultural artifacts are appropriated inside of literacy classrooms as sources of engagement, critical consciousness, and social action. These include more traditional texts such as novels, poems, and plays, but I have also been interested in the uses of popular culture (i.e., critical hip-hop, spoken word poetry, film, television, corporate and social media outlets) (Morrell et al., 2013).

I have also been interested in the textual and cultural production inside of literacy classrooms as students participate in community-centered research and outreach projects. I conducted a series of critical ethnographies where I followed high school and college students over a 12-year period from 1999 to 2011. I was interested in how students from historically nondominant backgrounds were able to utilize traditional and multimodal genres to distribute academic research to multiple audiences. These studies, in the tradition of the ethnography of communication, documented the literacy events over a multi-year trajectory as they related to textual consumption, research reports, essays, critical memoirs, spoken word poems, hip-hop songs, a hip-hop opera, plays, and documentary scripts (Morrell, 2008a,b; Mirra et al., 2016).

Critical Literacy and Urban Youth: Pedagogies of Access, Dissent, and Liberation (Morrell, 2008a,b) offers an interrogation of critical theory developed from my work with young people in classrooms, neighborhoods, and institutions of power. Through cases, an articulated process, and a theory of literacy education and social change, I attempt to extend the conversation among literacy educators about what constitutes critical literacy while also examining implications for practice in secondary and postsecondary American educational contexts. This book is distinguished by its weaving together of theory and practice. I begin by arguing for a broader definition of the “critical” in critical literacy—one that encapsulates the entire Western philosophical tradition as well as several important “Othered” traditions ranging from Postcolonialism to the African-American tradition. Next, I look at four cases of critical literacy pedagogy with urban youth: teaching popular culture in a high school English classroom; conducting community-based critical research; engaging in cyber-activism; and doing critical media literacy education. Lastly, I return to theory, first considering two areas of critical literacy pedagogy that are still relatively unexplored: the importance of critical reading and writing in constituting and reconstituting the self, and critical writing that is not just about coming to a critical understanding of the world but that plays an explicit and self-referential role in changing the world. I conclude by outlining a grounded theory of critical literacy pedagogy and considering its implications for literacy research, teacher education, classroom practice, and advocacy work for social change.

I have incorporated African-Diaspora popular cultural studies and postcolonial studies into my conceptual framework for critical literacy pedagogy and praxis. One of the criticisms of critical pedagogy is that, in focusing on issues of class and social reproduction, it ignores or downplays the significance of race to experiences in America’s schools. Any examination of academic achievement globally reveals gross inequities in achievement for members of different racial groups, even when class is constant. My work has drawn upon African Diaspora Studies to theorize pedagogies of ethnic empowerment as well as academic achievement. For instance, in a chapter published in the book *Teach Freedom: Education for Liberation in the African-American Tradition*, I examined the role of rebel ethnic music in promoting universal critical literacy, paying particular attention to the relationships between critical hip-hop and roots reggae music and the literacy practices of African-Americans and Afro-Caribbeans (Morrell, 2008a,b). I have also examined the possible intersections between postcolonial theory and the development of teachers across the P-16 pipeline who simultaneously prepare their students for the worlds of professional employment and engaged citizenship and cultural production for racial and social justice in the Americas.

Toward a critical pedagogy of English language, literature, and literacy

I join a chorus of scholars like Carol D. Lee, Robert Scholes (Scholes, 1998), Geneva Smitherman (Smitherman, 1999), Keith Gilyard, Gerald Graff, Anne Gere, and Arthur Applebee who have been strong proponents of an increased critical pedagogical and critical literacy focus in English studies. These scholars span the Anglophone world from the United Kingdom and Canada to Australia, New Zealand, Jamaica and the Anglophone Caribbean to the United States. My ideas about the critical pedagogy of English have been largely formed by Paulo Freire (1970) and other adherents of what we have been calling critical pedagogy such as Antonia Darder (Darder, 1991), Peter McLaren (McLaren, 1989), Carlos Torres (Torres, 1998), and Angela Valenzuela. Paulo Freire, a Brazilian literacy educator and education advisor to the World Council of Churches believed that education held the potential to enhance conditions of either oppression or liberation. In *Pedagogy of the Oppressed*, written while in exile at Harvard in the late 1960s, Freire identified banking education as the primary method by which educational institutions maintain oppressive ideologies. According to his banking metaphor, schools or educational institutions treat students as empty repositories waiting to be filled by institutionally sanctioned knowledge, even if possessing that knowledge was detrimental to students’ sense of self.

Freire argued instead for a problem-posing education, an education rooted in the existential experiences of the people, in which they explored their own immediate surroundings and, in an attempt to become more present in understanding and reshaping their own reality, develop important academic and critical literacy skills.

In *The Art of Critical Pedagogy: Possibilities of Moving from Theory to Practice in Urban Schools* my co-author Jeff Duncan-Andrade and I (Duncan-Andrade and Morrell, 2008) explore critical pedagogy and its practical applications for the Teaching of English. It addresses two looming, yet under-explored questions that have emerged with the ascendancy of critical pedagogy in the educational discourse: (1) What does critical pedagogy look like in urban secondary English classrooms? and (2) How can a systematic investigation of critical work enacted in urban contexts simultaneously draw upon and push the core tenets of critical pedagogy? Addressing the tensions inherent in enacting critical pedagogy—between working to disrupt and to successfully navigate oppressive

institutionalized structures,—*The Art of Critical Pedagogy* seeks to generate authentic internal and external dialogs among English educators in search of literary texts and classroom practices that contribute toward a more socially just world.

My current manuscript, *Critical Pedagogy and Cultural Studies in Literacy Education*, considers the potential of conceptual and empirical work in critical pedagogy and cultural studies to inform, confront and transform many of the persistent challenges we presently face in literacy classrooms. The book begins with an examination of the historical antecedents of critical pedagogy, both from the Western philosophical tradition and “Othered” traditions such as the African-Diaspora and Latin American traditions and Postcolonialism. The book will then examine the theory and research of critical pedagogists such as Paulo Freire, Sonia Nieto (Nieto, 1992), Carlos Torres, Peter McLaren, Henry Giroux (Giroux, 1992), Antonia Darder, and bell hooks (hooks, 1994). The second half of the book focuses on cultural studies and, in particular, the critical uses of popular culture in secondary and college English classrooms. Examples include critical hip-hop and spoken word poetry, film, television, mass media consumption and production, youth participatory action and their implications for transformative work with the post-millennial generation.

I am joined in this work by scholars such as H. Samy Alim, Maisha Winn, Django Paris, Michele Knobel, Colin Lankshear, Alan Luke, Carmen Luke, Donna Alvermann, Kris Gutierrez, Mimi Ito, Jennifer Rowsell and many others who have shown the brilliant ways that youth engage popular culture as well as the exciting academic and social outcomes that accompany critical literacy pedagogies that engage digital media and youth popular culture. These scholars come from diverse fields like critical linguistics, cultural anthropology, the learning sciences, philosophy, ethnic studies and the educational sciences to explore the language of hip hop (Alim, 2006), sociocritical literacies (Gutierrez, 2008), overcoming digital divides (Morrell and Rowsell, 2020), spoken word poetry in city classrooms (Fisher, 2007), and everyday teen life in the era of digital media (Ito et al., 2019).

Powerful literacy education today and tomorrow

How do contemporary critical literacy scholars and educators build upon the amazing work that has come before? What do we need to risk today to make possible 50 years from now a movement for powerful literacy education? I like to think of beginning with an assessment of who we are at present. When do we begin to recognize, as a field, that we are a multiliterate, multilingual planet, and that one cannot extricate the political enterprise from the formal teaching of language and literacy? If we are not for social justice then we are for social injustice. There is no in-between.

What do we need to do to make tomorrow’s critical literacy education the norm today and not the exception in our classrooms? We need to start a new movement. Children’s Defense Fund President Marian Wright Edelman has often said, “If you don’t like the way the world is, you change it. You have an obligation to change it. Just do it one step at a time” and “Stop talking about what’s not possible.” Nobody has time for us to twiddle our thumbs and be hopeless. So if we don’t like the way things are, we have to become agents of change.

Our work becomes critical when we share it internally through our books, journals and our conference presentations. It becomes critical when we share it through online chats, through social media, on podiums, as part of activism. The best-kept secret in literacy is the daily genius in our primary and secondary classrooms across the globe. What does that genius look like?

Critical approaches to teaching the Africanist presence in literature

In this here place, we flesh; flesh that weeps, laughs; flesh that dances on bare feet in grass. Love it. Love it hard. Yonder they do not love your flesh. They despise it. They don’t love your eyes; they’d just as soon pick em out. No more do they love the skin on your back. Yonder they flay it. And O my people they do not love your hands. Those they only use, tie, bind, chop off and leave empty. Love your hands! Love them. Raise them up and kiss them. Touch others with them, pat them together, stroke them on your face “cause they don’t love that either. You got to love it, you! And no, they ain’t in love with your mouth. Yonder, out there, they will see it broken and break it again. What you say out of it they will not heed. What you scream from it they do not hear. What you put into it to nourish your body they will snatch away and give you leavins instead. No, they don’t love your mouth. You got to love it. This is flesh I’m talking about here. Flesh that needs to be loved. Feet that need to rest and to dance; backs that need support; shoulders that need arms, strong arms I’m telling you. And O my people, out yonder, hear me, they do not love your neck unnoosed and straight. So love your neck; put a hand on it, grace it, stroke it and hold it up. and all your inside parts that they’d just as soon slop for hogs, you got to love them. The dark, dark liver—love it, love it and the beat and beating heart, love that too. More than eyes or feet. More than lungs that have yet to draw free air. More than your life-holding womb ... hear me now, love your heart. For this is the prize.”

Toni Morrison (1987). (Morrison, 1987)

In *Playing in the Dark: Whiteness and the literary imagination* Nobel Laureate Morrison (1993) challenges us to stop averting our gaze to the Africanist presence that saturates American literature. To engage this presence, she argues, should be an act of pure joy for all. Her work, she testifies, is an act of love and joy. It is my hope, my dream that we all learn to love the Diaspora as much as Morrison challenges us to. That regardless of our immediate origins, we see in the eyes, the necks, the hearts and the souls of the Diaspora our own selves, our own pasts, our own possibilities. The Diaspora is big enough for us all! Morrison’s call is echoed by scholars in multicultural education (Jim Banks, Gloria Ladson-Billings, Sonia Nieto, Carl Banks, Christine Sleeter), critical race studies in education (Solorzano and Yosso, Ladson-Billings and Tate) and critical linguistics (Ofelia Garcia, Geneva Smitherman)

to make language and literacy instruction more responsive to the heterogeneity of our student bodies around the world. In what follows I hone in on what this might mean for critical literacy studies to more meaningfully engage the African Diaspora.

As profoundly as literacy studies have shaped my academic trajectory over the past 30 years, the African Diaspora has also formed (and transformed) me ontologically as well as intellectually. James (1989) describes the process of revolution as the process of people finding themselves. Part of my ontological development has resulted from my own search for an African Diaspora that was not very present in my formative schooling experiences. The fragments of the African Diaspora that were present in schooling and often in the larger public discourse did not feel familiar. Nor did they feel powerful.

Through my own travels, readings, and connections with scholars throughout the globe, I have discovered an African Diaspora more present and potent than I could have imagined as a schoolchild or as a burgeoning intellectual. Still, I often wonder as I peruse bookshelves, libraries, and university catalogs, why one has to seek so diligently to find true and authentic narratives of a population that collectively represents more than 20% of the planet. How do we ethically tolerate the academic invisibility and educational disinvestment (at every level including, and maybe most profoundly, at the university level)? More importantly, how do we make amends?

Were it a nation the African Diaspora would be the third largest country on earth, behind only China and India, and ahead of the United States. Taken together Africa and the African Diaspora represent 20% of the world's population. According to UNESCO projections, by 2050 40% of all births will occur in sub-Saharan Africa and by 2100 40% of the world's population could live in or hail immediately from Africa (UNESCO, 2017). This represents a significant and fast growing population that we know relatively little about from a scholarly perspective. What we do know is that the continent has the lowest life expectancies and some of the lowest literacy rates on the planet. We also know that its artists, its heroes, its priests and bishops, its educators, its cultural institutions, and its martyrs are largely unknown and understudied by the Western world. Hierarchies of knowledge, legacies of enslavement and exploitation, and academic structures that lean toward Europe have prevented the serious study of the African Continent and African Diaspora that are warranted.

Along with colleagues Kassie Freeman and Henry Levin, we developed the African Diaspora Consortium (ADC) realizing that, if we waited for an invitation to do so, one would never arrive. The African Diaspora consortium was created to impact positively the educational, economic, and artistic outcomes and opportunities of Black populations across the African Diaspora. ADC is unique in that it focuses its attention and develops partnerships in countries where African descendants were historically and culturally dispersed during the transatlantic slave trade and/or different migration periods (e.g., the Caribbean region, Europe, Latin America, and North America). By all accounts, there has been a striking lack of research and best practices devoted to dispersed African descendant populations. Yet, logically and intuitively, better understanding comparatively how dispersed populations navigated their locations in different and new worlds while developing art forms as modes of communication, interconnectedness, and education and economic uplift can hold great importance from an economy of scales perspective for countries globally. Increased understanding of both commonalities and differences in experiences as shaped by different immigrant origins and societal practices of their resident societies across nations can provide useful possibilities and outcomes.

Led by Dr. Kassie Freeman, former president of the Comparative International Education Society, ADC seeks to accomplish three main goals: (1) Identify via cross-disciplinary academic research the historic and cultural connections as well as common challenges and opportunities across the Diaspora; (2) Increase educational success and outcomes for African descendant populations at all levels globally, promoting and exchanging learning successes from partner countries; and (3) Utilize literature and the arts as a vehicle to enhance understanding and connectedness as historical and cultural uplift. Our greatest accomplishment to date is the development of academic content focused on the African Diaspora for the Advanced Placement (or AP) Research Seminar offered by the College Board in the US. The African Diaspora content represents the first offering of its kind by the College Board and is available to thousands of classrooms. The content focuses on colonialism, resistance, and everyday cultural practice and invites students of all backgrounds to draw on primary sources, popular culture to develop their own research and action projects related to the global African Diaspora.

Exploring race and justice in America through literature and popular culture

Lee (1993) set the gold standard for incorporating African-American literature and discourse genres into the formal teaching of English in North American central cities that were becoming increasingly African-American. In my own classroom in Oakland, California I was inspired by Lee and others to incorporate Richard Wright's (Wright, 1940) *Native Son* as a bridge to African American languages and literacies in my secondary English classroom. In work with 120 seniors across four sections of senior English I structured a mock trial where Bigger Thomas, the protagonist, would be put on trial. Richard Wright's *Native Son* was nested inside of a larger multimedia themed curriculum dealing with Race and Justice in American society. In addition to the Wright book students watched "A Time to Kill," they listened to popular music that dealt with issues of race in Imperialist society such as Bob Dylan's Hurricane, Bob Marley's Buffalo Soldiers (Marley and Sporty, 1978), and Lauryn Hill's *The Miseducation of Lauryn Hill* (Hill, 1998). The traditional capstone activity was a two-week mock trial although students also needed to develop a social action project. One of the years this became a community newsletter that used student-produced literature, personal narratives, and qualitative research to promote racial healing and, in another year, the students initiated dialogs of mutual understanding with local law enforcement officials.

I amassed a collection of casebooks across four sections of English over a period of 3 years (approximately 500 casebooks) to determine the connection between the oral arguments presented in the trial setting and the quality of arguments made by students in their written analysis of the trial. In the casebooks the students were asked to include: (1) All the questions prepared for the witnesses of their team and scripted responses; (2) All of the anticipated questions to be asked of their witnesses in cross-examination; (3) All of the questions prepared in advance for opposing witnesses in cross-examination; (4) The opening arguments; (5) All notes taken during the trial (trial notes are mandatory); and (6) A 5–7 page summation/analysis of the trial. The case books averaged between 50 and 60 pages. I was very interested in how the inclusion of oral language in the English classroom, and more specifically oral arguments, translated into robust written arguments. I was also interested in how students' discussions of a popular film like *A Time to Kill* would influence their analysis of a text written in the 1940s. My findings indicated that the volume and quality of students' arguments increased and I began to use similar practices in the College Writing course I was teaching at a local university.

Using Langston Hughes' *Dream Deferred* (Hughes, 1994) as a launching text, I designed a project where high school and college students created youth participatory action research projects in response to the poem. One class decided to study the educational and social disinvestments in the nation's young people, what they described as the killing of dreams. This idea of connecting students' lives to dreams deferred became the focal point of campus movements that ultimately became national social movements that remain relevant today. I can still remember one of my undergraduates at UCLA promising the class that, for her final project, she was going to start a national movement. By the end of the term we had witnessed a series of civil disobedience protests in Senate offices and a response from the president of the United States. Another of my students at UCLA ran into then governor Arnold Schwarzenegger at a Starbucks in Santa Monica and informed him of the work of our class. She invited him to class and he, in true Arnold fashion, accepted the invitation and engaged our undergraduates in an honest and robust dialog around issues of equity and social justice. I have been heartened by the role of the study of literature and contemporary popular culture in facilitating dynamic spaces for critical social thought and counter-textual production; and in my own work I study the connections between these literature-based projects and students' academic literacy learning, their identity development, and their involvement in social action.

Critical hip-hop and literacy studies

Literacy studies should concern itself with the study of hip-hop as a critical postindustrial expression of youth culture throughout the African Diaspora. Several notable scholars who have made these connections include: Maisha Fisher, H. Samy Alim, Marc Lamont Hill, Chris Emdin, Emery Petchauer, Lauren Kelly, and Jamila Lyiscott. My interest in critical hip hop in literacy classrooms dates back to the mid 1990s. While heeding the concerns expressed in George Lipsitz's (Lipsitz, 1997) *Dangerous Crossroads: Postmodernism, Popular Music, and the Poetics of Place* about corporatized popular culture as existing in a dialectic tension between co-optation and postcolonial counterhegemony, I was also influenced by Rose's (1994) classic *Black Noise*, where she theorized critical hip-hop as a youth-initiated critique of urban post-industrialism. This view comported with my own reading of hip-hop as a voice of agency and resistance and, through lyricists I was studying at the time like Lauryn Hill, Rakim, and Chuck D, I argued to my less than enthusiastic dissertation committee that hip-hop represented a social theorizing and critical public pedagogy. My aim was to write a study that focused on the poetics of hip-hop, similar to the work of Imani Perry (Perry, 2004) at Princeton University. However, I was unprepared for how much I would be drawn into the pedagogy of hip-hop through my work with the high school students I worked with every day in an East Oakland high school.

Synthesizing literature from critical pedagogy, sociocultural psychology, and cultural studies with popular cultural texts and experiences from actual English classroom practice, I embarked upon a series of studies that explored the critical teaching of hip-hop and other forms of youth popular culture as a viable strategy to increase academic and critical literacies in urban secondary English classrooms. Relying on scholarship which views youth popular culture as a powerful, but oftentimes underutilized point of intervention for schools, these authors discuss the impact of using youth popular culture to reconnect with otherwise disenfranchised schooling populations. I attempt to rebut criticisms associated with the teaching of popular culture by showing how teachers can simultaneously honor and draw upon the sociocultural practices of their students while also adhering to state and national standards. Further, the work hopefully demonstrates the social relevance, academic worthiness and intellectual merit of hip-hop music and culture.

I have taught hip-hop as a period of poetry; as an inspired youth resistance to urban post-industrialism (Rose, 1994). Alongside other periods of poetry such as Elizabethan, Romantic, Harlem Renaissance, and Civil Rights Poetry student research groups compared and contrasted contemporary hip-hop texts with classic poetry texts. They then led an hour presentation/discussion among their peers about the two texts. Students also wrote a 5–7 page critical interpretations of a song of their choice and an anthology of ten poems and presented their poetry to their peers.

Other teachers and members of the larger society perceived many of the students in my classrooms as functionally illiterate and lacking in intellect. These detractors relied on traditional school-based measures (past school performance, test scores) to support their claims and their resultant behavior toward (more like against) these youth. These assessments, however, ran counter to my observations of our students' sophisticated literacy practices that accompanied their participation in hip-hop culture. The same students portrayed as uninterested in literacy would come into our classrooms capable of reciting from memory the lyrics of entire rap albums. They voraciously read popular hip-hop magazines and transcribed song lyrics, and several carried their own rap

composition books. Even those students who were not so deeply involved with hip-hop culture regularly listened to popular radio stations that programmed predominantly hip-hop and accessed magazines and CDs. To tap into this literacy-rich youth cultural activity, we developed a seven-week poetry unit that paired hip-hop texts with canonical works of poetry. The goal of the unit was to play on students' heavy investments in hip-hop to create deeper understandings of school-based forms of poetry. We aimed to help students see the timelessness of the literary themes present in both canonical texts that they were mostly unfamiliar with and some of the music they listened to daily. We made explicit the fact that the literacy skills and knowledge that they exhibited in their interactions with hip-hop texts were not far removed from the skills required to succeed in analysis of poetry. After pairing eight traditional poetry texts with eight rap songs, we divided our classes into groups of four or five students. Each group was responsible for preparing a class lesson on its pair of works. They were given relative autonomy for the form of their presentation, but each presentation was to have the following three elements:

1. Analysis of the literary themes of each work (we gave students a list of common themes but encouraged them to develop others as they saw them emerge in the work)
2. Comparative analysis of the two works (how are the works similar, how are they different, how do the authors use similar or different devices to deliver their message?)
3. Guiding questions for class discussion (we had modeled complex, open-ended questions throughout the year and encouraged students to design similar questions)

In addition to preparing the group presentation, each student was expected to prepare an analysis of the other seven poem/song pairings to participate in class discussions. We had worked hard all year on developing a classroom culture where student participation in dialogs about literary themes was normalized. This, along with the use of hip-hop texts that drew students' attention, made the prospect of out-of-class preparation more likely, as the class came to enjoy challenging and broadening the literary analysis of their peers. While my work has focused on the uses of hip hop in classroom contexts, others have explored the linguistic dexterity of hip hop (Alim, Lyiscott), hip hop as a source of consciousness and activism (Hill, Kelly), and the uses of hip hop in teacher education (Petchauer, etc.). Indeed, #HipHopEd is an amazing community of in-school and out of school artists and educators who imagine uses of hip hop in cultural work with young children and adolescents (Morrell, 2008a,b).

Critical media pedagogies

In the first two decades of the 21st century literacy studies have become increasingly interested in digital media. This makes good sense as the lives of today's youth are saturated with media from television and film, to social media and video games, and the millennial generation spend the majority of their waking lives as consumers and producers of media. This interaction has both positive and potentially harmful effects. On one hand, youth are acquiring sophisticated skills via their media participation that could be useful to transition to academic achievement, professional employment, and civic life in the 21st century. However, major corporations who want to develop loyal consumers and therefore play into the insecurities of young people also heavily influence the media that youth consume. Media are also prone to send negative messages and they reinforce stereotypes and dominant value systems that could have deleterious effects on uncritical consumers.

The power of new media in the lives of young people cannot be denied and, therefore, media education needs to be more effectively implemented into the K-16 English curricula in the United States. Other nations such as Canada and Great Britain have called for media education, but their versions of media education have largely entailed the study of media texts. While this is important, we must recognize the democratizing nature of media as sites of production. And media education has not necessarily advocated a *critical stance* toward media production, nor has it looked at the considerations for historically underserved youth and their relationships toward media, past and present. There are, however, a recent generation of scholars who are looking to hip hop culture (Alim, 2006), film (Morrell, 2004), radio (Soepp and Chavez, 2010), spoken word poetry (Fisher, 2007) video gaming (Gee, 2003), and social media (Ito, 2009) as sites of youth participatory culture (Jenkins, 2009) engagement, agency, and resistance. While still wary of the potentially ill effects of media use, this latter group is much more optimistic about critical media literacy as an act of a critical and culturally responsive approach to literacy instruction.

My work theorizes critical media production among urban youth as a strategy for increasing literacy development and civic engagement. Drawing from a combination of philosophy, critical theory, and empirical research on media pedagogy with urban teens across inside of an inner city high school, I explore the applications of critical media pedagogy across the humanities disciplines noting how everyday teachers and university faculty bring new media into their teaching and the impacts of the new media on students' learning, achievement, and engagement.

I was influenced by the sociology of culture and cultural studies, particularly the work of Raymond Williams and John Storey (Storey, 2018) in thinking about approaches to the reading and re-writing of media found in the critical media pedagogy studies. Cultural theorist Raymond Williams (1995) articulates three components of culture that are essential to any thorough analysis of the subject.

1. *Ideal*, in which culture is a state or process of human perfection in terms of absolute or universal values.
2. According to the *documentary* component, culture is the body of intellectual and imaginative work, in which human thought and experience are recorded.

3. Finally, the third, or *social*, component of culture is a description of a particular way of life, which expresses certain meanings and values not only in art and learning, but also in institutions and “ordinary” behavior.

I am also interested in critical media analysis and youth popular cultural production as acts of counter hegemony. Hegemony, for Gramsci is a political concept developed to explain the absence of socialist revolutions in Western capitalist democracies (Gramsci and Hoare, 1957). The concept of hegemony is used by Gramsci to refer to a process in which a dominant class does not merely rule a society, but leads it through the exercise of moral and intellectual leadership. In this sense, the concept is used to suggest a society in which, despite oppression and exploitation, there is a high degree of consensus and a large measure of social stability; a society where the subordinate groups and classes appear to support meanings which bind them to, and incorporate them into, the prevailing structures of power. Hegemony is organized by those that Gramsci designates as the organic intellectuals.

In this approach, popular culture is seen as a site of struggle between the forces of resistance of subordinate groups in society, and the forces of incorporation of dominant groups in society. Popular culture in this usage is not an imposed “mass culture,” or a “people’s culture,” it is more of a terrain of exchange between the two. The texts and practices of popular culture move within what Gramsci calls a “compromise equilibrium.” Those who look at popular culture from a neo-Gramscian perspective tend to see it as a terrain of ideological struggle between dominant and subordinate classes, or dominant and subordinate cultures. Using a neo-Gramscian analysis, popular culture is what men and women make from their active consumption of the texts and practices of the culture industries. Youth cultures, for example, are able to appropriate for their own purposes and meanings the commodities that are commercially provided. For instance, in popular musical genres such as Reggae and hip-hop, it is possible to have anti-capitalist politics articulated in the economic interests of capitalism. The music may be lubricating the very system that it seeks to condemn. It may exist as an expression of oppositional politics which produces certain political and cultural effects in a form that is of financial benefit to the dominant culture. Cultural theorists must be aware of the simultaneous possibilities of the making of popular culture for subordinate groups. It has the potential of empowerment and resistance, but it can also lead to passivity and consumption of the hegemonic ideals promoted by the organic intellectuals of the dominant class.

More recently, we have expanded the critical media pedagogy work to articulate a new critical theory of multiliteracies that encompasses 4 types of digital engagement: (a) critical digital consumption, (b) critical digital production, (c) critical distribution, and (d) critical digital invention. We make the argument that a new critical theory of multiliteracies needs to account for each of these types of digital engagement but that, ultimately, we must move beyond theorizing our youth as passive consumers or even critical users of digital technologies toward the project of facilitating youth communities of digital innovation (Mirra et al., 2018).

YPAR and critical literacies for social action

There are thirty-five of us huddled in a corner of the Sheraton in New Orleans, away from the presentation rooms, trying to find a quiet place in the half hour or so we have before we present. The high school students pose for pictures with their peers and their teachers. Dressed up in suits, they look like they’re headed to work or church; instead, these high school students are about to deliver a presentation to the American Educational Research Association articulating the research on civic engagement and urban educational reform that they have been conducting for an entire school year. Ten minutes prior to show time we convene in our usual circle. Stragglers amble out of bathrooms tucking in shirts and checking their hair. The seniors step up for our traditional pep talk and offer their words of advice and encouragement. Their use of the collective pronoun reminds me again of what we are capable of creating as a society if we only had the will, if we only paid attention, and if we only fully believed in the humanity of these students and the neighborhoods and families from which they came. Everybody wants to feel a part of something special; everyone wants to contribute. We are dreamers all unless of course, those dreams through time and heartache have been deferred or extinguished altogether.

This brief vignette comes from a presentation to the American Educational Research Association (AERA) by high school students who were participants in the Council of Youth Research, a project I directed for 12 years (1999–2011) where high school students developed Youth Participatory Action Research (YPAR) projects in the neighborhoods and communities of Los Angeles. Our research group has written extensively on the connection between participation in this work and critical literacy development, civic engagement, community transformation, educational policy advocacy, and digital literacy learning.

YPAR is considered a subcategory of Participatory Action Research (PAR) and represents, “a radical shift from traditional research because of *what* it investigates and *who* does the investigating” (Morrell, 2006, pp. 112–113). YPAR is “critical,” “participatory” and “action-oriented” in that it “challenges the hierarchy of knowledge production and changes the relationship of knowledge producers to knowledge consumers” (Morrell, 2006, p. 3). While PAR is not a new research approach to, a critical perspective and approach with youth as the subjects rather than the objects, certainly is.

Morrow and Torres (1995) discuss how participatory action research gained recognition in Latin America during the 1970s through the Symposium of Cartagena on Critical Social Science Research in April 1977 and “the influence of Brazilian educator Freire (1970) who engaged marginalized populations of Brazilian peasants as collaborators, researchers, and activists” (Morrell, 2006, p. 6). His revolutionary pedagogy helped ordinary people develop critical literacy and inquiry skills to address structures of power. The idea of employing inquiry for advocacy is, as Morrell (2006) says, “as old as the tradition of inquiry itself” and reminds us that, “This is important to keep in mind amid contemporary conversations about quality, validity, and rigor in social scientific research” (p. 6).

McIntyre (2000) offered three principles that guide most participatory action research projects: (a) the collective investigation of a problem; (b) the reliance on indigenous knowledge to better understand that problem; and (c) the desire to take individual and/or collective action to deal with the stated problem.

Today there exists a flourishing community of educators who are using YPAR with students to investigate and intervene in social injustices around the world. This is a dynamic and inspiring body of work that can link YPAR work to academic achievement, positive identity development, and community engagement. Cammarota and Fine (2008) edited the first book that specifically refers to this work as Youth Participatory Action Research. I was fortunate enough to participate in that collective effort, which essentially presented this growing movement to the larger critical educational studies community. Scholars such as Shawn Ginwright, Beth Rubin, Eve Tuck, Ben Girshner, Maria Elena Torre, Ernest Morrell, Augustin Romero have produced volumes in the ensuing years that have shaped our understanding of YPAR and its contributions to an educational field that is increasingly interested in engagement, youth production, and social action.

Toward a praxis of universal critical literacy and revolutionary love

Toni Morrison reminds us that, if there is a book that you want to read, but it hasn't been written yet, then you must write it. For those of us in the academy we can take this quote quite literally. However, I would like to consider it collectively as a challenge and an opportunity to further integrate our field with the words and worlds of youth and communities, with media and popular culture, and with the cultural and literary histories of the Anglophonic African Diaspora.

It is my dream that this collective work, this as yet unwritten book, speaks to the silences that have plagued us and to those who have been left searching, in the silence, for recognition, for affirmation, and for historical significance. Yet, I also imagine this collective work as an invitation to all, as a source of celebration, and as a call to become co-participants in the making of bold, robust, and globally representative disciplinary futures. Put simply, the work matters. We matter. When I think of that April, nearly 30 years ago and the call of critical literacy studies, I am humbled by what this field has made possible for me and I am even further humbled by what remains to be accomplished.

References

- Alim, S., 2006. *Roc the Mic Right: The Language of Hip Hop Culture*. Routledge, New York.
- Barton, D., Hamilton, M., 2000. *Local Literacies: Reading and Writing in One Community*. Routledge, London.
- Bell, D., 1992. *Faces at the Bottom of the Well: The Permanence of Racism*. Basic Books, New York.
- Cammarota, J., Fine, M. (Eds.), 2008. *Revolutionizing Education: Youth Participatory Action Research in Motion*. Routledge, New York.
- Crenshaw, K., Gatonda, N., Peller, G., Thomas, K. (Eds.), 1995. *Critical Race Theory: The Key Writings That Formed the Movement*. The New Press, New York.
- Darder, A., 1991. *Culture and Power in the Classroom: A Critical Foundation for Bicultural Education*. Greenwood Publishing, Santa Barbara, CA.
- Duncan-Andrade, J., Morrell, E., 2008. *The Art of Critical Pedagogy: Possibilities for Moving From Theory to Practice in Urban Schools*. Peter Lang, New York.
- Fisher, M., 2007. *Writing in Rhythm: Spoken Word Poetry in Urban Classrooms*. Teachers College Press, New York.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York.
- Gee, J., 2003. *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave-Macmillan.
- Giroux, H., 1992. *Border Crossings: Cultural Workers and the Politics of Education*. Taylor and Francis, New York.
- Gramsci, A., Hoare, Q., 1957. *Selections From the Prison Notebooks*. Lawrence and Wishart, Chadwell Heath, UK.
- Gutierrez, K., 2008. Developing a sociocritical literacy in the third space. *Read. Res. Q.* 43 (2), 148–164.
- Hall, S., 1997. *Representation: Cultural Representations and Signifying Practices*. Sage, Thousand Oaks, California.
- Heath, S.B., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, Cambridge, UK.
- Hill, L., 1998. *The miseducation of Lauryn Hill*. In: *On the Miseducation of Lauryn Hill*. Ruffhouse, Philadelphia.
- hooks, b., 1994. *Teaching to Transgress*. Routledge, New York.
- 1902–1967 Hughes, L., 1994. *The Collected Poems of Langston Hughes*. Knopf : Distributed by Random House, New York.
- Ito, M., et al., 2019. *Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning With New Media*. MIT Press, Cambridge.
- Ito, M., 2009. *Engineering Play: A Cultural History of Children's Software*. The MIT Press, Cambridge, MA.
- James, C.L.R., 1989. *The Black Jacobins: Toussaint L'Ouverture and the San Domingo Revolution*. Vintage, New York.
- Jenkins, H., 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. The MIT Press, Cambridge, MA.
- Lee, C., 1993. Signifying as a Scaffold for Literary Interpretations: The Pedagogical Implications of an African American Discourse Genre. *National Council of Teachers of English*.
- Lipsitz, G., 1997. *Dangerous Crossroads: Postmodernism, Popular Music and the Poetics of Place*. Verso, New York.
- Marley, B., Sporty, K., 1978. *Buffalo soldier*. In: *On Confrontation*. Tuff Gong Studios, Kingston, Jamaica.
- McIntyre, A., 2000. Constructing meaning about violence, school, and community: participatory action research with urban youth. *Urban Rev.* 32, 123–154.
- McLaren, P., 1989. *Life in Schools: An Introduction to Critical Pedagogy in the Foundations of Education*. Longman, Harlow, UK.
- Mirra, N., Garcia, A., Morrell, E., 2016. *Doing Youth Participatory Action Research: Transforming Inquiry for Researchers, Educators, and Students*. Routledge, New York.
- Mirra, N., Morrell, E., Filipiak, D., 2018. From digital consumption to digital invention: toward a new critical theory and practice of multiliteracies. *Theor. Pract.* 57, 12–19.
- Morrell, E., Rowsell, J. (Eds.), 2020. *Stories From Inequity to Justice in Literacy Education: Confronting Digital Divides*. Routledge, New York.
- Morrell, E., Garcia, V., Lopez, J., Duenas, R., 2013. *Critical Media Pedagogy: Achievement and Justice in City Schools*. Teachers College Press, New York.
- Morrell, E., 2004. *Linking Literacy and Popular Culture: Finding Connections for Lifelong Learning*. Christopher Gordon, Norwood, MA.
- Morrell, E., 2006. Critical participatory action research and the literacy achievement of ethnic minority groups. In: *55th Annual Yearbook of the National Reading Conference*, vol. 55, pp. 60–78.
- Morrell, E., 2008a. *Critical Literacy and Urban Youth: Pedagogies of Access, Dissent, and Liberation*. Routledge, New York.
- Morrell, E., 2008b. *Rebel musics: African Diaspora popular culture and critical literacy pedagogies*. In: Payne, C., Strickland, C.S. (Eds.), *Teach Freedom: Education for Liberation in the African-American Tradition*. Teachers College Press, New York.

- Morrison, T., 1987. *Beloved*. Alfred A. Knopf, New York.
- Morrison, T., 1993. *Playing in the Dark: Whiteness and the Literary Imagination*. Verso, New York.
- Morrow, R.A., Torres, C.A., 1995. *Social Theory and Education: A Critique of Theories of Social and Cultural Reproduction* (SUNY Series, Teacher Empowerment and School Reform). State University of New York Press, Albany, NY.
- Nieto, S., 1992. *Affirming Diversity: The Sociopolitical Context of Multicultural Education*. Allyn & Bacon, Incorporated, New York.
- Perry, I., 2004. *Prophets of the Hood: Politics and Poetics in Hip Hop*. Duke University Press, Durham, NC.
- Rose, T., 1994. *Black Noise: Rap Music and Black Culture in Contemporary America*. Wesleyan Press, Middletown, CT.
- Scholes, R., 1998. *The Rise and Fall of English: Reconstructing English as a Discipline*. Yale University Press, New Haven, CT.
- Smitherman, G., 1999. *Talkin That Talk: Language, Culture, and Education in African America*. Routledge, New York.
- Soepp, E., Chavez, V., 2010. *Drop That Knowledge: Youth Radio Stories*. University of California Press, Berkeley and Los Angeles, CA.
- Solórzano, D., Yosso, T., 2002. A critical race counterstory of race, racism, and affirmative action. *Equity Excell. Educ.* 35 (2), 155–168.
- Storey, J., 2018. *Cultural Theory and Popular Culture: An Introduction*, eighth ed. Routledge, New York.
- Torres, C., 1998. *Democracy, Education, and Multiculturalism: Dilemmas of Citizenship in a Global World*. Rowman and Littlefield.
- UNESCO, 2017. Africa Department: Demographic Growth. Retrieved from. <http://www.unesco.org/new/en/africa-department/priority-africa/operational-strategy/demographic-growth/>.
- West, C., 1993. *Race Matters*. Beacon Press, Boston.
- Williams, R., 1995. *The Sociology of Culture*. University of Chicago Press, Chicago, IL.
- Wright, R., 1940. *Native Son*. Harper and Brothers, New York.

Youth poetry and literacies in remaking the world

Korina M. Jocson, York University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

On pedagogy and the politics of poetry	235
On participatory cultures and new media literacies	236
On youth media and translocal practices	236
Taking a step back, reflecting on the limits of poetry	237
Trauma, poetry, and literacies	237
Beyond humanist approaches in youth poetry	238
Futures in/of/with youth poetry	240
Acknowledgments	240
References	240

Poetry is a much-valued form of expression. For the past 25 years, poetry has opened up possibilities in teaching and learning and expanded scholarship in literacy studies. Three dimensions of this research based on my own inquiry are (a) pedagogy and the politics of poetry, (b) participatory cultures and new media literacies, and (c) youth media and translocal practices. It is also important to discuss emerging lessons about youth poetry and questions that pertain, for example, to trauma-informed literacies and new materialisms in youth studies. For this entry, I have culled together previous writings to illustrate studies. For this entry, I have culled together previous writings to illustrate some relevant perspectives on youth poetry. Lastly, I point to future directions in civic literacies, speculative education, multimodality, and critical research methodologies in further shaping our understanding of youth poetry. My hope is to build on connective threads toward more fruitful conversations.

On pedagogy and the politics of poetry

Given the persisting perceptions and treatment of young people in school and society, particularly young people of color from marginalized communities, it has been crucial to turn to youth themselves to understand the ways in which poetry mediates their everyday realities. My earlier studies focused on how young people draw on poetry as a medium to help make sense of their social words, including the persons they are becoming and the desires they have for their present and future lives (Jocson, 2006a,b). Tapping into the power of poetry and what young people have to say renders possibilities for educators, youth advocates, cultural workers, researchers, and policymakers alike. In my work in classrooms and community settings, I have witnessed the influence of poetry and how poetry can be used as a pedagogical tool to expand views about young people and youth cultures, inform academic and critical literacies, and challenge static notions of teaching and learning (Jocson, 2008). Examples are detailed in my book *Youth Poets: Empowering Literacies In and Out of Schools*, which sheds some light on treating poetry as a social practice among youth to reveal complexities of teaching and learning in the context of a university-school partnership through June Jordan's Poetry for the People (P4P). Central to P4P's pedagogical approach, then and now, has been to support youth's participation in poetry writing and to create opportunities for critical literacy and civic engagement.

Writers all over the world have felt a kinship with poetry and at times have utilized poetry as a democratic means to respond to political upheavals. Poetry pushes boundaries to raise popular consciousness by speaking truth to power. It serves as a medium for truth-telling, a vital element in political activism and social transformation. In the US poets and writers and their historical legacies have been a driving force in cultural and social movements. Youth poets, in particular, are at the center of these movements and over the years have increasingly participated in literary arts organizations where writing workshops and public performances help to showcase their written-verbal artistry and community activism (Jocson and Rivera, 2010; Kim, 2013). In my studies, I have seen youth's increasing interest in poetry within the broader contexts of their lives to shape what and how they write, for whom it matters, and with whom they share.

With influences of popular culture, poetry has become a more inclusive medium of expression in schools and carries with it historical legacies of African oral traditions and storytelling. In the 1960s and 1970s, for example, poetry as a black art form played a significant role in helping to disseminate the politics of the Civil Rights Movement. Poetry readings served as esthetic spaces for sharing messages of hope and struggle. Poetry advanced the Civil Rights Movement through public performances within institutional settings and local community venues, including spoken word and open mics. Poetry's revival in the 1990s and into 2000s drew upon earlier periods of the Black Arts Renaissance and today has reached global audiences. In the field of education, scholars have offered a range of perspectives on the power and potential of youth poetry (Fisher, 2007; Sablo Sutton, 2004; Watson, 2013; Weinstein and West, 2012). With the advent of new technologies, poetry has exploded from the page through chapbooks and zines, radio broadcasts, television shows and commercials, films, podcasts, and various other social media outlets. There is no doubt

that poetry has (re)emerged stronger than ever and has become more publicly accessible. It is this cultural phenomenon to which I now turn.

On participatory cultures and new media literacies

Today's creative expressions blur the lines between youth cultural production and participatory politics. Indeed, there are myriad ways of participating in civic life. It is no longer a surprise that many young people are making videos, taking photographs, creating websites, blogging, assembling mash ups, and sharing them online for others to view. From a new media literacies perspective, it is argued that an ethos of collaboration, participation, and distributed expertise shapes how individuals see themselves in the world and interact with each other as afforded by digital technologies (Lankshear and Knobel, 2006). At the center of this ontological and technocultural shift are youths who engage in alternative and activist practices (Lievrouw, 2011).

Such practices reflect what Cohen and Kahne (2012) call participatory politics, in which interactive, peer-based acts have the ability to reach large audiences, fashion agendas through dialog, and exert greater agency through the circulation of information and the production of original content, both online and offline. In this light, the notion of participatory politics is seen as wielding voice and influence on issues of public concern without being guided by deference to elites or formal institutions. Participatory politics are interest-driven and voluntary. In my continued work with secondary and tertiary students, I have been moved by the potential of participatory politics not only as voluntary, but also as pedagogically operational in critical education (Anyon, 2009a,b; Apple et al., 2009). What happens when participatory politics become a key component of the classroom experience? What pedagogical considerations are necessary to enable both students and educators to embrace varied uses of technology toward alternative and activist practices? In what ways can the curriculum be invigorated by interest-driven and peer-based acts within the confines of school, university, and other formal institutions? Philip and Garcia (2013) remind us of the centrality of pedagogy and the continued need for a dynamic relationship between students and teachers. As such, youth poetry is consistent with growing forms of youth cultural production and participatory politics in the context of new media literacies (Jocson, 2013, 2015). It takes many forms especially given the pace of changing technologies. In classrooms and beyond, it is important to consider how youth leverage new media literacies at the intersection of youth poetry and youth media in order to understand pedagogical possibilities.

On youth media and translocal practices

One of the exciting practices in youth media is remixing cultural material. Remix is the idea that young people are blurring genres of alternative media by building on each other's work and using technology as dialogical tools of power to combat racial injustice, surveillance, and police brutality, among others. The forging of cultural texts as found in multimedia poems and other genres of alternative media has implications for pedagogical innovations in education. Building on these perspectives on remix, the term *translocal* within literacy studies has been theorized in various ways. Scholars have pointed to the limits of the local and asserted the material consequences associated with literacy at a more global level (Brandt and Clinton, 2002). Others have viewed artifacts as central to literacy research and have suggested the importance of attending to the material consequences that artifacts bear in the social world. For example, Braynham and Prinsloo (2009) trace the objects that circulate within and between sites of social human interaction as "things" which are seen as necessary components of social networks and practices (also see Latour, 2005). This is similar to what Lam (2009) and Medina (2010) discuss as translocal literacy practices among youth in a digitally mediated world and the cultural flows that happen through migration and media exchange. Examples such as instant messaging and reading literature in transnational contexts are not only artifactual but also build on cultural imaginaries toward learning and literacy development (see Medina and Wohlgend, 2014). It has also been argued that, in grassroots literacies, texts produced in local economies become translocal documents with different possible uptakes (Blommaert, 2008). More specific to remix in youth cultural production and participatory politics, translocal assemblage (McFarlane, 2009) offers as an analytic frame for place-based social movements with the following orientations:

- (1) Exchange of ideas, knowledge, practices, materials, and resources across sites;
- (2) Building capacity to exceed connections between other groups or places in the movement beyond a node or point in history;
- (3) Relations between and within sites that signify a doing or performance but at the same time are vulnerable to collapse or to reassembling in other forms.

In other words, translocal assemblage provides a way to understand the production and dissemination of youth media work toward the exchange of ideas and knowledge, building capacity and connections, and performance as action. Taking into account existing texts to create new ones (or rewriting the familiar) is central to the process. Multimedia poems, or poems rendered with image, sound, and related effects, display layers of intertextuality that signal a kind of assembling or (re)assembling of cultural material across space and time. For example, the spoken word poem "Barely Audible" performed by then-high-schooler Chinaka Hodge about issues of gentrification in Oakland, California, became the basis for a video produced by NYU Tisch School filmmakers Katherine Copeland Anderson, Eli Jacobs-Fantauzzi, and Vivian Wenli Lin, about similar sociohistorical conditions in Brooklyn, New York (see Jocson with Jacobs-Fantauzzi, 2013). Or the spoken word poem performed by then-college-student Adriel Luis about

gender and standards of beauty became a personal critique of media's depiction of women, particularly women of color, in a video by then-high-schooler and budding multimedia artist Karen Lum (see [Jocson, 2005](#)).

Theoretically, the intertextual nature of media production opens up doors for exchange with similar yet different possible uptakes. Pedagogically, by building on existing media products, it renders the convergence of cultural material as part of translocal assemblage. In my work in higher education, the college classroom has provided a context for instantiations of translocal assemblage in furthering the discourse on racial justice and related topics (see [Jocson with Carpenter, 2019](#)). Students enrolled in my courses have had opportunities to create their own alternative media to render the social, economic, cultural, and political context of their/our time. I discuss related examples in *Youth Media Matters: Participatory Cultures and Literacies in Education* ([Jocson, 2018](#)).

There are many more examples on social media that can build on translocal assemblage practices. They demonstrate a way of speaking back to highlight students' voices about racialized experiences across geographic lines. These works reflect cultural flows as fluid and dynamic. Materially and discursively, these works are indicative of translocal assemblage that point to the exchange of ideas, knowledge, practices, materials, and resources across sites ([McFarlane, 2009, 2012](#)), and the ways in which the connections between them and the potential for reassembling cultural material into different forms of alternative media can be relevant to teaching, learning, and critical education ([Cohen and Kahne, 2012](#); [Soep, 2014](#)). Young people are finding civic resources and telling stories about their political becoming. That requires designing learning environments that support the making of public spheres.

Youth media continues to push our thinking about how youth are coming together in local–distant contexts, participating in the new media landscape, and inserting themselves more actively online and offline to raise awareness about issues that matter in their lives. Some become more visible than others, which may pave the way more quickly for other forms of remix and cultural flows. Indeed, youth are using digital media technology with different tools for different purposes across educational settings using tactics of participatory politics ([Soep, 2014](#)). With an understanding of the relational impact of new technologies on teaching and learning, it is important to see translocal assemblage as the reinterpretation of cultural imaginaries and to ask the following question: How can we as educators and youth advocates create the kinds of spaces, practices, and conditions toward educative possibilities that can have real and tangible social impact?

In thinking about possible answers, educators and youth advocates must also participate in the making and remaking of cultural imaginaries, a kind of public pedagogy ([Giroux, 2004](#)) that extends the who and the how of pedagogy to the ways in which the cultural–political ([Youdell, 2011](#)) can be intertwined in the daily life of school and other educational contexts. This is more crucial than ever as we have seen increasing levels of participatory politics among youth in recent times. The translocal links different forms of struggle and politics of difference in cultural and social movements, where the translocal becomes part of transnational discourses and vice versa. Multimedia poems and other forms of media production offer young people ways of writing the politicized/ing (collective) self. Indeed, many young people are drawing on their voices and talents to challenge normalized discourses, disrupt dominant narratives, and act as concerned members of society. They are performing identities, allegiances, and oppositions that reach wide audiences, to engage in communicative action and to mobilize the public concerning racial justice.

It is no accident that various collectives across locales have come together as part of a historical continuum to remix cultural material with a commitment to social alliance and social action. Alternative media does not just surface because more and more young people are engaging in participatory politics. Tactics of participatory politics such as creating and circulating information via social media platforms often require a continuous process of mediation between people's uses of technology and communicative action to push social and cultural change in certain directions. That is, alternative media in the form of multimedia poems help to incite the exchange of ideas and capacity building to reassemble cultural material into different genres of participatory politics. Translocal practices also enable learning focused on equity and justice-oriented approaches in and beyond the classroom.

Taking a step back, reflecting on the limits of poetry

As much as it is about truth-telling, youth poetry and its many forms are also about remembering. In many classrooms young people are often invited to write from/about their lived experiences that may invoke traumas or the relationships within them ([Jean-Denis and Jocson, 2021](#)). Here, I pause to consider what they may bring to light for deepening our understanding of youth poetry, and what else we as educators and researchers might consider beyond humanist approaches in youth studies ([Albright and Jocson, 2021](#)). The lessons point to key questions.

Trauma, poetry, and literacies

Poetry within trauma-informed literacies has been influential to understanding youth writing. In recent years, trauma-informed practice has become a buzzword throughout many public school districts. The mandates for classroom teachers include remaining vigilant for signs of trauma among youth. But what happens when the trauma stems from the very system that is supposedly designed to keep youth safe? For educators, it is imperative to deepen our understanding of the myriad ways that trauma might be expressed in classrooms and how it manifests through/despite histories of colonialism, enslavement, and dehumanization.

Increasing interests in and applications of trauma-informed practice in education guide us toward scholarship in literacy studies to open up conversations about pedagogy. It is well noted that trauma-informed care has often focused on the individual or on the social emotional learning and resources needed in support of "difficult" students. Rather than extend this approach, there is a call to

shift from trauma-informed care to healing in order to more pointedly address historical conditions and sociopolitical processes that sustain different forms of violence in young people's lives.

While trauma-informed practice is important, [Ginwright \(2018\)](#) in "The Future of Healing" argues that it is incomplete. The focus on individual versus collective experience tends to center on individual pain (trauma), not well-being (healing). Put simply, the focus on "fixing" the person does not get at the root causes of trauma (i.e., context, social and economic conditions). Drawing on his extensive work in urban education and with black youth in the San Francisco Bay Area, [Ginwright \(2016\)](#) in *Hope and Healing in Early Education* encourages educators to consider healing justice practices and suggests ways of supporting youth through complex forms of trauma and combining healing with hope toward social change. Ginwright and colleagues engage in community organizing as part of the healing process. Ginwright proposes healing-centered approaches (radical healing) to facilitate quality of life at the level of both individuals' well-being and the community's well-being—that is, to attend to and restore the economic, political, and social conditions that allow people to thrive. Aligned with healing-informed practices are the limits of socio-emotional learning which, in their view, does not address social forces that shape the health and wellness of communities of color. [Camangian and Cariaga \(2021\)](#) call for a process of humanization, "a reciprocal, mutually anti-oppressive process of self and collective care and development in the context of social transformation" (p. 6). As such, these studies particularly with/for youth and communities of color indicate the limits of trauma-informed care and the potential of humanizing pedagogy in education.

The inextricable link between trauma, poetry, and literacy is important to interrogate. In literacy studies, [Dutro and Bien \(2014\)](#) examine the role of trauma in classrooms. Using the metaphor of the speaking wound, they consider the adult-authored representations of lived experience and how they position students within/through stories of trauma. [Dutro \(2019\)](#) also explicates the importance of critical witnessing and the prevailing ways that trauma may fuel damage narratives about students and their families. It is another reminder that [Tuck \(2009\)](#) admonishes and warns against damage or pain narratives, which are built on notions of brokenness as a "pathologizing approach" (p. 413), reinforcing deficit perspectives particularly among indigenous communities and across communities of color. Relatedly, [Wissman and Wiseman \(2011\)](#) have found in their studies that writing poetry offers an opportunity for students to engage in "narrative control" or "a way of using language to claim the right to name their own experience" (p. 243) as they speak-write in school classrooms about difficult experiences, including family conflict, particularly when conditions at home or constrained relationships limit their conversations. For educators, classrooms ought to consider critical witnessing as a way of paying attention to everyday testimonies while resisting deficit narratives as students receive the support they need.

In more recent works, [Baxter and Sealey-Ruiz \(2020\)](#) demonstrate the importance of poetry as an art form to "mirror Black resistance" against dehumanization and offer a window into collective struggle in line with the black radical tradition. In particular, they illustrate through their own poetry how words heal, or how poetry becomes a medium through which to "bare our soul, release our pain and sadness, and restore ourselves" ([Baxter and Sealey-Ruiz, 2020](#), p. 312). Because as noted time and again "poetry is not a luxury; it is a vital necessity of existence" ([Lorde, 1984](#), p. 37). Because poetry is a site of struggle, a consciousness-raising tool to speak back through "the system of language that excludes her name and all the names of her people" ([Jordan, 1985](#), p. 5), a site of radical openness ([hooks, 1989](#)). To utilize poetry in the classroom, then, is to provide a critical space for healing. We already know this to be significant as various studies on written and spoken word poetry in classrooms or after school programs have promoted a range of pedagogical possibilities with youth's voice at the center (see [Fisher, 2007](#); [Jocson, 2008](#); [Kim, 2013](#); [Morrell, 2008](#); [Sablo Sutton, 2004](#); [Watson, 2013](#); [Weinstein and West, 2012](#)). Healing-centered approaches widen the learning and teaching spectrum for youth engagement in educational contexts.

Keeping in mind a critical space for healing, it is important for educators to step back from trauma-informed practice. When we ask young people to write about trauma, it is very possible that the task may actually be more harmful than we think, not just in the moment but also in other moments to come. Perhaps young people desire something else or would prefer to exercise refusal. While it matters to be heard, seen, validated, and affirmed, it also matters to just be, to exist outside the limits of categories. In an adult-driven world where young people are often constrained by rules and norms, the tendency for many young people is to rely on doing things out of expectations (sometimes against their good sense) in order to be "heard, seen, validated, and affirmed." As [Evans-Winters \(2019\)](#) contends, it is very possible that young people may refuse or may not yet be ready to write about particular moments in their lives, that some young women's "experiences are hers alone to experience and decide if they should be shared" (p. 105). There might be a need for discretion to understand time-space or context differently. There might be a need to rethink current approaches or to steer clear of hurried decisions about what works and what doesn't.

Beyond humanist approaches in youth poetry

In my work, if poems on the page were strictly treated from a humanist, sociocultural lens, then there would be missed opportunities for considering the moments still folding at the time of the writing. What are those moments, and in what ways do they shape (how we view or treat) youth's literacies across various social and educational spaces? In what ways can they better illuminate text and context in youth poetry, or literacy practices among youth? Most salient are sociocultural approaches focused on the construction of words and meanings through dialogic encounters ([Bakhtin, 1935/1981](#)). Specific to youth literacies, for example, poems produced by youth, inclusive of spoken word performances, are influenced by others' voices and writing styles with some utterances drawn directly from existing literary texts. A new materialist approach, attentive to (non)human relationships and moments still unfolding, can help to illustrate how we as researchers are always a part of the phenomenon being studied.

A re-reading of youth poetry is without a re-reading of our own processes of becoming as writers, thinkers, and collaborators within youth studies. For instance, while attuned to dialogism and the ways utterances are lifted from others' words or derived from literature, I recognize the assemblages that influence one's writing and how they materially and discursively shape the text produced. What might a new materialist approach to youth poetry yield to help expand theory and practice in literacy studies? How might a consideration for literacies as assemblage shift how we understand the text produced or the context in which it is produced? How might a consideration of our own processes of becoming influence how we view text and context in youth studies? For example, 17-year-old Carolyn wrote a poem in response to a critical race moment (see [Jocson, 2011](#)). The poem is as follows:

Can you see me?
 When you think black
 You think guns going glack glack
 Sending shots through your neighborhood
 Stores
 There's no turning back
 You wack, you scared
 Can you compare
 To all the lashes I got across
 My back
 Seek
 The curse in my eyes
 The flare of my nose
 The adrenaline in my chest
 The grinch in my teeth
 The bitch in my breath
 Why the hell
 You gotta be so
 Difficult
 Listen to me when I tell you
 What's on my mind
 The truth
 Twisted up but I spit it out
 So let it be known
 What the guns in their hands is all about
 You
 Don't see me
 You see right through me
 You want me locked up as much as the next man
 Do I look like a hoodlum to you?
 I'm not the black you know
 I'm the black you will know
 So I ask
 Once more
 Can you see me?

Similar to Carolyn on a separate occasion, I, too, wrote a poem in response to a critical race moment. Everyday encounters with race and language point to raciolinguistic ideologies ([Flores and Rosa, 2015](#)) that construct what is deemed appropriate linguistic practices in academic settings. The poem titled "academic dialogue" is as follows:

i readandreadandreadandread still refuse to regurgitate the meaning of classic modernist paradigmatic functional discourse associated with those i know don't look like me pressed to articulate whatever sorries may come ambiguous words are just that words
 i writeandwriteandwriteandwrite but still can't find my own way to describe the relationship between polisocioeconomicsubandsuperstructure without calling on marx verbal marks 'cause the words i'm looking for really just tell it like it is spoken so i talkandtalkandtalkandtalk what may come in knots and communicate face to face vocabulary that makes sense settle once and for all freedom among us as words flow from our criticallyconscioustongues.

These two poems serve as examples to echo curricular material from June Jordan's Poetry for the People program (as discussed above), in particular writing to a prompt on racial profiling that many youth poets, students, and student-teacher-poets including

myself encountered in the program. The poems also serve as a heuristic for what possibilities may come when everyday experiences become living entities for analysis and exchange.

In stepping back to rethink the moments that shaped what is revealed and not revealed in the poems, I have asked myself (and perhaps you—the reader—might do the same), “How do these particular instances of writing, of everyday living, of racialized and gendered experience continue to inform the ways I engage youth literacies in my work? What might be missed when my own processes of becoming and enfolding possibilities remain in the sidelines of theory and research?” The “knots” within the poems remain relevant in this very moment, may even resonate in multiple ways, and are worth considering (again). Processes of becoming offer ways of further understanding youth poetry and expanding methodological approaches in literacy studies.

Futures in/of/with youth poetry

There is continued excitement in the field of literacy studies. From the purview of youth poetry, I am inspired and compelled by contemporary scholarship on youth civic literacies and civic reasoning (Mirra and Garcia, 2020; Lee et al., 2021), speculative education and fiction writing (Nasir et al., 2021; Toliver, 2021; Thomas, 2019), and multimodality in youth engaged research (Lam et al., 2021; Ybarra, 2021). These areas not only push the way of thinking about or the potential of youth poetry across learning environments, including the digital, but also offer methodological openings for how we do inquiry with considerations for what young people have to say, their use of alternative modes of storytelling, and how they share their perspectives situated within social, cultural, political, historical, and geographic/spatial contexts, as well as the conditions in which they live, thrive, and struggle. Pursuing forms of poetic inquiry will require serious considerations for method-making and innovative approaches espousing arts-based critical methodologies to reimagine education (McKittrick, 2021; Rashid and Jocson, 2021). This means engaging less prescriptive and more creative (often speculative) endeavors in pursuit of otherwise worlds (Crawley, 2016).

From my experience monitoring youth’s engagement with poetry in the US, I am inspired and compelled to lean in (a bit) closer to hear what young people have to say, inciting ideas about how education might respond to serve all students. Youth poet laureate Amanda Gorman at the 2020 US Presidential Inauguration delivered an incisive message reaching the hearts of many around the globe. Her words from “The Hill We Climb” now extend far beyond that particular moment in time. Her written poems fill many other pages of books, headline social media sites, and are featured on various websites of poetry organizations (Gorman, 2021a,b). Her poetry has also been published as a children’s book, *Change Sings*, in collaboration with bestselling illustrator Loren Long to enliven each storied line. Civic literacies along with the multimodal dynamics of poetry and the speculative nature of storytelling are embedded. Such are examples of youth poets remaking the world and shaping future directions in literacy studies.

Acknowledgments

This entry is made possible by long-standing relationships and co-learning over the years with youth poets, artists, teachers, and partners in education. I am grateful to co-authors and collaborators for opportunities to engage youth poetry through interdisciplinary perspectives, some of which are included here. Readers are encouraged to learn more from the cited works.

References

- Albright, T., Jocson, K.M., 2021. Theorizing with assemblage: text and context in youth studies. In: Matias, C. (Ed.), *Handbook of Critical Theoretical Research Methods in Education*. Routledge, New York, pp. 212–225.
- Anyon, J., 2009a. Critical pedagogy is not enough: social justice education, political participation, and the politicization of students. In: Apple, M., Au, W., Gandin, L. (Eds.), *The Routledge International Handbook of Critical Education*. Routledge, New York, pp. 389–395.
- Anyon, J. (Ed.), 2009b. *Theory and Educational Research: Toward Critical Social Explanation*. Routledge, New York.
- Apple, M., Au, W., Gandin, L. (Eds.), 2009. *The Routledge International Handbook of Critical Education*. Routledge, New York.
- Bakhtin, M., 1935/1981. *The Dialogic Imagination: Four Essays* by M. Bakhtin. University of Texas Press, Austin, TX.
- Baxter, G., Yolanda Sealey-Ruiz, Y., 2020. In the Black Radical Tradition: poetry as a praxis for healing and resistance in education. *Res. Teach. Engl.* 55 (30), 311–321.
- Blommaert, J., 2008. *Grassroots Literacy: Writing, Identity, and Voice Central Africa*. Routledge, London, UK.
- Brandt, D., Clinton, K., 2002. Limits of the local: expanding perspectives on literacy as a social practice. *J. Lit. Res.* 34, 337–356. https://doi.org/10.1207/s15548430jlr3403_4.
- Braynham, M., Prinsloo, M. (Eds.), 2009. *Literacies, Global and Local*. John Benjamins Publishing, Amsterdam.
- Camangian, P., Cariaga, S., 2021. Social and Emotional Learning is Hegemonic Miseducation: Students Deserve Humanization Instead. *Race Ethnicity and Education*, pp. 1–21. <https://doi.org/10.1080/13613324.2020.1798374>.
- Crawley, A., 2016. *Blackpentecostal Breath: The Aesthetics of Possibility*. Fordham University Press, New York.
- Cohen, K., Kahne, J., 2012. *Participatory Politics: New Media and Youth Political Action*. YPP Network/MacArthur Foundation, Oakland, CA.
- Dutro, E., 2019. *Vulnerable Heart of Literacy: Centering Trauma as Powerful Pedagogy*. Teachers College Press.
- Dutro, E., Bien, A., 2014. Listening to the speaking wound: a trauma studies perspective on student positioning in schools. *Am. Educ. Res. J.* 51 (1), 7–35.
- Evans-Winters, V., 2019. *Black Feminism in Qualitative Inquiry: A Mosaic for Writing Our Daughter’s Body*. Routledge.
- Fisher, M.T., 2007. *Writing in Rhythm: Spoken Word in Urban Classrooms*. Teachers College Press.
- Flores, N., Rosa, J., 2015. Undoing appropriateness: raciolinguistic ideologies and language diversity in education. *Harv. Educ. Rev.* 85 (2), 149–171.
- Giroux, H., 2004. Cultural studies, public pedagogy, and the responsibility of intellectuals. *Commun. Crit. Cult. Stud.* 1 (1), 59–79. <https://doi.org/10.1080/1479142042000180926>.
- Ginwright, S., 2016. *Hope and Healing in Urban Education: How Urban Activists and Teachers are Reclaiming Matters of the Heart*. Routledge.

- Ginwright, S., 2018. The future of healing: shifting from trauma informed care to healing centered engagement. Medium. <https://medium.com/@ginwright/the-future-of-healing-shifting-from-trauma-informed-care-to-healing-centered-engagement-634f557ce69c>. (Accessed 28 June 2021).
- Gorman, A., 2021a. *Change Sings: A Children's Anthem*. Penguin.
- Gorman, A., 2021b. *Call Us What We Carry*. Viking.
- hooks, b., 1989. Choosing the margin as a space of radical openness. *Framework* 36, 15–23.
- Jean-Denis, A.S., Jocson, K.M., 2021. "Because I won't ever forget": toward livingness in youth poetry. *J. Juvenilia Stud.* 4 (1), 41–50. <https://doi.org/10.29173/jjs75>.
- Jocson, K.M., 2005. "Taking it to the mic": pedagogy of June Jordan's Poetry for the people and partnership with an urban high school. *Engl. Educ.* 37 (2), 44–60.
- Jocson, K.M., 2006a. "Bob Dylan and Hip Hop": intersecting literacy practices in youth poetry communities. *Writ. Commun.* 23 (3), 231–259.
- Jocson, K.M., 2006b. "There's a better word": urban youth re-writing their social worlds through poetry. *J. Adolesc. Adult Lit.* 49 (8), 700–707.
- Jocson, K.M., 2008. *Youth Poets: Empowering Literacies in and Out of Schools*. Peter Lang.
- Jocson, K.M., 2011. Poetry in a new race era. *Daedalus* 140 (1), 154–162.
- Jocson, K.M., 2013. Remix revisited: critical solidarity in youth media arts. *E-Learn. Digit. Media* 10 (1), 68–82.
- Jocson, K.M., 2015. New media literacies as social action: the centrality of pedagogy in the politics of knowledge production. *Curric. Inq.* 45 (1), 30–51.
- Jocson, K.M., 2018. *Youth Media Matters: Participatory Cultures and Literacies in Education*. University of Minnesota Press, Minneapolis, MN.
- Jocson, K.M., Rivera, T., 2010. Toward a theory of poemness: cultural politics and transformative pedagogies. In: Leonardo, Z. (Ed.), *Handbook of Cultural Politics and Education*. SensePublishers, Boston, MA, pp. 527–540.
- Jocson, K.M., Carpenter, A., 2019. Translocal assemblage and the practice of alternative media toward racial justice: a pedagogical perspective. *Crit. Stud. Educ.* 60 (2), 245–262. <https://doi.org/10.1080/17508487.2016.1234493>.
- Jocson, K.M., Jacobs-Fantauzzi, E., 2013. "Barely audible": a remix of poetry and video as pedagogical practice. In: Jocson, K.M. (Ed.), *Cultural Transformations: Youth and Pedagogies of Possibility*. Harvard Education Press, Cambridge, MA, pp. 13–31.
- Jordan, J., 1985. *On Call: Political Essays*. South End Press.
- Kim, R.H., 2013. "Never knew literacy could get at my soul": on how words matter for youth, or notes toward decolonizing literacy. *Rev. Educ. Pedagog. Cult. Stud.* 35 (5), 392–407.
- Lam, W.S.A., 2009. Multiliteracies on instant messaging in negotiating local, translocal, and transnational affiliations: a case of an adolescent immigrant. *Read. Res. Q.* 44 (4), 377–397. <https://doi.org/10.1598/RRQ.44.4.5>.
- Lam, W.S.A., Smirnov, N., Chang, A., Easterday, M., Rosario-Ramos, E., Doppelt, J., 2021. Multimodal voicing and scale-making in a youth-produced video documentary on immigration. *Res. Teach. Engl.* 55 (4), 340–368.
- Lankshear, C., Knobel, M., 2006. *New Literacies: Everyday Practices and Classroom Learning*. Open University Press, Maidenshead.
- Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press, New York, NY.
- Lee, C.D., White, G., Dong, D. (Eds.), 2021. *Educating for Civic Reasoning and Discourse*. National Academy of Education, Washington, DC.
- Lievrouw, L., 2011. *Alternative and Activist New Media*. Polity, Malden, MA.
- Lorde, A., 1984. *Sister Outsider: Essays and Speeches*. Crossing Press.
- McFarlane, C., 2009. Translocal assemblages: space, power, and social movements. *Geoforum* 40 (4), 561–567. <https://doi.org/10.1016/j.geoforum.2009.05.003>.
- McFarlane, C., 2012. *Learning the City: Knowledge and Translocal Assemblage*. Wiley-Blackwell, West Sussex, UK.
- Medina, C., 2010. "Reading across communities" in biliteracy practices: examining translocal discourses and cultural flows in literature discussions. *Read. Res. Q.* 45 (1), 40–60. <https://doi.org/10.1598/RRQ.45.1.3>.
- Medina, C., Wohlwend, K., 2014. *Literacy, Play and Globalization: Converging Imaginaries in Children's Critical and Cultural Performances*. Routledge, New York, NY.
- Mirra, N., Garcia, A., 2020. "I hesitate but I do have hope": youth speculative civic literacies for troubled times. *Harv. Educ. Rev.* 90 (2), 295–321.
- McKittrick, K., 2021. *Dear Science and Other Stories*. Duke University Press.
- Morrell, E., 2008. *Critical Literacy and Urban Youth: Pedagogies of Access, Dissent, and Liberation*. Routledge.
- Nasir, N., Bang, M., Yoshikawa, H., 2021. Possible futures: what might we accomplish in 25 years? Phi Delta Kappan. <https://kappanonline.org/reimagining-education-25-years-spencer-nasir-bang-yoshikawa/>.
- Philip, T., Garcia, A., 2013. The importance of still teaching the iGeneration: new technologies and the centrality of pedagogy. *Harv. Educ. Rev.* 83, 300–319.
- Rashid, M., Jocson, K.M., 2021. Postcolonial m/othering: poetics of remembering and writing as an invitation to rememory. *Cult. Stud. Crit. Methodol.* 21 (5), 401–409. <https://doi.org/10.1177/15327086211028675>.
- Sablo Sutton, S., 2004. Spoken word: performance poetry in the black community. In: *What They Don't Learn in School: Literacy in the Lives of Urban Youth*. Jabari Mahiri, Peter Lang, pp. 213–242.
- Soep, E., 2014. *Participatory Politics: Next-Generation Tactics to Remake Public Spheres*. The MIT Press, Cambridge, MA.
- Thomas, E.E., 2019. *The Dark Fantastic: Race and the Imagination from Harry Potter to the Hunger Games*. NYU Press, New York, NY.
- Toliver, S.R., 2021. Freedom dreaming in a broken world: The Black radical imagination in Black girls' science fiction stories. *Research in the Teaching of English* 56 (1), 85–106.
- Tuck, E., 2009. Suspending damage: a letter to communities. *Harv. Educ. Rev.* 7 (3), 409–428.
- Watson, V., 2013. Censoring freedom: community-based professional development and the politics of profanity. *Equity Excell. Educ.* 46 (3), 387–410.
- Weinstein, S., West, A., 2012. Call and responsibility: critical questions for youth spoken word poetry. *Harv. Educ. Rev.* 82 (2), 282–302.
- Wissman, K.M., Wiseman, A.M., 2011. "That's my worst nightmare": poetry and trauma in the middle school classroom. *Pedagogies* 6 (3), 234–249.
- Ybarra, M.G., 2021. "Because we have lived it": Chicana/Latina Youth Multimodal Literacies in youth participatory action research. *Read. Res. Q.* <https://doi.org/10.1002/rrq.455>.
- Yodell, D., 2011. *School Trouble: Identity, Power and Politics in Education*. Routledge, New York, NY.

Youth out-of-school (critical) literacies, texts, and audiences

Kathleen A. Hinchman, Syracuse University, Syracuse, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Youth Out-of-school (critical) literacies, texts, and audiences	242
Youth's literacies	242
Youth's texts	244
Youth's audiences	245
Implications	246
References	246

Youth Out-of-school (critical) literacies, texts, and audiences

Across the globe and at every moment of each day, one can find youth reading, composing, listening to, and viewing many different forms of print, multimodal, and video texts. They consume, compose, and critique texts when ideas matter to them, demonstrating multiple, often quite varying literacies around family, face-to-face and digital friends, schoolmates, fan groups, and workplace colleagues. This means that the literacy practices they exercise cannot be painted with the same brush: their literacies, texts, and audiences differ by their ages, gender identifications, races, classes, interests, experiences, geographies, resources, and cultures. Their literacies, texts, and audiences reflect the privileges, oppressions, aspirations, disappointments, inclusions, and exclusions demonstrated and experienced by their respective social groups.

For half a century, scholars have documented the social nature of youth's out-of-school literacy practices, distinguishing them from the academically-driven practices of school by purpose and not necessarily physical locale, since youth deploy such literacies when they are in school, too (Hull and Schultz, 2002). The not-so-subtle parenthetical play on words in this essay's title, Youth Out-of-School (Critical) Literacies, Texts, and Audiences, refers, first, to the idea that youth's out-of-school literacies, texts, and audiences are central, aka critical, to how they draw on literacy practices to situate themselves in non-academic physical and social spaces. The word, critical, also references the literacies youth use to judge explanations offered in texts they consume. The parentheses signal the tentative, fluid character of youth identities and their literacy practices, which shift shape with evolving pop culture, technology, and affinity groups. In the following sections, I summarize what is known about youth literacies, texts, and audiences. This is followed by a foray into implications for research, instructional practice, and policy, and a conclusion.

Youth's literacies

Curwood (2013) captured one category of youth literacies in an analysis of the fan fiction composed by one youth to connect distinctive features of the *Harry Potter* series with those of *The Hunger Games* series, popular fantasy novels and film series of the time. Her study participant, 16-year-old Cassie, drew on various elements of the ten texts of both literature series to design videos and interactive texts, share news in print and podcasts, and manage user accounts for a site dedicated to *Harry Potter*. Because of her well-known fan fiction prowess, Cassie was invited by another fan fiction website manager to write about an assigned character from *Harry Potter*, Colin Creevey, designing the story format, storyline, and outcome. Curwood shared Cassie's response to this invitation:

Cassie felt it was an incredible honor to be asked to contribute Colin's story. As a fan herself, she knew that her readers would be smart, be concerned about authenticity, and have high expectations of any *Harry Potter* and *Hunger Games* remix. As a writer, Cassie was motivated by the opportunity to learn from mentor texts penned by J.K. Rowling and Suzanne Collins. She specifically talked about her goals as a writer, and she readily drew on resources that existed within the fandoms and affinity spaces. Moreover, Cassie was eager to share her work with a global audience. Through the process of remixing, Cassie deepened her knowledge of the *Harry Potter* and *Hunger Games* canons at the same time that she honed her craft as a writer (pp. 88–89).

Curwood's description of Cassie's response acknowledges the youth's written fiction composing literacies, including purposeful uses of mentoring texts, narrative structure, and sense of audience, skills Cassie acquired by writing for and consuming the fan fictions of similar websites.

Millions of youth across the globe demonstrate such literacies in fan fiction compositions, along with acumen in an array of other social media contexts. Despite this, US youth are among those who are described by some as possessing deficient literacy skills

en masse (Salinger, 2011). This stance reflects one common understanding of the word, literacy, which refers to individuals' ability to read and write as reflected by standardized assessment performance. Brian Street (1984) labeled this view universalist or autonomous because of its contention that the most successful readers draw on early language experiences and school reading and writing instruction to learn uniform skills in predictable stages (Chall, 1983). According to this view, literacy instruction should focus on teaching and measuring these skills. When students do not follow the expected pattern of skill acquisition, they are deemed deficient. Based on tasks that involve reading and answering questions about small samplings of texts, the results of these assessments serve as reductive proxies for the more purposeful, wide-ranging literacies of youth like Cassie.

The word, youth, is typically used to refer to young people who are past puberty but not yet adults (Lesko, 2012). Like its frequent synonym, adolescent, the construct arrived with the industrial revolution and the dawn of the field of psychology at the turn of the last century (Hall, 1905). Common usage of both terms confers a transitional, "not yet" status that often results in advice to "wait until you are an adult to understand the real world." This generalization does not acknowledge the stance I take in this chapter, which embraces as important what youth learn from childhood interactions and the communities within which they find themselves. This perspective also embraces the ways youth literacy practices shape and are shaped by the evolving, intersecting identities youth adopt, enact, and shift frequently, most with so-called real-world repercussions that depend on how and with whom they see themselves at any given moment (Alvermann and Moje, 2019; Gee, 2015b; Smith et al., 2022).

Scholars who embrace the out-of-school literacies referenced in this chapter's title hold more complex, socially situated views about the nature of literacy. These views acknowledge that uses of communicative symbolic practices, including those involved in reading, writing, and viewing, are shaped by the purposes for which they are developed and deployed—a social practices viewpoint which has been gaining acceptance since the argument was first made in the 1980s (Scribner and Cole, 1981; Street, 1984; Wertsch, 1991). For instance, the United Nations Educational, Scientific and Cultural Organization (UNESCO) has adopted the multidimensional, purpose-driven realization that, "The concept of literacy has evolved from basic reading, writing and numeracy skills to broader notions such as functional literacy and a foundation for lifelong learning" (UNESCO, 2019). This view is also aligned with the US National Academies of Sciences, Engineering, and Medicine (2018) conceptualization of all learning as socially and culturally situated, a situatedness that means literacy practices differ across social and cultural groups (Barton and Hamilton, 2000; Cazden et al., 1996).

There are multiple and evolving theoretical explanations of the social nature of literacy (Burnett et al., 2014). The underlying premises of social theories of literacy hold that very young children drive their own literacy acquisition due to their immersion and help communicating in the socially- and culturally-driven verbal and symbol systems that surround them in families and communities (Lee et al., 2020; Sulzby and Teale, 1991). They learn the code breaking, meaning making, text use, and text analysis skills and strategies of print literacies with school instruction even as they simultaneously develop out-of-school literacy practices that vary across social and geographic contexts (Freebody and Luke, 1990). They refine their literacy skills and strategies with their use of online and offline school texts as well as in the texts found in their homes and favored community settings (Nelson et al., 2020; Vetter et al., 2021), skills and strategies that, like Cassie's, are not well represented by standardized assessments (Brandt, 2015; Mislevy, 2016; Willis, 2012).

Social perspectives toward youth literacies acknowledge how such literacies are embedded with the shifting, evolving identities they enact at home, in school, and with peers (Blaikie, 2021; Paris, 2011; Vetter et al., 2021). From their earliest years, they develop literacies from interactions with family and other caregivers using such shared texts as directions, cereal boxes, religious texts, children's books, and community signs (Heath, 1983). These literacies continue to develop as youth's social worlds expand, sometimes from reading and writing in school, but also from other affinity groups, including church groups (McMillon, 2017; Sarroub, 2011), writing groups (Blackburn, 2003; Watson and Beymer, 2019), and gangs (Moje, 2000). Youth become aware of multiple intersecting linguistic and social systems. They learn to navigate these systems within and across assemblages of social, cultural, linguistic, political, and geographic, often conflicting boundaries, including transnational and translingual boundaries (Baker-Bell, 2020; Guerrero-Arias, 2021; Iwase, 2014; Kinloch et al., 2020; Price-Dennis and Muhammad, 2021; Skerrett and Omogun, 2020; Tomei and Hollington, 2020). They develop dialects, languages, and literacies for the use and production of the texts they use most frequently because it is in their interest to do so (Creese and Blackledge, 2010; Nkadameng and Makalela, 2015; Stroud and Wee, 2012).

Youth also develop literacies in familiar settings that can allow them to note whose perspectives are represented, or not, when texts are brought into communities, as well as the implications of such insights (Braden, 2020). These critical literacies assume multiple forms, depending on where and with whom youth spend their time (Smith, 2020): Cassie, at the beginning of this section, demonstrated her understanding of perspective as she remixed elements from various literary mentor texts. Vasquez and her colleagues described such critical literacies as "looking at an issue or topic in different ways, analyzing it, and suggesting possibilities for change and improvement" (Vasquez et al., 2019, p. 300). Freire (1985) offered the following explanation:

The act of reading cannot be explained as merely reading words since every act of reading words implies a previous reading of the world and a subsequent rereading of the world. There is a permanent movement back and forth between "reading" reality and reading words—the spoken word too is our reading of the world. We can go further, however, and say that reading the word is not only preceded by reading the world, but also by a certain form of writing it or rewriting it. In other words, of transforming it by means of conscious practical action. For me, this dynamic movement is central to literacy (p. 18).

Youth's texts

Like youth literacies before the dawn of the social media age (Finders, 1997; Moje, 2000), youth's present-day text production and consumption are social, wide-ranging, and rapidly evolving, facilitated by the global connections made possible by the Internet. As an example of one youth's novel mix of media and ideas, Wargo (2018) described how one youth, Andi, used a GoPro video camera and iPhone voice memos to compose a multimodal text. Wargo summarized:

As a form of resistance, Andi composed a text that used sound to amplify the silences that formal education presented to her as a lesbian-identifying Latina. Through walking the city and writing it with sound, she presented a topography and depth of experience that disoriented the listener and onlooker. From upside-down visuals to the ambient acoustics of homophobia and misogyny, Andi delivered a multimodal walk that is both textually interpretive and affectively felt... a practice that worked to reorient the space, place, and perception of schooling; who she was as a survivor; and how she navigated personal issues of injustice (pp. 21–22).

Wargo illustrated how Andi's multiple and evolving literacies were used to compose a compelling, identity-threaded new form of text, composed because the ideas mattered.

The Chinese are generally credited with inventing paper and the earliest forms of printing, however, human invention of written language occurred multiple times even before the invention of paper, with cave paintings, strings, and shells (Ferrara, 2022). Gutenberg's refinement of an efficient printing press in the mid-Fifteenth Century initiated a proliferation of printed texts that continues today (Kurlansky, 2016). The term, text, is commonly used today to describe books, most often school textbooks. However, scholars exploring the social nature of literacy in the late Twentieth Century began to apply a broader definition, using the term to refer to multiple forms of symbolic representations that readers decode, interpret, and compose and including print, video, photographs, paintings, and music, and including digital and mixed, multimodal forms (Fish, 1980; McLuhan, 1966).

Contemporary youth texts evolve with society and include print on paper or published in digital form, film, or ever-changing social media platforms, such as Facebook, Instagram, SnapChat, and YouTube, that are accessible from computers, smart phones, tablets, gaming systems, and libraries (Boyd, 2014; Boyd and Brock, 2014; Coiro et al., 2008; Lankshear and Knobel, 2011). When viewed as social and cultural representations, such wide-ranging forms as memoirs, comics, anime, and multimodal composition illustrate how youth's consumption preferences and composing practices are threaded with identity (Rogers and Marshall, 2012; Rowsell and Pahl, 2007). Alvermann and Sanders (2019) described these literacies as "new in the sense that social, economic, cultural, intellectual, and institutional changes are continually at work" (p. 1), noting that youth who grew up in the digital age often enact multiple identities in varied forums, easily shifting literacy identities as they move among affinity groups.

To counter periodic claims of today's youth's a- and illiteracy, literacy scholars who are interested in social practices perspectives have devoted much attention to capturing youth's purposeful and prolific use and production of texts outside of school. Studies have described youth 'zines (Finders, 1997; Rogers and Winters, 2010), gang tags (Moje, 2000; MacGillivray and Curwen, 2007), self-advocating essays and poetry (Blackburn, 2003; Gudjonsson, 2018), and religious texts (Sarroub, 2011; Weyand and Juzwik, 2020). Social media have become an especially important site for youth's collaborative and inventive text use (Burnett and Merchant, 2020), including instant messaging (Lam, 2009; Lewis and Fabos, 2005), multiplayer online gaming (Abrams and Lammers, 2017; Grzesiak, 2018), anime (Chandler-Olcott and Mahar, 2003; Fukunaga, 2006), songs (DeRoo and Watson, 2020), hip-hop poetry (Castillo, 2021; Alim, 2011), fanfiction (García, 2013; Lammers, 2016), drama (Kim and Omerbašić, 2017), and TikTok videos (Jerasa and Boffone, 2021; Nam, 2020). The literacies invoked by these varied texts are typically developed and shared among group members, often across geographical and political boundaries (e.g., Ellison, 2014).

Texts take on complex social meanings tied to where, when, and who produces and consumes them (Collins and Blot, 2003). Their forms are shaped by the uses individuals find for them and the affinity group literacies used to compose, distribute, and read them. Composers produce them from their world experiences, and readers interpret them through lenses they develop from their own experiences. Social disparity can result in inequitable participation in society when only some individuals have more access than others and, as a result, are invited, or not, to learn the literacies needed to access a group's texts. This means that, despite the ubiquity of youth's literacies, each youth's access to the literacies tied to their aspirations is uneven at best and selectively oppressive at worst (Alvermann, 2016; Bloome and Green, 2015; Gee, 2015b).

As noted earlier in this piece, the term, critical, can be used to describe texts youth consume and produce as well as the ways they can and, many would say, should consume texts they encounter each day. This is especially true of popular media texts generated by commercial entities, such as Meta, Twitter, and TikTok, that do always not have youth's best interests at heart (Luke, 2019). When youth consume and produce texts to address purposes, topics, and tasks they find compelling, meaningful, and familiar, they are more likely to be able and willing to weigh issues of accuracy, perspective, and representativeness. They are also more likely to search out texts to enlighten themselves about the chasms and possibilities of competing perspectives, whether the texts are videos, novels, or fanfiction websites (Guzzetti and Lesley, 2016).

Youth's audiences

When [Muhammad and McArthur \(2015\)](#) conducted a study to explore girls' perceptions of cultural media representations of Black people, one young woman's response recognized that such representations did not seem to understand her and others in their study group, saying:

Who can tell our stories best but us? We determine how we think of ourselves, not by what other people think of us. But our actions affect others' perceptions. If you think that we are not portrayed in media badly, you obviously don't spend enough time watching television, reading magazines, or listening to music. Other people who are not black think we are always out to commit violent actions. That's just media. We can take control by starting our own representations of who we are as young black girls leaving an impact on other young girls' lives (pp. 137–138).

Muhammad and McArthur explained that the young woman, Dahlia, “wrote back against such views in the last lines expressing that she and other Black girls can reclaim authority to set their own standards of representation” (p. 138). All the young women in this study wrote for similar purposes, to address an audience of people like them and the world at large to counter negative media stereotypes that did not consider them accurately as an audience.

Audience, the final central construct of this *Encyclopedia* chapter, is a term that is commonly used to describe those who read, view, or listen to others' compositions. As such, connecting with audience is a consideration for composers across media. Social perspectives toward literacy added nuance to scholarly understandings of how audience works. [Fish \(1980\)](#) long ago introduced the idea of instantiated understandings—written communication—that occurs when individuals can draw on shared cultural knowledge of symbol systems. Audiences are composers' interpretive communities across modalities. Youth serve as audiences for others' texts and, in turn, compose texts for print, multimodal, or digital audiences that emulate those texts that they most consume.

Youth become audiences to a seemingly infinite range of texts as producers and consumers. Their audiences and authors of texts they consume may include individuals known, unknown, near, and far. Youth are primary users of a range of communication technologies and social media, showing more savvy in their selected venues than many adults and offering a blueprint of possibilities in new forms of text. They are often the first to recognize the potential of new digital fora, becoming early adopters of Facebook, and then Twitter, Instagram, Snapchat, YouTube Gamer platforms, and TikTok as outlets for their compositions and consumption ([Alvermann, 2004](#)).

Youth's literacies are, thus, tied to their service as audience for others' literacy enactments. Their literacy practices are honed among individuals who serve as one another's audiences in person, print, or digital media. Youth, thus, serve as audiences for others' videos, memes, and blog posts or for similar texts by unidentified others who participate in their social networks. They draw on varied text forms and social media platforms according to their understandings of the conventions of their varied social groups. Practices evolve according to family, school, and community cultural practices, traditions of worship, or shared genuine or manipulative interest in music or reading genres, video and tabletop games, jobs, or hobbies, with groups shaping their practices to suit communication purposes. Variations among groups evolve from differences in groups' purposes for communicating. As a result, youth's literacies outside school are in a near constant state of fluctuation, along with their identities, as they enter and exit social groups. As they expand their networks they are increasingly likely to be duped, at least until they learn more nuanced aspects of communication within these networks ([Garcia, 2020](#); [Gee, 2015a](#); [Kress, 2003](#)).

Youth consume and produce texts to experiment with enacting varied identities, even producing texts for which they are their own audiences as they try on different personas ([Lewis et al., 2007](#)). Many youth write to each other and the world at large to demonstrate their identities in response to stereotypes of others, including popular media, as in [Muhammad and McArthur's \(2015\)](#) study at the beginning of this section. Researchers have also recorded how youth performed texts they created to disrupt observed LGBTQ-related inequities ([Schey, 2020](#)) and drew on religious and secular texts to mediate life as Yemeni American females ([Sarroub, 2002](#)) and Christian fundamentalists ([Eakle, 2007](#); [Skerrett, 2014](#)). This research was followed by additional work exploring how social affiliations, including race, gender, social class, dis/ability, and culture, and the intersections of such affiliations (e.g., Black girls, cognitive disabilities of those in poverty), have shaped youth's texts and audiences historically and in present times ([Collins and Ferri, 2016](#); [Kendi and Blain, 2021](#); [Kinloch, 2015](#); [Muhammad and Haddix, 2016](#); [Tatum, 2008](#); [Ybarra, 2020](#)).

Youth continue to refine literacy practices over time to address audiences of both real and imagined affinity groups. They draw on myriad media and popular culture discourses that are shaped, in turn, by innumerable commercial and political interests ([Alvermann, 2004](#); [Alvermann et al., 2003](#)), literacies that may be coopted by such global transmedia enterprises as those involving Harry Potter, Star Wars, Minecraft, and Fortnite ([Fung et al., 2019](#); [Kraidy, 2017](#); [Lemke, 2009](#); [Ünver, 2018](#)), which, in turn, inform local cultural, political, and heritage projects ([Burkholder, 2017](#); [Freeman and Proctor, 2018](#)). Some youth figure out how to restructure cooptation by composing elaborate counter-narratives, as in fanfiction characters drawn from popular literature ([Duggan, 2021](#); [Lammers, 2016](#); [Sugihartati, 2020](#)), videogame paratexts posted on YouTube and featuring gamer footage with player commentary ([Burwell and Miller, 2016](#)), and as exemplified by the young woman's composition at the beginning of this section. These exercises of critical literacies were described by [Rogers and colleagues \(2014\)](#) as “expressions of resistance to the inheritance of the broken promises of democratic citizenship and their ability to imagine new possibilities for public engagement” (p. 2).

Implications

Along with other aspects of language and communication, youth develop literacies they need for successful communication within their varied, often intersecting, affinity groups. The texts youth generate and consume are myriad and ever-changing. They develop critical literacies as they situate themselves in communities, develop insight about how others' texts are situating them, and gain the sense of audience needed to navigate these experiences successfully. These literacies are critical, too, because they allow youth to navigate within and across social groups. When youth initiate communication and their intended audience does not respond as they predicted, it is likely because there is something about the speech community, modality, or symbol systems they do not understand.

One of current society's greatest youth stereotypes is that they are "all" on the most popular social media software platform *du jour*, chasing each other from Facebook to Snapchat to Instagram to YouTube to TikTok as adults in their life share their applications. Predictably, reality is more complex, and some youth have more success than others with the shifting involved in navigating multiple communities of practice (Knobel and Lankshear, 2007; Vasudevan et al., 2014). Others elect to be less visible in some contexts than others, and still others are forced out of groups when hegemonic group practices promote bullying and other forms of exclusion (Nilan et al., 2015).

Researchers who study out-of-school literacies have engaged in valiant efforts to understand intersections of literacies, texts, audiences, and identity that are invoked by affinity groups. Digital exclusion scholars have begun to describe the relative nature of intersections of personal circumstances, social group processes, and exclusionary social structures. Gaining the insider access needed for nuanced insights about why some aspiring group members are marginalized or excluded has been more problematic, sometimes resulting in ill-conceived efforts to include individuals who are marginalized because of gender identification, race, social class, or disability that worsen the marginalization (Setchell et al., 2021). More work in this area is needed if youth are to develop abilities needed to access to all their aspirational groups (Helsper, 2017).

There have been multiple calls for importing youth's out-of-school literacies to school to heighten youth's school engagement. Alvermann, for instance, argued against separating out-of-school and in-school literacies because she saw educators organizing multiliteracies projects (e.g., book clubs, web searches) that invoked youth's motivation and self-efficacy with the same level of engagement as has often been observed in youth's out-of-school literacy practices (Alvermann and Moore, 2011). Alvermann and colleagues (2012) have also made the same argument in reverse, demonstrating how youth's engagement with so-called out-of-school literacy practices rival the most impactful in-school literacy practices, erasing the divide between the two settings.

Even though popular media frequently suggests this, it is probably as bad an idea for educators to appropriate youth's out-of-school literacies for use in school as it has been for adults to appropriate youth's social media networks, chasing youth to ever newer, more exclusive platforms. However, educators would also be foolish to underestimate or fail to appreciate the literacy skills and strategies youth develop within their outside-of-school affinity groups. Pedagogies that offer youth the chance to utilize their community-based literacies to shape their own learning—as well as the learning of the educators who serve them—can also help us to reshape schools and literacies in new, more relevant, inclusive ways (Yosso, 2005). Scholars have begun to offer multiple examples of educators are doing just that, individuals who design classroom inquiry to incorporate youth's resources to reshape school tasks and literacies, inviting these youth to be the next generation of scientists, historians, game engineers, literary theorists, and plumbers, reshaping literacies and the youth's aspirations (e.g., Alford, 2021; Chandler-Olcott et al., 2016; Hull and Stornaiuolo, 2010; Learned, 2018; Mills et al., 2018; Mirra and Garcia, 2021; Moje, 2015; Muhammad, 2015; Pandya et al., 2022; Unsworth et al., 2022; Vasudevan, 2020). Educators need multiple models of instruction that incorporates such insights for various youth age groups across disciplines. Policymakers need to embrace such models in lieu of the reductive, assessment-driven approaches to school that are now in vogue.

There is danger in inviting youth to engage in such inquiry without also figuring out how to trigger their deployment of critical literacies. Such literacy practices can help them to understand, accept, or reject various sides of any new argument in any context with empathy and insight. Luke (2019) is among those who have recommended that educators invite youth to critique all texts, including those that represent shared points of view and those in opposition, by drawing on such questions as, "How does language, text, discourse and information make a difference? For whom? In what material, social, and consequential ways? In whose interests? According to what patterns, rules, and in what institutional and cultural sites?" (p. 351). Such critical literacies will allow youth to shape and navigate their own and other's participation in all aspects of all the world, including school, in more relevant and equitable ways.

References

- Abrams, S.S., Lammers, J.C., 2017. Belonging in a videogame space: bridging affinity spaces and communities of practice. *Teach. Coll. Rec.* 119 (12), 1–34. <https://doi.org/10.1177/016146811711901101>.
- Alford, J., 2021. *Critical Literacy with Adolescent English Language Learners: Exploring Policy and Practice in Global Contexts*. Routledge, New York, NY.
- Alim, H.S., 2011. Global ill-literacies: Hip hop cultures, youth identities, and the politics of literacy. *Rev. Res. Educ.* 35 (1), 120–146. <https://doi.org/10.3102/0091732X10383208>.
- Alvermann, D.E., 2004. Media, information communication technologies, and youth literacies: a cultural studies perspective. *Am. Behav. Sci.* 48 (1), 78–83. <https://doi.org/10.1177/0002764204267271>.
- Alvermann, D.E., 2016. Literacies and identity. In: Peters, M. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore. https://doi.org/10.1007/978-981-287-532-7_116-1.

- Alvermann, D.E., Moje, E.B., 2019. A relational model of adolescent literacy instruction: disrupting the discourse of "every teacher a teacher of reading". In: Alvermann, D.E., et al. (Eds.), *Theoretical Models and Processes of Literacy*. Routledge, New York, NY, pp. 362–380.
- Alvermann, D.E., Moore, D.W., 2011. Questioning the separation of in-school from out-of-school contexts for literacy learning: an interview with Donna E. Alvermann. *J. Adolesc. Adult Lit.* 55 (2), 156–158. Available at: <https://www.jstor.org/stable/41309671>. (Accessed 10 April 2022).
- Alvermann, D.E., Sanders, R.K., 2019. Adolescent literacy in a digital world. In: Hobbs, R., et al. (Eds.), *The International Encyclopedia of Media Literacy*. Wiley-Blackwell, Hoboken, NJ, pp. 1–6.
- Alvermann, D.E., Xu, S.H., Carpenter, M., 2003. Children's everyday literacies: intersections of popular culture and language arts instruction. *Lang. Arts* 81 (2), 145–154.
- Alvermann, D.E., et al., 2012. Adolescents' web-based literacies, identity construction, and skill development. *Lit. Res. Instr.* 51 (3), 179–195. <https://doi.org/10.1080/19388071.2010.523135>.
- Baker-Bell, A., 2020. *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*. Routledge, New York, NY.
- Barton, D., Hamilton, M., 2000. Literacy practices. In: Barton, D., Hamilton, M., Ivanic, R. (Eds.), *Situated Literacies: Reading and Writing in Context*. Routledge, New York, NY, pp. 7–14.
- Blackburn, M., 2003. Exploring literacy performances and power dynamics at the loft: "queer youth reading the world and the word". *Res. Teach. Engl.* 37 (4), 467–490. Available at: <https://www.jstor.org/stable/40171600>. (Accessed 10 April 2022).
- Blaikie, F. (Ed.), 2021. *Visual and Cultural Identity Constructs of Global Youth and Young Adults: Situated, Embodied and Performed Ways of Being, Engaging and Belonging*. Routledge, New York, NY. <https://doi-org.libezproxy2.syr.edu/10.4324/9781003055822>.
- Bloome, D., Green, J., 2015. The social and linguistic turns in studying language and literacy. In: Rowsell, J., Pahl, K. (Eds.), *The Routledge Handbook of Literacy Studies*. Routledge, New York, NY, pp. 19–34.
- Boyd, D., 2014. *It's Complicated: The Social Lives of Networked Teens*. Yale University Press, New Haven, CT.
- Boyd, F.B., Brock, C.H. (Eds.), 2014. *Social Diversity within Multiliteracies*. Routledge, New York, NY, pp. 154–167.
- Braden, E.G., 2020. Navigating black racial identities: literacy insights from an immigrant family. *Teach. Coll. Rec.* 122 (13), 1–26. <https://doi.org/10.1177/016146812012201310>.
- Brandt, D., 2015. *The Rise of Writing: Redefining Mass Literacy*. Cambridge University Press, Cambridge, MA.
- Burkholder, C., 2017. "Before Occupy Central, I wasn't concerned": examining participatory visual research for social change with Hong Kong-based Filipina youth. *Lang. Lit.* 19 (3), 56–73. <https://doi.org/10.20360/G2737B>.
- Burnett, C., Merchant, G., 2020. *Undoing the Digital: Sociomaterialism and Literacy Education*. Routledge, New York, NY.
- Burnett, C., et al. (Eds.), 2014. *New Literacies Around the Globe*. Routledge, New York, NY.
- Burwell, C., Miller, T., 2016. Let's play: exploring literacy practices in an emerging videogame paratext. *E-Learn. Digit. Media* 13 (3–4), 109–125. <https://doi.org/10.1177/2042753016677858>.
- Castillo, C.E., 2021. *Rap as Multiliteracy: Understanding Hip Hop Literacies across Spaces*. Ph.D. Thesis. The University of Wisconsin-Madison.
- Cazden, C., et al., 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Chall, J.S., 1983. *Stages of Reading Development*. McGraw-Hill, New York.
- Chandler-Olcott, K., Doerr, H.M., Hinchman, K.A., Masingila, J.O., 2016. Bypass, augment, or integrate: how secondary mathematics teachers address the literacy demands of standards-based curriculum materials. *J. Lit. Res.* 47 (4), 439–472. <https://doi.org/10.1177/1086296X16632040>.
- Chandler-Olcott, K., Mahar, D., 2003. "Tech-savviness" meets multiliteracies: exploring adolescent girls' technology-mediated literacy practices. *Read. Res. Q.* 38 (3), 356–385. Available at: <https://www.jstor.org/stable/4151823>. (Accessed 10 April 2022).
- Coiro, J., et al., 2008. Central issues in new literacies and new literacies research. In: Coiro, J., et al. (Eds.), *Handbook of Research on New Literacies*. Lawrence Erlbaum Associates/Taylor & Francis Group, New York, NY, pp. 1–21.
- Collins, J., Blot, R., 2003. *Literacy and Literacies: Texts, Power, and Identity*. Cambridge University Press, Cambridge.
- Collins, K., Ferri, B., 2016. Literacy education and disability studies: reenvisioning struggling students. *J. Adolesc. Adult Lit.* 60 (1), 7–12. <http://doi.10.1002/jaal.552>.
- Creese, A., Blackledge, A., 2010. Translanguaging in the bilingual classroom: a pedagogy for learning and teaching. *Mod. Lang. J.* 94 (1), 103–115. <https://doi-org.libezproxy2.syr.edu/10.1111/j.1540-4781.2009.00986.x>.
- Curwood, J.S., 2013. Fan fiction, remix culture, and the Potter Games. In: Frankel, V.E. (Ed.), *Teaching with Harry Potter*. McFarland & Co., Jefferson, NC, pp. 81–92.
- Deroo, M.R., Watson, V.W., 2020. "Air I breathe": songwriting as literacy practices of remembrance. *J. Literacy Res.* 52 (2), 158–179. <https://doi.org/10.1177/1086296X20915518>.
- Duggan, J., 2021. "Worlds...[of] contingent possibilities": genderqueer and trans adolescents reading fan fiction. *Televis. N. Media* 1527–4764. <https://doi.org/10.1177/15274764211016305>.
- Eakle, A.J., 2007. Literacy spaces of a Christian faith-based school. *Read. Res. Q.* 42 (4), 472–510. Available in: <https://doi.org/10.1598/RRQ.42.4.3>.
- Ellison, T.L., 2014. Digital ontologies of self: two African American adolescents co-construct and negotiate identities through the Sims 2. *Digit. Cult. Educ.* 6 (4), 334–351.
- Ferrara, S., 2022. *The Greatest Invention: A History of the World in Nine Mysterious Scripts*. Translated from the Italian by Todd Portnowitz. Farrar, Straus and Giroux, New York, NY.
- Finders, M.J., 1997. *Just Girls: Hidden Literacies and Life in Junior High*. Teachers College Press, New York, NY.
- Fish, S.E., 1980. *Is There a Text in This Class?: The Authority of Interpretive Communities*. Harvard University Press, Cambridge, MA.
- Freebody, P., Luke, A., 1990. Literacies programs: debates and demands in cultural context. *Prospect* 5 (3), 7–16.
- Freeman, M., Proctor, W. (Eds.), 2018. *Global Convergence Cultures: Transmedia Earth*. Routledge, New York, NY.
- Freire, P., 1985. Reading the world and reading the word: an interview with Paulo Freire. *Lang. Arts* 62 (1), 15–21.
- Fukunaga, N., 2006. "Those anime students": foreign language literacy development through Japanese popular culture. *J. Adolesc. Adult Lit.* 50 (3), 206–222. <https://doi.org/10.1598/JAAL.50.3.5>.
- Fung, A., Pun, B., Mori, Y., 2019. Reading border-crossing Japanese comics/anime in China: cultural consumption, fandom, and imagination. *Global Media China* 4 (1), 125–137. <https://doi.org/10.1177/2059436419835379>.
- García, J.D., 2013. *The Chican@ Hip Hop Nation: Politics of a New Millennial Mestizaje*. Michigan State University, East Lansing, MI.
- García, A., 2020. Gaming literacies: spatiality, materiality, and analog learning in a digital age. *Read. Res. Q.* 55 (1), 9–27. <https://doi.org/10.1002/rrq.260>.
- Gee, J.P., 2015a. Discourse analysis of games. In: Jones, R.H., Chik, A., Hafner, C.A. (Eds.), *Discourse and Digital Practices*. Routledge, New York, pp. 30–39.
- Gee, J.P., 2015b. *Social Linguistics and Literacies: Ideology in Discourses*. Routledge, New York.
- Grzesiak, M., 2018. Personal Brand Creation in the Digital Age. Palgrave Macmillan, Cham, CH. Available at: <https://link.springer.com/book/10.1007/978-3-319-69697-3>. (Accessed 10 April 2022).
- Gudjonsson, B.E.H., 2018. "I wish color had no history": an exploration of identity through student performance poetry. *Ethnogr. Educ.* 13 (3), 322–339. <https://doi.org/10.1080/17457823.2017.1422134>.
- Guerrero-Arias, B.E., 2021. "What's Blackness got to do with all this?" Racializing literacy assemblages. *Read. Res. Q.* 56 (4), 659–672. <https://doi.org/10.1002/rrq.437>.
- Guzzetti, B., Lesley, M. (Eds.), 2016. *Handbook of Research on the Societal Impact of Digital Media*. IGI Global, Hershey, PA.
- Hall, G.S., 1905. *Adolescence: Its Psychology and its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion and Education (Volume 2)*. D. Appleton and Company, New York, NY.
- Heath, S.B., 1983. *Ways with Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.

- Helsper, E.J., 2017. The social relativity of digital exclusion: applying relative deprivation theory to digital inequalities. *Commun. Theor.* 27 (3), 223–242. <https://doi.org/10.1111/comt.12110>.
- Hull, G., Schultz, K., 2002. *School's Out: Bridging Out-Of-School Literacies with Classroom Practices*. Teachers College Press, New York, NY.
- Hull, G.A., Stormaiuolo, A., 2010. Literate arts in a global world: reframing social networking as cosmopolitan practice. *J. Adolesc. Adult Lit.* 54 (2), 84–96. <https://doi.org/10.1598/JAAL.54.2.1>.
- Iwase, M., 2014. *New Literacies, Japanese Youth, & Global Fast Food Culture: Exploring Critical Youth Agencies*. Ph.D. Thesis. Simon Fraser University.
- Jerasa, S., Boffone, T., 2021. BookTok 101: TikTok, digital literacies, and out-of-school reading practices. *J. Adolesc. Adult Lit.* 65 (3), 219–226. <https://doi.org/10.1002/jaal.1199>.
- Kendi, I.X., Blain, K.N., 2021. *Four Hundred Souls: A Community History of African America 1619–2019*. Random House, New York, NY.
- Kim, G.M., Omerbašić, D., 2017. Multimodal literacies: imagining lives through Korean dramas. *J. Adolesc. Adult Lit.* 60 (5), 557–566. <https://doi.org/10.1002/jaal.609>.
- Kinloch, V., 2015. *Harlem on Our Minds: Place, Race, and the Literacies of Urban Youth*. Teachers College Press, New York, NY.
- Kinloch, V., Burkhard, T., Graham, D., 2020. Storying youth lives: centering equity in teaching and teacher education. *Int. J. Qual. Stud. Educ.* 33 (1), 66–79. <https://doi.org/10.1080/09518398.2019.1678779>.
- Knobel, M., Lankshear, C. (Eds.), 2007. *A New Literacies Sampler*. Peter Lang, New York, NY.
- Kraidy, M.M., 2017. The projectile image: Islamic State's digital visual warfare and global networked affect. *Media Cult. Soc.* 39 (8), 1194–1209. <https://doi.org/10.1177/0163443717725575>.
- Kress, G., 2003. *Literacy in the New Media Age*. Routledge, New York, NY.
- Kurlansky, M., 2016. *Paging through History*. W.W. Norton & Company, New York, NY.
- Lam, W.S.E., 2009. Multiliteracies on instant messaging in negotiating local, translocal, and transnational affiliations: a case of an adolescent immigrant. *Read. Res. Q.* 44 (4), 377–397. <https://doi.org/10.1598/RRQ.44.4.5>.
- Lammers, J.C., 2016. "The Hangout was serious business": leveraging participation in an online space to design Sims fanfiction. *Res. Teach. Engl.* 50 (3), 309–332. Available at: <https://www.jstor.org/stable/24889924>. (Accessed 10 April 2022).
- Lankshear, C., Knobel, M., 2011. *New Literacies: Everyday Practices and Social Learning*, third ed. Open University Press/McGraw-Hill, Maidenhead & New York, NY.
- Learned, J.E., 2018. Doing history: a study of disciplinary literacy and readers labeled as struggling. *J. Literacy Res.* 50 (2), 190–216. <https://doi.org/10.1177/1086296X17746446>.
- Lee, C.D., Meltzoff, A.N., Kuhl, P.K., 2020. The braid of human learning and development: neuro-physiological processes and participation in cultural practices. In: Nasir, N.S., et al. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, New York, NY, pp. 24–43.
- Lemke, J., 2009. Multimodal genres and transmedia traversals: social semiotics and the political economy of the sign. *Semiotica* 2009 (173), 283–297. <https://doi.org/10.1515/SEMI.2009.012>.
- Lesko, N., 2012. *Act Your Age!: A Cultural Construction of Adolescence*, second ed. Routledge, New York, NY.
- Lewis, C., Fabos, B., 2005. Instant messaging, literacies, and social identities. *Read. Res. Q.* 40 (4), 470–501. <https://doi.org/10.1598/RRQ.40.4.5>.
- Lewis, C., Enciso, P.E., Moje, E.B. (Eds.), 2007. *Reframing Sociocultural Research on Literacy: Identity, Agency, and Power*. Routledge, New York, NY.
- Luke, A., 2019. Regrounding critical literacy. In: Alvermann, D.E., Unrau, N., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Reading*, seventh ed. Routledge, New York, NY, pp. 349–361.
- MacGillivray, L., Curwen, M.S., 2007. Tagging as a social literacy practice. *J. Adolesc. Adult Lit.* 50 (5), 354–369. Available at: <https://www.jstor.org/stable/40015478>. (Accessed 10 April 2022).
- McLuhan, H.M., 1966. *Understanding Media. Signet*, New York, NY.
- McMillon, G.T., 2017. School–university–community collaboration: building bridges at the water's edge. *J. Adolesc. Adult Lit.* 60 (4), 375–381. <https://doi.org/10.1002/jaal.604>.
- Mills, K., Stormaiuolo, A., Smith, A., Pandya, J.Z. (Eds.), 2018. *Handbook of Writing, Literacies, and Education in Digital Cultures*. Routledge, New York, NY.
- Mirra, N., Garcia, A., 2021. In search of the meaning and purpose of 21st-century literacy learning: a critical review of research and practice. *Read. Res. Q.* 56 (3), 463–496. <https://doi.org/10.1002/rq.313>.
- Mislevy, R.J., 2016. How developments in psychology and technology challenge validity argumentation. *J. Educ. Meas.* 53 (3), 265–292. <https://doi.org/10.1111/jedm.12117>.
- Moje, E.B., 2000. "To be part of the story": the literacy practices of gangsta' adolescents. *Teach. Coll. Rec.* 102 (3), 651–690. <https://doi.org/10.1177/016146810010200306>.
- Moje, E.B., 2015. Doing and teaching disciplinary literacy with adolescent learners: a social and cultural enterprise. *Harv. Educ. Rev.* 85 (2), 254–278. <https://doi.org/10.11763/0017-8055.85.2.254>.
- Muhammad, G.E., 2015. Searching for full vision: writing representations of African American adolescent girls. *Res. Teach. Engl.* 49 (3), 224–247.
- Muhammad, G.E., Haddix, M., 2016. Centering Black girls' literacies: a review of literature on the multiple ways of knowing of Black girls. *Engl. Educ.* 48 (4), 299–336.
- Muhammad, G.E., McArthur, S.A., 2015. "Styled by their perceptions": black adolescent girls interpret representations of Black females in popular culture. *Multicult. Perspect.* 17 (3), 133–140. <https://doi.org/10.1080/15210960.2015.1048340>.
- Nam, Y., 2020. "It's TikTok with no c!": the need for critical literacy curricula in the elementary classroom. *Int. J. Crit. Media Lit.* 2 (1), 103–119. <https://doi.org/10.1163/25900110-00201006>.
- National Academies of Sciences, Engineering, and Medicine, 2018. *How People Learn II: Learners, Contexts, and Cultures*. National Academies Press, Washington, DC. Available at: <https://nap.nationalacademies.org/catalog/24783/how-people-learn-ii-learners-contexts-and-cultures>. (Accessed 10 April 2022).
- Nelson, E.L., Perry, M., Rogers, T., 2020. Introducing offlineness: theorizing (digital) literacy engagements. *J. Lit. Res.* 52 (1), 101–107. <https://doi.org/10.1177/1086296X19898003>.
- Nilan, P., Burgess, H., Hobbs, M., Threadgold, S., Alexander, W., 2015. Youth, social media, and cyberbullying among Australian youth: "sick friends". *Soc. Med. Soc.* 1 (2). <https://doi.org/10.1177/2056305115604848>.
- Nkadameng, S., Makalela, L., 2015. Identity negotiation in a super-diverse community: the fuzzy languaging logic of high school students in Soweto. *Int. J. Sociol. Lang.* 2015 (234), 7–26. <https://doi.org/10.1515/ijsl-2015-0002>.
- Pandya, J.Z., et al. (Eds.), 2022. *The Handbook of Critical Literacies*. Routledge, New York, NY.
- Paris, D., 2011. *Language across Difference: Ethnicity, Communication, and Youth Identities in Changing Urban Schools*. Cambridge University Press, Cambridge.
- Price-Dennis, D., Muhammad, G.E. (Eds.), 2021. *Black Girls' Literacies: Transforming Lives and Literacy Practices*. Routledge, New York, NY.
- Rogers, T., Marshall, E., 2012. On the road: examining self-representation and discourses of homelessness in young adult texts. *J. Adolesc. Adult Lit.* 55 (8), 725–733. <https://doi.org/10.1002/JAAL.00087>.
- Rogers, T., Winters, K.L., 2010. Textual play, satire, and counter discourses of street youth 'zining practices'. In: Alvermann, D.E. (Ed.), *Adolescents' Online Literacies: Connecting Classrooms, Digital Media, and Popular Culture*. Peter Lang, New York, NY, pp. 91–108.
- Rogers, T., Winters, K., Perry, M., LaMonde, A., 2014. *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the Lives of Adolescents*. Routledge, New York, NY.
- Rowell, J., Pahl, K., 2007. Sedimented identities in texts: instances of practice. *Read. Res. Q.* 42 (3), 388–404. <https://doi.org/10.1598/RRQ.42.3.3>.
- Salinger, T., 2011. Addressing the "Crisis" in Adolescent Literacy. American Institutes for Research, Washington, DC. Available at: <https://www2.ed.gov/programs/sclp/finalcrisis.pdf>. (Accessed 10 April 2022).
- Sarroub, L.K., 2002. In-betweenness: religion and conflicting visions of literacy. *Read. Res. Q.* 37 (2), 130–148. Available at: <https://www.jstor.org/stable/748154>. (Accessed 10 April 2022).

- Sarroub, L.K., 2011. Discontinuities and differences among Muslim Arab-Americans: making it at home and school. In: Dantas, M.L., Manyak, P. (Eds.), *Home-school Connections in a Multicultural Society*. Routledge, New York, NY, pp. 92–109.
- Schey, R., 2020. 'Youths' literacy disidentifications in a secondary classroom: contesting transphobia through humor in role-playing. *Teach. Coll. Rec.* 122 (7), 1–42. <https://doi.org/10.1177/016146812012200712>.
- Scribner, S., Cole, M., 1981. Unpackaging literacy. In: Whiteman, M.F. (Ed.), *Writing: The Nature, Development, and Teaching of Written Communication*. Lawrence Erlbaum, Hillsdale, NJ, pp. 71–87.
- Setchell, J., Barlott, T., Torres, M., 2021. A socio-emotional analysis of technology use by people with intellectual disabilities. *J. Intellect. Disabil. Res.* 65 (2), 149–161. <https://doi.org/10.1111/jir.12796>.
- Skerrett, A., 2014. Religious literacies in a secular literacy classroom. *Read. Res. Q.* 49 (2), 233–250. <https://doi.org/10.1002/rrq.65>.
- Skerrett, A., Omogun, L., 2020. When racial, transnational, and immigrant identities, literacies, and languages meet: black youth of Caribbean origin speak. *Teach. Coll. Rec.* 122 (13), 1–24. <https://doi.org/10.1177/016146812012201302>.
- Smith, P., 2020. Silencing invisibility: toward a framework for Black immigrant literacies. *Teach. Coll. Rec.* 122 (13), 1–42. <https://doi.org/10.1177/016146812012201301>.
- Smith, P., Lee, J., Chang, R., 2022. Characterizing competing tensions in Black immigrant literacies: beyond partial representations of success. *Read. Res. Q.* 57 (1), 59–90. <https://doi.org/10.1002/rrq.375>.
- Street, B.V., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Stroud, C., Wee, L., 2012. *Style, Identity and Literacy: English in Singapore*. Multilingual Matters, Bristol.
- Sugihartati, R., 2020. Youth fans of global popular culture: between prosumer and free digital labourer. *J. Consum. Cult.* 20 (3), 305–323. <https://doi.org/10.1177/1469540517736522>.
- Sulzby, E., Teale, W., 1991. Emergent literacy. In: Barr, R., et al. (Eds.), *Handbook of Reading Research*, vol. II. Longman, New York, NY, pp. 727–757.
- Tatum, A., 2008. Toward a more anatomically complete model of literacy instruction: a focus on African American male adolescents and texts. *Harv. Educ. Rev.* 78 (1), 155–180. <https://doi.org/10.17763/haer.78.1.6852t5065w30h449>.
- Tomei, R., Hollington, A., 2020. Transatlantic linguistic ties: the impact of Jamaican on African youth language practices. *Linguistics Vanguard* 6 (s4). <https://doi.org/10.1515/lingvan-2019-0048>.
- UNESCO, 2019. Five Decades of Literacy Work. Available at: <https://en.unesco.org/themes/literacy-all/five-decades>. (Accessed 10 April 2022).
- Unsworth, L., et al., 2022. *Multimodal Literacy in School Science: Transdisciplinary Perspectives on Theory, Research and Pedagogy*. Routledge, London.
- Ünver, H., 2018. *Global Networking, Communication and Culture: Conflict or Convergence? Spread of ICT, Internet Governance, Superorganism Humanity and Global Culture*. Springer International Publishing, Gewerbestrasse, CH.
- Vasquez, V.M., Janks, H., Comber, B., 2019. Critical literacy as a way of being and doing. *Lang. Arts* 96 (5), 300–311.
- Vasudevan, L., 2020. Three moments in teaching with/in arts practices. *Res. Teach. Engl.* 55 (2), 192–196.
- Vasudevan, L., Kerr, K.R., Hibbert, M., Fernandez, E., Park, A., 2014. Cosmopolitan literacies of belonging in an after-school program with court-involved youths. *J. Adolesc. Adult Lit.* 57 (7), 538–548. <https://doi.org/10.1002/jaal.288>.
- Vetter, A., et al., 2021. Asking Teens about Their Writing Lives: The Writing Identity Work of Youth. *Literacy Research and Instruction*, pp. 1–24. <https://doi.org/10.1080/19388071.2021.2008559>.
- Wargo, J.M., 2018. Writing with wearables? Young children's intra-active authoring and the sounds of emplaced invention. *J. Lit. Res.* 50 (4), 502–523. <https://doi.org/10.1177/1086296X18802880>.
- Watson, V.W., Beymer, A., 2019. Praisesongs of place: youth envisioning space and place in a literacy and songwriting initiative. *Res. Teach. Engl.* 53 (4), 297–319.
- Wertsch, J.V., 1991. A sociocultural approach to socially shared cognition. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC, pp. 85–100. <https://doi.org/10.1037/10096-004>.
- Weyand, L., Juzwik, M.M., 2020. Schooling activist evangelical literacy: speaking, writing, and storying Christian faith in dialogue with public secondary literacy curriculum. *Ling. Educ.* 55, 100789. <https://doi.org/10.1016/j.linged.2019.100789>.
- Willis, A.I., 2012. *Reading Comprehension Research and Testing in the US: Undercurrents of Race, Class, and Power in the Struggle for Meaning*. Routledge, New York, NY.
- Ybarra, M.G., 2020. "We have a strong way of thinking... and it shows through our words": exploring mujerista literacies with Chicana/Latina youth in a community ethnic studies course. *Res. Teach. Engl.* 54 (3), 231–253.
- Yosso, T.J., 2005. Whose culture has capital? a critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91. <https://doi.org/10.1080/1361332052000341006>.

The construction of gender in literacy

Barbara J. Guzzetti^a and Katina Zammit^b, ^a Arizona State University, New College of Interdisciplinary Arts & Sciences, Humanities, Arts & Cultural Studies, Department of English, Phoenix, AZ, United States; and ^b Western Sydney University, School of Education, Penrith, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	250
Perspectives on gender and sex	250
Troubling social models of gender and sex	251
Gender, schooling and literacy	252
Assessments of male and females' engagement in literacy	253
Critiques of gender differences on literacy assessments and their interpretations	254
Literacy as a cognitive process or performance and the role of gender	255
Gender differences on US National Assessments	255
Gender differences in international assessments	255
Literacy as social practice and the role of gender	256
Gender and socio-cultural literacies	256
Social and cultural constructions of gender	257
Gender and literacy practices: those who identify as female	257
Gender and literacy practices: those who identify as male	258
Literacy practices: beyond the binary	258
Recommendations for future research on gender and literacy	259
References	259

Introduction

What is meant by “gender” in the extant research on literacy? How has gender been examined in literacy research? To answer these questions requires an examination of the construct of gender and an integrative overview of literacy research. In doing so, we highlight the ways in which gender has been treated in research that examines literacy as a cognitive achievement or performance (e.g., Reilly, 2012), as well as a social or sociocultural practice (e.g., Gee, 1992; Street, 1996). We present studies that describe the literacy practices of males and females, particularly those studies in which other social factors or subjectivities that intersect with gender were considered, those in which gender was considered beyond gender binaries, and those that were analyzed from a gender framework.

To begin, we present an overview of research on gender and sexuality, problematizing the social models and consider how the discourses of education can reproduce stereotypical, heteronormative binary gendered identities. We present findings from international reports and views of scholars from various nations, defining how gender has been viewed in the extant research and deconstructing those views. In the next section, we examine how educational discourse and practices influence the engagement of males and females in literacy and critique the gender differences in literacy assessments and their interpretation. We then view literacy from a cognitive perspective as a product or process of cognition by discussing gender as an influence on literacy achievement and by presenting results of national and international survey research reports of literacy achievement and engagement analyzed and compared by gender, as well as criticisms of those results and their interpretations in regard to gender implications. In the final section of the report, we examine literacy as a social practice utilizing research conducted from the perspective of literacy as a social or sociocultural practice (Gee, 1992; Street, 1996). We consider gender and literacy practices of those who identify as female, male and beyond the binary. We conclude by identifying gaps in the literature on gender and literacy and suggest directions for future inquiries.

Perspectives on gender and sex

Conceptualizations of gender and sex within societies are generally associated with the biological or physical state of an individual. With rare exceptions, an individual is identified at birth as either biologically male or female despite the range of bodily morphologies and chromosomal configurations (Paechter, 2007). Physical characteristics classify an individual as either male or female. From this perspective, gender is synonymous with biological sex which typically alludes to gender binaries as an individual being male or female, which in turn leads to both gender and sex being misconstrued as fixed. Sexuality, as with gender, can be fluid but can include “many different elements besides gender (fantasies, emotions, power, biology, ability, props, preferred partners, acts, milieus and so on)” (Persson et al., 2020, p. 57 citing Sedgwick, 1993).

Gender stereotypes from a binary view of gender attribute certain characteristics as “natural” or “innate” to men and to women. For instance, men are thought to be aggressive, forceful, independent, rational, strong, active and decisive. In contrast, women are considered to be kind, helpful, beautiful, emotional, weak, passive and concerned about others (Francis, 2010; Francis and Paechter, 2015; Kollmayer et al., 2018; Meier and Diefenbach, 2020). These stereotypical beliefs aligning gender and sex to one’s personal capabilities, thoughts, behaviors, roles and interests can affect an individual’s education, cognition, career or life choices (Kollmayer et al., 2018).

From a social learning theory perspective, gender is a “relatively stable binary construction reproduced via societal institutions” (Francis, 2010, p. 477). Individuals learn how to “be” a boy or girl and a man or woman by interacting with their communities of practice (Paechter, 2007; Wenger, 1999), including their family members, teachers and peers, and fellow employees or employers at work. They learn to perform in heteronormative gender-appropriate ways through these dominating discourses. Francis (2010) noted that gender is “a construction that is socially dominant; and which authors itself as a unitary, total, inevitable, and natural, system ... and is commonly understood and agreed as a ‘true’, uncontested account” (p. 479). Through the doing of binary gender, differences between girls and boys and women and men are created which are not natural, essential or biological, but are entrenched in social norms, practices and rules (Callahan and Nicholas, 2019; Subrahmanian, 2005). Gender is thus misconstrued as fixed and synonymous with binary conceptualizations. It should also be noted that sex, like gender, is not exclusively male/female but can also be fluid.

With the act of categorization comes the problem of reproducing gender binaries. Labels are associated with different performances of gender with the heteronormative as the basis for categorization which can override individual identities (Callahan and Nicholas, 2019). The labels individuals employ may constrain an individual’s identity construction but can also de-emphasize difference in different generations (Persson et al., 2020). The proliferation of labels has occurred alongside “newer discourses [which] have emerged that emphasize ‘sameness’, seeking to represent gays, lesbians, bisexuals, *trans** and other queer people as good citizens deserving of equality and access to the same rights and social institutions as heterosexuals—still different, but not that different” (Persson et al., 2020, p. 57). Persson et al. (2020) note that those born in the 1990s seem to negotiate a “spectrum of possibilities” which were “less about difference from a norm (in a Derridean sense) than about a deeper sense of specificity” (p. 66) and that could also “be read as being less about departures from hegemonic heterosexuality, than about the production of abundant identities that set themselves apart from older mononormative templates” (p. 67).

Families and teachers may reinforce traditional and stereotypical views of gender, resulting in girls and boys either conforming or nonconforming and being confrontational. Educational environments from pre-school to the university are places where “binary discourses produce gendered bodies” (Callahan and Nicholas, 2019, p. 707). From a young age these binaries are reinforced through media, advertising and in interactions between adults and children: the language they use, the behaviors and actions they encourage and discourage, the clothes they choose for children to wear and censure for wearing, the toys they purchase, to the color of the clothing and their rooms (Paechter, 2007). This surveillance and disciplining of children’s gendered bodies continues to be reinforced with the practices of school as “children’s bodies are regulated and controlled, often in ways that are gendered, so that they may have to wear particular forms of clothing, sit in specific ways, at particular times and in designated places and, at times, to act more as a collective than as an individual body” (Paechter, 2007, p. 61). However, it is not just the adults that police gender “appropriate” behavior, peers also pressure others to conform. As Paechter (2007, p. 64) comments:

Children come to nursery or school with their existing understandings of masculinity and femininity, developed in the communities of practice of which they have previously been members. They are also influenced by the values and perceptions of the adults in the setting, staff and peer conceptions of what is ‘natural’ behaviour, images from the media, and as they grow older, the structures of school. They use these, in conjunction with the perceptions of other newcomers and the constructions dominant within the peer group, to construct group understandings of the rules of being a ‘proper’ boy or girl in this context.

Troubling social models of gender and sex

In the 1980s and 1990s, feminist and post-structural theorists challenged binary constructions of gender and the categorization of certain behaviors as “masculine” or “feminine” as precipitated around social inequalities (Francis and Paechter, 2015). Gender came to be viewed as socially constructed by acknowledging that gender “operates across individual, interactional and institutional levels of social life” (Callahan and Nicholas, 2019, p. 706). Gender is “not a property of bodies or something originally existent in human beings ... it is a product and process of various social technologies, institutional discourses, epistemologies, and critical practices, as well as practices of daily life” (de Laurentis, 1987, p. 3 cited in Sanford, 2005, p. 304). Butler (1990), as cited in Callahan and Nicholas (2019) noted that “gender functions as a discourse that produces what it names” (p. 706).

This challenge to the accepted “naturalistic” view of gender facilitated a fuller understanding of and attention to gender complexity and fluidity, but still focused on the performance of gender by gender-appropriate bodies (Francis, 2010; Francis and Paechter, 2015). Connell (1995) drew attention to a variety of behaviors associated with performed masculinity with hegemonic masculinity or the legitimization of men’s dominant position in society and subordination of women. Researchers (Francis, 2010; Francis and Paechter, 2015; Paechter, 2006, 2007) began to use the plural—“masculinities” and “femininities”—to acknowledge that there is not a single form of masculinity or femininity, but rather competing, contradictory, and overlapping forms (Kehler, 2010). Paechter (2007) used the terms “masculinities” and “femininities” to mean the varying ways in which individuals “do”

boy or girl or man or woman and how individuals through their behaviors and attitudes, actions, thoughts, and dispositions demonstrate to themselves and to others how they are male or female.

However, these labels of masculine or feminine are problematic as they imply a binary exists and that there are actions, abilities, and behaviors that can be characterized as masculine or as feminine. This practice can continue to essentialize or stereotype individuals and social constructions of gender. Any binary sets up one category as superior with the other as subordinate; as a result, they are not equal (Callahan and Nicholas, 2019). Francis and Paechter (2015) commented, “female” non-gender-traditional expression is not read as “doing masculinity” or male non-traditional constructions as “doing femininity”; rather, researchers tend to identify and analyze gender-traditional behavior (p. 778).

Gender attribution acknowledges that the performance of gender is enacted through shared knowledge of social norms and signs, such as clothing, language and other social cues (Callahan and Nicholas, 2019; Francis and Paechter, 2015). These performances allow individuals to claim attributions of gender which “will be enacted in a differential way between, for example, transmen and transwomen, with the former more likely to have their gender claims legitimated and accepted than the latter” (Francis and Paechter, 2015, p. 781). Children’s experimentations with gender have become more acceptable although this experimentation is more socially acceptable with girls, such as enacting ‘tomboy’ identities, than it is acceptable for boys (Francis and Paechter, 2015). Adolescents also experiment with gender labels beyond the binary as they attempt to find one that captures their gender or sexual subjectivity (Persson et al., 2020). Experimenting with gender could be viewed as “some boys do girl as their way of being a boy (inhabiting or living that male body) and vice versa” which may be considered transgressive because the performance does not match the expectations of the body (Paechter, 2012, p. 239).

Sex role theories view masculinity and femininity as being located solely within male and female bodies whereas gender relational theorists see gender as more “fluid” (Connell, 1995; Skeggs, 1997; Skelton and Francis, 2003). Rather than seeing boys and girls in stereotypical ways (girls do “x”, boys do “y”) which is characteristic of sex role theories, gender relational theorists argue that gender is essentially constructed in relation to and in opposition to the other (that is, masculinity is what femininity is not) and is influenced by context (Connell, 1995; Skelton and Francis, 2003). Skelton and Francis (2003) argue that more recent thinking demands a consideration of the multidimensionality of identity whereby masculinities and femininities are viewed as shaped by social class, sexuality, religion, age, ethnicity and other social factors. Simply focusing on gender denies the influence of other subjectivities or social factors on behaviors, representations and identity formation, including a person’s socio-economic background, culture, race, ethnicity, sexuality, religion, family, and schooling (Martino and Kehler, 2007).

Gender, schooling and literacy

Gender as a social construct may affect what is and can be learned in and out of school and what is considered to be out-of-bounds (Sanford, 2005; Subrahmanian, 2005). The expectations of school personnel often reinforce stereotypical views about the interests and behaviors of boys and girls and how they learn which may limit students’ engagement and career choices in a discipline or subject (Retelsdorf et al., 2015; Sanford, 2005). These expectations may also influence students’ self-concepts by presuming differences in abilities and behaviors according to gender which in turn may impact students’ academic performance. Teachers can make a difference in students’ achievement (Hattie, 2003), but teachers’ judgements of students’ academic performance can also be influenced by their students’ (mis)behavior, particularly when teachers award higher grades to girls than to boys for the same performance (Buchman et al., 2007; Meier and Diefenbach, 2020). An individual’s self-confidence is based on experience; therefore, lower achievement in reading and writing may result in less motivation for engaging in literacy practices (Kollmayer et al., 2018; Meier and Diefenbach, 2020). As Retelsdorf et al. (2015) noted, “teachers acting upon gender stereotypes could—consciously or unconsciously—shape social interactions in class by, for example, creating a warm and challenging atmosphere for students from positively stereotyped groups and a cold and less challenging environment for challenging students from negatively stereotyped groups” (p. 188).

Being a good student or “ideal student” has been aligned by teachers and schools to fit stereotypical feminine characteristics, including listening and attending to the teacher and to others, working hard, being diligent, being conscientious, working collaboratively, and being articulate (Martino and Kehler, 2007). Enacting the ideal feminine student may not bring benefits for young women in post-school life, however, as being a submissive girl in school settings may disadvantage females in the workplace where assertion, leadership, and initiative are valued (Beaman et al., 2006). Being a compliant student can also be problematic for boys as being compliant is a subordinate position which is typically understood as a feminine posture (Paechter, 2012). These views construct a hegemonic gender binary, essentializing or stereotyping gender, and limit individuals’ conceptions of masculinity and femininity, of doing and being a boy or a girl or a male or a female (Paechter, 2012; Skelton and Francis, 2011). When scrutinized, the “ideal (neo-liberal) pupil” is actually associated with binary and stereotypical male and female characteristics (Skelton, 2010; Skelton and Francis, 2011).

Performing successful masculinity or being the “proper boy” often entails devaluing academic achievement, particularly literacy achievement. Academically successful boys can rework this image while still maintaining their hetero-sexualized gender position by other accomplishments and behaviors, such as being good at sports and playing computer games with friends (Skelton and Francis, 2011). The “all-rounded” boy incorporates and negotiates feminine traits in producing a “powerful mode of masculinity” (Skelton and Francis, 2011, p. 471). These “feminine” traits are accepted because they carry social capital within schools and in workplaces.

Boys typically need to navigate what is “allowed” and what counts as masculine while girls are more enabled to explore behaviors aligned with masculinities (Paechter, 2012).

It has been tacitly accepted in education and in the media that “naturally” girls would do better in reading and language-related subjects and boys would perform better in subjects like mathematics and physics (Barrs, 2000). In the latter decades of the 20th century, the reasons for girls’ lack of confidence were identified as aspects of their schooling that formed the “hidden curriculum” (Spender, 1982), including representations of males and females in books and textbooks (or their lack of representation) and teachers’ differential attitudes and expectations in their interactions with boys and girls (Callahan and Nicholas, 2019; Skelton, 2010). For example, researchers, teachers, and students have noted the lack of representation of women and girls in science textbooks, particularly in illustrations that showed only men in science careers (Guzzetti and Williams, 1996a,b), sending subtle messages to students that science is the work of men.

In addition to their differential expectations, teachers’ educational practices, co-educational classes, activities, and texts may not meet diverse students’ needs. Due to the large number of white and middle-class female teachers who may themselves identify as *trans*, queer etc. in American schools, curriculum and instruction may be considered as too feminine to gain boys’ interest and motivate them to succeed (Bean and Ransaw, 2013; Martino and Kehler, 2007; Skelton, 2010). “Boy-friendly” approaches based on gender stereotypes of their interests and learning styles, such as providing boys with books about sports and kinaesthetic learning approaches, may let boys down as these interventions frequently do not challenge boys intellectually (Skelton and Francis, 2011). Kosciw et al. (2020) noted that students who identify as lesbian, gay, bisexual, transgender, or queer (LGBTQ+) typically experience a negative environment in school and are at greater risk of underperforming than their heterosexual peers. Students of color and LGBTQ+ students typically do not see themselves represented in the curriculum and consequently feel alienated from literacy engagement and performance.

Teachers’ stereotypical or gendered expectations for literacy performance for boys and for girls may result in differential literacy achievement and engagement. Researchers have suggested that teachers can reinforce gender stereotypes by having lower expectations for students’ literacy performance and mirror the stereotypical and limiting views and values presented regularly in the public media as accepted and unchallenged (Sanford, 2005, p. 307). The rhetoric of boys as disadvantaged can feed into “teachers’ taken-for-granted assumption that boys require more time in class, require more discipline, are behavior problems, are poor at expressing their emotions, and are not good at reading” (Lingard, 2003, p. 43). Sanford (2005) noted that teachers’ expectations reinforced stereotypical views of boys’ and girls’ literacies, such as who read more or who read less, and who were risk takers or not. These preconceptions can also influence the direction of teachers’ encouragement to students and the types of assignments they accept which in turn can differentially influence their students’ literacy engagement, attitudes, and performance.

It is therefore important to examine how literacy education can serve all students across all areas of the curriculum and contribute to all students’ success (Lingard, 2003; Meier and Diefenbach, 2020) and to re-examine notions of who is successful and who is not based on gender stereotypes and under-examined achievement test scores. For example, scholars have noted that girls cannot be considered to be “success stories” when their lack of self-confidence is considered rather than simply their academic achievement (Skelton, 2010). Girls are not necessarily educationally conforming or advantaged; they may not conform to acceptable classroom behaviors they are often noted for, such as being compliant and quiet when they are being assertive and active (Callahan and Nicholas, 2019; Francis, 2010). As Skelton and Francis (2011) found, despite perceptions to the contrary, there are successfully literate boys who do engage in classroom literacy activities, including discussions, reading, writing of emotionality, and empathizing with literary characters. Boys can be as prepared as girls to develop elaborate contexts for their stories. Therefore, education policies and practices should not be driven by “one size fits all” or offer easy solutions that reinforce dominant masculinity or femininity (Martino and Kehler, 2007).

Home and family conditions may also influence gender gaps in literacy engagement and achievement. These factors include differential time that parents can provide to support and reinforce their children’s learning if they are middle class or above versus if they are working class or poverty stricken (National Education Association, 2012–2015). Students of color are more likely to live in poverty than their white counterparts. For example, research from the National Center for Children in Poverty (<http://www.nccp.org>) has indicated that 62% of Latino children live in low-income families as compared to 31% of White children (Jiang et al., 2015). Latino children are less likely than other children to live in homes with a head of household with a high school diploma (Kids Count, 2014) and may lack role models of adults engaging in literacy practices and events that are aligned with school-based literacy practices (see Williams, 1998; Brice-Heath, 1983 for discussion of cultural and socio-economic family vs. school literacy practices). Researchers (e.g., Brooks, 2002; Ransaw, 2019) have identified the importance of parents who read to their children and engage with them in literacy activities as a positive influence on their children’s school literacy performance and engagement.

Assessments of male and females’ engagement in literacy

Literacy engagement is associated with and influences literacy achievement (Brozo et al., 2014). From a gender binary perspective, gender disparities have been noted in literacy engagement or involvement in literacy activities and processes that can develop knowledge, skills, and attitudes toward reading, writing and understanding and composing texts. For example, researchers analyzing the results of PISA assessments in five countries noted that not only do boys achieve less in reading than do girls, they also engage less in reading than do girls; globally and generally, girls are more engaged in reading and are better readers than boys (Brozo et al., 2014).

This gender gap in literacy engagement continues into adulthood; survey research by the Pew Research Center found that women in the United States reported reading more books than did men while gender gaps still persisted, but were less pronounced for other types of reading materials (Perrin, 2019). These findings with print reading did not carry over to reading online, however, at least not for adolescents—analysis of gender differences in online reading engagement on the 2009 ePISA revealed there were no gender differences in most countries in how engaged 15-year-old girls and boys were in online reading (Wu, 2014). While Wu (2014) found there were no gender differences in social or information-reading engagement overall, he suggested that the smaller gap in online reading might be related to boys' greater confidence and girls' higher anxiety in the electronic reading environment (p. 268). Wu also commented on the inconsistency between his finding that girls' navigation skills, measured in terms of relevant pages visited, were better than boys' in comparison to research in the online search strategies that boys have better self-reported search (or navigation) skills; given this inconsistency, he recommends that future studies examine whether such effects exist (p. 268).

Critiques of gender differences on literacy assessments and their interpretations

Numerous researchers have questioned generalizations based on the results of literacy assessments (e.g., Kehler, 2016; Meier and Diefenbach, 2020; Reilly, 2012). Meier and Diefenbach (2020) noted that the PISA's Gender Report stereotyped girls as "passively driven by parental, educational and/or general social forces beyond their control" (p. 638). Girls' mathematical issues were considered as caused by external factors, but boys' literacy issues were considered as caused by innate or biological differences and behavioral variables. These gendered interpretations have conformed to middle-class stereotypes of girls and were not deconstructed to consider the cultural and social factors that influence achievement (Reilly, 2012).

It is also important to interrogate not just the perception of a gap in literacy achievement based on male-female binaries, but to consider which boys and which girls are disadvantaged; small performance gains for mostly middle-class girls should not be interpreted as "large gains for all girls outperforming all boys" in literacy achievement (Lingard, 2003, p. 38). Lingard (2003) cautioned the interpretation of performance data by gender that essentializes the category of "boys" in informing educational practices and policies. Such discussions ignore the influences of social factors or subjectivities, such as race, ethnicity, age, culture, sexual orientation, and social class, that intersect with gender and can influence literacy achievement, performance, and engagement. The identities of boys—like girls—are not a single, coherent and undifferentiated group.

Rather, a more nuanced interpretation of test scores should consider social factors or subjectivities, such as social class, race, and geography, i.e., rural versus urban location and other social factors (Froese-Germain, 2006; Reilly, 2012) as doing so troubles sweeping generalizations. Re-examining boys' test scores has demonstrated that not all boys are failing, and education is not failing all boys. For example, although middle-class boys have been challenged by middle-class girls in "high status subjects", they still perform relatively well; middle-class boys have demonstrated higher achievement than working-class girls in literacy (Martino and Kehler, 2007) and they have outperformed boys in poverty on writing assessments (Disenhaus, 2015).

It is also important to consider assessment results for each gender in terms of race or ethnicity, as well as financial and social capital, when considering literacy achievement (Buchman et al., 2007). For example, reading achievement was lower for Hispanic and Black students than for white students on the 2017 and 2019 National Assessments of Educational Progress (NAEP Report Card, 2019). This result confirms the observation that African American boys are more likely to attain lower academic achievement than their white peers (Neal, 2003). Compared to 1998 and to 1992, however, average scores were higher in 2019 for all racial/ethnic groups (White, Black, Hispanic, and Asian/Pacific Islander students). In 2019, average reading scores were lower compared to the previous assessment in 2017 for almost all student groups at grade 8, finding that average scores were lower on the 2019 assessment than on the 2017 assessment for White, Black, Hispanic, American Indian/Alaska Native, and students of Two or More Races. Research sponsored by The Schott Foundation for Public Education (2011) predicted a 65% high school graduation rate for Latino males compared to 80% for White males for 2012–2013. This study also showed that 81% of Hispanic 4th graders were not proficient in reading as compared to the national average of 66%.

Hence, researchers have cautioned against forming educational policies and practices based on generalizations about gender differences in literacy achievement. Scholars have critiqued how the growing panic about boys is framed by educational systems in countries like the United States, the United Kingdom, Australia, and Canada based on alarmist reports in the media portraying boys as victims of radical feminists who support girls at the expense of boys' achievement (Kehler, 2016). Christina Hoff Sommers' (2000/2013) book, *The war against boys: How misguided policies are harming our young men*, was a reaction to a focus on increasing educational outcomes for girls and women given the limitations of the education system and society to adequately meet females' needs. The premise of the "war against boys" was that an emphasis on girls' education on improving the breadth of study of girls in mathematics, science, and engineering-related subjects and degrees and increasing the job opportunities that are higher paid had had the desired impact, but that there remained groups of boys who were disadvantaged because they were underachieving or also not reaching their potential. The "war against boys" was fueled by the media and by public discourse which viewed boys as the "new" disadvantaged given the "perceived" improvements for girls and women in education and work which is only a perception and not the case for all girls or women (Martino and Kehler, 2007; Skelton, 2010).

It is within this discourse that education and schooling has been framed as failing boys. This position has been critiqued as an over-simplified perspective on a far more complex situation, reflecting essentialized notions of boys and boys' learning styles and reinforcing entrenched heteronormative behaviors and roles based on dominant male stereotypes (Kehler, 2010; Lingard, 2003; Lingard et al., 2009). Foregrounding this dominant construction of boys and masculinity can also be considered as discriminatory as it "can bring enormous pressure to bear on boys significantly affecting their performance in school"

(Subrahmanian, 2005, p. 406). The rhetoric of boys as the new disadvantaged feeds into “teachers’ taken-for-granted assumption that boys require more time in class, require more discipline, are behavior problems, are poor at expressing their emotions, and are not good at reading” (Lingard, 2003, p. 43).

Scholars have cautioned educators to be mindful of “campaigns” that consider one gender over another at face value. In considering boys’ and girls’ achievement in literacy, Skelton and Francis (2011) recommended that educators need to “interrogate the media’s and their government’s emphasis on ‘boys’ as a homogenous group and to question the language suggesting a ‘crisis’ in boys’ education” (p. 473). In addition, as teachers are encouraged to focus on “poor” boys, it is important that girls’ disadvantages do not become invisible (Sanford, 2005). Lingard (2003) and Lingard et al. (2009) caution that it is a question of “which girls and which boys” are advantaged or disadvantaged that requires educational reform rather than engaging in a victim/victor dichotomy.

Literacy as a cognitive process or performance and the role of gender

Literacy as a cognitive performance focuses on a set of cognitive skills that people either possess or do not possess. These discrete skills can be taught in an appropriate sequence and are needed for learning but which for many are considered the basis of issues in literacy learning and performance. From this perspective, reading is closely associated with efficient use of perceptual skills, automatic word recognition, decoding skills and comprehension (Anstey and Bull, 2004; Farrell, 2012). As an individual acquires the necessary skills they then can apply them at the appropriate time. It takes place in the head as the individual constructs and represents meaning, usually in print-based texts (Luke et al., 2010; Street, 1984). Emphasis is on “working out the answer, engaging in the thinking process and knowing how to use comprehension skills” (Anstey and Bull, 2004, p. 127). It doesn’t take into consideration the socio-cultural background of students, the funds of knowledge they bring to school nor how these influence individual students’ understanding or achievement (Street, 1984).

Large-scale achievement assessments are based on a view of literacy as a cognitive performance employing survey research where there is only one correct answer. These assessments are intended to allow for comparisons in achievement across time with results typically analyzed for subgroups by age, gender, race, and other social factors. As such these assessments are conducted from a binary view of gender and continue to reinforce gendered assumptions and expectations of literacy achievement.

Gender differences on US National Assessments

In the United States, achievement assessments have tended to indicate that boys have higher test scores in mathematics while girls have higher test scores in English, specifically reading (Buchman et al., 2007; Kollmayer et al., 2018). Gender differences favoring girls have been consistently demonstrated in students’ performance and in students’ motivation (Kollmayer et al., 2018). Researchers have discovered that in the early years of schooling, girls and boys have similar performances in reading, but gender differences in scores increase as students move through school and persist on college entry tests, such as the SAT (Buchman et al., 2007). Literacy researchers have also noted a gender gap in writing (Newkirk, 2000) as well as in reading achievement (Watson et al., 2010) discovering that, from a binary perspective, boys consistently lag behind girls in literacy performance.

The most widely recognized literacy assessment that examines gender differences in literacy achievement conducted in the United States is the National Assessment of Educational Progress, (NAEP), a project of the National Center for Education Statistics (NCES), a section of the United States Department of Education. The NAEP is a national achievement survey used to track student achievement over time in 4th, 8th, and 12th grades at the state and national levels by measuring students’ achievement in reading, writing, mathematics, science, and other subjects. Results are reported after each assessment as The Nation’s Report Card.

Analyses of these assessments have consistently demonstrated that girls outperform boys in literacy. Reilly et al. (2018) examination of data from three decades of NAEP assessments conducted from 1988 to 2015 in reading achievement and from 1988 to 2011 in writing performance, demonstrated that girls showed significantly higher reading and writing scores than boys in every wave of the assessment at every grade. The gap between girls’ and boys’ achievement in reading was greatest at grade 12 with the magnitude of these gender performance differences demonstrated to be higher in writing than in reading. The authors listed five possible explanations of these gender differences in literacy achievement, including differential maturation rates; differences in lateralization of brain function; greater male variability effect; gender differences in language competencies and behaviors; and gender stereotyping of reading and writing as feminine pursuits. Reilly et al.’s (2018) analyses and five possible explanations highlight the limitations of large-scale assessments that view literacy as a cognitive process, reinforcing binary gendered perspectives of literacy achievement.

Gender differences in international assessments

These gender gaps in literacy achievement are also evidenced in educational assessments and outcomes internationally, which similarly reinforce binary discourses associated with achievement. On a global scale, The United Nations Educational Scientific and Cultural Organization (UNESCO), a program of the United Nations dedicated to providing global leadership in education; strengthening educational systems worldwide; and responding to global challenges through education and gender equality (<https://en.unesco.org/themes/education>) reported that boys are more likely than girls to underachieve and drop out of school (UNESCO, 2012). Boys of color and those from poor or working-class families were found to be more likely than their White or middle-class peers to be academic underachievers compared to their female peers (UNESCO, 2014).

UNESCO defines literacy as “the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts” (Montoya, 2018) and characterizes literacy as “beyond a set of reading, writing and counting skills [and as] a means of identification, understanding, interpretation, creation, and communication in an increasingly digital, text-mediated, information-rich and fast-changing world” (<https://en.unesco.org/themes/literacy>). From that definition, UNESCO has conducted a Literacy Assessment and Monitoring Program (LAMP) every 5–10 years in six countries (El Salvador, Kenya, Mongolia, Morocco, Niger and Palestinian Autonomous Countries), discovering there were 773.5 million adults identified globally who were categorized as functionally illiterate in 2013. In contrast to other UNESCO findings, LAMP found of that number, 63.8% were women. 123.2 million children were considered illiterate and 61.3% of them were girls with the highest rates of illiteracy evidenced in South and West Asia and Sub-Saharan Africa (UNESCO, 2013). While there may be international global comparisons for differences in literacy achievement for boys and girls, in some countries females do not have access to education at all.

The International Association for the Evaluation of Educational Achievement also administers a global achievement survey analyzed by gender, the Progress in International Reading Literacy Study (PIRLS), which commenced in 2001 and has been administered every 5 years as a comparative international measure of 4th-grade students’ reading comprehension. This assessment is undertaken so countries can understand how successful they are in preparing students to read, understand, and interpret information (Warner-Griffin et al., 2017). The 2016 achievement survey was conducted in 58 countries and included an additional ePIRLS as a measure of online reading comprehension. Gender gaps in achievement have been revealed by these assessments with girls consistently outperforming boys in reading comprehension across nations (except for Macao, SAR, and Portugal where achievement in 2016 was equal) in 2011 and 2016 (National Center for Education Statistics, 2011; Warner-Griffin et al., 2017).

The Organization for Economic Co-operation and Development (OECD) conducts an additional international literacy achievement assessment analyzed by gender, the Program for International Student Assessment (PISA). This assessment is conducted to assess the academic performance of 15 year-old students across nations, comparing students’ achievement in literacy, mathematics, and science (Office for Economic Co-operation and Development (OECD), 2014, 2019). Although a gender analysis was not reported for the 2018 assessment published in 2019, the 2015 PISA Gender Report (Office for Economic Co-operation and Development OECD, 2015) examined academic performance by gender by using data from the 2012 PISA. On this assessment, girls achieved lower scores in mathematics than did boys while boys achieved lower scores in literacy than did girls. The advantage of boys in mathematics was consistently three to four times lower than the advantage of girls in literacy; the mean score was 11 points in favor of boys in mathematics and 38 points in favor of girls in literacy. The 2009 PISA assessment demonstrated that girls outperformed boys in reading literacy consistently across all nations with the United States showing the largest gender gap in reading achievement (Reilly, 2012). Analysis of gender differences in online reading on the 2009 ePISA revealed that girls performed better in metacognitive strategies and online navigation than did boys (Wu, 2014). However, caution needs to be applied by educators and governments in employing these gender-based assessment reports for changes to educational policy and practices as they report from a binary perspective and limited, if any, social considerations.

Literacy as social practice and the role of gender

A plethora of research on gender and literacy has been conducted from a view of literacy as a social or sociocultural practice (Gee, 1992; Street, 2014). Research from this perspective examines literacy as more than just a set of cognitive skills that people either possess or do not possess, but as a social activity or what people do. This perspective allows for the study of how and why people engage in literacy practices or events. Barton (1994) explained, “Literacy is not simply knowing how to read and write a particular script but applying this knowledge for specific purposes in specific contexts of use” (p. 24). Literacy is a part of a wider set of communicative practices that embrace oral, written, and visual communication and communicating for a purpose. People engage in literacy practices within their social networks that serve as resources for exchanging knowledge and skills in which individuals are treated as equals and not as “literate” or “illiterate”. In this view, literacy is always contextualized and situated within specific sociocultural settings. Individuals’ literacy practices are embedded in their social goals and cultural practices (Barton et al., 1999). Meaning is made through situated practices and interpretations, designing texts by using resources from multiple modes and representational features in everyday life (Jewitt, 2014).

Gender and socio-cultural literacies

Literacy research on gender conducted from a social or sociocultural perspective tends to recognize that there is no single “literacy” but “literacies”. There is not one universal application of literacy; rather, there are different literacy practices for different groups (e.g., members of specific organizations or occupations), for different kinds of activities (e.g., engaging in education or practicing religion), and for different social and institutional contexts (e.g., informal meetings or academic occasions). These literacies are not confined to reading, writing, listening and speaking, but include a range of multimodal textual productions, including aural, visual, print, and digital communications. Multiple modalities suggest the interrelationship between various modes of meaning, including written and spoken words, pictures, graphics, and auditory and spatial modalities (Cope and Kalantzis, 2009). Practices of the “new literacies” are those literacy practices that are chronologically new and constitute socially recognized ways of generating,

communicating, and negotiating meaningful content (Lankshear and Knobel, 2011) and are characterized by an ethos of participation, collaboration, and distribution of content (Knobel and Lankshear, 2001). These practices may be in print or digital form.

Social and cultural constructions of gender

Research conducted from the perspective of literacy as a social or sociocultural practice was informed by social and cultural constructions of gender. Such theories recognize the nuanced and relational context of lived experience where individuals have agency to overcome hegemonic views of gender and challenge dominant notions of gender and literacy performance (Bean and Harper, 2007). Gender theorists, such as Judith Butler, (1990, 2004), noted that the performance of gender is fluid and may change within contexts and social settings or over time. Masculinity and more broadly, gender, are social constructions rather than biological designations, constituting identity performances amenable to critique and change (Bean and Harper, 2007). The terms “masculinities” and “femininities” recognize myriad ways of doing gender that extend beyond gender binaries.

Feminist and gender theorists (e.g., Stanley and Wise, 1993) recognized that other subjectivities or social factors, such as race, culture, social class, age, sexual orientation, and geographical location interact with gender and influence gender representations in literacy practices. Intersectionality or intersectional feminist theory, a perspective developed by Crenshaw (1989, 2016) and applied in literacy studies of girls and women (e.g., Hinchman and Alvermann, 2019), extended this theory, noting that unique points of oppression stem from the intersection of two or more of these subjectivities.

Gender and literacy practices: those who identify as female

Since the advancement of the view of literacy as a social practice in the mid-1990s numerous noteworthy studies, undertaken from a frame of binary gender at the time, identified and described the literacy practices of those who identify as girls and women in school, home, and out-of-school settings, describing the functions that literacy served in their lives as females; we highlight a few of them here that focused on their consumption and production of texts. One of the first of these that examined how girls consumed texts was Margaret Finders' (1996) investigation of preteen girls who subscribed to, read, and shared popular commercial magazines aimed at young women. Her study identified how teen girls use literacy for their own purposes; how literacy practices assist girls in monitoring and achieving social roles and statuses; and how social groups regulate literacy practices. Finders' research extended Meredith Cherland's earlier (1994) study of how preteen girls read fiction to construct their identities, one of the first studies in literacy to illustrate gender as both a cultural and a social construction. These studies were informed by ethnographic research from the prior decade conducted by Shirley Brice Heath (1983) that examined local communities' ways with language and literacy and described how girls' literacy practices were limited and shaped by local cultural norms.

The concomitant development of new technologies, the advent of Web 2.0 capabilities, and an emerging line of inquiry on the new literacies stimulated researchers' exploration of girls' textual production with various participatory media. Barbara Guzzetti (2006) described how a young woman was marginalized by males in participating in a male-dominated online discussion forum for punk rock fans. Olcott and Mahar (2003) explored early adolescent girls' construction of identity texts in online communities creating fan fiction, fan art, and web sites, and their participation in electronic mail discussions that focused on Japanese animation (anime), finding that a girl who had benefited from online mentoring herself was hesitant to share her expertise with HTML coding and become a mentor for other girls for fear of being labeled “geeky”. Elisabeth Hayes (2013), again from a binary perspective, described how the identities and literacies practices of video gaming continue to be gendered with more boys playing games and participating in the gaming community than girls and games having more genres and gendered storylines that appeal to boys. Guzzetti and Gamboa (2005) investigated adolescent girls' online journaling, discovering that writing online facilitated the girl's identity formation and independence, empowering them to express intense emotions and ideas in a safe space for an authentic audience they designated. Guzzetti and Gamboa (2004) also described how adolescent girls authored and edited zines, do-it-yourself, self-published alternatives to commercial magazines, compiling entries from their classmates to deconstruct stereotypical notions of gender and social class performance. Barbara Guzzetti, 2013 investigated women's zines as a global literacy practice, describing how marginalized women in prison authored zines as a strategy to resist oppression of difference and loss of voice stemming from incarceration and from society's views of women doing time.

Literacy researchers have also included girls and women of other marginalized populations in their studies of identity texts. A foregrounding study by Pat Rigg (1985) identified gendered cultural beliefs about women's roles that limited a low-income woman who was a monolingual Spanish-speaking Mexican migrant worker in gaining basic literacy skills of reading and writing in either English or Spanish, but eventually learning to write her name which she wrote on all her possessions, a simple text of identity created by a literacy practice that allowed her agency and voice. Decades later, Mellinee Lesley (2012) explored preadolescent Latina and African American girls from low-income backgrounds writing in an environmentally safe “third space” between home and school, discovering that girls need instruction in critical literacy skills to interrogate texts and resist dominant and discriminatory ideologies in writing and reading online and off line. Rubinstein Avila (2007) described how an immigrant girl from the Dominican Republic learned to do gender by reading popular magazines. Muhammad and McArthur (2015) investigated black adolescent girls who identified and wrote against stereotypical portrayals of black girls as sexualized and as angry, loud and violent, images that dominate popular culture media and echoed the need for critical media literacy. Jong Kyeong Jung (2019) investigated Korean American women's participation in an online discussion forum, writing online to actively respond to the #MeToo and #WithYou digital social movements, engaging in practices of the literacies to effect social change and refute gendered notions of Asian women

as passive and silent. Barbara Guzzetti (2022) described how a Navajo woman created semiautobiographical stories in personal zines and social media to refute notions of Native Americans as a dying people, change stereotypical notions of Navajo women as squaws or princesses, preserve the Diné language, and share cultural customs through stories with contemporary settings.

Gender and literacy practices: those who identify as male

Following the development of the theory of literacy as a social practice in the mid-1990s, numerous noteworthy studies, also undertaken from a frame of gender as binary, identified and described the literacy practices of boys and men in school, home, and out-of-school settings, describing the functions that literacy served in their lives as males; we highlight a few of them here that focused on their consumption and production of texts. William Brozo (2013) described how early adolescent boys bridged the gap between their in-school instruction and their out-of-school interests by bringing in music lyrics and writing their own song lyrics during word study. In another investigation exploring a boy's reading of and reaction to popular culture texts, superhero comics were compared to classical literature, Cassidy and Kehler (2019) found an adolescent boy noticed the gendered images of hypersexualized women in these texts. In a study of adult men's online literacies, Erik Jacobson (2019) discovered toxic masculinity in men's discussions in online sports-news comment boards that were dominated by men, including expressions of homophobic language, explicit misogyny, and cuckoldry as insult that men used to demean others, describing how some men pushed back against these protean discourses in their online writing.

These studies are a reminder of the importance of the role of masculinity in relation to literacy, particularly as reasons why and how boys and men engage or do not engage in literacy practices and how they marginalize others as a result of their views. As noted by Smith and Wilhelm (2002) in their comprehensive review of the research on boys and literacy, the Australian scholar, Wayne Martino (1995) argued that hegemonic versions of masculinity in which boys see literacy as feminized and define maleness as being "not female" are inconsistent with being literate. Researchers have also found that adolescent boys' views of masculinity as superordinate over females plays a role in marginalizing girls' contributions to class discussions when boys interrupt and ridicule girls in small-group and whole-class discussions, and dominate talk about science concepts and texts (Guzzetti and Williams, 1996a,b).

Other researchers exploring the literacy practices of boys and men examined the multiple subjectivities of gender, race, and culture in their literacy development and practice. David Kirkland (2013) described how Black men engaged in the production and construction of texts, such as song lyrics and poetry that refute stereotypical notions of Black men and the enactment of Black masculinity as hypermasculine and cold. Theodore Ransaw (2019) interviewed African American fathers, finding that they enjoyed reading to their children daily and engaged them in a variety of literacy practices, including reading newspapers and the Bible and using dictionaries and thesaurus, as well as Smartphones and apps for literacy development and support, refuting stereotypical notions of African American fathers as distant and uninvolved in their children's literacy development. Guzzetti and Foley (2017) detailed how a Chicano man authored and edited multiple issues of a zine, *Rad Dad*, that refuted stereotypical notions of men and the masculinities displayed by men of Latin descent as macho/machismo, highlighting their active parenting. Guzzetti and Lesley (2017) detailed how an African American man refuted stereotypes of African American men as aloof hypermasculine and by writing a zine that presented alternative constructions of masculinity and fatherhood. Barbara Guzzetti (2019) described how a Navajo man created fictional stories in zines that illustrated the view of two spirits (male and female) and four genders (feminine female; masculine female; masculine male and feminine male) that inhabit each person, a theory held by some Navajos.

Literacy practices: beyond the binary

A growing line of inquiry has been undertaken reflecting gender and gender performance in literacy development and practice beyond binary views of gender. Jennifer Ingrey (2013) investigated the online writing practices of youth participating in a teen forum on Laura's Playground, a "Transgender Umbrella Resource Community Online", (<http://www.outproudandhealthy.org/2013/lauras-playground>), describing a discussion thread billed as a short story authored by a transgender teen living as boy/man that traced his transition journey. Simon et al. (2019) described how adolescents constructed a music video that critically rewrote a *trans* young adult novel, resisting stereotypical portrayals of female characters in the story and challenging gender expectations.

Mollie Blackburn (2002) described the literacy performances (a poem, *Pride*, and a video, texts of resistance against heterosexism and homophobia) created by an African American teenager who self-identified as a lesbian in the context of a youth-run center for LGBTQ+ youth. Jon Wargo (2017) explored how LGBTQ youth navigated Tumblr, a social media platform known for fostering spaces for socially marginalized users, describing the processes, modes, and social tactics youth used when writing the #socialjusticewarrior self, such as reblogging others' posts to combat and critique homophobia; remixing to advance racial justice; and tagging posts to educate others. Wargo (2015) also explored how LGBTQ youth constructed digital photos of themselves (selfies) and their online writings that served as textual artifacts of identity on Tumblr that allowed them to centralize their sexuality and gender. Oakley (2016) similarly explored the role of Tumblr in disrupting hegemonic public discourses associated with nonbinary genders and sexual orientations, including "heterosexual" as the default sexuality. She found that users "experience freedom specifically related to gender and sexuality" (p. 7) and constructed their identities through their creative use of Tumblr affordances particularly through a combination of bios: About Me pages, blog posts and tags on the posts.

Recommendations for future research on gender and literacy

Our review of the extant research on gender and literacy has led us to recommend that more investigations be conducted that focus on the literacy practices of diverse populations, and from a non-binary position. More studies need to be conducted on how members of marginalized groups engage in literacy endeavors and for what purposes they do so. Few studies have paid careful attention to issues of gender, race, social class, and sexualities and how these intersectional identities shape and are shaped by literacy development and practices. Without such nuanced studies that attend to these subjectivities, there is a danger of essentializing.

Relatively few studies focusing on gender and literacy that also attended to race and culture have been conducted. More distinctions need to be made among members of various racial and cultural groups, as well. We are reminded, for example, that studies of a Navajo man's literacy practices do not represent all Native American men just as studies of Korean American women engaging in literacies do not represent all Asian American women. Therefore, there is a relative paucity of research on the literacy practices of men and women of various Asian descent or those from Indigenous or Native American groups. Additional inquiries are also needed to explore the ways in which members of the LGBTQ+ community represent and explore gender and sexualities through literacies. It is our hope that future researchers will fill these gaps in the extant research and undertake studies that promote understandings of various ways of doing gender, race, and social class to advance a line of inquiry for gender and social justice.

References

- Anstey, M., Bull, G., 2004. *The Literacy Labyrinth*. Pearson, Sydney, Australia.
- Barrs, M., 2000. Gendered literacy? *Lang. Arts* 77 (4), 287–293.
- Barton, D., 1994. *Literacy: An Introduction to the Ecology of Written Language*. Blackwells, Oxford, UK.
- Barton, D., Hamilton, M., Ivanic, R., 1999. *Situated Literacies: Reading and Writing in Context*. Routledge, London.
- Beaman, R., Wheldall, K., Kemp, C., 2006. Differential teacher attention to boys and girls in the classroom. *Educ. Rev.* 58 (3), 339–366. <https://doi.org/10.1080/00131910600748406>.
- Bean, T., Harper, H., 2007. Reading men differently: alternative portrayals of masculinity in contemporary young adult fiction. *Reading Psychol.* 28 (1), 11–30.
- Bean, T.W., Ransaw, T., 2013. Masculinity and portrayals of African American boys in young adult literature. In: Guzzetti, B.J., Bean, T.W. (Eds.), *Adolescent Literacies and the Gendered Self*. Routledge, New York, NY, pp. 22–29.
- Blackburn, M., 2002. Disrupting the heteronormative: exploring literacy performances and identity work with queer youth. *J. Adolesc. Adult Literacy* 46 (4), 312–324.
- Brice Heath, S., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press, Cambridge, MA.
- Brooks, G., 2002. What Works for Children With Literacy Difficulties? The Effectiveness of Intervention Schemes. DTES Research Report No. RR380. London, UK.
- Brozo, W., Sulkunen, S., Shiel, G., Garbe, C., Pandian, A., Valtin, R., 2014. Reading gender and engagement: lesson from five PISA countries. *J. Adolesc. Adult Literacy* 57 (6), 584–593.
- Brozo, W., 2013. Outside interests and literate practices as contexts for increasing engagement and critical reading for adolescent boys. In: Guzzetti, B.J., Bean, T.W. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY, pp. 3–12.
- Buchman, C., Diprete, T., McDaniel, A., 2007. Gender inequalities in education. *Annu. Rev. Sociol.* 34, 1–50.
- Butler, J., 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, New York, NY.
- Butler, J., 2004. *Undoing Gender*. Routledge, New York, NY.
- Callahan, S., Nicholas, L., 2019. Dragon wings and butterfly wings: implicit gender binarism in early childhood. *Gend. Educ.* 31 (6), 705–723. <https://doi.org/10.1080/09540253.2018.1552361>.
- Cassidy, J., Kehler, M., 2019. "Holy gendered resource, Batman!": examining the broader applications of comics and superhero fiction beyond their restrictive relationship with boys. In: Guzzetti, B.J., Bean, T., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, NY, pp. 79–92.
- Cherland, M., 1994. *Private Practices: Girls Reading Fiction and Constructing Identity*. Taylor & Francis, New York, NY.
- Connell, R., 1995. *Masculinities*. Allen & Unwin, Sydney.
- Cope, W., Kalantzis, M., 2009. Multiliteracies: new literacies, new learning. *Pedagogies* 4 (3), 164–195.
- Crenshaw, K.W., 1989. Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *Chicago Legal Forum* 140, 139–167.
- Crenshaw, K., October 2016. The Urgency of Intersectionality. TEDWomen 2016. Available. http://www.ted.com/talks/kimberle_crenshaw_the_urgency_of_intersectionality.
- Disenhaus, N., 2015. Boys, Writing and the Literacy Gender Gap: What We Know What We Think We Know. University of Vermont eGraduate College Dissertations and Theses, p. 30. Retrieved. <https://scholarworks.uvm.edu/graddis/330/>.
- Farrell, M.L., 2012. *Reading Assessment: Linking Language, Literacy, and Cognition*. John Wiley & Sons, Hoboken, NJ.
- Finders, M., 1996. Queens and teen zines: early adolescent females reading their way toward adulthood. *Anthropol. Educ. Q.* 27 (1), 71–89.
- Francis, B., 2010. Re/theorising gender: female masculinity and male femininity in the classroom? *Gend. Educ.* 22 (5), 477–490. <https://doi.org/10.1080/09540250903341146>.
- Francis, B., Paechter, C., 2015. The problem of gender categorization: addressing dilemmas past and present in gender and education research. *Gend. Educ.* 27 (7), 776–790. <https://doi.org/10.1080/09540253.2015.1092503>.
- Froese-Germain, B., 2006. Educating boys: tempering rhetoric with research. *McGill J. Educ. Revue Des Sciences De l'éducation De McGill* 41 (2), 145–154.
- Guzzetti, B.J., 2006. Cybergirls: negotiating social identities on cybersites. *E-Learn. Digital Media* 3 (2), 158–169.
- Gee, J.P., 1992. *The Social Mind: Language: Ideology and Social Practice*. Bergin and Garvey, New York, NY.
- Guzzetti, B.J., 2013. Girl's zines as a global literacy practice. In: Guzzetti, B.J., Bean, T.W. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY.
- Guzzetti, B.J., 2019. Diverse men making media: creating cultural (re)constructions of gender and race. In: Guzzetti, B., Bean, T., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, NY, pp. 193–207.
- Guzzetti, B.J., 2022. "Stories of a Healing Way": a Navajo woman's media production for cultural representation and identification. *Read. Res. Q.* 57 (1), 131–148. <https://doi.org/10.1002/rq.377>.
- Guzzetti, B.J., Gamboa, M., 2004. Zines for social justice: adolescent girls writing on their own. *Read. Res. Q.* 39 (4), 408–437.
- Guzzetti, B.J., Gamboa, M., 2005. Online journaling: the informal writings of two adolescent girls. *Res. Teach. Engl.* 40 (2), 168–206.

- Guzzetti, B.J., Foley, L., 2017. Rad Dad: a Chicano man reconstructing masculinity and fatherhood through zines. In: Ortlieb, E., Cheek, E. (Eds.), *Addressing Diversity in Literacy Instruction*. Emerald Press, Bingley, UK, pp. 89–114.
- Guzzetti, B.J., Lesley, M., 2017. "Surviving": an African American man reconstructing masculinity through literacy. In: Ransaw, T.S., Majors, M. (Eds.), *International Race and Education Series: Emerging Issues and Trends in Education*, vol. 3. MSU Press, East Lansing, MI, pp. 57–85.
- Guzzetti, B.J., Williams, W.O., 1996a. Changing the pattern of gendered discussion: lessons from science classrooms. *J. Adolesc. Adult Literacy* 40 (1), 2–11.
- Guzzetti, B.J., Williams, W.O., 1996b. Gender, text and discussion: examining intellectual safety in the science classroom. *J. Res. Sci. Teach.* 33 (1), 5–20.
- Hattie, J., 2003. Teachers Make a Difference: What is the Research Evidence? Paper presented at The Building Teacher Quality: What Does the Research Tell Us? Melbourne, Australia. http://research.acer.edu.au/research_conference_2003/4/.
- Hayes, E., 2013. A new look at girls, gaming and literacies. In: Guzzetti, B.J., Bean, T. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY, pp. 101–108.
- Hinchman, K., Alvermann, D.E., 2019. How gender and intersectionality inform adolescent literacy. In: Guzzetti, B.J., Bean, T.W., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, NY, pp. 105–116.
- Hoff Sommers, C., 2000/2013. *The War Against Boys: How Misguided Policies Are Harming Our Young Men*. Simon & Schuster, New York, NY.
- Ingrej, J., 2013. Gendered subjectivities in online spaces: the significance of genderqueer youth writing practices in a global time. In: Guzzetti, B.J., Bean, T.W. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY, pp. 140–149.
- Jacobson, E., 2019. Performing and resisting toxic masculinities on sports news comment boards. In: Guzzetti, B.J., Bean, T., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, NY, pp. 169–181.
- Jewitt, C. (Ed.), 2014. *The Routledge Handbook of Multimodal Analysis*. Routledge, London.
- Jiang, Y., Ekono, M., Skinner, C., 2015. Basic Facts About Low Income Children. National Center for Children in Poverty, Columbia University, New York, NY. Available. http://www.nccp.org/publications/pub_1099.html.
- Jung, J.K., 2019. Transnational women's online literacies: writing as social action. In: Guzzetti, B.J., Bean, T., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Erlbaum, New York, NY, pp. 182–192.
- Kehler, M., 2010. Boys, books and homophobia: exploring the practices and policies of masculinities in school. *McGill J. Educ.* 45 (3), 351–370. <https://doi.org/10.7202/1003567ar>.
- Kehler, M., 2016. Who will save the boys? In: Guzzetti, B.J., Bean, T. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY, pp. 182–192.
- Kids Count, 2014. Kids count data book. An Annual Report on How Children Are Faring in the United States, the Annie E, 25th ed. Casey Foundation, Baltimore, Maryland. Available. <http://www.aecf.org/resources/the-2014-kids-count-data-book/>.
- Kirkland, D., 2013. Inventing masculinity: young black males, literacy and tears. In: Guzzetti, B.J., Bean, T.W. (Eds.), *Adolescent Literacies and the Gendered Self: (Re)constructing Identities Through Multimodal Literacy Practices*. Routledge, New York, NY, pp. 131–139.
- Knobel, M., Lankshear, C., 2001. Cut, paste and publish: the production and consumption of zines. In: Alvermann, D.E. (Ed.), *Adolescents and Literacies in a Digital World*. Peter Lang, New York, NY, pp. 131–145.
- Kollmayer, M., Schober, B., Spiel, C., 2018. Gender stereotypes in education: development, consequences, and interventions. *Eur. J. Dev. Psychol.* 15 (4), 361–377. <https://doi.org/10.1080/17405629.2016.1193483>.
- Kosciw, J.G., Clark, C.M., Truong, N.L., Zongrone, A.D., 2020. The 2019 national school climate survey: the experiences of lesbian, gay, bisexual, transgender, and queer youth in our nation's schools. GLSEN, New York.
- Lankshear, C., Knobel, M., 2011. *New Literacies*. Open University Press, McGraw Hill.
- Lesley, M.K., 2012. *Invisible Girls: At Risk Adolescent Girls Writing Within and Beyond School*. Peter Lang, New York, NY.
- Lingard, B., 2003. Where to in gender policy in education after recuperative masculinity politics? *Int. J. Incl. Educ.* 7 (1), 33–56. <https://doi.org/10.1080/13603110210143626>.
- Lingard, B., Martino, W., Mills, M., 2009. Boys and Schooling: Beyond Structural Reform. Palgrave Macmillan, Basingstoke, UK.
- Luke, A., Woods, A., Dooley, K., 2010. Comprehension as social and intellectual practice: rebuilding curriculum in low socioeconomic and cultural minority schools. *Theory Into Pract.* 50, 157–164. <https://doi.org/10.1080/00405841.2011.558445>.
- Martino, W., 1995. Deconstructing masculinity in the English classroom: a site for reconstituting gendered subjectivity. *Gen. Educ.* 7 (2), 205–220.
- Martino, W., Kehler, M., 2007. Gender-based literacy reform: a question of challenging or recuperating gender binaries. *Can. J. Educ.* 30 (2), 406–431.
- Meier, M.D., Diefenbach, H., 2020. The OECD between political and scientific agendas—a critique of the 2015 PISA gender report. *Gen. Educ.* 32 (5), 626–645. <https://doi.org/10.1080/09540253.2018.1471198>.
- Montoya, S., 2018. Defining Literacy. UNESCO Institute for Statistics. Retrieved. http://gamli.uis.unesco.org/wp-content/uploads/sites/2/2018/12/4.6.1_07_4.6-defining-literacy.pdf.
- Muhammad, G., McArthur, S., 2015. "Styled by their perceptions": black adolescent girls interpret representations of black females in popular culture. *Multicult. Perspect.* i, 133–140.
- National Center for Education Statistics, 2011. Progress in International Reading Literacy Study. Retrieved. <https://nces.ed.gov/surveys/pirls/pirls2016/index.asp>.
- National Assessment of Educational Progress, 2019. NAEP Report Card. Retrieved. <https://www.nationsreportcard.gov/reading/nation/groups?grade=8>.
- National Education Association, 2012–2015. Identifying Factors that Contribute to Achievement Gaps. The National Education Association, Gaithersburg, Maryland. Available. <http://www.nea.org/home/17413.htm>.
- Neal, L., 2003. The effects of African American movement styles on teachers' perceptions and reactions. *J. Spec. Educ.* 37 (2), 98–120.
- Newkirk, T., 2000. Misreading masculinity: speculations on the great gender gap in writing. *Language Arts* 77 (4), 294–300.
- Oakley, A., 2016. Disturbing Hegemonic Discourse: Nonbinary Gender and Sexual Orientation Labeling on Tumblr. *Social Media and Society*, pp. 1–12. <https://doi.org/10.1177/2056305116664217>. July–September.
- Office for Economic Co-operation and Development (OECD), 2014. PISA 2012 Results: What Students Know and Can Do—Student Performance in Mathematics, Reading and Science. OECD. Retrieved from. https://www.oecd-ilibrary.org/education/pisa-2012-results-what-students-know-and-can-do-volume-i-revised-edition-february-2014_9789264208780-en.
- Office for Economic Co-operation and Development (OECD), 2015. *The ABC of Gender Equality in Education*. PISA, OECD Publishing, Paris.
- Office for Economic Co-operation and Development (OECD), 2019. PISA 2018 Results: What Students Know and Can Do. OECD. Retrieved. <https://www.oecd.org/pisa/publications/pisa-2018-results-volume-i-5f07c754-en.htm>.
- Olcott, K., Mahar, D., 2003. "Tech-Saviness" meets multiliteracies: exploring adolescent girls' technology mediated literacy practices. *Read. Res. Q.* 38 (3), 356–385.
- Paechter, C., 2006. Masculine femininities/feminine masculinities: power, identities and gender. *Gen. Educ.* 18 (3), 253–263.
- Paechter, C., 2007. *Being Boys, Being Girls: Learning Masculinities and Femininities*. Open University Press, Maidenhead, UK.
- Paechter, C., 2012. Bodies, identities and performances: reconfiguring the language of gender and schooling. *Gen. Educ.* 24 (2), 229–241. <https://doi.org/10.1080/09540253.2011.606210>.
- Perrin, A., September 26, 2019. Who doesn't read books in America?. In: FactTank: News in the Numbers. Pew Research Center. Retrieved. <https://www.pewresearch.org/fact-tank/2019/09/26/who-doesnt-read-books-in-america/>.
- Persson, A., Newman, C.F., Rasmussen, M., Marshall, D., Cover, R., Aggleton, P., 2020. Queerifying notions of "difference" among two generations of Australians who do not identify heteronormatively. *Sex. Cult.* 24, 54–71. <https://doi.org/10.1007/s12119-019-09625-3>.

- Ransaw, T., 2019. Reading fatherhood: the importance of fathers in children's literacy development. In: Guzzetti, B.J., Bean, T., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Erlbaum, New York, NY, pp. 208–220.
- Reilly, D., 2012. Gender, culture and sex-typed cognitive abilities. *PLoS One* 7, 7. Retrieved. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0039904> 58.
- Reilly, D.L., Neuman, D.L., Andres, G., 2018. Gender differences in reading and writing achievement: evidence from the national assessment of educational progress. *Am. Psychol.* 74 (4), 445–458.
- Retelsdorf, J., Schwartz, K., Asbrock, F., 2015. "Michael Can't Read!" Teachers' gender stereotypes and boys' reading self-concept. *J. Educ. Psychol.* 107 (1), 186–194. <https://doi.org/10.1037/a0037107>.
- Rigg, P., 1985. Petra: learning to read at 45. *J. Educ.* 167 (1), 129–140.
- Rubinstein Avila, E., 2007. From the Dominican republic to drew high: what counts as literacy for Yanira Iara? *Read. Res. Q.* 42 (4), 568–589.
- Sanford, K., 2005. Gendered literacy experiences: the effects of expectation and opportunity for boys' and girls' learning. *J. Adolesc. Adult Literacy* 49 (4), 302–315. <https://doi.org/10.1588/JAAL.49.4.4>.
- Schott Foundation for Public Education, 2011. Schott 50 State Report on Public Education and Black Males. The Schott Foundation for Public Education, Cambridge, MA. Retrieved. <http://blackboysreport.org/national-summary/>.
- Simon, R., Hicks, B.L., Walkland, T., Gallagher, B., Evis, S., Baer, P., 2019. Breaking gender expectations: adolescents' critical rewriting of a trans young adult novel. In: Guzzetti, B.J., Bean, T.W., Dunkerly Bean, J. (Eds.), *Literacies, Sexualities and Gender: Understanding Identities From Preschool to Adulthood*. Routledge, New York, NY.
- Skeggs, B., 1997. *Formations of Class and Gender*. Sage, London.
- Skelton, C., Francis, B., 2003. Introduction: boys and girls in the primary classroom. In: Skelton, C., Francis, B. (Eds.), *Boys and Girls in the Primary Classroom*. Open University Press, Maidenhead, pp. 3–25.
- Skelton, C., 2010. Gender and achievement: are girls the "success stories" of restructured education systems? *Educ. Rev.* 62 (2), 131–142. <https://doi.org/10.1080/00131910903469536>.
- Skelton, C., Francis, B., 2011. Successful boys and literacy: are "literate boys" challenging or repackaging hegemonic masculinity? *Curric. Inq.* 4 (4), 456–479. <https://doi.org/10.1111/j.1467-873X.2011.00559.x>.
- Smith, M., Wilhelm, J.D., 2002. *"Reading Don't Fix No Chevys": Literacy in the Lives of Young Men*. Heineman, Portsmouth, New Hampshire.
- Spender, D., 1982. *Invisible Women: The Schooling Scandal*. Writers and Readers Publishing, London.
- Stanley, L., Wise, S., 1993. *Breaking Out Again: Feminist Ontology and Epistemology*, 2nd ed. Routledge, London.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Street, B., 1996. Literacy and power? Paper originally written for an international seminar on "literacy and power?" held in Harare, Zimbabwe August, 1995. *Open Lett.* 6 (2), 7–16. Available. <https://search.informit.com.au/documentSummary;dn=970909190;res=IELAPA>.
- Street, B., 2014. *Social Literacies: Critical Approaches to Literacy Development, Ethnography and Education*. Routledge, New York, NY.
- Subrahmanian, R., 2005. Gender equality in education: definitions and measurements. *Int. J. Educ. Dev.* 25, 395–407. <https://doi.org/10.1016/j.ijedudev.2005.04.003>.
- UNESCO, 2012. *Education for All: Global Monitoring Report: Teaching and Learning: Achieving Quality for All*. UNESCO, Paris, France.
- UNESCO, 2013. *Adult and Youth Literacy: Fact Sheet*. UNESCO Institute for Statistics.
- UNESCO, 2014. *Education for All: Global Monitoring Report: Teaching and Learning: Achieving Quality for All*. UNESCO, Paris, France.
- Wargo, J.W., 2015. Every Selfie Tells a Story: LGBTQ Youth Lifestreams and New Media Narratives as Connective Identity Texts. *New Media & Society*, pp. 1–19.
- Wargo, J.M., 2017. #donttagyour hate: reading collecting and curating as genres of participation in LGBT youth activism on Tumblr. *Digital Cult. Educ.* 9 (1), 14–31.
- Warner-Griffin, C., Liu, H., Tadler, C., Herget, D., Dalton, B., 2017. Reading Achievement of U.S. Fourth-Grade Students in an International Context: First look at the Progress in International Reading Literacy Study (PIRLS) 2016 and ePIRLS 2016 (NCES 2018-017). National Center for Education Statistics, Washington, DC. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018017>.
- Watson, A., Kehler, M., Martino, W., 2010. The problem of boys' literacy underachievement: raising some questions. *J. Adolesc. Adult Literacy* 53 (5). <https://doi.org/10.1598/JAAL.53.5.1>.
- Wenger, E., 1999. *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, Cambridge, UK.
- Williams, G., 1998. Children entering literate worlds: perspectives from the study of textual practices. In: Christie, F. (Ed.), *Literacy and Schooling*. Routledge, London, pp. 47–73.
- Wu, J.Y., 2014. Gender differences in lone reading engagement, metacognitive strategies, navigation skills and reading literacy. *J. Comput. Assist. Learn.* 30, 252–271.

Literacy at the intersections of race, gender, and sexuality

Latrise P. Johnson^a and Mollie V. Blackburn^b, ^aThe University of Alabama, Tuscaloosa, AL, United States; and ^bThe Ohio State University, Columbus, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Learning with(in) white and straight places	262
Literacy normativity, Black queer literacies, and restorative literacies: theoretical considerations	262
The literate lives of queer students of color: empirical scholarship	263
Sites of normativity	263
Subverting sites of normativity	264
Consequences of queering	265
Sites of promise: implications for teachers and students	266
References	267

Learning with(in) white and straight places

Many US institutions founded in the name of citizenry, faith, education, and “public good” rely on racist, classist, and cishetero-normative ways of becoming in order to create productive members of society. Primarily, racial and Christian hegemony operate within societal institutions that shape a pervasive and systematic set of values and beliefs that, in turn, teach the public about who is good, what is (in)appropriate, and who deserves access. Systems of racial, class, gender, and sexuality oppression make possible the normalization of practices that disrupt educational and social opportunities for African American and LGBTQ people. Further, these systems legislate, miseducate, and relegate them to the margins of society and acceptance. Many schools in the US, for example, reify hierarchical structures that support the reproduction of ideas and ideals in order to maintain order, impose regimented activity, and surveil text, bodies, and actions that threaten that order. These structures make possible the “academic” discourses that deny students’ funds of knowledge (Moll et al., 1992); discourses of standardization that marginalize language; discipline discourses that push Black and Brown bodies out of classrooms and schools; and discourses of silence where queer, trans, differently abled, and people with other nonconforming identities are denied spaces to exist and thrive with(in) the structures of school.

Through racist, heteronormative, and homophobic discourses (Blackburn, 2005), schools (re)create practices of violence against the “othered” while teachers inherit violent language that regards particular bodies, namely Black and queer, as dangerous and pathological. This violence is apparent with(in) literacy practices that “correct” African American Language speakers, only teach and center canonical (and canonical-like) texts, and stifle identity- and meaning-making through writing and other forms of literacy participation. Such practices perpetrate violence in the name of standardization and “academic” achievement while perpetuating the identities, values, and beliefs embedded within those practices. Namely, that ways of knowing and being that fall outside of straight, white, American English, and Christian are subject to oppression, injustice, and inequity.

According to Johnson (2017), “Educators are in need of theoretical and pedagogical models in which restorative literacies are enacted in order to undo the violence against and silencing of [African American] queer youth that happens in US schools” (p. 30). We draw on Pritchard’s (2017) notions of normative and restorative literacies for Black queer people to offer such a theoretical model, and we review empirical literature to explore how and where literacy normativity impacts the literate lives of Black queer youth, and other queer youth of color in the US. Moreover, we illuminate the ways that the youth represented in these studies used literacy to link their racial, gendered, and sexual identities in ways that challenged normativity. We illuminate the ways they engaged in restorative literacies as a promising pathway toward enacting pedagogies and practices of transformation and healing for themselves, as Black and queer youth of color specifically, and all youth, writ large.

Literacy normativity, Black queer literacies, and restorative literacies: theoretical considerations

In his book on Black queer literacies, Pritchard (2017) defines literacy normativity as “the use of literacy to create and impose normative standards and beliefs onto people who are labeled alien or other through textscapes that are experienced as painful because they do damage or inflict harm” (p. 28). Pritchard asserts that

[I]literacy normativity steals emotional resources from people, wounding people through texts. Moreover, literacy normativity places limits on what is considered literacy and who is considered literate. Not only does literacy normativity make possible harmful discourses that “marginalize, ostracize, and condemn people for their identities and other ways of being” (p. 28).

It also operates to leave out, dismiss, and correct voices of the “othered.” In intellectual spaces, (e.g., academia, scholarship, schools), literacy normativity plays out in the ways that prioritize particular voices that echo white, heteronormative, standardized, and middle-class ways of reading, writing, using language and other literacy events (see [Heath, 1983](#)) and dismisses opposition to normativity as extracurricular, nonscholarly, artistic, and/or personal.

Queer pedagogies provide the space to reimagine literacies in that it leans on criticality and invites us to think about queerness in relation to teaching the English Language Arts. In their positioning queerness as more than being connected to queer texts, students, or teachers, [Waite \(2017\)](#) considers “queer forms” as:

non-normative and category-resistant forms of writing that move between the critical and the creative, the theoretical and the practical, the rhetorical and the poetic, the queer and the often invisible normative functions of the classrooms (p. 6).

Such underpinnings provide us with ways to interrogate Black queer literacies and understand them within the context of literacy teaching and literacy practice. In order to argue the harm of literacy normativity, we consider Black Queer Literacies as antinormative praxis ([Pritchard, 2017](#)) as they provide an intersectional landscape for understanding how, when, and for what purposes literacies are racialized, gendered, and sexualized. In addition, Black queer literacies provide space to interpret radical praxis of freedom, love, healing, and transformation in the face of “social, political and cultural circumstances in which [people reading and writing, doing/being literacy from the margins] are targeted for debasement, degradation, and in many cases, death” (p. 37). They are, in Pritchard’s words, restorative.

Whereas literacy normativity’s pervasive nature has a hold on all people reading and writing ([Yagelski, 2011](#)), especially in school, restorative literacies are used to subvert dominant ways of doing literacy and being literate as well as to “remak[e] those emotional resources people need for living, especially love, and retur[n] them to work in the best interest of the individual and others” ([Pritchard, 2017](#), p. 24). Restorative literacies invite literacy participation that centers the self ([Johnson, 2017](#)), asserts control over one’s life and voice with, through, and in literacy action and participation ([Johnson, 2017](#); [Pritchard, 2017](#)). Thus, they are associated with choices about language and identities ([Daniell, 1999](#); [Pritchard, 2017](#)).

In the following sections, we focus on Black queer literacies to consider sites of both normative and restorative literacy practices for queer youth of color. We use empirical scholarship to illuminate literacy practices that perpetuate as well as challenge racism and cisheteronormativity and conclude with a discussion that addresses the question(s): What are the consequences of taking up queer in sites of learning? And how does taking up queer (in literacy practice) benefit all students?

The literate lives of queer students of color: empirical scholarship

We turn to educational literacy research about Black queer students, as well as other queer students of color, to explore what normative and restorative literacies look like in the lives of these young people as they navigate spaces in and out of schools. We explore how scholars have interrogated spaces, texts, and literacy practices in order to illustrate how normativity and restoration look in the lives of Black queer youth and other youth of color. Additionally, we analyze this literature in order to illuminate the ways that normativity of spaces, texts, and literacy practices work to dehumanize, debase, and destroy people who identify as queer and how transformation and liberation are possible in and through restorative literacies.

Sites of normativity

In educational literacy scholarship focused on Black queer students, normative literacies are reported as taking place in schools. [Schey \(2020\)](#), for example, introduces readers to Imani, a “young cis Black woman” who identifies “variously, as bi, bi-curious, and questioning” (p. 17); Schey reported that when queer topics were raised in Imani’s sophomore humanities class, they often led to conflict, which provoked the teachers to put an end to them. [Johnson \(2020\)](#) introduces readers to Nilah, an “African American female who named her sexuality as fluid” (p. 641). Johnson noted that there were places at Nilah’s high school where “students are taught not to love who they are, to feel ashamed” (p. 635). Even in the “LGBTQ continuation school,” where [Cruz \(2013\)](#) conducted a study of queer youth of color, the “[c]urriculum that dealt with LGBTQ issues was additive, or surface ... an in-depth examination of LGBTQ literature and lives” (p. 452) was not prioritized. Instead “worksheets dominated most units and lesson plans, [and] very little teaching happened” (p. 453). Not surprisingly, given such realities, Justine, an African American lesbian, explained to [Blackburn \(2002/2003\)](#) that there was “nothing in [her] school journal” (p. 316).

This does not mean that normative literacies only happen in school. They don’t. Such literacies for queer youth can happen at home, when a family member, particularly a parent, embodies homophobia. We saw evidence of this when Nilah, who was in [Johnson’s \(2020\)](#) Creative Writing class, where Johnson provided Nilah with a “safe space for her to question her sexuality” (p. 645). Nilah’s mother called the principal and insisted that Johnson “stop filling her daughter’s head with ‘ideas’” (p. 644). Nilah’s mother’s intervention essentially ended the teacher-student relationship between Johnson and Nilah in which Nilah could explore her sexuality through journal writing, and Johnson could listen, read, and support her. Kira, who identified as a biracial dyke, had

a journal where she grappled with her sexuality, too, but when her foster mother found it, her mother mistreated her until Kira left home (Blackburn, 2003). In these cases, normative literacies were at the homes of these youth.

Social media can also be a site for normative literacies. Consider Zeke, an “African American gay male” (Wargo, 2017a, p. 564). In his high school’s Gay Straight Alliance meeting, Zeke shared a photo he posted on Tumblr that was of a basketball jersey with his name and the year he would graduate from high school. It certainly conveyed a certain version of Black masculinity, as Wargo explores. Zeke explained to Wargo that he used the post “as a way to identify himself as ‘... not one of those loud gays’” (p. 566), indeed in and beyond the post, Zeke performed “gestures of cisgendered masculinity purposefully erasing the queerness” (p. 566). Clearly, normative literacies are not limited to schools.

Subverting sites of normativity

However, neither are schools limited to normative literacies. Indeed, we see evidence of this fact above, when we get a glimpse into Johnson’s classroom in the predominantly Black school where she works as a professor-in-residence. In her Creative Writing class, Johnson (2020) encouraged Nilah to explore questions of her sexuality in her journal and “listened (and read) and provided support for [Nilah’s] struggles” (p. 644). Indeed, Johnson was explicit about her purpose to “inform, awaken, transform, mediate, support, advocate and nurture” in her class. She was deliberate about her purpose to begin and end with love. Clearly, Johnson (2020, 2017) embraces and embodies restorative literacies.

Even in the absence of such powerful teachers, queer Black students engage in restorative literacies (Blackburn, 2002/2003; Cruz, 2013; Schey, 2020). Even when teachers did not invite them in, as Johnson did, students in these studies transformed their experiences in ways that aligned with their identities and ways of being. They rejected normative literacies that excluded or erased them and created restorative literacies that represented and elevated them. There were risks associated with such actions, and they did it anyway. Justine, for example, produced an assignment for her television class in which she read a lesbian-themed poem written by a lesbian author and complemented the poem with images of women who have loved women over decades and, finally, an image of her girlfriend and herself (Blackburn, 2002/2003). Kira wrote a gay-themed poem, story, and book report to fulfill school assignments, even when she was barely able to attend school because of the homophobia she encountered both at school and home (Blackburn, 2003). Cruz (2013) documented a video poem workshop in the high school where she worked. In this workshop, at least six queer students of color produced, shared, and discussed video poems that highlighted their sexualities, but Cruz focused on one Latino student in particular. Peter John Cord explored how his twin brother engaged in survival sex and died of complications related to HIV/AIDS only the year before. Cruz (2013) states that

[cen]tering the stories of LGBTQ youth in video means that students begin with their own stories, write their own books, read each other’s poetry, create their own cultural artifacts. To center the histories of LGBTQ lives is ... about providing a counternarrative ... where young people are not without agency but instead reveal narratives of resiliency and hope. It is life affirming to use video to tell a critical story of LGBTQ youth of color survival and resilience (p. 453).

Such literacies are, in other words, restorative.

In these situations, for these young people, teachers assigned particular formats but invited students to select topics. Choosing a queer topic can be risky, particularly in cisheteronormative schools, as most are. Therefore, it is worth turning our attention to times when teachers initiated queer topics in classrooms and considering the consequences, particularly for Black queer youth. You will recall, from above, that Imani’s sophomore humanities class was one where queer topics were raised by the teachers. Although discussions pertinent to these topics were often shut down, as indicated above, they were not always. For example, the teachers initiated a discussion of “North Carolina’s House Bill 2 (HB2)—more commonly known as the *bathroom bill*—a law then in effect that mandated transphobic state practices and prohibited municipal legal protections for LGBTQ people” (Schey, 2020, p. 17). The class was to debate the bill, and Imani was assigned to assume the role of lawmaker George Holding and therefore argue in support of the bill. One might expect this assignment to epitomize normative literacies, and it could have, but Imani rejected that possibility and asserted a subversive alternative: “She inhabited the role, but she interjected a critical difference ... She used humor and hyperbole, resulting in parody ... [to] reveal the absurdity of Holding’s transphobic views” (pp. 20–21). Thus, Imani shifted what could have been an LGBTQ-inclusive lesson with a normative impact into one with a restorative, and in Schey’s (2020) words, “liberatory” (p. 32), effect.

We also see evidence of restorative literacy events outside of school. Johnson (2017), for example, facilitated an after-school writing club at the same school where she taught Creative Writing. In this club, Black queer students “navigate[d] and disrupt [ed] heteronormativity” and “(re)claim[ed] who they [we]re through their writing” (p. 13). Ava was among those students. He identified as an African American bisexual male, and in his journal, he:

critiqued the heteronormative expectations of the people around him and bore witness to his inability to embody his true self. He used the act of writing to trace the pain of being still, not moving in the ways that made him who he was (p. 25).

Sanavia was also among these students. She identified as a gay African American gender-nonconforming female who said, “writing in my journal just makes me feel better” (p. 25). According to Johnson,

The act of writing the self freed [Sanavia] from the constraints of writing and allowed her to explore her experiences and confront the painful feeling associated with those experiences. Writing in this way also enabled her to restore her sense of self as she admitted that she was able to make it despite her experiences and pain (p. 27).

Anika, another student in the club, identified as a Black queer female. She found, in her journal, that she could “keep it real ... even when it hurts” (p. 27). In this context, she “rewrote herself—her love for another female and her gender nonconforming dress—as normal” (p. 28). Thus, these Black queer youth were able to engage in restorative literacies through their journal writing in an afterschool club.

For some Black queer youth, it is their private journals that created such space. While there was nothing in Justine’s school journal, as noted above, she kept a robust personal journal. It was, in her words, “just [her]” (Blackburn, 2002/2003, p. 316). She read a poem she wrote in it at a literacy group called “Story Time” in an LGBTQ-youth center. The poem represented a threatening interaction at a Pride parade, but in the poem, she was safe. She could fight back. When she read it aloud to her peers at the youth center, they were respectful, they laughed, they applauded. They participated in her restoration via literacies.

In the same youth center and group, a group of students read “Am I Blue?” (Coville, 1994). Among those in the group were Katrina, a young African American trans woman, and Trey, a young African American gay man. In the story is a character named Butch who is violently homophobic but who eventually is outed as gay. This prompted a discussion between these two teens, in which Trey talked about someone at school who reminds him of Butch. This classmate “bothers” Trey, even though he is not out at school. Katrina assured Trey that the classmate was gay. Trey did not seem to accept Katrina’s interpretation until she said, “I’m not going to lie, I did it too. Before I came out. I used to pick on people who were flamey, I used to pick on them ... just because it was so hard for me to come out” (Blackburn, 2005, p. 410). By reading and discussing this story, Katrina was able to help Trey see that the abuse he was experiencing in school revealed as much about the abuser as it did about him. In this way, Katrina helped restore Trey in this queer-friendly context, even if not in school.

This is not the case for all queer-friendly places, which, like any places, can be racist. Kira, for example, used to go in read at a café and bookstore in the “gayborhood,” as an alternative to going to school, before the youth center opened. She would read, seeking images of herself in an effort to figure herself out. She read biographies and autobiographies, and one time she was particularly enthusiastic about a collection of lesbian monologues, since she facilitated the “Thespians and Drama Queens” group at the youth center. She found it at the local feminist and queer bookstore, but when she realized that there was only one monologue about a Black woman, and it was a slave story, she decided the book did not represent her at all. It was, in her words, “lame” (Blackburn, 2003, p. 8).

Online communities provide another space in which Black queer youth can seek restoration by engaging in literacy practices, even though we know that can also be normative, as we saw evidence of in Zeke. But Zeke, later, posted an account of his “First High School Crush” on social media (Wargo, 2017a, p. 572). In doing so, he centered his sexuality and gender and used technology to “speak back to and provide counter-narratives of being schooled while holding the youth identity of ‘black urban student’” (p. 569). Ultimately, Zeke, “maneuvered across topographies of visual representation to write a self that felt most real” to him (p. 576). Further, Gabe, who identified as an African American and gender queer, livestreamed a photo mashup of his face on a model wrapped around Justin Bieber. They explained, “I want to show people my sexuality. This is the only way I can” (Wargo, 2017b, p. 150). The importance of this was underscored by their saying, “you can’t be Black and gay [at this school]; you unfortunately just have to pick one” (p. 151). According to Wargo, Gabe, among others, used “multimodal (counter) storytelling” to “stitch[h] together hurt, angst, desire, and (in)equality” in ways that “rendered knowing and being and for being known differently” (p. 158). In these ways, for these young people, multimodal literacies were also restorative ones.

When we reflect on these young people’s literacy events and practices in and out of schools, we see not that some events and practices—even when they are LGBTQ-inclusive—or even places—even when they are queer-friendly—are catalysts for, or homes to, either normative or restorative literacies. Even a normative practice can be reworked to be restorative, but so too can a restorative one fall prey to normativity. It is in the multiple, variable, and fluid nature of literacy practices that there is both danger and promise. These practices, normative and restorative, provide sites of inquiry where we may reimagine literacy teaching and participation in order to disrupt fixed notions of the personal and the public and what of that “belongs” in the classroom.

Consequences of queering

Pritchard (2017) and Waite (2017) foreground the queer in their studies of literacy in general and writing in particular, respectively. Moreover, the queer undergirds the studies of Blackburn, Cruz, Johnson, Schey, and Wargo. We see evidence of this in the multiplicity, variability, and fluidity of literacies and identities as well as the consequences of such intersections. We see it in the disruptions of norms, whether they are norms of literacies or identities. In some cases, the disruptions of literacy and identity happened

simultaneously. And if we look more closely, we see resistance to cisheteronormativity in the students' critical, reflective, and embodied engagements with literacies.

Ava wrote about wanting to dance in the halls (Johnson, 2017), and then, inspired by the band practicing in a nearby room, he embodied this desire. He

offered a couple of eight-count majorette-style dance moves where he flung his arms and rolled his head and neck just like the school's [dancers]. ... He ended with a move that landed him on the floor, one leg bent at the knee, the other straight out and the rest of his body flat on the floor.

Johnson (2020, p. 643).

Those in the afterschool club with him smiled and cheered, and Ava sat down and "wrote in his journal quietly" (p. 643). Ava's resistance, through his dancing and writing, interrupts institution-sanctioned bodies of knowledge meant to "codify and compartmentalize what we know from what we do, from what we are, from the lived experience of our bodies" (Waite, 2017, p. 18). Instead, Ava used movement and writing to embody and reflect on his identity with/in a literacy community that supported his identity-making.

Imani, in the example above, not only assumed the role of a white transphobic lawmaker, as she was assigned to for her class's debate about HB2 (Schey, 2020). She parodied him. She mocked him. She made it perfectly obvious that the view he held offended her. They were, in her view, absurd. And she conveyed all of this to her large class without breaking character. Imani used her body to teach and assumed an intelligibility that, according to Waite, "is to be seen through the lens of what already is, or what already appears to be" (p. 35). In other words, she uses her body and being as a way to critique the absurdity of policies used to disenfranchise trans and other historically marginalized people.

Katrina selected and read aloud the role of the "fairy godfather," in the literacy group at the youth center, and her performance was a delight (Blackburn, 2005). But after the performance, when the group discussed the story, Katrina revealed the ways she had embodied homophobia to protect herself, which was, we imagine, a painful revelation, but she did it to help Trey understand why someone was doing this to him, just as Butch had done in the story. Her painful revelation provoked by the story was a gift to Trey. Her reflection made real the consequences of not just homophobia, but the ways that it affects an entire community.

By looking closely at youth engaged in literacy practices, we see how the spaces—in and out of classrooms, online and face-to-face—are improved by queering. We see how the queered writing of these young people is of particular significance as they explored and asserted their identities. We see how the literacies, themselves, are strengthened when queered. But, more importantly, we see how this queering benefits Ava, Katrina, and Imani. We see how this queering restores them.

Sites of promise: implications for teachers and students

Teachers can be a part of making such restoration possible, for Black queer youth, for queer youth of color, but also for all of their students. We believe that normative literacies are damaging to all who are required to limit themselves to them. While they might be more damaging for some people than others, and more damaging in some situations than others, they are damaging. In contrast, restorative literacies have healing potential for all of those who engage with them. To be clear, we do not believe that there are a clear set of literacies that are "normative" and therefore "bad" and another set that are "restorative" and therefore "good." If queer theory has taught us anything, it is that dichotomies fail us. Instead, there are literacies; likely any of them have the potential for danger, the potential for promise. Different literacies will be engaged with and even embodied differently by different people. As teachers, we are responsible for knowing our students and for knowing which literacy practices and events have the potential for promise, the potential of restoration for our students, for all of our students, including but not limited to our Black, queer students.

Literacy educators who strive to make their classrooms and schools queer by engaging their students in literacy practices and events that are restorative can get guidance from the scholars who worked with Imani, Nilah, Justine, and Zeke, among others. Educators must reflect on their own biases (Blackburn, 2005), not only against LGBTQ people but also people of color. They must own these biases and then work every moment to disown them. Kendi (2019) describes antiracism as relying on a rejection of other forms of oppression, such as homophobia and transphobia. Further, he describes antiracism as a sticker rather than a tattoo; that is to say, that it is something that we must claim for ourselves and work to make happen every moment of every day; it is not something that we accomplish and then check off of our list. Likewise, taking an antinormative stance requires that teachers rely on a rejection of discourses of standardization that relegate Black and Brown students to the margins of educational opportunity. Educators must understand that even though they are doing antiracist and antinormative work, not everyone is. Because of that reality, they must work tirelessly to ensure that the queer youth of color in their classrooms and schools are protected from the curriculum, disciplinary, and actual violence that often accompanies racism, homophobia, and transphobia. Of course, it is not enough to protect students from violence, but it is imperative.

In such a context, literacy educators can engage students in "writing that connects the self and the world," which Johnson (2017) describes as including language "to question the social construction of the self and challenge the status quo in an effort to discover alternative paths for self- and social development" (p. 30). In this way, Johnson asserts, literacy teachers can "invite students to compose themselves in ways that enact their queer identities" (p. 30). Educators might also take the initiative to "design literacy

learning opportunities that explicitly center queer-inclusive curricular content" (Schey, 2020, p. 35). By doing this, teachers take on the responsibility of having chosen to include queer content as well as queer forms (Waite, 2017) which relieves students of the potential homophobic and transphobic consequences for having made such choices themselves.

Beyond this, literacy educators must prepare queer students of color "to interrogate perspectives, produce countertexts, and imagine a more just world" (Johnson, 2017, p. 30). In short, they must prepare their students to use literacies for anti-racist and queer activism. Johnson (2017) challenges literacy educators to encourage students "to write the self ... in ways that allow them to recognize and fight injustice" (Johnson, 2017, p. 30) and to "provide literacy practice aimed at empowering [students] to use their voices against racism, homophobia, and injustices while learning to be in a world where these are realities" (Johnson, 2017, p. 30). Wargo (2017b) encourages literacy teachers to pay attention to the ways "historically marginalized youth ... use an array of communicative resources to remix, remediate and redistribute stories of inequity" (p. 158). Schey (2020), however, notes that sometimes when queer students of color engage in activism in their classrooms, others push back against them, resulting in conflict. Sometimes, teachers are so uncomfortable with this conflict that they terminate the discussion, which essentially terminates the activist effort. Furthermore, Schey suggests that educators

reflect on how they respond when conflict arises in the classroom, especially around topics of sexual and gender diversity. ... attempt to understand what motivates their pedagogical decisions and how these decisions affect youth. ... [Do they] reify oppressive values and practices ... or ... invite people to resist, contest, and even transform them (p. 34).

Moreover, Schey (2020) asks that teachers "attempt to stick with conflict—even though it is often incredibly uncomfortable—when it arises around contentious topics such as queerness" and to "consider whose comfort and safety is maintained when conflict is regularly shut down" (p. 35). Together, these scholars challenge educators to prepare students to assert themselves, but also let them do it, in their classrooms, even when it makes them uncomfortable.

We argue that restorative literacies, which are, at the core, anti-racist and queer, benefit all students—certainly queer students of color, but not solely queer students of color. We argue that all students benefit from being protected against violence, even violators, who are made lesser as they embrace violent ideals and commit violent acts. All students benefit from teacher support to write, read, and use language in ways that permit students to come to know more about themselves in relation to the world around and beyond them. This is an advantage that many students have already and one that queer students of color are owed. All students benefit, too, from being prepared to assert themselves and to undo oppressions, whether they are oppressors or oppressed; either way, their dignity suffers in a context where oppression reigns. Enacting practices and pedagogies grounded in restorative literacy requires a commitment to improving the lives of Black, queer students, but also obliges belief that in their value even more broadly speaking, not only because we could all benefit from self-definition, self-care, and self-determination, but also because we all benefit from being a part of a community where everyone is free.

References

- Blackburn, M.V., 2002–2003. Disrupting the (hetero)normative: exploring literacy performances and identity work with queer youth. *J. Adolesc. Adult Lit.* 46 (4), 312–324.
- Blackburn, M.V., 2003. Losing, finding, and making space for activism through literacy performances and identity work. *Penn GSE Perspect. Urban Educ.* 2 (1). <https://urbanedjournal.gse.upenn.edu/archive/volume-2-issue-1-spring-2003/losing-finding-and-making-space-activism-through-literacy-perfo>.
- Blackburn, M.V., 2005. Disrupting dichotomies for social change: a review of, critique of, and complement to current literacy scholarship on gender. *Res. Teach. Engl.* 39 (4), 398–416.
- Coville, B., 1994. Am I blue? In: Bauer, M.D. (Ed.), *Am I Blue?: Coming Out From the Silence*. HarperTrophy, New York, pp. 1–16.
- Cruz, C., 2013. LGBTQ youth of color video making as radical curriculum: a brother mourning his brother and a theory in the flesh. *Curric. Inq.* 43 (4), 441–460.
- Daniell, B., 1999. Narratives of literacy: connecting composition to culture. *Coll. Compos. Commun.* 50 (3), 393–410.
- Heath, S.B., 1983. *Ways With Words*. Teachers College Press, New York.
- Johnson, L.P., 2017. Writing the self: black queer youth challenge heteronormative ways of being in an after school writing club. *Res. Teach. Engl.* 52 (1), 13–33.
- Johnson, L.P., 2020. Professor in residence model for queering the edges of school: a project in humanization and love. *J. Qual. Stud. Educ.* 33 (6), 634–648.
- Kendi, I.X., 2019. *How to Be Antiracist*. Bodley Head, London.
- Moll, L.C., Amanti, C., Neff, N., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31 (2), 132–141. <https://doi.org/10.1080/00405849209543534>.
- Pritchard, E.D., 2017. *Fashioning Lives: Black Queers and the Politics of Literacy*. Southern Illinois University Press, Carbondale.
- Schey, R., 2020. Youths' literacy disidentifications in a secondary classroom: contesting transphobia through humor and role-playing. *Teach. Coll. Rec.* 122 (070312), 1–42.
- Waite, S., 2017. *Teaching Queer: Radical Possibilities for Writing and Knowing*. University of Pittsburgh Press, Pittsburgh.
- Wargo, J.M., 2017a. "Every selfie tells a story ...": LGBTQ youth livestreams and new media narratives as connective identity texts. *New Media Soc.* 19 (4), 560–578.
- Wargo, J.M., 2017b. Designing More Just Social Futures or Remixing the Radical Present? English teaching: Practice & critique.
- Yagelski, R.P., 2011. *Writing as a Way of Being: Writing Instruction, Duality, and the Crisis of Sustainability*. Hampton Press, New York.

Literacy access, the law, and racism in the United States: a critique

Arlette Ingram Willis, University of Illinois, Urbana Champaign, College of Education, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Understanding the roots of literacy in the United States	268
African, Indigenous, and Spanish literacies	269
Codifying anti-literacy through US courts	270
Anti-black racism and white supremacy	270
Civil War Amendments and equal rights	272
Systemic racism in politics and education	273
Codifying racism	274
President Lyndon B. Johnson (1963–1969) reign	274
Systemic racism in education	275
State court cases	276
Looking forward	276
References	277

Understanding the roots of literacy in the United States

Ancient civilizations of African, Asian, Indigenous, and Latinx people who inhabited the Americas before European settler-colonialists arrived had developed literate histories and educational systems, including colleges and universities (or equivalents), centuries before the Common Era. As well, each society possessed literacies used by its members to communicate with one another. Nearly half of Africans brought to the United States came from Senegambia (Senegal, Gambia, Guinea-Bissau, and Mali) and West-central Africa (Angola, Congo, Democratic Republic of Congo, and Gabon). These regions in Africa established universities, such as the University of Timbuktu, Mali, (982 CE). China also established imperial colleges during the Western Han dynasty (206 BC–9 AD), and Mayans established colleges in Mexico, College of Santa Cruz de Tlatelolco (1536). These societies were literate and richly cultured, intellectual, sophisticated, and complete. The idea that only Europeans were civilized, and they alone possessed literacy, is a false narrative created to obfuscate cultural, educational, linguistic, and literary histories of Black, Indigenous, and People of Color (BIPOC). The false narrative was developed to depict non-White ancient civilizations as primitive, their people as less-than-human, and to justify European settler-colonialism's castration, terrorism, theft, rape, slaughter, and violence perpetrated against BIPOC. The false narrative also is used to deflect from the ideological and legal barriers to literacy access created by settler-colonialists and generations of dominant power elites that followed.

In this historical moment, the US is repudiating the false narrative and revealing the original American sin of racism and exposing the Big Lie: in the US all people are created equal, citizens receive due process, and have equal rights and privileges under the law. We have a moral responsibility to reveal and correct the past as we move forward. This chapter is premised on the fundamental idea that literacy is a human right, meaning that literacy is necessary to function in society. I interrogate literacy history and use counter-narratives of BIPOC to examine the intersection of access to literacy, laws, and racism. The result is a case study of how systemic institutional and structural racism pre-empt the literacy progress of BIPOC.

The text is framed as a critique, in the Foucauldian (1988) sense, where "a critique is not a matter of saying that things are not right as they are. It is a matter of pointing out on what kinds of assumptions, what kinds of familiar, unchallenged, unconsidered modes of thought the practices we accept rest" (p. 154, cited in Gore and Morrison, 2001, p. 568). I identify an ideology of White supremacy as underpinning the systematic denial of literacy access. Ansley (1997) remarks that White supremacy, represents

a political, economic, and cultural system in which whites overwhelmingly control power and material resources, conscious and unconscious ideas of white superiority and entitlement are widespread, and relations of white dominance and non-white subordination are daily reenacted across a broad array of institutions and social settings. (p. 592)

I extend his description to the legal system by examining the intersection of laws, literacy, and racism "to better understand how race functioned [and functions] in conjunction with educational policy" (Child and Klopotek, 2014, p. 6), and how White supremacy is "reconstitutive" (Crenshaw, 2020). Central to this examination is exposing White supremacy within law, literacy, and society. The text hovers over the early years of nation-building to demonstrate attempts to codify, justify, legalize, normalize, and politicize White supremacy. To anchor such a huge undertaking, I draw on US Supreme Court law as a guidepost and BIPOC narratives as counter-hegemonic responses. Briefly reviewed, are three landmark Supreme Court cases (*Dred Scott v Sanford*, *Plessy v Ferguson*, *Brown v Board*) as they represent laws and legal decisions established by the country's penultimate court. They also expose

how “racism was supported by the rule of law and perpetrated by constitutional mandate” (Browne-Marshall, 2013, p. xx). Further, I include a review of selective state and local laws that address race to provide context about racial inequality and its effects on the lives of individuals, families, and communities.

A single chapter cannot adequately represent centuries of sociohistorical contexts, laws, literacy, and racism; thus this chapter provides a limited and selected review. The chapter sections include early U. S. history (laws, literacy, and racism), followed by like-minded discussions in the 20th and 21st centuries. I situate events within narratives about the lived experiences of BIPOC, in which “situatedness posits that the activity of the mind (in all its cognitive, emotional, perceptual, and social capacities) is the causal product of bodily, social, environmental, and cultural interactions and that these factors serve as both the implicit causal backdrop and ontological basis for the mind” (Costello, 2014, p. 1757).

Thus, situatedness helps to reveal the human toll of White supremacy expressed in BIPOC counter-narratives and disentangles how barriers to literacy were supported by federal, state, or local laws. BIPOC counter-narratives unveil how literacy was conceptualized, framed, and practiced. Ideally, in a chapter that addresses race, all races would be equally and sufficiently covered; however, space limitations prevent a thorough retelling.

African, Indigenous, and Spanish literacies

When European settler-colonists arrived on the shores of the “new world,” millions of Indigenous people called the land home. Dunbar-Ortiz (2014) emphasizes the goal of European settler colonialism was “a genocidal policy ... to terminate their existence as peoples—not as random individuals” (p. 6). Castillo (2015) also details how Catholic priests, under the auspices of the Spanish government, took possession of Indigenous lands in the West, enslaved Indigenous people, and forbade teaching Indigenous people literacy. McCarty and Lee (2014) note that Indigenous people historically have been considered sovereign tribal nations, and acknowledge “tribal sovereignty is inherent, predating the US Constitution, but is also recognized within the Constitution and in treaties and case law. The cornerstone of the tribal-federal relationship is a legally and morally codified relationship of *trust responsibility* that is both voluntary and contractual” (p. 102, emphasis in the original).

Irrespective of the Constitutional (1787) mandate, Congress sanctioned a Civilization Plan (1789) to acquire Indigenous lands, use missionaries to convert Indigenous nations to Christianity, and teach English to Indigenous people. By 1819, the Choctaw nation avowed its interest and commitment to education, established and maintained schools, and taught students to read and write in their native language as well as English (Spring, 1996, p. 14). The Cherokee people also understood the value of education and literacy.

Sequoyah (1760–1843), a full-blood Cherokee, invented the Cherokee syllabary, that “empowered those who spoke the language to read and write” (Parins, 2013, p. xiv). Many Cherokee people learned to read and write: “in a matter of weeks, large numbers of the population became literate in their native language” (p. xiv). The rapid literacy acquisition by 1826, was based on the use of Cherokee symbols, unlike attempts to replicate the Cherokee language using the English (Roman) alphabet. The Cherokee syllabary, “consisted of 86 characters, each of which represented a Cherokee syllable. Speakers ... could memorize the characters, and within days they could read anything written using the syllabary” (Parins, 2013, p. 32). The Cherokee nation published their first newspaper the *Cherokee Phoenix* in 1828, in Cherokee and English, among Cherokee living in New Echota. Parins (2013) notes the Cherokee syllabary “became a gateway to a world of intellectual possibilities and had profound effects on the political, social, economic, and educational affairs of the Cherokee Nation that continue to this day” (p. xiv). It created an opportunity for communication among the Cherokee and helped “to preserve their culture and heritage, and to prove to the world that their society was viable and distinct” (Parins, 2013, p. xvi). Many Cherokee people also were able to read and write in English. (My paternal grandmother was a full-blood Cherokee of the Echota tribe.) The settler-colonialists and missionaries mounted a campaign to delegitimize the Cherokee syllabary, in support of the English alphabet. Given that literate Cherokee people were not dependent on White settler-colonialists to communicate, govern, or transact business, White settler-colonialist experienced a loss of power and control. It should also be noted that while the Chickasaw and Cherokee nations developed literacy, other Indigenous nations believed the “word” was sacred, “present,” and should not be written or preserved in written form.

The Indian Removal Act (1830) also was passed during this period as it forced the migration of five nations of Indigenous people (Cherokee, Chickasaw, Choctaw, Creek, and Seminole) from their fertile farmland in the Southeast to land west of the Mississippi. When the Indigenous peoples arrived, they were confronted by White colonizers who did not accept them as equally human and ignored the Constitutional mandate: “to protect or enhance tribal assets (including fiscal, natural, human, and cultural resources) through policy decisions and management actions” (Wilkins & Lomawaima quoted in McCarty & Lee, p. 102). Indigenous people were not alone in their quest for freedom or education, nor were they alone in their quest to acquire literacy.

African people were first brought to North America, now Florida in 1565, by a Spanish conquistador. The false narrative, however, begins with the arrival of African people in Virginia (1619), in large part because it creates a connection to Northern Europe. Initially, African people were brought as indentured servants, people who could work a determined number of years and purchase their freedom. In less than 50 years, however, their status changed when Slave Codes were created to restrict the lives of enslaved African people: “stripped of their names, religions, identities, nationalities, and legal status as human beings” (Browne-Marshall, 2013, p. 3). The Slave Codes justified, and legalized White supremacy as enslaved Africans were not considered people but the property of their owners, and chattel slavery had no end.

Codifying anti-literacy through US courts

As the fledgling nation struggled to enact a democracy, a deal was struck during the Constitutional Convention (1787), between the Northern and Southern states for representation in the House of Representatives. State representatives would be determined by counting enslaved Black people as three-fifths of a (White) person (Three-Fifth's Compromise). The lives of enslaved Black people were regulated and surveilled by White people, they were not permitted to assemble (unless a White person was present), own fire-arms, testify in a court of law, or leave the premises of a White person without permission. They also were not permitted to: marry, own property, learn (or teach someone) to read or write, nor read literature that might excite them and encourage rebellion. Black people, however, sought an education—as [Anderson \(1988\)](#) argues “the denial of opportunities for schooling has translated into the denial of access to literacy” (pp. 35–36).

Most of the Southern states established laws prohibiting enslaved Black people from learning to read, and from teaching others to read. In 1818, North Carolina's General Assembly (1818), enacted an anti-literacy law, later revised and entitled “A Bill to Prevent All Persons from Teaching Slaves to Read or Write, the Use of Figures Excepted” (1830):

Whereas the teaching of slaves to read and write, has a tendency to excite dissatisfaction in their minds, and to produce insurrection and rebellion, to the manifest injury of the citizens of this State:

... any free person, who shall hereafter teach, or attempt to teach, any slave within the State to read or write, the use of figures excepted, or shall give or sell to such slave or slaves any books or pamphlets, shall be liable to indictment ...

... That if any slave shall hereafter teach, or attempt to teach, any other slave to read or write, the use of figures excepted, he or she may be carried before any justice of the peace

n.p., [Legislative Papers \(1830–31 Session of the General Assembly\)](#).

To further discourage access to literacy, enslaved Black people were described as subhuman and “uneducable.”

Black and White people broke laws and taught Black people to read and write. Harriet Ann Jacobs (1813–1897) who was born as an enslaved person in Edenton, North Carolina, learned to read and write when she was taught by her first White mistress. In her narrative, *Incidents in the life of a slave girl: Written by herself* (1861), Jacobs acknowledged that she also taught enslaved Black people to read and write. In a touching example, she described teaching an elderly enslaved man, Uncle Fred:

I asked him if he didn't know it was contrary to law; and that slaves were whipped and imprisoned for teaching each other to read. This brought the tears into his eyes. “Don't be troubled uncle Fred,” said I. “I have no thoughts of refusing to teach you. I only told you of the law, that you might know the danger, and be on your guard.” ... I selected a quiet nook, where no intruder was likely to penetrate, and there I taught him his A, B, and C. Considering his age, his progress was astonishing (p. 112).

With the support of White abolitionist, Lydia M. Child, Jacobs published her narrative that includes a description of the horrors of slavery, a strong pronouncement of Christianity, and fervent support of the importance and value of literacy. In some Northern communities, however, Black children attended schools and learned to read and write.

After the Massachusetts Supreme Court (1783) case, the *Commonwealth v. Jennison* abolished slavery, Black school children in Boston, Massachusetts, were permitted to attend public schools with White children. However, this was often fraught as Black children faced racial harassment and “physical and emotional discrimination that led to the creation in 1798 of the Smith School, a private school for Blacks” ([Browne-Marshall, 2013](#), p. 20). Benjamin F. Roberts brought a lawsuit against the City of Boston's School Committee on behalf of his young daughter, Sarah, *Roberts v Boston* (1850), so she could attend the White school close to their home. The presiding Chief Justice in the case, Lemuel Shaw, ruled in favor of the Boston School Committee noting how people perceive racial differences “is not created by law, and probably cannot be changed by law” ([Massachusetts Historical Society, n.p.](#)). [Browne-Marshall \(2013\)](#) maintains the Roberts' lawsuit “would form the cornerstone of future court decisions legally segregating children in public schools” (p. 21). This judicial decision also normalized anti-Black racism, anti-literacy, and unequal access to education.

Anti-black racism and white supremacy

The Supreme Court codified, justified, legalized, and normalized White supremacy as well as anti-Black racism in the *Dred Scott v Sandford* (1856) case. Mr. Scott argued that he was a citizen of the US, and therefore the rights and privileges of citizenship adhere.

Chief Justice Taney provided a (White) commonsensical understanding to an interpretation of the [Declaration of Independence \(1776\)](#) and the US Constitution, by first posing a question, then supplying a response:

Can a negro, whose ancestors were imported into this country, and sold as slaves, become a member of the political community formed and brought into existence by the Constitution of the United States, and as such become entitled to all the rights, and privileges, and immunities, guaranteed by that instrument to the citizen?

Moreover, the Court's majority opined deeply-held beliefs about the superiority of White people and the inhumanity of Black people:

The words "people of the United States and "citizens" are synonymous terms, and mean the same thing

... they [Black people] were at that time considered as a subordinate and inferior class of beings, who had been subjugated by the dominant race, and, whether emancipated or not, yet remained subject to their authority, and had no rights or privileges but such as those who held the power and the Government might choose to grant them

They had for more than a century before been regarded as beings of an inferior order, and altogether unfit to associate with the white race, either in social or political relations; and so far inferior, that they had no rights which the white man was bound to respect; and that the negro might justly and lawfully be reduced to slavery This opinion was at that time fixed and universal in the civilized portion of the white race. It was regarded as an axiom in morals as well as in politics, which no one thought of disputing, or supposed to be open to dispute; and men in every grade and position in society daily and habitually acted upon it in their private pursuits, as well as in matters of public concern, without doubting for a moment the correctness of this opinion

... in the eyes and thoughts of the men who framed the Declaration of Independence and established the State Constitutions and Governments ... that a perpetual and impassable barrier was intended to be erected between the white race and the one which they had reduced to slavery, and governed as subjects with absolute and despotic power, and which they then looked upon as so far below them in the scale of created beings,

They perfectly understood the meaning of the language they used, and how it would be understood by others; and they knew that it would not in any part of the civilized world be supposed to embrace the negro race, which, by common consent, had been excluded from civilized Governments and the family of nations, and doomed to slavery.

Given Taney's explanation of commonly understood discourse among White people in his opinion, permitted The *Dred Scott* decision to codify and justify the ideology of White supremacy as a widely accepted belief. Importantly, it also attempted to legalize and federalize White supremacy using the Supreme Court.

Legally, the overt White supremacy expressed in the *Dred Scott* law was overturned after the Civil War (1861–1865), by a series of laws: Civil Rights Law (1866), and the Civil War Amendments ([Const. amend. XIII](#), abolished slavery; [Const. amend. XIV](#), defined equal rights, due process, & citizenship; and [Const. amend. XV](#) extended voting rights to Black men). Among the changes were greater access to literacy, confirming Black people's belief that access was a key component to freedom and liberation. During Reconstruction, several agencies worked at the federal and state level to extend literacy access: the American Missionary Association (AMA) and the Bureau of Refugees, Freedmen, and Abandoned Lands (Freedmen's Bureau). The latter was established by Congress under the Freedmen's Bureau Acts (1865–1866), and initially designed to address the needs of newly freed Black people. The agencies shared a focus on literacy, "to meet their [Black people's] economic, educational, political, and social obligations and construct a barrier to further exploitation" ([Cabral, 2006](#), p. 13). Wishing to please Southerners, Congress at the insistence of President Andrew Jackson, included the needs of White people in these acts. Their decision began a pattern within the federal government to only make a provision for Black people if White people also were (are) accommodated.

There also were immeasurable numbers of Black and White teachers and volunteers who sought to improve literacy access following the Civil War. Among those were Black women, some of whom had normal school and college teacher training, and others who were less well-educated but dedicated to providing literacy access to Black people and racial uplift. The notion of racial uplift encompassed the "value and significance of literacy education as a necessary first step toward full freedom ... learning to read and write as well as other aspects of self-improvement and self-sufficiency were eventually categorized under the heading of uplift" ([Cabral, 2006](#), p. xi). BIPOC understood literacy was not limited to the ability to read and write or communicate, they had read the world around them to survive; literacy was a matter of individual and collective freedom and liberation. Acquiring literacy would lead to increased knowledge and Black people believed that literacy was "the most significant weapon to ward off deliberate, calculated, and dehumanizing confrontations with the White power structure and members of the local White community who attempted to take advantage of the Blacks' lack of basic literacy skills" ([Cabral, 2006](#), p. 14). Administrators from the AMA and Freedmen's Bureau, as well as teachers, periodically published field reports that stressed Christian faith, agency progress, students' progress, and the importance of literacy.

Many of the Black women teachers were often activists and aware of the horrors of slavery as well as the destitute conditions under which formerly enslaved Black people lived. Black women activist teachers understood the importance that Black parents placed on having their children acquire literacy. Charlotte “Lottie” Bridges Forten Grimké (1837–1914) exemplified this group of Black women teachers. She was born into a free, financially secure, and respected abolitionist Black family in Philadelphia, Pennsylvania. Moreover, she was educated by private tutors before attending Higginson Grammar School. She was the only Black student at Salem Normal School, and later taught at Higginson. Charlotte was committed to racial uplift and advancing education for Black people and, as a well-published author, she often wrote about racial injustice. While teaching on St. Helena Island, South Carolina, she published a two-part article about her experiences in *The Atlantic Monthly*.

After the lessons, we used to talk freely to the children, often giving them slight sketches of some of the great and good men. ... Miss T. told them the story of the brave old man who had died for them. I told them about Toussaint, thinking it well they should know what one of their own color had done for his race. They listened attentively, and seemed to understand (Forten, 1864, n.p.).

Grimké emphasized the eagerness of Black people in the community to acquire an education. A similar sentiment was expressed by Alvord (1980), a Freedmen Bureau supervisor

Throughout the entire South an effort is being made by the colored people to educate themselves. In the absence of other teaching they are determined to be self-taught; and everywhere some elementary text-book, or fragment of it, may be seen in the hands of negroes. They quickly communicate to each other what they already know, so that with very little learning many make take to teaching.

Quoted in Brosnan (p. 235).

Johnson (1991) also mentions that during Reconstruction, Black people often preferred schools established by other Black people as they were “distrustful of whites, while whites were disdainful of black teachers and dismayed at the popularity of black-controlled schools among black families who preferred them to the white-managed counterparts, even when the latter were free and the former charged tuition” (p. 90). Legally, the ratification of the Civil War Amendments (13, 14, 15, 16) extended to all citizens the rights and privileges guaranteed in the Constitution (sans voting rights for women). The right to equal education was an important opportunity as Black people understood: “literacy was fundamental to their survival in post-war society, southern black men and women energetically campaigned for a system of education” (Brosnan, 2019, p. 238). Black people, with few resources and often with the support of sympathetic White people, managed to create Black academies, colleges, industrial training schools, normal schools, and universities.

Civil War Amendments and equal rights

The 13th Amendment (1865) included the “exception clause” that stated: “Neither slavery nor involuntary servitude, except as a punishment for crime whereof the party shall have been duly convicted, shall exist within the United States, or any place subject to their jurisdiction” (n.p.). Given that many locales did not permit Black people to participate in courts of law, i.e., *Ohio History Central* (1807) did not permit Black people to testify against a White person, the clause was used to unjustly incarcerate Black people who became a free source of labor for well-connected White people. Despite the passage of Civil Rights Laws many White Southerners also continued to embrace White supremacy and ignored the laws. They created the notion of “states rights,” also known as “a states right to continue slavery,” to circumvent the law. For example, they established Black Codes and Jim/Jane Crow laws to reinstate aspects of slavery: economic dependency, disenfranchisement, mass incarceration, racial discrimination, and segregation. Although the 15th Amendment (1870) prohibited the federal government and each state from denying a citizen the right to vote based on that citizen’s “race, color, or previous condition of servitude,” disenfranchisement was rampant.

Asian Americans also endured unjust treatment at the hands of White settler-colonialists. Asian Americans and American Pacific Islander (AAPI) were the focus of several laws, for example the *National Archives* (2021) noted: “the Chinese Exclusion Act of 1882, ...was the first significant law restricting immigration to be passed by the US Congress (cited in Kim and Hsieh, LRA Statement against Anti-Asian Violence 2021). Through years of anti-Chinese discrimination, this law increased tensions for new immigrants” (Kim and Hsieh, 2021, para 3). Anti-Asian racism erupted in 1886, as White vigilantes sought to deport hundreds of people of Chinese descent. Some people were killed, others deported, and still others sought legal redress. Although there was a trial, no one was held accountable. Anti-Asian violence continued and was supported by federal law in “the Immigration Act (1917),” a law that barred all immigrants from the “Asiatic zone.” The act also is known as the “Literacy Act” – as it required immigrants to pass an English language literacy test for entrance into the US (Kim and Hsieh, 2021).

During this period, the federal government established and withdrew treaties supported by the Dawes (1888) and Curtis (1898) Acts as Indigenous nations were removed from their land. Missionaries from various religious denominations sought to convert the Indigenous people to Christianity and to teach them English in missionary and boarding schools. Indigenous children were forcibly removed from their families, and communities, prohibited from using cultural ways of knowing and home languages, and forced to assimilate to White American ways (Biava, 1990). In boarding schools an “... ‘appropriate’ education was... that which eradicated Indianness or promoted Anglo values and ways of communicating” (Brayboy, 2005, p. 437). Pratt, who created an Indian boarding school, stated, “the only good Indian is a dead one ... I agree with the sentiment, but only in this: that all the Indian there is in the race should be dead. Kill the Indian in him, and save the man” (quoted in Browne-Marshall, 2013, p. 47). Indigenous children were

required to disassociate themselves from their language and heritage and learn the attitudes, beliefs, behaviors, languages, and values of White settler-colonialists.

The federal government was slow to grant citizenship status to Indigenous people and people of Mexican descent. After signing The Treaty of Guadalupe Hidalgo (1848) that ceded present-day Arizona, New Mexico, and Texas, as well as parts of California, Colorado, Nevada and Utah, the US government failed to uphold agreements (Article VIII and Article IX) whereby landowning people of Mexican descent retained their land. When the Mexican economy became depressed there was a massive migration (1872–1910), as Mexican people migrated to the United States. Gallegos (1992) mentions that Mexican people possessed a literate society before they were forced to learn English literacy in the US, as Mexican communities established schools and taught literacy skills in Spanish.

Another landmark Supreme Court case, *Plessy v Ferguson* (1896), occurred less than 50 years after Dred Scott. The case began with the *Louisiana Separate Car Act* (1890) which required Black people to sit in separate cars from White people. Homer Plessy believed he was legally permitted to sit in the White section because he claimed that he was 7/8s White. In a landmark Supreme Court decision, the Court ruled that racial segregation was constitutionally legal, under the “separate but equal” doctrine. Speaking for the majority, Justice Brown argued, “a statute which implies merely a legal distinction between the white and colored races has no tendency to destroy the legal equality of the two races, or reestablish a state of involuntary servitude” (p. 543). He added that the 14th Amendment was not written “to enforce the absolute equality of the two races before the law,” because the amendment “could not have been intended to abolish distinction based upon color, or to enforce social, as distinguished from political, equality, or a commingling of the two races upon terms unsatisfactory to either” (p. 544). Drawing on the *Robert’s* case, the Court’s majority ruled:

[The] establishment of separate schools for white and colored children... has been [deemed] a valid exercise of the legislative power even by courts of States where the political rights of the colored race have been longest and most earnestly enforced (p. 544).

Justice John Harland lodged the lone dissent:

The white race deems itself to be the dominant race in this country. And so it [is], in prestige, in achievements, in education, in wealth, and in power. So, I doubt not, it will continue to be for all time, if it remains true to its great heritage, and holds fast to the principles of constitutional liberty. But in view of the constitution, in the eye of the law, there is in this country no superior, dominant, ruling class of citizens (p. 559).

The case launched a political, judicial, societal, and racial juggernaut: a federal law that codified and legalized racial inequality to exploit BIPOC and privilege White people, or as Prendergast (2002) remarked, the decision to legalize segregation highlighted “the country’s most visible symbol of White supremacy” (p. 210). The doctrine of “separate but equal,” and the Jim/Jane Crow laws legalized under it, were contrary the US Constitution’s declarations of equal protection, due process, and the rights of citizenship under the Constitution.

Systemic racism in politics and education

The 19th century Supreme Court rulings and racially exclusionary laws extended into the 20th century when people of Asian and Pacific Islander descent entered the US, through Hawaii, and Angel Island, California. Their children, like BIPOC before them, were required to attend segregated schools and learn ‘American’ attitudes, beliefs, culture, language, and values. In this manner, literacy access was disproportionately based on economics, ethnicity, geography, language, and race. Irrespective of a person’s access, literacy was assessed and measured as if all people had the same access.

White supremacy reconstituted and extended itself in educational testing by embracing eugenics, or the inheritability of traits, herein intelligence. Galton (1883) defined *eugenics* as, “the study of the agencies under social control that seek to improve or impair the racial qualities of future generations either physically or mentally” (n.p). His belief that Nordic European people represented the apex of intellect, and African people the base, gave rise to scientific racism, the notion that race is biological. Scientific racism is “a deliberate attempt to justify and protect systems that allows exploitation ... to reap economic and political rewards” (Thompson, quoted in Paludi and Haley, 2014, p. 1697). Educational experts used scientific racism to conduct early literacy studies based on the culture and language of ‘native born’ White Americans to the exclusion of all other ethnic/linguistic/racial groups, then, assessed all students by Eurocentric standards. The application of scientific racism in educational testing demystifies how “an ever-changing European-American Protestant standard [that] is the unachievable goal placed on people of color and poor Whites ... placing people of color beneath Whites in all aspects of life” (Browne-Marshall, 2013, p. 5). Thus, allowing educational researchers to conclude that the performance of BIPOC on standardized achievement and intelligence tests was proof of White superior morals and intelligence. Educational psychologists and members of eugenics organizations (Terman, Thorndike, Yerkes) also worked in concert with the federal government during World War I, to develop tests for Army recruits largely, but not exclusively, based on “verbal skills” or literacy. Not surprisingly, BIPOC—who for over two centuries had been denied access to literacy—did not perform

as well as their White peers. The poor performance of BIPOC was celebrated as scientific “proof” of the intellectual superiority of White people. These early studies concealed White supremacy within literacy theorizing and testing by claiming to be culturally-neutral, objective, and scientific measures.

Codifying racism

The judicial ruling of *Plessy v. Ferguson*, “separate but equal education,” resurfaced in three Supreme Court cases. First, *Gong Lum v. Rice* (1927) ruled a Chinese American girl was not White, and thus could not attend a White school. Similarly, students of Mexican descent were not part of the California Education Code that required segregated schools for students of Black, Asian (Chinese, Japanese, or Mongolian), and Indigenous descent. Second, students of Mexican-descent won the right to an equal education in *Mendez v. Westminster* (1947). The U. S. Supreme Court stated that Mexican Americans and “all other groups” have equal protection under the 14th Amendment. Third, the landmark case of *Brown v. Board of Education of Topeka, Kansas* (1954, 1955). The Brown case was part of a class-action lawsuit brought by aggrieved parents on behalf of their Black children who were forced to attend segregated schools. Mr. and Mrs. Brown believed their daughter, Linda a first-grader, should not walk past a White school, and through a series of railroad crossings to attend a Black school. The psychological trauma inflicted on Black children who attended segregated schools appeared to be a key component of the Brown case (Browne-Marshall, 2013). The justices argued that education:

is a principal instrument and awakening the child to cultural values, preparing him for later professional training, and in helping him adjust normally ... it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education. Such an opportunity, where the state has undertaken to provide it, is it right which must be made available for all on equal terms.

Brown (1954, p. 493).

The US Supreme Court unanimously passed Brown, ruling a violation of the equal protection clause of the 14th Amendment, thus dismantling the doctrine of “separate but equal.” Chief Justice Warren declared, “In the field of public education, the doctrine of ‘separate but equal’ has no place. Separate facilities are inherently unequal” (*Brown v. Board of Education I*, 1954, p. 495). *Brown II* (1955) maintained that the Court’s notion of “all deliberate speed” – was permitted compliance on terms set by local school boards.

While many Black people sought integration of public schools, some leaders like Martin Luther King, Jr., did not. He argued, “white people view black people as inferior ... People with such a low view of the black race cannot be given free rein and put in charge of the intellectual care and development of our boys and girls” (quoted in Kendi, 2019, p. 173). At the state and local level, school boards did not rush to comply with the law, trying various ways to subvert the federal order by removing White students from public schools, or moving to all-White suburbs to avoid integration. The implementation of *Brown* hinged not only on federal law but on the willingness of state and local officials to enforce the law. Drawing on White supremacy and admonitions of “states” rights, the process to end racial discrimination in public education was undermined.

President Lyndon B. Johnson (1963–1969) reign

A decade later, the 14th Amendment and the *Civil Rights Act* (1964) effectively struck down Jim/Jane Crow laws and legally ensured human and civil rights were extended to all US citizens. The Civil Rights Act of 1964 (Sec. 601; 78 Stat. 252; 42 U.S.C. 2000d) was part of a multipronged federal socioeconomic plan implemented by President Lyndon B. Johnson. His strategy sought to link political and social reform efforts to civil rights legislation and enforcement, as well as to educational reform and funding. Under this plan, the 1965 *Elementary and Secondary Education Act* (ESEA) was designed to work in concert with the Civil Rights Act to address racial and social inequality: “Congress would not have taken this step [passing ESEA] had Title VI [of the Civil Rights Act] not established the principle that schools receiving federal assistance must meet uniform national standards for desegregation” (*US v. Jefferson Co. Bd. Of Ed.*, 372, E. 836, fifth Cir. 1966, as cited in Frankenberg and Taylor, 2015). ESEA provided financial support to address economic, racial, and social discrimination as well as political avenues for educational access and opportunity, specifically for BIPOC students, students living in poverty, and students with disabilities (ESEA, Sec. 201). In addition, under Johnson’s plan voting rights—that had been restrained, suppressed, and violated in many states—needed to be extended to all people. Voter suppression took on many forms and one tactic used to disenfranchise Black people were literacy tests. Although the period of Reconstruction had improved access to literacy for Black people; Jim/Jane Crow laws were established to retard and slow down equal access to education. Through the efforts of individuals and institutions, some Black people were taught to read and write. Septima Poinsette Clark (1898–1987), for example, was a civil rights advocate, member of the NAACP, and teacher who worked with Myles Horton of Highlander Folk School. They created Citizenship Schools to teach Black adults literacy and to register to vote in South Carolina. Bernice Robinson, Septima’s cousin, secretly taught Black people to pass the “literacy test.” They believed that teaching literacy

should not try to isolate people’s learning of reading, writing, and critical thinking from the community in which they reside; the historical roots from which they follow; the language they use in defining their existence; and the politics and economics that shape and reshape their critical thinking/reflections of themselves and others.

Peavy (1993, p. 213).

The Citizenship Schools were successful in teaching literacy, but the literacy test required a person to read and explain any portion of the South Carolina state constitution. Local deputy registers could ask any question and the veracity of responses was decided at the register's discretion. The responses of Black people were deemed unacceptable.

In the summer of 1964, Freedom Schools also were organized by a variety of Black community, political, and student groups to improve the quality of education received by Black students in the rural South. A network of over 40 schools, for example, served two thousand Black students in Mississippi from July–August 1964, where improving the literacy skills of students was a major component. Students learned basic grammar, improved reading comprehension, and used writing for expression and protest. The Black and White Freedom School teachers were a mix of classroom teachers, college students, and volunteers... [who] sought to build students' self-esteem, linked cultural understandings to the curriculum, and encouraged students to use their voice to protest racial injustice. Citizenship and Freedom Schools were grassroots efforts as well as political extensions to provide and ensure Black people had access to literacy.

Systemic racism in education

The passage of the 1964 Civil Rights Act, 1965 Voting Rights Act, and the ESEA, helped to usher in the politicization of literacy in education. Congress charged experts to find the best way to teach beginning reading in the First-Grade Reading Studies (Bond and Dykstra, 1967). They requested a review of studies that could address cultural, economic, linguistic, and racial differences among students. The experts solicited studies that applied experimental research methods in the study of beginning reading programs; however, studies that focused on students who were Black, lived in urban or low-income areas, and students who were Spanish dominant, were discarded (Chall, 1967). The final report was positioned as apolitical, color-blind, and objective, although it applied and extended scientific racism through experimental testing to early reading. The study established a pattern that: focused on the learning needs of White students; denigrated BIPOC communities, families, and students; and used federal funding for research to validate White intellectual superiority. It also idealized US students as English dominant, middle-to-upper class, with above-normal abilities, and White (Willis, 2008). An equal education for students of Asian descent was the grounds for the *Lau v Nichols* (1972) case as the rights of children of Chinese ancestry were violated, as outlined in the 1964 Civil Rights Act, when English was used as a language of instruction. The sole use of English for instruction denied students who were not proficient in English an equal opportunity for an education.

In the *Plyler v Doe* (1982) case, a class-action suit on behalf of undocumented children of Mexican descent, the US Supreme Court ruled in favor of the plaintiff's case. They cited the equal protection clause of the 14th Amendment, and acknowledge a person's psychological well-being, and highlighted the importance of literacy. Justice Brennan presented the Court's opinion

illiteracy is an enduring disability. The inability to read and write will handicap the individual deprived of a basic education each and every day of their lives. The inestimable toll of that deprivation on the social, economic, intellectual, and psychological well-being of the individual, and the obstacle it poses to individual achievement, make it most difficult to reconcile the cost or the principle of a status-based denial of basic education with the framework of equality embodied in the Equal Protection Clause.

In addition to new laws, a series of Congressional education reports, *A nation at risk* (1983), *Goals 2000*, and *America 2000*, helped to chart the nation's future educational goals.

Several reports specific to literacy requested by Congress, also were published: *Becoming a Nation of Readers: The Report of the Commission on Reading* (Anderson et al., 1985), *Beginning Reading* (Adams, 1990), *Preventing Reading Difficulties* (Snow et al., 1998), and the *National Reading Panel Report* (2000). The politicization of education spawned new federal agencies, the National Center for Educational Statistics, and the National Assessment of Education Progress (NAEP), among others. With Congressional support, cyclically, NAEP assessed school children's reading (and writing) and produced the *Nation's Reading Report Card*. The results are disaggregated by gender, race, and social class, to reveal reading results for Black, Latinx, and Indigenous students are below Asian, Asian Pacific Islander, and White students. Educational researchers and political stakeholders consistently pointed to the trustworthiness and veracity of the experimental/scientific methods, and ignored how the processes support/ed White supremacy.

As a candidate for US President, Texas governor, George W. Bush, drew on the Civil Rights era discourse in his bid for the presidency. He declared "education is the greatest civil rights issue of our time," and that the nation faced a "reading crisis." He championed improving the reading outcomes for BIPOC and low-income students by highlighting the results of their reading performance as described in the *National Reading Panel Report* (2000) and NAEP data. Under President George W. Bush, Congress reauthorized ESEA as the No Child Left Behind Act (2001), along with Title I (*Improving the Academic Performance of the Disadvantaged, Part A, Basic Program*), Title III (*Language Instruction for Limited English Proficient and Immigrant Students*), Title VII (*Indian, Native Hawaiian, and Alaska Native Education*), and *Individuals with Disabilities Education Act* (IDEA), and later IDEIA in 2004. NCLB required states to assure adequate yearly progress was made by economically disadvantaged students, students from major racial and ethnic groups, students with disabilities, and students with limited English proficiency. Reading was recognized as pivotal to the effectiveness of NCLB and the national focus on reading resulted in significant increases in federal research funding. Underpinning the push for increased testing and support of scientific/evidenced-based results is the ideology of White supremacy, recast as a reasonable response to the literacy crisis. The failure of NCLB and subsequent ESEA authorizations, *Race to the Top* (2009) and *Every Child*

Succeeds Act (2015), to improve the literacy educational performance of BIPOC students indicates federal and legal remedies for racial injustice are undermined by systemic institutional racism.

State court cases

In the *Ella T. v State of California* case students in three different elementary schools won literacy for all California students based on the 14th Amendment. The plaintiffs' lawyers argued that California was doing a poor job of teaching students to read, and the performance of economically poor students on literacy assessments was a violation of the CA constitutional mandate, to provide equal access to education. An award of \$53 million was allocated to provide resources to the state's 75 lowest-performing elementary schools. In a similar case, a suit was brought on behalf of Black students in Detroit, MI in *Gary B. v. Snyder* (2016), also called the "right to read" lawsuit, followed by the *Gary B. v. Whitmer* case. Citing several alleged violations of the 14th Amendment as well as racial discrimination in violation of Title VI of the Civil Rights Act of 1964, 42, U.S.C., and 34 C.F.R. 1000.3(b)(2), the plaintiffs' lawyers compared the education received by the student Black plaintiffs to the education of their White students or peers. Their central argument was that Black students were not receiving an equal and equitable education and emphasized the reading instruction was inadequate and unable to propel them to a meaningful life or future. To be clear, these cases hinged, in part, on whether reading is a fundamental and constitutional right. After a series of appeals, the Sixth Circuit US Court of Appeals ruled that "students have a right to 'a basic minimum education,' which the court defined as giving students the opportunity to learn to read" (Peak and Hanford, 2020, n.p.). The ruling granted \$94.4 million to the Detroit Public Schools to improve reading programs, however, reading was not declared a Constitutional right.

In the nearly 70 years since *Brown*, BIPOC continue to legally demand the right to equal education, although what is considered an equal education remains undefined. In recent state-legal victories, the desire for an equal education has centered on access to literacy, especially appropriate reading instruction. Appropriate reading instruction is needed for all students to perform well on standardized reading tests required under reauthorizations of ESEA. However, a review of state laws and policies found that there are 30 states with reading retention laws, requiring students to read on grade level by the end of third grade or be retained. In Michigan, for example, during the 2015–2016 legislative session, MCL 380.1280f was passed and requires students to be able to read and write on grade level by the end of third grade. The devil is in the details, as there are accompanying exceptions that waive the requirement for students identified as dyslexic (the majority of whom are White and middle-to-high income), but not for students identified as SLD (the majority of whom are Black, Indigenous, and Latinx and low-to middle income). These laws ignore that historically the BIPOC students attend schools that have been underfunded and under-resourced (Browne-Marshall, 2013; Frankenberg et al., 2019; Gabriel, 2020), and are most often retained. Supporters of the legislation argue that the failure to read on grade level by the end of third grade is a predictor of later reading success and high school completion. DeLissovoy (2016) believes as positioned, their argument perpetuates "White ingenuity and generosity, and at the same time constructing interrogations of this narrative as unreasonable and inadmissible" (p. 353). There are numerous studies that do not support the finding of failure to read on grade level by the end of third grade increases the high school dropout rate, and acknowledge a disproportionate percentage of BIPOC students are retained under these state laws. Opponents submit the new laws reconstitute White supremacy—under the guise of "states' rights" to make educational decisions and the use of standardized state literacy tests as a cover for racist literacy policies. Historically, literacy has been weaponized to "remove people of color from economic, political, and social competition" (Browne-Marshall, 2013, p. xxxvi), and the most recent laws, as well as selective waivers, continue the process.

Beginning in 2020, a spate of new state laws was established, allegedly in response to the teaching of critical race theory, although it is not a theory taught in schools. The reasons given for new state laws (labeled as decency or parents' bill of rights'), ranged from barring the discussion and teaching of US history as systemically racist, to concerns about hurting the feelings of White students. In addition, under cloaks of "decency" some school districts have begun a massive overhaul of literature, primarily banning books that discuss issues of anti-Semitism; ethnicity/race/racism; and LGBTQAI+ issues. By contrast, the American Psychological Association (2021) admitted to racist actions made by, and, on behalf of the organization in the past. They acknowledged their support of scientific racism and White supremacy, provided a detailed chronology of past racist actions, recounted the failure to acknowledge and publish scholarship by Black researchers, and apologized for in/actions taken in their self-interest that harmed people of color. Moreover, APA resolved to improve, and presented a series of action steps to be implemented. Despite the stunning admission, federal and state laws, as well as local school board policies have been indefensibly silent about the announcement and have not announced any change in laws or policies to create more equitable standardized reading tests.

Looking forward

This history of literacy access in the US critiques the ideological roots of White supremacy and the false narrative that frames the founding and the democratic stance of the nation: the importance of European culture, language, literacy, and values; and the unfettered mischaracterizations about BIPOC as less-than-human, unintelligent, and without culture, language, literacy, and value. Glaude (2016) argues these ideas are reflected in US democracy as a fundamental belief that White people are more valued than all other people. His idea is noteworthy given the interpretation of the Declaration of Independence and the US Constitution, as well as US Supreme Court rulings that opine White people are the superior race and White people should dominate society. The discourse of the documents and US Supreme court cases appears to confirm the belief about the superiority of White people

over all other races. Harris (1993) theorizes that the sustainability of these ideas is premised on believing that, “whiteness meets the functional criteria of property” (p. 1731). From her perspective, Whiteness has value and the reality of Whiteness is tangible, “the law has accorded holders of whiteness the same privileges and benefits according to holders of other types of property ... it includes exclusive rights of possession and use and disposition” (p. 1731). Wilkerson (2020) believes America’s problems are deeper and prefers the notion of caste, “it’s much harder to look into the darkness in the hearts of ordinary people with unquiet minds, needing someone to feel better than, ...the toxins can infect succeeding generations” (p. 267). Beyond beliefs, customs, laws, and statutes the US legal system consistently has failed to reckon with the history of economic, social, and racial injustices while seeking to codify and normalize White supremacy.

As this examination of the nation’s literacy history and review of several federal and state laws reveals, there is a repeated pattern of compromising literacy equity for BIPOC, while simultaneously promoting a façade that new approaches, assessments, curricula, instruction, and laws, will ensure equitable literacy and racial progress. Attempts to federalize racial equality suggest that racial equality under the law is an insufficient response, as innumerable legal and political maneuvers successfully have dismantled, ignored, and sidestepped laws that provide racial reconciliation and the opportunity for equal education and literacy access. For nearly four centuries, literacy access for BIPOC has been criminalized, denied, derailed, and restricted. The counter-narratives in this chapter describe how (a) some people from Africa, Asia, and Central America (Mexico) were literate when they arrived in North America; (b) Sequoyia created a Cherokee syllabary and proved that literacy is not defined by, nor synonymous with, Whiteness; (c) Black people defied dehumanizing characterizations and laws to acquire, and, teach literacy; (d) literacy has been weaponized to subvert equal access to education; (e) literacy is a pathway to freedom and liberation and not a broad set of skills; and (f) BIPOC extended literacy to all people irrespective of economic status, ethnicities, and genders. In the twentieth century, White supremacy evolved within education following the Civil Rights Act and the establishment of ESEA. Congressional requests to better improve the literacy performance of BIPOC, for instance, resulted in experts drawing on scientific racism and experimental methods to socially control concepts, criteria, definitions, instructional approaches, materials, and measures of literacy. Every legal effort has failed to make a significant improvement in BIPOC literacy performance as measured by standardized achievement tests. Recent state court victories in California and Michigan confirm the Constitutional right to equal access to literacy. As a nation, however, we bear witness to a history of legal processes that have obfuscated the potential for authentic literacy access and educational equality. As a nation, it is important to acknowledge that rich literacy histories among BIPOC existed long before European settler colonialists invaded North America, and to understand that White supremacy informs and frames the false narrative used to deny literacy access to BIPOC. There is a moral imperative to create a novel narrative that: ensures equal education; defines literacy as a human right; and values the dignity of all humans as well as their cultures, ethnicities languages, races, and knowledge systems. Looking forward, literacy is paramount to anchoring US democracy, equality, equity, freedom, and justice.

References

- Act for the Protection of the People of Indian Territory, 1898. Curtis Act. <https://nativeheritageproject.com/2013/08/17/the-curtis-act-of-1898/>.
- Adams, M.J., 1990. *Beginning to Read: Thinking and Learning About Print*. MIT Press.
- Alvord, J.W., 1980. *First Semi-annual Report on Schools and Finances of Freedmen*. January 1. AMS Press (original published in 1866).
- An Act to Provide for the Allotment of Lands in Severalty to Indians on the Various Reservations, Dawes Act, <https://www.ourdocuments.gov/doc.php?flash=true&doc=50>.
- Anderson, R.C., Hiebert, E., Scott, J., Wilkinson, I., 1985. *Becoming a Nation of Readers: The Report of the Commission on Reading*. Center for the Study of Reading.
- Anderson, J., 1988. *The Education of Blacks in the South, 1860–1935*. University of North Carolina Press.
- Ansley, F.L., 1997. White supremacy (and what we should do about it). In: Delgado, R., Stefancic, J. (Eds.), *Critical White Studies: Looking behind the Mirror*. Temple University Press, pp. 592–595.
- Biava, C., 1990. Native American Languages and Literacy: issues of orthography choice and bilingual education. *Kans. Work. Pap. Ling.* 15 (2), 45–59.
- Bond, G., Dykstra, R., 1967. The cooperative research program in first-grade reading instruction. *Read. Res. Q.* 2, 5–142.
- Brayboy, B.M.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37 (5), 425–446.
- Brosnan, A., 2019. “To educate themselves”: southern Black teachers in North Carolina’s schools for the freedpeople during the Civil War and reconstruction period, 1862–1875. *Am. Nineteen Century Hist.* 20 (3), 213–248.
- Brown v. Board of Education of Topeka, 347 U.S. 483, 1954.
- Brown v. Board of Education II, 349 U.S. 294, 1955.
- Browne-Marshall, G.J., 2013. *Race, Law, and American Society 1607–Present*, second ed. Routledge.
- Cabral, L.B., 2006. *Dissertation: Letters From Four Antebellum Black Women Educators to the American Missionary Association, 1863–1870* (Blanche Virginia Harris, Clara C. Duncan, Sallie Louise Daffin, Edmonia Goodelle Highgate. University of Massachusetts, Boston).
- Castillo, E., 2015. *A Cross of Thorns: The Enslavement of California’s Indians by the Spanish Missions*. Craven Street Books.
- Chall, J., 1967. *Learning to Read: The Great Debate*. McGraw-Hill.
- Child, B.J., Klopotek, B. (Eds.), 2014. *Indian Subjects: Hemispheric Perspectives on the History of Indigenous Education*. School for Advanced Research Press.
- Chinese Exclusion Act, 1882.
- Civil Rights Act of 1866, Pub. L. 14, Stat. 27–20, 1870.
- Civil Rights Act of 1964 § 7, 42 U.S.C. § 2000e et seq, 1964.
- Costello, M., 2014. *Encyclopedia of Critical Psychology*. https://doi-org.proxy2.library.illinois.edu/10.1007/978-1-4614-5583-7_470.
- Crenshaw, K.W., 2020. Twitter. <https://twitter.com/sandylocks/status/1276571389911154688>.
- DeLissovoy, N., 2016. Race, reason, and reasonableness: toward an “unreasonable” pedagogy. *Educ. Stud.* 52, 346–362.
- Dred Scott v. Sandford, 60, U. S. 394. (1856).
- Elementary and Secondary Education Act, Public Law 89-10, 1965. <https://www2.ed.gov/about/offices/list/oi/nonpublic/esearchauth>.
- Ella T. v State of California.
- Every Child Succeeds Act, <https://www.ed.gov/essa?src=m>.

- Dunbar-Ortiz, R., 2014. *An Indigenous Peoples' History of the United States*. Beacon.
- Forten, C., 1864. Life on the sea islands. *Atl. Mon.* https://www.theatlantic.com/magazine/archive/1864/05/life-on-the-sea-islands/308758/?gclid=CjwKAjw9vn4BRBaEi-wAhOmudCwol-Sxb77NcgDrndQGvxoVAtiql5EJl0-XXnB6AyyPfBi_cee1BoChYgQAvD_BwE.
- Foucault, M., 1988. The subject and power. *Crit. Enq.* 8, 777–795.
- Frankenberg, E., Taylor, K., 2015. ESEA and the Civil Rights Act: an interbranch approach to desegregation. *Russell Sage Foundation J. Soc. Sci.* 1, 32–49.
- Frankenberg, E., Ed. J., Ayscue, J.B., Orfield, G., 2019. Harming our common future: America's segregated schools 65 years after Brown. Center for Civil Rights. <https://www.civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/harming-our-common-future-americas-segregated-schools-65-years-after-brown>.
- Gabriel, R.E., 2020. Converting privatization: a discourse analysis of dyslexia policy narratives. *Am. Educ. Res. J.* 57 (1), 305–338. <https://doi.org/10.3102/0002831219861945>.
- Gallegos, B., 1992. *Literacy, Education, and Society in New Mexico, 1693–1821*. University of New Mexico Press.
- Galton, F., 1883. *Inquiries Into Human Faculty and Its Development*. Macmillan.
- Glaude, E.S., 2016. *Democracy in Black: How Race Still Enslaves the American Soul*. Broadway Books, New York, NY.
- Gong L.v. Rice, 275 U.S. 78, 1927.
- Gore, J., Morrison, K., 2001. The perpetuation of a (semi-) profession: challenges in the governance of teacher education. *Teach. Teach. Educ.* 17 (5), 567–582. [https://doi.org/10.1016/S0742-051X\(01\)00014-2](https://doi.org/10.1016/S0742-051X(01)00014-2).
- Harris, C., 1993. Whiteness as property. *Harv. Law Rev.* 106, 1707–1791.
- Massachusetts Historical Society. Long Road to Justice: Sarah C. Roberts vs. City of Boston, <http://www.longroadtojustice.org/topics/education/sarah-roberts.php>.
- Individuals with Disabilities Education. Improvement Act of 2004, Pub L. No., 08-446, 118.
- Jacobs, H.A., 1861. *Incidents in the Life of a Slave Girl: Written by Herself*. Post.
- Johnson, W.B., 1991. A Black teacher and her school in Reconstruction darien: The correspondence of Hettie Sabattie and J. Murray Hoag, 1868–1869. *The Georgia Historical Quarterly* 75 (1), 90–105.
- Kendi, I.X., 2019. *How to Be an Anti-racist*. One World.
- Kim, J., Hsieh, B., 2021. LRA Statement against Anti-Asian Violence. <https://lira.memberclicks.net/assets/docs/LRA%20Statement%20against%20Anti-Asian%20Violence.pdf>.
- Legislative Papers, 1830–31 Session of the General Assembly, <https://www.ncpedia.org/anchor/bill-prevent-all-persons>.
- Louisiana Separate Car Act, 1890.
- McCarty, T.L., Lee, T.S., 2014. Critical culturally sustaining/revitalizing pedagogy and indigenous education sovereignty. *Harv. Educ. Rev.* 84 (1), 101–124.
- National Archives, 2021. Chinese Exclusion Act Records at the National Archives at Seattle. www.archives.gov/seattle/finding-aids/chinese-exclusion-actLiteracy. Research Association.
- National Reading Panel Report, 2000. *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction* (NIH Publication No. 00-4769). U.S. Government Printing Office.
- No Child Left Behind Act of 2001, Pub. L. Mo. 107-110, 115 Stat. 1425, 2002.
- Ohio History Central, 1807. *Black Laws of 1807*. http://ohiohistorycentral.org/w/Black_Laws_of_1807.
- Paludi, M.A., Haley, S., 2014. *Encyclopedia of Critical Psychology*. Retrieved, July 2020. https://doi-org.proxy2.library.illinois.edu/10.1007/978-1-4614-5583-7_394.
- Parins, J.W., 2013. *Literacy and Intellectual Life in the Cherokee Nation, 1820–1906*. University of Oklahoma Press.
- Peak, C., Hanford, E., 2020. Is Learning to Read a Constitutional Right? Retrieved, August 2020. <https://www.apmreports.org/episode/2020/04/30/is-learning-to-read-a-constitutional-right>.
- Peavy, L., 1993. Promoting liberation literacy: a grassroots solution. In: *Cultural Keepers: Enlightening and Empowering Our Communities: Proceedings of the First National Conference of African American Librarians*. Black Caucus of the American Library Association, pp. 212–215.
- Plessy v. Ferguson, 1963 U. S. 537, 1896.
- Plyler v. Doe, 457 U.S. 202, 1982.
- Prendergast, C., 2002. The economy of literacy: how the Supreme Court stalled the civil rights movement. *Harv. Educ. Rev.* 72 (2), 206–229.
- Roberts v. City of Boston, 5 Cush. 198, 1850.
- Snow, C.E., Burns, M.S., Griffin, P. (Eds.), 1998. *Preventing Reading Difficulties in Young Children*. National Academies Press.
- Spring, J., 1996. *The Cultural Transformation of a Native American Family and Its Tribe 1763–1995: A Basket of Apples*. Lawrence Erlbaum Associates.
- U.S. Const. amend. XIII.
- U.S. Const. amend. XIV.
- U.S. Const. amend. XV.
- U.S. Constitution, 1787. Article I, Section 2. Three-Fifths Compromise.
- U.S. Declaration of Independence, (1776) <https://www.archives.gov/founding-docs/declaration-history>.
- U.S. Department of Education, 2009. Race to the Top program: Executive summary <https://files.eric.ed.gov/fulltext/ED557422.pdf>.
- Wilkerson, I., 2020. *Caste: The Origins of Our Discontents*. Simon Schuster.
- Willis, A.I., 2008. *Reading Comprehension Research and Testing in the US: Undercurrents of Race, Class, and Power in the Struggle for Meaning*. Lawrence Erlbaum.

Learning disabilities' unsettling intersections: disrupting learning, cultural, socioeconomic, and environmental deficits

Adai A. Tefera^a and Alfredo J. Artiles^b, ^a University of Arizona, Tucson, AZ, United States; and ^b Stanford University, Stanford, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

A historical sketch of tensions in the conceptualization of learning disabilities	279
Intersectional framing of SLD: moving beyond a unitary approach	281
Disrupting double bind identities: the promise of intersectionality	281
Dangers of the unitary approach: history, epistemic cultures and ideologies of difference	282
Why culture matters in SLD identification practices	283
Reflections for future SLD research	284
Acknowledgments	285
References	286

This chapter presents a critical review of the category specific learning disabilities (SLD) in the United States (US).¹ Students with SLD make up the largest disability category in the US, consisting of 33% of approximately seven million students with disabilities (NCES, 2020). The use of the adjective “specific” and the plural noun to describe this condition (“disabilities”) speak to the complexities of this category. Indeed, SLD covers several language and psychological domains as reflected in the federal government definition:

a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken, or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. [SLD] does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of intellectual disability, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

U.S. Department of Education (2018).

The conceptualization of SLD has a long history of tensions and ambiguities between biological origins and the role of contextual and cultural influences. In fact, as reflected in the exclusionary clause, the current definition suggests that SLD is orthogonal to cultural, environmental, and socioeconomic factors. These uncertainties trickle down to policies and identification practices that complicate the work of school professionals, which in turn, might mediate fluctuations in SLD prevalence over time. Moreover, the ambivalent connections between biology and context in the SLD definition have likely contributed to a history of racial and linguistic disparities in identification rates.

We argue that the conceptualization of SLD can benefit from an interdisciplinary analysis of the interlocking of learning with racial, cultural, and socioeconomic differences. Rather than keeping them apart, as if learning happens in places and times devoid of sociocultural influences, we advance an intersectional critique challenging the field to consider these complexities in the conceptualization of this condition. For this purpose, we present a brief historical overview of SLD, illuminating the connections to biological deficiencies and making visible the ways sociocultural differences (e.g., race) have historically been imbricated in the evolution of this concept. Next, we present an intersectional analysis showing how the notion of SLD has assumed a universalistic standpoint that has been largely shaped by federal policies that de-contextualize learning differences and focus on a unitary approach that focuses on SLD as the defining feature of this condition. Despite recurrent debates about the intersections of SLD with race, language and social class, the federal definition continues to exclude such intersections. We then demonstrate the importance of embracing a cultural vision of learning in SLD identification practices. We end with brief reflections for future SLD research.

A historical sketch of tensions in the conceptualization of learning disabilities

To begin, we recognize the importance of understanding how SLD emerged as a disability category, the ways it has evolved and developed over time, and how the early roots of SLD have contributed to contemporary challenges. Although SLD as a construct

¹It is important to note that we approach this chapter with a focus on SLD in the United States but recognize that many of the same challenges in the global North related to a focus on biological causes and ambiguous definitions are troubling the global South as well (Grünke and Cavendish, 2016). We also acknowledge the problematic influence of the global North regarding SLD definitions and identification processes that have encroached upon and exacerbated inequities in the global South (Kalyanpur, 2020).

was not adopted in US education policy until 1975, its origins can be traced back to nineteenth century Europe when the medical establishment focused primarily on the biological and neurological deficits of individuals facing challenges related primarily to the spoken language. A related research strand focused on reading disorders assumed to be linked to biological causes (congenital or hereditary factors) (Hallahan et al., 2013). At the time, studies included a focus on soldiers with brain injuries and their effect on language development. Building on this early "brain science," by the 1900s, research expanded to the US where the focus included children placed in facilities due to what was then described as "mental deficiencies." Today these children would most likely be identified with intellectual disabilities. As Danforth (2009) described, in the US, during the early twentieth century there was increasingly a focus on children who were believed to display forms of brain "dysfunction" despite these children not actually having experienced brain injuries. Werner and Strauss (1941), for instance, crafted behavioral profiles of these students by type of etiology, namely brain injury (exogenous) and poor learning backgrounds (endogenous). This was a key moment in the conceptualization of SLD given the assumption that learning differences observed in children were due primarily to either biological factors (brain injuries) or cultural (environmental) deprivation. These early views about problems related to "language, memory, behavior, and learning difficulties initiated lasting ideas and also deep conceptual tensions that continued in multiple forms through the developing field of learning disability research for many years" (Danforth, 2009, p. 29). In this sense, the medical model was pivotal given the focus on language and learning difficulties and critical to its development as a category of disability; but as we explain below, we have also witnessed intermittent attention to a cultural deprivation hypothesis over time.

While the focus in the early 1900s was primarily on observing and explaining learning differences, this was followed by a period of focusing on diagnoses, assessment tools, and interventions to address mainly reading disabilities. Approaching the mid 20th century, Monroe (1932) offered the notion of a discrepancy between actual and expected achievement to conceptualize reading disabilities which has had a lasting impact on the definition of SLD (Hallahan et al., 2013). It was not until 1963, however, that Sam Kirk was credited with coining the term *learning disabilities* in his address to parents of "perceptually handicapped" children. He defined the condition as:

[A] retardation, disorder, or delayed development in one or more of the processes of speech, language, reading, writing, arithmetic, or other school subject resulting from a psychological handicap caused by a possible cerebral dysfunction and/or emotional or behavioral disturbances. It is not the result of mental retardation, sensory deprivation, or cultural and instructional factors

Kirk as cited in Hallahan et al. (2013, p. 22).

It is interesting that Kirk considered "instructional" influences as an exclusionary factor in the early definition, which serves as a reminder of the importance of educational opportunity in shaping learning difficulties. A year later, in 1964, the Learning Disabilities Association of America (LDA) was founded and four years later, the Council for Exceptional Children established the Division for Children with Learning Disabilities.

It is important to note that the early federal definition of SLD stressed the biological underpinnings of this disability, grounded in cognitive and behavioral models. Furthermore, the impetus for the definition's exclusionary clause was to differentiate this population from other groups for the purposes of securing resources. This is not surprising since the recognition of cultural identities (e.g., poor, disabled, racial groups) in the pursuit of justice agendas is linked to the redistribution of resources (Fraser, 2007). Thus, advocates worked to secure the recognition of learners with disabilities to obtain education funding and entitlements. This is tricky work because redistribution decisions are generally attached to particular categories in the chessboard of identity politics. Thus, advocates supporting the inclusion of SLD in federal policy were concerned that other groups struggling with reading—i.e., racially minoritized students and students living in poverty—might be included in this category, thus increasing the number of eligible beneficiaries and boosting the costs associated with the emerging policy (Martin, 2020).

We should note, however, that other cultural-historical forces have been at play in efforts to exclude "environmental, cultural, or economic disadvantage" from the definition of SLD—what Kirk first described as "sensory deprivation or cultural factors." These influences were already present during the early years of brain injury and learning impairment studies and have been intricately interwoven in a longstanding project of "othering" certain groups in US society. For instance, the notion of "backwardness" emerged in the early 1900s to describe children who were of "normal intelligence" but did not meet what was deemed adequate progress in school (Franklin, 1980). It was common to blame cultural and socioeconomic factors as primary explanations for "backwardness." As Franklin (1980) argued, "backwardness" was linked to how researchers later framed learning disabilities in an attempt "to control sources of cultural diversity within the population in the name of cultural homogeneity and the hegemony of the native middle class" (p. 18). In subsequent decades, a distinctive view of success and competent performance has been enforced in educational systems which has consistently marginalized low-income and racialized groups, including recent immigrants (Deschenes et al., 2001).

A related cultural-historical influence is the idea of "culture of poverty" (and its cognate "cultural deprivation") which emerged in the 1960s at the height of racial tensions and the struggle for Civil Rights for people of color in the US. These notions were used to describe mostly minoritized groups, but also all poor people, who were perceived to be "severely handicapped by home environments that lacked environmental stimuli" (Sleeter, 1986, p. 49). This characterization was consistent with the historical trope of

deficit thinking to describe and treat communities of color (McDermott and Vossoughi, 2020). The culture of poverty model assumes that the self-perpetuating mechanisms through which culture impairs children's learning are socialization or deprivation processes. In this view:

[c]hildren either learn skills and dispositions [in their cultural communities] that are not useful for school learning ... [socialization] or the children's culture prevents them from learning certain skills, habits, or values that prepare them for success in school (deprivation). The premise in the cultural deprivation hypothesis is that living under certain conditions (e.g., poverty) exposes children to cultural practices that limit the acquisition of normative bodies of knowledge, dispositions, as well as skills, and limits access to experiences that are valued by the dominant society. In this view, "the culture of poverty" deprives children of sound developmental experiences and becomes destiny ... [T]hese two understandings of the roles of culture in learning are problematic on two counts: they rely on limited definitions of culture and lack theoretical specificity.

Artiles et al. (2010, p. 291).

These stratifying undercurrents have permeated the history of US society and have certainly influenced the development of categories like SLD, including its exclusionary considerations. Unfortunately, the culture of poverty has resurfaced, again (McDermott and Vossoughi, 2020). Meanwhile, the definition of SLD has continued to be revised over the years, though the core features of this category including a discrepancy model (between child potential and achievement) have remained in place, as reflected in the current definition presented in the introduction of this chapter.

The longstanding tension between (assumed) biological causes and exclusionary cultural/socioeconomic considerations has also persisted and has contributed to conceptual ambiguities and dilemmas in identification practices. For instance, the SLD population (of mostly white middle class students) grew substantially in the early years of the federal category. Over time, however, African American learners migrated to this category from the intellectual disability classification, in part due to litigation related to bias in identification (Donovan and Cross, 2002). In fact, the representation of low-income African Americans and Native Americans in the SLD category grew in the subsequent decades after its inception in several geographical regions, despite the definition's exclusionary clause.

This historical context sets the stage for contemporary challenges related to SLD, particularly given the disregard for how racial, socioeconomic, and cultural differences and contextual factors intersect with this disability. In the next section, we critique this logic by offering insights from intersectionality to overcome the shortcomings of a universalistic approach.

Intersectional framing of SLD: moving beyond a unitary approach

We use an intersectional lens to discuss key policy and definitional issues related to SLD that have beleaguered the field. These issues have created conceptual ambiguities and blind spots that have consequential ramifications for diagnosis and research practices. More specifically, we draw from intersectionality theory (Crenshaw, 1989, 1991) to critique the unitary approach (Hancock, 2007) to policy, focusing specifically on the limitations of the exclusionary clause in the SLD definition. The unitary approach privileges a "universalistic perspective in which one 'variable' or marker of difference, such as [cognition], race or class, is assumed to be both more important than the others and stable" (Artiles, 2013, p. 337). This universalistic approach leads to what Hancock (2007) refers to a "logic of group solidarity" (p. 65), or the assumption that a single, unitary category "serves to bind people into a political group based on a uniform set of experiences" (p. 65). In this case, the assumption is that psychological or biological factors are the defining features of learning disabilities.

In order to advance an intersectional framing of SLD, we critique the unitary approach that informs SLD scholarship and policy. For this purpose, we discuss (a) how the unitary approach shapes the construction of double bind identities; (b) the limits of the unitary approach associated with the erasure of historical legacies of inequity, the reliance on epistemic cultures that ignore culture and other markers of difference, and the lack of attention to ideologies of difference that may perpetuate the racialization of SLD despite the exclusionary clause; and (c) why the SLD field needs to engage with robust views of culture to enrich its knowledge base. We end the chapter with a brief reflection on future directions for SLD research.

Disrupting double bind identities: the promise of intersectionality

The history of the SLD exclusionary clause reveals efforts that were made to exclude race, culture, poverty, and environmental factors from consideration in identification practices. Yet, despite the clause, SLD has been racialized over time. This means that although a diagnosis affords educational opportunities to students with SLDs, new inequities can unwittingly emerge for students located at the intersection of ability, racial, socioeconomic, and cultural differences as reflected in the racialization of disabilities. Indeed, the quest for educational equity for students with disabilities has been fraught with paradoxes and inequities that create double bind identities for students as disabled *and* racialized (Artiles, 2011). Both categories bring stigma and disadvantage, and a complicating factor is that each category rests on disparate notions of justice—students with disabilities have sought *differential* treatment through individualization, while racialized learners have struggled for *equal* treatment through the abolition of segregation and equal access to resources. This situation has become more complicated as students with disabilities now aspire to be treated the same (via the standards movement) *and* differently (through individualization) (McLaughlin, 2010). In the end, these students acquire double

bind identities and are positioned in contradictory circumstances as the remedy for one group—e.g., special education—can unwittingly have detrimental consequences for other groups—e.g., students of color—through placement in more segregated settings or with limited access to the general education curriculum or related services than their white counterparts *with the same disability label*; thus, muddling the effects of well-intentioned efforts to expand equity.

An intersectionality lens exposes the limitations of the unitary approach (Crenshaw, 1989, 1991), revealing gaps and silences in single-category approaches to research, policy, and the law. To illustrate, policies focused on racial inequality tend to emphasize the experiences of Black males while gender equity policies implicitly focus on white women; thus, marginalizing the needs and experiences of African American women (Crenshaw, 1991). Additionally, an intersectional approach reveals how interlocking oppressions at the macrolevel (e.g., racism and ableism) become enacted in the day-to-day lives of people of color (Connor et al., 2016). In this way, intersectionality enables us to examine the different experiences and opportunities afforded to students of color and white students *with the same disability label*. Thus, from a structural intersectionality perspective, learners located at the intersection of racialized and disabled groups have qualitatively different experiences than students not positioned at this juncture (i.e., students with disabilities or racialized learners). From a political intersectionality standpoint, racialized students with SLD are given identities that pursue different (often conflicting) equity agendas, as explained in the preceding paragraph. For multiply marginalized learners, conflicting equity agendas contribute to schools acting as sites of dispossession where schools become more focused on surveillance and control, stealing essential time, resources, and relationships from students (Fine and Ruglis, 2009; Tefera, 2019).

As Hankivsky and Cormier (2011) note, intersectionality, similar to other critical frameworks, reveals the ways laws and policies are not neutral nor are they experienced in the same way by all people, contributing to key differences that have to be taken into account when developing, enacting, and evaluating public policy. To this end, intersectionality promises to disrupt these double bind identities by demonstrating how education policies often fall short in addressing the needs of learners at the intersection of multiple categories of difference, which is evidenced in the federal SLD definition. A growing number of scholars are demonstrating how policies can be written and enacted in a way that considers intersectionality (Hankivsky and Cormier, 2011). For instance, Bishwakarma et al. (2007) provide guidance on how to systematically integrate intersectionality in the policy-making process by structuring the problem in a way that considers how various social groups may experience a defined policy differently within specific contexts. In this way, there is a focus on the intersectional nature of the problem and the ways corresponding policies can more appropriately meet the intersectional lives of learners.

Dangers of the unitary approach: history, epistemic cultures and ideologies of difference

The unitary approach to define SLD raises difficult questions about this field and has important implications for research and practice. First, the exclusive attention to psychological factors erases problematic historical antecedents of contemporary conceptions of student competence. For instance, while practice dictates that teachers hold expectations for students that are “grade appropriate,” students who do not respond within an “acceptable” range are often regarded as falling outside of the prescribed “norm” (Reid and Valle, 2004). We must remember, however, the idea of the “normal child” is historically intertwined with race and other forms of difference in complex ways. After all, “disability has refused to remain linguistically stable On the one hand, disabled populations remain segregated and marginal throughout history. On the other hand, cultures persist in creating difference into a paramount trope of the human condition” (Mitchell, 2009, p. ix). For example, people of African descent were viewed as sub-human based on biological assessments such as cranium measurements that were used to justify the degradation of African Americans through policies and practices, including slavery, racial segregation, and unequal opportunities (Valencia, 1997). The history of IQ testing is also fraught with questions of cultural bias that disadvantage students of color and reduces their educational opportunities. Alas, the historical entanglements of race and ability are enmeshed in schools today (Artiles, 2019). Racial disparities in identification, discipline, and the segregation of students of color with disabilities are just some of the enduring injustices in US schools and the SLD field cannot afford to ignore these historical legacies (Skiba et al., 2016). We must remember that societal struggles surrounding inclusion and segregation of distinctive groups signal critical tensions that constitute its very core. For this reason, intersectional analysis of SLD should not be concerned only with the history of this condition, but with unveiling “what is most hidden under this history” (Stiker, 2009, p. 19).

Second, the requirements of the exclusionary clause call for attention to specific sociocultural markers. Yet, most SLD research has ignored these considerations. For instance, there have been efforts to expand research practices through more comprehensive descriptions of research samples (e.g., intersections of race/ethnicity, language, socioeconomic status). The Council for Learning Disabilities (CLD) and the National Joint Committee on Learning Disabilities (NJCLD) set minimum standards for including demographic markers of SLD students within research (Reed et al., 2012). Nonetheless, published studies continue to overlook how social categories intersect with SLD, particularly for Native American, Asian, and emergent bilingual learners (Reed et al., 2012).

Ironically, more than half of states in the US (33 out of 50) identified emergent bilinguals for SLD at a higher rate than the national average, with identification rates as high as 65–70% of emergent bilinguals in some states identified with SLDs (WIDA, 2017). Furthermore, when students of color are disproportionately identified with SLD, they are more likely to be placed in segregated special education classrooms compared to white students with the same diagnosis (Fierros and Conroy, 2002). These are unsettling practices that demonstrate the ways students of color with SLD can experience the label and presumed accompanying

supports in differential ways, despite the presence of the exclusionary clause. Given the lack of attention to exclusionary categories in research reports, should we assume that researchers consider these factors irrelevant? What factors maintain the erasure of environmental, cultural or economic considerations in research papers in the SLD field? These questions challenge the SLD field to take a critical look at its own knowledge production practices and take corrective measures where needed.

Third, how do practitioners manage to racialize the SLD category in certain locations despite the conceptual blinders encoded in the definition's exclusionary clause? Let us consider, for example, the case of Response to Intervention (RTI). RTI is a tiered system that includes assessments and interventions to support student achievement and address behavioral challenges. RTI provides opportunities for states and school districts to use alternative processes that circumvent the limits of the discrepancy formula approach to determine whether a student has SLD. Although The Individuals with Disabilities Education Act (IDEA) does not require schools to adopt RTI, many schools now use *both* the discrepancy model and RTI. Initially hailed as potentially addressing persistent equity challenges, including racial disproportionality, uncertainties and misunderstandings about how to implement RTI with fidelity in school systems serving heterogeneous populations and how to diagnose SLD with evidence from RTI have been reported from various states (Deshler, 2015).

For example, in a study examining the relationship between RTI and SLD disproportionality, Bouman (2010), found that although there was a decrease in the overall number of students identified with SLD, there was an increase in the disproportionate identification of African American students with learning disabilities. Similarly, Hartlep and Ellis (2012) found that RTI did not significantly reduce the likelihood of SLD identification, including for African American and Native American students who remained disproportionately represented. Thus, the intersectional aversive logic of the SLD definition and unitary focus on interventions and identification practices ultimately fall short in addressing key equity challenges that have plagued the field. How do we explain the persistence of the racialization of SLD across various identification models? Are there *ideological assumptions about the competence of students of color* that permeate these identification practices (Artiles, 2020)? An intersectional analytic framework, with its attention to historical legacies and hierarchies of power in social categories, promise a useful resource to examine these puzzling trends.

Why culture matters in SLD identification practices

We explained in the history section above the impetus for the exclusionary clause, particularly the focus on "environmental, cultural, or economic disadvantage." Two factors contributed to the inclusion of the clause. First, a concern with group recognition and access to resources. In fact, the terms included in the clause were code for "poor and racialized groups." The goal was to keep these groups outside of the SLD prevalence to protect budget allocations in the federal policy. Second, we situated the choice of the clause terms in the context of the 1960s in which the notion of "culture of poverty" and "cultural disadvantaged" were in vogue. These terms were loaded with ideological assumptions that kept structural inequalities invisible while they perpetuated the pathologizing of students of color (Scott, 2007). Here we return to this history to clarify why we need attention to culture in SLD, but also why it is necessary to avoid colonial conceptions of culture.

First, we underscore the importance of recognizing the cultural nature of human learning and development (Rogoff, 2003). The underlying view of culture in this perspective is dynamic, historically grounded and instrumental. A basic premise is that culture is *not* destiny; people *reproduce* the cultural practices learned at their homes and communities, but also *produce* new cultural practices through improvisation and borrowing from other groups—the latter point is significant because it prevents the stereotyping of cultural group members. Learning is cultural because human beings learn through mediation with cultural (psychological and material) artifacts and develop multiple and hybrid cultural repertoires as they participate in various communities (family, peer, school, recreational, etc.) (Artiles, 2015).

Unfortunately, rather than focusing on the cultural resources that students bring to the classroom and the cultural tools they deploy in their learning processes, culture and context are overlooked in the diagnosis of SLD. This results in an identification process that is largely driven by a vision of a-cultural human beings and that is reinforced by a system fixated on technical compliance with federal and state identification laws—i.e., driven by rules and regulations that tend to standardize subjects.

A cultural vision of learning compels professionals to understand students' sociocultural backgrounds and contexts when engaged in assessment and teaching processes. Yet, these considerations are often neglected in SLD identification practices (Fletcher and Navarette, 2011) or addressed with a unitary logic in which cognitive or biological considerations are paramount. In contrast, the perspective we subscribe to avoids stereotyping of students based on static proxies of culture such as racial background. Instead, it directs attention to the intersectional repertoires of practices and tools that individuals curate over time as they participate in multiple cultural communities.

Thus, a Latinx girl in middle school would likely develop repertoires through participation in her nuclear and extended families, her peer affinities with school, neighborhood and social media communities, as well as through her participation in other groups (e.g., sports, recreation, cultural, volunteer activities). In all of these social spheres are intersections of gender, linguistic, racial, ethnic, social class, and sexual orientation identities. This young learner still has primary affiliations in which she has acquired the core of her cultural practices, but she has likely curated hybrid repertoires derived from her multiple affinities. A *structural intersectional analysis* of her repertoires would assume that identity intersections place her, depending on the situation, in a particular social location; sometimes affording privileges, other times closing opportunities; thus, making her experiences as a Latinx adolescent both comparable and different in some respects to other peers of similar and distinct racial/ethnic identities.

In turn, a *political intersectionality* analysis would examine the political agendas of her various identities, searching for potential contradictions and documenting the consequences. For children and youth who are members of subordinated groups that may have conflicting agendas—e.g., racialized and disability groups—the consequences are considerable. As Crenshaw (1991) explained, “[t]he need to split one’s political energies between two sometimes-opposing groups is an additional dimension of intersectional disempowerment” (p. 360). This is why it is imperative that student school performance is informed by intersectional analyses that transcend stereotypical conceptions of culture and identity that often reify negative portrayals of students of color.

There is a second, equally significant, conception of culture that is relevant to the SLD field. This view focuses on the *cultural practices of institutions* (Gallego and Cole, 2001). For instance, schooling cultural practices rest on assumptions about what constitutes student competence, what curriculum content is worth prioritizing, what question formats and types of student behavior are acceptable to gauge learning. Instructional routines, ways of participating in classroom discourse, and seating arrangements also reflect the institutional culture of schools. Similarly, deficit-laden ideologies that assign static and stereotypical views of culture to students of color often permeate school culture. The idea of “culture of poverty” which sustains a damaged imagery of students of color (Scott, 2007) is an enduring example.

It is not uncommon to see this kind of ideology underlying explanations of the racialization of disabilities. For example, in a study we conducted with colleagues (Tefera et al., under review) we examined suburban school districts frequently cited for racial disproportionality in the identification and disciplining of multiply marginalized learners. We found that leaders and educators subscribed to deficit perspectives and a damage imagery related to the cultural practices of the increasing number of families of color who moved into the school district. We documented how deficit tropes were enacted given educators’ perceptions that the historically “high expectations” and “high achievement” of the school district was interrupted by families of color who, while searching for greater opportunities for their children, embodied cultures that were deemed to be misaligned with the culture of high expectations in the district. In a predominately white school district, African American students’ multiple identities (e.g., gender, sexuality, income, etc.) and cultural assets were ignored, with the key assumption being that African American students were unprepared and thus did not belong in the district. One of the key consequences, therefore, was that educators’ practices were left relatively unexamined since the problem was in the (essentialized) culture of students, families, and communities who were determined to be in need of changing and conforming to the expectations of the district.

Institutional cultural practices are also encoded in RTI protocols. The unitary focus of the SLD definition aligned with RTI’s uses of interventions and monitoring strategies that target cognitive skills devoid of the cultural repertoires that students of color bring from their communities may worsen inequities for multiply marginalized learners with SLD. These challenges make it essential to disrupt “the cultural inclination to focus on what is wrong with individual children ... [and] seek data showing children more skilled than schools have categories or time to notice, describe, diagnose, record, and remediate” (McDermott et al., 2006, p. 15). A number of scholars are focused on the cultural assets and practices that encourage a dynamic relationship between students’ cultural repertoires and school cultural practices (Bang et al., 2016; Gutiérrez and Rogoff, 2003). It is therefore essential that the assessments, interventions, instructional methods, and educational goals account for the complex intersecting identities and cultural practices that students bring to schools and classrooms.

Reflections for future SLD research

Our intersectional critique of SLD raises questions about difference and equity that are often neglected in this field. The ambiguity of the current definition and the ways it has been enacted have contributed to a situation in which learning difficulties associated with poverty and racialized status are presumably kept apart; yet we continue to see students from such groups differentially placed in this category. This paradox is connected to the dual nature of disability as an object of protection and a tool of stratification (Artiles, 2011), and it is sometimes perpetuated by the very legislation created to disrupt disability inequalities. Mitchell (2009) illustrates this point regarding the American Disabilities Act (1990), explaining “[l]egislation is concerned only that disabled people *pass the threshold to invisibility*. We reference the ADA as an indicator of our society’s commitment to social justice while failing to achieve demonstrable evidence of its successful implementation” (p. xiii, emphasis added). We must break these circuits of miseducation through the deployment of intersectional analytic practices, especially because educational equity tools (i.e., special education) can be unwittingly used to make students of color further invisible.

We must also note that although this chapter focuses on SLD in the US, a growing number of scholars are discussing the hegemonic narrative of disability identification practices and policies in other regions of the world (Grech, 2015; Kalyanpur, 2020; Werning et al., 2016). This is particularly noticeable in efforts to transfer disability knowledge across national contexts, particularly from the global North to the South.² Indeed, the global North and South are not merely “a geographical

²We do not frame the terms Global North and South based on a binary logic. Rather, we use the terms “to delineate power, resource, epistemological and other differentials, that though not unshiftable, and though not localized, embrace a substantial portion of the world living in a scenario of profound geopolitical asymmetries, poverty and isolation confronting deeply entrenched centers of concentrated wealth and power accumulated historically and perpetuated in times of coloniality. There is not one global South, but there are indeed many global ‘Souths.’ Furthermore, the global South is not only present in, but it also lives within the global north” (Grech, 2015, Terminology section, para. 3).

delineation of space, but a historical and sociological context that has traditionally excluded and disadvantaged certain communities of people" (Kalyanpur, 2020, p. 296). Kalyanpur (2022), for example, demonstrates the troubling parallels between SLD in the US and India, asking key questions about the role of power in the construction of SLD in India, including *Why is the label [SLD] itself based on the American definition? Why does the emergence of this category [SLD] look similar to that in the US?*" (p. 16). These questions are relevant and important not only to India but many parts of the global South that are adopting disability approaches from the global North, particularly given the ambiguities and intersectional shortcomings of SLD definitions described in this chapter.

At the same time there are important advancements in the field of SLD. Researchers, for instance, have mapped specific areas of the brain involved in aspects of reading (e.g., recognition of single words) and are documenting that the brain anatomy, physiology, and functionality of individuals with dyslexia differ from the brain of individuals without dyslexia (Artiles et al., 2020). Intervention studies are showing that some of the under-activated brain reading areas can be activated.

Most importantly, recent brain studies have shown that successful interventions change both physiology and anatomy broadly throughout the extended reading networks, including the prefrontal cortex, the right hemisphere, and in medial areas (cingulate cortex and insula) as well Researchers have interpreted these changes not as "normalization" but as *compensation*—the reading disabled reader's brain is compensating for weaknesses by *recruiting new areas of the brain into a different kind of reading network*. This suggests there is *great diversity in the parts of the brain that reading requires, and great diversity among individuals in the kinds of reading brains they develop*

Artiles et al. (2020, p. 166, emphases added).

Therefore, a key insight from recent neuroscience research is that "the brain's potential is not simply a matter of preprogrammed specialized modules but depends crucially on culturally organized experience" (Artiles et al., 2020, p. 162). For this reason, it is necessary to frame human development as constituted by bio-cultural influences.

Additionally, scholars are advancing research on how to engage in an intersectional policy approach. An intersectional policy approach would consider not only multiple categories of difference within the SLD category, but also situating policies within historical and social contexts. Furthermore, an intersectional policy approach links "educational outcomes with the context of people's lives—that is, where they live and how this is linked to job and food security, unemployment, and health insurance" (Hankivsky and Cormier, 2011, p. 227). In this sense, an intersectional policy approach recognizes and aims to address the interlocking of multiple structures of oppression that shape students' lives and opportunities. There is also growing scholarship in the US and abroad focused on redefining the boundaries of intersectionality in policy in ways that consider the voices of marginalized learners (Robert and Yu, 2018). Indeed, an intersectional policy approach that considers multiple categories of difference, context, interlocking of oppressions, and student voice holds important promise for future SLD research.

New advancements have also been made in the types of interventions that embrace the complexities of culture and students' intersectional lives. For example, Bal et al. (2018) demonstrate how to redesign interventions related to behavior that includes a focus on culture with Culturally Relevant Positive Behavioral supports (CR-PBIS). Although Bal et al. (2018) focus on an intervention related to behavior, it is equally critical that interventions focused on SLD such as RTI expand in ways that embrace students' multidimensional lives and dynamic notions of culture. Specifically, CR-PBIS focuses on bringing school community members and stakeholders together, including students and their families, to reflect on patterns in data, including disciplinary outcomes to improve practices and initiate transformation within a situated context. CR-PBIS is one of few interventions designed to explicitly consider students' intersectional lives, culture, and context to improve equity, and provides an important roadmap for similar advancements related to SLD interventions.

In summary, in this chapter we discuss the roles of biological factors and cultural influences that have been significant in the history of SLD. The former has dominated causal explanations while the latter has been taken up in ambiguous ways—i.e., at some point as a means to secure resources, other times as expressions of historical archetypes that advance the othering of students and communities of color. In addition, we outline the importance of embracing an intersectional framing of SLD by disrupting the unitary approach that informs SLD scholarship and policy. Specifically, we discuss the limits of the unitary approach associated with the erasure of a history of inequities, including culture and other markers of difference. We also demonstrate the need for the field of SLD to engage with more robust views of culture to enrich the current knowledge base. Finally, we believe that important advancements in scholarship focused on the biological layers of SLD, an intersectional approach to policy, and more dynamic engagement with equity and culture in interventions hold promise for enriching current conceptualizations and practices related to the field of SLD.

Acknowledgments

We are grateful to David Connor, Scot Danforth, Michael Gerber, and Todd Fletcher for their insights and expertise regarding the history of SLD in the US.

References

- Artiles, A.J., Kozleski, E., Trent, S., Osher, D., Ortiz, A., 2010. Justifying and explaining disproportionality, 1968-2008: a critique of underlying views of culture. *Except. Child.* 76, 279-299.
- Artiles, A.J., Rose, D., González, T., Bal, A., 2020. Culture and biology in learning disabilities research: legacies and possible futures. In: Nasir, N.S., Lee, C., Pea, R., de Royson, M.M. (Eds.), *Handbook of the Cultural Foundations of Learning*. Routledge, pp. 160-177.
- Artiles, A.J., 2011. Toward an interdisciplinary understanding of educational equity and difference: the case of the racialization of ability. *Educ. Res.* 40, 431-445.
- Artiles, A.J., 2013. Untangling the racialization of disabilities: an intersectionality critique across disability models. *Du. Bois Rev.* 10, 329-347.
- Artiles, A.J., 2015. Beyond responsiveness to identity badges: future research on culture in disability and implications for RTI. *Educ. Res.* 44, 1-22.
- Artiles, A.J., 2019. Re-envisioning equity research: disability identification disparities as a case in point. *Educ. Res.* 48, 325-335.
- Artiles, A.J., 2020. Interrogating Ideology & Ontology Circuits in Inclusive Education: Reflections on a Future Research Program [Virtual Event]. Universidade do Estado de Santa Catarina, Florianópolis, Brazil.
- Bal, A., Afacan, K., Cakir, H., 2018. Culturally responsive school discipline: implementing learning lab at a high school for systemic transformation. *Am. Educ. Res. J.* 55 (5), 1007-1050.
- Bang, M., Faber, L., Gurneau, J., Marin, A., Soto, C., 2016. Community-based design research: learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind Cult. Activ.* 23 (1), 28-41.
- Bishwakarma, R., Hunt, V.H., Zajicek, A., 2007. Educating dalit women: beyond a one-dimensional policy formulation. *Himalaya* 27 (1), 27-38.
- Bouman, S.H., 2010. Response-to-Intervention in California Public Schools: Has It Helped Address Disproportionality Placement Rates for Students With Learning Disabilities? A Dissertation Submitted to Claremont Graduate University.
- Connor, D., Ferri, B., Annamma, S., 2016. *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press.
- Crenshaw, K., 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, pp. 139-167.
- Crenshaw, K., 1991. Mapping the margins: intersectionality, identity politics, and violence against Women of Color. *Stanford Law Rev.* 43 (6), 1241-1299.
- Danforth, S., 2009. The Incomplete Child: An Intellectual History of Learning Disabilities. Peter Lang.
- Deschenes, S., Cuban, L., Tyack, D., 2001. Mismatch: historical perspectives on schools and students who don't fit them. *Teach. Coll. Rec.* 103, 527-547.
- Deshler, D.D., 2015. Moving in the right direction, but at what speed and how smoothly? *Remedial Spec. Educ.* 36 (2), 72-76.
- Donovan, M.S., Cross, C.T. (Eds.), 2002. *Minority Students in Special and Gifted Education*. National Academy Press.
- Fierros, E.G., Conroy, J.W., 2002. Double jeopardy: an exploration of restrictiveness and race in special education. In: Losen, D.J., Orfield, G. (Eds.), *Racial Inequity in Special Education*. Harvard Education Press, pp. 39-70.
- Fine, M., Ruglis, J., 2009. Circuits and consequences of dispossession: the racialized realignment of the public sphere for U.S. youth. *Transform. Anthropol.* 17 (1), 20-33.
- Fletcher, T., Navarette, L.A., 2011. Learning disabilities or difference: a critical look at issues associated with the misidentification and placement of Hispanic students in special education programs. *Rural Spec. Educ. Q.* 30 (1), 30-38.
- Franklin, B.M., 1980. From backwardness to LD: behaviorism, systems theory, and the learning disabilities field historically reconsidered. *J. Educ.* 162 (4), 5-22.
- Fraser, N., 2007. Re-framing justice in a globalizing world. In: Lovell, T. (Ed.), *(Mis)recognition, Social Inequality, and Social Justice*. Routledge, pp. 17-35.
- Gallego, M.A., Cole, M., 2001. Laboratory of Comparative Human Cognition, 2001. Classroom cultures and cultures in the classroom. In: Richardson, V. (Ed.), *Handbook of Research on Teaching*, fourth ed. American Educational Research Association, pp. 951-997.
- Grech, S., 2015. Disability and Poverty in the Global South: Renegotiating Development in Guatemala. Palgrave Macmillan.
- Grünke, M., Cavendish, W.M., 2016. Learning disabilities around the globe: making sense of the heterogeneity of the different viewpoints. *Learn. Disabil.* 14 (1), 1-8.
- Gutiérrez, K., Rogoff, B., 2003. Cultural ways of learning: individual traits of repertoires of practice. *Educ. Res.* 32 (5), 19-25.
- Hallahan, D.P., Pullen, P.C., Ward, D., 2013. A brief history of the field of learning disabilities. In: Swanson, H.L., Harris, K.R., Graham, S. (Eds.), *Handbook of Learning Disabilities*, second ed. The Guilford Press, pp. 15-32.
- Hancock, A.-M., 2007. When multiplication doesn't equal quick addition: examining intersectionality as a research paradigm. *Am. Polit. Sci. Assoc.* 5 (1), 63-79.
- Hankivsky, O., Cormier, R., 2011. Intersectionality and public policy: some lessons from existing models. *Polit. Res. Q.* 64 (1), 217-229.
- Hartlep, N.D., Ellis, A.L., 2012. Just what is response to intervention and what's it doing in a nice field like education? A critical race theory examination of RTI. In: Gorlewski, J., Porfilio, B., Gorlewski, D. (Eds.), *Using Standards and High-Stakes Testing for Students: Exploiting Power With Critical Pedagogy*. Peter Lang, pp. 87-108.
- Kalyanpur, M., 2020. Disrupting the narrative of universality of inclusive education: the new marginalization of low-income, English Language Learners in India. *Educ. Forum* 84, 296-308.
- Kalyanpur, M., 2022. *The Emergence of the Category of Learning Disability in the US and India*. Palgrave Macmillan.
- Martin, E., Personal Communication, November 29, 2020.
- McDermott, M., Vossoughi, S., 2020. The culture of poverty, again. *Diaspora Indig. Minor. Educ.* 14 (2), 60-69.
- McDermott, R., Goldman, S., Varenne, H., 2006. The cultural work of learning disabilities. *Educ. Res.* 35 (6), 12-17.
- McLaughlin, M.J., 2010. Evolving interpretations of educational equity and students with disabilities. *Except. Child.* 76 (3), 265-278.
- Mitchell, D.T., 2009. Foreword. In: Stiker, H.J. (Ed.), *A History of Disability*. University of Michigan Press.
- Monroe, M., 1932. *Children Who Cannot Read*. The University of Chicago Press.
- National Center for Education Statistics, 2020. The Condition of Education. Students With Disabilities. Retrieved online from: https://nces.ed.gov/programs/coe/indicator_cgg.asp.
- Reed, D.K., Sorrells McCray, A., Cole, H.A., Takakawa, N.N., 2012. The ecological and population validity of reading interventions for adolescents: can effectiveness be generalized? *Learn. Disabil. Q.* 36 (3), 131-144.
- Reid, K., Valle, J.W., 2004. The discursive practice of learning disability: implications for instruction and parent-school relations. *J. Learn. Disabil.* 37 (6), 466-481.
- Robert, S.A., Yu, M., 2018. Intersectionality in transnational education policy research. *Rev. Res. Educ.* 42, 93-121.
- Rogoff, B., 2003. *The Cultural Nature of Human Development*. Oxford University Press.
- Scott, D.M., 2007. *Contempt & Pity: Social Policy and the Image of the Damaged Black Psyche, 1889-1996*. University of North Carolina Press.
- Skiba, R., Artiles, A.J., Kozleski, E.B., Losen, D., Harry, B., 2016. Risks and consequences of over-simplifying educational inequities. *Educ. Res.* 45, 221-225.
- Sleeter, C.E., 1986. Learning disabilities: the social construction of a special education category. *Except. Child.* 53 (1), 46-54.
- Stiker, H.J., 2009. *A History of Disability*. University of Michigan Press.
- Tefera, A.A., Artiles, A.J., Kramarczuk, C., Aylward, A., Alvarado, S. (under review). The Aftermath of Disproportionality Citations: An Interdisciplinary Re-framing of Racial Disparities in Special Education.
- Tefera, A.A., 2019. Listening to and learning from the perspectives and experiences of Black and Latinx students with disabilities: examining the challenges and contradictions of high-stakes testing policies. *Urban Rev.* 51 (9), 1-20.

- US Department of Education, 2018. IDEA. Sec. 300.8 (c) (10). Retrieved online from. <https://sites.ed.gov/idea/regs/b/a/300.8/c/10>.
- Valencia, R.R., 1997. *The Evolution of Deficit Thinking: Educational Thought and Practice*. RoutledgeFalmer.
- Werner, H., Strauss, A.A., 1941. Pathology of figure-background relation in the child. *J. Abnorm. Soc. Psychol.* 36 (2), 236–248.
- Werning, R., Artilles, A.J., Engelbrecht, P., Hummel, M., Caballeros, M., Rothe, A., 2016. Keeping the Promise? Contextualizing Inclusive Education in Developing Countries. Klinkhardt.
- WIDA, 2017. Identifying ELLs With Specific Learning Disabilities: Facts, Advice, and Resources for School Teams. Retrieved online from file. [Users/adaitefera/Downloads/FocusOn-Identifying-ELLs-with-Specific-Learning-Disabilities.pdf](#).

Borderlands: unique sources of transnational language and literacy for educational research

Rosalind Horowitz, International Discourse & Literacy Studies, Culture and Learning Sciences, Departments of Interdisciplinary Learning and Teaching, Educational Psychology, College of Education and Human Development, The University of Texas at San Antonio, San Antonio, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Background	288
Characterizing the traversing of borderlands	288
Crossing borders: increasing opportunities for language, literacy, and identity self-positioning	289
Border affect and emotional encounters	290
Recent growth of hispanic population and in educational research	290
Part I Language and literacy educational research on the Texas-side of the U.S. Mexico border	290
Youth who regularly cross the United States–Mexico border	290
A research study of Border-Hispanic Adolescents	291
Rampant stereotypes of border lands and youth populations	291
Research to uncover realities of United States–Mexico border adolescent populations	291
Crossing from Mexico to the United States-side of the border – and returning to Mexico	292
Part II Bedouin border life in the Negev of Israel	292
New measures for study of border populations?	292
Oral testimonials of border life and authentic memory accounts of language conflicts	293
Hierarchical language perspectives	293
Cultural assimilation through name changes	293
Language functions and self-perceptions of identity	293
Part III Constructing a unique teacher education for border adolescents	293
The border as bridge for transnational and international language and literacy	293
Shifts in educational research and pedagogy	293
References	294

Background

Part I of this review addresses the U.S.–Mexico Borderland/La Frontera, The Frontier, the new mestiza, culture and increasing language uses or mixing on the border. The Texas-Mexico Borderland is at the same time a land mark for acquiring new literacies. In Part II, comparisons are made to other border communities, particularly those border communities surrounding Israel, in Beer-sheba, land near the Gaza Strip, that is of interest to geographers and historians studying Bedouins and multi-cultural communities. Part III offers proposals for constructing a unique teacher education that will be vital for designing language and literacy for border populations.

Crossing borders holds potential for transforming of the Self and constructing a new, multilayered identity. However, this opportunity is also fraught with challenges which educators and researchers must address. In this essay, attention to Borderlands contribute to our understanding of how adolescents conceptualize who they are and will become as they grow into adulthood while encountering more than one culture. The goal of the essay is to extend educational research to what has been a too long-neglected investigation of unique language and literacy education for border populations.

Characterizing the traversing of borderlands

Since adolescents growing up today on the United States-side of the Texas-Mexico Border (e.g. in El Paso, Brownsville, Laredo) travel routinely to and from Mexico, opportunities abound for interacting with family, friends, in neighborhoods in urban and rural Mexico. Cross-linguistic and cognitive-cultural shifts have begun to enter institutions of learning and classroom practices (Alvarez, 1995; Bhabha, 1998). On the United State -side of the border, “standard” English is typically perceived as the prime vehicle for literacy and upward mobility by teachers and students-alike. On the Mexico-side of the border, (e.g. in Neuvo Laredo, Metamoros) family interaction is in a first language, typically Spanish, with intergenerational registers or vocabulary. Spanish is viewed as the mark of cultural affiliation. What Joshua Fishman (1966) referred to in his seminal book, as a *Language Loyalty*. Cognitive-cultural shifts with intersections of cultures, in presentation of the self and voice, are common in oral discourse on each side of the border. These language shifts, with related cognitive realignments, are frequently a source of internal or externally expressed identity conflicts that our research team and others have addressed (Anzaldúa, 1987; Barrera, 2007; Horowitz, 1995, 2013, 2023; Hudson, 2014; Rodriguez et al., 2019). Adolescents here and elsewhere are typically engaged in exploring their identity, as they encounter language conflicts and/or mental struggles based on an “internal” and “external” landscape (Csikszentmihalyi and Larson, 1984; Jiménez,

2004). More recently, [Dyrness and Sepúlveda \(2020\)](#) describe Latinx youth with a “border thinking” with a cognitive perception that they are framed as “outsiders”.

Gloria Anzaldúa opens her seminal book with a portrait of living on the Texas—U.S.-Border, an account that may reflect the mind-set and emotions of many border residents. She describes her Border-identity in the following wrenching way:

I am a border woman. I grew up between two cultures, the Mexican (with a heavy Indian influence) and the Anglo (as a member of a colonized people in our own territory). I have been straddling that tejas-Mexican border, and others, all my life. It's not a comfortable territory to live in, this place of contradictions. Hatred, anger and exploitation are the prominent features of this landscape.

Preface to first edition, 1987, paragraph 2.

Relevant to the research about adolescent literacy habits to be reported below, Anzaldúa conveys how reading of books and oral language were central to her life.

This book, then, speaks of my existence. My preoccupations with the inner life of the Self, and with the struggle of the Self amidst adversity and violation; with the confluence of primordial images, with the unique positions consciousness takes at these confluent streams; and with my almost instinctive urge to communicate, to speak, to, write about life on the borders, life in the shadows.

Books saved my sanity, knowledge opened the locked places in me and taught me first how to survive and then how to soar ... The switching of ‘codes’ in this book from English to Castillian Spanish to the North Mexican dialect to Tex-Mex, to a sprinkling of Nahuatl to a mixture of all of these, reflects my language, a new language—the language of the Borderlands.

Preface to the first edition, 1987, paragraph 5.

Crossing borders: increasing opportunities for language, literacy, and identity self-positioning

The crossing of borders may be a life-changing event. It serves a variety of psychological functions. First, the “Crossing of a Border” forces adolescents and youth to interact with multiple-cultures and peoples—concomitantly, ways of speaking and thinking. The youth who cross borders, “transborder students” have access to multiple language styles, registers, and literacies—from newspapers, magazines, videos, genres of the oral within writing, such as poetry or songs, and music genres. Borders are the bridge for transitions in identity forming by hybrid cultures ([Canclini, 2005](#)). Historically, individuals remained on one side of a border whereby crossing a border was infrequent and burdensome. Or it was legally prohibited. With use of modern transportation, the opening of borderlands, and traversing through media, travel to and across borders is manageable and a frequent occurrence, especially among adolescents ([Brown et al., 2002](#); [De la Piedra, 2010](#)) bearing the energy and mindset. Transborder-students often live in Mexico but attend school in the U.S. given determination for a good education, coupled with possession of proper documentation. Still other transborder-students develop penpal programs in border spaces ([Perez, 2010](#)).

Second, Borders, themselves, are fluid. They change. They may be marked by bridges, or a fence, wall, line, a strip, river contributing to *how* border crossers traverse across nations. Youth frequently, rapidly respond to change and intersections. For example, the Russian-Border became Poland. It, later, was occupied by Russia ([Stoklosa, 2012](#)). With different leadership and regimes, borders have shifted whereby travel expanded or was constrained depending upon the leadership. With border populations and activity shifting, so have governmental expectations of travelers and traveler expectations of the self. Today, borders are often associated in the news with immigration restrictions, monitored by a border control ([Berry et al., 2006](#)). They have been characterized by a visual presentation on a map but also by a mental map that resurfaces in one’s mind.

This demarcation on a *physical map* or a *mental map* is referred to as a Border. In this chapter, the term Border refers not only to the physical separation of land but also encompasses the socio-cultural and psychological dimensions of the separation of space and land that persists within oneself. The Border is a means for adolescent “positioning” of one-self, what [Nelson et al. \(2016\)](#), [Nelson et al. \(2008\)](#) refer to as forming an “identity positionality”. This is a means for positioning oneself in one or multiple cultures, history, or local situation, to be described in more detail below. This positioning may influence one’s thinking about a community, friend-peer groups and the “other” ([West et al., 2019](#)). Positioning oneself *around* or *on* borders may influence attitudes toward a language, an audience, or the pursuit or neglect of higher schooling. Early on, high school students decide they are “more” Spanish-than English-centered. Their language of choice may contribute to whom they interact with most readily, and where they encounter daily friendships or disheartening conflicts ([Kosmutzky and Patty, 2016](#)). It may also contribute to their expectations for literacy—what language they read or how they write or speak ([Lam, 2009](#)). Wherever is “home,” determines oral and literate practices ([Guerra, 1998](#)) and spoken and written modes of meaning ([Halliday, 1987](#)). Maya [Hickmann \(2003\)](#) and [Hickmann and Robert \(2006\)](#) highlight challenges of space and time across languages. [Horowitz \(2017\)](#) describes adolescent literacy achievement from a global perspective and research that characterizes language conflicts and how high levels of literacy appear to be declining worldwide. While useful, such research often neglects to consider the uniqueness of specific border populations, the goal of this discussion.

Border affect and emotional encounters

The Borderlands of interest here are also associated with emotions, as [Anzaldúa projects \(1987\)](#). Border lands have historical, cross-generational meanings that provoke the emotions of pain, joy, spirituality, passions. In Israel, the land takes on a “holiness” and is treasured by Jews, Christians, and Muslims. Each group takes pilgrimages to historical sites ([Berger, 2011](#)). Wars, battles conveyed in the Bible, by cultures of the past are reiterated today with distinct meanings in written news reports, liturgy, or poetry. Today, battles are fought throughout the world for possession of land and control of borders. Along the Gaza Strip, conflicts have resulted in near-daily missile attacks on Israel, which cause significant trauma to its peoples.

Recent growth of hispanic population and in educational research

[Gándara \(2017\)](#) who served as Coordinator of the Civil Rights Project at UCLA, and Chair of the Working Group on Education for the California-Mexico initiative highlights the rich promise of Latino students. Without doubt, Latinos comprise the largest minority group of the U.S. population. In 2017, one of four K-12 students in the U.S. were Latino/a, with principle locations in California, New York, and Texas. She argues that more students have left the United States than have entered as immigrants, with 90% born in the U.S. (2017, p. 4). One may turn to numerous scholarly journals, to witness the increase in educational research of Hispanic populations and literacy education in schools as in the *Journal of Literacy Research* or the *Reading Research Quarterly* or *Reading and Writing Quarterly*. In fact, Hispanic populations have become the largest minority in the U.S. with significant variations in languages, lifestyle, and travels.

Part I Language and literacy educational research on the Texas-side of the U.S. Mexico border

The Border Literacy Project at The University of Texas at San Antonio included study of a high school on the Texas-side of the Texas-Mexico Border, where Hispanic adolescents travel daily by foot across a bridge, for a sheer 20 min to Mexico ([Horowitz, 2006](#); [Horowitz et al., 2009](#)). We have been interested in (a) how this travel influences transitional, transnational cognition (b) what positive language or literacy interactions and dire conflicts emerge among adolescents who are speaking English and/or Spanish when they travel across boundaries and (c) literacy choices that students make based on aspirations. This research has been reported in *A Bilateral Perspective on Mexico-U.S. Migration* ([Horowitz, 2012](#)), distributed through The University of Texas at San Antonio Mexico Center and at international conferences including the World Congress on Comparative Education (WERA) held in Istanbul, Turkey ([Horowitz, 2010](#)) and [Horowitz \(2013\)](#) discussion of realignment of language and transcultural identities (Guanajuato, Mexico).

There are many ways youth use language expressions to characterize and navigate movement across space and daily life. They use language and gesture as references-to-space to understand a problem or resolve cognitive dissonance (See Horowitz, 2021). Levinson proposes, “... different species paint mental pictures of their spatial environment using some pretty specialized sensory equipment” (2006, p. 221). In the case of humans, and youth per se, autobiographical memories, we have learned, are based in geographic locations in space and time. This includes capturing geographic space in a house, one’s block on a street, and bridges that lead to new spaces. Of late, language representing space is seen as the embodiment of language and contexts by social-psychologists and speech communication theorists, not to be ignored.

The theoretical work of Lev Vygotsky (1929/1994, 1978) highlighted the importance of culture and social context for child development, language, and literacy. This work is of great significance here for our purposes. Vygotsky’s theories about language created in social-situational contexts have led researchers across the world to look more closely at discourse patterns, functions of language in space—and, in our instance, among boundaries between nations. How adolescents use language within distinct space across a border may capture their mental attitudes and beliefs. [Saldivar \(1997\)](#) proposes a remapping of cultural studies with attention to borders.

Youth who regularly cross the United States–Mexico border

There are few scholarly studies of border crossings. Testimonial narratives, however, are surfacing from historical diaries, interviews, and surveys. Little is known about the oral and literate practices of migrant adolescent (13–19) and young adult (20–27) populations who travel across a border that would be useful for educators in developing pedagogical practices. [Gabaccia and Ruiz \(2006\)](#) report some of the largest immigrant groups across the world are from Mexico, China, and India and often, today, cross borders to return to their country of origin in order to maintain contact with family and friends. We have not found case studies that address the higher educational needs of those youth who regularly cross a border, in this instance, from the United States to Mexico—sometimes missing schooling and opportunities for advanced education or simply dropping-out of school ([Kennedy and Roudometof, 2002](#)). Some studies consider multi-lingual literacies in transnational digitally mediated contexts ([Lam and Rosario-Ramos, 2009](#)) and narratives told across borders ([Sanchez, 2007](#)). With the rise of the uses of the internet, transnational communication is rampant particularly among adolescents and youth.

A research study of Border-Hispanic Adolescents

We chose one high school to study on the United States-Mexico Border that had a reputation for serving Border-Hispanic Adolescents, successfully. While teaching at a Hispanic Serving Institution in South Texas, I interacted with our students from U.S. border towns who travel transnationally and regularly, to no end, to further their education. They helped identify the high school we studied—indicating its unique attributes in serving students who often travel to and from Mexico. The school was located two-and-one-half hours to the closest city. In 2008, based on Texas Education Association (TEA) data, 26% of the town was in poverty, with women working in education or health occupations and men largely in construction. When we first entered the school, in 2008–9, 99.9% of the students were Hispanic. There was only one Anglo among the 1591 students! At that time, the graduation rate included 83% of the students.

In 2019, ten-years later, there were only slight changes, based on Texas Education Association (TEA) data. There was a slight drop in Hispanic students, 95% of the school, with 95% economically-disadvantaged. However, we found 96% of the students had graduated, rather than 83% in 2008, an impressive increase for high school graduation—a rarity in urban schools in the U.S. where there is high secondary drop out and only 50% or more have graduated. See research by Robert Balfanz (2008) on what he terms, “U.S. high-school dropout factories.”

The high school we studied was characterized, in multiple ways, as one where youth were achieving, overcoming poverty-based struggles, and developing what Norman Garmezy, University of Minnesota (Garmezy, 1991 and personal communication) has coined as “Resilience.” There were key factors that our research team regarded as influential in the education of these high school students.

The teachers in the high school on this United States-side of the Texas-Mexico Border were residents of the neighborhood surrounding the school, and expressed, in informal interviews with this researcher, an unyielding loyalty to the adolescents, dedication to community and faith in the school. This is because they grew up as students in the very high school that we studied, many promising to “give-back” to the school in which they were raised. This school, on the border, built resilience in its students, for one, symbolized by the mascot, *El Toros*, a sculpture of a fierce bull, visible outside of the building. The students sought to use this bull and the school setting strategically, in ways that Anzaldúa characterized as a place to distinguish—“us” from “them” (1987, p. 25).

The principal of the high school was charismatic. He created an atmosphere that was energized. This principal projected a rapport with students that was captivating. We observed his greeting of students with handshakes, smiles, and verbal connectivity, particularly so, with the males in the hallways. Teachers followed his dialogue style, repeating similar acts of communication, rapport, and emotional support to connect to the adolescent students.

Rampant stereotypes of border lands and youth populations

The adolescents on the border of Texas-and-Mexico are portrayed as subject to poverty, violence through gangs, and physical or mental illnesses. This space is *not* characterized by the “Good American Life.” These negative generalizations abound on social media, news reports, and in the movie industry. From these labels, there is a perception of human deficiencies and limitations of possibilities for work and human growth. David Berliner (2017), former President of the American Educational Research Association, characterizes poverty as an unending culprit for linguistic constraints with deep learning foils. He shares how poverty is overlooked in developing educational research and practices.

Research to uncover realities of United States—Mexico border adolescent populations

The students we interviewed in a focus group, who attended the high school of interest, on the United States side of the border, faced many obstacles with trepidation often promulgated by poverty. However, what was amazing to our team was the drive and motivation that we heard for advancing learning and chosen careers. In fact, one might argue that poverty here was an impetus for sustaining achievement. An English Department Chair selected the youth we interviewed—for a focus group during the school day. At our request, they were selected because they were (a) engaged in frequent travel across the border (daily, weekly, or monthly; (b) identified as below the established poverty line; (c) were voracious readers; d) were highly verbal, based on verbal assessment data coupled with qualitative measures, and (e) willing to talk openly about their reading-life habits. We wanted to know when they read, how they read, and why they read.

A focus group was created, with four students attending the one high school and routine travelers across the Border to Mexico. Details of this study are reported in Horowitz (2012). One female student, Nava (all names have been changed), bilingual, reported travel across the border to visit her grandmother and grandfather and to eat in her favorite restaurant. She described reading profusely—from Shakespeare’s *Taming the Shrew* to the Bible daily, newspapers, and health books. Her wish was to become a pediatrician. She hoped to attend a research university in Texas. Her parents only speak and read Spanish. Another female student, Janie, described travel twice per month to Mexico. She viewed reading as central to her life. As a voracious reader, she conveyed buying “a minimum of two to three books per week” although Jamie revealed “her family frowns on it.” She described reading in a fast-food restaurant when there were no customers, hiding the book under the counter. Her joy was described as teen fiction and romance novels. She hoped to study Psychology at a four-year college. The two males in our focus group also talked about reading and their aspirations. Ed noted travel to Mexico to the Orthodontist once per month or to the beach. He preferred Manga books—Japanese graphic novels. Less sure than the females about what he will do after high school, Ed conveyed he may pursue technical work or the education field. Finally, David described travel three times per week across the border to visit his father who is a preacher of a large church. His travel was to go to *La Feria*, the fair, to buy CD’s, to shop or for less expensive medicine in Mexico. He conveyed purchasing books every week, usually fiction. He will pursue business administration until 40 and then write novels. In sum,

our focus group of high school adolescents showed strong reading interests, a commitment to distinct goals, and were articulate in English about their aspirations. As transnational travelers, these students were gaining fluency in speaking and reading English and Spanish.

Crossing from Mexico to the United States-side of the border – and returning to Mexico

Interestingly enough, a number of adolescents and youth have crossed from the U.S. only to return to live in Mexico. Research by Professor Tatyana Kleyn (2017), a Russian-born, immigrant who moved to the U.S., investigates the dilemmas these returnee youth face once living in Mexico and the psychological conflicts faced in the U.S.—resulting in a return to Mexico. In some cases, those Americans in Mexico seek to and do return to the States because they conveyed in interviews by Kleyn as seeing themselves as more “American” than “Mexican”. Surprisingly, many U.S. students living in Mexico, who returned to Mexico, arrive with their families, but have limited Spanish and cultural affiliations that will serve them well in Mexico, resulting in a return to the U.S. We have no data for educators about this population that I could locate at the time of this essay.

Part II Bedouin border life in the Negev of Israel

Another project at The University of Texas at San Antonio is in collaboration with Kaye Academic College and Dr. Doron Narkiss in Beersheba, Israel, about 20-miles from the Gaza Strip. Dr. Narkiss taught a *Seminar in Postcolonial Literature* – a course that examined minority versus majority issues of Identity and Inclusion, at Kaye Academic College, a teacher training institution. Students are international but largely Bedouins. They were asked to read Gloria Anzaldúa’s chapter (1987) “How to Tame a Wild Tongue” and respond re-telling about their own “wild tongue” story. Anzaldúa describes a trip to the dentist and the challenge of “dental (personal) extractions”, a metaphor used to describe how language of origin is often extracted, literally removed, from a people, in this case, the Spanish language is blocked. At UT-SA, we asked undergraduate Hispanic students in teacher training to read the same pages and to write about a language conflict and obstacle they too experienced in or near a border town and how it was resolved. Interestingly, they were many similar accounts of “language extractions” across these two locales, although I highlight the Israeli border accounts due to space limitations.

The largely female, future teachers at Kaye Academic College, Bedouins, Russians, Europeans, and other Israelis, live in and around Beersheba, a frontier of South Israel near Gaza, and are pursuing teaching as a career. The Bedouins, the largest local group, have been described as a “nomadic” culture. They are *not* required to serve in the Israeli Army, but many do volunteer to serve, much like the Druze of Israel; the Druze are another unique cultural group with multifaceted identities faced much like the Bedouins with constant questions of identity, citizenship, and patriotism (Nisan, 2010). To elaborate, the Bedouins in this region, like many, of United States Hispanic border populations, have ties to family outside of Israel in Gaza and cross a border for visits with relatives. Prior to 1948, Bedouins in the Negev were known as Arabs of Beersheba. In Beersheba, a city founded in 1900, where they have lived and grown. Since 1966, they engaged in livestock rearing and farming. Many live within tents or other insecure housing. They are attached to the land, living in hamlets, or Bedouin groups. Unemployment has been high, with 29% for men, and 80% unemployment for women. Some Bedouin have become more urbanized but it was noted by Kaye administration that over half of the families were in poverty and 60% of children in poverty. Border patrol with Gaza has been loose but with time more stringent in border crossings. The Bedouins cherish their land, and there has been constant struggle over land rights with the Israeli government that have been discussed by scholars and advocates of indigenous border populations (Elsana, 2021).

On my trips to Israel, and in 2017, I began developing a partnership between The University of Texas at San Antonio and Kaye Academic College in Beersheba. I visited college classrooms comprised largely of Bedouin youth from Beersheba, preparing to become educators in Israel. There were also Bedouin faculty helping to transition the students to a new mental framework that would serve them well in teaching. The women were struggling with their identity, as a Bedouin woman, and their life in a multi-cultural Israel, as an educator. While many did not travel to Gaza regularly, there were family and friends there with whom they expressed fierce loyalties and dynamic connections. Thus, one may recognize many commonalities between our Bedouin residents and our U.S. Mexico residents.

New measures for study of border populations?

What questions need to be asked about Border Populations? What measures do we use for understanding the motivations, interests, and literacy and educational aspirations of the population addressed in this chapter? Do we need new measures of individual learning within cultural-social spaces and practices that will inform instruction in classrooms (Brittain, 2009)? There are additional sources of information to acquire: (a) autobiographical and ethnographic accounts of those pursuing the teaching profession from Bedouin communities are needed; (b) historical accounts of the changes among Bedouin communities or hamlets are necessary in order to help students make a transition to current Israeli living and teaching; (c) what requires intervention is a tricky matter, as we want to respect the values and cultural practices of the Bedouin females and males alike. How do we sustain that respect? And (d) studies by comparative-education researchers are warranted. Are there specific methods for teaching that are successful with particularly first-generation college youth who cross borders—physically and psychologically world-wide? Such research is needed.

Oral testimonials of border life and authentic memory accounts of language conflicts

The Borderland research reported in this chapter used Narrative Testimonials and Memory Accounts to learn about the border language and discourse conflicts that our adolescents and college youth encountered (See Markowitsch and Welzer, 2009; Riessman, 2008). The following content themes suggest avenues for future educational research.

Hierarchical language perspectives

Anzaldúa's book chapter triggered painful thoughts expressed in writing by our immigrants in Gaza.

One female student at Kaye College highlighted a *hierarchy of attitudes* toward languages: "cultural superiority finds its way into every discourse. When Anzaldúa's teacher told her to speak American, she was telling her to think and feel American as well, and when any English speaker goes abroad and assumes that they will be speaking English, they are saying the same thing."

Another female wrote about *the metaphor of the dentist*, extractions, and non-frequently used "tongues". This becomes a metaphor, as noted by one female: "Anzaldúa uses a word game with the word 'tongue.' The word 'tongue' refers to a language dialect or pronunciations of power, and also to the tongue in our mouth. Also, in order to speak and let our voices be heard we must physically use the tongue in our mouth."

Cultural assimilation through name changes

Another female student, Russian, noted how her name was changed to an Israeli name for cultural assimilation: "I showed my drawing to the class and told everyone that my name is Katya, then the teacher asked me to stop my presentation and told me: 'Katya is not an Israeli name, from now on your name is going to be Ketty.'" And so she did to the other Russian pupils, by the end of the day we all had new Israeli names."

Still another male also wrote about name changes: "Many of the Russian and Ethiopians were also asked to change their names to more Israelis names and to work on their accent and pronunciation."

Language functions and self-perceptions of identity

A Russian male student living on the Israeli border of Gaza writes: "In this chapter [Anzaldúa] I could not but think about the question: Do we create language or does language create us? My answer to this question after reading is that language creates us. Our language is one of the things that create our identity and when we are ashamed of it we are actually ashamed of ourselves. That's why Anzaldúa writes in her book 'I am a language'".

Another student in Beersheba writes about his interpretation of the violence associated with language extractions: "After giving the dramatic and a bit violent description of the visit at the dentist, Anzaldúa in *Borderlands: La Frontera* (1987) quotes Ray Awyn Smith: "who is to say that robbing a people of its language is less violent than war" (p. 75) when right after that she starts telling her own story of struggling with the "language terrorists" who try to extinct any trace of her language and later on she shares her own and her people's "fight of pursuing the dream that one day their language will be legitimate, acceptable, approved and recognized by all."

Part III Constructing a unique teacher education for border adolescents

The border as bridge for transnational and international language and literacy

Among the various metaphors of the Border, for our purposes, *The Bridge* is the most significant and cross-culturally meaningful. As a *Bridge*, Borders allow for traveling across space, for transnational human contacts, the resolution of verbally expressed conflicts, and development of the prominence of language and literacies uses for specific societal and educational purposes. With time, as Michael Halliday noted (1973), certain language functions and styles do become prominent/necessary for the adolescent as they enter adulthood and the teaching profession. Border travel can become the valuable beginning of the adolescent living in multiple cultures, use of multiple languages, and understanding of the functions of multi-literacies— that make meaning of life, across international cultures.

Shifts in educational research and pedagogy

Teachers who serve borderland populations require a unique education and training. Jiménez et al. (2017) call for creating culturally responsive teachers of Hispanic and bilingual students learning English. This essay elaborates upon content, empathy, and teacher cognitive skills needed. Taking into account the aforementioned research, professional education and development should also include,

1. The history and current practices of the culture on the border, with its internal and external language conflicts;
2. The effective uses of bi-lingual communication in classrooms, by teachers and students;
3. The understanding of the daily and work life of the population on the border, and the obstacles given cultures face;
4. The role of family and emotional attachments to a language, literacies, and the culture of the population relevant to teaching unique border populations; and

5. The adolescent's creatively viable and distinctly "new aspirations" that influence future educators—teachers and administrators.

Such research will help inform the classroom curriculum and strategies that are used and sustained.

The recent AERA-OECD International Study of 8 countries (2021) presented by the American Educational Research Association over Zoom provides information about teaching practices in mathematics but is relevant beyond mathematics, to other disciplines of instruction. Professor Frederick K.S. Leung of the University of Hong Kong on the AERA-OECD Panel asked astute questions with an eye toward comparative education: What is good teaching that is universal? What is unique to a country? And most importantly, he has asked what is good teaching in a cultural setting? Borderlands capture more than one culture and provide avenues for advancing inter-cultural and language/discourse blending in unprecedented ways that we must address.

How do teacher-educators develop a literacy curriculum for those who cross a border, move from one nation to another? There are tremendous individual differences. In my research and that of others, we have found some border crossers prefer to preserve their first language and develop it. Others prefer continued growth in their first language with development in a second language, often English. Still another option is some individuals prefer to develop both a first and second language, recognizing that a second language may be an avenue for better work and life opportunities. These preferences are also evidenced in the literacy education, reading and writing practices. All of this makes the teacher's role a complex one. It requires knowing student backgrounds, individual needs and interests, and an ability to navigate multiple languages and literacies in cross-national contexts.

References

- Alvarez, R.R., 1995. The Mexican-U.S. border. The making of an anthropology of borderlands. *Annu. Rev. Anthropol.* 24, 47–70.
- American Educational Research Association (AERA-OECD), March 11, 2021. The Global Teaching Insights Video Study: Methodology, Data, and Research Opportunities. <https://www.aera.net/Events-Meetings/The-Global-Teaching-Insights-Video-Study-Methodology-Data-and-Research-Opportunities>.
- Anzaldúa, G., 1987. *Borderlands/La frontera: The New Mestiza*, second ed. Aunt Lute Books, San Francisco, CA.
- Balfanz, R., 2008. Do dropout factories really exist and if so what can be done about them?. In: Paper presented at the American Educational Research Association Conference. New York.
- Barrera, E., November 28, 2007. Giving voice at the margins: counter-discourse, Reggaeton, and critical literacy. In: Presented at The National Reading Conference. Austin, TX.
- Berger, P., 2011. Jewish-Muslim veneration at pilgrimage places in the Holy Land. *Relig. Arts* 15 (1–2), 1–60.
- Berliner, D.C., 2017. Poverty's powerful effects on reading achievement and the achievement gap. In: Horowitz, R., Samuels, S.J. (Eds.), *The Achievement Gap in Reading: Complex Causes, Persistent Issues, Possible Solutions*. Routledge, Taylor & Francis, New York, pp. 23–37.
- Berry, J.W., Phinney, J.S., Sam, D., Veddar, P., 2006. *Immigrant Youth in Cultural Transition: Acculturation, Identity, and Adaptation across National Contexts*. Lawrence Erlbaum, Mahwah, New Jersey.
- Bhabha, H.K., 1998. Cultures in between. In: Bennett, D. (Ed.), *Multicultural States. Rethinking Difference and Identity*. Routledge, New York, pp. 29–36.
- Brittain, C., 2009. Transnational messages: what teachers can learn from understanding students' lives in transnational social spaces. *High Sch. J.* 100–114.
- Brown, B.B., Larson, R., Saraswathi, T.S., 2002. *The World's Youth: Adolescence in Eight Regions of the Globe*. Cambridge, New York.
- Canciani, N.G., 2005. *Hybrid Cultures: Strategies for Entering and Leaving Modernity*. University of Minnesota Press, Minneapolis (Original work published 1990).
- Csikszentmihalyi, M., Larson, R., 1984. *Being Adolescent. Conflict and Growth in the Teenage Years*. Basic Books, New York.
- De la Piedra, M.T., 2010. Adolescent worlds and literacy practices on the United States—Mexico border. *J. Adolesc. Adult Literacy* 53, 575–584.
- Dyness, A., Sepúlveda III, E., 2020. *Border Thinking: Latinx Youth Decolonizing Citizenship*. University of Minnesota Press, Minneapolis.
- Elsana, A., 2021. Managing the tensions between service and advocacy: the case of AJEEC social change service Organization, Naqab, Israel. *Int. J. Soc. Welfare (Wiley Online Library)*.
- Fishman, J., 1966. *Language Loyalty in the United States*. Mouton, The Hague.
- Gabaccia, D.R., Ruiz, V.L., 2006. Migrations and destinations: reflections on the histories of US immigrant women. *J. Am. Ethn. Hist.* 26 (1), 3–19.
- Gándara, P., 2017, Spring. The potential and promise of Latino students. *The American Educator. Educating Latino students in an age of uncertainty* 41 (1), 4–11, 42–43.
- Garnezy, N., 1991. Resiliency and vulnerability to adverse developmental outcomes associated with poverty. *Am. Behav. Sci.* 34 (4), 416–430.
- Guerra, J.C., 1998. *Close to Home: Oral and Literate Practices in a Transnational Mexican Community*. Teachers College Press, New York.
- Halliday, M.A.K., 1973. *Explorations in the Functions of Language*. Elsevier, New York.
- Halliday, M.A.K., 1987. Spoken and written modes of meaning. In: Horowitz, R., Samuels, S.J. (Eds.), *Comprehending Oral and Written Language*. Academic Press, London and New York, pp. 55–82.
- Hickmann, M., 2003. *Children's Discourse. Person, Space, and Time across Languages*. Cambridge University Press, New York.
- Hickmann, M., Robert, S. (Eds.), 2006. *Space in Languages. Linguistic Systems and Cognitive Categories*. John Benjamins, Amsterdam, The Netherlands.
- Horowitz, R., 1995. Orality and literacy. The uses of speech in written language by bilingual and bicultural writers. In: Rubin, D. (Ed.), *Composing Social Identity in Written Language*. Erlbaum, Hillsdale, NJ, pp. 47–74.
- Horowitz, R., January 14, 2006. Social and academic issues on the border. In: Presented at the Mexico-Texas Border Literacy Conference. Laredo, TX.
- Horowitz, R., June 14, 2010. Border literacy: how do border crossings affect reading practices and identity of adolescents residing on the U.S. Mexico border?. In: Presented at the XIV World Congress on Comparative Education, Bogazici University, Istanbul, Turkey.
- Horowitz, R., 2012. Border crossing: geographic space and cognitive shifts in adolescent language and literacy practices. In: Romo, H., de la Calleja, C., López, O. (Eds.), *A Bilateral Perspective on México-US Migration*. IETEC-Arana Editores, pp. 147–164.
- Horowitz, R., November, 2013. Crossing the border. How physical space and place contribute to adolescent realignment of language, loyalties, and life aspirations in border language and culture. In: *Translanguage, Transcultural, and Transnational Symposium. WERA Meeting with COMIE*. Mexican Society for Educational Research. Guanajuato, University, Guanajuato, Mexico.
- Horowitz, R., 2017. Adolescent brain wars: adolescent proficiency, performance, and achievement from a competitive global perspective. In: Horowitz, R., Samuels, S.J. (Eds.), *The Achievement Gap in Reading. Complex Causes, Persistent Issues, Possible Solutions*. Routledge Taylor & Francis, New York.
- Horowitz, R. (Ed.), 2023. *The Routledge Handbook of International Research on Writing*. Routledge/Taylor & Francis, New York.
- Horowitz, R., Prior, L., Dart, M.O., Peralta, A., April 13, 2009. Border literacy: highly engaged adolescent learners and readers on the Texas-Mexico Border. In: Presented at the American Educational Research Association Conference. San Diego, CA.
- Hudson, C.B., 2014. *Crossing the United States-Mexico Border: Effects on Literacy Practices and Identity* (Masters thesis submitted to The University of Texas at San Antonio).
- Jiménez, R.T., 2004. Literacy and the identity development of Latina/o students. In: Ruddell, R.B., Unrau, N.J. (Eds.), *Theoretical Models and Processes of Reading*, fifth ed. Newark, DE: International Reading Association, pp. 116–132.

- Jiménez, R., David, S., Pacheco, M., Risko, V.J., Pray, L., Fagan, K., Gonzales, M., 2017. Creating responsive teachers of hispanic and bilingual students learning English. In: Horowitz, R., Samuels, S.J. (Eds.), *The Achievement Gap in Reading. Complex Causes, Persistent Issues, Possible Solutions*. Routledge, Taylor & Francis, New York, pp. 38–56.
- Kennedy, P., Roudometof, V., 2002. *Communities Across Borders: New Immigrants and Transnational Cultures*. Routledge, New York.
- Kleyn, T., 2017. Centering transborder students: perspectives on identity, language, and schooling between the U.S. and Mexico. *Multicult. Perspect.* 19 (2), 56–84.
- Kosmutzky, A., Patty, R., 2016. Transcending borders and traversing boundaries: a systematic review of the literature on transnational, offshore, cross-border and borderless higher education. *J. Stud. Int. Educ.* 20 (1), 8–33 (Is it 2015 or 16).
- Lam, W.S.E., 2009. Multiliteracies on instant messaging in negotiating local, translocal, and transnational affiliations: a case of an adolescent immigrant. *Read. Res. Q.* 44 (4), 377–397.
- Lam, W.S.E., Rosario-Ramos, E., 2009. Multilingual literacies in transnational digitally mediated contexts: an exploratory study of immigrant teens in the United States. *Lang. Educ.* 23 (2), 171–190.
- Levinson, S., 2006. *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge Press, New York.
- Markowitsch, H.J., Welzer, H., 2009. *The Development of Autobiographical Memory*. Psychology Press, New York.
- Nelson, N., Barrera, E.S., Skinner, K., 2008. Border literacy: positionality, mestizaje, and Tejano/a counter discourse. In: Presented at the International Conference on Text, Speech, and Context, Querétaro, Mexico.
- Nelson, N., Barrera, E.S., Skinner, K., Fuentes, A.M., 2016. Language, culture, and border lives: Mestizaje as positionality/Lengua, cultura y vidas de frontera: el mestizaje como posicionalidad. *Cultura y Educación* 28 (1), 1–41. <https://doi.org/10.1080/11356405.2014.980121>.
- Nisan, M., 2010. The Druze in Israel: questions of identity, citizenship, and patriotism. *Middle East J.* 64 (4), 575–596.
- Perez, R., November 2010. Narratives from the other side: the revelations and dynamics of a bi-national penpal program in border spaces. *Child Geogr.* 8 (4), 353–361.
- Riessman, C.K., 2008. *Narrative Methods for the Human Sciences*. Sage Publications, Los Angeles, CA.
- Rodriguez, J.A., Gonzalez, L., Myers, R., Horowitz, R., 2019. Borderland language conflicts: a study of cross-national interactions & conflict resolutions. *UT-SA J. Undergrad. Scholar. works* (6). Retrieved from. <https://research.utsa.edu/ug-Journal/archive/Vol. 6. Html>.
- Saldívar, J.D., 1997. *Border Matters: Remapping Cultural Studies*. University of California Press, Berkeley, CA.
- Sanchez, P., 2007. Cultural authenticity and transnational Latina youth: constructing a meta-narrative across borders. *Ling. Educ.* 18 (3–4), 258–282.
- Stoklosa, K., 2012. Neighborhood relations on the Polish borders: the example of the Polish-German, Polish-Ukrainian, and Polish-Russian border regions. *J. Borderl. Stud.* 27 (3), 245–255.
- Vygotsky, L.S., 1929/1994. The problem of cultural development of the child. In: Van Der Veer, R., Valsiner, J. (Eds.), *The Vygotsky Reader*. Blackwell, Cambridge, MA, pp. 5–72.
- Vygotsky, L.S., 1978. *Mind in Society*. Harvard University Press, Cambridge, MA.
- West, G.B., Tinari, N.M., Zhang, B., March 2019. Youth Constructing Transnational Identities: Mapping Transmodal Resources in Narrative Positioning. American Association of Applied Linguistics, Atlanta, Georgia.

Culture and literacy: situated social literacies

Mastin Prinsloo^a and **Lara-Stephanie Krause-Alzaidi^b**, ^a School of Education, University of Cape Town, Rondebosch, South Africa; and ^b Institute of African Studies, Leipzig University, Germany

© 2023 Elsevier Ltd. All rights reserved.

Great divides: literacy, culture and social development	296
Collapsing the great divide	297
Varying “ways with words”	298
Literacy within socially organized practices	299
Methodologies and research constructs	300
Second phase socio-cultural studies of literacy	301
Local literacies	301
Schools	301
Universities	302
Workplaces and adult-learning	302
Prison, religion	302
Unsettling the local: mobility, migration, multilingualism	303
The turn to multiple modalities and digital literacies	304
Posthumanism, materiality and networks of literacy	304
References	305

Great divides: literacy, culture and social development

The socio-cultural focus in Literacy Studies grew as a field of research in the second half of the 20th century out of conflicting views on what literacy is, including claims and counter-claims about the role and importance of reading and writing in social development, industrialization and the rise of the West. The debates that gave rise to the field took place while modernization theory was influential, which viewed Western, and particularly US American capitalism and industrialization as the model of social development for everybody else (e.g., in [Rostow, 1960](#); [McClelland, 1967](#); [Inkeles and Smith, 1974](#)). Also implicated were human capital theories linking education to economic growth ([Schultz, 1961](#); [Becker, 1964](#)). These theoretical perspectives emerging in the context of Cold War politics sought to identify key “social variables” that could be manipulated to foster development toward modernization. One of the important theoretical questions relating to literacy and language raised by this literature was whether there existed a *literacy divide* which lay at the fault-line between “developed” and “backward” parts of the world and whether it accounted for that divide, to any extent. Literacy was often closely correlated in such research with income, wealth and health at an individual and a societal level. Richer and healthier regions of the world, wealthier nation states, and individuals were seen to have higher levels of literacy than their less prosperous or healthy counterparts. It was then concluded ([Goody and Watt, 1975](#)), and, indeed, is still often concluded, explicitly or implicitly, that literacy itself was the technology that made the difference. The socio-cultural orientation to literacy studies, also known as the New Literacy Studies ([Gee, 1996](#)), grew in critical reaction to these highly generalized and teleological views on literacy and development.

The sources for such “great divide” theories of literacy and of society more generally were popularized in the mid-20th century through the writings, in particular, of anthropologist Jack Goody, social theorist Walter Ong, as well as the psychological arguments of David Olsson. Goody, Ong and Olsson all drew on historian Eric Havelock’s work on the ancient Greeks and literacy, particularly on his *Preface to Plato* ([1963](#)). According to Havelock, the Greeks of around 650–550 BC took the pictographic, ideographic and syllabic writing systems that had been under development from around 3000 BC on the part of the Egyptians, Sumerians and lastly the Phoenicians and transformed them into the alphabetical system of writing, adding letters for vowels to the consonant and syllable signs of earlier script forms. This move allowed, as he claimed, for written symbols to at last closely represent the meaningful sounds (phonemes) of (the Greek) language. He viewed this as a pivotal moment in the rise of the West, as it finally allowed language to be written down clearly and unambiguously, transforming what literacy was and what effects it had, making possible philosophy, historical study and scientific thought for the first time.

[Ong \(1982\)](#), following this example, drew a sharp division between what he called literate cultures and oral cultures. He embraced the Havelockian idea that “writing restructures consciousness” (the title of Ong’s fourth chapter) and is the very source of logic, while speech is transient and evanescent—no sooner is something said than its sounds fade away. Alphabetic writing, in contrast, he argued, provides a clear record that can be thought about and returned to, without it having changed in the interim. The effect on thought and consequently on social organization were profound, he claimed, resulting in a great divide at every level of social organization between oral and literate cultures. [Olson \(1977\)](#) similarly argued that alphabetical literacy initiates a unique and

pivotal development in human cognition, changing language from “utterance” to “text,” where the uncertainties of speech are supplanted by the precision of writing, with language increasingly able to stand as an unambiguous or autonomous representation of meaning. Goody (1975) described literacy as “the technology of the intellect” which set the human mind free from the constraints of the oral and which could only be held back (“restricted”) from fostering universal enlightenment, reason and democracy by certain distorting features of non-western societies, such as social secrecy, caste-based restrictions, rote learning of literacy in religious schools and other “oral residues in a literate culture” (Goody, 1975, 14). In Goody’s (1987) analysis, dynamic literate cultures could be distinguished from conservative oral ones, as well as from societies with what Goody called “restricted” versions of literacy (India, China and in Arab countries, for example): Such restricted literacy, where literacy activity was distorted by the dominance of pre-literate forms of social organization, did not allow literacy’s transformative social potential to do its work in producing modern, progressive and industrializing societies.

Street (1984) identified this thesis on literacy put forward by these “great divide” theorists, as he labeled them, as the “autonomous model of literacy,” due to these claims about the context-independent, or autonomous effects that literacy is said to have. He pointed out that such a sharp division between orality and literacy recycled discredited divisions that had been made in social theory between “modern” and “primitive” societies, placing western societies on the side of progress and other societies in the camp of backwardness and restriction. Indeed, these “great divide” arguments included claims that the advent of alphabetical literacy in the West had marked the move from prehistory to history; from primitive or traditional to civilized or modern societies; from pre-logical to logical and analytical thinking; from aggregative, redundant and “copious” spoken forms to abstract, concise and decontextualized uses of language.

These dichotomous and decontextualized claims by what Goody had called the “literacy thesis” might appear one-sided and biased now (and Goody later revised these claims) but their influence was and remains considerable. Take for example the importance attached to contemporary large-scale standardized literacy testing that continues to be organized, nationally and internationally, or the UNESCO “literacy league tables” still published annually, which claim to compare “literacy levels” between countries, as if something precise, reliable, generalizable and of importance is displayed in these data, something upon which societal decisions can be based—including how investments should be made in education.

Collapsing the great divide

In *Literacy in Theory and Practice* (1984) Street argued that many of the characteristics which great divide scholars attributed to “literate societies” alone were either part of the intellectual framework of any society or the outcomes of complex social processes—not simply effects of literacy in itself. He wrote that literacy was not so transparent and democratized a practice in early Greece as was being claimed. The Greek histories and philosophies were driven by sectional interest, rather than disinterested logic and clarity. The ostensibly “politically neutral” model of literacy put forward by Goody and colleagues relied upon a rhetoric of individual and social developmentalism that celebrated certain western literacy practices as universally normative. Street’s analysis of literacy joined a social analysis of power relations as well as language and literacy ideologies to an orientation to the cultural production of meaning and values in particular settings. Therefore, he argued, consequences that ensue from literacy are neither “neutral” nor effects of literacy on its own but are instead variable. They depend on the nature of the myriad literacy activities that play out in social life and are integral components of larger social practices. The “literacy bits,” Street argued, cannot be studied as if they have effects of their own, separate from the larger social “goings-on.” There were no empirical grounds for assuming an automatic, causal or universal relationship between literacy and social development of various kinds. Rather, different histories of exposure to certain ways of communicating, valuing, reading and writing yielded different forms of reading and writing as practice. The latter evolve and are enacted in contexts involving particular relations and structures of power, values and beliefs. His ideological model of literacy thus sees “literacy” as a shorthand term for practices rooted in social, cultural and political contexts, such that the relationship of writing to language and to thought was different in different contexts. Therefore, no general arguments could be made about the effects of the introduction or presence of writing.

Theories about a “great divide” between literacy and orality, and between literate and oral individuals and cultures, Street argued, reduced the rich diversity of socio-cultural life-forms to a one-dimensional scale where one version of literacy was set as the standard in an ideological and biased way. Street (1984) explained that why he called this counter-position an “ideological” rather than a “cultural model” of literacy was that “culture” had become too static a term and he wished to distinguish his approach from one that saw a world of independent one-language-one-literacy units. Using “ideological” (rather than “cultural”) signaled precisely that there are always contests over the meaning and the use of literacy practices. He saw culture “as a verb,” “an active process of meaning-making and contest over definition, including its own meaning” (Street, 1993, 25).

The origins of Street’s thoughts on literacy lay in the research that he conducted in the 1970s in North East Iran, in fruit-growing villages, especially in the village of Cheshmeh, near the city of Mashhad on the border with Afghanistan. Unlike what he read in Goody and in the arguments of UNESCO and the Iranian government, who were campaigning to remove illiteracy in the villages, Street found various literacy activities in the village, despite outside assumptions about widespread illiteracy in such environments. He saw literacy being used in commercial activities on the part of farmers and traders, in the Islamic religious center or *Maktabs* where the Koran was read and learned, and in children’s schooling. Literacy taught in schools thus turned out to be only one of the literacies that people drew on in the village, and school literacy had a less than dominant role there, because the other literacies

were part of the everyday social practices of influential villagers. The latter were not usually those who had attended school but older men who had attended the Maktab. Koranic literacy instruction is stereotyped for not providing proper literacy because it is simply memorization of passages, but Street found interesting variety and complexity instead. The texts were differently organized on the page compared to western linear writing, the writing was inserted in different forms, angles and in varying relationship with other units of text, so that students learned that reading is not just about language written down, but that organization of text also carried meaning in particular ways.

Despite the complexity of the literacy practices that Street (1984) identified in Cheshmeh, outsiders, including literacy campaigners, assumed villagers were illiterate and living in the dark; and that without the literacy of standardized instruction, they would remain cognitively, socially and culturally deficient. Street instead noted that the people in the village might well have failed tests for literacy, developed with schooling models in mind. Yet, they were using reading and writing on a regular basis in situated and locally appropriate ways. Standards for literacy were a function of local practices, and a test run by outsiders would not reflect that. "If these complex variations in literacy which were happening in one small locale were characterized by outside agencies—State education, UNESCO, literacy campaigns—as 'illiterate', might this also be the case in other situations too?" Street (2001, 6) asked, concluding that the then dominant assumptions about literacy, both in development circles and in academia, were wrong. They treated literacy as unitary—an independent variable that could be applied with predictable effects across contexts with consequences for social and economic progress, cognitive development and democracy. For example, modernization economists attempting to quantify literacy's developmental capacity had claimed that a 40% literacy level was necessary for economic take-off, and 80% for rapid industrialization (Street, 2001). Street identified the prevalence of what he called *scriptism*—a view of the influence of writing on the conceptualization of speech—a belief in the superiority in various respects of written over spoken language and the view that some uses of language are more "context-dependent" or "objective" than others. He stressed the importance of going beyond a narrow focus on provision and instruction in literacy and language research and in developmental planning. Ignoring the situated and variable nature of language and literacy practices has led to a fundamentally flawed set of assumptions about language, literacy and society in much of the developmental literature, he insisted.

Varying "ways with words"

A second strong impetus for conceptualizing literacy as variable socio-cultural practice in local settings came from Heath's (1983) widely read study of south-eastern communities in the USA. Heath (1982, 1983) questioned why Black students were failing in the recently desegregated schools in Southern Carolina and carried out ethnographic research into the literacy, language and learning in three neighboring communities. She was informed by a socio-linguistic orientation to language, particularly by Dell Hymes' approach to researching language and communication through the ethnographic study of specific speech communities. Hymes' work contrasted with Chomsky's (1975) focus on language competence as a biologically evolved capacity of humans, best studied in abstraction from social context and individual performance. Hymes, instead, emphasized the social and cultural dimensions of language, insisting that language should be studied with regard to how meaning was produced through social interaction in specific contexts, and how what counted as *communicative competence* was specific to particular speech communities. Intelligibility was for him not only a complex function of features of linguistic form (phonological, lexical, syntactic), but also of norms of interaction and conduct in conversation and in writing. He suggested that "one think of a community (or any group, or person) in terms, not of a single language, but of a repertoire," comprising a set of ways of speaking with "speech styles, on the one hand, and contexts of discourse, on the other, together with the relations of appropriateness obtaining between styles and contexts" (Hymes, 1996, 33). Membership in a speech community, he argued, consists in sharing one or more of its ways of speaking—that is, not in knowledge of a speech style (or any other purely linguistic entity, such as a language) alone, but in knowledge of appropriate use as well.

Heath (1983) shared Hymes' concern for the communicative dimensions of linguistic interaction. She wanted to understand the socio-historical context that shaped distinct communicative orientations, as well as to study how literacy was part of such communicative practices. She found that children in three socio-culturally divergent settings of the town (Roadville, Trackton and "the townspeople"), while being socialized within their local communities, were acquiring particular ways of using language and narrative in social communication which were incompatible with those of schooling in specific ways. As babies, children in Trackton (a working-class black community) were almost always held in their waking hours and were constantly in the midst of ongoing human communication, verbal and non-verbal. Adults did not engage children as conversational partners in "baby talk," where they simplify their language, as was practice among the townspeople. Trackton adults did not sit and read to children or direct questions at pre-verbal children (such as "Why are you crying?"). Adults chose verbally competent conversational partners (other adults or older children) and talked *about* the infant to them. Children were discouraged from answering the questions of outsiders and were rarely asked "What is X?" questions (Heath, 2001, 320). Instead, they were asked analogical questions which called for non-specific comparisons of one item, event, or person with another (e.g., "What's that like?") (Heath, 2001, 333). Though children could answer such questions, they rarely named the specific feature(s) which made two items or events alike. They learned to tell stories by rendering a context and inviting the audience's participation to join in the imaginative creation of a story. Children were encouraged to draw attention to themselves and to entertain by way of the display of particular kinds of associative verbal dexterity. Group negotiation and participation was a prevalent feature of the narrative performances of adults and children.

Heath reported that in school, Trackton children commonly had trouble adapting to the social interactional rules for literacy events. They did not respond well to direct questions, as there was no precedence for this interactive pattern in their home settings. Their teachers ignored their developed abilities to metaphorically link events and situations and recreate scenes. After the elementary years, when their imaginative skills and verbal dexterity could really pay off, they had often failed to gain the necessary written composition skills needed to translate their analogical skills into a channel teachers could accept.

The working-class white children in Roadville, whose families had been mill-workers in the region for several generations, were brought up into ways of knowing that were very different from those of the Trackton children, but also incompatible with the communicative forms and practices that school teachers took to be the norm. This was in spite of the abundance of reading materials in the neighborhood homes, in the form of magazines, newspapers and books. Roadville was a religiously conservative (Christian) community whose children reportedly struggled at school despite the high value attached to reading and writing. Heath traced in fine-grain ethnographic detail the communicative practices into which they were raised instead. Children were read to from books that emphasized nursery rhymes, alphabet learning and simplified bible stories. They were encouraged to find a clear moral message in each account they read or gave. They were discouraged from elaborating on accounts, giving perspectival or emotional readings, or fictionalizing real events, a practice associated with lying among parents who saw the Bible, for example, as containing literal meaning and guidance for everyday living. They were good at well-ordered practices, involving the keeping of time and space limits, and good at delivering factual recounts, as opposed to interactive and/or intertextual accounts. They struggled with the kinds of evaluative, decontextualizing and recontextualizing narrative practices valued in school, and therefore generally struggled to do well.

In contrast, Heath (1983) suggested that the way children of both White and Black “townspeople” grew up talking, communicating and relating to narrative and print in their middle-class homes matched well the practices of school and the forms that literacy took there. From an infant age the children encountered the question formats that characterize so much of school interactive discourse (those of initiation-reply-evaluation sequences, for example), labeling habits, fictionalizing practices, practices of relating book knowledge to real life experience, and such interactive behavior as learning to listen and wait in response to particular cues from their parents and carers. They learned patterns of interaction in relation to texts which prepared them for those of schooling. They were encouraged to take up positions in relation to the texts they were responding to, to present “accounts” rather than “recounts.”

Heath concluded that what counted in effective communication was not a generalized competence (e.g., being able to “speak English” or “code and decode letters”) but a situated, communicative competence embedded in acquired, “deep” cultural knowledge and learned models of using language and making meaning in specific ways. Heath’s work made the same case as Street that there are multiple ways of taking and making meaning in reading and writing practices. The selection of but one way as normative in school and in formal institutions, meant that people whose ways were different from the norm had to continuously struggle to adapt to that standard. She criticized the “great divide” distinction between literacy and orality because it placed undue importance on the medium of communication at the expense of its social purpose.

Heath (1983, 2001) suggested that teachers should learn more about the communicative practices and traditions that children from different socio-cultural, linguistic and socio-economic backgrounds brought to school contexts, and to treat children’s language and social behavior as manifestations of the rules of communicative competence they had acquired successfully in their communities. She argued that dichotomization between oral and literate traditions was “a construct of researchers, not an accurate portrayal of reality across cultures” (Heath, 2001, 339). Her study instead showed the culturally diverse ways of acquiring knowledge and developing cognitive styles. Trackton and Roadville children were “disabled” in the schooling context not because of any general, individualized cognitive deficit but because their practices of everyday life had initiated them into differing sets of expectations regarding intimates, non-intimates, and how they relate to one another—in speech acts as well as in reading and writing activities.

Literacy within socially organized practices

A third influence on the socio-cultural, or social practices model of literacy was Sylvia Scribner and Michael Cole’s ground-breaking research (1978; 1981) of literacy and cognition in Liberia, West Africa which found that cognitive skills associated with literacy varied dramatically in relation to the wider social practices within which literacy was embedded.

Scribner and Cole were among the co-editors of the influential English language version of Vygotsky’s *Mind in society: The development of higher psychological processes* (Vygotsky, 1978). Concerned to elaborate on the social dimensions of cognition, they drew on Vygotsky’s claims that interpersonal/inter-mental (or social) processes are the necessary condition for the emergence of individual/intra-mental (psychological) processes. Luria (1979), a contemporary and colleague of Vygotsky, had carried out a study of the cognitive effects of literacy acquisition by running tests among persons living in remote parts of Uzbekistan in Central Asia, some of whom had learned to read and write and some who had not. Luria reported significant differences in response to tests of cognitive processes across the two samples and concluded that these differences were a direct outcome of whether or not the research subjects had learned to read and write.

Luria did not, however, distinguish between the cognitive effects of literacy and other possible causes of cognitive change, such as participation in schooling practices. Scribner and Cole (1981), in contrast, were concerned to distinguish between the effects of literacy as distinct from the effects of schooling and other social practices. Their research team studied in detail the cognitive and social consequences of literacy in a different setting—Liberia—where three different scripts and literacy traditions were present:

school literacy in English, a religious literacy in Arabic script and an indigenous script used by individuals for letter writing and record keeping in the local language. Some of the local Vai people were able to read and write in one, some in two of these scripts, and some were unable to read or write in any of them.

Over a period of 4 years and extended follow-up work, Scribner and Cole, with research assistants, gathered ethnographic and survey-based descriptions of language and literacy use and also ran a battery of tests of Vai persons' cognitive, perceptual and conceptual processes, including tests for abstraction, memorization, categorization and verbal explanation skills, so as to study the uses and consequences of literacy in these three different scripts, languages and contexts. They were able to distinguish in their analysis between "literacy effects" and "school effects," and showed that the cognitive attributes previously associated with literacy (by Luria and others) were not products of literacy itself but variable outcomes of particular social practices like schooling or urban living.

The failure of literacy to yield consistent cognitive effects across all three scripts and literacies (and the inconsistency of schooling effects on measured cognitive outcomes) led Scribner and Cole to conclude that "schooling and literacy are not synonymous" and that "literacies are in fact highly differentiated" (1981: 132). This made them question the tendency by many writers to "discuss literacy and its social and psychological implications as though literacy entails the same knowledge and skills whenever [and wherever] people read and write" (1981: 132). They found that cognitive skills associated with literacy varied dramatically depending on whether people's literacy experiences were school, religious or community activities. Scribner and Cole made the case that there was nothing special about literacy in general as regards its consequences for cognitive skills. Rather, as had been argued by Vygotsky (1978), cognitive processes had a direct basis in socio-cultural activity.

They thus challenged "great divide" claims that literacy acquisition produced certain cognitive changes, regardless of the context of learning. They made the important argument that literacy is not a general technology that is the same thing with the same consequences regardless of what the contexts of its acquisition might be. Instead, they claimed that literacy was always constituted within socially organized practices.

Methodologies and research constructs

Gee (1999, 1996, 2003) provided detailed examples of what literacy and language looked like from a discourse-oriented socio-cultural perspective. His examples show that an interpretation of any utterance or text always involves an interplay of linguistic, semiotic and social frames of interpretation. The sources of meanings in a text then don't lie in the word itself but in the social matrix within which a piece of writing is produced and understood. He made clear that reading involves the making of meaning with available resources, and such meaning is not made or taken mechanically one word at a time, but rather with a sense of contextual specificity drawn from the wider text and context that is being produced, invoked or interpreted.

Take something so simple as the following sentence: "The guard dribbled down court, held up two fingers, and passed to the open man" (Gee, 2003, 29). While one might seem to know every word in the sentence, you can't make sense of the piece without knowing in detail about basketball as a game, that is, as a particular sort of social practice. Gee showed that the same applies to any sentence or text. Gee's point is that what is important in understanding and meaning-making is not words (oral or written) themselves, but the larger and specific sociocultural coordinations in which they gain their significance.

Key terms in the study of literacy as situated, socio-cultural practice have been "literacy event" and "literacy practices." Literacy events are moments when "written language is integral to the nature of participants' interactions and their interpretive processes and strategies" (Heath, 1982, 93) and literacy practices are the more general sociocultural framing that gives significance to particular acts. Literacy events, in Heath's conception, included those moments when inscription or decoding of text is featured—but not necessarily centrally. Central was instead the configuration of action, talk and text in multiple and socially varying ways. "Events are observable episodes which arise from practices and are shaped by them. The notion of events stresses the situated nature of literacy, that it always exists in a social context" (Barton et al., 2000, 8).

The concept of literacy practices incorporates literacy events as empirical occasions to which literacy is integral and analyses them in terms of the models or preconceptions that make people decide who does what, where and when, as far as reading and writing is concerned. Epistemologically, the concept of literacy as a social practice provides the frame for an analysis of meaning making. Methodologically, the approach has been grounded in linguistic ethnography and has drawn on discourse analysis, socio-cultural models of cognition and various strands of sociolinguistics and social theory for its analytical work. Widely referred to as The New Literacy Studies (Gee, 1996) the approach has influenced at least four generations of researchers (Baynham and Prinsloo, 2009). Typically, researchers have observed or recorded particular literacy events at their site of research and then tried to understand the wider discursive framings and social practices that produce such events in their particular form and shape. This "eventness" as a feature of literacy practices has partially been disturbed by later work, however, which has questioned its application to translocally-networked writing activities such as online, digital or virtual, literacy-linked communicative activity and other kinds of electronically enabled social activity. The construct of practices has similarly received critical attention for being so all-encompassing that it lacks analytical precision as it is referring to both routinized social activities as well as reflexive adaptations, revisions and meshings of such routines and genres (Baynham and Prinsloo, 2009).

Second phase socio-cultural studies of literacy

The criticisms of “great divide” theories of literacy in these early shaping studies described above, and in their development of an alternative focus on literacy as variable social practice, were supported by detailed research studies that accompanied and followed them, from various parts of the world. Such research included localized ethnographic studies, many of them of non-Western, non-middle-class communities and localities, as in [Scollon and Scollon \(1981\)](#), [Finnegan \(1977\)](#), [Kulick and Stroud \(1993\)](#), [Besnier \(1995\)](#), [Prinsloo and Breier \(1996\)](#), and [Barton and Hamilton \(1998\)](#), among others. Besnier described how letter writing among a group of South-Pacific islanders was culturally marked out as a genre resource for communicating intense and effusive emotional regard for often distant relatives, friends and intimates whereas such displays were not culturally appropriate in face-to-face spoken communication. Kulick and Stroud described how engagements with literacy in a Papua New Guinea village were shaped by the high social value given to humility and “indirectness” in communication over explicitness, resulting in various kinds of novel uses for literacy. These uses would appear as unusual or perplexing to those without insider knowledge of how meanings were constructed from indirect clues as to the intent of individual writers, rather than from explicit statements. In both cases the established literacy practices of these communities contradicted the views of literacy as a “technology of the intellect,” in [Goody’s \(1975, 1\)](#) words, that operated as a culturally-neutral and transparent conduit of meaning and was displayed, at its best, in essay form, where the values of “clarity, brevity and sincerity” held sway ([Scollon and Scollon, 1981](#)).

Taking a similar, localized, ethnographic research approach, [Barton and Hamilton \(1998\)](#) observed community members in Lancaster, England and asked them to reflect on their literacy practices. Echoing Heath’s work, they showed that as important as family practices are for children’s literacy development, these practices take place in larger community contexts that influence family activities. The researchers drew a distinction between dominant (institutionalized) and vernacular (self-generated, everyday) literacies. Vernacular knowledge was seen to be local, procedural and minutely detailed. Literacy was not an explicit focus of everyday activities but literacy elements were implicit in most activities and were used to get things done, including the learning of a martial art, paying the bills, organizing a musical event or finding out about local news. When questioned about them, people did not always regard their vernacular literacies as real reading or real writing as they were embedded in other activities, like shopping, writing to a relative, paying an invoice or applying for something. They therefore did not carry the same status as more conventionally recognized literacy activities like the reading of literature or “school literacy” activity. Indeed, some vernacular literacies were deliberately hidden, because they were private or oppositional, including secret notes and love letters, comics and fanzines. The researchers concluded that much talk in everyday life that they studied was in fact talk about texts or shaped by documents or textual practices and that reading, writing and writing artifacts were very much part of the “glue” of social life in these communities.

These and similar studies gave evidence of cross-cultural differences with regard to uses of literacy and also showed that reading and writing differed across different activities and contexts among the same groups of people. Even within the same site, at work or at the home, there can also be different meanings of reading, writing and literacy, because people participate in a range of contexts, or textually mediated social worlds, within that setting, that assume differing forms of social interaction.

Local literacies

In the wider research, such cross-cultural differences were made visible across various institutional contexts, including schools, workplaces, religious settings, prisons, and in a range of home, community and leisure contexts:

Schools

As regards literacy in educational settings, influenced by the earlier work, a wider number of studies examined the ways of being, knowing, valuing, acting, speaking and attitudes to writing that children developed in their home and other out-of-school environments and which met up with the discourses and practices of schooling ([Street and Street, 1991](#); [Cook-Gumperz, 2006](#); [Gee, 1996](#); [Dyson, 1993](#); [Freebody and Welch, 1993](#); [Wagner, 1993](#); [Maybin, 1994](#); [Street, 1995](#); [Gregory, 1996](#); [Lankshear, 1997](#), among many others). Drawing variously on Bakhtin, Bourdieu, Vygotsky, Hymes, Giddens and other social theorists, they pointed to the significant variety and cultural specificity of the early literacy practices of children across different social contexts. Those studies also showed that the ways meaning was shaped through text in school settings was more compatible with the way some children had learned to make meaning, and less compatible with the “ways of knowing” of children from other backgrounds. Schooling was also producing a narrow and rigid definition of literacy as severed from social dialog, and as a “decontextualized knowledge validated through text performances,” ([Cook-Gumperz, 2006, 27](#)). Schools and adult literacy classes taught literacy as a skill which was primarily about coding and de-coding rather than teaching how to construct meaning through texts in those particular ways that were favored in schooling and other educational settings. Another central finding was that misinterpretations of what children were doing and misrecognition of what communicative resources they were bringing to schooling, on the part of administrators and teachers, tended to happen with reference to models of cultural deficit, where differences in children’s ways were then automatically understood as lacks or absences.

Universities

A further direction of work developed under the rubric of “academic literacies” in university and tertiary education contexts (e.g., Ivanic, 1998; Lea and Street, 1998; Canagarajah, 2002; Lillis, 2003, 2006; Thesen and van Pletzen, 2006; Jacobs, 2010). These scholars drew attention to power relations and identity processes around student writing in relation to institutional practices, highlighted the diversity of writing practices and genres across disciplines and recognized that students bring their histories, identifications, resources and commitments to writing. The academic literacies approach sees writing and text production as socially situated practices and activities of diverse kinds and looks at extending the range of semiotic resources—linguistic, rhetorical and technological—that are allowed and valued in contemporary universities. This orientation sees that the complexities shaping writing as an activity can be productively engaged with by university teachers rather than having them treat particular examples of student writing as evidence of some lack or deficiency. It also shows university practices to be malleable as they can be reshaped to include and capitalize on the particularities of local and regional contexts and the differences that students bring to institutions of higher learning.

Workplaces and adult-learning

Looking at workplace settings, other Literacy Studies researchers showed that workplace skills, including print-related skills, develop and flourish or remain undeveloped, as a result of social rather than purely technical processes (Gee et al., 1996; Farrell, 1997, 1999; Scheeres, 2007; Scholtz and Prinsloo, 2001). Learning is structured and defined by social relations and social practices taking place on shop floors, training sites, and other work contexts, whereas “stand-alone” literacy classes that focused on coding and reading skills remained disconnected from the actual uses of reading and writing in work settings. Researchers showed that work-related literacy practices, for both low-level and high-level workers, required them to draw on repertoires of situated ways of using reading, writing, language and other semiotic resources in order to carry out a work-related activity. Hull (1997) noted in a list of observed “literacy functions” in factory settings, that very few fell into the category of “basic” (i.e., relatively simple self-contained tasks like copying, labeling, keyboarding, tallying). The continuum of literacy that can be observed expands to include categories in which the purposes of literacy are first more complex—using literacy to explain, taking part in discourse around texts, participating in the flow of information, problem solving—and then also more obviously connected with issues of power—using literacy in the exercise of critical judgment, using literacy to acknowledge, exercise or resist authority.

Prinsloo and Breier (1996) in several studies in South Africa on the social uses of literacy in situated community and work settings, found that the literacy taught in adult education classes that they observed was not related to the literacy practices that people engaged with outside of the classes. Accordingly, they questioned the value of a pedagogy which focused on the transfer of disembodied coding and decoding skills. The adult literacy night schools were seen to promote school-like pedagogical practices which were both new to the adult learners (since they had little experience of schooling) and had little connection to the actual literacy activities in their lives.

Prison, religion

Wilson’s (2000) study of literacy in prisons similarly made the case that literacies were multiple and context-dependent. Literacy/ies in prison could not be comprehensively studied without taking into account the social circumstances of which they were a part. Her findings pushed toward a more contextualized and situated view of multiple literacies and away from prescriptive instructional models, such as the “school” perspective, the “training” perspective or the “standards” perspective. She identified various distinct literacy activities that inmates were regularly engaging in, for example frequent signature writing, official bureaucracy on reception into and release from the jail, graffiti, letters to family and friends, subversive notes to acquaintances in the jail, reading court depositions, reading official prison documentation, commissary forms, complaints forms, visits requests, tattoos, appeals, poetry, educational pursuit, books and magazines.

Kapizske (1995) studied the textual politics and practices of a 7th day Adventist community in North Australia, focusing on processes of theological (church), familial (home) and educational (school) normalization of community members into regulated ways of hearing and speaking, reading and writing, being and believing, which constructed and constrained what could and could not be articulated and enacted by believers. Once again, the “literacy bits” could not be sensibly studied or made sense of separately from the larger social “goings-on” in which they were embedded.

Examples from a wider literature that accompanied and followed these studies include Papen’s (2005) study of tourism, governmentality and literacy in Namibia; Robinson-Pant’s (2001) account of literacy and development among women in Nepal which focuses on the processes by which women in Nepal acquire literacy and deploy its use for their own purposes; Kalman’s (1999) study of mediated literacy practices in Mexico City; Maddox and Esposito’s (2012) research around literacy inequalities and social distance in Nepal; Achen and Openjuru’s (2012) research on language and literacy as globalized practices in the poorer residential areas of Kampala, Uganda; Pahl and Rowsell’s application of these insights to classroom work (2012); Kell’s (2008) study of literacy and housing disputes near Cape Town; and Prinsloo and Stein’s studies of early literacy in homes and schools in Cape Town and Gauteng, South Africa (Stein and Slonimsky, 2006; Stein and Mamabalo, 2005; Prinsloo and Stein, 2004; Prinsloo, 2004).

These studies, show, among other things, that people labeled “illiterate” within older models of literacy may, from a more culturally sensitive viewpoint, be seen to make significant use of literacy practices for specific purposes and in specific contexts. Accordingly, the boundary between literate/non literate is less obvious than individual “measures” of literacy and testing might suggest.

Unsettling the local: mobility, migration, multilingualism

Though it is not a new phenomenon by any means, migration or the movement of people between regions and localities has attracted considerable attention since the 1980s due to social, technological, and geopolitical developments around “globalization” phenomena. People and literacy resources, along with languages and other semiotic resources are more frequently seen as border-crossing phenomena than in the past, signaled by the visibility of large-scale migrations and the increasingly multi-ethnic and multilingual nature of urban communities in major centers around the world. Such mobilities contributed to raising the question of literacy in relation to language and particularly to multilingualism that was taken up in a number of early studies and has become a topic of growing importance and attention.

Baynham (1995) examined the way that Moroccan migrants in London shifted between the communicative modes of text and talk in social interaction among each other, while assisting each other with language and literacy challenges. Bilingual talk around monolingual text in school and community settings is, indeed, characteristic of most multilingual social contexts, as was seen in Gregory and Williams (2000), a study of literacy based on long term ethnographic engagement within the Bangladeshi settlement in East London, UK. Martin-Jones and Jones (2000) presented multiple studies of ways in which languages and literacies shaped, and at times perpetuated, asymmetrical power relations, especially between minority language groups and dominant language and literacy practices for migrants in host countries. The studies ranged from local life worlds to formal education and from community based to institutionalized literacy practices, providing systematic descriptions of literacy and language practices across local language groups in schools, institutional and bureaucratic settings, homes and local communities. The authors called these practices multilingual, rather than bilingual, to draw attention to the multiple social languages and the range of repertoires, communicative purposes, paths, and ways in which people acquired and developed literacies in these settings. While these studies often centered around the experiences of migrants, they retained a territorialized focus on local communities. But the heightening awareness of movement inevitably raised questions for the local character of studies of literacy as socio-cultural practices.

Importantly, in earlier days both Street and Hymes had already strongly rejected a notion of discrete cultural communities with stable languages and literacies. They were highly critical of “a ‘Herderian’ conception, from 19th Century Europe, of the world as composed of individual language-and-culture units” (Hymes, 1974, 59). Street’s study of village literacies in Iran was indeed provoked by an awareness of the tensions produced by the interface of more localized and everyday literacies, on the one hand and national and more bureaucratized literacies, on the other. So criticisms of an excessively local focus in earlier socio-cultural studies can certainly be overstated, but they did help in drawing more widespread attention to fluid and changing ways with texts and language. Brandt and Clinton (2002) for example, in a paper titled “The Limits of the Local,” criticized the “local-community” focus in earlier literacy studies. Baynham and Prinsloo argued for considering “the transcontextualized and transcontextualizing potentials of literacy” (Baynham and Prinsloo, 2009, 4), suggesting that researchers focus on locally situated literacy and language practices “with a sense of how remote sites and remote literate practices shape and constrain local literacy practices,” but also of how wider practices and routines were expressed and refigured in more local contexts. Blommaert (2010, 2008) and Kell (2015, 2011) discussed processes of recontextualization where texts move out of settings of situated practices into broader, more bureaucratic or formalized settings, with particular and varying effects. Warriner (2007), Lam and Warriner (2012), and North (2018) among others examined the transnational nature of migrant workers’ lives and the complex ways in which literacy is threaded through their social and material practices.

These studies point to a more complex conceptualization of context and of the relationship between the local and less-local (or global) within literacy research. Such re-conceptualization draws on Massey’s (2005) idea of context as both locally and socially produced and simultaneously and complexly connected, involving processes of cultural production shaped by social actors from disparate locations and where standardized practices from larger and more bureaucratic spheres encounter localized social practices that reshape these more global forms. In socio- and applied linguistics language is being rethought of more as social, material and multimodal practices (Heller, 2007) and not just as sets of disembodied “linguistic resources.” This development draws inspiration from the way that literacy has been rethought as practices since the 1980s.

Krause and Prinsloo (Krause and Prinsloo, 2016; Prinsloo and Krause, 2019; Krause, 2021) examine the notion of fluid languaging in classroom talk that is neither Standard English nor Standard Xhosa—the two official languages formally expected in the primary school in Khayelitsha, the biggest black township in Cape Town, South Africa. Here, everyday languaging practices are heterogeneous and often fluid. Speakers assemble linguistic resources without restrictions by the limits of one or more named language. Despite their fluid linguistic skills, purified, standardized Xhosa or English is mostly not a comfortable element in their repertoire and is not a local community resource in Khayelitsha. Because their fluid urban languaging resourcefulness does not help students when they encounter writing practices through standard monolingual English, they are at a disadvantage when they test for all their subjects in Standard English examinations. Teachers in turn are under pressure from the Education Department to produce high pass rates and adopt strategies to help students by coaching them on the specifics of the test. The arrival of this centrally distributed paper at the school thus induces a series of creative, localized re-shaping processes, involving translating and exam practice activities, aimed at making it (barely) manageable for most learners to pass and by extension to not get their teachers into trouble with educational authorities, who do not see the linguistic complexities they have to handle in their classrooms.

Studies like these manage to show how language and literacy practices are always nested within other socio-cultural and material practices, some new and some old, forming a nexus of practices (Scollon and Scollon, 2007), a configuration or assemblage of tools and actions with various conventions and histories associated with them which come together to form recognizable sequences of actions. To facilitate the study of such complex nexuses or assemblages, much methodological and conceptual work is under way. One innovative idea has been what Bartlett and Vavrus (2014) called a “vertical case study approach.” In this approach, a localized focus is replaced or accompanied by attention to the wider activities, practices and “goings-on” that shape the local. It still remains, however, a socio-cultural approach, committed to understanding literacy and language practices, and the social policies which shape them in institutions like schools, as “a deeply political process of cultural production engaged in and shaped by social actors in disparate locations who exert incongruent amounts of influence” (Bartlett and Vavrus, 2014: 132). It involves a translocal methodology that rejects the selection of clearly bounded research and treats aspects of school literacy practices that are shaped by wider socio-political dynamics, including the activities of curriculum boards, examination units and other units, as sites of shifting cultural production by actors through social practices.

The turn to multiple modalities and digital literacies

Under conditions of increasing online writing practices since the later 1980s researchers began to note how (sometimes moving) images and sound accompanied written texts to a degree seldom focused on before. Multimodality became a central topic in Literacy Studies from the 1990s, shaped, firstly, by Gunther Kress’s work. As Gee (2001) described it, distinct sites of social practice or “semiotic domains,” as he named them, including first person-shooter-video games, advertising, cellular biology, midwifery and surgery incorporated multiple modalities (e.g., oral or written languages, images, equations, symbols, sounds, gestures, graphs, artifacts, and so forth) to communicate distinctive types of messages. Multimodal, trans-semiotic literacy and language studies have grown exponentially since. For want of space and because they are dealt with in detail elsewhere in this Encyclopedia, we will move on here (but see also Prinsloo and Baynham, 2013, vol. 3 for a selection from the digital literacies literature).

Posthumanism, materiality and networks of literacy

Among the most recent developments in the socio-cultural study of literacy have been the influences of “new materialism” and post-humanist studies. We do not have space to discuss this new direction in detail here and it is covered elsewhere in this Encyclopedia. So we will limit our remarks here to some points about how these new directions relate to and contrast with earlier literacy studies research. In brief, these new directions are responses to what have been identified as further “great divides” to those of literacy/orality, literate culture/oral culture, as analyzed and deconstructed in the earlier days of the socio-cultural approach to the study of literacy (Street, 1984). Latour (2005) identified these further and perhaps deeper divides as those between nature and society, between human and non-human and between knowledge and being. Questioning the validity of these long-held dualisms, contemporary research that draws on posthumanist and new materialist theorizations attempts to undo related divides, between meaning and materiality, macro and micro, social and technical, along with nature and culture (Law, 2009, 147). To point to some research examples:

Leander and Boldt (2013) follow a 10-year-old boy through one day as he engages in reading and playing with text from Japanese manga, focusing on the sensations and movements of the body in the moment-by-moment unfolding or emergence of activity. This nonrepresentational approach describes literacy-related activity not as projected toward some textual end point but as living its life in the ongoing present, forming relations and connections across signs, objects, and bodies. They see such activity as saturated with affect and emotion and contrast the child’s animated literacy-filled play with his disinterested underperformance in school literacy, where his engagement is not invoked. Leander and Burris (2020) are concerned how human lives are entwined with computers and computational systems. They quote a study of student reading and how integral algorithms are to modern texts and to constructing readers: “What we read and how we read, and, more importantly, how we are conditioned to imagine ourselves as readers, is increasing determined by algorithms that operate underneath of the surface of texts.” They invoke the image of a world where “technologies are active agents, recruiting and enrolling humans and where subjectivity and agency are not merely given in advance, but are relational achievements involving people and things.”

Lindgren (2019, 6) examines children’s learning and development by “paying attention to the myriad of relationships going on between children, materials and environments in preschool.” She finds that children constantly seem to be preoccupied with establishing relationships with their environment, as well with other people, with books, pens, paints, clay, water and paper. She describes children as emergent in a relational field, where non-human forces are equally at play in constituting children’s becomings and for her “the child figures as the embodiment of flexibility and transformation through its multiple becomings with the world.” Similarly, Hackett and Somerville (2017) focus on literacy, bodies and materiality, focusing on how children engage with text on a sensual level, foregrounding “non-representational aspects of language and communication.” They focus on movement and sound in their study “as more-than-human practices through which the world forms itself, and young children’s sounding and movement happen in relation to this” (2017, 377).

Young children's literacies are seen not only as embodied sensory experiences but embedded in and inseparable from their entanglement with the world. Zapata et al. (2018) start from the perspective that knowing, being, becoming, and doing are co-constituted in a world of lively relations between humans and nonhumans. They focus on (a) writing as intra-activity; (b) writing as trans-corporeal, porous, and lively; and (c) writing as translanguaging assemblages of humans and non-humans, attending to "the entanglements of writing and its multiple productions rather than focusing on human pursuits alone" (2018, 479). They see this rethinking of writing as a matter of ethics, recasting subjectivities, where the "subject does not exist ahead of or outside language [and materials] but is a dynamic, unstable effect of language/discourse and cultural practice."

These studies describe themselves as being energized by the ethical orientation at the heart of the materialist shift. But in the work done so far ethics comes to replace socio-cultural and political analysis, with a result that the politics involved get obscured. This is a worry that Kuby and Rowsell (2017, 288) encounter in responses to their work: "What about power?" or "What about social justice and issues of race and/or economic inequities?" They respond by seeing "much hopefulness in posthumanism and are energized by the ethical orientation at the heart of this work." But how such an individualized and voluntaristic ethics deals with complex political concerns is an unanswered question.

The problem here seems to be that the focus on human/material loses sight of the diversity of humans and their situatedness with regard to the phenomena in question. The reappraisal of matter intended by the new materialism and post-humanism aims to provide a different account of agency and ontology, that fundamentally affects socio-cultural and political categories and concepts. But, in its application, if it loses its hold and attention to situated socio-cultural and political specificity, it loses a very important edge. While laudably critical of the anthropocentrism of preceding approaches, the human-material focus also needs to ask how *particular versions of the human* come into being within phenomena, rather than assuming a generic construct of "the human" in sensuous engagement with vibrant material. It needs to continue asking questions about causes and consequences of socio-economic and political inequalities, questions about the making of humanity as a category of practice—across lines of race, coloniality, migrants and borders, gender, class, commodification and place, to name just some forms of boundary-making and discrimination, attention to which is at risk of being lost, at least in the first wave of posthumanist research. The Literacy Studies focus on situated socio-cultural practices points to a route back to those concerns.

References

- Achen, S., Openjuru, G., 2012. Hollywood in Uganda: local appropriation of trans-national English-language movies. *Lang. Educ.* 26 (4), 363–376.
- Bartlett, L., Vavrus, F., 2014. Transversing the vertical case study: a methodological approach to studies of educational policy as practice. *Anthropol. Educ. Q.* 45 (2), 131–147.
- Barton, D., Hamilton, M., 1998. *Local Literacies: Reading and Writing in One Community*. Routledge, London.
- Barton, D., Hamilton, M., Ivanic, R. (Eds.), 2000. *Situated Literacies: Reading and Writing in Context*. Routledge, London.
- Baynham, M., Prinsloo, M. (Eds.), 2009. *The Future of Literacy Studies*. Palgrave Macmillan, Basingstoke.
- Baynham, M., 1995. *Literacy Practices*. Longman, London.
- Becker, G., 1964. *Human Capital*. Columbia University Press, New York.
- Besnier, N., 1995. *Literacy, Emotion and Authority: Reading and Writing on a Polynesian Atoll*. Cambridge University Press, Cambridge, England.
- Blommaert, J., 2008. *Grassroots Literacy: Writing, Identity and Voice in Central Africa*. Routledge, London.
- Blommaert, J., 2010. *The Sociolinguistics of Globalization*. Cambridge University Press, Cambridge.
- Brandt, D., Clinton, K., 2002. Limits of the local: expanding perspectives on literacy as a social practice. *J. Lit. Res.* 34 (3), 337–356.
- Canagarajah, A., 2002. Multilingual writers and the academic community: towards a critical relationship. *J. Engl. Acad. Purp.* 1, 29–44.
- Chomsky, N., 1975. *Reflections on Language*. Pantheon Books, New York.
- Cook-Gumperz, J., 2006. The social construction of literacy. In: Cook-Gumperz, J. (Ed.), *The Social Construction of Literacy*, second ed. Cambridge University Press, Cambridge, pp. 1–18.
- Dyson, A., 1993. *Social Worlds of Children Learning to Write in an Urban Primary School*. Teacher's College Press, New York.
- Farrell, L., 1997. Literacy, schooling and the workplace revolution. *Engl. Aust.* (119–120), 56–62.
- Farrell, L., 1999. Reconstructing Sally: narratives and counter-narratives around work education and workplace restructure. *Lit. Numeracy Stud.* 9 (1), 5–26.
- Finnegan, R., 1977. *Literacy and Orality*. Blackwell, Oxford.
- Freebody, P., Welch, A.R. (Eds.), 1993. *Knowledge, Culture, and Power: International Perspectives on Literacy Policies and Practices*. Falmer Press, London.
- Gee, J., Hull, G., Lankshear, C., 1996. *The New Work Order: Behind the Language of the New Capitalism*. Allen & Unwin, London.
- Gee, J., 1996. *Social Linguistics and Literacies*. Falmer Press, London.
- Gee, J., 1999. An Introduction to Discourse Analysis: Theory and Method. Routledge, London.
- Gee, J., 2001. Learning in semiotic domains: a social and situated account. In: Schallert, D., Hoffman, L., Maloch, J.V., Worthy, B.E., Fairbanks, C.M. (Eds.), *51st Yearbook of the National Reading Conference*. NRC, San Antonio, Texas, pp. 23–32.
- Gee, J., 2003. Opportunity to learn: a language-based perspective on assessment. *Assess. Educ. Princ. Pol. Pract.* 10 (1), 27–46.
- Goody, J., Watt, I., 1975. The consequences of literacy. In: Goody, J. (Ed.), *Literacy in Traditional Societies*. Cambridge University Press, Cambridge, pp. 27–68.
- Goody, J. (Ed.), 1975. *Literacy in Traditional Societies*. Cambridge University Press, Cambridge.
- Goody, J., 1987. *The Interface Between the Written and the Oral*. Cambridge University Press, Cambridge.
- Gregory, E., Williams, A., 2000. *City Literacies: Learning to Read Across Generations and Cultures*. Routledge, London.
- Gregory, E., 1996. *Making Sense of a New World: Learning to Read in a Second Language*. Paul Chapman, London.
- Hackett, A., Somerville, M., 2017. Posthuman literacies: young children moving in time, place and more-than-human worlds. *J. Early Child. Lit.* 17 (3), 374–391.
- Havelock, E.A., 1963. *Preface to Plato*. Harvard University Press, Cambridge, Mass.
- Heath, S., 1982. Protean shapes in literacy events: ever-shifting oral and literate traditions. In: Tannen, D. (Ed.), *Spoken and Written Language: Exploring Literacy and Orality*. Ablex, Norwood, N.J., pp. 91–118.
- Heath, S., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.

- Heath, S.B., 2001. What no bedtime story means: narrative skills at home and school. In: Duranti, A. (Ed.), *Linguistic Anthropology: A Reader*. Blackwell, Oxford, pp. 318–342. First printed in *Language in Society*, (1982) 11(1): 49–76.
- Heller, M., 2007. In: *Bilingualism: A social approach*. Palgrave Macmillan, New York.
- Hull, G. (Ed.), 1997. *Changing Work, Changing Workers: Critical Perspectives on Language, Literacy and Skills*. State University of New York Press, NY.
- Hymes, D., 1974. *Foundations in Sociolinguistics*. Tavistock, London.
- Hymes, D., 1996. *Ethnography, Linguistics, Narrative Inequality: Toward an Understanding of Voice*. Taylor & Francis, London.
- Inkeles, A., Smith, D., 1974. *Becoming Modern: Individual Change in Six Developing Countries*. Harvard University Press, Cambridge.
- Ivanic, R., 1998. *Writing and Identity*. John Benjamins, Amsterdam.
- Jacobs, C., 2010. Collaboration as pedagogy: consequences and implications for partnerships between communication and disciplinary specialists. *South. Afr. Ling. Appl. Lang. Stud.* 28 (3), 227–237.
- Kalman, J., 1999. *Writing on the Plaza: Mediated Literacy Practices Among Scribes and Clients in Mexico City*. Hampton Press, Cresskill, NJ.
- Kapitske, C., 1995. *Literacy and Religion: The Textual Politics and Practice of Seventh-Day Adventism*. John Benjamins, Amsterdam.
- Kell, C., 2008. Making things happen: meaning making, agency and literacy in housing struggles in South Africa. *J. Dev. Stud.* 44 (6), 892–912.
- Kell, C., 2011. Inequalities and crossings: literacy and the spaces-in-between. *Int. J. Educ. Dev.* 31 (6), 606–613.
- Kell, C., 2015. Ariadne's thread: literacy, scale and meaning-making across time and space. In: Stroud, C., Prinsloo, M. (Eds.), *Literacy, Language and Diversity: Moving Words*. Routledge, London, pp. 72–91.
- Krause, L., Prinsloo, M., 2016. Translanguaging in a township primary school: policy and practice. *South. Afr. Ling. Appl. Lang. Stud.* 34 (4), 347–357.
- Krause, L.S., 2021. *Relanguaging Language From a South African Township School*. Multilingual Matters, Bristol.
- Kulick, D., Stroud, C., 1993. Conceptions and uses of literacy in a Papua-New Guinea village. In: Street, B. (Ed.), *Cross-Cultural Approaches to Literacy*. Cambridge University Press, Cambridge, pp. 33–61.
- Lam, W., Warriner, D., 2012. Transnationalism and literacy: investigating the mobility of people, languages, texts, and practices in contexts of migration. *Read. Res. Q.* 47 (2), 191–215.
- Lankshear, C., 1997. *Changing Literacies*. Open University Press, Buckingham.
- Latour, B., 2005. *Reassembling the Social: An Introduction to Actor–Network–Theory*. Oxford University Press, Oxford.
- Law, J., 2009. Actor network theory and material semiotics. In: Turner, B. (Ed.), *The New Blackwell Companion to Social Theory*. Wiley–Blackwell, West Sussex, pp. 141–158.
- Lea, M., Street, B., 1998. Student writing and faculty feedback in higher education: an academic literacies approach. *Stud. High. Educ.* 23 (2), 157–172.
- Leander, K., Boldt, G., 2013. Rereading “a pedagogy of multiliteracies”: bodies, texts and emergence. *J. Lit. Res.* 45 (1), 22–46.
- Leander, K., Burris, S., 2020. Critical literacy for a posthuman world: when people read, and become, with machines. *Br. J. Educ. Technol.* 51 (4), 1262–1276.
- Lillis, T., 2003. Student writing as “academic literacies”: drawing on Bakhtin to move from critique to design. *Lang. Educ.* 17 (3), 192–207.
- Lillis, T., 2006. Moving towards an “academic literacies” pedagogy: dialogues of participation. In: Ganobcsik-Williams, L. (Ed.), *Teaching Academic Writing in UK Higher Education: Theories, Practice and Models*. Palgrave MacMillan Publishing, London, pp. 30–45.
- Lindgren, T., 2019. The figuration of the posthuman child. In: *Discourse: Studies in the Cultural Politics of Education*. Taylor & Francis Online. <https://www.tandfonline.com/doi/full/10.1080/01596306.2019.1576589>.
- Luria, A., 1979. *The Making of Mind: A Personal Account of Soviet Psychology*. Harvard University Press, Cambridge, MA.
- Maddox, B., Esposito, L., 2012. Literacy inequalities, mediation and the public good: a case study of physical proximity and social distance in Nepal. *Br. J. Sociol. Educ.* 1–19.
- Martin-Jones, M., Jones, K. (Eds.), 2000. *Multilingual Literacies*. John Benjamins, Amsterdam.
- Massey, D., 2005. *For Space*. Sage, London.
- Maybin, J. (Ed.), 1994. *Language and Literacy in Social Practice*. Multilingual Matters, Clevedon.
- McClelland, D., 1967. *The Achieving Society*. Free Press, New York.
- North, A., 2018. Domestic work, learning and literacy practices across transnational space. *Int. Stud. Sociol. Educ.* 27 (2–3), 217–238.
- Olson, D., 1977. From utterance to text: the bias of language in speech and writing. *Harv. Educ. Rev.* 47, 257–281.
- Ong, W., 1982. *Orality and Literacy*. Methuen, London.
- Pahl, K., Rowsell, J., 2012. *Literacy and Education: Understanding the New Literacy Studies in the Classroom*, second ed. Sage, London.
- Papen, U., 2005. *Adult Literacy as Social Practice: More Than Skills*. Routledge, London.
- Prinsloo, M., Baynham, M., 2013. *Literacy Studies*, vols. 1–5. Sage, London.
- Prinsloo, M., Breier, M., 1996. *The Social Uses of Literacy*. John Benjamins, Amsterdam.
- Prinsloo, M., Krause, L., 2019. Translanguaging, place and complexity. *Lang. Educ.* 33 (2), 159–173.
- Prinsloo, M., Stein, P., 2004. What's inside the box? Children's early encounters with literacy in South African classrooms. *Perspect. Educ.* 22 (2), 67–84.
- Prinsloo, M., 2004. Literacy is child's play: making sense in Khwezi Park. *Lang. Educ.* 18 (4), 291–304.
- Robinson-Pant, A., 2001. Why Eat Green Cucumber at the Time of Dying: Exploring the Link Between Women's Literacy and Development-A Nepal Perspective. UNESCO Institute of Education. “UNESCO”. <https://unesdoc.unesco.org/ark:/48223/pf0000123606>.
- Rostow, W., 1960. *The Stages of Economic Growth: A Non-communist Manifesto*. Cambridge University Press, Cambridge.
- Scheeres, H., 2007. Talk and texts at work: beyond language and literacy skills. *Lit. Numeracy Stud.* 15 (2), 15–18.
- Scholtz, S., Prinsloo, M., 2001. New workplaces, new literacies, new identities. *J. Adolesc. Adult Lit.* 44 (8), 710–713.
- Schultz, T.W., 1961. Investment in human capital. *Am. Econ. Rev.* 51, 1–17.
- Scollon, R., Scollon, S., 1981. *Narrative, Literacy and Face in Interethnic Communication*. Ablex, Norwood, NJ.
- Scollon, R., Scollon, S., 2007. Nexus analysis: refocusing ethnography on action. *Ling. Ethnogr.* 11 (5), 608–625.
- Scribner, S., Cole, M., 1978. Literacy without schooling: testing for intellectual effects. *Harv. Educ. Rev.* 29 (2), 96–106.
- Scribner, S., Cole, M., 1981. *The Psychology of Literacy*. Harvard University Press, Cambridge, MA.
- Stein, P., Mamabolo, T., 2005. “Pedagogy is not enough”: early literacy practices in a South African school. In: Street, B. (Ed.), *Literacies Across Educational Contexts: Mediating Teaching and Learning*. Caslon, Philadelphia.
- Stein, P., Slonimsky, L., 2006. An eye on the text and an eye on the future: multimodal literacy in three Johannesburg families. In: Pahl, K., Rowsell, J. (Eds.), *Travel Notes From the New Literacy Studies: Instances of Practice*. Multilingual Matters, Clevedon, UK, pp. 118–146.
- Street, J.C., Street, B., 1991. The schooling of literacy. In: Barton, D., Ivanic, R. (Eds.), *Writing in the Community*. Sage, Newbury Park, CA, pp. 10 6–131.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Street, B. (Ed.), 1993. *Cross-Cultural Approaches to Literacy*. Cambridge University Press, Cambridge.
- Street, B., 1995. *Social Literacies: Critical Perspectives on Literacy in Development, Ethnography and Education*. Longmans, London.
- Street, B. (Ed.), 2001. *Literacy and Development: Ethnographic Perspectives*. Routledge, London.
- Thesen, L., van Pletzen, E., 2006. *Academic Literacy and the Languages of Change*. Continuum, London.
- Vygotsky, L., 1978. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.

- Wagner, D., 1993. *Literacy, Culture and Development: Becoming Literate in Morocco*. Cambridge University Press, Cambridge.
- Warriner, D., 2007. Introduction to special issue, "transnational literacies: immigration, language learning and identity". *Ling. Educ.* 18 (3/4), 201–214.
- Wilson, A., 2000. There's no escape from third-space theory—borderland discourse and the in-between literacies of prison. In: Barton, D., Hamilton, M., Ivanic, R. (Eds.), *Situated Literacies: Reading and Writing in Context*. Routledge, London, pp. 54–69.
- Zapata, A., Kuby, C., Johnson, J., 2018. Encounters with writing: becoming-with posthumanist ethics. *J. Lit. Res.* 50 (4), 478–501.

Literacies, language & schooling: teaching and learning the language arts

David Bloome^a, Hilary Janks^b, and Allison Wynhoff Olsen^c, ^aThe Ohio State University, Department of Teaching and Learning, Columbus, OH, United States; ^bThe University of the Witwatersrand, School of Education, Johannesburg, South Africa; and ^cMontana State University, Bozeman, MT, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	308
On the tension between defining literacy as autonomous skills and defining literacy as social processes and ideological	309
Definitional contrasts between autonomous and ideological literacy	309
The social impact of autonomous and ideological models of literacy in the classroom	310
From a view of language as bounded, codified and socially stratified to a view of language as fluid and variable in relation to social purpose and context	312
The expulsion of African-American language from the classroom	312
Promoting linguistic diversity in language pedagogy	313
From defining epistemology and rationality in the language arts as autonomous to dialectical and dialogic social practices	314
Implicit classroom communication about knowledge and rationality	314
Epistemological regimes	315
Final comments	316
References	316

Introduction

Scholars and educators have shifted conversations regarding the teaching of the language arts over the last four decades. Conversations had predominately been about “best practices,” what was the most efficient way to teach the skills of reading, writing, and using language. Increasingly, they now focus on how curriculum, instruction, assessment, and the everyday lives of teachers and students in language arts classrooms might be defined. Defining has not been a simple matter of working toward a definition, but rather the conversations are perhaps best characterized as centered around a series of tensions that continuously evolve.

A review of language arts curricula in primary and secondary grades across countries shows the field is mostly organized into the areas of reading, writing, speaking, and listening (e.g., in the US, the [Common Core State Standards Initiative, 2022](#); in South Africa, the [Curriculum and Assessment Policy Statement Standards, 2022](#)). Reading and writing are commonly defined as literacy skills. *Language* is understood as speaking and writing the dominant variety of the language using “correct” grammar. *Schooling* is defined as institutional spaces for educating children in the language arts, among other subjects. Debates among educational researchers and educators have focused on what constitutes these literacy and language skills and how best to school students in these skills (e.g., [Neuman and Dickinson, 2001](#)). It is within this framework that the constructs of scope and sequence for acquiring reading skills, writing skills, and language “correctness” are derived, as well as best instructional and assessment practices. So widespread and taken-for-granted in education policy, practice, and research are these conceptions of literacy, language, and schooling, that alternative conceptions are marginalized, if considered at all.

Inherent to this view of literacy and language learning are the binaries of success/failure and educable/ineducable. These, and similar binaries, promulgate deficit theories of learning to read, write and be schooled; a child either masters the target skills given appropriate instruction or by definition there exists a deficit. When difficulties in acquiring target skills are manifest by large numbers of children within a particular racial, cultural, linguistic, socio-economic, or neurodiverse group, the difficulties are seen by some educators and educational researchers as indications of deficits produced by characteristics inherent to that group (for further discussion see [Dudley-Marling, 2007](#); [Smitherman, 2015](#)).

Definitions of literacy and language in schooling contribute to the production of social identities including racially inscribed identities. For example, consider an instructional conversation captured by [Piestrup \(1973, pp. 54–55\)](#) in which a first grade African-American student has been asked to orally read a sentence and is repeatedly interrupted by the teacher.

Child 1: “Dey,
Teacher: Get your finger out of your mouth
Child: “Call ...
Teacher: Start again.
Child: “Dey call, “What i’ it? What is it?”
Teacher: What’s this word?
Child 2: Dey
Child 1: Dat
Child 3: Dey

Child 4: (Laughs)

Child 1: Dey

Teacher: Look at my tongue. They.

Child 1: "They"

Teacher: "They." Look at my tongue (Between her teeth).

Child 1: They

Teacher: That's right. Say it again.

Child 1: They

Teacher: "They." OK. Pretty good,

The child, who is a speaker of African American language (also called Black language, African-American English, Ebonics, among other labels), is confused and seemingly so are other children in the class. Using the phonemic system of his native language, he is pronouncing the word correctly and both he and the teacher recognize what the word is that he is orally reading. The teacher's correction has the effect of telling him that he is wrong when he has rightly identified the word, perhaps resulting in the child's questioning of himself and his competence to read and understand. There is also a subtle but powerful message that the way that he, his family, and his community speak is wrong.

This brief classroom episode of orally rendering a word reveals powerful schooling processes that create a "person" whose language, culture, family, and community are constructed in "deficit." This view of reading, as the ability to pronounce words in isolation, arranges people hierarchically along a scale of good to bad readers and simultaneously along a scale of people who do or do not speak a language that is acceptable in school. These hierarchies apply not just to the child, but to the communities from which they come. It is not the teacher alone who is implicated in such constructions of personhood, so too are governmental policies about how to teach reading, reading programs and assessments produced by large corporations, and supervisory observations of teachers that surveil their adherence to a school district's policies for teaching reading.

As the example and discussion above show, definitions of language, literacy, and schooling have consequences for children, teachers, communities, and what happens in classrooms. Here, we focus on three tensions in how language, literacy, and schooling have been defined.

1. The tension between defining literacy and language as decontextualized, autonomous, cognitive skills and defining literacy and language as ideologically inscribed social processes;
2. The tension between a view of language as bounded, codified and socially stratified and a view of language as fluid and variable in relation to social purpose and context.
3. The tension between defining epistemology in the language arts as autonomous, given and fixed, and defining knowledge as social constructions grounded in the material realities of the everyday worlds in which people live.

As prelude to our discussion of these tensions, we offer these four caveats: first, these tensions do not get resolved through empirical sciences per se nor through compromise, but rather, evolve through dialog; second, describing a definitional issue as a tension is not to imply that there is equal weight of warranted, evidenced, or ethical argumentation or legitimacy on both sides of the tension; third, power relations limit what can be said, who can participate, and what arguments can be made and are pervasive; and fourth, both we and the field more generally, struggle in discussing these tensions because of a lack of a shared language adequate to the task of unpacking.

On the tension between defining literacy as autonomous skills and defining literacy as social processes and ideological

Beginning prominently in the 1950s, the initial teaching of reading became overtly politically controversial (Allington, 2002; Larson, 2007) with some right-wing political leaders and pundits making specious arguments connecting particular definitions of reading and reading research with encouraging the undermining of Western culture and encouraging communism (e.g., Schlafly, 1997). Conversations about how to teach reading moved from the domain of educators and educational researchers to the broader public, mass media, pundits, and politicians. The political attacks and debates on the teaching of reading, literature, and the language arts by conservative groups were taken up by elected politicians who implemented government policies aligned with those conservative positions most notably by defining how reading could be taught and what research would be funded. In brief, the advocacy for direct phonics instruction, the teaching of "correct" grammar, and a restricted canon of literature for P-12 education was and is part of a broader culture war, an ideological battle that implicates government as well as the education, literacy and language research communities (for further discussion see Pearson, 2004).

Definitional contrasts between autonomous and ideological literacy

Street (1984) has labeled the view of reading, writing, and language use as skills as an autonomous model, which posits literacy (reading and writing) as a "neutral," context-independent technology that requires the acquisition of "universal" skills. Within the context of schooling, the autonomous model defines the teaching of literacy as the teaching of decontextualized skills (e.g.,

the teaching of decoding phonic skills) assessed atomistically, and the evaluation of learners in terms of their acquisition of the list of targeted skills. An autonomous model of literacy also continues to dominate educational research.

The autonomous model can be contrasted with what Street (1984) has labeled an ideological model of literacy. Street argued that all models of literacy are ideological, including autonomous models (despite their proponents' claim that autonomous models were universal and not ideological). An ideological model of literacy defines literacy as social processes involving written language that vary in relation to social and cultural situations and contexts. An ideological model of literacy recognizes that students acquire particular shared ways of using written language from their participation in their social and cultural lives at home, in their communities, and in school. These contextualized and situated ways of using written language can be called literacy practices (social practices involving the non-trivial use of written language, Street, 2000). Literacy practices can be understood as shared ways of using written language for engaging in recognizable types of social events that are materially bounded and enacted. As students and teacher come together in a literacy event such as a reading lesson, they do not enact an idealized literacy practice *per se*; rather, literacy practices get refracted. That is, the students and teacher bring to the classroom event their experiences and understanding of pertinent shared literacy practices and then together revise, modify, change, and transform those literacy practices (For further discussion of literacy practices see Street, 1995; Baynham and Prinsloo, 2009).

Consider the classroom event described earlier in which the teacher "corrects" a student's oral reading of "they" as "dey". From an autonomous model of literacy perspective, what is at issue is the student's acquisition of the skills for orally decoding the graphic sign "they" and his ability to render the sounds according to the norms of the dominant language variety, namely White standardized American English. From an ideological model of literacy, the question to ask would be what social practices are being enacted in this social event and what social ideologies and power relations are those social practices indexing? For example, consider the classroom event displayed earlier in which the teacher and students enact a discourse pattern of teacher initiation—"What is this word?"—followed by student response—"Dey"—followed by teacher evaluation—"Look at my tongue. They". This interactional pattern is a widespread classroom social practice often referred to as Initiate-Response-Evaluation (Mehan, 1979). As social practices are enacted within specific events, the social practice is refracted (modified) to address the specific situation and social context. In this case, the specific social situation includes racially-aligned language variation and power relations, a school district policy that emphasizes phonics-based reading instruction requiring "correct" and fluent oral rendition of "basic" words, interpersonal tensions between the teacher and the student and between the school and the local African-American community, as well as a broader social, political and historical context of segregation and a denial of civil and human rights to people of African descent in the United States. What is being taught and learned in this social event is not a set of decontextualized skills but rather a series of social practices about how written language is to be used in such classroom events, about social relations including who the teacher and student are (inclusive of their racial identities), about power relations between the institution of schooling and local communities, about who can claim belonging to the space of that classroom and school, and about the institutionally assumed qualities of the kinds of persons to be found in such locations.

The tension between defining literacy as autonomous versus defining literacy as social and ideological also plays out in multimodal and digital communication. Kress (2010), among others, drew attention to the ways in which print texts increasingly include visual images and spatial organization and online digital texts include a wide range of modes: images, moving images, sound, subtitles, speech, bodies and proxemics etc. Being able to read a text involves incorporating modes in addition to printed words. While multimodal texts have been ubiquitous long before digital communication, the widespread use of digital communication has highlighted the importance of navigating multimodal digital texts. Statistics show that the use of digital technologies is widespread and popular globally, especially among younger generations. As such, educators and educational researchers have, *ipso facto*, argued that the literacy skills needed for using digital technologies should be taught in school as part of the language arts—another iteration of the autonomous model (Baker, 2020).

From a social practices and ideological view of literacy, it is important to situate digital technologies within the complex, situated social contexts of their use. One needs to ask—how is any technology associated with written language (whether it is a stone wall and a burnt stick, paper and pencil, a sophisticated personal computer connected to the internet, or a networked system of school computers) being used, to do what, by whom, when, where, and how, within what social and historical contexts? Simultaneously, one needs to ask—Who has access to which technologies? What does the choice of technology accomplish? What modes do different technologies enable? What does a chosen mode mean, signify, and do?

The social impact of autonomous and ideological models of literacy in the classroom

Banks (2006) situates the tension between autonomous models of literacy and social practice/ideological models of literacy in African American history and its resistance to racism, oppression and Whiteness. He writes:

... rather than answer either/or questions about whether technological advancement and dependence leads to utopia or dystopic, whether technologies overdetermine or have minimal effects on a society's development, or whether people (especially those who have been systematically excluded from both the society and its technologies) should embrace or avoid those technologies, African American history as reflected through its rhetorical production shows a group of people who consistently refuse to settle for the limiting parameters set by either/or binaries. Instead, African Americans have always sought "third way" answers to systematically racist exclusions, demanding full access to and participation in American society and its technologies on their own terms, [using their own discursive practices] and working to transform both the society and its technologies, to ensure that not only Black people but all Americans can participate as full partners (pp. 2–3; original emphases).

The issues raised by Banks are not limited to the United States (for example, see [Faloye and Ajayi, 2021](#); [Nemer and Gray, 2019](#)). While access to new technologies has created what has come to be called “the digital divide”, access is not just a technical matter, but rather more complexly as part of relations of power (e.g., [Hoffman et al., 2001](#); [Strover, 2014](#)). These relations of power manifest themselves not simply as a matter of hands-on keyboards but more so as a tension over who one can be, where one can be, how one can be, what one can do, and whose literacy practices count within what social contexts including those social contexts that substantively involve the use of digital technologies.

For example, consider a 10th grade language arts classroom in which the students are exploring “The American Dream” from the perspectives of minoritized people (see [Bloome et al., 2019a](#), for a more detailed discussion of this classroom event.) They had been to a local museum to see an exhibit on the topic, generated essential questions to research, completed the research, presented to the class on their chosen essential question, and then engaged in a Socratic circle (a whole class peer discussion group) on a question the teacher posed, “What is the American Dream?” The students were from racially and religiously diverse working-class neighborhoods. Throughout the instructional unit, the students used a range of digital technologies including their cell phones with a broad range of apps (including word processing, internet search engines, Facebook, a broad range of other social media apps, and more), chrome books, a classroom desktop computer, smart board, and a school-based WIFI system. A few students also had digital notebooks. While some of the students did not have digital communication devices or WIFI access at home, many did. Throughout the instructional unit, regardless of the activity, students were often multi-tasking on their digital devices—interacting with others in the class or the teacher while also engaged in their academic work.

During the instructional unit the students saw a photograph by Louis Agassiz. Originally, the students believed that Louis Agassiz was Black and that his photographs were meant to show the painful conditions in which Black people lived. It was only when one of the students was researching him on the internet that she found out that he was White and a promoter of scientific racism. She told the teacher and included a photograph of Louis Agassiz in her PowerPoint presentation along with some other pictures he took, which she used to explain what scientific racism was. This led to a heated discussion in their Socratic Circle between a few of the White students and the Black students (and other students of color) about discrimination versus racism. The discussion affected future class discussions and social and friendship relationships.

Even a surface level look at the classroom events described above shows students’ resourceful and agentive use of digital technology, using apps, websites, and social media to explore their different concerns. Although the school district had implemented a series of filters that prevented students accessing particular types of web sites, when students wanted to access a particular web site that was blocked by the school’s chrome books or its Wi-Fi filters, they would use their smart phones to access those web sites. As [Hill \(2017\)](#) has noted in her study of young Black women’s digital composing in classrooms and in their homes and communities, students more freely employ their imaginations and bring their own experiences and thoughts to their composing outside of teacher surveillance.

While nearly all students used popular social media sites (e.g., Facebook, Wikipedia, Instagram, TikTok), their use of other web sites and networks was differentiated by race, heritage, and religion. Although the students had similar digital devices and tools, their experiences in digital communication and their social practices of use varied largely aligned with their different social identity positions.

Digital communication is not simply a matter of new technologies that require a set of autonomous, technical skills to engage with them—it is fundamentally ideological. As with print-based literacy, the social events that occur through digital communication involve tensions between the use of those social practices derived from the social, historical, cultural, and linguistic experiences of dominant groups and those social practices grounded in the experiences of marginalized groups. It should come as no surprise that the power relations of race play themselves out in and through digital communication; nor should it come as a surprise that marginalized students find ways to engage with digital communication that reflect who they are and how they are in the world. We believe that equity in literacy education requires us to embrace such diversity. Language arts needs a vision of literacy that is inclusive of students’ different ways with words ([Heath, 1983](#)) what Banks calls “discursive practices,” that is open to changes in the communication landscape, and that prepares students to understand and use the technologies and semiotic resources needed to meet the literacy challenges of different, ever-changing social contexts.

From the perspective of an ideological model of literacy, autonomous models are a non-sequitur. They are themselves a particular kind of ideological model that universalizes particular set of ways of using written language (the literacy practices of a socially, culturally, and politically dominating group) and denies the legitimacy and even existence of the literacy practices of other social, cultural, and linguistic groups. It is in this sense that a universalizing ideology obfuscates how it produces deficit theories aligned racially, culturally, linguistically, socio-economically, and in terms of neurodiversity. The tension that characterizes the schooling of literacy is not between autonomous and ideological models of literacy. Rather, the tension is between the privileging and hegemony of those literacy practices associated with dominant social, culturally, racial, linguistic, and economic communities and their enactment in classroom literacy events and, on the other side, recognizing, validating, and employing those literacy practices associated with a broad range of communities including marginalized communities. It is a tension not only about how to teach reading and writing but also about the system of social identities that students are allowed to envision and that are assigned in schools ([Worham, 2004](#)).

From a view of language as bounded, codified and socially stratified to a view of language as fluid and variable in relation to social purpose and context

Language arts education, with few exceptions, is explicitly designed to give students access to the dominant variety of a language, so that students learn to use it “correctly” in speaking and writing. This normalizes and standardizes the variety of the language spoken in school and dominant society as the home language of powerful elites, despite the fact that in some countries the home language of the majority of students is not the dominant language (for example, in South Africa, Zulu is the language spoken by most people, English and Afrikaans are minority languages in terms of the number of speakers; yet English is the dominant language with regard to schooling and government; Posel et al., 2020).

The restricting of language arts education to the dominant language of the society (or nation-state) is also, in part, the result of conceptualizing languages, dialects, and creoles as bounded, integral entities. Such a conception derives from a monolingual ideology that ignores the fluidity in use among named languages (Garcia and Wei, 2014; Makoni and Pennycook, 2007) as well as among dialects of the same language. Research shows that languages are not, metaphorically speaking, hermetically sealed; language contact is ubiquitous and is one of the main drivers of ever-present language change (Crystal, 2003). Although what counts as “correct” varies according to context, curriculum policy in language arts education around the world, with few exceptions, mandates the teaching of the dominant variety of the language, elevated by codification in prescriptive grammars and dictionaries. In brief, a majority of students world-wide are discouraged from using the full extent of their linguistic repertoires in language arts classrooms. A consequence of organizing the language arts curriculum around the trope of a dominating language is that speakers of other languages or varieties are constructed as lacking even if they speak several languages. At issue in the context of language arts education is that the rejection of non-dominant languages/varieties, multilingualism, and translanguaging from language arts classrooms marginalizes students who are speakers of these excluded languages, makes difficult their participation in instructional activities, and erases their voices (McKinney, 2017). In the case where the language of the school is dominant and historically colonial, the use of local, indigenous languages constitutes a decolonizing act of language resistance (e.g., Ramanathan, 2014; Reyhner, 2010).

The expulsion of African-American language from the classroom

The defenestration of African-American Language from language arts education in the United States, which has taken place in the historical context of the enslavement of African peoples and their subsequent segregation and oppression, delegitimizes the language and the culture (Baldwin, 1979; Scott et al., 2009). Sociolinguists have argued that African-American Language is a language in its own right and not a dialect of White Mainstream English (also referred to as Standard English) (Alim and Smitherman, 2012; Baker-Bell, 2020). This argument is based on linguistic evidence of the systematicity of African American Language and its distinct phonological, syntactic, and semantic systems (Ball et al., 2003). In the United States with its multi-racial population and its history of enslaving people of African descent and subsequent racial segregation, the exclusion and erasure of African American Language from the classroom can be viewed as a continuation of that racialized history, its violence, and the state-sponsored effort to delegitimize the full humanity and personhood of people of African heritage. That is, as Williams (1977) notes, “A definition of language is always, implicitly or explicitly, a definition of human beings in the world” (p. 21).

As such, the defenestration of African American Language from the language arts classroom can be viewed as a defenestration of African American people, culture, history, community, and children from the classroom. Following Baker-Bell (2020) recognizing African-American language as a language in its own right, (a) draws attention to anti-Black linguistic racism, (b) creates awareness of the vitality of African-American Language and the literacies and literature of African-American people, including students, (c) asserts students’ rights to their own language in the classroom (d) asserts the full personhood of children (and others) of African heritage, and (e) challenges the foundations of language arts education that takes for granted the exclusiveness of monolingualism and White standardized English.

Teaching the dominant language as the language of language arts is often justified as providing “the same education for all students” ensuring equal access to the economic, educational, social, and cultural opportunities in society. Yet, with such an approach that ignores inherent relations of power that are embodied in language (Bourdieu, 1991), students not only are denied access but also have their personhood threatened. This tension is at the heart of what Janks (2010) refers to as the “access paradox.”

There are curricular and instructional alternatives to language arts pedagogies that are monolingual-centric. These include:

- Translanguaging pedagogies which are inclusive and respectful of the fact that the language(s) students speak are inextricably bound up with their sense of identity community and belonging (e.g., Canagarajah, 2013; Makalela, 2015; McKinney, 2017; Baker-Bell, 2020).
- Third Space pedagogies in which local literacies of students from non-dominant communities are celebrated and constitute a basis for transformative spaces for new learning (e.g., Gutierrez, 2008);
- Cultural modeling pedagogies for teaching literature (e.g., Lee, 2004) in which students’ uses of language outside of the classroom and their local literacies are used to frame and define literary study;

- Early reading and literacy pedagogies that encourage students to use their own home languages to read and write about their own communities, culture, families, and lives;
- Curricular units for the study of language variation and for the field study of language diversity in students' own communities (e.g., Egan-Robertson and Bloome, 1998).

Promoting linguistic diversity in language pedagogy

Language study within the language arts classroom can supplant monolingualism with a curriculum that emphasizes linguistic diversity as a productive resource (rather than as a problem). For example, the curricular and pedagogical approach known as Language Awareness (Hawkins, 1987) encourages students to think about the differences between inflected and analytic languages, stress-timed and syllable-timed languages, cognate and non-cognate languages, and so forth. For example, in South Africa it helps if one understands the difference in the vowel phonemes of African languages and English. Because many African languages have only five vowel sounds, it is difficult for African students to recognize and produce the distinctions between long and short vowels in English (e.g., ship and sheep). Similarly, because African languages use tonal variation to produce different meanings, many English speakers cannot hear or make these distinctive sounds.

Because the language of language arts education is generally taught as a bounded entity, grammar instruction (also called language study) focuses prescriptively on the grammar of the dominant variety. Moreover, the study of grammar is often treated as a stand-alone, decontextualized aspect of language arts in education policies and standards (Zuidema, 2005). We note that there is an ongoing debate about whether grammar instruction should focus on skills acquired through drills, examine linguistic structures within model sentences and texts, or be supplanted by attention to rhetoric emphasizing how grammar affects meaning (Myhill, 2018).

Linguists make a distinction between prescriptive and descriptive grammars. Prescriptive grammar imposes standards and rules for how people *should* talk and write, including what tense to use, whether or not to use the third person, when to use the passive voice, the correct tense and so forth. Mandating the correct pronunciation of a word is also prescriptive. The grammar taught in language arts classrooms is predominantly a prescriptive grammar. These prescriptions are usually based on formal written language and not on how people actually speak in everyday life.

By contrast, descriptive grammar is based on how people actually use language in communication without a judgment of correctness. Descriptive grammars therefore recognize variation in how language is actually used, and they accommodate language change. Change in language happens slowly but can be seen in process.

Beyond the "correct" use of the dominant language, schools are also concerned that students learn to use language appropriately. We use "appropriately" here, to refer to the complex issues of ethics, group social identity, and power relations that underpin what Hymes (1972) described as "the rules of use" and what Cameron (1995, 2000) has labeled "verbal hygiene". Verbal hygiene refers to the broader social, cultural, and political processes by which the state, social groups and social institutions regulate language appropriateness. The cultural variation in language use (e.g., swearing, turn taking processes, acceptable volume, eye contact) together with cultural variation as to the appropriate choice of language variety, means that students are often expected to unlearn the verbal hygiene practices of home and community and subscribe to those mandated by institutional power relations. Appropriateness is then policed by teachers or others with institutional authority.

Cameron also notes the complex dynamics of verbal hygiene in the promulgation of so-called "politically correct" talk. Using sexist language as an example, Cameron says,

There is nothing trivial about trying to institutionalize a public norm of respect rather than disrespect ... through linguistic choices: in the context of addressing or referring to someone, words are deeds. Compare "hey, bitch!" with "Excuse me madam" (1995, p. 143).

Having seen the success of feminist interventions to abandon the sexist use of "he" as a generic pronoun in English, we now see the pronoun system slowly returning to an older use of "they" for both third person singular and plural, to accommodate people who do not confirm to the gender binary instantiated in "he" and "she".

Cameron (1995) notes that verbal hygiene includes language that obfuscates reality giving Newspeak, as described by Orwell (1949) in his novel 1984, as an example. In our current post-truth era, one would assume that language arts education has a role to play in helping students understand the interests served when the power of language is used to obfuscate reality by generating and spreading lies, alternative facts, and conspiracy theories (especially within the context of social media). Yet rarely is verbal hygiene part of the core of language study in the language arts classroom.

The typical emphasis in language study on form, prescriptive grammar, monolingualism, and dominant language can be productively supplanted by a focus on how language choices yield meaning (e.g., Myhill, 2018; Kamler, 2001) and how language produces and maintains power relations (Fairclough, 1989). Following Halliday on language as "meaning potential" (1985), attention can be focused on how users produce meaning by selecting from the options available in a language (e.g., tense, voice, sequence, vocabulary) or from a broader range of languages and language varieties in their linguistic repertoires. Of course, language

users—whether they are speakers, writers, bloggers, social media users, actors—not only select from the options available in and across languages, they also create new options for making meaning (as occurs in the creation of new words, new grammatical structures, new discourses, mash-ups, intertextualities, new images, etc.).

From defining epistemology and rationality in the language arts as autonomous to dialectical and dialogic social practices

Questions of epistemology and rationality have long been debated in philosophy. Such questions are also central to language arts education (Giroux, 1984; Magrini, 2010; Willis, 2017) although they are rarely explicitly referenced in education policy, practice or research. These questions are essential to argumentation, which is ubiquitous in language arts classrooms, especially in the teaching of literature and in the writing of essays. Manifest in the teaching of argumentation is a growing tension between stances toward epistemology and rationality that are given, autonomous, and formulaic, and approaches that are dialogic and dialectic. The former is typified by what is commonly taught as the five-paragraph essay (also known as the three-pronged essay): the writer presents a thesis or claim in the first paragraph, writes three paragraphs of supporting evidence, and ends with a fifth, concluding paragraph. In this and similar formulaic approaches to argumentation, students begin with a claim or thesis and then seek out supporting evidence. The ethical obligation inherent in argumentation oriented to supporting an *a priori* claim is advocacy for the selected thesis or claim. Within this framework, students' written arguments are often assessed by their inclusion of the structural components of an argument.

In some classrooms, two students may generate antithetical claims but if each has evidence and warrants to support their claim, both are evaluated as good arguments (a relativistic epistemology). In other classrooms, arguments are set in competition with each other, the winner having the stronger evidence and warrants (similar to how lawyers are viewed as making arguments in support of their clients). When students are taught that arguments are to be won or are relativist, they lose opportunity for nuanced dialog (Clifton, 2017). As such, students may view knowledge and rationality as universal instead of understanding knowledge and rationality as a social, relational process that acknowledges knowledge as fully embodied in relation to how we are in a social world (Newell et al., 2015).

Implicit classroom communication about knowledge and rationality

The use of the five-paragraph argumentative essay in primary and secondary classrooms does not usually involve explicit discussions about what counts as knowledge or what is rational. In a social and political context in which assertion substitutes for argumentation and empirical facts and material realities have the same status as "alternative facts" and fantastical conspiracy theories, many teachers find it challenging enough to teach students that claims need evidence. Nonetheless, even in such classroom situations, definitions of epistemology (Newell et al., 2014) and rationality (Wynhoff Olsen et al., 2013) are being taught, albeit implicitly. For example, consider a high school classroom in which the students are writing arguments about whether Atticus Finch, the protagonist in the novel *To Kill a Mockingbird* (Lee, 1988), is a hero. In response, the students pick a side and then search the literary text for three pieces of evidence that support their view. What counts as knowledge is text from the novel. The experiences the students themselves have had with race and racism are not even considered as possible evidence in support of their claim. What counts as epistemology and rationality exists autonomously, in isolation from the particular life-experiences the students have had. Imagine that this classroom consists of only White students who have grown up in a nearly all-White, *de facto* segregated community and school system. Imagine also that these students have repeatedly heard from popular media and other social institutions in which they participate (e.g., church, family) that racism was a problem in the United States, but that racism no longer exists. In their argumentative essays, students may include as a given truth (as not needing evidence or backing) the condition that racism was a problem "back then" (at the time in which the novel was set) but that it is no longer a problem. Their own personal experiences are such that there is no challenge or alternative to that belief. Thus, it appears to them "reasonable" and "rational" to build an argument based on the premise that racism was a problem but no longer is.

Now imagine that the students in the class are racially diverse and have grown up in a racially diverse community and that some of the students in the class (and some of their family members) have experienced racism, including racist violence. As the students share their theses and the evidence they are using to support their theses, classroom conversations may be characterized by disagreements and by the sharing of different experiences leading students to question the theses they have had. Perhaps the classroom dialog results in students having more sophisticated theses and more complex arguments. As such, the argument being made eschews the selection of a claim followed by the searching only for supporting evidence; and, by contrast, engages the students in exploration of the diverse kinds of evidence in search of a thesis that all the evidence supports. The point here is that what counts as knowledge and what counts as rational are socially situated even when on the surface they may not appear to be so.

Consider another high school classroom in which the students are reading and discussing Hemingway's short story, "Indian Camp" (1924). The students offer various interpretations of what happens in the story and what those events mean. The teacher requests that students support their interpretations and assignments of meaning by asking, "What did you find as evidence?" Implicitly, the teacher is asking the students to take a stance regarding what can and should count as knowledge (epistemology) and what

is reasonable (rationality) in their interpretations of the literary text and in their assessments of others' interpretations. Literary scholarship on "Indian Camp" characterizes the short story as a "Coming of Age" story (e.g., Young, 1964). One way to approach the teaching of "Indian Camp" would be to tell the students what literary scholars have decided about the story's meaningfulness. Students might be held accountable to craft arguments about the meaning of the story that are aligned with what literary scholars have decided. In such a case, epistemology and rationality are located outside of the classroom and outside of students' experiences; epistemology and rationality are, respectfully, autonomous and universal.

Consider an alternative classroom scenario: Imagine that the teacher has asked students to read the story and then make an argument about what kinds of dominance they found. Perhaps some students respond to the task by quoting text from the story, assuming, without reflection, that only textual evidence counts. Perhaps the history they have had in language arts classroom discussions has neither sanctioned other kinds of evidence nor asked students to explain how a particular bit of textual evidence warrants an interpretation. Perhaps a particular quotation from the literary text does not seem to the other students nor the teacher as a reasonable warrant for a particular interpretation or perhaps they all take the stance that each person is allowed their own interpretation regardless of whether the quote seems to others as reasonable evidence for the interpretation. Implicit in the relationship between evidence and an interpretation—the warranting—are assumptions about what is reasonable and what is not (Wynhoff Olsen et al., 2013).

Alternatively, imagine that the students get into a heated debate about whether a particular bit of text is evidence of, for example, misogyny. Some students might argue that the use of a pen knife by the doctor to perform a Caesarian birth and the ignoring of the woman's screams are evidence of a disrespect toward women. Other students might argue that the doctor needed to ignore the woman's screams in order to save her life. The teacher might orchestrate the classroom conversation so that the students cannot ignore addressing each other's arguments, pushing the students to explore and unpack the literary text in more detail and to bring their own knowledge and experiences about misogyny into the argument (experiences that may vary greatly depending largely on the gender of the student). Implicit in this classroom discussion are questions about what counts as evidence (epistemology) and what counts as reasonable warranting (rationality). Is it reasonable and rational for some students to ignore the evidence that other students have presented about the misogyny they have experienced and that they see in the story? Such ignoring might be characterized as a definition of irrationality (cf., Habermas, 1984). Returning to the teacher's decision to ask students to read the short story within the framework of dominance, would it be rational to read the story as a "coming of age" story that ignores misogyny and racial oppression especially given that it is part of the social context in which students live? The issue here is not just what counts as epistemology and rationality in the two classroom lessons described, but rather that what happens in such classroom events is the implicit socialization of students to institutionally supported definitions of epistemology and rationality.

Epistemological regimes

In their study of the teaching and learning of argumentation in high school English language arts classrooms, Newell et al. (2014) organize epistemologies into four groupings: given, relativist, critical, and dialectical. Given epistemologies locate knowledge in authorities who represent (are delegates of) Truth. An exemplar of a given epistemology is New Criticism. Students acquire the Truth from their teachers through didactic pedagogies (what Freire, 1972b, calls a banking model of education). Relativist epistemologies locate knowledge in the individual and in the individuals interaction with texts. An exemplar of a relativist epistemology is Reader-Response Theory (e.g., Rosenblatt, 1938, 1978). Students acquire knowledge by exploring their own meaning-making grounded in their own experiences as they interact with texts. Critical epistemologies locate knowledge in the interrogation of power relations that have resulted in civil, social, and economic inequities grounded in racial, gender, language, and class ideologies of hierarchy (e.g., Appleman, 2014). Dialectical epistemologies locate knowledge in the tensions that arise when people engage each other in dialog grounded in the material realities of their lives.

Note that each epistemology involves relations of power in locating authority for what counts as knowledge and what counts as reasonable. For example, there may be a top-down power relation between "authorities" in the field of literary scholarship (typically prominent literary scholars at elite universities) and the teacher, and between the teacher and the students. The correct interpretation of a literary text is established by elite authorities and the language arts teacher holds students accountable to acquiring that interpretation with its given evidence and given warrants. By contrast, power relations associated with relativist epistemologies may appear equal and democratic among the readers of a literary text as each reader has the power to articulate their own interpretation using their own definition of evidence and own rationality. Thus, the power relation between teacher and students is one in which the student is accountable to the teacher for articulating an individual interpretation that seems to the teacher to be authentic to the student's own individual experiences as applied to the literary text. Critical epistemologies requires teachers to provide students with an analytic framework that, in part, derives from a set of critical social theories and, in part, from the student's own experiences of oppression and exploitation. As such, the power relation between the teacher and student is one in which the student has authority (power) for interpreting a literary text; but in generation of an interpretation students are held accountable to teachers for using a critical framework and for using a particular set of their experiences. Dialogic epistemologies aspire to power-relations as held jointly and collectively by people (teachers and students) who are in dialog with each other. However, rarely is it the case that dialog occurs in a social context devoid of power relations and thus interpretations that are generated and articulated through dialog reflect the extant power relations.

The classroom dynamics described in this chapter exemplify the movement from epistemology and rationality as given and autonomous to epistemology and rationality as socially situated, dialogic, and dialectic (e.g., Bloome et al., 2019b). Defining epistemology and rationality as dialogic and dialectic, explicitly or implicitly, pushes students to explore and consider all pertinent evidence and potential warrants before generating a claim. Students must also engage others in dialog about the evidence, warrants, and possible claims to understand others' perspectives producing either a consensus or a situationally and historically qualified argument. Such moves push against ideological isolation and echo chamber effects (Mirra, 2018): two byproducts of insulated thinking, social media algorithms for funneling biased information, and an inundation of "alternative facts" and baseless conspiracy theories.

In the classroom described earlier in which the students were discussing "Indian Camp," there was a disagreement among the students about whether there is male dominance over females or White dominance over indigenous people in the short story (as if both were not simultaneously possible). As the students talk with each other, their insights expand (about both the story and about being human more generally). While not all of them come to agreement, their arguments become more sophisticated as they hear and have to engage divergent perspectives. Regardless of whether students eventually come to a consensus in building an argument about the meaning of the short story, what is being taught and learned is that definitions of epistemology and rationality are socially situated, dialogic, and dialectical. We also note that expanded views of rationality promote an ethics of care (Gilligan, 1982) that encourage people to consider themselves in relation to others and as such, take on the experiences of others (both within our time and across generations) as they engage in language arts curriculum.

Final comments

We have described evolving tensions in literacy, language and schooling manifest in language arts education. Heuristically, these tensions occur at both a surface level and at deeper levels. At a surface level, there are evolving tensions in defining literacy, language, and schooling as skills to be acquired with the inherent consequence of generating deficit theories. These deficit theories are aligned with power relations defined racially and linguistically and in terms of class, gender, sexuality, and neurodiversity. Teachers are often forced into prescriptive or restricted curricular choices, spanning across selections and considerations of texts; how to work with a whole class and/or create space for individual reading; and how to invite students to be both consumers and producers of texts and knowledge. Literacy, language and schooling exist in the tension between the pain and suffering of structural and institutional inequality and the ideal of literacy, language and schooling as democratizing practices restructuring power relations. At deeper levels, the teaching of the language arts lies in the tensions between: (a) defining literacy and language as decontextualized, autonomous skills to defining literacy and language as social practices and ideologies; (b) monolingual-centered and hierarchical conception of language to language as multilingual, and non-elitist; and (c) defining epistemology and rationality as context-independent versus socially situated and context-dependent.

How these tensions are addressed have consequences for people in their daily lives, for social institutions, for government, and for the pursuit of racial justice and educational equity. Literacy, language, and schooling are deeply entangled with ways of being, doing, believing and valuing that have profound social effects. Through and through, how the tensions are addressed are ethical and moral issues. Whether it is the storming of the US Capitol, the mass shooting/murder of people of Asian heritage, the murder of African-American people, genocides around the world, among other state and institutionally organized acts of violence and oppression, there has been an ethical and moral failure of individuals, communities, social institutions, and government to address the language and literacy events that have preceded these atrocities. It is also an educational failure. However, it is also the case that language arts education and language arts research have the opportunity to provide people, their communities, and their social institutions with an understanding of how language, literacy and schooling might be used to create futures that resist deficit theories, emphasize mutuality and respect for others, and pursue racial justice and dignity. Following Freire, language arts teachers and educational researchers can frame language arts and literary education to ensure that students are "creative subjects" (1972a, p. 29) with the power to use language and literacy to "name the world, to change it" (Freire, 1972a, p. 61), and who accept the responsibility to care for both the word and the world in their splendid diversity.

References

- Alim, H.S., Smitherman, G., 2012. *Articulate While Black: Barack Obama, Language and Race in the US*. Oxford University Press, Oxford, UK.
- Allington, R., 2002. *Big Brother and the National Reading Curriculum: How Ideology Trumped Evidence*. Heinemann, Portsmouth, NH.
- Appleman, D., 2014. *Critical enclutters in secondary English*. Teachers College Press, New York.
- Baker, E.A., 2020. *Wag the Dog—Digital-Literacies Narrative*. Presential Address, Literacy Research Association Conference, December. Available at: <https://www.youtube.com/watch?v=Avzup21ZnA4&t=52s>.
- Baker-Bell, A., 2020. *Linguistic Justice. Black Language, Literacy, Identity and Pedagogy*. Routledge, New York.
- Baldwin, J., 1979. If Black English isn't a language, then tell me, what is? N. Y. Times. Available at: https://archive.nytimes.com/www.nytimes.com/books/98/03/29/specials/baldwin-english.html?_r=1.
- Ball, A., Makoni, S., Smitherman, G., Spears, A.K. (Eds.), 2003. *Black Linguistics: Language, Society and Politics in Africa and the Americas*. Routledge, New York.
- Banks, A., 2006. *Race, Rhetoric, and Technology*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.
- Baynham, M., Prinsloo, M. (Eds.), 2009. *The Future of Literacy Studies*. Palgrave Macmillan, London.

- Bloome, D., Brown, A., Kim, M., Tang, R., 2019a. Languageing personhood in classroom conversation. In: Beach, R., Bloome, D. (Eds.), *Languageing Relations Across Social Worlds*. Routledge, New York, pp. 235–254.
- Bloome, D., Newell, G., Hirvela, A., Lin, T.-J., Brady, J., Ha, S.Y., Kwak, S., Seymour, M., Thanos, T., VanDerHeide, J., Wynhoff Olsen, A., 2019b. *Dialogic Literary Argumentation in High School Language Arts Classrooms*. Routledge, New York.
- Bourdieu, P., 1991. *Language and Symbolic Power* (J.B. Thompson, Trans.). Polity Press, Cambridge.
- Cameron, D., 1995. *Verbal Hygiene*. Routledge, London.
- Cameron, D., 2000. *Good to Talk?* Sage, London.
- Canagarajah, S., 2013. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge, New York.
- Clifton, J., 2017. *Argument as Dialogue Across Difference: Engaging Youth in Public Literacies*. Routledge, New York.
- Common Core State Standards Initiative, 2022. Retrieved on April 4, 2022, from <http://www.corestandards.org>.
- Crystal, D., 2003. *English as a Global Language*. Cambridge University Press.
- Curriculum and Assessment Policy Statement, 2022. Retrieved on April 4, 2022, from [https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements\(CAPS\).aspx](https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements(CAPS).aspx).
- Dudley-Marling, C., 2007. Return of the deficit. *J. Educ. Controversy* 2, 1. Available at <https://cedar.www.edu/jec/vol2/iss1/5>.
- Egan-Robertson, A., Bloome, D. (Eds.), 1998. *Students as Researchers of Culture and Language in Their Own Communities*. Hampton Press, Cresskill, NJ.
- Fairclough, N., 1989. *Language and Power*. Longman, London.
- Faloye, S.T., Ajayi, N., 2021. Understanding the impact of the digital divide on South African students in higher educational institutions. *Afr. J. Sci. Technol. Innov. Dev.* 1–11.
- Freire, P., 1972a. *Cultural Action for Freedom*. Penguin, Harmondsworth.
- Freire, P., 1972b. *Pedagogy of the Oppressed*. Penguin, Harmondsworth.
- García, O., Wei, L., 2014. *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan, London.
- Gilligan, C., 1982. In a Different Voice: Psychological Theory and Women's Development. Harvard University, Cambridge: MA.
- Giroux, H.A., 1984. Rethinking the language of schooling. *Lang. Arts* 61 (1), 33–40.
- Gutiérrez, K.D., 2008. Developing a sociocritical literacy in the third space. *Read. Res. Q.* 43 (2), 148–164.
- Habermas, J., 1984. *The Theory of Communicative Action: Vol. 1. Reason and the Rationalization of Society* (T. McCarthy, Trans.). Beacon Press, Boston, MA.
- Halliday, M.A.K., 1985. *An Introduction to Functional Grammar*. Arnold, London.
- Hawkins, E., 1987. *Awareness of Language: An Introduction*. Cambridge University Press, Cambridge, UK.
- Heath, S.B., 1983. *Ways With Words*. Cambridge University Press, Cambridge, UK.
- Hemingway, E., 1924. *Indian Camp*. Available at <https://www.nrcs.net/Downloads/Indian%20Camp.pdf>.
- Hill, H., 2017. *A Study of Black Adolescent Females Writing in an Urban Public School*. Doctoral Dissertation. The Ohio State University.
- Hoffman, D.L., Novak, T.P., Schlosser, A.E., 2001. The evolution of the digital divide: examining the relationship of race to Internet access and usage over time. In: Compaine, B.M. (Ed.), *The Digital Divide: Facing a Crisis or Creating a Myth*. MIT Press, Cambridge, MA, pp. 47–97.
- Hymes, D.H., 1972. On communicative competence. In: Pride, J.B., Holmes, J. (Eds.), *Sociolinguistics. Selected Readings*. Penguin, Harmondsworth, pp. 269–293.
- Janks, H., 2010. *Literacy and Power*. Routledge, London.
- Kamler, B., 2001. *Relocating the Personal: A Critical Writing Pedagogy*. SUNY Press, Albany, NY.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Taylor and Francis, London.
- Larson, J. (Ed.), 2007. *Literacy as Snake Oil: Beyond the Quick Fix*, second ed. Lang, New York.
- Lee, H., 1988. *To Kill a Mockingbird*. Grand Central Publishing, New York.
- Lee, C.D., 2004. African American Vernacular English as resource in cultural modeling classrooms. In: Ball, A., Freedman, S.W. (Eds.), *Bakhtinian Perspectives on Language, Literacy, and Learning*. Cambridge University Press, Cambridge, pp. 129–147.
- Magrini, J., 2010. How the conception of knowledge influences our educational practices: "Toward a philosophical understanding of epistemology in education". *Curric. Matters* 6, 6–27.
- Makalela, L., 2015. Moving out of linguistic boxes: the effects of translanguaging strategies for multilingual classrooms. *Lang. Educ.* 29 (3), 200–217.
- Makoni, S., Pennycook, A., 2007. Disinventing and reconstituting languages. In: Makoni, S., Pennycook, A. (Eds.), *Disinventing and Reconstituting Languages*. Multilingual Matters, Bristol, pp. 1–41.
- McKinney, C., 2017. *Language and Power in Post-colonial Contexts: Ideologies in Practice*. Routledge, London.
- Mehan, H., 1979. *Learning Lessons: Social Organization in the Classroom*. Harvard University Press, Cambridge, MA.
- Mirra, N., 2018. *Educating for Empathy*. Teachers College Press, New York.
- Myhill, D.A., 2018. Grammar as a meaning-making resource for improving writing. *L1 Educ. Stud. Lang. Lit.* 18, 1–21.
- Nemer, D., Gray, K.L., 2019. Reproducing hierarchies or resisting domination: exploring the gendering of technology spaces in the favelas. *Gend. Technol. Dev.* 23 (1), 76–92.
- Newman, S.B., Dickinson, D. (Eds.), 2001. *Handbook of Early Literacy Research*. The Guilford Press, New York.
- Newell, G., Van Der Heide, J., Wynhoff Olsen, A., 2014. High school English language arts teachers' argumentative epistemologies for teaching writing. *Res. Teach. Engl.* 49 (2), 95–119.
- Newell, G.E., Bloome, D., Hirvela, A., Lin, T.-J., VanDerHeide, J., Wynhoff Olsen, A., Buescher, E., Goff, B., Kim, M., Ryu, S., Weyand, L., 2015. *Teaching and Learning Argumentative Writing in High School English Language Arts Classrooms*. Routledge, New York.
- Orwell, G., 1949. *Nineteen Eighty-Four*. Penguin, London.
- Pearson, P.D., 2004. The reading wars. *Educ. Pol.* 18 (1), 216–252.
- Piastrop, A.M., 1973. *Black Dialect Interference and Accommodations of Reading Instruction in First Grade*. University of California Language and Behavior Research Lab, Berkeley (ERIC # ED119113).
- Posel, D., Hunter, M., Rudwick, S., 2020. Revisiting the prevalence of English: language use outside the home in South Africa. *J. Multiling. Multicult. Dev.* 1–13.
- Ramanathan, V., 2014. Overcoming colonial policies of divide and rule: postcolonialism's harnessing of the vernaculars. *Rev. Res. Educ.* 38 (1), 290–311.
- Reyhner, J., 2010. Indigenous language immersion schools for strong Indigenous identities. *Herit. Lang. J.* 7 (2), 138–152.
- Rosenblatt, L., 1938. *Literature as exploration*. Appleton-Century, New York.
- Rosenblatt, L., 1978. *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, Southern Illinois University Press.
- Schlaflay, P., 1997. Why education is Americans' number-one worry. *Contemp. Educ.* 68, 213–216.
- Scott, J.C., Straker, D.Y., Katz, L. (Eds.), 2009. *Affirming Students' Right to Their Own Language*. Routledge, New York.
- Smitherman, G., 2015. African American language and education: history and controversy in the twentieth century. In: Bloomquist, J., Green, L.J., Lanehart, S.L. (Eds.), *The Oxford Handbook of African American Language*. Available at <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199795390.001.0001/oxfordhb-9780199795390-e-51>.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge: UK.
- Street, B., 1995. Social literacies: critical approaches to literacy in development, ethnography, and education. Longman, London.
- Street, B., 2000. Literacy events and literacy practices. In: Martin-Jones, M., Jones, K. (Eds.), *Multilingual Literacies: Comparative Perspectives on Research and Practice*. John Benjamins, Amsterdam, pp. 17–30.
- Strover, S., 2014. The US digital divide: a call for a new philosophy. *Crit. Stud. Mass Commun.* 31 (2), 114–122.
- Williams, R., 1977. *Marxism and literature*. Oxford University Press, Oxford.

- Willis, A.I., 2017. Re-positioning race in English language arts research. In: Lapp, D., Fisher, D. (Eds.), *Handbook of Research on Teaching the English Language Arts*, fourth ed. Routledge, New York, pp. 30–56.
- Worham, S., 2004. The interdependence of social identification and learning. *Am. Educ. Res. J.* 41 (3), 715–750.
- Wynhoff Olsen, A., Ryu, S., Bloome, D., 2013. (Re)constructing rationality and social relations in the teaching and learning of argumentative writing in two high school English language arts classrooms. In: Dunston, P.J., Fullerton, S.K., Bates, C.C., Stecker, P.M., Cole, M.W., Hall, A.H., Herro, D., Headly, K.N. (Eds.), *62nd Yearbook of the Literacy Research Association Yearbook*. Literacy Research Association, Altamonte Springs, FL, pp. 360–377.
- Young, P., 1964. *Ernest Hemingway*. Minnesota University Press, St. Paul, MN.
- Zuidema, L., 2005. Myth education: rationale and strategies for teaching against linguistic prejudice. *J. Adolesc. Adult Literacy* 48 (8), 666–675.

Reading comprehension in troubled times

Gina N. Cervetti^a and P. David Pearson^{b, a} Educational Studies, School of Education, University of Michigan, Ann Arbor, MI, United States; and ^b Instructional Science, Graduate School of Education, University of California, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Reading comprehension in troubled times	319
Theoretical traditions	320
Critical literacy	320
Global literacies	321
Digital multimodal literacy	321
Practices needed to engage as text users/analysts and critics	322
Practice 1: unearthing ideologies in text	322
What is this practice	322
Why is it important in this historical moment?	322
What do we know about pedagogy?	323
Practice 2: engaging multiple perspectives and positioning	323
What is this practice?	323
Why is it important in this historical moment?	323
What do we know about pedagogy?	324
Practice 3: orientation toward action	325
What is this practice?	325
Why is it important in this historical moment?	325
What do we know about pedagogy?	325
Practice 4: evaluating text	326
What is this practice?	326
Why is it important in this historical moment?	326
What do we know about pedagogy?	326
Conclusion	327
References	328

Reading comprehension in troubled times

We write this chapter in the year 2021, at a time in which a litany of political, social, and health crises continue to confront the global community. In light of these challenges—and those that lie ahead—we consider (more aptly reconsider) how reading research speaks to this moment: What does it have to say about the skills (and knowledge) needed to make decisions as global citizens in an era and a world filled with social injustice, misinformation, and all too little cooperation and understanding within and across a host of boundaries?

In the main, the vast body of existing scholarship on teaching reading comprehension has not strayed far from the project of making meaning of individual texts for the purpose of temporally proximal engagement in discussions and written activities. Over the last five decades, reading researchers have produced a vast body of scholarship on teaching reading comprehension. Within these studies, comprehension instruction is often characterized as a portfolio of practices that help readers navigate the meanings that reside at the intersection of the reader, the text, and the context. By and large, the interventions described in this extensive body of research have focused on developing the reader's skills and knowledge, so they can better uncover—or construct a version of—meanings inscribed in text. These interventions have largely avoided a committed engagement with context (see [Pearson and Cervetti, 2015](#)). Instruction has focused on helping readers locate or recall information from texts, distinguish main ideas from supporting details, and make relatively shallow inferences among the ideas in a given text and between those ideas and their background knowledge.

Thus, reading comprehension, in the way we currently study, describe, and teach it, typically gets at what texts say (understanding and recalling explicitly stated key ideas and details), and what texts mean (where and how those ideas connect to one another and connect to existing background knowledge). Occasionally, reading comprehension instruction also involves engaging students in conversations about how texts work (when we study authors' craft and matters such as word choice, point of view, and shades of meaning), and even how language works. Rarely, comprehension instruction engages students in learning about what texts do (when we analyze and evaluate how texts and authors position readers and their cultural assets, and perspectives—privileging some and marginalizing others, or how we can use and apply the insights and knowledge we gain from our understanding to DO something out in the world). Research on this classic version of reading comprehension instruction has been summarized many

times over the last 35 years (Pearson and Cervetti, 2017; Duke et al., 2021; Pearson and Fielding, 1991; Pressley, 2001; Tierney and Cunningham, 1984; Wilkinson and Son, 2011).

As we considered the extent to which current reading research and instruction helps to prepare students to address pressing global challenges, we turned to a set of theoretical and pedagogical resources. First, we identified three loosely coupled traditions, each of which offers opportunities to reimagine comprehension beyond the boundaries of textual meaning making: Critical Literacy, Global Literacy, and Digital/Multimodal Literacy.

Second, to ground our work even more directly in pedagogical matters, we turned to Freebody and Luke's Four Resources Model (Freebody and Luke, 1990; Luke and Freebody, 1999), especially as it has been reframed to invoke a more critical view of what is involved in comprehension (Luke et al., 2011), or reimagined to apply to the rapidly expanding representations of meaning we encounter in digital/multimodal "texts" as well as conventional print (Serafini, 2012). The Four Resources Model describes a set of roles that readers play as they interact with text. The first two roles—the code breaker and the meaning maker—correspond to much of existing research on the nature of reading comprehension. But it is the final two roles that serve as the basis of our inquiry in this chapter. In the text user/analyst or analyst role, readers approach texts with an eye to how the construction of a text relates to its uses (e.g., how genre and discourse shape the meanings of inscribed messages) and how readers might use texts in the world outside. In the text critic role, readers consider the ideologies, interests, and voices at play in texts, as well as voices that are missing.

For the pedagogical emphasis, we looked to research that bears on these final two roles. In expanding reading comprehension to include the world that is represented in texts and the ways that texts are both shaped by and shape the world, the Four Resources Model orients reading, and thus, reading instruction, toward the purposes that inspire this essay—engaging in the global project of creating a safer and more equitable world. In particular, we draw on research around three pedagogical traditions—critical literacy, global literacies, and digital literacies—to consider what we do and do not know about how the skills necessary to enact the text user/analyst and text critic roles might be cultivated through ELA instruction.

As we searched for research about the nature and efficacy of attempts to use critical, global, and digital literacies to support students in taking on text user/analyst and text critic roles, we mainly found cases of individual classrooms and teachers that examined classroom practice with the "up close and personal" lenses of ethnography and discourse analysis. Hence, we shifted our goal away from attempting to generalize about efficacy (although we do refer to summaries of efficacy studies available) to accumulating and describing a rich and diverse sample of thick descriptions of this kind of instruction in action. This places limits on generalizations about the relative efficacy of different approaches within this body of instructional research. We are content, at least for now, to provide an account of what teachers can accomplish with different approaches to achieving the major goals of critical, global, and digital literacies. It is important to note that while this work can take place in different disciplinary contexts, we limited our investigation to ELA instruction, crossing over to science or social studies if and only if research in ELA was unavailable and the focus of a study was clearly on understanding or critiquing text in those disciplines.

Theoretical traditions

As a starting point for this inquiry, we briefly discuss critical literacy, global literacies, and digital multimodal literacy as three broad traditions that can inform the roles of text user/analyst and critic. We then discuss four practices that are valued within these traditions and describe what we know about their instruction from existing research.

Critical literacy

Critical literacy has roots in critical social theory, with its interest in the formation of a more just world through the practice of critique and social action. Critical literacy takes up these concerns in relation to the ways texts are regarded, interpreted, and. Texts are viewed as socially constructed and designed to present particular perspectives and worldviews that are necessarily selective and interested. In addition, these representations are viewed as having real world consequences as they shape understandings about culture and society and sustain dominant interests (Luke, 2018).

Grounded in the writings of the Frankfurt School Critical Theorists (see Bronner and Kelner, 1989; Freire, 1970), and critical postmodernists (e.g., Peters and Lankshear, 1996), critical literacy concerns itself with understanding the ways in which texts and discourses operate in social institutions to represent the world, construct identities, and position readers (Luke, 2000). Critical literacy is also concerned with devising ways of reconstructing these texts and, hence, the world. O'Brien (2001), a classroom teacher, wrote of her critical perspective:

[Texts] create ideologically-motivated versions of the nature of the social world and the complex, shifting unequal power relationships—e.g., those based on race or gender. From this perspective, texts are not fixed and complete objects to be comprehended, but opportunities for discussion, argument, and challenge as well as for enjoyment, information, and pleasure (p. 40).

According to McLaughlin and De Voogd (2004), critical literacy uses careful examinations of power relations in text to understand social problems in their complexity and to seek alternative explanations and underrepresented perspectives. The ultimate goal

of critical literacy is to prepare learners to understand power relations in texts and in the world so they can take action in the interest of creating a more just society (Vasquez, 2017).

McLaughlin and De Voogd (2004) describe a number of different instructional approaches for helping students develop a critical stance, including:

- Engaging in problem posing by questioning the author's message and representations using questions like, "what does this author want us to think?" and "how might alternative perspectives be represented" (p. 53).
- Juxtaposing different texts and images in order to consider multiple perspectives represented within and across texts.

Global literacies

A second, more recent perspective on literacy education that bears on the roles of text user/analyst and text critic is global literacies (Rizvi, 2009; Tierney, 2019; Tierney and Pearson, 2021). Dwyer's (2016) definition of global literacies includes elements of critical literacy with a stronger focus on cultural knowledge. Dwyer argues that information technologies are creating connections and interactions with diverse cultures and traditions. Global literacies involve the ability to analyze contemporary and historical issues from a perspective of understanding and respecting diversity.

Dwyer describes different sets of capabilities that students need in order to be globally literate, including:

- Ability to investigate the world and analyze complex global issues.
- Ability to be critical consumers of media (think critically, reflect, form judgments, and ask questions from multiple perspectives)
- Ability to conduct intercultural dialog across multimodal and print platforms
- Ability to hold an ethic around diversity and ethical orientations

Tierney (2019) describes global reading as a shift toward supporting meaning makers in "moving beyond their own cultural bubbles and self-interests to weighing information from multiple informants and multiple sources" (p. 42). This involves educational practice that leans toward multiculturalism and multilingualism, engaging students with global and cross-cultural issues. It also involves engaging students with diverse stakeholders and texts as a way of exploring different viewpoints, perspectives, and positionalities. These ways of reading invite students to step outside of their own views to imagine what the world, and the texts they read, might look like to others. Further, global reading develops an orientation that moves even beyond critique to action—a disposition to change the world in the name of social justice or equity and an ability to analyze complex global issues with an eye to solutions.

Yoon et al. (2018) describe an instructional framework that brings together critical literacies, global education, and multicultural education into a unified approach for using literacy instruction to develop responsible global citizens who see themselves as belonging to a world community that extends beyond local and national boundaries. In developing the framework, Yoon et al. invoke a "critical cosmopolitanism" that includes a focus on critical engagement with global issues and a focus on social transformation through action. The framework includes guidance for teachers on helping students:

- Develop an interconnected view of the world by engaging with world events and issues, using critical questions to understand the complexities of these issues (e.g., air pollution) and the ways that they reveal global interconnectedness
- Form connections between the global issues they investigate and their everyday lives and activities by reading about and investigating local issues (e.g., fracking) and personal practices (e.g., what they eat, the media they consume) that have global connections
- Analyze texts from global and cross-cultural perspectives by, for example, bringing these perspectives to bear on canonical literature, comparing Eastern and Western classic texts, and analyzing how advertisements from around the world invoke cultural values
- Learn about and engage in social and political actions to promote social justice.

Digital multimodal literacy

As we have implied, digital/multimodal (hereafter truncated to digital) literacy is often intertwined with both critical and global literacies. Digital literacy focuses on new ways of representing meaning associated with modern technologies and modes of communication. In the "world on paper," that prevailed a century ago (until the advent of movies and television), text is an inherently static mode of representing language and ideas. This world comprises symbols and images that are "fixed" on the paper on which they are printed where they can stand still for scrutiny. Not so in the year 2021; the world on screen seldom stands still. Besides the dramatic shift in our ideas about what text is and how it works, the explosion in modes of representation over the last century, especially the last quarter century, has dramatically altered our definition of what counts as text and how meaning can be represented in our everyday communications.

In terms of multimodality of representation, to the *static* forms already outlined (and more or less assumed in our earlier analysis of pedagogical principles) we must now add dynamic forms (Coiro, 2020). Movies and videos have been with us for almost a century and a half, but with the widespread influence of the internet, *dynamic* representations have become more prominent if not ubiquitous. Meaning potentials are no longer "fixed" on the printed page. Movement, via hyperlinks, can easily transport

the “reader” to almost any point or space within or across representations, whether the representations are themselves static or dynamic. Additionally, the digital setting adds another dimension of complexity because of its capacity to house, juxtapose, and even intermingle all of the modes. Then add the capacity of a digital setting to support both comprehension (readers have to be able to *read* within and across these modes) and composition (readers have a wider array of modes in which to *write* their ideas and interpretations), and the possibilities become even more complex, but also richer.

The capacity of digital, multimodal media to juxtapose so many forms and ideas redoubles the need for greater scrutiny in order to analyze and evaluate how texts and their authors (attempt to) shape our lives as readers in this new world of representation. Along with the wide array of possibilities and affordances that come with this explosion in what counts as a text, or at least modalities of representations, come a dramatically recontextualized, set of practices, strategies, and skills. [Fitzgerald et al. \(2020\)](#) note that “New technologies in particular require us to rethink how readers make meaning, as digital texts afford semiotic opportunities that are different from print texts ([Jewitt, 2005](#))” (p. 6).

In her conceptual essay about digital reading, [Coiro \(2020\)](#) notes key patterns of practices related to digital reading. Both involve the capacity of readers to shift rapidly and nimbly between or among:

- Complex and overlapping processes, including navigation, evaluation, and integration that require readers to move nimbly within and across multiple media (modes of representation) within digital environments, and
- At least two “reading gears,” including what we might call quick reading (browsing, linking, and scanning) and slow reading (the deep, reflective, and evaluation that is the essence of analysis and critique).

We side with [Serafini \(2012\)](#) in arguing that the analyst (with its emphasis on analyzing how writers use form to serve function) and critic (emphasizing the evaluation of the assumptions and consequences of text) roles are even more important in digital than the print formats. Our claim is based on the sheer volume of digital information that requires sorting as well as the lack of any broadly accepted conceptual or rhetorical standards for publishing on the internet. When the world of text is neither curated nor governed, when we live in the democratization of information (some have labeled it the “wild west”), it is important to make sure that every reader brings analysis and critique to the text.

Practices needed to engage as text users/analysts and critics

In the remainder of this essay, we detail a set of four practices derived from critical literacy, global literacies, and digital, multimodal literacies that address the call for students to embody the text user/analyst and text critic roles and that have the potential to do so in a way that develop students’ abilities to engage with contemporary dilemmas that have roots in social, political, and ideological tensions. For each, we describe the nature of the practice, why it might be important in this historical moment, and what we know about the nature and efficacy of its instruction.

Practice 1: unearthing ideologies in text

What is this practice

From the perspective of critical literacy, meanings are inscribed in text, but they are neither singular nor fixed. From critical perspectives, all texts and readings are embedded within normative relations of power and value and reflect the cultural positions and ideologies of the writer and reader ([Luke, 2000](#)). Thus, texts are both composed and comprehended from particular points of view and for particular purposes. Authors use the manipulable features of text (lexical, syntactic, discourse, structural, etc.) to privilege certain meanings, to position readers, and to support particular social relations through their portrayals of the world ([Luke, 2000](#)). Critical literacy seeks to develop students’ ability to read against the text by examining the contexts and features of texts’ construction and by carefully dissecting the meanings that are communicated. These examinations of the ideological dimensions of text are closely aligned with the Four Resources role of the text critic, and involve understandings about text construction and genre that are aligned with the text user/analyst role.

Why is it important in this historical moment?

Through interrogations of power and representation in text, critical literacy raises issues of social justice and equity, particularly related to the marginalization, disenfranchisement, and disadvantage of individuals, groups, or communities ([Vasquez, 2000](#)). From a critical perspective, it is vital to both examine the ways that language and literacy function in powerful ways—in texts and in the world—to produce and reproduce race, ethnicity, social class, and gender positions and to include marginalized perspectives in the school curriculum ([Moje et al., 2000](#)).

Critical literacy practices encourage students to see texts as both fundamentally ideological and also as open to a number of possible readings, including personal, cultural, critical, and resistant readings. Once students recognize that texts are mere representations of reality and that these representations are social and cultural constructions, they have a greater opportunity to take more powerful positions with respect to these texts—to resist, reject, deconstruct, and/or reconstruct them in ways that are more consistent with their own experiences in the world or in ways that create possible visions of a more just world. These ways of reading have the potential to prepare students to engage in discussions around important social issues and to resist false and problematic representations of the world. They also have the potential to establish a foundation for social action ([Kunnath and Jackson, 2019](#)).

What do we know about pedagogy?

Leonardo (2004) identifies critical classroom dialog as a key pedagogical feature, which includes asking students to critique difficult social problems. Many examples of critical literacy teaching include this kind of dialog, focused on texts and around texts and issues in children's schools and the world.

Critical teachers use a variety of approaches to encourage students to take a critical stance toward text, but critical practice is typically characterized by textual analysis, dialog, and questioning or problem posing. For example, Simpson (1996) documented one teacher's attempt to nurture an understanding among her seventh grade students that characters are not real but are constructed by authors, and that stories are not reflections of reality but selective versions of it. This teacher continually foregrounded the author's role in constructing the text, posing questions such as, "What does the author want you, the reader, to think and feel about particular characters or events? How does he or she achieve this?" Students compared their readings of these texts and examined the ways that the authors positioned them to respond to the texts in particular ways through choices about language, point of view, and other textual features.

Leland et al. (2005) document one first grade teacher's use of the entire range of roles in the Four Resources Model to implement a critical literacy curriculum in her classroom. The teacher used story time to engage students in critical conversations about homelessness and racial injustice. The teacher observed that her students became more aware of social issues, formed intertextual connections as they read, and began to ask more critical questions of the texts they read. She also observed an unusual environment of "acceptance and good will" in her classroom (p. 261).

Even very young children can analyze and question the portrayals of people and communities offered in texts. Labadie et al. (2013) examined the implementation of critical literacy read alouds in a multicultural and multilingual kindergarten classroom. Using texts that highlighted the effects of social class and poverty in different settings, including school and social settings, the teacher engaged students in extended conversations around a recurring set of themes, such as job loss, hunger, and homelessness. Students were also given opportunities to express their thoughts through drawing and writing. Labadie, Pole, and Rogers found that the children began to challenge some common societal views about social class. For example some students talked about how even people who have jobs can struggle financially.

Stevenson and Beck (2017) report on a qualitative study of a summer literacy program for intermediate and middle school children of migrant farmworkers. The program was designed around sociocultural and critical approaches. Students watched a documentary, read YA novels, and read children's books representing the lives of migrant farmworkers and featuring Mexican heritage protagonists. These books addressed a wide range of aspects of migrant life, including challenges related to poverty, working conditions, and discrimination. The students then created their own illustrated narratives about growing up as migrants. The researchers documented students' emerging critical consciousness (or "emergent conscientization") as they began to see their own experiences through the sociopolitical lenses and issues raised by the texts. Some students represented critical understandings of their lives as migrants in the narratives they developed, writing about the bullying they experienced and issues of poverty and child labor. Several of the narratives also included episodes that modeled how teachers and others could help them in overcoming some of these challenges.

Understandings about the constructed nature of texts, the power of texts to offer representations of the world, and the ability of readers to interrogate those representations are perhaps the most fundamental aspects of a critical orientation toward reading. Leveraging these understandings involves stepping into the text user/analyst and text critic roles offered in the Four Resources Model. At the heart of these roles is a move beyond "what the text says" or "what the text means" to consideration of "what the text does" as it describes lives, ideas, and events.

Practice 2: engaging multiple perspectives and positioning

What is this practice?

The ability to approach texts from multiple perspectives is a fundamental aspect of a critical stance (McLaughlin and De Voogd, 2004). Lewison et al. (2002) describe the multiple perspectives approach to critical literacy as asking us "to imagine standing in the shoes of others—to understand experience and texts from our own perspectives and the viewpoints of others and to consider these various perspectives concurrently" (p. 383).

Yoon et al. (2018) describe the development of cross cultural perspective taking skills as central to global literacies. One of these skills is the ability to "deconstruct their own ways of thinking" and analyze texts from diverse perspectives (p. 210). Thus, perspective taking in critical and global literacies involves both stepping into other perspectives and interrogating how one's own experiences shape interpretations of text and the world. This is a recognition that readers, like writers, are influenced by their experiences and their assumptions about social positionality (e.g., gender, race, age, and social class) (Moje et al., 2000).

Why is it important in this historical moment?

Critical educators commonly ask students to reflect on multiple, competing perspectives and examine competing narratives. Some critical literacy theorists suggest that the ability to represent alternative perspectives on information or representations in text may make readers more resistant to manipulation (Leonardo, 2004), an important skill in the current climate of misinformation (Vosoughi et al., 2018).

Others have suggested that perspective taking can build empathy and respect for difference. Pereira and Campos (2014) describe how the affordance of literacy in "offering the lives of others for our contemplation" and the practice of critical analysis and

identification with characters can build student's empathy (p. 359). In one representative study, Mabry and Bhavnagri (2012) demonstrate how non-immigrant students who were encouraged to engage in perspective taking were able to internalize and reflect on characters' experiences while engaging with children's literature about immigrant families. The students formed connections between their lives and those of the characters and expressed empathy and understanding of the immigrant children's feelings and experiences. Mabry and Bhavnagri (2012) point out that perspective taking is critical in diverse societies and foundational to cross-cultural communication. In an increasingly connected world with global challenges like the health and environment we are currently confronting, cross-cultural empathy and communication are essential. The capacity to assume different perspectives can, and we think should, promote both more resistant and more empathetic readings of text, enabling readers to be both more open and more closed to alternative readings of the word and the world, depending on the experiences and purposes readers bring to the reading event.

What do we know about pedagogy?

A number of articles by and for teachers describe instruction aimed at helping students use texts to take on new perspectives, develop deeper understandings of complex issues, and cultivate empathy. For example, Newstreet et al. (2019) describe how Shearer engaged her class of diverse sixth grade students with books about Muslim people and Muslim areas of the world. The students engaged in discussion as they read the texts, and they used internet research to augment their understandings. Their responses involved taking on the perspectives of individuals like those they read about. For example, after reading a true story about a refugee family, the students wrote about how they would feel if they had to leave their homeland. Some students also used the reading to reflect on their own experiences related to immigration. The authors saw evidence that students developed greater empathy in letters they wrote to pen pals in a rural Bangladeshi primary school. The authors argue that developing empathy in this way contributes to understanding and reduces intolerance.

Several studies have examined the use of critical conversations around read alouds can invite very young children to challenge dominant ideologies in text and take on new perspectives. In the Labadie, Pole, and Rogers study (discussed above), kindergarten children began to challenge conventional narratives about issues related to social class, including the nature and causes of financial struggles. Kim and Cho (2017) studied a preschool classroom in Korea where the teacher worked to use literacy activities to foster multiple perspectives and challenge the dominant ideologies inherent in children's books. The instruction took place through read alouds of classic children's books, including *Little Red Riding Hood* and *Snow White and the Seven Dwarves*. Choosing familiar books was seen as a way to engage in critical examination and new ways of seeing the stories. The children were guided to attend to marginalized voices in the stories, to bring their own perspectives to bear, and to challenge the author's message by "talking back" to the text (One example: seeing *Jack and the Beanstalk* from the giant's perspective.) The children used talk and drawing to reinterpret and recreate the stories. The authors present evidence that the children could examine ideological choices in the stories when guided by the teacher's problem-posing questions about whose viewpoint was expressed and who was silenced. For example, with *Jack and the Beanstalk*, the children speculated about the giant's point of view and life circumstances to better understand its frustration with Jack. Their drawings captured these new perspectives, showed empathy for the giant, and challenged the right/wrong duality in the story.

A number of studies describe how children's literature can be used to help students' critically consider their own experiences and others' and the cultural norms and policies that shape those experiences (e.g., Braden, 2019; Kim et al., 2020). Braden (2019) describes a study of third grade Latinx students' discussion of literature on themes related to immigration. The students engaged in read alouds, discussions, and blogging around critical multicultural literature over two months. In addition, students viewed images, watched videos, and had a visit from community members who shared experiences related to immigration. At the time of the study, the students, who were mainly of Mexican heritage, were grappling with new immigration policies in their state, which were impacting their communities and families. In their discussions, the students took up issues around anti-immigrant sentiments, the use of the term illegal, and the family separation associated with immigration. They used the stories and their own personal experiences to reflect on—and reimagine—immigration policy. They also considered different perspectives on immigration and developed alternatives to some of the dominant narratives about immigrants.

Kim et al. (2020) engaged Hispanic kindergarteners in literature-based explorations of gender and family composition that were often unfamiliar to them (Interviews with the children's parents revealed that the parents were concerned that reading LGBT themed books to their children might cause them to grow up to be gay or lesbian.) The students read, discussed, and wrote about picture books about gender non-conforming families, gender norms, and gender identities. Students were initially confused by the books and showed a preference for traditional family structures. The teacher asked students to consider their own experiences and others', to consider what makes a family, and to think about similarities and differences among different families, and to form connections between their own families and the unfamiliar families represented in the books. The children were initially unfamiliar with gender non-conforming families, but came to understand different family structures and to recognize similarities and differences with their own experiences.

As they challenge dominant and often hidden ideologies in texts and consider multiple perspectives, readers step beyond traditional forms of meaning making to embody the analytic (text user) and critical (text critic) roles described in the Four Resources model. These roles invite them into more agentic stances relative to text and more empathetic relationships with individuals and communities whose experiences differ from their own.

Practice 3: orientation toward action

What is this practice?

Critical approaches emphasize the importance of reading for real purposes, including the purpose of creating, maintaining, or expanding social change (Borsheim-Black et al., 2014). To set a foundation for action, critical teaching should help students read the world around them by examining socially relevant issues and current social practices (e.g., McLaren, 1998). Critical teaching urges students to critique current conditions, asking why we do things in the way that we do (Leland and Harste, 2000), what the consequences are for doing things in these ways, and how things might be done differently.

Nichols (2007) offers a vision of education that regards children as social agents and citizens in the present, rather than as citizens of the future. She advocates for literacy as a form of participation in the communities in which they are already members—"the spaces in which children live, work, and play" (p. 123).

Taking children's citizenship seriously means providing educational environments in which educators: model, explicitly teach and provide opportunities for the expression of views; actively listen and respond to children's viewpoints and encourage children to listen and respond to each other; hold up for question adults' views about children; explain that adults sometimes write these views down and that these texts impact on children's lives; and translate adults' texts for children so that children can actively engage with these texts (p. 128).

Moreover, critical literacy moves students from text critics to actors who "undertake transformative social action" in order to build a more just society (Janks and Vasquez, 2011, p. 1). This orientation toward action offers a natural extension of the role of text critic as described in the Four Resources Model.

Why is it important in this historical moment?

The goal of critical literacy is to develop students' agency to address inequities they experience in their own lives, to talk back to texts, to open themselves to the stories of others, and to become decision-makers in the classroom and in their communities outside of school. For many critical educators, the goal is to teach students to use literacy to work for social justice and equity in their lives and the lives of others (Heffernan and Lewison, 2000).

What do we know about pedagogy?

Several studies and reports from the classroom demonstrate the potential for young children to engage in critical inquiry of texts and move from text to the practice of social action. Powell et al. (2001) describe an inquiry-based project undertaken by a group of fourth-graders as an example of critical literacy. The authors described the student's efforts to investigate and ultimately prevent strip mining on Kentucky's Black Mountain. Though the project was not undertaken within an explicitly critical framework, the authors use the project to "illustrate the transformative potential of critical literacy." In another example, Sweeney (1997) describes a set of activities undertaken by her fourth-grade class around the South African elections of 1994, including the culmination of their work in a play about apartheid written by the students and performed for the other classes in their school building.

Curwen et al. (2019) report on a qualitative case study of fifth grade students and their three teachers (at a mainly white upper class school) as they participated in a critical literacy approach to studying the topic of slavery. The instruction focused on:

- Understanding how events are not isolated, but contextualized in particular social, historical, cultural and economic contexts.
- Interrogating the construction and evolution of race and racism
- Examining personal connections with historical events
- Drawing connections between the historical and contemporary—in this case, forms of slavery, and the underlying systems that have perpetuated them
- Interrogating multiple perspectives
- Taking action

The authors report that students were able to form connections across time and difference, recognizing common patterns, such as the human tendency to seek to dominate others. They also came to see themselves as agentic global citizens who could reshape the world through their own actions. The students produced activism-oriented projects in which they engaged in actions, such as writing letters to political leaders about modern day forms of slavery.

Gatto (2013) describes an ethnographic study of a critical literacy unit in her fourth grade classroom. She prompted her students to "do something" about the school lunches they complained about. The students engaged in problem-posing and investigations, many involving the interrogation of real world texts (e.g., letters, menus, blog comments, and newspaper editorial). The students developed a documentary to show to the school board and engaged in other forms of activism to improve the quality of school lunches. Gatto reports that students became more aware of the relationship between language, including text, and power and learned to question and take action on issues related to their lives.

As we suggested at the outset of our account of practice #3, when students move to this practice, they cross over the line from critic to actor in using their understanding and critique of text to engage in transformational social change. This has always been at least an implicit commitment of critical literacy, but in these troubled times, it has become an explicit commitment.

Practice 4: evaluating text

What is this practice?

As we have argued in unpacking the first three skills, texts are never neutral; to the contrary, they are always imbued with ideology that we must learn to read from multiple perspectives in order to take actions to improve the world. In order to enact the first three skills, we need a fourth—the capacity of individuals to apply a consistent set of criteria in order to evaluate the trustworthiness and utility of the information inscribed in texts.

Why is it important in this historical moment?

Paramount among the possibilities and problems that the digital/multimodal world brings to readers of all ages and especially to our classrooms is the issue of the trustworthiness of text information; it isn't that trustworthiness is not required in the world on paper; it's that the expansion in representational affordances and distributed authority for publication render the quest for trustworthiness more salient and more challenging in digital environments. Tom Freedman, NY Times columnist, writing in 1999, recognized this dilemma early on in our dalliances with the internet:

There is a dawning realization that the Internet ... is different from radio, television and newspapers in that it is totally open, interactive technology-but with no built-in editor, publisher, censor or even filters. With one mouse click, you can wander into a Nazi beer hall or a pornographer's library, hack the NASA computers or roam the Sorbonne library, and no one is there to stop or direct you. You interact with the network naked. The only really effective filters are the values, knowledge and judgment that your kid brings to the Web in his or her own head and heart (p. 15-A).

More recently and in a similar vein, the Organization for Economic Co-operation and Development (OECD, 2014), noted that readers can no longer rely on publishers who have traditionally acted as gatekeepers of textually-presented information; thus the responsibility for evaluating information has now become the responsibility of readers. The implication for this essay is transparent: Whatever we have said about the need for critique for garden variety texts goes double or triple for texts on the internet—the ultimate democracy of information. But it is doubtful that even an observer and prognosticator as insightful as Freedman could have anticipated, or even imagined, the blatant disregard for rules of evidence, argument, relevance, or trustworthiness in conveying information that we have witnessed in the last 5–10 years around the world, especially in the United States.

What do we know about pedagogy?

An internet search for the string, “research about teaching students to evaluate text for relevance and credibility” will yield thousands of hits and endless armchair advice and lesson plans about how to engage students in text evaluation, but precious little in the way of research assessing the efficacy or consequences of the recommended instruction. If one adds English Language Arts to the search string, the yield for research is even thinner, probably because science and social studies scholars have worried more about what has come to be called “sourcing” (assessing the accuracy and credibility/trustworthiness of evidentiary text sources) than have scholars who study literary interpretation, where relevance may be more important than credibility or accuracy for selecting a textual evidence to support a claim in an explanation or an argument a reader is constructing.

In one recent review of the scant literature available to assess pedagogical attempts to improve students' ability to evaluate texts on criteria such as relevance, accuracy, and trustworthiness/credibility, Fitzgerald and colleagues found very few pedagogical studies (Fitzgerald et al., 2020). In one exception, with college students reading civics texts, McGrew et al. (2019) found that training to judge text credibility improved scores on a civics on-line reasoning assessment.

In another recent review of intervention studies designed to improve students' sourcing skills (Brante and Stremse, 2018), which they viewed as attending to “metadata” or features of text sources (such as who created it, what affiliations they might have, for what purpose and audience they created it, and where and when it was published or disseminated) that influence readers' assessment of relevance, accuracy, and trustworthiness of the information in texts. From a much larger database, they found 18 studies that were relevant to their charge. The majority of the studies, especially those in elementary and middle school, included a focus on evaluating the credibility of the sources, most often by asking students to provide a rank-ordered list of the credibility of the sources, with modeling, group discussion, and worksheets serving as the primary instruction drivers. A few of the studies (e.g., Zhang and Duke, 2011) focused explicitly on examining credibility and requiring students to provide explanations for rankings of credibility. Very few studies went the next step by asking students to use their evaluation of sources as a criterion for interpreting the issues and questions they were trying to address. Nor did many of the studies focus much on motivation in spite of evidence (List and Alexander, 2017) that balancing ability with motivation is a characteristic of successful document evaluation. Six of the eight interventions that yielded large effect sizes for either judging credibility or interpretation relied heavily on teacher-led interventions. The authors concluded that more emphasis is needed on source evaluation of multiple documents representing different perspectives, as well as providing explicit instruction for judgments of both source relevance and credibility.

We searched, as we have with the other three crosscutting practices, for classroom cases documenting what students are capable of doing when given opportunities to learn these critical skills and dispositions. However, our search yielded very few classroom cases that focused explicitly on evaluating either single or multiple text sources on criteria such as perspective, relevance, accuracy, and trustworthiness/credibility. So we broadened our search, for this practice only, to include both older students and pedagogical experiments that relied on quantitative methods of data analysis. One of the most encouraging studies with college students (Wiley

et al., 2009), involved an instructional heuristic, SEEK (source, evidence, explanation, knowledge) taught to college students. When source evaluation for relevant and trustworthy evidence was a key component of instruction, students demonstrated greater text comprehension and better explanations (in post reading explanatory essays) of scientific phenomena that were the focus of their inquiries. In an experiment with elementary students, Zhang and Duke (2011) taught intermediate grade students a specific approach to evaluating internet sources. The routine (WWWDOT) for evaluating the relevance and credibility of internet sources focused on “who” wrote information (plus credentials), “why” they wrote it, “when” was it written or updated, “does” it meet my needs (relevance), “organization” of the site, “to do” list for follow up. The WWWDOT framework intervention, which consisted of four 30 min sessions, made students, “more aware of the need to evaluate information on the Internet for credibility and better able to evaluate the trustworthiness of Web sites on multiple dimensions.” (p. 145). However, when compared to a business-as-usual control group, WWWDOT students did not perform better in overall judgment about the quality of the information on websites.

Looking across this set of studies for Practice #4, Evaluating Text, our conclusions have a very different tone and voice than do those for Practices #1–3, more than likely because the evidence we were able to find is driven by more conventional outcomes, such as surveys and assessment tasks from which scores are derived and evaluated statistically. Additionally, the studies that have closely examined text evaluation tend to come from the liberal humanist “critical reading” tradition of critique, which is more focused on assessing the internal validity of texts (are the assumptions clear, the claims testable, the evidence linked to the claims through a solid line of reasoning). We found almost no text evaluation studies, even individual classroom cases, that emanated from the critical literacy perspective that is more sensitive to the external validity of texts, where texts are held to account for questions like, whose voices are privileged or marginalized (or silenced), or how power is assigned differentially. Notice that these questions are very similar to those we used to characterize what it means to enact the first practice—unearthing ideologies in text. This means that to evaluate these external faces of text involves unearthing their ideologies. So, to enact practice #1 is to enact this face of practice #4.

But the overall conclusion to be drawn is similar to practices #1–3: We know something about the efficacy and promise of this practice but we have a lot more to learn. Specifically, it is clear that (a) students can learn to apply evaluation criteria such as relevance (to my inquiry), trustworthiness/credibility, and perspective to texts they encounter in the context of completing a project, and (b) that when students assume a critical stance toward text, their understanding and ability to use what they have learned is enhanced.

Conclusion

The theory and research discussed in this chapter point to promising practices for supporting students in developing the kinds of reading practices that engage real world dilemmas, seed cross-cultural competence, and have the potential to bring about a more just world through social action and global cooperation. What emerges from this review is a vision of students who, from a very young age, move beyond traditional forms of reading comprehension to ask about the implications of texts for how we view and act upon the world. Reading comprehension is reenvisioned in a way that brings social dilemmas and injustices (alluded to, countered in, or raised by texts) to the fore and treats them as problems to be solved, rather than distant or unsolvable inevitabilities. It invites students to consider the commonality and connectedness of human experience, including but not limited to the worlds that students encounter on paper and on the screen.

More research is needed to guide this work. As we noted at the outset, there is very little research on the pedagogy, and even less on the efficacy, of various approaches to preparing students to enact the text user/analyst and text critic roles. Moreover, this new research must attend to variation in the characteristics of students, texts, and contexts. Even more important, we need to know what students can do when critical approaches become the norm rather than the one-off exception (they get it from a single teacher or for a weeklong unit in January).

Several pedagogical themes rise to the top—problem posing around text, engaging in rich discussions about the construction of the world in texts, using perspective taking to understand the experiences of others, and evaluating the trustworthiness of texts. Even so, much remains to be learned about the educational experiences that build these practices: What kinds of texts evoke critical conversations at different developmental levels? How can teachers guide students into these ways of reading and reasoning? How will students know whether the evidence they have gathered is convincing in supporting their case?

We also need to better understand the outcomes of this kind of instruction. This requires comprehension researchers to envision, develop, and validate a wider range of outcomes. Possible outcomes must include critical and intertextual reasoning with text, evidence of considering multiple perspectives, and, perhaps most important, evidence of transfer to the ways that students read, engage with digital media, and act outside of school. Longitudinal research is also needed to document whether this kind of textual work produces tangible outcomes for individuals and communities who enact them as a matter of everyday cultural practice.

One possible implication of this review is the addition of a fifth reader role: the text actor. This proposed role builds on the text user/analyst and text critic by envisioning a reader who uses text to inspire and inform social action in alignment with the aspirations of critical literacy. Students take on this role when they come to see themselves as having agency to change the world and when they use reading (and other literacy skills) to produce tangible actions toward positive changes in their communities, as did the students in the Curwen et al. (2019) study. This role builds on a wealth of international scholarship describing a global and decolonizing approach to education, especially in the areas of social studies and civics (e.g., Abdi et al., 2015; Andreotti and de Souza, 2011). For example, Silva (2016) describes a first grade class in which the teacher guided students to engage in actions, inspired by

their readings, around environmental justice. Students created materials to encourage their neighbors to recycle and compost. The text actor asks, what perspectives, policies, and practices need to change in order to address the social problems raised by this text? How can I take action to produce meaningful change? This role offers a view of reading instruction that positions the ultimate ends of the instruction as the development of engaged citizens who use text as a resource for problem-solving and participation. In other words the job of comprehension for troubled times is not finished until that comprehension informs and catalyzes action toward a safer and more just global community.

References

- Abdi, A.A., Shultz, L., Pillay, T. (Eds.), 2015. *Decolonizing Global Citizenship Education*. Sense.
- Andreotti, V., de Souza, L.M.T.M., 2011. *Postcolonial Perspectives on Global Citizenship Education*. Routledge, New York.
- Borsheim-Black, C., Macaluso, M., Petrone, R., 2014. Critical literature pedagogy: teaching canonical literature for critical literacy. *J. Adolesc. Adult Lit.* 58 (2), 123–133.
- Braden, E.G., 2019. Do you have papers? Latinx third graders analyze immigration policy through critical multicultural literature. *Child Lit. Educ.* 50 (4), 464–480.
- Brante, E.W., Stenise, H.I., 2018. Sourcing in text comprehension: a review of interventions targeting sourcing skills. *Educ. Psychol. Rev.* 30 (3), 773–799.
- Bronner, S.E., Kellner, D., 1989. *Critical Theory and Society: A Reader*. Routledge.
- Coiro, J., 2020. Toward a multi-faceted heuristic of digital reading to inform assessment, research, practice, and policy. *Read. Res. Q.* 56 (1), 9–31.
- Curwen, M.S., Ardell, A., MacGillivray, L., 2019. Hopeful discourse: elementary children's activist responses to modern-day slavery grounded in systems thinking. *Lit. Res.* 68, 139–161.
- Duke, N.K., Ward, A.E., Pearson, P.D., 2021. The science of reading comprehension. *Read. Teach.* 74 (6), 663–672.
- Dwyer, B., 2016. Teaching and learning in the global village: connect, create, collaborate, and communicate. *Read. Teach.* 75 (1), 131–136.
- Fitzgerald, M., Higgs, J., Palincsar, A., 2020. New media, new literacies: implications for reading for understanding. In: Pearson, P.D., Palincsar, A.S., Biancarosa, G., Berman, A.I. (Eds.), *Reaping the Rewards of Reading for Understanding*. National Academy of Education.
- Freebody, P., Luke, A., 1990. Literacies programs: debates and demands in cultural context. *Prospect* 5 (7), 7–16.
- Freedman, T.L., 1999. *Kids Need Values for Internet World*. The Post and Courier, p. 15A.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Seabury Press.
- Gatto, L., 2013. "Lunch is gross": gaining access to powerful literacies. *Lang. Arts* 90 (4), 241–252.
- Heffernan, L., Lewison, M., 2000. Making real-world issues our business: critical literacy in a third-grade classroom. *Prim. Voices K-6* 9 (2), 15–21.
- Janks, H., Vasquez, V., 2011. Editorial: critical literacy revisited: writing as critique. *Engl. Teach. Pract. Critiq.* 10, 1–6.
- Jewitt, C., 2005. Multimodality, "reading," and "writing" for the 21st century. *Discourse* 26 (3), 315–331.
- Kim, S.J., Cho, H., 2017. Reading outside the box: exploring critical literacy with Korean preschool children. *Lang. Educ.* 31 (2), 110–129.
- Kim, S.J., Kwangok, S., Corolla, P.A., 2020. Living together in a diverse world: conversations about the stories of two mommies and daddies among Mexican-descent bilingual kindergarteners. *Multicult. Learn Teach.* 15 (1), 1–14.
- Kunnath, J.P., Jackson, A., 2019. Developing student critical consciousness: twitter as a tool to apply critical literacy in the English classroom. *J. Media Lit. Educ.* 11 (1), 52–74.
- Labadie, M., Pole, K., Rogers, R., 2013. How kindergarten students connect and critically respond to themes of social class in children's literature. *Lit. Res. Instr.* 52 (4), 312–338.
- Leland, C.H., Harste, J.C., 2000. Critical literacy: enlarging the space of the possible. *Prim. Voices K-6* 9 (2), 3–7.
- Leland, C.H., Harste, J.C., Smith, K., 2005. Out of the box: critical literacy in a first-grade classroom. *Lang. Arts* 82 (4), 257.
- Leonardo, Z., 2004. Critical social theory and transformative knowledge: the functions of criticism in quality education. *Educ. Res.* 33 (6), 11–18.
- Lewison, M., Flint, A.S., Van Sluys, K., 2002. Taking on critical literacy: the journey of newcomers and novices. *Lang. Arts* 79 (5), 382–392.
- List, A., Alexander, P.A., 2017. Cognitive affective engagement model of multiple source use. *Educ. Psychol.* 52, 182–199.
- Luke, A., Freebody, P., 1999a. A map of possible practices: further notes on the four resources model. *Pract. Prim.* 4 (2), 5–8.
- Luke, A., Freebody, P., 1999b. Further notes on the four resources model. *Read. Online* 1–4.
- Luke, A., Dooley, K., Woods, A., 2011. Comprehension and content: planning literacy in low socioeconomic and culturally diverse schools. *Aust. Educ. Res.* 38 (2), 149–166.
- Luke, A., 2000. Critical literacy in Australia: a matter of context and standpoint. *J. Adolesc. Adult Lit.* 43 (5), 448–461.
- Luke, A., 2018. Regrounding critical literacy. In: Alverman, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*. Routledge, pp. 349–361.
- Mabry, M., Bhavnagari, N.P., 2012. Perspective taking of immigrant children: utilizing children's literature and related activities. *Multicult. Educ.* 19 (3), 48–54.
- McGrew, S., Smith, M., Breakstone, J., Ortega, T., Wineburg, S., 2019. Improving university students' web savvy: an intervention study. *Br. J. Educ. Psychol.* 89 (3), 485–500.
- McLaren, P., 1998. Revolutionary pedagogy in post-revolutionary times: rethinking the political economy of critical education. *Educ. Theor.* 48 (4), 431–462.
- McLaughlin, M., De Voogd, G., 2004. Critical literacy as comprehension: expanding reader response. *J. Adolesc. Adult Lit.* 48 (1), 52–62.
- Moje, E.B., Young, J.P., Readece, J.E., Moore, D.W., 2000. Reviving adolescent literacy for new times. *Perennial and millennial issues*. *J. Adolesc. Adult Lit.* 43 (5), 400–410.
- Newstreet, C., Sarker, A., Shearer, R., 2019. Teaching empathy: exploring multiple perspectives to address Islamophobia through children's literature. *Read. Teach.* 72 (5), 559–568.
- Nichols, S., 2007. Children as citizens: literacies for social participation. *Early Years* 27 (2), 119–130.
- O'Brien, J., 2001. Children reading critically: a local history. In: Comber, B., Simpson, A. (Eds.), *Negotiating Critical Literacies in Classrooms*. Erlbaum.
- OECD, 2014. Education at a Glance: OECD Indicators. <http://www.oecd.org>.
- Pearson, P.D., Cervetti, G.N., 2015. Fifty years of reading comprehension theory and practice. In: Pearson, P.D., Hiebert, E.H. (Eds.), *Research-Based Practices for Common Core Literacy*. Teachers College Press, pp. 1–24.
- Pearson, P.D., Cervetti, G.N., 2017. The roots of reading comprehension instruction. In: Israel, S.E. (Ed.), *Handbook of Research on Reading Comprehension*, second ed. The Guilford Press, New York, NY, pp. 12–56.
- Pearson, P.D., Fielding, L., 1991. Comprehension instruction. In: Barr, R., Kamil, M., Mosenthal, P., Pearson, P.D. (Eds.), *Handbook of Reading Research*, vol. 2. Longman.
- Pereira, I.S.P., Campos, A., 2014. Turning ideologies inside out: developing young readers' empathy for critical voices in narrative fiction. *Appl. Financ. Econ.* 21, 359–370.
- Peters, M., Lankshear, C., 1996. Postmodern counternarratives. In: Giroux, H.A., Lankshear, C., McLaren, P.L., Peters, M. (Eds.), *Counternarratives: Cultural Studies and Critical Pedagogies in Postmodern Spaces*. Routledge.
- Powell, R., Cantrell, S.C., Adams, S., 2001. Saving Black Mountain: the promise of critical literacy in a multicultural democracy. *Read. Teach.* 54, 772–781.
- Pressley, M., 2001. What should comprehension instruction be the instruction of? In: Kamil, M.L., Rosenthal, P.B., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research*, vol. III. Lawrence Erlbaum, pp. 545–561.
- Rizvi, F., 2009. Toward cosmopolitan learning. *Discourse* 30 (3), 253–268.
- Serafini, F., 2012. Expanding the four resources model: reading visual and multi-modal texts. *Pedagogies* 7, 150–164.
- Silva, J.M., 2016. Reading, writing, and revolution: facilitating social activism in first grade. *Soc. Stud.* 107 (5), 1–8.

- Simpson, A., 1996. Critical questions: whose questions? *Read. Teach.* 50 (2), 118–127.
- Stevenson, A., Beck, S., 2017. Migrant students' emergent conscientization through critical, socioculturally responsive literacy pedagogy. *J. Lit. Res.* 49 (2), 240–272.
- Sweeney, M., 1997. "No easy road to freedom": critical literacy in a fourth-grade classroom. *Read. Writ. Q.* 13 (3), 279–290.
- Tierney, R., Pearson, P.D., 2021. *A History of Literacy Education: Waves of Research and Practice*. Teachers College Press.
- Tierney, R.J., Cunningham, J.W., 1984. In: Pearson, P.D., Barr, R., Kamil, M.L., Mosenthal, P. (Eds.), *Handbook of Reading Research*. Longman, pp. 609–655.
- Tierney, R., 2019. Notes on global reading critical cultural traversals, transactions and transformations. In: Misiasek, L.I. (Ed.), *Exploiting the Complexities in Global Citizenship Education*. Routledge, pp. 38–68.
- Vasquez, V., 2000. Our way: using the everyday to create a critical literacy curriculum. *Prim. Voices K-6* (2), 8–13.
- Vasquez, V., 2017. *Critical Literacy*. Oxford Research Encyclopedia of Education. Oxford University Press.
- Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. *Science* 359 (6380), 1146–1151.
- Wiley, J., Goldman, S.R., Graesser, A.C., Sanchez, C.A., Ash, I.K., Hemmerich, J.A., 2009. Source evaluation, comprehension, and learning in Internet science inquiry tasks. *Am. Educ. Res. J.* 46 (4), 1060–1106.
- Wilkinson, I.A.G., Son, E.H., 2011. A dialogic turn in research on learning and teaching to comprehend. In: Kamil, M.L., Pearson, P.D., Moje, E., Afflerbach, P. (Eds.), *Handbook of Reading Research*, vol. IV. Erlbaum, New York, pp. 359–387.
- Yoon, B., Yol, O., Haag, C., Simpson, A., 2018. Critical global literacies: a new instructional framework in the global era. *J. Adolesc. Adult Lit.* 62 (3), 1–10.
- Zhang, S., Duke, N.K., 2011. The impact of instruction in the WWWDOT Framework on students' disposition and ability to evaluate websites as sources of information. *Elem. Sch. J.* 112 (1), 132–154.

Collaborative Reasoning and children's development of language and thought

Shufeng Ma^a, Richard C. Anderson^b, and Jingjing Sun^c, ^a Department of Educational Psychology, East China Normal University, Shanghai, China; ^b Center for the Study of Reading, University of Illinois at Urbana-Champaign, Champaign, IL, United States; and ^c Department of Teaching and Learning, University of Montana, Missoula, MT, United States

© 2023 Elsevier Ltd. All rights reserved.

Collaborative Reasoning: a new learning format	331
How Collaborative Reasoning promotes the development of language	331
Listening and reading comprehension	331
Speaking and writing proficiency	331
Language coherence	332
Academic vocabulary	332
Pragmatic awareness	332
How Collaborative Reasoning promotes the development of thought	332
Argumentative reasoning	332
Analogical reasoning	333
Multilink causal reasoning	333
Transfer of thinking and reasoning	333
The social mechanism of Collaborative Reasoning	334
CR fosters a zone of proximal development	334
CR creates opportunities for exploratory talk	334
CR promotes collaborative learning	334
CR engages students and teachers in the social exchange of ideas	334
Conclusion	335
References	335

This chapter describes an approach to collaborative learning called Collaborative Reasoning that facilitates the simultaneous development of language and thought. Collaborative learning generally refers to “a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” among a group of people (Roschelle and Teasley, 1995, p. 70). Collaborative learning creates a social setting that involves life-like problems and challenging learning tasks for students to develop and practice language and thinking skills. Compared to teacher-centered learning, where students have little control over when they can speak and what they speak about, collaborative learning offers students more autonomy in how they approach problem solving, as well as opportunities to evaluate their own thinking, rather than depending on the teacher to evaluate their contributions (Dillenbourg et al., 2009).

In teacher-centered environments, language and thinking are often viewed separately and taught from different angles. The typical approach to language teaching is to facilitate the acquisition of pronunciation, spelling, rules of grammar, vocabulary, syntax, and pragmatics, especially when the students are bilingual learners (August et al., 2014). Teacher-led direct instruction predominates in language teaching throughout the world due to its efficiency and perceived effectiveness (Adams and Engelmann, 1996; Chall, 2000; Hattie, 2009). It is common for students to recite words, complete grammar exercises, produce monologues, and enact short dialogs presented in the textbook, without ever actually engaging in social interaction (Michaels et al., 2010).

Similar practices are observed in the teaching of thinking skills. Cognitive skills are usually taught in elementary and secondary classrooms by having students solve well-structured close-ended problems, repeat a set of routine procedures, or practice the “transfer of knowledge” from one form of standardized test to another (Stigler and Hiebert, 2000).

A commonsensical, but dubious assumption is that students will be able to think and reason well once they have acquired basic language skills. Thinking and reasoning, however, do not necessarily improve with the growth in basic elements of language. Cummins (1979) found that, even after a person had become proficient in basic interpersonal communication, it usually takes another 3–5 years for them to be able to understand content knowledge and express academic ideas.

According to sociocultural theory, thinking is developed through the internalization of social practices (Vygotsky, 1978). Language plays an important role during the socializing practice of thinking. It provides the rudimentary tool for people to express themselves and engage in dialogs with more competent others so that they can complete tasks beyond their current developmental levels (Vygotsky, 1978). Sociocultural theory emphasizes that the ultimate goal of learning is not to pass tests in school but to become competent members of a community of practice (Lave and Wenger, 1991), and therefore “their [teachers’] effort was not just to teach the same things more effectively, but to have knowing be a product of the students’ individual and social intellectual activity” (Greeno, 1989, p. 137).

Collaborative learning has been shown to provide an environment where language and thinking can develop simultaneously. Three decades of research on collaborative learning has demonstrated its benefits in promoting students’ deep comprehension of

texts, language production, and a variety of higher-order thinking and reasoning skills (Mercer and Dawes, 2008; Murphy et al., 2009; Resnick and Schantz, 2015; Reznitskaya et al., 2009). To explore how collaborative learning fosters children's development of language and thinking, this chapter introduces a specific learning format called *Collaborative Reasoning* (CR). In the following sections, we describe key characteristics of CR, review studies that document the effects of CR on children's language and thinking, and discuss the mechanism by which collaborative learning could foster the development of language and thought simultaneously.

Collaborative Reasoning: a new learning format

Collaborative Reasoning is a special form of collaborative learning that consists of open discussion of significant, controversial issues managed by students themselves (Anderson et al., 1998). Students in a small group, normally between five to eight students, read a short story together and then have a discussion of a Big Question based on the story. The Big Question presents a dilemma that does not have immediate right or wrong answers. The dilemma involves a practical, ethical, or socio-scientific issue, such as whether zoos are good places for animals or whether a grizzly bear that has hurt a visitor to a National Park should be put down. Students are encouraged to argue for their position on the issue, use evidence from the story or life experience to support their arguments, and present counter-arguments to challenge their peers' positions or rebuttals to defend their own positions. They are expected to listen carefully to each other, challenge each other respectfully, and engage everyone in the group throughout the discussion.

An important finding is that child leaders emerge who help groups have socially satisfying and intellectually productive discussions. Li et al. (2007) identified five types of leadership moves in children's CR discussions: turn management, argument development, planning and organizing, topic control, and acknowledgment. These moves have been observed among elementary school children in the United States (Li et al., 2007), as well as children in China (Gao et al., 2015). The different types of leadership moves serve as building blocks of the leadership schema (Li et al., 2007), through which children learn to understand the social structure of collaborative groups, the roles fulfilled by different group members, the function of leadership moves, the complementary relationship between leading and following, and the resolution of conflicts during interaction.

As children take on the major responsibility for regulating their own discussions, teachers step away from being the center of the classroom and instead become discussion facilitators. They can prompt students to explain or clarify their positions, elaborate their ideas with details and evidence from text or personal experiences, or acknowledge students' ideas or appropriate social and collaborative manners (Chinn et al., 2001). In addition to facilitating discussions, teachers play a major role in reminding students of the discussion norms at the beginning, and at the end debriefing students through group reflection on what went well during the discussion, and what could be changed for the better (Baker et al., 2017). This reflection involves students' active critique of the content of the discussion as well as the manner through which they engaged with each other.

Initially developed in the United States, CR has been implemented successfully in several Asian countries—China, South Korea, and Malaysia. Despite cultural differences, children everywhere seem to be able to quickly adapt to this new collaborative learning format and enjoy the freedom in reasoning with peers.

How Collaborative Reasoning promotes the development of language

A body of research has shown that CR promotes children's development of various aspects of language including listening comprehension, reading comprehension, speaking and writing proficiency, language coherence, academic vocabulary, and pragmatic awareness.

Listening and reading comprehension

In a study involving 75 fifth-grade Spanish-speaking English language learners from a low-income school district in the United States, Zhang and her colleagues found significant improvement in listening comprehension and reading comprehension after the children had eight CR discussions during a four-week instructional intervention (Zhang et al., 2013). A similar study was conducted with 106 fourth-graders from working-class families in China (Cheng et al., 2015). Children's performance on a reading comprehension test (Progress in International Reading Literacy Study; PIRLS) was examined before and after the children participated in eight CR discussions. The results showed that children in collaborative groups were superior in evaluating and analyzing information as compared control students who continued recitation-based instruction. Students did not differ in recalling information or making simple inferences. These studies indicate that CR improves children's in-depth processing of information and higher-level reading comprehension.

Speaking and writing proficiency

After participating in CR, children produce a larger volume of language, produce longer words, generate a greater variety of words, and use more complex syntax in spoken and written language, all suggesting higher expressive language proficiency. Zhang et al. (2013) study found that bilingual children used more complex English syntax in their oral narratives and produced more diverse English vocabulary in their reflective essays after participating in CR discussions. Cheng et al. (2015) found that the volume, complexity, and fluency of students' oral Chinese all significantly increased from the first to the last CR discussion. Ma'rof (2014) investigated the oral and written English of 430 eight-grade Malaysian students whose home languages included Chinese,

Tamil (a language of south India), or Indonesian, as well as Malaysian. After six weeks of whole class discussions or small-group CR discussions in English, students who had small-group CR discussions showed significantly better English reading fluency, increased oral and written English production, as well as greater English syntactic complexity and vocabulary diversity. Across several studies CR has enhanced children's language skills regardless of whether they were speaking their first language, learning a second language, or they were limited by scarce resources at home and school.

Language coherence

The coherence of students' talk and writing, as indicated by their use of logical conjunctions and other coordinating words and phrases, increases after participating in CR discussions. A connective serves as the bridge to tie together different ideas. For example, children use logical or causal connectives (e.g., *because, if, so, so that*) to indicate causal relationships between arguments, and analogical or hypothetical connectives (e.g., *just like, what if, so as*) to compare two scenarios or map prior knowledge to new information (Lin et al., 2015). Morris et al. (2018) conducted a discourse analysis to examine children's use of logical conjunctions in classroom dialog. They found that logical connectives were five times more frequent in CR discussion dialog than recitation discourse on the same topic in comparable classrooms. The results indicate that children's talk is more complex and connected during CR discussion as compared to teacher-managed instruction.

Academic vocabulary

Children's development of language is manifested in gains in their ability to make productive use of advanced and abstract vocabulary. *Productive use* means children can employ the words when talking or writing, not just choose correct answers to multiple-choice questions. An important type of advanced vocabulary is *general academic vocabulary*, words used in many disciplines that are uncommon in ordinary conversation. Ma et al. (2017b) investigated children's spontaneous productive use of academic vocabulary. The results indicated that students who had six-weeks of CR discussions produced more general academic words, such as *increase, reduce, and protect*, in an oral interview on a different topic which they had not studied than students who studied the same curriculum via teacher-led direct instruction for six weeks. The difference could be traced to the fact that students in collaborative groups had abundant practice using academic vocabulary words during CR discussions.

Pan et al. (2020) further investigated the role of collaborative discussion in children's acquisition and productive use of general academic vocabulary words. Multilevel logistic regression analyses with crossed random factors involving 229 fifth graders and 325 words showed that the general academic vocabulary words children employed when writing essays was directly related to the frequency the child and the child's peers used the words during discussions, after controlling for prior knowledge of the words, the frequency of the words in the instructional materials, and the children's reading level. Thus, the reason that children were able to produce more academic vocabulary after having CR discussions is that they learned the words by listening to their classmates use the words and then using the words themselves.

Pragmatic awareness

Latawiec et al. (2016) investigated children's use of metadiscourse in reflective essays they wrote after participating in CR discussions. Metadiscourse refers to "how we use language out of consideration for our readers or hearers based on our estimation of how best we can help them process and comprehend what we are saying" (Hyland, 2017, p. 17). Metadiscourse is using linguistic markers to organize what you have to say and convey accurate meanings to your audience. Latawiec et al. (2016) found that children in collaborative groups "have greater concern for how readers will take their arguments and greater appropriation of argument-enhancing formal elements," while students who experienced teacher-led recitation "more frequently linked propositions together with simple additive conjunctions" (Latawiec et al., 2016, p. 23). This result suggests that CR students have developed a better sense of pragmatics and more awareness of effective communication.

How Collaborative Reasoning promotes the development of thought

We have elaborated on the impact of Collaborative Reasoning on children's development of language. A commonly asked question is—does language proficiency constrain one's development of thinking? Language is conventionally regarded as the symbolic expression of thought, so although language skills are considered to provide the means for an individual to express thoughts and ideas, thought is not necessarily limited by impoverished language. Several studies have shown that CR improves children's development of higher-order thinking and reasoning skills including argumentative reasoning, analogical reasoning, and multilink causal reasoning, and even children low on measures of language proficiency are able to transfer these forms of reasoning to other contexts.

Argumentative reasoning

Students' argumentation skills significantly improve during CR discussions (Reznitskaya et al., 2009). Children take positions, draw evidence from reading materials, provide reasons and justifications, and defend their positions with counterarguments and rebuttals. Students establish their own argument strategies or emulate other people's argument moves during CR discussions. Improvement in skills of argument has been documented with Chinese and Korean students (Dong et al., 2008; Kim et al., 2011) and Spanish-speaking English language learners (Zhang et al., 2013) as well as mainstream American children. Skill in argumentation has been found to snowball within groups and transfer between different contexts, which suggests the development of an argument

schema (Anderson et al., 2001). Children are able to transfer the skills of argument they learn in CR discussions to independent essay writing on a topic they have not discussed (Dong et al., 2008; Kim et al., 2011; Reznitskaya et al., 2001).

The open participation structure of CR discussions helps to ensure that students socialize their thinking. Students adopt such argument stratagems as "What do you think, [CLASSMATE]?" to maintain participation equity (Anderson et al., 2001). Those who are shy and reluctant to give their position in front of the whole group have more chances to speak due to encouragement from emergent child leaders (Li et al., 2007; Sun et al., 2017). When the students themselves have control over turn taking and topic, they are more likely to develop a sense of agency and engage in self-directed thinking and reasoning (Sun et al., 2021).

Analogical reasoning

In addition to learning argument skills, CR discussions accelerate children's ability to use analogies to validate their viewpoints. Analogical reasoning refers to establishing relationships between apparently dissimilar concepts based on overlapping deep structure. Lin et al. (2012) analyzed the dialog in 54 CR discussions and found that children's analogical reasoning grew steadily from the first to ninth CR discussion both in quantity and quality. The frequency of novel analogies increased within discussions and previously-used analogies were adopted by others once generated by innovative children (Lin et al., 2012). The growth of analogical reasoning and the spread of analogies within discussion groups suggest that children improved their ability to recognize structural similarities between different concepts.

Multilink causal reasoning

Not only does CR promote children's understanding of structural similarities among different concepts, it enhances their ability to connect concepts into interlocking causal chains, an ability we call *multi-link causal reasoning* (Ma et al., 2017a). Compared to students who were taught through teachers' direct instruction, students who participated in CR discussions created more causal chains and longer causal chains in independent writing about the topic they had studied for six weeks. A similar pattern was found in a near-transfer task, in which students were interviewed about a different but conceptually related topic. Analysis of the interviews showed that students in the CR condition generated longer causal chains and were more likely to identify causal relations and relational concepts (Ma et al., 2021; Ma and Anderson, 2015). In a far-transfer task, students were asked to tell a story based on a picture book, a task that involved no conceptual overlap with the topic students had studied for six weeks. Children who participated in CR discussions connected the successive events in the pictures into a coherent story based on the causal relationships, while children who received direct instruction made fewer attempts to link successive events and simply described the story scenes shown on each picture (Ma et al., 2017a).

Transfer of thinking and reasoning

Besides multi-link reasoning, students have also been able to transfer other higher-order thinking skills developed in CR, including decision-making and creative problem solving, in post-CR activities. Decision making has been defined as "the process of making choices among competing courses of action" (Beyth-Marom et al., 1991, p. 20). Zhang et al. (2016) examined children's decision-making ability in reflective essay written after experiencing a socio-scientific curriculum about wolf management through CR discussions. Compared to students in direct instruction classrooms or uninstructed control classrooms, students in CR classrooms were more likely to write individual essays that recognized more than one side of a dilemma, considered a greater variety of reasons, and weighed the relative importance of competing reasons. These findings suggest that as a result of participating in CR discussions children acquire competencies in analyzing the pros and cons of different options in the process of making moral decisions or public policy decisions.

While writing a reflective essay about a dilemma resembles what is required of active participants in a CR discussion of a controversial issue, Sun et al. (2017) investigated 252 fifth graders from mid-east China solving non-routine math problems, a process with few obvious features in common with CR discussions of moral and practical dilemmas. Nonetheless, students who had CR discussions before solving the math problems produced significantly better problem solutions than comparable control peers and had more positive feelings toward the problem-solving experience (Sun et al., 2017). The explanation was that children acquired leadership skills from participating in the CR discussions that are needed for harmonious and productive group work, and that they were able to employ these skills in navigating the cooperative problem-solving activity. Effective leadership served as a mediator for groups to reach better problem solutions.

In another study enrolling 474 seventh-grade students from a rural middle school in northern China, Ma and Zhou (2021) found that students who participated in an eight-day instructional intervention infused with CR discussions performed better in a creative problem-solving task than students in the control group. CR students proposed more solutions than those in the control group, offered a greater variety of solutions, and their solutions were more distinctive and judged to be more useful in a real-life situation. Collaborative Reasoning promoted children's creative solutions to social problems, suggesting that the experience of socializing with peers through CR may improve their ability to deal with social conflict in real life (Lin et al., 2022).

To summarize, during CR discussions, students freely express their views about the best answer to the Big Question and generate reasons to support their views. They challenge the reasoning of others in several ways, such as by asking clarification questions to promote further thinking when others' arguments seem incomplete or problematic. CR facilitates the emergence of novel ideas by providing ample opportunities for the social exchange of ideas among children, which improves their ability to generate solutions to several kinds of problems. Children are encouraged to draw on their own experiences, which probably facilitates the reconstruction of their mental models. Children learn to consider multiple points of view and weigh the pros and cons before making choices.

The social mechanism of Collaborative Reasoning

We have elaborated on the impact of CR on children's development of language and thinking separately. In this section we will consider the two facets of development together. We will attempt to explain the social mechanism of collaborative learning that fosters the development of language and thought, using CR discussion as one particular context for development. Our thinking is that language is fundamentally designed for interpersonal communication, and opportunities for interpersonal communication are maximized during collaborative discussions with peers.

In his sociocultural theory, Vygotsky proposed the concept of "zone of proximal development," which describes the difference between a person's actual developmental level and their potential developmental level that may be achieved with the help of others (Vygotsky, 1978). In this process of CR discussions, children receive guidance from more competent agents (e.g., peers with higher language proficiency as well, of course, as from their teachers and in other contexts from their parents) and gradually internalize it, making it distinctively their own. As a result, children develop their own ways of using a language and construct their personal understanding of the world through internalizing the conversations from the social cultural environment. During CR, this socialization process has several features that may contribute to children's rapid development of language and thought, including the free and open format of discussion, the exploratory nature of talk, the constructive interaction among people with different expertise, and the guided participation scaffolded by more competent peers and occasionally by adults. We will elaborate each feature below.

CR fosters a zone of proximal development

First of all, CR creates an open social context for children to exchange language and thoughts. Unlike traditional teacher-led instruction in which children are following steps established by the teachers to solve problems, CR naturally forms a zone of proximal development. The problems that children encounter in CR do not have definite answers or predetermined solutions, so they have to rely on each other to solve the problems. Hatano and Inagaki (1984) identified two types of expertise in learning: routine expertise and adaptive expertise. Routine expertise enables a learner to perform a task following well-structured procedures or a set of predetermined rules; whereas adaptive expertise refers to the ability to flexibly transfer what has been acquired in the past to new situations through accommodation and integration (Hatano and Inagaki, 1984). The process of solving an open-format, non-routine problem involves a series of high-level mental operations, such as retrieving relevant background knowledge, generating initial hypothesis, developing procedures to test hypothesis, drawing conclusion from the results of these procedures, and if necessary, revisiting the initial hypothesis to evaluate the results (Hmelo-Silver and DeSimone, 2013). The open-format of CR discussions enables children to constantly reflect on their thinking and modify these thoughts and ideas based on other people's inputs. Through socializing with others, children are able to develop a better understanding of the problem and accommodate what they know to the current uncertain situation.

CR creates opportunities for exploratory talk

Second, children have the opportunity to engage in "exploratory talk" during CR discussions, instead of having "cumulative talk" or "disputational talk." Disputational talk refers to the situation when speakers cannot reach an agreement but have to end up with individual decision making. Cumulative talk is the opposite; it means that speakers agree on the same opinion without critically responding to each other's thinking. During CR discussions, information is shared with everyone in the group and every member takes responsibility for decision making (Anderson et al., 2001; Chinn et al., 2001; Reznitskaya et al., 2009). Speakers in CR discussions "engage critically but constructively with each other's ideas" (Mercer, 2004, p. 146), which indicates the exploratory nature of the talk. Exploratory talk provides stimulates the socialization of thinking, and enables children to incorporate external information into existing schemas.

CR promotes collaborative learning

Third, CR creates an environment for children with different areas of expertise to communicate freely and interact constructively. Vygotsky (1978) believed that people are more likely to complete tasks that they cannot accomplish independently when they collaborate with more capable others. Previous studies have compared the situation of children completing complex tasks alone or completing tasks with others, and found that children who complete tasks in collaboration with others perform better than children who complete tasks alone (Azmitia, 1988; Fawcett and Garton, 2005). Following peers' modeling and explanation, students with lower academic competency are able to engage in higher-level thinking. When reminded of gaps, inconsistencies, or misconceptions in their statements, students have the chance to revise their original thinking. In collaborative learning, children with low-competency may spontaneously use their high-ability peers as a resource, asking them for help or confirming whether their thinking is on the right track (Garton and Pratt, 2001). Through constructive interaction, children often are able to incorporate complex language forms used by linguistically more advanced students, and internalize higher-order thinking and reasoning skills modeled by more competent peers.

CR engages students and teachers in the social exchange of ideas

Finally, teachers guide and support students participating in CR discussions, rather than talking a lot themselves. In CR discussions, teacher's role is not to cover the content knowledge, but to promote students' thinking and reasoning. Teachers' instructional moves include: [1] prompting for position, reason, evidence, clarification, alternative perspective, or counterarguments; [2] modeling the process of making arguments, defending positions, or challenging other people's opinions so that students learn ways to articulate

their thinking; [3] encouraging students to express their opinions freely and to challenge each other when they have a different opinion; [4] giving students the freedom to decide what to say and when they can say it; [5] bringing up unexplored topics, [6] challenging when the group too quickly reaches a consensus or only considers one side of an argument; and [7] summing up what has been said and how the argument has developed so that the group members can maintain a shared understanding (Chinn et al., 2001; Nguyen-Jahiel et al., 2007). With these instructional moves, teachers provide scaffolding as needed while maintaining students' autonomy to regulate their own learning.

We need to emphasize that teachers do not talk very much during CR discussions. Sometimes a whole discussion can take place without the teacher saying a word. The discussion belongs to the children. When the discussion is going well the children are in charge and can do most things for themselves.

Conclusion

Research over the past three decades shows that participation with small groups of classmates in Collaborative Reasoning discussions improves various aspects of children's language and thought. Aspects of language that are enhanced include listening comprehension, reading comprehension, and proficiency in speaking and writing. Children acquire and are able productively use academic vocabulary words; they tell stories that are more complete and coherent; they write essays in which they mark relationships among ideas with connective words and phrases. Research studies demonstrate that CR improves children's higher-order thinking and reasoning, including promoting the habits of mind to formulate complete arguments; make decisions that consider both sides of issues, take into account many reasons, and weigh the importance of reasons; and construct chains of reasoning to envision possible consequences of courses of action. Even children low on measures of language proficiency are often able to acquire these forms of reasoning and transfer them to other contexts.

References

- Adams, G., Engelmann, S., 1996. Research on Direct Instruction: 25 Years Beyond DISTAR. ERIC.
- Anderson, R.C., Chinn, C., Waggoner, M., Nguyen, K., 1998. Intellectually stimulating story discussions. In: Osborn, J., Lehr, F. (Eds.), *Literacy for All: Issues in Teaching and Learning*. Guilford, New York, pp. 170–186.
- Anderson, R.C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S.Y., Reznitskaya, A., Tillmanns, M., Gilbert, L., 2001. The snowball phenomenon: spread of ways of talking and ways of thinking across groups of children. *Cognit. Instr.* 19 (1), 1–46.
- August, D., McCardle, P., Shanahan, T., 2014. Developing literacy in English language learners: findings from a review of the experimental research. *Sch. Psychol. Rev.* 43 (4), 490–498.
- Azmitia, M., 1988. Peer interaction and problem solving: when are two heads better than one? *Child Dev.* 59 (1), 87–96.
- Baker, A.R., Lin, T.-J., Chen, J., Paul, N., Anderson, R.C., Nguyen-Jahiel, K., 2017. Effects of teacher framing on student engagement during collaborative reasoning discussions. *Contemp. Educ. Psychol.* 51, 253–266.
- Beyth-Marom, R., Fischhoff, B., Quadrel, M.J., Furby, L., 1991. Teaching decision making to adolescents: a critical review. In: Baron, J., V Brown, R. (Eds.), *Teaching Decision Making to Adolescents*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 19–59.
- Chall, J.S., 2000. *The Academic Achievement Challenge: What Really Works in the Classroom?* Guilford Press, New York, NY.
- Cheng, Y., Zhang, J., Li, H., Anderson, R., Ding, F., Nguyen-Jahiel, K., Shu, H., Wu, X., 2015. Moving from recitation to open-format literature discussion in Chinese classrooms. *Instr. Sci.* 43 (6), 643–664.
- Chinn, C.A., Anderson, R.C., Waggoner, M.A., 2001. Patterns of discourse in two kinds of literature discussion. *Read. Res. Q.* 36 (4), 378–411.
- Cummins, J., 1979. Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Work. Pap. Biling.* 19, 198–205.
- Dillenbourg, P., Järvelä, S., Fischer, F., 2009. The evolution of research on computer-supported collaborative learning: from design to orchestration. In: *Technology-Enhanced Learning: Principles and Products*. Springer, pp. 3–19.
- Dong, T., Anderson, R.C., Kim, I.-H., Li, Y., 2008. Collaborative reasoning in China and Korea. *Read. Res. Q.* 43 (4), 400–424.
- Fawcett, L.M., Garton, A.F., 2005. The effect of peer collaboration on children's problem-solving ability. *Br. J. Educ. Psychol.* 75 (Pt 2), 157–169.
- Gao, X., Cheng, Y., Li, H., Zhang, J., Ding, F., Nguyen-Jahiel, K., Anderson, R.C., Shu, H., Wu, X., Zheng, M., Cui, Y., 2015. Emergence and development of primary school children's leadership in group discussions [in Chinese]. *J. Psychol. Sci.* 38 (3), 612–617.
- Garton, A.F., Pratt, C., 2001. Peer assistance in children's problem solving. *Br. J. Dev. Psychol.* 19 (2), 307–318.
- Greeno, J.G., 1989. A perspective on thinking. *Am. Psychol.* 44 (2), 134–141.
- Hatano, G., Inagaki, K., 1984. Two courses of expertise. *Res. Clin. Cent. Child Dev. Annu. Rep.* 83, 27–36.
- Hattie, J., 2009. *Visible Learning: A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge, London, England.
- Hmelo-Silver, C.E., DeSimone, C., 2013. Problem-based learning: an instructional model of collaborative learning. In: Hmelo-Silver, C.E., Chinn, C.A., Chan, C.K.K., O'Donnell, A.M. (Eds.), *The International Handbook of Collaborative Learning*. Routledge, New York, NY, pp. 370–385.
- Hyland, K., 2017. Metadiscourse: what is it and where is it going? *J. Pragmat.* 113, 16–29.
- Kim, I.H., Anderson, R.C., Miller, B., Jeong, J., Swim, T., 2011. Influence of cultural norms and collaborative discussions on children's reflective essays. *Discourse Process.* 48 (7), 501–528.
- Lataviec, B.M., Anderson, R.C., Ma, S., Nguyen-Jahiel, K., 2016. Influence of Collaborative Reasoning discussions on metadiscourse in children's essays. *Text Talk* 36 (1), 23–46.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Li, Y., Anderson, R.C., Nguyen-Jahiel, K., Dong, T., Archodidou, A., Kim, I.H., Kuo, L.-J., Clark, A.M., Wu, X., Jadallah, M., Miller, B., 2007. Emergent leadership in children's discussion groups. *Cognit. Instr.* 25 (1), 75–111.
- Lin, T.-J., Anderson, R.C., Hummel, J.E., Jadallah, M., Miller, B.W., Nguyen-Jahiel, K., Morris, J.A., Kuo, L.-J., Kim, I.-H., Wu, X., Dong, T., 2012. Children's use of analogy during collaborative reasoning. *Child Dev.* 83 (4), 1429–1443.
- Lin, T.-J., Anderson, R.C., Jadallah, M., Nguyen-Jahiel, K., Kim, I.H., Kuo, L.-J., Miller, B.W., Logis, H.A., Dong, T., Wu, X., Li, Y., 2015. Social influences on children's development of relational thinking during small-group discussions. *Contemp. Educ. Psychol.* 41 (1), 83–97.

- Lin, T.J., Kraatz, E., Ha, S.Y., Hsieh, M.Y., Glassman, M., Nagpal, M., Sallade, R., Shin, S., 2022. Shaping classroom social experiences through collaborative small-group discussions. *Br. J. Educ. Psychol.* 92 (1), 131–154.
- Ma, S., Anderson, R.C., 2015. Transfer of Multilink Causal Reasoning in Collaborative Group Work [Paper presentation]. Annual Meeting of the American Educational Research Association, Chicago, IL, United States.
- Ma, S., Zhou, Q., 2021. Improving Creative Problem Solving Through Collaborative Reasoning [Paper presentation]. Annual Virtual Meeting of the American Educational Research Association.
- Ma, S., Anderson, R.C., Lin, T.-J., Zhang, J., Morris, J.A., Nguyen-Jahiel, K., Miller, B.W., Jadallah, M., Scott, T., Sun, J., Grabow, K., Latawiec, B.M., Yi, S., 2017a. Instructional influences on English language learners' storytelling. *Learn. Instr.* 49, 64–80.
- Ma, S., Zhang, J., Anderson, R.C., Morris, J., Nguyen-Jahiel, K.T., Miller, B., Jadallah, M., Sun, J., Lin, T.-J., Scott, T., Hsu, Y.L., Zhang, X., Latawiec, B., Grabow, K., 2017b. Children's productive use of academic vocabulary. *Discourse Process.* 54 (1), 40–61.
- Ma, S., Anderson, R.C., Lin, T.-J., Morris, J.A., Sun, J., 2021. Aspects of the Social Construction of Causal Reasoning [Unpublished manuscript]. Department of Educational Psychology, East China Normal University.
- Ma'rof, A.M., 2014. Think, Talk, Read and Write Better English: Improving L2 Literacy Skills of Malaysian Schoolchildren Through Collaborative Reasoning (Publication No. 3711056) [Doctoral dissertation. University of Illinois at Urbana-Champaign. ProQuest Dissertations Publishing.
- Mercer, N., Dawes, L., 2008. The value of exploratory talk. In: *Exploring Talk in School: Inspired by the Work of Douglas Barnes*. Sage, London, pp. 55–72.
- Mercer, N., 2004. Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *J. Appl. Ling.* 1 (2), 137–168.
- Michaels, S., Connor, M.C.O., Hall, M.W., Resnick, L.B., 2010. *Accountable Talk Sourcebook: For Classroom Conversation That Works*. Institute for learning, University of Pittsburgh, Pittsburgh, PA.
- Morris, J.A., Miller, B.W., Anderson, R.C., Nguyen-Jahiel, K.T., Lin, T.J., Scott, T., Zhang, J., Sun, J., Ma, S., 2018. Instructional discourse and argumentative writing. *Int. J. Educ. Res.* 90, 234–247.
- Murphy, P.K., Wilkinson, I.A.G., Soter, A.O., Hennessey, M.N., Alexander, J.F., 2009. Examining the effects of classroom discussion on students' comprehension of text: a meta-analysis. *J. Educ. Psychol.* 101 (3), 740–764.
- Nguyen-Jahiel, K.T., Anderson, R.C., Waggoner, M., Rowell, B., 2007. Using literature discussions to reason through real-life dilemmas: a journey taken by one teacher and her fourth-grade students. In: Horowitz, R. (Ed.), *Talking Texts*, first ed. Routledge, New York, NY, pp. 187–205.
- Pan, M., Anderson, R.C., Ma, S., 2020. Acquiring General Academic Vocabulary Through Collaborative Discussion [Paper presentation]. Annual Virtual Meeting of the American Educational Research Association.
- Resnick, L.B., Schantz, F., 2015. Talking to learn: the promise and challenge of dialogic teaching. In: *Socializing Intelligence Through Academic Talk and Dialogue*. AERA, Washington DC, pp. 441–450.
- Reznitskaya, A., Anderson, R.C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., Kim, S., 2001. Influence of oral discussion on written argument. *Discourse Process* 32 (2–3), 155–175.
- Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R.C., Nguyen-Jahiel, K., 2009. Collaborative reasoning: a dialogic approach to group discussions. *Camb. J. Educ.* 39 (1), 29–48.
- Roschelle, J., Teasley, S.D., 1995. The construction of shared knowledge in collaborative problem solving. In: *Computer Supported Collaborative Learning*. Springer, Berlin, pp. 69–97.
- Stigler, J.W., Hiebert, J., 2000. The teaching gap: best ideas from the world's teachers for improving education in the classroom. *J. Curric. Stud.* 32 (6). Simon and Schuster.
- Sun, J., Anderson, R.C., Perry, M., Lin, T.J., 2017. Emergent leadership in children's cooperative problem solving groups. *Cognit. Instr.* 35 (3), 212–235.
- Sun, J., Anderson, R.C., Morris, J.A., Lin, T.-J., Miller, B.W., Ma, S., Nguyen-Jahiel, K., Scott, T., 2021. Children's Engagement and Affect During Collaborative Learning and Direct Instruction. Department of Teaching and Learning, University of Montana [Unpublished manuscript].
- Vygotsky, L.S., 1978. *Mind and Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Zhang, J., Anderson, R.C., Nguyen-Jahiel, K., 2013. Language-rich discussions for English language learners. *Int. J. Educ. Res.* 58 (217), 44–60.
- Zhang, X., Anderson, R.C., Morris, J., Miller, B., Nguyen-Jahiel, K.T., Lin, T.J., Zhang, J., Jadallah, M., Scott, T., Sun, J., Latawiec, B., Ma, S., Grabow, K., Hsu, J.Y.L., 2016. Improving children's competence as decision makers: contrasting effects of collaborative interaction and direct instruction. *Am. Educ. Res. J.* 53 (1), 194–223.

Literature: critical imagination and young adult readers remaking worlds

Patricia Enciso^a and Beth Krone^b, ^a The Ohio State University, Columbus, OH, United States; and ^b Kennesaw State University, Kennesaw, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Remaking worlds	337
Introduction	337
Critical imagination, play and figured worlds	338
Uncertainty and imagination	338
Figured worlds	339
Play and figured worlds	339
Critical imagination and a transformative activist stance	339
Defining critical imagination in literary reading	339
World making	340
World making for presence, belonging, and collective action	340
Asserting presence	340
Belonging in possible futures	340
Collective engagement in digital worldmaking	341
Summary	341
Positioning and repositioning	342
Writing selves into texts	342
Rewriting spaces	342
Repositioning through ruptures	342
Language play	343
Freeing social signs	343
Figurative language	343
Conclusion	343
References	344

Remaking worlds

In a superhero storytelling project led by the authors, in a US midwestern middle school (ages 12–14), Kiana, Rene, Melissa, Lin and Aidan worked on building their superhero team. Each student had developed their own superhero, and Rene worked on drawing these superheroes on an iPad as the group thought through what their five superheroes might do in their neighborhood, the Mount, where their story was set. They talked about their superheroes heading to a different (real) neighborhood in their city, “where we go to get the best Halloween candy.” This led to a discussion of how the superheroes might trick or treat in their real superhero costumes. “I swear I’m cool,” Kiana noted, embodying her superhero.

Rene then changed the topic, “They should live in a universe where everyone has quirks and 30% of people do not have super-powers.” “Maybe we should make this like a fandom,” he mused. Aidan responded with a warning, “Do you know if we publish this we would not get paid a lot because it’s all five of us?”

Although at the time, we did not know what Rene was referencing in his mentions of “quirks,” we learned in exit interviews that the term was recycled from Rene’s deep knowledge of the Japanese anime and manga series and global franchise, *My Hero Academia* (MHA). The MHA storyworld posits a central character who, unlike most of his school-age peers, has no apparent superpower but begins to discover underlying strengths for resisting evil. Rene reported that he had been writing fanfiction about this series for a while on the platform, Wattpad, and that he had had over a thousand readers. “I like to create things in general,” Rene explained. “You decide how [the characters] act, what they do. It gives you a sense of power because you have control over the situation and stuff.”

Introduction

In their influential conceptualization of “restorying,” Thomas and Stornaouiolo (2016) identify six ways youth use fanfiction sites to critically and collectively revise the character representations of much loved stories such as *The Hunger Games* and *The Hobbit*, to better represent their identities as Brown and Black youth who want to feel present in these storyworlds. Thomas, for example, describes one youth’s practice of “race-bending” Galadriel, the magnificent elven queen from *The Lord of The Rings*, by using digital media tools to create an image of a dark brown-skin woman draped in a gem-filled robe. The image was posted to a fanfic website where she garnered praise for the beauty of her design. Similarly, youth in the vignette above extend their pleasure in viewing MHA

from knowing every detail of the original characters' personalities and powers to creating their own anime characters whose inner and outer conflicts reflect who they are in a school and society that often denies or threatens their differences.

As illustrated by these examples, restorying entails "a participatory textual practice in which young people use new media tools to inscribe themselves into existence" (Thomas and Stornaoui p. 313). Drawing on Rosenblatt's (1978) transactional theory of literary reading, restorying centers readers as story makers, whose references to their intertextual and multi-sensory knowledge, experiences, and multi-modal meaning-making bring stories to life, under guidance by the text. Whether the text is presented as an oral history, counterstory (Solorzano and Yosso, 2001), folktale, joke, meme, novel, comic, manga, picturebook, audio book, anime, YouTube video, or film, today's youth are using digital tools, expansive content, and transmedial platforms to support their drive to read, remake, and be recognized within and beyond the narratives they value.

In this chapter, we highlight research with youth whose participatory textual practices may be understood as both restorying and, within a broader sociocultural framework of interpretation and cultural production, as enactments of critical imagination. Scholars and educators who study children's and young adults' interpretations of literature and related story forms have long recognized that imagination underwrites readers' and storytellers' play with and transformations of meaning and language as they shape and reshape story worlds, and simultaneously transform their sense of self among others (Dyson, 1996; Enciso, 1994, 2003; Hooks, 1991; Winn, 2010). As early as the 1930s, Vygotsky proposed that play and imagination are necessary for displacing or releasing the immediate material and social reality so that a new temporary world might be evoked. He recognized that just as children must be able to displace reality in order to create new meanings about the world through improvised play, they must also regard abstract symbols as representations for images, actions, and feelings in order to create meaning from texts. Further, story interpretations are a form of improvised meaning-making. As Smagorinsky (2001) notes, Vygotsky (1987) understood that literary experiences cannot be predetermined or finalized but rather sensed, initially, as a "cloud of meaning" expressed and transformed through a "shower of words" (Vygotsky, 1987, p. 281). Literacy, in this view, entails constant movement between symbols and lived experiences, subject to historically formed meanings and relations of power throughout the interpretive process (Bakhtin, 2010). As a ubiquitous and transformative cultural practice, however, imagination is not well understood or explicitly figured into literary response theories and story making pedagogies. Along with Rosenblatt's transactional theory, restorying is one of the clearest conceptualizations of imagination-in-action through which youth use the gaps in representation to not only complete a text but also displace dominant images by using digital and analog tools to experiment and play with whom they might become within a wide range of textual storylines, characters, and settings.

In the following sections, we draw on contemporary sociocultural theory, including transformative activist theories of imagination to outline key features of imagination in literature and literacy education and its critical potential for youth's opportunities to author their own worlds. Throughout, we highlight research that illustrates youth story reading and story making as critical imagination and restorying practices. We categorize these studies in relation to three overarching processes in story reading and story making: (1) world building (2) perspective taking and positioning and (3) language play. We conclude with proposals for centering critical imagination in literature education and research.

Critical imagination, play and figured worlds

In this section we describe foundational theories of imagination (Vygotsky, 2004), figured worlds (Holland et al., 2001) and play (Vygotsky, 2004) and we explain the interrelation between these critical sociocultural processes and Stetsenko's (2017) transformative extension of Vygotskian theory. We then explain why an understanding of literary reading should be expanded to encompass imagination and critical sociocultural processes in theory building and pedagogy.

Uncertainty and imagination

In Vygotsky's (2004) sociocultural theory of change, imagination is at the center of the production of all cultural life. In this view, imagination is not a practice of conjuring something out of nothing in service of creative performance, but rather an ongoing effort to resolve the dissonances that arise in everyday events, images, or meanings (Pelaprat and Cole, 2011). When we are faced with uncertainty or ambiguity, or differences in the socially constructed meanings about who we are, whether in storytelling or everyday life, we experience a pause (in microseconds or longer) during which we "fill the gap" with culturally-historically formed meanings, images, experiences, and projections of what has been and what might be possible—as disruptions of presumed colonizing continuities from past to present to future social conditions (Pelaprat and Cole, 2011; Torres and Medina, 2021; Zuckerman, 2021). Similarly, Rosenblatt and other literary scholars have argued that the indeterminacy of language and the necessarily incomplete representation of life in stories, create openings or gaps to be elaborated on by readers (Herman, 2009; Iser, 1972; Suleiman and Crosman, 2014). Stories, in this sense, unsettle what is taken for granted by proposing alternative visions of past and future landscapes, by allowing us to consider "what if" and "what might be" as characters face decisions and readers, alongside characters, pause to consider the implications of different actions for their own humanity (Bruner, 1986).

Figured worlds

Disrupting worlds through narrative speculation may also be understood through the concept of figured worlds, developed by Holland et al. (2001) in their ethnographic studies of improvisational action and the use of mediating artifacts that introduce change in taken for granted daily interactions. Holland et al., argue that our lives are “figured” by our own and others’ ongoing participation in accepted ways of talking about, relating to, narrating, and interpreting events:

By “figured world” we mean a socially and culturally constructed realm of interpretation in which particular characters and actors are recognized, significance is assigned to certain acts, and particular outcomes are valued over others. These collective “as-if” worlds ... mediate behavior and ... inform participants’ outlooks (p. 32).

Although figured worlds inform participants’ outlooks and may appear to be thoroughly predictable and stable, all parts of the constructed “as if” realm are also open to improvisation and change. As Rene suggests in the opening vignette, beyond the figured world of school-based literature education in which he might be presumed to be a passive learner who has yet to acquire interpretive skills, he can be seen as a popular fanfiction author in the figured worlds of fandom and peer story making where innovations and risk-taking are valued.

Play and figured worlds

Holland et al. (2001) align with Vygotsky’s analysis of play as a primary practice for freeing ourselves from the material constraints of the present, “in order to construct a new reality” (Vygotsky, 2004, p. 11). Through play, youth assert ambiguous meaning about their everyday relationships and ways of being seen by others. Play, however, is not an escape, but rather a way to examine reality by “staging” how the world works or by reworking parts of the world to discover what else might be possible. Vygotsky’s famous example of two children who are sisters, playing “sisters”, illustrates our human capacity to step outside of a figured world to both enjoy the doubled meaning of being and performing as ourselves, while also looking in on what “sisters” means within the boundaries of a figured world (Vygotsky, 2016). In Holland et al. (2001) words, practices of imagination and improvisation displace familiar fields of vision in service of an emerging figured world. Similarly, in their work with young children playing with popular media narratives, Medina and Wohlwend (2014) understand play as a practice of remediating their surroundings to create “as if” worlds “where alternative sets of premises make additional or contradictory identity texts possible” (p. 29).

Critical imagination and a transformative activist stance

Stetsenko’s (2017) scholarship on Vygotsky’s revolutionary theories of social-historical change magnifies the significance of contributing to new worlds, within moment to moment “gaps” and within play and narrative spaces. For Stetsenko, human development depends on orienting toward justice and diversity, toward our histories and our hopes for the future. A passive or unexamined participation in everyday life maintains oppressive relationships and social conditions. Stetsenko’s theory, defined overall as a “transformative activist stance” is a powerful description of youth agency as they move texts through possible meanings and worlds, with an intention to open more space for their own and others’ dignity, creativity, and visions for the future. As Lewis et al. (2021) argue, drawing on Enciso (2017, 2014, 2007) critical analysis of power and history in imagined futures, “imagination is not an individual, spontaneous act of free will, but is constrained by sociocultural conditions that open or close the meaning of artifacts, language, images, and social positions” (Lewis et al.; p. 153). Storytelling and storymaking fill these gaps through play and rehearsals of possible new relationships that forecast liberatory futures while contributing to potentially altered relations in the present.

Defining critical imagination in literary reading

Conceptualizing literary reading through the lens of critical imagination broadens the scope of how we understand youth’s play with and through the mediating means of texts. First it collapses the binary between the consumption and production of stories, by understanding literary reading as (re)shaping figured worlds. It acknowledges that students’ movements within and around stories is agentic work through which they may restory and replay dominant narratives and imagine alternate possibilities outside of these narratives. We argue that by opening up space to multiple narrative modes and platforms, literary study may offer youth the potential to be more agentic authors of their reading and actions in textual and societal contexts. Critical imagination permits a broad understanding of literary reading as a practice of deconstructing and reconstructing social realities and reconfiguring these realities through speculation and play with worlds, social positions, and language. We describe guiding research in these three areas (worlds, positions, and language) in the following sections. This review highlights the literacy practices students are already using to shift the ways their realities are represented and read, and argues that such practices should be at the center of literary instruction in schools.

World making

Youth not only read fiction but also use fiction to establish new routes through their everyday and imagined worlds. The opening vignette, for example, points to the storytellers' movements from an international franchise, MHA, to their own constructed story-world of superheroes and "quirks", and into a real neighborhood "where they get all the good Halloween candy." Here, youth are not only interpreting a beloved story world, but also annotating their everyday places and pathways. Moment to moment interactions and worldmaking among youth reveal the interplay of youth interests, concerns, and joys as they critically (re)imagine and restore the intersections of their everyday and fictional worlds. Literacy researchers have long recognized the power of youth to remake places and their identities through linguistic and symbolic means (Maira and Soep, 2004; Rogers et al., 2014). We focus, in the next section, on research that directly engages literary and other narrative forms for worldmaking.

World making for presence, belonging, and collective action

In our overview of worldmaking and critical imagination, we consider research with youth who are remaking narrative worlds, sometimes but not always facilitated by adults, in schools and communities, for purposes of (1) asserting their presence in story-worlds and related discourses of social life (2) deepening belonging in the present while projecting new social structures for possible futures and (3) constructing collective civic engagement through shared story fandom.

Asserting presence

In an international study of the intersections between children's immigrant experiences and their explorations of Shaun Tan's *The Arrival* (2007), Arizpe et al. (2014) invited children to talk together, draw, and question what it looks and feels like to be a newcomer, entering an unfamiliar world. Although the researchers were primarily concerned with documenting the many literary reading strategies and cultural resources children mobilize to construct rich interpretations of an image-rich, wordless novel, we are interested in the ways children speculated and built worlds together as they clarified and interrogated what it means to "be an immigrant." For example, in the USA-based research, Martinez-Roldán and Newcomer (2011) describe two students, Almah and Ryan's, drawing and storytelling as "performative narratives," during which they imagined themselves in places and times like those of the central character in *The Arrival*. Throughout their dialog about their own drawings and the detailed, surreal images of the world experienced by the book's central character, the two children listened to one another's descriptions and entered one another's imagined worlds to question the significance of details in one another's drawing, describe their feelings, and hypothesize about how the world of passports and offices and lines of people matter in the process of immigration. By inviting children to assert their presence in Tan's storyworld, the researchers assumed an opening or gap for critical imagination, where multiple possible futures were valued, potentially contributing to a fuller explanation of the motives and challenges faced by immigrants who are forced to follow often baffling (and violent) international policies and pathways to entry.

In a storytelling project similar to the one referenced in the opening event, Enciso and Krone (2022) proposed to a racially and ethnically diverse group of 12 and 13-year-olds that they create superhero stories set in their neighborhood, thus inviting youth to assert their presence in fantastic landscapes with imagined powerful resources, through their knowledge of the history and materiality of their streets. Drawing on their love of DC and anime, two project members, Lee and Sebro created an antihero and hero who battled one another across the long corridor of a main thoroughfare, cutting through their neighborhood, and leaving destruction in their wake, as well as hope for the end of gentrification. Lee described his character, Timeless, as an AI robot who made himself, "from the rejected cars, parts, and tore up buildings he scrounges from highway ditches, alleys, and junkyards." Lee seems to imagine his character at the intersection of the pathways between his material experience of the past and the possibility of future equity and change. This is one of multiple stories created by youth over 3 years, as they returned almost weekly to elaborate on their local stories, by reaching for and referencing their extensive knowledge of superhero narratives. The adults in this study actively introduced and sustained a practice of critical imagination among youth by referring to the youth as authors and artists who were contributing to the creative environment of a writers' room in the service of creating superhero stories for their community. For example, instead of asking if youth had finished their scene or character portrait, adults would invite youth to summarize last week's work and ask what they wanted to develop during the current session. Adults would also scribe youth ideas; a practice that lowered adult status and freed youth to focus on ideas rather than the mechanics of recording and editing. Adults' respect for youth knowledge and creativity supported their capacities to narrate their everyday worlds as they also referenced stories they most enjoyed while developing the themes of friendship, mutual care, safety, community stability, and mental health.

Belonging in possible futures

Through practices of critical imagination, "a play world may be taken, ultimately and profoundly, into the "real world", to shape the sites of everyday activity" (Holland et al., 2001, p. 239). By moving through the everyday world, with storytelling as a guide and resource for valuing and projecting future care and strength, youth and adults have been able to collectively build intergenerational knowledge and an orientation to self- and community love (Anzaldúa, 2004; Rowell and Butler, 1997; Woodson, 2003). Two recent research projects exemplify worldmaking for belonging in possible futures.

Literature and education scholar, Toliver (2020, 2021a,b,c), has studied the extensive philosophies and projected futures of Black women who have created poetry, dance, literature and music to understand and elevate their vision of survival and thriving despite the persistence of anti-Black violence. Toliver situates and facilitates young Black girls' storytelling (2020, 2021a,b) in

relation with the works of Black feminist and Afrofuturist writers whose stories engaged them in creating counterstories within “an experiential portal that guides readers to reflect on the current state of the world, to hypothesize about the trajectory of society, and to challenge any possible future that continues the subjugation of Black people” (2021a,b,c; p. 132). In her research with Black girls, Toliver sustains a space for their movement between a figured order of racial hierarchy and what is imaginable and actionable beyond these figurations. Toliver recognizes that, “stories ... are a means to reaffirm the humanity of the storyteller because they share their lives with those willing to listen” (2020, p. 1). “Ultimately,” she continues, “the Black narrative tradition imbues truth in the tale, as Black storytellers draw from history, real-life experience, personal feeling, and imagination to create intricate testimonies for their listeners to witness” (Toliver, 2021a,b, p. 2). For Toliver, the relationships among teller, present adults, and storied listeners and witnesses make possible a sacred space for youth to imagine themselves into a future where Blackness, across time, is loved and fully alive in a society designed exquisitely by them for their belonging and freedom.

The second study is described by the author, Medina (2021), as a “(re)membering” (Dillard, 2012) of her own embodied explorations and performances of possible decolonial Puerto Rican futures alongside the anticolonial literacy work of young artist activists “who use their arts to enact, re-create, and theorize with hidden and marginalized knowledges ... through playful engagements with social realities” (p. 137). Medina describes the improvisational performance of a well-known Puerto Rican youth troupe, Jóvenes del 98, as they deconstruct capitalist and evangelical Christian discourses that reinforce practices of colonizing and subjugating decolonial histories and voices. Working as both a memoirist of her own artist/activist identities and as a chronicler of the troupe’s improvisations, Medina shows how youth constructed a montage of episodes, prayers, pronouncements, and projections that drew on colonial and neoliberal social discourses while keeping open the interpretive space through which their embodied, critical imagining and knowledge of colonial and decolonial worlds could collide, rearrange, and thoroughly disrupt “the complex social ecology and trajectories of colonialities” (p. 151). Like other scholar-artist-activists (Sealey-Ruíz, 2020; Dillard, 2021), Medina’s research exemplifies the necessary excavation of educators’ personal ways of knowing and unknowing, through our embodied experiences of the colonizing discourses and materialities we seek to disrupt. By creating a (re)membered story of improvisation, Medina illustrates the labor and pleasure of not only naming but embodying stories, as we improvise our ways into belonging in more equitable and just possible futures.

Collective engagement in digital worldmaking

Youths’ agency as world makers through digital tools and platforms has been the subject of international literature and literacy studies for more than two decades (De Kloet and Fung, 2016; Knobel and Lankshear, 2008; Koh and Benson, 2011; Lewis and Fabos, 2000; Maira and Soep, 2004; Pathak-Shelat and DeShano, 2014; Rogers et al., 2014; Stornaiuolo and Whitney, 2018). Scholars of popular media have closely followed the ways youth and adults have collectively developed platforms for extending stories, games, and films for their multi-layered, global immersive interaction (de Castell and Jenson, 2003) and, more recently, for repurposing their fandom for civic imagination and engagement (Jenkins et al., 2020). Several international projects have recognized youth fandom as a collective and connected practice for organizing social justice initiatives (Ito et al., 2015). One of the more striking, literature-based online collectives, named the Harry Potter Alliance (recently renamed Fandom Forward), was initiated by fans who used their affinity for the Harry Potter series to “bring together narrative elements from social justice and the Harry Potter content world. For example, their annual book drive to support libraries is called ‘Accio Books,’ referencing a charm in the Harry Potter story that summons objects from a distance” (Ito et al.; p. 18). Their website also provides toolkits in support of campaigns to address “voting, body image, marriage equality, fair trade, literacy, child slavery, disaster relief, and hunger” (p. 20).

Working with the idea of “hacking” and “remixing” storyworlds, Dando et al. (2019) developed an 8-week project with five New York City high school girls of color who were assisted by Black and African diaspora “professional artists, designers and educators to develop tangible artifacts that critically imagine a future” (p. 157) that, as many authors of comics, plays, and films have done, rewrite the legacy of white supremacy. The focus of their redesign was Wakanda, the fictional, Afrofuturist home of the uncolonized and highly technologically, politically and artistically sophisticated African people featured in the Black Panther comics. The project leaders and artists invited participants “to create their Wakanda—to bring histories, family experiences, art, textiles, style, and technology together to create a new home that would address the problems they face and elevate the values they espouse” (p. 158). Throughout the project, the girls read, listened to music and participated in discussions about Black artistic and cultural expression, and began creating medallions that symbolized their understanding of complex cultural histories and trajectories within and projected beyond the everyday experiences of racism and colonization. Aida, a 15-year-old participant, used a 3-D tool to create a medallion that centered the continent of Africa with lines extending out to symbolize the connections that western societies should recognize. She also expressed anger about silences and the absence of voices, so that “they stay silent and the only people who really talk are the wealthy!” (p. 159). Dando et al. facilitated a project of worldmaking inspired by Wakanda’s images and ideology that clearly drew participants and artists together in a celebration of critical imagination, where playing with digital materials, artifacts, sound, symbols, and stories opened dialog, personal discovery and new forms of representation for inscribing the self into past, present and future collectives.

Summary

Practices of critical imagination reduce the pull of binaries and hierarchies, to create a space for exploring the possible relationships between textual storyworlds and everyday realities. Youth may mobilize their everyday resources, along with their political insights about inequities and injustices to create new ways of organizing for a future they want to bring forward. The three dimensions and related studies of worldmaking described here—asserting one’s presence, belonging in possible futures, and collective

engagement—illustrate the interdependent and overlapping ways youth and researchers value stories as speculative framings for refiguring ways of being understood in fictional and real worlds.

Positioning and repositioning

Attention to positioning in and through storyworlds, reveals the way stories momentarily free social actors from the constraints of the present, creating alternative identity trajectories. In the opening vignette, Kiana uses the space produced by the story to engage in such repositioning, as she briefly toggles into the first-person role of her superhero who is trick or treating in their superhero costume. “I swear I’m cool,” she playfully performed, speaking to an imagined person passing out candy. Aidan also plays with possible new positions as he projects a future in which this story is published and produced and students are valued as authors and creators. “Do you know if we publish this we would not get paid a lot because it’s all five of us?” Aidan cautioned. As noted above, it was only in exit interviews that we realized this position of published author was not a future projection, but rather something Rene was already stepping into the digital platform, Wattpad. Stories, this group suggests, open alternate positionings in not only imagined but also in material worlds.

Writing selves into texts

As students refigure worlds through story, they actively step into new positions, getting unstuck from static roles and categorizations. We highlight how students use stories and storytelling to construct new communities for themselves and others as they build windows (Sims-Bishop, 1990) into texts, new media, and school spaces. In addition to Thomas’ work on restorying (2016), her recent text, the *Dark Fantastic* (2020), describes how popular young adult literature and media like *The Vampire Diaries* and *The Hunger Games* narratively perpetuate violence and harm against particularly girls of color, using discriminatory tropes and practices from everyday life. However, Thomas also describes how fans are reshaping these literary worlds, using apps like Wattpad, to insert themselves into problematic storyworlds in ways that flip and counter this harm. Low’s (2017) study describes similar work students conducted in a “Comics Inquiry Community,” an after-school club in a Philadelphia school, where color-blindness was the norm. Low details how students’ talk about superheroes like Miles Morales in *Spiderman* and Kamala Khan in *Ms. Marvel* opened up discussions of race that had previously been silenced. In these discussions, students challenged the white male norm position in superhero texts and society and imagined new possibilities for whom gets to be a superhero, taking on interpretative positions as narrative creators and social agents. Both of these studies suggest that literary reading entails not only identifying the narrow subject positions available in dominant narratives and existing genres, but also deliberate work to form new textual subject positions that better represent readers’ desires, identities and experiences.

Rewriting spaces

Students may use storytelling and storyworlds to (re)position themselves in not only textual spaces but also school spaces that have historically excluded them. Campano et al. (2013) describe how students they worked with in an over-tested and over-surveilled elementary school used a variety of texts including Louis Sachar’s *Holes*, to assert their agency as “organic intellectuals” (Gramsci, 1971) who contested dehumanization and marginalization. These students shifted from critics to creators as they engaged in their own library-building projects that communicated different norms for justice and inclusion in their community. Johnson’s (2014, 2017) work also demonstrates how storytelling and writing may open safe spaces on the “edges” (2014) of school for students to share, explore, and affirm their identities. Working from her own position as a professor-in-residence at West High School, Johnson (2017) describes Writing@West, an elective program in which Black queer youth wrote about themselves in ways that were not supported in their academic coursework. Johnson describes students’ literary production in this elective course as epistemological and ontological, as students in this space shared stories that asserted their humanity and challenged dominant ways of knowing in school. In Writing@West, Johnson explains “students not only named their experiences, but were involved in rewriting a world where they could enact several aspects of their identities” (Johnson, 2017, p. 29). Literary study, in this view, actively attends to the world-making that occurs when students are invited to write themselves into texts and schools.

Repositioning through ruptures

Studies have also highlighted the potential of rhetorical ruptures or gaps in literary reading (Torres and Medina, 2021; Schey and Blackburn, 2019). Such gaps create disjunctures that challenge the sociocultural frameworks students use to make sense of both stories and worlds. Schey and Blackburn (2019) describe the experience of reading *Brooklyn, Burning*, with students in a LGBTQ themed literature course in a charter arts high school. Neither the main character in this text nor this character’s love interest is gendered, a rupture that, Schey and Blackburn argue, offered productive potential when students were supported in engaging in dialog to think through this gap. When the class gave sustained attention to this surprising textual feature, students began to shift their own positionalities by deconstructing dominant expectations around gender and sexuality both in and outside of this text. Like Schey and Blackburn, Torres and Medina (2021) analyze the transformative potential of textual ruptures. In their analysis of Puerto Rican children’s books about Hurricane Maria, they identify several literary devices and rhetorical patterns that challenge a western way of understanding this event and open decolonial readings of both Hurricane Maria and Puerto Rico broadly. While previous studies examine how students use texts to shift roles, these studies highlight how texts may push readers into new interpretative positions, encouraging them to imagine different ways of reading not only texts and characters but also worlds.

Language play

Within story spaces, words and signs are open to unexpected meanings and usages. For example, in our opening vignette, and throughout his larger storytelling Rene drew language from *My Hero Academia*, using elements like “quirks” from this textual universe as tools that he could add to and reshape in his own created world. As Rene noted in his exit interview, “I like to create things in general. You decide how [the characters] act and what they do. It gives you a sense of power because you have control over the situation and stuff.” In addition to readers’ practices of world-making and repositioning, we are interested in how youth broadly assume power over their circumstances by creating new meanings of words and signs within imagined storyworlds.

Freeing social signs

Youth use play to free social signs from previous meanings and propose new ones, refiguring worlds in the process. Conceptualizing what counts as literary work through the lens of critical imagination permits an attention to how youth are already engaging in such play in small scale acts of linguistic innovation and subversion. Following scholars who have highlighted the way young people creatively use and refashion language to challenge and bypass white western linguistic norms (Richardson, 2013; Baker-Bell, 2020; Martínez and Morales, 2014; Machado and Gonzales, 2020), we see the language play that happens within storytelling as literary work. Orellana’s (2016) work with B-Club, an after-school program grounded in sociocultural theories of imagination and play, exemplifies the relationship between language play and storytelling. Youth created norms, rules and routines for their space, creating opportunities for play and performance and for engaging in novel ways with each other and their material surroundings. Instead of conceptualizing literate activities as practice for a future assessment, students in B-club used language and stories to accomplish real action within their local and global communities. Within this space, Orellana points out the shifts in language and semiotics when students tried to get things done, using processes of critical imagination to meet real ends in the immediate future. Within this translingual, transcultural group, such shifts often took the form of translanguaging and working across and between linguistic borders. However, Orellana also indicates how material signs, like a detergent box became a soapbox for youth speakers, who adopted and performed new playful meanings for material and language within this space.

Figurative language

Figured world-making also has a relationship to the study of figurative language that has been a long-standing feature of literary reading in ELA classrooms. Studies have demonstrated youth’s figurative meanings may function as critical tools and that students’ esthetic transactions with literature hold sociopolitical weight (Cai, 2008; Lewis, 2000; Wissman, 2018). Following this work, we wonder how students are agentially reshaping language through figurative meanings and readings, challenging on a microscale, the larger frames or heuristics we use to read our worlds. A pivotal study in this area is Lee’s (2007) exploration of students’ literary reading in a Chicago area high school, where Lee and others instituted curricula based on principles of cultural modeling. Lee describes how the African American students she worked with already used figurative meaning-making practices in their everyday language and communication, a repertoire of linguistic and literate skills that she drew into students’ literary work. Working from this foundation, students in her classroom moved beyond literal meanings to open up figurative interpretations of language in texts that included Toni Morrison’s *Beloved*. Through figurative meaning-making students in this classroom shifted the frames their peers were employing to re-figure these textual worlds. Everett’s (2018) recent study of a summer writing program she facilitated for high school students on critical metaphor also describes how youth use figurative language to shift readings of their worlds. Everett focuses on one student, Shawn, who purposefully used metaphor to critique the school to prison pipeline that criminalizes Black youth. For Shawn, metaphor was more than “a mere literary device” (p. 44) but rather a way to challenge how his audience understood school. Across both of these studies, students’ figurative language use was not bound to literary texts but rather used as a critical sociocultural tool to reimagine words and signs within real worlds.

Conclusion

In this reconceptualization of literary reading, we assert the pressing need for youth to imagine together and create worlds in which they are agentic and bold. We argue for an expanded understanding of literary reading and interpretation that encompasses practices of critical imagination as both esthetic and political work. Stetsenko and Ho’s (2015) analysis of play as cultural practice best summarizes the relationship between imagining and producing worlds: “Play is a pathway for children to discover themselves and the world in a bidirectional spiral of mutual becoming” (p. 225). Researchers and educators who recognize the “serious work” (Stetsenko and Ho, 2015) of play as a practice of exploration and critical imagination for refiguring selves and worlds, will be able to support youth worldmaking, positioning and repositioning, and refigurations of language. Within their worldmaking, we can support young people as they not only move between their experiences and the world of the story, but also strive to be seen and understood alongside peers making the same effort. As “worldmakers” youth and adults reflect on and project their belonging in past, present, and future worlds, by reimagining oppressive social discourses through improvisation and storytelling, in performances and published works for audiences who are more than passive listeners or readers, but rather accompanying witnesses to their vision and (re)membering. Working with digital platforms and tools, youth are also mobilizing their passion for story themes and inventions to organize for social change, and by doing so, they directly challenge any assumption that reading is an isolated individualistic escape from reality—or a form of school work, subject to categories of achievement and failure.

Critical imagining supports a view of readers as “social actors” (Holland et al., 2001) who use stories, language, and artifacts to reposition themselves and others and construct alternative identity trajectories for possible ways of being, knowing, and doing in the present and the future. As described in the opening of this chapter, youth are attuned to and critical of narrow subject positions and therefore, agentially restory these representations through practices of inscription that displace and redefine characterizations, settings, and plots in ways that reflect their own passions and identities. Youth also interrogate and rewrite stories, including superhero narratives, to relocate themselves not only in storyworlds but also in their everyday school spaces, where they position themselves as authors and curators. Researchers and educators can also support dialogic inquiry with youth as they examine the ruptures created by authors and artists, through which they call into question taken for granted social positionings. Finally, by reconceptualizing literary reading as a practice of critical imagination we can better recognize language play as an especially literary form of world refiguration. Youth enjoy and produce linguistic meanings and forms by freeing language from singular or limited meanings, as they also assume power to redefine and manage meanings in real and imagined worlds. Researchers and educators can pay attention to and support youth in disrupting limiting discourses about whose language and whose meaning has value in shaping storyworlds and everyday relationships.

Youth who are witnessing the political violence of book banning and white supremacist ideology on the doorsteps of their homes, libraries, and schools, need a harbor made by adults who will safeguard a play space for their imagined present and future selves. When literary reading is understood as a practice of critical imagination, youth are centered as agentic players with worlds and words, whose capacities for restorying, worldmaking, repositioning, and linguistic reshaping, realign them with stories, where they may enter into new worlds, take new positions, reassign meanings and question what is possible.

References

- Anzaldúa, G., 2004. *Borderlands/La Frontera*. Aunt Lute Press.
- Arizpe, E., Colomer, T., Martínez-Roldán, C., 2014. Visual Journeys Through Wordless Narratives: An International Inquiry With Immigrant Children and the Arrival. A&C Black.
- Baker-Bell, A., 2020. *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*. Routledge.
- Bakhtin, M.M., 2010. *The Dialogic Imagination: Four Essays*. University of Texas Press.
- Bruner, J., 1986. *Actual Minds, Possible Worlds*. Harvard University Press, Cambridge, MA.
- Cai, M., 2008. Transactional theory and the study of multicultural literature. *Lang. Arts* 85 (3), 212–220.
- Campano, G., Ghiso, M.P., Sanchez, L., 2013. “Nobody knows the ... amount of a person”: elementary students critiquing dehumanization through organic critical literacies. *Res. Teach. Engl.* 98–125.
- Dando, M.B., Holbert, N., Correa, I., 2019. Remixing Wakanda: envisioning critical afrofuturist design pedagogies. In: *Proceedings of FabLearn*, pp. 156–159, 2019.
- De Castell, S., Jenson, J., 2003. OP-ED serious play. *J. Curric. Stud.* 35 (6), 649–665.
- De Kloet, J., Fung, A.Y., 2016. *Youth Cultures in China*. John Wiley & Sons.
- Dillard, C.B., 2012. Learning to (Re)member the Things We’ve Learned to Forget: Endarkened Feminisms, Spirituality, and the Sacred Nature of Research and Teaching. Peter Lang.
- Dillard, C., 2021. *The Spirit of Our Work: Black Women Teachers (Re) Member*. Beacon Press.
- Dyson, A.H., 1996. Cultural constellations and childhood identities: on Greek gods, cartoon heroes, and the social lives of schoolchildren. *Harv. Educ. Rev.* 66 (3), 471–496.
- Enciso, P., 1994. Cultural identity and response to literature: running lessons from “Maniac Magee”. *Lang. Arts* 71 (7), 524–533.
- Enciso, P., 2003. Reading discrimination. In: *Making Race Visible: Literacy Research for Cultural Understanding*, pp. 149–177.
- Enciso, P., 2007. Reframing history in sociocultural theories: toward an expansive vision. In: Lewis, C., Enciso, P., Moje, E. (Eds.), *Reframing Sociocultural Research on Literacy: Identity, Agency, and Power*. Routledge, New York, pp. 49–74.
- Enciso, P., 2014. *Sociocultural Theory & Imagination*. Invited presentation at the University of Minnesota.
- Enciso, P., 2017. Stories lost and found: mobilizing imagination in literacy research and practice. *Lit. Res.* 66 (1), 29–52.
- Enciso, P., Krone, B., 2022. Storytelling and proleptic gaps: reimagining inequities in the Mount. *Res. Teach. Engl.* 56 (4).
- Everett, S., 2018. Untold stories: cultivating a consequential approach to writing with a black male student through a critical approach to metaphor. *Res. Teach. Engl.* 53 (1), 34–57.
- Gramsci, A., 1971. The modern prince. In: *Selections From the Prison Notebooks*, pp. 123–205.
- Herman, D., 2009. *Basic Elements of Narrative*. John Wiley & Sons.
- Holland, D., Lachicotte Jr., W., Skinner, D., Cain, C., 2001. *Identity and Agency in Cultural Worlds*. Harvard University Press.
- Hooks, B., 1991. Narratives of struggle. In: Mariani, P. (Ed.), *Critical Fictions: The Politics of Imaginative Writing*. Bay Press, pp. 53–61.
- Iser, W., 1972. The reading process: a phenomenological approach. *New Lit. Hist.* 3 (2), 279–299.
- Ito, M., Soep, E., Kligler-Vilenchik, N., Shresthova, S., Gamber-Thompson, L., Zimmerman, A., 2015. Learning connected civics: narratives, practices, infrastructures. *Curric. Inq.* 45 (1), 10–29.
- Jenkins, H., Peters-Lazaro, G., Shresthova, S. (Eds.), 2020. *Popular culture and the civic imagination: case studies of creative social change*. NYU Press, New York.
- Johnson, L., 2014. “Can I come in here?” Winston’s discovery of edge-of-school spaces and meaningful literacy engagement. *Appl. Financ. Econ.* 21 (3), 201–214.
- Johnson, L.P., 2017. Writing the self: black queer youth challenge heteronormative ways of being in an after-school writing club. *Res. Teach. Engl.* 13–33.
- Knobel, M., Lankshear, C., 2008. Remix: the art and craft of endless hybridization. *J. Adolesc. Adult Lit.* 52 (1), 22–33.
- Koh, A., Benson, P., 2011. Exploring pedagogies in the popular culture and education nexus. *Pedagogies* 6 (2), 123–129.
- Lee, C.D., 2007. *Culture, Literacy, and Learning: Taking Bloom in the Midst of the Whirlwind*. Teachers College Press.
- Lewis, C., 2000. Critical issues: limits of identification: the personal, pleasurable, and critical in reader response. *J. Lit. Res.* 32 (2), 253–266.
- Lewis, C., Fabos, B., 2000. But will it work in the heartland? A response and illustration. *J. Adolesc. Adult Lit.* 43 (5), 462–469.
- Lewis, C., Crampton, A., Scharber, C., 2021. The sociocultural role of imagination in critical digital literacy. *Pedagogies* 16 (2), 154–172.
- Low, D.E., 2017. Students contesting “colormuteness” through critical inquiries into comics. *Engl. J.* 19–28.
- Machado, E., Gonzales, G.C., 2020. “I can write in my language and switch back and forth”: elementary teacher candidates experiencing and enacting translanguaging pedagogies in a literacy methods course. *Lit. Res.* 69 (1), 211–229.
- Maira, Soep, 2004. *Youthscapes: The Popular, the National, the Global*. University of Pennsylvania Press.
- Martínez, R.A., Morales, P.Z., 2014. ¿Puras Groserías?: rethinking the role of profanity and graphic humor in Latin@ students’ bilingual wordplay. *Anthropol. Educ. Q.* 45 (4), 337–354.
- Martínez-Roldán, C.M., Newcomer, S., 2011. “Reading between the pictures”: immigrant students’ interpretations of “the arrival”. *Lang. Arts* 88 (3), 188–197.
- Medina, C.L., 2021. Barruntos: youth improvisational work as anticolonial literacy actionings in Puerto Rico. *Res. Teach. Engl.* 56 (2), 132–154.

- Medina, C., Wohltwend, K., 2014. *Literacy, Play and Globalization: Converging Imaginaries in Children's Critical and Cultural Performances*. Routledge.
- Orellana, M.F., 2016. *Immigrant Children in Transcultural Spaces: Language, Learning, and Love*. Routledge.
- Pathak-Shelat, M., DeShano, C., 2014. Digital youth cultures in small town and rural Gujarat, India. *New Media Soc.* 16 (6), 983–1001.
- Pelaprat, E., Cole, M., 2011. "Minding the gap": imagination, creativity and human cognition. *Integr. Psychol. Behav. Sci.* 45 (4), 397–418.
- Richardson, E., 2013. Developing critical hip hop feminist literacies: centrality and subversion of sexuality in the lives of Black girls. *Equity Excell. Educ.* 46 (3), 327–341.
- Rogers, T., Winters, K.L., Perry, M., LaMonde, A.M., 2014. *Youth, Critical Literacies, and Civic Engagement: Arts, Media, and Literacy in the Lives of Adolescents*. Routledge.
- Rosenblatt, L., 1978. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Southern Illinois University, Carbondale, IL.
- Rowell, C.H., Butler, O.E., 1997. An interview with Octavia E. Butler. *Callaloo* 20 (1), 47–66.
- Schey, R., Blackburn, M., 2019. Queer ruptures of normative literacy practices: toward visualizing, hypothesizing, and empathizing. *Res. Teach. Engl.* 54 (1), 58–80.
- Sealey-Ruiz, Y., 2020. *Love From the Vortex & Other Poems*. Kaleidoscope Vibrations, LLC.
- Sims Bishop, R., 1990. Windows, mirrors, and sliding glass doors. *Perspectives* 6 (3), ix–xi.
- Smagorinsky, P., 2001. Rethinking protocol analysis from a cultural perspective. *Annu. Rev. Appl. Ling.* 21, 233–245.
- Solorzano, D.G., Yosso, T.J., 2001. Critical race and LatCrit theory and method: counter-storytelling. *Int. J. Qual. Stud. Educ.* 14 (4), 471–495.
- Stetsenko, A., 2017. *The Transformative Mind: Expanding Vygotsky's Approach to Development and Education*. Cambridge University Press., Massachusetts.
- Stetsenko, A., Ho, P.C.G., 2015. The serious joy and the joyful work of play: children becoming agentive actors in co-authoring themselves and their world through play. *Int. J. Early Child.* 47 (2), 221–234.
- Stornaiuolo, A., Whitney, E.H., 2018. Writing as worldmaking. *Lang. Arts* 95 (4), 205–217.
- Suleiman, S.R., Crosman, I. (Eds.), 2014. *The Reader in the Text: Essays on Audience and Interpretation*, vol 617. Princeton University Press.
- Tan, S., 2007. *The Arrival*. Arthur A. Levine Books, New York.
- Thomas, E.E., Stornaiuolo, A., 2016. Restorying the self: bending toward textual justice. *Harv. Educ. Rev.* 86 (3), 313–338.
- Toliver, S.R., 2020. Can I get a witness? Speculative fiction as testimony and counterstory. *J. Lit. Res.* 52 (4), 507–529.
- Toliver, S.R., 2021a. Afrocarnival: celebrating black bodies and critiquing oppressive bodies in Afrofuturist literature. *Child Lit. Educ.* 52 (1), 132–148.
- Toliver, S.R., 2021b. Freedom dreaming in a broken world: the black radical imagination in black girls' science fiction stories. *Res. Teach. Engl.* 56 (1), 85–106.
- Toliver, S.R., 2021c. *Recovering Black Storytelling in Qualitative Research: Endarkened Storywork*. Routledge.
- Torres, F.L., Medina, C.L., 2021. Cuentos combativos: decolonialities in Puerto Rican books about María. *J. Lit. Res.* 53 (2), 242–264.
- Vygotsky, L.S., 1987. Thinking and speech. In: *The Collected Works of LS Vygotsky*, vol. 1, pp. 39–285.
- Vygotsky, L.S., 2004. Imagination and creativity in childhood. *J. Russ. East Eur. Psychol.* 42 (1), 7–97 (First published 1967).
- Vygotsky, L.S., 2016. Play and its role in the mental development of the child. *Int. Res. Early Child. Educ.* 7 (2), 3–25 (First Published 1966).
- Winn, M.T., 2010. "Our side of the story": moving incarcerated youth voices from margins to center. *Race Ethn. Educ.* 13 (3), 313–325.
- Wissman, K.K., 2018. Teaching global literature to "disturb the waters". *Engl. Educ.* 51 (1), 17–48.
- Woodson, J., 2003. Who can tell my story. *Stories matter: the complexity of cultural authenticity in children's literature* 41–45.
- Zuckerman, E., 2021. 27 "Without my city, where is my past?". In: *Popular Culture and the Civic Imagination*. New York University Press, pp. 281–290.

Humanizing language and literacy in education with dramatic inquiry

Brian Edmiston, The Ohio State University, Columbus, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Dramatic inquiry pedagogy	347
Dramatic imagination	348
Transforming classroom reality	349
Dramatic conventions	349
Dramatic dialogic inquiry	349
Different approaches to dramatic inquiry	350
Four approaches	350
Framing	350
Humanizing engagement and purpose	350
Engagement with relational agency	351
Purpose and audience	351
Humanizing classroom communities	351
Humanizing students' connections with the curriculum	352
Transforming classroom communities	353
Humanizing relationships and identities	353
Transformative crucible	354
Transformative Dialogue	354
Humanizing identities	355
Conclusion	355
References	356

What else is drama about other than examining “what it is to be human” in the safety of the made event?

Heathcote (2009, n.p.).

There is an urgent need internationally to humanize education as an antidote to the pervasive drift toward dehumanizing standardization. Dramatic approaches have long been used in classrooms to contextualize language and literacy learning and teaching. I propose a rediscovery of the radical pedagogy of Dorothy Heathcote (1925–2011), a renowned British educator, who viewed language and literacies as integral to a holistic approach to education. In this essay, I develop an argument for how the pedagogy Dorothy Heathcote invented could be used to transform education.

I introduced the over-arching term *dramatic inquiry* to characterize Heathcote's deceptively simple yet highly engaging and complex inquiry-based approaches to drama as education (Aitken, 2021; Edmiston, 2010, 2014). I argue that she was a visionary educator whose practice has been over-simplified in its application, under-theorized, and under-appreciated for its humanizing potential. Heathcote was a pioneer in advocating for teachers to educate by dramatizing, dialoguing, and inquiring with young people rather than merely using drama to facilitate student learning determined by an official curriculum or pre-determined standards. I argue that dramatic inquiry pedagogy should be documented, theorized, and researched in ways that recognize its complexity, its humanizing approach to language and literacy teaching and learning, and its transformative potential in any classroom. I illustrate, using a recent example of my own teaching (Edmiston and Towler-Evans, 2022).

Dorothy Heathcote was a humanizing educator. For her, education should always be concerned with the human condition. She stressed that, “All my ideas ... are rooted in the remarkable variations of human encounters” (Heathcote, 2009, n.p.) as teachers explore with young people “what sort of society we want” both within and beyond the walls of a classroom (Heathcote, 1983, p. 693). Heathcote pioneered the use of drama, along with the other arts, to humanize classroom communities and educational pursuits. For over half-a-century she showed how young people and topics can come to life in schools that too often deaden learning and teaching (Bolton and Heathcote, 1999; Heathcote and Bolton, 1995; Johnson and O'Neill, 1984; O'Neill, 2014; Wagner, 1999). As discussed by Lacey and Wooland (2009), her work was grounded in dramatic form and could be considered radical theater practice. Her influence became and remains global. For example, her work has inspired teacher networks in New Zealand (Aitken, 2021) and the UK (<https://www.mantleoftheexpert.com>), summer schools in Jordan (<http://qattanfoundation.org/en/qattan/media/news/ninth-drama-education-summer-school-concluded>), and international conferences (<https://stayhappening.com/e/dorothy-heathcote-now-E3LUS030H4S3>).

Heathcote's dramatic inquiry pedagogy has engaged groups of people of all ages—from children both not yet in school and those in classrooms, to managers in industry, to adolescents in detention facilities and hospital residents classified as mentally handicapped. She focused on events and endeavors that were meaningful to them. She celebrated people's humanity and rejected deficit dehumanizing assumptions and practices within schooling and in society. She approached education holistically,

emphasizing that what teachers wanted young people to learn and how they hoped they might be in the world could not be separated from the students' actual experiences in classroom communities.

Heathcote's intention was always to create rigorous yet joyful communities of critical inquiry into narratives that honor the complexity and dignity of human struggles. These narratives emerge from actual and imagined societies depicted in literature, historical documents, and contemporary accounts encountered via media, in person, and in dramatized events. Though her practice might look very different in varying situations, Heathcote's educational values and ethical principles were constant. She treated every person with dignity as a whole human being of equal value within any group because of their unique life experiences, imagination, energy, and potential to contribute productively to all classroom activities. She worked with young people as colleagues using active and dramatic approaches to entice and include all students in meaningful ventures. Working collaboratively as an artist, she created contexts in which everyone could form more complex understanding of their own and others' humanity (and inhumanity) as they extended their everyday experiences in imagined events, reflected via extended dialogue, and formed pedagogical relationships within an inclusive classroom community. Heathcote was respectful of all teachers and honored the complexity of educators' lives, while being adamantly opposed to mindless, standardized, or oppressive schooling tasks and practices as well as to institutionalized segregation whether on the basis of age, ethnicity, race, native language, test results, or any presumed ability or disability. She wanted teachers to know what they stood for. She spent her professional career, as a professor and guest teacher in classrooms world-wide, advocating for transformative changes in education believing that schools should be reformed, not at the directive of bureaucratic or corporate interests, but rather in response to the personal, social, and cultural human needs of young people and of a democratic society.

Heathcote died in 2011 before the rise of dehumanizing post-truth uses of social media, polarizing attacks on mainstream news, and authoritarianism on both sides of the Atlantic that fueled Brexit and Trumpism. Having lived through depressions, wars, and assaults on democracy, Heathcote would have taken a long view encouraging educators working within institutional constraints to continue building vital communities of truthful humanizing education whether in person or online. She would have reminded teachers that dramatic inquiry can be transformative when premised on equality, while critiquing social and structural inequities and struggling to imagine and be more responsive to the world as others perceive it. She would have advocated for dialogue among competing viewpoints intending to develop ethical frameworks for understanding and promoting the common good within and beyond each classroom.

Dramatic inquiry pedagogy

It is humanity—ours and that of our students too, in all its complexity—and justice that are, and should be, at the core of teaching and teacher education.

Philip et al. (2019, p. 260).

Dramatic inquiry is a holistic humanizing pedagogy that contextualizes and extends language uses and literacy practices in communities within cultural worlds that are actual and imagined. Heathcote rejected education reduced to teachers presenting abstractions or individualized cognitive tasks detached from the embodied, affective, and social lives of actual people within cultures past, present, and future. She proposed and demonstrated a radical humanizing teaching stance across her practice: present any subject contextualized in events in the lives of actual and imagined people in society that can be dramatized, interpreted, celebrated, and critiqued from multiple competing positions via dialogue among young people and adults and across time in collaborative inquiry. Language and literacies are always contextualized in the dialogic interactions, meaningful purposes, and humanizing relationships among young people located within, and shaped by, human events experienced not only in the actual community of the classroom but also created and dramatized from communities within imagined and actual worlds.

Dramatic inquiry pedagogy aligns with established and recent humanizing theoretical conceptualizations of language and literacy education. A sociocultural and historical framework has long been used to conceptualize how dramatic approaches enhance holistic social, cultural, and critical approaches to language and literacy education (Britton, 1972; Davis et al., 2015; Edmiston and Enciso, 2002; Medina and Campano, 2006; Moffett, 1987; Schneider et al., 2006; Wagner, 1996; Wolf et al., 1997). Conceptualizing language use as dialogic meaning-making (Bakhtin, 1981) humanizes education in the sense that across social and cultural verbal and non-verbal interactions people are "learning to learn, think and thrive in the context of working with multiple perspectives and ultimate uncertainty" (Wegerif, 2015). Using dramatic inquiry to engage in reading and composing texts can be understood as inherently imaginative (Enciso, 1996, 2017). Dramatic inquiry embraces the uncertainty of open-ended inquiry in improvised dialogic interactions from competing critical positions on events (Edmiston and Enciso, 2002; Enciso et al., 2016). Rather than reducing language and literacies to representations of what individuals know and understand, human beings can be conceptualized more holistically as "doing-knowing-beings" who are "becoming" language users in society by what we are doing together and how we are being in relation to others as well as by what we are coming to know and understand, collectively as well as personally (Stetsenko, 2017). The recent affective turn, applying affect theory to literacy studies, recognizes what Heathcote always assumed about language and literacy: teachers must become "alive to its affect" as well as "the ethics" of the "visceral experiences" on living bodies (Leander and Ehret, 2019, p. 9). Thus, as knowing-doing-beings we are always sensing, feeling, and acting as well as thinking. A

linguaging perspective has been proposed to stress the social, emotional, embodied, and relational nature of language as social action (Beach and Bloome, 2019; Beach and Beauchemin, 2019). Following Heathcote (1983), as Beach and Beauchemin (2019) put it, “a person cannot be a person alone but rather enacts being human through [building] their connections and relationships with others” (p. 4). Conceptualizing dramatic inquiry as humanizing education embraces the recent posthumanism orientation toward literacy. As Kuby et al. (2019) make clear, posthumanism is not a rejection of a humanistic commitment to, for example, equality, solidarity, community-bonding, social justice, and equity. Posthumanism is “a materialized way of knowing/ becoming/doing the world and (producing) literacies” (p. 6). It assumes that, “humans and non-humans are intra-active agents in the world’s becoming” (p. 8). Dramatic inquiry provides educators with tools to embody, experience, and interpret interactions among human, other-than-human, material, and even structural dimensions of events in actual and imagined societies as young people inquire into who they might become in relation to others.

I propose three significant educational benefits of dramatic inquiry pedagogy that I use to structure the remainder of the essay: humanizing engagement and purpose; humanizing classroom communities; and humanizing relationships and identities. First, I outline two core transformative and humanizing dimensions: dramatic imagination and dramatic dialogic inquiry. Throughout, I weave an illustrative example of my recent teaching in a UK classroom using dramatic inquiry with 10- and 11-year-old children who were already focused on the topic of people who are refugees.

Dramatic imagination

Dramatic inquiry is dramatic in the sense that, like the pretend play of children and adult role-play, young people and teachers collectively use dramatic imagination to embody, enact, and improvise interactions within co-constructed events. Heathcote stresses that as teachers plan, they use dramatic imagination to envision other worlds to be explored by dramatizing events (Heathcote, 2009). I extend the term to include how people collectively enact what they imagine in dramatic inquiry. Dramatic imagination exponentially expands our capacity for dialogue about any subject. In contrast to personal imagination that is invisible to others, with dramatic imagination people of any age embody a collective use of social imagination to interact as if social interactions in imagined events are actually happening. Over time, a coauthored imagined world is created, contained and constrained by fictionalized social situations.

The significance of dramatic imagination as a collaborative use of imagination contextualizing social uses of language and literacy extends the understanding that playing as imagination in action is pivotal in early social and language development (Vygotsky, 1978). Dramatic imagination can engage students in extensive playful, yet serious, collaborative explorations of the actual and possible challenges, responsibilities, responses, and consequences faced by people of all ages interacting in situations across time and space located as if in another cultural world. As Heathcote (2010) stresses, “Dramatic activity is concerned with the crises, the turning points of life, large and small, which cause people to reflect and take note” (p. 74).

Dramatic inquiry can focus on understanding, authoring, and humanizing text. Dramatic imagination extends Bakhtin’s core idea of the dialogic imagination of speakers, writers, and readers (Bakhtin, 1981; Edmiston and Enciso, 2002; Edmiston, 2014; Edmiston and Beach, 2017). With dramatic imagination, dialogic interactions are dramatized. For Bakhtin, writers and readers imagine and compose dialogue among the language genres that people use in different contexts for different purposes. Students working with teachers use contextualizing language to compose events and make sense of text located within meaningful encounters in a fictionalized world. With dramatic imagination, young people and teachers collaboratively compose and revise as they improvise the coauthoring of shared text for agreed or implicit purposes and audiences within events located in a society. When reading is understood as dialogue in imagination with authors and characters to interpret text then with dramatic imagination young people can make public, social, and more extended, dialogue that for readers is usually private inner speech.

July 2019, Newcastle, UK. I am in St John’s Primary School with Iona Towler-Evans, who worked closely with Dorothy Heathcote and is a free-lance dramatic inquiry consultant. We are with Ian Gorman and his class of 10- and 11-year-old children. They are in the midst of a term-long dramatic inquiry-based project he initiated focused on the topic of refugees. With afterschool programming and high parental involvement, the school is a vibrant hub in the local community within a working-class area of the city. A few Asian and Afro-Caribbean Black children are students in each classroom that is predominantly White. Refugees seeking asylum status have been settled throughout the UK though not locally. Ian is committed to social justice education. Having led the children in a project related to local homeless people, he has chosen the topic of refugees that had become a controversial issue around the 2016 Brexit referendum. He has begun to share accumulated information and stories in a critique of the dire situation for people fleeing wars in Africa and the Middle East. Using dramatic inquiry, James has created a fictionalized town with the children so as to explore cross-cultural challenges of resettling people from European camps. On the wall is a map, writing, and drawings all created and used by the children, as if they were a town council. Talking informally with children it becomes clear that they are invested in imagining places, amenities, and businesses in the town and conducting council meetings as they prepare for the arrival of refugees. At lunch Ian asks us to lead in the afternoon. With little time for planning, we gather up some materials aware that we will have to mostly improvise yet knowing we can be guided by the inquiry that has already been established by the teacher.

After introductions, Ian asks the children to tell us about their project. I field comments from eager children. Soon, sensing a loss of focus, Iona begins talking quietly with girls sitting near her. She holds a sheet, making soothing sounds as if she is holding a baby. Girls are keen to rock and pretend to care for the child. I shift everyone’s focus to look as if we are seeing a photograph and then a video of a mother and baby, with possible siblings, waiting for a place to begin their lives again. We listen in to hear the mother’s thoughts: “I am so tired. I just want to find a new home.”

Transforming classroom reality

Dramatic imagination is inherently transformative and humanizing. With social imagination, students and teachers may experience from the relative positions of multiple other people, or non-humans, and interact as if they were those others located elsewhere. Shared experiences of young people and adults are consistent with the social expectations of an imagined world but are no longer perceived as bounded by the physical world nor limited to individual experiences. Dramatic imagination is used not to escape reality, simulate another's reality, or critique from a privileged position, but rather to dialogue from multiple competing positions and thereby transform our classroom reality.

Dramatizing events creates holistic experiences and depictions of events that are interpreted for their significance. Though there are no physical consequences of embodied imagined actions, people may move and interact in community events that are social as well as individual and experienced affectively as well as intellectually. Indeed, it is only when dramatizing stereotypical schooling that students might sit silently at desks working on individual cognitive tasks fixed by a predetermined agenda and monitored by a teacher.

Dramatic inquiry uses dramatic imagination to extend shared classroom experiences by making the power dynamics of teacher-student relationships more equal so as to collectively and collaboratively dramatize events. When all in effect imagine, "What if the world were different?" then dramatic inquiry can be used to embody imagined possibilities of alternative ways of acting in, being in, and understanding the world.

Dramatic conventions

Heathcote invented dozens of drama *conventions* each of which harnesses personal and dramatic imagination to play with time and space. Using these conventions, students interact with teachers to give form to significant experiences and representations of other people and other-than-human aspects of places, which may then be interpreted (Heathcote, 1980b). These conventions can be used to create the experience of forming physical and social space as well as stopping or reversing or fast-forwarding time into alternative pasts, parallel presents, and possible futures. Heathcote used the metaphor of the hologram, a 3-D image of a whole object, to stress how a carefully chosen and depicted encounter can open into a collaborative exploration of facets of a whole topic. She stressed the significance of people interacting in what she called "now time," as if those we "stand in for" were actually present here-and-now and could speak and be spoken to as subjects. For example, a woman and child who are refugees. The affordance of Heathcote's convention of a teacher- or people-in-role means that adults may interact with students as if all are within an imagined event. This was a radical departure from more limited uses of school drama as performance or role-play exercises. Teachers and/or students could now extend dialogue to include people as if in non-realistic times and places, for example, as if in a photograph, a person's memory, or a conceptual drawing. Other innovative conventions include actual or drawn artifacts used as icons to represent the possessions of a person or objects in an event. For example, a sheet to represent a baby's blanket or a letter held by the baby's mother. For Heathcote, conventions are ways to signify and be in dialogue with anything that is "other" than us so as to humanize our experience and interpretation of events that may be concerned with other-than-human creatures, artifacts, environments, and societies as well as particular people.

Dramatic dialogic inquiry

Dramatic inquiry is dialogic inquiry into aspects of the human condition extended by dramatic imagination. The significance of dialogic inquiry in language and literacy education was proposed by Gordon Wells: "To create the conditions that promote sustained student engagement and a willingness to engage in discourse that aimed at increasing understanding." Teachers need to "create communities of inquiry in which all are encouraged to ask 'real' questions and together attempt to construct answers" (Wells, 2001, p. 1). When classroom life is organized for inquiry it is not structured for individual pursuit but rather shaped to become a shared endeavor in which everyone in a classroom community may be in dialogue with everyone else, focused on facing people's concerns and resolving problems inherent in a shared project (Fecho and Clifton, 2017; Wells, 1999). Dramatic inquiry is dialogic inquiry into life events that may be dramatized.

Addressing the children, as if we are the town council, I state, "I understand we've agreed to take in some refugees. Is that right?" There is general nodding. Having asked the children to talk with those near them about what had been agreed at the last council meeting I speak with the boy who says he's taken on the role of the mayor. I learn that he can't quite remember what had happened. Listening in to the others talking it is clear that how they might resettle the newly arriving refugees and who would be responsible remains an open question.

I switch strategies to using a continuum strategy with the class. I ask the children to place themselves in a line to show how strongly they felt about finding a place to live in the town for the woman with the baby. We look at Iona again as if she is in a photograph on her asylum application and wonder how she might be feeling. Children respond out loud including: "Sad" "Help with her baby."

The study of any content in dramatic inquiry is always contextualized by dialogue in relation to events in the lives of people. With dramatic inquiry, young people encounter and work to resolve challenging complex problems that other people are facing while being addressed as contributors to activities focused on inquiring into concerns related to those problems. Focus centers on creating shared meaning, while acknowledging contested meaning, within an ongoing inquiry into human concerns. Teachers can respond to students' questions and interests intending to engage in meaning-making dialogue that may focus subsequent inquiry.

Different approaches to dramatic inquiry

Here I review four approaches to dramatic inquiry and how the concept of framing informs the pedagogy.

Four approaches

Heathcote invented four related ensemble approaches to dramatic inquiry pedagogy: Process Drama, Rolling Role, Mantle of the Expert, and the Commission Model (Heathcote, 2002). *Process Drama* can involve one or several sessions in which a teacher artfully structures activities to focus a group on creating and exploring events in an imagined drama world such as the world of a story (O'Neill, 1995). The *Rolling Role* approach was created to be used primarily by teachers working collaboratively across different classrooms (Davis, 2016). *Mantle of the Expert* and the *Commission Model* are more complex approaches envisioned to continue over weeks or months of extended and deepening dramatic inquiry that may focus on any aspects of curriculum. Below, I elaborate on and provide examples of these two approaches using the frequently used abbreviations mantle and commission.

With a mantle approach, young people and teachers create an imagined community as a fictionalized team with a focus on helping other people with a problem (Aitken, 2021; Heathcote and Bolton, 1995; mantleoftheexpert.com; mantleoftheexpert.co.nz; Taylor, 2016). For example, as if we are a town council planning for the arrival of people who are refugees. Students gradually take on a mantle of expertise shared by an imagined team of people so as to frame their collaborative inquiries over time. In a mantle, young people work as a team to grapple with resolving problems faced by fictionalized people who need help (often referred to as “clients”). In my extended example for this essay, we collectively imagined a fictionalized scenario: a town council is being asked by the UK Refugee Council to help with resettlement. In a commission, students are commissioned by actual people to grapple with real-life problems in an extended project (Heathcote, 2003; mantlenetwork.com/commission-model). For example, in 2016–17 Iona (my co-teacher in the extended example herein) set up a commission with older students in South Wales: a representative from the local library invited young people to create a documentary film on the life of Richard Burton, the actor who had been born and grown up locally (Edmiston and Towler-Evans, 2022). With both these approaches young people may at any time dramatize events: with mantle they interpret as a fictionalized team whereas with a commission they do so as part of their ongoing collaborative work as students.

Framing

Heathcote applied the sociological concept of framing to dramatic inquiry (Heathcote, 1980b). In addition to taking on various social roles of imagined people in dramatized events, she recognized the advantage of adopting a unifying position of assumed responsibility for the welfare of others. A consistent dramatic framing became a core dimension of her mantle approach: collectively imagining that they are a team of professionals or other people facing challenging problems, young people are more open to accepting the need to work collaboratively and take on grown-up responsibilities in activities focused on resolving those problems. Significantly, the people needing help in grappling with those problems personalize the audience for any activities, contextualizing language use and literacy practices in life situations rather than in the context of schooling.

Over time, as young people work as an ensemble through the continued use of mantle or commission approaches to dramatic inquiry, they may acquire established social and literacy practices (Lave and Wenger, 1991). In mantle or commission projects, dramatic imagination is used to explore events from multiple positions in the lives of a myriad of people in past, present, or fictionalized cultural worlds that contextualize ongoing inquiry into a topic. Now the focus of dramatic inquiry extends beyond single texts. When students dramatize and seek to understand events chosen to resemble challenging episodes in the everyday lives of professionals or other people with grown-up responsibilities, within, and at times beyond the classroom, young people encounter, use, and over time may become more proficient at using disciplinary languages and literacies (Moje, 2015) in meaningful and relevant human contexts within communities of practice. Indeed, when guided by a shared humanizing vision of a desired future, what young people imagine may become more of a reality in the changing classroom community.

Humanizing engagement and purpose

Humanizing pedagogy respects and uses the reality, history, and perspectives of [all] students as an integral part of educational practice.

Bartolomé (1994, p. 173).

With dramatic inquiry, teachers can engage every young person, regardless of age or apparent ability, with novel activities employing language and text related to imagined events. Students are consistently positioned with the power to shape events and the authority to interpret.

Heathcote lamented that for students schooling can feel like a perpetual “waiting room” where the institution is denying them “the power to influence” what is happening (Heathcote, 1980a, p. 132). In contrast, with dramatic inquiry the people imagined in fictionalized events have purposes that sustain and deepen student engagement over time when young people identify with the challenges those imagined people face in that cultural world. Everyone playfully but intentionally engages when shifting back-and-forth in imagination between co-creating and co-interpreting; between foregrounding interacting as if all are other people experiencing engaging life events elsewhere and foregrounding interacting as students focused on the meaning of those actions.

Engagement with relational agency

What Anne Haas Dyson has shown about play extends to all young people in dramatic inquiry activities: they are “interwoven with their agency” both “individual and collective” (Genishi and Dyson, 2014, p. 229). Texts can be encountered and created for uses in communities within cultural worlds beyond the school. When young people use language and participate in social literacy practices in events intending to achieve something that they desire and in response to the possible consequences of actions from the perspectives of others they act with relational agency (Edwards, 2005). Dramatic inquiry can establish a norm that the views of other people on any event are worthy of being considered and answered in ongoing engaging dialogue. Young people extend dialogue about particular human concerns as they take up competing points of view on people’s actions while facing the possible consequences for themselves and others.

Most children place themselves at the “I’m sure that the council should find her a home in our town” end of the continuum. Others are in the “I’m not sure” middle. Two boys, Tom and Neill, are alone on the “I don’t think the council should find her a home in our town” end of the continuum. When I ask children to explain their position, the positive reasons range from “She needs our help” to “She has nowhere to live.” Tom says, “They probably have disease.” Neil adds, “They can go somewhere else. They don’t have to come here.” The boys allude to racist anti-immigrant discourses circulating at the time. Seeking to promote some dialogue, I ask for responses. The children basically restate their ideas. While I sense a collaborative sense of community belonging among all students, I feel that continuing to ask questions might reify monologic reactions rather than promote dialogic interactions. It is time for more dramatization.

Heathcote introduced two ideas to conceptualize teaching that promotes engagement with relational agency.

First, teachers work to “awaken the spectator” in every student as all look for the “significance” in the work (Heathcote, 1980a). Greene (1995) believed that young people need the arts so as to be “wide-awake” in school not passively receiving information, but rather, “thinking about their condition in the world, to inquire into the forces that appear to dominate, to interpret the experiences they are having from day to day” (pp. 43–44). Heathcote embraced James Britton’s (1972) vision of humanizing literacy pedagogy. People must learn to use language both as participants and as spectators; we participate to communicate while as spectators we reflect to interpret for meaning. Acting with dramatic imagination, students shift positions to become spectators on their own interactions both as if in an imagined world and as students in the classroom community.

Second, Heathcote advised teachers to actively press students to focus on the “implications” of what people are doing in any event (Heathcote, 1984a). She used drama conventions to shift positions in space while stopping, slowing, reversing, or doubling the experience of time. In doing so she created openings to dialogue focused on looking beneath the surface of imagined people’s acts for likely inferences, causes, consequences, and connections to other possible events all related to whatever concerns or questions guide ongoing inquiry into facets of life. Shifting among different people’s relative positions on events, young people may reflect on the social or cultural significance and implications for others, as well as oneself, and of the consequences of any collectively imagined actions.

Purpose and audience

“Why are we doing this?” When students have explicit and engaging answers to this implicit question then they contribute to classroom activities aware of a purpose that they care about. Dramatic inquiry activities have inherently collaborative purposes: we dramatize events intending to explore together a life-affirming and meaningful topic, such as settling people seeking asylum status. All verbal and non-verbal social interactions and literacy practices have actual and/or imagined people as audiences. As Britton (1972) put it, there are no “dummy runs.”

Purpose feeds extended engagement. Teachers engage students by introducing meaningful content, not only to initially attract attention, but also to cultivate shared interest and promote community commitment to a life-affirming project expected to take time. Young people’s sense of purpose is strengthened when they align with and collectively share the purpose of people framed collectively to address and resolve a pressing concern.

Humanizing classroom communities

What am I doing at this present moment within my society which is of any use? How will what happens in this time contribute to this individual, this group of people, my community, my nation, the world of people and objects?

Heathcote (1984b, p. 171).

A more humanizing classroom community is being created with dramatic inquiry pedagogy as interactions become social practices that establish humanizing classroom cultural norms. Using dramatic inquiry changes “how we do things in our classroom community.” It shifts what’s “normal” about what it means to be a student and how people create new knowledge with language. Everyone contributes to each activity in an unfolding classroom community thereby cocreating a shared world that is both real and imagined.

Education humanizes when the classroom community culture changes so that every young person feels they belong because they are being addressed as already sufficiently competent, capable, or knowledgeable in relation to activities and possible actions that they care about. This is in contrast to feeling that schooling is, in effect, saying to them, “one day you will be good enough to really do [something worthwhile], but not today.” As Heathcote puts it, young people’s relationships with peers and teachers may be transformed when, in relation to an imagined future that they care about, they hear: “Today you are good enough and you really are going to do it” (Heathcote, 1984b, p. 178).

Young people consistently create and experience as an ensemble while also interpreting and critiquing the meaning of any text in dialogue between different perspectives on events from any of the multiple available positions of actual and imagined people. The full range of human emotions from joy to despair or from fear to courage may be experienced and shared in relation to collectively imagined and embodied events. At the same time, abstract ideas, processes, literacies, and more of the complexity of contexts in any subject matter come to life as young people contribute to the inclusive practices of collaborative dramatization of human events.

Ian has identified illustrations from a graphic novel he is planned to show the children before reading the story. When he projects a colored picture of a helicopter lifting someone out of a boat, the children are captivated, perhaps wondering what else might happen. Ian, Iona, and I are able to provide answers to their many questions. Yes, they are refugees. Yes, the helicopter could be British. Yes, the RAF could bring them to the UK. No, not everyone might be saved. I ask the children if they’d like to set up that event. They are eager to “show it like it was.” We move to the carpeted area to imagine boarding and crowding into a small boat each of us with only a backpack filled with precious possessions. We playfully dramatize navigating the waves, being the storm coming, feeling the boat tip, and staring into the water. After practicing again, we agree to make it scarier by imagining it is night. We make the sounds of a helicopter over the sounds of wind and rain. We close our eyes so as to sense the moment. I pause to ask if they want everyone to survive. A few say yes, but most are adamant that some would drown. The tone has become solemn as I ask them to choose whether or not the person that they are imagining will make it. They stand up if they want to imagine being rescued. They stay sitting if they want to not be hoisted into this helicopter. They slide off the carpet if they want to imagine drowning. There is giggling at first but not when we repeat. Tom and Neil both stand to be rescued. I look them in the eye, smiling, yet addressing them with a serious tone: “So you want a community to give you refuge ... You could stay on the boat and take your chances?” They don’t move. Then turning to everyone standing I speak as if I am in the helicopter: “Look down. Don’t forget those we had to leave behind.” If we’d had time, we would have dramatized events in the implicit storylines. Having established what could happen we enact the rescue again but now with Iona and me pretending to be helicopter pilots narrating the event.

For Heathcote, the antidote for dehumanization is not moralizing but humanizing over time with “preventative education.” Heathcote sees racism, sexism, and all systemic inequalities, as like “weeds waiting to become part of a young person’s value system” (Bolton and Heathcote, 1999, p. 129). Heathcote stresses how humanizing values are being formed when the emerging culture of the classroom community is “respectful.” Humanizing values engender an ongoing “context of caring” and “nurture” in which “the language of honoring [all] people can be regularly learned and practiced” so that honoring others becomes normalized as a regular classroom social practice (Bolton and Heathcote, 1999, p. 131).

Humanizing students’ connections with the curriculum

With dramatic inquiry, language and literacy is always encountered and used in context within cultural worlds. Simultaneously, teachers focus on building connections and relationships between young people’s ongoing classroom experiences in the unfolding world of the classroom and the cultural worlds from which different students come as well as with the lives of people in the cultural worlds of whatever content is to be studied. Students are not addressed as generic or idealized. Pedagogy is intended to be culturally and historically relevant, responsive, and sustainable in relation to the actual young people in the room (Ladson-Billings, 1995; Paris, 2012; Muhammed, 2020).

Dramatic imagination may be conceptualized as creating a permeable membrane between the lives of people in a particular classroom community located in a school within the world of education and people’s lives in communities located in cultural worlds beyond the classroom that must be imagined to be experienced. The curriculum as content studied in school, for Heathcote, should always be a living encounter between the actual people in a classroom and the lives of other people either met in person or who must be imagined yet are accessible via dramatic imagination. This is echoed in Dyson (1993) who argued nearly thirty years ago that curriculum should be “permeable” not based on “standards.” Dyson’s research shows why: “the worlds of teachers and children come together in instructionally powerful ways” that includes “the interplay between teachers’ and children’s language and experiences.” Dyson, like Heathcote, is emphatic: creating a “shared world is essential for the growth of both oral and written language” including sense-making and awareness of audience. Forming a shared world not only in reading together but also in playing together creates a humanizing community: “it is essential as well if teachers and children are to feel connected to, not alienated from, each other” (p. 1).

The understanding of young people becomes more complex when their actual experiences extend in imagination to connect with other people’s ways of making sense of different cultural worlds. Dramatic inquiry activities must “protect students into experience” and not leave them emotionally exposed (Heathcote, 1978, p. 153). Because dramatized experiences are both real and imagined, students choose to engage in actions anticipating any physical and emotional affects. For example, agreeing to pretend to be in a sinking boat when they know they are not going to drown even though they may imagine the feeling of panic when floundering in the water as they move their bodies. Thus, in the way that biological membranes (like skin) protect and regulate what moves in and out of an organism, teachers can use dramatic imagination with groups to create shared experiences that focus collaborative inquiry in relation to facets of events that are experienced holistically yet selectively.

Dramatic inquiry creates a what is ~ what if? permeable space-time for new ways of being in the world to emerge. Not only are the possibilities of “what if?” informed by people’s existing knowledge and past experiences of “what is” but over time the inverse can also be true; humanizing experiences of imagined events can transform the assumptions, values, and possibilities of everyday classroom life.

Transforming classroom communities

Maxine [Greene \(1995\)](#) argues that community, “has to be achieved by persons offered the space in which to discover what they recognize together and appreciate in common ... infused by imaginative awareness that enables those involved to imagine alternative possibilities of their own becoming and of their group’s becoming ... reaching toward some common world” (p. 39). If society is to become a more morally just and caring democratic society, Greene argues that people must be consciously wide-awake so as to engage in critical inquiry. She rejects the notion of passive socialization. For her, the arts and humanities can “awaken the imagination” and allow us to actively and socially explore the human condition to “bring human beings in touch with themselves” (1977, p. 124).

Anna [Stetsenko \(2017\)](#), the Vygotsky scholar and sociocultural theorist, similarly rejects passive and individualistic assumptions about education. She calls for teachers to “transform” the status quo by envisioning with students and bringing into being a desired classroom community. Rather than be content with students only “adapting” to established ways of being while “assimilating” knowledge and skills, Stetsenko calls for teachers to adopt with students a “transformative activist stance.” When power is shared in dramatic inquiry, following Stetsenko, teachers actively work with students as a community, addressing them as potential “contributors” to all activities not just participants. Established power dynamics can shift so as to collectively imagine and bring into being a more humanizing classroom community. Young people’s agency can be transformative as well as relational: as teachers and students contribute to activities then they are both shaping the cultural world and being changed by interactions ([Stetsenko, 2019](#)).

However, as [Lensmire \(2011\)](#) argues, a commitment to transformative change can be overly serious. Lensmire calls for a joy that dramatic inquiry can bring. He puts it this way: “in order to criticize and remake the world [we] need to play (with ideas, with each other) in order to experience something better—a something better than throws the present shortcomings into bold relief. Seriousness can undermine truthfulness. Criticism may require laughter” (pp. 125–126). Playfulness is inherent in dramatic inquiry pedagogy. With dramatic inquiry we create events that are both joyful and serious and in which possible ways of being in the world are not only referenced but may be embodied and experienced in encounters located in a time and space previously only talked about: “Time and space are rearranged in ways which show their contingency and indissolubility. All of this is done in a mood of celebration and laughter” ([Robinson, 2011](#), n.p.).

With dramatic inquiry, teachers and young people can bring into being radically different and less hierarchical ways of being in community. Ways that open up more possibilities for how to interact and create understanding in dialogue that may critique existing assumptions by playing around with possibilities. In dramatic imagination we can create playful, ever-changing, “carnavalesque” spaces. For [Bakhtin \(1984\)](#), these are “characterized by freedom, equality and abundance [where] rank is abolished and everyone is equal” ([Robinson, 2011](#), n.p.). Unlike traditional performance spaces, the carnivalesque spaces of dramatic inquiry are “communal, with no boundary between performers and audience.” The social consequences for interactions are profound: “It creates a situation in which diverse voices are heard and interact, breaking down conventions and enabling genuine dialogue.” The traditional teacher–student status divide is destabilized, it “eliminates barriers among people created by hierarchies, replacing it with a vision of mutual cooperation and equality.” Teachers may represent children while students may depict professional adults ([Robinson, 2011](#), n.p.). As [Heathcote](#) argued, everyone’s “stature” as a person is now more important than one’s social “status” in everyday life ([Heathcote, 1971](#), p. 69). Anyone may engage the world and dialogue taking up the relative positions of people implicated by events in an imagined world. The permeable space of dramatic inquiry becomes “a zone in which new birth or emergence becomes possible.”

Some children want to imagine a different outcome for their character. I direct another playful dramatization that has both laughter and moments of stillness. Some children pretend to be hoisted into the helicopter, others are not, with some becoming those who rescue or greet with blankets, food, and water. Iona again takes up a pose as the woman with the baby as I narrate how she has been one of those rescued by the helicopter and eventually flown to a camp in England where they are now.

A classroom community formed with dramatic inquiry is a playful group intentionally focused on collaboratively and artfully exploring how the world could be different. Everyone is a potential shape-shifter able to explore aspects of other cultural worlds from multiple relative positions by imagining events from the standpoints of different people within collectively imagined encounters. Unlike the riotous herd behavior at the US Capitol on January 6, 2021, in a carnivalesque community there is dialogic space for young people and teachers to inquire into both humanizing and dehumanizing facets of socially imagined experiences in the lives of other people as all improvise dialogue. With dramatic inquiry, as people explore and experience aspects of other people’s lives, we are transforming our personal and shared experience and our understanding of how people might be in the world.

Humanizing relationships and identities

Humanizing approaches involve the building of relationships of care and dignity ... involving reciprocity and respect ... feeling valued by a worthy witness.

[Paris and Winn \(2014, p. xvi\)](#).

Dramatic inquiry pedagogy can humanize language and literacy education for students in classrooms by enriching the relationships among young people and teachers and the social practices related to texts that establish cultural norms in a classroom community. Young people can relate to topics of concern, what is “other” than them, by creating and dramatizing representations of other people and non-human dimensions of anything that becomes the subject of their shared inquiry. Young people are (as much as possible) actively contributing to all activities and engaging in social practices that humanize their relationships with one another, with adults in the room, and with people in other cultural worlds. As Heathcote puts it, young people in a classroom community ideally are forming both a collective sense of “we” as well as a growing trusting connection with each teacher and peer, all of whom they identify as “colleagues.” As such, in dramatic inquiry, most actions are “we-for-us” (Edmiston, 2014). Over time, young people may humanize their relational identities as students in a classroom community and as people in the world.

Dramatic inquiry is premised on equality and thus the potential for dialogue with all people within and beyond the physical classroom. Like the philosopher, Jacques Rancière (1991), Heathcote always began with a “presupposition of equality.” Rather than working toward equality, Heathcote began with a presumption that everyone’s way of being, along with their ideas and interpretations, are already equally worthy of being recognized, taken seriously, and responded to. How, as teachers, we relate to some people must not be more privileged than how we relate to others.

I ask the children to come back to where we’d been in the room 20 min earlier and to place themselves again on the continuum. Everyone has moved closer toward the “I’m sure that we should find her a home in our town” end, including Tom and Neil who have moved slightly.

Transformative crucible

Heathcote introduced the metaphor of a crucible to characterize the personal and social transformative potential of dramatic inquiry (Heathcote, 1983). Like a crucible that may be used to heat, and with stirring transform, existing materials into new substances, dramatic inquiry extends predictable classroom interactions by introducing alternative ways of being and knowing. Dramatic inquiry interactions may feel challenging to young people, even though they can be fictionalized and transformed into unexpected experiences, novel ways of relating, and unanticipated realizations. However, having experienced being in the crucible changes people. As the anthropologist, Dorothy Holland puts it in relation to play worlds, “we may return to the everyday with an altered subjectivity, an altered sense of who we are.” Over time, dramatizing events in other worlds can “help us envision new identities.” Over time, we are playfully coauthoring a “refigured” community in which “new worlds and new identities are created” (Holland et al., 1998, p. 238, 239, 250).

When young people dramatize novel encounters that contextualize competing ways of making sense of events this is felt as dramatic tension. Dramatic tension not only engages people to experience events but is also like the heat in a crucible that allows competing positions on events to be experienced as intermingling and in dialogue with each other rather than remaining distinct. In the crucible people can have heated interactions without being angry. Dramatic inquiry fosters collective exploration not personal disagreement. Though heat may be generated via dramatic tension, in social interactions as if in imagined events students’ experiences have built-in emotional distance since any uncomfortable disagreements, like any comforting agreements, are about topics in an imagined world and not focused directly on everyday life. Heathcote stresses the significance of the indirect approach of dramatic inquiry:

It may be impossible for young children to recognize their own value system dictating “this is how life is,” but they *can* learn to look for a deep level of explanation of behavior in the characters within a fiction.

Heathcote (1984a, p. 142).

Heathcote used framing to increase (or decrease) emotional distance from events yet focus on responsibility and accountability for the consequences of actions (Eriksson, 2011). Students and teacher can focus collectively on how people might respond to questions or concerns as part of an ongoing inquiry. A teacher’s interactions and use of dramatic conventions is like stirring different viewpoints together so as to promote and press for transformative dialogue with novel perspectives, between students’ views, and within each young person.

I stand by the boys. I ask Tom what he’s thinking: “She might have caught a disease in the boat.” Despite some groans from other children, I support his position: “It’s true. People can catch diseases and I don’t want to catch one. Do you?” I make a face of disgust as I address the question to his peers. There are several shakes of the head. “We could take her to the hospital to be tested” suggests one of the girls, “My dad works there.” She moves toward Iona and others join her. “What do you think, Tom?” After several embodied exchanges showing how this might happen, I press him: “So, if we were able to test her and found she didn’t have any diseases, what do you think we should do?” He nods: “We should find her a home.”

Transformative Dialogue

Dialogue is more than talk. It is thinking together in response to authentic questions and concerns (Mercer, 2004; Nystrand, 1997). Dialogue can involve experiencing at the same time different ways of being in and making sense of the world (Bakhtin, 1981). With dramatic imagination, young people and teachers collectively embody events in the narrative world of any text as if those events are

actually happening in a physical, social, and cultural world. Further, they both experience and interpret life events from the multiple and competing perspectives of any of the people depicted in the narrative worlds of any printed, narrated, embodied, composed, or digital texts.

Dialogue is transformative when experienced as a back-and-forth receiving-and-giving of recognition and ideas. Dialogue in the potentially transformative crucible of dramatic inquiry can be understood as occurring in a mutually created “Third” space created “between” two (or more) competing perspectives on events (Benjamin, 2018). In this Third space, the psychologist Jessica Benjamin (2018) conceptualizes that “the self is changed by having to struggle with the other’s difference” (p. 3). A changing self is in contrast to a static, unchanging, or narcissistic self that is either unaffected or more entrenched by an encounter with another. For each person, the self can be “projecting into the not-me” that may include “its own vulnerability” (p. 14) creating space for an “internal confrontation” (p. 19). Ideally there is “mutual responsiveness and transformation” (p. 3). Transformation can be both personal and in relation to others. As Benjamin puts it, in the Third “our attachment to all beings as part of the whole is honored as real One’s humanity depends not only on the respect one receives but the quality of the recognition one gives, that indeed one’s own dignity is fostered by giving recognition” which is “reciprocal recognition” (p. 234).

With dramatic inquiry, young people, and teachers, can imagine and then dialogue from the previously unconsidered position of any person, including positions judged as oppressive or otherwise dehumanizing. In dialogue, young people may begin to question some existing assumptions though they may not. The experience of intermingling dialogue contextualized by interactions between actual and/or imagined people with different relational positions, may create feelings of discomfort that especially over time can lead to transformative, humanizing change in students’ ways of being, doing, and knowing (Edmiston and Beach, 2017).

I turn to Neil, who looks a bit distracted: “How do you feel about that idea?” It is almost the end of the day and the children will have to leave soon. I am unsure if Neil has been listening. I look at my watch saying we’ll have to end soon. “So, what about this home for her. Where could it be?” I muse. “We could build it” says one of the girls. Though there is no time to consider the complexity of this idea I am enthusiastic: “We could! I wonder where?” “We could build it here!” the girl replies. The idea might seem fanciful. However, these children had actually witnessed a derelict field adjacent to the school be transformed into a sports field for the school and the community along with a building for indoor activities. Children had been included in school-wide dialogue related to the initial planning ideas. I move as if lifting a heavy load: “I could carry bricks. Can anyone help?” A couple of the children join me. “Does anyone have any tools?” I make the sound of a drill. Though some gravitate toward packing up, most children join in pretending to build. I turn to Neil, “Do you have any tools we might borrow?” “I have a hammer.” I am appreciative: “Could you help us hammer in these nails?” As the children start to line up to leave, Neil and I are laughing as we pass an imagined hammer back-and-forth and pretend to strike nails. Ian quietens the children to thank us. I nod: “Unfortunately, we can’t come tomorrow. I wish we could see the home the people in the town built. I wonder what name was written over the door?” There is a sense of completion when one of the girls who’d sat close to Iona suggests, “How about ‘The Helping Home?’” No one disagrees.

Humanizing identities

Our social and cultural identities are how we identify with groups and communities of other people and thus act in relation to them to humanize (or dehumanize) relationships. Students’ identities, like teachers’ identities, are dynamic, not fixed (Holland et al., 1998). They shape and are shaped by the shared values and cultural norms of a community. Everyone’s subjective experience of being in the classroom community is affected by, and affects, how people come to assume they should act relative to other people in the world within and beyond the school. Additionally, dramatic, dialogic interactions over time can shape how students identify ethically (Edmiston, 2000, 2008). The significant social and personal transformation of relationships that is possible via the humanizing experiences, activities, social practices, and outcomes of dramatic inquiry can transform identities (Medina and Campano, 2006). Thus, young people may begin to change what it means to them to identify with others as “students” in the classroom as well as how they identify as “people” in the world beyond the school all of which shapes how they may act in society.

When young people and educators are focused on intentionally realizing in the present, aspects of a collectively imagined common good tied to a vision of a hopeful future, then they are acting to transform relationships both within the classroom community and with people in communities beyond the school. Over time, dramatic inquiry pedagogy can make a difference to the changing relational identities of young people as humanizing students, readers, writers, language users, and people.

Over time, dialogic interactions among students and teachers can create possibilities for each young person to become a more grown-up human being in the world in terms of taking on new responsibility and acting in relation to others with more awareness of other viewpoints and of the consequences of actions (Biesta, 2014). Rather than reify facets of established abstract ideas, young people may clarify and extend their understanding of life including realizing more about their own value system as they contextualize existing knowledge and assumptions in back-and-forth humanizing responses to different perspectives on challenging concrete problems faced by people in imagined events.

Conclusion

Freedom dreaming is imagining worlds that are just, representing people’s full humanity, centering people left on the edges, thriving in solidarity with folk [folks] from different identities who have struggled together for justice.

Love (2020, p. 103).

Marcus Crum (2020), in his concluding column in the special issue of the *Journal of Literacy Research*, “Black Lives Matter in Literacy Research,” argues that teaching, along with research, should be evaluated to determine whether they are “supporting human well-being or undermining human well-being in our racialized societies” (p. 544). Further, he stresses that such evaluation is only possible when, “we rethink and redefine what it means to be human and what it means to be ourselves.” However, doing so requires that we recognize that, “Who we imagine others as depends a great deal on who we imagine ourselves to be” (p. 544).

In dramatic inquiry we draw on everyone’s knowledge, experience, and skills developed in the past and we are aware of the present status quo. Yet, with dramatic inquiry the future for students is not predetermined adaptation. With dramatic imagination we can embody and experience a more humanizing future as if we were already there. Within long-term projects young people are able to develop literacy knowledge and critical practices as they use language to enact grown-up ethical responsibilities both toward other people being served and to one another within classroom communities. With shared experiences of alternative ways of being in the world we can be in dialogue about how to get there more often. It is no longer a case of teaching with a “one day you’ll be good enough” stance in relation to content that is always “over there.” Using dramatic inquiry we begin with attention to our here-and-now relationships with young people addressed as already good enough yet all on a shared imaginative journey to becoming more humanizing people in the world (Heathcote, 1978, p. 178).

I advocate for dramatic inquiry pedagogy because, like Love (2020), I am “imagining life not death, imagining seeing the world, and imagining freedom ... to preserve humanity” (p. 102). Dramatic inquiry creates a sense of an ongoing dance between what is and what if. The status quo—what is generally accepted as the way schooling is—can be in a continual dialogue with alternative ways of being educated in a classroom community—what if our classroom was a more humanizing community? Dramatic approaches need not be tangential to education but via dramatic inquiry pedagogy may become established ways of envisioning and bringing into daily classroom life how we all want to be in the world so that the world might become a better place.

References

- Aitken, V., 2021. “Real in All the Ways that Matter”: Weaving Learning Across the Curriculum With Mantle of the Expert. New Zealand Council for Educational Research, Auckland.
- Bakhtin, M.M., 1981. In: Holquist, M. (Ed.), *The Dialogic Imagination: Four Essays*. University of Texas Press, Austin.
- Bakhtin, M.M., 1984. In: Emerson, C. (Ed.), *Problems of Dostoevsky’s Poetics*. University of Minnesota Press, Minneapolis.
- Bartolomé, L., 1994. Beyond the methods fetish: toward a humanizing pedagogy. *Harv. Educ. Rev.* 64 (2), 173–195.
- Beach, R., Beauchemin, 2019. *Teaching Language as Action in the ELA Classroom*. Routledge, New York.
- Beach, R., Bloome, D., 2019. *Language Relations for Transforming the Literacy and Language Arts Classrooms*. Routledge, New York.
- Benjamin, J., 2018. *Beyond Doer and Done to: Recognition Theory, Intersubjectivity and the Third*. Routledge, New York.
- Biesta, G.J.J., 2014. *The Beautiful Risk of Education*. Press, Paradigm.
- Bolton, G., Heathcote, D., 1999. *So You Want to Use Role Play?* Stoke-on-Trent: Trentham.
- Britton, J., 1972. *Language and Learning*. Penguin, London.
- Crum, M., 2020. “If Black lives matter in literacy research,” then take this racial turn: developing racial literacies. Special issue, *Black Lives Matter in Literacy Research*. *J. Liter. Res.* 52 (4), 530–552.
- Davis, S., Ferholt, B., Clemson, H.G., Saut-Mari, J., Marjanovic-Shane, A., 2015. *Dramatic Interactions in Education: Vygotskian and Sociocultural Approaches to Drama, Education and Research*. Bloomsbury, New York.
- Davis, S., 2016. *Learning that Matters: Revitalizing Heathcote’s Rolling Role for the Digital Age*. Sense, Boston.
- Dyson, A.H., 1993. *Negotiating a Permeable Curriculum: On Literacy, Diversity, and the Interplay of Children’s and Teachers’ Worlds*. Concept Paper No. 9. ERIC Number: ED359548.
- Edmiston, B., Beach, R., 2017. Dialogic eventful teaching through dialogic conversations and dramatic inquiry. In: Durst, R.K., Newell, G.E., Marshall, J.D. (Eds.), *Language, Literacy, and Learning: Arthur Applebee and the Future of Inquiry in English Education*. Routledge, New York.
- Edmiston, B., Enciso, P., 2002. Reflections and refractions of meaning: dialogic approaches to classroom drama and reading. In: Flood, J., Lapp, D., Squire, J., Jensen, J. (Eds.), *The Handbook of Research on Teaching the English Language Arts*. Simon and Schuster Macmillan, New York, pp. 868–880.
- Edmiston, B., Towler-Evans, I., 2022. *Humanizing Education With Dramatic Inquiry: In Dialogue With Dorothy Heathcote’s Transformative Pedagogy*. Routledge.
- Edmiston, B., 2000. Drama as ethical education. *Res. Drama Educ.* 5 (1), 63–84.
- Edmiston, B., 2008. *Forming Ethical Identities in Early Childhood Play*. Routledge, London.
- Edmiston, B., 2011. Chapter five: dramatic inquiry. In: Beach, R., Campano, G., Edmiston, B., Borgman, M. (Eds.), *Literacy Tools in the Classroom: Teaching Through Critical Inquiry, Grades 5-12*. Teachers College Press, New York.
- Edmiston, B., 2014. *Transforming Teaching and Learning With Active and Dramatic Approaches: Engaging Students Across the Curriculum*. Routledge, New York.
- Edwards, A., 2005. Relational agency: learning to be a resourceful practitioner. *Int. J. Educ. Res.* 43, 168–182. <https://doi.org/10.1016/j.ijer.2006.06.010>.
- Enciso, P., Edmiston, B., Volz, A., Lee, B., Shivashankar, N., 2016. “I’m tryin to save some lives here”: critical dramatic inquiry in the aftermath of Hurricane Katrina. In: Jocson, K., Simon, R. (Eds.), *Special Issue: Youth Critical Literacy and Social Justice, English Teaching: Practice and Critique*, vol. 15, pp. 333–354, 3.
- Enciso, P., 1996. Why engagement in reading matters to Molly. *Read. Writ. Q.* 12 (2), 174–194.
- Enciso, P., 2017. Stories lost and found: mobilizing imagination in literacy research and practice. *Liter. Res.* 66, 29–52.
- Eriksson, S., 2011. Distancing at close range: making strange devices in Dorothy Heathcote’s process drama teaching political awareness through drama. *Res. Drama Educ.* 16 (1), 101–123. <https://doi.org/10.1080/13569783.2011.541613>.
- Fecho, B., Clifton, J., 2017. *Dialoguing Across Cultures, Identities, and Learning: Crosscurrents and Complexities in Literacy Classrooms*. Routledge, New York.
- Genishi, C., Dyson, A.H., 2014. Chapter 19: play as the precursor for literacy development. In: Blaise, M., Edwards, S. (Eds.), *SAGE Handbook of Play and Learning in Early Childhood*. SAGE Publications, United Kingdom, pp. 228–251.
- Greene, M., 1977. *Landscapes of Learning*. Teachers College Press, New York.
- Greene, M., 1995. *Releasing the Imagination: Essay on Education, the Arts, and Social Change*. Jossey-Bass, San Francisco.
- Heathcote, D., Bolton, G., 1995. *Drama for Learning*. Heinemann, Portsmouth, NH.
- Heathcote, D., 1971. Subject or system? In: Johnson, L., O’Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 61–79, 1984.
- Heathcote, D., 1978. Drama and the mentally handicapped. In: Johnson, L., O’Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 148–153, 1984.

- Heathcote, D., 1980a. Material for significance. In: Johnson, L., O'Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 126–137, 1984.
- Heathcote, D., 1980b. Signs and portents. In: Johnson, L., O'Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 160–169, 1984.
- Heathcote, D., 1983. Learning, knowing, and languaging in Drama. *Lang. Arts* 60 (6), 695–701.
- Heathcote, D., 1984a. Drama as context for talking and writing. In: Johnson, L., O'Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 138–147, 1984.
- Heathcote, D., 1984b. The authentic teacher and the future. In: Johnson, L., O'Neill, C. (Eds.), *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London, pp. 170–199, 1984.
- Heathcote, D., 2002. Contexts for active learning: four models to forge links between schooling and society. *J. Drama Educ.* 19 (1), 1–14.
- Heathcote, D., 2003. A Vision Possible: The Commission Model of Teaching. *Drama Magazine*, Winter, pp. 16–24.
- Heathcote, D., 2009. Mantle of the Expert: My Current Understanding. Unpublished Keynote address, Weaving our Stories International Mantle of the Expert Conference Hamilton, New Zealand.
- Heathcote, D., 2010. Dramatic imagination. *J. Drama Educ.* 26 (2), 14–30.
- Holland, D., Skinner, D., Lachicotte, W., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge, MA.
- Johnson, L., O'Neill, C. (Eds.), 1984. *Dorothy Heathcote: Collected Writings on Education and Drama*. Hutchinson, London.
- Kuby, C., Spector, K., Johnson Thiel, J., 2019. *Posthumanism and Literacy Education: Knowing/Becoming/Doing Literacies (Expanding Literacies in Education)*. Routledge, New York.
- Lacey, S., Wooland, B., 2009. *New Theat.* Q. 8 (29), 81–91.
- Ladson-Billings, G., 1995. *Dreamkeepers: Successful Teachers of African American Children*. Jossey-Bass, San Francisco.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Leander, K., Ehret, C., 2019. *Affect in Literacy Learning and Teaching: Pedagogies, Politics and Coming to Know (Expanding Literacies in Education)*. Routledge, New York.
- Lensmire, T., 2011. Too serious: learning, schools, and Bakhtin's carnival. In: *Bakhtinian Pedagogy*. Peter Lang, New York, pp. 117–128.
- Love, B., 2020. We Want to Do More than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom. Beacon Press.
- Medina, C., Campano, G., 2006. Performing identities through drama and teatro practices in multilingual classrooms. *Lang. Arts* 83 (4), 332–341.
- Mercer, N., 2004. Sociocultural discourse analysis: analyzing classroom talk as a social mode of thinking. *J. Appl. Ling. Prof. Pract.* 1 (2), 137–168.
- Moffett, J., 1987. *Teaching the Universe of Discourse*. Boynton/Cook.
- Moje, E.B., 2015. Doing and teaching disciplinary literacy with adolescent learners: a social and cultural enterprise. *Harv. Educ. Rev.* <https://doi.org/10.17763/0017-8055.85.2.254>.
- Muhammad, G., 2020. *Cultivating Genius: An Equity Framework for Culturally and Historically Responsive Literacy*. Scholastic.
- Nystrand, M., 1997. *Opening Dialogue*. Teachers College Press, New York.
- O'Neill, C., 1995. *Drama Worlds: A Framework for Process Drama*. Heinemann, Portsmouth, NH.
- O'Neill, C. (Ed.), 2014. *Dorothy Heathcote on Education and Drama: Essential Writings*. Hutchinson, London.
- Paris, D., Winn, M., 2014. *Humanizing research: Decolonizing qualitative inquiry with youth and communities*. Sage.
- Paris, J., 2012. Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41 (3), 93–97.
- Philip, T.M., Souto-Manning, M., Anderson, L., Horn, I., Carter Andrews, D., Stillman, J., Varghese, M., 2019. Making justice peripheral by constructing practice as "core": how the increasing prominence of core practices challenges teacher education. *J. Teach. Educ.* 70 (3), 251–264.
- Rancière, J., 1991. *The Ignorant Schoolmaster: Five Lessons in Intellectual Emancipation*. Stanford University Press, Stanford, CA.
- Robinson, A., 2011. In theory: carnival against capital, carnival against power. *Ceasefire Mag.* <https://ceasefiremagazine.co.uk/in-theory-bakhtin-2/>.
- Schneider, J., Crumpler, T., Rogers, T., 2006. *Process Drama and Multiple Literacies*. Heinemann, Portsmouth, NH.
- Stetsenko, A., 2017. *The Transformative Mind*. Cambridge University Press, Cambridge, UK.
- Stetsenko, A., 2019. Radical-transformative agency: continuities and contrasts with relational agency and implications for education. *Front. Educ.* 17. <https://doi.org/10.3389/educ.2019.00148>.
- Taylor, T., 2016. *A Beginner's Guide to Mantle of the Expert: A Transformative Approach to Education*. Singular, Norwich, UK.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Wagner, B.J., 1996. *Drama and the Language Arts: What Research Shows*. Heinemann, Portsmouth, NH.
- Wagner, B.J., 1999. *Dorothy Heathcote: Drama as a Learning Medium*, second ed. Calendar Island, Portland, ME.
- Wegerif, R., 2015. *Dialogic: Education for the Internet Age*. Routledge, New York.
- Wells, G., 1999. *Dialogic Inquiry: Towards a Socio-cultural Practice and Theory of Education*. Cambridge University Press, Cambridge, UK.
- Wells, G., 2001. *Action, Talk, and Text: Learning and Teaching Through Inquiry*. Teachers College Press, New York.
- Wolf, S., Edmiston, B., Enciso, P., 1997. Drama worlds: places of the heart, head, voice and hand in dramatic interpretation. In: Flood, J., Heath, S.B., Lapp, D. (Eds.), *The Handbook for Literacy Educators: Research on Teaching the Communicative and Visual Arts*. Simon and Schuster Macmillan, New York, pp. 492–505.

Arts-based practices for literacy & language learning

Douglas Fisher and Nancy Frey, San Diego State University, Health Sciences High, San Diego, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	358
What are the arts?	358
Oral language development	359
Phonological awareness and phonics through the arts	359
Phonological awareness	360
Phonics	360
Reading fluency and comprehension	361
Vocabulary learning	361
Writing	362
Conclusion	363
References	363

Introduction

Learning about art, whether that be appreciation, creation, or critique, is worth time in schools. Eisner (1965), a champion of arts-based education, wrote, "I would like to propose that the art education curriculum be built along three major lines of focus. These three foci are *the productive aspects of art, the critical aspects, and the historical aspects*" (p. 11). As such, students should have access to art specialists who can lead lessons designed to develop student's artistic abilities. But that's not the point of this chapter; instead, we focus on the role that the arts can play in literacy and language learning.

What are the arts?

Wikipedia defines art as follows: Art is a diverse range of human activities in creating visual, auditory or performing artifacts, expressing the author's imaginative, conceptual ideas, or technical skill, intended to be appreciated for their beauty or emotional power (<https://en.wikipedia.org/wiki/Art>). In school systems, we typically think of art along the following lines:

- Visual art, which include the traditional fine arts such as drawing, painting, ceramics, photography, sculpture, works in wood, and mixed media.
- Drama or theater, which include performance and interpretation of performances.
- Music, which includes singing and the use of instruments.
- Dance, which involve a range of movement experiences and the interpretation of movement.
- Media arts, which includes photography, imaging, sound, animation, video, web design, graphic design, virtual design, and interactive design.

As we have noted before, these are each worthy of instructional attention in their own right. For example, students in Singapore experienced a series of lessons (including dance, music, drama and visual arts) focused around the concept of space (Bautista et al., 2016). As the authors noted, "the arts are often regarded as areas for enrichment or even entertainment within schools" but they demonstrated the learning of students within the arts themselves (p. 624). Students in a school in Turkey used the arts to explore their personal histories (Coşkun, 2017). As one of the students noted, "I used to read Derrida, since high school, and I was seriously obsessed with his theory about deconstruction. What Derrida said was that in order to understand the whole we needed to deconstruct it. When you analyze something into its pieces you actually understand it better, from pieces to whole" (p. 360). Our point is sharing these two studies of arts education, from the thousands that are published worldwide, is to convey the idea that art is a discipline in and of itself. We do not want to imply that art is only useful in service to other areas. Having said that, there is fairly extensive literature base about the ways in which the arts impact literacy and language learning.

Consider the very young child who sees an image, possibly a drawing or sculpture. The young child wants to talk about it, ask questions about it, and understand it. All the while, this child is using and developing language. Jessie's mom took him to a museum just after his third birthday. Each Tuesday, different museums in our community are free. On this particular day, Jessie and his mom visited the Timken Museum of Art (www.timkenmuseum.org). As they visited, Jessie noticed Jean-Honoré Fragonard's painting *Blindman's Buff* (*Le Colin-Maillard*), c. 1775–80. As he looked at it, he started saying colors: *green, white, blue*. He then pointed to the dress a woman was wearing and said *red*. His mom responded, "Honey, we call that pink. It's like red but softer and has more white in it. Do you see it?" He started pointing to other parts of the painting saying *pink, pink, pink*. He pointed to another

part of the painting and said *baby*. He then asks, “what are they doing?” pointing to the man with a blindfold. His mom responds, “I think that they are playing a game. He doesn’t look scared and the day is very nice, so I think it’s something fun that they are going to do.” Their conversation continues, but we’ll pause here to consider the literacy aspects of this interaction.

To be sure, Jessie was learning to appreciate and analyze art, but he was also developing vocabulary, applying knowledge about words to images, asking questions, engaging in a conversation, and building background knowledge that he could use later.

Oral language development

One of the things that stands out in Jessie’s interaction with his mom at the museum is his development of oral language. And oral language is a critical aspect of literacy development. As many have noted, human brains are hardwired to speak and listen but not read and write. Instead, every brain must learn to read and write (e.g., Wolf, 2008). The importance of oral language in learning to read and write cannot be overstated. Consider the following two quotes:

- “Oral comprehension sets the ceiling on reading comprehension” (Biemiller, 2003, n.p.).
- “Reading and writing float on a sea of talk” (Britton, 1983, p. 11).

Both convey the power of oral language and the impact that strong oracy has on literacy development. Yet, oral language has been, and continues to be, a neglected aspect of literacy instruction and has been for some time (Holbrook, 1983; Rossiter et al., 2010).

There are any number of studies that provide tools for educators to foster students’ oral language skills. We will highlight several that involve the arts as these are easy and fun ways to engage students in speaking and listening. For example, Brouillette (2012) highlights the use of creative drama to develop oral vocabulary skills for students in kindergarten through second grade who were learning English as an additional language in US schools. As one of the teachers noted, “When we acted it out, more of the kids were engaged, even the kids who couldn’t say the words yet. At the beginning of the year, they copy movements. Eventually, they understand how to say that, using language” (p. 144). Similarly, Stinson (2015) noted the value of drama in improving students’ oral language development with elementary-aged students in Australia. In this study, parents noted increased self-confidence in their children as well as increased engagement in school because the children “have something to say” (p. 303).

Music can also be used to foster students’ speaking and listening skills (Mobbs and Cuyul, 2018). In fact, these educators offer recommendations for song selection to develop students’ oral language skills, including the clarity of the lyrics, vocabulary choices, grammar structures, and cultural sensitivity.

In addition to helping to develop students’ oral language skills, music promotes multiculturalism among students in the classroom (Paquette and Rieg, 2008). As DiDomenico (2017) notes, music can be integrated into the classroom such that listening skills are developed. Cornett (2007) discusses several research-based findings that generally show that music-based instruction helps to increase students’ self-concept, language, cognitive development, listening skills, critical thinking, and social skills. She claims: “Every musical experience that we offer our students affects their brains, bodies and feelings. In short, it changes their minds permanently” (p. 463).

In addition, students can talk about fine arts to build their speaking and listening skills. In our documentation of regular museum visits, students’ oral language skills improved as their teachers invited conversations and discussions about the various exhibits that the students attended (Pumpian et al., 2006). In this investigation, over 1000 students from a historically low-performing school spent one week each month of the school year in a museum, learning content through the arts available in that museum. Their literacy skills improved significantly as they interacted with each other, the arts, and their teachers. In fact, the school was recognized for the increase in the reading performance of students and was identified as a distinguished school in California.

As we will see in the remainder of this entry, the arts can develop students’ language and literacy skills. Let’s start with early aspects of reading, which are phonological awareness and phonics.

Phonological awareness and phonics through the arts

Children learn to read through acquisition and consolidation of the sounds, symbols, and meaning of language. The sounds and their corresponding written symbols require coordination of key areas of the left hemisphere of the brain responsible for discriminating speech sounds, phonological awareness, and decoding words (Dehaene et al., 2005). Because the brain is not hardwired to read, instruction during the primary grades focuses on the coordination of these systems.

The sounds and written symbols of a language are constrained skills, which are “small sets of knowledge that are mastered in relatively brief periods of development” (Paris, 2005, p. 184). The smallest unit of sound (phonemes) vary by language. English has 44 phonemes, while Spanish has 25. The written symbols also vary. The English alphabet’s 26 letters and letter patterns (graphemes) represent the phonemes. The ability to accurately map the sounds of a language onto the graphs of the language is phonics knowledge, completed in typically developing children by the age of 8 or 9.

Phonological awareness

Phonological awareness is the ability to decipher and manipulate speech sounds, syllables, and rhymes. Adams' (1990) proposed five levels of phonological awareness:

Level 1: Hearing rhymes and alliteration

Level 2: Oddity tasks, such as figuring out which word in a string does not rhyme

Level 3: Blending words and splitting syllables

Level 4: Orally segmenting words

Level 5: Manipulation tasks that require the ability to delete, substitute, add, and reverse phonemes.

Music is a natural fit for teaching phonological awareness (PA) and is widely utilized for this purpose. Music enhances the acquisition of early language skills, especially through singing, chants, clapping, and the use of rhythmic instruments (Fisher, 2001; Gromko, 2005). A study of music education on the phonological awareness of kindergarteners found that children with higher exposure outperformed same-age peers on measures of manipulation and rhyming. In addition, their rhythm ability correlated to their PA skills (Moritz et al., 2013). Similarly, Gromko (2005) found that 4 months of literacy-based music education resulted in statistically significant gains among kindergarten students.

Singing and movement appear to have positive effects on the phonological awareness of young children. Auditory discrimination of pitch is associated with elements of PA, particularly the ability to segment and manipulate phonemes (Lamb and Gregory, 1993). There is emerging evidence of the link between the cerebellum, which controls movement, and the pre-frontal cortex, which governs cognition, in the development of academic skills (Piek et al., 2008). Callcott et al. (2014) evaluated the effectiveness of a movement and music literacy program of action-based songs that incorporated fine and gross motor practice, cross-lateral coordination, and rhythm on Australian pre-primary students. The children who participated in the movement and music literacy program outperformed peers on measures of phonological awareness and developmental spelling. Rule et al. (2006) found that students in grades 1 through 3 who participated in a movement and language program improved their PA at higher levels than those who attended a traditional pull-out literacy intervention program. The combination of singing songs, body percussion or instruments to keep the beat, and dance movements require an organization of aural, temporal, and spatial reasoning similar to those needed in reading (Gromko and Poorman, 1998).

Music and movement are commonly incorporated into classroom phonological awareness activities (Adams et al., 1997). Songs and chants that are brief and repeated so that they become familiar routines form the core of these experiences. The developmental need of young children to play as a means to learn dovetails well with the word games, rhymes, and movement featured in many PA programs. Multisensory games that allow children to hear slight shifts in language, often to humorous effect, are utilized to coordinate with finger play. Nursery rhymes such as "Hickory Dickory Dock" and "Old McDonald Had a Farm" are readily adapted to include gestures, clapping, and movement. Other examples include incorporating children's names in chants, foot stomping to count out syllables, and simple chants, and songs that use repetitive lyrics that change slightly (e.g., "There Was an Old Woman Who Swallowed a Fly").

Phonics

In practice, phonological awareness instruction is coupled with phonics instruction to build decoding and word recognition skills. Thus, lessons are often a combination of instruction about letters and words that accompany the songs, chants, and movement. A study of Canadian 4- and 5-year-olds found a correlational relation between musical skills, phonological processes, and early reading ability (Anvari et al., 2002).

The alphabetic principle is integral to early decoding, as children need to accurately and rapidly recognize letters and their associated sounds (Chall, 1999). Many language arts curricula include multisensory songs designed to teach students the alphabetic principle. Although not extensively researched, DiLorenzo et al. (2011) did show that a multisensory alphabet program was effective for typically progressing students, special education students, and others deemed to be at risk academically.

An innovative dance-based reading program was utilized with more than 700 first-grade students to improve their alphabetic and phonics skills (McMahon et al., 2003). The researchers used an experimental methodological design to assess the impact of the 20 lessons on students' code knowledge (consonant and vowel recognition). Over a 3-month period, students used their bodies to make the shapes of letters and letter combinations of sounds. Members in the experimental group who had originally scored lower than control group students outperformed them at the end of 3 months, a result that the researchers noted "the program took low-performing readers and turned them into significantly better readers" (p. 119).

The use of music for the purpose of developing early reading skills in word knowledge has been documented as effective as intervention for struggling readers (Register et al., 2007). Song lyrics have been used to improve decoding and word recognition skills with middle school students with learning disabilities (Hines, 2010). The researcher noted that popular songs can be motivational for adolescents, and that the salience of music makes recollection of lyrics easier (Morrongiello and Roes, 1990). Students in the study selected popular songs with the teacher and rehearsed them over a period of days using printed lyrics. Difficult words were highlighted and practiced until the student could read all of the lyrics accurately. Words that the student had difficulty with were then used to create rimes and word ladders to build the student's decoding skills and promote transfer of learning to new words. For example, *fake* was expanded to include the rimes *make*, *bake*, and *shake*. Word ladders, which require a one-letter change with each word, helped students read through the entire word. Thus, *sell* became *spell*, *spill*, *spin*, and *spit*. Students who participated in the program made gains in their decoding skills.

The arts can be an effective vehicle for teaching early reading skills with young children. Many early literacy scholars argue that these approaches should be combined with systematic phonological awareness and phonics curriculum designed to ensure that acquisition of these constrained skills is not left to chance (Juel and Minden-Cupp, 2000; Snow et al., 1998). These early decoding skills are essential for reading, but do not represent all facets needed. Fluency and comprehension are also necessary for meaning making.

Reading fluency and comprehension

Reading fluency is generally comprised of two factors: rate and prosody. Rate is the speed at which a person reads. Too slow, and meaning is lost. Too fast, and meaning is lost. It appears that there is a just right, or Goldilocks, rate for reading. It is generally understood that readers who “sound good” when reading aloud are also good comprehenders (e.g., Kuhn and Stahl, 2003; Hudson et al., 2009). In their study of 98 kindergarteners and 170 first graders in Korea, Kim et al. (2014) noted that there “developmental relationships among listening comprehension, word reading fluency, text reading fluency, and reading comprehension” (p. 94) and that these relationships changed as students reading levels increased. Interestingly, in this Korean study, “listening comprehension largely explained variation in reading comprehension” (p. 94) for early readers but becomes less explanatory as text reading fluency begins to play a role in understanding.

But rate is not the only factor that impacts fluency. Prosody, or the expressiveness with which a person reads, also contributes to understanding of the text. Prosody involves the appropriate intonation, rhythm, and emphasis given to words and sentences when they are being read. Typically, prosody instruction focuses on reading aloud, but the goal is that students internalize this type of reading and develop an inner speech while reading. It seems that readers, when reading silently, use prosody in ways that are similar to what they would do while reading aloud, thus the Implicit Prosody Hypothesis (Fodor, 2002). Importantly, the ways that readers use these prosodic representations affect their interpretation of the text (Breen, 2014). As Breen et al. (2016) have noted, there is a relationship between prosody and comprehension through high school. In their study of 112 elementary school students in Spain, Calet et al. (2017) noted that prosody training was the most effective method to improve both fluency skills (speed, accuracy, and prosody) and reading comprehension.

It should be clear that fluency is an important aspect of literacy. But where do the arts enter the picture? In their review of research on students with disabilities, Stevens et al. (2017) noted the relationship between reading fluency and comprehension, emphasizing that repeated reading remains the most effective way to improve reading fluency. In their study of bilingual students in Vietnam, Gorsuch and Taguchi (2008) noted the value of using repeated readings to increase reading rates, which in turn impacted reading comprehension. The question is how to get students to read again and again to improve their fluency, both in terms of rate and prosody.

That’s where the arts enter. Students are likely reluctant to read texts again unless they see a clear purpose for it. Preparing for a performance, whether that be a dramatic interpretation, signing, or readers’ theater, seems to invite students to read texts over and over, thus improving their fluency. As Young et al. (2016) noted, performances allow students to practice with a purpose. In their study of second graders, integrating performances into instruction resulted in statistically significant gains in fluency. Similarly, Iwasaki et al. (2013) described a musical method in which students learned a song and performed it each week. Each day, students interacted with the song lyrics and practiced. The result of the integrate of music was a significant increase in reading scores. With older students, Biggs et al. (2008) examined the use of an interactive singing software with middle schoolers in the United States. Their findings suggest that fluency improved as a result of reading and re-reading texts in preparation for singing. Further, Paquette and Rieg (2008) described the value of using music to develop literacy skills in young children, especially their fluency as they sign and develop a sense of pace.

Similarly, readers theater has been used encourage repeated reading and improve fluency (Keehn, 2003). Readers’ theater is a dramatic reading without costumes, props, scenery, or special lighting. Students use their vocal expressions to help the audience understand the story. Typically, students read and re-read the script in preparation for the performance. In some cases, students create the script from a longer text. As Martinez et al. (1998) noted, using drama and theater in this way improved students’ fluency both in terms of rate and prosody. Evidence from the use of readers’ theater with high school students in Turkey suggests that students enjoyed the experience and that increased their participation, self-confidence, self-control, and curiosity of the subject while also encouraging repeated reading and improved reading fluency.

In the case of fluency, it appears that many of the effective approaches in literacy are based on experiences with the arts. As students read to perform or engage in repeated readings and singing of songs, their rate and prosody improve as does their overall reading comprehension. And, as evidenced in several of the studies reviewed, students enjoyed the experience having the arts integrated into their instruction.

Vocabulary learning

Word knowledge is a significant predictor of comprehension (Yildirim et al., 2011). Students need to learn a lot of words every year in school if they are to understand the texts that they are expected to read. Thankfully, we all learn many words incidentally, as we read and listen to others (Swanborn and de Glopper, 1999). Other words need to be taught (Biemiller, 2011), particularly in the

context of reading and writing. One of the challenges with vocabulary learning is that it is unconstrained (Paris, 2005). Unlike phonological awareness and phonics, and to a certain extent fluency, which have a more finite level of learning, vocabulary continues to develop across our lifetimes. That growth occurs in both depth and breadth. As we get older, our vocabularies expand as we deepen our understanding of words, often learning more obscure meanings, historical derivations, and newer usages.

There are a number of evidence-based approaches to teaching vocabulary (Hiebert, 2019). In this entry we will focus on the role that the arts can play in developing students' word knowledge (LaBrocca and Morrow, 2016). In their study in Iran, Gorjian et al. (2012) developed authentic art experiences that included visual arts, literature, drama, music, and dance and studied the impact that these experiences had on learners' vocabulary. They noted significant increases in word knowledge and attribute those changes to the experiences students had with art. As they noted, "Working with something that is concrete such as art to something abstract such as language encourages students to link their own experiences" (p. 349).

Charalambous and Yerosimou (2015) describe the use of drama, music, and media to improve students' vocabulary knowledge in a Greek school located in London. As they noted, "In situating the focus on creative and artistic process through which to develop the language at focus, we observed that students' interest in language learning was enhanced and their motivation to come to school and learn Greek was developed" (p. 380). Their study demonstrates the ways in which the arts can be used to introduce and reinforce vocabulary knowledge in a multilingual setting.

In the United States, Keehn et al. (2008) studied the use of readers theater with middle school students. The students engaged with short stories using scripts developed by the researchers. They compared students who participated in weekly readers theater with those who did not and noted that the vocabulary of those who did doubled compared with those who did not.

Also in the US, Fisher and McDonald (2001) describe the integration of music to build students' phonics and vocabulary skills. As they note, music invites student to meet words, and unique uses of words, that they might not otherwise encounter and, in doing so, students vocabulary increases.

Finally, Montgomerie and Ferguson (1999) describe the ways in which drama can be used to build students' knowledge. They were specifically interested in developing students' critical literacy skills. Their students were between the ages of four and eight and lived in the United Kingdom. As the authors noted, the use of process drama provided a rich context for language development including the construction of argument and adoption of different ways of speaking depending on circumstances. In doing so, students vocabularies were extended.

Writing

The language arts are categorized as receptive and expressive domains. Listening and reading are receptive in nature, in the sense that they require the ability to comprehend received messages. Speaking and writing are expressive language skills, as a message must be produced in forms that can be understood by others. Writing is the apex of language production and as such requires consolidation and coordination of multiple systems to encode, including composition, grammar, spelling, and idea generation (MacArthur et al., 2016).

Writing is itself an expression of art, especially in the field of drama. The texts of plays and films begins its life in written form as an original or adapted script. In this context, art is story and can serve to both provoke writing as well as to be an outcome of the process.

Participatory photography is a research method approach used with learners as a way to illuminate their lives and experiences (Singhal et al., 2007). Our own work with struggling ninth grade writers led to us to using participatory photography through a photo essay project accompanied by original writing (Frey, 2003). Using their cameras, they shot scenes from their lives outside of school. Once printed, each student curated their collection of photographs to compose a story using visual and written text. Wise-man et al. (2017) reported on a similar approach with third grade students in a Literacy Through Photography project. The students engaged in three projects: self-portraits, African proverbs, and community photographs. A cross-case analysis of interviews with the students surfaced themes of the application of visual strategies for communication, social collaboration with peers, and connections to community experiences.

As the availability of visual tools has expanded, digital storytelling (DST) has received increased attention. Noguerón-Liu and Hogan (2017) used multimedia visual tools in a cross-generational project of immigrant youth and adults from the Michoacán state in Mexico to examine transnational similarities and differences existed among them. Using a similar participatory methodology, they drew themes about composition processes in light of the migratory histories of families, and the its application to literacy learning in adolescents' US classrooms to strengthen writing. The use of DST appears to boost writing skills of students. Because digital storytelling involves many aspects of writing at the preproduction level, including planning, determining the flow of the story, and listing the visuals to be used, conventional composition improves. As an example, Sarica and Usluel's (2016) study of 59 s grade students in Turkey revealed that writing skills improved after DST, as did their visual memory skills.

The use of visuals in graphic novels can be used to prompt and scaffold writing with young students as well as adolescents. Paneling in sequential storytelling is essential for the flow of ideas. Pantatelo's (2013) study of the use of paneling in the composition of original graphic novels composed by Canadian fourth grade students revealed that their decisions about color, size, and shape of the panels in the texts they created were deliberate and translated to other written forms. In some cases, original graphic novels are created by students as a means to improve writing skills and foster motivation in writing. That was the case in a study of the use of graphic novels by Caribbean high school students using both their home language and English (Manning-Lewis, 2019).

Our studies of the use of graphic novels with high school students have revealed similar findings. One study was done with students in World History and English classes (Frey and Fisher, 2010). In English, they were introduced to observational skills using visuals to hone their eye for detail. Several times a week, they wrote short responses to a visual image displayed by the teacher. The World History teacher soon adopted a similar method, using photographs and paintings of historical events. Not only did the amount of writing increase over time, but content analyses conducted by the teachers revealed that the quality of their writing improved.

The use of visual arts as both a means for expression and as a tool to prompt writing continue to develop as access to tools becomes more widespread. Student and teacher access to videos, photographs, and museums has made it possible for writing classrooms to bring a wider range of masterpieces in art to students. In an increasingly visual world that relies on images to tell real and imagined narratives, their use in schools will likely only expand.

Conclusion

As we have noted, art is important in and of itself. As Cummings noted, “Viewing or creating art can be a form of mindfulness, experiencing life in the moment while noticing your feelings without judging” (2020, p. E2). As such, students should experience art for arts’ sake. We are not arts specialists but rather educators who understand that the arts contribute to human thought in a number of ways. In this entry, we focused on the ways in which the arts can contribute to literacy and language development. Of course, teaching literacy is not limited to the arts. There are a vast number of instructional approaches useful in developing students’ prowess in reading, writing, speaking, listening, and viewing. It’s just that the arts can add a bit of flair to literacy learning. It may just be that the cliché is true: A picture is worth a thousand words. So, let’s tap into every possible tool we have to build students’ literacy lives.

References

- Adams, M.J., Foorman, B., Lundberg, L., Beeler, T., 1997. Phonemic Awareness in Young Children: A Classroom Curriculum. Paul H. Brookes.
- Adams, M.J., 1990. *Beginning to Read: Thinking and Learning About Print*. MIT Press.
- Anvari, S.H., Trainor, L.J., Woodside, J., 2002. Relations among musical skills, phonological processing, and early reading ability in preschool children. *J. Exp. Child Psychol.* 83 (2), 111–130.
- Bautista, A., Tan, L.S., Ponnusamy, L.D., Yau, X., 2016. Curriculum integration in arts education. *J. Curric. Stud.* 48 (5), 610–629.
- Biemiller, A., 2003. Oral Comprehension Sets the Ceiling on Reading Comprehension, Electronic ed. *American Educator*. Retrieved July 5, 2020 from: <https://www.aft.org/periodical/american-educator/spring-2003/oral-comprehension-sets-ceiling-reading>.
- Biemiller, A., 2011. Vocabulary: needed if more children are to read well. *Read. Psychol.* 24 (3–4), 323–335.
- Biggs, M., Homan, S., Dedrick, R., Rasinski, T., 2008. Using an interactive singing software program: a comparative study of middle school struggling readers. *Read. Psychol.* 29 (3), 195–213.
- Breen, M., Kaswer, L., Van Dyke, J.A., Krivokapic, J., Landi, N., 2016. Imitated prosodic fluency predicts reading comprehension ability in good and poor high school readers. *Front. Psychol.* 7, 1026. <https://doi.org/10.3389/fpsyg.2016.01026>.
- Breen, M., 2014. Empirical investigations of the role of implicit prosody in sentence processing. *Lang. Linguist. Compass* 8, 37–50.
- Britton, J., 1983. Writing and the story of the world. In: Kroll, B., Wells, E. (Eds.), *Explorations in the Development of Writing Theory, Research, and Practice*. Wiley, New York, pp. 3–30.
- Brouillette, L., 2012. Advancing the speaking and listening skills of K–2 English language learners through creative drama. *TESOL J.* 3 (1), 138–145.
- Calet, N., Gutiérrez-Palma, N., Defior, S., 2017. Effects of fluency training on reading competency in primary school children: the role of prosody. *Learn. Instr.* 52, 59–68.
- Calcott, D., Hammond, L., Meryl Hill, S., 2014. The synergistic effect of teaching a combined explicit movement and phonological awareness program to preschool aged students. *Early Child. Educ. J.* 43 (3). <https://doi.org/10.1007/s10643-014-0652-7>.
- Chall, J.S., 1999. Some thoughts on reading research: revisiting the first-grade studies. *Read. Res. Q.* 34, 8–10.
- Charalambous, C., Yerosimou, M., 2015. Drama, music and media in heritage language learning. *J. Educ. Cult. Soc.* 6 (2), 370–381.
- Cornett, C.E., 2007. *Creating Meaning Through Literature and the Arts: An Integration Resource for Classroom Teachers*. Pearson Merrill Prentice Hall, Upper Saddle River, N.J.
- Coşkun, N., 2017. Self-history project in visual arts education. *Int. J. Educ. Through Art* 13 (3), 349–367.
- Cummings, D.E., 2020. Now More Than Ever, the World Needs Art. *San Diego Union-Tribune*, p. E2.
- Dehaene, S., Cohen, L., Sigman, M., Vinckier, F., 2005. The neural code for written words: a proposal. *Trends Cogn. Sci.* 9 (7), 335–341. <https://doi.org/10.1016/j.tics.2005.05.004>.
- DiDomenico, J., 2017. Effective integration of music in the elementary school classroom. *Inq. Educ.* 9 (2), 1–17. Retrieved from: <https://digitalcommons.nl.edu/ie/vol9/iss2/4>.
- DiLorenzo, K., Rody, C., Bucholz, J., Brady, M., 2011. Teaching letter-sound connections with picture mnemonics: Itchy’s Alphabet and early decoding. *Prev. Sch. Fail.* 55 (1), 28–34. <https://doi.org/10.1080/10459880903286763>.
- Eisner, E., 1965. Curriculum ideas in a time of crisis. *Art Educ.* 18 (7), 7–12.
- Fisher, D., McDonald, N., 2001. The intersection between music and early literacy instruction: listening to literacy. *Read. Improv.* 38, 106–115.
- Fisher, D., 2001. Early language learning with and without music. *Read. Horiz.* 42 (1), 39–49.
- Fodor, J.D., 2002. Psycholinguistics cannot escape prosody. In: *International Conference on Speech Prosody 2002* (Aix-En-Provence: ICSPA).
- Frey, N., Fisher, D., 2010. Graphic novels: composing with sequential art in high school English and history. *N. Engl. Read. Assoc. J.* 45 (2), 9–15.
- Frey, N., 2003. A picture prompts a thousand words: creating photo essays with struggling writers. *Calif. Engl.* 8 (5), 16–21.
- Gorjian, B., Hayati, A., Barazandeh, E., 2012. An evaluation of the effects of art on vocabulary learning through multi-sensory modalities. *Proc. Technol.* 1, 345–350.
- Gorsuch, G., Taguchi, E., 2008. Repeated reading for developing reading fluency and reading comprehension: the case of EFL learners in Vietnam. *System* 36 (2), 253–278.
- Gromko, J., Poorman, A., 1998. The effect of music training on preschoolers’ spatial-temporal task performance. *J. Res. Music Educ.* 46, 173–181.
- Gromko, J.E., 2005. The effect of music instruction on phonemic awareness in beginning readers. *J. Res. Music Educ.* 53 (3), 199–209. <https://doi.org/10.1177/002242940505300302>.
- Hiebert, E.H., 2019. *Teaching Words and How They Work*. Teachers College Press, New York.

- Hines, S., 2010. Name that word: using song lyrics to improve the decoding skills of adolescents with learning disabilities. *Teach. Except. Child.* 43 (1), 16–21.
- Holbrook, H.T., 1983. Oral language: a neglected language art? *Lang. Arts* 60, 255–258.
- Hudson, R.F., Pullen, P.C., Lane, H.B., Torgesen, J.K., 2009. The complex nature of reading fluency: a multidimensional view. *Read. Writ. Q.* 25, 4–32.
- Iwasaki, B., Rasinski, T., Yildirim, K., Zimmerman, B.S., 2013. Let's bring back the magic of song for teaching reading. *Read. Teach.* 67, 137–141.
- Juel, C., Minden-Cupp, C., 2000. Learning to read words: linguistic units and instructional strategies. *Read. Res. Q.* 35 (4), 458–492. <https://doi.org/10.1598/RRQ.35.4.2>.
- Keehn, S., Harmon, J., Shoho, A., 2008. A study of readers theater in eighth grade: issues of fluency, comprehension, and vocabulary. *Read. Writ. Q.* 24 (4), 335–362.
- Keehn, S., 2003. The effect of instruction and practice through readers theatre on young readers' oral reading fluency. *Read. Res. Instr.* 42 (4), 40–61.
- Kim, Y., Park, C.H., Wagner, R.K., 2014. Is oral/text reading fluency a "bridge" to reading comprehension? *Read. Writ.* 27 (1), 79–99.
- Kuhn, M.R., Stahl, S.A., 2003. Fluency: a review of developmental and remedial practices. *J. Educ. Psychol.* 95, 3–21.
- LaBrocca, R., Morrow, L.M., 2016. Embedding vocabulary instruction into the art experience. *Read. Teach.* 70 (2), 149–158.
- Lamb, S., Gregory, A., 1993. The relationship between music and reading in beginning readers. *Educ. Psychol.* 13 (1), 19–27.
- MacArthur, C., Graham, S., Fitzgerald, J., 2016. *Handbook of Writing Research*, second ed. Guilford.
- Manning-Lewis, T., 2019. I hate writing: making a case for the creation of graphic novels in the Caribbean English classroom to develop students' creative writing skills. *Appl. Financ. Econ.* 26 (4), 392–404.
- Martinez, M., Roser, N.L., Strecker, S., 1998. "I never thought I could be a star": a readers' theatre ticket to fluency. *Read. Teach.* 52 (4), 326–334.
- McMahon, S.D., Rose, D.S., Parks, M., 2003. Basic reading through dance program: the impact on first-grade students' basic reading skills. *Eval. Rev.* 27 (1), 104–125.
- Mobbs, A., Cuyul, M., 2018. Listening to music: using songs in listening and speaking classes. *Engl. Teach. Forum* 56 (1), 22–29.
- Montgomerie, D., Ferguson, J., 1999. Bears don't need phonics: an examination of the role of drama in laying the foundations for critical thinking in the reading process. *Res. Drama Educ.* 4 (1), 11–20.
- Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., Wolf, M., 2013. Links between early rhythm skills, musical training, and phonological awareness. *Read. Writ.* 26, 739–769.
- Morrongiello, B.A., Roes, C.L., 1990. Children's memory for new songs: integration or independent storage of words and tunes. *J. Exp. Child Psychol.* 50, 25–38.
- Noguerón-Liu, S., Hogan, J.J., 2017. Remembering Michoacán: digital representations of the homeland by immigrant adults and adolescents. *Res. Teach. Engl.* 51 (3), 267–289.
- Pantaleo, S., 2013. Paneling "matters" in elementary students' graphic narratives. *Lit. Res. Instr.* 52 (2), 150–171.
- Paquette, K.R., Rieg, S.A., 2008. Using music to support the literacy development of young English language learners. *Early Child. Educ. J.* 36 (3), 227–232.
- Paris, S.G., 2005. Reinterpreting the development of reading skills. *Read. Res. Q.* 40 (2), 184–202.
- Piek, J.P., Dawson, L., Smith, L., Gasson, N., 2008. The role of early fine and gross motor development on later motor and cognitive ability. *Hum. Mov. Sci.* 27, 668–681.
- Pumpian, I., Fisher, D., Wachowiak, S. (Eds.), 2006. *Challenging the Classroom Standard Through Museum Education: School in the Park*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Register, D., Darrow, A.A., Standley, J., Swedberg, O., 2007. The use of music to enhance reading skills of second grade students and students with reading disabilities. *J. Music Ther.* 44, 23–37.
- Rossiter, M.J., Derwing, T.M., Manimtim, L.G., Thomson, R.I., 2010. Oral fluency: the neglected component in the communicative language classroom. *Can. Mod. Lang. Rev.* 66 (4), 583–606. <https://doi.org/10.3138/CMLR.66.4.583>.
- Rule, A.C., Dockstader, C.J., Stewart, R.A., 2006. Hands-on and kinesthetic activities for teaching phonological awareness. *Early Child. Educ. J.* 34 (3), 195–201.
- Sarica, H.Ç., Usluel, Y.K., 2016. The effect of digital storytelling on visual memory and writing skills. *Comput. Educ.* 94, 298–309.
- Singhal, A., Harter, L.M., Chitnis, K., Sharma, D., 2007. Participatory photography as theory, method and praxis: analyzing an entertainment-education project in India. *Crit. Arts* 21 (1), 212–227.
- Snow, C., Burns, M., Griffin, P., 1998. *Preventing Reading Difficulties in Young Children*. National Academies Press.
- Stevens, E.A., Walker, M.A., Vaughn, S., 2017. The effects of reading fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: a synthesis of the research from 2001 to 2014. *J. Learn. Disabil.* 50 (5), 576–590.
- Stinson, M., 2015. Speaking up about oracy: the contribution of drama pedagogy to enhanced oral communication. *Engl. Teach. Pract. Critiq.* 14 (3), 303–313.
- Swanborn, M.S.L., de Gloppe, K., 1999. Incidental word learning while reading: a meta-analysis. *Rev. Educ. Res.* 69 (3), 261–285.
- Wiseman, A.M., Pendleton, M.J., Christianson, C., Nesheim, N., 2017. A case study of struggle and success: profiling a third grader's reading and writing in a multimodal curriculum. *J. Lang. Lit.* 13, 55–69.
- Wolf, M., 2008. *Proust and the Squid. The Story and Science of the Reading Brain*. Harper.
- Yildirim, K., Yildiz, M., Ates, S., 2011. Is vocabulary a strong variable predicting reading comprehension and does the prediction degree of vocabulary vary according to text types. *Educ. Sci. Theor. Pract.* 11 (3), 1541–1547.
- Young, C., Valadez, C., Gandara, C., 2016. Using performance methods to enhance students' reading fluency. *J. Educ. Res.* 109 (6), 624–630.

Which words to teach: review and reflection

Elfrieda H. Hiebert, TextProject, Santa Cruz, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

What is the size of the lexicon that students will encounter in school texts?	366
How does research on word features clarify which words to teach?	366
Features of words that influence students' word knowledge	366
Groups of words that might benefit from instruction	367
What has research shown about the types of words	367
That benefit from vocabulary interventions?	367
What research might support efficacious vocabulary instruction?	368
The need for vocabulary curricula	369
Core vocabulary	369
Vocabulary unique to literary and informational texts	369
The development of databases and assessments	370
Databases	370
Assessments	370
Conclusion	371
References	371

The [Oxford English Dictionary \(2005\)](#), the reference source for the English language, has entries for 616,500 word forms. This number does not capture all possible word meanings for a single entry, as illustrated by the six meanings for the entry *message* and the seven for *acquire*. Evidence is also unequivocal that the breadth and depth of individuals' vocabulary influence their textual comprehension (e.g., [Wagner and Ridgewell, 2009](#)). The lexicon is large, the size of students' vocabulary is consequential for learning, and time in classrooms is short. As educators, we are confronted with the question: Which of the many words in English should be taught?

An immediate answer to this question is that direct instruction is only one mechanism for acquiring vocabulary. Students learn the meanings of thousands of words as they interact in their communities. The word consciousness approach posits that if students are immersed in classrooms where vocabulary is rich, then they develop a love of words that results in a deep and engaged vocabulary ([Scott et al., 2008](#)). Meanings of words are also acquired through reading because many words appear more frequently in texts than in conversations ([Nagy et al., 1987](#)). Strategies such as the use of the textual context to establish unknown word meanings are another means of acquiring vocabulary ([Nagy et al., 2014](#)).

Even when all these forms of vocabulary instruction—word consciousness, strategies, and extensive reading—occur in English/Language Arts (ELAs) classrooms, there remain words that require direct instruction, such as those with complex meanings (e.g., *annihilation*) and with multiple meanings (e.g., *current*). [Nagy and Hiebert \(2011\)](#) proposed a framework for considering which words to teach in ELA instruction. Since that review was completed, numerous vocabulary interventions have been conducted. Further, advances in digital technology have meant access to large corpora of school texts and databases that allow key word features in these corpora to be established ([Brysbaert et al., 2014](#); [Kuperman et al., 2012](#)). This review updates research on word selection for direct instruction. Specifically, four topics are addressed: (a) the size of the lexicon of school texts, (b) features of words that influence the acquisition of word meaning, (c) findings of interventions that provide insight into vocabulary acquisition, and (d) questions that remain unanswered.

The research findings and conclusions of this chapter, while focused on word selection in classrooms where English is the primary language, have applicability to vocabulary instruction in other languages. The English lexicon may be somewhat larger than those of other languages, because of its historical roots in both Anglo-Saxon and Latin. However, most languages follow a pattern that has been described as Zipf's law ([1935](#)), which postulates that the frequency of any word in a corpus of natural language utterances is inversely proportional to its rank in a frequency table. That is, a relatively small group of words accounts for the majority of the words used in the conversations and writing of most languages, while a sizable portion of the lexicon accounts for only a small portion of the total words. Among the 100 languages that [Mehri and Jamaati \(2017\)](#) examined, all did not comply stringently to the frequencies proposed by Zipf but all followed this pattern of a small group of words making up the majority of words and a large group appearing infrequently. [Milton \(2009\)](#) concluded that, while Zipf's law may not be a perfect description of language, it is useful for describing the number of words required to reach specific thresholds in mastering a lexicon. Determining which words merit use of instructional time from among the many words in the lexicon of any language needs to be strategic, if students are to have the capacity to comprehend and learn from the complex and varied texts of the 21st century.

What is the size of the lexicon that students will encounter in school texts?

All 616,600 of the unique word forms in the OED are not expected to occur in school texts by any means. Numerous words represent the technical terms of highly specialized disciplines, whereas a sizable group consists of archaic words. However, if even a relatively small portion of the entire corpus appears in school texts, then students are likely to encounter many words.

In 1921, Thorndike identified 10,000 words as the appropriate focus for school instruction and, subsequently, 30,000 words (Thorndike and Lorge, 1944). Since Thorndike's analyses, which depended on manual vocabulary computation, computers have aided researchers. Numbers of unique words that students are predicted to encounter in school texts vary as a function of the size and contents of the corpora and include figures such as 154,941 (Zeno et al., 1995) and 86,471 (Carroll et al., 1971). The number of words that are predicted to appear at least once or more per million, however, is fairly consistent across databases: 19,928 (Zeno et al., 1995), 17,281 (Carroll et al., 1971), and 19,440 (Thorndike and Lorge, 1944).

Even addressing 20,000 words in the school curriculum is a massive undertaking. Nagy and Anderson (1984) suggested that when root words share transparent meanings with inflected and derived word forms, each instance of a word family may not need to be treated as a distinctive entity for vocabulary instruction. On the basis of a representative sample of Carroll et al.'s (1971) database, Nagy and Anderson predicted that school texts from grades 3 through 9 contained approximately 88,500 distinct word families.

Hiebert et al. (2018) used Nagy and Anderson's (1984) morphological lens to establish the number of word families that account for the majority of words in school texts. Hiebert et al. found that a group of approximately 6000 words that appear 10 times or more per million words of text (Zeno et al., 1995) could be reduced to 2500 word families with a shared root word. These word families accounted for 91.7% of the total words in texts described as exemplars of appropriate texts from grades 1 through 12 (National Governors Association & Council for Chief State School Officers, 2010).

In a related analysis, Hiebert (2017) identified an additional 3000 word families among words with frequencies from one to nine appearances per million words in the Zeno et al. (1995) database. This second level of word families accounted for an additional 4.5%, on average, of the total words in the exemplar texts. That is, approximately 5500 word families accounted for 96.7% of the total words in exemplar school texts from grades 1 through 12.

The findings on the core vocabulary (Hiebert, 2017; Hiebert et al., 2018) were replicated in an analysis of the words in four ELA programs from grades 1 through 5 (Graves et al., 2019). In this analysis, 4000 simple word families (i.e., words grouped by shared inflected endings within the 6000 most-frequent words) accounted for 90% of the words in texts. This analysis provided insight into the remaining 10%. This relatively small percentage of total words consisted of more than 20,000 unique words. Furthermore, 34% of the words in this group appeared a single time in the 2 million words in the database.

In summary, a core group of words accounts for the majority of words in the texts. In addition, students will encounter thousands of rare words in school texts. These rare words are often those that convey unique knowledge in informational texts or give narrative texts their texture and quality. A perspective on word selection must address both rare words and frequent words.

How does research on word features clarify which words to teach?

The typical approach to vocabulary instruction has been to select six to eight words from a text for a week of instruction (Stahl, 1999). For example, in a current ELA third-grade program (Beers et al., 2020), six words from *The Storyteller's Candle* (Gonzalez, 2012) were identified for a week of instruction: *chimed*, *concluded*, *flickered*, *gallant*, *preparations*, and *slender*. It is essential that students are able to understand the words in a text to comprehend its content (Brett et al., 1996), but the basis for choosing these words for instruction is unclear. Will words such as *chimed* and *flickered* be more challenging for most third graders than *rehearsed* and *applause*, which are also in the text and relate more directly to the story's theme of a celebration organized in an immigrant community?

Recent research on the influence of word features on vocabulary learning can aid in identifying the words that will most benefit from instruction. Two topics are discussed in this section: (a) features that influence knowledge of word meaning and (b) groups of words that might benefit from instruction.

Features of words that influence students' word knowledge

The frequency with which words appear in texts has long been recognized as an influence on vocabulary knowledge (Ryder and Slater, 1988). Students' prior experiences with words or familiarity has also been recognized as having an influence on vocabulary (Balota et al., 2001). These two criteria have been prominent in widely used instructional frameworks (Beck et al., 2013; Biemiller, 2010) and in vocabulary interventions (Scott et al., 2006).

Frequency and familiarity, although critical, are not the only features on which words vary. Nagy and Hiebert (2011) identified additional features of words as affecting vocabulary acquisition, including morphological relatedness, semantic relatedness, concept difficulty, and appearances across content areas (dispersion). Other word features have been shown as influential in the understanding of word meanings, such as degree of abstractness, concreteness (Schwanenflugel et al., 1992), and polysemy (the number of meanings associated with a word; Miller and Lee, 1993).

Multiple variables, rather than single ones, are likely to influence students' ability to learn and retain word meanings. In one such analysis of multiple word features, Cervetti et al. (2015) examined the relationship of seven word features—length, part of speech, polysemy, frequency, morphological frequency, domain specificity or dispersion, and concreteness—to second to fourth graders' knowledge of science vocabulary before and after an integrated science and literacy program. Three features were predictive of word knowledge (pretest score) at two or more grade levels: frequency, polysemy, and length. Frequency and polysemy explained students' vocabulary growth scores (posttest controlling for pretest) at two of three grade levels.

The availability of recent databases for word features such as age of acquisition (age at which words typically appear in children's oral environments; Kuperman et al., 2012) and concreteness-abstractness (Brysbaert et al., 2014) has meant an increase in the comprehensiveness of analyses of word factors on vocabulary knowledge. Hiebert et al. (2019) used these two databases, as well as frequency of word and of word family members, number of letters and of syllables, domain specificity, polysemy, and part of speech, to identify predictors of students' vocabulary performances in two large-scale projects that used distinctive assessments. In the first study, 28,901 students from grades 2 to 12 took grade-appropriate forms of a multiple-choice assessment of words predicted to occur with moderate to high frequency. Word frequency and age of acquisition predicted students' performances.

In the second study (Hiebert et al., 2019), 6976 fourth and fifth graders, 52% of whom were English learners, responded to six multiple-choice questions that each represented an aspect of word knowledge (metacognition, semantics, morphology, syntax, definitions, and polysemy). As in the first study, age of acquisition predicted student performance, but number of letters and syllables and word location (which word frequency confounded as words on the assessment became progressively less frequent) were also significant. In addition, English learners' vocabulary knowledge was significantly influenced by the concreteness of words, as evidenced by superior performances on concrete words.

Recent research has extended insight into the variables that influence the rate and retention of vocabulary learning. However, the lack of availability of databases for critical variables hampers acquiring a comprehensive perspective on word features. For example, conceptual difficulty has been shown to be a strong influence on students' learning of word meanings in context (Nagy et al., 1987; Shefelbine, 1990), but to date, no large-scale database for establishing conceptual complexity exists.

Groups of words that might benefit from instruction

The features of words require attention but so does the role of words in texts and curriculum (Nagy and Hiebert, 2011). In ELA, individual texts have been the most frequent basis for word selection, which is viewed as necessary for the comprehension of those texts. At the same time, questions can be raised as to whether words from texts can be selected to support generalizations about vocabulary use. To date, no curriculum has been designed specifically around the text of ELA instruction.

Attention to a vocabulary curriculum in ELA is essential, especially with the increase of informational texts in ELA programs. The words of informational texts appear to be treated in much the same way as those of narrative texts (Norris et al., 2008). The need for and potential avenues of exploration for identifying a vocabulary curriculum for the distinctive words in ELA programs' typical informational and narrative texts will be explored in the final section of this review.

One group of words that has begun to be treated as a distinctive vocabulary curriculum in ELA are general academic words (Nagy and Townsend, 2012). General academic words are not specific to a discipline; rather, they are words such as *category*, *consists*, and *specific* that are used to describe and clarify content-specific words in texts. Many of these words are distinguished by their abstractness and multisyllabic structure (e.g., Gardner and Davies, 2014).

Numerous research projects addressing general academic words have occurred over the past 15 years. The Word Generation (WG) program represents the most extensive and sustained research project of general academic language (Lawrence et al., 2012). A recent report from this project was based on the performances of 7752 fourth through seventh graders over a 2-year program (Jones et al., 2019). After 1 year of the treatment, significant overall impacts were found for academic language or reading comprehension for grades 4–5 or for grades 6–7. In year 2, the WG elementary curriculum had a statistically significant impact on all outcome variables of interest for the elementary cohort. The effect sizes for distal measures of academic language and comprehension ($ES = 0.06$ and $ES = 0.15$, respectively) were significant for the elementary group in the second year, whereas only the effect for distal reading comprehension ($ES = 0.16$) was significant for the middle-school students.

Various explanations have been given for the relatively modest effects of large-scale projects of general academic vocabulary, including the need for longer interventions. Facility with the abstract, polysyllabic nature of most general academic words likely requires exposure over an extended period. Further, these words are typically part of content texts that can be dense with complex subject-specific words. The role of general academic words as part of understanding subject-specific words requires attention, as does the level of exposure in texts.

What has research shown about the types of words

That benefit from vocabulary interventions?

Meta-analyses of vocabulary interventions have consistently reported that direct instruction promotes students' learning of instructed words and their comprehension of texts with those words (e.g., Marulis and Neuman, 2010; Stahl and Fairbanks, 1986; Wright and Cervetti, 2017). The finding that students learn what they are not taught is unsurprising (Fisher et al., 1981),

but when an average of 18 min of instructional time is devoted to each word in an intervention (Cervetti et al., 2019), a compelling question relates to the generalization of this learning to students' overall vocabulary and comprehension.

Consistent transfer of results from interventions to knowledge of words or comprehension beyond the taught vocabulary appears to occur at preschool and kindergarten levels, where vocabulary is assessed through receptive and productive oral language (Marulis and Neuman, 2010). For grades 1 and beyond, where assessments typically require students to read independently, results have not been as consistent or positive as those for preschool and kindergarten interventions. In an early meta-analysis of vocabulary instruction, Stahl and Fairbanks (1986) reported an effect size of 0.30 on a global measure of comprehension. In a subsequent meta-analysis with more conservative calculations of effect size and criteria for study inclusion than in Stahl and Fairbanks's (1986) analysis, Elleman et al. (2009) found no effects of interventions on transfer or distal comprehension outcomes.

Similarly, Wright and Cervetti (2017) found only 2 of 36 vocabulary interventions to have significant effects on distal measures of comprehension. Both of these studies were distinctive in that they taught words in semantic clusters (Beck et al., 1982; Nelson and Stage, 2007). Elleman et al. (2009) reported a significant effect for distal measures of vocabulary, but their basis for this conclusion was not a comprehensive sample of vocabulary interventions with distal vocabulary outcomes. In a subsequent comprehensive investigation of studies with vocabulary outcomes, Cervetti et al. (2019) reported a lack of significant differences of vocabulary interventions on distal vocabulary outcomes.

When effects of vocabulary instruction are insignificant on generalized vocabulary and comprehension assessments, researchers often cite the limitations of standardized assessments and the size of the English lexicon. These factors cannot be dismissed, but at the same time, little attention has been paid to the features of words in interventions and the relationship of word features to student performance. That is, might results be tied to the words that are taught?

To address this question, Hiebert and Nagy (2021) examined words taught in 15 of the 19 interventions in Cervetti et al.'s (2019) meta-analyses of vocabulary interventions. Each of the 1640 words was examined on 11 features: word length, frequency, frequency of morphological family, age of acquisition, concreteness-abstractness, polysemy, part of speech, dispersion, first grade of appearance in text, conceptual complexity, and semantic category. The studies were clustered into four grade bands: K, 1–2, 3, and 4–5.

Words at all grade bands were similar on all measures except for frequency, age of acquisition, and first grade of appearance. Words at K were more frequent than those at other grade bands, but even at this point, more than half the words were predicted to occur less than once per million words of text. Age of acquisition and first grade of appearance increased across the grade bands, but differences between students' actual age and age of acquisition and grade of first appearance were large at the K and grade 1–2 bands.

The research to date indicates students can learn words through instruction, even when words are unlikely to be in students' oral language or in the texts they read. However, when words are not used frequently in current language contexts, the carryover to students' overall vocabularies and comprehension appears limited. A question to be addressed in the next section of this review is whether instruction of words that may foster the comprehension of specific texts can be extended to words with similar functions in future texts that students will encounter. That is, can infrequent words such as *flickered*, a target word for a text in a core reading program (Beers et al., 2020), be used to illustrate the role of action and motion, even of inanimate objects, in narrative texts?

What research might support efficacious vocabulary instruction?

There will never be a perfect list of words to teach in English. The English lexicon is complex, but several conclusions are possible from this literature review as to the types of words that are likely to support increased textual comprehension. First, a small but growing body of evidence can be used to make fairly accurate predictions regarding which words will be relatively straightforward for students to understand. Further, tools are available so that these findings can be tested and elaborated with additional databases of students' vocabulary performances.

The availability of data on what features likely influence vocabulary knowledge can be used to distinguish among the many rare words that students are likely to encounter in text. Hayes et al. (1996) predicted approximately 3 rare words for every 100 words of text. That means the number of words in a text with 1444 words such as *The Storyteller's Candle*, which students likely have not encountered before, far exceeds six target words. Furthermore, current data can be used to predict which of these words merits direct instruction. Of the six target words from *The Storyteller's Candle* chosen for instruction (Beers et al., 2020), three are highly concrete and easily described with a word that is similarly concrete but has an age of acquisition at the kindergarten level and is easily decodable: *thin-slender*, *ring-chimed*, and *twinkle-flickered*. The three remaining target words are somewhat abstract, but one, *gallant*, has an easily decodable synonym, *brave*. The features of the other two abstract words—*concluded* and *preparations*—are polysyllabic, members of rich word families, somewhat complex to define, likely to appear in content areas, and central to the text's theme, which make them good candidates for instruction. A similar analysis can be applied to the other 35 words in the text, which are predicted to be relatively rare in text.

Findings from the existing research provide some guidance for curriculum developers in selecting words that merit direct instruction. At the same time, questions about which words require direct instruction and the nature of this direct instruction far outnumber answers at this point. In this final section, two issues are addressed that are central to supporting strong vocabularies for all students: (a) the need for an ELA vocabulary curriculum LA and (b) the need for databases for features such as conceptual complexity and polysemy and for assessments that give insight into students' knowledge of specific portions of the lexicon.

The need for vocabulary curricula

Students apply their knowledge of words in the context of specific texts, making the presence of vocabulary in texts a necessary and important consideration in the direct instruction of vocabulary in ELA. However, a topic that has not been considered within ELA is whether the words that are selected from specific texts can represent underlying constructs characteristic of genres. That is, can words for direct instruction be chosen on the basis of curriculum to ensure transfer to words in other texts? General academic vocabulary is one group of words that has begun to be treated as a curriculum. Other potential candidates for vocabulary curriculum can be identified. One such candidate pertains to the words that account for the majority of words in text or the core vocabulary. A second consists of the less frequent words that are unique to literary texts and the informational texts that are common in ELA.

Core vocabulary

As discussed earlier, a relatively small portion of the lexicon accounts for a majority of the words in texts. Within ESL, this core vocabulary is viewed as an essential foundation of proficient reading in English (Coady, 1997). Researchers in the field of ESL have the size of the lexicon required to read particular kinds of texts. According to Nation (2006), novel reading requires 9000 word families, and newspapers require 8000 word families. In the instruction of young children who are learning to read in English, however, attention to the core vocabulary has been limited to the high-function words that make up approximately 1% of the core vocabulary (Hiebert et al., 2018).

Students who cannot automatically recognize the core vocabulary appear to be challenged in learning words while reading, even less complex words (Cain et al., 2004; Shefelbine, 1990). Several interventions directed at increasing students' automaticity with the core vocabulary have demonstrated transfer to generalized comprehension measures. This finding has been reported with English learners (August et al., 2020), struggling readers (Fehr et al., 2011), and ESL students (Tozcu and Coady, 2004). The finding of transfer to generalized comprehension measures from core vocabulary interventions varies from research of direct vocabulary instruction (Elleman et al., 2009; Wright and Cervetti, 2017) where, as Hiebert and Nagy (2021) found, target words are typically infrequent in students' oral and written language.

Much needs to be understood, however, about core vocabulary interventions. Students may be familiar with many common words but may be challenged by their multiple meanings or their ability to make connections to morphological relatives. Furthermore, valid assessments are essential to ensure interventions do not cover words with which students are familiar. The present state of core vocabulary assessments is addressed in the final section of this review.

Vocabulary unique to literary and informational texts

As a result of guidelines from the National Assessment of Educational Progress (2009) and the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), the number of informational texts has increased substantially in many ELA classrooms. The design of ELA vocabulary curriculum, then, must address the words in informational as well as narrative texts.

Informational and narrative texts convey different types of content, which leads to the expectation that their vocabularies would vary. However, the selection of words for different genres requires more than assumptions. Several unique vocabulary features in narrative and informational texts have been documented. First, Gardner (2004) showed that, when common words (i.e., 2000 most frequent words) were eliminated, the unique words in narrative and informational texts on the same themes rarely overlapped.

Second, the less frequent vocabularies of the two genres come from distinctly different semantic clusters. In a comparison of target vocabulary in ELA and science programs, Hiebert and Cervetti (2012) reported that target words in ELA texts fell into clusters such as character roles, traits and ways of communicating. The majority of target words in science texts came from categories such as the natural environment and machines and materials.

Third, the cohesiveness of target words for individual selections followed unique patterns in the two genres. The target words for a science passage were cohesive, as illustrated by the words chosen for selections on sound and light that uniformly pertained to physical attributes of things, events, and experiences. By contrast, few of the words in a typical ELA passage fell into the same semantic cluster. Returning to the words selected for instruction with *The Storyteller's Candle*, none of the six target words (*chimed*, *concluded*, *flickered*, *gallant*, *preparations*, *slender*) came from the same semantic cluster.

Although the target words in ELA texts' individual passages represented an array of semantic clusters, words from similar clusters (e.g., communication, emotions, and attitudes) were targeted for most selections. Target words for selections that followed *The Storyteller's Candle*, for example, came from similar clusters (e.g., *graceful* (*slender*), *brilliant* (*gallant*)). This similarity of semantic clusters across narrative selections raises the possibility that instruction might focus on the function of particular words in literary texts rather than only on the meanings of individual words. That is, *slender* and *gallant* (and, subsequently, *graceful* and *brilliant*) could be used to demonstrate the crucial role of character traits in a story's plot and resolution.

An unresolved issue in the discussion of semantic clusters is the degree to which the words in these clusters interfere with one another in meaning. In the ESL field, vocabulary learning can be hampered when the meanings of words are too close. Using forms of nonsense words to simulate the learning of words in a nonnative language, Tinkham (1997) compared the learning of college sophomores regarding semantic clusters (e.g., *apple*, *pear*, *nectarine*, *peach*, *apricot*, *plum*), thematic clusters (e.g., *cave*, *explore*, *dark*, *stalactite*, *drip*, *cool*), and groups of unrelated words (e.g., *paint*, *funeral*, *recipe*, *market*, *uncle*, *ice*). Results led Tinkham to conclude semantic clustering hinders learning, whereas thematic clustering facilitates vocabulary learning.

In working with young learners who are learning to read in English rather than adult learners in ESL, the associative learning emphasized in ESL may not characterize the learning of semantic categories in that at least some of the words, such as the words in Tinkham's (1997) fruits category, are likely in young learners' vocabularies. Evidence for this suggestion that the learning of semantic clusters may function differently with young learners who are learning to read in English comes from Beck et al. (1982), which is one of the few vocabulary interventions to show an effect for distal comprehension measures (Wright and Cervetti, 2017).

In Beck et al.'s (1982) study, members of semantic groups were connected but not exact synonyms. For example, *glutton* was taught in relation to words related to eating (e.g., *nourishment*, *famished*). In a follow-up study, Stahl et al. (1992) taught the word *glutton* with words more closely associated in meaning (e.g., *obese*). The Stahl et al. study did not produce significant results, leading the authors to discourage the use of semantic clustering. It could be argued that Beck et al.'s categories are more thematic than semantic, whereas Stahl et al.'s were more semantically connected. Semantic and thematic word clustering as a mechanism for expanding students' vocabulary requires further investigation.

The target words for a selection in an informational text are of the type that Tinkham (1997) described as thematic such as *tomb*, *pharaoh*, and *pyramid* in a text on new ways of studying mummies. Several studies that have used a thematic approach in word selection have shown positive outcomes (Lubliner and Smetana, 2005; Simmons et al., 2010).

Word selection in informational texts in core reading programs does not appear to take advantage of these thematic connections. The pattern that Norris et al. (2008) reported well over a decade ago is apparent in a core reading program of the 2020s (Beers et al., 2020), as illustrated by the words chosen for a biographical text on an urban farmer (Briggs, 2016): *scarce*, *pollution*, *factories*, *vats*, *crowded*, *greenhouses*, and *vertical*. Some of the target words pertain to the theme, but additional theme-related words appear in the text (e.g., *composting*, *hoophouses*) and likely merit instruction attention more than words such as *crowded* and *scarce*.

The descriptive informational articles and biographies in ELA programs typically do not have the technical vocabulary of content-area textbooks. In a unit on recycling in a core reading program (Beers et al., 2020), words that might be expected in science lessons such as *biodegradable*, *polyurethane*, and *deforestation* do not appear, but words such as *plastic*, *recycled*, and *energy* do. ELA is not the sole context for the development of complex, content-area vocabulary, which is built through discussions, lessons, and activities of content areas. However, the manner in which the unique words of informational texts in ELA can be used to develop background knowledge on various topics that create the contexts of narrative texts and are germane to content areas requires consideration.

The development of databases and assessments

To increase our knowledge about and support for the best possible vocabulary instruction, work is needed in regard to expanding databases and creating assessments that clarify the portions of the lexicon with which students are and are not knowledgeable.

Databases

New databases have supported understandings of how word features influence vocabulary acquisition and retention, specifically age of acquisition and concreteness. But there are other word features that have been shown to challenge students, most notably conceptually complex words (Nagy et al., 1987) and polysemous words (Cervetti et al., 2015). In both the Nagy et al. and Cervetti et al. studies, the researchers developed coding systems that they used to manually code a limited number of words. Fitzgerald et al. (2017) proposed a system for conceptual categories, but this effort has not moved beyond experimental use to date. The existing large-scale database of polysemy—WordNet (Miller, 1998)—lacks precision.

Nagy et al. (1987) found a simple two-point measure could be used to predict conceptually complex words. It may be possible to tag large groups of words for conceptual complexity through a crowdsourcing technique similar to that used in recent databases of age of acquisition and concreteness-abstractness (Brysbaert et al., 2014; Kuperman et al., 2012). The development of tools that tag word features such as conceptual complexity and polysemy must be a priority of future vocabulary research.

Assessments

Vocabulary assessments, Pearson et al. (2007) observed, have been structured to compare students' vocabulary across groups, not to describe students' knowledge of specific bodies of words. Over the past decade, federal funds have been allocated to producing and validating new vocabulary measures. Several vocabulary instruments went through rigorous validation procedures as part of the Reading for Understanding (RFU) initiative. These assessments covered the range from K–2 (Foorman et al., 2015a) to grades 3–12 (Foorman et al., 2015b; Sabatini et al., 2015). Additionally, several individual vocabulary assessment projects were developed over the past decade (Goodwin et al., 2020; McKeown et al., 2017).

The measures designed over the past decade have several features that distinguish them from previous vocabulary measures. First, emphasis in these assessments is on morphology. Indeed, the grades 3 to 12 vocabulary subtest of the Florida Center for Reading Research's Reading Assessment (FRA) exclusively assesses morphologically related words (Foorman et al., 2015b). The Morphology subtest of the Educational Testing Service's (ETS) Reading Inventory and Scholastic Evaluation (RISE) assessment focuses on prefixes and/or suffixes attached to a root (Sabatini et al., 2015). Second, general academic vocabulary is emphasized. The *Core Academic Language Skills Assessment* is devoted entirely to academic language (e.g., general academic vocabulary; Barr et al., 2019), whereas one of the RISE subtests (Sabatini et al., 2015) assesses general academic vocabulary and content-area vocabulary, although separate scores do not appear to be given for the two vocabulary groups.

However, these new emphases in vocabulary assessments have not been accompanied by information on portions of the lexicon with which students are adept or challenged. Similar to previous vocabulary assessments, the new generation of assessments continues to provide information on students' performances relative to those of their peers. Scores of "proficient" on the FRA and RISE vocabulary subtests fail to indicate whether students will be able to establish the meanings of unknown words in typical grade-level ELA texts.

The ability to discern the meanings of words within the context of texts assumes proficiency with the bulk of words that make up the text—the core vocabulary. Assessment of students' proficiency with different bands of the core vocabulary has not been a focus of researchers and educators concerned with the literacy learning of young students learning to read in English. By contrast, this topic has been central in assessments within the ESL field (Schmitt et al., 2020). The *Vocabulary Levels Test* (VLT; Nation, 1990; Schmitt et al., 2001) typifies this commitment. The VLT establishes students' vocabulary knowledge at four frequency levels: 2,000, 3,000, 5,000, and 10,000 word families. Three items represent every 100 words within a 1000-word band. If a student gets half the words correct at the 2000-word level, then the number is multiplied by 10 (to represent all ten 100-word bands) to give students a score of 150. This score, however, does not indicate whether a student's correct responses came in the initial bands of the 1000 word families or were distributed intermittently across the bands.

Words are chosen randomly to represent a word-frequency band on the VLT, which means highly familiar words (e.g., *hat*, *boy*) are as likely to appear, as are less familiar words (e.g., *station*, *office*). For young adults in ESL programs, features such as familiarity, concreteness, and length may not influence vocabulary as they do with children learning to read English. An assessment—the Insight assessment's vocabulary subtest—aims to establish students' knowledge of words in the core vocabulary based on findings on word features to identify target words (Hiebert et al., 2019). Levels represent bands of words according to frequency, but words have been eliminated that are highly familiar and concrete and easily decoded. In the form of the CVA, this represents words with frequencies of 20–40 appearances per million—for example, words such as *cake*, *soap*, and *cup* are not assessed, but words such as *contrast* and *distribute* are. Administrations of the assessment have proven to predict comprehension as well as to support students' development by providing them with appropriate instruction.

Conclusion

Over 60 years ago, Gates (1961, 1962) showed that second graders knew over 75% of the third and fourth grade words that were recommended for instruction in core reading programs. That is, 2 years before these words would be taught to students, students had knowledge of them. Thirty years later, Stallman et al. (1989) came to a similar conclusion. Both second and fifth grade cohorts performed at levels of over 70% on words from subsequent grades. A similar study has not been conducted in the intervening years, but analyses of words selected for instruction (Hiebert and Cervetti, 2012) and interventions (Hiebert and Nagy, 2021) suggest choices about word selection for vocabulary instruction have not necessarily benefited from the insights of Gates and Stallman et al.

Current research provides directions for selecting words that may benefit from direct instruction. At the same time, many questions remain about the nature of this instruction. This review did not address the different methods of direct instruction. In particular, differential instruction for words with different features has not been addressed. Some words (e.g., *slender*, *chimed*) may require only a quick reference to a known synonym, whereas other conceptually complex words (e.g., *pollution*, *visionary*) may benefit from a lesson on semantically or thematically connected words, as well as morphologically related ones. The role that facility with the core vocabulary plays in ensuring students can apply strategies in the context of meaning-making of text also requires attention.

Also critical to the future of a global community are issues of applicability to and adaptivity of recommendations for word selection to the many languages used in literacy instruction around the world. Analyses of the texts used for early reading instruction in Egypt suggest the need for a theoretical framework in the selection of words to teach in Arabic (Wolsey et al., in press). Before conclusions can be made about word selection in other languages, substantial research is required on existing practices and the distribution of words in schoolbooks of different languages. It would be hoped that a review on word selection in the years to come will show that such research has been conducted and that solutions have been identified that support the goal of ensuring students have prowess in the vocabulary needed to comprehend the complex texts of the global communities of the 21st century.

References

- August, D., Uccelli, P., Artzi, L., Barr, C., Francis, D.J., 2020. English learners' acquisition of academic vocabulary: instruction matters, but so do word characteristics. *Read. Res. Q.* <https://doi.org/10.1002/rq.323>.
- Balota, D.A., Pilotti, M., Cortese, M.J., 2001. Subjective frequency estimates for 2,938 monosyllabic words. *Mem. Cognit.* 29 (4), 639–647.
- Barr, C.D., Uccelli, P., Phillips Galloway, E., 2019. Specifying the academic language skills that support text understanding in the middle grades: the design and validation of the core academic language skills construct and instrument. *Lang. Learn.* 69 (4), 978–1021.
- Beck, I.L., McKeown, M.G., Kucan, L., 2013. *Bringing Words to Life: Robust Vocabulary Instruction*. Guilford Press.
- Beck, I.L., Perfetti, C.A., McKeown, M.G., 1982. Effects of long-term vocabulary instruction on lexical access and reading comprehension. *J. Educ. Psychol.* 74 (4), 506–521.
- Beers, K., Armstrong-Carroll, J., Clemens, N., Cunningham, A., Hougen, M.C., Jago, C., Palmer, E., Probst, R.E., Templeton, S., Washington, J., 2020. *HMH Into Reading*. Houghton Mifflin Harcourt.
- Biemiller, A., 2010. *Words Worth Teaching: Closing the Vocabulary Gap*. McGraw-Hill SRA.
- Brett, A., Rothlein, L., Hurley, M., 1996. Vocabulary acquisition from listening to stories and explanations of target words. *Elem. Sch. J.* 96 (4), 415–422.
- Briggs, J.B., 2016. *Farmer Will Allen and the Growing Table*. Readers to Eaters.

- Brysbaert, M., Warriner, A.B., Kuperman, V., 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behav. Res. Methods* 46 (3), 904–911.
- Cain, K., Oakhill, J., Lemmon, K., 2004. Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge, and memory capacity. *J. Educ. Psychol.* 96 (4), 671–681. <https://doi.org/10.1037/0022-0663.96.4.671>.
- Carroll, J.B., Davies, P., Richman, B., 1971. *The American Heritage Word Frequency Book*. Houghton Mifflin.
- Cervetti, G.N., Fitzgerald, M., Hiebert, E.H., Hebert, M., 2019. The Influence of Vocabulary Instruction on Transfer Measures of Word Learning and Word Solving [Poster Presentation]. Annual Meeting of the Society for the Scientific Study of Reading, Toronto, ON.
- Cervetti, G.N., Hiebert, E.H., Pearson, P.D., McClung, N.A., 2015. Factors that influence the difficulty of science words. *J. Lit. Res.* 47 (2), 153–185.
- Coady, J., 1997. L2 vocabulary acquisition. In: Coady, J., Huckin, T. (Eds.), *Second Language Vocabulary Acquisition*. Cambridge University Press, pp. 273–290.
- Elleman, A.M., Lindo, E.J., Morphy, P., Compton, D.L., 2009. The impact of vocabulary instruction on passage-level comprehension of school-age children: a meta-analysis. *J. Res. Educ. Eff.* 2 (1), 1–44. <https://doi.org/10.1080/19345740802539200>.
- Fehr, C.N., Scholin, S.E., Davison, M.L., 2011. The effect of online, large-scale vocabulary instruction on listening and reading vocabulary. *Infor. Technol. Educ. Soc.* 12 (1), 5–18.
- Fisher, C.W., Berliner, D.C., Filby, N.N., Marliave, R., Cahen, L.S., Dishaw, M.M., 1981. Teaching behaviors, academic learning time, and student achievement: an overview. *J. Classr. Interact.* 17 (1), 2–15.
- Fitzgerald, W.J., Elmore, J., Kung, M., Stenner, A.J., 2017. The conceptual complexity of vocabulary in elementary-grades core science program textbooks. *Read. Res. Q.* 52 (4), 417–442.
- Forman, B.R., Petscher, Y., Schatschneider, C., 2015a. Florida Center for Reading Research (FCRR) Reading Assessment (FRA): Kindergarten to Grade 2 Technical Manual. Retrieved from. <https://eric.ed.gov/?id=ED580134>.
- Forman, B.R., Petscher, Y., Schatschneider, C., 2015b. Florida Center for Reading Research (FCRR) Reading Assessment (FRA): Grades 3 through 12 Technical Manual. Retrieved from. <https://files.eric.ed.gov/fulltext/ED580133.pdf>.
- Gardner, D., 2004. Vocabulary input through extensive reading: a comparison of words found in children's narrative and informational reading materials. *Appl. Linguist.* 25 (1), 1–37. <https://doi.org/10.1093/applin/25.1.1>.
- Gardner, D., Davies, M., 2014. A new academic vocabulary list. *Appl. Linguist.* 35 (3), 305–327.
- Gates, A.I., 1961. Vocabulary control in basal reading material. *Read. Teach.* 15 (2), 81–84.
- Gates, A.I., 1962. The word recognition ability and the reading vocabulary of second and third-grade children. *Read. Teach.* 15 (6), 443–448.
- Gonzalez, L., 2012. *The Storyteller's Candle*. Lee & Low Books.
- Goodwin, A.P., Petscher, Y., Tock, J., McFadden, S., Reynolds, D., Lantos, T., Jones, S., 2020. Monster, PI: validation evidence for an assessment of adolescent language that assesses vocabulary knowledge, morphological knowledge, and syntactical awareness. *Assess. Eff. Interv.* <https://doi.org/10.1177/1534508420966383>.
- Graves, M.F., Elmore, J., Fitzgerald, J., 2019. The vocabulary of core reading programs. *Elem. Sch. J.* 119 (3), 386–416.
- Hayes, D.P., Wolfer, L.T., Wolfe, M.F., 1996. Schoolbook simplification and its relation to the decline in SAT-verbal scores. *Am. Educ. Res. J.* 33 (2), 489–508.
- Hiebert, E.H., 2017. In Pursuit of a Parsimonious Vocabulary Curriculum in Middle School And beyond. Paper presentation. Annual Conference of the Literacy Research Association, Tampa, FL.
- Hiebert, E.H., Cervetti, G.N., 2012. What differences in narrative and informational texts mean for the learning and instruction of vocabulary. In: Baumann, J., Kame'enui, E. (Eds.), *Vocabulary Instruction: Research to Practice*, second ed. Guilford Press, pp. 322–344.
- Hiebert, E.H., Goodwin, A.P., Cervetti, G.N., 2018. Core vocabulary: its morphological content and presence in exemplar texts. *Read. Res. Q.* 53 (1), 29–49.
- Hiebert, E.H., Nagy, W.E., 2021. Which Words to Teach? An Analysis of the Target Words in Vocabulary Interventions [Paper presentation]. Annual Conference of the Society for the Scientific Study of Reading (Online Conference).
- Hiebert, E.H., Scott, J.A., Castaneda, R., Spichtig, A., 2019. An analysis of the features of words that influence vocabulary difficulty. *Educ. Sci.* 9 (1), 8.
- Jones, S.M., LaRusso, M., Kim, J., Yeon Kim, H., Selman, R., Uccelli, P., Snow, C., 2019. Experimental effects of Word Generation on vocabulary, academic language, perspective taking, and reading comprehension in high-poverty schools. *J. Res. Educ. Eff.* 12 (3), 448–483.
- Kuperman, V., Stadthagen-Gonzalez, H., Brysbaert, M., 2012. Age-of-acquisition ratings for 30,000 English words. *Behav. Res. Methods* 44 (4), 978–990.
- Lawrence, J.F., Capotosto, L., Branum-Martin, L., White, C., Snow, C.E., 2012. Language proficiency, home-language status, and English vocabulary development: a longitudinal follow-up of the Word Generation program. *Biling. Lang. Cognit.* 15 (3), 437–451.
- Lubliner, S., Smetana, L., 2005. The effects of comprehensive vocabulary instruction on Title 1 students' metacognitive word-learning skills and reading comprehension. *J. Lit. Res.* 27 (2), 163–200.
- Marulis, L.M., Neuman, S.B., 2010. The effects of vocabulary intervention on young children's word learning: a meta-analysis. *Rev. Educ. Res.* 80 (3), 300–335.
- McKeown, M., Deane, P., Scott, J., Krovetz, R., Lawless, R., 2017. *Vocabulary Assessment to Support Instruction: Building Rich Word-Learning Experiences*. Guilford.
- Mehri, A., Jamaati, M., 2017. Variation of Zipf's exponent in one hundred live languages: a study of the Holy Bible translations. *Phys. Lett.* 381 (31), 2470–2477.
- Miller, G.A., 1998. *WordNet: An Electronic Lexical Database*. MIT Press.
- Miller, L.T., Lee, C.J., 1993. Construct validation of the Peabody Picture Vocabulary Test—Revised: A structural equation model of the acquisition order of words. *Psychol. Assess.* 5 (4), 438–441.
- Milton, J., 2009. Measuring Second Language Vocabulary Acquisition. *Multilingual Matters*. <https://doi.org/10.21832/9781847692092>.
- Nagy, W.E., Anderson, R.C., 1984. How many words are there in printed school English? *Read. Res. Q.* 19 (2), 304–330.
- Nagy, W.E., Anderson, R.C., Herman, P.A., 1987. Learning word meanings from context during normal reading. *Am. Educ. Res. J.* 24 (2), 237–270.
- Nagy, W.E., Carlisle, J.F., Goodwin, A.P., 2014. Morphological knowledge and literacy acquisition. *J. Learn. Disabil.* 47 (1), 3–12.
- Nagy, W.E., Hiebert, E.H., 2011. Toward a theory of word selection. In: Kamil, M.L., Pearson, P.D., Moje, E.B., Afflerbach, P. (Eds.), *Handbook of Reading Research IV*. Routledge, pp. 388–404.
- Nagy, W., Townsend, D., 2012. Words as tools: learning academic vocabulary as language acquisition. *Read. Res. Q.* 47 (1), 91–108.
- Nation, I.S.P., 2006. How large a vocabulary is needed for reading and listening? *Can. Mod. Lang. Rev.* 63 (1), 59–82.
- Nation, I.S.P., 1990. *Teaching and Learning Vocabulary*. Newbury House.
- National Assessment of Educational Progress, 2009. *Reading Framework for the 2009 National Assessment of Educational Progress*. U.S. Department of Education.
- National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010. *Common Core State Standards for English Language Arts & Literacy in History/social Studies, Science, and Technical Subjects*. <http://www.corestandards.org/the-standards>.
- Nelson, J.R., Stage, S.A., 2007. Fostering the development of vocabulary knowledge and reading comprehension through contextually based multiple meaning vocabulary instruction. *Educ. Treat. Child.* 30 (1), 1–22.
- Norris, S.P., Phillips, L.M., Smith, M.L., Baker, J.J., Weber, A.C., 2008. Learning to read scientific text: do elementary school commercial reading programs help? *Sci. Educ.* 92 (5), 765–798.
- Oxford English Dictionary, 2005. Oxford Dictionaries. Oxford University Press.
- Pearson, P.D., Hiebert, E.H., Kamil, M.L., 2007. Vocabulary assessment: what we know and what we need to learn. *Read. Res. Q.* 42 (2), 282–296.
- Ryder, R.J., Slater, W.H., 1988. The relationship between word frequency and word knowledge. *J. Educ. Res.* 81 (5), 312–317.
- Sabatini, J., Bruce, K., Steinberg, J., Weeks, J., 2015. SARA reading components tests, RISE forms: technical adequacy and test design. ETS Research Report Series 2015 (2), 1–20.
- Schmitt, N., Nation, P., Kremmel, B., 2020. Moving the field of vocabulary assessment forward: the need for more rigorous test development and validation. *Lang. Teach.* 53 (1), 109–120.

- Schmitt, N., Schmitt, D., Clapham, C., 2001. Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Lang. Test.* 18 (1), 55–88.
- Schwanenflugel, P.J., Akin, C., Luh, W.M., 1992. Context availability and the recall of abstract and concrete words. *Mem. Cognit.* 20 (1), 96–104.
- Scott, J.A., Lubliner, S., Hiebert, E.H., 2006. Constructs underlying word selection and assessments tasks in the archival research on vocabulary instruction. In: 55th Yearbook of the National Reading Conference. NRC, pp. 264–275.
- Scott, J.A., Skobel, B.J., Wells, J., 2008. *The Word-conscious Classroom. Building the Classroom Writers and Readers Need.* Scholastic.
- Shelfbline, J.L., 1990. Student factors related to variability in learning word meanings from context. *J. Read. Behav.* 22 (1), 71–97.
- Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Willson, V., Byrns, G., 2010. A comparison of multiple-strategy methods: effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. *J. Res. Educ. Eff.* 3 (2), 121–156. <https://doi.org/10.1080/19345741003596890>.
- Stahl, S.A., 1999. *Vocabulary Development.* Brookline Books.
- Stahl, S.A., Burdge, J.L., Machuga, M.B., Stecyk, S., 1992. The effects of semantic grouping on learning word meanings. *Read. Psychol.* 13 (1), 19–35. <https://doi.org/10.1080/027027192130103>.
- Stahl, S.A., Fairbanks, M.M., 1986. The effects of vocabulary instruction: a model-based meta-analysis. *Rev. Educ. Res.* 56 (1), 72–110.
- Stallman, A.C., Commeyras, M., Kerr, B., Reimer, K., Jimenez, R., Hartman, D.K., Pearson, P.D., 1989. Are “new” words really new? *Lit. Res. Instr.* 29 (2), 12–29.
- Thorndike, E.L., 1921. *The Teacher's Word Book.* Teachers College, Columbia University.
- Thorndike, E.L., Lorge, I., 1944. *The Teacher's Word Book of 30,000 Words.* Teachers College, Columbia University.
- Tinkham, T., 1997. The effects of semantic and thematic clustering on the learning of second language vocabulary. *Sec. Lang. Res.* 13 (2), 138–163.
- Tozcu, A., Coady, J., 2004. Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. *Comput. Assist. Lang. Learn.* 17 (5), 473–495.
- Wagner, R.K., Ridgewell, C., 2009. A large-scale study of specific reading comprehension disability. *Perspect. Lang. Lit.* 35 (5), 27–31.
- Wolsey, T.D., Karkouti, I.M., Hiebert, E.H., Seoud, D.A.E., Abadzi, H., & Abdelkhalek, F. (in press). *Texts for Reading Instruction and the Most Common Words in Modern Standard Arabic: An Investigation.* Reading and Writing.
- Wright, T.S., Cervetti, G.N., 2017. A systematic review of the research on vocabulary instruction that impacts text comprehension. *Read. Res. Q.* 52 (2), 203–226.
- Zeno, S., Ivens, S.H., Millard, R.T., Duvvuri, R., 1995. *The Educator's Word Frequency Guide.* Touchstone Applied Science Associates.
- Zipf, G.K., 1935. *The Psycho-biology of Language: An Introduction to Dynamic Philology.* Houghton Mifflin, Boston, MA.

Spelling: theory, assessment, and pedagogy

Shane Templeton, University of Nevada, Reno, NV, United States

© 2023 Elsevier Ltd. All rights reserved.

Explorations in the development of spelling from emergent to proficient spellers: findings from research reflecting stage, phase and alternative theoretical approaches	374
Emergent and beginning spellers	376
Transitional spellers	376
Intermediate spellers	378
Proficient spellers	378
Implications and recommendations for assessment and developmentally-based instruction	379
Assessment	379
Instruction	380
A closer look at developmental constraints	382
Explorations in other languages	383
A concluding perspective	383
References	383

In recent decades spelling, or *orthographic* knowledge, has moved beyond being considered as simply a foundational skill for writing. It has acquired a more impressive cachet, claiming a broader linguistic, cognitive, and psychological terrain than it occupied up to the mid-20th century (Frith, 1980; Berninger, 1994, 1995; Brown and Ellis, 1994; Templeton and Morris, 2000). The pedagogical implications of this are broader as well, in part because research over the last several decades has consistently shown that orthographic knowledge is a significant if not a primary determinant in individuals' ability to *read* as well as to *write* words (Ehri, 2000, 2014; Perfetti, 2007; Perfetti et al., 1997; Rastle and Davis, 2008; Templeton and Bear, 1992/2011; Treiman and Kessler, 2014). Moreover, in English-speaking societies, a number of educators and scholars have pointed out that the spelling system of English is not a failed means of representing individual sounds, or phonemes, but rather a more logical system that works to represent both sound *and* meaning (Chomsky, 1970; Francis, 1958; Henderson, 1985; Templeton, 1979b; Venezky, 1999). The consequences for pedagogy of this research and better understanding of the logic of the system have been significant: In addition to its role as a foundational skill of written communication, spelling development now interfaces with word analysis/decoding and vocabulary instruction (International Literacy Association, 2019; Templeton and Bear, 2018; Templeton, 2020; Townsend et al., 2016).

This reconceptualization of spelling knowledge and its instruction also evolved because of the increasingly sophisticated analysis of children's and older students' spelling attempts as they interacted with the system. This work began with the seminal studies of Chomsky (1970) and Read (1971, 1975) and continued with the "Virginia Studies" of Edmund Henderson and his students at the University of Virginia (Beers and Henderson, 1977; Henderson, 1981, 1985, 1990; Henderson and Beers, 1980; Invernizzi et al., 1994; Templeton and Bear, 1992/2011) as well as Ehri (1980, 1992/2011, 1995, 2005, 2014), Ehri and Wilce (1987) and Rebecca Treiman (Bourassa and Treiman, 2008; Treiman, 1993; Treiman and Kessler, 2014). In recent years, the analysis of errors as well as their possible instructional implications is reflected in the work of a growing number of researchers (e.g., Adoniou, 2014; Bahr et al., 2012; Daffern et al., 2015; Devonshire et al., 2013; Masterson and Apel, 2010).

This chapter will address (a) the theoretical and research contributions to our understanding of how learners develop knowledge of the various types of information that the spelling system of English represents; and (b) the implications for instruction of these contributions for spelling and for its broader role in processing written language.

Explorations in the development of spelling from emergent to proficient spellers: findings from research reflecting stage, phase and alternative theoretical approaches

There are two components to spelling knowledge: (a) word-specific—knowledge of the spelling of individual words; and (b) general—knowledge about how the spelling system reflects phonology or sound, patterns of letters within syllables and between syllables, and meaning, or morphology. Treiman (2017) described learning to spell as "learning both about the visual appearance of written words and about how written words symbolize words' linguistic forms" (p. 645). Alphabetically-based orthographies may be characterized as *shallow* or *deep* (Bear et al., 2003; Katz and Feldman, 1983; Liberman et al., 1980). Shallow orthographies have a more direct, transparent, and consistent relationship between letters and the sounds they represent and are more quickly learned (Viise et al., 2011), while deep orthographies have a more opaque relationship. In the latter instance, the sounds that individual letters and groups of letters represent often depend on particular words and their structure or *morphology*. Compare, for example, the different sounds represented by the letters *m-e-d-i-c* in the words *medicine*, *medic*, and *medicinal*. Despite the variation in sounds, the spelling of the base *medic* does not vary. This is one reason English spelling has been described as *morphophonemic* (Francis, 1958;

Venezky, 1970, 1999) and more logical than commonly assumed: “There is considerable order in the spelling system. It is not a haphazard chaotic frenzy of letter-sound anomalies . . .” (Templeton, 1991, p. 186). Bowers and Bowers (2017) describe English orthography as “a morphophonemic system in which spellings have evolved to represent sound (phonemes), meaning (morphemes), and history (etymology) in an orderly way” (p. 2). What deep orthographies concede in ease of decoding print to sound may be reclaimed in ease of decoding print to meaning: words that are visually, or orthographically, similar are usually similar in meaning as well (Singson et al., 2000; Templeton, 1979a; Templeton and Scarborough-Franks, 1985).

For the past several decades, the theoretical and research landscape in spelling has been defined by *stage* or *phase* theories on the one hand and a number of alternative theories on the other—for example *overlapping waves* (Kwong and Varnhagen, 2005), *multiple linguistic repertoire* (Apel et al., 2004); *conjoint* (Bahr et al., 2009), *integration of multiple pattern* (Treiman and Kessler, 2014), and *triple word form* (Daffern et al., 2015). The explorations of invented spelling by Chomsky (1970) and Read (1971) provided a key for Edmund Henderson, one of the intellectual founders of modern stage theory in the development of orthographic knowledge, and his students to unlock the logic of the spellings of first graders (Henderson et al., 1972; Beers and Henderson, 1977), and later, of students throughout the elementary grades and beyond (Henderson and Beers, 1980; Templeton and Bear, 1992/2011). The concept of “stages” suggested by these Virginia studies as an epistemological framework emerged in part from the influence at the time of the Piagetian perspective in language and psychology (Piaget, 1973; Piaget and Inhelder, 1969) and Brown’s (1973) stages of language development. Suggesting early on that orthographic knowledge supports both reading and writing, Henderson (1980) described how the concept of stages reflected “a progressive differentiation of orthographic knowledge which in time underlies the abstract relational features by which words are identified in text and produced in written form” (p. 10).

From the beginning, however, Henderson was clear about the connotation of the term “stages.” In his classic text *Teaching Spelling* (1985) he emphasized, “The developmental stages of word knowledge, like the evolutionary periods of the language itself, are somewhat arbitrary divisions. Language change is continuous, and continuous, too, is the learner’s progress as he or she gradually masters English spelling” (p. 40). A few years later Henderson observed that “Developmental stages on any dimension of human behavior are always proximate statements with gray areas between them” (1992/2011, p. 24). Henderson’s students always emphasized the notion of stages was not “stringent” or rigid, and that more than one source of information or strategy could be available at different points in development. For example, Schlagal (2001) commented, “These stages describe common developmental themes seen across many children. They should not be taken as rigid, airtight, and invariant categories” (p. 153). The reason for emphasizing this point is because, following upon the stage theorists’ initial work, much of the research into spelling development assumed the Virginia studies’ stage-based perspective implied that learners focus *exclusively* first on sound, then on pattern, and only eventually on meaning, or morphology (e.g., Daffern et al., 2015; Devonshire et al., 2013; Treiman, 1993). Treiman and Cassar (1997), for example, stated, “Theories that treat learning to spell in terms of a sequence of stages that involve qualitatively different types of information are too simple. We must acknowledge that spellers, even young ones, are able to use various types of knowledge” (p. 78). The fundamental criticism of stage models, then, was that phonological influences do not extend to the later phases of spelling development, and morphological influences do not appear earlier in development. This point, however, was never in question; the current author observed that “[L]earners draw from *multiple sources of information* . . . the degree to which these different sources are used varies as learning progresses” (Templeton, 2003, p. 744, emphasis added) and that the developmental progression “involves complex processes in which a *number of different strategies* are called upon throughout the learning process” (Templeton and Morris, 1999, p. 105, emphasis added).

Throughout the course of continuous development, however, Henderson and his students noted that there appeared to be periods during which learners’ spelling attempts reflected their negotiation primarily, though not exclusively, of alphabetic/letter name information, within-word pattern information, or meaning/morphological information. Other stage theorists observed this developmental tendency (e.g., Nunes and Bryant, 2009), and some (Larkin and Snowling, 2008; Marsh et al., 1980; Frith, 1985) did argue for a more exclusive focus on sound first, morphology later. The work of Ehri (1992/2011), investigating the development of the ability to *read* words, interestingly complemented the work in developmental spelling; her developmental phases in reading words reflected stage theorists’ stages in spelling words. Ehri’s earlier work was framed within the stage paradigm (Ehri and Wilce, 1985, 1986), though she later explained, “We prefer the concept of phase, which is a less stringent way to characterize periods of development than the concept of stage. One phase may overlap with the next phase . . . Each phase highlights a characteristic of word learning that becomes prominent” (Ehri and McCormick, 1998, p. 140), and that prominent characteristic determines what is primarily taught.

To frame the insights from spelling research over the last few decades, two studies serve as appropriate bookends. Working within the Virginia stage theory paradigm, Schlagal’s comprehensive study of the “constancy and change” in spelling development (1992/2011) in grades one through six in the United States revealed the same error tendencies as in earlier studies described in Henderson and Beers (1980) and extended analyses to the upper elementary grades. Error types that remained constant for students through the grades included, for example, phonologically-based spellings of ambiguous vowels (digraphs such as *au*, diphthongs such as *oi*, r-colored such as *ir* and *ur*) and syllable juncture conditions determining whether or not consonants are doubled (*batted*, *admitted*, *orbited*). More recently, working within the Triple Word Form paradigm, Daffern’s comprehensive study (2017) with Australian students in years one through six found the same developmental trends that Schlagal had found for phonologically-based and for morphological spellings. The investigative focus in both stage and alternative paradigms in both studies emphasizes how learning at any point is not *only* sound, pattern, or meaning; as Templeton (2020) summarized, “Clarifying this distinction between the *predominant* focus of learners and an *exclusive* focus helps to bridge the misunderstanding between alternative and stage theorists” (p. 319).

In the following sections, focus will be on the nature of the findings from investigations by both stage and alternative theorists/researchers, and how developmental insights underlie the hierarchical, simple-to-more complex progression in spelling knowledge that integrates aspects of phonological, pattern, and morphological information. Although there will be some allusion to instructional influences, the following section will look more specifically at implications for instruction.

Emergent and beginning spellers

Even before letters appear, children's early writing attempts reflect characteristics of the print to which they are exposed (Ferreiro and Teberosky, 1982). Gradually, when letter-like figures and letters do appear, their frequency is related to how often they occur in written materials (Kessler et al., 2013; Pollo et al., 2009; Treiman, 2018; Treiman et al., 2018). With increasing exposure to print, being read to, and children's own attempts at writing, phonemic awareness begins to emerge. A number of researchers (e.g., Goswami, 2001) suggested that phonemic awareness develops through instruction in alphabetic orthography. At a fundamental level, this insight illustrates Olson's observation that "We introspect our language in terms of the categories laid down by our script" (1996, p. xii). The existence of things called "letters" and their consistent features present young children with the challenge of figuring out how they work and behave, and as children are guided by teachers and literate others in the examination of this behavior they develop awareness of and concept for an individual speech sound, or phoneme. "Print freezes the continuous stream of speech into perceptually manipulable blocks," Templeton (1980) observed, "and begins to pull a child's tacit knowledge about words to the surface where an explicit understanding can begin to develop" (p. 30).

Invernizzi and Buckrop (2018) found that learning letter names enables children to learn the sounds the letters make because most letter names contain their sound; for example, "B" contains/buh/, "M" contains/m/, "S" contains /s/. Spaces on the page, together with an understanding of beginning sounds in syllables, help the learner become aware of the printed word unit and how it matches to the spoken word unit; once this rudimentary concept of word in text is established, learners become aware of and learn how letters within the printed word unit correspond to sounds, and full phonemic awareness is developed (Bowling and Cabell, 2018; Flanigan, 2007; Morris et al., 2003a). "[T]he word begins 'to stand still' for analysis, [and children] can attend to other letter-sound properties within the word unit (e.g., the ending consonant)" (Morris et al., 2003, p. 321). Importantly, throughout this process children's *writing* plays a pivotal role; invented spelling powerfully supports the development of phonemic awareness (Copp et al., 2019; Ouellette and Sénéchal, 2017; Ouellette et al., 2013). This is first evident during the late emergent/early letter name stage (Bear and Templeton, 2000). During this time and throughout the beginning and transitional stages, the physical act of writing in development is critical, integrating the tactile and visual with sound (Planton et al., 2013; Tan et al., 2005). Developing orthographic knowledge is supported by shared neural substrates for both writing and reading (Longcamp et al., 2016; Richards et al., 2009).

The stage designation *letter name* to characterize the initial alphabetic stage arose from how young children use letter names to spell words (Read, 1971; Henderson, 1981) and then how they extended this understanding to how letters can represent sounds other than their names. Letter names are also cues, in other words, to spelling the sounds learners want to represent but which do not themselves have letter names to represent them. For example, toward the end of the kindergarten year a child wrote "ILIK SeT INDR MI FAVR TRe" (I like sitting under my favorite tree). There are letters whose names spell the long vowel sounds the child hears: LIK, MI, FAVR, TRe. She has not yet learned, however, that the letter *i* spells /i/, so in the word *sitting* she selects the letter name for *e* – /ē/ – because this sound is a cue to the sound she wants to spell: /ē/ is closer in articulation than any other vowel letter name to /i/. This decision reveals the child's tacit phonological knowledge (Read, 1971, 1975), and is an ontogenetic reflection of the long/short vowel pairings in English prior to the Great Vowel Shift (Templeton, 1976). As this child begins to learn the conventional spelling of short vowel sounds—an alphabetic rather than a simple letter name understanding—this information is now applied: MY Sitr WUS MEN TO ME WIN SHE STO MY TETY BER ("My sister was mean to me when she stole my teddy bear") (Gehsmann and Templeton, 2022). This understanding that a letter's name can also stand for a sound other than that name underlies beginners' correct representation of the plural marker *s* when it is pronounced /z/ as in *mitins* (*mitten*s), the earliest reflection of morphology in young children's spelling. Rather than reflecting an explicit awareness of its morphological function, however, children's correct spelling may come from the same understanding that a single letter can represent more than one sound, just as the letter *a* can represent /ā/ and /ă/.

Transitional spellers

Children's reading will present them with words the spelling of which represent more advanced patterns that do not reflect sequential left-to-right letter-sound pairings—for example, final *e*'s, vowel pairs or digraphs such as *ai*, *oa*, *ea*—and teachers usually address these patterns in "decoding" lessons. Both exposure to more advanced patterns through reading and explicit instruction in decoding these patterns influences children's spellings, and they grow from their linear, alphabetic expectation for how words are spelled to a relational, context-dependent expectation; that is, spellings and sounds that occur before and after a particular spelling determine that spelling (Invernizzi, 1992/2011; Schlagal, 1992/2011; Treiman and Kessler, 2006). These context-dependent patterns will reveal that English orthography is more predictable than often thought in its representation of sound. Treiman (1993, 2018) and Treiman and Kessler (2014) have explored the development of learners' awareness and understanding of this probability of particular spellings occurring in particular contexts. In describing the effects of this *statistical learning*, Treiman and Kessler (2006) observed "On this view, researchers might expect contextual effects to strengthen gradually over time, in line with children's

increasing opportunities to make the necessary observations. Researchers might also expect children to honor some contextually conditioned patterns earlier than others, with variations in speed of learning reflecting the statistical properties of the input" (p. 643). This describes the developmental continuum identified years earlier by stage theorists (Henderson and Beers, 1980; Henderson and Templeton, 1986).

When position or context of a spelling is taken into account, a greater degree of regularity emerges. For example: The *ai* spelling for long *a* occurs primarily in the middle of an English word (*rain*), never at the end; when long *a* occurs at the end of a word it is usually spelled *ay* (*stay*), less frequently *ey* (*they*) or *igh* (*sleigh*). The spelling of /j/ in single-syllable words is dependent on the quality of the preceding vowel; it is *ge* if the vowel is long (*page*) and *dge* if the vowel is short (*badge*). The sound /ch/ is usually spelled *tch* after a short vowel (*snitch*) but *ch* after a long vowel (*beach*). When children's spellings reflect their attention to these relational patterns, spelling instruction may systematically address them (Bear et al., 2020; Read and Treiman, 2013). For example, most children in grade two reveal their understanding of the function of silent *e* to signal a preceding long vowel sound through the correct spellings of *hope* and *shine*, but overgeneralize this knowledge through spellings such as *wate* for *wait*, *coche* for *coach*, and *frite* for *fright*. They can benefit from systematic spelling instruction in vowel digraphs.

In the absence of instruction addressing certain graphotactic patterns, the role of *tacit* or *implicit* learning based on their statistical probabilities has been investigated (Henderson and Chard, 1980; Treiman, 2017; Treiman and Wolter, 2018). Treiman and Kessler (2006) found that 8- to 10-year-olds were likely to spell /ē/ as *ea* rather than *e* before a final *d*; they were likely to spell the /ō/ sound with *a* after an initial *w*, as in *wander* (Kessler and Treiman, 2001). Because these patterns were not explicitly taught, it is likely that learners are tacitly picking up on these spellings based on their occurrence in print. Notably, younger learners—those whom stage theorists would describe as beginning, letter name/alphabetic spellers—are not sensitive to these types of context-dependent spellings.

At first described as the "vowel transition" stage, Henderson later suggested the term "within-word pattern," and that designation has remained in subsequent research. Both stage and phase theory emphasize that, although spelling instruction focuses primarily on single-syllable spelling patterns, this knowledge can be applied to decoding and spelling polysyllabic words. Explicit spelling instruction of the latter, however, will as explained in the next section not occur until a bit later. Early in the development of stage theory, Henderson observed that "some understanding of *base root constancy* probably accompanies the learning of the base vowel patterns" (Henderson, 1980, p. 143; emphasis added). Henderson and Templeton (1986) addressed what alternative theorists later referred to as the "overlapping" of different types of information when they observed that, during the within-word pattern stage "children also begin to deal with the third ordering principle of spelling, meaning. Readiness ... is first apparent when children begin to spell correctly the past tense ending *-ed*, despite its variation in sound" (p. 309). In addition, this is when "Teachers present the concept of base words, that is, that a base can stand alone after prefixes and suffixes are removed ... and pupils are able to learn the consistent spellings and meanings of many prefixes and suffixes" (pp. 309–310). This is because most of these affixes are predictably spelled according to pattern; at this stage, spelling many words of more than one syllable simply entails blending separate single-syllable morphemes together that are spelled according to pattern. Unlike the next significant stage of development in which spelling changes are often required when separate syllables and morphemes are conjoined, at this stage the process is more straightforward: for example, *un-* + *lock* + *-ed* = *unlocked*; *re-* + *turn* + *-ing* = *returning*.

Later studies of the development of morphological knowledge rediscovered Henderson's insight; for example, Robertson and Deacon (2019) concluded that "the relatively consistent patterns of inflectional markers in the early elementary grades ... provide a stable base" (p. 1053) for later morphological awareness and development across the grades. Deacon and Dhooge (2010), however, observed that it was not until grade four that spellings "maximally reflected" what they termed the "root consistency principle" (p. 1005). This also reflected a developmental tendency that stage theorists had observed many years before (e.g., Zutell, 1979, 1980). These findings have underscored the importance of a balance between spelling and vocabulary instruction. Deacon and Bryant suggested that "Teaching young children about root morphemes, to use magic in spelling both magical and magician, for example, may offer a simple way for children to learn about the spellings and meanings of quite complicated words" (2005, p. 387). Although this emphasis on explicitly presenting derivationally-related word families earlier than has traditionally been the case in the reading and vocabulary curricula of English-speaking countries is to be applauded, expecting students immediately or shortly thereafter to apply in their *spelling* all of the featural complexity those relationships reflect is not realistic (see below). For example, the spelling of the *-al* and *-ian* suffixes would remain a challenge for transitional spellers because there are not many exemplars in the children's reading that include those spellings, much less the accompanying context-dependent juncture spellings when they are affixed to base words. In the case of /shun/ there are not as many exemplars of derivational relationships with a corresponding orthographic representation that will help learners differentiate when to use one spelling versus a different spelling, as for example *illustrate/illustration*, *express/expression*.

Importantly, a span develops during the within word pattern stage between the ability to *read* many words accurately and the ability to *spell* those words correctly (Ahmed et al., 2014; Ehri, 2014; Morris et al., 1995a,b). As Ehri (2000) observed, "The amount of information to be drawn from memory is greater for spellers who must produce *multiple letters correctly sequenced* than for readers who must produce single pronunciations and meanings for written words" (p. 19; emphasis added). This in part is because learners' sensitivity to context-dependent orthographic patterns "emerges earlier in reading than in spelling" (Treiman and Kessler, 2006, p. 650). This is a reflection of the common psychological universal - the difference between *recognition* and *production*.

Intermediate spellers

As learners consolidate spelling pattern knowledge within single-syllable words they come to spell a majority of those words correctly. They often commit errors, however, at syllable junctures in words of two or more syllables and within unaccented syllables: *triping* for *tripping*, *slamed* for *slammed*, *leter* for *letter*; and later, *sailer* for *sailor*, *middel* for *middle*, *forchinet* for *fortunate*. These types of errors indicate a learner is in the *syllables and affixes* stage (Bear et al., 2020). Though the simple inflectional endings *-ed* and *-ing* come to be spelled correctly during the within word pattern stage, the conditions that determine the spelling of junctures when these inflectional endings are affixed to the base word are not understood and consistently applied until this stage: there may be no change (*claiming*), consonant doubling (*tapping*), or e-drop (*taping*). The spelling at syllable junctures signals important phonological information; in *tapping* the second *p* keeps the vowel sound in the first syllable short—otherwise the vowel would be long, as in *taping*. These conditions are then extended beyond affixation to the spelling of junctures within base words: *supper*, *super*.

There are of course “exceptions” to the vowel-consonant-vowel syllable juncture pattern that signals a long vowel. When these occur, a word-final spelling may be the clue (Berg, 2016; Treiman and Wolter, 2018). For example, consonant letters rarely double before *-ic*, *-it*, or *-ish* (*panic*, *comic*; *visit*, *merit*; *finish*, *punish*). Because vocabulary study at this level includes a more systematic focus on Greek and Latin roots, exploring the etymology underlying the spelling of many words that end with *-ic*, *-it*, or *-ish* often reveals a Latin root that must retain its spelling in the face of a syllable-juncture “rule”: The *n* in *finish* and *punish* does not double because the spelling of the roots *fin* and *pun* must be retained (“end”; “punish”). This type of attention helps learners realize that “exceptions” may not in fact be exceptions, but upon closer inspection, are revealed to follow patterns as well, in this case, motivated by context (spelling of an ending) and etymology.

During the latter period of this stage, learners’ spellings reflect their increasing attempts to negotiate spelling-meaning connections across words that are related through derivational morphological relationships. These connections are primarily those in which there are few if any phonological and orthographic changes within the base (e.g., *excite*-*excitement*; *ease*-*easy*-*uneasily*). English orthography becomes critical in the more expansive exploration of derivational morphology in vocabulary growth; the ability to consistently spell more advanced relationships follows later in the Proficient stage (Templeton, 1979a; Templeton and Scarborough-Franks, 1985; Zutell, 1980).

Proficient spellers

At this stage individuals become quite competent spellers and their errors reflect higher-order aspects of English orthography, for example *oppisition/opposition*, *accomodate/accommodate*, *circumfrence/circumference*, *advisible/advisable*, *dominance/dominance*, *irelevant/irrelevant*, *privelege/privilege*. A number of researchers have consistently found that knowledge of derivational morphological patterns in spelling follow an extended developmental arc across many grades (e.g., Treiman and Kessler, 2014; Robertson and Deacon, 2019; Templeton and Scarborough-Franks, 1985). The present author found that even at grades 6 and 8 students were still uncertain about spellings in which a phonological shift across derivationally-related words occurs, particularly where vowel reduction occurs as in *incline-inclination* (Templeton, 1979a; Templeton and Scarborough-Franks, 1985), a finding confirmed by subsequent studies (Robertson and Deacon, 2019; Treiman and Cassar, 1997; Tyler and Nagy, 1989). Treiman and Cassar (1997) observed, for example, that “the ability to take advantage of complex derivational relations, such as the relation between *courage* and *courageous*, is a late development” (p. 168). Robertson and Deacon summarized how “derivational morphology [in spelling] has a far more protracted period of development” (p. 1053). Nagy et al. (2006) concluded that the “role of morphology in the writing system [spelling] is grasped later than basic insight that new words can be analyzed as combinations of familiar parts. This pattern is consistent with Bear et al.’s (2004) developmental approach to spelling” (pp. 144–145).

Proficient spellers explicitly understand and apply spelling-meaning connections shared by words with more transparent spelling and sound relationships, such as *courage/courageous* and *ignite/ignition*. Most do not notice, however, more opaque relationships among words unless they are pointed out (Templeton, 1989); for example, those in which phonological changes occur without changes in spelling of the base or root (*confer/conference*, *custody/custodian*, *mandate/mandatory*, *facile/facilitate*, *spectator/inspection*, *purge/expurgate*, *territory/subterranean*) and those in which these changes occur with changes in spelling (*provoke/provocation*, *reclaim/reclamation*, *sustain/sustenance*). Almost all of these higher-level processes involve derivational morphology and classical Greek and Latin word roots, the understanding of which is also embedded within advancing vocabulary knowledge and increasing potential for understanding the role of etymology in vocabulary and spelling. At this level, a generalizable knowledge of higher-order derivational morphophonological rules including stress assignment in longer polysyllabic words is dependent on higher-order orthographic knowledge (Moskowitz, 1973; Templeton, 1979a). Rather than bringing phonological competence to bear on the examination of orthographic structure—the challenge for younger readers and spellers—the process at this level is reversed, and “the spelling system may influence the generalization of this higher order phonological knowledge” (Templeton, p. 256).

As both stage and alternative theorists have noted, throughout development the spelling of words can bring to explicit levels of awareness their morphology, and thereby their meaning (*magic-magician*). There may be developmental constraints, however, that preclude correctly spelling all of the new polysyllabic words learners will encounter (see below). At the proficient stage, these constraints are no longer a significant factor. Students can explore how spelling and morphology interact, as well as determine pronunciation, across successively more opaque relationships; for example in families such as *convene*, *convention*, *covenant*, *coven*, and *manual*, *manuscript*, *manufacture*, *manumission*, *manure*. Apparent anomalies help to expand understanding; having learned that

the Latin root *dec* means “ten” in words such as *decimeter* and *decade*, confusion at first results from attempting to apply this knowledge to the words *deciduous* and *decadence* in which *dec* does not mean “ten.” Exploration of how the letter sequence is parsed, however, reveals that two spellings of the Latin root meaning “fall”—*cid* and *cad*—combine with the prefix *de-* to express the sense of “falling away” (in the first sense, leaves falling away; in the second, a moral falling away). Again, such investigation is clearly within the purview of vocabulary development, but at this developmental level the role of orthographic representation of morphology may play a significant role in facilitating this more advanced morphological development.

The key difference between stage theories and the alternative theories that developed in response to them is the degree to which different sources of information are emphasized in instruction: What are the types of phonological, orthographic, and morphological information that teachers should be sure are available to learners at different points along the developmental continuum, and to what degree should learners be held accountable for applying these types of information consistently and correctly in their spelling? Stage theorists have noted that, although learners may be sensitive to several types of information as they spell, they consistently demonstrate productive understanding of certain types of information before others. This is acknowledged in fact by most alternative theorists as well, as their following observations attest:

All the spelling knowledge threads are useful at all grades, *although we may expect some to dominate at certain grades*
 IIA (2019, p. 9, emphasis added).

There may be greater explicit understanding of some concepts before others, e.g., understanding of magic “e” may precede understanding of when to apply the -ed rule
 Critten and Pine (2009, p. 104).

it seems that some phonological skills may be necessary before other specific phonological and morphological skills emerge
 Daffern (2017, p. 322).

In summary, most spelling theorists and researchers—stage, phase, conjoint, repertoire, overlapping waves, integration of multiple pattern, Triple Word Form—concur that the acquisition of orthographic knowledge in alphabetically-based writing systems follows a developmental trajectory that reflects the increasing complexity of the spelling system (Bahr et al., 2020; Templeton and Bear, 2018). The farther along this trajectory learners advance, the more information from phonological, pattern, and morphological influences are they able to draw and integrate so as to construct their understanding. It is time to turn to the specific implications for learning this constructive, integrative process.

Implications and recommendations for assessment and developmentally-based instruction

Gehsmann et al. (2019) emphasized that assessment, and instruction based on that assessment, should be “informed by a deep understanding of the developmental milestones associated with learning to read and write proficiently” (p. 2). This understanding, described in the first section of this chapter, underlies the discussion of spelling assessment and instruction in this section. Because English spelling represents phonological, pattern, and meaning/morphological information, how spelling instruction may be situated in a broader word study curriculum that includes phonics/decoding and vocabulary instruction will be touched upon.

Assessment

The cognitive psychologist Charles Perfetti observed, “Spelling and reading use the same lexical representation” (1993, p. 170) and therefore “spelling is the purest indicator of lexical quality” (1997, p. 30). His observation built on Henderson’s earlier conclusion (1980) regarding spelling knowledge underlying both reading and writing (Perfetti, 1986). Thus, assessing a student’s spelling should provide insight into what types of information that student is using to *read* as well as to *spell* words.

Children’s writing before they know what words are reveals their developing conceptions of print and the writing system to which they are exposed (Puranik et al., 2011; Zhang et al., 2017). They begin to incorporate letter-like features and actual letters in their writing, and eventually use letter names as cues to the “pieces” of spoken language they wish to represent (Beers and Henderson, 1977; Mann et al., 1987; Read, 1971; Treiman and Wolter, 2020). Children’s invented spelling is a very good predictor of reading success in first grade (Morris et al., 2003b; Mann, 1993), and often, “a better predictor than an oral phoneme segmentation task” (Morris et al., 2017). Its assessment value lies in the nature of the task, which is arguably more ecologically valid than a number of phoneme segmentation tasks; as Morris et al. (2017) pointed out, “the letter-by-letter spelling process slows down the process of phoneme segmentation” (p. 310). When writing *sit*, for example, the letter *s* and its sound, /s/, remain in place on the page, in effect freezing the first sound in the word, allowing the child to attend to other sounds, /i/ and /t/, in the word. This process, spelling letter-by-letter, enhances the ability to perceive and attend to individual sounds within a syllable or word.

A number of spelling assessments have been developed that attempt to identify the types of phonological, orthographic/pattern, and morphological information that learners are attempting to negotiate (Bear et al., 2020; Daffern, 2017; Ganske, 1999; Masterson and Apel, 2010; Silliman et al., 2006; Treiman et al., 2016). Stage and most alternative theorists agree that instruction should usually be grounded in the types of spelling errors that learners make as they develop knowledge of the spelling system, but they differ somewhat in the conclusions they draw as to the degree to which the errors should inform the focus of spelling instruction. Given the importance of assessment in determining the type, source, and degree of information that should be specifically included in spelling instruction, the assessment and interpretation perspectives of stage and some alternative theories are examined here.

Consider the spelling errors of a fifth grader who is struggling with reading and spelling (Masterson and Apel, 2010):

LOWFS (<i>loaves</i>)	CRETICES (<i>criticize</i>)
FRUTE (<i>fruit</i>)	MUZISHUN (<i>musician</i>)
CHOSE (<i>choice</i>)	CATIONS (<i>cautious</i>)
BLOSE (<i>blouse</i>)	

This student's errors fall across several categories: vowel patterns and diphthongs within single-syllable words; words of two or more syllables, including those that reflect more advanced features of derivational morphology, for example, *musician* and *criticize*. This profile is usually a reflection of developmentally-inappropriate instruction but may also suggest a learning disability; regardless, the developmental sequence holds for struggling readers and children identified as dyslexic as well as typically-developing readers (Bourassa and Treiman, 2008; Worthly and Invernizzi, 1990). Some alternative theorists suggest that all these types of errors be addressed in spelling instruction (Masterson and Apel, 2010). From a stage theory perspective, let's examine why this may not be advisable.

The errors in the first column represent the predominant type of information this student is trying to negotiate; the errors are characteristic of within word pattern spellers (Bear et al., 2020). This student is still learning about ambiguous, less-frequently-occurring vowel patterns as well as diphthongs, and spelling instruction should focus on one-syllable words with these patterns. The misspelled words in the second column reflect attempts to spell higher-level derivational patterns. These morphological relationships should be pointed out and discussed in terms of *vocabulary* relationships. In contrast to repertoire theory, however, the student's spelling errors in these words should not lead us to expect that this student should study and learn to correctly *spell* a wide range of higher level, derivationally-related words. Schlagal (2001) noted that attempts to do so "invites rote memorization of words, an inefficient and often impermanent form of learning" (p. 160).

Morris and his colleagues offered further support for focusing spelling instruction at the appropriate developmental level (e.g., Morris, 2011, 2021; Morris et al., 1995, 1986). Investigating the learning and retention of specific spelling words at students' instructional and frustration levels they found that, despite focused study, words at students' frustration levels (such as *cautious* and *criticize* for the student above) most likely would remain consistently misspelled. There is simply not enough underlying orthographic knowledge to support memory for the correct spelling of many such words, much less abstracting an understanding of the underlying pattern. In addition, students have often not yet learned to read these words. As noted earlier, this span between learning to read words and eventually to spell those words correctly reflects the basic psychological principle of correct recognition preceding correct production. This is why expecting students to learn to recognize new words *and* to learn and retain correct spelling of those words is to teach at their spelling frustration level—very little learning occurs. And for students who are struggling, "Reading disabilities can be prolonged and/or exacerbated" (Invernizzi and Hayes, 2011, p. 205).

Instruction

Based on students' spelling errors on well-constructed assessments, the stage model of development provides insight into the instructional balance between which aspects or components of orthography may be primarily addressed in *spelling*, and when, and which may be addressed in *word analysis/decoding* and *vocabulary*. Words at spelling frustration levels and the types of information they represent can be the focus for word-analysis or decoding instruction.

Decoding instruction—pointing out "what's going on" in words (such as the role of silent *e*, for example)—establishes the signpost. When these features begin to appear in children's spelling, either in their own compositions or on spelling inventories, they are an indication that children are ready to explore associated patterns in their spelling instruction. Vocabulary instruction may include attention to derivationally-related words, noting semantic relationships, for example, between *music-musician* and *critic-criticize*. If we try, however, to get learners to *spell* such words correctly (often believing such work will reinforce memory for the words) we will inadvertently require them to try to work at a frustration level. The *instructional* level work builds and reinforces the understandings into which the new patterns will fit.

What might engaging instructional-level activities look like? Emphasizing the importance of teaching children at all levels the logic of the spelling system, Bowers and Bowers (2017) observed, "Psychological theories of learning highlight the importance of organizing information in meaningful ways, as well as the importance of explanatory questioning in which students generate plausible explanations as to why some stated fact is true, as well as monitoring and reflecting on their learning" (p. 15). Bowers

and Kirby's *structured word inquiry* (2010) reflects this organization and focus, as do *focused contrasts* involving word sorting or categorizing words according to different criteria (Bear et al., 2020). Interestingly, this latter approach is not new. Botel (1964) observed that "we are not concerned with teaching rules, but rather with manipulating patterns, encouraging more accurate observations, and extending collections of words that belong together for some analogical reason" (p. 192). This perspective was advanced by Russell Stauffer (1969, 1970) at the University of Delaware reading clinic and later by his student, Edmund Henderson, at the McGuffey Reading Center at the University of Virginia (Henderson, 1981). Many years later, other researchers such as Nunes and Bryant (2009) and Treiman (2017) emphasized how conditional or context-dependent spellings are best learned through contrasting different cases. "Providing children with structured opportunities to discover patterns in written words," Treiman concluded, "can not only teach them about the patterns but also *increase their interest* in why words are spelled as they are" (2017, p. 649, emphasis added). Cartwright (2012) emphasized sorting activities as a way for educators to "provide multiple opportunities to think and talk about both sounds and meanings to foster children's flexible thinking about those elements of print" (p. 33).

Following are three examples illustrating how focused contrasts in word sorting activities can lead to insight and generalization about the logic underlying the English spelling system:

For transitional spellers and readers, the context-dependent spelling of the final /k/ sound is explored by sorting words such as *pack*, *broke*, and *cook* into columns according to final /k/ spelling:

<i>kick</i>	<i>bike</i>	<i>look</i>
<i>pack</i>	<i>take</i>	<i>shook</i>
<i>sock</i>	<i>strike</i>	<i>cook</i>
<i>duck</i>	<i>broke</i>	<i>took</i>

Through comparison/contrast, and if necessary the teacher's facilitative questioning (Bear et al., 2020), learners realize that *ck* is preceded by a short vowel sound, *ke* by a long vowel sound, and *k* by *oo*.

Later in the development of transitional spellers and readers, the context-dependent spelling of the final /ch/ sound is explored through sorting and examining words such as *catch*, *fetch* and *reach*, *coach*:

<i>catch</i>	<i>reach</i>
<i>pitch</i>	<i>coach</i>
<i>fetch</i>	<i>speech</i>
<i>match</i>	<i>screech</i>
<i>hutch</i>	<i>roach</i>

Students learn that a short vowel signals the following /ch/ sound is spelled *tch*; a long vowel signals *ch*. Exceptions such as *rich* and *much* are few, though they occur frequently in print and that frequency supports memory for spelling.

For intermediate spellers, the following sort explores four different spellings of the final stable syllable /el/. By sorting words into categories according to the different spellings *-al*, *-el*, *-le*, or *-il*, this examination leads students toward awareness and application of statistical probabilities as well as more advanced morphological connections that support learning and remembering the different spellings:

<i>whistle</i>	<i>channel</i>	<i>legal</i>	<i>pupil</i>
<i>bicycle</i>	<i>level</i>	<i>normal</i>	<i>fossil</i>
<i>puzzle</i>	<i>label</i>	<i>local</i>	<i>pencil</i>
<i>angle</i>	<i>quarrel</i>	<i>formal</i>	
<i>needle</i>	<i>angel</i>	<i>metal</i>	
<i>stumble</i>	<i>travel</i>		
<i>struggle</i>	<i>nickel</i>		
<i>article</i>			

After sorting the words according to the spelling of the final syllable, students are asked "Which /el/ spelling occurs most often? Least often?" Asking this question makes explicit the principle that all spellings are not equally probable, and being aware of the most frequent or frequent spellings can be of help. Interestingly, for the *-al* spellings a meaning connection may be made to related words in which the final syllable is clearly pronounced, providing a key to remembering the unaccented *-al* spelling:

<i>legal</i>	<i>legality</i>
<i>local</i>	<i>locality</i>
<i>formal</i>	<i>formality</i>
<i>normal</i>	<i>normality</i>
<i>metal</i>	<i>metallic</i>

The words that provide the clue for the *-al* spelling may be unfamiliar to the students, but by presenting them the teacher not only clears up the uncertain spelling, but importantly, expands their vocabularies as well.

A number of alternative theorists emphasized that the role of *meaning* and *etymology* in the spelling system should be addressed from the beginning of instruction (e.g., Ramsden, 1993; Adoniou, 2016; Bowers and Bowers, 2017; International Literacy Association, 2019). In so doing, children will therefore learn from the outset that the spelling system is more regular and logical because, these theorists argue, it is *not* primarily based on sound. For example, silent letters that are etymological markers explain the silent *w* in *two*: *twin*, *twice*, *twenty* (Devonshire and Fluck, 2010); by grouping *do*, *does*, *done*, and *don't* together the meaning relationship they share despite the different pronunciations of *do* becomes apparent (Bowers and Bowers, 2017); stories that explain other silent letters should be shared—in *know* and *knew* the letter was once pronounced, and the original Old English spelling was *cnawan* (Adoniou, 2016). Etymology also includes information about form, or graphotactics: English words cannot end with the letter *v* because hundreds of years ago, if *v* were doubled to mark the short vowel in words such as *have* and *love*, it would have looked like a *w*, so the practice of adding an *e* took hold.

Although it is probably the case that young children can benefit from learning that there are other reasons besides representing sound for the way words are spelled, results are mixed regarding the degree to which this additional information is incorporated (Devonshire et al., 2013; Wolter and Dilworth, 2014). At the very least, however, these few examples may help children become aware of the important *process* of etymology. This awareness may help in developing an important and facilitative mindset that there is a logic to the spelling of English, and it can help to explain what on the surface appear to be exceptions and irregularities (Adoniou, 2014; Daffern and Critten, 2019).

A closer look at developmental constraints

The apparent developmental constraints on the depth and range of orthographic understanding may be explained in part by developing brain structure and function (Sussman et al., 2016). Advances in neuroscience provide insight into understanding cognitive development and psychological processes including the learning and processing of written words. Beginning at about age 8 and reaching its peak between the ages of about 11–15, neuroanatomical development may help more specifically explain the consistent findings regarding learners' understanding and application of progressively more abstract orthographic relationships among phonology, pattern, and morphology (e.g., Brown, 2017; Brown et al., 2012; Johnson, 2011; Liebig et al., 2017; Perrachione et al., 2016; Schlaggar and McCandliss, 2007). Brown and colleagues note that "developmental alterations in the fundamental brain tissue biology and neural architecture are much more tightly linked to chronological age throughout childhood and adolescence" than was previously known (2017, p. 4). This may also underlie the significant increase in the functioning of working memory noted by Gathercole and Baddeley (2014) that occurs between 10 and 12 years of age, in turn supporting learners' engagement with longer and more orthographically complex words.

This research helps to explain why younger learners are aware of morphology and their attempted spellings reflect this awareness, but that it is not until a few years later that they are able to extend this understanding to consistently and correctly spell derivationally-related words in which phonological shifts occur, whether or not these shifts are accompanied by orthographic change (*electric-electrician*; *explode-explosion*). The generalizability, therefore, of Deacon and Bryant's (2005) suggestion regarding the use of *magic* in spelling both *magical* and *magician* as being a "simple way for children to learn about the spellings and meanings of quite complicated words" (p. 387) may not be as productive as one might hope (Rastle, 2019). Moreover, a large number of words reflecting more advanced derivational morphological relationships reflect increasingly opaque semantic relationships (e.g., *custody-custodian*; *facile-facility*; *pugnacious-impugn*). By about 9–10 years of age, however, understanding and application of morphological processes in vocabulary (e.g., Anglin, 1993) and in spelling begin to grow significantly. As Templeton (1979a) concluded, "Thus, a more efficient correspondence between an individual's sound system and the orthography may be, as Carol Chomsky (1970) expressed it, a 'late intellectual product'" (p. 262).

This should not, however, suggest a rigid deterministic framework for learning. Seeds for more advanced processes may be sown along the developmental path even at the outset—for example *magic* and *magician* and Greek roots such as *geo* and *micro* (Freeman et al., 2019)—but expectations for significant growth may be realistically tempered by developmental constraints and the consequent allocation of valuable instructional time and engagement to orthographic patterns and processes that firmly establish the foundation for later growth. As was discussed, most transitional students' word study, for example, should address spelling patterns that they are ready to examine in single-syllable words, because these words and their patterns and words occur with considerable frequency in reading materials at this level. The patterns occur in two-syllable words as well, so students' understanding of these patterns may be applied in decoding words of more than one syllable: For example, understanding the *ai* spelling in single-syllable words such as *rain* may be applied in decoding two-syllable words such as *daily*. Examining patterns through comparison

and contrast, discussion, and writing ensures that they will become internalized, making more rapid recognition and production possible.

In this latter regard, because orthographic knowledge underlies reading as well as writing, there is a body of research addressing the visual perception of words that reflects and explains what spelling stage theorists have found regarding the importance of establishing an appropriate spelling instructional level and the constraints on attempting to learn beyond that level. Research investigating the development of the perceptual span in reading has consistently supported the characterization of word knowledge as a developmental process, advancing in stages (Rayner et al., 2011); importantly, this knowledge underlies the development of fluency in reading (Chace et al., 2005; Clifton et al., 2007; Gehsmann et al., 2017; Li et al., 2014). Templeton and Bear (2018) summarized, “Readers’ underlying orthographic knowledge determines their ability to pick up an initial image based on the visual display of print, then to plan and move to the next fixation point ... the farther along the developmental [orthographic] continuum they have progressed ... the more letters they will perceive during a fixation” (p. 214). Again, the importance of appropriate, developmentally-based instruction is emphasized, particularly at the earlier stages. Teaching at these developmental levels involves “clarifying emerging concepts through the careful study of words that illustrate those concepts” (Schlagal, 2001, p. 160). In the context of engaging instructional activities such instruction provides sustained examination of simpler patterns without undue distraction of other visual and/or phonological and morphological factors. This helps learners abstract and solidify pattern knowledge that will drive reading fluency through directing and focusing examination of words during reading.

Explorations in other languages

Until the turn of this century, spelling theory and pedagogy evolved primarily from research with native English speakers learning English orthography. Since then, however, research in other languages with alphabetic orthographies has contributed significantly to the developmental landscape (e.g., Awramiuk, 2014 [Polish]; Dai et al., 2013 [Arabic]; Defior et al., 2008 [Spanish]; Diamanti et al., 2018 [Greek]; Ford et al., 2018 [Spanish]; Joshi and Aaron, 2006 [over two dozen languages and their orthographies]; Helman et al., 2015 [Spanish]; Ravid, 2012 [Hebrew]; Sénéchal, 2000 [French]; Shany and Share, 2011 [Hebrew]; Treiman and Kessler, 2021; Viise et al., 2011 [Estonian]). Though beyond the scope of this chapter, readers are referred to promising work in a number of other languages and orthographies. For example, consonant doubling patterns in Dutch follow a similar developmental trajectory as in English (Hilte and Reitsma, 2011), and derivational morphological relationships exist in a number of Romance languages that also follow similar developmental trajectories, as with the silent-sounded alternation pattern in French (*ouvert-ouverte*; Sénéchal, 2000) and consonant doubling with assimilated prefixes (*atterrir* = *ad-* + *terrir*; “to land”; Fejzo, 2016).

Work in languages with non-alphabetic scripts, including those that are nonlinear, is increasingly informing the developmental perspective (Daniels and Share, 2018; McBride, 2016 [Chinese]; Share, 2008; Shen and Bear, 2000 [Chinese]; Tong et al., 2009 [Chinese]; Winkler and Iemwanthong, 2010 [Thai]; Zhang et al., 2017 [Chinese]). Young children’s invented spelling attempts in a number of different languages and scripts is a particularly promising area of investigation (e.g., Kwak, 2007 [Korean heritage]; Martins et al., 2013 [Portuguese]; Mirzaei, 2003 [Farsi]; Yaden et al., 2015; Yaden and Tsai, 2011 [Chinese]; Zhang et al., 2017 [Chinese]). There is a growing body of research investigating the spelling of emergent bilinguals in English (Helman, 2016; Kiernan and Bear, 2018), and developmental spelling assessments have been developed for native Spanish, Korean, and Mandarin Chinese speakers (Bear et al., 2020; Helman et al., 2012; Shen and Bear, 2000).

A concluding perspective

As writing and reading increasingly occur in fluid and fractal digital spaces, the question is often raised concerning the future of spelling instruction: Won’t spell-check remove the need for attention to spelling—indeed, the need to attend to the written structure of words at all? The reconceptualization of the role of orthographic knowledge in the enterprise of literacy suggests not (Templeton, 2006). Because spelling knowledge underlies reading as well as writing, written words and the information their orthography encodes will still need to be examined and learned. In the shorter term, as social messaging encourages a more flexible orthographic license, might that chip away at standard spelling? Probably no more than the common practice in several languages of intentionally misspelling words in advertising to heighten notice of and attention to a product.

The precision of thought that a visible language offers, regardless of the orthography in which it is expressed, serves narrative and functional writing alike, affording examination and reflection. The burgeoning lexicons of digital worlds offer interesting possibilities for studying the degree to which students’ attention to words—their spellings, polysemies, connotations, and nuances—may be increasing. Will morphological processes become more transparent and intriguing as new words are coined? Orthography will be a pivotal platform from which such inquiries—by both researchers and students—may be launched.

References

- Adoniou, M., 2014. What should teachers know about spelling? *Literacy* 48 (3), 144–154.
Adoniou, M., 2016. *Spelling it Out*. Cambridge University Press, Cambridge, UK.

- Ahmed, Y., Wagner, R., Lopez, D., 2014. Developmental relations between reading and writing at the word, sentence, and text levels: a latent change score analysis. *J. Educ. Psychol.* 106 (2), 419–434.
- Anglin, J.M., 1993. Vocabulary development: a morphological analysis. *Monogr. Soc. Res. Child Dev.* 58 (10) (Serial No. 238).
- Apel, K., Masterson, J., Hart, P., 2004. Integration of language components in spelling. In: Wilkinson, L., Silliman, E. (Eds.), *Language and Literacy Learning in Schools*. Guilford Press, New York, pp. 292–318.
- Avramiuk, E., 2014. Invented spelling—a window on early literacy. *EDUKACJA* 5 (130), 112–123.
- Bahr, R.H., Silliman, E.R., Berninger, V., 2009. What spelling errors have to tell about vocabulary learning. In: Wood, C., Connelly, V. (Eds.), *Contemporary Perspectives on Reading and Spelling*. Routledge, London, pp. 109–129.
- Bahr, R.H., Silliman, E.R., Berninger, V.W., Dow, M., 2012. Linguistic pattern analysis of misspellings of typically developing writers in grades 1–9. *J. Speech Lang. Hear. Res.* 55 (6), 1587–1599.
- Bahr, R.H., Silliman, E.R., Berninger, V.W., 2020. Derivational morphology bridges phonology and orthography: insights into the development of word-specific spellings by superior, average, and poor spellers. *Lang. Speech Hear. Serv. Sch.* 51 (3), 640–654.
- Bear, D., Templeton, S., 2000. Matching development and instruction. In: Padak, N., Rasinski, T. (Eds.), *Distinguished Educators on Reading: Contributions that Have Shaped Effective Literacy Instruction*. International Reading Association, Newark, DE, pp. 363–376.
- Bear, D.R., Templeton, S., Helman, L., Baren, T., 2003. Orthographic development and learning to read in different languages. In: Garcia, G.C. (Ed.), *English Learners: Reaching the Highest Level of English Literacy*. International Reading Association, Newark, DE, pp. 71–95.
- Bear, D.R., Invernizzi, M., Templeton, S., Johnston, F., 2004. *Words Their Way: Word Study for Phonics, Vocabulary, and Spelling Instruction*, third ed. Pearson, Upper Saddle River, NJ.
- Bear, D.R., Invernizzi, M., Templeton, S., Johnston, F., 2020. *Words Their Way: Word Study for Phonics, Vocabulary, and Spelling Instruction*, seventh ed. Pearson, New York, NY.
- Beers, J.W., Henderson, E.H., 1977. A study of developing orthographic concepts among first graders. *Res. Teach. Engl.* 11, 133–148.
- Berg, K., 2016. Double consonants in English: graphemic, morphological, prosodic and etymological determinants. *Read. Writ.* 29 (3), 453–474.
- Berninger, V.W. (Ed.), 1994. *The Varieties of Orthographic Knowledge I: Theoretical and Developmental Issues*. Kluwer, Dordrecht, The Netherlands.
- Berninger, V.W. (Ed.), 1995. *The Varieties of Orthographic Knowledge II: Relationships to Phonology, Reading, and Writing*. Kluwer, Dordrecht, The Netherlands.
- Botel, M., 1964. What linguistics says to this teacher of reading and spelling. *Read. Teach.* 18 (3), 188–193.
- Bourassa, D.C., Treiman, R., 2008. Morphological constancy in spelling: a comparison of children with dyslexia and typically developing children. *Dyslexia* 14, 155–169.
- Bowers, J.S., Bowers, P., 2017. Beyond phonics: the case for teaching children the logic of the English spelling system. *Educ. Psychol.* 52 (2), 124–141.
- Bowers, P.N., Kirby, J.R., 2010. Effects of morphological instruction on vocabulary acquisition. *Read. Writ.* 23 (5), 515–537.
- Bowling, E.C.C., Cabell, S.Q., 2022. Developing readers: understanding concept of word in text development in emergent readers. *Early Child. Educ. J.* 47 (2), 143–151.
- Brown, G.D.A., Ellis, N.C. (Eds.), 1994. *Handbook of Spelling: Theory, Process and Intervention*. John Wiley & Sons, Chichester, GB.
- Brown, T.T., Kuperman, J.M., Chung, Y., et al., 2012. Neuroanatomical assessment of biological maturity. *Curr. Biol.* 22 (18), 1693–1698.
- Brown, R., 1973. *A First Language: The Early Stages*. George Allen & Unwin, London.
- Brown, T.T., 2017. Individual differences in human brain development. *WIREs Cognit. Sci.* 8 (1–2), 1–8.
- Cartwright, K.B., 2012. Insights from cognitive neuroscience: the importance of executive function for early reading development and education. *Early Educ. Dev.* 23 (1), 24–36.
- Chace, K.H., Rayner, K., Well, A.D., 2005. Eye movements and phonological parafoveal preview: effects of reading skill. *Canadian J. Exp. Psychol.* 59 (3), 209–217.
- Chomsky, C., 1970. Reading, writing, and phonology. *Harv. Educ. Rev.* 40 (2), 287–309.
- Clifton, C., Staub, A., Rayner, K., 2007. Eye movements in reading words and sentences. In: van Gompel, R.G., Fischer, M.H., Murray, W.S., Hill, R.L. (Eds.), *Eye Movements: A Window on Mind and Brain*. Elsevier, Amsterdam, Netherlands, pp. 341–371.
- Copp, S.B., Cabell, S.Q., Invernizzi, M.A., 2019. Kindergarten teachers' use of writing scaffolds to support children's developing orthographic knowledge. *Literacy Res. Instr.* 58 (3), 164–183.
- Critten, S., Pine, K.J., 2009. Viewing spelling in a cognitive context: underlying representations and processes. In: Wood, C., Connelly, V. (Eds.), *Contemporary Perspectives on Reading and Spelling*. Routledge, London, pp. 92–108.
- Daffern, T., Critten, S., 2019. Student and teacher perspectives on spelling. *Aust. J. Lang. Literacy* 42 (1), 40–57.
- Daffern, T., Mackenzie, N.M., Hemmings, B., 2015. The development of a spelling assessment tool informed by Triple Word Form theory. *Aust. J. Lang. Literacy* 38 (2), 72–82.
- Daffern, T., 2017. Linguistic skills involved in learning to spell: an Australian study. *Lang. Educ.* 31 (4), 307–329.
- Dai, J., Ibrahim, R., Share, D.L., 2013. The influence of orthographic structure on printed word learning in Arabic. *Writ. Syst. Res.* 5, 189–213.
- Daniels, P.T., Share, D.L., 2018. Writing system variation and its consequences for reading and dyslexia. *Sci. Stud. Read.* 22 (1), 101–116.
- Deacon, S.H., Bryant, P., 2005. The strength of children's knowledge of the role of root morphemes in the spelling of derived words. *J. Child Lang.* 32, 375–389.
- Deacon, S.H., Dhooge, S., 2010. Developmental stability and changes in the impact of root consistency on children's spelling. *Read. Writ.* 23, 1055–1069.
- Defior, S., Alegria, J., Titos, R., Martos, F., 2008. Using morphology when spelling in a shallow orthographic system: the case of Spanish. *Cognit. Dev.* 23 (1), 204–215.
- Devonshire, V., Fluck, M., 2010. Spelling development: fine-tuning strategy-use and capitalising on the connections between words. *Learn. InStruct.* 20, 361–371.
- Devonshire, V., Morris, P., Fluck, M., 2013. Spelling and reading development: the effect of teaching children multiple levels of representation in their orthography. *Learn. InStruct.* 25, 85–94.
- Diamanti, V., Gouladrin, N., Campbell, R., Protopapas, A., 2018. Dyslexia profiles across orthographies differing in transparency: an evaluation of theoretical predictions contrasting English and Greek. *Sci. Stud. Read.* 22 (1), 55–69.
- Ehri, L.C., McCormick, S., 1998. Phases of word learning: implications for instruction with delayed and disabled readers. *Read. Writ. Q.* 14 (2), 135–163.
- Ehri, L.C., Wilce, L.S., 1985. Movement into reading: is the first stage of printed word learning visual or phonetic? *Read. Res. Q.* 20, 163–179.
- Ehri, L.C., Wilce, L., 1986. The influence of spellings on speech: are alveolar flaps /d/ or /t/? In: Yaden Jr., D.B., Templeton, S. (Eds.), *Metalinguistic Awareness and Beginning Literacy: Conceptualizing What It Means to Learn to Read and Write*. Heinemann Educational Books, Portsmouth, NH, pp. 101–114.
- Ehri, L.C., Wilce, L., 1987. Does learning to spell help beginners learn to read words? *Read. Res. Q.* 22, 47–65.
- Ehri, L.C., 1980. The influence of orthography on readers' conceptualization of the phonemic structure of words. *Appl. Psycholinguist.* 1, 371–385.
- Ehri, L.C., 1995. Phases of development in learning to read words by sight. *J. Res. Read.* 18 (1), 116–125.
- Ehri, L.C., 2000. Learning to read and learning to spell: two sides of a coin. *Top. Lang. Disord.* 20 (3), 19–36.
- Ehri, L.C., 2005. Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9 (2), 167–188.
- Ehri, L.C., 2011. Review and commentary: stages of spelling development. In: Templeton, S., Bear, D.R. (Eds.), *Development of Orthographic Knowledge and the Foundations of Literacy: A Memorial Festschrift for Edmund H. Henderson*. Routledge/Taylor & Francis, New York, pp. 307–332 (Original work published 1992; Hillsdale, NJ: Lawrence Erlbaum Associates).
- Ehri, L.C., 2014. Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18 (1), 5–21.
- Fejzo, A., 2016. The contribution of morphological awareness to the spelling of morphemes and morphologically complex words in French. *Read. Writ.* 29 (2), 207–228.
- Ferreiro, E., Teberosky, A., 1982. *Literacy Before Schooling*. Heinemann, Exeter, NH.
- Flanigan, K., 2007. A concept of word in text: a pivotal event in early reading acquisition. *J. Literacy Res.* 39 (1), 37–70.
- Ford, K., Invernizzi, M., Huang, F., 2018. The effect of orthographic complexity on Spanish spelling in grades 1–3. *Read. Writ.* 31, 1063–1081. <https://doi.org/10.1007/s11145-018-9828-7>.
- Francis, W.N., 1958. *The Structure of American English*. Ronald Press, New York.

- Freeman, N., Townsend, D., Templeton, S., 2019. Thinking about words: first graders' response to morphological instruction with an emphasis on Greek and Latin roots. *Read. Teach.* 72 (4), 463–473.
- Frith, U. (Ed.), 1980. *Cognitive processes in spelling*. Academic Press, London.
- Frith, U., 1985. Beneath the surface of developmental dyslexia. In: Patterson, K., Marshall, J., Coltheart, M. (Eds.), *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading*. Lawrence Erlbaum Associates, London, pp. 301–330.
- Ganske, K., 1999. The developmental spelling analysis: a measure of orthographic knowledge. *Educ. Assess.* 6, 41–70.
- Gathercole, S.E., Baddeley, A.D., 2014. *Working Memory and Language*. Psychology Press.
- Gehsmann, K., Templeton, S., 2022. *Teaching Reading and Writing: The Developmental Approach (preK-8)*, second ed. Pearson, New York.
- Gehsmann, K., Tousley, M.E., Spichtig, A.N., Pascoe, J.P., Ferrara, J.D., 2017. Relationship between students' stages of orthographic knowledge and reading efficiency [Abstract]. *J. Eye Movement Res.* 10 (6), 209. <https://doi.org/10.16910/jemr.10.6>.
- Gehsmann, K., Spichtig, A., Pascoe, J., Ferrara, J., Tousley, E., 2019. The confluence of orthographic knowledge and reading efficiency in proficient and non-proficient fourth- and fifth-grade readers. Presented at Society for the Scientific Study of Reading, Toronto, ON.
- Goswami, U., 2001. Early phonological development. In: Neuman, S.B., Dickinson, D.K. (Eds.), *Handbook of Early Literacy Research*. The Guilford Press, New York, pp. 111–125.
- Helman, L.A., Bear, D.R., Templeton, S., Invernizzi, M., Johnston, F., 2012. *Words Their Way With English Learners*, second ed. Pearson/Allyn & Bacon, Boston.
- Helman, L.A., Delbridge, A., Parker, D., Arnal, M., Jara Mödinger, L., 2015. Measuring Spanish orthographic development in private, public, and subsidised schools in Chile. In: *Assessment in Education: Principles, Policy & Practice*. <https://doi.org/10.1080/0969594X.2015.1038217>.
- Helman, L. (Ed.), 2016. *Literacy Development With English Learners: Research-Based Instruction in Grades K-6*, second ed. Guilford Press, New York.
- Henderson, E.H., Beers, J. (Eds.), 1980. *Developmental and Cognitive Aspects of Learning to Spell: A Reflection of Word Knowledge*. International Reading Association, Newark, DE.
- Henderson, L., Chard, J., 1980. The reader's implicit knowledge of orthographic structure. In: Frith, U. (Ed.), *Cognitive Processes in Spelling*. Academic Press, London, pp. 85–116.
- Henderson, E.H., Templeton, S., 1986. A developmental perspective of formal spelling instruction through alphabet, pattern, and meaning. *Elem. Sch. J.* 86, 305–316.
- Henderson, E.H., Estes, T.H., Stonecash, S., 1972. An exploratory study of word acquisition among first-graders at mid-year in a Language Experience Approach. *J. Read. Behav.* 4, 21–30.
- Henderson, E.H., 1980. Developmental concepts of word. In: Henderson, E.H., Beers, J. (Eds.), *Developmental and Cognitive Aspects of Learning to Spell: A Reflection of Word Knowledge*. International Reading Association, Newark, DE, pp. 1–14.
- Henderson, E.H., 1981. *Learning to Read and Spell: The Child's Knowledge of Words*. Northern Illinois Press, DeKalb, IL.
- Henderson, E.H., 1985. *Teaching Spelling*. Houghton Mifflin, Boston.
- Henderson, E.H., 1990. *Teaching Spelling*, second ed. Houghton Mifflin, Boston.
- Hilte, M., Reitsma, P., 2011. Effects of explicit rules in learning to spell open- and closed-syllable words. *Learn. Instruct.* 21, 34–45.
- International Literacy Association, 2019. *Teaching and Assessing Spelling [Literacy Leadership Brief]*. International Literacy Association, Newark, DE.
- Invernizzi, M., Buckrop, J., 2018. Reconceptualizing alphabet learning and instruction. In: Cassano, C.M., Dougherty, S.M. (Eds.), *Pivotal Research in Early Literacy: Foundational Studies and Current Practices*. The Guilford Press, New York, pp. 85–110.
- Invernizzi, M., Hayes, L., 2011. Developmental patterns of reading proficiency and reading difficulties. In: McGill-Franzen, A., Allington, R.L. (Eds.), *Handbook of Reading Disability Research*. Routledge/Taylor & Francis, New York, pp. 196–207.
- Invernizzi, M., Abouzeid, M., Gill, J.T., 1994. Using students' invented spellings as a guide for spelling instruction that emphasizes word study. *Elem. Sch. J.* 95, 155–167.
- Invernizzi, M., 2011. The vowel and what follows: a phonological frame of orthographic analysis. In: Templeton, S., Bear, D.R. (Eds.), *Development of Orthographic Knowledge and the Foundations of Literacy: A Memorial Festschrift for Edmund H. Henderson*. Routledge/Taylor & Francis, New York, pp. 105–136 (Original work published 1992; Hillsdale, NJ: Lawrence Erlbaum Associates).
- Johnson, M.H., 2011. Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cognit. Neurosci.* 1, 7–21.
- Joshi, R.M., Aaron, P.G. (Eds.), 2006. *Handbook of Orthography and Literacy*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Katz, L., Feldman, L.B., 1983. Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *J. Exp. Psychol. Learn. Mem. Cognit.* 9, 157–164.
- Kessler, B., Treiman, R., 2001. Relationships between sounds and letters in English monosyllables. *J. Mem. Lang.* 44, 592–617.
- Kessler, B., Pollo, T.C., Treiman, R., Cardoso-Martins, C., 2013. Frequency analyses of prephonological spellings as predictors of success in conventional spelling. *J. Learn. Disabil.* 46 (3), 252–259.
- Kiernan, D., Bear, D.R., 2018. What spelling tells us about the orthographic development and word study instruction with emergent bilingual secondary students. *Read. Psychol.* 39 (3), 227–252.
- Kwak, J., 2007. *Writing development of Korean heritage children [University of Southern California UMI]*. In: *Dissertation Abstracts International, Section A: Humanit. Soc. Sci.*, vol. 67, p. 3707, 10.
- Kwong, T.E., Varnhagen, C.K., 2005. Strategy development and learning to spell new words: generalization of a process. *Dev. Psychol.* 41 (1), 148–159.
- Larkin, R.F., Snowling, M.J., 2008. Morphological spelling development. *Read. Writ. Q.* 24, 363–376.
- Li, X., Bicknell, K., Liu, P., Wei, W., Rayner, K., 2014. Reading is fundamentally similar across disparate writing systems: a systematic characterization of how words and characters influence eye movements in Chinese reading. *J. Exp. Psychol. Gen.* 143 (2), 895–913.
- Liberman, I.Y., Liberman, A.M., Mattingly, I.G., Shankweiler, D.L., 1980. Orthography and the beginning reader. In: Kavanagh, J.F., Venezky, R.L. (Eds.), *Orthography, Reading, and Dyslexia*. University Park Press, Baltimore, pp. 137–153.
- Liebig, J., Froehlich, E., Morawetz, C., et al., 2017. Neurofunctionally dissecting the reading system in children. *Dev. Cognit. Neurosci.* 27, 45–57.
- Longcamp, M., Velay, J.L., Berninger, V.W., Richards, T., 2016. Neuroanatomy of handwriting and related reading and writing skills in adults and children with and without learning disabilities: French-American connections. *Pratiques* 171–172.
- Mann, V., Tobin, P., Wilson, R., 1987. Measuring phonological awareness through the invented spellings of kindergarten children. *Merrill-Palmer Q.* 33, 365–391.
- Mann, V., 1993. Phoneme awareness and future reading ability. *J. Learn. Disabil.* 26, 259–269.
- Marsh, G., Friedman, M., Welch, V., Desberg, P., 1980. The development of strategies in spelling. In: Frith, U. (Ed.), *Cognitive Strategies in Spelling*. Academic Press, New York, pp. 339–353.
- Martins, M.A., Albuquerque, A., Salvador, L., Silva, C., 2013. The impact of invented spelling on early spelling and reading. *J. Writ. Res.* 5 (2), 215–237.
- Masterson, J.J., Apel, K., 2010. Linking characteristics discovered in spelling assessment to intervention goals and methods. *Learn. Disabil. Q.* 33, 185–198.
- McBride, C.A., 2016. Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educ. Psychol. Rev.* 28, 523–549.
- Mirzaei, A., 2003. *Development of orthographic knowledge in a consonantal script: children's invented spellings in Farsi [University of Southern California UMI]*. In: *Dissertation Abstracts International, Section A: Humanit. Soc. Sci.*, vol. 63, pp. 4226–4227, 12.
- Morris, D., Nelson, L., Perney, J., 1986. Exploring the concept of "spelling instructional level" through the analysis of error-types. *Elem. Sch. J.* 87, 181–200.
- Morris, D., Blanton, L., Blanton, W.E., Nowacek, J., Perney, J., 1995a. Teaching low-achieving spellers at their "instructional level". *Elem. Sch. J.* 96, 163–178.
- Morris, D., Blanton, L., Blanton, W., Perney, J., 1995b. Spelling instruction and achievement in six classrooms. *Elem. Sch. J.* 96, 145–162.
- Morris, D., Bloodgood, J.W., Lomax, R.G., Perney, J., 2003a. Developmental steps in learning to read: a longitudinal study in kindergarten and first grade. *Read. Res. Q.* 38, 302–328.
- Morris, D., Bloodgood, J., Perney, J., 2003b. Kindergarten predictors of first- and second-grade reading achievement. *Elem. Sch. J.* 104 (2), 93–109.

- Morris, D., Trathen, W., Perney, J., Gill, T., Schlagal, R., Ward, D., Frye, E.M., 2017. Three DIBELS tasks vs. three informal reading/spelling tasks: a comparison of predictive validity. *Read. Psychol.* 38 (3), 289–320.
- Morris, D., 2011. Interventions to develop phonological and orthographic systems. In: McGill-Franzen, A., Allington, R.L. (Eds.), *Handbook of Reading Disability Research*. Routledge, New York, pp. 279–288.
- Morris, D., 2021. *A Studebaker's Memoir: Preventing Reading Failure in Grades 1-3*. Fieldstream Publications, Boone, NC.
- Moskowitz, B.A., 1973. On the status of vowel shift in English. In: Moore, T. (Ed.), *Cognitive Development and the Acquisition of Language*. Academic Press, New York, pp. 223–260.
- Nagy, W., Berninger, V.W., Abbott, R.D., 2006. Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *J. Educ. Psychol.* 98 (1), 134–147.
- Nunes, T., Bryant, P., 2009. *Children's Reading and Spelling: Beyond the First Steps*. Wiley-Blackwell, London: Chichester, West Sussex, UK.
- Olson, D., 1996. *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading*. Cambridge University Press, New York.
- Ouellette, G., Sénéchal, M., 2017. Invented spelling in kindergarten as a predictor of reading and spelling in grade 1: a new pathway to literacy, or just the same road, less known? *Dev. Psychol.* 53 (1), 77–88.
- Ouellette, G., Sénéchal, M., Haley, A., 2013. Guiding children's invented spellings: a gateway into literacy learning. *J. Exp. Educ.* 81, 261–279. <https://doi.org/10.1080/00220973.2012.699903>.
- Perfetti, C.A., Rieben, L., Fayol, M. (Eds.), 1997. *Learning to Spell: Research, Theory, and Practice Across Languages*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Perfetti, C.A., 1986. George Graham Memorial Lecture. University of Virginia, May.
- Perfetti, C.A., 1993. The representation problem in reading acquisition. In: Gough, P.B., Ehri, L.C., Treiman, R. (Eds.), *Reading Acquisition*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 145–174.
- Perfetti, C.A., 1997. The psycholinguistics of spelling and reading. In: Perfetti, C.A., Rieben, L., Fayol, M. (Eds.), *Learning to Spell: Research, Theory, and Practice Across Languages*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 21–38.
- Perfetti, C.A., 2007. Reading ability: lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383.
- Perrachione, T.K., Del Tufo, S.N., Winter, R., et al., 2016. Dysfunction of rapid neural adaptation in dyslexia. *Neuron* 92, 1383–1397.
- Piaget, J., Inhelder, B., 1969. *The Psychology of the Child*. Basic Books, New York.
- Piaget, J., 1973. *The Child and Reality: Problems of Genetic Psychology*. Viking Press, New York.
- Planton, S., Jucla, M., Roux, F.E., Démonet, J.F., 2013. The "handwriting brain": a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex* 49 (10), 2772–2787.
- Pollo, T.C., Kessler, B., Treiman, R., 2009. Statistical patterns in children's early writing. *J. Exp. Child Psychol.* 104 (4), 410–426.
- Puranik, C.S., Lonigan, C.J., Kim, Y.S., 2011. Contributions of emergent literacy skills to name writing, letter writing, and spelling in preschool children. *Early Child. Res. Q.* 26 (4), 465–474.
- Ramsden, M., 1993. *Rescuing Spelling*. Southgate Publishers, Sandford, UK.
- Rastle, K., Davis, M.H., 2008. Morphological decomposition based on the analysis of orthography. *Lang. Cognit. Process.* 23 (7/8), 942–971.
- Rastle, K., 2019. The place of morphology in learning to read in English. *Cortex* 116, 45–54. <https://doi-org.unr.idm.oclc.org/10.1016/j.cortex.2018.02.008>.
- Ravid, D.D., 2012. *Spelling Morphology: The Psycholinguistics of Hebrew Spelling (Literacy Studies 3)*. Springer, New York, NY.
- Rayner, K., Pollatsek, A., Ashby, J., Clifton, C., 2011. *The Psychology of Reading*, second ed. Psychology Press, New York.
- Read, C., Treiman, R., 2013. Children's invented spelling: what we have learned in forty years. In: Piattelli-Palmarini, M., Berwick, R. (Eds.), *Rich Languages From Poor Inputs*. Oxford University Press, Oxford, UK, pp. 195–209.
- Read, C., 1971. Preschool children's knowledge of English phonology. *Harv. Educ. Rev.* 41, 1–34.
- Read, C., 1975. Children's Categorizations of Speech Sounds in English (Research Report No. 17). National Council of Teachers of English, Urbana, IL.
- Richards, T.L., Berninger, V.W., Fayol, M., 2009. fMRI activation differences between 11-year-old good and poor spellers' access in working memory to temporary and long-term orthographic representations. *J. Neurolinguistics* 22 (4), 327–353.
- Robertson, E.K., Deacon, S.H., 2019. Morphological awareness and word-level reading in early and middle elementary school years. *Appl. Psycholinguist.* 40 (4), 1051–1071. <https://doi-org.unr.idm.oclc.org/10.1017/S0142716419000134>.
- Schlagal, B., 2001. Traditional, developmental, and structured approaches to spelling: review and recommendations. *Ann. Dyslexia* 51, 147–176.
- Schlagal, R., 2011. Patterns of orthographic development into the intermediate grades. In: Templeton, S., Bear, D.R. (Eds.), *Development of Orthographic Knowledge and the Foundations of Literacy: A Memorial Festschrift for Edmund H. Henderson*. Routledge/Taylor & Francis, New York, pp. 31–52 (Original work published 1992; Hillsdale, NJ: Lawrence Erlbaum Associates).
- Schlaggar, B.L., McCandless, B.D., 2007. Development of neural systems for reading. *Annu. Rev. Neurosci.* 30, 475–503.
- Sénéchal, M., 2000. Morphological effects in children's spelling of French words. *Canadian J. Exp. Psychol.* 54, 76–85. <https://doi.org/10.1037/h0087331>.
- Shany, M., Share, D.L., 2011. Subtypes of reading disability in a shallow orthography: a double dissociation between accuracy-disabled and rate-disabled readers of Hebrew. *Ann. Dyslexia* 61, 64–84.
- Share, D.L., 2008. On the anglocentricities of current reading research and practice: the perils of overreliance on an "outlier" orthography. *Psychol. Bull.* 134 (4), 584–615.
- Shen, H.H., Bear, D.R., 2000. Development of orthographic skills in Chinese children. *Read. Writ.* 13, 197–236.
- Silliman, E.R., Bahr, R.H., Peters, M.L., 2006. Spelling patterns in preadolescents with atypical language skills: phonological, morphological, and orthographic factors. *Dev. Neuropsychol.* 29, 93–123.
- Singson, M., Mahony, D., Mann, V., 2000. The relation between reading ability and morphological skills: evidence from derivational suffixes. *Read. Writ.* 12, 219–252.
- Stauffer, R., 1969. *Directing Reading Maturity as a Cognitive Process*. Harper & Row, New York.
- Stauffer, R., 1970. *The Language-Experience Approach to the Teaching of Reading*. Harper & Row, New York.
- Sussman, D., Leung, R.C., Chakravarty, M.M., Lerch, J.P., Taylor, M.J., 2016. The developing human brain: age-related changes in cortical, subcortical, and cerebellar anatomy. *Brain Behav.* 6 (4).
- Tan, L.H., Spinks, J.A., Eden, G.F., Perfetti, C.A., Siok, W.T., 2005. Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. U. S. A.* 102 (24), 8781–8785.
- Templeton, S., Bear, D.R. (Eds.), 2011. *Development of Orthographic Knowledge and the Foundations of Literacy: A Memorial Festschrift for Edmund Henderson*. Routledge/Taylor & Francis, New York (Original work published 1992; Hillsdale, NJ: Lawrence Erlbaum Associates).
- Templeton, S., Bear, D.R., 2018. Word study, research to practice: spelling, phonics, meaning. In: Lapp, D., Fisher, D. (Eds.), *Handbook of Research on Teaching the English Language Arts*, fourth ed. Routledge/Taylor & Francis, New York, pp. 207–232.
- Templeton, S., Morris, D., 1999. Questions teachers ask about spelling. *Read. Res. Q.* 34, 102–112.
- Templeton, S., Morris, D., 2000. Spelling. In: Kamil, M., Mosenthal, P., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research*, vol. 3. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 525–543.
- Templeton, S., Scarborough-Franks, L., 1985. The spelling's the thing: older students' knowledge of derivational morphology in phonology and orthography. *Appl. Psycholinguist.* 6, 371–389.
- Templeton, S., 1976. *The Spelling of Young Children in Relation to the Logic of Alphabetic Orthography*. Paper Presented at the 26th Annual Meeting of the National Reading Conference, Atlanta, Georgia, December.
- Templeton, S., 1979a. Spelling first, sound later: the relationship between orthography and higher order phonological knowledge in older students. *Res. Teach. Engl.* 13, 255–264.

- Templeton, S., 1979b. The circle game of English spelling: a reappraisal for teachers. *Lang. Arts* 56, 789–797.
- Templeton, S., 1980. What is a word? In: Henderson, E.H., Beers, J.W. (Eds.), *Developmental and Cognitive Aspects of Learning to Spell: A Reflection of Word Knowledge*. International Reading Association, Newark, DE, pp. 15–35.
- Templeton, S., 1989. Tacit and explicit knowledge of derivational morphology: foundations for a unified approach to spelling and vocabulary development in the intermediate grades and beyond. *Read. Psychol.* 10, 233–253.
- Templeton, S., 1991. Teaching and learning the English spelling system: reconceptualizing method and purpose. *Elem. Sch. J.* 92, 183–199.
- Templeton, S., 2003. Spelling. In: Flood, J., Lapp, D., Squire, J.R., Jensen, J.M. (Eds.), *Handbook of Research on Teaching the English Language Arts*, second ed. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 738–751.
- Templeton, S., 2006. Dispelling spelling assumptions: technology and spelling, present and future. In: McKenna, M., Labbo, L., Kieffer, R., Reinking, D. (Eds.), *Handbook of Literacy and Technology*, second ed. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 335–339.
- Templeton, S., 2020. Stages, phases, repertoires, and waves: learning to spell and read words. *Read. Teach.* 74 (3), 315–323. <https://doi.org/10.1002/trtr.1951>.
- Tong, X., McBride-Chang, C., Shu, H., Wong, A.M., 2009. Morphological awareness, orthographic knowledge, and spelling errors: keys to understanding early Chinese literacy acquisition. *Sci. Stud. Read.* 13 (5), 426–452.
- Townsend, D., Bear, D., Templeton, S., Burton, A., 2016. The implications of adolescents' academic word knowledge for achievement and instruction. *Read. Psychol.* 37 (8), 1119–1148. <https://doi.org/10.1080/02702711.2016.1193579>.
- Treiman, R., Cassar, 1997. Spelling acquisition in English. In: Perfetti, C.A., Rieben, L., Fayol, M. (Eds.), *Learning to Spell: Research, Theory, and Practice Across Languages*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 61–80.
- Treiman, R., Kessler, B., 2006. Spelling as statistical learning: using consonantal context to spell vowels. *J. Educ. Psychol.* 98, 642–652. <https://doi.org/10.1037/0022-0663.98.3.642>.
- Treiman, R., Kessler, B., 2014. *How Children Learn to Write Words*. Oxford University Press, New York, NY.
- Treiman, R., Kessler, B., 2021. Statistical learning in word reading and spelling across languages and writing systems. *Sci. Stud. Read.* 1–11. <https://doi.org/10.1080/10888438.2021.1920951>.
- Treiman, R., Wolter, S., 2018. Phonological and graphotactic influences on spellers' decisions about consonant doubling. *Mem. Cognit.* 46, 614–624.
- Treiman, R., Wolter, S., 2020. Use of letter names benefits young children's spelling. *Psychol. Sci.* 31 (1), 43–50.
- Treiman, R., Kessler, B., Pollo, T.C., Byrne, B., Olson, R.K., 2016. Measures of kindergarten spelling and their relations to later spelling performance. *Sci. Stud. Read.* 20, 349–362.
- Treiman, R., Kessler, B., Boland, K., Clocksin, H., Chen, Z., 2018. Statistical learning and spelling: older prephonological spellers produce more wordlike spellings than younger prephonological spellers. *Child Dev.* 89 (4), 431–443.
- Treiman, R., 1993. *Beginning to Spell*. Oxford University Press, New York.
- Treiman, R., 2017. Learning to spell words: findings, theories, and issues. *Sci. Stud. Read.* 21 (4), 265–276.
- Treiman, R., 2018. Statistical learning and spelling. *Lang. Speech Hear. Serv. Sch.* 49 (3S), 644–652.
- Tyler, A., Nagy, W., 1989. The acquisition of English derivational morphology. *J. Mem. Lang.* 28, 649–667.
- Venezky, R.L., 1970. *The Structure of English Orthography*. Mouton, The Hague.
- Venezky, R.L., 1999. *The American Way of Spelling: The Structure and Origins of American English Orthography*. Guilford Press, New York.
- Viise, N.M., Richards, H.C., Pandis, M., 2011. Orthographic depth and spelling acquisition in Estonian and English: a comparison of two diverse alphabetic languages. *Scand. J. Educ. Res.* 55 (4), 425–453.
- Winkel, H., Iemwanthong, K., 2010. Reading and spelling acquisition in Thai children. *Read. Writ.* 23, 1021–1053.
- Wolter, J.A., Dilworth, V., 2014. The effects of a multilingual morphological awareness approach for improving language and literacy. *J. Learn. Disabil.* 47 (1), 76–85.
- Worthy, M.J., Invernizzi, M., 1990. Spelling errors of normal and disabled students on achievement levels one through four: instructional implications. *Ann. Dyslexia* 40 (1), 138–151.
- Yaden Jr., D.B., Tsai, T., 2011. Learning how to write in English and Chinese: young kindergarten and first grade children explore the similarities and differences in writing systems. In: Gort, M., Bauer, E. (Eds.), *Early Biliteracy Development: Exploring Young Learners' Use of Their Linguistic Resources*. Routledge, New York, NY, pp. 55–84.
- Yaden Jr., D.B., Rueda, R., Tardibuo, J., Martinez-Yaden, C., Mirzaei, A., Scott-Weich, B., Tsai, T., 2015. Interviewing young children using psycho-and micro-genetic design methodology to assess understandings of reading and writing: the promise and challenge. In: Saracho, O.N., Spodek, B. (Eds.), *Contemporary Perspectives in Early Childhood Education, Handbook of Research Methods in Early Childhood Education*, 2. Information Age Publishing, pp. 309–348.
- Zhang, L., Yin, L., Treiman, R., 2017. Chinese children's early knowledge about writing. *Br. J. Dev. Psychol.* 35 (3), 349–358.
- Zutell, J., 1979. Spelling strategies of primary school children and their relationship to Piaget's concept of decentration. *Res. Teach. Engl.* 13 (1), 69–80.
- Zutell, J., 1980. Children's spelling strategies and their cognitive development. In: Henderson, E.H., Beers, J.W. (Eds.), *Developmental and Cognitive Aspects of Learning to Spell*. International Reading Association, Newark, DE, pp. 52–73.

The art and science of reading fluency and fluency instruction

Timothy Rasinski^a, Chase Young^b, David D. Paige^c, and Kasim Yildirim^d, ^a School of Teaching Learning and Curriculum Studies, Kent State University, Kent, OH, United States; ^b Sam Houston State University, Huntsville, TX, United States; ^c Northern Illinois University, Department of Curriculum & Instruction, DeKalb, IL, United States; and ^d Muğla Sıtkı Koçman University, Muğla, Turkey

© 2023 Elsevier Ltd. All rights reserved.

Defining fluency	388
Automaticity—The link to word recognition	389
Prosody—The link to comprehension	389
A brief history of fluency	390
Fluency as science	390
Automatic processes	390
Prosodic processes	391
Fluency to comprehension	391
Fluency as art	392
Fluency—Beyond English	392
Assessing reading fluency	393
Conclusion	394
References	394

Reading fluency, as a conceptual component of proficient reading and an instructional concept, has had a history of neglect and misunderstanding. It has been largely ignored in the so called Reading Wars (Castles et al., 2018; Goldenberg, 2020; Wyse and Bradbury, 2022) and is missing from the Simple View of Reading (Hoover and Gough, 1990). Whole Language perspectives of reading also seem to miss it as well. And yet, research has consistently demonstrated that reading fluency is critical to success in learning to read, that it is applicable at all grade levels and ages, and that many students who are not proficient in reading manifest notable difficulties in reading fluency. In this paper, we hope to explain the nature of fluency, its importance for success in reading, and how it can be taught and assessed in ways that are authentic and engaging for students and teachers.

After years of neglect, fluency returned to the mainstream professional conversation on reading with a paper by Richard Allington (1983) published in *The Reading Teacher* entitled Fluency, the Neglected Reading Goal. In the piece, Allington argued that reading fluency or, more importantly, a lack of fluency in reading is easy to detect. One needs simply to listen to a less-than-fluent reader read orally for a few minutes and recognize the student's difficulty with fluency. Even if all the words are read correctly and understood by the reader, the word-by-word, monotone reading provides ample evidence that the reader is experiencing difficulty. Yet, Allington argued, although it is easily diagnosed, it is seldom treated in actual instruction. His paper put out a call for fluency to be made a more significant part of reading, that should have equal status with word recognition (phonics) and comprehension.

Since that time, reading fluency has gained some degree of recognition among researchers and practitioners. Still, at the same time, there continues to be a diminution of fluency with respect to other reading competencies. For example, in her series of media articles, educational journalist Emily Hanford (2018, 2019) has made the case that the reason for the lack of improvement in literacy outcomes for students in the United States has been because of the way that phonics or word recognition is taught. She argued that a lack of systematic phonics as well as an overemphasis on cueing-based approaches to word recognition is the primary reason why large numbers of students fail to achieve proficiency in reading. Suggesting that a single aspect of reading instruction is at fault for generations of struggling readers is shocking. Such a claim also indirectly places less importance on other components of reading, including fluency, which continues to struggle for recognition as an essential component for full proficiency in reading.

Defining fluency

The simple view of reading (Gough and Tunmer, 1986; Hoover and Gough, 1990) suggests that the two major competencies involved in proficient reading are word decoding and language/listening comprehension. While acknowledging these important and distinct components, fluency, in a sense, links the two so that reading can become an integrated whole. Pikulski and Chard (2005) described fluency as a bridge from word recognition accuracy to text comprehension. Reading fluency is broadly defined as the ability to read the words in texts effortlessly and with expression, essentially sounding like authentic oral language.

More specifically, reading fluency is defined by its two major components: automaticity and prosody. Automaticity is the link to word decoding as it refers to the ability to recognize words automatically or effortlessly (LaBerge and Samuels, 1974). Prosody, or expression in oral reading, is the ability to read with a level of expression that reflects the meaning of text read (Schreiber, 1980).

Automaticity—The link to word recognition

It is not enough for readers to read the words they encounter in reading text accurately—they need to read the words automatically. In a landmark paper, LaBerge and Samuels (1974) posited that all readers have a limited amount of attention or cognitive resources. If readers have to use too much of their resources to decode the words in a text, less is available for the more important task in reading—comprehension. Such readers are marked by their slow, laborious, word-by-word reading of texts.

The goal should be for readers to read the words in texts not just accurately, but also *automatically*. When the words in the text are identified automatically, readers use a minimal amount of attention so as to deploy more of their cognitive resources to textual comprehension. For many readers, comprehension suffers not because of insufficient cognitive resources, but because those resources were depleted in the employment of word recognition. These are the same readers who are able to understand texts read to them—when someone else takes on the task of decoding the words, their cognitive resources can be deployed to making meaning.

As in developing automaticity in any endeavor, automaticity in reading is accomplished through practice. There are three types of practice that are worth noting here; wide reading practice, repeated (close or deep) reading practice, and assisted reading practice (Kodan and Akyol, 2018; Petersen-Brown et al., 2021; Reed et al., 2019; Therrien, 2004; Wexler et al., 2010).

Prosody—The link to comprehension

As automaticity is part of the conceptual fluency bridge that links to word recognition, prosody completes the bridge by linking fluency to comprehension. Prosody refers to a reader's ability to read aloud with an appropriate level of expression that reflects and even expands the meaning of the text (Young et al., 2021). This component is important for several reasons. First, prosodic renderings of text are more aesthetically pleasing. Hearing a story, poem, or speech read aloud expressively is arguably far more engaging than listening to a choppy and monotone read aloud. This is not only true for oral reading, but also for silent reading. There is evidence that those who read aloud with good expression also read silently with similar expression (Gross et al., 2014). Someone who is a fluent reader or speaker is likely to be described as a person who uses her or his voice to convey meaning to a listener when speaking or reading orally. Consider the following sentence:

Thomas borrowed my new car.

This declarative sentence describes an action performed by Thomas. However, emphasis on a single word can add implied meaning to the sentence.

Thomas borrowed my new car. (Thomas, not Raymond, borrowed my car.)

Thomas *borrowed* my new car. (Thomas did not steal my car, he only borrowed it.)

Thomas borrowed *my* new car. (Thomas didn't borrow your car, he borrowed mine.)

Thomas borrowed my *new* car. (Thomas didn't borrow my old car, he borrowed the new one.)

Thomas borrowed my new *car*. (Thomas didn't borrow my new book, he borrowed my car.)

Prosody allows the reader to comprehend a text at a more nuanced level than the text itself allows. Scholars have argued that prosody in reading also assists the reader in identifying critical and meaningful phrase boundaries that are not marked by punctuation (Schreiber, 1980, 1987, 1991; Schreiber and Read, 1980; Whalley and Hansen, 2006). Again, prosody allows the reader to infer information that is not explicitly stated in the passage.

Like automaticity, prosody is highly correlated with reading comprehension (Benjamin and Schwanenflugel, 2010; Miller and Schwanenflugel, 2006; Sabatini et al., 2019; Whalley and Hansen, 2006). In a longitudinal study, Miller and Schwanenflugel (2008) found that first graders who read with more adult-like prosody were more likely to be proficient in reading comprehension by third grade. Although these studies cited are relatively recent, Goodman (1964) realized decades ago that more prosodic readers were often better readers in general.

Prosody is not an issue solely for oral reading. Many adult proficient readers indicate that they also hear themselves when reading silently. Indeed, several studies have found that readers at the third, fourth, fifth, and eighth grade levels who read orally with good prosody also tend to be good comprehenders when reading silently (Daane et al., 2005; Pinnell et al., 1995; Rasinski et al., 2009). Conversely, these same studies have also found that readers who read with limited prosody (in a more monotone and word-by-word manner) also manifest poorer comprehension when reading silently.

Prosody is related to proficient and meaningful reading—oral and silent. So how do readers develop their prosody in reading? As noted with automaticity, prosody is developed through wide and deep reading. However, when prosody is emphasized, the goal of wide and deep (repeated) reading is to achieve an expressive oral rendering of the passage that reflects the meaning of the passage.

A brief history of fluency

Early accounts of reading fluency were largely related to elocution and proper pronunciation during oral reading. Oral reading dominated the classroom prior to around 1925, emphasizing that students should read aloud accurately and with good expression through drill, practice, and correction (Hyatt, 1943). However, the work of Huey (1908) and others (e.g. Judd and Buswell, 1922) led to classroom reading instruction where silent reading became the dominant form. Indeed, through the last half of the 20th century, oral reading was largely confined to the ubiquitous and widely criticized round robin reading where teachers indiscriminately and without warning called on individual students to read during group reading instruction. It is unclear when oral reading, beyond the round robin format, worked its way back into the classroom; unlike the silent reading revolution incited by Huey and others, there was not a mass call to have students read aloud. Most likely, however, oral reading came back in the form of reading instruction and interventions. Several seminal studies in the 1960s and 70s required oral reading by students as a means of remediation; namely, the Neurological Impress Method (Heckelman, 1969) and the method of repeated readings (Samuels, 1979). In the 1980s even more reading intervention-type studies documented classroom practices that required oral reading, such as paired reading (Topping, 1989). Although the frequency of oral reading in some contexts had increased, at this point in time it had not necessarily become a part of the general or core curriculum.

As mentioned previously, Allington (1983) had deemed reading fluency a neglected reading goal. Important in this paper was his seemingly obvious recommendation that students who were identified as disfluent should receive targeted fluency instruction and not simply additional instruction in other areas such as phonics. This call for fluency instruction was not only aimed toward special education and remedial reading teachers, but to general classroom teachers. Reading fluency instruction was not simply a remediation tactic, but a good approach to developing a strong foundation for reading development and proficiency (Rasinski, 2010).

More recently, based on its review of scientific studies of reading fluency, the National Reading Panel (2000) identified reading fluency to be a major component of the reading process. The notion of fluent reading became increasingly important, and rather than waiting for a student to need remediation in reading fluency, instruction began to be introduced into the general elementary curriculum to develop all students' reading fluency. Most of the instructional approaches are based on the goals of improving automaticity and prosody and include various iterations of repeated reading and assisted reading such as readers theater (Griffith and Rasinski, 2004; Martinez et al., 1999), poetry recitation, and performance of songs (Iwasaki et al., 2013; Wilfong, 2008), making movies (Young and Rasinski, 2013), and performing speeches (Young and Rasinski, 2022).

Fluency as science

Reading fluency has a strong, theoretical foundation accompanied by a large scientific research base of historic and newly developed instructional practices that impact reading comprehension.

Automatic processes

Instant or automatic recognition of a large corpus of words is critical to fluent reading. A process is considered automatic when it operates quickly, autonomously, and requires no or minimal conscious awareness, meaning it can operate in the presence of other processes requiring full attention (LaBerge and Samuels, 1974; Rawson and Middleton, 2009). In their seminal study, LaBerge and Samuels (1974) proposed that word recognition automaticity is necessary to allow the reader to focus their limited attention on comprehension. The authors hypothesized that repeated encounters with a word results in the strengthening of the connections necessary to pronounce (and understand) the word to the point that recognition becomes automatic. Other theories hypothesized that repeated encounters with a word made it available in long-term memory as a single-step identification process (Logan, 1988; Logan and Klapp, 1991). In either instance, the end result is that it is the development of a large inventory of words that are recognized automatically and independently from context that enables one to become a fluent reader (Stanovich, 1980; Torgesen, 2004; Torgesen and Hudson, 2006).

A logical question regarding word reading asks how do the word recognition processes critical to fluent reading become automatic? The orthographic mapping theory (Ehri, 2005, 2014) hypothesizes that learning to efficiently read words in an alphabetic language requires that the individual letters within words become anchored in the reader's lexical memory by their associated

sounds. As children acquire awareness of the 44 sounds (phonemes) in the English language, they connect these sounds to the 26 individual letters and their various combinations. Through extensive practice of reading text, these connections between orthographic features within words and their associated sounds become bonded in long-term memory. In sum, children do not memorize how to pronounce strings of individual letters for the tens of thousands of words they will learn. This makes practice reading text critical to word reading and fluent reading processes (Anderson et al., 1988; Cunningham and Stanovich, 1991, 1998; Elley, 2000; Taylor et al., 1990).

Developing readers achieve word recognition automaticity in the same way that other automatic processes in life are developed—through wide, deep, and assisted practice. Wide reading refers to the common classroom practice of reading a text once followed by discussion and extension. This routine is repeated with each new text. A general purpose of wide reading is to increase the volume of reading by having students read one new text after another. This is the type of reading done by most proficient readers. Studies have demonstrated that students' achievement in reading is related to the amount of time spent reading (Postlethwaite and Ross, 1992; Samuels and Wu, 2004).

Deep or close reading is more commonly referred to as repeated reading (Samuels, 1979). Deep reading occurs when a student is asked to read a single text multiple times until fluency on that passage is achieved. The initial reading of a text for developing and struggling students who have not yet achieved automaticity in their word recognition is often marked by characteristics of disfluent reading. The slow, halting reading that characterizes less than automatic word recognition will have a detrimental effect on the reader's comprehension. However, when that text is read multiple times the reader will begin to incrementally approach fluency on that passage.

Repeated practice improves the performance of the actual text practiced. The real value of deep or repeated reading, however, is manifested when students move on to new and previously unread passages. What students gained from the repeated reading of one passage generalizes, at least in part, to new passages.

Assisted reading involves a student reading a text while simultaneously listening to a fluent oral rendering of the same text (Rasinski, 2010). Assisted reading can take a variety of forms (Rasinski, 2010) – choral reading where members of a group of readers support one another (Paige, 2011a,b), paired or duolog reading in which a more fluent reader reads with a less fluent reader (Heckelman, 1969; Topping, 1987, 1989), technology- or audio-assisted reading in which student reads and listens to a pre-recorded oral reading of the text (Carbo, 1978; Chomsky, 1976), and captioned video recordings in which the text heard on the video is also represented in captions on the video screen. Repeated and assisted reading can often be implemented in tandem where early readings of one text are done with assistance, but that assistance diminishes with subsequent readings of the text. A number of studies and reviews of research on fluency have shown that word recognition accuracy, automaticity, comprehension, and attitude toward reading improve with repeated and assisted readings (Chard et al., 2002; Dowhower, 1994; Kuhn and Stahl, 2003; Rasinski et al., 2011; Stevens et al., 2017).). Wide, deep, and assisted reading are foundational to any effective fluency program or intervention.

Prosodic processes

Peter Schreiber (1980, 1991; Schreiber and Read, 1980) provided a complementary explanation for the positive results from repeated readings. He argued that through repeated practice of a text students develop a greater awareness of the prosodic or melodic features of oral reading and speech. That is, through repeated readings, they were learning to embed in their reading the expressive and intonational features of oral speech that help to mark phrase boundaries within and between sentences and convey meaning. Dysfluent readers tend to read in a slow, word-by-word manner that does not lend itself well to prosodic and syntactically appropriate phrased reading that carries meaning. Through repeated readings, even disfluent readers are more able to capture the prosodic and syntactic essence of the text, thus improving the surface-level processing of the passage as well as text comprehension. As with spoken language, a reader uses expression and phrasing to bring a text to life making prosody important to understanding written language whether reading silently or aloud (Ashby, 2006; Cowie et al., 2002; Erikson, 2010; Fodor, 1998; Miller and Schwanenflugel, 2006; Schreiber, 1980, 1991; Schwanenflugel et al., 2004). By using prosodic features, a reader is able to bracket syntactic and semantic features of the text to improve reading comprehension (Kuhn and Stahl, 2003; Miller and Schwanenflugel, 2006; Paige et al., 2014; Rasinski, 1985). Multiple researchers in the US and other countries have shown that prosody allows for additional understanding of text beyond that accounted for by other factors. Indeed, a significant body of research over the past two decades (Benjamin and Schwanenflugel, 2010; Daane et al., 2005; Klauda and Guthrie, 2008; Miller and Schwanenflugel, 2006; Paige et al., 2012, 2014; Pinnell et al., 1995; Rasinski et al., 2009; Valencia et al., 2010; Veenendaal et al., 2014, 2016) has demonstrated the connection between prosody and comprehension.

Fluency to comprehension

An abundance of research has demonstrated that fluent reading benefits textual understanding whether one is reading aloud or silently (Carlson, 2009; Frazier et al., 2006; Rasinski et al., 2011). While fluent reading is often helpful to comprehension, it does not cause reading comprehension and does not always overcome or compensate for other associated cognitive processes. An example of this was reported by Swanson and O'Connor (2009) who found that increasing reading fluency in students with poor working memory did not increase reading comprehension.

However, for many readers improving reading fluency has been shown by multiple researchers over several decades to improve understanding of text. In a meta-analysis of 15 studies using repeated reading to improve fluency, Therrien (2004) found that fluency improved reading comprehension nearly one-half of a standard deviation. Multiple authors have found reading fluency to be consistently associated with one-fourth to over one-half of the variance in reading comprehension across elementary, middle, and secondary students (Daane et al., 2005; Fuchs et al., 2001; Jenkins, Fuchs, Espin, van den Broek and Deno, 2003a, 2003b; Kuhn and Stahl, 2003; Paige, 2011a,b; Paige et al., 2014; Pinnell et al., 1995; Schatschneider et al., 2004; Stecker et al., 1998; Wang et al., 2019; Young et al., 1996).

Reading fluency reflects the efficient operation of lower-level reading processes, that in addition to phonemic awareness include letter-sound knowledge, the application of that knowledge to decoding words, and the accumulated inventory of the reader's sight-word vocabulary (Hoover and Gough, 1990; Perfetti and Hart, 2002). The lexical quality hypothesis (Perfetti and Hart, 2002) posits that textual understanding is based on the extent to which these processes operate efficiently to connect word pronunciations to their meanings stored in the reader's long-term memory. This makes word recognition and meaning inextricably linked (White et al., 1990). Because attention, working memory, and executive function are dependent on cognitive resources, poor reading fluency requires the reader to direct a disproportionate amount of those resources to lower level decoding processes rather than to the higher level task of integrating the author's message with their background knowledge to create meaning (Kintsch, 1998; Walczyk et al., 2004).

In summary, the science supporting the importance of reading fluency to proficient reading is rich and deep. Ensuring students become fluent readers results in better understanding of text, students who are more motivated to read, and a higher likelihood of a successful academic outcome (Paige et al., 2019). Both scientific conceptualizations of reading fluency, automatic processing of the surface-level features of text, and the ability to attend to the prosodic and syntactic features of text while reading, appear compelling and are considered central elements of contemporary conceptions of reading fluency (Kuhn and Stahl, 2003).

Fluency as art

A scientific approach to understanding fluency is clearly important, however it may not be sufficient for reading instruction that aims to develop proficient, engaged, and lifelong readers. Although science informs what children need to learn at various stages of their reading acquisition, it does not fully answer the question of how that teaching of those learning outcomes should occur. We suggest that scientifically-based fluency instruction can and should be delivered in artful and authentic ways that students find engaging and worthwhile. We propose that successful teaching of reading blends both science of reading and teacher art or craft.

What do we mean when we say that teaching is an art? Art can be defined as the expression or application of human creative skill and imagination, producing works to be appreciated primarily for their emotional power. In its broadest sense, "art" means "to arrange." Art is created when someone arranges things found in the world into a new or different design or form. Thus, teachers are artists when they arrange instruction to evoke an aesthetic or emotional response from the audience: the students. Science of reading scholar and advocate Mark Seidenberg (as cited in Goldstein, 2020) acknowledged that "The science that you need to know [reading acquisition] is good. The science on how to teach it [reading] effectively is not."

Fluency is the ability to read the words in texts automatically and with a level of expression (prosody) that reflects the meaning of the text. As described earlier, two of the scientifically validated methods for developing fluency are assisted reading and repeated reading. Assisted reading involves students reading a text orally while simultaneously listening to a fluent rendering of the same text. Repeated reading involves multiple readings of a text until students can read it at a level that achieves or approaches fluent reading. Taking an artful approach, teachers need to consider authentic situations in which students engage in assisted and repeated reading for the dual purposes of increasing automatic and expressive reading of the text? One answer lies in performance. Whether actors rehearsing for a theatrical performance, poets practicing their poems for a performance at a coffee house, or authors reading over a selection of their book in anticipation of a public reading, the rehearsal often involves assisted and repeated reading. The essential difference is in the purpose of the assisted and repeated readings: to communicate with and engage an audience. Similarly, choral reading, an assisted reading strategy where students read in unison with the teacher, allows students to support one another and provides a sense of the rhythm and feel for the language of a text reflecting the author's art. Results of studies of artful approaches to fluency have been promising. For example, studies of fluency instruction using readers theater, poetry, and oratory have led to significant and substantive improvements in fluency, word decoding, and even comprehension (Griffith and Rasinski, 2004; Young and Rasinski, 2009, 2018).

Fluency—Beyond English

A growing body of evidence points to reading fluency in English as an important factor in student reading success. While there is a growing recognition of the importance of fluency in English, the research on fluency of connected text reading in other languages is less robust. Less is known about the extent to which reading fluency is a significant competency in learning to read languages other than English.

Understanding that reading fluency is an important reading competency in other languages may lead to a better understanding that reading fluency is a more universal variable across many written languages. In a longitudinal study (Kim, 2015) involving 143

Korean speaking children text reading fluency was found to be significantly associated with reading comprehension. Later, Kim (2020) examined the relationship of component reading competencies to reading comprehension in Korean with 201 first-grade children. The study's findings revealed that text reading fluency had a moderate but significant correlation ($r = 0.51$) to reading comprehension. Unlike English, the Korean language uses an alphasyllabary writing system and is relatively transparent with consistent phoneme-grapheme correspondences (Kim, 2011).

A cross-linguistic study conducted by Baker et al. (2012) aimed to investigate whether initial status and growth rates in reading fluency in Spanish and English significantly predicted reading comprehension within and across languages. Participants were 173 Spanish-speaking English learners (EL) in first grade, 156 ELs in second grade, and 142 ELs in third grade across four schools providing a paired bilingual reading program. Results from regression and path analyses indicated that students' initial scores and growth on reading fluency were strong and direct predictors of reading comprehension within the same language, but not across other languages. This suggests that reading fluency is a good predictor of reading comprehension in Spanish as well as English. Additionally, other studies at various grade levels in Spanish have found that reading fluency to be an important component of proficient reading (e.g., Calet et al., 2015; Crosson and Lesaux, 2010).

The Turkish language is orthographically transparent. Most of the studies among Turkish language users have shown strong relationships between fluency and comprehension (e.g., Basaran, 2013; Bastug and Akyol, 2012; Kocaarslan, 2019; Yildirim et al., 2020; Yildirim and Rasinski, 2014). Yildirim, Rasinski, and Kaya (2017, 2019) found positive relationships between reading fluency and comprehension for narrative texts for grades 5 through 8 and grades 6 through 8 for expository reading comprehension.

Studies in Greek have reported positive relationships between fluency and reading comprehension and literacy acquisition (e.g., Georgiou et al., 2012; Georgiou et al., 2008; Manolitsis, Grikorakis, and Georgiou, 2017). Padeliadu and Antoniou (2014), for example, investigated reading comprehension in Greek, a language that is less transparent than English, with 1070 elementary and secondary students. The findings indicated significant positive correlations between word decoding and fluency with all comprehension measures at each grade level.

Liao et al. (2020) aimed to identify the predictors of Chinese reading and literacy skills among Chinese school children in Taiwan. The Chinese language, which has an opaque orthography, employs one of the most complex orthographic systems in the world (McBride, 2016). In the study, 182 first-grade students were recruited and structural equation modeling analysis found that reading comprehension at Grade 3 was predicted by reading fluency and literacy skills at Grade 2. Other studies in Chinese have also revealed the importance of reading fluency for reading comprehension at different grade levels (e.g., Zhao, Kwok, Liu, Liu, and Huang, 2017; Joshi et al., 2012; Li and Wu, 2015; Yeung et al., 2016).

These studies, among others, of reading fluency in languages other than English suggest that fluency in reading is a significant indicator of reading proficiency in various language and orthographies, transparent as well as opaque. However, more research needs to be done within and across languages to extend our knowledge regarding the relationship and role of reading fluency with reading comprehension in still other language contexts.

Assessing reading fluency

As with other phenomenon, how reading fluency is defined reflects how it is assessed (Andreassen and Bråten, 2010; Best et al., 2008; Nation and Snowling, 1997). For many years, reading fluency has been conceptualized as text processing where a student's word identification accuracy and automaticity, usually termed Oral Reading Fluency (ORF) is measured as a reading rate or words-correct-per-minute (WCPM) metric (Deno, 1985; Hudson et al., 2005). The most common process for assessing ORF is quite simple. The student is asked to read a grade level text orally for 1 min. The examiner marks any uncorrected errors made in the reading and then determines the number of words read correctly. That score is then compared against previously established norms for grade level and time within the school year (e.g., Hasbrouck and Tindal, 2017). Despite its ease and simplicity of administration, the ORF score has proven to be strongly related to measures of reading comprehension (Deno, 1985; Fuchs et al., 2001). However, while the ORF/WCPM metric captures the automaticity component of fluency it does not provide a measure of prosody.

One major concern with frequent use of the ORF/WCPM score is that it provides students (and teachers) with an implicit notion that automaticity in word recognition is achieved through exercises in increasing students' reading speed (Samuels, 2007; Rasinski, 2010). As students are taught explicitly to read fast, their ORF/WCPM scores, and thus fluency, should improve. The problem is that students may come to incorrectly define proficient reading as merely fast reading. Missing from this line of thought is that ORF/WCPM is a consequence, not a cause, of automatic word recognition. For most fluent readers, automaticity in word recognition is caused through authentic wide, deep, and assisted reading.

While reading prosody can be assessed using computer-based, spectrographic technology, a more common approach is to closely observe a student reading a passage aloud, and then evaluate the reading for the qualities of prosodic reading using a rating scale in the form of a rubric (Benjamin and Schwanenflugel, 2010). Several rubrics have been used with the two most common being a scale used in studies of the National Assessment of Educational Progress (NAEP (Daane et al., 2005; Pinnell et al., 1995)) and the *Multidimensional Fluency Scale* (MDFS (Zutell and Rasinski, 1991); available at www.timrasinski.com > Resources). The NAEP scale is a four-indicator rubric where a 1 represents word-by-word reading and a 4 indicates fully developed, fluent reading of the text. The MDFS on the other hand, evaluates four domains of fluent reading including reading pace, smoothness, phrasing, and expression/volume. Each of these domains is then evaluated on a 1–4 scale with a 4 indicating full development. While the NAEP has a range of 4, the range of the MDFS is 4–16. Using generalizability and decision analysis

techniques, studies by Moser et al. (2014) and Smith and Paige (2019) have found that both measures exhibit high reliability. Smith and Paige also specify a protocol for rater training. Research has also found significant correlations between observer ratings of prosody and students' overall reading proficiency (Daane et al., 2005; Pinnell et al., 1995; Rasinski et al., 2009).

Standardized assessments of reading fluency include the *Gray Oral Reading Test -5* (Wiederholt and Bryant, 2012), the advantage of which allows the comparison of a reader to a large sample of peer readers. A widely used program is the *Dynamic Indicator of Basic Literacy Skills* (DIBELS; Good and Kaminski, 2005) that assess a variety of literacy competencies, including automaticity.

Although fluency assessment has traditionally been conducted on a one-to-one basis between a student and a teacher, technology has emerged that harnesses voice recognition technology that can be used to assess various aspects of fluency. Most of these assessments are available from corporations specializing in district-wide assessments. Such assessments may become more widely available as technology advances and less commercially-driven entities enter the field.

Conclusion

The emerging evidence from scientific studies is that reading fluency is a complex, foundational, and essential competency for proficient reading in English as well as other written languages. As such, intentional instruction in reading fluency must occur to ensure reading success for all students. Moreover, assessment of fluency and progress monitoring of students' growth in fluency need to also be part of such curricula. Questions however remain as to how to provide the most efficacious instruction fluency. Among these questions are the following:

- When should fluency instruction begin?
- What are the appropriate grade (age) levels for fluency instruction to occur?
- How can fluency instruction be based on the science of fluency as well as addressing the need for such instruction to be authentic and artful? Indeed, the increase in available reading fluency strategies integrated into daily classroom literacy blocks warrants further research and development of engaging activities that bolster reading fluency, which can lead to improved proficiency in reading.
- To what extent do instructional approaches aimed specifically at increasing reading rate lead to improved fluency and improved overall reading achievement.
- What text types and genre, if any, are most ideal for fluency instruction?
- What level of text, in relation to the reading level of the student, is most appropriate for fluency instruction?
- What role do/can parents and families be expected to play in developing fluency?
- How can fluency instruction be effectively delivered in virtual/online instructional environments?
- To what extent can fluency assessment be delivered through digital and remote learning programs and environments?
- To what extent is fluency a significant literacy competency beyond English and the languages mentioned in this paper?
- To what extent is fluency an issue beyond the elementary grades?
- What are the norms for word recognition automaticity (words correct per minute) beyond grade 8?
- Are there ceiling levels of fluency development after which fluency instructional need not be mandated?
- Although the term "words-correct-per-minute" includes word accuracy and automaticity and excludes prosody, many in the field use it to refer to the construct of fluency. To clarify this confounding of indicators, Paige and Magpuri-Lavell (2012) have proposed the term *accumativity* in place of words-correct-per-minute.
- How best can fluency information be communicated to parents and other constituent groups?

In addition to these questions that focus on delivering instruction in fluency to students, future research may wish to address questions about the nature of fluency and our evolving understanding of reading fluency. For example, Cartwright et al. (2019) make the case that reading fluency itself involves more than automaticity in word recognition and prosody in oral reading. They argue that executive functioning, and in particular graphophonological-semantic cognitive flexibility plays a significant role in fluent reading.

Regardless of future developments in our understanding and instruction of reading fluency, it is clear that it can no longer be ignored. Intentional and dedicated instruction in reading fluency must become a non-negotiable part of any effective elementary reading curriculum.

References

- Allington, R.L., 1983. Fluency: the neglected reading goal. *The Read. Teach.* 36, 556–561.
- Anderson, R.C., Wilson, P.T., Fielding, L.G., 1988. Growth in reading and how children spend their time outside of school. *Read. Res. Q.* 23, 285–303.
- Andreassen, R., Braten, I., 2010. Examining the Prediction of Reading Comprehension.
- Ashby, J., 2006. Prosody in skilled silent reading: evidence from eye movements. *J. Res. Read.* 29, 318–333.
- Baker, D.L., Park, Y., Baker, S.K., 2012. The reading performance of English learners in grades 1–3: the role of initial status and growth on reading fluency in Spanish and English. *Read. Writ.* 25, 251–281.
- Basaran, M., 2013. Reading fluency as an indicator of reading comprehension. *Kuram ve Uygulamada Egitim Bilimleri* 13, 2277–2290.
- Bastug, M., Akyol, H., 2012. The level of prediction of reading comprehension by fluent reading skills. *J. Theor. Educ. Sci.* 54, 394–411.
- Benjamin, R.G., Schwanenflugel, P.J., 2010. Text complexity and oral reading prosody in young readers. *Read. Res. Q.* 45, 388–404.

- Best, R.M., Floyd, R.G., McNamara, D.S., 2008. Differential competencies contributing to children's comprehension of narrative and expository texts. *Read. Psychol.* 29, 137–164.
- Calet, N., Gutiérrez-Palma, N., Defior, S., 2015. A cross-sectional study of fluency and reading comprehension in Spanish primary school children. *J. Res. Read.* 38, 272–285.
- Carbo, M., 1978. Teaching reading with talking books. *Read. Teach.* 32, 267–273.
- Carlson, K., 2009. How prosody influences sentence comprehension. *Lang. Linguist. Compass* 3, 1188–1200.
- Cartwright, K.B., Marshall, T.R., Huemer, C.M., Payne, J.B., 2019. Executive function in the classroom: cognitive flexibility supports reading fluency for typical readers and teacher-identified low-achieving readers. *Res. Dev. Disabil.* 88, 42–52.
- Castles, A., Rastle, K., Nation, K., 2018. Ending the reading wars: reading acquisition from novice to expert. *Psychol. Sci. Publ. Interest* 19, 5–51.
- Chard, D.J., Vaughn, S., Tyler, B., 2002. A synthesis of research on effective interventions for building fluency with elementary students with learning disabilities. *J. Learn. Disabil.* 35, 386–406.
- Chomsky, C., 1976. After decoding: what? *Lang. Arts* 53, 288–296.
- Cowie, R., Douglas-Cowie, E., Wichmann, A., 2002. Prosodic characteristics of skilled reading: fluency and expressiveness in 8–10 year-old readers. *Lang. Speech* 45, 47–82.
- Crosson, A.C., Lesaux, N.K., 2010. Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text-reading fluency in English. *Read. Writ.* 23, 475–494.
- Cunningham, A.E., Stanovich, K.E., 1991. Tracking the unique effects of print exposure in children: associations with vocabulary, general knowledge, and spelling. *J. Educ. Psychol.* 83, 264–274.
- Cunningham, A.E., Stanovich, K.E., 1998. The impact of print exposure on word recognition. In: Metsala, J., Ehri, L. (Eds.), *Word Recognition and Beginning Literacy*. Erlbaum, Mahwah, NJ, pp. 235–262.
- Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M.J., Oranje, A., 2005. *Fourth-grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading*. U.S. Department of Education, Institute of Education Sciences, Washington, DC.
- Deno, S.L., 1985. Curriculum-based measurement: the emerging alternative. *Except. Child.* 52, 219–232.
- Dowhower, S.L., 1994. Repeated reading revisited: research into practice. *Read. Writ. Q.* 10, 343–358.
- Ehri, L., 2014. Orthographic mapping in the acquisition of sight word reading, spelling, memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21.
- Ehri, L.C., 2005. Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188.
- Elley, W.B., 2000. The potential of book floods for raising literacy levels. *Int. Rev. Educ.* 46, 233–255.
- Erekson, J.A., 2010. Prosody and interpretation. *Read. Horiz.* 50, 80–98.
- Fodor, J.D., 1998. Parsing to learn. *J. Psycholinguist. Res.* 27, 339–375.
- Frazier, L., Carlson, K., Clifton, C., 2006. Prosodic phrasing is central to language comprehension. *Trends Cognit. Sci.* 10, 244–249.
- Fuchs, L.S., Fuchs, D., Hosp, M.K., Jenkins, J.R., 2001. Oral reading fluency as an indicator of reading competence: a theoretical, empirical, and historical analysis. *Sci. Stud. Read.* 5, 239–256.
- Georgiou, G.K., Parrila, R., Papadopoulos, T.C., 2008. Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *J. Educ. Psychol.* 100, 566–580.
- Georgiou, G.K., Torppa, M., Manolitsis, G., Lyytinen, H., Parrila, R., 2012. Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Read. Writ.* 25, 321–346.
- Goldstein, D., 2020. An Old and Contested Solution to Boost Reading Scores: Phonics. *New York Times*.
- Good, R., Kaminski, R., 2005. *Dynamic Indicators or Basic Early Literacy Skills*, sixth ed. University of Oregon, Eugene, OR.
- Goodman, K., 1964. The linguistics of reading. *Elem. Sch. J.* 64 (8), 355–361.
- Gough, P., Tunmer, W., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10.
- Griffith, L.W., Rasinski, T.V., 2004. A focus on fluency: how one teacher incorporated fluency with her reading curriculum. *Read. Teach.* 58, 126–137.
- Goldenberg, C., 2020. Reading wars, reading science, and English learners. *Read. Res. Q.* 55, 131–144.
- Gross, J., Millett, A.L., Bartek, B., Bredell, K.H., Winegard, B., 2014. Evidence for prosody in silent reading. *Read. Res. Q.* 49, 189–208.
- Hanford, E., 2018, September 10. *Hard Words: Why Aren't Kids Being Taught to Read?* APM Reports. Retrieved from: <https://www.apmreports.org/episode/2018/09/10/hard-words-why-american-kids-arent-beingtaught-to-read>.
- Hanford, E., 2019, August 22. *At a Loss for Words: How a Flawed Idea Is Teaching Millions of Kids to Be Poor Readers*. APM Reports. Retrieved from: <https://www.apmreports.org/episode/2019/08/22/whats-wrong-how-schools-teach-reading>.
- Hasbrouck, J., Tindal, G., 2017. *An Update to Compiled ORF Norms*. Technical Report No. 1702. Behavioral Research and Teaching, University of Oregon, Eugene, OR.
- Heckelman, R.G., 1969. A neurological impress method of reading instruction. *Acad. Ther.* 4, 277–282.
- Hoover, W.A., Gough, P.B., 1990. The simple view of reading. *Read. Writ.* 2, 27–160.
- Hudson, R.F., Lane, H., Pullen, P., 2005. Reading fluency assessment and instruction: what, why, and how? *Read. Teach.* 58, 702–714.
- Huey, E.B., 1908. *The Psychology and Pedagogy of Reading*, fifth ed. MIT Press, Cambridge, MA.
- Hyatt, A.V., 1943. *The Place of Oral Reading in the School Program: Its History and Development from 1880–1941*. Teachers College Press, New York.
- Iwasaki, B., Rasinski, T., Yildirim, K., Zimmerman, B., 2013. Let's bring back the magic of song for teaching reading. *Read. Teach.* 67, 137–141.
- Jenkins, J.R., Fuchs, L.S., Espin, C., van den Broek, P., Deno, S.L., 2003a. Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learn. Disabil. Res. Pract.* 18, 237–245.
- Jenkins, J.R., Fuchs, L.S., Espin, C., van den Broek, P., Deno, S.L., 2003b. Sources of individual differences in reading comprehension and reading fluency. *J. Educ. Psychol.* 95, 719–729.
- Joshi, R.M., Tao, S., Aaron, P.G., Quiroz, B., 2012. Cognitive component of componential model of reading applied to different orthographies. *J. Learn. Disabil.* 45, 480–486.
- Judd, C.H., Buswell, G.T., 1922. *Silent Reading: A Study of the Various Types*. University of Chicago Press, Chicago, IL.
- Kim, Y.G., 2015. Developmental, component-based model of reading fluency: an investigation of predictors word-reading fluency, text reading fluency, and reading comprehension. *Read. Res. Q.* 50, 459–481.
- Kim, Y.-S.G., 2011. Considering linguistic and orthographic features in early literacy acquisition: evidence from Korean. *Contemp. Educ. Psychol.* 36, 177–189.
- Kim, Y.-S.G., 2020. Toward integrative reading science: the direct and indirect effects of reading. *J. Learn. Disabil.* 1–23.
- Kintsch, W., 1998. *Comprehension: A Paradigm for Cognition*. Cambridge University Press, Cambridge, UK.
- Klauda, S.L., Guthrie, J.T., 2008. Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321.
- Kodan, H., Akyol, H., 2018. Effects of choral, repeated and assisted reading strategies on reading and reading comprehension skills of poor readers. *Educ. Sci.* 43, 159–179.
- Kocaarslan, M., 2019. The effects of reading rate, accuracy and prosody on second grade students' oral retelling. *Acta Psychol.* 197, 86–93.
- Kuhn, M.R., Stahl, S.A., 2003. Fluency: a review of developmental and remedial practices. *J. Educ. Psychol.* 95, 3–21.
- LaBerge, D., Samuels, S.J., 1974. Toward a theory of automatic information processing in reading. *Cognit. Psychol.* 6, 293–323.
- Li, L., Wu, X., 2015. Effects of metalinguistic awareness on reading comprehension and the mediator role of reading fluency from grades 2 to 4. *PLoS One* 10, 1–16.
- Liao, C.H., Kuo, B.C., Tsao, C.J., Mok, M.M.C., 2020. Predictors of Chinese reading and literacy skills among Chinese school children: a 3-year longitudinal study. *Educ. Psychol.* 40, 838–855.
- Logan, G.D., 1988. Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527.
- Logan, G.D., Klapp, S.T., 1991. Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *J. Exp. Psychol. Learn. Mem. Cognit.* 17, 179–195.

- Manolitsis, G., Grigorakis, I., Georgiou, G.K., 2017. The longitudinal contribution of early morphological awareness skills to reading fluency and comprehension in Greek. *Front. Psychol.* 8, 1793.
- Martinez, M., Roser, N., Strecker, S., 1999. "I never thought I could be a star": a Readers Theatre ticket to reading fluency. *Read. Teach.* 52, 326–334.
- McBride, C.A., 2016. Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educ. Psychol. Rev.* 28, 523–549.
- Miller, J., Schwanenflugel, P.J., 2006. Prosody of syntactically complex sentences in the oral reading of young children. *J. Educ. Psychol.* 98, 839–843.
- Miller, J., Schwanenflugel, P.J., 2008. A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Read. Res. Q.* 43, 336–354.
- Moser, G.P., Sudweeks, R.R., Morrison, T.G., Wilcox, B., 2014. Reliability of ratings of children's expressive reading. *Read. Psychol.* 35, 58–79.
- Nation, K., Snowling, M., 1997. Assessing reading difficulties: the validity and utility of current measures of reading skill. *Br. J. Educ. Psychol.* 67, 359–370.
- National Reading Panel, 2000. Report of the National Reading Panel: Teaching Children to Read. Report of the Subgroups. U.S. Department of Health and Human Services, National Institutes of Health, Washington, DC.
- Padeliadu, S., Antoniou, F., 2014. The relationship between reading comprehension, decoding, and fluency in Greek: a cross-sectional study. *Read. Writ. Q.* 30, 1–31.
- Paige, D.D., 2011a. "That sounded good!" Using whole-class choral reading to improve reading fluency. *Read. Teach.* 64 (6), 435–438. <https://doi.org/10.1598/RT.64.6.5>.
- Paige, D.D., 2011b. Engaging struggling adolescent readers through situational interest: a model proposing the relationships among extrinsic motivation, oral reading fluency, comprehension, and academic achievement. *Read. Psychol.* 32, 395–425.
- Paige, D.D., Magguri-Lavell, T., 2012. Reading fluency in the middle and secondary grades. *International Electronic Journal of Elementary Education* 7, 83–96.
- Paige, D.D., Rasinski, T., Magguri-Lavell, T., 2012. Is fluent, expressive reading important for high school readers? *Journal of Adult & Adolescent Literacy* 56, 67–76.
- Paige, D.D., Rasinski, T., Magguri-Lavell, T., Smith, G.S., 2014. Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *J. Literacy Res.* 46, 123–156.
- Paige, D.D., Smith, G.S., Rasinski, T.V., et al., 2019. A path analytic model linking foundational skills to grade 3 state reading achievement. *J. Educ. Res.* 112, 110–120.
- Perfetti, C.A., Hart, L., 2002. The lexical quality hypothesis. In: Verhoeven, L.T., Elbro, C., Reitsma, P. (Eds.), *Precursors of Functional Literacy*, vol. 11. John Benjamins Publishing, Amsterdam, the Netherlands, pp. 189–213.
- Petersen-Brown, S., Johnson, M.E., Bowen, J., Lundberg, A.R., Nelson, J.D., Williamson, A.A., Wiswell, J.M., 2021. Is repeated reading evidence-based? A review of the literature. *Prev. Sch. Fail.* 65, 379–391.
- Pikulski, J.J., Chard, D.J., 2005. Fluency: bridge between decoding and reading comprehension. *Read. Teach.* 58, 510–519.
- Pinnell, G.S., Pikulski, J.J., Wixson, K.K., et al., 1995. *Listening to Children Read Aloud*. U.S. Department of Education, Office of Educational Research and Improvement, Washington, DC.
- Postlethwaite, T.N., Ross, K.N., 1992. *Effective Schools in Reading: Implications for Policy Planner*. International Association for the Evaluation of Educational Achievement, The Hague.
- Rasinski, T.V., 1985. A Study of Factors Involved in Reader-Text Interactions that Contribute to Fluency in Reading. Unpublished doctoral dissertation, The Ohio State University.
- Rasinski, T.V., 2010. *The Fluent Reader: Oral and Silent Reading Strategies for Building Word Recognition, Fluency, and Comprehension*, second ed. Scholastic, New York.
- Rasinski, T.V., Reutzel, C.R., Chard, D., Linan-Thompson, S., 2011. Reading fluency. In: Kamil, M.L., Pearson, P.D., Moje, E.B., Afflerbach, P.P. (Eds.), *Handbook of Reading Research*, vol. IV. Routledge, New York, pp. 286–319.
- Rasinski, T.V., Rikli, A., Johnston, S., 2009. Reading fluency: more than automaticity? More than a concern for the primary grades? *Literacy Research and Instruction* 48, 350–361.
- Rawson, K.A., Middleton, E.L., 2009. Memory-based processing as a mechanism of automaticity in text comprehension. *J. Exp. Psychol. Learn. Mem. Cognit.* 35, 353–370.
- Reed, D.K., Zimmermann, L.M., Reeger, A.J., Aloe, A.M., 2019. The effects of varied practice on the oral reading fluency of fourth-grade students. *J. Sch. Psychol.* 77, 24–35.
- Sabatini, J., Wang, Z., O'Reilly, T., 2019. Relating reading comprehension to oral reading performance in the NAEP fourth-grade special study of oral reading. *Read. Res. Q.* 54, 253–271.
- Samuels, S.J., 1979. The method of repeated readings. *Read. Teach.* 32, 403–408.
- Samuels, S.J., 2007. The DIBELS tests: is speed of barking at print what we mean by reading fluency? *Read. Res. Q.* 42, 563–566.
- Samuels, S.J., Wu, Y.-C., 2004. How the Amount of Time Spent on Independent Reading Affects Reading Achievement: A Response to the National Reading Panel. Accessed at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.539.9906>. (Accessed 23 July 2020).
- Schatschneider, C., Buck, J., Torgesen, J.K., Wagner, R.K., et al., 2004. A Multivariate Study of Factors that Contribute to Individual Differences in Performance on the Florida Comprehensive Reading Assessment Test. Technical Report No.5. Florida Center for Reading Research, Tallahassee.
- Schreiber, P.A., 1980. On the acquisition of reading fluency. *J. Read. Behav.* 12, 177–186.
- Schreiber, P.A., 1987. Prosody and structure in children's syntactic processing. In: Horowitz, R., Samuels, S.J. (Eds.), *Comprehending Oral and Written Language*. Academic, New York, pp. 243–270.
- Schreiber, P.A., 1991. Understanding prosody's role in reading acquisition. *Theor. Pract.* 30, 158–164.
- Schreiber, P.A., Read, C., 1980. Children's use of phonetic cues in spelling, parsing, and maybe-reading. *Bull. Ort. Soc.* 30, 209–224.
- Schwanenflugel, P.J., Hamilton, A.M., Kuhn, M.R., Wisenbaker, J.M., Stahl, S.A., 2004. Becoming a fluent reader: reading skill and prosodic features in the oral reading of young readers. *J. Educ. Psychol.* 96, 119–129.
- Smith, G.S., Paige, D.D., 2019. A study of reliability across multiple raters when using the NAEP and MDFS rubrics to measure oral reading fluency. *Read. Psychol.* 40, 34–69.
- Stanovich, K.E., 1980. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Read. Res. Q.* 16, 32–71.
- Stecker, S.K., Roser, N.L., Martinez, M.G., 1998. Understanding oral reading fluency. In: Shanahan, T., Rodriguez-Brown, F.V. (Eds.), *47th Yearbook of the National Reading Conference*. National Reading Conference, Chicago, IL, pp. 295–310.
- Stevens, E., Walker, M., Vaughn, S., 2017. The effects of fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: a synthesis of the research from 2001–2014. *J. Learn. Disabil.* 50, 576–590.
- Swanson, H.L., O'Connor, R., 2009. The role of working memory and fluency practice on the reading comprehension of students who are dysfluent readers. *J. Learn. Disabil.* 42, 548–575.
- Taylor, B.M., Fry, B.J., Maruyama, G.M., 1990. Time spent reading and reading growth. *Am. Educ. Res. J.* 27, 351–362.
- Therrien, W.J., 2004. Fluency and comprehension gains as a result of repeated reading. *Remedial Spec. Educ.* 25, 252–261.
- Topping, K., 1987. Paired reading: a powerful technique for parent use. *Read. Teach.* 40, 604–614.
- Topping, K., 1989. Peer tutoring and paired reading. Combining two powerful techniques. *Read. Teach.* 42, 488–494.
- Torgesen, J.K., 2004. Bringing it all together: from phonemic awareness to fluency. In: Paper Presented at the CORE Literacy Leadership Summit. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.365.9335&rep=rep1&type=pdf>.
- Torgesen, J.K., Hudson, R.F., 2006. Reading fluency: critical issues for struggling readers. In: Samuels, S.J., Farstrup, A.E. (Eds.), *What Research Has to Say about Fluency Instruction*. International Reading Association, Newark, DE, pp. 130–158.
- Valencia, S.W., Smith, A., Reece, A.M., Li, M., Wixson, K.K., Newman, H., 2010. Oral reading fluency assessment: issues of construct, criterion, and consequent validity. *Read. Res. Q.* 45, 521–536.
- Veenendaal, N.J., Groen, M.A., Verhoeven, L., 2014. The role of speech prosody and text reading prosody in children's reading comprehension. *Br. J. Educ. Psychol.* 84, 521–536.
- Veenendaal, N.J., Groen, M.A., Verhoeven, L., 2016. Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades: a longitudinal perspective. *Sci. Stud. Read.* 20, 189–202.

- Walczyk, J.J., Marsiglia, C.S., Johns, A.K., Bryan, K.S., 2004. Children's compensations for poorly automated reading skills. *Discourse Process* 37, 47–66.
- Wang, Z., Sabitini, J., O'Reilly, T., Weeks, J., 2019. Decoding and reading comprehension: a test of the decoding threshold hypothesis. *J. Educ. Psychol.* 111, 387–401.
- Wexler, J., Vaughn, S., Roberts, G., Denton, C.A., 2010. The efficacy of repeated reading and wide reading practice for high school students with severe reading disabilities. *Learn. Disabil. Res. Pract.* 25, 2–10.
- Whalley, K., Hansen, J., 2006. The role of prosodic sensitivity in children's reading development. *J. Res. Read.* 29, 288–303.
- White, T.G., Graves, M.F., Slater, W.H., 1990. Growth of reading vocabulary in diverse elementary schools: decoding and word meaning. *J. Educ. Psychol.* 82, 281–290.
- Wiederholt, J.L., Bryant, B.R., 2012. *Gray Oral Reading Test*, fifth ed. Pro-Ed, Austin, TX.
- Wilfong, L., 2008. Building fluency, word recognition ability, and confidence in struggling readers: the poetry academy. *Read. Teach.* 62, 4–13.
- Wyse, D., Bradbury, A., 2022. Reading wars or reading reconciliation? A critical examination of robust research evidence, curriculum policy and teachers' practices for teaching phonics and reading. *Rev. Educ.* 10, 1–53.
- Yeung, P.S., Ho, C.S.H., Chan, D.W.O., Chung, K.K.H., 2016. A componential model of reading in Chinese. *Learn. Individ. Differ.* 45, 11–24.
- Yildirim, K., Rasinski, T., 2014. Reading fluency beyond English: investigations into reading fluency in Turkish elementary students. *Int. Electr. J. Elem. Educ.* 7, 97–106.
- Yildirim, K., Cetinkaya, F.C., Ates, S., Kaya, D., Rasinski, T., 2020. Testing the KAPS model of reading comprehension in a Turkish elementary school context from low socio-economic background. *Educ. Sci.* 10, 90.
- Yildirim, K., Rasinski, T., Kaya, D., 2017. Fluency and comprehension of expository texts in Turkish students in grades through eight. *Educ. Sci.* 42, 87–98.
- Yildirim, K., Rasinski, T., Kaya, D., 2019. Fluency and comprehension of narrative texts in Turkish students in grades 4 through 8. *Education 3-13* 47, 348–357.
- Young, A.R., Bowers, P.G., Mackinnon, G.E., 1996. Effects of prosodic modeling and repeated reading on poor readers' fluency and comprehension. *Appl. Psycholinguist.* 17, 59–84.
- Young, C., Rasinski, T., 2009. Implementing readers theatre as an approach to classroom fluency instruction. *Read. Teach.* 63, 4–13.
- Young, C., Rasinski, T., 2013. Student-produced movies as a medium for literacy development. *Read. Teach.* 66, 670–675.
- Young, C., Rasinski, T., 2018. Readers Theatre: effects on word recognition automaticity and reading prosody. *J. Res. Read.* 41, 475–485.
- Young, C., Rasinski, T., Landreth, S., 2021. Principles of effective reading fluency instruction. In: Parsons, S., Vaughn, M. (Eds.), *Principles of Effective Literacy Instruction*. Guilford Press, New York, NY, pp. 68–78.
- Young, C., Rasinski, T., 2022. Tiered fluency instruction: from science of reading to effective reading fluency instruction and interventions. KDP. <https://amzn.to/3uXHaQ8>.
- Zhao, J., Kwok, R.K., Liu, M., Liu, H., Huang, C., 2017. Underlying skills of oral and silent reading fluency in Chinese: perspective of visual rapid processing. *Front. Psychol.* 7, 1–11.
- Zutell, J., Rasinski, T.V., 1991. Training teachers to attend to their students' oral reading fluency. *Theory Into Pract.* 30, 211–217.

Expanding meaningfulness for struggling readers

Gay Ivey and Karole-Ann Friddle, University of North Carolina-Greensboro, Greensboro, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Reading as a transformative human experience	399
Meaningful reading for struggling readers	399
Talk about books, expanding relationships, and shifting identities	400
Emotional and behavioral self-regulation	400
Social imagination	400
Intellectual engagement	401
Agency over future life narratives	401
Happiness	402
Increasing reading competence in meaningful, engaged reading experiences	402
Reading strategically	402
Reading minds	403
Comprehending together	403
Arranging for reading to be meaningful for struggling readers	403
Arranging for autonomy	403
Arranging for a sense of competence	404
Arranging for a sense of relatedness	404
Conclusions	405
References	405
Young adult literature references	406

In his eighth grade year in a United States school, Calvin enthusiastically chose to read *The Killer Angels* (Shaara, 1974), a fictionalized account of the Battle of Gettysburg in the American Civil War. In reflecting on the narrative, a complex interplay of shifting perspectives from opposing sides of the war, Calvin marveled, “I know that just one person wrote this, but it just [doesn’t] seem that way.” Having had access to the minds, motives and emotions of multiple characters through his reading, he concluded, “Everybody was in the wrong [in the war], both sides.” The “everybody” included members of the Confederacy, whose legacy, more than 150 years later, remains a stronghold in some communities in the American South, with lingering ties to the institution of slavery and to the perpetuation of White supremacy. Calvin’s experience is remarkable not just because this book won the Pulitzer Prize, not just because Calvin had historically scored poorly on mandated, standardized tests of reading achievement, and not just because it was not required reading. Importantly, the culture of the Confederacy permeated Calvin’s immediate and extended family, some of whom participated in re-enactments of Civil War events, making his implied rebuke of this legacy as remarkable as having made sense of such a complex book.

Calvin’s story exemplifies tensions in discourse on struggling readers as they are represented in research over the past several decades. On the one hand, attention to children like Calvin who experience difficulty with reading in school has centered historically on the questions “What skills are students lacking?” and “What are the most effective ways to teach those skills?” Inherent in these questions and corresponding research are the beliefs that there is vital knowledge about reading that all children must gain, that identifying what children are missing and filling the holes will edge children closer to competence, and that the way to gain essential skills is through particular kinds of instruction delivered in specific sorts of arrangements. From this view, what children know and do not know is defined by external criteria and can be accurately measured by an adult or a computer, and what children do is wait passively to be filled with vital skills once their areas of need are determined.

On the other hand, and certainly less visible in educational policies, is the notion that children and young adults considered “struggling readers” might be viewed otherwise in contexts where they participate in literate activities outside of the academic norm. This perspective suggests the identity of struggling readers is culturally and socially constructed, by both school and the larger society (Alvermann, 2001; Hall, 2009; Kontovourki, 2012; Learned, 2016). Accordingly, negative identities around reading can be reversed when instruction honors students’ everyday literacies, invoking resources from their personal, cultural, linguistic, and social lives (Learned, 2018; Skerrett, 2012; Skerrett and Bomer, 2011).

For Calvin, both of these perspectives are relevant. We suspect that he did, indeed, struggle while reading *The Killer Angels* in multiple ways (e.g., reading fluency, vocabulary), and that if he had to take a test of how well he understood the book based on standard criteria, he might fail, and that this would be the case even in texts that better matched his grade level in terms of difficulty. We also believe that this was a rich and positive experience for Calvin, and that in this context, drawing from his cultural knowledge, he is a motivated, sophisticated reader. To ignore either of these realities would likely hamper his development, both as a reader and a human being.

Across age levels and gender, reading is positively associated with socio-emotional dimensions of a person's life (Batini et al., 2021). We propose that for school-aged struggling readers, like all readers, the development of competence in reading need not—indeed, should not—be separate from the development of their emotional, relational, and intellectual lives. When Calvin's teacher arranged for him to take up a personally meaningful literacy experience, he elevated himself intellectually and morally. Expansions of these dimensions of his development enabled him to navigate a difficult text, and likely future complexities in reading. In turn, a growing sense of competence will make it more possible for him to pursue questions of personal, social, and cultural significance. Central to this chain of consequences, though, is that Calvin read to shape the person he was becoming, not to improve his reading test scores.

We situate our perspective on struggling readers in the context of *actual reading* rather than instruction about or in service of reading. Although there is research documenting the role of the act of reading, particularly in large volume, to reading improvement (Allington and McGill-Franzen, 2021), we would draw attention instead to the both the activity within engaged reading and the consequences of engagement in reading that include, but also transcend, conventional achievement, such as the expansion of children's socioemotional capacities, their self-regulation, and their moral agency (Ivey and Johnston, 2013). Engagement in narratives, including the processing that happens while in a book and after the book is closed, along with ongoing mental conversations with characters and with others outside of books, are fertile contexts for growth across the breadth of human development, as well as growth in reading.

Reading as a transformative human experience

It likely the case that, particularly in school, struggling readers do not experience reading as a generative activity that enriches their intellectual lives and overall well-being. In contrast, it is a source of emotional stress (Kos, 1991) and a trigger for diminishing self-worth (Skaalvik, 2004). For instance, they may be preoccupied by social comparisons in terms of competence and assigned tasks (Skaalvik, 1993), as well as by strategies to hide from peers their difficulties with reading (Skaalvik, 2004). Even those students who demonstrate a degree of resilience and agency are often not able to overcome their positioning or to feel valued and socially included in classrooms (Enriquez, 2014; Learned, 2016).

Not surprisingly, reading difficulties in children and young adults are associated with negative behaviors in school (Morgan et al., 2008), high levels of anxiety (Casey et al., 1992; Willcutt and Pennington, 2000), low self-esteem and self-concept (Boyes et al., 2018; Humphrey and Mullins, 2004), somatic complaints (Arnold et al., 2005), and social problems (Goldston et al., 2007). Elementary grades struggling readers in one study were three times more prone to depression than their peers (Maughn et al., 2003), and even younger children who do not read well feel angry, sad, lonely, distractible, and unpopular (Morgan et al., 2012). Difficulties with reading have been linked not only to dropping out of school, but also to suicide (Daniel et al., 2006).

We take the position, though, that these relationships between reading difficulties and overall wellbeing are exacerbated by the instructional and discursive environment students experience. We propose that both low reading ability and poor mental and social health in students can be reversed by first identifying why a child or young adult might want to read in the first place.

Meaningful reading for struggling readers

We suspect that few children who are identified in school as struggling readers have the opportunities to experience reading meaningfully, as in the example of Calvin we described earlier. Calvin was a participant in research (Ivey and Johnston, 2013, 2015) conducted in a school in which all eighth-grade English teachers decided to forgo conventional reading instruction that focused on the building of skills and strategies in favor of a curriculum driven by students' own reading interests. In short, they made available to students a vast array of high-interest young adult books, allowed the students to choose what to read, set aside time in class to read and talk about books, and attached no assignments to students' reading. The teachers had originally made that decision from a perspective that attending to student needs would motivate them to read, with an eye toward improving reading achievement, a decision that would pay off on both counts. Students from these classrooms interviewed at the end of one school year indicated that they read substantially more and that they were engaged in their reading, rather than merely compliant, and that for many, this was a notable change. As one student put it, "Last year I didn't read at all. The only book I read is *The Outsiders* (Hinton, 1967) ... but now I read whenever I get the chance" (Ivey and Johnston, 2013, p. 261). The engagement though, was not merely behavioral, but also emotional, social, and intellectual. One lower achieving reader gave an example:

Before I Fall (Oliver, 2010) is one of those, like the first book I've ever truly gotten into, like one of those books where you could really picture yourself in that scene. Like you feel like you're with them ... I used to hate to read. I didn't really read at all. I'd sit there and turn the pages when the teacher looked at me, and that was pretty much it (p. 261).

As teachers had hoped, students' standardized test scores also rose, with more students passing compared to the previous school year, while passing rates for all students in the state compared to the previous year remained constant. This was the case also for important subcategories of students, including economically disadvantaged students, non-White students, and boys.

Although students in these classrooms were aware of how their reading skills increased, and also how they expanded their reading repertoires in terms of genre, read more critically, knew more about books, and had become better writers due to their reading, these were not the outcomes about which they were most enthusiastic and most forthcoming. They were clear that the most significant consequences for them had to do with how reading and the activity around reading had enriched their broader lives. For struggling readers, this meant not only gains in reading competence, but also, as discussed below, the cultivation of healthier mental and social worlds.

Talk about books, expanding relationships, and shifting identities

Students who had never read voluntarily, much less talked about books, reported widespread conversations through and about books—both in and out of school, with peers, teachers, and family members. These interactions helped to shift some relationships. For instance, books became tools for children to open up lines of communications with their parents, as one student explained, “Me and my mom have more things to talk about now. When she asks me about school, I can say, ‘I read a book,’ give her details about it” (Ivey and Johnston, 2013, p. 262). Other students reported making new friends over books. This was especially meaningful for struggling readers in the classes, who expanded their social networks to include peers with whom they might not typically interact. An academically advanced student noted, for instance, talking regularly with Roberto, who historically struggled in reading: “I never would have thought he was a bookworm.” She added that from her view, and due to all the reading and conversations, that “everyone has gotten so much smarter.”

The conversations, along with shifts in relationships and identities, were made possible for struggling readers because rather than having to read a book assigned to all students regardless of the difficulty level or personal relevance, or having a book selected from an intervention package, according to their “level,” per an assessment and readability formula, they chose books based on how personally meaningful the experience would be for them. Like Calvin, who we mentioned earlier, Kadeem was a struggling reader who was most engaged in books that challenged his own perspective about cultural phenomena he knew well. He read several memoirs of reformed gang members, convinced classmates to read some of the same books, and was frequently mentioned by more capable readers as a source of help when they encountered difficulty while reading. Individual differences in reading typically focus on varying degrees of competence and the comparisons between so-called good and poor readers. In contrast, the individual perspectives and everyday experiences brought to bear on the conversations about books in classrooms that prioritize meaningfulness, including those students who struggle with reading, are valuable resources for dialogic engagement, thus deeper understanding not just of books, but of selves and other (Johnston and Ivey, 2016). In such a situation, the lines between struggling and proficient reader become blurred.

Emotional and behavioral self-regulation

Students reported that the experience of reading itself had a calming effect. One girl disclosed that before eighth grade she “got sent to the office so many times” for behavioral issues. In contrast, she said, “But then when I started reading, like, I don’t get involved in so much drama anymore, so I just keep to myself, so I don’t get in trouble.” Similarly, a male classmate explained, “[Books] give me something like to do when I’m [fidgety]. I’ll just turn to the next page and read, instead of getting in trouble.”

Social imagination

For certain, the engaging books provided a useful distraction from anxiety and worries that might have otherwise resulted in trouble for students. But the experiences with the books also made them think differently about the impact of their words and actions on others. The social worlds of narratives, and particularly the mental lives of characters who have captured the interest and investment of readers, are fertile sites for understanding the experiences of others (Mar and Oatley, 2008).

Eighth-grade students in the Ivey and Johnston (2013, 2015) studies routinely described taking up residence in the minds of characters and developing increased capacity to imagine the thinking and emotions of others, as well as the logic behind social interactions, both inside and outside of books. For some students struggling academically and socially, this opportunity made them feel less alone, but at the same time, able to reflect on their experiences, as in this example from a girl whose family was also experiencing economic struggle:

Tyrell (Booth, 2010) it made me think ... cause I was kind of raised in that kind of lifestyle ... I looked at it like it was kind of hard, you know that? But seeing how it was for somebody else, and seeing how it was for them to live it every day. But to read about someone’s life like that, it made me think ... even though I didn’t have too many things when I was little, I should still be appreciative that I had my parents in my life.

Access to characters’ minds who were both like and unlike themselves also forced them to reckon with their own social responsibility. A common theme was gaining a heightened awareness of the power of words, in some cases, specific to actual relationships in their own lives, for instance, “In part of the book that I was read, there was this girl, and her mom was all rude to her, and ... it

kind of had me feeling [really] bad, cause I was rude to my aunt, and my situation could have been worse.” Students indicated a growing propensity to anticipate the consequences of what they said, and to regulate accordingly, as one student explained:

Like, you never know if what you say to someone will be offensive. Like one little word could change somebody’s point of view over a second. They could take it the wrong way. So, I think that [*Thirteen Reasons Why* (Asher, 2007)] taught me a whole lot. Don’t say nothing that you think someone else might take in a different way.

In related research, Schrijvers et al. (2019) demonstrated that students who were asked to explore the inner psychological worlds of characters in fiction, then discuss their reading experiences with peers, reported improved insights into others as compared with students who experienced traditional literature instruction.

Intellectual engagement

Similar to Calvin’s experience with *The Killer Angels*, other eighth grade struggling readers learned that critical to knowing something well is realizing there are new perspectives to uncover. Having the opportunity to read complex books with multiple, shifting narrators—books not typically assigned in skills-oriented reading intervention programs—certainly accelerated this possibility for many students. As one student put it, “Now I like books like that the give more than one view of a story ... I don’t follow one view ... I can make up my own if I read two.” These students came to believe that their understandings were limited unless they sought out other ways of thinking. Most frequently, this meant recruiting others to read for the purpose of getting their thoughts on the moral complexities of characters’ decisions. One student suggested that getting another’s perspective was “automatically” her response to finishing a book. The purpose was not to ascertain the basic plot or main idea, but instead to further problematize the issues. Students regularly gravitated toward books they called “disturbing” due to moral ambiguities and uncertainties. A good example of their growing propensities to intellectualize their understandings of the narratives was in a small group conversation about *Living Dead Girl* (Scott, 2008), in which a young girl is abducted, abused, and held captive by an adult male. While at first the talk centered on the emotional pain of the main character, the group began to question, explicitly, what had caused her kidnapper to commit such violent acts, for example, the possibility that he had been abused in his past. The point was not to excuse the behavior, but to see all sides. Students reported that reading these sorts of books and talking to peers in this way helped them to become generally more open-minded.

Agency over future life narratives

Children and young adults, like all other humans, have the capacity to consider their own lives and exercise some agency about who they will become and how their lives will go (Sugarman and Martin, 2011). Through experiencing meaningful narratives and the social and moral dilemmas of the characters within them, and by recruiting others for conversations about these issues, the eighth-grade students in the engagement-focused classrooms began to see that they could actively structure their futures. For instance, Mason, who had long struggled with school reading, started the school year with the goal merely to “try to get on a better reading level and to read as much books as I can.” By mid-year, though, he was choosing books with characters and situations that mirrored his own, along with a group of other boys who struggled with reading and who, like Mason, had become involved in gangs. They routinely read and talked through and about stories and memoirs that brought to light the consequences of gang involvement for a person’s life trajectory. Mason reflected within a small group of these boys:

There’s a lesson in life from *Life in Prison* (Williams, 1998) and *If I Grow Up* (Strasser, 2008). There are lots of opportunities, but you have to choose the right one. Like Tookie Williams, if he wasn’t in prison, he might not have been executed. Then Deshawn [from *If I Grow Up*], he had the opportunity to get an education, but he did something different.

Kadeem, a friend of Mason’s, had a similar conversation with another group of boys, about *Yummy: The Last Days of a Southside Shorty* (Neri, 2010), a graphic novel chronicling real events in Chicago in the 1990s that resulted in the death of an 11-year-old boy with gang connections. Kadeem, who had disclosed that he was frequently left home alone, expressed sympathy for the Yummy: “He didn’t have a mom or a dad or anything. He lived with his grandma, and his grandma had all [those] kids. Nobody didn’t pay any attention to him.” When Rico lamented, “[Yummy] was just a kid” when he died, Kadeem responded, “If Yummy had lived, I think he would have changed his life.” The possibilities of both a sense of control over your life, as well as redemption, were common threads in students’ talk about and reflections on reading.

The books themselves, including both characters and dialog, also served as touchstones for some students and how they wished to design their futures. For, instance, In debriefing with peers on a conversation he had recently with his parents about academics and his future, Akeem quoted a relevant line from *The Last Shot* (Feinstein, 2005), months after he had read it: “[The character said], ‘I’m going to be the only [family member] to ever graduate and go to college’ ... And he does get good grades. He said, ‘I made my mom happy with me.’ And then she said, ‘You keep it up.’”

Reading and talking allowed the students to role-play in a safe space, taking on the situations characters face, mulling over complex decisions, experiencing the possible consequences ahead of real-life time, and adjusting for the future accordingly. As one student put it, “We’ve been reading things about people our age and the things they go through and the choices they make ... It kind of helped, because we go through those things, too, and it helps us learn about ourselves.”

Happiness

Students in the Ivey and Johnston (2013) study frequently described their experiences as “fun,” but their choices to read sometimes disturbing narratives do not parallel what is often called pleasure reading. At the end of the school year, for instance, one struggling reader who consistently read books about characters enduring psychological, emotional, and sexual trauma confessed she ended the year in a different state of mind than when she began: “I’m happier, but I can’t explain it.” For this student and others, this shift was likely connected to becoming engaged in the social worlds of books and in meaningful conversations with others. Some students reported that books and conversation had helped them overcome grief, feelings of loss, loneliness, and others indicated they had developed feelings of gratitude, optimism for the future, and intellectual competence.

It is difficult to how imagine common instructional interventions for struggling readers—focused on the building of reading skills and strategies—might lead to, even indirectly and over time, this level of personal satisfaction and overall wellbeing. How, though, can children and young adults who are experiencing difficulty with literacy, though, grow their competence within reading that matters to them? We will suggest below that it is this personal significance that catalyzes growth.

Increasing reading competence in meaningful, engaged reading experiences

The development of reading proficiency has been studied primarily in contexts where students have limited autonomy and agency, where particular skills and strategies to be learned are chosen for them, practiced in texts that have no relevance to their lives, with no connections to how they might use them to change their situations either inside or outside of school. This is not to question the value of experimental interventions, but instead to make the point that research offers much less at the moment on how literacy might develop when children and young adults have control over their own learning and when it is possible for them to explore personally and socially meaningful activities through reading.

There is enough, however, to provide a clear window into the possibilities. We can start with research exploring the degree to which particular variables influence the motivation to read. A meta-analysis by Guthrie and Humenick (2004) demonstrated that having choice, interesting texts available, and a mastery-oriented culture were strongly correlated with wanting to read, and that the opportunity to collaborate was also moderately influential. But these same variables had similarly robust effects on reading comprehension. Examining the processes of reading when a reader experiences a sense of autonomy and relevance—when they are engaged rather than merely compliant—offers a glimpse into how and what kinds of proficiencies children are building.

Reading strategically

Engaged reading itself is strategic reading (Guthrie et al., 2012). When children’s reading choices are based on how meaningful they find a text, rather than how easy or hard it is to read, they will do whatever is necessary to solve problems they encounter in their reading, thus expanding their strategic knowledge (Ivey, 2019). In many cases, these strategic moves mirror the same cognitive strategies demonstrated through research on expert readers and that comprise popular reading interventions, such as self-questioning, rereading, inferring, and summarizing. For instance, a less-experienced reader engaged in *Glimpse* (Williams, 2012) shared that while reading she “had a lot of questions.” What helped, she explained, was to “write [the questions] down on sticky notes, then I go back later and see if I can answer them ... after I’ve read more.” What makes this book puzzling is its unreliable narrator, the younger sister of a psychologically disturbed teenager, whose naivete limits her interpretation of key events. The reader is required to make inferential leaps based on their own life experiences and world knowledge. This student did not act strategically merely to ensure comprehension of the book. She was most interested in figuring out what would make the older sister of the narrator, who was around her own age, attempt suicide multiple times. Weeks after finishing the book, she told a teacher who had also read the book, “I always suspected some kind of abuse [from the mother].” In synchrony with the realization that reading might be a tool for thinking about issues arising in her everyday life and relationships, though, this same student, who described having been in remedial reading classes because of low achievement, reported also her awareness of her growing competence as a reader.

The *Glimpse* example also brings life to the finding that struggling readers might not avoid difficult books if they are meaningful, because high interest in a text buffers the potential negative effects of text difficulty (Fulmer et al., 2015; Fulmer and Frijters, 2011). As a result, students are more willing to try to read complex books, more likely to persevere, and also more inclined to create strategies specific to books, as this student described:

I go through the book and see what’s in it, to see if it has parts or weird titles of chapters. And I looked through this book, and it has like 6 parts, and the titles of the chapters are lyrics from songs. I do that with most of my books. I write the parts down on a piece of paper and use it as I read.

Ivey (2014, p. 68).

Reading minds

Earlier we described how readers engaged in narratives, particularly fiction, experience vicariously the minds and emotions of others as they access characters' thoughts. From the perspective of meaningfulness, social imagination has strong relational benefits for children and adults, as they develop an awareness of the social consequences of their own language and actions. This seems particularly true when readers are deeply immersed in the social world of the narrative and when they feel connected to characters (Carpenter et al., 2018). The capacity to imagine the thoughts and experiences of others, particularly through books, is not limited to older students or the highest achieving readers. Even young children not yet able to decode words can recognize emotions and infer characters' thoughts in picture books in tandem with the development of important social skills (Lysaker, 2019a).

From a reading development perspective, there are different, but parallel benefits. There is a growing body of work that provides consistent evidence for how theory of mind, a related phenomenon, accounts for individual differences in reading comprehension (e.g., Atkinson et al., 2017; Dore et al., 2018; Guajardo and Cartwright, 2016; Lecce et al., 2021). Lysaker (2019b) pointed out that the relational aspects of reading ought to help both redefine the properties of reading and the expansiveness of reading as a human activity. She noted, for example, "Relational experiences become important resources for comprehending and composing and can be seen as valuable contributions to comprehending activity, much like the notion of background knowledge is currently viewed" (p. 79). By the same token, she suggested, that deep comprehension would be indicated by "meaning making like emotional involvement, imagination, moral work, and relational insight" (p. 79).

Comprehending together

Engaged readers in the company of other engaged readers view each other as resources for expanding meaning (Ivey and Johnston, 2015). They often choose to read because they want to participate in conversations with friends who appear engaged, as one seventh-grade struggling reader explained: "I watch them while they read and I see how much they enjoy their book, and I wonder what it's about. I like reading after them." In such groups, it is not uncommon to take confusion to a friend, as one student described:

Like I talk to Magdalena about what happens in the books. The books she hasn't read, I give her a little bit of [information] about it. She can read it when she's done with her book. And some parts I have to get off my mind and talk through with her. And she helps me get through parts I'm stuck at. I'm either confused or it's bugging me. Like, then, I'll go back to the book, and I figure out what has happened, then I go back and tell her about it (Johnston and Ivey, 2016, p. 11).

Even students who are broadly viewed as less-able readers, when they have had meaningful experiences with particular books, can offer assistance to peers, even those normally viewed as more able. In short, reading problems are not viewed as individual problems, but as challenges to be sorted out within the community.

Unfortunately, conventional classrooms designed for struggling readers have probably not yet taken the dialogic turn in reading comprehension instruction (Wilkinson and Son, 2011) that is evident in recent advances in research. But examples from more atypical classrooms prioritizing socially engaged reading (e.g., Ivey and Johnston, 2013) suggest that struggling readers not only learn to comprehend text better, but also, together, make sense of selves, other, and their relational, epistemological, and moral lives. For instance, in a conversation about a book in which a character is the victim of racial slurs, a Mexican American student announced he might become violent if he were similarly taunted. His classmate, an African American student, suggested that he might take up a more peaceful response, citing an example shown by a character in a different book the boys had read previously. Engaged reading communities such as the one occupied by these boys, also demonstrate, through examples like this, that when students are engaged, we might view deep comprehension as arising not just from a reader-text relationship, but instead, from a network that includes self, text, and other.

Arranging for reading to be meaningful for struggling readers

Altering the academic, personal, and social trajectories for children and young adults who experience the most difficulty learning to read, regardless of whether the contributing factors are differences within individuals or due to insufficient prior instruction, requires making it possible, first, for them to realize and pursue their own motivations around reading. This is most likely when students experience, simultaneously, a sense of autonomy, a sense of competence, and a sense of relatedness (Ryan and Deci, 2000). Children are also more emotionally aroused, thus more likely to learn, when their needs of a sense of autonomy, competence, and relatedness are met (Streb et al., 2015). In addition, they are less stressed (Reeve and Tseng, 2011), and are also more likely to develop healthier identities (La Guardia, 2009).

Arranging for autonomy

Readers will experience a sense of autonomy when their reading is self-initiated and self-directed, in other words, when they decide when, what, and why to read. For the least experienced readers, addressing this need is frequently oversimplified. For students who

have been engaged in reading on limited occasions or who have faced struggle without the payoff of meaningfulness, gaining a sense of autonomy is a process. Although having choices and interesting reading materials available is important to students' motivation to read (Ivey and Broadus, 2001), those who have read less might not yet know the range of personally and socially significant encounters texts have to offer them, and have not had many opportunities to choose and discuss books (Merga, 2018). For instance, Roland, a seventh-grade boy who was classified in school as learning disabled, told his teachers that he wanted books on vintage automobiles, one of his personal interests. Although the photographs had temporary appeal, the books had little new information to offer, and he quickly grew bored. What eventually engaged him was a novel he heard a classmate discussing, in which the main character was suffering through his mother's illness. This was a topic that was relevant in Roland's own life at that time, but that he could not have articulated or imagined as a topic of book when his teachers asked what he wanted to read.

Roland's example also highlights additional ways to productively complicate the notion of autonomy for struggling readers, starting with the importance of a vibrant social context. Interests are often shaped through social interaction (Bergin, 2016), and struggling readers can often identify what they find meaningful from hearing peers talk about the texts they are reading. Related, it is sometimes the desire to become involved in the rich conversations around text, to understand others, or to work through a personal problem, rather than a book itself, that channels the motivation to read. One group of eighth-grade students, for example, after hearing a classmate tearfully describe how a book with a gay character made her think about strained relationships with her own family, decided to read that same book in order to in order to better understand their classmate's life (Ivey and Johnston, 2015).

Importantly, a sense of autonomy will be undermined when assignments and tests are linked to student learning experiences (Assor et al., 2002), a situation frequently faced by struggling readers, particularly when teachers are concerned about monitoring progress in their reading development. Open literacy activities, such as self-selected reading without having to write a personal response, completing a worksheet, or answering comprehension, are not only more conducive to motivation (Turner, 1995), but also to making possible a broader range of outcomes that are relevant to individuals, rather than uniform, and also that touch the breadth of human development rather than reading growth alone.

Arranging for a sense of competence

There are common practices that certainly undermine a sense of competence for less-experienced readers, starting with common reading assignments. Requiring all students in a classroom—regardless of individual differences in reading fluency, personal interest, and knowledge about the text topic—is especially problematic. In addition to shattering a sense of autonomy, such situations accentuate differences in competence and perpetuate feelings of inadequacy for struggling readers. Equally troublesome are assignments and teacher-generated questions that demand one particular correct response and that fail to draw from students' cultural knowledge and everyday lives. Having a range of reading materials, varying in topic, genre, and difficulty level, is fundamentally necessary for students to experience reading as productive and manageable.

Struggling readers whose goal is to read in order to grapple with the complexities of being human, though, choose texts on how interesting those texts are rather than how easy or difficult. However, it is also quite possible to experience a sense of competence in contexts where encountering difficulty while reading is viewed as normal, and even expected, for all readers, regardless of achievement levels. Dialogic classrooms built on the idea that confusion is a normal property of reading are contexts with minimal risk for less experienced readers to seek assistance.

Johnston and Ivey (2016) described a classroom that commonly featured intertextual conversations. Because students came to these discussions having read different texts, rather than one text assigned for all, struggling readers were able to make unique contributions. For example, when the teacher was reading aloud to the students a book in which a troubled teenaged character needed to confide in a trustworthy adult, student Deandra asked, "Would that be like the person that the counselor was to Sonia [a character from a different book]?" (p. 214). The teacher then expanded Deandra's questions by inviting other students to describe similar relationships from books they had read. After several examples, she said to the class, "I love how Deandre made us think about the word for that—confidante" (p. 214), thus rendering Deandra an expert and fostering in her a sense of competence, but also increasing the likelihood that her peers might see her that way. This is much less likely to happen in a classroom where the object of reading is to get the main idea of one particular text.

Arranging for a sense of relatedness

With respect to her classroom, struggling reader Anna stated:

I think we've become a lot closer because we have the same taste in books, and we know we're on the same page. And our thoughts, when one person finishes off the other person starts and picks up where the other person left off.

Engaged readers, like most human beings, are fundamentally social creatures. The kind of community Anna describes is not likely to be found in reading interventions for struggling readers that rely primarily on interacting with a computer or in instructional arrangements in which children must compete and compare rather than collaborate. To be sure, it is unlikely to be found in classrooms where students finding little meaningful in the reading to talk about.

The kinship with classmates and teachers that Anna described was cultivated because of the limitless opportunities to talk through and about books in her classroom. As one student in Anna's class put it, "[Books] help us communicate in a way. It kind of helps you make new friends almost. People I would never have talked to last year, we've become like friends through books." Realizing the need for a sense of belonging, the teacher of this classroom intentionally turns children toward each other. One of the students gave this example: "Tell Tyra about [your book]"—[the teacher] said that to Charity, and Charity came over and explained the book to me and I was like, 'I really want to read it,' and that's how we became friends, through that book" (Johnston et al., 2011, p. 237).

It is also the quality of the talk that matters. A sense of relatedness is made possible in a reading classroom by having materials children can relate to. When children and young adults have personally meaningful texts they want to discuss, they open the possibility of revealing information about themselves in the process, deepening relationships, as one engaged reader put it, "I don't know how to put it, but it's just that talk, you get to talk between each other, and that makes you feel you can trust that person more" (Ivey, 2019, p. 12). The openness and encouragement to talk, which is likely in stark contrast to many traditional reading programs designed for struggling readers, changes the relational properties of the classroom, and thus individuals' engagement in learning.

Conclusions

We have proposed that a productive way forward for how we perceive and teach struggling readers ought to focus on the breadth of children's and young adults' development, not just their literacy skills. We suggested that when meaningful reading within a socially rich learning environment is prioritized, struggling readers will be motivated to expand their knowledge about reading, but in the service of expanding the richness of their personal and social lives. Doing so will require a reconceptualization of so-called reading interventions, away from skill-oriented programs wherein children practice their reading, so to speak. In meaningful interventions, reading would be resituated not as the goal of instruction now, in hopes of academic and economic survival later, but as a tool children use for building happier, healthier lives in the present.

Our proposal does not negate the role of explicit instruction in the conventional skills of reading, although we suspect that, as our examples have demonstrated, much less would be required, since engaged reading involves self-teaching and social problem-solving. Curricula driven by a fixed scope and sequence of skills would be rendered less relevant, because engaged readers learn skills in the order in which they become meaningful to them rather than in the order prescribed by an instructional program.

Most of our examples have centered on early adolescence. We are cautious when we use the phrase "struggling readers" in general, but for young children, who we expect to be still developing before and throughout the primary grades, we are especially reserved. We do suspect, though, that many reading difficulties might be prevented, if early instruction was designed to shape in positive ways children's socio-emotional, intellectual and literate lives simultaneously. The research linking literacy practices and socio-emotional growth includes even the youngest of children (Batini et al., 2021). Research that fleshes out these relationships might also attend to how children learn code-related information, such as phonics and spelling, with and through meaningful activity and when they feel a sense of agency. Writing is seldom invoked in early reading interventions, but children find it meaningful when it is centered in their social worlds and grounded in the cultures of childhood (Dyson, 1993) rather than the rules and conventions of the adult writing world. There is already some evidence that children's invented spellings, which facilitate children's meaning-making on their own terms as they compose, accelerates children's abilities to read and spell words (Ouellette and Sénéchal, 2017).

For decades, struggling readers have been famously associated with the Matthew effects phenomenon (Stanovich, 1986) caused by individual differences in the acquisition of reading competence. In short, children who learn to read early, read more, catalyzing a reciprocal relationship between their reading skill development and their cognitive growth. For children who experience learning to read, the opposite occurs, thus setting up a rich-get-richer and poor-get-poorer scenario separating competent and less competent readers. We would suggest that limited opportunities for struggling readers to reap personal, relational, and intellectual benefits of meaningful reading adds another dimension to problem. A reversal of both problems, and possibly a bounty of growth across the breadth of human development that we have yet to fully know, must start with the reading and social activity that really matters to students in their everyday lives.

References

- Allington, R.L., McGill-Franzen, A., 2021. Reading volume and reading achievement: a review of recent research. *Read. Res. Q.* 56, 231–238.
- Alvermann, D.E., 2001. Reading adolescents' identities: looking back to see ahead. *J. Adolesc. Adult Lit.* 44, 676–690.
- Arnold, E.M., Goldston, D.B., Walsh, A.K., Reboussin, B.A., Daniel, S.S., et al., 2005. Severity of emotional and behavioral problems among poor and typical readers. *J. Abnorm. Child Psychol.* 33, 205–217.
- Assor, A., Kaplan, H., Roth, G., 2002. Choice is good, but relevance is excellent: autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *Br. J. Educ. Psychol.* 72, 261–278.
- Atkinson, L., Slade, L., Powell, D., Levy, J.P., 2017. Theory of mind in emerging reading comprehension: a longitudinal study of early indirect and direct effects. *J. Exp. Child Psychol.* 164, 225–238.

- Batini, F., Luperini, V., Cei, E., Izzo, D., Toti, G., 2021. The association between reading and emotional development: a systematic review. *J. Educ. Train. Stud.* 9, 12–48.
- Bergin, D.A., 2016. Social influences on interest. *Educ. Psychol.* 51, 7–22.
- Boyes, M.E., Tebbutt, B., Preece, K.A., Badcock, N.A., 2018. Relationships between reading ability and child mental health: moderating effects of self-esteem. *Austral. Psychol.* 53, 125–133.
- Carpenter, J.M., Green, M.C., Fitzgerald, K., 2018. Mind-reading motivation: individual differences in desire to perspective-take influence narrative processing. *Sci. Study Lit.* 8, 211–238.
- Casey, R., Levy, S.E., Brown, K., Brooks-Gunn, J., 1992. Impaired emotional health in children with mild reading disability. *J. Dev. Behav. Pediatr.* 13, 256–260.
- Daniel, S.S., Walsh, A.K., Goldston, D.B., Arnold, E.M., Reboussin, B.A., et al., 2006. Suicidality, school dropout, and reading problems among adolescents. *J. Learn. Disabil.* 39, 507–514.
- Dore, R.A., Amend, S.J., Golinkoff, R.M., Hirsh-Pasek, K., 2018. Theory of mind: a hidden factor in reading comprehension? *Educ. Psychol. Rev.* 30, 1067–1089.
- Dyson, A.H., 1993. *Social Worlds of Children: Learning to Write in an Urban Primary School*. Teachers College Press.
- Enriquez, G., 2014. Embodiments of “struggle”: the melancholy, loss, and interactions with print of two “struggling readers”. *Anthropol. Educ. Q.* 45, 105–122.
- Fulmer, S.M., Frijters, J.C., 2011. Motivation during an excessively challenging reading task: the buffering role of relative topic interest. *J. Exp. Educ.* 79, 185–208.
- Fulmer, S.M., D’Mello, S.K., Strain, A., Graesser, A.C., 2015. Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemp. Educ. Psychol.* 41, 98–110.
- Goldston, D.B., Walsh, A., Arnold, A.M., Reboussin, B., Daniel, S.S., et al., 2007. Reading problems, psychiatric disorders, and functional impairment from mid- to late-adolescence. *J. Am. Acad. Child Adolesc. Psychiatr.* 46, 25–32.
- Guajardo, N.R., Cartwright, K.B., 2016. The contribution of theory of mind, counterfactual reasoning, and executive function to pre-readers’ language comprehension and later reading awareness and comprehension in elementary school. *J. Exp. Child Psychol.* 144, 27–45.
- Guthrie, J.T., Humenick, N.M., 2004. Motivating students to read: evidence for classroom practice that increase motivation and achievement. In: McCardle, P., Chhabra, V. (Eds.), *The Voice of Evidence in Reading Research*. Paul Brookes Publishing, pp. 329–354.
- Guthrie, J.T., Wigfield, A., You, W., 2012. Instructional contexts for engagement and achievement in reading. In: Christenson, S., Wylie, C., Reschly, A. (Eds.), *Handbook of Research on Student Engagement*. Springer, pp. 675–694.
- Hall, L., 2009. Struggling reader, struggling teacher: an examination of student-teacher transactions with reading instruction and text in social studies. *Res. Teach. Engl.* 43, 286–309.
- Humphrey, N., Mullins, P.M., 2004. Self-concept and self-esteem in developmental dyslexia. *J. Res. Spec. Educ. Needs* 2, 1–13.
- Ivey, G., Broadus, K., 2001. “Just plain reading”: a survey of what makes students want to read in middle school classrooms. *Read. Res. Q.* 36, 350–377.
- Ivey, G., Johnston, P.H., 2013. Engagement with young adult literature: outcomes and processes. *Read. Res. Q.* 48, 255–275.
- Ivey, G., Johnston, P.H., 2015. Engaged reading as a collaborative transformative practice. *J. Literacy Res.* 47, 297–327.
- Ivey, G., 2014. The social side of engaged reading for young adolescents. *Read. Teach.* 68, 165–171.
- Ivey, G., 2019. Engaging possibilities: reinvigorating the call for research on reading. *Literacy Res.* 68, 25–44.
- Johnston, P., Ivey, G., 2016. Discursive contexts, individual differences and reading. In: Afflerbach, P. (Ed.), *Handbook of Individual Differences in Reading: Reader, Text and Context*. Routledge, pp. 209–233.
- Johnston, P., Ivey, G., Faulkner, A., 2011. Talking in class: remembering what is important about classroom talk. *Read. Teach.* 65, 232–237.
- Kontovourki, S., 2012. Reading leveled books in assessment-saturated classrooms: a close examination of the unmarked processes of assessment. *Read. Res. Q.* 47, 153–171.
- Kos, R., 1991. Persistence of reading disabilities: the voices of four middle school students. *Am. Educ. Res. J.* 28, 875–895.
- La Guardia, J.G., 2009. Developing who I am: a self-determination approach to the establishment of healthy identities. *Educ. Psychol.* 44, 90–104.
- Learned, J., 2016. “The behavior kids”: examining the conflation of youth reading difficulty and behavior problem positioning among school institutional contexts. *Am. Educ. Res. J.* 53, 1271–1309.
- Learned, J., 2018. Classroom contexts and the construction of struggling high school readers. *Teach. Coll. Rec.* 120, 1–47.
- Lecce, S., Bianco, F., Hughes, C., 2021. Reading minds and reading texts: evidence for independent and specific associations. *Cognit. Dev.* 57, 1–14.
- Lysaker, J., 2019a. *Before Words: Wordless Picture Books and the Development of Reading in Young Children*. Teachers College Press.
- Lysaker, J., 2019b. Expanding meaningfulness in literacy research: possibilities of a relational perspective. *Literacy Res.* 67, 64–85.
- Mar, R.A., Oatley, K., 2008. The function of fiction is the abstraction and simulation of social experience. *Perspect. Psychol. Sci.* 3 (3), 173–192.
- Maughn, B., Rowe, R., Loeber, R., Stouthamer-Loeber, M., 2003. Reading problems and depressed mood. *J. Abnorm. Child Psychol.* 31, 219–229.
- Merga, M., 2018. Silent reading and discussion of self-selected books in the contemporary classroom. *Engl. Aust.* 53, 70–82.
- Morgan, P.L., Farkas, G., Tufis, P.A., Sperling, R.A., 2008. Are reading and behavior problems risk factors for each other? *J. Learn. Disabil.* 41, 417–436.
- Morgan, P.L., Farkas, G., Wu, G., 2012. Do poor readers feel angry, sad, and unpopular? *Sci. Stud. Read.* 16, 360–381.
- Ouellette, G., Sénéchal, M., 2017. Invented spelling in kindergarten as a predictor of reading and spelling in Grade 1: a new pathway to literacy, or just the same road, less known? *Dev. Psychol.* 53, 77–88.
- Reeve, J., Tseng, C., 2011. Cortisol reactivity to a teacher’s motivating style: the biology of being controlled versus supporting autonomy. *Motiv. Emot.* 35, 63–74.
- Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78.
- Schrijvers, M., Jannsen, T., Fialho, O., De Maeyer, S., Rijlaarsdam, G., 2019. Transformative dialogic literature teaching fosters adolescents’ insight into human nature and motivation. *Learn. Instruct.* 63, 101216.
- Skaalvik, S., 1993. Ego-involvement and self-protection among slow learners: four case studies. *Scand. J. Educ. Res.* 37, 305–315.
- Skaalvik, S., 2004. Reading problems in school children and adults: experiences, self-perceptions and strategies. *Soc. Psychol. Educ.* 7, 105–125.
- Skerrett, A., Bomer, R., 2011. Borderzones in adolescents’ literacy practices: connecting of out-of-school literacies to the reading curriculum. *Urban Educ.* 46, 1256–1279.
- Skerrett, A., 2012. “We hatched in this class”: repositioning of identity in and beyond a reading classroom. *High Sch. J.* 95, 62–75.
- Stanovich, K.E., 1986. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21, 360–407.
- Streb, J., Keis, O., Lau, M., Hille, K., Spitzer, M., Sosic-Vasic, Z., 2015. Emotional engagement in kindergarten and school children: a self-determination theory perspective. *Trends Neurosci. Educ.* 4, 102–107.
- Sugarman, J., Martin, J., 2011. Theorizing relational agency. *J. Constr. Psychol.* 24, 283–289.
- Turner, J.C., 1995. The influence of classroom context on young children’s motivation for literacy. *Read. Res. Q.* 30, 410–441.
- Wilkinson, I.A.G., Son, E.H., 2011. A dialogic turn in research on learning and teaching to comprehend. In: Kamil, M.L., Pearson, P.D., Moje, E.B., Afflerbach, P. (Eds.), *Handbook of Reading Research*, vol. 4. Routledge, pp. 359–387.
- Willcutt, E.G., Pennington, B.F., 2000. Psychiatric comorbidity in children and adolescents with reading disability. *J. Child Psychol. Psychiatr.* 41, 1039–1048.

Young adult literature references

- Asher, J., 2007. *Thirteen Reasons Why*. Razorbill.
- Booth, C., 2010. *Tyrell Scholastic*.
- Feinstein, J., 2005. *The Last Shot*. Alfred A. Knopf.
- Hinton, S.E., 1967. *The Outsiders*. Penguin.

- Neri, G., 2010. *Yummy: The Last Days of a Southside Shorty*. Lee & Low.
- Oliver, L., 2010. *Before I Fall*. HarperCollins.
- Scott, E., 2008. *Living Dead Girl*. Simon Pulse.
- Shaara, M., 1974. *The Killer Angels*. David McKay Company.
- Strasser, T., 2008. *If I Grow up*. Simon & Schuster.
- Williams, S., 1998. *Life in Prison*. Chronicle.
- Williams, C.L., 2012. *Glimpse*. Simon & Schuster.

The meaning(s) of dyslexia

Richard L. Allington^{a,b} and Rachael Gabriel^{a,b}, ^a University of Tennessee, Knoxville, TN, United States; and ^b University of Connecticut, Storrs, CT, United States

© 2023 Elsevier Ltd. All rights reserved.

A brief etymological history of the term “dyslexia”	408
Varying definitions of dyslexia	409
Challenges related to assessment and categorization	410
Competing theories as to the cause of dyslexia	411
Perceptual issues	411
Difficulties with phonological processing	411
Research in neuroscience and dyslexia	411
Early identification	411
Conclusions	412
References	412

Dyslexia refers typically to the difficulties that some children have in learning to read. Though its definition varies widely and has been contested and refined from different angles in different decades and contexts (c.f. [Elliot, 2020](#)), it most often refers to a profile of difficulties learning to read that, taken together, describe a profile of difficulty learning to read. While dyslexia has been studied by researchers for over 100 years there still exists debate over (a) whether dyslexia is a unique form of reading disability that differs from the most common reasons for reading difficulties and (b) how best to intervene to address the instructional needs of children identified as dyslexic ([Worthy et al., 2018](#)). There is also concern over the size of the dyslexic population, the basis necessary for identifying a child as dyslexic, and over whether there exist any proven instructional strategies for effectively teaching dyslexic children to read. Much of the concern has a long history, as [Gabriel \(2019\)](#) has noted,

Dyslexia is a contested construct within and across the fields of education, medicine and psychology. Debates about its definition, etiology, prevalence, diagnostic criteria, treatment and relationship to theories of literacy development have continued for more than a century (p. 343).

In this entry we hope to offer a broad portrait of the concerns that have been raised about the topic of dyslexia with a particular focus on the meanings ascribed to the word.

A brief etymological history of the term “dyslexia”

If we backtrack roughly 125 years, we will find that the term dyslexia was first used by [Berlin \(1887\)](#), to describe a condition that had earlier been called “word blindness.” A small number of scholars had been writing about word blindness since 1676 ([Hinshelwood, 1917](#)). Almost everything written about dyslexia prior to 1930 was written by a medical doctor. “Dyslexia” emerged as an educational problem around the turn of the 20th century as education became compulsory across Europe and then in the US. With more young people attending school and engaging in early literacy development, both educators and psychologists were ushered into specializing in the identification and remediation of dyslexia. As awareness of and engagement with students with difficulty learning to read expanded, so did the multiple definitions and understandings that were offered to explain dyslexia. In fact, today there are definitions of a dyslexic child available from pediatricians, neurologists, educational psychologists, special educators, general education personnel, as well as differing definitions from different educational and medical agencies at both local and national levels. Each of several professions, as well as individual researchers, have offered differing definitions of dyslexia. These differences largely center on whether dyslexia refers to any difficulty learning to read, only those that are unforeseen, or to a specific profile of reading difficulty, which itself is described variably. However, the presence of multiple definitions of dyslexia has created its’ own set of problems, as noted by several scholars:

Without an agreed-on definition that can be implemented reliably and validly, understanding the nature, causes, and best treatments for reading disability is unlikely.

[Brown Waesche et al. \(2011, p. 296\).](#)

It is difficult [to define dyslexia] because the field has been unable to produce a universally accepted definition that is not imprecise, amorphous, or difficult to operationalize.

Elliott and Grigorenko (2014, p. 5).

We would add that any estimate of the incidence of dyslexia in a population must also be taken lightly since there is no agreed upon standard for diagnosis, much less any research-based support for any estimate of its prevalence.

Varying definitions of dyslexia

The contested field surrounding the topic of dyslexia is, perhaps, best summarized by Elliott and Grigorenko (2014) in their book-length treatment of the problems confronting anyone interested in the topic. They argue that at least four major and differing definitions of dyslexia exist and each definition aligns itself with a particular profession. Elliott and Grigorenko begin by noting that dyslexia has been seen as either different from or synonymous with a variety of other terms, such as “specific reading disability,” “learning disability in reading,” “reading disorder,” “severe reading disability,” and “unexpected reading difficulty.” We would also note that most research reporting on individuals with dyslexia labels draws a sample of dyslexic subjects from a population of children that has experienced difficulty in acquiring reading proficiency alongside their peers. The size of the lag in reading development in children varies depending upon the research team. Evidence of dyslexia was, until recently, seen in a child whose reading proficiency was substantially lower than a child’s IQ would suggest it should be. However, the IQ/achievement discrepancy definition of dyslexia was dropped in the US under IDEA 2004, and today dyslexia is seen as a possibility for explaining why some children, regardless of their IQ, have significant difficulty learning to read. In particular, reading difficulties related to word recognition and spelling, which will impact other areas of reading such as fluency (via difficulties in word recognition) and therefore (via practice reading and exposure to text) vocabulary, comprehension and engagement. Shaywitz and Shaywitz (2013), for instance, noted that “Perhaps the most consistent and enduring core of the definition is the concept of dyslexia as unexpected difficulty in reading” (p. 643). The word “unexpected” in the Shaywitz’s definition also appears in the International Dyslexia Association (IDA)’s “consensus definition” which is the definition most widely adopted by states in the US:

Dyslexia is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge.

Early evidence from states that adopted dyslexia-specific policies shows that the concept of dyslexia as an “unexpected” difficulty has great consequence for whom is identified as dyslexic and whose difficulty is assumed to be related to circumstance, race, gender or class. For example, Odegard et al. (2020) found “systematic demographic differences in whether a student is identified with dyslexia by schools even when using universal screening.” Students that attended schools with a higher percentage of minority students, with larger percentages of students with lower literacy skills, or students from minoritized backgrounds were significantly less likely to be assigned a dyslexia label in school, even when behavioral characteristics of dyslexia were observed. In other words, where reading difficulty is expected based on demographics, students are much less likely to be identified as dyslexic than those attending schools where the same difficulty is not expected.

Likewise, using data from the UK Millennium Cohort Study, Knight and Crick (2021) reported that “factors seemingly unrelated to the clinical aspects of dyslexia influence whether a child is identified as dyslexic in England and Wales.” In particular, gender, season of birth, socio-economic class and parental education were found to be significant predictors of the dyslexia label. They argue that a “dyslexic sub-group” has been created within the population of poor readers with questionable “reliability, validity and moral integrity of the allocation of the dyslexia label.”

Without prescribing a specific definition, Tunmer and Greaney (2010) offered a set of criteria for any definition of dyslexia. They included the following three components that must be present when identifying dyslexia as the cause of a child’s difficulties with reading acquisition. The three components necessary were (a) persistent difficulty in learning to read, (b) the child has access to evidence-based reading instruction, and (c) inclusionary factors. However, their criteria included that the child had had access to evidence-based reading instruction, in either or both, the classroom and intervention reading lessons. We mention this if only because this factor has been neglected in most other definitions and in the evaluations conducted to identify children as dyslexic.

It seems then that difficulty, in fact, persistent difficulty, in learning to read even when provided evidence-based reading lessons is necessary for a child to be identified as a dyslexic child. However, in our experience, evidence of access to high-quality reading lessons is almost never documented in diagnostic case reports identifying a child as dyslexic. Instead, the identification seems to simply require that the child has had the opportunity to participate in reading lessons, but has not much benefitted from this participation.

We accept that dyslexia might be used to describe a child who has experienced difficulties in acquiring reading proficiency when provided high-quality reading instruction. The issue with this definition is that different scholars would likely define and measure both “difficulties” and “high quality instruction” differently.

If dyslexic children differ from other struggling readers in some way, then the question remains as to what sort of differences the definition includes. In addition, there is little agreement on what features constitute “high quality” reading instruction. This lack of clarity in defining dyslexia seems to underlie the suggestions that dyslexia is a common affliction and affects up to 20% of the population (Shaywitz, 2005) or that dyslexia is uncommon and affects less than 1% of the population (Pressley and Allington, 2014). Mazzocco and Grimm (2013), for instance, used nonword reading scores below the 10th percentile as their criterion for identifying dyslexic students. But even 10 children out of every 100 children (10%) seems too high a cutoff point for identifying dyslexics. Those who prefer to think that dyslexia is common have argued that up to 20% of American children are afflicted with dyslexia (Shaywitz, 2005). Those who think that dyslexia is uncommon and affects less than 1% of the population hold a much different view of the disability labeled dyslexia (Pressley and Allington, 2014).

By merely adopting, revising or creating a different definition of dyslexia one can categorize any reader as dyslexic or nearly all readers as non-dyslexic. The question, then, is who benefits from setting criteria in a particular way? In some settings, the label ensures students receive access to more and better instruction, in others, it relegates their study of reading to programs that are a poor match for their existing strengths, difficulties and identities (c.f. Frankel et al., 2021; Sleeter, 2018). In some cases, the label explains a specific area of difficulty, thus allowing adults to support the student toward high academic expectations despite a delay in reading development. In other cases, the label itself lowers expectations and support considerably. The potential impact of understanding reading difficulty as dyslexia may be considered as varied as the definition itself given the social and instructional realities that greet children over time and across geographic regions.

Challenges related to assessment and categorization

Still, in the US alone, experts have suggested everything from a spectrum theory that everyone is a little bit dyslexic, but some are more than others; or that everyone is dyslexic until they learn to read, all the way to much more specific and stringent criteria that limits projected prevalence to less than 1% of the population (see Protopapas, 2019 for a discussion). This range of ways to understand what dyslexia is and is not is also visible within the category itself. Scholars have identified multiple subtypes of dyslexia using different methods of categorization (Katzir et al., 2008). One method is related to etiology, or whether the set of difficulties learning to read are developmental or acquired through stroke or brain injury. Another is by profiling specific difficulties. For example, within the area of phonological processing, which is most often associated with developmental learning disabilities rather than later injuries, there are two types of difficulties: those related to difficulty related to Rapid Automatic Naming (RAN) or the ability to quickly recognize and associate a symbol with a sound and Phonological Awareness or the ability to hear and manipulate the sounds in words (c.f. Katzir et al., 2008; Vander Stappen and ReyBroeck, 2018). One may, of course, have both difficulties at once, which is referred to as “double-deficit” or “mixed” dyslexia. Finally, given that dyslexia can be understood as having high heritability, and a complex genetic structure, neuroscientists have attempted to examine the heterogeneous relationships between genes, brain activation patterns, and reading behavior (c.f. van Zuijlen et al., 2013; see Eicher and Gruen, 2013 for a review). Amid all these possibilities for understanding and defining dyslexia, the US has seen a wave of state-level legislation enshrining definitions, screening procedures and instructional responses into state codes in the years between 2010 and 2020 (Zirkel, 2020).

For example, the International Dyslexia Association’s (IDA) definition (see above) has been adopted by and/or has informed state-specific definitions and those of other similar societies across the globe. IDA is a US.-based organization, originally called “The Orton Society,” for those with an interest in the methods of multisensory direct instruction described by Dr. Samuel Orton and Anna Gillingham in the 1930s. This formerly niche organization grew its membership in the US significantly between 2010 and the present. Its “consensus definition,” created by a panel in the early 2000s, focuses on features like a neurobiological (innate) origin, of dyslexia as well as its relationship to phonological processing and its “unexpected” nature (IDA, 2002; see Gabriel, 2018 for a discussion). The challenge with the unexpected nature of the difficulty is, as we discuss later, deciding whether and when we might expect the experience of having difficulty learning to read signals a child is dyslexic.

Today many scholars argue that a key characteristic of dyslexia is difficulties with single word decoding. Testing the ability to recognize individual words eliminates the sorts of contextual cues (semantic and syntactic) that a child might use to assist in identifying individual words were she reading a passage of continuous text (Vellutino et al., 2004). The problem with this view of dyslexia is that reading is essentially extracting meaning from printed text, a task quite different from identifying words (or nonwords) in isolation. Of course, a difficulty in identifying individual words in a passage would typically mean that comprehension would suffer. On the other hand, the ability to identify all the words in a passage does not ensure that the reader has understood what was just read. At this point, one can see the basic problem with defining dyslexia and with operationalizing that definition.

Competing theories as to the cause of dyslexia

Perceptual issues

Orton (1925) noted that dyslexic children often reversed letters and words. This was explained through his theory of dyslexia which suggested that mixed cerebral dominance was at the base for children experiencing reading difficulties and that this mixed dominance would account for the word and letter reversals he observed in some children who struggled with acquiring literacy. However, roughly 40 years later Vellutino (1979) published a book length review of the evidence for visual perceptual problems as the cause of dyslexia. He demonstrated that the reversals that Orton had noted, were in fact, common among beginning readers and that some children had difficulty resolving the issue that *b* and *d*, while reversals of each other, were still two different letters.

The study of dyslexia began to broaden in the 1980s and continues today with not only physicians and educators interested in the problem of unexpected difficulties in learning to read but also psychologists, neurologists, ophthalmologists, lawyers, and members of other professions. This broadening of the fields of interests presents the possibility that sooner rather than later there will be a breakthrough that clarifies the many areas around dyslexia where debates by opposing parties occur. But for now:

We have found ourselves in a position in which science has presided over the production of a vast volume of research studies in the field of reading difficulty, but the translation of this into a coherent body of scientific knowledge that appropriately informs legislation, and clinical and educational practice, continues to be absent.

Elliott (2020, p. 565).

This persistent problem is understood variably as a problem of knowledge translation (e.g., Solari et al., 2020), the site and nature of research (e.g., Milner, 2020) and engagement with issues of implementation (Woulfin and Gabriel, 2020). Despite these persistent problems in the translation of the vast volume of research in the area from varying fields and perspectives, research on dyslexia continues to evolve in nature and focus.

Difficulties with phonological processing

With the demise of visual perceptual-based definitions, the 1990s saw the rise of problems with phonological tasks as the reason for dyslexia. At the same time, though, problems with phonological tasks were identified as common among virtually all children experiencing difficulties with reading acquisition (Adams, 1990).

Research in neuroscience and dyslexia

Most recently, the neurological processes involved in reading have begun to be addressed but today we have only a very small body of neuroscience research available and will require substantially more neuroscience studies to inform us reliably (e.g., Munzer et al., 2020). Technology to support neuroimaging and genetic research has developed rapidly over the past decades. New technologies have created opportunities to identify and confirm differences in brain anatomy, organization and functioning between participants with and without dyslexia labels, and between participants with dyslexia labels before and after instruction or instructional intervention.

Family studies, particularly studies of twins, have shown both how reading difficulties are heritable and how environmental and behavioral factors influence their trajectories (Church et al., 2021). Molecular genetic studies, those aimed at identifying specific genes that predict or are associated with dyslexia, have identified many candidate genes, but no isolated region or specific sequence as of yet (Carrion-Castillo et al., 2013). Similarly, genetics research has been used to understand rates and likelihood of heritability as well as the possibility of early and accurate screening for predispositions toward reading difficulty, and a framework for studying epigenetic and protective factors that prevent reading difficulties (Erbeli et al., 2021).

We estimate that the concept of protective factors is likely to be the most generative area of research in the future because of its potential to contribute to understandings about early prevention of difficulties rather than intervention once difficulties have risen to the level of accurate identification in school settings. For example, researchers have recently investigated whether there is a protective or preventative roles for rhythm play, musical training and shared reading with 2 and 3 year old children with family histories of dyslexia (c.f. Cancer and Antonetti, 2022; see also: Huss et al., 2011). These and other similar studies could significantly reduce the percentage of students who experience difficulty learning to read. Investigations of the potential for prevention is especially important because, despite dramatic growth in technological capability and feasibility, as Church et al. (2021) have noted, "Neural and genetic studies do not prescribe how children should be taught to read, but these studies have underscored the critical role of early intervention and ongoing support" (p. 1).

Early identification

In the early 2020s, some researchers argue that early screening and intervention is paramount so that dyslexic children should be identified at an early age so as to not waste time as a severe reading disability develops (Snowling, 2013). Apps that can screen for

early signs of difficulty and tools identification in Kindergarten or PreK are currently under development (Perera et al., 2016). However, the research on such early identification is problematic when there has been little consensus on which developmental factors need to be assessed to identify a young child as a dyslexic. Without both an accepted and substantial research base identifying the common preliterate components that will be needed to ensure that those children helped by early intervention would have had much difficulty without the early intervention plus strong evidence that the condition of dyslexia has been avoided by those children who received the early intervention.

Scanlon et al. (2005), for instance, compared the progress struggling young readers made in reading acquisition under three conditions. First, one group of struggling readers received a text emphasis intervention. A second group received a phonological skills intervention and a third group was the control group of struggling readers who received neither intervention. The results showed that both treatment groups of children benefitted and both intervention groups outperformed the control group children on the word reading and the text comprehension assessments. However, the text emphasis children outperformed the phonological emphasis children on the word identification assessment. Thus, the group concluded,

It is clear that many children who were identified as poor readers at the beginning of first grade responded well to the instruction they received, regardless of the type of instruction provided.

Scanlon et al. (2005, p. 224).

While this study makes it clear that early intervention can benefit struggling young readers, it is unclear about the implications for children identified at age 5 as dyslexic. This because the Scanlon et al. (2005) study provided early intervention to all young children who were found to perform below average on a reading assessment, thus they had low-achieving young children, not necessarily a group of young children identified as dyslexic.

A major limitation in the research on dyslexic children replicates the Scanlon et al. (2005) study—using low achieving students but not necessarily students who had been identified as dyslexic students. A major reason for this limitation is the lack of a consensus on how one would identify a dyslexic child. However, with no universally accepted definition of what constitutes dyslexia, developing criteria for including/excluding a struggling reader from a pool of low-achieving students becomes an impossible task. We should not be surprised, then, that studies purporting to be studies of dyslexic children are too often more simply studies of struggling readers. With no consensus on a definition of dyslexia, one can never be sure whether the subjects in the study identified as dyslexic children were children who, in fact, met the three criteria listed by Tunmer and Greaney (2010) or not. The same holds for estimates of the number (or percentage of the population) of dyslexics in a class, a school, a school district, a state, or nationally.

Conclusions

Perhaps, in the future we will be able to be quite precise on what constitutes a dyslexic reader, other than difficulties with reading acquisition. We write this because, whether a condition called dyslexia exists as a natural, identifiable pattern of human variability or not, we know of too many children who struggle with reading acquisition and never truly become achieving readers. But for now, we will argue along the same lines as did Elliott and Grigorenko (2014) who noted that there were many definitions of dyslexia, that there exists no widespread consensus for an operational definition even where the term is defined, and no good reason to prefer one definition over another. Following them we would also argue that substituting the term reading difficulty for dyslexia makes sense to us in this time of limited agreement on the definition or the defining characteristics of dyslexia.

References

- Adams, M.J., 1990. *Beginning to Read: Thinking and Learning About Print*. MIT Press, Cambridge, MA.
- Berlin, R., 1887. *A Specific Kind of Word Blindness*. Bergman, Wiesbaden, Germany.
- Brown Waesche, J.S., Schatschneider, C., Maner, J.K., Ahmed, Y., Wagner, R.K., 2011. Examining agreement and longitudinal stability among traditional and RTI-based definitions of reading disability using the affected-status agreement statistic. *J. Learn. Disabil.* 44, 296–307.
- Cancer, A., Antonietti, A., 2022. Music-based and auditory-based interventions for reading difficulties: a literature review. *Heliyon* 8 (4), e09293. <https://doi.org/10.1016/j.heliyon.2022.e09293>.
- Carrion-Castillo, A., Franke, B., Fisher, S.E., 2013. Molecular genetics of dyslexia: an overview. *Dyslexia* 19 (4), 214–240.
- Church, J.A., Grigorenko, E.L., Fletcher, J.M., et al., 2021. The role of neural and genetic processes in learning to read and specific reading disabilities: implications for instruction. *Read. Res. Q.* 1–17. <https://doi.org/10.1002/rrq.439>.
- Eicher, J., Gruen, J., 2013. Imaging-genetics in dyslexia: connecting risk genetic variants to brain neuroimaging and ultimately to reading impairments. *Mol. Genet. Metab.* 110 (3), 201–212. <https://doi.org/10.1016/j.ymgme.2013.07.001>.
- Elliott, J.G., 2020. It's time to be scientific about dyslexia. *Read. Res. Q.* 55 (S1), 561–575.
- Elliott, J.G., Grigorenko, E.L., 2014. *The Dyslexia Debate*. Cambridge University Press, London.
- Erbil, F., Rice, M., Paracchini, S., 2021. Insights into dyslexia genetics research from the last two decades. *Brain Sci.* 12 (1), 27.
- Frankel, K.K., Deanna Brooks, M., Learned, J.E., 2021. A meta-synthesis of qualitative research on reading intervention classes in secondary schools. *Teach. Coll. Rec.* 123 (8), 31–58.
- Gabriel, R., 2018. Dyslexia legislation: a brief history. *J. Read. Recovery* 17 (2), 25–34.
- Gabriel, R.E., 2019. Converting to privatization: a discourse analysis of dyslexia policy narratives. *Am. Educ. Res. J.* 57 (1), 355–358.

- Hinshelwood, J., 1917. Congenital Word Blindness. H. K. Lewis, London.
- Huss, M., Verney, J.P., Fosker, T., Mead, N., Goswami, U., 2011. Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex* 47 (6), 674–689. <https://doi.org/10.1016/j.cortex.2010.07.010>.
- International Dyslexia Association, 2002. Consensus Definition Project. Dyslexiaida.org. Retrieved from. <https://dyslexiaida.org/definition-consensus-project>.
- Katzir, T., Kim, Y.-S., Wolf, M., Morris, R., Lovett, M., 2008. The varieties of pathways to dysfluent reading: comparing subtypes of children with dyslexia at letter, word, and connected text levels of reading. *J. Learn. Disabil.* 41, 47–66. <https://doi.org/10.1177/0022219407311325>.
- Knight, C., Crick, T., 2021. The assignment and distribution of the dyslexia label: using the UK Millennium Cohort Study to investigate the socio-demographic predictors of the dyslexia label in England and Wales. *PLoS One* 16 (8), e0256114. <https://doi.org/10.1371/journal.pone.0256114>.
- Mazzocco, M.M.M., Grimm, K.J., 2013. Growth in rapid automatized naming from grades K to 8 in children with math or reading disabilities. *J. Learn. Disabil.* 46 (6), 517–533.
- Milner, H.R., 2020. Disrupting racism and whiteness in researching a science of reading. *Read. Res. Q.* 55 (S1), S249. <https://doi.org/10.1002/rq.347>.
- Munzer, T., Hussain, K., Soares, N., 2020. Dyslexia: neurobiology, clinical features, evaluation and management. *Transl. Pediatr.* 9 (Suppl. 1), S36.
- Odegard, T.N., Farris, E.A., Middleton, A.E., Oslund, E., Rimrodt-Frierson, S., 2020. Characteristics of students identified with dyslexia within the context of state legislation. *J. Learn. Disabil.* 53 (5), 366–379.
- Orton, S.T., 1925. Word-blindness in school children. *Arch. Neurol. Psychiatr.* 14 (5), 581–615.
- Perera, H., Shiratuddin, M.F., Wong, K.W., 2016. Review of the role of modern computational technologies in the detection of dyslexia. In: *Information Science and Applications (ICISA) 2016*. Springer, Singapore, pp. 1465–1475.
- Pressley, M., Allington, R.L., 2014. *Reading Instruction That Works: The Case for Balanced Teaching*, fourth ed. Guilford, New York.
- Protopapas, A., 2019. Evolving concepts of dyslexia and their implications for research and remediation. *Front. Psychol.* 10, 2873. <https://doi.org/10.3389/fpsyg.2019.02873>.
- Scanlon, D.M., Vellutino, F.R., Small, S.G., Fanuele, D.P., Sweeney, J.M., 2005. Severe reading difficulties—can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality* 13 (4), 209–227.
- Shaywitz, S., 2005. *Overcoming Dyslexia: A New and Complete Science-Based Program for Reading Problems at Any Level*. Vintage Books, New York.
- Shaywitz, S.E., Shaywitz, B.A., 2013. Making a hidden disability visible: what has been learned from neurological studies of dyslexia. In: Swanson, H.L., Harris, K.R., Graham, S. (Eds.), *Handbook of Learning Disabilities*, second ed. Guilford, New York, pp. 643–657.
- Sleeter, C.E., 2018. Why is there learning disabilities of a critical analysis of the birth of the field in its social context. In: *The Formation of School Subjects*. Routledge, pp. 210–237.
- Snowling, M.J., 2013. Early identification and interventions for dyslexia: a contemporary view. *J. Res. Spec. Educ. Needs* 13 (1), 7–14.
- Solari, E.J., Terry, N.P., Gaab, N., Hogan, T.P., Nelson, N.J., Pentimonti, J.M., Petscher, Y., Sayko, S., 2020. Translational science: a road map for the science of reading. *Read. Res. Q.* 55 (S1), S347–S360. <https://doi.org/10.1002/rq.357>.
- Tunmer, W., Greaney, K., 2010. Defining dyslexia. *J. Learn. Disabil.* 43 (3), 229–243.
- Vander Stappen, C., Reybroeck, M.V., 2018. Phonological awareness and rapid automatized naming are independent phonological competencies with specific impacts on word reading and spelling: an intervention study. *Front. Psychol.* 9, 320. <https://doi.org/10.3389/fpsyg.2018.00320>.
- van Zuijen, T.L., Plakas, A., Maassen, B.A., Maurits, N.M., van der Leij, A., 2013. Infant ERPs separate children at risk of dyslexia who become good readers from those who become poor readers. *Dev. Sci.* 16 (4), 554–563. <https://doi.org/10.1111/desc.12049>.
- Vellutino, F.R., 1979. *Dyslexia: Theory and Research*. MIT Press, Cambridge, MA.
- Vellutino, F.R., Fletcher, J.M., Snowling, M.J., Scanlon, D.M., 2004. Specific reading disability (dyslexia): what have we learned in the past four decades? *JCPP* 45 (1), 2–40.
- Worthy, J., Svrcek, N., Daly-Lesch, A., Tily, S., 2018. “We know for a fact”: dyslexia interventionists and the power of authoritative discourse. *J. Lit. Res.* 50 (3), 359–382.
- Woulfin, S., Gabriel, R.E., 2020. Interconnected infrastructure for improving reading instruction. *Read. Res. Q.* 55, S109–S117.
- Zirkel, P., 2020. Legal developments for students with dyslexia. *Learn. Disabil. Q.* 43 (3), 127–139.

A constructivist lens of eye-tracking in special populations

Ryan L. Nelson^a, Laura E. Arrington^a, Narges Firouzshahi^a, Patricia Arias Hunt^a, Amanda Percle^a, Holly Damico^a, and Jack S. Damico^b, ^aUniversity of Louisiana at Lafayette, Lafayette, LA, United States; and ^bUniversity of Colorado Boulder, Boulder, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

A constructivist lens of eye-tracking in special populations	414
Understanding eye-tracking—key components & terms	415
Presentation of eye-tracking data	415
Overview of theory	416
A behaviorist perspective on eye-tracking	416
Eye-tracking through a social-constructivist lens	417
Major contributions of eye-tracking research to understanding literacy	417
Strategic visual sampling	417
The nonlinearity of scan paths	417
Normative eye movement patterns from reading	418
Nonlinear characteristics of eye movements in reading	418
Variables influencing fixation location	418
Patterns of regression	419
Variations in eye movement due to text difficulty and reader purpose	419
Which words are fixated?	419
Limitations and conclusion	420
Design and Initial Procedure	420
Data processing and analysis	420
Interpretations	421
References	421
Further reading	422

A constructivist lens of eye-tracking in special populations

A number of methods have been utilized as researchers, educators, and special service providers have sought to better understand how readers marshal resources to make sense of print. Among these, eye movement research has long been used in the investigation of how individuals visually sample print during reading and reading related activities (Paulson and Goodman, 1999). For well over 100 years, studies of eye movements have captivated researchers seeking to better conceptualize reading. With the advancement of technology, eye-tracking has become an increasingly accessible data source for research into the underlying cognitive processes associated with reading (Starr and Rayner, 2001; Rayner, 1978; Liwanag et al., 2017) and into the strategies readers employ while transacting with text (Krstić et al., 2018; Percle et al., 2020). The cost reductions resulting from technological innovation and accessibility of reliable equipment have enabled eye tracking research to consider reading behaviors in many different languages and cultures across the world (e.g., Hermena, 2021; Cop et al., 2017; Hung, 2021; Jian, 2015; Zang et al., 2011). Just as an audio recorder functions to capture verbal productions, in a parallel manner, eye-tracking systems are the recorder of eye movements. The researcher's contextualized expertise in understanding and interpreting behaviors is required to gain meaningful insights from the movements of the eye.

A variety of extensions of eye-tracking technologies for use with special populations assist in accomplishing multiple purposes. These applications are too expansive to address in a single entry. Therefore, we feel it appropriate to state the parameters through which we are conceptualizing the topic of eye tracking in special populations. Our consideration of special populations includes those individuals with a diagnosis of impairment sufficient to qualify for special services or individualized education plans. While we accept that many reject the deficit-based practice of assigning educational labels of impairment in favor of recognition of cognitive diversity, the literature in eye-tracking and the practices of educational professionals continue to largely operate under these frameworks. In this entry, we apply the theoretical lens of constructivism to interpret relevant findings in eye-tracking research with consideration to their differing applications. This enables us to apply a lens focused on meaning-making through which to consider the available literature.

We also confine our view of eye-tracking to studies engaged in illuminating the reading process. We purposefully avoid the literature focused on using technology to navigate assistive and augmentative communication devices. We also omit eye-tracking use as interventions, such as eye scan-path training or eye strengthening exercises. Instead, we focus on the literature informing the literacy process as it is realized in the lives of readers, including persons with special needs. The application of eye-tracking technology to the study of reading and the methodologies implemented in the investigation apply to all readers, regardless of efficiency. Generally

speaking, the patterns of eye movements in special populations largely match typically developing individuals of similar proficiency (Arya and Feathers, 2012; Nelson et al., 2008; Seassau and Bucci, 2013; Krstić et al., 2018). Because studies of reading apply to all readers, we can employ and integrate the findings of studies exclusive of special populations to this demographic.

Understanding eye-tracking—key components & terms

Several general terms are important for considering eye-tracking research, namely: fixation, saccade, and regression. Most eye-tracking systems capture the angle of the fovea, a flat part of the eye responsible for the greatest amount of visual acuity, as it relates to the stimuli or visual scene. This is done through a process of identifying the reflection of harmless infrared light off the cornea and the resulting variation in pupil diameter changes. Each of these points are translated into coordinates capable of being graphed. Through a calibration process, the point of gaze can be appended onto images of a scene and interpreted as locations upon which the individual looks. When a person's point of gaze remains within a narrowly defined window for a predetermined length of time, the person is said to have produced a fixation (Rayner, 1997).

Fixations are moments when the eye is relatively unmoving. It is during fixations when the processing of visual stimuli is believed to be possible. Understanding of fixations as times when the eye is relatively stable enough to process visual information comes from tachistoscopic studies of the length of time necessary for an individual to identify randomly presented print symbols (Huey, 1968). *Saccades* are the ballistic-like, rapid movements the eyes produce as they move from one fixation to the next. Attention to the duration of fixations or distance covered by saccades between fixations continues to be used by researchers making inferences about literacy processes. Additionally, in the eye-tracking research associated with reading, the term *regression* is used for eye movements produced when readers direct the eyes in any backwards movement (i.e., backwards in relation to the directionality of text within the language). These three terms essentially represent graphed coordinates in the same way the audio recordings represent sound wave productions. Their utility as research indices are dependent upon the investigator's interpretation of their significance. In social sciences of literacy, such interpretations are always driven by a theoretical stance (Geertz, 1996).

Presentation of eye-tracking data

As with all social science research, it is difficult to represent findings related to complex phenomena in written form. Doing so requires the researchers to present a static representation of dynamic processes. The challenge is to represent where the eye was, when it was there, and how the mind considered the visual input. Data is presented as a static image due to the limitations of print media. Because the goal is to represent the pattern of visual sampling, researchers present eye-tracking data several ways.

In Fig. 1, adapted from Paulson (2002), the reader's eye-movements are represented by a number positioned beneath the word on which the reader fixated. The numbers provide the sequence of fixations. For example, line 308 begins with the last three letters of the word "continued" ("ued showing right up until the day when, with a muffled"). The first fixation occurs at the beginning of the word "showing." Fixation two occurs in the middle of the next word, "right." Fixation three occurs at the end of the word "up," at which point the reader regresses back to the end of the word "showing." The reader then directs the eyes past the next two words, producing the fifth fixation on the word "until." In this style of data presentation, we can also visualize differences in length of fixation. Here, the duration of the fixation is conveyed directly below the numbered fixation. For example, fixation two (750 ms) was more than three times the length of fixation four (200 ms).

As technology has improved, so has the presentation of eye-movement data. Eye tracking systems now enable fixations to be superimposed onto a static image. Fig. 2 (an unpublished image from Nelson et al.'s 2008 study) shows a bilingual Spanish/English child with Down's Syndrome eye movements as she read from a page of a children's book. Presentations of eye data, like that found in this example, allow researchers to envision how the reader incorporated aspects of the illustration with the visually sampled text. The fixations are represented by blue dots and the sequence is demonstrated through numbering of dots and connected via red lines. Dot size conveys length of fixations, with larger dots reflecting longer fixation durations.

In Fig. 3, a similar method is used by Hung (2019). Here, fixations are represented by a red circle superimposed over the text. Durations are again represented by the size of the circle. Often, the choice of how to present the data is driven by the purpose of the study and the information the researcher is trying to convey. These decisions are always informed by the theoretical lens guiding the researcher's question(s).

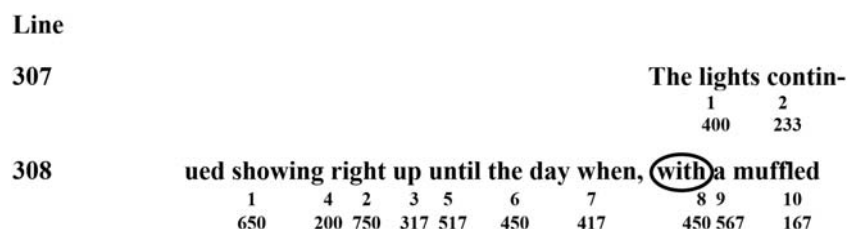


Fig. 1 Representation of sequence and duration adopted from Paulson (2002).

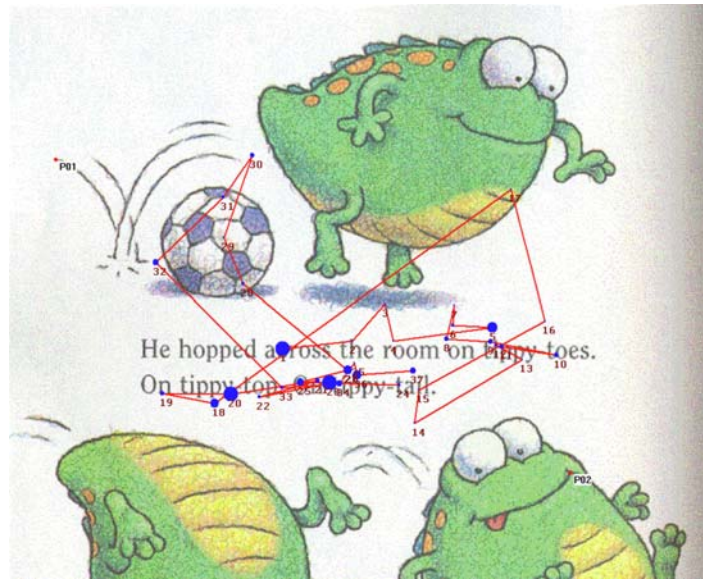


Fig. 2 . Fixation sequence from unpublished image by Nelson et al. (2008).



Fig. 3 Representation of fixations from Hung (2019).

Overview of theory

Theories provide explanatory and guiding expectations of how and why a phenomenon occurs. Researchers use these theories to both determine how to investigate issues and interpret findings. While the underlying operations of the brain are consistent, the ways in which we understand those operations differ vastly depending upon theoretical orientation. Theory can serve to illuminate or obscure, as theoretical orientations shade every aspect of our understanding of neurological processes, including reading (Tracey and Morrow, 2017). Two theoretical orientations profoundly impacting literacy research and pedagogy are behaviorism and constructivism.

A behaviorist perspective on eye-tracking

The foundation of behaviorism, where component skills were expected to build upon one another, could seem a logical basis upon which to form models of literacy (Smith, 1999). American psychological and educational research applying eye-tracking technologies to literacy was overwhelmingly influenced during the mid-twentieth century by the theoretical model of behaviorism (Paulson and Goodman, 2008; Goodman et al., 2016). The impact of this theoretical lens of eye-tracking, as applied to special populations, has been described and the limitations detailed (Damico et al., 2021). Eye-tracking techniques appeared to fit seamlessly into models oriented to the quantification of observable behaviors. Historically, plotting the reflection of harmless infrared light off of a reader's cornea enabled researchers to create a visual representation of the path followed while viewing a text. The data points monitored by eye-trackers captured fixations, as well as the sequenced path of inspection (Rayner, 1998, 2009).

With eye-tracking technology as a data collection tool, inferences were then made about how visual information was utilized by the reader. The theoretical assumptions guiding eye-tracking research from this lens become transparent when selection of stimuli and the expected responses were considered. Adopting a behaviorist theoretical stance, many eye-tracking studies of reading have chosen to present decontextualized symbols, words, and sentences where speed and accuracy in identification were the required outcomes necessary for determining whether or not the individual viewed the stimuli. Rarely considered with behaviorist approaches, were the types of reading engaged in by individuals in the classroom and certainly not in day-to-day life. Missing were studies of readers transacting with complete texts for the purposes of understanding information and integrating it into their prior background knowledge. From studies of eye-tracking, interpreted through the lens of behaviorism, assumptions about how children interacted with texts have resulted in advocacy of practices inconsistent with biological and physiological abilities of both the brain, and its handmaiden, the eye (Strauss et al., 2009; Goodman et al., 2016).

Eye-tracking through a social-constructivist lens

When a different theoretical model is applied, the same data reveal substantially more interpretive adequacy in accounting for how the brain directs the eyes in the process of making sense of print. Social-constructivist theory provided this contrastive view, and with it, a different understanding of neurological processes that are built from its purpose emerged. From the social-constructivist perspective, human interaction with the environment is one led by an overarching march toward understanding. Humans are driven to make sense of their environment, a search for meaning that is reliant upon and driven by the meanings we construct from the knowledge we have and obtain. The mind drives this process by making predictions about the world around and directing the senses to obtain confirming or disconfirming information. Paulson (2000) and others after him (Duckett, 2001; Freeman, 2001; Liwanag et al., 2017; Nelson et al., 2008; Perde et al., 2020), applied eye-tracking technology to the study of reading starting from a social-constructivist view of human learning.

Their assumption that reading was a meaning-making phenomenon informed their selection of stimuli as well as the analysis they employed to interpret the visual sampling patterns of readers (Goodman et al., 2016). The juxtaposition of eye movements occurring within the transaction of reading complete passages, offered insights into instances where readers produced unexpected responses. These unanticipated responses or oral reading miscues (Goodman, 1967) revealed how readers strategically drew upon systems of background knowledge. An investigatory integration of eye-tracking technology with the study of oral reading miscues (EMMA) offered greater interpretive adequacy. This enabled researchers to better account for the background knowledge readers brought to the texts as well as offering insights into the cognitive processes engaged in as visual sampling occurred.

Because theoretical approaches offer a lens through which research can be understood, considering research findings from the perspective that best accounts for the data yields greater understanding of phenomena in question. Contributions of the 100-plus years of eye-tracking research can be more adequately integrated to assist in understanding how readers make use of visual information if the theoretical lamp of social constructivism is applied to the data. Similar interpretive adequacy and supportive evidence better informs the implications of this research on pedagogical practice. In the sections that follow, we discuss contributions of eye-tracking through the lens of social-constructivism to the understanding of reading in general and the extensions of it to special populations.

Major contributions of eye-tracking research to understanding literacy

Strategic visual sampling

A significant contribution of eye-tracking technology to reading research, including variables associated with special populations, is the understanding that readers visually sample print through a progressive series of rapid, ballistic movements (i.e., saccades) and moments of relative stability (i.e., fixations) (Rayner, 1998, 2009). How readers sample visual information may run contrary to the general belief that eyes move continuously and smoothly across the page during reading only pausing occasionally at difficult moments. The history of eye-tracking technology applied to reading research has demonstrated that the process is, in fact, not always continuous or smooth (see the variable scan paths in Figs. 1 and 2). Eye movement research has revealed that a reader systematically alternates between fixations and saccades (Rayner, 1998; Hung, 2021). Furthermore, it is only during fixations that a reader is able to clearly focus on print (Strauss et al., 2009). Current technology typically defines the accuracy of identifying fixations as falling within 0.5° of the visual viewing angle or about two-three alphabetic characters (Rayner, 2009). That means, when the font size and distance between print and reader are constant, the technology can identify, very narrowly, the specific letters of a word the reader is looking at.

Operationalizing fixations in this way is relevant to reading researchers because it is at the point of fixation when the fovea has the largest degree of visual acuity. Outside of the fovea, the area referred to as the parafovea, acuity quickly diminishes. A reader must reposition the fovea (i.e., produce a saccade) in order to visually sample print outside of a fixation (Schotter and Rayner, 2013). However, current belief is that vision is suppressed during saccades leaving the reader, in effect, blind during these ballistic movements, though mental processing does continue (Schotter and Rayner, 2013; Irwin, 2004). Because of these facts, eye movement research has focused largely on what information is available to readers during fixations and on understanding how the mind is able to direct the reader, via saccades, from one fixation to the next (Hung, 2021; Schotter and Rayner, 2013; Rayner et al., 2006; Strauss et al., 2009).

The nonlinearity of scan paths

Interestingly, readers do not visually inspect every word while reading as demonstrated by the words not fixated on in Figs. 2 and 3. In fact, there is great variability between readers and within readers on how many words in a passage they will fixate (Paulson and Freeman, 2003). A number of factors contribute to this variability including reading proficiency, predictability of text, and purpose for reading (Krstić et al., 2018; Paulson and Henry, 2002; Rayner, 1998, 2009). Nevertheless, readers typically fixate between 60 and 80% of the words while reading (Paulson, 2000; Rayner, 1998; Just and Carpenter, 1987). That means that 20%–40% are not fixated, they are visually skipped. Paulson (2000) also demonstrated that a reader would vary the number of words fixated from sentence to sentence and connected this to the predictability and complexity of the text. For example, one of his participants fixated 50% of words in one sentence only to fixate 77% of the words in a sentence later on in the same passage. All the while, the reader verbally produced every word in the passage accurately (Paulson, 2000).

While many studies describe the fact that readers visually skip a significant portion of the words in a passage when reading silently and out loud, readers in both conditions report perceiving that they had visually inspected every word. Furthermore, in the case of reading aloud, in many instances readers actually verbally produce the words that are not fixated (Paulson, 2005; Schotter and Rayner, 2013). On those occasions when the reader's verbal productions differed from the print, eye movement research has shown that this verbalized word was typically visually inspected prior to the verbal production (Paulson, 2005; Paulson and Henry, 2002; Percle et al., 2020). The difference in these instances was on how long the words were fixated, how many fixations a specific word received, the order in which the reader fixated the words, and the questions the reader was asking of the text in the moment. The questions of overall fixations, fixation durations, and saccade length are considered global measures of reading difficulty in the study of eye movements. All of these variables are clearly influenced by how difficult the text is to the reader (Rayner et al., 2006; Schotter and Rayner, 2013; Seassau and Bucci, 2013), a finding with great relevance to special populations (Clark et al., 2021; Krstić et al., 2018). Eye movement research has focused on understanding how these variables reveal the processing efforts of a reader. It is therefore important to consider the general pattern of these variables produced by readers in order to understand how visual input contributes to processing of print.

Normative eye movement patterns from reading

Consistent with a social constructivist theoretical stance, our review of the literature in addition to our research and clinical experience indicates that variations in eye movement patterns of special populations reflect proficiency in reading rather than symptomatology of impairment (Damico et al., 2021; Goodman et al., 2016; Tracey and Morrow, 2017). That is, readers qualifying under special population guidelines demonstrate eye movement patterns consistent with neurotypical readers of similar reading proficiency (e.g., Ferguson et al., 2019; Nelson et al., 2008). Therefore, when considering normative data's application to special populations, educators and researchers should adapt interpretations to readers of similar proficiency.

Overall, normative data for average fixation duration have been identified for both silent and oral reading. In neurotypical individuals there appears a general decreasing length of fixation duration as readers develop into adulthood. Rayner and Juhasz (2006) reported general changes in eye movement data as neurotypical readers progress across early grade levels toward adulthood and implied greater reading proficiency. The average fixation duration time for silent reading in the early grades ranges between 348 and 425 ms in children 8–9 years of age (Tiffin-Richards and Schroeder, 2015) to an average duration for adults of 225–250 ms (Rayner and Juhasz, 2006). Oral reading shows a similar pattern with a consistent decrease in average duration time across from childhood to adulthood (Rayner, 2009; Seassau and Bucci, 2013).

Saccades last on average between 20 and 50 ms and comprise an average distance of six-nine characters (Rayner and Juhasz, 2006). Just like in oral and silent reading fixation duration times, there is general variability between the distance covered by readers during saccades. In oral reading, readers average approximately six-seven letters per saccade and seven-nine letters during silent reading. Rayner (1998) attributes this variability to the fact that the movement of the eyes is forced to slow as they wait for verbal production to catch up. This eye-voice gap, or eye-voice span, suggests that in most instances, the eyes are slightly ahead of the voice. Further, as the complexity of the reading process increases, studies have shown that the reader's fixation duration times increase and the distance between saccades decreases (Schotter and Rayner, 2013; Rayner et al., 2006; Rayner, 2009; Paulson and Freeman, 2003). Of relevance for understanding the reading process is the notion revealed in the fixation duration data. As demonstrated by differing fixation durations in Figs. 1–3, readers do not spend the same amount of time on each word. The length of visual inspection appears to be closely related to the reader's comprehension of the passage (Rayner et al., 2006; Paulson, 2000, 2005; Schotter and Rayner, 2013). In instances where comprehension difficulties arise, a number of studies have documented increases in fixation durations and a reduced saccade distance; a pattern also observed in beginning and impaired readers (Ashby et al., 2005; Rayner, 2009; Paulson, 2000; Duckett, 2002; Nelson et al., 2008; Percle et al., 2020; Clark et al., 2021).

Nonlinear characteristics of eye movements in reading

A second contribution of eye movement research to the understanding of reading is the realization that reading is not a strictly linear process (Hung, 2021). The history of eye movement research has informed us that readers do not begin at the first letter of the first line and (in English) read smoothly in an undeviating manner from left to right, top to bottom (Nelson et al., 2008; Paulson and Freeman, 2003; Paulson and Henry, 2002; Rayner, 1998, 2009). A simple illustration of this pattern of behavior comes from where the initial and final fixations produced by readers usually appear. Typically, the first fixation produced by a reader occurs approximately five-seven letter spaces from the beginning of the line of text. Returning to Fig. 1, we see the first fixation in line 308 occurring on the fifth letter and the final fixation occurring on the fourth to last letter. This would suggest the reader did not focus on the first few or last several letters of the passage they were reading.

Variables influencing fixation location

This data is consistent with the previously discussed notion that readers do not fixate every word in the text. Just as there exists great variability across readers and texts as to the exact number of words that receive fixations, there is also variability as to where in the words those fixations occur. The proficiency of the reader and the predictability of the text again contribute to fixation location within the word (Rayner et al., 2006). Carpenter and Just (1983) observed that while the majority of words are fixated, readers skip a significant percentage of words (60%–85% of words). This phenomenon appears to be informed by a wide range of variables, (Flurkey et al., 2008; Rayner, 1998; Percle et al., 2020). The tendency is for readers to fixate in the middle of words 4–10 letters long

if they are presented in isolation; however, the patterns become much more complex when words are presented in context. In fact, Paulson and Henry (2002) demonstrated that the purpose of the reading, in addition to word length and context, influences where the reader will fixate. Paulson and Henry found that readers reading a passage for the purpose of retelling the passage fixated in significantly different places within words and the overall text than they did when reading a passage for the purpose of answering multiple choice questions. How and where the reader visually inspected the words in the passage was directly influenced by the nature of the task.

Patterns of regression

Further demonstration that the reading process is not a strictly linear phenomenon is informed by the production of regressions. Regressions are moments when a reader's eyes move backwards in the text. While most of a reader's fixations move progressively forward through the reading passage, approximately 10%–15% of fixations in neurotypical literate adults can be classified as regressions. Approximately 28% of fixations produced by beginning readers in first Grade can be counted as regressions with their overall average fixation duration times equaling nearly 1/3 of a second (Rayner and Juhasz, 2006). Rayner (2009) further shows that beginning readers produce more than twice as many fixations per 100 words than adults. As the reader matures and assumedly becomes a more proficient reader, systematic reduction in all three of the represented indices of eye movement are observed. By adulthood, the fixation duration times, average number of fixations per 100 words, and percentage of regressive fixations appears to stabilize. The fact that readers across all grade levels and ages consistently produce a substantial percentage of backwards eye movements, while skipping some words and fixating multiple times on other words, further suggests that the reading process is not strictly linear. Further, there is a great variability as to where on the word a reader will fixate both on initial and successive fixations as well as on regressions (Rayner, 1998).

Variations in eye movement due to text difficulty and reader purpose

The percentage of regressions and the total fixation duration times on a single word are global measures of reading proficiency and are clearly influenced by the reader's construction of meaning through transactions with the text (Percle et al., 2020). More difficult texts typically lead to more regressive saccades, shorter saccades, and an increase in duration and number of fixations. The number of fixations classified as regressions increases when the reader encounters particularly long words or experiences comprehension difficulties or linguistic confusion (Rayner, 2009; Just and Carpenter, 1987). Beginning readers and readers identified as dyslexic also demonstrate a greater number of regressions and are consequently more likely to produce an increased number of overall fixations, average fixation durations, and overall gaze duration times (Ashby et al., 2005; Rayner, 2009). That is, the eye data reveal how less expert readers, regardless of label, tend to sample less efficiently as they search to construct meaning.

Additionally, the type of reading material and the reader's purpose in reading it influence these measures (Pollatsek and Rayner, 1989). The ways in which readers interact with text are individualized. They are selected purposefully through a predictive process based upon the reader's unique experiences with and knowledge of the world. Paulson and Henry (2002) demonstrated this individual power of purpose. In this study, researchers asked students to read from both a standardized test battery, which claimed to reflect authentic reading, and a reading passage, where the purpose was simply to retell the passage at the conclusion of the reading. The researchers found that while the percentages of fixations and number of words fixated were similar across conditions, it was visibly clear that the same reader transacted with the two passages in very different ways. The eye-tracking data from the testing scenario contained significantly more fixations and the reader bounced between the questions and the portion of the text from which the questions came. Contrasted with this were the eye movements during the authentic reading condition. Here, data reflected a smoother progression through the text with fewer regressive eye movements. The reader's fixations resulted in a higher number of regressions during the standardized assessment. The conclusions drawn from this and other studies (Percle et al., 2020) suggest that the purpose of the transaction with text greatly influences how the reader samples from the passage.

Which words are fixated?

A third contribution of eye movement research to the study of reading involves an increased understanding of how the brain directs the eyes to visually inspect specific types of information (Strauss et al., 2009). Emerging research demonstrates an evolving understanding of how environmental stimuli are obtained and understood by the brain. Central to this shifting paradigm is a developing understanding of how predictions drive interactions. These predictions are individualized as they are developed purposefully based on each individual's unique previous experiences and interactions with the world. In turn, the brain purposefully directs sampling of sensory information from the environment to confirm or determine a need to alter these predictions.

This movement toward understanding of the brain as selective and purposeful in its transactions with the environment finds a parallel in the application of this theoretical lens to the phenomenon of reading and other literacy practices. Applied to reading, individuals make predictions about the text and gather supporting information. This model described as a "sociopsychoneurolinguistic" transaction is demonstrated and supported by eye-tracking research (Goodman et al., 2016, p. 35). The written symbols constructing text, while phenomenal in their significance, are no different than other stimuli the brain interacts with in that they are primarily a source of information. The brain transacts with text through this same purposeful sampling of stimuli for the same meaning-making purpose of confirming predictions. These complex neurological actions are internal; however, they are not fully invisible to researchers. Eye-tracking methodology, when combined with indices of comprehension, allows us to visualize this meaning-making process in unique ways.

After studying readers engaged in sentence and essay reading tasks, Fisher and Shebilske (1985) concluded words not fixated were still processed. That is, the reader still makes sense of the specific skipped word. For example, function words prior to content words are more likely to be skipped while still being processed. This is especially evident when the content word is longer than the overlooked word (Drieghe et al., 2007; Rayner, 2009). Paulson (2000) demonstrated that readers clearly make predictions based upon their expectations and this is observed in number and location of the fixations produced. When a reader is able to predict the text, they require less of the visual information. Conversely, when the text is less predictable and more challenging for the reader, they fixate on those portions of the text that give them the information necessary to process the passage. For example, readers will fixate on content words about 85% of the time while fixating on function words only 35% of the time (Carpenter and Just, 1983; Duffy et al., 1988; Rayner, 2009).

Paulson (2000), looking at oral reading of longer, more complete passages, found his readers fixated on 79% of content words and 46% of function words. Paulson argues the pattern of fixating more on content words than function words reveals the processing patterns of the reader. He proposes that when readers skip function words, it is most likely because they have constructed the syntactic pattern of the sentence and need to focus on the semantic information conveyed. Additionally, function words are often short and represent fewer grammatical options for a reader. This may be why two-three letter words are fixated only 25% of the time while words of eight letters or more are almost always fixated and in fact more likely to receive multiple fixations (McConkie et al., 1989; Rayner et al., 1996). From these data it is clear that visual inspection of lines of text is not a uniform and smooth progression from left to right; rather, visual sampling in reading is a strategic act of focusing on the areas of print that will provide the reader with useful information.

Limitations and conclusion

Despite the obvious advantages of eye-tracking methods, there are specific limitations and methodological issues that need to be considered, specifically when considering applications to special populations. For decades, the cost of equipment needed for eye-tracking and the time needed to become proficient made practical applications to educational settings unrealistic. In recent years, technological advances have minimized these limitations; however, others remain. We choose to highlight three additional categories of Design and Initial Procedures, Data Processing and Analysis, and Interpretations.

Design and Initial Procedure

Determining gaze location is one of the earliest steps that helps with processing eye-tracking data and comes with its own challenges. Eye-trackers require calibration at the beginning of the task which can be disadvantageous when working with special populations due to challenges with attention and comprehension of motor commands. To manage this issue, it is suggested to use dynamic and musical animations during the calibration process to assist in attracting the individual's attention. Nevertheless, in research studies some individuals fail to complete the calibration task and need to be excluded (Venker and Kover, 2015). This exclusionary practice is inappropriate for classrooms where inclusion is a significant priority. Furthermore, there appears to be an art to tracking individuals. Not all individuals are able to be tracked. For example, children wearing glasses may have lenses incompatible with the eye-tracker's infrared light system. Individuals on medication may have delayed pupillary responses. These issues and others make calibration a challenge (i.e., Nyström et al., 2013). Recalibrating at different points during data collection can help minimize these errors at expense of time and it also breaks up the nature of the reading task and may potentially impact comprehension. Educators and researchers seeking to answer questions using eye-tracking technology must consider these variables in light of the contexts in which they work.

Data processing and analysis

A second limitation of eye-tracking specific to reading research in special populations considers data processing and analysis. Individuals from special populations demonstrate similar eye movement patterns as neurotypical readers of matching competence. Limitations in processing and analysis of eye-tracking in special populations therefore are largely methodological and logistical. For example, manual offline coding of eye-tracking data is done frame-by-frame and by hand and requires extensive training, especially when it comes to special populations. Due to the challenging nature of obtaining eye-tracking data from certain special populations, they are mostly excluded from the studies, preventing the examinations of some of the factors that led to their deficits (Venker and Kover, 2015). In many instances, the challenges qualifying individuals for classification as a member of a special population make data collection difficult and time consuming. For example, when working with children who have sensory issues, use of head mounted eye tracking equipment can lead to intermittent data collection. Similarly, children who experience cognitive limitations may have difficulty following instructions or staying engaged in the experimental tasks. During analysis, researchers must understand and take care to account for variability such as this. Additionally, these challenges in working with special populations may lead to data collection and analysis stages occurring over longer periods of time, resulting in lengthy time lapses, coder turnover, and reduction in the reliability of the data (Wass et al., 2014; Venker and Kover, 2015).

When it comes to data processing by eye-tracking devices, on the other hand, researchers should be aware of the default processing algorithms that define temporal and spatial parameters (i.e., defining limits of fixations and saccades). These algorithms may allow poor measurement qualities to mask the true eye movement patterns by developing undesirable relationships between the quality of the data and the length of fixation (Wass et al., 2013). There are options available to customize these algorithms but modifying the algorithms may dramatically change a study's results. In order to account for or control the behavioral and procedural challenges, often the nature of the reading tasks must be modified in order to collect useable data. Consequently, the ecological validity and authenticity of what the reading task represents is altered. The researcher's theoretical stance is typically what governs the acceptability of the modifications needed to work with special populations.

Interpretations

Eye movements are reflective of both the text variables and the reader, and each reader produces unique patterns of movements and miscues based on the organization of the text and how the text is being approached by the reader. Since reading is active and constructive, readers modify the text instead of modifying their minds. Eye movement research offers researchers and educators insight into the ways readers interact with texts. As demonstrated throughout this entry, interpretation of eye-tracking is dependent upon a researcher's theoretical understanding. Given purpose-driven interactions readers bring to texts, reading research, including that involving eye tracking, should be centered around authentic reading tasks. Additionally, in order to make sense of the eye tracking data, a reader's eye movements must be considered alongside some measure of their comprehension. When comprehension is included with eye tracking data, a map of readers strategies as they interact with texts can be constructed. Understanding the strategies employed by all readers can have significant implications on special populations whose paths to literacy success are often unique and nuanced (Nelson et al., 2020).

References

- Ashby, J., Rayner, K., Clifton Jr., C., 2005. Eye movements of highly skilled and average readers: differential effects of frequency and predictability. *Quart. J. Exper. Psychol. Sect. A* 58 (6), 1065–1086.
- Arya, P., Feathers, K.M., 2012. Reconsidering children's readings: insights into the reading process. *Read. Psychol.* 33 (4), 301–322.
- Ashby, J., Rayner, K., Clifton Jr., C., 2005. Eye movements of highly skilled and average readers: differential effects of frequency and predictability. *Quart. J. Exper. Psychol. Section A* 58 (6), 1065–1086.
- Carpenter, P.A., Just, M.A., 1983. What your eyes do while your mind is reading. In: *Eye Movements in Reading*. Academic Press, pp. 275–307.
- Clark, C., Nelson, R., Damico, J.S., Damico, H., Weill, C., Arrington, L., Percle, A., 2021. Shared Reading and EMMA. In: Liwanag, M.P., Kim, K., Martins, P. (Eds.), *Understanding Literacy Using Eye Movement Miscue Analysis in a Global World*. DIO Press, Inc, NY.
- Cop, U., Dirix, N., Drieghe, D., Duyck, W., 2017. Presenting GECO: an eyetracking corpus of monolingual and bilingual sentence reading. *Behav. Res. Methods* 49, 602–615.
- Damico, H., Damico, J.S., Nelson, R., 2021. Literacy and Literacy Impairments. *The Handbook of Language and Speech Disorders*, pp. 237–265.
- Drieghe, D., Desmet, T., Brysbaert, M., 2007. How important are linguistic factors in word skipping during reading? *Br. J. Psychol.* 98 (1), 157–171.
- Duckett, P.d.B., 2001. First-grade Beginning Readers' Use of Pictures and Print as They Read: A Miscue Analysis and Eye Movement Study. The University of Arizona.
- Duckett, P., 2002. New insights: eye fixations and the reading process. *Talk. Points* 13 (2), 16–21.
- Duffy, S.A., Morris, R.K., Rayner, K., 1988. Lexical ambiguity and fixation times in reading. *J. Mem. Lang.* 27 (4), 429–446.
- Ferguson, H.J., Black, J., Williams, D., 2019. Distinguishing reality from fantasy in adults with autism spectrum disorder: evidence from eye movements and reading. *J. Mem. Lang.* 106, 95–107.
- Fisher, D.F., Shebilske, W.L., 1985. There is more that meets the eye than the eyemind assumption. *Eye Move. Human Inform. Proc.* 149–158.
- Flurkey, A.D., Paulson, P.A., Goodman, K.S. (Eds.), 2008. *Scientific Realism in Studies of Reading*. Routledge.
- Freeman, A.E., 2001. The Eyes Have it: Oral Miscue and Eye Movement Analyses of the Reading of Fourth-Grade Spanish/English Bilinguals. Doctoral dissertation. The University of Arizona.
- Geertz, C., 1996. *After the Fact: Two Countries, Four Decades, One Anthropologist*, vol. 5. Harvard University Press.
- Goodman, K.S., 1967. Reading: A Psycholinguistic Guessing Game, vol. 6. *Literacy Research and Instruction*, pp. 126–135 (4).
- Goodman, K., Fries, P.H., Strauss, S.L., 2016. *Reading-the Grand Illusion: How and Why People Make Sense of Print*. Routledge.
- Hermena, E.W., 2021. Manipulating interword and interletter spacing in cursive script: an eye movements investigation of reading Persian. *J. Eye Reading Res.* 14 (1), 1–14.
- Huey, E.B., 1968. *The Psychology and Pedagogy of Reading*. The MIT Press.
- Hung, Y.N., 2019. Fifth grade students reading a Chinese text with embedded errors: an eye movement miscue analysis study. *Read. Psychol.* 40, 397–424.
- Hung, Y.N., 2021. The science of reading: the eyes cannot lie. *Int. J. Educ. Literacy Stud.* 9 (4), 26–31.
- Irwin, D.E., 2004. Fixation location and fixation duration as indices of cognitive processing. *Interf. Lang. Vis. Action: Eye Move. Vis. World* 217, 105–133.
- Jian, Y.-C., 2015. Fourth graders' cognitive processes and learning strategies for reading illustrated biology texts: eye Movement measurements. *Read. Res. Q.* 51 (1), 93–109.
- Just, M.A., Carpenter, P.A., 1987. *The Psychology of Reading and Language Comprehension*. Allyn & Bacon.
- Krstić, K., Šoškić, A., Ković, V., Holmqvist, K., 2018. All good readers are the same, but every low-skilled reader is different: an eye-tracking study using PISA data. *Eur. J. Psychol. Educ.* 33 (3), 521–541.
- Liwanag, M.P.S.U., Martens, P., Martens, R., Pelatti, C.Y., 2017. Examining a reader's meaning-making process of picture books using eye movement miscue analysis. *Liter. Res. Theory, Method, Pract.* 66 (1), 248–263.
- McConkie, G.W., Kerr, P.W., Reddix, M.D., Zola, D., Jacobs, A.M., 1989. Eye movement control during reading: II. Frequency of refixating a word. *Percept. Psychophys.* 46 (3), 245–253.
- Nelson, R.L., Damico, J.S., Smith, S.K., 2008. Applying eye movement miscue analysis to the reading patterns of children with language impairment. *Clin. Linguist. Phon.* 22 (4–5), 293–303.
- Nelson, R., Damico, J.S., Damico, H.L., Lynch, K., Abendroth, K.J., Weill, C., Arrington, L.E., Percle, A., 2020. Reading trajectories in children with language disorders: capturing variability of change over time. *J. Interactional Res. Commun. Disord.* 11 (2), 171–193.
- Nyström, M., Andersson, R., Holmqvist, K., Van De Weijer, J., 2013. The influence of calibration method and eye physiology on eyetracking data quality. *Behav. Res. Methods* 45 (1), 272–288.

- Paulson, E.J., 2000. Adult Readers' Eye Movements during the Production of Oral Miscues. The University of Arizona.
- Paulson, E.J., 2002. Are oral reading word omissions and substitutions caused by careless eye movements? *Read. Psychol.* 23, 45–66.
- Paulson, E.J., 2005. Viewing eye movements during reading through the lens of chaos theory: how reading is like the weather. *Read. Res. Q.* 40 (3), 338–358.
- Paulson, E.J., Freeman, A.E., 2003. *Insight from the Eyes: The Science of Effective Reading Instruction*. Heinemann.
- Paulson, E.J., Goodman, K.S., 1999. *Influential Studies in Eye-Movement Research* (Reading Online).
- Paulson, E.J., Goodman, K.S., 2008. Re-Reading Eye-Movement Research. *Scientific Realism in Studies of Reading*, pp. 25–47.
- Paulson, E.J., Henry, J., 2002. Does the Degrees of Reading Power assessment reflect the reading process? An eye-movement examination. *J. Adolesc. Adult Liter.* 46 (3), 234–244.
- Percle, A., Arrington, L., Flurkey, A.D., Damico, H., Weill, C., Damico, J., Nelson, R.L., 2020. Illuminating the complexity of oral reading fluency: a multiple lens approach. *Liter. Res. Theory, Method, Pract.* 69 (1), 358–376.
- Pollatsek, A., Rayner, K., 1989. Reading. In: Posner, M.I. (Ed.), *Foundations of Cognitive Science*. MIT Press, Cambridge, 1989.
- Rayner, K., 1978. Eye movements in reading and information processing. *Psychol. Bull.* 85 (3), 618.
- Rayner, K., 1997. Understanding eye movements in reading. *Sci. Stud. Read.* 1 (4), 317–339.
- Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124 (3), 372.
- Rayner, K., 2009. The 35th Sir Frederick Bartlett Lecture: eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62 (8), 1457–1506.
- Rayner, K., Juhasz, B., 2006. Reading processes in adults. In: Brown, K. (Ed.), *Encyclopedia of Language & Linguistics*, vol. 2, pp. 373–378 (10).
- Rayner, K., Sereno, S.C., Raney, G.E., 1996. Eye movement control in reading: a comparison of two types of models. *J. Exp. Psychol. Hum. Percept. Perform.* 22 (5), 1188.
- Rayner, K., Chace, K.H., Slattery, T.J., Ashby, J., 2006. Eye movements as reflections of comprehension processes in reading. *Sci. Stud. Read.* 10 (3), 241–255.
- Schotter, E.R., Rayner, K., 2013. Eye movements in reading. *Eye Track. Audiovis. Transl.* 83–102.
- Seassau, M., Bucci, M.P., 2013. Reading and visual search: a developmental study in normal children. *PLoS One* 8 (7), e70261.
- Smith, F., 1999. Why systematic phonics and phonemic awareness instruction constitute an educational hazard. *Lang. Arts* 77 (2), 150–155.
- Starr, M.S., Rayner, K., 2001. Eye movements during reading: some current controversies. *Trends Cognit. Sci.* 5 (4), 156–163.
- Strauss, S.L., Goodman, K.S., Paulson, E.J., 2009. Brain research and reading: how emerging concepts in neuroscience support a meaning construction view of the reading process. *Educ. Res. Rev.* 4 (2), 021–033.
- Tiffin-Richards, S.P., Schroeder, S., 2015. Word length and frequency effects on children's eye movements during silent reading. *Vis. Res.* 113, 33–43.
- Tracey, D.H., Morrow, L.M., 2017. *Lenses on Reading: An Introduction to Theories and Models*. Guilford Publications.
- Venker, C.E., Kover, S.T., 2015. An open conversation on using eye-gaze methods in studies of neurodevelopmental disorders. *J. Speech Lang. Hear. Res.* 58 (6), 1719–1732.
- Wass, S.V., Smith, T.J., Johnson, M.H., 2013. Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behav. Res. Methods* 45 (1), 229–250.
- Wass, S.V., Forssman, L., Leppänen, J., 2014. Robustness and precision: how data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy* 19 (5), 427–460.
- Zang, C., Liversedge, S.P., Bai, X., Yan, G., 2011. Eye movements during Chinese reading. In: Liversedge, S.P., Gilchrist, I.D., Everling, S. (Eds.), *The Oxford Handbook of Eye Movements*. Oxford University Press, Oxford, UK, pp. 961–978.

Further reading

- Goodman, K., Fries, P.H., Strauss, S.L., 2016. *Reading-the Grand Illusion: How and Why People Make Sense of Print*. Routledge.
- Liwanag, M.P., Kim, K., Martins, P., 2021. *Understanding Literacy Using Eye Movement Miscue Analysis in a Global World*. DIO Press, Inc, NY.
- Paulson, E.J., Freeman, A.E., 2003. *Insight from the Eyes: The Science of Effective Reading Instruction*. Heinemann.

Issues and current trends of research in braille for people with visual impairments: questions and implications

Sunggye Hong, University of Arizona, Department of Disability and Psychoeducational Studies, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

History of braille	423
Braille system and contractions	423
Adoption of braille in the United States	425
Unified English Braille (UEB)	425
Current practices for braille instructions	427
Reading development	427
Braille writing	428
Spelling	429
Research implications of the future of braille	429
Is it necessary to retain braille contractions and its governing rules?	430
Why is it so slow to read braille?	431
Why is braille reading cognitively and tactually demanding?	431
Conclusions and future directions	432
References	432

Communication is an integral component of successful academic achievement. The amount of available knowledge and information has exponentially increased over the last several decades. Many types and forms of literacy media have been devised or invented but there is no medium other than braille that has been most widely adopted among people with visual impairments.

This chapter discusses the importance of braille literacy by reviewing existing literature and suggesting ways that research activities in braille reading be implemented. First, the history of braille along with key literature and events associated with adoption of braille code is discussed. Second, current issues and significant aspects of braille reading are noted. Finally, questions about the future of braille along with emerging topics of braille research are suggested.

History of braille

Students with visual impairments have unique needs when compared to print readers with regards to literacy development. Just as the primary medium of reading and writing for sighted students is print, that of students with visual impairments is braille.

Although Louis Braille did not live to see braille adopted as the official medium for people with visual impairments, it has become internationally adopted soon after his death (Bogart, 2009; Irwin, 1970). Braille enabled people with visual impairments to appreciate literacy skills, which is an important tool to participate equally in this society. Though braille is only used dominantly among people with visual impairments, efforts had been given in developing and creating alternative tactual methods of reading and writing long before its invention.

They took the form of writing on wax tablets, providing letters carved in wood, teaching a blind person to follow the shape of letters in grooved woodblocks, forming letters by wire, using knots of varying thicknesses and in varying distances on a thread, forming letters by pins stuck into felt or similar material, pinpricking letters tangible in relief, and so forth (Lowenfeld et al., 1969).

Practical means of beginning of literacy for people with visual impairments did not commence, however, until the invention of braille by a Frenchman named Louis Braille.

Braille system and contractions

Braille is comprised of 26 letters, punctuations, and contractions. Braille has two horizontal and three vertical dots in a cell. The combination of six dots can express sixty-three different symbols. Table 1 presents 26 alphabet letters and commonly used punctuations. A limitation introduced by the lack of variations that can be expressed by a single cell required to use a dot formation for more than a letter. For example, dot 14, presenting the letter c is used to write 3 or the word “can” depending on the context where it is used.

Braille takes more space than print. In 18th century when braille was first created, production cost was directly related to space of a material and Lewis Braille came up with a set of contractions, which were comprised of letter combinations, different dot formations, or combinations of both. Contractions were believed to make braille reading and writing faster. Several studies have been conducted to answer these questions. Nilsson (1982) examined the reading speed of braille. After comparing reading speed using

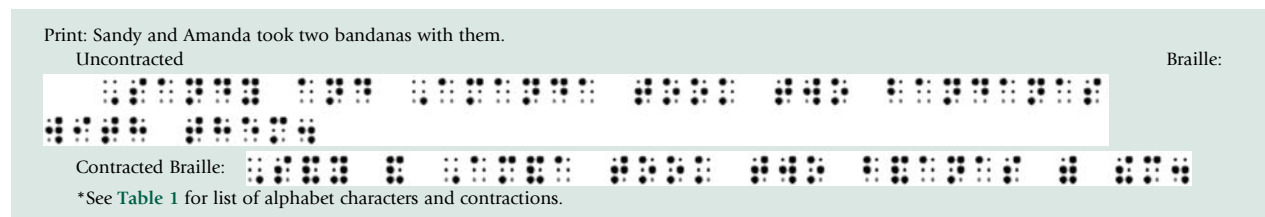
Table 1 Braille alphabet letters and punctuation signs.

A L P H A B E T	a	⠁	NUMBERS	1	⠼	Typeform	italic symbol	⠠	P U N C T U A T I O N	comma ,	⠂
	b	⠃		2	⠽		italic word	⠡		period .	⠆
	c	⠉		3	⠿		italic passage	⠢		apostrophe ’	⠄
	d	⠇		4	⠻		italic terminator	⠣		colon :	⠒
	e	⠑		5	⠝		bold symbol	⠨		dash _	⠤
	f	⠑		6	⠺		bold word	⠦		long dash --	⠤⠤
	g	⠒		7	⠼		bold passage	⠧		Exclamation mark !	⠆
	h	⠈		8	⠼		bold terminator	⠧		hyphen -	⠤
	i	⠑		9	⠻		underline symbol	⠨		question mark ?	⠆
	j	⠑		0	⠼		underline word	⠦		Semicolon ;	⠒
	k	⠑	INDICATORS				underline passage	⠧		ellipsis ...	⠒⠒⠒
	l	⠑		Numeric	⠼		underline terminator	⠧		forward dash /	⠤
	m	⠓		letter	⠼		script symbol	⠨		backward dash \	⠤
	n	⠓		word	⠼		script word	⠨		opening outer quotation mark “	⠆
	o	⠑	Capital	passage	⠼		script passage	⠨		closing outer quotation mark ”	⠆
	p	⠑		capital terminator	⠼		script terminator	⠨		opening inner quotation mark ‘	⠆
	q	⠑								closing inner quotation mark ’	⠆
	r	⠑	Grade 1	symbol	⠼						
	s	⠑		word	⠼						
	t	⠑		passage	⠼						
	u	⠑		Grade 1 terminator	⠼						
	v	⠑									
	w	⠑									
	x	⠑									
	y	⠑									
	z	⠑									

uncontracted and contracted braille, the author concluded that contracted braille demonstrated a higher mean score of reading speed, suggesting that contracted braille contributed to the faster reading of braille. While participants in the study were all contracted braille readers and having them to read uncontracted braille would negatively influence their reading speed and materials used for the study were back-translated from contracted to uncontracted braille, findings of this study supported the idea that contractions would make braille reading speed faster. A measure that may warrant more scientific findings would need to use words that were otherwise typically not contracted. On the other hand, [Troughton \(1992\)](#) conducted a research study to determine the actual differences of reading speed between uncontracted and contracted braille. On this study where 125 people participated, she used materials that were chosen from words that could not be contracted. Selected uncontracted words allowed the researcher to isolate potential noise that could be introduced by using back-translated materials. For example, since “dog” does not contain any contraction, regardless of previous knowledge with contractions, readers were not disadvantaged. If “people” was provided as an example, those who already learned contracted braille should read spelled out version of the word that was different from their conventional usage. She found that 113 participants read faster when using uncontracted braille and only 11 participants were faster in contracted braille. This finding supports the idea that faster reading speed in uncontracted braille is expected when materials are carefully selected.

Another study that added a significant finding into this topic was conducted by [Lorimer and Tobin \(1979\)](#). They conducted a study to identify a relationship between the number of contractions and reading speed, by employing four different combinations of braille contractions. Three of them reduced the number of contractions, and one expanded the code. Their findings suggested that reduction of contractions in codes did not influence students’ reading speed. On the other hand, the code with a greater number of contractions was responsible for lowering the rate of reading speed. When considering that [Lorimer and Tobin \(1979\)](#) and [Troughton \(1992\)](#) controlled unexpected variables introduced by back-translation and selection of participants, it is hinted that the use of uncontracted braille did not negatively influence reading speed.

A major advantage of contracted braille to consider is its space saving and reduction of production costs. [Kederis et al. \(1965\)](#) found that 31% of space could be saved by using contracted braille and, when including numbers, punctuation, and special signs, 26.5% of space could be saved by employing contracted braille. [Durre \(1996\)](#) randomly selected 39,717 words to analyze space-saving features of contracted braille. The result suggested that contracted braille took up to 20.42% less space than print. What is also important to note is that the top-ranked 100 contractions contributed to 95% of total space-saving. For example, as shown below, commonly used contractions such as “and” “with” or “the” contribute greatly in reducing the space, when compared to print.



A couple of points to consider are the extent to which the space-saving features of contracted braille can be implemented in actual braille materials. For example, children's books, including textbooks for early graders, are not as thick as those used by older students and adults, and they often include many tactually adapted drawings and pictures of objects which may cancel space-saving due to the use of contractions. In addition, braille production's costs and processes have evolved and changed over the years and it can be concluded that the identified level of space-saving would be accountable for marginal differences. The use of refreshable braille displays, which are designed to express braille electronically by raising and lowering different combinations of pins in braille cells, may be contributable for further reducing the gaps between uncontracted and contracted braille. Fig. 1 illustrates a diagram of a refreshable braille display and its mechanism for producing braille.

Adoption of braille in the United States

Braille became the official medium for students at the Institute for the Blind in Paris in 1854, 2 years after the death of Louis Braille. Later, braille was adopted in the United States in 1860 and became the official medium at the Missouri School for the Blind. At the time when braille was introduced in the United States, several other tactile codes were already in competition with one another. This conflicts between braille and other systems (including the Moon type, New York dots, and Boston type) are known as "The War of the Dots" (Irwin, 1955, p. 1), "The Battle of the Types" (Farrell, 1956, p. 105), and "The Battle of the Contractions" (Lowenfeld et al., 1969, p. 8).

New York Point, one of the most widely accepted tactile code for people with visual impairments other than braille, used one to four pairs of points set side by side, each containing one or two dots. New York Point was devised by William Wait, a teacher in the New York Institute for the Education of the Blind, and its composition was done through adding additional dots vertically. This was known as variable base. New York Point used very few capitals and common punctuation signs. Fig. 2 shows lower case letters in New York Point.

Another code used at the time, though not as popular as New York Point, was modified braille. The idea of the code was to assign commonly used letters for braille formation with less number of dots. Farrell (1956) reported that the Modified system could be considered as the most scientific method of arranging dots. "The War of the Dots" was, in most cases, a battle only between braille and the New York system. In 1909, New York Public Schools decided to adopt braille as an official medium for reading and writing of people with visual impairments. In 1916, agreement settled on English Braille and this is mainly due to its capacity for expressing punctuations and the large amount of written material available in English Braille.

Unified English Braille (UEB)

Since then, literary braille code maintained most of its rules and contractions while conveying mathematical and scientific notations has evolved in the form of the Nemeth Code, a code for Braille Mathematics and Scientific Notation. Developed in 1946, Nemeth code was aimed at creating set of experiences of readers that printed text would be translated accurately and technical literature would be presented and read in braille (Nemeth, 1972). Some examples of the Nemeth code are shown in Table 2. Although Nemeth code did not go through extensive revisions, discussions about literary braille with regards to its ambiguity and necessity of additional signs and letters were beginning to emerge (Cranmer and Nemeth, 1991; Jackson et al., 1993).



Fig. 1 A refreshable braille display in acting with a hand, showing the position of reading braille.



Fig. 2 Lower case letters in New York Point.

Table 2 Nemeth code numeral signs, signs of operation, and signs of fraction.

NUMBERS	1		OPERATION SIGNS	Plus	+	
	2			Minus	-	
	3			Multiplication: Cross (Cartesian product)	x	
	4			Multiplication: Dot		
	5			Division (divided by)	÷	
	6			Plus or Minus	±	
	7			Minus or Plus	∓	
	8			Plus Followed by Minus	+ -	
	9			Minus Followed by Plus	- +	
	0			Minus Followed by Minus	--	
FRACTION SIGNS	Fraction Lines (over, divided by): Horizontal Simple Fraction Line			—		
	Fraction Lines (over, divided by): Diagonal Simple Fraction Line			/		
	Fraction Lines (over, divided by): Horizontal Complex Fraction Line			—		
	Fraction Lines (over, divided by): Diagonal Complex Fraction Line			/		

The initial discussion of a unified braille code took place through the Braille Authority of North America (BANA), the standardizing body of English Braille orthography in the United States and Canada (Bogart et al., 2000). The primary goal of a unified English braille was to better facilitate the use of technology in the production of braille and to reflect changes in printed letters and signs (BANA, 2012). The International Council on English Braille (ICEB), began work to develop a unified braille code (Bogart et al., 2000). Unified English Braille (UEB), was adopted the US in 2016 (D'Andrea, 2013).

UEB makes a number of significant changes to the original English braille. Changes in contractions and ways that the contractions are used, punctuation marks, ways that computer symbols are written, rules for capital sign and indicators for expressing changing text attributes, use of numbers in literary braille, and so on are a few examples that the UEB brought to many braille readers. Potential impacts of the adoption of UEB will need to be studied from multiple perspectives.

Current practices for braille instructions

Most students who are visually impaired may use one medium such as print or braille to read and write, but some are taught both when their visual condition is likely to change or when students are unable to read print efficiently (Lusk and Corn, 2006). In recent years, teachers of students with visual impairments have begun conducting learning media assessments to help in determining the appropriate medium for reading and writing. Learning media assessment, initiated by Koenig and Holbrook (1995), combines the primary use of sensory channels, measures of reading speed, and use of an informal reading inventory through observations to identify the student's primary medium for literacy. This section reviews key research studies related to instructional aspects of braille reading, writing, and their implications.

Reading development

Numerous differences exist in literacy development of students with visual impairments, and it will be necessary to examine common stages of reading and ways that such stages can be extended to braille reading. For students who are in the early stage of reading, it requires additional intervention to make sure that they have sufficient preliteracy experience to begin a formal braille reading program. Wormsley (1997) noted that early exposure to braille was important, regardless of readiness of a child with visual impairments, toward official reading intervention. For instance, environmental modifications such as adding braille labels on common household objects, placing name cards at home, and making signages available in braille could allow children to explore braille in their daily lives (Drezek, 1999; Koenig and Holbrook, 2000). Unlike their sighted peers whose family members can read and write print, many family members may not read and write braille and it would be difficult for them to provide similar support (Craig, 1996).

Craig (1996) reported that, among 112 study participants, approximately one-third of them could read braille. Limitations paused by lack of environmental braille and lack of family involvement in creating such environment for students with visual impairments will serve as an obstacle for braille rich environment and call for the needs of early intervention by teachers of students with visual impairments. Otherwise, children with visual impairments may be deprived of developing early literacy experiences.

An interesting aspect of literacy development and their relationship with braille is that decoding skills might be influenced by contracted braille. Not only do they require sufficient understanding of the language, but they also are exposed to an additional challenge. The letter *c* has a different pronunciation, depending on where it is located in a word. *C* in the word *cab* makes a *k* sound but, in the word *ice*, it makes an *s* sound. Before appropriate development of understandings of phonetic rules, students may be confused by the fact that the letter *c* can be used as in *can* and in *cat*. Though many students with visual impairments have adjusted to such variances, it cannot be assumed that confusion associated with contractions does not create additional barriers in literacy development.

O'Donnell and Wood (1999), proposed five stages of reading: emergent reading, initial reading, transitional reading, basic literacy, and refinement. The authors stated most preschoolers and kindergartners could be classified in the emergent reading stage, which functions as the gateway to literate life. During this stage, students familiarize themselves with print. In the second stage, students learn the relationship between print and spoken language. Furthermore, students expand their knowledge to unfamiliar words. The next stage is aimed at reading more fluently and making sense out of reading. In the stage of basic reading, students work to learn from reading. Students read in order to gain knowledge from materials. The final stage is for refining previously learned reading competencies as well as learning thinking competencies.

Examining the importance of reading and literacy development of students who use braille as the primary medium suggests close collaboration between teachers of students with visual impairments and classroom teachers. Wittenstein (1994) suggested that the introduction of braille code only to preservice teachers of students with visual impairments led to negative attitude toward braille. Teaching methods of braille along with ways for collaboration with classroom teachers are the key component of positive attitude of teachers of students with visual impairments toward braille. Furthermore, according to the U.S. Department of Education, Office of Special Education and Rehabilitative Services, it is essential that a student receives ongoing support from a professional, trained for becoming proficient in Braille reading and writing (U.S. Department of Education, 2000). Yet, as mentioned, braille code does not correspond directly with printed letters. Collaborations with reading specialists or classroom teachers will be the key component of successful intervention for reading and literacy development (Munro and Munro, 2013).

Braille writing

Another difference is in the way that braille materials are produced and the method that braille writing is taught. Students using braille as a primary mode of reading and writing begin using a Perkins braillewriter, a device that embosses braille characters on paper with the use of six key entry. [Erin and Wright \(2011\)](#) reported that 39 students who participated in their study used the Perkins braillewriter until second grade. Only two students used assistive technology options for producing braille materials. The Perkins braillewriter requires more physical strength as well as mastery of processes such as loading braille paper, which is different from print users where the fine motor skills are reinforced for handwriting. Alternatively, a device called the *slate and stylus* is used for writing braille by students, but is criticized for its mechanism where dots are inverted between writing and reading and users need to write from left to right to compensate such differences. [Fig. 3](#) and [Fig. 4](#) show a picture of a slate and stylus and a Perkins braillewriter. More and more teachers actively adopt assistive technology options in writing braille, using braille notetakers, electronic devices that use refreshable braille displays and six key entry.

On the Perkins braillewriter and the braille notetakers, Braille letters and contractions are created by pressing combinations of six keys, except on some braille notetakers where a qwerty keyboard system is used. The letter “a” requires pressing only a single key, whereas the symbol for the word “for” is written by pressing all six keys simultaneously. Insufficient pressure on a key or misplacement of a finger on key combination can result in braille errors. For example, the letter “f” can become the letter “p” only if one dot “dot3” is added. Careful assessment needs to be made so that mechanical and syntactic errors are differentiated. Newly developed assistive technology options such as braille notetakers can alleviate certain challenges associated with the physical nature of the Perkins braillewriter and can provide instantaneous feedback to students through speech.

Existing studies reporting the use of assistive technology options for writing braille by students with visual impairments suggest satisfaction and motivation. However, cautions should be given where few studies found quality outcomes and user performance when assistive technology options were deployed. Discussions on the observed differences of an assistive technology option such as electric braillewriters go back to a study conducted by [Koenig and Ashcroft \(1983\)](#). They found no differences in writing speed or accuracy. More recently, [Bickford and Falco \(2012\)](#) reported results comparable with the previous study. They found no clear advantage in outcomes for the early use of electronic braille notetakers for writing. They noted meaningful progress in writing samples for students who produced written material electronically, nonetheless the measurement was not significant. Another study focusing on satisfaction and motivation of students using assistive technology options was conducted by [Cooper and Nichols \(2007\)](#). They



Fig. 3 A diagram of a Perkins brailier with a hand positioned to push keys.

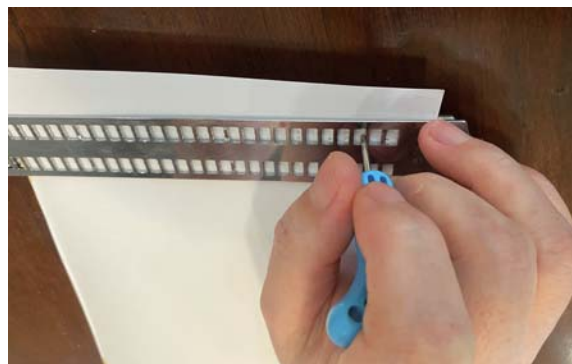


Fig. 4 A user writes braille using a slate and stylus.

reported positive perceptions by teachers of students with visual impairments and students who used the Mountbatten electronic braillewriter. No measures of writing performance were reported as part of the study.

A more current investigation by D'Andrea (2010) looked into ways that technology options were utilized by students with visual impairments. She reported that 77 high school students with visual impairments were interviewed for the study. Participants used between four and eight devices to help them accomplish their tasks. Tasks such as writing letters, labeling, or editing, were accomplished by using braille. On the other hand, they indicated to use print for sighted instructors and classmates. As students grew older, more auditory tools and speech were adopted for accomplishing their literacy tasks. D'Andrea (2010) asserted the need for more research on writing by students with visual impairments in the areas such as decreased use of slate and stylus and technology use. She also suggested the need for additional research in the processes of writing and development of authorships among students with visual impairments.

Spelling

Research findings about braille readers demonstrate relatively minor differences in spelling performance compared to print readers. Some studies even report better performance in spelling by students who use braille. A major concern to focus here is the use of contractions in instruction and testing of spelling.

There has been significant concern about the effects of contracted braille on spelling skills of braille reading students. "By using contractions, we will gradually lose our ability to spell the words properly" (Nilsson, 1982, p. 199). Johnson (1989), for example, reported that those who learned uncontracted braille in the early grades exhibited better spelling skills than those who received instruction using contractions. The findings implied that students using contracted braille may not develop their full potential in spelling due to the difficulty of contractions. For example, the word "would" in braille is contracted as "wd". If a student has never learned how to spell "would" and repeatedly written "wd" in braille, one may not correctly spell the word correctly. Godshall (1961) conducted a study to examine spelling skills of students with visual impairments, ranging from third grade to ninth grade. The author concluded that students using braille were above grade level when asked to write in uncontracted braille. Two standardized tests were administered to 227 students, attending to residential school programs. Students did better when they were asked to write in uncontracted braille (Godshall, 1961). The implication of this study is significant. The result suggests one of two things: teachers in classrooms are usually giving spelling tests in uncontracted braille and students might be accustomed to the practice, or students were able to master full contractions and rules of contracted braille.

Koenig (1987) confirmed better spelling performance of students who read braille. In his study, spelling skills of the two younger groups outperformed peers who were using print. Though the high school group was similar to print readers, these findings were supported by subsequent studies that found average to superior spelling skills by braille readers (Clark-Bishke and Stoner, 2009; Clark and Stoner, 2008; Emerson et al., 2009; Grenier and Giroux, 1997). Emerson et al. (2009, p. 618) stated, "A surprising finding was that spelling was a strength for these young braille-reading children, regardless of whether they used contractions or not ...".

Writing errors related to the braille code are described separately in some studies. Koenig (1987) reported the lower four dots of the braille cell such as "be, by, to, in, and was" accounted for substantial source of observed errors. It should be noted, however, that some of these contractions were eliminated in UEB. Erin and Wright (2011) presented that the majority of errors collected from samples produced by students were unrelated to the braille code. Phonetic errors that reflected irregular spellings were the primary source of errors, where a dot or dots were incorrectly placed on the wrong side of the braille cell.

Though significant differences exist in braille and cautions should be given for ways that literacy development of students with visual impairments are ensured, it is important that collaborations be facilitated between teachers of students with visual impairments, classroom teachers, and reading specialists. Familial involvement is equally important in creating braille rich environment so that proper pre-braille skills and foundation of literacy can be constructed.

Research implications of the future of braille

Several topics about braille and its implications for reading and writing development of students with visual impairments have been discussed. However, a number of other topics pertaining to the future of braille exist. The decline in the number of users who use braille as their primary medium has been reported in the existing literature. For example, Miller (2002) indicated that the number of braille readers has been decreased over the years. American Printing House for the Blind (2018) indicated that approximately 10% or about 5000 students with visual impairments identified braille as their primary medium. Rodenberg (1955) reported that more than 8000 children and 20,000 adults used standard English braille in the US and equivalent number of users existed in England. While it is difficult to ascertain the exact number of people with visual impairments in the US, it is clear that the proportional representation of braille users within the visual impairment population has declined.

Though a particular reason for the decline has not been identified, several potential factors might contribute to the overall implications of such observation. First, people who are elderly or who have multiple impairments are included when calculating the total number of people with visual impairments. Often, people who are elderly may not feel the need to use braille regularly, or their physical conditions may not allow them to use braille as the primary medium of reading and writing. Second, a lack of teachers for students with visual impairments may make it difficult for students with visual impairments to receive adequate instruction in braille learning. For example, In the 40th Annual Report to Congress on the Implementation of the Individuals with Disabilities

instruction and examinations of the roles that braille contractions play in braille learning be conducted to identify potential impacts of the use of contractions in braille. Now that adoption of refreshable braille display and electronic text negate the impacts of space saving and cost reduction of braille contractions, well designed research studies on this topic will be of great help in setting up an important foundation of the code.

Why is it so slow to read braille?

Braille is often considered to be slow when compared with print reading. Though many factors may contribute to the discrepancy, existing studies accuse inherent difference of utilized sensory channels as a potential cause. Visual readers read as many as 15 characters at a fixation while braille requires character level contact. Consequently, the amount of information retrieved through the visual sense is larger than that received through touch. For example, Foulke (1991) suggested that braille readers read 65–76 words per minute at the junior high school level and 80–90 words per minute at the senior high school level. Nolan and Kederis (1969) observed similar outcome and reported that the oral reading speed of students with visual impairments was 90 words per minute. Lowenfeld et al. (1969) reported reading rates of 116 words per minute for students in residential schools and 149 words for students in a public school setting.

Foulke and Smith (1973) and Troxel (1967) further examined this issue by creating an experience where print readers could only see a single character. Troxel (1967) gave participants a tube, with which only one character is seen at a time. Most participants reported that it was very difficult to read character by character. Some participants who reached a comfortable reading level exhibited 65–75 words per minute. Troxel (1967) concluded that reading character by character might explain the slow rate of braille reading.

Reported slow reading speed of braille led to some studies where efforts to increase the speed and to identify an interventional strategy for braille reading were implemented. Findings of such studies were inconsistent. Furthermore, the majority of studies was criticized by the lack of a control group and too many uncontrolled variables, caused by small sample size. For example, Umsted (1970) reported that faster reading speed was observed when providing additional practices in the area of short-form words and contractions among 72 participants. Participants in the better performing group gained as many as 25 words per minute, leading to the conclusion that the interventional strategy was effective and contributed to the improvement of reading speed. A different result was reported by Flanigan and Joslin (1969) where reading speed of 26 participants was increased but with more reading errors. They indicated a short span of training as a potential cause of this discrepancy. Other studies (Crandell and Wallace, 1974; Henderson, 1967; Olson et al., 1975) were conducted to delineate an effective means of providing interventional training for slow reading rate of braille readers. Repeated reading received attention from researchers as an instructional method in improving braille reading speed. For example, Stanfa and Johnson (2017) reported that three participants on their study showed notable increase in the reading speed when trained with repeated reading. Yet, none of the studies yielded significant differences mainly because the gain was not statistically significant or because extraneous variables were not controlled.

It is suggested that there should be a large-scale study to establish reliable speed of braille reading, which will serve as a foundation for designing and implementing interventional strategies for tactile recognition, mechanics of braille reading, and differences of employed modality (Mousty and Bertelson, 1985; Wright et al., 2009). In addition, factors and predictors of braille reading and associated variables need to be examined closely.

Why is braille reading cognitively and tactually demanding?

While a substantial body of literature exists on vision and its function for reading, touch has received relatively little attention. Millar (1997) identified passive or active aspects of touch. In active touch, one's hands, arms, and body are used to control the types of information obtained through the channel. On the other hand, information is gathered by cutaneous receptors which receive information when the skin is touched by an object in passive touch. It is known that fingertips have the highest tactile acuity (Millar, 1997). Fingertips exhibit the most accurate discrimination of a two-point threshold (to distinguish one object from another), which can identify as little as two millimeters of difference. It is interesting to note that braille dot density is around the borderline of this threshold. In the developmental stage, tactile discrimination follows this order.

1. Large solid geometric shape
2. Flat (puzzle pieces) figures
3. Embossed-dot geometric figures
4. Embossed-dot line figures
5. Braille figures (Kershman, 1976–1977)

Development of tactile sense begins with large object recognition before progressing to small and fine objects (Kershman, 1976–1977; Barraga and Erin, 1992). A child explores from large objects to small objects, from simple objects to complex objects, and from general objects to detailed objects. A child is ready to be exposed to braille when these skills have been sequentially developed (Kershman, 1976–1977). Characteristics such as passive touch versus active touch, the function of different hemispheres when reading braille, sensitivities of different body parts and their tactual acuity, and registration of the tactile sense to memory have received the most attention from researchers (Foulke, 1964; Kusajima, 1974; Millar, 1997; Nolan and Kederis, 1969).

Another area of discussion is about tactual sensitivity, applicable to people who have learned braille adventitiously or those who have other health conditions. A paucity of qualified personnel to work in braille coupled with the lack of understandings in this area

of research makes it difficult to devise effective programs and instructional strategies. Both perceptual and cognitive factors underlying proficiency in reading braille need to be studied, and discussions about characteristics of touch and its implication should be well understood.

Conclusions and future directions

This chapter examined the importance of braille literacy. Through literature review and analysis, this chapter suggested current issues in braille research and ways that research activities in braille would inform the field with regards to the unique needs of reading for people with visual impairments. First, the history of braille along with key literature and events associated with adoption of braille code was discussed. Second, current issues and significant aspects of braille reading were examined. Finally, questions about the future of braille along with emerging topics of braille research were presented.

First, though braille contractions have served an important role, a well-designed longitudinal study on the impact of braille literacy has not been conducted. When considering that spelling, reading speed, reading mechanics, and reading development are intimately related to contractions, the field needs to ponder the influence of braille contractions and their long-term outcomes for reading. The scope of such efforts may not be limited only to a unidimensional application of braille contractions but should expand to sophisticated implications of the roles of braille contractions for literacy development. Functional use of braille by people with multiple disabilities, refreshable and electronic braille, and a decline in the number of braille readers in the United States call for quick responses, informed by research.

Second, development of additional technology options for delivering braille will be of interests to many stakeholders. Though refreshable braille displays allowed readers with visual impairments to gain access to computerized braille, their applications are yet restricted to line by line display of braille text. Diagrams, tables, and pictures often require much larger display surface and current technology options cannot seamlessly accommodate such demands. Multi-line braille displays and haptic displays can potentially provide much more flexible options and development efforts and research activities to document these needs should be carried.

Third, it is disappointing that Many students with visual impairments who use braille exhibit slow reading speed when compared to their sighted counterparts. It is unclear whether slower rate of reading demonstrated by braille users are due to differences in visual comprehension or limitations associated with braille. However, what is more crucial is the need of a study, documenting the severity of such reading delays. While anecdotal evidence of slow reading of students with visual impairments have been reported, we do not know how fast or how slow a child using braille would read by grade level, to date, there is no data to substantiate reliable and accurate braille reading speed of people with visual impairments. In part, standardized braille speed by grade and or age may be very difficult to establish because of diverse visual acuity, experiences in reading, and additional disabilities. However, the lack of baseline data for braille reading speed makes it difficult to establish appropriate plans for mediating problems associated with slow reading speed.

There are notable concerns as to whether braille will eventually be replaced by new technology such as screen and print reading devices, which convert text into spoken words, or will become obsolete by rapid development of technology. However, no one argues that word processors make print obsolete. People with visual impairments need to have varying options for their literacy media. Everyone has the right to have access to information in their preferred format. For many, it is braille.

References

- American Printing House for the Blind, 2018. Annual Research Report. Louisville KY.
- Ashcroft, S.C., 1960. Errors in Oral Reading of Braille at Elementary Grade Levels. Ph.D. University of Illinois.
- Barraga, N.C., Erin, J.N., 1992. Visual Handicaps and Learning. Pro-Ed, Austin, TX.
- Bickford, J.O., Falco, R.A., 2012. Technology for early braille literacy: comparison of traditional braille instruction and instruction with an electronic notetaker. *J. Vis. Impair. Blind.* 106 (10), 679–693.
- Bogart, D., 2009. Unifying the English braille codes. *J. Vis. Impair. Blind.* 103 (10), 581–583.
- Bogart, D.E., Cranmer, T.V., Sullivan, J.E., 2000. Unifying the braille codes. In: Dixon, J. (Ed.), *Braille Into the Next Millennium*, first ed. National Library Service for the Blind and Physically Handicapped, Library of Congress, Washington, DC, pp. 160–181.
- Bozeman, L.A., Zebehazy, K., 2014. Personnel preparation in visual impairment. In: Sindealar, P.T., McCray, E.D., Brownell, M.T., Lignuagris-Kraft, B. (Eds.), *Handbook of Research in Special Education Teacher Preparation*, first ed. Routledge, New York, pp. 353–368.
- Braille Authority of North America, 2012. The Evolution of Braille: Can the Past Help Plan the Future? Online. Available at: http://www.brailleauthority.org/article/evolution_of_braille-part1.pdf.
- Clark, C., Stoner, J.B., 2008. An investigation of the spelling skills of braille readers. *J. Vis. Impair. Blind.* 102 (9), 553–563.
- Clark-Bischke, C., Stoner, J.B., 2009. An investigation of spelling in the written compositions of students who read braille. *J. Vis. Impair. Blind.* 103 (10), 668–679.
- Cooper, H.L., Nichols, S.K., 2007. Technology and early braille literacy: using the Mountbatten Pro Brailier in primary-grade classrooms. *J. Vis. Impair. Blind.* 101 (1), 22–31.
- Craig, C.J., 1996. Family support of the emergent literacy of children with visual impairments. *J. Vis. Impair. Blind.* 90, 194–200.
- Randell, J.M., Wallace, D.H., 1974. Speed reading in braille: an empirical study. *N. Outlook Blind* 68, 13–19.
- Cranmer, T.V., Nemeth, A., 1991. A uniform braille code. *Braille Monitor* 377–383.
- D'Andrea, F.M., 2010. Practices and Preferences Among Students Who Read Braille and Use Assistive Technology. Ph.D. University of Pittsburgh.
- D'Andrea, F.M., 2013. Unified English Braille: the future of braille in the United States. *J. Vis. Impair. Blind.* 107 (3), 243–246.
- Drezek, W., 1999. Emergent braille literacy with move, touch, read. *J. Vis. Impair. Blind.* 93, 104–107.
- Durre, I.K., 1996. How much space does grade 2 braille really save? *J. Vis. Impair. Blind.* 90, 247–251.

- Emerson, R.W., Holbrook, M.C., D'Andrea, F.M., 2009. Acquisition of literacy skills by young children who are blind: results from the ABC Braille Study. *J. Vis. Impair. Blind.* 103 (10), 610–624.
- Erin, J.N., Wright, T.S., 2011. Learning to write in braille: an analysis of writing samples from participants in the Alphabetic Braille and Contracted (ABC) braille study. *J. Vis. Impair. Blind.* 105 (7), 389–401.
- Farrell, G., 1956. *The Story of Blindness*. Harvard University Press, Cambridge.
- Flanigan, P.J., Joslin, E.S., 1969. Patterns of response in the perception of braille configurations. *New Outlook* 63, 232–244.
- Foulke, E., 1964. Transfer of a complex perceptual skill. *Percept. Mot. Skills* 18, 733–740.
- Foulke, E., 1991. Braille. In: Heller, M.A., Schiff, W. (Eds.), *The Psychology of Touch*. Lawrence Erlbaum Associates., Hillsdale, NJ, pp. 219–238.
- Foulke, E., Smith, T., 1973. Reading print a letter at a time. In: Foulke, E. (Ed.), *The Development of an Expanded Reading Code for the Blind* (Pt. 2). Bureau of Education for the Handicapped, Washington, DC, pp. 82–87. Final Rep., Project 7-1185.
- Godshall, L.W., 1961. Analysis of Spelling Behavior of Braille Reading Blind Children. Unpublished master's thesis. George Peabody College for Teachers.
- Grenier, D., Giroux, N., 1997. A comparative study of spelling performance of sighted and blind students in senior high school. *J. Vis. Impair. Blind.* 91 (4), 393–400.
- Harley, R.K., 1969. Comparison of Several Approaches for Teaching Braille Reading to Blind Children. George Peabody College for Teachers, Nashville, TN. Final report.
- Henderson, F.M., 1967. The Effect of Character Recognition Training on Braille Reading. Unpublished specialist in education thesis. George Peabody College.
- Irwin, R.B., 1955. *As I Saw It*. American Foundation for the Blind, New York.
- Irwin, R.B., 1970. *The War of the Dots*. American Foundation for the Blind.
- Jackson, J., Bogart, D., Caton, H., 1993. The unified braille code: some myths and realities. *J. Vis. Impair. Blind.* 87 (10), 395–396.
- Johnson, L., 1989. Let's teach grade one braille. *J. Vis. Impair. Blind.* 83, 491–493.
- Kederis, C.J., Siems, J.R., Hayes, R.L., 1965. A frequency count of the symbology of English braille Grade 2, American usage. *Int. J. Educ. Blind* 15, 38–46.
- Kern, D., Bean, R.M., 2018. ILA 2017 standards: key notions, challenges, and opportunities for middle and high school classroom teachers. *J. Adolesc. Adult Lit.* 62 (1), 89–94.
- Kershman, S.M., 1976–1977. A Hierarchy of Tasks in the Development of Tactual Discrimination. Part Two. Education of the Visually Handicapped, pp. 107–115.
- Koenig, A.J., 1987. A Study of Expressive Writing Skills of Blind Students Including Partial Replication of the National Assessment of Educational Progress Third Writing Evaluation, Ph.D. Vanderbilt University.
- Koenig, A., Ashcroft, S.C., 1983. Initial assessment of student use of the electric Perkins Brailier. *Educ. Vis. Handicapped* 15 (2), 59–66.
- Koenig, A.J., Holbrook, M.C., 1995. Learning Media Assessment of Students With Visual Impairments: A Resource Guide. Texas School for the Blind and Visually Impaired, Austin.
- Koenig, A.J., Holbrook, M.C., 2000. *Foundations of Education*, second ed. AFB Press, New York.
- Kusajima, T., 1974. Visual Reading and Braille Reading: An Experimental Investigation of the Physiology and Psychology of Visual and Tactual Reading. American Foundation for the Blind, New York.
- Lorimer, J., Tobin, M.J., 1979. Experiments with modified Grade 2 braille codes to determine their effect on reading speed. *J. Vis. Impair. Blind.* 73, 324–328.
- Lowenfeld, B., Abel, G.L., Hatlen, P.H., 1969. *Blind Children Learn to Read*. Charles C. Thompson, Springfield, IL.
- Lusk, K.E., Corn, A.L., 2006. Learning and using print and braille: a study of dual-media learners, Part 1. *J. Vis. Impair. Blind.* 100 (10), 606–619.
- Millar, S., 1997. *Reading by Touch*. Routledge, New York.
- Miller, S., 2002. Practice makes perfect. *Future Reflections* 21 (2), 39–40.
- Mousty, P., Bertelson, P., 1985. A study of braille reading: 1. Reading speed as a function of hand usage and context. *Q. J. Exp. Psychol. Sec. A* 37 (2), 217–233.
- Munro, M.P., Munro, H.R., 2013. Infusion of print literacy methodology into braille instruction for students with visual impairments. *J. Blind. Innov. Res.* 3 (2).
- Nemeth, A., 1972. The Nemeth braille code for mathematics and science notation: 1972 revision. In: *Produced in Braille for the Library of Congress*. National Library Service for the Blind and Physically Handicapped by the American Printing House for the Blind.
- Nilsson, B., 1982. Braille: a medium for everybody? *J. Vis. Impair. Blind.* 76, 197–199.
- Nolan, C.Y., Kederis, C.J., 1969. Perceptual Factors in Braille Word Recognition. American Foundation for the Blind, New York.
- Olson, M.R., Harlow, S.D., Williams, J., 1975. Rapid reading in braille and large print: an examination of McBride's procedures. *N. Outlook Blind* 69, 392–395.
- O'Donnell, M.P., Wood, M., 1999. *Becoming a Reader*, second ed. Allyn & Bacon, Needham Heights, MA.
- Rex, E.J., Koenig, A.J., Wormsley, D.P., Baker, R.L., 1994. *Foundations of Braille Literacy*. American Foundation for the Blind, New York.
- Rodenberg, L.W., 1955. *The Story of Embossed Books for the Blind*. American Foundation for the Blind, New York.
- Snyder, T.D., Dillow, S.A., Hoffman, C.M., 2008. Digest of Education Statistics, 2007. NCES 2008-022. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.
- Snyder, T.D., De Brey, C., Dillow, S.A., 2019. Digest of Education Statistics 2017, NCES 2018-070. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.
- Spungin, S.J., 2003. Cannibalism is alive and well in the blindness field. *J. Vis. Impair. Blind.* 97 (2), 69–71.
- Stanfa, K., Johnson, N., 2017. Improving reading fluency in braille readers using repeated readings. *J. Blind. Innov. Res.* 7 (1), 1.
- Troughton, M., 1992. "One is Fun": Guidelines for Better Braille Literacy. Dialatype, Brantford, Ontario.
- Troxel, E.A., 1967. Experiments in tactile and visual reading. In: *IEEE Transactions on Human Factors in Electronics*, HFE-8, pp. 261–263.
- Umsted, R.G., 1970. Improvement of Braille Reading Through Code Recognition Training. Doctoral Dissertation, George Peabody College for Teachers. Ann Arbor, MI, University Microfilms, No. 71-4255.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services, 2000. Educating blind and visually impaired students: policy guidance. Fed. Regist. 65 (111), 36586–36594.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services, 2018. 40th Annual Report to Congress on the Implementation of the Individuals With Disabilities Education Act, 2018. Available at: <https://www2.ed.gov/about/offices/list/osers/frame-work/osers-framework-9-20-2018.pdf>.
- Wittenstein, S.H., 1994. Braille literacy: preservice training and teachers' attitudes. *J. Vis. Impair. Blind.* 88 (6), 516–524.
- Wittenstein, S.H., Pardee, M.L., 1996. Teachers' voices: comments on braille and literacy from the field. *J. Vis. Impair. Blind.* 90, 201–209.
- Wormsley, D.P., 1997. Learning to read, reading to learn. In: Wormsley, D.P., D'Andrea, F.M. (Eds.), *Instructional Strategies for Braille Literacy*, first ed. AFB Press, New York, pp. 57–110.
- Wright, T., Wormsley, D.P., Kamei-Hannan, C., 2009. Hand movements and braille reading efficiency: data from the alphabetic braille and contracted braille study. *J. Vis. Impair. Blind.* 103 (10), 649–661.

Autism, literacies, and social communication: a framework of inclusive new literacies

Bessie G. Stone and Beth Sagers, Queensland University of Technology, Faculty of Creative Industries, Education and Social Justice, Brisbane City, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

New Literacy Studies	434
What are New Literacy Studies?	434
New Literacy Studies and literacy education	435
Inclusive literacy education	435
Inclusive goals linked to literacy research and education	435
What is the autism spectrum?	436
The autism spectrum: embracing New Literacy Studies	436
Autism, literacies, and social communication	436
Student literacy profiles and autism	436
Characteristics of home literacy and autism	436
Literacies and social communication: closely related constructs	437
Literacy and social communication success and post-school outcomes	437
Autism and literacy challenges: medical and social perspectives	437
Medical model perspective: framing literacy challenges	438
Social model of disability: embracing students' strengths and interests	438
Social model of disability: illuminating and removing literacy barriers	438
Supporting inclusion of students on the autism spectrum through new literacies	438
Literacies: online, digital, and multimodal social communication	439
Autism research: social communication beyond physical contexts	439
Literacies and multimodal social communication	439
Autism and multimodal approaches to literacies and social communication	439
Inclusive literacies and autism: video games	440
Social communication through literacies of video games	440
Autism, social communication, and literacies of video games	440
Social communication and literacy motives and benefits of video gaming	441
Multimodal potential through gaming literacies	441
Multimodal potential of gaming literacies for autism	441
Conventional and functional literacy skills: virtual and informal educational contexts	442
A framework of inclusive new literacies	442
Inclusive new literacies: New Literacy Studies	442
Online multiplayer games: a model of inclusive new literacies	442
Inclusive new literacies: multimodal social communication	443
Inclusive new literacies: concepts of inclusive education	443
Research recommendations: literacies and social communication	443
References	444

New Literacy Studies

What are New Literacy Studies?

New Literacy Studies emphasize a broadened and newer way of defining the contemporary and evolving nature of literacies, that is influenced by digital communication and newer technologies (Street et al., 2017). New Literacy Studies is defined as a body of work in which there is an understanding of multiple literacies that are changeable, multimodal, and aligned with social practices across time, place, and space (Chakrabarty, 2020). A key focus of New Literacy Studies is making connections between literacies that are situated across formal and informal learning environments, and physical and virtual spaces (Alvermann and Robinson, 2017). The term literacies is embraced in this paper as an inclusive way to use different modes of communication to support literacy education (Fig. 1).

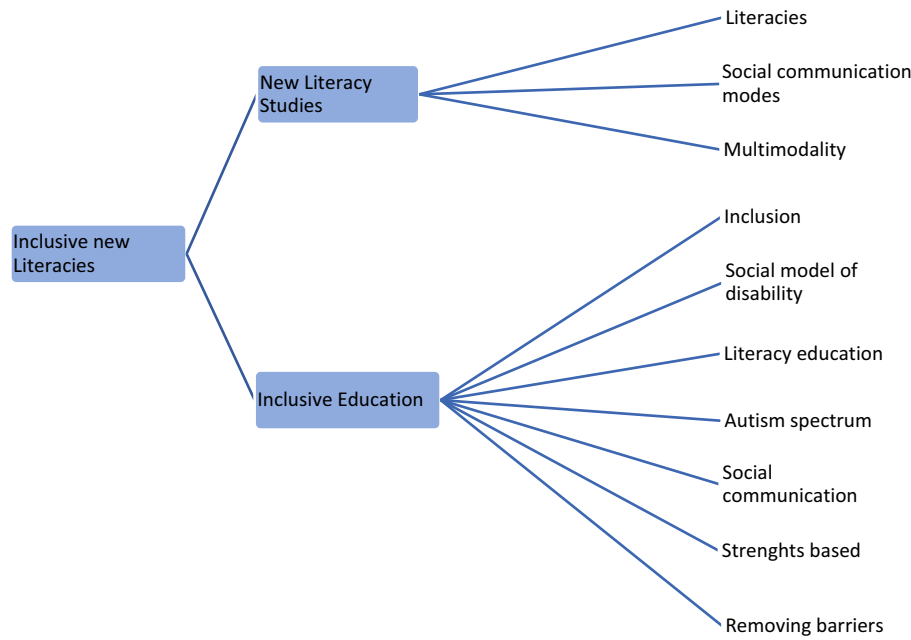


Fig. 1 Conceptual framework of inclusive new literacies.

New Literacy Studies and literacy education

Recently, New Literacy Studies has influenced a renewed interest in all aspects of literacy teaching and literacy education (Chakrabarty, 2020; Street, 2013). In addition, there had been an emergence of new literacies, which suggests that the characteristics and mechanics of literacies are situated and are different across diverse contexts (Mills, 2010). Students' access to contemporary literacy practices have been dictated by the developments of a variety of digital technologies within social contexts that have facilitated the development of literacy skills (Gee and Hayes, 2011). Literacies are also mediated by the availability of and interrelationship between spaces, devices, and other people to interact with (Oliver, 2016).

Newer understandings of literacies have evolved with the emergence of digital literacies (Jones, 2012), global literacies (Yoon, 2016), technology literacies (Thomas, 2011), virtual literacies (Merchant et al., 2014; Zhao et al., 2021), information literacies (Mackey and Jacobson, 2008), and visual literacies (Gitsaki, 2015). These and other types of literacies are considered inclusive if they facilitate opportunities for repeated practice; reciprocal, motivational, collaborative, and cooperative learning with peers; and technology-based engagements, such as with digital texts and virtual adventures (Boon et al., 2013). However, literacies for students with disabilities and diverse learners is less researched and understood in New Literacy Studies, especially for students on the autism spectrum.

Inclusive literacy education

Inclusive literacy learning and support for diversity are valuable notions in New Literacy Studies and literacy research (Dovchin and Pennycook, 2017; Luke et al., 2011). This type of learning involves having insights into offline and online social practices of all learners and remoulding school policies, culture, and practices to meet students' diverse needs (Booth and Ainscow, 2011; Jorgensen and Lowrie, 2011). Inclusive literacy education entails cognitive processes that are mediated by social practices, technologies, narratives, complex arrays of rules, and cultural knowledges. Within this context, students of all races, cultures, socioeconomic backgrounds, abilities, and linguistic diversities have the right to access scaffolded learning, participation, and resources that support their education (Booth and Ainscow, 2011; Liasidou, 2015; Ravet, 2011). Therefore, in contemporary literacy research, educators are encouraged to build on notions of inclusive education and focus on differentiated support for students and their diverse needs (Ashman, 2014).

Inclusive goals linked to literacy research and education

The United Nations Convention on the Rights of Persons with Disabilities (UNCRPD), declares that all students have the right to inclusive education (Committee on the Rights of Persons with Disabilities, CRPD, 2016; United Nations, 2018). Drawing on the guidelines of UNCRPD, Article 24, Comment Number Four (CRPD, 2016), literacy researchers and literacy educators need to understand and address:

- strategies by which diverse students can be supported through the literacies of information and communication technologies to meet their needs and targeted goals,
- ways to minimize challenges to literacy development and social practices, of all students, and
- ways to make reasonable accommodations for the literacy practices of social groups such as students with autism.

Core to inclusive literacy research is to draw attention to issues of diversity, inequality, and interventions, and to understand the real and daily challenges to social inclusion and barriers in literacies that students experience (Dovchin and Pennycook, 2017). Opportunities to fully participate within communities through semiotic resources, new knowledge, and skills in literacies evoke the notion of inclusion (Warschauer and Tate, 2017). However, little is known about how newer literacies of students with autism may be situated within the context of New Literacy Studies.

What is the autism spectrum?

The autism spectrum is a neuro-developmental condition characterized by two main diagnostic criteria: (i) persistent challenges of social communication and (ii) restricted and repetitive patterns in interests, activities, or behaviors (Bishop et al., 2013). The prevalence rates in students diagnosed on the spectrum have increased over the past decades (Australian Bureau of Statistics, 2020; Centers for Disease Control and Prevention, 2020). Given this increase, a timely synthesis of knowledge and understanding about the relationship between literacies and social communication for this population of students is needed.

The autism spectrum: embracing New Literacy Studies

There is a role for understanding literacies for students on the autism spectrum that is informed by New Literacy Studies. Gee (2007) argues that literacy is not unitary but is multiple because distinct types of literacies are needed to read, interpret, and write various types of texts. From this perspective, we acknowledge students on the spectrum may have different ways of approaching communicating, reading, writing, and acting with texts. Students have multiple means of feeling, interacting with, and developing knowledge as they engage in social practices with various print and literacies in online and offline contexts (Street et al., 2017). It is therefore important to consider New Literacy Studies as a way to support the inclusion and diverse literacy practices for students on the autism spectrum. With the growing number of students on the spectrum (Christensen et al., 2016), it is important to develop inclusive frameworks for engaging students in literacy education and supporting their social communication.

Autism, literacies, and social communication

Educators need to make connections between literacies in social practices that are situated across diverse physical and virtual spaces, and informal and formal learning environments (Alvermann and Robinson, 2017; Feiler et al., 2007). New Literacy Studies enable understandings of how bridges can be built among conventional literacies, home literacies, school literacies, and online and offline social contexts. To date, a large body of autism researchers have focused on conventional literacies and literacy skills for students with autism within formal education contexts. For example, the relationship between levels of autism, students' emergent-literacy skills, and decoding abilities have been subjects of investigation (Clendon et al., 2021; Dynia et al., 2017; Hume et al., 2020; Jacobs and Richdale, 2013). Researchers of emergent literacies and autism have considered students' receptive language demands, and factors in physical environments when trying to assess students' verbal communication skills (Clendon et al., 2021; Kwok et al., 2015). Autism research within New Literacy Studies may enhance understanding of how digital technologies such as computers can be used to target and teach literacy skills such as reading and writing of letters, syllables, words, sentence construction, sentence imitation, verbal expression, phonological awareness, spelling, and receptive identification and matching of pictures and vocabulary words (Ramdosset al., 2011).

Student literacy profiles and autism

Students' literacy profiles, comprehension abilities, academic knowledge, reading abilities, and receptive, expressive, and written language, and communication skills are also considered important areas of literacy focus in autism studies (Bailey et al., 2017; McIntyre et al., 2021). Literacy profiles of students with autism have been conceptualized along a continuum of diverse adjustments where learners may receive comprehensive support, intensive support, moderate support, or limited support for conventional literacies and school literacies (Hume et al., 2020).

Characteristics of home literacy and autism

Researchers of conventional literacies and literacy skills have also investigated characteristics of home literacy and home literacy environments in students on the spectrum (Dynia et al., 2014; Lucas and Norbury, 2018). Home literacies are essential to access educational curriculum, and are associated with educational attainment, independent reading, variations in reading ability, oral

language ability, strength and weaknesses in letter name identification and print functions, parental encouragement and scaffolding, and parent-child engagement (Lanter et al., 2012; Lucas and Norbury, 2018). Home literacy strengths and challenges of students on the spectrum have been addressed through shared book-reading practices and interventions (Fleury et al., 2021).

Educators are urged to discern how home literacies can be aligned with classroom literacies and used as pedagogical tools to achieve learning outcomes (Lucas and Norbury, 2018; Yeh and Mitric, 2020). It is inferred that, if schools have a tendency to value and promote only their own perspectives of literacies, then the exclusion of out-of-school and home literacy practices across families of students on the spectrum may restrict student literacy and social communication development (Feiler et al., 2017; Vigdor et al., 2014). New Literacy Studies provides a context for broadening understanding of students' social practices, literacy profiles, and literacy skills that are situated across diverse formal and informal learning environments.

Literacies and social communication: closely related constructs

Social communication, a key area of challenge for students on the spectrum, includes direct human engagement and contact with others during physical or virtual play and verbal exchange (Chen and Tsai, 2016; Hughes et al., 2013; Locke et al., 2016). A growing body of autism research shows that social communication and literacy development, interactions in literacy events, and literacy experiences are intertwined and closely related constructs (Ricketts et al., 2013). Social communication skills facilitate critical thinking, discussions, collaboration, and presentations through the use of grammar, figurative language, and vocabulary (McIntyre et al., 2021). Autism scholars have focused on how students communicate with each other within specific contexts and across time and space and through various literacies (Vasudevan et al., 2018; Warschauer and Tate, 2017).

More specifically, the link between social communication and literacies have been shown through studies that focus on shared reading with adults who allow students on the spectrum to develop social communication through engagement in acts such as vocabulary building abilities, listening, uttering, inference, commenting, expressing, comprehending, asking questions, and promoting interaction (Bean et al., 2019; Boyle et al., 2019; Fleury and Hugh, 2018; Fleury et al., 2021). The relationship between literacy skills such as decoding skill and comprehension and language competence has also been discussed in the research (Lucas and Norbury, 2018). Given the strong link between literacies and social communication, inclusive literacies frameworks could be embraced to develop better understanding of improving outcomes for students with autism who display literacy strengths but who may need individualized supports for social communication (Hume et al., 2020).

Literacy and social communication success and post-school outcomes

Students' literacy success impacts on communication, participation, and well-being beyond formal education years (Westerveld and Paynter, 2021). The long-term social communication, academic, behavioral, emotional, employment, and economic outcomes of all students, including those on the spectrum may be better if they achieve literacy benchmarks (Whalon et al., 2009). The acquisition of social communication and literacy skills are fundamental to developing relationships with others, sharing and receiving information, expressing interests, needs and desires, engagement and participation in important aspects of life including formal and informal education contexts, community, family, relationships, and employment opportunities (Light and McNaughton, 2012). Taken together, the research shows literacy and social communication skills are integral to success post-school years across various content areas and contexts.

By the time students reach high school, they are expected to have achieved basic reading, writing, and comprehension skills and social knowledge (Boyle et al., 2019). A focus on students' achievements, strengths, interest, and motivation may be important for students who are at increased risk of experiencing challenges with the development of literacy skills that are understood to underpin success in social communication (Vasudevan et al., 2018; Warschauer and Tate, 2017). While success in literacy education is critical for all students across early childhood to adolescence, many students with autism may experience more social communication and literacy challenges than their peers (Clendon et al., 2021; Westerveld and Paynter, 2021). Students with autism need to develop social communication skills to limit bullying, peer rejection, loneliness, and to develop friendship qualities, and classroom social networks, acceptance, centrality, companionship, and reciprocity during and after formal education (Carrington, 2017).

Given that students' social communication abilities and successful outcomes may be influenced by literacy activities, students who have limited social communication skills may benefit from participating in literacy activities with instructional support (Kluth and Chandler-Olcott, 2008). The closely related constructs between literacies and social communication as well as challenges that students on the spectrum may face, suggest it may be preferable to embrace inclusive frameworks that have implications to support them to successfully develop their strengths and a repertoire of social communication skills as they continue to engage with high-interest literacy practices across diverse formal and informal educational contexts.

Autism and literacy challenges: medical and social perspectives

The literacy challenges of students with autism has been researched and understood from a clinical and medically based perspective and through the social model lens of disability (Waltz, 2013). Attention is drawn to the medical versus social model perspectives of autism and literacy challenges to highlight the need to build an inclusive literacy framework.

Medical model perspective: framing literacy challenges

Within autism research, literacy challenges for students on the spectrum have been linked with delays in early literacy skills such as decoding, phonological processing, and reading comprehension, and challenges with verbal language skills and expressive and receptive language (Boyle et al., 2019; Davidson and Ellis Weismer, 2014; Kwok et al., 2015; Tager-Flusberg and Kasari, 2013; Westerveld and Paynter, 2021). From a medical perspective, challenges in literacies may be associated with factors such as impact of autism traits, students' lack of social understanding, limited social communication skills, and restricted interests and repetitive behaviors (American Psychiatric Association, 2013, see also Ricketts et al., 2013). Challenges with coexisting intellectual disabilities (Westerveld et al., 2020), theory of mind (see for example, Baron-Cohen, 2000; Baron-Cohen et al., 2013; Westerveld and Roberts, 2017), and executive functioning (Lai et al., 2017) have also been linked with literacy challenges for this cohort of students.

Jordan et al. (2010) argue that a medical perspective can guide understandings that literacy skills of learners on the spectrum are affected by their learning styles, intellectual abilities, personalities, and patterns of weaknesses. Humphrey and Lewis (2008) support the view that students are sometimes segregated, embarrassed, or ignored, and subjected to unnecessary exclusion where literacy education is heavily influenced by an autonomous model of literacy and the medical model of disability. A deficit perspective may result in students being excluded from literacy opportunities and denied resources to acquire and develop these skills (Keefe and Copeland, 2011). This paper embraces New Literacy Studies to build an inclusive framework for understanding literacy skills, strengths, and challenges experienced by students on the spectrum.

Social model of disability: embracing students' strengths and interests

Despite medical perspectives of literacies, much of the inclusive education literature embrace the fundamental principle of inclusion that supports diversity and promotes the notion that all learners have valuable practices, strengths, individual differences, and social and cultural backgrounds for literacy classrooms (Carrington and MacArthur, 2012; Comber, 2007; McIntyre et al., 2011; Tomlinson and Imbeau, 2010). Using the perspective of a social model of disability, literacy models that emphasize the skill deficits students may experience are being rejected (Waltz, 2013). Instead, the social model of disability enables a focus on building on students' strengths and abilities rather than the perspective that literacy challenges result from the characteristics of autism spectrum and are innate to the students (Durell, 2014; Patel et al., 2014). A social model of disability requires inclusive frameworks in education with multiple approaches to literacies informed by collaboration with a range of stakeholders, optimization of assessment processes, and a thorough knowledge of students' literacy profiles including their interests, strengths, and needs, and prioritizing adjustments and supports (Clendon et al., 2021; Winter-Messiers, 2007).

Social model of disability: illuminating and removing literacy barriers

Through a social model of disability lens, barriers for literacy education are socially, institutionally, and environmentally constructed (Oliver, 2013; Woods, 2017). Through this lens, Disability may be perceived in literacy education to result from factors such as the persistent attitudinal barriers of others in the educational environments, societies, and communities that restrict students from fully, equitably, and effectively learning, engaging, and participating with others. Other factors that may contribute toward a literacy disability are inadequate resources, funding, and opportunities to promote interactions and learning.

Students can develop literacy skills when educators remove barriers, understand students' strengths, provide learning opportunities, and develop in-depth assessments to inform literacy instruction (Clendon et al., 2021). Supporting literacy inclusion for students on the autism spectrum (who may be at risk of experiencing challenges with the development of literacy skills and may perform slower and lower than their typically developing peers in literacy areas) requires identify and removing barriers to literacy learning (Wei et al., 2011, 2014). For example, social, economic, and physical barriers to literacy education need to be identified and removed in virtual spaces to enhance opportunities for social communication (Gee, 2015). New inclusive frameworks within New Literacy Studies may broaden understandings of barriers that students on the spectrum face in online and virtual contexts and how students can be supported through inclusive practices, that target their interests, strengths, and needs.

Supporting inclusion of students on the autism spectrum through new literacies

Newer definitions of literacies for students on the autism spectrum have shifted from instructional practices and discrete skill instruction to more inclusive approaches that acknowledge potentials incorporated in online practices (Stone, 2018). Literacies for students on the autism spectrum now means acknowledging the challenges students may experience and support in an integrated way their abilities, strengths, skills, and competencies to motivate students to view, read, write, speak, listen, design, and construct meaning from electronic texts and information and communication technologies in independent, collaborative, and appropriate ways (Baynard, 2010; Stone et al., 2018, 2019). Consideration of how the use of new literacies can be maximized to support inclusion of students on the autism spectrum, and other diverse learners to meet their targeted goals and achieve their optimum levels of educational outcomes is needed (Barr et al., 2008; Boche and Henning, 2015).

Literacies: online, digital, and multimodal social communication

The New Literacy Studies has broadened perspectives that literacy practices in online virtual settings are linked with new ways of social communication (Gee, 2015; Garcia, 2017; Ito et al., 2009; Peters et al., 2013; Quandt and Kröger, 2014). Many students are active participants in social communication processes through digital literacies as opposed to being passive recipients of knowledge, who are taught that they need to develop and use social communication skills in physical contexts, such as classrooms and playgrounds (Beavis, 2014; Locke et al., 2016). Many students' social communication has expanded from physical contexts to virtual contexts, and from a reliance on language to using digital tools, game engines, and virtual worlds that are created by game designers (Gee, 2015b). Improvements in literacies are also associated with opportunities that students have to communicate with each other through the use of new technologies, such as online computer activities (Genlott and Grönlund, 2013; McCreery et al., 2015). Therefore, there is now increased recognition that social communication is an essential process that should be integrated into literacy education through digital media platforms such as: video games, Email, music video, smart phone, and social media (Light and McNaughton, 2012).

Autism research: social communication beyond physical contexts

Until recently, much of autism research into students' social communication and literacy skills has focused on students' verbal communication skills in physical contexts. In the past, challenges with social communication skills in students with autism have been manifested, observed, and assessed in physical contexts such as through physical and verbal face-to-face play in physical environments and between peers (Peters et al., 2013). Similarly, challenges with social communication and literacies have been associated with students' speech, language, and sensory perception (Clendon et al., 2021; Westerveld and Paynter, 2021). However, a greater understanding of students' proficiencies is needed beyond their spoken language because when students on the spectrum display social communication needs and limitations within narrow definitions of literacy, they may become marginalized. Newer understandings of literacy practices and social communication are necessary to move beyond physical contexts, and medical and conventional perspectives.

Notions of social communication are expanding in unprecedented ways through digital and virtual spaces and screen-based literacy texts to include forms of literacy practices (Garcia, 2017; Ito et al., 2009; Peters et al., 2013). These notions need to be embraced on a wider scale for students with autism. The continued development of inclusive education, technologies, and digital communication give rise to an increased number and types of contemporary digital practices, literacies, and online and virtual environments that serve as platforms for students on the spectrum to acquire opportunities and skills to socially communicate. For example, Zhao et al. (2021) found that cognitive training based on virtual reality was very appealing to students on the spectrum and that within this context students' social communication and speech improved.

All students need to be given opportunities to practice, generalize, and transfer essential skills beyond virtual environments, across different contexts with different people, and in different activities, until eventually they develop proficiency (Cope et al., 2000; Gee, 2015; Rivalland, 2004). Opportunities and platforms to develop social communication skills in online and virtual environments may enable students on the spectrum to benefit from digital technologies in their development of social communication and literacy skills without the unpredictability and challenges of physical-world and face-to-face interactions that could result in increased anxiety. The New Literacy Studies is embraced to broaden understanding of students' literacy proficiencies and needs, adjustments, and differentiated support considerate of and responsive to students' diverse social communication needs beyond physical contexts (Clendon et al., 2021; Fleury et al., 2021; Foreman and Arthur-Kelly, 2017; Hume et al., 2020).

Literacies and multimodal social communication

A 21st century description of literacies acknowledges and embraces students' ability to use modes of communication in a variety of contextual forms for social communication (Alvermann, 2009; Tompkins, 2014). Improvements in literacies are associated with multiple contexts and opportunities where students have to socially communicate with each other through new modes of instruction (Genlott and Grönlund, 2013; McCreery et al., 2015; Street et al., 2017).

Researchers have shown links between students' online literacy practices and their development in social communication (Merchant et al., 2014). Multimodal approaches to literacies and social communication allow students to gain meaning and communicate in the widest sense through—visual, oral, gestural, linguistic, musical, kinesthetic, and digital ways (Alvermann, 2009). Nonnis and Bryan-Kinns (2021) agree and have proposed methodological approaches that embrace diversity and promote designs to support students' use of music in favorable environments to spontaneously socially communicate and engage with peers. Given that students have a diverse range of needs, they should be supported through a multiplicity of social communication modes, such as speech, writing, gestures, images, and sounds. Multimodality is defined as the combination of resources for multiple semiotic resources of meaning making and communication (Kress, 2017).

Autism and multimodal approaches to literacies and social communication

As some students on the spectrum engage in various online digital media and digital texts, an inclusive framework within New Literacy Studies should be encouraged to broaden understanding of their interests in a range of social communication modes which

facilitate literacy and social communication development. Recently, concepts of literacies and social communication for students on the spectrum have become attached to the theory of multimodality (Stone et al., 2018, 2019). New literacies for students on the spectrum encompass digital acts of social communication that combine speech, gaze, sound, gesture, facial expression, intonation, and body positioning. Findings from Nonnis and Bryan-Kinns (2021) revealed that interactive technology can be used by students on the spectrum as opportunities for social communication through eye-contact, turn-taking, sharing of objects and space. Social communication for students with autism has also been researched through arm and hand movement (Braddock and Hilton, 2016), image such as illustrations and drawings (Peters et al., 2013; Rogers, 2013), gesture (Dindar et al., 2016), touch (Gentry et al., 2015), body posture (Doody and Bull, 2013), body movement (Lucas and Norbury, 2018), body language (Peterson et al., 2015), speech (Katz and Girolametto, 2015), writing (Asaro-Saddler, 2014; Caron, 2016; Geither and Meeks, 2014), and sound (Boddaert et al., 2004; Russo et al., 2009). These findings confirm consensus among researchers that students now use a wider range of social communication modes and express themselves through the multimodal symbols available to them within their online, virtual, and digital literacy engagements more than they used to.

Autism researchers have continued to seek multimodal ways to understand social communication in the literacy practices of students with autism. Multimodal elements in virtual worlds have been used as realistic platforms and tools for enhancing social communication for students on the spectrum. For example, a study by Finke et al. (2015) highlights that during video game play, students with autism use diverse social communication modes such as speech and/or vocalizations, gestures, tapping tablet or smartphone, sign language, communication board or book, picture exchange communication systems, and computer with speech output.

Although many students on the spectrum may have a desire to socially communicate through the use of digital technologies, they may experience challenges interpreting meaning from verbal and nonverbal social communication modes such as interpreting facial expression, gesture, and body language (Atkinson, 2009; Bekele et al., 2013; Doody and Bull, 2013; Kaartinen et al., 2012). Considering the importance of multimodal online and virtual literacies for social communication proficiencies and success, the New Literacy Studies is important to inform research into online multimodal literacy practices of students with autism.

Inclusive literacies and autism: video games

There are many digital literacies that students on the spectrum engage with but for the purpose of this paper, video games, such as those played in online multiplayer gaming contexts will be the focus. Video gaming is defined as the use of specific game systems, through digital devices such as computers, tablets, and cellular phones, where players experience entertainment and enjoyment as they engage alone or in the company of other players (Finke et al., 2015). These digital texts are sophisticated literacy practices that are part of the popular culture of students (Beavis, 2014; Gallup et al., 2016; Stone, 2018).

Social communication through literacies of video games

New Literacy Studies have influenced interest in the way students' literacy and social communication are shaped by video gaming. Newer understandings of literacies have enabled broadened ways of defining the contemporary and evolving nature of social communication through students' interests in video games. The gap in autism research and the New Literacy Studies literature draws attention to the need for a conceptual framework to better understand and describe the video gaming literacy practices of students on the spectrum and to address students' social communication strengths and needs within an inclusive context.

Autism, social communication, and literacies of video games

A study into the prevalence of screen-based media use among youths with autism suggests that the majority of youths on the spectrum (64.2%) spent most of their free time playing video games (Mazurek et al., 2012). Students with autism have been reported to use their interest in video gaming to have conversations with their peers about video gaming, and consequently show improvement in communication skills (Winter-Messiers, 2007). In the study reported by Gallup et al. (2016), the researchers explored aspects of social communication in socialization of adolescents with autism who engaged in multiplayer online role-playing games. The participants reported that they preferred the ease with which they could talk with other gamers through virtual medium that facilitated communication through voice talk. The findings from this study suggest video games that are played online by students on the spectrum provide platforms and opportunities for students to engage in conversation and build social connections with peers, friends, and unknown players in virtual communities.

Students on the spectrum may use sophisticated digital literacy and social communication skills to enthusiastically seek out others to actively work together and proficiently solve complex problems within the game context (Gallup et al., 2016). For example, Stone et al. (2018, 2019) investigated how social interactions within online multiplayer gaming contexts provide opportunities for students with autism to socially communicate and collaborate with online gamers, and to talk about digital creations in classroom and virtual world spaces. Similarly, findings from the study of Finke et al. (2015) revealed that students on the spectrum talked about playing video games with familiar partners, unknown people, and valuable friends and engaged in social communication through positive social experiences. The authors encourage professionals to consider how videogame play may be used to build and teach new social communication skills to students on the spectrum. Findings from these and other studies suggest an

increasing number of students with autism are likely to prefer online multiplayer gaming as opposed to engaging in physical play with peers or offline single player gaming (Gallup et al., 2016; Kuo et al., 2014; Lucas and Norbury, 2018; Mairena et al., 2019; Mazurek and Wenstrup, 2013; Stone et al., 2019). Further connections to New Literacies Studies are relevant in showing that the interests of students with autism to socially communicate permeate their literacy practices beyond the boundaries of physical spaces, to include digital contexts such as video games (Marsh et al., 2016; Richards and Burn, 2014).

Social communication and literacy motives and benefits of video gaming

Literacies of digital games such as Minecraft are useful to motivate student engagements in social communication through literacy learning opportunities (Marcon, 2013; Marcon and Faulkner, 2016). To illustrate further, research by Mazurek et al. (2015) suggests video gaming may provide opportunities for students with autism to enjoy reading game stories and commentaries on certain issues and story elements of fantasy and adventure. A study conducted by Finke et al. (2018), shows students on the spectrum can be attentive to social communication during video gaming, and develop literacy skills such as writing, speaking, reading, and spelling. Researchers have also observed during a brief group intervention for students with autism, teaching sportsmanship skills through video games and video-gaming technologies facilitate learning of social communication skills (Ferguson et al., 2013).

Social communication benefits through video game design have also been the focus of research. Video games have been designed to help students on the spectrum construct, recognize, and follow emotions through 3D avatar facial expressions (Abirached et al., 2011; Finke et al., 2017). Terlouw et al. (2021) developed a multiplayer virtual escape room to facilitate direct and playful social communication between students on the spectrum and their peers. Ben-Sasson et al. (2013) found that games with a technological touch increased positive social communication for students on the spectrum. There is a gap in the New Literacies Studies for building on these findings and enhancing understandings of social communication for students with autism, through video game design.

Unique literacy and social communication benefits and motives of engaging with video games play include personalized learning, social initiation, creativity, story design, and social connection (Finke et al., 2018; Gallup et al., 2016; McEvoy, 2016). Likewise, social communication in virtual environments of online video gaming is a necessary part of a healthy lifestyle to help students on the spectrum overcome anxiety and stress and limits the chance for social isolation, and feeling depressed and lonely (Gallup et al., 2016). Social communication in game play for students on the spectrum may be influenced by factors such students' age, expressive communication abilities, and social awareness (Ben-Sasson et al., 2013). Given these benefits and motives, the New Literacies Studies is considered useful to better understand potentials of video games for the literacy and social communication of students on the spectrum.

Multimodal potential through gaming literacies

Video gaming literacies embody several modalities, with offerings that enable students to navigate virtual spaces, perform actions as gamers, and create and gain specific meanings from communities they socially interact within (Gee, 2015; Quandt and Kröger, 2014; Vance, 2017). Students as gamers combine language, social practices, and multimodal semiotic resources through avatars' actions and the use of conversational and visual literacies (Gee, 2007, 2015). Multimodal potentials of gaming literacies to facilitate social communication between student gamers can also be understood through the games' digital tools such as chat boxes, voice chats, and video chats. These tools hold the potential to support social communication skills through immediate and constant written, spoken, visual, gestural, and aural feedback among players (Cheng and Huang, 2012; Fullen, 2012; Gee, 2015). When video gaming, students' social communication is also influenced by music, in-game sounds, and sound effects heard; and virtual images, gestures, and actions viewed (Lemke, 2017). Students as gamers use these modes of communication to engage in sophisticated dialogs and collaborate to create, socialize, problem-solve, strategize, and engage in ways that may be challenging in face-to-face contexts. As a result, game-based literacy learning has made its way into schools through forms of social communication such as discussions and play (Beavis et al., 2015).

The use of digital and video gaming literacies, such as in online multiplayer gaming activities, provide opportunities for students to socially communicate with each other through a variety of modes (Genlott and Grönlund, 2013). Descriptions of inclusive literacy practices therefore must integrate multimodal aspects of social communication instead of a sole emphasis on behavioral observations and use of verbal communication in physical environments (Quandt and Kröger, 2014). Video gaming technologies and features can support things that are important for social communication with friends and other members of gaming communities with whom students on the spectrum engage (Boyd et al., 2015). This is important to consider within New Literacy Studies as students with autism continue to spend more time engaged with electronic screen media than any other leisure activity (Howard and Patti Ducoff, 2008).

Multimodal potential of gaming literacies for autism

An expansion of video gaming technologies, gaming information, and increasingly globalized affinity spaces means that students with autism are enhancing their abilities to socially communicate with others through new multimodal systems and digital literacies of video games. As gamers, students on the spectrum have demonstrated conscious and simultaneous interpretation of a variety of social communication modes (Stone et al., 2018, 2019). In this context, students cumulatively immerse themselves in conversations

and interactions over extended timescales, through in-game and vocally produced sounds, spoken words, images, written texts, and body movements. To illustrate, key findings from [Hopkins et al. \(2011\)](#) suggest social communication skills may be improved in students on the spectrum through computer-based interactive avatar intervention if students are provided with opportunities to practice attending to eye gaze, recognizing faces and emotions, and discriminating facial expressions.

The review of literature shows that physical contexts do not represent all the literacy and social communication benefits for students on the spectrum through video games such as those played in online multiplayer contexts ([Gallup et al., 2016](#)). Inclusive approaches for students on the spectrum to develop social communication through the literacies of online multiplayer games may include: (i) opportunities for them to design games' elements and mechanics; and (ii) opportunities to display social interaction behaviors, including descriptive behaviors, social smiling, visual contact, and vocalization directed toward game characters and people ([Malinverniet al., 2014](#); [Stone, 2018](#)).

Despite this knowledge, only a few studies have explored how video game play, such as online multiplayer, help foster literacy and social communication skills for students on the spectrum through interactions with peers, friends, and family members ([Mazurek et al., 2015](#); [Terlouw et al., 2021](#)). Although multimodal social communication is a very important area of literacy learning for students on the spectrum, few studies provide a glimpse into the multimodal potential of online video gaming environments to support students' development of social communication skills in inclusive ways ([Stone et al., 2018, 2019](#)). The New Literacy Studies is embraced to draw attention to how video games' multimodal resources support or hinder social communication skills that may ultimately support greater inclusive participation in literacy and successful outcomes beyond classroom walls.

Conventional and functional literacy skills: virtual and informal educational contexts

A vast body of research in New Literacy Studies and autism has examined conventional and functional literacy skills including decoding, encoding, reading, and writing skills in offline-physical environments ([Boyle et al., 2019](#); [Davidson and Ellis Weismer, 2014](#); [Kwok et al., 2015](#); [Tager-Flusberg and Kasari, 2013](#); [Westerveld and Paynter, 2021](#)). Far fewer studies have examined these skills in online virtual environments in formal and informal educational contexts. These skills are important for reciprocal social communication in online digital text-based mediums such as social media that rely on screen text-based communication ([Abirached et al., 2011](#)). Despite research on conventional and emerging literacy skills of students on the spectrum, there remains a relative lack of understanding regarding online and virtual inclusive literacy practices for this population of students. The review of literature has shown that to explore only clinical and conventional assumptions about the literacy practices and social communication of students on the spectrum in physical and offline settings provides a partial and inhibited perspective of autism, literacies, and social communication.

A framework of inclusive new literacies

[Stone \(2018\)](#) presents a conceptual framework for inclusive new literacies that highlights potentials for using digital and gaming literacies such as those of online multiplayer games as inclusive resources within formal educational contexts. The conceptual framework is expanded from a review and integration of research, discourses, concepts, and theoretical assumptions from the fields of New Literacy Studies and inclusive education.

Inclusive new literacies: New Literacy Studies

The field of New Literacy Studies has reflected aspects of access to literacy education and has provided implications to support the social communication of students. Therefore, the framework draws on the digital strand of the New Literacy Studies and builds on the works by theorists within this field of education to highlight the notion students' online literacy practices are often linked to their social communication development which are manifested in virtual and physical environments and online activities and digital contexts ([Gee, 2015](#); [Hayes and Duncan, 2012](#); [Merchant et al., 2014](#); [Prensky, 2007](#); [Richards and Burn, 2014](#); [Street et al., 2017](#)). As [Stone \(2018\)](#) explains, a focus on inclusive new literacies for students on the spectrum is necessary because the New Literacy Studies literature has not explicitly addressed how to remove barriers to literacy learning and participation and how reasonable accommodations could be made for all learners, including students with autism, according to the inclusive educational principles of the *UNCRPD Article 24, Comment Number Four* ([CRPD, 2016](#)).

Online multiplayer games: a model of inclusive new literacies

Online multiplayer games are selected as a model of inclusive new literacies because they are embodied in the realm of many sophisticated literacies that can meet the diverse needs of learners when required ([Beavis et al., 2012](#); [Finke et al., 2015](#); [Gallup et al., 2016](#); [Garcia, 2017](#); [Oakley, 2017](#); [Stone, 2018](#); [Whitton, 2013](#)). Furthermore, describing online multiplayer games as inclusive new literacies, within the context of New Literacy Studies, can broaden understandings of how connections can be made among online and offline social practices, home literacies, and school literacies ([Feileret al., 2007](#); [Stone et al., 2018, 2019](#)).

This notion has not been adequately embraced in the literature despite research indicating that the popularity of online multiplayer games among students, including those on the spectrum has soared over the past two decades ([Beavis, 2014](#); [Nagygyörgy et al., 2013](#)). For example, to date few studies have used an inclusive multimodal framework that examines how the use of online

multiplayer games can develop meaningful literacy and social communication skills for students with autism (Stone et al., 2018, 2019). The increasing interest of students with autism in online gaming and virtual environments (Merchant et al., 2014) raises important consideration for video games such as those played in online multiplayer format to be understood as inclusive new literacies.

Inclusive new literacies: multimodal social communication

The notion of inclusive new literacies is consistent with the consideration that literacy resources can be designed, manipulated, created, and reconstructed across multimodal platforms and multimedia such as those of online multiplayer games to make new meaning and new experiences for social communication (Kress, 2013; Luke et al., 2011). The framework integrates aspects of new literacies, which emphasizes the importance of multimodal social communication in virtual and physical contexts rather than focusing only on the social, academic, symbolic, and face-to-face uses of spoken and written language (Kress, 2017). Within the inclusive new literacies framework, we highlight that as students on the spectrum engage with digital technologies, social communication becomes more complex intensifying the multimodal benefits and challenges they experience. Descriptions of social communication extend deeper than words and physical contact among peers for enjoyment, to include multimodal aspects of online literacy practices, such as through the semiotic resources offered through online multiplayer games (Gee, 2015).

A review of autism literature, literacies studies, and multimodal research (Alvermann, 2009; Beavis et al., 2012; Keefe and Copeland, 2011; Oakley, 2017) provides better understanding of how engagements of students on the spectrum with inclusive new literacies enabled them to gain authentic meanings for social communication through oral, written, visual, gestural, audio, and digital ways. The framework embraces integration of a broad and flexible group of skills, abilities, strategies, and competencies that motivates students to communicate with others for example, through speech, written texts, images, gestures, and sounds within the games' context as students navigate virtual and physical spaces (Finke et al., 2015; Stone et al., 2018, 2019).

Inclusive new literacies: concepts of inclusive education

The framework is also embedded within the field of inclusive education, which contributes recent understandings of autism spectrum, inclusion, literacy education, and social communication for this population of students. The framework draws on social model perspectives of disability, to give attention to students with and without medical diagnoses opportunities, platforms, and support to engage in online and virtual literacies through all available resources and in multimodal ways. As Mitchell et al. (2007) imply, scaffolded support in virtual environments might be achieved through guidance, prompts, and options for appropriate multimodal engagements.

Inclusive education within the framework means barriers should be illuminated and removed, and changes should be made to provide appropriate and necessary resources and adjustments for students to competently participate in social communication and quality literacy experiences in virtual and physical contexts after adult, peer, technological, or virtual guidance is withdrawn. Stakeholders are obligated to provide resources, and differentiated exemplary literacy instructions, and strategies and interventions that suit the diversity in students' strengths, interests, skills, abilities, learning styles, and individual characteristics (Saggers et al., 2011).

Twenty first century inclusion means that all students need resources and opportunities to develop their literacy learning, social communication, and digital capabilities regardless of their medical diagnosis, and cultural, economic, and geographical backgrounds (Oakley, 2017; Price-Dennis et al., 2015). Regardless of the personal, physical, and social constraints to literacies and social communication, supporters for inclusive education have emphasized the importance of providing all students, including those on the spectrum, with environments where their diverse situated literacy experiences and social communication are respected (Ainscow and Sandill, 2010; Carroll, 2015; Malinverni et al., 2014).

The framework supports the notion educators and researchers should divert from the conventional notions that literacy is a mental set of abilities and functional skills that reside in the minds of students, such as those on the spectrum. It argues literacies should be embraced as inclusive currencies to help students socially communicate within various contexts, such as home and school to ensure access to online, virtual, and home literacy resources; making-meaning and continuity in learning within and across inclusive environments (Comber and Barnett, 2003; Lucas and Norbury, 2018; Martini and Sénéchal, 2012). This framework is timely, given that it expands inclusive understandings of the online and offline literacy practices of 21st century students on the spectrum (Reynolds et al., 2011; Stone, 2018).

This paper broadens the original conceptualization of inclusive new literacies with a renewed focus on the alignment of online and virtual literacies and social communication for students with autism (Stone, 2018; Stone et al., 2019). By embracing a newer definition of literacies for students on the spectrum this means the term is not independent of inclusive guidelines, social communication, and virtual and online environments and digital technologies, such as online multiplayer games (Thomas, 2011).

Research recommendations: literacies and social communication

This paper demonstrates that New Literacy Studies has a role in broadening understandings of how students with autism can be supported through inclusive practices for literacy educational success as they engage with digital technologies in virtual worlds. The conceptual framework highlights a gap in the New Literacy Studies literature to make way for digital games and literacy research to build inclusive validity by taking into account strengths and challenges within social practices and online literacies of students on the spectrum (Bossavit and Parsons, 2018). Further research is needed to understand how educators can access students' knowledge,

strengths, and experiences and so build on the literacy skills that students bring from their homes, online literacy practices, and virtual contexts (Bourgonjon et al., 2010; Martini and Sénéchal, 2012).

This paper reviewed pertinent literature around the literacies and social communication of students on the spectrum and presented a conceptual framework of inclusive new literacies to fill a gap within the New Literacy Studies. The framework builds the foundation for continuing research using digital and gaming literacies of virtual environments to support positive social communication outcome of students on the spectrum. Future research within the fields of New Literacy Studies and inclusive education must address benefits, barriers, and challenges students on the spectrum experience in virtual contexts so up-to-date conceptual frameworks and approaches can be implemented to improve their social communication and better meet the social communication needs for this population of students. Although the research shows that video games embody many literacies and modes for social communication further research is needed about how other digital and virtual literacies align with and influence inclusive notions of social communication for students with autism.

References

- Abirached, B., Zhang, Y., Aggarwal, J., Tamersoy, B., Fernandes, T., Carlos, J., 2011. Improving Communication Skills of Children With ASDs Through Interaction With Virtual Characters. Presented at: IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH). Braga, Portugal.
- Ainscow, M., Sandill, A., 2010. Developing inclusive education systems: the role of organisational cultures and leadership. *Int. J. Incl. Educ.* 14, 401–416.
- Alvermann, D.E., Robinson, B., 2017. Youths' global engagement in digital writing ecologies. In: Mills, K., Stornaiuolo, A., Pandya, J., Smith, A. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Taylor & Francis, New York, NY, pp. 161–172.
- Alvermann, D.E., 2009. Sociocultural constructions of adolescence and young people's literacies. In: Christenbury, C.L., Bomer, R., Smagorinsky, P. (Eds.), *Handbook of Adolescent Literacy Research*. Guilford Press, New York, NY, pp. 14–28.
- American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders. DSM-5*, fifth ed. American Psychiatric Association, Arlington, VA.
- Asaro-Saddler, K., 2014. Self-regulated strategy development: effects on writers with autism spectrum disorders. *Educ. Train. Autism Dev. Disabil.* 49, 78–91.
- Ashman, A.F., 2014. *Education for Inclusion and Diversity*, fifth ed. Pearson Australia, Melbourne, Australia.
- Atkinson, A.P., 2009. Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia* 47, 3023–3029.
- Australian Bureau of Statistics, 2020. Autism in Australia. Retrieved from. <https://www.abs.gov.au/statistics/health/disability/disability-ageing-and-carers-australia-summary-findings/2018#autism-in-australia>.
- Bailey, B., Arciuli, J., Stancliffe, R., 2017. Effects of abracadabra literacy instruction on children with autism spectrum disorder. *J. Educ. Psychol.* 109, 257–268.
- Baron-Cohen, S., Tager-Flusberg, H., Lombardo, M., 2013. *Understanding Other Minds: Perspectives From Developmental Social Neuroscience*, third ed. Oxford University Press, Oxford, England.
- Baron-Cohen, S., 2000. Theory of mind in autism: a fifteen year review. In: Baron-Cohen, S., Tager-Flusberg, H., Cohen, D.J. (Eds.), *Understanding Other Minds: Perspectives From Developmental Cognitive Neuroscience*. Oxford University Press, pp. 3–20.
- Barr, A., Gillard, J., Firth, V., Scrymgeour, M., Welford, R., Lomax-Smith, J., Bartlett, D., Pike, B., Constable, E., 2008. Melbourne declaration on educational goals for young Australians. In: Ministerial Council on Education, Employment, Training and Youth Affairs. Ministerial Council on Education, Employment, Training and Youth Affairs.
- Baynard, L., 2010. Literacy for the digital generation: enabling students to develop 21st-century skills through real-world chemistry (Science Sampler). *Sci. Scope* 34, 32.
- Bean, A.F., Perez, B.I., Dymia, J.M., Kaderavek, J.N., Justice, L.M., 2019. Book-reading engagement in children with autism and language impairment: associations with emergent-literacy skills. *J. Autism Dev. Disord.* 50, 1018–1030.
- Beavis, C., O'Mara, J., McNeice, L., 2012. *Digital Games: Literacy in Action*. Wakefield Press, Adelaide, Australia.
- Beavis, C., Muspratt, S., Thompson, R., 2015. "Computer games can get your brain working": student experience and perceptions of digital games in the classroom. *Learn. Media Technol.* 40, 21–42.
- Beavis, C., 2014. Games as text, games as action. *J. Adolesc. Adult Literacy* 57, 433–439.
- Bekele, E., Zheng, Z., Swanson, A., Crittendon, J., Warren, Z., Sarkar, N., 2013. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. *IEEE Trans. Visual. Comput. Graph.* 19, 711–720.
- Ben-Sasson, A., Lamash, L., Gal, E., 2013. To enforce or not to enforce? The use of collaborative interfaces to promote social skills in children with high functioning autism spectrum disorder. *Autism* 17, 608–622.
- Bishop, S.L., Hus, V., Duncan, A., Huerta, M., Gotham, K., Pickles, A., Kreiger, Buja, A., Lund, S., Lord, C., 2013. Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. *J. Autism Dev. Disord.* 43, 1287–1297.
- Boche, B., Henning, M., 2015. Multimodal scaffolding in the secondary English classroom curriculum. *J. Adolesc. Adult Literacy* 58, 579–590.
- Boddaert, N., Chabane, N., Belin, P., Bourgeois, M., Royer, V., Barthelemy, C., Mouren-Simeoni, M.-C., Philippe, A., Brunelle, F., Samson, Y., Zilbovicius, M., 2004. Perception of complex sounds in autism: abnormal auditory cortical processing in children. *Am. J. Psychiatr.* 161, 2117–2120.
- Boon, R.T., Spencer, V.G., Deshler, D.D., 2013. *Adolescent Literacy: Strategies for Content Comprehension in Inclusive Classrooms*. Brookes, Baltimore, MD.
- Booth, T., Ainscow, M., 2011. *Index for Inclusion: Developing Learning and Participation in Schools*, third ed. Centre for Studies on Inclusive Education, Bristol, England.
- Bossavit, B., Parsons, S., 2018. Outcomes for design and learning when teenagers with autism codesign a serious game: a pilot study. *J. Comput. Assist. Learn.* 34, 293–305.
- Bourgonjon, J., Valcke, M., Soetaert, R., Schellens, T., 2010. Students' perceptions about the use of video games in the classroom. *Comput. Educ.* 54, 1145–1156.
- Boyd, L.E., Ringland, K.E., Haimson, O.L., Fernandez, H., Bistarkey, M., Hayes, G.R., 2015. Evaluating a collaborative iPad game's impact on social relationships for children with autism spectrum disorder. *ACM Trans. Access. Comput.* 7, 1–18.
- Boyle, S., McNaughton, D., Chapin, S., 2019. Effects of shared reading on the early language and literacy skills of children with autism spectrum disorders: a systematic review. *Focus Autism Other Dev. Disabil.* 34, 205–214.
- Braddock, B.A., Hilton, J.C., 2016. Arm and hand movement in children suspected of having autism spectrum disorder. *Commun. Disord. Q.* 37, 148–159.
- Caron, J.G., 2016. *Effects of Adapted Instruction on the Acquisition of Letter-Sound Correspondences and Sight Words by Pre-adolescent/Adolescent Learners With Complex Communication Needs and Autism Spectrum Disorders* (Doctoral thesis).
- Carrington, S., MacArthur, J., 2012. *Teaching in Inclusive School Communities*. Wiley, Brisbane, Australia.
- Carrington, S., 2017. Inclusive education. In: Plows, V., Whithurn, B. (Eds.), *Inclusive Education: Making Sense of Everyday Practice*. Sense, Rotterdam, Holland, pp. 233–248.
- Carroll, C., 2015. Kids With Autism Shouldn't Spend Their School Days in Survival Mode. *Chronicle-Herald*.
- Centers for Disease Control and Prevention, 2020. Autism Spectrum Disorder. Data and Statistic. Prevalence. Retrieved from. <https://www.cdc.gov/ncbddd/autism/facts.html>.
- Chakrabarty, D., 2020. Theories of the new literacy studies (NLS). *Res. J. English Lang. Literature* 8, 1–8.
- Chen, F., Tsai, C., 2016. A light fingertip touch reduces postural sway in children with autism spectrum disorders. *Gait Posture* 43, 137–140.

- Cheng, Y., Huang, R., 2012. Using virtual reality environment to improve joint attention associated with pervasive developmental disorder. *Res. Dev. Disabil.* 33, 2141–2152.
- Christensen, D.L., Baio, J., Braun, K.V.N., Bilder, D., Charles, J., Constantino, J.N., Daniels, J., Durkin, M.S., Fitzgerald, R.T., Kurzius-Spencer, M., Lee, L.-C., Pettygrove, S., Robinson, C., Schulz, E., Wells, C., Wingate, M.S., Zahorodny, W., Yeargin-Allsopp, M., Yeargin-Allsopp, M., 2016. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012. *Morb. Mortal. Wkly. Rep.* 65, 1–23.
- Clendon, S., Paynter, J., Walker, S., Bowen, R., Westerveld, M.F., 2021. Emergent literacy assessment in children with autism spectrum disorder who have limited verbal communication skills: a tutorial. *Lang. Speech Hear. Serv. Sch.* 52, 165–180.
- Comber, B., Barnett, J., 2003. *Look Again: Longitudinal Studies of Children & Literacy Learning*. Primary English Teaching Association, Sydney, Australia.
- Comber, B., 2007. Assembling dynamic repertoires of literate practices: teaching that makes a difference. In: Marsh, J., Bearn, E. (Eds.), *Literacy and Social Inclusion: Closing the Gap*. Trentham Books, Stoke-on-Trent, England, pp. 115–131.
- Committee on the Rights of Persons with Disabilities, 2016. General Comment No. 4 (2016), Article 24: Right to Inclusive Education. Retrieved from: <http://www.refworld.org/docid/57c977e34.html><http://www.refworld.org/docid/57c977e34.html>.
- Cope, B., Kalantzis, M., New London Group, 2000. *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, London, England.
- Davidson, M., Ellis Weismer, S., 2014. Characterization and prediction of early reading abilities in children on the autism spectrum. *J. Autism Dev. Disord.* 44, 828–845.
- Dindar, K., Korkiakangas, T., Laitila, A., Kärrä, E., 2016. Facilitating joint attention with salient pointing in interactions involving children with autism spectrum disorder. *Gesture* 15, 372–403.
- Doody, J., Bull, P., 2013. Evidence for impaired verbal identification but intact nonverbal recognition of fearful body postures in asperger's syndrome. *J. Autism Dev. Disord.* 43, 1652–1661.
- Dovchin, S., Pennycook, A., 2017. Digital metroliteracies: space, diversity, and identity. In: Mills, K., Stornaiuolo, A., Pandya, J.Z., Smith, A. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Taylor & Francis, New York, NY, pp. 211–222.
- Durell, S., 2014. How the social model of disability evolved. *Nursing Times* 110, 20–22.
- Dynia, J.M., Lawton, K., Logan, J.A., Justice, L.M., 2014. Comparing emergent-literacy skills and home-literacy environment of children with autism and their peers. *Top. Early Child. Spec. Educ.* 34, 142–153.
- Dynia, J.M., Brock, M.E., Justice, L.M., Kaderavek, J.N., 2017. Predictors of decoding for children with autism spectrum disorder in comparison to their peers. *Res. Autism Spectr. Disord.* 37, 41–48.
- Feiler, A., Andrews, J., Greenhough, P., Hughes, M., Johnson, D., Scanlan, M., Yee, W.C., 2007. *Improving Primary Literacy: Linking Home and School*. Taylor & Francis, Hoboken, NJ.
- Ferguson, B.R., Gillis, J.M., Seveler, M., 2013. A brief group intervention using video games to teach sportsmanship skills to children with autism spectrum disorders. *Child Fam. Behav. Ther.* 35, 293–306.
- Finke, E.H., Hickerson, B., McLaughlin, E., 2015. Parental intention to support video game play by children with autism spectrum disorder: an application of the theory of planned behavior. *Lang. Speech Hear. Serv. Sch.* 46, 154–165.
- Finke, E.H., Wilkinson, K.M., Hickerson, B.D., 2017. Social referencing gaze behavior during a video game task: eye tracking evidence from children with and without ASD. *J. Autism Dev. Disord.* 47, 415–423.
- Finke, E.H., Hickerson, B.D., Kremkow, J.M.D., 2018. "To be quite honest, if it wasn't for videogames I wouldn't have a social life at all": motivations of young adults with autism spectrum disorder for playing videogames as leisure. *Am. J. Speech Lang. Pathol.* 27, 672–689.
- Fleury, V.P., Hugh, M.L., 2018. Exploring engagement in shared reading activities between children with autism spectrum disorder and their caregivers. *J. Autism Dev. Disord.* 48, 3596–3607.
- Fleury, V.P., Whalon, K., Gilmore, C., Wang, X., Marks, R., 2021. Building comprehension skills of young children with autism one storybook at a time. *Lang. Speech Hear. Serv. Sch.* 52, 153–164.
- Foreman, P., Arthur-Kelly, M., 2017. *Inclusion in Action*, fifth ed. Cengage Learning Australia, Melbourne, Australia.
- Fullen, M., 2012. *Stratosphere: Integrating Technology, Pedagogy, and Change Knowledge*. Pearson Canada.
- Gallup, J., Duff, C., Seriani, B., Gallup, A., 2016. An exploration of friendships and socialization for adolescents with autism engaged in massively multiplayer online role-playing games (MMORPG). *Educ. Train. Autism Dev. Disabil.* 51, 223–237.
- Garcia, A., 2017. Space, time, and production: games and the new frontier of digital literacies. In: Mills, K., Stornaiuolo, A., Pandya, J.Z., Smith, A. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Taylor & Francis, New York, NY, pp. 198–210.
- Gee, J.P., Hayes, E.R., 2011. *Language and Learning in the Digital Age*. Taylor and Francis, Hoboken, NJ.
- Gee, J.P., 2007. What Video Games Have to Teach Us About Learning and Literacy. Palgrave Macmillan, Basingstoke.
- Gee, J.P., 2015a. *Social Linguistics and Literacies: Ideology in Discourses*, fifth ed. Routledge, New York, NY.
- Gee, J.P., 2015b. *Unified Discourse Analysis: Language, Reality, Virtual Worlds, and Video Games*. Routledge, Oxon, England.
- Geither, E.J.B., Meeks, L., 2014. *Helping Students With Autism Spectrum Disorder Express Their Thoughts and Knowledge in Writing: Tips and Exercises for Developing Writing Skills*. Kingsley, London, England.
- Genlott, A.A., Grönlund, Å., 2013. Improving literacy skills through learning reading by writing: the iWTR method presented and tested. *Comput. Educ.* 67, 98–104.
- Gentry, T., Kriner, R., Sima, A., McDonough, J., Wehman, P., 2015. Reducing the need for personal supports among workers with autism using an ipod touch as an assistive technology: delayed randomized control trial. *J. Autism Dev. Disord.* 45, 669–684.
- Gitsaki, C., 2015. *Current Issues in Reading, Writing and Visual Literacy Research and Practice*. Cambridge Scholars, Newcastle, England.
- Hayes, E.R., Duncan, S.C., 2012. *Learning in Video Game Affinity Spaces*. Lang, Bern, Switzerland.
- Hopkins, I., Gower, M., Perez, T., Smith, D., Amthor, F., Wimsatt, C., Biasini, F., 2011. Avatar Assistant: improving social skills in students with an ASD through a computer-based intervention. *J. Autism Dev. Disord.* 41, 1543–1555.
- Howard, S., Patti Ducoff, A., 2008. Electronic screen media for persons with autism spectrum disorders: results of a survey. *J. Autism Dev. Disord.* 38, 1499–1508.
- Hughes, C., Bernstein, R.T., Kaplan, L.M., Reilly, C.M., Brigham, N.L., Cosgriff, J.C., Boykin, M.P., 2013. Increasing conversational interactions between verbal high school students with autism and their peers without disabilities. *Focus Autism Other Dev. Disabil.* 28, 241–254.
- Hume, K., McIntyre, N., Tomaszewski, B., Odom, S., 2020. Stability of literacy profiles of adolescents with autism spectrum disorder and associations with stakeholder perceptions of appropriate high school support needs. *Lang. Speech Hear. Serv. Sch.* 52, 209.
- Humphrey, N., Lewis, S., 2008. What does "inclusion" mean for pupils on the autistic spectrum in mainstream secondary schools? *J. Res. Spec. Educ. Needs* 8, 132–140.
- Ito, M., Horst, H., Bittanti, M., Stephenson, R., Lange, P., Pascoe, C., Robinson, L., 2009. *Living and Learning With New Media: Summary of Findings From the Digital Youth Project*. MIT Press, Cambridge, Mass.
- Jacobs, D.W., Richdale, A.L., 2013. Predicting literacy in children with a high-functioning autism spectrum disorder. *Res. Dev. Disabil.* 34, 2379–2390.
- Jones, R., 2012. *Understanding Digital Literacies a Practical Introduction*. Taylor and Francis, Florence, KY.
- Jordan, A., Glenn, C., McGhie-Richmond, D., 2010. The supporting effective teaching (SET) project: the relationship of inclusive teaching practices to teachers & beliefs about disability and ability, and about their roles as teachers. *Teach. Educ.* 26, 259–266.
- Jorgensen, R., Lowrie, T., 2011. Both ways strong: using digital games to engage Aboriginal learners. *Int. J. Incl. Educ.* 17, 130–142.
- Kaartinen, M., Puura, K., Mäkelä, T., Rannisto, M., Lemponen, R., Helminen, M., Salmelin, R., Himanen, S.-L., Hietanen, J.K., 2012. Autonomic arousal to direct gaze correlates with social impairments among children with an ASD. *J. Autism Dev. Disord.* 42, 1917–1927.
- Katz, E., Girolametto, L., 2015. Peer-mediated intervention for pre-schoolers with an ASD: effects on responses and initiations. *Int. J. Speech Lang. Pathol.* 17, 565–576.

- Keefe, E.B., Copeland, S.R., 2011. What is literacy? The power of a definition. *Res. Pract. Persons Severe Disabil.* 36, 3–4.
- Kluth, P., Chandler-Olcott, K., 2008. *A Land We Can Share: Teaching Literacy to Students With Autism*. Paul H Brookes Publishing, Baltimore, MD.
- Kress, G., 2013. Multimodality a Social Semiotic Approach to Contemporary Communication. Taylor and Francis, Hoboken, NJ.
- Kress, G., 2017. What is a mode? In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*, second ed. Routledge, New York, NY, pp. 60–75.
- Kuo, M.H., Orsmond, G.I., Coster, W.J., Cohn, E.S., 2014. Media use among adolescents with autism spectrum disorder. *Autism* 18, 914–923.
- Kwok, E.Y.L., Brown, H.M., Smyth, R.E., Oram Cardy, J., 2015. Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Res. Autism Spectr. Disord.* 9, 202–222.
- Lai, C.L.E., Lau, Z., Lui, S.S.Y., Lok, E., Tam, V., Chan, Q., Cheng, K.M., Lam, S.M., Cheung, E.F.C., 2017. Meta-analysis of neuropsychological measures of executive functioning in children and adolescents with high-functioning autism spectrum disorder. *Autism Res.* 10, 911–939.
- Lanter, E., Watson, L.R., Erickson, K.A., Freeman, D., 2012. Emergent literacy in children with autism: an exploration of developmental and contextual dynamic processes. *Lang. Speech Hear. Serv. Sch.* 43, 308–324.
- Lemke, J., 2017. Multimodality, identity, and time. In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*, second ed. Routledge, New York, NY, pp. 165–175.
- Liasidou, A., 2015. *Inclusive Education and the Issue of Change: Theory, Policy and Pedagogy*. Palgrave Macmillan, Basingstoke, England.
- Light, J., McNaughton, D., 2012. Supporting the communication, language, and literacy development of children with complex communication needs: state of the science and future research priorities. *Assist. Technol.* 24, 34–44.
- Locke, J., Shih, W., Kretzmann, M., Kasari, C., 2016. Examining playground engagement between elementary school children with and without autism spectrum disorder. *Autism* 20, 653–662.
- Lucas, R., Norbury, C.F., 2018. The home literacy environment of school-aged children with autism spectrum disorders. *J. Res. Read.* 41, 197–219.
- Luke, A., Dooley, K., Woods, A., 2011. Comprehension and content: planning literacy in low socioeconomic and culturally diverse schools. *Aust. Educ. Res.* 38, 149–166.
- Mackey, T.P., Jacobson, T.E., 2008. *Using Technology to Teach Information Literacy*. Neal-Schuman, New York, NY.
- Mairena, M.Á., Mora-Guiard, J., Malinverni, L., Padillo, V., Valero, L., Hervás, A., Pares, N., 2019. A full-body interactive videogame used as a tool to foster social conducts in children with autism spectrum disorders. *Res. Autism Spectr. Disord.* 67, 101438.
- Malinverni, L., Mora-Guiard, J., Padillo, V., Valero, L., Hervás, A., Pares, N., 2014. An inclusive design approach for developing video games for children with autism spectrum disorder. *Comput. Hum. Behav.* 71, 535–549.
- Marcon, N., 2013. Minecraft as a powerful literacy prompt in the secondary English classroom. *Idiom* 49, 35–37.
- Marcon, N., Faulkner, J., 2016. Exploring Minecraft as a pedagogy to motivate girls' literacy practices in the secondary English classroom. *Engl. Aust.* 51, 63–69.
- Marsh, J., Plowman, L., Yamada-Rice, D., Bishop, J., Scott, F., 2016. Digital play: a new classification. *Early Years Int. J. Res. Dev.* 36, 242–253.
- Martini, F., Sénéchal, M., 2012. Learning literacy skills at home: parent teaching, expectations, and child interest. *Can. J. Behav. Sci.* 44, 210–221.
- Mazurek, M.O., Wenstrup, C., 2013. Television, video game and social media use among children with an ASD and typically developing siblings. *J. Autism Dev. Disord.* 43, 1258–1271.
- Mazurek, M.O., Shattuck, P.T., Wagner, M., Cooper, B.P., 2012. Prevalence and correlates of screen-based media use among youths with autism spectrum disorders. *J. Autism Dev. Disord.* 42, 1757–1767.
- Mazurek, M.O., Engelhardt, C.R., Clark, K.E., 2015. Video games from the perspective of adults with autism spectrum disorder. *Comput. Hum. Behav.* 51, 122–130.
- McCreery, M.P., Vallett, D.B., Clark, C., 2015. Social interaction in a virtual environment: examining socio-spatial interactivity and social presence using behavioral analytics. *Comput. Hum. Behav.* 51, 203–206.
- McEvoy, L., 2016. *Understanding Socialization for Young Adults With High Functioning Autism Spectrum Disorder Utilizing Massively Multiplayer Online Role-Playing Games*. ProQuest Dissertations Publishing.
- McIntyre, E., Hulan, N., Layne, V., 2011. *Reading Instruction for Diverse Classrooms: Research-Based, Culturally Responsive Practice*. Guilford Press, New York, NY.
- McIntyre, N.S., Tomaszewski, B., Hume, K.A., Odom, S.L., 2021. Stability of literacy profiles of adolescents with autism spectrum disorder and associations with stakeholder perceptions of appropriate high school support needs. *Lang. Speech Hear. Serv. Sch.* 52, 209–224.
- Merchant, G., Gillen, J., Marsh, J., Davies, J., 2014. *Virtual Literacies: Interactive Spaces for Children and Young People*. Routledge, New York, NY.
- Mills, K., 2010. *The Multiliteracies Classroom. New Perspectives on Language and Education*. Multilingual Matters, Bristol, England.
- Mitchell, P., Parsons, S., Leonard, A., 2007. Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *J. Autism Dev. Disord.* 37, 589–600.
- Nagygyörgy, K., Urbán, R., Farkas, J., Griffiths, M.D., Zilahy, D., Kökönyei, Mervó, B., Reindl, A., Ágoston, C., Kertész, A., Harmath, E., Oláh, A., Demetrovics, Z., 2013. Typology and sociodemographic characteristics of massively multiplayer online game players. *Int. J. Hum. Comput. Interact.* 29, 192–200.
- Nonnis, A., Bryan-Kinns, N., 2021. Oily: a tangible for togetherness. *Int. J. Human Comp. Stud.* 153, 102647.
- Oakley, G., 2017. Engaging students in inclusive literacy learning with technology. In: Milton, M. (Ed.), *Inclusive Principles and Practices in Literacy Education*. Emerald, Bingley, UK, pp. 159–176.
- Oliver, M., 2013. The social model of disability: thirty years on. *Disabil. Soc.* 28, 1024–1026.
- Oliver, M., 2016. Students' day-to-day engagements with technologies: rethinking digital literacies. *Irish J. Technol. Enhanced Learn.* 1, 27–33.
- Patel, V.B., Preedy, V.R., Martin, C.R., 2014. *Comprehensive Guide to Autism*. Springer, New York, NY.
- Peters, B., Forlin, C., McInerney, D., Maclean, R., 2013. Social interaction and cooperative activities: drawing plans as a means of increasing engagement for children with an ASD. *Int. J. Whole Sch.* 9, 61–86.
- Peterson, C.C., Slaughter, V., Brownell, C., 2015. Children with autism spectrum disorder are skilled at reading emotion body language. *J. Exp. Child Psychol.* 139, 35–50.
- Prensky, M., 2007. *Digital Game-Based Learning*. Paragon House, St. Paul, MN.
- Price-Dennis, D., Holmes, K.A., Smith, E., 2015. Exploring digital literacy practices in an inclusive classroom. *Read. Teach.* 69, 195–205.
- Quandt, T., Kröger, S., 2014. *Multiplayer: The Social Aspects of Digital Gaming*. Routledge, London, England.
- Ramdoss, S., Mulloy, A., Lang, R., O'Reilly, M., Sigafos, J., Lancioni, G., Didden, R., El Zein, F., 2011. Use of computer-based interventions to improve literacy skills in students with autism spectrum disorders: a systematic review. *Res. Autism Spectr. Disord.* 5, 1306–1318.
- Ravet, J., 2011. Inclusive/exclusive? Contradictory perspectives on autism and inclusion: the case for an integrative position. *Int. J. Incl. Educ.* 15, 667–682.
- Reynolds, M., Wheldall, K., Madeline, A., 2011. What recent reviews tell us about the efficacy of reading interventions for struggling readers in the early years of schooling. *Int. J. Disabil. Dev. Educ.* 58, 257–286.
- Richards, C., Burn, A.N., 2014. *Children's Games in the New Media Age: Childlore, Media and the Playground*. Ashgate, Burlington, VT.
- Ricketts, J., Jones, C.G., Happe, F., Charman, T., 2013. Reading comprehension in autism spectrum disorders: the role of oral language and social functioning. *J. Autism Dev. Disord.* 43, 807–816.
- Rivalland, J., 2004. Oral language development and access to school discourses (a case study of language acquisition). *Aust. J. Lang. Literacy* 27, 142–158.
- Rogers, L., 2013. *Visual Supports for Visual Thinkers: Practical Ideas for Students With Autism Spectrum Disorders and Other Special Educational Needs*. Kingsley, London, England.
- Russo, N., Zecker, S., Trommer, B., Chen, J., Kraus, N., 2009. Effects of background noise on cortical encoding of speech in autism spectrum disorders. *J. Autism Dev. Disord.* 39, 1185–1196.
- Saggers, B., Hwang, Y.-S., Mercer, K.L., 2011. Your voice counts: listening to the voice of high school students with autism spectrum disorder. *Australas. J. Spec. Educ.* 35, 173–190.

- Stone, B., Mills, K.A., Sagers, B., 2018. Online multiplayer games for the social interactions of children with autism spectrum disorder: a resource for inclusive education. *Int. J. Incl. Educ.* 23, 209–228.
- Stone, B., Mills, K.A., Sagers, B., 2019. Multiplayer games: multimodal features that support friendships of students with autism spectrum disorder. *Australas. J. Spec. Incl. Educ.* 43, 69–82.
- Stone, B., 2018. The Affordances of Online Multiplayer Games for the Social Interactions of Middle-Primary School-Aged Students With ASD. PhD thesis. Australian Catholic University.
- Street, B., Phal, K., Rowsell, J., 2017. Multimodality and new literacy studies. In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*. Routledge, New York, NY, pp. 227–237.
- Street, B., 2013. Literacy in theory and practice: challenges and debates over 50 years. *Theory Into Pract.* 52, 52–62.
- Tager-Flusberg, H., Kasari, C., 2013. Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Res.* 6, 468–478.
- Terlouw, G., Kuipers, D., van't Veer, J., Prins, J., Pierie, J., 2021. The development of an escape room-based serious game to trigger social interaction and communication between high-functioning children with autism and their peers: iterative design approach. *JMIR Ser. Games* 9, e19765.
- Thomas, M., 2011. *Deconstructing Digital Natives: Young People, Technology, and the New Literacies*. Taylor and Francis, Hoboken, NJ.
- Tomlinson, C.A., Imbeau, M.B., 2010. *Leading and Managing a Differentiated Classroom*. ASCD, Alexandria, VA.
- Tompkins, G.E., 2014. *Literacy for the 21st Century: A Balanced Approach*. Pearson, Harlow, England.
- United Nations, 2018. Convention on the Rights of Persons With Disabilities. Retrieved from: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- Vance, B., 2017. Video games and multimodality in first-year composition. *Coll. English Assoc. Critic* 79, 120–134.
- Vasudevan, L., Rodriguez Kerr, K., Salazar Gallardo, C., 2018. Digital young and educational justice. In: Mills, K.A., Stornaiuolo, A., Smith, A., Pandya, J.Z. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Routledge, London, pp. 263–274.
- Vigdor, J.L., Ladd, H.F., Martinez, E., 2014. Scaling the digital divide: home computer technology and student achievement. *Econ. Inq.* 52, 1103–1119.
- Waltz, M., 2013. *Autism: A Social and Medical History*. Palgrave Macmillan, Basingstoke, UK.
- Warschauer, M., Tate, T., 2017. Digital divides and social inclusion. In: Mills, K., Stornaiuolo, A., Pandya, J.Z., Smith, A. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Taylor & Francis, New York, NY, pp. 63–75.
- Wei, X., Blackorby, J., Schiller, E., 2011. Growth in reading achievement of students with disabilities, ages 7 to 17. *Except. Child.* 78, 89–106.
- Wei, X., Christiano, E.R.A., Yu, J.W., Wagner, M., Spiker, D., 2014. Reading and math achievement profiles and longitudinal growth trajectories of children with an autism spectrum disorder. *Autism* 19, 200–210.
- Westerveld, M., Paynter, J., 2021. Introduction to the forum: literacy in autism-across the spectrum. *Lang. Speech Hear. Serv. Sch.* 52.
- Westerveld, M.F., Roberts, J.M.A., 2017. The oral narrative comprehension and production abilities of verbal preschoolers on the autism spectrum. *Lang. Speech Hear. Serv. Sch.* 48, 260–272.
- Westerveld, M.F., Paynter, J., Brignell, A., Reilly, S., 2020. No differences in code-related emergent literacy skills in well-matched 4-year-old children with and without ASD. *J. Autism Dev. Disord.* 50, 3060–3065.
- Whalon, K.J., Al Otaiba, S., Delano, M.E., 2009. Evidence based reading instruction for individuals with autism spectrum disorders. *Focus Autism Other Dev. Disabil.* 24, 3–16.
- Whitton, N., 2013. *Cases on Digital Game-Based Learning: Methods, Models, and Strategies*. Information Science Reference, Hershey, PA.
- Winter-Messiers, M.A., 2007. From tarantulas to toilet brushes: understanding the special interest areas of children and youth with asperger syndrome. *Remedial Special Educ.* 28, 140–152.
- Woods, R., 2017. Exploring how the social model of disability can be re-invigorated for autism: in response to Jonathan Levitt. *Disabil. Soc.* 32, 1090–1095.
- Yeh, E., Mitric, S., 2020. Bridging activities: social media for connecting language learners' in-school and out-of-school literacy practices. *Int. J. Comput. Assist. Lang. Learn. Teach.* 10, 48–66.
- Yoon, B., 2016. *Critical Literacies Global and Multicultural Perspectives*. Springer, New York, NY.
- Zhao, J., Zhang, X., Wang, C., Yang, J., 2021. Effect of cognitive training based on virtual reality on the children with autism spectrum disorder. *Curr. Res. Behav. Sci.* 2.

Shifting from disciplinary versus content area literacy to a transdisciplinary approach

Thomas W. Bean^a, Amy Wilson-Lopez^b, and Kristen Gregory^c, ^aOld Dominion University, Darden College of Education and Professional Studies, Department of Teaching and Learning, Norfolk, VA, United States; ^bAdolescent Literacy, Utah State University, Emma Eccles Jones College of Education and Human Services, Logan, UT, United States; and ^cEast Carolina University, College of Education, Department of Elementary and Middle Grades, Greenville, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	448
Vignette	448
The future forecast	449
A brief history and epistemological issues	449
Disciplinary literacy and content area literacy	449
Examples of disciplinary literacy in practice	450
Examples of content area literacy in practice	450
A brief history—some caveats	450
Ongoing research in adolescent disciplinary literacy	451
Project Based Learning	451
A learner centered approach	451
Transdisciplinary literacy and deep comprehension	452
Post-secondary disciplinary literacy	452
Critical cosmopolitanism literacy and transdisciplinary literacy	453
References	454

Introduction

This article explores the tensions between disciplinary literacy and content area literacy by considering where the field of adolescent and college literacy should go given dramatic and fast moving technological elements, societal demands to address big questions (e.g. climate change), and the changing nature of learning. We start with a vignette set in 2030 inspired by the research of futurists (Diamandis and Kotler, 2020) and others. Following this look into the future of teaching and learning, we take up a brief history of disciplinary and content area literacy, along with a look at current research in the area and a critique based on efforts to envision what students in the future are likely to experience as they acquire transdisciplinary capital (Crow and Dabars, 2020).

Vignette

You teach high school environmental science in a collaborative core team of other faculty at your school from mathematics, history, English, art, music, and computer science as well as dedicated staff to help with technology innovation. The year is 2030 and big problems ranging from climate change to plastics pollution are on your and your students' radar. Today, you are heading to the Brock Environmental Center on the Chesapeake Bay where you and your students will learn more about the Great Pacific Garbage Patch plastics pollution, and inventions aimed at ameliorating these environmental threats. "Entitled Plastics and Microplastics: A Pressing Global Issue," today's workshop will prompt your students to think creatively about possible ways to change their dependence on plastics. One of the early inventions to capture plastic waste from the Great Pacific Garbage Patch, where 1.3 trillion items of plastic drifted between Hawaii and California, was created by a 25-year old Dutch entrepreneur, Boyan Slat. He envisioned a huge floating barrier that could use the ocean's current to corral plastic waste. You are excited to see what newer inventions are being created to reduce pollution from plastics.

Your electric, autonomous self-driving vehicle arrives from its down town parking garage to transport you to the Center as you drink your morning coffee and enjoy the view of the Lynnhaven River in Virginia Beach. Your ridesharing car is 80% less expensive than owning a vehicle and that's allowed you to purchase a condo. Unlike your college dorm room, your condo is quiet, well equipped with technology for creating lessons and units that students find engaging. You know that many of your students come to school from families struggling to make ends meet, juggling multiple jobs, and trying to carve out time and space to read and talk about their kids' interests. Fortunately, schools have embraced transdisciplinary curriculum where faculty core teams create lessons that center on questions students generate in relation to their lives that span poverty, health care, adequate housing, and

a host of other issues. These social issues are a huge part of the curriculum and today's field trip is centered on possible futures in an environment that is increasingly threatened.

The Brock Environmental Center is a model for other construction projects aimed at being sensitive to increasing threats from sea level rise and ecological extinction of many species. Constructed 16 years ago, the Center is a vanguard for green construction. At 10,000 square feet, it is fully independent and off the power grid. The rooftop photovoltaic solar panes generate 60% of the Center's energy needs and wind turbines add the remaining 40% needed to operate the Center. Rain cisterns capture and filter water for drinking and composting toilets provide human waste for rain gardens. The building itself rises 14 feet above sea level so that it may stay dry during flooding.

Most importantly for your students today, mini-courses like the one on plastics pollution encompass earth science, biology, history, art, English/writing, math, chemistry, civics, economics, government, computer science, and responsible citizenship. Environmental restoration projects include oyster, wetland, and other habitat restoration. You think, yeah, this will be a good day and a great launch pad to engage your high school students in creative ways to tackle some of these issues.

As your 20 min ride to the center continues, you daydream about your students having access to large data sets that chart trends in plastics production, use, and disposal in real time. Drawing from their understandings of data science, your students can then evaluate the extent to which scientists' current methods of measuring plastics and their impacts is adequate after learning that previous methods of measuring microplastics underestimated their presence by five to seven orders of magnitude (Brandon et al., 2020). Your students can document and reflect on their own plastics usage and compare it to that of other people. They can develop high-leverage solutions to the problem, such as through engineering new packaging for food transportation and storage rather than plastics. They could make art and spoken word poetry to engage the community in the issue. In addition, they could create multimodal narratives that move hearts and minds.

The future forecast

In the vignette set in 2030, global issues including pandemics, climate change, sea level rise, destruction of ocean coral reefs, and a host of other problems have shifted the ability of individual disciplines to tackle these complex problems. Rather, transdisciplinary teams are already reconfiguring how students and faculty organize to be able to tackle dynamic problems. "Novel transdisciplinary configurations represent institutional experiments that can recalibrate the course of inquiry and engage the application of research" (Crow and Dabars, 2020, p. 369).

At the school site level, newer school buildings are designed with collaboration in mind, as well as project based elements (e.g. gardens, rain catchment, ponds, and areas to observe these features). The expectation is that out of this transdisciplinary curriculum, new theories and practices will emerge to address global challenges. "Our collective survival as a species may be contingent on our capacity to collaborate across disciplinary boundaries" (Crow and Dabars, 2020, p. 383).

Now that we have taken a brief look into the future of education, we are going to backtrack a bit to consider where we have been in the realm of adolescent literacy and some of the tensions that appear in arguments favoring disciplinary literacy and content area literacy. In the vignette, our environmental science teacher is part of a transdisciplinary core team at the high school, collaborating across disciplinary silos in 2030. But, as we will see, this process is already underway and badly needed given a multitude of challenging world problems such as COVID-19.

A brief history and epistemological issues

Disciplinary literacy and content area literacy

As we think about learning and knowledge acquisition in content area classrooms, a key issue is the particular epistemological stance adopted by a teacher and students. In essence, some classrooms treat knowledge as a set of factoids to be covered and regurgitated. This is unfortunate because we are then in danger of developing future citizens who may well be able to survive shallow high stakes tests while unable to creatively tackle challenging global problems (Bean and Dunkerly-Bean, 2016).

In the next section we define disciplinary literacy and content area literacy and include some examples that illustrate these practices in action. Both are important as students navigate an increasingly complex array of text forms that include traditional print and, increasingly, multimodal forms including video, podcasts, virtual classrooms and beyond.

In disciplinary literacy "students use literacy to engage in goals and practices that are unique to each discipline" (Wilson-Lopez and Bean, 2017, p. 3). For example, in history students can evaluate the truth value of historical information across multiple sources.

In content area literacy, "Students learn reading and writing processes that are common across disciplines" (Wilson-Lopez and Bean, 2017, p. 2). For example, taking notes on a text or video, creating a graphic organizer of key ideas in a content area reading assignment, and other general meaning making approaches are ones that can be applied to a variety of content areas.

In the sections that follow we provide some examples of both disciplinary literacy and content area literacy practices.

Examples of disciplinary literacy in practice

Disciplinary literacy recognizes and capitalizes upon the particularities and discourse elements of various fields (Wilson-Lopez and Bean, 2017). For example, the particular ways in which an environmental scientist studies oceanic reefs, threatened with extinction, provides a model for reading and writing about how an expert might address looming environmental problems.

Reading in science as a discipline focuses on reading across multiple forms of texts including graphs, data charts, observational field notes, as a host of Internet based information (Greenleaf and Hinchman, 2020). Inquiry in science acknowledges its tentative nature as new information is discovered. In the classroom, this means that students need multiple opportunities to discuss readings and navigate multiple text forms to study issues like water pollution deeply.

Mathematics literacy practices uses formulas, graphs, symbols, and algorithms to engage students in tackling real life problems (Enderson and Colwell, 2021). For example, staghorn coral reefs in the Bahamas are dying off. At what rate is this happening and what are the possible causes of this movement toward extinction (e.g. climate change) (Trammell, 2018)? By using the Internet and applying mathematical literacy approaches to real life inquiry, students can use mathematics these tools in problem solving and discussion of possible solutions.

Many recommendations regarding disciplinary literacy have been developed based upon studies of professionals within disciplines, such as history (Wineburg, 1991), English language arts (Rainey, 2017), or physics (Bazerman, 1985). These studies have indicated that disciplinary practitioners interpret and evaluate texts using discipline-specific frameworks, and they use these texts in the context of discipline-specific knowledge construction practices. By implication, within each discipline, students can gain practice in interpreting and evaluating texts using disciplinary frameworks, and they can use these texts to solve problems or generate knowledge using methods practiced by professionals in those disciplines.

Examples of content area literacy in practice

Instead of emphasizing the interpretive frameworks and practices that are specific within a single discipline, content area literacy offers an array of strategies (e.g. graphic organizers) that compliment and increase the various concepts students encounter in multiple fields including history, mathematics, robotics, artificial intelligence, virtual reality, engineering, English, art, music, computer science and so on. While there have been debates comparing and contrasting the advantages and disadvantages both approaches offer teachers, Wilson-Lopez and Bean (2017) note that: "Content area literacy and disciplinary literacy are not mutually exclusive approaches to literacy instruction" (p. 5). Rather, traditional approaches such as notetaking, recording lectures for later study, creating graphic organizers and summaries of longer articles are all within the realm of content area reading practices. Thus, both disciplinary literacy and content area literacy can be mutually supportive of students' learning.

In an effort to shed light on the state of content area literacy amid increasing calls for attention to disciplinary literacies, Collin (2014) used Basil Bernstein's theory of the pedagogic device to deconstruct arguments for both approaches. Bernstein's research examined the sociology of school knowledge that revealed that discourses were recontextualized as they moved from one field to another. For example, the character of a discourse in one context (e.g. university level biology) is transformed when it is applied in a different context (e.g. high school biology). Collin notes: "Each effort to reset distributive rules is shaped by ideological beliefs about what schools should do for whom and why (p. 314)". In essence, in high school, students may well balance university like biology concepts with science issues applied to everyday life. In essence, a blended version of content area and disciplinary literacy are in operation in this context. And, as we noted earlier, content area literacy and disciplinary literacy are not mutually exclusive domains.

A brief history—some caveats

Nevertheless, as we write this an increasing number of studies and articles center on the unique dimensions of various disciplines, staying on a course that keeps disciplinary literacy in disciplinary silos (Dunkerly-Bean and Bean, 2016). If we compare disciplinary literacy and content area literacy, disciplinary literacy is the new kid on the block in terms of research exploring its ongoing contributions. Content area literacy influenced our field as far back as the early 1900s with an emphasis on approaches to reading and studying in the disciplines (Moore et al., 1983). During the 1970s, Hal Herber published the landmark textbook, *Teaching Reading in the Content Areas*. Hal Herber's doctoral students at the Syracuse University Reading and Language Arts Center published a series of research monographs that explored various reading strategies including a substantial body of work on graphic organizers (see Barron, 1969; Moore et al., 1983). For a more complete history of both content area and disciplinary literacy (see also Dunkerly-Bean and Bean, 2016).

In a compendium of nearly 50 years of literacy research across the disciplines, Scott et al. (2018) conducted a systematic analysis of trends spanning 1969–2017. The time frame of this body of work encompassed pivotal documents including the 1970 publication of Harold Herber's landmark textbook, as well as the National Assessment of Educational Progress (NAEP). In all, these researchers began by examining 3413 articles and their abstracts, paring this down to 53 studies that met key criteria, and ultimately arriving at 29 studies that seemed to impact the field. Not surprisingly, emergent themes included teachers' perceptions that supported the integration of content area literacy practices, as well as reported resistance to doing so. The researchers pointed to the need for additional research mapping writing within the disciplines.

Rather than rehash academic debates surrounding content area and disciplinary literacy, we want to focus this analytic argument on an increasing body of research in disciplinary literacy. In addition, we add a look at the future of disciplinary literacy from ongoing research in future studies that pay close attention to the fields of Artificial Intelligence, Robotics, and the increasing use of technology to rapidly cull information in high-speed searches.

Indeed, as we are writing this the challenges of living with a global pandemic, racism, and many peaceful protests in support of #BlackLivesMatter, along with calls for police reform occupy the news. In addition, families are home schooling children, managing online learning, and balancing their own efforts to return to jobs where they have been laid off or furloughed. Thus, while we ponder the future it is in a context where the unknown looks toward a vaccine and a possible future that is just and peaceful. Toward that end, a significant body of extant research helps inform our practices in disciplinary literacy.

Ongoing research in adolescent disciplinary literacy

An ever-expanding body of research is providing a better understanding of various epistemological approaches to disciplinary literacy.

Project Based Learning

Project Based Learning (PBL) offers an epistemological approach that provides space for students to tackle pressing problems like pollution and environmental issues (Wurdinger, 2016).

For example, Windschitl (2019) reported on 12 years of research with biology teachers aimed at transforming teacher-centered discussions into student-centered, face-to-face deliberation around scientific data. To make this work clearer, the following hypothetical example illustrates how students might engage with multiple disciplinary literacy practices while exploring the environmental problem of an invasive species.

Nutria, a large, herbivorous, semiaquatic rodent occupy wetlands where they eat and destroy marsh grasses that shelter areas from erosion. They were originally brought to the United States in the late 19th century to be raised for pelts. When that business failed, they were released into local waterways. If students in Baton Rouge, Louisiana were studying research about an invasive species like Nutria they could first view a YouTube clip providing a visual display of Nutria eating away at the grassy area of a watershed, and visit a local watershed to collect data related to the destruction of the landscape, and compare these data with local ecologists' data. In addition to YouTube clips, students could consult texts that talk about the Nutria's role in wetland erosion.

In accordance with Windschitl's recommendations, the students would hold vibrant discussions in the context of explaining phenomena, gathering known facts and data, and drawing from their everyday knowledge (e.g. Nutria do not attack people and they are nocturnal feeders). In this curriculum model, students are expected to use source material (including online searches), to challenge competing ideas, find evidence and defend their knowledge to other students, and, ultimately, build and evaluate collective knowledge and methods relative to problems of importance (e.g. evaluating wildlife biologists' solutions for eradicating nutria).

In thinking about curriculum design in the academic disciplines, Deng (2013) argued that it is far too easy to center students' efforts on concept learning devoid of sociocultural dimensions. "Curriculum development framed by this approach ignores the interests, attitudes, and feelings of learners" (p. 42). Deng especially worried that issues of social reform will be left out of disciplinary literacy. A learner-centered approach may be preferable to a disciplinary approach in Deng's view.

A learner centered approach

A good example of this stance is a middle school social studies unit that takes a social justice focus where middle school students explored their own local history via multimodal complex text sets (Kucan et al., 2019). The unit focused on events related to the Hill District in Pittsburgh, Pennsylvania set in the 1930s and 1940s. Two 8th grade social studies teachers engaged their students in an exploration of the Hill District neighborhood, a vibrant African-American community. It was often called "Little Harlem" because of its rich cultural offerings, but in the 1960s urban renewal destroyed the neighborhood to make room for a Civic Arena. A new highway was constructed that effectively cut off the neighborhood from downtown.

Students were able to consult an online digital repository to study the Hill's history and ultimate demise. This research helped them understand current race relations, economic conditions, and housing conditions in the present day. During the course of the 15-lesson unit, students read primary source material and raised their own questions about the Hill, then and now. Their multimodal projects highlighted the creative ways the Hill residents developed their own businesses and entertainment venues because they were denied access to the downtown businesses. Students discussed and undertook projects illustrating the neighborhood at the peak of its history, offering claims and evidence to support their research. As a culminating activity they argued for a museum that would honor the Hill's rich history. In essence, their voices propelled this research and related projects, chronicling events that were both local and beyond, still occurring in many cities where project housing is being gentrified to make way for urban renewal.

Engineering, too, is a discipline that lends itself to learner-centric approaches wherein youth can imagine new possibilities for their communities and work to effect change. Wilson-Lopez and Minichiello (2017), for example, described middle school youth who sought to improve their communities through designing a safer parking lot for their middle school. Guided by their technology and engineering teacher, they gathered, analyzed, and represented qualitative and quantitative data related to parking lot usage; interviewed community members and other stakeholders in their home languages regarding their experiences; analyzed aerial images of the existing parking lot and used computer-aided drawing (CAD) tools to design new solutions; identified and weighed criteria

and constraints needed for a successful solution; and ultimately used rhetorically powerful images and words to communicate their proposed solutions to a community council after rounds of revision and feedback from their teacher.

While we may imagine that engineers design things with function primarily in mind, the field is much more complicated than that. Cognitive scientist [Norman \(2004\)](#) has spent his life paying attention to how people respond to everyday products (e.g. washing machines). When we encounter a bad design, like a washing machine that leaks, we become angry. In contrast, when an invented object works (e.g. iPhone text messaging), we smile. The problem with bad designs in the machine world is that these flaws frequently cause errors in use.

As suggested by these examples, disciplinary literacy and learner-centric approaches that embrace social reform are not mutually exclusive. On the contrary, youth can use disciplinary practices (e.g. analyzing a variety of primary sources to determine multiple perspectives), as well as more generalizable content area literacy strategies (e.g. revising word choice based on feedback), as tools for addressing social issues of personal, local, and global significance. Literacy pedagogies, as embedded within different content areas, can support youth in using disciplinary lenses and approaches, as well as hybrid approaches that cross disciplines, as well as community and school boundaries. Increasingly, the most pressing problems of today and tomorrow demand transdisciplinary approaches to envision and enact better solutions, and this demand can be met in part through creative teachers enacting innovative literacy pedagogies oriented toward change within and across disciplines.

Transdisciplinary literacy and deep comprehension

Transdisciplinary literacy can be practiced naturally within Project Based Learning. Transdisciplinary literacy taps into multiple disciplinary literacy knowledge spanning more than a single discipline. Student Project Based Learning projects are increasingly in play across content areas including English Language Arts, history and social studies, the sciences and mathematics ([Spires et al., 2016](#)). Based on [Dewey's \(1927\)](#) and [Kilpatrick's \(1918\)](#) early work on the value of student inquiry, and more recent work in Project Based Learning (e.g. [Wurdinger, 2016](#)), [Spires et al. \(2019\)](#) describe a five step process for instituting inquiry projects in various content areas.

In order to truly “own” and comprehend content at a deep level, students can engage in the following 5 steps over the span of a unit: (1) Ask a compelling question; (2) Gather and analyze resources; (3) Creatively synthesize claims and evidence; (4) Critically evaluate and revise findings; and, (5) Share, publish and act on a problem.

For example, in a 10th grade biology class, students used the five inquiry steps to explore local water ecology on a North Carolina river ([Spires et al., 2019](#)) and engaged in disciplinary literacy practices specific to science, English language arts, mathematics, and civics. Students practiced testing pH levels (how acidic the water is with pH values less than 7 indicating acidity) in the school pond as an initial starting point before creating their own compelling question: “How does water pollution affect the ecosystem of the North Carolina Upper Neuse River?” The river is located in the Piedmont area of North Carolina. It provides Raleigh’s major drinking and recreational water source. The river has experienced pollution, resulting in acidity and related algae blooms.

The biology students, organized in teams of four, used online resources to construct annotated bibliographies. They evaluated the credibility of sources that indicated agricultural run-off was responsible for the river’s current pollution. They constructed an Animoto YouTube clip as a public service announcement of their findings and presented this production to other stakeholders including parents and members of the community. The authors note that: “When students address real-world problems that interest them, they are more likely to interact with information inquisitively and curiously, leading to authentic learning incomes” (p. 159). In essence, a well-constructed inquiry project that includes transdisciplinary literacy experiences leads to deep comprehension.

Transdisciplinary literacy practices can be applied to teamwork across disciplines within community college and college reading practices. While silos remain, some promising research suggests that faculty would be willing to embrace more collaboration ([Bean et al., 2018](#)).

Post-secondary disciplinary literacy

In this section we take up the role of post-secondary disciplinary literacy, including community colleges and higher education more generally. In its current form disciplines within higher education often cluster together in their own silos while content area literacy practices advance generic strategies ([Dunkerly-Bean and Bean, 2016](#); [Draper, 2010](#)). For example, recording data in a scientific notebook is typically discipline specific while creating a Venn diagram can be applied to information in multiple content areas. The reality in post-secondary settings is that there is often limited interaction and collaboration between faculty across disciplines, literacy faculty, and literacy professionals.

While researchers ([Gregory and Bean, 2020](#); [Heller, 2010](#)) argue that disciplinary literacy skills are essential in post-secondary and professional settings, “it has yet to catch fire in the content fields of postsecondary education or in developmental reading circles” ([Armstrong and Stahl, 2017](#), p. 114). In order for disciplinary literacy to “catch fire,” the status quo of college disciplinary isolation must change ([Bean et al., 2018](#)). Increasing transdisciplinary collaboration across disciplines, reducing isolated work in silos, participating in collaborative professional development, and curriculum redesign are all steps that can be taken toward changing this status quo ([Bean et al., 2018](#); [Gregory et al., 2019](#); [Gregory and Bean, 2020](#)).

For example, Gregory and Bean (2020) examined disciplinary literacy practices specific to the trades from a sociocultural perspective. In community college mechatronics and diesel mechanics courses, for example, disciplinary literacy was viewed as a social practice influenced by the discipline's discourse (Dunkerly-Bean and Bean, 2016). In order to diagnose and solve problems related to ships, circuits, and instrumentation, students immersed themselves in the robust literacy and disciplinary practices of the field. Through this process, they moved toward becoming members of the discipline, in essence "insiders" in a field of practice (Paulson and Holschuh, 2018). From a sociocultural perspective, as students engaged in disciplinary literacy practices, such as using manuals to repair diesel engines, communicating with clients, or collaboratively repairing broken circuits, they were acquiring critical social capital (Bourdieu, 1986). "Indeed, their ability to flourish in the trades rests on knowing the discourse and procedures for solving problems in the field" (Gregory and Bean, 2020, p. 4).

Faculty across higher education have a great opportunity to provide disciplinary literacy support and instruction for their students. However, they often do not have the background in literacy theory and pedagogy that is greatly needed to best support their students with disciplinary literacy. They may view literacy as basic reading and writing rather than the intricacies of reading discipline-specific texts (e.g. manuals) and writing discipline-specific documents (e.g. lab reports) (Gregory and Bean, 2020). Community college faculty, in particular, have limited time, if any, to learn disciplinary teaching strategies due to high course loads (Gregory and Colclough, 2018).

Both disciplinary and transdisciplinary professional learning communities can provide a space for faculty to explore disciplinary literacy practices and teaching practices (Bean et al., 2018; Bergman, 2014; Furco and Moely, 2012; Jacobs, 2005, 2007). The essential difference is that in disciplinary professional learning communities faculty in this model come from a single discipline (e.g. English) while in a transdisciplinary model, multiple disciplines collaborate to update curriculum and research (Crow and Dabars, 2020).

Jacobs (2005) emphasizes the importance of creating such spaces for faculty to collaborate on neutral ground while sharing leadership and responsibility for increasing student success. In such a space, discipline and literacy faculty can work together to expand their knowledge of disciplinary literacy and how to effectively make a curricular change to better prepare students to be career ready, critically socially minded, and well-versed in disciplinary and transdisciplinary problem solving. Further, transdisciplinary learning communities could provide a space for project-based inquiry and cross-curricular initiatives that would expand students' experience with and knowledge of transdisciplinary literacy.

Critical cosmopolitanism literacy and transdisciplinary literacy

Cosmopolitanism has a long, and dynamic history, most recently applied to education and our increasingly global and interconnected world where students can, via the Internet, communicate and collaborate over vast distances (Bean and Dunkerly-Bean, 2020; Hansen, 2014; Stornaiuolo and Nichols, 2019). In its most recent iteration, critical cosmopolitan literacy (CCL) can be defined as a set of practices to engage students in the critical examination of global and local issues of power, access, and social justice (Bean, 2016). Across multiple disciplines, CCL propels student conversations into a larger sphere where issues of basic human rights, marginalization and oppression can be interrogated.

CCL can drive Project-Based Inquiry to expand beyond nation state borders to address global issues including pandemics, poverty, water and sanitation initiatives, climate change, immigration, public health, and indigenous funds of knowledge (Spires et al., 2019). At the core of this cosmopolitan view is a predisposition that encourages caring toward others from both local and global contexts (Appiah, 2008). As part of their project-based inquiry, students engage in transdisciplinary literacy skills and strategies.

For example, one of the United Nations 17 sustainable development goals is centered on access to clean water for people. An 11th grade English teacher, Ms. Bianchi, engaged her students in argumentative writing centered on expanding their prior knowledge gleaned from natural resource conflict studies in other disciplinary classes including science (Spires et al., 2019). To begin this 6-week unit, Ms. Bianchi and her students read the novel, *A Long Walk to Water* (2010) by Linda Sue Park that features two 11-year-olds, Salva, a boy, and Nya, a girl navigating political conflict in South Sudan as they search for clean water.

Following this initial common reading, students worked in teams with 3–5 students in each team that met 3 times each week in their 60 min block schedule followed by 30 min of teacher led lessons or formative assessments. Each team generated their own compelling questions to start this phase of the unit. For example, one team explored: "How does access to clean water and adequate sanitation impact women's health" (Spires et al., p. 10).

To actually engage her students in intercultural global dialog, Ms. Bianchi brought in experts on global water and sanitation from different countries into her classroom both virtually (via the Internet) and in person to work with the student teams. Thus, the teams had access to a representative from the Organization of Water for Sudan, a civil engineer from the Philippines, and a waterborne disease specialist from Australia, as well as other experts from China. To prepare for their conversations with these experts in the field, students consulted online biographies and published research bibliographies for each expert.

Ultimately, each student team created an annotated bibliography, an argumentative writing piece to weigh warrants and claims about water and sanitation from a global perspective, and a multimodal digital product using Google Docs where they had team space to collaborate. In line with cosmopolitanism and CCL, they had to consider the credibility of various sources of information.

For their final projects they could choose from a variety of multimodal applications including Animoto, Voice Thread, and others. Their audience consisted of community stakeholders, family, and friends. They also had an open access site for posting their work. In addition, they established a Water for South Sudan donation site as part of a worldwide effort to raise money for drilling more wells in South Sudan. This organization was founded by Salva Dut, the main character in *A Long Walk to Water*, bringing this 6-week global unit full circle.

We would be remiss if we did not note why cosmopolitan theory and related activism is important. As we write this chapter the world is in the midst of a global pandemic brought on by the rapidly transmitted Covid-19 Virus that has shuttered businesses, restricted travel, and sequestered students and their families in voluntary and involuntary quarantines. In essence, cosmopolitanism embraces global citizenship and transnational efforts to solve complex problems (Appiah, 2006). For example, access to clean water is a global issue. The [Global Water Project](http://globalwater.org/global-projects) (<http://globalwater.org/global-projects>) centers its efforts on Central and South America, and Africa. This includes hand washing stations for schools, spring catchment systems, hand dug wells, and the involvement of local communities. This is transdisciplinary work around daunting issues that connect all of us beyond nation state borders. All facets of education need to be embracing a future that is global and interconnected and badly in need of creative, critical thinkers.

References

- Appiah, K.A., 2006. *Cosmopolitanism: Ethics in a World of Strangers*. W. Norton & Company, New York.
- Appiah, K.A., 2008. Education for global citizenship. *Yearbk. Natl. Soc. Stud. Educ.* 107 (1), 83–99.
- Armstrong, S.L., Stahl, N.A., 2017. Communication across the silos and borders: the culture of reading in a community college. *J. Coll. Read. Learn.* 47 (2), 99–122.
- Barron, R.F., 1969. The use of vocabulary as an advance organizer. In: Herber, H.L., Sanders, P.L. (Eds.), *Research in Reading in the Content Areas: First Year Report*. Syracuse University Reading and Language Arts Center, New York, pp. 34–48.
- Bazerman, C., 1985. Physicists reading physics: schema-laden purposes and purpose-laden schema. *Writ. Commun.* 2, 3–23.
- Bean, T., Dunkerly-Bean, J., 2016. At the intersection of creativity and civic engagement Adolescents' literacies in action. *J. Adolesc. Adult Lit.* 60 (3), 247–253.
- Bean, T.W., Dunkerly-Bean, J.M., 2020. Cosmopolitanism critical literacy and youth civic engagement for human rights. *Pedagogies* 262–278.
- Bean, T., Gregory, K., Dunkerly-Bean, J., 2018. Disciplinary literacy. In: Flippo, R., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 89–97.
- Bean, T.W., 2016. Digital media and cosmopolitan critical literacy: research and practice. In: Guzzetti, B., Lesley, M. (Eds.), *Handbook of Research on the Societal Impact of Digital Media*, pp. 46–68.
- Bergman, L., 2014. The research circle as a resource in challenging academics' perceptions of how to support students' literacy development in higher education. *Can. J. Action Res.* 15 (2), 3–20.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, New York, pp. 242–258.
- Brandon, J.A., Freibott, A., Sala, L.M., 2020. Patterns of suspended and salp-ingested microplastic debris in the North Pacific investigated with epifluorescence microscopy. *Limnol. Oceanogr. Lett.* 5 (1), 46–53.
- Collin, R., 2014. A Bernsteinian analysis of content area literacy. *J. Lit. Res.* 46 (3), 306–329.
- Crow, M.M., Dabars, W.B., 2020. *The Fifth Wave: The Evolution of American Higher Education*. John Hopkins University Press, Baltimore, MA.
- Deng, Z., 2013. School subjects and academic disciplines: the differences. In: Luke, A., Woods, A., Weir, K. (Eds.), *Curriculum, Syllabus Design and Equity: A Primer and Model*. Routledge, New York, pp. 40–53.
- Dewey, J., 1927. *The Public and Its Problems*. Shallow, Athens, OH.
- Diamandis, P.H., Kotler, S., 2020. *The Future is Faster Than You Think: How Converging Technologies are Transforming Business, Industries, and Our Lives*. Simon & Shuster, New York.
- Draper, R.J., 2010. (Re) Imagining Content-Area Literacy Instruction. Teachers College Press, New York.
- Dunkerly-Bean, J., Bean, T.W., 2016. Missing the savoir for the connaissance: disciplinary and content area literacy as regimes of truth. *J. Lit. Res.* 48 (4), 448–475.
- Enderson, M.C., Colwell, J., 2021. Considering possibilities to promote disciplinary literacy instruction in mathematics. *J. Adolesc. Adult Lit.* 64 (6), 683–692.
- Furco, A., Moely, B.E., 2012. Using learning communities to build faculty support for pedagogical innovation: a multi-campus study. *J. High. Educ.* 83 (1), 128–153.
- Global Water: Overcoming Hunger, Disease, and Poverty. <http://globalwater.org/globalprojects/>.
- Greenleaf, C., Hinchman, K., 2020. Expanding teaching and learning with disciplinary texts: the case of reading and science. In: Moje, E.B., Afflerbach, P.P., Enciso, P., Lesaux, N.K. (Eds.), *Handbook of Reading Research*, vol. V. Routledge, New York.
- Gregory, K., Bean, T., 2020. Trades-based literacy: community college faculty's exploration of disciplinary literacy. *J. Coll. Read. Learn.* 51 (2).
- Gregory, K., Colclough, M., 2018. Community college discipline faculty perceptions of role as literacy educators. *Inquiry* 21 (1), 1–15.
- Gregory, K., Bol, L., Bean, T., Perez, T., 2019. Community college faculty's attitudes and self-efficacy with literacy instruction in the disciplines. *J. Behav. Soc. Sci.* 6 (1), 14–28.
- Hansen, D.T., 2014. Cosmopolitanism and cultural creativity: new modes of educational practice in globalizing times. *Curric. Inq.* 44 (1), 1–14.
- Heller, R., 2010. In praise of amateurism: a friendly critique of Moje's "Call for Change" in secondary literacy. *J. Adolesc. Adult Lit.* 54 (4), 267–273.
- Jacobs, C., 2005. On being an insider on the outside: new spaces for integrating academic literacies. *Teach. High. Educ.* 10 (4), 475–487.
- Jacobs, C., 2007. Mainstreaming academic literacy teaching: implications for how academic development understands its work in higher education. *S. Afr. J. High. Educ.* 21 (7), 870–881.
- Kilpatrick, W.H., 1918. *The Project Method*, vol. 19. Teachers College Record, New York, pp. 319–335.
- Kucan, L., Rainey, E., Cho, B.Y., 2019. Engaging middle school students in disciplinary literacy through culturally relevant historical inquiry. *J. Adolesc. Adult Lit.* 63 (1), 15–27.
- Moore, D.W., Readence, J.E., Rickelman, R.J., 1983. An historical exploration of content area reading instruction. *Read. Res. Q.* 18 (4), 419–438.
- Norman, D.A., 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Perseus Books Group, New York.
- Park, L.S., 2010. *A Long Walk to Water*. Houghton Mifflin, New York.
- Paulson, E.J., Holschuh, J.P., 2018. College reading. In: Flippo, R.F., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 27–41.
- Rainey, E.C., 2017. Disciplinary literacy in English language arts: exploring the social and problem-based nature of literary reading and reasoning. *Read. Res. Q.* 52 (1), 53–71.
- Scott, C.E., McTigue, E.M., Miller, D.M., Washburn, E.K., 2018. The what, when, and how of preservice teachers and literacy across the the disciplines: a systematic literature review of nearly 50 years of research. *Teach. Teach. Educ.* 73, 1–13.
- Spire, H.A., Kerkhoff, S.N., Graham, A.C.K., 2016. Disciplinary Literacy and inquiry: teaching for deeper content. *J. Adolesc. Adult Lit.* 60 (2), 151–161.

- Spires, H.A., Himes, M.P., Paul, C.M., Kerkhoff, S.N., 2019. Going global with project-based inquiry: cosmopolitan literacies in practice. *J. Adolesc. Adult Lit.* 62 (1), 51–64.
- Storaiuolo, A., Nichols, T.P., 2019. Cosmopolitanism and Education. *Oxford Research Encyclopedia of Education*. Oxford University Press, USA.
- The Chesapeake Bay Brock Environmental Center-International Living Future Institute. Retrieved from living-future.org.
- Trammell, P., 2018. *Journey to the Ragged Islands: Sailing Solo Through the Bahamas*. Middletown, DE: P. Trammell.
- Wilson-Lopez, A., Bean, T.W., 2017. Content Area and Disciplinary Literacy: Strategies and Frameworks. International Literacy Association, Newark, DE. Retrieved from. <https://literacyworldwide.org>.
- Wilson-Lopez, A., Minichiello, A., 2017. Disciplinary literacy in engineering. *J. Adolesc. Adult Lit.* 61 (1), 7–14.
- Windschitl, M., 2019. Disciplinary literacy versus doing school. *J. Adolesc. Adult Lit.* 63 (1), 7–13.
- Wineburg, S., 1991. Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *J. Educ. Psychol.* 83 (1), 73–87.
- Wurdinger, S.D., 2016. *The Power of Project-Based Learning: Helping Students Develop Important Life Skills*. Rowman & Littlefield, New York.

(Mis)understanding adult literacy

Erik Jacobson, Montclair State University, Montclair, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Adult literacy as a problem	456
The workforce development model	456
Life skills	457
International surveys of adults' literacy skills	457
Adult literacy as a paradigm	458
Functional literacy	458
Adult literacy and identity formation	458
Non-valorized practices	459
Adult literacy as political praxis	459
Adult literacy and political struggle	459
Problematizing technology	460
Conclusion	461
References	461

Despite being the focus of research and policy for over a century, what is meant by the term *adult literacy* remains a bit unclear. In part, this is due to the fact that examinations of adult literacy have been approached from a number of different directions. Formal adult education programs have tended to conceptualize literacy as a set of skills and proficiencies, with the corresponding goal being establishing or enhancing those skills. By contrast, anthropologists and others interested in the role of culture in human development have looked to what adults do with reading and writing to gain insights into different communities and contexts. Finally, there is a well-established tradition of educators and activists making connections between literacy and systemic inequities a central part of their analysis. An examination of each approach will suggest some of the ways that adult literacy has been both understood and misunderstood, and what implications might be for those who are interested in this population of learners.

Adult literacy as a problem

In educational policy and planning discourse, the term *adult literacy* is most often associated with concerns about the reading, writing and numeracy proficiency of adults. Across the globe, municipalities, states and nations fund programs designed to help adults improve their skills, complete compulsory education and/or prepare for employment. Outside of the public sector, there are also volunteer and community-based organizations of various sizes who train tutors and advocate on adult learners' behalf. Because the focus is on the struggle that some adults experience, in this context *adult literacy* functions like a contronym. That is, an interest or commitment to *adult literacy* typically signifies a concern about its potential absence.

The workforce development model

As with many other educational issues, analyses of the state of adults' literacy skills are informed by broader ideological tensions. Within the United States, Canada, the UK and other parts of Europe, support for adult education programs has tended to be framed in economic terms. For example, it is suggested that because employers cannot fill positions due to what has been called a "skills gap," funding adult education programs is a necessary investment in the economy (National Commission on Adult Education, 2008; Goodman et al., 2015). The expectation is that programs will help adults as potential employees develop the skills and dispositions to make them attractive to employers. This is consistent with the logic of human capital development – adults here are understood to be a resource to be cultivated and utilized, and their literacy skills are thus evaluated in terms of economic outcomes. Within this framing, adults with limited literacy skills are blamed for both their own economic vulnerability and their nation's. Indeed, Brine's (2006) review of European Commission policy statements on lifelong learning suggests that, "low knowledge-skilled workers are not only those at risk, they are increasingly constructed as the risk" (p. 649).

In the United States, this commitment to a workforce development model has explicitly shaped federal policy regarding federal support of adult education since the passage of the Workforce Investment Act in 1998 (see Jacobson, 2017, for a summary). This law was updated in 2014 and renamed the Workforce Investment and Opportunity Act, but its core elements remain the same. For example, key outcome metrics for program evaluation still include students gaining employment, retaining employment and/or receiving an increase in wages. In keeping with these goals, adult education providers are expected to help create career pathways for their students and to make the content of their courses responsive to demands of particular local industries (e.g., shipping and logistics, manufacturing, etc.). In what are called Integrated Education and Training (IET) Programs, students combine work on their

literacy and/or English skills with targeted vocational training. This approach is consistent with what has been referred to as *contextualized instruction* - connecting classroom work to adults' lives outside of class, rather than studying subjects in an abstract manner.

As noted above, this neoliberal workforce approach has also been taken up in other countries, often resulting in a shift away from more holistic or humanistic models of adult education. For example, *The Skills for Life* agenda in the United Kingdom has also been criticized for having a narrow focus on employment and for making the ability to access social welfare services dependent upon being "employable" (Hamilton and Pitt, 2011; Allatt and Tett, 2019). Similar criticisms have been made of the *National Skills Foundation Strategy* in Australia (Yasukawa and Osmond, 2019). In Canada, the language of accountability policies is explicit with regards to the goals. Programs are required to report on "training and skills development instead of learning and education" in order to continue receiving funding (Thériault, 2019, p. 19). Here again, it is only in their position as potential employees that adults are deemed worthy of support.

Although the idea of a skills gap is popularly used to explain economic volatility and vulnerability, research suggests that the situation is more complex. Indeed, rather than struggling to find skilled workers for their open positions, some employers report having a hard time finding skilled workers who are willing to accept the wages being offered (Goldstein, 2012; Pellizzari and Fichen, 2013). Here, rhetoric about a pervasive lack of skills is used as cover for a refusal to pay appropriate wages (Gilpin, 2014; Levine, 2013). Many of these same employers also note that they are reluctant to hire workers who have "union-type experience" (Gilpin, 2014, p. 17). Rather than literacy skills being the issue, it is a matter of disposition, as an interest in organizing co-workers makes a potential worker less attractive to the employer. Adult literacy is thus misrepresented and subsequently misunderstood as the problem, moving the analysis away from macroeconomic factors that shape the labor market.

Life skills

Outside of the economic sphere, adult literacy understood in its contronymic form is also associated with other negative outcomes. For example, there is evidence that lower levels of literacy are associated with poorer health (Dewalt et al., 2004). This includes being diagnosed with more advanced stages of prostate cancer (Baquet et al., 1991; Tobias-Machado and Cooperative Brazilian Uro-Oncology Group, 2013), being less likely to access screenings like pap smears and mammograms (Scott et al., 2002) and increased risks of hospitalizations (Baker et al., 2002). In response, adult educators have worked to contextualize literacy instruction within public health efforts (Santos and Paasche-Orlow, 2019), including programs devoted to smoking cessation, diabetes awareness and COVID prevention. These projects are designed to provide adults with the skills necessary to become and stay informed about their own health and the health of those they care about.

Similarly, there is concern about some adults' ability to understand and work with important financial documents. Here again, ideological differences shape the identification of the problem and potential responses to it. For example, Lusardi's (2019) discussion of financial literacy focuses on the ability to do interest rate calculations, understanding inflation and managing risk diversification. This misunderstands or misrepresents the economic conditions most adults find themselves in, as the investment class remains a small fraction of global society. By contrast, the *Financial Literacy and Education Commission in the United States* (2016) pays particular attention to the needs of economically vulnerable populations. The focus of their work is increasing adults' ability to budget, plan for unexpected outcomes and access additional resources when in crisis. For both health and financial literacy efforts, programs vary as to how much they attend to structural versus personal dimensions. Healthy eating may be framed as an individual choice, or an issue of affordable access to healthy foods. Financial literacy may be conceptualized as a matter of surviving with limited means, or developing a strategy to change the economic system that produces that scarcity for exploited populations.

In recent times, there has also been a growing concern about the ability of adults to use computer or digital resources (National Skills Coalition, 2020; Yoon et al., 2020). At times, this has been suggested to mainly be a function of age. For example, nearly two decades after being coined by Prensky (2001), the *Digital Immigrant and Digital Native* dichotomy lives on in popular imagination and policy discourse. Digital Natives – that is to say, younger people – are described as having a natural facility with technology because they were born into a digital world. Digital Immigrants – older adults – were born prior to the advent of the digital age and are thought to be behind the times and struggling to navigate a new reality. In addition to the problematic use of *immigrant* as a metaphor for having a lack of skills or awareness, this assumption perpetuates a deficit model that educators rightly reject in other contexts. Although it is certainly true that some adults find it harder to pick up digital skills, it is not inherent to being older. Many older people learn to use technology on their own terms, and they generally don't need schooling to do so (Selwyn et al., 2006). Indeed, research suggests that in addition to motivation, adults need time and connectivity, and these are not equally distributed. Promoting the idea that age is the key factor distracts from issues with inequitable resource allocation that are endemic to capitalism.

International surveys of adults' literacy skills

Perhaps the clearest indication that adult literacy is considered a problem to be solved is the attention paid to international surveys of adults' skills. For example, the Program for the International Assessment of Adult Competencies (PIAAC) conducted by the Organization for Economic Co-operation and Development (OECD) examines adults' ability to read and understand texts, use mathematics and solve problems (OECD, 2021). Nations are celebrated or critiqued depending upon their place in the rankings and poor results can be part of calls for additional funding for education programs (European Association for the Education of Adults, 2014). Consistent with the human capital approach, Schwerdt et al. (2020) use PIAAC results to investigate how investments in

adult education positively impact Gross Domestic Product. However, the same results can be used to call attention to what social factors might play a role in poor outcomes [Trawick \(2019\)](#). Thus, if adult literacy is understood to be a problem, does it need to be solved for the sake of the economy, or to improve the lives of adults facing precarity?

Adult literacy as a paradigm

The construct of *adolescent literacy* provides a helpful contrast to adult literacy. Central to this concept is the idea that there is something distinct about the reading and writing practices of adolescents. In part, this can be ascribed to the developmental stages they are passing through. Just as important, however, are particular literacy practices that are associated with the young ([Herrera et al., 2016](#)). This includes both formal activities they are expected to engage in (e.g., the genre of school-based essays) and informal activities they do by choice (e.g., chatting while playing online games, reading and writing zines, texting, etc.). By expanding the scope of analysis beyond institutionally valorized practices, researchers have attempted to get a sense of the lifeworld of adolescents and how literacy plays a role in the construction of adolescent identity ([Knobel, 1999](#); [Ito et al., 2019](#)). One goal of this research is to better support adolescent literacy development in formal settings by re-imagining what can be accomplished.

Functional literacy

In comparison, work that has attempted to understand the role that literacy plays in the construction of adult identity is oddly bifurcated. One influential approach is that of *functional literacy*. As with health literacy and financial literacy, the goal here is to move away from decontextualized lessons by focusing on helping adults accomplish literacy-related tasks in their everyday lives. Tests and surveys of functional literacy include various genres of texts adults engage with (e.g., menus, maps, receipts and forms) and include tasks such as locating information, making inferences and connecting text ([White, 2011](#)). Not surprisingly, this is contested terrain. The skills adults need in order to be seen as functioning depend upon how adulthood is conceptualized. Even projects intended to expand beyond simply economic outcomes may work within narrow frameworks. For example, in the 1990's, the Equipped for the Future project in the United States promoted a multi-dimensional vision of functional literacy, identifying literacy skills adults needed in their roles as workers, family members, community members and as learners. Although broader than a purely economic model, this framework still centered adult literacy practices around what are seen as productive roles or what the adult is expected to contribute. There was an implicit prioritization of what adults' responsibilities were to their family, their community and their employer. The focus was squarely on what adults *need* to do with literacy, not what they might *want* to do with it.

In sharp contrast, adult literacy as conceptualized by researchers working outside of the field of formal literacy pedagogy is much broader and more sensitive to the diverse ways that adults use texts. Indeed, the social practice turn in literacy studies was founded in large part upon studies of adults' literacy practices. [Scribner and Cole's \(1981\)](#) work with the Vai, [Street's \(1984\)](#) work in Iran and [Heath's \(1983\)](#) work with families in United States all study the ways that adults use literacy in their daily lives, and how those practices are shaped by a variety of sociocultural and socioeconomic factors. In [Barton and Hamilton's \(1998\)](#) survey of the literacy practices of adults in Lancaster, in addition to "getting things done in the community" (e.g., organizing meetings, creating newsletters, making flyers for yard sales), reading and writing were taken up as part of "leisure and pleasure" (e.g., getting lost in a book, writing stories, being a fan or collector). This research helps emphasize the fact that adults engaging in literacy practices are not completely defined by their roles as workers and economic beings. Moreover, as with adolescents, adults' literacy practices are also a part of expressions of identity.

Adult literacy and identity formation

For example, there is a rich body of literature that examines the nature of African American women's literacy practices. Researchers have explored how shared social rules and expectations about reading and writing shape relationships between African American women (and relationships between African American women and others) (e.g., [Lanehart, 2002](#); [Baugh and Smitherman, 2007](#)). Because these cultural practices are diverse, there is not a single "African American women's literacy practice." However, [Richardson \(2003\)](#) helpfully suggests that the concept of African American female literacies "refers to the ways of knowing and acting and the development of skills, vernacular expressive arts and crafts that help females to advance and protect themselves and their loved ones in society" (p. 77). Although not reducible to linguistic structures, studies have looked at how African American women's literacies are informed by their use of African American Vernacular English ([Redd and Webb, 2005](#)), particular rhetorical structures ([Ball, 1992](#); [Lee, 2007](#)), and the strategic use of silence ([Franklin and Dowdy, 2005](#), citing the work of [Rybarczyk and Bellig, 1997](#)). Indeed, what is left unsaid can function as a powerful critique of racial capitalism, patriarchy and the privileging of so-called "mainstream" American English. Silence should not be mistaken for the inability to speak. A key implication of this is that what is left unwritten needs to be included in considerations of literacy practices.

The role that literacy plays in immigrant communities also highlights how adults often adapt to new circumstances. For example, [Kalmar \(2001\)](#) detailed the experiences of a group of Spanish speaking migrant workers who created their own English orthography in the absence of formal English classes. In Nara, Japan, I met with Nepali immigrants who had done the same by using the Napali alphabet to represent the sounds of Japanese. In both cases, immigrant adults were creatively refashioning existing linguistic and textual resources. Studies during the early days of the expansion of the internet identified the ways that immigrant communities

utilized online portals and sites to stay connected their countries of origins (Lam, 2006; Yao, 2009). Here, actual immigrant adults are displaying the type of facility with digital resources that is supposed to be difficult for metaphorical digital immigrants, further highlighting the problematic nature of the trope.

Adult literacy practices also provide insight into the ways that literacy can be central in the creation of new communities based on other common interests, such as hobbies and fandom. For example, although writing and distributing zines is often associated with younger readers and writers, some of the earliest pioneers included women writing *Star Trek* fan fiction in the early 1970s (Bacon-Smith, 1992). The members of this community found each other through distributing the stories they were writing and bonded over their shared love of the show. Current research into zine production has examined how men are exploring new ways to express their gender and racial identities (Guzzetti, 2019) and there is a flowering of zines made by and for the transgender community (e.g., *Trans Affirming Magical Care* – a zine made by transgender Harry Potter fans in response to J.K. Rowling's public statements about transgender politics). Over the decades, zine creators have adapted their methods, moving from mimeograph machines, to copiers and now to the internet. Again, rather than being behind technological change, some communities of adults have modeled how to take full advantage of new affordances.

Non-valorized practices

The non-valorized nature of these practices is consistent with the vulnerability of certain adult populations. Kalman (1999) documented the work of scribes who set up temporary shop on the edges of public plazas in Mexico to earn a living reading and writing letters for those who cannot. Here, literacy skills provide for informal employment and way to manage a difficult situation. Even more precarious are those adults who write messages on paper or cardboard as a means of asking for assistance from passersby. These signs are vulnerable to both the elements and the authorities. For example, police in Alabama (in the United States) made a “quilt” of the signs they had taken from unsheltered people (Burke, 2019). This action demonstrates both a clear lack of empathy on the part of law enforcement and the stigmatized nature of this literacy practice. Their literacy skills may be one of the few resources these adults have, but their written calls for help often become an object of scorn.

By contrast, non-valorized literacy practices of adults who are not in such precarious positions face fewer sanctions. For example, it is a common practice in meetings for adults to surf the web or to send emails and texts that are not work-related. While doing so, these workers may engage in mock participation, and such dissimulations are accepted to be part of white collar working life. In fact, this type of activity is part of a long tradition of workers reclaiming bits of time or resources away from the employer. de Certeau (1984) identifies it as “la perruque,” and notes that while formally discouraged, it is informally expected to take place. Another example of a non-valorized literacy practice that produces fewer sanctions is posting hostile comments on discussion boards. For example, in a study of interactions on sports news message boards, Jacobson (2019) noted the prevalence of racist, misogynistic and homophobic discourse. Although some comments did receive pushback, any stigma or scorn toxic posters faced did not rise to the level of signs being taken from the unsheltered.

An expansive view of adult literacy points to how the literacy practices of adults can be used to amplify voices or silence them, create or disturb community, and recognize or dehumanize each other. In this way, adult literacy serves as a paradigm for the social model of literacy, as literacy can only be understood within complex and dynamic sociocultural and socioeconomic contexts. Being an adult is complicated, and adult literacy practices are no less so. Although adopting this perspective expands our sense of what formal programs can take up (Jacobson et al., 2003), just as importantly it highlights the multiple ways that adults continue to develop their literacy skills and practices in a variety of informal and non-formal contexts (Brandt, 2001; Purcell-Gates, 1995). People, regardless of age, want to be engaged in play, study, activity and work that is relevant and meaningful.

Adult literacy as political praxis

Although there is ample evidence of the negative impact neoliberalism has had on the provision of adult education (and education more generally), it should not be understood as a historical anomaly. For example, in Graff's (1991) analysis of seismic educational shifts in the 1800s, the introduction of mass schooling in Canada and other places was a reaction on the part of the ruling class to the fact that the working class was becoming literate on its own terms. Rather than being some kind of benevolent gesture or measure of social progress, the expansion of formal education can be read as an attempt to direct and control the nature of the working class' literacy activities. The curriculum of the time included frequent invocations of religious themes and a call for people to accept their place in the established social hierarchy. As with the legal prohibition against teaching enslaved people how to read and write in the United States, this domestication of working class literacy was a means to maintain the status quo.

Adult literacy and political struggle

Not surprisingly, states undergoing revolutionary changes have understood increases in adult literacy skills as a key part of their struggle. National literacy campaigns in the Soviet Union (Clark, 1995), Cuba (Herman, 2012) and Viet Nam (Anderson, 1995) explicitly connected literacy education to political action. Smaller groups working toward social justice have also engaged in these types of coordinated efforts. For example, in the United States, enslaved African American risked their lives to teach themselves and others how to read. Nearly a century later, literacy programs were an important part of the civil rights movement (e.g., the

Citizenship Schools founded by Esau Jenkins and Septima Clark in South Carolina). Similar efforts were undertaken in Japan, where adult literacy was an important part of the struggles of both the Buraku and Korean communities against discrimination (Asano, 1990; Mori, 1995). Freire (1985) documents similar projects in Brazil and Guinea Bissau. In each case, adult literacy instruction relied upon a commitment to praxis – the cycle of taking informed action on the world, critically reflecting on the results and developing theory that informs subsequent actions. This approach to literacy shifts focus away from isolated evaluations of skills (e.g., decoding, encoding) or accounts of diversity in cultural practices (e.g., how particular religious communities organize around sacred texts) and toward the impact adults' myriad literacy practices have on their community and the world.

In contrast to atomized accounts of literacy skills, adult literacy understood as praxis is necessarily realized in collective activity (Osawa, 1990). An influential demonstration of this is the work of the Combahee River Collective (Taylor, 2017), who were a non-formal study group active in Boston in the 1970 dedicated to developing a Marxism fully informed by the experiences of African American women. Per Barton and Hamilton (1998), they were using literacy to “get things done in the community,” but in this case, their goal was communal liberation from the interlaced oppressions they faced (e.g., racism, sexism, capitalism). As with other self-directed learning projects (Tough, 2002), the group itself selected the texts for study and generated the questions that guided their analysis. One result of their work was a publicly shared statement that articulated the concepts of identity politics and intersectionality before they became more widely known. The Collective's activities made clear the ideological dimensions of what were described above as African American women's literacy practices – the common concern for protecting their loved ones and shared understandings of the world informed a distinctive vision of change. In current times, adult education projects in Mexico and Central America dedicated to local and regional decolonization efforts demonstrate similar approaches. Rather than embracing decontextualized concepts of literacy or imported discourses of liberation, activists there are drawing on and articulating indigenous epistemologies to make their own way forward (Meyer and Alvarado, 2010). Similar efforts are underway in South America, where groups of workers who share a commitment to horizontalism are experimenting with new forms of non-hierarchical adult education (Sitrin, 2012).

Problematizing technology

The near ubiquity of the internet and social media has created new affordances for adults to take action on the world via literacy. For example, Jung's (2019) studied how a Korean American women's community used an online forum to take action around sexual harassment and assault. A key was their strategic use of posts, petitions and social media hashtags to get their message across. As another example, Palestinian hacktivists (Skare, 2016) have moved their resistance to Israeli occupation online by attacking government and corporate websites. Though they differ in scale and technical complexity, what these efforts share is an origin in informal groups and self-study. Learning how to take political action in cyberspace is not taught in schools. Because digital spaces can create barriers to participation, particularly for those with certain impairments, just getting online can be a profoundly political act. Ellcessor (2016) provides an account of disability activists simultaneously adapting available resources and arguing for wider accessibility as an equity issue. In each of these cases, what has been referred to as digital literacy involves the use of digital resources for fully engaged praxis, rather than the passive consumption of information.

The ubiquity of technology has also been identified as a key feature of an expanded surveillance state, and it creates a need for shifts in praxis. For example, Eubanks (2012) has reported on efforts of women in an adult literacy class to examine how they are being tracked and monitored by the social welfare system in the United States when they use electronic cards to access resources. This became the first step in them developing a “critical technological citizenship” that drew on their own experiences with technology in order to advocate for fair housing, welfare and education policies. Similarly, adult learners who contributed to a technology themed issue of *The Change Agent* (a newspaper dedicated to writings by adult learners focused on social justice issues), balanced praise for devices like cell phones and tablets with concerns about the environmental and personal costs. For example, dos Santos (2013) explained to other adult learners how cookies work as part of Google searches and why users get targeted advertising. dos Santos explains in clear terms, “In a way, you are the product. You— your history, your habits, your interests, etc.— are being sold to companies that want to sell you things” (p. 44). Rather than coming from a formal study of technology, this critique is a result of the same type of popular education programs that Eubanks described.

The concern with how technology aids surveillance extends beyond the internet, as researchers and activists have been tracking how we are being read by public places that we pass through, via security cameras, facial recognition software, GPS on cell phones and other means. McCullough (2013) suggests that, “Citizens have a right to engage one another and the built world they inhabit in ways that are unmediated, uninstructed, unscripted and undocumented” (p. 273). To preserve that right, Brunton and Nissenbaum (2016) describe what they refer to as the strategy of obfuscation, “the deliberate addition of ambiguous, confusing, or misleading information to interfere with surveillance and data collection” (p. 1). More directly, those working in activist communities are careful to use communication platforms that are less liable to be hacked and monitored, and teach each other how to stay as safe as possible. For this reason, so-called digital literacy is not just about getting online and getting connected, it includes being able to manage when and where you want to be connected and monitored.

An addition to expanded means of surveillance, there is a concern that the way the internet and social media have developed has made it more difficult to determine the veracity of what we read, see and hear. Digital literacy is thus linked to information literacy, with the goal of helping users distinguish facts from incorrect reports, propaganda or outright lies. This is true not only of those who are labeled adult literacy learners, but those with college educations and advanced degrees. It is also not a new issue. Purcell-Gates (2007) suggested that in the run up to the second Gulf war, the behavior of many politicians and thought leaders in the United

States suggested a lack of the kind of skills that adults in literacy programs are asked to work on improving (e.g., reading critically, connecting texts, etc.). Although it is true that it can be difficult to navigate a complex mediascape, the poor information literacy of the credentialed class can also be understood to be a form of a gnosis (Moffett, 1988) - an unwillingness to comprehend or to consider ideas that challenge one's own views.

Conclusion

There is ample evidence that adults who have no issues with decoding take on a stance of willful ignorance and refuse to engage with texts that threaten their understanding of the world. In Moffett's case, parents and others in the US in the 1970's were enraged that their children were being given textbooks that had critiques of the United States in them (e.g., reporting on US war crimes abroad and systemic racist violence at home). These fights over the content of textbooks are once again on the rise in the United States, with some parents sharing the same desire to remove mention of anything that runs counter to ideas of American exceptionalism. This is happening at the same time significant numbers of adults deny the reality of the climate crisis that is already upon us and the existence of the COVID pandemic. However, when the PIAAC and other national surveys are used to identify potential problems with adult literacy that create risk for individuals and society, this deep-seated and ideologically motivated desire not to know is not part of the evaluation. Given the seriousness of the multiple crises the global community faces, rather than learners in adult literacy programs, it is those with degrees, credentials and valorized literacy skills that are *the* risk, and they are putting the rest of the world *at* risk.

When conceived of as a problem with skills or as paradigmatic of social models of literacy, *adult literacy* functions as an odd type of marked case. Indeed, the majority of literacy activity carried out across the world is conducted by adults, and at any given time, adults are engaged in a nearly countless number of activities. This includes writing poetry, plays and novels, presenting proposals for research or business, reporting on the news and drafting laws and policies. These activities typically do not factor in discussions of adults' reading and writing.

Instead, the focus is on what are thought of as commonplace practices, such as the writing of shopping lists, correspondence or reading the labels of medicine bottles. This narrow view of adult literacy means that the term is only applied to certain populations, typically those who need to work on their skills or those who are looking to make it through another day. If attention to adult literacy continues to be misunderstood as a means of propping up the current economic system, we are likely to perpetuate existing inequalities. However, if we understand adult literacy as the praxis of all adults, and not just those in programs, we might be able to make the profound changes we need to survive as a species.

References

- Allatt, G., Tett, L., 2019. The employability skills discourse and literacy practitioners. In: Tett, L., Hamilton, M. (Eds.), *Resisting Neoliberalism in Education: Local, National and Transnational Perspectives*. Policy Press, Bristol, pp. 41–54.
- Anderson, B., 1995. *Imagined Communities*. Verso, London.
- Asano, T., 1990. Buraku literacy movement. In: Buraku Liberation Research Institute (Ed.), *The Literacy Work and Discrimination in Japan*. Kaihou Shuppan Sha, Osaka, pp. 2–8.
- Bacon-Smith, C., 1992. *Enterprising Women: Television Fandom and the Creation of Popular Myth*. University of Pennsylvania Press, Philadelphia, PA.
- Baker, D.W., Gazmararian, J.A., Williams, M.V., et al., 2002. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am. J. Publ. Health* 92, 1278–1283.
- Ball, A., 1992. Cultural preference and the expository writing of African-American adolescents. *Writ. Commun.* 9 (4), 501–532.
- Barton, D., Hamilton, M., 1998. *Local Literacies*. Routledge, New York, NY.
- Baugh, J., Smitherman, G., 2007. Linguistic emancipation in global perspective. In: Alim, H.S., Baugh, J. (Eds.), *Talkin' Black Talk: Language, Education, and Social Change*. Teachers College Press, New York, NY, pp. 115–131.
- Baquet, C.R., Horn, J.W., Gibbs, T., Greenwald, P., 1991. Socioeconomic factors and cancer incidence among Blacks and Whites. *JNCI (J. Natl. Cancer Inst.)* 83 (8), 551–557.
- Brine, J., 2006. Lifelong learning and the knowledge economy: those that know and those that don't—the discourse of the European Union. *Br. Educ. Res. J.* 32 (5), 649–665.
- Brandt, D., 2001. *Literacy in American Lives*. Cambridge University Press, Cambridge.
- Brunton, F., Nissenbaum, H., 2016. *Obsfucation*. MIT Press, Cambridge, MA.
- Burke, M., 2019. Alabama police officers pose with "homeless quilt" made from panhandling signs. NBC News. Retrieved on June 21, 2022 from: <https://www.nbcnews.com/news/us-news/alabama-police-officers-pose-homeless-quilt-made-cardboard-signs-n1108776>.
- Clark, C., 1995. Literacy and labour: the Russian literacy campaign within the trade unions, 1923–1927. *Eur. Asia Stud.* 47 (8), 1327–1341.
- De Certeau, M., 1984. *The Practice of Everyday Life*. Steven F. Rendall (Trans.). University of California Press, Berkeley, CA.
- Dewalt, D.A., Berkman, N.D., Sheridan, S., Lohr, K.N., Pignone, M.P., 2004. Literacy and health outcomes: a systematic review of the literature. *J. Gen. Intern. Med.* 19 (12), 1228–1239.
- dos Santos, R., 2013. You are the product! *Change Agent* 37 (1), 44.
- Elcissor, E., 2016. *Restricted Access: Media, Disability and the Politics of Participation*. New York University Press, New York, NY.
- Eubanks, V., 2012. *Digital Dead End: Fighting for Social Justice in the Information Age*. MIT Press, Cambridge, MA.
- European Association for the Education of Adults, 2014. *PIAAC-OECD Survey of Adult Skills: A Wake-Up Call for Europe*. Author, Brussels.
- Financial Literacy and Education Commission, 2016. *Promoting Financial Success in the United States: National Strategy for Financial Literacy, 2016 Update*. Author, Washington, DC.
- Franklin, W.J., Dowdy, J.K., 2005. Storytelling. In: Dowdy, J.K. (Ed.), *The Readers of the Quilt*. Hampton Press, Cresskill, NJ, pp. 119–136.
- Freire, P., 1985. *The Politics of Education* (D. Macedo, Trans.). Bergin and Garvey, New York, NY.
- Gilpin, T., 2014. Skills gap a convenient myth. *Labor Notes*. Retrieved June 22, 2022 from: <http://www.labornotes.org/2014/01/skills-gap-convenient-myth>.

- Goldstein, A., 2012. Rare agreement: Obama, Romney, Ryan all endorse retraining for jobless—but are they right? Retrieved on June 22, 2022 from: <http://www.propublica.org/article/rare-agreement-obama-romney-ryan-endorse-retraining-for-jobless-but-are-the>.
- Goodman, M., Sands, A., Coley, R., 2015. America's Skills Challenge: Millennials and the Future. Educational Testing Service, Princeton, NJ.
- Graff, H., 1991. *The Literacy Myth*. Transaction, New Brunswick, NJ.
- Guzzetti, B., 2019. Diverse men making media: creating cultural (re)constructions of gender and race. In: Guzzetti, B., Bean, T., Dunkerly-Bean, J. (Eds.), *Literacies, Sexualities and Gender*. Routledge, New York, NY, pp. 193–207.
- Hamilton, M., Pitt, K., 2011. Challenging representations: constructing the adult literacy learner over 30 years of policy and practice in the United Kingdom. *Read. Res. Q.* 46 (4), 350–373.
- Heath, S.B., 1983. *Ways with Words*. Cambridge University Press, Cambridge.
- Herman, R., 2012. An army of educators: gender, revolution and the Cuban literacy campaign of 1961. *Gend. Hist.* 24 (1), 93–111.
- Herrera, S., Truckenmiller, A., Foorman, B., 2016. Summary of 20 Years of Research on the Effectiveness of Adolescent Literacy Programs and Practices. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast, Washington, DC.
- Ito, M., Baumer, S., Bittanti, M., Boyd, D., Cody, R., Stephenson, B.H., Horst, H.A., Lange, P.G., Mahendran, D., Martinez Pascoe, C.J., Perkel, D., Robinson, L., Sims, C., Tripp, L., 2019. *Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning with New Media*. MIT Press, Cambridge, MA.
- Jacobson, E., Degener, S., Purcell-Gates, V., 2003. *Creating Authentic Literacy Materials for the Adult Education Classroom*. World Education, Boston, MA.
- Jacobson, E., 2017. The workforce investment and opportunity act: new policy developments and persistent issues. In: Belzer, A. (Ed.), *New Directions for Adult and Continuing Education*, 155. Wiley Periodicals, Hoboken, NJ, pp. 19–27.
- Jacobson, E., 2019. Performing and resisting toxic masculinities on sports news comment boards. In: Guzzetti, B., Bean, T., Dunkerly-Bean, J. (Eds.), *Literacies, Sexualities and Gender*. Routledge, New York, NY, pp. 169–181.
- Jung, J.K., 2019. Transnational women's online literacies. In: Guzzetti, B., Bean, T., Dunkerly-Bean, J. (Eds.), *Literacies, Sexualities and Gender*. Routledge, New York, NY, pp. 182–192.
- Kalman, J., 1999. *Writing on the Plaza*. Hampton Press, Cresskill, NJ.
- Kalmar, T.M., 2001. *Illegal Alphabets and Adult Bilingualism: Latino Migrants Crossing the Linguistic Border*. Lawrence Erlbaum, Mahwah, NJ.
- Knobel, M., 1999. *Everyday Literacies: Students, Discourse and Social Practice*. Peter Lang, New York, NY.
- Lam, E., 2006. Re-envisioning language, literacy, and the immigrant subject in new mediascapes. *Pedagogies: Int. J.* 1 (3), 171–195.
- Lanehart, S., 2002. *Sista, Speak! Black Women Kinfolk Talk About Language and Literacy*. University of Texas Press, Austin, TX.
- Lee, C., 2007. *Culture, Literacy, and Learning: Taking Bloom in the Midst of a Whirlwind*. Teachers College Press, New York, NY.
- Levine, M., 2013. *The Skills Gap and Unemployment in Wisconsin: Separating Fact from Fiction*. University of Wisconsin-Milwaukee, Center for Economic Development, Milwaukee, WI.
- Lusardi, A., 2019. Financial literacy and the need for financial education: evidence and implications. *Swiss Journal of Economics and Statistics* 155 (1), 1–8.
- McCullough, M., 2013. *Ambient Commons: Attention in the Age of Embodied Information*. MIT Press, Cambridge, MA.
- Meyer, L., Alvarado, M. (Eds.), 2010. *New World of Indigenous Resistance*. City Lights Books, San Francisco, CA.
- Moffett, J., 1988. *Storm in the Mountains: A Case Study of Censorship, Conflict, and Consciousness*. Southern Illinois University Press, Carbondale, IL.
- Mori, M., 1995. Educational activities in the community. In: Hirasawa, Y., Nabeshima, Y. (Eds.), *Dowa Education: Educational Challenge Toward a Discrimination Free Japan*. Kaihou Shuppan Sha, Osaka, pp. 51–57.
- National Commission on Adult Literacy, 2008. *Reach Higher America: Overcoming the Crises in the US Workforce*. Author, Washington, DC.
- National Skills Coalition, 2020. *Applying a Racial Equity Lens to Digital Literacy*. Author, Washington, DC.
- Organization for Economic Co-operation and Development, 2021. *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies*. OECD Publishing, Paris.
- Osawa, T., 1990. The Kotobuku literacy school. In: Buraku Liberation Research Institute (Ed.), *The Literacy Work and Discrimination in Japan*. Kaihou Shuppan Sha, Osaka, pp. 18–20.
- Pellizzari, M., Fichen, A., 2013. A New Measure of Skills Mismatch: Theory and Evidence from the Survey of Adult Skills (PIAAC). OECD Social, Employment and Migration Working Papers, No. 153. OECD Publishing, Paris.
- Prensky, M., 2001. Digital natives, digital immigrants. *Horizon* 9 (5), 1–6.
- Purcell-Gates, V., 1995. *Other People's Words: The Cycle of Low Literacy*. Harvard University Press, Cambridge, MA.
- Purcell-Gates, V., 2007. What's it all about? Literacy research and civic responsibility. In: Rowe, D., Jiménez, R., Compton, D., Dickenson, D., Kim, Y., Leander, K., Risko, V. (Eds.), *The 56th Yearbook of the National Reading Conference*. National Reading Conference, Oak Creek, WI, pp. 1–13.
- Redd, T., Webb, K., 2005. *African American English: What a Writing Teacher Should Know*. National Council of Teachers of English, Urbana, IL.
- Richardson, E., 2003. *African American Literacies*. Routledge, New York, NY.
- Rybarczyk, B., Billig, A., 1997. *Listening to Life Stories: A New Approach to Stress Intervention in Health Care*. Springer, New York, NY.
- Santos, M., Paasche-Orlow, M., 2019. Special supplement: health literacy and adult basic education. *HLP: Health Literacy Research and Practice* 3, s88–s90.
- Scott, T.L., Gazmararian, J.A., Williams, M.V., Baker, D.W., 2002. Health literacy and preventive health care use among Medicare enrollees in a managed care organization. *Med. Care* 40, 395–404.
- Schwerdt, G., Wiederhold, S., Murray, T.S., 2020. *Literacy and Growth: New Evidence from PIAAC*. National Center for Education Statistics (NCES) of the U.S. Department of Education, Washington, DC.
- Scribner, S., Cole, M., 1981. *The Psychology of Literacy*. Harvard University Press, Cambridge, MA.
- Selwyn, N., Gorard, S., Furlong, J., 2006. *Adult Learning in the Digital Age*. Routledge, New York, NY.
- Sitrin, M., 2012. *Everyday Revolutions: Horizontalism and Autonomy in Argentina*. Zed Books, New York, NY.
- Skare, E., 2016. *Digital Jihad: Palestinian Resistance in the Digital Era*. Zed Books, London.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Taylor, K.Y., 2017. *How We Get Free: Black Feminism and the Combahee River Collective*. Haymarket Books, Chicago, IL.
- Thériault, V., 2019. Accountability literacies and conflictual cooperation in community-based organisations for young people in Québec. In: Tett, L., Hamilton, M. (Eds.), *Resisting Neoliberalism in Education: Local, National and Transnational Perspectives*. Policy Press, Bristol, pp. 13–26.
- Tobias-Machado, M., Cooperative Brazilian Uro-Oncology Group (CBUG), 2013. Association between literacy, compliance with prostate cancer screening, and cancer aggressiveness: results from a Brazilian screening study. *Int. Braz. J. Urol.* 39 (3), 328–334.
- Tough, A., 2002. *The Iceberg of Adult Informal Learning* (Working Paper No 49). The Research Network for New Approaches to Adult Lifelong Learning (NALL), Toronto, ON.
- Trawick, A., 2019. The PIAAC literacy framework and adult reading instruction. *Adult Literacy Education* 1 (1), 37–52.
- White, S., 2011. *Understanding Adult Functional Literacy*. Routledge, New York, NY.
- Yao, A., 2009. Enriching the migrant experience: blogging motivation, privacy, and offline lives of Filipino women in Britain. *Clin. Hemorheol. and Microcirc.* 14 (3).
- Yasukawa, K., Osmond, P., 2019. Adult basic education in Australia: in need of a new song sheet? In: Tett, L., Hamilton, M. (Eds.), *Resisting Neoliberalism in Education: Local, National and Transnational Perspectives*. Policy Press, Bristol, pp. 195–207.
- Yoon, H., Jang, Y., Vaughan, P.W., Garcia, M., 2020. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. *J. Appl. Gerontol.* 39 (1), 105–110.

College reading

Norman A. Stahl^a, Sonya L. Armstrong^b, and James R. King^c, ^a Department of Curriculum and Instruction, Northern Illinois University, DeKalb, IL, United States; ^b Department of Curriculum and Instruction, Texas State University, San Marcos, TX, United States; and ^c Literacy Studies, University of South Florida, Tampa, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

History	463
Curricular structures, past and present	465
Instructional foci	466
Assessment	466
Professional associations	466
Theoretical assumptions	467
Perspectives focused on learner attributes	468
Lifespan literacy	468
Funds of knowledge	468
Perspectives focused on language/text valuing	468
Academic literacy, disciplinary literacies, and professional literacies	468
Vernaculars	468
Multimodality and new literacy studies	469
Perspectives focused on the expectations and purposes of the college context	469
Contextualization	469
Entextualization	470
Positioning theory	470
Conclusions and future directions	470
References	471

The term “college reading,” as used in this entry, is being applied on two parallel levels. It is first meant to describe a professional field of researchers, theoreticians, and practitioners whose work focuses on, generally, learning with and through text at the postsecondary education level. In a second sense, college reading is a broad descriptor for a range of constructs focused on the academic literacy expectations and demands required for a learner’s successful journey through stages of postsecondary/tertiary experience. The culture of reading at a particular institution of higher education, whether a liberal arts college, a community college, a technical college, or an elite university, will determine whether or not reading is taught, and, if so, how: explicitly within disciplinary offerings, through pre-credit remedial or developmental courses, or as non-course-based options in learning assistance centers or tutorial programs. On the other hand, considerations for college reading also determine whether a learner’s competencies in reading are expected to grow in both breadth and depth from the learner’s immersion in a rich academic content and under the guidance of a highly knowledgeable faculty.

In the nascent 21st century, the term “college reading” (CR) requires purposeful framing, as it can simultaneously be used to represent the pervasive culture of literacies throughout an institution as well as specific curricular offerings charged with explicitly training students to be effective and efficient learners. Over time, CR has been used to describe instruction that is remedial and also instruction that is developmental in the sense that every student, regardless of level or status within the institution, benefits. And, CR has and continues to refer to both a narrow, skills-based focus on reading competency as well as a broader focus on higher-order, strategic learning processes.

Our discussion of college reading must begin with an acknowledgment of the differences in tertiary and postsecondary education across international contexts. Because other nations either exclude any reading instruction in higher education or include elements of reading embedded in informal educational contexts such as adult and community literacy, and not necessarily in a model of developmental education in formal postsecondary contexts (as in the US), much of the historical discussion that follows is US-centric.

History

The study of reading in postsecondary institutions of the US begins with the founding of Harvard College in 1636 (Stahl and King, 2018; Wyatt, 1992). Then, continuing for more than two centuries, subject matter in colleges across the country concentrated intensely upon a limited number of fields such as philosophy, theology, and classical studies. Books had great monetary value and though not widely distributed were found in libraries for use by faculty and senior-level students. Students might also belong to literary societies that maintained libraries of academic books, though often of a more secular nature. Admission to college was seen as proof that students could read (Leedy, 1958).

When land grant colleges and the state colleges/universities began appearing beyond the eastern shores, these institutions regularly found that incoming students lacked sufficient competency in the basic skills. This awareness led to schools such as the University of Wisconsin establishing preparatory departments (preparatory programs at essentially high school levels), where admitted students would be prepared to handle the rigors of the college-level courses offered in the institutions (Arendale, 2010).

Subsequently, the emergence of US universities modeled after institutions such as Germany's universities, including Johns Hopkins in the latter 1800s, shifted the mission of higher education to a focus on research, physical and life science, social science, and English literature. In these newer institutions, libraries became more active in their agency for scholarship. With a greater breadth in the curriculum, the role of reading and literacy competence levels required to read across multiple fields now took on new meaning and importance. The rigid and narrow curriculum of the past, with reliance on recitation, eventually gave way to survey courses, lectures, extensive reading with new texts, collateral reading, and class tests.

Several formative research studies, curricular innovations, and textbook publications across the latter 1890s and first decades of the 20th century gave rise to the nascent field of college reading:

- Abell's (1894) first research study with college readers,
- Hinsdale's (1900) *The Art of Study*,
- McMurry's (1909) *How to Study and Teaching How to Study*,
- Koopman's (1913) qualitatively oriented study of students' reading behaviors,
- Moore's (1915) implementation of a "how to study" course at Harvard, and
- Kerfoot's (1916) authorship of *How to Read: Techniques of Reading Improvement*.

Furthermore, studies by early psychologists on reading habits, skimming, speed, comprehension, work study skills, and eye movement research, as well as the appearance of standardized tests led to an understanding that not all college students were reading-ready for college demands, at least as measured by the new learning science embodied in these new tests. Indeed, Leedy (1958) has productively designated the period from 1910 to 1920 as the *formative years* for the field. Courses had a heavy emphasis on how to study with collateral work in reading.

By the latter 1920s, interest in the "reading problem" with 1st-year college students had reached a new level. A new student clientele was coming to college. Orientation courses (Strang, 1937; Traxler, 1938) that often included some coverage of work study skills emerged. Book (1927a,b) determined that students experiencing academic difficulties at the University of Indiana were deficient in reading study skills and that they would benefit from taking a "how to study" course. Parr (1930) conducted the first national survey of reading and study courses, and while the sample was limited (as was response), information became available on the operation of programs across the nation. In the same year Eurich (1929) and Haggerty and Eurich (1929) released the *Minnesota Speed of Reading Test* and a *Test of Reading Comprehension* for the college population. What was likely the first summer bridge program focusing on work study skills was delivered at the University of Buffalo.

With the 1930s students were expected to read effectively and efficiently with an ever-increasing volume of texts. Assessment of students' reading abilities was now often undertaken at their time of admission. The nation's center for research and praxis was The Ohio State University (OSU) where, under the leadership of Luella Cole Pressey and Sidney Pressey, one of the first contextualized study courses was offered (see Stahl and Armstrong, 2020). As programs sprung up throughout the country so did early use of technology designed to measure and/or promote students' efficient eye movements in reading (e.g., the Harvard and the Iowa reading films, the metronoscope, the ophthalmograph, and the tachistoscope). It became a period during which the field of college reading received a certain degree of acceptance and even stature, in part, because of the growth in research and the knowledge that not all students were predicted to be able to handle the reading load (Leedy, 1958).

The second world war had direct influence upon college reading instruction, demonstrated by the STAR Program at OSU. Robinson (1943) found that all students, even the best and brightest, could benefit from study skills instruction. During that same decade, two early attempts integrated reading and learning techniques in content courses that took place in the field (City College of New York in educational psychology, and College of St. Scholastica in English). College reading pioneer James M. McCallister (1942) authored *Purposeful Reading in College* (1942), a model for the texts to come in the 1980s. The first part of each of McCallister's chapters presented instruction on reading and study skills, while in the second section students would undertake comprehension tasks and rate improvement with paragraphs from contemporary college content area texts. Frances Oralind Triggs (1943) released *Remedial Reading: The Diagnosis and Correction of Reading Difficulties at the College Level*, which was the first book to train reading/learning specialists for the postsecondary field. In 1946, Robinson released *Effective Study*, with its focus on research-driven, higher-level study skills such as SurveyQ3R. As such, *Effective Study* became the model for books across the next three decades. The close of the 1940s decade hosted debates on speed versus comprehension, mechanized reading instruction versus more natural approaches, and controlled versus uncontrolled instruction.

The design of programming for the 1950s was greatly defined by the Servicemen's Readjustment Act of 1944 (i.e., the GI Bill of Rights), as thousands of veterans entered or reentered postsecondary institutions. Reading load expectations for college students were greater than anytime previous in the history of higher education. Yet, colleges were becoming increasingly aware that numerous students were not prepared to handle the rigors of such reading and learning requirements. Surveys from the previous decade and those conducted during the 1950s demonstrated that support programs differed greatly in mission, institutional support, and in their designs. Faculty charged with the teaching of reading and study skills from across the west convened to promote research and discuss praxis leading to the founding of the Southwest Reading Conference (now the Literacy Research Association) in 1952. Later on, and with similar rationale, the College Reading Association (now the Association for Literacy Educators

and Researchers) was formed in 1958. Federal legislation (National Defense Act of 1958) further expanded the population of college-going individuals. At Harvard University, William Perry found that entering students could not write a short sentence about the content of a chapter from a college-level textbook, nor could many students identify a well-written essay response. Perry's (1959) subsequent report to the Harvard faculty was used by programs across the nation as a rationale for widely implementing developmental reading programs. Finally, in 1958, Leedy authored the foundational work *A History of the Origin and Development of Instruction in Reading Improvement at the College Level*.

In the 1960s, a time of cultural deconstruction of hierarchy, numerous colleges adopted open-admissions programs. Congress passed legislation titled the Higher Education Act of 1965 with grants for undergraduates. From the earlier decade, the pervasive influence of the *Evelyn Wood Reading Dynamics* (Maxwell, 1979; Wood, 1960) in the popular culture created an emphasis on speed reading within college reading programs, also making use of the EDL Controlled Reader, the Craig Reader, the Perceptoscope, the Tachomatic 500, and the Hoffman Readers (see Berger and Peebles, 1976). Furthermore, the learning theories of behaviorism, exemplified in Skinner's (1957) *Verbal Behavior*, led to increased use of programmed reading skill builders in CR. In fact, Lowe (1968) discovered that on some campuses there could be multiple programs offering college reading and study skills training.

In the 1970s the fruition of student success programming arrived with the birth of the learning skills centers across the country. Many of the leaders of this movement possessed a qualifying academic background in the reading/learning field. While Martha Maxwell (1966) disseminated a plan for a campus-wide program, Frank Christ (1971) received the credit for the first learning assistance center in the nation at Long Beach State University. Programs were resourced by institutions providing traditional college reading programs and counseling centers that provided students with diagnostic assessment, computer-assisted instruction, tutoring, and mini-courses associated with reading and study skills development.

With the full impact of the cognitive revolution in the 1980s and on into the 1990s, postsecondary researchers in the fields of education, cognitive psychology, linguistics, and university developmental education programs began to focus on interactive, contextual, and generative theories as related to students' mastery of strategic reading and learning processes, in preference to the simple description or effectiveness of a reading and/or study skills course. Impactful research and related training programs were centered at the University of Georgia and Georgia State University. Textbooks such as McWhorter (1980) and Nist and Diehl (1985) employed simulation models that embedded reading and learning strategies, contextualizing instruction in stand-alone, precredit reading courses, with somewhat standardized developmental education programs. Bartholomae and Petrosky (1986) released a groundbreaking text for reading/writing-centered approaches that described the theory and praxis associated with the Basic Reading/Basic Writing course from the University of Pittsburgh, which continues to impact the delivery of current integrated reading and writing (IRW) classes (Armstrong et al., 2018).

Martha Maxwell encapsulated the status and positioning of the field at the beginning of the new millennium in the first edition of *The Handbook of College Reading and Study Strategy Research* (Maxwell, 2000). She described the continuing debate on reading and learning strategies required for success in college, as well as where and how the relevant instruction should be delivered. Indeed, more widely, there was also disagreement on the very definition and the theoretical foundations of reading, particularly whether reading was a process, a product, or a discipline, or a mix of all three? After experiencing a decade when numerous programs were removed from 4-year institutions, Maxwell questioned whether CR coursework had actually been successful. She was one of the first professionals to point out that mandated gatekeeping courses associated with the developmental education practices of pseudo-diagnosis (testing), placement, and remediation led to stigmatizing students, increasing drop-out rates, lowering of self-confidence, and delaying progress toward the degree objective. She further highlighted that under-trained or mis-trained instructors continued to plague the CR field. On the more positive side and with prescience regarding the coming of disciplinary literacies, Maxwell pointed out that all students, not just those placed into developmental education, required assistance in mastering the texts used in content fields, where text expectations frequently exceed learners' functional competence. Finally, it was proposed that more viable alternatives to traditional courses might include adjunct and paired classes, supplemental instruction, tutoring services, and emerging technological supports that might promote student learning competencies that had emerged across the previous two decades.

Curricular structures, past and present

Depending on definitions for CR, as well as the overall theoretical and pedagogical model adopted, a number of curricular approaches may be relevant. Historically, in the US, CR has often been enacted within traditional remedial and developmental education courses. Many course models have existed over the years, including stand-alone reading courses, strings of stacked prerequisite courses, reading courses paired with content area courses, integrated reading and writing courses, among others. In the past, such courses were intimately connected with developmental education and academic assistance across higher education. However, over the past four decades, state-level mandates in the US have all but removed such interventions—including centralized developmental education units—at 4-year universities, deferring such courses to community colleges (e.g., Ohio Revised Code, 2017; Tennessee SB 7006, 2010).

In line with these moves to relegate reading instruction to 2-year institutions, much criticism has been aimed at stacked, stand-alone reading and learning courses at the postsecondary level in the US (Bailey et al., 2010; Edgecombe et al., 2013; Rutschow et al., 2019). Therefore, as part of a much larger reform of community college education, significant curricular and structural changes for CR are underway. Most significantly, individual reading and learning courses are being minimized or even eradicated, yielding to

preferences for curricular structures that integrate or blend reading with other content. Specifically, integrated reading and writing (IRW), a course model originally designed at the University of Pittsburgh (Bartholomae and Petrosky, 1986), has been revisited as a way to accelerate students through developmental education courses. Similarly, contextualized reading and learning within content courses or career technical education courses is a popular reform, with wide-scale models such as the State of Washington's I-BEST leading the way (Jenkins et al., 2009; Wachen et al., 2011; Zeidenberg et al., 2010). Another new structure is corequisite courses (Daugherty et al., 2018) that embed a support training (often an IRW course) within a college-level course, such as required 1st-year composition.

Historically, much of the work of CR in the US has tended to be course-based. However, other non-course-based curricular structures have also been utilized. These have included reading labs, workshops hosted by learning assistance centers, and students' individual work with reading specialists in tutoring centers. Other such specialized programs that include some aspect of reading and learning instruction include bridge programs, college transition courses, and adjunct courses (Johnson and Carpenter, 2000).

Beyond the US, course-based interventions for reading are less prevalent in tertiary educational settings as these types of interventions often exist only in situations where students have not passed admissions exams. For instance, in Sahito et al. (2017) investigation of students in a university in Pakistan, a newly created remedial English course was explored from the perspective of both students and teachers.

Instructional foci

Regardless of the particular curricular structure, much reading and learning instruction at the postsecondary level tends to focus on issues related to reading comprehension. Although the theoretical model applied will undoubtedly shape instructional focus, specific instructional foci are related to the course type. The traditional remedial reading course foci tend to emphasize more discrete skills, such as identification of a text's main idea and supporting details, vocabulary development, procedural aspects of reading strategies, and use of context clues based on the belief that there was a necessary "developmental" ladder of skills for reading competence (Stahl and Armstrong, 2018). More prevalent, however, are developmental and transitional reading approaches that tend to focus on strategic reading with an emphasis on the cognitive, metacognitive, and affective correlates that impact reading comprehension (Holschuh and Lampi, 2018; Holschuh and Paulson, 2013; Simpson and Nist, 2002). Specific instructional foci in these courses include strategic vocabulary building, context-dependent reading and learning strategies, metacognitive strategies, flexibility and self-regulation, and an emphasis on reading across contexts (disciplinary literacy). With the developmental education model found in most community colleges, identified students were not allowed to register in general education courses before successfully completing any requirements association. Hence, courses were generally taught via a simulation model that lacked direct contextualization and had questionable transfer value to courses in which the students would enroll in future terms (Stahl et al., 2020). In the second decade of the 21st century this model has been found to be on the decline. Because the majority of published literature on CR tends to be practical in its focus, a wealth of explanation exists on both general and specific instructional practice (e.g., Armstrong and Newman, 2011; Nist and Simpson, 2000; Simpson et al., 2004).

Assessment

Generally speaking, to reference assessment as it relates to college reading and learning means to reference placement testing, or an initial screening of a student's readiness for college-level reading (Flippo et al., 2018). Historically, such placement tests have helped determine whether a student places directly into college-level courses or is instead first referred to a developmental reading course. As one of the key contributors to the present critique of stand-alone developmental reading courses, single, cognitively oriented placement tests that determine students' placement into or out of developmental reading have largely become the focus of conversations related to assessment. For many years, scholars of college reading and learning have called for multiple placement measures that incorporate a mix of cognitive and affective instruments (e.g., Bliss, 2001; Maxwell, 1997; Roueche and Kirk, 1973). These proposed revisions also include more holistic assessment of learners' reading experiences or preparedness (i.e., secondary school overall GPA, and grades in target courses), calls that have been largely ignored until the most recent era, much to the detriment of the field of CR.

A less-common type of assessment used within college reading is a more fine-grained diagnosis. Diagnostic assessment in reading at the college level typically occurs within disability resource centers, tutoring centers, or reading labs, and tends to be disconnected from the student's coursework (Flippo et al., 2018). A less-common type of assessment is program evaluation, which can provide information on student progress. Such data can inform instructional changes or provide information on the overall success of a program (Armstrong and Williams, 2019; Simpson, 2002).

Professional associations

Although there are a number of professional associations that focus on the larger field of literacy, such is not currently the case for postsecondary reading and learning instruction. Yet, that has not always been so. In 1953, the Southwest Reading Conference for Colleges and Universities (SWRC) was organized, and later, in 1958, the College Reading Association (CRA) was established. Both

groups had a designated focus on CR. Then in 1966, the Western College Reading Association (WCRA) was founded. After decades, the SWRC eventually became the Literacy Research Association (<https://literacyresearchassociation.org>); CRA became the Association of Literacy Educators and Researchers (<https://www.aleronline.org/default.aspx>); and WCRA became the College Reading and Learning Association (<https://www.crla.net>). Although each association has greatly expanded its respective mission and audience, postsecondary literacy is still a welcome topic for conference presentations and for journal articles. The one rather small organization that specifically focuses on college reading and learning was found within the International Literacy Association as a special interest group known as College Literacy and Learning. Unfortunately, professional literacy associations tend toward PK-12 orientations with the effect of marginalizing the field of college reading and learning within the greater literacy field.

Beyond literacy-oriented associations, college reading and learning professionals regularly participate in associations including the Association of Colleges for Tutoring and Learning Assistance (<http://actla.info/contact-us/>), the Association for the Coaching & Tutoring Profession (<https://www.myactp.com>), the National College Learning Center Association (<https://nclca.wildapricot.org>), and the National Organization for Student Success (<https://thenoss.org>).

Beyond the US, college reading and learning specialists may belong to the Federation of European Literacy Associations (<https://www.literacyeurope.org>), or to the member associations of the International Consortium of Academic Language and Learning Developers (<https://icalld.wordpress.com>), the Association for Academic Language and Learning (Australia) (<https://www.aall.org.au>), the Learning Specialists Association of Canada (<http://lsac.wildapricot.org>), the Association for Learning Development in Higher Education (United Kingdom) (<https://aldinhe.ac.uk>), the Association of Tertiary Learning Advisors of Aotearoa (New Zealand) (<http://atlaanz.org>), and the Scottish Higher Education Learning Developers (<https://www.scotelas.org/about-us>), among others.

Theoretical assumptions

As with reading and literacy learning at any stage or in any context, there are many theories that intersect within reading and literacy at the college level. However, a thorough review of foundational texts pertaining to theories and models of reading fails to identify a model specific to the college context. There are, of course, models that touch upon or partially address the college context. For one, Holmes' (1948) substrata factor theory initially grew out of his dissertation with college students but evolved into a more generalized model across time. Also, Chall's (1983) stages of reading development model included a stage 5: construction and reconstruction that focused on college and adult readers (also see Heron, 1989) although it was more directly oriented to the earlier stages of reading mastery. At this point, although many models may have implications for reading at the college level, none identified takes into account the particulars of the college context, culture, forms of language, or student or faculty philosophies of learning. Such a model is needed to serve as a touchstone for researchers to evaluate research or for practitioners to evaluate praxis. Hence, a proposal for an evolving model for college reading follows.

First and foremost, any theoretical model of college reading must proceed from an understanding of the wide range of learner backgrounds within literacy instruction at the postsecondary level. Second, this understanding of learner diversity cannot be divorced from a deep awareness of the demands of the educational context. More specifically, an understanding of the potential difficulties caused by a mismatch of primary and secondary Discourses (Gee, 2001) are important aspects of learning and literacy theory in this context. These theories must necessarily include a focus on the restrictions and limitations of system-bound languages, and not fetishize assumed learner deficits. And, third, any theoretical perspective that positions reading as merely and exclusively a set of decontextualized skills should be understood as ultimately being theoretically unsupported and pedagogically ineffective. Rather, reading at the college level is reading to learn and reading to do, both in the immediate learning context and for the future.

En route to a future, more complete model, we propose eight complementary and often overlapping theoretical accounts that may provide the framing for explaining, evaluating, and researching reading in the college context, and might also further determine the field of college reading. These eight accounts are as follows:

- Lifespan Literacy (e.g., Alexander, 2006; Kerfoot, 1916)
- Funds of Knowledge (González et al., 2006)
- Academic, Disciplinary, and Professional Literacies (Shanahan and Shanahan, 2008)
- Vernacular Literacies (Ivanic et al., 2009)
- New Literacies—Multimodal Pedagogy (Jewitt, 2014; Knobel and Lankshear, 2007; Street, 1995)
- Contextualization (Perin, 2011)
- Entextualization (Baumann and Briggs, 1990; Leppänen et al., 2013)
- Positioning Theory (Harré and van Langenhove, 1999; McVee et al., 2019)

We acknowledge that these theoretical constructs often share concepts and content used in building theories. For the sake of clarity, these constructs can be placed into three general categories: those that focus on learner attributes (Lifespan Literacy and Funds of Knowledge), those that focus on the language and text practices valued and privileged in college contexts (Academic, Disciplinary, and Professional Literacies; Vernaculars; and New Literacies), and those that focus on the specific expectations and purposes of the college context (Contextualization, Entextualization, and Positioning Theory).

Perspectives focused on learner attributes

Lifespan literacy

We are all, always, learning to read. Alexander's (2006) position is that we develop as readers and learners across our lifetimes (see also Kerfoot, 1916). Consequently, learners benefit from focused and explicit literacy support across their lives, not just in the primary grades, as so often assumed. Beginning postsecondary students are faced with a number of developmental and personal transitions (relational, social, cultural, and academic), not the least significant of which is a literacy transition to the academic literacy practices and expectations of higher education and ultimately to a particular discipline. A lifespan literacy perspective is therefore an essential component of any model of college reading as it recognizes that reading instruction cannot merely be relegated to the earlier grade levels.

Funds of knowledge

Funds of knowledge (González et al., 2006) are construed as the cultural capital learners bring to the literacy events of formal educational contexts. Students' funds of knowledge can be described as academic and personal background knowledge, accumulated life experiences, skills, knowledge used to navigate everyday social contexts, and world views structured by broader cultural, historical, and political social forces (Kiyama and Rios-Aguilar, 2018). Because students all have prior experiences, they also have variable examples, data, stories, and ways of understanding the world that influence their conceptual understandings, or schema, about the world. Working from a viewpoint of Whorf's cultural and linguistic relativism (Carroll, 1956), certain cultural schemata are valued more highly than others; thus, power to determine the value of knowledge and knowledge systems resides with those in charge.

However, when instructors are aware of cultural differences that constitute different ways of knowing, they can benefit from engaging with the resulting diversity that is present in every classroom. Likewise, students who perceive that their instructors value what they each uniquely know, self-esteem, risk-taking, and achievement can be enhanced. But these alternative ways of seeing, knowing, and doing knowledge acquisition must be valued multiply, in their many forms, habits, and vocabulary for knowing, not just a convenient nod toward holidays, recipes, and costume parades. Specific to the writing of college students, Canagarajah (2011) suggests translanguaging approaches engage with funds of knowledge within transactional relationships and operate from negotiated achievements of meaning. Pacheco and Miller (2015), as well as Wiley (1996), suggest the same approach, even of younger students. The inclusion of funds of knowledge necessitates a democratic learning stance.

Perspectives focused on language/text valuing

Academic literacy, disciplinary literacies, and professional literacies

The literacy practices, routines, and values of professionals have been adopted as aspirational goals for textualizing in college reading through the construct of academic literacies (Lea and Street, 2006), and, more recently, disciplinary literacies (Shanahan and Shanahan, 2008; see also Bean et al., 2018). Whereas an academic literacies model, as explained by Lea and Street (2006) "is concerned with meaning making, identity, power, and authority, and foregrounds the institutional nature of what counts as knowledge in any particular academic context" (p. 369), disciplinary literacies as a perspective values the tools (processes and text types) that are specific to particular disciplines, such as history, chemistry, and mathematics (Shanahan et al., 2011). Even as college students learn strategic forms for writing, reading, and studying, the very reading habits and rhetoric guiding these processes differ across different subject areas. The tools used for communication purposes are driven by various domain-specific epistemological perspectives (Lee and Spratley, 2010; Moje, 2007; Shanahan and Shanahan, 2012). Although the disciplinary literacy literature stops short of defining what counts as a text, it does refine notions of what is valued, based on communication norms within a discipline. As evidenced by its very name, but also by the foci of so many expert reader studies (history, chemistry, mathematics; Shanahan et al., 2011), it is clear that what is privileged in disciplinary literacy are reading and writing in traditional academic disciplines. However, the college context spans well beyond traditional academic areas, particularly in community college and technical college contexts. Indeed, in areas of career technical education, the literacy practices present a different set of needs that go beyond traditional text structures and demands requiring unique competency and practices might be classified as professional literacies (Armstrong et al., 2021).

Vernaculars

Literacy learning should be in a constant state of change in its accommodation to the individuals practicing with it. Likewise, the building blocks and metaphors for literacy also flex with use. One specific approach to individuals' literacy development is the area of vernacular literacies. Vernacular refers to the everyday uses of language by participating individuals (see also Gee, 2001 and the construct of primary and secondary discourses), which makes it similar to funds of knowledge, but is more explicitly focused on language.

Respect for vernacular language and literacy recognizes and values the productive literacy practices students have learned and used at home. Therefore, college learning would use and value the "home talk" of its participants as resources in learning. Used by Ivanič et al. (2009), the term vernacular literacy in college learning discourses urges us to value and consider college students'

non-academic versions as communication that teachers accept and negotiate. The successful use of “the vernacular” inside of more restricted rhetorical practice of formal schooling would be successful border crossing. Ivanic also uses the term “border literacies” or

reading and writing practices in other domains of students’ lives-home, work, community-that are, or have the potential to be, situated also in the educational domain, and may therefore be mobilized to support student learning (p. 22).

From our uptake of the recommendations made by Ivanic et al. the use of vernacular language must be layered in social or learning contexts, so that from a pedagogical perspective, the vernacular evolves.

One model in which such a shift has been successfully conceptualized is translanguaging (Bernfeld, 2020; Garcia and Sylvan, 2011; Inoue, 2020; Young et al., 2018). Like recommendations for vernacular literacy, translanguaging “workshops” language use, substituting gateways of “correct usage” with shared communication as the goal. All language resources (vernacular, dialect, registers, as well as additional languages) are deployed to solve communication issues. Interestingly, Otheguy et al. (2015) suggest that it may be more productive to stop thinking in terms of discrete languages and instead focus on the collection of undifferentiated language aptitude and resources carried by all participants, students as well as teachers. In any case, students will certainly benefit from fusing their vernacular literacy in combination with those literacies valued by schooling (Ivanic et al., 2009). But it is not clear how those same students learn to do so, nor if specific instructors will be amenable, or even permitted for vernacular and less formal literacies in school contexts.

Multimodality and new literacy studies

More recent sociocognitive approaches to literacy have increased the purview of texts with multimodality and with New Literacy Studies. Mainly concerned with characteristics of the texts, multimodality focuses on the uses and creation of texts, while pointing out the various modes (image, sound, gesture, etc.) that convey meaning. Texts are broadly reconstituted as organized semiotic systems that communicate information via recognizable, shared, and predictable signs. This semiotic perspective is in stark contrast with literacies based on more traditional alphabetic texts. Of course, traditional texts were also always multimodal, with images, tables, as well as the meaning-making practices of their readers. The meaning-making processes of traditional texts (particularly, traditional paper-and-pencil, or keyboard-and-printer texts) are not lost in multimodal ensembles but elaborated and enhanced by texts from other modalities. Yet, with the inclusions of multiple modes and multiple materials of conveyance, little attention to postsecondary, and especially career technical education contexts, has registered in the studies and theorizing emanating from multimodal stances.

Primarily concerned with the *uses* of texts, that is, how texts are used in different social contexts, New Literacy Studies (NLS) derives from systematic, qualitative inquiries into the use of textual materials. From the premises of NLS (Knobel and Lankshear, 2007), text has also been impacted by multimodal influences to include many different kinds of texts that include visual texts, audio recordings, performances, as well as a plethora of objects that present information to “readers.” But even more importantly, the collection of social practices, which came to be known as literacy events (Heath, 1982), has problematized the very notion of autonomous literacies, such that literacy practices are now seen as socially grounded (Street, 1995). Although much of the progress in NLS has come from systematic ethnographic research into social practices with literacy, no recognizable body of work has been focused on postsecondary cultures. Indeed, the social practices with various objects in these fields are ripe for inquiry in the ways the objects are used as texts. This has simply not happened. One promising foray into text/object explication comes from researchers and writers in anthropology and performance studies (see section on **Entextualization**, below).

Perspectives focused on the expectations and purposes of the college context

Contextualization

Mistaken assumptions regarding the work of college reading in instructional contexts point to skills-oriented and therefore remedial practices. Such assumptions, however, ignore the actual expectations and purposes for reading in the college context. A functional context approach calls for development of curriculum that focuses literacy within real-world contexts. Such an approach facilitates students’ learning within the instructional setting and encourages transfer of learning to situations beyond the setting as well as the course itself (see Goldberg, 1951; Roueche and Roueche, 1993; Shoemaker, 1960; Sticht, 1975a,b, 1997; Sticht et al., 1987). Recently under *The Workforce Innovation and Opportunity Act* (2014), Integrated Education and Training models concurrently provide literacy activities with career technical training in specific occupational clusters in models such as I-BEST (see Hirschy et al., 2011; Zeidenberg et al., 2010). The use of contextualization must by its nature include the construct of intentional transfer as the purpose for learning particular literacy practices is two-pronged: learning for the immediate course goal of content mastery and learning for the long-term purpose of civil and workplace success.

Entextualization

The term was first coined by [Baumann and Briggs \(1990\)](#) in performance theory to capture the idea that performers must assess a situation (e.g., read the situation) in order to gauge their current performance and plan for the next action in their sequence of presentation. In effect the performers turn the setting into a text and read it in order to know how to proceed.

In the career technical education environment, composers of messages (communication of multiple varieties) must similarly “suss out” which objects can be productively used as texts (any item that carries needed information), how to read (process) the texts chosen, and then how to use the information with the projected audience. The items that occur in such contexts (any context for that matter) that have pertinent information and can be productively, reliably read, are called entexts. These are the objects like a drip bag, a cam shaft, a thermometer, a hemline, a recipe, all of which are interpreted (read) for vital information in the conduct of literacy events within the disciplinary contexts. Far from assumptions of discrete-skills instruction and remediation of deficiencies, then, reading at the college level requires sophistication in analysis and interpretation. It is, in short, a juggling act of intellectual creativity with a variety of information-bearing objects.

Positioning theory

We include one final perspective here that must be considered in conversations specific to the context and purpose of college reading. For this, we circle back to the very first perspective in our list of eight: lifespan literacy. It is the unfortunate present—and historical—reality that college reading carries a low status as both a field of professionals and as a needed area of instruction in postsecondary contexts. That reality must be captured in any model of college reading.

In descriptions of the social contexts that house disciplinary literacies and professional literacies is the notion of users’ positioning. [Harré and van Langenhove’s \(1999\)](#) positioning theory, a hierarchically organized theory of power distribution, analyzes individuals’ relative positions within bounded social contexts. The theory can be broadened to consider organizations, institutions, or fields of inquiry ([McVee et al., 2019](#)). Indeed, history demonstrates that across the past five decades since the coming of the developmental education field that college reading as a field and as a faculty have suffered from low prestige and lack of security within higher education. Scholarly elitism, academic elitism, or the privileging of traditional general education disciplines continue to define “academic literacies” as the fields with prestige, with physical books serving as the prestigious texts. Such bias is alive and well in US higher education.

Additionally, positioning must be considered from the perspective of the relationship of students as readers with faculty as conveyors of text and text practices. This is particularly relevant if students are expected to be fully text ready from the first day of a course rather than faculty anticipating the need to apprentice students across the duration of the course to develop the necessary disciplinary literacy and/or professional literacy competencies so as to achieve student learning objectives. It is problematic, to say the least, that reading is simultaneously low status to the point of invisibility, yet also the single most widely used learning mode in college contexts.

Conclusions and future directions

We suggest that while the postsecondary community at large may have expectations of the need for reading competency, it is not cognizant of reading’s foundational role in the culture of the institution and most certainly as represented by the constructs covered here. Though these constructs do not yet offer a comprehensive explanation for reading at the postsecondary level, the perspectives we have offered are intended to begin to describe the present state of college reading and to begin to shape pedagogical needs for acknowledging the specific—yet varying—reading and literacy demands, expectations, and practices represented across the postsecondary context.

The field of college reading has work to do to debunk myths that reading is mastered in the primary grades and to reverse assumptions that suggest it is socially acceptable to extend writing and mathematical instruction into the college years, but that doing the same with reading is simply remediating deficits. Specifically, especially if there is to be an intentional shift toward more assets-based approaches to creating student-ready campuses (as opposed to the more deficit-oriented notion of “college-ready” students), then more research is needed to effectively shift social and political assumptions that reading instruction at the postsecondary/tertiary level—and thus a field of professional experts—is unnecessary. This work is necessarily multifaceted, given the complex and varied nature of both the field of CR as well as the actual realities of reading in college worldwide.

First, broader explorations of developmental and remedial education globally will necessarily shape these conversations; key examples of this have been ongoing in Italy ([De Paola and Scoppa, 2014](#); [Sgobbi, 2019](#)). Second, problematizing the notion of ‘readiness’ has certainly been a staple of this work, as [Luoch’s \(2017\)](#) study of students at the United States International University-Africa in Kenya indicates. Third, more focused visibility needs to be drawn specifically to reading. This will include both a focus on the literacy experiences students bring to postsecondary/tertiary education as well as the reading demands student encounter within higher education. Toward the former focus, studies with Turkish university students have focused on reader profiles and habits ([Alan, 2020](#); [Erdem, 2015](#); [Odabaş et al., 2008](#); [Solmaz, 2018](#)) whereas [Trigos-Carrillo \(2019\)](#) has explored the need for academic literacies to take into account family and community literacies for first-generation college students in Latin America. Toward the latter focus, studies of 1st-year university students in Belgium have emphasized the literacy expectations students will face ([Heeren et al., 2021](#); [Van Houtven et al., 2013](#)).

Regardless of the particular context, reading is a key success factor for learners in nearly all formal postsecondary/tertiary education. This, coupled with the realities of the increasingly noticeable internationalization of higher education, is indication that more work on reading in college is needed and that perspectives from experts doing the work of CR globally must do better toward collaboration.

References

- Abell, A., 1894. Rapid reading: advantage and methods. *Educ. Rev.* 8, 283–286.
- Alan, Y., 2020. Türkiye'deki Üniversite Öğrencilerinin Kitap Okuma Profilleri. *Söylem Filoloji Dergisi* 5 (1), 278–303. <https://doi.org/10.29110/soylemdergi.720634>. Available at.
- Alexander, P.A., 2006. The path to competence: a lifespan developmental perspective on reading. *J. Literacy Res.* 37 (4), 413–436. Available from. https://journals.sagepub.com/doi/pdf/10.1207/s15548430jlr3704_1.
- Arndale, D.R., 2010. Access at the Crossroads: Learning Assistance in Higher Education. Jossey-Bass, San Francisco.
- Armstrong, S.L., Newman, M.C., 2011. Teaching textual conversations: intertextuality in the college reading classroom. *J. Coll. Read. Learn.* 41 (2), 7–22.
- Armstrong, S.L., Williams, C.A., 2019. Reconceptualizing the work of placement: toward a culture of inquiry in developmental reading. *J. Coll. Literacy Learn.* 45, 54–68.
- Armstrong, S.L., Williams, J., Stahl, N.A., 2018. Integrated reading and writing. In: Flippo, R.F., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York. <https://doi.org/10.4324/9781315629810>. Available from.
- Armstrong, S.L., Stahl, N.A., King, J.R., 2021. What does it mean to be college ready for career technical education? *Community Coll. Rev.* 49 (2), 131–155. <https://doi.org/10.1177/0091552120982031>. Available from.
- Bailey, T., Jeong, D., Cho, S., 2010. Referral, enrollment, and completion in developmental education sequences in community colleges. *Econ. Educ. Rev.* 29 (2), 255–270. <https://doi.org/10.1016/j.econedurev.2009.09.002>. Available at.
- Bartholomae, D., Petrosky, A., 1986. Facts, Artifacts and Counterfacts: Theory and Method for a Reading and Writing Course. Boynton/Cook, Upper Montclair.
- Baumann, R., Briggs, C., 1990. Poetics and performance as critical perspectives in language and social life. *Annu. Rev. Anthropol.* 19, 59–88. https://doi.org/10.1007/978-1-349-92299-4_41. Available from.
- Bean, T.W., Gregory, K., Dunkerly-Bean, J., 2018. Disciplinary reading. In: Flippo, R.F., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 89–97. <https://doi.org/10.4324/9781315629810>. Available from.
- Berger, A., Peebles, J.D., 1976. Rates of Comprehension: An Annotated Bibliography. Newark, International Reading Association.
- Bernfeld, T., 2020. El Futuro Es Multilingüe: moving from monolingual resistance to engagement with multilingual texts. *J. Coll. Read. Learn.* 50 (2), 113–126. <https://doi.org/10.1080/10790195.2020.1737596>. Available from.
- Bliss, L.B., 2001. Assessment in developmental education. In: Farmer, V.L., Barham, W.A. (Eds.), *Selected Models of Developmental Education Programs in Higher Education*. University Press of America, Lanham, NY, pp. 359–386.
- Book, W., 1927a. How Well College Students Can Read, vol. 26. School and Society, pp. 242–248.
- Book, W., 1927b. Results Obtained in a Special “How to Study” Course Given to College Students, vol. 26. School and Society, pp. 529–534.
- Canagarajah, S., 2011. Codemeshing in academic writing: identifying teachable strategies of translanguaging. *Mod. Lang. J.* 95 (3), 401–417. <https://doi.org/10.1111/j.1540-4781.2011.01207.x>. Available from.
- Carroll, J.B. (Ed.), 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. MIT Press, Cambridge. Available from. <https://ia802605.us.archive.org/7/items/languagethought00whor/languagethought00whor.pdf>.
- Chall, J., 1983. *Stages of Reading Development*. McGraw-Hill, New York.
- Christ, F., 1971. Systems for learning assistance: learners, learning facilitators, and learning centers. In: Christ, F. (Ed.), *Interdisciplinary Aspects of Reading Instruction: Proceedings of the 4th Annual Conference of the Western College Reading Association*. Western College Reading Association, Whittier, pp. 32–41. <https://doi.org/10.1080/24699365.1971.11669422>. Available from.
- Dagherty, L., Gomez, C.J., Carew, D.G., Mendoza-Graf, A., Miller, T., 2018. Designing and Implementing Corequisite Models of Developmental Education: Findings From Texas Community Colleges. RAND Corporation, Santa Monica. Available at. https://www.rand.org/content/dam/rand/pubs/research_reports/RR2300/RR2337/RAND_RR2337.pdf.
- De Paola, M., Scoppa, V., 2014. The effectiveness of remedial courses in Italy: a fuzzy regression discontinuity design. *J. Popul. Econ.* 27, 365–386.
- Edgecombe, N.D., Cormier, M.S., Bickerstaff, S.E., Barragan, M., 2013. Strengthening Developmental Education Reforms: Evidence on Implementation Efforts From the Scaling Innovation Project (CCRC Working Paper No. 61). Community College Research Center, New York. <https://doi.org/10.7916/D8QV3JJ3>. Available at.
- Erdem, A.R., 2015. A Research on Reading Habits of University Students: Sample of Ankara University and Erciyes University. In: *Procedia—Social and Behavioral Sciences*, vol. 174, pp. 3983–3990.
- Eurich, A., 1929. *A Speed of Reading Test for College Students*. University of Minnesota Press, Minneapolis.
- Flippo, R.F., Armstrong, S.L., Schumm, J., 2018. Reading tests. In: Flippo, R.F., Bean, T. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 340–380.
- García, O., Sylvan, C., 2011. Pedagogies and practices in multilingual classrooms: singularities in pluralities. *Mod. Lang. J.* 95, 385–400. <https://doi.org/10.1111/j.1540-4781.2011.01208.x>.
- Gee, J.P., 2001. Reading as situated language: a sociocognitive perspective. *J. Adolesc. Adult Literacy* 44 (8), 714–725. <https://doi.org/10.1598/0710.04>. Available from.
- Goldberg, S., 1951. *Army Training of Illiterates in World War II*. Teachers College, Columbia University, New York.
- González, N., Moll, L., Amanti, C. (Eds.), 2006. *Funds of Knowledge: Theorizing Practice in Households, Communities, and Classrooms*. Erlbaum, Mahwah. <https://doi.org/10.4324/9781410613462>. Available from.
- Haggerty, M., Eurich, A., 1929. *A Test of Reading Comprehension*. University of Minnesota Press, Minneapolis.
- Harre, R., van Langenhove, L. (Eds.), 1999. *Positioning Theory*. Blackwell Publishers, Malden.
- Heath, S.B., 1982. *Ways With Words*. Cambridge University Press, New York.
- Heeren, J., Speelman, D., De Wachter, L., 2021. A practical academic reading and vocabulary screening test as a predictor of achievement in first-year university students: implications for test purpose and use. *Int. J. Biling. Educ. Biling.* 24 (10), 1458–1473. <https://doi.org/10.1080/13670050.2019.1709411>. Available at.
- Heron, E., 1989. *The Dilemma of College Reading Instruction: A Developmental Analysis*. Ed.D. Harvard University.
- Hinsdale, B., 1900. *The Art of Study*. American, New York.
- Hirschy, A., Bremer, C., Castellano, M., 2011. Career and technical education (CTE) student success in community colleges: a conceptual model. *Community Coll. Rev.* 39, 296–318. <https://doi.org/10.1177/0091552111416349>. Available from.
- Holmes, J., 1948. *Factors Underlying Major Reading Disabilities at the College Level*. Ph.D. University of California, Berkeley.
- Holschuh, J.P., Lampi, J.P., 2018. Comprehension. In: Flippo, R.F., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 118–142.
- Holschuh, J.P., Paulson, E.J., 2013. *The Terrain of College Developmental Reading [Invited White Paper]*. College Reading & Learning Association. Available from. <http://www.crla.net/publications.htm>.

- Inoue, A., 2020. Teaching antiracist reading. *J. Coll. Read. Learn.* 50, 134–156. <https://doi.org/10.1080/10790195.2020.1787079>. Available from.
- Ivanic, R., Edwards, R., Barton, D., Martin-Jones, M., Fowler, Z., Hughes, B., Mannion, G., Miller, K., Satchwell, C., Smith, J., 2009. *Improving Learning in College: Rethinking Literacies Across the Curriculum*. Routledge, London.
- Jenkins, D., Zeidenberg, M., Kienzl, G., 2009. Educational Outcomes of I-BEST, Washington State Community and Technical College System's Integrated Basic Education and Skills Training Program Findings From Multivariate Analysis. CCRC Working Paper No. 16. Community College Research Center, New York. Available at. <https://ccrc.tc.columbia.edu/publications/i-best-multivariate-analysis.html#:~:text=Students%20participating%20in%20I-BEST%20achieved%20better%20educational%20outcomes,and%20make%20point%20gains%20on%20basic%20skills%20tests>.
- Jewitt, C. (Ed.), 2014. *The Routledge Handbook of Multimodal Analysis*. Routledge, London.
- Johnson, L.L., Carpenter, K., 2000. College reading programs. In: Flippo, R.F., Caverly, D.C. (Eds.), *Handbook of College Reading and Study Strategy Research*, second ed. Erlbaum, Mahwah, pp. 321–363.
- Kerfoot, J., 1916. *How to Read*. Houghton Mifflin, Boston.
- Kiyama, J.M., Rios-Aguilar, C. (Eds.), 2018. *Funds of Knowledge in Higher Education: Honoring Student's Cultural Experiences and Resources as Strengths*. Routledge, New York.
- Knobel, M., Lankshear, C., 2007. *New Literacies Sampler*. Peter Lang, New York.
- Koopman, H., 1913. *How Students Actually Read*, vol. 23. Education, pp. 563–569.
- Lea, M.R., Street, B.V., 2006. The “academic literacies” model: theory and applications. *Theory Into Pract.* 45 (4), 368–378. https://doi.org/10.1207/s15430421tip4504_11. Available from.
- Lee, C.D., Spratley, A., 2010. *Reading in the Disciplines: The Challenges of Adolescent Literacy*. Carnegie Corporation of New York, New York.
- Leedy, P., 1958. *A History of the Origin and Development of Instruction in Reading Improvement at the College Level*. Ph.D. New York University.
- Leppänen, S., Kytölä, S., Jousmäki, H., Peuronen, S., Westinen, E., 2013. Entextualization and resemiotization as resources for identification in social media. In: Sargeant, P., Tagg, C. (Eds.), *The Language of Social Media: Communication and Community on the Internet*. Palgrave, Basingstoke, pp. 112–138.
- Lowe, A., 1968. The reading improvement programs of Florida: institutions of higher learning, 1966–67. In: Schick, G., May, M. (Eds.), *Multidisciplinary Aspects of College-Adult Reading: 17th Yearbook of the National Reading Conference*. Milwaukee: National Reading Conference, pp. 149–155.
- Luoch, T.O., 2017. Readiness: the key factor in remedial teaching and learning. *Eur. J. Soc. Behav. Sci.* 18 (1), 25–41. <https://doi.org/10.15405/ejbs.203>. Available at.
- Maxwell, M., 1966. *An Individualized College Learning Laboratory. Reading Improvement*, vol. 4, pp. 5–6.
- Maxwell, M., 1979. *Improving Student Learning Skills*. Jossey-Bass, San Francisco.
- Maxwell, M., 1997. *Improving Student Learning Skills: A New Edition*. H & H Publications, Clearwater.
- Maxwell, M., 2000. Forward. In: Flippo, R.F., Caverly, D.C. (Eds.), *Handbook of College Reading and Study Strategy Research*, second ed. Erlbaum, Mahwah, pp. vii–xii.
- McCallister, J., 1942. *Purposeful Reading in College*. Appleton-Century-Croft, New York.
- McMurry, F., 1909. *How to Study and Teaching How to Study*. Houghton Mifflin, Cambridge.
- McVee, M.B., Silvestri, K.N., Barrett, N., Haq, K.S., 2019. Positioning theory. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, pp. 381–400.
- McWhorter, K., 1980. *College Reading and Study Skills*. Little Brown, Boston.
- Moje, E., 2007. Developing socially just subject-matter instruction: a review of the literature on disciplinary literacy teaching. *Rev. Res. Educ.* 31, 1–44. <https://doi.org/10.3102/0091732x07300046001>. Available from.
- Moore, E., 1915. *An Experiment in Teaching College Students How to Study*, vol. 2. School and Society, pp. 100–107.
- Nist, S., Diehl, W., 1985. *Developing Textbook Thinking: Strategies for Success in College*. D. C. Heath, Boston.
- Nist, S.L., Simpson, M.L., 2000. College studying. In: Kamil, M., Barr, R., Mosenthal, P., Pearson, P.D. (Eds.), *Handbook of Reading Research*, third ed. Erlbaum, Mahwah, NJ, pp. 645–666.
- Odabaş, H., Odabaş, Z.Y., Polat, C., 2008. Üniversite Öğrencilerinin Okuma Alışkanlığı: Ankara Üniversitesi Örneği. *Bilgi Dünyası* 9 (2), 431–465.
- Otheguy, R., García, O., Reid, W., 2015. Clarifying translanguaging and deconstructing named languages: a perspective from linguistics. *Appl. Ling. Rev.* 6, 281–307. <https://doi.org/10.1515/applirev-2015-0014>. Available from.
- Pacheco, M., Miller, M., 2015. Making meaning through translanguaging in the literacy classroom. *Read. Teach.* 69, 533–537. <https://doi.org/10.1002/trtr.1390>. Available from.
- Parr, F., 1930. The Extent of Remedial Reading Work in State Universities in the United States, vol. 31. *School and Society*, pp. 547–548.
- Perin, D., 2011. Facilitating student learning through contextualization: a review of evidence. *Community Coll. Rev.* 39, 268–295. <https://doi.org/10.1177/0091552111416227>. Available from.
- Perry, W., 1959. Student use and misuse of reading skills: a report to the faculty. *Harv. Educ. Rev.* 29, 193–200.
- Robinson, F., 1943. *Study Skills of Soldiers in ASTP*, vol. 58. *School and Society*, pp. 398–399.
- Robinson, F., 1946. *Effective Study*. Harper Row, New York.
- Roueche, J.E., Kirk, R.W., 1973. *Catching Up: Remedial Education*. Jossey-Bass, San Francisco.
- Roueche, S., Roueche, J., 1993. Literacy development: foundation for general education. In: Johnson, J.L. (Ed.), *New Directions for Community Colleges*, vol. 40. Jossey-Bass, San Francisco, pp. 31–37.
- Rutschow, E.Z., Cormier, M.S., Dukes, D., Zamora, D.E.C., 2019. The Changing Landscape of Developmental Education Practices: Findings From a National Survey and Interviews With Postsecondary Institutions. Center for the Analysis of Postsecondary Readiness, New York. Available at. <https://postsecondaryreadiness.org/wp-content/uploads/2019/11/changing-landscape-developmental-education-practices.pdf>.
- Sahito, Z., Siddiqui, A., Khawaja, M., Shaheen, A., Saeed, H., Laghari, S.H., 2017. Teaching of remedial English and the problems of the students: a case of university of Sindh, Jamshoro, Sindh, Pakistan. *World J. Engl. Lang.* 7 (1), 1–10.
- Sgobbi, F., 2019. The effectiveness of remedial courses: an analysis of freshmen in industrial engineering. *Educ. Econ.* <https://doi.org/10.1080/09645292.2021.1974343>. Available at.
- Shanahan, T., Shanahan, C., 2008. Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harv. Educ. Rev.* 78, 40–59. <https://doi.org/10.7763/haer.78.1.v62444321p602101>. Available from.
- Shanahan, T., Shanahan, C., 2012. What is disciplinary literacy and why does it matter? *Top. Lang. Disord.* 32, 7–18. <https://doi.org/10.1097/tld.0b013e318244557a>. Available from.
- Shanahan, C., Shanahan, T., Mischia, C., 2011. Analysis of expert readers in three disciplines: history, mathematics, and chemistry. *J. Literacy Res.* 43, 393–429. <https://doi.org/10.1177/1086296x11424071>. Available from.
- Shoemaker, H.A., 1960. The functional context method of instruction. *IRE Trans. Educ. E-3* (2), 52–57. <https://doi.org/10.1109/te.1960.4322128>. Available from.
- Simpson, M.L., Nist, S.L., 2002. Encouraging active reading at the college level. In: Pressley, M., Collins-Block, C. (Eds.), *Comprehension Instruction: Research Based Best Practices*. Guilford, New York, pp. 365–379.
- Simpson, M.L., Stahl, N.A., Francis, M., 2004. Reading and learning strategies: recommendations for the 21st century. *J. Dev. Educ.* 28 (2), 2–4, 6, 8, 10–12, 14.
- Simpson, M.L., 2002. Program evaluation studies: strategic learning delivery model suggestions. *J. Dev. Educ.* 26, 2–10.
- Skinner, B., 1957. *Verbal Behavior*. Copley, Acton.
- Solmaz, M., 2018. Üniversite Öğrencilerinin Okuma Alışkanlığı Üzerine Bir Araştırma: Yüzüncü Yıl Üniversitesi Örneği. *Türkiye Sosyal Araştırmalar Dergisi* 22 (2), 603–622.
- Stahl, N.A., Armstrong, S.L., 2018. Re-claiming, re-inventing, and re-reforming a field: the future of college reading. *J. Coll. Read. Learn.* 48 (1), 47–66.

- Stahl, N.A., Armstrong, S.L., 2020. So much more than SQ3R: a life history of Francis P. Robinson. *Read. Psychol.* 41 (4), 287–321. <https://doi.org/10.1080/02702711.2020.1768979>. Available from.
- Stahl, N.A., King, J.R., 2018. History. In: Flippo, R.F., Bean, T.W. (Eds.), *Handbook of College Reading and Study Strategy Research*, third ed. Routledge, New York, pp. 3–26.
- Stahl, N.A., Armstrong, S.L., Hewett, E., 2020. Contextualization models for community college reading and study strategies: an integrative literature review. *Community Coll. J.* 45 (10), 735–755. <https://doi.org/10.1080/10668926.2020.1785355>. Available from.
- Sticht, T.G., Armstrong, W.B., Hickey, D.T., Caylor, J.S., 1987. *Cast-Off Youth: Policy and Training Methods From the Military Experience*. Praeger, New York.
- Sticht, T.G., 1975a. *Reading for Working: A Functional Literacy Anthology*. Human Resources Research Organization, Alexandria.
- Sticht, T.G., 1975b. *A Program of Army Functional Job Reading Training*. HumRRO-FR-WD-CA-75-7. Human Resources Research Organization, Alexandria.
- Sticht, T.G., 1997. *The Theory Behind Content-Based Instruction. Focus on Basics, Content -Based Instruction*, vol. 1 (D). Available from. <http://www.ncsall.net/?id=433>.
- Strang, R., 1937. The improvement of reading in college. *Engl. J.* 26, 548–559.
- Street, B., 1995. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Longman, New York.
- Sunset for State Operating Subsidies for Remedial Courses, 2017. Section 3345.061. Ohio. Available at. <https://codes.ohio.gov/ohio-revised-code/section-3345.061>.
- Tennessee Senate Bill No. 7006, 2010. Public Acts 2010. Available at. <https://www.tbr.edu/sites/tbr.edu/files/media/2015/01/Complete%20College%20TN%20Act%202010%20-%20signed.pdf>.
- Traxler, A.E., 1938. The freshman curriculum and methods of teaching. In: *From High School to College*, N.E.A, pp. 89–94. Research Bulletin, No. 2.
- Triggs, F., 1943. *Remedial Reading: The Diagnosis and Correction of Reading Difficulties at the College Level*. University of Minnesota Press, Minneapolis.
- Trigos-Carrillo, L., 2019. Community cultural wealth and literacy capital in Latin American communities. *Engl. Teach. Pract. Critiq.* 19 (1), 3–19. <https://doi.org/10.1108/ETPC-05-2019-0071>. Available at.
- Van Houtven, T., Peters, E., Van den Branden, K., 2013. Analyzing student teachers' academic literacy needs: a qualitative analysis of Flemish first-year teacher trainees' needs. *Lang. Learn. High. Educ.* 2 (2), 267–292.
- Wachen, J., Jenkins, D., Van Noy, M., 2011. Integrating basic skills and career-technical instruction: findings from a field study of Washington state's I-best model. *Community Coll. Rev.* 39 (2), 136–159. <https://doi.org/10.1177/0091552111406108>. Available at.
- Wiley, T., 1996. *Literacy and Language Diversity in the United States*. Center for Applied Linguistics, Washington, DC. Available from ERIC Document Reproduction Services ED 440 557.
- Wood, E.N., 1960. A breakthrough in reading. *Read. Teach.* 14, 115–117.
- Workforce Innovation and Opportunity Act of 2014, 29 U.S.C. § 3101 et seq, (2014). Available at <https://uscode.house.gov/view.xhtml?path=/prelim@title29/chapter32&edition=prelim>.
- Wyatt, M., 1992. The past, present, and future need for college reading courses in the US. *J. Read.* 36 (1), 10–20.
- Young, V.A., Barrett, R., Young-Rivera, Y., Lovejoy, K., 2018. *Other People's English: Code-Meshing, Code-Switching, and African American Literacy*, revised and updated ed. Parlor Press, Anderson, SC.
- Zeidenberg, M., Cho, S., Jenkins, D., 2010. *Washington State's Integrated Basic Education and Skills Training Program (I-BEST): New Evidence of Effectiveness*. CCRC Working Paper No. 20. Community College Research Center, New York.

Research and practice in literacy teacher education: imagining, designing, and engaging in “what could be”

James V. Hoffman, Language and Literacy Studies, Department of Teacher Education and Administration, The University of North Texas, Denton, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Characteristics and principles of design research	474
Research into practice	474
Research in practice	475
Literacy teacher preparation	475
Literacy mentoring in hybrid spaces	476
Coaching with CARE	476
The BETTER project in Mozambique	477
Considerations and cautions in a design research agenda	478
Imagination in design research	478
Collaboration vs. scale in design research	478
Publishing in the academy	478
Conclusion	478
References	479

Herb Simon, of noble laureate fame, wrote often regarding the intersections of research, science, and practice in the realm of higher education (Simon, 2019). He distinguished between research that explores “what is” as contrasted with research that engages with “what could be.” The latter form of research he framed as driving the “sciences of the artificial.” “Artificial” here refers to the construction of things—not artificial in the sense of fake or unreal. Simon argued that engineering, for example, is one of the sciences of the artificial. Engineers, for example, make things in their research such as computer scientists building more powerful algorithms for data processing. Simon was not opposed to sciences that explore the “what is”—as in a psychologist’s efforts to understand the mind or a geologist’s efforts to understand the earth’s elements through research. He was just opposed to the lack of attention to and misrepresentation of the sciences of the artificial as derivative and less theoretical.

The field of education, for all kinds of historical reasons, has been primarily rooted in the academic discipline of psychology—a field of inquiry focused on how the mind works. I will make the argument here, that this historical tie has sometimes diverted educators from a much more powerful path toward research, science, and practice. Education is an artificial science. As educators, we strive to improve on the tools, systems, structures that serve the needs of children, youth and adults through design and redesign. We can draw on the knowledge generated in the traditional sciences (e.g., psychology, sociology, anthropology, child development), but our research could become much more powerful through a focus on improving practices using a design lens.

I will argue for a “what could be” vision for the relationship between research, practice, and literacy teacher preparation. In constructing our argument for a break from the traditions of psychological research, I will offer examples of research from my own work most often conducted in association with research colloquies. The focus on literacy and teacher education reflects my history as a literacy teacher educator. I subscribe to a philosophy of teacher preparation as a space for transformative change that impacts practice—not “teacher training” that prepares the next generation of teachers to replicate what is. Consistent with the call to make this *Encyclopedia* more international in perspective, I will offer examples of our own design research efforts both in the United States and in Southern Africa. This vision offers alternatives for literacy teacher educators to re-imagine who we are in the construction of scientific knowledge and the roles we play. This vision does not come with a promise of quick fixes to the challenges we face in education. A design perspective is typically iterative with progress realized in small increments.

Characteristics and principles of design research

Research into practice

In 2020, I participated in the Literacy Research Association’s online conference in this COVID era. Here are some variations on a theme I heard again and again: “How do we get teachers to do what research says is best practice?” “Why don’t teachers pay more attention to research?” “Research shows that what most teachers are doing in their classrooms does not reflect best practices from research.”

Underlying these kinds of questions is the view that it is the work of researchers to study processes like literacy and make claims that practitioners are expected to take up and use. This is research into practice. Sub-themes of these questions include:

- Compared to other approaches and methods, my research shows mine is better (than X, Y or Z);

- Do it with fidelity (just the way we did it in my research); and,
- Trust me ... what that other researcher is claiming to be true is not.

The subtext in all of these statements is about researchers as more informed, more capable, even more intelligent than practitioners. The subtext is also about the view of “best practices” as a real thing. Of course, this is not a new theme, but it has become more consequential during a time of mandates from national (e.g., No Child Left Behind and Reading First) and state initiatives (e.g., legislation in Texas regarding the training of all teachers in the LETRS program). Is it any wonder that practitioners become exhausted in their efforts to sort out sometimes competing claims, and become increasingly cynical in the face of the research treadmill coming to regard researchers as out of touch with reality? Is it any wonder that practitioners regard themselves as objects in a system that offers no space for them to contribute?

Research in practice

Just the simple adjustment from “research *into*” to “research *in*” practice leads us to consider a very different perspective on research and practice—one that is very much aligned with Simon’s sciences of the artificial. This version of research and practice was explored by Lewin (1946) in the form of action research. Lewin was very much about understanding complex systems (like teaching and learning) as they can be studied in context—not by isolating and attempting to control variation. For some time, this notion of action research was taken up as “teacher research” (Cochran-Smith and Lytle, 1993). Unfortunately, this well-intended shift became a euphemism for “second class” research—not theoretical and not worthy of attention. Fortunately, there has been a shift away from the term “teacher research” to “practice-based research” (Sailors and Hoffman, 2019). This shift is more than just relocating traditional educational studies into teaching and learning context (e.g., comparing method a and b), but to a different paradigm for research referred commonly as design research (e.g., Bakker, 2018; Van den Akker et al., 2006) and in other circles as “improvement sciences” research (see, Bryk, 2015; Lewis, 2015). According to Gutiérrez et al. (2017),

... research is being reconceived in ways that decenters the researchers, and reframes the aims, goals, and outcomes of research. In particular, this body of work foregrounds important principles about knowledge production and tries to advance a new generation of methods with new sensibilities (p. 36).

To be clear, this contrast between design research and traditional psychological research is not about a qualitative vs. quantitative tools and methods. Design research is guided by empirical data gathered and interpreted with the same level of rigor as in traditional psychological research (Sailors and Hoffman, 2019). Some of the important characteristics of design research include:

- Studies emerge through problem posing in practice.
- Studies are situated in the context of practice (nothing is controlled artificially).
- Studies are directed toward the improvement of outcomes.
- Studies are iterative with changes introduced to increase effectiveness—some of these interactions are planned in advance and others might emerge.
- Data collection and interpretation processes are continuous and rigorous.

Design research does not focus on the “best” only on the “bettering” or “improving” of practices. Design research also takes a fundamentally different stance from traditional research related to generalizability. In traditional research, probability theory is used to judge, within confidence parameters, whether a finding observed in a particular study can be generalized to a larger context (for a critical analysis of the assumptions made in this argument, see Carver (1978)). In contrast, design research strives to identify “design principles” to guide not prescribe similar efforts in other contexts. Replication of a study in other contexts is expected to yield adaptations to that different context and the discovery of, or modification of principles.

Simon writing on the sciences of the artificial reached across professional worlds and spawned a movement that is still growing. At the Environmental Design Research Association conference in 1969, Simon explained:

1. That representation of design problems holds the key to their solution.
2. That the complexity of design can be tamed by representing problems in smaller more manageable and potentially well-defined sub-problems.
3. That the packaging of information that goes on the human memory systems is universal, and
4. That the style of the designer is a function of her process (Akin, 2001).

We would argue that education, literacy teaching and teacher preparation, are best situated under this umbrella of artificial sciences and that Simon’s points made to the environmental research community apply equally well to our work as literacy teacher educators.

Literacy teacher preparation

To illustrate this process of design research, I will draw on three programs of research we and other close colleagues have been directly involved with over the past 20 years. I identify as university-based teacher educator. I study my own practices with the goal of improving teacher preparation and, ultimately, the transformation of literacy teaching in schools.

Literacy mentoring in hybrid spaces

University based teacher educators have struggled over decades with the “great divide” between academic preparation and field experiences (Horn and Campbell, 2015). Studies in the late 1960s and early 1970s revealed that the most common model for literacy teacher preparation involved a combination of foundational courses and methods courses typically delivered in lecture format on university campuses followed by one semester of field experiences (i.e., student teaching). Research into the effectiveness of these programs revealed that the student teaching experience shaped the preservice teachers into the traditional practices in classrooms and away from the innovative methods espoused in the academic courses. Over the next several decades there was an increase in early field experiences that were more carefully associated with the methods and philosophies introduced in the academic courses (Zeichner, 2010; Zeichner et al., 2012). While a significant step forward, Britzman (2012) has cautioned that a simple increase in uncritical placements in field settings as part of preservice teacher preparation may only serve to further compromise the transformative stance for change through the “practice makes practice” phenomenon.

In our recent work, we have begun to explore the promise of hybrid spaces as contexts for introducing transformative practices. What began as “reading tutoring” in a single methods course in our teacher preparation program gradually expanded into literacy mentoring as a core set of literacy teaching experiences that spanned the entire program. Over the course of more than ten studies—each study an iteration building on the previous work, I articulated a set of design principles for mentoring that are evidence based and inform our practice (Hoffman and Mosley Wetzel, 2017). With each iteration we found ourselves dealing with more complex issues—as in the hybrid mentoring spaces as contexts for introducing practices not found in the typical classroom placements (e.g., work with inquiry) to specifically address the kinds of concerns raised by Britzman. We do not have a “best” mentoring program for others to reproduce, but we do have a set of principles grounded in data that have grown throughout practice that others can explore and take up in their own context. These design principles are as follows:

1. Design the structure for the mentoring experience explicitly and intentionally to encourage engagement with transformative teaching practices.
 - Strive for a balance between a planned structure (models and expectations for work in certain areas and methodologies) and flexibility (spaces for being responsive) and for the preservice teacher to modify and adapt in their work with students.
 - Provide ample time in sessions to allow for relationship building and deep engagement with content.
 - Consider multiple, varied (e.g., location, participants, focus areas), and cumulative mentoring experiences across the program of preparation.
 - Connect explicitly toward concurrent and future classroom clinical experiences in literacy.
 - Provide spaces for innovative practices that include work with critical literacies, multimodal literacies, and digital literacies.
2. Build on the intellectual and socio-cultural assets and identities of the mentor and the mentee(s).
 - Focus on the importance of building a trusting and caring relationship between the mentor and mentee to encourage risk-taking.
 - Emphasize assessment practices that are appreciative of, building on, and responsive to the resources that mentee brings to the mentoring context.
 - Attend carefully and explicitly to the work of identity building for both the mentor and the mentee (“I am a reader who ...” “I am a writer who ...” “I am a teacher who”
3. Attend to the mediating processes that can support growth in a particular context.
 - Be present as teacher educators leading this work are present and interact directly in both academic coursework and in mentoring through coaching support.
 - Reflect in, on and for practice to encourage to support growth.
 - Nurture a community of practice that brings preservice teachers and teacher educators in spaces to grow together.
 - Connect mentoring directly to academic coursework and provide contexts for exploring pedagogical content knowledge, culturally responsive material, and culturally sustaining teaching frameworks.

The fact that we have articulated these design principles, does not mean our work is finished. We continue to explore how we can explicitly support our preservice teachers in moving the transformative practices into their early classroom field experiences and clinical teaching.

Coaching with CARE

As we moved our teacher preparation practices more deeply into the field in both hybrid spaces and classroom-based practicum experiences, we encountered a new set of contradictions that demanded design work. The coaching support offered by cooperating teachers to their preservice teachers was often coming from an evaluative model that reflected the ways in which these teachers were being coached (clinically supervised) by their administrators. Even the coaching support from the university supervisors was often coming from an authoritative space—just different from the cooperating teachers in the referencing point for what was “good practice.” Preservice teachers working in this context were often left to sort through the conflicting messages and walk the tightrope safely through to program completion.

We saw coaching support as critical in our mentoring work with preservice teachers but found the need to go beyond the evaluative “two glows and a grow” approach (e.g., “Here are two things you did that I liked and here is one thing you need to change . . .”) to explore more reflective models of coaching that would carry over into classroom teaching and life-long teacher development. Across a series of studies, we explored iterations that involved such changes as more intentional work with our preservice teachers on the

coaching model; preparing cooperating teachers to carry forward this model into classroom experiences; video supported coaching; and coaching practices that involved cooperating teachers, university supervisors and preservice teachers working collaboratively. We refer to this model as “Coaching with CARE” (Wetzel et al., 2017).

Through our studies of Coaching with CARE, all drawing on a design research basis, we inferred a set of design principles for coaching and mentoring preservice teachers (Mosley Wetzel et al., 2020).

Existing literacy teaching practices in the placement classroom must be critically examined as a starting point

Cooperating Teachers building a deep understanding of their own teaching practices through inspection, observation, and reflection.

Developing trusting relationships through caring for/with preservice teachers

Caring with and caring for are necessary elements in developing trusting relationships for coaching.

Reflecting within a community

Coaching includes reflecting for practice (before teaching), during practice, and after practice. Reflecting around action is a process that includes the ongoing conversations about literacy teaching and learning that are part of the double loop and triple loop reflection.

Using video and student work to focus on students as a source of feedback

Build a tool kit for exploring tensions in literacy teaching—video and student work—as a means to hold practice still. The use of video in coaching we call RCA translates directly to retrospective video analysis (RVA), a tool that practicing teachers can use in their classroom (Mosley Wetzel et al., 2017).

Building practices for disruption and critical conversations to build more equitable classrooms

We have drawn more and more on narrative approaches to analysis along with discourse analysis to understand the ways that teachers make space for narratives in their coaching and how those narratives become opportunities for critical work around the topics of literacy and social justice.

As with our work in mentoring in hybrid spaces our work with the coaching support process is unfinished. What has become increasingly intriguing to us is the impact of these mentoring experiences on the classroom teaching practices of cooperating teachers in their interactions with their own students.

The BETTER project in Mozambique

Over the past 20 years, we have been involved in multiple research and development projects related to literacy teaching in schools. The Ithuba Project in South Africa (see <https://www.utsa.edu/ithuba/Frames/Home.html>) and the READ Malawi Project (see <https://www.utsa.edu/readmalawi/>). Both projects focused on the development of instructional (supplementary) books for children in official languages to support literacy learning (Hoffman et al., 2016; Sailors et al., 2009, 2013a,b, 2014, 2010, 2012; Sailors and Kaambankadzanga, 2017; Shin et al., 2020). In all this work we have been conscious of our position of power (through control over resources) regarding teaching practices “imported” from the outside into local contexts. We have resisted, what Easterly (2006) described as the “planner” model for international development work where solutions to problems posed by outsiders are implemented on a broad scale. Our goal and process has been to take what Easterly described as a “seeker” stance that involves recognition and collaboration with local expertise. All of the materials we have worked to develop in the Ithuba, and Read Malawi books were authored by local educators, edited locally, and printed in countries. All the staff development work associated with these materials was planned and organized collaboratively with local partners.

The BETTER project in Mozambique marked a shift in our work from materials development to teacher preparation (Sailors and Hoffman, 2019). While the larger BETTER project is focused on promoting interactive methodologies in teacher preparation, our small portion of this project has focused on supporting hybrid mentoring practices with one local teacher training institute. We brought forward into this context the design principles we had developed in our mentoring research. Through our intensive work with one local teacher preparation site over 3 years, we were able to develop a model for mentoring in hybrid spaces that reflected the local spaces, context, and resources. We developed videos and other support materials (e.g., assessment tools) that could be used in the work at additional sites across the country.

Our work in Mozambique has built on and extended our understanding of mentoring in hybrid spaces in two areas. First, the bulk of our work in literacy teacher preparation has been in the area of English. In Mozambique, Portuguese was not the home language for either then mentors or the mentees. It was rare that there was even a shared home language between the home languages of the mentor and mentee. The design principle emerging for us is the need for careful negotiation between the mentor and mentee between indigenous and the “official” targeted language as literacy practices are explored and developed.

A second, emerging design principle related to the insider/outsider expertise and positionality with respect to leadership and guidance. We were extremely fortunate to have partnered with a local teacher educator who quickly stepped into a leadership role in the mentoring project. She adapted and improved the process for both the mentoring and the preparation of the mentors.

She became the leading voice in work with other sites around the country that were interested in adopting the mentoring model. Our voice and experiences were influential but not central to the process as it was taken up and expanded.

Considerations and cautions in a design research agenda

There is nothing special in our work with design. There are other literacy researchers doing design work (see [Gutiérrez, 2018](#); [Reinking, 2021](#); [Bradley and Reinking, 2011a,b](#)). There are also educators engaged in design research without “naming” their work as research, without guidance in how to study their professional work in ways that contribute to our knowledge of practice. Evaluating studies, in the past, were regarded as not theoretical and of lessor value than traditional research paradigms. This view was disrupted by [Reinking and Bradley’s \(2008\)](#) careful articulation of formative design research. In thinking toward a future driven by design research, there are three considerations we would raise for deliberation.

Imagination in design research

The more we have worked with the “what could be” stance toward our research, the more we have become aware of the critical role of imagination in the design process ([Sailors, 2018](#)). Imagination in design research demands that we step outside of the assumptions that have led to the identification of a particular “problem” for study into a more expansive process of problem posing. [Gutiérrez and Penuel \(2014\)](#), for example, describe imagination in relation to design research in critical spaces of transformation. They contrast this approach with traditional approaches to research. Studying the “social life of interventions” moves us away from imagining interventions as fixed packages of strategies with readily measurable outcomes and toward more open-ended social or socially embedded experiments that involve ongoing mutual engagement (p. 20). The focus is on seeing everyday practices in new (and critical) ways and then using design research to explore the “what could be.”

Collaboration vs. scale in design research

Issues of scale are of importance in educational research that draws on traditional psychological models. Large is better, more trusted, and more consequential. Findings from large studies (e.g., with huge sample sizes, systematic sampling, and associated generalizability across contexts) are viewed as the gold standard for research (see [Torgerson and Torgerson, 2001](#)). Design studies, in contrast tend to be small in scale and local. Context is not eliminated or controlled but studied for its influence. The power of design research comes in the interaction of researchers across different contexts engaging with similar challenges and similar design paths. Robust design principles emerge across these contexts. We are currently working as part of a collaborative of researchers at ten different universities engaged in the study of literacy mentoring in hybrid spaces and the impact of preservice teacher development. We have developed shared instruments to use across our sites. We have flexible clusters of institutions researching different contexts for mentoring. We are looking across sites at our findings to generate design principles that can be used to inform practices. This is new and exciting territory for those of us raised in the context of experimental research traditions.

Publishing in the academy

Design research will be of no consequence if it is not shared broadly in scholarly circles. While there are important examples of design research published in academic journals (e.g., [Neuman, 1999](#); [Neuman and Dwyer, 2011](#); [Roskos et al., 2011](#)), these publications tend to adopt of formative design experimental approach to design research (close to traditional psychological research) rather than more expansive versions of design research such as in Engeström’s developmental work research ([Engeström, 2016](#)).

There has been promising work in the scholarly literature. For example, *The Journal of Operations Management* created a new “design science department.” In addition, [Van Aken et al. \(2016\)](#) wrote an inaugural essay describing the qualities and characteristics of what they term DSR (Design Science Research). According to [Van Aken et al. \(2016\)](#), the key questions of validity and relevance are “(a) How strong is the evidence that the design will produce the desired results? (i.e., pragmatic validity); and (b) In what way does the design make a valuable contribution to addressing a significant field problem or exploiting a promising opportunity? (i.e., practical relevance)?” (pp. 1–2). They argue that DSR is a research strategy with an aim toward generic design and the workings of the design. They explain that the differences with more common research design is in at the strategic level not the tactical level. “There are, in principle, no differences at the tactical level of methods for data gathering and analysis” ([Van Aken et al., 2016](#), p. 8). We can aspire toward a future in educational journals may take a similar stance in promoting design research efforts.

Conclusion

Miriam Lewin, Kurt Lewin’s daughter, uses two of Lewin’s most quoted statements related to research and practice in her celebration of his work: “There is nothing as practical as a good theory.” And “If you want to truly understand something you have to try to change it” ([Lewin, 1998](#)). These statements, and Lewin’s work generally, embody the relationship of research and practice that I

have attempted to articulate here in relating research and practice. I would only extend his observations to include that: *there is nothing more theoretical than good practice*.

The sciences of the artificial should be about change and improvement. Reeves et al. (2011) present the case that while educational research has expanded exponentially over the past decades, so has the evidence of failure across the educational system continued to expand the differences between the haves and the have nots. How can this be? Because, they argue, we have been doing the wrong kind of research. The future is to be found, as we have claimed here, in design research.

References

- Akin, O., 2001. Simon says: Design is representation. Unpublished manuscript. School of Architecture, Carnegie Mellon University, Pittsburgh, PA. Retrieved from. <https://users.metu.edu.tr/baykan/arch467/Readings/AradSimon.pdf>.
- Bakker, A., 2018. Design Research in Education: A Practical Guide for Early Career Researchers. Routledge.
- Britzman, D.P., 2012. Practice Makes Practice: A Critical Study of Learning to Teach. SUNY Press.
- Bradley, B.A., Reinking, D., 2011a. Enhancing research and practice in early childhood through formative and design experiments. *Early Child. Dev. Care* 181 (3), 305–319.
- Bradley, B.A., Reinking, D., 2011b. A formative experiment to enhance teacher-child language interactions in a preschool classroom. *J. Early Child. Lit.* 11 (3), 362–401.
- Bryk, A.S., 2015. Accelerating how we learn to improve. *Educ. Res.* 44 (9), 467–477.
- Carver, R., 1978. The case against statistical significance testing. *Harv. Educ. Rev.* 48 (3), 378–399.
- Cochran-Smith, M., Lytle, S.L. (Eds.), 1993. *Inside/Outside: Teacher Research and Knowledge*. Teachers College Press.
- Easterly, W., 2006. Planners vs. searchers in foreign aid. *Asian Dev. Rev.* 23 (2), 1–35.
- Engeström, Y., Lompscher, J., Rückriem, G. (Eds.), 2016. Putting Activity Theory to Work: Contributions From Developmental Work Research, vol. 13. Lehmanns Media.
- Gutiérrez, K.D., 2018. Social design-based experiments: a proleptic approach to literacy. *Lit. Res.* 67 (1), 86–108.
- Gutiérrez, K.D., Peniel, B., 2014. Relevance to practice as a criterion for rigor. *Educ. Res.* 43 (1), 19–23.
- Gutiérrez, K.D., Cortes, K., Cortez, A., DiGiacomo, D., Higgs, J., Johnson, P., et al., 2017. Replacing representation with imagination: finding ingenuity in everyday practices. *Rev. Res. Educ.* 41 (1), 30–60.
- Hoffman, J.V., Mosley Wetzel, M., 2017. Exploring some design principles for tutoring in preservice teacher preparation. *Read. Writ. Q.* 33 (4), 348–363.
- Hoffman, J.V., Sailors, M., Aguirre, S.H., 2016. Thinking globally in literacy instruction: making a difference in the world. *Read. Teach.* 70 (2), 143–148.
- Horn, I.S., Campbell, S.S., 2015. Developing pedagogical judgment in novice teachers: mediated field experience as a pedagogy for teacher education. *Pedagogies* 10 (2), 149–176.
- Lewin, K., 1946. Action research and minority problems. *J. Soc. Issues* 2 (4), 34–46.
- Lewin, M.A., 1998. Kurt Lewin: his psychology and a daughter's recollections. In: *Portraits of Pioneers of Psychology*, vol. III, pp. 105–120.
- Lewis, C., 2015. What is improvement science? Do we need it in education? *Educ. Res.* 44 (1), 54–61.
- Mosley Wetzel, M., Maloch, B., Hoffman, J.V., 2017. Retrospective video analysis: a reflective tool for teachers and teacher educators. *Read. Teach.* 70 (5), 533–542.
- Mosley Wetzel, M., Maloch, B., Hoffman, J.V., 2020. Design principles for coaching toward reflection and transformation in literacy preservice teacher education. *Lit. Res.* 69 (1), 248–265.
- Neuman, S.B., 1999. Books make a difference: a study of access to literacy. *Read. Res. Q.* 34 (3), 286–311.
- Neuman, S.B., Dwyer, J., 2011. Developing vocabulary and conceptual knowledge for low-income preschoolers: a design experiment. *J. Lit. Res.* 43 (2), 103–129.
- Reeves, T.C., McKenney, S., Herrington, J., 2011. Publishing and perishing: the critical importance of educational design research. *Australas. J. Educ. Technol.* 27 (1).
- Reinking, D., 2021. *Design-Based Research in Education: Theory and Applications*. Guilford Publications.
- Reinking, D., Bradley, B.A., 2008. *Formative and Design Experiments: Approaches to Language and Literacy Research*. Teachers College Press, New York.
- Roskos, K., Burstein, K., You, B.K., Brueck, J., O'Brien, C., 2011. A formative study of an e-book instructional model in early literacy. *Creativ. Educ.* 2 (01), 10.
- Sailors, M., 2018. Re-imagining teacher education. In: *Theoretical Models and Processes of Literacy*. Routledge, pp. 430–448.
- Sailors, M., Hoffman, J.V., 2019. Mentoring in transformative hybrid spaces: preservice teacher preparation and literacy learning in Mocambique. *J. Adolesc. Adult Lit.* 63 (2), 127–133.
- Sailors, M., Kaambankadzanja, D., 2017. Developing a culture of readers: complementary materials that engage. *Read. Teach.* 71 (2), 199–208.
- Sailors, M., Hoffman, J.V., Condon, M., 2009. The challenges of developing leveled texts in and for developing countries. In: Heibert, Sailors (Eds.), *Finding the Right Texts*. Guilford, New York, pp. 70–85.
- Sailors, M., Makalela, L., Hoffman, J.V., Frey, N., 2010. Opportunity matters: the Ithuba writing project in South African schools. *Voices Middle* 18 (1), 8.
- Sailors, M., Hoffman, J.V., Chiloru, H., Kaambankadzanja, D., Mapondera, A., Betha, F., 2012. Complementary reading materials development: successes and challenges of the Read Malawi Programme. *Pula vol.* 26 (2), 171–186.
- Sailors, M., Hoffman, J., Wilson, T., Villarreal, L., Peterson, K., Chiloru, H., Saka, T., 2013a. Implementing a school-wide reading program in Malawi: a case study of change. In: 62nd Yearbook of the Literacy Research Association, pp. 220–232.
- Sailors, M., Martinez, M., Villarreal, L., 2013b. Teacher authored supplementary reading materials in South Africa. *Bookbird* 51 (4), 35–46.
- Sailors, M., Hoffman, J.V., David Pearson, P., McClung, N., Shin, J., Phiri, L.M., Saka, T., 2014. Supporting change in literacy instruction in Malawi. *Read. Res. Q.* 49 (2), 209–231.
- Shin, J., Sailors, M., McClung, N., Hoffman, J.V., Pearson, P.D., Kaambankadzanja, D., Mwale, L., 2020. Access to local books: the effects of Read Malawi from a children's rights perspective. *Literacy* 54 (3), 74–90.
- Simon, H.A., 2019. *The Sciences of the Artificial*. MIT press.
- Torgerson, C.J., Torgerson, D.J., 2001. The need for randomised controlled trials in educational research. *Br. J. Educ. Stud.* 49 (3), 316–328.
- Van den Akker, J., Gravemeijer, K., McKenney, S., Nieveen, N. (Eds.), 2006. *Educational Design Research*. Routledge.
- Van Aken, J., Chandrasekaran, A., Halman, J., 2016. Conducting and publishing design science research: inaugural essay of the design science department of the Journal of Operations Management. *J. Oper. Manag.* 47, 1–8.
- Wetzel, M.M., Hoffman, J.V., Maloch, B., 2017. *Mentoring preservice teachers through practice: a framework for coaching with CARE*. Routledge.
- Zeichner, K., 2010. Rethinking the connections between campus courses and field experiences in college-and university-based teacher education. *J. Teach. Educ.* 61 (1–2), 89–99.
- Zeichner, K., Payne, K., Brayko, K., 2012. Democratizing knowledge in university teacher education through practice-based methods teaching and mediated field experience in schools and communities. *Harv. Educ. Rev.* 1–46.

Sensory literacies: the full sensorium in literacy learning

Kathy A. Mills^a and Beryl Exley^b, ^a Institute for Learning Sciences & Teacher Education, Brisbane Campus, Brisbane, QLD, Australia; and ^b School of Education and Professional Studies, Griffith University, Mount Gravatt, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Sensory literacies: senses, the body, and literacy practices	480
Privileging of vision in theories of communication	481
Embodied cognition and sensory studies: origins of sensory literacies	481
Use of the full sensorium in literacy and digital media practices	482
Implications of sensory literacies across the life course	483
Pedagogical challenges and possibilities	484
Conclusion	484
Acknowledgment	484
References	485
Further reading	486

Glossary

- Aurality** Relating to the sense of hearing
- Embodied cognition** Cognition is shaped by interactions involving the whole body
- Embodiment** The state of being corporeal, tangible, and embodied
- EFL** English as foreign language
- Haptics** Movement of the hands
- Kinesis** Movement of the body
- Ocularcentrism** The privileging of vision over other senses
- Olfaction** Relation to the sense of smell
- Senses** Faculties through which stimuli is perceived from outside or inside the body
- Sensorium** Human sensory apparatus or faculties as a whole
- Sensory orchestration** The senses working together
- Touch** To bring the hands and fingers into contact with someone or something

Sensory literacies: senses, the body, and literacy practices

There is a long history of literacy research that has yet to explore the fundamental basis for human thought and language—the senses, embodied activity, and real-world referents (Gibbs, 2005). A sensory literacies paradigm is a recent approach that foregrounds the role of the senses and embodiment in literacy and communication practices—social practices that vary across cultures, contexts, times, and technologies. The sensory literacies approach was coined and described in the volume “Literacy Theories for the Digital Age” (Mills, 2016). A sensory approach to literacies advances understandings of literacy as a social practice with a focus on the body and the sensorium. However, it does so without taking for granted the role of the mind and the technologies that support literacy practices. Both the mind and body are seen as integral to literacy practices, consistent with embodied cognition (Mills, 2016). Theorists of embodied cognition have shown that even when abstract mental processes are in operation, the mind draws on integral sensorimotor processes that operate in the background to orchestrate images, sounds, movements, and memories (Wilson, 2002). In other words, even the so-called “thinking” part of literacy practices is body based.

A sensory literacies approach follows a move toward sensory scholarship across many disciplines (Stoller, 1997), influenced by sensory studies (Howes, 1991, 2003; Howes and Classen, 2014), sensory ethnography (e.g., Pink, 2015), and other tributaries of thought on literacy and the body in education (e.g., Mills et al., 2013; Scollon and Scollon, 2003). The theoretical basis of the sensory literacies approach follows a sensorial turn that has emerged in many fields of research, including, but not limited to, cultural anthropology (e.g., Howes, 2014), neuroscience (e.g., Murray and Wallace, 2012), cultural geography (e.g., Boswell, 2017), marketing (e.g., Hulten, 2011), law (e.g., Sculthorpe, 2014), sensory methodologies (Pink, 2015), and architecture (Pallasmaa, 2012).

A sensory literacies approach concerns the embodied and emplaced relations between humans, materials, texts, and technologies in the social practices of literacy. For the most part, the senses in education have been seen within the ambit of psychology, while literacy education has been somewhat ambivalent in terms of the senses, concerned primarily with the teaching of language

as text (Majid and Levinson, 2011). While some multimodal theorists have examined the visual semiotics of texts for many decades (e.g., Kress and van Leeuwen, 2006), recent work has examined embodiment and the senses in text encoding and decoding across a range of modes and media. For example, Mills et al. (2013) explored the under-researched role of locomotion and kinesics in children's film making. This work drew the attention of Howes (2016), a sensory anthropologist, who acknowledged the usefulness of such applications to theorizing literacy practices. Theories of touch and haptics (e.g., Bezemer and Kress, 2014) are relatively emergent within a sensory turn in literacy research, in part influenced by the move to touchscreen technologies, and supported by seismic changes to the use of touch and haptics in textual practices in the 21st century (e.g., Crescenzi et al., 2014; Walsh and Simpson, 2014).

Few would disagree with the assertion that "without our bodies—our sensing abilities—we do not have a world" (Arola and Wysocki, 2012, 3). A sensory literacies approach has opened up different ways of teaching, researching, and theorizing literacy practices in education, highlighting the way in which literacy practices are embodied, multisensorial, and interconnected to sensorial and placed action that differs across cultures and social contexts. A sensory literacies approach similarly accounts for the role of the body, the senses, and the mind in relation to reading, writing, listening, viewing, and creating meaning in digital and non-digital ways.

Privileging of vision in theories of communication

Western epistemic beliefs have for some time prioritized empiricism, which is established on the central assumption that observable knowledge—what can be repeatedly seen and confirmed with the eyes—is truth. This privileging of what can be perceived through vision has not been exempt from language and literacy studies, where ocularcentrism—the partiality toward the vision over other senses—has relegated the so-called "lesser modes" and "lesser senses," such as taste (van Leeuwen, 1997), smell or olfaction (Wasikul et al., 2015), kinesis or movements of the lower limbs and feet (Mills and Dooley, 2019), and even sound (Ceraso, 2014) outside the central purview of language and literacy teaching and research. While written and spoken words feature prominently in literacy curricula, and to a lesser extent, images and gestures in texts and performances, the marginalization of the so-called lesser senses of bodily movement and locomotion (Ingold, 2011), can be understood as part of the broader influence of empiricism and ocularcentrism across a range of disciplines (Howes, 1991; Porteous, 1990).

Theorists of sensory studies have provided enlightening critiques of the visual and textual bias of the Western episteme, recognizing that our cultural bias toward vision becomes apparent when researchers examine how "life is lived in other cultural settings" (Howes, 1991, 3). There are numerous examples in non-Western cultures that ascribe varied meanings to sensory experiences that are richly illustrative of multisensorial and socially organized meanings that matter, and that extend well beyond vision (see, for example, Feld, 1991; Pandya, 1993). Classen (1999) writes about the Desana people of the Colombian rainforest whose social and cosmic understandings evolve from the sounds of bird calls, the color and scent of flowers, and the taste of fruit. For the Warao of Venezuela, herbal plants of similar appearance used in pharmacopoeia are identified by their smell, because relying on visual means alone may be fatal (Wilbert, 1987). The point is not so much to establish an *anti-visualist* approach to literacy, but rather, to develop an approach to language, literacies, and communication that is able to reflect upon its sensory assumptions and hierarchies. A sensory literacies approach is catalyzing new understandings of the role of the full sensorium in the social practices of literacy, and the specialized roles and interconnectedness of the senses and modalities that are mobilized in meaningful and embodied ways—ways that are constantly shifting across cultures and time.

The privileging of vision in learning and communicating has far-reaching negative consequences. Multisensory research from the cognitive sciences has shown that restricting the use of the senses to vision results in perceptual processing deficits, as well as reduced efficiency for memory, learning and communication (Shams and Kim, 2012). Certain facts and types of knowledge are well-suited to visual representation, but other kinds of knowledge about the world are better suited to multisensorial combinations of modes (Stein, 2012). For example, in response to questions from his young Fijian students about the use of mangrove sap in making tapa cloth, the classroom teacher talked about the inherent benefits of taking the children for a walk to the mangroves to "witness the reality of it" and extend their initial understandings (Carss et al., 2015, 35). Consequently, privileging ocularcentrism in literacy research and practice limits avenues for students to communicate varied kinds of knowledge (Klatzky and Lederman, 1988).

A sensory approach to literacies had drawn attention to the interrelationships among the senses that are vital to meaning making. Ingold (2000) argues that "... the eyes and the ears should not be understood as separate keyboards for the registration of sensation, but as organs of the body as a whole" (268). Sensory approaches attend to multiple senses and their interrelationships in meaning making, also referred to as sensory orchestration—the senses working together. Literacy as an embodied practice involves aurality, touch, sight, movement, taste, smell and other forms of bodily awareness that are still being explored in research, such as thermoception, a sense of temperature, and proprioception, the awareness of one's body in space.

Embodied cognition and sensory studies: origins of sensory literacies

The theory of sensory literacies involves a paradigm shift from traditional views of language and communication in cognitive psychology that emphasis the neural base of language (Gibbs, 2005), to the theory of embodied cognition that examines the grounding of language in kinesthetic experiences (Wilson, 2002). The view that cognitive, social, and practical enactment of literacy is body-based draws upon the growing recognition that human perception, thought, and language are fundamentally shaped by embodied experience (Haas and McGrath, 2018). Literacy practices do call upon mental activities, such as the act of writing, which

can be considered as a vehicle for human thought; yet our thoughts constitute an array of internalized sensorimotor experiences and functions, operating covertly to simulate or reflect upon external events using visual, auditory, and kinesthetic imagery and memories (Friend and Mills, 2021). In other words, literacy practices, including the mental and practical work of reading, writing, viewing, designing and so forth, involve embodied cognition. Literacy is not the achievement of the mind inside a disconnected body, nor the body performing mindlessly, but involves a synergistic interrelationship between mind and body (Mills, 2016). Likewise, the body is central to the practical enactment of literacy practices in terms of their corporeality. As Scollon and Scollon (2003, 45) argue, “Our bodies as humans cannot fail to anchor us in the real, physical world in which we are performing as social actors.” Teachers capitalize upon this, seeing children’s physical cues to storybook read alouds as a means by which to make assessments of their ability to comprehend the text (Luke, 1992).

Cognitive theorists distinguish between two important forms of embodied cognition—online and offline. These are not to be associated with being connected to the internet. Rather, online embodied cognition refers to mental activity in which the mind is operating while the sensing body simultaneously interacts with a physical or virtual environment. Offline embodied cognition is the use of sensorimotor processes to support abstract, mental tasks, but operating in a covert way (Wilson, 2002). What is remarkable about human thought and language is that these ostensibly abstract processes in offline embodied cognition are also body-based (Friend and Mills, 2021; Gibbs, 2005). This is because learners have previously internalized sensorimotor functions to reproduce memories of external events, images, movements, music, language, and other resources that were learned or perceived sensorially. A classic example is mental calculations, where operations that were previously supported by counting on one’s fingers become pictured in the mind’s eye. Sometimes this can be accompanied by subtle movements, like twitching, but once the sensorimotor activity has been fully internalized, the mind can run a mental simulation of the movement (Wilson, 2002).

This is significant because the senses are not only vital to the corporeal aspects of literacy, such as turning the pages of a book, gesturing with hands, or writing with a pen, but to the abstract facets of communication, such as the internalized sensorimotor simulation of smells, sights, events, and ideas about the world that are described and remixed as one composes a written text. In contemporary literacy environments, where communication is often digitally mediated, literacy practices require augmented forms of online (practical) and offline (abstract) sensorial and embodied practices. Similarly, listening is not just the practice of attending to audible words or sounds in order to make meaning of them. Ceraso (2014) explains how other parts of the body, such as the stomach, throat and legs, can be engaged during a “sonic encounter” (102).

Likewise, even many words in language have a sensory basis because things seen, heard, felt, tasted, smelled, and perceived form the physical basis of meanings that are later mapped onto abstract metaphors. Sweetser (1990), for example, showed how polysemous vocabulary, or words having many meanings, in Indo-European languages that are sensorial, experiential, and physical, become extended or mapped semantically to become metaphors that have non-physical, abstract, and intellectual meanings. For example, “keep my ear to the ground, going in with your eyes open; dip your toe in; stick your neck out.” There are many examples in language etymology where bodily experiences undergird semantic meaning making. One can also detect the dominance of vision in the wide usage of the word “seeing” as “knowing” in English, which is also the case in Russian (Andrews, 1995; Gibbs, 2005). Examples in English are “I see what you mean; it’s my observation that ...; I’ve watched this unfold; From my viewpoint; Look at it this way; It’s in our line of sight.” The sense of sight is used in manifold ways as a metaphor for expressing and communicating cognitive understanding. Thus, embodiment and the senses are not only observably central to literacy practices, they are also central to metaphoric reasoning and many instances of semantic interpretation in the very etymology and evolution of language.

Use of the full sensorium in literacy and digital media practices

Literacy practices, digital and non-digital, involve the full sensorium—the human sensory apparatus or faculties as a whole. Literacy practices constitute the work of the mind and body in its entirety, involving interrelationships between different modalities that do not merely overlap, but which operate together as a synergetic system in social action. For example, creating a video can involve an array of embodied creative processes, including first drawing or visualizing a storyboard or sequence of still images, walking, and moving to different locations with the camera to gain different viewpoints, and filming scenes from different bodily positions. Dramatic performances may be included in the film, while scenes may involve narrating with the voice. Editing the film is typically enacted while seated at a computer to juxtapose audio, still or moving images, transitions, subtitles and other compositional elements (Mills et al., 2013). A broadened repertoire of digital technologies to mediate textual practices across modes, likewise, accompanies different forms of material and bodily engagement in literacy practices.

At the same time, the integration of the senses in literacy practices is neither new nor unique to practices mediated by digital technologies. For example, Indigenous cultures have long valued the integration of multiple ways of knowing about the world. Constance Classen concluded in her observation of the senses in Indigenous cultures, “Each of the senses has a vital role to play in the acquisition of knowledge about the world” (Classen, 1999, 269). Likewise, in Andean culture, the textile technology of the *quipu*—a set of knotted, colored cords suspended from string—communicated information three-dimensionally, with each color, position and size of the knots encoding different meanings (Howes, 2016). The creation of *quipu* was simultaneously an art form, a mnemonic device and a form of Inka writing, while each *quipu* maker had a signature style or shorthand, making the accurate decoding and interpretation of *quipu* from a vanished civilization near impossible. Ascher (2002, 133) argues that *quipu* was multi-sensorial, “simultaneously tactile and visual, and probably more.” The point is, that in older civilizations, knowing, and communicating about the world has involved multiple senses.

However, the increasing role of digital technologies in literacy and communication is unprecedented (Mills, 2010), and with these transformations have arisen new sensory affordances of technologies involving novel uses of haptics, movement, and other modes that are having a monumental impact on recreation, social life, and the workforce. For example, young people imitate rhythmic gestures, movements, and facial expressions with lip-syncing in short video sharing and social media platforms (Simsek et al., 2018). Virtual reality headsets and hand-controlled sensors open up a wider range of applications for communicating through kinesthesia—the sensory awareness of the position and movement of the body, particularly of the hands and arms (Mills and Brown, 2021). When used with programs such as Google Tilt Brush, creating three-dimensional, virtual artwork can be created in an immersive, whole-body, multisensorial environment (Mills et al., 2022).

Since time immemorial, literacy practices have been multisensorial, and at different times, a narrower range of sensory practices have been privileged over all others. Compare the Gutenberg era beginning in 1450 to the wide-ranging digital texts options of the current period. The convergence of electronic media with mobile technologies has seen the rise of interactive literary apps and other online forms of online reading on the internet and social web, with changing materially demanding different sensorial practices (Mills et al., 2018). Researchers have drawn attention to new typologies of interactive touch in e-book reading, such as the use of extratext functions (e.g., pressing an audio button to listen), and intratext interactivity, such as enacting part of the action of the plot or characters through tapping, swiping, shaking, or drawing thumb and fingers together or apart in pinching movements (Zhao and Unsworth, 2017). These digital versions of books, or other long form reading material, such as electronic magazines, journals, blogs, and news articles, are not rendering conventional printed book reading obsolete, rather, they contribute to the broadened variety of sensorial engagements for new reading practices.

Similarly, the sensoriality and materiality of writing practices are swiftly changing, even though writing has always been an embodied practice (Mills, 2016). Writing has been aptly called “language by hand” by educational psychologists to draw attention to both the serious cognitive work and the fine-motor skills (Berninger et al., 2009). Haas and McGrath (2018) explain how writing movements trigger activation of the brain, regardless of whether an individual actually provided a motor response. The act of writing by hand brings modes of visual and proprioception together in sensorial orchestration. It has long been established that sensory modalities and the creation of neural structures support the ongoing development of literacies (Haas and McGrath, 2018; James and Gauthier, 2006; Matsuo et al., 2003). In addition to the complex and difficult-to-observe neurological relations between a writer’s mind and body-based memories (e.g., Wilson, 2002), few would dispute that writing is embodied in observable, material, and corporeal ways as “writers pick up and chew their pencils, they rest their hands on keyboards, they hunch over manuscripts ... push keys or click buttons” (Haas, 1996, 226). Writers cannot produce text other than through the work of their bodies, while at the same time sensorial processes of writing involve more than the corporal interaction of minds, bodies, hands, mouths, eyes, pens, papers, voices and computers, employing a plethora of cultural and social practices that vary in nuanced ways across writing communities, from graffiti artists who inscribe walls, to children who write using a stylus and touchscreen at school (Mills, 2016). Theorists have similarly drawn attention to the fact that the embodiment of writing is different for able bodies than those whose bodies may not conform to ableist norms, challenging assumptions about the writing body (e.g., Dolmage, 2012).

Literacy practices are continually shifting, and with these changes come modified forms of embodied meaning. Each new social practice, such as using a virtual reality headset and sensors for the first time, involves learning new forms of embodied engagement in meaning making. Specialized competency can involve years of training. The enactment of literacy practices, like many social acts across all spheres of life, constitutes an array of embodied knowledge that may be acquired from birth, (e.g., early vocalizations, speech) knowledge that is taught and practiced in the socialization of cultural communities and societies, and which may become systematized over time. Likewise, the embodied nature of familiar literacy practices can become taken-for-granted or backgrounded with familiar and repeated use as socially learned repertoires of action (Mills, 2016; Mauss, 1979). In short, a sensorial approach to the integration of the senses moves away from disembodied views of literacy practices, while also resisting the hegemony of sight.

Implications of sensory literacies across the life course

The physical transformations of the senses that occur over one’s lifespan follow a progression, beginning in the womb. Medical scientists have documented reactions of the fetus in utero and pre-mature babies to confirm that humans of gestational age develop their sense of taste first, transmitting information from their taste buds to the central nervous system during the last trimester of prenatal development (Forestell, 2017). The fetus is receptive to and responsive to auditory and tactile senses, even if they are not aware of these perceptions (Bradley and Mistretta, 1975). The fetus is also sensitive to touch, sound, and smell. Visual receptors do not receive adequate stimulation until after birth (Bradley and Mistretta, 1975).

During gestation and after birth, the human brain continually integrates information from the difference senses to provide a “unified perception of the world” (King, 2007, 3), a process that is essential for directing attention and controlling movement of the body (Denervaud et al., 2020). Although the human brain has the capacity to selectively decrease amplification of some senses to concentrate on other senses (Petrini et al., 2020), the success of this undertaking is dependent on an individual’s capacity to bind together the most appropriate multisensory signals (King, 2007).

The specific development of the senses diverges in important ways in response to cognitive maturation, and environmental and contextual factors. For example, infants have a pronounced sense of olfaction, using smell to locate food sources (such as the mother’s breast), and being able to recognize the smell of their mother. An adaptive behavior study undertaken by Cavazzana et al. (2016) found that by age 5, young children were able to mediate the affective strength of olfaction with social information. Environmental factors responsible for different rates of multisensory development include diet, nutrition, and healthcare (Crafter

et al., 2019), as well as contextual factors, such as one's socio-cultural experiences. This means that the repertoire of sensory literacies evolves in particular and context dependent ways, coming to the fore in different combinations for different reasons over one's lifetime. For example, experiences during infancy shape the way the visual and auditory inputs interact to determine the responses of the neurons in later life (Wallace and Stein, 2007), and thus, the repertoire of practices an individual has available for a range of literacy encounters.

In old age, the adage "use it or lose it" takes hold, drawing attention to the need to maintain the use of the senses for communicating through ongoing sub/conscious practice. The rate of deterioration of each sense is not uniform. For example, when an additional language is acquired in adolescent or adulthood, passive recognition of vocabulary extends over relatively long periods of time. In contrast, for the same individual, "without consistent use, speaking fluency decays rapidly" (Hertzog, 2009, 162).

Pedagogical challenges and possibilities

The use of the full sensorium in communication is crucial for supporting and strengthening an individual's social competence, formal learning, and vocational expertise (Petrini et al., 2020). Growth of capabilities and competence with the senses in literacy practices is not tied exclusively to biological processes, but is also dependent upon a dynamic process of co-construction with others in their physical and virtual environment.

The need for a wide repertoire of sensorial ways of perceiving and communicating is important in the early years, as it is in the years of schooling and in vocational pursuits (Denervaud et al., 2020). Some Australian scholars have documented the reduction of multisensory learning environments in early learning classrooms due to the pressures of national assessments in the early years of schooling (Somerwil et al., 2020). Teachers working with diverse student groups, such as those with differently abled bodies, are challenged to think beyond conventional teaching styles. Mashimo (2019) reports on the enactment of a multisensorial pedagogy that involved "hands-on educational workshops" instead of the more traditional vision-centered approach to teaching art history. His work documents the challenges of this adjustment for the sighted students in his classes. Similarly, a recent study exploring the use of a sensory approach to learning and teaching in English as Foreign Language (EFL) lessons for Pakistani students documented significant growth in the development of students' speaking skills (Ghani et al., 2015).

Conclusion

While cognitive science has examined the neural basis for human language functioning, it is prudent to also consider research on language and cognition, giving due acknowledgment to the role of the senses and the body, including everyday kinesthetic experiences. Gibbs (2005, 159) has argued that ignoring this thinking "has seriously undermined scientific understanding of the relations between mind and body, and more specifically, linguistic meaning, communication, and embodiment." Literacy practices, whether involving reading, composing texts, performing, viewing, listening, or combinations of these and other capacities involves the corporeality of one's body that is situated in a physical, social, and cultural environment. For example, even the most attentive listener in a seemingly quiet audience will make subtle eye movements, postural shifts, and other fine-motor responses that often occur subconsciously, and which may be interpreted by the speaker. Conversely, the absence of a bodily response, such as a blank stare from a participant in a conversation, can also speak volumes (Gibbs, 2005).

However, perhaps more critically, all human thought involves previously internalized sensorimotor functions to reproduce memories of concepts and external happening, words, images, tastes, textures, smells, sounds, songs and ensembles of other resources that were learned or perceived sensorially. Literacy practices require the attention of a knowing mind, and a perceiving and sensing body. Literacies can be conceived as sets of transposable dispositions that are consciously or unconsciously acquired through one's socialization over the life course, embodied in one's mind and muscle memory in often taken-for-granted ways, and which are not independent of the senses (Mills, 2016). Certainly, vision plays a central role in many literacies, but not inherently so, and more often, multiple senses operate together, some foregrounded, others backgrounded.

In sum, this chapter has challenged the belief in certain fields of Western philosophy, cognitive science, linguistics, and literacy theory, that language and meaning making is a predominantly abstract, psychological process divorced from the body, and likewise, the understanding that the role of the senses and the body is merely backgrounded in language and communication. In contrast, we have demonstrated here through evidence from embodied cognition and sensory studies that human bodily and sensorial experiences are the fundamental basis for thought and language, which occurs alongside cognition as corporeal bodies interact with materials and others in tangible, cultural, and social worlds. At a time in history when "standard English" literacy valued in schooling has, for the most part been reduced, taught and tested in global and national standardized tests as an atomistic set of skills, a sensory literacies perspective draws attention to the way in which language and literacies are not best understood and researched as disembodied and purely abstract activities of the mind, but as intimately amalgamated with embodied action.

Acknowledgment

This research is supported by an Australian Research Council Australian Future Fellowship (FT180100009) funded by the Australian Government. The views expressed herein are those of the authors and are not necessarily those of the Australian Research Council.

References

- Andrews, E., 1995. Seeing is believing: visual categories in the Russian lexicon. In: Cortini-Morava, E., Goldberg, B. (Eds.), *Meaning as Explanation*. Mouton de Gruyter, Berlin, pp. 637–677.
- Arola, K.L., Wysocki, A., 2012. *Composing (Media) = Composing (Embodiment)*. Utah State University Press, Logan, Utah.
- Ascher, R., 2002. Inka writing. In: Quilter, J., Urton, G. (Eds.), *Narrative Threads: Accounting and Reaccounting in Andean Khipu*. University of Texas Press, Austin, pp. 103–115.
- Berninger, V.W., Richards, T., Abbott, R., 2009. The role of the hand in written idea expression. In: Alamargot, D., Bouchard, J., Lambert, E., Millogo, V., Beaudet, C. (Eds.), *Proceedings of the International Conference de la France au Quebec: l'Ecriture dans tous ses états*. Poitiers, France.
- Bezemer, J., Kress, G., 2014. Touch: a resource for making meaning. *Aust. J. Lang. Literacy* 37 (2), 77–85.
- Boswell, R., 2017. Sensuous stories in the Indian Ocean Islands. *Senses Soc.* 12 (2), 193–208.
- Bradley, R.M., Mistretta, C.M., 1975. Fetal sensory receptors. *Physiol. Rev.* 55 (3), 352–382.
- Carss, W., Tamata, A., Exley, B., 2015. Catering for cultural and linguistic diversity: using teacher created information texts. *Literacy Learn. Middle Years* 23 (2), 29–39.
- Cavazzana, A., Wesarg, C., Parish-Morris, J., Lundstrom, J.N., Parma, V., 2016. When pre-schoolers follow their eyes and older children follow their noses: visuo-olfactory social affective matching in early childhood. *Dev. Sci.* 21, 1. <https://doi.org/10.1111/desc.12507>.
- Ceraso, S., 2014. (Re)educating the senses: multimodal listening, bodily learning, and the composition of sonic experiences. *Coll. Engl.* 77 (2), 102–123.
- Classen, C., 1999. Other ways to wisdom: learning through the senses across cultures. *Int. Rev. Educ.* 45 (3/4), 269–280.
- Crafter, S., Maunder, R., Solusby, L., 2019. *Developmental Transitions: Exploring Stability and Change Through Lifespan*. Taylor & Francis, Milton.
- Crescenzi, L., Jewitt, C., Price, S., 2014. The role of touch in preschool children's learning using iPad versus paper interaction. *Aust. J. Lang. Literacy* 37 (2), 86–95.
- Denervaud, S., Gentaz, E., Matusz, P.J., Murray, M.M., 2020. Multisensory gains in simple detection predict global cognition in schoolchildren. *Sci. Rep.* 10, 1394.
- Dolmage, J., 2012. Writing against normal: navigating a corporeal turn. In: Arola, K.L., Wysocki, A. (Eds.), *Composing Media Composing Education*. Utah State University Press, Logan, UT, pp. 110–126.
- Feld, S., 1991. Sound as symbolic system: the Kaluli drum. In: Howes, D. (Ed.), *The Varieties of Sensory Experience: A Sourcebook in the Anthropology of the Senses*. University of Toronto Press, Toronto, pp. 79–99.
- Forestell, C.A., 2017. Flavor perception and preference development in human infants. *Ann. Nutr. Metabol.* 70 (3), 17–25.
- Friend, L., Mills, K.A., 2021. Towards a typology of touch in multisensory makerspaces. *Learn. Media Technol.* <https://doi.org/10.1080/17439884.2021.1928695>.
- Ghani, M., Nawaz, S., Asif, M., 2015. Developing speaking skills through sensory activities: an empirical study conducted on Pakistani EFL learners. *J. Educ. Res.* 18 (2), 84–95.
- Gibbs, R.W., 2005. *Embodiment and Cognitive Science*. Cambridge University Press, Cambridge, UK.
- Haas, C., McGrath, M., 2018. Embodiment and literacy in a digital age. In: Mills, K.A., Stornaïoulo, A., Smith, A., Pandya, J.Z. (Eds.), *Writing, Literacies, and Education in Digital Cultures*. Routledge, New York, NY, pp. 125–135.
- Haas, C., 1996. *Writing Technology: Studies on the Materiality of Literacy*. Lawrence Erlbaum Associates, New Jersey.
- Hertzog, C., 2009. Use it or lose it: an old hypothesis, new evidence, and an ongoing controversy. In: Bosworth, H.B., Hertzog, C. (Eds.), *Aging and Cognition: Research Methodologies and Empirical Advances*. American Psychological Association, pp. 161–179.
- Howes, D., Classen, C., 2014. *Ways of Sensing: Understanding the Senses in Society*. Routledge, London, UK.
- Howes, D. (Ed.), 1991. *The Varieties of Sensory Experience: A Sourcebook in the Anthropology of the Senses*. University of Toronto Press, Toronto, ON.
- Howes, D., 2003. *Sensual Relations: Engaging the Senses in Culture and Social Theory*. The University of Michigan press, Ann Arbor, MI.
- Howes, D., 2014. Anthropology and multimodality: the conjugation of the senses. In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*, second ed. Routledge, London, UK, pp. 225–236.
- Howes, D., 2016. Foreword: the sensory turn in literacy theory and practice. In: Mills, K.A. (Ed.), *Literacy Theories for the Digital Age: Social, Critical, Multimodal, Spatial, Material and Sensory Lenses*. Multilingual Matters, Bristol, UK.
- Hulten, B., 2011. Sensory marketing: the multi-sensory brand-experience concept. *Eur. Bus. Rev.* 23 (3), 256–273.
- Ingold, T., 2000. *The Perception of the Environment: Essays in Livelihood, Dwelling and Skill*. Routledge, London.
- Ingold, T., 2011. *Being Alive: Essays on Movement, Knowledge and Description*. Routledge, Abingdon, UK.
- James, K.H., Gauthier, I., 2006. Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia* 44, 2937–2949.
- King, A.J., 2007. Coordinating different sensory inputs during development. *J. Neurophysiol.* 97 (1), 3–4.
- Klatzky, R.L., Lederman, S.J., 1988. The intelligent hand. *Psychol. Learn. Motiv.* 21, 121–151.
- Kress, G., van Leeuwen, T., 2006. *Reading Images: The Grammar of Visual Design*, second ed. Routledge, London.
- Luke, A., 1992. The body literate. *Ling. Educ.* 4, 107–129.
- Majid, A., Levinson, S.C., 2011. The senses in language and culture. *Senses Soc.* 6 (1), 5–18.
- Mashimo, Y., 2019. Multi-sensorial pedagogy for Art history education: integrating the collective wisdom of people who are blind and how low vision to reconsider conventional academic norms. *J. Lit. Cult. Disabil. Stud.* 13 (3), 305–322.
- Matsuo, K., Kato, C., Okada, T., Morita, T., Glover, G., Nakai, T., 2003. Finger movements lighten neural loads in the recognition of ideographic characters. *Cognit. Brain Res.* 17, 263–272.
- Mauss, M., 1979. *Techniques of the body* (B. Brewster translator). In: *Sociology and Psychology: Essays by Marcel Mauss*. Routledge and Kegan Paul, London, pp. 95–123.
- Mills, K.A., Dooley, J., 2019. Sensory ways to indigenous multimodal literacies: hands and feet tell the story. In: Rennie, J., Harper, H. (Eds.), *Literacy Education and Indigenous Australians: Theory, Education and Practice*. Springer, Berlin, Germany, pp. 33–50.
- Mills, K.A., Brown, A., 2021. Immersive virtual reality (VR) for digital media making: transmediation is key. *Learn. Media Technol.* 1–22. <https://doi.org/10.1080/17439884.2021.1952428>.
- Mills, K.A., Comber, B., Kelly, P., 2013. Sensing place: embodiment, sensoriality, kinesis, and children behind the camera. *Engl. Teach. Pract. Critiq.* 12 (2), 11–27.
- Mills, K.A., Unsworth, L., Exley, B., 2018. Sensory literacies, the body and digital media. In: Mills, K.A., Stornaïoulo, A., Smith, A., Pandya, J.Z. (Eds.), *Writing, Literacies, and Education in Digital Cultures*. Routledge, New York, NY, pp. 26–36.
- Mills, K.A., Unsworth, L., Scholes, L., 2022. *Literacy for Digital Futures: Mind, Body, Text*. Routledge.
- Mills, K.A., 2010. A review of the digital turn in the New Literacy Studies. *Rev. Educ. Res.* 80 (2), 246–271.
- Mills, K.A., 2016. *Literacy Theories for the Digital Age: Social, Critical, Multimodal, Spatial, Material and Sensory Lenses*. Multilingual Matters, Bristol, UK.
- Murray, M.M., Wallace, M.T. (Eds.), 2012. *The Neural Bases of Multisensory Processes*. CRC Press, Taylor and Francis, London, UK.
- Pallasmaa, J., 2012. *The Eyes of the Skin: Architecture and the Senses*. John Wiley and Sons, West Sussex.
- Pandya, V., 1993. *Above the Forest: A Study of Andamanese Ethnoanemology, Cosmology, and the Power of Ritual*. Oxford University Press.
- Petrini, K., Denis, G., Love, S.A., Nardini, M., 2020. Combining the senses: the role of experience- and task-dependent mechanisms in the development of audiovisual simultaneity perception. *J. Exp. Psychol. Hum. Percept. Perform.* <https://doi.org/10.1037/xhp0000827>.
- Pink, S., 2015. *Doing Sensory Ethnography*, second ed. SAGE, London.
- Porteous, D., 1990. *Landscapes of the Mind: World of Sense and Metaphor*. University of Toronto Press, Toronto, ON.
- Scollon, R., Scollon, S.W., 2003. *Discourses in Place: Language in the Material World*. Routledge, London.
- Sculthorpe, A., 2014. *A Courtroom With a View: When a Justice System Senses for Itself*. Concordia University, Montreal, Canada.

- Shams, L., Kim, R., 2012. Cross-modal facilitation of unisensory learning. In: Stein, B.E. (Ed.), *The New Handbook of Multisensory Processing*. MIT Press, Cambridge, MA, pp. 475–482.
- Simsek, B., Abidin, C., Brown, M.L., 2018. Musically and microcelebrity among girls. In: Abidin, C., Brown, M.L. (Eds.), *Microcelebrity Around the Globe*. Emerald Publishing Limited, pp. 47–56.
- Somerwil, T., Klieve, H., Exley, B., 2020. Preschool educators' readiness to promote children's emotional competence. *Asia Pac. J. Res. Early Childhood Educ.* 14 (2), 135–158.
- Stein, E. (Ed.), 2012. *The New Handbook of Multisensory Processes*. MIT Press, Cambridge, MA.
- Stoller, P., 1997. *Sensuous Scholarship*. University of Pennsylvania Press, Philadelphia, PA.
- Sweetser, E., 1990. From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure. Cambridge University Press, Cambridge, UK.
- van Leeuwen, T., 1997. Taste in the framework of a semiotics of materiality. In: Piroelle, A. (Ed.), *La Representation Sociale Du Gout*. Coloradoc, Dijon, pp. 149–162.
- Wallace, M.T., Stein, B.E., 2007. Early experiences determines how the senses will interact. *J. Neurophysiol.* 97 (1), 921–926.
- Walsh, M., Simpson, A., 2014. Exploring literacies through touch pad technologies: the dynamic materiality of modal interactions. *Aust. J. Lang. Literacy* 37 (2), 98–106.
- Waskul, D., Vannini, P., Wilson, J., 2015. The aroma of recollection: olfaction, nostalgia, and the shaping of the sensuous self. *Senses Soc.* 4 (1), 5–22.
- Wilbert, W., 1987. The pneumatic theory of female Warao herbalists. *Soc. Sci. Med.* 25 (10), 1139–1146.
- Wilson, M., 2002. Six views of embodied cognition. *Psychon. Bull. Rev.* 9 (4), 625–636.
- Zhao, S., Unsworth, L., 2017. Touch design and narrative interpretation: a social semiotic approach to picture book apps. In: Kucirkova, N., Falloon, G. (Eds.), *Apps and Pre and Primary-School-Aged Children: International Perspectives and Empirical Evidence*. Routledge, London, UK, pp. 89–102.

Further reading

- Gershon, W.S., Appelbaum, P.M., 2020. *Sonic Studies in Educational Foundations: Echoes, Reverberations, Silences, Noise*. Routledge Research in Education. Routledge, London.
- Gibbs, R.W., 2005. *Embodiment and Cognitive Science*. Cambridge University Press, Cambridge, UK.
- Haas, C., 1996. *Writing Technology: Studies on the Materiality of Literacy*. Lawrence Erlbaum Associates, New Jersey.
- Howes, D., Classen, C., 2014. *Ways of Sensing: Understanding the Senses in Society*. Routledge, London, UK.
- Howes, D., 2003. *Sensual Relations: Engaging the Senses in Culture and Social Theory*. The University of Michigan press, Ann Arbor, MI.
- Ingold, T., 2011. *Being Alive: Essays on Movement, Knowledge and Description*. Routledge, Abingdon, UK.
- Mills, K.A., 2016. Literacy Theories for the Digital Age: Social, Critical, Multimodal, Spatial, Material and Sensory Lenses. *Multilingual Matters*, Bristol, UK.
- Mills, K.A., Unsworth, L., Scholes, L., 2022. *Literacy for Digital Futures: Mind, Body, Text*. Routledge, UK.
- Sefton-Green, J., Erstad, O., 2017. Researching "learning lives"—a new agenda for learning, media and technology. *Learn. Media Technol.* 42 (2), 246–250.
- van Leeuwen, T., 1997. Taste in the framework of a semiotics of materiality. In: Piroelle, A. (Ed.), *La Representation Sociale Du Gout*. Coloradoc, Dijon, pp. 149–162.

Expansive, dynamic, and critical new literacies: widening views of digital and multimodal learning

Elena Forzani^a and Jill Castek^b, ^aLiteracy Education, Boston University, Wheelock College of Education & Human Development, Boston, MA, United States; and ^bTeaching, Learning and Sociocultural Studies, University of Arizona, College of Education, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Examining the new literacies of online of reading and research	488
Characteristics of an internet environment: affordances and constraints	488
Characteristic 1: constancy of change	489
Characteristic 2: Assumption of multiliteracies and online mobility, creation, and networking	489
Characteristic 3: easier, cheaper, and faster access to information	489
Characteristic 4: fluidity across meaning-making environments	490
Shifting social values and opportunities for social change	490
Four 21st century social values	490
Participation & collaboration	490
Criticality, responsibility, and social change	490
Multiplicity, hybridity, and flexibility	491
Creativity, innovation, and problem-solving	491
New literacies social values reflected in educational standards around the world	491
Under-realized potentials	491
Criticality and justice	491
Self-criticality	492
Digital equity	492
Awareness of data practices	492
Interconnectedness of literacies online	493
A call to action — looking toward the future	493
Avenues for research	493
Research centered on criticality and social change	493
Research centered on digital identity and citizenship	494
Practical considerations in our lives and in everyday contexts	494
Teaching and learning	494
References	495
Relevant website	496

Since the late 1990s and early 2000s, connected, digital technologies that utilize the internet have expanded and changed the ways we engage with the world (Greenhow et al., 2009). Perhaps one of the most prominent ways has been to change our literacy practices. This change has had substantial consequences for K-12 learning, which requires the use of literacy practices across all disciplines (Leu et al., 2013). The increasing growth of digital technologies has given rise to multimodal methods of communicating, has enabled and demanded global participation and collaboration, and has expanded the spaces that we inhabit as we connect and communicate. These shifts are affecting the ways we read, write, interpret, and create. New literacies are needed to navigate the internet, and they require the nimble use of skills, strategies, and mindsets to navigate online in everyday contexts (Jacobs and Castek, 2018, 2022). Moreover, new literacies are necessary when locating online information, networking, or using texts, tools, and interfaces to accomplish personal and professional goals.

Literacy is central to all disciplines, and changes in literacy practices have implications for the ways we design and engage in instruction, teaching, and learning across all fields of study. In this chapter, we consider how reading and writing change on the internet and the implications of such changes. We also discuss some of the ways that reading and writing—or, interpretation and creation (Forzani and Ly, 2022)—shifts online. To do so, we first begin with an overview of theoretical work that has impacted the ways we think about new literacies with the internet. We then discuss characteristics inherent in an internet environment and the affordances, constraints, and implications associated with these characteristics. Next, we examine how the internet has shifted our social values and the ways in which these changing social values are now reflected in curriculums around the world. Finally, we consider under-realized potentials in the field and avenues for future research.

Examining the new literacies of online of reading and research

The online world is social, and the definition of literacy is evolving due to the prevalence of digital texts, tools, and applications that are ubiquitous within daily interactions (Leu et al., 2013). Online networks, search engines, friends, and family are places people turn to when seeking information or attempting to solve a problem in everyday life (Jacobs and Castek, 2018, 2022). Information seeking involves deep levels of critical thinking when examining different websites and resources, determining which ones are most useful, and deciding which information within any given resource is relevant to the purpose (Goldman et al., 2013). In addition to thinking critically about information, learners also need to rapidly shift perspective and adapt their skills to complex situations they have not encountered before. Kiili et al. (2018a,b) and Coiro et al. (2019) frame online reading as an activity that involves curiosity, inquiry, and learning how to learn within ever-changing digital environments.

The digital world introduces greater complexity of meaning-making, as there are now more texts, text types, tools, modalities, creators, and a greater range of information quality online (Coiro, 2020). Such complexity can be both supportive and challenging. On one hand, this variety of literacies allows people to have some choice and control over the kinds of literacies they use and when and where they might “enter into” a literacy activity. For example, someone might choose to learn about news through videos if they cannot read in a particular language. In this sense, multiple modes and text types can be more engaging and accessible. People now have to learn not just how to interpret ideas within a single literacy but also how to use combinations of different literacies across multiple, rather than single, texts. At the same time, people are required to learn how to create using many different modes, text types, environments, contexts, and tools resulting in more complex and more challenging meaning-making processes (Kress, 2003, Chapter 4; Coiro, 2020).

A dual level theory frames two levels of new literacies: lowercase (*new literacies*) and uppercase (*New Literacies*). The term lowercase new literacies refers to a specific literacy area or technology, such as the literacy practices used within social media platforms or even within a particular social media platform such as Twitter. Lowercase new literacies also can refer to a particular area of study, such as data literacies, the literacies of social media, or networking literacies for navigating learning environments, career possibilities, or among multilingual individuals from around the globe. Understandings about particular lowercase new literacies can inform our understandings of other lowercase new literacies. To capture such similarities across lowercase new literacies, Leu et al. (2013) use the term uppercase New Literacies to refer to the commonalities across lowercase new literacies. Importantly, they describe eight central principles of uppercase New Literacies theory that capture current, common patterns (Leu et al., 2013, p. 1158). We summarize these to include:

1. The Internet is this generation’s most important and widely used technology for literacy and learning within our global community.
2. People need new practices to be able to understand and use the texts and tools associated with the internet.
3. New literacies are deictic, meaning that meaning-making with each new literacy is dependent upon the context in which a given technology is used.
4. New literacies are multiple, multimodal, and multifaceted, making them more complex in some ways compared to traditional literacies.
5. Critical literacies are central to new literacies practices.
6. New forms of knowledge, and new kinds of literacy strategies, are needed when using new literacies.
7. New Literacies require new social practices.
8. Teachers’ roles are somewhat different, in new literacies learning contexts.

Characteristics of an internet environment: affordances and constraints

Leu et al. (2013) argue that the internet is the “defining technology” of our time (p. 1158). Today, the internet plays a central role in our personal and working lives throughout our lifespans, from the time we are very young children to when we are the oldest of adults (Anderson, 2019; Rideout and Robb, 2019, 2020). The internet and other digital technologies have changed the ways we engage with literacy practices because new technologies offer different *affordances* and *constraints*, for both theory and practice. Such affordances and constraints have implications for the ways we teach students to take up different literacy practices.

Here, we use the term, *affordance*, to refer to a possibility for action posed by objects in the world; similarly, we use the term, *constraint*, to refer to a limitation for action (Gibson, 1977; Hammond, 2010; Ryder and Wilson, 1996). An *affordance* extends our capabilities as human beings, and a *constraint* restricts such capabilities. In this sense, technologies such as the internet offer numerous affordances. We also can view features, or characteristics, of the internet as affordances which themselves transform our capabilities in specific ways. Yet, affordances offered by one technology might also carry certain constraints. Such affordances and constraints have led to changes in our literacy practices and, as we argue here and elsewhere, in our social values as a result (Forzani et al., 2020).

Each new technology has different characteristics that introduce both affordances and constraints; thus, as new technologies emerge, our literacy practices change as we work to leverage available affordances. Here, we describe four central characteristics

of the internet as a literacy space, particularly as compared to more traditional, offline paper-based literacies. In so doing, we also discuss the affordances and constraints that arise from each characteristic. These characteristics include: 1) constancy of change; 2) assumption of multiliteracies; 3) easier, cheaper, and faster information access; and 4) fluidity across meaning-making environments.

Characteristic 1: constancy of change

One important characteristic of the internet is the constancy of change inherent in this literacy space (Leu et al., 2013). As new needs arise in our environments, we can address them with new tools. Digital technologies, and especially the internet, allow us to do this quickly and in real time. As a result, the focus in learning and work contexts shifts from learning particular skills or technologies to learning how to flexibly adapt to new technologies.

This constancy of change, however, creates an important constraint for theory. As Leu et al. (2013) point out, change poses a problem to the theorizing of new literacies, as it is hard to develop theory, which is relatively stable across time, around something that is not stable across time. To address this issue, Leu et al. (2013) pose the dual-level theory of new literacies, described above, that takes into account this constancy of change as one element that is stable across time. Such a dual-level theory allows us to consider what might be stable or common across time and across different lowercase new literacies while also being flexible enough to accommodate change as new literacies emerge.

Characteristic 2. Assumption of multiliteracies and online mobility, creation, and networking

Another fundamental characteristic of an internet environment is that multiliteracies are the norm rather than the exception. The term, multiliteracies, in this context refers to the full range of ways to share and create meaning across different cultural, social, or domain-specific contexts. These forms of meaning-making are more flexible and multimodal than paper-based environments, where it is harder to use multiple modes, texts, and activities, and with multiple people, at the same time. Digital environments, however, are conducive to multiliteracies in the broadest sense. In reflecting on multiliteracies, we draw on the New London Group (1996), Kress (2003, Chapter 4), and Coiro (2020) when referring to multiple modes, text types, number of texts, activities, environments, contexts, and people. Digital technologies make it easier and faster to use multiple modes of communication (e.g., linguistic, visual, aural, spatial, gestural) at once. Additionally, different people around the world are able to make meaning across modes and contexts giving broader access to meaning-making practices to a greater number of people. Such multimodal affordances have also resulted in an important focus on semiotics rather than on reading and writing per se (Kress, 2003, Chapter 4). In a narrow sense, the terms “reading” and “writing” might refer to interpretation and creation that occurs with linguistic text. However, given that different modes use different forms of text (visual, aural, spatial, gestural as well), the terms “reading” and “writing” are no longer sufficient. Therefore, here, we use the terms “interpretation” and “creation” to capture the meaning-making processes that occur across different modes and when multiliteracies are centered (Forzani and Ly, 2022).

One of the affordances of multiliteracies is that there are more access points and meaning-making opportunities for more people across potentially common modal lines. For example, visual symbols and video can be interpreted by nearly anyone across the globe, making it easier for people who might not share the same language to communicate. Although visual signs and symbols are culturally situated and might be interpreted differently by people from different cultural communities, they are also more readily available and interpretable than alphabetic or other linguistic signs.

Another affordance of such multiliteracies is the integration of both interpretation and creation processes. Rather than each of these two categories of practices occurring at separate times and in separate activities, in digital environments, these practices typically occur in tightly integrated iterations, and throughout the same activity.

Beyond affordances, one of the constraints of multiliteracies is the greater complexity of information and of the literacy practices one needs to use in order to make meaning from such information. Because there are now more texts, text types, and modes, the possibilities for different combinations, and the number of aspects within each activity, are greater, resulting in a potentially more complex meaning-making experience (Coiro, 2020). Despite this, however, multiliteracies also can offer support, since there are different possible ways into a text and into meaning-making for any given person.

Characteristic 3: easier, cheaper, and faster access to information

A third characteristic of the internet and of other digital environments is easier, cheaper, and faster information access. Today, 93% of Americans report that they use the internet (Perrin and Atske, 2021); worldwide, more than half of the global population is digitally connected (Johnson, 2021). People around the world can now access an immense amount of information at any time and place with the relatively fixed cost of an internet connection and a device with which to access it. In contrast to the cost of books and other paper materials that provide information, the internet is relatively easy to access, cheap, and fast. People do not need to spend time locating stores or libraries that have finite materials available and do not need to constantly purchase new materials as people want them or as they become available. Rather, people can, without leaving their homes, and in a matter of minutes, identify and locate needed information using a search engine.

One affordance of easy access is expanded, but sometimes problematic, participation and collaboration. Easier access could mean that more people can participate online. Yet, there are still people who do not have high-quality, fast, and affordable internet access. When access to online information is shared across different countries and cultures, the result could lead to greater global and multilingual interconnectedness. Particularly important is the notion that those who might not have been able to participate in certain spaces in the past, because they did not have access to them, may now have greater opportunity to access such spaces digitally. Although digital spaces often reflect societal structures, people may have greater opportunity to participate in online spaces that anyone can access.

One constraint of increased information access is the need for greater criticality. Because anyone may publish whatever they wish online, with few restrictions and vetting processes in place, there is a high degree of diversity of text types and quality on the internet (Britt and Gabrys, 2001; Rouet et al., 2011). Such information diversity can make it challenging for readers to make sense of texts and to determine who and what to believe. Thus, online contexts require readers to forefront critical reading, making the information vetting process largely the responsibility of individuals rather than of institutions.

Characteristic 4: fluidity across meaning-making environments

A fourth characteristic of digital spaces such as the internet is the ability to fluidly move in and out of participation contexts as one negotiates identities across in-person and digital environments. One example of this is the ability to use different modes of communication and the affordance of technology tools to quickly change text from one mode to another (e.g., the “talk to text” feature on a phone that allows a person to use the microphone to compose an email or text message, which then gets translated from aural to linguistic mode). Although the email gets composed by the user in one mode (aural), it gets sent to another user in a different mode (linguistic). A similar process might occur across environments. A user might have an experience in an in-person context and then move to the internet, just as when collaborators have a meeting in-person and then send a follow-up email. In our day-to-day lives, we often move across paper and digital environments in seamless ways.

Such fluidity of modes and environments has implications for learning contexts and for identity and digital presence (e.g., for how we position ourselves and our identities online; see Semington et al., 2017). Students, teachers, and learners of all ages are challenged to consider their identities across in-person and digital environments and to consider the extent to which such embodied identities align (or not).

Shifting social values and opportunities for social change

Digital affordances on the internet have both enabled and demanded a shift in what we position here and elsewhere (see Forzani et al., 2020; Leu et al., 2013) as social values. In particular, we outline four social values of the 21st century that have emerged from the rise of the internet, and particularly since the rise of Web 2.0 in or around the year 2000 (Greenhow et al., 2009).

Four 21st century social values

Participation & collaboration

The internet has enabled more people to both access and produce information. As a result, a greater number of people, across geographic, economic, and other divides, have been able to participate in society through the use of the internet (Jenkins, 2006). Not only can people access information through the internet, but they also can participate in a host of local and global social networks, such as Facebook, Instagram, and TikTok. Participation in organizations and other groups often involves some kind of connection through the internet. Moreover, people can communicate and collaborate across social divides. Although this does not always occur, and many internet contexts mirror existing social structures in the physical world, people also have opportunities to participate in ways that they might not otherwise have had when face-to-face.

Undergirding a focus on participatory cultures is a move toward embracing a global understanding of the world, including an openness to diverse cultures, perspectives, and ways of knowing, doing, and being. In order to communicate and collaborate across different cultures, languages, and perspectives in productive ways, people need to be open and willing to listen and understand other ways of thinking and communicating. Particularly with growing concerns about increasing social and political polarization, spurred by the rise of digital media (Dvir-Gvirman et al., 2016; Van Aelst et al., 2017), the ability to intentionally seek out multiple perspectives and to work toward solutions to problems that will work for multiple different stakeholders becomes an important skill.

Criticality, responsibility, and social change

Digital participation has demanded a focus on criticality. Here, we use the term criticality to refer to a spectrum of perspectives and practices that involve questioning and judging information and social structures and taking action toward social change (see, for example, Garcia et al., 2020). On one end of the spectrum, a person might simply question information rather than accept it as true without examining it first. Such questioning, however, can lead to judging information, which can lead, at the far end of the spectrum, to taking action toward changing existing realities and social structures.

Because anyone can publish whatever they wish on the internet, there are few vetting processes. This means that people have to evaluate information for themselves, determining the extent to which information is relevant and credible for their purposes. In the past, institutional vetting processes, such as those undertaken by publishers, libraries, and even book stores, have resulted in less diversity of text types and text quality, particularly for those texts most widely available. Today, however, it is largely up to individuals to vet information for themselves. Additionally, because of the affordances of multiple modes and texts in digital environments like the internet, people also have to understand, scrutinize and evaluate the ways in which different text types and text features affect information quality. With the rise of the internet, and, especially, of Web 2.0, criticality has become a central way of approaching and engaging with the world (Leu et al., 2013).

In addition to the need for criticality, the internet also has enabled people to critique social structures and to take action to change those social structures. The internet allows for access to information as well as for the formation of social groups—often across economic, geographic, and global boundaries that would not have been possible without the internet. Our ability to network in an unfettered manner allows us to take action to change power dynamics.

People have a responsibility to create and use information in critical, ethical ways. Because it is incredibly easy to create and share information rapidly, people have the responsibility to check credibility, interpret potential bias, and to consider, for themselves, the impact on other individuals and communities before doing so.

Multiplicity, hybridity, and flexibility

Because digital technologies allow for more modes, texts, text types, and even activities, our social values have shifted from embracing single to multiple forms of text and communication. With this shift comes a focus on hybridity, or working and moving within different combinations of modes, texts, practices, and communities, and sometimes in novel ways. Along with the need to be able to understand and make meaning from and with multiple modes and text forms across multiple, diverse cultures and people comes the need for flexibility. Such flexibility means that people need to be able to not only move among different modes, texts, and cultures but also must be able to be open to all of these different possible combinations. Additionally, people have to become flexible at moving across contexts—for example, beginning an activity in an in-person environment and moving to a digital environment. Similarly, rather than being focused in particular disciplines or fields, people are often working collaboratively in groups to solve interdisciplinary problems.

Creativity, innovation, and problem-solving

As the internet has allowed for information to be produced and copied quickly, a focus on creativity and innovation has arisen. Because the internet and other digital technologies allow for the rapid distribution of information, and because they now allow for text to be created through AI technologies, those who have new ideas and texts bring added value that cannot be created by AI (see, for example, Hancock et al., 2020). Therefore, our society places value on innovative and creative ideas that cannot simply be developed or reproduced by technology.

Moreover, rather than simply memorizing information, we now value accessing and using information in the service of problem-solving (Jacobs and Castek, 2018, 2022). Often, this means approaching informational problems with an inquiry framework—and being open to multiple, possible solutions—rather than approaching problems as if there is one, right solution. In an information-based society, people need to be able to locate, organize, evaluate, synthesize, and communicate information in creative and innovative ways as they work toward solutions (Leu et al., 2013).

New literacies social values reflected in educational standards around the world

Educational standards are beginning to reflect these changes in social values. Large curricular frameworks, such as the International Society for Technology in Education (ISTE, 2018), Partnership for 21st Century Learning (Battelle for Kids, 2019), NCTE's Position Statement on Adolescent Literacy (NCTE, 2018), and the Common Core State Standards (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010), now include practices that reflect such social values. In schools, course work centers on activities that support the development of such practices and values. Rather than being focused on individual work that occurs in isolated disciplinary areas, these new standards focus on developing students' capacity for collaboration, critical thinking, digital citizenship, innovation, and the ability to work together to solve complex, interdisciplinary problems such as those we are faced with in the real world. In this way, the affordances and constraints of digital literacies and the internet have shifted the kinds of practices we teach in schools and the ways in which these practices get taken up in learning and instruction as educators aim to support students to be prepared to work in a digital world.

Under-realized potentials

Criticality and justice

One of the most important under-realized potentials in the digital literacies field is a focus on criticality and justice (O'Byrne, 2019). Because the internet enables more people to access, interpret, use, and create information, more people—across diverse cultural, social, and economic divides—now have the ability to critique and evaluate the world. Such critique and evaluation often leads to action. Once people identify problems with existing ways of thinking, living, doing, and being, they often want to act to address

those problems. Moreover, the internet enables social action in important ways, since it allows people to gather together across more traditional divides, finding strength in numbers. The internet also allows people to take action across existing social structures, since people with an internet connection can easily send email, share information, make and carry out specific plans by communicating on social networks, websites, and through private networks.

Despite the ways the internet facilitates social change, teachers need to take up the call to become a change agent. By virtue of their role within communities, many educators have opportunities to engage in democratic discourse that the internet helps facilitate, such as engaging in debates and advocating for more equitable global solutions (Bourn, 2016). Although some classrooms have begun to incorporate projects that address social change, many have not. In fact, such social change is not often written into curricular frameworks. On the contrary, social change is sometimes seen by those in power (and those who create curriculums) as something to be kept out of our schools.

Self-criticality

Another important under-realized potential is the need to be self-critical of our own digital literacies practices. As it becomes easier and easier to quickly share and reproduce information, we have to consider the extent to which we are sharing reliable information and the impact that the information might have on other people. For example, before sharing information, we might ask ourselves how the information might affect communities of people other than those of which we are a part. This is especially relevant in a global society in which we have greater opportunity to interact with many different communities and people through different internet contexts.

Additionally, while the internet widens the communities of people with whom we interact, it also can narrow available choices. The algorithms of many social networking sites, such as TikTok and Facebook, are actually designed to limit, in scope, the kinds of content we see. Moreover, because individuals decide to whom to be connected in social networking platforms, people often are connected to other people that are similar to them. This means that the content we see is more likely to be narrowly defined because we are exposed to fewer outside perspectives. Such a phenomenon is sometimes referred to as being inside a digital echo chamber (see Dubois and Blank, 2018). Thus, rather than having to interact with people less like us in our in-person lives, we can choose to interact only with those who share perspectives similar to our own.

However, on the internet, people also have unprecedented access to multiple different kinds of information and perspectives and can choose the media and information that they read and believe. Recent research suggests that people with diverse media interaction can avoid digital echo chambers within a high-choice environment such as the internet, if they choose to do so (see, for example, Dubois and Blank, 2018).

Researchers and practitioners across the education sector are called to move toward more expansive and deeper acts of social justice in order to affect social change. As educators, we must intentionally make efforts to seek out diverse media and information and to critique—and to thoughtfully engage in—our own digital literacies practices. Because the default in online spaces is often for us to be exposed only to perspectives similar to our own, intentional effort is required for us to seek out, and to be open to, ideas other than the ones we already hold. Creating purposeful practices and habits of mind for learners of all ages to do this then becomes important in our digital society, especially for the purpose of achieving digital equity and digital inclusion.

Digital equity

Digital equity is a condition to strive for; yet, according to Reder's (2015) analyses, there is a consistent lack of digital equity based on race, gender, and national origin. The education field cannot ignore these vast racial inequities. Although education is seen by many as a pathway toward personal achievement and transformation, it is not a level playing field for all (Jacobs et al., 2022). Equitable educational practices must be rooted in a commitment to understand and build on the skills, practices, values, and brilliance of the very communities that have long been marginalized (Ladson-Billings and Tate, 1995). Equity processes require that educators be simultaneously conscientious of social realities and inequities while creating opportunities for systemic change. Equitable processes provide opportunities for quality learning experiences in a multiplicity of settings in ways that position learners as knowledgeable and creative people who work in service of each other (Oguilve et al., 2021). It is only when learners have access to learning new literacies, and opportunities to apply them in flexible ways, that individuals can be fully digitally included. Together, we can strive to connect with learners and communities with the explicit goal of disrupting oppressive power structures, flattening dynamics, and leveling the playing field in teaching, learning, leadership and research. Flattening hierarchies requires educators to re-distribute authority in education (Castek and Manderino, 2020). These changes require long-term and consciously collaborative efforts.

Awareness of data practices

The Internet offers access to large amounts of data and information. As individuals engage in online experiences, the Internet is extracting large amounts of data from us. Our data is being used to influence our everyday decision-making and by corporations who make marketing decisions about our preferences. With the amount of personal data about our choices collected with every click on any website, and with wearables collecting data on where we go, and how much we move, we need greater efforts to understand data privacy and to learn data literacies. Sternkopf and Mueller (2018) suggest that learning about data is a continuous

journey that creates the ability to identify, understand, interpret, create, communicate, and compute pieces of information (data) to develop one's knowledge and ability to participate fully in society. From a more critical standpoint, educators must carefully examine the power relationships that exist between those who produce the data, those who control the data, and those who use the data to produce the narratives that shape the social, cultural, and political world to influence economic decisions. The power to collect and interpret data can be seen as a process (Bloome et al., 2005), revealing how data can be used to either persuade, coerce, or bias a particular point of view.

In order to foster critical data literacies, we advocate for engaging learners in conversations that question who gets counted and who is not represented in data (Sheldon & Castek, accepted). As educators, we need to examine the critical and social practices to make sense of experiences in the digital world of data. This includes how we individually and collaboratively engage in creative thinking and critical questioning to learn, problem-solve, and examine the power relationships that exist between those who produce data, those who control data, and those who use the data to influence decision making.

Interconnectedness of literacies online

Digital literacies are complex, multiple, and diverse (Coiro, 2020). Heitin (2016) suggested that there are three dimensions to digital literacies: (a) searching and locating digital content (e.g., Leu et al., 2013), (b) creating digital content (e.g., Hobbs, 2017), and (c) communicating or sharing digital content (e.g., Spiers and Bartlett, 2012). Digital content refers to any product and/or information in a digital form. Within these digital forms, digital literacies, reading, writing, viewing, content creation, and communication can be viewed as tightly intertwined. As we search for and locate information, for example, we might also be creating new digital content and sharing the content that we find, along with the newly created content that resulted. In this way, meaning-making processes integrate seamlessly, within a single task, across all three of Heitin's (2016) dimensions of digital literacies. Now and in the future, digital literacies process are likely to become even more interdependent as more digital texts, tools, and networks emerge and the digital landscape continually changes.

A call to action — looking toward the future

Avenues for research

Based on the under-realized potentials we identified, we outline two key areas for future research: (a) research on critical digital literacies (see, for example, Pangrazio and Selwyn, 2019)—especially those studies that focus on social change (see, for example, Garcia et al., 2020)—and (b) research on digital identity and citizenship (see, for example, Choi et al., 2018; Gleason and von Gillern, 2018; Lauricella et al., 2020). Importantly, these two research areas are synergistic in that both support a more globalized, democratized, and equitable society in which people are both open and critical to their own and others' perspectives. This is a departure from some past social realities in which it has been productive, and sometimes even beneficial, for people to be closed to perspectives other than their own. In our pre-internet society, most communities interacted mainly with others within their own locality and only very little with other communities. In such contexts, the need to understand multiple and diverse perspectives was less important. However, in a globalized society, this need is paramount. Thus, future research requires us to work toward supporting learners to negotiate meaning across multiple perspectives, and especially across perspectives that differ from their own.

Research centered on criticality and social change

Research in digital literacies needs to take up issues of social justice and social change. The field of critical media literacy has begun to do this, examining the ways youth interact with media and the ways they critically interpret and construct media to purposefully resist existing, oppressive social narratives (Graves et al., 2020). Such research has begun to work toward effecting social change from the ground up by engaging youth in critical media literacy instruction focused on taking action toward social change using media platforms (see, for example, Guldin et al., 2021). In particular, those who study online reading and information evaluation need to consider the ways in which power structures in society affect students' online reading practices, and the ways in which students can use online reading and research to investigate, and work to change, social issues since the internet is an environment highly conducive to this work.

However, more research is needed to better understand how to support learners to engage with information and perspectives that differ from their own. In particular, we need to better understand how to support learners to be open to, but critical of, both their own and others' perspectives. Some research has begun to examine how readers' own identities and prior stances on issues, as well as their epistemic beliefs, or beliefs about how knowledge is constructed (see Kuhn and Weinstock, 2002), affect the ways in which they approach newly encountered perspectives and information. For example, Kiili et al. (2018a,b) found that students' prior stance on an issue affected their ability to evaluate online information. Cho, Woodward, Li, and Barlow (2017) found that online readers with more, compared to less, critical and flexible stances toward knowledge learned more from reading and were better able to generate questions after reading.

Importantly, future research that supports learners to be critical of their own and others' perspectives will require researchers to draw on not just educational research but also on research in the social sciences, which has begun to examine the psychology behind people's practices in digital literacies environments, such as how one's identities and beliefs affect perceptions of information such as those encountered in social media environments (see, for example, [Pereira et al., 2021](#)). Such studies allow us to better understand how to support learners to be critical of their own and others' ideas so that we can work toward social change across differing perspectives.

Research centered on digital identity and citizenship

Relatedly, many curricular frameworks around the world such as the International Society for Technology in Education ([ISTE, 2018](#)), Partnership for 21st Century Learning ([Battelle for Kids, 2019](#)), and the Common Core State Standards ([National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010](#)), now include digital identity and citizenship standards. For example, the ISTE digital citizenship standard focuses on supporting students to practice safe, ethical, and responsible information use.

In a post-truth era, responsible information use includes carefully evaluating the credibility of the information found on the internet rather than assuming it to be true. Learners must therefore thoughtfully consider their own and others' identities as they are engaging with multiple perspectives around the world. Work is emerging in understanding how students and teachers construct their digital identities and how they negotiate their identities within and across face-to-face and digital environments (see, for example, [O'Byrne and Hunter-Doniger, 2021](#)), but more work is needed in this area.

Practical considerations in our lives and in everyday contexts

The American Library Association suggests that digital literacy is the ability to use information and communication technologies to find, evaluate, create, and communicate information ([Digital Literacy Task Force, American Library Association, 2012](#)). In this chapter, we use the term literacies in the plural as a way to recognize that digital literacies refers to a composite set of skills, mindsets, and social practices that are needed to navigate online in everyday contexts and to use novel resources, tools, and interfaces in efficient and flexible ways. Literacies in our digital world are continually changing and are part of the constellation of literacy practices necessary for civic, economic, and personal engagement in all facets of our lives. We advocate for literacy educators to embrace a broader view of literacies that are neither digital nor paper-based but instead are flexible, reciprocal, and recursive across online and offline spaces. In the process, educational institutions need to expand the role of digital technologies in school, outside of school, and in our world at large and work in concert with educational partners locally, nationally, and globally linking the work of teachers, administrators, school personnel, families, and global communities.

Teaching and learning

Literacy practices are pluralistic and comprise more than reading, writing, listening, and speaking processes that have historically been core parts of educational spaces. Likewise, knowledge is not acquired through a linear process ([Kot, 2011](#)). Instead, learning is fluid and dynamic, flowing through a network of individuals, communities, societies, technologies, and artifacts ([Goldie, 2016](#)). [Watulak et al. \(2018\)](#) consider the interconnection of knowledge, people, information sources, technologies, and communities. Through this lens, literacies are mobile and evolve across spaces, times, modes, disciplines, media, and contexts ([Stornaiuolo et al., 2017](#)).

Literacies will inevitably shift over time as technologies rapidly emerge today, tomorrow, and into the future. Implications of constant change have profoundly impacted the learning materials we design, instructional pedagogies we employ, and the learning environments that transverse online and offline learning spaces. To help learners become fully literate, educators need to create multiple opportunities to integrate digital literacies, multimodal approaches, texts and tools for design, creation, play, and problem-solving across the curriculum both in and out of school. For example, co.lab is a freshman program in the Leyden School District. <https://sites.google.com/leyden212.org/colab/co-lab>. Their interdisciplinary approach to learning and collaborative teaching spaces foster inquiry, leadership, communication, and reflection. Their integrated approach to curriculum design connects all the disciplines and supports learners' digital literacies as they discover the passions that fuel their learning. Unit themes include identity, adversity and resilience, empathy, and global contributions, which are woven throughout project-based instructional approaches that culminate in community engagement events. Learners design advocacy campaigns, publish models and displays, create films, and illustrate learning through multimedia learning products that advance their self, school, and/or communities. What is salient in the co.lab context is the emphasis on ways critical literacies create opportunities to conduct inquiry and encourage the sharing of ideas digitally beyond traditional, paper-based forms. Such efforts illustrate the possibility of enacting robust efforts to incorporate digital literacies instruction widely to expand ways of reading, learning, collaborating, networking and sharing knowledge. More illustrative examples are needed to envision expansive and dynamic new literacies educational transformation. Expansive conceptions of new literacies practices widen our views of digital and multimodal learning, expanding the ways in which we conceive of meaning-making, both now and into the future.

References

- Anderson, M., 2019. Mobile Technology and Home Broadband 2019. Pew Research Center. <https://www.pewresearch.org/internet/2019/06/13/mobile-technology-and-home-broadband-2019/>.
- Battelle for Kids, 2019. Partnership for 21st century learning: framework for 21st century learning. <https://www.battelleforkids.org/networks/p21/frameworks-resources>.
- Bloome, D., Carter, S.P., Christian, B.M., Otto, S., Stuart-Faris, N., 2005. Discourse Analysis and the Study of Classroom Language and Literacy Events: A Microethnographic Perspective. Lawrence Erlbaum, Mahwah, NJ.
- Bourn, D., 2016. Teachers as agents of social change. *Int. J. Dev. Educ. Global Learn.* 7 (3), 63–77. <https://files.eric.ed.gov/fulltext/EJ1167813.pdf>.
- Britt, M.A., Gabrys, G.L., 2001. Teaching advanced literacy skills for the World Wide Web. In: Wolfe, C.R. (Ed.), *Learning and Teaching on the World Wide Web*. Academic, San Diego, CA, pp. 73–90. <https://doi.org/10.1016/B978-012761891-3/50007-2>.
- Castek, J., Manderino, M., 2020. Teaching and learning in a digital world: digital literacies for disciplinary learning. In: Kuhn, M., Dreher, M. (Eds.), *Developing Conceptual Knowledge through Oral and Written Language*. Guilford, New York, NY.
- Cho, B.-Y., Woodward, L., Li, D., Barlow, W., 2017. Examining adolescents' strategic processing during online reading with a question-generating task. *Am. Educ. Res. J.* 54 (4), 691–772. <https://doi.org/10.3102/0002831217701694>.
- Choi, M., Cristol, D., Gimbert, B., 2018. Teachers as digital citizens: the influence of individual backgrounds, internet use and psychological characteristics on teachers' levels of digital citizenship. *Comput. Educ.* 121, 143–161. <https://doi.org/10.1016/j.compedu.2018.03.005>.
- Coiro, J., 2020. Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Read. Res. Q.* <https://doi.org/10.1002/rrq.302>. Advance online publication.
- Coiro, J., Dobler, E., Pelekis, K., 2019. *From Curiosity to Deep Learning: Personal Digital Inquiry in Grades K-5*. Stenhouse Publishers.
- Digital Literacy Task Force American Library Association, 2012. Digital literacy. Retrieved from. <https://literacy.ala.org/digital-literacy/>.
- Dubois, E., Blank, G., 2018. The echo chamber is overstated: the moderating effect of political interest and diverse media. *Inf. Commun. Soc.* 21 (5). <https://doi.org/10.1080/1369118X.2018.1428656>, 729–245.
- Dvir-Gvirsman, S., Tzfati, Y., Menchen-Trevino, E., 2016. The extent and nature of ideological selective exposure online: combining survey responses with actual web log data from the 2013 Israeli Elections. *New Media Soc.* 18 (5), 857–877. <https://doi.org/10.1177/1461444814549041>.
- Forzani, E., Ly, C.N., 2022. Beyond multimodality to multiplicity: developing more equitable and relevant literacy learning spaces for young children. *Read. Teach.* x (x), 1–10. [https://doi.org/10.1002/\(ISSN\)1936-2714](https://doi.org/10.1002/(ISSN)1936-2714).
- Forzani, E., Corrigan, J., Slomp, D., 2020. Reimagining literacy assessment through a new literacies lens. *J. Adolesc. Adult Literacy* 64 (3), 351–355. <https://doi.org/10.1002/jaal.1098>.
- García, P., Fernández, C., Okonkwo, H., 2020. Leveraging technology: how black girls enact critical digital literacies for social change. *Learn. Media Technol.* 45 (4), 345–362. <https://doi.org/10.1080/17439884.2020.1773851>.
- Gibson, J.J., 1977. The theory of affordances. In: Shaw, R., Bransford, J. (Eds.), *Perceiving, Acting and Knowing: Toward an Ecological Psychology*. Erlbaum, pp. 67–82. https://monoskop.org/images/2/2c/Gibson_James_J_1977_The_Theory_of_Affordances.pdf.
- Gleason, B., von Gillern, S., 2018. Digital citizenship with social media: participatory practices of teaching and learning in secondary education. *Educ. Technol. Soc.* 1, 200–212.
- Goldie, J.G.S., 2016. Connectivism: a knowledge learning theory for the digital age? *Med. Teach.* 38 (10), 1064–1069.
- Goldman, S.R., Lawless, K., Manning, F., 2013. Research and development of multiple source comprehension assessment. In: Britt, M.A., Goldman, S.R., Rouet, J.-F. (Eds.), *Reading—from Words to Multiple Texts*. Routledge, New York, NY, pp. 180–199.
- Graves, D., Kelly, L., McArthur, S., 2020. Introduction: teaching for critical consciousness at the intersection of critical media literacy and Hip Hop education. *Int. J. Crit. Media Lit.* 2, 1–8. <https://doi.org/10.1163/25900110-00201001>.
- Greenhow, C., Robelia, B., Hughes, J., 2009. Web 2.0 and classroom research: what path should we take now? *Educ. Res.* 38, 246–259. <https://doi.org/10.3102/0013189X09336671>.
- Guldin, R., Madison, E., Anderson, R.C., 2021. Writing for social justice: journalistic strategies for catalyzing agentic engagement among Latinx middle school students through media education. *J. Media Lit. Educ.* 13 (2), 71–85. <https://doi.org/10.23860/JMLE2021-13-2-6>.
- Hammond, M., 2010. What is an affordance and can it help us understand the use of ICT in education? *Educ. Inf. Technol.* 15 (3), 205–217. <https://doi.org/10.1007/s10639-009-9106-z>.
- Hancock, J.T., Naaman, M., Levy, K., 2020. AI-mediated communication: definition, research agenda, and ethical considerations. *J. Comput. Mediat. Commun.* 25, 89–100. <https://doi.org/10.1093/jcmc/zmz022>.
- Heitin, L., November 2016. What is digital literacy? Education Week. Retrieved from. <https://www.edweek.org/teaching-learning/what-is-digital-literacy/2016/11>.
- Hobbs, R., 2017. *Create to Learn: Introduction to Digital Literacy*. John Wiley & Sons, Hoboken, NJ.
- International Society for Technology in Education (ISTE), 2018. ISTE standards for education leaders. Retrieved from. <https://www.iste.org/standards>.
- Jacobs, G.E., Castek, J., 2018. Digital problem solving: the literacies of navigating life in the digital age. *J. Adolesc. Adult Lit.* 61 (6), 681–685. <https://doi.org/10.1002/jaal.745>.
- Jacobs, G.E., Castek, J., 2022. Collaborative digital problem-solving: power, relationships, and participation. *J. Adolesc. Adult Lit.* 65 (5), 377–387. <https://doi.org/10.1002/jaal.1216>.
- Jacobs, G.E., Castek, J., Harris, K., Vanek, J., 2022. Examining the perspectives of adult working learners and key stakeholders using critical race theory. *High Educ. Skills Work. Based Learn.* <https://doi.org/10.1108/HESWBL-08-2021-0158>.
- Jenkins, H., (with Clinton, K., Purushotma, R., Robison, A.J., & Weigel, M.). (2006). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. Chicago, IL: John D. and Catherine T. MacArthur Foundation. Retrieved from https://www.macfound.org/media/article_pdfs/JENKINS_WHITE_PAPER.PDF
- Johnson, J., 2021. Internet usage worldwide—statistics and facts. Statista. <https://www.statista.com/topics/1145/internet-usage-worldwide/#dossierKeyfigures>.
- Kiili, C., Leu, D.J., Marttunen, M., Hautala, J., Leppänen, P.H.T., 2018a. Exploring early adolescents' evaluation of academic and commercial online resources related to health. *Read. Writ.* 31, 533–557. <https://doi.org/10.1007/s11145-017-9797-2>.
- Kiili, C., Leu, D., Utraiainen, J., Coiro, J., Kannianen, L., Tolvanen, A., Lohvansuu, K., Leppanen, P., 2018b. Reading to learn from online information: modeling the factor structure. *J. Lit. Res.* 50, 304–334. <https://doi.org/10.1177/1086296X18784640>. Retrieved from.
- Kot, R., 2011. The challenges to connectivist learning on open online networks: learning experiences during a Massive Open Online Course. *Int. Rev. Res. Open Dist. Learn.* 12, 19–38.
- Kress, G., 2003. *Literacy in the New Media Age*. Taylor & Francis. <https://doi.org/10.4324/9780203299234>.
- Kuhn, D., Weinstock, M., 2002. What is epistemological thinking and why does it matter? In: Hofer, B.K., Pintrich, P.R. (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*. Lawrence Erlbaum, pp. 121–144.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Lauricella, A.R., Herdzina, J., Robb, M., 2020. Early childhood educators' teaching of digital citizenship competencies. *Comput. Educ.* 158, 1–9. <https://doi.org/10.1016/j.compedu.2020.103989>.
- Leu, D.J., Kinzer, C.K., Coiro, J., Castek, J., Henry, L.A., 2013. New literacies: a dual level theory of the changing nature of literacy, instruction, and assessment. In: Alvermann, D.E., Unrau, N.J., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Reading*, sixth ed. International Reading Association, Newark, DE, pp. 1150–1181.
- National Council of Teachers of English [NCTE], 2018. A call to action: what we know about adolescent literacy instruction. <https://ncte.org/statement/adolescentliteracy/>.

- National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010. Common Core State Standards for English Language Arts and Literacy in History/social Studies, Science, and Technical Subjects. http://www.corestandards.org/wp-content/uploads/ELA_Standards1.pdf.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>.
- Oguilve, V., Wen, W., Bowen, E., Abourehab, Y., Bermudez, A., Gaxiola, E., Castek, J., 2021. Community making: an expansive view of curriculum. *J. Curric. Stud. Res.* 3 (1), 69–100. <https://doi.org/10.46303/jcsr.2021.8>.
- O'Byrne, I., 2019. Educate, empower, advocate: amplifying marginalized voices in a digital society. *Contemp. Issues Technol. Teach. Educ.* 19 (4), 640–669. <https://www.learnedlib.org/primary/p/188279/>.
- O'Byrne, W.I., Hunter-Doniger, T., 2021. Pre-service educators developing a digital identity. *TechTrends* 65 (4), 444–453. <https://doi.org/10.1007/s11528-021-00617-5>.
- Pangrazio, L., Selwyn, N., 2019. "Personal data literacies": A critical literacies approach to enhancing understandings of personal digital data. *New Med. Soc.* 21 (2), 419–437. <https://doi.org/10.1177/1461444818799523>.
- Pereira, A., Harrison, E.A., Van Bavel, J.J., 2021. Identity concerns drive belief: the impact of partisan identity on the belief and dissemination of true and false news. <https://doi.org/10.31234/osf.io/7vc5d>.
- Perrin, A., Atske, S., 2021. 7% of Americans Don't Use the Internet. Who Are They? Pew Research Center. <https://www.pewresearch.org/fact-tank/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/>.
- Reder, S., 2015. Digital inclusion and digital literacy in the United States: a portrait from PIAAC's survey of adult skills. Retrieved from. https://static1.squarespace.com/static/51bb74b8e4b0139570ddf020/t/551c3e82e4b0d2fede6481f9/1427914370277/Reder_PIAAC.pdf.
- Rideout, V., Robb, M.B., 2019. The common sense census: media use by teens and tweens. Retrieved from. <https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-teens-and-tweens-2019>.
- Rideout, V., Robb, M.B., 2020. The common sense census: media use by kids age zero to eight. https://www.commonsensemedia.org/sites/default/files/uploads/research/2020_zero_to_eight_census_final_web.pdf.
- Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M., Dinet, J., 2011. The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learn. Instruct.* 21 (2), 205–219. <https://doi.org/10.1016/j.learninstruct.2010.02.007>.
- Ryder, M., Wilson, B., 1996. Affordances and constraints of the internet for learning and instruction. In: Proceedings of Selected Research and Development Presentations at the 1996. National Convention of the Association for Educational Communications and Technology, Indianapolis, IN.
- Semingson, P., O'Byrne, I., Mora, R.A., Kist, W., 2017. Social scholarship and the networked scholar: researching, reading, and writing the web. *Educ. Media Int.* 54 (4), 360–372. <https://doi.org/10.1080/09523987.2017.1391525>.
- Sheldon, L. & Castek, J. (accepted). New Literacies: expanding the conversation to include data literacies. In P. Schultz & K. Muis (Eds.), *Handbook of Educational Psychology* (fourth ed.). American Psychological Association.
- Spires, H.A., Bartlett, M.E., 2012. Digital Literacies and Learning: Designing a Path Forward (Friday Institute White Paper Series No. 5). Friday Institute for Educational Innovation, North Carolina State University, Raleigh.
- Sternkopf, H., Mueller, R.M., 2018. Doing good with data: development of a maturity model for data literacy in non-governmental organizations. In: Proceedings of the 51st Hawaii International Conference on System Sciences. <https://doi.org/10.24251/hicss.2018.630>.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Lit. Res.* 49 (1), 68–91.
- Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C.H., Matthes, J., Stanyer, J., 2017. Political communication in a high-choice media environment: a challenge for democracy? *Ann. Int. Commun. Assoc.* 41 (1), 3–27. <https://doi.org/10.1080/23808985.2017.1288551>.
- Watulak, S.L., Woodard, R., Smith, A., Johnson, L., Phillips, N., Wargo, K., 2018. Connected teaching and learning in K-16+ contexts: an annotated bibliography. *Contemp. Issues Technol. Teach. Educ.* 18 (2), 289–312.

Relevant website

<https://sites.google.com/leyden212.org/colab/co-lab>.

Digital media arts and critical media education: an alchemy of creative literacy practices

Michelle Cannon, The UCL Knowledge Lab, Institute of Education, University of London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction & context	497
A snapshot of media education in the UK	498
The critical and the creative: an alchemy of practice	498
Le Cinéma: Cent Ans de Jeunesse	499
Blending drama and digital animation	500
Blending promenade performance, literature and digital media	500
Concluding thoughts	501
Acknowledgment	501
References	501

Introduction & context

A strong body of critical scholarship has been accumulating in developed nations over the past decades supporting film and media education as core components of formal school curricula from the primary years through to students' early-to-mid teens (Buckingham, 2003; Hoechsmann and Poyntz, 2012; Cannon, 2018). In many instances, however, this scholarship fails to translate into policy and practice. In England for example, the country in which accounts of practice in this chapter are located, a narrow back-to-basics agenda prevails in state education that ignores media learning. Watching, analyzing and producing media texts are often framed as introductory routes to literacy work perceived as legitimate. Either this, or engagements with film, moving image and animation remain in the world of informal literacy enrichment, once the "real work" of established literacy has been covered.

This view of media literacy as a secondary or "residual" category (Potter, 2012) contrasts with other nations' progressive curricular reforms that incorporate everyday digital media cultures and media literacy programs as vital components of a more relevant and dynamic curriculum (Kupiainen, 2013). It seems that, if media literacy practices are accommodated at all in the English state school system, they are accommodated mainly in the informal spaces of extracurricular learning. While these peripheral approaches are welcomed, they are by no means enough nor sufficiently widespread to equip children and young people for adequate media learning in the digital age. This chapter reflects on the importance and relevance of critical and creative media education across the curriculum from the primary years of learning onwards. Digital media production processes and creative media arts are seen as central dimensions of a wider and more diverse conceptualization of literacy, one that is in step with multiple communicative modes, digital meaning-making practices and the complex moving image and media landscapes in which most children and young people in developed countries are immersed.

The rationale for a multimodal approach to literacy (Jewitt, 2008) is based on the premise that there is a need to equip our youth, not only with digital skills and knowledge narrowly linked to employability, but also with a critical grasp of the communicative means with which they make sense of themselves and others, and the mediated sociocultural circumstances in which they live their lives. To conceive of literacy as a set of static reading, writing and literary appreciation skills, is simply not enough for full and meaningful participation in late-modern society. Following Gee (2015) and the tradition of New Literacy Studies (NLS), and Street's social literacies work (2003), this chapter advocates for school experiences that are responsive to the contingencies of everyday living and being. These are literacy practices conceived as fluid, embodied and dynamic (Potter and McDougall, 2017), and have more to do with the relational and the sociomaterial conditions of learning (Burnett and Merchant, 2020a,b; Hawley, 2021; Parry and Taylor, 2021). In this view, pedagogies shift to less hierarchical and more inclusive models of practice, and where collaborative practices in so-called "third spaces" come to the fore (Potter and McDougall, 2017). Of primary importance in these spaces is an approach to literacy that builds on learners' existing interests, knowledge and cultural affiliations, thus a critical embrace of youth media worlds and popular culture is as important as the introduction of a range of canonical texts.

The author argues that as the noise of "educational technology solutionism" proliferates (Selwyn, 2016), and as the use of commercial digital platforms consolidates in school contexts (Williamson and Hogan, 2020), there is a need to press pause on issues of efficient data management, to give more thought to the relevance of school curricula to learners' lived lives. Many schools are sold all-encompassing "ed-tech solutions" by technology companies promising to improve schools' and pupils' "performance". It is as if what matters most is the mobilization of one-size-fits-all "what works" ideologies (Macgilchrist et al., 2021), with contestable ends that expunge the complex heterogeneity of school contexts and mask the uneven distribution of basic educational resources. The global pandemic has brought to light many of the pre-existing inequalities that permeate school structures, not least on questions of digital access. This is the moment for remote learning for all to be improved, and for learners of all ages to be authorized to look both inwards for reflection on their media habits and cultural agency, and outwards toward an analysis of the invisible

ideologies that give structure to their media worlds and that shape their very identities (Buckingham, 2008, 2019). The proposition is that as children and young people make, edit and disseminate film and moving image texts of relevance to them, they engage in a level of praxis that develops a particular disposition generative of positive and lasting learning moments. This blend of reflexive engagement with production, and then the sharing of creative outputs is framed as a kind of alchemy.

The reader will be presented with three case studies, involving children ranging in age from 10 to 15-years-old in informal school contexts. The first involves film-making in an international film education program based in Paris, the second relates to a school drama and digital animation project in London, and the third describes a dramatic, multimodal promenade performance from the *Beowulf Sydney Project* in Australia (Cameron et al., 2018). These arts-based approaches to literacy (Burn et al., 2014) foster pupils' feelings of empowerment, performance and artful manipulation of multiple forms of media. It is argued here that these practices not only build confidence in the ability to learn, but esthetic sensitivities toward the moving image, soundscapes, prop-making, puppetry, the senses, and the smooth orchestration of time and space. Less is written about the latter modes that shape the temporal and the spatial, preoccupied as many are, and understandably so, with the meaning-making affordances of the audiovisual (Bazalgette, 2000). Young participants in these projects engage in productive, playful and reflexive practices that explicitly question creative choices, and in the mobilization of rhetorical devices to produce particular effects (Burn and Durran, 2007; Cannon et al., 2018). Experiences such as these arguably function as stepping stones to children's incipient critical understanding of the constructed nature of their media worlds. Crucially, and by extension, sustained practical media work of this nature may start to have a bearing on the dispositions that inform their capacity to have a stake in these worlds.

A snapshot of media education in the UK

Media education in England is in a parlous state. The other UK nations—Wales, Northern Ireland and Scotland—have more advanced ideas on what constitutes a broad curriculum inclusive of digital media and the expressive arts up to 14 years of age. Twenty years ago, UK government interest in creative media arts was very much in evidence, and direct funding for projects was made available over the course of a decade (Parker, 2013). Cultural partnerships, and research therein, flourished between schools, artists, practitioners and arts agencies (Thomson et al., 2012). However, more conservative policy-makers have since failed to build on these developments, and as Buckingham (2014) critiques, even actively sought to remove any mention of digital media from the curriculum in England. Concurrently, as a result of strong industry lobbying, the subjects of computer-programming and computational thinking were hastily created and rolled out to primary schools across the country in an effort to fill workforce gaps. Such instrumental interventions were regarded as an evolution of the ICT curriculum (Information, Communication and Technology), but also oversaw the loss of the Communication component. Media literacy educators saw the demise of creative media arts initiatives as a missed opportunity for progressive education to gain a foothold in the transformation of established literacy practices.

It will be clear from what has been detailed above that literacy is an ideological and political football and the powers that mold its contours and content are the same powers that constrain the possibilities for learners to represent and express themselves using a variety of different symbolic resources. In the move from page to screen, this chapter asserts that the written word is just one of a range of media, and the pen just one of a range of tools, with which to wield meaning and participate socially and culturally. Indeed, arguments have been made based on the articles of the UNCRC (United Nations Convention on the Rights of the Child, 1989), that it is children and young people's right to be offered as rich a palette as possible of digital tools, languages and representational practices to express their identity (Cannon et al., 2022; Cannon, 2023). Such an approach is regarded as more egalitarian and accessible for diverse and democratic communities of learning. The reader will detect a strong socio-political dimension to the ideas on literacy presented in this chapter, and the instrumental ways in which it can be re-fashioned and mobilized in the interests of the state, without necessarily appropriate thought being given to the varied needs and interests of young learners. It is with this imbalance in mind that I turn to theory to bolster the case for a wider literacy and the development of learners as media literate practitioners.

The critical and the creative: an alchemy of practice

The theoretical framing of this account is rooted in two academic disciplines, the one being Cultural Studies (Hall, 1997; du Gay et al., 2003), the other New Literacy Studies (Street, 2003; Gee, 2015). The scope of this piece of writing does not allow an in-depth foray into these disciplines, but by way of summary, I offer these working definitions for the purposes of grounding the accounts of practice to follow. In the Cultural Studies tradition, attention is given to the ways in which texts and artifacts are produced, circulated and consumed within a matrix of power relations. These social dynamics and the multiple ways meanings are constructed, foreground the analysis of discourse, texts and historical context. As we enter the second decade of the 21st century, digital media are the primary means of communication and everyday cultural exchange, thus, a vision of literacy as related to productive practices and the situatedness of social interactions becomes more salient.

We can start to see the connections between the principles of Cultural Studies and New Literacy Studies: they are concerned with an impulse to continually experiment and question, and to create and collaborate on new knowledge and perspectives. Critical social thought and creative identity work (Potter, 2012) are less privileged in institutional literacy practices, whereas these two pillars of social progress are what binds the two academic traditions on which this account draws. It is proposed here that film

and moving image education is one means through which a contextualized sense of self can be explored, embodied and discussed (Cannon et al., 2014). The act of producing and displaying moving image texts with a soundtrack is suggestive of a process of alchemy (or “une petite alchimie” as described by Natalie Bourgeois, in Cannon, 2018, p. 180), where the relationship between children and their films shifts. In theoretical terms, an “(im)material” (Burnett et al., 2012) set of localized meanings conceived in a private space, translates into diffuse meanings in a communal space, and the ability to sense and articulate this transmutation, is a powerful learning moment.

Le Cinéma: Cent Ans de Jeunesse

Film literacy is a discrete discipline but for the purposes of this chapter, it is framed here as a subset of media literacy, as it addresses the ways in which our cultural repertoires, whatever the origin, become valued resources to be brought to light, and then critically and creatively re-worked through the medium of film. I first introduce an international film education program—Le Cinéma: Cent Ans de Jeunesse (*Cinema: a hundred years “young”*, henceforth CCAJ) based at the Cinémathèque française in Paris, and its collaboration with The British Film Institute (BFI) in London. The program has been running for over 25 years and the BFI has developed its own national version of the program. In this account, I describe workshops and events with 10–11-year-old pupils at a north London primary school, a more detailed description of CCAJ’s practices on an international scale, can be found in Reid (2019).

The overall purpose of CCAJ is for young people and educators to watch, interpret, discuss and practice making films, where film is conceived as an art form with intrinsic worth, as opposed to a medium with strictly entertainment or instrumental value. Participants are given a series of simple and highly structured production exercises, culminating in a “film essai”—which is an 8 min short film exploring the theme for that year. The program has other distinguishing features: themes revolve around film language; there is no appeal to the tropes of Hollywood glitz, festivals or competition; the age range of participants is from 7 to 18; and where possible some of the young film-makers from around the world unite in Paris for summer screenings of their work. A key component of the program is the focus on the esthetic affordances of film as an art form. Other familiar dimensions of educational filmmaking, such as acting, narrative structure, script-writing and dialog are de-emphasized in favor of visual and cinematic storytelling. But perhaps the most important characteristic is the invitation to film-makers of all ages to discuss their authorial choices after the screening, in a public multilingual Q & A. Amid a group of peers, young creators are given opportunities to articulate their craft explicitly, to express the challenges they faced in the production process, and to exercise their critical muscle in a safe receptive arena.

CCAJ actively works at maintaining the balance between theory and practice, and between reflection and material experimentation. This key principle means that explicit attention is given to the reading and writing of film. Past themes on cinematic grammar have included: camera movement, color, light, depth of field, mise-en-scène, the weather, games and playing, or the more conceptual hiding and revealing in film. In 2013–2014 for example, the theme was “le plan-séquence”, or the long take. This is defined as one sustained shot, moving or static, that frames a sequence of events, and was inspired by the century old practices of French film pioneers, the Lumière brothers. Indeed, continuities with cultural histories are maintained by watching specific clips from a wide range of nations and times past, with viewer-learners’ sensibilities attuned to one particular film esthetic. Natural elements in the local environment, tropes from popular cultural texts, and qualities within the cinematic frame are in constant dialog with each other, combining to develop the digital artistry of students.

My account in this volume focuses on orientation to learning that illustrates the kinds of sophisticated cognitive and imaginative thought processes analogous with regular literacy practices.

Learners are offered opportunities to voice creative decision-making, on the basis that confident speaking drawing on collective experience, is a vital aptitude that is often over-looked in conventional literacy practices. Film screenings are integral to the design of CCDJ, in that they are a moment of celebration, a part of the media-making process during which audiences become co-collaborators in the film’s biography. Burn (2009) proposes that the making and showing of film and media texts transforms the maker in empowering ways contributing to experiences of cultural engagement that might otherwise elude them. As the young makers assemble in front of the screen, even when words evade some of the less vocal pupils, data suggests that most feel a unique sense of ownership and authorship.

More detail on the intricacies of the CCAJ program and the analysis of the children’s making experiences can be found in Chapter 5 of Cannon (2018), what I wish to demonstrate below is the significance of the post-screening Q&A. Having answered the same film brief, the audience of peers are imaginatively and cognitively engaged in each other’s work. Questions are posed and creative challenges are discussed in an open arena. In most cases, the quality of peer-to-peer attention at this point is almost palpable as the normative hierarchies are overturned and autonomous learners express themselves from a position of knowing through experience, rather than being told by those in authority. The level of sophistication in shooting and editing that it is possible for young film-makers to achieve, and for young audiences to notice, is such that intellectual curiosity is piqued and artistic detail starts to surface.

The reader may remember that the theme for the particular year in question was the long take, where students were to experiment with the nature and composition of a long take and how they can be achieved in a school environment with an iPad and no steady-cam or professional equipment. An example of nuanced understanding and creative problem-solving was evinced when one pupil explained to the attentive audience how her group enacted one of the stipulated camera movements in the brief. According to the CCAJ brief, the camera had to bifurcate, meaning that it had to depart momentarily from the main action, move autonomously through the space of the story world, and then re-join the action. The children decided to achieve this by leaving one dramatic scene in a classroom and joining another in the playground. They shot the long take through a series of doors and along corridors using

a school dinner trolley as a dolly, with the camera secured on top with tape and a small tripod. Along the way the group figured out how to open double doors and negotiate flights of stairs while filming so as to enter the playground and re-join the actors to effect a smooth transitional moment. This kind of initiative on the fly is typical of the kind of practical decision-making that filmmaking can occasion. Over a period of sustained practice, the pupils deal with constraints and amass an archive of audiovisual vocabulary to tell their stories with impact and artistry. Reid (2019) argues that to be literate is to fully participate in a culture, the language, the values, and the meaning-making and sharing practices of a community, and such dialogic screening events are one way of rehearsing these capacities.

Blending drama and digital animation

Digital animation is a mode of meaning-making that remains undeveloped in many mainstream school environments, which is lamentable when considering the ways in which it can enlarge orthodox literacy and visual arts practices from an early age. Digital animation can be slotted into classroom practices across the curriculum, whether in the context of storytelling or the illustration of scientific processes, it combines the use of film language and hands-on craft principles. The next case study demonstrates how animation can be brought into the realm of dramatic performance to enhance mood and atmosphere, while providing important additional roles for pupils in the iconic event that is the production and performance of a school play. Researchers from the UCL Institute of Education in London were interested in the relationships between drama, literacy and digitally mediated animated texts. This project, entitled *Adventures in Drama and Digital Animation*, was funded by the United Kingdom Literacy Association (UKLA), and involved a group of 10-year-old animators collaborating with practitioners from an internationally recognized theater company, called 1927.

1927's expertise in animated backdrops was central to the investigation and provided inspiration for a school production of *The Lion King*. Practitioners ran workshops on how to make articulated silhouette puppets for use in animations, deploying a lightbox, the Stop Motion app and simple editing software. The class as a whole suggested key points in the narrative where an animation would be most appropriate, such as the wilder beast stampede, a dream sequence and the fall of Mufasa, and they then set about making 2D animal puppets out of black card and thin spirals of wire. A smaller group of children were shown how to turn the puppets into animated characters in evocative narrative sequences. Each sequence was a separate movie file, and these were inserted into a Power point presentation, where the slides served as multimodal backdrops.

There were many interesting outcomes in this project, the extent of which had not been foreseen at the outset, namely: the different backstage roles that emerged as the play and peripheral activities progressed; the responsible sense of ownership and agency that was displayed by certain children perceived as disruptive; the leveling cultural recall, as all participants in this project could quote from the original Disney animation and contribute to the overall effect; and the esthetic and affective value of composing with material objects, moving images, music, drama and technology. From a teacher's perspective, what was overwhelmingly the case was the children's willingness to partake in painstakingly detailed craftwork to achieve particular visceral and emotional effects. This work proved to be an exciting gateway for learners described as low literacy achievers, to become cultural producers in their own right: they were offered another medium through which to re-interpret their funds of knowledge, and opportunities to gain mastery of particular digital technologies. If high stakes hand-written texts are the gold standard of literacy competence, these had been exchanged for hand-wrought displays of nuanced understanding through performative spectacle in a creative and educational context.

Blending promenade performance, literature and digital media

A further example of the ways in which media can knit the disciplines together is seen in a project for 11–15-year-olds that took place in Sydney, at the Australian Theater for Young People (Cameron et al., 2018) during a week-long school holiday workshop. The project was an "intermedial mashup" of drama, drawing, visual arts, ancient text, a 2 m-tall large-scale puppet, video game design, voice acting, digital and live dramatic creation processes, and a 20 min promenade performance for family, friends and Periscope streaming. The show re-animated, re-mediated and re-interpreted the classic Anglo-Saxon epic poem, *Beowulf*, as an integrated live arts extravaganza, complete with warriors, dragons, monsters, mobile bluetooth speakers, and the digital projection of a video game designed by children as the backdrop.

The *Beowulf Sydney Project's* interdisciplinary, drama-making model builds on Burn (2013) and Bryer et al.'s (2014) work on revisiting classic literary texts such as *Macbeth*, *Beowulf* and Edgar Allan Poe's *The Raven*, and understanding them in the context of film-making and video game-making. The latter scholars' rule-bending aligns with what Cameron et al. (2018, p. 113) call a bold "proto-pedagogy" that welcomes hybrid forms and complex collaborative classroom practices. It showcases young people's willingness to engage in emergent new forms of performative meaning-making, where they were able to: flip between genres and co-construct a hypermedia narrative; alternate roles on stage and share responsibility; take charge of the rules and conventions of game-design; see hitherto obscure connections between the ludic, the canonical, the representational, the real and the virtual; and finally the youth in this project relished in the real-time appreciation of the audience's response to the "liveness affect" (Cameron et al., 2018, p. 138). School shows strike a chord in most of us, often etched in memory as moments in childhood that in some way transcend the norms. In this project, the mingling of media and visual arts, literary forms, technology and the unconventional proximity

between audience and performers disrupted traditional boundaries. The disposition to experiment playfully with expressive modes is part of what constitutes critical media literacy (Cannon, 2018), rehearsing as it does the means by which a comfortable and meaningful collusion with audiences is achieved.

Concluding thoughts

It has been put forward that educational film and media production activities illustrate the kinds of situated cultural practices that encourage dynamic perspectives on literacy, but they are more than this. The momentum that was generated was a gestalt of greater significance than the media texts themselves. Pedagogies in “third spaces”, meaning those interstitial learning spaces that are neither school nor home, are permeable and allow for informal and perhaps previously unrecognized skills and abilities to emerge and be valued (Potter and McDougall, 2017). Indeed, CCAJ principles and those underpinning the drama, and mixed media performance projects are dismissive of prescribed ranking and categorizing systems; rather the aim in this work with young people is for a parity of understanding, and a co-dependence between apprentices and facilitators. In other words, project work that involves integrated arts and intermedial practices often generate a kind of wisdom that attends to the value of qualitative, esthetic and conceptual relationships, and an appetite for responsibility and risk-taking.

Grammar or spelling inaccuracies in print are a common occurrence for some learners, and this is damaging and disempowering if reading and writing are the sole means of representational story-telling at their disposal. Story-making with media enables fluency in a language that is already familiar, and any sense of failure is just part of the process of iterative review and redraft. By widening the palette of available tools and media with which to make meaning and authorizing the incorporation of popular cultural motifs in literacy work, school leaders are participating in more democratic, relevant and inclusive approaches to learning. These are safe spaces for children and young people to practice oracy and rhetorical skills, to debate ambiguous meanings, and to offer rationales for creative actions. It was suggested earlier that there is a sweet spot of learning that can be described as “une petite alchimie”, at the moment when film, media and performative texts are released into the interpretive ether and become separate entities. All the skill and open-ended possibility thinking that go into these works make a ripple in the world, and recognition of this impact emboldens the learner.

Learners get to try things out and act on hunches in intermedial territories both familiar, in terms of media consumption, and strange in terms of chances at concrete practice. Eisner (2005), in the context of art education, called this the exercise of “judgment in the absence of rule”, and for him this was one of the most important attributes of arts learning. Eisner’s thinking applies just as much in the field of contemporary media literacy, and just as much to primary learners as it does to older students. Making critical evaluations of one’s own and others’ decisions, and of the effects on others of these decisions, is an indispensable part of the quest for social transformation and human flourishing. It is incumbent on school leaders and policy-makers to refresh and reset literacy practices, in support of more sophisticated and ambitious agendas, in the interests of young learners’ participation in democratic futures.

Acknowledgment

The author would like to thank the The United Kingdom Literacy Association (UKLA) for funding the Adventures in Drama and Digital Animation project, and also to thank academic colleagues at the UCL Institute of Education, Dr. Sara Hawley, Dr. Theo Bryer and Rebecca Wilson for their expert collaboration in this project. A co-written article is forthcoming in the UKLA journal *Literacy* in 2023.

References

- Bazalgette, C., 2000. A stitch in time: skills for the new literacy. *Engl. Educ.* 34 (1), 42–49.
- Bryer, T., Lindsay, M., Wilson, R., 2014. A take on a Gothic poem: tablet film-making and literary texts. *Changing English: Studies in Culture and Education* 21 (3), 235–251.
- Buckingham, D., 2003. *Media Education: Literacy, Learning and Contemporary Culture*. Polity, Cambridge, UK.
- Buckingham, D., 2008. Introducing identity. In: Buckingham, D. (Ed.), *Youth, Identity and Digital Media*. The John D. And Catherine T. MacArthur Foundation Series on Digital Media and Learning. MIT Press, Cambridge, MA, pp. 1–22.
- Buckingham, D., 2014. The success and failure of media education. *Media Educ. Res. J.* 4 (2), 5–17.
- Buckingham, D., 2019. *The Media Education Manifesto*. Polity Press, Cambridge.
- Burn, A., Durran, J., 2007. *Media Literacy in Schools: Practice, Production and Progression*. Paul Chapman, London.
- Burn, A., Potter, J., Reid, M., 2014. Media arts, digital culture and education. *Media Educ. Res. J.* 5 (1), 5–14.
- Burn, A., 2009. *Making New Media: Creative Production and Digital Literacies*. Peter Lang, New York.
- Burn, A., 2013. Computer games on the playground: ludic systems, dramatised narrative and virtual embodiment. In: Willett, R., et al. (Eds.), *Children, Media and Playground Cultures*. Palgrave Macmillan UK, London, pp. 120–144.
- Burnett, C., Merchant, G., 2020a. *Undoing the Digital: Sociomaterialism and Literacy Education*. Routledge, Abingdon.
- Burnett, C., Merchant, G., 2020b. Literacy-as-event: accounting for relationality in literacy research. *Discourse* 41 (1), 45–56.
- Burnett, C., Merchant, G., Pahl, K., Rowsell, J., 2012. The (im)materiality of literacy: the significance of subjectivity to new literacies research. *Discourse* 35 (1), 90–103.
- Cameron, D., Anderson, M., Wotzko, R., Brater, E., Batty-Taylor, M., 2018. *Drama and Digital Arts Cultures*. Bloomsbury, London.
- Cannon, M., Bryer, T., Lindsey, M., 2014. Media production and disruptive innovation: exploring the interrelations between children, tablets, teachers and texts in subject English settings. *Media Educ. Res. J.* 5 (1), 16–31.

- Cannon, M., Connolly, S., Parry, R., 2022. Media literacy, curriculum and the rights of the child. *Discourse: Studies in the Cultural Politics of Education* 43 (2), 322–334.
- Cannon, M., Potter, J., Burn, A., 2018. Dynamic, playful and productive literacies. *Changing English: Studies in Culture and Education* 25 (2), 183–199.
- Cannon, M., 2018. Digital Media in Education: Teaching, Learning and Literacy Practices With Young Learners. Palgrave Macmillan, London.
- Cannon, M., 2023, forthcoming. Making, feeling and moving among media: a pupil's right. In: de Abreu, B. (Ed.), *Media Literacy, Equity and Justice*. Routledge, Abingdon.
- Du Gay, P., Hall, S., Janes, L., Koed Madsen, A., MacKay, H., Negus, K., 2003. *Doing Cultural Studies: The Story of the Sony Walkman*. Sage Publications, London.
- Eisner, E., 2005. *Reimagining Schools: The Selected Works of Elliot W. Eisner*. Routledge, Abingdon.
- Gee, J.P., 2015. The new literacy studies. In: Rowsell, J., Pahl, K. (Eds.), *The Routledge Handbook of Literacy Studies*. Routledge, Abingdon, pp. 35–48.
- Hall, S., 1997. *Representation: Cultural Representations and Signifying Practices*. Sage Publications, London.
- Hawley, S., 2021. Doing sociomaterial studies: the circuit of agency. *Learn. Media Technol.* <https://doi.org/10.1080/17439884.2021.1986064>.
- Hoechsmann, M., Poyntz, S.R., 2012. *Media Literacies: A Critical Introduction*. Wiley-Blackwell, Oxford.
- Jewitt, C., 2008. Multimodality and literacy in school classrooms. *Rev. Res. Educ.* 32, 241–267.
- Kupiainen, R., 2013. *Media and Digital Literacies in Secondary School*. Peter Lang, Oxford.
- Macgilchrist, F., Potter, J., Williamson, B., 2021. Shifting scales of research on learning, media and technology. *Learn. Media Technol.* 46 (4), 369–376.
- Parker, D., 2013. *Creative Partnerships in Practice: Developing Creative Learners*. Bloomsbury, London.
- Parry, B., Taylor, L., 2021. Emergent digital authoring: playful tinkering with mode, media, and technology. *Theory Into Pract.* 60 (2), 148–159.
- Potter, J., McDougall, J., 2017. *Digital Media, Culture and Education: Theorising Third Space Literacies*. Palgrave Macmillan, London.
- Potter, J., 2012. *Digital Media and Learner Identity: The New Curatorship*. Palgrave Macmillan, London.
- Reid, M., 2019. Film, arts education, and cognition: the case of le cinéma cent ans de jeunesse. In: Hermansson, C., Zepernick, J. (Eds.), *The Palgrave Handbook of Children's Film and Television*. Palgrave Macmillan, New York, pp. 469–484.
- Selwyn, N., 2016. *Is Technology Good for Education?* Polity Press, Cambridge.
- Street, B., 2003. What's "new" in New Literacy Studies? Critical approaches to literacy in theory and practice. *Curr. Issues Comp. Educ.* 5 (2), 77–91.
- Thomson, P., Hall, C., Jones, K., Sefton-Green, J., 2012. *The Signature Pedagogies Project: Final Report*. The Cambridge Primary Review Trust, Newcastle-upon-Tyne.
- United Nations Convention on the Rights of the Child (UNCRC), 1989. OHCHR Office of the High Commissioner for Human Rights, Geneva.
- Williamson, B., Hogan, A., 2020. Commercialisation and Privatisation in/of Education in the Context of Covid-19. ISSUU, Education International, Brussels.

Responding to and creating multimodal texts

Richard Beach^a and Blaine E. Smith^b, ^a University of Minnesota, Minneapolis, MN, United States; and ^b University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Concepts constituting responding to and creating multimodal texts	504
Social semiotics	504
Transmediation	504
Affordances	505
Emotions	506
Engaging in multimodal response and production in time and space	506
Unfolding vs. text-centric	506
Creating “digital geographies” of places	507
Fostering responding to and creating multimodal texts in the classroom	507
Responding to multimodal texts	508
Responding to images and illustrations	508
Responding to graphic novels and comics	509
Responding to digital videos	509
Creating multimodal texts	509
Using remix in creating multimodal texts	510
Digital storytelling	511
Engaging in meta-reflection and assessments of multimodal practices	512
Implications for instruction and assessment	513
References	513
Relevant websites	517

Glossary

Affordances How certain digital tools are employed to enact multimodality, collaboration/interactivity, ubiquity/authenticity, shared productivity, portability, and personalization/adaptivity in responding to and creating multimodal texts

Social semiotics Theory of how people define and interpret the social and cultural meanings of multimodal images and signs based on certain interest and ideologies

Transmediation The process of translating meaning across different sign systems to create new meanings, for example, transforming an oral narrative into a digital story through images or video

Much of literacy learning continues to privilege print literacies in ways that fail to recognize the marked shift in literacies constituted by multimodal digital practices shaping meaning-making within and outside of schools (Kress, 2010; Siegel, 2006; West, 2019). Students are increasingly responding to and creating multimodal digital texts—graphic novels, digital comics/storytelling, websites, podcasts, blog posts, videos, or digital art/music, given that adolescents average daily screen time is 7 h and 22 min (Common Sense Media, 2019). Understanding this shift entails redefining literacies based on multimodality theories that involve combining visual, audio, embodied actions, and print text for responding to and constructing meanings based on specific social and cultural norms (Serafini, 2011).

Engaging in responding and creating multimodal texts requires that students learn to employ a metalanguage for describing or naming specific elements of their texts. It also entails valuing popular culture texts by going beyond binaries between popular culture and canonical literary/mainstream texts to create engaging activities for critical literacy response to both types of texts. For example, in creating a digital comic for responding to *The Outsiders* (Hinton, 1967), students were just as focused on their creative literary responses to the novel than on the technological aspects of creating the digital comic (Weissman and Costello, 2014).

This engagement in the interactive, creative production of multimodal texts goes beyond focusing on visual/digital design to also applying critical media literacy practices related to how texts position audiences (Newell et al., 2016; Serafini, 2011). Serafini (2015) conceptualizes analysis and production of multimodal texts as not only “visual objects” but also as “multimodal events by focusing on elements of visual grammar related to perspectives, composition, use of frames, motifs, method of representation, associations, audience positioning, etc.” (p. 414).

This use of different modes also entails employing different visual/social literacies that are layered (Abrams and Russo, 2015; Chisholm and Whitmore, 2018). These layers include embodied literacies associated with ways of being (Claxton, 2015; Edmiston, 2014) and emotional literacies (Lewis and Crampton, 2016). For example, students responding to Anne Frank’s diary employed

visual responses by sharing images in a classroom, embodied reactions through the use of tableaux performances, and oral/written responses regarding their emotions associated with the Holocaust (Chisholm and Whitmore, 2018).

In doing so, students engaged in four different levels of responses. At Level 1, they viewed the texts to infer various meanings of the texts. At Level 2, students selected one photo to respond to the embodied meanings related to the people in the photo, such as the expression of their eyes to convey certain emotions—application of “embodied” and “emotional” literacies. At Level 3, students applied specific essential questions associated with the curriculum to respond to these texts. At Level 4, students shared their responses with peers in ways that fostered alternative perspectives on the multimodal texts (Chisholm and Whitmore, 2018).

Understanding the use of different modes also entails analyzing how people design and assemble multimodal texts over time using a range of various tools (Roswell, 2013). In her case study analyses of filmmakers, clothes/web/videogame, designers, book illustrators, musicians, composers, stage designers/actors, architects, and journalists, Roswell (2013) found that they drew on their childhood use of specific modes for creating multimodal texts inspired by previous workers’ texts. These experts also perceived their work as highly collaborative and that they could teach multimodal design practices to others.

Concepts constituting responding to and creating multimodal texts

There are several different concepts relevant to understanding how people construct meaning in responding to and creating multimodal texts.

Social semiotics

Underlying theories of multimodality is the concept of *social semiotics* related to defining the social and cultural meanings of different signs (Beemer and Kress, 2008; Hull and Nelson, 2005; Smith, 2019; Unsworth, 2008). Social semiotics theory focuses on using semiotic affordances of specific modes, for example, how an image assumes unique meanings within a particular setting as a “semiotic landscape” based on a participant’s knowledge and attitudes (Hewitt, 2008, 2014; Pantaleo, 2014). As described by Jewitt (2014), the “meanings in any mode are always interwoven with the meanings made with those of all other modes co-present and co-operating in the communication event” (p. 15). Modes are socially-constructed, and audiences interpret these modes’ meanings based on their ways of being and becoming. For example, students may associate the meaning of status-symbol images in advertising of cars, clothes, or beauty products related to their class or gender identities (Gee and Hayes, 2011).

Audiences also define social semiotic meanings based on their alliances with their social groups’ discourses, tastes, or ideological stances (Adami, 2015). For example, the homepages of two British food blogs, The Diary of a Frugal Family (FF) (<https://www.frugalfamily.co.uk>) and Thinly Spread (TS) (<https://thinlyspread.co.uk>), differ according to the appeal to different social classes. The FF blog employs a traditional linear, vertical blog format associated with print texts. The TS blogs use a more complex, modular layout consistent with online newspapers related to more contemporary online texts. The esthetic appeal of FF is “playful, for children, chaotic, spontaneous, amateur and authentic [while] TS instead shapes itself as minimalist, professional, carefully balanced and planned [can] be associated ... with the tastes of different class fragments and lifestyles” (p. 59).

Transmediation

A related concept is *transmediation* is associated with the translation of meaning across different sign systems to create new meanings, requiring students to engage in innovative, critical thinking (McCormick, 2011; Siegel, 2006). Transmediation involves going beyond reproducing knowledge or ideas, for example, retelling a story through drawings, to transforming meanings as in how drawing a storyboard for responding to or creating a video of a story requires thinking about multimodal implications (Mills, 2015; Smith et al., 2016b). Transmediation also involves redefining one’s intentions to exploit the affordances of the use of specific digital tools. For example, students creating a digital video using drawings in a storyboard consider the affordances associated with thinking about how those drawings mediate employing digital video/editing tools in creating a video.

Transmediation also involves a focus on knowing how to transfer multimodal features—images, video, audio, music, text, etc., across texts based on the use of particular design and esthetic practices (Beach et al., 2017). Through responding to visual effects, gutters, and speech bubbles panels in the graphic novel, *Around the World* (Phelan, 2014), students then transferred that experience to create their graphic novels (Napoli and Sychterz, 2015).

Transferring multimodal features as “transmedia navigation” involves students’ use of meta-language to reflect on reasons for employing certain features to achieve desired audience responses (Jenkins, 2006). In creating a video based on clips from a Romeo and Juliet film (Luhman, 1996), students employed meta-language to describe how they were using specific shots and angles to generate their video of a scene from the play (Burn and Durran, 2006; cited in Unsworth, 2011). Students may also reflect on how they employ certain similar practices across modes. For example, a student noted how in writing stories, she manipulates the development of time to generate suspense in the same manner as in creating music based on shifts in time (DePalma and Alexander, 2015).

It is also useful to consider how traditional print literacies apply to multimodal literacies related to transfer across different contexts. For example, learning to engage in the close reading of print texts may transfer to responding to multimodal texts, while close reading of multimodal texts transfers to reading print texts (McDermott et al., 2017).

Transfer across different contexts is more likely to occur when students are engaged in collaborative learning (Smith, 2017). College students used their Whiteboards to collaboratively critique and generate alternative versions of Play-Doh ads (Sobko et al., 2020). This focus on the use of multimodal digital tools expands notions of collaborative learning to include “(1) the convergence of human and non-human actants, (2) within a synchronous or asynchronous network, (3) facilitating and facilitated by shifts in knowledge and thinking, (4) in service of multiple and competing goals” (Sobko et al., 2020, p. 50).

Affordances

Another key concept related to how objects, artifacts, or digital tools serve to create multimodal meanings of texts is *affordances*. Affordances of navigational tools allow for automatic steering of airplanes, boats, or cars. Similarly, the affordances of particular objects, artifacts, or digital tools provide a means for engaging in specific actions or meaning-making (Harste and Kress, 2012). In responding to multimodal texts, for example, a website, students are drawing on the affordances of multimodality, collaboration/interactivity, ubiquity/authenticity, shared productivity, portability, and personalization/adaptivity for responding to how images, music, videos, or text on a website allow students to generate particular meanings (Beach, 2012) (see Fig. 1 below).

Students acquire knowledge of the multimodal affordances associated with the use of particular apps for responding to and creating texts (Beach and O’Brien, 2015a; Hutchison and Beschoner, 2014). For example, the ShowMe app allows students to import images to their devices to draw on an image and record an audio description of their drawing to share online. Students studying genetics in a science course created doodle images using ShowMe to portray cell mutations for presentations to their class (Beach and O’Brien, 2015a).

Students also draw on the affordances of digital tools to engage in collaborative, multimodal sharing of responses to and creating texts (Burnett and Merchant, 2018a,b; Leander and Boldt, 2013; Smith, 2019). Digital tools provided semiotic affordances associated with collaborative visual focus, shared embodied participation, and accessibility constituting creative, collaborative interactions between users (Newell et al., 2016; Rogers and Lindley, 2004; Wohlwend, 2017). 5-and-6-year-olds used iPads for doodling images in pairs using the iOS or Android Kids Doodle app.

In contrast to drawing doodles with paper and pen, students were more likely to engage in a shared gaze and physical proximity when working with the iPad than with their drawing (Sakr, 2018). Children were more engaged with the Kids Doodle app’s use given their surprise with shifting displays and color than with their drawings.

Students employed Diigo yellow sticky notes to collaboratively add and respond to each other annotations for responding to a text about measurements of emissions in a science class (Castek et al., 2014). Students perceiving each other’s sticky notes as including useful comments served as a sensory appeal. Students then wanted to pry open their peers’ notes to read each other’s annotations, leading to interactions between students in a collaboration conversation.

Students also employ screencasting apps such as ShowMe or VoiceThread to collaboratively respond with annotations to images or video clips (Beach and O’Brien, 2015a,b). Students enacted collaborative relations using VoiceThread to respond to students’ creation of images of clay figures to portray a chicken (Walsh, 2010).

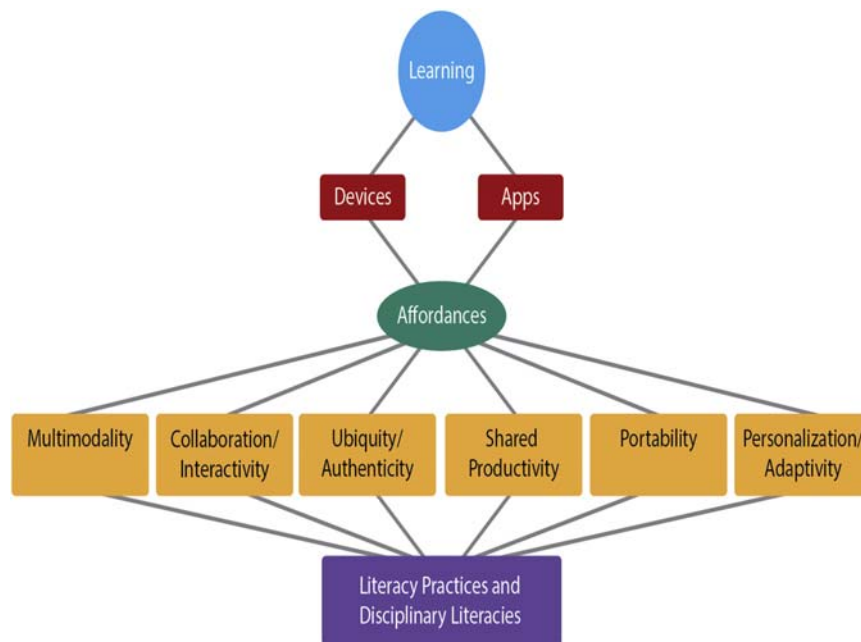


Fig. 1 Affordances constituting use of multimodal tools for responding to and creating texts.

Emotions

Students associate certain emotions as actions with the meanings of material images, objects, or physical locations in responding to and creating multimodal texts associated with continuities across time and space to address questions, such as “What is important to us? What is salient? What matters? What is desirable? What is judged to be normal versus novel? What possible relationships that we can imagine do we feel to be more probable or less so?” (Lemke, 2012a, p. 65).

A focus on resonance involves determining how certain practices are employed and shared across spaces and time based on how emotions serve to build connections across time as space (Stornaiuolo et al., 2017). For example, readers acquire the meaning of Harry Potter over time and space in responding to book covers, films, computer games, and commercial artifacts (Lemke, 2012b).

In creating animations using iPad apps, students identified how they portrayed the characters’ emotions by using specific visual images to engage their audiences (Mills and Unsworth, 2018). To do so, they learned to employ metalanguage for describing the expression of emotions such as sadness, anger, or joy through words, images, gestures, or embodied actions. Mills and Unsworth (2018) noted how their characters’ body language conveyed certain emotions as students drew and revised their animations to express those emotions.

Engaging in multimodal response and production in time and space

Researchers studying responding to and creating multimodal texts as literacy events recognize the need to perceive literacy events as open, porous, and continually evolving in time and space (Hollett et al., 2018; Pyles et al., 2019; Scollon and Scollon, 2003). Focusing on the open-ended nature of events leads to the analysis of a “literacy-as-event” enacting of positive relations through the use of different modes (Burnett and Merchant, 2018a,b, p. 9).

One challenge in thinking about response and creation of digital texts is the issue of a “text-centricity” framing of the meaning of texts as failing to consider how the meaning of texts is enacted through users unfolding, novel activity as shaped by emotions and embodied actions (Leander and Boldt, 2013). These “text-centric or discourse-centric perspectives on the production of subjectivity ignore movement and sensation [that] prompts us to interpret backward from texts to practices—to infer literacy practices from textual forms, even when other aspects of social practice are clearly evident” (Leander and Boldt, 2013, p. 33). The result is that “episodes of practice become dominated and misread as they are sifted and pushed through textual funnels such that much that characterizes practice become sorted out, pushed aside, ignored” (p. 33). In their analysis of two male adolescents responding to manga texts, the meaning of their responses derived from their social interactions enacted through unfolding, indeterminate, embodied actions.

Unfolding vs. text-centric

Participation in social media often involves unfolding, decontextualized interactions across time and space (Beach and Tierney, 2016). Posting videos online triggers unfolding responses, as evident in the example of an adolescent female posting a video on TikTok posing the question as to whether math is based on any reality resulting in 1.3 million viewing and millions more with reposting on Twitter (Haigney, 2020). This momentary, unfolding interaction is part of “the natural life cycle of a thing on the internet—a viral video, a meme, a piece of writing—to come loose from its context and take on a plethora of meanings ... this is not so much the relaxing of context as the willful ignorance of it” (Haigney, 2020, p. 8).

These momentary, unfolding experiences with multimodal texts suggests the need to study the response and production of multimodal digital texts within the context of users’ actual, unfolding, embodied activity instead of adopting a “text-centric” perspective. Analyzing these experiences involves adopting a “transliteracy” perspective to examine practices for responding to and creating multimodal texts as “practices of placemaking—ways of making meaningful spaces for learning through interactions with others, with things, and with the natural and built environment” (Stornaiuolo et al., 2017, p. 78).

Students also engage in these experiences by transferring their literacy practices in reading print texts to reading multimodal texts such as graphic novels based on images, captions, panels, and speech balloons (Gillenwater, 2014) as well as practices for remixing and recontextualizing digital texts in creative, multimodal ways (Domingo, 2014). For example, a student composed a digital story using Snapchat to portray images of his experiences of daily everyday events (Wargo, 2015).

Students participating in virtual spaces also engage in alternative multimodal practices for interacting with others in time and space. Young students interacted in the three-dimensional, virtual world of “Barnsborough,” as a deserted town in which they, as avatars, engage in online chat as they move around different spaces and discover clues as to why the town is deserted (Burnett and Merchant, 2014). One student, John, used the virtual chat log to interact with avatars in the town and talk with his lived-world peers in the classroom as different layered literacies based on his previous experiences in various virtual and lived worlds. Analysis of this virtual world suggests the importance of examining movements across different times and spaces associated with the virtual and blurred material (Burnett and Merchant, 2014).

In another example of an experience of multimodality in time and space (Leander and Hollett, 2013), a team of literacy researchers shared their responses to the graphic novel *Th3 Anomaly: Crossing the Rubicon* by the author/artist, Landry (2016) (Leander et al., 2017). Rather than a text, Landry created the novel in the form of 300 paintings housed in an art gallery to create an interactive space for viewers’ embodied responses to the paintings. As one researcher (Roswell) noted, in this novel “modes are

culturally shaped and subject to cross-cultural ways of knowing and seeing. This presents a challenge to the text-based stance of multimodality. Instead, multimodal meaning-making can be linked to the movement of bodies in space” (Leander et al., 2017, p. 108).

Creating “digital geographies” of places

Redefining notions of virtual spaces and time entails recognizing how people engage in “mediatized worlds” (Hepp and Krotz, 2014, p. 8), constituted by virtual space and time transcending the local through interactions on a global level (Beach and Tierney, 2016). Students employ digital tools or apps to generate alternative, multimodal representations as “digital geographies” mediating their mobile participation within and across specific sites and spaces, as evident in the use of GPS tools on devices/phones related to “neogeography” multimodal mapping practices (Burnett et al., 2014; Hollett et al., 2018, p. 152).

Users of mobile device apps such as FourSquare and Yelp can also create digital texts based on specific locations associated with layering annotations or texts to location sites (Hollett et al., 2018). Students employed VR tools such as Google Street View to engage in virtual field trips associated with studying a specific location related to what they are studying. For example, the “Poetry 4 U Project” connects poems to specific locations to enhance audiences’ sense of place associated with a particular poem (Berry and Goodwin, 2013).

For the Visual Communication Design program at the Manukau college campus, South Auckland, New Zealand, students wrote poems or selected a poem related to a specific location with the Manukau region associated with certain emotions, autobiographical events, cultural associations, etc. (Sinfeld and Cochrane, 2018). They also created images/videos, using AR tools such as Wikitude and Google Street View to add data to the images/videos related to mapping their location. Students noted that using these tools enhanced audience interactivity with their productions and how they experienced novel, innovative use of tools to generate multimodal experiences. Instructors at the RMIT School of Education, Australia, created a virtual school space, Lathner Primary, that allows preservice teachers to interact with teachers and their peers regarding their instructional practices (Faulkner and Latham, 2010). Students at RMIT University share their multimodal digital texts using the Climate Action Adventure! web application that combines virtual learning contexts with social media platforms and climate data to address climate justice issues (Rousell et al., 2017).

Students add data or information to Wikimapia.org or OpenStreetMap that involves users and audiences in “new forms of reading and writing the world—a new ethos that goes beyond the digital media space afforded by previous, sedentary technologies” (Hollett et al., 2018, p. 155). They can also experience multimodal representations of climate change effects related to extreme weather events, droughts, melting glaciers, sea rise, or flooding (Beach and Smith, 2019; Starsielski and Walker, 2016). For example, in the Visualizing Change project, students created multimodal, 3D representations of change over time. They could then portray melting glaciers or sea rise based on a relatively rapid rate of change (O’Neill and Smith, 2014). They also used Google Earth Timelapse tool to analyze changes in glaciers melting over three decades. Students at the University of Lisbon, Portugal, created digital art for portraying their concerns about climate change (<https://artforadaptation.com>) (Bentz, 2020).

Teachers have also employed virtual reality (VR) and alternative reality (AR) tools to have students engage in a virtual spatial experience that can then transfer to their analysis of lived-world topics or events. For example, in an “ecomobile” project, students collected data from a virtual pond to then analyze the water quality of a virtual pond to then apply to their classroom analysis of water quality (Kamarainen et al., 2013). Preservice teachers at the University of Arizona studying climate change visited the Biosphere 2 (B2) lab biosphere2.org for studying the earth’s ecological systems and biospheres (Beach et al., 2020). They collected audio, video, and 360 interactions for their digital storytelling productions for use in their teaching to create a video of their tour, as well as developed a lesson on growing food on Mars.

Fostering responding to and creating multimodal texts in the classroom

Fostering students’ engagement in responding to and creating multimodal texts in the classroom entails adopting practices beyond literacy instruction based on print-based literacies associated with adopting alternative activity systems (Kalantzis and Cope, 2012; Shanahan et al., 2014). This involves analyzing how the affordances of tools serve to achieve certain outcomes within a community constituted by adopting different roles and adherence to specific rules. For example, students in a fifth-grade science class studying the environmental effects of acid rain used Hyperstudio for users to click on images to access information about these effects. A ninth-grade teacher used a music video to model the creation of multimodal poems using audio and visual tools (Shanahan et al., 2014).

Smith (2014) reviewed 76 research studies on instructional activities for fostering adolescents’ multimodal composition. Students experienced higher engagement levels when they could choose their topics associated with a focus on popular culture, music, manga/anime, fanfiction, or video games associated with personal issues in their lives. They benefited from collaborative production through learning to assist each other in using digital tools to achieve a group’s goals. Students from non-dominant backgrounds used their multimodal productions to portray their lives in different cultural contexts in ways that were not possible through traditional, print-based literacies.

Smith’s review also found that teachers and peers assume important roles in providing scaffolding related to the use of digital tools for generating multimodal productions by acquiring a metalanguage for reflecting on the use of tools. Students also learned to

engage in recursive production/composing processes associated with planning, videotaping, and editing over time. However, both teachers and students in these research studies experience tensions and challenges associated with acquiring production practices requiring the need to not romanticize their experiences as an idealized alternative to traditional print-based instruction (Hull, 2003).

For teaching *Macbeth* based on the theme of curses, students generated and shared video presentations that drew on clips from movies portraying curses (Marlatt, 2019a,b). In the classroom, the teacher adopted a more facilitator role instead of a lecturer role through modification and redefinition of instruction. Simultaneously, one challenge related to using multimodal digital tools in the classroom has to do with the disparities in online access and the use of digital devices/apps related to economic and class inequalities (Mills, 2010).

Responding to multimodal texts

In responding to online digital texts, students are engaged in clicking links to explore other texts in a dynamic, unfolding process so that texts, consistent with a transactional theory of response (Rosenblatt, 1993) are no longer isolated, decontextualized entities (Pet et al., 2014). Navigating multimodal texts involves interacting with a text by using different pathways or processes based on a text's design features related to layout, subheads, visual images/graphics, etc. Responding to multimodal texts involves adopting the roles of "(1) navigator, (2) interpreter, (3) designer, and (4) interrogator" images related to the use of a text with a particular social, ideological, or cultural context (Serafini, 2012, p. 153).

Loh and Sun (2019) found that students who enjoy reading were more likely to use their phones or devices for reading e-books or articles. In contrast, students who did not enjoy reading were more likely to use phones/devices to interact on social media or play games. Students are also increasingly responding to and creating texts within and outside the classroom through the use of multimodal practices, instead of relying solely on print texts, requiring a marked shift in instruction that builds on the use of these multimodal practices (Beemer and Kress, 2008; Cartner and Hallas, 2020; West, 2019).

Responding to images and illustrations

Students benefit from instruction on inferring intertextual links between images and texts through understanding how images and language portray events. In contrast to reading instruction in the Western world based primarily on print words/texts, children in China acquire initial literacy practices primarily through immersion in images in texts designed to socialize them about schooling (Bin and Freebody, 2010).

For responding to children's books or graphic novels, students learn how to draw on their semiotic knowledge of images to recontextualize the meaning of images and illustrations related to books' or novels' narrative development (Lysaker, 2018; Pantaleo, 2013). In responding to the wordless book, *The Other Side* (Woodson and Lewis, 2001), a fourth-grade student engaged in intersubjective retelling of images portraying events as a co-author based on his lived-world experiences associated with a relational sense of belonging with the characters (Lysaker and Wessel-Powell, 2019). Middle-school students responded to the use of images in two children's books to support the story development based on how the books positioned them as readers (Unsworth and Macken-Horarik, 2015).

In addition to responding to images in texts, students also use images to enact their responses, for example, generating drawings or sketches on large pieces of paper for responding to *The Odyssey* (Zoss, 2019). For responding to a picture book, *The Red Piano* (LeBlanc and Barroux, 2009) portraying a young girl in a Chinese re-education camp, students used an iPad camera to create photos of pages from the book. They then used the Popplet app to generate written responses to images within boxes to connect the boxes to specific aspects of the book's images. They may also use audio responses, as did students using iPads to add the images and sounds to a description of a windstorm from the novel, *The Big Dry* (Davis, 2013), including writing their descriptions of their images/sounds (Zammit, 2016).

After freewriting about a poem, students created multimodal responses to the poem by listing words or phrases related to a poem's perceived tone (Pet et al., 2014). They then selected images to create tableaux that arranged images to convey their experience of a poem's tone leading to writing another freewrite comparing their initial response to their new response (Pet et al., 2014). They then used these responses to create their metaphor poems. Their use of multimodal responses and their creation of poems fostered visual/audio experiences, intertextual, affective connections, multiple layers of meaning, and enhanced discussions.

Students also use artwork—for example, for responding to *The Pearl* (Steinbeck, 1947), by creating a wire sculpture with a hand holding some pearls to portray characters' obsession with obtaining the pearls (White, 2019). To portray conflicts between the two brothers in the novel *Fighting Ruben Wolfe* (Zusak, 2000), a student drew images that included boxing gloves and a heart to portray the brothers' love for each other (White, 2019). To portray changes in the main character, Melinda, over time in *Speak* (Anderson, 1999), students used layering of translucent tracing paper with images of drawing of trees that portray rougher and harsher aspects of the tree to show Melinda coping with difficulties in her relationships over time (Hartman et al., 2019).

As part of their participation in literature circles, students engaged in producing short video productions for responding to texts such as *Into the Wild* (Krakauer, 1996) about how the central person, Christopher, made mistakes in exploring the Alaskan wilderness resulting in his death (Marlatt, 2019a,b). Students created a satiric film posing as a guide for surviving in the wild as a mock documentary portraying mistakes made by the main character.

Students also draw on their prior experiences and expectations with multimodal texts to engage with current texts, for example, knowledge of both multimodal and game literacies based on specific design features for engaging in actions as they move through interactive literary texts or games (Hsu and Wang, 2010; Partington, 2010; Wyatt-Smith and Elkins, 2008). Students responding to the Delight Games interactive fiction app drew on game social semiotic practices for making narrative choices based on their previous experience with multimodal game semiotics, including the use of embodied actions related to their participation as avatars (Hall, 2019).

Responding to graphic novels and comics

Graphic novels and comics continue to be popular for young readers, particularly given how the combination of images and words, along with the use of panels and gutters, serve to scaffold and motivate readers' experience (Botzakis et al., 2017; Griffith, 2010). In response to a traditional print novel versus graphic novel format, sixth graders preferred the graphic novel format (Moeller, 2016).

Readers learn to infer how multimodal, visual features of light, color, lines, and shadow serve to engage readers as co-authors (Smetana et al., 2009; Low, 2012). Teachers have also employed graphic novels versions of canonical literature, such as the graphic novel version of *Othello* (Zarate, 2011), that provides visual scaffolding to help students respond to the play (Wolfe and Kleijwegt, 2012). Analysis of eighth-grade students' think-aloud responses to multimodal features in *Episode 2: Italy in Inanimate Alice* (Pullinger and Joseph, 2012), found that students focused on the images and sounds/music as a means of understanding the text (Karchmer-Klein and Shinas, 2019).

For responding to graphic novels in a college course, students created video essays that included content from both the novels and online video essays related to engaging images, minimal use of written text, and use of audio/voice to convey key points (Miller, 2019). In a related study, students received instruction in design/art elements with layout, perspectives, framing, color, typeface, framing, panels, speech bubbles, and physical construction of graphic novels (Pantaleo, 2013). For responding to the graphic novel, *Sidekicks* (Santat, 2011), students focused on using color, point of view/perspectives, or exaggerated character poses/expressions related to enhancing narrative plot intensity and mood (Pantaleo, 2013). Students used Comic Life to create comic book versions of the novel *The Outsiders* (Hinton, 1967) in ways that enhanced their aesthetic transactions with the novel (Weissman and Costello, 2014).

Responding to digital videos

A survey of 1600 US 8–18-year-olds found that they devote an average of 1 h daily to viewing online videos (Rideout and Robb, 2019). This raises the question of whether, in viewing these videos, students are focused not only on the use of cinematic video techniques but also on adopting a critical response to underlying cultural or political stances adopted in these videos. In one study, students in an EFL class in a German high school responded to digital stories on YouTube portraying undocumented youth in the US and how they used the media to challenge conservative political framing of immigration issues (Ruhe, 2020). Students responding to three digital stories used emojis and comments to share what they learned about undocumented youth's immigrant's lives; they also searched online posts by undocumented youth. The teacher then provided students with "mode cards" to have students identify different modes employed in the videos related to embodied movements and cinematic techniques, auditory language, captions, and sound related to how these modes serve to convey specific political messages regarding immigration issues.

Students could engage in critical responses to videos using questions from the Center for Media Literacy (<https://www.medialit.org>):

1. Authorship: "Who created this message?"
2. Format: "What creative techniques are used to attract my attention?"
3. Audience: "How might different people understand this message differently than me?"
4. Content: "What values, lifestyles and points of view are represented in, or omitted from, this message?"
5. Purpose: "Why is this message being sent?" (McClain, 2016, p. 40).

Creating multimodal texts

Students also employ tools and apps for creating multimodal texts, such as visual images in argumentative writing (Roque, 2012). In their systematic review of 70 studies in secondary schools, Smith et al. (2021) found that teachers integrate a wide variety of multimodal projects, digital videos, multimodal presentations, and visual projects podcasts, and webpages being the most prevalent. Creating multimodal texts requires instruction for students who were accustomed to writing text-based arguments on Google Docs and their lack of experience with using digital tools on how to employ multimodal approaches to generate their PSA. However, students indicated that they were more engaged with completing the assignment than writing print-based argumentative essays given the higher levels of creativity associated with making choices about multimodal options and that their audiences were more apt to access their productions. At the same time, there was no evidence that their experience in generating multimodal arguments transferred to writing an argumentative essay (Howell et al., 2017).

Teachers in Nepal participated in a project involving creating videos about environmental issues in their communities, for example, lack of water, pollution, and adverse effects of garbage (Walker, 2017). In reflecting on their experiences, these educators

noted that it was through the process of producing videos itself that enhanced their critical awareness of sustainability challenges as a form of conscientization (Freire, 1970) as contrasted with simply studying these challenges, implying that the experience of multimodal production enhances the need to make change in society. Similarly, students who created videos portraying fake news examples were more likely to acquire knowledge about fake news and a sense of political efficacy than students who studied these videos or students who studied the topic of fake news (Geers et al., 2020).

Several studies have shown how students can analyze literature through creating multimodal texts (Beach et al., 2017; Curwood and Gibbons, 2009; Jocius, 2013). By composing through multiple modes, Smith (2018, 2019) found that multilingual youth could explore literary themes and the affective dimensions of literature. For example, one group of 10th-grade bilingual students expressed what they viewed as the main theme of “oppression versus freedom” in the short story, Harrison Bergeron (Vonnegut, 1961), multimodally in a digital video. They filmed creative scenes of themselves wearing masks and layered voiceover narration with dramatic music to emphasize their interpretation. Through multimodal literary analyses, students also had the semiotic space to express their identities, appeal to different audiences, and make cultural connections to other media (Smith, 2019) (see Fig. 2 below).

Teachers have students employ zines as hand-drawn texts about students’ lives as an alternative to the written essay, in ways that “give them a chance to play—with form, voice, and message ... that leads them to share a highly curated version of their lives” (Collopy, 2020, p. 28). As one teacher noted, “I think [zines] inhabit this interesting space between public and private and between permanent and ephemeral that is still a meaningful one” (Collopy, 2020, p. 29). (Aragon et al., 2018, includes extensive resources on teaching zines).

For sharing multimodal narrative writing, students employ story-sharing apps such as Wattpad or Figment for adding or responding to others’ narratives (Korobkova and Collins, 2018). On these sites, older adolescents perceived themselves more as readers and novice writers than the professional writers on these sites than younger adolescents who perceived themselves more as writers who use their peer groups to identify their writing ability (Korobkova and Collins, 2018).

Using remix in creating multimodal texts

A primary practice in creating multimodal texts involves using remix of texts through alternative modes (Lankshear and Knobel, 2011). For engaging in a remix, students are drawing on how genres and memes operate in specific social contexts to play with or subvert these genres and memes (Bowen and Whithaus, 2012).

Generating remixed texts involves consideration of how different modes interact with each other in different ways. In a professional development seminar, teachers created multimodal memoir writing using a remix that included images, video clips, music, or audio podcasts to portray specific events or episodes from their memoir (Meixner et al., 2018). Teachers employed four different interactions between modes: (1) “symmetrical: two modes share the same message,” (2) “contradictory: two modes connote opposite meanings; establish tension,” (3) “complementary: two modes interdependent; fill in the gaps,” and (4) “metatextual: two modes explicitly show the act of composing: highlight design” (p. 500).

The increased use and popularity of YouTube videos have led to remixing and re-appropriation of video content related to “spreadability” (Jenkins et al., 2013), often based on producers and audiences’ shared knowledge of specific genres or memes. This shared knowledge leads to generating unique remixes of videos within a genre or meme (Adami, 2014). For example, a video, “The Harlem Shake,” of a man dancing to a song followed by a jump cut to others dancing spawned the Harlem Shake Meme with thousands of video responses from around the globe, including critiques of the expropriation of African American cultural practices by Whites (Koehler, 2013).

One primary mode in remix involves audio as music, voiceovers, sounds, etc. that contribute to the meaning of a remix based on “embodied knowledge” constituting the meaning of specific spaces (Gershon, 2011). Productions on the Youth Radio podcast as



Fig. 2 Scene from a video created by 10th-grade students who analyzed literary themes from a short story (Smith, 2019).

“soundscapes” includes remix use of music, sound effects, spoken-word poetry, and digital photography as “converged literacy” (Soep and Chávez, 2010, p. 47).

In creating remix video book trailers, one student employed both slow and fast/loud music to accompany her trailer’s portrayal of Holden’s experiences in *The Catcher in the Rye* (Salinger, 1951). Another student included images of African Americans living in poverty for responding to *Native Son* (Wright, 2005) (Baize, 2019). Students also draw on experiences with popular music videos to create remixes based on embodied experiences and images portrayed in music videos (Bazalgette and Buckingham, 2013). For example, a group of high school students created a video depicting a student being transformed into a character in a graphic novel world to engage in a dream of his peers dancing to popular music video songs in different spaces in the school (Beach and Swiss, 2010).

Students also employ multimodal remix productions for civic engagement related to addressing issues in their school or community (Mirra and Garcia, 2017). Students in one school created a bulletin board about the portrayal of women of color in the media that included texts with QR code links to a YouTube video remix for display in their school’s hallway (DeJaynes and Curmi-Hall, 2019). The students also enacted counter-storytelling through photos of themselves with bar codes over their eyes and words associated with stereotypes associated with their racial identities. For example, one Asian student included the terms, “Martial Arts,” “Straight A’s,” “Anti-Social,” “Dragon Lady,” “Small Eyes,” “Glasses,” “Chinese,” next to her photo (p. 304).

Researchers have also identified different types of remixing practices. Analysis of teachers interacting with each other on a MOOC course using Twitter, Google + Community, Facebook, and blog posts employed four types of remix practices (Smith et al., 2016a,b). “Bursting” and “drifting” were associated with the proximity and pace of remixing. When teachers received a newsletter that contained images with some prompts, “What sense do you make of the story? If you were to put words to it, what would it say? What songs would it sing?” one teacher created a remix of these images using his images and adding some music. Other teachers also created their remixes for enacting collaborative stances. Teachers in the course also engaged in “drifting” through innovative, experimental uses of apps, such as the use of emojis and animated GIFs with images.

Teachers also employed “leveraging” and “turning” practices that served to enact social relationships as well as disrupting power hierarchies. They used “leveraging” practices by employing tweetchats and Google Hangouts and making explicit how they work collaboratively as a community by using these tools. There were also instances of “turning” related to ruptures regarding the overall purpose of their interaction. One teacher posted images associated with immigration, leading to debates about the relevancy of that issue within their CLMOOC community.

In designing instruction for engaging students in remix, teachers employed an authoring tool (Strategy Tutor, developed by CAST, Inc.) that involved remixing different prompts, media, and agents (Dalton and Smith, 2014). In an analysis of 26 teachers who published their lessons to a publicly available database, Dalton and Smith (2014) found that teachers integrated a wide variety of modes in media in service of literature-based or content-based learning goals. Three design approaches and goal orientations emerged. Most teachers provided students with a multimodal immersion experience to enhance their understanding and engagement with literature or other subject-matter learning. These teachers took advantage of the media on the Internet, linking to websites that included at least bimodal representations, and more typically, multimodal representations.

Digital storytelling

Students also engage in digital storytelling through the use of resources from the Center for Digital Storytelling (<http://www.storycenter.org>), Digicom Learning (<http://digicomlearning.com>), StoryCenter (<http://www.storycenter.org>), or the University of Houston’s Center for Digital Storytelling (<http://digitalstorytelling.coe.uh.edu>). In creating digital stories, students draw on remix practices to combine images, video clips, audio, and text to create settings, characters, and narrative development.

Middle school students working in teams as part of the Project Imagine the Future (<http://imaginefuture.org>) assumed roles of writers, scientists, and designers to collaboratively create digital multimodal science fictions remixing comics, pictures, videos, infographics, and Scratch often related to climate change (Jiang et al., 2019; Smith and Shen, 2017). A key component of their productions has to do with students’ collaboration by assuming roles related to their expertise in using certain multimodal literacy practices in ways that helped them envision future identities. One student who assumed the role of scientist noted that “I want to be a forensic scientist. My character [in the sci-fi narrative] also became a scientist after saving the world” (Smith et al., 2019, p. 54) (See Fig. 3 below of students’ book covers).

Fifth-grade students composed digital stories in response to reading historical fiction novels (Kesler et al., 2016). They employed certain multimodal tools to portray their responses, including examples of other stories, use of storyboards, websites for free access to images, music/sounds, and digital cameras/voice recorders. One group responding to *The Gold Cadillac* (Taylor, 1987) created a digital story based on country and Chicago blues music and drawings and slides to portray a family’s move from Ohio to the segregated world of Mississippi in the 1950s. The slides include images of signs of “WHITES ONLY!” “COLORED WAITING ROOM,” and “WELCOME TO MISSISSIPPI/IT’S LIKE COMING HOME” (p. 59).

Students creating a digital story for *Roll of Thunder, Hear My Cry* (Taylor, 1997) portrayed children in the Logan family having to walk three and a half miles to and from their school while the White children road a school bus (<http://tinyurl.com/koo9hlp>). One of the images is a drawing of the White school bus driver frowning through the bus window as he drives the bus down a muddy dirt road in a storm to imply his sense of power. A subsequent slide portrays the bus going by the Logan children, who then scramble into a gully to avoid the mud splashing on them with music from *Jaws* playing in the background.

A research review identified reports of the use of digital storytelling and photovoice exhibits for participation in civic engagement projects by documenting issues in their schools or communities, leading to addressing those issues (Greene et al., 2016, 2018).



Fig. 3 Multimodal science fictions created by middle school students that explored socioscientific issues (imaginefuture.org).

These reports include students challenging stereotypes in schools (Rogers et al. (2014); declining physical conditions in neighborhoods in Detroit (Graham et al., 2013), Pittsburgh (Teixeira, 2015), and Latinx communities in southwest Texas (Guajardo et al., 2016); homelessness in Los Angeles (Garcia et al., 2015); and the experience of Native American students in Alaska schools (Wexler et al., 2013). They also include examples of middle school students of color engaged in a YPAR project to document how they engage in making change (Pellegriano et al., 2016).

Teachers in the Palm Springs school district reported using digital stories to generate texts as both fiction and nonfiction to enhance writing, learn collaboration, assess learning, and use projects to connect to their communities (Vu et al., 2019). Surveys of students who engaged in creating these digital stories based on this instruction indicated higher technology self-efficacy levels than peers who did not create digital stories. These students learned to employ a wide range of digital tools and apps and engaged in digital production practices. Students creating a digital story in response to citizen surveillance issues, as portrayed in 1984 (Orwell, 1950), demonstrated more insightful responses to the novel and retained more information about the novel than students who wrote an essay about the novel (Vu et al., 2019).

Engaging in meta-reflection and assessments of multimodal practices

For teaching composition, Jody Shipka (2011, 2012) problematizes equating multimodality using only certain kinds of new-media digital texts, for example, use of a digital story or blog post. She posits that using only certain kinds of digital texts leads to “freezing” writing as a static entity instead of writing as a set of unfolding literacy practices for engaging in action (Prior, 2009).

Given that students engage in multiple rhetorical decisions as they compose texts, Shipka asserts the importance of focusing on and having students engage in meta-reflection awareness to make explicit how they select and employ certain multimodal practices. She favors giving students problems that “require students to consider how the contexts in which texts participate shape the way those texts are received and responded to ... [through] attending to how language, combined with still other representational systems, mediates communicative practice” (2012, p. 78). For example, one of her students, Muffie, a dancer, worked on creating a video about her dancer peers (Shipka, 2011). To reflect on her planning processes, Muffie generated a set of drawings depicting instances in which, for example, she viewed a similar, previously made documentary by another student about dance and how she attempted to organize her peers for engaging in dance.

For reflecting on multimodal video production in a college composition course, a student envisioned how her appeals to online audiences related to her need to make revisions with those audiences in mind (VanKooten, 2016). She noted that “editing is revision; if there is one thing I have learned through English, it is that revision is as important as actually writing the paper” (p. 68). The student also described the importance of her shifting her use of music: “my favorite part [of the video], I swear, is the part where there’s music and it fades away. I like the seriousness of the moment, and I think I found a new appreciation for a shift in a paper” (p. 68).

Students can engage in this meta-reflection on how their multimodal texts achieve certain uptakes related to engagement in certain rhetorical context using criteria based on the Multimodal Assessment project: artifact (finished product), contexts, substance (content, quality, and significance of ideas presented), process management and skills, and habits of mind (National Writing

Project, 2011). For recording think-aloud reflections about processes for creating multimodal texts, students can use screencasting apps such as ShowMe or Explain Everything (Abrams, 2014).

Using criteria for meta-reflection requires going beyond framing context as conforming to specific formalist criteria to how their multimodal texts serve to create contexts for building relations with audiences (Gallagher, 2014). For example, a student reflected on creating a multimodal e-portfolio about the limitations of traditional schooling as itself a digital book with an introduction, cover, preface, and tabs (Gallagher, 2014). She noted how the rhetorical context could “cause the viewer to interact with images of historical and more outdated photographs of the education system as well as more modern ones” (p. 7).

Teachers can also create videos to model self-assessments. In a course for teachers, teachers generated blog posts regarding how they would employ multimodal tools for their instruction by reflecting on how the multimodal course design itself fosters their learning (Cartner and Hallas, 2020). Teachers noted the value of the visible course design based on criteria for their self-reflection associated with outcomes and activities based on teachers’ practical needs, for example, an activity in which they addressed two questions related to teaching with multimodal texts.

Implications for instruction and assessment

One of the major challenges for instruction and assessment is that much of schooling continues to revolve around verbocentric, print literacy practices, and evaluations in ways that marginalize responding to and creating multimodal texts that are now prevalent in students’ lives outside of school (DePalma and Alexander, 2015; Quinlan and Curtin, 2017; Ranker, 2015).

Another instructional challenge for engaging students in responding to and creating digital media is that approximately fifteen to sixteen million students, particularly for students in rural and/or low-income homes, and 300,000 to 400,000 K-12 teachers, lack adequate Internet access and/or owning devices needed for responding to or creating digital media (Chandra et al., 2020). For example, 26% of Latinx, 30% of Black, and 35% of Native American students lack access and/or devices. Having adequate access involves having a minimum of 25/3 Mbps (download/upload speeds) given that it takes 3 min to load a 30-min video in contrast to 9 min for a slow 10/3 Mbps. This suggests the need for both state and federal governments as well as digital/media companies to provide students with free or subsidized access to address these inequalities.

Assessing students’ multimodal responses and productions entail the development of criteria that differs from use in assessing responses to and production of print texts, for example, assessing students’ creation of advertising videos based on criteria related to attention, interest, desire, and action (Silseth and Gilje, 2019). In responding to the same assignment, one student employed more verbocentric writing consistent with his need to conform to the teacher’s evaluation based on composing process models, while another student employed more multimodal productions through a video consisting of thousands of photo images that were less consistent with the teacher’s criteria (Jocius, 2017).

One factor limiting a shift to multimodal instruction is that some large-scale standardized testing driving school curriculum is based primarily on print-based literacies. For example, while the Australian National Curriculum includes a strong emphasis on teaching multimodal literacies, the reading tests of the Australian National Assessment Program in Literacy and Numeracy (NAPLAN) contained little focus on multimodal literacies (Unsworth et al., 2019). The *Making News Today* project at the University of Wollongong, Australia, employed an alternative pre-post assessment of multimodal practices that includes images and texts from news and media. Students responded to open-ended questions requiring students to generate written answers on a website leading to assessing students learning overtime based on ratings of the extent to which students engaged in formulating and supporting claims and used intertextual connections (Brown et al., 2010).

This increased use of assessment that includes responding to and creating multimodal/digital texts will only enhance an increased instructional focus on responding to and creating multimodal/digital texts central to students’ daily lives within and outside of school.

References

- Abrams, S.S., Russo, M.P., 2015. Layering literacies and contemporary learning. *J. Adolesc. Adult Literacy* 59 (2), 131–135.
- Abrams, S.S., 2014. Integrating Virtual and Traditional Learning in 6–12 Classrooms: A Layered Literacies Approach to Multimodal Meaning Making. Routledge, New York, NY.
- Adami, E., 2014. “Why did dinosaurs evolve from water?”: (In)coherent relatedness in YouTube video interaction. *Text Talk* 34 (3), 239–259.
- Adami, E., 2015. Aesthetics in digital texts beyond writing: a social semiotic multimodal framework. In: Rijlaarsdam, G., Archer, A., Breuer, E. (Eds.), *Studies in Writing: Vol. 30, Multimodality in Writing*. Brill, Leiden, pp. 43–62.
- Anderson, L.H., 1999. *Speak*. Farrar Straus Giroux, New York.
- Aragon, K., Adams, D., Braun, J., Fernhout, E., Huddleston, J., McElroy, K., et al., 2018. Teaching With Zines. Retrieved from. <http://bit.ly/2H4W25j>.
- Baize, J., 2019. Classics in their own “words”: analytical remixes in a land of essays. *J. Adolesc. Adult Literacy* 62 (6), 625–633.
- Bazalgette, C., Buckingham, D., 2013. Literacy, media and multimodality: a critical response. *Literacy* 47 (2), 95–102.
- Beach, R., O’Brien, D., 2015a. Using Apps for Learning Across the Curriculum: A Literacy-Based Framework and Guide. Routledge, New York, NY.
- Beach, R., O’Brien, D., 2015b. Fostering students’ science inquiry through app affordances of multimodality, collaboration, interactivity, and connectivity. *Read. Writ. Q.* 31 (2), 119–134. <https://doi.org/10.1080/10573569.2014.962200>.
- Beach, R., Smith, B.E., 2019. Using digital tools for studying about and addressing climate change. In: Sullivan, P.M., Lantz, J.L., Sullivan, B.A. (Eds.), *Handbook of Research on Integrating Digital Technology With Literacy Pedagogies*. IGI Global, Hershey, PA, pp. 346–370.

- Beach, R., Swiss, T., 2010. Digital literacies, aesthetic, and pedagogies involved in a digital video production. In: Albers, P., Sanders, J. (Eds.), *Literacies, the Arts, and Multimodality*. National Council of Teachers of English, Urbana, IL, pp. 300–320.
- Beach, R., Tierney, R.J., 2016. Toward a theory of literacy meaning making within virtual worlds. In: Israel, S. (Ed.), *Handbook of Research on Reading Comprehension*, vol. 2. Routledge, New York, pp. 135–164.
- Beach, R., Castek, J., Scott, J., 2017. Acquiring processes for responding to and creating multimodal digital productions. In: Hinchman, D., Appleman, D. (Eds.), *Adolescent Literacies: A Handbook of Practice-Based Research*. Guilford Press, New York, pp. 292–309.
- Beach, R., Boggs, G., Castek, J., Damico, J., Panos, A., Spellman, R., Wilson, N., 2020. Fostering preservice and in-service ELA teachers' digital practices for addressing climate change. *CITE* 20 (1). Retrieved from: <https://tinyurl.com/y7zylobb>.
- Beach, R., 2012. Uses of digital tools and literacies in the English language arts classroom. *Res. Sch.* 19 (1), 45–59.
- Beemer, J., Kress, G., 2008. Writing in multimodal texts: a social semiotic account of designs for learning. *Writ. Commun.* 25 (2), 166–195. <https://doi.org/10.1177/0741088307313177>.
- Bentz, J., 2020. Learning about climate change in, with and through art. *Climatic Change* 16 (2), 1595–1612. <https://doi.org/10.1007/s10584-020-02804-4>.
- Berry, M., Goodwin, O., 2013. Poetry 4 U: pinning poems under/over/through the streets. *New Media Soc.* 15 (6), 909–929.
- Bin, B., Freebody, P., 2010. Image, genre, voice, and the making of the school-literate child: lessons from multiliteracy teaching in China. In: Cole, D.R., Pullen, D.L. (Eds.), *Multiliteracies in Motion: Current Theory and Practice*. Routledge, New York, NY, pp. 42–58.
- Botzakis, S., Savitz, R., Low, D.E., 2017. Adolescents reading graphic novels and comics: what we know from research. In: Hinchman, D., Appleman, D. (Eds.), *Adolescent Literacies: A Handbook of Practice-Based Research*. Guilford Press, New York, NY, pp. 310–322.
- Bowen, T., Whithaus, C., 2012. "What else is possible": multimodal composing and genre in the teaching of writing. In: Bowen, T., Whithaus, C. (Eds.), *Multimodal Literacies and Emerging Genres*. University of Pittsburgh Press, Pittsburgh, pp. 1–12.
- Brown, I., Lockyer, L., Caputi, P., 2010. Multiliteracies and assessment practice. In: Cole, D.R., Pullen, D.L. (Eds.), *Multiliteracies in Motion: Current Theory and Practice*. Routledge, New York, NY, pp. 191–206.
- Burn, A., Durran, J., 2006. Digital anatomies: analysis as production in media education. In: Buckingham, D., Willett, R. (Eds.), *Digital Generations: Children, Young People, and New Media*. Erlbaum, Mahwah, NJ, pp. 273–293.
- Burnett, C., Merchant, G., 2014. Points of view: reconceptualising literacies through an exploration of adult and child interactions in a virtual world. *J. Res. Read.* 37 (1), 36–50.
- Burnett, C., Merchant, G., 2018a. Affective encounters: enchantment and the possibility of reading for pleasure. *Literacy* 52 (2), 62–69.
- Burnett, C., Merchant, G., 2018b. Literacy-as-event: accounting for relationality in literacy research. *Discourse* 41 (1), 1–12.
- Burnett, R.E., Frazee, A., Hanggi, K., Madden, A., 2014. A programmatic ecology of assessment: using a common rubric to evaluate multimodal processes and artifacts. *Comput. Compos.* 31 (1), 53–66.
- Cartner, H., Hallas, J., 2020. Aligning assessment, technology, and multiliteracies. *E-Learn. Digital Media* 17 (2), 131–147. <https://doi.org/10.1177/2042753019899732>.
- Castek, J., Beach, R., Cotanch, H., Scott, J., 2014. Exploiting the affordances of multimodal tools for writing in the science classroom. In: Anderson, R.S., Mims, C. (Eds.), *Handbook of Research on Digital Tools for Writing Instruction in K-12 Settings*. IGI Global, Hershey, PA, pp. 80–101.
- Chandra, S., Chang, A., Day, L., Fazlullah, A., Liu, J., McBride, L., Mudalige, T., Weiss, D., 2020. Closing the K–12 Digital Divide in the Age of Distance Learning. *Common Sense Media*, San Francisco, CA. Boston, Massachusetts: Boston Consulting Group.
- Chisholm, J.S., Whitmore, K.F., 2018. *Reading Challenging Texts: Layering Literacies Through the Arts*. National Council of Teachers of English, Routledge/Urbana, IL, New York, NY.
- Claxton, G., 2015. *Intelligence in the Flesh: Why Your Mind Needs Your Body Much More than It Thinks*. Yale University Press, New Haven, CT.
- Collopy, T., 2020. Zines in the classroom: finding an audience of one—or 100. *Council Chronicle* 29 (3), 26–29.
- Common Sense Media, 2019. *The Common Sense Census: Media Use by Tweens and Teens, 2019*. Author. Retrieved from: <https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2019>.
- Curwood, J.S., Gibbons, D., 2009. "Just like I have felt": multimodal counternarratives in youth-produced digital media. *Int. J. Learn. Media* 1 (4), 59–77.
- Dalton, B., Smith, B.E., 2014. Teachers' lesson design as remix: composing with Internet resources and a smart authoring. In: Ferdig, R.E., Pytash, K.E. (Eds.), *Exploring Multimodal Composition and Digital Writing*. IGI Global Publishers, Hershey, PA, pp. 116–134.
- Davis, T., 2013. *The Big Dry*. Harper Collins, Sydney, NSW.
- DeJaynes, T., Curmi-Hall, C., 2019. Transforming school hallways through critical inquiry: multimodal literacies for civic engagement. *J. Adolesc. Adult Literacy* 63 (3), 699–309.
- DePalma, M.J., Alexander, K.P., 2015. A bag full of snakes: negotiating the challenges of multimodal composition. *Comput. Compos.* 37, 182–200.
- Domingo, M., 2014. Migrating literacies: multimodal texts and digitally enabled text making. *Text Talk* 34 (3), 261–282.
- Edmiston, B., 2014. *Transforming Teaching and Learning With Active and Dramatic Approaches: Engaging Students Across the Curriculum*. Routledge, New York, NY.
- Faulkner, J., Latham, G., 2010. A virtual school for rethinking learning. In: Cole, D.R., Pullen, D.L. (Eds.), *Multiliteracies in Motion: Current Theory and Practice*. Routledge, New York, NY, pp. 239–255.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Seabury Press, New York, NY.
- Gallagher, C.W., 2014. Staging encounters: assessing the performance of context in students' multimodal writing. *Comput. Compos.* 31, 1–12.
- Garcia, A., Minkler, M., Cardenas, Z., Grill, C., Porter, C., 2015. Engaging homeless youth in community-based participatory research: a case study from Skid Row, Los Angeles. *Health Promot. Pract.* 5, 18–27.
- Gee, J.P., Hayes, E.R., 2011. *Language and Learning in the Digital Age*. Routledge, New York, NY.
- Geers, S., Boukes, M., Moeller, J., 2020. Bridging the gap? The impact of a media literacy educational intervention on news media literacy, political knowledge, political efficacy among lower-educated youth. *J. Media Literacy Educ.* 12 (2), 41–53.
- Gershon, W., 2011. Embodied knowledge: sounds as educational systems. *J. Curriculum Theor.* 27 (2), 66–81.
- Gillenwater, C., 2014. Reading images: the phenomenon of intertextuality and how it may contribute to developing visual literacy with advanced placement English/language arts students. *J. Ethnogr. Qual. Res.* 8 (4), 251–263.
- Graham, L.F., Reyes, A.M., Lopez, W., Gracey, A., Snow, R.C., Padilla, M.B., 2013. Addressing economic devastation and built environment degradation to prevent violence: a photovoice project of Detroit Youth Passages. *Community Literacy J.* 8 (1), 41–52.
- Greene, S., Burke, K.J., McKenna, M.K., 2016. "When words fail, art speaks": learning to listen to youth stories in a community photovoice project. In: Greene, S., Burke, K.J., McKenna, M.K. (Eds.), *Youth Voices, Public Spaces, and Civic Engagement*. Routledge, New York, NY, pp. 235–258.
- Greene, S., Burke, K.J., McKenna, M., 2018. A review of research connecting digital storytelling, photovoice, and civic engagement. *Rev. Educ. Res.* 88 (6), 844–878.
- Griffith, P.E., 2010. Graphic novels in the secondary classroom and school libraries. *J. Adolesc. Adult Literacy* 54 (3), 181–189.
- Guajardo, F., Guajardo, M., Cantu, M., 2016. Where are they now? An intergenerational conversation on the work of the Llano Grande Center for Research and Development. In: Greene, S., Burke, K., McKenna, M. (Eds.), *Youth Voices, Public Spaces, and Civic Engagement*. Routledge, New York, NY, pp. 169–185.
- Haigney, S., 2020. Do viral videos really tell you anything about today's teens? *N. Y. Times Mag.* 7–8. Retrieved from: <https://tinyurl.com/y2f9urky>.
- Hall, L.A., 2019. "I didn't enjoy reading until now": how youth and adults engage with interactive digital texts. *Res. Teach. Engl.* 54 (2), 109–130.
- Harste, J.C., Kress, G., 2012. Image, identity, and insights into language. *Lang. Arts* 89 (3), 205.
- Hartman, P.M., Berg, J., Schuler, B., Knauer, E., 2019. Using artistic response strategies meaningfully in the English language arts classroom. In: Macro, K.J., Zoss, M. (Eds.), *A Symphony of Possibilities: A Handbook for Arts Integration in Secondary English Language Arts*. National Council of Teachers of English, Urbana, IL, pp. 121–138.
- Hepp, A., Krotz, F. (Eds.), 2014. *Mediatized Worlds: Culture and Society in a Media Age*. Palgrave Macmillan, New York, NY.
- Hewitt, C., 2008. Multimodality and literacy in school classrooms. *Rev. Res. Educ.* 32 (1), 241–267. <https://doi.org/10.3102/0091732x07310586>.

- Hewitt, C., 2014. An introduction to multimodality. In: Hewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*, second ed. Routledge, London, pp. 15–30.
- Hinton, S.E., 1967. *The Outsiders*. Viking, New York, NY.
- Hollett, T., Phillips, N.C., Leander, K.M., 2018. Digital geographies. In: Mills, K.A., Stornaiuolo, A., Smith, A., Pandya, J.Z. (Eds.), *Handbook of Writing, Literacies, and Education in Digital Cultures*. Routledge, New York, NY, pp. 148–172.
- Howell, E., Butler, T., Reinking, D., 2017. Integrating multimodal arguments into high school writing instruction. *J. Literacy Res.* 49 (2), 181–209.
- Hsu, H., Wang, S., 2010. Using gaming literacies to cultivate new literacies. *Simulat. Gaming* 41 (3), 400–417.
- Hull, G.A., Nelson, M.E., 2005. Locating the semiotic power of multimodality. *Writ. Commun.* 22 (2), 224–261.
- Hull, G.A., 2003. At last: youth culture and digital media: new literacies for new times. *Res. Teach. Engl.* 38 (2), 229–233.
- Hutchison, A., Beschoner, B., 2014. Using the iPad as a tool to support literacy instruction. *Technol. Pedagog. Educ.* 24 (4), 407–422. <https://doi.org/10.1080/1475939X.2014.918561>.
- Jenkins, H., Ford, S., Green, J., 2013. *Spreadable Media: Creating Value and Meaning in a Networked Culture*. New York University Press, New York.
- Jenkins, H., 2006. Confronting the Challenges of Participatory Culture: Media Education for the 21st Century. The MacArthur Foundation, Chicago. Retrieved from. <http://tinyw.in/xB75>.
- Jewitt, C. (Ed.), 2014. *The Routledge Handbook of Multimodal Analysis*. Routledge, New York, NY.
- Jiang, S., Smith, B.E., Shen, J., 2019. Examining how different modes mediate adolescents' interactions during their collaborative multimodal composing processes. *Interact. Learn. Environ.* 29 (5), 1–14. <https://doi.org/10.1080/10494820.2019.1612450>.
- Jocius, R., 2013. Exploring adolescents' multimodal response to the *Kite Runner*: understanding how students use digital media for academic purposes. *J. Media Literacy Educ.* 5 (1), 310–325.
- Jocius, R., 2017. Good student/bad student: situated identities in the figured worlds of school and creative multimodal production. *Literacy Res.* 66, 198–214. <https://doi.org/10.1177/23813369177>.
- Kalantzis, M., Cope, B., 2012. *New Learning: Elements of a Science of Education*, second ed. Cambridge University Press, New York, NY.
- Kamarainen, A.M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M.S., et al., 2013. EcoMOBILE: integrating augmented reality and probeware with environmental education field trips. *Comput. Educ.* 68, 545–556.
- Karchmer-Klein, R., Shinas, V.H., 2019. Adolescents' navigation of linguistic and nonlinguistic modes when reading a digital narrative. *J. Res. Read.* 42 (3–4), 469–484.
- Kesler, T., Gibson, L., Turansky, C., 2016. Bringing the book to life: responding to historical fiction using digital storytelling. *J. Literacy Res.* 48 (1), 39–79. <https://doi.org/10.1177/1086296X16654649>.
- Koehler, S., 2013. The Problematics of the Fake Harlem Shake. Retrieved from. <http://tinyurl.com/c7r7paf>.
- Korobkova, K.A., Collins, P., 2018. The variety of user experiences: literacy roles and stances on story-sharing platforms. *J. Adolesc. Adult Literacy* 62 (4), 387–399.
- Krakauer, J., 1996. *Into the Wild*. Random House, New York, NY.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, New York, NY.
- Landry, D., 2016. Th3 Anomaly: Crossing the Rubicon. Author. <https://th3anomaly.weebly.com/explor32.html>.
- Lankshear, C., Knobel, M., 2011. *New Literacies: Everyday Practices and Social Learning*, third ed. McGraw-Hill, New York, NY.
- Leander, K., Boldt, G., 2013. Rereading "A pedagogy of multiliteracies": bodies, texts, and emergence. *J. Literacy Res.* 45 (1), 22–46. <https://doi.org/10.1177/1086296X12468587>.
- Leander, K., Hollett, T., 2013. Designing new spaces for literacy learning. In: Dunston, P., Fullerton, S.K., Bates, C.C., Stecker, P.M., Cole, M.W., Hall, A.H., et al. (Eds.), 62nd Yearbook of the Literacy Research Association. Literacy Research Association, Altamonte Springs, FL, pp. 29–42.
- Leander, K.M., Aziz, S., Botzakis, S., Ehret, C., Landry, D., Roswell, J., 2017. Readings and experiences of multimodality. *Literacy Res.* 66, 95–116.
- LeBlanc, A., Barroux, 2009. *The Red Piano*. Wilkins Farago, Albert Park, VIC.
- Lemke, J.L., 2012a. Thinking about feeling: affect across literacies and lives. In: Erstad, O., Sefton-Green, J. (Eds.), *Identity, Community, and Learning Lives in the Digital Age*. Cambridge University Press, New York, NY, pp. 57–69.
- Lemke, J.L., 2012b. Multimedia and discourse analysis. In: Gee, J.P., Handford, M. (Eds.), *The Routledge Handbook of Discourse Analysis*. Routledge, New York: NY, pp. 79–89.
- Lewis, C., Crampton, A., 2016. Emotion as mediated action in doing research on learning. In: *Methodological Advances in Research on Emotion and Education*. Springer, New York, NY, pp. 137–150.
- Loh, C.E., Sun, B., 2019. "I'd still prefer to read the hard copy": adolescents' print and digital reading habits. *J. Adolesc. Adult Literacy* 62 (6), 663–672.
- Low, D.E., 2012. "Spaces invested with content": Crossing the "gaps" in comics with readers in schools. *Child Lit. Educ.* 43 (4), 368–484.
- Luhrman, B., 1996. *Romeo & Juliet*. United States. 20th Century Fox.
- Lysaker, J.T., Wessel-Powell, C., 2019. Comprehending as a manner of living. In: Beach, R., Bloome, D. (Eds.), *Language Relations for Transforming the Literacy and Language Arts Classroom*. Routledge, New York, NY, pp. 172–192.
- Lysaker, J.T., 2018. *Before Words: Wordless Picture Books and the Development of Reading in Young Children*. Teachers College Press, New York.
- Marlatt, R., 2019a. "I didn't say, 'Macbeth,' it was my Google Doc!": a secondary English case study of redefining learning in the 21st Century. *E-Learn. Digital Media* 16 (1), 46–62. <https://doi.org/10.1177/2042753018817544>.
- Marlatt, R., 2019b. "Ditch the study guide": creating short films to analyze literature circle texts. *J. Adolesc. Adult Literacy* 63 (3), 311–321.
- McClain, J.M., 2016. A framework for using popular music videos to teach media literacy dialogue. *Interdiscip. J. Pop. Cult. Pedagog.* 3 (1), 38–46.
- McCormick, J., 2011. Transmediation in the language arts classroom: creating contexts for analysis and ambiguity. *J. Adolesc. Adult Literacy* 54 (8), 579–587.
- McDermott, P., Falk-Ross, F., Medow, S., 2017. Using the visual and performing arts to complement young adolescents' "close reading" of texts. *Middle Sch. J.* 48 (1), 27–33.
- Meixner, E., Peel, A., Hendrickson, R., Szczeck, L., Bousum, K., 2018. Storied lives: teaching memoir writing through multimodal mentor texts. *J. Adolesc. Adult Literacy* 62 (5), 495–508.
- Miller, C., 2019. Using Graphic Novels and Video Essays to Critically Understand Socio-political Events. Retrieved from. <https://tinyurl.com/uc3jmv9>.
- Mills, K.A., Unsworth, L., 2018. iPad animations: powerful multimodal practices for adolescent literacy and emotional language. *J. Adolesc. Adult Literacy* 61 (6), 609–620.
- Mills, K.A., 2010. Shrek meets Vygotsky: rethinking adolescents' multimodal literacy practices in schools. *J. Adolesc. Adult Literacy* 54 (1), 35–45. <https://doi.org/10.1598/JAAL.54.1.4>.
- Mills, K.A., 2015. *Literacy Theories for the Digital Age: Social, Critical, Multimodal, Spatial, Material and Sensory Lenses*. Channel View Publications, Bristol, UK.
- Mirra, N., Garcia, A., 2017. Civic participation reimagined: youth interrogation and innovation in the multimodal public sphere. *Rev. Res. Educ.* 41 (1), 136–158. <https://doi.org/10.3102/0091732X17690121>.
- Moeller, R.A., 2016. A question of legitimacy: graphic novel reading as "real" reading. *J. Adolesc. Adult Literacy* 59 (6), 709–717.
- Napoli, M., Sychterz, T., 2015. Graphic novels come alive in a sixth grade classroom. In: Yenika-Agbaw, V., Sychterz, T. (Eds.), *Adolescents Rewrite Their Worlds: Using Literature to Illustrate Writing Forms*. Rowman & Littlefield, Lanham, MD, pp. 148–162.
- National Writing Project, 2011. Planning to Test the Framework in Practice. Retrieved from. <http://digitals.nwp.org/resource/2751>.
- Newell, R., Dale, A., Winters, C., 2016. A picture is worth a thousand data points: exploring visualizations as tools for connecting the public to climate change research. *Cog. Soc. Sci.* 2 (1), 1201885. <https://doi.org/10.1080/23311886.2016.1201885>.
- O'Neill, S.J., Smith, N., 2014. Climate change and visual imagery. *Wiley Interdiscip. Rev.* 5 (1), 73–87.
- Orwell, G., 1950. 1984. Signet, New York, NY.
- Pantaleo, S., 2013. Paneling "matters" in elementary students' graphic narratives. *Literacy Res. Instr.* 52 (2), 150–171.

- Pantaleo, S., 2014. Reading images in graphic novels: taking students to a "greater thinking level". *Engl. Aust.* 49 (1), 38–52.
- Partington, A., 2010. Game literacy, gaming cultures, and media education. *Engl. Teach. Pract. Critiq.* 9 (1), 73–86.
- Pellegrino, A., Zenkov, K., Gallagher, M., Long, L., 2016. Picturing new notions of civic engagement in the US: youth-facilitated, visually-based explorations of the perspectives of our least franchised and most diverse citizens. In: Greene, S., Burke, K., McKenna, M. (Eds.), *Youth Voices, Public Spaces, and Civic Engagement*. Routledge, New York, NY, pp. 25–49.
- Pet, S.R., McVerry, G., O'Byrne, W.I., 2014. Multimodal response and writing as poetry experience. In: Ferdig, R.E., Pytash, K.E. (Eds.), *Exploring Multimodal Composition and Digital Writing*. IGI Global, Hershey, PA, pp. 201–225.
- Phelan, M., 2014. *Around the World*. Penguin, New York, NY.
- Prior, P., 2009. From speech genres to mediated multimodal genre systems: Bakhtin, Voloshinov, and the question of writing. In: Bazerman, C., Bonini, A., Figueiredo, D. (Eds.), *Genre in a Changing World*. WAC Clearinghouse & Parlor Press, Anderson, SC, pp. 17–34. <https://wac.colostate.edu/docs/books/genre/chapter2.pdf>.
- Pullinger, K., Joseph, C., 2012. *Inanimate Alice*. Bradfield Company, London.
- Pyles, D.F., Rish, R.M., Warner, J., 2019. *Negotiating Place and Space Through Digital Literacies Research and Practice*. Information Age Publishing, Charlotte, NC.
- Quinlan, A., Curtin, A., 2017. Contorting identities: figuring literacy and identity in adolescent worlds. *Ir. Educ. Stud.* 36 (4), 457–470. <https://doi.org/10.1080/03323315.2017.1362352>.
- Ranker, J., 2015. Redesigning the everyday: recognizing creativity in student writing and multimodal composing. *Lang. Arts* 92, 359–365.
- Rideout, V., Robb, M.B., 2019. *The Common Sense Census: Media Use by Tweens and Teens, 2019*. Common Sense Media, San Francisco, CA.
- Rogers, Y., Lindley, S., 2004. Collaborating around vertical and horizontal large interactive displays: which way is best? *Interact. Comput.* 16 (6), 1133–1152.
- Rogers, T., Winters, K., LaMonde, M., 2014. *Youth, Critical Literacies, and Civic Engagement: Arts, Media, and Literacy in the Lives of Adolescents*. Routledge, New York, NY.
- Roque, G., 2012. Visual argumentation: a further reappraisal. In: van Eemeren, F.H., Garssen, B. (Eds.), *Topical Themes in Argumentation Theory: Twenty Exploratory Studies*. Springer Publications, Amsterdam, pp. 273–288.
- Rosenblatt, L.M., 1993. The transactional theory: against dualisms. *Coll. Engl.* 55 (4), 377–386.
- Roswell, J., 2013. *Working With Multimodality: Rethinking Literacy in a Digital Age*. Routledge, New York, NY.
- Rosell, D., Cutter-Mackenzie, A., Foster, J., 2017. Children of an Earth to Come: speculative fiction, geophilosophy and climate change education research. *Educ. Stud.* 53 (6), 654–669.
- Ruhe, S., 2020. Fostering (digital) media literacy skills and global citizenship in the EFL classroom: digital stories of undocumented youth. In: Farrelly, R., Krulatz, A., Neokleous, G. (Eds.), *Handbook of Research on Cultivating Literacy in Diverse and Multilingual Classrooms*. IGI Global, Hershey, PA, pp. 125–145.
- Sakr, M., 2018. Multimodal participation frameworks during young children's collaborative drawing on paper and on the iPad. *Think. Skills Creativ.* 29, 1–11. <https://doi.org/10.1016/j.tsc.2018.05.004>.
- Salinger, J.D., 1951. *The Catcher in the Rye*. Little, Brown, and Company, Boston, MA.
- Santat, D., 2011. *Sidekicks*. Scholastic Inc, New York, NY.
- Scollon, R., Scollon, S.W., 2003. *Discourses in Place: Language in the Material World*. Routledge, New York, NY.
- Serafini, F., 2011. Expanding perspectives for comprehending visual images in multimodal texts. *J. Adolesc. Adult Literacy* 54 (5), 342–350.
- Serafini, F., 2012. Expanding the four resources model: reading visual and multimodal texts. *Pedagogies. Int. J.* 7 (2), 150–164.
- Serafini, F., 2015. Multimodal literacy: from theories to practices. *Lang. Arts* 92 (6), 412–423.
- Shanahan, L.E., McVee, M.B., Bailey, N.M., 2014. Multimodality in action: new literacies as more than activity in middle and high school classrooms. In: Khosrow-Pour, M. (Ed.), *Computational Linguistics: Concepts, Methodologies, Tools, and Applications*, vols. I–III. IGI Global, Hershey, PA, pp. 1315–1333.
- Shipka, J., 2011. *Toward a Composition Made Whole*. University of Pittsburgh Press, Pittsburgh, PA.
- Shipka, J., 2012. Including, but not limited to, the digital: composing multimodal texts. In: Bowen, T., Whithaus, C. (Eds.), *Multimodal Literacies and Emerging Genres*. University of Pittsburgh Press, Pittsburgh, PA, pp. 73–89.
- Siegel, M., 2006. Rereading the signs: multimodal transformations in the field of literacy education. *Lang. Arts* 84 (1), 65–77.
- Silseth, K., Gilje, O., 2019. Multimodal composition and assessment: a sociocultural perspective. *Assess. Educ. Princ. Pol. Pract.* 26 (1), 26–42. <https://doi.org/10.1080/0969594X.2017.1297292>.
- Sinfield, D., Cochrane, T., 2018. Augmenting visual design: designing the changing classroom. *E-Learn. Digital Media* 17 (5), 1–23. <https://doi.org/10.1177/2042753018773769>.
- Smetana, L., Odelson, D., Burns, H., Grisham, D.L., 2009. Using graphic novels in the classroom: engaging deaf students with a new genre. *J. Adolesc. Adult Literacy* 53 (3), 228–240.
- Smith, B.E., Shen, J., 2017. Scaffolding digital literacies for disciplinary learning: adolescents collaboratively composing multimodal science fictions. *J. Adolesc. Adult Literacy* 61 (1), 85–90.
- Smith, A., West-Puckett, S., Cantrill, C., Zamora, M., 2016a. Remix as professional learning: educators' iterative literacy practice in CLMOOC. *Educ. Sci.* 6 (1), 1–19. <https://doi.org/10.3390/educsci6010012>.
- Smith, B.E., Kiili, C., Kauppinen, M., 2016b. Transmediating argumentation: students composing across written essays and digital videos in higher education. *Comput. Educ.* 102, 138–151.
- Smith, B.E., Shen, J., Jiang, S., 2019. The science of storytelling: middle schoolers engaging with socioscientific issues through multimodal science fictions. *Voices from the Middle* 26 (4), 50–55.
- Smith, B.E., Pacheco, M.B., Khorosheva, M., 2021. Emergent bilingual students and digital multimodal composition: a systematic review of research in secondary classrooms. *Read. Res. Q.* 56 (1), 33–52.
- Smith, B.E., 2014. Beyond words: a review of research on adolescents and multimodal composition. In: Ferdig, R.E., Pytash, K.E. (Eds.), *Exploring Multimodal Composition and Digital Writing*. IGI Global Publishers, Hershey, PA, pp. 1–19.
- Smith, B.E., 2017. Composing across modes: a comparative analysis of adolescents' multi-modal composing processes. *Learn. Media Technol.* 42 (3), 259–278.
- Smith, B.E., 2018. Composing for affect, audience, and identity: toward a multidimensional understanding of adolescents' multimodal composing goals and designs. *Writ. Commun.* 35 (2), 182–214.
- Smith, B.E., 2019. Mediatonal modalities: adolescents collaboratively interpreting literature through digital multimodal composing. *Res. Teach. Engl.* 53 (3), 197–222.
- Sobko, S., Unadkat, D., Adams, J., Hull, G., 2020. Learning through collaboration: a networked approach to online pedagogy. *E-Learn. Digital Media* 17 (1), 36–55. <https://doi.org/10.1177/2042753019882562>.
- Soep, E., Chávez, V., 2010. *Drop that Knowledge: Youth Radio Stories*. University of California Press, Berkeley, CA.
- Starosielski, N., Walker, J. (Eds.), 2016. *Sustainable Media: Critical Approaches to Media and Environment*. Routledge, New York.
- Steinbeck, J., 1947. *The Pearl*. Viking Press, New York, NY.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Literacy Res.* 49 (1), 68–91.
- Taylor, M.D., 1987. *The Gold Cadillac*. Penguin, New York, NY.
- Taylor, M.D., 1997. *Roll of Thunder, Hear My Cry*, vol. 2. Penguin, New York, NY.
- Teixeira, S., 2015. "It seems like no one cares": participatory photo mapping to understand youth perspectives on property vacancy. *J. Adolesc. Res.* 30 (3), 390–414.
- Unsworth, L., 2008. Multiliteracies and metalanguage: describing image/text relations as a resource for negotiating multimodal texts. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Erlbaum, New York, NY, pp. 377–405.
- Unsworth, L., 2011. *New Literacies and the English Curriculum*. Continuum, New York, NY.

- Unsworth, L., Macken-Horarik, M., 2015. Interpretive responses to images in picture books by primary and secondary school students: exploring curriculum expectations of a "visual grammatics". *English in Education* 49 (1), 56–79.
- Unsworth, L., Cope, J., Nicholls, L., 2019. Multimodal literacy and large-scale literacy tests: curriculum relevance and responsibility. *Aust. J. Lang. Literacy* 42 (2), 128–139.
- VanKooten, C., 2016. "The video was what did it for me": developing meta-awareness about composition across media. *Coll. Engl.* 79 (1), 57–80.
- Vonnegut, K., 1961. Harrison Bergeron. *Magazine of Fantasy and Science Fiction* 21 (4), 5–10.
- Vu, V., Warschauer, M., Yim, S., 2019. Digital storytelling: a district initiative for academic literacy improvement. *J. Adolesc. Adult Literacy* 63 (3), 257–267.
- Walker, G., 2017. Insider windows in Nepal: a critical pedagogy for environmental communication. In: Milstein, T., Pileggi, M., Morgan, E. (Eds.), *Environmental Communication Pedagogy and Practice*. Routledge, pp. 167–182.
- Walsh, M., 2010. Multimodal literacy: what does it mean for classroom practice? *Aust. J. Lang. Literacy* 33 (3), 211–239.
- Wargo, J.M., 2015. Spatial stories with nomadic narrators: affect, Snapchat, and feeling embodiment in youth mobile composing. *J. Lang. Literacy Educ.* 11 (1). Retrieved from <http://tinyw.in/Sz03>.
- Weissman, K., Costello, S., 2014. Creating digital comics in response to literature: aesthetic, aesthetic transactions, and meaning making. *Lang. Arts* 92 (2), 103–117.
- West, J., 2019. Using new literacies theory as a lens for analyzing technology-mediated literacy classrooms. *E-Learn. Digital Media* 16 (2), 151–173.
- Wexler, L., Gubrium, A., Griffin, M., DiFulvio, G., 2013. Promoting positive youth development and highlighting reasons for living in Northwest Alaska through digital storytelling. *Health Promot. Pract.* 14 (4), 617–623.
- White, A.M., 2019. Exploring the possibilities and tensions of visual responses to literature. In: Macro, K.J., Zoss, M. (Eds.), *A Symphony of Possibilities: A Handbook for Arts Integration in Secondary English Language Arts*. National Council of Teachers of English, Urbana, IL, pp. 98–120.
- Wohlwend, K., 2017. Chasing literacies across action texts and augmented realities: E-books, animated apps, and Pokémon go. In: Burnett, C., Merchant, G., Simpson, A., Walsh, M. (Eds.), *The Case of the iPad*. Springer, New York, NY, pp. 49–66.
- Wolfe, P., Kleijwegt, D., 2012. Interpreting graphic versions of Shakespearean plays. *Engl. J.* 101 (5), 30–36.
- Woodson, J., Lewis, E.B., 2001. *The Other Side*. Equinox, Sheffield, UK.
- Wright, R., 2005. *Native Son*. Harper, New York, NY.
- Wyatt-Smith, C., Elkins, J., 2008. Assessing new literacies in science and mathematics. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Routledge, New York, NY, pp. 899–940.
- Zammit, K., 2016. Responding to literature: iPads, apps and multimodal text creation. *Literacy Learn. Middle Years* 24 (2), 8–16.
- Zarate, O., 2011. *Othello*. Can of Worms Press, London, UK.
- Zoss, M., 2019. Teaching high school English with drawings and large-scale visual products. In: Macro, K.J., Zoss, M. (Eds.), *A Symphony of Possibilities: A Handbook for Arts Integration in Secondary English Language Arts*. National Council of Teachers of English, Urbana, IL, pp. 139–156.
- Zusak, M., 2000. *Fighting Ruben Wolfe*. Scholastic, New York, NY.

Relevant websites

Biosphere 2 (B2), University of Arizona Lab, biosphere2.org.

Center for Digital Storytelling, www.storycenter.org.

Digicom Learning, digicomlearning.com.

Digital Story for Roll of Thunder, Hear My Cry, tinyurl.com/koo9hlp.

Episode 2: Italy Inanimate Alice, inanimatealice.com/adventures/episode-2-italy.

Google Earth Timelapse Tool Earthengine, <http://google.com/timelapse>.

Media Representations of Climate Change, goo.gl/4dgv23.

New Learning Site on Multimodal-Based Instruction, newlearningonline.com/literacies/chapter-8.

Poetry 4 U Project, tinyurl.com/y7hvx8m8.

Project Imagine the Future, imaginefuture.org.

StoryCenter, www.storycenter.org.

Teaching With Zines, bit.ly/2H4W25j.

TechScaffold for Searching for Apps, techscaffold.com.

University of Houston Center for Digital Storytelling, digitalstorytelling.coe.uh.edu.

Website for Using Apps for Learning Across the Curriculum Book Usingipads, pbworks.com.

Electracy

Gregory L. Ulmer, University of Florida, Gainesville, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

The designing of the disaster	518
Heuretics	520
Theopraxis	521
The visceral turn	522
Cabaret academy	524
Theming mystory	526
References	527

A man's work is nothing but this slow trek to rediscover, through the detours of art, those two or three great and simple images in whose presence his heart first opened.

Albert Camus.

The designing of the disaster

The Institution of education will be reimagined for a post-COVID-19 world. The cascading disasters of 2020 created a classroom “catastrophe” jump from one paradigm to another: from print to digital media. After decades of gradual assimilation of computing within schooling, suddenly classroom instruction worldwide went online, exposing the incommensurability of literate pedagogy and digital technology. Electracy (the digital apparatus) dates from the beginning of the Industrial Revolution, in the later eighteenth century, which is to say that one of its identifying features is the Anthropocene (the epoch in which human activity began to modify the planetary environment). Here is the responsibility to be confronted in an electracy education: nature and culture have merged; *the cause of world catastrophe is us*, humanity itself. There have been four apparatus (the term is both singular and plural) in the history of civilization, beginning in the paleolithic era some 80,000 years ago: orality 10,000 years, (Western) literacy 3000 years, and electracy 300 years ago. The insight of grammatology (the history and theory of writing), is that technologies of language and communication are one part of a tripartite apparatus (desiring machine): part ideological and part technological. The operation of apparatus in history was not noticed until the twentieth century, when the cumulative effects of electronic civilization raised awareness of the limits of (Western) literacy, the nature of its historical shape, emerging from orality in Classical Greece, and giving rise to “secondary orality” (as Walter Ong called it) in modern times.

“Apparatus” (*dispositif*) is a term favored by the *Tel Quel* school in France (1960–1982), the journal that published works by many French poststructuralists. The French provided most of apparatus philosophy, while the Toronto school (Marshall McLuhan, Eric Havelock, Walter Ong among others) provided the scholarship. The prototype apparatus of literacy includes three interrelated but independent inventions, one of which is technological (alphabetic writing). After the Greeks acquired alphabetic writing (about 1000 BCE), Plato created the first school, the Academy in Athens (387 BCE), opening a “chora,” a region of new reality within the dominant orality of his era, with an original metaphysics of science (mathematics). Within the Academy the students began inventing the operating practices of literacy, codified by Plato’s student, Aristotle, who founded his own academy, the Lyceum. The third dimension of invention involves identity experience, individual and collective. Havelock (1963) argued that individual selfhood and the related political order of democracy are as much inventions of literacy as alphabetic writing and school. Socrates is a syncretic figure, partly oral, partly literate, the first person to experience Self. This history is familiar, noted to make the point that each apparatus produces an original metaphysics (world view account of reality) including *technology*, *institution formation*, and *identity behavior*. Metaphysics accounts for space, time, and cause, and directs all resources of a civilization to manage ultimate cause in the interests of human thriving. Paleolithic was animistic, regarding everything as alive, using magic to promote survival (life/death). Orality identified divine cause, creating religion and the institution of the church to address right/wrong behavior mandated by God. Literacy identified nature as cause (*Physis*), creating true/false logics of science to manage material laws. Electracy opens a new chora, with humanity as cause, including a metaphysics of attraction/repulsion. This history should be taken into account when organizing education native to the digital apparatus.

The optimistic side of this history is that the digital apparatus is invented, as an actualization of human desire, and not merely an affordance determined by technology: humanity and technology evolve together in a dialectical dynamic. The caveat is that whatever the digital apparatus is becoming, it is not more literacy. Is it any easier for a literate scholar to accept electracy than it was for a church theologian to accommodate science? The trial of Galileo is a cautionary tale. Walter Ong warned against such terms as “oral literature,” a phrase he called “preposterous,” because it masked the incommensurability of oral and literate productions (Ong, 1982). The same must be said about “digital literacy,” not in its narrow usage for learning “best practices” of authoring online, but as a frame for understanding education native to an apparatus whose causality is not accounted for by magic, religion, or science. The complementary relationship of digital literacy and digital humanities with electracy is evident in this context, in that the functions of the former up until now have been to adapt computing technologies to the metaphysics of school, similar

to religions of the book adopting alphabetic writing but not scientific metaphysics. Reviews of digital humanities discuss how information sciences augment critical thinking native to literacy. There is an analogy between concept formation in literacy and data mining. The concept as a category of reason was discovered in patterns made accessible through scanning the oral record visualized in writing. Justice is the first concept, created in the *Republic* when Plato went beyond Hesiod's topical collections of dramatic situations of (in)justice, such as the anger of Achilles in the *Iliad*, to ask, "what is justice itself?" Distance reading (text data mining), introduced by Franco Moretti, similarly is revealing new patterns in the digitized archives of print not discernible to scholarship. Apparatus history suggests that these patterns hold major unforeseeable consequences for knowledge creation.

James Watt's steam engine (1775), The French Revolution (1789), and Immanuel Kant's *Critique of Judgment* (1790), are three threshold events marking the inception of electracy. *Climate Change* stands in for all *hyperobjects* (Morton, 2013) created by and exceeding control of scientific (literate) civilization, with posthuman significance (in which human survival and thriving are not prioritized). The apparatus of science (including literate schooling, individual selfhood and the democratic nation state) has become a victim of its own success. Scientific civilization produced a generalized sprawl, overrunning every physical and informational category and limit, resulting in an historical crisis of the formless (addressed in the recent field of border theory). Virilio (2003) testified through his Museum of Accidents to the impasse created by digital speed, light-speed technologies creating real-time feed-back loops collapsing all dimensions into Now, overwhelming the time and space essential to critical thinking. Virilio's description of the claustrophobic closure of the world upon itself due to the dimension collapse he called *dromosphere* (dromos = -speed) has become accessible to individual intuition in conditions of home quarantine required by the pandemic. The quarantine serves as collective existential epiphany.

Hopes for engineering solutions through technological progress are delusions, Virilio observed, since technology itself has become a primary driver of the Anthropocene. Every invention includes its own accident: the invention of the car includes the car wreck. Virilio warned of the *General Accident*, that happens everywhere simultaneously, with disasters such as Chernobyl representing harbingers. The Internet Accident has happened already, with the pollution of the public sphere by cyberwarfare, contributing significantly to the erosion of democracy. The COVID-19 pandemic is the most complete instance of a General Accident to date, exposing the limitations of the existing world order.

Virilio lamented that dependence on computational algorithms of calculation in all areas of productivity eliminated *analogy* as a resource of thought and imagination.

Digital technology is a filter that is going to modify perception by means of a generalized morphing, and this in real time. We are faced with something which is more than the failure of the traditional static arts, both visual and plastic: we are faced with the failure of the analogical in favor of calculation and the numerology of the image. Every sensation is going to be digitized or digitalized. We are faced with the reconstruction of the phenomenology of perception according to the machine.

Lotringer and Paul (2005, p. 66).

Virilio's concern recalls Plato's warning in *Phaedrus*, that writing would destroy living memory. Virilio suggests *digital imaging will destroy perception itself*. Plato's Academy began inventing a new metaphysics adequate to the artificial memory of writing. Electracy must do the same for the constructed (fictional, deep fake) phenomenology of the digital apparatus.

The more familiar usage of "apparatus" is in the phrase "Ideological State Apparatus," (ISA) used by Louis Althusser to describe the institutions responsible for educating citizens to become productive members of society (Althusser, 2001). Ideology represents the imaginary relationship of individuals to their real conditions of existence (which are not knowable directly). The alternative to ISA is RSA, the Repressive State Apparatus (the police, for example). The difference between managing social cooperation by ideas versus force was manifested in the protests following the murder of George Floyd. Citizens are *interpellated* (invited, solicited, called) to their preestablished position in the social order, according to the culture, customs and traditions of identity categories (the WASP order in the tradition of patriarchy). Althusser's analogy for the iconic *gest* of this call, especially telling in our context, was that of a police officer approaching a group, commanding "hey you there!" and the one who turned around to answer is interpellated. We recognize our place in the system.

The authors of a recent special topic issue on Digital Humanities for *PMLA* declared ideological critique of identity as a priority:

We reassert a fact that we believe to be self-evident: that DH must concern itself deeply with race, gender, disability, economic and linguistic access, and other intersecting axes of power embedded in our materials and methods as demanded by this troubled moment in the world.

Booth and Posner (2020, p. 10).

Critique of ideology, and critical analysis in general, are inherently literate, applying analytical reason to the Western tradition itself and against humanism as a compromised metaphysics. There are good reasons for this critique within literacy. Electracy calls for an *aspect shift*, however, to regard the "humanities" in DH from the point of view of apparatus, as resources for creativity across all disciplines, essential to education's response to the Anthropocene. Hans Blumenberg's insight into the relation of secularization to Christianity applies to our project also, which *reoccupies* the questions of the traditional humanities while replacing the conventional answers (Blumenberg, 1983). The insight of apparatus is that the four primary institutions of ISA—Family, Church, School, Media—are precisely the primary institutionalizations of apparatus (Paleo, Oral, Literate, Electrate), each originating in its own epoch,

with its own metaphysics, coexisting historically now, and in need of attunement into a tempered “stack.” The challenge of electracy pedagogy is to discover how we are positioned within these institutions and how to take control of this positioning on behalf of individual and collective well-being.

Heuretics

Students learn electracy by helping to invent it. Having the same root as “eureka,” heuretics is classified as “obsolete” in the O.E.D., defined as the use of theory to invent, as distinct from hermeneutics, the use of theory for interpretation. The difference between hermeneutic humanities and heuretics is similar to Thomas Kuhn’s distinction between normal and revolutionary science (Kuhn, 2012). The logic of normal science has been simulated in AI expert systems, for addressing routine problems whose solutions are codified within the paradigm of the discipline. When anomalies accumulate to the point of putting the paradigm in question, revolutionary science emerges in original hypotheses formulated by certain creative individuals. Holton (1973), historian of science, studies such creative figures to understand the ability to think beyond a paradigm. Holton noted that science includes three primary dimensions. Two of the dimensions are taught in the disciplines—those concerning how knowledge is verified (normal science): empirical facts (the archive of theory-dependent materials); analytical operations (mathematics, logics for manipulating the materials)—object of study, and method of study.

The third dimension (not taught in most schools) concerns how knowledge is discovered, refocusing on the subject who invents, dependent on analogies outside the range of algorithms (hence Virilio’s concern about the elimination of analogy in favor of algorithms). This difference has received attention, especially in periods following great calamities, in the hopes of finding a logic of creativity, so far with little success. Holton named this third discovery dimension “thematic,” to indicate its origins in the presuppositions, the dispositions specific to individuals, constituted in part by received cultural archetypes, and in part by idiosyncratic experience acquired in childhood, which is to say that creativity is grounded in that order of common sense resistant (so far) to AI simulation. The thematic order manifests itself in an intuition of reality, a dispositional feeling about how the world works or how one wants it to be. Original hypotheses, it seems, are part of a larger order of positioning, “position” being “thesis,” one’s standpoint or stand: diathesis (state of mind), aisthesis (sensory perception). Northrop Frye referred to literature not as fictional but “hypothetical,” an observation that calls attention to literature as an archive of themata, expressing underlying intuitions about reality. The core of most of Anton Chekhov’s stories, for example, is said to convey his life feeling of the mutual isolation of human beings and the impossibility of understanding each other. A rich vocabulary surrounds “thesis,” including Heidegger’s *Gestell* (enframing), a synonym for *dispositif* (apparatus). Heidegger (1977) remarked the ontological importance of disposition, which is essentially what he meant by *Dasein* (being-in-the-world). If digital technologies give us GPS (physical location), electracy takes up the challenge of metaphysical orientation, EPS (Existential Positioning System). This context of thesis becoming thematic reveals what is at stake in our aspect shift regarding ISA, in that the recognition experience of ideological interpellation is a clue also to one’s unique disposition (*Dasein*). One imagines within a specific hegemonic cultural order in the same way that one speaks a native language. Disposition is the site of human cause, chora of electracy metaphysics.

A figure guiding reoccupation or retrieval (cf. McLuhan and McLuhan, 1988) of the humanities is Amartya Sen (Nobel Prize laureate in Economics, 1998), who argues for an economics of well-being (Sen, 2009). The pragmatic takeaway is that happiness is literally profitable, although the main point is to expand the measures of productivity. Sen defines justice as an inalienable human right to develop one’s capabilities to their fullest potential. A society that supports individuals’ capacity to choose their own life possibility will prosper economically (Sen, 2009). The capability movement, taken up also by the philosopher Nussbaum (2011), retrieves virtue ethics, as Sen acknowledges, and even the tradition of human “faculties.” The understanding of being in terms of human virtues, powers, abilities, extends from Plato to the present, in a heritage including Aristotle’s *Entelechy*, Leibniz’s *Monad*, Spinoza’s *Conatus*, to Heidegger’s *Dasein* and *Ereignis* (event). *Conatus* (*striving to persevere in one’s own being*) perhaps best defines what is at stake in electracy pedagogy, taking ownership (enownment) of our responsibility for the Anthropocene as *the creativity accident*. This pedagogy is personal, with students receiving the tradition as guide to their own capabilities. In our aspect shift we focus on the core lesson of how the faculties (virtues) operate within culture, repeated in many variations, with the same structure throughout history.

Paul Gauguin’s experience in Tahiti introduces the traditional questions to be reoccupied in electracy pedagogy. The moment is 1897, late in his career, in a state of mind (diathesis) informed by despair and disillusionment due to debt and illness. Gauguin decided to paint his final testament and then commit suicide. He painted a masterpiece, but the suicide attempt failed. The title of the fresco-size painting is, “Where Do We Come From? What Are We? Where Are We Going?” The art critic Hal Foster declared this work to be the catechism of modernism itself, the title drawn from Gauguin’s upbringing in the Catholic faith. Reading from right to left, the scene depicts three groups of figures in Gauguin’s Fauvist style of portraiture, expressing the passage from birth to death, with each grouping symbolically associated with one of the three questions of the catechism. It is a visualization of the force of being, of *Conatus*, that Derrida calls “trace.”

The assignment structuring the curriculum of disposition is for students to discover variations on these questions and ultimately to answer them relative to their own situations (EPS). An unpacking of the catechism opens a vector through the entire Western tradition (expandable to other traditions as well). Plato established a comprehensive version of the capabilities, including individual and collective dimensions (microcosm and macrocosm). The *Republic*, Plato’s dialogue outlining the ideal city of literacy, correlates the three faculties or virtues—Reason, Will (Spirit), Appetite—with Head (rulers), Heart (guardians), Viscera (workers). Aristotle updated the virtues into their definitive form: *Theoria* (knowledge of necessity, nature’s laws known through science); *Praxis*

(ethics and politics, not amenable to science because involving human choice); *Poiesis* (aisthesis: craft, making, *techne*). In Aristotle's *Entelechy* entities become what they (already) are through time, actualizing their potential, realizing their form in matter. Acorns become oaks in time. The art of education is responsible for the actualization of human potential (Sen's economics doubles as a theory of education). Human essence is happiness (Sen agrees), but it may be achieved only collectively, in a polis (Republic), through human choice and productivity. The portmanteau term *theopraxis* indicates the requirement of human intelligence to integrate the virtues into one performance. *We wanted Happiness but produced Anthropocene*. Such is the fundamental aporia (tragedy) of civilization that makes humanity uncanny, Heidegger said.

Immanuel Kant revised Aristotle in his three Critiques outlining the limits, functions, and *relationships* of capability. The Critiques treat Understanding (Pure Reason, *Verstand*) dealing with laws of nature; Reason (Practical Reason, *Vernunft*) treats moral freedom. Kant's innovation, the third Critique, the Judgment of Taste, promoted aesthetics, the capacity to enjoy beauty, the experience of pleasure and pain, the dimension of sensory appetite, to equal status as a faculty, not traditionally acknowledged. Kant's Copernican revolution is an aspect shift, regarding Aristotle's categories not as in nature but rather as human capabilities. The categories space-time-cause are in us, not the world, enabling all experience, thus positioning humanity as responsible for what it makes (the Anthropocene). Kant set the agenda for electracy pedagogy, assigning to aesthetic judgment the task of bridging the irreducible gap, the abyss, separating the actualizations of pure and practical reason, that is, science and religion, laws of nature and human freedom. We may recognize in this *aporia* the condition to which *fetish logic* responds, the human capacity to support contradiction, to oscillate between aspects: *I know what is the case, but still I believe otherwise*. As we learned from Wittgenstein's exercises with the duck-rabbit optical illusion (the Necker Cube Gestalt shift), this seeming glitch in the sensorium is also a power, precisely the power of toggling aspects or points of view, in conditions oriented not by identity (as in literacy) but difference.

Electracy can be understood in terms of Kant's hypothesis that aesthetic judgment is the key to attuning the faculties and related ISA institutions, an argument reiterated more recently by Hannah Arendt. Kant anticipated Gauguin's catechism, asking in the context of his three Critiques: *What can I know? What should I do? What may I hope?* (how does the world work; how should I behave in such a world; how do I feel about this behavior). The mythological version of the questions dramatizes the aporia. The theme is the Judgment of Paris, the contest of the Golden Apple, awarded to the most desirable goddess by Paris of Troy. The competition was among Athena (Wisdom), Hera (Political Power), Aphrodite (Fertility), whom we recognize as our three faculties personified. Each goddess bribes Paris, who decides in favor of Aphrodite in order to possess Helen, the most beautiful mortal woman, leading to the Trojan war disaster. Plato's solution was that of time and process, a passage of ascent beginning with sexual desire, which draws one into the customs of marriage and social relations, subsequently through love to recognize Beauty itself, the pure form (Idea). Botticelli, commissioned by the Medici in Renaissance Florence to visualize this Neoplatonic lesson, attempted to condense the passage, genesis, or trace of being into a scene of the birth of Venus out of sea foam. Michel Serres adopted Aphrodite rising from the sea as the emblem of *chaosmos* (order emerging from complexity)—the scene of *chora*.

Theopraxis

Theopraxis is a portmanteau term to signify the importance of syncretizing the abilities beyond thinking: *theoria*, *praxis*, *poiesis*. The practical application of theopraxis begins with testing the mode of *reflective judgment* that Kant proposed for integrating dissociated faculties, such as in the extreme breakdown of the sublime, when circumstances overwhelm both the archive of concepts and the productive imagination, as has become the norm in formless conditions of alienation. As a symptom of apparatus disorientation, the sublime sets the agenda of electracy to augment imagination (*poiesis*) to the same degree that literacy augmented theoretical reason. Reflective judgment is the ability to create an original category for a multiplicity of perceptions out of the pure relations immanent in the materials. The modality of reflective judgment (as opposed to determinate judgment operating through existing concepts) is pleasure-pain in the experience of beauty or ugliness, and attraction-repulsion in the sublime, the breakdown of perception and loss of intuition. The sublime is defined as the simultaneous experience of attraction and repulsion, the defining axis of electracy, the equivalent for the digital apparatus of true/false for literacy, right/wrong for orality, life/death for paleo. Categories formed without concepts by aesthetic design support sense-making native to electracy. Kant anticipated a new mode of autopoietic (self-organizing) causation (*chaosmos*).

The consequences of Kant's aspect shift are manifested in the revolution across the arts, their inward turn to perform their own formal "critiques" as operations in their own right, and as extensions of the human sensorium rather than as mimetic representations. Theopraxis may be observed in Alan Kay's creation of the Graphical User Interface (GUI). Participating in the development of the personal computer in Xerox Park in the 1970s, Kay recognized the need to harness all the capabilities orienting people in experience. His interface design engages the Symbolic dimension of language (*Theoria*) in the keyboard and code; the Enactive order (*Praxis*) of action, movement, the hand control of a mouse; the Imaginal (*Poiesis*) in the visual figures of the eye with windows and icons. Donald Norman, Director of the Design Lab at UC San Diego, developed a program of experience or emotional design, integrating the three faculties, articulated as: Reflective (intellect, useful); Behavioral (effective, usable); Visceral (appearance, desirable). Testing proved a correlation between usability and desirability. We are back at Sen's vision, that happiness has economic value.

A final point to make about capabilities concern the "popcycle," the stack or cognitive (conative) map, the EPS by which citizens orient themselves in reality. The Anthropocene is sublime, a dromosphere of collapsed dimensions, meaning that the challenge of electracy education is to restore orientation by overcoming alienation that broke the continuity between individual experience of agency and collective forces of energy in the world, everything from global capitalism to quantum physics that exceed individual intuition. The popcycle (transversal circulation of invention through the apparatus) proposes to restore the micro/macrocsm

continuity through insight into recapitulation, witnessing the relationship between phylogeny and ontogeny, the correspondence or homology between the four institutions of apparatus and the body with its three virtues. The old question regarding whether history has a direction (if not a goal), in this context, must be answered affirmatively in observing this correlation of institution and virtue, just as Plato proposed (although without his hierarchy), and as assumed within traditional cultures of world civilizations (the mandala principle). By the time children are five years old, they have been interpellated into and become native practitioners of four different metaphysics, four discourses (epistemes), collectively generated by human capability, but fatally unsynchronized.

We are born into a family, the institution created in the paleolithic (goddess culture), in which we gain control of our body. It is worth pausing to consider this primal scene of instruction of infants that Julia Kristeva called the semiotic chora, in which imagination is formatted: weaning from the breast, sensory-motor standing and walking, bowel control, learning to speak. Erikson (1982) described the developmental vectors relating this early apprenticeship in embodiment with the social orders of knowledge, law, play. These behaviors produce “part-objects” emerging in transitional space (mother-infant fusion, the first chora), relating oral-anal-genital-voice-gaze appetites, informing all subsequent learning and behavior. EPS orientation augments this embodied foundation of human intelligence, setting an (impossible?) agenda for AI, or even artificial theopraxis (AT). Apparatus education continues with children taken to church, mosque, synagogue, temple, and otherwise introduced to moral norms of right and wrong (orality, praxis). By age four children enter formal schooling, introduced to the metaphysics of reason, evidence-based thinking of true/false, the laws of nature and history (literacy, theoria). Meanwhile, at some point early on, children enter into media culture through the screens transmitting entertainment into the home, projecting the corporate commodity metaphysics native to electracy. We take this assemblage for granted, and yet the conflict and contradictions among the institutions are patent from the beginning, not only within individuals and among the members of a family, but historically in the hot and cold wars of McWorld and Jihad. The 9/11 attack summarizes the macrocosmic conflict, integrating as it does religious motivation (orality), capitalist target (electracy), and scientific technology (literacy), intended to destroy bodies (paleo). This event symptomizes the disaster, the chaos of capabilities, well documented in infinite variation in the cultural archive.

The visceral turn

What is the metaphysics native to electracy? This question brings us to the challenge of apparatus as a frame for education, concerning what makes it “preposterous” to characterize our epoch as “digital literacy.” The point of estrangement is that *electracy institutionalizes human visceral appetite*. It does for appetite what literacy did for reason, orality for will, paleo for the body itself, which is to say that the institution of electracy functions with a reality principle of attraction/repulsion (the experience French psychoanalysis calls *jouissance*), superseding right/wrong, true/false, or even life/death. This *differend* (constitutive incommensurability) is the aporia (impasse), the barrier inhibiting our ability to counter the Anthropocene, which is resistant to the concerns of families, churches, and schools. *The institution of appetite is the corporation*, especially the entertainment corporation, but in general the function of the corporation as prosthesis of collective sensory viscera, organized by attraction/repulsion. Given the differend separating school from corporation, the immediate question concerns the terms of a pedagogy capable of attuning or tempering the full apparatus cycle (theopraxis).

Kant’s Judgment of Taste is well-named to introduce the heuristics of electrate pedagogy. The “discovery” of America by Columbus (whose statues are now coming down around the world), was an accident of the sense of taste, emblem for the Anthropocene. The story is familiar but unappreciated as symptom of apparatus shift. Columbus was seeking a new route to India to allow Spain access to trade in pepper, a trade that, expanded to include all spices, became the first hyperobject. Pepper was one of the most valuable commodities in the world over its six-thousand-year history. The movement to rename Columbus Day as Indigenous Peoples Day is well-meant, to acknowledge the genocide that accompanied the event, but it would be more apt to call it World Pepper Day, ritualized by abstaining from visits to the grocery store Spice Aisles for a week. It has to be recognized that people love pepper in electracy the way they love God in orality or truth in literacy. It is further proof of Virilio’s insight into the General Accident. Genocide of Indigenous People is the pepper accident.

The Spanish Conquista was carried out primarily by private enterprise licensed by the Crown. The first corporation was the British East India Company, created in 1600 (the Dutch counterpart just two years later), establishing a monopoly to engage in the spice trade. This institutional invention involved the legal innovation of limited liability, which solved the problem of risk management with respect to raising the enormous amount of capital necessary for the business, giving us the stock market. This invention has a long history, associated with inventions in mathematics, especially statistics and probability, motivated by attempts to hedge bets by gamblers. The history of capital tells of a continuing struggle to escape the mimesis of risk, with the latest installment being the “credit default swap” that produced the bubble and crash of the Great Recession (2007–2009). Finance still awaits its own Jackson Pollock, although cryptocurrency bears watching as a site of invention.

When Galileo faced the Inquisition over his claim that the earth moves around the sun, he assured the theologians that his interest was only with the laws of nature, material entailments leaving in place their divine cause. The theologians refuted Galileo’s toy telescope with their own instrument of proof, the rack. The choros of spirit and nature, created within orality and literacy respectively, actualize the potentiality of Praxis and Theoria. Gregory Bateson acknowledged this apparatus differend when he observed that the past two and a half millennia of civilization could be accounted for by two terms: *sacrament* (religion); *entropy* (science) (Bateson, 1979). A third term is needed to mark the emergence of electracy. That term is “commodity,” the nature of which is evoked by identifying El Chapo (Shorty) as the Galileo of electracy: drug lord, ruthless head of the Sinaloa Cartel that supplied eighty

percent of the US market in cocaine, heroin, meth, and marijuana—the contemporary spice trade—earning some one hundred billion dollars annually. People love drugs, and the mayhem associated with the trade is the drug accident.

Drugs are a commodity, commodity being an invention of corporate capitalism as prosthesis of human appetite. It is abject by standards of the previous apparatus. Such is the conventional critique of the corporation, as in the epic theater of Bertolt Brecht, that the new identity behavior of the digital apparatus is the mafia (collectively) and the prostitute (individually). Benjamin (1999) identified the poet Charles Baudelaire as the first modern person because he experienced himself as a commodity and the city as a market. El Chapo as the Galileo of electracy makes explicit the differend, the barrier that makes electracy as metaphysics unacceptable to school, church, family. And yet it is us, our accident, our Anthropocene. The presidency of Donald Trump holds some of the same lessons, manifesting paradigmatic abjection. His credentials qualifying him as the first electrate (in the worst sense) president is his status as a billionaire of shady real-estate deals and a reality TV star. Most telling is his trading in conspiracy theories and science denial, introducing a politics that is literally uncanny in its exposure of the unofficial dimension of social life kept out of view until now. His administration exposed the limits of literate collective identity and the need to reimagine the political and ethical order itself, beyond selfhood and democracy into brand franchising in a way that promotes well-being rather than disaster.

Our post-COVID reimagining requires an aspect shift, to accept the sublime experience of attraction/repulsion (*jouissance*) as a capability, a primary human faculty, to take up the project inventing synergy among the faculties and their institutions. Let us review the elements of electracy as metaphysics (technology, institutional form, and identity behavior). *Technology* is manufacturing, the machine age of the Industrial Revolution, and the associated inventions of recording—photography, telegraphy, phonography and the like, evolving into the Information Revolution. The corporation is the *institution*, realizing the capitalist mode of production. Our apparatus analogy helps notice the relevance of certain events, such as recent Supreme Court decisions (Citizens United, 2010) confirming the standing of corporations as persons, whose speech is money and as such protected by the Constitution. The analogy is with Thespis, the legendary inventor of tragedy (a transitional form moving from ritual to theater): Thespis was the first actor to step out from the ritual chorus and speak as an individual. This event in the invention of selfhood in literacy has its equivalent in the emergence of the corporation as a collective entity or macro-subject, a kind of Leviathan, whose appearance stimulated the formation of its counterpart, the Proletariat. Critics have noted in this context that corporation behaviors match FBI profiles of sociopaths (a diagnosis applied to Donald Trump himself). In apparatus terms we understand that what corporate Leviathans are doing is not reasoning or worshipping but *enjoying*. They are not so much speaking as twerking with money. This enjoyment is not an aberration but a human capacity suppressed in previous apparatus and becoming hegemonic in electracy.

What of the *identity* dimension of electracy? The new behavior emergent in electracy is consumerism: *brand* supplementing soul and self, modeled in celebrity stardom—the experience of oneself as an image. The commodity form separated exchange value from use value (a necessary step in the creation of any language mode). The first to exploit design thinking native to electracy were advertisers, similar to the sophists at the beginnings of literacy using dialectic for purposes of deception rather than wisdom. Given that the commodity has exchange value, advertisers realized they were selling not the steak (use value, the substance in literate categories) but the sizzle (the accidents of taste, the aesthetics of sensation). Such is the aspect shift in ontology, from a reality of essence and substance, to a reality of accidents and flesh. The reason advertising is so superior to critique, Walter Benjamin observed, is its recognition that *what matters is not what the moving red neon sign says but the fiery pool reflecting in the asphalt* (not the concept but the atmosphere). The culture wars in contemporary politics are about sizzle.

Advertising emblems support desiring in the same way that logical arguments support reasoning. The ad evokes a promise of happiness by triggering a fantasy, an event of electracy. What inferential entailments are to arguments fostering understanding, colors textures shapes are to designs triggering moods (feeling). The insight of electracy is that these moods are an independent capacity equal to and apart from logic or morality, with equal power to organize civilization. The greatest conversion of a civilization since Rome converted to Christianity is that of Christian and ultimately global society to consumerism. Philosophers warn of the dangers of this conversion, in that the commodity form disseminated in entertainment media, circulating part-objects, has tapped directly into human libidinal economy, industrializing and homogenizing the embodied energy source of imagination (Stiegler, 2013). The experience of electracy seen from the point of view of literacy is the series of *Matrix* films.

Electracy opens a new chora governed by a native ontology, a new understanding of *thing*, as theorized in psychoanalysis. Humanity itself is metaphysical cause in electracy, and psychoanalysis is the best account we have of visceral cause. Aristotle is credited with inventing the literate thing (substance) and codifying in his categories the ten statements possible to make about anything using declarative propositions, limited to the neutral modality of true/false. These categories remain in force today, condensed into the five *Ws* of journalism: *Who What When Where Why*, manipulated within alphabetic media, vehicle of reason, and dismissed in electracy as fake news. Meanwhile, Aristotle's truth tables, joined in a bachelor machine with Leibniz's binary numbers and Tesla's logic gate, meet in electronic computing to write true/false at light speed. Literacy put true/false into a machine and gifted it to electracy, to be applied in emergent ways, just as orality put speech into writing and gifted it to literacy. The object of electracy is *das Ding*, to use Freud's term, naming object relations (it is a relation, not a substance). Kristeva (1982) revised Freud and Lacan to situate the formation of this new kind of thing in the pre-Oedipal period from birth to eighteen months, the period of sensorial apprenticeship. In this period the child's first object forms, *the transitional object* as Winnicott called it, simultaneously within and outside, parent and child at once, a moebius topology for a homothetic imagination (extimacy). The energy of electrate chora comes from the initial event of satisfaction undergone in this potential space of infancy, giving rise to the fantasy of lack retrospectively constructed as defense against anxiety. The fantasy of the *object-cause of desire* (not what is desired, but desire's cause), Lacan's *objet petit a* or object @ (the little other, embodied in part-objects, produced by and separated from the fragmented

body), inspired by Winnicott's transitional object, circulates among bodies through media, structuring deep memory, experienced as *attraction/repulsion orienting the reality of electracy*. This cause is not logical but catalyzed by fiery pools of aesthetic design. The political order today must be reimagined in terms of this aesthetic articulation of need, demand, and desire (Three Graces of psychoanalysis). The political order in every apparatus emerges in relations formed at the intersection of the faculties (Fukuyama, 2002). This intersection is the fundamental interval requiring augmentation in AI, AT, even artificial unconscious. Again, it is worth repeating, the digital apparatus augments the commodified object @ just as the alphabetic apparatus augmented the propositional thing.

The insight of electrate ontology is that the universal exchange enabled by the general equivalent of commodity form supports this visceral dimension of human imagination (*poiesis*), potentially activated by any trivial item or detail invested with libidinal energy (its prototype is PTSD). It is the motivational "butterfly effect" of chaos theory. The change from one apparatus to another includes a shift in the functioning of the human sensorium—dominance of the hand in paleo, ear in orality, and eye in literacy. Kant's taste registered the new priority of the contact senses over the distance senses of the previous orders. The libidinal chora clarifies that the electrate sensorium is organized by what Lacan calls *lamella*, a fantasy organ unifying the erogenous zones of the body. Lacan explicitly compared his topology of the subject, the Borromean rings (knot) relating the Real, Symbolic, and Imaginary (RSI), held together by the drive of object @, with Aristotle's *theopraxis*, thus confirming the consistency of the capability structure through history, self-evident also in Freud's Ego, Superego, and Id (principles of reality, morality, and pleasure). The pivot of our theme is that, as Heidegger promised, *where danger is, grows the saving power also*. The faculty of imagination is a *pharmakon* (both gift and poison).

Cabaret academy

The pivot into the saving power of electrate education was the creation of cabarets in Montmartre on the outskirts of Paris in the 1880s. It is useful to place this innovation in the context of education, to appreciate its significance for the apparatus. Our framing question concerns how our civilization may reimagine itself against self-destruction. The historical innovation is the French response to losing the Franco-Prussian war of 1870–1872. The treaty ceded Alsace-Lorraine back to the Germans (it had traded hands numerous times). The French did not respond with the creation of a new kind of university, as the Germans had after their humiliating defeat by Napoleon, but instead with a cabaret scene, which hosted (accidentally) the invention of electrate metaphysics. Alsatian patriots began meeting in Montmartre, a site of bohemian counterculture opened in the industrializing city, populated with the abject class Marx dubbed the Lumpenproletariat (the dregs of society). It was a vice district, supplying cheap wine, women, and song, disreputable, as theater districts always have been. The American equivalent is Storyville in New Orleans, scene of the invention of jazz, the music of modernism. The patriots met in the backrooms of bistros, invoking the spirit of Rabelais, *Gargantua and Pantagruel*, especially the utopian Abbey of Thelema, with its motto, "Do what you want." They invoked the *Esprit Gaulois* of carnivalesque wit. The emblem of this inaugural movement was a painting by Manet, one of the founders of modernist art, called *Le Bon Bock* (A Good Glass of Beer), depicting a drinker in Alsatian attire. The cabaret scene of the 1880s included such venues as Le Chat Noir and Le Lapin Agile where artists gathered and formed the Hydropaths and the Incoherents, forerunners of the modernist avant-garde. Modernism actualizes the visceral turn in civilization.

The new state of mind (disposition) nurtured in this environment was *fumisme*, from *fumer*, to smoke (chimney-sweeping, figuratively lackadaisical, slacker, shirker). *Fumisterie* is humbug, farce, parody. It is worth noting that Anna O. (Bertha Pappenheim), one of the first patients to undergo psychoanalysis, referred to the "talking cure" as "chimney sweeping." The bohemian partiers mocked the entire Western tradition for amusement, of which Marcel Duchamp's modified readymade, *L.H.O.O.Q.* (a punning acronym saying "she has a hot ass"), is the prototype—a cheap reproduction of Leonardo's Mona Lisa defaced by a mustache. The poetics of Duchamp's oeuvre remains (for now) the one "style" beyond the powers of AI simulation (Manovich and Arielli, 2022), perhaps calling for revision of the Turing test: not whether a machine can think, but laugh. This new institution culminates in the Cabaret Voltaire, opened in Zurich during World War I, hosting the invention of Dadaism. *Cabaret is the academy of electracy, and Dadaism is its logic*—the Surrealist bachelor machine, exemplified by the meeting of an umbrella and sewing machine on an operating table (collage sampling of an archive). The analogy with the invention of literacy clarifies the micro-level of metaphysical invention. The basic device of literacy is dramatized in Plato's dialogues when Socrates asks an interlocutor to define a term (for Euthyphro to define "impiety" for example). Definition is a practice native to concept formation in literacy, unknown to illiterate interlocutors. Definition (outlining verbally the shape of idea) exposes contradictions in reasoning, with the principle of noncontradiction, identity, the excluded middle, as the primary operating rule of logic.

Our analogy prompts us to look for the equivalent micro-device articulating visceral mood functioning within digital media. This device was invented in the modernist arts, the exemplar being Paul Cézanne's *passage*. Cézanne's "faceting" of planes (point of departure for Cubism subsequently) introduced a device for working with sprawl conditions (formless loss of shape, disappearance of idea, blurring of figure-ground relations). The discovery included a new way to represent three dimensions on a two-dimensional surface, by exploiting *the little sensation* of aesthetic perception apart from concepts (enabling reflective judgment). His post-mimetic space was created by addressing the *involuntary* perceptual operation of the human sensorium—cool colors recede and warm colors advance. The importance of this optical effect in art history reminds us that cinema is based on the *phi* phenomenon, the optical illusion by which still images are made to move. Lacan adopted it as analogy of the Phallus illusion in the unconscious. Electracy emerges within the gathering of modernist arts, cinema, and psychoanalysis around the involuntary operations of perception.

The equivalent device in electracy of noncontradiction in literacy is that bringing this little sensation into design opens a practice for managing the autotropic experience of human visceral orientation in the world. Sensation metaphysics, augmented in digital technology, builds out the human demand for self-stimulation, the appetites of sensory organs (the lamella) engaged through consumption of every kind, from drugs, coffee, spice, to body-to-body replication of behavior accomplished by watching sports, movies, pornography (accounted for by the discovery of mirror neurons). The insight of libidinal economy is that at the heart of this drive for satisfaction is the lost impossible object @, fusion engine of the Anthropocene (human dis/satisfaction). The autotropic body must be educated to thrive in the light-speed feed-back loops of dromosphere. This program is part of the aspect shift away from critique to heuristics. The shift is from exclusive attention to communication within the "I-s/he" system, to the "I-I" system. "The place of auto-communication in the system of culture," the Russian semiotician Yuri Lotman observed, "is far more significant than is commonly supposed" (Lotman, 2000, p. 21).

In electrate ontology *flesh* replaces Aristotle's substance, flesh being the term the phenomenologist Maurice Merleau-Ponty used to name the dimension of being revealed in the works of modernist painters such as Cézanne and Paul Klee. Marcel Proust dramatized the signature scene of the little sensations in his novel, *In Search of Lost Time*. The little sensation triggers involuntary memory for Marcel when he bites into the madeleine, the tea biscuit he first tasted as a child, served by his beloved grandmother. The incident is catalytic, which is the electrate equivalent of analytic in literacy, provoking in Marcel's case *an experience of happiness* that motivated the composition of his novel in search of the origins of this feeling. Merleau-Ponty adopted Proust's memory as methodology. These events of flesh testify to the existence of a visceral *value*, manifested in aesthetic qualities: intensity, vividness, participation, at the level of primary being, *the feeling of being alive*. The modernist arts originating in cabaret entertainments researched and experimented with the formal properties of design addressing the human sensorium, to augment these values and their catalytic effects on memory and imagination. In the context of apparatus, Virilio's warning about digitizing phenomena becomes an exact description of flesh actualized within electrate metaphysics. *Electracy designs flesh*, just as literacy writes speech.

A reason for insisting on the cabaret setting of the invention of the vanguard aesthetic revolution is to notice its context not only as entertainment, but the setting of vice, the attractions and dis gusts of desire and appetite. Electracy undertakes a revaluation of all values (Nietzsche). The (impossible) imperative today: *reimagine vice*. The revaluation of appetite is most explicitly embodied in the themed casinos of Las Vegas, such as the Venetian or the Bellagio, family-friendly capitals of vice. Walt Disney, codifying if not inventing the theme park in Disneyland, is the Plato of electracy, actually building his model polis devoted to happiness and knowledge. To appreciate this analogy, we have to consider the micro-level of design. What is the sensation-operator of Disneyland? Sergei Eisenstein, inventor of cinematic montage, praised Disney for his revelation of the power of the *plasmatic line* (plasma = shape, equivalent of Eidos, Gestalt, Idea). Eisenstein had in mind the first animated sound cartoon, Disney's *Steamboat Willie* (1928). Eisenstein spoke for his peers in his admiration for Mickey Mouse, whose popularity at the time was rivaled only by Charlie Chaplin. The key for electrate metaphysics is that this line, originating as the line of depiction in the caves of the paleo apparatus, is not restricted to the naturalism of photography, but exceeds all mimesis to articulate an imagined fictional dimension in which shape may be distorted and reassembled according to its own formal potential guided by imagination. The energy of this line is what Derrida (1974) identified as trace, referenced in the *gramme* of grammatology (the trace of electracy). Most telling in the context of capability is the comment by designer John Hench, responsible for the portrait celebrating the 25th anniversary of the creation of Mickey Mouse, that the cartoon figure in graphic terms is a symbol of life, expressed in the circular curves of the mouse (the overlapping circles that resonate with the Borromean knot of Lacan's subject topology). "The curves suggested fecundity, regeneration, babies and breasts, the lost safety of the womb, the life impulse itself" (Snow, 2019, p. 115). Aphrodite wins again. This world mouse personifies electrate thing. Hench characterized Disneyland as the architecture of reassurance, even if for its critics Disneyland is judged as the happiness accident.

The analogy between Plato and Disney affords a systematic comparison of apparatus: (1) Academy, school/Disneyland, theme park. (2) Science, reason/Entertainment, fantasy. (3) Dialogue, discourse/Cartoon, animation. (4) Socrates, dialectic/Mickey Mouse, gag. (5) Proof, persuasion/Event, enjoyment. The list could continue, as a way to appreciate the historical inevitability of Disneyization, the theming of everyday life now well advanced globally, as the electrate equivalent of secularization by which literacy displaced orality in the social organization of experience. The functional point is to note the opening of chora in each case, of a place for science in school within an oral society; a place for fantasy in theme parks within a literate society. The predictive value of the analogy is to note the trajectory of history, that theming will have become hegemonic in guiding collective thriving (or not), including global policy formation. The cabaret origins suggest that theming of the world is also the bohemianizing of lifestyle: Lumpenproletariat as macro-subject. The specific character of electrate chora is manifested in Disney's intention that entering his park was like Alice passing through the looking glass into another world (before Mickey Mouse, Disney created a series of films mixing live photography and animation based on the Alice story). Slavoj Žižek observed that the best emblem of Lacan's part-object is the grin of the Cheshire Cat.

The place opened to institutionalization by electrate chora has been present in the structure of narrative from the beginning of storytelling. A prototype is *The Wizard of Oz*, which serves as an electrate equivalent of Plato's allegory of the cave. Dorothy begins at home in Kansas, confronting a quotidian life problem—having her dog Toto taken from her to be put down. The tornado that blows her and the house into Oz is the disaster we are generalizing as the Anthropocene, the accident of human cause. The passage from the first to the second act of the narrative, from Kansas to Oz, transports Dorothy into *the other scene*, like Alice through the mirror, the scene Freud called the unconscious, a realm with its own rules and tests for the protagonist, as documented in Joseph Campbell's *Hero with a Thousand Faces*. The bonus in the case of *The Wizard of Oz* is that the companions Dorothy encounters following the yellow brick road—Scarecrow, Cowardly Lion, and Tin Woodman—personify the three capabilities (thinking, willing, feeling),

whose powers are in a state of privation (Aristotle's *steresis*), potential only, virtual, unrealized, like those of students on the quest of education. The lesson of the allegory is that this fictional other scene is now being realized, materialized, objectified in digitalized perception, *constructing a virtual dimension* compensating for the collapse of time and space in dromosphere. In the coming community, reality will be virtual, a dimension imagined until now in the modality of *impossible*, evoked only indirectly. Realization of this virtual order is a pharmakon, manifested equally in Otaku (Japanese) fan culture and QAnon conspiracy movements.

Theming mystory

Disneyland as a hyperreal Republic of fantasy provides instruction in the nature of plasmatic design, a dimension of fiction but not mimetic: a *simulacrum* of place. A Yale professor of Architecture observed this quality of Fantasyland, describing the cable car ride above the castle and through the paper-mache Matterhorn. "Now nobody thinks that mountain is the Matterhorn or even a mountain, or that those bobsleds are loose upon its slopes—slopes being on the outsides of mountains. Yet the experience of being in that space is a real one and an immensely exciting one, like looking at a Piranesi prison or escalating in the London Underground" (Sklar, 1967, p. 80). The comparison with Piranesi is most telling, the highest compliment coming from an Architect, given Piranesi's status as the ultimate designer of impossible architecture.

Another instance of this original dimension, relevant to our example of Gauguin's failed attempt to live his fantasy Tahiti, is a psychologist's description of the effect of the Tiki Room. "In a fake hut, fake parrots play-sang not very estimable tunes, but the colors were a riot of rainbows and the parrots moved their beaks in precision—now this group, now that, never faltering, always surprising. The great totems (Tiki) in various corners of the hexagonal room broke into mobile faces, singing and chanting, and soon the songs of men and birds were joined by the songs of flowers. It was like a moment from dimly remembered, complicated dreams" (Sklar, 1967, p. 111). The psychologist reported feeling himself and observing in those around him a reverence that he identified as *satori*. He credits the imagineers with having created an *anti-bomb*, equal in technological and imaginative skill to the atomic bomb haunting everyone during this period (fantasy against anxiety). The Tiki Room displayed a Polynesian theme, a simulacrum of Tahiti. It is a confrontation of two energies, libidinal and atomic. Gauguin's misunderstanding of the relation of art and fantasy is remedied in the AI GAN (generative adversarial network), capable of generating photorealistic images from diagrammatic sketches, dubbed GauGAN by its developers (Leach, 2022). AI is becoming heuritic, as it must to democratize augmented imagination. The lesson is not (only) to admire a document of someone else's fantasy, but to explore one's own in real time. Meanwhile, we can imagine a postmodern scene: "Gauguin visits the Tiki Room."

Paul Virilio declared Disney to be more radical, more avant-garde in his productions, than the modernist painters themselves. Appreciation of his claim requires recognizing in Disney the cabaret spirit of *fumisme*, expressed in design terms through caricature and camouflage (exaggeration and deception). Disneyland as simulacrum is a caricature of Plato's *Republic*. Plato declared there are two kinds of image: copies (likeness, representation), and simulacra (appearance). Caricature as device of simulacrum produces the hyperreal, exploiting the plasmatic line to *seem more real than reality itself*, with Disneyland as the prototype of hyperreality, the dimension opened in electrated chora (Baudrillard, 1983). The paradox is that a caricature of a famous person is more recognizable than a photograph, having to do with the nature of perception and memory. This lesson has been learned by Elvis impersonators, who report that the Elvis effect requires only three gestures (lip sneer, hip swivel, rubber leg motion). Las Vegas casinos apply experience design to create Venice (The Venetien), that is more like Venice than Venice itself, according to tourist promotions. The Venetien is a simulacrum of Venice, not a representation (it is a copy for which there is no original). The theme effect is produced by providing just a few exaggerated features, accidents of flesh, enough to catalyze memory and imagination, to solicit personal disposition into participation, recognition, which is to say, to interpellate. These attunements are to digital media what logical entailment is to alphabetic writing. The little sensation is the enthymeme of electracy. The implications for solving the problem of information overload are important: *bachelor machine meets simulacrum in a game engine supporting intuition of dromosphere*. Here is an assignment for electrated AI.

Interpellation brings us back to the dialectic of danger and saving power, in noting the echo of theming in Holton's definition of the discovery side of science as *thematic*. This resonance of terms is not accidental, prompting a shorthand version of electrated pedagogy as *designing a personal theme park*. Disney carried around in his mind a vision of the park for many years before it was built. It is a modern update of a memory palace, like the one designed (but never built) by Camillo beginning in 1519 for learning all of Cicero. Personal memory theaters were a necessary feature of literate education up to the invention of print, along with commonplace books for recording a personal archive of information and topical arguments. The ground level of the theater was a familiar street, like that of one's hometown, along which were placed at regular intervals striking images (second level) associated with knowledge, the information to be memorized (third level). When students orated, they walked in imagination through this scene. Today we know that beneath this mnemonic street is not a beach (as the Situationist said), but the infant body of semiotic chora.

Mainstreet USA, entrance to Disneyland, is based on Disney's idealized recollection of Marceline, Missouri, his hometown, generalized to evoke turn of the century America. The street leads to the town square at the center of the park, a "hub" (vortex), from which radiate the four lands: Frontier, Tomorrow, Adventure, and Fantasy—past, future, nature, imaginal. Together the four scenes constitute Disney's version of the modernist catechism, and the themes are his models for answers. The park as a whole was designed by filmmakers working for Disney's studio, not architects, objectifying a kind of visit to *the other scene* created by movies. Christian Metz argued that the experience of watching movies is regressive, producing a state of potential satisfaction recalling the original transitional space of early childhood (catalyzing semiotic chora). In the TV series Disney made with the ABC network leading up to the opening of the park in 1955, narrated by Disney himself, each land is associated with specific stories,

registering Disney's personal themes. Frontierland is introduced through the life of Davy Crockett; Tomorrowland features the rocket designs of Wernher von Braun; Fantasyland by fairy tales, especially the folktales of Uncle Remus; Adventureland by "reality" adventures in nature. America emerges as tourist macro-subject in its Disneyland memory palace.

Disney's creation may function as a relay for students exploring and composing a representation of their own themata, serving as a contemporary cognitive (or conative) map. Investigating the thematic dimension of knowledge creation, Holton (1973) identified an *image of wide scope* (wide image) organizing the imaginations of those responsible for scientific revolutions. Holton identified the wide image as a pattern emerging within the *oeuvre* of innovators. The pattern was constituted by four or five images, anchored in a scene of childhood memory. When a paradigmatic theory becomes anomalous, it is replaced hypothetically during the discovery phase by the innovator's unique wide image, formatting imagination. The wide image is vehicle for the object @ (cause of desire), making it the metaphysical operator of electracy. This connection between body and design is what digital technologies natively augment. The prototype is the career of Albert Einstein, and the memory he recorded in his autobiography, of the gift of a compass from his father when Albert was four years old. The uncanny symmetry between the compass and the forces of electromagnetism makes Einstein's case paradigmatic. The wide image is just that: a vehicle and a tenor. The vehicle for Einstein is the compass; the tenor (the theme) is "invariance." The theory of relativity has been called a theory of invariance, naming Einstein's primal intuition about reality. *Vehicle*: the compass needle pointing north; *Tenor*: the speed of light. Subsequently, analysts have identified wide images in the careers of hundreds of figures across all fields of culture (Briggs, 1990).

The genre of *mystory* was created to enable students to discover and design their wide image at the beginning of their productive lives (Ulmer, 2003). *Mystory* retrieves the pedagogical tradition of memory palace from manuscript culture, to connect a primal scene of childhood experience with a disciplinary aporia. It is an investigation of interpellation, registering the position one *recognizes* emotionally (punctum) in each of the primary institutions of apparatus: family (home); mythology (entertainment); history (community); career (discipline as anagogy). The wide image diagram emerges from the topological invariances registered across the four theses. A wide image is a *fable* in Nietzsche's account, replacing the dialectic of appearance/reality (truth) with the central event of electrate learning: *the intersection of an aphorism of thought with an anecdote of life*. In this bachelor machine, external macro events allegorize individual intuitions—the disaster in me (my Anthropocene). Such is the event of theopraxis made possible online: a convergence of human faculties and their institutional projections into one holistic "world." The caveat is that the state of mind native to this virtual dimension was modeled in the first novel of modernism, *Madame Bovary*—Bovarysme (Serres, 1994).

The life of James Joyce exemplifies theming the ISA instructive for *mystory*. Joyce oriented himself in his existence (EPS) through identification with four figures. *History* (School, Community): "Young Joyce's imagination was captured by the cause of Irish freedom, whose most prominent spokesman at the time was Charles Stewart Parnell, a national hero who suffered a tragic fall. He became in the Irish imagination the type of the betrayed hero" (Litz, 1972). *Church* (Religion): Joyce's formal education took place in schools run by the Jesuit order. His most important religious experience occurred at Belvedere College, when he was elected Prefect of the Sodality, that is, "head of a group of students who banded together for the purposes of devotion and mutual help." Although Joyce broke with the Church, as Litz observed, this stance carried over to his view of the artist as secular priest. *Family* (Paleo): The defining problem of Home for Joyce was his ambivalence toward his father, whose chief interest in life was "jolification." "The inefficiency of Joyce's father and his wasteful habits gradually undermined family finances and family solidarity." *Career*: At age eighteen Joyce published a review of Ibsen's play, *When We Dead Awaken*. In a letter he sent to Ibsen, the student Joyce explained the most important reason for his admiration: "How your willful resolution to wrest the secret from life gave me heart and how in your absolute indifference to public canons of art trends and shibboleths you walked in the light of your inward heroism" (Joyce, in Litz 24).

Joyce's EPS compass: Parnell, priest, his father, Ibsen. These figures manifest Holton's *themata*, perhaps summarized in the completed sentence of Ibsen's title, expressing Joyce's own life mood: "When we dead awaken, we will see that we have never lived." Such is the epiphany structuring Joyce's best-known *Dubliners* story, "The Dead." What makes life worth living? The wide image is an epiphany relative to this question. Poetic epiphany is the syllogism of electracy. If human capability is the cause of Anthropocene, human creativity is the remedy. Students are introduced to electracy by composing a *mystory*, to discover and design their own theming of capabilities, the target of marketing but also source of creativity, perhaps to play their wide image against disaster in AI game architecture. Our best chance collectively against the Anthropocene (creativity as pharmakon), the diagrams of a coming reimaged community will have originated in the transitional objects of children.

References

- Althusser, L., 2001. *Lenin and Philosophy and Other Essays*. Trans. Ben Brewster. Monthly Review Press, New York.
- Bateson, G., 1979. *Mind and Nature: A Necessary Unity*. Bantam, New York.
- Baudrillard, J., 1983. *Simulations*. Trans. Paul Foss, Paul Patton and Philip Beitchman. Semiotext(e), New York.
- Benjamin, W., 1999. *The Arcades Project*. Trans. Howard Eiland and Kevin McLaughlin. Harvard University Press, Cambridge, MA.
- Blumenberg, H., 1983. *The Legitimacy of the Modern Age*. Trans. Robert M. Wallace. MIT Press, Cambridge, MA.
- Booth, A., Posner, M., 2020. "Introduction—the materials at hand," special topic: varieties of digital humanities. *PMLA* 135, 1.
- Briggs, J., 1990. *Fire in the Crucible: The Self-Creation of Creativity and Genius*. Tarcher, Los Angeles.
- Derrida, J., 1974. *Of Grammatology*. Trans. Gayatri Chakravorty Spivak. Johns Hopkins University Press, Baltimore.
- Erikson, E.H., 1982. *The Life Cycle Completed: A Review*. W. W. Norton, New York.
- Fukuyama, F., 2002. *The End of History and the Last Man*. HarperCollins, New York.

- Havelock, E.A., 1963. Preface to Plato. Grosset & Dunlap, New York.
- Heidegger, M., 1977. The Question Concerning Technology and Other Essays. Trans. William Lovitt. Harper Colophon Books, New York.
- Holton, G., 1973. Thematic Origins of Scientific Thought. Harvard University Press, Cambridge, MA.
- Kristeva, J., 1982. Powers of Horror: An Essay on Abjection. Trans Leon S. Roudiez. Columbia University Press, New York.
- Kuhn, T., 2012. The Structure of Scientific Revolutions. The University of Chicago Press, Chicago.
- Leach, N., 2022. Architecture in the Age of Artificial Intelligence. Bloomsbury, New York.
- Litz, A.W., 1972. James Joyce, revised ed. Hippocrene Books, New York.
- Lotman, Y.M., 2000. Universe of the Mind: A Semiotic Theory of Culture. Trans. Ann Shukman. Indiana University Press, Bloomington.
- Lotringer, S., Paul, V., 2005. The Accident of Art. Trans. Michael Taormina. Semiotext(e), New York.
- McLuhan, M., McLuhan, E., 1988. Laws of Media: The New Science. University of Toronto Press, Toronto.
- Manovich, L., Arielli, E., 2022. Artificial Aesthetics: A Critical Guide to AI, Media and Design. <http://manovich.net/index.php/projects/artificial-aesthetics>.
- Morton, T., 2013. Hyperobjects: Philosophy and Ecology After the End of the World. University of Minnesota Press, Minneapolis.
- Nussbaum, M.C., 2011. Creating Capabilities: The Human Development Approach. Harvard University Press, Cambridge, MA.
- Ong, W., 1982. Orality and Literacy: The Technologizing of the Word. Methuen, New York.
- Sen, A., 2009. The Idea of Justice. Harvard University Press, Cambridge, MA.
- Serres, M., 1994. Atlas. Julliard, Paris.
- Sklar, M., 1967. Walt Disney world: background and philosophy. In: Harrison "Buzz" Price Papers, vol. 160. <https://stars.library.ucf.edu/buzzprice/160>.
- Snow, R., 2019. Disney's Land: Walt Disney and the Invention of the Amusement Park That Changed the World. Scribner, New York.
- Stiegler, B., 2013. What Makes Life Worth Living: On Pharmacology. Trans. Daniel Ross. Polity Press, Malden, MA.
- Ulmer, G.L., 2003. Internet Invention: From Literacy to Electracy. Longman, New York.
- Virilio, P., 2003. Unknown Quantity. Thames & Hudson, New York.

Artificial intelligence in literacy

Dom Massaro^a and Jared Bernstein^b, ^a Department of Psychology, University of California, Santa Cruz, CA, United States; and
^b Analytic Measures Inc., Palo Alto, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to Artificial Intelligence and Literacy	529
Components of reading skill	529
Spoken and written language vs. formal/informal	530
Artificial Intelligence (AI) delineated	530
Traditional, current, future approaches to reading instruction	531
Traditional reading instruction	531
Pre-digital reading instruction	531
Current AI tools applied in current contexts	532
The extended mind (& AI-in-loop, human-in-loop)	532
Technology aided assessment	533
Automating Oral Reading Fluency (ORF) assessment	533
Moby.Read: an automated ORF test	534
Automated tutoring	534
Phonics with Baldi	535
Read With Me!	535
Understand My World	536
Reading practice	536
Research foundations for future AI	536
Electronic databases	536
Mining databases to optimize reading complexity	537
Mentoring by social robots	537
Predicting the future	538
Sudowriters	539
Challenges to AI's role in literacy	539
Limitations in embodying us	539
Retrospective	540
References	540

Introduction to Artificial Intelligence and Literacy

Literacy involves knowledge and skills, which develop and deepen over time. This includes knowledge of print conventions on paper and on screen, of how characters relate to words, and of how words and other language structures relate to meaning. Full adult literacy entails some facility in applying this knowledge at a useful rate that yields accurate understanding of the intent and content of a variety of text forms—from tables and graphs, to recipes, and to extended narratives and advanced textbooks.

This article addresses some current impacts on literacy instruction that artificial intelligence is having now and some foreseeable impacts that AI can have in the future as literacy evolves. The focus here is on early literacy development in young children with no perceptual or cognitive impairments, and on instruction as practiced in school and at home in a North American context. The impacts are illustrated within a three-period timeframe: traditional practice, current AI applications in practice, and future applications and practice. Each time frame is illustrated with reference to a triad of Teacher, Student, and Student Performance. **Fig. 1** shows a simple instruction cycle within the basic triad structure.

The article does not address any recent controversies in reading pedagogy or current approaches to situate reading instruction and literacy development within social or cultural contexts. It also does not focus on the impact or amelioration of physical or cognitive disabilities that affect reading development. The assumption will be that most reading instruction has practical goals and teachers typically work with an eclectic set of resources that may be commonly at hand.

Components of reading skill

A simple theory of reading [historically associated with Charles Perfetti ([Perfetti and Hogaboam, 1975](#)) and Philip Gough ([Gough and Tunmer, 1986](#))] posits that reading skill has two main components, decoding (D) and listening comprehension (LC), and that some level of skill in both is required to achieve useful reading comprehension (RC). Decoding is the process of transforming

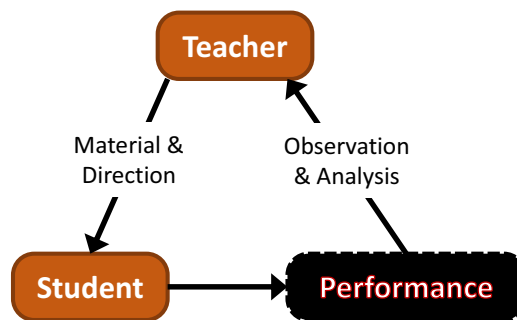


Fig. 1 A basic reading instruction cycle schematized within the teacher-student-performance triad.

a sequence of one or more characters into a likely phonological sequence and/or into a lexical item, that is, into a word. Listening comprehension is the process of transforming a spoken utterance, such as a conversational turn, into a likely sequence of linguistic forms (words, phrases, sentences), and with these linguistic forms, into the meanings that the speaker intended to convey. The theory is summarized in one simple equation: $RC = D \times LC$.

Because the operation between D and LC is multiplication, if either D or LC is equal to zero, then RC will also be zero. A more general term that subsumes decoding is Word Recognition. Word recognition includes decoding, as such, along with direct apprehension of word forms, known as sight-word recognition. Both components are needed for reading, and each of the two main components of reading skill in turn has several components, listed below for convenience.

- Word Recognition: knowledge of (alphabet, letter-to-sound relations, alphabetic and non-alphabetic conventions, affixes and their conventional spelling rules, etc.).
- Listening Comprehension: knowledge of (sound-to-lexical map, phrase structure, semantics, rhetorical conventions, etc.)

Spoken and written language vs. formal/informal

Learning to understand speech is never an issue as it is for written language. Conceivably our evolutionary history of spoken language relative to written language is responsible for this difference (Merlin, 1991; Pinker, 2007). Writing was originally met with some skepticism from early philosophers who thought that our humanity was compromised because memory was less essential with writing than with speech.

Table 1 illustrates examples of communication media in which the formality of the language and its modality can be categorized independently of one another. Although we usually think of informal spoken language and formal written language, a TED talk is formal and spoken whereas texting is informal and written. For example, the formal nature of picture books allows the writer to make more deliberate word and grammatical choices that are not always feasible in child-directed speech (CDS). CDS and adult-directed speech (ADS) requires a spontaneity (Grice, 1975) that necessarily limits word choice and allows deixis to substitute for various words and grammatical constructions.

Artificial Intelligence (AI) delineated

Note that machine learning (ML) is different from AI, and we can consider how the two relate to reading. Note that both ML and AI are relatively recent terms, so their meanings are still in flux. Briefly put:

AI refers to automated actions that seem appropriate to biological intelligence;

Table 1 Gives a taxonomy of the potential independence of language modality and formal (non-conversational) versus informal (conversational) dialog.

	<i>Spoken</i>	<i>Written</i>
Formal	TED talks, lectures, MOOCs	Books, articles Newspapers
Informal	Conversations, interviews TV media dialogs	Texting, messaging Email

ML is an engineering field that develops technology to build AI systems.

ML refers to algorithmic processes that operate on data sets to produce algorithms that cluster, classify, or identify patterns in new (unseen) data sets. Typically, an algorithm infers a function (or a model) that assigns labels to new data from an analysis of labeled training data. For example, thousands of transcribed voice recordings are analyzed by an ML procedure, which produces an algorithm that transcribes new voice recordings.

Artificial Intelligence (AI) is an ability of automated systems to perform tasks that until recently required human or other biological information processing, often including signal analysis, decision, and sometimes, verbal and/or mechanical response. People think of AI as the use of automated systems that perform tasks which traditionally have required human, or at least vertebrate, intelligence. For example, cars were designed to be driven by people, but now can drive themselves as well as an average human driver. As we delve into AI and reading, note that self-driving cars are possible in part because many millions of miles of car travel (on the road with multiple sensors) have been logged and analyzed with machine learning algorithms. Obtaining and analyzing large databases for machine learning has been essential in autonomous vehicle development, as in other domains. Thus, we expect that large databases will be necessary for the development of AI applications in reading too.

Traditional, current, future approaches to reading instruction

The remainder of this article is divided into three sections:

- Traditional reading instruction,
- Current reading instruction as augmented with AI tools/applications, and
- Future reading instruction, including future literacy practices more generally.

The first section is a brief overview that may help a reader locate the needs that current AI tools and applications address and the benefits that accrue from these. In the second section, we learn that most of the current AI tools do not disrupt the flow of reading instruction. Instead, they facilitate current practice and render it more fun, more efficient, more accurate, or more effective. AI may radically change the future of reading and reading instruction altogether. The last section describes a few current research initiatives that may further improve current instructional practice and a foreseeable future in which AI capabilities that are likely to change the nature of information access generally and change the nature of reading development and instruction along the way.

Traditional reading instruction

Pre-digital reading instruction

In the United States, between 2000 and 2020, reading instruction in schools largely followed patterns that were informed by a model of reading that was set forth in a report from the NICHD (2000), which [Tim Shanahan \(2006\)](#) usefully summarized in 2006, setting out a rough set of five main layers of skill/instruction that support successful reading. These five layers are shown in [Fig. 2](#).

We will assume that students have sufficient phonemic awareness to start learning to read. First comes the establishment of print awareness skills—the student can recognize and name all the lower- and uppercase letters and say what their most common sound value is. Then, students are introduced to the most common sight words (e.g., the, in, can, no, how), and to the basics of phonics, or letter-sound correspondences, going beyond the most common sound value for each letter. These foundational skills are usually established before the “Instruction Loop” shown in [Fig. 3](#) gets started.

When the instruction loop starts, teachers direct students to read from leveled text material—starting with very simple materials that have a picture and only a few words or a short sentence per page. The teacher then informally, monitors the literacy

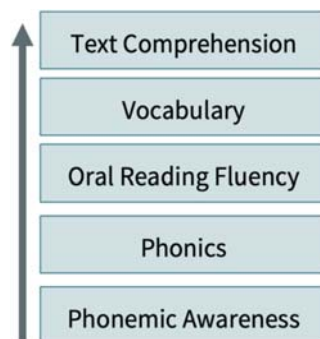


Fig. 2 Five elements of reading.

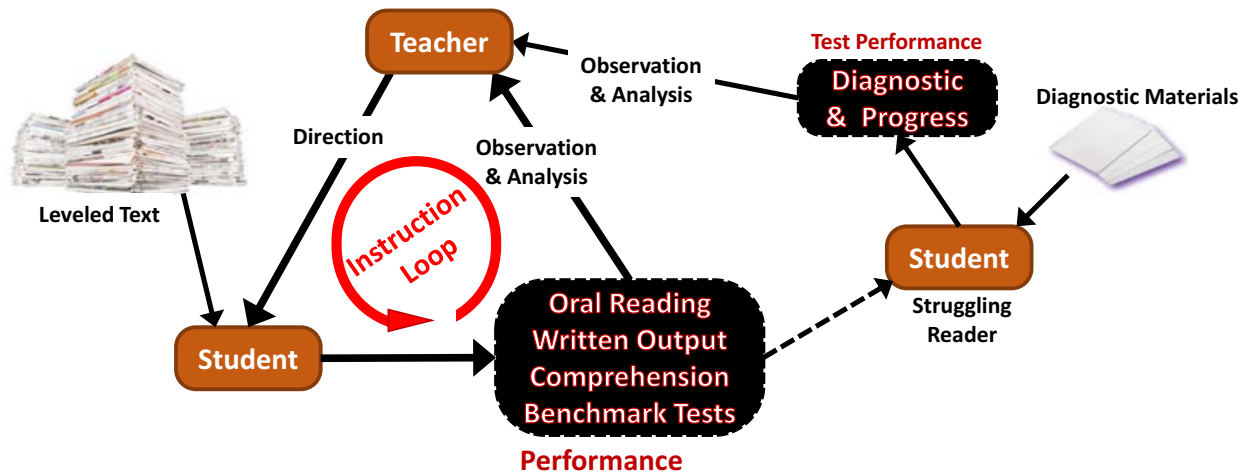


Fig. 3 A fuller version of a typical reading instruction cycle in grades 1–3, the dashed line shows an added path for students who are identified as struggling readers. Two static resources are shown: leveled text and diagnostic materials.

development of each student by observing as the student performs in a variety of tasks that may include reading aloud, writing, understanding texts, and sometimes in periodic formal benchmark tests that may be required of all students. As time goes by, if students perform at about the level expected for their age and grade, teachers gradually increase the difficulty of the leveled text material that students read, and the instruction loop continues, perhaps over 3–5 years, until students can read efficiently enough to understand and enjoy age-appropriate stories and books, and can independently read grade-appropriate academic text to learn new material.

However, for struggling readers who perform below expectation on benchmark tests, who are not performing at grade level, or who simply lag behind the rest of the class, a teacher or a reading specialist may administer a set of formal tests and score an individual student's reading performance. Diagnostic analysis of the individual test performance helps determine which, if any, of the component reading skills may be hindering the student's reading development. A teacher or reading specialist will then select suitable materials and give them to the struggling student with directions and encouragement, as needed.

Given reasonable curriculum materials and sufficient time with a qualified teacher over a 3- or 4-year period, most young readers will emerge with basic literacy skills and a solid foundation for gaining broader and deeper literacy with the practice they encounter in late elementary and middle school (grades 5–9, or ages 9–13 years). By the end of 12th grade (age 17–18), after many hours of reading in the instruction loop and then across the curriculum and in life outside school, many skilled readers can comprehend text at a rate of 200 words per minute or more (*Average Reading Speed, 2022*). This is faster than the rate of most spoken language they encounter. The next section reviews (2021) AI applications that extend the range or the nature of literacy instruction, improve the efficiency of reading development, or improve accuracy in measuring reading and literacy-related skills.

Current AI tools applied in current contexts

The extended mind (& AI-in-loop, human-in-loop)

Artificial intelligence (AI) is performed by machines, but it usually involves a human participant. Some applications of artificial intelligence can be viewed as contributing to the extended mind (*Clark and Chalmers, 1998*), which involves the situational milieu we inhabit. We now take for granted that many AI tools and apps facilitate a variety of human cognitive processes, as they become more accurate and commonplace. In the period 2000–2020, several technologies have off-loaded complex cognitive demands, including, complex information retrieval, text-to-speech synthesis, automatic speech recognition (ASR), and automatic language translation.

When early text-to-speech (TtS) synthesizers were first commercialized in the late 1970s, they were crude but sufficiently intelligible for successful commercial application in reading machines for the blind (*Klatt, 1987; History of TtS, 2022; van Santen et al., 1997*). By 40 years later, TtS systems had achieved wide acceptance in many applications as the systems evolved to emit highly intelligible and much more natural sounding speech. Speech synthesis is still viewed as an AI technology, and two core components of TtS are products of machine learning: first, the text-analysis methods for word disambiguation and for construction of prosody; and then the augmented ASR technology that finds optimal sequences of signal snippets to concatenate for clear and natural sounding words.

Going back further, the invention of visual symbols and print changed what was possible and might be considered an early instance of AI for language and communication of information. The magic of making physical marks that could be seen at a future time by others enabled communication in a manner previously impossible—it allowed one person to communicate to another person across time. Writing preceded the recording of spoken language and it leaves out some of the collateral information in

spoken language, such as personal identity and voice-borne expression of emotion. On the other hand, written medium also enabled important representations of information such as tables and graphs that are not easily understood in their spoken form.

Viewing AI as an elaboration of the extended mind might put AI in a more acceptable framework in that it is then a product of the human mind that has the human mind as an essential partner in various applications. As Neil Stephenson, the legendary science fiction writer has stated, “Ideally we ask AI to function as an aid or assistant in our journey toward knowledge” ([Michael Shermer interview, 2021](#)). This framework doesn’t eliminate the recent fearful premonitions of runaway AI but perhaps puts them in a better light.

Technology aided assessment

Given that live teacher time is limited, robots (and we believe virtual tutors) can substitute for the teacher just as effectively if not more effectively. Our ideal virtual tutor would have assessment built into the tutoring regimen, and the goal is to have the student as engaged with the testing as she is with the reading and mentoring itself. Engagement of a virtual agent or a physical robot has been studied in a variety of learning contexts, but not systematically in reading and learning to read. Although lacking in systematic reading contexts, there are helpful findings on refining measures of students’ engagement, how various verbal and nonverbals cues from the agent are picked up, and how engagement can fluctuate during the interaction. If researchers and teachers explore virtual or robot mentors for reading instruction, there is a wealth of guidelines in the literature ([Belpaeme et al., 2018](#)).

Students don’t necessarily enjoy being tested particularly if they are somewhat unskilled, as most learners by definition are. Furthermore, it’s been found that in simple multiple-choice tasks in which a student selects one response from multiple options, many students respond so quickly that they could not have completely understood the question and the gallery of answers. One answer to this challenge has been proposed: the Technology Enhanced Item or TEI ([Bryant, 2017](#); [TEIs, 2021](#)), with creative testing formats. An example of a TEI might ask the student to highlight the relevant text in answering a question. In a very large systematic study, TEIs have shown the lowest rapid guessing rate, which traditionally reflects lack of student engagement: 10 times more often on multiple choice items than on TEIs.

Electronic tutoring systems endorse intelligent programming to create innovative formats in the testing environment. A student might be asked to drag and drop content from a text to answer the question. Another possibility would be to highlight the text or select multiple items from the text. In this sense TEIs offer the promise that students can be kept engaged in a reading environment even though assessment is part of that reading environment. Given that there is strong evidence that TEIs improve engagement, we can push the envelope even further by requiring more sophisticated interactions from the students such as open-ended answers or having the student actually perform a simulation or instructing the virtual mentor to perform a simulation. Ideally, an engaging tutoring session in reading would blur any existing boundaries between learning and assessment.

One challenge in applying social robots to develop reading skills and to include intermittent testing is to maintain the student’s engagement with the tutor. There is some indication that students’ interest diminishes when the novelty effect evaporates. One key to maintaining interest is to establish a meaningful relationship between the student and the tutor, and to reinforce the student’s memory of this relationship ([Kasap and Magnenat-Thalmann, 2012](#)). Using RoboTutor, [Boote et al. \(2021\)](#) tracked the tutee’s face to monitor engagement, and the partially successful goal of predicting whether the student would complete one of the exercises. This offers one source of information among many others such as properties of the voice, response latencies, and accuracy in the task that could increase the effectiveness of monitoring engagement. Interestingly, they were able to define an early bailout when the child didn’t like the activity and a later bailout which occurred during the interactions with the tutor. This distinction is a valuable one because different types of intervention would be necessary for success.

Automating Oral Reading Fluency (ORF) assessment

Periodic assessment of reading performance is important for guiding instruction, but a recurring problem is that students spend too much time in testing, and teachers spend too much time on activities that surround testing—in-service training, administering and scoring tests, and then interpreting results ([Bernstein et al., 2017](#)). Thus, automatic systems to administer and score ORF performances have gained acceptance and proliferated in the period from about 2015 to 2022 and onward. This has been accomplished with student-friendly online presentation of material in a browser, which is combined with AI capabilities on servers in the cloud.

The server-side AI comprises child-appropriate ASR and automatic scoring and scaling of the recorded read-aloud responses. Such systems can now (2021) return accurate ORF scores and running records of a student’s performance within one second after test completion. The most salient advantages of current automated ORF tests are their convenience and reliability. They are convenient for the student, because the tests can be self-administered at the student’s convenience, and an automated ORF test may occupy less student time. Such tests are more reliable in that they apply acceptance and scoring rules quite consistently, and they score readers without favoritism toward individual readers or groups. Automated ORF measurement has also been shown to be indifferent to irrelevant dialect differences and reading pronunciations of readings by speakers of English as a second language ([Balogh et al., 2012](#); [Bernstein et al., 2017](#)).

A key disadvantage in current automated testing and scoring is the absence of real-time system-initiated interaction with the student during the test administration. Such interactions from the system might include interruptions to encourage or give guidance, or simply to present the test-texts at a more appropriate level so they are within the student’s zone of proximal development.

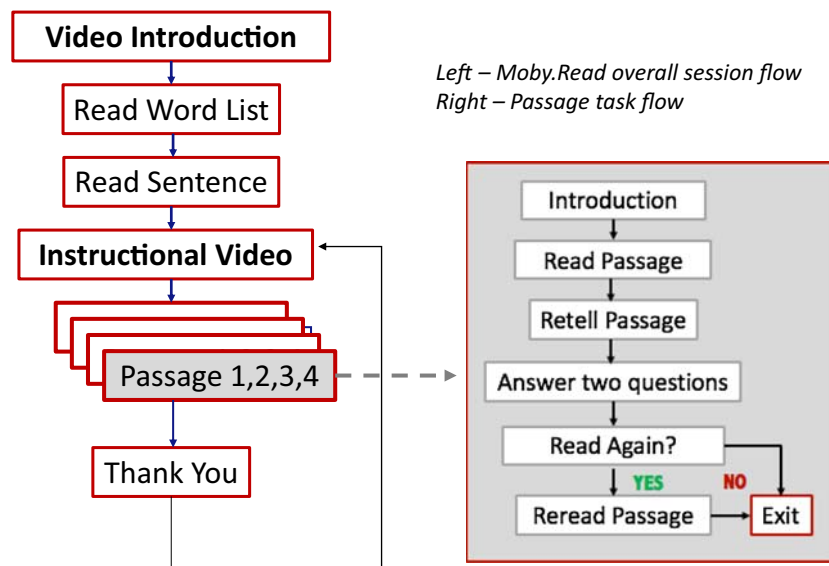


Fig. 4 Session and passage presentation flow for each passage (1 practice, 3 scored) in Moby.Read.

Moby.Read: an automated ORF test

The *Moby.Read*TM (2022) system is a fully automated and self-administered oral reading fluency (ORF) benchmark assessment designed and built in 2016–18 and introduced commercially in January 2019. *Moby.Read* is designed for benchmarking early reading performance three times a year. Inside the *Moby.Read* system, an ASR was optimized for children’s readings and their spontaneous speech in retelling text passages, and an augmented natural language processing (NLP) module enables immediate score reporting of five key oral reading fluency skills:

1. Reading Level,
2. Accuracy,
3. Accurate Reading Rate (in Words Correct Per Minute),
4. Comprehension, and
5. Expression.

During the assessment, as shown in *Fig. 4*, students read four passages aloud, summarize the passage content, and respond to short-answer questions using their own voice. *Moby.Read* embeds model readings and opportunities for students to “go back” and re-read a passage to hone their oral reading skills (see *Fig. 4*). An assessment takes about 12 min and runs in Chrome or as a native app on iPads.

Scores are available immediately so teachers can apply the most current possible information to individualize reading instruction. Additionally, audio recordings allow teachers to diagnose performance issues and to share them with reading specialists, parents, and with the students themselves (see *Fig. 3*). *Fig. 5* gives a sample presentation flow for each passage (1 practice, 3 scored) in *Moby.Read*.

Automated tutoring

Once interactive screen time became feasible, various commercial and school curricula of speech and reading tutoring were implemented in electronic exercises in speech, phonics, and vocabulary (*Phonics Exercises*, 2022). These involved very little AI other than utilizing speech and graphics within a given instructional-game design. Typically, children learn the alphabet and then how to sound out various spelling patterns. This method isn’t foolproof, however, as experienced by a grandmother with her grandchild sounding out a word written under a picture of an insect. The child expertly sounded out bug for the word but the word was actually *ant*. Another method has children spell various sounds. One commercial game reinforcing the sound of speech is *iSPY* in which the child is given a sound and matches words that begin with that sound (*iSPYPhonics Fun*, 2021).

Although most early reading instruction is grounded in phonics, a case can be made for increasing the beginning reader’s awareness of the language’s orthography. It seems reasonable that drill and practice in phonics also heightens a child’s awareness of these spelling patterns to facilitate letter and word recognition during reading. There is evidence that expert readers and children learning to read use spelling constraints to more easily identify letter patterns and words (*Venezky and Massaro*, 1979, 1987; *Massaro and Hestand*, 1983).

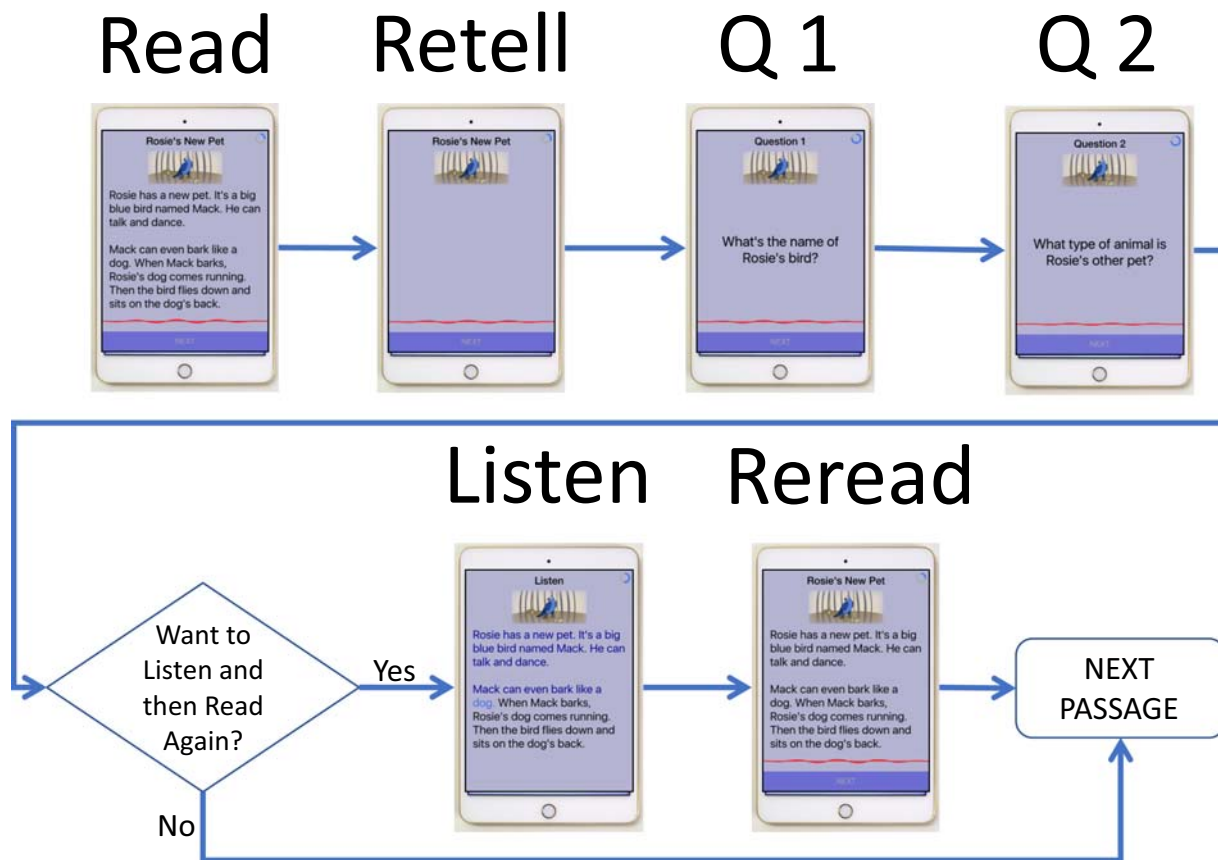


Fig. 5 A sample presentation flow for each passage (1 practice, 3 scored) in Moby.Read.

Phonics with Baldi

Given dramatic progress in creating a rich sensory world in educational domains, it is surprising that most current decoding exercises do not exploit multiple modalities. Given that seeing the speaker's face can contribute significantly to understanding and learning language (Massaro, 1998), and developments in computer animation of talking heads, it is now feasible to carry out phonics instruction and assessment bimodally. Various characteristics of the vocal tract can also be viewed from different perspectives, which might help the child (or second language learner) master the phonics and spelling of the language being learned.

The 3D computer-animated talking person, Baldi, was developed to produce accurate visible speech (Massaro et al., 2005; Baldi Tutoring, 2022a). Using this technology, a game was developed called Phonics with Baldi. This involves a matching game in which the child is required to match the animated pronunciation of a sound with a possible spelling. In the matching game, a successful match is rewarded by eliminating the matching pair from the display. The goal of the child is to complete all of the matches and move on to higher levels.

Read With Me!

Picture books are a foundation for a child's inauguration into a literate world (Gurdon, 2019). Although seldom put to rigorous experimental tests, they appear to be a prevailing influence on children's cognitive and linguistic development. Content analysis of picture books and child directed speech (CDS) documents that the linguistic and cognitive sophistication of picture books far exceeds that found in CDS (Massaro, 2015, 2017; Montag et al., 2015; Williams et al., 2022). Although electronic TV media consume more of the child's time, picture books best demonstrate to the child that a static medium can elicit language and ideas that can engage their attention. The caregiver-child dyad appears to be a prominent characteristic of the impact of reading these books to engage children.

What might be surprising is that this shared picture book reading contributes very little to a child's acquisition of the mechanics of reading (Massaro, 2012, 2017). By this we mean phonics, letter recognition, and written word recognition. Children do learn that books contain words and stories (or at least stories) that can transport them to another place and time. However, picture books aren't designed to highlight written language but rather primarily illustrate in pictures the story being told. Because of the emphasis on pictures, the print is often either considered unimportant or embellished to look less like print and more like picture design. The design of picture books encourages children to focus on the pictures rather than the words in shared reading-aloud situations (Evans and Saint-Aubin, 2013). Perceiving words requires focused viewing of the words. The print in most children's picture books is not

usually appropriate for a child learning the mechanics of reading (There is also some evidence, however, that children can learn some reading mechanics with picture books that emphasize the legibility of print, [Massaro, 2017](#); [Piasta et al., 2012](#)).

We might ask how AI can be utilized to complement the shared reading experience of picture books in order to reinforce reading mechanics. To promote learning the mechanics of reading during shared picture book reading, Dom Massaro created an Apple iPad app called *Read With Me!* (2022). At various times during shared reading, the adult reader can choose to display some text from the physical book in clear font on the iPad (*Read With Me! Demo, 2022*).

To use the app, the caregiver and the child choose one of their favorite books from the app's library. The caregiver reads the book to the child, both of them enjoying the rich sharing of emotion and adventure. *Read With Me!* adds to this experience by allowing the adult reader to dictate a sentence from the book to initiate the presentation of this sentence on the iPad screen. The app uses a robust method of ASR because it knows which book is being read.

The application allows the caregiver and child to share attention between the picture book and the text presented on the iPad. The child should have a good view of the screen when the dictated sentence is present. The iPad can be held by the caregiver, placed on a table, or placed inside a transparent holder sewn into a shirt that the caregiver would wear (*Read With Me! Demo, 2022*). This experience allows the child to experience written language in the manner that they experience spoken language. The *Read With Me!* app adds print without requiring the caregiver to provide explicit instruction in prereading skills. Many caregivers tend to have limited knowledge of or motivation to teach appropriate literacy skills. It remains to be tested whether this app or similar ones can help bootstrap a child's reading in the same manner that the speech toddlers hear aids spoken language acquisition.

Understand My World

Ideally, we would embrace an AI tutor reinforcing the mechanics of reading while simultaneously the child is primarily engaged in new and active learning. Using advances in AI, one example app, *Understand My World* (2022), immerses learners in spoken and written language as they are learning about their world. Equipped with a camera, microphone, and an internet connection, the student has direct access to unlimited information. Images captured by the camera are labeled and can be defined in both speech and writing. Participants can also speak and record their speech or ask questions. The app then presents the dictated words or answers to questions in either written or both spoken and written form. Some example questions are "Tell me about Sesame Street." "Who is Ben Stiller?" "How many miles in 7 km?"

Because evaluation is central to language interventions, the app records a history of the student's interactions with all three functionalities. This history can be utilized in several ways, such as recording time on task, the vocabulary encountered, the student's speech, and the content of the student's interactions. One possible evaluation, for example, would be to compare the student's mastery of the content experienced within the app to comparable content the student experienced outside the app.

Reading practice

Jack Mostow and colleagues, situated within a campus with a premier speech science department, developed a computer program that uses speech recognition to listen to children read. Beginning roughly 3 decades ago, their core technology was ASR, well before it evolved to have the quality it has today ([Mostow, 2013](#); *Project LISTEN's Reading Tutor, 2022*). The goal was to use ASR to listen to children read and guide each student individually in oral reading. Their reading tutor proved successful in a number of independent studies when compared to independent reading practice and even tutoring by certified teachers ([Mostow, 2013](#)). Dependent measures included word identification, oral reading fluency, and comprehension.

More recently, a number of reading tutors have been deployed. Reading with companion robots will definitely become more commonplace, with their prices within the range of other appliances such as TVs. In one representative study ([Sauppé and Mutlu, 2014, 2015](#)), a reading companion/mentor robot, Minnie, interacted with children during the child's reading. Minnie was not programmed with the latest speech technology as in applications assessing oral reading fluency but she did make book suggestions, guided and monitored the child's reading goals, and made comments to support their understanding and interest in reading and science. In the near future, we expect assessment and instruction to be embedded in tutoring interactions. Given Minnie's portability, she could provide time on task for reading at home as well as in school. Based on their pilot studies ([Sauppé and Mutlu, 2014, 2015](#)), children appeared to make a strong social connection with Minnie, which agrees with findings in other domains. We found this to be true of Baldi, an animated tutor, for both hard of hearing and children on the spectrum ([Massaro, 1995](#)).

Research foundations for future AI

Electronic databases

Some of the major achievements of AI are critically dependent on acquiring and analyzing electronic information. We all know how the internet has consolidated many unfathomable sources spanning the spectrum of knowledge. Students and professionals alike can easily access Wikipedia and other websites to find the latest state of knowledge in almost any domain. The AI programming to make this information efficiently retrievable and useable has led to impressive intelligent AI systems such as IBM's Watson, which was able to beat the champion Ken Jennings in a Jeopardy competition (*Jeopardy, 2022*).

AI's intelligence has increased exponentially in effectiveness hand-in-hand with computer's speed and memory. Data mining recording and analyzing linguistic content has also grown exponentially. We've learned that the amount and quality of language children experience is central to the path to literacy ([Hart and Risley, 1995](#)). This observation has motivated researchers to

determine important academic words to use in schooling (at perhaps the expense of non-essential words, acknowledging that time-on task is limited).

Children are talked to, read to, and in many cases watch TV Media. Different estimates of the amount of input 2–5-year-old children experience from these three different contexts are roughly 150 min for both CDS and TV Media, and only about 15 min of parents reading (picture books) to their children (Media, 2022). Massaro (2015; unpublished) analyzed potential vocabulary and grammar differences in these three linguistic environments. Content from each of these three media were analyzed to record their vocabulary and the reading level. To determine vocabulary differences among the three media, the three samples were assessed against two independent measures of frequently used spoken and written words (Brysbaert and New, 2009; COCA, 1990–2012; SUBTLEX, 2021). Children experienced over twice as many rare word types with picture books than with the other two media (Massaro, 2015; Williams et al., 2022).

The formal nature of picture books allows the writer to make more deliberate word and grammatical choices that are not always feasible in CDS and written in TV scripts. As mentioned previously, the spontaneous nature of conversational speech constrains word and grammar selections and encourages use of gesture and deixis. Using Hiebert's (2005) list of academic words, these rare words actually corresponded to important words for schooling. Picture books had about 2 or 3 times the number of these academic words than did Parent Speech and TV Media. Examples of academic words found in the Picture Book database but not in the other two media are *centers*, *heated*, *lowering*, and *winding*.

Words are also more valuable for schooling if they have high dispersion, that is, found in different content areas. Picture Books provided about twice as many high dispersion words relative to the other two media. Examples of words with a high dispersion value in picture books that did not occur in CDS and TV Media were *behave*, *circulation*, *exchange*, *intricate*, *strain*, and *switch*. Using data mining, we learned that picture book reading provides valuable vocabulary for future schooling.

Mining databases to optimize reading complexity

Paving the road to literacy would ideally have linguistic and cognitive complexity of spoken language that is representative of material that the child will be reading in school. We expect that more complexity would be valuable. Grade level readability measures (Readability Formulas, 2021) were used to determine the reading grade level of the three media just discussed, as well as ADS. Children's picture books had a higher reading grade level than CDS and TV Media, as well as ADS. The data mining revealed that reading picture books to children is beneficial, as it appears to familiarize them with more challenging language and a more extensive vocabulary than they would otherwise encounter (e.g., van Kleeck et al., 2003).

Mentoring by social robots

A robot track man that calls balls and strikes in baseball more accurately than human experts is a prototypical instance of artificial intelligence (AI Umpire, 2022). It mimics intelligence of humans but in this situation also does it better. It picks up information from the trajectory and location of the ball and computes if each pitch is in the strike zone. Given high-quality sensors and a simple AI algorithm, an accurate call is assured, significantly better than a live umpire.

Imagine a child learning to read with a personalized robot companion with knowledge of the child's reading level, including vocabulary and grammar, and general knowledge. In addition, the robot will have goals including one to advance the child's literacy, and will continuously learn about the child and the appropriate tutoring required to achieve the goals. With this knowledge, the robot can then anticipate where the child could be tripped up by the text and could modify and even embellish it appropriately to make it compatible with the understanding of the child. In this case, the robot would provide the affordances necessary for acquired literacy. This would also exploit the zone of proximal development described by Vygotsky (2022).

Two exploratory studies provide some incentive to determine whether social robots are facile enough to encourage and aid a child at various levels of reading skill, such as phonics, vocabulary learning, and fluent reading aloud. Michaelis and Mutlu (2018) found that a social robot engaged early adolescent children in a variety of guided reading activities. The children claimed that the robot helped with comprehension and provided them with motivation to read. Yueh et al. (2020) used a social Robot Julia to facilitate library literacy activities with elementary school children.

There has been a plethora of studies evaluating the effectiveness of robots as tutors in a variety of domains other than reading instruction and assessment (Belpaeme et al., 2018). According to a Belpaeme et al. meta-analysis of hundreds of studies, tutoring robots have been cognitively and affectively effective in many different domains. Initially, there seemed to be an advantage to having physically-present robot tutors in place of virtual tutors on a screen. A meta-analysis of 33 experimental studies found that "a co-present robot to be more persuasive, receive more attention, and be perceived more positively than a virtual simulation of the agent being tele-present" (e.g., presented on a screen, Li, 2015). However, there is also more recent evidence that virtual tutors can be just as persuasive as co-present tutors (Deng et al., 2019).

Although it is not certain that any difference between virtual and physically present tutors would occur with a reading tutor, it would be more practical and less costly to program tutors on screens than to provide every child with a physically present robot (Deng et al., 2019; Li, 2015). Students might also demand that their robot be unique in both physical and psychological characteristics. Advances in Augmented and Virtual Reality might diminish this gap between physically-present and on-screen social robots so that it would be more feasible for each student to have a personalized robot. On the other hand, we might learn that something

like physical presence is important, perhaps a carryover from the child sharing picture books with loved ones. The final word about virtual tutors and co-presence has yet to be uttered or written but their destiny will have a role in reinforcing literacy.

Although this review covered normative reading instruction and practice, it is worth mentioning that robotic tutors should have even greater potential for children with challenges such as hearing loss or existing on the autism spectrum. Baldi, introduced earlier, proved a successful tutor for both these groups of children. One major advantage of Baldi, like any robot, is his availability 24/7 and programmed extreme patience and encouragement. In one-on-one tutorials, both populations of children learned and retained new vocabulary and grammar. Children on the spectrum, who tended not to look at faces, also learned to perceive visible speech more accurately. Using a within-subject design, the research also convincingly showed that Baldi's intervention was responsible and not some other explanation such as time on task. Demonstrations of children in the tutoring situation are available ([Baldi Tutoring, 2022b](#)).

Predicting the future

While developing Baldi, a computer animated talking head, to provide accurate visible as well as audible spoken language, showing it was effective in communicating spoken language, and applying it to situations in language learning ([Massaro et al., 2005](#)), [Massaro \(1995\)](#) wrote a science fiction scenario on a future without written language. English had become the world's dominant language and societies could no longer afford to continue English literacy. Literacy learners had difficulty conquering the irregularity of English's spelling, with many inconsistencies between spoken words and their spelling. A popular example is that it would be acceptable to spell fish as ghoti (gh as /f/ in tough, o as /I/ in women, and ti as /s/ in nation). Critiques and revisions of English writing have been advocated by illustrious celebrities including George Bernard Shaw and Mark Twain ([English Spelling, 2022](#)).

As with most antecedents of change, there were other factors leading to literacy's demise. Having overpopulated the earth with diminished raw materials, it was convenient to use existing books as insulation for high density housing. Most importantly, however, was the technological development of computer-animated characters with realistic speech and language. Although Face Time, audible books, and podcasts were not yet on the horizon, the prediction was that communication was primarily spoken whether between live participants or with animated characters.

Traveling from this 30-year-old prediction of the future to today, written language might still be in jeopardy. English no longer has to be the dominant language, so English will no longer be necessary for successfully navigating the international world. Huge increases in computational power and wireless communication with small portable devices, efficient algorithms in speech synthesis, spoken language recognition, and language translation are currently available for many languages and many more around the corner (Google Read Along is currently available in 180 countries and supports 9 languages). Given this technology, real-time translation will allow access to any language, so one only requires their family's language learned from birth.

Spoken language might again gain prominence after just a few centuries of text dominance since the invention of the printing press. Being lost in a book or going to that special place with text may have a shorter lifespan than currently anticipated. As observed by Farhad Manjoo, "Yet there are just as many books that achieve a resonance via the spoken word that their text alone cannot fully deliver" ([Manjoo, 2021](#)). The narrator can add life to the writing and capture what [Jaynes \(1976\)](#) described as the voice capturing, or even kidnapping the listener's attention. As technology advances, the audio book will continue to be embellished creating sound effects, surround sound, and features yet to be imagined. For memoirs, your favorite celebrities, and stand-up comedians, the likely choice would be to listen (and perhaps watch if possible). In 2019, almost one in four Americans listened to an audiobook in the previous year, only slightly less than the number who read an electronic book ([Reading or Listening to Books, 2022](#)). In his dystopian novel of the future, not anticipating advances in speech technology. Neal Stephenson's *Diamond Age* (1995) found an essential role for live narrators of texts (called ractors) in contrast to the many technological advances in his multiverse.

For many of us, an illiterate society is a sobering thought. There must exist convincing reasons that written language has a role to play in the future. As one possible reason would be efficiency: we read at least twice the rate as we utter and hear spoken language. But many of us have already learned that various audio and video applications have options to speed up (or slow down) the play-back rate. Text-to-speech can be easily understood at much faster than normal, rates comparable to skilled silent reading. Podcasts (now in the millions) might be predicting the future. Yes, texting is pervasive but many texters use speech recognition to send these messages, and speech could be used to understand them. We can expect to have a technology translating adumbrated texts to coherent speech, ensuring spoken language its dominance. We think of text demanding much less memory storage but given advances in spoken language technology, the electronic footprint of text to speech can be reduced to be more similar to text.

Understanding tables, graphs and equations will demand some written form. As [Manjoo \(2021\)](#) observed, "I have also had trouble listening to dense, especially technical books." Written language won't disappear but for many, spoken communication will be preferred. We might fall back on print when the material is dense and complex, and simply to take a break from a voice that may be commanding more obedience than we are willing to fork over. But we predict the voice (even if it is synthetic) will carry the day. We know also that the real voice coming from the body will have a game-winning advantage, conveying emotion from multiple sensors including heart rate, sweat glands, and facial expressions that are lost in print. The voice will have a multimodal advantage, when it is in YouTube video guides of information to solve the most esoteric challenge. With one giant leap beyond virtual and augmented displays, an AI implant will connect us with what we need now because it is in constant communication with our own hardware. We are, however, humble in knowing how many predictions go awry.

Our private reading might be preempted by listening, but we won't want to sacrifice reading aloud to our children. This became particularly poignant when we observed the richer language our preschoolers experience in picture books relative to CDS and TV Media (Massaro, 2015; Williams et al., 2022). To preserve this tradition, the pictures in picture books can exist as they do now or they can be presented on screens. Since it is unlikely that we will memorize the many stories, the caregiver could listen along with the child or would have to somehow shadow the audio book while reading it to the child. But the caregivers could listen along with the child, embellishing the story in a variety of ways as they see appropriate. Thus, although listened to rather than read, picture books will continue to enrich our toddlers' brain with the richness of formal language.

A potential danger is that dominance of spoken language might lower the quality of our stories. That is, we know writing is mostly rewriting following Nabokov's prescription that reading is re-reading and this dictum might fail to impress the spoken language storyteller that scripting is re-scripting. Books will not have the formal language we value if the author doesn't have the opportunity to embellish and revise their drafts. With the expected advancement of controls over spoken language, the author could choose to stick with the spoken medium (especially if she is illiterate). The first generation experiencing the impending demise of written language will stick with writing even though the majority of "readers" will be listeners.

Sudowriters

After cautioning storytellers about AI's impending demise of written language, it might seem ironic to question whether AI could aid in our storytelling. To sample what AI might contribute, we input our initial first draft sentence in a just discovered AI writing application [Sudowrite \(2022\)](#).

The intersection of Artificial Intelligence (AI) and Literacy should certainly be a show stopper for the casual reader.

Sudowrite is powered by (and we mean powered by) a supercomputer with 285,000 processors to learn from the existing multi-verse of content using, of course, a multilayered neural network. Given a single topic sentence, it spits out some additional text corresponding to a variety of genres that include our senses and various metaphoric domains. Here's what we were able to select from Sudowrite's output delivered more or less instantaneously to our seedling first sentence.

One of the most advanced pieces of technology in existence, artificial intelligence has evolved to reach comprehension for both spoken and written language. It is now able to understand simple questions, like "what is the definition of the word 'cat'?", and can tell you what it is, or anything else you'd ask of it.

This technology is still in its infancy, but artificial intelligence is quickly improving. More and more, it's becoming more engrained into our daily lives.

The AI's purpose is to read things, but it is not a physical piece of machinery. Despite being made of machinery, it has no physical form.

The concept of Artificial Intelligence is that, if you have enough information, you can teach it to learn. Eventually, you will have an intelligence that is more sophisticated than humans. At the moment, artificial intelligence is still in its infancy, but it's quickly becoming more common.

Although the program expects multiple iterations of feeding each output back to Sudowrite, we stopped here. Our gracefully extended deadline became hard and it was somewhat eerie to include Sudowrite as a co-author.

Challenges to AI's role in literacy

Limitations in embodying us

Should we expect a social robot to embody all of the unique characteristics of live tutors anytime soon? Some are optimistic that this goal is within reach (Kurzweil, 2021), whereas others resound a strong negation (Moore, 2016). Linguistic exchanges create a milieu of anxiety shared by both parties, and we are uncertain how to embody this characteristic into robots (Rowe, 2020). Reducing this anxiety during exchanges in conversations and even narratives is married to rewards. A non-linguistic example involves the little hit of dopamine that the infant receives when she wiggles her foot to jiggle the overhead happy birds mobile. Similarly, an interlocutor thrives on resolving uncertainty in the conversational exchange (Ramscar and Port, 2015, 2016).

However, we can expect that an artificial tutor who only approximates embodiment of a human tutor to be effective in guiding the child or student to learn both language and new knowledge. Can an AI companion ever achieve the linguistic and cognitive

capacities that humans find so natural? Christiansen and Chater (2022) make a strong case that humans are unique in pursuing an iterative language game in which utterances take on meaning whereas AI systems are nothing more than mindless fact-collecting automata that don't have an inkling of what's up. Moore (2016) argues that improvements in speech (and animation/robotics) technologies can outpace a social robot's cognitive and linguistic capacities. This "uncanny valley" would dismantle productive interactions because participants would overestimate the social robot's capability. Humans are fairly robust interlocutors, however, and we might be well-equipped to converse productively with unequal partners. Real life successful examples include interlocutors with language and cultural differences, as well as children, chimpanzees, and dogs. Children are also successful in conversations with single word utterances (Herr-Israel and McCune, 2011; Massaro, 2020). The infamous chimpanzee Nim convinced his caregivers that they were productively conversing even while they were implicitly cuing him for appropriate replies (Massaro, 2020).

Although we are both predicting the future, one of us might actually be correct. As evidence for his pessimism, Moore in 2016 claimed that intelligent assistants, such as Siri, were not very popular but the demographics have changed significantly—another instance in which it is difficult to predict the future. As we write this, Siri receives about a billion questions a week; in 2016 Alexa had only 130 interactive skills but now the number is well above 100,000; Google's Home assistant is also popular and of course many people google search by voice and often receive voice replies or links to spoken and video answers. It's expected that over half of the US households have or will soon have smart devices that allow human-machine dialogs via voice. We might also expect that these devices will become more embodied, with virtual agents or even a physical robotic presence such as the intriguing applications of Furhat (2021).

We can expect that virtual tutors rather than physically-present robots would be prevalent because sensory motor interactions are much easier to program virtually than in actual physical substance. Successful interactions between humans and machines have certainly increased dramatically in the last few years, with no sign of diminishing. We find it more than natural to address Siri or Alexa with a plethora of questions, for example. Also, there is active research and applications extending human-robot interactions to be embodied (Furhat, 2021). As social robots incorporate multisensory perception and action, we can expect much more rewarding dialogs and outcomes.

Retrospective

Artificial Intelligence and Literacy is expected to continue to provide a synergistic partnership in both reading instruction, reading, and access to written sources spanning the range of human knowledge. Improved machine learning algorithms, technical advances, and computational speed and efficiency are responsible for these impressive developments. Surprisingly for one of the authors was the realization that speech science was the primary AI contributing to this progress. Cognitive Science tended to partition research in separate speech and reading bins, and most research adhered to this separation. Much earlier, St. Augustine might have noticed the tight relationship between spoken and written language when he was surprised observing his mentor, St. Ambrose, reading silently without moving his lips (Manguel, 1996). Bochner and Bochner (2009) make a strong case against print alone accessing, acquiring, and utilizing linguistic input, based on deaf individuals learning to read (see also Lichtenstein, 1998; Yang, 2008). Some phonological source is needed which is most easily obtained via hearing. Other sources such as visible speech and touch from the face, and cued speech with the hands can help but appear to fall short of audible speech.

One reason that reading may have an advantage over listening is that since writing elicits that little voice in your head, it has two sources of information relative to just one. Utilizing multiple sources of sensory and contextual information in language processing has been a hallmark of recent theories (e.g., Massaro, 1998). Having speech, even if it's internal, as well as the written word on the page, anchors the language processor in more elaborate constraints than either modality alone (There are also rare instances of synesthesia in which speech can elicit written text, as in the mnemonist studied by Luria (1968, 1987). Writing also reduces language change which might be somewhat problematical in a speech limited metaverse.

For reading instruction, AI can automate reading experiences that can combine instruction and assessment in a seamless manner. AI can also optimize what the child is reading in terms of reading difficulty and important content to be learned. Assessment can also be made more ecologically valid and engaging by creative forms of integrated testing during reading. AI also can easily embody a reading tutor to help maintain the student's interest in reading and learning. With individually guided tutoring, reading to learn curricula can exploit the knowledge that has been organized and consolidated in various domains.

References

- AI Umpire, 2022. <https://sportsnaut.com/robot-umps-to-call-balls-strikes-in-triple-a-in-2022/>.
- Average Reading Speed, 2022. <https://scholarwithin.com/average-reading-speed>.
- Baldi Tutoring, 2022a. <https://cpb-us-e1.wpmucdn.com/sites.ucsc.edu/dist/0/158/files/2022/02/raylay.mov>.
- Baldi Tutoring, 2022b. <https://www.youtube.com/watch?v=BeWxjMhAm-s>; https://www.youtube.com/watch?v=mDfKdRh5h_g.
- Balogh, J., Bernstein, J., Cheng, J., Van Moere, A., Townshend, B., Suzuki, M., 2012. Validation of automated scoring of oral reading. *Educ. Psychol. Meas.* 72 (3), 435–452.
- Belpaeme, T., et al., 2018. Guidelines for designing social robots as second language tutors. *Int. J. Soc. Robot.* 10 (3), 325–341.
- Bernstein, J., Cheng, J., Balogh, J., Rosenfeld, E., 2017. Studies of a self-administered oral reading assessment. In: *Proceedings of the 7th ISCA Workshop on Speech and Language Technology in Education*, Stockholm, Sweden, pp. 180–184.
- Bochner, J.H., Bochner, A.M., 2009. A limitation on reading as a source of linguistic input: evidence from deaf learners. *Read. Foreign Lang.* 21 (2), 143–158. ISSN 1539-0578.

- Boote, B., Agarwal, M., Mostow, J., 2021. Early Prediction of Children's Task Completion in a Tablet Tutor Using Visual Features. The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21). Pp. 15761-15762.
- Bryant, W., 2017. Developing a strategy for using technology-enhanced items in large-scale standardized tests. *Pract. Assess. Res. Eval.* 22, 1. <https://doi.org/10.7275/70yb-dj34>. Available at: <https://scholarworks.umass.edu/pare/vol22/iss1/1>.
- Brysbaert, M., New, B., 2009. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* 41, 977-990.
- Christiansen, M.H., Chater, N., 2022. *The Language Game: How Improvisation Created Language and Changed the World*. Basic Books, New York.
- Clark, A., Chalmers, D., 1998. The extended mind. *Analysis* 58 (1), 7-19.
- COCA, 1990-2012. Corpus of Contemporary American English (COCA). (1990-2012). Words Added Between 1990 and 2012. Retrieved from: <http://corpus.byu.edu/coca/>.
- Deng, E., Mutlu, B., Mataric, M.J., 2019. Embodiment in socially interactive robots. *FNT Robot.* 7 (4), 251-356. <https://doi.org/10.1561/23000000056>.
- English Spelling, 2022. <https://deepblue.lib.umich.edu/bitstream/handle/2027.42/62008/words.html>.
- Evans, M.A., Saint-Aubin, J., 2013. Vocabulary acquisition without adult explanations in repeated shared book reading: an eye movement study. *J. Educ. Psychol.* 105 (3), 596-608. <https://doi.org/10.1037/a0032465>.
- Furhat, 2021. <https://furhatrobotics.com/blog/5-ways-social-robots-are-innovating-education/>.
- Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6-10. <https://doi.org/10.1177/074193258600700104>.
- Grice, H.P., 1975. Logic and conversation. In: Martinich, A.P. (Ed.), *Philosophy of Language*. Oxford University Press, New York, NY, pp. 165-175.
- Gurdon, M.C., 2019. *The Enchanted Hour: The Miraculous Power of Reading Aloud in the Age of Distraction*. Harper, New York.
- Hart, B., Risley, T., 1995. Meaningful Differences in the Everyday Experience of Young American Children. Paul H. Brookes, Baltimore, CA.
- Herr-Israel, E., McCune, L., 2011. Successive single-word utterances and use of conversational input: a pre-syntactic route to multiword utterances. *J. Child Lang.* 38 (1), 166-180. <https://doi.org/10.1017/S0305000909990237>.
- Hiebert, E.H., 2005. In pursuit of an effective, efficient vocabulary curriculum for the elementary grades. In: Hiebert, E.H., Kamil, M. (Eds.), *The Teaching and Learning of Vocabulary: Bringing Scientific Research to Practice*. LEA, Mahwah, NJ, USA, pp. 243-263.
- History of TTS, 2022. <https://www.kurzweiledo.com/about-kurzweil/history-of-text-to-speech.html>.
- iSPYPhonics Fun, 2021. <https://www.thriftbooks.com/series/i-spy-a-book-of-picture-riddles/37646/>.
- Jaynes, J., 1976. *The Origin of Consciousness in the Breakdown of the Bicameral Mind* (PDF). Houghton Mifflin, New York. ISBN 0-618-05707-2.
- Jeopardy, 2022. <https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html>.
- Kasap, Z., Magnenat-Thalmann, N., 2012. Building long-term relationships with virtual and robotic characters: the role of remembering. *Vis. Comput.* 28 (1), 87-97. <https://doi.org/10.1007/s00371-011-0630-7>.
- Klatt, D., 1987. Review of text-to-speech conversion for English. *J. Acoust. Soc. Am.* 82 (3), 737-933. <https://doi.org/10.1121/1.395275>. Bibcode:1987ASAJ...82..737K. PMID 2958525.
- Kurzweil, R., 2021. <https://www.inverse.com/article/30913-8-staggering-predictions-ray-kurzweil>.
- Li, J., 2015. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Comput. Stud.* 77, 23-37. <https://doi.org/10.1016/j.ijhcs.2015.01.001>.
- Lichtenstein, E.H., 1998. The relationships between reading processes and English skills of deaf college students. *J. Deaf Stud. Deaf Educ.* 3, 80-134.
- Luria, A.R., 1968. *The Mind of a Mnemonist*. Washington Square Press, New York.
- Luria, A.R., 1987. *The Mind of a Mnemonist. A Little Book About a Vast Memory*, With a New Foreword by Jerome S. Bruner. Harvard University Press, Cambridge, MA.
- Manguel, A., 1996. *A History of Reading*. <https://www.historyofinformation.com/detail.php?entryid=4341>.
- Manjoo, F., 2021. <https://www.nytimes.com/2021/10/06/opinion/audiobooks-better-than-reading.html>.
- Massaro, D.W., 1995. From speech-is-special to talking heads: the past to the present. In: Solso, R.L., Massaro, D.W. (Eds.), *The Science of the Mind: 2001 and Beyond*. Oxford University Press, New York, pp. 203-220.
- Massaro, D.W., 1998. *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. MIT Press, Cambridge, Massachusetts.
- Massaro, D.W., 2012. Acquiring literacy naturally: behavioral science and technology could empower preschool children to learn to read naturally without instruction. *Am. Sci.* 100, 324-333.
- Massaro, D.W., 2015. Two different communication genres and implications for vocabulary development and learning to read. *J. Lit. Res.* 47 (4), 505-527. <https://doi.org/10.1177/1086296X15627528>.
- Massaro, D.W., 2017. Reading aloud to children: benefits and implications for acquiring literacy before schooling begins. *Am. J. Psychol.* 130 (1), 63-72. <https://doi.org/10.5406/amerjpsyc.130.1.0063>.
- Massaro, D.W., 2020. My words fly up, my thoughts remain below. Words without thoughts never to heaven go. Review of why chimpanzees can't learn language and only humans can (Leonard Hastings Schoff lectures) by Herbert S. Terrace. *Am. J. Psychol.* 133, 378-392.
- Massaro, D.W., Hestand, J., 1983. Developmental relations between reading ability and knowledge of orthographic structure. *Contemp. Educ. Psychol.* 8, 174-180.
- Massaro, D.W., Ouni, S., Cohen, M.M., Clark, R., 2005. A multilingual embodied conversational agent. In: Sprague, R.H. (Ed.), *Proceedings of 38th Annual Hawaii International Conference on System Sciences (HICSS'05)* (CD-ROM, 10 Pages). IEEE Computer Society Press, Los Alamitos, CA.
- Media, 2022. <https://www.common-sense-media.org/>.
- Merlin, D., 1991. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press, Cambridge, MA.
- Michaelis, J.E., Mutlu, B., 2018. Reading socially: transforming the in-home reading experience with a learning-companion robot. *Sci. Robot.* 3 (21). Springer robotics.sciencemag.org).
- Moby.ReadTM, 2022. www.analyticmeasures.com/moby-read.
- Montag, J.L., Jones, M.N., Smith, L.B., 2015. The words children hear: picture books and the statistics for language learning. *Psychol. Sci.* 26 (9), 1489-1496. <https://doi.org/10.1177/095679761559436>.
- Moore, R.K., 2016. Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction. *International Workshop on Spoken Dialogue Systems (IWSDS)* 13-16 January 2016, Riekkonlinna, Finland.
- Mostow, J., 2013. Lessons from project LISTEN: what have we learned from a reading tutor that listens?. In: *Artificial Intelligence in Education*, pp. 557-561.
- Perfetti, C., Hogaboam, T., 1975. The relationship between single word decoding and reading comprehension skill. *J. Educ. Psychol.* 67, 461-469.
- Phonics Exercises, 2022. <https://www.spellingcity.com/phonics.html>.
- Piasta, S.B., Justice, L.M., McGinty, A.S., Kaderavek, J.N., 2012. Increasing young children's contract with print during shared reading: longitudinal effects on literacy achievement. *Child Dev.* 83, 810-820. <https://doi.org/10.1111/j.1467-8624.2012.01754.x>.
- Pinker, S., 2007. *The Language Instinct: How the Mind Creates Language*. Harper, New York.
- Project LISTEN's Reading Tutor, 2022. <https://www.ri.cmu.edu/project-listens-reading-tutor/>.
- Ramscar, M., Port, R., 2015. Categorization (without categories). In: Dawbroska, E., Divjak, D. (Eds.), *Handbook of Cognitive Linguistics*. DeGruyter Mouton, Berlin, Germany, pp. 75-99.
- Ramscar, M., Port, R., 2016. How spoken languages work in the absence of an inventory of discrete units. *Lang. Sci.* 53, 58-74.
- Read With Me! Demo, 2022. <http://psyntificmind.com/wp-content/uploads/2012/07/Keagen2.mov>.
- Read With Me! 2022. <https://apps.apple.com/us/app/read-with-me/id494909666>.

- Readability, 2021. Automatic Readability Checker, a Free Readability Formula Consensus Calculator. Readability Formulas: Automatic Readability Checker. <https://readabilityformulas.com/free-readability-formula-tests.php>.
- Reading or Listening to Books, 2022. <https://www.bbc.com/culture/article/2020104-audiobooks-the-rise-and-rise-of-the-books-you-dont-read>.
- Rowe, B., 2020. The marvelous gadget. Cognitive gadgets: the cultural evolution of thinking by Cecilia Heyes. *Am. J. Psychol.* 133 (3), 341–378 (Fall 2020).
- Sauppé, A., Mutlu, B., 2014. Effective task training strategies for instructional robots. In: *Proceedings of Robotics: Science and Systems (RSS 2014)*.
- Sauppé, A., Mutlu, B., 2015. Effective task training strategies for human and robot instructors. *Aut. Robots* 1–17. <https://doi.org/10.1007/s10514-015-9461-0>.
- Shanahan, T., 2006. *The National Reading Panel Report: Practical Advice for Teachers*. Learning Point, Naperville, Illinois.
- Stephenson, N., 1995. *Diamond Age*.
- Stephenson, N., 2021. Michael Shermer Interview.
- SUBTLEX, 2021. <https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus>.
- Sudowrite, 2022. <https://gpt3demo.com/apps/sudowrite>.
- TEIs, 2021. <https://edulastic.com/blog/technology-enhanced-items-showcase/>.
- Understand My World, 2022. <https://apps.apple.com/us/app/understand-my-world/id1298250806>.
- van Kleeck, A., Stahl, S.A., Bauer, E.B., 2003. *On reading books to children: parents and teachers*. Lawrence Erlbaum Associates, Mahwah, NJ.
- van Santen, J.P.H., Sproat, R.W., Olive, J.P., Hirschberg, J., 1997. *Progress in Speech Synthesis*. Springer. ISBN 978-0-387-94701-3.
- Venezky, R.L., Massaro, D.W., 1979. The role of orthographic regularity in word recognition. In: Resnick, L.B. (Ed.), *Theory and Practice of Early Reading*, vol. 1. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 85–107.
- Venezky, R.L., Massaro, D.W., 1987. Orthographic structure and spelling-sound regularity in reading English words. In: Allport, A., MacKay, D., Prinz, W., Scheerer, E. (Eds.), *Language Perception and Production: Relationships Between Speaking, Reading and Writing*. Academic Press, pp. 159–179.
- Vygotsky, L., 2022. https://en.wikipedia.org/wiki/Zone_of_proximal_development.
- Williams, A., Tudjinovic, E., Baylon, T., Alarcon, J., Bose, R., Massaro, D.W., 2022. Mining Big Data to Analyze Influences on Children's Language Acquisition.
- Yang, J.H., 2008. Sign language and oral/written language in deaf education in China. In: Plaza-Pust, C., Morales-Lopes, E. (Eds.), *Sign Bilingualism: Language Development, Interaction and Maintenance in Sign Language Contact Situations*. John Benjamins, Amsterdam, pp. 297–331.
- Yueh, H.P., Lin, W., Wang, S.C., 2020. Reading with robot and human companions in library literacy activities: a comparison study. *Br. J. Educ. Technol.* 51, 1884–1900.

Evaluating, supporting and enhancing Internet search activity: a review of research, with implications for pedagogy

Colin Harrison, Emeritus Professor of Literacy Studies in Education, University of Nottingham, Nottingham, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	543
Methodology	544
Results: three broad research themes	544
Research into Internet users' interaction processes	544
Search completion research	546
Research into students' criticality on the internet	547
Pedagogical implications	548
Conclusions	548
References	549

Introduction

Defining internet literacy is not a simple matter. In their review of research in an area related to Internet literacy, namely media literacy, [Potter and Thai \(2019\)](#) noted that across the 88 studies they reviewed, definitions of the term “media literacy” were idiosyncratic and difficult to categorize. Definitions were not offered in 25% of studies, and in over half of the studies, the authors presented their own unique definition. For Potter and Thai, this variability raised issues of content validity across the field of media literacy studies, and made offering a meta-analysis very challenging. The field of Internet literacy is equally difficult to categorize, since it has been recognized from the outset that Internet literacy is deictic: that is, by its very nature it is something that changes and is redefined according to context ([Leu, 2000](#); [Coiro et al., 2008](#)). The definition of Internet literacy has thus been continuously shifting, as new technologies change the nature of social practices, as the tools we use for reading and writing change, as the very nature of what we call “information” changes, and as the skill-set needed to recruit these technologies for our own benefit mutates and broadens.

The population for whom “internet literacy” might be necessary has also broadened significantly. In the 1980s, ARPANET and NSFNET, the forerunners of what came to be called “the Internet,” connected only universities and university researchers. In the 1990s, and following the invention of the WorldWideWeb ([Berners-Lee and Cailliau, 1990](#)), access to the Internet began to expand exponentially, first to all university students and their professors, then to all library users, then to school students and their teachers, and finally since the early 2000s, to all parents and their children. In many nations, Internet adoption now spans the full age-range. In 2018, Ofcom, the UK government’s telecoms regulator, reported that 69% of 3–4-year-olds went online for nearly 9 h per week ([Ofcom, 2018](#)), while in the US the Pew Research Center reported that 67% of those over the age of 65 were Internet users ([Anderson and Perrin, 2017](#)).

The ability to search and use the Internet wisely is therefore a skill that is essential not only for students, but for every citizen, and we know already that this ability is spread unevenly across populations. As Hargatti has argued, “there is a spectrum of differences associated with Internet usage once people have crossed the access divide and have become users” ([Hargittai et al., 2019](#)). We know for example that those from more privileged backgrounds tend to have higher Internet skills ([Ono and Zavodny, 2007](#)), and also that those who get their news mainly from social media are both less knowledgeable and more likely to encounter false or unproven claims ([Pew Research Center, 2020](#)). Tim Berners-Lee set up the WorldWideWeb with the intention of creating a platform for the free and open worldwide exchange of information ([Berners-Lee and Cailliau, 1990](#)). Today, many users of the Internet do not see themselves as searching for “information”: their goal may be to watch a movie, to buy a product, or to contact friends using social media. However, many mouse “clicks” not only send a request for information, they also trigger targeted and unrequested advertising, or information that is untrustworthy. Given this scenario, the issue of how we evaluate, support and enhance the ability of the citizens of tomorrow to use the Internet wisely is one of critical importance, and it is not surprising that it has become an important area of research.

As Becker found in her longitudinal study of undergraduates’ Internet use ([Becker, 2003](#)), despite a plethora of skill development programs across educational levels, even college students continued to give uncritical primacy to Google results, and if anything, were likely to place more trust in unedited online search results than in print resources that had been through an editorial process. If we share the view of Becker and other researchers who assert that Internet literacy is important not only in education but across the lifespan ([Leu, 2000](#); [DiMaggio and Bonikowski, 2008](#); [Harrison, 2018a](#)), this raises two important questions: how should we analyze students’ and others’ Internet use, and on which skills should pedagogy be focused, as we seek to prepare students to

face the future in this “post-truth” world (Keyes, 2004)? The present review seeks to address these questions, and will attempt an overview of research in the field, looking particularly at the impact of technological changes over the period from 1990, when the World Wide Web was invented (Berners-Lee and Cailliau, 1990), to the present day.

Methodology

In order to review the field systematically, a search with Google Scholar using the term <Internet search behaviour> (using both English and US spellings for the word “behaviour/behavior”) for papers published since 1990 yielded over 800,000 results. A phrase-constrained search for <“Internet search ‘behaviour/behavior’”> narrowed this down to under two hundred results, with many of the research papers focusing not only on library skills but also on two other areas: adults searching for health information, and marketing and adult consumer behavior. The areas of adult health and marketing continue to be the focus of much Internet research, but this review will focus more on research into student activity and pedagogy. From the year 2000 onwards, research studies dealing with Internet surveillance and authoritarianism began to be published, and these were clearly relevant to the present review. In order to focus more closely on children’s and students’ Internet search skills, the search phrase <“Internet searching skills”> was used. This yielded six hundred results (after false positives were excluded), ranging from the years 1995–2020, five of which were themselves research reviews or meta-analyses (Hsieh-Yee, 2001; Hargittai, 2002; Coiro and Dobler, 2007; Litt, 2013; Leeder and Shah, 2016). These research reviews offered citations leading to a further four hundred papers on Internet searching strategies and skills, and together these groups of papers formed the basis of the research reported here.

Results: three broad research themes

The three broad themes emerged from the research review: these were interaction processes, search completion outcomes, and dialogic criticality. These three themes represent roughly chronological areas of evolving research focus, which in turn reflect developments in both technology and pedagogy over a thirty-year period. Understandably enough, the early studies focused on interaction processes, with methodologies closely resembling those of papers on library searching that had preceded them. These early papers generally categorized students’ search strategies, but also reported users’ beliefs and opinions about what was in the 1990s a novel (and often frustrating) extension to library provision. Between 1990 and 2000, Internet content massively expanded, access to the Internet became easier, upload and download speed increased, and as platforms became more stable, researchers’ attention shifted toward our second theme: the outcomes of student search. Many papers reported on the efficiency of student search, with search success and search time emerging as important variables. Beginning in the early 2000s, research began to focus on our third theme, user criticality, as increased use of the Internet by younger students, combined with teachers’ awareness of the unreliability of many Internet sources, prompted more attention to factors such as how students made judgments about the material they encountered, how they collaborated in enquiry, and how they synthesized information from multiple sites.

The three broad research themes are represented diagrammatically in Fig. 1, below, which also shows details of sub-themes that will be discussed in the sections that follow.

Research into Internet users’ interaction processes

The first papers on students’ Internet searching were almost all published in library science or computing journals, and focused on interaction processes. One of the earliest was that of Catledge and Pitkow (1995), which captured client-side log files of students using the Mosaic web browser, and described search patterns such as general browsing versus targeted searching, and went on to note whether search results went unrecorded, or were saved or tagged. There was no author comment on the presumed skill or efficiency of the students who were using the system, though the authors expressed excitement about the possibility of future server-side software that might alter page design the fly, based on user history or other characteristics. Hsieh-Yee’s review (2001) of 66 papers published between 1995 and 2000 on Web search behavior confirmed that most of the early studies focused on simply describing and categorizing aspects of search activity. Attempting a meta-analysis of these studies is difficult because of the great variety of search engines. More than two thousand different search engines were in use during this period, and there was also the problem of relatively poor coverage of Internet content. In 1997 it was estimated at no more than 16% of Internet sites were covered by search engines (Lawrence and Giles, 1999). Hsieh-Yee made the valuable and prescient point that a search engine’s user model influenced both the nature and outcome of an Internet search.

This type of descriptive research was extended by Hölscher and Strube (2000) to include both interviews and prior knowledge data, contrasting “expert” and “newbie” search activity, and focusing on search terms (those of experts were on average twice as long as those of newbies). This more descriptive research, based on interviews, focus groups and self-report, carried on into the 2000s, with the work of Joo et al. (2000) and the studies of Tsai and Tsai (2003) on the relationship between self-efficacy and Web search success making a useful contribution. Litt’s (2013) authoritative and comprehensive review of research gave summaries of a further six studies that used self-report measures as part of an assessment of online activity.

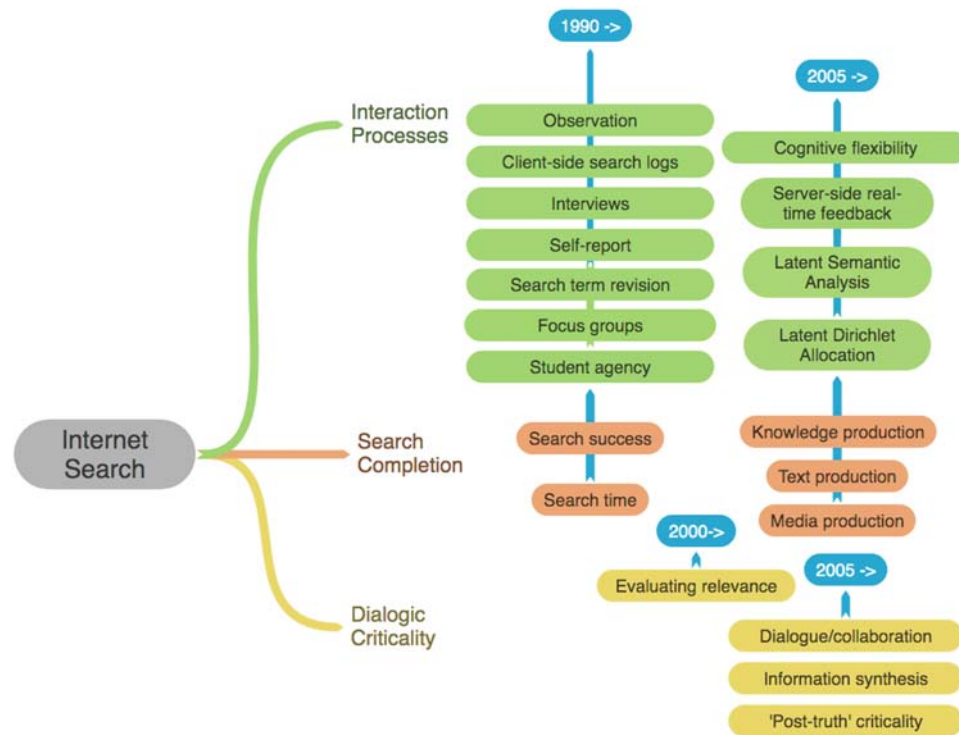


Fig. 1 Themes of research into Internet activity that emerged from the review of research.

The themes of self-regulation and motivation have continued to receive attention, with Forzani, Leu, Guthrie and their co-authors arguing (2020) that although there may be many similarities between online and offline reading, affective factors such as intrinsic motivation, engagement, perceived self-efficacy, and beliefs related to self-improvement may be so different in online reading as to require a different set of questionnaire approaches. This study, which drew upon the responses of nearly 2000 participants, found that students' perceived Internet self-efficacy most strongly predicted success on an online comprehension task. The authors explained this by suggesting that online reading perhaps requires a greater degree of self-monitoring (and the ability to ignore distractions such as advertisements and extraneous links), compared with offline reading.

Developments in cognitive psychology have also had an impact on research into interaction processes, with DeSchraver's study (2012) bringing together a number of these perspectives. DeSchraver is one of many researchers who used Rand Spiro's Cognitive Flexibility Theory (2012) to capture some of the dynamic shifts of attention and focus in Internet searching, and he also foregrounded Walter Kintsch's construction-integration model (1998) to inform his discussion of how students bring together prior knowledge and emerging components of textual understanding to synthesize and create new knowledge in the ill-structured domain of learning online. These theoretical perspectives from Spiro and Kintsch offer a much deeper and more complex account of the adaptive searching and synthesis that take place on the Internet than had been possible in the 1990s.

Developments in computer power and software have also had a significant impact on more recent work on interaction processes. Research using server-side data-collection and feedback has brought new tools to the analysis of web searching, making use of client-side output such as what search terms were used and how much verbatim text was stored, but then going on to estimate how well source texts were summarized, and how well complex concepts were understood. One interesting tool that has been used in the analysis of user-produced output is Latent Semantic Analysis (LSA). LSA is a natural language processing technique that uses statistics related to factor analysis to represent a word, a sentence or a whole text as a vector, a single point in n -dimensional semantic space (Landauer et al., 1998). The power of the technique comes from the way in which it can enable comparisons between the vectors of texts that have similar theme or topic, for example a source text and a summary of that text, even if the verbatim semantic overlaps are few. For example, LSA knows that the following two sentences are closely related thematically, even though they have only the word *the* in common: 1. *Spy aircraft fly up to the stratosphere.* 2. *A pilot can take a surveillance plane high into the sky.* LSA takes account of meaning, and not simply verbatim matching between the words that make up texts; it is for this reason that it is described as a "latent" analysis.

In 2005, Omar and her colleagues (2005) applied LSA to analyze both search terms and student summaries of web content, and also used LSA to estimate text coherence as well as thematic similarity. Sung and colleagues (2016) used LSA as part of an online system that aimed to help sixth grade students in China improve their ability to write summaries. Their system used LSA to compare draft summaries of science texts with those of an expert, and the system then gave feedback in real time on a sentence-by-sentence

basis, with high-scoring summary sections (i.e., those with LSA scores close to those of the original) shown in black, and low-scoring sections shown in gray text. The highlighting changed in real time as the draft summaries were improved. The researchers found that providing the online LSA-driven feedback was associated with improved summary writing.

Another approach to generating server-side real-time feedback came from the work of Dascalu and colleagues (Dascalu et al., 2015; Gutu et al., 2016), who developed software called *ReaderBench* that uses natural language processing tools to generate automated estimates of student understanding. *ReaderBench* also moves into the domain of Computer Supported Collaborative Learning using the relatively new tool of Latent Dirichlet Allocation (Blei et al., 2003) for analyzing students' online interactions as they worked together to understand new concepts. LDA attempts to isolate the thematic content of dialogs, differentiating between scholarly and social chat themes, and estimating (and offering the students formative feedback on) their understanding. More recent studies using *ReaderBench* have looked more closely at student participation, with an in-depth longitudinal analysis of online course participation based on students' discourse and measures of text cohesion (Sirbu et al., 2019).

Search completion research

Clearly, search completion is always the implied goal of Internet searching, but it is worth examining research that has focused particularly on this aspect of Internet use. As we have already mentioned, Internet searching in the 1990s was hampered by the sheer variety of Web search tools, and by the inability of search engines to cope with spelling inaccuracies on the one hand, and search results that had usually been written for adult specialist audiences on the other. Borgman's et al. (1995) study represented a detailed and sympathetic account of the problems in Internet search completion faced by children, who understandably reported that the more complex the search, the more difficulty they had in selecting appropriate search terms to match their research goals. Similarly, many searches led to a screen with alphabetized keyword display systems, and this was a problem for many elementary school students, as were interfaces that required students to include Boolean operators (for example AND, OR, and NOT). In the 1990s, many searches led not to a specific page, but to a hierarchically organized data base, which many students found difficult to navigate. Fidel and colleagues reported (1999) that in general, high school students were beginning to become competent Internet searchers, and appreciated encountering pictures as well as text results, but that their searches often led to data-centered results rather than user-centered results.

An interesting study by Bilal (2002) reported on children's success in using the Yahhooligans! Web search engine on two types of task: teacher-assigned and self-chosen. Overall, students were more successful on the self-chosen tasks. One of the reasons for failure in self-chosen tasks, however, was problems with spelling: even in 2002, search engines were unable to deal well with misspellings, even close approximations, such as "veterinarians" for "veterinarians", or "phsychology," for "psychology". Perhaps understandably, unsuccessful searches involved more browsing than was the case for successful searches. An interesting Dutch study by Van Deursen and Van Dijk (2009) looked at the digital divide between groups of adult Internet users. The researchers made the important point that having the technology to access information is not the same thing as having the ability to locate information, and they made a detailed study of the Internet skills of a stratified random sample of 109 adult participants, finding that on average, adults were able to complete only two-thirds of closed tasks looking for information on a government web site, and fared even worse on more open challenges, such as finding out what tax or pension benefits they might be entitled to. Perhaps surprisingly, the amount of time participants spent online each week was not a significant predictor of task success; only educational level correlated highly with tasks completed successfully. Finally, they made the important point that studies based on school or university students may paint an over-optimistic picture of likely Internet search success across the whole population, and that even with universal Internet access, for people with lower information and strategic skills it will become harder and harder to use the Internet successfully to gain access to key information, resources and support for living.

As Internet use expanded, studies that looked closely at search completion also began to focus on the use made of search outcomes. Early Internet tools led users to nothing more than databases or library catalogs which were often essentially graphical interface versions of the Dewey catalog system. Gradually attention turned from whether students located the information they needed to whether they simply copied verbatim, or turned to other forms of text production. Eshet's (2004) study identified a range of new information skills demanded by the Internet, including "photo-visual literacy" (for example, reading instructions from graphical displays); using digital reproduction to create new, meaningful materials from existing ones; constructing knowledge from nonlinear, hypertext navigation; evaluating the quality and validity of information; and having a mature and realistic understanding of the "rules" that prevail in the Cyberspace (which included, for example, vigilance in relation to spam and virus detection).

Eshet's study of new aspects of digital production as a result of Internet searching has many parallels with the work of Don Leu and his collaborators at the University of Connecticut. Leu's work over a twenty-year period (Leu, 2000; Forzani et al., 2020) began with the assertion that learning on the Internet was "deictic," that is, it was fluid, dynamic, mutable, and defined according to social as well as pedagogical context. He made the point that generalizing from research in new technologies is by definition a risky undertaking, for the simple reason that as technologies evolve, it becomes increasingly difficult to make the assumption that terms such as "Internet search" actually refer to the same activity as that researched some years before. Coiro (Coiro and Dobler, 2007; Coiro, 2011, 2020) has carried Leu's approach forward with case studies and large-scale investigations into the ways in which reading on the Internet differs from reading traditional printed text, putting particular emphasis on the more complex demands that the Internet makes of readers in three areas (a) the application of prior knowledge, (b) inferential reading strategies, and (c) self-regulated reading processes. Both Leu and Coiro stress that individual agency is a factor that needs close attention: motivation,

cognitive preferences, student-centered learning and equity issues are all important factors in Internet learning. Finally, and this point builds a useful bridge to our third sub-section: a deictic perspective implies the need for new forms of critical thinking and reasoning, and also a pedagogical model that has the flexibility to adapt to new ways of teaching, and new forms of knowledge.

Research into students' criticality on the internet

As Leu indicated in his account of the deictic nature of Internet learning, the traditional forces that have exercised control over the accuracy of information distributed through print publication do not exist in hyperspace. As this fact became clearer, researchers turned their attention to the extent to which Internet users were adopting a critical stance toward the results of their search. Hirsh's (1999) paper was one of the first detailed case-study investigations of children's decisions about the relevance criteria used during Internet search. Her findings included student frustration with very slow Internet response times (this was 1999), but Hirsh also identified a number of useful relevance criteria, which included: authoritativeness, accessibility, interest level, readability, novelty (i.e., results different from those of other searches), peer interest, and currency/topicality. However, the issue of authority had only been foregrounded by two out of the ten participants.

Edwards and Bruce (2002) put forward an action-research model for Internet searching that drew upon Schön's (1987) reflective practice principles, and foregrounded the importance of criticality in evaluating a number of factors: the likely accuracy of sources, the author's intentionality, and whether the information was up-to-date. Among the papers that have already been cited, those of Eshet (2004) and Van Deursen and van Dijk (2009) included variables related to evaluating credibility. Leeder and Shah (2016) explored this area in depth, with treatment conditions that included evaluating sources in relation to estimates of the author's expertise or trustworthiness, and they concluded that embedding structured support for critical evaluation in a task had a demonstrably positive effect on student search and behavior outcomes.

In their extensive case studies of online reading comprehension strategies, Coiro and Dobler (2007) reported on the skilled use of monitoring and evaluating search strategies used by some sixth-graders (11-year-olds), often with rapid cycles of self-regulated high-level cognitive processing, shown as they navigated across multiple sites, using different search engines.

Similarly, Zhang and Duke (2011) reported on the success of an intervention with 242 students that used a framework called WWWDOT to encourage a systematic and critical approach to Internet searching. The framework encourages students to think about at least six dimensions when evaluating a website: (1) Who wrote this and what credentials do they have? (2) Why was it written? (3) When was it written and updated? (4) Does this help meet my needs? (5) Organization of website; (6) To do list for the future. This approach prompted the students to ask questions about the author's credentials and purpose, and the relevance of the site for the task in hand. The final character in their framework, the "T", encouraged the students to make a "to do" list, for example planning ahead to widen the lens of their research, or consult an expert. However, while their students did become more aware of the need to evaluate, they did not immediately make better judgments than the students in the control group. One possible reason for Zhang's slightly disappointing results is that the outcome measure was based on fourth and fifth grade individuals carrying out a task in isolation.

A recognition of the fact that working alone to solve an Internet enquiry problem may be too challenging and too lacking in feedback for many students has led to a number of approaches that in different ways seek to scaffold deeper learning. Such approaches have foregrounded the value of collaborative rather than individual work, and have broadened the pedagogic lens from simple search tasks to more comprehensive enquiry learning, based on more worthwhile and engaging tasks. It is for this reason that I have characterized this approach as focusing on dialogic criticality.

In supporting the case for collaborative online enquiry, Castek et al. (2012) made the point that the Internet is not only a rich context for learning, it is also a much more complicated one, with an almost infinite range of decision points. They reported that learners were more efficient in conducting online enquiry when they had engaging and worthwhile tasks, and when they worked in pairs to co-construct meaning.

Researchers who have focused on collaborative learning using the Internet believe that small group discussion can encourage more confident and critical evaluations, even with quite young students. Dwyer (2010), for example, combined the socio-cognitive perspectives of Vygotsky (1980, 2012) and the small-group *Thinking Together* approaches of Mercer (2009), when she set up a project to improve the reading and Internet searching capabilities of elementary school students in a high-poverty area of Dublin. Dwyer worked with students in groups of three to tackle Internet search tasks, and she reported on their moving from an uncritical "snatch and grab" approach to a "skillful investigation" level, where the relevance of each search result was assessed by systematically critiquing evidence, and matching this to the task goal. Dwyer ensured that each member of the triad took on a specific role: that of Questioner, Navigator or Summarizer, in order to promote more effective inquiry-based learning.

Harrison (2018b) adapted this approach in his study in England of elementary school triads working collaboratively to determine the trustworthiness and relevance of Internet sites, and he identified nine strategies for enhancing Critical Internet Literacy. These included "Be alert! Be suspicious!", "Read between the lines," "Make late decisions," "Integrate information across sources," and "Make joint decisions". Collaborative learning is not always popular in schools, because it can be difficult to identify the contribution of individuals, but if Dwyer and Harrison are correct in their assertion that collaborative enquiry produces better outcomes, this perhaps indicates a need to develop better tools for assessing such work rather than hold back learning because such tools are not yet available.

Dialogic criticality cannot easily be developed with “snatch and grab” Internet search tasks. As Castek and Beach argued (2013), learning to behave like a scientist in Internet enquiry requires the formulation of hypotheses, gathering data, assessing plausibility, and presenting evidence in compelling ways. Such approaches, they argue, require three things: first, collaboration between peers to collect and assess competing claims; second, multimodality, to present evidence in a variety of formats such as text, concept maps, graphs, photographs or movies; and thirdly, shared productivity, with outcomes presented not only to peers and teachers, but to a wider audience. Fortunately, there is now software that supports such approaches, such as collaborative concept-mapping tools, shared voice-notes or Post-Its, and shared highlighting apps for marking source texts or collaboratively written reports.

The studies that have been summarized above on collaborative and critical Internet enquiry indicate that this will continue to be a rich and important area of future research.

Pedagogical implications

As this review has made clear, in the field of educational technology there is a reciprocal relationship between research and pedagogy. In 1996, the US National Education Technology Plan aimed to connect every student in every classroom to a national fiber-optic “information superhighway,” and the authors made it clear that the proposed two billion dollar program was underpinned by research: “a growing body of research shows that use of technology in the classroom can increase student achievement significantly” (United States Office of Educational Technology, 1996, p. 11). In the UK, from 1989 to 2010, the National Council for Educational Technology, and its successor, the British Educational Communications and Technology Agency, piloted, organized the procurement of, and evaluated new technologies such as laptop computers, smartboards, as well as setting up a National Grid for Learning, which delivered terabytes of firewall-protected curriculum resources to schools. But as Sarah Younie pointed out (2006), the rate at which pedagogy changes as a result of educational policy, even if it is underpinned by sound research, can be very slow. The US National Education Technology Plan of 2016 also acknowledged this, and encouraged stronger partnerships between research and pedagogy: “To make better use of the existing body of research literature, however, educators and researchers will need to work together to determine the most useful dissemination methods for easy incorporation and synthesis of research findings into teachers’ instructional practices” (U.S. Department of Education, 2016).

This review has foregrounded many areas in which research has implications for practice, as educators contemplate how pedagogy will need to adapt to a world in which knowledge is mediated by the Internet. In the area of interaction processes, it is clear that student agency will be an increasingly important factor: The Internet can be an unforgiving and lonely place, and research tells us that students who are highly motivated, and who have a sense of personal involvement in their search activity, will learn more and become more expert. In the future, it is clear that server-side tools, of which LSA is just one example, will become even more powerful, and will inform and deepen the real-time learning of students who are learning on the WorldWideWeb. Search completion is becoming less important than knowledge production, and although writing for the teacher as the implied audience has traditionally been the expected outcome of learning, media production with a potentially worldwide audience will be much more common in the future.

Finally, what I have called dialogic criticality will threaten the traditional authority of the textbook: the teacher’s voice will not be the only contender for authority. In science as well as in the humanities, “truth” will be regarded as in many respects as provisional, and dialogic partners in looking for certainty will include Internet documents and online experts, as well as the students’ own classmates. In such an environment, the ability to cope with uncertainty, to synthesize information from multiple sources, and to judge provenance, relevance and trustworthiness will be paramount. However, none of the findings reported here suggest that the teacher has no place in a classroom of the future. The Internet will not displace the teacher. If anything, research over the past thirty years suggests that teachers will have an absolutely critical role to play in supporting the citizens of tomorrow to learn how to manage the Internet with confidence, criticality, and skill.

Conclusions

In 2001, Hseih-Yee argued that Web search behavior was “fertile ground for research.” This has certainly proved to be the case, and as Leu argued twenty years ago (2000), that research has had to adapt to an ever more rapidly evolving technological environment. The present review has attempted to superimpose some thematic cohesion on a very broad and complex range of research from the fields of computer science, library and information science, education, sociology and psychology, and much of that research has been valuable in moving forward both the field and the pedagogy associated with it. However, at the same time as search engines have become more rapid, more powerful, more user-friendly and more tolerant of spelling errors, the Internet has also become more dangerous. Today the Internet is viewed by businesses as primarily a commercial tool, and by many users as simply a source for gaming and entertainment (Search engine marketing statistics, 2020).

When Tim Berners-Lee set up the Internet as an open platform for exchanging information, he could hardly have predicted that it would become an ideological and political battleground, on which information provided by trustworthy sources had to compete for attention against social media which feed on dogmatism and amplify prejudice. Nearly thirty years after he invented the WorldWideWeb, Berners-Lee (2018) expressed alarm at three trends, all of which threaten not only the Web, but humanity itself. These are the ubiquity of “fake news,” the rise of personal surveillance over the Internet, and cyber-warfare: the attacks by nations attempting

to damage, corrupt or halt the Internet activity of other nations. On a more positive note, however, he pointed out that the problems we face are complex and large, but they were created by people, and they can be fixed by people too. As he put it "I want the web to reflect our hopes and fulfill our dreams, rather than magnify our fears and deepen our divisions."

It will need firm political will as well as sound educational policy to enable the educational and liberating potential of the Internet to be realized. But hopefully the research described in this review makes it clear that we already have many conceptual and computational tools that can enable the adults of tomorrow to harness the potential of the Internet, as well as pedagogical approaches that can support teachers in making those tools available, and helping to ensure that they are used effectively.

References

- Anderson, M., Perrin, A., 2017. Tech Adoption Climbs Among Older Adults. Pew research center, pp. 1–22.
- Becker, N.J., 2003. Google in perspective: understanding and enhancing student search skills. *New Rev. Acad. Librarian* 9 (1), 84–99.
- Berners-Lee, T., 2018. The Web Is under Threat. Join Us and Fight for it. Available from: <https://webfoundation.org/2018/03/web-birthday-29/> (Accessed 28 September 2020).
- Berners-Lee, T., Cailliau, R., 1990. World-Wide Web: Proposal for Hypertext. CERN Memo, November. Geneva, Switzerland.
- Bilal, D., 2002. Children's use of the Yahoo! search engine. III. Cognitive and physical behaviors on fully self-generated search tasks. *J. Am. Soc. Inf. Sci. Technol.* 53 (13), 1170–1183.
- Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3 (4–5), 993–1022.
- Borgman, C.L., Hirsh, S.G., Walter, V.A., Gallagher, A.L., 1995. Children's searching behavior on browsing and keyword online catalogs: the Science Library Catalog project. *J. Am. Soc. Inf. Sci.* 46 (9), 663–684.
- Castek, J., Coiro, J., Guznick, L., Bradshaw, C., 2012. Examining peer collaboration in online inquiry. In: *The Educational Forum*, vol 76. Taylor & Francis Group, pp. 479–496. No. 4.
- Castek, J., Beach, R., 2013. Using apps to support disciplinary literacy and science learning. *J. Adolesc. Adult Literacy* 56 (7), 554–564.
- Cattledge, L.D., Pitkow, J.E., 1995. Characterizing browsing strategies in the world-wide web. *Comput. Netw. ISDN Syst.* 27 (6), 1065–1073.
- Coiro, J., 2011. Predicting reading comprehension on the Internet: contributions of offline reading skills, online reading skills, and prior knowledge. *J. Literacy Res.* 43 (4), 352–392.
- Coiro, J., 2020. Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Read. Res. Q.* <https://doi.org/10.1002/rq.302>.
- Coiro, J., Dobler, E., 2007. Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Read. Res. Q.* 42 (2), 214–257.
- Coiro, J., Knobel, M., Lankshear, C., Leu, D.J., 2008. Central issues in new literacies and new literacies research. In: *Handbook of research on new literacies*, pp. 1–21.
- Dascalu, M., Trausan-Matu, S., McNamara, D.S., Dessus, P., December 1, 2015. ReaderBench: automated evaluation of collaboration based on cohesion and dialogism. *Int. J. Comput. Support Collab. Learn.* 10 (4), 395–423.
- DeSchryver, M., 2012. Toward a Theory of Web-Mediated Knowledge Synthesis: How Advanced Learners Used the Web to Construct Knowledge about Climate Change Behavior. Michigan State University, Educational Psychology and Educational Technology.
- DiMaggio, P., Bonikowski, B., 2008. Make money surfing the web? The impact of Internet use on the earnings of US workers. *Am. Socio. Rev.* 73 (2), 227–250.
- Dwyer, B., 2010. Scaffolding Internet Reading: A Study of a Disadvantaged School Community in Ireland. Doctoral dissertation. University of Nottingham.
- Edwards, L.S., Bruce, C., October 1, 2002. Reflective internet searching: an action research model. *Learn. Organ.* 9 (4), 180–188.
- Eshet, Y., January 2004. Digital literacy: a conceptual framework for survival skills in the digital era. *J. Educ. Multimedia Hypermedia* 13 (1), 93–106.
- Fidel, R., Davies, R.K., Douglass, M.H., Holder, J.K., Hopkins, C.J., Kushner, E.J., Miyagishima, B.K., Toney, C.D., 1999. A visit to the information mall: web searching behavior of high school students. *J. Am. Soc. Inf. Sci.* 50 (1), 24–37.
- Forzani, E., Leu, D.J., Yujia Li, E., Rhoads, C., Guthrie, J.T., McCoach, B., 2020. Characteristics and validity of an instrument for assessing motivations for online reading to learn. *Read. Res. Q.* 56, 761–780.
- Gutu, G., Dascalu, M., Trausan-Matu, S., Dessus, P., 2016. ReaderBench goes online: a comprehension-centered framework for educational purposes. Available from: <https://hal.archives-ouvertes.fr/hal-01374884/>.
- Hargittai, E., Piper, A.M., Morris, M.R., 2019. From internet access to internet skills: digital inequality among older adults. *Univ. Access Inf. Soc.* 18 (4), 881–890. <https://doi.org/10.1007/s10209-018-0617-5>.
- Harrison, C., 2018a. Critical internet literacy: what is it, and how should we teach it? *J. Adolesc. Adult Lit.* 61 (4), 461–464.
- Harrison, C., 2018b. Defining and seeking to identify critical Internet literacy: a discourse analysis of fifth-graders' Internet search and evaluation activity. *Literacy* 52 (3), 153–160.
- Hargittai, E., 2002. Second-level digital divide: differences in people's online skills. *Clin. Hemorheol. and Microcirc.* 7 (4).
- Hirsh, S.G., 1999. Children's relevance criteria and information seeking on electronic resources. *J. Am. Soc. Inf. Sci.* 50 (14), 1265–1283.
- Hölscher, C., Strube, G., June 1, 2000. Web search behavior of Internet experts and newbies. *Comput. Network.* 33 (1–6), 337–346.
- Hsieh-Yee, I., June 1, 2001. Research on Web search behavior. *Libr. Inf. Sci. Res.* 23 (2), 167–185.
- Joo, Y.J., Bong, M., Choi, H.J., June 1, 2000. Self-efficacy for self-regulated learning, academic self-efficacy, and Internet self-efficacy in Web-based instruction. *Educ. Technol. Res. Dev.* 48 (2), 5–17.
- Keyes, R., 2004. *The Post-truth Era: Dishonesty and Deception in Contemporary Life*. Macmillan.
- Kintsch, W., 1998. *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Landauer, T.K., Foltz, P.W., Laham, D., 1998. Introduction to latent semantic analysis. *Discourse Process.* 25, 259–284.
- Lawrence, S., Giles, C.L., 1999. Accessibility and distribution of information on the web. *Nature* 400, 107–109.
- Leader, C., Shah, C., 2016. Practicing critical evaluation of online sources improves student search behavior. *J. Acad. Librarian* 42 (4), 459–468.
- Leu, D.J., 2000. Literacy and technology: deictic consequences for literacy education in an information age. In: Kamil, M.L., Mosenthal, P.B., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research, Vol III*. Erlbaum, Mahwah, NJ, pp. 743–770.
- Litt, E., June 2013. Measuring users' internet skills: a review of past assessments and a look toward the future. *New Media Soc.* 15 (4), 612–630.
- Mercer, N., 2009. Developing argumentation: lessons learned in the primary school. In: *Argumentation and Education*. Springer, Boston, MA, pp. 177–194.
- Ofcom, 2018. Children and Parents: Media Use and Attitudes Report. Available from: https://www.ofcom.org.uk/_data/assets/pdf_file/0024/134907/children-and-parents-media-use-and-attitudes-2018.pdf. (Accessed 28 October 2020).
- Omar, N., Higgins, C., Harrison, C., 2005. July. Investigating an approach for online reading assessment. In: *Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05)*. IEEE, pp. 526–530.
- Ono, H., Zavodny, M., 2007. Digital inequality: a five country comparison using microdata. *Soc. Sci. Res.* 36 (3), 1135–1155.
- Pew Research Center, July 2020. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable. Available from: <https://www.pewresearch.org/topics/social-media/>. (Accessed 17 October 2020).
- Potter, W.J., Thai, C.L., 2019. Reviewing media literacy intervention studies for validity. *Rev. Commun. Res.* 7, 38–66.

- Schön, D.A., 1987. *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. Jossey Bass.
- Search engine marketing statistics, 2020. Available from: <https://99firms.com/blog/search-engine-statistics/#gref>. (Accessed 8 January 2020).
- Sirbu, M., Dascalu, M., Crossley, S., McNamara, D., Trausan-Matu, S., 2019. Longitudinal analysis and visualization of participation in online courses powered by cohesion network analysis. In: Lund, K., Niccolai, G.P., Lavoué, E., Hmelo-Silver, C., Gweon, G., Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings*, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, vol. 2. International Society of the Learning Sciences, Lyon, pp. 640–643. Available from: <https://repository.isls.org/handle/1/4476>. (Accessed 25 October 2020).
- Spiro, R.J., 2012. Cognitive flexibility and hypertext: theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In: *Cognition, education, and multimedia*. Routledge, pp. 177–220.
- Sung, Y.T., Liao, C.N., Chang, T.H., Chen, C.L., Chang, K.E., 2016. The effect of online summary assessment and feedback system on the summary writing on 6th graders: the LSA-based technique. *Comput. Educ.* 95, 1–18.
- Tsai, M.J., Tsai, C.C., 2003. Information searching strategies in web-based science learning: the role of Internet self-efficacy. *Innovat. Educ. Teach. Int.* 40 (1), 43–50.
- United States Office of Educational Technology, 1996. *Getting America's Students Ready for the 21st Century: Meeting the Technology Literacy Challenge: A Report to the Nation on Technology and Education*. US Department of Education.
- US Department of Education, 2016. *Future Ready Learning: Reimagining the Role of Technology in Education*.
- Van Deursen, A.J., van Dijk, J.A., April 1, 2009. Improving digital skills for the use of online public information and services. *Govern. Inf. Q.* 26 (2), 333–340.
- Vygotsky, L.S., 1980. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Vygotsky, L.S., 2012. *Thought and Language*. MIT Press.
- Younie, S., 2006. Implementing government policy on ICT in education: lessons learnt. *Educ. Inf. Technol.* 11 (3–4), 385–400.
- Zhang, S., Duke, N.K., September 1, 2011. The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *Elem. Sch. J.* 112 (1), 132–154.

E-texts and e-books: constructing meaning with and from texts in an online and digital world

Bernadette Dwyer, Institute of Education, School of Language, Literacy & Early Childhood Education, Dublin City University, Dublin, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	551
Emergence of the digital landscape	551
Constructing meaning with and from text in a digital world	552
The key differences of moving from print to screen	552
Skills, strategies, dispositions and social practices for reading both online and digital texts	553
The reader as an assembler	554
The reader as a fixer (adapted from)	554
The reader as a builder (adapted from)	554
The reader as a responder	555
The reader as a composer	555
Internet inquiry and research	555
Forming goals and asking questions	555
Generating search terms and investigating search results	555
Locating information	556
Critical evaluation of online information	557
Communicating and synthesizing information	557
Internet inquiry and research: classroom based research	557
E-book and digital texts	558
Global literature circles	558
The bi-directional nature of reading and response using e-book texts	558
Choice of text matters	559
The ecology of learning in classrooms	559
References	560

Introduction

The [RAND Reading Study Group \(RRSG\) \(2002, p. 11\)](#) describe reading comprehension as “the process of simultaneously extracting and constructing meaning through interaction and involvement with written language.” This interaction involves the reader, the text, the activity and the sociocultural context. Although the report does refer briefly to “electronic text” much has changed in the world and indeed the world of literacy since this report was published.

This chapter considers the changes to reading when you move from page to screen. It explores a range of metaphors suggesting the nested stances of the reader in engaging with online and digital texts. A definition of digital literacy is then offered to underpin the chapter. Next, the key stages in Internet inquiry and emerging research conducted in this field are explored. Following this, research examining reader response using e-books and digital texts, the reading writing connection, and the importance of text choice is presented. Finally, the implications for teaching and learning in the 21st century classroom are considered. First, the emergence of the digital landscape is documented.

Emergence of the digital landscape

We are indeed the connected generation. The Internet is flattening our world ([Friedman, 2005](#)) and almost 66% of the world population is now online ([Internet World Statistics, 2022](#)). The internet and other digital technologies, including smartphones, smart home, and wearable technologies, provide us with instantaneous and immediate access to information, entertainment, and community and affinity groups. During the COVID-19 pandemic technology provided us with a means to stay connected while staying apart through “zooming” with friends, working online with colleagues and providing online classes for our students.

Nevertheless, the pandemic brought into sharp focus issues of equity, equality of opportunity and quality in access to technology as people in different geographical locations and socio-economic classes struggled to gain access to infrastructure, resources, capacities

and reliable broadband connectivity (Scheerder et al., 2017). The pandemic highlighted systemic inequalities for student access to the Internet between low level and higher level income countries (United Nations Children's Fund and International Telecommunication, 2020) and structural inequalities within countries where an urban rural divide was observed for access to infrastructure and connectivity for remote learning (Barron Rodríguez and Cobo, 2022). For parents and caregivers, the importance of well-developed literacy and digital literacy skills were highlighted. For example, in Ireland parents and caregivers with low levels of skills in both literacy and digital literacy skills struggled to support their children during home schooling (National Adult Literacy Agency (NALA), 2020).

The "seamy underworld" (Deibert, 2020, p. 3) and shortcomings of the online world, and social media in particular, was revealed especially with regards to some toxicity in social media commentary, concerns about privacy infringements through surveillance, and the emergence of "fake news" and manipulation of political events. It is clear as boyd contends that the internet "mirrors and magnifies the good, bad, and the ugly" of everyday life (cited in Turner, 2012, p. 180).

Young people are living in media-rich households and either have, or have access to, smartphones and other technologies (Anderson and Jingjing, 2018). They are constantly online (Meeker, 2015) using technology for socialization, leisure time pursuits, education and civic participation (Lorenz and Kapella, 2020). While generally positive about their uses of technology for connectivity, connection, and entertainment, teens point to negative aspects, such as anxiety about social media, online bullying and the lack of in-person connectivity (Anderson and Jingjing, 2018). However, the widespread belief that young people are sophisticated users of technology is not reflected in the research literature (Bennett et al., 2008; Twining, 2021). We need to support our students as they grapple with the complexity and demands placed on the reader when reading digital or online texts. The section which follows explores the changing landscape as readers construct meaning with and from digital and online texts.

Constructing meaning with and from text in a digital world

The Four Resources Model ((1) reader as code breaker, (2) reader as text participant, (3) reader as text user and (4) reader as text analyst) as conceptualized by Freebody and Luke (1990) placed the reader at the heart of the reading process with the power to shape the reading transaction drawing on knowledge, thinking, culture, and experiences. In further iterations, Freebody and Luke (2003) clarified that these roles were not fixed but rather were fluid practices or resources that helped the reader to construct meaning from text. Each of the roles was necessary but the power of the model lay in the totality of the roles embedded in each other as readers accommodated each role. Of particular concern to Freebody and Luke was an overreliance in literacy instruction in schools on the reader as code breaker and text participant. They noted that this reductionist approach to literacy in schools and society ran

the risk of developing young people who are bankrupt consumers when it comes to resourcing the complex demands of contemporary societies, who need to place themselves in a kind of receivership that can manage for them interpretations of the conditions of their lives.

Freebody and Luke (2003, p. 64).

To contend with the complex demands and challenges of contemporary societies our students need to be curious about the world, they need to be critical, empathetic, and creative thinkers, and they need to be highly skilled collaborators and communicators. Students are living concurrently in a print and a digital world and we need to privilege both in teaching and learning. However, research suggests that literacy is more complex when you move from print to screen (Afflerbach and Cho, 2009; Coiro and Dobler, 2007; Kuiper and Volman, 2008). The section which follows considers the key changes to reading in both online and digital e-book multimodal formats.

The key differences of moving from print to screen

Successful print and digital reading share a number of commonalities and key skills, such as rapid decoding and word recognition to aid reading fluency; and specific strategies like, monitoring understanding, identifying and locating information, and evaluating text (Afflerbach and Cho, 2009; Pressley and Afflerbach, 1995). However, reading online and reading digital e-book texts introduces both challenges and opportunities for the reader and may require deeper levels of higher order processing skills, strategies, social practices and dispositions.

First, text is no longer the dominant mode and information or narrative may be presented in multiple ways including text, image, graph, braille, performance, video or audio modes. Moreover, print text is stable and fixed whereas online and digitized text is fluid. Second, the interaction between the reader and text may be more dynamic, opportunistic, and interactive in an online or digitized environment (McEneaney, 2006). Third, reading online is more effortful in "realizing and constructing potential texts to be read" (Afflerbach and Cho, 2009, p. 209). For example, it is the reader rather than the author who determines a circuitous pathway through a labyrinth of websites. Reader agency is heightened and it is the reader's interests that determine the reading path through

active decisions making processes. Fourth, digital texts in electronic formats (e-books) afford a myriad of possibilities to provide a “scaffolded digital reading environment” (Dalton and Proctor, 2008, p. 303) which is responsive, flexible and supportive to the needs of readers and writers (Rose and Meyer, 2002). Embedded multimodal supports include, a text-to-speech functionality, audio narration, an inbuilt dictionary, hotspots, links, and a customizable background and font size. Such embedded supports provide an interactive learning experience, but may also be irrelevant and distracting particularly for struggling (Dalton, 2014) and younger readers (Bus et al., 2019; Rowe and Miller, 2017). Finally, reader control and choice is heightened. While this may empower the skillful reader it may be daunting for a less proficient reader as they take responsibility for their own reading and learning in a shifting environment.

So a range of skills, strategies, social practices and dispositions (3SD) (Dwyer, 2020) are used by readers and required by our students in engaging in Internet inquiry and research and developing reading comprehension with digital e-books. The section which follows considers these in turn.

Skills, strategies, dispositions and social practices for reading both online and digital texts

The reader needs to be both skilled and strategic in an online and digital environment. Whereas strategies are conscious, effortful, deliberate, active and purposeful actions on the part of the reader when constructing meaning from text, skills are characterized by automaticity, fluency, effortlessness and effectiveness, often without the conscious control of the reader (Afflerbach et al., 2008; Alexander, 2006; Pressley, 2000). These terms are sometimes used interchangeably in the literature (Afflerbach et al., 2008). Paris et al. (1983) describe the active nature of reading strategies as being “skills under consideration” (Paris et al., 1983, p. 295). The online and e-book reader needs to be both skilled and strategic and needs to switch seamlessly (Afflerbach et al., 2008, p. 371) between the automaticity of the skill to the effortful use of a reading strategy. This happens in print-based environments when a student meets a demanding text. It is more heightened in an online environment where a student is in a constant state of flux moving from the automaticity of highly developed skills to conscious application of strategies as they encounter more demanding texts, modes and digital tools.

The online reading comprehension skills and strategies include activating prior knowledge, making connections, asking questions, making inferences, monitoring understanding, determining important ideas, evaluating, summarizing, synthesizing, and communicating. Dispositions for reading are important to consider in a digital or online environment. Studies (Dwyer, 2010; Leu et al., 2004; Putman, 2014) point to issues relating to self-regulation, curiosity, resilience, self-efficacy, persistence, intrinsic motivation and an ability to set learner-centered goals. A classroom culture that promotes a sense of community helps students to develop these dispositions. Social practices within such a classroom culture include working collaboratively with others, supporting each other’s thinking and problem solving to create and apply solutions in a creative way.

Digital literacy involves more than the functional use of digital devices and is the “ability to creatively engage in particular *social practices* to assume appropriate *social identities* and to form and maintain *social relationships*” (italics in original) (Jones and Hafner, 2012, p. 12 cited in Pangrazio et al., 2020). Combining the skills, strategies, social practices and dispositions needed to be digitally literate digital literacy can be described as follows:

a complex orchestration of a range of skills, strategies, dispositions and social practices to leverage technology to access, acquire, comprehend, analyze, evaluate, create and communicate knowledge in multiple modes and through multiple representations to and with a variety of audiences and in a variety of contexts for a range of purposes.

Dwyer (2020).

This definition is offered tentatively in the knowledge that literacy is a deictic construct (Leu, 2000) that is constantly evolving and morphing. Likewise, the notion of being digitally literate is better considered as becoming digitally literate in an ever-changing world (Reinking, 1998). The definition stresses the importance of developing skills, strategies, dispositions and social practices (3SD) to leverage technology in meaningful ways through a complex orchestration of processes to enhance literacy and learning in the classroom. This can only happen when teachers adopt socially mediated practices to create authentic learning environments which gives primacy to curriculum goals and learning outcomes for students before considering the meaningful use of technology to enhance literacy and learning in the classroom (Mishra and Koehler, 2006).

Other definitions stress that being digitally literate requires both cognitive and technical skills and the ability to critically examine how texts are constructed and produced, whose voices are heard and whose voices are silenced (ILA, n.d). The National Council of Teachers of English (NCTE) (2019) emphasize the importance of equitable participation to engage effectively in a networked, globalized world as people consume, curate and create actively across different contexts. Coiro (2020, p. 4) conceptualizes digital literacy as involving a “process of integration and construction in social and cultural practices” that is enacted for particular purposes and in diverse contexts.

In sum, a range of skills, strategies, social practices, and dispositions need to be orchestrated, intertwined and embedded (each necessary but insufficient individually) when reading online or in a digital e-book format. Over many years, I have been working in elementary school classrooms with teachers and students as they build online Internet inquiry skills and read and respond to e-texts

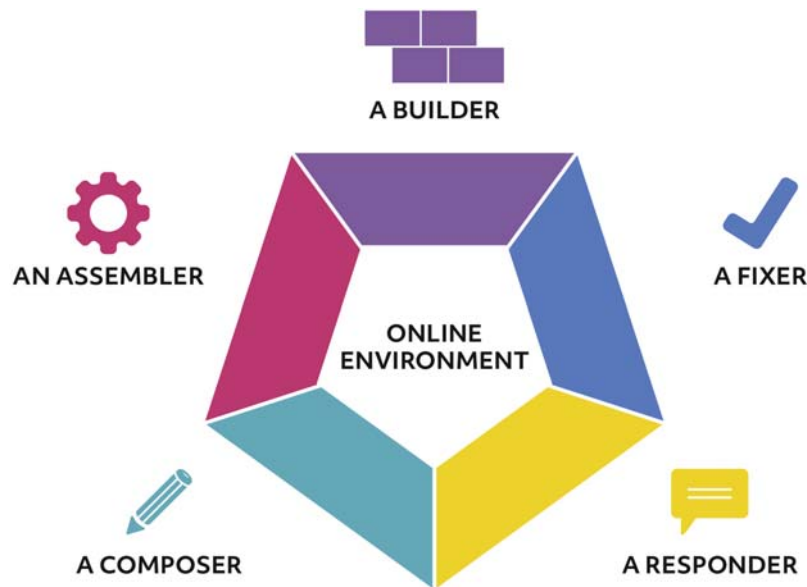


Fig. 1 The nested stances of the online and digital reader as an assembler, builder, fixer, responder, and composer.

and e-books. Analysis of many hours of observations, interviews and captured data of online screen reading suggests that the reader adopts a bricolage of nested stances when engaged with these multimodal formats. These stances can be conceptualized using the metaphors of the reader as an assembler, builder, fixer, responder, and composer (Dwyer, 2010, 2016; Dwyer and Harrison, 2018). See Fig. 1.

The reader as an assembler

The reader as assembler views the reader as an active participant in the reading process. Unlike print based text in a book which is fixed and stable, online and digitized text is fluid and opportunistic. For example, in an online environment, the reader must assemble the text to be read by creating search terms and investigating search results.

There is no predetermined route and the reader carves a unique pathway through a nonlinear, multimodal, online and digital space. Multiple reading paths are possible given the compositional nature and spatial arrangement of multimodal texts (Kress, 2010). The reader assembles the resources needed to support reading and has the power to shape the reading transaction with media and learning supports.

The reader as a fixer (adapted from Pearson, 2009)

The reader as a fixer draws on the levels of declarative, procedural, and conditional knowledge (Paris et al., 1983) that the strategic reader brings to the reading process.

In an online environment the reader engages in active decision making processes while making intertextual links (Hartman, 1995) across multiple websites and multiple modes of representation. In a digital e-book format the reader actively customizes the reading space through a layering of embedded supports like text-to-speech functionality, audio narration, an inbuilt dictionary, and external link supports like video and podcasts. A difficulty for the reader is “remaining in the reading moment” (Dobler, 2015, p. 487) and avoiding the cognitive overload when transitioning from one mode to another.

The reader as a builder (adapted from Pearson, 2009)

The reader as a builder assimilates prior experiences and prior knowledge sources with experiences from their own cultures and communities in transacting with the text (Anderson, 2004).

In an online environment the reader must supplement this range of prior knowledge sources to include navigational and online user application knowledge (e.g. activating browser features) and online informational text structure knowledge (e.g. inferring the hidden content of hyperlinks (Coiro and Dobler, 2007)). In an e-book the reader constructs meaning by fusing together a range of interrelating modes and embedded links while concurrently avoiding the “bells and whistles” that are either irrelevant or distracting.

The reader as a responder

The reader as a responder needs to construct meaning with, from and against texts (Janks, 2019). Given that anyone can publish to an online forum, the online reader as a responder acts as a gatekeeper, curator and editor of online information. This calls for heightened levels of critical thinking, critical evaluation, and media literacy skills.

The online reader must assess the credibility, authority, reasonableness, believability, author agenda, and applicability of information in relation to their inquiry purpose. In an e-book format the reader can customize and personalize the reading experience, and can respond in real time through digital thinkmarks. Digital thinkmarks are text annotations and notes which capture fleeting thoughts and responses as the reader engages with text (Dwyer and Larson, 2014).

The reader as a composer

The reader as a composer must curate, organize, summarize, and synthesize information across multiple websites and multiple modes in e-books to create, remix and communicate information to others or respond to e-books and share with others in multiple modes.

Dalton (2014) has noted that e-books and e-texts are changing the reading palette through inbuilt supports. The reader as a composer needs both receptive and expressive skills. Indeed, when reading online or digital texts it is difficult to separate the bi-directional nature of response and composition. Collaboration is an integral part of digital composing as students work with distributed expertise, in pairs or small groups (Smith, 2019).

In sum, reading in an online and digital space involves a bricolage of skills, strategies resources, and practices as the reader adopts the nested stances of assembler, fixer, builder, responder and composer when transacting with such texts.

Internet inquiry and research

Research in the area of Internet inquiry is driven by a constellation of theories. These theories include multiliteracies, (The New London Group, 2000); New Literacies Studies and multimodality (Kress, 2010; Rowsell et al., 2019); and New Literacies (Coiro et al., 2008; Leu et al., 2004). These theories provide lenses to explore the complexity of Internet inquiry from cognitive, cultural, social and affective dimensions.

The ubiquitous nature of the Internet and other ICTs is influencing both what we read and how we read information and subsequently shaping our social, political, cultural, economic and civic lives. Many students are “hanging out and messing around” (Ito et al., 2009) with technology in their out-of-school lives composing, remixing and uploading content to Face-book and TikTok, connecting with others and social positioning on Twitter, WhatsApp and Instagram and using the Internet for entertainment purposes. However, research suggest that our students are struggling to use the Internet for research and academic purposes. Moreover, Agosto (2002, p. 17) noted the “satisficing” behavior of students, (i.e. finding sufficient information to satisfy their information needs), as they browse, click, surf, and skim online. Meanwhile educators are striving to find meaningful ways to embed digital technologies to enhance and deepen student learning, to scaffold and support their students online, to challenge and nurture student creativity, and to develop students’ skills in collaboration, critical thinking and communication during the Internet inquiry and research cycle.

The key stages during Internet inquiry and research (Dwyer, 2010; Leu et al., 2004) include forming goals and asking questions, generating search terms, investigating search results, locating and critically evaluating information and communicating information to others. Fig. 2 shows this recursive cycle including key skills in each stage. We will now consider these stages in turn.

Forming goals and asking questions

Conducting internet inquiry and research in a complex, ill-structured online domain (Spiro et al., 2004) requires resilience, persistence, self-efficacy, self-regulation, and cognitive flexibility. If our students are to experience success during Internet inquiry and research they need to construct engaging questions to provide a purpose for their inquiry, set a context for problem solving, and establish a goal for learning (Castek and Dwyer, 2018).

Developing situational and in turn personal interest in the topic is key. Creating an authentic inquiry-based learning opportunity, or setting a purpose or task, based for example, on the lived experiences of students in their communities, on a real world global to local issue, on a field trip, or expert visit builds situational interest. This can spark students’ motivation, curiosity, and civic engagement. Student self-generated questions drive student interest and curiosity (Kuiper and Volman, 2008). In classrooms overt modeling by the teacher will enable students to ask higher level what if, why, how, and “I wonder” questions. In addition, encouraging student to collaborate to formulate and post questions to an online share board like Padlet (<https://padlet.com>) provides an audience of peers and helps students to generate more challenging questions.

Generating search terms and investigating search results

Generating and choosing effective key search terms involves active decision making and executive functions. Students input search terms either through typing or through voice search on a search engine. Vocabulary knowledge is key for the successful generation of

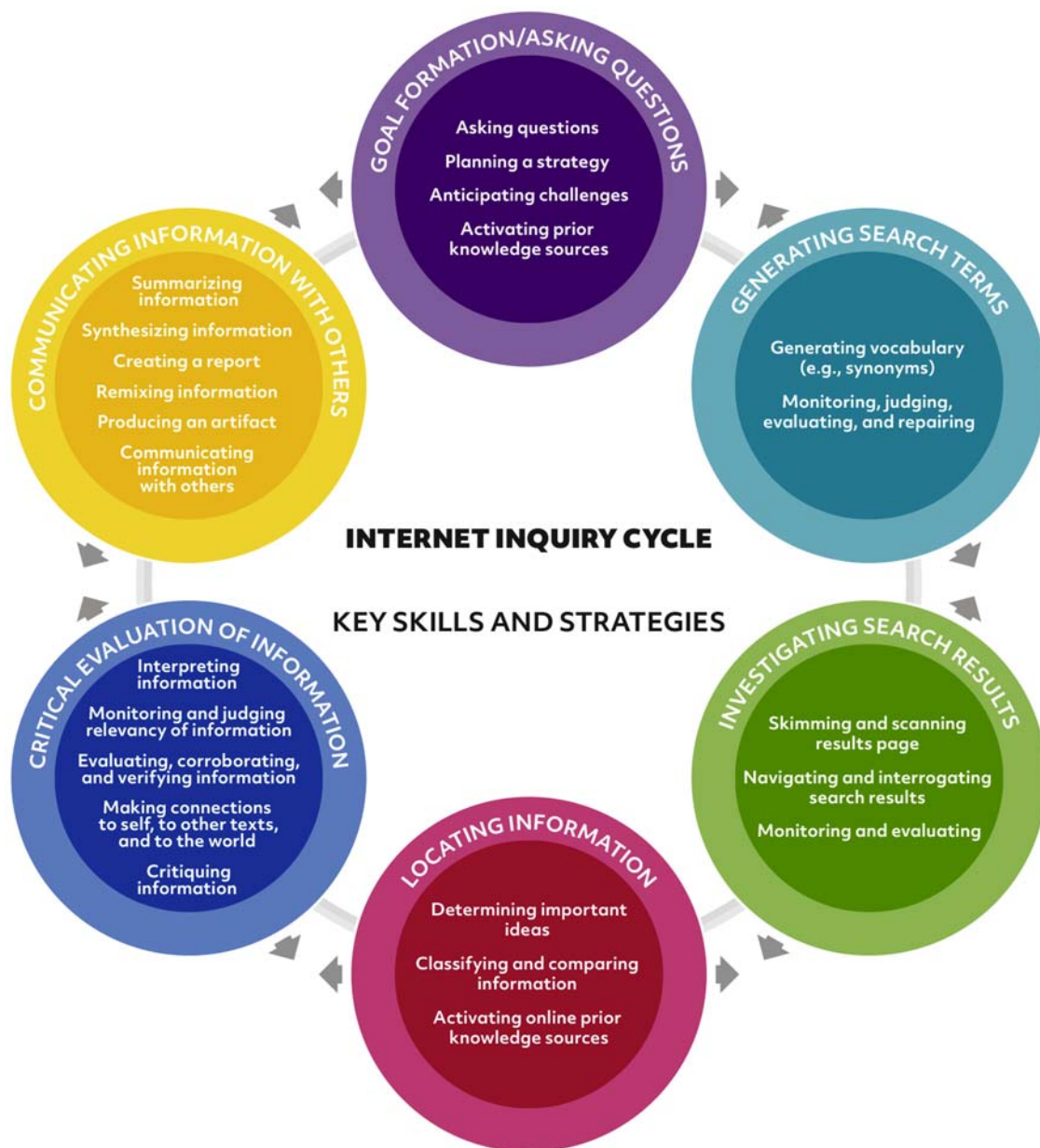


Fig. 2 Internet inquiry and research: key skills and strategies.

a just right key term for Internet inquiry. Monitoring, judging, evaluating, and revising search terms requires sufficient vocabulary knowledge related to the domain and an ability to generate synonyms related to the topic.

Key skills in interrogating search results include adopting a questioning stance to navigate the blurb and the URL with a critical eye, to evaluate the source of the information and critique the relevance and reliability of both to the task or inquiry at hand. Research suggests that elementary school students rarely go beyond the first or second result returned by the algorithm of the search engine adopting a snatch and grab or limited strategic scrolling approach rather than a speedy and skillful investigation of search results (Dwyer, 2010).

Locating information

The ability to engage effectively in the process of Internet inquiry and research through effective key word generation and investigation of search results facilitates the location of information related to the inquiry. Locating information involves skimming, scanning, and filtering information across multiple websites to separate the “wheat from the chaff” to build and update knowledge on-the-hoof in the malleable moments of an online search.

Locating information involves components, such as locating, transforming, synthesizing, and evaluating. These are cognitively demanding comprehension processes requiring dispositions and mind-sets to avoid the cognitive overload often experienced by readers in an online environment who expect information to be a mere click away.

Critical evaluation of online information

Anyone can post information on the Internet. Print-based media, with a five century start on online media, have gatekeepers, such as peer review and editors. The Internet has shifted the burden of review onto the reader to assess the accuracy, trustworthiness, objectivity and credibility of the information presented (Metzger, 2007).

Research suggests that online readers (children, adolescents, and adults) are struggling with the task (Cho et al., 2018; Goldman et al., 2012; Stanford History Education Group, 2016). A National Literacy Trust survey in the UK (2018) noted that 2% of children and young people have the critical literacy skills to ascertain if a news story is fake or real. Moreover, almost 50% of this cohort expressed anxiety about their inability to spot such stories.

Critical evaluation, critical thinking and critical multimedia skills are required to evaluate, corroborate, and interrogate the reliability and veracity of the information presented online. The online reader must “read with and against the text” (Janks, 2019, p. 561) to interrogate the perspectives presented, to question how the text is built and how the text defines the world with regard to gender, power and social issues. Elementary and middle school students draw on limited prior experiences and world knowledge (often referred to as common sense knowledge) to assess and evaluate online information. Therefore, critical evaluation skills may be developmental by nature. Students need to raise their antennae to develop a questioning stance and a healthy skepticism toward online information.

Communicating and synthesizing information

Summarizing and synthesizing information pose difficulties for our students (Dole et al., 1991). These difficulties are compounded in an online environment as students navigate across websites and modes within those websites, using short term cognitive energy to sift, determine importance, condense, and transform information.

It is important for our students to reflect on the processes of Internet inquiry and research within the community of learning in their classroom. Further, students can represent their learning and knowledge through multiple means of representation using a range of digital tools to apply their knowledge in new and creative ways. The Internet and other digital technologies expand our vision for audience beyond the walls of the classroom to a global audience.

Internet inquiry and research: classroom based research

Recent research in the area of Internet inquiry and research has focused on authentic, developmentally appropriate, inquiry-based (Castek and Dwyer, 2018; Coiro et al., 2016; Spire et al., 2019) or project-based learning (PBL) opportunities (Chen and Yang, 2019). These approaches share a number of commonalities. The focus is on curriculum alignment, global connectedness, challenging real-world problems or the lived experiences of students. Students engage in sustained inquiry, problem solving and critique which leads to action and the development of a solution, artifact, or a public product. Learning communities are established which privilege student choice and agency as students acquire and build knowledge. Social practice, collaboration and reflection among students are keys to success.

Many of the studies are transdisciplinary where students explore topics as mathematicians, historians, geographers, scientists and engineers engaging in the distinct social, semiotic, cognitive and linguistic processes of these disciplines. Student curiosity, personal and situational interest is built through reading, writing, talking, and doing. Some of the approaches assign roles to students which empower students, ensure equitable participation, and play to the strengths of individual students. For example, in the Castek and Dwyer (2018) study students engaged with the roles of questioner, navigator, responder, designer and engineer. Both print and digital texts are privileged and multiple means of expression and representation are encouraged through the use of a range of digital tools and modes.

A meta-analysis on PBL research (Chen and Yang, 2019) demonstrated a medium to large effect on students’ achievement compared with traditional classroom approaches. Recent results from a number of randomized control trial PBL studies, conducted in elementary schools, show increased achievement and significant gains in science knowledge and social and emotional learning for students (Krajcik et al., 2021) and social studies and literacy (Duke et al., 2021) when PBL is rigorously applied.

These studies show the importance of exploring real world issues in developmentally appropriate and creative ways to provide meaningful opportunities to learn across multiple disciplines (Castek and Manderino, 2017). In the section that follows we will look at the construction of meaning as students engage with e-books and digital texts.

E-book and digital texts

Rosenblatt's reader response theory (1938, 1978) has been seminal in shaping our understanding of how readers construct meaning when transacting with a text. As noted by Reinking (2008, p. 1178) reading online or on digital texts move beyond literacy events to "instantiate literacy practices because they are overtly and consistently social acts." Such literacy practices are shaped not only by the personal response of the reader but also by context, culture and identity (Rowse et al., 2019). In turn the attributes and identity of the reader/learner (*who*) as they interact with a text (*what*), in a particular context (*where*), and for what reason (*why*) shapes the reading transaction (Singer and Alexander, 2017, words in italics added).

The *who*, *what*, *when*, *where* and *why* has shaped much of the recent research work carried out around the use of e-books and e-texts in classrooms and other settings. The section which follows considers reading transaction within a community of readers; the bi-directional nature of reading and response; and the importance of choice and culturally relevant texts to engage and motivate students. As readers transact with text they draw on textual content, and their knowledge and experiences by means of the text base and situation model respectively (Kintsch, 1998; Singer and Alexander, 2017). When readers engage in a community of readers they have multiple resources at the reading construction site (Kucan and Beck, 1997) that may well stimulate and extend the individual's thinking and advance their learning and understanding.

Global literature circles

Literature circles (Daniels, 2002) or Book Clubs (Raphael et al., 2004) are systematic student-led, small group discussions around literary texts. Students adopt interchangeable roles as they read, such as artful artist, passage picker, connector or questioner, and meet to discuss their response in small groups. Simpson (2014, p. 2) describes these roles as "like shaking the crystals in a kaleidoscope. With each change of role the image the student has of a book is fractured and remade to reveal a new pattern of meaning." However, these roles may become staid and formulaic and not lead to the grand conversations envisaged by Eeds and Wells (1989).

In a series of studies (Dwyer and Larson, 2014; Larson and Dwyer, 2015) sixth grade students from the USA and Ireland read e-book versions of the same young adult novels on digital reading devices. These exploratory studies sought to investigate what happens when students from diverse cultures and backgrounds interact and collaborate in an online global literature circle when reading and responding to an e-book text. First, the students crafted a personal response by annotating the text and creating digital thinkmarks as they read. These digital thinkmarks captured the fleeting thoughts and response moments of students as they read providing an immediacy and physicality to their response (Walsh, 2013). The digital thinkmarks provided thinking time for the students and insights into their meaning making processes.

Next, the students met face-to-face in their own contexts and discussed their response. These discussions aided the students in posting questions and discussion prompts onto an asynchronous message board. The message board served as a springboard for a global literature circle where all students responded and engaged in threaded discussions around the chosen texts. Students posted responses using emoticons, font color, highlighting features and text size to establish their personal identity. They built social connection, affinity and a participatory culture (Jenkins, 2006) by discussing issues related to culture, education, religion, and personal interests while concurrently posting their responses to the text on the discussion threads. Through discussion their response to text evolved from immersion, to involvement, to interpretation as students drew on cultural, experiential, esthetic, cognitive, critical, and empathetic responses.

Analysis of data sources suggested that reading in a community of readers, in a global digital environment, deepened reader response and enhanced engagement. Further the studies suggested a blurring of lines, in a digital reading environment, between reader and writer, consumer and producer, and author and audience as the students responded to texts through the creation of blog posts, digital posters and videos using a range of digital tools. In the section which follows we consider reader response with e-books.

The bi-directional nature of reading and response using e-book texts

It is difficult to write about reading without discussing response. Exploratory and descriptive studies (Dalton et al., 2015; Smith, 2018; Smith et al., 2020) conducted in the area of e-book reader response and composition share a number of common threads.

First, multimodal composition draws on socio-cultural perspectives viewing literacy practices as socially constructed and mediated. In addition, this work is also rooted in social semiotic theories (Kress, 2010) stressing that various modes are integral to constructing meaning. Making space for students to appropriate a range of modes of response through a range of digital tools engenders critical engagement, creativity, and agency.

Second, the Dalton et al. (2015) studies in fifth grade noted the importance of a teacher scaffolded digital composing environment where the students composed retellings using PowerPoint. In retrospective interviews students expressed both a design intentionality and a metamodel awareness of how modes gelled together to create meaningful responses to folk tales. In a review of the literature of emerging bi-lingual students (Smith et al., 2020) the authors described how digital multimodal composition supported identity expression and agency as students appropriated the unique resources of multiple modes when composing.

Third, collaboration among students is important as students work in partnerships building on the strengths of each other in unique ways. For example, Smith (2018) explored three pairs of culturally and linguistically diverse students analyzing literature relating to the Vietnam War. Students responded to the literature by collaboratively designing a website, creating hyperlinked

PowerPoints and creating a podcast response from a character in the literature. The technical and multimodal composing expertise was distributed and negotiated among the pairs as they adopted various roles (designer, creative visionary and technical expert) depending on their individual strengths.

Choice of text matters

When students' lives, interests and experiences are reflected in the texts used in classrooms personal interest and investment in these texts also grows. The texts chosen should provide opportunities for deep personal and emotional connection and space for critical thinking and response as students engage with social, political and cultural issues (McLaughlin and DeVoogd, 2004). The oft referenced metaphors of windows, mirrors and sliding glass doors (Bishop, 1990) encapsulates the importance of culturally relevant texts (Sharma and Christ, 2017) where students' lives and community are present and relevant in classroom pedagogies and in their learning (Ladson-Billings, 2014; Paris and Alim, 2017). Readers should see themselves, their "funds of knowledge" (Moll et al., 1992, p. 133) from their cultures and experiences mirrored in the texts they read. Moreover, readers are transported through doors and windows into other worlds, other cultures, other experiences, other ways of being. Studies reflect the importance of culturally relevant texts that mirror, reflect, and validate cultural norms and reader experiences (Wood and Jocius, 2013) and that challenge representations of disability, stereotypes and difference by creating collaborative spaces for critical conversations around texts and in turn promoting community action projects to promote social justice (Jocius and Shealy, 2018).

In the chapter we have explored Internet inquiry and reader response as students construct meaning with and texts in a digital world. The final section considers the implications for teaching and learning within the ecology of learning in the classroom.

The ecology of learning in classrooms

In the report entitled *Students, Computers and Learning: Making the Connection* (OECD, 2015) the authors note that the promise of technology to enhance teaching, and in turn, literacy and learning in schools has not been realized

The impact of technology on education delivery remains sub-optimal, because we may overestimate the digital skills of both teachers and students, because of naïve policy design and implementation strategies, [and] because of a poor understanding of pedagogy.

OECD (2015, p. 4).

The approach to the integration of technology in schools has, in essence, been a "build it and they will come" approach. Serious consideration has not been to the goals of such integration or the trajectory by which curriculum and pedagogy could be aligned with the meaningful use of technology to enhance both (Harrison, 2015). The approach also ignores the "multiple realities" (Labbo and Reinking, 1999, p. 478) of the classroom learning ecology. The classroom learning ecology involves curriculum and pedagogies and a complex interplay and interdependency between teachers and their students, between students and their peers, in a social learning environment. It also involves relationships with school administrators, parents and care-givers, the broader community and societal and political influences.

We need to recognize the professionalism of teachers and acknowledge their professional literacy practices, and their own positions and histories as professional educators. Teachers need to engage with professional learning opportunities that are sustained, on-site, and customized to the literacy needs of the school community. In such learning communities, teachers can develop the pedagogies and strategies required to provide the scaffolded learning environments that students require as they interact with print texts, e-texts on the Internet and e-books. Working with the Internet and other digital technologies repositions the role of the teacher from the "most knowledgeable other" in the classroom to a co-creator and co-learner with students. This promotes a mutual interdependence and respect between teacher and student.

A classroom culture that encourages collaboration among students both shapes and influences how learners make sense of the world around them. These learning environments place student curiosity at the center of the learning experience and involves students in structured cross-disciplinary inquiries often based on developmentally appropriate, diverse, compelling, authentic real world issues. Students pose questions, investigate, wonder, design solutions in collaboration with others, build and act on new understanding and knowledge and reflect on new conceptual knowledge within a community of learners involving both teachers and students. It is important that students develop the "cosmopolitan dispositions and habits of mind" (Hull and Stornaiuolo, 2010, p. 89) to connect, create, collaborate, and communicate within and across a global community (Dwyer, 2018).

We need a robust research-to-practice focus that is firmly grounded in classrooms and other learning spaces to see what works and in what circumstances (Bradley and Reinking, 2011). As researchers we should work in collaboration and partnership with teachers and their students to drive assessment practices, curriculum design, learning spaces, and pedagogies. Policy development and curriculum must be firmly informed by this research.

References

- Afflerbach, P., Cho, B.Y., 2009. Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In: Israel, S.E., Duffy, G.G. (Eds.), *Handbook of Research on Reading Comprehension*. Routledge, New York, pp. 69–90.
- Afflerbach, P., Pearson, P.D., Paris, S.G., 2008. Clarifying differences between reading skills and reading strategies. *Read. Teach.* 61 (5), 364–373.
- Agosto, D.E., 2002. Bounded rationality and satisficing in young people's Web-based decision-making. *J. Am. Soc. Inf. Sci. Technol.* 53, 16–27.
- Alexander, P.A., 2006. The path to competence: a lifespan developmental perspective on reading. *J. Lit. Res.* 37, 413–436.
- Anderson, R.C., 2004. Role of reader's schema in comprehension, learning, and memory. In: Ruddell, R.B., Unrau, N.J. (Eds.), *Theoretical models and processes of reading*, 5th. International Reading Association, Newark, DE, pp. 594–606.
- Anderson, M., Jingjing, J., 2018. *Teens, Social Media and Technology*. Pew Internet Research, USA.
- Barron Rodríguez, M.B., Cobo, C., 2022. COVID-19 and education in the global South: emergency remote learning solutions with long-term implications. In: Vincent-Lancrin, S., Cobo Romani, C., Reimers, F. (Eds.), *How Learning Continued During the COVID-19 Pandemic: Global Lessons From Initiatives to Support Learners and Teachers*. OECD Publishing, Paris. <https://doi.org/10.1787/897cd0d8-en>.
- Bennett, S., Matton, K., Kervin, L., 2008. The “digital natives” debate: a critical review of the evidence. *Br. J. Educ. Technol.* 39 (5), 775–786.
- Bishop, R.S., 1990. Mirrors, windows and sliding glass doors. *Perspectives* 6 (3).
- Bradley, B.A., Reinking, D., 2011. Revisiting the connection between research and practice using formative and design experiments. In: Duke, N.K., Mallette, M.H. (Eds.), *Literacy Research Methodologies*. Guilford Publications, New York, pp. 188–211.
- Bus, A.G., Sari, B., Takas, Z.K., 2019. The promise of multimedia enhancement in children's digital storybooks. In: Kim, J.E., Hassinger-Das, B. (Eds.), *Reading in the Digital Age: Young Children's Experiences With E-Books*. International Studies With E-Books in Diverse Contexts. Springer, Cham, Switzerland, pp. 45–58.
- Castek, J., Dwyer, B., 2018. Think globally act locally: teaching climate change through digital inquiry. *Read. Teach.* 71 (6).
- Castek, J., Manderino, M., 2017. A planning framework for integrating digital literacies and disciplinary learning. *J. Adolesc. Adult Lit.* 60 (6), 697–700.
- Chen, C.-H., Yang, Y.-C., 2019. Revisiting the effects of project-based learning on students' academic achievement: a meta-analysis investigating moderators. *Educ. Res. Rev.* 26, 71–81.
- Cho, B.-Y., Woodward, L., Li, D., 2018. Epistemic processing when adolescents read online: a verbal protocol analysis of more and less successful online readers. *Read. Res. Q.* 53 (2), 197–221.
- Coiro, J., Dobler, E., 2007. Exploring the comprehension strategies used by sixth-grade skilled readers as they search for and locate information on the Internet. *Read. Res. Q.* 42, 214–257.
- Coiro, J., Knobel, M., Lankshear, C., Leu, D.J., 2008. Central issues in new literacies and new literacies research. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 1–21.
- Coiro, J., Castek, J., Quinn, D.J., 2016. Personal inquiry and online research: connecting readers in ways that matter. *Read. Teach.* 69 (5), 483–492.
- Coiro, J., 2020. Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Read. Res. Q.* 56 (1), 9–31.
- Dalton, B., Proctor, C.P., 2008. The changing landscape of text and comprehension in the age of the new literacies. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 297–324.
- Dalton, B., Robinson, K.H., Loworn, J.F., Smith, B.E., Alvey, T., Mo, E., et al., 2015. Fifth-grade students' digital retellings and the Common Core: modal use and design intentionality. *Elem. Sch. J.* 115 (4), 548–569.
- Dalton, B., 2014. E-texts and e-books are changing the literacy landscape. *Phi Delta Kappan* 96 (3), 38–43.
- Daniels, H., 2002. *Literature Circles: Voice and Choice in Book Clubs and Reading Groups*. Stenhouse, Portland, Maine.
- Deibert, R.J., 2020. *Reset: Reclaiming the Internet for Civil Society*. House of Anansi Press Inc, Canada.
- Dobler, E., 2015. E-Textbooks: a personalised learning experience or a digital distraction? *J. Adolesc. Adult Lit.* 58 (6), 482–491.
- Dole, J.A., Duffy, G.G., Roehler, L.R., Pearson, P.D., 1991. Moving from the old to the new: research on reading comprehension instruction. *Rev. Educ. Res.* 61 (2), 239–264.
- Duke, N.K., Halvorsen, A.-L., Strachan, S.L., Kim, J., Konstantopoulos, S., 2021. Putting PjBL to the test: the impact of Project-Based Learning on second graders' social studies and literacy learning and motivation in low-SES school settings. *Am. Educ. Res. J.* 58 (1), 160–200.
- Dwyer, B., Harrison, C., 2018. *Using Technology to Improve Reading and Learning*. United Kingdom Literacy Association, UK.
- Dwyer, B., Larson, L., 2014. The writer in the reader: building communities of response in digital environments. In: Pytash, K.E., Ferdig, R.E. (Eds.), *Exploring Technology for Writing and Writing Instruction*. IGI Publications, USA, pp. 202–220.
- Dwyer, B., 2010. *Scaffolding Internet Reading: A Study of a Disadvantaged School Community in Ireland* (Unpublished doctoral dissertation). University of Nottingham, Nottingham, UK.
- Dwyer, B., 2016. Engaging all students in internet inquiry. *Read. Teach.* 69 (4), 383–389.
- Dwyer, B., 2018. Teaching and learning in the global village: connect, create, collaborate, and communicate. *Read. Teach.* 70 (1), 131–136. <https://doi.org/10.1002/trtr>.
- Dwyer, B., 2020. *Internet Research and Inquiry Cycle*. Support Materials for Teachers. National Council for Curriculum and Assessment (NCCA), Dublin. Retrieved from: <https://www.curriculumonline.ie/getmedia/494184f0-d17b-468d-8173-23c4e5513d4d/Internet-research-and-inquiry-cycle.pdf>.
- Eeds, M., Wells, D., 1989. Grand conversations: an exploration of meaning construction in literature study groups. *Res. Teach. Engl.* 23 (1), 4–29.
- Freebody, P., Luke, A., 1990. Literacies programs: debates and demands in cultural context. *Prospect* 5 (7), 7–16.
- Freebody, P., Luke, A., 2003. Literacy as engaging with new forms of life: the “four roles” model. In: Bull, G., Anstey, M. (Eds.), *The Literacy Lexicon*, second ed. Frenchs Forest Pearson Education, NSW, pp. 51–66.
- Friedman, T.L., 2005. *The World is Flat: The Globalised World in the Twenty-First Century*. Penguin, London, UK.
- Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., Brodowinska, K., 2012. Comprehending and learning from Internet sources: processing patterns of better and poorer learners. *Read. Res. Q.* 47 (4), 356–381. <https://doi.org/10.1002/RRQ.027>.
- Harrison, C., 2015. We're closing the digital divide: now let's work on closing the teleological divide. In: Spiro, R.J., Deschryver, M., Hagerman, M.S., Morsink, P.M., Thompson, P. (Eds.), *Reading At a Crossroads? Disjunctures and Continuities in Current Conceptions and Practices*. Routledge Taylor & Francis Group, New York, pp. 162–171.
- Hartman, D.K., 1995. Eight readers reading: the intertextual links of proficient readers reading multiple passages. *Read. Res. Q.* 30 (3), 520–561.
- Hull, G.A., Stornaiuolo, A., 2010. Literate arts in a global world: reframing social networking as cosmopolitan practice. *J. Adolesc. Adult Lit.* 54 (2), 85–97.
- International Literacy Association (ILA). n.d. *Literacy Glossary*. Retrieved from: <https://www.literacyworldwide.org/get-resources/literacy-glossary>.
- Internet World Statistics, 2022. *Internet Usage Statistics*. The Big Picture: World Internet Users and Population Stats. Retrieved from: <http://www.internetworldstats.com/stats.htm>.
- Ito, M., Baumer, S.,ittanti, M., Boyd, D., Cody, R., Herr Stephenson, B., Tripp, L., 2009. *Hanging Out, Messing Around, Geeking Out: Living and Learning with New Media*. MIT Press, Cambridge, MA.
- Janks, H., 2019. Critical literacy and the importance of reading with and against a text. *J. Adolesc. Adult Lit.* 62 (5), 561–564.
- Jenkins, H., 2006. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. White paper for the Mac Arthur Foundation. Retrieved from: www.digitallearningmacfound.org.
- Jocius, R., Shealy, S., 2018. Critical book clubs: re-imagining literature reading and response. *Read. Teach.* 71 (6), 691–702.
- Kintsch, W., 1998. *Comprehension: A Paradigm for Cognition*. Cambridge University Press, Cambridge, UK.

- Krajcik, J., Schneider, B., Miller, E., Chen, I.-C., Bradford, L., et al., 2021. Assessing the Effect of Project-Based Learning on Science Learning in Elementary Schools. Lucas Education Research: George Lucas Educational Foundation, USA.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, New York, NY.
- Kucan, L., Beck, I.L., 1997. Thinking aloud and reading comprehension research: inquiry, instruction and social interaction. *Rev. Educ. Res.* 67 (3), 271–279.
- Kuiper, E., Volman, M., 2008. The web as a source of information for K-12 education. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Erlbaum, Mahwah, NJ, pp. 267–296.
- Labbo, L.D., Reinking, D., 1999. Negotiating the multiple realities of technology in literacy research and instruction. *Read. Res. Q.* 34 (4), 478–492.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Larson, L., Dwyer, B., 2015. Digging deeper with reader response: using digital tools to support comprehension of literary texts in online learning environments. In: Rasinski, T., Pytash, K.E., Ferdig, R.E. (Eds.), *Preparing Teachers to Teach Reading Using Technology*. Solution Tree Publications, USA, pp. 121–129.
- Leu, D.J., Kinzer, C.K., Coiro, J., Cammack, D., 2004. Towards a theory of new literacies emerging from the Internet and other information and communication technologies. In: Ruddell, R.B., Unrau, N. (Eds.), *Theoretical Models and Processes of Reading*, fifth ed. DE: International Reading Association, pp. 1570–1613.
- Leu, D.J., 2000. Literacy and technology: deictic consequences for literacy education in an information age. In: Kamil, M.L., Mosenthal, P.B., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research*, vol. III. Erlbaum, New York, pp. 743–770.
- Lorenz, T., Kapella, O., 2020. Children's ICT Use and Its Impact on Family Life-Literature Review. Retrieved from: <https://www.digigen.eu/>.
- McEneaney, J.E., 2006. Agent-based literacy theory. *Read. Res. Q.* 41 (3), 352–371.
- McLaughlin, M., DeVogd, G., 2004. Critical literacy as comprehension: expanding reader response. *J. Adolesc. Adult Lit.* 48 (1), 52–62. <https://doi.org/10.1598/JAAL.48.1.5>.
- Meeker, M., 2015. Internet Trends 2015-Code Conference. Retrieved from: <http://www.kpcb.com/internet-trends>.
- Metzger, M.J., 2007. Making sense of credibility on the web: models for evaluating online information and recommendations for future research. *J. Am. Soc. Inf. Sci. Technol.* 58 (13), 2078–2091.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a new framework for teacher knowledge. *Teach. Coll. Rec.* 108 (6), 1017–1054.
- Moll, L.C., Amanti, C., Neff, D., Gonzalez, N., 1992. Funds of knowledge for teaching: using a qualitative approach to connect homes and classrooms. *Theor. Pract.* 31, 132–141.
- National Adult Literacy Agency (NALA), 2020. *Literacy for Life*. NALA Policy and Research. NALA, Dublin.
- OECD, 2015. *Students, Computers and Learning: Making the Connection*. PISA, OECD Publishing. <https://doi.org/10.1787/9789264239555-en>. Retrieved from.
- Pangrazio, L., Godhe, A.-L., Ledesma, A.G.L., 2020. What is digital literacy? A comparative review of publications across three language contexts. *E-Learn. Digit. Media* 17 (6), 442–459. <https://doi.org/10.1177/2042753020946291>.
- Paris, D., Alim, H.S. (Eds.), 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York, NY.
- Paris, S.G., Lipson, M.Y., Wixson, K.K., 1983. Becoming a strategic reader. *Contemp. Educ. Psychol.* 8, 293–316.
- Pearson, P.D., 2009. The roots of reading comprehension instruction. In: Israel, S.E., Duffy, G.G. (Eds.), *Handbook of Research on Reading Comprehension*. Routledge, New York, pp. 3–31.
- Pressley, M., Afflerbach, P., 1995. *Verbal Protocols for Reading: The Nature of Constructively Responsive Reading*. Erlbaum, Hillsdale, NJ.
- Pressley, M., 2000. What should comprehension instruction be the instruction of? In: Kamil, M.L., Mosenthal, P., Pearson, P.D., Barr, R. (Eds.), *Handbook of Reading Research*, vol. III. Erlbaum, Mahwah, NJ, pp. 545–562.
- Putman, S.M., 2014. Exploring dispositions toward online reading: analyzing the survey of online reading attitudes and behaviors. *Read. Psychol.* 35 (1), 1–31. <https://doi.org/10.1080/02702711.2012.664250>.
- Raphael, T.E., Florio-Ruane, S., George, M., Hasty, N.L., Highfield, K., 2004. *Book Club Plus! A Literacy Framework for the Primary Grades*. Small Planet Communications, Lawrence, MA.
- Reinking, D., 1998. Synthesising technological transformations of literacy in a post-typographic world. In: Reinking, D., McKenna, M., Labbo, L., Kieffer, R. (Eds.), *Handbook of Literacy and Technology: Transformations in a Post-typographic World*. Erlbaum, Mahwah, NJ, pp. xi–xxx.
- Reinking, D., 2008. Thoughts on the Lewis and Fabos article on instant messaging. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Erlbaum, Mahwah, NJ, pp. 1175–1187.
- Rose, D.H., Meyer, A., 2002. *Teaching Every Student in the Digital Age: Universal Design for Learning*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Rosenblatt, L.M., 1938. *Literature as Exploration*. Appleton-Century-Croft, New York.
- Rosenblatt, L.M., 1978. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Southern Illinois University Press, Carbondale, IL.
- Rowe, D.W., Miller, M.E., 2017. The affordances of touchscreen tablets and digital cameras as tools for young children's multimodal, multilingual composing. In: Burnett, C., Merchant, G., Simpson, A., Walsh, M. (Eds.), *The Case of the iPad: Mobile Literacies in Education*. Springer, London, England, pp. 159–178.
- Rowell, J., Kress, G., Pahl, K., Street, B., 2019. The social practice of multimodal reading: a new literacy studies–multimodal perspective on reading. In: Ruddell, R.B., Alvermann, D.E., Sailors, M., Unrau, N. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, London, pp. 514–532.
- RAND Reading Study Group (RRSG), 2002. *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. RAND, Santa Monica, CA.
- Scheerder, A., van Deursen, A., van Dijk, J., 2017. Determinants of Internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. *Telematics Inf.* 34 (8), 1607–1624. <https://doi.org/10.1016/j.tele.2017.07.007>.
- Sharma, S.E., Christ, T., 2017. Five steps towards successful culturally relevant text selection and integration. *Read. Teach.* 71 (3), 295–307.
- Simpson, A., 2014. *Responding to Literature: Talking About Books in Literature Circles*. Primary English Teaching Association Australia (PETAA), Australia.
- Singer, L.M., Alexander, P.A., 2017. Reading across mediums: effects of reading digital texts and print texts on comprehension and calibration. *J. Exp. Educ.* 85 (1), 155–172.
- Smith, B.E., Pacheco, M.B., Khorosheva, M., 2020. Emergent bilingual students and digital multimodal composition: a systematic review of research in secondary classrooms. *Read. Res. Q.* 56 (1), 33–52.
- Smith, B.E., 2018. Composing for affect, audience, and identity: toward a multidimensional understanding of adolescents' multimodal composing goals and designs. *Writ. Commun.* 35 (2), 182–214.
- Smith, B.E., 2019. Collaborative multimodal composing: tracing the unique partnerships of three pairs of adolescents composing across three digital projects. *Literacy* 53 (1), 14–21.
- Spires, H.A., Himes, M.P., Medlock Paul, C., Kerkhoff, S.N., 2019. Going global with project-based inquiry: cosmopolitan literacies in practice. *J. Adolesc. Adult Lit.* 63 (1), 51–64.
- Spiro, R.J., Coulson, R.L., Feltovich, P.J., Anderson, D.K., 2004. Cognitive flexibility theory: advanced knowledge acquisition in ill-structured domains. In: Ruddell, R.B., Unrau, N.J. (Eds.), *Theoretical Models and Processes of Reading*, fifth ed. International Reading Association, Newark, DE, pp. 640–654.
- Stanford History Education Group, 2016. *Evaluating Information: The Cornerstone of Civic Online Reasoning: Executive Summary*. Author, Stanford, CA. Retrieved from: <https://sheg.stanford.edu/upload/V3LessonPlans/Executive%20Summary%202011.21.16.pdf>.
- The National Council of Teachers of English (NCTE), 2019. *Definitions of Literacy in a Digital Age*. Position Statement. Retrieved from: <https://ncte.org/statement/nctes-definition-literacy-digital-age/>.

- The New London Group, 2000. A pedagogy of multiliteracies designing social futures. In: Cope, B., Kalantzis, M. (Eds.), *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, London, pp. 9–37.
- Turner, F., 2012. A conversation with Danah Boyd, Microsoft Research. *Televis. N. Media* 13 (2), 177–185.
- Twining, P., 2021. Making sense of young people's digital practices in informal contexts: the digital practice framework. *Br. J. Educ. Technol.* 52 (1), 461–481. <https://doi.org/10.1111/bjet.13032>.
- United Nations Children's Fund and International Telecommunication Union, 2020. How Many Children and Young People Have Internet Access At Home? Estimating Digital Connectivity During the COVID-19 Pandemic. UNICEF, New York. <https://data.unicef.org/resources/children-and-young-people-internet-access-at-home-duringcovid19/>.
- Walsh, M., 2013. Literature in a digital environment. In: McDonald, L. (Ed.), *A Literature Companion for Teachers*. Primary English Teaching Association Australia, Marrickville, NSW, pp. 181–194.
- Wood, S., Jocius, R., 2013. Combating "I hate this stupid book!": black males and critical literacy. *Read. Teach.* 66 (8), 661–669. <https://doi.org/10.1002/trtr.1177>.

Digital literacies, multimodality and the engagement of adolescents who are disengaged or disenfranchised in school

David G. O'Brien[†] and Deborah R. Dillon, University of Minnesota, Minneapolis, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	563
A brief history of work with disenfranchised adolescents	564
The Jeff Literacy Lab	564
The curriculum	565
The Reading and Writing Intervention Class	565
The curriculum	566
The Frederick Douglass High School Disciplinary Literacy Project	566
The curriculum	567
Theoretical frameworks that have informed our work	567
What we have learned over decades of work	569
Print-centrism Dwarfs multimodal textuality	569
The print text and textbook are intractable	569
The importance of engagement is neglected	569
There is hope	569
References	569

Introduction

The longer adolescents are in school, the lower their reading performance. Empirical, longitudinal data bear this out. Traditional measures of reading like the National Assessment of Educational Progress (NAEP) show that adolescents have not improved in over 50 years of U.S. NAEP assessments ([National Center for Educational Statistics \[NCES\], 2019](#)). The 2019 Reading Report Card indicates that slight improvements for eighth graders in 2017 are trending down and slight improvements in the reading of grade 12 students, evident 5 years ago, have been also trending down. Multiple data profiles from assessments such as NAEP and The Program for International Student Assessment (PISA), indicate that so-called “achievement gaps” between the top and bottom deciles and the top and bottom quartiles of the SES distribution have been impervious to national policies intended to address the achievement gap disparities ([Hanushek et al., 2020](#)). And even though SES disparities have shown smaller gaps for early adolescents, the gaps widen for high school students. This is puzzling if SES is constant. Hanushek et al. offer two explanations for the puzzle: Teacher salary compression for high school teachers that might result in poorer instruction, and elementary grade directed policies like No Child Left Behind (NCLB) that have impacted elementary instruction and neglected later adolescents. We offer a more nuanced explanation for the trend described above that also forms the rationale for our decades of research on engaging students with multimodal digital texts.

Our research, as well as that of other scholars, indicates that the longer students are in school, the more negatively they feel about texts and their ability to comprehend them ([O'Brien et al., 2007](#)). As reading ability increases for most readers during elementary grades, once they reach adolescence, levels of reading engagement decline ([Eccles et al., 1993](#); [Smith et al., 2012](#); [Snow, 2002](#)). By the time students reach middle school, they have come to dislike textbooks and assignments connected to textbooks. Students relate to the relentless barrage of bland reading assignments only through producing products tied to grades—skimming reading assignments to answer questions at the ends of chapter sections; writing short papers summarizing or explaining key concepts; defining vocabulary terms from chapter lists; or filling out study guides. Adolescents who have been most recently characterized as “struggling” with reading are even less likely to engage with school texts ([O'Brien et al., 2007](#)). And even the most competent readers are unlikely to engage deeply with texts, or even expect to understand the ideas, language, or to immerse themselves in the reading experiences that some authors might expect.

School informational textbooks lack real authors—that is, the people who write them are not trying to make connections with the readers who carry the books around in their backpacks as dead weight. These students occasionally flip through pages, or even try to read the texts with enough purpose and determination to understand what authors might have intended. But the textbook authors usually write for publishers to glean royalties, not to engage students. They often produce inconsiderate texts ([Armbruster, 2016](#))—texts that are inconsiderate of readers who are trying to make sense of them. Considerate texts, on the other hand, are texts intentionally designed to engage learners, to stimulate curiosity, and to ultimately facilitate learning. Think about how many of

[†]Deceased Author

these considerate texts you encountered through school. Likely not many that you recall as stimulating and engaging. We can think only of young adult literature in this engaging category.

The notion of text considerateness, an important concept since the early 1980s, has been replaced with the idea of text complexity (Hiebert, 2018). Whereas text considerateness posits that even competent readers cannot adequately process poorly written (inconsiderate) texts, the latter notion of text complexity holds students responsible for understanding even the most inconsiderate texts. Yet many complex texts are simply inconsiderate. Text complexity gives authors a free pass; they are not held accountable for creating texts that lack accessibility and are approached by students with trepidation. Instead, the Common Core State Standards (CCSS) (National Governors Association, 2010) put the onus of responsibility on students and teachers—teachers are charged with preparing students to read increasingly complex texts as these students progress through the grades. Some complex texts may be well written rather than inconsiderate, but they are still not accessible because of the overall readability, dense vocabulary, and the esoteric topics they target.

The CCSS subscribe to the strategies/intervention fallacy: With strategies and scaffolded instruction, all students are assumed to be capable of reading texts that become more and more difficult at each grade level. But the efficacy of strategies and interventions holds mainly for the most successful “strategic” readers. Students who are more likely to struggle with reading earlier in their academic careers become increasingly disengaged as the texts become more complex—they feel disenfranchised from school, less hopeful of ever accessing the curriculum, and more and more excluded from the academic discourses across the curriculum. Thus, reading performance is flat over decades, the achievement gaps are sustained, and policies that led to targeted interventions fail.

Even if educators could improve “achievement and proficiency” in reading (O’Brien et al., 2008), there is little effort to improve motivation and engagement with school texts. In our work over the last 30 years, we have taken the position that disengagement from school is almost totally tied to the limits of print texts and print-centric instruction (O’Brien et al., 2007, 2001). Some readers, regardless of assessments and interventions aligned with these assessments, are not successful at reading print, and the more failure they experience, the less likely they are to engage with print texts. Indeed, the relatively recent history of the study of specific learning disabilities tied to language and literacy processing, including the recent resurgence of interest in dyslexia, highlights this fact (Soriano-Ferrer and Morte-Soriano, 2017). But the assessments, the etiologies that explain struggles, and the entire discourse around success in reading, is tied to print and equating print with text.

Add to this population of “struggling” readers—who struggle with reading for a range of reasons including documented learning disabilities—the group of capable readers who are simply not motivated enough to engage with school texts. In spite of the relatively rapid proliferation of digital media, the availability of digital tools to produce multimedia products, and the global span of social networks and other ways to connect with an expansive audience of the Internet, schools have been relatively impervious to digital advances that afford students who are disenfranchised with print to explore a range of multimodal texts that would invite them into the potential new ways of learning in school.

In this article we explore—through a brief history of three key projects discussed in chronological order—the frameworks that support the curricula and instructional approaches that can re-engage disenfranchised adolescents using the affordances of digital multimodal texts. First, we ground each project in the theoretical frameworks that informed the work before presenting curriculum exemplars of how the respective frameworks were enacted. Finally, we highlight the chronology of theoretical frameworks informing our work over three decades and synthesize what we have learned from this work in terms of the redesign of curricula and instruction for inviting disengaged adolescents into the excitement of school through multimodal affordances.

A brief history of work with disenfranchised adolescents

In this section we discuss three key projects that span the early 1990s up to our most current work at the time of this writing. We present an overview of each project before describing and discussing specific curricula exemplars of the work.

The Jeff Literacy Lab

In the early 1990s, we were on the faculty of Purdue University in West Lafayette, Indiana. One of our local graduate students, Rebecca Springs, a Lafayette, Indiana Jefferson High School teacher, called us and asked if we had a way to use computers to better assess students who were “at risk” in reading. This classification of students included about 20% of the population of the school’s enrollment of about 2300 students who had tested near the lowest quartile on a state-mandated reading assessment. This phone call led to our next project—one that spanned 7 years and laid the foundation for the topic of this article. We agreed to meet with Rebecca and her colleague, Dave Stith, to discuss the issue. We met over a series of summer meetings to devise a plan to completely redesign the curriculum for struggling readers at Jeff High. A sign with the label, *Remedial Reading Lab*, complete with white lettering etched into black plastic, was a stigmatizing symbol for people entering that space in the building. The first official enactment of the new project was to cover that sign with a brightly colored and inviting message we devised: *Jeff High Literacy Lab: Using Computers and Multimedia to Engage Students in Literacy*.

We decided, over a series of lunches in various restaurants over the summer, to morph the typical print-based curricula into one based mostly on “multimedia projects.” Rebecca and Dave remained good colleagues who became friends through the duration of the project. By the end of the summer of 1992, we were ready to launch the *Jeff High Literacy Lab*, with the new laminated sign

covering the old sign on the door. We applied for financial support from a local corporate foundation to purchase new tech tools including computers, scanners, printers, and CD drives. The grant proposal was not funded, but the principal of the high school liked our plan so much that he found money in the district's technology fund to get us started. We merely had to promise to push up the reading scores on the Indiana Statewide Test for Educational Progress (ISTEP). We immediately promised and then wondered how we actually could raise standardized achievement test scores on assessments consisting of print passages with multiple choice questions, using our new multimedia curriculum.

The curriculum

We constructed a series of broad project themes that required the students to do research and create multimedia presentations within the themes to share with us, their parents, and fellow students. In the beginning, we drew upon work defining intermediality (Tucker and Dempsey, 1991)—specifically the idea that text (print) could be used in a non-linear fashion with graphics, video and audio, and in a non-linear semiotic system to allow the creation of, and critical analysis of information. Students used tools like HyperCard or SuperCard to represent ideas in a non-linear fashion. In the earliest days of the internet, students started to tediously make Web pages. We were years away from reading about and incorporating multiliteracies frameworks (e.g., The New London Group, 1996), but we were reading about media use and critical media literacy in schools (e.g., Luke and Roe, 1993). Some of the projects were collaborative and students could select partners or group members they wanted to work with. Hence, projects had multiple components where students had choice: They could pick topics within broad themes; they could pick the combination of tools they wanted to use to create media and choose the best media for representing aspects of their ideas; and they could choose who they wanted to collaborate with on projects.

In the early 1990s, the instructional practices in the *Jeff Literacy Lab* were ahead of many of the new literacies frameworks that would eventually explain why the project was effective. But the work in the *Jeff Literacy Lab* was in tune with contemporary research in the fields of motivation and achievement motivation. For example, we incorporated frameworks that supported the importance of choice (Nicholls, 1978, 1989) and establishing self-efficacy related to completing tasks and establishing the importance and value of goals (Zimmerman et al., 1992).

We have used a particular *Jeff Literacy Lab* project over the years called the *Violence Project*. This portion of the curriculum was an exemplar among *Jeff Literacy Lab* projects because of the richness of data it yielded. We describe it again here for the same reason because exemplars from various students' work illustrates both the kinds of engagement it engendered and its overall impact on students. The broad theme of that project was this: Explore the impact of media on adolescents, using a range of media to illustrate the impact.

One group of students created an audio narrative of a collage of photos of 18-year-old Pennsylvania teen Michael Shingledecker who tried to mimic in real life, what high school football players in the movie, "The Program" (Ward, 1993) did: In a dare, they laid down in the middle of a highway so traffic would pass directly over them. In the real-life enactment of this stunt, Shingledecker was killed instantly. After a number of similar incidents were reported around the country, the death-defying scene was removed from the film. In the print text that students wrote to accompany the photo collage, they explained the scene in the film, articulating the consequences of life imitating art and the power of media in promoting the real-life tragic enactments.

In order to convey the tragedy of the death of Michael Shingledecker, the students did not tell the story with print as it had been shared in newspapers. Rather, they used audio narration with images of Shingledecker's yearbook photo, next to a photograph of his friends at his graveside. In the yearbook photo he gazes out from the page to his future of infinite possibilities; in the photo next to this his high school friends hugged each other in grief at the graveside. The message of this project: media, in this case, popular films, can change your life forever.

Through the use of numerous tech tools to create projects like the *Violence Project* over a period of over 6 years, students who typically did not engage with print literacies enthusiastically created and shared multimedia projects that included print, library and internet research on topics, and writing products. The net effect, for students who attended the *Jeff Literacy Lab* on a regular basis, was that their reading performance on the state achievement test (ISTEP), went up. But most importantly, many of the students who had given up on thinking of themselves as competent readers or writers, believed that they were improving as literacy learners and looked forward to coming to school—mainly due to the projects in *Literacy Lab* (see O'Brien et al., 2001 for more information on this project).

The Reading and Writing Intervention Class

This project spanned the 2004 through 2007 academic years, commencing after we joined the faculty at the University of Minnesota, Twin Cities. Research at local junior high site was initiated by our colleague, Rick Beach, and we saw facets of the work as a logical extension to what we had done on the topic of using multimedia projects to engage disengaged adolescents in the *Jeff Literacy Lab*. There were a number of syntheses of this new project (e.g., Beach and O'Brien, 2008), but the focus on how the *Reading and Writing Intervention Class* engaged disenfranchized adolescents appears mainly in O'Brien et al. (2007).

The purpose of the study was to examine how struggling seventh and eighth grade adolescents engaged with a curriculum co-taught by two teachers with different approaches. One teacher took up traditional English Language Arts (ELA) approaches like reading and discussing young adult novels; the other teacher engaged students in new literacies practices, mostly through the use of digital tools and media projects. As a continuation of previous work (O'Brien, 1998, 2003; O'Brien et al., 2007, 2001), we wanted to know how the students engaged with the two different curricula, one organized around traditional lessons and one organized around media projects.

The key component we drew upon from previous work in the *Jeff Literacy Lab* was to study how, overall, students improved their self-efficacy and perceptions of competency after years of feeling unsuccessful and disengaged with print texts in English language arts classes. In the *Jeff Literacy Lab* project, we studied how students, typically unengaged and disenfranchized from print-centric curricula, attained personally relevant goals and developed agency and perceptions about competence as they collaborated on media projects. In previous research we studied how the use of digital media literacies in the production of multimedia projects engaged those students who were typically unengaged. In the newer work in the *Reading and Writing Intervention Class* there were two foci: (a) studying how engagement with new literacies practices (Lankshear and Knobel, 2003, 2006, O'Brien and Bauer, 2005), particularly media projects, improved adolescents' literacy performance and sense of competence; and (b) studying their engagement when they had opportunities to participate in activities that grounded them more in their sense of place and within their popular media culture (O'Brien et al., 2007).

The curriculum

As the title of the project implies, the students engaged in activities in which reading and writing were integrated both through traditional approaches, and also via new literacies practices through the use of digital tools or in relation to students' experiences in digital spaces like playing video games. The *Reading and Writing Intervention Class* was designed to support students who had been classified as struggling readers and to boost their achievement in both reading and writing so the students could pass the state high stakes assessments required to eventually graduate from high school.

On the traditional side, students read young adult novels and participated in guided reading with various strategies (e.g., mind-mapping activities). The teacher read orally while students followed and interspersed this whole-class oral reading with discussion, dramatization of stories, and questions. The focus on reading and engaging with a range of activities related to young adult novels was exclusively print based.

Within the new literacies approach (the *Intervention Class*), students participated in a range of projects. They also engaged in writing activities like digital productions of stories, producing comics with *Comic Life*, writing wikis, storyboarding, and writing and producing a radio play—complete with dramatization and sound effects. In addition, they produced PowerPoint presentations about topics such as their favorite video games, or young adult novels they had read. One of the most extensive projects was the *Sense of Place* project in which students collaborated on multiple assignments where they engaged in new literacies practices based in their multiple place-based identities associated with communities, neighborhoods, and school. This focus was taken up by Beach (e.g., see a later synthesis by Beach, 2014) and was an inquiry into engaging students based on connecting them with places and spaces of their experiences which are tied to their own and others' social and cultural constructions of "figured worlds" (Holland et al., 1998).

An example of the *Intervention Class* was a project called *Sense of Place*. This curriculum focused on students constructing a prototype of an imaginary town that they modeled after Lake Wobegon from the long-running Garrison Keillor's *A Prairie Home Companion* radio show. Students devised stories based on events that occurred in the town and publish their stories in the school district Moodle site. Another project featured students in self-selected groups as they constructed VoiceThread presentations in which they represented their favorite places in school with images and constructed narratives. They often picked the favorite place among all places they represented, and using the VoiceThread Comment tool either wrote a print entry or selected audio of video modalities to explain why they picked a particular place and how they felt about it. In studying the project one of the things we looked at was why students picked a particular modality to construct narratives about school places and how they supposed that representation of their feelings about school places would be taken up by peers.

The *Intervention Class Sense of Place* project provided evidence of how students, who are typically disenfranchized, found multimodal new literacies practices enabled by collaborative projects engaging—clearly more engaging than typical print-centric practices. Certainly, the labeling associated with terms like struggling and intervention, and the broader deficit discourse that constructs one's identity as a learner once they have been identified for special classes, had a negative impact on the students. But at the end of the *Reading and Writing Intervention Class*, students reported that they perceived that they were improving their reading and writing regardless of what print-centric state and school mandated assessments showed.

The Frederick Douglass High School Disciplinary Literacy Project

In this project, a collaborative team of six University of Minnesota faculty members and graduate research assistants worked with 5 Minneapolis high school social studies teachers to collaboratively design instructional units that integrated disciplinary literacy practices into the curriculum (Lemanski et al., 2020). The *Disciplinary Literacy Project* (DLP) started when we were contacted in the fall of 2014 by high school history teachers who stated that they needed support in helping students in their International Baccalaureate (IB) program understand the complex texts in the curriculum. Due to a strong social justice orientation at the school, the IB program, previously restricted to students who have strong academic records, was opened up to a range of students representing the demographic of the building. The school of 1100 students included 92% students of color, and 90% of students coming from poverty. Before the DLP redesign of the curriculum mostly white students and a small portion of students of color who identified as college bound were enrolled in the IB program.

After an initial meeting in the fall of 2014, a partnership was established that worked together on the DLP for the next 5 years. The partnership was designed to include bi-weekly collaborative meetings (Ortmann et al., 2015; O'Brien et al., 2017) in which we discussed general issues related to the DLP collaboration, informed by our shared readings and what we had observed in classes that would inform the redesign. After 2 years of establishing this “baseline” of the current teaching practices, student engagement and disengagement with the content, the use of texts and discourse surrounding the texts, and student motivation, we were ready to implement the new approaches.

The curriculum

Our baseline data included studying the IB Diploma Program, specifically teachers and students enrolled in the IB history courses at Frederick Douglass. This work was first comprehensively documented by Van Deventer (2018) in a doctoral dissertation based on the project. In the 2-year, *Diploma Program History, Higher Level* we started to work with two teachers and multiple classes they taught. We added one teacher who taught multiple classes of a pre-IB course titled US History. We later added an IB Global Politics class, which we will use as an exemplar below. The core team of the early work included the three teachers and one university faculty member and two graduate research assistants (GRAs) who continued to work together to meet in collaborative sessions to share ideas, to discuss classroom observations, and to plan specific modifications to units.

In the fall of 2016, we received funding to expand the DLP and plan three more years of activities. We were able to enlist the help of three more faculty, a social studies educator, a historian, and another literacy researcher, as well as fund assistantships for three GRAs. We started to ramp up the DLP curriculum re-design by including funded summer institutes to work on designing units for upcoming academic years. With the added expertise of a historian, we started to design lessons around a model we eventually termed the “uncoverage” model, based on the work of Calder (2006). Our “uncoverage” design drew upon Calder’s critique of the notion that teaching means coverage—passing on information to students, assigning textbook reading, lectures, and then assessing these as decontextualized. We started to work in collaborative “uncoverage” sessions (Lee et al., 2020a; Lemanski et al., 2020) and developed essential questions (EQs) that, if answered by students using a range of resources, would be an indicator of engagement and learning. A key part of the “uncoverage” framework was that we would only select, write, or adapt texts that would address the essential questions. And the texts that were selected would be multimodal, depending on positive affordances.

One example of how this DLP collaborative planning worked was in an IB Global Politics class. The unit topic was the war in Syria. The teacher, Dawn, due to “uncoverage” sessions planning, was more likely to only use primary source material she had tailored to be more accessible by shortening the length of the pieces, substituting in more current or less technical vocabulary, and framing the text with literacy scaffolding and historical thinking prompts (Lee et al., 2020b). In addition, also due to “uncoverage” planning, Dawn systematically drew from multimodal sources like videos, maps, diagrams, and even used a sort of kinesthetic reinforcement consisting of the use of token manipulatives representing countries and proxy groups that could be moved around on a map of Syria.

All of these components were designed into the unit because they allowed students to respond to each of three EQs. An illustrative instructional approach Dawn used was to project, using a document camera, portions of a key text and a scholarly article we wrote to be more accessible by reducing the readability by putting in more signaling and using more graphics. During these sessions, she would model for students how she found information that might answer the EQs and annotate important points. She often accomplished this goal by talking aloud to model her thinking related to disciplinary thinking categories we were working with.

How did *The Fredrick Douglas Disciplinary Literacy Project*, as exemplified by Dawn’s unit design and teaching, continue the program of research we started in the 1990s? The students in the IB program, most of whom would be disenfranchized by a typical IB curriculum, were motivated and engaged when we reconfigured the curriculum to include multimodal texts tailored to EQs, used student-selected collaborative groups, and promoted teacher scaffolding of both literacy and disciplinary thinking. This DLP, which we explain in terms of “strategic engagement,” was a main focus of Lee’s (2020) doctoral dissertation study, and was included as a component in multiple presentations we made during the duration of the project. We elaborate on how the DLP extended the previous work in the theoretical frameworks section below.

Theoretical frameworks that have informed our work

In retrospect, the *Jeff Literacy Lab* project started because of our interest in the relation between motivation and engagement and literacy, specifically the connection between ability perceptions and achievement (Anderman and Maehr, 1994; Dillon, 1989; Pintrich and Blumenfeld, 1985). From a socio-cognitive view, students identified as being “at risk”—the term of choice for struggling readers in the early 1990s—were viewed as having little chance of success from middle grades onward due to early negative perceptions about ability. The research suggested that negative perceptions of ability could be a stronger prediction of success or lack thereof than actual evidence of success (Dweck and Leggett, 1988; Nicholls, 1978, 1989). As learners progress through school, they have an increasing awareness of ability as being separate from effort; and ability is viewed as a stable factor. The result: If you do not think you have ability, and your added effort does not yield the outcomes that ability does, why keep trying on tasks on which you expect to fail?

From this perspective, disengagement can be viewed as a defensive posture, designed to avoid feelings of incompetence. *The Jeff Literacy Lab* students represented the population of students who learned early in their academic careers that even if they exerted

effort, they did not read as well as their peers. Their population was also defined as the students whose lack of success was amplified each year they were in school. By the time they entered high school and were referred to a program called *Remedial Reading*, they were more likely to view it as a stigma than a reasonable attempt to boost their achievement and perceptions of competence.

Shortly after the inception of the *Jeff Literacy Lab* project, we were influenced by the reading engagement theme of National Reading Research Centers (NRRC) (Alvermann and Guthrie, 1993). The focus of the NRRC represented a major shift away from assuming that everyone could and would read if they received adequate strategies instruction and scaffolding, to an examination of why students chose not to engage in reading even if they could “read.” The NRRC goal of improving motivation to read was key in our research that would span not only the *Jeff Literacy Lab* but other projects over the next 25 years. A publication that synthesizes our perspective on the role of motivation to engage adolescent learners has been revised over the years, and was recently updated in O’Brien et al. (2022).

At the inception of the *Jeff Literacy Lab* project in 1992, we believed that two aspects of what were then termed “multimedia oriented” projects were sufficient to instill higher motivation and, to our way of thinking from the research, more engagement in literacy for “at-risk” adolescents. These two aspects included: (a) technology novelty effect based on an assumption that students would simply be motivated to engage in literacy if it gave them access to a lab with computers and other tools that most of the school did not have; and (b) students’ beliefs that using all of these high-tech tools was better than most of the print-based tasks they had come to dislike and feel incompetent using. Some of our earliest work on using multimodal texts to engage typically disenfranchized students (O’Brien, 1998; O’Brien and Dillon, 2002) was guided by this key question: How can we disrupt the cycle of disengagement and diminishing perceptions about ability through the use of media projects? The earliest publication in which we synthesized the theoretical foundations and instructional implications of the *Jeff Literacy Lab* work appeared in O’Brien (2003), but much of the description of the project can be found in O’Brien et al. (2001).

By the time we started on the *Reading and Writing Intervention Class* research, our work drawing from social cognitive perspectives on motivation and achievement motivation was complemented by our colleague, Rick Beach’s, focus on using activity theory (Cole and Engeström, 1993). Instead of studying only how media projects contributed to the engagement of students due to the technology novelty effect and the modal affordances of using digital tools, we started to look at learners’ dimensions of practice. This perspective focused on how students used mediating artifacts within activity systems, and how the use of these artifacts made sense within the classroom community to achieve the goals of specific projects. The broader goal was the same: If students saw more value in, and connection to, their personal lives with work in school, they would be more motivated to engage in that work.

At this point we had also moved from social cognitive perspectives into sociocultural perspectives, and the basic premise that new literacies practices based in digital tools and digital spaces are shaped by, and contribute to, the social contexts and discourses in which adolescents participate (Barton et al., 2000; Gee, 1996). Adolescents have increasing opportunities to move beyond print to engage in media-rich activities within new literacies frameworks (Alvermann and Hagood, 2000; Gee, 1996; Kamil et al., 2000; Kist, 2005; Lankshear and Knobel, 2003). Using and producing media was not only an alternative to the restrictions of, and adverse reactions to print that many disengaged students experienced, but it also presented an opportunity to be part of participatory culture enabled by a wider and wider range of online communities (Jenkins, 2009), and to gather wider audiences with which to share ideas and feelings and to unleash hidden passions. By the time the *Reading and Writing Intervention Class* project in the mid 2000s was developed, our focus was on how multimediating and the process of using and producing media (Lankshear and Knobel, 2003) served to not only transform learners’ sense of competence, but as a way to make sense of the world (O’Brien, 2012; O’Brien and Ortmann, 2017; Hinchman and O’Brien, 2019).

The later interest in the connection between multimodality and learning across disciplines for us, as well as many other educators, was catalyzed by the publication of *The New London Group’s* (1996) theory of multiliteracies. In contrast to the print-centric notion of learning from textbooks mapped to curriculum guides that reproduce stilted notions of culture and learning, the New London Group steered educators away from the local to multicultural diversity and global language and discourses and toward situated learning grounded in students’ lives and experiences.

As we started the *Frederick Douglass High School Disciplinary Literacy Project* in the 2013, we developed a framework O’Brien et al. (2017) based on the notion of multimodality that moves us from writing and print to visual images—moving from the “page to the screen” (Kress, 2003): In understanding and constructing meaning, drawing from a wide range of semiotic tools, learners are more able to navigate and create understanding by using a range of modes (Bezemer and Kress, 2015). Sometimes a short, written description may engage learners; at other times, only a graphic design affords the meaning that authors intend. Sometimes a photograph might enable the emotive dimension needed to engage learners; at other times, only a video with audio narration would suffice. In the *Frederick Douglass High School Disciplinary Literacy Project* and the framework that guided it, we realized that unlike previous frameworks that relied on scaffolding students to read print, we had to move to the use of multimodal texts and multimodal notions of textuality to support students’ engagement and learning. And, as we noted, in order to decide what texts to include in planning instruction, and the modality of those texts, we drew from Shulman’s (2005) and Ciccone’s (2012) notion of signature pedagogies as applied to teacher education, with a particular focus on texts, textuality, and learning, what counts as knowledge, and approaches to critiquing and dismantling rigid disciplinary approaches to using print texts. As previously described, for particular notions of unit planning, we drew from Calder’s (2006) extension of these ideas as applied to the discipline of history.

What we have learned over decades of work

In this final section, we reflect on what we have learned. Here, we will summarize without reference back to the projects or the particular theoretical frameworks from which we have grounded the work. Our parting thoughts are framed as both reflections and conclusions.

Print-centrism Dwarfs multimodal textuality

Education will continue to move “from the page to the screen.” There are almost infinite possibilities in engaging almost all students by exploring the digital textualities—the many ways of weaving meaning of, and emotional connection to school disciplines and topics within those often mundane and boring subject areas. Over the decades, the adolescents we have worked with, within projects where teachers and students helped us redesign learning, have been engaged, creative, and have redefined their perceptions of competence. But, we have not been able to generalize these approaches and students’ experiences developed in the relatively controlled settings of our projects to broader school learning cultures and across disciplines. That said, we believe that many of the approaches and ideas we developed are transferable. Overall, if educators do not address motivation and engagement of learners and new digital textualities, we fear that some of the disengaging aspects of signature pedagogies that have prevailed over the years will persist into the foreseeable future.

The print text and textbook are intractable

In concert with the point we just made, there is little evidence that print and its typical institutionalized artifacts like textbooks and worksheets will disappear any time soon. They are fading, gradually, but often the digital multimodal replacements are not carefully considered in terms of either engagement or learning potential. This is evidenced by the rapid, often unreflective movement of school curricula online during the COVID 19 pandemic (initially occurring in March 2020). Literacy itself is still typically compartmentalized as a set of skills rather than situated practices, many of which are related to structures and modes of learning. The notions of situated multiliteracies is conceptually challenging for teachers trying to use it to design units and lessons tied to broader curricula goals.

The importance of engagement is neglected

To date, our work and related work of other scholars appears to have had little impact on what is defined as engagement and learning in school, particularly as it is tied to literacies. The Institution of the School and the technocratic rationality that shores up dominant notions of assessment, achievement, and mapping and meeting standards, seldom draws from research on engagement related to the artifacts and learning processes of adolescents.

There is hope

Our recent experience with COVID has reshaped our thinking. Districts and schools in the US and many other countries have taken notice of the increased levels of disengagement among many students moved out of face-to-face environments and daily interactions with peers and teachers. But the more positive complement to this disengagement is that educators have been forced to pay more attention to ways to engage students with digital tools, multimodal texts, and creative ways of building online learning communities. More parents are also aware of what disengagement looks like in their own children, and how various teachers, schools, and districts have explored digital multimodal ways to keep students engaged online. Time will tell, but this experience of forced online, digital learning, particularly how it can be used to engage learners, may persist in hybrid future notions of schooling.

In this article, we explored some broader issues of student disengagement tied to institutionalized assessment and instructional practices, print-centrism articulated in textbooks, and policies that define success or failure in school. We described three projects that we constructed to address disengagement through the intersection of digital literacies, multimodality, and disciplinary learning, woven together with practices used to engage learners. Through this chronology of work, we reflected briefly on what we have learned, and offered some conclusions and possible predictions of where this work is headed. As many others build on our scholarly efforts, and as teaching and learning become more coherently tied to multimodal textuality, we hope that this brief chronology helps ground future progress in welcoming disenfranchized learners into the excitement of schooling.

References

- Alvermann, D., Guthrie, J., 1993. Themes and Directions of the National Reading Research Center: Perspectives in Reading Research. National Reading Research Center, Athens GA.
- Alvermann, D., Hagood, M., 2000. Critical media literacy: research, theory, and practice in “New Times”. *J. Educ. Res.* 93 (3), 193–205.
- Anderman, E., Maehr, M., 1994. Motivation and schooling in the middle grades. *Rev. Educ. Res.* 64 (2), 287–309.
- Armbruster, B., 2016. Matching readers and texts: the continuing quest. In: Lapp, D., Flood, J., Farnan, N. (Eds.), *Content Area Reading and Learning*, first ed. Routledge, pp. 47–64.
- Barton, D., Hamilton, M., Ivanic, R. (Eds.), 2000. *Situated Literacies: Reading and Writing in Context*. Routledge.

- Beach, R., O'Brien, D., 2008. Popular culture texts in the classroom. In: Leu, D., Coiro, J., Knobel, M., Lankshear, C. (Eds.), *Handbook of Research on New Literacies*. Lawrence Erlbaum, pp. 775–804.
- Beach, R., 2014. Engaging students through place-based education. *Ubiquity* 1 (1), 122–137.
- Bezemer, J., Kress, G., 2015. *Multimodality, Learning and Communication*. Routledge.
- Calder, L., 2006. Uncoverage: toward a signature pedagogy for the history survey. *J. Am. Hist.* 92 (4), 1358–1370.
- Ciccone, A., 2012. *Exploring More Signature Pedagogies: Approaches to Teaching Disciplinary Habits of Mind*. Stylus.
- Cole, M., Engeström, Y., 1993. A cultural-historical approach to distributed cognition. In: Solomon, G. (Ed.), *Distributed Cognitions: Psychological and Educational Considerations*. Cambridge University Press, pp. 1–46.
- Dillon, D.R., 1989. Showing them that I want them to learn and that I care about who they are: a microethnography of the social organization of a secondary low-track English-reading classroom. *Am. Educ. Res. J.* 26 (2), 227–259.
- Dweck, C., Leggett, E., 1988. A social cognitive approach to motivation and personality. *Psychol. Rev.* 95 (2), 256–273.
- Eccles, J., Wigfield, A., Harold, R., Blumenfeld, P., 1993. Age and gender differences in children's self- and task perceptions during elementary school. *Child Dev.* 64 (3), 830–847.
- Gee, J., 1996. *Social Linguistics and Literacies: Ideologies in Discourse*. Routledge/Farmer.
- Hanushek, E., Peterson, P., Talpey, L., Woessmann, L., 2020. Long-Run Trends in the U.S. SES-Achievement Gap. NBER Working Paper No. 26764. National Bureau of Economic Research. Retrieved March 30, 2023, from <http://hanushek.stanford.edu/publications/long-run-trends-us-ses-achievement-gap>.
- Hiebert, E., 2018. *Text Complexity Systems: A Teacher's Toolkit*. Text Project and University of California.
- Hinchman, K.A., O'Brien, D.G., 2019. Disciplinary literacy: from infusion to hybridity. *J. Lit. Res.* 51 (4), 525–536.
- Holland, D., Lachicotte, W., Skinner, D., Cain, C. (Eds.), 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press.
- Jenkins, H., 2009. *Confronting the Challenges of Participatory Culture*. The MIT Press.
- Kamil, M., Intrator, S., Kim, H., 2000. The effects of other technologies on literacy and literacy learning. In: Kamil, M., Rosenthal, P., Pearson, P., Barr, R. (Eds.), *Handbook of Reading Research*, third ed. Erlbaum, pp. 771–788.
- Kist, W., 2005. *New Literacies in Action*. Teachers College Press.
- Kress, G., 2003. *Literacy in the New Media Age*. Routledge.
- Lankshear, C., Knobel, M., 2003. *New Literacies: Changing Knowledge and Classroom Learning*. Open University Press.
- Lankshear, C., Knobel, M., 2006. *New Literacies: Everyday Practices and Classroom Learning*, second ed. Open University Press.
- Lee, Y., Lemanski, L., O'Brien, D.G., 2020a. Disciplinary Literacy: A Study of Students' Discipline-Specific Motivation [Paper presentation]. Annual Meeting, American Educational Research Association, San Francisco, CA (Conference online due to COVID-19).
- Lee, Y., Lemanski, L., Van Deventer, M., O'Brien, D.G., 2020b. Leveraging collaborative expertise: social studies teachers' perspectives of disciplinary literacy instruction. *Lit. Res. Instr.* 59 (1), 1–22.
- Lee, Y., 2020. *Exploring Reading Motivation and Engagement in Discipline-Specific Classes*. Doctoral dissertation. The University of Minnesota. <https://conservancy.umn.edu/handle/11299/218702>.
- Lemanski, L., Lee, Y., O'Brien, D.G., Dillon, D.R., Poch, R., 2020. Collaborative Expertise: Literacy in the Disciplines From the Inside Out [Paper presentation]. Annual Meeting, American Educational Research Association, San Francisco, CA (Conference Online due to COVID-19). <http://tinyurl.com/vq9w5vg>.
- Luke, C., Roe, K., 1993. Introduction to special issue: media and popular cultural studies in the classroom. *Aust. J. Educ.* 37 (2), 115–118.
- National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010. *Common Core State Standards*. <http://www.corestandards.org/about-the-standards/>.
- Nicholls, J., 1978. The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Dev.* 49 (3), 800–814.
- Nicholls, J., 1989. *The Competitive Ethos and Democratic Education*. Harvard University Press.
- O'Brien, D.G., Bauer, E., 2005. New literacies and the institution of old learning. *Read. Res. Q.* 40 (1), 120–131.
- O'Brien, D.G., Dillon, D.R., 2002. *Motivation* [Unpublished professional development materials, Minnesota Reading Excellence Act]. College of Education and Human Development, University of Minnesota.
- O'Brien, D., Ortmann, L., 2017. Disciplinary literacy: a multidisciplinary synthesis. In: Hinchman, K.A., Appleman, D. (Eds.), *Adolescent Literacy: A Handbook of Practice-Based Research*. Guilford, pp. 182–198.
- O'Brien, D.G., Springs, R., Stith, D., 2001. Engaging at-risk students: literacy learning in a high school literacy lab. In: Moje, E., O'Brien, D. (Eds.), *Constructions of Literacy: Studies of Teaching and Learning in and Out of Secondary Schools*. Lawrence Erlbaum, pp. 105–123.
- O'Brien, D.G., Beach, R., Scharber, C., 2007. "Struggling" middle schoolers: engagement and literate competence in a reading writing intervention class. *Read. Psychol.* 28 (1), 51–73.
- O'Brien, D.G., Stewart, R., Beach, R., 2008. Proficient reading in school: traditional paradigms and new textual landscapes. In: Christenbury, L., Bome, R., Smagorinsky, P. (Eds.), *Handbook of Adolescent Literacy Research*. Guilford Press, pp. 80–97.
- O'Brien, D.G., Dillon, D.R., Avery, P.G., Poch, R., Ortmann, L.L., VanDeventer, M.M., Hylton, R., 2017. A Disciplinary Literacy Framework Fostering Equity, Access, and Agency in a History International Baccalaureate Curriculum [Conference paper presentation]. Annual Meeting, American Educational Research Association, San Antonio, TX.
- O'Brien, D.G., Dillon, D.R., Lee, Y., Norton, C., 2022. The role of motivation in engaged reading of adolescents. In: Hinchman, K., Sheridan-Thomas, H. (Eds.), *Best Practices in Adolescent Literacy Instruction*, third ed. Guilford, pp. 40–61.
- O'Brien, D.G., 1998. Multiple literacies in a high school program for "at-risk" adolescents. In: Alvermann, D. (Ed.), *Reconceptualizing the Literacies in Adolescents' Lives*. Erlbaum, pp. 27–49.
- O'Brien, D.G., 2003. Juxtaposing traditional and intermedial literacies to redefine the competence of struggling adolescents. *Read. Online* 6 (7) [Website no longer available]. <http://www.readingonline.org/newliteracies/obrien2/>.
- O'Brien, D.G., 2012. "Struggling" adolescents' engagement in multimediating: countering the institutional construction of incompetence. In: Alvermann, D.E., Hinchman, K.A. (Eds.), *Reconceptualizing the Literacies in Adolescents' Lives*, third ed. Routledge, pp. 71–91.
- Ortmann, L., O'Brien, D.G., Van Deventer, M., 2015. The Role of Disciplinary Literacies to Engage Nontraditional IB History Students [Conference paper presentation]. Annual Meeting, Literacy Research Association, Carlsbad, CA.
- Pintrich, P., Blumenfeld, P., 1985. Classroom experience and children's self-perceptions of ability, effort, and conduct. *J. Educ. Psychol.* 77 (6), 646–657.
- Shulman, L., 2005. Signature pedagogies in the professions. *Daedalus* 134 (3), 52–59.
- Smith, J., Smith, L., Gilmore, A., Jameson, M., 2012. Students' self-perception of reading ability, enjoyment of reading and reading achievement. *Learn. Indiv. Differ.* 22 (2), 202–206.
- Snow, C., 2002. *Reading for Understanding: Toward a RAND Program in Reading Comprehension*. Rand Corporation.
- Soriano-Ferrer, M., Morte-Soriano, M., 2017. Teacher perceptions of reading motivation in children with developmental dyslexia and average readers. *Procedia* 237, 50–56.
- The National Center for Educational Statistics, 2019. *The Nation's Report Card: Reading*. <https://nces.ed.gov/nationsreportcard/reading/>.
- The New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–93.

- Tucker, S., Dempsey, J., 1991. Semiotic Criteria for Evaluating Instructional Hypermedia [Conference paper presentation]. Annual Meeting, American Educational Research Association, Chicago, Illinois. <https://files.eric.ed.gov/fulltext/ED337155.pdf>.
- Van Deventer, M., 2018. A Design-Based Research Intervention on Motivating Teachers to Feel Capable of Designing and Implementing Effective Disciplinary Literacy Instruction. [Doctoral dissertation, The University of Minnesota]. <https://conservancy.umn.edu/handle/11299/215213>.
- (Director) Ward, D.S., 1993. The Program [Film]. Touchstone Pictures.
- Zimmerman, B., Bandura, A., Martinez-Pons, M., 1992. Self-Motivation for academic attainment: the role of self-efficacy, beliefs and personal goal setting. *Am. Educ. Res. J.* 29 (3), 663–676.

Digital literacies & multimodal learning: writing in digital formats

Blaine E. Smith, College of Education, University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Prominent theoretical frameworks in multimodal composing research	572
Sociocultural theories of multimodal composing	572
Multiliteracies	573
Social semiotics	573
Research on adolescents and multimodal composition	573
Agentive identity expression	574
Reshaping classroom spaces	574
Developing as multimodal designers	575
Multimodal composing-to-learn	575
Research on multimodal composing processes	576
Collaborative multimodal composing	576
Mediated by composing tools	576
Recursive process and modal preferences	577
Affect-driven	577
New areas of exploration	578
Multifaceted view of multimodal composing processes	578
Multimodal meaning-making	578
Integration into schools	578
Multimodal methods	579
Critical stance toward multimodal composing	579
Conclusion	579
References	579

Writing today is often multimodal, involving the fluid interweaving of visuals, sounds, movements, and texts in digital formats, such as videos, webpages, blogs, social media posts, and podcasts (to name a few). Over the past two and a half decades, researchers have begun to examine how an expanded view of writing across multiple modes and in digital environments affects youths' communication and learning in a variety of contexts. This growing body of research has explored how youth multimodally compose online for their own purposes outside of school (Black, 2009; Hull et al., 2010; Kim, 2018; Lam, 2009) and how it can be effectively integrated into the classroom (Cimasko and Shin, 2017; Dalton, 2020; Jocius, 2013; Miller and McVee, 2013).

This chapter surveys the landscape of research focused on adolescents and multimodal composing. First, I summarize three theoretical perspectives that have been influential in guiding the field thus far. Next, I describe four prominent themes in the extant research, particularly how multimodal composing fosters (a) agentive identity expression, (b) the reshaping of classroom spaces, (c) the development of multimodal designers, and (d) multimodal composing-to-learn. I then synthesize the research by examining youths' processes when composing with multiple modes, including the collaborative, mediated, recursive, and affective nature of digital writing. This chapter concludes with a discussion of new directions for research on youth and digital multimodal composing.

Prominent theoretical frameworks in multimodal composing research

Three primary theoretical perspectives have influenced research on youth and multimodal composing: sociocultural theories of composition (Gee, 2012; Prior, 2006), multiliteracies (New London Group, 1996), and social semiotics (Jewitt, 2009; Kress, 2003, 2010). In the following, each of these interrelated perspectives is briefly described with examples of how they have been employed in multimodal composing studies.

Sociocultural theories of multimodal composing

Multimodal composing research has been guided by sociocultural theories of literacy, particularly New Literacy Studies, which recognizes that meaning-making is always positioned within specific social contexts and that it is these contexts that give meaning to literacy practices (Barton and Hamilton, 1998; Gee, 2012). Moving away from "autonomous" models that view literacy purely as a cognitive activity, emphasis is placed on literacy within contexts that reflect particular worldviews and accepted practices (Heath, 1983; Scollon and Scollon, 1981; Street, 1984). As described by Gee (2007), "Literacy has no effects—indeed, no meaning—apart from particular cultural contexts in which it is used, and it has different effects in different contexts" (p. 82).

A sociocultural view of multimodal composing emphasizes how the process is influenced by a constellation of social, cultural, material, historical, technological, and individualized factors (Prior, 2006; Vygotsky, 1978). Foundational to this perspective is the understanding of meaning-making as mediated action between composers and their cultural tools (Wertsch, 1998). Multimodal research has employed this perspective to understand how tools—both material and intellectual (Vygotsky, 1978)—are leveraged and negotiated by students during collaborative composing (Bruce, 2009; Gilje, 2011; Ranker, 2008). Additionally, some studies explore how composing tools and modes are imbued with certain constraints and affordances that affect students' processes, collaborations, and meaning-making (Jocius, 2013; Leander and Lovvorn, 2006; Smith, 2017b).

Other research employing a sociocultural view has examined how multiple modes serve as conduits through which youth agentively represent themselves (Hull and Nelson, 2005; Lewis and Fabos, 2005) and participate in social action (de los Ríos, 2017; Prior, 2006; Vasudevan et al., 2010). As described in more depth later, a large body of this work examines how youths' identities are sedimented (Rowse and Pahl, 2007) through their layering of modes—expressing personal goals, interests, and cultural and linguistic backgrounds (Alvermann, 2008; Kim, 2018; Pacheco and Smith, 2015; Rowe and Miller, 2016).

Multiliteracies

Vital to a multiliteracies framework (Cope and Kalantzis, 2000; New London Group, 1996) is the understanding that all meaning-making is multimodal, including linguistic, visual, audio, gestural, and spatial elements, which are in dynamic interaction with one another during communication. In high-tech, globalized, and culturally diverse workplaces, the New London Group (1996) emphasized how a multiliteracies approach aids students in dealing with modern literacy demands, fostering, in particular, critical engagement and empowering meaning-making. The multiliteracies emphasis on the saliency of cultural and linguistic diversity involves providing opportunities for composers to connect to their unique lifeworlds—finding their own voices and leveraging their cultural knowledge and experiences.

Multiliteracies pedagogy (New London Group, 1996) emphasizes how *design* is a key process for multimodal meaning-making, particularly the cyclical relationship among Available Designs, Designing, and The Redesigned. While Designing, composers make new use of old materials (Available Designs) to create The Redesigned. Based on this framework, numerous studies have examined youths' purposeful use of modes when designing for different purposes and audiences (Chandler-Olcott and Mahar, 2003; Dalton, 2020; Mills, 2010; Thibaut and Curwood, 2018). Other studies have drawn from youths' perspectives to understand how their goals and creative visions are achieved in their modal designs (Dalton et al., 2015; Smith, 2018).

A multiliteracies framework has also been employed in research examining how to scaffold students' multimodal composing processes in the classroom (Ajayi, 2015; Callow, 2006; Mills, 2006). For example, Tan and Guo (2009) illustrated how a teacher employed a multiliteracies pedagogy by providing students with overt instruction, critical framing, and situated practice (New London Group, 1996). Other studies have also examined how students can learn the requisite *metalanguages* to effectively communicate through multiple modes (Mills, 2009; Shanahan, 2013).

Social semiotics

A social semiotics (Jewitt, 2009; Kress, 2003, 2010) framework also emphasizes how various modes are integral to communication. Modes are sets of socially and culturally shaped resources for making meaning. They are not fixed or universal, but fluid and created through social processes within specific communities (Kress and van Leeuwen, 2001). This perspective elucidates how meaning is created through the synergistic interaction between different modes. Linguistic modes are no longer privileged in communication but, instead, are one part of a multimodal palette of semiotic possibilities for composers (Kress, 2010). The unique interweaving of modes creates generative messages that no single mode can express independently (Jewitt, 2009; Stein, 2008). Thus, endless combinations and intersemiotic relations can be crafted by combining different modes (Unsworth, 2008).

Social semiotics also elucidates how modes are shaped by sociocultural factors that influence how they are employed in communication. A mode carries with it specific communicative affordances, histories, and possibilities for constructing meaning, which also interact with and contribute to the constructed multimodal message (van Leeuwen, 2004). Multimodality studies have examined how youth draw on the affordances of specific modes (e.g., visuals, sounds) in their digital products (Dalton et al., 2015; Hull and Nelson, 2005; Jocius, 2013; Shanahan, 2013). For example, research has illustrated how students explain the affordances of nonlinguistic modes for conveying emotions or abstract concepts (de los Ríos, 2018b; Skerrett, 2018), whereas other students prefer the stability and linearity of writing for expressing ideas (Smith, 2017b). Social semiotics has also been instrumental in studies examining how youths orchestrate modes for varying purposes and audiences (Ho et al., 2011; Lenters and Winters, 2013; Toohey et al., 2012).

Research on adolescents and multimodal composition

In the following, I describe prominent themes in the research on multimodal composing and adolescents, including (a) agentive identity expression, (b) the reshaping of classroom spaces, (c) the development of multimodal designers, and (d) multimodal composing-to-learn.

Agentive identity expression

The majority of studies on multimodal composing illustrate how youth explore and express aspects of their identities through digital projects. By having expanded opportunities to convey ideas with multiple modes, students are engaged during the meaning-making process and often use their compositions as a means to agentively (re)present themselves and their cultures.

Numerous studies emphasize the ways in which composing with visuals, sounds, and texts offers adolescents *multiple points of entry* (Jewitt, 2008) into a composition for expressing, negotiating, and affirming their identities (Danzak, 2011; Goulah, 2015; Honeyford, 2013b; Hughes and Morrison, 2014). The semiotic freedom of choosing and combining different modalities creates space for students to represent aspects of their lives in individualized ways (Kim, 2018; Poveda, 2012; Wilson and Boatright, 2011; Zenkov et al., 2011). Students' purposeful infusion of their identities into their work ranges from showcasing out-of-school interests and skills (Gynne and Bagga-Gupta, 2015; Skerrett, 2012; Wiltse, 2015), to making popular culture connections (Smith, 2018; Wilson et al., 2012), and sharing their cultural histories and practices (Anderson and Macelroy, 2017; Pacheco and Smith, 2015; Pandya et al., 2015).

A growing body of research illustrates how multilingual students, in particular, employ their cultural and linguistic repertoires while orchestrating modes to feature and bridge aspects of their transnational identities (Jocson, 2012; Mills et al., 2016). Skerrett (2012, 2013) demonstrates how transnational students activate their multiliterate and multilingual capacities across lifeworlds to construct positive identities in school. Similarly, Kim (2018) describes the nuanced ways a middle school student "drew upon her entire semiotic repertoire" (p. 39) to share her transnational Korean-American identity. Her multimodal composing involved the meshing of images, music, gestures, dance, drawings, text, and speech to communicate with different Korean and American audiences. Other studies (Honeyford, 2014; Noguérón-Liu and Hogan, 2017) report how multimodal projects support students in juxtaposing, comparing, and contrasting different cultural practices from their transnational identities in empowering ways.

Research also suggests that students can agentively (re)position themselves through multimodal projects (Anderson et al., 2017; Pandya et al., 2015; Paris, 2010; Vasudevan et al., 2010). Studies show how students' impactful self-authoring can be achieved through crafting multimodal *counternarratives* (de los Ríos, 2018a; Price-Dennis and Carrion, 2017; Walsh, 2009). For example, Ajayi (2015) describes how three female students in Nigeria were able to "rewrite themselves" through their digital projects and question the social production of gender (p. 238). Cummins et al. (2015) research on identity texts with Indigenous students illustrates multiple examples of students shaping identities through collaborative multimodal projects. The authors suggest that creating identity texts through personal narratives that interweave multiple modalities "constitutes a counter-discourse that repudiates the devaluation of identity that is frequently embedded in educational structures and relationships" (Cummins et al., 2015, p. 558). When examining the multimodal practices of Black male adolescents, Vasudevan (2006) explains how "digital and visual modalities make it possible to perform and author new selves that are not only resistant to dominant images but also offer new sights of inquiry and exploration" (p. 214).

This substantial body of multimodal research underscores how the nature of multimodal composing offers new opportunities for students to strategically showcase aspects of their lives. Through the freedom of composing with multiple modes, semiotic space is opened up for students to achieve multiple co-occurring composing goals—students share ideas along with their identities.

Reshaping classroom spaces

Kress (2003) asserts that multimodal literacies can offer "an inversion in semiotic power" (p. 3) that allows composers to express themselves in innovative and empowering ways not typically afforded by traditional literacies. Research demonstrates how the integration of digital multimodal projects often reshapes classrooms by allowing students to leverage their cultural and linguistic backgrounds (Noguérón-Liu and Hogan, 2017; Vasudevan et al., 2010), reach authentic audiences (Kim, 2018; Pacheco and Smith, 2015), promote social justice (de los Ríos, 2018; Walsh, 2009), and share their out-of-school expertise (Gynne and Bagga-Gupta, 2015; Skerrett, 2012; Wiltse, 2015).

Along with multiple entry points, studies situated in schools illustrate how multimodal composition offers meaningful platforms for adolescents to engage with audiences across time and space (DeBruin-Parecki and Klein, 2003; Price-Dennis and Carrion, 2017), thus challenging rigid physical and temporal boundaries that define classrooms. Studies describe digital projects that help students to connect with their local communities and family members (Noguérón-Liu and Hogan, 2017; Pacheco et al., 2017), as well as with international audiences (Lantz-Andersson et al., 2016; Toohey et al., 2012). Pacheco and Smith (2015), for example, documented how students living in the Southeastern United States drew upon Spanish, Vietnamese, Bahdini, and English linguistic resources in an English-medium enrichment eighth grade class. The authors described how, at times, students used English to communicate with English-speaking peers, but also used non-English resources to communicate with family members not present within the classroom, thus positioning this language as a valuable semiotic resource. This work and other research (e.g., Kim, 2018) shows how the digital and mobile nature of multimodal composition can afford the use of varied multilingual resources, thus reshaping the classroom's semiotic ecology (Scollon and Scollon, 2004).

In addition to shaping what counts as legitimate composing in "traditional English literacy classrooms" (Smythe and Neufeld, 2010, p. 488), multimodal composition can reshape the classroom as a locus for social justice and activism. Research on multimodal composition suggests ways that composing can restructure classrooms as spaces for building community (Bell et al., 2011; Honeyford, 2013a; Van Sluys and Reinier, 2006), enacting social action (Albers and Frederick, 2013; de los Ríos, 2017; Price-Dennis and Carrion, 2017), and critiquing unequal social structures (Vasudevan, 2006; Walsh, 2009). For example, de los

Ríos (2018) demonstrated how Latinx bilingual adolescents created short Vine videos to post on social media that “ma[d]e meaning of power and privilege in their lives, their communities, and problematize[d] greater oppressive discourses” (p. 11).

Studies also show how traditional power hierarchies can be disrupted through the integration of technology in the classroom. Teachers often take on the role of facilitator while students collaboratively learn together. In particular, adolescent composers who possess certain technical skills are positioned as “experts” in the classroom and able to guide their peers (Oldaker, 2010; Smith, 2017b).

Developing as multimodal designers

Some studies illustrate how students develop as intentional designers through the multimodal composing process. Often with careful teacher scaffolding, students learn to leverage the unique communicative affordances of multiple modes when conveying meaning for distinct purposes and audiences. Additionally, research emphasizes how multimodal composition offers students new and different possibilities for communicating ideas.

Targeted scaffolding strategies support students in gaining understandings of multimodality, which they then, in turn, apply to their own digital projects. Through explicit instruction, students learn metalanguages (New London Group, 1996) and principles of design, along with examining how different modes can be combined for different rhetorical purposes (Ajayi, 2015; Dalton et al., 2015; Tan and Guo, 2009). Callow (2006), for example, explained how sixth grade students were explicitly taught visual concepts and grammars (Kress and van Leeuwen, 1996) that informed their own designs. Some studies emphasized the importance of students critically analyzing multimodal examples before composing (Ajayi, 2015; Callow, 2006; Mills, 2006; Ranker, 2008). These focused analyses—often involving multiple viewings of real-world and/or student-created examples—allow students to recognize and discuss the multidimensional nature of multimodal meaning-making.

Research also demonstrates how students develop a semiotic awareness of the affordances of different modalities during their composing processes (Burke and Hardware, 2015; Davis and Reed, 2003; Hepple et al., 2014; Vincent, 2007). In particular, students identified the communicative potentials of visuals, sounds, movements, and texts for communicating specific messages and observed how they could be combined for generative effects (Callow, 2006; Smith, 2018; Tan and Guo, 2009). Ho et al. (2011), for example, examined the perspectives of 158 seventh grade students in Singapore after constructing multimodal virtual museums. Students reported a growing awareness of the impact of multimodal forms of expression, including viewing visuals and color as playing significant roles in meaning-making, and an increased understanding of how modes conveyed multisensory dimensions in their virtual museums.

Some research illustrated how students articulated intentional uses of specific modes in their designs for targeted communicative effects, purposes, and audiences (Lenters and Winters, 2013; Kim, 2018; Smith, 2018; Tan and Guo, 2009; Toohey et al., 2012). Through retrospective design interviews, Dalton et al. (2015) detailed how fifth grade students who created multimodal folktale retellings expanded on their modal decisions in nuanced ways. In particular, students demonstrated a *metamodal awareness* by connecting their use of modes to story grammar elements (e.g., setting, character, action, problem, and resolution). Student explanations included labeling multimodal elements in relation to rhetorical functions (e.g., “poor music,” “happy color”) and describing how they wanted modes to “go together” in their designs (Dalton et al., 2015).

Research on multimodal composing in schools emphasizes how students feel they can communicate in new and different ways not typically afforded by traditional academic writing (Davis and Reed, 2003; Hepple et al., 2014; Vincent, 2007). In their seminal study, Hull and Nelson (2005) described how Randy a teen participating in an afterschool program, was able to construct a “different system of signification and a different kind of meaning” through his multimodal poetry. Hull and Nelson “argue[d] that the meaning that a view or listener experiences is qualitatively different, transcending what is possible via each mode separately” (2003, p. 251). These studies emphasize how multimodal composing offers students a rich communicative palette from which to share complex ideas in innovative and personalized ways.

Multimodal composing-to-learn

A growing body of recent research illustrates how students can work toward academic learning goals through multimodal composing in schools. This emerging research points to the potential of *multimodal composing-to-learn* (Smith, 2019b): that is, nonlinguistic modes serving as valuable tools for thinking (Wertsch, 1998) to support students’ literacy and content learning.

Most of the research focused on academic learning is situated in English Language Arts (ELA) classrooms where students interpret and analyze literature through digital multimodal projects (Ajayi, 2015; Davis and Reed, 2003; Lenters and Winter, 2013; Pacheco and Smith, 2015; Pacheco et al., 2017; Smith, 2017b, 2018; Tan and Guo, 2009). Smith (2019b), for example, illustrated how 10th grade culturally and linguistically diverse students analyzed literature through creating multimodal hypertext presentations, podcasts, and videos. In particular, Smith (2019b) found that students drew upon visuals and music to conceptualize literary themes as well as elucidate specific vocabulary and literary devices.

In addition to interpreting and responding to literature, some research in ELA contexts examined relationships between multimodal composing and academic writing. Many of these studies emphasize the constraints teachers face and how they attempt to infuse technology in the curriculum while concurrently designing projects with standards and state-wide assessments in mind (Goulah, 2015; Lenters and Winters, 2013; Tan and Guo, 2009). Research suggests multiple modes offer students opportunities to develop ideas and understand important aspects of writing, including narrative structures and discursive strategies (Dalton et al.,

2015; Hepple et al., 2014; Honeyford, 2013b; Pandya et al., 2015). For example, Vandommele et al. (2017) assessed the integration of multimodal composition on L2 learners' academic writing development in Belgium. Their multilevel analysis revealed that students who created a multimodal website had greater growth related to the development of writing skills compared to students who participated in their usual writing tasks. Specifically, students who composed multimodally in school showed growth with regard to six measures (complexity and text length in one task and communicative effectiveness, content, lexical diversity, and text length in the other task).

Lastly, a few studies have explored how students engage with content beyond ELA through creating digital multimodal projects (Cummins et al., 2015; Goulah, 2015; Rao, 2015; Zheng et al., 2014). Cummins et al. (2015) research on identity texts demonstrated how multilingual students could work toward the dual goals of identity expression and content learning. They provided a variety of examples to this end, ranging from junior high emergent bilingual students learning about the migration patterns of Canadian geese through designing a digital book, to students creating photo stories that focused on learning number sense problems, concepts of time, and financial literacy. There has also been a handful of studies exploring how students can problem-solve and engage with issues of climate change through creating digital multimodal compositions (Beach and Smith, 2020; Goulah, 2015; Jiang et al., 2019).

Research on multimodal composing processes

While the majority of research on youth and multimodal composing has focused on the positive outcomes previously described, less research has closely examined multimodal composing processes. These studies detail how multimodal composing can be a collaborative, mediated, recursive, and affective process for youth.

Collaborative multimodal composing

Across contexts, research portrays multimodal composing as a social and collaborative process. When creating outside of school, adolescents are involved in online participatory cultures (Jenkins, 2008) where their work is shared with receptive audiences (Ito et al., 2010). Research in schools shows that students typically collaborate in pairs or small groups—building on each other's strengths, discussing their design decisions, and providing one another with feedback (Bruce, 2009; Ho et al., 2011). Students often undertake a distinctive piece of the project on their own and then integrate their contribution into the overall composition (Goodman, 2003; Smith, 2019a; Wikan et al., 2010).

Digital literacies research also emphasizes the critical role of student interaction as a mediating factor during collaboration. For example, Wikan et al. (2010) described how the complex and multilayered nature of multimodal texts requires intensive interaction during the composing process. Smythe and Neufeld (2010) found that students were more likely to offer substantial feedback compared to more traditional writing assignments. Students questioned their peers' rhetorical moves and made their semiotic decisions transparent through discussions (Bruce, 2009). Multimodal collaborators are often exposed to their peers' alternative thinking processes, which can, in turn, allow the adaptation of their own thinking processes (Beach et al., 2016). Finally, research illustrates how interactions (both online and verbal) during multimodal composing can particularly support the composing processes of emergent bilingual students who often translanguage, discuss design decisions, and consider larger audiences (Goulah, 2015; Ho et al., 2011; Pacheco and Smith, 2015; Smith et al., 2017).

Mediated by composing tools

Research illustrates how composing tools also mediate the multimodal composing process in impactful ways (Gilje, 2011; Halverson, 2013; Smith, 2017b). The unique features of digital tools and composing programs often shape how students design. For example, Ranker (2008) found that when composing a video, fifth graders used the tool bar function in Video Studio Editor, which displayed the layers of each mode, including sounds, voiceovers, and images, to help organize their work and see the semiotic relationship between modes. Ranker concluded that the visual resources made available by the composing tool affected students' processes and allowed them to see the "interactive synergy" between modes (2008, p. 228). Similarly, Jocius (2013) examined how high school students created multimodal retellings by choosing one of two digital formats—a multimodal PowerPoint presentation or digital video. Findings revealed that the compositional tool greatly influenced the types of modes students employed to create their retelling and the various communicative weights each mode carried.

Some studies also illustrate how adolescent composers can be restricted by the technology at hand when creating their visions (Gilje, 2011; Leander and Lovvorn, 2006; Smith, 2017b). Gilje (2011) described how three high school students in Norway interacted with an editing tool while creating a short fiction film. At two different stages in the editing process, when trying to blur an image and rework a sound clip, the group could not achieve the effects they "actually wanted" (Gilje, 2011, p. 52), thus becoming frustrated and giving up on achieving their specific vision.

Recursive process and modal preferences

Research that closely examines multimodal composing illustrates it as a complex and recursive process with students skillfully traversing modes (Bruce, 2009; Smith, 2017b). Bruce (2009) found that high school students composed in a nonlinear fashion when creating digital videos. Students overlapped the videotaping and editing processes—often beginning to edit while still filming other parts of their videos. As a result of these ground-breaking findings, Bruce (2009) developed a video composition process model conceptualized as a continuum on which composers recursively shuttled back and forth between visually conceptualizing ideas they wanted to videotape and evaluating how they would go about creating those ideas.

Smith (2017b) conducted a comparative analysis with urban 12th grade students composing across three distinct multimodal projects in response to literature. Through the development of multimodal composing timescapes (Fig. 1), the findings revealed how students' rapid traversals of modes were mediated by the convergence of the digital tools, the multimodal genre, and students' modal preferences and skills. For open and flexible composing tools (e.g., PowerPoint and website programs), where modes could come into contact in endless ways, students exhibited similar recursive composing patterns and often relied on the same mode (e.g., visuals or writing) for entering the process and building new content. Additionally, students viewed their preferred modes as carrying unique affordances (Kress, 2010) for communicating their creative visions. For example, one student described preferring to use words to precisely express her ideas, whereas another student shared how "sometimes there really are no words to put into effect what I'm feeling, so pictures just really help me express myself a lot more" (Smith, 2017b, p. 272).

Similarly, Smith et al. (2017) found that emergent bilingual adolescents also demonstrated and expressed different modal preferences (i.e., visuals, music, or text) when composing. Students launched into creating their projects through a preferred mode and worked with that mode first to create a foundation before layering other modes. Students' cross-modal traversals often increased with time as they became more comfortable and skilled with multimodal composing. Furthermore, the authors found that recursion occurred on multiple interrelated levels simultaneously—across modes, across phases of the process, and across sections of students' multimodal projects.

Affect-driven

Recent research also explores how adolescents' multimodal literacy practices are driven by subconscious affective intensities (Ehret and Hollett, 2014; Lenters, 2016; Schmidt and Beucher, 2020; Wargo, 2015) and interactions with material artifacts (Zapata and Horn, 2017). This work challenges the New London Group's (1996) view of composers as intentional designers (Leander and Boldt, 2013). For example, Smith (2017a) described the "noncognitive" and affective dimensions of 18-year-old Emily's multimodal online practices as she wrote a series of young adult gothic novels. Emily's work was often "operating at the margins of her consciousness" as her narratives bled into her everyday life and she experienced affinity toward her characters (p. 13). Digital

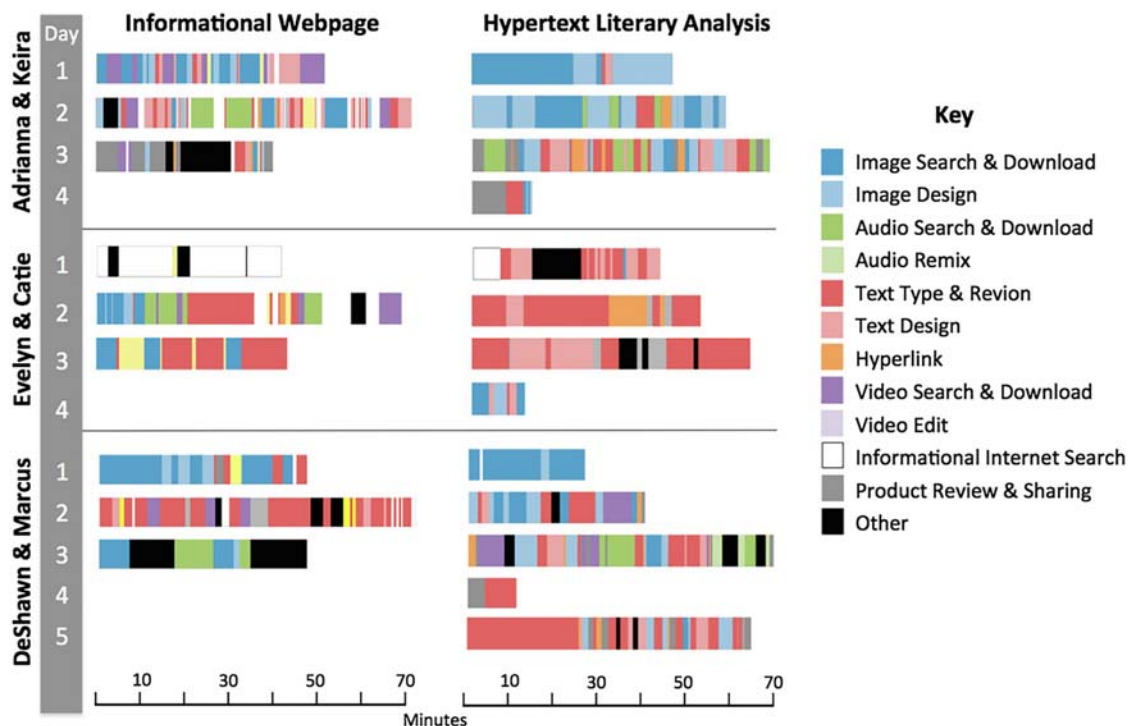


Fig. 1 Multimodal composing timescapes depict how students traverse different modes and their modal preferences when creating digital projects.

literacies research from this viewpoint, which often focuses on out-of-school contexts, complements research on multimodal composing processes by expanding the possibilities of how adolescents are inspired to create with digital tools.

New areas of exploration

Research has made important advances in terms of understanding how youth can agentively express their identities (Halverson, 2010; Skerrett, 2013), leverage their cultural and linguistic backgrounds (Honeyford, 2014; Pandya et al., 2015), and communicate in innovative ways (Hull and Nelson, 2005; Smith, 2018) through multimodal composing. A large body of studies has shown how the integration of digital writing in the classroom can disrupt engrained hierarchies (Smythe and Neufeld, 2010) and center students' collaborative meaning-making (Beach and O'Brien, 2015).

Given the current empirical landscape, in the following, I discuss new areas of exploration for research on youth and multimodal composing, including gaining a multifaceted understanding of composing processes, exploring innovative meaning-making through multiple modes, and examining the effective integration of digital tools in the classroom. Furthermore, I discuss the need for a wider variety of methodologies and a critical perspective in multimodal composing research.

Multifaceted view of multimodal composing processes

Only a handful of studies have closely examined youths' processes with regard to meaning-making across modes (Bruce, 2009; Ranker, 2008; Smith, 2017b; Smith et al., 2017), illustrating it as a collaborative, mediated, recursive, and affect-driven process. More needs to be uncovered about the processes of multimodal composition at a finer grain of detail. In particular, the following questions need to be explored: What meaning-making processes occur as students compose across modes? How do the processes of multimodal composition differ based on mode, digital format, purpose, and genre? How is multimodal composing similar to and different from print-based writing processes? What is the role of collaboration during the composing process? So far, only one model has been advanced for understanding the process of video composition (Bruce, 2009); more models need to be developed that acknowledge the great diversity in composers, contexts, and digital tools.

Additionally, research has presented sharp divisions between the portrayal of youths' design intentionality in opposition with a subconscious and emotionally-laden multimodal composing process. The field should continue working toward disrupting isolated conceptions of multimodal literacy practices to understand the complex nuances and multiplicities involved in multimodal composing. This includes understanding how multimodal composing can be shaped based on a blend of co-occurring factors. Research focused on the nuances of multimodal composing processes from multiple perspectives and grain sizes would help educators in understanding the complexity involved and how to best support students in the classroom.

Multimodal meaning-making

Research shows emerging evidence that expressing meaning through multiple modes in digital environments can mediate student learning (Cummins et al., 2015; Vandomelle et al., 2017; Zheng et al., 2014). Most of the research focused on academic learning is situated in ELA classes, demonstrating how multimodal composing can support the interpretation of literature, rhetorical elements, and academic writing. Limited studies have explored learning in other content areas. More research is needed on multimodal composing-to-learn (Smith, 2019b) and on understanding the possibilities of multiple modes for mediating and transforming academic learning (Moje, 2009; Yi, 2017). Future work might ask how the process of representing and moving meaning across different sign systems serves as a tool for thinking (see Siegel, 1995) or how digital multimodal products represent authentic disciplinary meaning-making (see Manderino and Casteck, 2016). Other valuable questions to explore could include the following: Can the knowledge and skills developed through multimodal means be traced to other contexts and genres? How is the process of meaning-making across modes different from monomodal communicative processes? Can student learning through multimodal composing be built upon and traced to other contexts and genres (e.g., academic writing)?

In agreement with others (Dalton, 2013; Grapin, 2019; Siegel, 2006), there needs to be an expanded view of meaning-making in the classroom. To that end, future research should examine the unique communicative affordances of multimodal composing for supporting students in developing and representing their ideas. Furthermore, we should reexamine the privileged position of writing in and continue to explore meaning-making and learning potential possibilities through multiple modes of expression.

Integration into schools

Future research is also needed on how multimodal projects can be effectively integrated into diverse classrooms. Some studies provide recommendations for scaffolding, including explicitly teaching metalanguages and principles of design and asking students to critically analyze multimodal examples to understand the communicative effects when modes are combined for different purposes (Ajayi, 2015; Callow, 2006; Dalton et al., 2015; Tan and Guo, 2009). Additional research should address how teachers can integrate technology to meet the needs of students, including scaffolds that foster intentional designing, collaboration, and audience awareness (Dalton, 2013; de Oliveira and Smith, 2019; Jocius, 2020). Very little research thus far has examined best practices for assessing students' multimodal compositions (cf. Anderson and Kachorsky, 2019) or the quality of students' designs.

An additional area underexplored in the literature is how to prepare teachers for these rich pedagogies (Hundley and Holbrook, 2013). More attention should be given to supporting teachers in developing an understanding of multimodal pedagogies, as well as in expanding their own skills as digital multimodal designers.

Multimodal methods

Multimodal composing studies have overwhelmingly relied on qualitative research methods, including case studies, ethnographies, and action research (see Smith, 2014; Smith et al., 2021 for reviews). These methods have been generative in terms of illuminating students' perspectives on the affordances of multimodal composing, their detailed composing processes, and design nuances in their digital products. However, with only a handful of studies employing design-based research, mixed methods, or quantitative methods, a wider variety of research methodologies are needed to gain a multidimensional understanding of youths' multimodal composing. It would also be beneficial for more research to trace youths' multimodal practices across formal and informal settings, as well as in material and virtual spaces (e.g., Ehret and Hollett, 2014; Stornaiuolo et al., 2017). Finally, multimodal methods for closely analyzing or displaying modal processes need to continue to be developed (e.g., Anderson et al., 2017; Domingo, 2012; Hull and Nelson, 2005; Kim, 2018). New methods for capturing, analyzing, and presenting research on multimodal composition need to be innovated to fully understand youths' complex multimodal composing processes and products.

Critical stance toward multimodal composing

As is evident from this review, research on multimodal composing has overwhelmingly reported the various positive aspects for students, including engagement, identity expression, and the reshaping of classroom pacing. Echoing the concerns of others (Hull, 2003; Jocius, 2018; van Leeuwen, 2015), future research must also take a critical stance toward multimodal composition and consider who participates in these practices, where the practices are enacted, and how the practices shape different learning opportunities. Only a few studies have explicitly addressed the challenges associated with implementing multimodal composing pedagogies and, similarly, the ways in which these pedagogies might constrain students' learning, collaborations, or communication (Hellmich et al., 2021; Mirra and Garcia, 2020; Paris, 2010; Ware, 2008). Mills (2006) cautioned that students might have "differential access to multiliteracies"; for instance, Anglo-Australian, middle-class students could have more opportunities to engage in productive multimodal practices than "those who were culturally or socio-economically marginalized" (p. 305). Future research is needed that critically examines how multiple factors—including the learning context and students' access to technology—might also pose barriers to empowerment and effective multimodal meaning-making.

Conclusion

Important strides have been made in the past two and a half decades toward understanding the pedagogical promise of digital multimodal composing for culturally and linguistically diverse youth. This scholarship has revealed the many unique benefits of digital writing for adolescents, including agentive identity expression, the reshaping of classroom spaces, and the development of multimodal designers. Building upon what we have learned thus far, it is important for research in this area to continue critically exploring the complex processes involved when youth compose with multiple modes. Future scholarship should also continue expanding research methods, instructional contexts, and scaffolding models in order for educators to leverage the unique meaning-making possibilities of digital multimodal composing in the classroom.

References

- Ajayi, L., 2015. Critical multimodal literacy: how Nigerian female students critique texts and reconstruct unequal social structures. *J. Literacy Res.* 47 (2), 216–244.
- Albers, P., Frederick, T., 2013. "We teach who we are": a study of two Latino transformative educators. *TESOL J.* 4 (2), 233–260.
- Alvermann, D., 2008. Why bother theorizing adolescents' online literacies for classroom practice and research? *J. Adolesc. Adult Literacy* 52 (1), 8–19.
- Anderson, K.T., Kachorsky, D., 2019. Assessing students' multimodal compositions: an analysis of the literature. *Engl. Teach. Pract. Critiq.* 18 (3), 312–334.
- Anderson, J., Macleary, V., 2017. Connecting worlds: interculturality, identity and multilingual digital stories in the making. *Lang. Intercult. Commun.* 17 (4), 494–517.
- Anderson, K.T., Stewart, O.G., Kachorsky, D., 2017. Seeing academically marginalized students' multimodal designs from a position of strength. *Writ. Commun.* 34 (2), 104–134.
- Barton, D., Hamilton, M., 1998. *Local Literacies: Reading and Writing in One Community*. Routledge, London.
- Beach, R., Newell, G.E., Vanderheide, J., 2016. A socio-cultural perspective on writing development. *Handbook of Writing Research*, 2nd ed. Guilford Press, New York, pp. 88–101.
- Beach, R., O'Brien, D., 2015. Fostering students' science inquiry through app affordances of multimodality, collaboration, interactivity, and connectivity. *Read. Writ. Q.* 31 (2), 119–134.
- Beach, R., Smith, B.E., 2020. Using digital tools for studying about and addressing climate change. In: Sullivan, P.M., Lantz, J.L., Sullivan, B.A. (Eds.), *Handbook of Research on Integrating Digital Technology With Literacy Pedagogies*. IGI Global, Hershey, PA, pp. 346–370.
- Bell, A., Ewald, M., Lynch, M.R., Zenkov, K., Fink, L., 2011. Student perspectives on quality teaching: words and images. *Voices Middle* 19 (1), 32–40.
- Black, R.W., 2009. Online fan fiction, global identities, and imagination. *Res. Teach. Engl.* 43 (4), 397–425.
- Bruce, D., 2009. Writing with visual images: examining the video composition processes of high school students. *Res. Teach. Engl.* 43 (4), 426–450.
- Burke, A., Hardware, S., 2015. Honouring ESL students' lived experiences in school learning with multiliteracies pedagogy. *Lang. Cult. Curric.* 28 (2), 143–157.
- Callow, J., 2006. Images, politics and multiliteracies: using a visual metalanguage. *Aust. J. Lang. Literacy* 29 (1), 7–23.
- Chandler-Olcott, K., Mahar, D., 2003. "Tech-savviness" meets multiliteracies: exploring adolescent girls' technology-mediated literacy practices. *Read. Res. Q.* 38 (3), 356–385.

- Cimasko, T., Shin, D.S., 2017. Multimodal semiotization and authorial agency in an L2 writing classroom. *Writ. Commun.* 34 (4), 387–413.
- Cope, B., Kalantzis, M. (Eds.), 2000. *Multiliteracies*. Routledge, London.
- Cummins, J., Hu, S., Markus, P., Kristiina Montero, M., 2015. Identity texts and academic achievement: connecting the dots in multilingual school contexts. *Tesol Q.* 49 (3), 555–581.
- Dalton, B., Robinson, K.H., Loworn, J.F., Smith, B.E., Alvey, T., Mo, E., et al., 2015. Fifth-grade students' digital retellings and the common core: modal use and design intentionality. *Elem. Sch. J.* 115 (4), 548–569.
- Dalton, B., 2013. Multimodal composition and the common core state standards. *Read. Teach.* 66 (4), 333–339.
- Dalton, B., 2020. Bringing together multimodal composition and maker education in K–8 classrooms. *Lang. Arts* 97 (3), 159–171.
- Danzak, R.L., 2011. Defining identities through multiliteracies: EL teens narrate their immigration experiences as graphic stories. *J. Adolesc. Adult Literacy* 55 (3), 187–196.
- Davis, H., Reed, Y., 2003. Assessing multimodal texts in multilingual classrooms. *Perspect. Educ.* 21 (1), 101–118.
- de los Rios, C.V., 2017. Picturing ethnic studies: photovoice and youth literacies of social action. *J. Adolesc. Adult Literacy* 61 (1), 15–24.
- de los Rios, C.V., 2018a. Bilingual Vine making: problematizing oppressive discourses in a secondary Chicana/Latina studies course. *Learn. Media Technol.* 43 (4), 359–373.
- de los Rios, C.V., 2018b. Toward a corridista consciousness: learning from one transnational youth's critical reading, writing, and performance of Mexican corridos. *Read. Res. Q.* 53 (4), 455–471.
- de Oliveira, L.C., Smith, B.E. (Eds.), 2019. *Expanding Literacy Practices Across Multiple Modes and Languages for Multilingual Students*. Information Age, Charlotte, NC.
- DeBruin-Parecki, A., Klein, H.A., 2003. Stvaranje prijatelja/making friends: multimodal literacy activities as bridges to intercultural friendship and understanding. *J. Adolesc. Adult Literacy* 46 (6), 506–513.
- Domingo, M., 2012. Linguistic layering: social language development in the context of multimodal design and digital technologies. *Learn. Media Technol.* 37 (2), 177–197.
- Ehret, C., Hollett, T., 2014. Embodied composition in real virtualities: adolescents' literacy practices and felt experiences moving with digital, mobile devices in school. *Res. Teach. Engl.* 48 (4), 428–452.
- Gee, J., 2007. *Social Linguistics and Literacies*. Routledge, London.
- Gee, J., 2012. *Social Linguistics and Literacies: Ideology in Discourses*, fourth ed. Routledge, New York.
- Gilje, Y., 2011. Working in tandem with editing tools: iterative meaning-making in filmmaking practices. *Vis. Commun.* 10 (1), 45–62.
- Goodman, S., 2003. *Teaching Youth Media: A Critical Guide to Literacy, Video Production and Social Change*. Teachers College Press, New York.
- Goulah, J., 2015. Climate change and TESOL: language, literacies, and the creation of eco-ethical consciousness. *Tesol Q.* 51 (1), 90–114.
- Grapin, S., 2019. Multimodality in the new content standards era: implications for English learners. *Tesol Q.* 53 (1), 30–55.
- Gynne, A., Bagga-Gupta, S., 2015. Language in the twenty-first century: exploring varieties and modalities in literacies inside and outside learning spaces. *Lang. Educ.* 29 (6), 509–526.
- Halverson, E.R., 2010. Film as identity exploration: a multimodal analysis of youth-produced films. *Teach. Coll. Rec.* 112 (9), 2352–2378.
- Halverson, E.R., 2013. Digital art making as a representational process. *J. Learn. Sci.* 22 (1), 121–162.
- Heath, S.B., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.
- Hellmich, E., Castek, J., Smith, B.E., Floyd, R., Wen, W., 2021. Student perspectives on multimodal composition in the L2 classroom: tensions with audience, media, learning, and sharing. *Engl. Teach. Pract. Critiq.* <https://doi.org/10.1108/ETPC-07-2020-0082>.
- Hepple, E., Sockhill, M., Tan, A., Alford, J., 2014. Multiliteracies pedagogy: creating claymations with adolescent, post-beginner English language learners. *J. Adolesc. Adult Literacy* 58 (3), 219–229.
- Ho, C.M., Nelson, M.E., Müller-Wittig, W., 2011. Design and implementation of a student-generated virtual museum in a language curriculum to enhance collaborative multimodal meaning-making. *Comput. Educ.* 57 (1), 1083–1097.
- Honeyford, M.A., 2013a. Critical projects of Latino cultural citizenship: literacy and immigrant activism. *Pedagogies* 8 (1), 60–76.
- Honeyford, M.A., 2013b. The simultaneity of experience: cultural identity, magical realism and the artefactual in digital storytelling. *Literacy* 47 (1), 17–25.
- Honeyford, M.A., 2014. From aquí and allá: symbolic convergence in the multimodal literacy practices of adolescent immigrant students. *J. Literacy Res.* 46 (2), 194–233.
- Hughes, J.M., Morrison, L., 2014. The impact of social networking and a multiliteracies pedagogy on English language learners' writer identities. *Writ. Pedagog.* 6 (3), 607–631.
- Hull, G.A., Nelson, M., 2005. Locating the semiotic power of multimodality. *Writ. Commun.* 22 (2), 224–261.
- Hull, G.A., Stornaiuolo, A., Sahni, U., 2010. Cultural citizenship and cosmopolitan practice: global youth communicate online. *Engl. Educ.* 42 (4), 331–367.
- Hull, G.A., 2003. Youth culture and digital media: new literacies for new times. *Res. Teach. Engl.* 38 (2), 229–233.
- Hundley, M., Holbrook, T., 2013. Set in stone or set in motion? Multimodal and digital writing with preservice English teachers. *J. Adolesc. Adult Literacy* 56 (6), 500–509.
- Ito, M., Horst, H., Bittanti, M., Boyd, D., Herr-Stephenson, B., Lange, P., et al., 2010. *Hanging Out, Messing Around and Geeking Out: Kids Learning and Living in New Media*. MIT Press, Cambridge, MA.
- Jenkins, H., 2008. *Convergence Culture: Where Old and New Media Collide*, updated ed. University Press, New York.
- Jewitt, C., 2008. Multimodality and literacy in school classrooms. *Rev. Res. Educ.* 32 (1), 241–267.
- Jewitt, C. (Ed.), 2009. *The Routledge Handbook of Multimodal Analysis*. Routledge, New York, NY.
- Jiang, S., Smith, B.E., Shen, J., 2019. Examining how different modes mediate adolescents' interactions during their collaborative multimodal composing processes. *Interact. Learn. Environ.* <https://doi.org/10.1080/10494820.2019.1612450>.
- Jocius, R., 2013. Exploring adolescents' multimodal responses to the Kite Runner: understanding how students use digital media for academic purposes. *J. Media Literacy Educ.* 5 (1), 310–325.
- Jocius, R., 2018. Becoming entangled: an analysis of 5th grade students collaborative multimodal composing practices. *Comput. Compos.* 47, 14–30.
- Jocius, R., 2020. The CLICK model: scaffolding multimodal composing for academic purposes. *Lang. Arts* 97 (3), 146–158.
- Jocson, K.M., 2012. Youth media as narrative assemblage: examining new literacies at an urban high school. *Pedagogies* 7 (4), 298–316.
- Kim, S., 2018. "It was kind of a given that we were all multilingual": transnational youth identity work in digital translanguaging. *Ling. Educ.* 43, 39–52.
- Kress, G., van Leeuwen, T., 1996. *Reading Images: The Grammar of Visual Design*. Routledge, London.
- Kress, G., van Leeuwen, T., 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Edward Arnold, London.
- Kress, G., 2003. *Literacy in the New Media Age*. Routledge, London, UK.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, New York.
- Lam, W.S.E., 2009. Multiliteracies on instant messaging in negotiating local, translocal, and transnational affiliations: a case of an adolescent immigrant. *Read. Res. Q.* 44 (4), 377–397.
- Lantz-Andersson, A., Vigmo, S., Bowen, R., 2016. Students' frame shifting—resonances of social media in schooling. *Learn. Media Technol.* 41 (2), 371–395.
- Leander, K.M., Boldt, G., 2013. Rereading "A Pedagogy of Multiliteracies": texts, identities, and futures. *J. Lit. Res.* 45 (1), 291–340.
- Leander, K.M., Loworn, J., 2006. Literacy networks: following the circulation of texts, bodies, and objects in the schooling and online gaming of one youth. *Cognit. Instruct.* 24 (3), 291–340.
- Lenters, K., Winters, K.-L., 2013. Fracturing writing spaces: multimodal storytelling ignites process writing. *Read. Teach.* 67 (3), 227–237.
- Lenters, K., 2016. Riding the lines and overwriting in the margins: affect and multimodal literacy practices. *J. Literacy Res.* 48 (3), 280–316.
- Lewis, C., Fabos, B., 2005. Instant messaging, literacies, and social identities. *Read. Res. Q.* 40 (4), 470–501.
- Manderino, M., Castek, J., 2016. Digital literacies for disciplinary learning: a call to action. *J. Adolesc. Adult Literacy* 60 (1), 79–81.
- Miller, S.M., McVee, M.B. (Eds.), 2013. *Multimodal Composing in Classrooms: Learning and Teaching for the Digital World*. Routledge.

- Mills, K.A., Davis-Warra, J., Sewell, M., Anderson, M., 2016. Indigenous ways with literacies: transgenerational, multimodal, placed, and collective. *Lang. Educ.* 30 (1), 1–21.
- Mills, K.A., 2006. "We've been wastin' a whole million watchin' her doin' her shoes": situated practice within a pedagogy of multiliteracies. *Aust. Educ. Res.* 33 (3), 13–32.
- Mills, K.A., 2009. Multiliteracies and a new metalanguage for the moving image. In: AARE: Australian Association of Research in Education Conference 2008.
- Mills, K.A., 2010. The Multiliteracies Classroom. *Multilingual Matters*.
- Mirra, N., Garcia, A., 2020. In search of the meaning and purpose of 21st-century literacy learning: a critical review of research and practice. *Read. Res. Q.* 56 (3), 463–496.
- Moje, E.B., 2009. Standpoints: a call for new research on new and multi-literacies. *Res. Teach. Engl.* 43 (4), 348–362.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Noguerón-Liu, S., Hogan, J.J., 2017. Remembering Michoacán: digital representations of the homeland by immigrant adults and adolescents. *Res. Teach. Engl.* 51 (3), 267–289.
- Oldaker, A., 2010. Creating video games in middle school Language Arts classroom: a narrative account. *Voices Middle* 17 (3), 19–26.
- Pacheco, M.B., Smith, B.E., 2015. Across languages, modes, and identities: bilingual adolescents' multimodal codemeshing in the literacy classroom. *Biling. Res. J.* 38 (3), 292–312.
- Pacheco, M.B., Smith, B.E., Carr, S., 2017. Connecting classrooms and communities with language and technology: a multimodal codemeshing project. *Voices From the Middle* 24 (3), 65–69.
- Pandya, J.Z., Pagdilao, K.C., Kim, A.E., Marquez, E., 2015. Transnational children orchestrating competing voices in multimodal, digital autobiographies. *Teach. Coll. Rec.* 117 (7), 1–32.
- Paris, D., 2010. Texting identities: lessons for classrooms from multiethnic youth space. *Engl. Educ.* 42 (3), 278–292.
- Poveda, D., 2012. Literacy artifacts and the semiotic landscape of a Spanish secondary school. *Read. Res. Q.* 47 (1), 61–88.
- Price-Dennis, D., Carrion, S., 2017. Leveraging digital literacies for equity and social justice. *Lang. Arts* 94 (3), 190–195.
- Prior, P., 2006. A sociocultural theory of writing. In: MacArthur, C.A., Graham, S., Fitzgerald, J. (Eds.), *Handbook of Writing Research*. Guilford Press, New York, pp. 54–66.
- Ranker, J., 2008. Composing across multiple media. *Writ. Commun.* 25 (2), 196–234.
- Rao, K., 2015. Universal design for learning and multimedia technology: supporting culturally and linguistically diverse students. *J. Educ. Multimedia Hypermedia* 24 (2), 121–137.
- Rowe, D.W., Miller, M.E., 2016. Designing for diverse classrooms: using iPads and digital cameras to compose eBooks with emergent bilingual/biliterate four-year-olds. *J. Early Child. Literacy* 16 (4), 425–472.
- Rossell, J., Pahl, K., 2007. Sedimented identities in texts: instances of practice. *Read. Res. Q.* 42 (3), 388–404.
- Schmidt, K.M., Beucher, R., 2020. Affective intensities: emotion, race, gender and the push and pull of bodies. *Engl. Teach. Pract. Critiq.* 19 (4), 403–416.
- Scollon, R., Scollon, S.W., 1981. *Narrative, Literacy, and Face in Interethnic Communication*. Ablex, Norwood NJ.
- Scollon, R., Scollon, S.W., 2004. *Nexus Analysis: Discourse and the Emerging Internet*. Routledge, London, UK.
- Shanahan, L.E., 2013. Composing "kid-friendly" multimodal text: when conversations, instruction, and signs come together. *Writ. Commun.* 30 (2), 194–227.
- Siegel, M., 1995. More than words: the generative potential of transmediation for learning. *Can. J. Educ.* 20 (4), 455–475.
- Siegel, M., 2006. Rereading the signs: multimodal transformations in the field of literacy education. *Lang. Arts* 84 (1), 65.
- Skerrett, A., 2012. Languages and literacies in translocation: experiences and perspectives of a transnational youth. *J. Literacy Res.* 44 (4), 364–395.
- Skerrett, A., 2013. Building multiliterate and multilingual writing practices and identities. *Engl. Educ.* 45 (4), 322–360.
- Skerrett, A., 2018. Learning music across transnational school settings. *J. Literacy Res.* 50 (1), 31–51.
- Smith, B.E., Pacheco, M.B., de Almeida, C.R., 2017. Multimodal codemeshing: bilingual adolescents' processes composing across modes and languages. *J. Sec Lang. Writ.* 36, 6–22.
- Smith, B.E., Pacheco, M.B., Khorosheva, M., 2021. Emergent bilingual students and digital multimodal composition: a systematic review of research in secondary classrooms. *Read. Res. Q.* 56 (1), 33–52.
- Smith, B.E., 2014. Beyond words: a review of research on adolescents and multimodal composition. In: Ferdig, R.E., Pytash, K.E. (Eds.), *Exploring Multimodal Composition and Digital Writing*. IGI Global, Hershey, PA, pp. 1–19.
- Smith, A.R., 2017a. Bare writing: comparing multiliteracies theory and nonrepresentational theory approaches to a young writer writing. *Read. Res. Q.* 52 (1), 125–140.
- Smith, B.E., 2017b. Composing across modes: a comparative analysis of adolescents' multimodal composing processes. *Learn. Media Technol.* 42 (3), 259–278.
- Smith, B.E., 2018. Composing for affect, audience, and identity: toward a multidimensional understanding of adolescents' multimodal composing goals and designs. *Writ. Commun.* 35 (2), 182–214.
- Smith, B.E., 2019a. Collaborative multimodal composing: tracing the unique partnerships of three pairs of adolescents composing across three digital projects. *Literacy* 53 (1), 14–21.
- Smith, B.E., 2019b. Mediational modalities: adolescents collaboratively interpreting literature through digital multimodal composing. *Res. Teach. Engl.* 53 (3), 197–222.
- Smythe, S., Neufeld, P., 2010. "Podcast time": negotiating digital literacies and communities of learning in a middle years ELL classroom. *J. Adolesc. Adult Literacy* 53 (6), 488–496.
- Stein, P., 2008. Multimodal instructional practices. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D. (Eds.), *Handbook of Research on New Literacies*. Lawrence Erlbaum Associates, New York, pp. 871–898.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Literacy Res.* 49 (1), 68–91.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Tan, L., Guo, L., 2009. From print to critical multimedia literacy: one teacher's foray into new literacies practices. *J. Adolesc. Adult Literacy* 53 (4), 315–324.
- Thibaut, P., Curwood, J.S., 2018. Multiliteracies in practice: integrating multimodal production across the curriculum. *Theory Into Pract.* 57 (1), 48–55.
- Toohy, K., Dagenais, D., Schulze, E., 2012. Second language learners making videos in three contexts. *Lang. Lit.* 14 (2), 75–96.
- Unsworth, L., 2008. Multiliteracies and metalanguage: describing image/text relations as a resource for negotiating multimodal texts. In: Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), *Handbook of Research on New Literacies*. Erlbaum, New York, NY, pp. 377–405.
- van Leeuwen, T., 2004. *Introducing Social Semiotics: An Introductory Textbook*. Routledge, London.
- van Leeuwen, T., 2015. Multimodality in education: some directions and some questions. *Tesol Q.* 49 (3), 582–589.
- Van Sluys, K., Reinier, R., 2006. Seeing the possibilities: learning from, with, and about multilingual classroom communities. *Language Arts* 83 (4), 321–331.
- Vandommele, G., Van den Branden, K., Van Gorp, K., De Maeyer, S., 2017. In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch. *J. Sec Lang. Writ.* 36, 23–36.
- Vasudevan, L., Schultz, K., Bateman, J., 2010. Rethinking composing in a digital age: authoring literate identities through multimodal storytelling. *Writ. Commun.* 27 (4), 442–468.
- Vasudevan, L., 2006. Making known differently: engaging visual modalities as spaces to author new selves. *E-Learn. Digital Media* 3 (2), 207–216.
- Vincent, J., 2007. Writing and coding: assisting writers to cross the modes. *Lang. Educ.* 21 (2), 141–157.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Walsh, C., 2009. The multimodal redesign of school texts. *J. Res. Read.* 32 (1), 126–136.
- Ware, P., 2008. Language learners and multimedia literacy in and after school. *Pedagogies* 3 (1), 37–51.
- Wargo, J.M., 2015. Spatial stories with nomadic narrators: affect, snapchat, and feeling embodiment in youth mobile composing. *J. Lang. Literacy Educ.* 11 (1), 47–64.
- Wertsch, J.V., 1998. *Mind as Action*. Oxford University Press, Oxford, England.
- Wikan, G., Mølster, T., Faugli, B., Hope, R., 2010. Digital multimodal texts and their role in project work: opportunities and dilemmas. *Technol. Pedagog. Educ.* 19 (2), 225–235.
- Wilson, A.A., Boatright, M.D., 2011. One adolescent's construction of native identity in school: "speaking with dance and not in words and writing". *Res. Teach. Engl.* 45 (3), 252–277.

- Wilson, A.A., Chavez, K., Anders, P.L., 2012. "From the Koran and Family Guy": expressions of identity in English learners' digital podcasts. *J. Adolesc. Adult Literacy* 55 (5), 374–384.
- Wiltse, L., 2015. Not just "sunny days": aboriginal students connect out-of-school literacy resources with school literacy practices. *Literacy* 49 (2), 60–68.
- Yi, Y., 2017. Establishing multimodal literacy research in the field of L2 writing: let's move the field forward. *J. Sec Lang. Writ.* 38, 90–91.
- Zapata, A., Van Horn, S., 2017. "Because I'm Smooth": material intra-actions and text productions among young, Latino picturebook makers. *Res. Teach. Engl.* 51 (3), 290–316.
- Zenkov, K., Harmon, J., Bell, A., Ewaida, M., Lynch, M.R., 2011. Seeing our city, students, and school: using photography to engage diverse youth with our English classes. *Engl. Educ.* 43 (4), 369–389.
- Zheng, B., Warschauer, M., Hwang, J.K., Collins, P., 2014. Laptop use, interactive science software, and science learning among at-risk students. *J. Sci. Educ. Technol.* 23 (4), 591–603.

Digital games and literacies

Jennifer Jenson and Danielle Burrell-Kim, University of British Columbia (UBC), Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	583
What literacies?	584
Enacted and emergent narrative understandings	584
Character action, design, context and game rules	584
“Transformative” literacy actions	584
Digital games as multimodal systems	584
Procedural literacy	585
School-based literacies and digital games	585
Perspectives on gaming—a positive or negative influence on learning	585
Gaming fosters complex thinking and increases motivation	586
The construction of paratexts	586
Can gaming increase literacy performance?	586
Gaming promotes the development of “soft-skills”	586
Digital literacies & digital games	587
Conclusion	587
References	588

Introduction

Digital games are, like film before them, a massive world-wide entertainment industry, one that continues to capture the attention of both youth and adults. As the strongest grossing creative medium of the 21st century (Witkowski, 2021), digital games and their adjacent media (Twitch.TV, YouTube) have been valued more for their entertainment than for their educational uses and impacts. That said, digital games have for quite some time been considered an integral part of digital literacies, included in discourses around “21st century skills and competencies” that focus on how video games provide contexts for problem solving and collaboration (Dede, 2009; Garneli et al., 2017).

Literacy has always referred to mastery over the means by which culturally significant information is coded (de Castell et al., 1981), and nowadays a great deal of culturally significant information, impacting social, economic and political practices, is coded in the form of digital games. On that argument, digital gameplay and game making merit consideration as critical components of contemporary digital literacy practices.

There are two related views of games and digital literacy that we will overview in this entry. The first argues that there are new and emerging literacies distinctive to digital games that can only be acquired through playing. The primary questions provoked by this view is what literacies are recognizably *unique* to digital games and how are these acquired informally, outside of school contexts? A related question is how to support those game-based literacy practices in school contexts. This in turn connects to the relatively conventional second view of games’ contributions to literacy, arguing that *traditional* literacy skills can be developed through playing digital games. This viewpoint provokes its own question: how does playing digital games support traditional literacy skills such as reading and writing? What “constellation of literacy practices” (Steinkuehler, 2007, pp. 299–300) does playing digital games support?

For our purposes, digital games can be separated into broad categories:

- Those built especially for educational purposes,
- “Serious games,” which can include games for health, learning games, and games whose goals and/or outcomes are defined by the player and/or game designer as “serious,”
- “AAA” games (or “commercial off the shelf games”),
- Casual, often mobile games (like Angry Birds, Candy Crush, or Plants vs. Zombies), and
- “Indie” games, meaning those designed by smaller, sometimes one-person games studios.

Many of these game types have been implemented in classroom settings, in a variety of contexts, to support educational outcomes that range from language learning and digital literacies to health and mathematics. Overall, the impacts of learning through digital games have been “inconclusive” at best (Tobias and Fletcher, 2012; Young et al., 2012), and yet digital games remain a core touchstone for discourses on 21st century learning, in and out of school.

What literacies?

Enacted and emergent narrative understandings

Part of the appeal of digital games is their profound cultural influence. Outside of classroom contexts, however, and from the standpoint of digital literacies, video games constitute a unique set of communicative media that incorporates linguistic and multimodal forms. Game-based literacies have most typically been understood through comparative analysis—what difference and similarities to print-based literacies are invoked through playing video games and interacting with their attendant paratexts and media? One direct method of comparison is to explore “reading” and “writing,” a game’s narrative, as playing games involves making choices within a set of narrative affordances both enabled and limited by the rule structures of the game. In other words, as a player progresses through the game and makes choices: through those in-game choices, players “write” their own narrative within the constraints of the game, tacking between what Jenkins (2004) has referred to as “enacted” and “emergent” narratives. For Jenkins (2004), emergent narratives are those that a player chooses as they move through the game environment, with these narratives varying depending on what the player does. Enacted narratives are those that the game provides to the player in the form of “backstories” (about characters), cut scenes, and environmental storytelling, in effect, what the game does. The point is, digital games afford complex storytelling environments that “that the player inscribes with his or her own intent” (Steinkuhler, 2010, p. 61).

Character action, design, context and game rules

Narratives and narrative understanding are but one aspect of digital game literacies. Others that have been explored include those by Apperly and Christopher (2012), include:

1. “Actions,” which “define *how* the space(s), and the objects in it, will be used by the players” (p. 118)—that is, how the character moves, what it can and cannot do, and so on;
2. “Design,” including the ability to customize playable characters (or avatars), as well as modding (i.e., modification), and programming customized functions and game mechanics;
3. The “situation” or “context” of the game—i.e., where players are playing, with whom, and using what technologies; and
4. “Systems”: a game is a system with a particular set of rules and mechanics that are learned through playing the game. The focus on procedural or system-level understanding of how games work is, relatedly, something that Ian Bogost’s work has focused on. He argues that: “Playing video games is a kind of literacy. Not the literacy that helps us read books or write term papers, but the kind of literacy that helps us make or critique the systems we live in” (Bogost, 2008, p. 136).

“Transformative” literacy actions

Similarly, but different in its emphasis on sociality through play, Tekinbas (2008) views “gaming literacy” as: (a) not only understanding systems, but how to “transform” them through, for example, modding; (b) making use of “out of game” resources, digital and non-digital (game guides, FAQs, walkthroughs); (c) understanding the social norms of play; and (d) collaboration. Each of these understandings of video games as a form of literacy emphasizes their unique properties as a representational medium distinctively different from the linear logics and command and control systems of books or film. Video games, too, are rule-bound systems, but ones which player agency is paramount, where “playing” involves degrees of choice and interactivity that readers and viewers are not afforded in other media.

Entwined in the above examples are a combination of both “reading” (playing) and “writing” (making) digital games. Playing successfully demands fluency in decoding the games narratives, characters, rule systems, operations and mechanics. Designing and making digital games demand skill in encoding, are provide ways of taking up “producer”-like relations to video games (Gee, 2004). Making games, it has been argued, supports “hands-on” learning that allows students to explore what it game making entails (Kafai and Burke, 2016) or make use of increasingly accessible game development tools that introduce them to the programming logics and paradigms, integral to computational literacies (Denner et al., 2019).

Several studies have found that programming games increases familiarity with computational concepts (Hsu and Wang, 2018; Kazimoglu et al., 2012; Lee et al., 2013; Leonard et al., 2016) and enhances problem-solving abilities as students learn to write and debug their own code (Akcaoglu, 2014; Allsop, 2016; Hamlen, 2017). Important as it is, Kafai and Burke (2016) have pointed out that game design is not just about learning programming logics, but is also about bringing together and integrating multimodal resources like images, texts, sounds, rule structures, and narratives. Andrew Burn argues, similarly, that the process of making games is both technological and conceptual: “For Vygotsky, tools were also semiotic: they involved the manipulation and creation of signs; and though his work pre-dated multimodal theory, in some ways it anticipated it” (2021, p. 67). In the next section, we overview the centrality of multimodality to understanding video games and literacies.

Digital games as multimodal systems

As Gunther Kress has persuasively argued, all communicative forms are multimodal (Kress, 2010). Video games are no different, and entwined within the acknowledgment of playing games/reading games as engaging in literacies is the view that digital games

inescapably include multiple modalities—image, text, sound, gesture, embodiment—in a variety of combinations. For example, if you are playing in a massively multiplayer online game such as *World of Warcraft*, your screen will include the game world, with its elaborate narrative backstory and whatever the avatar you control is doing. It will also display a plethora of choices available to you, of weapons, clothing, appearance, capabilities, depending on the character you are playing and what level you are in the game. You will be controlling your avatar through a keyboard or game controller, and those keystrokes and/or button pressing will result in a virtual embodiment of some kind of action. Game player action must be built up from predefined elements (walking, running, sword fight) that can be made operational only once players learn to decode, all at once, visually/acoustically conveyed environmental requirements like “*Zombies! Run!*” and encode that game-based imperative into both skillful and strategic manual operation of color-coded icons on a controller, using keypresses guided by symbols (find the images and Stick em in here) which, resonant of the alphabet, have themselves no meaning—nor does digital code. And in that kind of learning, a recognizable literacy comes into play.

As for kinds of oral communicative competencies developed through your play, you could also be either in real-time voice chat with your guild (other players who group together within the gameworld), or text chatting, or both, in your own language, or not, and the list goes on. The point is, video games, especially for highly engaged players, present a complex, changing, and multimodal environment. The semiotic score of meaning-making then happens across multiple convergent channels of information communicated during gameplay, and all of those semiotic resources combine to compose the distinctive forms of meaning expressible in and through digital games and play.

Video game literacies include digital and nondigital “meaning making that hinges on player recognition and understanding of various multimodal sign systems” (Abrams, 2016, p. 2319), and theories of multimodality can help us to better understand the complexities of digital games (Hawreliak, 2018), including what types of learning occurs through playing games.

Procedural literacy

A rhetorical analysis of videogames as a medium has been introduced by Michael Mateas and Ian Bogost, whose analytical frame for visual and multimodal modes of communication, Bogost (2008, 2011) introduces “procedurality” as a form of communication and, thus, literacy in video games. Procedural rhetoric refers to the structures and mechanisms through which games communicate and express meaning, and to the accomplishment of persuasion through the coding and process-oriented rules of the game. The required and prevented procedures as well as procedures that are rewarded or reprimanded work together to express ideas and arguments in a mode distinctive of, and arguably unique to, video games (Bogost, 2011; Hawreliak, 2018).

Hawreliak’s (2018) example is *Dys4ia*, an independent game created by Anna Anthropy (2012). *Dys4ia* is a short, proceduralist game which explores the experiences of Anna as a transwomen and her journey to begin using estrogen. Her story is told through short mini games accompanied by text. Each minigame represents some aspect of her experience such as the first one in which the player is asked to fit a block through a maze it does not fit. Next to it is the text “I feel weird about my body.” The player often faces minigames that are impossible to win and receives no instruction on how to play through each one, to better represent the road-blocks and challenges Anthropy faced, and, perhaps, evoking the sensation of at times feeling lost and helpless at the lack of guidance and help along the way during his quest to transition. At the end, the player is presented with a game of pong in which they are supposed to remove all the bricks blocking their way. However, only one layer of bricks will disappear while others remain, indexical of the seemingly irremovable barriers transwomen face in society.

Mateas (2005) uses the term “procedural literacy” as the ability to understand “how code operates as an expressive medium” (p. 102). To Mateas (2005), procedural literacy includes the ability to read and write code. For Hawreliak (2018) and Bogost (2008, 2011), procedural literacy is about understanding what the code does “as a representation of culture as much as that of dynamic systems” (Bogost, 2007, p. 252). From this perspective, players can and often do gain procedural literacy when playing, whether through repeated play or purposeful, critical reflection on a game’s procedural rhetoric and its embedded ideologies.

School-based literacies and digital games

As digital games gained popularity, scholars started to question what kind of benefits games might have in relation to school-based literacies. As illustrated above, a wide range of diverse literacies are exercised through playing digital games. However, school-based literacies, most often referring to the ability to read and write, and the applicability of using digital games in the classroom, have been of particular interest to researchers. Digital games have also been used to encourage the learning of second languages (Sauro, 2021) as well as the digital literacies which are viewed as essential for academic and career success in the 21st century (Gee, 2004; O’Mara, 2017; Rowan, 2017).

Perspectives on gaming—a positive or negative influence on learning

However, game-based learning has been a topic of a significant epistemological divide within both education and popular media. Video games are often portrayed as inherently violent, negative influences on youth (Ferguson et al., 2008; Ferguson, 2015). Identities of laziness and immaturity are often ascribed to avid gamers (Bergstrom et al., 2014), and the nature of gaming as an addiction

has been well-debated among parents, teachers, and scholars alike (Markey and Ferguson, 2017). In contrast, arguments can also be made for the positive impacts of gaming on literacy and learning, given that as players interact with games, they are required to engage in a range of literacy practices, to varying extents. They must exercise traditional, school-based literacies such as reading to navigate menus, understand tutorials and item descriptions, or interact with dialog options. Simultaneously, digital games may offer an avenue for building digital literacies such as typing, utilizing databases, and understanding common digital interface systems.

Gaming fosters complex thinking and increases motivation

Another argument related to games and literacies views digital games as a substitute for other media such as books or films in language, novel studies, or media curricula, with the underlying assumption that the innovative and interactive medium of games should increase motivation and engagement. Through this lens, digital games can be viewed as powerful communication engines, capable of helping players engage with complex ideas in increasingly unique ways. Dezuanni and O'Mara (2017) argue that games open a door to impassioned learning within and outside of the classroom.

As Gee (2004) has also argued, when players are deeply engaged in play, they gain opportunities for meaning making, identity construction, and the acquisition of technical registers, all of which qualify as “advanced” levels of literacy. Dezuanni and O'Mara (2017) provide the example of literacies gained when playing *Minecraft*, describing insider knowledge, or the ability to understand complex systems specific to the game and communicate about the game collaboratively. In this way, the literacies exercised through *Minecraft* mimic those academic literacies in which students are expected to maintain in-depth knowledge on complex fields and communicate said knowledge using subject-specific terminology.

The construction of paratexts

As players become invested in games, they may engage in creating or consuming paratexts. Paratexts refer to individual or official game-brand content surrounding a game. These paratexts include guidebooks, online forums, mods, memes, fanfiction, fan art, and much more, and engaging with paratexts is a significant part of how players participate in gaming culture and affinity groups online (Burwell and Miller, 2016; Consalvo, 2007). The creation or consumption of these paratexts often rely heavily on school-based literacies such as reading and writing, and they may be constructed independently or through collaboration or interaction with others (Apperly and Walsh, 2012). While there has been some work that argues that playing/reading a game, or exploring its related paratexts might support school-based literacy practices (Apperly and Walsh, 2012; Burwell and Miller, 2016; Sauro, 2021), there is still a paucity of empirical in-school research on the subject.

Can gaming increase literacy performance?

While paratexts may be viable contexts for literacy and language production, more debate has centered on whether playing games increases players' literacy or not, and there is little consensus among scholars regarding the results of research on games and literacy. Steinkhueler et al. (2009) explored using digital games to increase the literacy of grades 6–9 students who were found to read at a substantially lower grade level than their own. The researchers found that by engaging in MMO games, paratext utilization and creation, and an online affinity group directed by researchers, participants were able to read and comprehend texts written at a grade 12 level. Furthermore, some participants scored much higher on reading assessments and demonstrated higher levels of engagement when utilizing texts related to digital games (Steinkhueler et al., 2009). However, some participants showed a complete unwillingness to engage in classroom activities and assignments that did not focus on digital games (Steinkhueler, 2010). So, while playing commercial games may have some benefits related to school-based literacies, even when played outside of school, it is crucial to delineate the difference between school-based skills and school performance: even if playing games helps students cultivate various literacies.

Gaming promotes the development of “soft-skills”

Examples of using a commercial off-the-shelf game for subject-matter learning are studies that use games to bridge barriers, focused on the use of massively multiplayer online games (MMOGs) as environments that support second language (L2) learning. A review of literature in the area by Jabbari and Zohreh (2019) argues that MMOGs are “low language anxiety” spaces that promote communication and socialization through the games' affordances. For example, an MMOG such as World of Warcraft (WoW) has opportunities for players to interact through text chat and/or via voice chat, and game play is supported by paratexts (WoW lore and background, forums, and fan fiction) and other resources such as game walkthroughs and live streaming of game play. The review concludes that “MMOGs encourage learners to get actively involved in L2 socialization and collaborative interactions with other PCs [player characters] to perform a variety of goal-oriented tasks within and beyond game contexts” (Jabbari and Eslami, 2019, p. 106). The point is that MMOGs are being used in educational contexts globally, to support language learning, with studies taking place in such places as Turkey (Alp and Patat, 2015), South Korea (Lee and Gerber, 2013), Spain (Rama et al., 2012), and the United States (Newgarden et al., 2015) in a variety of languages.

Digital literacies & digital games

Digital games have also been used to teach digital literacies both in and outside of the classroom. Increasingly, many career paths and educational systems require high levels of digital literacy. Gee (2012) refers to “premium digital literacies” or the “ability to use specialist/technical language connected to digital tools” (p. 419). While specialized digital skills may be gained in a variety of ways, coding, as we noted above, is one digital literacy that is highly valued in many career and educational institutions. O’Mara (2017) illustrates that students can begin coding as early as grade three, engaging with complex digital literacies through the context of game making. Making digital games also requires students to engage in collaborative learning and exercise school-based literacies like reading and writing as they create narratives for their games (O’Mara, 2017; Rowan, 2017). Introducing students to coding and digital literacy at a young age is especially relevant in efforts to diversify game development, where women and trans folks make up a very small percentage of workers and an even smaller percent of upper management (Edwards et al., 2014).

In addition to using commercial games as a tool and context of learning, educational games have also been created and studied in classrooms. Educational and/or serious games are classified as those meant explicitly for imparting information (Hawreliak and Lemieux, 2020). The exact definition of a serious game has been debated. For Michael and Chen (2006) it is a game that places learning over fun, and thus a serious game may or may not inspire entertainment. While mainstream games may give way to learning and address many significant topics, the possible learning outcomes are not explicitly addressed. Serious games, however, by Dörner et al.’s (2016) definition have at least one explicit learning goal. For example, serious games are: “a digital game created with the intention to entertain and to achieve at least one additional goal (e.g., learning or health)” (Dörner et al., 2016, p. 2). Such educational games take many forms and cover a range of topics such as vocabulary, reading, math, health, and geography, and are usually subject-matter specific. Children’s games may focus on traditional, school-based literacies and beginning digital literacies such as typing. Educational games, like dissection simulators, have been created to provide interactive environments for otherwise unattainable experiences in some schools. There are even some serious games that are created independently from educational institutions yet are designed with the aim to educate. These often revolve around critical social issues like gender dysphoria in Anna Anthropy’s *Dys4ia*.

The aim of serious games is to prioritize education over fun. Where traditional games may inspire impassioned learning, they are not focused on learning or learning outcomes, but rather entertainment and sales. One benefit of serious games is their ability to offer an embodied, simulated and interactive experience, that has learning entwined with play. Where students may traditionally read or watch movies about a topic, educationally focused games give them an opportunity to actively engage through modes unavailable in other mediums. In fact, serious games may be purposely monotonous, or emulate PCs’ powerlessness to give players a glimpse at the lived realities represented within the game. Director Luc Bernard discusses the intentionality of a lack of gamification and feelings of powerlessness in the upcoming game *A light in the darkness*, that follows a family of Jewish immigrants in France during 1942. *A light in the darkness* was created as an educational game about the holocaust in France. In particular, Bernard (2021) aims to highlight the seemingly small, compounding acts which led to the mass arrest of Jewish people in France in 1942. To emulate the powerlessness that Jewish people faced, all choices are removed from the player throughout the game. Players interact within their environments between many cut scenes, fulfilling the required tasks in order to progress. As Bernard (2021) describes some of these tasks, like registering with the police as a Jewish person, require compliance without resistance to those who seek to oppress or eradicate the player characters. In this way, the aim of the game is to create an immersive experience which can help players develop a greater understanding and empathy with the many people who were victims of the holocaust. To this emotional and psychological education is superadded the historical education of highlighting many of the real experiences and perspectives of holocaust survivors (Bernard, 2021). Digital games can offer perspectives and interactions unavailable through other literacies. Any lessons imparted are likely to be amplified through being actively experienced by the player in the form of their surrogate character, rather than through the more passive reception of traditional media. This is perhaps video games’ most powerful mechanic when considered as educational tools: if recited lessons are never as impactful as those learned from events personally experienced, then no medium, no other literacy, possesses as much learning potential as video games, for no other medium places the player/reader/learner in the active and volitional role of the character to whom these in-game event lessons happen.

Conclusion

Digital games modify and change the constructs that, in other domains, have been recognized as literacies. But the risk with all literacies, digital games included, is that we take analytical constructs and transform them into categories with prescriptions attached to them (i.e., you should do this, you need to learn this through playing a game). Instead, digital games allow us new ways of viewing what count as “literacies” that might also be useful for learning, understanding, for example, how games work not only at the level of code but as personally and socially formative, and as culturally processual. In other words, digital games can help us understand, much like literature before them, something about the cultural moments we are in, and who we are in those moments, or those experiences no longer available to us except through historical accounts. This understanding can be guided by the constraints and affordances that those digitally-enabled processes impart through play.

There is yet another way of thinking about the entanglement of literacies and video games; and that is, as we have hoped to show here, to engage with video games as a literacy practice akin to such traditional practices as the analysis of literature or poetry. From this perspective, video games would be approached not as one-dimensional products for entertainment, but from the critical

viewpoint deserving of a highly relevant and widely consumed creative medium transformative both through digital play and through digital design. As Jagoda (2020) argues, video games “constitute a unique medium for thinking through and affecting our historical present from the inside” (p. 219). To appreciate video games as sites of creativity, affect, critique, deep engagement, multimodality, play, and culture in the 21st century, inside or outside school walls, is to acknowledge video games as new tools for making meaning. The moment we find ourselves in, brings with it not only new literacies that demand innovative understandings of what those literacies entail, and how they might best be cultivated, but as well new opportunities for transforming how we use digital tools, including digital games, to help creatively and responsibly evolve pedagogy and learning.

References

- Abrams, S.S., 2016. Videogaming and literacies. In: Peters, M. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, pp. 2318–2323.
- Akcaoglu, M., 2014. Learning problem-solving through making games at the game design & learning summer program. *Educ. Technol. Res. Dev.* 62 (5), 583–600.
- Allsop, Y., 2016. A reflective study into children's cognition when making computer plans. *Br. J. Educ. Technol.* 47 (4), 665–679.
- Alp, P., Patat, E., 2015. Gameplay in Italian: the impacts of World of Warcraft on second language learners in Turkey. *Int. J. Technol. Learn.* 22 (3), 1–9.
- Anthropy, A., 2012. *Dys4ia*. Newgrounds, Pennsylvania.
- Apperley, T., Walsh, C., 2012. What digital games and literacy have in common: a heuristic for understanding pupils' gaming literacy. *Literacy* 46 (3), 115–122.
- Bergstrom, K., Fisher, S., Jensen, J., 2014. Disavowing “that guy.”. *Convergence* 22 (3), 233–249.
- Bernard, L., 2021. The Light in the Darkness—An Educational Historical Experience About the Holocaust in France [online]. Available at: <https://news.xbox.com/en-us/2021/11/16/the-light-in-the-darkness-an-educational-historical-experience-about-the-holocaust-in-france/>. (Accessed 26 March 2022).
- Bogost, I., 2007. *Persuasive Games: The Expressive Power of Videogames*. The MIT Press, Cambridge, Mass.
- Bogost, I., 2008. The rhetoric of video games. In: Tekinbas, K.S. (Ed.), *The Ecology of Games: Connecting Youth, Games, and Learning*. MIT Press, Boston, Mass, pp. 117–139.
- Bogost, I., 2011. How to Do Things With Videogames. University of Minnesota Press. <https://doi.org/10.5749/j.ctttmwd>.
- Burn, A., 2021. *Literature, Videogames and Learning*. Routledge, London. <https://doi.org/10.4324/9781003025597>.
- Burwell, C., Miller, T., 2016. Let's play: exploring literacy practices in an emerging videogame paratext. *E-Learn. Digit. Media* 13 (3–4), 109–125. <https://doi.org/10.1177/2042753016677858>.
- Consalvo, M., 2007. *Cheating: Gaining Advantage in Videogames*. The MIT Press, Cambridge, Mass.
- de Castell, S., Luke, A., MacLennan, D., 1981. On defining literacy. *Can. J. Educ.* 6 (3), 7–18. <https://doi.org/10.2307/1494652>.
- Dede, C., 2009. Comparing Frameworks for “21st Century Skills”. Cambridge, MA. Available at: http://www.watertown.k12.ma.us/dept/ed_tech/research/pdf/ChrisDede.pdf.
- Denner, J., Campe, S., Werner, L., 2019. Does computer game design and programming benefit children? A meta-synthesis of research. *ACM Trans. Comput. Educ.* 19 (3). <https://doi.org/10.1145/3277565>. Article 19.
- Dezuanni, M., O'Mara, J., 2017. Impassioned learning and Minecraft. In: Beavis, C., Dezuanni, M., O'Mara, J. (Eds.), *Serious Play: Literacy, Learning and Digital Games*. Taylor and Francis, pp. 36–48.
- Dörner, R., Göbel, S., Kickmeier-Rust, M., Masuch, M., Zweig, K., 2016. *Serious Games: Foundations, Concepts, and Practice*. Springer International Publishing, Switzerland.
- Edwards, K., Weststar, J., Meloni, W., Pearce, C., Legault, M.J., 2014. Developer Satisfaction Survey. International Game Developers Association.
- Ferguson, C.J., Rueda, S.M., Cruz, A.M., Ferguson, D.E., Fritz, S., Smith, S.M., 2008. Violent video games and aggression: causal relationship or byproduct of family violence and intrinsic violence motivation? *Crim. Justice Behav.* 35 (3), 311–332. <https://doi.org/10.1177/0093854807311719>.
- Ferguson, C.J., 2015. Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. *Perspect. Psychol. Sci.* 10 (5), 646–666. <https://doi.org/10.1177/1745691615592234>.
- Garnell, V., Giannakos, M., Chorianopoulos, K., 2017. Serious games as a malleable learning medium: the effects of narrative, gameplay, and making on students' performance and attitudes. *Br. J. Educ. Technol.* 48 (3), 842, 818.
- Gee, J.P., 2004. *What Video Games Have to Teach Us About Learning and Literacy*, first ed. Palgrave Macmillan.
- Gee, J.P., 2012. The old and the new in the new digital literacies. *Educ. Forum* 76 (4), 418–420. <https://doi.org/10.1080/00131725.2012.708622>.
- Hamlen, K.R., 2017. General problem-solving styles and problem-solving approaches in video games.
- Hawreliak, J., Lemieux, A., 2020. The semiotics of social justice: a multimodal approach to examining social justice issues in videogames. *Discourse* 41 (5), 723–739. <https://doi.org/10.1080/01596306.2020.1769936>.
- Hawreliak, J., 2018. *Multimodal Semiotics and Rhetoric in Video Games*. Routledge, New York.
- Hsu, C.-C., Wang, T.-I., 2018. Applying game mechanics and student generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Comput. Educ.* 121, 73–88.
- Jabbari, N., Eslami, Z.R., 2019. Second language learning in the context of massively multiplayer online games: a scoping review. *ReCALL* 31 (1), 92–113.
- Jagoda, P., 2020. *Experimental Games*. The University of Chicago Press, Chicago.
- Jenkins, H., 2004. Game design as narrative architecture. In: Wardrip-Fruin, N., Harrigan, P. (Eds.), *First Person: New Media as Story, Performance, and Game*. The MIT Press, Cambridge, Mass, pp. 118–130.
- Kafai, Y.B., Burke, Q., 2016. *What making video games can teach us about learning and literacy*. The MIT Press, Cambridge, Mass.
- Kazimoglu, C., Kiernan, M., Bacon, L., Mackinnon, L., 2012. A serious game for developing computational thinking and learning introductory computer programming. *Procedia-Soc. Behav. Sci.* 47, 1991–1999.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, New York, NY.
- Lee, Y.J., Gerber, H., 2013. It's a WoW world: second language acquisition and massively multiplayer online gaming. *Multimed. Assist. Lang. Learn.* 16 (2), 53–70.
- Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O.S., Hubert, T., Almughayrah, S., 2016. Using robotics and game design to enhance children's self-efficacy, STEM attitudes, and computational thinking skills. *Journal of Science Education and Technology*, 25 (6), 860–876. <https://doi.org/10.1007/s10956-016-9628-2>.
- Markey, P.M., Ferguson, C.J., 2017. Internet gaming addiction: disorder or moral panic? *Am. J. Psychiatr.* 174 (3), 195–196. <https://doi.org/10.1176/appi.ajp.2016.16121341>.
- Mateas, M., 2005. Procedural literacy: educating the new media practitioner. *Horizon* 13 (2), 101–111. <https://doi.org/10.1108/10748120510608133>.
- Michael, D., Chen, S., 2006. *Serious games: Games that educate, train, and inform*. Thomson Course Technology PTR, Boston, MA.
- Newgarden, K., Zheng, D., Liu, M., 2015. An eco-dialogical study of second language learners' World of Warcraft (WoW) gameplay. *Lang. Sci.* 48, 22–41. <https://doi.org/10.1016/j.langsci.2014.10.004>.
- O'Mara, J., 2017. Narratives come to life through coding: digital game making as language and literacy curriculum. In: Catherine, B., Michael, D., O'Mara, J. (Eds.), *Serious Play: Literacy, Learning and Digital Games*. Routledge, New York, N.Y., pp. 102–114.
- Rama, P.S., Black, R.W., van Es, E., Warschauer, M., 2012. Affordances for second language learning in World of Warcraft. *ReCALL* 24 (3), 322–338. <https://doi.org/10.1017/S0958344012000171>.

- Rowan, L., 2017. Playing, making and analysing games: cases of assessment and serious play. In: Catherine, B., Michael, D., O'Mara, J. (Eds.), *Serious Play: Literacy, Learning and Digital Games*. Routledge, New York, N.Y., pp. 186–206.
- Sauro, S., 2021. Online Fanfiction for Language Teaching and Learning. *ALSIC : Apprentissage Des Langues Et Systèmes d'Information Et De Communication*. <https://doi.org/10.4000/alsic.5763>.
- Steinkhuebler, C., Compton-Lily, C., King, E., 2009. Reading in the context of online games. In: *Proceedings of the 9th International Conference of the Learning Sciences—Volume 1 (ICLS '10)*. International Society of the Learning Sciences, pp. 222–229.
- Steinkhuebler, C., 2007. Massively multiplayer online games as a constellation of literacy practices. *E-learning* 4 (3), 297–318.
- Steinkhuebler, C., 2010. Digital games and literacies. *J. Adolesc. Adult Lit.* 54 (1), 61–63.
- Tekinbas, K.S., 2008. In: Tekinbas, K.S. (Ed.), *The Ecology of Games: Connecting Youth, Games, and Learning*. MIT Press.
- Tobias, S., Fletcher, D., 2012. Learning from computer games: a research review. In: De Wannemacker, S., Vandercruysse, S., Clarebout, G. (Eds.), *Serious Games: The Challenge*. Springer, Berlin, Germany, pp. 6–17.
- Witkowski, W., 2021. Videogames are a Bigger Industry Than Movies and North American Sports Combined, Thanks to the Pandemic. MarketWatch. Available at. <https://www.marketwatch.com/story/videogames-are-a-bigger-industry-than-sports-and-movies-combined-thanks-to-the-pandemic-11608654990>.
- Young, M.F., Slota, S.T., Cutter, A., Jaletter, G.C., 2012. Our princess is in another castle: a review of trends in serious gaming for education. *Rev. Educ. Res.* 82 (1), 61–89. <https://doi.org/10.3102/0034654312436980>.

Transliteracy and the digital media: theorizing media and information literacy

Divina Frau-Meigs, University Sorbonne Nouvelle, Paris, France

© 2023 Elsevier Ltd. All rights reserved.

Definitions and epistemological diversity in global perspective	590
Augmenting communication: mediatization rebooted by the increased emergence of participation	591
Augmenting information: documentarization displacements	591
Augmenting code: datafication beyond computing	592
The rise of literacy discourses	593
Transliteracy as a frontier field: between pedagogy and policy	593
Defining transliteracy and its stakes	593
Epistemology and core concepts and themes	595
Emergent themes and research areas	597
The rise of the platform society and new consumption and protection competences	597
Information troubles and new participation competences	597
Sustainability challenges and data governance competences	598
Innovative research postures and methodologies: data sciences and cognition for transdisciplinary designs	599
Future of MIL and transliteracy for transformative education	600
Revisiting the school curriculum	600
Institutionalization of MIL and its risks	601
Conclusion	601
References	601
Relevant websites	604

Media and Information Literacy (MIL) is no longer the shrinking poor relation of more established literacies, due to major digital changes that emerged during the World Summit on the Information Society (2003–05) and were confirmed by the UNESCO Fez Declaration (2011). It recombines in new ways the relatively staid strands of media literacy and information literacy and comes in many guises, using low and high tech, often in extra-curricular activities and projects, in non-formal popular education settings. What is new is the need to twine those strands more closely and to integrate yet another dimension of information, data and its attendant algorithms as they drive social media and other upcoming immersive media. What needs to be done is to scale up MIL in curricular activities, in the formal education sector, with innovative pedagogical and research postures for future-proofing the XXIst century needs for a fully literate citizenry ready to face global uncertainties and crises.

What is lacking and what can be added to the current research is a change of focus, less on the skills and practices and more on the competences, finalities and values of such an education. What is also lacking is a full-fledged articulation of the notion of “information” with communication as it has evolved with digital convergence. This convergence enables mediatization (the ways in which networked media penetrate all cultural spheres) to mesh with documentarization (the ways in which all our records can be traced and interconnected) and datafication (the ways in which all our data transform and predict our social actions). The new challenges posed by globalization also call for a new restructuring of the notion of MIL, incorporating computing and documentation beyond the computer and library contexts to which they tend to be bound (Serres, 2012). For a clearer and larger vision of the issues at stake, this change of paradigm needs to propose theories for MIL that take into account its epistemological complexity, as a “transliteracy” (Frau-Meigs, 2012) responsive to the felt experience of most people in the digital culture. The Covid-19 pandemic has accelerated the awareness of such necessity to develop informational resilience among the public at large.

Definitions and epistemological diversity in global perspective

The field of MIL is made of sedimented accretions because of the extreme lability of media themselves. It has been theorized by stages, with early epistemologies organized around the triptych of production, representation and audience (Masterman, 1985; Buckingham, 2007; Potter, 2013), all calling attention to the socialization and social learning conveyed by media in society. The pedagogical postures have ranged from encouraging hands-on approaches and learning by doing with the media (Hobbs, 2010) to laying the emphasis on critical thinking and civic agency (von Feilitzen and Carlsson, 2003; Kotilainen et al., 2010; Carlsson, 2019; Kellner and Share, 2005). Such an approach has transitioned fairly easily to new media and social media with the adjunction of participation to the initial triptych focusing on production, representation and audience (Jenkins et al., 2006; Denouel et al., 2014; Jehel, 2014). But it has failed to embrace the full significance of information uses and practices and in particular disinformation, with the attendant social disorders of polarization and radicalization. The early euphoric view of the information networks as

“surfing” opportunities for an ever growing mediatization of society —attached to increased democratization worldwide—have failed to discern the “mining” risks of documentarization and datafication on media. This implies to revisit mediatization, documentarization and datafication in light of their impact on socialization and on MIL.

Augmenting communication: mediatization rebooted by the increased emergence of participation

The contours of media have shifted and impacted almost every aspect of human interaction, in most of the world, via the Internet networks and platforms. This has complexified “mediations” (Martin Barbero, 1987) as youth and adult lives are mediated by all sorts of media, including self-made media, like selfies or stories on social media like Instagram, Snapchat, etc. These pure-player media cohabit with pre-digital ones, with series and other long narrative forms that emerged on radio and television and are still quite alive, with extensions to videogames (Frau-Meigs, 2021c). The current media culture is striking by its enveloping multi-layered dimensions:

- It is multimodal with multiple affordances (uniting modalities of oral, print, (mass) media, participatory media to create remixed “texts” and “spectacles”);
- It is convergent and transversal (multiple platforms and settings, constant crossover, transmedia navigation, virtual environments, etc.);
- It is variable and multileveled (various points of entry, whether basic communication, mediated consumption, or complex levels of creation);
- It is ubiquitous, mobile and just-in-time (tethered selves and always-on amateurization of media production as an everyday reality);
- It is an immersive sensory environment with embedded learning affordances (how we take advantage of the learning capacities is key to media and information literacy) (Poyntz et al., 2020).

This complex ubiquity entails that these media formats have been enmeshed in other spheres of life, beyond entertainment, play and news, leading to “mediatization” that refers to the ways in which globalized and networked media penetrate the spheres of work, labor, trade, politics, religion and education and have a transformative impact on how these spheres operate (Hepp, 2013; Hjarvard, 2013). They also restructure them, with disruptive digital strategies that can displace whole economic sectors, as Amazon did to book retail first, before dealing with the whole retail sector and moving to online video streaming with Amazon Prime. The same could be said of Facebook or Google as social media and search engines are not just used for search and interaction but as gateways to all kinds of services (booking, buying, advertizing ...).

Mediatization processes show how media matter for cultural production and consumption but also for education and employment, as digital modes of communication can lead to direct actions on the real world. These new mediated processes require new mediated forms of socialization as they affect identity formation (Krotz, 2009; Erstad and Sefton-Green, 2015; Livingstone and Haddon, 2009), but also privacy, security and agency. They also affect citizenship as they enable participatory strategies that develop individual and collective engagement in public life (Kupiainen and Sintonen, 2009; Johansson and Hvinden, 2007).

Mediatization has affected representation and citizenship in positive and negative ways, affecting especially the notion of “truth” as sources of knowledge and information are being questioned by social media, with opinions painted as “alternative facts” competing with verified facts. The communication processes that have resulted from uses and practices point to societal issues with implications for democratic societies such as “fake news” and propaganda, echo chambers and filter bubbles, “narrowcasting” to specific audience segments known as “online communities” with trolling and astroturfing strategies. This has led to offline consequences such as the rise of populism, authoritarianism, radicalization and polarization (Tucker et al., 2018).

Mediatization has affected media literacy, as it needed to follow the media it examines critically. The augmentation of participation afforded by mass media and mass social media has added new competences to deal with the transmedia and transborder nature of the networks as well as “information troubles” (Wardle and Drekscham, 2017). Understanding online participation, with its opportunities for activism and civic engagement has become key (De Abreu et al., 2017; Mihailidis, 2018). New kinds of subjectivities have been taken into accounts, linked to emotions and cognitive biases and also to creative industries that require flexible, adjustable and mobile learners and workers, capable of acting confidently in a multicultural environment (Rizvi, 2009; Rizvi and Lingard, 2010; Frau-Meigs and Torrent, 2009).

Augmenting information: documentarization displacements

The contours of information have also moved in the digital era. Historically, it encompassed the production and referencing of (mostly paper) documents with a strong stress on their verifiability so as to ensure knowledge construction and transmission (Chartier, 1993; Goody, 1986). This early documentarization of the world used print formats as vehicles for cataloging, indexing and contextualizing documents to provide better access to their contents and to their context (Pédaque, 2007). The print formats have been displaced by digital files coming in all sorts of standards (word, pdf, ppt, jpeg ...). This re-documentarization entails both the digital treatment of old documents with the added functionalities afforded by ICTs (reconstruction, traceability, interactivity) and a different organizational process (search engines, online deposits, open sources, recommendation systems, online collections and archives) with record management and continuous updating via metadata. Thesauri and indexes have given way

to ontologies, “tags” and search engines, with attendant consequences on knowledge construction and transmission. The platformization of global search engines such as Google, Baidu, Bing, Yahoo! or Yandex testifies to the power of search, especially allied with navigators such as Chrome, Firefox or Safari as access points to a wealth of documents.

The construction, composition and publication of such documents have been made available to all thanks to networks and personal computers as they have become dematerialized and turned into distinct pixels (Salaun, 2007) that allow for the infinite transformation of documents, including their falsification. Wikipedia in many ways exemplifies this change, from an online encyclopedia of yore to portals and hyperlinks, curated by professionals and amateurs. Documentarization has also increased participation in the construction and evolution of such documents, with increased sociability and relational interactions applied to information sharing and curating. This has led to new practices of curation with tools such as Pearltrees, Pocket or Paper.li that have modified the order of recommendation, displacing the vertical gate-keeping of culture with a more horizontal peer-to-peer system (Serres, 2012). Participation has benefited from it with the immediate sharing of publications in the public space, and the confrontation of texts of all sorts to many others, visible by all sorts of viewers and readers (Salaun, 2007).

Documentarization has also affected democratic societies, potentially transforming humans as documents among many others (Ertzscheid, 2009). This prolongs the mediated identity formation (drivers license, identity cards, birth certificates, social security numbers ...) with a digital identity summed up by the collection of traces left by navigating, searching and browsing online. URL (Uniform Resource Location) and IRL (In Real Life) modes are both part of socialization and identity formation, be it voluntary or involuntary. It has led to information troubles of a different kind, due to the ease with which documents can be modified, morphed, falsified. It has contributed to “fake news” and disinformation, putting at risk scientific knowledge as well as documentation and library sciences.

Documentarization has also affected media literacy, as information literacy has been aggregated to it by various UNESCO initiatives such as the MIL Alliance or AMIL (ex-GAPMIL) since 2011. MIL brings together the potential of search and the potential of media, to encapsulate the necessary uses and practices of the social media platforms. It also aims at making more critical citizens, capable of assessing truth and trust, aware of the risks and opportunities of online publication and life long learning (von Feilitzen and Carlsson, 2003; Machin-Maestromatteo et al., 2013).

Augmenting code: datafication beyond computing

Datafication pushes mediatization and documentarization one step further, as much information is emanating from processed data. The traces left by search, navigation and other online interactions create a series of data that can be turned into big data and used to profile not only users but also actions. Datafication (Cukier and Mayer Schoenberger, 2013) is the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis. Once more, all spheres are touched, and invisible processes can be turned into data thanks to algorithms that can then monitor, analyze and most importantly optimize such actions. Data can thus be turned into information, via computerized processes that convert human interactions into new forms of value as data formats can be put to social, political and economical uses. Datafication is the source for the main business model of search engines and social media: psycho-metric micro-targeted advertising and can be used for propaganda and disinformation of the general public (Benkler et al., 2018).

Datafication affects socialization as social media platforms like Facebook or Instagram can trace friendships and use them to market other services (Casilli, 2008; Cardon, 2015; Boulrier, 2019). Less positive activities such as surveillance or social credit affect citizenship as well, with censorship, behavioral nudging or electoral manipulation via “fake news.” Netflix for instance moved from its core business in mail order, subscription-based video rentals (DVD and Blu-ray) to a full production system, with recommendations and global trending via its algorithm CineMatch that uses “machine learning” technology and a huge database to make its algorithms learn according to how users consume videos and to recommend to them what is “trending now.” Beyond recommendation systems, datafication is used to inform content, to train machines to behave like humans, for services in insurance, banking, social science research, that are increasingly managed by bots and metadata crawlers.

Datafication has an impact on socialization and identity formation because of the creation of the “quantified self” (Lupton, 2016) that does not necessarily bear resemblance with reality but generates metrics based on information provided willingly or unwillingly. Metrics are also connected to citizenship in that they show engagement. It prolongs human as documents with life mining options as data crunching becomes the norm. Even friendships are « datafied » via the « likes » of Facebook (Cukier and Mayer Schoenberger, 2013).

Datafication has also affected democratic societies, with deepfakes (AI-driven visual and audio content intended to deceive), dataveillance (undetected monitoring of users’ activities across platforms) and all sorts of privacy trespassing strategies (Van Dijck, 2014). This phenomenon is increased by the geopolitics of the internet as huge transnational companies protected by California law, but also increasingly Chinese law, shape a “platform society,” with the sense that they are organizing public and private life and cultural values (Van Dijck et al., 2018). Everyday life is modulated by the global ecosystem they create via algorithms and data, beyond virtual worlds (Srnicek, 2016). Beyond the pure player services they offer such as search engines and browsers, data servers and cloud computing, social networks and app stores, they also are operating other economic domains regardless of their nature as in the case of Uber (for taxis) or Airbnb (for hotels), with their own payment systems and engagement metrics. In so doing they often bypass national sovereignty. The platform ecosystem has incorporated mass media outlets and used its capacities to manipulate political situations, as revealed by the Cambridge Analytica scandal in 2017 (Manokha, 2018).

Datafication as *digital literacy* also reshapes the contours of media literacy. Often seen in contradistinction to computer sciences (focused on programming languages and informatics) it covers a more or less stable set of uses and practices in relation to digital tools and platforms, without necessarily understanding their design, their functioning and their finalities (Eshet-Alkalai, 2004; Wing, 2006; Baron and Bruillard, 2001; Drot-Delange and Bruillard, 2012). The core issues and key words become those dealing with transparency, accuracy, safety and privacy issues on the one hand and public good, social responsibility and democratic control on the other hand (Van Dijk et al., 2018; Frau-Meigs, 2021b; Van Dijk and Van Deursen, 2014). As they are not inscribed into the algorithmic processes operated by the platforms, they need to be imparted to citizens so that they are not treated as mere users.

The rise of literacy discourses

These three strands have each created an array of practices, pedagogies and policies that have led to many new literacies (Landry, 2017). Mediatization has led to news literacy, visual literacy, transmedia literacy. Documentarization has led to information literacy, search literacy, and datafication to digital literacy, algorithmic literacy, AI literacy. Globalization and the emergence of a global youth culture in connection to connected networks has also played a role in the intensification of literacy discourses and appended descriptors, instead of education discourses (Frechette, 2002).

This shift from education to literacy reflects the attempt at conceptualizing informal practices and uses that are not formally included in school curricula or integrated in official programs (Frau-Meigs, 2021b). Turning such set of practices and uses into a literacy has helped them penetrate the education sphere, as in the case of France where a report by the CNUM (2013), the National council on digital matters, opened the door for the imported term “littératie.” Frenchifying the term smoothed the path of resistance and shifted mindsets away from early discourses of cultural distinctions (Bourdieu, 1979) or class accommodations (Hoggart, 1957; Williams, 1958) associated with literacy. Some countries like the USA combine both, and deal with media literacy education (Hobbs, 2011), blending the ability to identify, interpret and create media contents with communicating and computing.

This shift conveys the pressure to respond to the needs of the platform society, as all sectors are pervaded with new workplace sets of skills, soft and hard (Hoechsmann and Poyntz, 2012; Frau-Meigs and Hibbard, 2016). MIL can thus be seen either as critically fostering such a set of skills or as catering and pandering to platform society requirements for subjectivities that are non critical of the new ecosystem. As schools are required to go digital and increase the digital competences of their students as future workers, so is MIL solicited to deal with values and concepts out of the instrumental realm, in a sort of division of labor (Fastrez and De Smedt, 2012).

So the literacy discourses are not neutral and reveal tensions in the field, around the necessity to “adapt” to the digital culture or to “cope” with it, at all costs for instance, that has technologically deterministic overtones. This can create confusion or chaos or even dilute the strength of MIL. But such profusion can also be constructive as the piling up of such literacies may finally shake the pre-digital school curricula current standstill and create enough momentum to break the status quo and allow MIL as part of the first curricula (Lee, 2020). It can also reveal the full scope of MIL as many shades of MIL are at work in diverse regions of the world from “educational communication” in Latin America (Aparici, 2010; Kaplun, 1998; Soares, 2019) to “multi-literacies” in the Nordic countries (Selber, 2004). The competing paradigms for learning show tensions between transmission (top-down knowledge acquisition in school institutions), co-construction (bottom-up development of competences via shared tools for mediation) and participation (horizontal and vertical circulation of knowledge within communities) (Frau-Meigs, 2012, 2017). The common goal remains the same though, fighting illiteracy and, additionally, illettronism that could be defined as the lack of transliteracy, with the cultural and digital divides that come with such lack.

At the international level, the tensions between these different forms of literacy are also at play, often in disputes over the articulation of computer sciences and ICT training, especially in schools where most countries have made the historic choice not to teach computing, except in optional and vocational sections. This has increased the gap between information as a science and information as a culture to the point that the hard sciences (physics, chemistry, maths, biology, etc.) tend to reorganize themselves around “information sciences” that exclude the humanities and therefore the political, social and economical dimensions of information cultures, most evident in the heavily adopted STEM initiatives. This rift has stalled the implementation of fully transliterate solutions. There has been no acknowledgment of effective uses, their representations in the users’ minds and their potential for creativity and empowerment, which could be seen as way of naturalizing digitization of society in a frictionless manner, without acknowledging the highly-media driven content and erasing the critical perspective on the evolution of ICTs (Frau-meigs, 2013, 2015).

Transliteracy as a frontier field: between pedagogy and policy

Defining transliteracy and its stakes

These three strands (mediatization, documentarization and datafication) are twined and are experienced as seamless by users in the practice of their everyday lives and experiences. A displacement in competences is visible that combines knowledge constructions and (self) representations with individual skills. Broad skills are needed for leisure but also for work and political participation; more specific skills are needed to deal with training and education within the new industrial frame of creativity and innovation (Drotner, 2011) that replaces the reproduction and transmission model of the pre-digital era in all sectors of society, via mediatization (Frau-Meigs et al., 2017).

This paradigm change requires researchers to revisit the territories of “information” as they are more and more closely related to datafication (computer literacy), documentation (information literacy) and communication (media literacy), where the term “information” gathers added meanings (news, documents and data). This move implies to spell out the double meaning of “transliteracy” in today’s context of convergence:

1. The ability to embrace the multimedia layout that encompasses skills for reading, writing and counting with all the available tools (from paper to image, from book to wiki);
2. The capacity to navigate through multiple domains that includes the ability to search, to evaluate, to test, to validate, and to modify information according to its relevant contexts of use (as code, news and document) (Frau-Meigs, 2012).

The rise of transliteracy itself is an emerging phenomenon, associated to the rise of literacy discourses. The term was coined by Liu (2012), director of the Transliterations Project in the Department of English at the University of California Santa Barbara (USA) and was focused on the practices of online reading. The same year Thomas et al. (2007) at the Institute of Creative Technologies (IOCT) at De Montfort University (Leicester, UK), founded the Production and Research in Transliteracy (PART) group, that focused on online writing practices. From its origins focused on English teaching the terme moved to embrace the convergence of communication, information and computer sciences, with the pioneering work of the research group LIMIN-R and the TRANSLIT project piloted by Frau-Meigs (2013). The emphasis was no longer on reading and writing but on the way the three information cultures required a fully-fledged set of competences for navigating the digital world (Delamotte et al., 2014).

The “trans-” prefix aims at conveying the notion that learning is “transverse,” that it is set across several domains, that are trans-media, transdata, transdocuments and even translingual and transborders, creating configurations that are sustainable though they may seem unlikely if each piece is taken separately. It allows for simulations, immersions, visualizations that can be learning events, produced in a non linear, non transmissive manner. The term also refers to digital popular culture and its “metaverse” (contraction of meta universe) a virtual world described in Neal Stephenson’s novel *Snow Crash* (1992). This virtual world created by a computer program, hosts a community of users who interact via their avatars, in anticipation of what would become Second Life (2003), one of the first online metaverses where, as in other virtual immersive environments people can work, play and learn. When applied to education, what is important to keep in mind is scalability (more and more people can aggregate to it, like a networked structure), usability (the user experience is central and has to feel friendly, like a relational good) and the prevalence of code and protocols to build architectures and interfaces and connected objects. Facebook’s decision to call itself META, in 2021, reflects the paradigm shift and the potential focus on participation added to education.

Defining and modeling transliteracy consists in adopting the convergence of three main types of literacies: media literacy, information literacy to which computer literacy is added for a full understanding of the techno-cultural constraints of the digital world. Its aim is to contribute to the construction of an analytical framework on the status of information in today’s culture, from the perspective of the communities of interpretation and practice that elaborate the three converging fields of information and library sciences, media and communications studies and computer engineering for social uses. Ultimately, this modeling process can help understand “information cultures” (not reduced to computer use). It can be of benefit to all researchers interested in the digital humanities as well as the communities of practice around young people (teachers, librarians, webmasters, ...) in their training and practice, on the basis that information and its communication are fueling knowledge development in the XXIst century (Salvador, 2019; Frau-Meigs, 2019b).

This hypothetical model was tested during the ANR Translit project, that was ground-breaking because of its transdisciplinary nature and its integrative purpose, taking learning events as they were in disjointed spaces (libraries, classrooms, virtual platforms) and conjoining them by looking at the seamless uses by young learners. This led to a construct for literacy with information at its core, within and outside schools, for social and pedagogical uses, with empowerment and engagement as learning outcomes (see Fig. 1). Three meta domains of competences were found: cultural and socialization competences for media literacy (or infomedia), organizational and editorial competences for information literacy (or infodocs) and operational and problem solving competences for data literacy (or infodata).

Within each of the three meta-dimensions, subsets for each could be broken down into a series of competences: criticism, citizenship, communication for info-media, search, publication and navigation for info-docs and protection, creativity and participation for info-data. Each of these competences can be implemented according to the four components of competences: skills, attitudes, knowledge and values. This framework for transliteracy, though elaborated separately, showed some overlap with the newly minted digital competences framework in Europe, that mention five major dimensions: information, communication, safety, creativity and problem solving (DIGCOMP, 2013, 2018, 2022). Unesco adopted a composite multi-faceted MIL approach not unlike transliteracy as “MIL constitutes a composite set of knowledge, skills, attitudes, competencies and practices that allow effectively access, analyze, critically evaluate, interpret, use, create and disseminate information and media products with the use of existing means and tools on a creative, legal and ethical basis” (2013). The notion remains that media cannot be diluted in the digital, because the convergence of mediatization, documentarization and datafication have to be thought per se, critically and creatively. Both may coexist, though they describe different realities and operate on different scales, but they both signal a paradigm shift. Transliteracy as the understanding of how the three information cultures are generated, processed, and packaged for audiences and communities provides a basis for mastering both the limitations and opportunities of the digital environment and acts as a foundation to use information, evaluate its relevance and contribute cogent contents and comments in the digital ecosystem.

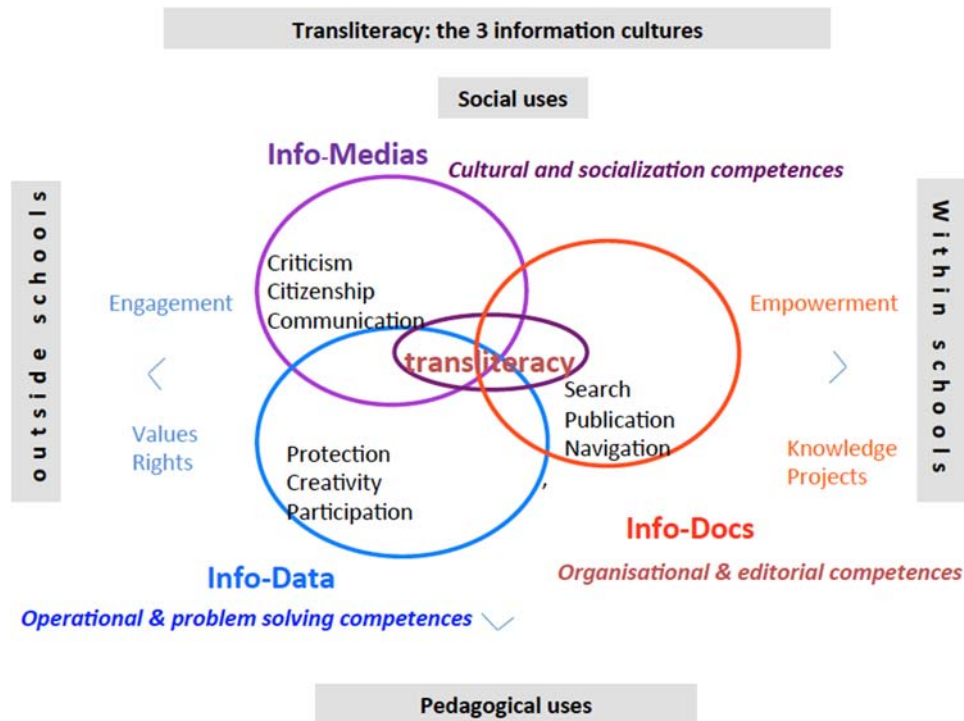


Fig. 1 Transliteracy model: the three information cultures (Frau-Meigs, 2013).

Epistemology and core concepts and themes

Being aware of the centrality and of the inevitability of these literacies makes it possible to consider how transliteracy allows for a better knowledge of the processes associated to computer sciences, information and communication sciences and library sciences. Such a stance implies an analysis based both on the contents and their specificities (representations, formats, audiences) and on the processes and modes of access characteristic of data sciences (coding, storing and file-sharing, recommending, predicting). The operational and functional competences make it possible to understand the impact of technological design as well as the actors behind such design and the actors at the end process of this design, with their finalized needs and goals.

So transliteracy has a broad remit within MIL: responding to mediatization, documentarization and datafication, with key words that resonate with each other, beyond pre-digital notions while including them. So to initial core notions of production, representation, audiences and citizenship, others need to be added: participation, communities, networks, governance and creativity (see Fig. 2). These notions touch other boundaries, in the research on education sciences, youth studies, communication and

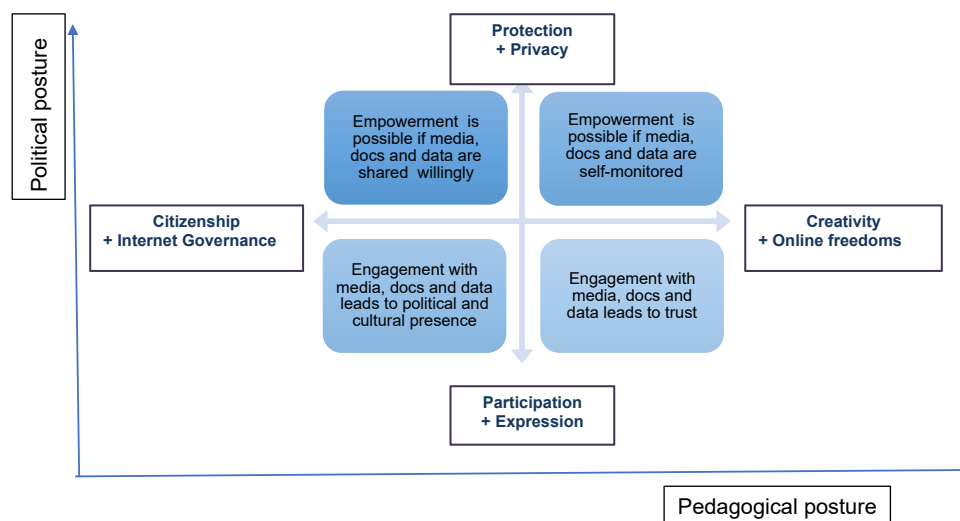


Fig. 2 MIL epistemology (Frau-Meigs, 2017).

information sciences and data studies. But what distinguishes MIL and transliteracy is the double axis of political project and pedagogical posture, in “a call to praxis, the fundamental demand that knowledge in the field orient to action in the world, the critical and creative act of responding” (Poyntz et al., 2020) ... and not just coping.

The new media and social media have reinforced participation and added this notion to representation, as well as they have added communities to pre-digital audiences. This more important place of participation has always been present in practical pedagogies and learning by doing scenarios but has been unleashed by interactive media forms and social media platforms. This has led to a “participatory culture” (Jenkins et al., 2006) with creativity and collaboration as key notions. Consequently, early notions of MIL such as protection from risks and safety have been less prevalent, though they are still present in the sphere of data privacy. Additionally online participation has created new conditions for civic agency, engagement and empowerment, encouraging citizens to take part in decision shaping and decision-making (Johansson and Hvinden, 2007). This has added an extra layer to the core notion of citizenship in MIL as citizens need to understand the construction of an array of subject-positions and expression of opinions because that globalization and transborder networks have loosened nationality and brought citizens to exert multi-level loyalties, to their self, their local community, their country, their region (Mäkinen, 2012).

The result is a rather holistic vision of MIL supported by transliteracy, with its recognizable epistemology or unique set of concepts and competences (see Fig. 3). A rapid synthesis of various existing competences frameworks shows that they can be summed up as the 7Cs: Comprehension, Criticism, Creativity, Citizenship, Consumption, Cross-Cultural Communication and Conflict Resolution (Frau-Meigs, 2012). The first three Cs are mostly found within schools while the four other Cs tend to be found outside schools. They reflect the MIL epistemology, with some competences aligning on pedagogy, some aligning on policy. They promote ethics and move MIL beyond current policies for IT or e-skills that are mainly operational. In a holistic manner, they recombine complex problem solving with technical, social and soft skills (Frau-Meigs, 2017).

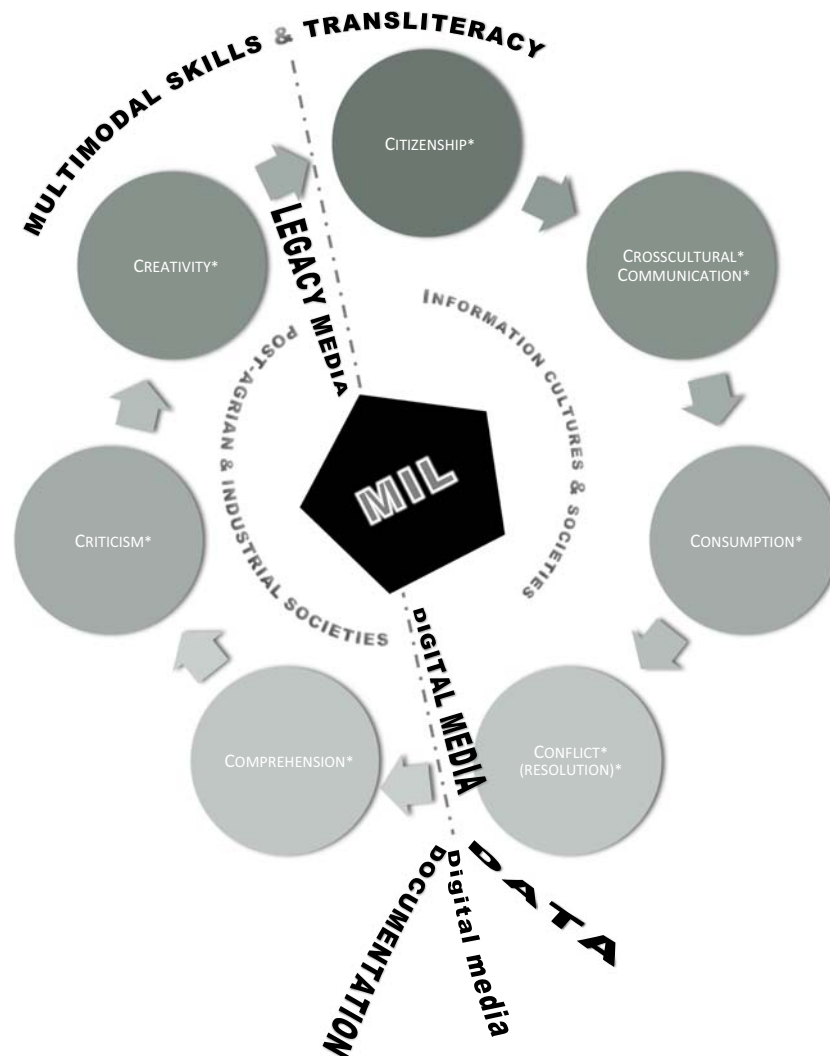


Fig. 3 MIL as transliteracy and 7Cs (Frau-Meigs, 2017).

This set of constructs marks transliteracy as a frontier research with its characteristic paradigm shifts in phases of crisis, not unlike those of scientific revolutions (Khun, 1970). It deals with issues that are highly controversial, such as data privacy (mass surveillance, profiling, nudging), or online freedoms (hate speech, anonymity). It uses methodologies and concepts that are atypical in the general field of education but also in information and communication sciences. It challenges the dominant paradigms of schooling as transmission of knowledge only and breaks the fourth wall of the school and of the screen by allowing online expression and participation, with recommendation systems among peers and not just teachers. It shares with most frontier research the common feature of potentially transforming our understanding of knowledge construction and co-construction as well as identity formation.

Emergent themes and research areas

The new research themes that are related to transliteracy as frontier research call on knowledge and critical understanding of mediatisation, documentarization and datafication and their convergence. The field of MIL is very rich in resources with a praxis-oriented tradition, that is aligned with the felt needs of learners and users (Frau-Meigs et al., 2020, 2021). In spite of the rise of creativity and participation, issues of representation, production, reception and protection are still very much key, as a multiplicity of stories, selfies and series are broadcast via screens, as well as fake news that built on stereotypes to fuel polarization and hate speech. So production, language, audience and citizenship are still very much in need of critical thinking, with identity, gender and class as elements for dealing with disinformation for instance. But also new issues surrounding data, privacy and online well-being require critical and creative attention.

This moves MIL and transliteracy away from a mere understanding of “the press” (Gonnet, 1995) to learning about mass and social media as part of growing up in media saturated information cultures. New directions for research have emerged from the need to respond to crises and rapid changes while providing understanding and resilience, to resist anxieties and fears caused by excessive algorithmic surveillance and to clear democratic doubts caused by migrant crises, populism, social polarization and radicalization. The existential threat posed by climate change as well as pandemics such as Covid-19 also call for ways of nurturing innovation in research design and pedagogy.

The rise of the platform society and new consumption and protection competences

The rise of the platform society is affecting MIL in several ways, be it the understanding of the audiovisual fiction production line or the new consumption habits of globalized youth audiences. On the production side, research on Netflix as a paradigmatic company in the field of audiovisual creations, points to the way data and algorithms are used to predict the preferences of the audience (via Cinematch, Netflix’s algorithm). Many “over the top” (OTT) audiovisual platforms are configuring an “algorithmic taste” based on the mechanisms of Artificial Intelligence to mold and design contents supposedly tailored to consumer tastes and preferences (Halpin and Striplas, 2016).

This in turn affects culture and the control of the cultural and creative industries: a disruption of the pre-digital studio and network Hollywood-style system and the rise of “Hollyweb” with a different kind of control and power. On the surface it feels like more freedom of choices for users, under the surface more control and opaque processes that monitor users unbeknownst of their own will and act on their bodies, subjectivities and territories (Roberge and Seyfert, 2016; Gendler, 2016). Datafication is at work in this cultural change with new modes of perceiving the world and perceiving oneself (Van Dijck, 2014; Elmer et al., 2015), with questions about algorithmic recommendations and the future of audiovisual consumption, the degrees of freedom and conformity within algorithmic tastes and the opportunities and risks for audiences and their socialization to storytelling. A characteristic example is FOMO, the “Fear of Missing Out,” that points to the needs for social belonging (Przybylski et al., 2013) and the strive for being up to date in tendencies and recommendations that will also fuel conversations on social media.

Socialization is driving the reconfiguration of the GAFAM, as exemplified by Amazon’s acquisition of Metro-Goldwyn Mayer in 2021, to lay hands on a century long catalog of movies, hence proceeding to its redocumentarization. They are all striving to enrich their storytelling and their user-aggregated storytelling, each adding functionalities such as “reels” or “stories” to their social media. “Who owns what” is part of key knowledge for a better understanding of the future of communication and of cinema and audiovisual content, important for all to know as it may lead to improved competences in media consumption and algorithmic literacy. It becomes crucial to pay attention to traces, to navigation history and to cookies as well as requesting transparency in the collection of data and their uses by the GAFAM (Google, Apple, Facebook, Amazon, Microsoft) and the BATX (Baidu, Alibaba, Tencent and Xiaomi).

Information troubles and new participation competences

The rise of the platform society also impacts hard information, not just the soft socialization via fiction. Datafication and the mass sharing of news has led to information troubles (Wardle and Drekscham, 2017; Frau-Meigs, 2019a,c). The spread of disinformation has been connected to documentarization, with the fabrication and manipulation of authentic documents and contents in other

contexts, as exemplified by deepfakes, where faces and voices can be swapped from one body to another. Disinformation mechanisms prey on human biological, cognitive, and emotional biases that render citizens vulnerable to manipulation but they also play on publicity and the business model of social media that drives traffic for profit, based on the popularity of news rather than their quality.

Transliteracy can help raise levels of visual literacy and news literacy in the transverse dimensions of the Internet (transnational, transmedia, translanguages). It builds on critical thinking competences to consider how people make sense of information, how they understand the conditions of its production and amplification and how they are affected by disinformation, as it impacts their participation in democratic life and their trust in institutions, including the media (Ashley et al., 2017). Citizenship for empowerment and engagement is necessary for socially sustainable societies that can counter disinformation in day-to-day life and responsive to lived experiences (Boyd, 2017; Mihailidis and Thevenin, 2013).

Being able to discriminate various elements of the information factory and its attendant disinformation versant, implies to know about content, context and visual-image characteristics, and also about source (partisan sources, ingroup/outgroup sources), alongside characteristics of users (e.g., personality traits, values, prior knowledge/exposure, and media-engagement tendencies) and social norms (e.g., peers' reactions). This implies knowledge also of filter bubbles and echo chambers to better understand the role of online communities of practice (rather than audiences). These elements solicited transliteracy competences in critical thinking, but also conflict resolution as disinformation can lead to polarization and radicalization (Frau-Meigs, 2021b). Users thus need to realize that commercial media platforms are being used as de facto public spaces but without the public service obligations of public media.

Robust models to build informational resilience are also necessary as illustrated in project Youcheck! that associated a fact checking visual tool for journalists with inclass interventions in four European countries. It checked visual credibility and used advanced data science approaches to analyze the multicultural datasets in novel ways to build a quantitative model for fighting disinformation effects. Preparing interventions for students can be done by scaffolding strategies that enable them to access and assess credible news, so as to develop their competences in transliteracy, that is to say their ability to navigate between media, data and online documents, some of them fake, some of them accurate. The findings confirm that a 2 h intervention may have a significant effect on young people's abilities to determine the credibility of fake images and videos. They also show that sensitizing them to fact-checking tools changed their behavior as to their use to evaluate news credibility. Additionally, the intervention brought attention to image design and cognitive strategies such as lateralization in order to build heuristics for visual literacy among young people (Nygren et al., 2021).

Sustainability challenges and data governance competences

The two major crises of the XXIst century relate to sustainability and human well being. This points to focus of transliteracy on eco-media and on care praxis. It also points to governance reconsidered as it is about societal decision-making beyond governments, at global level, with multi-actors (Hajek and Wagenaar, 2003). The extensive use of epidemiological models during the covid-19 crisis and those used for systemic climate change provide a common thread for transliteracy, data-visualization, as a technique to describe the present and predict the future. Such models have been very active in the media with the "flatten the curve" injunction that evolved from it, showing that they do not only produce knowledge about complex systems, but they also frame societal decision-making.

Drawing the lessons from the disinfodemic that came together with the pandemic, a full information ecology is called for, to help solve environmental issues and foster sustainable development. Eco-media literacy (Rust et al., 2016; López, 2014, 2020) construes sustainability as environmental sustainability (weather reports, resource management, air pollution, etc.) and its public accountability which implies the need for open public spaces for the expression of all sorts of concerns for the future (climate change, digital pollution, migrations ...). Eco-media calls attention to the "anthropocene" (Crutzen and Stoermer, 2000; Chakrabarty, 2018), the recognition that humans have been modifying the environment with their mass of artifacts, including those needed to convey information like satellites, servers, clouds, screens, optic fiber cables and rare earth minerals (Maxwell et al., 2015; Walker and Starosielski, 2016).

Transliteracy can draw attention to the materiality of information and its ecological footprint as well as its mindprint, that is the representations about environmental issues as they are mainly constructed by media. Such constructs affect our beliefs and attitudes about climate change, with disinformation alleging that unlimited commercial freedom should not tolerate any regulation about the environment (Frau-Meigs, 2021a). Denying human responsibility is tantamount to « greenwashing », which has been apparent in advertising in particular and calls attention to the political economy of media ownership, as they rely on corporate ads or belong to large corporations, and promote unsustainable lifestyles.

MIL and transliteracy bring attention to the competences needed to decode such brainwashing, and engaging people's agency in not accepting seemingly innocuous metaphors such as "clouds" or "surf" that in fact obfuscate the reality of their polluting practices. Care, as the art of solicitude, calls attention to the needs for a full internet governance, with transparency, accountability, eco-justice and eco-responsibility in a networked society (Gilligan, 1982; Fine and Tronto, 2020). It produces robust counter-narratives to the current models of globalization that are exploitative and stifle the voice of indigenous people or young people, such as those expressed in the Fridays for Future (fridaysforfuture.org) movement (Redpath et al., 2018; Vieira and Tenreiro-Vieira, 2016).

These changes point to citizenship and how it is affected by data and its necessary governance (Brown and Katz, 2019; Brännback et al., 2017). Transliteracy can then rely on design thinking to embed its value systems in decision-shaping if not decision-making. It is paramount to understand how data can serve as evidence, and to make sense of different inferences from data, especially when they take the shape of predictive modeling or causal inference, that may sustain a critical discourse on future crises. This governance of data comes with care as data-based decision-making can create biases because that algorithms are editorialized by humans, with consequences on human rights and social justice, with populations at risk such as migrants or minorities that can be marginalized or discriminated against (D'Ignazio and Klein, 2020). Venues like the Internet Governance Forum (IGF) have stressed this need for competences in so called “digital education,” using policy-making with a high degree of inclusion of various stakeholders from the public, private and civic sectors (Frau-Meigs and Hibbard 2016). MIL has always relied on this multistakeholder approach to train and sensitize young people and adults to information cultures, creating a kind of governance from the bottom up, that is sensitive to rapid change and social innovation (Frau-Meigs et al., 2017).

Competences in rights and responsibilities online, activism for GDPR and facial surveillance, they all point to the fact that “the pedagogical is political” (Frau-meigs, 2021a). Mastering the proliferation of apps and immersive environments, as well as the robotic and prosthetic presence of non-human agents and connected objects is going to be essential to build a sense of community and well-being. This will not be the responsibility of individual users alone, but also of the platforms, to develop a sense of citizenship in a common planetary ecosystem.

Innovative research postures and methodologies: data sciences and cognition for transdisciplinary designs

Transliteracy affords many opportunities for theoretical and methodological developments that are much-needed in the area of converging literacies and in the face of rapid changes. Among the research questions, the issue of competences and their relation to the finalities of information search and use is crucial. Another key research question is the new collective dynamics, on the scale of the school and of the territory, that may be set up with the development of transliteracy in learning events with varied configurations, exploring the articulation between the individual and the collective, in the digital space as much as in the physical space. The political and educational options that might appear to reach a fully-fledged transliteracy are also key in education research, even in repurposing education around information multi-literacies as is currently being done in Finland (Palsa, 2020).

Transliteracy lends itself to transdisciplinary designs (not just interdisciplinary, between the social sciences), across social sciences and data sciences. There is much hope in evidence-based development of intelligent digital tools for assessing critical thinking skills and design experiments that iteratively scale innovations into practice. Such experiment was conducted in the YouVerify! project, where a Massive Online Open Course (MOOC) on “Disinformation Step by Step” was produced to scale up MIL contents, tools and pedagogical scenarios and make them available to participants at large (youverify.eu/).

Data sciences, in combination with cognition and cognitive neurosciences, belong naturally with transliteracy as it strives to clarify the relationship between data and media. Using datafication to conduct research on the rise of the platform society, information troubles and ecomedia sustainability is yet another manner to understand the way data function and bolster the mediatisation of society. For instance, the InVID plug-in used in YouCheck! and YouVerify! projects offers “metadata” tools and Social Network Analysis tools that not only help identify if images are fake or not but also how data are being used in disinformation and, in reverse engineering, to debunk disinformation. Such projects provide methods that can be effective for joint analysis of multidimensional behavioral and social patterns. They have already been used effectively with predictive models of susceptibility to the effects of disinformation (Hand, 2020; Mittelstadt et al., 2016; Vosoughi et al., 2018).

In this light, transliteracy lends itself to socio-cognition and cognitive neuroscience approaches as they shed light on socialization processes since they consider the learners’ interactions with their environment (including the media artifacts) and the multiplicity of interactions with their environment (Salomon, 1994; Perkins, 1995). Cognition posits that the learners are capable of articulating knowledge (content level) and pertinent information search (access, validation) in such a way as to construct their own understanding. Cognition comes in many shapes, like epistemic cognition that enables critical thinking (Greene et al., 2016), multiple source use (Verhoeven and Perfetti, 2008), and multimodal text and image processing (Leu et al., 2015). It takes into account the learners’ epistemic beliefs to build epistemic maturity, to turn information into knowledge and build a coherent representation of multiple texts (Frau-Meigs and Bossu, 2018). Techno-cognitive approaches are helping deal with disinformation and cognitive biases, with the use of smart AI-driven tools (Lewandosky et al., 2017).

Cognitive approaches also make it possible to deal with the emotional load of participation (Kahan, 2013, 2017; Kahne and Bowyer, 2017). Learning theories that describe mechanisms of how emotions and learning processes interact can be used to design those transliterate materials par excellence that are serious games. Although the literature on utilizing games to promote critical thinking is still sparse, research suggests promising results in this domain of gamification (Kapp, 2012; Sanchez et al., 2016). Games are undergoing a paradigmatic shift toward constructionist gaming in which students design games and generate game content for learning instead of just consuming games (Kafai and Burke, 2015; Moon and Liu, 2019; Shute and Sun, 2020). Participatory science can also be used for crowdsourced game content for instance (Lukyanenko et al., 2016). The Play Your Role project engaged in just such experiment to train young people to fight hate speech, with pedagogic itineraries that integrate cognitive and emotional solutions to support critical thinking and civic engagement (savoirdavenir.net/play-your-role).

Information troubles have pointed to the need to pay attention to images and emotions and how compelling they are in understanding their contribution to human development and cultural phenomena such as online series, selfies, streams and stories. Due to their high emotional charge, audio-visual messages seem to be more persuasive and more easily spread than textual content. Battling false visual information requires not only special attention and a new kind of information literacy but also automatic tools that facilitate detection of media manipulations. It implies the mastery of interactive visualization techniques in the face of prevalence of datavisualizations in media (Hohman et al., 2018). This could foster a rapprochement of the MIL community with the fact-checking community as they are developing image forensics with algorithms and smart tools to respond to the growing need to detect and prevent possible misuse of fake media content (Nygren et al., 2021).

Future of MIL and transliteracy for transformative education

In transformative education, learning involves the transformation of the learners' frames of reference (mindsets, worldviews) and the critical thinking and epistemic maturity to reflect on the learning processes even more than the educational outcomes and outputs (Mezirow, 1997). In times of crises and uncertainty, the need to project oneself and constantly update one's knowledge, seeing the strengths and distortions in one's emotions, re-assessing one's values and decisions requires "forwardance" (Frau-Meigs, 2015), the reasoned and agile use of the "affordances" (Gibson, 1977). To foster this forwardance, MIL as transliteracy needs to become part of the first curriculum and to be inserted fully in education policies at large, with proper scaffolding (Yelland and Masters, 2007).

Revisiting the school curriculum

In most countries, MIL is not part of the core school curriculum. It tends to be acknowledged as a cross-cutting topic, not even a discipline, sometime confused as online civics, to be appended to language learning or history classes. But with the rise of the platform society, information troubles and sustainable emergencies, such an approach is cracking at the seams, as MIL epistemology embraces new notions and themes. MIL needs to be both a transversal topic, just like reading and writing are needed for conducting math experiments or science tests, and a core topic per se, just as reading and writing are needed for basic literature and language learning. Only a trans-disciplinary approach such as the one afforded by MIL can help change the curriculum, with new core notions and fields of study around sustainability and information troubles, promoting a pedagogical posture that fosters learner-centered projects rather than top-down programmatic contents.

This decision is political because the risk of ignoring such changes may marginalize education via schools and precipitate education via platforms. Educational institutions need to develop their own platforms and networks, so that digital tools are not used with their current commercial missions, as discovered in the case of Zoom during the Covid-19 pandemic that turned out to have no policy for young learners' data protection, not complying with a certain number of international instruments with the best interests of young people and citizens in mind, such as the General Data Protection Regulation (GDPR). MIL can have structuring effects to help platformize education in an ethical manner, with teachers as influencers and students as active followers and content aggregators.

Putting MIL in the first curriculum, from primary to high school, is crucial to provide citizens with notions of credibility, authenticity, authority, accountability and transparency and for understanding information processes, data patterns, predictive algorithms and social network storytelling (Frau-Meigs, 2021b). Combatting information disorders and building informational resilience with quick healthy reflexes or heuristics for young people to alleviate their doubts about the credibility of a piece of information is as important as helping them use their creativity with their own series, selfies and stories. Critical thinking practices that involve using digital tools to solve problems and resolve conflicts by "learning by doing" and trial-and-error methods are key to forwardance and the care it requires to address issues of socialization and identity construction in relation to information cultures.

For doing so, the training of trainers is essential if schools want to scale up their curriculum to XXIst century subject matters. The lack of preparedness of the school systems at the local and national levels is hampering transformative learning and creating much frustration, among teachers as well as among learners. This requires to address the disconnect between universities and teacher training at large, as higher education plays a role in taking MIL to the school curricula. Currently MIL research and pedagogy tends to be conducted in theoretical fields that are outside education, in particular information and communication sciences and media studies, where it is considered as an ancillary domain to loftier topics such as political communication. Places like the Media Education Research section of the International Association for Media and Communication Research (IAMCR) have been building bridges to make this frontier field visible and remedy the situation. UNESCO, with its unitwin network of chairs in Media and Information Literacy and Intercultural Diversity (MILID) chairs, has pushed for a similar agenda. This has led to the production of MIL handbooks that try to offer constructs and inspirational narratives for MIL, to provide for its own storytelling legitimacy and testify to the vibrancy of the frontier field in many diversified settings and contexts, in intercultural dialog (Yildiz and Keengwe, 2015; De Abreu et al., 2017; Frau-Meigs et al., 2020).

Institutionalization of MIL and its risks

MIL is getting a fair amount of traction as countries are becoming aware of the disruptive and destructive uses of information troubles, as the covid-19 pandemic became a disinfodemic and a matter of public health and social justice. Due to the increase in information troubles and their impact on democratic societies, since 2018, but climaxing in the January 2021 US Capitol Insurrection and the banning of President Trump from social media, MIL is being institutionalized, albeit outside schools systems. For instance the European Union has incorporated MIL in DG-Connect, the division dealing with the EU digital agenda (rather than locating it in DGEAC, the division of education and culture). The 2018 Directive on Audiovisual Media Services makes it an obligation of the member states, in association with an event, European Media Education Week. The EU plan of action to combat online disinformation has led to the creation of the European Digital Media Observatory (EDMO) and 10 national and regional hubs. Additionally, social media platforms have been invited to promote MIL initiatives, which they have promptly added to their repertoire, as a kind of Media Accountability System (MAS) and a way of preventing state intervention (Frau-Meigs, 2019a). Facebook has created a chair for MIL research in the for profit Lille school of journalism in 2020. Google has provided 25 million euros for the MIL fund in the EU in 2021. However, encouraging social media platforms to get involved in MIL initiatives presents a risk of privatizing its public service mission with citizenship at its core.

Such institutionalization brings forth, reflexively, the issue of governance, ... applied to MIL itself. It draws attention to the need for frontier research in policy-making around MIL (Frau-Meigs et al., 2017). The tendency seems to be to use "sensible" practices to move fast, from the bottom up and capitalize on existing initiatives while promoting their larger diffusion and transfer, as in the case of Youcheck! and YouVerify!. The tendency also seems to encourage multi-stakeholder strategies and co-operation among social actors, including the for-profit sector, for providing many resources and materials. While being a positive sign of interest and diversity, the risk is the fragmentation of MIL in a myriad initiatives that can blur the need for a coherent competences framework and epistemology, if not diluting MIL into Digital competences frameworks. The role and emphasis of MIL in policies and the actors' positions in supervising and implementing such policies is therefore key to its capacity to effect transformative change in education. MIL governance should play a key part in public policy to ensure that its political project remains aligned with its pedagogical postures, in the public interest.

Conclusion

MIL as transliteracy can be seen as a rapid response to global changes and crises, not just as engagement with sustainability but also as forwardance, as an ecological manner of dealing with information cultures as they shape everyday life. From a rather ancillary and marginal area of non formal education it is emerging as a field that opens new horizons and new vistas to pursue, with enticing designs and experiments that have immediate consequences on people's experiences and attitudes. It can potentially help the education institutions transition effectively to fulfill the needs of the digital age, with ethics and values as part of the design. It also can contribute to education scholarship and development at large, with compelling notions and themes that shed new light on our ways of learning, as emotions, cognitive biases and multimodal texts and spectacles and stories drive the motivation to learn and to construct knowledge.

References

- Aparici, R. (Ed.), 2010. *Educomunicación: más allá del 2.0*. Gedisa, Barcelona.
- Ashley, S., Maksl, A., Craft, C., 2017. News media literacy and political engagement, what's the connection? *J. Media Lit. Educ.* 9 (1), 79–98.
- Baron, G.L., Bruillard, E., 2001. Une Didactique de l'informatique ? *Revue Française de Pédagogie* 135, 163–172.
- Benkler, Y., Faris, R., Roberts, H., 2018. Network propaganda: manipulation, disinformation, and radicalization in American politics. *Oxf. Scholarsh.* <https://doi.org/10.1093/oso/9780190923624.001.0001>.
- Boulier, D., 2019. *Sociologie du numérique*. Colin, Paris.
- Bourdieu, P., 1979. *La distinction. critique sociale du jugement*. minuit, Paris.
- boyd, D., 2017. Did Media Literacy Backfire? <https://points.datasociety.net/did-media-literacy-backfire-7418c084d88d>.
- Brännback, M., Nikou, S., Bouwman, H., 2017. Value Systems and Intentions to interact in social media: the digital natives. *Telematics Inf.* 34 (4), 365–381.
- Brown, T., Katz, B., 2019. *Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation*. Harper Business, New York.
- Buckingham, D., 2007. Digital media literacies: rethinking media education in the age of the internet. *Res. Comp. Int. Educ.* 2 (1), 43–55.
- Cardon, D., 2015. *A quoi rêvent les algorithmes: nos vies à l'heure des big data*. Le Seuil, Paris.
- Carlsson, U. (Ed.), 2019. *Understanding MIL in the Digital Age, a Question of Democracy*. Nordicom, Goteborg.
- Casilli, A., 2008. *Les liaisons numériques. vers une nouvelle sociabilité?* Seuil, Paris.
- Chakrabarty, D., 2018. Anthropocene time. *Hist. Theor.* 57 (1), 5–32.
- Chartier, R. (Ed.), 1993. *Pratiques de la lecture*. Payot, Paris.
- CNN, 2013. *Citoyen d'une société numérique: accès, littératie, Mediations, pouvoir d'agir: pour une nouvelle politique d'inclusion*. Cnum, Paris.
- Crutzen, P.J., Stoermer, E.F., 2000. The "anthropocene". *Glob. Change Newsl.* 41, 17–18.
- Cukier, K., Mayer Schoenberger, V., 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt, NY.
- De Abreu, B., Mihailidis, P., Lee, A., Melki, J., McDougall, J., 2017. *International Handbook of Media Literacy Education*. Routledge, London.
- Delamotte, E., Liquète, V., Frau-Meigs, D., 2014. La translittératie, ou la convergence des cultures de l'information: supports, contextes et modalités. *Spirale* 53, 145–156.
- Denouël, J., Granjon, F., Aubert, A., 2014. *Médias numériques & participation. Entre engagement citoyen et production de soi*. Mare & Martin, Paris.

- DIGCOMP, 2013, revised 2018, 2022. A Framework for Developing and Understanding Digital Competence in Europe. European Commission Joint Research Centre, Institute for Prospective Technological Studies, Brussels. <https://ec.europa.eu/jrc/en/digcomp>.
- D'Ignazio, C., Klein, L.F., 2020. *Data Feminism*. The MIT Press, Cambridge.
- Drot-Delange, B., Bruillard, E., 2012. Éducation aux TIC, cultures informatique et du numérique: quelques repères historiques. *Études de communication* 38, 69–80.
- Drotner, K., 2011. The cult of creativity: opposition, incorporation, transformation. In: Sefton Green, J., Thomson, P., Bresler, L., Jones, K. (Eds.), *The Routledge International Handbook of Creative Learning*. Routledge, London, pp. 72–80.
- Elmer, G., Langlois, G., Redden, J., 2015. *Compromised Data: From Social Media to Big Data*. Bloomsbury, London.
- Erstad, O., Sefton-Green, J. (Eds.), 2015. *Identity, Community, and Learning Lives in the Digital Age*. Cambridge University Press, Cambridge.
- Ertzscheid, O., 2009. L'homme est un document comme les autres: du World Wide Web au World Life Web. Hermès, Paris, pp. 33–40. CNRS-Éditions.
- Eshet-Alkalai, Y., 2004. Digital literacy: a conceptual framework for survival skills in the digital era. *J. Educ. Multimedia Hypermedia* 13, 93–106.
- Fastrez, P., De Smedt, T., 2012. Une description matricielle des compétences en littératie médiatique. In: Lebrun-Brossard, M., Lacelle, N., Boutin, J.-F. (Eds.), *La littératie médiatique multimodale. De nouvelles approches en lecture-écriture à l'école et hors de l'école*. Presses de l'Université du Québec, Québec, pp. 45–60.
- Fine, M., Tronto, J., 2020. Care goes viral: care theory and research confront the global COVID-19 pandemic. *Int. J. Care Caring* 4 (3), 301–309.
- Frau-Meigs, D., Bossu, A., 2018. La maturité épistémique dans les MOOC: le cas du « Pas à Pas » du projet ECO. *Revue française des sciences de l'information et de la communication* [En ligne] 12. <https://doi.org/10.4000/rfsic.3366>. <http://journals.openedition.org/rfsic/3366>.
- Frau-Meigs, D., Hibbard, L., 2016. *Education 3.0 and Internet Governance: A New Global Alliance for Children and Young People's Sustainable Digital Development* Global Commission on Internet Governance Series. Chatham House, London.
- Frau-Meigs, D., Torrent, J. (Eds.), 2009. *Mapping Media Education Policies in the World: Visions, Programmes and Challenges*. UN-Alliance of Civilizations-UNESCO, New York.
- Frau-Meigs, D., Velez, I., Flores-Michel, 2017. *Public Policies in Media and Information Literacy in Europe: Cross-Country Comparisons*. Routledge, London.
- Frau-Meigs, D., Kotilainen, S., Pathak-Shelat, M., Hoechsmann, M., Poyntz, S.R. (Eds.), 2020. *Handbook of Media Education Research*. Wiley, London.
- Frau-Meigs, D., Osuna-Acedo, S., Marta-Lazo, C. (Eds.), 2021. *MOOCs and the Participatory Challenge. From Revolution to Reality*. Springer, Heidelberg.
- Frau-Meigs, D., 2012. Transliteracy as the new research horizon for media and information literacy. *Media Stud.* 3 (2), 14–27.
- Frau-Meigs, D., 2013. Final Report. ANR translit. <http://www.univ-paris3.fr/anr-translit-243520.kjsp>.
- Frau-Meigs, D., 2015. *Augmented Media and Information Literacy (MIL): how can MIL harness the affordances of digital information cultures?* In: Kotilainen, S., Kupiainen, R. (Eds.), *Media Education Futures*. Göteborg: Clearinghouse, pp. 13–26.
- Frau-Meigs, D., 2017. *Media and Information Literacy (MIL): taking the digital social turn for online freedoms and education 3.0*. In: Waisbord, S., Tumber, H. (Eds.), *Routledge Companion to Media & Human Rights*. Routledge, London.
- Frau-Meigs, D., 2019a. Qui a peur des fake news? *La Documentation française*, Paris.
- Frau-Meigs, D., 2019b. Créativité, éducation aux médias et à l'information, translittératie: vers des humanités numériques. *Quaderni* 98, 87–106.
- Frau-Meigs, D., 2019c. Information disorders: risks and opportunities for digital media and information literacy? *Media Stud.* 10 (19), 11–27. <https://hrcak.srce.hr/ojs/index.php/medijske-studije/view/392>.
- Frau-Meigs, D., 2021a. Media and information literacies for a living world. In: Wasko, J., Swartz, J. (Eds.), *LIFE: A Transdisciplinary Inquiry*. Intellect/University of Chicago Press, Chicago.
- Frau-Meigs, D., 2021b. Lasting lessons learnt from the “fake news” crisis: MIL as the 1st curriculum. In: Thésée, G., Hoechsmann, M., Carr, P.R. (Eds.), *Democracy 2.0*. Brill, Boston.
- Frau-Meigs, 2021c. La nouvelle norme culturelle sur écrans: Le cyber baroque comme fait culturel total. *Nectart* 13.
- Frechette, J., 2002. *Developing Media Literacy in Cyberspace: Pedagogy and Critical Learning for the Twenty-First-Century Classroom*. Praeger, New York.
- Gendler, M., 2016. Datos, Algoritmos, Neutralidad de la Red y Sociedades de Control. In: Ríos Rojo, C. (Ed.), *¿Nuevos paradigmas de vigilancia? Miradas desde América Latina*. Fundación Vía Libre, Córdoba.
- Gibson, J.J., 1977. The theory of affordances. In: Shaw, R., Bransford, J. (Eds.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Erlbaum Associates, Hillsdale, NJ.
- Gilligan, C., 1982. *In a Different Voice: Psychological Theory and Women's Development*. Harvard U. P., Cambridge, Mass.
- Gonnet, J., 1995. *De l'actualité à l'école. Pour des ateliers de démocratie*. Colin, Paris.
- Goody, J., 1986. *The Logic of Writing and the Organization of Society*. Cambridge UP, Cambridge.
- Greene, J.A., Sandoval, W.A., Bråten, I. (Eds.), 2016. *Handbook of Epistemic Cognition*. Routledge, London.
- Hajek, M., Wagenaar, H. (Eds.), 2003. *Deliberative Policy Analysis: Understanding Governance in the Network Society*. Cambridge University Press, Cambridge.
- Hallinan, B., Striplas, T., 2016. Recommended for you: the Netflix Prize and the production of algorithmic culture. *Indiana University New Media Soc.* 18 (1), 117–137.
- Hand, D., 2020. *Dark Data. Why What You Don't Know Matters*. Princeton University Press.
- Hepp, A., 2013. *Cultures of Mediatization*. Polity, Cambridge.
- Hjarvard, S., 2013. *The Mediatization of Culture and Society*. Routledge, London.
- Hobbs, R., 2010. *Digital and Media Literacy: A Plan of Action*. The Aspen Institute, Washington.
- Hobbs, R., 2011. *Digital and Media Literacy, Connecting Culture and Classroom*. Corwin Press, Thousand Oaks.
- Hoechsmann, M., Poyntz, S.R., 2012. *Media Literacies: A Critical Introduction*. Wiley-Blackwell, Malden/Hichester.
- Hoggart, R., 1957. *The Uses of Literacy*. Penguin, London.
- Hohman, F., Kahng, M., Pienta, R., Chau, D.H., 2018. Visual analytics in deep learning: an interrogative survey for the next frontiers. *IEEE Trans. Vis. Comput. Graph.* 25 (8), 2674–2693.
- Jehel, S., 2014. Éducation aux médias comme démarche participative et heuristique. Analyser les modalités de distanciation face aux programmes de télé réalité. In: Loicq, M., Rio, F. (Eds.), *Les jeunes: acteurs des médias. Participation et accompagnements*. Les éditions du Centre d'études sur les jeunes et les médias, Paris.
- Jenkins, H., Purushotma, R., Clinton, K., Weigel, M., Robison, A.J., 2006. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. The John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning, MacArthur Foundation, Chicago.
- Johansson, H., Hvinden, B., 2007. Opening citizenship. Why do we need a new understanding of social citizenship. In: Hvinden, B., Johansson, H. (Eds.), *Citizenship in Nordic Welfare States, Dynamics of Choice, Duties and Participation in a Changing Europe*. Routledge, London, pp. 3–18.
- Kafai, Y.B., Burke, Q., 2015. Constructionist gaming: understanding the benefits of making games for learning. *Educ. Psychol.* 50, 313–334.
- Kahan, D.M., 2013. Ideology, motivated reasoning, and cognitive reflection. *Judgm. Decis. Mak.* 8 (4), 407–424. <https://doi.org/10.2139/ssrn.2182588>.
- Kahan, D.M., 2017. Misconceptions, misinformation, and the logic of identity-protective cognition. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.2973067>.
- Kahne, J., Bowyer, B., 2017. Educating for democracy in a partisan age: confronting the challenges of motivated reasoning and misinformation. *Am. Educ. Res. J.* 54 (1), 3–34.
- Kaplún, M., 1998. *Una pedagogía de la Comunicación*. Ediciones de la Torre, Madrid.
- Kapp, K., 2012. *The Gamification of Learning and Instruction*. Pfeiffer, San Francisco, CA.
- Kellner, D., Share, J., 2005. Toward critical media literacy: core concepts, debates, organizations and policy. *Discourse* 26 (3), 369–386.
- Khun, T., 1970. *The Structure of Scientific Revolutions*. U press, Chicago.
- Kotilainen, S., et al., 2010. *Media Literacy Education. Nordic Perspectives*. Nordicom, Göteborg.
- Krotz, F., 2009. Mediatization. A concept with which to grasp media and societal change. In: Lundby, K. (Ed.), *Mediatization. Concept, Changes, Consequences*. Peter Lang, New York, pp. 21–40.

- Kupiainen, R., Sintonen, S., 2009. *Medialukutaidot, osallisuus, mediakasvatus*. Palmenia. Helsinki University Press, Helsinki [Media literacies, participation, media education].
- Landry, N., 2017. Articuler les dimensions constitutives de l'éducation aux médias. *tic&société* 11 (1), 7–45.
- Lee, A., 2020. The development of media literacy in Chinese societies: from grassroots efforts to institutional support. In: Frau-Meigs, et al. (Eds.), *Handbook of Media Education Research*. Wiley, London.
- Leu, D.J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., Timbrell, N., 2015. The new literacies of online research and comprehension: rethinking the reading achievement gap. *Read. Res. Q.* 50 (1), 37–59.
- Lewandowsky, S., Ecker, U.K.H., Cook, J., 2017. Beyond misinformation: understanding and coping with the “post-truth” era. *J. Appl. Res. Mem. Cognit.* 6 (4), 353–369.
- Liu, A., 2012. Translittératies: le big bang de la lecture en ligne (2011). In: Frau-Meigs, D., Bruillard, E., Delamotte, E. (Eds.), *L'éducation aux cultures de l'information*. Available at: <http://www.ina-sup.com/ressources/dossiers-de-l'audiovisuel/les-e-dossiers-de-l'audiovisuel/e-dossier-leducation-aux-cultures>.
- Livingstone, S., Haddon, L., 2009. EU Kids Online. *Z. für Psychol. / J. Psychol.* 217 (4), 236–239.
- López, A., 2014. *Greening Media Education: Bridging Media Literacy With Green Cultural Citizenship*. Peter Lang, New York.
- Lopez, A., 2020. Expanding ethics to the environment with ecomedia literacy. In: Frau-Meigs, D., Kotilainen, S., Hoerschmann, M., Shelat, M., Poyntz, S. (Eds.), *Handbook of Media Education Research*. Wiley, New York.
- Lukyanenko, R., Parsons, J., Wiersma, Y., Maddah, M., 2016. Participatory design for user-generated content: understanding the challenges and moving forward. *Scand. J. Inf. Syst.* 28 (1), 37–70.
- Lupton, D., 2016. *The Quantified Self*. Polity Press, Cambridge.
- Machin-Mastromatteo, J.D., Lau, J., Virkus, S., 2013. Participatory action research and information literacy: revising an old new hope for research and practice. *Commun. Comput. Inf. Sci.* 397, 48–53.
- Mäkinen, K., 2012. Ohjelmoidut eurooppalaiset. Kansalaisuus ja kulttuuri EU-asiakirjoissa. Jyväskylän yliopisto, Jyväskylä. *Nykykulttuurin tutkimuskeskuksen julkaisuja* 111. [Programmed Europeans: Citizenship and culture in EU documents].
- Manokha, I., 2018. Surveillance: the DNA of Platform Capital—the case of Cambridge Analytica put into perspective. *Theor. Event* 21 (4), 891–913. <https://www.muse.jhu.edu/article/707015>.
- Martin-Barbero, J., 1987. *De los medios a las mediaciones: Comunicación, cultura y hegemonía*. Gustavo Gili, Mexico.
- Masterman, L., 1985. *Teaching the Media*. Comedia Publishing Group, Abingdon, Oxon.
- Maxwell, R., Lager Vestberg, N., Raundalen, J., 2015. *Media and the Ecological Crisis*. Routledge, New York.
- Mezirow, J., 1997. Transformative learning: theory to practice. *N. Dir. Adult Cont. Educ.* 74, 5–12.
- Mihailidis, P., Thevenin, B., 2013. Media literacy as a core competency for engaged citizenship in participatory democracy. *Am. Behav. Sci.* 57 (11), 1611–1622.
- Mihailidis, P., 2018. Civic media literacies: re-imagining engagement for civic intentionality. *Learn. Media Technol.* 43 (2), 152–164.
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L., 2016. The ethics of algorithms: mapping the debate. *Big Data Soc.* 3 (2).
- Moon, J., Liu, Z., 2019. Rich representations for analyzing learning trajectories: systematic review on sequential data analytics in game-based learning research. In: *Data Analytics Approaches in Educational Games and Gamification Systems*. Springer, Singapore, pp. 27–53.
- Nygren, T., Guath, M., Axelsson, C.A., Frau-Meigs, D., 2021. Facing visual fake news with a professional fact-checking tool in education in France, Romania, Spain and Sweden. *Information* 12 (5), 201.
- Palsa, P., 2020. Searching for common ground: multiliteracy and the curricular consistency in the Finnish education system. In: Frau-Meigs, D., Kotilainen, S., Pathak-Shelat, M., Hoerschmann, M., Poyntz, S.R. (Eds.), *Handbook of Media Education Research*. Wiley, London.
- Pédaque, R., 2007. La redocumentarisation du monde. *Bulletin des bibliothèques de France (BBF)* 4, 122–123.
- Perkins, D.N., 1995. L'individu plus. Une vision distribuée de la pensée et de l'apprentissage. *Revue française de pédagogie. Recherches en éducation* 111, 57–72.
- Potter, J.W., 2013. *Media Literacy*, seventh éd. University of California, Santa Barbara.
- Poyntz, S., Frau-Meigs, D., Hoerschmann, M., Kotilainen, S., Shelat, M., 2020. Introduction: media education research in a rapidly changing media environment. In: Frau-Meigs, D., Kotilainen, S., Pathak-Shelat, M., Hoerschmann, M., Poyntz, S.R. (Eds.), *Handbook of Media Education Research*. Wiley, London.
- Przybylski, A.K., Murayama, K., Dehaan, C.R., Gladwell, V., 2013. Motivational, emotional, and behavioral correlates of fear of missing out. *Comput. Hum. Behav.* 29 (4), 1841–1848. ISSN 0747-5632.
- Redpath, S.M., et al., 2018. Games as tools to address conservation conflicts. *Trends Ecol. Evol.* 33 (6), 415–426.
- Rizvi, F., Lingard, B., 2010. *Globalizing Education Policy*. Routledge, London.
- Rizvi, F., 2009. Globalization and policy research in education. In: Ryan, K.E., Cousins, J.B. (Eds.), *The SAGE International Handbook of Educational Evaluation*. Sage, Los Angeles, pp. 3–18.
- Roberge, J., Seyfert, R., 2016. *Algorithmic Cultures Essays on Meaning, Performance and New Technologies*. Routledge, London.
- Rust, S., Monani, S., Cubitt, S., 2016. *Ecomedia: Key Issues*. Routledge, London.
- Salaun, M., 2007. La redocumentarisation, un défi pour les sciences de l'information/re-documentation: a challenge for information studies. *Etudes de communication* 30, 13–23.
- Salomon, G., 1994. *Interaction of Media, Cognition and Learning*. Lawrence Erlbaum, Mahwah.
- Salvador, X.-L. (Ed.), 2019. *Humanités numériques: vers l'institutionnalisation*, Quaderni 98.
- Sanchez, E., Young, S., Jouneau-Sion, C., 2016. Classcraft: from gamification to ludicization of classroom management. *Educ. Inf. Technol.* 20 (2), 497–513.
- Selber, S., 2004. *Multiliteracies for a Digital Age*. SIU Press, Carbondale.
- Serres, A., 2012. Dans le labyrinthe. Évaluer l'information sur internet. C&F éditions, Caen.
- Shute, V., Sun, C., 2020. Games for assessment. In: Plass, J.L., Mayer, R.E., Homer, B.D. (Eds.), *The Handbook of Game-Based Learning*. MIT Press, Cambridge, pp. 491–512.
- Soares, I. de O., 2019. Educommunication landmarks in Latin America: what should be considered in the last 50 years. In: Mateus, J.-C., Andrada, P., Quiroz, M.T. (Eds.), *Media Education in Latin America*. Routledge, London.
- Srnicek, N., 2016. *Platform Capitalism*. Polity Press, Cambridge and Malden.
- Thomas, S., et al., 2007. Transliteracy: Crossing Divides. *First Monday*, p. 12 [online].
- Tucker, J.A., Guess, Barbera, Vaccari, Siegel, Sanovich, Stukal, Nyhan, 2018. Social media, political polarization, and political disinformation: a review of the scientific literature. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3144139>.
- UNESCO, 2011. *Fez Declaration*.
- Van Dijck, J.A., Van Deursen, A.J., 2014. *Digital Skills: Unlocking the Information Society*. Palgrave Macmillan, New York.
- Van Dijck, J., Poell, T., de Waal, M., 2018. *The Platform Society: Public Values in a Connective World*. OUP, Oxford.
- Van Dijck, J., 2014. Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. *Surveill. Soc.* 12 (2), 197–208.
- Verhoeven, L., Perfetti, C., 2008. Advances in text comprehension: model, process and development. *Appl. Cognit. Psychol.* 22, 293–301.
- Vieira, R.M., Tenreiro-Vieira, C., 2016. Fostering scientific literacy and critical thinking in elementary school education. *Int. J. Sci. Math. Edu.* 14 (4), 659–680.
- von Feilitzen, C., Carlsson, U. (Eds.), 2003. *Promote or Protect?: Perspectives on Media Literacy and Media Regulations*. The International Clearinghouse on Children/Youth and Media, Göteborg.
- Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. *Science* 359 (6380), 1146–1151.
- Walker, J., Starosielski, N. (Eds.), 2016. *Sustainable Media: Critical Approaches to Media and Environment*. Routledge, London.

- Wardle, C., Drekscham, H., 2017. Information Disorder. Council of Europe, Strasbourg.
- Williams, R., 1958. Culture and Society. Columbia University Press, New York.
- Wing, J., 2006. Computational thinking. Commun. ACM 49 (3), 33–35.
- Yelland, N., Masters, J., 2007. Rethinking scaffolding in the information age. Comput. Educ. 48, 362–382.
- Yildiz, M.N., Keengwe, J., 2015. Handbook of Research on Media Literacy in the Digital Age. IGI global, NY.

Relevant websites

- ANR TRANSLIT, <http://translit.univ-paris3.fr/en>.
- Council of Europe Digital Citizenship Education, <https://www.coe.int/en/web/digital-citizenship-education/home>.
- Digcomp, <https://ec.europa.eu/jrc/en/digcomp>.
- Savoir Devenir, Savoirdevenir.net/ressources.
- Transliteracies Research Project, <http://transliteracies.english.ucsb.edu/category/research-project/>.
- UNESCO, Communication and Information Division, www.unesco.org/new/en/communication-and-information/media-development/media-literacy/mil-as-composite-concept/.
- UNESCO, Curriculum for Teachers, <http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/media-and-information-literacy-curriculum-for-teachers/>.
- UNESCO, Institute for Information Technologies in Education, <https://iite.unesco.org/mil/>.

Critical literary perspectives on children's and young adult literature

Theresa Rogers and Caroline Hamilton-McKenna, Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	605
Contested borders: childhood, youth, and literature for young people	606
Boundaries of literature for young people	606
Critical literary theories	607
Reader response theories	607
Cultural studies and cultural materialisms	608
Feminisms and children's literature	609
Gender, sex, and sexuality	609
Posthumanism	610
Postcolonial theory	611
Activist lenses: critical race, critical disabilities, and ecocriticism	611
Critical race theory	612
Critical disabilities studies	612
Ecocriticism	613
Conclusion	614
References	614
Children's and YA literature	619

Introduction

Literature provides a space for readers to imagine and gain insight into the human experience. [Rosenblatt \(1938/1995\)](#) eloquently described literary engagement as a means of rich exploration, of “*living through*” (p. 38). Long before literary theorists and teachers of literature fully embraced reader-oriented perspectives, she argued for the role of literature in imagining the lives of others, examining our own lives, and tracing our responses back out into the world. As educators of literacy and children's literature, we similarly advocate fostering students' engagement with texts for young people, drawing on the wide array of critical literary perspectives that have emerged over the past sixty or so years. We find that critical lenses expand potential meanings, enrich classroom dialog, and offer students powerful tools for navigating the interrelated webs of real world and textual spaces. Reading literature through critical frameworks continues to be an exploration—perhaps an even more dynamic one.

Early on in a graduate course for prospective and practicing teachers, writers, and researchers of children's and young adult literature developed by the first author, we introduce critical literary perspectives by asking students what it might look like to analyze a picturebook using a variety of frames. We provide the class with a brief overview of selected key literary theoretical perspectives—beginning with the traditional lens of archetypal analysis, followed by more contemporary lenses, such as reader-response, cultural materialism, feminisms and queer theory, posthumanism, postcolonialism, and critical activist approaches. We then assign one perspective to each of several small groups of students, along with a copy of *Where the Wild Things Are* ([Sendak, 1963](#)). We ask each group to simply “play” with their assigned theory—without expectation of sophisticated responses. Invariably, however, we find that students are quite adept at rereading the picturebook through a critical lens, as illustrated in the following excerpted examples of their work:

Archetypal: From this perspective, we thought Max could be seen the classic hero—setting out on a quest, and returning transformed. The mother can be seen as the nurturer to whom he returns (symbolized by the warm meal waiting).

Feminist: Besides Max himself, we identified two “others” in this text, the wild things and “presence” of the mother. Max wears an androgynous wolf suit that could be seen as a rejection of the gender roles in his mother's house. When he goes off to the land of the wild things—a place where bodies, sexualities, and genders are unclear (e.g., long hair, phallic tails, no clothes)—notions of adult male heterosexual authority are disturbed. But he returns to the more normative gender politics of home.

Postcolonial: When we looked through this lens, we saw Max as one who goes out to conquer others. He becomes the ruler and the wild things are treated as others, as inferior, as less than human, possibly ascribed with some “primitive” qualities. We also saw themes of assimilation, yet with

a paradoxical sense of both helping and hurting in this line: "Oh please don't go, we'll eat you up we love you so."

Continuing in this playful spirit, we spend the remainder of the course embarking on a more in-depth journey of some of the literary lenses that are often applied to Children's (CL) and Young Adult (YAL) Literature in its various formats. For the purposes of this chapter, we selectively share these contemporary literary theoretical perspectives and the work of key scholars in the field to illustrate the richness of these approaches. Along the way, we offer up occasional examples of how we teach these lenses in order to bring into focus some of the key concepts illustrate our underlying pedagogical argument—that critical engagement with literature for young people can enrich and expand students' understandings of both texts and their everyday worlds. But first, it is important to situate this scholarship in light of the contested spaces of childhood and youth, and the slippery boundaries of the literature for and about them.

Contested borders: childhood, youth, and literature for young people

Critical approaches to analyzing literature for young people are inextricably bound to the contested borders of childhood and youth. Over time, competing Western models of childhood have shifted with changing attitudes, understandings, and values—from philosopher John Locke's notion of the child as a tabula rasa to Jean-Jacques Rousseau's vision of an unspoiled, more natural state of being. Central to these debates has been the shared acknowledgment that children are somehow distinct from adults, despite little agreement on what a child, or childhood, actually is (e.g., Ariès, 1962; Sanchez-Eppler, 2013). By the end of the 19th century, the English Romantic poets had fortified the symbolic child as vulnerable beacon of virtue amidst the corrupting influences of the industrialized world (Gubar, 2009)—a pervasive figurative link between youth and innocence still seen today (Aitken, 2001).

The 20th century continued to reshape how children and young people are perceived and studied. Early on, Piaget's (1970) theory of development conceptualized childhood as a sequence of stages—a linear "temporal passage" culminating in adult rationality and autonomy (James and Prout, 2015; Venn and Walkerdine, 1978). Later critiques of developmental models opened up more culturally-situated paradigms (Duane, 2013). Interdisciplinary scholars within the then emerging field of childhood studies retheorized childhood as a complex, contradictory social construction that varies across cultures and historical eras. They also began to question the ambiguous borders of adolescence—the liminal period between childhood and adulthood (James, 1986) described by psychologist Hall (1904) as a rebellious time of "storm and stress" (p. xiii). By the 1990s, adolescence—like childhood—had emerged as an equally elusive and fluid category, encompassing a range of experiences of degrees of independence (Lesko, 2012).

Despite these shifts, Romantic constructions of innocence remain at the heart of debates over childhood and youth. The presumed causal relationship between loss of innocence and increased experience as one transitions from child to adult is still widely used as justification for the watchful maintenance and policing of childhood borders (Marshall, 2012). Moreover, these discursive constructions of youth—those that are often sustained by and in children's literature—tend to limit the normative bounds of innocence to a "White, middle class, and heterosexual" (Marshall, 2012, p. 298) frame.

Boundaries of literature for young people

The history of children's literature has been similarly troubled by a "lack of definitional consensus" (Hintz, 2020) and disputes over what children's literature does, who it is for, and what form and content it presents to readers (e.g., Gubar, 2011; Nodelman, 2008). By the early 20th century, Anglo-American children's literature had blossomed as a profitable niche market and professional field, in part due to the growth of print culture, juvenile divisions of publishing houses, and the popularity of playful (as opposed to didactic) texts purporting to serve the needs of actual children (e.g., Demers, 2009; Mickenberg and Vallone, 2011; Postman, 1982).

Yet as childhood studies and children's literature gained scholarly traction in the 1980s and 1990s, critics began to challenge the fundamental notion that the child was a universally identifiable consumer marked by playful innocence (e.g., Higonnet, 1998; Kincaid, 1992). Perhaps the most influential of these literary censures came in 1984, when Jacqueline Rose asserted that children's literature was "impossible." Whereas child-centered narratives were widely assumed to represent, know, and speak to and for children, Rose argued that children's literature was instead fueled by adult agendas and desires—thereby casting the child "as an outsider to its own process" (Rose, 1984, p. 2).

Rose's contention has since prompted scholars to reckon with how children's literature is defined, regulated, and censored by adults (e.g., Clark, 2003). Some scholars maintain the impossibility of defining it in the first place (e.g., Rudd, 2010). Others, such as Nodelman (2008) and Hunt (2005b), concede the situated cultural meanings of children's literature while still seeking to generalize its unique features. As Hunt (2005b) maintains, children's books "are written for a different audience, with ... different ways of reading; equally, children experience texts in ways which are often unknowable, but which many of us strongly suspect to be very rich and complex" (p. 3).

The disputed bounds of CL likewise infuse definitional borders of YAL. However, debates over innocence and youth play out differently when applied to adolescents and, in turn, the literature marketed to them. While innocence as a childlike attribute exemplifies earnestness and asexuality, innocence in the adolescent sphere is related to conceptions of youth as a time of precariousness and vulnerability that may require guidance, monitoring, and regulation (Marshall, 2012).

These competing views of what constitutes childhood, adolescence, and children's and young adult literature inevitably influence what kinds of texts are produced for young audiences, how that literature is read, and how various literary theoretical perspectives are applied to those works. While this chapter presents some of the ways in which scholars have approached these literary spaces, our review of theoretical lenses and their applications to children's and young adult literature is necessarily selective. For example, we begin in the latter half of 20th century, when examinations of literature for young people shifted to include discussions of readers and cultural discourses and practices. As educators, we also tend to underscore how literary perspectives can deepen the practical ways in which teachers, librarians, and writers explore texts with young readers. Most of all, we are invested in how children and adolescents themselves imaginatively engage with, resist, and negotiate literary worlds in relation to their own lived experiences.

Critical literary theories

Critical literary perspectives are as old as philosophy, religion, and literature. However, it wasn't until the 20th century that literary theory became a profession separate from philosophy, largely adopting secular literature as its subject. In the West, early approaches to modern literary analysis were often structural in nature. These included archetypal analysis, exemplified by Frye's (1957) *Anatomy of Criticism*. Drawing on the work of psychological Carl Jung, Frye (1957) argued that specific literary archetypes connect one literary work to another. Frye identified children's literature in particular as having a common archetypal pattern, in which characters embark on and emerge from a transformative adventure (Nodelman and Reimer, 2003). By the mid-20th century, other approaches, such as New Criticism, grew in popularity—effectively reshaping how literature was read in schools. New Criticism emphasized the connection between the structure and meaning of literary works. Proponents argued that literary studies should focus solely on the text—not on the responses of readers, the author's intention, or any other historical, cultural contexts (Wimsatt and Beardsley 1946, 1954).

By the early 1970s, literary theorists had begun to shift away from structural approaches toward various reader-oriented frameworks instead (e.g., Jauss, 1970). Educational researchers likewise began drawing upon “response to literature” perspectives (Purves and Beach, 1972; Purves et al., 1990); the Dartmouth Conference in particular noted the role of personal expression in literature classrooms (Dixon, 1967). Reader-response perspectives came into fuller prominence later, with the publication of Tompkins' (1980) anthology, *Reader-Response Criticism: From Formalism to Post-Structuralism*, in which contributors (e.g., Bleich, 1978; Fish, 1980; Iser, 1978) turned to readers and their readings as the primary source of meaning. As Tompkins (1980) argues in her introduction, reader-response is not a unified theory, but a collection of post-structural perspectives focused on the interplay of meaning in texts. Despite internal disagreements among reader-response critics, Tompkins (1980) notes, “they are united in one thing: their opposition to the belief that meaning inheres completely and exclusively in the literary text” (p. 201).

In our own course, we often begin discussing this shift from structural or text-oriented approaches to more reader-oriented perspectives with a reading of Williams' (1938/1991) imagist poem, “This is just to say.” Inevitably, our students approach the poem with a variety of tools and frameworks. Many seek examples of symbolism—since that is so often what they have been taught to do—but they also bring their experiences, their appreciation of the imagery, and their sense of the ironic tone (a noticing that often leads to conversations about reading from a gendered or cultural perspective). Through this exercise, we illustrate the many reader-based strategies that guide interpretations of literary works; Williams' poem facilitates a brief glimpse into the myriad ways individual readers make sense of even the smallest of texts. Following this reading, we share a brief history of reader-response theories, from the seminal literary and educational scholarship drawing on Fish (1980) and Rosenblatt (1938/1995) to more contemporary research that takes into account a broader range of texts, cultures, and communities.

Reader response theories

Fish's reader-oriented perspective, articulated in his now well-known work, “Is There a Text in this Class?” (1980), rejects the text as the primary source of meaning in favor of readers. He notes that the play or instability of texts and meaning is not worrisome given the constraints on interpretive strategies or approaches available to readers in any given “interpretive community;” our common beliefs and practices about how meaning is created results in ample amounts of shared meanings. While this view that meaning resides with readers resonated with studies on the psychology of reading and the earlier work on response to literature, it was also a clear rejection of the work of the New Critics and their focus on textual interpretation.

Along with reader-response critics, literary theorists who had long understood this phenomenon were rediscovered around this time, most notably Rosenblatt (1938/1995, 1978/1994). Rosenblatt focused not only on the reader's meaning making, but also on the rich interactions (what she called transactions) between readers' whole worlds and what Tolkien (1964) called the “secondary worlds” of literary texts. Her work memorably evokes a sense of reading as the creation of a dynamic, alternative reality; through their active participation or performance in the alternative worlds of texts, readers co-create “the poem” that results (Rosenblatt, 1978/1994).

The confluent rise reader-response theory, response to literature, and cognitive psychology also reshaped the focus of educational theory and instruction (Rogers, 1999). Key educational studies in the 1980s and 90s drew on reader response theory to examine the dynamic ways in which readers enter and engage in the secondary world(s) of literary texts (Benton, 1983; Enciso,

1992; Galda, 1982) as well as the ways readers (re)construct those worlds both inside and outside of classrooms (e.g., Many and Cox, 1992; Sipe, 2000). A new generation of qualitative response studies (e.g., Hickman, 1983; Hynds, 1989; Rogers, 1991) explored how instructional contexts and discourses create various interpretive spaces for students' responses to literature (e.g., through conversation, performance, art, and writing).

Reader-response criticism has similarly influenced scholarship on how young people engage with visual texts, such as picture-books and graphic novels. Sipe (1998) and Low (2012) draw on Wolfgang Iser's theory of the gaps in literature, along with social semiotic theory (Kress and Van Leeuwen, 2006), to explore how words, images, and other visual features (e.g., gutters and panels) work together to influence meaning making. Drawing on those same frameworks, Styles and Arizpe (2001) provide a particularly striking example of how a range of young readers employed a sophisticated visual grammar to read, across gaps and images, the emotional and visual subtleties of Browne's 1992 picturebook, *Zoo* (see also Arizpe and Styles, 2016). Related work on reader responses to visual texts continues to flourish with multimodal and digital turns in CL and YAL, together with shifts toward post-modernism and popular culture perspectives (Pantaleo, 2008; Serafino and Pantaleo, 2022).

By the 1990s, much of the work on reader-response criticism in education began to give way to cultural perspectives (e.g., Rogers and Soter, 1997). In recognition of the diverse interpretive communities within classrooms—and the influences of cultural backgrounds and perceptions in the process of meaning making—scholars focused more on how cultural discourses circulate across texts and readers, taking into account issues of power and systemic social inequity. As Marxist critic Eagleton (1983) argued, interpretive communities are actually sites of struggle, wherein “certain meanings are elevated by social ideologies to a privileged position” (p. 132). While readers within communities may draw on previously learned conventions of reading and writing, they also are situated culturally (e.g., along race, class, and gender lines; Rabinowitz, 1987)—taking up various or even contradictory identity or subject positions. As such, many CL and YAL scholars began to look to cultural and political perspectives in their analyses of literary works and how young people engage with these texts.

Cultural studies and cultural materialisms

The origins of cultural studies can be traced back to the work of theorists such as Hoggart (1957), who studied literacy as part of a larger network of life practices that could not be separated into “high” and “low” culture (Williams, 1958). In applying this to literature, cultural studies theorists take the perspective that literary works exist within vast network of cultural texts and are not necessarily accorded higher status than, say, a newspaper. Rather, all texts and authors construct and reflect political and cultural meanings. This perspective also reveals literary canons as social and cultural constructs, reflective of the tastes and influences of dominant groups. It would be unwieldy in this space to try to capture all the different strands of cultural studies and cultural materialisms (e.g., New Historicism, Marxist perspectives, the works of The Frankfurt School, British Cultural Materialisms, American Multiculturalism, and Foucauldian theory), so we focus on just a few.

For instance, Marxist literary perspectives, like feminist perspectives (as we discuss in next section), interrogate texts for underlying ideologies related to material inequality and social reality. They link reading with raced, classed, and gendered processes that cannot be fully understood in the absence of analyses of power and privilege. Marxist perspectives examine the ways divisions of labor lead to inequities, inevitably toward revolution. They raise questions about values, relationships between the fate of individuals and societies, and the dialectics of morality and social justice. Zipes (2002) applies Marxist historical materialism to fairy tales by locating their origins in political, class-based struggles rife with exploitation, hunger, and injustice—as well as the human drive to remake the world. It is only when the tales are appropriated by upper class and bourgeois writers, he argues, that they begin to reflect more conservative views (Zipes, 2002).

Cultural literary theorists have also drawn on the writings of Michel Foucault. Though Foucault himself resisted being defined as poststructuralist or postmodern (he openly criticized Marxist theory), his work often echoes Marxist critiques of social structures—linking ideologies to social relations and systems of power. However, inconsistent with Marxist tenets, Foucault believed such change occurs in smaller increments over time, stemming from localized forms of resistance and rupture rather than large-scale revolutions. As he outlined in *Discipline and Punish* (Foucault and Sheridan-Smith, 1979), individuals are spatially and behaviorally regulated and disciplined (i.e., into subordinate, “docile bodies,” p. 135) through institutional and discursive methods of surveillance and punishment. In effect, read through a Foucauldian lens, subjectivities are produced through discourse—though individuals still possess the agency and ability to resist.

With the students in our course, we examine how Foucauldian discourse analysis has been productively taken up by CL and YAL scholars to examine how dystopian works in particular produce ways of thinking and speaking about the child and adolescent in society. Latham (2004), for instance, provides a Foucauldian reading of *The Giver* (Lowry, 1993), tracing the tension between community surveillance and individual resistance to power in the novel as well as highlighting the ambiguity of its ending. Connors (2017) draws on Foucault's concept of the all-seeing Panopticon to analyze *The Hunger Games* (Collins, 2008)—a core text for our course that we use to explore many of the key literary perspectives that follow. Equating Panem with the Connors (2017) argues that the potential of resistance lies within characters' everyday tactics (de Certeau and Rendall, 1984) and subversive acts (see also Hubler's 2014 historical materialist readings of Collins' novels).

These rich perspectives offered by cultural studies and cultural materialisms on the ways ideology and power inhere in texts form the basis of many other critical perspectives on CL and YAL, including the various theoretical underpinnings for feminist frameworks, readings of gender and sexuality, and posthumanist, postcolonial, and critical activist approaches.

Feminisms and children's literature

The revolutionary work by feminist theorists, critics, and writers beginning in the late 1970s (e.g., Gilbert and Gubar, 1979; Heilbrun, 1973; Showalter, 1985) gave rise to the scholarship on the politics of gender and representation applied to children's literature today. Throughout that formative decade, feminist theory gained traction in the US and Europe as readers and critics began to "resist" (Fetterley, 1978) the ubiquity of conventional gender representation in popular texts, particularly in Grimm, Anderson, and Perrault fairy tales (Lieberman, 1972). Writers and illustrators such as Margaret Atwood, Jane Yolen, and Angela Carter took up this revisionist mantle as well, publishing so-called fractured fairy tales and retellings that sought to destabilize the textual perpetuation of patriarchal systems, beliefs, and practices (see Zipes, 1986).

Many have discussed how these early retellings did not so much change gendered scripts as merely reverse them with superficial alterations (Davies, 1989; Paul, 2005; Trites, 1997), such as switching the protagonist from a hero to a heroine. Consequently, Paul (2005) writes, the ostensibly empowered female heroines of the 1970s and 80s "tended to be more like men tricked out in drag" (p. 120). Moreover, revisions of popular fairy tales tended to center on white, cisgender girls and their cisgender male counterparts, mirroring the common critique that second wave feminism was singularly concerned with the plight of white, heterosexual middle-class women (Marshall, 2021).

We use such contentious fairy tale retellings to broach the topic of gendered cultural discourses in our course, focusing on variants of *Little Red Riding Hood* (based on the Perrault and Grimm versions of the tale). Trina Schart Hyman's romantic 1983 retelling of the Grimm version, for instance, privileges discourses of femininity that value obedience and morality: Little Red coyly vows she will "never stray from the path again." In contrast, texts based on the Perrault version tend to emphasize his more cautionary tone. While Coady (1991) creates an ominous world of shadows and claustrophobic framing, Montresor (1991) jars readers with gruesome, wordless illustrations of Little Red entombed in the wolf's bloated stomach. Moon's (2002) urban photo-essay is even more explicit, starkly visualizing, in black-and-white, a Little Red rape narrative. As we explore these retellings with students, we discuss the nuanced, contradictory, and interrelated discursive bounds of girlhood, sexuality, innocence, and violence.

By the late 1980s and early 1990s, children's literature scholarship had blossomed beyond fairy tales to examine how a range of classic and contemporary CL presented grooming sites for "gendered selves" (Davies, 1989, p. x). Work by Vallone (1995), Higgonnet (1998), and Beverly Lyon Clark (e.g., Clark and Higgonnet, 1999) sought to peel back the seemingly innocuous veneer of dichotomous gender constructions in children's texts and material culture. Black feminist theorists and writers had expanded literary feminist analyses to embrace the intersections of gender with other facets of identity, such as race and class (e.g., Hooks, 1992). Influential work by postcolonial theorists (e.g., Spivak, 1988) increased scholarly attention to how normative constructions of gender also interconnect with those of race and power in texts for young people. For example, in their respective mappings of the Victorian "golden age" of Anglophone CL, Reynolds (1990) and Nelson (1991) reveal how immovable constructions of gender and childhood lay at the heart of White colonialist societies.

As feminist readings began to more closely align with the broadening field of gender studies (see Butler, 1990, 1993), new strands of scholarship explored how children's literature contributes to the fluid and relational delineations of multiple gender identities and expressions. Butler's (1990) transformative notion of performativity reconceptualized gender and identity "as something that we 'do' rather than 'are'" (Mallan, 2009, p. 5)—sparking investigations into how gendered subjects are produced or resisted through literary representations. Coupled with a renewed focus on embodiment in texts for young people (Nikolajeva, 2016), such work laid the foundation for queer criticism and interrogations of gender, sex, and sexuality. As Trites (2018) notes, by the turn of the 21st century, children's literature scholarship approached gender as an interplay of "discourse, performance, and the material" (p. 142)—inviting examinations of presumed cis- and heteronormativity in texts for young people.

Gender, sex, and sexuality

Building on influential work by Butler, queer theorist Sedgwick (1985, 1993), and discussions of sexuality, innocence, and desire in children's literature by Rose (1984), Higgonnet (1998), and Kincaid (1992), early queer readings of children's and adolescent literature uncovered the largely ignored presence of gender and sexuality in classic texts. In his role as guest editor for the 1999 edition of the *Lion and the Unicorn*, Kenneth Kidd memorably noted how "sexuality seems to be virtually everywhere in children's literature and media" (p. v), inviting scholars to "hold open the speculative space of sexuality as well as gender, think carefully about the place of race, class, and national identity in that space, and see what we learn" (p. vii). Key to this opening for queer and trans criticism of CL and YAL was muddying cultural divisions between adults and children, maturity and innocence, as well as contentions that sexuality serves as the dividing line between children's and young adult fiction (see Coats, 2011).

For instance, citing scholarship on childhood queerness (Bruhm and Hurley, 2004; Stockton, 2009), Pugh (2011) noted the complex portrayals of sexuality found within popular children's texts, from the Oz (Baum and Denslow, 1900) series to *Harry Potter* (Rowling, 1997). Much of children's literature carries what Pugh (2011) termed an "inherent paradox"—a tense endorsement of both sexual innocence and heterosexuality for its implied child reader. Other queer readings have focused on recognizing the rich albeit unacknowledged history of gay, lesbian, queer, and trans representation. While Kidd (1998) maintains that the emphasis on homosocial relationships in classic American fiction is "decidedly queer" (p. 114), Abate's (2008) explorations of tomboyism trace the cultural evolution of gender-nonconforming characters in texts like *Little Women* (Alcott, 1868/1983); see also Abate and Kidd (2014).

As queer representation in CL and YAL arguably became less ambiguous in the 1990s and 2000s, queer readings also began to focus on identifying to what extent certain patterns and tropes either destabilize or reinscribe hegemonic categories of gender and sexuality. Intersectional feminist, queer, and trans scholars note how children's and young adult literature have long been used as pedagogical tools, instructing youth, parents, and educators on issues of gender, sex, and sexuality (Mallan, 2009; Trites, 2000). As such, new novels can contribute to the increased diversity and complexity of representation—or reify assimilatory norms and values. For example, picturebooks purporting to represent rainbow families for young readers, such as *Heather Has Two Mommies* (Newman and Souza, 1989), have been critiqued as conforming to heteronormative tropes of White, middle-class suburbia (Mason, 2020). According to Taylor (2012), such texts celebrate queer families—as long as they visually and behaviorally reflect “homonormative” standards.

Scholars have similarly observed how the explosion of YAL in the early 2000s featuring LGBTQ+ characters and plotlines mostly still adhered to a conservative YA “coercive tradition” (Kokkola, 2013, p. 5) of punishing or demonizing adolescent sexuality. These issues continue to be especially fraught for trans characters; as Mason (2021) notes, despite the increase in trans representation in North American texts, the dearth of more global and non-Eurocentric trans and genderqueer childhoods in literature still pose challenges to decolonizing efforts in this field. Pini et al. (2018) further underscore how even when trans characters are central to North American texts—as in Katcher's (2009) Stonewall Award-winning novel, *Almost Perfect*—heteronormative, cisnormative, and transphobic perspectives persist as the default and acceptable positions.

Nevertheless, in recent years, scholars have also looked to CL and YAL for hopeful if not idealistic visions of queer figures, spaces, and communities. While some have considered the potential for intersectional queer futurities (e.g., Matos, 2019; Pattee, 2008), others have explored how queer and feminist approaches might interrelate with those of posthumanist lenses (Flanagan, 2011), moving away from post-structural criticism to embrace posthuman ontologies of being.

Posthumanism

Posthumanism is one of the later topics we cover in our course, as CL and YAL scholars frequently combine posthuman lenses with other critical frameworks including feminisms, ecocriticism, and critical disabilities studies. The “post” in posthumanism signals its ontological break from longstanding humanist conceptions of humans as self-contained, exclusive, and bounded (Nayar, 2014). Rather than privileging human subjectivity, posthumanists seek instead to disrupt biological and social hierarchies—focusing on the interrelated, dynamic agencies of organisms and that flow through complex systems and networks (e.g., Pepperell, 1995; Wolfe, 2009). In this way, posthumanists share ecocritical and critical disabilities notions of human minds and bodies “becoming-with” discursive and material worlds (Haraway, 2008); the human is seen not as exceptional or dominant, but instead as interconnected in body, behavior, and mind to other species and the environment (Tarr and White, 2018).

Much of the critical posthumanist approaches to children's literature draws on work by Haraway (1985, 1991, 2003) and Braidotti (1994, 2013, 2019) to explore issues of embodiment and hybridity. Those building on Haraway's (1991) feminist exploration of the positive transgressive potential of cyborgs have focused on the “unstable body of the adolescent and the hybrid bodies of monsters, cyborgs, and other ‘nonhumans’” (Harrison, 2018, p. 7) featured in YA dystopias, science fiction novels, and children's picturebooks and fantasy texts (e.g., Hintz et al., 2013; Jaques, 2015). Many of these scholars have also detailed how contemporary CL and YAL often serve to underscore the destructive potential of technological advancement, depicting the posthuman body as a site of fear (Kokkola, 2018). Ostry (2004) maintains that cyborgs in YA science fiction dystopias such as Anderson's (2002) novel, *Feed*, Westerfeld's (2005) *Uglies* series, and Dickinson's 1988 *Eva* ultimately emphasize the dangers of bioengineered, mechanistic futurities. By satirizing and demonizing the mechanical Other as freakish, unnatural, and artificial, such texts can reify fixed assumptions about the superiority and innate qualities of humans (Jaques, 2021).

Posthuman approaches to CL and YAL thus delve into tensions and cultural anxieties about the murky, unregulated bounds of cyborg futurities. To explore these common, wavering uncertainties in texts for young people with our own students, we pair this lens with a rereading of *The Hunger Games* trilogy (Collins, 2008, 2009, 2010), considering how the figure of the mockingjay in particular might exemplify a posthumanist blurring of human, animal, and machine. As Guanio-Uluru (2017) argues, the mockingjay shifts in symbolic significance as the series catalogs the varying political and social implications of hybrid beings—vacillating between an emblem of love, hope, and loss to a source of destructive, militaristic power.

Other scholars focus more decidedly on the liberatory potential of posthumanist subjects in both classic and contemporary CL and YAL, borrowing from Braidotti's (1994) Deleuzian construction of the nomad (e.g., Flanagan, 2014; Reynolds, 2007). Entities conventionally thought of as nonhuman (e.g., inorganic materials, plants, environments, animals, and toys) thereby become textual sites of embodied and conceptual leakage, signaling opportunities for reinvention. As Tarr and White (2018) argue, posthumanism is “not so much anti-humanism as it is a re-envisioning” (p. xvi)—opening up new possibilities for conceptions of childhood, adolescence, and adulthood.

For instance, in her readings of robots and toys in children's texts such as *Winnie-the-Pooh* (1926) and *Pinocchio* (1883/1996), Jaques (2015) notes that toys not only trouble the boundaries of being, they also expose the artificial bounds of humanity itself. Similarly, disputing Ostry's (2004) readings of contemporary YA science fiction, McCulloch (2016b) notes how Revis's (2011, 2012, 2013) *Across the Universe* trilogy offers “a hopeful empowerment of adolescents that can inspire tomorrow's citizens to form ethical and empathetic relations with others” (p. 75). Contending that posthuman texts invite multiple boundary crossings and queer readings, McCulloch further describes their transgressive potential to rupture parameters of race, ethnicity, sex, and gender in defiance of

globalized hegemony (McCulloch, 2016b). This activist reading of the posthuman has been similarly embraced in recent years by numerous scholars who meld posthumanist frameworks with other critical approaches to embodiment, subjectivity, culture, and the environment, particularly through feminist, queer, and, as we discuss below, ecocritical lenses (e.g., Duckworth and Guanio-Uluru, 2022).

Postcolonial theory

To introduce postcolonial theory in our course, we ask students to read the graphic novel, *Pemmican Wars* (2017), by Métis writer Katherena Vermette—the first volume in her series, *A Girl Called Echo*. The novel follows 13-year-old Echo Desjardins as she struggles with loneliness both at home and at school. As she learns about Métis communities in a history class, she begins to day-dream—time-slipping in and out of the Saskatchewan prairie during the era of fur trade wars in the early 1800s. Our students appreciate the easy interplay of word and text and how themes of dislocation, identity, liminality, and (re)connections to place and heritage are visually invoked.

We choose Vermette's contemporary text to highlight the central tenet of postcolonial theory—that the practices and effects of colonialism are not over, but ongoing, and perpetuated both systemically and through everyday acts and cultural materials (e.g., literary texts). Although the “post” in postcolonialism has been critiqued for seeming to imply otherwise (e.g., Churchill, 2012), it is perhaps helpful to characterize postcolonial readings—and postcolonial texts themselves—as part of an active process to dismantle colonialist ideologies and universalist accounts of the human experience (Ashcroft et al., 1995). As McGillis (2000) notes, “Postcolonialism is a decolonizing activity, rather than a *fait accompli*” (p. 223).

The aims of postcolonial approaches to literary analysis include interrogating the discursive strategies of colonizing societies and reclaiming the history, cultural contributions, and global significance of those individuals and communities who have been erased, silenced, or dismissed (Appleman, 2015). As theorized by Said (1979, 1993), Hulme (1986), Spivak (1988, 1990) and Bhabha (1994), colonialist discourses promulgated in literary texts mobilize certain representations of land and Other in order to naturalize white colonial sovereignty.

Many have explored how Western texts for young people have served and continue to uphold these colonialist regimes and worldviews (Kutzer, 2000; Reimer, 2008; McGillis, 2000). Early evocations of postcoloniality and CL by Rose (1984) and Nodelman (1992) notably compared children to silenced and exploited subjects in relation to the “colonizing” mythologies and interests of adults. Rebutting this view, postcolonial scholar Bradford (2011a) maintains that conflating children with colonized people not only privileges White, middle-class consumers of children's texts, but also deflects attention away from the realities and lasting effects of colonization. In her own extensive analyses, Bradford (2001, 2007) situates postcolonial discussions of CL within white settler literary traditions (e.g., those stemming from the US, Canada, Australia, and South Africa), analyzing novels by non-Indigenous White authors, such as Speare's (1983) *Sign of the Beaver*, to chronicle how colonial meanings permeate even ostensibly anti-colonial texts (Bradford, 2001).

Other postcolonial readings have considered the long-established links between children's literature and nationalist agendas, noting how depictions of “home” frequently uphold paradigms of citizenship bound by markers of race, class, and gender (e.g., Kelen and Sundmark, 2013; Reimer, 2008). Such discussions have extended to explorations of transnationalism, cosmopolitanism, and multiculturalism in children's texts and literacy practices (McCulloch, 2016a; Wu et al., 2013). While some have questioned the literary romanticization of multilingual, hypermobile child elites in an age of globalization (Bradford, 2011b; Snell, 2014), others have drawn on postcolonial theory to examine themes of hybridity and transculturation as characters negotiate national, linguistic, and cultural borders (Brochin and Medina, 2017; Christensen, 2013; González, 2018).

Postcolonial theory has also informed anticolonial and decolonizing strategies in schools and literature classrooms. As Sung et al. (2017) argue, despite the large body of US immigrant-themed narratives published between 2000 and 2006, contemporary novels are still “replete with postcolonial narratives that foster mental images of dominance over immigrant, migrant, and refugee people” (p. 45). Drawing on Ladson-Billing's (1998) call for critical race studies in education, Sinclair (2018) illustrates how the careful selection of texts, such as Older's (2015) YA novel, *Shadowshaper*, might begin to decolonize English literature classrooms where Eurocentric Whiteness is the norm. As Bradford (2021) notes, “postcolonialism now aligns itself with critical race theory and whiteness studies, thus enabling comparative and historicized approaches that consider both national literatures and the global conditions in which they are produced” (p. 153).

Activist lenses: critical race, critical disabilities, and ecocriticism

Reflecting the growth of scholarship and the widespread recognition of CL and YAL across academic fields, a number of critical readings are structured around overtly activist objectives. Many examples, such as those featured at the 2013 *Journal of Language and Literacy Education* inaugural Activist Literacies conference, wed children's literature scholarship with educational aims—considering how a range of youth media “is both a tool of and for activism” (Graff, 2013). As Bradford (2017) maintains, critical reading does not “comprise a body of rules or strategies, but is always situated and contingent, calling on us to be alive to how texts for children position their readers, the ideologies they advocate, and the worlds they imagine” (p. 26). Below are three examples of how activist

readings have emerged through applications of critical race theory, critical disabilities studies, and ecocritical frameworks to texts for young people.

Critical race theory

Since its inception in the 1970s, critical race theory (CRT) has become a core discipline in US law schools, used to interrogate how legal structures and legislation contribute to the systemic oppression of Black and Indigenous peoples, people of color, and immigrants to the United States (e.g., Mills & Unsworth, 2018; Crenshaw, 1995; Crenshaw et al., 2019). From a legal and political activist standpoint, the central objective of a CRT lens is to understand the embedded nature of racism in order to address it through structural change, legal reform, and other democratic interventions. As Delgado and Stefancic (2017) note, “unlike traditional civil rights discourse, which stresses incrementalism and step-by-step progress, critical race theory questions the very foundation of the liberal order” (p. 3).

Building on this legal foundation, Ladson-Billings and Tate (1995) reshaped CRT as a tool for confronting the ongoing oppression and marginalization of groups within and across educational systems and institutions (see also: Ladson-Billings, 1998; Tate, 1997). According to Solórzano (1998), a CRT lens in education foregrounds five defining elements: “(1) The centrality and intersectionality of race and racism; (2) the challenge to dominant ideology (3) the commitment to social justice; (4) the emphasis of experiential knowledge; and (5) the importance of interdisciplinary perspectives” (Osorio, 2018, p. 94). As it has advanced and undergone critiques in education and humanities scholarship, CRT has sparked a number of other racialized “Crits”—such as Black-Crit, LatCrit, AsianCrit, and TribalCrit—that seek to more precisely address and complicate how racism affects and is experienced by intersectional group identities (e.g., Brayboy, 2005; Chae, 2004; Dumas and Ross, 2016; Osorio, 2018; Solórzano and Yosso, 2002; Writer, 2008).

The tenets of CRT in the context of children's literature scholarship remain equally wedded to these social justice objectives, often advocating a critical renegotiation of texts in the interest of what Thomas and Stornaiuolo (2016) term “textual justice.” Many of these explorations seek to detail how specific racialized groups have shaped or have been represented in literature for children and youth (Acevedo, 2015; Bishop, 2007; García, 2021; Reese, 2008)—while at the same time lamenting the dearth of children's publications written by, and for, Indigenous people and people of color. As CL author Myers (2014) opined: “Books transmit values. They explore our common humanity. What is the message when some children are not represented in those books?” (para. 18).

CRT and other “racecrit”-based readings of children's literature, including whiteness studies (e.g., Engles and Kory, 2014b; Sams and Boyd, 2019), examine racist or “colorblind” ideologies and discourses in texts and media (e.g., Cappiccie et al., 2012; Ishizuka and Stephens, 2019); counter-storytelling techniques that reclaim experiences of oppressed groups (e.g., Brooks, 2009; Tulino et al., 2019); and the fusion of CRT-based literary scholarship with transformative, antiracist teaching practices and policies (e.g., Johnson, 2018; Lesley, 2017). For instance, Cueto and Brooks (2019) illuminate the counter-storytelling strategies employed in visual texts—arguing that picturebooks like *I am Enough* (Byers and Bobo, 2018) and *Uptown* (Collier, 2000) dismantle the hegemonic White gaze and “embody the potential to initiate, support, and promote change” (p. 41). Addressing ways graphic life writing might be used as a tool for literature educators, Marshall (2016) similarly notes how certain texts can serve as powerful media for representing diverse realities and demonstrating antiracist curricula, pedagogy, and everyday practices.

Turning attention to the inherent racism of narrative structures, Thomas (2019) pairs CRT with Afrofuturism to reconsider the disparaging metaphorical use of Blackness in contemporary speculative and high fantasy texts. In her discussion of the character Rue from *The Hunger Games* (Collins, 2008), she notes the overwhelmingly negative reaction by some readers and viewers when Amandla Stenberg was cast in the film adaptation. As Thomas argues, the revelation of Rue's Blackness was viewed as transgressive, even unwelcome, precisely because cultural scripts of Whiteness, Blackness, and fantasy genres rely upon a Black/White, Villain/Hero dichotomy. The hostile rejection of Rue by some fans underscores the endemic nature of what Thomas terms “the dark fantastic”—the cyclic rejection of racially-marked characters in service of the presumed White qualities of “goodness, beauty, innocence, and truth.”

Critical disabilities studies

At the heart of disability studies—an interdisciplinary field of academic scholarship and activism—lies the tension between medical and social models of disability (Davis, 1995, 2013). Whereas medical definitions tend to pathologize disability as a personal affliction, the social models that arose during the Civil Rights campaigns of the 1960s differentiated cognitive or physical impairments from the disabling practices imposed by society (Hunt, 1966; Shakespeare, 2013). More recent debates have spawned “interactionist” perspectives, drawing from Butler (1993) and Foucault and Sheridan-Smith (1979) to further define disability as the ongoing production of material bodies as they interact with the institutions, policies, and norms of social environments (Baglieri and Lalvani, 2020; Hall, 2016). As with social models, interactionist views recognize the embodied experiences of individuals while also espousing notions of disability as a political identity for social change.

In keeping with social and interactionist models, applications of critical disability studies frameworks in literature scholarship question normative constructions of bodies and minds (i.e., “ableist” discourses and assumptions; see Garland-Thomson, 1997; Campbell, 2009). Literature scholars often cite the use of disability as a metaphor, or narrative device—marking an individual's deviance against which the “hegemony of normalcy” (Davis, 1995, p. 44) might be narratively constructed and maintained. Terming this phenomenon the “narrative prosthesis,” Mitchell and Snyder (2000) argue that disability is frequently invoked as shorthand

for moral depravity, weakness, or the superhuman capacity for resilience and fortitude (as in the controversial trope of “supercrips;” see Schalk, 2016; Shapiro, 1994). Further, they note, literary canons—featuring work such as Charles Dickens’ *A Christmas Carol* (1843), and John Steinbeck’s *Of Mice and Men* (1937)—shape and normalize dominant narratives of disability more than policies (Snyder and Mitchell, 2007).

The literary tendency to equate disability with deficiency, criminality, or disease has been especially pervasive throughout fictional portraits of childhood. As Hollander (2018) argues, the conflation of disability, morality, and class adopted by eugenicists in the early 20th century led to the construction of dystopic childhood narratives, wherein fictional disabled children embodied fears over the future of nationhood. Ableist portrayals of disabled children being “cured” were also common scripts in 19th and early 20th century novels for children, evident in work such as Johanna Spyri’s *Heidi* (1880/2012), Eleanor Porter’s *Pollyanna* (1913), and Frances Hodgson Burnett’s *The Secret Garden* (1911) (Nikolajeva, 2016). Reductive portraits of disabled childhood have served as symbols of either a crumbling empire or, alternatively, a society on the mend—equating a child’s return to able-bodiedness with a nation’s own potential for an economically productive future.

While analyses of CL and YAL continue to catalog dehumanizing depictions in classic and contemporary work (e.g., Aho and Alter, 2018; Meyer, 2013), critical disabilities scholars also consider the productive and pedagogical possibilities of characterizations of disabilities, even well-worn tropes (e.g., Dunn, 2015; Murray, 2012). Detailing how the “autistic savant” detective in Haddon’s 2003 novel, *The Curious Incident of the Dog in the Night-Time*, is inherently problematic, Resene (2016) observes how that same characterization of a neurodiverse mystery-solver is employed to an empowering effect in Dowd’s (2007) novel, *The London Eye Mystery*. As Resene and others argue, the inclusion of CL and YAL characters with “alternative corporealities” (McRuer, 2006, p. 146) can generate greater empathy and understanding, challenging misconceptions and encouraging readers to share their own experiences (Kersten, 2018).

Wheeler (2019) further argues that protagonists with disabilities produce and reflect profound political and cultural change—redefining whose stories and perspectives are valued in public spaces. Extending this transformative potential to the classroom, Kleekamp and Zapata (2019) note how diverse literary representations of disabilities can provide teachers with tools to create more inclusive learning spaces. To this end, scholars have also fused disability studies lenses with feminist, queer, or critical race approaches, “cripping” explorations of intersectional nonnormative bodies and minds (e.g., Annamma et al., 2018; Kafer, 2013). As Markotić (2018) details in her analyses of Canadian YAL, the social constructions and embodied experiences of gender, queerness, and corporeal difference can be narratively linked in complex, destabilizing ways as adolescent characters negotiate multiple boundaries of identity.

Ecocriticism

Like the other activist lenses under discussion, literary ecocritical frameworks are used to explore themes and patterns of representation as well as to evaluate how texts might guide readers toward activist goals (Buell, 1995, 2005; Garrard, 2012). Inspired by feminist theory and African American studies scholars, Cheryll Glotfelty (e.g., Glotfelty and Fromm, 1996) is routinely credited with shaping the development of “ecocriticism” (a term coined by William Rueckert in 1978) as a literary discipline. As she notes, this interdisciplinary work arose from an interest in how textual analysis of landscape and place might enrich environmental studies and broaden awareness of ecological issues (Glotfelty and Balaev, 2012).

In terms of children’s literature scholarship, ecocritical analyses seek to examine the relationships between young protagonists and their landscapes and environments—at the same time considering the didactic potential of those texts to promote conservationist efforts and ecojustice (e.g., Dobrin and Kidd, 2004; Gaard, 2009). Many scholars cite the ongoing Western association of the Romantic child with the natural world (Reimer, 2013), stemming from 19th century English and Anglo-American novels. According to Darcy (1995), the personal and spiritual growth afforded by rural idylls in works such as Hodgson Burnett’s *The Secret Garden* (1911) and Kenneth Grahame’s *The Wind and the Willows* (1908) underscore the nascent conservationist agendas of literature published in the wake of the Industrial Revolution. These interconnected themes of nostalgia for childhood and the environment reverberate throughout 20th and 21st century children’s texts as well—albeit with more overt appeals for individual or collective action, as exemplified in Seuss’s (1971) *The Lorax* (e.g., Teorey, 2014; Zicht, 1991) and contemporary YA dystopias (Steenkamp, 2015).

Ecocritical scholars are also concerned with how literary conceptions of nature—like childhood—are mutable and inconstant, varying within texts themselves and across genres, eras, and regions and cultures of the world. As Dobrin and Kidd (2004) note, “the ‘wild’ is hardly a neutral concept but an ever shifting rhetorical and political term in need of specific elaborations. Neither are terms like nature or the environment free-floating or self-evident” (p. 2). Focusing mainly on Nordic literatures, the editors and contributors of the 2018 volume *Ecocritical Perspectives on Children’s Texts and Cultures* unpack nature as an umbrella term for a range of esthetics and topics, including landscapes, vegetation, animals, and humans (Goga et al., 2018). In his chronicle of the literary child-nature connection, Hollindale (2021) likewise details how constructions of nature in Anglophone literature have undergone various shifts in tone and intent—from pastoral idylls in *Robinson Crusoe* (Defoe, 1719) to survival stories of savagery or vulnerability in *Lord of the Flies* (Golding, 1954) and *Eva* (Dickinson, 1988).

In recent years, ecocriticism in children’s literature scholarship has broadened alongside other critical cultural lenses (e.g., Duckworth and Guanio-Uluru, 2022). Scholars have combined ecocritical readings with feminist, postcolonial, and posthumanist frameworks examine how texts perpetuate or resist traditionally gendered, imperialist, and anthropocentric perspectives of the natural world. While some have detailed how wilderness coincides with constructions of masculinities and femininities (e.g., Kidd,

2004; Newtown, 2010; Stephens, 2002), others have adopted ecofeminist readings (Gaard et al., 2013) to note how vulnerable bodies of women and youth function as symbols of at-risk ecologies (Curry, 2013; Elbert, 2014).

Ultimately, ecocritical frameworks seek to highlight texts, worldviews, and approaches to reading that inspire social change—and develop youth as local and global ecocitizens (see also: Echterling, 2016a,b; Massey and Bradford, 2011; Moriarty, 2021). For instance, Soulioti (2021) discusses how ecocentric plots and visuals in Greek picturebooks address the real-world destruction of Mediterranean forest ecosystems while still conveying “positive messages on what can plausibly be achieved” (para. 4). Turning to Appalachia and *The Hunger Games* (Collins, 2008), Harde (2018) similarly argues that the continued destruction and rejuvenation of Katniss’s body interrelate with the environmental exploitation of District 12, offering “an eco-political conclusion that might suggest to young readers their own physical and political engagements with their homelands” (p. 69). As climate justice campaigns intensify and U.N. accounts of climate change grow increasingly dire, ecocriticism continues to grapple with literature’s educational and political role in environmental awareness and activism.

Conclusion

Our selective journey through the array of critical frameworks that can be brought to bear in reading CL and YAL necessarily excludes many perspectives that have been productively mobilized to interpret literature for young readers. For instance, scholars have examined children’s and YA literature through Lacanian psychoanalytic lenses (e.g., Coats, 2004; Wilkie-Stibbs, 2002) to explore how the figurative child relates to fractured instantiations of the self in the modern era. An emerging area of critical scholarship, spatiality, draws on the geographical concepts and frameworks of Massey (1994, 2005), Tuan (1977), and others to interrogate the contingency of youth participation and belonging in certain spaces and places—underscoring themes of exclusion and displacement as well as the ways fictional youth negotiate the cultural and physical borders of their worlds (e.g., Glenn, 2019; Hamilton-McKenna and Rogers, 2021; Murray and Cortés-Morales, 2019). And, broadening the notion of textual spaces, scholars interested in transmedia (Jenkins, 2006; Kress, 2010), or cross-media narratives, investigate how the ever-changing, interactive reader engagement with story—as well as the story itself—travel and blur in potentially creative ways across fluid media adaptations and modalities (e.g., Bullen et al., 2018; Hamer, 2017). Many of the students we work with choose to draw on these lenses as well to think about their own scholarship and teaching.

For the final project in our course, we ask students to present analyses of literary works or to develop or adapt literature units for classrooms—encouraging them to draw on the myriad critical perspectives or even weave together frameworks that best fit their chosen texts. One student, a middle grade English teacher, designed a unit on selected speculative “cli-fi” texts for young people, illustrating how reading environmental literature through ecofeminist and postcolonial perspectives might empower students to take productive action toward creating more sustainable, equitable communities. Another teacher of English language learners drew on critical race theory and Thomas’ (2019) treatise, *The Dark Fantastic*, to explore the misrepresentation, silencing, or complete absence of East Asian peoples and cultures in YAL and popular texts. As evidence of the normalized dearth of Asian representation in Anglophone fantasy, he shared his surprise upon discovering a prominent character with East Asian heritage in the *Harry Potter* series (Rowling, 1997). An English teacher revised a grade 12 literature unit in which he taught Cheyenne and Arapaho author Orange’s (2018) novel, *There*. This teacher had noticed the reliance on stereotypes among the students when interpreting the novel, so he proposed incorporating postcolonial and critical race theory perspectives to draw attention to the ways Orange actually subverts stereotypes through counternarratives.

What we find most encouraging are the ways in which these early scholars and teachers readily take up the potential of critical perspectives to deepen the conversations about children’s and young adult literature in their work and classrooms. The critical approaches to literature provide educators with a rich set of tools to more fully engage young people in drawing insightful connections across narratives, their own lives, and their cultural and societal contexts. As we have discussed, critical and activist lenses can also contribute to the ongoing project of facilitating students’ thoughtful considerations of the future—offering up literary and interpretive spaces as sites of creative reimaginings, civic engagement, and social justice.

References

- Abate, M.A., 2008. *Tomboys: A Literary and Cultural History*. Temple University Press.
- Abate, M.A., Kidd, K., 2014. *Over the Rainbow: Queer Children's and Young Adult Literature*. University of Michigan Press.
- Acevedo, M.V., 2015. The portrayal of Puerto Ricans in children's literature. *Bookbird* 53 (2), 4–11. <https://doi.org/10.1353/bkb.2015.0038>.
- Aho, T., Alter, G., 2018. “Just like me, just like you”: narrative erasure as disability normalization in children's picture books. *J. Lit. Cult. Disabil. Stud.* 12 (3), 303–319. <https://doi.org/10.3828/jlcs.2018.24>.
- Aitken, S.C., 2001. *Geographies of Young People: The Morally Contested Spaces of Identity*. Routledge.
- Annamma, S.A., Ferri, B.A., Connor, D.J., 2018. Disability critical race theory: exploring the intersectional lineage, emergence, and potential futures of DisCrit in education. *Rev. Res. Educ.* 42 (1), 46–71. <https://doi.org/10.3102/0091732X18759041>.
- Appleman, D., 2015. *Critical Encounters in Secondary English: Teaching Literary Theory to Adolescents*, third ed. Teachers College Press.
- Ariès, P., 1962. *Centuries of Childhood*. Cape.
- Arizpe, E., Styles, M., 2016. *Children Reading Picturebooks: Interpreting Visual Texts*, second ed. Routledge.
- Ashcroft, B., Griffiths, G., Tiffin, H., 1995. *The Post-colonial Studies Reader*. Routledge.
- Baglieri, S., Lavani, P., 2020. *Undoing Ableism: Teaching About Disability in K-12 Classrooms*. Routledge.

- Benton, M., 1983. Secondary worlds. *J. Res. Dev. Educ.* 16 (3), 68–75.
- Bhabha, H.K., 1994. *The Location of Culture*. Routledge, Oxford and New York.
- Bishop, R.S., 2007. *Free Within Ourselves: The Development of African American Children's Literature*. Greenwood Press.
- Bleich, D., 1978. *Subjective Criticism*. The Johns Hopkins University Press.
- Bradford, C., 2001. *Reading Race: Aboriginality in Australian Children's Literature*. Melbourne University Press.
- Bradford, C., 2007. *Unsettling Narratives: Postcolonial Readings of Children's Literature*. Wilfrid Laurier University Press, Waterloo.
- Bradford, C., 2011a. The case of children's literature: colonial or anti-colonial? *Glob. Stud. Child.* 1 (4), 271–279. <https://doi.org/10.2304/gsch.2011.1.4.271>.
- Bradford, C., 2011b. Children's literature in a global age: transnational and local identities. *Nord. J. ChildLit Aesthet.* 2, 20–34.
- Bradford, C., 2017. The critical reading of children's texts: theories, methodologies, and critiques. In: Johnson, H., Mathis, J., Short, K.G. (Eds.), *Critical Content Analysis of Children's and Young Adult Literature: Reframing Perspective*, pp. 16–27. <https://doi.org/10.4324/9781315651927-6>.
- Bradford, C., 2021. Postcolonial. In: Nel, P., Paul, L., Christensen, N. (Eds.), *Keywords for Children's Literature*, second ed. New York University Press.
- Braidotti, R., 1994. *Nomadic Subjects: Embodiment and Sexual Difference in Contemporary Feminist Theory*. Columbia University Press.
- Braidotti, R., 2013. *The Posthuman*. Polity Press, Cambridge.
- Braidotti, R., 2019. *Posthuman Knowledge*. Polity Press.
- Brayboy, B.M.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37, 425–446.
- Brochin, C., Medina, C., 2017. Critical fictions of transnationalism in Latinx children's literature. *Bookbird* 55 (3), 4–11. <https://doi.org/10.1353/bkb.2017.0036>.
- Brooks, W., 2009. An author as a counter-storyteller: applying critical race theory to a Coretta Scott King Award book. *Child. Lit. Educ.* 40 (1), 33–45. <https://doi.org/10.1007/s10583-008-9065-9>.
- Bruhm, S., Hurley, N. (Eds.), 2004. *Curiouser: On the Queerness of Children*. U of Minnesota Press, Minneapolis.
- Buell, L., 1995. *The Environmental Imagination: Thoreau, Nature Writing, and the Formation of American Culture*. Harvard University Press.
- Buell, L., 2005. *The Future of Environmental Criticism: Environmental Crisis and Literary Imagination*. Blackwell.
- Bullen, E., Rutherford, L., Urquhart, I., 2018. Cinema and the book: intermediality in the invention of Hugo Cabret. In: *The Lion and the Unicorn (Brooklyn)*, vol. 42, pp. 73–87. <https://doi.org/10.1353/uni.2018.0005>.
- Butler, J., 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge.
- Butler, J., 1993. *Bodies that Matter: On the Discursive Limits of "Sex"*. Routledge.
- Campbell, F.K., 2009. *Contours of Ableism: The Production of Disability and Abledness*. Palgrave Macmillan.
- Cappiccie, A., Chadha, J., Lin, M.B., Snyder, F., 2012. Using critical race theory to analyze how Disney constructs diversity: a construct for the baccalaureate human behavior in the social environment curriculum. *J. Teach. Soc. Work* 32 (1), 46–61. <https://doi.org/10.1080/08841233.2012.640252>.
- Chae, H.S., 2004. Talking back to the Asian model minority discourse: Korean-origin youth experiences in high school. *J. Intercult. Stud.* 25, 59–73. <https://doi.org/10.1080/07256860410001687027>.
- Christensen, N., 2013. Contemporary picturebooks in the Nordic countries: concepts of literature and childhood. In: Ommundsen, A.M. (Ed.), *Looking Out and Looking in: National Identity in Picturebooks of the New Millennium*. Novus Forlag, pp. 183–194.
- Churchill, W., 2012. There is nothing "post" about colonialism: on the continuing reality of colonization and the implications of terminological denial. In: *A Decolonizing Encounter: Ward Churchill and Antonia Darder in Dialogue*, vol. 430. Counterpoints, pp. 11–38.
- Clark, B.L., 2003. *Kiddie Lit: The Cultural Construction of Children's Literature in America*. Johns Hopkins University Press.
- Clark, B.L., Higdon, M.R., 1999. *Girls, Boys, Books, Toys: Gender in Children's Literature and Culture*. Johns Hopkins University Press.
- Coats, K.S., 2004. *Looking Glasses and Neverlands: Lacan, Desire, and Subjectivity in Children's Literature*. University of Iowa Press.
- Coats, K.S., 2011. Growing up, in theory. In: Wolf, S.A., Coats, K., Enciso, P., Jenkins, C.A. (Eds.), *The Handbook of Research on Children's and Young Adult Literature*. Routledge, New York, pp. 315–329.
- Connors, S.P., 2017. "I have a kind of power I never knew I possessed:" Surveillance, agency, and the possibility of resistance in YA dystopian fiction. *Study Scrutiny* 2 (2), 1–23. Retrieved from: <https://journals.shareok.org/studyandscrutiny/issue/view/133>.
- Crenshaw, K., 1995. *Critical Race Theory: The Key Writings That Formed the Movement*. New Press.
- Crenshaw, K., Harris, L.C., HoSang, D.M., Lipsitz, G. (Eds.), 2019. *Seeing Race Again: Countering Colorblindness Across the Disciplines*. University of California Press.
- Cueto, D., Brooks, W.M., 2019. Drawing humanity. In: Johnson, H., Mathis, J., Short, K.G. (Eds.), *Critical Content Analysis of Visual Images in Books for Young People*. Routledge, pp. 41–58. <https://doi.org/10.4324/9780429426469-3>.
- Curry, A., 2013. *Environmental Crisis in Young Adult Fiction: A Poetics of the Earth*. Palgrave Macmillan.
- Darcy, J., 1995. The representation of nature in the *Wind in the Willows* and the *Secret Garden*. In: *The Lion and the Unicorn*, vol. 19, pp. 211–221, 2.
- Davies, B., 1989. *Frogs and Snails and Feminist Tales*. Allen & Unwin.
- Davis, L.J., 1995. *Enforcing Normalcy: Disability, Deafness, and the Body*. Verso.
- Davis, L.J., 2013. *The Disability Studies Reader*, fourth ed. Routledge.
- de Certeau, M., Rendall, S., 1984. *The Practice of Everyday Life*. University of California Press.
- Delgado, R., Stefancic, J., 2017. *Critical Race Theory: An Introduction*, third ed. NYU Press, New York.
- Demers, P. (Ed.), 2009. *From Instruction to Delight: An Anthology of Children's Literature to 1850*, third ed. Oxford University Press.
- Dixon, J., 1967. *Growth Through English*. National Association for the Teaching of English.
- Dobrin, S.I., Kidd, K.B. (Eds.), 2004. *Wild Things: Children's Culture and Ecocriticism*. Wayne State University Press.
- Duane, A.M. (Ed.), 2013. *The Children's Table: Childhood Studies and the Humanities*. University of Georgia, Athens.
- Duckworth, M., Guanio-Uluru, L., 2022. *Plants in Children's and Young Adult Literature*. Taylor and Francis. <https://doi.org/10.4324/9781032066356>.
- Dumas, M.J., Ross, K.M., 2016. "Be real Black for me": imagining BlackCrit in education. *Urban Educ.* 51 (4), 415–442. <https://doi.org/10.1177/0042085916628611>.
- Dunn, P.A., 2015. *Disabling Characters: Representations of Disability in Young Adult Literature*. Peter Lang.
- Eagleton, T., 1983. *Literary Theory: An Introduction*. Blackwell.
- Echterling, C., 2016a. How to save the world and other lessons from children's environmental literature. *Child. Lit. Educ.* 47 (4), 283–299. <https://doi.org/10.1007/s10583-016-9290-6>.
- Echterling, C., 2016b. Postcolonial ecocriticism, classic children's literature, and the imperial-environmental imagination in the *Chronicles of Narnia*. *J. Midwest Mod. Lang. Assoc.* 49 (1), 93–117.
- Elbert, M., 2014. At home in nature: negotiating ecofeminist politics in Heidi and Pollyanna. In: Harde, R., Kokkola, L. (Eds.), *Eleanor H. Porter's Pollyanna: A Children's Classic at 100*. University of Mississippi Press, pp. 96–118.
- Enciso, P., 1992. Creating the story world: a case study of a young reader's engagement strategies and stances. In: Many, J., Cox, C. (Eds.), *Reader Stance and Literary Understanding: Exploring the Theories, Research, and Practice*. Ablex, pp. 75–102.
- Engles, T., Kory, F., 2014. "What did she see?": the White gaze and postmodern triple consciousness in Walter Dean Myers's *Monster*. *Child. Lit. Assoc. Q.* 39 (1), 49–67. <https://doi.org/10.1353/chq.2014.0017>.
- Fetterley, J., 1978. *The Resisting Reader: A Feminist Approach to American Fiction*. Indiana University Press.
- Fish, S.E., 1980. *Is There a Text in This Class?: The Authority of Interpretive Communities*. Harvard University Press.

- Flanagan, V., 2011. Girl parts: the female body, subjectivity and technology in posthuman young adult fiction. *Fem. Theor.* 12 (1), 39–53. <https://doi.org/10.1177/1464700110390596>.
- Flanagan, V., 2014. *Technology and Identity in Young Adult Fiction: The Posthuman Subject*. Palgrave Macmillan.
- Foucault, M., Sheridan-Smith, A., 1979. *Discipline and Punish: The Birth of the Prison*. Vintage Books.
- Frye, N., 1957. *Anatomy of Criticism: Four Essays*. Princeton University Press.
- Gaard, G., 2009. Children's environmental literature: from ecocriticism to ecopedagogy. *Neohelicon* 36 (2), 321. <https://doi.org/10.1007/s11059-009-0003-7>.
- Gaard, G.C., Estok, S.C., Oppermann, S. (Eds.), 2013. *International Perspectives in Feminist Ecocriticism*. Routledge. <https://doi.org/10.4324/9780203520840>.
- Galda, L., 1982. Assuming the spectator stance: an examination of the responses of three young readers. *Res. Teach. Engl.* 16 (1), 1–20.
- Garcia, E., 2021. Latina feminist agency: manifestations of a new Mestiza consciousness in Gloria Anzaldúa's children's books. *Child. Lit. Assoc. Q.* 46 (2), 111–124. <https://doi.org/10.1353/chq.2021.0027>.
- Garland-Thomson, R., 1997. *Extraordinary Bodies: Figuring Physical Disability in American Culture and Literature*. Columbia University Press.
- Garrard, G., 2012. *Ecocriticism*, second ed. Routledge, New York.
- Gilbert, S.M., Gubar, S., 1979. *The Mad Woman in the Attic: The Women Writer and the Nineteenth-Century Imagination*. Yale University Press.
- Glenn, W., 2019. Theories of space, place, and navigational identity: turning Inside Out and Back Again in the exploration of immigration. In: Ginsberg, R., Glenn, W. (Eds.), *Engaging With Multicultural YA Literature in the Secondary Classroom*. Routledge, pp. 113–121. <https://doi.org/10.4324/9780429053191>.
- Glotfelty, C., Balaev, M., 2012. The formation of a field: ecocriticism in America—an interview with Cheryll Glotfelty. *PMLA* 127 (3), 607–616. <https://doi.org/10.1632/pmla.2012.127.3.607>.
- Glotfelty, C., Fromm, H., 1996. *The Ecocriticism Reader: Landmarks in Literary Ecology*. The University of Georgia Press.
- Goga, N., Guanilo-Uluru, L., Hallås, B.O., Nymes, A. (Eds.), 2018. *Ecocritical Perspectives on Children's Texts and Cultures: Critical Approaches to Children's Literature*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-90497-9>.
- González, A.B., 2018. *Postcolonial Approaches to Latin American Children's Literature*, first ed. Routledge. <https://doi.org/10.4324/9781315647999>.
- Graff, J.M., 2013. Children's literature as tools of and for activism: reflections of JoLLE's inaugural Activist Literacies conference. *J. Lang. Lit. Educ.* 9 (1), 136–143. Available at <http://jolle.coe.uga.edu/wp-content/uploads/2013/05/Childrens-literature-tool-for-activism.pdf>.
- Guanilo-Uluru, L., 2017. Katniss Everdeen's posthuman identity in Suzanne Collins's hunger games series: free as a mockingjay? *eunesse: Young People, Texts, Cultures* 9 (1), 57–81. <https://doi.org/10.1353/jeu.2017.0012>.
- Gubar, M., 2009. *Artful Dodgers: Reconceiving the Golden Age of Children's Literature*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195336252.001.0001>.
- Gubar, M., 2011. On not defining children's literature. *PMLA* 126 (1), 209–216. <https://doi.org/10.1632/pmla.2011.126.1.209>.
- Hall, G.S., 1904. *Adolescence: Its Psychology and Its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion and Education*. D. Appleton and Company.
- Hall, A., 2016. *Literature and Disability*. Routledge. <https://doi.org/10.4324/9781315726595>.
- Hamer, N., 2017. The design and development of the picture book for mobile and interactive platforms: you get to BE Harold's purple crayon. In: Hamer, N., Nodelman, P., Reimer, M. (Eds.), *More Words About Pictures*. Routledge, pp. 63–80. <https://doi.org/10.4324/9781315621814-5>.
- Hamilton-McKenna, C., Rogers, T., 2021. Reenvisioning space, mobilities and public engagement with young adult literature. *Engl. Teach.* 20 (1), 64–77. <https://doi.org/10.1108/ETPC-06-2020-0055>.
- Haraway, D.J., 1985. A manifesto for cyborgs: science, technology, and socialist feminism in the 1980s. *Social. Rev.* 80, 65–107.
- Haraway, D.J., 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. Free Association.
- Haraway, D.J., 2003. *The Companion Species Manifesto: Dogs, People and Significant Otherness*. Prickly Paradigm Press.
- Haraway, D.J., 2008. *When Species Meet*. University of Minnesota Press.
- Harde, R., 2018. "[I]t's my skin that's paid most dearly": Katniss Everdeen and/as the Appalachian Body. In: Harde, R., Kokkola, L. (Eds.), *The Embodied Child: Readings in Children's Literature and Culture*. Routledge, pp. 57–69.
- Harrison, J., 2018. Posthuman power: the magic of hybridity in the Harry Potter series. *Child. Lit. Assoc. Q.* 43 (3), 325–343. <https://doi.org/10.1353/chq.2018.0037>.
- Heilbrun, C.G., 1973. *Toward a Recognition of Androgyny*. Knopf.
- Hickman, J., 1983. Everything considered: response to literature in an elementary school setting. *J. Res. Dev.* 16 (3), 8–13.
- Higonnet, A., 1998. *Pictures of Innocence: The History and Crisis of Ideal Childhood*. Thames & Hudson.
- Hintz, C., 2020. *Children's Literature*. Routledge.
- Hintz, C., Basu, B., Broad, K.R. (Eds.), 2013. *Contemporary Dystopian Fiction for Young Adults: Brave New Teenagers*. Routledge.
- Hoggart, R., 1957. *The Uses of Literacy: Aspects of Working-Class Life, With Special References to Publications and Entertainments*. Chatto and Windus.
- Hollander, A., 2018. Liberty in the age of eugenics: non-normative bodies in Fabian socialist children's fiction. In: Harde, R., Kokkola, L. (Eds.), *The Embodied Child*. Routledge, pp. 127–140.
- Hollindale, P., 2021. *Nature*. In: Nel, P., Paul, L., Christensen, N. (Eds.), *Keywords for Children's Literature*, second ed. New York University Press.
- Hooks, B., 1992. *Black Looks: Race and Representation*. South End Press.
- Hubler, A.E., 2014. *Little Red Readings: Historical Materialist Perspectives on Children's Literature*. University Press of Mississippi.
- Hulme, P., 1986. *Colonial Encounters: Europe and the Native Caribbean, 1492–1797*. London.
- Hunt, P., 1966. *Stigma: The Experience of Disability*. G. Chapman.
- Hunt, P., 2005a. Introduction: the expanding world of children's literature studies. In: Hunt, P. (Ed.), *Understanding Children's Literature: Key Essays From the Second Edition of the International Companion Encyclopedia of Children's Literature*. Routledge, pp. 1–14.
- Hunt, P. (Ed.), 2005b. *Understanding Children's Literature: Key Essays From the Second Edition of the International Companion Encyclopedia of Children's Literature*. Routledge.
- Hynds, S., 1989. Bringing life to literature and literature to life: social constructs and contexts of four adolescent readers. *Res. Teach. Engl.* 23 (1), 30–61.
- Iser, W., 1978. *The Act of Reading: A Theory of Aesthetic Response*. Johns Hopkins University Press.
- Ishizuka, K., Stephens, R., 2019. The cat is out of the bag: orientalism, anti-blackness, and white supremacy in Dr. Seuss's children's books. *Res. Divers. Youth Lit.* 1 (2). <https://sophia.stkate.edu/rdyl/vol1/iss2/4>.
- James, A., 1986. Learning to belong: the boundaries of adolescence. In: Cohen, A.P. (Ed.), *Symbolising Boundaries: Identities and Difference in British Cultures*. Manchester University Press.
- James, A., Prout, A. (Eds.), 2015. *Constructing and Reconstructing Childhood: Contemporary Issues in the Sociological Study of Childhood*, third ed. Routledge. <https://doi.org/10.4324/9781315745008>.
- Jaques, Z., 2015. *Children's Literature and the Posthuman: Animal, Environment, Cyborg*. Taylor and Francis. <https://doi.org/10.4324/9780203581490>.
- Jaques, Z., 2021. *Posthuman*. In: Nel, P., Paul, L., Christensen, N. (Eds.), *Keywords for Children's Literature*, second ed. New York University Press.
- Jauss, H.R., 1970. Literary history as a challenge to literary theory. *New Lit. Hist.* 2 (1), 7–37. <https://doi.org/10.2307/468585>.
- Jenkins, H., 2006. *Convergence Culture: Where Old and New Media Collide*. New York University Press.
- Johnson, L.L., 2018. Where do we go from here? Toward a critical race English education. *Res. Teach. Engl.* 53 (2), 102–124.
- Kafer, A., 2013. *Feminist, Queer, Crip*. Indiana University Press, Bloomington.
- Kelen, C., Sundmark, B., 2013. *The Nation in Children's Literature: Nations of Childhood*. Routledge.
- Kersten, S., 2018. "We are just as confused and lost as she is": the primacy of the graphic novel form in exploring conversations around deafness. *Child. Lit. Educ.* 49 (3), 282–301. <https://doi.org/10.1007/s10583-017-9323-9>.

- Kidd, K., 1998. Introduction: lesbian/gay literature for children and young adults. *Child. Lit. Assoc. Q.* 23 (3), 114–119. <https://doi.org/10.1353/chq.0.1284>.
- Kidd, K., 1999. Editor's introduction. In: *The Lion and the Unicorn*, vol. 23, pp. v–viii, 3.
- Kidd, K., 2004. *Making American Boys: Boyology and the Feral Tale*. University of Minnesota Press.
- Kincaid, J., 1992. *Child-Loving: The Erotic Child and Victorian Culture*. Routledge.
- Kleekamp, M.C., Zapata, A., 2019. Interrogating depictions of disability in children's picturebooks. *Read. Teach.* 72 (5), 589–597. <https://doi.org/10.1002/trtr.1766>.
- Kokkola, L., 2013. Fictions of Adolescent Carnality: Sexy Sinners and Delinquent Deviants. John Benjamins.
- Kokkola, L., 2018. The embodied child: an introduction. In: Harde, R., Kokkola, L. (Eds.), *The Embodied Child: Readings in Children's Literature and Culture*. Routledge, pp. 1–20. <https://doi.org/10.4324/9781315101262>.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge.
- Kress, G., van Leeuwen, T., 2006. *Reading Images: The Grammar of Visual Design*. Routledge.
- Kutzer, M.D., 2000. *Empire's Children: Empire and Imperialism in Classic British Children's Books*. Garland.
- Ladson-Billings, G., 1998. Just what is critical race theory and what's it doing in a nice field like education? *Int. J. Qual. Stud. Educ.* 11 (1), 7–24. <https://doi.org/10.1080/095183998236863>.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Latham, D., 2004. Discipline and its discontents: a foucauldian reading of the giver. *Child. Lit. Stud.* 32 (1), 134–151. <https://doi.org/10.1353/chl.2004.0017>.
- Lesko, N., 2012. *Act Your Age: A Cultural Construction of Adolescence*, second ed. Routledge. <https://doi.org/10.4324/9780203121580>.
- Lesley, N., 2017. *Fictions of Integration: American Children's Literature and the Legacies of Brown v. Board of Education*. Taylor and Francis. <https://doi.org/10.4324/9781315472294>.
- Lieberman, M.R., 1972. "Some day my prince will come": female acculturation through the fairy tale. *Coll. Engl.* 34, 383–395.
- Low, D.E., 2012. "Spaces invested with content": crossing the "gaps" in comics with readers in schools. *Child. Lit. Educ.* 43 (4), 368–385. <https://doi.org/10.1007/s10583-012-9172-5>.
- Mallan, K., 2009. *Gender Dilemmas in Children's Fiction*. Palgrave Macmillan UK. <https://doi.org/10.1057/9780230244559>.
- Many, J., Cox, C., 1992. *Reader Stance and Literary Understanding: Exploring the Theories, Research, and Practice*. Ablex.
- Markotić, N., 2018. You're queer? That's so lame!: queering disability in Brian Francis's *Fruit: a Novel about a Boy and His Nipples* and Mariko Tamaki's *(You) Set Me on Fire*. *J. Lit. Cult. Disabil. Stud.* 12 (3), 337–352. <https://doi.org/10.3828/jlcs.2018.26>.
- Marshall, E., 2012. Innocence. In: Lesko, N., Talburt, S. (Eds.), *Keywords in Youth Studies: Tracing Affects, Movements, Knowledges*, pp. 295–299. <https://doi.org/10.4324/9780203805909>.
- Marshall, E., 2016. Counter-storytelling through graphic life writing. *Lang. Arts* 94 (2), 79–93.
- Marshall, E., 2021. Gender. In: Nel, P., Paul, L., Christensen, N. (Eds.), *Keywords for Children's Literature*, second ed. New York University Press.
- Mason, D., 2020. What having two mommies looks like now: queer picture books in the twenty-first century. In: op de Beeck, N. (Ed.), *Literary Cultures and Twenty-First-Century Childhoods*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-32146-8_6.
- Mason, D., 2021. Trans. In: Nel, P., Paul, L., Christensen, N. (Eds.), *Keywords for Children's Literature*, second ed. New York University Press.
- Massey, D., 1994. *Space, Place, and Gender*. University of Minnesota Press.
- Massey, D., 2005. *For Space*. Sage.
- Massey, G., Bradford, C., 2011. Children as ecocitizens: ecocriticism and environmental texts. In: Mallan, K., Bradford, C. (Eds.), *Contemporary Children's Literature and Film: Engaging With Theory*. Palgrave Macmillan, pp. 109–126.
- Matos, A.D., 2019. A narrative of a future past: historical authenticity, ethics, and queer Latinx futurity in Aristotle and Dante discover the secrets of the universe. *Child. Lit.* 47 (1), 30–56. <https://doi.org/10.1353/chl.2019.0003>.
- McCulloch, F., 2016a. Contemporary British Children's Fiction and Cosmopolitanism. Routledge. <https://doi.org/10.4324/9781315738512>.
- McCulloch, F., 2016b. "No longer just human": the posthuman child in Beth Revis's across the universe trilogy. *Child. Lit. Assoc. Q.* 41 (1), 74–92. <https://doi.org/10.1353/chq.2016.0008>.
- McGillis, R., 2000. Introduction. In: McGillis, R. (Ed.), *Voices of the Other: Children's Literature and the Postcolonial Context*. Routledge, pp. xix–xxii.
- McRuer, R., 2006. *Crip Theory: Cultural Signs of Queerness and Disability*. New York University Press.
- Meyer, A.E., 2013. "But she's not retarded": contemporary adolescent literature humanizes disability but marginalizes intellectual disability. *Child. Lit. Assoc. Q.* 38 (3), 267–283. <https://doi.org/10.1353/chq.2013.0035>.
- Mickenberg, J., Vallone, L. (Eds.), 2011. *The Oxford Handbook of Children's Literature*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195379785.001.0001>.
- Mills, K.A., Unsworth, L., 2018. The multimodal construction of race: a review of critical race theory research. *Lang. Educ.* 32 (4), 313–332. <https://doi.org/10.1080/09500782.2018.1434787>.
- Mitchell, D.T., Snyder, S., 2000. *Narrative Prosthesis: Disability and the Dependencies of Discourse*. University of Michigan Press.
- Moriarty, S., 2021. Modeling environmental heroes in literature for children: stories of youth climate activist Greta Thunberg. In: *The Lion and the Unicorn*, vol. 45, pp. 192–210, 2.
- Murray, S., 2012. From Virginia's sister to Friday's silence: presence, metaphor, and the persistence of disability in contemporary writing. *J. Lit. Cult. Disabil. Stud.* 6 (3), 241–258.
- Murray, L., Cortés-Morales, S., 2019. *Children's Mobilities: Interdependent, Imagined, Relational*. Palgrave Macmillan UK. <https://doi.org/10.1057/978-1-137-52114-9>.
- Myers, W.D., March 15, 2014. Where are the people of color in children's books? The New York Times. <https://www.nytimes.com/2014/03/16/opinion/sunday/where-are-the-people-of-color-in-childrens-books.html?smid=url-share>.
- Nayar, P.K., 2014. *Posthumanism*. Polity.
- Nelson, C., 1991. *Boys Will Be Girls: The Feminist Ethic and British Children's Fiction, 1857–1917*. Rutgers University Press.
- Newtown, M., 2010. *Savage Girls and Wild Boys: A History of Feral Children*. Faber and Faber.
- Nikolajeva, M., 2016. Recent trends in children's literature research: return to the body. *Int. Res. Child. Lit.* 9 (2), 132–145. <https://doi.org/10.3366/ircl.2016.0198>.
- Nodelman, P., 1992. The other: orientalism, colonialism, and children's literature. *Child. Lit. Assoc. Q.* 17 (1), 29–35.
- Nodelman, P., 2008. *The Hidden Adult: Defining Children's Literature*. Johns Hopkins University Press.
- Nodelman, P., Reimer, M., 2003. *The Pleasures of Children's Literature*, third ed. Allyn and Bacon.
- Osorio, S.L., 2018. Border stories: using critical race and Latino critical theories to understand the experiences of Latino/a children. *Race Ethn. Educ.* 21 (1), 92–104. <https://doi.org/10.1080/13613324.2016.1195351>.
- Ostry, E., 2004. "Is he still human? Are you?": young adult science fiction in the posthuman age. In: *The Lion and the Unicorn*, vol. 28, pp. 222–246. <https://doi.org/10.1353/uni.2004.0024>, 2.
- Pantaleo, S., 2008. *Exploring Student Response to Contemporary Picturebooks*. University of Toronto Press, Toronto, ON.
- Pattee, A.S., 2008. Sexual fantasy: the queer utopia of David Levithan's *Boy meets boy*. *Child. Lit. Assoc. Q.* 33 (2), 156–171. <https://doi.org/10.1353/chq.0.0003>.
- Paul, L., 2005. Feminism revisited. In: Hunt, P. (Ed.), *Understanding Children's Literature*, second ed. Routledge, pp. 114–127.
- Pepperell, R., 1995. *The Post-human Condition*. Intellect Books.
- Piaget, J., 1970. Piaget's theory. In: Carmichael, L., Mussen, P.H. (Eds.), *Carmichael's Manual of Child Psychology*, third ed., vol. 1. Wiley, pp. 703–732.
- Pini, B., Keys, W., Riggs, D.W., 2018. Transphobic tropes and young adult fiction: an analysis of Brian Katcher's *Almost Perfect*. In: *The Lion and the Unicorn*, vol. 42, pp. 57–72. <https://doi.org/10.1353/uni.2018.0004>, 1.
- Postman, N., 1982. *The Disappearance of Childhood*. Delacorte Press.

- Pugh, T., 2011. *Innocence, Heterosexuality, and the Queerness of Children's Literature*. Routledge.
- Purves, A.C., Beach, R., 1972. *Literature and the Reader: Research in Response to Literature, Reading Interests, and the Teaching of Literature*. National Council of Teachers of English.
- Purves, A.C., Rogers, T., Soter, A., 1990. *How Porcupines Make Love II: Readers, Texts, Cultures in the Response-Based Literature Classroom*. Longman.
- Rabinowitz, P.J., 1987. *Before Reading: Narrative Conventions and the Politics of Interpretation*. Cornell University Press.
- Reese, D., 2008. Indigenizing children's literature. *J. Lang. Lit. Educ.* 4 (2), 59–72.
- Reimer, M., 2008. Home Words: Discourses of Children's Literature in Canada. Wilfrid Laurier University Press, Waterloo, ON.
- Reimer, M., 2013. The child of nature and the home child. *Jeunesse* 5 (2), 1–16. <https://doi.org/10.1353/jeu.2013.0012>.
- Resene, M., 2016. A "curious incident": representations of autism in children's detective fiction. In: *The Lion and the Unicorn*, vol. 40, pp. 81–99. <https://doi.org/10.1353/unl.2016.0000>, 1.
- Reynolds, K., 1990. *Girls Only?: Gender and Popular Children's Fiction in Britain*. Harvester Wheatsheaf, pp. 1880–1910.
- Reynolds, K., 2007. *Radical Children's Literature: Future Visions and Aesthetic Transformations in Juvenile Fiction*. Palgrave Macmillan.
- Rimmerede, H.E., 2018. "Who will save us from the rabbits?" Problematizing nature in the anthropocene. In: Goga, N., Guanio-Uluru, L., Hallås, B.O., Nyrnes, A. (Eds.), *Ecocritical Perspectives on Children's Texts and Cultures: Critical Approaches to Children's Literature*. Springer International, pp. 207–222. https://doi.org/10.1007/978-3-319-90497-9_13.
- Rogers, T., 1991. Students as literary critics: the interpretive experiences, beliefs, and processes of ninth-grade students. *J. Read. Behav.* 23 (4), 391–423. <https://doi.org/10.1080/10862969109547751>.
- Rogers, T., 1999. Literary theory and children's literature: interpreting ourselves and our worlds. *Theor. Pract.* 38 (3), 138–146. <https://doi.org/10.1080/00405849909543845>.
- Rogers, T., Soter, A.O., 1997. *Reading Across Cultures: Teaching Literature in a Diverse Society*. Teachers College Press.
- Rose, J., 1984. *The Case of Peter Pan or the Impossibility of Children's Fiction*. MacMillan.
- Rosenblatt, L.M., 1994. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Southern Illinois University Press. Original Work Published 1978.
- Rosenblatt, L.M., 1995. *Literature as exploration*. In: *Modern Language Association of America*, fifth ed. First Edition Published 1938.
- Rudd, D. (Ed.), 2010. *The Routledge Companion to Children's Literature*. Routledge.
- Said, E.W., 1979. *Orientalism*. Vintage Books.
- Said, E.W., 1993. *Culture and Imperialism*. Vintage.
- Sams, B., Boyd, A.S., 2019. Parties, pranks, and privilege: reading *Looking for Alaska* through the lens of critical whiteness. In: Malo-Juvera, V., Hill, C. (Eds.), *Critical Explorations of Young Adult Literature*. Routledge, pp. 202–214. <https://doi.org/10.4324/9780429322914>.
- Sánchez-Eppler, K., 2013. In the archives of childhood. In: Duane, A.M. (Ed.), *The Children's Table: Childhood Studies and the Humanities*. University of Georgia, Athens, pp. 213–237.
- Schalk, S., 2016. Reevaluating the supercrip. *J. Lit. Cult. Disabil. Stud.* 10 (1), 71–86.
- Sedgwick, E.K., 1985. *Between Men: English Literature and Male Homosocial Desire*. Columbia University Press.
- Sedgwick, E.K., 1993. *Tendencies*. Duke University Press.
- Serafino, F., Pantaleo, S., 2022. *The Evolution of Children's Literature: Metafictional Devices and Digital Mediation (This Volume)*.
- Shakespeare, T., 2013. The social model of disability. In: Davis, L. (Ed.), *The Disability Studies Reader*, fourth ed. Routledge, pp. 214–221.
- Shapiro, J.P., 1994. *No Pity: People With Disabilities Forging a New Civil Rights Movement*. Times Books.
- Showalter, E. (Ed.), 1985. *New Feminist Criticism: Essays on Women, Literature, and Theory*. Patheon.
- Sinclair, M., 2018. Decolonizing ELA: confronting privilege and oppression in textual spaces. *Engl. J.* 107 (6), 89–94.
- Sipe, L.R., 1998. How picture books work: a semiotically framed theory of text-picture relationships. *Child. Lit. Educ.* 29 (2), 97.
- Sipe, L.R., 2000. The construction of literary understanding by first and second graders in oral response to picture storybook read-alouds. *Read. Res. Q.* 35 (2), 252–275.
- Snell, H., 2014. Outward bound: adventures in cross-cultural reading and global citizenship in North American young adult literatures. *Child. Lit. Assoc. Q.* 39 (2), 252–274. <https://doi.org/10.1353/chq.2014.0031>.
- Snyder, S., Mitchell, D., 2007. Disability haunting in American poetics. *J. Lit. Cult. Disabil. Stud.* 1 (1), 1–12. <https://doi.org/10.3828/jlcsd.1.1.2>.
- Solórzano, D., 1998. Critical race theory, racial and gender microaggressions, and the experiences of Chicana and Chicano scholars. *Int. J. Qual. Stud. Educ.* 11, 121–136.
- Solórzano, D., Yosso, T., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8, 23–44.
- Soulioti, D., 2021. Ecocritical insights: contemporary concerns about forest ecosystems in a Greek picturebook. *Child. Lit. Educ.* <https://doi.org/10.1007/s10583-021-09451-y>.
- Spivak, G.C., 1988. Can the subaltern speak? In: Nelson, C., Grossberg, L. (Eds.), *Marxism and the Interpretation of Culture*. University of Illinois Press, pp. 21–78.
- Spivak, G.C., 1990. *The Post-colonial Critic: Interviews, Strategies, Dialogues*. Routledge, New York, NY.
- Steenkamp, E., 2015. Sinister ecology: space, environmental justice, and belonging in Jenny Robson's *Savannah 2116 AD*. In: Cecire, M.S., Field, H., Finn, K.M., Roy, M. (Eds.), *Space and Place in Children's Literature, 1789 to the Present*. Ashgate, UK, pp. 95–110. <https://doi.org/10.4324/9781315610108>.
- Stephens, J., 2002. *Ways of Being Male: Representing Masculinities in Children's Literature and Film*. Routledge.
- Stockton, K.B., 2009. *The Queer Child: Or Growing Sideways in the Twentieth Century*. Duke University Press.
- Styles, M., Arizpe, E., 2001. A gorilla with "grandpa's eyes": how children interpret visual texts - a case study of Anthony Browne's "zoo". *Child. Lit. Educ.* 32 (4), 261.
- Sung, Y.K., Fahrenbruck, M.L., López-Robertson, J., 2017. Using intertextuality to unpack representations of immigration in children's literature. In: Johnson, H., Mathis, J., Short, K.G. (Eds.), *Critical Content Analysis of Children's and Young Adult Literature*. Routledge, pp. 44–60. <https://doi.org/10.4324/9781315651927>.
- Tarr, A., White, D.R., 2018. *Posthumanism in Young Adult Fiction: Finding Humanity in a Posthuman World*. University Press of Mississippi.
- Tate, W., 1997. Critical race theory and education: history, theory, and implications. *Rev. Res. Educ.* 22, 195–247.
- Taylor, N., 2012. U.S. children's picture books and the homonormative subject. *J. LGBT Youth* 9 (2), 136–152.
- Teorey, M., 2014. The Lorax and Wallace Stegner: inspiring children's environmental activism. *Child. Lit. Educ.* 45 (4), 324–339. <https://doi.org/10.1007/s10583-014-9222-2>.
- Thomas, E.E., 2019. *The Dark Fantastic: Race and the Imagination From Harry Potter to the Hunger Games*. New York University Press.
- Thomas, E.E., Stornaiuolo, A., 2016. Restorying the self: bending toward textual justice. *Harv. Educ. Rev.* 86 (3), 313–338. <https://doi.org/10.17763/1943-5045-86.3.313>.
- Tolkien, J.R.R., 1964. *Tree and Leaf*. Unwin Books.
- Tompkins, J.P., 1980. *Reader-Response Criticism: From Formalism to Post-structuralism*. Johns Hopkins University Press.
- Trites, R.S., 1997. *Waking Sleeping Beauty: Feminist Voices in Children's Novels*. University of Iowa Press.
- Trites, R.S., 2000. *Disturbing the Universe: Power and Repression in Young Adult Literature*. University of Iowa Press.
- Trites, R.S., 2018. *Twenty-First Century Feminisms in Children's and Adolescent Literature*. University of Mississippi Press.
- Tuan, Y., 1977. *Space and Place: The Perspective of Experience*. University of Minnesota Press.
- Tulino, D., Krishnamurthy, S., Fall, M., Browne, S., 2019. Resisting anti-blackness through counternarratives. *Engl. J.* 109 (2), 32–38.
- Vallone, L., 1995. *Disciplines of Virtue: Girls Culture in the Eighteenth and Nineteenth Centuries*. Yale University Press.
- Venn, C., Walkerdine, V., 1978. The acquisition and production of knowledge: Piaget's theory reconsidered. *Ideol. Conscious.* 3, 67–94.
- Wheeler, E.A., 2019. *Handiland: The Crippiest Place on Earth*. University of Michigan Press.
- Wilkie-Stibbs, C., 2002. *The Feminine Subject in Children's Literature*. Routledge. <https://doi.org/10.4324/9781315023359>.
- Williams, R., 1958. *Culture and Society*. Chatto & Windus, pp. 1780–1950.
- Williams, W.C., 1991. *The Collected Poems: Volume 1, 1909–1939*. New Directions. Original Poem Published 1938.

- Wimsatt, W.K., Beardsley, M.C., 1946. The intentional fallacy. *Sewanee Rev.* 54 (3), 468–488.
- Wimsatt, W.K., Beardsley, M.C., 1954. *The Verbal Icon: Studies in the Meaning of Poetry*. Methuen.
- Wolfe, C., 2009. What is Posthumanism? University of Minnesota Press.
- Writer, J.H., 2008. Unmasking, exposing, and confronting: critical race theory, tribal critical race theory and multicultural education. *Int. J. Multicult. Educ.* 10 (2), 1–15.
- Wu, Y., Mallan, K., McGillis, R. (Eds.), 2013. *(Re)imagining the World*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-36760-1>.
- Zicht, J., 1991. In pursuit of the Lorax: who's in charge of the last Truffula seed? *EPA J.* 17 (4), 27–30.
- Zipes, J. (Ed.), 1986. *Don't Bet on the Prince: Contemporary Feminist Fairy Tales in North America and England*. Gower.
- Zipes, J., 2002. *Breaking the Magic Spell: Radical Theories of Folk and Fairy Tales*, Revised ed. University of Kentucky Press.

Children's and YA literature

- Alcott, L.M., 1983. *Little Women*. Bantam Classics. Original Work Published 1868.
- Anderson, M.T., 2002. *Feed*. Candlewick Press.
- Baum, L.F., Denslow, W.W., 1900. *The Wonderful Wizard of Oz*. G. M. Hill Co.
- Browne, A., 1992. *Zoo*. Julia MacRae.
- Byers, G., Bobo, K.A., 2018. *I Am Enough*. HarperCollins.
- Burnett, F.H., 1911. *The Secret Garden*. Heinemann.
- Coady, C., 1991. *Red Riding Hood*. ABC.
- Collier, B., 2000. *Uptown*. Henry Holt.
- Collins, S., 2008. *The Hunger Games*. Scholastic.
- Collins, S., 2009. *Catching Fire*. Scholastic.
- Collins, S., 2010. *Mockingjay*. Scholastic.
- Collodi, C., 1996. *The Adventures of Pinocchio*. Oxford World's Classics. Original Work Published 1883.
- Defoe, D., 1719. *The Life and Strange Surprising Adventures of Robinson Crusoe*. William Taylor.
- Dickens, C., 1843. *A Christmas Carol*. In *Prose*. Being a Ghost Story of Christmas. Chapman & Hall.
- Dickinson, P., 1988. *Eva*. Gollancz.
- Dowd, S., 2007. *The London Eye Mystery*. David Fickling Books.
- Golding, W., 1954. *Lord of the Flies*. Faber and Faber.
- Grahame, K., 1908. *The Wind in the Willows*. Charles Scribner's Sons.
- Haddon, M., 2003. *The Curious Incident of the Dog in the Night-Time*. Doubleday.
- Hyman, T.S., 1983. *Little Red Riding Hood*, first ed. Holiday House.
- Katcher, B., 2009. *Almost Perfect*. Delacorte Books.
- Lowry, L., 1993. *The Giver*. Houghton Mifflin.
- Milne, A.A., 1926. *Winnie-the-Pooh*. Methuen & Co.
- Montresor, B., 1991. *Little Red Riding Hood*. Doubleday.
- Moon, S., 2002. *Little Red Riding Hood*. Creative Ed.
- Newman, L., Souza, D., 1989. *Heather Has Two Mommies*. Alyson Books.
- Older, D.J., 2015. *Shadowshaper*. Scholastic.
- Orange, T., 2018. *There There*. Knopf.
- Porter, E., 1913. *Pollyanna*. L.C. Page.
- Revis, B., 2011. *Across the Universe*. Penguin.
- Revis, B., 2012. *A Million Suns*. Penguin.
- Revis, B., 2013. *Shades of Earth*. Penguin.
- Rowling, J.K., 1997. *Harry Potter and the Philosopher's Stone*. Bloomsbury.
- Sendak, M., 1963. *Where the Wild Things are*. Harper and Row, New York.
- Seuss, D., 1971. *The Lorax*. Random House.
- Speare, E.G., 1983. *The Sign of the Beaver*. Houghton Mifflin.
- Spyri, J., 2012. *Heidi*. Vintage Classics. Original Work Published 1880.
- Steinbeck, J., 1937. *Of Mice and Men*. Covici Friede.
- Vermette, K., Henderson, S.B., Yaciuk, D., 2017. *A Girl Called Echo*. In: *Pemmican Wars*, vol. 1. Highwater Press.
- Westerfeld, S., 2005. *Uglies*. Simon Pulse.

The evolution of children's literature: metafictional devices and digital mediation

Frank Serafini^a and Sylvia Pantaleo^b, ^a Literacy Education & Children's Literature, Mary Lou Fulton Teachers College, Arizona State University, Phoenix, AZ, United States; and ^b Language & Literacies, Faculty of Education, University of Victoria, Victoria, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	620
Postmodern and digital turns in children's literature	620
Metafictional literature for children	622
Digitally mediated literature for children	623
Multimodality	624
Materiality	625
Interactivity	626
Paratextual features	627
Narrative structures and metalepsis	628
Transmedial narratives and storytelling	629
Concluding remarks	630
References	630
Children's literature references	632

Introduction

The effects of the postmodern and digital *turns* in children's literature have supported the publication and distribution of picturebooks and novels that contain multimodal designs, metafictional devices, digitally mediated interfaces, varying levels of interactivity, unique paratextual materials, diverse narrative structures, metaleptic transgressions, and transmedial components (Aarseth, 1997; Sargeant, 2015; Sipe and Pantaleo, 2008). Throughout this chapter, we use the term *metafictional* to refer to picturebooks that exhibit metafictional features or elements, and the term *digitally mediated narratives* as an umbrella term to refer to a wide range of textual structures, genres, and formats, including novels and picturebooks across digital devices and platforms. Due to the length limitations of this publication, we focus on narrative examples, only briefly referring to informational texts when appropriate (Fig. 1).

Postmodern and digital turns in children's literature

Throughout history, social, technological, and cultural changes have impacted the ecology of the picturebook and other formats of children's literature (Lewis, 2001; Wolf et al., 2011). The broader historical, social, and cultural movement referred to as postmodernism has also been used to explicate changes in children's literature. Although confusion, vagueness and lack of consensus surrounds the meaning and scope of the term *postmodernism* (Coles and Hall, 2001; Hutcheon, 1980; Sim, 2011; Sipe and Pantaleo, 2008), scholars have developed lists of features or concepts associated with postmodernism (Hassan, 1997; Lewis, 2001) and then applied these characteristics to various disciplines and cultural activities. In children's literature scholarship, features of postmodernism have been used by researchers to analyze and discuss picturebooks and young-adult novels, as well as to explore students' responses and interpretations of such literature (e.g., Allan, 2018; Groenke and Youngquist, 2011; Pantaleo, 2008a,b; Serafini, 2005; Sipe and Pantaleo, 2008).

Dresang (1999) proposed her *Radical Change Theory* to explain how the fundamental changes she observed taking place in children's and young-adult literature were reflective of the "connectivity, interactivity and access in the digital world" (p. 14). Indeed, advances in book design software, printing technologies, and digital interfaces have significantly impacted the evolution of formats, designs, and structures of digital and print-based children's literature in general, and picturebooks specifically.

With the advent of computer technologies, it was only a matter of time before digital platforms, the world wide web, software applications, and electronic devices affected the features, formatting, designs, and types of interactivity associated with the publication of picturebooks and other narratives on digital platforms (Serafini et al., 2015; Yokota and Teale, 2014). Over the past two decades, interactive storybooks on CD-ROM have given way to picturebook apps and subscription-based platforms, offering picturebooks and multimodal novels that can be purchased on software developer websites or in an "app store." The expansion of these digitally mediated narratives was accelerated in 2010 with the introduction of the iPad by Apple Computers and the subsequent development of software specifically designed to offer a range of narratives and informational texts to a younger audience. Across these digital platforms and devices, children's literature became readily available as electronic or *e-books* and as self-contained software applications or *picturebook apps*.

The introduction of digital tablets required readers to navigate narrative texts in different ways when compared with the requirements of reading print-based picturebooks in codex formats. Electronic devices, touch screens, and digital software designs continue

- a) “overly obtrusive narrators who directly address readers and comment on their own narrations” (McCallum, 1996, p. 397) (Anstey, 2002; Athanasiou-Krikelis, 2018; Georgakopoulou, 1991; Lodge, 1992; Nikolajeva & Scott, 2001; Silva-Díaz, 2018; Stephens & Watson, 1994; Waugh, 1984)
- b) multiple narrators or character focalisers (McCallum, 1996; Silva-Díaz, 2018; Stephens & Watson, 1994; Viswanath, 2018;)
- c) manifold or multiple narratives (Goldstone, 1998; McCallum, 1996; Silva-Díaz, 2018; Stephens & Watson, 1994; Trites, 1994; Viswanath, 2018; Waugh, 1984) or multistranded narratives (i.e. “two or more interconnected narrative strands differentiated by shifts in temporal and spatial relationships, and/or shifts in narrative point of view” McCallum, 1996, p. 406)
- d) narrative framing devices (e.g., stories within stories, “Chinese-box structures,” and “characters reading about their own fictional lives ... self-consuming worlds or mutually contradictory situations ... a nesting of narrators” Waugh, 1984, p. 22 & 30; Athanasiou-Krikelis, 2018; Viswanath, 2018)
- e) disruptions of traditional time and space relationships in the narrative(s) (Georgakopoulou, 1991; Goldstone, 1998; Lodge, 1992; McCallum, 1996; Silva-Díaz, 2018; Stephens & Watson, 1994; Viswanath, 2018; Waugh, 1984)
- f) metalepsis – “situations where characters and narrators change places, or shift from one plane of being to another” (Stephens & Watson, 1994, p. 44); disruption to diegetic levels of narration (McCallum, 2004; Athanasiou-Krikelis, 2018; Silva-Díaz, 2018; Viswanath, 2018)
- g) nonlinear and nonsequential plots including narrative discontinuities (Goldstone, 1998; McCallum, 1996; Stephens & Watson, 1994; Viswanath, 2018; Waugh, 1984)
- h) intertextuality (Anstey, 2002; Athanasiou-Krikelis, 2018; McCallum, 1996; Nikolajeva & Scott, 2001; Silva-Díaz, 2018; Waugh, 1984)
- i) “parodic appropriations of other texts, genres and discourses” (Stephens & Watson, 1994, p. 44) (Athanasiou-Krikelis, 2018; McCallum, 1996; Silva-Díaz, 2018; Waugh, 1984)
- j) typographic experimentation (Athanasiou-Krikelis, 2018; McCallum, 1996; Stephens & Watson, 1994; Waugh, 1984)
- k) “mixing of genres, discourse styles, modes of narration and speech representation” (McCallum, 1996, p. 397), including “people prose” (Goldstone, 1998; Stephens & Watson, 1994; Silva-Díaz, 2018; Waugh, 1984)
- l) “a pastiche of illustrative styles” (Anstey, 2002, p. 447; Silva-Díaz, 2018) or literary pastiche (Viswanath, 2018; Waugh, 1984)
- m) “new and unusual design and layout, which challenge the reader’s perception of how to read a book” (Anstey, 2002, p. 447)

Fig. 1 List of metafictional devices. Modified version of list of metafictional devices originally published as “Appendix A: List of Metafictional Devices,” in *Exploring Student Response to Contemporary Picturebooks* by Sylvia Pantaleo © University of Toronto Press 2008 (pp. 191–192). Reprinted with permission of the publisher.

- n) excess (i.e. “testing limits – linguistic, literary, social, conceptual, ethical, narrative” Lewis, 1990, p. 144; Silva-Díaz, 2018)
- o) illustrative framing, including *mise-en-abyme* (i.e. “a text – visual or verbal – embedded within another text as its miniature replica” Nikolajeva & Scott, 2001, p. 226; Athanasiou-Krikelis, 2018)
- p) description of the creative process making readers “conscious of the literary and artistic devices used in the story’s creation” (Goldstone, 1998, p. 50) (Georgakopoulou, 1991; McCallum, 1996; Waugh, 1984)
- q) “indeterminacy in written or illustrative text, plot, character or setting” (Anstey, 2002, p. 447) (Athanasiou-Krikelis, 2018; Goldstone, 1998; Lewis, 1990; McCallum, 1996; Silva-Díaz, 2018)
- r) “availability of multiple readings and meanings for a variety of audiences” (Anstey, 2002, p. 447) (McCallum, 1996; Nikolajeva & Scott, 2001; Stephens & Watson, 1994)

Fig. 1 (continued).

to mediate the interactions between readers and texts and offer narratives for children in new and unique configurations (Sargeant, 2015). The digital book market, including picturebook apps, dwells somewhere between the children’s book market and the general entertainment industry, forcing researchers to recontextualize digitally mediated texts across numerous fields of inquiry and academic domains (Turrión, 2014).

Rather than focusing on heuristic labels such as the postmodern or digital turn to document cultural changes and continuities in children’s literature (Hutcheon, 2002), we begin with a brief explication of the metafictional devices and digitally mediated influences on the world of children’s literature and discuss how these influences have affected the forms, designs, features, and narrative structures of contemporary children’s literature. In addition, we explore a range of *meta-structures* that occur across print-based and digitally mediated texts, focusing specifically on picturebooks, and how these meta-structures work across a range of children’s literature.

Metafictional literature for children

The term *metafiction* was coined by Gass (1970, p. 65), referring to narratives that explicitly break their own mimetic illusion. While Gass alluded specifically to fictional narratives, as is discussed below, metafiction is a transmedial phenomenon evident in paintings, television shows, plays, movies, video games, and apps. Although examples of metafictional texts have existed since the beginning of literary publishing, metafictional writing “began to run rampant in the 1960s” (Hutcheon, 1980, p. 1). Hutcheon (1980) defined metafiction as “fiction about fiction—that is, fiction that includes within itself a commentary on its own narrative and/or linguistic identity” (p. 1). She used the term metafiction to “denote a technically definable literary entity” (Hutcheon, 1980, p. 2), and described narrative texts that exhibit “textual self-awareness” (p. 1) as narcissistic. Waugh (1984), whose scholarship on metafiction is referred to extensively in the literature on this topic, described metafictional writing as writing that “self-consciously and systematically draws attention to its status as an artifact in order to pose questions about the relationship between fiction and reality” (p. 2).

Although metafiction has often been associated with postmodernism, as stated previously, postmodernism is a broader intellectual, historical and cultural phenomenon, and metafiction is “an approach, or set of devices, for undermining expectations or for exposing the fictional nature of fictions” (Lewis, 2001, p. 94). However, as stated by Lewis (2001), “one can see why metafictional devices are essential to the postmodernist enterprise, with its sustained attack on all manifestations of authoritative order and unity” (p. 94). Indeed, many of the features or characteristics of postmodernism and metafiction are similar in nature: “narrative fragmentation and discontinuity, disorder and chaos, code mixing and absurdity” as well as openness, playfulness and parody (McCallum, 1996, p. 400).

Essentially, metafictional texts draw “the attention of readers to how texts work and to how meaning is created through the use of a number of specific devices” (Pantaleo, 2008b, p. 11). In addition, metafictional techniques are generally used in combination and the synergy of numerous devices serves to augment or intensify the fictional status and self-conscious nature of a text. Furthermore, the devices are not mutually exclusive; intersections exist among many of the techniques, and several specific devices can be

incorporated by more general metafictional techniques such as “boundary-breaking” (Lewis, 1990) or “accentuates the reader’s presence and makes the reading process more salient” (Athanasidou-Krikelis, 2018, p. 3).

Many scholars have analyzed and critically examined the nature of metafictional devices in selections of children’s literature. For example, Athanasidou-Krikelis (2018), Viswanath (2018), Schwenke-Wyile (2006) and Pantaleo (2004a,b,c, 2010b) have explored the metafictional nature of a variety of picturebooks published in English. Silva-Díaz (2018) has provided examples of picturebooks written in languages other than English that feature metafictional devices. Examination of picturebooks with metafictional devices reveals a gradation in the number of techniques featured in the books, as well as the contribution of the devices to the overall quality of the book and to the reading transaction. In picturebooks, metafictional devices range from being tangential to integral to the story.

Many novels for children and youth also feature metafictional devices. *Save Me a Seat* (Weeks and Varadarajan, 2016), *Wonder* (Palacio, 2012), and *Hello Universe* (Kelly, 2017), are examples of metafictional novels for elementary and middle school students that feature several character focalizers and narrative strands. In the novel *Operation Frog Effect* (Scheerger, 2019), prose, comics, letters, free verse poetry, play script and journal entry format are some of the discourse styles of representation used by the eight character-focalizers to communicate their point of view. In *Project Mulberry* (Park, 2005), chapters interspersed with dialog between the main character, Julia, and her author provide readers with a perspective on Park’s creative writing process.

According to McCallum (1996), metafictional devices can contribute to readers’ esthetic reading experiences offering “a heightened sense of the status of fiction as an elaborate form of play, that is a game with linguistic and narrative codes and conventions” (p. 398). Indeed, while simultaneously providing the pleasure of non-conformity, metafictional devices can contribute to students’ narrative competence (Prince, 2003) by “laying bare the mechanics” of its fiction-making (Hutcheon, 1980, p. 63). By “involving readers in the production of textual meanings, metafictional devices can implicitly teach literary and cultural codes and conventions, as well as specific interpretive strategies, and hence empower readers to read more competently” (McCallum, 2004, p. 588). Many picturebooks with metafictional devices afford readers opportunities for pleasurable reading transactions as they puzzle through the multiple layers of meaning and ambiguity and engage in interpretive strategies in a playful and game-like manner (Pantaleo, 2004a, 2008b). In addition to facilitating readers’ higher-level thinking skills, picturebooks with metafictional devices can contribute to the development of their visual literacy competence, as well as enhance their literacy and literary understandings, including the growth of a repertoire of narrative structures (Pantaleo, 2004b, 2008a, 2010a; Silva-Díaz, 2018; Viswanath, 2018). Finally, developing metafictional awareness can assist readers as they experience these devices in other print and digital texts both within and out of school.

Digitally mediated literature for children

All forms of *digitally mediated literature* require digital software, an electronic platform and device to be experienced. Whether the narrative presented is a digital translation or adaptation of an analog or print-based picturebook or novel, sometimes referred to as a *digital immigrant* (Schwebs, 2014), or a narrative that was originally created electronically, sometimes referred to as a *digital native* or *native e-lit* (Stichnothe, 2014), a tablet, smartphone, or computer screen mediates the narrative, serving as a digital interface between the story and the reader.

In an early definition of digital literature, Holtzman, in his book *Digital Mosaics* (1997), offered the following characteristics of these textual forms: (1) discontinuity—no pre-determined reading paths, (2) interactivity—demands active participation, (3) dynamism—generates a wide range of production possibilities, (4) ethereal worlds—no material aspects of the work (except devices for accessing it), (5) ephemeral worlds—suggesting the work is manifested only as it is executed, and (6) virtual communities. The stability, accessibility, and continuity of conventional narratives are challenged in digital environments requiring different ways for readers to interact and experience various digitally mediated narratives.

Regardless of the labels used to describe digitally mediated literature for children, including e-books, enhanced e-books, or picturebook apps, digitally mediated narratives incorporate various media that require a range of analytical frameworks for investigating the storytelling mechanisms and structures used in these multimodal entities. Scholars interested in the social, technological, and cultural factors that affect the forms, structures, and components of contemporary narratives for children need to consider a wider array of analytical frameworks to better understand how these diverse texts work and how they are experienced by young readers (Kachorsky et al., 2016).

Calls for developing a common metalanguage to define and describe the various components, affordances, and multimodal aspects of digitally mediated literature for children has been put forward since these texts were first introduced and investigated (Al-Yaqout and Nikolajeva, 2015). Although there remains little agreed upon terminology for these components and affordances, it is important to advocate for the continual refinement of a vocabulary for describing and investigating the continual evolution of digitally mediated narratives for children.

The types of digitally mediated literature for children available today range from basic *digitized* adaptations of print-based texts to more complex hyperlinked, transmedial, and interactive *cybertexts* that require readers to make decisions about the navigation and structure of the story. Along this *digital spectrum* (Mackey, 2015), types of digitally mediated literature for children range across a wide variety of formats, interactional affordances, and digital features and structures (Serafini et al., 2015). Demarcating the similarities and differences among these various types of digitally mediated narratives is important for understanding the mediational devices used by book publishers, experienced by young readers, and investigated by children’s literature scholars.

Mackey (2011) highlighted several concepts that have emerged in the development of digitally mediated narratives. These concepts suggest a lack of permanence, a blurring of traditional textual boundaries, and the constantly changing nature of digital

texts, and may be used to better understand the new narrative formats, structures, and spaces that are being developed in digitally mediated literature.

Unlike an e-book, which is basically a digitized version of a print-based picturebook, a picturebook *app* is a type of software application that consists of picturebook content in a digital form that users navigate using a touch-screen interface (Sargeant, 2015). Along with the features associated with basic e-books, such as digitally displayed text and images, navigational buttons, and home screens, picturebook apps offer interactional features that expand the choices, potential reading paths, and experiences of the reader. Many picturebook apps also include voice over narration, background music, video clips, sound effects, animation, and game-like features, thus creating an experience that sometimes resembles playing a digital game rather than reading a narrative. Most of the features are optional and can be engaged or turned off.

The introduction of digital tablets, including Apple's iPad and Amazon's Kindle Fire enabled picturebooks to be made available in full-color and opened up the market for more publishers and designers to offer their print-based texts across electronic formats and platforms. Recent innovations in software design have made most e-books available across a variety of devices and offered on various *app stores* depending on the device purchased, predominantly Android and iOS software platforms.

Several publishing or design companies, including *Nosy Crow*, *Loud Crow*, *Oceanhousemedia*, and *Winged Chariot Press* offered some of the first picturebook apps that went beyond simple scanned or e-book formats, offering enhanced features including voice-overs, interactive features, musical backgrounds, and hyperlinks. In addition, *Nosy Crow* and *Oceanhousemedia* have created proprietary software engines for further development of digitally mediated picturebooks.

In addition to the previously mentioned affordances, there are limitations associated with digitally mediated literature, including the challenge of sharing digital texts with other readers due to software restrictions, a shortened shelf-life caused by frequent updates associated with improvements in particular operating systems, and the initial expense of purchasing, charging, and maintaining an electronic device. Some of the biggest challenges with digitally mediated narratives involve the archiving and retrieval of electronic texts, and the continuing evolution of the software itself. From a publication and archival perspective, when print-based books go out of print, old copies are still available to read or purchase.

When a digital book or picturebook app is updated, or the operating system on the supporting device is updated, no older version remains available to archive or analyze. In fact, many of the original picturebook apps analyzed and mentioned in this chapter have become unavailable due to system upgrades, making one's financial outlays for them irretrievable. Sharing digital e-books and picturebook apps is not as easy as handing a print-based picturebook to a friend or storing one on the library shelf, as many of these apps remain proprietary to the device on which they were purchased.

In the following sections of this chapter, we address a selection of meta-structures that include a variety of metafictional devices rendered in digitally mediated literature for children. Each meta-structure is instantiated in different ways across print-based and digital media and affects the readers' experiences and the structures and features inherent in the narratives themselves.

Multimodality

Picturebooks, textbooks, websites, screen-based apps, digital texts, museum exhibits, television programs, and billboards are just a few examples of the wide variety of multimodal texts young readers experience in contemporary society. Multimodal texts are composed of a variety of *semiotic resources* that are used to represent and communicate meaning across particular social contexts. According to theories of social semiotics, the semiotic resources of a community are defined as socioculturally and historically shaped materials or modes for making meaning (Bezemer and Kress, 2016). All available modes, such as written and oral language, images, gestures, sculpture, architecture, and music include various semiotic resources, organizing principles, limitations and affordances, which affect the possibilities for communication and representation in today's world.

Kress (2015) uses the term multimodal ensemble to refer to "a designed complex of different modes" (p. 57) and emphasizes how "the meaning of the whole arises out of the contribution of each part in its interaction with all other parts" (p. 57). Indeed, although in multimodal ensembles, analysis can focus specifically on the meaning-making affordances of each mode, it is essential to understand and consider how intersemiotic combinations and relationships among modes constitute an integrated whole (Jewitt, 2014; O'Halloran, 2004).

In general, multimodality and social semiotics are concerned with how people generate meanings in transaction with multimodal texts, including written language, visual images, and design features, from a variety of perspectives to meet the requirements of particular social contexts (Serafini, 2014).

The multimodal nature of print-based children's literature extends beyond picturebooks. For example, an examination of a collection of multimodal novels by Reid and Serafini (2018) revealed the variation, and complexity of use and relationship among multiple modes to communicate and represent meaning in these texts. In their qualitative content analysis of the multimodal covers of 21 young-adult titles in both graphic novel and novel format, Kachorsky and Reid (2020) noted how the distinct differences in the images and design elements on the covers communicated both overt and symbolic messages about the intended audience and the role of literary elements of each format of text. Furthermore, they described how publishers established a relationship between the young-adult novel and the graphic novel through the use of "design features and branding devices" on the covers (Kachorsky and Reid, 2020, p. 314).

Whereas, print-based picturebooks draw on visual images, design elements, and written language as modes of meaning making, children's picturebook apps expand the types of modalities presented by including animation, sound effects, music, and embodied

user interactions to orchestrate complex ensembles of meaning in addition to text and image (Serafini et al., 2016). The transition of picturebooks from printed codex into a digitally mediated format is a reflection of the shaping of semiotic resources to accommodate emerging ways of constructing meaning in the digital age.

Digitally mediated texts are created using computer processing codes or binary codes. These codes are the same regardless of whether they are used to represent language, images, videos, gaming modules, or navigational hotspots on an electronic device. It is this common ground in basic digital coding that has allowed publishers and picturebook app designers to easily incorporate multimodal elements into electronic texts and applications. Although images and text and other multimodal elements appear in different forms on the tablet, computer or smartphone, the common building blocks at the level of digitization readily affords designers of picturebook apps the ability to blend a range of modalities in their apps and e-books.

Materiality

The materiality of print-based picturebooks refers to the physical and sensory qualities of materials and design structures used in creating traditional picturebooks. In the opinion of Alaca (2018), the materiality of picturebooks is “a third narrative system, alongside text and images ... that deepens children’s understanding of a story, initiate[s] play, and make[s] connection to the real world” (pp. 66–67). Indeed, the materiality of print-based picturebooks contributes significantly to the reading experience—to readers’ expectations, appreciation, interpretations and evaluation of and responses to books in both print and digital forms.

Nikolajeva (2008) emphasized how substantial changes to the materiality of picturebooks, such as inclusion in “anthologies where covers are absent, [and] several double spreads placed on the same page” (p. 59), can ruin their “esthetic totality” (p. 59). Furthermore, some aspects of the materiality of hardback editions of picturebooks (e.g., stamping) are impossible to reproduce in paperback editions. In addition, some publishers choose not to replicate all of the peritextual features (e.g., endpapers) in the paperback edition. As physical and esthetic entities, these design and publication decisions affect the meaning-making potential of picturebooks.

The affordances and constraints for making meaning are affected by the material boundaries of this format of literature. In fact, playing with the material form of books, including the size and shape, has a long history in literature. However, advances in technology have simultaneously shaped innovations in printing, which have expanded options for the types of materials used in the production and the overall design of picturebooks.

In some of their metafictional picturebooks, creators Emily Gravett and Lauren Childs have incorporated “the materiality of the book into the narrative” (Do Rozario, 2012, p. 151). In addition to drawing attention to the materiality of books themselves, both Gravett and Childs have explored “the kinds of relationships that form between readers and books through the act of reading” in some of their creations (Do Rozario, 2012, p. 153). In the metafictional picturebooks *Wolves* (2005) by Gravett and *Who’s Afraid of the Big Bad Book?* (2002) by Childs, the creators concurrently “make visible the power and constraints of the bound page” and “interrupt and interrogate the static qualities” of print-based texts (Mackey, 2008, p. 115). Gravett’s overall book design and playful use of material aspects in *Wolves*, specifically the peritextual features, draw readers’ attention to the status of the picturebook as an artifact (Hutcheon, 1980) and position readers inside a book system (Mackey, 2008).

Interestingly, Mackey (2008) suggested children’s awareness of the physicality of the book has been heightened because of their experiences with digital texts and formats (p. 103). The materiality of picturebooks, the physicality of these esthetic objects, cannot be replicated with digital books, interfaces, and devices. Furthermore, comparison of the nature and thus the affordances of materiality of print-based picturebooks to the (im)materiality of digital picturebooks or apps by Al-Yaqout and Nikolajeva (2015) revealed fundamental differences that affect the design of the book, “the esthetic experience of the narrative” (p. 3), and the perceptual and sensory reading experience.

The different ways picturebook narratives are mediated across analog and digital platforms and devices should not be conceptualized as a neutral or innocent form of mediation that simply represents the narrative without affecting a reader’s experience. The differences between print-based texts and digitally mediated texts extend beyond material considerations and affect the meanings constructed by readers and the experiences they have with these texts. Whether the tactile, navigational, and experiential demands associated with digitally mediated literature deters or supports children’s engagements compared with print-based counterparts has been investigated using cognitively-based experimental studies designed to identify variables that affect readers’ experiences and levels of comprehension (Kucirkova, 2017).

In contrast, Hateley (2013) suggests the size and shape of digital tablets (iPad and Kindle Fire) does not necessarily differ all that much from the physical experience of reading a print-based or codex picturebook. Drawing on *skeumorphs* (Sargeant, 2015), virtual features that mimic print-based experiences, digitally mediated picturebooks offer visual aids for navigating and progressing through digital texts, including the sounds of pages turning, visual renderings of pages turning, and the opening and closing of virtual book covers. These particular visual features of digitally mediated narratives help readers connect to their previous experiences with print-based texts.

Some features of children’s literature are present in both print-based and digitally mediated instantiations, including pages, page turns, chapters, covers, titles, characters, and settings, while other features are present exclusively in one or the other forms, including book spines, size and dimensional characteristics, and different types of paper used in printed texts, and video clips, sound effects, and gaming modules for digitally mediated texts. Essentially, picturebooks and picturebook apps have unique ways of *being* texts and have created their own specific materiality and forms of mediating interfaces (Henkel, 2018).

Movements of the reader's fingers across the touch-sensitive screen creates an illusion of materiality that differs from the material object of a print-based picturebook. In order to activate components of an interactive or enhanced picturebook, readers are required to swipe, tap, scroll, and press visual triggers or "hotspots" to activate available features and hyperlinks. Whereas, in print-based picturebooks the story progression is controlled by the author, illustrator and designer, with digitally mediated literature the reader has options and can manipulate the story sequence or progression by making gestural moves that changes the structure of the story.

Kurcikova (2019) asserted that digitally mediated literature has both internal and external material properties. External properties include types of haptic engagement, where readers tap, swipe, pinch, stretch, and drag virtual elements across a scratch-resistant glass interface (Wohlwend, 2017). Internal material properties refer to digital story elements, including multimedia and graphic design elements, as well as the digital coding that drives the software application. Additionally, Henkel (2018) described how the materiality of picturebook apps should be analyzed across three dimensions, including the intrinsic, embodied material qualities, the physicality of the app as an embedded device and interface, and the mode of existence or cultural scope of materiality of the app (p. 340).

Augmented reality apps combine the presence of print-based books with the features of digital devices. Drawing on the affordances of a camera lens located on a digital device, augmented reality software is used to locate *visual triggers* that sets in motion additional video and audio components that can be viewed on the digital device. These digitally rendered visual and textual features can be overlaid on the environment in which the reader is located, allowing the reader to blend the virtual world of the text with the world in which they live.

Moonbot Studios, the work of acclaimed children's author and illustrator William Joyce, developed an augmented reality software application, called *Imag-N-O-Tron*, that was designed to work with his print-based picturebooks, including *The Numberlys* (Joyce, 2014) and *The Fantastic Flying Books of Mr. Morris Lessmore* (Joyce, 2012). By holding the digital device above the print-based picturebook, the camera detected unique visual markers in the print book and set in motion additional narrative features that could be viewed and manipulated on the digital device. Unfortunately, at the time of this writing many of the augmented reality and picturebook apps are no longer available due to updated operating systems. In a personal communication, Joyce stated the expense of constantly updating these picturebook apps was prohibitive for his company due to sales of the app being less than originally predicted.

Interactivity

Kucirkova (2017) generated an integrative framework for theorizing and researching interactivity in children's digital books. Based on a content analysis of one hundred children's digitally mediated interactive narrative picturebooks, Kucirkova (2017) identified five categories of interactivity and listed "the design elements that can be incorporated" (p. 1176) into digital books for each category. Only the interactivity category of synesthesia, "a sense experience or sense impression, derived from reading a book" (Kucirkova, 2017, p. 1176), is pertinent to a discussion of print-based books. Visual presentation, touch manipulation, and audio presentation are three key ingredients or elements of synesthesia.

Reading print-based literature involves holding a book, moving the eyes and turning the pages. In pop-up books and some board and trade books, readers can pull flaps, touch some tactile material or manipulate movable parts such as the newspaper in *Fairytale News* (Hawkins and Hawkins, 2004) or the postcards in *Meerkat Mail* (Gravett, 2006). In Pilkey's (2016) graphic novels and his *Captain Underpants* (2010) books, he includes examples of "flip-o-rama", which is a method of animation where pictures appear to move when readers flip the right-hand page back and forth. Other design features in books requiring physical interaction by readers are gatefolds. Some picturebooks have a single gatefold (e.g., *The Getaway*, Vere, 2006) or a double gatefold (e.g., *Window*, Arbona, 2020; *Who's Afraid of the Big Bad Book?*, Child, 2002) that readers must open to follow the progression of the narrative. Finally, the nature of reading aloud of print-based literature is a key factor in the sensory experience of the reading event. Typographical features of the written text provide readers with information about how to interpret the visual voice of typefaces. In addition, when reading silently one's inner voice can be considered as an aspect of audio presentation.

Many of the metafictional devices included in contemporary picturebooks, "demand a greater degree of attentiveness and interaction ... [as readers] must decide whether to 'point and click' here or there with their eyes and their minds" (Dresang, 1999, p. 114). One example of an interactive format is the inclusion of manifold or multiple narratives. A second narrative may be told entirely through illustrations (e.g., *Something From Nothing*, Gilman, 1992); or parallel stories may be told with images and written language (e.g., *Artie and Julie*, Chih-Yuan, 2008). Readers must assume more active interpretive roles as they make connections between the multiple narratives.

Other metafictional devices that create an interactive format include authors commenting on the story or directly addressing characters (e.g., *Snappy the Alligator (Did Not Ask to Be in this Book!)*, Falatko and Miller, 2016); or obtrusive characters directly addressing readers and/or the author (e.g., *An Undone Fairy Tale*, Lendler and Martin, 2005); or a character, who is also the narrator, addressing readers (*This Book Will Not be Fun.*, Dunlap and Tallec, 2017). Reader interactivity is also heightened by the metafictional devices of nonlinear and nonsequential plots including narrative discontinuities (e.g., *Interrupting Chicken*, Stein, 2010); disruptions to the time and space of a diegetic (e.g., *Ivan the Terrier*, Catalanotto, 2007); texts requiring the reading of more than one set of words (e.g., speech bubbles and narrative prose) (e.g., *Knuckle Heads*, Holub and Slack, 2008); and indeterminacy (e.g., *The Three Pigs*, Wiesner, 2001).

With respect to digitally mediated literature, the term interactivity has been used to suggest digital formats and interfaces offer more interactive features or potential within specific tableaux (similar to the page spread in print-based picturebooks), across transitional features including embedded hotspots and skeumorphs, or within enhanced features like hyperlinks, gaming components, on-demand dictionaries and embedded video clips. From these perspectives, interactivity has been defined as the potential for readers to participate in modifying the form or content of a mediated environment (Serafini et al., 2015), as the characteristics of a particular medium or platform (Schwebs, 2014), as a non-trivial effort on behalf of the reader (Aarseth, 1997), and as the difference among conventional, active, and interactive participation (Turrión, 2014).

Although interactivity remains an elusive concept (Bucy, 2004), various typologies of interactivity have been proposed to elucidate the differences among actual and perceived interactivity (Voorveld et al., 2011), potential, enacted, and required interactivity (Serafini et al., 2015), and passive, active, control-based, and joining forms of interaction (Adami, 2014). In addition, Unsworth (2006) has suggested a continuum of digitally mediated literature that ranges from basic e-books and linear forms of digital texts, through interactive and hyperlinked texts, to more interactive transmedial narratives. Turrión (2014) has suggested the digital, multimodal, non-linear, and transmedial properties of digitally mediated literature can be conceptualized as a rupture of discursive or narrative linearity that is more commonplace among digital texts than print-based counterparts. A basic distinction to be made here is the difference between apps that require readers to do or add something to the story for the story to progress and those where the reader is simply tapping a hotspot on the touchscreen to allow the narrative to unfold in a predetermined sequence. Those apps that require basic reader interactions are closely related to the “choose your own adventure stories” where readers select from a finite number of options as the story unfolds.

In one of the most elaborate analyses and conceptualizations of interactivity, Ryan (2004) constructed a typology of interactive features that fall along two primary axes: (1) external vs. internal interactions, and (2) exploratory versus ontological interactions. External interactions feature a traditional reader experiencing a narrative from a distance whereas internal interactions involve readers interacting as characters, or avatars, “inside” a particular story. Exploratory interactions are involved with looking at, but not changing the nature of the narrative, whereas ontological interactions require decisions or actions that cause changes to the narrative itself (Ryan, 2004).

The multi-touch screen that is featured across digital tablets is a medium-specific affordance that allows or drives readers' engagement with the features of a digital text. According to Rafaeli (1988), the feeling of meaningful dialog between reader and device offered by a touch screen offers a responsiveness and level of control that provides agency for the reader through their interactions with digitally mediated literature. Whereas Lindhe (2010) refers to these interactions or modes of engagement as an *aesthetic of tactility* that embodies experience with digital devices, Soyland and Gulliksen (2019) describe the touch screen of a digital tablet as a form of *virtual materiality* where the virtual elements of a story can be manipulated but not actually touched.

Kurcikova (2019) has advocated for moving past previous conceptions that see interactivity in terms of binary distinctions between reader and text and suggests using “third space” theories to reconceptualize interactivity as a more inclusive, comprehensive aspect of the transaction among readers, texts, and interactions. In either case, the question of whether interactivity is an aspect or feature of the text itself or of the transactions between readers and texts remains a tension in the scholarship on interactivity.

Paratextual features

The concept of paratext was originated by Gérard Genette (1997). He described the paratext as a threshold, an “undefined zone between the inside and the outside, a zone without any hard and fast boundary on either the inward side (turned toward the text) or the outward side (turned toward the world's discourse about the text)” (Genette, 1997, p. 2). Genette (1997) noted how the time “period, culture, genre, author, work, and edition” (p. 3) affected the nature of paratexts.

According to Genette, the overarching category of the paratext is constituted by *epitexts* and *peritexts*. Epitexts, which can be either *private* or *public*, are located outside a book: “any paratextual element not materially appended to the text within the same volume but circulating, as it were, freely, in a virtually limitless physical and social space” (Genette, 1997, p. 344). Genette described two categories of private epitexts: the confidential epitext (e.g., private correspondence with another) and the intimate epitext (e.g., diary or journal). Examples of public epitexts include public critiques or statements (written or oral) about the work by the publisher or creator; lectures, interviews or podcasts by the creator; digital texts associated with a book such as previews or a book trailer; print-based or digital reviews of the work by the creator; and public responses to critics by the creator. Other public epitexts include creators' websites and blogs, publishers' websites and social networking sites.

The entire spatial and material zone of the peritext “is the direct and principal (but not exclusive) responsibility of the publisher” (Genette, 1997, p. 16). Indeed, many authors and illustrators of picturebooks design the publisher's peritext to function as narrative text (Do Rozario, 2012). Other terms used to refer to the elements of the publisher's peritext include *front matter* (Doonan, 1993; Sipe, 2008) and *publishing information* (Sipe, 1998). Peritextual elements in picturebooks include the dust jacket, case (front and back hard covers), endpapers, frontispiece, full title page, half-title page, dedication page, address of publisher, copyright date, price, typesetting style, choice of paper, table of contents, prefaces, intertitles, and footnotes. The inclusion, location and intention of specific peritextual elements varies across picturebooks.

Emily Gravett's intentional design of the peritextual features of *Wolves* (2005) communicates the metafictional nature of the picturebook. Indeed, the synergy of several peritextual elements, including the double binding effect, create additional “versions” of the picturebook. However, the “various versions are never fully separate from each other, coexisting as the one material narrative”

(Do Rozario, 2012, p. 156). Readers must attend closely to visual and textual details on the dust jacket, case, endpages and front matter as these peritextual features are not only significant parts of the narrative, but also they emphasize the physicality of the book as a book and as work of fiction with multiple narratives.

In *Paratexts: Thresholds of Interpretation*, Genette (1997) noted how his nascent theorizing would provide a “provisional service” to others (p. 15). Indeed, in addition to criticizing Genette’s restricted meaning of the term “text” (print books) and his focus on literary works, scholars have applied the concept of paratext to analyses and discussions of audiovisual forms and digitized media (Birke and Christ, 2013; Kümmerling-Meibauer, 2013; Leavenworth, 2015; McCracken, 2013; Stanitzek, 2005). The discussion of the paratextual landscape by Brookey and Gray (2017) reveals the diverse range of topics that academics have explored using the concept of paratext.

Traditional paratextual features are instantiated in unique ways when considering digitally mediated literature for children. McCracken (2013) asserted paratextual features can no longer be studied as singular fixed objects, like those associated with print-based picturebooks, rather elements associated with digitally mediated texts exist temporally and spatially within a particular reading experience. Drawing upon empirical studies of particular digitally mediated picturebooks in lieu of Genette’s (1980, 1988) original theoretical formulations, McCracken (2013) asserted inward and outward pathways from the central text as better ways of conceiving paratextual features.

McCracken (2013) suggested researchers should reconceptualize paratextual features as *centrifugal and centripetal vectors*, rather than focusing on traditional conceptions of *epitext* and *peritext*. Centrifugal vectors potentially draw readers outside the text proper, including publisher-created materials, blogs, reader commentaries, advertisements, affinity spaces associated with the story, and other intertextual links, while centripetal vectors modify the reader’s experience on inward vectors or the internal elements of the text. The centripetal vectors may include readers’ ability to modify the displayed language of the text, selecting how the book shall be read, and choosing to add voiced narration or background music provided in the picturebook app. These paratextual features are more often found in digitally born apps rather than stories adapted from original print-based versions.

The first peritextual element encountered when reading a digitally mediated book or app is the icon used to identify the text on a tablet interface, often accompanied by an abbreviated title. Various icons feature logos from the text or a publisher’s logo or trademark. These standardized icons are very different from how a reader first encounters a print-based text, as regarding size, shape, and material form. The opening screen of picturebook apps and storybooks often features instructions on how to operate or navigate the app or how to proceed through the text, including “read by myself, read to me, or read automatically” options. In order to see the front cover on a digital book, the reader may have to “swipe back” since the default start page is often not the cover.

The peritextual literacy framework (PLF) created by Witte et al. (2019) can be used to identify, comprehend and evaluate peritextual elements in print-based texts and multimodal digital texts across a range of media types and fields. The heuristic focuses on Genette’s (1997) interest in the functional characteristics of paratextual elements: “the functions that the message aims to fulfill (*to do what?*)” (p. 4). The PLF features six types of functions: production, promotional, navigational, intratextual, supplemental and documentary (Gross and Latham, 2017). Although the PLF provides educators with a scaffold “to promote critical thinking in the use, analysis, and production of works” (Gross, 2019, p. 11), fundamentally the PFL draws attention to the importance and multifaceted nature of peritextual information.

Narrative structures and metalepsis

Through multiple experiences with oral and written stories, individuals construct schemata or cognitive representations of story structures, elements and genres. In Genette’s (1980) “systematic theory of narrative” (p. 7), he used the word *story* to signify narrative content; the term *narrative* to denote the “narrative text itself” (p. 27), the oral or written discourse; and the word *narrating* to describe “the producing narrative action and, by extension, the whole of the real or fictional situation in which the action takes place” (p. 27). These terms serve to differentiate between story and plot and convey how various discourses can be used in the telling of a narrative. The terms *fabula* and *sjuzet*, originating in Russian Formalism (Buchanan, 2018; Herman, 2005; Prince, 2003), can also be used to consider how texts are structured and how they achieve their effects (Pantaleo, 2004a). *Fabula* refers to the story, the sequence of events and situations that constitute the story—the chronological order of narrated actions and circumstances. In contrast, *sjuzet* refers to the order and composition of how the *fabula* or story is constructed and narrated (Buchanan, 2018)—the plot. Genette (1980) also wrote about relations between story time and narrative time (p. 33).

Other terms important to understanding the narrative structure of texts include *diegesis* and *metalepsis*. *Diegesis* is “the narrative world, whether fictional or nonfictional” (Phelan and Rabinowitz, 2005, p. 544), in which “the situations and events narrated occur” (Prince, 2003, p. 20). Genette (1980) wrote about the existence of various diegetic levels in narratives and coined the term *metalepsis* to describe a transgression between the “sacred frontier between two worlds [or narrative levels], the world in which one tells, [and] the world of which one tells” (p. 236). As a narrative structural device, *metalepsis* serves to create ambiguity and to increase the complexity of narrative representations by breaching conventional “relationships and hierarchies between characters, texts, authors, illustrators and readers” (McCallum, 2008, p. 181).

Genette (1988 as cited in Pier, 2011) and other individuals have refined and extended his original ideas about the concept of *metalepsis* (see Hanebeck, 2017 and/or Pier, 2016 for a comprehensive discussion of *metalepsis*). Indeed, scholars have written about the occurrence of *metalepsis* across other literary genres, video games, virtual reality technology (Ryan, 2004), animation (Feyersinger, 2010), film, television, theater, digital fiction (Bell, 2014), comics and graphic novels, and art (Kukkonen and Klimek,

2011). As well, taking up a transmedial perspective on the concept reveals how metalepsis is not exclusive to high culture, or innovative written works (Pier, 2011), or postmodern texts.

Scholars have generated taxonomies to distinguish different types of metalepsis (Fludernik, 2003; Hanebeck, 2017; Pier, 2016; Ryan, 2004; Wolf, 2013) (e.g., rhetorical, ontological, and epistemological) (see Pantaleo, 2019 for discussion); created descriptive frameworks to portray the nature of diegetic transgressions (Bell and Alber, 2012; Hanebeck, 2017; Kukkonen, 2011; Pier, 2016) (e.g., ascending, descending and horizontal); and discussed multiple possible effects of metalepses (Bell and Alber, 2012; Feyersinger, 2010; Genette, 1988; Kukkonen, 2011). However, Kukkonen (2011) emphasized the need to consider the relationships between the effects of metalepsis and types of media due to the wide-ranging "spectrum of effects and functions" of metalepsis in popular culture (p. 12). In addition, the effects or potential interpretations of metalepses must be contextualized more broadly to include consideration of reader/viewer and text elements and dynamics. As noted by Hanebeck (2017), "metaleptic texts, performances or artifacts (typically) have a metaleptic *potential* that needs to be *realized* by a reader or viewer" (p. 27).

Although picturebooks can be metafictional in nature without featuring metaleptic transgressions, the breaching of ontological narrative boundaries or levels in texts disrupts their fictional illusion, drawing attention to their artifact status (Waugh, 1984). Metaleptic transgressions of the boundaries between the telling and the told in picturebooks can occur independently or concurrently through written language, images or layout. Some researchers have engaged in content analysis of particular picturebooks and described the nature of the breaching of barriers between diegetic levels (e.g., Pantaleo and Sipe, 2012). Analysis of the kinds of metaleptic transgressions in a collection of contemporary picturebooks (Serafini and Reid, 2019) revealed three main categories, each composed of several subcategories: "(1) rhetorical transgressions only, (2) ontological transgressions only, and (3) a blending of rhetorical and ontological transgressions" (p. 267). The findings from such analyses reveal the diversity and complexity of devices that can be used by picturebook creators to contravene narrative levels.

Digitally mediated children's literature may further expand the possibilities of narrative structures by increasing narrative complexity, fragmentation, non-linearity, and the changing nature of reader interactions with digital media. Hutcheon (2013) proposed three modes of engagement with digitally mediated narratives, namely *telling*, *showing*, and *interacting*, with the last mode of *interactive* being the most salient for this discussion. Digitally mediated children's literature provides a platform for changing the basic nature of how readers interact with narrative structures.

The ontological and rhetorical boundaries found in traditional, print-based literature are dissimilar to those found in digital spaces that easily construct intertextual connections across storyworlds, software that allows readers to record themselves and add to the story through the inclusion of video or audio clips, characters that talk directly to the reader when prompted or touched, and cyber textual features that require readers to play a non-trivial role in the story's development (Aarseth, 1997). These affordances of digitally mediated literature for children provide opportunities for both ontological transgressions of the narrative structure, when the boundaries between reader and text are transgressed by characters addressing the reader directly, narrators breaking the fourth wall, and readers reproducing themselves as avatars and immersing themselves in the narrative itself.

Turion (2014) distinguished between *real* and *simulative* types of participation with digitally mediated narratives. Real participation requires what Aarseth (1997) called *non-trivial interaction*, in other words the reader has to interact with elements and features of the electronic narrative in ways beyond simply turning the page for the story to unfold. In *The Monster at the End of the Book* (Stone, 2003), an early revised version of a well-known print-based picturebook, Grover, the main character in the story directly addresses the reader and instructs them not to turn the page. In this ontological transgression, the reader is made conscious of the materiality and fictionality of the narrative world in the digital narrative. The app version of the story maintains the metafictional play, making the users' interactions more realistic, blurring the boundaries between the narrative and real worlds.

In the augmented reality version of *The Flying Books of Mr. Morris Lessmore* (Joyce, 2012), the reader uses an electronic tablet to bring selected pages of the narrative to life, crossing the boundaries of fictional storyworlds and the world of the reader. These types of ontological transgressions, specifically ascending metalepses where characters leave the storyworld and "enter" the reality of the reader, are directly supported by the affordances of digitally mediated literature in ways print-based texts are not (Serafini and Reid, 2019). Although ontological and rhetorical transgression may occur in both print and digital formats, the affordances provided by digital interfaces and software make these metaleptic boundary crossings easier to produce and design.

In a general sense, all forms of hypertext represent a form of rhetorical transgression as the selected links take the reader from one text to another (Ryan, 2004). These hyperlinked actions may or may not change the diegetic levels of the story and may or may not transgress ontological boundaries as readers connect to different narratives. The path a reader follows when interacting with digitally mediated literature may wander from text to text and across ontological boundaries in similar fashion to animated cartoons, video games, and other cyber texts have done for years.

Transmedial narratives and storytelling

One of the unique features of digitally mediated literature is the potential for transmedial storytelling or narratives that work across a range of texts and media to offer narratives stories across a variety of devices, platforms and modalities. Some forms of digitally mediated literature working across a range of digital platforms to extend the narrative experience are referred to as *transmedial narratives* (Jenkins, 2006). Transmedial narratives are based on the systematic dispersal of integral elements of a storyworld across multiple delivery channels which make a unique, original contribution to a coordinated entertainment experience (Jenkins,

2006). Drawing on a media conscious narratology (Ryan, 2004), transmedial narratives use a range of media, including television, printed text, internet websites, CD-ROM, and audio-visual resources to render a narrative experience.

Transmedial narratives have been researched from a range of theories and academic perspectives including media studies (Ellestrom, 2019), narratology (Thon, 2016), participatory culture (Jenkins, 2013), and affinity groups or spaces (Gee, 2007). Transmedial narratives can be theorized as a digitally mediated form of narrative, as a form of new media (Ryan, 2004), a type of narrative structure and design (Mackey, 2011), a form of adaptation or translation of fictional narratives (Kerchy, 2018), and as a storytelling mechanism (Zheng, 2016). In addition, the evolution of transmedial narratives suggest the need to consider the possible social, cultural, economic, and pedagogical implications for these new texts, platforms, structures, and designs (Ryan, 2004).

Jenkins (2006) asserted transmedial narratives are not always based on individual characters or specific plots but rather complex fictional worlds which can sustain multiple interrelated characters and stories. Some key aspects of transmedial storytelling are the collaborative nature of the work of storytelling, including authors, publishers, and web designers and the relationship between the storytellers and the readers as they navigate the online space that furthers character dynamics, plot, and setting. In many ways, the evolution of digital forms of storytelling and transmedial narratives represent an expansion of the ways we define texts in the digital age and the blurring of boundaries regarding what constitutes a fictional narrative.

One of the first series of transmedial narratives developed for children was the Skeleton Creek stories developed by Cormack (2009, 2016). In this series, readers must purchase a print-based novel and navigate to selected pages in each chapter to get a website address and password to then go online to view video clips offering additional narrative elements from second point of view of the character in the video, then navigate back to the book to continue. Neither the print-based book, nor the video narrated elements offer the complete narrative. A reader must access both narrative representations to experience the story.

It has been suggested that one of the key reasons transmedial narratives have been recently produced in larger numbers, aside from technological advances and possibilities in storytelling, are the economic incentives for large corporations to capitalize on these intersections of well-known characters and storyworlds (Jenkins, 2013; Aarseth, 1997; Ellestrom, 2019; Jenkins, 2013; Kerchy, 2018; Zheng, 2016). By creating different points of entry for consumers, the entertainment industry can take advantage of comic book readers buying movie tickets, and toy manufacturers offering prominent characters from narratives as action figures and so on.

Concluding remarks

The analysis of both metafictional picturebooks and digitally mediated literature requires researchers to consider their manifestation as textual and material objects, their embedment in a particular medium or interface, and the ways various platforms, codices, or digital devices mediate the narrative experience (Henkel, 2018). Working across a range of academic domains, researchers have taken up research on metafictional and digitally mediated literature for children that involves theories associated with narratology, multimodality, gaming, interactivity, and social semiotics. The scholarly literature on metafictional and digitally mediated literature for children continues to expand as researchers draw upon wider ranges of theories and research methods.

Some areas in need for continued investigation include the following: children's responses and opinions about how material affordances of print picturebooks and digitally mediated literature contribute to or affect their reading experiences; the tactile, navigational, and experiential demands associated with digitally mediated literature and consideration of how these demands deter or support children's engagements compared with print-based counterparts; the role of peritextual elements across print and digitally mediated literature; the range of diverse structures included in these complex narratives; and the pedagogical possibilities these texts bring to various educational settings.

References

- Aarseth, E.J., 1997. *Cybertext: Perspectives on Ergodic Literature*. The Johns Hopkins University Press, Baltimore, MD.
- Adami, E., 2014. What's in a click?: a social semiotic framework for the multimodal analysis of website interactivity. *Vis. Commun.* 14 (2), 133–153.
- Alaca, I.V., 2018. Materiality in picturebooks. In: Kümmerling-Meibauer, B. (Ed.), *The Routledge Companion to Picturebooks*. Routledge, London, UK, pp. 59–68.
- Allan, C., 2018. Postmodern picturebooks. In: Kümmerling-Meibauer, B. (Ed.), *The Routledge Companion to Picturebooks*. Routledge, London, UK, pp. 201–208.
- Al-Yaqout, G., Nikolajeva, M., 2015. Re-conceptualising picturebook theory in the digital age. *Nord. J. ChildLit Aesthet.* 6 (1), 1–7.
- Athanasiou-Krikelis, L., 2018. Defining metafiction: authorship and readership in Emily Gravett's picturebooks. *Lion Unic.* 42 (1), 1–19.
- Bell, A., Alber, J., 2012. Ontological metalepsis and unnatural narratology. *J. Narrat. Theory* 42 (2), 166–192.
- Bell, A., 2014. Media-specific metalepsis in 10:01. In: Bell, A., Ensslin, A., Rustad, H.K. (Eds.), *Analyzing Digital Fiction*. Routledge, London, UK, pp. 21–38.
- Bezemer, J., Kress, G., 2016. *Multimodality, Learning and Communication: A Social Semiotic Frame*. Routledge, London, UK.
- Birke, D., Christ, B., 2013. Paratext and digitized narrative: mapping the field. *Narrative* 21 (1), 65–87.
- Brooke, R., Gray, J., 2017. "Not merely para": continuing steps in paratextual research. *Crit. Stud. Mass Commun.* 34 (2), 101–110.
- Buchanan, I., 2018. *A Dictionary of Critical Theory*, second ed. Oxford University Press. <https://doi.org/10.1093/acref/9780198794790.001.0001>.
- Bucy, E.P., 2004. Interactivity in society: locating an elusive concept. *Inf. Soc.* 20 (5), 373–383.
- Coles, M., Hall, C., 2001. Breaking the line: new literacies, postmodernism and the teaching of printed texts. *Reading* 35 (3), 111–114.
- Do Rozario, R.-A.C., 2012. Consuming books: synergies of materiality and narrative in picturebooks. *Child Lit.* 40, 151–166.
- Doonan, J., 1993. *Looking at Pictures in Picture Books*. The Thimble Press, Stroud, Glos.
- Dresang, E., 1999. *Radical Change: Books for Youth in a Digital Age*. The H. W. Wilson Company, New York.
- Ellestrom, L., 2019. Transmediation: some theoretical considerations. In: Salmose, N., Ellestrom, L. (Eds.), *Transmediations: Communication Across Media Borders*. Routledge, London, pp. 1–13.

- Feyersinger, E., 2010. Diegetic short circuits: metalepsis in animation. *Animation* 5 (3), 279–294.
- Fludernik, M., 2003. Scene shift, metalepsis, and the metaleptic mode. *Style* 37 (4), 382–400.
- Gass, W., 1970. Philosophy and the form of fiction. In: Scholes, R. (Ed.), *The Philosopher Critic*. The University of Tulsa, Tulsa, OK, pp. 50–66.
- Gee, J.P., 2007. *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave, New York.
- Genette, G., 1980. *Narrative Discourse: An Essay in Method* (J. Lewin, Trans.). Cornell University Press, Ithaca, NY.
- Genette, G., 1988. *Narrative Discourse: Revisited* (J. Lewin, Trans.). Cornell University Press, Ithaca, NY.
- Genette, G., 1997. *Paratexts: Thresholds of Interpretation* (J. Lewin, Trans.). Cambridge Press, New York, NY (Original Work Published 1987).
- Groenke, S.L., Youngquist, M., 2011. Are we postmodern yet? Reading “Monster” with 21st-century ninth graders. *J. Adolesc. Adult Literacy* 54 (7), 505–513.
- Gross, M., Latham, D., 2017. The peritextual literacy framework: using the functions of peritext to support critical thinking. *Libr. Inf. Sci. Res.* 39 (2), 116–123.
- Gross, M., 2019. Grounding our work theoretically: the peritextual literacy framework. In: Witte, S., Latham, D., Gross, M. (Eds.), *Literacy Engagement Through Peritextual Analysis*. American Library Association, Chicago, IL, pp. 3–15.
- Hanebeck, J., 2017. *Understanding Metalepsis: The Hermeneutics of Narrative Transgression*. De Gruyter, Berlin, Germany.
- Hassan, I., 1997. Toward a concept of postmodernism. In: Waugh, P. (Ed.), *Revolutions of the Word: Intellectual Contexts for the Study of Modern Literature*. Arnold, New York, pp. 95–101.
- Hateley, E., 2013. *The Value of Books: The Fantastic Flying Books of Morris Lessmore and the Social Hieroglyphic of Reading*. Paper Presented at the 40th Annual Children's Literature Association Conference, Biloxi, MS.
- Henkel, A.Q., 2018. Exploring the materiality of literary apps for children. *Child Lit. Educ.* 49, 338–355.
- Herman, D., 2005. Histories of narrative theory (1): a genealogy of early developments. In: Phelan, J., Rabinowitz, P.J. (Eds.), *A Companion to Narrative Theory*. Blackwell Publishing Ltd., Oxford, UK, pp. 19–34.
- Holtzman, S., 1997. *Digital Mosaics: The Esthetics of Cyberspace*. Simon & Schuster, New York.
- Hutcheon, L., 1980. *Narcissistic Narrative: The Metafictional Paradox*. Wilfrid Laurier University Press, Waterloo, ON.
- Hutcheon, L., 2002. *The Politics of Postmodernism*, second ed. Routledge, London, UK.
- Hutcheon, L., 2013. *Narcissistic Narrative: The Metafictional Paradox*. Wilfrid Laurier University Press, Waterloo, ON.
- Jenkins, H., 2006. *Convergence Culture: Where Old and New Media Collide*. New York University Press, New York.
- Jenkins, H., 2013. *Reading in a Participatory Culture: Remixing Moby Dick in the English Classroom*. Teachers College Press, New York.
- Jewitt, C. (Ed.), 2014. *The Routledge Handbook of Multimodal Analysis*, second ed. Routledge, London, UK.
- Kachorsky, D., Reid, S., 2020. The covers of young adult novels and their graphic novel adaptations: a qualitative multimodal content analysis. *Literacy Res.* 69, 303–391.
- Kachorsky, D., Aguilera, E., Gee, E., Serafini, F., 2016. Expanding analytical perspectives on children's picturebook apps. *Literacy Res.* 1–15.
- Kerchy, A., 2018. Translation and transmedia in children's literature. *Bookbird* 56 (1), 4–9.
- Kress, G., 2015. Semiotic work: applied linguistics and a social semiotic account of multimodality. *AIRA Rev.* 28 (1), 49–71.
- Kucirkova, N., 2017. An integrative framework for studying, designing and conceptualizing interactivity in children's digital books. *Br. Educ. Res. J.* 43 (6), 1168–1185.
- Kucirkova, N., 2019. Theorizing materiality in children's digital books. *Libri Liberi* 8 (2), 279–292.
- Kukkonen, K., Klimek, S. (Eds.), 2011. *Metalepsis in Popular Culture*. De Gruyter, Berlin, Germany.
- Kukkonen, K., 2011. Metalepsis in popular culture: an introduction. In: Kukkonen, K., Klimek, S. (Eds.), *Metalepsis in Popular Culture*. De Gruyter, Berlin, Germany, pp. 1–21.
- Kümmerling-Meibauer, B., 2013. Paratexts in children's films and the concept of meta-filmic awareness. *J. Educ. Media Memory Soc.* 5 (2), 108–123.
- Leavenworth, M.L., 2015. The paratext of fan fiction. *Narrative* 23 (1), 40–60.
- Lewis, D., 1990. The constructedness of texts: picture books and the metafiction. *Signal* 62, 131–146.
- Lewis, D., 2001. *Reading Contemporary Picture Books: Picturing Text*. RoutledgeFalmer, New York.
- Lindhe, C., 2010. Bildseendet fods i fingertopparna: Om en ekfras for den digitale tidsaldern. *Nordic J. Visual Cult.* 1, 1–12.
- Mackey, M., 2008. Postmodern picturebooks and the material conditions of reading. In: Sipe, L.R., Pantaleo, S. (Eds.), *Postmodern Picturebooks: Play, Parody, and Self-Referentiality*. Routledge, London, UK, pp. 103–116.
- Mackey, M., 2011. The case of the flat rectangles: children's literature on page and screen. *Int. Res. Child. Lit.* 4 (1), 99–114.
- Mackey, M., 2015. Digital picturebooks. In: Arizpe, E., Styles, M. (Eds.), *Children Reading Picturebooks: Interpreting Visual Texts*, second ed. Routledge, New York, pp. 169–181.
- McCallum, R., 1996. Metafiction and experimental work. In: Hunt, P. (Ed.), *International Companion Encyclopedia of Children's Literature*. Routledge, New York, NY, pp. 397–409.
- McCallum, R., 2004. Metafiction and experimental work. In: Hunt, P. (Ed.), *International Companion Encyclopedia of Children's Literature*, second ed., vol. 1. Routledge, New York, NY, pp. 587–598.
- McCallum, R., 2008. Would I lie to you? Metalepsis and modal disruption in some “true” fairy tales. In: Sipe, L.R., Pantaleo, S. (Eds.), *Postmodern Picturebooks: Play, Parody, and Self-Referentiality*. Routledge, New York, NY, pp. 188–192.
- McCracken, E., 2013. Expanding Genette's epitext/peritext model for transitional electronic literature: centrifugal and centripetal vectors on kindles and iPads. *Narrative* 21 (1), 105–124.
- Nikolajeva, N., 2008. Play and playfulness in postmodern picturebooks. In: Sipe, L.R., Pantaleo, S. (Eds.), *Postmodern Picturebooks: Play, Parody, and Self-Referentiality*. Routledge, London, UK, pp. 53–74.
- O'Halloran, K.L. (Ed.), 2004. *Multimodal Discourse Analysis: Systemic Functional Perspective*. Continuum, London, UK.
- Pantaleo, S., Sipe, L.R., 2012. Diverse narrative structures in contemporary picturebooks: opportunities for children's meaning-making. *J. Child. Lit.* 38 (1), 6–15.
- Pantaleo, S., 2004a. The long, long way: young children explore the fabula and syuzhet of shortcut. *Child Lit. Educ.* 35 (1), 1–20.
- Pantaleo, S., 2004b. Young children interpret the metafiction in Anthony Browne's voices in the park. *J. Early Child. Literacy* 4 (2), 211–233.
- Pantaleo, S., 2004c. What do four voices, a shortcut and three pigs have in common? Metafiction! *Bookbird* 4 (1), 4–12.
- Pantaleo, S., 2008a. Ed Vere's the Getaway: starring a postmodern cheese thief. In: Sipe, L.R., Pantaleo, S. (Eds.), *Postmodern Picturebooks: Play, Parody, and Self-Referentiality*. Routledge, New York, NY, pp. 238–255.
- Pantaleo, S., 2008b. *Exploring Student Response to Contemporary Picturebooks*. University of Toronto Press, Toronto, ON.
- Pantaleo, S., 2010a. Developing narrative competence through reading and writing metafictional texts. *Literacy Res. Instr.* 49 (3), 264–281.
- Pantaleo, S., 2010b. Emily Gravett's postmodern picturebook wolves. *J. Child. Lit.* 36 (1), 51–59.
- Pantaleo, S., 2019. Metalepsis in elementary students' multimodal narrative representations. *Res. Teach. Engl.* 54 (1), 8–30.
- Phelan, J., Rabinowitz, P.J., 2005. Glossary. In: Phelan, J., Rabinowitz, P.J. (Eds.), *A Companion to Narrative Theory*. Blackwell Publishing, Oxford, UK, pp. 542–551.
- Pier, J., 2011. Afterword. In: Kukkonen, K., Klimek, S. (Eds.), *Metalepsis in Popular Culture*. De Gruyter, Berlin, DE, pp. 268–279.
- Pier, J., 2016. Metalepsis (revised version). In: Hühn, P., et al. (Eds.), *The Living Handbook of Narratology*. Hamburg University, Hamburg, Germany. Retrieved from: <http://www.lhn.uni-hamburg.de>.
- Prince, G., 2003. *Dictionary of Narratology*, revised ed. University of Nebraska Press, Lincoln, NE.
- Rafaeli, S., 1988. Interactivity: from new media to communication. *Sage Annu. Rev. Commun. Res.* 16, 110–134.
- Reid, S., Serafini, F., 2018. More than words: an investigation of the middle-grade multimodal novel. *J. Child. Lit.* 44 (2), 32–44.
- Ryan, M.-L., 2004. Metafictional machines. *Semiotica* 150 (1), 439–469.
- Sargeant, B., 2015. What is an ebook? what is a book app? and why should we care? an analysis of contemporary digital picture books. *Child. Lit. Educ.* 46 (4), 454–466.
- Schwebs, T., 2014. Affordances of an app. *Nordic J. Child Lit. Aesthetics* 5, 1–11.

- Schwenke-Wyile, A., 2006. The drama of potentiality in metafictional picture books: engaging in pictorialization in *Shortcut*, *Ohh-la-la* and voices in the park (with occasional assistance from A. Wolf's *True Story*). *Child Lit. Assoc. Q.* 31 (2), 176–196.
- Serafini, F., Reid, S., 2019. Crossing boundaries: exploring metaleptic transgressions in contemporary picturebooks. *Child Lit. Educ.* 51, 261–284.
- Serafini, F., Kachorsky, D., Aguilera, E., 2015. Picturebooks 2.0: transmedial features across narrative platforms. *J. Child. Lit.* 41 (2), 16–24.
- Serafini, F., Kachorsky, D., Aguilera, E., 2016. Picturebooks in the digital age. *Read. Teach.* 69 (5), 509–512.
- Serafini, F., 2005. Voices in the park, voices in the classroom: readers responding to postmodern picture books. *Read. Res. Instruct.* 44 (3), 47–65.
- Serafini, F., 2014. *Reading the Visual: An Introduction to Teaching Multimodal Literacy*. Teachers College Press, New York.
- Silva-Díaz, M.C., 2018. Picturebooks and metafiction. In: Kümmerling-Meibauer, B. (Ed.), *The Routledge Companion to Picturebooks*. Routledge, London, UK, pp. 69–80.
- Sim, S. (Ed.), 2011. *The Routledge Companion to Postmodernism*, third ed. Routledge, London, UK.
- Sipe, L.R., Pantaleo, S. (Eds.), 2008. *Postmodern Picturebooks: Play, Parody, and Self-Referentiality*. Routledge, London, UK.
- Sipe, L.R., 1998. Learning the language of picturebooks. *J. Child. Lit.* 24 (2), 66–75.
- Sipe, L.R., 2008. *Storytime: Young Children's Literary Understanding in the Classroom*. Teachers College Press, New York, NY.
- Soyland, L., Gulliksen, M.S., 2019. Sense-making through touch interaction with a picturebook app. *Universitetsforlaget* 10 (1), 1–12.
- Stanitzek, G., 2005. Texts and paratexts in media (Ellen Klein Trans.). *Crit. Inq.* 32 (1), 27–42.
- Stichnothe, H., 2014. Engineering stories? A narratological approach to children's book apps. *Nordic J. ChilLit Aesthetics* 5. <https://doi.org/10.3402/blft.v5.23602>.
- Thon, J.-N., 2016. *Transmedial Narratology and Contemporary Media Culture*. University of Nebraska Press, Lincoln, NE.
- Turrion, C., 2014. Multimedia book apps in a contemporary culture: commerce and innovation, continuity and rupture. *Nordic Journal of ChilLit Aesthetics* 5.
- Unsworth, L., 2006. *E-Literature for Children: Enhancing Digital Literacy Learning*. Routledge, London.
- Viswanath, T., 2018. Issues of control and agency in staged metafictional picture books. *Child Lit. Assoc. Q.* 43 (1), 67–83.
- Voorveld, H.M., Neijens, P.C., Smit, E.G., 2011. The relation between actual and perceived interactivity. *J. Advert.* 40 (2), 77–92.
- Waugh, P., 1984. *Metafiction: The Theory and Practice of Self-Conscious Fiction*. Methuen, New York, NY.
- Witte, S., Latham, D., Gross, M. (Eds.), 2019. *Literacy Engagement through Peritextual Analysis*. American Library Association, Chicago, IL.
- Wohlwend, K., 2017. Toddlers and touchscreens: learning “concepts beyond print” with tablet technologies. In: Meyer, R.J., Whitmore, K.F. (Eds.), *Reclaiming Early Childhood Literacies: Narratives of Hope, Power, and Vision*. Routledge, New York, pp. 64–75.
- Wolf, S., Coats, K., Enciso, P., Jenkins, C.A. (Eds.), 2011. *Handbook Of Research on Children's and Young Adult Literature*. Routledge, New York, NY.
- Wolf, W., 2013. “Unnatural” metalepsis and immersion: necessarily incompatible? In: Abler, J., Nielsen, H.S., Richardson, B. (Eds.), *A Poetics of Unnatural Narrative*. The Ohio State University Press, Columbus, OH, pp. 113–141.
- Yokota, J., Teale, W.H., 2014. Picture books and the digital world. *Read. Teach.* 67 (8), 577–585.
- Zheng, Y., 2016. Anything new here in story apps? A reflection on the storytelling mechanism across media. *Libri Liberi* 5 (1), 55–76.

Children's literature references

- Arbona, M., 2020. *Window*. Kids Can Press, Toronto, ON.
- Catalanotto, P., 2007. *Ivan the Terrier*. Simon & Schuster Children's Publishing, New York.
- Chih-Yuan, C., 2008. *Artie and Julie*. Heryin Books, Alhambra, CA.
- Childs, L., 2002. *Who's Afraid of the Big Bad Book?* Hyperion Books for Children, New York.
- Dunlap, C., Tallec, O., 2017. *This Book Will Not Be Fun*. Random House, New York.
- Falatto, J., Miller, T., 2016. *Snappy the Alligator (Did Not Ask to Be in This Book!)*. Viking, New York.
- Gilman, P., 1992. *Something From Nothing*. Scholastic, Richmond Hill, ON.
- Gravett, E., 2005. *Wolves*. Macmillan Children's Books, London.
- Gravett, E., 2006. *Meerkat Mail*. Macmillan Children's Books, London.
- Hawkins, C., Hawkins, J., 2004. *Fairytale News*. Candlewick Press, Cambridge, MA.
- Holub, J., Slack, M., 2008. *Knuckle Heads*. Chronicle Books, San Francisco.
- Joyce, W., 2012. *The Fantastic Flying Books of Mr. Morris Lessmore*. Atheneum, New York.
- Joyce, W., 2014. *The Numberlys*. Atheneum, New York.
- Kelly, E.E., 2017. *Hello, Universe*. Greenwillow Books, New York.
- Lendler, I., Martin, W., 2005. *An Undone Fairy Tale*. Simon & Schuster, New York.
- Palacio, R.J., 2012. *Wonder*. Knopf Books for Young Readers, New York.
- Park, L.S., 2005. *Project Mulberry*. Houghton Mifflin Company, New York.
- Pilkey, D., 2016. *Dogman*. Graphix, New York.
- Scheerger, 2019. *Operation Frog Effect*. Yearling, New York.
- Stein, D.E., 2010. *Interrupting Chicken*. Candlewick Press, Somerville, MA.
- Stone, J., 2003. *The Monster at the End of the Book*. Golden Books, New York.
- Vere, E., 2006. *The Getaway*. Puffin Books, London.
- Weeks, S., Varadarajan, G., 2016. *Save Me a Seat*. Scholastic, New York.
- Wiesner, D., 2001. *The Three Pigs*. Clarion Books, New York.

Global literature for young people: critical issues and social responsibilities

Kathy G. Short^a and Junko Yokota^b, ^a University of Arizona, Tucson, AZ, United States; and ^b National Louis University, Chicago, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

The challenge of defining global literature	633
The possibilities of global literature in the lives of readers	634
The social responsibility of readers	635
The social responsibility of bookmakers	636
The challenge of availability and access	636
The challenge of accuracy and authenticity	637
The challenge of cultural representations	639
The social responsibility of influencers	640
Researchers of global children's literature	640
University instructors of global children's literature	640
Educators in schools and libraries	641
Resources on global literature	641
Conclusion	642
References	642

The life of Jella Lepman epitomizes the belief that global literature has the potential to open children's minds as a means of transforming society. A German Jew, Lepman fled to London when Hitler came to power, returning almost immediately after the war in response to an invitation to establish cultural programs for children. The decision to return to Germany in 1945 was extremely difficult, given that millions of her people had been killed, but she believed that if children, who only knew their wartime country, were to have a future they needed books to view the world in new ways. Her plea for donations brought in 4000 books from publishers in 14 countries for the first book exhibition, leading to the founding of the International Youth Library (IYL) in Munich. The following year, Lepman founded the International Board on Books for Young People (IBBY), inspired by Eric Kästner's (1949) *Conference of the Animals*, in which animals come together to work toward a common purpose. Seventy years later, the IYL has over 650,000 books and IBBY has 81 countries as members, supporting projects that touch the lives of thousands of children worldwide. The idea that books can build bridges of understanding reaches around the world, built by Lepman's (2002) capacity for forgiveness and compassion and her belief in the transformative potential of global literature.

Lepman's life is a reminder that adults influence what and how children read, particularly in opening children's minds through global literature to provide understandings that go beyond their life experiences. The influence of adults on children's options and positions as readers cuts across all aspects of creating and marketing books and inviting children to connect with these books. In organizing this chapter on the state of the art in global literature, we decided to use the lens of social responsibility to consider children as readers and the adult bookmakers and influencers who surround these books. First, however, the debate about how to define global literature is addressed along with why these books matter in the lives of readers.

The challenge of defining global literature

The inclusion of global children's literature in this encyclopedia assumes that a shared definition exists within the field and across global settings. The reality is that a range of terms are used to label this body of literature, even within the same country, and that global literature is not always viewed as a separate category. Global literature is often not in common usage in countries, such as New Zealand, in which only a small number of books are local with the majority coming from other countries. Their focus is on identifying local books for their children with everything else labeled as children's literature. In contrast, in Western countries like the United States (US), the majority of books are local, in the sense of published within that country by authors from that country about that country, so that only a small number of books are set in other places around the world. In these situations, global literature is an important category to ensure that children have the opportunity to immerse themselves into story worlds set in global cultures. A further complication is that in some countries, such as Germany, differentiating a set of books as global is viewed as unnecessary. So even the recognition of a set of books labeled global literature is not shared globally, let alone a definition.

In this chapter, we define global literature as any book set in a country or global culture outside of the reader's own culture. For a Canadian, any book set outside Canada is considered global, no matter who is the author. A book set in Poland by a Canadian author is global literature as is a fantasy written by a Japanese author in Japanese and then translated into English and/or French for Canadian children. The definition of which books are considered global thus depends on a reader's sense of identity, which is further complicated by children who have more than one cultural identity. Immigrant and refugee children often have multiple

identities, such as both Greek and Syrian, as do “third culture kids” (Pollock and Van Reken, 2001) who spend their childhood moving frequently to new global locations and who often never live in either parent’s country of citizenship.

International literature is one type of global literature, consisting of books originally created and published in a country outside of the reader’s location and then made available in that reader’s country. For a Canadian child, an international book might be published first in Australia in English and then adapted for English publication in Canada, or a book published in Spanish in Chile might be translated into English and/or French in Canada. The important distinction is that these books were first published for children of a country in their own language by an author from that country, before being imported and either translated or adapted for children in a new country.

World literature is a term often used in secondary schools and universities to refer to canonical literature considered exemplary for a geographical region. These books are usually adult classics and are read as high-quality literature representative of those cultures. We are not including world literature since our focus is children’s and young adult literature.

Finally, multicultural literature is sometimes used to refer to global literature. More commonly, multicultural literature indicates the diversity of cultures *within* a country or geographical location, not across global settings. In the US, multicultural literature includes books about groups who are marginalized, stereotyped, or underrepresented by the dominant European-American culture within the US. Multicultural literature thus includes groups marginalized based on race, ethnicity, religion, language, ability, gender identity, sexual orientation, social class, and so on. Since this chapter highlights diversity across global cultures, rather than within a national boundary, multicultural literature is not included.

This chapter thus focuses on global literature, recognizing that this definition includes books and bookmakers from many cultural locations and is based in a definition of the reader’s cultural location. This complexity of defining global literature also acknowledges differing cultural conceptions of global literature’s centrality based on each country’s history of creating literature for children.

The possibilities of global literature in the lives of readers

As long-term advocates of global literature, we are committed to the value of these books in adding possibilities to the lives of readers, both children and the adults who interact with them around these books, including educators, parents, and librarians. We have seen firsthand how global literature can expand readers’ views and encourage open-mindedness toward ways of living and thinking that differ from their own. At the same time, we are painfully aware that asking readers to read outside their comfort zones holds danger and can establish misunderstandings and perpetuate stereotypes.

The possibilities of global literature are grounded in theories of story as a primary act of mind, as the way in which we bring meaning to our experiences and make sense out of the chaos of our daily lives (Rosen, 1986). We create stories in our minds to understand our experiences, so that these stories do not just recount those events but also interpret them. Stories are thus what makes us human—the nature of a life is that it is a story. A web of interconnected stories function in our minds to create our views of the world as we continuously add and rearrange these stories and their connections (Short, 2012).

An understanding of story as world-making underlies the potentials of reading globally, especially in encouraging narrative imagination by inviting readers into story worlds to experience how people, live, feel, and think in global cultures. Through literature, readers immerse themselves into story worlds, providing experiences that go beyond surface-level tourist information to deeper cultural values and beliefs. As readers engage with characters in these story worlds, they develop emotional connections and empathy as well as knowledge about a culture. They learn from, not merely tolerate, those whose views and ways of living differ from their own. They also come to know their own home cultures better through exploring worlds beyond their homes. These understandings grow out of recognizing the common values they share with those living in unfamiliar cultural communities along with the unique differences that distinguish each culture. Experiences with global literature can encourage a stance of openness to how people view themselves, not just how we view them (Short et al., 2016).

Some books feel comfortable to readers even when that book is in an unfamiliar cultural context, and so readers are encouraged to connect with children from those places, providing a sense of community across global borders. Other books stretch readers to go beyond what they know and expand their worldviews to new perspectives, information, and insights that highlight the diversity and richness of the global world (Yokota and Teale, 2017). Still other books, however, are so disconnected from readers’ literary and life experiences that they respond with discomfort and rejection. By surrounding these books with additional experiences and texts, supportive contexts can be created for building understanding, instead of dismissal.

The potentials of global literature thus must be balanced with recognizing the dangers of reading globally and addressing these dangers through our social responsibilities as readers, bookmakers, and influencers (Short, 2019). Social responsibility is a system of ethics from the field of business that argues decisions and actions must be ethically considered along with financial gain, specifically if an action or decision causes harm to people, society or the environment (Chandler, 2016). Rosenblatt (1994) extends this concept, arguing that responsibility as a reader and writer balances individual voice with group responsibility. The right to expression of voice must be weighed against the consequences for others and their right to have their voices heard. Social responsibility thus combines the conviction and enthusiasm of each person’s ideas with an open mind to multiple perspectives and needs in the broader community.

Building on Rosenblatt, theorists use the lens of cosmopolitanism to discuss our responsibility as inhabitants of the cosmos. Hansen (2011) views this orientation as balancing reflective loyalty of the known with reflective openness to the new to create

a philosophy *for* and *of* life rather than about life. Choo (2014) argues literature supports readers in using imagination to critically engage in a stance of hospitable openness to others. Literature functions as a tool for accessing the thinking and beliefs of another person or community, becoming an entry for dialogic inquiry about issues that need to be addressed in order to live fully with diverse peoples in the world. This chapter builds on this idea to explore the responsibilities of bookmakers, readers, educators, researchers, reviewers, and others in relation to global literature.

Before we move into these discussions, we acknowledge our responsibilities as authors of this chapter, recognizing the influence of our backgrounds on our views of the world. Both of us are academics from the field of education in universities in the United States, who also were educators in schools. Our expertise is children's and young adult literature and literacy, with a focus on global and international books and a passion for the ways these books can transform the lives of readers. We are both active in international organizations and award committees, especially IBBY. Kathy has always lived in the US but works with primary teachers around the world to create contexts for dialogue and inquiry around global literature and with researchers on critical content analysis. She is the founding director of a center on global literacies and literatures, Worlds of Words, that offers many resources to educators. Junko was born and raised in Japan, attended one year of high school in the US, and returned for college before becoming an adult immigrant. She is the founding director of the Center for Teaching through Children's Books, which supports teachers and librarians in considering the ways books are central to their teaching. Her research, teaching, and speaking takes her around the world to work with teachers, librarians, and book creators as well as literacy and literature advocates.

Just as we identify our positionality, readers of this chapter are encouraged to examine their perspectives around issues of privilege, specifically the awareness that some are granted unearned privileges that others are denied because of how history and society enact value and oppression due to the cultural group in which people are perceived to belong. Awareness of these societal inequities challenges us to think, read, and face the world to examine how those perspectives influence our decisions in what and how we read, and how we can go beyond talk to take action in addressing these inequities. As authors and readers of this chapter, most of us bring the privilege of advanced education and membership in a community of scholars and so must consider our privilege and resulting responsibilities.

The social responsibility of readers

Although much of the scholarship on global literature focuses on textual authenticity and representations, we believe that global books matter because of their potential for readers. This chapter, thus, first addresses the social responsibilities of readers before considering the responsibilities of bookmakers and adults who influence the field. This section is organized around issues related to balancing critical loyalty to one's belonging with critical openness to the new as readers (Hansen, 2011), and evaluating the extent to which readers share a cultural match with a book (Johnson, 2016).

Without this positioning of social responsibility, the danger is that readers judge a book based on their culture as the norm, not recognizing that they are reading from their own cultural locations. Cultural mismatches can occur due to unfamiliar story structures, differing uses of literary elements, unfamiliar illustration styles, unknown terminology or names, and lack of knowledge about specific contexts and events. Readers may struggle with books from cultures that use more inference, description, or action than books from their culture. These differences can result in readers rejecting a book as unreadable or feeling condescension and superiority based on their cultural lens, particularly readers who read from positions of privilege due to their country's political, economic, or social status. This positioning can establish stereotypes if readers focus too much on difference and judge characters as exotic or strange. Readers can also fall into the trap of universalizing, failing to recognize the ways in which another culture is unique and differs from their own and so viewing people from that culture as the same "on the inside," a position of colorblindness.

An understanding of the social responsibility of readers is based in a dynamic view of culture, which Geertz (1973) defines as the shared patterns that set the tone, character, and quality of people's lives. These patterns are webs of significance that act as shared codes of meaning between people. González (2005) extends this definition to highlight culture as practice, including what people do, what they say about what they do, and why they do it, which she labels as funds of knowledge. Lived experiences are thus interwoven with identity markers to create cultural locations from which readers interact with books and the world. Cultural locations are dynamic, constantly shifting as readers move across time and place, but remain attached to a core of loyalty to the cultural resources and values in their hearts.

Brooks and Browne (2012) argue that readers engage in culturally situated response, a home from which they make sense of the world and of texts through reading as a transactional process in a dialogue between readers and texts. Responses to a book are always culturally situated and so readers must examine how their cultural locations influence their understandings and the perspectives they bring to meaning-making. Readers, therefore, must be aware of the mismatch or gap between the cultural community portrayed in a book and their own cultural locations and reflect critically on *why* they have a particular response to a book. This awareness involves both naming and reflecting on their cultural locations as readers to critique their biases and assumptions.

Reflective loyalty to the known (Hansen, 2011) challenges readers to search for emotional connections with characters instead of focusing only on personal connections. Encouraging readers to make personal connections to a global book where the experiences and settings are unfamiliar can lead readers to judge that culture against their own, resulting in perspectives of pity and self-importance. Instead, an emphasis on empathy, on finding something within one's self related to the emotions of characters, such as loneliness or fear, invites a stance of listening and connection through "feeling as" (Brown, 2018; Cuff et al., 2016). In contrast, sympathy involves "feeling for" and leads to feeling sorry for someone's misfortunes and wanting to make things better

for them, fueling disconnection and pity from a stance of privilege. Empathy does not assume that one person can fully stand in the shoes of someone else, but that a reader can make an emotional connection to a character whose experiences differ from one's own. More importantly, empathy invites readers to read from a position of shared vulnerability, rather than from a position of privilege.

Reflective openness to the new provides a further challenge (Hansen, 2011). Instead of expecting personal connection, the challenge is for readers to read from a stance of becoming more conscious and committed to others and to moving beyond self to responsible engagement to the other. Instead of reading for pleasure, readers engage with books out of obligation and responsibility to others and accept that the reading may be intellectually discomforting. Socially responsible readers focus on learning from the experiences and perspectives of those whose cultures differ from their own, so they can learn from, not merely tolerate, difference. Further, readers recognize that worlds unfamiliar to them are possible worlds that exist and hold value. Readers thus balance reason with emotion to open themselves to learning from the new (Rosenblatt, 1938). This position of learning from difference may lead readers to engage in further research on cultural contexts and events because of their own recognized lack of knowledge related to a particular book; for example, readers who do not understand the deep roots of racial tension in the US.

Choo (2014) argues that the reader must be willing to engage the other and focus on difference, rather than uniformity, through imaginatively entering the other's world and taking an other-centered orientation. This stance of hospitable imagination involves an orientation of openness *without limit* where readers struggle to understand new perspectives and experiences without imposing judgment from their cultural locations. Choo notes that readers need a home from which to be hospitable, but their openness to the new means they imagine a new reality while still holding on to their own sense of self. Although readers do not let go of themselves to host the other, they are willing to be influenced by that new reality. In opening their minds and immersing themselves in new experiences through literature, readers remain willing to shift their perspectives and accept the possibility of changing their cultural locations as readers and human beings. Wahlström (2014) notes this stance is a tension between resistance and receptivity for readers in a continuous process of both *being* and *becoming*.

Heise (2008) extends this stance of openness to include the universe, earth, and nature, a position of eco-cosmopolitanism. This orientation challenges human supremacy through recognition of the interconnectedness and codependence of humans and the natural world. Heise argues that humans and nonhumans are part of planetary imagined communities, global networks sharing a responsibility based in a sense of planet rather than a sense of place. Her views extend the social responsibility of the reader to becoming open to the experiences of nature through a stance of empathy and learning from nature.

The social responsibility of bookmakers

The words and illustrations in children's books are extensively critiqued in scholarly literature around issues of accuracy, authenticity, and representation. Our focus on bookmakers recognizes that a book is created by a team that includes the author, illustrator, translator, agent, editor, art director, publicist, and others in the publishing company. This team has the same social responsibilities as readers in reflecting on their cultural locations and being open to the other as they create and promote a book. They are responsible to research details of everyday life and language for a global community, understand the core beliefs and values of that community, and explore existing books and media images to determine the currently available representations of that culture. They consider how to balance the individual voice of creative expression within a book with the social good of society, both locally and globally, so that a book does not do harm through misrepresentations of a global culture.

The challenge of availability and access

The availability of global literature is dependent on the books selected for publication, translation, and distribution across global boundaries. Books from countries publishing large numbers of books in English, including the US, UK, Australia, New Zealand, and Canada, dominate the global market, so much so that in many countries the largest percentage of books are translated from English into the local language, providing only a few books that reflect children's own culture. Joosen (2020) states that translation policies have dramatically affected availability, resulting in English books being read all over the world while the rest of the world remains silent and invisible for readers. Seventy to ninety percent of the books available in countries outside the US and Europe are by European and American authors with few books from these countries crossing linguistic, political, and cultural divides to become available in Western markets (O'Sullivan, 2011).

Sixty years ago, Haviland (1961) wrote that European bookstores had too many imports from the US, indicating a problem with the flow of US books. Decades later, this flow has spread to other continents beyond Europe, and the reverse flow of books into the US continues to be significantly lower. English language books dominate because of the accessibility of the English language. The world reads in translation what is first published in English, but English-language countries see little of what the world is publishing in other languages, especially lesser-known languages. For many years in China, picturebooks originally published in English were translated for the Chinese market in enormous numbers, especially award-winning US books. This trend was followed by pairing a Chinese author with an established award-winning illustrator from a Western country. Recently, more books are created by Chinese authors and illustrators; yet it is the rare few that make their way into the English market. This problem of flow is compounded because acquisition editors and publisher teams typically do not know multiple languages and rely on reader-reports from those who read that language and interpret the appropriateness of a book for a new audience.

In examining recent global publication trends, one major upswing is books on trauma. The details of the trauma vary from abuse to child labor to the aftermath of natural disasters and war, from refugees to pandemics to systemic racism. Such books go beyond the individual traumas of death and divorce in earlier books to recognize how worldwide traumas impact children. One major increase due to the global migrant crisis is racially and culturally diverse books reflecting the experiences of refugees. Previously, many of the books addressing diversity were viewed as uniquely an American problem. The displacement of millions of Syrian refugees into areas without much diversity, compounded by large numbers of displaced peoples in other parts of the world, has created an interest in refugee books. Most, however, are written from the perspective of encouraging European and American children to act with compassion toward refugees, instead of depicting children who are displaced acting with their own agency (Vassiloudi, 2019).

Another trend internationally is a rise in the range and style of nonfiction books, especially picturebooks as a place in which the telling of information has thrived. For example, Polish authors and illustrators Aleksandra Mizelińska, Daniel Mizeliński, Piotr Socha, and Wojciech Grajkowski are creating books in a lesser-known language that have become widely translated international sensations. This trend has led to nonfiction categories in illustration competitions and entire bookfairs on nonfiction, such as the Bologna Bookfair where there is a separate category for nonfiction and a 2019 nonfiction conference in St. Petersburg.

One interesting trend is the rise in international collaborations, influenced by international bookfairs and illustration competitions, where publishers become familiar with the work of a wider range of global bookmakers. International competitions such as those sponsored by the Nami Concours in Korea, China Shanghai International Children's Book Fair, and Chen Bochui International Children's Literature Award broaden the possibilities for collaboration. Some authors and illustrators are published first in countries other than their own or published simultaneously across multiple countries. In addition, authors from one country are brought together with illustrators from another, creating a collaboration through a shared international editorial vision.

Books with strong visual images, including picturebooks, graphic novels, and heavily illustrated middle grade novels such as those by Shaun Tan and Liniers, often travel more easily across global boundaries. These books hold appeal because children today are immersed in a visual culture in which images are central to their experiences and interactions, influencing what they know and how they think about the world (Duncum, 2002). Books with interconnected media platforms, including books, films, film, animation, computer games, fanfiction, cosplay, and collectibles, maximize audience engagement and create an increasingly elaborate fictional reality for characters such as Harry Potter, the Moomins, and Alice in Wonderland (Kerchy, 2018). Media provides a new cultural layer and supports more diverse ways of communicating, thinking, and creating.

Another global trend is culturally generic books with universal themes and no location in a specific place and time. Culturally generic books, such as *The Very Hungry Caterpillar* (Carle, 1969), are easy to translate across cultural boundaries and have a large market because they do not challenge readers to move beyond their cultural views. In contrast, culturally specific books combine universal themes with a specific cultural setting in the text or illustrations, and so challenge readers to immerse themselves in another culture and way of thinking. These books, however, can sometimes be difficult for readers in new cultures to understand and so may have smaller market appeal. Uehashi's Moribito series, for example, is very popular in Japan and widely enjoyed by people in Asia, Europe, and Latin America, winning multiple awards. The style of storytelling and grounding in Japanese mythology, however, led to poor sales within the US. The dilemma facing authors from a small country is that writing a generic book can result in large worldwide markets, providing them with financial success, while a culturally specific book connects to experiences of children in their country, but may not be picked up for translation and broader distribution.

Another related issue of availability is the mass marketing of books and related products. Many Western publishing houses are owned by large international entertainment corporations for whom books are only one small segment of their market, and who therefore focus on the child as a consumer rather than as a reader (Hade and Edmondson, 2003). These corporations hope to attract children to any product related to a book that carries the brand's name, such as toys, clothing, games, and videos. These large multimedia conglomerates have a globalizing influence by promoting certain books, such as the Harry Potter series, and their related mass-produced goods across borders, leading to less diversity and more uniformity that threatens regional literature.

The influence of market-driven decisions is challenged by small presses whose mission is cultural diversity and whose decisions about balancing profit margins and losses allow them to concentrate on making the best books, instead of being ruled only by sales and marketing (Settingington, 2016). In some countries, these efforts are furthered by funding from governments and arts organizations to support the publication of children's books about local and global cultures. In Taiwan, for example, the federal government provides funding for local publishers to publish books by Taiwanese authors (Bradford, 2011), some of which are translated for global markets. The efforts and success of these presses have encouraged large presses to take more risks in publishing books set in global cultures.

The availability of a global book within a country does not mean that children get equal access to that book. Who has access varies by the price of books and socioeconomic levels within a country, which determines whether books are affordable. Another issue is whether public libraries and bookstores are present in communities and whether books are easily available in schools and libraries. The availability of a book in different formats, such as an e-book as well as a physical book, also influences access. Finally, access is connected to influencers who promote certain books, such as reviewers, booksellers, bloggers, and award committees.

The challenge of accuracy and authenticity

The social responsibility of bookmakers includes accuracy in details of everyday life and language and authenticity of core values. Accuracy can appear to overemphasize small insignificant details to those who are not from that culture, but those who identify with that culture feel a sense of violation and a lack of respect when details are wrong. Authenticity is more widely discussed because of

the connection to values at the heart of a culture. Although multiple values and ways of living exist within every culture, authenticity focuses on believability, specifically whether a character acts or thinks in ways viewed as possible by those who live within a cultural community. Mo and Shen (2003) point out that although all members of a culture do not share the same values, the themes, beliefs, and actions within a book should fit within the range of values seen as acceptable within that culture.

Authenticity can be viewed through the lens of readers' experiences, particularly when readers from within a culture identify and feel affirmed that what they are reading rings true and is believable, while readers from another culture are able to immerse themselves into the culture and learn something of value about that culture (Fox and Short, 2003). Authenticity also includes the absence of stereotypes or the challenge of commonly held stereotypes about a culture within a book (Yoo-Lee et al., 2014).

Instead of simplistic outsider/insider distinctions, authenticity is more productively viewed as a critical lens for considering complex issues, such as the cultural locations of authors, illustrators, and translators as the most visible members of a bookmaking team. The cultural locations of those who create global literature and the impact of these locations on books indicate the complexity of authenticity and why reducing authenticity to insiders and outsiders or ingroup and outgroup are problematic. These complex cultural locations include:

- Authors and illustrators who live in a global culture and write about that culture for children of that culture, with the books then translated for readers in a new global cultures, such as Maria Parr, Cao Wenxuan, Jairo Buitrago, and Nahoko Uehashi.
 - Translated books are considered authentic sources of lived experiences and perspectives from within a culture. One issue, however, is which perspective or group within a country is the focus and the bookmakers' relationships with that group. Also, readers cannot assume that a translated book was written by bookmakers about their own culture. Translated books can be difficult for children from a new culture to understand due to lack of familiarity and so the translator influences the extent to which a book retains the source culture while also meeting the readability needs of the new target culture.
- Authors and illustrators who live and work across global cultures and regularly move between several cultures, including the culture reflected in their books, such as Suzy Lee, Duncan Tonatiuh, and Yuyi Morales.
 - Their knowledge of multiple cultures provides in-depth understandings of the cultural setting of the book and the audiences reading that book.
- Authors and illustrators who are immigrants and write about their country of origin, such as Ifeoma Onyefulu, Rashin Kheiriyeh, Kashmira Sheth, and Jiang Hong Chen.
 - These bookmakers often access cultural values and knowledge along with awareness of their audience in their current location. Their knowledge, however, can be dated, based in childhood memories, rather than contemporary life in that culture, or, depending on their age at the time of immigration, they may have few experiences on which to draw. Immigrants also tend to write about conflict and violence in their country of origin since many were forced to flee their countries, providing a single perspective of life in that place that can be overgeneralized.
- Authors and illustrators who draw on family heritage to write about the family's country of origin but have not lived in that country, such as Linda Sue Park, Nnedi Okorafor, Margarita Engle, and Adib Khorram.
 - These bookmakers can often tap into core values and beliefs of a culture based on their experiences within their families and have access to resources from their communities and families. They can also lack knowledge about cultural details and need to engage in extensive research.
- Authors and illustrators who live in another country for a significant time and create books from that experience, such as Kate Banks, Holly Thompson, and Katherine Marsh.
 - Significant experiences within a culture, particularly over a long time period, can support in-depth understandings of core values and ways of living as well as access to resources within that culture. Some people, however, live within closed communities and keep themselves separate from the global culture of their location, continuing to judge that culture through their own cultural lenses.
- Authors and illustrators who research a cultural community with which they lack experience, although they may visit that place as part of their research, such as Gill Lewis, J. L. Powers, and Eliot Schrefer.
 - These bookmakers are the most vulnerable to issues of authenticity because they lack in-depth experiences from which to draw. Some, however, engage in extensive research, spend time in a place, and develop relationships with people from that culture to open themselves to that cultural perspective. Also, some are reporters, historians, and anthropologists who draw on their research skills.
- Authors and illustrators who work in collaboration with someone from a global community, such as Omar Mohammed and Victoria Jamieson or Elizabeth Laird and Sonia Nimr.
 - One recent trend is for an author or illustrator who is not from the global culture to collaborate with someone from that culture, so the book is co-created with the goal of accuracy and authenticity in representations.

Simplistic distinctions of insiders and outsiders to a culture are thus no longer relevant in a world where people are constantly on the move. The cultural locations of authors, illustrators, and translators do matter as does their awareness of how those locations influence their work. One trend is for authors and illustrators to include extensive backmatter in books and information on their websites to indicate their cultural locations, relationships, and research. That same information on translators is much more difficult to locate. Ultimately, the authenticity of a book depends on bookmakers' ability to engage in hospitable imagination by opening themselves to understanding and hosting cultural perspectives that are not their own.

One continuing issue across bookmakers is that the audience and markets for most global books is dominated by English-speaking and European countries. Children from a global culture outside of these countries are rarely the intended audience; they are being written *about* rather than *for*. The characters in books set in Africa and Latin America, for example, are frequently viewed through the eyes of a Western child as audience.

The translator is a highly influential member of the bookmaking team for a translated book, yet sometimes is not recognized as such, especially in picturebook translation. Translators are accorded high regard in some countries, while in other countries the translator is not given that respect and may not even be named. Picturebook translators in Japan are highly regarded, such as well-known author, Momoko Ishii, who both writes and translates for children with her work so highly regarded that she is the focus of a book on the child appeal of her translations (Takeuchi, 2014). In contrast, the award-winning *Hiroshima No Pika* (Maruki, 1980) was done as a “rough translation” for the purpose of selling rights, and the American editor (who did not speak Japanese) smoothed out the translation. Three different English translations of this book in three countries have wording choices with significant political impact (Yokota et al., 2007).

Translations allow readers of one language to read books originally written in an unfamiliar language. The goal of the translator is for readers of the translation to have an emotional and aesthetic response parallel to that of readers of the original text (Grossman, 2010). The decisions made by translators are complex with enormous impact on accessibility as well as interpretation for the new audience. The translator must consider verb tense, proper names, humor, idioms, dialogue, what is considered acceptable in illustration, and the voice in which a story is narrated, to name only a few of the many ways a translator interprets a book in a new language for a new audience (Lathey, 2006). The translator must constantly consider foreignizing versus domesticating to balance retaining specific cultural details with changes so readers from the target audience can make sense of the book (Eco, 2001). Children should recognize a book as coming from another culture and expect to struggle with understanding since the culture is unfamiliar, but not to the point that readers are frustrated and abandon the book. What is important to recognize is that translation changes the book’s meaning and to acknowledge that translated books have three creators—the author, the illustrator, and the translator—each of whom has colleagues who support their roles.

Although some believe the language of pictures is universal, illustrations have cultural nuances in visual communication. The limited text in picturebooks may mean that each sentence’s structure, punctuation, font, and placement on the page communicates in a visual way. Illustrations may need to be called to the reader’s attention or explained in ways that are not necessary in one culture (Oittinen, 2008), given that a lack of understanding illustrations can lead to misunderstanding the book. Miyake (2002) argues that picturebooks can be difficult to translate because the simpler the texts, the deeper the roots in culture.

The challenge of cultural representations

A book can be accurate and authentic but problematic in terms of representation, which refers to the relationship of an individual book to the range of available images about a global culture across books and media. When a visual image such as camels and sand in the Middle East, a characterization such as oppressed Muslim girls, or a topic such as the potato famine in Ireland, are overrepresented with few other images available, another book with the same image contributes to existing stereotypes. Issues of representation also arise when only a few books exist about a cultural community, leading to those books viewed as representing the entire culture. Although a book can cause individual harm through inaccuracy and inauthenticity, representation leads to collective harm through the dominance of images that together create stereotypes even when the individual books are not problematic.

One representational issue is power relationships and agency, particularly who solves a problem, such as in refugee books where a child befriends a refugee child within a Western context as a place of safety. These issues of power are also evident in comparison books, where a child’s life in India or an African country is compared to a child’s life in the West, portraying one as a life of poverty in remote rural areas and the other as a life of financial means in a modern urban city. This portrayal is further complicated in that these books portray a comparison of a white child with a child of color, who is found “lacking” by readers, since the audience is typically Western children reading from a position of privilege.

Because Europe and the US dominate worldwide markets, books from these countries that portray life in Latin American and African countries are usually created by bookmakers from the West. Often non-Western countries are portrayed through historical fiction and traditional literature, with few contemporary depictions, and through rural and village settings, with few urban portrayals, leading to misrepresentations and stereotypes of these cultures as set back in time (Gultekin and May 2018; Short, 2018). For example, in the US, middle grade novels about China are primarily set during the time of the Cultural Revolution and picturebooks about Mexico are overwhelmingly set in rural areas of small villages and ranches.

These issues of accuracy, authenticity, and representation in books can establish stereotypes, instead of opening minds to the rich diversity of our world. When readers see a person or event as one thing over and over, they come to see that person or event in a limited way. Adichie (2009) argues that people can be reduced to a single commonplace narrative of who we think they are, rather than who they actually are, what she calls “the dangers of a single story.” Bishop (1990) points out that readers need books as windows into the lives of others as well as mirrors for their own lives. Readers, however, need windows into many rooms to understand a culture and so bookmakers need to be aware of which windows are already plentiful and easy to see through, which are boarded up and difficult to locate, and which windows cannot be found in books. Another complexity is that some windows into a culture need curtains because certain images are not appropriate for specific audiences, such as images of the prophet Mohammed by Sunni Muslims (Gultekin and May 2019) or Hopi stories meant to be told only in a particular season and that should not be written down in books accessible at any time.

The social responsibility of influencers

Adults act as intermediaries throughout the process of creating books and determining the roles of books in children's reading lives. Adults influence whether children have access to books, which books are recommended, and what types of interactions with books are encouraged. Although adults continuously act for children, influencing their reading, their role with global literature is heightened since these books often require more support due to children's unfamiliarity with certain cultures. Adult influencers thus have a social responsibility to readers, whether their role is a bookseller, reviewer, blogger, librarian, classroom teacher, university scholar, or researcher.

Reviewers and bloggers in each country provide perspectives and critiques of new global books that influence purchasing decisions and the perspectives of adults who use these books with children or adults. In addition, the participation of educators and scholars on book award committees leads to public recognition of books named as award winners or on award lists that greatly increase the likelihood of that book's success in being made available to children. Other adult influencers include researchers, university instructors, and educators. These influencers can counteract negative aspects of trends in global markets through affecting how consumers spend their money on books.

Researchers of global children's literature

Researchers both reflect and influence which books receive attention and the perspectives of scholars and educators about those books. Research related to global literature has been cross-disciplinary and the location of researchers within their fields of scholarship influences their research questions and methodology. The three major fields are education, humanities/literary studies, and library science, but researchers from many other fields are involved including childhood studies, history, linguistics, and cognitive psychology (Wolf et al., 2010).

The primary emphasis of research on global literature remains textual analysis of words, images, and design to look closely at "What is this text about?" by interpreting content through a theoretical lens and close analysis of narrative patterns, stylistic devices, and design features. The central goal of this research is examining literature as representations of human experience and asking questions about the significance of what is happening in a book (Stephens, 2015). Much of this research occurs through a critical theory lens, such as critical race theory, postcolonialism, childism, or eco-criticism, used to frame the research process, questions, analysis, and findings. The researcher uses a deep understanding of a specific critical theory to draw key tenets from that theory as a lens for the analysis of a book or set of books in order to think with theory throughout the research process (Jackson and Mazzei, 2012). The methodology of critical reading in literary criticism (Bradford, 2016) and critical content analysis in education (Johnson et al., 2016, 2019) dominates current research, although there are some studies that instead use content analysis to identify themes or characteristics across a set of books, such as the ethnicity or gender identities of characters in award-winning books.

García-González and Deszcz-Tryhubczak (2020) argue that posthumanist and new materialist philosophies provide paradigms for new openings in research that counter the emphasis on language and discourse that dominates the field. New materialism rejects binary categories to emphasize the relational and the complex and highlight processes of becoming rather than states of being as a frame for research with global literature, instead of about this literature. They encourage scholars to "let go of the researcher's gaze and acknowledge our exposure to the relationalities of texts, readers, affects, and matter" (p. 57) and provide examples of research studies from this philosophical perspective.

The International Research Society for Children's Literature (IRSCL) is an international scholarly organization that promotes literary research in children's literature through a biennial World Congress and a scholarly journal, *International Research in Children's Literature*. The organization has members from over 40 countries and seeks to facilitate research across countries as well as the scholarly exchange of research and discussion of professional and theoretical issues. Other journals that publish research in English on the textual analyses of global children's and young adult literature include *Bookbird: An International Journal of Children's Literature* (IBBY), *Papers: Explorations into Children's Literature* (Australia), *Jeunesse: Young People, Texts, Cultures* (Canada); *Barnboken: Journal of Children's Literature Research* (Sweden), *Filoteknos* (Poland), *Libri & Liberi: Journal of Research on Children's Literature and Culture* (Croatia), *The Lion and the Unicorn* (US), *Children's Literature in Education* (US), and *Research on Diversity in Youth Literature* (US).

Although there are some studies of the interactions of readers around global children's literature, this area of research is limited and needs continued expansion, including children's interactions with translated books. National journals in the field of literacy and literature, such as the *Journal of Children's Literature*, or journals with an interdisciplinary focus, such as *Bookbird*, are the most likely places to find this research.

University instructors of global children's literature

Children's literature courses are located in different disciplines within a university, reaching out to students who have differing relationships and interests in these books. Sometimes children's literature is taught within education for students who are teachers or preparing to be teachers. Their interest is not only in becoming familiar with the books and ways of connecting readers with books, but also in developing a critical lens for their reading and that of their students. Courses located within the humanities tend to take a literary criticism perspective with a focus on deep textual analysis of a small group of books, often those considered classics, as separate from children as readers, and are often taught as English literature courses. Courses located within library science tend to take more of a survey approach to acquaint students with a wide range of books to recommend to readers. Courses in the

humanities and library science may be taught for majors in those areas or as service courses for teacher education majors. The perspectives and knowledge that students develop around global literature thus differ by the location of the course.

At the undergraduate level, global literature is frequently one topic or unit within a broader survey course on children's literature, no matter where that course is located, often leading to a surface knowledge of the books and bookmakers and limited understandings of how the cultural lenses of readers influence their understandings. In teacher education courses, global literature is often taught as a tool for teaching something else, such as literacy, social studies, mathematics, or science, overlooking literature as a value in and of itself. One strong emerging area of focus is the use of global children's literature in the teaching of English as another language (Bland, 2013) as well as in the teaching of world languages. In contrast, in-depth courses of global literature support much deeper understandings of the complex issues of culture surrounding books and readers.

Educators in schools and libraries

The social responsibilities of readers and bookmakers inform the work of educators, particularly in their awareness of the potentials and dangers of specific books in relation to the cultural locations of readers and bookmakers. Teachers and librarians, therefore, have responsibility for gaining deep knowledge of the cultural perspectives that students bring to their reading so they can anticipate the ways in which students may struggle or connect with a book set within a cultural community. This knowledge supports educators in creating learning contexts that encourage dialogic inquiry around literature. These contexts ensure that readers never read a book alone, but instead read alongside other books and engage in dialogue with each other to develop a critical lens to bring to reading events.

One social responsibility of educators is to offer instructional engagements that provide a context in which readers can position themselves and the book, learning to question *why* they have a particular response. These engagements include ones that help readers become aware of, and reflect on, their cultural locations, such as creating a heart map of the values they hold most closely in their hearts. This awareness is further supported by engagements that challenge readers to examine the match or mismatch between their cultural location and that of bookmakers, especially the author, illustrator, and translator, as well as to research where a book was originally published and by whom. By engaging readers in strategies for researching the cultural locations of bookmakers and providing frameworks of questions readers can ask themselves related to accuracy, authenticity, and representation, educators encourage readers to develop a critical lens from which to approach their reading of any book.

When students read a global book alone, their tendency is to view that book as representative of an entire global culture, unless they have life experiences within that culture. If instead a book is embedded in a unit of inquiry and read alongside other books and materials, they use intertextuality to consider how texts play off each other and create a context within which they can make connections for constructing understandings of a book. Instructional engagements with paired books and text sets help readers view a book within a broader representational context for a cultural community and can create a sense of connection and difference across global cultures. Pairing a book that is culturally distant with a book that is culturally near may build conceptual bridges. Readers can also be encouraged to research additional information in the library or on the internet to place a book within a broader context.

Another aspect of social responsibility for educators is creating contexts that encourage intercultural dialogue between readers. Dialogue involves readers walking alongside and talking with, not at, others to wrestle with issues instead of walking on the surface of those issues (Freire, 1970). Dialogue combines critique and inquiry, as readers pose questions, challenge the status quo, and collaboratively consider multiple viewpoints on an issue. Freire (1970) argues that dialogue has the most potential for societal transformation by combining love, humility, faith in others, hope, and critical thinking (p. 72). Dialogue strategies, such as webbing issues or selecting a memorable quote, encourage readers to reflect deeply on a book in order to find issues they consider significant, instead of being guided by what a teacher has already identified.

Resources on global literature

For those who want to become more involved in the field of global children's literature, a wide range of resources exist, including international organizations, bookfairs, awards, and scholarly overview projects. In addition, there are many national organizations and awards, but our focus is on those that are international in scope.

Although most countries have libraries that house their country's literature, only a handful strive to internationalize their holdings to showcase the world's literature for children. The most significant is the International Youth Library, which offers international opportunities within and beyond their location in Munich, such as traveling exhibitions that make their way around the world. "Hello, Dear Enemy" has been traveling the world for more than two decades, with updated iterations of new books. The IYL Fellowship program brings together researchers from around the world who rely on the world's largest collection of international books to conduct research and weave together international understandings. The White Ravens Recommendation List features notable books from around the world that are singular in distinction and deserve international awareness.

The International Board on Books for Young People is an organization that brings together those involved in the creation and dissemination of literature while supporting activities for literacy promotion. Eighty-one national sections share common goals of building relationships to promote literature and literacy, encouraging international understanding, and sharing ideas and resources in regions of high need, along with administering awards that feature the highest recognition for achievement. The Hans Christian Andersen Award highlights a nominee list of authors and illustrators regarded by different countries as their best; the winners feature

those recognized as the best among the best. The IBBY Honour Books for writing, illustration, and translation are nominated by member nations who recognize their own best literature; likewise, each IBBY award shows selection and support by their own countries. What counts as “best” is an elusive definition made difficult across cultural perceptions but left to the decision of each country and then judged by an international jury; this two-tiered system balances the national with the international. Every two years, a World Congress is held in a host country allowing members from around the world to come together to share book creation and research, celebrate literature and literacy initiatives, and engage in discussions to further collaborative efforts of mutual goals.

First held in 1964 with European exhibitors, the Bologna Book Fair has blossomed to participants from all corners of the world reaching large international audiences of visitors and publishers. The primary goal is for the sale and purchase of foreign rights and the introduction of authors, illustrators, and translators across international borders, along with increasing opportunities for digital media and communication. The physical exposure to books and their creators and opportunities to engage in discussion make international book fairs a unique place to see what the world is publishing. However, this fair is dedicated to selling rights, and what sells rules decisions about displays in booths. For example, Disney, Nickelodeon, and other media-driven books are prominently displayed whereas more literary works are seen by appointment. Often, the emphasis is on the rights to generic books with artwork that transcends cultural specifics and minimal text that is easy to translate. Sell-ability dominates over unique distinctions. Although the Bologna Bookfair is the oldest and largest international fair, other large book fairs occur around the world, such as in Shanghai, Moscow, and Guadalajara.

Book awards are another resource for locating global literature, particularly works of notable quality. The IBBY Hans Christian Andersen Award is the oldest lifetime achievement award. What makes it distinct is that nominees are named by individual member nations, and an international jury selects from among those nationally nominated authors and illustrators. In a sense, the nominees are themselves winners of the nomination from their own countries. The Astrid Lindgren Memorial Award is presented by Sweden and is considered a generous monetary gift in honor and memory of Sweden’s most famous writer, Astrid Lindgren. Nominations of authors, illustrators, or literacy organizations are made by officially ordained nominating bodies from around the world, and, because they are not limited to their own countries, some are nominated simultaneously multiple times.

The most well-established awards from book fairs are the Bologna Book Fair Illustration Exhibition that goes on tour for two years and the Bologna Ragazzi Award for best book design. The China Shanghai International Children’s Book Fair has the Golden Kite illustration competition and exhibition as well as the Chen Bochui Literary Award. Some countries value children’s literature so much that they host international exhibitions with competitions to encourage the highest quality literature for children. Biennale of Illustration Bratislava (BIB) has the support of the Slovak Ministry of Culture, and the cooperation of embassies from around the world. What is striking to note is that submissions from English language countries are few in number, providing a more global scope.

Finally, the perspectives of scholars, especially literary critics, are available in handbook and encyclopedia projects. *The Routledge Companion of International Children’s Literature* (Stephens et al., 2018) is organized around broad themes to highlight scholarship from Africa, Asia, the Caribbean, and Latin America and to reconsider Euro-centric scholarship often held as the standard of the profession. *The Edinburgh Companion to Children’s Literature* (Beauvais and Nikolajeva, 2017) focuses on current directions in children’s literature scholarship, contemporary trends in the literature, and future unmapped territories. In contrast, *Handbook of Research on Children’s and Young Adult Literature* (Wolf et al., 2010) is a collection of interdisciplinary essays on literary reading, texts and context across the fields of literacy education, English literature, and library sciences, written by scholars and children’s authors.

Conclusion

In writing this state of the art in global children’s literature, we are aware of the caution offered by Fraser’s analysis of the state of the field of international literature in 1981. Fraser critiqued the work of scholars, noting that even though they avidly advocated for this literature and its merits, their work was based in the belief that books for readers in English language countries originate in the English language. Fraser’s own analysis, however, failed to recognize that her examples were European, indicating her assumption that international was synonymous with European literature. We are aware that, in the future, our own analysis will be shown to be limited by our current world views, experiences and conversations in the field.

For us, global literature holds significance because of how those books play out in the lives of children, both in helping them critically examine their cultural assumptions and values and in opening themselves to alternate values and perspectives. Global books influence how children view the world and its complexities as they figure out their places and functions in the world they have come to know. Children’s engagements with literature hold the possibility of transforming their world views through understanding their current lives and imagining beyond themselves. If their reading only mirrors their own views and cultural communities, children have difficulty envisioning alternative ways of thinking and being, essential to the societal transformation that Lepman (2002) dreamed about so many years ago.

References

- Beauvais, C., Nikolajeva, M. (Eds.), 2017. *The Edinburgh Companion to Children’s Literature*. Edinburgh University Press, Edinburgh.
- Bishop, R.S., 1990. Mirrors, windows, and sliding glass doors. *Perspectives: Choosing and using books for the classroom* 6 (3), ix–xi.

- Bland, J., 2013. *Children's Literature and Learner Empowerment: Children and Teenagers in English Language Education*. Bloomsbury, London.
- Bradford, C., 2011. Children's literature in a global age: transnational and local identities. *Barnelitterært Forskningstidsskrift* 2 (1), 20–34.
- Bradford, C., 2016. The critical reading of children's texts. In: Johnson, H., Mathis, J., Short, K.G. (Eds.), *Critical Content Analysis of Children's and Young Adult Literature*. Routledge, New York, pp. 16–27.
- Brooks, W., Browne, S., 2012. Towards a culturally situated reader response theory. *Child Lit. Educ.* 43, 74–85.
- Brown, B., 2018. *Dare to Lead*. Random House, New York, NY.
- Carle, E., 1969. *The Very Hungry Caterpillar*. Penguin, New York, NY.
- Chandler, D., 2016. *Strategic Corporate Social Responsibility*. Sage, Thousand Oaks, CA.
- Choo, S., 2014. Cultivating a hospitable imagination. *Curric. Inq.* 44 (1), 68–89.
- Cuff, B., Brown, S., Taylor, L., Howatt, D., 2016. Empathy: a review of the concept. *Emotion Review* 8 (2), 144–153.
- Duncum, P., 2002. Visual culture art education. *Int. J. Art Des. Educ.* 21 (1), 15–23.
- Eco, U., 2001. *Experiences in Translation*. University of Toronto Press, Toronto, ON.
- Fox, D., Short, K., 2003. *Stories Matter: The Complexity of Cultural Authenticity in Children's Literature*. National Council of Teachers of English, Urbana, IL.
- Fraser, J.H., 1981. Internationalism and the children's literature community in the United States. *Libr. Q.* 51 (1), 54–67.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York, NY.
- García-González, M., Deszcz-Tryhubczak, J., 2020. New materialist openings in children's literature studies. *Int. Res. Child. Lit.* 13 (1), 45–60.
- Geertz, C., 1973. *The Interpretation of Cultures*. Basic Books, New York, NY.
- González, N., 2005. Beyond culture: the hybridity of funds of knowledge. In: González, N., Moll, L., Amanti, C. (Eds.), *Funds of Knowledge*. Routledge, New York, NY, pp. 29–46.
- Grossman, E., 2010. *Why Translation Matters*. Yale University, New Haven, CT.
- Gultekin, M., May, L., 2019. Children's literature as fun-house mirrors, blind spots, and curtains. *Read. Teach.* 73 (5), 627–635.
- Hade, D., Edmonson, J., 2003. Children's book publishing in neoliberal times. *Lang. Arts* 81 (2), 135–144.
- Hansen, D., 2011. *The Teacher and the World: A Study of Cosmopolitanism as Education*. Routledge, London.
- Haviland, V., 1961. The flow of children's books from country to country. *Top News* 18 (2), 7–14.
- Heise, U., 2008. *Sense of Place and Sense of Planet*. Oxford, New York, NY.
- Jackson, A., Mazzei, L., 2012. *Thinking with Theory in Qualitative Research*. Routledge, New York, NY.
- Johnson, H., 2016. The social responsibility of the reader. In: Johnson, H., Mathis, J., Short, K.G. (Eds.), *Teaching Globally: Reading the World through Literature*. Stenhouse, Portsmouth, NH, pp. 25–42.
- Johnson, H., Mathis, J., Short, K.G., 2016. *Critical Content Analysis of Children's and Adolescent Literature*. Routledge, New York, NY.
- Johnson, H., Mathis, J., Short, K.G., 2019. *Critical Content Analysis of Visual Images in Literature for Young People*. Routledge, New York, NY.
- Joosen, V., 2020. Just listen? Silence, silencing, and voice in the aesthetics, reception, and study of children's literature. *Int. Res. Child. Lit.* 13 (1), 111–125.
- Kästner, E., 1949. *The Conference of Animals*. McKay, London.
- Kerchy, A., 2018. Translation and transmedia in children's literature. *Bookbird* 56 (1), 4–9.
- Lathey, G. (Ed.), 2006. *The Translation of Children's Literature: A Reader*. Multilingual Matters, Clevedon.
- Lepman, J., 2002. *A Bridge of Children's Books: The Inspiring Autobiography of a Remarkable Woman*. Trans. by E. McCormick. The O'Brien Press, Dublin.
- Maruki, T., 1980. *Hiroshima No Pika*. Lothrop, New York, NY.
- Miyake, O., 2002. A study on the cultural background of picture books: when Japanese picture books are translated into English. *Bulletin of the Association for Studies of Picture Books* 4, 37–45.
- Mo, W., Shen, M., 2003. Accuracy is not enough: the role of cultural values in the authenticity of picture books. In: Fox, D., Short, K.G. (Eds.), *Stories Matter: The Complexity of Cultural Authenticity in Children's Literature*. National Council of Teachers of English, Urbana, IL, pp. 198–212.
- O'Sullivan, E., 2011. Comparative children's literature. *PMLA* 126 (1), 189–196.
- Oittinen, R., 2008. Audiences and influences: multisensory translations of picturebooks. In: González Davies, M., Oittinen, R. (Eds.), *Whose Story? Translating the Verbal and the Visual in Literature for Young Readers*. Cambridge Scholars, Newcastle, pp. 3–16.
- Pollock, D.C., Van Reken, R., 2001. *Third Culture Kids*. Nicholas Brealey, Boston, MA.
- Rosen, H., 1986. *Stories and Meanings*. NATE, London.
- Rosenblatt, R., 1938. *Literature as Exploration*. Modern Language Association, Chicago, IL.
- Rosenblatt, R., 1994. The transactional theory of reading and writing. In: Ruddell, R.B., Ruddell, M.R., Singer, H. (Eds.), *Theoretical Models and Processes of Reading*. International Reading Association, Newark, DE, pp. 1037–1092.
- Settingington, K., 2016. How Canada publishes so many diverse children's literature. *Sch. Libr. J.*
- Short, K.G., 2012. Story as world making. *Lang. Arts* 90 (1), 9–17.
- Short, K.G., 2018. What's trending in children's literature and why it matters. *Lang. Arts* 95 (5), 287–298.
- Short, K.G., 2019. The dangers of reading globally. *Bookbird* 57 (2), 1–11.
- Short, K.G., Day, D., Schroeder, J. (Eds.), 2016. *Teaching Globally: Reading the World through Literature*. Stenhouse.
- Stephens, J., 2015. Editorial: critical content analysis and literary criticism. *Int. Res. Child. Lit.* 8 (1), v–viii.
- Stephens, J., Belmiro, C., Curry, A., Lifang, L., Motawy, Y., 2018. *The Routledge Companion to International Children's Literature*. Routledge, London.
- Takeuchi, M., 2014. Ishii Momoko no honyaku wa naze kodomo o hikitsukeruno ka/Why do Momoko Ishii's translations attract children? Minerva Publishing, Tokyo.
- Vassilousi, V., 2019. International and local relief organizations and the promotion of children's and young adult refugee narratives. *Bookbird* 57 (2), 35–49.
- Wahlström, N., 2014. Toward a conceptual framework for understanding cosmopolitanism on the ground. *Curric. Inq.* 44 (1), 113–132.
- Wolf, S., Coats, K., Enciso, P., Jenkins, C. (Eds.), 2010. *Handbook of Research on Children's and Young Adult Literature*. Routledge, New York, NY.
- Yokota, J., Teale, W., 2017. Striving for international understanding through literature. *Read. Teach.* 70 (5), 629–633.
- Yokota, J., Jorenby, M., Adachi, S., Teale, W., 2007, August. One Japanese original text, multiple English versions: critical discourse analysis of *Hiroshima No Pika* and differences in impact on reader response. In: Presented at the 18th Biennial Congress of the International Research Society for Children's Literature, Kyoto, Japan.
- Yoo-Lee, E., Fowler, L., Adkins, D., Kim, K., Davis, H., 2014. Evaluating cultural authenticity in multicultural picture books. *Libr. Q.* 84 (3), 324–347.

Children's literature in Latin America today: a historical perspective

Lilian Alvarez-Arellano^a, Carlos Rubio-Torres^b, Marcelo Emilio Bianchi-Bustos^c, and Donovan Herrera-Santillán^d, ^a Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico; ^b Universidad de Costa Rica, San José, San Pedro, Costa Rica; ^c Academia de Literatura Infantil, Buenos Aires, Argentina; and ^d Universidad Nacional Autónoma de México, Mexico City, Mexico

© 2023 Elsevier Ltd. All rights reserved.

Overview	644
When and how did it start?	644
Who are some of its main authors and what are its historical and current trends?	647
What is its relationship to other literatures of the world?	653
Discussion: the concept of Latin America in relation to children's literature	655
Latin America's social configuration and literary heritage	655
African legacy	656
Alternative terms to name the region, and their consequences for children's literature	656
How many "roots"?	657
Other denominations of the area, and their consequences for children's literature	658
Conclusions	659
Further Reading	659
Cited works of Latin American children's and youths literature	660
Brief selection of relevant childrens literature studies	661

Overview

When and how did it start?

In this region, increasing social awareness about the importance of books for young readers, the growth of the book publishing industry, as well as the intervention of governments through reading policies, massive book purchases for public libraries and schools, as well as widespread primary instruction set the basis for children's literature to develop with force toward the end of the 20th century.

If we consider industrially produced and commercially distributed texts for children in European languages and traditions, we can say children's literature in Latin America started during the early 16th century with a few titles imported from Europe and even fewer written and printed in Spanish and Portuguese America. The area witnessed sustained growth of local production of children's literature since the early 19th century. Nowadays it is the sector of the publishing business which has most potential growth.

But if we are to consider those orally produced and transmitted texts in the indigenous languages and traditions of the region, however, we must acknowledge that children's literature had a much earlier inception. If we take into account those oral or written texts which originally belonged to European, African or Asian literatures and lore, but were adopted and developed locally, a rich heritage children's literature in Latin America is unveiled. The region is a living repository of multicultural and multilingual corpora that interact in several ways. The texts of this complex heritage have social uses beyond reading, for children's literature is performed in play, community rituals, feasts and celebrations in the forms of dance, prayers and incantations, singing, drama and storytelling.

There was, since the beginning of colonial rule during the 16th century, a local production geared to the instruction of aborigines in the language and religion of their conquerors, often in bilingual editions. Among the first printed books in Spanish America were catechisms in Nahuatl, the language of the Aztec, and in Quechua and Aymara, languages of Perú, as early as 1539 and 1583, respectively. The first textbook to be used in America among young readers presented a selection of Aesop's *Fables* translated into Latin and Nahuatl, in the first higher education school for young natives in the Continent, the Colegio de la Santa Cruz in Santiago de Tlatelolco, established in 1536.

Noteworthy in this early production are educational efforts to produce pictographic prayer books in which two very different communication systems, world views and ways of representation are combined (see Figs. 1 and 2). Instead of words, the reader finds drawings set in the lines of a page to be read top-down and from left to right, as in a European book. These pictographic texts communicated Christian concepts through images which the aborigines could read metaphorically. The colonial policy of imposing one way of reading—virtually nullifying interest in the semantics of these visual languages and literally destroying monuments and codices—meant that Latin American children lost cognitive access to a vast repertoire of indigenous texts that are extant.

During the 1st centuries of colonial rule, reading books for beginners, prayer books, fables and adaptations of Greek and Latin texts were mostly imported from the Metropolises. The prohibition of novels in Spanish dominions; the unequal introduction of the printing press and of universities in the region and restrictions imposed to printing and the free expression of ideas explain in part the early limited development of imaginative, pertinent and locally situated literature for children and the tardiness of the



Fig. 1 Peter of Ghent. *Pictographic catechism*, ca. 1525-1528. Hand painted. Notice long-beaked bird to signify the Holy Spirit. Identify European soldier and friar depicted.

formation of what could be called a literate society. Neither Spain nor Portugal had a strong tradition of books for children, so what circulated in their New World colonies were mostly adaptations and translations of French works such as Reyre's *Lami des enfants*, Fenelon's *The Adventures of Telemachus*, a Christian rendering of Odysseus' son's quest or, more rarely, the works of some of the imaginative authors of the French Enlightenment (see Fig. 3).

The first Latin American novel, José Joaquín Fernández de Lizardi's *El Periquillo Sarniento* (1816) dealt with the problems of being born and educated in the decadent final years of colonial rule in New Spain (see Fig. 4). Among highlighted problems related to rearing children that the novel analyzes is the scarcity of good reading books for them (see Fig. 4).

Nevertheless, not all the flow of new ideas depended upon books. Among other strategies, staged plays were used in the Spanish and Portuguese dominions to spread Christianity. Through mystery and Nativity plays, and later "*Comedias de magia*" (magic plays), elements of medieval theater and later a selected corpus of Gilded Age and Baroque characters and plots entered popular spectacles and culture (Carranza Vera and Silva Martínez, 2017, 10–13). In Brazil Jesuit José de Anchieta, the "father of Brazilian literature," catechized Indians through plays he wrote to that purpose. The scenography resources with which such works were staged opened the way for eerie ambiances and plots, imagination, magic, miracles and phantasy. This had effects on popular children's culture. It also paved the way for puppet theater companies for children in different countries of Latin America along the 19th and 20th centuries. *Pedro y su teatrino maravilloso* pays homage to this activity (Rubio, 1993).

Local production of children's books during colonial times was forged in European molds and had religious, moral and educational ends to which an amiable style, typography and illustration contributed gradually as technology advanced. Children's literature is a phenomenon that in Latin America started to flourish in the early 19th century, as part of the cultural revival called for and spurred by Independence movements, the first of which was that of Haiti in 1794 and last that of Cuba in 1898.

Newspapers and books during the 19th century were seen as vehicles to instill in children civic ideals such as freedom and equality as well as love of country and of learning, by providing readers with accounts of natural sites, cities of splendor, memorable historical facts and the admirable lives of heroes of each nation and of the region. Atahualpa, Cuauhtémoc, Andrés Bello, Simón Bolívar, José María Morelos and Benito Juárez, to name a few, were paid homage in children's literature. Contemporarily, the ever-present cult of heroes in children's literature incorporates not only individuals or groups considered political heroes as Francisco

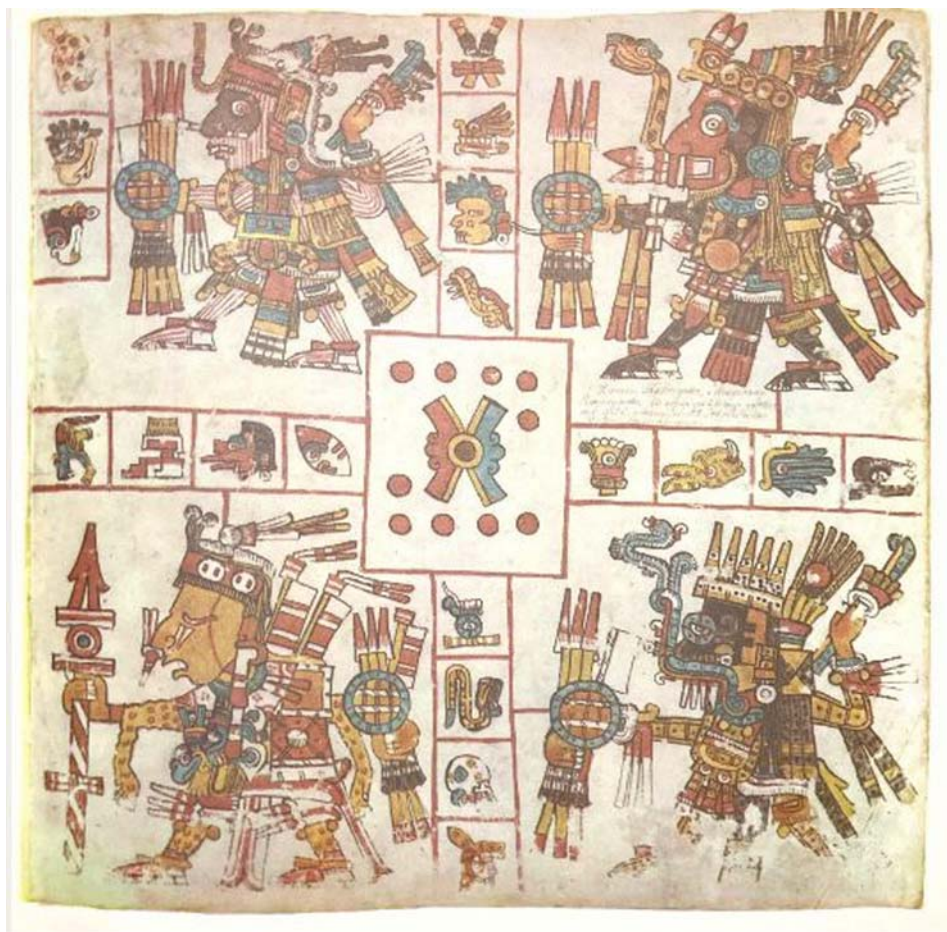


Fig. 2 Anonymous. *Borgia Codex*. XVth century. Page to be read clock-wise.

Miranda or César Chavez, but also those of other walks of life such as sports and music, as well as writers, dancers and other artists. *Tito Puente, rey del mambo* (Brown, 2013) is a case in point.

Printed children's literature appeared in different countries with varying paces and vigor. Most probably the first periodical publication for children in the region was *El Correo de los Niños* (1813), edited by Wenceslao Sánchez de la Barquera in New Spain; *El Periquito* (1868), was one of the first Cuban children's periodicals (Guadarrama Olivera, 2000); *El Atomo* (1888) was the first children's periodical in Guatemala (Lozano Alvarado, 114–116). Mostly short lived, initiatives for children during the 19th century suffered from economic and political instability that thwarted cultural and economic initiatives.

Throughout the 19th century periodicals and books for children in Latin America followed French, Portuguese, Spanish and sometimes English or North American models, gradually granting more space to local news and themes, authors, and illustrations, contributing in this way to the creativity, originality and relevance of children's literature. The local industry geared to produce books for children was slowly built in Latin America, where European entrepreneurs played a seminal role as printers, editors and distributors. In time, in some countries, together with imported merchandise, the basis to cater to young readers within a local publishing industry nourished by national authors and local printing shops was established. For example, several businesses coexisted in Brazil during the second half of the 19th century: the locally owned and operated *Livreria do Povo*, of editor Pedro Quarasma, the German owned Laemmert, the French Garnier—which operated from Paris in different points of Latin America—and the business of Francisco Alves, a Portuguese publisher (Fioravanti, 2013). The first printing shop and bookstore for children was José Rosas' *Imprenta y Librería de los Niños* (1873–75) in Mexico City (see Fig. 5).

The longest-lived initiatives in children's periodicals in the area are *El Peneca* (1908–1960), printed in Chile and read in several countries of South America, and the monthly *Billiken*, founded in Argentina by Constancio C. Vigil, which has been published since 1919 to our day, an unparalleled feat in the region. Starting in 1950 this monthly magazine was exported to Uruguay, Paraguay, Venezuela, Puerto Rico and Bolivia. It evolved with the times, dodging and sometimes supporting or turning a blind eye to censorship from authoritarian regimes, struggling for a share of the market in competition with North American comic books translated and printed in Mexico by Editorial Novaro since post-war time onwards (Díaz Rönnner, 2000, p. 523), when the United States endorsed the development of cultural initiatives in the publishing and film industries in that country as a buffer to communism.

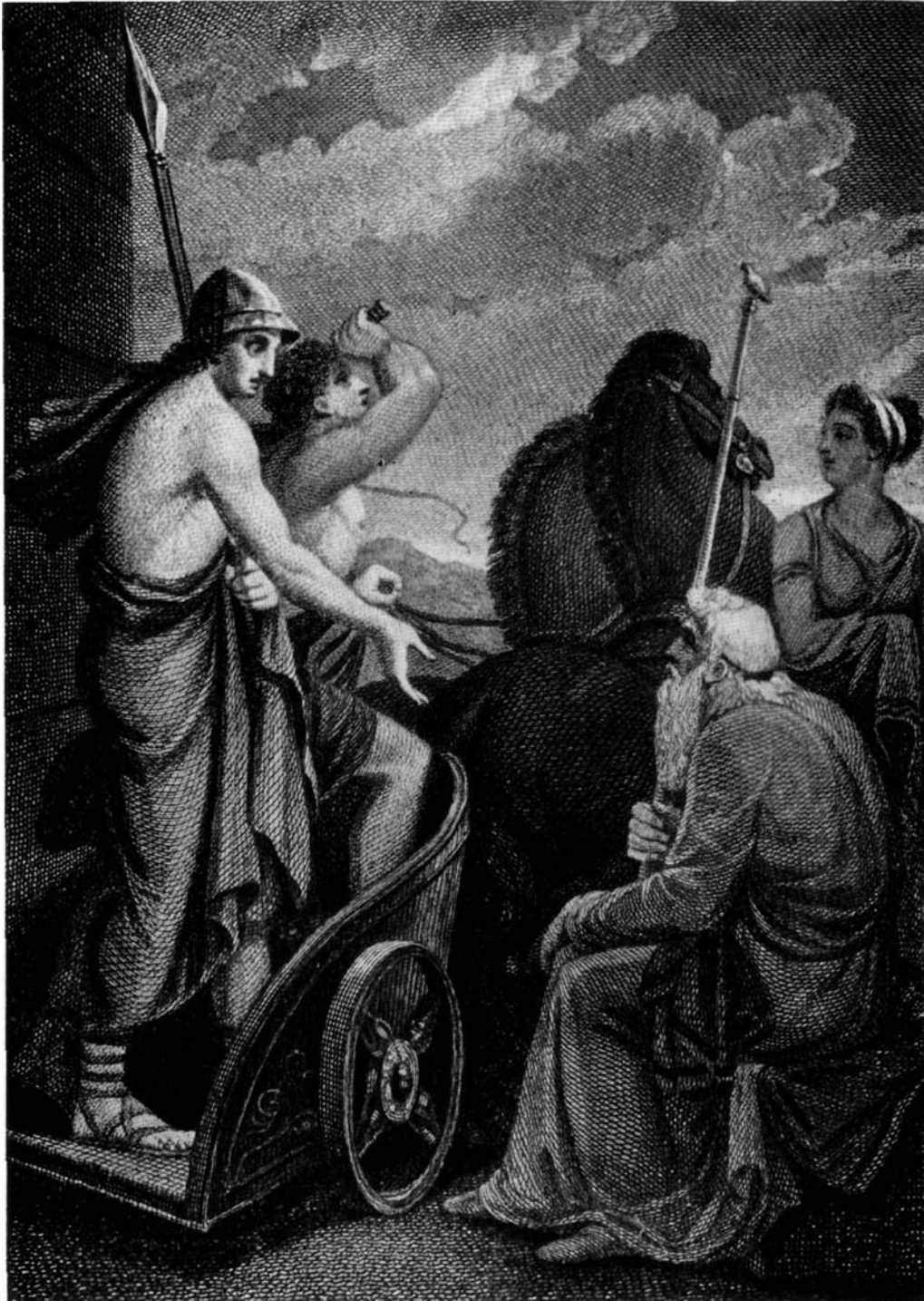


Fig. 3 François Salignac de la Mothe Fenélon. *Les Aventures de Télémaque, fils, d'Ulysse*, ca.1699. Metal engraving.

Who are some of its main authors and what are its historical and current trends?

Writers, editors and illustrators of children's literature in Latin America strive to find and offer their own voice and esthetic values in a multicultural and plurilingual literary tradition woven with many threads, in a society with numerous needs. Increasingly public and private cultural, educational and welfare institutions have participated in projects geared to encourage the creation and promotion of children's literature. There is today an increased specialization in professions related to publishing in this field.

Although there are in fact several precursors of children's literature as it is understood and practiced in each country today, certain authors captured the public's imagination as the founders of each country's production. Such is the case of Rafael Pombo



Fig. 4 José Joaquín Fernández de Lizardi. *El Periquillo Sarniento*, Mexico, Imprenta de Galván, 1816. Metal engraving.

and his fables in Colombia; Alberto Figueredo Pimentel (1869–1914), Alexinha de Malghães Pinto (1870–1921) and [José Bento Renato] Monteiro Lobato (1882–1946) in Brazil; José Rosas Moreno (1838–1883) in México; José Martí (1853–1895) in Cuba; Horacio Quiroga (1878–1937), Juana de Ibarborou (1892–1979) and José María Firpo in Uruguay (1916–1979); José Sebastián Tallón (1904–1954) and María Elena Walsh (1930–2011) in Argentina; Carmen Lyra (1887–1949), Luis Carlos Sáenz (1889–1983) and Adela Ferreto in Costa Rica; Concepción Leyes de Chávez (1891–1985) and Josefina Plá (1903–1999) in Paraguay; Carlota Carvallo de Nunes (1909–1980) in Perú; Salomé Ureña (1850–1897) in the Dominican Republic; Elvira Santa Cruz Ossa (1886–1960) in Chile; Manuel J. Calle (1866–1918) in Ecuador; Adela Zamudio (1854–1928) and Nataniel Aguirre (1843–1888) in Bolivia; Daniel Armas (1897–1983) and Marilena López (1902–1980) in Guatemala; in Panamá Luisita Aguilera Patiño (1914 to our day), Hersilia Ramos de Argote (1910–1991), Gonzalo Brenes (1907–2003) and Rogelio Sinán (1902–1994); the contemporaries Salarrué (1899–1975) and Claudia Lars (1899–1974) in El Salvador; in Puerto Rico, María Cadilla (1884–1951); Teresa de la Parra (1889–1936) and Rafael Rivero Oramas (1904–1992) in Venezuela; José Trinidad Reyes (1797–1855) in Honduras.

The phenomenon of authors devoted to writing or illustrating for children or youths is, with exceptions, more a 21st century phenomenon. National, regional and international prizes have contributed to form a sizable group of children's literature authors and illustrators. In spite of insufficient translation and distribution of the works of Latin American authors for children both within

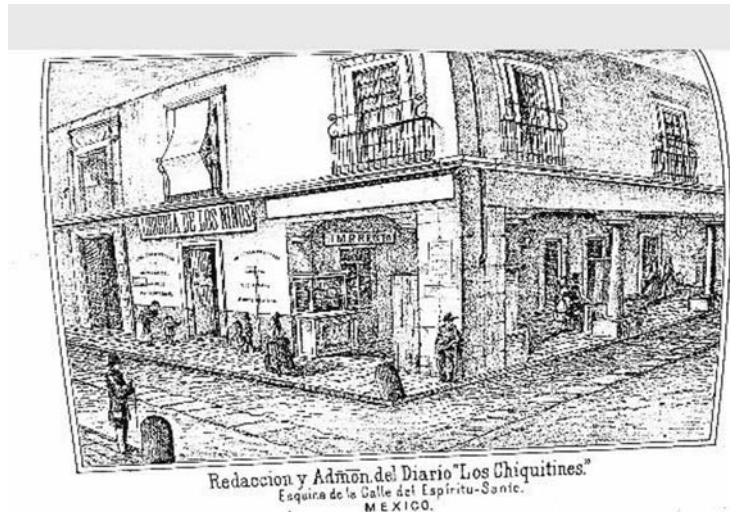


Fig. 5 Children's Printshop and Bookstore (1873), the first specialized enterprise in the region. Metal engraving. Source: José Rosas Moreno, *Los Chiquitines. Diario Infantil, Joco-serio, con Estampas y Caricaturas*. México: Imprenta y Librería de los Niños, 1874.

Latin America and in the world, contemporary writers and illustrators such as Maria Teresa Andruetto, Ana María Machado, Isol Ziraldo, José Mauro de Vasconcelos, Lydia Bojunga Nunes, Carmen Leñero, Adolfo Córdoba, Antonio Orlando Rodríguez, Vicky Ramos, Jorge Luján, Micaela Cherif, Martha Riva Palacio, Monica Doppert, Roger Mello and Fabrizio Van den Broek, to name only a few, have won international prizes that grant them recognition as artists in the region and beyond.

We find Nobel prize winners of Latin America who devoted part of their endeavors to young readers, as the *Rondas* of Gabriela Mistral, Pablo Neruda's monumental *Canto General* written for youths of the region, Rigoberta Menchú Tum's *El legado secreto* (2016), or Miguel Ángel Asturias' *Leyendas de Guatemala*. Canonical authors all over Latin America have written for children, as did Paraguayan Augusto Roa Bastos in *El pollito de fuego* (1974), Sergio Ramírez in Nicaragua with *La jirafa embarazada* (2013), Clarise Lispector in *Quase de verdade* (1978), or Jorge Amado's *O gato malhado e a andovinha Sinhá* (1948), to name a few. When the most famous authors have not engaged in children's literature, selections and sometimes adaptations of their works have been traditionally refunneled to children.

Children's literature and textbooks have been thought instrumental to achieve the social conditions adults strive for. Not uncommonly, authors have had religious, ecological, political or various instructional ends when writing for children. Trends in adults' affairs tend to shape the production of children's literature as is the case, for example, of sustainable growth nowadays (CERLALC, 2020). Thus, to our day texts for children have often had in view not only literary considerations but also or mainly educational as well as political contents and goals.

Different political ideologies and forms of didacticism have shaped children's literature in Latin America, site of invasions, revolutions, counter-revolutions and dictatorships, each with consequences for the development of the lives of children and the books that are available for them to read. Regimes have seen in literature a venue to attain unconditional support and in youngsters a vehicle for social change, renovation or control.

The will to control the ideological orientation of texts for children to read and learn is well entrenched in Latin American culture, at least that of Spanish origin. The earliest colonial policies for children's education prohibited instruction devoid of religious orientation: "*No instruir sin catequizar*" (no instruction without religious content). The first Council of 1555 in New Spain was earlier than the Council of Trent in what pertains to the control of printed texts, which specifically called for never teaching children to read without concurrently teaching them the basic principles of religion (Fernández de Zamora, 2008, p. 111).

Nationalism became one of the literary antidotes for colonialism and neo-colonialism. Thus, in 1901, for example, Heriberto Frías wrote about the Spanish, French and North American invasions of the 16th and 19th centuries in *Biblioteca del Niño Mexicano*, a series of 100 booklets, published by Maucci Hermanos, a Catalan publisher which had business in several countries of Latin America and Manila (see Fig. 6).

The 20th century revolutions of Mexico, Cuba and Nicaragua carried out impressive literacy campaigns and created series of children's books known as *Libros del Chapulín*, in Mexico, *Editorial Juvenil* and *Gente nueva* in Cuba, *Nueva Nicaragua* in that country. The 1944 Revolution in Guatemala produced the first modern children's magazine in that country: *Alegría*, which was censored by the counter-revolution (Maldonado de Masaya et al., 2011, pp. 12–13). The cases and modalities of censorship in Argentina, Chile, Colombia, Cuba, Guatemala, Mexico, Venezuela as well as Spain during the 20th century are well documented (Cerrillo and Sotomayor, 2016).

Other ideological orientations have acted upon children's literature in Latin America. Disney's overtake of children's imagination was decried in children's literature, for example, in *Ratón Pérez contra Ratón Miguelito* (see Fig. 7); it was dissected in *Para leer al Pato Donald* (Dorfman and Mattelart, 1971). Cultural diplomacy and policies of political propaganda from China and Russia reached



Fig. 6 Heriberto Frías. *Biblioteca del Niño Mexicano*, Mexico, Maucci Hermanos, 1910. Color lithography.

Latin America as systematic and useful arms of foreign policy (Ratcliff, pp. 53–79). Books for children, some by leading writers who exalted harmony in cooperative living, were in this way available in Spanish and Portuguese through literally millions of copies of beautifully illustrated Russian and Chinese books at very low or no cost from Progress Publications, based in Moscow, and Foreign Language Publishing House, in Beijing (Prashay, 2019).

Literature for children in Latin America tends to have this age group as protagonists, or recounts childhood memories, as well as the lives and works of individuals of historical relevance. It dwells on local myths or appropriations of international folklore. Poetry for children is practiced all over the region. Storybooks are favorites, and growingly novels, including graphic novels. Theater has been cultivated as a continuous undertone, with special predilection for puppets. Under the label “children’s lit” fall also, though less frequently, texts written by children. There are more works for young readers that are historical in perspective than those which offer a prospective or futuristic outlook. Outside picture books, few are abstract or conceptual. Texts for children tend to be more didactic than playful or poetic. Translations and adaptations of regional and international classics and contemporary works nourish what children read. There is a larger participation of women writers, editors and illustrators for children’s books than in others specialties, often in multiple capacities. Yuyi Morales, Beatriz Ferro, Isol, Ricardo Liniers and Christian Palacios are among those contemporary multifaceted professionals.

According to scholar Guerrero Guadarrama (2016), contemporary Latin American children’s literature critically re-reads old artistic molds subversively, irreverently, with humor and irony. It makes use of resources available in different genres and traditions to touch upon what traditional societies consider controversial topics when addressing children: divorce, adoption, mental health, sexuality, ageing, violence. It questions gendered behavior by playfully reversing or questioning roles. According to her, one of its most interesting current forms is the “unearthing of memory” in countries with a recent history of authoritarianism. Some have installed commissions of truth to reconstruct those periods, as is the case of Uruguay, Argentina, Chile, El Salvador, Guatemala, Perú, Paraguay, Colombia, Ecuador, Honduras and Brazil. Writers María Teresa Andruetto, Laura Devetach, Graciela Montes and María José Ferradas, among other winners of the S/M Iberoamerican Prize of Children’s and Youths’ Literature, have addressed part of their works to this reappraisal in a region where historical narratives account for a large part of all the writing geared to this age group.

Along the 20th century the notion that to achieve full literacy children need to find pleasure and some form of utility in reading texts that are written for them took force. Nowadays equally potent is the shared conviction that the texts children read should help them to perform freedom, gradual moral autonomy and a critical view—not obedience and conformity to unquestioned or unexamined social beliefs and values as they are traditionally embodied in historic or contemporary renditions of didactic literature.



Fig. 7 Miguel Ángel Bayardi. *Ratón Pérez contra Miguelito*. Illustrated by Dagoberto Dinorín, México, Secretaría de Educación Pública, 1945. Ratón Pérez is a traditional Spanish children's story character.

Given the role of children's literature in literacy, it is considered a key input in the development of psychomotor skills as well as in motivating students and enabling them to stay in school with good academic achievement. Furthermore, children's literature is expected to spark self-knowledge, critical thinking, social empathy and civic awareness, as well as to provide elements of personal, national and regional identity together with a sense of cultural and linguistic pride, in addition to providing models for healthy and sustainable lifestyles. This is a tall order for literature, although the availability of good texts together with skilled mediation can in some way contribute to individual actions and social policies that have these aims.

In Latin America the appreciation of the role of the “mediator” between child and book is evolving into a social and professional practice that involves librarians, teachers, parents and grandparents, older siblings, youths and others who are interested in contributing to those goals by creating reading communities in which participants hear stories, read out loud, share reading experiences and write their own texts so in that process they learn to select books, enjoy literature, express themselves orally and in writing, bond with fellow readers and have tools to deal with their everyday experiences. Mediation has a long history in the region, and its own heroes. One of them, Colombian rural librarian Luis Soriano, is the inspiration for Monica Brown and John Parra's picture book *Waiting for the biblioburro/Esperando el biblioburro* (2011). Argentina's writer María Cristina Ramos, winner of the Organization of Iberoamerican States and S/M Children's and Youths' Literary Award (2016) as well as runner up for the Hans Christian Anderson Award in 2020, is also a *mediadora*. Different institutions are engaged in mediation and cater to professionals and para-professionals who engage in it. The Pura Belpré Award honors the Porto Rican librarian, writer, folklorist and puppeteer who excelled in making available and palatable Latin American children's literature to Spanish speaking young readers at the New York Public Library.

Publishing for children is a way to tap a growing market nourished by national laws and social efforts for universal education. Children's literature in Latin America is, in that way, a product of global commercial and cultural exchanges that take place in asymmetrical conditions among different social groups as well as among hegemonic and peripheral nations. Nowadays, the books children read in Latin America come in large proportion from the United States, Spain, Portugal or those that are printed in regionally installed subsidiaries of American and European companies. Printing takes place in large numbers in China, Hong Kong and South Korea. In 2014, only 45% of authors of children's literature that circulated in Latin America were from the region (CERLALC, 2014, p. 62). The slow-down of publishing in Spain has been compensated by larger sales in Latin America (CERLALC, 2008, p. 88, 90–91).

Together with pedagogical and commercial practices, several factors have been and are central in configuring this field, among them the range of technological media and social practices available for writing, editing, printing, illustrating and distributing texts; general trends in the development of the fields of Literature, Education and Psychology, together with the goals of national education systems and their policies to extend reading and promote social agendas related to children's rights. There is a diversity of settings and conditions where children's lives develop: some of them live with the values and practices of modern societies while others dwell in more traditional, though evolving, settings. Access to education and books in those contexts is not the same. White Ravens' prize winner Claudio Aguilera's *9 kilómetros*, for example, describes the road a child has to walk to get to the local rural school; Chilean Marta Brunet's (1897–1967) and Salvadoran Salarrué's (1889–1975) works touch upon or center on *campesino* (peasant) stories.

Far from being idyllic, children's life in Latin America and the Caribbean is afflicted by poverty, malnutrition, conflict, exclusion, migration, child labor and parental informal employment or unemployment. According to UNICEF, in 2020 there were 188 million children and adolescents in the region, 7.9 million of which were migrants and 12 million were out of school systems (UNICEF/LACRO, 2020). Children's literature writers address these problems in various formats and genres: picture books, poetry, novels, stories, comic strips, songs and graphic novels in paper or electronic devices, in intermedial treatments, videogames and films all of which are venues for communication with young audiences while, at the same time, they are vessels to instill values and learning. In this sense children's literature can be examined for its social representations of childhood and youth.

There is a long tradition of visual communication of ideas in Latin America, both for adults and children, rooted in the ancient local traditions of mural painting, *stellae* and *codices*, a tradition which has had different modes, purposes and moments. Some of them are *aleluyas* (series of stamps with verses in the Spanish tradition), *monitos* (comic strips), *cuentos* (serial comic books in the Japanese, Franco-Belgian or North American traditions, as well as local talents and traditions as *Mafalda*, *La familia Burrón*, *Kalimán*, among others), *fotonovelas* (photographic novels) and, today, graphic novels.

In this last, flourishing genre, there are several lines of development. There are books that interpret for children, with the interplay of text and image, the lives and works of cultural heroes, like writers Jorge Luis Borges, Gabriel García Márquez or Juan Rulfo, in the works of Óscar Pantoja/Felipe Camargo and Nicolás Castell, alongside others that depict and analyze difficult historical moments of contemporary Latin American history, as *Tortas fritas de polenta* (Martinelli/Bayúgar, 2014), 2021; the anthology *Las otras islas* (Birmajer et al., 2012) about the war between Argentina and England over the Malvinas/Falklands islands; the short lived regime of Salvador Allende and the ensuing military coup that overthrew in *Los años de Allende* (Reyes and Elgueta, 2015). Japanese immigration is explained in *Los samurais de México* (Ueno/Sakuya, 2018). The challenges of growing up in Latin America are examined in works as *Diario de Un solo* (Bustos, 2016), or in *Virus tropical* (date), of Colombian-Ecuadorian author Powerpaola (Gaviria, 2018), available also in movie version. The genre engages with national and international literary traditions in *Mocha Dick: la leyenda de la ballena blanca* (Ortega, 2016), a mixed narrative of Moby Dick and the whale that appeared in Chilean Mapuche coast that inspired Melville.

Another form of visual communication that is steadily gaining authors and readers are *libros album* (picture books), narratives where artistically rendered images take the place of words. Social critique and understanding are furthered in graphic narratives as *Morro de favela* (Diniz, 2011), or in *Camino a casa*, that combines written and visual narratives (Buitrago, 2008). The small issues of a very young child's day are addressed in albums like ¡Él empezó!/He started! (Keselman, 2006). Other albums, an example of which is *¡Es una travesía esa raya!* (Leñero, 2001) have a more conceptual bend. “Album books” in Latin America, given their artistic qualities, have captivated adult readers too, as do graphic novels.

In Latin America, paper is still more prevalent than digital supports for children's literature, although there is an upward trend toward more use of digital supports, including those with some interactivity (CERLALC, 2014, 62). A silver lining of the recent Covid-19 pandemic is that large numbers of schooled children are now used to reading e-books. On the down side, exacerbated

poverty and unequal access to education due to this disease make inaccessible these media and even printed books to many young readers (OREALC et al., 2020).

What is its relationship to other literatures of the world?

Latin American children and writers are heirs to multiple literary traditions of different world cultures through the Spanish and Portuguese languages. The Spanish empire extended all the way to the Philippines; the Portuguese empire included Angola, Mozambique, Cape Verde, Goa and Macao. Furthermore, imbibed in Arabic and Sephardic traditions, rooted in classical Judeo-Christian and Greco-Roman texts as well as part of those of ancient Carthage, India and Persia, the Spanish and Portuguese literatures that arrived to the New World—at the time Jews and Muslims were expelled from Spain and some sought refuge in Portugal and America—are a rich mix of topics and genres that took root in novel soil.

The elites of the colonies and later independent states of Latin America have traditionally had full access to the currents, trends and authors of the Western canon, including children's repertoires, for more than 5 centuries, be it directly or in translations and reappropriations of texts for children or adult readers.

Popular culture receives the western canon and transforms it in a different way. One well-studied example of the new life given to ancient texts in an entanglement of literary traditions is *Fierabrás*, a *chanson de geste* originally composed in French toward the end of the 12th century, which was translated into Castilian as *Carlomagno y los doce pares de Francia* in 1521, immediately becoming immensely popular in Spanish and Portuguese colonies to the point that "it became part of the blueprint to build a new religiously and culturally hybrid society in the Americas and the Philippines" (Grinberg, 2012, p. 139).

The concept of opposition between two forceful characters found in the performed representation of the bloody and cruel Battle of Roncesvalles, fought in 778 by Charlemagne against Turkish forces, became a causeway for original binary pre-Hispanic conceptions all over the region, danced to music and dialog. It is likely that in the literary rendition of this battle young readers found a metaphor and the basis for their elders' performance of a tradition of rebelliousness, subversion, confrontation and taking action for a community (Flores, 2009, pp. 100–101).

All over America there are versions of tales of the battles and skirmishes between two animals, one large and mighty and the other small and meek, but crafty: *Macho Ratón* in Nicaragua, *Conejo* and *Coyote* in México, *Br'er Rabbit* in the United States, *El zorro y el sapo* in Perú, *Tío Conejo* and *Tía Boa* in Costa Rica, *Tío Tigre* and *Tío Conejo*, known to children in several countries. Most contemporarily, the tale has been retold in *Pancho Rabbit and the Coyote* (Tonatiuh, 2013) to address migration.

Commerce of ideas in the field of children's literature during the 19th and 20th centuries can be seen in the case of the development of British Arthur Mee's *The Children's Encyclopedia*, reworked and re-edited in the United States as *The Book of Knowledge* (Jackson et al., 1912) which had several editions in different languages and countries, including China and Russia, as well as *Tesouro da Juventude* for Luso-Brazilian readers and *El tesoro de la juventud* for Spanish speaking countries.

If we compare the *Tesouro* with *El Tesoro* we see that each is not only a version in Portuguese or Spanish of *The Book of Knowledge* but, from the viewpoint of this article, a quite different edition, given that each offered its own repertoire of children's literature. Furthermore, the Portuguese language edition did not have a permanent section on "Latin America" as did the Spanish language edition which featured it in each of its twenty volumes (Riesco, 2008, 198–225; de Oliveira, 2018, 105–128).

Different currents of migration have defined life in Latin America. Its vernacular linguistic diversity bears witness to the multiple origins and the resilience of its first inhabitants; afterward, there has been a continuous flow of European, African, Oceanian, Middle Eastern and Asian immigrants as the result of colonization from the 16th to the 19th centuries, and again during the 20th century, when Latin America was haven for those who fled two world wars and their human, economic and political consequences.

Many Latin American individuals and families have, in their turn—notably after the second half of the 20th century—migrated in search of a better life both within Latin America and to other regions, mainly North America and Europe, thus constituting linguistic communities abroad that among many other needs require reading materials for children in their mother tongues. The first Chicano novel, *Las aventuras de Don Chipote o cuando los pericos mamen* (Venegas, 1999) speaks of this kind of displacement. Contemporarily there are compelling accounts as *Migrar* (Mateo and Martínez, 2012) or *Somos como las nubes/We are like the clouds* (Argueta, 2016).

The widespread phenomenon of migration has spurred a market for Latin American children's books in regions outside of Latin America. There is a notable growth of authors who write in English for children whose mother tongue, tongue of heritage, tongue of instruction, or second language is Spanish or Portuguese. These authors take topics of both the Latin American traditions and the migrant and first-generation American experience together with elements of oral language as code-switching or sprinkling words in either of those languages into a text that is mainly in English. Several prestigious awards acknowledge and stimulate this production, among them The America's Award, first discerned in 1993. Among the recipients are Francisco Jiménez, Francisco X. Alarcón, Jorge Argueta, Alma Flor Ada and Gary Soto.

Latin America has contributed to world thinking and rethinking of fantastic, improbable, unthinkable and un-to-fore unimagined worlds. This region has been seen in world literature as the site of actual or potential utopia, a locus of the fight for freedom and the renovating winds of revolution, as well as of innumerable sources of wealth. It is also the origin of topics, genres, characters and settings.

Early to 19th century European texts that spoke of Latin America eventually made their way to European children's literature and in time reached the children and youths of America, who read about their land through eyes that were not theirs. For example, Lucien Biart's *Aventures d'un jeune naturaliste* (1869)—mentioned by Rudyard Kipling's character Kim as one of his favorite

readings—puts down local inhabitants. With examples like this in mind, writers, editors and educators have striven to portray Latin America and Latin Americans in a light that makes them justice. Readers have struggled for representational justice internally too. In 1996 two Costa Rican adolescents filed a complaint against the practice of reading in schools Joaquín Gutiérrez Manguel's children's classic about a black girl, *Cocorí* (Gutiérrez, 1947), a text translated to a myriad of languages. In 2005 a group of African-American opinion leaders officially decried the racist stereotypical visual rendition of *Memín Pinguín*, (Vargas Dulché, 1943) a black child character in a serial comic book. In both cases the polemic is still open.

The appearance and reception of Ruben Darío's *Prosas profanas* (1896) marked the coming of age of Hispanic American literature and gave authors of the region an equal footing with their European peers. Written for a girl who would later give birth to the Nicaraguan dictator Anastasio Somoza, Darío's *Margarita* is a beloved story in verse for children. *Realismo mágico*, during the second half of the 20th century, brought a distinctive form of fantasy to world literature. Earlier in the 20th century, Monteiro Lobato offered Brazilian children the series *Sítio do picapão amarelo* (1921–1944), where they could explore life in a farmland and imaginary worlds.

Poetic prose and poetry, with strong roots in Native, Spanish and Portuguese literatures, gained momentum in the area by contact with the Bengali Renaissance (Tagore), Maeterlinck, English nursery and nonsense rhymes, as well as with the work of the poetic avant-gardes. Thus, little by little books for children could be poetic, fun and entertaining, although in them science had a higher hand than fairies, which fluttered mainly around English children's stories and were practically banned from children's books, especially schoolbooks, in countries where governments embraced a Positivistic ideology, such as Brazil and Mexico, although, decades later, also in Cuba (Andricáin and Rodríguez, 2016, 393). Marina Colasanti, in Brazil, has given fairy tales a contemporary, feminist and ludic twist (Colasanti, 2015); Uruguayan writer Edna Iturralde revisits the Brothers Grimm in *Los hermanos que cosechaban cuentos de hadas* (Iturralde and Icaza, 2016) (Fig. 8).

It is not easy to gauge the recent production of books for children and adolescents in the region, nor to assess its quality, or even to have a reliable, comprehensive historical view of its development. Please see historical and critical works in "Brief selection of other relevant Children's Literature Studies" in References.

The longest standing seminar on children's literature is the *Cátedra de Literatura Infantil*, inaugurated in 1917 by Joaquín García Monge, editor of the journal *Repertorio Americano*, at the Teacher's University of Costa Rica, a double endeavor that steadily nourished the development of children's literature in the region.

Academies for the study and promotion of children's literature have later sprung up in Latin America. After founding the Uruguayan Academy of Children's Literature in that country, Sylvia Puentes de Oyénard established, in 2002, in Perú, the Latin American Academy of Children's and Youth's Literature, inviting scholars, teachers and librarians to set up an equivalent in each of their countries (Pizarro et al., 2020, p. 5).

Increasingly universities have become interested in children's literature as subject of teaching, research and dissertation topics. There are few regional scholarly periodical publications, among them, the award winning *LIIlbero* (see Fig. 9). For teachers,



Fig. 8 José Joaquín de Arriaga. *La ciencia recreativa*. Publicación dedicada a los niños y a las clases trabajadoras. Mexico, Tip. de J. M. Aguilar Ortiz, 1873. Lithography by Iriarte.

researchers and *mediadores*, IBBY produces *Revista Latinoamericana de Literatura Infantil y Juvenil* and Norma Publishing House, *Barataria*. The most recent magazine is IBBY—Latinoamerica's *Voces y tintas*, published in Spanish and English, which started in April 2021 with a regional scope and relative autonomy from publishing houses that are welcome.

Not all of studies or compendia are originated in or funded by universities. The role of national institutions geared to increase reading or non-governmental organizations and commercial publishing houses has been crucial in analyzing and disseminating information about Latin American children's books, authors and illustrators, as well as in fomenting the rise of new voices and texts through contests and publishing ventures. They have propelled the production of catalogs, yearbooks and reference books, reading guides, teaching materials and contests which have been helpful in fomenting children's literature and spreading knowledge about it in the area.

Discussion: the concept of Latin America in relation to children's literature

Placing a set of authors and works under the umbrella of "Latin American children's literature" has consequences that pertain to the chronology, the languages and the corpora that are to be studied, at the time it determines the choice of cultural elements to analyze this production.

Latin America's social configuration and literary heritage

Simon Bolívar dreamed of "*una patria grande*" (an all-encompassing homeland). "Latin America" is a term coined to proclaim the cultural unity of a vast geographical region. The term first mentioned appeared in 1836 (Ardao, 2000, p. 157) and was originally used in political, not geographical writings, to group and refer to "Latin" as opposed to "Anglo-Saxon" cultures in the American Continent.

This loosely defined concept—especially in the works of Uruguayan José Enrique Rodó (1871–1917), read with interest and fervor by youths of Spanish speaking "Latin America" during the first two decades of the 20th century—divided the population of the American Continent into two clusters and emphasized the clash of ideologies of the peoples of the "North" and those of the "South" due to their different languages and cultural traditions as well as their opposed religions (Roman Catholicism in Latin America and the Protestant Church among the Anglo-Saxon population of North America). Thus defined, the term provided an



Fig. 9 *LIJ-Ibero. Revista de Literatura Infantil y Juvenil Contemporánea*. Recipient of the *Hormiguita viajera* National and Latin American Award for Children's Literature in 2017.

ideological basis for French and Spanish interests in the region in an era of expansion of the United States and its Monroe Doctrine. In children's books, this is the time of North American companies such as Appleton and others which extended to Latin America.

Canada and Haiti, in spite of their French speaking populations, were not part of this first definition of "Latin America"; neither was Brazil, until 1889, after this country became a democracy and slavery was abolished.

In spite of its lack of precision and notwithstanding the fact that it does not take into consideration neither Native nor African ascendancy in the cultural definition of the area, the term "Latin America" served as a call for union among the thinkers and nations that embraced it.

The term emphasizes a legacy that is fruit of colonization, and does not take into account that as both its premise and aftermath the aboriginal population of the region was dispossessed and reduced to serfdom in order for European settlers to have a work force to build and give maintenance to roads, homes, hospitals and churches and at the same time organize the production of haciendas, plantations and mines. The cultural resistance of the local population decimated by disease, famine and work took the form of rebellion, frugal life, stoic endurance and persistence of vernacular languages, while poverty was exulted by early Franciscan friars and seen as natural in colonial society.

Dignity in poverty, as well as persistence in the use of mother tongues became part of peasant culture and rural children's lore as much as the oral literature that has survived and adapted to changing circumstances. Salient among what has been transmitted of this literature to our day are tales relative to Creation and the place of Humankind in the Universe; myths, fables, rituals, legends, riddles, poetry, mock battles and dances, which have specific genres as are Quechuan *wawaki*, lyrical dialogs and *urpi*, short poems (Hiflin, 1991, pp. 37–38).

Content of this lore has been transmitted in different ways. For example, the cosmogony and saga of the Quiché, *Popol Vuh*, written in colonial times by aboriginal authors, has been source for children's literature in adapted renditions, as that of Ermilo Abreu Gómez. In an indirect form of inspiration, an Inca knot system for communicating, *quipus*, and ancient tales were the basis for the *Cuentos peruanos* of María Wiese and Carlota Carvallo's tales for children, which honed into the ancient tales in *Rutsí* and *El pájaro niño* (Hiflin, 1991, p. 207). The figure of the traditional storyteller, *Amauta* among Andean cultures, often an aged person, persists in both oral and written traditions, as is the "*chasqui*," messenger who carried news in the Inca trail, name taken by a Bolivian magazine for children (Peña Muñoz, 2009, p. 531). Although probably altered in the early process of transmission from Indian languages to Spanish, there is record of early riddles, sermons and myths that extoll the role of virtue in childhood and youth, as *huehuetlatolli* in Nahuatl and *haravicu* in Quechua, as well as texts in non-alphabetic forms of representation that remind the beholder of the moral obligation to make the world a better place for all (Bonifaz Nuño, 1987, pp. 41–55).

In spite of long-standing acculturation processes, this endogenous or vernacular literature presents a set of narrative codes and value systems that differ in important ways from the Western Judeo-Christian and Greco-Latin canons, in this last case most notably in what pertains to community life and the relationship with nature. Specific genres have been identified, as the Pemon "*ekaré*," which designates a rumor or news of a true or testimonial historical fact, and "*pantón*," associated to fictional narrative (Peña Muñoz, 2009, p. 410). Contemporary renditions of this heritage are Ana Castillo's *Mi hija, mi hijo. El Águila. La Paloma: Un canto azteca/My Daughter, My son, The Eagle, the Dove: An Aztec Chant* (2000) and *Cuentos de los abuelos* (Bonifaz Nuño and Pizarro, 1999). Nevertheless, it is not the case that all countries publish children's literature in indigenous languages. In Ecuador, for example, in 2014 children's literature was not written in any of the national languages other than Spanish, in spite of a long tradition of ethnographic work with oral literature and local legends (Franco, 2014, pp. 235–236).

African legacy

Slavery was widespread and profitable in both Spanish and Portuguese dominions. Islands of the Caribbean and Brazil were important transit and end points of the slave trade that established international networks since early times—in the region from the early 16th to well into the end of the 19th century—as a means to have increased access to labor. With the work of the enslaved population, the area became highly productive. At the same time, the languages and traditions of different African peoples became part of the culture of the New World, and thus enriched the corpus of lullabies, songs, poetry, stories and fables for children with tales of wonder, narratives of suffering and resistance, an attitude of joy of living despite extreme hardship, and a form of spirituality that is different from Roman Catholicism, all underscored by the plight of slavery and the ensuing attitudes of rebellion and fight for freedom.

In the stream of high culture, there have been notable artists and intellectuals who have delved into the African heritage of Latin America, among them children literature authors as José Bento Monteiro Lobato, Mario de Andrade and Rogério Andrade Barbosa. Among contemporary works, examples of narratives about peoples of African cultural ascents are *Paquelé* (Llanes, 2006), which explores the life of a slave in colonial Cuba and *Y su corazón escapó para convertirse en pájaro* (Iturralde, 2001).

Alternative terms to name the region, and their consequences for children's literature

To refer centrally to aboriginal cultures in the region's identity José Carlos Mariátegui proposed the term "Indoamérica," taking as root the word "Indian," as the aboriginal population was called since colonial times. The *indigenista* discourse of Peruvians Mariátegui and José María Arguedas vindicated the rights and cultural values of the indigenous population and its cultural heritage as distinctive of the nation and the region.

Mariátegui's avant-garde magazine *Amauta*, first published in 1926, inspired essays on children's literature and children's stories rooted in vernacular traditions, respectively, by writers María Wiese and Carlota Carvallo—one of Peru's 20th century pioneer

children's writers, together with Francisco Izquierdo Ríos. Writers and illustrators found a wealth of prime materials to work with by latching onto the Arts and crafts movement that proposed research in folklore and Native traditions to find new, authentic venues for national creation, as did Elena Izcué. Later, in Bolivia, Rosalba Guzmán and Gaby Vallejo, salient in a group of talented writers and educators, created several initiatives and children's magazines that gave children access to Bolivian contents and cultural forms that were conceived as especially relevant and situated (Peña Muñoz, 2009, pp. 530–542). Atavistically attached to their land, it was hard to integrate Indians to industrial society and the proletariat. In what has become a classic of children's literature, *Pablo Yunque*, Peruvian César Vallejo portrays the hardships of poor urban children of Indian ancestry in societal tensions of race and class (Vallejo, 1981).

How many “roots”?

Twentieth century developments of the idea of Latin America gave way to that of a “double root”: Indian and European. The African heritage was rarely identified or acknowledged inasmuch the myth of the “double root” was entrenched. We can see this in the field of children's literature by examining the contents of *Lecturas clásicas para niños* (1924), a project of Latinamericanist José Vasconcelos that consisted in adapting a set of canonical classical works from different Eastern, Western and local traditions, in which those of African origin were not included.

In the 21st century, African ancestry has been claimed or conceived of as a “third root,” together with Native and European heritages. The foundational role gradually acknowledged to Native and African cultures, however, has not translated into equal access to human rights and economic prosperity for the Indian or Afrodescendant populations. These groups are nowadays equal in rights to people from European or mixed ascendancy, but are disproportionately afflicted by unequal access to well remunerated work or to full social, health and educational services; they are especially vulnerable to forms of violence, racism, classism and sexism that undermine their rightful exercise of human rights (Bello and Rangel, 2002, pp. 39–54).

In 2010, 63% of all children in the region were affected by poverty in some way, while the situation was even more critical among indigenous children, 88% of whom were affected, and three times more indigenous children living in extreme poverty suffer severe privation in access to education, safe drinking water and housing than other children (ECLAC and UNICEF, 2012, 4–6). Inequality in the area has increased in the last decade, and was worsened by the Covid-19 pandemic (ECLAC, 2016; ECLAC and UNOCHR 2020). The region's population, including children, is afflicted by what has been called “accumulative disadvantages,” related to gender, age, ethnicity, health and place of residence (UNESCO, 2020).

Nevertheless, the place of Indigenous and African cultures in Latin America is decisive and widespread.

An estimated 41.8 million indigenous people lived in Latin America in 2010 (World Bank, 2015, 25). Currently 420 languages are spoken and 99 language families can be discerned in Latin America, according to the *Atlas sociolingüístico de Pueblos Indígenas de America Latina*. Currently the Amazon is the region with the largest diversity of indigenous peoples (316 groups), followed by Mesoamerica, the Orinoco basin, the Andes and the Chaco region (World Bank, 2015, 24). The distribution of Indian languages does not coincide with current national boundaries. According to the mentioned Sociolinguistic Atlas of Latin America, one-fourth of all Indian languages are spoken in two or more countries, and Quichua extends to seven countries (Sichra, 2009, ix). The area is inhabited by 665 peoples with a total of 29, 464, 00 persons (Sichra, 2009, 65). Almost a fifth of these individuals have stopped speaking a native language.

There are national and local initiatives to respond to the reduction of speakers and vernacular languages: old and new texts are published so that children can practice their mother tongue or that of their elders. Bilingual editions are growingly available due to national legislation and educational programs that foster the cultural revival or strengthening of communities where native languages are or were spoken. In several countries native languages are recognized as national languages and there is a thrust to integrate Indian languages, knowledge and values, contemporarily understood as “native epistemologies,” into the general curriculum. Forerunner countries of these policies are Peru, Bolivia, Mexico and Ecuador.

Thus, Latin American children's literature in the 21st century is in part geared to fostering the national knowledge of aboriginal cultures by drawing mainly on myths, folktales, fables, riddles and depictions of life in rural areas, mainly in Spanish or Portuguese, but increasingly also in Indian language editions. Several national, regional and international institutions and publishing houses have devoted books or series to children. Nevertheless, it is not a practice that is sufficiently extended. In decolonizing critique, this move to appropriate indigenous lore has been seen as an attempt to make the settler indigenous to the land he occupies (Tuck and Yang, 2012, pp. 1–40).

In a different line of inquiry, one can find that though needed and welcome, children's books for indigenous population in vernacular languages are disproportionately centered on traditional folklore, generally omitting both new writers and different topics in the humanities, sciences, technology or current social issues, with scarce use of either endogenous or innovative literary forms, all of which would contribute to the expansion of horizons that is one of the expected outcomes of reading literature. Some initiatives are geared to compiling and publishing texts written by children, for example *Relatos y andanzas. Historias de niños y niñas de los pueblos originarios de Chile* (Cuevas Parra, 2002) or the games they play, as in CONAFE's *Así cuentan y juegan en la Huasteca* (CONAFE, 1982). In any case, as it is stated in an intergovernmental-sponsored document, “the poverty of indigenous children cannot be understood without considering the systematic impoverishment to which their peoples have been subjected” (del Popolo, 2012 p. 6). Literature can contribute to this understanding.

In what pertains to African cultures in Latin America, according to UNESCO, the estimated number of people enslaved in Africa and taken to America between 1500 and 1867 is 12.5 million. The descendants of this African diaspora are called “Afrodescendants”

in Latin America since 2001—after the Durban World Conference against Racism, Racial Discrimination, Xenophobia and Related Intolerance. In 2015 Afrodescendent population of the region amounted to 130 million people, that is, 21.5% of the total. In terms of absolute numbers, Brazil has the largest Afrodescendent population. In the year 2000, 45% Brazilians identified themselves as Black or Mulatto (Abramo and Rangel, 2019). Nevertheless, the role of African population in the national construction of identity has only recently been acknowledged in national laws and policies of the region.

In children's literature, "Black" and "slave" were, until very recently, treated as synonyms, which reinforced racist attitudes in young minds. In spite of much earlier abolition of slavery in other countries of Latin America, in Brazil it was legal until 1880; 1886 in Cuba. This tardiness explains and reinforces structural racism against Afrodescendent people in these countries and in the region, where during the 21st century African ancestry has been claimed or conceived of as a "third root," together with Indian and European heritage, as was stated above. This should mean making available books and visual narratives for children that portray African heritage and Afrodescendent experience in Latin America in an accurate, dignified and appealing manner, by exploring either the cultures and languages of Africa or their development in America.

In literature, although early authors of *negrista* poetry imitated, recreated or invented African rhythms and words, at the same time they made reference to drums and other percussion instruments, as did Juana Inés de la Cruz in New Spain, it was not until well into the 20th century that authors of African ascent like Nicolás Guillén and Emilio Balagas, in Cuba, wrote of African heritage from their own experience, or produced criticism that decried social exploitation that went beyond the attitude of commiseration of most of 19th century production. Avant-garde authors, among them Porto Rican Luis Palé Malós, experimented in poetry with African heritage rhythms. Trained anthropologists directed their inquiries to the African legacy to local folklore, language and spirituality. In Costa Rica, the Africanized adaptation of an Andalusian tale: "Ratón Pérez y la Cucarachita Mandinga" appears in *Cuentos de mi Tía Panchita* (Lyra, 1984; Cuesta de Vélez, 2013, 365); in Colombia, classic children's writer Rafael Pombo imitated words of African root when he named the protagonist of the fable "Mirringa Mirronga. La Gata Candonga"; *Silver people: Voices from the Panama Canal* (Engle, 2014) commemorates for young readers the "darker skinned" people who worked on the site more than 100 years ago.

In Brazil, Romantic writer Antonio Frederico de Castro Alves penned the poem *O navio negreiro* (Castro Alves, 1880) and the novel *Os escravos* (Castro Alves, 1883) while Modernist José Lins do Rego gave life to *Menino de engenho* (Lins do Rego, 1932). Aimé Césaire's *Discours sur le colonialisme* and the African movement of *Négritude* had resonance in Latin American culture (Clouthard, 31). Contemporarily, André Diniz's *Morro de favela* (Diniz, 2011) portrays, with the aid of color photography, the different lives of Blacks and Whites. Heloisa Pires Lima, Liz Zatz, Julio Emilio Braz and Rogerio Andrade Barbosa are writers who work on the line of dignified representation (Meléndez Moré, 2019, 13).

Nevertheless, Latin American authors, educators and publishers still have to find ways to responsibly, effectively and persuasively work literarily with multiracial, multicultural and multilingual individual and collective identities. There is currently a commitment to portray Indian communities and Afrodescendants in a way that is accurate and fair at the time it accounts for the diversity of historical and actual experiences and contributions of these populations to the culture and welfare of Latin American societies, although miserabilistic or folkloric approaches are not always avoided.

Other denominations of the area, and their consequences for children's literature

Other denominations of the region we have so far called "Latin America" have different political and practical implications. José Martí kept the basic distinction drawn by Rodó when he used the term "*Nuestra América*" (*our America*). Martí's conception can be clearly seen in his children's periodical *La Edad de Oro* (1889), where he admonishes readers to learn what is *theirs* and read *their* classics, as much as he calls on their parents to drink local wine, even if it is made of bananas. In his conception, pre-Hispanic cultures together with the fight for freedom and independence define the region, so children's literature must first provide that basis to later on open up to the world.

Right after the independence of Cuba from Spain, in lieu of "Latin America," James G. Blaine proposed the term "Pan America," which comprises North and South of the Continent, a logical extension of the notion of "America for Americans" of the Monroe Doctrine (Ardao, 2000, p. 159). American and International publishing houses produced series of stories about the Continent for children, among them both W. Jackson and Herder. The Pan American Bureau in 1945 put forth a "Series for Young Readers" as well as the compilation *Children's Books on Latin America* and *Materials in English on Latin American Literature*. In the same vein, the Biblioteca Billiken produced the "Blue collection," in which biographies of notable men and women of all the Continent were presented to children, a series which ran at least from 1939 to 1968, with five editions (Varela, 1994).

In view of the loss of its colonies in America, Spain and Portugal advanced the idea of "Hispanic America," or "Iberoamérica" on which "Hispanic" referred to the Hispanic Peninsula to encompass both Spain and Portugal, and their interests in the region. Contemporarily Spanish and Portuguese publishing companies—endorsing this conception—have activity in several countries of the region. This has consequences for the relationships among authors, publishing house and readers, because it induces metropolitan-style dominance in the distribution and ranking of works, as well as in its marketing strategies, which consist in the recognizable homogeneity among all products (Barrera Anderle, 2002). This apparently runs counter to what in our day has been hailed in the region as "Bibliodiversity," at least in one of its possible interpretations (XXVI Cumbre Iberoamericana, 2018, 15–16).

During the 20th century political and economic agreements and institutions redefined the scope of the term "Latin America." ECLAC grouped "Latin America" and "the Caribbean"—which already included former Spanish dominions in the Caribbean—to

include Antigua and Barbuda, Bahamas, Barbados, Belice, Guyana, Haití, Jamaica, Saint Lucia, Trinidad and Tobago, Suriname. Due to political considerations, the Organization of American States includes all the countries of the American Continent except Cuba.

The Regional Office of UNESCO for Latin America and the Caribbean, through CERLALC (Regional Center for the Support of Book Publishing) groups Latin American book production with that of Spain and Portugal, which includes titles in three Spanish national languages: Catalan, Gallegan and Basque (González Mendoza, 2021; González Mendoza and Wisenbart, 2019). By the same token, books in Spanish and Portuguese are written today and published in the United States of America, also produced to be exported to Latin American countries.

Conclusions

Children's literature in Latin America is today a growing industry and a social endeavor that is still woven into the goal of contributing to ensure that every child has the opportunity to reach adequate levels of reading skills and is thus prepared to exert his or her right to education, so it has been difficult to completely shake off didactic means and goals in favor of literary exploration of forms and values which are formative in their own way.

The study of the literature for children of the region depends on how it is defined, by whom and with what purpose. Within the region, it has been used mainly to convince children to embrace certain values and attitudes; from outside, Latin American children's literature has mainly served as a means to understand and sometimes either buttress or distort or redefine the identity of individuals and communities that are related in some way to those lands, and not only as a means of artistic communication with young readers. Taking into account that identity is dynamic, and that literature is the art of words and form, Latin American children's literature can be approached in more nuanced and diversified ways than those that in large measure have been used up to now. Authors, editors, illustrators and editors can continue to strive for what Alicia Borinsky has described as "invoking a region beyond literature through literature" (Borinsky, 1993 ix-x) and purport to speak to children of all that which authors, illustrators, parents, teachers, institutions or governments find valuable, noteworthy or worrying in the history, actual life or future of their country, their region or the world, Humankind and the Universe, through the means of art, making resource to local and universal genres, topics and styles in a way children and youths feel respectfully addressed, creatively engaged, esthetically moved, enlightened and free to examine and judge for themselves.

Great strength would be gained from easy access to both the historical and the new production of different countries, of which so far most circulates mainly within its own borders. The historical archive is practically inaccessible, despite serious efforts of national libraries and other institutions geared to providing bibliographic services. In general terms, a more equitable access to education and culture, and a fairer distribution of income would, undoubtedly, make life and literature better in the area. Offering more opportunities for all children and youths to exert their rights to read and write, to listen to and to tell stories, in individual, collective and inter-generational processes, in all the languages and traditions of the region, in conversations with literatures and languages of the world, would bolster and fairly distribute among the population skills for discernment, appropriation, mastery, enjoyment and creativity. This can only strengthen every person's capacity to create and enjoy children's literature, and to share it with the world.

Further Reading

- Abramo, L., Rangel, M., Nov. 15 2019. Children of African descent in Latin America. In: Newsletter on Progress Towards the Millennium Development Goals From a Child Rights Perspective. ECLAC. Online version.
- Andricain, S., Rodríguez, A.O., 2016. La censura de autores y libros de LJJ en Cuba. In: Cerrillo, P.C., Sotomayor, M.V. (Eds.), *Censuras y LJJ en el siglo XX (en España y 7 países latinoamericanos)*. Ediciones de la Universidad de Castilla-La Mancha, Cuenca.
- Ardao, A., 2000. Panamericanismo y latinoamericanismo. In: Zea, L. (Ed.), *América Latina en sus ideas. Siglo Veintiuno Editores/UNESCO*, México City, pp. 157–171.
- Barrera Anderle, V., 2002. Entradas y salidas del fenómeno literario actual o "alfaguarización" de la literatura hispano-americana. *Sincronía* (1). Online version.
- Bello, Á., Rangel, M., April 2002. La equidad y la exclusión de los pueblos indígenas y afrodescendientes en América Latina y el Caribe. *Revista de la Cepal* 76, 39–54.
- Bonifaz Nuño, R., 1987. El humanismo prehispánico. In: *El humanismo en vísperas del siglo XXI*. Universidad Nacional Autónoma de México, México, pp. 41–61.
- Borinsky, A., 1993. *Theoretical Fables: The Pedagogical Dream in Contemporary Latin American Literatures*. University of Pennsylvania Press, Philadelphia.
- Carranza Vera, C., Silva Martínez, A., Jul–Dec. 2017. Un ejemplo de comedia de magia en el teatro infantil de México: La almoneda del diablo (siglos XIX y XX). *Literatura Mexicana* XXIII, 9–34.
- CERLALC (Centro para el fomento del libro en América Latina, el Caribe, España y Portugal), 2008. *Catálogo de autores y libros de América latina*. Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba. CERLALC, Madrid.
- CERLALC (Centro para el fomento del libro en América Latina, el Caribe, España y Portugal), 2014. *El espacio latinoamericano del libro 2014*. CERLALC, Bogotá.
- CERLALC, April 2020. *Leer, imaginar, actuar. Catálogo CERLALC-IBBY de libros infantiles para el desarrollo sostenible*. CERLALC/IBBY, Bogotá.
- Cuesta de Vélez, C., 2013. La travesía de la Cucarachita Martínez y el Ratón Pérez. *Educere* 17 (57), 361–367.
- de Guatemala, D., 2018. XXVI Cumbre Iberoamericana de Jefes de Estado y de Gobierno. XXVI Cumbre Iberoamericana de Jefes de Estado y de Gobierno. *Compromiso Iberoamericano por el Desarrollo sostenible*, Guatemala City, Guatemala.
- del Popolo, F., September 2012. The right to well-being for indigenous children: progress to date and the current situation in Latin America. In: *Challenges. Newsletter on Progress Towards the Millennium Development Goals From a Child Rights Perspective*. Number 14. The Rights of Indigenous Children. ECLAC and UNICEF, Santiago, pp. 4–9.
- Del Solar Tola, P.P., 2019. "Aquí la muerte nos ronda a todos, Cuy": tiras cómicas y violencia política en el Perú. *LJLibero* 4 (7), 23–42.
- Díaz Rönnér, A., 2000. *Literatura infantil de "menor" a "mayor"*. In: Jitrik, N., Drucaroff, C.E. (Eds.), *Historia crítica de la literatura argentina*, vol. 11. Emecé, Buenos Aires, pp. 511–531.
- ECLAC, November 2016. *La matriz de la desigualdad social en América Latina*. ECLAC, Santiago.
- ECLAC, May 2020. *El desafío social en tiempos del Covid-19*. ECLAC, Santiago.

- ECLAC (Economic Commission for Latin America and the Caribbean)/Office of the United Nations High Commissioner for Human Rights (OHCHR), 2020. People of African Descent in Latin America and the Caribbean: Developing Indicators to Measure and counter Inequalities (LCT/TZ.2019/62). ECLAC, Santiago.
- Fernández de Zamora, R.M., 2008. Los concilios mexicanos promotores del libro y la lectura en el siglo XVI. *Investigación Bibliotecológica* 22 (45), 105–123.
- Fioravanti, C., 2013. Los precursores de Lobato. *Pesquisa FAPESP* 253.
- Flores, E., 2009. Carlomagno en cordel. Entre la tradición y el canon. In: Domenella, A.R., et al. (Eds.), *Homenaje a Yvette Jiménez de Báez*. El Colegio de México, Mexico, pp. 93–122.
- Franco, B., 2014. La literatura infantil como formadora de identidades nacionales, en Ecuador y Suramérica. *Saber, Ciencia y Libertad* 9 (2), 229–239.
- González Mendoza, J.D., 2021. El Espacio Iberoamericano del Libro 2020. CERLALC, Bogotá.
- González Mendoza, J.D., Wishenbart, R., 2019. El espacio iberoamericano del libro 2018. CERLALC, Bogotá.
- Grinberg, A., 2012. (Un) stable Identities: Impersonation, Conversion and Relocation in Historias del emperador Carlo Magno y los doce pares. University of California, San Diego, San Diego. PhD Dissertation.
- Guadarrama Olivera, H., 2000. Ildefonso Estrada y Zenea: un intelectual cubano en el puerto de Veracruz, 1871-1872. *Sotavento* 8, 71–93.
- Guerrero Guadarrama, L., 2016. Neosubversión en la LJ contemporánea: una aproximación a México y España. Ilustrado por Thales I. Aquino. *Textofilia/Universidad Iberoamericana*, México.
- Hiffin, D.D., 1991. La contribución de la cuentística de Carlota Carvallo a la literatura infantil peruana. A dissertation in Spanish Submitted to the Graduate Faculty of Texas Tech University in partial fulfillment of the Requirements for the Degree of Doctor of Philosophy.
- Maldonado de Masaya, N., Barco, F.L.M., Mendoza, B.L., 2011. Voces de la literatura infantil y juvenil de Guatemala. Universidad de San Carlos, Guatemala.
- Meléndez Moré, M.P., 2019. Literatura infantil con representaciones afro: promoción lectora y construcción identitaria. Universidad Veracruzana, Centro de Estudios de la Cultura y la Comunicación. Especialización en Promoción de la Lectura. Trabajo recepcional de la Especialización.
- Oliveira, B.J.de, 2018. Science and its appropriation in the twentieth century in Latin America. *Br. Soc. Hist. Sci.* Thames 3, 105–128.
- OREALC, ECLAC, UNESCO, 2020. Report Education in the Times of COVID-19. OREALC, Santiago.
- OREALC/UNESCO, June 2021. A un año de la pandemia. Continuidad educativa y evaluación en América Latina y el Caribe en 2021. UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean/Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, Santiago.
- Peña Muñoz, M., 2009. Historia de la literatura infantil en América Latina. Fundación S/M, Bogotá.
- Pizarro, C., Prina, Z., Bustos, M.B., 2020. In: Fernanda, M. (Ed.), *La gesta literaria de Sylvia Puentes de Oyenard en la difusión de la Literatura Infantil y Juvenil de América Latina*. Homenaje. Digital Book, Online Version. Macimiani in Leeme un cuento ediciones.
- Prashay, V. (Ed.), 2019. *The East Was Read. Socialist Culture in the Third World*. LeftWord Books, New Delhi.
- Ratiff, W.E., 1969. Chinese communist cultural diplomacy towards Latin America, 1949–1969. *Hisp. Am. Hist. Rev.* 49 (1), 53–79.
- Riesco, L., 2008. El maravilloso mundo de El Tesoro de la Juventud: apuntes históricos de una enciclopedia para niños. *Universum* 23 (1). Talca.
- Sichra, I. (Ed.), 2009. *Atlas sociolingüístico de pueblos indígenas en América Latina y el Caribe*. UNICEF and FUNPROIEB Andes, Cochabamba, Bolivia.
- Stinchcomb, D.F., 2001. *The Development of Literary Blackness in the Dominican Republic*. PhD Dissertation. The University of Tennessee, Knoxville. Published in 2004 as *The Development of Literary Blackness in the Dominican Republic* by University Press of Florida, 2004.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization* 1 (1), 1–40.
- UNESCO, 2020. *Global Education Monitoring Report 2020. Inclusion and Education. All Means All*. UNESCO, Paris.
- UNICEF/LACRO. *Children in Latin America and the Caribbean, 2020. Regional Overview on the Situation of Children and Adolescents*. UNICEF, Panama.
- Varela, M., 1994. *Los hombres ilustres del Billiken. Héroes en los medios y en la escuela*. Editorial Colihue, Buenos Aires.
- World Bank, 2015. *Indigenous Latin America in the Twenty First Century: The First Decade*. The World Bank Group, Washington.
- XXVI Cumbre Iberoamericana, 2018. *Agenda para Reforzar el Acceso Democrático al Libro, la Lectura y la Escritura* (Documento SEGIB-CERLALC). Secretaría General Iberoamericana/Secretaría para la Cooperación/Espacio Cultural Iberoamericano, Guatemala.

Cited works of Latin American children's and youths literature

- Alarcón, A., 1906. *De mi tierra y de mi alma*. Imp. Velarde, La Paz, Bolivia.
- Illustrator Alfaro, Ó., Osorio, E., 1990. *El Pájaro de Fuego y otros Cuentos*. Editorial Andrés Bello, Santiago de Chile. Illustrator.
- Amado, J., 1978. *O Gato Malhado e a Andorinha Sinhá*. Uma história de amor. Illustrated by Carybé. Publicações Dom Quixote, Lisbon. Originally written in 1948.
- Anonymous/version of Abreu Gómez, Ermilo, 2003. *Popol Vuh*. Antiguas leyendas del Quiché. Fondo de Cultura Económica, México.
- Argueta, J., 2016. *Somos como las nubes*/We are Like the Clouds. Illustrated by Alfonso Ruano. Ediciones SM, Madrid.
- Birmajer, M., et al., 2012. *Las otras Islas*. Antología. Prologue by Edgardo Esteban. Aguilar, Altea, Taurus, Alfaguara S.A., Buenos Aires.
- Brown, M., 2013. *Tito Puente*. Mambo King/Tito Puente Rey del Mambo. Illustrated by Rafael López. Harper Collins Publishers, California.
- Brown, M., Parra, J., 2011. *Waiting for the Biblioburro*. Tricycle Press, Canada.
- Buitrago, J., 2008. *Camino a casa*. Illustrated by Rafael Yockteng. Fondo de Cultura Económica, México.
- Bu(štos Mendoza), 2016. *Catalina, Diario de Un solo*. Tusquets, México.
- Castell, N., Pantoja, Ó., Camargo, F., 2017. *Borges, el Laberinto Infinito*. Rey Naranja Ediciones, Bogotá.
- With Her Illustrations Colasanti, M., 2015. *Mais de 100 histórias maravilhosas*. Global Editora, Sao Paulo.
- CONAFE, 1982. *Así cuentan y juegan en la Huasteca*. CONAFE, México.
- Cuevas Parra, P., 2002. *Relatos y andanzas. Historias de niños y niñas de los pueblos originarios de Chile*. Unicef/World Vision/Gobierno de Chile, Santiago de Chile.
- Dario, R., 1979. *Margarita*. Illustrated by Monika Doppert. Ekaré, Caracas. First published in 1907.
- de la Barquera, S., Wenceslao, J. (Eds.), 1813. *El Correo de los Niños*. Papel periódico sobre educación física, moral, civil y literaria. Imprenta de Juan Bautista de Arizpe, México.
- Author and Illustrator Diniz, A., 2011. *Photos by Mauricio Horta. Morro de Favela*. Edições Polvo, Lisbon.
- Dorfman, A., Mattelart, A., 1971. *Para leer al Pato Donald*. Ediciones Universitarias, Valparaíso.
- Engle, M., 2014. *Silver People: Voices Form the Panama Canal*. Harcourt Brace & Co., San Diego.
- Fernández de Lizardi, J.J., 1816. *El Periquillo Sarmiento*. Oficina de Alexandro Valdés, Mexico City.
- Fink, N., 2017. *Alfonsina Storni para niñas y niños*. Illustrated by Pitú Saá. Chirimbote, Buenos Aires.
- Frías, H., 1901. *Biblioteca del Niño Mexicano*. Maucii Hermanos, Mexico City.
- Gaviria, P., 2018. *Virus Tropical*. Sexto Piso, Mexico City.
- Gutiérrez Manguel, J., 1947. *Cocorí*. Illustrated by Coré (Mario Silva Ossa). Rapa Nui, Santiago.
- Iturralde, E., 2001. *Y su corazón escapó para convertirse en pájaro*. Illustrations by Santiago González. Loqueleto, Quito.
- Iturralde, E., 2013. *Los hermanos que cosechaban cuentos de hadas*. Illustrations by Roger Icaza. Alfaguara, Quito.
- Jackson, W.M., Thompson, H., Mee, A., Finley, J.H. (Eds.), 1912. *The Book of Knowledge*. The Grolier Society/London: The Educational book Co, New York.
- Keselman, P., 2006. *Gabriela*. With Her Illustrations. ¡Él empezó!. Ediciones SM, Buenos Aires.
- Leñero, C., 2001. *¡Es una travesía esa raya!*. Illustrated by Luis Fernando Enríquez Fondo de Cultura Económica, México.
- Lispector, C., 1978. *Quase de verdade*. Illustrated by Carla Iresta. Editora Rocco, Rio de Janeiro.
- Llanes, J.M., 2006. *Paquelé*. Illustrated by Julián Roldán Buenos Aires. Ediciones del Eclipse.

- Lobato, M., 2019. Sitio do picapao amarelo. In: A Chave do Tamanho. Projeto Livro Livre. Iba Mendes, Editor Digital, Sao Paulo. Originally published in 1941.
- Lyra, C., 1984. Cuentos de mi Tía Panchita. Editora Universitaria Centroamericana, San José. Originally published in 1920.
- Martinelli, A., 2014. Tortas fritas de polenta. Illustrated by Adolfo Guillero Bayúgar. Edicions de Ponent, Alicante.
- Menchú Tum, R., Liaño, D., 2016. El legado secreto. Illustrated by María Magaña. Editorial Universidad de Guadalajara, Guadalajara.
- Mistral, G., 2020. Rondas. In: Obra reunida. Selection and Research Gustavo Barrera Calderón/Carlos Decap Fernández/Jaime Quezada Ruiz/Magda Sepúlveda Eriz, vol. II. Ediciones Biblioteca Nacional, Biblioteca Nacional de Chile, Santiago, pp. 79–101.
- Neruda, P., 1950. Canto General. Illustrated by Diego Rivera and David Alfaro Siqueiros. Talleres Gráficos de la Nación, México.
- Ortega, F., 2016. Mocha Dick. Illustrated by Gonzalo Martínez. Planeta Comic, Santiago.
- Pombo, R., 1867. Cuentos pintados. D. Appleton and Co., New York. Published online as Cuentos Pintados y Cuentos Morales, preface by Beatriz Helena Robledo, Bogota: Ministerio de Cultura/Biblioteca Nacional de Colombia, 2015.
- Ramírez, S., 2013. La jirafa embarazada. Fundación Editorial Libros para Niños, Managua.
- Reyes, C., Elgueta, R., 2015. Los años de Allende. Editorial Hueders, Santiago.
- Roa Bastos, A., 1976. El Pollito en el Fuego. Illustrated by Juan Marchesi. Ediciones La Flor, Buenos Aires.
- Rubio, C., 1993. Pedro y su teatrino maravilloso. Illustrated by Vicky Ramos. Editorial Costa Rica, San José.
- Salarrué, 1945. Cuentos de cipotes. Editorial Nosotros, San Salvador.
- Tonatiuh, D., 2013. Pancho Rabbit and the Coyote: A Migrant's Tale. Harry N. Abrams, New York.
- Ueno, H., 2018. Los samurais de México. La verdadera historia de los primeros inmigrantes japoneses en Latinoamérica. Illustrated by Shibuya. Kyoto International Manga Museum, Kyoto and Mexico City.
- Vallejo, C., 1981. Pablo Yunque. Illustrated by Pablo Picasso. Editorial Girándula, Tegucigalpa, Honduras. Written in 1931 and posthumously published in 1951.
- Vargas Dulché, Y., Memín Pinguín. Illustrated by Alberto Cabrera. Mexico: Editorial Panamericana. From 5-X-1943 to 2014, approximately.
- Vasconcelos, J., 1924. "Prologue" in several authors. Illustrated by Roberto Montengro and Gabriel Fernández Ledesma. In: Lecturas Clásicas para Niños. Secretaría de Educación Pública, México.
- Venegas, D., 1999. La aventuras de Don Chipote o cuando los pericos mamen. Introduction by Nicolás Kanellos. Arte Público Press, Houston, Texas.

Brief selection of relevant childrens literature studies

- Amarilha, M., 2013. Alice que nao for au pais das maravillas: educar para leer ficcao na escola. Livraria da Fisica, Sao Paulo.
- Andruetto, A.M., 10 agosto 2007. Algunas cuestiones en torno al canon. Imaginaria. Revista quincenal sobre literatura infantil y juvenil (217).
- Bravo-Villasante, C., 1988. Historia y antología de la literatura infantil universal, tomo IV. Miñón, Valladolid.
- Bravo-Villasante, C., Antología e historia de la literatura infantil iberoamericana.
- Cabel, J., 1984. Literatura infantil y juvenil en Nuestra América: análisis y crítica. Centro de Investigación de la Literatura Infantil y Juvenil del Perú, Lima.
- Coelho, N.N., 1995. Dicionario Crítico de Literatura Infantil e Juvenil Brasileira. Secolos XIX e XX, 4ª ed. Editora de Universidade de Sao Paulo. first ed., 1983.
- Cordorés, M.T., Santos, A.C.de S., 2018. A contribuição da literatura infantil ao desenvolvimento da criança: um estudo de caso o Projeto Literatura de Biblioteca do SESC DOCA. Brasília II (2), 410–430.
- Cunha, M.A.A., 2006. Literatura infantil: teoria e pratica. Atica, Sao Paulo.
- Díaz, F.H., 2020. Sombras, censuras y tabús en los libros infantiles. Ediciones de la Universidad de Castilla-La Mancha, Cuenca.
- García Padrino, J., 2010. Gran diccionario de autores latinoamericanos de literatura infantil y juvenil. Fundación SM, Madrid.
- Garraón, A., 2004. Historia portátil de la literatura infantil. Alianza Editorial Mexican/SEP, México.
- González, A., 2018. Postcolonial Approaches to Latin American Children's Literatures. Routledge.
- Itzcovich, S.R., 1995. Veinte años no es nada: La literatura y la cultura para niños vista desde el periodismo. Ediciones Colihue, Buenos Aires.
- Lozano Alvarado, S., 2014. Revistas de literatura infantil en el Perú. In Crescendo 5 (1), 113–130.
- Rodríguez, A.O., 1994. Panorama histórico de la literatura infantil en América Latina y el Caribe. CERLALC, Bogota.
- Scorer, J., 2020. Comics Beyond the Page in Latin America. University College London, London.

Discovered and reconstructed: a perspective of Chinese children's literature over the last 100 years

Wei Kan, Faculty of Education, Beijing Normal University, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	662
Discovering the children: children's literature at beginning of 20th century	662
Politicalization of children's images: children's literature between 1950s and 1970s	664
Back to childhood and recreation: children's literature since 1980s	665
Multiple perspectives on understanding Chinese children's literature	666
Explorations in visualization in children's literature	667
Conclusions: popular and unique characteristics of Chinese realistic children's literature	669
References	670

Glossary

May Fourth Movement The May Fourth Movement of 1919 was a nationalist mass movement, considered as the beginning of Contemporary history, including in the field of literature. During the May Fourth era (1915–1925), Chinese intellectuals waged the New Culture Movement, questioning the relevance and validity of the Confucian tradition while searching for new ideas from abroad. Marxism and Communism—together with liberalism, anarchism, pragmatism, Bergsonism, etc.—entered China. Given its multifarious and multileveled significance, the May Fourth Movement sparked interest among scholars of different politico-cultural backgrounds, almost immediately from its occurrence into the early 21st century

Feng Zikai Chinese Children Picture Book Award is the first international Chinese children's picture book award. The award aims to honor writers and illustrators for their creation of distinctive Chinese children's picture books, to encourage publishers to publish original Chinese children's picture books, and to engage the public in valuing and reading original Chinese children's books

"Literature is a vehicle for conveying truth" (Wen-yi-zai-dao) This is the key of understanding Chinese literature and children's literature. This sentence was summarized by Zhou Dunyi (1017–1073), a philosopher of Song Dynasty: It means the text of a book, story or a poem is like a car that carries goods to their destination. Therefore, texts must serve to carry the doctrine; the ancient Chinese literary tradition emphasizes this focus on reality

Young Pioneers The Chinese Young Pioneers (referred to as "Young Pioneers") is an organization for Chinese children. It is an organization for children to learn about communism. All children whose age are 7 or 8 should join in the organization. The Communist Party of China entrusts the Communist Youth League of China (as CYLC) to directly lead the Chinese Young Pioneers. The Chinese Young Pioneers hold the Young Pioneers Congress (referred to as the "Youth Congress") every year, and elect new leaders of the large, medium and small teams. The slogan of the Young Pioneers is: honesty, bravery, liveliness and unity

Introduction

In the view of most Chinese scholars (e.g., Huang, 2021; Wang, 2009), children's literature in China emphasizes understanding through literary works, such as novels, prose and other works created, adapted, translated and published by adults for children's development. Literature plays the role of promoting children's spiritual life and esthetic development, shaping their ideology according to moral education, and transmitting knowledge. This perspective suggests that children's literature should be a multi-coexisting subject system that includes not only creating literature in a variety of styles and genres, but also draws on literary means to transform all kinds of knowledge, connecting children to vivid representations of the world that foster both understanding and feeling.

Discovering the children: children's literature at beginning of 20th century

There are ongoing debates about the origins of Chinese children's literature. Most researchers think that the birth of Chinese children's literature is best viewed against the background of China's modernization, which aimed to facilitate social progress by

importing Western technology, knowledge and culture. Elite intellectuals made great efforts to borrow from Western culture and literature so as to enlighten the Chinese public, with a special focus on children (Chen, 2013).

When did children's literature begin in ancient China? While there is no consensus, there are several, and some controversial, arguments about the origin of literature for children. Some researchers in the field of children's literature believe that Chinese children's literature originated in ancient China; for instance Wang (1999) points out that the collection *Searching for God [Sou Shenji]*, which was written in the 4th century by the historiographer Gan Bao, includes a chapter that was deemed "the most laudable and famous fairy tale" (Wang, 1999). However, the relatively limited view of children and childhood in Ancient China stifled the development of children's independent position in society, and the shortcomings of early schooling restricted the appearance of children's literature. Therefore, the role and status of children's literature in educational history had not been highlighted (Zhu, 1998). Children's literature had not yet become a kind of conscious literary category, and the related production of literature, development of theoretical criticism and the writing of literary history had not been established. This was the status quo without a substantial breakthrough until the beginning of 20 century. As Mao Dun (1896–1981), the famous modern Chinese writer, noted: "The terminology of children's literature began to be used only since 1919" (Mao, 1993:147). With the "discovery" of "children" during the turn of the 19th and 20th centuries, literary works specially created for children came into being; Chinese children literature was born from this moment.

In terms of the characteristics of children's literature in China, as contemporary educator Gu (2018) argues, Chinese traditional culture has to be scrutinized with an understanding of the way Confucianism provided the ideological foundation and value orientation, although Chinese traditional culture includes not only Confucianism, but also Taoism, Legalism, Mohism, etc. Under the influence of the Chinese cultural spirit and the psychological structure of national culture, Chinese children's literature has an important tradition – it represented the view that literature is the carrier of morality and ideology ("Wen-yi-zai-dao"). Therefore, literature with conveying morality (Zhou, Dunyi, The Confucian of Song Dynasty, 11 century) has deeply impacted children's readings, emphasizing that all kinds of text must serve as the carriers of doctrine since ancient China. Hence, the ancient Chinese literary tradition has also deeply impacted contemporary Chinese children literature, and naturally has had a profound influence on modern literature, and on the creation of Chinese children's literature in the 20th century (Wang, 2009). Modern Chinese children's literature writers seem to have an innate sense of mission, social responsibility, and a sense of "using the pen as the banner" as quoted by poet Ai, Qing, 1910–1996 (Ai, date, page) to serve the nation.

The birth of the category of Chinese children's literature is best viewed against the background of China's modernization, which aimed to facilitate social progress by importing Western technology, knowledge and culture. Elite intellectuals made great efforts to borrow from Western culture and literature so as to enlighten the Chinese public, with a special focus on children. Influenced by the literature of Japan and modern European literature, children in China had been begun to be recognized by writers and society at large (Wu, 2011, 2012). At the same time, Chinese children's literature began to look very different from the long historical period that came before. In this initial development of Chinese children literature, new literary qualities of modernity gradually emerged, which are intricately intertwined with a new cultural form. From the May Fourth Movement in 1919 to the present, Chinese children's literature has gone through a century of exploring its own way and it has demonstrated the ideological goal of cultivating its young people; that is, the 100 years of children's literature development has always been closely related to the destiny of the nation-state. The May Fourth Movement of 1919 not only opened up modern Chinese literature, but also the cradle of Chinese children's literature. As noted above, these beginnings of Chinese children's literature began with the discovery of "children"—those who had previously been considered "little adults".

As Wang (2016), one of the contemporary scholars in the field of children's literature argues, unlike western children's literature that takes fantasy as the basic foundation, Chinese children's literature has a distinct overall characteristic that runs through it since its birth; that is, the connection between children's literature and reality. In this sense, the history of Chinese children's literature is a history of realistic literature (Wang, 2016). This kind of literature is manifested in the description, reflection, judgment and imagination of reality, in the pursuit of realistic and vivid artistic effects, and by a focus on the cognitive and enlightening functions of literature. Moreover, as Wang points out, in a sense the development history of a children's literature is the evolution of the history of adults' view of children. In this way, children's literature texts, influenced by a particular view of childhood, will invariably include cultural assumptions that reflect the characteristics of the larger culture. In the 1980s, academic circles in China reached a consensus that children, or cultural assumptions about childhood, should be the basis for the study of children's literature; for example, Zhu (1998) argues that "the view of children should be the origin of children's literature" (p. 98). Similarly, Fang emphasizes childhood as the logical starting point of children's literature theory (Fang, 1993), and Wang argues that understanding of children is at the base of the esthetic origins of children's literature (Wang, 1999). Chinese children's literature, then, hews close to reality, with representative works rooted in the Chinese social and cultural context.

It was not until the 1920s that the Chinese people's emphasis on children's literature underwent a fundamental change and transformation. There are two very important articles, namely *How We Be Fathers Now* written by the famous writer Lu (1919) and *Children's Literature* written by Zhou Zuoren (1885–1967) in 1920. The publication of *Children's Literature* (Zhou, 1920) is undoubtedly a landmark, which first proposed the concept of "children's literature" in Chinese literature history. Since then, the attention to children's literature had been gradually expanded, and the amount of work identified as "children's literature" has become more culturally visible. Wei Shouyong and Zhou Houyu first proposed that children need their own literature in their work, *Introduction to Children's Literature* (Wei and Zhou, 1923). Influenced by John Dewey's child-centrism, they convened a group of teachers and created a large number of children's literature education books for young children and primary school students. Meanwhile, *Children's Literature Studies* (Zhu, 1924), *ABC of Children's Literature* (1929), *Literature on Children's Literature* (Zhou,

1932), *Researches on Children's Stories* (Chen, 1932), and other works on the definition, classification, positioning, and teaching methods of children's literature research and thinking provided preliminary theoretical bases for the construction of the history of Chinese children's literature at the beginning of the 20th century.

Along with theoretical explorations, representative works illustrate the vibrant creativity in the field of children's literature. *Scarecrow* by Ye Shengtao (1894–1988), published by the Commercial Press in 1923, was the first collection of short Chinese children fairy tales in China. The scarecrow in the story sees the grains of the old peasant women's labor being devoured by locusts, the fishermen's children freezing and starving on the boat, and the lower-class women who had the miserable fate of throwing herself into the river to commit suicide. All this was seen and heard by the Scarecrow, but he could neither shout nor run to save their lives. The only thing he could do was to wave his broken fan, which required the external force of the wind to move. This children's fairy tale deeply reflects the anxiety and pain in the Scarecrow's inner heart. He saw and heard these miseries, but the scarecrow couldn't do anything. The final outcome could only be one of helplessness and painfulness and to fall down to the ground. Scarecrow is a symbolic children's fairy tale character, who was regarded as providing a model for the creation of Chinese fairy tales, which reflected the unique realism in children's literature.

After Ye published his work, the second representative writer of Chinese original fairy tales was Zhang Tianyi (1906–1985). Zhang inherited and carried forward Ye Shengtao's original tradition of children's literature, pushing the spirit of realism to a new height. He directly writes about social changes, especially realistic social relations and struggles, in his works. Zhang Tianyi created two genial characters, Dalin and Xiaolin (Zhang, 1934). *Da Lin and Xiao Lin* tells the story of Da Lin and Xiao Lin's fantastical experiences to reveal the reality of life under social conflict. This long children's story, published in 1932, reflects Zhang Tianyi's view of children's literature as a depiction of the real world; Da Lin and Xiao Lin were fictional characters, but they have a deep basis in reality. The work takes the different fates of the two brothers as a clue, and through their experiences, they link the life of the upper and lower social strata and the contradictions and struggles of the two classes, conveying a sharp social criticism. Zhang Tianyi adds many fantasy elements and absurd plots, such as monsters that can turn child laborers into eggs and eat them, which makes the cartoonish depiction more interesting to children and fully demonstrates the author's fairy tale approach. It is a classic among the Chinese children fairy tales, and became the most popular fairy tale in the period of 1930s and 1940s. From the point of view of artistic innovation, his works with their exaggerated characters and witty language indicated that the era of Chinese fairy tales had begun to mature.

The best-selling collection of children's literature around this time is Bing Xin's (primitive name Xie Wanying, 1900–1999) *To Little Readers* (Bing, 1941) which is a collection of her letters written from 1923 to 1926. This book was published in a single volume from 1926 to 1941, and it has been reprinted 36 times. Besides this, the other most influential children's publication in this period was *Children's World*, which was founded and edited by Zheng Zhenduo in January 1922 and published by the Commercial Press. Bing Xin and Zheng Zhenduo created stories which idolized and eulogized the child—narratives in which the child may even have the power to save the (despairing) man. The pure child also took its place within the discovery of formerly marginalized figures. The child figure retained both an abstract quality and an allegiance to society's victims; an often mute, pathetic companion to the adult whose suffering was the narrative's primary concern.

The choice of realism in Chinese children's literature is not accidental, but an inevitable choice of the norms of the times and the evolution of Chinese society. From the very beginning, modern Chinese children's literature has been confronted with a very serious problem; that is, 20th century China was a disaster-ridden country.

It is in such a social background and cultural context that the Chinese modern children's literature was generated. "Enlightenment, national salvation, revolution, and war" are the panoramic perspective and key words for observing the century-old modern Chinese society.

Politicalization of children's images: children's literature between 1950s and 1970s

After the founding of Peoples' Republic of China, Chinese society entered into the collective action of building a new China. In the peaceful period of the 1950s and 1960s, the creation of children's literature might have turned to fantasy, poetic, and romantic literature. However, Chinese children's literature in the 1950s and 1960s was still dominated by realism with adults' interests and will as the interests and will of the children. Children's literature at that time mainly reflected realistic political issues and social expectations, which were related to the understanding of childhood at that time. This includes a certain pragmatism or even utilitarianism in the way literature was viewed at that time, which naturally also affected children's literature. Pragmatism, utilitarianism and children's literature are closely related. Under these circumstances, children were considered to be the successors of the new Communist Party, and children's literature writers and originators were obliged to cultivate these successors. For instance, during this period, the "red classics" of children's literature were produced. The representative works of children's literature were *The Little Hero Yulai* (Guan, 1955), *The Sparkling Red Star* (Li, 1972) and *Little Soldier Zhang Ga* (Xu, 1968). The children portrayed in these works were witty and brave, participated in the struggle of the turbulent times of war, and reflected the living conditions of children in the war years.

At the same time, the creation of children's literature in the 1950s and 1960s was linked to the work of the Young Pioneers, a youth organization for children run by the Communist Youth League founded in 1949. Children's literature authors cooperated with the tasks of the Young Pioneers and expressed the life of children in the organization. There were also a number of excellent children's literature works in the literature of the Young Pioneers. For instance, the novel *The Story of Luo Wenying* (Zhang, 1955), written by Zhang Tianyi in the 1950s, featured images of soldiers and counselors of the Young Pioneers. Meanwhile, many poems

mainly expressed the life of the Young Pioneers, such as "Annoyance of the little Captain" (Jin, 1955). In this way children's literature had become a tool to serve the Young Pioneers during 1950s–1970s.

It was under this pressure of the political context that researchers of children's literature published their works. Examples include Chen Bochui's work, *A Brief Discussion of Children's Literature* (Chen, 1957), Fang Jisheng's study, *Trial Discussion on Children's Literature* (Fang, 1957), Jiang Feng's *Speech on Chinese Children's Literature* (Jiang, 1960), He Yi's *Essays on Children's Literature* (He, 1960) and Lu Bing's work, *Literature for Educating Children* (Lu, 1962). These theoretical works continually examined the questions: What is children's literature?, How should children's literature be categorized?, and How should we teach children's literature?, among other issues, all infused with a strong sense of social responsibility.

In the 1950s and 1960, although Chinese children's literature was endowed with new characters and requirements on shaping children's images which served political purposes, it was still moving forward, forming a flourishing period. At this time, the creation of children's science fiction literature entered a short-lived but productive period. Influenced by science fiction literature of the Soviet Union, science fiction and science literature were regarded as a means to popularize scientific knowledge to children in China. Examples are the well-known *Little Swallow's Million Miles' Flight* (Qin, 1950), *Our Soil Mother* (Gao, 1955) and *Little Tadpole Looking for Mother* (Fang and Sheng, 1960). At the same time, there was increased interest in translation and the introduction of Soviet children's literature theory. In the first 20 years of the founding of the People's Republic of China, not including the Soviet children's literature papers scattered in various academic journals and periodicals, more than 25 types of children's literature studies were compiled and printed.

The other development in children's literature was its role in the reshaping of images of children in education. Some representative works of this time include, *Magic Pen Ma Liang* (Hong, 1956), *The Secret of the Treasure Gourd* (Zhang, 1956), and *The Harbor of Next Sailing* (Yang, 1958). *The Secret of the Treasure Gourd* tells the story of Wang Bao's fantasy that he has a treasure gourd and can obtain anything, but this treasure also brings him pain. So he throws away the treasure gourd and later finds out that it was a dream. He therefore corrected his shortcomings and studied hard. This work opens up a new way of writing fairy tales; it creates both the fantasy world and the real world, and satisfactorily solves the transition and connection between the two worlds. *The Harbor of Next Sailing* describes Tang Xiaoxi's peculiar experience in the port of "The Next Ship", which is cleverly conceived. The author uses the idea of time, seizing the "next time" excuse that children often use to waste time and creates a fairy tale world where there is no time and hour and, for children, it is a place "happier than happiness". But it's also a place where bad guys go about their ways. Yan Wenjing pays attention to fictional images from the perspective of children, and integrates game-like elements into the characters, so that the fairy tale image is suitable for children.

However, from the mid-1960s to the late 1970s, the huge political catastrophe of the so-called *Cultural Revolution* (1966–1976) resulted in all works in the field of children's literature and most writers and theorists being criticized, and the creation of children's literature ceased. Questions such as: what kind of characters should children's literature create?, what kind of children's personality should be cultivated?, what is the social function of children's literature?, should children's literature expose the dark side?, as well as whether children's literature should reflect the political struggle of the people, that been discussed earlier, were now overwhelmed by strong political criticism. Critics were not only unable to pose these exploratory questions, but were deprived of their right to defend themselves.

Back to childhood and recreation: children's literature since 1980s

After the end of the Cultural Revolution, children's literature magazines such as *Children's Literature* began to publish again. The creation of children's literature also became active after the reform and opening up of the country. With the deepening of ideological emancipation, the authors began to pay attention to the esthetic and entertainment value of fairy tales themselves, and penetrated into the inner world of children. They broke the shackles of the formulaic and conceptualization, and brought fairy tale creation into a new stage. At the beginning of the discussion, "child-centered" was still regarded as Deweyian theory, which had been vehemently criticized during 1950s–1970s. Due to the constraints and limitations resulting from closure of the state, there were no further in-depth studies in this progressive tradition. Since 1970s, some researchers have argued that the function of children's literature was to serve educational instruction, rather than just providing the literary works (Lu, 1979). A few years later, Cao Wenxuan proposed that *Children's literature is literature* (Cao, 2012), which represented the aspirations of many young writers and scholars to focus on children's literature. With time, more researchers and theorists also joined this way of thinking. The younger generation of writers who had experienced political turmoil now started to rethink the essentiality of children's literature. A group of young fairy tale writers emerged in the 1980s, holding high the esthetic banner of artistic innovation. These writers have become the backbone of fairy tale creation in the new era, and have exerted extensive influence on children. Representative works include Zhang Zhi's *Thunderbolt Beibei* (Zheng, 1986), Zheng Yuanjie's *Pipiru* and *Lu Xixi* (Zheng, 1989), Zhou Rui's *Second General*, Bing Bo's *Swing, Swing*, and Tang Sulan's *Stupid Wolf Story*. Most of these works respect children's personality development and focus on cultivating children's awareness of independent thinking, which is a great leap. The fairy tale works return to the nature of literature, and turn their attention to the family life and school that children are familiar with.

Hence, the 1980s opened the era of rewriting the history of children's literature. The children's literature history was now seen as worthy of recognition in the collection and arrangement of historical materials. Since then, the writing of the history of Chinese children's literature has gradually had a strong developmental trend. For instance, books such as *History of Contemporary Chinese Children's Literature* (Chen, 1989), and *Essays on Children's Literature* (Jiang, 1979) were published. Wang Quangen published his

work, *The Main Trend of Modern Chinese Children's Literature* (Wang, 2000). In the same year, Zhu Ziqiang's work, *Chinese Children's Literature and the Process of Modernization* (Zhu, 2000), triggered further discussions. In addition, Liu Xuyuan's, *A Brief History of Chinese Children's Literature 1916–1977* (Liu, 2013), and Fang Weiping's work *Four Decades of Chinese Children's Literature: A Historical Overview* (Fang, 2018) had a huge influence in the field of literature. From the conception of "children's literature" to the writing of the history of children's literature, scholars have re-focused on new problems and new phenomena in the developmental history of Chinese children's literature over the past century. The current studies have also included many achievements in the writing of children's literature history, the sorting out of children's literature trends and genres, and the elaboration of children's literature creation. As Fang argues, children's literature is a special form of literature; specifically, its uniqueness lies in the fact that it is mainly created by adults and the recipients are children. This constitutes the inherent contradictions and puzzles of children's literature: the cultural generation gap and the intergenerational conflicts can impede literary acceptance. Indeed, the key issue of children's literature is how children are viewed (Fang, 2018; see also Rose, 1984 for a similar argument in the West).

Since the reform and opening-up policy implemented during 1980s, along with the reflection and review of "childhood theory," the concept of "child-oriented" has taken a foothold in children's literature research. Chinese children's literature researchers began to explore children's literature theories with a holistic perspective to comprehensively understand children's literature. Recent studies reflect many new, deeply discussed themes, such as a Chinese-style children's literature theoretical framework, comparative children's literature, the study of fairy tale art space, the study of children's esthetic instruction, and the study of the narrative form of children's fiction. Moreover, there is a greater acceptance of children's literature as literature, and increased interest in the study of the motifs of children's literature (Zhang, 2006). These multi-faceted and multi-layered theoretical explorations have greatly expanded the theoretical space of Chinese children's literature. The continuous expansion of theoretical boundaries has also brought diversification of research methods. Particularly, recent researchers have borrowed from the fields of esthetics, anthropology, psychology, ethics and other disciplines to read and think about the works, embodying a strong esthetic and academic rationality (Wang, 2000; Liu, 2013). In general, the study of children's literature during this period has made considerable progress, and both the system and value of the discipline have strengthened and become more accepted. The theoretical research of Chinese children's literature realized the modernization of its research paradigm (Zhang, 2006).

Multiple perspectives on understanding Chinese children's literature

China is a vast country with abundant resources, and has nurtured a diverse and rich regional culture. Since the beginning of the 21st century, a vibrant area in academic research on children's literature has focused on regional culture and regional human factors, exploring the various aspects of the national characteristics of children's literature. From the perspective of regional culture, a series of children's literature histories with regional cultural characteristics have emerged. For example, *Re-discussions on Inner Mongolia Children's Literature* (Zhang, 2018) investigates children's literature in the inner Mongolia region. Another book, *Analysis of the Development of Anhui Children's Literature*, describes children's poems and poets in Anhui province (Liu, 2004). Other examples include Sun Hailang's study, *Review of Jiangxi Children's Literature in the New Era* (Sun, 2004), Zhang Zhongyi's research, *Observation of Henan Children's Literature* (Zhang, 1952), and Chen Zidian's book, *Introduction to Guangdong Contemporary Children's Literature* (Chen, 2005), that concentrate on Jiangxi, Henan and Guangdong area in the field of children's literature respectively. Clearly, the role of such regional children's literature history writing is to reconstruct the entire Chinese children's literature map with respect to regional perspectives.

Researchers of children's literature in China have also expanded their horizons to the field of children's literature created by ethnic minorities. Minority children's literature is a very important part of the treasure house of Chinese culture and oriental civilization. Ethnic children's literature embodies the ardent expectations and hopes of the elders of all ethnic groups for the new generation, while the traditional national virtues and the essence of national culture are naturally hidden deep within it. In the progress of history and the development of the times, national children's literature has become a unique accumulation and unique symbol of national culture. Zhang's research brings brushstrokes to the ways the villages and wastelands inhabited by ethnic minorities explore the unique spiritual representation and artistic form of ethnic minority children's literature (Zhang, 1952, 2018). In addition, Zawa's work, *Introduction to Tibetan Children's Literature* (Zawa, 2005) focuses on the "ethnicity" characteristics of Tibet minority children's literature, and organically integrates the psychological qualities of minority children and ethnic groups with the development trend of the times. As a whole, these studies enrich the nationality and folklore character of Chinese children's literature.

The tentacles of expansion of Chinese children's literature history have also extended to Hong Kong, Macao and Taiwan, which are integral to "Greater China" children's literature. *One Hundred Years of Hong Kong Children's Dreams (1920s–1950s)* edited by Zhou Mimi and *One Hundred Years of Hong Kong's Children's Dreams (1960s–1990s)* (Zhou, 1996) describe the basic aspects of Hong Kong children's literature and the scattered works of authors. The introduction also analyzes the progress of the history of children's literature in Hong Kong. Hong Wenqiong's study, *Taiwan Children's Literature History* (Hong, 1990) more clearly outlines the development of Taiwanese children's literature at different stages, and also explains the important works of writers in the form of special topics; for instance, Zhang Yongjian's work *The History of Chinese Children's Literature in the 20th Century* (Zhang, 2006) includes a special chapter, "Children's Literature in Taiwan, Hong Kong and Macau," which examines its unique historical development track and characteristics in more detail by means of overview and analysis of the writers' works. Regrettably, when outlining the development of children's literature in Hong Kong, Macao and Taiwan, the above researchers failed to analyze it in the context of the overall development of Chinese children's literature, resulting in a situation of separation and isolation.

Explorations in visualization in children's literation

As early as the Ming (1368–1644) and Qing dynasties (1636–1912), children's literary works with harmonious pictures and texts had appeared (Shen, 2012). In the early 20th century, some Chinese scholars had a deeper understanding of the relationship between picture stories and children's literature:

Pictures can sometimes form a story continuously, such as pictures in serial stories, and pictures in movies, etc. Words do the same thing. Pictures are an excellent way to use pictures to motivate children to appreciate literature.

Li (1962: 112).

In the 1920s, these picture stories became the new favorite in children's books, which had developed by leaps and bounds. While paying attention to children's literature, the May Fourth literati also attached great importance to the burgeoning picture stories in children's books, and many of the great writers also worked in the creative team of picture stories.

Zheng Zhenduo (1898–1958), who was a famous writer and poet, coined "picture story" as the new terminology for the journal *Children's World* in the first volume published in 1932. Meanwhile, Li Jinhui (1891–1967) writer of children's song and dance music, Xie Liuyi (1898–1945) writer and interpreter, Xu Dunggu (1892–1983) modern painter, and other artists and intellectuals made up picture stories for children. For instance, modern writer and educator Zhao Jingshen (1895–1985) wrote three representative picture stories, such as "A Pea," "Sister Mouse," and "To the Lilliputian Country" (Zhao, 1962). In his collected work, Zhao published his pioneering research *Theory on Children's Picture Story Theory* (Zhao, 1962).

At the same time, a representative figure who was regarded as one of a generation of masters of Chinese children's literature, Chen Bochui (1906–1997), wrote the article "Children's Literature Must Prosper and Develop" in his work, *Brief Discussion of Children's Literature* (Chen, 1982). In it, he stated:

Picture-story books are of course not 'decorating the facade' or helping to explain the content for children's learning, but as the subject to express ideas, which directly appeal to young children's senses more vividly than words. This can also be regarded as a feature of young children's literature.

Chen (1982: 79).

However, due to regressive notions of social development, the germination of children's picture-story books did not gain proper recognition again in China until the founding of the Republic of China. During 1950s and 1960s posters were made that were concise, expressive and stimulating; they were applied as a visual method to meet the need of political mobilization. Painters employed various political symbols in a unique way to express the ideology of the government. Deeply impacted by the genre, Chinese creators absorbed elements from Soviet posters and combined it with traditional Chinese New Year pictures, creating a large number of propaganda images with children as the theme. In the vast rural areas where educational resources were limited, children were encouraged to understand national policies, carry out health and disease prevention, and complete literacy tasks based on understanding these posters on the walls of schools and houses (Jiang and An, 2018). Given visual imagery is a powerful communication medium, children's propaganda posters had been centered on political tasks. For the state and society, there were no more important visual methods than the design of visual images representing the state and class to influence the psychological identity of children, and thus to generate cohesion to maintain the stability of power.

In contrast to the focus on children's propaganda, the 1980s saw a large number of art professionals devoting themselves to the creation of picture-story books. At the same time, the Chinese children literature was going through a period of great leaps forward, resulting in many excellent works of picture-story books for children; in fact, the publishing of picture-story books for children now flourished in unprecedented ways. In the 1980s, the publishing of children's picture-story books reached a peak in terms of variety, print volumes, and quality, which was the first developmental peak since the founding of New China. According to statistics, in 1980 more than 1000 topics of children's picture-story books and more than 400 million copies were published nationwide, and there were more than 2400 kinds, with more than 630 million copies just 2 years later (Chinese Publishing Almanac, 1983).

Influenced by the media and popular culture, the absorption and re-creation of popular culture began influencing illustration creation at this stage. Early on, Chinese children's picture-story books mainly borrowed the genre of Japanese popular culture while, at the same time, many artists were looking for inspiration from the fonts, graphics, colors and other elements in the Chinese traditional paintings works to refine artistic symbols.

As the creative element of illustration, representative visual cultural symbols were widely used. Some examples include: the picture book, *Fox Spirit in the Wasteland*, adapted by Cai Gao and based on a novelist in Qing Dynasty; and Pu Songling's (1640–1715) work, *Strange Tales from a Liao-zhai* (Pu, 1982), that depicts the fairy tale as a story of a child saving his mother and fighting wits and courage with a fox who pretends to be a human being, reflecting the uniqueness of Chinese storytelling's traditional cultural spirit. This latter work uses black and gray as the basic colors, highlighting the mysterious mood of the story. In another example of the use of color to portray mood, the protagonist in the picture, *Baoer*, wears red clothes, implying the belief that *Baoer* will win; the strong use of red and black emphasize the contest between good and evil, while the interlaced use of green, purple, orange and other colors provides beats, much like musical notes. The ingenious use of colors in these works together with

the design of the clothes for the characters and the architectural modeling of folk printmaking reflect the use of cultural symbols with Chinese characteristics.

At the beginning of the 21st century, due to the stimulation of the introduction of contemporary children's picture books from western countries and Japan, China's original children's picture books also gained new popularity. Even in ancient China, there was no such thing called a "picture book," since the creative form of coexistence of pictures and texts were a longstanding tradition; that is, there had always been a tradition of emphasizing the conveying and ideographic functions of pictures and texts. This creative tradition of complementary pictures and texts laid the foundation for the development of contemporary children's picture books (Xiang, 2017).

At the beginning of the 20th century, the child-centered view of children and childhood and concepts of children's literature emerged in unison. Children's literature writers were clearly aware of the existence of children's readers, and began to consciously write for children themselves, and a large number of children's literature works were published. Illustrations played a key role in these children's literature works. For instance, in Gao Shiqi's (1905–1988) popular science fairy tales, entitled *The Autobiography of Fungus* (Gao, 2017), and *The Food, Clothing, Housing and Transportation of Bacteria and Our Heroes Against the Enemy* (Gao, 2017), were illustrated with anthropomorphic images of bacteria and white blood cells by the creators. These titles were vivid, interesting and appealing. Similarly, the modern Chinese painter and cartoonist Feng Zikai (1898–1975) drew illustrations in Ye Shengtao's work, *Kaiming National Language textbooks* (Ye, 2018). Ye was a pioneer of Chinese literature education—his illustrations enhanced the visual sense of the textbook and stimulated students' enthusiasm for learning.

From the 1920s to the 1930s, a large number of children's picture stories were published in children's magazines. The most famous series picture stories were magazine named "Children's World" founded by Zheng Zhenduo (1898–1958), literary historian of Chinese vernacular literature. Some scholars pointed out that the budding of Chinese children's picture books should be the "columns of picture story" in the magazine, *Children's World*, published in 1922. These columns were published in forty-six serialized graphic stories which were believed to have been illustrated by Zheng, including the stories such as "Hippo Kindergarten" and "The Adventure of Two Little Monkeys," and were seen as suitable for children's reading given the comic strip format. Since these picture books were published in *Children's World*, many children's books were published with picture stories as an important component, such as *Little Kids* published in 1923, *Children's Morning News* in 1932, and *Children on Time* in 1939. These picture stories were developed with developmental psychology in mind, focusing on children's true feelings, and were in line with children's interests. These titles laid the foundation for the creation of modern children's picture books (Xiong, 2018).

In the 1940s, Zhang Leping (1910–1992), one of the pioneering manhua (Chinese comic) artists, created a series of works for young readers, such as *San Mao Joining in Army* (Zhang, 1940) and *San Mao Wandering Ji* (Zhang, 1940), also in the form of comic strips. Each page contains four pictures without text, relying on the changes across the panels to convey the storyline. Subsequently, in the following 20 years, *Journey to the West*, was adapted from an ancient classic work and created by Zhang (1983). In addition, other comic books, including *Little Doctor* created by Xiuwen, and *West Chamber* created by Wang Shuhui, appeared successively in 1940s, where they used pictorial language to tell one ongoing story. This large-scale publishing of comics has improved the status of children's picture books, and has also cultivated a group of readers for children's picture books, while becoming more popular.

After the founding of New China, children's literature became a tool for educating children for societal development. Under the influence of national policies, the number of graphic stories has not only increased greatly, but they have also been separated from magazines and adapted into discreet volumes. During this period, a very important work was born, namely Fang Yiqun's *The Carrots are Back* (Fang, 2008), which tells a warm story of small animals sending love to each other on a snowy day. It uses a combination of text and pictures, closer in form to the basic form of modern children's picture books.

In the 1960s, Taiwan began to introduce foreign picture book titles, mainly from the Japanese, which provided new references and a development impetus for its local picture book creation. In 1976, Taiwan's influential children's artist Zheng Mingjin, who was known as the godfather of picture books in Taiwan area, introduced Japanese children's picture books and also introduced the term "picture book" to Taiwan. In 1986, he published *On Japanese Children's Picture Books* (Zheng, 1986), which deepened the local people's understanding of the concept of picture books. Gradually, the publishing houses in mainland China also accepted the new term "picture book" (*Huiben*, 绘本, copied from the Japanese) as a replacement to the traditional term "picture-story book" (*lianhuan hua*) since 2010s (see Glossary in this chapter, Picture Book (*huiben*)).

At the beginning of 21st century, mainland China ushered in reform and, in opening up to the outside world, economic strength continued to improve and the publishing industry also flourished. Chinese elements and traditional cultural symbols were used by the writers and artists when they were engaging in creating children's picture books. Zhu Deyong, the cartoonist who comes from the Taiwan area, as well as other authors, adopted the form of "bubble painting", which uses speech bubbles with simple and interesting text within the four-frame comics. This unique painting style had been fostered by the children's literature field. Since the 2010s, Fubin Liao, who writes under the pseudonym "Jimmy" and who also comes from Taiwan area, has published picture books in mainland China inspiring fellow writers in mainland China with titles such as *Go Left, Go Right* (Liao, 2018) and *Listen to Jimmy Sing* (Liao, 2019). As children's literature scholars have pointed out, Jimmy fully adopts the narrative strategy of children's views, borrowing the perspective of children to observe the living faces of adults, and examining the adult world rationally and tenderly. Jimmy's picture books care for children's ontological world, explore childhood, and draw on the dynamic nature of dreams (Lin, 2017).

The creation of Chinese children's literature has emerged as a new field in the 21st century with the advent of the picture-book. On the one hand, this makes up for the long-standing gap in the field of Chinese literary creation in preschool children's literature and, on the other hand, it is also the result of the further deepening of the understanding of children by the creators of children's literature since the new century. The creation of children's picture books has undertaken many missions, including cultural inheritance, national identity, and esthetic cultivation (Kan, 2019).

At the beginning of the 21st century, picture books gained strong appeal in the children's literature market. Attracted by huge market interests, famous Chinese publishing houses, such as *Nanhai Publishing House*, *Tomorrow Publishing House* and *21st Century Publishing House*, have increased their efforts to introduce foreign picture books. Many excellent foreign picture books, including *The Tale of Peter Rabbit* (1901) written by Beatrix Potter from Britain, *The Cat That Lives a Million Times* (1998) by Yoko Sano from Japan, and *The Giving Tree* (1964) by Shel Silverstein, entered and blossomed in the Chinese market. The introduction of a large number of excellent children's picture books from Taiwan and abroad has not only enriched China's children's book market, but also opened a window for Chinese children's picture book creation, learning from the experience the Taiwanese and from foreign writers and illustrators (Hao, 2009).

In order to carry forward Chinese traditional culture and promote Chinese children's understanding of these images in picture books, the creators have accelerated their exploration of original Chinese children's picture books, resulting in marked achievements in the exploration of original picture books. In 2002, Jiangsu Children's Publishing House published the well-received series entitled *I'm Awesome: Picture Book for Children's Growth in Kindergarten*. From 2006 to 2008, Xiong Liang, Xiong Lei, Zhou Xiang, and Zhu Chengliang published many original books, which further sparked the trend of creating children's picture books. In 2009, the "Feng Zikai Picture Book Award" was founded and became a powerful encouragement in the further development and promotion of original picture books in China. In 2010, the Caterpillar Children's Book Center published the *Chinese Nursery Rhymes Series Picture Book*, which created a new period for the classic Chinese nursery rhymes. In 2012, Beijing Normal University Press took materials from traditional culture and launched the *Memory of China: Traditional Holiday Picture Book*. In 2013, People's Education Publishing House published the *Most Beautiful China Series Picture Book*, based on Chinese landscapes and customs. In 2015, Xinjiang Children's Publishing House launched the *Story of China Picture Book*, and in 2018, Chinese children's picture book creator Cao Wenxuan won the International Hans Christian Andersen Award. Under the efforts of the creators and promoters of Chinese original picture books, such works are becoming popular, influential and lucrative.

The difference between picture books and traditional children's reading materials in mainland China, as Peng (2012) argues, is that picture books convey the art of telling stories with drawings, which had been neglected by Chinese writers before the last century. The most important function of illustrations in picture books is to express the connotation of the story. The plots of the picture books are not only seen, but the pictures provide additional reading clues for children (Peng, 2012). The children are encouraged to imagine as they "read" the images, which convey both the storyline and the affective qualities. For a long time, the development of Chinese children's literature was not based on children's psychological and sociological features, but rather on the concerns of the adults' imagination and high expectations, which typically linked with political and social requirements. In contrast, the images and language of contemporary picture books, with the guidance of adults, helps children to recreate vivid and imaginative mental images for themselves. In the context of parent-child reading, the time and space created by the parent's voice and tone are valuable aspects of the reading process. Children not only hear the unique stories brought to life by the sounds, but also obtain meaning from the pictures. As Wang points out, is an indescribably beautiful process to acquire the stories from two different modes and experiences at the same time. Only in this way can the function of picture books be brought into full play; that is, on the basis of increasing children's sense of pictures, which can be transformed into a rich sense of language (Wang, 2011).

Conclusions: popular and unique characteristics of Chinese realistic children's literature

In the past 100 years, the development of Chinese children's literature shifted from a process of manipulating children to discovering children and finally to supporting children's development. In this process, children's literature was influenced by the impact of war and political influences. During 1930s, the children's literature was booming as a "political fairy tale", which was then transformed as the "children's literature in the war" during the World War II (1931–1945). Behind the transformation in children's literature was the view of children in the context of Chinese modern history, which resulted children's literature undergoing several twists and turns. Over the past more than 100 years, it has witnessed the struggle for the child-centered theory, the disputes between class and children, the debates among aims of education and literature, and the process of re-understanding children in culture in the 21st century. At the end of this 100-year exploration is the beginning of a return to the original purpose. In contemporary China, child-centered theory is still developing in schooling and family education; there has not yet been a Chinese child psychologist the likes of Piaget, nor a master of children's literature education. As this chapter illustrates, the development of Chinese children's literature is still facing the challenges of the times, and it is important to continue to reflect on the "child-centered theory", to understand the world of children, and to truly establish a new concept of "children" and children's literature in China.

References

- Bing, X., 1941. *To Little Readers*. Beixing Publishing House, Beijing.
- Cao, W.X., 2012. *Bronze and Sunflower*. People's Publishing House, Beijing.
- Chen, B.C. 1932. *Researches on Children's Literature*. Beixin Publishing House, Beijing.
- Chen, B.C., 1957. *A Brief Discussion of Children's Literature*. Yangzi River Publishing House, Shanghai.
- Chen, B.C., 1982. *Brief Discussion of Children's Literature*. Yangzi River Publishing House, Wuhan.
- Chen, Z.J., 1989. *History of Contemporary Chinese Children's Literature* (In Chinese). Hunan Youth & Children Publishing House, Changsha.
- Chen, Z.D., 2005. *Guangdong Contemporary Children's Literature*. Guangdong Higher Education Press, Guangzhou.
- Chen, A.L., 2013. *Children's Literature in China From the Perspective of Mass Culture* (In Chinese). Zhejiang University Press, Zhejiang.
- Chinese Publishing Association, 1983. *Chinese Publishing Almanac*. Press of Publishing and knowledge, Beijing.
- Fang, J.S., 1957. *Trial Discussion on Children's Literature* (In Chinese). Hebei People Publishing House, Baoding.
- Fang, W.P., 1993. *History of Criticism on Theories of Chinese Children's Literature*. Jiangsu Children Publishing House, Nanjing.
- Fang, Y.Q., 2008. *The Carrots are Back*. Children publishing House, Shanghai.
- Fang, W.P., 2018. *Four Decades of Chinese Children's Literature: A Historical Overview*. China Youth & Children Publishing House, Beijing.
- Fang, H.Z., Sheng, L.D., 1960. *Little Tadpole Looking for Mother* (In Chinese). Chinese Youth Publishing House, Beijing.
- Gao, S.Q., 1955. *Our Mother of Soil*. Chinese Youth Publishing House, Beijing.
- Gao, S.Q., reprint 2017. *The Popular Science Fairy Tales Created by Gao ShiQi*. Yangzi River Publishing House, Wuhan.
- Gu, M.Y., 2018. *The Culture Base of Chinese Education*. Shanxi Educational Publishing House, Taiyuan.
- Guan, H., 1955. *The Little Hero Yulai* (In Chinese). Youth & Children Publishing House, Beijing.
- Hao, G.C., 2009. *Why Good Picture Book Looks So Perfect?* Nanning. 21st Century Publishing House.
- He, Y., 1960. *Essays on Children's Literature* (In Chinese). Baihua Literature Publishing House, Tianjin.
- Hong, X.T., 1956. *Magic Pen Ma Liang* (In Chinese). Chinese Youth Publishing House, Beijing.
- Hong, X.T., 1990. *Taiwan Children's Literature History*. Anhui Children's Publishing House, Hefei.
- Huang, Y.S., 2021. *Courses for Children's Literature*. Zhejiang University Press, Hangzhou.
- Jiang, F., 1960. *Speech on Chinese Children's Literature* (In Chinese). Jiangsu Literature Publishing House, Nangjing.
- Jiang, F., 1979. *Essays on Children's Literature*. Hunan People's Publishing House, Changsha.
- Jiang, L., An, H., 2018. Analysis on the visual statement of political posters in Soviet Union. *J. Nat. Mus.* 10 (2), 113–121.
- Jin, J., 1955. *Annoyance of the little Captain*. Children's Publishing House, Beijing.
- Kan, W., 2019. Missions on creating of the children's picture-book (In Chinese). *Literature Constr.* 9 (2). Preface.
- Li, Z.P., 1962. *Discussions on Editing Public Children's Literatures for Kids*, in *Selected Papers on Children's Literature (1913–1949)*. Shanghai Children's Publishing House, Shanghai, p. 492.
- Li, X.T., 1972. *The Sparkling Red Star*. People's Literature Publishing House, Beijing.
- Liao, J., 2018. *Go Left, Go Right*, fourth ed. Taihai Publishing House, Beijing.
- Liao, J., 2019. *Listen to Jimmy Sing*, fourth ed. Taihai Publishing House, Beijing.
- Lin, L., 2017. Children's perspectives and values of texts: study on Jimmy's picture books. *Acad. J. Huaihai Coll.* 15 (09), 37–41.
- Liu, X.P., 2004. *Analysis of the Development of Anhui Children's Literature* (In Chinese). Anhui People's Publishing House, Hefei.
- Liu, X.Y., 2013. *A Brief History of Chinese Children's Literature 1916–1977* (In Chinese). Fudan University Press, Shanghai.
- Lu, X., 1919. *Collected Works of Luxun*. Tianjin People's Publishing House, Tianjin.
- Lu, B., 1962. *Essays on Children's Literature* (In Chinese). Youth & Children Publishing House, Beijing.
- Lu, B., 1979. *Literature for educating kids* (In Chinese). *J. Childrens Literature* 4 (2), 17–21.
- Mao, D., 1993. *Ramblings on Children's Literature in 60 years*. In: Dun, M., *The Complete Works of Mao Dun*. Huangshan Press, Hefei.
- Peng, Y., 2012. *Picture Books Should Read on This Way*. Jieli (Relay) Publishing House, Nanning.
- Pu, S.L. (Ed.), 1982. *Strange Tales From a Liaozhai*. People's Literature Publishing House, Beijing.
- Qin, Z.Y., 1950. *Little Swallow's Flight* (In Chinese). Youth & Children Publishing House, Beijing.
- Rose, J., 1984. *The Case of Peter Pan or the Impossibility of Children's Literature*. MacMillan, London.
- Shen, L.N., 2012. In: *The History and Current Publishing on Picture-Book in Mainland China* (In Chinese), vol. 12. Wide-Angle Publishing, pp. 45–50, 9.
- Silverstein, S., 1964. *The Giving Tree*. Harper & Row, New York.
- Sun, H.L., 2004. *Review of Jiangxi Children's Literature in the New Era*. Jiangxi Educational Publishing House, Nanchang.
- Wang, Q.G., 1999. Changes on views of children and transformation of children's literature in history. *J. Loudi Normal Coll.* 4 (1), 45–61.
- Wang, Q.G., 2000. *The Main Trend of Modern Chinese Children's Literature* (In Chinese). Chongqing Publishing Group, Chongqing.
- Wang, J.H., 2011. *Enjoying Picture Books With Children*. National Library Press, Beijing.
- Wang, Q.G., 2016. Realistic: the main wave of children's literature in hundred year. *J. Henan Soc. Sci.* 24 (6).
- Wang, Q.G., 2009. *Children's Literature Course*. Beijing Normal University, Beijing.
- Wei, S.Y., Zhou, H.Y., 1923. *Introduction to Children's Literature*. The Commercial Press, Shanghai.
- Wu, Q.N., 2011. Turning Changes in Chinese Children's Literature Theory: From "Child-Oriented" to "Creating for Children". *Reading News of China*, December 7.
- Wu, Q.N., 2012. *Cultural Interpretations on Chinese Children's Literature in 20th Century*. China Social Science Publishing House, Beijing.
- Xiang, L.D., 2017. The heritage and innovation of visualization in folk literature in mainland China (In Chinese). *J.Kunming Coll.* 12 (1), 54–60.
- Xiong, B.Z., 2018. *Memory on Children Hood: History of Chinese Children*. Guangxi Normal University, Guilin.
- Xu, G.Y., 1968. *Little Soldier Zhang Ga* (In Chinese). Youth & Children Publishing House, Beijing.
- Yang, W.J., 1958. *The Harbor of Next Sailing* (In Chinese). Chinese Youth Publishing House, Beijing.
- Ye, S.T., 1934. *Kaiming National Language Textbooks*. Shanghai Science and Technology Literature Press, Shanghai.
- Zawa, 2005. *Introduction to Tibetan Children's Literature*. Qinghai Ethnic Publishing House, Xining.
- Zhang, J.Y., 2018. *Re-discussions on Inner Mongolia Children's Literature* (In Chinese).
- Zhang, L.P., 1940. *Sanmao Joined the Army*. Sichuan Children Publishing House, Chengdu.
- Zhang, T.Y., 1955. *The Story of Luo Wenying*. Children's Press, Beijing.
- Zhang, T.Y., 1956. *The Secret of the Treasure Gourd* (In Chinese). Youth & Children Publishing House, Beijing.
- Zhang, T.Y., 1934. *Dalin and Xiaolin*. Reprint (2006).
- Zhang, T.Y., 1952. *The Story of Luo Wenying* (In Chinese). Youth & Children Publishing House, Beijing.
- Zhang, L.P., 1983. *Stories of Sanmao*. Anhui Children's Publishing House, Hefei.
- Zhang, Y.J., 2006. *The History of Chinese Children's Literature in 20th Century* (In Chinese). Liaoning Youth & Children Publishing House, Shenyang.
- Zhao, J.S., 1962. *On Children's Picture-Story*. In: Zhao, J.S. (Ed.), *Series Discussions on Folk Literature* (In Chinese). Hunan Publishing House, Changsha.

- Zheng, M.J., 1986. On Japanese Children's Picture Books. Qinlin International Publishing House, Taipei.
- Zheng, Y.J., 1989. The Complete Biography of Pipilu and Luxixi. Sichuan Children's Publishing House, Chengdu.
- Zhou, Z.R., 1920. Literature of Children. Children's Bookstore, Shanghai.
- Zhou, Z.R., 1932. On Children's Literature. Children Publishing House, Shanghai.
- Zhou, M.M., 1996. One Hundred Years of Hong Kong's Children's Dreams (1960s-1990s). In: Fang, W.P., (edn.), Series of Chinese Children's Literature, Theories III, Hope Publishing House, Taiyuan.
- Zhu, D.Y., 1924. Children's Literature Studies (In Chinese). Zhonghua Publishing House, Shanghai.
- Zhu, Z.,Q., 1998. Chinese Children Literature and Progress of Modernity. Zhejiang Children Publishing House, Hangzhou.
- Zhu, Z.,Q., 2000. Chinese Children's Literature and the Process of Modernization. Zhejiang Children's Publishing House, Hangzhou.

Children's literature in Greece and Cyprus

Petros Panaou, University of Georgia, Department of Language and Literacy Education, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	672
Historical development of Greek children's literature	672
The springtime of children's literature in Greece	673
Greek-Cypriot children's literature	673
The talent and influence of Eugene Trivizas	674
Traits and trends of Greek children's literature	675
References	676
Children's books	676
Further Reading	676
Secondary sources	676

Introduction

Greek and Greek-Cypriot literatures for children and young adults are rich, adventurous, and socially aware literatures. As such, they deserve to be translated and made available to readers across the globe. Young readers should not be deprived of the playful joy and global awareness that these books can bring into their lives. This paper contributes to informing international scholars and educators about Greek and Greek-Cypriot literatures for children and young adults, thus encouraging them to explore and share their worthwhile literary texts transnationally. The International Board on Books for Young People (IBBY) and its peer-reviewed journal *Bookbird* have been playing a crucial role as champions of Greek, Greek-Cypriot, and other diverse literatures from around the world. This paper draws heavily from multiple *Bookbird* and IBBY sources that have continually made the world aware of talented Greek and Greek-Cypriot creators and their work. My article in a 2018 special *Bookbird* issue that focused on children's literature from Greece, Cyprus, and Turkey (Panaou, 2018), along with other articles published in the same issue, form the basis for this paper.

Historical development of Greek children's literature

According to Anagnostopoulos (2004), *Old Stathis or Memories of My Childhood* (1858) by Leon Melas was the work that marked the true starting point of children's literature in Greece. Throughout the book, Old Stathis, a sweet and fatherly figure, tells children stories from the Greek tradition, history, and mythology, lovingly advising them on how to lead a life of virtue and become socially responsible citizens. Melas' work was highly influential in developing the educational and aesthetic framework for the development of Greek children's literature. During the 19th and well into the 20th centuries, ethical, religious, and national edification was the main focus of Greek texts addressed to children (Anagnostopoulos, 2004).

As interest in educational matters grew in Greece, children's literature benefited mainly through the proliferation of children's magazines. One of these magazines, *The molding of youth* [I diaplasis ton pedon] was founded in 1879 and influenced the production of children's literature for the following seventy years. It rallied around itself prestigious authors who published their work in the magazine and dynamically promoted children's literature in Greece (Anagnostopoulos, 2004). Circulating thousands of copies of each issue, *Diaplasis* included a broad spectrum of texts, from poems, to fables, to short stories, and novels in periodical installments. Scholars have praised the magazine's longevity, originality, content variety, artful illustration, and appeal to the child's psyche, as well as the masterful ways in which it used simple but aesthetically pleasing language to cater both to the child reader's instruction and delight (Patsiou, 1995, p. 103). During the same time, anthologies of children's poems for school use, as well as translations and adaptations of international works by authors such as Verne, Dickens, and Twain, offered inspiration for a Greek literary tradition of children's and young adult literature.

In the first decades of the twentieth century, historical events such as the First World War, the Balkan Wars, and the Asia Minor Catastrophe, along with conflicts around language issues and educational reform, shaped Greek education and children's literature (Anagnostopoulos, 2004). An important development was the pioneering and establishment of the historical novel as a children's literature genre by Penelope Delta (1874–1941) with works such as *For the motherland* (1909) and *A tale with no name* (1910). *For the motherland* is a historical novel that takes place during the Byzantine era and is packed with adventure, patriotic heroism, romance, and suspense. Even though *A tale with no name* is fantasy, it espouses similar ideals, such as heroism and patriotism. It is the story of a young prince, who saves his father's deteriorating kingdom and leads it victoriously through battle and toward rebuilding and peaceful prosperity. While these works were quite didactic, they also told moving and delightful narratives that engaged young readers' interest.

The springtime of children's literature in Greece

Tsilimeni (2004) identifies the predominance of comic books, illustrated classics, and cheap popular publications, along with an effort to provide children with books that were distinctively Greek, as the two main trends during the post-Second-World-War period. In the decades that followed, Greek children's literature began to mature, but it was not until the 1970s that it truly bloomed. It was at this time that literary production in Greece developed dramatically, both in quantity and quality, breaking away from didacticism and addressing the young reader's critical thinking with respect. Tsilimeni explains that the period from 1969 to 1980 is considered the springtime of children's literature in Greece. Two important factors for this springtime were the reinstatement of democracy in the country and the institution of Demotic Greek (the language spoken by the people at that time, rather than the archaic Katharevousa language) as the official state language. Other factors included the introduction of children's literature as a course/discipline at Greek Universities, the publication of the first studies on children's books, and the activity of literary societies such as the Circle of Greek Children's Books (the Greek national section of IBBY).

Ada Katsiki-Givalou (2002, p. 47) recognizes Alki Zei as one of the most influential authors in the development of children's literature in Greece from the '70s and onwards, especially after the collapse of the dictatorship in 1974. Givalou sees the transition to democracy, liberal educational theories and reforms, and cultural convergence with Europe as conditions that formed Greek authors' focus on child perspectives and their "conviction that the child is an equally important member of society, participating in the community's socio-political affairs, observing scientific and other developments, and forming his or her own personality and character" (p. 48).

Zei's works—including *To kaplani tis vitrinas* [Wildcat under glass] and *O megalos peripatos tou Petrou* [published in English as *Petros' War*]—have been widely translated and internationally published. *Wildcat*, which was first published in 1963, features child characters who experience fascism and dictatorship. In *Petros' War* (first published in 1971) children experience life in Athens, Greece under Italian and then German occupation during World War II. Zei was impressively skilled in authoring stories from a child's perspective (p. 48) as well as in incorporating delightful humor, even as she wrote about serious subjects, such as war, addiction, oppression, and intergenerational conflict (Freeman et al., 2004, p. 41). Here is how the author herself explained her commitment to child perspectives:

I believe, I am actually convinced, that a decisive factor for the success of a book for children is the writer's ability to truly assume the role of the child, i.e., to see everything through the eyes of a child who writes. I remember once seeing an English film, a detective movie, in which the main character was a ten-year-old boy; the entire film was shot from the point of view of the child, furniture appeared larger than normal, everything was above the child's eye level. This has turned into a guiding principle for me. If one succeeds in this, then the book will work, then everything will click into place.

Qtd. in Katsiki-Givalou (2002, p. 48).

In addition to Zei, Givalou mentions several other Greek authors who started to create critical and socially aware young characters in their stories during that time and beyond, including Zorz Sari, Litsa Psarafi, and Pantelis Kaliotsos. The main trends were toward realism, the projection of values with an international character, reference to space and technology, and to the natural environment. Novels became the dominant form and featured dynamic treatments of contemporary global problems, social and political events, and the relationship between humanity and the environment. Poetry also followed new paths in both form and content and children's theater developed radically, as it was freed from its close ties to education (Katsiki-Givalou, 2002).

According to Tsilimeni (2004), the didactic element virtually disappeared from Greek children's literature in the 1990s, the predominant elements becoming entertainment along with social, environmental, and international issues. To this day, contemporary Greek children's literature continues to respond to the call of our times by cultivating in young readers an optimism for life, an appreciation of love in human relationships, and the value of peaceful communication and co-existence between peoples and communities.

Greek-Cypriot children's literature

Interestingly enough, Greek children's literature on the island-nation of Cyprus also bloomed from the 1970s and onwards, but not due to the fall of the dictators in Greece. During that same time, tragic, traumatic events took place in Cyprus. A military coup that was directed by the Greek dictators was followed by a Turkish military invasion and occupation of a large part of the island. These events changed the course of Greek-Cypriot children's literature. "Ironically, one of the darkest moments in the modern history of Cyprus brought the booming development of its literature" (Panaou and Michaelides, 2010). Passing through stages of anger, mourning, and hope, numerous authors—with Kika Poulcheriou and Maria Pyliotou being the most prominent ones—wrote children's literature that was inspired by the tragic stories they had witnessed, the trauma that both children and adults had experienced, and the challenging times that followed.

In a special *Bookbird* issue that focused on Greek, Greek-Cypriot, and Turkish children's literature, Chatzianastasi (2018) looks at one aspect of this literary production, examining how representations of enclavement (the experience of small Greek-Cypriot communities living in isolation within the Turkish-occupied area) explore traumatic experiences and help create new frameworks of understanding trauma. And in their IBBY Congress presentation, Katsonis and Kyriakides (2008) examine fifteen story collections

(a total of 126 short stories by sixteen different authors) that revolve around the 1974 war and its implications for every aspect of life on the island. They identify four thematic strata in these stories:

- Cultivation of love for democracy as a response to the coup that preceded the war
- Cultivation of pacifism and anti-war stances, along with preservation of memories from the occupied lands
- Reference to the loss of beloved ones and the remembrance of missing persons
- Endorsement of peaceful coexistence between the Greek and Turkish communities on the island and encouragement of hope for a settlement of the Cyprus problem.

Katsonis and Kyriakides (2008).

Michaelides and Panaou (2010) view Greek-Cypriot children's literature as a minority literature, existing in the periphery of metropolitan Greek Children's Literature, striving to establish its place within the European and international literature. Greek literature published for children on this island has historically managed to progress in spite of illiteracy, colonialism, war, marginal cultural status, identity crisis, and language confusion (because of the marginalization of the Greek-Cypriot dialect); this persistence reveals the strength of the minority culture it is stemming from and the will of those involved to help it achieve the status it deserves (Panaou and Michaelides, 2010).

Antri Antoniou, a young author belonging to a new generation of prolific Cypriot creators that matured during the period following Cyprus' inclusion in the European Union, says "Cypriot children's literature merits recognition and respect. During the past couple of decades, a substantial number of young Cypriot authors have been writing texts that are of equal quality with good books authored by writers who live in Greece" (Charalambous, 2018). Indeed, from the beginning of the 21st century and onwards—a period that Michaelides and Panaou (2010) call "The Era of Europe" as Cyprus officially became an EU member state in 2004—an entire new cohort of Greek-Cypriot authors have been creating literature of impressive quality and innovation. Several of their works, including texts by Antri Antoniou, Anna Kouppanou, Marina Michaelidou-Kadi, Elena Perikleous, Demetra Charalambous, Despina Eracleous, Fytoula Vakana, and others have been successfully published in Greece, where they have also won prestigious awards.

Cyprus' 2022 Hans Christian Andersen author nominee Anna Kouppanou, for example, is celebrated both in Cyprus and abroad. Her novel *The incredible discovery of Sebastian Montefiore* was awarded the Republic of Cyprus State Award for Young Adult Literature as well as the Anagnostis (Reader) award in Greece. The same book was included in IBBY's Honor List in 2018. Kouppanou's literary work encourages interspecies coexistence and respect for difference, while calling for social change. She believes that children face much hardship, which often ensues from the way adults fail to understand and communicate with them. Her books strive to fill in this gap (Mathis and Panaou, 2021, p. 17). In her own words, her books aim "to offer a space in which children can experience moments of safety within a world that is just coming into being, a world that invites them to think for themselves and to change it—if they wish to do so" (p. 17).

Anna Kouppanou's work is often organized through a creative combination of the real and the imaginary, drawing on enchanted and mythical elements that take on new meanings. Along with *The incredible discovery of Sebastian Montefiore*, her works *The argonauts of time* (Republic of Cyprus State Award), *The day we broke the world* and *To the dark of darkness and beyond* contribute to the renewal of themes and narrative techniques in children's fantasy, not only in Cyprus but also in Greece. And with books such as *Grandpa, Superman, and a little cherry* and *The disappearance of Mrs. Papadakou* (nominated for the Hellenic Republic State Award; short-listed for the Republic of Cyprus State Award and the IBBY Greece Award), Kouppanou innovates realistic fiction, engaging important psychological, philosophical, and social issues while depicting childhood and adolescence in contemporary terms. In her acclaimed picturebook *O Fivos ke i falena* [Fivos and the whale], for instance, a young boy named Fivos gives shape to his overwhelming sense of fear, turning it into a whale and releasing it in the sea. In this manner, he makes room in his heart for other things/emotions.

Sandra Eleftheriou, a member of this new cohort of creators, was the first-ever illustrator to be nominated by the Cyprus branch of IBBY for the Andersen award. Like the rest of her cohort, playfulness is central in her work: "What instantly wins the reader's interest is the illustrator's playfulness. She plays with her drawings, with colors, figures, and shapes the way a child would play, and it's this playful mood that gives the text she is illustrating a new dimension" (Panaou and Mathis, 2019, p. 22). A small number of texts by these artists has been translated and/or included in international lists. The European influences on their work are evident, as are the influences from metropolitan Greece. Pavlos Ioannides writes: "Although Greek Cypriot literature reflects the way of life, history, feelings, hopes and dreams of islanders, it is strongly attached to, and affected by, the literature of Greece" (2002, p. 2).

The talent and influence of Eugene Trivizas

Perhaps no other Greek children's literature author has been more influential to his peers, in Cyprus and in Greece, than Eugene Trivizas. Christie Davies writes about Trivizas: "He is justly celebrated in his native Greece, where he has been described by reviewers and commentators as a 'miracle worker who brings to life a whole new world based on imagination,' 'a revitalizing force in Greek children's literature,' 'the most important heir of Rodari,' and 'a classic writer whose books belong in each and every library'" (2003, p. 23). Writing in 2003, Davies praises Trivizas' international appeal and predicts a stellar international career, which never came; at least not to the extent that critics had expected it would.

The most talented and prolific Greek author of our times (with more than 150 works) seems to currently only have one book in print in English. This is *The three little wolves and the big bad pig* (1994), illustrated by Helen Oxenbury and concurrently published in Greek and English. It has been translated in more than 20 languages, is an international bestseller, and has won multiple awards. *The three little wolves* retells the classic *Three little pigs tale*, reversing the roles of the main characters and subverting the moral universe of the story. Rather than ending with a harsh punishment of an evil adversary, Trivizas' imagines a world where beauty, friendship, and kindness prevail and function as agents of change. It is a masterfully told, humorous, and playfully subversive fractured fairy-tale, but Trivizas' other works—including *The last black cat* (2012), which has been published in English but is currently out of print—are equally deserving of being read across the globe. *The last black cat* is a fantasy novel set on a distant island, where the powers that be convince people that black cats and the bad luck they bring are to be blamed for everything that is wrong on the island. The main characters, a black cat and a young girl, fight against all odds to put an end to the persecution and violence that ensue. Reminiscing of the persecution of Jews by the Nazis, Trivizas' powerful story is a warning that "people forget and it doesn't take much for madness to return" (Trivizas, 2012, p. 358).

Davies attributes the greatness of Trivizas' work to three elements. First, to his gift for humorous play with risk, fear, and danger, as well as with incongruity and nonsense. Next, to the way in which his writing operates on multiple levels, engaging readers of diverse ages. And finally, to his ability to transform the familiar, spinning traditional stories, well-known images, and everyday words in unexpected and comic ways (Davies, 2003, p. 23). Yannicopoulou (2007) writes about the interactive, hypertextual elements in Trivizas' fiction, Kanatsouli (2012) praises the use of play in his work, and Athanasiou-Krikelis (2018) analyzes the brilliant ways in which Trivizas subverts power and authority.

Traits and trends of Greek children's literature

All of the above have been influencing Greek and Greek-Cypriot authors, even though one should not, of course, attribute all innovation and narrative playfulness to Trivizas. In fact, Davies argues that it was the openness of Greek children's literature to fluidity and experimentation that enabled Trivizas to thrive in the first place: "In a less adventurous, more restrictive, less aesthetically aware culture, a children's and adults' writer like Trivizas would not have been given this chance to succeed" (2003, p. 31). Writing about contemporary Greek illustrated books for children, Alexandra Zervou asserts that they engage in "an unending game of opposites: it is a meeting point of the world of adults and the world of children, of tradition and innovation, of dreams and real events" (2002, p. 6).

Tradition and innovation, the second game of opposites listed by Zervou, is also identified in the work of Greek illustrators by Elena Popova, who observes that they successfully merge classic styles with "the synthetic language of modern art" (2002, p. 40). Tradition and innovation, sometimes merging and sometimes struggling with each other, form a central feature of contemporary Greek literature for young readers. Kanatsouli and Tzoka discuss a selection of contemporary texts that take the traditional concept of Greek national identity and evolve it into a more "multicultural view of Greekness" (2005, pp. 32–33). Papantonakis (2004) explains how contemporary Greek science fiction for youth reconciles the seeming opposites of national and global awareness. Both of these articles, for instance, discuss stories in which the protagonists retain their Greek identity while upholding universal values and striving for the global good, sometimes collaborating with people of other ethnicities to achieve that.

Tsilimeni and Panaou (2011) divide contemporary Greek poets for children into those whose work could be described as traditional poetry and those who stray from traditional paths, in order to free their work from conventional forms, subjects, and functions. In the past few decades, for example, an increasing number of poets have been exploring nonsense, humor, and parody, and are constantly eager to play with language. Theti Chortiaty is an important poet who plays with all of the above. Chortiaty started her writing career with traditional poems, but soon turned to radically different kinds of poetry that make heavy use of language play, alliteration, neologism, portmanteau techniques, and so on. In the following sample of her work from her poetry collection *Pechnidolexa* [Playwords] (1986) she joyfully and nonsensically plays with homophones (underlined):

I giagia psomi zimoni
Ke sto spiti zi moni

[Grandma bread bakes]
[And in the house she lives alone]

Similarly, Politis (2018) places Greek children's literature at the end of the 20th century in a middle-ground between Modernism and Postmodernism, identifying renewed writing styles and an adaptability to the needs of present-day readers. Politis explains how authors like Manos Kontoleon, Eugene Trivizas, and Christos Boulotis use postmodern techniques to offer readers multiple perspectives and diverse realities. In his novel about the life of a family *To 33* [The 33], for instance, Kontoleon concludes a multifocal narrative that is characterized by frequent narrative ruptures by inviting the reader to provide the ending they prefer. And Trivizas' *Ta 88 Dolmadakia* [The 88 Dolmas] could be described as an entirely meta-fictional/experimental text as it relinquishes to the reader the responsibility of connecting its narrative parts and making them work. Depending on the narrative option they select at the end of each page, readers may create a new story every time they read the book (a total of 1921 different possible stories).

Vagelis Iliopoulos, the 2018 Hans Christian Andersen Greek author nominee, is also described as a writer who is “always experimenting with the form and meaning of words” and “does not hesitate to apply pioneering narration techniques” (Sundmark, 2017b, p. 22). And the 2018 illustrator nominee Christos Dimos observes the past with a sensitive eye to produce images that “enlighten present eyes about the future” (Adamopoulou qtd. in Sundmark, 2017a, p. 53). For contemporary book creators in Greece, this enlightenment about the future often requires critical questioning and social change. Maria Papayanni, Greece’s 2022 Andersen author nominee brilliantly summarizes this approach: “No, a book cannot change the world. But it can open Aeolus’s bag of wind; it can unleash new winds and give birth to new questions. A book can lure young readers to the forest of life, where they can search for answers that will lead to a more just society” (Mathis and Panaou, 2021, p. 25).

As mentioned earlier, the purpose of this chapter is not to cover the entire history or current state of Greek and Greek-Cypriot children’s literatures. It rather aims to introduce its readers to the production of Greek literary texts for young readers in Greece and Cyprus and convince them that young readers across the globe should not be deprived of the playful joy and social/global awareness that can be found in these books. Davies writes: “Trivizas’ work deserves to be more widely known. Who knows how many more potential international bestsellers are lurking in the treasure chest of this remarkable author?” (2003, p. 34). The same could be said about the wider treasure chest of Greek and Greek-Cypriot literary production for children and young adults.

References

Children's books

- Chortiat, T., 1986. *Pechnidolexa* [Playwords]. Kedros, Athens.
 Delta, P., 1909. *For the Motherland* [Gia tin patrida]. G.S. Veloni, London.
 Delta, P., 1910. *A Tale With No Name* [Paramithi choris onoma]. G.S. Veloni, London.
 Melas, L.G., 1858. *Old Stathis or Memories of My Childhood* [Gerostathis I Anamnis tis pedikis mou ilikias]. Printing House Lazarou, Athens.
 Trivizas, E., 1994. *The Three Little Wolves and the Big Bad Pig* [Ta tria mikra likakia]. Minoas, Athens.
 Trivizas, E., 2012. *The Last Black Cat* [I teleftea mavri gata]. Metechmio, Athens.

Further Reading

- Kontoleon, M., 1989. *To 33* [The 33]. Patakis, Athens.
 Kouppanou, A., 2009. *The Argonauts of Time* [I argonautes tou chronou]. Kedros, Athens.
 Kouppanou, A., 2015. *The Incredible Discovery of Sebastian Montefiore* [I apistefti anakalipsi tou Sebastian Montefiore]. Kedros, Athens.
 Kouppanou, A., 2017. *The Disappearance of Mrs. Papadakou and What Happened that Summer* [I eksafanisi tis K. Papadakou]. Patakis, Athens.
 Kouppanou, A., 2018. *Grandpa, Superman, and a Little Cherry* [O pappous, o Superman kai ena kerasaki]. Kedros, Athens.
 Kouppanou, A., 2019. *The Day We Broke the World—The Lost Kids Club, Vol. 1* [Ti mera pou spasame ton kosmo]. Patakis, Athens.
 Kouppanou, A., 2021. *Fivos and the Whale* [O Fivos ke I falena]. Patakis, Athens.
 Trivizas, E., 1997. *The 88 Dolmas* [Ta 88 Dolmadakia]. Illustrated by Rania Varvaki. Kalentis, Athens.
 Zei, A., 2011a. *Petros' War* [O megalos peripatos tou Petrou]. Metechmio, Athens.
 Zei, A., 2011b. *Wildcat Under Glass* [To kaplani tis vitrinas]. Metechmio, Athens.

Secondary sources

- Anagnostopoulos, V.D., 2004. Greece: from the beginnings to 1945 (Part V: national and international). In: *International Companion Encyclopedia of Children's Literature*. Routledge, Taylor & Francis Group, Abingdon.
 Athanasiou-Krikelis, L., 2018. Authority and power: the peculiar case of Ignatius and other picturebooks by Eugene Trivizas. *Bookbird* 56 (2), 43–51.
 Charalambous, D., 2018. Interview with Antri Antoniou, a prolific young writer from Cyprus. *Bookbird* 56 (2), 67–69.
 Chatzianastasi, M., 2018. Cultural expression and representation of the quotidian trauma of enclavement in literature for the young in Cyprus. *Bookbird* 56 (2), 36–42.
 Davies, C., 2003. A Greek bearing gifts: the literary talents of Eugene Trivizas. *Bookbird* 41 (3), 23–35.
 Freeman, E.B., et al., 2004. Alki Zei: author—Greece. *Bookbird* 42 (4), 41.
 Ioannides, P.F., 2002. Letter to the editors. *Bookbird* 40 (1), 2–3.
 Kanatsouli, M., Tzoka, T., 2005. Embracing multiculturalism through understanding “Greekness”: contemporary Greek books for children. *Bookbird* 43 (2), 30–37.
 Kanatsouli, M., 2012. Games inside books for young children. *Bookbird* 50 (4), 33–40.
 Katsiki-Givalou, A., 2002. Society, history, and the environment: contemporary concerns in Greek children’s literature. *Bookbird* 40 (1), 47–51.
 Katsonis, K., Kyriakides, E., 2008. Reflections of Cyprus history in novels for children and young adults. In: *Proceedings of the 31st IBBY World Congress*. IBBY, Copenhagen.
 Mathis, J., Panaou, P., 2021. Anna Kouppanou: author—Cyprus. *Bookbird* 59 (4), 17.
 Panaou, P., 2018. Come explore with me the literary treasures of Cyprus and Greece: a short introduction to Greek and Greek-Cypriot children’s literatures. *Bookbird: J. Int. Child. Lit.* 56 (2), 4–8.
 Panaou, P., Mathis, J., 2019. Sandra Eleftheriou: illustrator—Cyprus. *Bookbird* 57 (4), 22.
 Panaou, P., Michaelides, F., 2010. Greek-Cypriot children’s literature: a small literature and the challenges it faces in the big world of globalized publishing for children. In: *Proceedings of the 32nd IBBY World Congress*. IBBY, Santiago de Compostela.
 Papantonakis, G., 2004. Globalization, national awareness, and children’s literature: evidence from Greek children’s science fiction. *Bookbird* 42 (2), 37–43.
 Patsiou, V., 1995. *I Diaplasia ton Pedon 1879–1922* [The Molding of Youth 1879–1922]. Kastaniotis, Athens.
 Politis, D., 2018. Between modernism and postmodernism: Greek literature for children and youth in the last decades of the 20th century. *Bookbird* 56 (2), 52–56.
 Popova, E., 2002. Along the via Dolorosa of transition—and beyond: children’s book illustration in Bulgaria, Greece, Yugoslavia, and the former Yugoslav Republic of Macedonia. *Bookbird* 40 (1), 35–41.
 Sundmark, B., 2017a. Christos Dimos: illustrator—Greece. *Bookbird* 55 (4), 53.
 Sundmark, B., 2017b. Vagelis Iliopoulos: author—Greece. *Bookbird* 55 (4), 22.
 Tsilimeni, T., Panaou, P., 2011. Poetry for young people in Greece: themes, forms, types and trends. *Bookbird* 49 (2), 17–24.

- Tsilimeni, T., 2004. Greece: from 1945 to the present (Part V: national and international). In: *International Companion Encyclopedia of Children's Literature*. Routledge, Taylor & Francis Group, Abingdon.
- Yannicopoulou, A., 2007. Hypertextual elements in the interactive fiction of Trivizas. *Bookbird* 45 (1), 15–22.
- Zervou, A., 2002. Historicity and nature—text and image: the game of oppositions in Greek illustrated books. *Bookbird* 40 (1), 6–15. **This page intentionally left blank**

Educational concerns in Arabic picture books

Yasmine S. Motawy^a and **Susanne K. Abou Ghaida^{b, a}**, ^a Department of Rhetoric and Composition, The American University in Cairo, New Cairo, Egypt; and ^b University of Glasgow, Glasgow, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	678
The Arab picture book ecosystem and its opportunities	678
The Arab children's publishing scene	678
The scholastic landscape in Arab countries	680
Other active agents in the ecosystem	681
Trends and opportunities	681
Inside the book	683
Language	683
Poetry and rhyme	685
Illustration and type	685
Conclusion	686
References	686

Introduction

The history of children's literature is as strongly intertwined with the development of education and schooling in the Arab world, as it is in any other context. More specifically, the spread of mass schooling in the twentieth century made the decipherment of written text possible for a large majority of the populace, dramatically expanded the market for books, and contributed to the creation or enlargement of national middle classes to which most writers and buyers of books continue to belong. Conversely, the scholastic market strongly asserts itself as the main market for Arabic children's books today, especially as neither public libraries nor individual consumers currently constitute a critical buying mass. It is exactly this symbiosis and the tensions between consumer publishing and the scholastic book market that this article examines as it approaches the Arabic children's book, in particular the Arabic language picture book, as an educational artifact that rests at the heart of a complex ecosystem.

We break down this ecosystem in the first half of this article (Section **The Arab picture book ecosystem and its opportunities**). We start with an overview of the Arabic children's publishing field (Section **The Arab children's publishing scene**), the scholastic landscape in the Arab world (Section **The scholastic landscape in Arab countries**), other relevant actors (Section **Other active agents in the ecosystem**) as we advocate for greater coordination and cooperation among these various nodes, as well as make recommendations for reading promotion (Section **Trends and opportunities**). In the **Inside the book** section, we examine the form of Arabic picture books, including language, rhyme, illustration and design from a pedagogical angle.

The general consensus in the Arab world is that the function of children's literature is primarily as a socialization tool which provides an esthetic literary and artistic experience, teaches the children something about themselves, promotes linguistic, mental, moral and sensory development and aids emotional maturity and psychological adjustment to the world (Kenawy, 1994, 28). This entry does not survey or make commentary on the content of picture books at this time, as this is a sizable endeavor, and has been partially dealt with by Chèvre (2015), in her overview of 40 years of Arabic picture books in Egypt, Lebanon and Syria, and Motawy (2021) in her ideological analysis of three main themes in Egyptian picture books in the last two decades.

The Arab picture book ecosystem and its opportunities

The ecosystem depicted below, with the book reader at its center, pictorializes the various parties who act upon the reader and their relationships with one another, recognizing that many more relationships are possible, and that the qualitative dynamics between them require elucidation. This visualization facilitates the identification of weak nodes and potential relationships and flows that can improve the reader's access to, and utilization of quality picture books (Fig. 1).

The Arab children's publishing scene

Arabic is an official language in twenty-five countries. These include the twenty-two countries that are members of the Arab league as well as three other countries. There are also Arab diaspora communities as well as others invested in reading Arabic, learning the language or teaching it to their children, whether for religious, cultural, personal, or professional reasons. With the lack of comprehensive histories, statistics, and bibliographic documentation of Arabic books for children, a large-scale pan Arab scholarly collaboration is needed if the entire Arab world is to be covered. As such, in the interest of attainability of scope, this article focuses on the

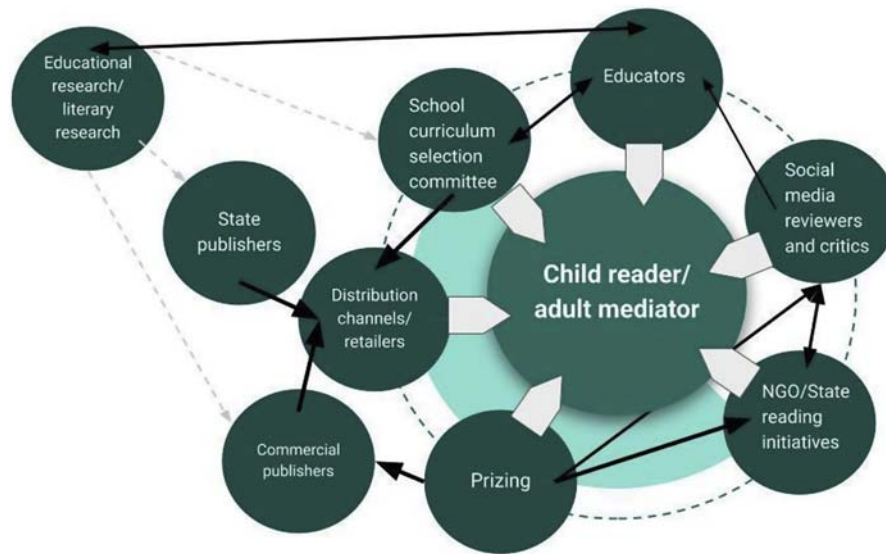


Fig. 1 The picture book ecosystem.

production of the main book publishing nations in the language bloc. We additionally focus on general trends in the field and attempt to be as inclusive as possible in our choice of examples.

The Arab book market is, on the one hand, an aggregation of twenty-two different national markets, each with its own particularities; on the other hand, there is also a regional market due to the status of Standard Arabic as the shared macrolanguage of reading and publishing. This makes the entire Arab world a potential market for an Arabic book and enables artistic collaborations, where the publisher, author, and illustrators may be from different Arab countries. Furthermore, different countries play unequal roles in publishing and buying Arab children's books although these roles may also change and shift in response to regional and international political and economic developments. In general, Egypt and Lebanon, notwithstanding economic and political shifts, are the two countries with the longest publishing and printing traditions as well as the major producers of Arabic children's books. In terms of book buying, the market currently skews toward the oil rich, member countries of the Gulf Cooperation Council (GCC), especially the Kingdom of Saudi Arabia, the United Arab Emirates and Kuwait. As a result, the markets in these countries are a main target for many publishers, alongside their national markets.

Starting in the new millennium, there has been a new wave of production of Arabic picture books (Chèvre, 2015; Motawy, 2021). By the mid 2000s, many new publishing houses were being founded as already established publishing houses began to take or renew their interest in publishing for this age group. This coincided with the emergence of a new generation of authors and illustrators, as new authors and illustrators continue to enter this field. What also played a role in this renaissance was institutional support, whether in the shape of the establishment of national and regional book prizes, or financial support for training workshops and projects in the field of children's literature and reading promotion although some of these programs have since ended. This new wave was especially prominent in countries such as Egypt, Lebanon, the United Arab Emirates and to a lesser extent Jordan, Palestine and the Kingdom of Saudi Arabia. While the war in Syria affected book production in the country, Syrian publishers, authors and illustrators still play an important role through their contribution to books published in other Arab countries or their establishment and/or relocation of publishing houses to other Arab or European countries.

Arab publishers for children tend to be either state institutions, private sector publishers or NGOs. Among Arab state publishers, the most prominent and influential is the General Egyptian Book Organization (GEBO), part of the Egyptian Ministry of Culture. It is one of Egypt's largest publishers, and additionally sponsors the "Family Library" initiative which subsidizes the republication of curated books at highly affordable prices. The GEBO behemoth aside, most prolific publishers of Arabic children's books tend to be private sector companies, both general interest publishers as well as specialized publishers. The majority of these publishers are small to medium-sized institutions, mainly family-owned and operated and include several one-woman shows (Stephan-Hachem, 2010). There is also a noticeable trend of publishing houses established and headed by children's authors (Stephan-Hachem, 2010; Kreidieh, 2015), such as Taghreed Najjar (Al Salwa Publishers, Jordan), Abeer Al-Taher (Dar Al-Yasmine, Jordan), Amal Farah (Dar Shagara, Egypt), Weam Ahmed (Aliareach, Egypt), Amal Naser Kojok (Dar Rummana, Lebanon), Nabiha Mheidly (Al Hadaek Group, Lebanon) and Arwa Khmayyis (Dar Arwa, Kingdom of Saudi Arabia). The most prominent NGO publisher is Tamer Institute for Community Education (Palestine).

While possessing a certain dynamism, the Arab publishing scene also faces a number of challenges, including a vulnerability to external events, instability, and distorted structures. Due to the relative ease and low cost of establishing publishing houses, new publishers may regularly emerge but existing ones also disappear or become inactive. In addition, there are often no clear divisions

of roles within publishing houses or between publishers and distributors. Many heads and founders of publishing houses carry out roles that should be assigned to editorial committees or marketing teams (Kreidieh, 2015, 193). More challengingly, most publishers are also the main distributors of their books, due to inefficient distribution networks, mistrustful relationships with distributors, or the very large cuts often taken by the latter, between 40 and 60% of the cover price (Arab Thought Foundation, 2008, 446; Stephan-Hachem, 2010, 111; Eschweiler and Goehler, 2010, 195). This places a great strain on publishers in terms of human and financial resources and often restricts marketing opportunities in other Arab countries to distributor and bookstore collaborations that come out of personal connections, co-publishing agreements with other publishers, or selling books directly to clients during Arab book fairs. The lack of a unified or centralized pan-Arab distribution system is further hampered by the lack of free movement between many countries, currency exchange rate differences which make a book affordable in one country but expensive in another, and high shipping costs (Arab Thought Foundation, 2008, 446; Stephan-Hachem, 2010, 111; Eschweiler and Goehler, 2010, 195). For these and other reasons, most print runs are in the range of 1000–3000 books and stocks often take several years to sell (Kreidieh, 2015, 138).

An additional bottleneck to book distribution is the not fully realized role of public libraries in the Arab countries. While most countries do have library networks, these do not keep pace with population increases, and often exclusively exist in large urban centers, leaving the rest of the country to be served by mobile seasonal libraries, if at all. Additionally, in countries with budget constraints, libraries suffer from outdated collections, untrained librarians, and a host of infrastructural issues. One of the findings in the research carried out by Kreidieh (2015) is the important role that libraries played in exposing many of the parents she surveyed in countries such as Lebanon and Tunisia to titles they would not have been aware of otherwise (221 and 233). This supports the argument made by Stephan-Hachem (2010) that the health of the book industry depends on all the elements of the chain of book production and distribution, including libraries (131).

This state of affairs has implications for book content and the equity of access to books. The challenges of selling even relatively limited print runs put many publishers in the position of needing to play it safe, sometimes at the expense of originality, to avoid potentially controversial content and to favor didactic sure-sells over more playful books. In addition, publishers evenly straddle considerations of authenticity versus regional commercial reach and often exclude from books any material that might be deemed too “local”, whether it be the inclusion of dialect, local minority languages, or references and landmarks that might only be familiar to residents of a particular country. In terms of access to books, this is highly dependent on geographical location and circumscribed by a number of factors, including the number of active publishers in the country in question, the extent to which existing distribution and cooperation links allow access to books from other Arab countries. Expectedly, those residing outside urban centers have far less access to bookstores, which hinder their ability to buy books, even ones published in their countries (Stephan-Hachem, 2010, 141–142; Eschweiler and Goehler, 2010, 197).

The scholastic landscape in Arab countries

In spite of the growth of mass schooling and attention to education following the independence movements that swept across the different Arab countries in the twentieth century, a number of Arab countries continue to struggle with various rates of adult and youth illiteracy to this day (UNICEF, 2021). With the diverse trajectories taken by the educational systems in the different countries, the scholastic landscape in the region is characterized by diverse curricular priorities and educational practices as well as a wide panorama of school categories. These include public or state schools; private school; hybrid forms (e.g. government-subsidized private schools), religious schools (e.g. Azhari religious schools in Egypt), branches of international schools, language schools, community schools, or schools tied to UN agencies that serve refugees in various countries. Within this context, differences in the role assigned to Arabic children’s literature and the access children would have to such books is affected by a number of factors, including the priorities of the national Arabic curriculum, the nature and pedagogical orientations of the school and available financial resources or subsidies.

Increasingly, the Arabic language and literature curricula in various countries are including the “promotion of voluntary reading” within their educational objectives. However, it is often mentioned alongside many other objectives that are more focused on literacy acquisition (decoding, comprehension, writing etc.) (UAE Ministry of Culture, 2020). This is an acknowledgment of the important role that reading and children’s literature can play in developing literacy, a fact underlined by a number of recent reports on Arabic literacy and reading pedagogy (Gregory et al., 2021; UAE Ministry of Culture, 2020, 362; USAID, 2019). It also relates to the dominant role assigned more generally to books to develop Arabic literacy, to the detriment of other potential roles of literature in general, and the visual literacy that picture books specifically enable. For the most part, the Arabic curriculum equates to the Arabic language textbook (UAE Ministry of Culture, 2020, 362; Gregory et al., 2021, 30; USAID, 2019, 58), and literary excerpts within scholastic textbooks and required readings, sometimes in the form of picturebooks specially created by curriculum designers to teach particular concepts, seem to constitute the totality of the literary experience of many school children with little time dedicated to independent reading (Gregory et al., 2021, 38–39; USAID, 2019, 18). Some children’s book publishers are also running their own textbook series and publishing other materials to teach literacy (e.g. Antoine’s Bookstore, Lebanon; Nahdet Misr, Egypt; the Hurouf imprint of Kalimat, the United Arab Emirates). Any deviation from this general rule is often an outcome of individual initiatives at the level of educators, schools, or voluntary involvement in external projects that seek to promote reading and children’s literature (see **Other active agents in the ecosystem** section below). Having said that, publishers are reporting an increase in the awareness among schools of the important role of Arabic children’s books and interest in acquiring such books in the past two decades.

Notwithstanding the untapped potential, the scholastic market remains of primary importance to children's book publishers, and better endowed private schools are more likely to curate school libraries, but by virtue of numbers alone, public schools purchase more textbooks. When there is a policy of using public funds to purchase books, ministries of education in countries like the UAE purchase large quantities of selected titles for school libraries or for individual students, with similar processes occurring in some other GCC countries. In less affluent countries, civil society or international aid organizations might furnish these although in a more inconsistent manner. Ultimately, access to high quality books falls along class as well as national lines, and school and class libraries are, for the most part, either absent or poorly resourced and under-utilized (UAE Ministry of Culture, 2020, 362–363, 372; Gregory et al., 2021, 37; USAID, 2019, 61 & 57).

Additionally, education in almost all Arab countries is multilingual, and proficiency in either English or French is highly valued. This may be the outcome of a colonial legacy, cultural globalization, or the influence of the sizable expat communities notably residing in the GCC countries (Zakharia, 2015). This has implications concerning the extent to which private schools are choosing to invest in promoting reading in Arabic and how it fares in comparison with their investment in teaching and promoting reading in European languages. Translation of works into Arabic accounts for a significant part of Arab children's publishing, although we argue that the last two decades of increasingly high quality original production, along with unfavorable currency rates (vis-a-vis Europe) in countries such as Egypt, Lebanon, and Sudan that have experienced devaluation, are turning this tide.

Other active agents in the ecosystem

Beyond publishers and the various actors involved in the educational system, there is a constellation of other organizations as well as individuals that carry out a range of roles, some of which aim to develop children's literature, bring books to children, and promote reading. These organizations could be public agencies, non-governmental organizations, or even academic or research centers. Examples include the National Sections of the International Board on Books for Young People (IBBY) in five Arab countries: Egypt, Jordan, Lebanon, the United Arab Emirates and Palestine. The activities organized by these organizations include book-gifting schemes, book prizes, research and conferences on children's literature, training workshops for authors, illustrators and publishers, the management of libraries, the acquisition of books, and the organization of storytelling and book-based activities with children. In addition to these organizations, there are individual actors including trainers, book critics, storytellers, and edutainment animators, many of whom count on work with private schools, public libraries and NGOs.

One prominent example of such an NGO is the Integrated Care Society which, in coordination with GEBO, launched "The Family Library" project and the "Reading for All" campaign in Egypt in 1991. These State funded initiatives helped establish countless libraries and print millions of copies of curated books, including children's books. The Suzanne Mubarak Prize for Children's Literature connected to these initiatives was highly instrumental in launching a new generation of writers and illustrators in the 1980s and 1990s. The campaign and prize ended with the ousting of former First Lady Suzanne Mubarak in 2011, although the Family Library project is still run by GEBO. Another example is the United Arab Emirates Board on Books for Young People (UAEBBY) which manages the prestigious Etisalat Award for Arabic Children's Literature and carries out training workshops with authors and illustrators among its many other activities. In addition to publishing books for children, Tamer Institute for Community Education in Palestine also organizes training workshops for authors, illustrators, librarians and animators, carries out book discussions and other activities with children, and supports academic research on children's literature.

There are some initiatives that directly seek to support reading at schools. For example, Ana Aqra' Association in Lebanon has succeeded in creating class and school libraries in public schools, training teachers on reading aloud and developing literacy and currently works with the Syrian refugee student population in the country. On a more regional level, the Arab Thought Foundation created a leveled reading scheme with its Arabi 21 project. The Arabi 21 label that categorizes books according to reading difficulty is sought by publishers as these make their books more attractive to schools. Another example is the Arab Reading Challenge launched by the Foundation of Mohammed Bin Rashid, the ruler of the Dubai Emirate, a reading marathon among school children with the purpose of promoting reading. Students keep passport-like reading journals of the books they have read and summarized in 1 year, with winners selected. The competition has undoubtedly galvanized schools to devote greater attention to reading, with millions of students in Arab countries and the diaspora participating, more than half of them elementary students (UAE Ministry of Culture, 2020, 370–372). However, the educational impact of this initiative has yet to be systematically assessed.

Trends and opportunities

From the overview presented so far, it becomes clear that there is, on the one hand, great potential and many passionate advocates for Arabic children's literature and reading promotion. However, this potential is undermined by the lack of sustained financial support for projects and book buying, a comprehensive vision related to the development of children's literature and reading promotion, or deliberate reading promotion policies. We also observe that a major issue in the Arabic picture book ecosystem is the inefficient coordination among its various parts which stands in the way of bringing the book and the child together. Some of the actors that we have discussed above have begun to specifically address this disconnect. In what follows, we will highlight both the current efforts toward greater flow as well as opportunities for collaboration.

1. One major issue is the serious shortage of information on books and where to find them available to children's book mediators (parents, teachers, librarians, animators). In the specific case of teachers, most of them do not receive a solid knowledge of Arabic children's literature, its importance in comprehensive literacy acquisition and innovative literacy strategies in their training (UAE Ministry of Culture, 2020, p. 362). The bridging of this gap becomes all the more critical with an increase in the number of children's books being published which threatens to overwhelm the buyer. As a result, all those who mediate between the book and child can benefit from information not only about new titles but about their quality, pedagogic and esthetic value, and suitability. This can be achieved through accessible resource centers, book lists, bibliographies, reading and activity guides as well as capacity building opportunities for book mediators. In the past, the IBBY National Sections in Egypt and Lebanon have produced some valuable book lists but not in a systematic or regular manner. The last comprehensive bibliography of Arabic children's books across the entirety of the Arab region was *ad-Daleel al-Bibliographi li-Kitaab at-Tifl al-'Arabi* (*The Bibliographical Guide to Arab Children's Books*) compiled by Faisal Hajji and which covers the years 1950–1998 (Hajji, 1990, 1995, 1999). Additionally, Bloomsbury Qatar published *The Golden Guide to Children's Books* in 2015 (Abed, 2015). Most recently, Dalil alqira'a alilajiya (*Bibliotherapy guide*) published by the UAE Board on Books for Young People also provides valuable bibliographic information (Mohamed, 2021).
2. In the same vein, a key issue is the lack of strong cooperative links between civil society initiatives and schools, particularly government or public schools. This is often due to the bureaucracy associated with gaining entry into these classrooms, the lack of resources to build and sustain these links, and the reliance of many of these initiatives on volunteers.
3. Organizations that manage children's book prizes should aim for a more direct relationship with schools and children. In addition to recognizing the work of authors, illustrators and publishers, prizes should also seek to promote reading among children. We take the positive examples of the Stichting Lezen Prize (the Netherlands), Le Prix des Incorruptibles (France) and the shadow Carnegie and Kate Greenaway Awards (the UK) in which those who decide on prize winners are children and young people. In the case of Stichting Lezen Prize (the Netherlands), meetings are also held between the child jurors and authors and illustrators while an extensive school cooperation is sought in order to ensure that the shortlisted works for the Carnegie Greenaway medals are read and discussed by children. There is a similar promising, and replicable Arab initiative, the Kitabi Award organized by the Arab Thought Foundation. Works are shortlisted by a selection committee for three age groups and two categories (fiction and nonfiction) and then read by children who ultimately decide the winners.
4. There is a further need to bridge the gap between the academy and practitioners as well as between academic disciplines. There is a disconnect between academics producing scholarly work on children's literature and literacy, on the one hand, and authors and teachers on the other hand. Children's literature needs to be an integral part of courses for practitioners that work with children, from child psychologists to teachers to mediators, while academics have much to learn from the practical experiences of those who have direct contact with children. In addition, scholars working in one field do not always cooperate with those in other fields, sadly reminiscent of the divide between book and child people observed by John Rowe Townsend over 40 years ago (1980, 199). For example, book critics often make generalizations about the child reader, their psychological needs and linguistic or cognitive abilities, but these pronouncements are rarely fortified by updated and context-specific data. For many authors, information about the developmental and linguistic abilities of children would be highly useful.
5. There is a need for more research on all aspects of children's literature and reading. In particular, we advocate for more empirical research on how picture books are used by teachers, librarians, and other mediators and how they are received by children, as very little is known about reception esthetics or impact. Ideally, such empirical research would be participatory in order to make up for the lack of tradition or support for practitioner enquiry, i.e. practitioners carrying out research on their practice to inform their work. Furthermore, an important link can be forged in the ecosystem if this research can be used to inform national scale educational policy. However, this research would be without value if it were not easily accessible. This is a major issue in Arab countries as much of the research coming out of the schools of education is archived in non-digital repositories and does not feed into academic research. In recent years, the Shamaa Arab Educational Information Network archives educational research, but theses, dissertations, and other resources are listed but not digitally accessible. The difficulties of archival research are bureaucratic in Egypt, while archives in countries such as Lebanon Iraq, and Palestine have been reduced by past or current destruction or looting (Byrne, 2020).
6. Children's authors and illustrators are perhaps the most vital actors in this ecosystem, but they may also be the least supported. This is especially the case of new writing and illustrating talent who are often self-taught and seek feedback through informal or online networks, since training courses, where they exist, are sporadic and brief. This results in work that is sometimes mediocre, derivative, and lacking in local identity, especially in the case of up and coming illustrators. To address this, there is a need for more long-term training courses, including specialized university programs as well as systems that incentivize and create opportunities for official and unofficial mentorship relationships between members of different generations of creators. Such relationships are less likely to grow organically as it is virtually impossible for a children's book illustrator or author to make a living wage solely from their work, and the income they make from public appearances, school visits, brief consultancies, and creative writing workshops is often negligible or non-existent hence the unlikelihood of them setting aside to do potentially monetized work for free. Nevertheless, it is important to acknowledge the role being played by prominent, prize-winning illustrators such as Raouf Karray, Hassan Zahreddine, and Maya Fidawi to nurture a new generation, either through teaching at universities or offering workshops.

7. Parental lack of interest in promoting any reading outside of what can directly advance academic achievement (Gregory et al., 2021, 26–27), and the fact that public libraries are not central to the lives of most children, poses challenges in promoting a reading culture in the Arab world. Again, the revival and multiplication of the multimodal library as a community hub, with trained book mediators can change the public perception of the utility of reading, as well as offset many social ills.
8. It is important to note that while Arabic and its various regional dialects are the dominant native languages in Arab countries, the Arab world hosts a large number of minority languages. Some of these minority languages (or rather specific sub-languages), such as Sorani Kurdish (Iraq), Tamazight (Algeria), and Armenian (Lebanon) have achieved the status of official languages in certain countries, but this is not the case of other languages, such as Nubian in Egypt. There is some production of children's literature in minority languages, whether these books are originally written in these languages or made available in them through translation. For example, many Algerian children's books would be available in Tamazight, while Kotot publishers in Lebanon publish Armenian children's literature. Nevertheless, this is an area of Arab publishing that bears developing.

Inside the book

After discussing issues related to the publishing, distribution, promotion, mediation and research on Arabic picture books, we turn to examine the Arabic picture book itself and its content. In particular, we will focus on language, an aspect that provokes more debate than any other in Arabic children's literature; rhyme and poetry as esthetic elements whose pedagogical uses require more attention; and illustration and type.

Language

In addition to the esthetic pleasures they afford, picture books are often considered for the role they may play in promoting literacy at the level of teaching letter and word recognition, building vocabulary, and developing the abilities of children to both comprehend and compose "correct" spoken and written texts. In the case of Arabic picture books, the discussions over this role take place against the backdrop of other ideological discussions related to the diglossic nature of the Arabic language.

To use the term popularized by the American linguist Charles Ferguson (1959), Arabic is a diglossic language in that there is great divergence between the standard written language and spoken dialects, which in themselves differ from one country and region to another. This divergence is significantly wider than that which exists in other languages such as English, and is mostly at the level of grammatical structure and vocabulary. In terms of vocabulary, there is, on one hand, great overlap with words that are either identical or very similar both phonetically and lexically between the written and colloquial varieties, but with equally significant differences (USAID, 2019, 10–12). These overlaps and differences are not stable, but shift depending on the Arabic dialect in question, and even written output in Standard Arabic is often peppered with word choices and expressions that reveal the nationality of the author.

In the oft-reprinted article "Living in Arabic", Edward Said observes that the "educated [Arab] person has two quite distinct linguistic personae in the mother-tongue" and describes his own linguistic path between the colloquial and Standard Arabic. He recognizes the difficulty of mapping the complexity of the relationship between the two tongues and "the way the language is lived in, deployed, and experienced by native speakers and users" onto a Western context (2004). For the most part, Arab children growing up in an Arab country will only begin to be exposed to Standard Arabic in a concentrated manner in school, unless they are also undergoing religious instruction entirely in Standard Arabic. In addition, Standard Arabic is the only language variety acknowledged in educational curricula. It is for this reason that children's books are considered primarily pedagogical tools to teach literacy, and there is some research evidence that reading children's literature helps children make the transition from colloquial to Standard Arabic (Gregory et al., 2021, 26). Arab children growing up in non-Arab countries do not have the same access to or instruction in Standard Arabic, and pick the dialect up passively from their families and possible communities.

Given this diglossic situation as well as the changes that Standard Arabic has undergone over centuries, authors and publishers have choices to make regarding the register or language level they write or publish in as well as in the incorporation (or not) of colloquial words or expressions. In this case, language ideologies figure in an important way. Mathilde Chèvre (2016) noted that children's authors had different viewpoints regarding the "proper" language to use in children's books and identified two main orientations. Some authors opt for an elevated and sometimes challenging language, the mastery of which allows the child to access Arab literary, religious, and historical heritage; others prioritize the promotion of reading for pleasure and favor a more accessible language that can act as a bridge between the spoken and the written.

Nowhere is the debate over language more ideologically fraught than when it comes to the use of or authorship of children's works in a colloquial variety. These disagreements are connected to wider debates over national as well as literary languages that go back for centuries and continue to flare up periodically (Al Jahiz in 800; Osman Sabri and Louis Awad in 1960s, Tawfiq al-Hakim in 1990s, Madiha Doss and Sohair Soukkary in 2000s). Since the start of the twentieth century, these debates have been shaped by attempts to define national identity, especially in times of political unrest or transition. The tension between a pan Arabist political orientation and political movements that emphasize a nation's singularity and separateness, and more recently, globalization and associated fears of cultural attrition and Western cultural imperialism, all have bearings on this discussion.

Given the current state of publishing and education in the Arab world, the sole use of Standard Arabic in children's books is certainly the more expedient choice. Arab Ministries of Education often show serious intolerance to the appearance of colloquial words and expressions in textbooks and children's books. In addition, most publishers also limit themselves to Standard Arabic, not only because schools are their major market but because this allows books to be more versatile, and to sell well in various parts of the Arab region.

While it is not unusual for authors or publishers producing works in colloquial dialects to be criticized, there have been some initiatives to produce books written completely in colloquial Arabic by Arab writers based both inside and outside the Arab world. Those in the diaspora are mainly responding to the absence of books accessible to their own children, whom they wish to be able to understand and communicate in a spoken colloquial but are not necessarily training to read Standard Arabic in the future. Examples of these are the US-based Qossas Stories, a publishing house established by the Palestinian journalist Rym Makhoul and her husband, who have published two books, each made available in the Levantine (spoken in countries such as Lebanon, Syria, Jordan and Palestine in the Western Mediterranean) and Egyptian dialects with the hope of expanding into other dialects. *Kan Ya Makan* (2020) by Riham Al-Shendy is also a well publicized self-published collection of well-known fairy tales retold in rhyme in the Egyptian dialect. The stories have also been "Egyptianized", with each tale set in a different part of the country and incorporating elements of local culture. Shendy has since published two picturebooks also in the Egyptian dialect while in 2021, the publishing house Liblib was launched with the aim of publishing children's literature in various Arabic dialects.

Given how heated the discussion between including (or writing completely in) and excluding colloquial young children's picture books can be, it is ironic that those on different sides of this linguistic debate both appear to consider diglossia as a problem to be solved or a zero sum game. We adopt a different position and view diglossia, differences in national dialects as well as the register within Standard Arabic as a resource and a site of possibility and experimentation. The "distinct" linguistic personae to which Said refers do not have to be in conflict. In addition, it is important to avoid simplistic, often false, binary thinking concerning these two language varieties, with Standard Arabic seen as beautiful but also difficult and potentially alienating while colloquial varieties are seen as more accessible and likely to appeal to children, but less eloquent and pedagogically valuable. Finally, there is the need to respect and accommodate the diverse abilities of children. Further understanding of these issues requires large scale studies as well as the accumulation of small scale studies across contexts and with different groups of children, in addition to the studies that currently exist.

As a result, we advocate moving past this debate in order to raise relevant questions with concrete implications and for multiple strategies to be experimented with and researched. For works primarily or exclusively in Standard Arabic; what choices can be made to balance the need to expand the child's vocabulary and introduce the structure of Standard Arabic in preparation for adult reading with presenting the child with a language that is not excessively distant from the one they use in everyday life or beyond their existing reading competence level? In addition, does the occasional use of colloquial words and expressions lead to Standard Arabic language attrition? What place is there for colloquialisms? A cursory research would reveal many classic works of Arabic literature contain colloquial language or reveal local usage without compromising their classic status.

Those who wish to use colloquial Arabic face a different set of questions. For instance; how does one "transcribe" the colloquial? Should this be done phonetically or should the colloquial be written in a manner similar to Standard Arabic? How does one take into account the diversity within dialects themselves without privileging certain dialects in ways that reproduce class and regional inequalities and stereotypes? In addition, if writing in colloquial language might make a book unmarketable in the larger Arab region, how can the publisher profit from national markets alone? What strategies can publishers adopt if they do not want to limit themselves to local markets? Some strategies that have been used to tackle these operational difficulties include publishing versions in other dialects or Standard Arabic as well as including glossaries at the start of the books such as in Samah Idriss's (2004) *A Boy from Beirut* series. This brings us to another, more ambitious question: how can books feed the Arab child's curiosity about other Arab dialects while remaining hospitable and inclusive?

In addition, it is important to move away from simplistic ways of viewing each language variety. Eloquence is not the sole quality of Standard Arabic but could equally be a feature of colloquial Arabic in much the same way that Standard Arabic can be accessible and playful. In addition, the use of colloquial language is not enough to make a book more appealing. Rather, what distinguishes one book from the other is depth or humor, innovative plots and language use that is audience-aware, and faithful to both the character and the setting.

The debate, too often, centers heavily on comprehensibility rather than the potential of language to esthetically engage the child. Therefore in lieu of asking: can it be understood? We might do well to ask: Does it paint an image? Does it move the child? Does it express the child's voice? Does it expand the child's horizons and perceptions? Does it give the child rich language with which to express the world? How does it fit with other elements of the text or the visual text? In certain books, there is a serious need to align level, subject matter, and form. The quantity of text and language level often exceed the abilities of the target reader of the picture book form it employs, even while the content is childish in some cases. This results in books that are not really likely to be used by mediators, as picture books selected for use with child readers are typically the ones that lie within the range of 500 words. Additionally, the child most likely to select picture books would not have the reading competence to navigate them or feel addressed by their content.

Poetry and rhyme

Arab children have access to poetry and rhyme in a number of texts. The most prominent of these are school textbooks which often include classic poems for children, such as those by Egyptian poet Ahmad Shawqi, or poems originally written for adults considered at the level of theme and language to be appropriate for children. Sometimes, they include original poems composed by textbook compilers, Ministry of Education officials, or poets and writers commissioned to the task. Children's magazines are another venue for poetry, such as *Tuta Tuta* (Lebanon); *Mizmar* (Iraq) and *Samir* (Egypt). In addition, there are a few poetry collections as well as rhymed stories published as standalone picture books. Last, but certainly not least, there are children songs, some of which have become regional classics (for example, the songs of Mohamed Fawzy) to the more modern ones sung by pop stars or actors (Nancy Ajram, Ahmed Helmy) or popularized by Arab satellite channels targeting children.

Some of the features that characterize certain children's poetry that has achieved canonical status is their heavily ideological and didactic nature, with heavy handed nationalistic and religious messaging packaged in an often artless delivery. In addition, they rely on a heavily romantic vision of childhood and repetitive themes, often related to nature, the nation, or the child's everyday life (Al-Safadi, 2008). In terms of language, some questions can also be raised as well in terms of imagery and voice. Many poets do strive to simplify and economize verbiage in their poems, including prominent Syrian poet, Suleiman Issa, as well as other big names such as Mohamed Jamal Amro (Jordan), Haidar Ghazi (Iraq), Jalil Khazaal (Iraq), etc. Given the relative scarcity of new established children's poets, these older poems continue to dominate and represent a useful corpus from which to selectively draw.

Another issue for educators is that the rather scattered manner in which individual poems are published may make them difficult for teachers to find and collect. Poetry collections can begin to address this gap. After the closure of its magazine *Tuta Tuta*, Dar Al-Hadaek started to publish collections of some poems previously published by the magazine, such as those by Jawdat Fakhreddine (Lebanon) and Bayan Safadi (Syria). It is hoped that similar efforts to collect poems originally published in other magazines may materialize. In general, collections of poetry are currently quite rare, and some of the few examples include *Hilm* [A Dream] by Sulafa Hijawi (2009) published by Tamer Institute, Palestine and *Qasr min Al-Rimal* [Sand Castle] by Haidar Ghazi (2012) published by Dar Ashjar, UAE.

In addition, a number of authors are publishing stories in rhyming texts as standalone picture books. However, some of these texts often compromise on musicality and meter, as well as clarity and logic when committing to the use of rhyme in a full length text. Some successful examples of rhymed stories include Alaa Mortada's *Amgad fi Akoun Aw la Akoun* [Amgad in ... To Be or Not To Be] (2018), a humorous tale of the efforts of Amgad, the titular character to find a vocation, often to disastrous ends (Dar Al-Balsam, Egypt) as well as *Daraja Ala Al-Hibaal* [A Bicycle on the Tightrope] (2016) by Elias Zougheib and illustrated by Hatem Fathy Ali, in which a rabbit and a rooster decide to leave the farm where they live to join the circus (Librairie Antoine, Lebanon). Hadil Ghoneim from Egypt uses Standard Arabic that draws from the colloquial lexicon to create the most modern and poetic alphabet book in the new wave 'id Milad' Harf Al-Daad [The Birthday Party of the Letter Daad] (2018). The republishing of some of Mahmoud Darwish's poems not originally intended for children as picture books by Tanmeya Publishing House (Egypt), such as *Fakkir bi-Ghayrak* [Think of Others] (2018) illustrated by Sahar Abdallah, may herald more efforts to repackage and introduce poems by prominent Arab poets to a younger generation. There are also some deliberate attempts to awaken the child's sense of the poetry around them, and encourage them to appreciate the intrinsic esthetic of the poetic Arabic language, and the earth, the seasons, and the sensual as source of inspiration and the emancipation of the imagination, such as the works of the Egyptian poet Mohammad Afifi Matar, whose experimental poetry collection *Catching pigeons* (Sayd al-Yamam) was published posthumously in 2002 by the Department of Culture and Information in Sharjah, the United Arab Emirates (Motawy, 2010).

Poetry and rhyme are under-exploited whether as language acquisition tools, for their ludic potential or as a vehicle for children to see themselves as authors of poetry. Some hoped for future directions in the area of poetry are certainly for more of it to be composed, published, and anthologized. Qualitatively, there is room for it to expand into new themes and away from purely romantic ideas of childhood and cliched images toward integrating the modern, the abstract, and the playful. There are ample opportunities to innovate and experiment at the level of form.

Illustration and type

It is difficult to provide any semblance of an overview of the illustrations in Arabic picture books for the sheer diversity in styles and media, ranging from the closely related to the rich Arabic tradition of visual and calligraphic arts to the clearly anime/manga-inspired, from the hyper-realistic to the more stylized. While some illustrators draw by hand, many others are resorting to digital means and the use of graphic tablets. The artistic backgrounds of illustrators are also diverse. Some come to the field of drawing for children via a training in either the visual arts (as is often the case with Egyptian illustrators such as Ahmed Soliman) or graphic design (the Sudanese Salah Elmur); many are also active in other art forms, such as comics (Lena Mirhij, Ahmed Wahba, Magdy Shafei) or animation (David Habchy) which they then creatively integrate into their work. The recent trend among Arab publishers to work with non-Arab illustrators, primarily from Iran but also from European and South American countries, has made Arabic picture books more globalized in content and visually similar to other international offerings at the expense of a local particularity. While illustrations in children's books have received much less research attention than textual content, exceptions include work by Chèvre (2015), Dinges (2011), Bizri (2011), Al-Zayyat (2011), Batran and Ihshayesh (2011), and Motawy (2021).

In this section, we will focus on one aspect of illustration, type, given the rich heritage of Arabic calligraphy styles. While innovation in traditional Arabic types is enmeshed in tension, the last two decades has seen reasonable experimentation in children's book types. Innovation in children's type and scholarly research on Arabic typography has come from children's book artists as well as academics/artists such as Huda AbiFares, Nadine Chahine, Bahia Shehab, and Mohamed Hassan.

While most picture books selected for scholastic purposes present the core text in traditional fonts, that are equivalent to Arial or Times New Roman, speech/thought bubbles, sounds, front and back matter are often where scripts are played with the most. Some incorporate stylized types that draw attention to the esthetics of Arabic calligraphy itself. Examples include *Waraqat Al-Hayat* [The Paper of Life] (2014) illustrated by Ahmed Soliman and James Rumford's *Silent Music* (2008; 2015) available in an Arabic translation. The type and the rest of the book are most likely to be coherent "iconotexts" when the author/illustrator are the same person, with some iconic handwritten types becoming as recognizable as Quentin Blake's handwriting for instance, such as those developed by the late Mohieddine Labbad (Egypt), Lena Merhej (Lebanon), Raouf Karray (Tunisia) and Walid Taher (Egypt). In some books that revive tradition, or break with it to experiment with the modern, calligraphy is often a character in the text, such as in many of the books published by Dar Onboz in Lebanon and Dar Al-Balsam in Egypt.

As design in children's books is still developing in the Arab world, designers are still finding their place in the ecosystem and the links and lines of communication are still developing between children's publishers and designers. Children's book publishers are also still navigating layout and design. Designers need to understand the specificity of the field, and as Arabic typography research gains momentum, for example through the creation of Typelab at the American University in Cairo in 2020, children's books need to be a part of this research in much the same way as children's illustrators were pioneers in this field.

Type is a clear pedagogical concern for its role in cultivating esthetic appreciation, enabling children to recognize their rich artistic heritage, and directing meaning making in reading. However, experimentation with form should not come at the expense of legibility. The work of Chahine, in particular, addresses accessibility, as her doctoral research tracks children's eye movement in reading Arabic texts to find fixations and saccades in the reading and hypothesize on the impact of complexity of Arabic types on the legibility of text for children. She argues that increased complexity reduces legibility and that the inclusion of short vowels slows reading (see also Gregory et al., 2021, 24 and USAID, 2019, 13–14). Chahine goes on to conclude, "however, the literature review has shown that early exposure to Modern Standard Arabic facilitates the process for children first learning how to read Arabic" (2012, 224). Another issue is that some publishers and designers may choose fonts that require longer fixations to decipher which also slows down the reading process.

Conclusion

We end this section by addressing the move toward the digital in children's books. Some Arab publishing houses, both established and new, such as Kadi and Ramadi, Aliareach, Nahdet Misr, Tuta Tuta and others are transforming picturebooks into digitally enhanced and more interactive ebooks. Other creators such as Lamsa and Bees Studio have produced digital stories that did not originate as books. There is an expansion of potential in the ecosystem and potential links to established Arab adult e-book apps such as Kotobi, Jamaloon, or iKitab as well as those specifically for children's books, such as Kutubee and Noory Books. Without the implementation of more technological and legal barriers to piracy, it is difficult to encourage platforms or other parties to invest in research on optimizing the Arabic picture book as an e-book.

In concluding this overview, our primary overriding recommendation is for greater synchronicity and interconnectedness in the picture book ecosystem. In terms of form, in addition to poetry, there is room for much more writing in the genres, fantasy, science fiction, non-fiction, and new approaches to the ever-popular social realism. Finally, experimentation in language use and type, as well as digitization, would benefit from closer coordination with those conducting participatory research in the Arab world.

References

- Abed, A., 2015. Golden Guide for Children's Books. Bloomsbury-Qatar Foundation, Doha.
- Al-Safadi, B., 2008. Shi'r al-attfal fi al-watan al-arabi [Children's Poetry in the Arab World]. Syrian Ministry of Culture, Damascus.
- Al-Zayyat, G., 2011. Surat Al-Tofl fi-Adab Al-Atfal: Al-Rosoum fi Kotob Al-Atfal Al-Musawwara Ala-Masriyah [The Image of the Child in Children's Literature: Illustrations in Egyptian Picture Books]. A Study Commissioned by the Arab Children's Literature and Reading Program at the Anna Lindh Foundation.
- Arab Thought Foundation (ATF), 2008. Al-Taqrir Al-"Arabi Al-Awwal li"-I-Tanmiya Al-Thaqafiyya [The First Arab Report on Cultural Development]. Arab Thought Foundation, Beirut.
- Batran, Y., Ihsayish, W., 2011. Surat Al-Tofl fi Rosoumat Kotob Al-Tofl Al-Falasteeni [The Image of the Child in the Illustrations of Palestinian Children's Books]. A Study Commissioned by the Arab Children's Literature and Reading Program at the Anna Lindh Foundation.
- Bizri, H., 2011. Surat Al-Tofl fi-Adab Al-Atfal: Al-Rosoum fi Kotob Al-Atfal Al-Manshura fi Lobnan [The Image of the Child in Children's Literature: Illustrations in Children's Books Published in Lebanon]. A Study Commissioned by the Arab Children's Literature and Reading Program at the Anna Lindh Foundation.
- Byrne, K., 2020. A Survey of Middle East Archives: Egypt. Wilson Center. www.wilsoncenter.org/blog-post/survey-middle-east-archives-egypt. (Accessed 29 December 2020).
- Chahine, N., 2012. Reading Arabic: Legibility Studies for the Arabic Script. Doctoral thesis. Leiden University.
- Chèvre, M., 2015. *Le poussin n'est pas un chien* [The Chick is Not a Dog]. Presse de L'IFPO, Paris.
- Chèvre, M., 2016. Lire le livre: Quelle langue arabe écrire pour que les enfants aiment lire? [Reading the book: Which Arabic language should we write so that children would want to read?] In: Majdalani, C., Mermier, F. (Eds.), *Regards sur l'Édition dans le Monde Arabe* [A Look at Publishing in the Arab World]. Editions Karthala, Paris, pp. 145–166.
- Dünges, P., 2011. Arabic Children's Literature Today: Determining Factors and Tendencies, vol. 126. PMLA/Publications of the Modern Language Association of America, Cambridge University Press, pp. 170–181, 1.

- Eschweiler, J.C., Goehler, A., 2010. Book distribution in the Arab world. *Publish. Res. Q.* 26 (3), 193–201.
- Ferguson, C., 1959. Diglossia. *Word* 15 (2), 325–340.
- Gregory, L., Taha Thomure, H., Kazem, A., Boni, A., El-Sayed, M.A.A., Taibah, N., 2021. Advancing Arabic Language Teaching and Learning: A Path to Reducing Learning Poverty in the Middle East and North Africa. International Bank for Reconstruction and Development/The World Bank, Washington, D.C.
- Hajji, F., 1990. *Ad-Daleel al-Bibliograafi li-Kitaab at-Tifl al-'Arabi* (The Bibliographical Guide to Arab Children's Books). Da'irat ath-Thaqafa wa-l-i'lam, Sharjah.
- Hajji, F., 1995. *Ad-Daleel al-Bibliograafi li-Kitaab at-Tifl al-'Arabi* (The Bibliographical Guide to Arab Children's Books). Da'irat ath-Thaqafa wa-l-i'lam, Sharjah.
- Hajji, F., 1999. *Ad-Daleel al-Bibliograafi li-Kitaab at-Tifl al-'Arabi* (Bibliographical Guide to Arab Children's Books). Da'irat ath-Thaqafa wa-l-i'lam, Sharjah.
- Kenawy, H., 1994. *Al-tifl wa Adab Al-attal* (The Child and Children's literature). Maktabat al-Anjlu al-Misriyah, Cairo.
- Kreidieh, S., 2015. Publishing and Marketing Lebanese Children's Books in the Arab world. PhD dissertation. Oxford Brookes University.
- Mohamed, E., 2021. In: *Dalil alqira'a allajiya* (Bibliotherapy guide). UAE Board on Books for Young People, Sharjah.
- Motawy, Y., 2010. Shaking up Egyptian children's literature: the stories of Muhammad Afifi Matar. *Int. Res. Child. Lit.* 3 (2), 217–221.
- Motawy, Y., 2021. *al-Sukun ma bayn al-Amwaj: Kutub al-Attal al-Musawara wal-Mujtama' al-Misri al-Mu'aser* (Stillness Between the Waves: Egyptian Children's Picturebooks and Contemporary Egyptian Society). Balsam Publishing House, Cairo.
- Said, E., 2004. *Living in Arabic*, vol. 677. Al Ahram Weekly.
- Stephan-Hachem, M., 2010. *Al-Kitaab Nashran wa Intishaaran* [Book Publishing and Distribution]. Dar Al-Nahda Al-'Arabiyya, Beirut.
- Townsend, J.R., 1980. Standards of criticism for children's literature. In: Chambers, N. (Ed.), *The Signal Approach to Children's Books*. Kestrel Books, London.
- UNICEF Global Database on Literacy Rate, 2021. Available at: <https://data.unicef.org/resources/dataset/education-data/>.
- United Arab Emirates Ministry of Culture and Youth, 2020. *Taqreeru Halati al-Lughati Al-Arabiyyati wa Mustaqbaliha* (Arabic Language Status and Future Report).
- USAID, 2019. Arabic Literacy and Numeracy State of the Art Conference Desk Review Final Report. Prepared independently by Dr. Helen Boyle on behalf of Social Impact Inc.
- Zakharia, Z., 2015. Language, conflict, and social change: Arabic bilingual education in the Middle East and North Africa. In: Wright, W.E., Boun, S., Garcia, O. (Eds.), *Handbook of Bilingual and Multilingual Education*. Wiley-Blackwell, Hoboken, NJ.

The scholarship of African Children's literature: issues of representation of critical voices in the conversation

Vivian S. Yenika-Agbaw, Phoebe L. Quaynor, and Jocelyn Amevuvor, Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	688
Critical evaluation of studies and essays/articles surveyed	689
Research and literary criticism on African children's literature	690
The scholarship of annotations	690
Representations of Africa in children's literature	691
Publishing of books with Black and Brown faces	692
Accessibility issues	693
Gaps in the scholarship	694
References	694

Introduction

We open this article by drawing attention to the complicated nature of what we consider literature about African children (LAAC) and/or African children's literature (ACL). Some scholars refer to LAAC or ACL simply as children's literature about Africa (Schmidt, 1981); while others refer to it as African children's literature (Osa, 1985). Nevertheless, Osa (1985) posits, "A true canon of written African children's literature is yet to emerge" (p. 753). We agree with him that African children's literature is a subgenre of African literature that is consistently ignored (Osa, 1985). We would add that the scholarship of LAAC and ACL is also often ignored. For the purpose of this article, we will refer to this genre of literature as ACL while acknowledging that the term can sound generic and all-encompassing—not necessarily capturing the complexity and plurality of stories in ACL. Still, we select this term to align with previous researchers of the field, like Osa (1985).

The terms ACL and LAAC are generic terms educators, librarians, and scholars often use to refer to literature that centers African children's stories, regardless of who is telling/writing the story and the indigenous or colonial languages the author/storyteller chooses to share the story. Oral African children's literature predates its written version, as folklore has existed in Africa for centuries. Thus, we consider this subgenre an authentic African children's literature. The audience for the most part is African peoples including its children. As Osayimwense Osa notes in his introduction to the last issue of the *Journal of African Children's and Youth Literature* (JACYL, 2007–2009—volumes 17 and 18), one of the first journals that focused on ACL,

For generations the people of Black Africa have told stories in their numerous and multitudinous indigenous tongues to children and young people and to adults as well, in their social, informal, or formal gatherings. Storytelling is not a new art for them but one as old as the continent itself (2007/9; p. 1).

He asserts further that: "The present body of African oral tales that have been committed to print are a very insignificant proportion of the total body of folktales on the African continent" (p. 2). However, the issue at stake in this review are the written stories, regardless of the genre. As we begin this review, we ask: For whom are the stories written, and what might be some possible rationale and/or implications for writing these stories as outsiders and for outsiders, as insiders for outsiders, and as outsiders for insiders? Finally, where does this position African children in the continent whose stories are being circulated across the globe?

Advocating for joint conversations and literary criticism that considers African, African American, and Caribbean children's literature, Osa (2007) notes that "Black people in the African diaspora have a common ancestor and a common heritage—Africa before its contact with the outside world and before slavery and its aftermath" (p. 4). In this chapter, we acknowledge essays that focus on these connections but also point out those that highlight the subtle differences that permeate these diverse groups of Black people and how their historical trajectories have contributed to shaping the contents and outcomes of the stories that are published in print or constructed in other modes for consumption by children within our ever-evolving global community. Thus, the "sociocultural connections" (Osa, 2009, p. 4) continue to exist, but diversity in how each Black group navigates space is also apparent and shaped by their unique sociocultural histories as Yenika-Agbaw (2013) demonstrates in "Black Cinderella: Multicultural Literature and School Curriculum."

When it comes to ACL that is published, there are many perspectives through which LAAC and ACL are crafted and distributed. For instance, LAAC and ACL are sometimes written by westerners and published in the west; or published in the continent and marketed to the west, published in the continent and distributed within the continent, or published in the west and reissued in Africa. Some of the scholarship surrounding ACL centers on the issues related to publishing and who is telling these stories. Now that we have provided some background on our perspectives of ACL or LAAC, we will focus on the main issue at hand for this review, scholarship on ACL or LAAC.

Early scholarship, such as anthropologist Schmidt's (1965) ["Children's books about Africa"] and (1975) *Children's books on Africa and their authors: An annotated bibliography*, where surveys of books for children available in the west are discussed, paved the way for serious scholarship on ACL. Schmidt's (1981) ground breaking book, *Children's fiction about Africa in English*, which focuses on fiction, built on this foundation. This body of scholarship was then expanded by other scholars and educators such as Osa (1985, 1987, 1995), Khorana (1994, 1998), MacCann and Maddy (2013, 2008, 1996), Yenika-Agbaw (1996, 2007), 1996) and Yenika-Agbaw and Napoli (2011). Of recent, other scholarly voices have joined the conversation; however, from our perspective, these contributions seem sporadic and, as Thomas (2020) notes, "The outlets of African scholarship in our field are slim [and] [l]ocation [of the scholar also] matters" (p. 11). We find it troubling that scholarly interest in LAAC and ACL seems more about tokenizing LAAC and ACL, and we hope this trend will eventually shift as genuine interest in diverse books for children and young adults continues to flourish.

Sadly, misinformation about Africa seems endemic even in the 21st century, a problem that had caught the attention of several North American Africanists during the mid-nineteenth century. According to Kuntz (1997), one such Africanist and a longtime jury of the Children's African Book Award (CABA) committee, on reviewing the curricula materials on Africa for K-12, a select group of educators noticed that "stereotypes and misinformation" were prevalent (p. 4). This was also confirmed by surveys later conducted in the 1970s by Africanists who observed that:

stereotypes generated in North America and Europe during the 19th and 20th centuries prevailed in textbooks and trade books of pre-collegiate readers. Misinformation about Africa was so pervasive that they had to spend most of their time countering stereotypes rather than administering other outreach services (pp. 3–4).

While there is progress in regard to the overall quality of books about Africa that are being published, stereotypes persist according to the body of scholarship that we read. What follows next is our critical evaluation of existing scholarship on the topic under study.

Critical evaluation of studies and essays/articles surveyed

This critical review discusses empirical and theoretical essays/articles based on studies conducted on African children's literature (ACL) between 1990 and 2020. We chose to focus on these three decades primarily because the 1990s saw the rise in scholarly interest in ACL. Other reasons include the publication of more books about African children and the continent, the emergence of Children's Africana Book Award (CABA) as a prestigious award, in its own right, recognized and housed by African Studies Association (ASA) and heavily supported by Howard University (a prestigious HBCU institution) as part of its outreach program. These decades also encompass the period when the public began seeing the proliferation of independent publishing houses, such as the highly successful Sub-Saharan Publishers in Accra, Ghana, and Lantana in the United Kingdom, that focus on African children's literature. Book fairs across the globe are now sensitive to the exclusion of quality books about Black people. Finally, the We Need Diverse Books movement also drew attention to the lack of books about Blacks more generally, including Black Africans. Thus, the early 21st century heralded an atmosphere of the modern globalization era that boosted transformative practices within the publishing industry and, as such, slowly paved the way for partnerships to blossom between some publishers in the west and those in Africa.

During these decades much has happened to change the landscape of the children's literature field drawing attention to ACL. For instance, the *Journal of African Youth Literature* (JACYL) gained more visibility, *Sankofa: The Journal of African Children's and Young Adult Literature* was launched by Morgan State University with Meena Khorana as the founding editor. An outreach program, Africaaccessreview.org (established by Brenda Randolph a librarian in 1989) gained in popularity as a site where educators could locate books on Africa reviewed by Africanists. Additionally, there now exist more books on the African diaspora that deliberately connect people of African descent with the continent. Because these have become readily available even though in low numbers, they are helping to expand our repertoire of what it means to be Black and African in the 21st century and across the globe.

With all of these exciting initiatives we were hoping that the scholarship on ACL would have evolved in similar ways, privileging voices of Africans both as scholars/researchers and authors in the west and in Africa. Unfortunately, this is not so, as one can deduce from Thomas' (2020) passionate and intellectually invigorating essay on this vexing issue. Thomas rightfully observes, "Black African and African diaspora literature, media and culture for youth and young adults has heretofore been largely out of the hands of scholars of African descent" (Thomas, 2020, p. 3). This therefore begs the question, why? Exploring this question, in her article, Thomas "trace[s] traditions of research and criticism of African diaspora young adult literature in four regions" including Africa, the

continent of interest in this article (p. 3). We are continuing to search for answers by considering research traditions and criticism across what is considered African children's and young adult literature—often lumped together in most of the studies. Thus, the following research questions guide our review:

- What topics seem to be of interest to researchers undertaking empirical studies and/or theoretical research on African children's literature?
- What do the researchers and scholars assert about their observations of how Africans are portrayed and analyzed in the literature?
 - *Sub-question:* How have studies about African children's literature analyzed their depiction in children's books?

Research and literary criticism on African children's literature

Studies and essays were located primarily from searches done in four databases: JSTOR, Google Scholar, Proquest and ERIC. We examined 50 articles/book chapters/books and what we present here is representative of the findings, in which we focus on a select 17 of the articles that encompassed the themes we consistently saw. Some essays first appeared in journals and subsequently as parts of books, for instance, Osa's (1987) "The rise of African children's literature" and Yenika-Agbaw's (1998) "Images of Africa in children's books: Replacing old stereotypes with new ones." We avoided book reviews, commentaries in blogs, and articles that focused on reading materials, such as excerpts for textbooks that serve as curricular materials.

Within the selected essays and book chapters we reviewed, we found a variety of discourses. Our first finding was that African children's literature, whether directly or indirectly, seem to be aligned with and influenced by the colonial experience, values and timeline. In other words, we often observed postcolonial or neocolonial lenses, as influenced by the colonial history of the African contexts. Depending on who is writing, children's literature about Africa is either described as promoting imperialistic, white supremacist values or is instrumental in restoring, reconstructing or reinforcing post-colonial cultural or national identity. This latter task of reconstruction happens in various ways.

The earlier articles such as those by Schmidt and Osa often began with a discussion about the various ways that children's literature is defined. The range of issues under this topic includes: Who is a child? How is childhood defined? Who is writing stories for children (implying adults as the spokesperson)? They also convey an urgency to decolonize the field for it to be truly considered African children's literature but also acknowledge the persistent challenges in this task of decolonization and reconstruction, including limited indigenous voices (creative writers and as scholars), publishing issues, market forces, audience, ideologies of readers and authors.

The representative essays selected for this review cover six major topics as noted below; however, our discussion will revolve around the first four, which we believe are quite pertinent given the realities of the modern global society:

- Annotations of books (Schmidt, 1975, 1981; Khorana, 1994)
- Issues of representations as informed by postcolonial and/or neocolonialism (Osa, 1983—Nigeria; Khorana, 1994, 1998—the continent; Yenika-Agbaw, 2007—west Africa; Yenika-Agbaw and Napoli, 2011—west Africa; Yenika-Agbaw, 2018—the Nso people of Cameroon; Maddy and McCann, 2013—South Africa); Heale and Williams (2013—white South Africans); Representation and environment (Jenkins, 2004); Representation of African cities (Kuntz, 1997);
- Publishing challenges (Osa, 1987, 2010; Joubert, 1991; Nwapa, 1997), and the dynamics of translation (Kruger, 2012); and
- Accessibility issues and the integration of African children's literature in school curriculum across the globe (Osa, 1987).
- Values encountered in African children's books published in the west (Fleishack and Meehan, 2014; Ntuli and Mafela, 2013);
- Quality of African children's literature (Williams and Heale, 2013)

Before we share our observations regarding critical conversations on ACL, we need to confirm our hypothesis of fewer indigenous scholars from the continent participating in these scholarly conversations. We would also like to draw attention to the fact that the bulk of the literary criticism emanates from scholars who are located in South Africa, followed by those in Nigeria. Even including those South African and Nigerian scholars, the contribution of African scholars to the ACL field is still lean, demonstrating the lack of scholarly voices coming from the African continent.

The scholarship of annotations

The scholarship of annotations is a good place to begin since many educators often complain about not being able to locate African children's literature in their local libraries and/or bookstores. Thus, Nancy Schmidt's pioneering book, (1975), *Children's books on Africa and their authors: An annotated bibliography* paved the way for scholars with interest in this topic. In her book, she discusses the kinds of stories that existed in print, focusing on who the authors were. While a good resource at the time, as it got publishers to pay attention, it was limited in scope. Almost two decades later, Khorana's (1994) *Africa in literature for children and young adults: An annotated bibliography of English language books* built on this foundation, expanding the entries to include not only a variety of genres but also an introduction that critically analyzed the overall quality of the books that were accessible to the public. Like Schmidt, Khorana had to canvas the streets looking for books by indigenes that were published in independent bookstores and so were not readily available in mainstream libraries or bookstores.

Representations of Africa in children's literature

By far the largest form of scholarship revolves around the issue of representations.

As early as 1981, when Schmidt's groundbreaking book caught the public's attention was published, it was already clear that biases existed in books that attempted to portray Africa and her children. In her book, instead of focusing on the authors as she did in the earlier annotation, she adopted a more critical lens, albeit from her anthropological perspective, pointing out the discrepancies in the way African characters were portrayed in the books around specific themes. Osa (1983) remarks in his review of Schmidt's 1981 book,

Schmidt identified a related problem inherent in the study of juvenile and other literature of Africa. The bibliographic task here is formidable, she points out, since the literature is written in many languages in Europe and in Africa... By thematically comparing and contrasting fiction written for Euro-American children with that written for African children, Schmidt shows that biases about Africa are more predominant in Euro-American books (p. 259).

This is how endemic stereotypes about Africa are. One would imagine that things would have improved years later after Khorana's annotated bibliography (1994) was published, but it wasn't so as her next book, *Critical perspectives on postcolonial African children's and young adult Literature* published in 1998 would demonstrate.

Maddy and McCann (1996) use a neocolonial lens to discuss common stereotypes that are pervasive in books about Africa, pointing that often times these books receive awards despite the misinformation that they spew about the continent and her children. They continue this line of inquiry in their 2008 book, *Neo-Imperialism in children's books about Africa: A study of contemporary fiction*. Situating their analysis within a historical context, they discuss biases in books by well-known western and/or South American authors such as Nancy Farmer and Isabel Allende. In this way they are able to label such practices as racist. In the front matter of their book, they note (referencing Jeffery Richards):

Children's literature is a useful instrument for imperialists as they preserve their faith in their imperial agendas. Stories for the young appear innocent, unconnected with the violence of political and cultural domination. However, these stories are largely improvisations on the Western notions that there exists a White man's burden, a concept reinforced when fictional characters are depicted in children's books as incapable of self-determination. White children absorb this myth easily from educational professionals and see no reason to doubt their own alleged superiority (unpagged).

In this particular book, they analyzed thirteen novels set in various African settings to illustrate their point. Their next book, *Apartheid and Racism in South African Children's Literature: 1985–1995* (2013) examines racism that permeates books for children set in South Africa, arguing that the racist apartheid ideology permeates the literature. All in all, representations of Black Africans remain the focus as it draws attention to injustices that spans decades within South Africa.

Yenika-Agbaw's (1998) "Images of West Africa in children's books: Replacing Old Stereotypes with new ones?" and her *Representing Africa in children's literature: Old and new ways of seeing* (2007), expand the dialog by exploring other forms of representations by outsiders and insiders who continue to denigrate the idea of being Black and African through a post-colonial lens. Focusing on West Africa, she was surprised at the realization that even images by renowned authors such as Cyprian Ekwensi perpetuated the idea of Africa as a jungle and Africans as savage. Her findings further complicated the discourse as it begged the question of who then can tell our children's stories? In her closing chapter, she draws attention to the idea of how colonized education seems to have distorted how west Africans regard themselves. Thus, through whose eyes might these authors be telling the stories, or are they simply practicing the culture of mimicry, which Bhabha (2004) defined as "ironic compromise" that is born of the desire to create a "recognizable Other" that results in a "difference that is almost the same" (p. 122).

Essays on representations do not remain only at the continental level as we have illustrated in the last paragraphs, some often target representations in specific African countries. For instance, Williams and Heale's (2013) further discuss the implications of having early books for South African children that center mainly white children. They postulate (2013), "The adult African in European children's fiction was either portrayed as a 'noble savage' or a 'stupid servant'. Any concept of 'noble' disappeared swiftly after South Africa came under National Unity Party" (p. 88). Other works of criticism that examined the contents of children's or young adult literature set in South Africa include Stadler's (2017) *South African Young adult literature in English: 2000–2014*. In his review of Sadler's book, Ellyn Jenkins remarks that her readings of the twenty-four novels "gives no close textual analysis and makes no judgment" (Jenkins, 2017, p. 118). This makes us wonder why she hesitates to give an opinion, especially since her approach described by herself is a "witness document" (p. 11) even though she utilizes several critical discourses, including postcolonialism.

Also surrounding children's literature in South Africa, Jenkins (2004) discussed the shifting focus of children's literature and conservation. His work was not specifically on representation, but he did discuss the transition of children's literature from featuring a white perspective that viewed Whites as having sole ownership of the land to now including and representing indigenous voices through folktales that represent the Black South African indigenous beliefs and practices in conservation. We see this as representation, as he is discussing the bringing in of an indigenous perspective of conservation and land ownership in children's literature.

In regards to representations of Nigerians, renowned author and publisher, [Nwapa \(1997\)](#), who struggled to cultivate local voices to participate in image production, was often frustrated at the quality of illustrations that appeared in locally produced books for children. This sentiment was also shared by Mabel Segun, considered the “doyen of children’s literature and champion of children’s literature in Nigeria” ([Hurs and Ryder, n.d.](#)). In “The importance of illustrations in children’s books,” (1988) she remarks that “Art can help a young child to discover his own identity and cultural heritage. This is particularly important in countries which have suffered from the effects of colonization and cultural imperialism ... Good illustrations can also help to correct wrong notions” (p. 25). Her remarks remind us of the caricature images of Blacks that were ingrained in earlier textbooks such as the *Oxford English Course Reader Four*. These books were written by colonial masters to teach English in “British Africa.” In the collection, a cumulative tale was accompanied by an illustration of a half-naked man, called Lakayana, holding a spear; hence the title of the tale, “See me Lakayana with my spear.” This image haunted some of Vivian’s (first author’s) childhood friends growing up in Cameroon. Her friends were teased and referred to as Lakayana by other classmates because of their clothing. Segun added that, “Too many Nigerian artists give illustrations unimaginative treatment. Having imbibed stereotypes ideas themselves, they perpetuate these stereotypes in their illustrations” (p. 26).

Another country that is focused on in terms of ACL literary criticism and representations is Ghana. For instance, in their essay, “Children’s literature in Ghana: A survey,” [Yitah and Komasi \(2009\)](#) examine the transition of the subject of children’s literature from colonial to postcolonial times. They conclude that children’s literature during the colonial era did not represent the sociocultural context of the Ghanaian child. During the post-independence period, while books gradually started reflecting some of the socio-cultural contexts, their overall quality remained wanting. This essay, therefore, while addressing issues of representations, also addresses publishing issues in ways similar to Mabel Segun’s on Nigerian representations.

Other essays, such as [Kunz \(1997\)](#), focus on representations of specific aspects of Africa. In her work, Kunz works to counterbalance the notion of Africa as one big rural community by examining the images of cities that prevail in selected books from the Cooperative Children’s Book Center (CCBC) collection at Wisconsin. Her findings confirmed her hypothesis of the prevalence of rural cultures in African children’s literature in relation to urban cultures. [Desai \(2011\)](#) also describes the themes and trends in African children’s literature as predominately related to village life, war and ethnic conflict, echoing some of the Kunz’s concerns on the lack of urban representations. The issue here, of course, is that when stories of war become the common narrative of African countries, people’s perceptions of Africa and African countries becomes one of instability and even barbarianism.

Publishing of books with Black and Brown faces

Why is it important to publish books with Black and Brown faces? Of course, within the field of children’s literature, the call for representation has become increasingly louder, beginning with Rudine Sims [Bishop’s \(1990\)](#) call for children’s books that work as mirrors, windows, and sliding glass doors, so that children can see themselves and other children’s experiences through children’s literature. This takes Vivian (first author) back to one of her earlier studies with her children on Lloyd Alexander and Trina Schart Hyman’s *The Fortune-Tellers*. As Vivian ([Yenika-Agbaw, 1996](#)) noted in the essay,

When my children first saw *The fortune-tellers*, their immediate reaction was, “What kind of book is this?”, Yenik, my then 7-year -old-son began twisting his nose in disgust, then suddenly burst out, “Oh, it’s Africa people!” This aroused their curiosity. Luma, the second son (age 6), grabbed the book quickly and began leafing through its pages. “No, he protested, It’s Cameroon people!” From the illustrations he was able to identify a specific group of people—his people—whose cultural experience he had actually participated in before coming to the United States (p. 28).

If this is not enough reason to publish quality books that are representative of all cultures, then we have a problem. Nevertheless, this continues to be a challenge, as some of the essays posit. In “The growth of African children’s literature,” [Osa \(1987\)](#) observes that translating African “oral narratives into print” can be a challenge, but argues that this challenge should not deter African authors from embarking on this creative journey. To him, “what is needed in Africa today is a concerted continent-wide effort by children’s literature scholars, educators, sociologists, and anthropologists to share their ideas. It is in this spirit of sharing that true interdisciplinary criticism of African children’s literature will slowly but surely emerge” (p. 322). [Yenika-Agbaw \(2007\)](#) also addresses this issue of publishing. Trying to get her children’s book manuscripts published have posed particular challenges. In the continent, one was locating an artist who considers illustrating for children worthwhile. Another is that, in the US, it has been a question of audience. For whom, was she writing? Rejection letters praise her strong sense of place, but do not think they can publish her work (perhaps because of the market forces, she speculates). Therefore, she knows this challenge is real.

While more books for children published in the continent are from South Africa, [Heale’s \(1998\)](#) essay argues that more could be done. He discusses the publishing trends and how this might also influence how children interact with books in “What publishers are publishing and what children want to read.” Since this essay was published, we have seen the emergence of Niki Daly, whose books are extremely popular in South Africa and do also have global appeal, especially his Jamila series—even when they seem to convey stereotypes. In an interview ([2020](#)) with Naomi Valenzuela, an intern at Catalyst press, when Daly was asked why he chose to make his protagonist a “black girl,” Daly replied, “Compared to the street life I could relate to as a child I found little that

interested me in the subdued streets and middle class culture I found myself. So, I turned to the township streets and the lives of the poor and working class to find settings for the stories that I could tell" (unpaged). From a critical stance, many scholars would regard this as cultural appropriation, as it comes across as though he is exploiting the misery of the Black working poor for his own material gain. We could also ask, would his books have been so successful were he Black?

When it comes to representing Brown and Black folks, the publishing industry continues to pose a challenge. Many are seeking creative ways to fill in the gaps in their respective countries, such as the recent success of the Golden Baobab Foundation of Education, started by Deborah Ahenkorah, who launched the Golden Baobab prize in 2016 to encourage local publishing in Ghana and the continent. A clever tweet from the Global Centre for Pluralism, which Deborah Ahenkorah retweeted on February 14, 2020, touches on the complex nature of racism and children's literature. It reads, "Imagine if Anne of Green Gables lived not in Avonlea P.E.I but in Harare, Zimbabwe, would she have become the household name she is today?" This tweet echoes several nuances about the publishing experience, highlighting race and geographical location as pivotal to the success of a story. The challenges that Osa (1987) identified several decades ago remain—continuing the need for market forces and inclusive voices that allow Africans to tell their stories, review these stories, and critique them according to their own research traditions.

Joubert (1991) describes the inequities of the dominant publishing model in South Africa which is mainly driven by profit and serves the minority white population. She advocates a budding alternative that is both ideologically and economically efficient and will attend to the literacy needs of the "other" South Africa that has been marginalized by the current industry. Traditional publishers are capable of this alternative equitable publishing under the right circumstances, as is educational publishing. Due to its segregated origins, there exists a black majority in the school system whose required textbook demands are met by these same conventional publishers. The true distribution of the population should be represented in the objectives of the entire publishing industry.

Essays on publishing in Kenya reveal other concerns. While Muringi (2014) attests to the existence of books in indigenous languages, his focus in "Transformations in Kenyan's Children's Prose" is on the social evolution of the publishing landscape vis à vis the evolution of the country itself. Like the other scholars, he opens with references to the lament of the dearth of books with local flavor. He then acknowledges that since research done by his predecessors, the children's literature landscape has gradually evolved. The entire article chronicles this development from a historical perspective starting from 1966 through 2008. After discussing ten carefully selected books, he concludes that the books portray a variety of representations reflecting their colonial, post-independence and postcolonial histories. Thus, to him, "as the Kenyan society changes, so does its children's literature" (p. 97). The essay, while implying publishing struggles, does not overtly address this, as the primary focus was on the evolution of the children's literature as a cultural artifact that is shaped by its cultural surroundings.

Accessibility issues

Accessibility issues continue to be a major challenge despite the numerous international book fairs, including the Ghanaian book fair that occurs in Africa. Vivian (first author) was privileged to have attended the book fair in Accra in 2019 as part of her role with the International Board on Books for Young people (IBBY). While she totally enjoyed the experience, she was also saddened at the overall quality of books on display. Leafing through some, the concerns that Nwapa (1997) and Segun (1988/2015) expressed continue to manifest. However, all is not lost as there are a few publishing houses such as the Sub Saharan Publishers, whose quality titles are rising to the top. With local illustrators stepping up to deliver, this publisher seems to be the hope for Ghana and most of English-speaking Africa, but accessibility issues remain, as it is difficult to get these quality titles outside of Africa. This is where Brenda Randolph's Africa Access organization steps in, forming a great partnership with authors/illustrators and publishers in Africa whose work the organization promotes in the United States.

In a YouTube conversation of their work, Brenda Randolph (Africa Access Outreach Director at the Center for African Studies at Howard University), and Dorcas Wepukhulu (Eastern and Western Africa Kenya-based Partner Development Coordinator at African Storybook), discussed the impact and evolution of African children's literature from the use of racist images through recycled images to more authentic narratives, like *Galimoto*. Of specific mention is the crop of second generation African American writers born in America, creating literature for American youth that is faithful to their parent's African heritage.

The seminar discussed online resources (www.africaaccessreview.org and www.africanstorybook.org) for finding and using African's children's literature, and the collaboration happening between teachers, libraries, authors, and publishers working together to provide constructive learning environments for students in both US and Africa. Africa Access, based in the United States, has extensive partnerships with Howard University, local libraries etc. in the sponsorship of the CABA awards among other initiatives.

African Storybook is an openly licensed digital publication endeavor based in South Africa. It was started as a response to the dearth of storybooks in Sub Saharan African languages by providing digital stories for individual, schools, or libraries to directly print and use. Users can also upload their own stories, translate existing stories or adapt provided images to create their own stories. Their goal is for knowledge and stories to be (multilingually) created and shared in the sub-Saharan community and beyond.

As a professional who has actively participated in selecting books for CABA awards, and who has represented Brenda Randolph at conferences or presented side-by-side her and Meena Khorana, the founding editor of *Sankofa*, Vivian can testify as to the tireless work that goes into seeking ways to publicize African children's literature globally and locally. With Vanessa Oyugi, a new outreach director now working closely with Brenda Randolph, the hope is that many more teachers, librarians, educators would even become more aware of this initiative.

While the work of the Africa Access Outreach largely opens doors of access between Africa and the US, similar efforts are being made elsewhere through individual partnerships. For example, Deborah Ahenkorah at Golden Baobab and Sarah Odedina, an editor for Pushkin Children's Books in the UK have teamed up to bring books created by African authors and illustrators to the UK. They work together, as Pushkin Children's Books, under the direction of Odedina, works to publish select titles from Ahenkorah's African Bureau Stories. Still, more efforts need to be made to create access to children's books coming from the African continent.

Now that we have taken time to unpack representation, publishing, and accessibility issues, we can say that there now seems to be a heightened awareness of this gross injustice within the fields of children's literature. In response, an international journal of the International Research Society for Children's Literature has created a section for African children's literature in their highly regarded journal, *International Research in Children's Literature*. Even *Children's Literature in Education*, which many of us consider extremely resistant to diverse voices, is stepping up to include pieces on Black children's literature. These are all baby steps and we hope that more will be done when it comes to including critical voices on African children's literature.

Gaps in the scholarship

We started off with two primary purposes for our review: to understand the kinds of studies that researchers are conducting or have conducted in regards to children's literatures about Africa; and to understand who the researchers behind these studies are. Overall, our findings reveal that the majority of the studies focused on representations and authenticity, and are grounded in postcolonial and neocolonial theories, shaped by anthropological and "witness" approaches and informed by general literary criticism guided by textual analysis.

As one can see, because of the endemic exclusion of Black African critical voices, research traditions and/or discourses in the conversation on the few books that are being published and hopefully read, include a heavy emphasis on representational issues. While this issue remains pivotal as we forge forward as a community, there are other areas that could benefit from more research. Some of these include:

- Literacies around books published and used in the continent;
- African children's responses to books published in their local settings;
- Dialog on the state of multilingual publishing of African CL in the continent and the market forces;
- African children as authors;
- African diaspora and African protagonists who navigate both spaces;
- Building of equitable partnerships with other stakeholders within the continent and across the globe. What makes these partnerships equitable or not?
- ACL scholarship from or about the other African subregions apart from South Africa, which seems to produce relatively higher amount of works.

We need to reiterate that this is not an exhaustive study; it is just a sample of the types of research and literary criticism on ACL that are currently available based on the four databases we researched. We would have liked to include critical pieces on African diasporan stories that feature African protagonists who shuttle back and forth between the continent and the diaspora, but we could not find much on this. As we conclude our chapter, we hope the landscape of African children's literature will improve to reflect the complex diversity within the continent and among its children including those who call the diaspora and Africa home. We look forward to more studies that critically examine this diversity. The current climate in the country demands that we educate ourselves on issues of otherness and how our perspectives on this might shape how we navigate texts about the other and make space within our scholarly and educational platforms for healthy conversations that enable us to better understand one another. But we must continue to study the books that are published locally, nationally, and globally by insiders, in-betweens, and outsiders.

References

- Deborah Ahenkorah Twitter, 2020. <https://twitter.com/missahenkorah?lang=en>.
- Bhabha, H., 2004. *The Location of Culture*. Routledge.
- Bishop, R.S., 1990. Mirrors, windows, and sliding glass doors. In: *Perspectives: Choosing and Using Books for the Classroom*, 6, pp. ix–xi, 3.
- Desai, C.M., 2011. The world as presented to US children: building bridges or confirming expectations? *Bookbird* 49 (1), 41–50.
- Fleishack, A., Meehan, T., 2014. Values in children's stories in South Africa: an exploratory discourse analysis. *J. Psychol. Afr.* 24 (6), 509–512.
- Heale, J., 1998. What publishers are publishing and what children want to read. *Bookbird* 36 (1), 46–50.
- Hurs, Ryder, (n.d.). Celebrating Nigeria's Literary Icons: Biography: Mabel Segun. <https://hursandryder.wordpress.com/mabel-segun/biography-mabel-segun/>.
- Jenkins, E., 2004. English South African children's literature and the environment. *Literator* 25 (3), 107–123.
- Jenkins, E., 2017. Reviewed work: South African young adult literature in English: 2000–2014 by Sandra Stadler. *Engl. Afr.* 44 (3), 115–119.
- Joubert, S., 1991. Publishing in another South Africa. *Afr. Book Publ. Rec.* 17 (1). <https://doi.org/10.1515/abpr.1991.17.1.9>.
- Khorana, M., 1994. *Africa in Literature for Children and Young Adults: An Annotated Bibliography of English Language Books*. Greenwood.
- Khorana, M., 1998. *Critical Perspectives on Postcolonial African Children's and Young Adult Literature*. Praeger.
- Kruger, H., 2012. *Postcolonial polysystems*. John Benjamins Publishing Company.
- Kuntz, P., 1997. *City Children in African Children's Literature*. ERIC database.

- MacCann, D., Maddy, Y., 2008. *Neo-imperialism in Children's Books About Africa: A Study of Contemporary Fiction*. Routledge.
- MacCann, D., Maddy, Y., 2013. *Apartheid and Racism in South African Children's Literature: 1985–1995*. Routledge.
- Maddy, Y., MacCann, D., 1996. African Images in Juvenile Children's Literature: Commentaries on Neocolonialist Fiction. *MacFarland*.
- Muringi, C., 2014. Transformations in Kenyan's Children's Prose. Online, pp. 88–98.
- Ntuli, C.D., Mafela, M.J., 2013. Children's literature and the realities of social and cultural conditions: a perspective on Niki Daly's not so fast Songololo. *S. Afr. J. Afr. Lang.* 33 (1), 39–43.
- Nwapa, F., 1997. Writing and publishing for children in Africa: a personal account. In: Granqvist, R., Martin, J. (Eds.), *Preserving the Landscape of Imagination: Children's Literature in Africa*. Rodopi, Atlanta, GA, pp. 265–275.
- Osa, O., 1983. Reviewed work: children's fiction about Africa in English by Nancy J. Schmidt. *Res. Afr. Lit.* 14 (2), 257–261. Special Issue on Modern African Fiction.
- Osa, O., 1985. The rise of African children's literature. *Read. Teach.* 28 (8), 750–754.
- Osa, O., 1987. The growth of African children's literature. *Read. Teach.* 41 (3), 316–322.
- Osa, O., 1995. *African Children's and Youth Literature*. Twayne.
- Osa, O., 2007–2009. Introduction. *JACYL* 1–19.
- Osa, O., 2010. The expanding universe of African children's literature: the why, the how, and the what of publishing in Africa about African children's literature. *J. Afr. Child. Youth Lit.* 19, 1–18.
- Q&A With Niki Daly, 2020. <https://www.catalystpress.org/qa-with-niki-daly/>.
- Schmidt, N., 1965. Children's literature about Africa. *Afr. Stud. Bull.* 61–70.
- Schmidt, N., 1975. *Children's Books on Africa and Their Authors: An Annotated Bibliography*. Africana.
- Schmidt, N., 1981. *Children's Fiction About Africa in English*. Conch Magazine Ltd, New York, NY.
- Segun, M., 1988/2015. The importance of illustrations in children's books. In: *Illustrating for Children*. CLAN, pp. 25–27. <https://www.yumpu.com/en/document/view/20150930/the-importance-of-illustrations-in-childrens-books-mabel-segun>.
- Stadler, S., 2017. *South African Young Adult Literature in English: 2000–2014*. Universitätsverlag.
- The American Library Association's IRC Africa Subcommittee hosted, 2020. *Accessing African Children's Literature Through Online Resources*. Accessing African Children's Literature Through Online Resources—YouTube.
- Thomas, E.E., 2020. Young adult literature for black lives: critical and storytelling traditions from the African diaspora. *Int. J. Young Adult Lit.* 1 (1).
- Williams, J., Heale, J., 2013. The growth toward a truly African quality in South African children's literature. *Bookbird* 51 (1), 87–93.
- Yenika-Agbaw, V., 1996. When an outside perspective touches the insider: a commentary on the illustrations of Lloyd Alexander's the fortune-tellers. *J. Child. Lit.* 22 (1), 28–30.
- Yenika-Agbaw, V., 1998. Images of West Africa in children's books: replacing old stereotypes with new ones? *N. Advocate* 11 (3), 203–218.
- Yenika-Agbaw, V., 2007. *Representing Africa in Children's Literature: Old and New Ways of Seeing*. Routledge.
- Yenika-Agbaw, V., 2013. Black Cinderella: multicultural literature and school curriculum. *Pedagog. Cult. Soc.* 22 (2), 233–250.
- Yenika-Agbaw, V., 2018. From orality to print: construction of Nso identity in folk tales. with Celia Abicalil Belmino, Alice Curry, Li Lifang, and Yasmine Motawy. In: Stephens, J. (Ed.), *The Routledge Companion to International Children's Literature*. Routledge, New York, NY.
- Yenika-Agbaw, Napoli, M., 2011. *African and African American Children's and Adolescent Literature: A Critical Guide*. Peter Lang.
- Yitah, H., Komasi, M., 2009. Children's literature in Ghana: a survey. *Child Lit.* 37 (1), 236–255.

Global meaning-making: critical analysis and advocacy for a diverse world

Robert J. Tierney^{a,b,c}, ^a University of British Columbia, Vancouver, BC, Canada; ^b University of Sydney, Sydney, NSW, Australia; and ^c Beijing Normal University, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	696
Dimensions of global meaning-making: key processes	697
Sites of contemplation	698
Analyses	700
Advocacy	700
Border-crossing across multitopic spaces	702
Being ethical, respectful, and responsive	702
Final reflections: key considerations for global meaning-making	703
The dance of global meaning-makers	703
Global meaning-makers as critical culturalists	704
Global meaning-making is contemplative, analytic, and advocacy-oriented	704
Closing summary	705
References	705

Introduction

Global meaning-making emanates from the view that there are important local and global issues that demand our consideration. These issues are tied to human rights concerns as they relate to social, cultural, economic, and health matters, as well as to the natural world. They pertain to matters of equity and discrimination, opportunity and access, responsibility and respect, and sustainability and health. The notion of global meaning-making supports responses to these issues that wed critical engagement with socio-political analyses and advocacy—enlisting understandings of globalism, multiculturalism, postcolonialism, and diversity (especially with regard to issues of race, indigeneity, and migrations). More broadly, it entails those socio-cultural views of meaning-making and critical theoretical and sociological interrogations of society (especially postcolonialism) that rose toward the turn of the last century.

At the local level, global meaning-making responses relate to [Ladson-Billings' \(1994\)](#) quest for “culturally relevant schools that ... provide educational self-determination, (and) honor and respect the student's home culture” (pp. 135–137). They are tied to notions of the cultured learner and teacher emanating from Canada (see the Québec Charter; [Ministère de l'Éducation, 2001](#)). As [Paré \(2004\)](#) suggested, “the cultured teacher is ... a broker who establishes links between students and the world” (p. 3). The tools employed in such a process similarly resonate with scholars of multiculturalism (especially translanguaging), who argue that all meaning-makers are multilingual as they enlist their ever-expanding repertoires of experiences to understand and contemplate meanings ([García and Kleifgen, 2020](#); [Gee, 2001](#); [Gutiérrez et al., 2009](#); [Lee, 2020](#); [NASEM, 2018](#); [Purcell-Gates, 2006](#)).

Befitting forms of shuttling back and forth been the local and wider world, global meaning-making also weaves together notions of cosmopolitanism (e.g., what has been described as public diplomacy) with a disposition toward respect for cultural diversity ([Hull and Stornaiuolo, 2014](#); [Stornaiuolo and Nichols, 2019](#)). As with [Rizvi's \(2009b\)](#) discussion of cosmopolitanism, global meaning-making is “a political philosophy, a moral theory and a cultural disposition” (p. 253). It is, as [Nussbaum \(1997\)](#) has suggested, a process of critical reflection and reflexivity—one that identifies with the global human community and engages one's ability to imagine across cultural differences. As [Luke \(2004\)](#) explained, this involves:

... exploring the conditions for intercultural and global intersubjectivity ... an engagement in globalized analyses that continually situate and resituate learners ... their local conditions, social relations and communities, in critical analyses of the directions, impacts and consequences of global flows of capital, bodies, and discourse (pp. 1438–1439, 1441).

Such work is consistent with advocates of local cultural expectations, as seen in [Giddens \(1999\)](#), who argues for a global dialectic, or [Park \(2017\)](#), who espouses “a paradigm shift ... with a de-imperialized, de-colonized, and de-Cold War mentality” (p. 760). As [Soysal \(1994\)](#) has argued:

The state is no longer an autonomous and independent organization closed over a nationally defined population. Instead, we have a system of constitutionally interconnected states with a multiplicity of membership. [Hence] ... the logic of personhood supersedes the logic of national

citizenship, [and] individual rights and obligations, which were historically located in the nation-state, have increasingly moved to a universalistic plane, transcending the boundaries of national states (pp. 164–165).

The commitment to the importance of diversity is consistent with the view of the world as pluralistic—comprised of coexisting, criss-crossing spaces with multicultural and multilingual aspirations as well as ecological goals (de Sousa Santos, 2007a,b, 2013; de Sousa Santos et al., 2007; Grigorov and Fleuri, 2012; Misiaszek, 2015).

Global meaning-making thus aligns with calls for responsive, ethical and proactive engagements with the socio-politics of planetary considerations (e.g., Abdi, 2015; Andreotti and de Sousa, 2011; Misiaszek, 2015; Torres, 2015). Accordingly, critical analysis and advocacy are seen as integral to global meaning-making—especially pursuits that espouse postcolonial ethical stances (e.g., Smith, 2005) and the liberatory educational aims of Freire (1973) and Boal (1993) to go beyond reading the *word* to reading the *world*. This is exemplified by the ongoing Indigenous concerns raised and activism pursued by First Nations, Maori, and Australian Aboriginal educators in New Zealand, Australia, and Canada, which honor and respect Aboriginal ways of knowing and languages (e.g., Battiste, 1998; Battiste and Henderson, 2000; Bishop, 1994; Nakata, 2001, 2004; Ocholla, 2007; Rigney and Hattam, 2018; Smith, 2000, 2015, 1999; Woods and Biermann, 2009). It is consistent with the design principles proposed by Gutiérrez (2011, 2016) for ecologically resistant and culture accommodating research that include the following:

(a) deep attention to history and historicity, including how they relate to resilience, sustainability, and equity; (b) a focus on re-mediating activity, that is, a focus on reorganizing the functional system rather than individuals; (c) employing a dynamic model of culture with an understanding that cultural and other forms of diversity are key resources for sustainability; (d) a persistent emphasis on equity across the design process in the theorization, design, and implementation; (e) an emphasis on resilience and change, where change implicates the individual, the collective, and the ecology; and (f) an end goal of transformation and sustainability.

Gutiérrez (2016, p. 192).

Dimensions of global meaning-making: key processes

Global meaning-making aligns with pursuits valuing heterogeneity over homogeneity; accommodation over assimilation; and contextually bound engagements over broad generalized prescriptions. These values are ensued as meaning makers engage with:

- Addressing pressing local and global issues;
- Interrupting existing frames, especially cultural hegemony;
- Decolonizing and opening spaces for the sovereignty of and respect for diverse cultures and their self-determination;
- Accommodating cultural dynamics within and across communities—via adaptation, translanguaging, fusion, and border-crossing;
- Reading self and deconstructing identity as one engages with diversity, indigeneity, and decolonization;
- Being mindful, and finding a higher moral plane;
- Interrogating truth/post-truth;
- Being activist or actionist.

Tierney (2018a,b).

Forms of participation may shift as meaning-makers traverse multiple public spheres, identities, and approaches in their contemplations, analyzes, and responses. As individuals contend with various cultures, this process may align with context-specific negotiations or navigations of cross-border meaning-making (e.g., García and Kleifgen, 2020; Horowitz, 2012; Nelson et al., 2016). This may involve “... an instantiation of boundary crossing where an individual deliberately and consciously pushes against society’s ideological constraints” (Gutiérrez et al., 2017, p. 53). Singularly or collectively, it may entail forms of consultation and leveraging as meaning-makers negotiate the norms of societies and communities—in respectful ways that do not defy or defile, unless intended to disrupt the systems antithetical to the pursuits themselves.

Undeniably, certain kinds of shuttle diplomacy are complex, especially as meaning-makers themselves engage in or affiliate with movements of social transformation as activists, allies, or cultural development workers (Bishop, 1994, 2009; Butler, 2011; Campano et al., 2010; San Pedro, 2018). Challenges to the dominant cultural interests will be ill-fated unless meaning-makers approach matters reservedly as they engage with and seek to shift the socio-political currents that they encounter.

Global meaning-making requires responding to local concerns with an eye to both local considerations and global matters. Global meaning-making can therefore take on diverse forms, move in different directions that stem from multiple perspectives, different interpretations and are shaped by various forms of reconciliation arise (e.g., in cases where intra- and intercultural developments and differences, and parsimony between local and global communities, are negotiated).

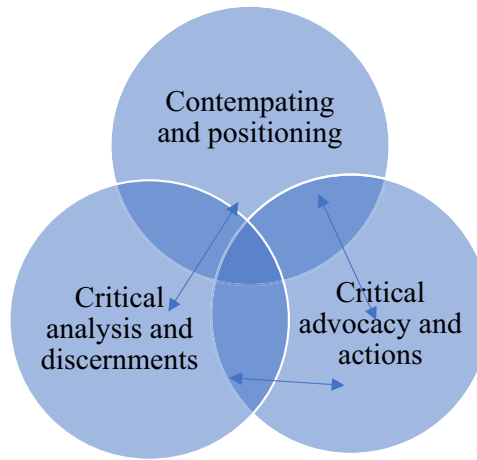


Fig. 1 Dimensions of global meaning-making.

Under the umbrellas of contemplation, analysis, and advocacy, **Fig. 1** attempts to capture the intersecting and iterative nature of global meaning-making—and how they interface with one another. As suggested the clusters entail:

- Adopting a contemplative countenance and probing positioning via socio-political currents, including observing, listening, interrogating, and enlisting multiple perspectives;
- Analyzing how people, events, and issues are represented and positioned by authors and audiences, and how their responses and perceptions enlist multiple perspectives and repertoires of experiences (including multilingualism and rhetorical analyses to discern socio-political matters and possibilities);
- Acting upon deliberations of discernments as one enlists forms of ethical, respectful, and responsive social consciousness as an advocate, activist, and ally.

Befittingly, global meaning-making is diverse rather than predetermined. Its dimensions are more intertwined than discrete, and more recursive than linear.

Sites of contemplation

Befitting this cross-cultural deliberation model, global meaning-making requires contemplation based upon investigations of global and local meanings in terms of their antecedents and constitutions—especially their historical, cultural, and ecological relevance to sociopolitical considerations. As global scholar and comparative educator, [Hayhoe \(2021\)](#) has suggested, cross-cultural deliberations are prompting individuals to adopt a “listening intellect” to engage with the significance of developments on a deeper level—“requiring us to open our minds and hearts to the other” (p. 13). And, as Hayhoe and others have contended (e.g., [Andreotti and de Sousa, 2011](#); [Smith, 1999](#); [Spivak, 1988, 1990](#); [Takayama et al., 2017](#)), being contemplative involves reflexivity, respectfulness, and self-consciousness; it requires staying attuned to what is presented, by whom to whom, and why, when, and where. It also requires an awareness of the intentions and possible effects of various meanings. It likewise demands forms of shuttle diplomacy as one delves into the worlds of others while remaining conscious of the intrusion of one’s own history and its potentially skewed perceptions. Global meaning-makers are not just seeking to understand or present themselves. They are examining ideas in terms of the various subtexts that they might represent or advance.

The basis for contemplations can be quite varied and mixed. They might be spurred by momentary reflections—regarding everyday self-monitoring, incidental exchanges, or news releases—or they can involve extended considerations of an issue or an event in chat rooms, in the media, or across other venues. They may emerge from conversations with friends, family, or community members; they may be derived from “across the aisle” exchanges with strangers or those previously identified as outsiders (e.g., from other cultures or opposing points of view). They might arise in different spaces and experienced as proximate or firsthand, distant or through secondary accounts, or a combination thereof. Oftentimes they reflect our everyday literacies as meaning-makers view or encounter various forms of media (e.g., films, television, news media, social media or websites). They might stem from events that are experienced or witnessed such as incursions between individuals, groups, communities or nations.

Nowadays many of our lives exist across multiple time zones with different communities. We engage in global exchanges on the condition of our worlds especially as we experience the events occurring. Sometimes these events are spurred by natural disasters or political upheaval or forms of terrorism; at other times they may be spurred by efforts to engage with or learn from one another, to aid or to join together in partnerships around regional or planetary concerns. Some might be based in global charters and declarations by global organizations such as the United Nations; others might stem from movements tied to human rights or environmental concerns. Certainly, a number stem from natural disasters such as earthquakes, droughts, hurricanes. Others involve

inequitable treatments such as women have experienced especially in terms of their education (United Nations (UN) Human Rights Council, 2017; United Nations Educational Social Cultural Organization (UNESCO), 2014; UNESCO, 2015a; UN Girls Education Initiative, 2017).

Take, if you will, the various forms of global meaning-making that have arisen recently with the events experienced in Europe. For example, within and across the various sectors of the world, Russia's invasion of Ukraine has led to intense contemplations, analyses and acts amid global considerations of democracy versus autocracy, sovereignty and self-determination together with analyses of motives and actions, responsibilities and ethics. Many of us have been a mix of witness and participant as we have engaged in forms of global meaning-making with these events.

Oftentimes we experience events as they unfold from global sports to acts of terrorism. Such was the case for those of us who grappled with the global news of the Australian white supremacist who killed some 51 Muslim worshipers at two mosques in Christchurch in 2019. The event brought to the surface a range of existing and longstanding issues and debates, including concerns over global terrorism and the rising white supremacist movement; the consequences of interconnectedness with global media; and the precipitousness of hate speech across tweets, blogs and political policy speeches (e.g., as disseminated by online figures from Pewdiepie to Donald Trump). Among the most poignant moments following the attacks occurred with the comments of New Zealand Prime Minister Jacinda Ardern, who reminded her community and the global community at large of the need for a moral compass that embraced each another as one, with all our differences. Her words sharply contrasted with the vitriol and negativity around migration elsewhere in the world. As she stated: "Many of those who would have been directly affected by this shooting may be migrants to New Zealand. They may even be refugees here. They have chosen to make New Zealand their home, and it is their home They are us" (New Zealand Herald, 2019, paras. 3–5).

Similarly, across several countries, especially those involving Indigenous populations, the uncovering of the gravesite of children taken from their families and placed in Residential Schools in Canada spurred global contemplations about the history of assimilation by Church and State (Callimachi, 2021; Tilmouth, 2021; Truth and Reconciliation Commission of Canada, 2015). In Australia particularly, the finding resonated with Australia's atrocious history of assimilation and genocide of Aboriginal populations—still manifested in its systemic discrimination and lack of recognition or respect for Indigenous rights and sovereignty. The events in Canada thus triggered contemplations mixed with analysis and advocacy as they contributed to a greater public reckoning with inequities and racism. For Australians, these stories were educative, provocative, and fueled support for Aboriginal demands for their own transformative change—serving as antecedents to a reconsideration the systemic forces that have shaped the Australian nation and culture, including the unacceptable racism that persists and is perpetuated by the lock that white settler privilege holds over Australian Aboriginal and Torres Strait Islander society.

In Australia, for example, these matters resonate with an awakening of settler society to their miseducation, ignorance and complicity with Australia's racist history. For Australian settlers, several books and movies exposed the history including: Willmott's (1988) *Pemulwuy: The Rainbow Warrior*—the story of Australian Aboriginal resistance to British occupation that was kept invisible to outsiders and across generations of Australian immigrants; Behrendt's (2004) modern day story, *Home*, in which she, an urban Aboriginal, visits her place of heritage in a remote rural area where her community experienced horrific treatment; or movies, such as the 2002 film *Rabbit-Proof Fence*, directed by Phillip Noyce and based upon the 1996 novel *Follow the Rabbit-Proof Fence* by Doris Pilkington Garimara, which follows three girls trying to return to their home following their abduction by government authorities. In Australia, they resonate with the argument presented books challenging misinformation and the faulty understandings of the ways of being of the Australian Aborigines such as Pascoe's (2014) *Dark Emu* (Pascoe, 2014) that described the sophistication of Australian Aboriginal agricultural pursuits or accompanied questioning of the western material terms for judging Aboriginal practices such as the book by Indigenous anthropologists and archeologists, Peter Sutton and Keryn Walshe who challenge Pascoe's measures of development with fuller discussions of Aboriginal life spiritually and ecologically as very advanced hunter-gatherers (Sutton and Walshe, 2021). And, with other commentaries (especially data on education, health, and economics, incarcerations and suicides), add to the need for scrutiny of the hegemonic lock that assimilation, racist and other limiting views holds on Australian society.

There are numerous contemporary texts expressing similarly challenging ideas that call for contemplation and cultural analyses. Take, if you will, *The Poet X* by Acevedo (2018)—a novel in verse examining the struggles of growing up as a Dominican American teenage girl. Or consider the post 9/11 tensions uncovered in *Home is Not a Country* by Safia Elhillo (2021). These are stories involving cultural and intergenerational spaces where authors prompt forms of border crossing. They are stories that readily question socio-political issues and currents, in overt and edgy ways that resonate with the everyday—affording readers opportunities to enlist their translanguaging repertoires of cultural perspectives, languages, and insights as they explore the worlds of diverse populations (e.g., Garcia, 2009; Garcia et al., 2017).

Matters being dealt with these books may be global but are not necessarily foreign. Baldwin's (1974) *If Beale Street Could Talk* moves readers across multiple perspectives, cultures, and worlds. It is a story that engages readers in complex social dynamics within and across families, communities, and cultures. Readers are prompted to actively transact with these cultures, bringing their own selves and their societal experiences into play as they contemplate the events and the cultural forces flowing through the text. And, as Baldwin seems to suggest, Beale Street is not just one place within one story. It represents what for many is a familiar story of differential treatment and injustice tied to race—concerns most recently being confronted by movements such as Black Lives Matter and tied to events such as the death of George Floyd at the hands of police in Minneapolis. Perhaps in tandem with such events, it is a story to contemplate and ponder, and to analyze and act upon.

Analyses

In terms of socio-political analyses, global meaning-making entails adopting different critical frames as vehicles for examining circumstances. Analyses might emanate from close observations of circumstances, or through needs assessments. They might entail examining the rhetoric of politicians and the subtext. Indeed, politicians' speeches and tweets can be sites that lend themselves to various analyses related to how politicians position themselves, key issues, their followers, and their opponents. For instance, in her university course on the use of rhetorical strategy in public discourse, [Mercieca \(2015\)](#), an American political rhetorician and educator, asks students to examine their world news and consider the strategies employed to either advance false storylines/narratives or to pivot away from, rather than confront, core issues. Without accusing news outlets of lies, she has her students screen the array of sources for instances of rhetoric crafted to disguise or override truth, perpetuate certain conspiracies, or sustain narratives that advance insurrection or violent mob behavior. Through this exercise, students engage in rhetorical analyses that address how authors present ideas and consider how audiences are shaped and perhaps coerced into believing and acting in certain ways ([Mercieca, 2015](#); [McMurray, 2021](#)).

The media is replete with analyses focused upon the representation of events, the warrants and evidence of claims being made as well as the motives and politics suggested, implied and verifiable. Consider the focus in recent years on Donald Trump or Vladimir Putin especially the analyses of their speeches and in terms of their goals and their positioning against what is viewed as known and possible agendas. For Donald Trump, it is often aligned with discussions of salespersonship such as Trump espouses in *Trump: The Art of the Deal* ([Trump and Schwartz, 1988](#)) or partisanship, and concern over truthfulness. For both Donald Trump and Vladimir Putin, their representation of history and rationalization of actions are reminiscent of Orwellian discussions of society *Nineteen Eighty-Four*, ([Orwell, 1947](#)) along with the politics ([Orwell, 1984](#)) of language discussed in [Orwell's \(1984\)](#) seminal article, "Politics and the English Language:"

Political language—and with variations this is true of all political parties, from Conservatives to Anarchists—is designed to make lies sound truthful and murder respectable, and to give an appearance of solidity to pure wind (p. 366).

Accordingly, analyses of events can occur at several levels and involve multiple frames and perspectives. Analyses might involve critical examinations of the politics of persuasion, evaluating claims in terms of foundation and evidence. They might entail confronting regressive or insular tendencies of a "post-truth" world ([Peters, 2017](#))—those ideas and rhetorical approaches that advance the suppression of environmental and social developments, or dismiss universal truths through the use of unfettered hyperbole, misrepresentation, oppressive personifications, and false equivalencies. They might address matters of local concern such as social services or individual rights. They might extend the scope to issues of choice and multicultural considerations, such as language use. They might enlist some of the books written in more than one language about cross-cultural engagements (e.g., *My Papi has a Motorcycle*, by Isabel Quintero) to invite and legitimize the multilingual repertoires that meaning-makers bring to explore, imagine, exchange experiences and perspectives, and respond to the issues represented. Such analyses may be spurred by the use of sociological frames (e.g., [Arendt, 1973](#); [Bourdieu, 1993](#); [Connell et al., 2017](#)); rhetorical analyses (e.g., [Burke, 1969](#); [Kubota and Lehner, 2004](#); [Lakoff, 2016](#); [Orwell, 1984](#); [Toulmin, 1958](#)); cross-cultural, critical race, feminisms and gender theory (e.g., [Butler, 2011](#); [Enciso, 1997, 2004](#); [Rogers, 2017](#); [Rogers and Soter, 1997](#)); as well as postmodernism (e.g., [Peters, 2017](#)) to examine how people and ideas are positioned and travel (e.g., [Leander and Sheehy, 2004](#); [Stornaiuolo et al., 2017](#)). Befitting global meaning-making such analyses entail cultural analyses that delve into potential subtexts and socio-cultural-political undercurrents.

Advocacy

Global meaning-making also involves advocacy ranging from local problem-solving to cross-cultural negotiations directed at leveraging. Indeed, the ideological foundation of critical global meaning-making pursuits is tied to pursuing change in ways that support local sustainable efforts or, if necessary, to challenge the hegemonies that advance forms of exceptionalism and privilege. As [Andreotti and de Souza \(2011\)](#) have suggested, there is a need for readers to interrogate ideas, especially in terms of systems. In particular, they suggest readers should be alert to both colonizing constructions of the world (e.g., hegemonic, ethnocentric, ahistorical, depolitic, paternalistic, and simplistic) as well as to their own complicities with the continuation of colonialist systems.

Global meaning-making therefore involves a critical countenance, growing out of a contemplation of and response to the systems at play in ways that advocate for and at times challenge their premises, roles, or influence. Building upon the critical work of [Spivak \(1988, 1990, 1999\)](#), [Andreotti and de Souza \(2011\)](#) call for a global responsiveness and imaginary that disrupts the normative and challenges the colonized status quo. Such work underscores the importance of identifying individual desires as a way to "de-mystify the fantasies behind them and to open up possibilities previously unintelligible to the invested self" ([Andreotti and de Souza, 2011](#), p. 225). As [Stein \(2017a\)](#) has similarly suggested, there is need for a decolonizing orientation,

to interrupt and denaturalize the epistemological and ontological frames of modern existence that are produced through violent and unsustainable relationships and processes— ... to reimagine and recreate what is possible—and to ask about the role of education in facilitating these other possibilities (p. 44).

Inherent to this critical global meaning-making perspective on advocacy is the view that all of us are planetary dwellers with a shared view of human, societal, and planetary rights, along with an interest in the sustainability of our diversity. To share this planet requires an interest in spaces that support cross-cultural engagements—those to which all can bring cultural capital that is distinctive to their meaning-making; where all can actively contribute to a global trajectory built upon support for diversity; where differences without violent intent are respected. Global meaning-making is wedded to decolonizing spaces and knowledge emancipation, where the cultural capital (knowledge and ways of knowing) of various meaning-makers is respected (de Sousa Santos, 2007a).

Global meaning-making involves opening up and supporting the quest for new spaces for contemplation, analysis, and transformation (i.e., of matters pertaining to issues of bias and respect for differences). As Kay (2018) described in his book, *Not Light, but Fire: How to Read Race Conversations in the Classroom*, global meaning-making might be construed as pursuing and opening up new spaces for contemplating complex issues such as racism and prejudice, akin to what have been described as “brave” spaces (Ali, 2017). Such spaces afford meaning-makers opportunities to relate their engagement to people and events within and across worlds—from neighborhood incursions such as shootings by the police, protests, or severed relationships among friends, to more global events. They open up opportunities to share experiences, reflections, and concerns. As translanguaging advocates argue, these spaces advance multiculturalism through pedagogical initiatives that engage meaning-makers’ linguistic and experiential repertoires; expand the community; and challenge the prevailing monolingualistic hegemony within systems (e.g., Garcia et al., 2017).

As advocates and activists, global meaning-makers act as cultural protagonists, siding with local and global diversity, responsibility, and equity against hegemonies that seek to overpower and discriminate. Global meaning-making entails advocating for an accommodation of differences rather than ignoring them. As Haluza-DeLay (2003) asserts: “Knowledge uncovers the oppressive structures and confronts power. However, it is not the ‘knowledge’ alone that does this, but the process by which the knowledge is taken up and used in the community, altering ‘common-sense’” (p. 86). In accordance with activism, global meaning-making involves.

thinking, talking about, researching and theorizing about what is going on, what they are going to do next and how to analyse the situations they face, whether in relation to attending a demonstration, a meeting, a confrontation with institutional forces or planning the next action or campaign.

Kinsman (2006, p. 134).

The activism associated with global meaning-making befits Gramsci’s (1982) view that scholars should be “organic” rather than “traditional intellectuals;” that is, they should engage with the public rather than focus on perpetuating their insular academic theories. This notion is not unlike Freire (1973) conscientização, or transformative change, or what Lather (1986) termed “catalytic”. Moreover, it is consistent with educators engaging in resistance and activism (e.g., Simon and Campano, 2013; Wandera, 2020) and all of us engaging a model of reading and inquiry that aspires to be transformative (Dion et al., 2020). As Goodwin (2012) suggests, it should be “an orientation to inquiry with an obligation to action” (p. 3), involving participation that is tied to local interests and pursued cyclically. Activism is integral to global meaning-making, and the goals of activism are forms of engagement that mobilize communities. As Hall (1978) and Choudry (2015) suggest, activism involves continuously testing, acting on, trying out, and repeating the cycle as one moves forward. It shifts the role of reader from that of a distant observer, critic, or director to a supporter and ally (Bishop, 2009).

Advocacy, activism, and respect entail supporting worlds with fewer walls—campaigning for fewer jurisdictional constraints that often dismiss, censor, or immobilize those who are “othered” within and across communities. Ideally, global meaning-making is not solitary or insular; engagements are participatory and expansive. It involves colleagues and collaborators with knowledges often left out of or overlooked by colonizing Western perspectives, such as those with Indigenous histories, migrant pasts, and cultural moorings from places such as Oceania, Asia, the Americas, and Africa. It seeks spaces that support participatory forms of reading that move beyond contemplation, analyses, understandings and interpretations to responses involving advocacy.

Such advocacy has been exemplified by research projects with youth, involving the use of social media for political purposes (and in conjunction with social movements). For example, Rogers et al. (2015) explored how homeless youth enlist new and critical literacies, especially the multimodal possibilities afforded by various digital media, to engage various publics and themselves in speaking out about societal issues and counter faulty public constructions of their identities. Their projects initiated and described the purposeful enlistment of a mix of genres and forms to foment reflective and responsive engagement, including forms of cultural criticism and civic engagement by youth. As Rogers et al. (2015) describe in their study, young people’s enlistment of digital resources supported “expressions of resistance to the inheritance of the broken promises of democratic citizenship and their ability to imagine new possibilities of public engagement” (p. 2). The youth they observed engaged in “juxtaposing ... hybridizing ... remixing” (p. 102) to create counternarratives and to speak for themselves. As Henry Jenkins and his colleagues describe in their book, *By Any Media Necessary: The New Youth Activism, Connected Youth and Digital Futures* (Jenkins et al., 2016), various innovative forms of youth-led activism demonstrate how young activists’ uses of media and networking might mobilize and engage others in coordinated pursuits that establish points of contact politically, culturally, and institutionally; stimulate the civic imagination; and

explore paths and possibilities in ways that they describe as vertical (from expert to novice, or in terms of institutional hierarchies) or horizontal (peer to peer). As others have suggested the power of forms of symbolic esthetic forms to raise consciousness and engagement in societal issues (e.g., Clover, 2013a,b; Clover and Stalker, 2007).

Border-crossing across multitopic spaces

As implied, being a global meaning-maker entails being multilingual—befitting models of a pluralist world or postmonolingualism (e.g., see Escobar, 2018; Perry, 2020; Perry and Pullanikkatil, 2019; Singh and Lu, 2020; Thiong'o, 1986, 2012). It requires agility and responsiveness, as well as potentially significant forms of accommodations. Such forms of renaturalization, community consultation, and cross-cultural understandings may be less than transparent unless one has language skills and informants (Ndimande, 2018).

As Massey (2005) suggests, spaces for such cross-cultural exchanges require Westerners to look inward and outward, within and across public and private spaces. These spaces operate through intercultural networks (Luke, 2011; Rizvi, 2021), following the bottom-up and sideways flows of people, ideas, and acts (Abdi, 2015; Campano et al., 2010; King, 2017; Tierney, 2017a,b). It may be that “uptakes” occur horizontally or syncretically across sites or boundaries, not unlike the translanguaging that occurs at borders (e.g., Gutiérrez, 2014; Gutiérrez and Rogoff, 2003; Nelson et al., 2016), or the line-stepping, or testing of boundaries, by individuals or groups. As Gutiérrez et al. (2017) describe:

Line-stepping is an instantiation of boundary crossing where an individual deliberately and consciously pushes against society's ideological constraints. Rather than seeing boundaries as static, we recognize their dynamism. By subtly identifying and testing a line, the line-stepper learns how and where lines are permeable and the available latitude in their enforcement. At times, youth will encroach the lines without going over them; at other times, they will cross the lines, attempting to ascertain the severity of the consequences of their boundary crossing (p. 53).

With these global meaning-making crossings and exchanges, there are fusions—in the form of mixing, remixing, adapting, and taking up ideas (e.g., by endorsing, refuting, protesting, or mocking). These may involve artistic renditions, or fusions tied to different epistemologies. With increased access, heightened bandwidth, and a variety of outlets, digital spaces offer global meaning-makers less-regulated venues to assemble, explore, share, and connect. When these fusions are not pursued respectfully, ambitions that are profit-oriented, commodifying, or politically-fueled can override either the diverse interests or the proprietary rights of individuals or communities. The role of the digital in society the power and the significance of digital affordances for global meaning-making should not be underestimated (Deuze et al., 2012; Luke, 2018). These literacies “serve as vehicles to name our worlds, interact with others, imagine, test ideas and change our world. The affordances have become as integral to meaning-making, as air is associated with breathing” (Beach and Tierney, 2016, p. 135). Various digital media can offer individuals and groups new vehicles for speaking out about societal issues, including matters of their identity (Beach and Smith, n.d.; Beach et al., 2017). For example, digital media may offer spaces that impose, override, colonize, standardize and restrict diverse cultural expressions and their agency. On the other hand, media platforms can also provide palates for widening the forms, as well as the reach of and possibilities for, existing or alternative engagements (e.g., Kim, 2016). Indeed, affordances can be prefabricated or customized with designs that are either barren or fertile sites for sharing and advancing ideas.

Being ethical, respectful, and responsive

Global meaning-making is not exempted from ethical considerations—especially in terms of respect for and responsibility toward others (e.g., Butler, 2011; King, 2017; Kristeva, 1991; Levinas, 1974/1981, 1993, 1995). To be contemplative, there is need to move outside of one's bubble—to engage with and consider others and the socio-political systems at play. Considerations of broader systems include examining how and what they regulate, and whose interests are served, compromised, or forfeited in the process. Also key to this contemplative process of global meaning-making and advocacy is the pursuit of actions that are respectful and supportive. In keeping with goals of examining and interrogating one's own motives and positioning, it is crucial that actions and engagements are not presumptuous or out of place with the rights and interests of others, especially in terms of self-determination and sovereignty.

To engage in global meaning-making requires occupying a moral ground that bridges across and within nations, identities, cultures, and communities (e.g., between the privileged and the marginalized; across gender and races; and between Indigenous and immigrant cultures). It encompasses a journey across complex and diverse terrain (ecologically, epistemologically, and ethically). As Kohlberg (1981) suggests, there may be a seventh level of moral reasoning, to which global reading and research should aspire—a level that looks more broadly at the ethics of decision-making, aligned with a fuller and ongoing consideration of the future of a society. As Appadurai (2005) comments, this requires more than adopting an ecumenical and generous approach; it requires suspending certainty and opening oneself up to debate, to differences, and to considering grassroots internationalism as a crucible for emergent new forms of global ethics. It is consistent with arguments made by de Sousa Santos et al. (2007)—and, more recently, those by Chinese and global scholars such as Shi Zhongying (Shi and Li, 2018; Tierney et al., 2021)—of the need to decenter western epistemologies as the metonym for knowledge (i.e., to instead position the west, as a range of constituent knowledges, among other knowledges).

To achieve these moral goals, there is a need for aspirational spaces where cross-cultural meaning-making can occur on multiple levels and proceed in a manner that is responsive, respectful, relevant, and energizing. As Maori scholars [Smith \(2000, 2015\)](#) and [Smith \(1999, 2005\)](#) have argued, cross-cultural engagements should proceed in a manner that is respectful of the histories, ways of knowing, needs, hopes, and values of all. The specific circumstances of “others” should be respected, including geographies of time and space, and the ecological systems’ local norms, self-realization, and self-determination. In these challenging times, global meaning-making is not straightforward—indeed, it contends with the kind of “tricky ground” that [Smith \(2005\)](#) described for researchers attempting to navigate within and across Indigenous spaces.

As such, global meaning-making requires deliberations relative to: (1) How authors are positioning themselves; the topics and socio-cultural features that they are exploring; and the reader; and (2) How readers might position themselves—or find themselves positioned—in terms of the ideas represented and their own responses. These deliberations extend the tenets for a meaning-maker’s assessment of their own understanding. Beyond the scope and quality of their recall, readers are instead engaged in contemplating their own responses as actors in the author-reader exchange, and with the socio-cultural frames in mind. This requires a cultural reading of the people, ideas, and events, especially in terms of their politics, biases, and assumptions; the interests that are being pursued; the historical context; and how they (i.e., people, ideas, events, information) are mixed. It entails recognition of oneself and others as both audience for and subject of comments or portrayals for different purposes.

Final reflections: key considerations for global meaning-making

Global meaning-making does not offer a scripted reading of the world to be applied in a singular or monolithic fashion. Its dynamic processes are by their very nature diversified, multilayered, and multifaceted, involving fusions and adaptations of ideas and styles. Global meaning-making involves complex negotiations that are not preset or standardized, but anchored in ethics—ethics aligned with respect for the local, the pursuit of reciprocity between local and global, and ecological eclecticism ([Tierney, 2018a,b, 2020a,b](#)). The dimensions of global meaning-making outlined here are more akin to a set of values and guidelines—like compass points for telescopes, in search not of the destination but of the path forward.

Global meaning-making is a call to override the current predilection to exist in cultural cocoons and digital bubbles. It breaks away from pretailored worlds governed by practices and policies that perpetuate insularity, and use divisive representations of others (explicitly or implicitly) in public rallies, rarified scholarly spaces, or sites for teaching and learning. Instead, it involves what [Hymes \(1990\)](#) describes as a kind of dialectic between insider-outsider perspectives. Global meaning-making is rarely solitary; instead, engagements are participatory. It involves multiple colleagues and collaborators with local knowledges, Indigenous histories, migrant pasts, and cultural moorings from a range of places—including Oceania, Asia, the Americas and Africa.

The dance of global meaning-makers

For the global meaning-maker, these meaning-making engagements involve crisscrossing global time and space. It not only involves forms of cultural studies but also participatory engagements, sometimes with the aid of cultural intermediaries. At one level, a reader’s sophistication in terms of engaging with these processes is relative. It is anchored in their pre-existing knowledge about people, places, and times; their adroitness and the tools available to support them as they move with others within and across borders of space and time; and their ability to adapt and adjust to shifting norms and expectations. At another level, global meaning-makers require knowledge of themselves as they step in, out, and to the side of worlds to observe and engage with others with respect. Global meaning-making thus requires self-interrogation of one’s own enculturation—a continual scrutinizing of one’s interests, activities, positionality, perspectives, and biases. As [Spivak \(1988\)](#) cautions, such self-examinations should be ongoing, lest they become aligned with the systems they purport to challenge.

Global meaning-makers should be especially careful not to position themselves as the saviors or champions of others. As [Abdi \(2015\)](#) warns, well-intended allies or advocates should be alert to those tendencies that merely give the appearance of eclecticism:

[There is] a European predestination to save non-cultured natives from themselves ... [and] we should not discount ... the need to see beyond the fog of the still problematically benevolent political correctness as the creators of the new scholarship are somehow oblivious in turning the gaze upon themselves and societies (p. 16).

This undoubtedly requires a study of self that seeks to challenge both self-righteous objectives to enfranchise as well as failures to self-implicate. Global meaning-makers should, as [Spivak \(1988, 1990\)](#) suggests, be contemplative as they reconcile their complicity with their own privilege, and adopt dispositions and approaches that are not presumptuous, colonizing, or recolonizing.

Global meaning-makers as critical culturalists

Indeed, global meaning-making entails reckoning with oneself and one's cultural ways of knowing as one journeys across borders with others, with and for the interests of all. It represents a mix of participatory literacy, promoting approaches that are cooperative, collaborative, and contrastive while being respectful and reciprocal. It befits a planetary view that is ecumenical and emancipatory. A key thesis undergirding the rationale for global meaning-making is the advancement of "other" alongside of "all," in concert with accommodation for (rather than assimilation of) differences; that is, the pursuit of eclecticism in support of a global complementarity or inter-operationality. It entails a turn from self-righteousness to critical reflexivity, and from imposition and imperialism to respect and restraint. As global meaning-makers develop cross-border understandings, they also challenge their insularity, ignorance, and historic depreciation of others. They interrogate their complicity with their own advancement—reckoning with their own assertions of global jurisdictional mandates, or engagement in misinterpreting, misrepresenting, or disregarding the sovereignty, languages, ways of knowing and cultures of others. Similar to what Cochran-Smith (2000) suggested in her discussion of racism, global meaning-makers need to interrogate their "own complicity in maintaining existing systems of privilege and oppression" (p. 186).

Global meaning-makers are therefore expected to read for socio-political currents and pursue critical forms of advocacies that are committed to advancing diversity. In so doing, they engage in an ethics of respect for others and support for personal sovereignties. At the same time, they are never untethered from their own histories, predispositions, and positioning (e.g., the influence of mainstream imperial forces of colonization, as seen in marketing strategies tailored to digital user profiles). To fuse self with others, global meaning-makers must engage with persona and ethos, the pragmatics of language use, and notions of identity over space and time. As they consider the diverse circumstances within and across countries and cultures, their communications may take the form of translanguaging and other means of criss-crossing meaning-making communities. As Willinsky (1998) suggested, meaning-makers should engage not as imperialists but as critical culturalists, working across borders with a view of themselves as foreigners in support of others (see also: Kristeva, 1991).

Global meaning-making is contemplative, analytic, and advocacy-oriented

As noted by postcolonialists, meaning-makers can too often find themselves confined to sites where forms of epistemological imposition and resocialization occur. These insular sites might be tailored to Eurocentric traditions that befit colonialist or assimilationist models rather than those espousing epistemological eclecticism, Indigeneity, and internationalism (Abdi, 2015; Connell, 2007; Nozaki, 2009; Takayama, 2009; Takayama et al., 2017). Unwittingly, everyday meaning-making pursuits and understandings can be permeated by media that silences concerns and reinforce complicitness with regulatory practices. By contrast, the pursuit of global meaning-making is a call for exchanges to flow across and within a futuristic, pluralistic world. Its approach to everyday literacy involves shared responsibilities as well as alternative spaces that afford expression and exert change. These ideas reflect Gutiérrez's (2008) discussions of a third space—an aspirational hybrid space wherein exchanges flow across cultural identities and positionalities, in accordance with the socio-political dynamics of students' worlds and their communities. Within this third space, improvisations within and across borders and identities serve to both empower engaged individuals and improve the group.

Global meaning-making requires support for and commitment to engaging with forms of border-crossings, as one steps across or out of line. It is more provocative than neutral, and more disruptive than disassociated. In that way, it is similar to forms of counter discourse, coupled with proactive engagements, such projects or praxis, directed toward transformative change. To these ends, global-meaning making processes of contemplation, analysis, and advocacy are crossed with the following combination of stances:

- Responsive: To be open to consider and address matters of human rights and planetary responsibilities that arise or exist locally and globally;
- Perspectival: To enlist different perspectives, especially those stemming from considerations of context and relevance, so that engagements are respectful, responsive, and proffer understandings of events that illuminate different understandings;
- Evaluative: To delve into different readings to consider the assumptions, norms, and tenets that serve as the bases for perspectives and understandings—that is, to bear responsibility for judging the ideologies represented in and by the text, including the systemic forces at play that undergird societal hierarchies and frame exchanges;
- Reflexive: To maintain responsibility for recognizing one's frames and their nature or potential for influence—especially in terms of limiting or skewing understandings; to acknowledge self-interest; and to respect the interests of others (i.e., systemic forces, Indigenous interests and ways of knowing, etc.);
- Proactive and Transformative: To promote and pursue agency, advocacy and transformative change (i.e., forms of systemic change that address the development needs of communities in ways that are respectful, organic and sustainable).

In other words, global meaning-making involves a mindfulness toward the world. It demands agency, responsibly and respectfully acting upon socio-political discernments in ways that are ethical and community-based. It entails acceptance and reconciliation that are informed by a planetary epoch outlook and embrace plurality and universal rights. And it requires reckoning with, challenging, and changing hegemonies, with a reverence for the sovereignty and various ways of doing by others. Indeed, it is the standard for transformative engagements globally—distinguished from engagements that seek merely to understand others, analyze and critique

socio-political systems, embrace multiculturalism, or promote culturally-responsive or reciprocal education or pronouncements. Whereas prior notions of meaning-making may have defined bringing life to the page as simply building on and from one's background knowledge and experiences, global meaning-making represents a shift in the intimacy of one's engagement with texts and the world of media. It suggests discernments and contemplations that move beyond the page as one considers possibilities, recognizes one's role in relation to others, and acts carefully, respectfully, responsibly. Further, as action-oriented meaning-making, this process engages readers in interrogating their own readings and the socio-political dispositions of the texts before responding—seeking to eschew actions that are well-meaning but perhaps presumptuous or may align with systems and hegemonies that undermine their intent.

Closing summary

Global meaning-making assumes all behaviors are political; that responsible and responsive meaning-making respects and serves the interests all of our worlds; and that meaning-makers' engagements seek to change, challenge, or mitigate unjust systems. This view builds upon discussions among the growing circle of literacy scholars invested in global thinking, especially those researchers who investigate translanguaging, hybridity, global mobility (e.g., Lam and Warriner, 2012; Nelson et al., 2016; Pieterse, 2005; Rizvi, 2009a; Robertson, 1985), and global citizenship (e.g., Torres, 2015; UNESCO, 2015b). It stems from the search for other spaces, as discussed by Gutiérrez (2008, 2011) and Gutiérrez et al. (1999), and the argument for a more generous countenance that is respectful of cultures (e.g., Campano et al., 2010; de Sousa Santos, 2007b, 2013; Singh et al., 2005; Stein, 2017b). It draws heavily upon issues of mobility—in terms of people, cultures, and literacies. It is consistent with the model of community-based literacy events and practices explored by Victoria Purcell-Gates (2006) and her colleagues (Purcell-Gates et al., 2011); explorations of the participatory dynamics of literacy across time and space (e.g., Dyson, 1988, 1989; Jenkins et al., 2009); and pursuits of epistemological diversity, especially Indigenous ways of knowing (e.g., Assié-Lumumba, 2017; Battiste and Henderson, 2000; Bishop, 1994; Connell, 2007; de Sousa Santos, 2007a; Nakata, 2001, 2004; Rigney and Hattam, 2018; Rigney et al., 2018). It also aligns with the transliteracy framework outlined by Stornaiuolo et al. (2017). This transliteracy approach explores dimensions such as emergence, uptake, resonance, and scale as a way of capturing “different kinds of relations among people and things—whether in horizontal, vertical, rhizomatic, or other relationships—and highlight(ing) people's literacy practices within and across systems that (re) produce, exacerbate, and/or challenge social inequities” (Stornaiuolo et al., 2017, p. 84).

Broadly speaking, global meaning-making comprises a triad of critical contemplation, analysis, and advocacy, stemming from the convergence of socio-cultural, critical, and globalist views. As noted, it builds upon socio-cultural views of reading (e.g., Garcia and Kleifgen, 2020; Lee, 2020; Purcell-Gates, 2006); discussions of globalism (Rizvi, 2009a; Robertson, 1985; Singh et al., 2005); and postcolonial epistemologies (e.g., Connell, 2007; de Sousa Santos, 2013; Said, 1979, 1993). Global meaning-makers are akin to researchers as they position themselves, grapple with their identities, and scrutinize texts amid the cauldron of global socio-political forces involved. They examine events, settings, characters, and issues from different perspectives as they observe and participate aesthetically, vicariously, efferently, and critically. Their worlds (virtual, real, read, or reported) involve complex social dynamics within and across individuals, families, communities, and cultures. Hopefully, in keeping with discussions of pluralism, they come to see how the worlds of others are also part of their own. They may contemplate and contribute to transformative engagements arising from disrupted complacencies, changed understandings, and questions regarding the systems that perpetuate perceived cultural and planetary affronts.

Global meaning-making involves cultural moves that go beyond comprehension approaches focused upon just reading for understanding. It moves beyond evaluation of texts in terms of their factual bases, or the quality of their arguments. Instead, it entails examining texts, positions, and cultures—as well as the people within them. It asks meaning makers to unmask, reflect, and act upon socio-political analyses. It is at the cusp of enacting change that embraces pluralism and does not slight diversity. It aligns with challenges to our critical illiteracies and calls for reading our worlds.

References

- Abdi, A.A., 2015. Decolonizing global citizenship education: critical reflections on the epistemic intersections of location, knowledge, and learning. In: Abdi, A.A., Shultz, L., Pillay, T. (Eds.), *Decolonizing Global Citizenship Education*. Sense, Rotterdam, The Netherlands, pp. 11–26.
- Acevedo, E., 2018. *The Poet X*. HarperTeen.
- Ali, D., 2017. Safe Spaces and Brave Spaces: Historical Context and Recommendations for Student Affairs Professionals. *NASPA Policy and Practice Series*, Issue 2, October. https://www.naspa.org/images/uploads/main/Policy_and_Practice_No_2_Safe_Brave_Spaces_DOWNLOAD.pdf.
- Andreotti, V., de Souza, L.M.T.M., 2011. *Postcolonial Perspectives on Global Citizenship Education*. Routledge, New York, NY.
- Appadurai, A., 2005. Grassroots globalization and the research imagination. In: Appadurai, A. (Ed.), *Globalization*. Duke University Press, Durham, NC, pp. 2–21.
- Arendt, H., 1973. *The Origins of Totalitarianism*. Harcourt Brace Jovanovich, New York, NY.
- Assié-Lumumba, N.T., 2017. The Ubuntu paradigm and comparative and international education: epistemological challenges and opportunities in our field. *Comp. Educ. Rev.* 61 (1), 1–21. <https://doi.org/10.1086/689922>.
- Baldwin, J., 1974. *If Beale Street Could Talk*. Dial Press, New York.
- Battiste, M., Henderson, J.Y., 2000. *Protecting Indigenous Knowledge and Heritage: A Global Challenge*. Pirich, Saskatoon, Saskatchewan, Canada.
- Battiste, M.A., 1998. Enabling the autumn seed: toward a decolonized approach to aboriginal knowledge, language and education. *Can. J. Native Educ.* 22 (1), 16–27.

- Beach, R., Share, J., Webb, A., 2017. Teaching climate change to adolescents: reading, writing, and making a difference. Routledge/National Council of Teachers of English., New York/Urbana, Illinois.
- Beach, R. and Smith, B.E. (n.d.) Responding to and creating multimodal texts. In: Tierney R.J., Rizvi F., Ercikan K. (eds.), International encyclopedia of education, Fourth Edition. Elsevier, Oxford. In press.
- Beach, R., Tierney, R.J., 2016. Toward a theory of literacy meaning-making within virtual worlds. In: Israel, S. (Ed.), Handbook of Research on Reading Comprehension, vol. 2. Guilford, New York, NY, pp. 135–164.
- Behrendt, L., 2004. Home. University of Queensland Press.
- Bishop, R., 1994. Initiating empowering research? N. Z. J. Educ. Stud. 29, 175–188.
- Bishop, A., 2009. Becoming an Ally: Breaking the Cycle of Oppression. Fernwood, Halifax, Nova Scotia, Canada.
- Boal, A., 1993. Theater of the Oppressed. Theatre Communications Group, New York.
- Bourdieu, P., 1993. Language & Symbolic Power (J. Thompson, Ed., & G. Raymond & M. Adamson, Trans.). Harvard University Press.
- Burke, K., 1969. A Rhetoric of Motives. University of California Press, Berkeley, CA.
- Butler, J., 2011. Precarious Life and the Obligations of Cohabitation. Address given at the Nobel Museum, Stockholm, Sweden [YouTube video]. Retrieved from: <https://www.youtube.com/watch?v=KJT69AQdDtg>.
- Callimachi, R., 2021. Lost lives, lost culture: the forgotten history of Indigenous boarding schools. N. Y. Times. November 17, 2021. <https://www.nytimes.com/2021/07/19/us/us-canada-indigenous-boarding-residential-schools.html>.
- Campano, G., Honeyford, M.A., Sánchez, L., Vander Zanden, S., 2010. Ends in themselves: theorizing the practice of university–school partnering through horizontalidad. Lang. Arts 87 (4), 277–286.
- Choudry, A., 2015. Learning Activism: The Intellectual Life of Contemporary Social Movements. University of Toronto Press, Toronto, Ontario, Canada.
- Clover, D.E., Stalker, J. (Eds.), 2007. The Arts and Social Justice: Re-crafting Adult Education and Community Cultural Leadership. NIACE, Leicester.
- Clover, D.E., 2013a. The art of environmental adult education: creative responses to a contemporary ecological imperative. Int. J. Talent Dev. Creat. 2, 53–64.
- Clover, D.E., 2013b. The art of adult education and community cultural development in Canada. In: Nesbit, T., Taber, N., Brigham, S., Gibb, T. (Eds.), Building on Critical Traditions: Adult Education and Learning in Canada. Thompson Educational Publishing, Toronto, pp. 204–214.
- Cochran-Smith, M., 2000. Blind vision: unlearning racism in teacher education. Harv. Educ. Rev. 70 (2), 157–190.
- Connell, R., Collyer, F., Maia, J., Morrell, R., 2017. Toward a global sociology of knowledge: post-colonial realities and intellectual practices. Int. Sociol. 32 (1), 21–37. <https://doi.org/10.1177/0268580916676913>.
- Connell, R., 2007. Southern Theory: The Global Dynamics of Knowledge in Social Science. Polity Press, Cambridge, UK.
- de Sousa Santos, B., Nunes, J.A., Meneses, M.P., 2007. Opening up the canon of knowledge and recognition of difference. In: de Sousa Santos, B. (Ed.), Another Knowledge is Possible. Verso, London, England, pp. xix–lxii.
- de Sousa Santos, B. (Ed.), 2007a. Another Knowledge is Possible. Verso, London, England.
- de Sousa Santos, B., 2007b. Beyond abyssal thinking: from global lines to ecologies of knowledge. Eurozine 33, 45–89.
- de Sousa Santos, B., 2013. Epistemologies of the South: Justice Against Epistemicide. Paradigm. <https://doi.org/10.4324/9781315634876>.
- Deuze, M., Blank, P., Speers, L., 2012. A life lived in media. Digit. Humanit. Q. 6 (1). <http://www.digitalhumanities.org/dhq/vol/6/1/000110/000110.html>.
- Dion, M.L., Rios, C.M.D., Leonard, K., Gabel, C., 2020. Research methodology and community participation: a decade of Indigenous social science research in Canada. Can. Rev. Sociol. 57 (1), 122–146. <https://doi.org/10.1111/cars.12270>.
- Dyson, A.H., 1988. Negotiating Among Multiple Worlds: The Space/Time Dimensions of Young Children (Report No. 15). California University Center for the Study of Writing, Berkeley, CA (ERIC Document Reproduction Service No. ED297334).
- Dyson, A.H., 1989. Multiple Worlds of Child Writers: Friends Learning to Write. Teachers College Press, NY.
- Elhilo, S., 2021. Home is Not Country. Make a World. Random House.
- Enciso, P., 1997. Negotiating the meaning of difference: talking back to multicultural literature. In: Rogers, T., Soter, A. (Eds.), Reading Across Cultures. Teachers College Press, pp. 13–41.
- Enciso, P., 2004. Reading discrimination. In: Greene, S., Abt-Perkins, D. (Eds.), Making Race Visible: Literacy Research for Cultural Understanding. Teachers College Press, pp. 149–177.
- Escobar, A., 2018. Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds. Duke University Press.
- Freire, P., 1973. Education for Critical Consciousness. Seabury Press, New York, NY.
- García, O., Kleifgen, J.A., 2020. Translanguaging and literacies. Read. Res. Q. 55 (4), 553–571. <https://doi.org/10.1002/rrq.286>.
- García, O., Johnson, S., Seltzer, K., 2017. The Translanguaging Classroom. Leveraging Student Bilingualism for Learning. Caslon, Philadelphia.
- García, O., 2009. Bilingual Education in the 21st Century: a Global Perspective. Basil/Blackwell, Malden, MA and Oxford.
- Gee, J.P., 2001. Reading as situated language: a sociocognitive perspective. J. Adolesc. Adult Lit. 44 (8), 714–725.
- Giddens, A., 1999. Runaway World: How Globalization is Reshaping Our Lives. Profile, London, England.
- Goodwin, S., 2012. Conducting Action Research in an “Evidence-Based” Policy Environment. Australian Qualitative Research Conference “Embodying Good Research—What Counts and Who Decides?”. Charles Darwin University, Australia.
- Gramsci, A., 1982. Selections From the Prison Books. Lawrence & Wishart, London, England.
- Grigorov, S., Fleuri, R., 2012. Ecopedagogy: educating for a new eco-social intercultural perspective. Visão Glob. 15 (1–2), 433–454. Retrieved from: http://www.academia.edu/4111655/Grigorov_S_and_Fleuri_R._2012._Ecopedagogy_educating_for_a_new_eco-social_intercultural_perspective._Visao_Global_UNOESC_Florianopolis.
- Gutiérrez, K.D., Rogoff, B., 2003. Cultural ways of learning: individual traits or repertoires of practice. Educ. Res. 32 (5), 19–25. <https://doi.org/10.3102/0013189X032005019>.
- Gutiérrez, K.D., Baquedano-López, P., Tejeda, C., 1999. Rethinking diversity: hybridity and hybrid language practices in the third space. Mind Cult. Activ. 6, 286–303.
- Gutiérrez, K.D., Morales, P.Z., Martínez, D.C., 2009. Re-mediating literacy: culture, difference, and learning for students from nondominant communities. Rev. Res. Educ. 33 (1), 212–245. <https://doi.org/10.3102/0091732X08328267>.
- Gutiérrez, K.D., Cortes, K., Cortez, A., DiGiacomo, D., Higgs, J., Johnson, P., Ramón Lizárraga, J., Mendoza, E., Tien, J., Vakil, S., 2017. Replacing representation with imagination: finding ingenuity in everyday practices. Rev. Res. Educ. 41 (1), 30–60. <https://doi.org/10.3102/0091732X16687523>.
- Gutiérrez, K., 2008. Developing sociocultural literacy in the third space. Read. Res. Q. 43, 148–164.
- Gutiérrez, K.D., 2011. Designing Resilient Ecologies: Towards a Human Science of Learning. Wallace Foundation Distinguished Lecture, presented at the annual meeting of the American Educational Research Association, New Orleans, LA. Retrieved from: <http://www.cmccg.com/Media/Synch/310442/170/player.html>.
- Gutiérrez, K., 2014. Syncretic approaches to literacy learning: leveraging horizontal knowledge and expertise. In: Dunston, P.J., Fullerton, S.K., Bates, C.C., Stecker, P.M., Cole, M., Hall, A., Herro, D., Headley, K. (Eds.), 63rd Yearbook of the Literacy Research Association. Literacy Research Association, pp. 48–60.
- Gutiérrez, K.D., 2016. Designing resilient ecologies: social design experiments and a new social imagination. Educ. Res. 45 (3), 187–196. <https://doi.org/10.3102/0013189X16645430>.
- Hall, B., 1978. Continuity in adult education and political struggle. Convergence 11 (1), 8–15.
- Haluza-DeLay, R., 2003. When the topic is racism: research and advocacy with a community coalition. Soc. Justice 30 (4), 77–90.
- Hayhoe, R., 2021. Cross-cultural understanding and the listening intellect. Universities Intellect. 1 (1), 12–20. Comparative Education Research Centre. <https://cerc.edu.hku.hk/universities-and-intellectuals/1-1/cross-cultural-understanding-and-the-listening-intellect/>.

- Horowitz, R., 2012. Border crossing: geographic space and cognitive shifts in adolescent language and literacy practices. In: Romo, H., Garrido de la Calejia, C., Lopez, O. (Eds.), *A Bilateral Perspective on Mexico-United States Migration*. Universidad Veracruzana and The University of Texas at San Antonio, pp. 147–164.
- Hull, G.A., Stornaiuolo, A., 2014. Cosmopolitan literacies, social networks, and “proper distance”: striving to understand in a global world. *Curric. Inq.* 44 (1), 15–44. <https://doi.org/10.1111/curi.12035>.
- Hymes, D., 1990. Emics, etics, and openness: an ecumenical approach. In: Headland, Y., Pike, K., Harris, M. (Eds.), *Emics and Etics: The Insider/Outsider Debate*. SAGE, Newbury Park, CA, pp. 120–126.
- Jenkins, H., Purushotma, R., Weigel, M., Clinton, K., Robison, A.J., 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. The MIT Press. <https://doi.org/10.7551/mitpress/8435.001.0001>.
- Jenkins, H., Shresthova, S., Gamber-Thompson, L., Kligler-Vilenchik, N., Zimmerman, A., 2016. *By Any Media Necessary: The New Youth Activism, Connected Youth and Digital Futures*. New York University Press.
- Kay, M., 2018. Not Light, But Fire: How to Read Race Conversations in the Classroom. Stenhouse.
- Kim, G.M., 2016. Transcultural digital literacies: cross-border connections and self-representations in an online forum. *Read. Res. Q.* 51 (2), 199–219. <https://doi.org/10.1002/rq.131>.
- King, J.E., 2017. 2015 AERA presidential address morally engaged research/ers dismantling epistemological nihilism in the age of impunity. *Educ. Res.* 46 (5), 211–222. <https://doi.org/10.3102/0013189X17719291>.
- Kinsman, G., 2006. Mapping social relations of struggle: activism, ethnography, social organization. In: Frampton, C., Kinsman, G., Thompson, A.K., Tilliczek, K. (Eds.), *Sociology for Changing the World: Social Movements/Social Research*. Fernwood, Black Point, Nova Scotia, Canada, pp. 133–156.
- Kohlberg, L., 1981. *Essays on Moral Development*. In: *Philosophy of Moral Development*, vol. 1. Harper & Row, San Francisco, CA.
- Kristeva, J., 1991. *Strangers to Ourselves*. Columbia University Press, New York, NY.
- Kubota, R., Lehner, A., 2004. Toward critical contrastive rhetoric. *J. Sec. Lang. Writ.* 13 (1), 7–27. <https://doi.org/10.1016/j.jslw.2004.04.003>.
- Ladson-Billings, G., 1994. *The Dreamkeepers: Successful Teachers of African American Students*. Jossey-Bass, San Francisco.
- Lakoff, G., 2016. Understanding Trump's Use of Language [Blog Post]. Retrieved from. <https://georgelakoff.com/2016/08/19/understanding-trumps-use-of-language/#more-5208>.
- Lam, W.S.E., Warriner, D., 2012. Transnationalism and literacy: investigating the mobility of people, languages, texts, and practices in contexts of migration. *Read. Res. Q.* 47, 191–215.
- Lather, P., 1986. Issues of validity in openly ideological research: between a rock and a soft place. *Interchange* 17 (4), 63–84.
- Leander, K.M., Sheehy, M., 2004. *Spatializing Literacy Research and Practice*. Peter Lang, New York, NY.
- Lee, C.D., 2020. Social and cultural diversity as lens for understanding student learning and the development of reading. In: Moje, E.B., Afflerbach, P., Lesaux, N.K., Enciso, P. (Eds.), *Handbook of Reading Research*, vol. V. Routledge, New York, pp. 37–56.
- Levinas, E., 1981. *Otherwise Than Being or beyond Essence* (A. Lingis, Trans.). Martinus Nijhoff, The Hague, The Netherlands (Original work published 1974.).
- Levinas, E., 1993. *Outside the Subject* (M. B. Smith, Trans.). Athlone Press, London, England.
- Levinas, E., 1995. *Ethics and Infinity* (R. Cohen, Trans.). Duquesne, Pittsburgh, PA.
- Luke, A., 2004. Teaching after the market: from commodity to cosmopolitan. *Teach. Coll. Rec.* 106 (7), 1422–1443.
- Luke, A., 2011. Generalizing across borders: policy and the limits of educational science. *Educ. Res.* 40, 367–377.
- Luke, A., 2018. Digital ethics now. *Lang. Lit.* 20, 185–198.
- Massey, D., 2005. *For Space*. SAGE, London, England.
- McMurray, 2021. Teaching in the Age of Disinformation: Propaganda and Conspiracy Theories are Everywhere. What's a Professor to Do? [Part of the After the Insurgency series.] The Chronicle of Higher Education. <https://www.chronicle.com/article/teaching-in-the-age-of-disinformation>.
- Mercieca, J., 2015. The Rhetorical Brilliance of Trump the Demagogue. The Conversation. Retrieved from. <https://theconversation.com/the-rhetorical-brilliance-of-trump-the-demagogue-51984>.
- Ministère de l'Éducation, 2001. *Teacher Training: Orientations, Professional Competencies*. Gouvernement du Québec, Québec, Canada.
- Misiasek, G.W., 2015. Ecopedagogy and citizenship in the age of globalisation: connections between environmental and global citizenship education to save the planet. *Eur. J. Educ.* 50 (3), 280–292. <https://doi.org/10.1111/ejed.12138>.
- Nakata, M., 2001. Another window on reality. In: Osborne, B. (Ed.), *Teaching Democracy and Diversity*. Common Ground Publishing, Altona, Australia, pp. 331–353.
- Nakata, M., 2004. Ongoing conversations about aboriginal and Torres Strait Islander research agendas and directions. *Aust. J. Indig. Educ.* 33 (2004), 1–6. <https://doi.org/10.1017/S1326011100600807>.
- National Academies of Sciences, Engineering, and Medicine (NASEM), 2018. *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24783>.
- Ndimande, B.S., 2018. Unraveling the neocolonial epistemologies: decolonizing research toward transformative literacy. *J. Lit. Res.* 50 (3), 383–390. <https://doi.org/10.1177/1086296X18784699>.
- Nelson, N., Barrera IV, E.S., Skinner, K., Fuentes, A.M., 2016. Language, culture and border lives: mestizaje as positionality/lengua, cultura y vidas de frontera: El mestizaje como posicionalidad. *Cult. y Educ.* 28 (1), 1–41. <https://doi.org/10.1080/11356405.2014.980121>.
- New Zealand Herald, 2019. Christchurch mosque shootings: New Zealand Prime Minister Jacinda Ardern describes one of “darkest days”. N. Z. Herald. <https://www.nzherald.co.nz/nz/christchurch-mosque-shootings-new-zealand-prime-minister-jacinda-ardern-describes-one-of-darkest-days/ARBHXCLURLXRCWK5KMXEVLNSYA/>.
- Nozaki, Y., 2009. Orientalism, the West and non-West binary, and postcolonial perspectives in cross-cultural research and education. In: Apple, M., Au, W., Gandin, L. (Eds.), *International Handbook of Critical Education*. Routledge, pp. 382–390.
- Nussbaum, M.C., 1997. *Cultivating Humanity: A Classical Defense of Reform in Liberal Education*. Harvard University Press, Cambridge, MA.
- Ocholla, D., 2007. Marginalized knowledge: an agenda for Indigenous knowledge development and integration with other forms of knowledge. *Int. Rev. Ethics* 7, 236–245. <http://www.i-r-i-e.net/inhalt/007/26-ocholla.pdf>.
- Orwell, G., 1947. *Nineteen Eighty-Four: A Novel*. Harcourt Brace.
- Orwell, G., 1984. Politics and the English language. In: Orwell, G. (Ed.), *The Orwell Reader: Fiction, Essays and Reportage*. Harcourt Brace Jovanovich, San Diego, CA, pp. 355–366.
- Paré, A., 2004. Professionalism and Teacher Education in Québec. Paper presented at the European Conference on Education Research, Crete, Greece.
- Park, J., 2017. Knowledge production with Asia-centric research methodology. *Comp. Educ. Rev.* 61 (4), 760–779. <https://doi.org/10.1086/693980>.
- Pascoe, B., 2014. *Dark Emu: Aboriginal Australia and the Birth of Agriculture*. Magabala Books Aboriginal Corporation.
- Perry, M., Pullanikkatil, D., 2019. Transforming international development. *Impact* 2019 (9), 30–32. <https://doi.org/10.21820/23987073.2019.9.30>.
- Perry, M., 2020. Pluriversal literacies: affect and relationality in vulnerable times. *Read. Res. Q.* <https://doi.org/10.1002/rq.312>.
- Peters, M.A., 2017. Education in a post-truth world. *Educ. Philos. Theor.* 49, 563–566. <https://doi.org/10.1080/00131857.2016.1264114>.
- Pieterse, J.N., 2005. *Globalization and Culture*. Routledge, London, England.
- Purcell-Gates, V., Perry, K.H., Briseño, A., 2011. Analyzing literacy practice: grounded theory to model. *Res. Teach. Engl.* 45, 439–458.
- Purcell-Gates, V., 2006. What does culture have to do with it? In: Maloch, B., Hoffman, J.V., Fairbanks, C.M., Worthy, J. (Eds.), *National Reading Conference 55th Yearbook*. National Reading Conference, pp. 43–59.
- Rigney, L., Hattam, R.J., 2018. Toward a Decolonizing Culturally Responsive Pedagogy? Paper presented at the American Educational Research Association annual meeting, New York, NY.

- Rigney, D.M., Hemming, J., Bignall, S., 2018. Indigenous "Posthumanisms": Relational Sovereignty, "Country," and Decolonizing Pedagogies. Paper presented at the American Educational Research Association annual meeting, New York, NY.
- Rizvi, F., 2009a. Global mobility and the challenges of educational research and policy. *Yearbk. Natl. Soc. Stud. Educ.* 108, 268–289.
- Rizvi, F., 2009b. Toward cosmopolitan learning. *Discourse* 30 (3), 253–268.
- Rizvi, F., 2021. Creating Dynamic Cultures in Education Post-Covid-19. Melbourne Graduate School of Education Dean's Lecture Series [YouTube video]. <https://www.youtube.com/watch?v=wMWkxUElFKM>.
- Robertson, R., 1985. Glocalization: time-space and homogeneity-heterogeneity. In: Featherstone, M., Lash, S., Robertson, R. (Eds.), *Global Modernities*. SAGE, London, England, pp. 25–44.
- Rogers, T., Soter, A., 1997. *Reading Across Cultures*. Teachers College Press, New York, NY.
- Rogers, T., Winters, K., Perry, M., LaMonde, A., 2015. *Youth, Critical Literacies and Civic Engagement*. Routledge, New York, NY.
- Rogers, R., 2017. Literacy Research, Racial Consciousness, and Equitable Flows of Knowledge. Presidential address at the Literacy Research Association, Tampa, FL.
- Said, E.W., 1979. *Orientalism*. Vintage Books.
- Said, E.W., 1993. *Culture and Imperialism*. Vintage.
- San Pedro, T., 2018. Abby as ally: an argument for culturally disruptive pedagogy. *Am. Educ. Res. J.* 55 (6), 1193–1232. <https://doi.org/10.3102/0002831218773488>.
- Shi, Z., Li, C., 2018. Bourdieu's sociological thinking and educational research in mainland China. In: Mu, G.M., Dooley, K., Luke, A. (Eds.), *Bourdieu and Chinese Education*. Routledge, pp. 45–61. <https://doi.org/10.4324/9781315104331-3>.
- Simon, R., Campano, G., 2013. Activist literacies: teacher research as resistance to the "normal curve". *J. Lang. Lit. Educ.* 9 (1), 21.
- Singh, M., Lu, S.Y., 2020. *Postmonolingual Critical Thinking: Internationalising Higher Education Through Students' Languages and Knowledge*. Routledge. <https://doi.org/10.4324/9780367810405>.
- Singh, M., Fenway, J., Apple, M.W., 2005. Globalizing education: perspectives from above and below. In: Apple, M.W., Fenway, J., Singh, M. (Eds.), *Globalizing Education: Policies, Pedagogies and Politics*. Peter Lang, New York, NY, pp. 1–31.
- Smith, L.T., 1999. *Decolonizing Methodologies*. Otago University Press, Dunedin, New Zealand.
- Smith, G.H., 2000. Protecting and respecting Indigenous knowledge. In: Battiste, M. (Ed.), *Reclaiming Indigenous Voice and Vision*. University of British Columbia Press, pp. 209–224.
- Smith, L.T., 2005. On tricky ground: researching the native in the age of uncertainty. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*. SAGE, Thousand Oaks, CA, pp. 85–108.
- Smith, G.H., 2015. Transforming Research: The Indigenous Struggle for Social, Cultural, and Economic Justice Within and Through Education [Video]. Presentation at the American Educational Research Association annual meeting. Retrieved from: http://www.aera.net/EventsMeetings/AnnualMeeting/PreviousAnnualMeetings/2015AnnualMeeting/2015AnnualMeetingWebcasts/TransformingResearchTheIndigenousStruggleforSocial,Cultural,andEconomicJusticeWithinandThroughEducation/ta_bid/15952/Default.aspx.
- Soysal, Y.N., 1994. *Changing Citizenship in Europe: Remarks on Postnational Membership and the National State*. University of Chicago Press.
- Spivak, G.C., 1988. *Can the Subalterns Speak?* Macmillan, Basingstoke, UK.
- Spivak, G.C., 1990. *The Post-colonial Critic: Interviews, Strategies, Dialogues*. Routledge, New York, NY.
- Spivak, G.C., 1999. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Harvard University Press.
- Stein, S., 2017a. Contested Imaginaries of Global Justice in the Internationalization of Higher Education (Unpublished MA Thesis). The University of British Columbia, Vancouver, Canada.
- Stein, S., 2017b. The persistent challenges of addressing epistemic dominance in higher education: considering the case of curriculum internationalization. *Comp. Educ. Rev.* 61 (1), S25–S50.
- Stornaiuolo, A., Nichols, T.P., 2019. Cosmopolitanism and education. *Oxford Research Encyclopedia of Education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.252>.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Lit. Res.* 49 (10), 68–91. <https://doi.org/10.1177/1086296X16683419>.
- Sutton, P., Walshe, K., 2021. *Farmers or Hunter Gatherers? The Dark Emu Debate*. Melbourne University Press, Melbourne.
- Takayama, K., Sriprakash, A., Connell, R., 2017. Toward a postcolonial comparative and international education. *Comp. Educ. Rev.* 61 (S1), S1–S24. <https://doi.org/10.1086/690455>.
- Takayama, K., 2009. Progressive education and critical education scholarship in Japan: toward the democratization of critical educational studies. In: Apple, M., Au, W., Gandin, L. (Eds.), *International Handbook of Critical Education*. Routledge, pp. 354–367.
- Thiong'o, N.W., 1986. *Decolonising the Mind: the Politics of Language in African Literature*. Currey.
- Thiong'o, N.W., 2012. *Globalectics: Theory and the Politics of Knowing*. Columbia University Press.
- Tierney, R.J., Smith, G.H., Kan, W., 2021. Global Literacies research diversity: a manifesto for change. *J. Lit. Res.* 53 (3), 1–19.
- Tierney, R.J., 2017a. Multitopia: global citizenship across diverse spaces. *Glob. Commons Rev.* 1, 75–81. Retrieved from: <https://independent.academia.edu/RobTierney/Papers>.
- Tierney, R.J., 2017b. Populism and literacy consciousness in uncertain global times. *J. Lit. Res.* 49, 305–308.
- Tierney, R.J., 2018a. Global educational research in western times: the rise and plight of Chinese educational research. *Front. Educ. China* 13 (2), 163–192. <https://doi.org/10.1007/s11516-018-0010-4>.
- Tierney, R.J., 2018b. Toward a model of global meaning making. *J. Lit. Res.* 50 (4), 397–422. <https://doi.org/10.1177/1086296X18803134>.
- Tierney, R.J., 2020a. Redefining reading as global meaning-making. *Engl. J.* 109 (3), 93–95.
- Tierney, R.J., 2020b. Notes on global reading: critical cultural traversals, transactions and transformations. In: Misiaszek, L.I. (Ed.), *Exploring the Complexities in Global Citizenship Education: Hard Spaces, Methodologies, and Ethics*. Routledge, New York, pp. 36–68.
- Tilmouth, W.P., 2021. The children's graves at residential schools in Canada evoke the massacres of Indigenous Australians. *Guardian*. June 29, 2021. <https://www.theguardian.com/world/2021/jun/29/the-childrens-graves-at-residential-schools-in-canada-evoke-the-massacres-of-indigenous-australians>.
- Torres, C.A., 2015. Global citizenship and global universities: the age of global interdependence and cosmopolitanism. *Eur. J. Educ.* 50 (3), 262–279. <https://doi.org/10.1111/ejed.12129>.
- Toulmin, S., 1958. *The Uses of Argument*. Cambridge University Press, Cambridge, UK.
- Trump, D., Schwartz, T., 1988. *Trump: The Art of the Deal*. Random House.
- Truth and Reconciliation Commission of Canada, 2015. *Canada's Residential Schools: The Final Report of the Truth and Reconciliation Commission of Canada*, vol. 6. McGill-Queen's University Press, Montreal/Kingston. https://ehpmh2mwo3.exactdn.com/wp-content/uploads/2021/01/Volume_6_Reconciliation_English_Web.pdf.
- UNESCO, 2015a. *EFA Global Monitoring Report 2015: Gender and EFA 2000–2015—Achievements and Challenges*. UNESCO, Paris.
- UNESCO, 2015b. *Global Citizenship Education: Topics and Learning Objectives*. Retrieved from: <http://unesdoc.unesco.org/images/0023/002329/232993e.pdf>.
- United Nations Educational Social Cultural Organization (UNESCO), 2014. *Girls' and Women's Right to Education: Overview of the Measures Supporting the Right to Education for Girls and Women—Reported by Member States*. UNESCO, Paris.
- United Nations Girls Education Initiative, 2017. *Addressing Threats to Girls' Education in Contexts Affected by Armed Conflict*. United Nations, New York.

- United Nations Human Rights Council, 2017. Realization of the Equal Enjoyment of the Right to Education for Every Girl. United Nations Human Rights Council, New York (A/HRC/35/11.).
- Wandera, D.B., 2020. Resisting epistemic blackout: illustrating Afrocentric methodology in a Kenyan classroom. *Read. Res. Q.* 55 (4), 643–662. <https://doi.org/10.1002/rq.283n>.
- Willinsky, J., 1998. *Learning to Divide the World: Education at Empire's End*. University of Minnesota Press.
- Willmott, E., 1988. *Pemulwuy: The Rainbow Warrior*. Bantam Press, London.
- Woods, G., Biermann, S., 2009. Sharing knowledge, sharing experience, sharing power. *J. Aust. Indig. Issues* 12 (1–4), 116–131.

Cosmopolitan literacies and the limits of understanding: possible directions in global literacy research

Amy Stornaiuolo and Rabani Garg, University of Pennsylvania, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	710
Cosmopolitan literacies and asymmetric ignorance: the limits of understanding	710
Cosmopolitan literacies as ways of worldmaking: navigating incommensurability	712
Worldmaking as practice: a palimpsest (case study)	713
Conclusions: cosmopolitan literacies, ethical relationality, and the practice of mutual recognition	714
References	715

Introduction

In recent years, scholars have theorized the “global” in literacy research by turning toward critical perspectives, especially those that deliberately decenter the west. Some examples include a global literacies manifesto challenging “assimilative tendencies, a politic of Western exclusiveness, and racism” (Tierney et al., 2021, pp. 294–295) and a transnational consortium of authors collaborating on an international handbook to reframe “critical literacy work in diverse sociocultural contexts to counter the often-Eurocentric foundations of its academic lineage” (Pandya et al., 2022, p. 4). Such work builds on foundational scholarship theorizing the ideologically, culturally situated, and mobile nature of global literacy practices (Brandt and Clinton, 2002; Street, 1984, 2003) but also foregrounds issues of power and globalization in relation to colonial histories (e.g., Lammers and Astuti, 2021; Mora and Cañas, 2020; Sánchez and Ensor, 2021; Yoon et al., 2018). We are interested in exploring how this critical turn to studying the global in literacy research articulates with global frameworks like cosmopolitanism, a theory often traced across western thinkers from Diogenes to Kant (Kleingeld and Brown, 2019) that focuses on how people negotiate conflicting identities and loyalties as they make meaning across cultural, linguistic, and ideological differences (Appiah, 2006; Hansen, 2011).

In literacy research, cosmopolitanism has been a fruitful direction for literacy scholars interested in global issues. In scholarship about cosmopolitan literacies, for example, researchers have explored how cosmopolitanism’s central tenets—the recognition of people’s shared humanity and commitment to learning from and with one another through cross-cultural dialogue—are both rooted in local practices and intimately tied to people’s multimodal literacies on the move (e.g., Bean and Dunkerly-Bean, 2020; Campano and Ghiso, 2011; Choo, 2018; Hull et al., 2010; Hull and Stornaiuolo, 2014; Juzwik and McKenzie, 2015; Smith and Hull, 2012; Spires et al., 2019; Sánchez and Ensor, 2020; Vasudevan, 2014). In recent years, literacy scholars have critiqued more liberal approaches to cosmopolitanism rooted in western, Enlightenment-centered visions of global togetherness for forwarding neutral, essentializing, and universalizing conceptions of multicultural belonging that can obscure, elide, or reproduce differences as much as celebrate them. Scholars have drawn on critical cosmopolitan perspectives (e.g., Canagarajah, 2012; Delanty, 2006; Mignolo, 2011; Robbins and Horta, 2017) to theorize how these differences have emerged, tracing how power asymmetries rooted in histories of racism, colonialism, and oppression have shaped the current global landscape and people’s literacy practices across national, linguistic, and cultural contexts (e.g., Campano and Ghiso, 2011; Canagarajah and de Costa, 2016; Sanchez and Ensor, 2021; Stornaiuolo, 2016; Stornaiuolo and Nichols, 2019; Yoon et al., 2018; Zaidi and Choo, 2021).

Leaning into this critical direction, we consider whether cosmopolitanism remains a useful lens for theorizing the global in people’s meaning making practices. Specifically, we explore what a decolonial approach to cosmopolitanism might offer the field of literacy research. At the center of this exploration is our concern about the nature—and limits—of understanding: How can people understand one another when they are positioned within colonial systems that “reproduce the blind epistemic ethnocentrism” that centers normative (white, western) ways of knowing as the default (Mignolo, 2002, p. 66)? We begin by delving into this question of what it means to know across asymmetrical systems of power—and whose epistemological ways of being in the world are recognized, centered, and foreclosed as a result.

Cosmopolitan literacies and asymmetric ignorance: the limits of understanding

In a digitally connected but increasingly polarized world, one of the most pressing questions facing educators and researchers revolves around how people come to understand one another. Scholars proposing cosmopolitanism as an ethical framework for addressing this challenge suggest that dialog represents a key strategy for developing mutual understanding of our shared humanity even while “rooted” in specific local histories (e.g., Appiah, 2006; Hansen, 2010). Dialog has been an important cornerstone for literacy researchers interested in tracing people’s cross-cultural meaning making practices, which Hull and Stornaiuolo (2014) discuss through the lens of cosmopolitan literacies, defined as “the cognitive, emotional, ethical, and esthetic meaning-making

capacities and practices of authors and audiences as they take differently situated others into account" (p. 17). "Cosmopolitan literacies" has been a generative frame for understanding the ethical dimensions of how people make meaning with and for others across cultural, linguistic, geographic, and ideological differences in a digital era, as people draw on circulating texts, stories, media, and their epistemic experiences as resources for self-imagining in relation to others in the world (cf. [Campano and Ghiso, 2011](#); [Stornaiuolo, 2015](#); [Vasudevan, 2014](#)). In other words, theories of cosmopolitan literacies highlight the question of how people *come to understand* locally and globally through multimodal meaning making practices.

In an effort to address critiques of cosmopolitanism, particularly around its Eurocentric roots and tendency toward color-blind universalism, literacy scholars have turned to critical approaches that center issues of power and privilege and offer possibilities for transforming our world(s) through collective action. [Stornaiuolo and Nichols \(2019\)](#), for example, draw on cosmopolitical thought (e.g., [Stengers, 2010](#); [Watson, 2014](#)) to foreground the historical, sociomaterial conditions that reproduce inequities and the "active political labor required to construct common worlds across difference" in light of these conditions (p. 2). [Zaidi and Choo \(2021\)](#) propose cosmopolitan critical literacy as "a more progressive and proactive way of empowering students to engage with the multiplicity of cultures and cultural differences in our global age" in culturally and linguistically diverse classrooms (p. 453). [Kerkoff \(2022\)](#) suggests that a critical cosmopolitan lens involves using literacy practices to dismantle hierarchies and take justice-oriented action from a global perspective. And [Yoon et al. \(2018\)](#), guided by [Freire's \(1970\)](#) critical pedagogy and scholarship on global and multicultural education, suggest that the transformation of the world requires understanding our interconnectedness in relation to others. While these and the many other critical approaches to global and cosmopolitan literacy practices do attempt to take account of histories and systems of power, we suggest that they do not—and perhaps cannot—unsettle one of the fundamental tensions of cosmopolitanism: its role as a colonizing project for imagining the "other" that continues to center white, western ways of knowing.

Framed as a means for the west to conceive of the "other," cosmopolitanism is deeply intertwined with histories of colonialism, with individual notions of personal rights that are "rooted in maintaining otherness. The fact that one needs rights continually extended works to inscribe one as perpetually foreign" ([Barton and Tan, 2020](#), p. 435). These beliefs about individual rights intertwine with visions of shared humanity to cast individuals from the English-speaking "West" in central positions (e.g., as hosts, inviting "global others") and position others in the margins (e.g., those who are invited, imagined in the hosts' own image, perceived as not fully human). Discourses of hospitality (guest and host) and inclusion (inviting people into existing spaces) that perpetuate this "othering" continue to permeate cosmopolitan thinking—scholars consider how to create cosmopolitan dispositions, expand access to resources, and give voice to others, all the while projecting their imaginaries on people from their own vantage points. [Cusicanqui \(2012\)](#) describes this projection as continuing to center white, western, normative perspectives, arguing that inclusion, while seemingly benign, represents "a mitigated and second-class citizenship that molds subaltern imaginaries and identities into the role of ornaments through which the anonymous masses play out the theatricality of their own identity" (p. 100). Discourses of individual rights, hospitality, and inclusion make it *seem* like we are working toward a more just, shared world but continue to center colonial systems that reproduce normative perspectives, without making any conceptual or structural shifts needed to dismantle power asymmetries or understand people on their own terms ([Chakrabarty, 2000](#); [Cusicanqui, 2012](#); [de la Cadena, 2015](#)).

One of the ways this fundamental tension plays out in global literacy scholarship is that scholars who draw on critical cosmopolitan frameworks can continue to reproduce colonizing structures *without even recognizing that they are doing so*. [Chakrabarty \(2000\)](#) calls this "asymmetric ignorance," as scholars from subaltern positions must recognize and take into account western ontologies, but those from majoritarian positions do not need to understand or even acknowledge other ways of knowing, being, or theorizing (and often lump together "non-western" histories as beneath notice). This "inequality of ignorance" ([Chakrabarty, 2000](#), p. 28) means that scholars embedded in colonial discourses and systems do not even know that they do not know or recognize the labor of the subaltern, instead assuming that their ways of knowing are the default—something [Mignolo \(2002\)](#) calls "blind epistemic ethnocentrism" (p. 66). Literacy researcher [Thomas \(2019\)](#) describes this phenomenon in terms of the "imagination gap"—the impossibility of imagining worlds different from the ones we inhabit—as it pervades people's meaning making practices, locating such a gap in racial metanarratives that cast white characters as protagonists whose perspectives are centered and either omit characters of color entirely or relegate them to stereotypes or caricatures that revolve around the white protagonist. In light of this asymmetric ignorance that limits people's capacity not only to imagine another's perspective but to realize that they are in fact inhabiting different worlds, we wonder: just as "inclusion" has limits ([Barton and Tan, 2020](#)), what might be the limits of understanding?

To unpack this question, we turned to decolonial theories of cosmopolitan worldmaking ([Carrington, 2015](#)) and indigenous theorizing of the onto-epistemological nature of understanding ([de la Cadena, 2015](#); [Todd, 2016](#)). These theories highlight that to make space for diverse others, to understand them, and to highlight their "epistemic privilege" ([Campano, 2007](#); [Campano and Ghiso, 2011](#); [Mohanty, 2000](#); [Moya, 2000](#)) requires recognition of the "(Western/colonized) epistemic tools and ontology [that] allow us to see things a certain way, and in ways that close certain ontology and epistemologies to us" ([de la Cadena, 2015](#), p. xxv). While the logic of coloniality ([Quijano, 2000](#)) has meant that western subjectivity is not only more legible but also deeply embedded in our texts, visual media, and ways of reading the world to both those who are from the west and those who are not ([Chouliaraki, 2016](#); [Ghiso and Campano, 2013](#); [Leonardo, 2018](#); [Mignolo, 2001](#); [Mignolo and Ennis, 2001](#)), the reverse may not be true. Non-western epistemes and ontologies are not always recognized or engaged with by those who are situated in hegemonic western spaces ([Todd, 2016](#)). Critical frameworks like cosmopolitanism can still end up offering a "conditional inclusion" to the other based on what is legible to those who are rooted in the western epistemes ([Cusicanqui, 2012](#), p. 100). While those from other epistemes are able to (or have had to) read the western world, those working with hegemonic western epistemes are often unable to do so, creating asymmetric ignorance ([Chakrabarty, 2000](#))—the recognition of which serves as a productive starting place for the reimagining of cosmopolitan literacies.

Cosmopolitan literacies as ways of worldmaking: navigating incommensurability

Given that people's ways of understanding and making meaning in the world are bound up with the epistemic tools and ontologies available to them, an important place to begin any reimagining of cosmopolitan literacies centers on the practice of worldmaking. Worldmaking recognizes that there are many different world versions and visions with each of these of interest and importance to those who are involved in its making (Goodman, 1978), allowing for a radical pluralism that does not distill differences (Ricoeur, 1982). Expanding on Goodman's (1978) definition of worldmaking, Ricoeur (1982) points out that "no version (of the world) is the basic one and the other derivative," asking us to discard the idea that our worlds are versions of one and the same "neutral and underlying world" (p. 108). We use these explications offered by Goodman (1978) to emphasize the need to *recognize* that we (all) occupy fundamentally different worlds. This *recognition* does not attempt to understand differences, but rather accepts that we may never fully know each other. Our understandings of our worlds are shaped by our onto-epistemologies (Barad, 2003, 2007; Canagarajah, 2019; Watts, 2013) that are not just different but that these differences exist along with similarities: making meaning together does not depend on sharing overlapping notions. As anthropologist de la Cadena (2015), who engaged in a 10-year ethnographic study with the Quechua people, shares—creating equivalences erases differences that are just too palpable (p. 3). de la Cadena (2015), like Glissant (1997), asks us to sit in comfort with this understanding that we may know each other only with intermittencies and that we are limited in how much we can really get to know another (p. 3). Working from the premise that we can never fully understand another also opens the possibilities of (or recognition of) other ways of knowing and being in the world. Worldmaking involves a re/making of worlds from the worlds already on hand (Goodman, 1978). This remaking by people invites us to think, to feel and move in the (their) world as a way to build connections between worlds—to imagine otherwise (Storinau, 2015). Being attentive to worldmaking practices of others offers the possibility of knowing each other and the world more (not better) and rethinking the tools that we use to know the world.

While people's worldmaking practices offer the possibility of extending our knowledge about others (and the world), we recognize that our reading/understanding of their world is always going to be incomplete. There is much to learn from our differences (Appiah, 2006), yet much will escape a mutual understanding. Cosmopolitan frameworks call attention to the importance of creating a sense of a shared world so that we continue to learn from one another (e.g., Appiah, 2006; Hansen, 2010). However, even with shared spaces and practices one may never fully understand another. Our communication with each other does not depend on us having shared understandings, or making our "different notions equivalent" but knowing that we can never fully comprehend the other (de la Cadena, 2015, p. 3); in other words, our worlds may be incommensurate. de la Cadena's (2015) emphasis on knowing more about each other versus knowing better offers a way to navigate differences. It highlights for us that while *we can't make claims to understanding another, we can recognize that there are other ways of moving, being, and knowing in the world*. This recognition does not foreground one knowledge over another; instead, it allows us to "rethink differences in ontological terms" (Foster and Reichman, 2015). "Making space" for (each) other necessitates being in comfort with differences and not working toward creating equivalences or building common grounds. Practices that ask of us to work toward mutual or shared understandings may inadvertently work toward erasing differences and foreclosing certain knowledges, inevitably making some knowledges less.

To address these incommensurabilities—and recognize and navigate the asymmetric ignorance that is involved in worldmaking—requires an explicitly decolonial approach. Mignolo (2018) points out that colonial differences are masked by language and signs, so the work of decoloniality is to understand what is hidden in such denotations (p. 186). Decolonial options involve "unveiling the logic of coloniality and the prospective task of contributing to build a world in which many worlds will coexist" (Mignolo, 2011, p. 54). For Carrington (2015), the innovative worldmaking practices and strategies that black diaspora communities have adopted to navigate spaces for and by white supremacy represents cosmopolitan worldmaking. He defines it as the "conscientious orientation toward the necessity of acquiring and deploying knowledge that will enable one to negotiate multiple fields of possibility" (p. 254). Through these worldmaking practices people fashion themselves and the worlds they inhabit (Dattatreya, 2020) as the center of their world—the "negotiating of visibility on one's own terms" (Carrington, 2015, p. 250). Carrington invites us to recognize these cosmopolitan practices as a "concentrated effort to intervene critically in the conditions of cultural production" (p. 246). Artist Kehinde Wiley's work that Carrington studies counters the racially coded and fetishistic images of white artists that presents "blackness as the other" (p. 250) by redefining how black bodies are represented in his art. This is a powerful subversion of the normative visual culture around black bodies.

These practices that allow people to tell a story (or restore a mainstream narrative; Thomas, 2019) offer a "horizon imbued with potentiality" (Munoz, 2009, in Govindarajan, 2018), yet it would be remiss not to acknowledge that through this worldmaking, others are not necessarily inviting everyone to make sense of or engage with it. We think here of Toni Morrison's mission to write for and to black audiences apart from the white gaze and outside of the white imaginary. To this we add that the non-western other occupies space through their cosmopolitan worldmaking practices and also offers the possibility of transformation in societies by opening up space to imagine and engage with other worlds on their terms. Though these might be seen as radically subversive acts, subversion might not always have been the intent (Govindarajan, 2018): not everything is oriented or responsive to white, western imaginaries. We present a case study of worldmaking to illustrate how worldmaking practices that challenge any molding of the imaginaries about the subaltern.

Worldmaking as practice: a palimpsest (case study)

We turn now to a case study illustrating this vision of worldmaking, focusing on an artist, “@vvrzone,” who juxtaposes anime characters with images from Kashmir on Instagram. The images they use are a part of the popular imagination of Kashmir, ones that invoke its beauty—the post card images speak to the metaphors often used for Kashmir, such as “Paradise on Earth,” “Jewel in India’s crown,” and so on—images that place Kashmir in a beautiful pristine past. There are also images of more modern elements of Kashmir life, from downtown streets to cafes and bedrooms. It is on these very images that the artist layers anime characters and other anime elements (their profile describes it as *Kashmir x Anime*). In the images we share here (and for which we have received permission from the artist), we see the juxtaposed scenes from a bucolic or modern vision of Kashmir remixed with anime characters (Figs. 1 and 2).

The remix is at once beautiful and modern, with anime bringing in a stateless quality (Naylor and Helford, 2014) to offer the viewer a place from which to view a Kashmir that is not a contested land. Though we are not delving into the nuances of the various anime characters that the artist layers onto each image (each anime character has their own world, politics, and possible imaginings), broadly, the stylized characters and the speculative nature of anime provides for an imaginative liberation that visually disconnects (Posadas, 2014) or loosens Kashmir from its history of oppression and the current militarized state. Cities and bodies mutually define each other. By juxtaposing anime pictures on images of a place that faces a turbulent past, present, and future, the city and the (anime) body produce, transform, and define each other (Thompson, 2007). In this speculative reimagination, the artist and their land move to a transnational identity, in conversation with and connected to other modern states and subjects (Suan, 2018). The artist themselves become a modern subject. A basic google search for Kashmiri art reveals pages of traditional



Fig. 1 Nature image from @vvrzone's *Kashmir x Anime* series.



Fig. 2 City image from @vvrzone's *Kashmir x Anime* series.

folk art—far from the artist’s imaginings of *Kashmir x Anime*. Their worldmaking practices offer, to us as readers, a powerful response to the colonial representations imposed on Kashmir and the Kashmiri subject by Neocolonial India (Carrington, 2015) and to those who are a part of this networked space of artists on Instagram. In the images created by @vvrzone, everydayness is disidentified with the prevailing concepts about Kashmir.

For Goodman (1978), worldmaking “always starts from worlds already on hand; the making is a remaking” (cited in Nünning and Nünning, 2010, p. 2). @vvrzone’s stylistic explorations of the writing, the remixing and juxtaposing of anime characters on the postcard-like-pictures, is a remaking of time and space from their location as a Kashmiri. The new image is, however, not built as “a tabula rasa but a palimpsest” (Carrington, 2015, p. 252), where the original image has been effaced to make place for new, but the traces of the older image remain (worlds are always made from other worlds) and informs the viewer of the historical context and setting in which the art is produced. Art’s style is shaped by historical facts, offers insights into a particular way of seeing the world and to the artist’s worldmaking practices (Carrington, 2015). The artist’s worldmaking practices in this case reconceptualized Kashmir for insiders and outsiders, but that reconceptualization is also dependent on where one is located in relation to those histories. For example, how Rabani made sense of these histories of Kashmir intersected with her own experiences of having lived there as a child in the early 80s in a military camp, her understandings about the complicated history of the military presence in the region, and her broader sense of the place formed through friends, colleagues, and activists from Kashmir and the media. For Amy, none of these histories surfaced in the palimpsest, with her white, western, US-centric lens investing these bucolic scenes with esthetic appreciation for a beautiful place and for the inventiveness of layering in anime characters but without any historical sense of either—a place of asymmetric ignorance that made any deeper interpretation challenging (she was not even aware of these histories until Rabani made her aware and even then could not imagine the historical, political, social, and cultural resonances and implications).

This incommensurability of “understanding” across different interpretations provides a compelling argument for what a decolonial approach to worldmaking offers researchers interested in cosmopolitan literacies. Our interpretations of @vvrzone’s art are our own representations, superimposing our worldviews over theirs. Yet the important point here is that @vvrzone, as an artist, is asserting their worldview, offering their interpretive lens to others as resources for meaning making. What we as viewers do with those representations certainly matters—as history points out, efforts to silence, distort, and appropriate people’s art and texts is rampant, particularly for individuals from historically oppressed groups. Yet, our interpretations also *do not* matter in another sense—@vvrzone’s perspectives exist separate from our responses to their work, whole on their own without our validation or interpretation of them. Part of what makes this turn to worldmaking so important is that we must learn to be comfortable with incommensurability, to what we don’t and maybe cannot ever know. This does not mean we cannot work toward mutual understanding, but even when that understanding does not emerge, we can work toward mutual recognition—recognition of others’ representations of their own worlds, on their own terms and for their own purposes; and recognition of our own ignorance of others’ worlds, and the asymmetrical contours of that ignorance, as shaped by histories of empire, colonization, racialization, and patriarchy.

Conclusions: cosmopolitan literacies, ethical relationality, and the practice of mutual recognition

We turn now to consider what this decolonial, ontological turn in cosmopolitanism might mean for scholars interested in cosmopolitan literacies. We have argued that the practice of mutual recognition is undergirded by a decolonial cosmopolitanism that Mignolo identifies as “a radical shift in the geopolitics of knowing and being” (p. 37), one that “dwells in the borders, in exteriority, in colonial difference” (p. 39) and emerges not from the center but from the margins: from artists like @vvrzone and Kehinde Wiley who are using the multimodal semiotic tools at their disposal to engage in a “concentrated effort to intervene critically in the conditions of cultural production” (Carrington, 2015, p. 246). As Carrington (2015) suggests, cosmopolitan worldmaking from the borders involves purposeful and thoughtful efforts not just to understand but to *know more*—which allows people to “negotiate multiple fields of possibility” from new positions (p. 254). Such a profound onto-epistemological shift to theorizing literacy practices from the borders—a move that involves shifting from seeking mutual understanding to attempting to know more—can reposition the global in more ethical relations in literacy research, we suggest.

Such a shift draws on the recognition that our scholarship is intimately tied up with who we are and how we know, and that these ways of being/knowing are shaped by colonial histories that contribute to asymmetric ignorance (Chakrabarty, 2000). This intertwining of knowing and being was illustrated powerfully in our case study of worldmaking and is described in the work of Barad (2007): “Practices of knowing and being are not isolable; they are mutually implicated. We don’t obtain knowledge by standing outside the world; we know because we are of the world. We are part of the world in its differential becoming” (p. 185). Difference represents the central organizing principle in how we come to mean together in the world—what Indigenous scholar and activist Donald (2009; cited in Todd, 2016) calls ethical relationality:

Ethical relationality is an ecological understanding of human relationality that does not deny difference, but rather seeks to more deeply understand how our different histories and experiences position us in relation to each other. This form of relationality is ethical because it does not overlook or invisibilize the particular historical, cultural, and social contexts from which a particular person understands and experiences living in the world (p. 18).

Ethical relationality, in Donald (2009) and Todd's (2016) scholarship, suggests a powerful way we might come to know more, by recognizing the differences in our worlds, tracing the contours of those differences, and seeking to know more through our differences.

One important insight for literacy educators and scholars interested in what ethical relationality means for scholarly practice involves recognizing the limits of understanding, what de la Cardena (2015) describes as the inevitable incommensurability of different worlds. The way we live with and move between different worlds as part of our everyday activity—what she calls “cosmo-life”—involves accepting translation between worlds as always incomplete: coming to terms with the fact that not everything can be translated or understood, and we do not need to seek univocal meaning or smooth out differences (p. 286). For literacy educators and researchers, this means accepting that people will create representations and manage visibility on their own terms. If representation always involves being entangled with other people's meaning making, our task as scholars is to take account of our positionality, recognizing that asymmetric ignorance (Chakrabarty, 2000) forecloses and fails to recognize certain ways of knowing/being. The only way to dismantle the ongoing colonial legacies that continue to enmesh us, anthropologist Todd (2016) argues, is to “face our complicity head on” (p. 15).

One way we might face our complicity with colonial epistemologies, Todd (2016) suggests, is to pay attention to our positions relative to others with whom we are in dialog—and to seek out whose voices may be missing from the conversation. For those drawing on indigenous scholarship, for example, she cautions scholars rooted in colonial, western epistemes to take care in how they build on and represent ideas, being careful not to ignore, romanticize, or appropriate indigenous thought but to trace histories of thought with respect for the communities and individuals involved: “Indigenous thinking must be seen as not just a well of ideas to draw from but a body of thinking that is *living and practiced by peoples with whom we all share reciprocal duties as citizens of shared territories (be they physical or the ephemeral)*” (p. 17, emphasis original). For scholars and educators interested in global and cosmopolitan literacies, practicing mutual relationality involves paying specific attention to the ways in which they are complicit in invoking global frameworks that may carry asymmetrical framings that perpetuate in a less visible way savior logics that continue to position the West (and English) as the dominant paradigm from which all “others” are imagined and understood. Critical self-reflexivity might entail asking how the metaphors, descriptions, and representations they choose can reproduce the white, colonial gaze; or who is doing the labor of translating across intermittencies; or how to engage with people's ideas and representations on their own terms. These fundamentally ethical efforts to engage in mutual recognition of people's literacy practices across different worlds—in an effort to understand or know more, and not necessarily better—represent a potential path for global literacy research.

References

- Appiah, K.A., 2006. *Cosmopolitanism: Ethics in a World of Strangers*. W.W. Norton & Company, New York.
- Barad, K., 2003. Posthumanist performativity: toward an understanding of how matter comes to matter. *Signs* 28, 801–831. <https://doi.org/10.1086/345321>.
- Barad, K., 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, Raleigh, NC.
- Barton, A.C., Tan, E., 2020. Beyond equity as inclusion: a framework of “Rightful Presence” for guiding justice-oriented studies in teaching and learning. *Educ. Res.* 49 (6), 433–440.
- Bean, T., Dunkerly-Bean, J., 2020. Cosmopolitan critical literacy and youth civic engagement for human rights. *Pedagogies* 15 (4), 262–278.
- Brandt, D., Clinton, K., 2002. Limits of the local: expanding perspectives on literacy as a social practice. *J. Lit. Res.* 34 (3), 337–356.
- Campano, G., 2007. *Immigrant Students and Literacy: Reading, Writing, and Remembering*. Teachers College Press, New York, NY.
- Campano, G., Ghiso, M.P., 2011. Immigrant students as cosmopolitan intellectuals. In: Wolf, S.A., Coats, K., Enciso, P., Jenkins, C.A. (Eds.), *Handbook of Research on Children's and Young Adult Literature*. Routledge, pp. 164–176.
- Canagarajah, S., 2012. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge.
- Canagarajah, S., 2019. Weaving the text: changing literacy practices and orientations. *Coll. Engl.* 82 (1), 7–28.
- Canagarajah, S., de Costa, P., 2016. Scales analysis, and its uses and prospects in educational linguistics. *Linguist. Educ.* 34, 1–10.
- Carrington, A.M., 2015. The cultural politics of worldmaking practice: Kehinde Wiley's cosmopolitanism. *Afr. Black Diaspora An Int. J.* 8 (2), 245–257.
- Chakrabarty, D., 2000. *Provincializing Europe: Postcolonial Thought and Historical Difference*. Princeton University Press, Princeton, NJ.
- Choo, S.S., 2018. The need for cosmopolitan literacy in a global age: implications for teaching literature. *J. Adolesc. Adult Lit.* 62 (1), 7–12.
- Chouliraki, L., 2016. Cosmopolitanism. In: Gray, J., Ouellette, L. (Eds.), *Media Studies*. New York University Press, New York.
- Cusicanqui, S.R., 2012. Ch'ixinakax utxiwa: a reflection on the practices and discourses of decolonization. *S. Atl. Q.* 111 (1), 95–109.
- Dattatreya, E.G., 2020. *The Globally Familiar: Digital Hip Hop, Masculinity, and Urban Space*. Duke University Press, Delhi.
- de la Cadena, M., 2015. *Earth Beings: Ecologies of Practice Across Andean Worlds*. Duke University Press, Durham.
- Delanty, G., 2006. The cosmopolitan imagination: critical cosmopolitanism and social theory. *Br. J. Sociol.* 57 (1), 25–47. <https://doi.org/10.1111/j.1468-4446.2006.00092.x>.
- Donald, D., 2009. Forts, curriculum, and Indigenous Metissage: imagining decolonization of aboriginal-Canadian relations in educational contexts. *First Nations Perspect.* 2 (1), 1–24.
- Foster, R.J., Reichman, D.R., 2015. Forward. In: de la Cadena, M. (Ed.), *Earth Beings: Ecologies of Practice Across Andean Worlds*. Duke University Press, Durham, pp. xi–xiii.
- Freire, P., 1970. *Pedagogy of the Oppressed*. M.B. Ramos, Trans. Continuum, New York, NY.
- Ghiso, M.P., Campano, G., 2013. Coloniality and education: negotiating discourses of immigration in schools and communities through border thinking. *Equity Excell. Educ.* 46 (2), 252–269. <https://doi.org/10.1080/10665684.2013.779160>.
- Glissant, É., 1997. For opacity. In: *Poetics of Relation*. University of Michigan Press, pp. 189–195.
- Goodman, N., 1978. *Ways of Worldmaking*. Hackett Publishing, Indianapolis.
- Govindran, R., 2018. *Animal Intimacies: Interspecies Relatedness in India's Central Himalayas*. University of Chicago Press.
- Hansen, D.T., 2010. Cosmopolitanism and education: a view from the ground. *Teach. Coll. Rec.* 112 (1), 1–30.
- Hansen, D.T., 2011. *The Teacher and the World: A Study of Cosmopolitanism as Education*. Routledge, New York.
- Hull, G., Stornaiuolo, A., 2014. Cosmopolitan literacies, social networks, and “proper distance”: striving to understand in a global world. *Curric. Inq.* 44 (1), 15–44.
- Hull, G.A., Stornaiuolo, A., Sahni, U., 2010. Cultural citizenship and cosmopolitan practice: global youth communicate online. *Engl. Educ.* 42 (4), 331–367.

- Juzwik, M.M., McKenzie, C., 2015. Writing, religious faith, and rooted cosmopolitan dialogue: portraits of two American evangelical men in a public school English classroom. *Writ. Commun.* 32 (2), 1–29.
- Kerkhoff, S.N., 2022. A pedagogical framework for critical cosmopolitan literacies. *Appl. Financ. Econ.* 1–23. <https://doi.org/10.1080/1358684X.2022.2042673>.
- Kleingeld, P., Brown, E., 2019. Cosmopolitanism. In: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2019 ed. <https://plato.stanford.edu/archives/win2019/entries/cosmopolitanism/>
- Lammers, J.C., Astuti, P., 2021. Calling for a global turn to inform digital literacies education. *J. Adolesc. Adult Lit.* 64 (4), 371–377. <https://doi.org/10.1002/jaal.1103>.
- Leonardo, Z., 2018. *Dis-orienting western knowledge*. Cambridge J. Anthropol. 36 (2), 7–20. <https://doi.org/10.3167/cja.2018.360203>.
- Mignolo, W.D., 2001. Coloniality of power and subalternity. In: Rodríguez, I., López, M.M. (Eds.), *The Latin American Subaltern Studies Reader*. Duke University Press Books, Durham, pp. 224–244.
- Mignolo, W.D., 2002. The geopolitics of knowledge and the colonial difference. *S. Atl. Q.* 101 (1), 57–96. <https://doi.org/10.1215/00382876-101-1-57>.
- Mignolo, W.D., 2011. *The Darker Side of Western Modernity: Global Futures, Decolonial Options*. Duke University Press, Durham.
- Mignolo, W.D., 2018. Colonial/imperial differences: classifying and inventing global orders of lands, seas, and living organisms. In: Mignolo, W.D., Walsh, C.E. (Eds.), *On Decoloniality: Concepts, Analytics, Praxis*. Duke University Press, New York, USA. <https://doi.org/10.1515/9780822371779>.
- Mignolo, W.D., Ennis, M., 2001. Coloniality at large: the western hemisphere in the colonial horizon of modernity. *CR* 1 (2), 19–54.
- Mohanty, S.P., 2000. The epistemic status of cultural identity: on beloved and the postcolonial condition. In: Moya, P.M., Hames-García, M.R. (Eds.), *Reclaiming Identity: Realist Theory and the Predicament of Postmodernism*. University of California Press, Berkeley, CA, pp. 29–66.
- Mora, R.A., Cañas, C., 2020. Literacy studies in Colombia: toward repositioning our field and the circulation of knowledge production. *Res. Teach. Engl.* 54 (4), 439–442.
- Moya, P., 2000. Introduction. In: Moya, P.M., Hames-García, M.R. (Eds.), *Reclaiming Identity: Realist Theory and the Predicament of Postmodernism*. University of California Press, Berkeley, CA, pp. 1–28.
- Muñoz, J.E., 2009. *Cruising Utopia: The Then and There of Queer Futurity; Sexual Cultures*. New York University Press, New York.
- Naylor, A., Helford, E.R., 2014. Introduction: science fiction anime: national, nationless, transnational, post/colonial. *Sci. Fict. Film Televis.* 7 (3), 309–314.
- Nünning, A., Nünning, V., 2010. Ways of worldmaking as a model for the study of culture: theoretical frameworks, epistemological underpinnings, new horizons. In: *Cultural Ways of Worldmaking*. De Gruyter, Berlin, Boston. <https://doi.org/10.1515/9783110227567.1>.
- Pandya, J.Z., Mora, R.A., Alford, J.H., Golden, N.A., de Roock, R.S., 2022. *The Handbook of Critical Literacies*. Routledge.
- Posadas, B.T., 2014. Remaking Yamato, remaking Japan: space battleship Yamato and SF anime. *Sci. Fict. Film Televis.* 7 (3), 315–342.
- Quijano, A., 2000. Coloniality of power, eurocentrism, and Latin America. *Nepantla* 1 (3), 533–580.
- Ricoeur, P., 1982. Review of ways of worldmaking. *Philos. Lit.* 4 (1), 107–120. <https://doi.org/10.2307/2215375>.
- Robbins, B., Horta, P.L., 2017. Cosmopolitanisms. In: Robbins, B., Horta, P.L. (Eds.), *Cosmopolitanisms*. New York University Press, pp. 1–20.
- Sánchez, L., Ensor, T., 2020. “We want to live”: teaching globally through cosmopolitan belonging. *Res. Teach. Engl.* 54 (3), 254–280.
- Sánchez, L., Ensor, T., 2021. Narrating global literacies: crossing borders of exclusion during a time of crisis. *J. Lit. Res.* 53 (2), 265–287.
- Smith, A., Hull, G.A., 2012. Critical literacies and social media: fostering ethical engagement with global youth. In: *Critical Digital Literacies as Social Praxis: Intersections and Challenges*. Peter Lang, pp. 63–84.
- Spires, H.A., Kerkhoff, S.N., Fortune, N., 2019. Educational cosmopolitanism and collaborative inquiry with Chinese and US teachers. *Teach. Educ.* 30 (4), 437–454. <https://doi.org/10.1080/10476210.2018.1506431>.
- Stengers, I., 2010. *Cosmopolitics I*. University of Minnesota Press, Minneapolis.
- Stornaiuolo, A., 2015. Literacy as worldmaking: multimodality, creativity and cosmopolitanism. In: *The Routledge Handbook of Literacy Studies*, vol. 37, pp. 561–571. <https://doi.org/10.4324/9781315717647.ch37>.
- Stornaiuolo, A., 2016. Teaching in global collaborations: navigating challenging conversations through cosmopolitan activity. *Teach. Teach. Educ.* 59. <https://doi.org/10.1016/j.tate.2016.07.001>.
- Stornaiuolo, A., Nichols, T.P., 2019. Cosmopolitanism and Education. *Oxford Research Encyclopedia of Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.252>.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge UP.
- Street, B., 2003. What’s “new” in New Literacy Studies? Critical approaches to literacy in theory and practice. *Curr. Issues Comp. Educ.* 5 (2), 77–91.
- Suan, S., 2018. Consuming production: anime’s layers of transnationality and dispersal of agency as seen in Shirobako and Sakuga-Fan practices. *Arts* 7 (3), 27.
- Thomas, E., 2019. *The Dark Fantastic: Race and the Imagination From Harry Potter to the Hunger Games*. NYU Press.
- Thompson, K.A., 2007. Performing visibility: freaknec and the spatial politics of sexuality, race, and class in Atlanta. *Drama Rev.* 51 (4), 24–46.
- Tierney, R.J., Smith, G.H., Kan, W., 2021. Global literacies research diversity: a manifesto for change. *J. Lit. Res.* 53 (3), 294–312. <https://doi.org/10.1177/1086296X211031262>.
- Todd, Z., 2016. An indigenous feminist’s take on the ontological turn: “ontology” is just another word for colonialism. *J. Hist. Sociol.* 29 (1), 4–22. <https://doi.org/10.1111/johs.12124>.
- Vasudevan, L.M., 2014. Multimodal cosmopolitanism: cultivating belonging in everyday moments with youth. *Curric. Inq.* 44 (1), 45–67. <https://doi.org/10.1111/curi.12040>.
- Watson, M.C., 2014. Derrida, Stengers, Latour, and subalternist cosmopolitics. *Theor. Cult. Soc.* 31 (1), 75–98.
- Watts, V., 2013. Indigenous place-thought and agency amongst humans and non-humans (first woman and sky woman go on a European tour!). *DIES* 2 (1), 20–34.
- Yoon, B., Yol, Ö., Haag, C., Simpson, A., 2018. Critical global literacies: a new instructional framework in the global era. *J. Adolesc. Adult Lit.* 62 (2), 205–214.
- Zaidi, R., Choo, S.S., 2021. Disrupting xenophobia through cosmopolitan critical literacy in education. In: Pandya, J.Z., Mora, R.A., Alford, J.H., Golden, N.A., de Roock, R.S. (Eds.), *The Handbook of Critical Literacies*. Routledge, pp. 447–455.

Literacy and international development

Daniel A. Wagner^a and Walid Hedidar^b, ^a Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States; and ^b Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

History and definitions of literacy	717
Life-span perspectives	717
Contemporary issues in literacy	718
Languages and literacies in contact	718
Technology and literacy	719
Literacy in times of crisis	720
Moving forward	720
References	721

History and definitions of literacy

Literacy is a complex and dynamic phenomenon that varies across history, language, and geography. There are competing views about the origins of literacy, which can be traced back thousands of years. For centuries, the formalization and educational roles of literacy as a tool for communication were dominated by a small portion of “educated” human society. However, over the years, many societies have experienced transitions from mostly illiterate to predominantly literate populations through various means involving both formal schooling and informal learning (in everyday practice). The present review considers literacy especially from an international development perspective.

UNESCO (2021)¹ estimates that about 86% of adults worldwide can read or write to some extent today. Nonetheless, the transition to widespread literacy has not been uniform across societies. Contexts vary greatly across the globe, and these have played a significant role in the motivation and support for acquiring literacy. For example, in Sweden, reading ability reportedly rose to above 90% by the mid-18th century as a result of the push from church and state on families to teach Bible reading at home. Among the Vai people of West Africa, literacy in the form of an indigenous script was developed and used for economic and personal written communication since the nineteenth century, a tradition that was transferred from one generation to another outside the structure of formal schooling (Scribner and Cole, 2013). This was also the case for many emerging nations, like Brazil, Tunisia and Turkey, that have moved to industrialize and enter the global market. In most of these contexts “being literate” entailed the ability to read and write in the country’s national (and post-colonial) language.

From these and other examples, it is clear that an understanding of social context is essential for understanding how literacy functions as a cultural phenomenon across the world. For decades, the framing of literacy in developing countries was strongly influenced by colonial epistemologies and the logic of the *mission civilisatrice* (Venezky et al., 1990). Within this perspective, literacy was defined principally in terms of the traits of the colonizers and elites, while the term “illiteracy” often characterized those who were poor and inferior. This is one reason why reading and writing in Western languages still constitutes the core of the global literacy agenda in which non-standard, non-European languages are mainly devalued (Wu et al., 2021). Only in recent decades has the field of literacy moved away from this restricted conceptualization.

The advent of modern public education, the increasing use of technology, and economic globalization are among the many factors that have contributed to changes in how literacy has been used over recent years. At a macro-level, the expansion of literacy work reflects the increased social complexities of our rapidly changing societies in the 21st century. At a micro-level, changes take place at different stages of an individual’s life through the learning of cognitive skills in and out of school, as well as through the diverse needs of everyday life.

Thus, formal definitions of literacy have varied over the years. A half-century ago, the United Nations thought of literacy as a dichotomous concept—you are either “literate” or “illiterate,” and that is how international statistics represented literacy in most countries (Collins, 1995; Wagner, 2011). Today, most specialists and governments think of literacy as a continuum of reading, writing, and numeracy skills that can be acquired in and outside of schools through training and practice. While some specialists today view literacy as how schools teach children how to read and write, others explain the degree of literacy acquisition as a consequence of (and contributor to) power, hierarchy, poverty and culture.

Life-span perspectives

Another way to think about literacy acquisition is to consider skills across the life-span. In the first few years of life, early language and literacy development has an impact on a child’s potential for future success. Studies have shown that skills developed during

¹The United Nations Educational, Scientific and Cultural Organization.

these years can affect future acquisition of other competences. As a result, there is a major movement toward support of early childhood development programs to provide children with basic cognitive, social, and emotional skills that set the foundation for learning literacy, numeracy and more advanced skills at later ages (UNICEF, 2019). Conversely, a lack of learning opportunities during the early years can increasingly hinder progress in the acquisition of higher order skills and future educational outcomes.

With the global growth of mass public education, primary schooling is the most common starting place for formal literacy learning, typically beginning with children between the ages of 5 and 7 years. Despite widely varying differences in primary schools across the world, the importance of literacy remains central to the curriculum across the globe (Wagner, 2018a,b). Increasing efforts are being made to assess literacy levels among primary and secondary school students internationally (Beeharry, 2021). Overall, the rapid expansion of access to primary schooling over the past few decades has resulted in a steady increase in literacy rates across the globe. Even so, increasing school access has not been able to address serious problems of low literacy in low-income countries, or among many poor children in middle- and high-income countries (Gove and Cvelich, 2011). In many low-income countries (e.g., Senegal, Bangladesh, South Africa and others), a large percentage of children are still unable to read with comprehension at the end of the fifth grade. This not only poses a serious challenge to children's future education, but will also have an impact on their potential contribution to economic growth.

Among youth (aged 15–24 years), literacy rates have been slowly improving over the past two decades due to an increased level of investment at the primary school level, notably from 83% in 1990 to 92% in 2021, according to UN statistics (UNESCO, 2021).² These literacy trends reflect increased access to primary and secondary education for younger generations and suggest that adult literacy rates will climb as youth replace older cohorts in the future. However, even now, more than 100 million youths lack basic literacy skills, mainly in two regions: South and West Asia and sub-Saharan Africa. These literacy rates are considerably lower among rural youth and girls in low-income countries. A major part of this problem results from young people who have never attended school or left school before acquiring strong literacy skills.

Adult literacy has improved somewhat for a majority of countries in the world over the past two decades. This is in part due to the ratification of the 2030 United Nations Sustainable Development Goals (SDGs), comprised of 17 goals adopted by its member states in 2015. SDG 4 is dedicated to education broadly, while its sub-goal SDG 4.6 aims to “ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy.” Even so, many countries failed to make significant progress over this same time period, a reflection of the limited national and international investments made, and also of the complexity of motivating adults to take part in such largely voluntary programs. Demographic changes, including the increase in access to primary and secondary schooling, will no doubt continue to improve adult literacy rates, but substantial low-literacy and illiteracy will likely persist well into the future.

Thus, in spite of the considerable efforts of governments, school systems and other agencies over many decades, achieving the improvement of, and equity in, literacy development has faced considerable challenges. Both *access* (such as student enrollment and retention) and *quality* (of instruction and student learning) are critical. Barriers to student enrollment and quality education exist, particularly for children growing up in poor contexts and in developing countries. These often include structural limitations, such as lack of school infrastructure, underpaid or inadequately trained teachers, and lack of appropriate textbooks. In youth and adult literacy programs, a major challenge is to meet the diverse needs of learners who are no longer in formal school programs.

Contemporary issues in literacy

Languages and literacies in contact

One of the key policy issues in global education is the language of instruction (LOI). In low-income countries, large numbers of children (sometimes the majority) come to school not knowing the LOI in the classroom (e.g., Hornberger, 2003). Educators and policymakers have long debated the role of LOI relative to local and indigenous languages, and more recently how literacies (and orthographies) come into contact in classrooms and beyond (Wu et al., 2021). On the one hand, these complex learning contexts require practical solutions that confront political realities—such as the need for proficiency in national and official languages—and an education system's ability to adapt to different languages and literacies along with effective ways of teaching. On the other hand, the cognitive and educational realities of how children learn to read must be acknowledged. Evidence suggests that children from minority language groups do the least well in school, irrespective of which languages are already spoken and which must be learned (Wagner, 2017).

Is there a best way to implement a LOI policy? There are no easy answers to this question, due to variations in national histories and traditions as well as uncertainties about the nature of globalization, language change, and language preferences (Valdiviezo, 2013). But there is mounting evidence that teaching students to read in their mother tongue is a crucial building block for effective learning (Hélot and Ó Laoire, 2011).

Education practitioners have attempted to take a more inclusive approach to students' mother tongues and cultures through policies of multilingual education (Trudell, 2016). In recent years, there has been a strong movement toward the use for early

²Official UN literacy rates are rarely based on direct measurements, but are rather estimates linked to schooling rates (Wagner, 2011).

reading (as early as the pre-primary years) of the child's first language (L1), as contrasted with a second language (L2), especially when considering students in disadvantaged communities who are from language minority groups. In a recent review, the [World Bank \(2021\)](#) found that 37% of children in developing countries are not being taught in L1, or about 250 million students worldwide. The same study supports the view that learning in L1 would allow children to have more time to gain initial proficiency in reading before moving on to L2, so that they could stay in school longer.

Even so, L1 does not always correspond to the mother tongue(s) spoken by the community at large. This is especially common with diglossia: a situation in which two or more dialects of the same language are used by some speakers in different contexts ([Ferguson, 1959](#)). Diglossia involves the use of two language varieties: one often designated as "High" and used for formal, written interactions; the other is the native variety designated as "Low" and used for informal, everyday interactions. In North African countries, for example, schooling in L1 is done in the "High" variety of Arabic (*Fusha*), as opposed to the local Arabic dialects (*Darija*). Teaching schoolchildren to read in a non-native variety of their language raises a set of unique challenges to learning and literacy in the early years. Within similar linguistic contexts, children are also often faced with an additional challenge of navigating multiple literacies at the same time. In Tunisia, for instance, children have to juggle not just the speech varieties between *Darija* and *Fusha*, but also the differing orthographies of the official L1 (Arabic) and L2 (French).

As noted above, the L2 used in schools typically corresponds to the language of the former colonizing country. This is the case for many former French colonies in Africa who have adopted French as their L2. Today, national education policies often require children to learn in languages that they and their teachers do not know well, an issue that contributes to low levels of reading achievement. Resistance to mother-tongue instruction by parents is also a recurring issue across contexts, and is often due to the parental views concerning national identity and socio-economic mobility. Still, language contexts are often highly dependent on local cultural variables so that, in some contexts, learning in a second language first does not always produce inferior results ([Wagner et al., 1989](#)).

Teachers play a major role in any language-sensitive education program. Often, teachers are not sufficiently trained to actually implement L1 reading policies because they were trained in government-run institutions in an official L2 language ([Piper and Miksic, 2011](#)). And, they may also lack the quality teaching and learning materials needed for teaching in L1. In other words, the ability of schools and teachers to implement a bilingual policy is necessarily limited by teachers' language competencies. Further, teachers' actual use of the school-mandated language of instruction may vary widely in practice, resulting in significant differences in students' language mastery by region and by instructor ([Ball et al., 2014](#); [Commeyras and Inyega, 2007](#)). In numerous countries, teacher deployment policies (i.e., putting teachers with particular language skills where they are most needed) often do not match the national language of instruction policy. For education systems with strongly centralized teacher recruitment and allocation processes, a structural shift in these processes is needed in order to ensure an adequate match between teacher's oral and written language skills and the levels and needs of the students ([World Bank, 2021](#)).

Overall, key questions around L1 and L2 languages and literacies remain unresolved: When should teaching of L2 reading begin if it is a nationally required language? Should L2 reading instruction co-occur with curricular content in L2? What are the best ways to support the transfer of reading skills from one language to another, and when should the transition from L1 to L2 occur ([Benson and Wong, 2015](#))? In contexts of conflict, fragility, or migration, how should language policies be shaped to accommodate students' learning? How can language of instruction policies be inclusive and responsive to the needs of students with special learning needs? The complexities of language of instruction continue to be a major challenge to reading development worldwide.

Technology and literacy

Technology has become indispensable for the ways in which we organize our lives. The [International Telecommunications Union \(2019\)](#) reports that 53.6% of the world's population has access to the Internet with a continuing upward trend in information and communications technologies (ICT) use and access. Additionally, almost all of the world's population lives within range of a mobile-cellular network. However, the lack of skills required for ICT use by students and learners of any age remains an obstacle to Internet access.

ICTs (including the Internet and mobile phones) are having a major impact on our thinking about literacy as well. Observational surveys ([OECD, 2019](#)) indicate that young learners in wealthier communities actively use websites, message boards, social media, and so on. If given a choice, they often prefer social interaction on the Internet or mobile devices instead of listening passively to an instructor or reading a textbook. Others have found that reading skills are significantly impacted by continuous interaction with web-based reading. It is also clear that young children with access to mobile devices have many new opportunities and ways to learn to be literate ([UNESCO, 2014](#)). In addition, ICTs provide a "lower barrier for entry" for learners to create texts themselves, rather than to simply consume written material.

If designed appropriately, ICTs can help to make the development of teaching and learning materials for language and literacy learning easier, cheaper, and faster. There are several existing tools that serve to create and adapt reading books for practice, textbooks, and teachers' guides. ICT-based reading programs have the capacity to determine appropriate linguistic scopes and sequences. They can also produce decodable text materials that children can use to supplement their learning. For example, ICT tools can follow student's learning sequentially, starting out with common and easier sounds and letters or symbols and moving up to less common and complex ones ([Castillo and Wagner, 2019](#)).

With recent developments in artificial intelligence, technology is gaining the ability to analyze spoken language and even correct pronunciations. Students can use such technologies to check their own performance and learn about the meaning of words, phrases and spoken or written discourses, and receive tailored and personalized tasks to their performance levels, of particular use in multi-level classrooms (Castillo et al., 2022). Additionally, in contexts where little is known about the languages spoken, technologies like data mining from cell phones or social media can help policy makers in mapping and tracing languages in particular communities or settings in a quick and cost-effective manner.

Despite these promising benefits, access to learning technologies remains limited in many communities in low-and-middle-income countries. Barriers to access are related to lack of infrastructure as well as to limited financial resources to equip schools with the internet, software and devices. In addition, teachers may not be trained or equipped with the resources to best use these tools in a way that supplements the curriculum. For example, beyond the classroom students might take their devices or programs home in order to continue their learning, but low levels of digital literacy among parents might limit students' ability to use technology effectively (UNICEF, 2020).

Taking advantage of the benefits of ICTs on literacy acquisition necessitates the presence of an ecosystem adapted to both curricula and instruction that supports the digital literacy of both students and teachers. Without such progress, ICTs will be most effective only to a select group, thereby widening the gap (or "digital divide") of inequality and further marginalizing the learning of children at the "bottom of the pyramid" (Wagner et al., 2018, 2022).

Literacy in times of crisis

The outbreak of the COVID-19 pandemic in 2020 has had a serious negative impact on the quality of education globally, and on literacy skills specifically. At the beginning of the pandemic, schools were closed in more than 190 countries, affecting the education of 90% of the world's student population of 1.57 billion (UNESCO, 2021). Overall, about 40% of learners living in remote areas, the poor, linguistic minorities and learners with disabilities in low-and lower-middle-income countries have not been adequately supported during the pandemic (UNESCO, 2021).

Many low-literate and non-literate children, youth and adults (especially women) have been disproportionately hit harder by the educational, social and economic impacts of the pandemic (UNESCO, 2021). In many countries, the response to the pandemic has involved replacing in-person instruction with virtual instruction, such as via TV, radio instruction, or social media to conduct virtual classes (Wagner et al., 2020; UNESCO, 2021). As a result, many countries have not been able to support remote learning, particularly in low-income countries (Vegas, 2020). The pandemic has clearly exposed the lack of preparedness of education systems to shift to distance teaching and learning. Even in places where remote instruction was technologically possible, there remained difficulties in developing pedagogical models and materials appropriate for interactive distance learning (UNESCO, 2020). As such, learning losses in literacy caused by the pandemic are projected to last well into the future (Angrist et al., 2021).

COVID-19 is not the only crisis facing education systems across the world. Save the Children (2021) estimates that the education of around 38 million children is disrupted each year by the climate crisis. As a result, there is a need for retaining and integrating programs that worked well during the pandemic into education systems on an ongoing basis. Such crises naturally result in negative consequences for the promotion of literacy.

Moving forward

In recent years, we have learned a great deal more about literacy and international development. This includes not only the acquisition of reading and writing skills but also more broadly the ways in which individuals communicate, create, and compute across cultures and contexts. Literacy encompasses the tools and competences needed for individuals to develop their knowledge and potential as well as participate in and contribute to the collective. The continuous reshaping of society and the emergence of new technologies have catalyzed the inception of multiple literacies such as information, computer, and digital literacy. These and other skills are create different possibilities of becoming and being literate. A lifespan approach provides a way to conceptualize literacy acquisition as a multidimensional and dynamic process from early childhood through adulthood.

Clearly, there are numerous challenges facing literacy development across the world. Inadequate language policies in low and middle-income countries often act as a barrier to children's journey of literacy acquisition. In schools in Morocco, for example, many children start most of their early schooling years either in a language completely different from their native language (as is the case with the *Amazigh* population) or in their native dialect (*Darija*). While for school children or adult literacy programs, the target literacy is typically in L2 (Modern Standard Arabic). One exception was an innovative and promising project that taught illiterate women who only spoke *Amazigh* to learn to read in that language using the Arabic orthography (Wagner and Maamouri, 2016).

Yet, in numerous multilingual but disadvantaged contexts, students often reach later grades with significant gaps in their L2 proficiency, which in turn affects their learning and academic advancement. The success of multilingual instruction lies in the support resources provided to students and teachers in building a foundation in L1 and then transitioning to L2. Capitalizing on pre-primary education to start L1 instruction, providing students with tailored curricula and books, and providing teachers with adequate training are all tools that research has shown to be effective in ensuring the success of multilingual language policies (Benson and Kosonen, 2021).

Overall, navigating multiple literacies is a common and complex task that requires adaptable language policies to ensure a better transition for children from one linguistic context to the next. When it comes to language policies, there is no one-size-fits-all solution.

With the COVID-19 pandemic still underway (as of this writing) and the emerging climate crisis becoming critical, the pressure on education systems will likely persist well into the future. Because literacy is fundamental for helping individuals and societies to navigate such crises, resilient education systems and practices are needed to ensure literacy acquisition can overcome these new challenges (Wagner et al., 2020).

As noted above, access to ICTs and remote learning is one promising example of how literacy learning can help to address such global problems. The shift to remote learning, which grew dramatically during the pandemic, has proven useful for ensuring the continuity of literacy learning, and led to increased access to ICTs among the poor in many countries. In the long run, governments, international organizations and civil society are finding new ways to leverage education technology innovations (Castillo et al., 2022; UNICEF, 2020b).

The importance and significance of literacy has been increasingly recognized over the years, even as it remains poorly acquired in many parts of the world today. Policy decisions about literacy development will become ever more relevant and more nuanced as societies evolve, and as our knowledge base about literacy continues to grow. The study and promotion of literacy will remain a central aspiration in educational planning and international development for many years to come.

References

- Angrist, N., de Barros, A., Bhula, R., Chakera, S., Cummskey, C., DeStefano, J., Floretta, J., Kaffenberger, M., Piper, B., Stern, J., 2021. Building back better to avert a learning catastrophe: estimating learning loss from COVID-19 school shutdowns in Africa and facilitating short-term and long-term learning recovery. *Int. J. Educ. Dev.* 84, 102397.
- Ball, J., Paris, S.G., Govinda, R., 2014. Literacy and numeracy skills among children in developing countries. In: Wagner, D.A. (Ed.), *Learning and Education in Developing Countries: Research and Policy for the Post-2015 UN Development Goals*. Palgrave Macmillan, New York, NY, pp. 26–41.
- Beeharry, G., 2021. The pathway to progress on SDG 4 requires the global education architecture to focus on foundational learning and to hold ourselves accountable for achieving it. *Int. J. Educ. Dev.* 82, 102375.
- Benson, C., Kosonen, K. (Eds.), 2021. *Language Issues in Comparative Education II: Policy and Practice in Multilingual Education Based on Non-dominant Languages*. Brill, Leiden.
- Benson, C., Wong, K.M., 2015. Development discourse on language of instruction and literacy: sound policy and Ubuntu or lip service? *Reconsidering Dev.* 4 (1), 1–16.
- Castillo, N., Wagner, D.A., 2019. Multilingual literacy transfer in rural South Africa: a technology-based impact study. *Int. Rev. Educ.* 65, 389–408.
- Castillo, N., Adam, T., Hassler, B., 2022. Improving the impact of educational technologies on learning within low-income contexts. In: Wagner, D.A., Castillo, N.M., Grant Lewis, S. (Eds.), *Learning, Marginalization, and Improving the Quality of Education in Low-Income Countries*. Open Book Publishers, Cambridge, UK.
- Collins, J., 1995. Literacy and literacies. *Annu. Rev. Anthropol.* 24 (1), 75–93.
- Commeys, M., Inyega, H.N., 2007. An integrative review of teaching reading in Kenyan primary schools. *Read. Res. Q.* 42 (2), 258–281.
- Ferguson, C.A., 1959. Diglossia. *Word* 15, 325–340.
- Gove, A., Cvelich, P., 2011. *Early Reading: Igniting Education for All*. Report by Early Grade Learning Community of Practice, revised ed. RTI, Washington.
- Hélot, C., Ó Laoire, M. (Eds.), 2011. *Language Policy for the Multilingual Classroom: Pedagogy of the Possible*. Multilingual Matters, Bristol, UK.
- Hornberger, N.H. (Ed.), 2003. *Continua of Bilinguality: An Ecological Framework for Educational Policy, Research, and Practice in Multilingual Settings*. Multilingual Matters, Bristol, UK.
- International Telecommunications Union, 2019. *Measuring Digital Development: Facts & Figures 2019*.
- OECD, 2019. *Education 21st Century Children*. OECD, Paris.
- Piper, B., Miksic, E., 2011. Mother tongue and reading: using early grade reading assessments to investigate language of instruction policy in East Africa. In: Gove, A., Wetterberg, A. (Eds.), *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy*. RTI, Research Triangle Park, NC, pp. 139–182.
- Save the Children, 2021. *Effect of Climate Change on Future Generations*. <https://www.savethechildren.org/us/what-we-do/emergency-response/climate-change>.
- Scribner, S., Cole, M., 2013. *The Psychology of Literacy*. Harvard University Press, Cambridge.
- Trudell, B., 2016. Language choice and education quality in Eastern and Southern Africa: a review. *Comp. Educ.* 52 (3), 281–293.
- UNESCO, 2014. *Reading in the Mobile Era: A Study of Mobile Reading in Developing Countries*. UNESCO, Paris.
- UNESCO, 2020. *International Literacy Day: Background Paper on “Youth and Adult Literacy in the Time of COVID-19: Impacts and Revelations”*. UNESCO, Paris.
- UNESCO, 2021. *International Literacy Day: Fact Sheet*. UNESCO, Paris.
- UNICEF, 2019. *A World Ready to Learn: Global Report on Pre-primary Education*. UNICEF, New York.
- UNICEF, 2020a. *How Many Children and Young People Have Internet Access at Home? Estimating Digital Connectivity During the COVID-19 Pandemic*. UNICEF, New York.
- UNICEF, 2020b. *School Mapping*. UNICEF. <https://www.unicef.org/innovation/school-mapping>.
- Valdiviezo, L.A., 2013. Vertical and horizontal approaches to ethnography of language policy in Peru. *Int. J. Sociol. Lang.* 2013 (219), 23–46.
- Vegas, E., 2020. *School Closures, Government Responses, and Learning Inequality Around the World During COVID-19*. Technical Report. Center for Global Development, Washington, DC.
- Venezky, R.L., Wagner, D.A., Ciliberti, B.S. (Eds.), 1990. *Toward Defining Literacy*. International Reading Association, Delaware.
- Wagner, D.A., 2011. What happened to literacy? Historical and conceptual perspectives on literacy in UNESCO. *Int. J. Educ. Dev.* 31, 319–323.
- Wagner, D.A., 2017. Children's reading in low-income countries. *Read. Teach.* 71 (2), 127–133.
- Wagner, D.A., 2018a. *Learning as Development: Rethinking International Education in a Changing World*. Routledge, New York.
- Wagner, D.A., 2018b. Literacy. In: Bornstein, M.H. (Ed.), *The SAGE Encyclopedia of Lifespan Human Development*. Sage Publishing, New York.
- Wagner, D.A., Maamouri, M., 2016. *Languages, Scripts and Reading Acquisition in Morocco: The Passerelle Approach for Improving Early Reading in Arabic*. ILI Technical Report. University of Pennsylvania, Philadelphia.
- Wagner, D.A., Spratt, J.E., Ezzaki, A., 1989. Does learning to read in a second language always put the child at a disadvantage? Some counterevidence from Morocco. *Appl. Psycholinguist.* 10 (1), 31–48.
- Wagner, D.A., Wolf, S., Boruch, R.F., 2018. *Learning at the Bottom of the Pyramid: Science, Measurement, and Policy in Low-Income Countries*. UNESCO, Paris.
- Wagner, D.A., Castillo, N.M., Grant Lewis, S., 2022. In: *Learning, Marginalization, and Improving the Quality of Education in Low-income Countries*. Open Book Publishers, London: Cambridge, UK. <https://books.openbookpublishers.com/10.11647/obp.0256.pdf>.

- Wagner, D.A., Castillo, N.M., Zahra, F.T., 2020. Global Learning Equity and Education: Looking Ahead: Background Paper for the Futures of Education Initiative. UNESCO Futures of Education Initiative, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000375000>.
- World Bank, 2021. Loud and Clear: Effective Language of Instruction Policies for Learning. World Bank, Washington, D.C.
- Wu, A.P., Gonsalves, S.L.R., Wagner, D.A., 2021. The impact of technology on a multilingual world: problems and opportunities. In: Lubin, I. (Ed.), *ICT and International Learning Ecologies: Representation and Sustainability Across Contexts*. Routledge, New York, pp. 42–60.

Regimes of literacy as regimes of truth about Africa: language ideologies and southern voices

Ashraf Abdelhay^a, Cristine Severo^b, and Sinfree Makoni^c, ^a Doha Institute for Graduate Studies, School of Social Sciences and Humanities, Doha, Qatar; ^b Department of Portuguese Language, Federal University of Santa Catarina, Campus Universitário Reitor João, Florianópolis, Brazil; and ^c Department of Applied Linguistics, Pennsylvania State University, College of the Liberal Arts, University Park, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	723
The African precolonial writing practices	724
Connecting colonial and contemporary regimes of truth with regimes of literacy	729
The tyranny of writing and the hegemonic ideology of alphabetic literacy	729
Orality and the testimonial narrative	730
Authorship: who is authorized to speak?	730
The African decolonial burden: the politics of citation and the southern voice	731
Concluding remarks	734
References	734

Introduction

In this chapter we discuss and problematize the colonial and contemporary connections between regimes of literacy—including ideologies of writing—and regimes of truth about Africa. This means considering the process of knowledge production, dissemination and reception as politically and economically controlled, contributing to a continuous remodeling of power relations. In the academic realm such power relations have historically characterized the relation North-South, being the North traditionally seen as the spoken subject and the South the silenced object: “the presumption is that people from the North offer through reason the ideas that make people of the South appear. Those from the South, thus, become passive recipients of such light. This, however, is a distortion of history and reality. It erases facts, and it cultivates dependency” (Gordon, 2021: 128). It overlooks the degree to which insights from the Global Souths have even during the period of imperialism been at the core of knowledge production of the Global North (Connell, 2019).

In this chapter we argue for a shift in a geography of reason (Gordon, 2021) that is able to expand the Western-Northern politics of truth that has characterized the production, validation and circulation of what counts as truth. We construe a “shift in a geography of reason” to be a conceptual framework which has multiple dimensions as we try to capture how knowledge about the Global Souths can be framed from a southern epistemological perspective and how and what the Global Norths look like from a perspective of the Souths. Following Makoni et al. (2022) we identify the following as the critical dimensions of a shift in a geography of reason:

1. Shifting from a closed to an open relational commitment
2. Thinking anew and creatively
3. Working on the basis of transindividual rather than isolated subject positions
4. Shifting the sites of the production of knowledge

By situating our arguments in African colonial contexts, we explore the role played by ideologies of literacy, mainly centered on the alphabetic script, in defining what counts as language—taken as fixed and named codes—, which consequently shaped the ideological beliefs about the nature of knowledge production and knowledge transmission which is carried out via extant codes. In African traditional contexts, if it is true that the elderly and ancestors are “the ones who give the first lessons in life, not only through experience but through the medium of stories, fables, legends, maxims, adages, and so on” (Hampaté Bâ, 1981: 179), we argue that such politics of knowledge transmission has been undermined by existing ideologies of literacy, schooling and truth-telling which are in conflict with long held views and practices of knowledge production and transmission. Our argument goes further and we argue that the shift in ideologies about language and literacy also impacted and changed what was regarded as knowledge and what knowledge was regarded as invalid or illegitimate.

This chapter is on literacy in Africa, colonial regime of literacy, regime of truth, authorship, standpoint, politics of knowledge production, oral tradition, politics of citation and editorial politics. The bigger proposition we are making goes beyond simply identifying insights in Africa; rather, we seek to understand how such insights affected the nature of literacy in the Global Norths as well. The critical point we are making is that we envisage “If African identity used to be constructed on the basis of European fictions, other narrative systems might construct it differently, the essential point being that in the future everyone can imagine and choose what makes him or her an African” (Mbembe, 2001: 17).

This chapter is structured as follows: first we briefly present a discussion of African precolonial writing practices; second, we analyze the connections between colonial/contemporary regimes of truth with regimes of literacy; finally, we address the politics of citation as part of the development of epistemologies of the South (Santos, 2014).

The African precolonial writing practices

We conceptualize literacy as a form of social practice tied to specific historical contexts in which it is used and valued. This means that literacy events should be studied “in relation to the larger sociocultural patterns which they may exemplify or reflect” (Heath, 1983: 74). Alphabetic literacy which is linked with formal education is just one instance, though the most powerful, of a situated action upon which the modern nation-state is built. Africa had its own list of literacy regimes and writing practices, some of which are among the oldest in the world (for a discussion see Nwosu, 2010; Juffermans et al., 2014). For instance, the list includes, but not restricted to, the forms of scripts associated with the Egyptian language, namely, the hieroglyphic, hieratic or Cursive, Coptic, and Demotic (Baines, 1983; Makar, n.d, see Fig. 1); Nsibidi which is a pictographic mode in Nigeria (Macgregor, 1909; Nwosu, 2010; Carlson, 2003, Ajibade et al., 2012, see Fig. 2); the Bamum syllabary in Cameroon (Kelly, 2018, Fig. 3); the Vai syllabary in Liberia (Dalby, 1967; Tuchscherer and Hair, 2002; Kelly, 2018, Fig. 4); the Kikakui script used by Mende of Sierra Leone (Tuchscherer, 1995, Fig. 5); the Ge’ez or Ethiopic script in Eritrea and Ethiopia (Asfaha et al., 2008, Fig. 7); Tifinagh in North Africa (Cline, 1953; El Aissati, 2014; Ali and Sedrati, 2016, Fig. 8); the Adinkra in Ghana (Temple, 2010, Fig. 9); the Meroitic script associated with the Nubian Kingdom of Kush in the South of Egypt around the 8th century BC (Griffith, 1913, 1916; Voogt, 2010; Nakanishi, 1980, Fig. 10).

Although most of the scripts above are old, other scripts are relatively recent. For example, Oberi Okaike is an indigenous (religious) script invented about 1930 by Akpan Udofia and Michael Ukpan and is used in the old Efik-Ibibio-speaking area in Nigeria (Akinnaso, 1996). Ismaaniya (Osmaniya) (or Somali script) was invented around 1920 by Osman Yusuf Kenadid, brother of Sultan Ali Yusuf Kandid in Somalia, and another non-Arabic script recently invented in Somalia is the Gadabuursi script by Sheikh Abdurahmaan Sheikh Nuur of the Gadabuursi clan (Lewis, 1958, Fig. 6). In addition to the native scripts listed above, there are also adapted scripts such as the Ajami which is a modified version of Arabic script used for writing in North Nigeria (Akinnaso, 1996).

It should be noted here that when we talk about African literacy traditions we are not just talking about other non-western forms of literacy and the local ways of “decoding” the scripts, but also, and most importantly, we are talking about different graphic modes of “reading the world” in Freire’s sense (2000). In other words, the focus is not exclusively on the linguistic function of these African

Letter	older Scripts			Coptic	
	Hiero-glyphic	Hieratic	Demotic	Cap	Small
Shai					
Fai					
Khai					
Hori					
Ganga					
Cheema					
Tee					

Fig. 1 Egyptian scripts: hieroglyphic, hieratic, demotic and coptic. Source: Maker (n.d).

Nsibidi	Names	Meaning
	Ekpe (Usime-Ekpe)	Indicating the Ekpe's strength and potential for anger and violence against anything when provoked.
	Uden	This is a harmless object of peace. It goes around pecking at food.
	Ubok	This outspread palm is a representation of the pioneers of the Ekpe society, as well as departed prominent members. It is for remembering important personages such as King James Eyo Honesty from Creek Town.

Fig. 2 Some Nsibidi Symbols from the Efik people of Calabar. Adapted from Ajibade et al. (2012: 302).

A	B	C	D	E	
					<u>yun</u> [.], les hôtes
					<u>tuə</u> [''], viens
					<u>yua</u> [''], que, dont
					<u>pe</u> [''], eux
					<u>tumə</u> [''], retire, extrais
					<u>kuə</u> [''], la toux
					<u>suən</u> [.], le caillou de pouzzolane
					<u>tuə</u> [''], penche
					<u>vua</u> [''], perdu, la perte
					<u>wux</u> [''], la mort
					<u>laam</u> [..], la forge

Fig. 3 Excerpted comparison of Bamum graphemes across five successive phases. Source: Dugast and Jeffreys (1950, 87), cited in Kelly (2018: 195).

writing signs but also on their symbolic sociolinguistic uses (Al-Kahtani, 1996; Suleiman, 2003). For instance, the Adinkra regime of literacy (in Ghana) is not a detached instrument of representing an independent world, it is rather “the rich cultural heritage of using pictorial images to convey the wisdom of Akan life” (Temple, 2010: 127). To take one example from the system, the symbol and meaning of “Sankofa” (going back to roots/the past) in Adinkra (see Fig. 9) is associated with or indexical of the Akan’s philosophy

᠋᠋᠋	a*	᠋᠋	ga	᠋᠋	la	᠋᠋	ra	᠋᠋	wa
᠋᠋᠋	é	᠋᠋	gé	᠋᠋	lé	᠋᠋	ré	᠋᠋	wé
᠋᠋᠋	ë	᠋᠋	gë	᠋᠋	lë	᠋᠋	rë	᠋᠋	wë
᠋᠋᠋	ɪ	᠋᠋	gɪ	᠋᠋	lɪ	᠋᠋	rɪ	᠋᠋	wɪ
᠋᠋᠋	ō	᠋᠋	gō	᠋᠋	lō	᠋᠋	rō	᠋᠋	wō
᠋᠋᠋	ö	᠋᠋	gö	᠋᠋	lö	᠋᠋	rö	᠋᠋	wö
᠋᠋᠋	ū	᠋᠋	gū	᠋᠋	lū	᠋᠋	rū	᠋᠋	wū
᠋᠋᠋	ba	᠋᠋	ha	᠋᠋	ma	᠋᠋	sa	᠋᠋	ya
᠋᠋᠋	bé	᠋᠋	hé	᠋᠋	mé	᠋᠋	sé	᠋᠋	yé
᠋᠋᠋	bë	᠋᠋	hë	᠋᠋	më	᠋᠋	së	᠋᠋	yë
᠋᠋᠋	bɪ	᠋᠋	hɪ	᠋᠋	mɪ	᠋᠋	sɪ	᠋᠋	yɪ
᠋᠋᠋	bō	᠋᠋	hō	᠋᠋	mō	᠋᠋	sō	᠋᠋	yō
᠋᠋᠋	bö	᠋᠋	hö	᠋᠋	mö	᠋᠋	sö	᠋᠋	yö
᠋᠋᠋	bū	᠋᠋	hū	᠋᠋	mū	᠋᠋	sū	᠋᠋	yū

Fig. 4 A phonetic chart of some Vai characters. Adapted from Massaquoi (1911: 459).

᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋
 ke nīlo tōgbo ngūm i su ye a ki ke
 ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋
 mo ga ngi wa ti gō si la bi na wo ngi ke nīhe
 ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋ ᠋᠋᠋
 mo gbo hu fa ngō ngō ngi ngi ti la ngi ti bu ti na Na

Fig. 5 Kikakui script. Source: Tuchscherer (1995: 177).

᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋
b	t	j	d	g	h	kh
᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋
h	d	r	s	sh	c	f
᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋
q	k	l	m	n	w	y
᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋	᠋᠋
a	i	u	o	e	aa	ii
᠋᠋	᠋᠋	᠋᠋				

Fig. 6 Somali script. Source: Lewis (1958: 141).

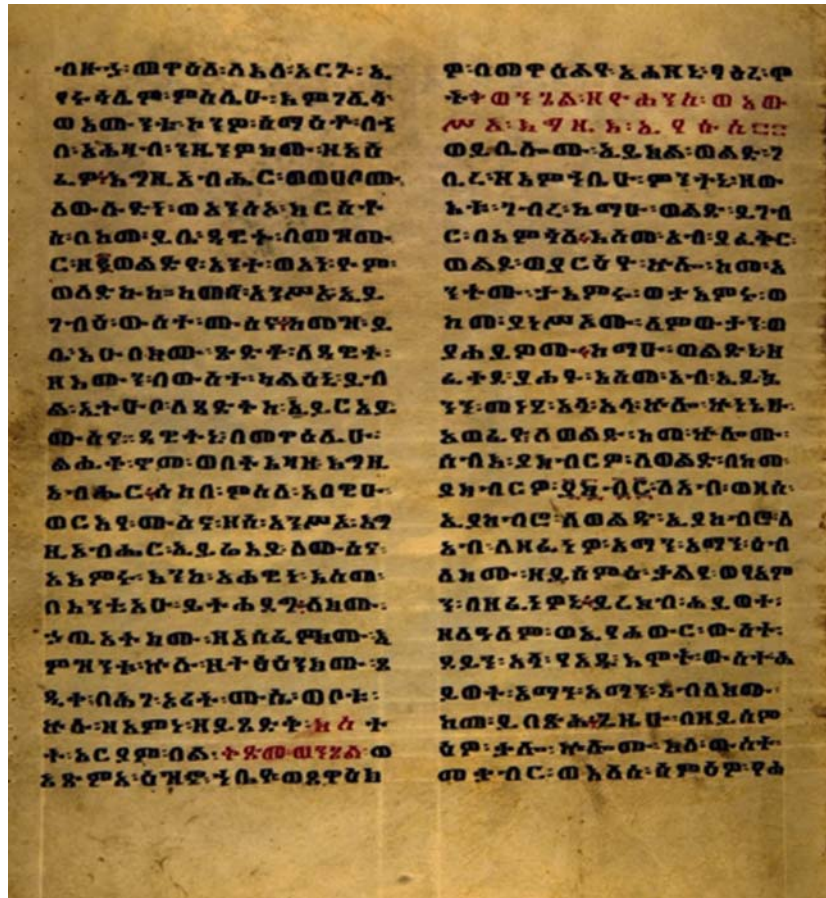


Fig. 7 A page of Ethiopic script from an Ethiopian manuscript containing. Source: DeGregorio Collection, Repository of the Bucknell University, <http://www.jstor.org/stable/community.2944750>.

o	θ	χ	χ ^u	Λ	E	◊	Η	Κ	Κ ^u	Θ
ya	yab	yag	yag ^w	yad	yad	yey	yaf	yak	yak ^w	yah
a	b	g	g ^w	d	ɖ	e	f	k	k ^w	h
[a]	[b/β]	[g/ɣ]	[g ^w]	[d/ð]	[ɖ]	[e]	[f]	[k/ç]	[k ^w]	[h]
Λ	Ϣ	χ	Ϫ	ς	I	H	Γ	I	◊	O
yah	yac	yax	yaq	yi	yaj	yal	yam	yan	yu	yar
h		x	q	i	j	l	m	n	u	r
[h]	[ʔ]	[x]	[q]	[i]	[j]	[l]	[m]	[n]	[u]	[r]
Q	Υ	Θ	Ø	C	†	E	U	ς	χ	ϣ
yar	yagh	yas	yaş	yac	yat	yat	yaw	yay	yaz	yaz
ʃ	gh	s	ş	c	t	ʈ	w	y	z	z
[ʃ]	[ɣ]	[s]	[s]	[ʃ]	[t/θ]	[ʈ]	[w]	[j]	[z]	[z]

Fig. 8 Neo-Tifinagh. Source: Ali and Sedrati (2016: 4).

Symbol	Name	Meaning
	Gye Nyame	Except God. A symbol expressing the omnipotence of God. Probably the most popular Adinkra symbol. It is featured on Ghana's largest-denomination banknote, the 200 cedi note.
	Sankofa	Go back and get it. A symbol of the wisdom of learning from the past to build for the future. From the Akan proverb, "Se wo were fi na wosan kofa a yenkyiri," meaning, "It is not taboo to go back for what you forgot (or left behind)."
	Sankofa	Another Sankofa symbol
	Adinkrahene	King of the Adinkra symbols. A symbol for authority, leadership, and charisma. Also a symbol for qualities associated with kings. Adinkrahene is reportedly the inspiration for the design of the other symbols.

Fig. 9 Adinkra symbols. Source: <https://adinkrasymbols.org/>.

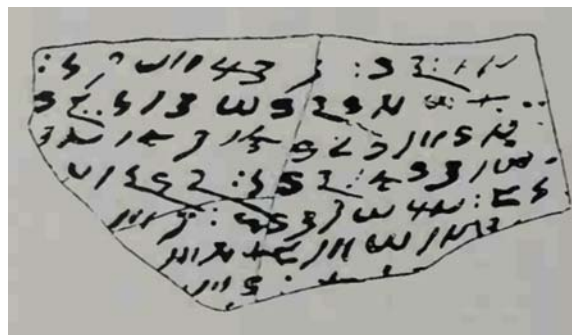


Fig. 10 The Meroitic writing of the Nubian kingdom in the south of Egypt about the 8th century BC. This form is one of two Meroitic scripts. Source: Nakanishi (1980: 104).

of life (on Sankofa's philosophy see Temple, 2010). It is thus a symbolic way of philosophizing which is historically constituted. Thus, it is by no means an "immobile" resource. Temple (2010) who studies Sankofa as a Diasporan practice in the USA, made the following note which deserves to be cited in full:

Sankofa practice is influenced by several orientations toward African consciousness: (a) as the legacy of natural cultural behaviors documented in its early usage by enslaved Africans who came to the Americas and in later usage, possibly, through epic memory; (b) as resistance with respect to rejecting Eurocentric language and world views and insisting on the relevance of using African conceptual possibilities to define and characterize African life in the contemporary era; and (c) as the symbolic gestures of Diasporan Africans interested in general forms of "returning to the source," or psychological steps toward Africanness. Beyond its usage by Africans, there is also the aspect of Sankofa appearing in non-African space due to cultural borrowing in an age of popular forms of diversity and multiculturalism.

Temple (2010: 128).

To further stress the issue of mobility, it is worth noting that some modern orthographies are developed from ancient African scripts with the purpose to teach currently used varieties of the ancient African languages to both native and non-native speakers not just in the "original" context of the language but also in the West (see Hashim, 2004). To illustrate, the Nubian Language Society (NLS) developed an Old Nubian-based phonemic orthography called "nobiin agii" for teaching Nobiin language to non-native speakers (Khalil, 2021). The NLS, among others, organizes Nobiin courses for non-native speakers in, among other contexts, Sudan, Egypt, and the USA. This and other orthographies are relatively mobile in the sense that they are largely developed and taught in the diaspora (see also Abdulmannan, 2015; for a detailed authoritative review of competing orthographies for writing Nubian languages see Hashim, 2004).

The above examples show that literacies in Africa cannot be conceptualized in purely instrumental terms. The symbolic dimension is always present. The instrumental function of literacy is generally associated with formal schooling or academic literacy, and it marks the shift from orally situated thinking to abstract, and literate thinking, and this shift is considered a defining feature of (western) modernity, and formal discourses associated with scientific disciplines (Street, 1995; Klein and Olson, 2001; Halliday and Martin, 1993). Meaning is viewed as embedded or contained within the word itself. Street (1984) terms this theoretical position the “autonomous model.” The perspective which supports this function is known in the literature as the “Great Divide” and is generally associated with the writings of McLuhan (1962), Goody and Watt (1968), Ong (1982), and Olson (1977) (for a review see Street, 1984; Klein and Olson, 2001). Alphabetic literacy has determined a significant portion of metalanguage which we use to convert language from a situated practice to an object of study; however, the cognitive implications drawn by the theorists of the Great Divide between orality and literacy are not universal. The use of these modes of communication is determined by their roles, goals, and learners’ needs in a society, though one mode may historically receive more emphasis than another (Al-Kahtani, 1996). For instance, although standard Arabic is largely associated with writing, it is conventionally used in specific communicative events as the default mode of expression (Ferguson, 1959). So, what determines the use of a highly codified medium is the context of situation and not the nature of the medium itself.

Connecting colonial and contemporary regimes of truth with regimes of literacy

In this section we aim at exploring the interconnection between the colonial regime of literacy and contemporary knowledge control mechanisms. The concept of regime of literacy (Kalantzis and Cope, 2015: 16) includes a political perspective that considers the “roles, rules, structures of governance and disciplinary logics” underlying the idea of literacy taken as a “regime of epistemic discipline, a moral economy.” We conjugate such idea with the discussion on how colonial regimes of truth (Foucault, 2003) about Africa contributed to invent an idea of Africa (Mudimbe, 1988). By regimes of truth we consider the types of discourse considered as truth to each society. In other terms, it is related to how each society forges specific ways of saying and validating its truth. Foucault (2003) identifies in Western societies five traits that characterize its regime of truth: (a) its relation with the scientific discourse; (b) its subjection to economic and political interests and forces; (c) its relation to a politics of diffusion and appropriation; (d) its relation with specific politics that regulates the way such discourses are distributed and institutionally validated; and (e) its relation with ideological struggles. In a nutshell, “truth is to be understood as a system of ordered procedures for the production, regulation, distribution, circulation and operation of statements” (Foucault, 2003: 132).

We argue that the regimes of truth about Africa were largely framed by specific discursive genres that also contributed to outline both the idea of what counts as language (Abdelhay et al., 2020; Gilmour, 2007; Makoni and Pennycook, 2006) and the discursive modes of expressing the truth about Africa (Severo and Makoni, 2020). Such discursive genres are based on the written word and its ideological role (Weth and Juffermans, 2018) that contributed to shape the ideas of language, knowledge transmission and truth. We argue that this scriptist logo-centric nature of knowledge has promoted particular types of knowledges and denigrated others. Examples of such ideologies based on the written word are missionary Linguistics and the schooling language ideology. We argue that there is a connection between concepts of language and the way the truth about Africa has been told and legitimated since colonial era until contemporary academic practices.

The tyranny of writing and the hegemonic ideology of alphabetic literacy

The introduction of alphabetic literacy in African contexts contributed to shape not only an ideology of language (Makoni and Meinhof, 2004), but to hierarchize languages and language practices in terms of the notions of prestige, citizenship, development and modes of telling the truth about Africa. We engage with the concept of “tyranny of writing” as a critical framework that contributes to elucidate the emergence and crystallization of “tensions and contradictions between the written and the spoken word, linguistic normativity, creativity and authenticity, and centers and peripheries in language practices” (Weth and Juffermans, 2018: 1). In this chapter we do not discuss the several concepts of literacies that have been under revision and expansion by the so-called New Literacy Studies (Street, 1984); rather, our objective is to elucidate how regimes of truth have contributed to reproduce language ideologies oriented by written scripts:

[the] ideology of literacy has significantly contributed to the construction of a particular “image” of Africa (imagined by this model of literacy as a bounded continent). By focusing on the cognitive (in)ability of a person to read and write, this ideological scheme of classification constructs what counts as “a normal person” (read: “modern”) in the process. The Eurocentric image of Africa is largely constructed through the variants of this technical discourse (i.e., a monolithic discourse on “literacy in Africa”).

Abdelhay et al. (2014: 2).

Such language perspective centered on ideologies of literacy has contributed not only to undermine what can be called as oral tradition in African contexts, but also to underestimate the role of African languages—and African discursive genres—in telling the truth about it. By inquiring about colonial and contemporary regimes of truth about Africa we draw attention to the role played by orality in shaping a sense of personhood (Wiredu, 2009: 13): “the influence of orality in philosophy need not be always negative. In

truth, sometimes an oral habit may be loaded with wisdom [...] this is so with regard to the African conception of a person." In this line of thought, we also consider that "Where writing does not exist, man is bound to the word he utters. He is committed by it. He is his word and his word bears witness to what he is" (Hampaté Bâ, 1981: 167). By problematizing how colonial and contemporary regimes of truth have been reinforced by specific regimes of literacy and vice versa, we spotlight the ethical and political dimension of such discussion.

Orality and the testimonial narrative

In the context of literary discursive regime, the testimonial narrative has been aesthetically recognized as a form of report of post-war survivors, whether in Europe, Latin America or Africa. As a discursive genre, "O testemunho, como sabemos, tem uma conotação política muito marcada que se traduz em permitir o uso da palavra por aqueles que tradicionalmente se encontram excluídos" ["The testimony, as we know, carries a marked political connotation by allowing the use of the word by those who are traditionally excluded"] (Reis, 2007: 79). We mention this as an example of discursive memory construction pervaded by the experience of historically invisible and silenced witnesses whose testimony has been transmitted across generations, helping to build a sense of truth. This means recognizing that "não se pode fazer uma história válida dos povos africanos sem a tradição oral" ["one cannot make a valid history of the African peoples without the oral tradition"] (Ki-Zerbo, 1969: 123). The testimonies have been used as central elements in the reconstruction of another history about Africa in which local voices work also as authors of such narratives. In this regard, Hrbek argues in favor of the reconstruction of South African history (1981: 124) from the perspective of "African voices":

the struggle now going on in South Africa in every field of human activity necessitates also a new approach to the sources; particular attention must be paid to the written evidence of the hard struggle of the Africans for their rights. Only research based on all the evidence and material will allow the writing of a truthful history of South Africa.

By discussing the role of oral tradition, testimony and personhood in the construction of narratives about what counts as Africa we also intend to problematize the Western and Northern regime of knowledge production. Such regimes of truth have been allied to regimes of modes of telling-writing which contributed to largely dictate the rules that determine not only the production of knowledge and its economic and political value, but also the language ideologies that frame the way we (Southern scholars) are supposed to tell-write the truth about Africa. We claim that contemporary scholarship has contributed to perpetuate colonial ideologies of literacy that overlaps ideas of knowledge, knowledge transmission, language and truth mainly centered not only on alphabetic literacy, but also on academic discursive genres broadly regulated by the academic publishing industry.

Authorship: who is authorized to speak?

In this sense, the concept of authorship regarding who is legitimated to tell-write about Africa also matters. Taking into account an ethical approach to this issue, we consider two interconnected elements: the individual and the collective dimensions of authorship, both referring to a given discursive function (Foucault, 1977) or standpoint (Hill, 2000). As a function of discourse, authorship refers to discursive memories and interpretive frameworks; it is concerned with how discourses are controlled, selected, organized and classified, which is strongly influenced by the material and political conditions underlining such discourses: "the function of an author is to characterize the existence, circulation, and operation of certain discourses within a society" (Foucault, 1977: 124). Similarly, the standpoint cannot be reduced to an individual's perspective, but is signals to specific modes of discursive functioning that is strongly connected to collective modes of living and interpreting the world: "Black women intellectuals have laid a vital analytical foundation for a distinctive standpoint on self, community, and society and, in doing so, created a multifaceted, African-American women's intellectual tradition" (Hill, 2000: 2–3).

The academic regime of knowledge reinforces a given authorship functioning based on politics of citation and editorial positions that tend to prioritize certain proper names—generally white men from the Global North—that work within neoliberal statistics of citation in which white men tend to cite other white men, research groups favor their peers, and the references tend to reproduce the same Eurocentric cadre of scholars (B. Makoni, 2021; Mott and Cockayne, 2007). Also noteworthy is the priority given to the English-speaking academic community, or to the translation and proofreading industry for English texts. For example, more than 90% of the information available in international databases, such as the Science Citation Index is in English, which attests that the language policy that guides the scientific and academic truth regime is based on use of the English language, on rhetoric and Anglo-style and on symbolic representations about language (Kirkpatrick, 2009). This fact matters when we think about the importance of the authorship from the Global South in the system of production and dissemination of knowledge to foster anti-racist, anti-sexist and geopolitically plural academic practices. In other words, attention to the citation interferes in the economy and geopolitics of knowledge itself:

Careful and conscientious citation is important because the choices we make about whom to cite—and who is then left out of the conversation—directly impact the cultivation of a rich and diverse discipline, and the reproduction of geographical knowledge itself.

Mott and Cockayne (2007: 955).

By connecting the discussion on regimes of truth and regimes of literacy, we intended to problematize the political framework that has defined and hierarchized concepts of language, ascribing to alphabetic writing and academic written genres not only social prestige, but also the power to tell-write the truth about the South and Southern people. We will problematize such aspect in the next section.

The African decolonial burden: the politics of citation and the southern voice

On the June 15, 2021 we submitted a manuscript entitled “The Discursive Politics of the Sudanese Revolution of December 19, 2018” for consideration for publication in the *Journal of Multilingual and Multicultural Development* (JMMD). On June 17th (after two days from the submission date), we received the following email from the General Editor of the JMMD, Professor Jean-Marc Dewaele:

Dear Dr. Abdelhay,

I regret to inform you that we have now considered your paper, but unfortunately feel it unsuitable for publication in Journal of Multilingual and Multicultural Development. You are now free to submit your manuscript to another journal.

Thank you for considering Journal of Multilingual and Multicultural Development. I hope the outcome of this specific submission will not discourage you from the submission of future manuscripts.

Yours sincerely,
Professor Jean-Marc Dewaele.
Editor.

Upon the receipt of this email, we responded with the following email (dated June 20, 2021):

Dear Professor Jean-Marc Dewaele,

Thank you for your email. We are disappointed with this editorial decision for two reasons, first because we are not told precisely the reason for not considering the paper for reviewing. It could have been very helpful had you told us about the reason. Second, the document of “Author’s Guidelines and General Information” which you attached also left us in the dark about the ground for your editorial decision. As we see it, our paper ticks three boxes: ethnicity and nationalism (the paper is about the December Revolution in Sudan, which it mainly addresses the issue of agency and gender in context of conflict); identity politics (the paper is framed from within the cultural politics to consider how language and identity in the context of the revolution are sites of conflict and struggle in Sudan); popular and folk sociolinguistics and usage (one of the fundamental goals of the paper is to critique the eurocentric discursive construction of the revolutions, and that is why we have used cultural politics as a framework to explore folk sociolinguistic constructions of the ‘same’ event); collective identity and its markers (to show markers of the different voices or identity we performed micro-analysis not as an isolated activity but as a way of addressing the wider issues referred to above). So we wonder how concretely our paper is “unsuitable for publication” in this journal if it addresses these required conceptual boxes?

This is not our first submission to the journal on a cultural politics of language in Sudan, we have published at least two papers on Sudan in this journal dealing with the politics of language, and this manuscript is part of our wider project which deals with the cultural politics in Sudan and other contexts. We are mentioning this point as a way of saying we cannot see a reason in the attached document for your editorial decision nor any ground for your statement that the paper is “unsuitable for publication.”

Since the Guidelines advise that if authors are unhappy about an editorial decision they can contact Professor John Edwards, we have cited him in this email.

Yours,
Ashraf Abdelhay (for Cristine Severo and Sinfree Makoni).

The Editor Jean-Marc Dewaele returned on 20 June with the following response:

Dear Ashraf (and co-authors),

I understand your disappointment about my decision not to send your paper out for review. The fact that your paper ticked the boxes did not automatically qualify it for being sent for external review. I noticed that the authors are respected scholars, and that the paper seemed well constructed. Also, I love contributions about Africa. However, I judged that this paper was unlikely to attract a high volume of future citations and my job as editor of JMMD is to boost its impact factor. You may disagree with me, and I fully understand, but in my 6 years as editor of IJBEB, I managed to double its impact factor, and I intend to do the same for JMMD.

I’m surprised John Edwards is still mentioned in the guidelines, as he is now Reviews Editor rather than General Editor, but I welcome his opinion in this matter!

Regards,
Jean-Marc Dewaele.
Editor Journal of Multilingual and Multicultural Development.

We responded to the email of Professor Jean-Marc Dewaele on 21 June with the following email:

Dear Professor Jean-Marc Dewaele,

We thank you for your transparency. We should let you know upfront that the linguistics we operate with, whether at the level of theory or practice, is precisely a negation of your editorial linguistic ideology. You are right that you might have succeeded in boosting the “impact factor” of the

International Journal of Bilingual Education and Bilingualism (IJBEB) and you might achieve the same “success” for the JMMD, but in the process of boosting the journals’ statistics and your editorial CV, you have effectively excluded and silenced “less known” scholarship from the Global South including the “voices” of the masses in the Arab world and Africa. If the default definition of sociolinguistics is the study of language as it is used by people on the ground (whether in Liverpool or Dongla in Sudan), the linguistics you are promoting through your editorial practice is Eurocentric to the core because it “erases” from view less known forms of (folk) linguistics outside Europe, let alone works which can be classified as “anti-linguistics.” It should have occurred to you that this understanding of the “impact factor” which leads an editor of a multilingual and multicultural journal to pre-determine whose voice should be heard is itself an effect of ideological practices such as yours. And the result is a corpus of monocultural and monolingual voices constructed through epistemological spaces paradoxically intended for the recognition and development of multiculturalism and multilingualism. This is the goal of JMMD as we understand it. Its goal is not to enhance its statistics of the impact factor at the cost of excluding “less known” works about less known sociolinguistic situations and topics. Otherwise, the journal should indicate this condition in the Guidelines which you shared with us, i.e., it should state that “potentially less citable work” will be deemed by the editor as “unsuitable for publication.” The process of peer reviewing which is the principle of validating knowledges in journals has now at your editorial hand become redundant for works which you believe (as part of your ideology) is (potentially) less cited. What is striking here is that you have even determined the future trajectory of “less known works” if they are published by the journal. And that is why with “confidence” you disqualify them from the very beginning as “not suitable for publication”! This metaphysical ideology is exclusionist because it effectively silences “less known” voices from the Global South using a politics of citationality to rationalize this symbolic violence. And worse, it reflects a particular (mis)understanding of knowledge as a quantifiably static “thing” whose trajectory can linearly be predetermined and known even before it is circulated or perceived. The question of “how” (how it will be perceived, and the effect it may produce on existing literature, etc.) is effectively erased. “Less known” works are not born “less known”; they are constructed as such through specific discursive planning practices including editorial policies and practices.

We know very well that publishing is a field of power (in Bourdieu’s sense). This is understandable, but it becomes an unfair game when a journal is gated by pre-judgmental editors against “less known” works and knowledges from Africa and the Middle East. However, this process becomes fair when editors grant “less known” works or contributions which deal with less “popular” topics, which tick the boxes, the opportunity to be heard (to be reviewed); to be treated objectively equally with other “works” from the Centers. We experienced this fair editorial practice by Professor John Edwards.

We will of course submit our paper to another journal, and we will make sure you are not its editor or “co-editor” simply because we can now expect the fate of our work or the works of colleagues from the Global South if you are the editor. Parallel to this process, we will continue our epistemological and “ethical” struggle against exclusionist practices in academic discourses and institutions, and we intend to take the case to the publishing house (Taylor and Francis). Our objective for pursuing the case further is, as we said, ethically motivated: If such an editorial practice is normalized as a policy, it will certainly have adverse consequences on scholars and scholarship from the Global South and other “less known” sociolinguistic contexts.

Yours,

Dr. Ashraf Abdelhay (for Dr. Cristine Severo and Prof Sinfree Makoni).

As an action-oriented step to change this editorial ideology, we filed a formal complaint to the publishing house (Taylor and Francis), and we cited the exchanges above in the light of which we contended that:

In the light of the above editorial justification of Professor Jean-Marc Dewael, we evidently believe that his editorial practice is discriminatory and exclusionist of scholarship from unknown sociolinguistic situations from Africa and the Middle East. The Editor admits that the paper is well constructed and written by established scholars and that its content ticks the boxes indicated in the Guide of the journal; however, he rejects to subject it to reviewing because he believes that it is not potentially citable. This justification is not “scientific” and we have never heard of it as a condition for publishing in a journal. As the Philosopher Karl Popper noted, scholarship proceeds by falsification from the margin, i.e., unknown cases can change the intellectual landscape. Most importantly, our paper is about the Sudanese December Revolution of 2018, which is a very recent development, and having it excluded from consideration for publication because it is relatively unknown context means excluding the contribution of local material-discursive situations to our understanding of mainstream concepts and theories which we use to understand those dynamics. Thus, we have an objective ground to believe that we have been unfairly discriminated against by Professor Jean-Marc Dewael because the work is from Africa.

We kindly ask that our complaint against this exclusionist and discriminatory editorial practice be given due consideration otherwise scholarship about new or unknown sociolinguistic conditions from Africa and the Middle East will be systematically marginalized and excluded by the journal.

Citation is the textual act of referring to another work as a way of recognizing the contribution of others in the process of knowledge production. And the impact factor is calculated as “the number of citations to papers published in a particular journal over a 2-year period divided by the total number of papers published by the journal in those 2 years,” and the purpose is that “this metric provides a standard through which librarians can compare journals to help in making decisions for purchasing subscriptions” (Marks et al., 2013: 611). It was never intended as a measure of the scientific quality of a research paper in a journal.¹ With this very precise scope in background, we can understand the “misuse of journal impact factors in scientific assessment” (to employ Marks et al., 2013 title). The above tabulated transcript is a real case of misuse of the impact factor by an editor of an established journal.

Our aim for citing the above institutional interaction with the editor of the JMMD is to show with concrete example how the issue of the “citation rates” in academic journals, which is the basis of calculating the “impact factor,” can be manipulated as a proxy for exercising epistemological racism against scholarship and scholars from the Global South. It is a way of raising the question of how our thinking about language and society has become so colonized by conditions such as the “citation rate” when it is pursued by journals as an end in itself, and the impact of this style of thinking on scholarly debates and on the limits of what Gordon (2021) terms the “geography of reason.”

Citation is a key metalinguistic device which forges a sense of collective construction and shared ownership of knowledge. However, it is simultaneously implicated in the cultural politics of stratification particularly when it is (ab)used in the academic discourse as a gatekeeping tool of excluding non-mainstream scholarship or works by writers of color or LGBTQ which deviate

¹San Francisco Declaration on Research Assessment, <https://sfidora.org/read/>.

from the established frame of conceptual expectations (Scharrón-Del Río, 2020). Marks et al. (2013), the editors of the *Journal of Traffic*, made the following significant observation:

Many investigators use the JIF [journal impact factor] in deciding where to submit their work for publication, opting to submit to “high impact” journals over journals, which though more appropriate, have a lower JIF value. More insidiously, many academics, employers, funding agencies and others, across the world have come to use the JIF of journals in which individuals publish to make decisions about hiring, academic promotion, funding, etc. Rather than judging the value of a paper based on its significance and the quality of the work presented, or judging our peers based on the aggregate of their individual merits, we judge them based on the average number of citations received to papers published in journals in which they publish. Publishers and editors of the journals themselves have taken note of this trend, and anxiously await the release of the journal JIF every June; many journals “game” the JIF by opting to favor publishing article types that tend to accumulate citations or by classifying some articles that will be cited by others.

Marks et al. (2013: 611).

Similarly, in the context of the Arab world, Amin (2021) noted that the main objective of research is no longer the production of scientific knowledge and the development of society through participation in policy making, but rather publishing in journals with a high impact factor. He added that researchers from the Arab world were consequently discouraged from working on issues related to their socio-political situations because “those in charge of these journals will most likely reject articles that will not achieve enough citations, thus weakening the impact factor of the journal” (Amin, 2021: 99). They were also enforced to abandon discursive methodologies, since high impact-factor journals favor positivist ones.

The editorial perspective in the example which we cited above shows how a particular editorial regime of textual rationality can render southern scholarship invisible by denying publishing works related to or published by scholars from the Global South. The discursive mechanism for this editorial act of erasure is the ideology of what counts as “a potentially citable work.” This institutionally supported Eurocentric and Orientalizing gaze becomes a grid for deciding what is “potentially citable” because it operates from within a monoglossic ideology of knowledge. Crudely, in some academic journals, the “citation rate” is both an effect and a proxy for peripheralizing the southern voice. This cultural politics of citation has serious implications and ramifications for scholars from the Global South including black female scholars (B. Makoni, 2021). For example, it disqualifies other historically situated epistemologies simply because they, as the above editorial position believes, cannot boost the impact factor of the journal and consequently the “CV” of the editor. In a paper entitled “bibliometrics as weapons of mass citation,” Molinié and Bodenhausen argued:

The allocation of resources for research is increasingly based on so-called “bibliometrics.” Scientists are now deemed to be successful on the sole condition that their work be abundantly cited. This world-wide trend appears to enjoy support not only by granting agencies (whose task is obviously simplified by extensive recourse to bibliometrics), but also by the scientists themselves (who seem to enjoy their status of celebrities) ... the overwhelming fashion of bibliometrics is largely due to the narcissistic mind of the scientists themselves. Their fascination for citation indexes often drives them beyond reason. Their obsession with their egos is much more perverse than the laziness of granting agencies.

Molinié and Bodenhausen (2010: 78).

We argue that although the above communicative process was with an individual scholar invested with editorial power, the encounter reveals structural pattern about what “counts as legitimate knowledge” within the hegemonic western regime of literacy and textuality. In other words, although the epistemic position of the editor is not shared by all western sociolinguistic journals and not even by members in its editorial board (who are normally denied voice), it is still systemic and hegemonic (for a discussion of the technical defects of the impact factor see Ramsden, 2009; Kiesslich et al., 2021). Afful and Janks (2013: 193) rightly argued that there “can be no doubt that citation in the academy is a politicized practice. In acquiring advanced academic literacy, students have to master the art of positioning themselves in relation to the work of others, so that they develop their own ‘scholarly identity’”. Thus, citation practices enact “in-built hierarchy of privileged academic sources,” and they define what counts as “knowledge” in a particular community of practice (Afful and Janks, 2013: 196). Reflecting on the systemic perpetuation of racism and sexism in the process of knowledge construction, B. Makoni (2021) argued that

gendered racial injustice is seen in publication rates, citation rates, and appointments to editorial boards. The underrepresentation of women in general, and black women in particular, discursively constructs scholarly enquiry as normatively white and masculine. The exclusion of nonwhite scholars partially emanates from a reliance on Euro-American theoretical frameworks that are applied, often uncritically, in other contexts as if the Euro-American experience were universal. The result is scholarship that is incongruous with local experiences and practices.

B. Makoni (2021: 48).

The result of processes of editorial erasure of the southern voice is an institutionalized construction of an “apartheid of knowledge in the academia” (Bernal and Villalpando, 2002). To drive the discussion home, journals which exclude works from/about the

Global South using the “citation rates” as an editorial rationalization, run the risk, in the words of B. Makoni (2021: 50), of “reproducing the power relations it critiques if the racialized relations that impact the global flows of power and knowledge remain unquestioned.”

It is worth stressing that some journals and editors have rejected this misuse of the JIF. At the 2012 meeting of the *American Society for Cell Biology* (ASCB), some editors of bioscience journals discussed the misuse of impact factor and suggested a number of recommendations to stem this misuse. The set of recommendations and the concerns became known as the “San Francisco Declaration on Research Assessment” (<https://sfdora.org/read/accessible> March 23, 2022). The “General Recommendation” made by the declaration reads: “Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions”.

Some sociolinguistics-oriented journals and editors also seriously started to widen its geography of reason by diversifying its editorial board and actively supporting and encouraging research from the Global South, and giving visibility to non-western languages. Ernst (2010: 90) suggested the creed of “discrediting ‘number games’ as incompatible with our goals of objectivity, credibility, fairness, and social responsibility, as researchers,” and called for establishing “on the Internet, a generally accessible Webpage to list agencies, journals, and individuals who regularly use and misuse bibliometric measures in their judgements.” He further argued that: It is only by this kind of active resistance to the follies of bibliometrics that our scientific self-respect and credibility can be saved. We should liberate our minds again to enable true creativity in view of long-term social benefits: “We certainly do not want to convert our precious universities into bureaucratic training centers for mindless citation hunters! Our institutions shall remain forever unbiased resorts of limitless human dignity and foresight” (Ernst, 2010: 90).

Concluding remarks

This chapter makes a case for the interconnection between the concept of authorship, the politics of knowledge production—which includes the politics of citation and impact factor—with the discussion on (colonial) literacy in African contexts. This means that the rules that dictate the discursive production of truth in Western societies are deeply inscribed into specific ideologies of literacy. From a historical and critical perspective we problematize such regimes of truth (Foucault, 2003) and “geography of reason” (Gordon, 2021) that have framed our understandings of language, writing and truth-telling in the academic field. Finally, concerning the ideologies of literacy and the politics of knowledge production on Africa, we follow Gordon’s (2021: 69) challenges regarding a shift in the geography of reason:

Shifting the geography of reason requires understanding that power should not be reduced to a single element but instead should be explored to its creative potential. If we acknowledge power as the ability to make things happen and securing access to the conditions of doing so, there is much proverbially to be done.

References

- Abdelhay, A., Asfaha, Y., Juffermans, K., 2014. African literacies: ideologies, scripts, education. In: Juffermans, K., Asfaha, Y., Abdelhay, A. (Eds.), *African Literacies: Ideologies, Scripts, Education*. Cambridge Scholars Publishing, Cambridge, pp. 1–62.
- Abdelhay, A., Makoni, S., Severo, C. (Eds.), 2020. *Language Planning and Policy: Ideologies, Ethnicities and Semiotic Spaces of Power*. Cambridge Scholars Publishing, Newcastle-Upon-Tyne.
- Abdulmannan, N., 2015. Nubian (Nobiin) Language and Grammar Book 1 (E-book). Maryland, USA, revised ed.
- Aful, J., Janks, H., 2013. The politics of citation: an analysis of doctoral theses across disciplines. *Critical Approaches to Discourse Analysis across Disciplines* 6 (2), 193–210.
- Ajibade, B., Ekpe, E., Bassey, T., 2012. More than fabric motifs: changed meaning of Nsibidi on the Efik ukara cloth. *Mediterr. J. Soc. Sci.* 3 (2), 297–303.
- Akinnaso, F.N., 1996. Vernacular literacy in modern Nigeria. *Int. J. Sociol. Lang.* 119, 43–68.
- Ali, A.A., Sedrati, A., 2016. Towards Internet domain names in Tifinagh. In: *International Conference on Information and Communication: Technologies for the Amazigh*, pp. 1–12. Rabat, Morocco.
- Al-Kahtani, A., 1996. Literacy from a linguistic and a sociolinguistic perspective. *Int. Rev. Educ.* 42 (6), 547–562.
- Amin, A., 2021. Social psychology in the Arab world and the dialectic between the local and the global [in Arabic]. *Idafat J.* 51/52, 83–107.
- Asfaha, Y.M., Kurvers, J., Kroon, S., 2008. Literacy and script attitudes in multilingual Eritrea. *J. Socioling.* 12 (2), 223–240.
- Baines, J., 1983. Literacy and ancient Egyptian society. *Man* 18 (3), 572–599.
- Bernal, D.D., Villalpando, O., 2002. An apartheid of knowledge in academia: the struggle over the “legitimate” knowledge of faculty of color. *Equity Excell. Educ.* 35 (2), 169–180.
- Carlson, A., 2003. Nsibiri, Gender, and Literacy: The Art of the Bakor-Ejagham (Cross River State, Nigeria). PhD Thesis. Indiana University, USA.
- Cline, W., 1953. Berber dialects and Berber. *SW. J. Anthropol.* 9 (3), 268–276.
- Connell, R., 2019. *The Good University: What Universities Actually Do and Why It’s Time for Radical Change*. Monash University Publishing Press.
- Dalby, D., 1967. A survey of the indigenous scripts of Liberia and Sierra Leone: Vai, Mende, Loma, Kpelle and Bassa. *Afr. Lang. Stud.* 8, 1–51.
- Dugast, I., Jeffreys, M., 1950. *L’Écriture des Bamum: sa naissance, son évolution, sa valeur phonétique, son utilisation*. Institut français d’Afrique Noire, Douala.
- El Aissati, A., 2014. Script choice and power struggle in Morocco. In: Juffermans, K., Asfaha, Y., Abdelhay, A. (Eds.), *African Literacies: Ideologies, Scripts, Education*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 147–177.
- Ernst, R., 2010. The follies of citation indices and academic ranking lists a brief commentary to “bibliometrics as weapons of mass citation”. *Chimia* 64 (1/2), 90.
- Ferguson, C., 1959. Diglossia. *Word* 15 (2), 325–340.
- Foucault, M., 1977. What is an author? In: Bouchard, D. (Ed.), *Language, Counter-Memory, Practice: Selected Essays and Interviews*. Cornell University Press, Ithaca, NY, pp. 113–138 (trans. D. Bouchard, and S. Simon).
- Foucault, M., 2003. In: Rabinow, P., Rose, N. (Eds.), *Essential Works of Foucault 1954–1989*. The New Press, New York.

- Freire, P., 2000, 1970. In: Bergman Ramos, M. (Ed.), *The Pedagogy of the Oppressed*. Continuum International Publishing Group.
- Gilmour, R., 2007. Missionaries, colonialism and language in nineteenth-century South Africa. *Hist. Compass* 5 (6), 1761–1777. <https://doi.org/10.1111/j.1478-0542.2007.00472.x>.
- Goody, J., Watt, I., 1968. The consequences of literacy. In: Goody, J., Watt, I. (Eds.), *Literacy in Traditional Societies*. Cambridge University Press, Cambridge, MA, pp. 27–68.
- Gordon, L., 2021. *Freedom, Justice, and Decolonization*. Routledge, New York.
- Griffith, F. (Ed.), 1913. *The Nubian Texts of the Christian Period*. Verlag der Königl. Akademie der Wissenschaften : Akademie der Wissenschaften, in Kommission bei G. Reimer, Berlin.
- Griffith, F., 1916. Meroitic studies. *J. Egypt. Archaeol.* 3 (1), 22–30.
- Halliday, M.A.K., Martin, J.R., 1993. *Writing Science: Literacy and Discursive Power*. The Falmer Press/Taylor & Francis Group, London.
- Hampaté Bâ, A., 1981. The living tradition. In: Ki-Zerbo, J. (Ed.), *General History of Africa I: Methodology and African Prehistory*. UNESCO, University of California Press, Berkeley, pp. 166–205.
- Hashim, M., 2004. Competing orthographies for writing Nobiin Nubian. In: Gilley, L. (Ed.), *Occasional Papers in the Study of Sudanese Languages*, vol. 9. SIL-Sudan, pp. 215–248.
- Heath, S., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. McGraw-Hill, New York.
- Hill, P.C., 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge, Abingdon.
- Hrbek, I., 1981. Written sources from the fifteenth century onwards. In: Ki-Zerbo, J. (Ed.), *General History of Africa I: Methodology and African PreHistory*. UNESCO, University of California Press, Berkeley, pp. 114–141.
- Juffermans, K., Mesfun, Y., Abdelhay, A. (Eds.), 2014. *African Literacies: Ideologies, Scripts, Education*. Cambridge Scholars Publishing, Newcastle upon Tyne.
- Kalantzis, M., Cope, B., 2015. Regimes of literacy. In: Heydon, R., Hibbert, K., Stooke, R. (Eds.), *Negotiating Spaces for Literacy Learning: Multimodality and Governmentality*. Bloomsbury Academic, pp. 15–24. <https://doi.org/10.5040/9781474257138.ch-002>.
- Kelly, P., 2018. The invention, transmission and evolution of writing: insights from the new scripts of West Africa. In: Ferrara, S., Valério, M. (Eds.), *Paths Into Script Formation in the Ancient Mediterranean*. Edizioni Quasar, Rome, pp. 189–209.
- Khalil, N., 2021. In: *Learn Nubiin: Book (1): Beginner's level (e-book)*. (n.p.).
- Kiesslich, T., Beyreis, M., Zimmermann, G., Traeger, A., 2021. Citation inequality and the journal impact factor: median, mean, (does it) matter? *Scientometrics* 126, 1249–1269.
- Kirkpatrick, A., 2009. English as the international language of scholarship: implications for the dissemination of “local” knowledge. In: Sharifian, F. (Ed.), *English as an International Language: Perspectives and Pedagogical Issues*. Multilingual Matter, Bristol, pp. 254–270.
- Ki-Zerbo, J., 1969. The oral tradition as a source of African history. *Diogenes* 17 (67), 110–124.
- Klein, P., Olson, D., 2001. Texts, technology, and thinking: lessons from the Great Divide. *Lang. Arts* 78 (3), 227–236.
- Lewis, I.M., 1958. The Gadabuursi Somali script. *Bull. Sch. Orient Afr. Stud.* 21 (1/3), 134–156.
- Macgregor, J.K., 1909. Some notes on Nsibidi. *J. R. Inst. Afr.* 39, 209–219.
- Makar, K.F., n.d. *The Coptic Language: Introduction*. St Mary's Hermitage Press.
- Makoni, S., Meinhof, U., 2004. Western perspectives in applied linguistics in Africa. *AILA Rev.* 17, 77–104. <https://doi.org/10.1075/aila.17.09mak>.
- Makoni, S., Pennycook, A. (Eds.), 2006. *Disinventing and Reconstituting Languages*. Multilingual Matters, Clevedon.
- Makoni, S., Kaaper-Marquez, A., Mokwena, L., 2022. In: *Language in the Global South*. Routledge, New York.
- Makoni, B., 2021. Black female scholarship matters: a black female's reflections on language and sexuality studies. *J. Lang. Sex.* 10 (1), 48–58.
- Marks, M., Marsh, M., Schroer, T., Stevens, T., 2013. Misuse of journal impact factors in scientific assessment. *Traffic* 14, 611–612.
- Massaquoi, M., 1911. The Vai people and their syllabic writing. *J. R. Afr. Soc.* 10 (40), 459–466.
- Mbembe, A., 2001. African modes of self-writing. *Ident. Cult. Polit.* 2 (1), 1–39.
- McLuhan, M., 1962. *The Gutenberg Galaxy*. The University of Toronto, Toronto, ON.
- Molinié, A., Bodenhausen, G., 2010. Bibliometrics as weapons of mass citation. *Chimia* 64, 78–89.
- Mott, C., Cockayne, D., 2007. Citation matters: mobilizing the politics of citation toward a practice of “conscientious engagement”. *Gend. Place Cult.* 24 (7), 954–973. <https://doi.org/10.1080/0966369X.2017.1339022>.
- Mudimbe, V.Y., 1988. *Invention of Africa: Gnosis, Philosophy, and the Order of Knowledge*. Indiana University Press, Bloomington.
- Nakanishi, A., 1980. *The Writing Systems of the World*. Charles E. Tuttle Company, Tokyo.
- Nwosu, M., 2010. In the name of the sign: the Nsibidi script as the language and literature of the crossroads. *Semiotica* 182, 285–303.
- Olson, D.R., 1977. From utterance to text: the bias of language in speech and writing. *Harv. Educ. Rev.* 47, 257–281.
- Ong, W.J., 1982. *Orality and Literacy: The Technologizing of the Word*. Met, London.
- Ramsden, J., 2009. Impact factors—a critique. *J. Biol. Phys. Chem.* 9, 139–140.
- Reis, L., 2007. Testemunho como Construção da Memória. *Cadernos de Letras da UFF* 33, 77–86.
- Santos, B.S., 2014. *Epistemologies of the South: Justice Against Epistemicide*. Routledge, New York.
- Scharrón-Del Río, M., 2020. Intersectionality is not a choice: reflections of a queer scholar of color on teaching, writing, and belonging in LGBTQ studies and academia. *J. Homosex.* 67 (3), 294–304.
- Severo, C., Makoni, S., 2020. African oral tradition, parrhesia and integrationism. In: Makoni, S., Bade, D. (Eds.), *Linguistic Ideologies, Sociolinguistic Myths and Discourse Strategies in Africa*. The International Association for the Integrational Study of Language and Communication, Hong Kong, pp. 13–34.
- Street, B., 1984. *Literacy in Theory and Practice*. Cambridge University Press, Cambridge.
- Street, B., 1995. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography and Education*. Longman, London.
- Suleiman, Y., 2003. *The Arabic Language and National Identity: A Study in Ideology*. Edinburgh University Press, Edinburgh.
- Temple, C., 2010. The emergence of Sankofa practice in the United States: a modern history. *J. Black Stud.* 41 (1), 127–150.
- Tuchscherer, K., Hair, P.E.H., 2002. Cherokee and West Africa: examining the origins of the Vai script. *Hist. Afr.* 29, 427–486.
- Tuchscherer, K., 1995. African script and scripture: the history of the Kikakui (Mende) writing system for bible translations. *Afr. Lang. Cult.* 8 (2), 169–188.
- Voogt, A.de, 2010. The meroitic script and the understanding of alpha-syllabic writing. *Bull. Sch. Orient Afr. Stud.* 73 (1), 101–105.
- Weth, C., Juffermans, K., 2018. *The Tyranny of Writing Ideologies of the Written Word*. Bloomsbury Publishing, New York.
- Wiredu, K., 2009. An oral philosophy of personhood: comments on philosophy and orality. *Res. Afr. Lit.* 40 (1), 8–18.

Literacy, colonialism and resistance in Latin America

Lina Trigos-Carrillo^a, Miriam Jorge^b, Luzkarime Calle-Díaz^a, and Rebecca Rogers^b, ^a Universidad del Norte, Barranquilla, Atlántico, Colombia; and ^b University of Missouri-St. Louis, St. Louis, MO, United States

© 2023 Elsevier Ltd. All rights reserved.

Colonialism and coloniality in Latin America	736
Literacy as a form of domination during colonialism in Latin America	736
The mechanisms of coloniality: missions and literacy campaigns	737
Literacy as a form of resistance in Latin America	737
Educación Propia: a dialogic educational system	738
Intercultural education initiatives	739
Intercultural bilingual education from within: IBE	740
Urban youth literacies: contemporary forms of resistance in Latin America	742
Continuing tensions	743
Visions for the future	744
Acknowledgment	744
References	744
Further reading	746

Colonialism and coloniality in Latin America

After the “Conquest,” or European invasion, of the Americas, a process called *colonialism* started in this part of the world. Colonialism is the process of dominance of the European sovereignty over the non-Western people who occupied the territory before the invasion (Kohn and Reddy, 2017). Spaniards and Portuguese dominated the region known now as Latin America, which goes from Mexico to Argentina. However, Latin America is not only a geographical region, it is also a “historical experience” consequential of the project of colonialism (Quijano, 2000). Anibal Quijano (2000) identifies two important elements that determined modern societies in Latin America: first, colonialism imposed new forms of labor based on capital that were the foundation of new patterns of power; and second, the idea of “race” was created to establish a hierarchical difference between the dominant and the dominated based on physical ideals and also human capabilities, such as reason and humanity (Walsh, 2010).

With the introduction of these two elements, new social identities were formed: “Spanish,” “Portuguese,” “Indians,” “Negroes,” and “Mestizos”. The idea of race was used to position people in the power structure of the new society; this division determined the division of labor and who was in control of the resources (Quijano, 2000). Whiteness was associated with power, control and domination; for instance, white people were the only social stratum who could receive a salary. The Latin American genocide was also based on race; during the invasion and colonial times millions of Indigenous and African descendants were killed because of their race (Menéndez, 1972). This system of domination has been known as the *coloniality of power*, grounded in the relationship between race, capitalism and colonialism (Walsh, 2010). While *colonialism* is a political, intellectual and historical process, *coloniality* is the pattern of power used since colonial times (Walsh, 2010). The coloniality of power is also part of the project of modernity, where knowledge was adapted to the needs of capitalism (Veronelli, 2015). In this project, objective reasoning is privileged over other forms of knowledge, and Indigeneity was converted into a problem, “the indigenous problem” (Walsh, 2010). Those who were “colonized” were considered as less-human, and their knowledge and language as less rational or not even human (Quijano, 2000).

As mestizos, who were born from Spanish men and Indigenous women, gained more privileges and power, they wanted to distance themselves from the Indigenous and African culture and aspired to become Europeans (Walsh, 2010). For that reason, *mestizaje* has also been described as a physical and cultural whitening process (Wade, 1997). Through time, *mestizaje* became a discourse of power. After the independence of the Latin American countries, the coloniality of power continued. The new countries depended financially on Europe, and they still aspired to be like Europe (Quijano, 2000). In the 19th Century, the mestizo discourse built stronger as emergent states started the formation of their national identities (Walsh, 2010). In the 21st Century, Latin America still experiences the lasting effects of coloniality and the coloniality of power.

Literacy as a form of domination during colonialism in Latin America

Before the arrival of Europeans to the Americas, approximately 1400 Indigenous languages, or Amerindian languages, organized in 123 linguistic families, were spoken in Latin America (Wagner, 1991). With slavery, other African languages came into contact with these Indigenous languages during colonial times, creating contact languages or creoles, such as the *criollo palenquero* in Colombia (Patiño Roselli, 2002). In addition, Spaniards spoke different Spanish dialects since they came from different territories in Spain, such as Andalucía, Castilla and Extremadura (Marimón Llorca, 2020). In face of this linguistic diversity, one of the main colonialism

projects was the erasure of linguistic diversity in order to eliminate Indigenous and African cultural identities. As a result, in the 21st Century, only around 160 Indigenous languages are spoken in Brazil, 140 in Peru, 70 in Colombia and 68 in Mexico. In this way, particular cultural identities such as those of the Quechuas, Chibchas, Aztecs, Mayas, Incas, etc. were reduced to a single racial category: "Indians" (Quijano, 2000).

Some of the Amerindian groups possessed an important literary tradition in pre-Columbian and colonial periods. For instance, the Aztecs and the Mixtecs had hieroglyphic writing and the Mayans had a systematic writing system (King, 1994; Wagner, 1991). Part of the colonialism project of cultural extermination included the prohibition of all forms of literacy in Amerindian languages. As a consequence, in many Indigenous groups, "writing is associated with the official and dominant language, Spanish, and, by extension, with schooling" (King, 1994, p. 2).

Colonialism equated literacy in the dominant language with knowledge and power. Spanish and Portuguese missionaries were in charge of the process of acculturation and assimilation of the Indigenous communities through education. Since the 17th Century, the Western European version of rationality was imposed as one in which "the 'inferior' 'races' are 'inferior' because they are 'objects' of study or of domination/discrimination, they are not 'subjects', and, most of all, they are not 'rational subjects'" (Quijano, 2000, p. 221). This idea grounded the ideology that non-Western people in the Americas needed to be "civilized".

This civilization process involved a process known as the *coloniality of language*, a dimension of the dehumanization process through racialization of "colonized" people (Veronelli, 2015). The Western idea of language was that only civilized populations could express knowledge through language and literacy (i.e., people who spoke Spanish and Portuguese languages, which had Latin roots). The consequence of the coloniality of language was that Indigenous and African descendants were considered as unable of dialogue. When Indigenous and African descendants learned Spanish or Portuguese, it was not with the intention of having a dialogue with them, but that they could follow orders and understand instructions (Veronelli, 2015).

The mechanisms of coloniality: missions and literacy campaigns

With these ideas in mind, various religious missions were sent to the Indigenous communities with the purpose of evangelizing and civilizing the native peoples. One example of this process was the Capuchino mission to the Sierra Nevada de Santa Marta in Colombia, where members of the Capuchino religious congregation established in the community to educate Arhuaco peoples into the dominant language and culture since 1916. After several complaints made by the Indigenous population and a long political struggle, the Capuchinos were expelled from the territory in 1982 (Duarte, 2018).

With the massification of education during the 20th Century, in the 1940s some Latin American governments (such as Ecuador, Dominican Republic, Mexico, Honduras, Guatemala, and Peru) instituted the literacy campaigns, aimed at reducing illiteracy rates across the countries (Rodríguez, 2009). In 1947, the United Nations Educational, Scientific and Cultural Organization (UNESCO) determined that "fundamental education" and basic literacy should be included in the development goals of the following decades in the region. However, during the 1940s and 1950s, literacy campaigns had a focus on Spanish and Portuguese as dominant languages and the development of basic literacy skills not as a means of liberation, but as a requirement to function in the dominant society or to "civilize" the "ignorant". During these decades, the literacy campaigns functioned in line with the dynamics of the coloniality of language.

Now that the two concepts: *coloniality of power* and *coloniality of language* have been introduced and explained in the Latin American context, we will present how literacy has become a form of resistance to these processes: a "decolonial" turn.

Literacy as a form of resistance in Latin America

... Soy lo que dejaron

Soy toda la sobra de lo que se robaron ...

Soy América Latina

Un pueblo sin piernas, pero que camina

Calle 13 (2011).

Domination, however, is not the "single story" (Adichie, 2018) of Latin America. The popular song *Latinoamérica* by Calle 13 (2011) reflects in a few words how the "remains" left by the European invasion did not stop the Latin American peoples to continue moving forward and resisting the consequences of coloniality. If Indigenous and African descendants were considered as unable of dialogue, it is through dialogue or dialogic expressions that Latin American communities have resisted the coloniality of language

(Veronelli, 2015). For instance, through: (a) dialogic pedagogy, as an act of sharing realities and seeking liberation among subjects under horizontal conditions; (b) transmodernity, as a process to overcome post-modernity, which is focused on European culture, and to highlight the potential of aspects outside of Western modernity, such as knowledge and practices that have been excluded by hegemonic models (Dussel, 2004, 2005); and (c) interculturality, as an appreciation of diversity in multiple contexts, including the cultural and linguistic identity of each subject and the idea of disposing homogenization processes (Bertely et al., 2015; Tinajero, 2010; Tinajero and Englander, 2011; Walsh, 2010).

Dialogic pedagogy pretends to transform oppressive situations derived from the processes of colonialism and coloniality through critical dialogue and critical consciousness (Freire, 2018). Literacy, in its multiple forms, can become a form of resistance when it is part of decolonial action. That is, when people engage with literacy practices in ways that recognize the geopolitics of knowledge, seeks to dismantle relations of coloniality in its material and discursive forms, and is used in the service of land repatriation and human rights (Tuck and Yang, 2012). One example of these processes are the literacy crusades in Cuba in the 1960s and Nicaragua in the 1980s, characterized by “the mass involvement of the people, and the commitment of time and resources at the national level” (Osborne, 1990, p. 2) to achieve social transformation and empowerment. Further, there is a long lineage of multi-literacies scholarship in Latin American which has resisted autonomous models of language and literacy (e.g., De la Piedra, 2009; Kalman and Street, 2013; Trigos Carrillo and Rogers, 2017; Zavala, 2002).

In the following section, we describe some Latin American initiatives and forms of resistance to the coloniality of language, acknowledging that many of them intersect with the three perspectives mentioned above. We do not pretend to be comprehensive, but to provide representative and innovative initiatives that work against the coloniality of language while we also recognize that a long path still needs to be walked to dismantle the effects of the coloniality of language and power in education.

Educación Propia: a dialogic educational system

Schooling has served as the main source of “acculturation” of native communities from colonial times until today. In general, schooling was not only used to force a new language into Indigenous communities, it also imposed a way of thinking, a more vertical view of the relation between school and communities, where teachers were seen as “the experts”, promoting authoritarian relations between teachers and students (Molina and Tabares, 2014). These vertical attitudes did not allow room for more dialogic relations between teachers and parents, which work against communal participation (CRIC, 2004). It is natural, then, that schooling has also become central to Indigenous struggles against the disappearance of their culture, language, and identities.

The notion of *Educación Propia* (translated as *own education* in Colombia and *differential education* in Mexico) is found across different Latin American countries as a pillar in their struggle for permanence. In Colombia, this concept emerged in the mid-seventies, as an initiative of the Consejo Regional Indígena del Cauca (CRIC), with the purpose of creating educational strategies to ensure cultural strengthening and survival. At that point, the national government and the Catholic church were in charge of education, which was not always in favor of the Indigenous communities’ interests. Bolaños and Tattay (2012) posited that the school could not continue being a space of disintegration and weakening of the cultural identity, so they sought to turn it into a strategy for the foundation, reconstruction and permanence of cultures. González (2012) adds that the traditional/modern school contributed to the discrimination and lack of recognition of the Indigenous identity.

Educación Propia emerged as a response to these claims, opening significant spaces for dialogic encounters in three main areas: (a) a dialogue between school and community, which strengthens a sense of belonging: the school belongs to the community; therefore, it becomes the core of the struggle of the people, a tool for awareness and organization (CRIC, 2004). This also includes a powerful encounter between the elderly and the younger generations; (b) a dialogue between Indigenous languages and Spanish, revitalizing the character of Indigenous languages as knowledge builders, and turning them into adequate means to build new knowledges, and establishing dialogues from a strengthened identity, whose interest is that both, the Indigenous language and Spanish, are appropriated adequately in their role of vehicular language and revitalize their own ways of communication in their bilingual character (Bolaños and Tattay, 2012); and (c) a dialogue between Western and local epistemologies, allowing places of intersection where mainstream systems (e.g., research processes, notion of higher education) would be adapted to meet the diversity and cultural needs of the communities, with the consideration of local knowledge and beliefs as fundamental basis for building new knowledge.

Cauca is a key example of resistance in Colombia (González-Terreros and Torres-Carrillo, 2020), however, it is not the only one. Other Indigenous peoples in the country have also led processes of empowerment and transformation in their literacy practices to make sure they can survive as they dialogue with mainstream epistemologies. The Arhuacos, in the region of Sierra Nevada de Santa Marta, have used language planning (*planificación del lenguaje*) (Fishman, 1991) to sustain and revitalize their Ika language even with the pressure of Spanish over their communities (Murillo, 2006). Through the approach of Educación Propia, they have transformed national educational policies into more relevant practices that include their traditions, needs, uses and customs, adequate for their cultural ethnic diversity. For instance, they have defined the Institutional Educational Project (PEI for its name in Spanish), which is mandatory for all schools across the country, as a Community Educational Project (PEC) (Poche Cuetocué, 2021), which is built with the participation of the elderly, mothers, and teachers, and has seven basic components: mother language, history, community, Spanish, geohistory, natural sciences, and mathematics. They also include areas, such as handcrafts, folklore, ancestral traditional history, and cultural legacy (Mendoza, 2010).



Fig. 1 Reading manual in Palenquero Creole. Pictures taken by Luzkarime Calle-Díaz in Palenque, Colombia.

Language planning, as stated by Kloss (1969), usually has two foci: corpus planning and status planning. Both entail efforts made by speakers to sustain or revitalize their own language and, therefore, identity. The former refers to the creation, modification or revitalization of language forms or codes (vocabulary, spelling, etc.). The latter has to do with the recognition of the social functions of a language and its position in relation to others. This is particularly important in the case of minoritized languages reclaiming spaces facing national governments or mainstream languages. In Latin America, language planning has served as a bottom-up alternative to resist homogenization processes in the face of the pressure exercised by Spanish, Portuguese, and English on Indigenous communities (Hornberger, 1997; McCarthy, 2018).

The Afro-Colombian creole spoken in Palenque, the first town of free slaves in America (Del Castillo, 1984; Dieck, 2008; Friedemann and Patiño Roselli, 1983), is another example of resistance through literacy. In the first two decades of the 21st Century, the community has organized around the formalization of their *Palenquero* and the strengthening of their culture and identity, with the support of foundations and governmental agencies. Literacy development strategies such as the publication of bilingual and monolingual books in their language (Fig. 1), the construction of a community library, and a community center where the elderly share their knowledge with the youth, are contributing to achieve their goal.

The Mikaná ethnic group in the Atlántico department also has used Educación Propia, and the foundations of the Sistema Educativo Indígena Propio (Comisión Nacional de Trabajo y Concertación de Educación para Pueblos Indígenas - CONTCEPI, 2013) to develop their “own indigenous educational model for the Malambo territory”, which aims at revitalizing their language and culture, and articulating it with the national education system, based on the pillars of unity, autonomy, territory, culture, spirituality, and interculturality (Peralta et al., 2019).

In sum, Educación Propia constitutes an alternative to dismantle mainstream education which seeks to impose values that are not part of the native peoples’ identities. Educación Propia aims at strengthening native peoples’ identity and sense of belonging. As Indigenous people say, in this education, we learn our own language, cultural values, traditions, myths, dances, forms of production, in other words, we learn to be Indigenous (Comisión Nacional de Trabajo y Concertación de Educación para Pueblos Indígenas - CONTCEPI, 2013, p. 19).

Intercultural education initiatives

Across Latin America, the concept of interculturality is seen as emerging from the struggles over political and linguistic rights within Indigenous communities (Tinajero and Englander, 2011). In Mexico, intercultural policies and practices have been instituted from the top-down and the bottom-up. By way of example, in 1992, the General Directorate of Indigenous Education (DGEI) produced a document that centered the voices and actions of Indigenous communities in their own education (Tirzo Gómez, 2021). It posited that Indigenous communities have the right to participate in their own development. However, despite national reforms, there have

still been many problems such as: a lack of infrastructure and funding to support the reforms; professionalization of teachers who speak indigenous languages; lack of appropriate pedagogical materials including access to the Internet and digital devices.

Consequently, one of the primary causes of the Zapatista rebellion starting in 1994 was the lack of substantive emphasis of Indigenous languages in the educational system (e.g., Tseltal and Tzotzil) (García and Velasco, 2012). In 1995, Mayan educators from the state of Chiapas created Unión de Maestros de la Nueva Educación para México (UNEM), a union for educators.

Nigh and Bertely (2018) reviewed three related, contemporary intercultural education projects designed in Chiapas by Indigenous educators. Each of these is considered to be an Indigenous led, intercultural inductive method (referred to as MMI, its name in Spanish). The goal of MMI is to formalize Indigenous pedagogies and knowledge. Examples of these projects include Laboratorios Socionaturales Vivos, Milpas Educativas (Living Socio-Natural Laboratories and Educational Milpas) and Laboratorios para la Vida (Laboratories for Life, a school garden project) (Nigh and Bertely, 2018; Sartorello, 2021).

Some of the unifying elements of these projects include the following: establishing relationships between education, cultural activity, and ecological awareness; expanding Indigenous knowledge, values, and meanings with Western scientific knowledge; inviting participants to engage in school gardens and educational milpas; and appreciating that dialogue is not free of conflicts (Nigh and Bertely, 2018, p. 5). In particular, what Gasché (2008) refers to as “cultural syntax” is vital to the efforts of linguistic and cultural preservation of Indigenous knowledge.

The UNEM (Unión de la Nueva Educación de México), formed in 1995 by Mayan educators, evolved into the Network of Intercultural Education, REDIIIN (Red de Educación Inductiva Intercultural), with over three hundred participating educators (Moreno Medrano and Corral Guillé, 2019). The network includes collaborations between Indigenous and non-Indigenous educators and across countries (Mexico and Brazil). This contribution documents some conditions that ensure the permanence of the altermunicipal educational project developed by the Inductive Intercultural Education Network in Mexico and Brazil.

Intercultural bilingual education from within: IBE

The notion of intercultural bilingual education (IBE) has gained importance throughout Latin America in the last decades. It emerged in response to continuous struggles from native communities to reclaim the status of their linguistic and identity diversity in different countries. This notion entails not only the right to create and implement their own educational practices, but also the dialogue between Indigenous communities and hegemonic epistemologies.

In a thorough review of IBE and the intercultural approach in Latin American educational systems, Corbetta et al. (2018) provided an overview of different models that have been put into place in these countries. The connection between interculturality and education became more visible after the acknowledgment of 11 Latin American countries (Argentina, Bolivia, Brazil, Colombia, Ecuador, Guatemala, Mexico, Nicaragua, Paraguay, Peru and Venezuela) of their multiethnic, plurilingual, and multicultural character in their national constitutions between 1990 and 2010. Chile and Panama did so through specific laws and a ministerial policy, respectively. Only four countries also included Afro-descendant communities in their Carta magna (Corbetta et al., 2018).

When analyzing the status of Indigenous and Afro descendant languages with respect to official languages, Corbetta et al. (2018, p. 64) found that countries like Bolivia, Colombia, Ecuador, Nicaragua, Peru, and Venezuela include vernacular languages (e.g., Kichwua, Suar, Guaraní, Quechua, Aimara, and others) as official in their (or parts of their) territories. In the case of Brazil, Indigenous communities are entitled to use their mother languages and their own learning methods. In countries like Mexico and Costa Rica, the state has to ensure the maintenance, preservation, and cultivation of their native languages. Panama and El Salvador declare them objects of study, conservation and/or divulgation. On the other hand, Guatemala, Venezuela, and Paraguay place the languages spoken in their territories as cultural patrimony of their nations.

Ecuador has been one of the referents in Latin America in terms of IBE. The struggle of more than 30 Indigenous peoples and nationalities across the country yielded its fruits with the current National Constitution, issued in 2008, by which the state recognized the intercultural and plurinational character of the nation. Ecuador has a Model of Intercultural Bilingual Education (MOSEIB), as part of the Ministry of Education policies, that currently generates guidelines for 17 nationalities, and 18 Indigenous peoples that co-live in its territories (Bastidas, 2015). Paronyan and Cuenca (2018) highlight some positive impact of the project in terms of attendance, promotion, and decrease of drop-out and repetition. It has also increased communal participation, and the intervention of parents and Indigenous families in the development of educational activities. Other important achievements have to do with teacher education programs in pedagogical institutions, as well as the incorporation of intercultural bilingual education programs in two important universities in the country (Paronyan and Cuenca, 2018). Some of the strategies put forward by IBE in Ecuador have centered around decolonization practices, as underlined by Conejo (2008, p. 75):

- Learning of Indigenous languages, as a process of pedagogical decolonization and interculturalization of children, starting with the Community Family Children's education that begins before their birth.
- The inclusion of bilingual teachers so that communication with students, and learning become fluid. This replaced the homogenic teaching of Spanish that worked against the postulates of interculturality.
- A rupture in the teaching of disciplines to allow for studies in Indigenous languages, history of Indigenous people, Andean literature, life sciences, Indigenous medicine; together with universal knowledge, technology and English as a strategy to access other cultural spaces.
- The transformation of traditional schools into Community Educational Centers (CEC), conceiving learning beyond the classroom, and turning the whole community into a space for learning and reflection, promoting the interculturalization of

learning spaces and methodologies. This includes workshops, agricultural tasks, institutions, meetings, *mingas*, plants, animals, ceremonial sites, rivers and mountains.

- Recognition of the elderly as the “great wise men” who share cultural knowledges. In the IBE view, these men are not considered illiterate, as they would in the Eurocentric, dominant perspective, considering they cannot read or write. On the contrary, they possess great knowledge and experience that can be passed on to their children and grandchildren.

IBE has also been used in Venezuela, as an alternative to resist the colonality of knowledge/power (Mignolo, 2003), and strengthen the identity of the Wayúu communities located in the northern urban context of Maracaibo. By means of learning communities, some efforts have been made to sustain their ancestral knowledge and native language, despite continuous challenges and tensions (Guerra and Finol, 2015).

Bolivia is one of the few countries in Latin America where the majority of the population is Indigenous, with presence both in rural and urban contexts (López, 2006). There are 36 Indigenous groups and 33 Indigenous languages across the country; three of them have been declared official languages: Quechua, Guaraní, and Aymara, due to their significant number of speakers. IBE emerged in 1989, sponsored by the National Ministry of Education and the United Nations International Children's Emergency Fund (UNICEF). It led to an educational reform prepared by the Equipo Técnico de Apoyo a la Reforma Educativa (ETARE) and issued in 1994 (López, 2005). By 2001, after a series of national educational policies that supported intercultural and bilingual education, 115,000 Quechua, Guaraní, and Aymara children had been included in this new curriculum, with a coverage of 2,300 rural schools (Albó and Anaya, 2004). According to Arispe (2008), some important achievements of IBE in Bolivia are: social participation in education, teacher education programs, curriculum management, and social appropriation of their own languages.

In Bolivia, as in other Latin American countries, IBE constitutes an important component in the formation of new mentalities and visions about the relationship State-Indigenous peoples (López, 2005). This has allowed indigenous groups to legitimately interact and interpellate the government beyond Indigenous issues of land and education, and toward the common good of the nation (López and Küper, 1999). Important events in the history of the country, such as La Guerra del Agua in 2000, and La Guerra del Gas in 2003 marked important milestones in the recognition of Indigenous rights and perspectives against neoliberal projects. These contributed to a process of structural transformation, which included the resignification of national educational policies toward the construction of decolonizing and intercultural education (Gómez et al., 2010). In discussing the advances and challenges on IBE in Bolivia, Zavala (2007) writes, “the implementation of IBE is the cornerstone of a new citizenship” (p. 38). The author goes on to describe three strategies that have allowed to rethink IBE in the country: (a) a strategy of revitalization of Bésiro language in Chiquitanía; (b) a strategy for the teaching of Spanish as a second language in the Guaraní zone; and (c) a strategy of social participation in education as part of the action of the Machaqá Amawt'a Foundation in an Aymara zone of the La Paz department. The first one is particularly interesting for the purpose of this paper because it is an initiative of an Indigenous group (Chiquitanos) that constitutes only the 3% of the Indigenous of Tierras Bajas in Bolivia. However, they have acquired important political force in favor of a language that has been in danger in relation to Spanish. This is a clear example of resistance in a scenario where usually only native languages that are spoken by a significant number of people have been considered by national governments.

In Peru, Quechua is the major indigenous language of the country with more than 3 million speakers, and it “has long been associated with social and political marginalization, economic poverty, and low educational achievement” (Zavala, 2019, p. 60). However, “many NGOs and Indigenous organizations have emphasized the importance of Indigenous knowledge and have considered how to incorporate it into the schooling process” (Zavala, 2014, p. 47). In 2003, the government promised IBE across the educational system (although the efforts to implement IBE started in the 1980s), guaranteeing education in a mother tongue and in Spanish as a second language for all the bilingual people (Trapnell and Neira, 2004). However, the implementation of IBE came with several challenges as many Indigenous populations (Quechua speakers and speakers of other Indigenous languages) have been excluded from these programs and forced to Spanish-only education; the IBE program is often limited to primary education and rural areas; and IBE initiatives are still peripheral and depend largely on the good will of local authorities (Trapnell and Neira, 2004).

While some IBE initiatives focus on language survival following traditional educational methodologies and a governmental agenda, in other initiatives, IBE is linked to the recovery of cultural practices, interculturality, and empowerment with a critical view to racism and discrimination (Zavala, 2014). In this process, literacy planning has been used to revive, maintain and spread the use of Quechua among Indigenous and non-indigenous Peruvians as a resource in a multilingual environment (Hornberger, 2012). After years working with Indigenous communities, Hornberger (2012) states that functional Quechua literacy contributes to “marginalized monolingual and bilingual Quechua speaking populations' fuller participation and empowerment in society” (p. 231). Based on the New Literacy Studies perspective, other scholars have recognized the hybrid literacy practices (alphabetic Quechua, vernacular language, and Spanish) of Indigenous and rural people in the Andes. De la Piedra's (2009) ethnographic work provides thick descriptions of hybrid and bilingual literacy practices of a Quechua community in the Peruvian Andes. This study invites us to broaden our understanding of literacy to include the multimodality of rituals, ceremonies, and living in harmony with the environment.

In other Peruvian contexts, literacy practices are used to question political dynamics around language and to claim a space for Quechua literacy as decolonial action. In particular, urban Quechua youth have constructed alternative versions of Quechuaness to question dominant IBE ideologies by carrying linguistic political actions (Zavala, 2019); “the activists' aim is not to ‘preserve’ languages within a dominant discourse of ‘inclusion’ of the neglected population; rather, they hope to make powerful claims of

belonging and exert new power relationships" (Zavala, 2019, p. 78). "Furthermore, students engage in literacy practices within an increasing elasticity of the concept of language. They deconstruct the association between language, community, identity and culture; they develop translingual practices; use language to change the world; draw on repertoires of social semiotic resources in contemporary contexts; and challenge some of the dichotomies of modernist linguistics (such as language and society; structure and use; native and non-native, etc.)" (Zavala, 2018, n.p.). These initiatives present an elaborated understanding of IBE as a complex concept that can be used for very different purposes.

Brazilian indigenous communities have been deprived of the most basic needs of successive exploitation since colonial rights. Still, resistance occurs in educational approaches (Monzilar, 2020) such as the case of the Kashinawa production of multimodal texts (Menezes de Souza, 2003), which create new relationships between speech, writing, and thought. As stated by Menezes de Souza (2003), the Kashinawa's texts come from complex synaesthetic attempts of transpositions of the wealthy and complexity of Indigenous orality to paper, as "voices on paper." The Kashinawa literacy practices perhaps represent a more or less permanent hybrid form of non-alphabetic writing.

Intercultural teacher education in Alto Xingu resists coloniality of language through challenging colonial values represented by the use of Portuguese in Indigenous schools. It is the case of Kamaiura-Aweti (Kamaiura, 2012) schooling, where classroom events are intertwined with community events. Indigenous teacher education programs, in normal schools, reflect literacy events and practices that are intertwined with conflicted acceptance of Portuguese as a *Lingua Franca* (Kamaiura, 2018). Multimodality allows for approaches to language teaching that considers drawing and storytelling as strategic to creating spaces for sharing traditions and making meaning of colonial experiences.

As we have seen in this section, IBE takes different forms and follows diverse purposes depending on the level of participation of the community and the use of decolonial practices. Although it does not come free of challenges, IBE initiatives offer opportunities to work against the dynamics of the coloniality of language and power.

Urban youth literacies: contemporary forms of resistance in Latin America

Resistance takes many forms in contemporary literacies in Latin America. Social movements continue promoting culturally sustaining approaches to education. Indigenous communities claim ownership of their epistemologies. Through movements that involve rap music, graffiti, blogging, and social media activism, Latin American urban youth represent social practices where literacies enhance the visibility of oppression against peripheral women and men.

Many times, Latin-American youth have been portrayed as deviant or in terms of deficiencies and absence of cultural capital, reading, and writing habits (Trigos-Carrillo, 2019). Bartlett et al. (2011) point that beyond school-based expressions of literacy have grown as an effect of globalization and technology. However, these new forms of literacy allow for social participation in ways that school-based literacy practices tend to delegitimize. Concerning youth literacies, schools still neglect the dehumanizing social and racial conditions experienced by the urban youth in Latin America, and the transformations demanded in forms other than colonial practices.

Latin American rappers, bloggers, social activists, and diverse youth collectives, movements, and leaders engage literacy practices that run counter-hegemonic approaches to education. Resistance is enacted through overlapping literacy practices that reflect contradictions and possibilities of creating new ways of knowing and doing. In Brazil, contemporary literacy research focusing on youth activism, social justice and structural racism, coins terms like *survival literacies* (letramentos de sobrevivência; Maia, 2018), literacies of re-existence (letramentos de reexistência; Souza, 2011) highlight the counter-hegemonic power of literacy practices in transgressive peripheral literacies.

Considering the myriad of metaphorical meanings for different literacy approaches, Windle et al. (2020) offer literacy scholars a pathway for breaking with established paradigms that divorce knowledge production about inequality from subjects and territories. The *transperipheral paradigm* (Windle et al., 2020) aims at connecting regional, national, and global peripheries in the production of knowledge that combines knowledge about, from, and between peripheral communities. Transperipheral literacies and projections offer a paradigm for literacy research that makes meaning of existence, resistances, and languaging in peripheral spaces, intertwined with experiences of exclusion in contemporary forms of oppression.

Hankin (2017), for example, discusses how Cuban hip hop represents a "marginocentric" and transnational flow of cultural capital and the decolonization of knowledge. Cuban hip hop represents an example of how 21st Century literacies in Latin America can be examples of resistance to unequal power relations expressed through the engagement with local experiences of racism, exclusion and re-existence(s) (Souza, 2011). See, for example, the Cuban hip hop group *Orishas* (2006, 2014). Latin American non-school literacy practices widen limited perspectives of literacy that do not consider literacies from below, where social, gender, racial, class, and territorial experiences create spaces for dissent and resistance.

Music, as contemporary literacy practice, has also constituted a tool of resistance for Afro descendants in Colombia. The Rap Folkórico Palenkero by *Kombilesa* (2017, 2020), a group of rappers from Palenque, is an example of using popular art to unveil issues of racism and discrimination in their own creole language. *Chocquibtown* (2010) is a very popular hip hop band from Chocó, the main Afro Colombian (and one of the poorest) region in the country, whose lyrics and artistic representations are used for reclaiming the recognition of their African heritage, their identity, and the key role of Afro descendants in the history of the country. They also raise awareness on systemic racism that still exists across Colombia.

In Brazilian urban contexts, literacy practices use non-school spaces to include the performance of the social identities of Black youth. Black youth experiences Brazil from historically neglected and violent spaces, and resistance takes the form of political and cultural activism in literacy practices. Since the writing of samba lyrics (Vianna and Chasteen, 2000) to contemporary rap music, youth has used art as forms of resistance (e.g., Sabotage, 2013). Present transnational experiences have strengthened the hip-hop culture in Brazilian peripheries. Souza (2011) observes that Black youth in the peripheries of São Paulo captures the social and historical complexity of languaging, and destabilize fossilized power relations through the combination of non-hierarchical multiliteracy practices in the hip-hop movement. Fanzines, rap lyrics, graffiti, and other hip-hop elements transcend schooled views of language and text. The result is that Black youth resists through performing their (re)existence in situated literacy practices.

Critical literacies movement became more robust in the past decade in Brazil. Applied linguists problematize the coloniality of teaching and learning languages in Brazilian contexts, in a movement that reflects some aspects of national guidelines for foreign language education. However, the threat to freedom of speech in schools, represented by claims such as “schools without party/escola sem partido” aim at denying the Freirean view of education as a political act. Critical literacies as an approach to language education are currently under the threat of a conservative, retrograde government that views education as mere knowledge transmission.

Continuing tensions

Despite the great cultural diversity and the significant achievements of regional movements against systemic racism and discrimination (Arocena and Sobottka, 2017), Indigenous and African descendant people in Latin America continue to experience the consequences of colonialism in the 21st Century, such as land displacement and violence (Reding, 2007).

Endogenous and hybrid literacies become central to resisting the coloniality of language, and of power. However, a myriad of tensions continues to exist in the search for more visibility and sustainability of own cultures, languages, and identities. To begin with, intercultural bilingual education (IBE) has several definitions according to the way people approach it. Governmental views of IBE seem to entail a one-way process, where Indigenous communities are required to dialogue with mainstream cultures, but not the other way around (Paronyan and Cuenca, 2018). For instance, in Peru, despite the efforts to revitalize the Quechua language, IBE discourse presents a “depoliticized, limiting and fictitious view of Quechua and Quechuaness” (Zavala, 2019, p. 60). In this sense, García and Velasco (2012) conclude, “as long as bilingual education in Indigenous languages/Spanish is considered only an Indigenous thing, and not a Mexican thing, it will not accomplish much ... Right now, non-Indigenous Mexicans remain largely apathetic toward the Indigenous presence, and bilingual education is entirely one-sided. Racism and linguisticism continue to play a part in the contemptuous attitude toward Indigenous peoples and their languages. This has been a long-standing problem in Latin America” (p. 15).

In addition, there are tensions between neoliberal/governmental functional views and critical perspectives of intercultural education (Ferrão Candau, 2010). While some Latin American governments have included intercultural education programs and initiatives, most of the time these programs are not intended to question the current socio-political model or power structures and are guided by neoliberal policies. However, to dismantle the coloniality of power, a critical perspective of intercultural education is needed, where historical inequities, unjust power relationships, and patterns of domination are questioned and transformed (Ferrão Candau, 2010). Further, intercultural education should not be only for an ethnic or linguistic group, it should be education for all so that we all understand the historical roots of racism and the importance of cultural difference in Latin American societies.

On a similar note, Indigenous communities continue to live under the fear of discrimination and oppression. When joining mainstream educational settings, using their language or any other cultural representations can be regarded as inferior. Some attitudes toward learning that are inherent to their culture: the value of observation, experience, respect for the elderly, importance of oral transmission and cooperation, might be interpreted differently from hegemonic views, where competition, memorization, verbalization, and discipline are the norm. This creates stereotypes that label Indigenous individuals as slow, timid, with scarce vocabulary, or passive in their learning process (Corbetta et al., 2018).

Additionally, language visibilization and revitalization has mainly worked bottom-up. Latin American governments and states have begun to acknowledge their rights only after long, often painful, struggles from native groups. However, it is not enough to promote language survival and Indigenous literacy if it is framed in what Zavala (2018) labels as “essentializing and ancestralizing ideologies of language and identity” (n.p.) that go hand in hand with the coloniality of language and modernist views of education.

On the other hand, tensions related to epistemic diversity arise when Indigenous and Afro descendant communities start to dialogue with hegemonic views. Notions such as research, academia, “higher” education are all influenced by western visions of knowledge, learning, and literacies. This has an impact on who dialogues with whom and by which means, therefore contributing to unbalanced flows of knowledge. This was clearly visible when conducting research for this paper. Finding literature about Indigenous literacies written by Indigenous people is scarce in Latin America. Science and research are still perceived from a Eurocentric, dominant perspective (Mignolo, 2013), which often translates into academic literacies (and superiority) being measured by, for instance, the number of written publications (Hanauer et al., 2019). This leaves aside the richness of other literacies that carry more value for Indigenous and Afro descendant peoples (Useda Sánchez, 2020). More openness is needed in terms of the hierarchies of literacies, and the superiority of written over oral languages. We argue for the inclusion of more opportunities for epistemologies of the South (Santos, 2015) to come to the front and have spaces of amplification.

Visions for the future

Literacy in Latin America has long been affected by the colonial experiences that shaped both tensions and resistance. In our visions for the future, literacy practices will come from diverse backgrounds, movements and technology innovations that will embrace decolonial perspectives. By incorporating multiple perspectives, needs and possibilities, literacy scholars and educators will continue engaging in decolonizing imbalanced power relations that lead to the dominance of certain forms of literacy over others. Literacy educators and scholars will develop fluid conceptions of literacies that include pathways to multimodal and hybrid learning and representation. We envision fluidity including academic literacies that affect the dissemination of knowledge and epistemologies. Transnational, transperipheral and multiple pathways to global and local connection mediated by digital technologies. Digital networks connecting learners, educators, scholars across historically contested or divided spaces will continue emerging as catalysts for literacies as resistance and disruption.

Literacies need to serve rather than threaten epistemological diversity and cognitive justice (alternative ways of understanding research, and educational processes, more visibility to Indigenous research in mainstream publications). Leaving coloniality behind requires creating spaces for understanding and disrupting the legacies of a violent and oppressive past that appropriated of Indigenous lands, Black bodies and erased multiple epistemologies, languages and cultures. The reclaiming of endogenous family and community knowledges as the basis of educational literacies and human development will lead to more ethical approaches to literacy education and research.

Finally, we understand that the scholarship of literacy will develop historical consciousness through resistant, transnational and critical literacy practices. As political acts, literacy educators and researchers will move literacies beyond the persistent oppression from colonial pasts.

Acknowledgment

We would like to acknowledge the support of Karen Colmenares, Colombian undergraduate research assistant.

References

- Adichie, C., 2018. *El Peligro de la Única Historia*. Penguin Random House, Bogotá.
- Albó, X., Anaya, A., 2004. La audacia de la educación intercultural y bilingüe en Bolivia. *Revista Andina* 38, 281–292.
- Arispe, V., 2008. *La Educación Intercultural Bilingüe en Bolivia*. PROEIB Andes, Cochabamba.
- Arocena, F., Sobottka, E., 2017. Diversidad cultural en América Latina. *Civitas. Revista de Ciências Sociais* 17 (2), 205–209.
- Bartlett, L., López, D., Mein, E., Valdiviezo, L.A., Hull, G., Villenas, S., 2011. Adolescent literacies in Latin America and the Caribbean. *Rev. Res. Educ.* 35 (1), 174–207. <https://doi.org/10.3102/0091732X10383210>.
- Bastidas, M., 2015. Educación Intercultural Bilingüe en el Ecuador: un estudio de la demanda social. *ALTERIDAD. Revista de Educación* 10 (2), 180–189.
- Bertely, M., Sartorello, C., Arcos, F., 2015. Vigilancia, cuidado y control étnico-político: Red de Educación Inductiva intercultural. *Descartors* 48, 32–49.
- Bolaños, G., Tattay, L., 2012. La educación propia: una realidad de resistencia educativa y cultural de los pueblos. *Educación y Ciudad* (22), 45–56.
- Calle 13. (September 27, 2011). *Calle 13 – Latinoamérica* [Video]. YouTube. <https://www.youtube.com/watch?v=DkFJE8ZdeG8>.
- Chocquibtown. (February 19, 2010). *De dónde vengo yo – Chocquibtown* [Video]. YouTube. <https://www.youtube.com/watch?v=yMS4J6Gp6e4>.
- Comisión Nacional de Trabajo y Concertación de Educación para Pueblos Indígenas - CONTCEPI, 2013. *Perfil del Sistema Educativo Indígena Propio*. CONTCEPI, Bogotá, Colombia.
- Conejo, A., 2008. Educación Intercultural Bilingüe en el Ecuador. *ALTERIDAD. Revista de Educación* 3 (2), 64–82.
- Consejo Regional Indígena del Cauca, 2004. *¿Qué pasaría si la escuela?...30 años de construcción de una educación propia*. Programa de Educación Bilingüe e Intercultural – PEBI, Popayán.
- Corbetta, S., Bonetti, C., Bustamante, F., Vergara, A., 2018. Educación intercultural bilingüe y enfoque de interculturalidad en los sistemas educativos latinoamericanos. *Avances y desafíos*. Naciones Unidas, Santiago.
- Del Castillo, N., 1984. El léxico negro-africano de San Basilio de Palenque. *Thesaurus* XXXIX (1–3), 80–170.
- De La Piedra, T., 2009. Hybrid literacies: the case of a Quechua community in the Andes. *Anthropol. Educ. Q.* 40 (2), 100–128.
- Dieck, M., 2008. La lengua de Palenque: avances en la investigación de su estructura gramatical. *Linguística y Literatura* 54, 133–146.
- Duarte, J., 2018. *Expulsión de los misioneros capuchinos por la comunidad arhuaco en la Sierra Nevada de Santa Marta*. Tesis de grado. Pontificia Universidad Javeriana, Colombia.
- Dussel, E., 2004. Sistema-mundo y Transmodernidad. In: Banerje, I., Dube, S., Mignolo, W. (Eds.), *Modernidades Coloniales*. Editorial El Colegio de México, México.
- Dussel, E., 2005. Transmodernidad e Interculturalidad (interpretaciones desde la Filosofía de la Liberación). Universidad Autónoma de México, México.
- Ferrão Candau, V.M., 2010. Educación intercultural en América Latina: distintas concepciones y tensiones actuales. *Estudios pedagógicos (Valdivia)* 36 (2), 333–342.
- Fishman, J., 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*, vol. 76. Multilingual matters Ltd.
- Freire, P., 2018. *Pedagogy of the Oppressed*. Bloomsbury publishing, USA.
- Fridemann, N., Patiño Roselli, C., 1983. *Lengua y Sociedad en el Palenque de San Basilio*. Instituto Caro y Cuervo, Bogotá.
- García, O., Velasco, P., 2012. Insufficient language education policy: intercultural bilingual education in Chiapas. *Diaspora, Indig. Minority Educ.* 6, 1–18.
- Gasché, J., 2008. Niños, maestros, comuneros y escritos antropológicos como fuentes de contenidos indígenas escolares y la actividad como puntos de partida de los procesos pedagógicos interculturales: un modelo sintáctico de cultura. In: Bertely, M., Gasché, J., Podestá, R. (Eds.), *Educando en la diversidad: Investigaciones y experiencias educativas interculturales y bilingües*. Abya-Yala/CIESAS, México, pp. 279–365.
- Guerra, S., Finol, L., 2015. Comunidad de aprendizaje para la Educación Intercultural Bilingüe en indígenas Wayúu. In: Ochoa, H., Torrenegra, M. (Eds.), *Construyendo Pensamiento Crítico en Venezuela III*. Universidad Bolivariana de Venezuela, Maracaibo, pp. 297–322.
- Gómez, L., Gómez, J., Pineau, F., Mora, A., 2010. La educación desde los pueblos indígenas de Bolivia. *Análisis del proyecto de Nueva Ley de Educación “Avelino Siñani y Elizardo Pérez”, como resistencia frente a la educación para la globalización neoliberal*. *Temas de Nuestra América* 26 (48), 117–136.
- González, M., 2012. La educación propia: entre legados católicos y reivindicaciones étnicas. *Pedagogía y Saberes* 36, 33–43.

- González-Terros, M.I., Torres-Carrillo, A., 2020. Popular education and self-education: dialogue between educational experiences in Cauca, Colombia. *Revista Colombiana de Educación* (80), 335–354.
- Hanauer, D.I., Sheridan, C.L., Englander, K., 2019. Linguistic injustice in the writing of research articles in English as a second language: data from Taiwanese and Mexican researchers. *Writ. Commun.* 36 (1), 136–154.
- Hankin, C.D., 2017. The (Latin) American underground: Havana and marginocentric hip hop. *Atl. Stud.* 14 (1), 82–98.
- Hornberger, N.H. (Ed.), 1997. *Indigenous literacies in the Americas: Language planning from the bottom up*. Mouton de Gruyter, Berlin.
- Hornberger, N., 2012. *Indigenous Literacies in the Americas. Language Planning from the Bottom up*. Mouton de Gruyter, New York.
- Kalman, J., Street, B., 2013. *Literacy and Numeracy in Latin America: Local Perspectives and beyond*. Routledge, NY.
- Kamaiurá, W., 2012. *Aweti e Tupi Guarani: Relações Genéticas e Contato Linguístico*. Unpublished Master's Thesis. Universidade de Brasília, Brasília, DF, Brasil.
- Kamaiurá, V.J.S., 2018. *Resistência e Luta: Questões de Letramentos e Identidades Docentes Indígenas*. Unpublished Master's Thesis. Universidade de Brasília, Brasília, DF, Brazil.
- King, L., 1994. *Roots of Identity: Language and Literacy in Mexico*. Stanford University Press.
- Kloss, H., 1969. *Research Possibilities on Group Bilingualism: A Report*. International Centre for Research on Bilingualism, Quebec.
- Kohn, M., Reddy, K., 2017. Colonialism. In: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2017 Edition. Available at: <https://plato.stanford.edu/archives/fall2017/entries/colonialism/>.
- Kombilesa M. (January 10, 2017). RFP: rap Folklórico Palenquero con Kombilesa Mí [Video]. YouTube. <https://www.youtube.com/watch?v=eBrByLi3Gt4>.
- Kombilesa M. (May 2, 2020). Kombilesa Mí los Peinados Video oficial [Video]. YouTube. <https://www.youtube.com/watch?v=tCy0HphFRYk>.
- López, L.E., 2005. *De Resquicios a Boquerones. La Educación Intercultural Bilingüe en Bolivia*. Plural Editores, La Paz.
- López, L.E., 2006. Desde arriba y desde abajo: Visiones contrapuestas de la educación intercultural bilingüe en América Latina. In: Gros, C., Strigler, M.-C. (Eds.), *Être indien dans les Amériques: spoliations et résistance*. Institut des Amériques, Paris, pp. 235–250.
- López, L.E., Küper, W., 1999. La educación intercultural bilingüe en América Latina: balance y perspectivas. *Revista Iberoamericana de Educación* 20, 17–85.
- Maia, J.O., 2018. Letramentos de sobrevivência em redes digitais: Caminhos possíveis na luta por direitos humanos. *Trab. Lingüística Apl.* 57 (2), 954–974.
- Marimón Llorca, C., 2020. *El español en América: de la conquista a la Época Colonial*. Biblioteca Virtual Miguel de Cervantes. Available at: http://www.cervantesvirtual.com/obra-visor/el-espaol-en-amrica-de-la-conquista-a-la-poca-colonial-0/html/00f4b922-82b2-11df-acc7-002185ce6064_2.html.
- McCarty, T.L., 2018. Community-based language planning: perspectives from Indigenous language revitalization. In: *The Routledge Handbook of Language Revitalization*. Routledge, pp. 22–35.
- Mendoza, C., 2010. La educación indígena ika, kankuama, nasa, wayúu y mokaaná fortalecen la interculturalidad en Colombia. *Revista Educación y Humanismo* 12 (19), 148–176.
- Menéndez, E., 1972. Racismo, colonialismo y violencia científica. *Revista Transformaciones* 47, 169–196.
- Menezes de Souza, L.M., 2003. Voices on paper: multimodal texts and indigenous literacy in Brazil. *Soc. Semiotic.* 13 (1), 29–42. <https://doi.org/10.1080/1035033032000133508>.
- Mignolo, W., 2003. *Historias locales/diseños globales – Colonialidad, conocimientos subalternos y pensamiento fronterizo*. Ediciones Akal, Madrid.
- Mignolo, W., 2013. Geopolitics of sensing and knowing: on (de) coloniality, border thinking, and epistemic disobedience. *Confero: Essays on Education, Philosophy and Politics* 1 (1), 129–150.
- Molina, V., Tabares, J., 2014. Educación Propia. Resistencia al modelo de homogeneización de los pueblos indígenas de Colombia. *Polis, Revista Latinoamericana* 13 (38), 149–172.
- Monzilar, E.B., 2020. Narrativa da educação indígena e da educação escolar indígena: a escola e o ensino do povo Balatiponé-Umutina. *Arquivos Analíticos de Políticas Educativas* 28 (77). <https://doi.org/10.14507/epaa.28.476>.
- Moreno Medrano, L., Corral Guillé, G., 2019. Metodologías inductivas interculturales para una pedagogía decolonial. *Sinéctica* (52), 00003. [https://doi.org/10.31391/s2007-7033\(2019\)0052-003](https://doi.org/10.31391/s2007-7033(2019)0052-003).
- Murillo, L.A., 2006. La planificación del lenguaje: herramienta de resistencia política y cultural de los indígenas arhuacos de Colombia. In: Terborgh, R., Landa, L. (Eds.), *Los retos de la planificación del lenguaje en el siglo XX*. UNAM, México, pp. 255–279.
- Nigh, R., Bertely, M., 2018. Indigenous knowledge and education in Chiapas, Mexico: an intercultural method. *Diálogos sobre educación: Temas actuales en investigación educativa* 9 (16), 1–20.
- Orishas. (September 30, 2006). Orishas – 537 Cuba [Video]. YouTube. <https://www.youtube.com/watch?v=L2eV20UhkQk>.
- Orishas. (February 26, 2014). Orishas “Represent Cuba” (feat. heather headley) [Video]. YouTube. <https://www.youtube.com/watch?v=ZBoqCDSrWi8>.
- Osborne, T., 1990. Politics and Education: The Nicaraguan Literacy Crusade. *Dissertations and Theses. Paper 2038*. <https://doi.org/10.15760/etd.2037>.
- Paronyan, H., Cuenca, M., 2018. Educación intercultural bilingüe en Ecuador: retos principales para su perfeccionamiento y sostenibilidad. *Transformación* 14 (3), 310–326.
- Patino Roselli, C., 2002. Sobre origen y composición del criollo palenquero. In: Moñino, Y., Schwegler, A. (Eds.), *Palenque, Cartagena y Afro-Caribe: Historia y Lengua*. Max Niemeyer Verlag, pp. 21–34. <https://doi.org/10.1515/9783110960228>.
- Peralta, P., Cervantes, V., Olivares, A., Ochoa, J., 2019. Educación propia de la etnia Mokaaná: Experiencia organizacional contemporánea. *Rev. Cien. Soc.* 25 (3), 88–100.
- Poche Cuetocué, C.A., 2021. *Los proyectos educativos comunitarios-PEC: una ruta para la consolidación del sistema educativo indígena propio-SEIP en el Cauca* [Undergraduate tesis, UNAD]. Repositorio UNAD. <https://repository.unad.edu.co/bitstream/handle/10596/42676/capoche.pdf?sequence=3&isAllowed=y>.
- Quijano, A., 2000. Coloniality of power and eurocentrism in Latin America. *Int. Sociol.* 15 (2), 215–232.
- Reding, S., 2007. Diversity and racism in América Latina. *Latinoamérica. Revista de Estudios Latinoamericanos* 44, 157–179.
- Rodríguez, L., 2009. Educación de adultos en la historia reciente de América Latina y el Caribe. *Revista Éfora* 3, 64–82.
- Sabotage. (July 30, 2013). Sabotage - Respeito é Pra Quem Tem [Video]. YouTube. https://www.youtube.com/watch?v=S0a-_eHXHM.
- Santos, B.S., 2015. *Epistemologies of the South: Justice against Epistemicide*. Routledge, London.
- Sartorello, S., 2021. Milpas Educativas: Entramados sacionaturales comunitarios para el buen vivir. *Rev. Mex. Invest. Educ.* 26 (88), 283–309.
- Souza, A.L.S., 2011. *Letramentos de reexistência - Poesia, grafite, música, dança: hip-hop*. Parábola Editorial, São Paulo, SP.
- Tinajero, G., 2010. Bicultural o intercultural? El discurso oficial y su apropiación por parte de los maestros indígenas. In: López-Bonilla, G., Pérez-Fragoso, C. (Eds.), *Discursos e identidades en contextos de cambio educativo*. Playa y Valdes Editores, Mexico, D.F, pp. 283–316.
- Tinajero, G., Englander, K., 2011. Bilingual-intercultural education for indigenous children: the case of Mexico in an era of globalization and uprisings. *Intercult. Educ.* 22 (3), 163–178.
- Tirzo Gómez, J., 2021. Las lenguas indígenas en la educación primaria en México. Una propuesta educativa intercultural. *LiminaR* 19 (2), 262–265.
- Trapnell, L., Neira, E., 2004. *Situación de la educación intercultural bilingüe en el Perú*. Consultancy for the World Bank (Working document).
- Trigos-Carrillo, L., Rogers, R., 2017. Latin American influences on multiliteracies: from epistemological diversity to cognitive justice. *Literacy Res. Theory Method Pract.* (66), 373–388.
- Trigos-Carrillo, L., 2019. Community cultural wealth and literacy capital in Latin American communities. *Engl. Teach. Pract. Critiq.* 19 (1), 3–19. <https://doi.org/10.1108/ETPC-05-2019-0071>.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decoloniz. Indigen. Educ. Soc.* 1 (1), 1–40.
- Useda Sánchez, H., 2020. Oralidad y saberes ancestrales en la crianza de las semillas de vida. *Cambios y Permanencias* 11 (1), 1627–1645.
- Veronelli, G.A., 2015. Sobre la colonialidad del lenguaje. *Universitas Humanística* 81 (81), 33–58.
- Vianna, H., Chasteen, J.C., 2000. *The Mystery of Samba: Popular Music & National Identity in Brazil*. University of North Carolina Press, Chapel Hill.
- Wade, P., 1997. *Gente negra, nación mestiza. Dinámicas de las identidades raciales en Colombia*. Uniandes/Universidad de Antioquia/Siglo del Hombre, Bogotá.

- Walsh, C., 2010. Raza, mestizaje y poder: horizontes coloniales pasados y presentes. *Crítica y emancipación* 3, 95–124.
- Wagner, C., 1991. Las lenguas indígenas de América (lenguas amerindias). *Revista Documentos Lingüísticos y Literarios UACH* 17, 30–37.
- Windle, J., Souza, A.L.S., do Nascimento e Silva, D., Zaidan, J.M., de Oliveira Maia, J., Muniz, K., Castro, S.L., 2020. Towards a transperipheral paradigm: an agenda for socially engaged research. *Trabalhos Em Lingüística Aplicada* 59 (2), 1577–1579. <https://doi-org.ezproxy.umsl.edu/10.1590/010318137496812202007061>.
- Zavala, V., 2002. Desencuentros con la Escritura: Escuela y comunidad en los Andes Peruanos. Red para el Desarrollo de las Ciencias Sociales en el Perú, Lima.
- Zavala, V., 2007. Avances y Desafíos de la Educación Intercultural Bilingüe en Bolivia, Ecuador y Perú. Ibis Dinamarca, La Paz.
- Zavala, V., 2014. What is Quechua literacy for? Ideological dilemmas in intercultural bilingual education in the Peruvian Andes. In: Prinsloo, M., Stroud, C. (Eds.), *Educating in Language and Literacy Diversity, Mobile Selves*. Palgrave Macmillan, London, pp. 45–72.
- Zavala, V., 2019. Youth and the repoliticization of Quechua. *Lang. Culture Soc.* 1 (1), 59–82.
- Zavala, V., 2018. Ideological battles over Quechua literacy in Perú: from the authority of experts to the innovation of youth. In: Bloome, D., Rowsell, J., Leung, C., Castanheira, A.L. (Eds.), *Re-Theorizing Literacy Practices: Complex Social and Cultural Contexts*. Routledge, London.

Further reading

- Castro Gómez, S., 2000. Ciencias sociales, violencia epistémica y el problema de la "invención del Otro". In: Lander, E. (Ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales. Perspectivas Latinoamericana*. CLACSO, Buenos Aires, pp. 191–213.
- De La Piedra, M.T., 2006. Literacies and Quechua oral language: connecting sociocultural worlds and linguistic resources for biliteracy development. *J. Early Child. Literacy* 6 (3), 383–406. <https://doi.org/10.1177/1468798406069800>.
- Grosfoguel, R., 2011. La descolonización del conocimiento: diálogo crítico entre la visión decolonial de Frantz Fanon y la sociología decolonial de Boaventura de Sousa Santos. *Formas-Otras: Saber, nombrar, narrar, hacer* 97–108.
- Guilherme, M., de Souza, L. (Eds.), 2019. *Global Languages and Critical Intercultural Awareness: The South Answers Back*. Routledge.
- Maglia, G., Schwegler, A., Correa, J.A., Ferrari, L., Helg, A., Moñino, Y., Navarrete, M.C., 2012. Palenque (Colombia): oralidad, identidad y resistencia. Pontificia Universidad Javeriana, Colombia.
- Maldonado-Torres, N., 2007. Sobre la colonialidad del ser: contribuciones al desarrollo de un concepto. In: Castro Gómez, S., Grosfoguel, R. (Eds.), *El giro decolonial: Reflexiones para una diversidad epistémica más allá del capitalismo global*. Siglo del Hombre, Colombia, pp. 127–167.
- Mignolo, W.D., 1986. La lengua, la letra, el territorio (o la crisis de los estudios literarios coloniales). *Dispositio* 11 (28–29), 137–160.
- Mignolo, W., 2008. El pensamiento des-colonial, desprendimiento y apertura: un manifiesto. *Revista Telar* (6), 7–38.
- Ortega, Y. (2019). CESE7-Ideologies of language and literacy education [Podcast]. 27 November. Available at: https://soundcloud.com/chasingencounters/ces2e7-ideologies-of-language-and-literacy-education?fbclid=IwAR21e_xIEzf6UupeOG5KawO_TVPIWgXy9PPvWS9I4o6gJRD5VGPr0eZlw.
- Pereira, A.C., 2019. School without party: around a discursive formation. *Perspectiva* 37 (3), 890–916. <https://doi.org/10.5007/2175-795X.2019.e55201>.
- Quijano, A., 1992. Raza, 'etnia', y 'nación': cuestiones abiertas. In: Forgues, R. (Ed.), *José Carlos Mariátegui y Europa: la otra cara del descubrimiento*. Amauta, Lima, pp. 757–775.
- Silko, L.M., 1995. An expression of profound gratitude. In: Katzenberger, E. (Ed.), *First World, Ha, Ha, Ha! the Zapatista Challenge*. City Lights Press, San Francisco, pp. 205–207.
- Zavala, V., 2008. Mail that feeds the family: popular correspondence and official literacy campaigns. *J. Dev. Stud. Liter. Ident. Social Change Interdiscip. Approach. Liter. Dev.* 44 (6), 880–891. <https://doi.org/10.1080/00220380802058248>.
- Zavala, V., 2015. "It will emerge if they grow fond of it": translanguaging and power in Quechua teaching. *Ling. Educ.* 32, 16–26. <https://doi.org/10.1016/j.linged.2015.01.009>.

Chinese literacy acquisition: an overview

Liqing Tao, Department of Curriculum and Instruction, School of Education, College of Staten Island, City University of New York, Staten Island, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	747
Language and writing system in China: a quick survey	747
Oral language in China: dialects and cross-dialect communications	747
Chinese writing system: character tones, formation principles of characters, and challenges	748
Answers to the challenges embedded in Chinese orthography	750
Writings and language beyond orthography: Wenyan Wen and Baihua Wen	752
Literacy acquisition in Chinese: highlighted features and associated research efforts	753
Conclusion: what we know and need to know more	755
Acknowledgments	756
References	756

Introduction

Literacy is informed by linguistic and social processes. While literacy and its acquisition are reflecting its corresponding language characteristics, they are undoubtedly also socially-based capacities and activities, particularly in their embodied social values and sociocultural environ including political and ideological tendencies (Santana, 2000; Venezky, 1996). The current chapter will give an overview of Chinese literacy in this broad sense, highlighting not only the linguistic features that underlie literacy and its acquisition in China, but also some related social forces and language practices that have impacted literacy conditions. Specifically, the chapter is divided into three sections: A description of language and writing systems in China, a summative highlight of literacy acquisition features and reformative and research efforts, and a short conclusion to offer some directions for further research that can have implications for literacy acquisitions and literacy policy. To make it accessible to the western readers accustomed to literacy in alphabetic language context, the current chapter will provide, whenever possible, a comparative frame of reference to English when discussing the literacy situation in China.

Language and writing system in China: a quick survey

Different writing systems capture their respective languages differently (Hooker, 1990). Orthographies of different writing systems are usually categorized by their different degrees of phonological transparency in their graphs to languages (Katz and Frost, 1992). In this sense, Chinese can be characterized as a deep orthography with tenuous signaling of its spoken language in its graphs. In addition, the Chinese writing system has two extra layers of complexity in its relationship with spoken language: its enduring role as the script for multiple regional dialects across China, and that the predominant form of writing used to be a highly literate form called Wenyan Wen with little connection to oral language (Norman, 1987). Both features further complicate the relationship between the written language and spoken language. In this subsection, we will briefly describe the spoken language situation, the characteristics of the writing system, the practices of literacy acquisition in the context of its writing system and the language situation.

Oral language in China: dialects and cross-dialect communications

Chinese oral language, like any native tongue in the world, is the language spoken by the Chinese people. But its relationship with literacy is more complicated than that in any alphabetic language writing system, which is mainly due to the fact that the relationship is mediated by the multiple dialects in China that are not mutually intelligible, and by the one writing system. Briefly speaking, aside from languages of national minorities, the main Chinese language is the one spoken by the Han nationals. Among the Han Chinese, there are seven major dialects: Mandarin, Min, Hakka, Yue, Wu, Gan, and Xiang dialects (Chen, 1999). A majority of more than 70% of Chinese speak a variant of Mandarin in the areas mostly north of the Yangtze River, whereas the other six dialects are mainly concentrated in the southern and southeast regions of China. Even though all the dialects share fundamentally similar grammar and semantic features, they do have drastically different pronunciations that make their users distinct. Without training or extended exposure to other dialects, a dialect user will be limited in his/her ability to communicate across dialect zones. For alphabetic language users, such oral language situation may be difficult to conceive, particularly when it comes to communications across dialect regions. For Chinese people, this is a common and practical issue and there are usually three ways to overcome the dialectal barrier between two dialect regions. Firstly, people can either receive focused training or pick up the dialect through

extensive exposure, as observed of emerging prominence of Cantonese (Yue dialect) words by Ichikawa and Komatsu (2008, pp. 184–188). Frequent exposures, both purposeful and unintentional, made the economically advanced Guangdong's Yue dialect more popular and familiar among other dialect users. Secondly, a common language for communication has always been in existence to bridge the communication barrier across dialect regions. It has always been the dialect spoken around the capital of the time (Wilkinson, 1998). The modern day common language—Putonghua, a mandarin dialect variant around Beijing region—started its development a little over a hundred years old, making it the latest effort to counter the adverse effect of dialects for communications over the vast territory of China.

Thirdly, people can communicate through written language. Literate people from different dialect regions can communicate with each other if they are unable to do so orally. Chinese script is unified and understood to anyone who can perform functional literacy to their own abilities such as reading and writing a letter regardless of different dialects. Thus, traveling to other dialect regions will not pose communication difficulties for people when they can rely on writing (Chu, 1973; Schirokauer, 1989). Occasionally, on-the-spot “pen talk” is also resorted to when oral communication is not successful (Chang and Chang, 1978), highlighting the communicative utility of the Chinese writing script as a saving feature for the diverse dialect situation.

In short, dialects have been in existence in China since antiquity, the communication obstacle has been a common daily phenomenon, making it both a problem and opportunity for literacy applications and reforms in China. The key to understand this lies in the fact that among all these mutually unintelligible dialects there is only one writing system, which we will discuss in some detail below.

Chinese writing system: character tones, formation principles of characters, and challenges

Scholars in linguistic and historical linguistics have studied the Chinese written language, and come up with definitions with degrees of difference in technicalities (See Qiu, 2009), but have consented to the general conclusion that it is a writing system different from an alphabetic system such as English (See Hansen, 1983). The present section is to highlight only these properties of the Chinese writing system that are relevant to literacy education and acquisition (See Tao and Qian, 2012a, pp. 2–7). All references to the phonological aspects of Chinese written language are made to the modern common language—Putonghua. And discussions of its orthographic nature are likewise about the modern Chinese unless otherwise indicated.

The Chinese writing system is a non-alphabetic system (Xu, 2009). Generally speaking, Chinese orthography can be considered ideographic in essence, consisting of mono-syllabic and tonal units called characters. Characters can be further categorized into simple characters and compound characters. Simple characters are stand-alone ones that have their own meanings such as mountain (山), water (水), person (人), or numbers (甲, 乙). They are few in number but very essential since they are still the most frequently used characters, helping as well constitute compound characters (Wilkinson, 1998). The compound characters are the majority (Zhou, Y. 2004), made up of two or more radicals and other components, some of which are still used as stand-alone simple characters. When written, all characters, simple or compound, take up an equal space, making them visually distinct individual units.

Since Chinese characters are non-alphabetic but fundamentally ideographic units, they are always pronounced as a single syllable entity (Zhou, Z. 2004). Phonemes as separate individual sub-units are never saliently pronounced by themselves in isolation, nor are they useful in offering help to readers and writers to identify and produce regular consistency at the phoneme level (Siok and Fletcher, 2001). For example, 人 (pronounced as/ren/) and 日 (pronounced as/ri/) are two simple characters. The pair shares the same beginning phoneme/r/. Yet, their physical forms contain no clues to that phoneme. Another pair of characters 人 (/ren/) and 从 (pronounced as/cong/) shares some semblance in their orthographic form. Yet, nothing at the phoneme level cues any phonetic connection between the two. In other words, knowing the characters and being aware of the constituting phonemes are not correlated with graphemes in the Chinese writing system. Therefore, while phonemic awareness is a very important insight and pre-requisite for learning English, it does not appear to be true in Chinese (see Holm and Dodd, 1996; Shu et al., 2006).

Modern Chinese as a tonal language has four tones at syllable level: level tone, rising tone, falling-rising tone, and falling tone (Chen, 1999). Each character has a tone attached to it. A number of characters would take two or more different tones often when they are used as homographs (see Table 1), which may also assume different pronunciations. But the majority of characters are of one tone only. This syllable level tonal occurrence is often difficult for non-native speakers to comprehend and acquire mainly due to the lack of tone specificity at the syllable level in languages such as English. But in Chinese tones are fixed to each character or syllable, making it mandatory to use the tones at syllable/character level for differentiation among characters of similar syllable pronunciations. It is part of phonological information for characters. However, characters themselves do not have tonal indicators when written, even the partially useful phonetic radicals are of little help here (Zhou, Y. 2004), making the tones of limited use in helping with character recognition. But as part of the phonological feature of every Chinese character, tones are explored by research as a type of implicit knowledge of the language that connects the oral with the script, and have been found to play a subtle role in phonological memory of characters (see Shu et al., 2008; Xu, 1991; Yip, 2014; Zhang and McBride-Chang, 2010).

Character formations are at the core of Chinese script (Qiu, 2009). According to traditional understanding and practices, there are six principles that reflect Chinese character construction (Xu, 121/1989), first clearly explained by Xu Shen as follows: pictographic characters (象形), simple ideographic characters (指事), compound ideographic characters (会意), semantic-phonetic characters (形声), loan characters (假借), and inter-explaining characters (转注). Except for the last two, the other four are structurally-based analyses of character construction (Su, 1997). Understanding how characters are formed is believed to be

Table 1 Samples of homographs with different tones and homographs of different pronunciations.

<i>Homographs with same pronunciations but different tones</i>		
<i>Chinese character</i>	<i>Pronunciation and tone (meaning)</i>	<i>Pronunciation and tone (meaning)</i>
好	hao3* (good, nice)	hao4 (like)
创	chuang1 (injury)	chuang4 (create)
漂	piao1 (float)	piao4 (beautiful)
Homographs with different pronunciations and different or same tones		
强	qiang2 (strong)	jiang4 (stubborn)
行	xing2 (move)	hang2 (line)
恶	e4 (evil)	wu4 (loathe)

NOTE: The asterisk (*) denotes that the pronunciation in pinyin is followed by a number that indicates one of the four tones in Mandarin Chinese.

helpful for Chinese learners (Zhang et al., 2000), even though knowledge of various structural information was found to be of different values at different stages of learning Chinese (Ho et al., 2003). Pictographic characters imitate the shape and look of the objects they used in ancient time and, no matter what script styles they evolve into, they still retain the abstract strokes of original characters, mostly providing learners with some lexical aid to access the meanings of the characters. Examples would be 日 (sun) and 月 (moon). These characters are not abundant but are usually among the first characters to be learned. They also serve today a very active role as radicals, often times in slightly altered variants, in a compound character in the last category of semantic-phonetic characters (Zhou, Y. 2004), making them very useful characters for beginning learners (Zhang et al., 2000).

Simple ideographic characters are also limited in number and symbolize concepts through their stroke structures or positions. For example, the characters 上 (up or above) and 下 (down or below) have a long bar with either a pointing post above or below the bar to signify the meanings. Such lexical suggestions, though, are more mnemonic than explicit meaning indicators, providing usually a memory support for beginning learners to access the characters. The next category of compound ideographic characters has an indirect indicative element but the meaningful radicals in them would combine to give a more complex and abstract meaning than simple ideographic characters can. For example, the characters 信 (trust) is expressing concepts that are born out of the fundamental radicals in them. The character “trust” is formed by a “person” radical and a “say” radical, signaling that “trust” should be something promised by someone in words (Qiu, 2009). This type of characters, acting very much like the previous two categories of pictographic and simple ideographic characters, is fundamentally symbolic and semantic in nature. But unlike the previous two, this type of characters is of dubious help for beginning learners. It should be noted that the commonly used characters of all these three categories culled together are no more than 500 in today’s writing. Most characters in current use belong to the next category: semantic-phonetic characters (Chen, 1999).

Semantic-phonetic characters refer to characters that have both semantic and phonetic radicals. Semantic radicals provide suggestions about meanings but in a broad manner, usually indicating a categorical property of the characters in question. Phonetic radicals likewise also provide some possible approximations but not with precision. Radicals, even with their broad suggestiveness, are still found to support learning to read in Chinese (Tong et al., 2017). To understand their functions, let us look at two pairs of examples: 铜 (copper, pronounced as/tong/) and 钢 (steel, pronounced as/gang/); 漂 (float, pronounced as/piao/) and 湖 (lake, pronounced as/hu/). The “metal” (钅) radical in the first two characters indicates semantically these two characters may belong to metals. The second pair of characters shares a “water” (氵) radical showing that the characters are water-related in meaning. These examples have provided a quick glimpse into how semantic radicals work to broadly hint at the property of their referenced things or concepts, offering a relative efficient semantic access for users via orthography without phonological mediation (Zhang et al., 2014). Phonetically for the all these sample characters here, they happen to have pronounced the same as their phonetic radicals standing alone in modern Chinese.

However, phonetic radicals, though useful in the earlier stages of learning characters (Tong et al., 2017), can be very ambiguous and are found to have an accuracy rate of only 39% for the pronunciations in modern Chinese characters (Zhou, Y. 2004), much lower than the estimated lexical suggestions of about 80% by semantic radicals (Tan and Perfetti, 1998), and not comparable to the phonological information in any spelled syllable in English (e.g., Ehri, 2014). Yet, since this type of semantic-phonetic characters makes up the majority of Chinese characters, these radical features do have effects on learners’ character recognition skills: learners are found to have developed an awareness of the positioning of common radicals in characters, and becoming more aware of possible suggestive phonological values of phonetic radicals and broad semantic categories (Tong et al., 2017), making awareness of the functions of radicals particularly important in learning Chinese (Ho et al., 2003).

Such phonetic ambiguity in characters indicates the complex relation between Chinese oral language and its writing system, presenting special challenges in learning Chinese script. Historically, Chinese writing is said to have been moving in the direction of characters having more and more phonetic indicative powers in the use of phonetic radicals (Zhou, Z. 2004), moving from “read-like” whole character approach in ancient Han time (206 BC–CE 220) to Fan Qie (onset-rime approximation) in Wei-Jin times (CE 220–280) (see Zhou, Z. 2004), to various Romanization efforts in modern time (Ichikawa and Komatsu, 2008). For Chinese, efforts to increase phonological transparency are expressed mainly by adding phonetic radicals to existing characters (Zhou, Z. 2004). In

other words, newly-minted characters tend to increase phonological information. But even such attempts using embedded phonological approximation over times come at a cognitive cost for learners. Since Chinese characters are stand-alone orthographic and syllabic units, any added phonetic radicals will increase the complexity within characters, making such phonetically enhanced characters more complex and consequently difficult to learn. We will return to this in the next paragraph. For now, we want to point out that the practice of adding phonological information has been a welcome effort in facilitating character recognition (Ho et al., 2003), and has seemed to have been adopted in modern China's elementary textbooks (Shu et al., 2003). But at the same time, thanks to the inconsistency of the practice as reflected in existing characters, such phonetic radicals, carrying different phonological information of different times due to oral language changes and other deviations, makes it not that transparent as its phonetic radicals would promise (Zhou, 2009), making phonological access through these phonetic radicals always an analogical approximation.

A related challenge in pursuing phonetic transparency through added radicals, both phonetic and semantic, is the increased orthographic complexity, an undesirable phenomenon for any script but particularly troubling for Chinese. Chinese characters act visually like English words, written with spaces to signal a separate character. However, for English, a word's visual complexity is usually expressed by the number of letters and syllables in it (Fry, 1989). The physical space a word occupies therefore varies, depending on the number of letters, and reflects that complexity. In contrast, Chinese characters use strokes, radicals in their formation, and always take up a similar square space visually. Its visual complexity is dependent on the number of strokes and radicals. For example, characters 日 (meaning "sun," with 4 strokes) and 曙 (meaning "dawn," with 17 strokes or two radicals) takes up an equal space in writing. The complexity is within a given space of any characters rather than without. This implies that strokes and radicals will change their sizes to fit into their positions in a character, a different approach than English in handling its visual complexity. Researchers have found that this complexity within a character must have had a reverse effect on learners' lexical processing in reading (Qian et al., 1994), and was certainly affecting the rate of writing as argued by Youguan Zhou (2004). Therefore, such a practice of moving toward phonetic transparency by adding radicals to create new characters is likely to add to the complexity and difficulty of characters, particularly challenging for beginners who are learning to pay more attention to the radicals within characters (Ho et al., 2003; Tong et al., 2017).

This, of course, has created more characters for the corpus of Chinese language, a phenomenon partially attested to by the steadily expanding dictionaries over times (Chen, 1999; Wilkinson, 1998). A quick and brief listing of the characters available in four major milestone dictionaries in China would give us some sense about the increase of characters over times. First Chinese dictionary *Shuo Wen Jie Zi* at about CE 100 contains 9,353 characters (Su, 1997). *Lei Pian* (Dictionary of Characters Categorized) compiled in 1066, following *Shuo Wen Jie Zi*'s format, has 31,319 characters in it (Liu, 2004). *Kang Xi Zi Dian* (Dictionary Compiled in the Kang Xi Period) of 1716, as one of the most consulted dictionaries for the past two hundred some years, lists 47,035 characters (Liu, 2004). The modern Chinese comprehensive dictionary *Han Yu Dai Zi Dian* published in 1990 includes 54,678 characters (Yin and Wang, 2008). The incremental accumulation of Chinese characters over time shows a clear upward trajectory, making any attempt to acquire the available characters a tremendous and literally impossible cognitive challenge.

Answers to the challenges embedded in Chinese orthography

In this section, we are focusing on the efforts to approach the afore-mentioned challenges to script acquisition posed by Chinese orthography, namely, the Pinyin adoption to make Chinese characters phonetically accessible, character simplification to reduce the orthographic complexity of some characters, and a constant number of characters in use to avoid the large number of existing characters for acquisition.

Pinyin, the latest in the series of Romanization of the Chinese writing system, is an officially endorsed system developed to index Chinese characters in roman alphabets. Unlike its ancient predecessors of "read-like" and Fan Qie, Pinyin moves in the direction of Romanization, literally meaning "spell the sound." It is premised on a common language, or the standard pronunciations (Zhou, Y. 2004), and, in turn, promotes the school use of Putonghua. The Pinyin system uses 25 Roman alphabetic letters without the letter "v" to represent 56 initials and finals (see Table 2) in characters according to the common language—Putonghua, the Mandarin dialect. As a sound indexing system, Pinyin provides a direct access to the phonological information of characters in Putonghua pronunciations. For example, "dou" and "liao" are directly spelling the sounds of two Chinese characters: 斗 and 料. Pinyin use, therefore, could alleviate the burden of remembering the unexpressed phonetic values of these two Chinese characters, reducing the phonetic ambiguity of Chinese characters. It is for this reason that the Pinyin system is now used in almost all beginning literacy learning texts and dictionaries in Mainland China (Zhang et al., 2000). All learners will move from Pinyin to characters in a rapid fashion (Si, 1982), a practice that has been found to have a sustained early start on reading and writing (Liu, 2005). But Pinyin in its current application does become an effective assistive means to help connect the standard oral language—Putonghua—with Chinese script for beginning learners (Lin et al., 2010). In this assistive spirit, even moderately extended period of using Pinyin in upper primary schools was found to have an encouraging effect when used together with characters (Liu, 2005). Yet, this solution to the challenge of phonetic ambiguity in the characters is not perfect mainly due to the homophonic phenomenon in Chinese script, which stands at an average estimate of somewhere between 6 and 40 homophones for each existing syllable in modern Chinese. The ambiguity would not be resolved without incurring the assistance of actual characters (Zhou, Y. 2004), a tendency literally having helped retain some orthographic complexity of characters as well as convincing the users and policy-makers alike that Pinyin would remain as a sound-indexing system rather than the written language.

Table 2 Twenty-five letters^a in Pinyin for 21 initials and 35 finals in Chinese^b.

25 letters							
a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, w, x, y, z							
21 initials							
b	p	m	f	d	t	n	l
g	k	h		j	q	x	
zh	ch	sh	r	z	c	s	
35 finals							
	i	u	ü				
a	ia	ua					
o		uo					
e	ie		üe				
ai		uai					
ei		uei					
ao	iao						
ou	iou						
an	ian	uan	üan				
en	in	uen	ün				
ang	iang	uang					
eng	ing	ueng					
ong	iong						

^aThe letter v is not present in the Pinyin system. In addition, a two-dot diacritic-added “ü” is an additional vowel which will change to the letter “u” when preceded by initials “j,” “q,” “x” and “y.”

^bThese 56 initials and finals cover all initial sounds and final sounds in Chinese characters but some rule-based changes will occur in actual onsets and rimes. Such rule-based changes would call in semi-vowels “y” and “w” as onsets for some medials and finals in specified contexts, and would also yield several rimes not listed in the finals in specified contexts. Refer for the specifics to [Institute of Linguistics CASS \(1979\)](#), pp. 1565–1567.

Modified from source [Institute of Linguistics CASS \(1979\)](#), pp. 1565–1567).

The orthographic complexity of characters has always been the challenge for learners and users. As pointed out, the complexity comes from its character formation approaches that continuously call for compounding characters through additional phonetic and semantic radicals. Such within-character complexity, while adding some phonetic consistency and semantic directions to characters, exacts a cognitive demand on the learners, forcing them to slow down to encode the complexity for character recognition. To counter such a challenge of increasing character complexity, Chinese script has traditionally adopted scriptural simplifications: script style changes and character simplifications. Script style changes occurred over long periods of time and were usually gradual, moving from archaic script styles with pictographic resemblance to ideographic abstraction, particularly expressed with streamlined radicals and straight and un-curved strokes ([Qiu, 2009](#)). The modern-looking clerical script style first emerged in popular use more than two thousand years ago ([Yin and Wang, 2008](#)). Script style changes (see [Table 3](#)) usually facilitate writing and should also enhance character recognition for readers. Character simplifications efforts meanwhile were seen throughout the history, including the most recent ones in the Mainland China that occurred in 1956 and 1964 simplifying 2236 frequently used characters all told by reducing their average number of strokes from 10.98 to 8.94, about a 20% reduction ([Zhou, Y. 2004](#)), a significant drop in the

Table 3 Chinese character script changes over history.

Character definitions	Oracle bone script samples (late Shang, West Zhou)	Small seal script samples (Warring States, West Han)	Clerical script samples [4th century BC(Qin)—4th c. CE]	Traditional script samples (modern time)	Simplified script samples (modern time)
Cowrie					
Vehicle					
Horse					
Tooth					
Music					

The ancient time periods are based on [Wilkinson \(1998\)](#).

complexity inside these characters. They were received by the society and many are still continued in current usage. However, the solution efforts to reduce the complexity of characters through simplification have also made it abundantly clear that simplification is not always a result of top-down decision, be they scholarly or administrative. They are usually based both on practice and on the script's internal features. For example, from a practice-based perspective, a second attempt of character simplification eventually failed after 1981 because that batch of 853 characters were too esoteric and not in circulation already (Qiu, 2009; Zhou, Y. 2004), indicating the power of popular use in the writing system development. Youguan Zhou (2004) has also pointed out that the internal features such as the informational load expressed by character complexity would play a role in eventual popular acceptance or rejection of a character, because an oversimplified character would lose its original information in the eyes of mature users (pp. 357–360). It is also likely that in the computer age, character complexity would be less burdensome in character recognition and production than before when computer input of Pinyin alone or coupled with character features allows for enhanced early ability of character recognition and writing (Liu and Zhang, 2006).

As to the issue of the ever-increasing number of characters, scholars have noted that the Chinese language itself has developed a practical response to it: rotating out the dated and obsolete characters, and retaining and using only a constant range of characters (Zhou, Z. 2004). In modern China, for example, estimates about characters in use from common entries in ordinary dictionaries, print-needed characters, and character frequency studies point to about a range of 5,000 to 6,000 common characters in addition to some extra characters that are rare, discipline-specific and seldom used. The addition of the latter might top out the modern character numbers at 9,000 at most (Zhang et al., 2000). For example, in a large-scale selection of frequently used characters sponsored by China's National Language Working Committee, researchers extracted the 2,500 highest frequency characters and the next 1,000 highest frequency characters, from a list of 15 publications of frequently used characters over a period of 58 years, and from the five other sources of commonly used characters, and then subjected these 3,500 characters to a two million character sample from three popular sources of newspapers, popular science magazines, and literature magazines. The most frequently used 2,500 characters have been found to have covered 97.97% of all characters used and the additional 1,000 have increased the coverage to 99.48% of all characters in the whole sample (see Appendix V in Yin and Wang, 2008, pp. 301–316), demonstrating the utility of these three thousand five hundred characters for being literately functional in the modern society. Students in Chinese schools are required to reach the similar range of numbers of characters for literacy levels: recognizing 3,000 characters at the end of elementary school and no fewer than 3,500 characters for finishing the whole nine years of compulsory education (Zhang et al., 2000). Knowledge of five to six thousand characters would qualify for advanced literacy in modern China (Wilkinson, 1998). There are some indications that historically the required characters have never been exorbitantly higher than the current range of number of characters (Wilkinson, 1998; Yin and Wang, 2008; Zhou, Y. 2004).

The dichotomy of increasingly more characters in dictionaries and a constant range of characters in use is showing a very interesting picture, particularly if Chinese character growth is claimed as going hand in hand with new concepts and expressions in the society (Huang and Meng, 2001). There is a fundamental reason for the number of characters to be kept within an adequate range. It comes primarily from the power of word formation by characters. To understand this, we need to understand the relationship between characters and words in Chinese. Linguists regard characters as morphemes, not words, because most of the meaningful expressions in modern Chinese are of multi-character (Wilkinson, 1998), thus making characters behave linguistically more like morphemes than words. In such multi-character words, characters serve as a meaningful component or a morpheme (Chen, 1999; Packard, 2015). For example, the expression “today” in Chinese is 今天 that contains two meaningful characters: “now” and “day.” They are seldom used alone as two stand-alone words now as they usually did in archaic Chinese (Qi, 1997). Instead, they usually combine with other characters to form related words such as 今晨 (this morning), 今昔 (now and then), 天桥 (overpass), or 天气 (weather). Therefore, in their forming the modern words, they behave more like compound words in English with free morphemes. Moreover, treating characters as morphemes overlooks the internal complexity we have shown when discussing the compound characters that carry semantic radicals. Take the compound characters 晨 and 桥 or simple characters 今 and 气 for examples. The first two contain radicals that can stand alone as simple characters such as 日 (sun), 辰 (celestial bodies), 木 (wood), 乔 (tall), and the last two can be combined into compound characters, with additional radicals, such as 矜 (pity) and 氡 (radon). These examples do not merely show the inadequacy of simply using an alphabet-based linguistic concept in Chinese without modifications, they also at the same time highlight why characters are so powerful in forming new words (Wilkinson, 1998). For example, all 626 newly coined expressions in 2010 in fact consist only of existing characters (see Hou and Zhou, 2011). Therefore, the power of characters to form words and expressions is tremendous, contributing to the constant range of characters required over times, and thereby, together with rotating out the obsolete characters and rotating in or creating new characters, providing a solution to the cognitive challenge of the increasing number of characters.

Writings and language beyond orthography: Wenyan Wen and Baihua Wen

We would be amiss if we focus only on the Chinese script and its features and relationship to oral language. Writing as an intellectual and social practice play a very important role too in how oral language is functionally captured in words. We will briefly talk about an archaic but enduring type of writing Wenyan Wen (literary language writing) in China and the reformed writing Baihua Wen (plain spoken language writing) in the current society. The intricate relationship between the written language and spoken language has clearly reflected the social, political, and cultural roles attributed to literacy.

Wenyan Wen had dominated Chinese writings ever since Han time, allegedly modeled after the Classic writings of Confucius's times (Chu, 1973). For two millennia, it was the official writing form for all government and court uses, the standard writing format adopted by men of letters, and the expected writing outcome of school educations. While it has multiple characteristics that distinguish it from the current writings (Li, 2011), the main feature relevant for our discussion here, *inter alia*, is its drastic separation from oral language (Mair, 1994). It does so not by adopting a different script as Latin in the case of Medieval Europe, but by using a way of writing, in the same script, that has very little resemblance to spoken language. While it is by no means a static and unchanging way of writing (Li, 2011), its distance from spoken language only grows larger over time. Given the relative quick-paced changes in oral language in terms of phonology, words, phrases, and grammar, Wenyan Wen did not keep up with the changes in new ways of expressions (Cao, 1952/2012), particularly in its continuous use of the main staples of empty (non-substantive) words/characters, archaic prepositions, pronouns, and sentence structures (Li, 2011). It retained its short, mostly single-character words, usually well-balanced four-character phrases, and terse sentences. Due to its distance from oral language, Wenyan Wen always required special and time-consuming training (Cao, 1952/2012). It is in this sense that Wenyan Wen was functioning as an additional layer that complicated the literacy situation in China in addition to the ideographic script and dialectal diversity. It should be pointed out that Wenyan Wen eventually lost its popularity in main-stream Chinese writings and trainings primarily due to this added complication to Chinese literacy acquisition (Zhou, 2009). In fact, it was called a dead "ancient language" when 20th century dawned (Zhou, 2009, p. 13).

The pronounced reason for Wenyan Wen to lose its traditional dominance is the national awareness of its obstruction to the popular literacy promotion and consequently its drag on the national modernization at the beginning of 20th century in China (Hu, 2005). In its stead, emerged Baihua Wen, writing that follows the spoken language (Zhou, Y. 2004). As a conscious counter measure, Baihua Wen allowed the learners to access the spoken language without having to go through the contrived structures and short dated single-character expressions, therefore presumably facilitating reading and writing processes. It should be noted that such shifts to oral language discourses have occurred continually through Chinese history, for example in the newly developed literary genres such as translated Buddhist literary narratives in the Tang time, plays in the Yuan time, and fictions in the Ming time (Cao, 1952/2012). The Baihua Wen promotion at the beginning of the 20th century was just more conscious as a national push for education reform and social and technology reform (Chen, 2014), and consequently replaced Wenyan Wen as the dominant writing for literature, textbooks, and newspapers (Hu, 2005). The solution it offers to deal with the exorbitant time spent to master Wenyan Wen has been obvious, making literacy more available to a broad populace (see Table 4). However, even though Baihua Wen is now the primary written language used in modern China, Wenyan Wen's influence and existence still persist. Not merely texts in Wenyan Wen is part of advanced Chinese curriculum in middle schools and beyond, numerous characters and idiomatic expressions from Wenyan Wen have retained their presence in Baihua Wen. In creative literature writings such as poems and essays, Wenyan Wen has been strongly argued for its existence because of its beauty, elegance, stately forms, and richness of cultural heritage (Chen, 2014; Qi, 1997).

As part of the national discourse about literacy, what Wenyan Wen and Baihua Wen have demonstrated is their historical role in literacy acquisition and literacy application. They show their vitality and traction in Chinese learning and use as tools for different utilities. Baihua Wen provides both a context for making a connection with oral language and a literate mentality of the populace, providing an alternative, though slighted by the learned circles throughout history, to the aloof style of Wenyan Wen, and emerging as a political answer as well as a practical means to dominate the writings when the country needs a new approach to solving its political crisis (Cao, 1952/2012; Lu, 1934/1973). Wenyan Wen, on the other hand, tends to have its own resilience in a national conversation about the utility of cultural heritage and hold its appeal as part of a needed diversity when the society is moving forward (Chen, 2014). In this sense, both are literacy tools with cultural, social, and political implications, called forth to perform their duties beyond mere pedagogical and practical considerations. Their parallel trajectories through history and the eventual succession in dominance in the 20th century, bring out the political, social, cultural, and educational shifts, reminding us again of the intricate relationship among the various functions of literacy.

Literacy acquisition in Chinese: highlighted features and associated research efforts

What stands out most of Chinese literacy acquisition is not that it is different from any alphabetic language acquisition. It requires knowledge of its writing script and its successful applications in readings and writings. But what is different from literacy acquisition of alphabetic languages is also clear: it stems from the nature of its ideographic script, and from how it has traditionally been used in the society and culture. In short, it comes from the unique distance between its writing script and spoken language, between its traditional written language and spoken language, and its social and cultural applications. These differences have been largely contributing to the complex literacy situation in China, and also to a great extent been the driving forces that move both literacy acquisition and research efforts. The literacy education in China tends to show the following features pertaining to overcoming these distances and some deep-rooted social practices.

First, literacy education usually starts with the common language and Pinyin (Zhou, Y. 2004). While the common language is still an acquired skill for many kids due to dialect situations, the widespread TV, radio, and other public media as well as current convenient means of traveling across the country have provided some solid experiential ground for the common language acquisition. Yet, in areas that Putonghua is not popularly used such as Hong Kong, learning Putonghua could be a challenge for children (Tong and Goh, 2008). Properly mastering Putonghua pronunciations should assist Pinyin learning, and in turn help

Table 4 Illustrations of some main differences between Wenyan Wen and Baihua Wen.

	Differences	Wenyan Wen	Samples (meaning or function)	Baihua Wen	Samples (meaning)
Word level	Characters as words	Usually single character words	民2 (people) 虎3 (tiger)	Two or multi-character words	人2民2 (people) 老3虎3 (tiger)
	Main meaning changes	Archaic (not available in modern writing)	微1 (if not) 夫2 (modal word)	Current main meaning (may exist in the past)	微1 小3 (tiny, small) 丈4夫1 (husband)
	Changes of word forms (personal pronouns as examples)	Archaic characters (few still used in modern Baihua Wen writing)	吾, 我, 尔, 余, 台, 朕, 汝, 尔, 若, 乃, 彼, 其, 或, 某, etc.	Limited and mostly different from the archaic ones	我, 你, 您, 他, 她, 它, 我们, 你们, 他们, 她们, 它们, etc.
Phrase level	Formulaic vs. free flow	Four-character phrases	悬象设教 (using embodied images in nature to educate and cultivate people)	Free flow expressions	利用天地的征象来教化民众 (This is a Baihua Wen translation of the Wenyan version)
Sentence level	Formulaic vs. free flow	Parallel sentences and phrases	神莫大于化道, 福莫长于无祸。 (Spiritual cultivation reaches its highest level when a person understands and is one with the Dao; happiness is at its best when a person can keep away from misfortunes)	Free flow expressions	最高的精神修养就是把道融入自身, 最好的幸福就是没有灾祸 (This is a Baihua Wen translation of the Wenyan version)
Grammar	Function words usage (auxiliary words at the end of an interrogative or exclamative sentence)	Plenty of archaic function words (belonging to empty/non-substantive words)	也, 矣, 乎, 诸, 哉, 耳, 焉, 邪, etc.	Limited uses of different function words	吗, 呢, 啊, 呀, etc.
	Complements	Complements can be omitted	民扶老携幼, 迎君 [于] 道中 (people came out with their elders and youngsters on the road to welcome the Lord)	The omitted complement is turned into an adverbial phrase "on" the road.	老百姓搀扶着老人, 携带着孩子, 在路上迎接孟尝君。 (this is a Baihua Wen translation of the Wenyan version)
	Object positions	Could be before verbs	不吾知 (do not know me)	Could not be before verbs	不知道我 (Baihua Wen translation)

NOTE: These characteristics are not absolute nor are they exhaustive. However, their combined effect is quite obvious in all the samples here: oral language-based and therefore lengthy Baihua Wen versus classic written language-based and terse Wenyan Wen.

Numbers in Table 4 designate one of the 4 tones in Mandarin Chinese.

children with early start of reading and writing (Liu, 2005), particularly for children from other dialect zones than that of Mandarin to develop Putonghua-based sensitivity to characters and expressions (Zhang et al., 2000). Second, Chinese text read-aloud has always been a salient practice of literacy acquisition in schools. Such practice is highly promoted in schools to assist the common language mastery, acquisition of the phonological values of characters, and building fluency in connecting writing with oral language (Zhou, 1999). It is also a traditionally endorsed and promoted practice in China (Tao and Zuo, 1997), not merely persisting for more than 2,000 years in Chinese literacy (Chi, 1998), but also is believed to be an effective way to experience the unfamiliar language elements and styles (Zhang, 1997), thereby building foundations for continuous reading at advanced stages where existing Wenyan Wen's proportion increases (Ye, 1980). Almost all texts still contain the requirement for recitation of certain texts as a push for text-familiarity (e.g., Shanghai Elementary and Middle School Textbook Reform Committee, 2000).

Third, based on an ideographic script, Chinese literacy learning has, from its inception, relied on using its available relationship between semantic and phonetic clues in its orthography to learn characters. Research has supported teaching children through exploring radicals in compound characters (Anderson et al., 2013). Such teaching practice hinges on maximizing the semantic categorical information and phonetic approximation information in characters, therefore encouraging analogical as well as reflective approaches to learning. This apparently is not too different from what Xu Sheng did 2,000 years ago with the explications of six principles of character formations (Su, 1997), allowing learners to make use of the structural components of characters to help alleviate the difficulty caused by the oral and written language distances. Indeed, in order to further connect the oral and written language, characters are also learned in meaningful contexts of words and texts, making students aware of the power of characters in forming meaningful words used in modern Chinese. Therefore, scholars recommend teaching characters for their generalization and teaching characters in the context of words and phrases, and the latter in the context of texts (Li, 1997; Tong and Zhang, 1999).

Finally, as writing units that do not bear explicit consistent phonetic symbols in their formation, characters require sufficient practices to be grasped (Si, 1982). For early literacy learners, learning to write characters usually goes along with learning to read characters, starting with tracing characters in the order in which strokes are written. This can help learners to grasp where to position the strokes within a given character in order to write it correctly, and helps with character differentiation and recognition via stroke pattern knowledge (Chen et al., 1996). In addition, knowing how to follow the sequence of strokes could facilitate character writing, not merely in speed as in cursive style writing in English, but also ensuring its procedural accuracy and esthetic presentation (Yin and Wang, 2008). Most of the characters students are required to learn to recognize are also required writing characters: two thousand five hundred characters out of three thousand (China's Department of Education, 2000). Such imitation writing practice is always accompanied by and extended to both formal and informal writing practices (Si, 1982; Zhang et al., 2000), supported by Pinyin use at the very beginning stage of learning (Liu, 2005). In recent times, computers have played an active role in both promoting students' early learning of reading and writing (Liu and Zhang, 2006).

As you may have already gleaned from the above examination of the Chinese written language, efforts, of research or praxis, in literacy acquisition have always been about the nature of the script and how to best bring out the power of the written language in its acquisition. In the history of Chinese literacy development, efforts were expended on simplifying characters from the archaic form to its modern-look characters with its typical straight-lines and regularly squared size no later than in the Warring States periods (475 BC–221 BC) (see Xu, 2009). Precursors to both modern efforts of Pinyin and Putonghua to connect the written language to an oral language had also surfaced then in the form of direct "sound-borrowings" (Cao, 1952/2012) and Yayan (Wilkinson, 1998). The differences between the intent of writing, written language (wen), and oral language (yan) had also been highlighted and debated in Mo Zi and Confucian works (Hu, 2005), projecting an early awareness and sensitivity in the cultural practices of the distances between oral and written language (Tao and Qian, 2012b). When Xu Sheng at around CE 100 came out with the explicit description of "loan characters" as a principle of character-formation (borrowing via sound-similarity), it highlighted the cultural awareness toward that yawning chasm between the written characters and oral language in reality (Xu, 121/1989). The history of Chinese literacy development afterward for about two thousand years has always reflected such understandings of the difference between oral and written expressions, from Han Dynasty's (206 BC–CE 220) application of the clerical character style, Weijin and NanBei Dynasties' (CE 220–581) broad adoption of Qieyun (onset-rime approximation) approach, Tang's (CE 618–907) popular oral style literary narratives, to the modern time efforts in Romanization, character simplification, and Putonghua (Zhou, Y. 2004). While most of these efforts, particularly the ancient ones, are indigenous in origin, they have often been informed by language influence from outside China, such as Sanskrit's influence on Qieyun and Buddhist translation on Tang's popular literary writings (Hu, 2005). As a cautious side note, it should also be added that centralized efforts from top down also seem to have a desired effect on literacy growth only if they are measured and rooted in practices as in the successful promotions of Pinyin, simplified characters and Putonghua. This persisting tension in literacy development created by distance between the oral language and the script, and between the popular but "vulgar" style and exclusive but aloof elite style must have reflected a dynamic vitality of the Chinese language to sustain its own growth by accommodating the input from outside and adjusting itself to meet the needs of social progress (see Huang and Meng, 2001; Yin and Wang, 2008).

The research efforts in the last quarter of the 20th century and first twenty years saw a steady and exciting rise of research into various cognitive, linguistic, and pedagogical aspects of Chinese written language acquisition (See Anderson and Chen, 2013), yielding much needed fine-grained understanding of many processes of Chinese acquisition as our preceding descriptions have shown. The focus of the research is primarily on the role characters play in literacy acquisition, and on the perceived distance between characters and oral expressions, offering insights into the phonological and morphological functions in the orthography of characters. While many areas of this line of research are still being explored, elucidated and discussed, there is no doubt that the current research, following the Western traditions of cognitive and psycholinguistic theories, has provided and continues to provide the necessary insight into our understanding of written languages as well as into pedagogical approaches catering to literacy education in general.

Conclusion: what we know and need to know more

The picture we have painted of Chinese literacy is straightforward. As an ideographic writing system, Chinese script shows its non-alphabetic peculiarities in emphasizing the converging clues from orthography, phonological, and semantic components in character acquisition, highlighting both the utility and accompanying uncertainty of using an opaque approach of analogy in the process. Over its history, the development of the Chinese writing system has reflected writing's relations to its oral language, accentuating the important roles social and political factors play in coordinating the relations. At the same time, it is clear that the ideographic system itself has also been open to efforts both from within and without to make its phonological connections with spoken language. Yet the degree of the phonological transparency in its script is not an absolute measure of its successful acquisition. In retaining its ancient ideographic nature, the Chinese writing system shows its ability to conform but also to persist, making such interventions as Pinyin and Putonghua effective only in the context of its fundamental characteristics and practical precedence.

Yet, many areas of Chinese literacy development and acquisition are still unclear. Two areas that call for research attention are highlighted here. The first is the relationship between literacy acquisition by learners outside Mandarin dialect zones and literacy acquisition by learners of Mandarin. Given the different distances of their oral dialects to the written language, we

would like to understand how and if the literacy acquisition is mediated by the dialects, particularly in the context of a common language. Our knowledge on this is still lacking even though most of the common language efforts and alphabetic push are partially based on the tacit assumptions of the obstructive distance between the oral and written language in dialect zones. Most of the studies on literacy have been done with either Mandarin or Yue dialect students (e.g., McDowell and Lorch, 2008; Xu, 1991; Yip, 2014), shedding some light on character acquisition with specific dialect features. However, systematic studies are still to come with regards to the possible interactions between dialects and the common language for literacy acquisition. A practical question of the dialect situation to literacy acquisition from a developmental angle seems to come naturally from the current parallel push for a common language and dialects, the latter for fear of losing local cultures (Erbaugh, 1995; Ichikawa and Komatsu, 2008). Without understanding how dialects interact with a common language in literacy acquisition, policy-makers' stances on literacy in dialect zones may be more persuaded by political and cultural affiliations than informed by literacy acquisition purpose, having implications both for beginning literacy acquisition in dialect zones and for popular literacy in general.

The second area that needs more attention is that of historical literacy research, including the early Chinese literacy. We have seen some efforts made in such directions (see Li and Branner, 2011; Tao and Reinking, 2020; Wang, 2020) providing some promising and innovative steps taken to explore early Chinese literacy practices. But due to the documents and records limitations pointed out by Li and Branner (2011), studies on literacy's past are still few and far between. Up to now, the most acknowledged studies on ancient Chinese literacy are still Rawski's 1979 *Education and Popular Literacy in Ch'ing China* in English, and Zhang Zhigong's 1964's book (1992 reprint) *On traditional Chinese education and texts with a bibliography of character texts and book reprints* in Chinese. It is obvious that our current knowledge of historical Chinese literacy is far from satisfactory. Of the numerous literacy areas usually explored in the western world (e.g., Heath, 1996; Venezky, 1996), Chinese historical literacy research is still to come. For example, we mentioned briefly some challenges and solutions surrounding the Chinese script development in China's history. But comprehensive studies are needed to examine from a literacy acquisition perspective what historical, linguistic, and socio-cultural factors interacted with the Chinese script to shape its continuous development on an ideographic path in spite of numerous encounters with cultures with alphabetic languages. Such an understanding will not only provide historical insight into a very important piece of literacy history, but would also inform our pedagogical practices and policy-making in current literacy education to cater to the nature of the script as well as to the accompanying social needs.

To conclude, the distances between Chinese written and oral languages will continue to play out in Chinese literacy acquisition and applications. Literacy acquisition and practices will continue to pose new challenges not merely about its writing system, but also raise questions about our theoretical frameworks through which we assess its uniqueness and idiosyncrasy and about our ability to influence its direction. In this sense, literacy acquisition and applications are never simply the uses of a writing system as a tool. They are also the powerful forces behind the vicissitudes of its change and the current form it takes.

Acknowledgments

My sincere thanks to Dr. Gene Fellner, Dr. David Kritt, and Dr. Li Zuo who have provided insightful comments and helpful suggestions to the several drafts and the final version of the current chapter.

References

- Anderson, R.C., Chen, X., 2013. Chinese reading development in monolingual and bilingual learners: Introduction to the special issue. *Sci. Stud. Read.* 17 (1), 1–4.
- Anderson, R.C., Ku, Y.-M., Li, W., Chen, X., Wu, X., Shu, H., 2013. Learning to see the patterns in Chinese characters. *Sci. Stud. Read.* 17 (1), 41–56.
- Cao, B., 1952/2012. *Zhong Guo Wen Zi De Yan Bian* [Evolution of Chinese Written Language]. Capital Economy and Commerce University Press, Beijing.
- Chang, R., Chang, M., 1978. *Speaking of Chinese*. Norton, New York, NY.
- Chen, Y., Allport, D.A., Marshall, J.C., 1996. What are the functional orthographic units in Chinese word recognition: the stroke or the stroke pattern? *Q. J. Exp. Psychol.* 49 (4), 1024–1043.
- Chen, P., 1999. *Modern Chinese: History and Sociolinguistics*. Cambridge University Press, New York, NY.
- Chen, J., 2014. *Lun Wen Yan Yu Bai Hua De Bian Zheng Guan Xi Ji Zhong Guo Xiang Da Wen Xue Zhi Yuan—Yi Zhou Sou Juan Wei Li* [on the dialectic relationship between Wenyan and Baihua and on the origin of Chinese Contemporary literature: a case study of Zhou Sou Juan]. *Qin Hua Chinese J.* 12, 309–371.
- Chi, X., 1998. *Zhou Guo Gu Dai Xiao Xue Jiao Yu Yan Jiu* [A Study of Elementary Education in Ancient China]. Shanghai Education Press, Shanghai.
- China's Department of Education, 2000. *Jiu Nian Yi Wu Jiao Yu Quan Ri Zhi Xiao Xue Yu Wen Jiao Xue Da Gong, Shi Yong Xiu Ding Ban* [Nine Year Full-Time Education Chinese Language Arts Curriculum: Elementary Schools, Revised Trial Edition]. People's Education Press, Beijing. Retrieved at: https://old.pep.com.cn/xiaoyu/jiaoshi/tbjx/kbjd/jxdg/201008/t20100818_663520.htm.
- Chu, Y., 1973. The Chinese language. In: Meskill, J. (Ed.), *An Introduction to Chinese Civilization*. Columbia University Press, New York, NY, pp. 587–615.
- Enri, L.C., 2014. Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18 (1), 5–21.
- Erbaugh, M.S., 1995. Southern Chinese dialects as a medium for reconciliation within Greater China. *Lang. Soc.* 24 (1), 79–94.
- Fry, E., 1989. Reading formulas—maligned but valid. *J. Read.* 32 (4), 292–297.
- Hansen, C., 1983. *Language and Logic in Ancient China*. The University of Michigan Press, Ann Arbor, MI.
- Heath, S.B., 1996. The sense of being literate: historical and cross-cultural features. In: Barr, R., Kamil, M.L., Mosenthal, P., David Pearson, P. (Eds.), *Handbook of Reading Research*, vol. II. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 3–25.

- Ho, C.S.H.H., Ng, T.T., Ng, W.K., 2003. A "radical" approach to reading development in Chinese: the role of semantic radicals and phonetic radicals. *J. Literacy Res.* 35 (3), 849–878.
- Holm, A., Dodd, B., 1996. The effect of first written language on the acquisition of English literacy. *Cognition* 59 (2), 119–147.
- Hooker, J.T., 1990. Introduction. In: *The Trustees of the British Museum, Compiled, Reading the Past: Ancient Writing Form Cuneiform to the Alphabet*. Barnes & Noble, Inc., New York, NY, pp. 6–13.
- Hou, M., Zhou, J. (Eds.), 2011. 2010 Han Yu Xin Ci Yu [New Chinese Words and Phrases of 2010]. The Commerce Press, Beijing.
- Hu, Q., 2005. *Zhong Guo Xiao Xue Shi* [A History of Chinese Philology]. Shanghai People's Press, Shanghai.
- Huang, Y., Meng, H., 2001. *Han Zi Fu Hao Xue* [Semiotics of Chinese Characters]. Shanghai Ancient Books Publishing House, Shanghai.
- Ichikawa, I., Komatsu, A., 2008. *Bai Nian Hua Yu* [Chinese in the Past Hundred Years]. Shanghai Education Press, Shanghai.
- Institute of Linguistics CASS, 1979. *Xian Dai Han Yu Ci Dian* [Modern Chinese Dictionary]. The Commerce Press, Beijing.
- Katz, L., Frost, R., 1992. Reading in different orthographies: the orthographic depth hypothesis. In: Frost, R., Katz, L. (Eds.), *Orthography, Phonology, Morphology, and Meaning*. Elsevier North Holland, Amsterdam, the Netherlands, pp. 67–84.
- Li, F., Branner, D.P. (Eds.), 2011. *Writing and Literacy in Early China*. University of Washington Press, Seattle, WA.
- Li, B., 1997. *Li Bing De Jiao Yu Wen Xuan* [Selections of Li Bingde's Writings on Education]. Educational Science Publishing House, Beijing.
- Li, Z., 2011. *Gu Han Yu Yu Fa Xue Shu Lie* [A Survey of the Grammar of the Ancient Chinese]. Jilin University Press, Changchun, Jilin.
- Lin, D., McBride-Chang, C., Shu, H., Zhang, Y.P., Li, H., Zhang, J., Aram, D., Levin, I., 2010. Small wins big: analytic pinyin skills promote Chinese word reading. *Psychol. Sci.* 21 (8), 1117–1122.
- Liu, Y., Zhang, D., 2006. ICT and Chinese literacy education: recent development. In: McKenna, M.D., Labbo, L.D., Kieffer, R.D., Reinking, D. (Eds.), *Int. Handb. Liter. Technol.*, vol. 2. Erlbaum, Mahwah, NJ, pp. 193–210.
- Liu, Y., 2004. *Zhong Guo Zi Dian Shi Lie* [A Summative History of Chinese Dictionaries]. Zhonghua Book Company, Beijing.
- Liu, Y., 2005. A pedagogy for digraphia: an analysis of the impact of pinyin on literacy teaching in China and its implications for curricular and pedagogical innovations in a wider community. *Lang. Educ.* 19 (5), 400–414.
- Lu, X., 1934/1973. *Meng Wai Wen Tan* [An outsider's comments on writings]. In: Chinese Departments of Fudan University and Shanghai Normal University (Ed.), *Lu Xun Zai Wen Xuan* [Selected Essays by Lu Xun]. Shanghai People Publishing House, Shanghai, pp. 122–141.
- Mair, V.H., 1994. Buddhism and the rise of the vernacular in East Asia: the making of national languages. *J. Asian Stud.* 53 (3), 707–751.
- McDowell, H.J., Lorch, M.P., 2008. Phonemic awareness in Chinese L1 readers of English: not simply an effect of orthography. *TESOL* 42 (3), 495–531.
- Norman, J., 1987. *Chinese*. Cambridge University Press, Cambridge.
- Packard, J., 2015. Morphology: morphemes in Chinese. In: Wang, W.S.-Y., Sun, C. (Eds.), *Oxford Handbook of Chinese Linguistics*. Oxford University Press, New York, NY, pp. 263–273.
- Qi, G., 1997. *Han Yu Xian Xiang Lun Cong* [Essays on Chinese Language]. Zhonghua Book Company, Beijing.
- Qian, G., Reinking, D., Wang, S., 1994. The effects of character complexity in recognizing Chinese characters. *Contemp. Educ. Psychol.* 19 (2), 155–166.
- Qiu, X., 2009. *Wen Zi Xue Gai Yao* [An Outline of Chinese Philology]. Commercial Press, Beijing.
- Santana, I.S., 2000. Literacy research in Latin America. In: Kamil, M.L., Mosenthal, P., David Pearson, P., Barr, R. (Eds.), *Handbook of Reading Research*, vol. III. Lawrence Erlbaum Associates, Mahwah, NJ, 41–25.
- Schirokauer, C., 1989. *A Brief History of Chinese and Japanese Civilizations*, second ed. Wadsworth/Thomason Learning, Belmont, CA.
- Shanghai Elementary and Middle School Textbook Reform Committee (Ed.), 2000. *Jiu Nian Shi Yi Wu Jiao Yu Ke Ben* [Chinese Textbooks For Nine Year Compulsory Education]. Shanghai Education Press, Shanghai.
- Shu, H., Chen, X., Anderson, R.C., Wu, N., Xuan, Y., 2003. Properties of school Chinese: implications for learning to read. *Child Dev.* 74 (1), 27–47.
- Shu, H., McBride-Chang, C., Wu, S., Liu, H., 2006. Understanding Chinese developmental dyslexia: morphological awareness as a core construct. *J. Educ. Psychol.* 98 (1), 122–133.
- Shu, H., Peng, H., McBride-Chang, C., 2008. Phonological awareness in young Chinese children. *Dev. Sci.* 11 (1), 171–181.
- Si, X., 1982. *Wo De Jiao Yu Sheng Ya* [My Teaching Career]. Shanghai Education Press, Shanghai.
- Stok, W.T., Fletcher, P., 2001. The role of phonological awareness and visual–orthographic skills in Chinese reading acquisition. *Dev. Psychol.* 37 (6), 886–899.
- Su, B., 1997. *Xu Shen Yu Shuo Wen Jie Zi* [Xu Shen and His Explication of Characters and Words]. Daxiang Press, Zhengzhou, Henan.
- Tan, L.H., Perfetti, C.A., 1998. Phonological codes as early sources of constraint in Chinese word identification: a review of current discoveries and theoretical accounts. *Read. Writ. Interdiscip. J.* 10, 165–200.
- Tao, L., Qian, G., 2012a. Historical perspectives on the Chinese written language and literacy education in China. In: Leung, C., Ruan, J. (Eds.), *Perspectives on Teaching and Learning Chinese Literacy in China*. Springer, New York, NY, pp. 1–17.
- Tao, L., Qian, G., 2012b. Literacy in Ancient China: a culturally and socially situated role in historical times. In: Leung, C., Ruan, J. (Eds.), *Perspectives on Teaching and Learning Chinese Literacy in China*. Springer, New York, NY, pp. 19–33.
- Tao, L., Reinking, D., 2020. The move toward literacy among Confucian scholars in ancient China. In: Hammond, M., Rose, J. (Eds.), *Edinburg Handbook of History of Reading*. Edinburg University Press, England, pp. 11–30.
- Tao, L., Zuo, L., 1997. Oral reading practice in China's elementary schools: a brief discussion of its unique roots in language, culture, and society. *Read. Teach.* 50 (8), 654–665.
- Tong, L., Zhang, Y., 1999. *Xiao Xue Shi Zi Jiao Xue Yan Jiu* [Studies on Teaching Character Acquisitions in Elementary Schools]. Guangdong Education Press, Guangzhou.
- Tong, H.K., Goh, Y.S., 2008. Challenges of Chinese language education in multi-lingual societies: Hong Kong and Singapore. *Int. Educ. Stud.* 1 (4), 44–52.
- Tong, X., Tong, X., McBride, C., 2017. Radical sensitivity is the key to understanding Chinese character acquisition in children. *Read. Writ.* 30, 1251–1265.
- Venezky, R.L., 1996. The development of literacy in the industrialized nations of the West. In: Rebecca, B., Kamil, M.L., Mosenthal, P., David Pearson, P. (Eds.), *Handbook of Reading Research*, vol. II. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 46–67.
- Wang, F., 2020. Reading for rule: Emperor Taizong of Tang and Qunshu zhiyao. In: Hammond, Mary (Ed.), *Edinburg Handbook of history of reading: early readers*. Edinburgh University Press, Edinburgh, pp. 31–53.
- Wilkinson, E., 1998. *Chinese History: A Manual*. Harvard University Press, Cambridge, MA.
- Xu, S., 121/1989. *Shuo Wen Jie Zi* [Explication of Characters and Words]. Zhonghua Book Company, Beijing.
- Xu, Y., 1991. Depth of phonological recoding in short-term memory. *Mem. Cognit.* 19 (3), 263–273.
- Xu, J.X., 2009. *Jian Ming Zhong Guo Wen Zi Xue* [A Concise Outline of Chinese Philology]. Zhonghua Book Company, Beijing.
- Ye, S., 1980. *Ye Shen Tao Yu Wen Jiao Yu Lun Ji* [Collection of Ye Shengtao's Writings on Chinese Education]. Educational Science Publishing House, Beijing.
- Yin, J., Wang, R., 2008. *Xian Dai Han Yu Wen Zi Xue* [Contemporary Chinese Philology]. Fudan University Press, Shanghai.
- Yip, M.C.W., 2014. What is the role of tone in the phonological similarity effect? *Psychol. Rec.* 64 (1), 115–122.
- Zhang, J., McBride-Chang, C., 2010. Auditory sensitivity, speech perception, and reading development and impairment. *Educ. Psychol. Rev.* 22 (3), 323–338.
- Zhang, T., Chen, L., Li, W., 2000. *Zhong Guo Dang Dai Han Zi Ren Du Yu Shu Xie* [Chinese Character Reading and Writing in Contemporary China]. Sichuan Education press, Chengdu, Sichuan.
- Zhang, Z., 1964/1992 reprint. *Chuan Tong Yu Wen Jiao Yu Jiao Cai Lun-Ji Meng Xue Shu Mu He Shu Ying* [On Traditional Chinese Education and Texts With a Bibliography of Character Texts and Book Reprints]. Shanghai Education Press, Shanghai.

- Zhang, L., 1997. *Yu Wen Su Zhi Jiao Yu Yan Jiu* [Studies on Teaching the Fundamentals of Chinese]. Hunan Education Press, Changsha, Hunan.
- Zhang, J., McBride-Chang, C., Wong, A.M.-Y., Tardif, T., Shu, H., Zhang, Y.P., 2014. Longitudinal correlates of reading comprehension difficulties in Chinese children. *Read Writ.* 27, 481–501.
- Zhou, Q., 1999. *Yu Wen Jiao Xue She Ji Lun* [On Pedagogy of Chinese Instruction]. Guangxi Education Press, Nanning, Guangxi.
- Zhou, Y., 2004. *Zhou You Guan Yu Yan Xue Lun Wen Ji* [A Collection of Zhou You Guan's Papers on Linguistics]. The Commerce Press, Beijing.
- Zhou, Z., 2004. *Wen Zi Yin Yun Xun Gu Jiang Yi* [Lectures on Chinese Phonology and Critical Exegesis of Ancient Texts by Zhou Zhu Mo]. Tianjian Ancient Books Press, Tianjing.
- Zhou, Y., 2009. *Han Zi He Wen Hua Wen Ti* [Chinese Characters and Cultural Issues]. People's Literature Publishing House, Beijing.

Chinese Minzu (民族) diversity and education

Sude^a, Mei Yuan^a, and Fred Dervin^{b,*}, ^a School of education, Minzu University of China, Beijing, China; and ^b Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

*Corresponding author

© 2023 Elsevier Ltd. All rights reserved.

Introduction	759
Preliminary notes: a short introduction to Chinese education to contextualize Minzu	760
Minzu as a complement to reflecting on languages and literacies in education	760
Minzu as a complex notion	761
Minzu as a political-economic and/or cultural-linguistic notion?	762
Minzu education as a mirror for education in other contexts?	763
How is Minzu education taking place?	764
Conclusions	767
References	768
Further reading	769

Introduction

As part of official trips made to China in the 1960s, visitors were taken to the Institute of Nationalities in Beijing, the predecessor to today's Minzu University of China (中央民族大学; MUC hereafter), whose mission is to promote research and education for and about China's "ethnic groups". The Chinese word *Minzu* (民族), which can be translated by means of different words in English as we shall explain, replaced that of *nationalities* in the name of the institution, since it is ambiguous in English. Minzu refers to the many and varied "ethnic groups" that compose China, of which the rest of the world knows very little, except for some of the groups which have been highly discussed in "Western" media in the 2020s (e.g. Uighurs in Northwest China). Most visitors to China in the 1960s mentioned their visit to the institution in their memoirs. This is the case of former Canadian premier Pierre Trudeau (1919–2000) who, during his visit to what he refers to as the Institute of Minorities, explained (Trudeau and Hébert, 1967: 48):

It is clear, too, that, among all the students in the Institute of Minorities, the most important, the best prepared, and the most senior are those who are enrolled in the political division and who aspire to become high officials of the Party and of the State when they return to their minorities some day. (At present the minorities provide only four per cent of the members of the Chinese Communist Party.) All the same, the Institute eagerly welcomes those who only wish to become musicians, poets, or dancers. And, far from assimilating these minorities as much as possible (some of them consist of fewer than five hundred people), the powerful Chinese State does all it can to preserve their national characteristics, and notably their languages.

It is asserted that this visit inspired Trudeau to engage with multicultural politics back in Canada (Qian, 2013: 54). Today a visit to MUC will also surprise the neophyte: the institution welcomes every year thousands of students from around the country who might speak languages other than the common language of China (Putonghua, often referred to as Mandarin), and share different worldviews and cultures (see Sude et al., 2020). This diversity is also omnipresent in China, although not so obvious to the untrained eyes. For example, strolling through the Gugong (故宫, i.e. the "Forbidden City" in Beijing), one can see old signs in both Chinese and Manchu (an East Asian Tungusic language from Manchuria in Northeast China) above the Gates and buildings. Every Chinese citizen can see the words "People's Bank of China" in Chinese and in four of China's Minzu languages (Mongol, Tibetan, Uighur and Zhuang) on banknotes in their pockets.

In this chapter we use the notion of Minzu to describe, discuss and problematize the situation of diversity in China, especially in relation to education. This chapter was written during the 2020–2021 pandemic, which was accompanied by many anti-Chinese sentiments in many parts of the world, with discussions about "Chinese minorities" included. However we often felt that critiques of the ways the Chinese are "dealing" with diversity were biased and misinformed. As specialists of interculturality and plurilingualism in education, we feel that we all need to listen to each other and to learn with each other. So-called Western diversity education (different forms of multicultural, intercultural, transcultural and even global education) is far from perfect and has not always proved to be successful. We are not claiming in what follows that Minzu education is a flawless solution or that it offers one-size-fits all recipe. Like all other educational perspectives, it has some "good" and "challenging" aspects. But we argue that by looking into it we can start reflecting on our own ways of "doing" diversity in education. This chapter was thus written with an open mind and it should be read with the same.

Preliminary notes: a short introduction to Chinese education to contextualize Minzu

It is fair to say that Chinese education is largely unknown in other parts of the world, although it has the largest education system in the world with around 10 million students taking the end of high school leaving certificate per year (named *gaokao*, 高考—approximately the entire population of the Nordic country of Sweden).

Often exclusively stereotyped in the media and in some research (Zhao, 2015) for being stressful, competitive, unfair and regimented with 15 million teachers, 300 million students in over 500,000 schools, Chinese education is far from a monolith.

As a reminder China has a population of over one billion, around 17 million newly-born every year, with a territory of about 9 million square kilometers. Just to get a sense of this let's take some examples: the distance between Harbin in the north of the country to Hong Kong is nearly 3000 km and flying time is about 5 h; the distance between Shanghai (East of China) and Kashgar (Westernmost city in China) is 5000 km and takes about 9 h. When we talk of China it is also important to mention the rural-urban divide, with approximately 50% of the population living in both. Up until recently most Chinese families followed the one-child policy—there were exceptions such as families living in the countryside and some Minzu groups—but in 2021, due to the aging population and labor shortage, the three-child policy was implemented.

Since 1986 education has been compulsory and free for 9 years (primary grades 1–6 and junior secondary grades 7–9) and 99% of all school-age children receive these 9 years. It is important to note that in China education is run by the state but decentralized to county-level governments. Pre-school, senior secondary and higher education are neither compulsory nor free, which can represent an economic burden for families since the state only supports very poor students. Early childhood education and care, which is highly privatized, has a small enrollment rate of 10%. Senior secondary education, to which 95% of junior secondary students continue, is divided into general and vocational secondary schools. About 40% of secondary students continue to higher education, a level of education which is very diverse and complex, with different kinds of institutions with varying reputation. Some universities such as Tsinghua and Peking University in Beijing are highly competitive, and are the only ones that can compete with top US universities in most Chinese families' dream of a perfect education and future career path. Interestingly, China is increasingly attracting a large number of international students in higher education (about half a million students per year pre-COVID-19).

For a big country like China, educational equality and equity are a priority. Narrowing the urban-rural gap and regional differences in terms of economic development is important. In rural areas programs such as the Rural Primary and Secondary Boarding Schools Project entitling students to stay in modern dormitories and the Comprehensive Improvement of the Basic School Conditions for Schools at Poor and Underdeveloped Areas, have been implemented for many years. The Special Teaching Post Plan for Rural School aims to attract teachers from urban contexts to teach in rural areas through excellent remuneration and benefits. Finally, long-distance education programs for rural primary and secondary schools have been established (see: Gu et al., 2017; Xue and Li, 2021).

In this chapter, we explore a perspective from Chinese education that has not been discussed in international research. China is a unitary multiethnic state, which means that among its over 1 billion inhabitants, individuals are classified according to different “Ethnic” Minzu groups—56 in total (for instance: Dai, Manchu, Tibetan, Uighur, Zhuang). The labeling of these Minzu groups is the product of a political classification project started in China in 1953 and finished in 1990. We note that different terms have been used in English to refer to these different groups: *ethnic minorities*, *minority groups*, *non-Han peoples*, *local peoples*, *nationalities*, *nationality groups*. In this chapter we use the Chinese term *Minzu* as it avoids misperceptions and inadequate connotations contained in words like *ethnicity*, *minority*, *nationality* in English and other languages, which can be politically marked differently in various parts of the world. We find this perspective to be interesting to unthink and rethink global engagement with languages and literacies in education.

Minzu as a complement to reflecting on languages and literacies in education

Current critiques of the various fields of diversity in education note that theories, methods and practices tend to be overly Euro-/Western-centric and reproduce unequal knowledge power relations (Amsler et al., 2020; Dervin and Jacobsson, 2022). Although diversity education is meant to help people think “otherwise”, try to open up to other ways of reflecting and being, and question the current knowledge and practices that are based on overtly dominating ideologies, what the fields often do is to impose ways of thinking about diversity and interculturality that tend to ignore the diverse ways of thinking about and “doing” them around the world.

A study by Peng et al. (2020) notes that the most quoted scholars in the field of interculturality in education are US- or UK-based, white and English speakers. For a field that strives to discuss and address the position of diverse individuals in education and beyond, this limits the spheres of intellectual engagement with, for instance, issues of equality-equity, social justice and encounters beyond borders. Looking at the list of dominating voices identified by Peng et al. (2020), one cannot help noticing that these Western scholars: (1) use similar concepts and notions in English, without questioning systematically their potential geopolitical and multilingual meanings or their “social lives” (e.g. *culture*, *community*, *respect*, *tolerance*, etc.); (2) emphasize cultural difference over similarity; (3) (somewhat naively) urge people to avoid conflicts and misunderstandings; (4) see diversity negotiations as something linear rather than complex and unstable; and finally, (5) avoid discussing economic, socio-political and ethical elements openly when dealing with interculturality (see Dervin, 2016; Dervin and Simpson, 2021).

Some localized forms of alternative, complementary and companion education have been identified in the literature, such as *interculturalidad* in South America, *interculturalité* in France, *tabunka kyōsei* (“multiculturalism and coexistence”) in Japan. Minzu, as a perspective from China, has not been discussed in international research on interculturality and diversity. The influence from outside China on the way Minzus have been classified and discussed in the 20th century is noted by Guo (2020: 17), who asserts: “[...] when foreign words and ideologies such as race, ethnicity, nation, nationalism, civilization and society were introduced into China, scholars at that time were struggling to understand these terms and how to relate them to the Chinese context”. The fact that translations of the word Minzu in English and other languages are multiform and somewhat confusing seems to derive from this mix of often a-contextualized ideologies that have penetrated the Chinese field of thought about Minzu since the early 20th Century.

To explain briefly the idea of Minzu we could say that it symbolizes a certain view of the heterogeneity of the Chinese nation and of the interaction between different groups which together form an entity called the *People's Republic of China*. As much as European nation-states were categorized as they stand today, through (re-)negotiation of (national) identities, (national) languages and (national) symbols, Minzus are the results of historical interactions between different kinds of peoples coming together. However, we note that, unlike many other places in the world, discussions of diversity and interculturality, via Minzu, are central in internal political and economic discourses in China and engagement with education.

We repeat: Getting acquainted with Chinese ways of discoursing about diversity, through Minzu, does not mean suggesting that we should copy and/or implement them in our respective contexts. Of course, there might be ideas that deserve being considered as inspiration, but it is up to the reader to decide what to do with them. We see considerations of Minzu as a way of allowing ourselves to reflect on how we *see* and *do* languages and literacies but also diversity and interculturality in our respective parts of the world—like looking in a mirror—taking stock of achievements, challenges and potential changes.

Minzu as a complex notion

Throughout its history, China witnessed interactions between different Minzu groups. Until 1911, China was ruled by the Manchu Minzu from Northeastern China (the last emperor Puyi from the 1987 film entitled *The Last Emperor* by Bertolucci was a member of the Manchu Minzu). According to Leibold (2016), it is around that time that the notion of Minzu appeared, influenced by Japan. As such, a Qing Dynasty scholar, Liang Qichao (梁启超, 1873–1929), is often reported as having translated the Japanese term *minzoku* (民族) into Minzu in the late 19th century, making the notion part of the discourse on the Chinese national state-building and of the nation-state. The Chinese word Minzu is composed of *min* (民) for “folk or common people” and *zu* (族) for “consanguinity or lineage”.

56 Minzu group identities and memberships are officially recognized by the Chinese State today. The Han majority represents about 91.5% of the population while the other 55 Minzu groups 8.5% (Zang, 2016: 1). In many cases the latter are often viewed as socio-economically, culturally and ethnically different from majority Han Chinese, but not necessarily “poorer” or “less educated” as is often wrongly suggested by some Chinese and foreign scholars. Minzu status appears on people’s identity documents and depends on their parents’ own status (e.g. Chinese Han, Chinese Korean, Chinese Uighur, etc.). Note that the idea of Minzu only applies to Mainland China and that in other parts of China such as Hong Kong, the concept of “ethnicity” refers to other categories, i.e. people from outside China, who speak a language other than Chinese/Cantonese (Pakistanis, Filipinos, etc.).

China has been described as consisting of different Minzu groups that interact with each other and have been integrated into the Chinese nation during its history. Different Minzu groups are scattered in different territories of the country, with many concentrated on border regions with, for instance, Kazakhstan, Korea, Mongolia, and Russia. There are also five provincial-level ethnic autonomous regions: Inner Mongolia Autonomous Region, Xinjiang Uygur Autonomous Region, Guangxi Zhuang Autonomous Region, Ningxia Hui Autonomous Region and Tibet Autonomous Region (Sude et al., 2020). We should bear in mind with Wang and Du (2018) that “ethnic minorities [Minzu] are not a group of homogenous people. The differences within ethnic minorities should not be ignored as well”. Many of the groups that compose these different areas share languages, worldviews and cultural aspects with groups on the other side of the borders. In interviews that we conducted with international students based at MUC, some students note these similarities. For example a student from Laos explained that his own language group (Laotian speakers) and the Bouyei language spoken by some Chinese Minzus in the Southern Guizhou Province of China are both Kra-Dai languages, tonal languages spoken in Southeast Asia, Southern China and Northeast India with around 93 million speakers, which are not necessarily unintelligible.

We also note that some Minzu minorities have different first languages on top of Mandarin Chinese, the common language of Mainland China and that some of their languages were divided politically at some point but might be very similar to other Minzu languages. For instance, we conducted a research interview with a specialist of the Bouyei who live in the South of China and number approximately 2.5 million people (Dervin and Yuan, 2021). When asked what language the Bouyei speak, the scholar discusses its similarities/differences and shared un/intelligibility with Zhuang from another Minzu. Zhuang is from the Tai family including Thai and Lao:

Excerpt 1

I(interviewer): If one person speaks Zhuang language, how much can you understand?

S(specialist): It depends on what region the person is from.

I: What if they are from Wuming [located in the capital of Guangxi Zhuang Autonomous Region of South China]?

S: I can't understand the Zhuang language in Wuming area, because it is very different from Bouyei language. But I can understand the Zhuang words of areas like Tianlin [County in the west of Guangxi] and so on.

I: How much can you understand? 50%?

S: I can understand 80%.

I: Can you understand 80%? Why?

S: Because the Zhuang language in these areas has a lot of the same vocabulary as Bouyei.

Zhuang is in fact a complex set of more than 12 languages spoken in the South of China, which are not mutually intelligible (Handel, 2019). This shows that the borders between Minzus are not always clear-cut and that complexity and hybridity seem to be strong components of what Minzus are from cultural and linguistic perspectives.

According to the Constitution and Law of the People's Republic of China on the Standard Spoken and Written Chinese Language, all Minzus are entitled to use and develop their own spoken and written languages (2004: Article 4): "All nationalities have the freedom to use and develop their own spoken and written languages and to preserve or reform their own folkways and customs." For instance, in Xinjiang, 10 different languages are used in administration, broadcasting and television, education, legislation and publishing (China Daily, June 30, 2019). It is necessary then to bear in mind that a Chinese person answering the question *what is your mother tongue?* might in fact not reveal their "real" mother tongue or the fact that they are bilingual from birth since many speak other languages as first languages (e.g. Mongolian, Korean, Tibetan, etc.). Instead, they might just mention Mandarin Chinese as their own mother tongue. Mandarin Chinese is the common national language of Mainland China that serves as a lingua franca and is widely used in education.

The notion of Minzu is thus much more complex than the mere naming of different "ethnic" and linguistic groups in China. For Zhao (2014: 31), when we try to express the idea of Minzu in English, "irrespective of which concept of Minzu we employ or which standpoint we take, we are only exchanging one Western model for another, without ever finding a way of identifying and expressing our own Chinese uniqueness".

Minzu as a political-economic and/or cultural-linguistic notion?

Politically, it was the Communist Party of China that made officially the country as a multi-Minzu and unified nation. The identification project, which took many decades, officialized Minzu identities and memberships. What is more, Zhang and Chen (2014: 400) explain that the 1950s marked "the establishment of Autonomous Regions for concentrated ethnic populations, the coexistence of an ethnic language school system and a Mandarin school system as options for ethnic students to choose from, and preferential treatments in minority students' access to college". All these were meant to promote cultural diversity, reforms and regional economic development as well as bridging between "majority" Han and other Minzus. Preferential treatment and policy privileges were widely made available for Minzus, in terms of employment, business development, and political representation. As far as education is concerned, remedial programs, preferential admissions into universities and boarding schools for Minzu students are common today.

In terms of research some Chinese scholars from different fields of research have problematized Minzu. This is the case of the anthropologist and sociologist Fei Xiaotong (费孝通, 1910–2005), whose theory on Minzu groups in Chinese history is influential. Fei argues that all the 56 Minzus share an overarching identity which is the result of centuries of cohabitation and interaction, regardless of their differences in terms of language and culture. The anthropologist and sociologist used the phrases "plural unity pattern" and "unified multiethnic state" (中华民族多元一體; Fei, 1988) to describe Minzu relations: Minzus are plural and diverse, yet, they are united by the fact that they form together the Chinese nation. The Chinese nation is thus both culturally pluralistic and

politically unified. For Fei, in order to coexist harmoniously, all Minzus need to be aware of and accept each other's unique characteristics and joint identity. Many Chinese scholars in education but also Chinese policymakers have embraced and implemented Fei's theory (Zhang and Chen, 2014).

However, some Chinese scholars have noted recently problems with the notion of Minzu. As such Zhang (2017) maintains that Minzu can have many different functions in Chinese society, which makes it somewhat polysemic. He argues (Zhang, 2017) that different definitions of Minzu seem to include:

- The United Nations' definition for "ethnic groups"
- The Chinese national definition based on the official categorization following the creation of the Republic of China
- An academic definition for "cultural groups"
- People's identity whereby they self-identify or are identified by others.

Another aspect of current debates is described by the same scholar (2017) who suggests, following Ma (2017), that the overemphasis on ethnic issues to deal with Minzus makes it too broad as it encompasses different elements such as cultural, social, religious but also economic issues (e.g. poverty). Ma (2017) had argued that autonomization of regions and preferential policies to boost certain Minzu groups' economic outlook, have led to the politicization of Minzu groups, which isolates them and creates, at times failed Minzu integration. Zhang and Chen (2014: 405) summarize this problem as follows: "(...) in the attempt to pacify minority groups and enhance social solidarity and national unity, the Chinese authorities had engaged in an active process of establishing, essentializing, and staticising cultural differences between groups". Zhang (2017) thus proposes that a focus on the cultural specificities of Minzu groups while bridging them with others, would ensure smoother relations between representatives of these groups. We note that discussions of culture and Minzus are common in scholarly circles, while linguistic aspects are somewhat left aside or embedded in discourses of "culture".

Minzu education as a mirror for education in other contexts?

It would be wrong to pretend that Minzu education is a label that refers to a unified approach to cultural and linguistic diversity in education in China. In fact, it refers to many and varied formal and informal aspects of Chinese education, related in/directly to Minzu.

As such different models of Minzu education have been identified in the literature. First, some Minzu groups can be educated in their first language in their home provinces, in Mandarin Chinese or in both languages (Leibold and Chen, 2014). Linguistically this means for example that a Mongolian student studying in a bilingual school gets to learn, use and work in three (very) different languages: Mongolian, Mandarin Chinese and English (which is compulsory). Second, three different types of primary and secondary schools are available: mono-ethnic Minzu schools, mixed Minzu schools, and mixed Han/other Minzu schools. Top students from Tibet and Xinjiang can also receive their middle and high school education in Putonghua in other provinces (so-called "inland" education). Leibold (2016) thus argues that China has the world's most extensive regime of minority affirmative action policies or "preferential policies" (优惠政策) in terms of admission to educational institutions and employment in state or public institutions, acknowledging the existence of educational inequality.

It is important to note that in terms of teaching, Minzu education concerns, on the one hand, specific forms of educational approaches, such as meeting the needs of Minzu students through bilingual education in Minzu regions, and Minzu higher education institutions like MUC in Beijing where Minzu cultures and languages are taught; and, on the other hand, education for all, with, for instance, the teaching of aspects of Minzu included in moral and political education in Chinese basic education.

This is confirmed in a review of what Zhang and Chen (2014) call "Chinese Multicultural Education":

- In Chinese citizenship education the centrality of both national and Minzu identities is emphasized. Strengthening all citizens' Chinese national identity (like in most other countries—see for instance the idea of *citizenship education* in Europe) is deemed essential to ensure social cohesion.
- Minority cultural and linguistic teaching is made available to Minzu students who have a first language other than Mandarin Chinese.
- School culture and as many teaching activities as possible should integrate students' cultural and linguistic backgrounds and experiences in basic education all over the country.

Theoretically, there have been attempts at problematizing and conceptualizing Minzu education. However, there is no unified theoretical underpinning of Minzu education in China, apart from the notion of "plurality within unity", which relates Chinese national unity and Minzu diversity. Chinese students tend to experience diversity in terms of language, culture, and sometimes worldviews/religions. Their identity can thus be situated on a continuum of "Chineseness" and "Minzu" (see Yang's (2017) study on different Tibetan students' identification).

Xing (2001: 580–581) is one of the few scholars to have tried to devise an "integrated multicultural education" containing both "plurality within unity" and (American) multicultural education (following American scholars like James Banks). His work was influenced by anthropologist Fei (1988) oft-quoted paper "The Chinese Nation with Multi-ethnic Groups" which contains a macro-theory of the formation and structural characteristics of the Chinese nation. His main analytical element of China as a unified multi-ethnic country is summarized in the phrase 一体多元 (*yiti duoyuan*, *united but pluralistic*). As mentioned earlier,

Fei's theory has been adopted by both education policymakers and scholars in China under the label "unified pluralist education" (多元一体教育, *duoyuan yiti jiaoyu*). Through this aspect Minzu students learn about their own Minzu's origins, history of formation and development while reflecting on the fact that their Minzu is interrelated with other Chinese Minzus through a common sense of national belongingness and unity.

Teng's (2010) own approach to "integrated multicultural education" was also directly inspired by Fei's work. In 2012 Teng described his approach as follows: "the education of a multi-ethnic country has the function not only to transfer the common human cultural achievements, but also, it should pass on the valued and time-honored cultural traditions of its dominant ethnic group as well as the ethnic minorities." (p. 165) The aims of this approach are to promote common economic development, cultural heritage, cross-cultural communication, the political equality of all Minzu groups, and the ultimate realization of national unity (Teng, 2010). For Teng (2012) all these elements concern all students in all Chinese schools, majority Han as well as other Minzu groups. As such Han students (and teachers) should get to know the languages and cultures of various Minzu groups too (Teng, 2012).

Zhang and Chen (2014) discuss a range of policy initiatives from 1987 onwards that represents a clear shift from education for different Minzu groups to what is referred to as "ethnic solidarity education". This kind of education is now included in moral and patriotic education in primary and middle schools and named "national unity education" (民族团结教育). Several textbooks are used to teach this subject and introduce students to *The Big Chinese Family* (primary school), *Ethnic Basics* (primary School), and *Basics of Ethnic Policy* (middle school).

Interestingly many Chinese scholars have used "American" multicultural education theories and practices to discuss the principles of Minzu education. For instance Zhang and Chen (2014) remark that there are both similarities and differences between diversity education from the "West" (multicultural/intercultural education) and Chinese Minzu perspectives. In terms of differences Zhang and Chen (2014) note that Chinese Minzu education is based on a different policy approach ("Pluralistic Unity of the Chinese Nation" with an emphasis on difference and similarity), and that its concerns have more to do with economic, social and cultural development rather than, for instance, issues of political equality.

As in other countries, national identity is seen as fundamental in Chinese Minzu education to ensure social cohesion and reduce potential Minzu tensions (Qian, 2010). More importantly, Minzu groups in China are indigenous peoples unlike many of the "targeted" populations in "Western" diversity education (e.g. multicultural/intercultural education), who e.g. migrated to the US or Europe. We might want to add that multicultural and intercultural education emerged from specific historical contexts. As such multicultural education was born out of the 1960s civil rights movements while intercultural education out of European decolonization in the 1970s (see Derwin and Jacobsson, 2022). Economic-politically speaking these differ from Chinese Minzu, whose sometimes disadvantaged position is said to derive from "the harsh climate, undesirable geographical conditions and economic underdevelopment in the border land they occupied" (Zhang and Chen, 2014: 403). In the "West", diversity education often relates to situations of education diversity in urban and large city contexts.

How is Minzu education taking place?

Engaging with the literature published in English, we note that there is a lack of an overview of what Minzu education encompasses. Fig. 1 represents an attempt at clarifying what this multifaceted label includes and lists perspectives that concern certain groups of students (e.g. bilingual education) and all Chinese students (e.g. national unity education).

The examples in this figure do not include informal education whereby, for instance, Chinese people learn about other Minzus. For example, in an interview with a Bouyei Minzu specialist (South of China) at MUC, we discussed the idea of the "sworn brothers

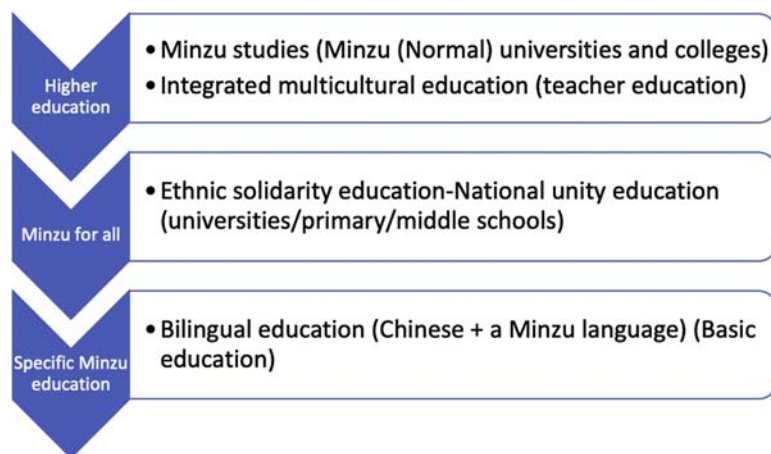


Fig. 1 Layers of Minzu education (examples) (Derwin and Yuan, 2021).

and sisters" tradition (結拜) (Dervin and Yuan, 2021), which creates some form of Godparent relations between people of different Minzus (e.g. Bouyei & Han/Miao). In his study of a Miao Minzu village, Wang (2018: 62) describes what "sworn brother-/sisterhood" consists in: "all sworn brothers come to give a hand for carrying firewood, while the housewife prepares a feast for coming brothers and their wives. They sit around eating, talking, laughing and playing poker until midnight. The gathering makes both cooperation among sworn brothers and progression emotion each other." This habit is qualified as "a basis for Minzu unity" by the Bouyei scholar whom we interviewed, which he confirms by claiming that "This custom would not exist if Minzu groups did not believe in each other". This Minzu inclusiveness from the Bouyei side is also illustrated by the Bouyei scholar's claim that Minzus learn each other's language to be able to interact with each other:

Excerpt 2

We learn the languages of other peoples in order to acquire a life skill of communicating with them. It is not surprising, for example, that Han people living in the Bouyei Minzu region are generally fluent in the Bouyei language, and that Bouyei people can also speak Chinese. In this way, there will be no misunderstandings between different Minzus because of language barriers. When I can speak your Minzu language, there is no barrier between us.

We discuss formal bilingual education in what follows, which plays a central part in Chinese Minzu education. Gao and Ren (2018) maintain that the 55 Minzu minority groups speak over 209 languages (e.g. Mongolian, Tibetan, Uyghur, Zhuang). In the past Chinese scholars have noted that many Minzu students stop learning their Minzu languages because they perceive them as "futile" in the labor market, especially when competing with other students with good Standard Chinese skills (e.g. Zeng, 2014). This has been perceived as counter to the plural unity pattern. This is why Minzu bilingual education has been put into place around the country. According to Zhou (2012) 21 Minzu groups have implemented some form of bilingual education for approximately 6 million Minzu students, whereby they get to learn three languages: their first language (a Minzu language), Chinese and English. Different types of bilingual education are offered in different parts of the country (Mańkowska, 2019): (1) a Minzu language is the main language of instruction from kindergarten to middle school, with Chinese taught as a separate subject; and (2) A Minzu language is used at the start of school education to help non-Chinese speakers easily understand the content. In higher grades, the percentage of Minzu language teaching is reduced, and the Chinese language predominates. It is important to note that thousands of textbooks are available in Minzu languages and consist of direct translations of textbooks published in Chinese.

At the time of writing, in a project that we were running about Science, Technology, Engineering and Mathematics (STEM) education in bilingual and Chinese-speaking schools the Mongolian city of Ordos (鄂尔多斯) (Yuan et al., under review), we discussed with primary school teachers how they negotiated multilingualism. In the Chinese-speaking schools, all teachers (be they Mongolian speakers or not) use exclusively Chinese to teach and may use Mongolian when they speak to individual students. Chinese is the school lingua franca and using other languages during teaching would be considered as discriminating against those who don't speak them. In the bilingual schools, the teachers and students use mostly Mongolian, yet, there are times when Mandarin Chinese is useful for both teachers and students. A technological instructor explained for instance that the review materials for national exams were limited in Mongolian so he had to use material produced in Mandarin Chinese. What is more, some teachers reported using code-switching between Mongolian and Mandarin Chinese for certain words that do not have "accepted" equivalents in Mongolian. The following excerpts are from fieldwork in two different Mongolian bilingual schools:

Excerpt 3

I: Right. Can you give me a specific example from your teaching experience? For example, if you speak to them in Mongolian, they may not understand you and you have to explain it in Chinese.

T: Like "dissolve".

I: Is there a special word for it in Mongolian?

T: Yes. And now most children watch cartoons that say "dissolve". They get it when they read it.

I: Maybe they have seen "dissolve" in cartoons and know what it means, but in Mongolian they may not know.

T: Yeah.

Excerpt 4

T: According to the actual situation of Mongolians, liberal arts, such as politics, history, language, etc., are more in line with the history of the Mongolian people. There are many technological aspects of Mongolian language, such as tractors. The Mongolian people do not have tractors. We still have to use Chinese or English.

To dissolve and *a tractor* are two examples that many teachers mentioned during the interviews.

In terms of teaching STEM, we were also curious about the use of ICT and written Mongolian during our visits to the schools. In China, unlike the country of Mongolia where it is written in Cyrillic, the traditional Mongolian script is used. Derived from the old Uighur alphabet, it is written in vertical lines right across the page. In our interview with a math teacher, he discussed the use of Mongolian in the iPads used in his school:

Excerpt 5

I: Is Mongolian the language of the iPads?

T: Yes. In the western area of China, we have developed a Mongolian software in 2019, and now we can directly use Mongolian for courseware. Originally the words that we write in Mongolian are “upside down”, so we needed to correct the typesetting after writing. Later, we had a training session. A lecturer from the West came to train us. We reported to him about the software problem. After he returned, he upgraded the software. Now writing Mongolian in the software isn’t “upside down” any longer, it is exactly the same as other software we use.

Besides languages, we also note that most teachers wish to pass on certain values that they consider to be specific to Chinese Mongolians to their students, while discussing STEM elements. In the following excerpt, a chemistry teacher makes a reference to the importance and protection of water by Mongolians:

Excerpt 6

I think about it ... Mongolian ... Um ... About the Mongols ... (Long pause) Well, it should be ... a frugal aspect of Mongolian culture. Because we have water in the fourth chapter, water is said to be all around us in that chapter, I just tell the students it’s frugality, and then don’t waste water. Perhaps that’s it.

Although this example might sound anecdotal, the teachers seemed to systematically include such elements in their classrooms and in their teaching. All of them admitted not having been professionally prepared for including these Mongolian Minzu elements or to support the students linguistically but that they “improvised” them for STEM education.

In general, one of the main issues faced by Minzu education is the lack of consistent teacher education and training focusing on Minzu. In autonomous regions of China, around 100 Minzu normal universities, normal colleges, or schools for teacher education and training have been established. Some Minzu universities and colleges located in Minzu areas also offer Minzu teacher education programs. Teacher training centers have been established and Minzu teacher education courses are also widely available in major Normal universities around China such as Beijing Normal University, Central China Normal University, and East China Normal University. These programs are opened to both Han and non-Han students. It is also important to note that inland universities have signed partnerships with universities in Minzu areas to offer assistance in terms of Minzu teacher education and training. For instance, inland experts are sent to Minzu areas to take part in teaching and research programs.

In general, Minzu teacher education and training appears scattered, and somewhat inconsistent. The emphasis appears to be on preparing teachers for specific Minzu groups rather than for mixed groups. What is more it is often unclear who gets trained: Han-majority to work with other Minzus and/or vice versa? Most review articles of Chinese Minzu education include very few references to this specific issue although forms of Minzu teacher education and training have been organized since the founding of the People’s Republic of China (Hawkins, 1978). A few articles have examined teacher education and training for teachers dealing with Minzu issues. For example, Ou and Du (2017)’s article shows that teachers interested in strengthening students’ Minzu identity can have

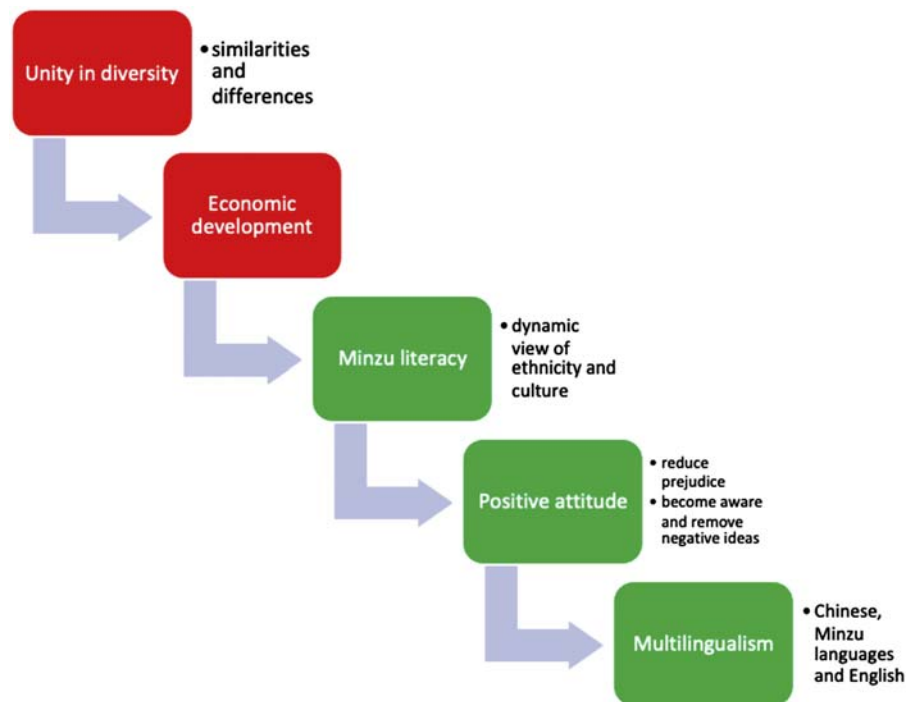


Fig. 2 Basic values of Minzu education (Dervin and Yuan, 2021).

a positive impact on their academic achievement. In her 2011 review article of a teacher education program in the Xinjiang Uyghur Autonomous Region, Liu proposes that knowledge of “multicultural education” should be added to the curriculum. Finally, in their 2014 paper entitled “Creating a multicultural curriculum in Han-dominant schools: The policy and practice of ethnic solidarity education in China”, Zhang and Chen (2014) look into teachers’ preparedness to implement ethnic solidarity education in a mix of (Minzu/non-Minzu) schools in Beijing. Their findings reveal that the schools incorporated Minzu elements in the school environment, in teaching and extra-curricular activities; however, the scholars note that “they tended to focus on the static cultural artifacts, such as food, music, dance and sports etc., which represented the museum-like knowledge about different ethnic groups, i.e. the so-called ‘living fossil’” (Zhang and Chen, 2014: 412). “Cultural heritage” rather than e.g. fluid and “living” perspectives on cultural complexity was highlighted. Zhang and Chen (2014) also assert that many teachers and school leaders found ethnic solidarity education to be unneeded.

In their extensive literature review of Minzu education from 2000 to 2018, Liu et al. (2020) claim that teacher education and training for Minzu is one of the six important topics in the field. The review seems to confirm that, in general, teacher education in China seems to lack systematic and coherent Minzu perspectives. Some scholars have made suggestions as to how to strengthen this aspect of Minzu teacher preparation: Zhang (2015) emphasizes the importance of critical thinking and creative thinking; Pu and Liu (2017) insist on creating a common sense of belonging by, for instance, teachers and students learning from each other through cultural exchange; and Zheng (2012) suggests that teachers explore the preferred learning styles of Minzu students. Developing multicultural and multilingual teaching strategies is deemed to be an important option. We note again that languages and literacies do not appear to be systematically discussed directly in research on teacher education and training but that they could be embedded in discourses of “culture”.

Conclusions

Real dialogue isn’t about talking to people who believe the same things as you.

Bauman (2016: n.p.).

At the end of this chapter, we note that Minzu education is a complex and yet rich perspective on “internal” diversity, which has been rarely discussed in global scholarship of languages and literacies. Like any approach to diversity in education, Minzu is based on political, scholarly and educational views of “us” and “them”, while trying to create a bridge between these entities. Although there are some similarities in terms of ideologies and practices with “Western” multicultural and/or intercultural education, Minzu education does have its own specificities. Figure 2 summarizes the basic values of Minzu education (Fig. 2):

Although the model might appear linear or hierarchical, all the different components are interrelated. These elements also appear to be overarching in the way multiple involvement with Minzu education has been organized in China. In red we indicate the basic politico-economic values around discussions of Minzu education: *unity in diversity* (i.e. we share similarities at the national level but differences at Minzu group levels), *economic development* (which is considered to be the key to equality among Minzus; see “preferential policies” for certain Minzus as a boost). The green elements indicate the principles of Minzu education observed nationally: *Minzu literacy* (the presence of Minzu representations and discourses in education for all; support for Minzu groups to become aware of their own Minzu specificities and similarities with other Minzus), *multilingualism* (Chinese, Minzu languages for Minzu language speakers and English as a global language) and a *positive attitude to all* (reducing prejudice, awareness and removing negative ideas about different Minzus).

To conclude this overview of Chinese Minzu diversity and education, we would like to renew our advice to examine the way languages and literacies are dealt with in contexts which are rarely discussed in diversity education (see Dervin and Jacobsson, 2022). Entering the world of Minzu requires adopting a different mindset and rethinking the way we use words, understand concepts and notions, but also policies and practices. What we noticed is that although concepts such as *social justice* are not omnipresent in discussions of Minzu in education, other words hint at similar ideologies about language and culture (e.g. the right to access to development).

References

- Amsler, S., Kerr, J., Andreotti, V., 2020. Interculturality in teacher education in times of unprecedented global challenges. *Educ. Soc.* 38 (1), 13–37.
- Bauman, Z., 2016. Social Media are a Trap. Interview With de Querol. *El Pais*. https://english.elpais.com/elpais/2016/01/19/inenglish/1453208692_424660.html.
- Dervin, F., Jacobsson, A., 2022. *Intercultural Communication Education. Broken Realities and Rebellious Dreams*. Springer, London.
- Dervin, F., Simpson, A., 2021. *Interculturality and the Political Within Education*. Routledge, London.
- Dervin, F., Yuan, M., 2021. *Revitalizing Interculturality in Education: Chinese Minzu as a Companion*. Routledge, London.
- Dervin, F., 2016. *Interculturality in Education*. Palgrave, London.
- Fei, X., 1988. The plural-unity structure of the Chinese nation (Zhonghua Minzu de duoyuan yiti geju). *J. Peking Univ. (Beijing Daxue Xuebao)* 4, 1–19.
- Gao, X., Ren, W., 2018. Controversies of bilingual education in China. *Int. J. Biling. Educ. Biling.* 22 (3), 267–273.
- Gu, M., Ma, J., Teng, J., 2017. *Portraits of Chinese Schools*. Springer, London.
- Guo, Z., 2020. *Changing Ethnicity: Contemporary Ethno-Politics in China*. Palgrave Macmillan, London.
- Handel, Z., 2019. *Language, Writing and Literary Culture in the Sinographic Cosmopolis*, vol. 1. Brill, Leiden.
- Hawkins, J., 1978. National-minority education in the People's Republic of China. *Comp. Educ. Rev.* 22 (1), 147–162.
- Leibold, J., Chen, Y.B. (Eds.), 2014. *Minority Education in China*. Hong Kong University Press, Hong Kong.
- Leibold, J., 2016. Preferential policies for ethnic minorities in China. In: Zang, X. (Ed.), *Handbook on Ethnic Minorities in China*. Elgar, London, pp. 165–188.
- Liu, Q., Çolak, F.Z., Agirdag, O., 2020. Characteristics, issues, and future directions in Chinese multicultural education: a review of selected research 2000–2018. *Asia Pac. Educ. Rev.* 21, 279–294.
- Ma, R., 2017. Reconstructing “nation” (Minzu) discourses in China. *Int. J. Anthropol. Ethnol.* 1 (8), 2–15.
- Mańkowska, A., 2019. Learning and teaching opportunities of ethnic minorities in multicultural China. Unity, diversity or critical pluralism? *Cult. Soc. Educ.* 1 (15), 69–81.
- Ou, Q., Du, N., 2017. Teachers’ multicultural awareness and the ethnic identity of minority students: an individual case study of a Hani student. *Front. Educ. China* 7, 212–226.
- Peng, R.-Z., Zhu, C., Wu, W.-P., 2020. Visualizing the knowledge domain of intercultural competence research: a bibliometric analysis. *Int. J. Intercult. Relat.* 74, 58–68.
- Pu, W., Liu, Y., 2017. *Lun minzu diqu jiaoshi de kuawenhua jiaoyu shiming jiqi hexin suyang* [One the intercultural education mission and the core qualities of the teachers in the ethnic regions]. *Guangxi shifan daxue xuebao (zhexue shehui kexueban)* 53, 109–114.
- Qian, M.H., 2010. Ethnic education should undertake the major historic task of ethnic unity and the possibility of and explorations into the localization of diversified education. *Chinese Educ. Soc.* 43 (5), 62–76.
- Qian, Z., 2013. *56 Ethnic Groups of China*. Beijing Language and Culture University Press, Beijing.
- Sude, Yuan, M., Dervin, F., 2020. *Introduction to Ethnic Minority Education in China: Policies and Practices*. Springer-Verlag Berlin An, Frankfurt.
- Teng, X., 2010. “多元文化整合教育”与基础教育改革 [“Multicultural integrated education” and curriculum reform of basic education]. *中国教育学刊* [J. Chinese Soc. Educ.] 11, 51–52.
- Teng, X., 2012. Cultural diversity and integrated multicultural education. *Front. Educ. China* 7 (2), 163–168.
- Trudeau, P.E., Hébert, J., 1967. *Two Innocents in Red China*. Oxford University Press, Toronto.
- Wang, W., Du, L., 2018. Education of ethnic minorities in China. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore.
- Wang, W., 2018. Researching education and ethnicity in China: a critical review of the literature between 1990 and 2014. *Front. Educ. China* 13 (2), 216–244.
- Xing, H., 2001. *Minzu Education Pedagogy*. Education Science Press, Beijing.
- Xue, E., Li, J., 2021. *Educational Equity Policy in China: Concept and Practice*. Springer, Frankfurt.
- Yang, M., 2017. *Learning to Be Tibetan: The Construction of Ethnic Identity at Minzu University of China*. Lexington Books, London.
- Yuan, M., Sude, Chen, N., Dervin, F. (under review) “I always combine Tiananmen Square with herdsmen’s cattle pens and sheepfolds”: The place of Minzu in Chinese STEM education (Ordos, Inner Mongolia).
- Zang, X. (Ed.), 2016. *Handbook on Ethnic Minorities in China*. Elgar, London.
- Zenz, A., 2014. *Neo-integrationism, Minority Education and Career Strategies in Qinghai*. Brill, P.R. China. Berlin.
- Zhang, D., Chen, L., 2014. Creating a multicultural curriculum in Han-dominant schools. *Comp. Educ.* 50 (4), 400–416.
- Zhang, J.Y., 2015. *Minzu diqu jiaoshi de wenhua mingxianxing jiqi shengcheng* [Cultural sensitivity of teachers in ethnic areas and its formation]. *Zhongguo chengren jiaoyu* 6, 80–82.
- Zhang, X., 2017. The paradigmatic crises in China’s Minzu studies: reflections from the perspective of human development. *J. Chinese Humanit.* 3 (7), 135–155.
- Zhao, X., 2014. *Zhongguo minzu yanjiu de kunjing jiqi fanshi zhuanhuan* [The dilemmas and paradigm shifts in Chinese nationality studies]. *Tansuo yu zhengming* 4, 29–35.
- Zhao, Y., 2015. *Who’s Afraid of the Big Bad Dragon?: Why China Has the Best (and Worst) Education System in the World*. Jossey-Bass Wiley, Hoboken, N.J.
- Zheng, Y., 2012. *Lijie xing wenhua yanjiuzhe: duoyuan wenhua lijie jiaoyu zhongde jiaoshi jueze* [Understanding-oriented culture researcher: the role of teachers in education for multicultural understanding]. *Minzu jiaoyu yanjiu* 23 (4), 70–74.
- Zhou, M., 2012. Historical review of the PRC’s minority/indigenous language policy and practice: nation-state building and identity construction. In: Beckett, G.H., Postiglione, G. (Eds.), *China’s Assimilationist Language Policy: The Impact on Indigenous/Minority Literacy and Social Harmony*. Routledge, New York, pp. 105–120.

Further reading

- Benton Lee, M., 2016. Sociological perspectives on ethnic minority teachers in China: a review of the research literature. *Diaspora Indig. Minor. Educ.* 10 (1), 55–68.
- Clothey, R., Hu, D.Y., 2014. The impact of a national goal driven higher education policy on an ethnic minority serving institution in China. *High. Educ. Pol.* 28, 353–368.
- Fang, W., 2019. *Modern Notions of Civilization and Culture in China*. Palgrave MacMillan, London.
- Longxi, Z., 1988. The myth of the other: China in the eyes of the West. *Crit. Inq.* 15 (1), 108–131.
- Wu, D.Y.-H., 1991. The construction of Chinese and non-Chinese identities. *Daedalus* 120 (2), 159–179.
- Xiong, W., Jacob, W.J., 2020. Ethnic minority-serving higher education institutions in the United States and China: a comparative case study of two institutions. *Asia Pac. Educ. Rev.* 21, 295–309.
- Yuan, M., Sude, Wang, T., Zhang, W., Chen, N., Simpson, A., Dervin, F., 2020. Chinese Minzu education in higher education: an inspiration for “Western” diversity education? *Br. J. Educ. Stud.* 68 (4), 461–486.
- Yuan, H., 2018. Educating culturally responsive Han teachers: case study of a teacher education program in China. *Int. J. Multicult. Educ.* 20 (2), 42–57.

The politics of literacy language and texts: the politics of literacy in Nigeria.

Chukwuemeka Eze Onukaogu^a, Joseph Onyema Ahaotu^b, Onyedikachi Abiodun-Ekus^c, and Kate Omenugha^d, ^a Reading Association of Nigeria, Ile-Ife, Osun, Nigeria; ^b English Language Teachers Association of Nigeria, Department of English Studies, University of Port Harcourt, Port Harcourt, Rivers, Nigeria; ^c Faculty of Education, University of Reading, Reading, United Kingdom; and ^d Nnamdi Azikiwe University, Awka, Anambra, Nigeria

© 2023 Elsevier Ltd. All rights reserved.

Introduction	770
Nigerian Indigenous education	771
Situating literacy in Nigerian education	773
Literacy and the Nigerian language ecology	774
The politics of literacy in Nigeria	774
The impact of missionaries	775
Impact of Islamic missionaries	775
Impact of Christian missionaries	775
The impact of government	776
Leaping into action: the impact of Nigerian academicians	777
Gestating and birthing literacy for life-long learning in Northern Nigeria	778
The literacy enhancement and achievement impact	779
Conclusion	783
References	784

Introduction

Literacy in foreign languages was introduced into the African continent in the 10th century AD and through the centuries thereafter, Portuguese, Arabic, and English languages assumed prominent roles in the politics of literacy in African countries such as Nigeria. This raised pertinent questions of the role of Indigenous languages in literacy dialectics in Nigeria (Fig. 1). The contentions have continued to generate concerns through the missionary, to colonial, and to post-colonial (contemporary) Nigerian educational experiences and attainment. They have often translated into the debate on which language should be used for education, and the dichotomy between the home language and the school language, with implications for educational achievement premised on the linguistic diversity in Nigeria. In the traditional (Indigenous) Nigerian system of education, the medium was the mother tongue (MT) while oracy was its pivot. In formal Nigerian education, which was introduced by foreigners, English is the primary medium while literacy is its pivot.

In the rest of this chapter, six key developments will be examined. First, we will analyze Indigenous Nigerian education before foreign systems of education were introduced into Nigeria. We will show how oracy promoted communication among Nigerians and served as the instrument through which Nigerians promoted their culture, science and technology in their various MTs. Second,



Fig. 1 Map of Nigeria representing the 36 states, the Federal Capital Territory, and the Northern and the Southern Regions. Courtesy: Chizoba Ikenwa (nigerianinfopedia@gmail.com), <https://nigerianinfopedia.com.ng/northern-states-in-nigeria/>.

we will examine the impact of formal education on literacy empowerment before the commencement of the Trans-Atlantic slave trade. Third, we will situate literacy in Nigerian education after the abolition of the slave trade. Fourth, we will examine the role politics plays in the literacy enhancement and achievement of Nigerians. Fifth, we will try to summarize the key measures being taken to gestate and birth literacy in Northern Nigeria. Finally, we will suggest a Literacy Enhancement and Achievement Paradigm (LEAP), a teacher/teacher, teacher/learner and learner/learner empirically proven effective literacy instructional and learning model as a paradigm shift in the promotion of literacy.

Nigerian Indigenous education

Nigerian Indigenous education is rooted in antiquity. Fafunwa (2004) has rightly observed that through Nigerian Indigenous education, children were empowered to acquire character, respect for others, physical skills, intellectual capabilities, appreciation of their religion and culture, etc. and oracy centrally underpinned the system of education. Writing on the value of oracy among the Ibos of Nigeria in his celebrated novel, *Things fall apart*, Achebe noted that “Among the Ibo the art of conversation is regarded very highly, and proverbs are the palm-oil with which words are eaten” (p. 6). Indigenous Nigerian education was transmitted orally through a guild system, the home, and a system of apprenticeship. The apprenticeship scheme as a form of entrepreneurship and skills education is still popular in contemporary Nigeria (Fig. 2).

In agriculture and home economics for instance, boys were trained by their fathers or older male members of their home or communities while girls were trained by their mothers or older female members of their communities respectively. In special professions like hunting, medicine, metal work (such as iron smiting, bronze and brass technologies) one had to belong to a guild before one could acquire the knowledge, attitudes, and skills that could make one excel in them. As could be seen in the activities of the Ekpe or Okonko guilds, those who did not belong were not privileged to benefit from the education which the guilds offered. This is because these guilds had their system of reading and writing—their kinds of “literacy”—that were not open to everybody. It is in this



Fig. 2 Master and apprentice: a young Nigerian boy learns tailoring from a skilled master. Courtesy: Rosemary Onyema-Ahaotu's Photo Album.

context that we can understand, for instance, the Insibidi writing system developed and used by members of the Ekpe guild in pre-colonial Igbo, Ibibio, and Annang communities of today's South Eastern and South Southern Nigeria. The guilds were also exclusive in other aspects of their activities, including dressing, perception of life, and culinary culture.

The leaders and teachers in their respective fields passed their knowledge, attitudes, and skills through the oral tradition to their students. It was through this mode that the people developed the science and technology that has amazed the world. For instance, the Nok Terra Cotta technology (Jemkur, 2006), which started about 500 BC, is considered one of the best and most sophisticated terra cotta technologies the world has ever practiced. The Terra Cotta art was part of an expansive civilization (National Geographic Society, 2020; Rupp et al., 2008) and its technology was practiced almost in all parts of present-day Nigeria except in the coastal regions of what is now the South-South region. Some historians have argued that the iron smiting technology that flourished in Awka, Anambra State and which provided the bulk of the metal wares used in agriculture, hunting, and warfare in the Gulf of Guinea and the Bight of Biafra is an offshoot of the Nok Terra Cotta technology. Similar arguments have also been postulated that the famous Brass and Bronze (Fig. 3) technologies that flourished in Benin and Ife had their origins from the Nok Terra Cotta technology (Onukaogu, 2008, 2010).

The tragedy of the Nigerian Indigenous education is that it was not allowed to thrive. Just as the Nigerian Indigenous education was refining and repositioning itself in the light of its contact with Arab Islamic education and Western European Christian education—especially in its synchronizing oracy with the emerging literacy, the Trans-Atlantic slave trade set in and further disrupted its development. One can then say that texting in Nigerian Indigenous education was oral while languaging was through the MT of the learner. Oracy however, needs to blend with literacy for education to maximally empower Nigerians for individual and national development. We share Oyedepi's (2007, p. 5) position that oracy has its limitations as stated in the following words:

Oracy was becoming inadequate and inefficient; inadequate because of the multiplicity of cultures that emerged and inefficient because the details in handing over cultures, history, economics and political practices from one generation to another cannot be retained in human memory without written records; hence it is said that when an illiterate dies it is as if a whole library is burnt.

Literacy languaging and texting forms enhance the preservation and transmission of knowledge and sustainable education (Swain, 2006). The political decision of choice of language for literacy has significant implications on the effectiveness of literacy empowerment as demonstrated in latter parts of this chapter.



Fig. 3 Benin Bronze sculpture Unfinished image of Queen Idia, mother of Oba Esigie (1504–1550). Courtesy: Mr. William Edosomwan, Sculptor, UNESCO Heritage Site, Benin City, Nigeria (+234 802 339 2043).

Situating literacy in Nigerian education

From the 10th to the 15th century AD, there were two sets of foreigners who impacted on the Nigerian Indigenous education, and their impact reshaped the education in the land. It added a new dimension to the oracy through the use of literacy (Onukaogu, 1998, 1999). The first set of foreigners was Arab Islamic missionaries who came in the 10th century AD and were concentrated in Northern Nigeria (Imam, 2012; Oroma and Wosu, 2018). They came primarily to convert the Indigenes to the Islamic faith using Arabic as their medium of communication (Trimingham, 1962; Ogbondah, 2002). The second set was Roman Catholic Jesuit Christian missionaries who came between 1515 and 1580 to the Benin Kingdom where they set up schools and churches (Ogbondah, 2002; Oroma and Wosu, 2018). They introduced Western literacy and emphasized the 3Rs—Reading, Writing, and Arithmetic.

Literacy has meant different things to some members of the Nigerian education orchestra. The Christian missions in the South taught literacy as the 3Rs. To the Christian missionaries, literacy was the pivot of formal education. They ensured that the graduates of their schools and colleges were literate in their MTs and in English. In contrast, the Islamic missions in the North did not teach the 3Rs in the Koranic schools they established. Instead, the Koran was read in Arabic to students who were required to memorize what was read to them. Thus, unlike their Christian school counterparts, the graduates of Koranic schools could neither read/write/compute in their MT nor in Arabic (Ajayi, 1965; Imam, 2012; Oroma and Wosu, 2018).

Generally, literacy is the skill of reading, writing and computation (Oyedemi, 2007). According to Opeyemi (2018, p. 150) “many definitions of literacy focus on the ability to read and write at appropriate levels.” Opeyemi agreed with English and Aers (2008, p. 32) that literacy “is the ability to read and write which makes a person literate, with varying degrees of fluency.”

The Nigerian Educational Research and Development Council (NERDC) holds the view that literacy is a skill. It is this view that informs its mass literacy strategy slogan “each one, teach one.” The Council’s position is based on the assumption that literacy does not require intellectual demands. That is why the Council promotes the assumption that an illiterate person can be taught to be literate by anyone who can read and write. However, some Nigerian scholars contend that “literacy, which used to be traditionally defined as the ability to read, write and calculate ... is man’s capacity to expand his own consciousness in a bid to exercise power over himself, his environment and his society at large” (Adedokun and Kayode, 2019, p. 91).

Literacy has been described as consisting of two models - the autonomous and the ideological models (Onukaogu, 2012; Street, 2006). Street (2006) compared the autonomous and ideological models and indicated that while the autonomous model of literacy presents its underlying cultural and ideological underpinning as neutral and universal; the ideological model challenges this view through research in new literacies that suggest that in practice literacy varies in contexts and cultures.

In the autonomous model, literacy is reading, writing and calculating without any intellectual, cultural and environmental impacts and requirements. In the ideological model, literacy is seen from a much wider perspective where the intellectual and social perspectives of learning, knowledge, attitudes, and skills are essential for its acquisition. Onukaogu observed that literacy has multiple meanings, and he also cited Rassol’s (1999) view that the definitions of literacy change in response to social transitions in history, ideology, or technology.

Indeed, many Nigerian scholars have argued that literacy is a set of knowledge, attitudes and skills for text processing. Ahaotu (2018), Aridegbe and Makinde (2014), Ayodele and Adedokun (2012), Adedokun (2008), and Orisawayi (2006) also shared the view that literacy has expanded in scope and some of them listed new literacies as digital literacy, critical literacy, computer literacy, information literacy, and media literacy. For Ahaotu (2018) literacy includes a variation of the communicative arts, including writing, calculating, cognitive processing of texts, and digital competences through which individuals and societies preserve their modes of interaction, learning, culture, and development. Promoting a synchronicity of the autonomous with the ideological models of literacy, Egbe (2019, pp. 22–23) posited that “literacy is dynamic, contextual and culture specific. This is why ... we speak of not just ‘literacies’ but new literacies given that literacies account for ever expanding dimensions of literacy from its traditional sense.”

UNESCO plays a leading role in global literacy empowerment since 1946 and has described literacy as follows:

Beyond its conventional concept as a set of reading, writing and counting skills, literacy is now understood as a means of identification, understanding, interpretation, creation, and communication in an increasingly digital, text-mediated, information-rich and fast-changing world.

UNESCO (2019, n. p.).

UNESCO’s perspective is that literacy empowers individuals to optimally comprehend the written text in various contexts and enables them to develop their knowledge and potentials to participate more meaningfully in both their community and the wider society. From this perspective, literacy is a critical element of both individual and national development (Ahaotu, 2018).

Both the autonomous and the ideological models of literacy share the view that literacy is a critical factor of human and societal developments. The perspectives of both these exponents and UNESCO’s (2019) definition of literacy may be synthesized into the following key points that are fundamental to literacy:

1. Written texts. Underlying literacy is the written text. The emphasis here is on the distinction between oracy and literacy. In contrast to literacy, oracy does not incorporate elaborate written text. There is no literacy without the written text, and so, it could be contended that without literacy written words have no meaning.

2. Written text diversity. The written text can be on printed materials, and can be reflected on diverse media. This diversity reflects the multiplication of media with the written words on them.
3. Intellectual and processing capabilities. Since the written or printed words have meanings and messages to convey, whoever wants to obtain the meaning or messages in them must have the knowledge, attitudes and skills to process the contexts of the written texts so that the person can identify, understand and interpret their meaning.
4. Synchronicity with oracy. The purposes of using the printed and written materials include creating, communicating and computing meaning. Obviously, oracy, sign language, and Braille may be utilized to share information garnered from the written or printed text given the various contexts in which individuals find themselves.
5. Knowledge, attitudes and skills enhancement. Processing the written text empowers the individual with the ability to identify, understand, interpret, create, communicate and compute meaning. These hallmarks of texting have the potentials of enhancing the knowledge, attitudes and skills of processing and giving information. That is why in the Nigerian situation, literacy is pivotal to the acquisition of knowledge, attitudes and skills in formal and non-formal school systems, where the written text is indispensable to knowledge, attitudes and skills acquisition and storage.

Literacy and the Nigerian language ecology

There are social pressures for the use of English for literacy practices in Nigeria and studies have identified a high demand for English among Nigerians (Ahaotu and Tobiahs, 2020; Famogbiyele, 2017; Egbe, 2014). This is understandable because English is Nigeria's major official language even though there are co-official languages. The NERDC's (2013) *National policy on education* (henceforth, *NPE*, 2013) provided that "Every child shall be taught in the mother tongue or the language of the immediate community for the first 4 years of basic education" (Section 1.8, *NPE*, 2013). The *NPE* stressed that "Government will ensure that the medium of instruction is principally the mother tongue or the language of the immediate community" (Section 2.16) for early childhood care development. However, Section 2.20 of the same document provides that "the medium of instruction in the primary school shall be the language of the environment for the first 3 years in monolingual communities. During this period, English shall be taught as a subject." Thereafter, the transition to an English-medium instruction will start from Primary 4 (P4) and above.

These declarations show degrees of inconsistencies. First, the policy assumes that there are many monolingual communities in Nigeria, which does not reflect the reality of the Nigerian multilingual context. Second, it is assumed that all the teachers have the linguistic competence and training to deliver instruction in the MT or the language of the immediate environment. Third, it is assumed that after 3 years of instruction in English in a second language context, children would have had sufficient proficiency in the language for transition to an all English medium instruction. These policy lacks practical implementation strategies and so, it leaves the classroom teacher in some dilemma; teachers are therefore likely to do what they think is feasible and realistic. Trudell (2018) agreed that there are also internal inconsistencies in the language provisions of the *NPE* and that, "... the *NPE* does not provide a clear, coherent, feasible language-in-education policy ... making them difficult to implement with fidelity" (p. 10). A study by Fadoro (2013) in three Western states of Nigeria found that the provisions of the *NPE* are violated extensively and the inherent inconsistencies of its provisions may have contributed to the violations.

Nigeria is a multilingual country with 526 foreign and local languages (Ethnologue, 2018). Since language remains the vehicle of instruction in the classroom, it is obvious that children with low proficiency in the school language are disadvantaged, especially in multilingual contexts where there is a difference between home language and school language. The importance of the Indigenous languages or MTs as the language of instruction is well documented (Wa Mberia, 2015). However, the high demand of English and its position in Nigeria create some dilemmas for its deployment in literacy. English competes with the three major Indigenous languages in Nigeria (Hausa, Igbo, and Yoruba). English competes with other Nigerian languages that have fewer speakers than the three major ones. English also competes with other foreign languages present in Nigeria. This results in Nigeria being a country with one of the most complex linguistic ecologies globally and increasing the dilemma of the choice of the language of literacy in Nigeria. On a positive note, there is a technical team charged with fashioning out a framework for developing a national language or languages for Nigeria. Bearing in mind that language issues are as volatile and contentious as religious issues in Nigeria, the tensions created by the multiplicity of languages in Nigeria and the inconsistencies identified in the language provisions in the *NPE* have serious implications on literacy development in Nigeria and, very importantly, heighten the politics of literacy in Nigeria as will be shown in subsequent sections of this chapter.

The politics of literacy in Nigeria

Politics, which determines the nature and scope of literacy in Nigeria, is influenced by three principal factors. First, there are the missionaries—Islamic and Christians—who used literacy to achieve their faith winning agenda in the country. Second, there is the Nigerian Federal Government whose literacy agenda is determined by the political party in control of the federal or state governments. Finally, there are the Nigerian scholars who through research and practice are endeavoring to reposition literacy in the Nigerian education system so that proven literacy views and best practices in the world can be adapted in the country.

The impact of missionaries

Impact of Islamic missionaries

We had noted earlier that the Islamic missionaries before and after the era of the Trans-Atlantic slave trade did not teach their students to read in their MTs. Literacy empowerment for critical and creative thinking was not available to the students. Since the students were not taught how to read in their MTs, they lacked the literacy empowerment of linking what they were taught to learn by heart with the day-to-day challenges in their communities. Besides, they were also not taught to read, write and compute in Arabic. If they had been taught to read, write and compute in Arabic, they would have been exposed to the enormous body of information and knowledge in Arabic which had helped to liberate Europe from the Dark Ages and launch the continent into the Renaissance. The Islamic missionaries in Northern Nigeria denied the people literacy empowerment so that they would remain in perpetual submission to their Koranic teachers and the feudal lords the teachers represent. Presently, according to the [UNICEF \(2015\)](#) Demographic Health Survey (DHS), 69% of the 13.2 million out-of-school children today in Nigeria are from the North.

The colonial administration also denied Northern Nigerians literacy empowerment by preventing Christian Missionaries from moving to the North and planting Western European formal education there. [Imam \(2012, p. 183\)](#) cited Fabunmi (2005) to contend that the "British policy of indirect rule restricted the activities of Christian missionaries in the predominantly Muslim Northern protectorate thereby, curtailing the spread of Western education" and also cited Ogunsola (1982) to indicate that this led to considerable education gap between Northern and Southern Nigeria. Obviously, the colonial administration endorsed the literacy disempowerment which the Islamic education posed to the generality of the children of the North. The grants-in-aid which were given to Christian missions to build schools in the South were not given for building and improving the Koranic schools in the North ([Imam, 2012](#)). The majority of children in the North were greatly handicapped; they could not perform literacy activities through their MTs, Arabic, and English. All they could show for their attending formal Koranic schools was the recitation of aspects of the Koran. The colonial administration thus sowed the seed of literacy disparity between the North and the South which today remains the bane of political, social and economic development of Nigeria.

Impact of Christian missionaries

Soon after the slave trade was abolished, Christian Missionaries came back to Nigeria to win converts to their faith and build schools, hospitals and craft centers in the country. Because of the restriction by the colonial administration, they concentrated in Southern Nigeria. By 1842 the first Christian primary school after the abolition of the slave trade was built in Badagry in present day Lagos state.

The formal education that the Christians introduced was anchored on autonomous literacy. The students were taught to read, write and compute in their MT. Because of the ease with which the students could read and write in their MT, and in English, formal education became interesting and very exciting to them. Consequently, there was an explosion in the number of students in formal education. The missionaries could not cope with the demands. They went to the Colonial Office and asked for financial assistance to build more schools. The Colonial Office offered to assist the missionaries only if they used English and not the MTs of the students in their teaching. Since the missionaries desperately needed funds in order to build more schools, they jettisoned the MT for English-medium instruction. This crippled the use of the MT as a medium of instruction in the mission institutions. Many communities in the South built their primary schools and handed them over to the Christian missions to administer.

When the Colonial Administration passed the Macpherson's Constitution in 1951, each of the regions was empowered to appropriate and use funds for education. As a result of this development, Western Nigeria started Universal Primary Education (UPE) in 1955 and Eastern Nigeria in 1957. Northern Nigeria went beyond UPE and gave enticing incentives to Northern children to go to school. However, the dislike in Northern Nigeria for Western formal education prevented the majority of the children from embracing the generous incentives. Enrollment and attendance at schools in the North were comparatively lower than it is in the South.

Since the 1940s the Christian Missions have built first rate primary and secondary schools in the South and the Middle Belt (present North Central States) areas of Nigeria. They also perfected the use of English as a medium of instruction. Their secondary and primary schools compared favorably with the best primary and secondary schools in the British Commonwealth. Literacy taught at the mission schools was now for literal, interpretive, creative, critical, life-long learning across the school curriculum. The graduates of the mission schools excelled in such external examinations as the University of London General Certificate of Education - Ordinary and Advanced Levels - and the University of Cambridge School Certificate Examination. The Colonial Administration established some secondary schools in the country to supplement the efforts of the Christian Missions. There were also few privately owned schools in the South. However, in terms of industry, honesty of purpose and academic excellence, the colonial administration and the few privately owned schools were no match for the mission schools. The literacy for critical, creative, life-long learning, across the school curriculum, and for independent living, which the mission schools had perfected and taught and which the government and private schools emulated, paid off handsomely. This trend continued till 1966 when the civil war broke out in Nigeria. We can say without equivocation that the 1940s-1960s marked the comparative golden age of Nigerian formal education. It was when literacy empowerment, as the pivot of formal education, was at its peak.

The impact of government

It has been shown in this chapter that the colonial government used literacy to socially control, repress the people, and make the people totally submissive to it (see [Stromquist, 2008](#); [Gaff, 1987a,b](#)). It promoted illiteracy in Northern Nigeria and stopped the use of MT in literacy development in Southern Nigeria. After the golden age of literacy in the country, the Nigerian Federal Government, by underfunding education, also hindered the ultimate development of literacy in the country. The Nigerian military government was more or less a surrogate of the feudal oligarchies in Northern Nigeria. The feudal oligarchies did not want the majority of the people in the country to have sound education, which could empower them with literacy and, by implication, liberate them from the shackles of feudalism. [Godonoo \(1991, p. 81\)](#) has observed that Nigerian politicians make “deceptive statements about literacy” and cited both [Ihejieta-Aharanwa \(1989\)](#) and [Brown and Tomori \(1979\)](#) for instances of misleading promises.

In the 1970s the Federal Military Government took over primary and secondary schools in the country. It vested the running of the schools to state governments. However, the state governments, through the underfunding of education, ran the schools down. As [Unoh \(1980\)](#) and [Omojuwa \(1984\)](#) have pointed out, many of the graduates of our primary and secondary schools all over the country can hardly read and write fluently in either English or in their MTs.

However, there were later efforts by some state governments, such as Rivers State in the South-South and Anambra State in the South-East of Nigeria, to redress the issue of government take-over of mission schools. In 2011 Anambra State established the Public Mission Partnership which returned mission schools formerly taken by the Federal Government to the original owners. Under the partnership, Anambra state government remained in charge of payment of salaries of teachers in the returned schools and also provided regular subventions to missions to manage the schools. However, the non-replacement of teachers due to attrition, retirements and deaths over the years have led to insufficient teachers in the schools, leaving the public schools grappling with insufficiency of teaching staff; which no doubt impinges on literacy. The public mission schools are dealing with the issue by permitting Parents Teachers Associations (PTAs) to recruit teachers as the need arises. Such employment has increased the burden of staff salary and caused an increase in tuition. Thus, the schools are becoming both elitist and out of the reach of the poor. A well-intended model is gradually being destroyed.

Generally, education in Nigeria is confronted with significant challenges that include poor funding, poor infrastructure, inadequate classrooms, inadequate teaching aids, inadequate and poor treatment of teachers, and many others. Successive governments have demonstrated a lack of political will to redress these problems and encouraged a growing contention that the government utilizes illiteracy as a means of impoverishing the masses and keeping them under manipulative control ([Oyedepi, 2007, p. 8](#)). The situation in Nigeria is compounded by the “attitude of the technocrats toward issues of literacy and adult education. Literacy has always been relegated to the background” ([Oyedepi, 2007, p. 8](#)). The deliberate neglect of education has made illiterate people living in poverty to accept their unfortunate state as normal. Indeed “it may also be true that the illiterate population have unsatisfactory or unacceptable attitudes toward literacy and education in general because, as they may argue, if their children who are lettered are unemployed what use would their own literacy be” ([Oyedepi, 2007, p. 9](#)).

The systemic disinterest in literacy has also been manifested in the level of support for language-based and literacy-related researches and projects in Nigeria. For instance, in the study “Profile of Education and Language in Nigeria from 1970 to 2019,” [Egbe and Ayeni \(2019\)](#) showed that beginning from the Primary Education Improvement Project (PEIP) in Northern Nigeria in 1970 and the Six-Year Primary Project popularly known as the Ife Primary Education Project also in 1970 in the South West, to the present, these projects have largely been funded by international donor agencies with very little support from local funding by Nigerian government. What this profile shows is that the past efforts to propagate literacy in Nigeria have met with failures because, as important as it is, literacy has always been ignored, poorly funded and not given appropriate position as a vital tool in the process of development and as the fulcrum of formal education ([Adedokun and Kayode, 2019](#)).

Apart from perceiving literacy from the narrow lens of being the ability of merely reading and writing, the NERDC does not seem to appreciate the centrality of reading in literacy. The Council has failed to realize that:

Reading is so interrelated with the total educational process that educational success requires successful reading. Experience has taught us that those who fail in school have failed first in reading. Giordano Bruno pointed out that if the first button of a man's coat is wrongly buttoned, all the rest are certain to be crooked. Reading is the first button in the garment of education.

[Dechant \(1977, p. 1\).](#)

Unfortunately, there is no serious provision for the teaching of reading in Nigerian schools. In an extensive study of how reading is taught in Nigerian secondary schools, [Kolawole](#) observed that 70% of teachers observed did not teach reading. 60% of the teachers felt that the teaching of reading should be handled by the English language teachers. The remaining 40% encourage their students to read if they want to pass examinations and “if they wished to do well in their academic work. Our library observation showed that close to 90% of the schools visited had no functional library; the 10% that indicated they had libraries had no adequate books in school subjects for students” ([Kolawole, 2009, p. 394](#)). As long as there are no qualified teachers who should supervise what is being taught, how it is being taught, and who should ensure that there are adequate materials for teaching and learning, optimal and maximal literacy empowerment cannot be attained in Nigeria.

The lack of political will to implement the Nigerian National Policy on Education (NPE) hinders the attainment of a robust and dynamic literacy position in the Nigerian educational system. For instance, the NPE stipulates that in the first 3 years of primary

school education, the MT of the child or the MT of the child's immediate community should be the medium of instruction. The Ife Primary Education Project that took place between 1970 and 1978 (Fafunwa et al., 1989) proved that when Nigerian children are taught through their MT, they learn all school subjects better than when they are taught through English. Unfortunately, the NERDC and both the Federal and State Ministries of Education do not have the political will to implement the MT medium policy and so straight-for-English teaching is the norm and not the exception.

The literacy rating of Nigeria is poor (UNESCO, 2012) and the efforts of the Nigeria government toward improving this poor literacy rating appear to be feeble in terms of policy making, policy implementation, and budgetary allocation (Onukaogu, 2012). According to UNICEF (2019), a fifth of the global population of out-of-school children are Nigerian. In 2009 regarding formal education, 69.1% Nigerians were literate. In 2015 the rate dropped to 59.6% showing a drop of 9.5%. In non-formal education in 2009 the literacy rate was 57% but in 2015 the rate dropped by 8%–49% (Opeyemi, 2018). Nigeria does not comply with UNICEF's recommendation that at least 26% of a nation's annual budget should be for education. For instance, in 2017, only 7.04% of the Nigerian annual budget was set aside for education. In 2018 it was also 7.4% while in 2019 it was 7.5% increasing by only 0.1%. The trend of 6–8% budgetary allocation has become a regular index of educational funding in Nigeria. Given this deliberate under funding, education in Nigeria requires significant intervention. During the golden age of education in the 1940s–1960s there were libraries in regional capitals and divisional headquarters all over the country. For instance, the Eastern Nigeria Library Board in addition to its regional and divisional libraries had mobile libraries, which operated on weekly basis between Enugu the regional headquarters and each divisional headquarter. It had a good book loan program in which book requests could be freighted from Enugu the regional capital to any part of the region. The British Council had big libraries open to the public at the regional and national capitals (Enugu, Ibadan, Kaduna, and Lagos) respectively. Secondary schools had first rate libraries. Such facilities are no longer available in the majority of schools and colleges in Nigeria.

During the golden era of Nigerian education, teachers were supported to attend local and foreign summer schools where they were retrained for more effectiveness and efficiency. Today, sponsored workshops, seminars, and summer schools for teachers are rare. It is clear that Nigeria's federal and most state governments do not care about their inefficient and ineffective planning of education and its negative impact on literacy. That is why Godonoo contended that:

Furthermore, educational planning and implementation in Nigeria has often been hampered with statistical constraints. Usually, planners and administrators do not have accurate figures with which to work For example, during the implementation of the Universal Primary Education scheme in Nigeria, it was projected that "2.3 million students would register during the first year of the program in 1976. It was however observed that over 8 million students registered." This difference between the projected and the actual figures created a lot of unforeseen problems in the areas of manpower requirements, classroom accommodation, and general facilities.

Godonoo (1991, pp. 84–85).

The Federal Ministry of Education took a significant step toward improving literacy in Nigeria in the project, Education for Change—A Ministerial Strategic Plan (2018–2020) by devoting a chapter to youth and adult literacy. The document admitted that this sector of education had been neglected over time and listed the reasons for the negligence to include poor and inadequate planning, lack of political will, and poor utilization of human and material resources, among others. The Ministerial Strategic Plan is an effort to refocus on both formal and informal education, which has been neglected for decades. The impact of the deliberate neglect of education has also affected the robust and dynamic development of literacy in Nigeria.

Leaping into action: the impact of Nigerian academicians

Academicians contribute to efforts to reposition literacy and overall education in Nigeria. They work from different professional group platforms and employ strategies that include: professional networking/development, engagement, advocacy, and intervention in collaborations with government and non-governmental agencies. Although their areas of focus differ, these associations generally promote aspects of education including literacy. They include the Academic Staff Union of Universities (ASUU), Reading Association of Nigeria (RAN), English Language Teachers Association of Nigeria (ELTAN), Nigerian Mathematical Society, Nigerian English Studies Association (NESA), and the Linguistics Association of Nigeria (LAN). RAN and ELTAN represent examples of groups of Nigerian academicians relentlessly endeavor to improve Nigeria's literacy rating and the quality of education available in schools.

The Reading Association of Nigeria (RAN) was formed in 1982 by a group of Nigerian academicians in order to reposition literacy to what it was during Nigeria's golden era of education. Through national and international conferences, workshops, round tables, seminars and symposia, RAN mitigates the challenges of literacy in Nigeria. The association convinced the Federal Ministry of Education to launch the Read Campaign; a project that promotes reading awareness nationwide, workshops, and roundtables for the professional development of classroom teachers.

In 2001 and 2004, RAN with the support of the World Bank undertook an extensive survey of literacy teaching and learning in all levels of the Nigerian school system. The survey showed that literacy was not taught in different levels of the school system except for adult literacy in the Departments of Adult Education in Nigerian universities. Literacy was only taught tangentially as listening, speaking, reading and writing in English language lessons in primary and secondary schools. Taught as atomistic skills, it was hoped that the learner would integrate them into literacy. It seems to us that many learners never integrate them.

The English Language Teachers Association of Nigeria (ELTAN) was founded in 2004 by some academicians through the efforts of Dr. Patience Ezinwoke and Dr. Felicia Moh to fill the gap in in-service teacher training and literacy development by forming a network of teachers of English at all levels of education in Nigeria. With the initial assistance of British Council, the association operated monthly workshops in which some members and other experts facilitate on chosen topics. The monthly workshops are organized at state chapter levels while an annual national workshop/conference holds between August and September every year. The basic networks include common regular group of-line events and online platforms, which are utilized in sharing resources and good professional practices. The wider networks include collaborative work with other organizations (RAN, British Council, the US Department of State, publishers, government agencies, etc.) and affiliations with associations of similar interest (Africa TESOL, IATEFL, etc.). Between 2005 and 2021, ELTAN has secured various grants from English Language Teaching Contact Scheme (ELTeCS), the US Department of State, British Council, etc. to execute national, regional, and international events that enhance literacy and the teaching and learning of English in Nigeria and other African countries.

ELTAN, in collaboration with British Council Nigeria, has played significant roles in promoting English language professional development and literacy via English in Nigeria. The association has partnered with British Council on literacy-based interventions in Northern Nigeria. RAN has also collaborated with ELTAN on many English language teaching (ELT) and literacy projects. ELTAN has active branches in twenty-nine states of Nigeria through which it drives literacy education. In 2018, ELTAN initiated the Language Resources Project; a collaboration of the Association, British Council, and the National Nomadic Education Commission (NNEC) to alleviate the challenges of teaching and learning literacy in the difficult circumstances of insecurity in North Eastern and North Western states in Nigeria. ELTAN has also enabled teachers and pupils in different parts of Nigeria to benefit from the Association's professional networking, sharing of good practice, workshops, conferences, publications, etc. and facilitated the cascading of RAN's LEAP through its chapter networks. The British Council is further enhancing the literacy empowerment content of its education promotion drive in the North through its Supporting Pedagogical Rigor in Networked Teaching (SPRINT) and its Stimulating Teacher Resourcefulness in Digital Education (STRIDE) in the Federal Capital Territory (FCT), Kaduna and Lagos. Through its collaboration with ELTAN in the North and sustained implementation of SPRINT and STRIDE, the North now even has the potential of not only equaling the South but also surpassing it in its literacy drives.

Gestating and birthing literacy for life-long learning in Northern Nigeria

While Koranic education emphasized rote learning where the pupils were taught recitation and were never given the literacy empowerment to read and write texts, store in writing what they had written and dialog with others regarding what they had written, their counterparts in schools in the rest of the country had such empowerment. This is because schools in the rest of the country had used literacy, the pivot of Western formal education in teaching and learning. The Federal Government of Nigeria and its development partners has taken deliberate measures to ensure that all Nigerian children irrespective of their geographical locations are optimally empowered with literacy for lifelong learning. However, some of Nigeria's development partners like the USAID, the World Bank, British Council, Action Aid, UNICEF, and UNESCO went the extra mile to promote literacy and Western education as instruments for attaining sustainable development in the North East and North West. We shall briefly mention some of the Nigerian development partners and some of the steps taken that have in recent times made the people of the North East and North West invest in Western formal education and are now vigorously implementing literacy empowerment in their schools.

In December 1989, the Federal Government of Nigeria established NNEC to provide formal Western education to children of herders. The idea was to make the nomadic people realize that formal Western education would enhance their productive capacity. In addition, the government wanted the herders to know that formal Western education could complement Koranic education. Specially trained personnel were attached to the herders who trained their children to not only read and write in their mother tongues but also learn conventional school subjects. The response to nomadic education was quite encouraging. Both the children of the herders and their parents came to appreciate formal education as a way of life. Because of the improvement in literacy learning, the mandate of NNEC was extended to also cover all migrant farmers and fishermen in the riverine areas of the country.

The Federal Government also set up the National Commission for Mass Literacy and Adult Non-Formal Education (NCMLANFE) in 1990. The basic aim of NCMLANFE was to empower illiterate Nigerian adults with literacy. Since the Christian missions in Nigeria had adult literacy programs for the illiterate adults in their respective congregations the focus of the NCMLANFE should have been more concentrated in the Northern part of the country where Christian missions had very few mass literacy programs for their illiterate adults. The activities of NCMLANFE failed to meet its objectives. For instance, it did not employ competent professionals who could handle the literacy development of adults. Besides, enough materials specific to the mother tongue literacy needs of the illiterate adults were not available to sustain the program. It appeared also that those assigned to teach the illiterate adults were not properly remunerated for their efforts. Thus, while the policy formulators of the program had mouth-watering incentives, those who actually did the teaching were demoralized. Thus, whether, in the North or South, the generality of illiterate adults were not interested in the mass literacy and adult non-formal education. However, advancement in literacy teaching and learning was made in formal education in Nigeria when the Federal Government launched the Universal Basic Education (UBE) program in 2004. The Universal Basic Education Commission (UBEC) was set up to oversee all aspects of the UBE at both the Federal and state levels. At the state level, the Universal Basic Education Board (UBEB) was set up to implement the policies of UBEC. Three percent (3%) of the federal budget was set aside for the implementation of UBE. The 6 years of primary and 3 years

of junior secondary school which were funded by the Federal Government have impacted very positively on literacy teaching and learning in the country. In order to further boost literacy teaching and learning in the North, the Federal Government has taken steps to integrate Koranic Schools in the North with conventional Western formal education. When this bold step of the Federal Government is fully implemented, literacy empowerment would be achieved in Northern Nigeria.

The World Bank has taken dynamic steps to promote literacy empowerment among teachers and learners in Nigeria, especially Northern Nigeria. Based on the success of its pilot program in some selected Northern states, the Bank has in its "Basic Education Service Delivery for All" (BESDA), decided to improve education in the North East and North West of Nigeria. Since the Bank has observed, "Enrollment rates in basic education are also marked by gender disparities and wide variation across zones of Nigeria, with the North East and North West demonstrating lower outcomes" (World Bank, 2017, p. 3), all the states in the North East and North West are covered in its BESDA initiative. The Reading Association of Nigeria (RAN) is partnering with the Bank in actualizing the literacy component of BESDA. RAN had collaborated with the Bank in implementing similar projects in Anambra, Ekiti and Osun states. We believe that if current efforts are sustained, the literacy rating of the North would be more positive in the future.

Since 2000, RAN has collaborated with the USAID in implementing its education programmes in the country. For instance, RAN has collaborated extensively with USAID in its Northern Education Initiative Plus (NEI Plus), a pilot project in Bauchi and Sokoto states designed to provide quality education especially for girls and out of school children in the two states. Apart from enabling the project partners to strengthen the abilities of local government authorities to better train and manage teachers, plan budgets, mobilize community support for learning, ensure timely distribution of learning materials, RAN in collaboration with NEI Plus established 400 Community Reading Centers or Hubs and 100 Reading Forums in schools in 10 Local Government Areas of each of Sokoto and Bauchi states. RAN also produced Community Reading Charter which captured the role of stakeholders in improving education in both states through literacy empowerment. All the materials produced for literacy teaching and learning are in both the mother tongue and in English. It is important to note that the establishment of a National Reading Research Center in Ado Bayero University, Kano state, is a result of USAID's investment and promotion of literacy in Northern Nigeria.

The literacy enhancement and achievement impact

RAN, in collaboration with the Anambra State Ministry of Education, has developed the Literacy Enhancement and Achievement Paradigm (LEAP). RAN is using LEAP as a paradigm shift to enhance literacy and overall education standard in Nigeria. LEAP is anchored on five strong pillars:

First, the thorough training of literacy teachers is mandatory because literacy is a discipline demanding multi layered knowledge, attitudes, and skills. Second, in LEAP an appropriate atmosphere must be deliberately created so that learners are surrounded with the atmosphere that entices and excites them to acquire and practice literacy. Group and individual practices for literacy are provided since interaction among members of literacy community is central to its acquisition. In order to ensure that Nigerian learners have the atmosphere that would make them learn and practice literacy, we had to empower classroom teachers in primary and secondary schools on literacy development. The empowerment was achieved by ensuring that the teachers were retrained on the knowledge, attitudes and skills required in literacy education. Since most of our primary and secondary school teachers were not taught literacy and how to teach it during their secondary and tertiary education, they could not effectively and efficiently teach it. LEAP is designed to fill the gap for them by developing their professional skills on literacy education.

We therefore organized a master-trainer workshop for two hundred and twenty Master Trainers drawn from the twenty-one educational zones in Anambra state. We brought in Dr. Alice Headlam to train the Master Trainers on what literacy education. Dr. Headlam is a freelance literacy consultant who was a Chair of the International Committee of the Whole Language Umbrella. She has tremendous experience teaching literacy in the Caribbean. She was assisted by RAN consultants - Professors C. Muodumogu, S. Ekpe, G. Egbe, K. Omenugha, C. Onukaogu, G. Oyinloye, G. Ofudu, and Dr. J. Ahaotu - who as past presidents of RAN and professors of Literacy Education and Communication Studies helped to guide Dr. Headlam to tailor her vast experiences and expertise to meet the needs of the Master Trainers (Fig. 4).

The Master Trainers' workshop lasted eighteen days; 8 h a day except for Sundays. The workshop had extensive hands-on activities. It made provisions for micro teaching by the Master Trainers under the supervision of Dr. Headlam and the RAN consultants. The workshop model of instruction was used, making use of mini lessons and activity periods, sharing and vetting what was practiced. It was therefore easy for the Master Trainers to practice what was taught and see the strengths and challenges in literacy teaching and learning. The workshop modeled how the Master Trainers can reinforce the strengths and address the challenges of their trainees and pupils.

At the end of the master trainer's workshop, they were sent to their respective education zones to replicate and reinforce the knowledge, attitudes and skills they had acquired at the workshop. We set up school clusters at the education zones where the Master Trainers replicated what they had learned. Each cluster consisted of six schools that were within walking distances from each other. The school that was at the center of the cluster was made a meeting point for the teachers in the cluster. The cluster training session lasted for 9 months during which two Master Trainers were sent to each cluster to train all the teachers there.

Each teacher in the state was given a copy of *Tool kit*, a resource material on literacy teaching and learning. It consists of commissioned articles on literacy teaching and learning written by veteran literacy education consultants drawn from Nigeria, the UK and the USA. Each teacher in the state was also given a copy of *Teaching Reading in Nigeria: A guidebook to theory and practice* edited by



Fig. 4 Dr. Headlam with some participants at the maiden LEAP workshop. Courtesy: RAN Executive album.

Professors C. Onukaogu, A. Arua, and T. Jegede. The book had commissioned chapters on various aspects of literacy written by veteran literacy specialists within and outside Nigeria. The book provides current and proven trends in literacy teaching and learning, which could empower the classroom teacher to make a positive difference in literacy teaching and learning in Nigerian.

Each teacher and each pupil in the state was given a literacy journal. The journal promoted variation of classroom interaction patterns—teacher/teacher, teacher/learner, and learner/learner. The journal thus provided a wonderful opportunity for learners and teachers to read, listen to what is read, write about what they see, hear, and above all interact as a community of literacy learners and teachers.

We had resource centers in each of the twenty-one education zones in Anambra state, each equipped with state-of-the-art resource materials on literacy teaching and learning. Each teacher in any cluster has access to the resource materials and could borrow books from there. We also had thirty-six monitors who supervised classroom teachers in order to ascertain the strength of the teachers and the challenges they faced in implementing what they had learned during their training at the cluster level. The monitors were trained to enable the teachers to overcome their challenges as well as reinforce their strengths. The monitors reported to the RAN consultants at the project headquarters.

From the report of the monitors we found that teachers were too pre-occupied with testing and so we brought in Professor Rob Tierney from the University of British Columbia, Canada to train the Master Trainers on Portfolio Assessment in a five-day intensive workshop (Fig. 5). At the end of the workshop, the Master Trainers proceeded to the clusters to replicate the training. Similarly, the report of the monitors, showed that teachers were deficient in the handling of challenged learners. We therefore had to bring in Dr. O. Abiodun-Ekus of the University of Reading, UK, to address the problem of teaching literacy to challenged learners in a three-day intensive workshop for the Master Trainers.



Fig. 5 Prof. Rob Tierney and Prof. Emeka Onukaogu (fourth & fifth from the right) with some RAN Project Consultants. Courtesy: LEAP project album.

We were able to make both teachers and learners participate actively in literacy teaching and learning through the following channels:

- (1) Competent classroom teachers who replicated what had been modeled in the Master Trainer's workshop;
- (2) Extensive resource materials which teachers had access to freely and fully; and
- (3) Effective patterns of interaction: teacher/learner, learner/learner and learner/teacher classroom interactions that were based on best practices and proven interactions strategies.

The third pillar of LEAP is achievement in which learners are encouraged to have as diverse opportunities as possible in using literacy. Before LEAP was introduced as a paradigm shift for the teaching of literacy and reading in Anambra state, the teaching of reading and literacy there was a nightmare to pupils. Reading was poorly taught and whole literacy was not taught to the pupils. After reading a given text, pupils were required to answer questions on the texts they read. It did not matter whether or not the texts were within the readability levels of the pupils and students or whether the contents of the text were culturally relevant to them. The teachers showed the learners their poor performance at the task/test. The reasons why the pupils performed so poorly in the tests were hardly explained to them. In addition, they were never taught how to read and comprehend texts generally. It was assumed by the teachers that once pupils and students can decode the words in a text, they were literate, and could therefore read and comprehend any text placed before them. The teachers were not familiar with the components of reading, the levels of reading, the levels of reading comprehension and the readability levels of the learners. The teachers were ill-equipped on how to teach reading and measure the performance of students and pupils when they read. The pupils were demoralized by their achievement and did not see any link between their reading/literacy lessons and their mathematics, science, history, and everyday life; they only saw reading and literacy in the context of English language.

When LEAP was introduced as a paradigm shift in the reading and literacy lessons, the narrative changed. Story-telling and retelling were introduced into reading and literacy lessons. In LEAP the teacher had the responsibility of searching for interesting and exciting stories. Using the principle of drama, the story would be told the pupils and they would be asked to retell the story orally and were never discouraged, derided, or criticized during an oral retelling session. When they had shown mastery of oral retelling, they were in practice sessions shown how to retell the stories in writing. Furthermore, the Directed Reading and Think Approach (DRTA) was used in enthusing the pupils to enjoy reading, listening, and writing. Exciting stories were read aloud to the pupils and students. They were asked such questions as—what do you think would happen next? Why do you say so?—that would make them participate in the Read Aloud. Such questions motivated them to predict what would happen next when the reading stopped at certain intervals.

The various methods adopted in LEAP allowed the students and pupils to participate actively in the reading and literacy lessons. Texts were also taken from the mathematics, science and technology textbooks. Using very exciting methods, such texts were jointly read by the pupils and their tasks jointly performed by them. The fun which the students derived when such texts were read aloud together in the class and also analyzed together made the pupils not only enjoy the lessons but also looked forward to having them. We took time to teach and model text processing strategies so that the pupils could read easily at the instructional and independent levels. In the area of comprehension, we modeled interpretive, critical, and creative comprehension strategies. We used diverse kinds of texts to model what we wanted the students to master. The use of journals for interactions among the learners and with their teachers was very interesting to them; fluency was emphasized rather than too much emphasis on grammatical accuracy. Later, the pupils were able to write their stories and autobiographies which we desktop-published for them (Fig. 6).



Fig. 6 Samples of learners writing that were desk-top published. Courtesy: LEAP Project Album.

One interesting hallmark of LEAP was that it de-emphasized testing. We adopted the Portfolio Assessment model in collecting information regarding the performance of each pupil. We did not compare the performance of any pupil with that of others. We monitored the performance of each pupil at the commencement of the year, at the mid-year and at the end of the year, based on the data we collected from the portfolio of each pupil or student. Each student also had a strong input regarding how s/he was assessed. At the end of each year, it was celebration galore as each learner has something to celebrate.

The fourth pillar of LEAP is a blend of teacher centeredness with learner centeredness, giving rise to learning centeredness. Thus, in LEAP both process and product synchronize into learning-centeredness. Three significant pendulum shifts are identifiable in general literacy and language teaching methodology. First, in the 1950s and 1960s the literacy teaching and learning pendulum swung to the extreme left where the teacher was the most significant element in literacy teaching and learning. This was a teacher-centered swing which emphasized the teacher as the center in literacy teaching and learning. The learner was treated as a tabula rasa; an empty slate that had nothing to offer in literacy learning and teaching. Given the overbearing and towering image of the teacher, what mattered to the learner was “what my teacher says,” while what mattered to the home of the learner was “what did your teacher say?” However, from the 1970s, there was another pendulum swing from the far left to the far right. This time the swing emphasized the learner. Literacy teaching and learning strove to be learner-centered. The emphasis was no longer on who does the teaching but on who does the learning. However, in LEAP there is now another pendulum swing; this time to the middle. This pendulum swing draws attention to the process of learning. Thus, while the teacher does the teaching and the learner does the learning, LEAP takes the position that if we can understand how the learner learns and uses literacy, we are more likely to address the needs of the learner. Learning-centredness is primarily an amalgam of teacher-centredness and learner-centredness. Thus, while learning-centredness draws attention to the process of learning, the other two emphasize the product of learning from the position of the teacher or from the position of the learner. For optimal performance in literacy teaching and learning, LEAP contends that the three approaches must symbiotically interplay and interrelate.

The fifth pillar of LEAP is celebration. Celebration is infused into LEAP in order to make the learner and the teacher realize that the society treasures their efforts at literacy learning and teaching. That is why at every point of learning and teaching both the learner and teacher are celebrated. That is also why the climax of celebration in LEAP is a Literacy Festival where the works of the literacy learners, teachers and literacy stakeholders at the home, the community, and school levels are celebrated (Fig. 7).

The efficacy of LEAP is reflected in the pilot study carried out in three local government areas (LGAs) of Anambra State—Orumba South, Awka South and Anambra East. The 7500 students who were taught during the pilot implementation of LEAP were tested in English, Mathematic, Basic Science and Technology. They had an average score of 30%, 25%, 25% and 20% respectively in the pre-test. Half way in the program, a mid-course test was administered to the students. They scored 60%, 60%, 50%, and 49% respectively. The post-test was administered at the project completion. Trainee performances in the four subjects were 80%, 85%, 70% and 80% respectively.

The impact of LEAP is felt within and outside the country. For instance, in 2018, Anambra State students won the gold medal in the Silicon Valley, the United States of America global competition of students with their production of an appliance for identifying fake drugs (Fig. 8). In 2019 Anambra State students won the bronze medal in international Festival of Science and Technology (IFEST). Since 2016, Anambra state students have appeared among the three best states in West African School Certificate



Fig. 7 Celebrating excellence with pupils who won prizes at the Literacy Festival From the left: Prof. Omenugha (Honorable Commissioner of Basic Education, Anambra state and contributing author to this chapter) Executive Governor of Anambra state, Mr. Willie Obiano, his wife, and Prof. C. E. Onukaogu (Chairman, Board of Trustees, RAN and lead author of this chapter) with three pupils who received prizes at a Literacy Festival. Courtesy: Prof. C. E. Onukaogu's Library.



Fig. 8 The Silicon Valley competition: Anambra State gold medal winners receiving their awards. Courtesy: Prof. C. E. Onukaogu's Library.

Examinations in Nigeria. Moreover, the state for three consecutive years won gold in the national debate championship organized by the Presidential Schools Debate of Nigeria.

One of the outcomes of LEAP in Anambra State was the improvement in the reading levels of the students. Before the commencement of LEAP, data showed that 76% of the students read at the frustration level. At this level, the students are able to decode some words in the text but cannot comprehend the text even if they receive assistance from a teacher. Only 14% of students read at the instructional level. At this level the students have some challenges in comprehending the text. However, with assistance from their teacher they can overcome the challenges. The data showed that 10% of the students read at the independent level. At this level the students can comprehend the text without any help from their teachers. LEAP engendered a reversal of the reading levels of the students and only 33% read at the frustration level, 47% at the instructional level and as many as 43% at the independent level after the project.

Given the successes of LEAP, the Anambra State Government with funding from the World Bank, collaborated with RAN to implement LEAP in all the primary schools in Anambra State. Also, the Ekiti State Government in collaboration with RAN is implementing LEAP in all the primary schools in Ekiti State in South-West Nigeria.

Conclusion

This chapter has traced the trajectories of literacy in Nigeria and how entwined it is with politics and other socio-cultural forces. A number of issues stand out in the discourse. First, the concept of literacy in Nigeria has undergone numerous changes—from oracy as an indigenous form of literacy to the traditional 3Rs (Reading, Writing, & Arithmetic) of Western literacy. The 3Rs, an autonomous model of literacy has also been interrogated, with the new literacies emerging. The new literacies are of the ideological model that derives from a wider perspective of intellectual and social perspective of learning, knowledge, attitudes, and skills. Second, in all these forms of literacies, politics has continued to play a role in suppressing or enhancing literacy. The politics of suppression is evident in literacy practices in Northern Nigeria where only recitation of Koran in Arabic was taught to the pupils—giving the people no opportunity to read, write or calculate (even in Arabic). This has kept the people under the dominion of the fiefdom. The long-term effect is that in the present times, the Northern Nigerians have remained predominantly illiterate; having not been exposed to literacy for life. This has also resulted in Northern Nigerians hating Western education and further remaining predominantly illiterate. The Christian missionaries in teaching the 3Rs did so, not in the MT, but in English, which became a language of domination of the local people. The colonial government used literacy to socially control and repress the people, making them submissive.

One would think that with independence and the taking over of schools from the missionaries by Nigerian government, a different narrative would emerge. Not only did the story not change, it appeared worse as many of the schools now put in the hands of government suffered neglect, dilapidation, and an almost total collapse of education and literacy. The government failed to give education the priority it deserves through poor policies, inadequate funding, mismanagement, and inability to address the mounting challenges and problems in education. Literacy was not prioritized as a subject to be taught in schools. The result is that teachers are ill prepared to teach the reading component of literacy and many students and pupils continue to read at frustration level. There have been attempts to bring new models of education such as Public-Mission model as seen in Anambra state of Nigeria. In this model, while the mission is in charge of administration of schools, the government pays the teachers. It is hoped that the model would enable the missions to bring back the golden era of education in Nigeria where teachers attended summer schools and libraries were functioning. The challenge has been that the government fails in the crucial responsibilities of effectively funding and supervising education, including teacher welfare and continuous professional development.

The Nigerian educational sector requires a political will and consistent, massive budgetary allocation by the government to improve it. Ebi and Ubi's (2017) analysis submitted that an implementation of UNESCO's recommended 26% expenditure on education in Nigeria would have had 19-times the impact on access to primary school enrollment, more than 2-times the impact on secondary school enrollment and 9-times the impact on access to tertiary education, and create an average of 10-times greater access to overall education in the country. These implementations could redress the decades of neglect and reposition formal education to desirable standards in the country. The efforts of government at various levels to provide free basic education and recent policy adjustments such as the Ministerial Strategic Plans (2018–2022) hold potentials of improving education in Nigeria if they are effectively implemented. ASUU has also relentlessly engaged the government and championed the cause of improvement in Nigerian universities. Also, RAN is collaborating with the World Bank, State Ministries of Education, State Universal Basic Education Boards and International Agencies like the USAID and DFID to highlight the central and pivotal role of literacy in formal education in Nigeria. Given the increasing interventions and many education initiatives in Nigeria, such as the World Bank's Better Education Service Delivery for All (BESDA), the USAID/DFID's NEI Plus Initiative, and others, there is hope of improvement in education and its literacy component in the country. Such improvement must take cognizance of the changing times which herald increasingly complex literary demands.

References

- Achebe, C., 1962. *Things Fall Apart*. Heinemann, London.
- Adedokun, M.O., Kayode, G.M., 2019. Literacy and improved living in Nigerian communities. *Sci. Res. J. VII (VIII)*, 91–99.
- Adedokun, M.O., 2008. *Introduction to Adult Education*. Honey Crown Educational Publishers, Lagos.
- Ahaotu, J.O., Tobiahs, I.R., 2020. A survey of gender politics and the use of English among students of the Department of English Studies, University of Port Harcourt. *Int. J. Lang. Lit. Gend. Stud.* 9 (1), 41–54.
- Ahaotu, J.O., 2018. Multiple roads to literacy: a professional association perspective. *JELTAN* 8, 97–108.
- Ajayi, A., 1965. *Christian Missions in Nigeria*. Longman, London.
- Aridegebe, O.E., Makinde, S.O., 2014. Literacy for Good Governance in Nigeria: Foundational Development. Paper Presented at the 14th Biennial National Conference of the Reading Association of Nigeria: 11–14th November, 2014 Ekiti State University Ado Ekiti, Nigeria.
- Ayodele, J.B., Adedokun, M.O., 2012. Towards achieving functional and adult literacy in Nigeria. *Eur. Sci. J.* 18 (5), 209–218.
- Dechant, E.V., 1977. *Psychology in Teaching Reading*. Prentice Hall Inc, London.
- Ebi, B.O., Ubi, P.S., 2017. Education expenditure and access to education: case study of United Nations Educational, Scientific and Cultural Organization Declaration in Nigeria. *Int. J. Econ. Financial Issues* 7 (5), 290–298.
- Egbe, G.B., Ayeni, A., 2019. Learning in English: the dilemma of the Nigerian child. *JELTAN* 8, 1–13.
- Egbe, G.B., 2014. Language preference among Nigerian undergraduates and the future of English. *Int. J. Engl. Lang. Transl. Stud.* 2 (4), 52–65.
- Egbe, G.B., 2019. *Of New Englishes and New Literacies*. Veritas University Abuja, Inaugural Lecture Series No.1. Veritas University Press, Abuja.
- Ethnologue, 2018. Nigeria. From: <https://www.ethnologue.com>. (Accessed 6 December 2018).
- Fadoro, J.O., 2013. The Nigerian language policy: theory or practice? In: Ndimele, O.-M., Yuka, L.C., Ilori, J.F. (Eds.), *Issues in Contemporary African Linguistics: A Festschrift for Oladele Awobuluyi*. M & J Grand Orbit Communications Ltd, Port Harcourt, pp. 29–38.
- Fafunwa, A.B., Macaulay, J.I., Sokoya, F.J.A., 1989. *Education in Mother Tongue: The Ife Primary Education Project*. University Press Ltd, Ibadan.
- Fafunwa, A.B., 2004. *History of Education in Nigeria*. NPS Educational Publishers, Ibadan.
- Famogbiyele, O.T., 2017. Nigerian senior secondary school students' attitude and motivation towards English language learning: case study of Northern and Southern geopolitical zones. *Asian J. Educ. Res.* 5 (4), 9–29.
- Federal Ministry of Education, 2018. *Education for Change: A Ministerial Strategic Plan (2018–2020)*. Abuja.
- Gaff, H., 1987a. *The Legacies of Literacy: Consistencies and Contradictions in Western Culture Society*. Bloomington, Indiana.
- Gaff, H., 1987b. *The Labyrinths of Literacy*. The Palmer Press, New York.
- Godonoo, P., 1991. The perceived role of literacy and its attendant problems in Nigeria. *Ufahamu* 19 (1), 80–91. Retrieved January 2020 from: <https://escholarship.org/uc/item/8vg3662x>.
- Imam, H., 2012. Educational policy in Nigeria: from colonial era to the post-independence period. *Ital. J. Sociol. Educ.* 1, 181–203.
- Inglis, F., Aers, L., 2008. *Key Concepts in Education*. Sage, London.
- Jemkur, J.F., 1999. *The Nok Culture Art in Nigeria 2,500 Years Ago*. University of Michigan: Prestel Publishers.
- Kolawole, C.O.O., 2009. The state of reading in selected secondary schools in Oyo State. *Afr. Res. Rev.* 3 (1), 388–398.
- National Education Research and Development Council, 2013. *National Policy on Education*, sixth ed. NERDC Press, Lagos.
- National Geographic Society, 2020. The Nok Culture. Retrieved July 2020, from: <https://www.nationalgeographic.org/encyclopedia/nok-culture/>.
- Ogbondah, L., 2002. *Historical Foundation of Western and Nigerian Education*. Harey Publication Coy, Port Harcourt.
- Omojuwa, R.A., 1984. A review of the reading problems in primary schools in Nigeria. *Lit. Read. Niger.* 2 (9), 97–207.
- Onukaogu, C.E., 1998. Functional and development-oriented literacy for sustainable national development. In: Obidi, S.S., Afolabi, E.R.I., Adelabu, M.A., Kobiowu, S.V. (Eds.), *A Book of Readings on Environment and Sustainable Development*. Ife Society of Education, Ife, pp. 270–284.
- Onukaogu, C.E., 1999. Effective and efficient ESL and Nigerian mother tongue literacy instruction: can literacy be a catalyst? *Lang. Cult. Curric.* 12 (2), 143–155.
- Onukaogu, C.E., 2008. Bilingualism and the attainment of sustainable development in multilingual Nigeria. *Forum Publ. Pol.* 1–20.
- Onukaogu, C.E., 2010. Towards the attainment of sustainable development of the Nigerian economy through the instrumentation of language and literature. *J. Sch. Lang.* 5, 1–41.
- Onukaogu, C.E., 2012. Electoral Literacy Empowerment and the Attainment of Sustainable Development: Quo Vadis Nigeria? Obafemi Awolowo University Ile-Ife, Inaugural Lecture Series 248 Obafemi Awolowo University Press, Ife.
- Opeyemi, A.O., 2018. Literacy structure, systems and policies in Nigerian primary education: a need for restructure. *Lit. Read. Niger.* 17 (1), 150–154.
- Orisawayi, D., 2006. Language, literacy, and national development. In: Ndimele, O.-M. (Ed.), *Language and Culture in Nigeria: A Festschrift for Okon Essien*. n. p.: NINLAN in collaboration with M & J Grand Orbits Communications Ltd, pp. 139–145.
- Oroma, A., Wosu, J.I., 2018. History of education and development of Nigeria society: an analytical approach. *Int. J. Res. Innov. Soc. Sci.* VII, 163–169.
- Oyediji, L., 2007. Literacy for governance in Nigeria. *Lit. Read. Niger.* 6, 31–37.
- Rupp, N., Ameje, J., Breunig, P., 2008. New studies on the Nok culture of Central Nigeria. *J. Afr. Archaeol.* 3 (2), 283–290. <https://doi.org/10.3213/1612-1651-10056>. Retrieved June 2020 from: https://www.researchgate.net/publication/228965010_New_Studies_on_the_Nok_Culture_of_Central_Nigeria.

- Street, B., 2006. Autonomous and Ideological Models of Literacy: Approaches From New Literacy Studies. Retrieved May, 2020 from. <https://www.Semanticscholar.org/paper/Autonomous-and-Ideological-Models-of-Literacy>.
- Stromquist, N.P., 2008. The politics of adult literacy: presumed and real effects. *Int. Multidiscip. Res. J.* 2, 88–101.
- Swain, M., 2006. Languaging agency and collaboration in advanced second language proficiency. In: Brymes, H. (Ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. Continuum Inc, London, pp. 95–108. cornell.edu/events/past/2011-2012/papers11/swain.pdf.
- Trimingham, L., 1962. *Islam in West Africa*. Cambridge University Press, Cambridge.
- Trudell, B., 2018. *Language and Education in Nigeria: A Review of Policy and Practice*. n. p.: British Council and UNICEF.
- UNESCO, 2012. Major Push for Adult and Youth Literacy in Nigeria. Retrieved May, 2020 from. http://www.unesco.org/new/en/member-states/single-view/news/major_push_for_adult_and_youth_literacy_in_nigeria/.
- UNESCO, 2019. Literacy. Retrieved May, 2020 from. <https://en.unesco.org/themes/literacy>.
- UNICEF, 2015. Education. Retrieved May, 2020 from. <https://www.Unicef.org/Nigeria/education>.
- UNICEF, 2019. Education. Retrieved May, 2020 from. <https://www.unicef.org/nigeria/education>.
- Unoh, S., 1980. Reading problems in secondary schools: some observations and research findings. *J. Lang. Art Commun.* 1 (1), 30–41.
- Wa Mberia, K., 2015. The place of indigenous languages in African development. *Int. J. Lang. Ling.* 2 (5), 53–60.
- World Bank, 2017. *Better Education Service Delivery for all Operation (BESDA)*. Report Number: 115391-NG.

Community-centered Indigenous language recovery, restoration, revitalization, and renewal

Candace Kaleimamoowahinekapu Galla^{a,b}, Amanda Holmes^d, Joshua Schwab Cartas^c, and William Rowluck^a, ^a Department of Language & Literacy Education, Faculty of Education, The University of British Columbia, Vancouver, BC, Canada; ^b Institute for Critical Indigenous Studies, Faculty of Arts, The University of British Columbia, Vancouver, BC, Canada; ^c Division of Art History and Contemporary Culture, Art Education, Nova Scotia College of Art and Design, Halifax, Nova Scotia, Canada; and ^d Independent Scholar, United States

© 2023 Elsevier Ltd. All rights reserved.

Co-authors as Indigenous scholar-practitioners	786
Lighting the way home	787
Elder praxis, intergenerational knowledge relations, and Indigenous language restoration, renewal, and resurgence	789
Intergenerational knowledge relations	789
Elder praxis	789
Indigenous language restoration, renewal, and resurgence	790
Indigenous language curriculum: issues and challenges	790
Media and technology	791
Documentation	792
Language-enabling technologies	792
Information and communication technologies	792
Technologies that support language acquisition	792
Social media platforms	792
Technology as a double-edged sword: considering benefits and drawbacks	793
Collaboration, networking & resources: engaging in Indigenous language recovery, restoration, revitalization, and renewal	793
Learning from and through relational exchange	793
Mohawk and Hawaiian	793
Hawaiian and Blackfoot	793
Language learning	794
References	795
Relevant Websites	796

Co-authors as Indigenous scholar-practitioners

As Indigenous scholar-practitioners from four distinct Indigenous communities who engage in the work of our traditional homelands—Hawai'i, Kanien'keha:ka, Lytton First Nation, Isthmus Zapotec—and in language revitalization efforts, we have many responsibilities that shape/determine how we participate in the various domains of which we are a part (e.g. the academy, schools, communities, cultural practices). It is important for us to articulate that although we have been educated and continue to work in Western institutions, our knowledge base is grounded in our respective Indigenous ways of knowing and being. We collectively continue to live and practice our lifeways even as we each currently reside away from our traditional homelands.

To privilege and promote our Indigenous languages, we have chosen to use our languages throughout the chapter (sometimes without a translation requiring the reader to make meaning through context, inquiry, research, and embodied understanding), employing and emphasizing Indigenous languages amidst the dominant language of English. We offer relevant websites to extend readers' understanding of the word or phrase. However, please note that these resources should be used responsibly with the utmost respect, as they are for the benefit of Indigenous Peoples of specific communities, first and foremost. Only when the language and knowledge are used *in relationship with* Indigenous Knowledge Keepers and Holders can one *begin* to understand the complexities of a concept embedded in Indigenous epistemologies, axiologies, and ontologies. We also recognize that by using a colonial language we are able to reach a much broader audience. Part of Indigenous Peoples' linguistic and cultural re-generation and resurgence is understanding what our kuleana (right, privilege, concern, responsibility) is to our homelands, and how we are accountable to our places, lands, and peoples.

Aloha. 'O Candace Kaleimamoowahinekapu Galla ko'u inoa. He Kanaka Hawai'i a Pilipino au. No Pāhala mai au ma ka moku-puni nui o Keawe. Ke a'o mai nei au i ka 'ōlelo Hawai'i i mea e ho'onaauao ai ia'u iho ma ke 'ano he Hawai'i a me ka ho'omau 'ana i ka 'ike Hawai'i no na hanauna e hiki mai 'ana. He poloheka hope ma ke kula nui o Kolumepia Pelekānia ma Wanekouwa i Kanakā i ka Hale Ho'ona'auao Ha'awina Kālailai 'Ōiwi (Kumu Pāheona) a me ka Māhele 'Ōlelo a Ho'ona'auao Heluhelu (Kumu Ho'ona'auao). Hello, my name is Candace Kaleimamoowahinekapu Galla. I am Hawaiian and Filipino from the island of Hawai'i. I am learning my language to strengthen my identity and to perpetuate the Hawaiian culture for the next generations. I am a professor

at the University of British Columbia in Vancouver, Canada in the Institute for Critical Indigenous Studies (Faculty of Arts) and the Department of Language and Literacy Education (Faculty of Education).

She:kon. My name is Amanda Holmes. I am Kanien'keha:ka on my Mother's side, and Highland Scottish on my Dad's side. I grew up near Onawi:io. I grew up without my language. I have been in the process of returning to and re-learning my language since my mid-twenties. I have had my Clan returned to me and it is hung around my neck. I am Turtle Clan. My community is Kanatsiohareke Mohawk Community in the Mohawk Valley.

Paduxi Na Laya Joshua Schwab Cartas. na naca binniza, Nacá baduhuni Gilberto Cartas Posada ne binilidxe Cartas Arias ne Ranchu Gubiña, ne Na Lilia Guzman Cartas laabee Ixtepec (San Geronimo) ne xtinne' binnilidxe Arias Cartas binni Ranchu Gubiña ne binilidxe de Vancouver and Montreal Canada na riniisi. Hello, my name is Joshua Schwab Cartas. I am of binniza (Zapotec)—the grandson of Gilberto Cartas Posada and family member of the Cartas Arias Family from Ranchu Gubiña, as well as the grandson of Lilia Guzman Cartas from Ixtepec (San Geronimo). I also have family in Vancouver and Montreal Canada where I grew up.

Nlaka'Pamux N'Chowa, Esquesht Pubska, Daquiyantsa Aguawn Kumcheen. I am Billy Rowluck and am a member of the Nlaka'Pamux Nation. My home community is Lytton First Nation, located at the conflux of the Thompson and Frazer Rivers in British Columbia. I entered public education in 1996 as a teacher's assistant and went on to earn a bachelor and masters of Education from the University of Saskatchewan. My experience has ranged from classroom teacher to principal in two northern schools in Saskatchewan. I am currently a doctoral student at the University of British Columbia in the Department of Language and Literacy Education and am interested in the representation of Indigenous ways of knowing within public-school curriculum.

Lighting the way home

Of the 7000 languages spoken in the world, 2680 languages are at risk (UN, 2018) of no longer being spoken or passed down to future generations. By contrast, "About 97% of the world's people speak about 4% of the world's languages; and conversely, about 96% of the world's languages are spoken by about 3% of the world's people" (Bernard, 1996, p. 142). Scholars in linguistic human rights, Skutnabb-Kangas and Phillipson (2010) warn that, "Almost all languages about to disappear would be indigenous, and most of today's indigenous languages would disappear, with the exception of very few ..." (p. 77).

This chapter discusses Indigenous language restoration, renewal, recovery, and resurgence, its spaces, contexts, practices, and possibilities, decolonizing spaces that hold the aspirations of their People for returning to their language, as their language lights the way home to Indigenous understandings of relationality, the interbraiding of cultural coherence, collective narrative memory (McLeod, 2007), and intergenerational relationship. Indigenous Peoples hold ancient connections to their languages and their lands; it is in relationship with their lands, waters, and natural world that their language and their own existence as a People have been formed and, in expressions of deep reciprocity, have given each other life. Widespread among Indigenous Peoples is the strong feeling and love they hold for their language, languages often understood as a sacred gift from their Creator, connecting them to their Ancestors, to their lands and lifeways, to their teachings and ethical principles, to their knowledge systems and ways of knowing-being since time immemorial; these are understandings that define *who they are* as a People.

From within this context of longevity on their lands – the intricate, intimate textures that form vast webs of relationship particular to home, land, language – the experience of colonization has wreaked havoc on Indigenous Peoples, languages, and lands, on the cycles and patterns of relationship that have maintained and sustained Indigenous "survivorship" (Vizenor, 1999) for generations. Indigenous communities continue to work with determination to re-awaken, restore, and renew their languages out of long histories of targeted, strategic language repression that has been a core feature of the settler-colonial project of Indigenous elimination and replacement (Wolfe, 2016). As Wolfe (2016) points out, settler colonialism destroys in order to replace, aiming for "more than summary liquidation, it strives for the dissolution of Native societies" (p. 33). As state policies pursued the dissolution of Native societies through coercion, control, and domination, the eradication and replacement of Native languages proved to be an effective tool in the settler colonial toolkit. Education became an increasingly important arena for settler colonial states to deal handily with "the Indian problem," as physical genocide became tempered by cultural genocide, with the "friends of the Indian" arguing for cultural genocide beginning in the latter half of the nineteenth century in the United States, with Canada soon following its lead. Recalling the blunt words of Philadelphia lawyer Henry Pancoast, "We must either butcher them or civilize them, and what we do we must do quickly" (Adams, 1995, p. 2).

In 1879, the Carlisle Indian Industrial School, which was to become the prototype of Indian boarding schools (and state-sponsored violence in the form of not only brutal physical, sexual, emotional abuse, but also of ethnocide and linguicide against Native children, their families, and communities), was founded under the direction of Colonel Richard Henry Pratt, a veteran of both the Washita massacre of the Cheyenne and the Red River campaign against the Comanche, Kiowa, Southern Cheyenne, and Arapaho, and in charge of the Indian prisoners at Fort Marion in Florida. Pratt participated in multiple contexts in the eradication and extermination of Indigenous people; he has become known for his now infamous operating motto: "Kill the Indian in order to save the Man" (Smith, 2009, p. 5).

In her work on the politics of language, language ownership, and the teaching of English to Native students during the late nineteenth century, Ruth Spack expands on settlers' racialized hierarchies and their technologies of linguistic and cultural replacement, where "Anglo Americans believed in the superiority of their own language and culture, and almost from the beginning they began to develop educational or civilization programs to impose the English way of life on the Native population" (Spack, 2002, p. 3). Soon, boarding or residential schools would become widespread across the U.S. (Lomawaima and McCarty, 2006) and Canada (Truth and Reconciliation Commission of Canada, 2015), as well as in Mexico (Dawson, 2012), India (Gupta and Padel, 2020), Australia (Commonwealth of

Australia, 1997), New Zealand (Walker, 2016), and other regions. While Indigenous children and communities have always found ways to resist enforced schooling (Lomawaima and McCarty, 2006), it is critical to remember that boarding schools were an act of aggression by settler states as “governments intentionally removed indigenous children to institutions as acts of colonial control, not assimilation” (Jacobs, 2006, p. 202), “as an official strategy of ... subjugation” (p. 204).

In international arenas, Indigenous Peoples have fought to have their fundamental rights as Peoples recognized by international bodies, including their rights to their lands and waters, languages, religious freedom, education, and traditional forms of governance. Finally, in 2007 after many long years of intense pressure, Indigenous Peoples from around the world succeeded in gaining the passage of the United Nations Declaration on the Rights of Indigenous Peoples by the UN General Assembly. Although it is beyond the scope of this chapter to provide an exhaustive summary of the national and international language policies and frameworks that impact Indigenous languages, we include several federal US and Canadian policies that attempt to address these countries’ long-standing practices of targeting Indigenous languages for extermination: in Canada, the Indigenous Languages Act (2019); in the United States, the Native American Languages Act (1990), the Esther Martinez Native Languages Programs Act (2006); and the long overdue repeal in 1986 of the Hawaiian Revised Statutes of 1896, a policy that prohibited Hawaiian language medium instruction and contributed to the widespread loss of the Hawaiian language, to name but a few.

The aim of this chapter is to provide a broad overview of Indigenous language renewal and regeneration from diverse perspectives, with resources that the reader might find useful. It should be noted that in any discussion of Indigenous languages, knowledge systems, and cultural practices that have been subjugated by colonization, deeply embedded differentials of power are also present. We are deliberately choosing the language of “community-centered” to describe and discuss Indigenous language renewal efforts that are emergent from within Indigenous communities – they are community-driven and community-led, where the needs of the community are centered (Evans, 2004; McDonald, 2003). Given long, extractive histories, where the tendrils of the academy have reached into Indigenous communities, this is a profoundly critical conversation (Battiste, 2013; Smith, 1999, 2012).

This chapter not only intends to disrupt the academy’s assumption about the necessity and centrality of Western academic research to Native communities’ efforts at revitalizing their languages, it also challenges the academy’s authority to define, delimit, and control the practice, theorizing, envisioning, and aspirations of Indigenous communities for the futurity of their languages. While the impact of colonization has indeed caused enormous cultural and linguistic disruption and upheaval for Indigenous Peoples, communities continue to fight for the continuance of their languages and cultural ways of knowing and being, defined on their own terms, within their own contexts of survival, persistence (Nicholas, 2014), resilience, resistance, self-determination, collective transformation, and resurgence (Galla et al., 2014). This chapter foregrounds self-determining, decolonizing, and transformative research methodologies, epistemologies, and practices of Indigenous language restoration, renewal, and regeneration from community-centered perspectives and orientations. We hope it chips away at the imbalance of power between Indigenous communities and Western academia in the field of language restoration and renewal. We explicitly and self-consciously toss the false Western construction of objectivity to the wind by supporting, promoting, privileging, highlighting, and amplifying Indigenous community language and cultural reclamation work, aligning ourselves with Indigenous community-based, community-centered desires, hopes, struggles, priorities, aspirations, and dreams.

Indigenous Peoples know well that the language we use holds power and contributes to the creation of the world in which we live. Thus, it is even more relevant to be explicit about the language choices we will be using in this chapter, and some of the context and meanings we intend by the particular choice of language we use. A focus on “extinction” characterizes mainstream discourse with regard to Indigenous languages. We align ourselves with Indigenous language advocates, scholars of the community and of the academy, who are contesting and resisting Western discursive framings of Indigenous languages in terms of *morbidity*, *extinction*, and *death*, including the framing of “revitalization” itself. We are thus following the lead of community as they re-frame Indigenous language “revitalization” as language awakening (Hobson, 2010), restoration, recovery, renewal, regeneration, and resurgence. Indigenous communities continue to resist and re-shape the discourse around their language in ways that are more relevant, meaningful, and coherent from within their own epistemologies and knowledge systems.

A few notes on language and discourse must be addressed as we set out on this collaborative thinking to bring some of the historical contexts, diversity, activity, and possibilities of Indigenous language renewal and regeneration to this work. We recognize that all too often, colonizing discourse frames Indigenous languages, shaping the perceptions and attitudes, the “orientations,” (Ruiz, 1984) around “language endangerment” (Dwyer et al., n.d.) that define and delimit Indigenous languages and their revitalization. We wish to cultivate some common ground with which to understand the contexts of Indigenous “language revitalization” or language awakening. We begin with the working definition of Indigenous Peoples used by the United Nations Permanent Forum on Indigenous Issues (UNPFII) as: “Practicing unique traditions, they retain social, cultural, economic and political characteristics that are distinct from those of the dominant societies in which they live. Spread across the world from the Arctic to the South Pacific, they are the descendants – according to a common definition – of those who inhabited a country or a geographical region at the time when people of different cultures or ethnic origins arrived. The new arrivals later became dominant through conquest, occupation, settlement or other means” (UNPFII Factsheet, n.d.).

Further, while acknowledging that efforts at Indigenous language renewal and regeneration emerge from within the particularities of Indigenous places and contexts, we join with Duarte and Belarde-Lewis (2015) who remind us that, “As Kanaka Maoli philosopher Manulani Aluli Meyer asserts, ‘specificity leads to universality.’ Understanding distinctiveness leads us to appreciate how ‘we are all the same, differently’” (p. 678). Indigenous communities continue to work with determination to re-awaken, restore, and renew their languages – and, as we get closer to our languages, there is a distinct sense that our living languages are also at work, waking us up to who we have always been, and who we are meant to be.

Elder praxis, intergenerational knowledge relations, and Indigenous language restoration, renewal, and resurgence

Reciprocity is a matter of keeping the gift in motion through self-perpetuating cycles of giving and receiving ... With their tobacco and their thanks, our people say to the Sweetgrass, 'I need you.' By its renewal after picking, the grass says to the people, 'I need you too.' Mishkos kenomagwen. Isn't this the lesson of grass? Through reciprocity the gift is replenished. All of our flourishing is mutual.

Kimmerer (2013, pp. 165–166).

Indigenous languages hold within them the endlessly renewing cycles of reciprocity and regeneration, carrying ancient teachings, intricately braided knowledge systems, philosophies, principles, and ethical, oral-relational ways of knowing-being that orient and instruct us about our relationships with each other and the natural world, from within the vast “multiverse” (Cajete, 1994, p. 75) of Indigenous contexts of diversity. From within Indigenous knowledge systems and ways of knowing-being, Indigenous Peoples know our languages to be a sacred gift, emergent from within ancestral pasts that are as old as the beginning of Time, giving voice to deep relationships they continue to maintain. It is in our languages that the beautiful and intimate ways we know our lands and waters and the beyond-human worlds, the ways they sound and speak to us, can be heard. Our languages and the knowledges they carry continue to shape our understandings of who we are and who we have been – and guide us to remember who we are meant to be.

Intergenerational knowledge relations

Situated within the particularities of Indigenous universes, these reciprocities of relationship between ourselves, our languages, our Ancestors, the coming generations, and the natural world become illuminated and ignited; and it is through these cycles of renewal and reciprocity that our languages call us into restoring and re-storying our ancient commitment to each other and to each strand of the vast web of relationship to which the People and their languages are connected. Indeed, “All of our flourishing is mutual” (Kimmerer, 2013, pp. 165–166). Indigenous languages tether their People to an intergenerational-relational matrix, to Ancestors and to future generations, and to the regeneration of ancestral memory into today through these relations of longevity and the continuities of intergenerational knowledge relations. Indigenous languages hold and carry the worldviews, teachings, principles, knowledge systems, collective narrative memories (McLeod, 2007), and ethical-relational ways of knowing/being of their People, from deep within their contexts. By restoring, renewing, and regenerating Indigenous languages, the matrix of Indigenous *relationality* is restored, renewed, regenerated. This vast relational web, this multiverse (Cajete, 1994), is the heart center of their People, with each filament connecting to the others through their language, their particular ways of understanding the meanings of who they are and how they are to be in the world. These are the relational worlds that Indigenous languages carry, and it is within these spaces and places of knowing and being as a People that we look to renew and re-generate our languages.

Elder praxis

Indigenous Elders have always held the center of Indigenous life, language, and knowledge, embodying the collective consciousness of their People – and they become ever more important as time passes. It is necessary to clarify the meaning of the term “Elder” when used within Indigenous contexts, as the idea creates some confusion for Western-oriented people, so a brief explanation is needed here. Indigenous Peoples hold a common understanding about the meaning and critical importance of Elders within Indigenous societies, as those who, because of their living and experience, have gained wisdom, widely recognized as spiritual wisdom, which is critical to the life of their People. They hold a revered position in Indigenous societies and are traditionally viewed with great respect. Elders carry the spark that nourish and sustain the social, spiritual, ceremonial, ethical, emotional, physical, intellectual, generational well-being of their People. They are located at a critical nexus of relationship between the generations, a space of exchange and reciprocity between Ancestors and the coming generations. Lakota Elder Rosalie Little Thunder always asserted, “Elders are the mirror of our Ancestors” (Little Thunder, personal communications to A. Holmes). The praxis of Elders (Holmes, 2018) holds the power to invigorate ancient knowledge systems within unfolding contexts of past, present, and future generations. From within their wisdom practices and position of deep respect, Elders hold the possibility for transforming Indigenous community aspirations for language restoration and renewal in culturally coherent ways and self-determining directions. Centering the guidance, leadership, and modeling of Indigenous Elders by re-establishing traditional intergenerational relationships of knowledge production and exchange has had transformative results in places like Kanatsiowhareke Mohawk Community and the work of Mohawk Elder Tom Porter, who has played a critical role in the restoration, renewal, regeneration, and resurgence of the Mohawk language and culture (Holmes, 2018) of the Rotinonhsionni, the Haudenosaunee, or Iroquois Confederacy. The praxis of Elders nurtures Indigenous community language and cultural regeneration efforts, supporting and guiding their communities to reflect deeply about their process and practice, to (re)conceptualize what they are doing, what is important, and how to get there.

Indigenous language restoration, renewal, and resurgence

Restoring traditional Indigenous protocols, processes, and practices of relationality among the generations by placing them at the center of movements for cultural renewal creates resurgent possibilities for the re-awakening of languages. Indigenous language revitalization is reawakening as critical decolonizing work, as communities re-conceptualize and *re-contextualize* Indigenous language recovery, restoration, and renewal and how to go about it from within their own center, placed within their own longevity of contexts, experiences, and practices of meaning, memory, and intergenerational knowledge relations. Indigenous languages mediate an interweaving web of Indigenous worlds and relationships, systems of knowledge, thought, philosophy, ethics, and practice – ways of knowing and being – and it is this understanding that illuminates possible pathways toward the regeneration and resurgence of Indigenous languages, pathways that simultaneously awaken and invigorate Indigenous relationality, where these relationships are lived, known, embodied. Indigenous epistemologies and the lived, everyday practice of relationality are carried within the language in culturally meaningful, deeply contextual, placed ways.

As Indigenous communities revitalize their languages, they are restoring and renewing connections within *their own* places of power, knowing, and deep cultural coherence, looking to those places that have always offered their people clarity and resilience through an intimacy born of longevity and relationship with lands, waters, the natural world, Elders, Ancestors, future generations, collective narrative memory (McLeod, 2007), and the spiraling cycle of intergenerational relations. Remembering, re-imagining, and re-envisioning the ways Indigenous languages root and are rooted within their own relational universe readies the ground for intergenerational renewal and resurgence. In restoring and renewing the relationships between Indigenous People and their languages, lands, ways of knowing, and practices of the collective, re-weaving and re-presence-ing to intergenerational relationships of Ancestors and coming generations, remembering from within the praxis and teachings of Elders, an intricately interconnected matrix of relations is restored. Re-centering intergenerational relations as critical to the process of language renewal and the restoration of Indigenous 'language ecologies' ensures that Elders are placed once again in the center, embodying the reciprocity between the generations and the living of ancestral wisdom ... and holding the transformative possibility of guiding Indigenous languages into resurgence.

Indigenous language curriculum: issues and challenges

Indeed, "to find life" and understand its manifestations within each individual self is the ultimate goal of all education.

Cajete (1999, p. 60).

This section is written with Indigenous and non-Indigenous educators and stakeholders in mind, for all those who have an interest in understanding the complex genocidal history of Indigenous education (Skutnabb-Kangas and Dunbar, 2010). Similar to Indigenous Peoples around the world, experience of education was constructed from within a non-Indigenous system, in a non-Indigenous instructional setting, taught by non-Indigenous teachers, using non-Indigenous resources and curricula. The formal and hidden curriculum reinforces several primary themes: school teachers were non-Indigenous and resources were determined by non-Indigenous people. Hampton (1995) states, "For most Indian students, now as in the past hundred years, Indian education means the education of Indians by non-Indians using non-Indian methods" (p. 6). Indigenous Peoples worldwide have undergone similar educational experiences. From Finland, Russia, China, Africa, Australia, Hawai'i, New Zealand and the United States, Indigenous Peoples have been, and continue to be, subjected to systems of schooling that are determined by dominant governments, religions, or more populous national ethnicities (Smith, 2009). In spite of this, resistance by Indigenous minorities to teach their own knowledge to their children has never abated.

Despite Canada's longstanding attempts to separate and eliminate Indigenous families and communities from their children's education, Indigenous Peoples have challenged these attempts. Indigenous Peoples have stated that,

Unless a child learns about the forces which shape him: the history of his people, their values and customs, their language, he will never really know himself or his potential as a human being. Indian culture and values have a unique place in the history of mankind. The Indian child who learns about his heritage will be proud of it. The lessons he learns in school, his whole school experience, should reinforce and contribute to the image he has of himself as an Indian.

National Indian Brotherhood (1972, p. 9).

This statement suggests that Indian people be directly involved in the creation of school curriculum alongside the Canadian government.

Boarding or residential schools have also been implemented in many regions throughout the world, with differences depending on colonial context. Given the abject failure of residential schools of the 19th and 20th centuries, their operation today in similar institutions around the world is cause for concern. In India, for example, the Kalinga Institute of Social Sciences (KISS) is one of India's large scale tribal boarding schools, which contains striking similarities to the former residential school models that were

designed to remove Indigenous children from their families, languages, and cultural identities in Canada and the United States. Like the Indian residential schools in Canada and the U.S., at KISS,

...most children have to adapt to a new language, usually the dominant state language. Exclusion of tribal languages from most of India's tribal schools has contributed to the diminishing use of several hundred ancient tribal languages, which some scholars have termed 'linguistic genocide' in similar contexts. This is a tremendous loss in that each such language encapsulates a unique worldview.

Gupta and Padel (2020).

Although Indigenous educators around the world have faced great challenges, they have continued to develop language and curriculum from within their own contexts and communities, incorporating locally developed resources that reflect the unique languages and identities of their people. In Aotearoa/New Zealand, Maori communities and educators have asserted their self-determination to educate their children in Māori with the start of Kōhanga Reo or language nest preschools in the 1980s. On the heels of the Māori, and following the amendment of the Hawai'i State Constitution in 1978 which states that the "State shall provide for a Hawaiian education program consisting of language, culture and history in the public schools. The use of community expertise shall be encouraged as a suitable and essential means in furtherance of the Hawaiian education program" (Legislative Reference Bureau, 2022), a small group of educators and parents who made a commitment to revitalize Hawaiian founded the 'Aha Pūnana Leo—a non-profit Hawaiian organization. This was the beginning of the Hawaiian language revitalization movement that modeled language nest preschools after their Māori cousins. Formal education in both settings of Aotearoa and Hawai'i have included Māori and 'Ōlelo Hawai'i language which has extended from infant-toddler programs, preschools, elementary, middle-school, and high-school, and college, university and professional degree programs.

The exclusion of local Indigenous Peoples from institutions of public education as educators, administrators, curriculum developers and the like, remains one of the most challenging deterrents to the implementation of language and curricular resources in schools. The responsibility for locating, designing, and delivering quality curricular resources for schools is often given to classroom teachers, under the direction of regional and federal education mandates. Due to the already burdensome workload of teachers, it is untenable for them to undertake the creation of curricular resources, with the result being that the majority of teachers look for ready-to-use resources. This leaves centralized education authorities, institutions, and/or corporations to design and deliver curriculum. Those who are called to develop these resources often have no experience in creating curriculum, and are frequently unaware of whether the resources they create are culturally relevant or responsive. While faced with formidable challenges in designing and delivering quality language and curricular resources for their people, Indigenous Peoples continue in their efforts to teach their children their languages, histories, and identities. The settler-colonial model of education is generally acknowledged to have failed. Indigenous Peoples need to be able to chart their own educational path – and these paths continue to be embarked upon by Indigenous educators around the world.

Media and technology

"For Aboriginal/Indigenous communities, it has always been traditional to utilize the latest technology...

—Hulleah J. Tsinhnahjinnie (Seminole/Muskogee/Diné).

Incorporating technology to support new strategies that encourage everyday language use and the preservation of ancestral languages, as many Indigenous Peoples, communities, activists, and scholars will tell you, is not a new phenomenon. However, the perception of technology as a "cultural nerve gas" (Krauss, 1992, p. 6) in relation to minority or endangered languages has dramatically shifted over the last several decades to a more optimistic one. Hawaiian language activist Candace Galla (2016) asserts that it is difficult to envisage the "survival of Indigenous languages in the 21st century without the intervention of digital technology" (p. 1137). This change in perception in large part has to do with the innovative ways Indigenous and minority language speaking communities *themselves* are using technologies to strengthen their language and ancestral ways of being (Schwab-Cartas, 2018; Galla, 2018).

Potential benefits of technology are easy to imagine, but they also come with risks and challenges that need to be considered carefully. Of course, it is hard to be prescriptive about the drawbacks and benefits of technology, as what may work for one community may not work for another. The term technology itself "is actually quite vague and could refer to everything from hardware to search algorithms to data formats or user interfaces or many, many other things" (Hieber, 2015, p. 346). There are many devices and applications that fall under the umbrella of technology that are being used by Indigenous communities for language revitalization purposes. Some of the technologies that have been used range from wax cylinder recordings to "analog and digital video and audio recording devices, as well as a constellation of computer, mobile, and Internet related technologies [apps], to capture, store, and make available to future generations important aspects of their languages, arts, and understanding" (Oppenneer, 2011, p. 1).

Many of the technologies, apps and software being used to promote and foster Indigenous language learning are overlapping in capabilities and tend to work in symbiosis, so it can be very difficult to isolate them from one another (i.e., smartphones or tablets). Therefore, to be as clear as possible, we will use the helpful classification structure developed by Hermes et al. (2016), in order to

help the reader navigate the vast number of technologies, as well as to assist readers in determining options that might be the most useful and beneficial option for their language community. (It should be noted that as technology is rapidly evolving, it is possible that some of the technologies mentioned here, especially apps, may no longer exist at the time of publication.) The categories developed by [Hermes et al. \(2016\)](#) are as follows:

Documentation

Technology has been used for over a century to document and preserve endangered languages and cultures. This has included such things as the wax cylinder/mechanical phonograph, voice recorder, camcorder, computer, and mobile devices such as tablets and smartphones. Documentation was once thought to be an activity reserved for linguists or Indigenous communities with the “proper” equipment, but documentation has become a major social activity among everyday mobile phone users, especially youth, who tend to use mobile phones as personal mobile logs ([Stald, 2008](#)), recording via pictures (selfies), sound (audio recording) and video (Snapchat/Instagram/TikTok) their everyday experiences, whether watching a child’s first steps or an interesting event or recording an Elder speaking an Indigenous language ([Schwab-Cartas, 2016](#)). As many scholars have noted, documentation alone cannot save a language ([Hinton and Hale, 2001](#); [McIvor, 2009](#); [Perley, 2012](#)), but it can be a rich resource with which to develop curriculum and teaching materials, while fostering language transmission to future generations ([Littlebear, 1996](#)).

Language-enabling technologies

Language-enabling technologies, on the other hand, enable users to successfully communicate in their Indigenous languages via fonts, adapted keyboards, online dictionaries (from simple definitions to talking dictionaries), unicode, and even through culturally-rooted, culturally-significant emojis, such as Haida emojis. These technologies are typically created by communities, language activists, and/or linguists working alongside Indigenous speaking communities as a means of addressing the fact that most technologies, apps, and software are designed to support dominant languages, rather than Indigenous or minority languages. However, due to ongoing Indigenous language activism and its support in Indigenous communities, major software developers have taken notice. Apple, for example, has included Cherokee syllabary on all their iOS devices (iPads, iPhones, and iPods) since 2010 ([Kemper, 2015](#)). This is one of the many ways that Indigenous Peoples and communities, particularly youth, have found to address technological disparities, while adapting digital technologies and applications to make them relevant to their local communities and lifeways.

Information and communication technologies

Information and communication technologies (ICTs) are an expansive and broad category referring to interactive media ([Begay, 2013](#)), and include a host of various platforms, ranging from radio, digital technology, Email, and databases in the form of YouTube, blogs, and vlogs to videotelephony. ICTs are not specifically designed by or with Indigenous communities in mind, yet they are being used by Indigenous and minority language speaking communities in creative ways to support language learning and revitalization alongside language-enabling technologies. Radio, sometimes overlooked in the Global North, continues to be an important form of technology being used by Indigenous communities around the world, particularly in places where internet connection is limited. For example, radio continues to be used by Binni Cubi—a Zapotec media collective in Southern Mexico—as a cross-generational and affordable way to broadcast not only local community news, but promote Isthmus Zapotec language learning. More recently, due the global COVID-19 pandemic, many Indigenous communities have increasingly embraced video telephone platforms, such as Zoom, Skype, Facebook, and Bluejeans to connect with their communities and learn from Elders and/or fluent speakers.

Technologies that support language acquisition

Social media platforms

Language acquisition technology includes social media platforms such as FaceBook ([Cru 2015](#)), Twitter ([Lillehaugen, 2019](#)), Instagram ([Cassels, 2019](#)), YouTube ([Cru, 2018](#)), and TikTok ([Holmes, 2020](#)), to name a few examples, but it can also include the development and use of specific mobile applications for the purpose of language learning. Apps are now seen as part of a larger phenomenon known as mobile learning (M-Learning), which is a “new type of learning based on mobile technology where students can follow up their education *anywhere, anytime*, and in any *form* using their mobile devices” ([Mohanna and Capus, 2012](#), p. 315). M-Learning allows the user to obtain and/or provide pedagogical content that can be accessed on mobile devices, smartphones, iPods, and tablets ([MacEntee et al., 2019](#)). The pedagogical content of many of these apps includes quizzes that incorporate flashcards, phrasebooks, and talking dictionaries to teach and test vocabulary while the more advanced apps include speech detection software that test pronunciation. An example of these apps is the popular Duolingo, which includes global languages, such as English and Spanish, but has recently included endangered languages such as Cherokee, Hawaiian, and Gaelic. The advantage and benefit of this type of learning via an app is that apps are a great way “to jump-start learning, to reawaken what can be heard, to allow for private practice, to focus on particular confusing conjugations or to investigate unknown constructs or appropriate registers” ([Hermes et al., 2016](#), p. 276).

Technology as a double-edged sword: considering benefits and drawbacks

It is easy to get excited and want to use the latest technology; and yet, digital technologies can also be a double-edged sword. Along with the benefits, there can also be drawbacks that need to be considered carefully by the community, individual, or researcher. The following are several relevant issues to consider before, during, and after using technology. First, no technology is or can be a replacement for an Elder or direct, face-to-face, embodied engagement with a language or an ancestral practice. Technology is, rather, a supplementary tool. For example, an online dictionary app in an Indigenous language can provide the learner with a word, but only an Elder can explain its significance, illustrate the proper context in which it is used, and check the pronunciation of the word. Another matter to consider is rapid obsolescence of technology and apps—the rapid pace and change of technology means that developers and/or communities using technology cannot continually update their technologies or software. The sustainability of these technologies must be considered, along with whether the community has access to internet or sufficient broadband in order to run these websites. Another factor that needs to be considered carefully by each community is how to safeguard and protect itself from the risk of cultural/ancestral knowledge theft, appropriation, or disrespectful usage of sacred knowledge. Further concerns also include risks of screen time on younger children, or other concerns, such as cyberbullying or sexting (Shariff, 2015). This is by no means an exhaustive list of concerns, nor is it prescriptive; concerns will vary from community to community and context to context.

Collaboration, networking & resources: engaging in Indigenous language recovery, restoration, revitalization, and renewal

In this section, we share some resources from Indigenous Peoples, communities, and schools from around the world that can serve as guides as we engage in Indigenous language work. Whether you are just starting your language journey, are looking for resources to support your community, or want to learn from the experiences of other Indigenous Peoples who have found ways that have worked for them in recovering and restoring their languages, we offer examples from our lived experiences.

For many Indigenous communities, intergenerational language transmission has been disrupted, resulting in varying degrees of language loss, shift, repression, and shame. And yet, even with these external and internal challenges, individuals, families, and communities have continued to find ways to live their Indigenous languages by learning cultural practices, attending cultural events, speaking with Elders, accessing archived recordings, and so forth, collaborating within community and with other Indigenous communities and peoples. The exchange of knowledge and language reflects critical Indigenous principles and practices that continue to be enacted, where successes and challenges are shared, best and inspired practices are contemplated, and strategies that will inform our next plan of action are remembered and (re)imagined.

Learning from and through relational exchange

Indigenous Peoples have long navigated ancient pathways of interaction and exchange, forming bonds of relationship with each other across vast distances that are indeed, in the deepest sense, global and transnational. These relational, ethical practices of interconnection exist as a profound resource from which to draw strength, nourishment, and sustenance, as Indigenous communities (re)navigate these well-worn routes through the treacherous waters of settler-colonial narratives, home toward their languages. The interconnections and relationships formed between Indigenous communities have had important and meaningful impacts on language recovery, restoration, revitalization, and renewal. We highlight several examples of these relational exchanges of learning that have emerged between the following Indigenous language communities:

Mohawk and Hawaiian

Sister Dorothy Karihwénhawe Lazore (Mohawk) among others have been credited for bringing attention and greater awareness to the vitality of the Mohawk language in the community (Rice, 2019). Lazore, a first language Mohawk speaker, was instrumental in establishing the Kahnawà:ke Survival School that immersed children in the Mohawk language. In the 1980s, when 'Aha Pūnana Leo ('APL)—a non-profit Hawaiian organization—was in its early stages of setting up Hawaiian language immersion schools, Lazore appeared before the Hawai'i Board of Education to advocate on behalf of the Hawaiian program, which has now become a nationally and internationally recognized model for Indigenous language revitalization. In 2016, the founding members of 'APL, along with the Mālama Honua Worldwide Voyage crew, honored this long-standing relationship that spans nearly 5000 miles and 40 years of dedication, advocacy, and commitment to Indigenous language education.

Hawaiian and Blackfoot

In 1994, nearly a decade after the Hawaiian language revitalization movement and the opening of Hawaiian language immersion schools, Darrell Kipp (Pikuni), founder of the Piegan Institute (Kipp et al., 2015), traveled to a bilingual conference where he met Hawaiian language immersion educators and founders of 'Aha Pūnana Leo. According to Kipp (2000), "They were the first people we ever met who knew what we were seeking." (p. 8). He and other tribal members would visit Hawai'i that year to witness first-hand children speaking Hawaiian language, a great achievement, considering that a decade earlier less than fifty children under the age of 18 spoke Hawaiian as their mother tongue. Their visit allowed them to experience, learn, and ask critical questions of the Hawaiian language educators, and would be a source of inspiration for an all-day Blackfoot language program. Kipp (2000) states,

"The Hawaiians let us stay with them long enough until we felt we were ready. They let us write to them, and they called us when they knew that times were bad" (p. 40). He continues, "Today, we are the best of friends and would not be where we are except for their friendship ... They were mentors, our support, and our guides." (p. 8).

These knowledge exchanges among Indigenous Nations are two examples of the ways Indigenous communities have worked together to support each other in the regeneration of their languages and their goals of creating speakers in the community. Many local, national, and global relationships have been cultivated in the effort to share pedagogical practices, guiding principles, resources, tools, and struggles related to language learning and teaching. As Kipp (2000) states, "There are those moments when you think, 'All I need to do is talk to someone. I got to make sure that I run this past someone.' Know that we'd be happy to be that someone" (p. 40). Extended networks of Indigenous language communities enable collaborative possibilities in language development, programming, and policy, where communities are not forced to "reinvent the wheel" and invent ideas from scratch. By participating together and establishing relational networks, communities are able to maximize their resources, learning from each other what has worked (and not worked) in their respective communities, to help inform and develop their own vision from within their own community context and local needs, aspirations, and priorities.

Language learning

Finding speakers, resources, and materials by which to learn an Indigenous language varies from one language/dialect to another. For those who are (re)learning their mother tongue, there may be additional, very real barriers that discourage this process, such as personal, family, and intergenerational trauma, shame, and repression, as well as language policies and ideologies. For those on a path of language restoration and recovery, we offer some meaningful ways to jumpstart your language journey, whether you are embarking on an individual or family endeavor. Building on the legacy of the International Year of Indigenous Languages (2019), as declared by the United Nations, and moving into the Decade of Indigenous Languages (2022–2032), let us be encouraged by the language memories that might resurface and the language knowledge gained, as these become more familiar with time.

1. **Names carry power** – Introduce yourself and others using traditional and ancestral names. Learn those of others, and refer to each other by our Indigenous names. Encourage your family, friends, and colleagues to refer to you by your traditional name. These names carry your story, history, and connection to land/place. (Kaleimamoowahinekapu)
2. **Positionality** – Learn how to introduce yourself and describe your relationships to people, land, and the more-than-human in your language. This can include your genealogy, traditional lands, waterways, and mountains. ('O Candace Kaleimamoowahinekapu Galla ko'u inoa. He Hawai'i au. *My name is Candace Kaleimamoowahinekapu Galla. I am Hawai'i/I am Hawaiian.*)
3. **Salutations** – Say your greetings and farewells, and use appropriate responses to each other in the language. (Aloha nō! Pehea 'oe? *Hello! How are you?*)
4. **Social Interactions** – Begin in-person, face-to-face, and digitally-mediated interactions in the language. Answer and end the phone call in the language. Record your voicemail greeting/message in the language. Begin and end your emails, texts, or chats in the language.
5. **Interior Design** – Consider writing words, common phrases, and actions on labels to display them in places that are occupied or frequented (for example, the kitchen, living room, office). Replace the labels with new ones as these words and phrases become a part of your lexicon and your vocabulary expands.
6. **Language domains** – Identify specific spaces or rooms in your home, frequented spots, and/or commit specific times in which you will spend immersed in the language (for example, at breakfast, on car rides, on the way to/from school or work).
7. **Cultural practice(s)** – Identify your favorite cultural practice or activity and learn words, phrases, and cultural protocols that are relevant to it. Use this vocabulary while engaging in the practice, and use it with other cultural practitioners and knowledge holders.
8. **Task-based activities** – Choose a task (for example, cooking, weaving), and learn how to complete the task in the language. Make this a part of the weekly schedule so that it becomes a customary practice.
9. **Song and Dance** – Learn songs, chants, and prayers in the language. Learn what is being expressed by the lyrics and be taken on a journey to experience the story. Or better yet, write your own lyrics, sing them, dance, and create choreography to the lyrics – share them!
10. **Other companions** – Speak to your pet(s) in the language. Most don't speak back, so try your language on them. They are great listeners and don't judge you!

Each of these practices can be adapted to work with language learners and speakers at any level. Whether you live in community or now reside in the big city or a different country, find language resources (speakers, books, media) to help support your journey. Remember to recognize, celebrate, and reward your progress, small and big. Safeguard Indigenous languages, protect these language spaces, and encourage others to join your movement with the intention to build language capacity so that we use Indigenous languages in all domains.

And most importantly, have fun – enjoy learning and speaking your Indigenous language! Through our ancestral languages, we begin to understand our relationships with and responsibilities to each other and our natural environment, which will guide us to a healthy and sustainable future.

References

- Adams, D.W., 1995. *Education for Extinction: American Indians and the Boarding School Experience, 1875-1928*. University Press of Kansas, Lawrence.
- Battiste, M., 2013. *Decolonizing Education: Nourishing the Learning Spirit*. UBC Press.
- Begay, W.R., 2013. *Mobile Apps and Indigenous Language Learning: New Developments in the Field of Indigenous Language Revitalization*. MA thesis. University of Arizona, Tucson.
- Bernard, H.R., 1996. Language preservation and publishing. In: Hornberger, N.H. (Ed.), *Indigenous Literacies in the Americas: Language Planning from the Bottom up*. Mouton de Gruyter, Berlin, pp. 139–156.
- Cajete, G., 1999. *Igniting the Sparkle: An Indigenous Science Education Model*. Kivaki Press, Durango, CO.
- Cajete, G., 1994. *Look to the Mountain: An Ecology of Indigenous Education*. Kivaki Press, Durango, CO.
- Cassels, M., 2019. Indigenous languages in new media: opportunities and challenges for language revitalization. *Working Papers of the Linguistics Circle* 29 (1), 25–43.
- Commonwealth of Australia, 1997. *Bringing Them Home: National Inquiry into the Separation of Aboriginal and Torres Strait Islander Children From Their Families*. Retrieved from: https://humanrights.gov.au/sites/default/files/content/pdf/social_justice/bringing_them_home_report.pdf.
- Cru, J., 2018. Micro-level language planning and YouTube comments: destigmatising Indigenous languages through rap music. *Curr. Issues Lang. Plann.* 19 (4), 434–452.
- Cru, J., 2015. Language revitalisation from the ground up: promoting Yucatec Maya on Facebook. *J. Multiling. Multicult. Dev.* 36 (3), 284–296.
- Dawson, A.S., 2012. Histories and memories of the Indian boarding schools in Mexico, Canada, and the United States. *Latin American Perspectives* 39 (5), 80–99.
- Duarte, M.E., Belarde-Lewis, M., 2015. Imagining: creating spaces for Indigenous ontologies. *Cat. Classif. Q.* 53 (5–6), 677–702. <https://doi.org/10.1080/01639374.2015.1018396>.
- Dwyer, A., Brezinger, M., and Yamamoto, A. (n.d.). UNESCO Program: Safeguarding of Endangered Languages. Intangible Cultural Heritage Section. Retrieved from http://193.48.145.249/Colloques/3L_2008/3LCourseMaterial/YamamotoDwyer_AAA_UNESCO.pdf.
- Evans, M., 2004. Ethics, anonymity, and authorship in community centred research or anonymity and the island cache. *Pimatisiwin* 2 (1), 59–75. Spring.
- Galla, C.K., Kawai'ae'a, K., Nicholas, S.E., 2014. Carrying the torch forward: Indigenous academics building capacity through an international collaborative model. *Can. J. Native Educ.* 37 (1), 193–217.
- Galla, C.K., 2016. Indigenous language revitalization, promotion, and education: function of digital technology. *Comput. Assist. Lang. Learn.* 29 (7), 1137–1151.
- Galla, C.K., 2018. Digital realities of Indigenous language revitalization: a look at Hawaiian language technology in the modern world. *Lang. Lite.* 20 (3), 100–120.
- Gupta, M., Padel, F., 2020. The Travesties of India's Tribal Boarding Schools. *Sapiens*. Retrieved from: <https://www.sapiens.org/culture/kalinga-institute-of-social-sciences/>.
- Hampton, E., 1995. Towards a redefinition of Indian Education. In: Battiste, M., Barman, J. (Eds.), *First Nations Education in Canada: The Circle Unfolds*. UBC Press, Vancouver, BC, pp. 5–46.
- Hermes, M., Cash, P.C., Donaghy, K., Erb, J., Penfield, S., 2016. New domains for Indigenous language acquisition and use in the USA and Canada. In: Coronel-Molina, S.M., McCarty, T.L. (Eds.), *Indigenous Language Revitalization in the Americas*. Taylor and Francis, pp. 269–291.
- Hieber, D.W., 2015. Review of endangered languages and new technologies. *Lang. Doc. Conserv.* 9, 344–350.
- Hinton, L., Hale, K. (Eds.), 2001. *The Green Book of Language Revitalization in Practice*. Brill.
- Hobson, J.R. (Ed.), 2010. *Re-awakening Languages: Theory and Practice in the Revitalisation of Australia's Indigenous Languages*. Sydney University Press.
- Holmes, A., 2018. *Geographies of Home, Memory, and Heart: Mohawk Elder Praxis, Land, Language, and Knowledge Woven in Place*. PhD dissertation. University of Arizona, Tucson.
- Holmes, C., 2020. Baby sharks in the desert: Lotta' little lights everywhere. *Rural Soc.* 1–17.
- Jacobs, M.D., 2006. Indian boarding schools in comparative perspective: the removal of Indigenous children in the United States and Australia, 1880–1940. In: Trafzer, C.E., Keller, J.A., Sisquoc, L. (Eds.), *Boarding School Blues: Revisiting American Indian Educational Experiences*. University of Nebraska Press, Lawrence, pp. 202–231.
- Krauss, M., 1992. The world's languages are in crisis. *Language* 68 (1), 4–10.
- Kemper, K.R., 2015. Cultural hybridity, resilience and the communication of contemporary Cherokee culture through mobile technologies. In: Dyson, L.E., Grant, S., Hendriks, M. (Eds.), *Indigenous People and Mobile Technologies*. Routledge, pp. 239–252.
- Kimmerer, R., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions, Minneapolis, MN.
- Kipp, D., 2000. *Encouragement, Guidance, Insights, and Lessons Learned for Native Language Activists Developing Their Own Tribal Language Programs*. Piegan Institute, Browning, MT.
- Kipp, D., DesRosier, J., Miyashita, M., 2015. Memoir and insights of Darrell R. Kipp. In: Reyhner, J., Martin, J., Lockard, L., Gilbert, W.S. (Eds.), *Honoring Our Elders: Culturally Appropriate Approaches for Teaching Indigenous Students*. Northern Arizona University, pp. 1–14.
- Legislative Reference Bureau, 2022. *The Constitution of the State of Hawai'i - Article X (Education)*.
- Lillehaugen, B.D., 2019. Tweeting in Zapotec: social media as a tool for language activists. In: Gomez Menjivar, J.C., Chacon, G.E. (Eds.), *Indigenous Interfaces: Spaces, Technology, and Social Networks in Mexico and Central America*. University of Arizona Press, pp. 201–226.
- Littlebear, R.E., 1996. A model for promoting Native American language preservation and teaching. In: Cantoni, G. (Ed.), *Stabilizing Indigenous Languages*. Northern Arizona University, pp. 234–239.
- Lomawaima, T., McCarty, T.L., 2006. *To Remain an Indian: Lessons in Democracy from a Century of Native American Education*. Teachers College Press, New York, NY.
- MacEntee, K., Burkholder, C., Schwab-Cartas, J., 2019. Cellphlms in public scholarship. In: Leavy, P. (Ed.), *The Oxford Handbook of Methods for Public Scholarship*. Oxford University Press, pp. 1–26.
- McDonald, J.A., 2003. *People of the Robin: The Tsimshian of Kitsumkalum: A Resource Book for the Kitsumkalum Education Committee and the Coast Mountain School District 82 (Terrace)*. CCI Press, Edmonton, AB.
- McIvor, O., 2009. Strategies for Indigenous language revitalization and maintenance. In: *Encyclopedia of Language and Literacy Development*. Canadian Language and Literacy Research Network, London, ON, pp. 1–12.
- McLeod, N., 2007. *Cree Narrative Memory: From Treaties to Contemporary Times*. Purich, Saskatoon, SK.
- Mohanna, M., Capus, L., 2012. Modeling of mobile learning. In: Sanchez, I.A., Isaias, P. (Eds.), *Proceedings—International Association for Development of the Information Society (IADIS) International Conference Mobile Learning*. IADIS Press, pp. 315–319.
- National Indian Brotherhood, 1972. *Indian Control of Indian Education*. National Indian Brotherhood. Policy paper.
- Nicholas, S., 2014. "Being" Hopi by "living" Hopi, redefining and reasserting cultural and linguistic identity: emergent Hopi youth ideologies. In: Wyman, L., McCarty, T., Nicholas, S. (Eds.), *Indigenous Youth and Multilingualism: Language Identity, Ideology, and Practice in Dynamic Cultural Worlds*. Routledge, New York, NY, pp. 70–89.
- Oppenheer, M., 2011. Using ICTs for Indigenous Cultural Preservation: Challenges and Strategies. The Ethnos Project. Retrieved from: <http://www.ethnosproject.org/using-icts-for-indigenous-cultural-preservation/>.
- Perley, B.C., 2012. Zombie linguistics: experts, endangered languages and the curse of undead voices. *J. Soc. Anthropol. Comp. Sociol.* 22 (2), 133–149.
- Rice, D., 2019. History of the Kanien'kéha Language in Kahnawà:ke. Retrieved from: <http://www.kahnawake.com/org/docs/2020-06-10-HistoryOfLanguage.pdf>.
- Ruiz, R., 1984. Orientations in language planning. *NABE: Journal of the National Association for Bilingual Education* 8 (2), 15–34.
- Schwab-Cartas, J., 2016. Living our language: Zapotec elders and youth fostering intergenerational dialogue through cellphone videos. In: MacEntee, K., Schwab-Cartas, J., Burkholder, C. (Eds.), *What's a Cellphilm? Integrating Mobile Phone Technology into Participatory Arts Based Research and Activism*. Sense Publishers, Rotterdam, NL, pp. 51–65.
- Schwab-Cartas, J., 2018. Keeping up with the sun: revitalizing Isthmus Zapotec and ancestral practices through cellphlms. *Can. Mod. Lang. Rev.* 74 (3), 363–387.

- Shariff, S., 2015. *Sexting and Cyberbullying*. Cambridge University Press.
- Skutnabb-Kangas, T., Dunbar, R., 2010. Indigenous children's education as linguistic genocide and a crime against humanity? A Global View. *Galdu Cala - Journal of Indigenous Peoples Rights* 1.
- Skutnabb-Kangas, T., Phillipson, R., 2010. The global politics of language: Markets, maintenance, marginalization or murder? In: Coupland, N. (Ed.), *Blackwell Handbooks in Linguistics: The Handbook of Language and Globalization*. Wiley-Blackwell, Sussex, pp. 77–100.
- Smith, A., 2009. Indigenous Peoples and Boarding Schools: A Comparative Study. Prepared for the Secretariat of the United Nations Permanent Forum on Indigenous Issues. Eighth Session. United Nations Economic and Social Council, New York, NY. https://www.un.org/esa/socdev/unpfi/documents/E_C_19_2009_crp1.pdf.
- Smith, L.T., 1999. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed, London.
- Smith, G.H., 2012. Interview: kaupapa Māori: the dangers of domestication. *N. Z. J. Educ. Stud.* 47 (2), 10–20.
- Spack, R., 2002. *America's Second Tongue: American Indian Education and the Ownership of English, 1860–1900*. University of Nebraska Press, Lincoln, NE.
- Stald, G., 2008. Mobile identity: youth, identity, and mobile communication media. In: Buckingham, D. (Ed.), *Youth, Identity, and Digital Media*. MIT Press, Cambridge, MA, pp. 143–164.
- Truth and Reconciliation Commission of Canada, 2015. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. Retrieved from: https://ehprnh2mwo3.exactdn.com/wp-content/uploads/2021/01/Executive_Summary_English_Web.pdf.
- United Nations, 2018. 2019 – International Year of Indigenous Languages. Retrieved from: <https://en.iyl2019.org>.
- UNPFII Factsheet. (n.d.) Who are Indigenous Peoples? Retrieved from https://www.un.org/esa/socdev/unpfi/documents/5session_factsheet1.pdf
- Vizenor, G.R., 1999. *Manifest Manners: Narratives on Postindian Survivance*. University of Nebraska Press, Lincoln, NE.
- Walker, R., 2016. Reclaiming Māori education. In: Hutchings, J., Lee-Morgan, J. (Eds.), *Decolonisation in Aotearoa: Education, Research and Practice*. NZCER Press, pp. 19–38.
- Wolfe, P., 2016. *Traces of History: Elementary Structures of Race*. Verso, London.

Relevant Websites

- Aha Pūnana Leo. <https://www.ahapunaleo.org>.
- Commonwealth of Australia: Bringing Them Home: National Inquiry to the Separation of Aboriginal and Torres Strait Islander Children From Their Families. https://humanrights.gov.au/sites/default/files/content/pdf/social_justice/bringing_them_home_report.pdf.
- Cultural Survival. <https://www.culturalsurvival.org>
- First Peoples' Cultural Council. <https://fpcc.ca>
- Hale Kuamo'o. Nā Puke Wehewehe 'Ōlelo Hawai'i (Hawaiian Dictionaries). <https://wehewehe.org/>.
- Indigenous children's education as linguistic genocide and a crime against humanity? A global view. *Journal of Indigenous Peoples Rights*. <https://www.afn.ca/uploads/files/education2/indigenousschoolseducation.pdf>.
- Legislative Reference Bureau. The Constitution of the State of Hawai'i. <https://lrb.hawaii.gov/constitution#articlex>.
- National Indigenous Languages Institute of Mexico. <https://www.gob.mx/inali>
- National Centre for Truth and Reconciliation. Residential School Timeline. <https://nctr.ca/exhibits/residential-school-timeline/>.
- New Zealand Council for Educational Research. Reclaiming Māori Education. <https://www.nzcer.org.nz/system/files/Chapter%201%20Ranginui%20Walker.pdf>.
- Northern Arizona University. Memoir and Insights of Darrell R. Kipp. <https://www2.nau.edu/~jar/HOE/HOE1.pdf>.
- NPR. American Indian Boarding Schools Haunt Many. <https://www.npr.org/templates/story/story.php?storyId=16516865>.
- Mokuola Honua. "Officially" What? The Legal Rights and Implications of 'Ōlelo Hawai'i. <https://www.mokuolahonua.com/resources/language-policy-and-advocacy/2019/2/18/officially-what-the-legal-rights-and-implications-of-lelo-hawaii-yk9gg>.
- Parliament of Canada. Indigenous Languages Act. <https://www.parl.ca/DocumentViewer/en/42-1/bill/C-91/first-reading>.
- Polynesian Voyage Society. Honoring Mohawk Language Leaders. <http://www.hokulea.com/hokulea-update-date/>.
- Polynesian Voyage Society: Mālama Hōnua Worldwide Voyage. <http://www.hokulea.com/worldwide-voyage/>.
- Te Kōhanga Reo. <https://www.kohanga.ac.nz/en/>.
- The National Native American Boarding School Healing Coalition. US Indian Boarding School History. <https://boardingschoolhealing.org/education/us-indian-boarding-school-history/>.
- The Piegan Institute. <https://www.pieganinstitute.org/our-history>.
- TreeTV/N2K. Need to Know: We are the Haudenosaunee. <https://www.youtube.com/watch?v=2DofTnRhm5o>.
- UNESCO. International Decade of Indigenous Languages. <https://en.unesco.org/idi2022-2032>.
- UNESCO. International Year of Indigenous Languages. <https://en.iyl2019.org>.
- United Nations. Declaration on the Rights of Indigenous Peoples. <https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html>.
- United Nations. Permanent Forum on Indigenous Issues. <https://www.un.org/development/desa/indigenouspeoples/unpfi-sessions-2.html>.
- UNPFII.(n.d.) Who are Indigenous Peoples? https://www.un.org/esa/socdev/unpfi/documents/5session_factsheet1.pdf.
- UNPFII. Indigenous Peoples and Boarding Schools: A Comparative Study. https://www.un.org/esa/socdev/unpfi/documents/IPS_Boarding_Schools.pdf.
- US Congress. Esther Martinez Native American Languages Preservation Act of 2006. <https://www.congress.gov/bills/109th-congress/house-bill/4766>.
- US Congress. Native American Languages Act. <https://www.congress.gov/bills/101st-congress/senate-bill/2167>.

(Just) languaging with *duende*: (re)inventing and innovating communicative practices as resistance, healing, and thriving in Indigenous contexts

Joaquin Muñoz^a, Candace Kaleimamoowahinekapu Galla^b, Leisy Wyman^c, and Perry Gilmore^c, ^a Department of Curriculum & Pedagogy, Faculty of Education, The University of British Columbia, Vancouver, BC, Canada; ^b Department of Language & Literacy Education, Faculty of Education, and Institute for Critical Indigenous Studies, Faculty of Arts, The University of British Columbia, Vancouver, BC, Canada; and ^c Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Kisizi: young children invent a Swahili pidgin in rural Kenya	799
The language	799
Symbolic performances and semiotic resistance	800
Fueling innovation: bringing forth embodied language practices	801
(Re)activating and (re)centering language in/to cultural practices	801
Mele and hula as embodied calls to action and solidarity	802
Inventive anti-colonial Indigenous social media	803
Something Else	803
Playing the way he felt it: Doggface208 “vibes indigenously”	804
Conclusion	806
References	806
Relevant websites	807

With idea, sound, gesture, the duende delights in struggling freely with the creator on the edge of the pit ... the duende wounds, and in trying to heal that wound that never heals, lies the strangeness, the inventiveness of a man's work.

... We have said that the duende loves the edge, the wound, and draws close to places where forms fuse in a yearning beyond visible expression.
Lorca (1933).

Scientists, linguists, philosophers and poets alike stand in awe and wonder at human language and the capacity for creativity and inventiveness in all its expressive communicative forms. From the vastly studied origins of human language, to the unending genesis of new languages (Riley, 2015), new language varieties (Mufwene, 2008), new styles (Eckert, 2012), new repertoires (Blommaert, 2010; Nassenstein and Hollington, 2016), and new modes of multimodal communication (Hollington and Nassenstein, 2018; Riley, 2015), this artful generativity both marks and is marked by emerging speech communities continuously creating and recreating individual and group social political identities. This deep, innovative, creative ability endlessly generates new ways to symbolically communicate our humanity through a vast range of fluid expressive forms, linguistic and communicative resources and artful performative practices. From individual's daily novel utterances, to newly emerging lexical and syntactic forms, language births, and original innovative multimodal communication systems, language never sits still. Even a child's language development and acquisition has been more aptly described as a process of continual language invention (Gleitman and Newport, 1995). Language is always changing; always dynamic; always newly performed, played with, invented, created, manipulated, (re)shaped, refined and reconstructed. It is always “language-in-motion” (Blommaert, 2010, p. 5). This fluidity describes “the shifting and ever-changing linguistic material, the ways this material is used by speakers, the speakers themselves, their social and cultural environments, people's identity practices, etc.” (Nassenstein and Hollington, 2016, p. 183).

Understanding its everchanging nature, Garcia and Wei (2014) prefer to use the verb “languaging,” introduced by Becker (1995) in his work with translation, rather than the noun “language.” They write,

... languaging is a better term to capture an ongoing process that is always being created as we interact with the world lingually. To learn a new way of languaging is not just to learn a new code, [...] it is to enter another history of interactions and cultural practices and to “learn a new way of being in the world” (p. 8).

They see language as “a set of resources, language ideologies and social constructions, the multimodality of communication among others” (p. 8).

In this essay, we favor the dynamic, complex and socially embedded view of language as always being invented, always in process, always meaning making, and always in context. Further, we recognize the esthetic, performative, emotive, relational, playful, multimodal and artful aspects of languaging too often ignored by linguistic studies that focus purely on language form and

structure. Influenced by and extending the work of Gilmore (2016, 2017), we view human language creativity in all its multimodal dynamic semiotic forms and communicative performances, as embodying what the poet Lorca (1933) describes as the miracle of the *duende* which he sees as the source of all “newly created things.” Gilmore (2016) writes,

Federico García Lorca ... has captured the mystery of the phenomenon of the creative process most profoundly. In his classic essay, *Theory and Play of the Duende*, based on a lecture he delivered in Buenos Aires in 1933, Lorca described the *duende* as the power and struggle of spontaneous creation, as a “burst of inspiration, the blush of all that is truly alive, all that the performer is creating at a certain moment.” Lorca writes, “The *duende*’s arrival always means a radical change in forms. It brings to old planes unknown feelings of freshness, with the quality of something newly created, like a *miracle* (emphasis mine).” ... In Lorca’s original lecture, he referred to the thousand boring lectures he attended in his studies in Madrid that made him “hungry for air and sunlight” and “covered by a light film of ash.” Instead of the intellectual rationality of scholarship, he turned to the *duende* to explain the unexplainable about the art of creation. He describes the *duende* as “a mysterious power which everyone senses but no philosopher explains” ... It is a heightened state of human reality with all its joys and sorrows, of spine-chilling authenticity The *duende*, especially associated with the evocative aesthetics of flamenco and the bullfight, has also been invoked to both inspire and explain a range of modern art forms in music, painting and poetry (pp. 55–56).

Lorca’s (1933/1998) notions of the *duende*, unlike the romantic inspiration of the angel or muse, depicts the dark ongoing struggle that both hurts and attempts to heal. Lorca writes, “With idea, sound, gesture, the *duende* delights in struggling freely with the creator on the edge of the pit. Angel and Muse flee, with violin and compasses, and the *duende* wounds, and in trying to heal that wound that never heals, lies the strangeness, the inventiveness of a man’s work” (p. 67).

In this essay provoked by the concept of the *wound that never heals*, we draw on Lorca’s notions of the miracle of *duende* to explore the creative processes in each of the illustrative communicative examples we present. We have selected three disparate examples of this dynamic communicative creativity and semiotic language innovation. In each of our examples, we highlight and extend Lorca’s conception of the *duende* by exploring themes of struggle with colonial oppression and a range of attempts to heal its injustices through newly creative human expressive forms. While Lorca focused on the inventiveness of the individual creator—the artist, the poet, the musician, the dancer—this discussion will extend the reach of the *duende* to examine collaborative, collective, and community communicative creativity and innovation. These social interactive creative processes demonstrate a significant layer of human inventiveness where the group or community—rather than an individual—comes together to struggle with the *duende* in order to generate a dynamic innovative capacity in the face of oppression and injustice. Further, we extend Lorca’s discussion of the *duende* by exploring the significant role play, humor, satire and irony can serve in creative resistance to oppression (Vizenor, 2008; Deloria, 1969; Bresnick, 2019; Gilmore, 2016; Wyman, 2012). Often, the “fueling of innovation” (Kahaunalele, 2020) can find its source in the struggle to *heal that wound that never heals*. The three stories that follow demonstrate *just* ways of creative *languaging* that resist and critique oppression, renew agentive identities, and establish community solidarity in the face of settler colonial injustice in Indigenous contexts. While an abundant range of artful creative expression is inspired by existential reflections about joy, love, birth, despair, grief and a myriad of other profound human life experiences, we focus exclusively on the inventiveness evoked by the wounds of, and resistance to, colonial domination, oppression, and injustices. Through pain, joy, humor, play, verbal art, dance, song and digital expression, each case dramatically calls out for *just languaging* with evocative spirit and social *duende*.

The first narrative describes the case of two 5-year-old boys, one white American and one Indigenous Samburu, who met and became inseparable friends on an isolated hillside in rural Kenya in 1975. Their controversial postcolonial relationship was created through and marked by their invention of *Kisisi*, an original Swahili pidgin that was understood and shared by just the two of them. Their irrepressible linguistic playful creativity creates a space for their friendship that transcends the existing language ideologies, marked colonial borders, and harsh inequities of economics, race, and settler colonial culture that engulfed all aspects of their daily lives (Gilmore, 2009, 2011, 2016).

In the second example, we examine the collective languaging and expressive practices of an Indigenous movement space (Eagle Shield et al., 2020; Paris, 2019), where we witness the creative mixture of enduring languaging in hula and mele that expresses collective life force, *duende*, and power in a struggle to protect sacred land during the current Hawaiian Renaissance. Traditional and contemporary practices of hula and mele, their intersections and cross-currents, meanings and poetics are explored through its innovation and inventiveness in response to controversial contemporary injustices despite pressures within and outside the community to conform to Western ideologies (Galla et al., 2015).

Drawing from anticolonial work with Indigenous literacies and creative writing, and holistic pedagogies (Horner et al., 2021; Muñoz, 2020), the third example explores Indigenous language practices and social media, largely influenced by youth, as an artful driving force in communicative practices that create digital spaces as sites where “new creative practices arise, are negotiated and serve as identity markers across widespread digital communities ...” (Hollington and Nassenstein, 2018). These transidiomatic practices (Jacquemet, 2005) are new forms of communication and new mediatized spaces (e.g., Facebook, Youtube, TikTok, Twitter, WhatsApp), with mobile people and mobile texts, that allow new syncretic ways to foster relationships and stay bonded across previously vast translocal life distances. These creative new communicative repertoires and codes include, and are a part of broader stylistic, esthetic, poetic, emotive, playful, humorous, satirical, and agentive choices that are strategically navigated in order to perform changing identities, assert sovereignty, affirm community, and enact collective resistance in the face of colonization and oppression (Gunn Allen, 1986; Berglund and Carlson, 2020; Gilmore and Wyman, 2013; Riley, 2015; Wyman et al., 2013, 2016).

Kisisi: young children invent a Swahili pidgin in rural Kenya

In this section, we highlight the unique processual documentation of the birth of a private language invented by two 5-year-old boys, Colin and Sadiki, who in 1975 by chance became neighbors and inseparable friends for 15 months on a remote and isolated multilingual hillside in postcolonial Kenya. Perry Gilmore, one of the authors of this essay and Colin's mother, describes these events in detail in her book, *Kisisi (Our Language): The Story of Colin and Sadiki* (2016). In her ethnography, she explores (1) the awe and mystery of the children's new language creation; (2) their agentive socio-political linguistic resistance to oppressive settler colonial language regimes, ideologies and policies; and (3) the evocative *duende* embodied in the verbal art, play, and performative transcendent healing power of their language use in this overtly harsh and racist context.

Colin, an English speaker, was the son of American researchers studying a troop of wild baboons. They lived in a research residence located on a 48,000-acre cattle ranch owned by a prominent settler colonial family. Sadiki, a Samburu speaker, was the son of Indigenous Samburu ranch workers who herded cattle on foot for meager wages and lived with their five children in small cramped laborer's quarters just yards away from Colin's family. Additional ranch workers from different tribes also lived on the hillside in workers' quarters. All social interactions between the residents in this multilingual contact zone were conducted in Swahili, the local lingua franca that held a dramatic history as a tool of colonial domination and power (Fabian, 1986; Gilmore, 2016; Myers-Scotton, 1978; Ndege, 2009; Skutnabb-Kangas, 2011). Swahili was the language of the dominant colonial enterprise. It was the language the colonials used to give orders to workers. It was not the language of personal relationships. Yet the boys were expected to use Swahili in all of their interactions. Instead, as the boys played together, they created a shared language that sounded Swahili yet was unintelligible to Swahili speakers. Their language, Kisisi, displayed characteristic features of pidgin, creole and emerging languages all over the world. These translanguaging practices (e.g., see García and Wei, 2014; Nassenstein and Hollington, 2016) are common in multilingual contact zones and have recently received much attention in discussions of superdiversity (Blommaert, 2010). However, the degree to which they expanded and elaborated their language was unique, especially for such young children. Their language invention demonstrates striking linguistic and sociolinguistic competencies of very young children who are too often highly underestimated in this regard. As these young boys playfully negotiated diverse linguistic ecologies and cultural spaces, they displayed abilities as highly effective language innovators creating a unique language with sophisticated original phonological, lexical and syntactic features (see Gilmore, 2011, 2016 for further discussion of these features).

The language

Kisisi shared many characteristics of pidgin languages all over the world. Pidgins arise in extended contact situations where individuals have no common language and develop a means of restricted communication that serve their limited needs. Extended language contact typically arises in a range of circumstances including "discovery," trade, conquest, slavery, migration, colonialism and nationalism, to name a few. Many of these contact conditions have historically marked unjust, often cruel and unequal power relationships. No one's first language, pidgins are characterized as simplified contact languages having pared down many of the inessential features of the different input languages. This simplification process is similar to the linguistic accommodation that occurs in a number of other linguistic circumstances where individuals do not share the same language competencies.

Gilmore (2016) found that the language shared the classic pidgin features including—a restricted lexicon, absence of articles, markers, inflection and copula (i.e., the verb to be), preverbal negation, polysemy (i.e., many meanings for one word), and circumlocution (i.e., figures of speech), to name a few. This finding was striking because it countered the pidgin/creole literature that claimed children learn pidgins from adults. Kisisi is a vivid example of children creating a pidgin rather than being taught by, or learning a pidgin from, adult speakers.

As the months passed, the boys' language began to exhibit characteristic features of more lexically and syntactically complex creole languages. Where a pidgin functions to serve limited language contact needs, a creole serves as a primary language in a speech community. As the creole process progresses over time, the original pidgin will undergo change in response to the growing needs of its speakers to communicate about a full range of human experiences with a richer lexicon and more grammatical complexity.

In Sadiki and Colin's language, there was a constantly expanding lexicon incorporating many new loan words, metaphoric compounds, and numerous neologisms (i.e., new words). Original articles, pronouns, and novel tense and aspect markers were introduced. All of these features are typical of elaborated creole languages rather than simplified pidgins. Creoles are used by large populations across multiple generations before accomplishing such language innovations. It is remarkable to discover two 5-year-old boys effectively creating and creolizing a pidgin in less than 15 months. An additional assumption in the creole literature is that adults are the innovators and children are characterized as the regulators. In this case, it is evident that the children were the innovators!

An illustration of their metaphoric creativity can be seen in the modification of several Swahili loanwords the boys combined to create new meanings. This process of "calquing" or compounding generated poetic imagery and lexical invention. For example, the Swahili word for sun is *jua*. Compounding two different Swahili words, the boys instead called the sun *kubwa moto* or *big fire*, displaying their rich metaphoric inventiveness. The word *choo* in Swahili has multiple meanings. It can refer to feces, urine, or the lavatory itself. The boys especially delighted in elaborating, over expanding and playing with their lexicon for toilet activities—often accompanied by farting noises and hilarious laughter. By compounding the Swahili word for water (*maji*) with *choo* (feces), they distinguished urinating (*maji choo*) from defecating (*choo*)—a significant and highly entertaining scatological distinction for 5-year-olds.

They also generated novel syntactic structures to express tense and aspect. For example, they created an original future tense marker by positioning the Swahili verb *kuja* (to come) as an auxiliary before the main verb. For example, “Mi *kuja* run” means “I will run.” The verb *kuja*, used as a newly invented future marker, was pragmatically helpful, enabling the boys to distinguish between play actions happening in the present, and imaginary future actions they were planning in their pretend play. They also used the Swahili verb *enda* (to go) as a preverbal auxiliary to mark on-going durative and iterative action as distinct from punctual action—allowing them to distinguish between someone shot dead on the spot, from someone dramatically dying slowly over time. The boys also introduced novel forms of articles, pronouns and markers. One fascinating example was their use of the Swahili word, *moja* (one), to function as a pronoun and an article. For example, they would say “*moja* road” for “the road.” Or they would refer to an individual thing or person as “*moja*” saying “*tagid moja*” or “take one.”

These examples of invented linguistic forms and structures of Kisisi demonstrate that 5-year-old children can and do understand and generate sophisticated grammatical functions and syntactic complexity in ways that our standard kindergarten curricula clearly do not recognize, and drastically underestimate.

Symbolic performances and semiotic resistance

The symbolic sociopolitical metacommunicative aspects of the performances of their language were also profound. The children’s vast differences in culture, race, and economics made their friendship highly controversial given the settler colonial socio-political ideologies and power hierarchies of the period. Yet despite these circumstances, Colin and Sadiki’s transgressive border crossing friendship grew into an intimate bond created through and marked by the invention of Kisisi. Their language helped them construct new identities, resist, transgress, transform and heal the marked colonial borders and harsh inequities that engulfed them and marginalized and excluded Indigenous Africans.

Colin and Sadiki’s isolated and remote rural situation contributed to their language innovation, collaborative language choices and social practices. By choosing to sustain, expand and develop their own private language, they resisted the dominant language ideologies that represented the hegemonic, racialized, postcolonial order of the newly independent Kenya. Their public use of their private language was their cultural critique of an oppressive regime in which their own cross-racial friendship was considered by many as a violation of social norms. The boys refused to docilely participate in the existing colonial order and rejected the Up-Country Swahili language that was designed and designated to keep that order, and the Indigenous population, in place. They resisted being socialized into a language ideology they rejected, and instead created a new language ideology that allowed a safe and celebrated space for their friendship. Their effective, border crossing agency was accomplished through loving verbal art and euphoric play. Their joy filled language practices challenged the oppressive colonial culture that surrounded them, identifying them as a distinct and separate speech community that valorized its own social justice values and allowed a space for their friendship.

Sadiki and Colin’s language began in response to a pragmatic need to understand each other in order to be playmates and friends. Their early genesis of an original, simplified Swahili pidgin served that immediate function, facilitating their play and budding friendship. As time went by and their close bonds deepened—even as Sadiki learned enough English and Colin enough Swahili to switch to using either of those languages—they continued to use, play with, expand and creolize Kisisi. Both English and Swahili signaled colonial power hierarchies and racial stratification that were inappropriate for two close and equal friends. Instead, Kisisi created a separate and just universe that allowed them to escape the unjust colonial order. Kisisi was fun. It was secret. It was theirs. It was an artistic verbal spectacle that surprised and captivated unsuspecting audiences. Their new language bonded them as much as it reflected their bonds. Through their language use, they carved a new exclusive and symbolically resonant healing space, a separate and safe universe for their friendship. Sadiki’s family saw the language as a gift from God, *Mungu*, because, as Sadiki’s mother said, “He saw how much they loved each other.” Seen as a sacred blessing, Kisisi carried a special spiritual evocative semiotic message, touching witnesses in a deeply emotive healing way. To mark this special spiritual bond, Sadiki’s grandfather traveled for miles from the Samburu Reserve to conduct a special ceremony making Colin Sadiki’s brother and one of their family.

Their wild-tongued (Anzaldúa, 1987) playful yet hegemonic defiance and social transgressions disturbed and disrupted racialized and strictly segregated postcolonial spaces. Where these ruptures occurred, the children opened spaces not only for themselves but for others whom they touched and changed by their friendship. Their newly created, ludic, and passionate communicative practices became a form of healing discursive “parrhesia,” a means of speaking truth to power, that confronted, transcended, and in some small ways transformed the harsh and brutal colonial world in which they lived. Their invention of Kisisi enabled them to question and challenge the hegemonic settler colonial ideologies and racist identities that attempted to define them, separate them and squelch their precious relationship—their *wound that never heals*. Their hybrid anti-language reflects their intentional revoicing of cultural elements and describes a process of ludic creativity that dialogically manipulates language, culture and identity for them and for others who witness and interact with them in their creative linguistically carved liminal, in-between, and subversive space. We see both Sadiki and Colin together, struggling with the *duende* on the edge of the pit, as they try to “heal the wound that never heals” with their risk-taking border crossing language practices—their creative languaging for moral justice against oppressive and persistent colonial powers.

Fueling innovation: bringing forth embodied language practices

From the onset of European colonization, raciolinguistic ideologies played an integral role positioning Indigenous Peoples in Africa, the Americas, and beyond as less than-human to justify projects of slavery and annihilation. Seen as the “Indian Problem,” Indigenous Peoples continue to be targeted in an effort to be extinguished and “absorbed into the body politic.” (Scott, 1920, p. 55 (L-3) & 63 (N-3)). As evidenced by the very existence of Indigenous Peoples today, the genocidal practice upon Indigenous bodies, hearts, spirits, and minds were not successful. Despite settler colonialism’s desire for assimilation, the prohibition of linguistic and cultural practices, the lived experiences and current realities as a result of forced removal and relocation, and other related oppressive systems, Indigenous Peoples have not only demonstrated resilience and survivance (Vizenor, 2008), but “thrivance.” Enacting thrivance calls for a resistance, perseverance, and endurance that looks toward an imagined, flourishing Indigenous future and healthy community. Through the reassertion of sovereignty and self-determination, Indigenous Peoples, communities, and nations have championed community-led and -centered initiatives, educational programing, and capacity building to effectively “break the chains that bind us to our colonial existences” (Alfred and Corntassel, 2005, p. 613). While Indigenous language transmission has been disrupted for generations, Indigenous Peoples throughout the world, are reaffirming their inherent human rights to protect their knowledges and languages from further desecration, exploitation, and commodification. The activation of “being Indigenous means thinking, speaking, and acting with the conscious intent of regenerating one’s indigeneity” (p. 614)—which includes the “maintaining of the tradition and the fueling of innovation” (Kahunaale, 2020, p. 541) through linguistic and cultural practices. One such exemplar that demonstrates innovation and inventiveness includes the recent creative contributions made by Kānaka Hawai’i specifically (and other cultural practitioners generally) who embody mele (song) and/or hula (dance) through ‘Ōlelo Hawai’i (Hawaiian language). While this section speaks of Hawai’i and Hawaiian language, the examples discussed may have applicability and similar pathways that parallel the experiences of other Indigenous nations, communities, and peoples. The struggles, challenges, opportunities, and successes are what binds us together. The illustrations described acknowledge, demonstrate and document language inventiveness that has been inspired by activism, sacrifice, and protection of Maunakea—also known as Mauna a Wākea—a sacred place/being that is significant to Hawaiian identity, well-being and security (Kānaka’ole-Kānahele as cited in Kumu Pono Associates, 1999).

The Thirty Meter Telescope (TMT)—a \$1.4 billion project—has garnered local, national and international attention as the largest telescope in the world to be built, in a location on Hawai’i Island that would further destruct and desecrate the living entity and deeply rooted spiritual connections to Hawaiian knowledge systems. The connection that Kānaka Hawai’i have to Maunakea can be appreciated and understood similarly by the relationships that peoples and communities have to their sacred lands—to name a few—Bears Ears (United States), Kaya Forests (Kenya), Uluru (Australia), Machu Picchu (Peru), Mount Fuji (Japan), and Mount Sinai (Egypt). More specifically, Whanganui River (New Zealand), Muteshekau Shipu (Canada), and Klamath River (United States) are natural landscapes and resources that have been granted the same legal rights as a person, thereby recognizing that the guardians of these sacred spaces will maintain their health and well-being according to specific local traditions and customs.

Though Maunakea has not yet been granted the same legal rights as a person, the Protect Maunakea movement has galvanized a present-day rebirth, restoration, renewal, resurgence, reclamation, and recommitment to Hawaiian ways of knowing, being, and doing—nearly 50 years following the Hawaiian Renaissance that propelled Hawaiian language medium education. This education in Indigenous movement spaces (Eagle Shield et al., 2020; Paris, 2019) has (re)activated the desire to (re)learn traditional practices grounded in Hawaiian epistemology that (re)ignites the flame within youth, adults, Elders; individuals, families, communities, diaspora, and beyond. Supporters near and far demonstrated support, aloha ‘āina and kapu aloha to Maunakea in various ways that include: witnessing, learning and participating in cultural protocols, mele and hula—through in-person visits and/or virtual engagement, attending community-led classes and university-credited courses on the mauna, sharing news and updates via social media using relevant hashtags, writing advocacy letters urging for the divestment from the TMT project, (re)familiarizing old and introducing new mele honoring Maunakea, and a virtual worldwide jam.

Some of these embodied and performative expressions have been passed down through generations retaining the same cadence, rhythm, style, sentiment, and/or gestural and foot movements, while others may have changed ever so slightly, or renewed, (re)created, and (re)imagined to depict the current contexts and realities. What remains consistent is the language—the lyrics of the mele which reveal just one layer of the Indigenous storywork. These documented and lived practices—preserved instances of oral and embodied practices—are demonstrations of genealogy(ies) of language and cultural practitioners who have resisted oppression and made concerted efforts to carry forward Hawaiian traditions and customs.

(Re)activating and (re)centering language in/to cultural practices

Hula is a traditional dance—an evolving tradition (Silva, 2014) of Hawai’i that requires a mele (song)—that is now practiced, learned, and taught to students of all ages and cultural backgrounds, not only in Hawai’i, but in many parts of the world (e.g., USA, Canada, Japan, Mexico, Australia, France, Germany, Nigeria, Tahiti). Unknown to many, Hawaiian language is the essence of hula that presents a window into the culture. “Hula is first a language that has a choreographical piece to it” (Taupōuri Tangaro as cited in Smithsonian, 2013, 0:39) that requires the living flesh to interpret it with truth (Lorca, 1933). However, due to the illegal overthrow in 1893 by the United States, along with a language bill HRS 298-2 in 1896 that prohibited Hawaiian language medium

instruction, many Kānaka Hawai'i did not grow up learning the language. The disruption of intergenerational knowledge and language transmission also applied to cultural practices like hula.

Fortunately, the Hawaiian language is being actively revitalized through community initiatives, school programing, and family and individual commitment, and is making an incremental return. Hula and mele are just some of the ways that Hawaiian language is being amplified. A concern, however, of this now globalized art is that the expansion of hula beyond Hawai'i and the plethora of digital technologies has resulted in learned choreographies that are mechanical in form, rather than a deep, heart learning and embodied understanding of the story that is based in language.

The Hawaiian language composed poetically and presented melodically in mele, carries the meaning in the art. Hula without the conventional use of its characteristic Indigenous language would not exist as a specific art form. The language encodes the values and knowledge of the Hawaiian people in a unique cultural way of being. As hula is learned and performed in many places outside of Hawai'i, it reinforces the importance of Hawai'i remaining true to the substance of hula's personification through its own Indigenous language.

Galla et al. (2015, 130–132).

It is even more critical now to (re)activate and (re)center Hawaiian language into hula to understand that hula is much more than entertainment and is not a stand-alone performative art that is void of language and culture. Not only is hula a site of linguistic memory, but “is inherently a site of cultural memory, not only in the act of performance, but in the entirety of its practices of archiving knowledge” (Stillman, 2001 p. 188) that endures generations—past, current, and future.

Along with this renewed energy to recover and reclaim mele and hula, cultural practitioners are creating new compositions and choreographies to employ activism and to document important events and lived experiences (Kahaunaele, 2020). This has resulted, for example, in mele from the Kūha'o Maunakea album that document and depict the reciprocal relationship that Kānaka Hawai'i have to Maunakea and vice versa, and hula that is choreographed to chronicle, personify, and illustrate a specific event, experience, emotion, and/or spiritual connection. These linguistic, literacy, and literary practices are forms of education that have adapted to the changing contexts of Kānaka Hawai'i who are creating new opportunities for language use for learners/speakers from newborns to elders. These multimodal opportunities ensure that there is thriving, cultural continuity and intergenerational knowledge transmission that responds to the needs of the diverse community that extends from Hawai'i born Kānaka Hawai'i to diaspora, and beyond.

Mele and hula as embodied calls to action and solidarity

Candace Galla, one of the authors of this essay recalls the summer of 2019, when there was a heightened awareness and tension regarding the TMT. Some of the poignant memories include livestreams of the arrests of kūpuna (Elders) who blocked the road to the summit of Maunakea in protest of TMT construction, and in protection of Maunakea, and the cutting of the hae Hawai'i (Hawaiian flag) by a law enforcement officer. Though she would bear witness through a laptop screen of her birthlands, she would engage in the mele and hula that kia'i (protectors) at Maunakea were leading. She had learned some of these mele and hula decades before by virtue of growing up in Hawai'i, being raised in a hula family, and learning 'Ōlelo Hawai'i. However, some of these mele were newly written, created, and composed as a way to respond to and document the change in tide that would live forever in the physical bodies, spirits, and virtual elements and realities, while contributing to an archive of embodied language practices.

Familiar language expressions were a comforting blanket, while new language phrases describing a not-so-foreign event of blatant disregard of Indigenous rights was met with hope and a desire to articulate and embody newly constructed mele that would resonate in the hearts of Kānaka Hawai'i, Indigenous Peoples, and allies across the globe. The recitation and invocation of mele summoned individuals, communities, and nations to stand in solidarity with Maunakea, while also calling upon ancestral genealogies and the more-than-human to witness, protect, and safeguard Maunakea and those kia'i who have taken on the privilege and responsibility to care for Maunakea.

One such mele that garnered attention was a collaboration between cultural practitioner Hinaleimoana Wong (also known as Kumu Hina) and musician Josh Tatofi—For the Lāhui (For the Nation), which was inspired by Tatofi's visit to Maunakea. This mele begins with describing the beloved flag waving high in the sky, calling on the lāhui to be bold, fearless, and steadfast, and ends with the words articulated on July 31, 1843 by Kamehameha III—Ua mau ke ea o ka 'āina i ka pono (The life of the land is perpetuated in righteousness)—which calls into memory the restoration of the Hawaiian kingdom following a 5-month long British occupation. The lyrical notes of this mele for this noted phrase—which is used as part of the State's motto—calls “[f]or the freedom of our land remains always in the truth of our people,” in turn giving strength and hope to Kānaka Hawai'i. The official music video included footage—a montage of photos and videos in black and white, and color of kia'i, kūpuna, community, law enforcement—capturing the sentiment, emotion, and connection to Maunakea, and documenting the continued occupation, injustice, and arrests. This mele inspired hula practitioners, like Galla who resides in Canada, to create new hula choreography that would be shared in-person and virtually in the United States, Canada, and Australia as a way to heal her wounds and to stand in solidarity with kia'i on Maunakea and other sacred lands that are under siege.

At and for Maunakea, mele and hula were used to invigorate and diffuse (Kahaunaele, 2020)—to resist, to energize, to empower, and to heal. A combination of old and new mele and hula graced the mountain in response to the violent acts which would be livestreamed for the world to see and experience. These intended and embodied expressions are examples of how Kānaka Hawai'i are continually languaging while acknowledging, respecting, and staying true to dynamic and cultural traditions.

Inventive anti-colonial Indigenous social media

Through wide-ranging, fluid expressive forms, Indigenous Peoples are resisting settler colonial violence and remaking representational politics, communicating humanity and community-building with inventive social media. Indigenous erasure, stereotyping, and misrepresentation in media naturalize colonial logics and “cycles of bias” that continue to harm Indigenous Peoples (Eason et al., 2018, p. 70; Meek, 2020). However, Indigenous influencers are showing Indigenous resistance, survivance, and thrivance through digital blogs, podcasts, arts, news, hashtag activism, research, and media movements. Related transidiomatic practices amplify Indigenous voices and engage publics in “changing the narrative,” about Indigenous Peoples to “reclaim Native truth,” (Reclaiming Native Truth, 2018; National Council of American Indians, 2019).

In 2020–21, alongside racial turmoil, violence, and COVID-19 pandemic devastation in Indigenous communities, we witnessed an explosion of Indigenous inventive languaging practices across social media platforms, in spite of entrenched digital divides (Howard and Morris, 2019). Two illustrative examples, discussed below, generated widespread vibrant expressions of heart and spirit among Indigenous communities in 2020–21, illuminating language creativity and *duende*.

Something Else

In the 2020 US elections, in spite of suppression and misrepresentation of Native American votes, grassroots mobilization and record levels of voter turnout in Native Nations strongly influenced voting trends and historic wins in national, state and local elections. Few news correspondents, however, mentioned Indigenous voters when reporting Presidential election polling data. Insultingly, CNN lumped together multiple non-white voting groups, including Indigenous voters, under the dehumanizing label “Something Else” (Native American Journalists Association, 2020). This erasure prompted a cascade of widespread critique and commentary by Native American, Indigenous and other communities across digital platforms.

For example, in an episode of *Natives React*—a YouTube vlog focused on Indigenous media and news—co-host Patrick Willie, a Navajo YouTube influencer, summed it up:

Guys welcome back to another episode of *Natives React: Something Else*, Natives. And what does that even mean? So basically, during this election time, CNN, they put up a graph detailing the different races that had gone through the exit poll. They broke down the race by white, Latino, Black, Asian and then Something Else [...] every Native American was talking about this being labeled as Something Else.

Patrickisnavajo (2020b).

Indigenous laughter and jokes have long affirmed solidarity within and across Indigenous communities, while countering “the horrors of history,” “the continuing impact of colonization,” “the biting knowledge that living as an exile in one’s own land necessitates,” and “shabby and tricky” offerings of the contemporary world, “that needs to be treated with laughter and ironic humor” (Gunn Allen, 1992, p. 158). Online, Indigenous content creators circulated humorous and serious memes, videos, and commentary to address this newest form of erasure, anti-Indigenous racism, and settler colonial equivalency lumping Indigenous Peoples with other marginalized groups (see Tuck and Yang, 2012).

Multiple content creators posted humorous Something Else TikTok video sketches. An example from Brett Mooswa, a Makwa Sahgaiehan First Nation influencer, starts as Mooswa completes a joke to someone offscreen left:

“So I says to Roger, you can’t put milk in it, it’s not even a cereal” [laughing]

An unseen speaker (also Mooswa) interrupts from offscreen front, asking in a high, tentative-sounding voice:

“Excuse me, are you Native American?”

Clearing his throat, Mooswa puts on a severe, “Hollywood Indian,” Richard Curtis photograph expression. Staring up into the distance, he answers in low, slow voice:

“No, I am ... Something Else”.

The offscreen, high-pitched voice replies, excitedly,

"Oh my God!"

The clip showcases Mooswa's linguistic flexibility and humor, echoing the anti-colonial comedy of Indigenous social media creators like the 1491s (Berglund, 2016; Goeman, 2011) and Indigenous parodies of white people (Basso, 1979). Above, Mooswa jokes in a relaxed, unmarked conversation, disrupting stereotypes of Indigenous North Americans as stoic, and settler assumptions of Indigenous North Americans as existing in the past (Berglund, 2016). He then performs a second speaker, using "mock white girl" voice (Slobe, 2018), higher pitch, and excessive politeness (Carris, 2011) to parody the racial microaggression of a white woman accosting him, and asking him about his "Native American"-ness. Clearing his throat to show himself shifting into performance mode, Mooswa dons the gaze and stance of a stereotypical, "mystical Indian" persona. He replies, using ponderous delivery and exaggerated mid-phrase pausing of "Hollywood Injun English"—a media-produced, imagined speech variety that indexes racialized assumptions about Indigenous North Americans (Meek, 2006, 2020). Back to offscreen mock white girl voice, Mooswa exclaims "Oh my God!" with exaggerated, hi-lo Valley girl intonation and phrasing (Slobe, 2018). Through dexterous style-shifting, Mooswa critiques Indigenous stereotypes and related settler colonial assumptions, as well as Indigenous performances of stereotypical indigeneity for others (Berglund, 2016). He also parodies a "hyperwhite" gendered and classed style associated with stereotypically "empty-headed," privileged white women (Eckert, 2008; Slobe, 2018) to critique whiteness and non-Indigenous surprise at encountering Indigenous Peoples and ways that Indigenous North Americans are racially "othered" in everyday interactions.

Like many Something Else jokes, Mooswa's clip further shows how commonplace interactions and struggles over Indigenous self-definition become even more absurd and dehumanizing when "Something Else" replaces terms like "Native American" and "Indigenous" that already obscure Indigenous North Americans' unique, diverse, and multiple identities. Using wide-ranging remixes of images, edited lines, and media references, Indigenous creators quickly, playfully, and pointedly connected Something Else to symbolic violence in old Westerns and children's movies, racial regulation in documents tracking blood quantum and tribal citizenship, DNA test-related challenges to Indigenous Peoples' ways of determining community membership, and colonization itself. Among other moves, they also remixed Something Else with ironic lines about indigeneity from Indigenous-produced media, as the fresh, multi-platformed, multimodal collective burst of expression rejected Something Else as more of the same old settler colonial identity politics.

Amid widespread election anxiety, creators also shared love and admiration for Indigenous strength in humor with posts like, "The world is all stressed out and Natives are just laughing all day about being labeled Something Else by CNN. I love us." As Sherry McKay, an Ojibway Anishinabe influencer from Sagkeeng First Nation concluded, in her own curated mix of Something Else posts, "That's just one of the 1000 reasons why I love being Indigenous, because we're resilient."

Natives React host Willie reflected,

It's becoming a situation where all of us Indigenous People are taking away the power of this Something Else, of being the Other. Like, "OK, sure," in a sarcastic manner. And it is pretty neat that we're not letting that define us, we're making fun of it, this is how we do Something Else is now being replaced with something deadly, talking about changing the narrative.

Patrickisnavajo (2020b).

Comments underscored the familiar, collective power of Indigenous content creators' spontaneous, critical humor. Remixing Something Else into "something deadly"—dangerously subversive and excellent—Indigenous social media creators helped a media movement "change the narrative," with inventive commentary, Indigenous connection, and community.

Playing the way he felt it: Doggface208 "vibes indigenously"

The digital universe provides growing modalities in which individuals create and express ideas and experience, drawing upon Indigenous strategies of remembering, reframing, intervening, representing, and connecting (Berglund, 2016; Smith, 2012) and Indigenous inventiveness through languaging. The viral TikTok from 2020 described below takes Blommaert's "language-in-motion" (2010, p. 5) figuratively and literally, through an interplay of linguistic and physical representations. TikTok is described as a place of "creativity and experimentation" (Anderson, 2020, p. 9) where users "play without adhering to visual styles and narratives" (Bresnick, 2019, p. 10). As part of the long tradition of Indigenous expressions of affirmation and community building, Indigenous TikTok creators are playing in the new environment, offering robust shows of creativity (In Focus, 2021).

One illustrative example is the viral TikTok "Morning Vibe" by Nathan Apodaca, also known by his TikTok handle, Doggface208. Posted on September 25, 2020, the evocative video shows Apodaca casually maneuvering his longboard on a major thruway, while singing along to the Fleetwood Mac song, "Dreams," drinking from a gallon-sized container of Ocean Spray Cran-Raspberry juice, all while, with seeming effortlessness, filming himself with his cell phone. Apodaca's video deeply touched something. Maybe the *duende*? In Lorca's words he seems to embody "the blush of all that is truly alive" (Gilmore, 2016, p. 55).

The TikTok embodies Apodaca's casual coolness, jamming out as he skateboards along the highway. He begins with a direct look at the camera. His slight smile and upward nod of recognition immediately ignites connection and familiarity; his shaved head and brown skin connote a person of color and communities claimed Apodaca as a kindred relative, referring to him as "an old school Chicano." Approximately 5 seconds into the video as he navigates his board, Apodaca turns his head away from the camera, exposing a large two feather tattoo behind his right ear stretching down to his neck, suggesting his Indigeneity. Indigenous Peoples claimed him in numerous posts, referring to his "deadly uncle vibes," (the honorific "uncle" is common in many Indigenous communities) as well as calling him "cousin" and "elder." In subsequent interviews, Apodaca identified himself as Northern Arapaho and Mexican.

The duende-filled "vibe" that so many connected to in the video weaves a deep tapestry. It is Indigenous survivance as an immediate action within a historic context. Apodaca evokes desire to survive, to be freshly alive, to move beyond the death, destruction and despair of recent years. One commenter noted that Apodaca's TikTok was "The One Positive Thing in 2020." The clip connects to a long history of resistance to settler colonial tactics. His skateboarding is part of a history of Native American skateboarding culture. His singing of "Dreams," evokes survivance and thrivance by centering Indigenous ways of being in the world. Apodaca's old truck, with 300 thousand miles on it, had broken down on his way to work. Undeterred, he grabbed his longboard and his juice and continued riding on to his destination. The challenges that morning make his "vibe" all the more a celebrated example of thrivance.

The 84 million views of Apodaca's clip is only one indicator of the video's impact. Many news stories explored the phenomenon of the video. The videos have been liked and commented on thousands of times, as well as inspiring multiple memes and copy videos, including one by a Fleetwood Mac band member. Many commenters describe the "vibe" of the video, noting its "Chill," "At Peace," and bringing "the world together" esthetic. Apodaca's "vibe" resonated powerfully with many in Indigenous communities.

The Natives React channel dedicated an episode (Patrickisnavajo, 2020a) to celebrate Apodaca, opening with the hosts drinking from the now iconic gallon sized Ocean Spray Cran-Raspberry juice bottles and "vibing Indigenously." The hosts share a humorous and moving video montage of the different ways Indigenous Peoples reinvented the video—lip syncing to Dreams while horseback riding, ice skating, and riding bikes. "I think it is so awesome to see the Native Community come together ... it warms my heart" (Patrickisnavajo, 2020a).

With so many Native communities connecting to this "Rez Anthem," the lyrics themselves offer a poetic lens of *duende* to consider,

Now, here you go again

You say, you want your freedom

Well, who am I to keep you down

It's only right that you should

Play the way you feel it

Nicks (1977).

Willie notes the video is a "reminder to enjoy life, to not take it so seriously (Patrickisnavajo, 2020a)." As he calls out to "play the way you feel it" maybe Apodaca is encouraging Indigenous Peoples to be who they are in the ways they decide; to assert their own Indigenous identity, authenticity and sovereignty. Apodaca's embodiment of Indigeneity offers an opportunity to explore representation, as social media practices of Indigenous Peoples illustrate displays of culture.

Outside observers may create or reify artificial signs of authenticity and identity of Indigenous Peoples, imposing biased views, relying on master narratives (Takaki, 2012) and single stories that restrict Indigenous expressions of being. Concepts of authenticity are "embedded ... within several colonialist systems ... held hostage by fossilized romantic notions of cultural purity and essentialist origins" (Urrieta, 2017, p. 256). Other than the feathers in Apodaca's tattoo, there is little in the video which would strike community outsiders as "authentic" Indigenous cultural practice. Nevertheless, Apodaca was almost immediately recognized and claimed by Indigenous Peoples. Apodaca's liberating authenticity is deeply felt by all audiences as they glide with him—immersed in his five senses—playing the way he feels it.

Indigenous content creators assert Indigenous “presence over absence, deracination, and oblivion,” (Vizenor, 2008, p. 1), and use humor to “laugh at”—and love—“themselves, laugh at others and hold all aspects of life together without letting anybody drive them to extremes” (Deloria, 1969, p. 53), renouncing “the unbearable sentiments of tragedy, and the legacy of victimry” (Vizenor, 2008, p. 1). Their offerings feed creative glory and Indigenous life force into the world.

Conclusion

The quest for creative expressive forms, inspiration and *duende* is a driving force in all the arts and evokes a deep basic human yearning for passion, love, joy, meaning, escape and freedom from the challenges of life. While the *duende* might be manifest in a wide variety of meaningful ways in Indigenous life, in this essay we have chosen to highlight some of the particular ways Indigenous Peoples and communities draw on the *duende* to struggle creatively with the profound and lasting wound of colonial oppression. In each of these three stories, the spirit of *duende* is evoked in Indigenous contexts by *the wound that never heals*. In each case the newly generated, “freshly created” social communicative practices bring radical change in forms of affirmation, healing and transcendent resilience to resist the enduring pain and trauma of settler colonial oppression.

Almost 50 years ago two young children playfully and lovingly invented a liberating language that defiantly transformed a rigid racialized colonial space in a remote Indigenous Kenya hillside community. The traditional cultural practices of Hawaiian mele and hula were poignantly, powerfully and politically (re)created and innovatively transformed with healing grace and evocative spirit as Kānaka Hawai‘i, Indigenous Peoples, communities, and nations stood in solidarity with *kia‘i* on the mauna. Indigenous social media content creators engaged in widespread transidiomatic practices of resistant healing humor, irony, survivance and thriving with the reversal of a narrative of erasure in the global #Something Else phenomenon and the hypnotic mesmerizing “cool vibe” of Apodaca’s viral video. Each example highlights the powerful struggles and effective agency of *just languaging* with *duende* as healing communicative practices in the face of settler colonial oppression. The poignance, passion, pain and pleasures of the *duende* are powerfully manifest through social semiotic communicative inventive practices in a battle for survival, pride, joy and community renewal. The heightened state of emotion, expression and heart in these examples inspires human resilience, healing creativity, liberating freedom and innovative *just languaging* in the *duende’s* spirit of evocation.

References

- Alfred, T., Cornsattel, J., 2005. Being Indigenous: resurgences against contemporary colonialism. *Gov. Oppos.* 40 (4), 597–614.
- Allen, P.G., 1992. *The sacred hoop: Recovering the feminine in American Indian traditions*. Beacon Press.
- Anderson, K.E., 2020. Getting Acquainted With Social Networks and Apps: It is Time to Talk About TikTok. *Library Hi Tech News*.
- Anzaldúa, G., 1987. *Borderlands/La Frontera: The New Mestiza*. Ann Lute Books, San Francisco, CA.
- Apodaca, N., 2020. Morning Vibe. Tik Tok. https://www.tiktok.com/@420doggyface208/video/6876424179084709126?lang=en&is_copy_url=1&is_from_webapp=v2.
- Basso, K., 1979. *Portraits of the Whiteman*. Cambridge University Press, Cambridge.
- Becker, A.L., 1995. *Beyond Translation: Essays Toward a Modern Philology*. University of Michigan Press, Ann Arbor, Michigan.
- Berglund, J., Carlson, B., 2020. *Indigenous Peoples Rise Up: The Global Ascendancy of Social Media Activism*. Rutgers University Press, New Brunswick.
- Berglund, J., 2016. “I’m just as indian standing before you with no feathers popping out of my head”: critiquing indigenous performativity in the youtube performances of the 1491s. *AlterNative* 12 (5), 541–557.
- Blommaert, J., 2010. *The Sociolinguistics of Globalization*. Cambridge University Press, Cambridge.
- Bresnick, E., 2019. Intensified Play: Cinematic Study of Tiktok Mobile App. Research Gate. Retrieved from: http://www.researchgate.net/publication/335570557_Intensified_Play_Cinematic_study_of_TikTok_mobile_app.
- Carris, L., 2011. La voz gringa. Latino stylization of linguistic (in)authenticity as social critique. *Discourse Soc.* 22 (4), 474–490.
- Deloria, V., 1969. *Custer Died for Your Sins: An Indian Manifesto*. MacMillan Publishers, New York.
- Eagle Shield, A., Paris, D., Paris, R., San Pedro, T., 2020. *Education in Movement Spaces: From Standing Rock to Chicago Freedom Square*. Routledge.
- Eason, A., Brady, L., Fryberg, S., 2018. Reclaiming representations & interrupting the cycle of bias against Native Americans. *Daedalus* 147 (2), 70–81.
- Eckert, P., 2008. Where do ethnolocks stop? *Int. J. Biling.* 12 (1–2), 25–42.
- Eckert, P., 2012. Three waves of variation study: the emergence of meaning in the study of variation. *Annu. Rev. Anthropol.* 41, 87–100.
- Fabian, J., 1986. *Language and Colonial Power: The Appropriation of Swahili in the Former Belgian Congo 1880–1938*. University of California Press, Berkeley, Ca.
- Galla, C.K., Galla, L., Keawe, D., Kimura, L., 2015. Perpetuating hula: globalization and the traditional art. *Pac. Arts* 14 (1–2), 129–140.
- Garcia, O., Wei, L., 2014. *Translanguaging, Language and Bilingualism in Education*. Palgrave MacMillan, New York.
- Gilmore, P., Wyman, L., 2013. An ethnographic long look: language and literacy over time and space in Alaska Native communities. In: Hall, K., Cremin, T., Comber, B., Moll, L. (Eds.), *International Handbook of Research on Children’s Literacy, Learning and Culture*. Wiley Blackwell, pp. 121–138.
- Gilmore, P., 2009. Our language: a friendship that transcended linguistic and cultural borders. In: Shweder, R., Bidell, T., Dailey, A., Miller, P., Model, J. (Eds.), *The Child: An Encyclopedic Companion*. University of Chicago Press, pp. 546–547.
- Gilmore, P., 2011. We call it “our language”: a children’s Swahili pidgin transforms social and symbolic order on a remote hillside in up-country Kenya. *Anthropol. Educ. Q.* 42 (4), 370–392.
- Gilmore, P., 2016. *Kisiki (Our Language): The Story of Colin and Sadiki*. Wiley- Blackwell, Hoboken, New Jersey.
- Gilmore, P., 2017. Ethnography with love and *duende*: the promise of ethnography in a democratic society. *Anthropol. Educ. Q.* 48 (4), 349–353.
- Gleitman, L.R., Newport, E.L., 1995. The invention of language by children: environmental and biological influences on the acquisition of language. In: Gleitman, L.R., Liberman, M. (Eds.), *Language: An Invitation to Cognitive Science*. MIT Press, Cambridge, pp. 1–24.
- Goeman, M., 2011. Introduction to indigenous performances: upsetting the terrains of settler colonialism. *Am. Indian Cult. Res. J.* 35 (4), 1–18.
- Gunn Allen, P., 1986. *The Sacred Hoop: Recovering the Feminine in American Indian Traditions*. Beacon Press, Boston, MA.
- Hollington, A., Nassenstein, N., 2018. African youth language practices and social media. In: Ziegler, A. (Ed.), *Youth Languages*. De Gruyter, Berlin, pp. 807–828.

- Horner, M., Muñoz, J., Petrone, R., 2021. Ni keetawmwi mooshahkinitounawn: lifting up representations of Indigenous education and futures in the marrow thieves. *Res. Divers. Youth Lit.* 4 (1), 7.
- Howard, B., Morris, T., 2019. Tribal Technology Assessment: The State of Internet Service on Tribal Lands. Arizona State University, Phoenix, AZ.
- In Focus, 2021. A Look at Indigenous Creators Taking #Native Tik Tok by Storm. Retrieved from: <https://www.youtube.com/watch?v=Qt1PJvgq-Eg>. (Accessed 3 September 2021).
- Jacquemet, M., 2005. Transidiomatic practices: language and power in the age of globalization. *Lang. Commun.* 25, 257–277.
- Kahaunaele, K., 2020. Mele and “ōlelo Hawai‘i on the mauna. *Biography* 43 (3), 541–550.
- Kumu Pono Associates, 1999. Mauna Kea Oral History Study and Archival Literature Research. <https://www.kumupono.com/ethnographic-studies/>.
- Lorca, F.G., 1933/1998. Theory and play of the duende. In: Maurer, C. (Ed.), *In Search of Duende*. New Directions, New York, pp. 56–72.
- Meek, B., 2006. And the injun goes “how!”: representations of American Indian English in white public space. *Lang. Soc.* 35, 93–128.
- Meek, B., 2020. Racing Indian language, languaging an Indian race: linguistic racisms and representations of Indigeneity. In: Alim, H.S., Reyes, A., Kroskrity, P. (Eds.), *The Oxford Handbook of Language and Race*. Oxford University Press, New York, Oxford, pp. 369–397.
- Mufwene, S., 2008. *Language Evolution: Contact, Competition and Change*. Continuum International Publishing Group, New York.
- Muñoz, J., 2020. Waldorf as an educational path in native America. *J. Am. Indian Educ.* 59 (2–3), 195–208.
- Myers-Scotton, C., 1978. Language in East Africa: linguistic patterns and political ideologies. In: Fishman, J.A. (Ed.), *Advances in the Study of Societal Multilingualism*. Mouton, The Hague, pp. 719–760.
- Nassenstein, N., Hollington, A., 2016. Global repertoires and urban fluidity: youth languages in Africa. *Int. J. Sociol. Lang.* 242, 171–193.
- National Congress of American Indians, 2019. *Becoming Visible: A Landscape Analysis of State Efforts to Provide Native American Education for All*. Washington, DC.
- Native American Journalists Association, 2020. NAJA Demands Cnn Apologize for Using “Something Else” to Describe Native Voters. NAJA. Website. <https://najanewsroom.com/2020/11/05/naja-demands-cnn-apologize-for-using-something-else-to-describe-native-voters/>.
- Ndege, P.O., 2009. Colonialism and Its Legacies in Kenya. Fulbright-Hays Group Lecture, Moi University, Nairobi, Kenya.
- Nicks, S., 1977. *Dreams. On Rumors*. Warner Bros.
- Paris, D., 2019. Naming beyond the white settler colonial gaze in educational research. *Int. J. Qual. Stud. Educ.* 32 (3), 217–224.
- Patrickisnavajo, 2020a. Fleetwood Mac Skateboarding Guy is Indigenous!—Natives React #26. YouTube. https://www.youtube.com/watch?v=zJh_Blbymyl.
- Patrickisnavajo, 2020b. CNN Calls Native Americans “Something Else” Natives React #29. YouTube. <https://www.youtube.com/watch?v=DWAg8PV2lnc>.
- Reclaiming Native Truth, 2018. *Research Findings: Compilation of All Research*. Echo Hawk Consulting & First Nations Development Institute.
- Riley, K., 2015. New and emergent languages. In: Bonvillian, N. (Ed.), *The Routledge Handbook of Linguistic Anthropology*. Routledge, Taylor and Francis, NY, NY.
- Scott, D.C., 1920. Library and Archives Canada, Record Group 10, vol. 6810, file 470-2-3. Testimony Before the Special Committee of the House of Commons Examining the Indian Act Amendments of 1920, vol. 7, pp. 55 (L-3)–63 (N-3).
- Silva, K., 2014. A contemporary response to increasing mele performance contexts. In: Carroll, J., McDougall, B., Nordstrom, G. (Eds.), *Huihui: Navigating Art and Literature in the Pacific*. University of Hawai‘i Press, Honolulu, pp. 24–45.
- Skutnabb-Kangas, T., 2011. Language ecology. In: Ostman, J., Verschueren, J. (Eds.), *Pragmatics in Practice*. John Benjamins, Amsterdam, pp. 177–198.
- Slobe, T., 2018. Style, stance, and social meaning in mock white girl. *Lang. Soc.* 47, 541–567.
- Smith, L.T., 2012. *Decolonizing methodologies: Research and Indigenous Peoples*, (Second Ed.). Zed Books.
- Smithsonian, 2013. *Hula: Preserving Native Hawaiian Language and Culture*. Retrieved from: <https://youtu.be/upK-xODtd50>.
- Stillman, A., 2001. Re-membering the history of the Hawaiian hula. In: Mageo, J.M. (Ed.), *Cultural Memory: Reconfiguring History and Identity in the Postcolonial Pacific*. University of Hawai‘i Press, Honolulu, pp. 187–204.
- Takaki, R., 2012. *A Different Mirror: A History of Multicultural America*. Back Bay Books, New York.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolonization: Indigeneity, Education & Society* 1 (1), 1–40.
- Urrieta, L., 2017. Identity, violence, and authenticity: challenging static conceptions of indigeneity. *Lat. Stud.* 15 (2), 254–261.
- Vizenor, G., 2008. *Survivance: Narratives of Native Presence*. University of Nebraska Press, Lincoln.
- Wyman, L., McCarty, T., Nicholas, S., 2013. *Indigenous Youth and Multilingualism: Language Identity, Ideology, and Practice in Dynamic Cultural Worlds*. Routledge, New York, pp. 1–25.
- Wyman, L., Galla, C., Jimenez-Quispe, L., 2016. Indigenous youth language resources, educational sovereignty and praxis: connecting a new body of language planning research to the legacy of Richard Ruiz. In: Hornberger, N. (Ed.), *Honoring Richard Ruiz and His Work on Language Planning and Bilingual Education*. Multilingual Matters, Bristol, UK, pp. 396–430.
- Wyman, L., 2012. *Youth Culture, Language Endangerment and Linguistic Survivance*. Multilingual Matters, Clevedon.

Relevant websites

- 1491s Official Website. 1491s Official Website, <http://www.1491s.com>.
- Adiche: The Danger of a Single Story, https://www.ted.com/talks/chimamanda_ngozi_adichie_the_danger_of_a_single_story?language=en.
- All My Relations Podcast: Blood Quantum, <https://podcasts.apple.com/us/podcast/ep-11-love-in-the-time-of-blood-quantum/id1454424563?i=1000456295864>.
- American Bar Association: Suppression of Indigenous Rights, https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/voting-rights/how-the-native-american-vote-continues-to-be-suppressed/.
- BBC: New Zealand River That Became a Person, <https://www.bbc.com/travel/article/20200319-the-new-zealand-river-that-became-a-legal-person>.
- Brett Mooswa TikTok, <https://www.tiktok.com/@brettstoise?lang=en>.
- Civil Beat: Hashtag Heard ‘Round the World, <https://www.civilbeat.org/2015/04/hashtag-heard-round-the-world-social-media-fuels-tmt-protests/>.
- Crosscut: Native Vote Data, <https://crosscut.com/politics/2020/12/shining-light-native-vote-data>.
- Cultural Survival: Yurok Nation Nature Rights, <https://www.culturalsurvival.org/news/yurok-nation-just-established-rights-klamath-river>.
- Cutting of the Hae Hawai‘i, <https://www.facebook.com/100001276351017/videos/2512898462096008/>.
- Database: Na Mele No Maunakea, https://docs.google.com/spreadsheets/d/1n628TQV6m6xNX7HX_QNhhWoJcob8yH_vEd8ly1bl6-s/edit?fbclid=IwAR2yA4U71qYxol3CJWwgBgVhqpUAtoWHTsn3AcNfunekdQt5zoPhWqngUJ#gid=0.
- Doggface208: Morning Vibe, https://www.tiktok.com/@420doggface208/video/6876424179084709126?is_copy_url=0&is_from_webapp=v1&sender_device=pc&sender_web_id=6894292046585267717.
- Doggface208 Morning Vibe Commentary, <https://www.tiktok.com/@420doggface208/video/6876424179084709126?lang=en>.
- Doggface208 TikTok, <https://www.tiktok.com/@420doggface208?>
- ExtraTV: Doggface208 Interview, <https://www.youtube.com/watch?v=fwBCYwyOx8>.
- Facebook: Letter Writing and Advocacy Work, <https://www.facebook.com/keoladonaghy/posts/10158779630291808>.
- Facebook: Preserved Oral and Embodied Practices, <https://www.facebook.com/AdamKeaweM/posts/4114694008648130>.
- Federico Garcia Lorca, <https://www.youtube.com/watch?v=ll99nCRNoFg>.
- For the Lāhui Song, https://kanaekana.net/wp-content/uploads/2019/09/For_The_Lahui.pdf.

Hawai'i Review, <http://hawaiiireview.org>.
IllumiNatives, <https://illuminatives.org>.
InFocus: A Look at Indigenous TikTok, <https://www.youtube.com/watch?v=Qt1PJvgq-Eg>.
Indian Country Today, <https://indiancountrytoday.com>.
Indigenous Peoples of Africa: The Forgotten Peoples? https://www.achpr.org/public/Document/file/Any/achpr_wgip_report_summary_version_eng.pdf.
Indigenous Podcasts to Follow, <https://indiginews.com/okanagan/top-ten-indigenous-podcasts-to-follow>.
Interview With Perry Gilmore, <https://almagottlieb.com/2017/05/interview-with-perry-gilmore-about-kisili-our-language-the-story-of-colin-and-sadiki/>.
Kim TallBear: Native American DNA, <https://www.upress.umn.edu/book-division/books/native-american-dna>.
Kisili Book Talk, [https://www.kaltura.com/index.php/extwidget/preview/partner_id/1909371/uiconf_id/36513372/entry_id/0_0zc6vj3/embed/auto?&flashvars\[streamerType\]=auto](https://www.kaltura.com/index.php/extwidget/preview/partner_id/1909371/uiconf_id/36513372/entry_id/0_0zc6vj3/embed/auto?&flashvars[streamerType]=auto).
KTVB: Doggface208 Story, <https://www.youtube.com/watch?v=4DXeXocagSc>.
Kūha'o Maunakea Music Project, <https://www.bigislandvideonews.com/2019/09/30/kuhao-maunakea-music-project-to-benefit-ku-kiai-mauna-movement/>.
Library Guides: Aloha 'Āina, <https://guides.westohu.hawaii.edu/c.php?g=977248&p=7065788>.
Lorca's Duende, <https://www.youtube.com/watch?v=4LidTVzW96U>.
Morning Vibe TikTok Copy, <https://www.youtube.com/watch?v=iKkNuAJBo-w>.
National Council of American Indians Report: Becoming Visible, https://www.ncai.org/policy-research-center/research-data/prc-publications/NCI-Becoming_Visible_Report-Digital_FINAL_10_2019.pdf.
Native Appropriations, <http://nativeappropriations.com>.
Native Arts and Culture Foundation, <https://www.nativeartsandcultures.org>.
Natives React: Doggface208 Episode, https://www.youtube.com/watch?v=zJh_Blbymyl&t=467s.
Protecting Maunakea Day 3, <https://www.facebook.com/watch/?v=485669822180209>.
Pūliima 2019 Women in Language Workshop (Australia), <https://www.facebook.com/watch/?v=1505814296276162>.
Pū'uhuluhulu University, <https://puuhuluhulu.com/learn/university>.
Reclaiming Native Truth, <https://rnt.firstnations.org>.
Sacred Land: Mauna Kea, <https://sacredland.org/mauna-kea-united-states/>.
Sacred Land: Bears Ears, <https://sacredland.org/bears-ears/>.
Sacred Land: Kaya Forests, <https://sacredland.org/kaya-forests-kenya/>.
Sacred Land: Machu Picchu, <https://sacredland.org/machu-picchu-peru/>.
Sacred Land: Mount Fuji, <https://sacredland.org/mount-fuji-japan/>.
Sacred Land: Mount Sinai, <https://sacredland.org/mount-sinai-egypt/>.
Sacred Land: Uluru, <https://sacredland.org/uluru-kata-tjuta-australia/>.
Sherry McKay TikTok, <https://www.tiktok.com/@sherry.mckay>.
Society for Pidgin and Creole Linguistics Webpage, <https://sites.google.com/site/societypidgncreolelinguistics/home>.
The Thirty Meter Telescope Project, <https://www.tmt.org>.
The Conversation: Rights for Nature, <https://theconversation.com/rights-for-nature-how-granting-a-river-personhood-could-help-protect-it-157117>.
Theory and Play of Duende, <https://www.youtube.com/watch?v=gtPij2xRcM>.
University Courses: Mauna, https://docs.google.com/spreadsheets/d/15MVemAQDq_-qsdrM5iduweDeo4CHV-69FpOo7yWZLE/edit?fbclid=IwAR3qnwplHpJMuc2Nu1V8oeXzMzp304tJpW7kgh1MQGTn1ChOpvCPUuPUNw#gid=0.
YouTube: Jam for Maunakea, <https://www.youtube.com/watch?v=k1Ul5xp4PTg&t=6s>.
YouTube: Natives React—CNN Episode, <https://www.youtube.com/watch?v=DWAg8PV2Inc>.

Mobilizing prefigurative hybrid literacies for decolonial resistance: three transnational examples

Astrid N. Sambolin Morales^a, Ericka Stauffert-Reyes^b, and Gerald Campano^b, ^a Kent State University, OH, United States; and ^b University of Pennsylvania, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	809
Prefigurative hybrid literacies	810
Transnationalism, precarity, and decolonial resistance	810
Puerto Rico	811
Imperialism and resistance in “La Isla del Encanto”	811
Youth literature, Puerto Rico, and the diaspora	811
La Borinqueña: celebrating and nurturing transnational Identities	812
Mexico	812
Historical roots of Ralámuli’s present context	812
The Ralámuli Nation’s literacies for resistance	813
The Philippines	814
Context of imperialism and resistance	814
“We call her Ina Bai”: mobilizing young adult fiction and storytelling	815
Concluding thoughts	815
References	815

Introduction

San Juan, Puerto Rico - Verano de 2019

Puerto Ricans engaged in a 3-week protest to oust then-governor Ricardo A. Roselló. Protesters gathered in front of La Fortaleza, the governor’s residence, and, in one of the more memorable acts of resistance, read “bedtime stories” to the Puerto Rican Constitution. This literal and symbolic act emerged as a challenge to the assaults Puerto Rico’s Police combat squad, la Fuerza de Choque, had consistently committed around 11 p.m. each night, claiming protestors’ constitutional rights expired after 11 p.m. because “La Constitución Duerme/the Constitution sleeps.” Feminist organizations and children’s authors read aloud the U.S. Constitution and children’s books to the Puerto Rican Constitution (1952) as bedtime stories. After three weeks of protests across the island, Roselló announced his resignation.

Paraphrased from [García \(2021\)](#).

We start this piece with a snapshot of *prefigurative hybrid literacies* in action. In the example above, Puerto Rican citizens took to the streets and used the United States constitution and Puerto Rican children’s literature to enact critiques of callous, abusive policies and practices by the Puerto Rican (pro-U.S. statehood) government. Although the reading of bedtime stories stemmed as a response to la Fuerza de Choque’s violent assaults, the acts also speak directly to the multiple violent, oppressive practices the governor and other high-level government officials engaged in as Puerto Ricans on the archipelago struggled with an ongoing economic crisis and the aftermath of Hurricanes Irma and Maria. Hundreds of leaked messages between Roselló and government officials showed the illicit use of public resources and funds for partisan purposes, misuse of public funds and time for non-government work, and misogynist, classist, disparaging remarks against the Puerto Rican community. By bringing in literacy practices considered as belonging in the private sphere to a public space, participants engaged in hybrid literacy practices to challenge oppressive, dehumanizing practices by those in power. Bedtime stories and protests, practices typically regarded as belonging to separate social contexts, are thus taken up together to generate new discourses—a third space—and *prefigure* the equitable, just society Puerto Ricans wish and deserve to live in.

The following chapter provides three examples of prefigurative hybrid literacies mobilized by transnational, subaltern communities in Puerto Rico, Mexico, and the Mindanao region of the Philippines, respectively. Despite their disparate geographical locations, these three contexts share histories of imperialism and colonization, their communities experiencing ontological precarity through constant exposure to economic insecurity, injury, violence, and forced migration ([Butler, 2004](#)). This piece will illustrate how these communities’ engagement with prefigurative hybrid practices allows disenfranchised, nondominant ([Gutiérrez, 2008](#); [Gutiérrez et al., 2009](#)) people to create third spaces ([Gutiérrez, 2008](#)) by drawing from multiply-situated knowledges. Such an

approach nurtures decolonial imaginaries (Pérez, 1999; Torres and Medina, 2021) “resulting in geographical and ideological repositioning” (Schechter, 2012, p. 2) that delinks these communities and lands from the dominant colonial US imaginary.

Prefigurative hybrid literacies

In this piece, we use the term *literacies* to capture local, multiple, and changing competencies and practices of language usage in the service of interacting with and manipulating texts, which shape and are shaped by sociocultural practices, contexts, mediums of representations, and users’ identities (Jocson, 2005/2006). Moreover, Jocson (2005/2006) defines *hybridity* as “the convergence and (re)emergence of cultural practices that create or lead to alternative possibilities in individual or social but goal-oriented activities such as learning and production” (p. 51). We also wish to highlight the term’s “community-based yet experimental, radical, but not static or dogmatic” nature (Gomez-Peña, 1996, p. 12). By using the term *hybridity*, we hope to capture how “we are not the product of just one culture; that we have multiple and transnational identities; that we contain a multiplicity of voices and selves, some of which may even be contradictory” (Gomez-Peña, 1996, p. 12).

Birr Moje (2013) describes hybrid literacies as intersecting knowledges, expressions, and ways to make meaning:

Various scholars have taken up the ideas from hybridity theory and applied them to the idea that the literate practices valued in home, community, peer groups, and schools can come together in unpredictable ways to produce new literate practices or conventional practices (both everyday and academic) used for unconventional purposes.

Birr Moje (2013, p. 361).

This idea of two practices coming together to create a new practice (as opposed to merely a combination of the two) is what links hybridity to third space (Gutiérrez, 2004). Third spaces center “in-betweenness” and the intersection of differing practices, identities and possibilities to create something new (Birr Moje, 2013). As with the opening example, hybridity involves “the integration of competing knowledges and discourses” (Birr Moje, 2013, p. 361); in a third space, “what seem to be oppositional categories can actually work together to generate new knowledge, new Discourses, and new forms of literacy” (Birr Moje, 2013, p. 362). These new discourses, forms of knowledge and literacies stem from tensions and ruptures (Jocson, 2005/2006).

Finally, we use the term *prefiguration*, which has links to decolonial resistance. Walia (2013) in particular notes:

Decolonization is a framework that offers a positive and concrete prefigurative vision. Prefiguration is the notion that our organizing reflects the society we wish to live in—that the methods we practice, institutions we create, and relationships we facilitate within our movements and communities align with our ideals.

Decolonizing Movement Borders section, para. 3.

Moreover, Walia (2013) argues that “decolonization movements create an alternative to power through committed struggle against settler colonialism, border imperialism, capitalism, and oppression, as well as through concrete practices that center other ways of laboring, thinking, loving, stewarding, and living” (Decolonizing Movement Borders section, para. 5).

When we refer to *prefigurative hybrid literacies*, then, we mean integrated, oftentimes competing knowledges and discourses that generate new knowledges/discourses in the service of decolonial resistance, thus working to dismantle oppressive, racist systems that fail to align with subaltern communities’ ideals and hopeful futures. Through this decolonial resistance, subaltern communities work toward “disrupt[ing] the dominant ordering schemas of modern society, specially the binaries of colonizer/subaltern and citizen/alien, identity pairings that usually map on to a white/black (or white/nonwhite) racialized social imaginary in the United States” (Schechter, 2012, p. 4). The literacies we describe in this chapter are a form of prefigurative pedagogy as well (Campano, 2013); they were staged not only to educate a broader audience about societal injustices, but also to collectively embody alternative decolonial forms of sociality premised on interdependence in the midst of oppression.

Transnationalism, precarity, and decolonial resistance

Transnationalism captures social, economic, and political attachments individuals maintain in both their home nation-state and the nation-state(s) to which they have moved. Linkages between two or more different countries occur through remittances, marriages, media consumption, communications, and travel, to name a few (Foner, 1997; Morawska, 1999; Portes et al., 1999). Transnationalism is a concept closely linked to identity, since (a) the attachments (im)migrants maintain in multiple nation-states are often the result of (perceived) commonalities in terms of language, customs, and cultures in multiple sites, and; (b) individuals and communities must negotiate their identities across different contexts (Vertovec, 2001). Çağlar (2001) points out (im)migrants “weave their collective identities out of multiple affiliations and positionings and link their cross-cutting belongingness with complex attachments and multiple allegiances to issues, peoples, places, and traditions beyond the boundaries of their resident nation-states” (p. 610).

However, despite the constant flow of goods, ideas, and communication made possible by technological advancements, political borders have influenced the physical (im)mobility of nondominant, racialized subjects. [Walia's \(2021\)](#) work examines how empire, capitalism, and human-made disasters simultaneously displace vulnerable, oppressed populations as they limit their mobility through colonial policies and technologies that serve neoliberalism tenets. Moreover, empire and capitalism actively construct marginalization and oppression, unequally ascribing value to and protecting certain bodies over others ([Kasmir, 2018](#)).

This piece will highlight the potential of prefigurative hybrid literacies as both sites of (de)colonial ([Mignolo, 1995](#)) resistance and mediums to nurture decolonial imaginaries ([Campano et al., 2020](#); [Pérez, 1999](#)). We will focus on examples across three different contexts—Puerto Rico, Mexico, and the Mindanao region of the Philippines—with shared histories of imperialism and colonial oppression, illustrating how nondominant communities engage in prefigurative hybrid literacies to (a) equitably redistribute access to information transnationally and (b) encourage transnational allyship toward liberatory, humanizing ends ([Luke, 2012](#)).

Puerto Rico

Imperialism and resistance in “La Isla del Encanto”

Puerto Rico became a US colony in 1898, when the United States followed an expansionist agenda that included the acquisition of Spanish ultramarine territories after the Spanish-American War. Race became an organizing principle for the US nation-state during this period of capitalist global expansion ([Singh, 2004](#)), and Puerto Rico is a prime example of this dynamic, relegated to a liminal space between inclusion and exclusion. As [Burnett and Marshall \(2001\)](#) document, “the Insular Cases, decided between 1901 and 1922, invented and developed the idea of unincorporated territorial status in order to enable the United States to acquire and govern its new ‘possessions’ [Puerto Rico, Guam, and the Philippines] without promising them either statehood or independence” (p. 2).

As an unincorporated territory, the island has been the subject of more than 100 bills introduced in the US Congress that attempt to (re)establish the island’s connection to the US in terms of citizenship, nationality, and naturalization, resulting in what [Venator-Santiago \(2017\)](#) terms “inclusionary exclusion” (p. 38). Puerto Ricans became US nationals through the 1917 Jones Act and obtained limited autonomy as a commonwealth in 1952 ([Duany, 2002](#)). However, the US created a system with provisions to limit the island’s agency and economic growth ([Burnett and Marshall, 2001](#), p. 10). As members of a commonwealth, Puerto Ricans can elect a governor and local public officials, but they are unable to vote in presidential elections and have no autonomy over “currency, defense, external relations, communications, the postal service, social security, and interstate commerce” ([Godreau and Bonilla, 2021](#), p. 511). The professed intention of protecting and bringing civilization to Puerto Ricans served to give the US nation-state the high moral ground as it actively worked to exploit the land and its residents, maintaining the island’s unequal, vulnerable status in the US polity.

[Duany \(2002\)](#) argues Puerto Rico is a “postcolonial colony” that, despite US intervention, has a strong national identity (i.e., cultural nationalism). Moreover, Puerto Ricans occupy a unique position in the US polity. According to [Duany \(2002\)](#), “Puerto Rico’s migratory patterns are distinctive for: (a) the island’s “persistent colonial condition,” (b) “the sheer magnitude” of its diaspora, and (c) the circular and return patterns of migration in which Puerto Ricans have engaged since the 1940s” (p. 13).

[Lloréns’ \(2018a\)](#) argues that “in Puerto Rico, societal vulnerabilities result from a long history of colonial subjugation, economic hardship, environmental injustice, infrastructural neglect, and, at the local island level, a broken rule of law” (p. 156). She identifies the 1980s as a critical decade when the archipelago’s vulnerability increased dramatically. Social problems such as crime, poverty, and unemployment as well as greater socio-economic disparity increased, and the trend of privatization began ([Lloréns, 2018a](#)). Unfortunately, the socio-economic problems resulting from the archipelago’s colonial status were only exacerbated after the land-fall of Hurricanes Irma and María in 2017, a series of earthquakes in 2020, and the global pandemic still in effect at the time of this writing. As [Lloréns \(2018b\)](#) concludes, Puerto Rico’s current economic crisis is compounded by “the social forces of five centuries of colonization by Spain and then the United States, which largely constructed the island as a site of extraction, with its people at times invisible and at others dispensable” (p. 158).

Youth literature, Puerto Rico, and the diaspora

[García \(2021\)](#) documents how youth literature and literacies have had a prominent role in Puerto Rican politics and activism. Puerto Rican authors on the archipelago and in the diaspora have historically used texts to nurture a strong sense of cultural nationalism and make sense of Puerto Rico and Puerto Ricans’ relationship to the US. For example, [Haiman \(2014\)](#), author of *Tulipán: The Puerto Rican giraffe*, tackles what it means to be Puerto Rican, “representing migration to young readers from necessary evil to exciting inevitability to, finally, an asset for creative and political freedom” ([García, 2021](#), p. 206). Youth literature and literacies also often center a feminist transnationalist lens that highlights collectivity, survivance, and resiliency in the face of dehumanizing, oppressive conditions, continuing “a legacy of feminist and Afro and Indigenous pedagogies through engaging audiences in stories about self-determination and community solidarity as opposed to individual heroism” ([García, 2021](#), p. 196).

After Hurricanes Irma and María in 2017, Puerto Ricans lived without access to electricity, internet, and other communication technologies. Puerto Rican authors heavily relied on “Afro-Boricua literacy practices of storytelling theaters and cultivating community critical literacies, understanding that government institutions often fail to support communities of color” ([García, 2021](#), p. 28).

Torres and Medina (2021) examined 4 children's books published after Hurricanes Irma and Maria: *Mi Isla Bella/Mi Isla Hermosa* (Pastrana-Andino, 2018), *Por Ahí Viene el Huracán* (Rexach-Olivencia, 2018), *La Cucarachita Martina y el Huracán* (Marichal-Lugo, 2019), and *Thiago y la Aventura del Huracán* (Arroyo-Pizarro, 2018). They concluded these children's books illustrate communal and activist practices toward epistemic justice and foreground "decolonial ruptures that reclaim local knowledges," thus "delinking the island from a dominant colonial US landscape and imaginary" (Torres and Medina, 2021, p. 260).

I (Astrid) engage in this work as a Puerto Rican who was born and raised on the island but had no exposure to texts such as these as a child. As an adult, I finally felt what it was like to have characters in stories share my cultural and linguistic background. More importantly, I was exposed to characters and stories that centered power, strength, and resistance, and I wonder about the effect texts such as these would have had on me at a younger age. I also engage in this work as a mother who is grateful for the opportunity to share these stories and messages with my daughter. The examples in this section demonstrate Puerto Rico's rich history of prefigurative hybrid literacies, where discourses centering decolonial resistance intersect with the genre of children's literature, creating and sustaining stories and culturally and linguistically diverse practices that draw from multiply-situated knowledges and disrupt the dominant colonial US imaginary.

La Borinqueña: celebrating and nurturing transnational Identities

Edgardo Miranda-Rodríguez's comic series, *La Borinqueña*, was created in 2016 and features Marisol Rios de la Cruz, an Afro-Boricua student from Brooklyn who goes to live with her grandparents in San Juan, Puerto Rico to complete her graduate studies as an environmental studies major. While in Puerto Rico, Marisol sets in motion a series of events that lead to Taíno goddess Atabex granting her superpowers, including the ability to control storms and create portals that allow her to travel freely across borders.

Miranda-Rodríguez's work centers a diasporic intersectionality that "not only acknowledges but actually privileges the Puerto Rican migrant experience" (García, 2018, p. 72). Castillo points out how, in *La Borinqueña's* case, "the soul and spirit of the island of Puerto Rico flows through Marisol's veins *because* [emphasis added] of her upbringing away from the island as well as her return to it to seek out her roots" (Castillo, 2021, pp. 1023–1024). This positioning encourages transnational allyship between Puerto Ricans on the island and the diaspora in particular. Moreover, Marisol provides a counter-narrative Puerto Rican hero, since the protagonist deviates from the typical White, middle-class male perspective, instead foregrounding Marisol's intersectionality as an Afro-Boricua woman from a working-class, bilingual background (García, 2018). Marisol's costume gives the illusion that the character is draped in the Puerto Rican flag, she translanguages—using the named languages of English and Spanish to navigate the world—and she maintains ties in both New York and Puerto Rico. In this way, the character embodies "a diasporic Puertoricanness [that] serves as an anticolonial tool to subvert stereotypical, disempowering national myths about races, gender, and class" (Castillo, 2021, p. 1023). In short, the text *prefigures* the character's in betweenness as a source of strength to be valued and nurtured. In addition, it represents a "decolonial, anti-patriarchal turn in the Puerto Rican imagination" (García, 2021, p. 203).

Marisol's backstory and the acquisition of her powers is firmly situated in multiply-situated knowledges spanning space and time, making it a prime example of hybrid literacies in action:

Atabex shows Marisol the Taíno resistance to Spanish domination; the moment pro-independence activist Mariana Bracetti sewed the flag of Lares, the first symbol of Puerto Rican anti-colonial resistance in the 1868 rebellion against Spain; Puerto Rican soldiers fighting in the U.S. Army unfurling the Puerto Rican flag while in a trench; Nationalist Party leader Pedro Albizu Campos, convalescing from skin burns that one scholar has found were caused by covert radiation in the 1950s; and the Young Lords marching through the streets of New York City in the 1970s.

García (2018, p. 79).

This example of prefigurative hybrid literacies draws from various histories and knowledges to delink the island's history from the US imaginary and *prefigure* a community of resistance fighting for liberation (García, 2018).

Finally, another important component in Marisol's story is the sense of hope and possibility at the center. García (2021) argues Miranda-Rodríguez's work allows readers "reimagining and restoring the present" (p. 74) to engage in collective dreamwork (p. 71) or *prefiguration* involving Puerto Ricans on the island and diaspora. *La Borinqueña* is rife with references to anti-colonial Puerto Rican culture, including excerpts from poetry written by Julia de Burgos, a Puerto Rican poet and activist for women and Afro-Caribbean writers; lyrics from the revolutionary national anthem; and a costume featuring the light blue color used in the original Puerto Rican flag, later changed to navy to better resemble the US flag and gain favor with the federal authorities (García, 2018). As García (2018) argues, "by including diverse literary allusions and locations as important cultural signposts of Puerto Ricanness, Miranda-Rodríguez suggests that racial mixture, anti-colonial resistance, and cross-locational mobility and adaptability are part and parcel of the Puerto Rican experience" (p. 79).

Mexico

Historical roots of Ralámuli's present context

Each part of Mexico has a specific history of the Original Peoples' re-existence (Walsh and Mignolo, 2018) in the face of and in resistance to domination by the military and religious conquerors of the Spanish Empire and, later, by the Mestize Mexicans,

who arguably constitute the official Mexican nationality. In some parts, this resistance has had notes of coexistence and cultural syncretism, as in the case of the Nahua Peoples in the center of the country during colonial times (Escalante Gonzalbo, 2020). In other latitudes of Mexico, the Original Peoples, such as the Yaqui People in the current state of Sonora, have faced the government's attempts of extermination and genocide (Turner, 1969). And in other corners, the Original Peoples have managed to counteract domination from colonial times to the current 20th century through the isolation provided by Mother Nature, as in the case of the Ralámuli in our contemporary Chihuahua state. This section will focus on the Ralámuli nation's fight, spanning at least 70 years, and the ways in which they use their literacies to resist the dispossession of their lands. To this end, it is necessary to give a brief historical contextualization of the theory of racial formation in Mexico and of the relations between the various nations that make up Mexico (i.e., the relations of the Original Peoples with the Mestizes). This will help us better understand the great feat of the Ralámuli people in their struggle to survive in the lands they have inhabited for millennia.

Before the European colonizers arrived to Abya Yala in 1492, the Original Peoples had been established for more than 35,000 years (Escalante Gonzalbo, 2020) in what we now call Mexico. Beyond the linear temporality of modernity/coloniality (Mignolo and Walsh, 2018), these tribes had developed unique civilizations in harmony with Mother Earth, in regions formerly known as Oasisamerica, Aridoamerica and Mesoamerica. However, the arrival of the Spanish Crown and the Catholic Church conquerors to Abya Yala permanently disturbed the being and knowing of Original Peoples across the continent. Between 1521 and 1821, the political demarcation we now call Mexico was a colony of the Spanish Empire called New Spain. Its highly diverse population was constituted by Native Americans who were dispossessed of their ancestral land, Spaniard colonizers, and enslaved sub-Saharan Africans. Social relationships in this Colonial period were marked by a caste and racial system that allowed, under the auspices of the Spanish Crown and the Catholic Church, the admixing of cultures and bodies of Native, European and African peoples. By 1810, the New Spain's population was 39.5% mixed (Lagunas, 2018).

The mixed population was regulated by the Crown under an intricate hierarchical classification of ethno-racial categories known as the caste system, and the categories varied between 14 and 20 (Palacios, 2012, pp. 38–39). In this system of corporal regulation, the Mestizes were the offspring of a Spaniard male and an Indigenous woman and constituted the third most powerful ethnic group during the Colony, behind the Spanish and the Criollos. Under this racial formation theory, "the meanings of 'Spanishness,' 'Indianness,' [sic] and 'mestizeness' [sic] were formulated from a specific historical context and within a complex web of power relations where social identity and status were constantly negotiated" (Palacios, 2012, p. 28). Accordingly, the notion of Indigenous people was constructed as an amalgamation of various identities and tribes, and as a common group in need of Christian salvation. The ethno-racial category of Black people derived from the dominant ideology associating dark skin with both the biblical story of Noah's curse of Ham and Aristotle's view of this group of people as natural slaves (Palacios, 2012, p. 34). Within each racial category, additional caste distinctions and exceptions determined peoples' rights and obligations (Palacios, 2012).

After 300 years of this colonial period, the new country of Mexico continued to use systemic racism to shape its identity as a nation-state in the 19th and 20th Centuries (Palacios, 2012). Following the Independence from Spain, Mexican Mestizes were in a privileged position that allowed them, in time, to take political and cultural control of the newborn country and to create the official national identity under the thesis of mestizefilia. Rooted in legacies of coloniality, systemic racism and biological determinism that uphold Eurocentric epistemologies, mestizefilia created a cultural homogeneity as the standard of what it means to be Mexican (Palacios, 2012), negating the diversity of cultures, cosmovisions, languages, and skin colors that actually inform Mexico. Mestizefilia was also used by the founding fathers of Mexico to eliminate the cultural presence of contemporary Indigenous cultures, relegating them to Mexico's past (Palacios, 2012) and "defin[ing] the indigenous [sic] as 'problems' in need of paternalistic intervention" (Palacios, 2012, p. 71).

The Ralámuli Nation's literacies for resistance

I (Ericka) engage with the Ralámuli nation's literacies as a Mexican Mestiza born and raised in the Northern land we now call Chihuahua, that once belonged to the Original Peoples referred to as Conchos. I have been a witness to the pride and resistance the Ralámuli people have as they face the power imbalance in their relationship to Mestizes, and I'm choosing to use the privilege I have through academia to denounce systemic racism in Mexico and to honor and aid the Ralámuli nation in their fight for their ancestral lands.

In line with theories of civic nationhood (Santos, 2013), the article number two of Mexico's Political Constitution states the Mexican nation is unique, indivisible, and with a pluricultural composition. However, Indigenous peoples across Abya Yala demand that, along with civic nationhood, states should also recognize ethnic-cultural nationhoods, i.e., a kind of nationhood founded on a collective sense of belonging, and sharing of the same symbolic and cultural universe (Santos, 2013). Under this premise, we can think of Mexico as a plurinational state, where Indigenous nations, like the Ralámuli's, are constantly using hybrid literacies to create coalitions with Mestize Mexicans—members of the official Mexican nation—in resisting the push of modernity/coloniality's battles against them.

These battles have social, political, economic, and historical expressions. For instance, the economy of the country has been grounded on capitalistic extractivism since the Colonial period, with disrespect toward Mother Earth. As I write this, the federal government is carrying out the ecocide of the Selva Maya in the Southern states, in favor of a highly ambitious tourism megaproject. At the North of the country, in the land of the Ralámulis, Mother Earth is being harassed through-out her forest by mining and tourism companies. In what follows, we will focus on one example illustrating how members of the Ralámuli nation are using their

prefigurative hybrid literacies to resist and re-exist (Walsh, 2020) in front of a tourism development's efforts to take their ancestral land in the town of Mogótavo.

Mogótavo is located in the Copper Canyon, in a zone highly coveted by Mestize people for tourism purposes. Diverse Mestize businesses and politicians have been attempting to displace the Ralámulis from Mogótavo since 1909. The struggle has intensified since 1980, when the Ralámulis legally requested their land. Currently, 16 members of the Ralámuli nation, including 3 autochthonous authorities, are facing legal prosecution by politicians with property titles who wish to dispossess the community of their lands.

In the multiple rallies they organize in highways and cities, the Ralámuli people of Mogótavo use their hybrid literacies to oppose their land dispossession. The ancestral dances, songs, and music that have been practiced during millennia to honor Mother Earth in the more private spheres of home and community are now being displayed in a more public sphere. By bringing these literacies to a Mestize audience, the Ralámuli people critique and resist oppressive, dehumanizing, imperial practices and systems that wish to eradicate their histories, multiply-situated knowledges, and their very being. As in our opening vignette, ancestral dance/music and protest are brought together and generate a third space that *prefigures* the equitable, just society the Ralámulis wish and deserve to live in. Moreover, they use their literacies to create a kind of transnational coalition with Mestize people, members of the official Mexican nation, as they advocate for themselves. Beyond the presence of Mestize activists and lawyers in these manifestations and in the courtroom, this coalition building is particularly visible in an object Ralámulis have incorporated from Mestize literacies into their rallies: the banner. These objects, material agents of Ralámulis' hybrid literacies, are written in hispanicized Ralámuli language over small pieces of cardboard, or in Spanish on plastic, printed banners.

The case of the resistance and re-existence of the Ralámuli nation through their hybrid literacies can be read as a counterstory to mestizofilia. Moreover, these practices nurture decolonial imaginaries that reposition and delink community and land from the dominant colonial US imaginary. As mestizofilia depicts Indigenous peoples in Mexico as an obstacle to progress, the Ralámuli nation has learned from the literacies of the Mestizes to defend their land and their cosmovision from intruders. The strength and honor of this nation, as well as their struggles against systemic racism, have echos everywhere in the modern world, and join the small hopes (Walsh, 2020) of colonized, subaltern subjects that together crack the matrix of modernity/coloniality.

The Philippines

Context of imperialism and resistance

The Philippines, although on the other side of the globe, shares a colonial history with Puerto Rico and Mexico. Composed of more than 7000 islands and a multitude of ethnolinguistic communities, the Spanish delineated its political borders in the mid-1500s. "Las Islas Filipinas," named after King Philip II, remained a part of Spain for roughly three centuries. The ideological justification for colonization was primarily religious: to "save the souls" of the native inhabitants, who were deemed inferior. The violent imposition of Christianity worked in tandem with economic plundering through land ownership and labor exploitation. It is important to note that, prior to colonization, and for many indigenous people in the Philippines to this day, land is part of a collective commons, a source of food and communal well-being. As a result, the Philippines has a long history of native resistance to colonial occupation symbolized by iconic figures of the Spanish colonial period, including Lapu Lapu, Gabriela Silang, Jose Rizal, the Katipunan, and Emilia Aguinaldo.

In 1898, the United States "acquired" the Philippines, along with other territories such as Puerto Rico, Guam, and Samoa, from Spain after the Spanish American War (Immerwahr, 2019). The Filipinos resisted their new colonizers and, by some estimates, roughly one million Filipinos were killed during the Philippines-American War (1899–1902). While brutalizing Filipino insurgents through a scorched-earth campaign, American imperialists simultaneously employed education and literacy as a soft tool of hegemony to civilize "America's Little Brown Brothers" through "benevolent assimilation." Around 1903, roughly 500 Thomasite Educators from the United States arrived in the Philippines to establish a public education system in English. One such educator, Carter G. Woodson, drew on his experiences in the Philippines to theorize the miseducation of Black students on the continental United States, who also did not have a curriculum that connected to their history and culture (Woodson, 1933/1999). The colonial project has thus employed not only conquest and genocide but epistemicide—the destruction of indigenous knowledge systems—and the stratification of humanity into those who are human and those who are positioned as sub-human and perpetually behind in the journey toward modernity (Santos, 2014).

Although the Philippines became an independent nation in 1946, many argue that it remains to this day subject to neocolonial economic dominance by Western Countries and multinational corporations. To complicate matters even further, many communities in the Philippines, including the Lumad (an umbrella term for the indigenous people of the southern island of Mindanao) and the Moro (Muslim) may be considered inhabitants of a settler colonial society. In this case, the settlers include Filipinos from the North, as well as an autocratic State that seizes indigenous lands for corporate mining and logging. The Lumad and Moro are displaced and terrorized through practices such as "red tagging" (unjustly accusing civilians of being communist to justify persecution and murder) and the harassment of Lumad schools for striving to preserve indigenous knowledge and value systems that are tied to the land.

“We call her Ina Bai”: mobilizing young adult fiction and storytelling

Just as literacy can be a weapon of oppression, it can also be a tool for anti-colonial liberation. Jose Rizal, for example, wrote novels that helped Filipinos imagine a world free of Spanish domination (Campano and Ghiso, 2011). In contemporary times these literary legacies continue. For example, *We Call Her Ina Bai: How Strong Women Are Made* (Santiago and Valle, 2021) is a poignant and gripping narrative of the struggle and resilience of the indigenous Lumad people of the Mindanao region of the Philippines. Fighting to protect their ancestral land and way of life from the plunder of multinational corporations, government soldiers, and local paramilitary squads, the Manobo community of the Pantaron Mountain Range have endured settler colonial displacement, state sanctioned terror, and inadequate access to health care and education. In this fraught context, leaders from the community have nonetheless risen to build solidarity within and beyond the Lumad people in order to resist their oppression. *We Call Her Ina Bai* tells the story of two of these powerful leaders, Sharmaine and her grandmother, Bai Bibyaon Ligkayan Bigkay, or Ina Bai, and is itself an outgrowth of this solidarity and collective effort.

Written in lucid and often poetic prose, the book is in the spirit of indigenous testimonials such as *I, Rigoberta Menchú* (Menchú and Burgos-Debray, 1992), which focuses on the Mayan people of Guatemala. The testimonial genre tells communal history through the voices and histories of individuals, and in doing so provides a counter-story to “official” accounts and erasures perpetuated by government propaganda. *We Call Her Ina Bai* is a text that illustrates hybrid literacies at play, since it weaves back and forth in both time and perspective between Ina Bai and Sharmaine. In doing so, the text artfully captures both the power of intergenerational, ancestral knowledge and the importance of adapting to new times and growing together as a community.

I (Gerald) read *We Call Her Ina Bai* from my position as an education scholar at a university in the Global North. I have never been to the Philippines, but I have ancestral roots in Mindanao from my grandfather, who migrated to the United States as a colonial subject at the beginning of the 20th Century during the American imperial occupation of the archipelago. I never imagined that a part of the history of Mindanao and its people would ever become accessible in a Young Adult nonfiction novella that would reach my hands here in the States. I had read about the predicament of the Lumad, but only as mediated by organizations such as Human Rights Watch, not the humanizing history as told directly from Lumad women leaders themselves, such as the women of Sabokahan.

We Call Her Ina Bai diversifies the Children’s Literature Canon because it manifests the Lumad tradition of collective storytelling in the contemporary format of YA Literature. It is a hybrid literacy practice; as the prologue states: “this story is woven in the beads and fabric of the lives of thousands of Lumad of Pantaron” (p. 10). The text confronts difficult topics, such as the horrific violence brought upon the Manobo and their right for self-defense and self-determination, head on. It is also unequivocal in regarding the role of the oppressors and their motivation: greed. Yet, it also points to another possible, more harmonious, and interdependent world—one which draws on indigenous ways of knowing and being, embodied in the tender relationship of the two protagonists.

The issues raised by *We Call Her Ina Bai* are not just relevant to one small corner of the planet; they have universal resonance. All human beings have a relationship to extractive capitalism, environmental degradation, mass displacement and migration, patriarchy, and legacies of colonialism, whether they recognize this relationship or not. All students and teachers have an obligation to grapple with what it means to teach the truth, especially when the truth is a target of dictators, strongmen, and local politicians. As such, the counter-story set forth by *We Call Her Ina Bai* serves to both redistribute the knowledges, histories, and truths circulating in this community—Lumad women leaders, youth activists, western academics, and many more—transnationally and foster a more transnational allyship toward liberatory, humanizing ends. Such an approach serves to prefigure a community of difference joined together by a desire for a more just world.

Concluding thoughts

The three examples discussed in this chapter illustrate different possibilities for disrupting hegemonic social and intellectual formations within contexts shaped by imperialism and colonial oppression. In all three examples, prefigurative hybrid literacies provide spaces to nurture decolonial imaginings and create/sustain transnational allyships toward a more just world. As Walia (2013) argues, “enacting a politics of decolonization [] necessitates an undoing of the borders between one another” (Decolonizing Movement Borders section, para. 4). In writing this piece, we hope to encourage “nurturing, emancipatory and expansive social relations and identities, forged in and through the course of struggle” (Walia, 2013, Decolonizing Movement Borders section, para. 1). It is through an understanding of the ways in which different communities create counterstories, engage in collective dreamwork, and nurture decolonial imaginaries that we can forge creative, radical, transformative solutions and partnerships toward social justice and equity.

References

- Arroyo-Pizarro, 2018. Thiago y la aventura del huracán. EDP University.
- Birr Moje, E., 2013. Hybrid literacies in a post-hybrid world: making a case for navigating. In: Hall, K., Cremin, T., Comber, B., Moll, L. (Eds.), *International Handbook of Research on Children’s Literacy, Learning and Culture*. Wiley-Blackwell, pp. 359–372.
- Burnett, C.D., Marshall, B., 2001. *Foreign in a Domestic Sense: Puerto Rico, American Expansion, and the Constitution*. Duke University Press.
- Butler, J., 2004. *Undoing Gender*. Routledge. <https://doi.org/10.4324/9780203499627>.

- Çağlar, A.S., 2001. Constraining metaphors and the transnationalisation of spaces in Berlin. *J. Ethnic Migrat. Stud.* 27 (4), 601–613. <https://doi-org.proxy.library.upenn.edu/10.1080/13691830120090403>.
- Campano, G., 2013. Literacy in/for action: prefigurative pedagogies and collective knowledge projects. *J. Literacy Compos. Stud.* 1 (2), 78–85.
- Campano, G., Ghiso, M.P., 2011. Immigrant students as cosmopolitan intellectuals. In: Wolf, S., Coates, P., Enciso, P., Jenkins, C. (Eds.), *Handbook on Research on Children's and Young Adult Literature*. Lawrence Erlbaum, pp. 164–176.
- Campano, G., Nichols, T.P., Player, G., 2020. Multimodal critical inquiry: nurturing decolonial imaginaries. In: Elizabeth, B., Moje, E.B., Afflerbach, P., Enciso, P., Lesaux, N.K. (Eds.), *Handbook of Reading Research*, vol. 5. Routledge, pp. 137–152.
- Castillo, M., 2021. Superheroine latinidad: the diasporic identities of America Chavez & La Borinqueña. *J. Popular Cult.* 54 (5), 1012–1030.
- Duany, J., 2002. *Puerto Rican Nation on the Move: Identities on the Island and in the United States*. The University of North Carolina Press.
- Escalante Gonzalbo, P., 2020. El México antiguo [Ancient Mexico]. In: El Colegio de México, A.C. (Ed.), *Nueva historia mínima de México [New Minimal History of Mexico]*, first Epub Interactive ed. El Colegio de México, A.C., pp. 4–48.
- Foner, N., 1997. What's new about transnationalism? *New York immigrants today and at the turn of the century. Diaspora* 6 (3), 355–375.
- García, I., 2018. Diasporic intersectionality: colonial history and Puerto Rican hero narratives in 21: the story of Roberto Clemente and La Borinqueña. In: Aldama, F. (Ed.), *The Routledge Companion to Gender, Sex and Latin American Culture*. Routledge, pp. 71–82.
- García, J., 2021. Side by Side: US Empire, Puerto Rico, and the Roots of American Youth Literature and Culture. University Press of Mississippi.
- Godreau, I., Bonilla, Y., 2021. Nonsovereign racecraft: how colonialism, debt, and disaster are transforming Puerto Rican racial subjectivities. *Am. Anthropol.* 123 (3), 509–525.
- Gomez-Peña, G., 1996. The new world border: prophesies, poems & loqueras for the end of the century. *City Lights*.
- Gutiérrez, K., 2004. Literacy as laminated activity: Rethinking literacy for English Learners. *National Reading Conference Yearbook*. (101–114). National Reading Conference.
- Gutiérrez, K., 2008. Developing a sociocritical literacy in the third space. *Read. Res. Q.* 43 (2), 148–164. <https://doi.org/10.1598/RRQ.43.2.3>.
- Gutiérrez, K.D., Morales, P.Z., Martínez, D.C., 2009. Re-mediating literacy: culture, difference, and learning for students from nondominant communities. *Rev. Res. Educ.* 33 (1), 212–245. <https://doi.org/https://doi.org/10.3102/0091732X08328267>.
- Haiman, A., 2014. *Tulipán: The Puerto Rican Giraffe*.
- Immerwahr, D., 2019. How to Hide an Empire: A History of the Greater United States. Farrar, Straus and Giroux.
- Jocson, K., 2005/2006. Examining hybrid literacies of critical youth poets. *Int. J. Learning* 18 (2), 49–55.
- Kasir, S., 2018. Precarity. *The Cambridge Encyclopedia of Anthropology*. Retrieved. (Accessed 14 April 2022).
- Lagunas, Z., 2018. Población, migración y mestizaje en México: Época prehispánica-época actual [Population, Migration and Admixing in Mexico: Pre-Hispanic Period-Current Period]. Instituto Nacional de Antropología e Historia.
- Lloréns, H., 2018a. Imaging disaster: Puerto Rico through the eye of Hurricane María. *Transform. Anthropol.* 26 (2), 136–156. <https://doi.org/10.1111/traa.12126>.
- Lloréns, H., 2018b. Ruin nation: in Puerto Rico, Hurricane María laid bare the results of a long-term crisis created by dispossession, migration, and economic predation. *NACLA Rep. Am.* 50 (2), 154–159. <https://doi.org/10.1080/10714839.2018.1479468>.
- Luke, A., 2012. Critical literacy: foundational notes. *Theory Into Pract.* 51 (1), 4–11. <http://www.jstor.org/stable/41331066>.
- Marichal-Lugo, T., 2019. *La Cucarachita Martina Y El Huracán*. Colección Cuento Que Te Cuento.
- Menchú, R., Burgos-Debray, E., 1992. I, Rigoberta Menchú: An Indian Woman in Guatemala. A. Wright, Trans. Verso.
- Mignolo, W., 1995. *The Darker Side of the Renaissance: Literacy, Territoriality, and Colonization*. University of Michigan Press.
- Mignolo, W., Walsh, C.E., 2018. *On Decoloniality: Concepts, Analytics, Praxis*. Duke University Press.
- Morawska, E., 1999. The New-Old Transmigrants, Their Transnational Lives, and Ethnicization: A Comparison of 19th/20th and 20th/21st C Situations. European University Institute Working Papers EUF No. 99/2, Florence.
- Palacios, A., 2012. (Dis)Claiming Mestizofilia: Chicana/os Disarticulating Euromestizaje. Doctoral dissertation, University of California, Berkeley. UC Berkeley Electronic Theses and Dissertations. <https://escholarship.org/uc/item/5jn8937q>.
- Pastrana-Andino, I., 2018. *Mi Isla bella/Mi Isla Hermosa*. Aparicio Distributors.
- Pérez, E., 1999. *The Decolonial Imaginary: Writing Chicanas Into History*. Indiana University Press.
- Portes, A., Guarnizo, L.E., Landolt, P., 1999. The study of transnationalism: pitfalls and promises of an emergent research field. *Ethn. Racial Stud.* 22 (2), 217–237.
- Rexach-Olivencia, L., 2018. Por Ahi Viene El Huracán. Here comes the hurricane. Editorial Destellos.
- Santiago II, J.R., Valle, M., 2021. We Call Her Ina Bai: How Strong Women Are Made (E. Muntimbuwan, Illus.). Sabokahan Unity of Lumad Women.
- Santos, B. de S., 2013. Cuando los excluidos tienen derecho: Justicia indígena, plurinacionalidad e interculturalidad [When the excluded have the right: indigenous justice, plurinationality and interculturality]. J. Aguilló Bonet & J. L. Exeni Rodríguez (Trans.). In: Santos, B. de S., Grijalva Jiménez, A. (Eds.), *Justicia indígena, plurinacionalidad e interculturalidad en Ecuador [Indigenous Justice, Plurinationality and Interculturality in Ecuador]*, second ed. Ediciones Abya Yala, pp. 13–50. <https://ebookcentral-proquest-com.proxy.library.upenn.edu/lib/upenn-ebooks/detail.action?docID=3217275>.
- Santos, B. de S., 2014. *Epistemologies of the South: Justice Against Epistemicide*. Routledge.
- Schechter, P.A., 2012. *Exploring the Decolonial Imaginary: Four Transnational Lives*. New Palgrave Macmillan.
- Singh, N., 2004. *Black is a Country: Race and the Unfinished Struggle for Democracy*. Harvard University Press.
- Torres, F.L., Medina, C.L., 2021. Cuentos Combativos: decolonialities in Puerto Rican books about María. *J. Literacy Res.* 53 (2), 242–264. <https://doi.org/10.1177/1086296X211009294>.
- Turner, J.K., 1969. *Barbarous Mexico*, first digital ed. University of Texas Press <http://hdl.handle.net/2027/heh.02804>.
- Venator-Santiago, C., 2017. Mapping the contours of the history of the extension of U.S. citizenship to Puerto Rico, 1898–present. *Centro J.* 29 (1), 38–55.
- Vertovec, S., 2001. Transnationalism and identity. *J. Ethn. Migr. Stud.* 27 (4), 573–582. <https://doi.org/10.1080/13691830120090386>.
- Walia, H., 2013. *Undoing Border Imperialism*. AK Press.
- Walia, H., 2021. *Border and Rule: Global Migration, Capitalism, and the Rise of Racist Nationalism*. Haymarket Books.
- Walsh, C., 2020. ¿Interculturalidad y (de)colonialidad? Gritos, grietas y siembras desde Abya Yala [Interculturalidad and (de)colonialidad? Cracks, shouts, and sowings from Abya Yala]. In: Romero Losacco, J. (Ed.), *Pensar distinto, pensar de(s)colonial [Distinct thinking, de(s)colonial thinking]*, first digital ed., pp. 139–178 Fundación Editorial El perro y la rana. <http://www.elperroylarana.gob.ve/pensar-distinto-pensar-descolonial/>.
- Walsh, C.E., Mignolo, W., 2018. Introduction. In: Mignolo, W., Walsh, C.E. (Eds.), *On Decoloniality: Concepts, Analytics, Praxis*. Duke University Press, pp. 1–12.
- Woodson, C.G., 1933/1999. *The Mis-education of the Negro*. 12th Media Services.

Language choice, technology and international education

Nathan M. Castillo^a, Daniel A. Wagner^b, Ghaida S. Alrawashdeh^a, and Aanchal Gidra^b, ^a College of Education, University of Illinois at Urbana-Champaign, Champaign, IL, United States; and ^b Graduate School of Education, University of Pennsylvania, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Language choice and education	817
Technology and language choice in education planning	818
A framework for language, technology, and education	819
Language context	819
Purposes	820
End-users	820
Technology	820
Design solution	820
Applying the ICT4E framework in the context of language diversity	820
The expanding role of technology in education	821
Student agency and empowerment	821
Inclusion	821
Content development	821
Adaptability through improved data capture	821
Assessment	822
Conclusion	822
References	822

The UN Sustainable Development Goal #4 (SDG4) calls for all children to complete a full cycle of schooling of sufficient quality to acquire the necessary skills for lifelong learning (United Nations, 2015). Yet, evidence points to pervasive deficiencies in global reading and math skills development among children and adolescents in many low-and middle-income (LMIC) countries. For instance, UNESCO (2017) estimates that over 600 million children and adolescents across the world are not acquiring basic skills in reading and math, despite several years of schooling. Other data illustrates the impact of COVID-19 on further expanding deficiencies in skills development globally. Patrinos et al. (2022) examined evidence from thirty-five international studies of documented learning loss among school-aged children during the first two years of the pandemic, with a particular focus on math and reading skills. Their analysis revealed an average loss equivalent to almost half a year of learning (Patrinos et al., 2022). Deficiencies in skills development and subsequent learning loss prompted by school closures is impacted by variations across many background characteristics. Among such variations, home (or mother-tongue) language has long been understood to play a critical role in students' ability to learn from the educational content they receive in schools (Trudell, 2016; Walter and Benson, 2012).

In this review, historical and international trends around language and policy planning are considered along with trends in the use of technology for mediating and supporting language and education with a particular emphasis on LMICs. The intersection of language choice, technology and education is one that has become increasingly pertinent as countries endeavor to enhance both access to and quality of education around the world.

Language choice and education

How languages are chosen—or “language choice”—for learning in schools, and more generally, is a multidimensional issue, ranging from the personal (such as one's feelings of identity) to the broadest (such as in national and international language policies). At the level of local societies and cultures, language preservation, especially among displaced communities and minority groups, is a current driving force behind supporting mother-tongue language use at home or in school. Such language choice for minority and disadvantaged groups often reflects a community desire for children to communicate, and eventually read and write, in their home language. For example, refugees may wish to maintain their culture and heritage through written (or spoken) language, and this can be a powerful motivator for maintaining mother-tongue literacy and education (Kelcey et al., 2022).

Concerns about the child's home language and the language of instruction (LOI) in schools in official language(s) has been a matter of academic and policy attention over many years. In numerous post-colonial low- and middle-income countries (LMICs) there may be a misalignment of home and school languages, as exemplified by the fact that both teachers and students alike may not be able to fully engage with the curriculum due to their inability to effectively communicate in the dominant or official languages (Ball et al., 2014; Cummins, 1980). As a consequence, school systems in these countries have been reconsidering their LOI policies.

While national policymakers are more intentional today about school language policies, challenges remain to effective multilingual learning in the classroom. One dimension is parent and teacher preference for dominant languages given their increasing importance to economic globalization (Benson, 2004; Wagner, 2018a). Another factor is the lack of educational resources available in mother-tongue languages due in part to low cost-benefit valuations of book publishers (Piper et al., 2016b). A further impediment to mother tongue language use is a result of LOI policies that put colonial languages into place over decades in support of Western expansionist policies. Thus, language choice can and does often include ways to ensure a community's access to expanded economic opportunities, such as by supporting the use of international or dominant language competence, but often at the expense of mother-tongue language instruction.

Another key dimension of language choice policy concerns the use of mother-tongue language (L1) in order to support dominant (L2) bilingualism in classrooms. In multilingual communities, where there is a single dominant international language (e.g., French in Senegal), the use of L1 to support learning in L2 is understood to be a sound policy choice as research indicates that children use their familiar L1 competence to learn an L2 (Benson and Kosonen, 2021). In order to better understand the dimensions of language choice, India and South Africa provide useful historical examples.

In India, language policy and language choice in schooling must take into account cultural, historical, and socio-economic variations along with changes in national educational policy. In present-day India, 96% of the mother-tongues are spoken by only 4% of the population (Narayana Singh et al., 2022). While over two centuries of colonialism strongly introduced English to the country, only an estimated 10% of the population speaks English fluently today. Colonization—both during and after—did not reduce the use of 22 official Indian languages, alongside hundreds of “tribal” and mother tongue languages (Daswani, 2001). In 1964, the Indian government decided to introduce into the school system what has come to be referred to as the “three-language formula”. This earlier policy, endorsed through the 2020 National Education Policy, requires all children to receive schooling in two Indian languages and English during their first ten years of schooling (NEP, 2020). Within this model, a learner's home language should remain the language of instruction through grade 5, but preferably through eight years of schooling where possible (Wagner et al., 2010).

In South Africa, several policies were adopted to address poverty and inequality in an attempt to reduce the impact of the country's oppressive apartheid era. The 1997 Language in Education Policy promotes home language teaching and learning in each of the 11 official languages. Similarly, the 2007 National Framework for Quality Education in Rural Areas emphasized better learning in historically under-resourced schools (SAHRC, 2012). More recently, the National Development Plan 2030 builds upon the country's current development agenda and places quality basic education as a key to achieving prosperity and equity (National Planning Commission, 2012). Preservation of local South African languages has been central to the country's post-apartheid Bill of Rights, which states that students have the right to receive education in an official language or their language of choice (Wu et al., 2021). Despite these collective policy efforts, implementing quality learning in home languages remains a persistent challenge for South African classrooms, and which disproportionately impacts poor, rural communities (Chetty et al., 2017).

Overall, language choice and educational policy depend on a multitude of factors. The overarching ecological environment that enables the use of different languages is an important driving force in determining language choice. In this context, technology becomes one of the important routes through which language choice is realized in school and society.

Technology and language choice in education planning

Information and communications technologies (ICTs) for education is increasingly viewed as a useful support for learning across the globe (World Bank, 2016). When aligned with the broader educational ecosystem, technology is particularly useful in support of learning and literacy in LMICs (Wagner, 2018a). Research suggests that when technology is leveraged to support language contexts, learning can be increased due to faster and automated analyses of learning data, and improved ways to be inclusive of those in need (Piper et al., 2016a; Castillo and Wagner, 2019).

One early ICT-based attempt of learning support for linguistic minorities in LMICs was developed for numeracy instruction in Nicaragua (Searle et al., 1976). This approach, called interactive radio instruction (IRI), was designed to deliver interactive learning to remote areas through audio broadcast with songs, dramatizations, and games (Ho and Thukral, 2009). During a scheduled time, teachers would turn on the radio and guide students through narrated pedagogical activities as a whole class. More recently, synchronous broadcast delivery has been substituted with digital audio files projected through portable speakers that give teachers the opportunity to pause and or repeat certain parts of the audio lesson as needed. Still prominent in many regions today, this approach supports quality learning by combining culturally relevant content through language and local context with interactive instruction. For instance, an early numeracy IRI project implemented in the Cordillera region of Paraguay reinforced key learning concepts in Guaraní, the indigenous language prominent across the region (Naslund-Hadley et al., 2014). Story lines and games for the Paraguay program also included references to local customs and artifacts that celebrated Guaraní heritage. Educational content delivered through interactive audio has served as an effective model for student-centered learning with technology, and has also been pedagogically supportive for teachers with limited access to ongoing training (Damani and Mitchell, 2020).

Closures of schools following the COVID-19 outbreak in 2020 resulted in a global shift toward remote learning and a greater use of ICTs in many countries (OECD, 2020). This trend produced further learning inequality stemming from a digital divide that resulted in unequal access to remote learning opportunities and resources. For instance, nearly all high-income countries provided remote learning opportunities (90%) and mainly online (Vegas, 2020). By contrast, less than twenty-five percent of low-income

countries offered remote learning (Vegas, 2020). While broadcast (either radio or television) is a more viable approach for remote learning in low-income and rural contexts (UNICEF, 2020), the digital divide created unequal access to learning opportunities for millions of students without connectivity.

Whatever the mode of delivery, there is no doubt that the use of technology has been instrumental in providing students with access to resources during the pandemic, and this will likely continue well into the future. Yet, evidence on the effectiveness of technology for education initiatives beyond mere access is mixed (Hinostroza et al., 2014). This finding is due, in part, to the fact that technology use for learning improvement has evolved into a much broader field in recent years. The breadth of design and implementation with technology has created serious debates over definitions and what can be claimed as evidence. For instance, claims relying on observational anecdotes or broad definitions of evidence without any clear indication of impact on learning gains are being replaced by empirical data about how ICTs are utilized to change learning processes and improved skills development (Wagner et al., 2014). One outcome from recent experimentation with technology and learning is clear: mere access alone will not produce consistent and positive differences. The widely cited educational technology intervention, One Laptop per Child (OLPC), underscores this outcome by producing a modest detectable impact on increased computer use and a cognitive measure of verbal fluency, but little or no impact on learning outcomes (Cristia et al., 2012). One explanation of such limited outcomes produced by OLPC is a lack of emphasis on locally relevant content. In order for initiatives to promote quality learning with technology, it is necessary to align their design and implementation with the broader language context.

A framework for language, technology, and education

One useful framework for examining the interaction between language, technology, and education derives from the ICT for education (ICT4E) framework proposed by Wagner et al. (2014). The framework originally comprised four distinct but overlapping components: intervention purposes, intended end users, devices, and context. When this framework is applied, a design solution for effective deployment of technology in educational settings is produced. Fig. 1 illustrates an adaptation of this framework that features the language elements within a specific learning context in conjunction with the other components.

Language context

In this review, we put the emphasis on *language* context, so that aspects of language choice can be understood in depth. By considering the language characteristics of a given context, the other components of the framework can be better applied to support quality learning. Language characteristics may include cultural relevance, breadth of socio-linguistic diversity, attention to and motivation for language learning, as well as broader support for language development within a learning environment. By invoking context, it is important to assure that local stakeholder engagement be taken into account. Typically, information about language characteristics in a given context are provided by local experts, either working with the government or as independent advisers. For example, the Khan Academy utilized support from the state government of Punjab in India to launch their ICT4E platform in the Punjabi language. A select group of teachers localized the content while the state education department set up the necessary infrastructure in school computer labs and provided ongoing teacher support through WhatsApp groups.

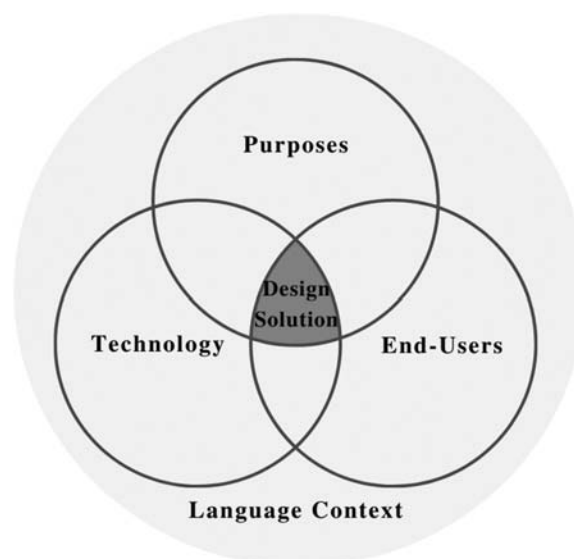


Fig. 1 A language-centered ICT4E framework. Adapted from Wagner et al. (2014) and Wagner (2018a,b).

Purposes

The intended purposes of leveraging technology to support learning quality within specific sociolinguistic groups must be described. For instance, one clear purpose would be to provide learners with a choice to develop foundational skills in reading and writing, usually in the home language (L1), and to support the transfer of those skills into an additional language (L2). Language choice was a driving design factor behind a digital early reading program in rural South Africa, which produced positive effects on fluency and comprehension gains as discussed in further detail in the next section (Castillo and Wagner, 2019). Another purpose might be to build upon language- and culturally-specific content to enhance both engagement and comprehension of learning materials such as with the Paraguay example described previously.

End-users

A second component of the framework considers the intended end-users, which could include learners who might have been linguistically marginalized (for instance, due to lack of content in their local languages), displaced populations (either internally or refugee status), and/or differently-abled (cognitively or physically) students. By providing learning materials that are tailored to individual learner abilities as well as access to open-source content and multi-sensory resources, this group of learners may gain access to knowledge that would otherwise be unavailable to them. For example, the iMlango initiative in Kenya aims to improve the educational outcomes of marginalized girls in math using adaptable learning content developed by Whizz Education. Within this project, learners move through a series of math activities that is continuously adjusted, based on data provided by the learners and derived from the accuracy of their responses (Adam et al., 2022).

End-users may also include teachers and parents. Teachers may be introduced to new curricula and learning material to further improve their pedagogical skills (Wagner, 2017b), or to connect them to other teachers to share resources and best practices. In a similar way, solutions can be designed to empower parents to take a more active role in their child's education. Ultimately, the effectiveness of any design solution hinges on the extent to which they are built with specific end-users in mind.

Technology

By leveraging appropriate technologies, whether hardware or software, along with end-users and clear purposes, new options for learning can be undertaken. For example, technology can range from conventional devices like radios and televisions to mobiles and other wearable devices. Also conceivable are digital tools such as automatic transcription, and speech synthesis. Web 2.0 comprises online tools that encourage and support learner's participation in the creation, uploading and sharing of digital artifacts. Examples could include blogs, podcasts, wikis and media editors (e.g., Conole and Alevizou, 2010). ICT4E enterprise Siyavula Education in South Africa, for instance, took into account the available resources and developed a machine learning-powered platform that provides personalized learning content that can be accessed via feature phones. Siyavula is also sensitive to the affordability concerns of its end-users and partnered with Africa's largest telecommunications provider, MTN, to allow learners all over South Africa to access its platform free of charge (Adam et al., 2022).

Design solution

The three circles in Fig. 1 converge in a way that considers each component of the framework with respect to their independent contributions. Together, these components enable the design of educational initiatives that (ideally) draw on optimal integration of technology within the broader language context. Application of this framework in the design and implementation of technology solutions for language diversity and education help provide a more coherent understanding as to how solutions can be scaled and adapted, as well as why some fail to become sustainable over time.

Applying the ICT4E framework in the context of language diversity

We have outlined some approaches to language planning in different country contexts. Here, building again on the cases of India and South Africa, we illustrate how applying the ICT4E framework enables better support of language and education policies.

In India, the StoryWeaver project, initially designed by Pratham, a non-governmental organization (NGO), represents an ICT4E approach aligned with India's three-language policy. This project aims to fill a gap in the quality of L1 material by providing a digital repository of local language stories for children, who have been associated with alienation from schooling and low literacy rates among rural Indian children (ASER, 2019). StoryWeaver's online repository consists of over 45,000 openly-licensed stories in more than 300 languages through national and international collaborations. Due to connectivity limitations, StoryWeaver allows users to download and read stories offline. In 2016, Pratham led a pilot to deliver content in English and Hindi through an audio call back service via phones. In four days, 40,000 calls were placed to 5000 children in the state of Delhi (Butcher et al., 2019). Multilingual content for StoryWeaver was developed by Pratham Books publications, and by users themselves who can upload content.

A second case example is the Bridges to the Future Initiative (BFI), a digital literacy platform created in collaboration with South African NGO Molteno Institute for Language and Literacy. BFI was developed in response to South Africa's Right to Education Policy to promote choice among socio-linguistic minority children. BFI investigated whether access to guided and culturally relevant digital reading content in three home languages could improve early grade reading among primary school children in rural Limpopo, South Africa (Castillo and Wagner, 2019). An important component of this intervention was the development of a multi-media program for early reading in collaboration with local experts to ensure linguistic accuracy and cultural relevance. BFI

offered sequential learning material in multiple L1 languages, providing learners the ability to choose their language of instruction and to engage with the interactive reading material at their own pace. Another underlying factor aligned with the framework was the development of software that could accommodate existing digital resources in participating schools rather than introducing new technology. Findings were associated with nearly three-quarters of a year of additional growth in oral reading fluency, as well as a two-fold increase in reading comprehension among the children who received the ICT-based learning material (Castillo and Wagner, 2019).

By applying the ICT4E framework to these two examples, we can see how a particular purpose (quality content in L1 for India and language choice within multilingual classrooms for South Africa) tailored for a specific end-user (early grade learners) and organized around existing digital resources (such as a free call back service linked to household phones or computer labs already in schools) lead to better learning results. Further, each of the framework components were informed by the overarching language context that included the use of local dialects, motivation for developing the solution (building out a digital book repository or improving early grade reading), and degree of support for the intervention (international collaboration with community and ministry buy-in).

The expanding role of technology in education

Schools can use technology as part of a strategic approach to better engage and support student learning. Here we examine how the expanding use of appropriate technology can help improve the process of learning and teaching, as well as create an inclusive language environment for students and schools.

Student agency and empowerment

The appropriate use of technology creates opportunities for improved student agency. Instead of being passive recipients of instruction, learners in open systems shape their own experiences according to individual goals and interests. This can be achieved by tracking progress along stated goals, and by engaging with the quality of their own work through guided reflection. For instance, many multilingual students struggle to participate on equal terms in the classroom as they confront the dual challenge of mastering grade-level content while learning to learn in L2 (Piper et al., 2016b). Language learners may find it helpful if tasks are broken down into manageable steps such as where hints are provided when needed in their L1 along with the option of rephrasing or translating into L2. In Lebanon, the ICT4E enterprise Augmental offers students the choice to rephrase questions if they are having trouble understanding them. By offering flexible options, such systems give students increased control over key aspects of their educational progress, leading to greater self-efficacy and engagement in their learning (Jääskelä et al., 2021).

Inclusion

Work on the interaction between language and technology sometimes overlooks the diverse needs of differently-abled students. Applying inclusive learning principles like the Universal Design for Learning (UDL) can help to ensure that all learners, regardless of their individual abilities, have the opportunity to use technology for learning (King-Sears, 2009). Embedded assistance features like text-to-speech and speech-to-text, can make learning more accessible to a broader group of learners in multilingual settings (Wood et al., 2018). The game-based learning website Mangahigh has created a text-to-speech engine to convert mathematical notation into readable text that is spoken by the user's browser (Adam et al., 2022).

Another inclusive design practice is access to content in local sign languages and braille for students with hearing impairments or poor vision. For instance, the Ministry of Education in Kenya provided TV and virtual broadcast of stories in Kenyan Sign Language as part of their remote learning policy during school closures due to the COVID-19 pandemic (All Children Reading, 2020).

Content development

When the COVID-19 pandemic led to abrupt school closures, schools faced challenges in providing specialized content and resources for multilingual learners (U.S. Department of Education, 2021). A recent review explored initiatives across LMICs that are incorporating digital personalized learning (DPL) approaches to assist with such challenges (Adam et al., 2022). One central feature is the ability to create learning content alongside local counterparts. This approach improves the ability to generate material that is aligned with curricular standards in multiple languages while ensuring cultural relevance within specific learning contexts. For instance, the ICT4E firm Angaza Elimu engages both local teachers and government in content creation. Learning content for the DPL platform was created by practicing teachers with guidance by a content evaluator from the Kenya Institute for Curriculum Development (KICD). Angaza Elimu program developers then followed the national policy on curriculum design to align the content within the platform.

Adaptability through improved data capture

Technology has made possible the seamless collection and sharing of student data (Wagner, 2018a,b). By combining students' learning preferences, background information (including their home language), and their prior knowledge, content sequences

can be tailored so that they build on and reinforce skills students have previously mastered. Further, the use of digital learning systems can combine predictive analytics with data visualization features that enable a shift in the teacher's role from instructor to learning facilitator. For instance, dedicated dashboards can provide teachers with real-time insight into how their students are performing and allow teachers to identify their strengths and learning gaps. With this information, they can make informed decisions and determine the best course of action to support their students. Pedagogical guidance can also be shared with other stakeholders involved in students' learning, including administrators and parents.

Some systems that deploy AI-based models come with inherent risk that may reproduce inequalities such as underlying biases based on language, gender or ethnicity (Baker and Hawn, 2021). Biases like these are not always apparent in the educational field (Anderson et al., 2019). Stakeholders building AI models thus need to examine the data for evidence of bias. Recruiting a diverse group of experts that are trained to analyze the algorithm output, exposing the algorithm to a more even-handed distribution of examples, can mitigate some biases (Adam et al., 2022). One such initiative is UNICEF's Policy guidance on AI for children (Virginia Dignum et al., 2021), which provides recommendations for building government and private sector policies and systems that uphold children's rights.

Assessment

Technology can also facilitate deployment of hybrid assessments that combine localization with speed and transparency of collected data, at lower stakes (Wagner, 2011). For instance, *Tangerine* is an open-source data collection software that provides insights about student learning by digitizing the Early Grade Reading and Mathematics Assessments (RTI International, 2018). A key benefit of this app is its use of automated data validation features to reduce missing values and other data entry errors common in paper-based assessments. Experimentation in Kenya within a broader digital literacy program was associated with improved monitoring and support to teachers through sustained assessment of learners in early grades (Olubendi, 2019; Strigel, 2014); the app has been deployed in over sixty countries and adapted in more than one hundred languages worldwide.

Conclusion

Multilingualism in education is evolving rapidly as a function of increased globalization, internal and external migration, climate change and more (Wagner, 2018a,b). Concurrently, the rise of new technologies with enabled digital design has led to an increased capability of integrated technology use for improving education through language choice.

Even so, lessons drawn from earlier technology interventions should guide future designs. For instance, early efforts emphasized devices and often overlooked local needs and language context resulting in limited empirical impact (Hinostroza et al., 2014). Despite mixed results in earlier efforts, a growing number of initiatives related to the SDGs emphasize the use of technology as a key aspect of future development efforts (Vinueza et al., 2020). With this growing interest, deployment of digital tools for international education development must center on local context, including language choice.

Further, more emphasis should be placed on leveraging resources to affect change, while disseminating evidence more broadly as a global community. For instance, UNICEF's Reimagine Education Initiative <https://www.unicef.org/reimagine/education> sees digital learning as a component of essential services for all children. The program calls for stronger public-private partnerships to expand connectivity, lower the cost of digital services, and create opportunities for young people to play a role. Similarly, the EdTech Hub (Hennessy et al., 2021) is a global research partnership that aims to disseminate evidence about technology in education to promote more learning across LMICs. While still early in their development, these initiatives, and others like them, have the potential to make a lasting impact.

In this review, we provide a conceptual and empirical basis for the use of technology that supports learning in contexts where the language of instruction is in policy contention in LMICs—usually in spaces where educational content in local languages and dialects has been limited. All the same, current efforts to produce digital solutions that support students learning in multilingual contexts, while currently limited in scope, is gaining in attention. In the coming years, technology will serve an expanding role in improving the processes of learning and teaching while promoting more inclusive and supportive language environments for learning. Such efforts will likely constitute an important contribution to achieving the UN Sustainable Development Goal for education.

References

- Adam, T., Alam, A., Alrawashdeh, G., Castillo, N.M., El-Serafy, Y., Sivanandan, M.S., Tiwari, P., 2022. Trends in Digital Personalized Learning in Low and Middle-Income Countries. Office of Global Insight and Policy. UNICEF, New York.
- All Children Reading, 2020. eKitabu's Signed Digital Story Time Included in Kenya Ministry of Education's Distance Learning Program. All Children Reading: A Grand Challenge for Development. <https://allchildrenreading.org/news/ekitabus-digital-story-time-brings-hope-to-out-of-school-children-across-kenya-2/>.
- Anderson, H., Boodhwani, A., Baker, R.S., 2019. Assessing the fairness of graduation predictions. Proceedings of the 12th International Conference on Educational Data Mining, pp. 488–491.
- ASER, 2019. Annual Status of Education Report (Rural) 2018. ASER Centre, p. 336. <https://learningportal.iiep.unesco.org/en/library/annual-status-of-education-report-rural-2018>.
- Baker, R.S., Hawn, A., 2021. Algorithmic Bias in Education. Preprint. EdArXiv. <https://doi.org/10.35542/osf.io/pbmzv>.

- Ball, J., Paris, S.G., Govinda, R., 2014. Literacy and numeracy skills among children in developing countries. In: Wagner, D.A. (Ed.), *Learning and Education in Developing Countries: Research and Policy for the Post-2015 UN Development Goals*. Scopus, pp. 26–41. <https://doi.org/10.1057/9781137455970>.
- Benson, C., 2004. Do we expect too much of bilingual teachers? Bilingual teaching in developing countries. *Int. J. Biling. Educ. Biling.* 7 (2–3), 204–221.
- Benson, C., Kosonen, K., 2021. Language Issues in Comparative Education II: Policy and Practice in Multilingual Education Based on Non-dominant Languages. Brill, Leiden.
- Butcher, N., Levey, L., von Gogh, K., 2019. Good Stories Don't Grow on Trees: A Guide to Effective Costing of Storybooks in the Global South | NBA. Early Literacy Resource Network. <https://www.earlyliteracynetwork.org/content/good-stories-dont-grow-trees-guide-effective-costing-storybooks-global-south>.
- Castillo, N.M., Wagner, D.A., 2019. Early-grade reading support in rural South Africa: a language-centred technology approach. *Int. Rev. Educ.* 65 (3), 389–408. <https://doi.org/10.1007/s11159-019-09779-0>.
- Chetty, M., Moloi, M.Q., Poliah, R.R., Tshikoro, J., 2017. The SACMEQ IV Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education. Department of Basic Education, p. 56 (Republic of South Africa).
- Conole, G., Alevizou, P., 2010. A Literature Review of the Use of Web 2.0 Tools in Higher Education—Open Research Online. HEA Academy. <http://oro.open.ac.uk/23154/>.
- Cristia, J.P., Ibararán, P., Cueto, S., Santiago, A., Severin, E., 2012. Technology and Child Development: Evidence from the One Laptop Per Child Program. IDB-WP-304; IDB Workin Paper Series. Interamerican Development Bank. <https://publications.iadb.org/publications/english/document/Technology-and-Child-Development-Evidence-from-the-One-Laptop-per-Child-Program.pdf>.
- Cummins, J., 1980. The cross-lingual dimensions of language proficiency: implications for bilingual education and the optimal age issue. *Tesol Q.* 175–187.
- Damani, K., Mitchell, J., 2020. Radio: A Rapid Evidence Review. EdTech Hub. <https://doi.org/10.5281/ZENODO.4556932>.
- Daswani, C.J., 2001. Language Education in Multilingual India. UNESCO, Delhi.
- Hennessy, S., Jordan, K., Wagner, D.A., 2021. Problem Analysis and focus of EdTech Hub's Work Technology in Education in Low- and Middle-Income Countries. Working Paper. London: UKaid/World Bank. Pps.1–73. https://www.repository.cam.ac.uk/bitstream/handle/1810/319313/Problem20Analysis_Technology20in20Education_Revised20Final_1.pdf?sequence=1
- Hinojosa, J.E., Isaacs, S., Bougroum, M., 2014. Information and communications technologies for improving learning opportunities and outcomes in developing countries. In: Wagner, D.A. (Ed.), *Learning and Education in Developing Countries: Research and Policy for the Post-2015 UN Development Goals*. Palgrave Macmillan US, pp. 42–57. https://doi.org/10.1057/9781137455970_3.
- Ho, J., Thukral, H., 2009. Tuned into Success: Assessing the Impact of Interactive Radio Instruction for the Hardest to Reach. Education Development Center. Chrome-extension. https://pdf.usaid.gov/pdf_docs/PCAAB855.pdf.
- Jääskelä, P., Heilala, V., Kärkkäinen, T., Häkkinen, P., 2021. Student agency analytics: learning analytics as a tool for analysing student agency in higher education. *Behav. Inf. Technol.* 40 (8), 790–808. <https://doi.org/10.1080/0144929X.2020.1725130>.
- Kelcey, J., Guven, O., Burde, D., 2022. Education on the move: how migration affects learning outcomes. In: Wagner, D.A., Castillo, N.M., Grant Lewis, S. (Eds.), *Learning, Marginalization, and Improving the Quality of Education in Low-Income Countries*. Open Book Publishers, Cambridge, UK. <https://www.openbookpublishers.com/product/1423>.
- King-Sears, M., 2009. Universal design for learning: technology and pedagogy. *Learn. Disabil. Q.* 32 (4), 199–201. <https://doi.org/10.2307/27740372>.
- Narayana Singh, U., Singh, R., Banerjee, P., 2022. India: learning challenges for the marginalised. In: Wagner, D.A., Castillo, N.M., Grant Lewis, S. (Eds.), *Learning, Marginalization and Improving the Quality of Education in Low-Income Countries*. Open Book Publishers, Cambridge, UK, pp. 293–308. <https://www.openbookpublishers.com/product/1423>.
- Naslund-Hadley, E., Parker, S.W., Hernandez-Agramonte, J.M., 2014. Fostering early math comprehension: experimental evidence from Paraguay. *Global Educ. Rev.* 1 (4), Article 4. <https://ger.mercy.edu/index.php/ger/article/view/53>.
- NEP, 2020. Policy Document Released by Government of India. Retrieved from. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English.pdf.
- National Planning Commission, 2012. Our Future—Make it Work: National Development Plan, 2030. National Planning Commission. https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-work.pdf.
- OECD, 2020. Keeping the Internet up and Running in Times of Crisis (OECD Policy Responses to Coronavirus (COVID-19)) [Policy]. <https://www.oecd.org/coronavirus/policy-responses/keeping-the-internet-up-and-running-in-times-of-crisis-4017c4c9/>.
- Olubendi, D., 2019. A Framework for Evaluating the Outcome of Use of ICT on Early Grade Literacy Assessment; A Case Study of Tusome Literacy Programme in Kenya. Thesis. University of Nairobi. <http://erepository.uonbi.ac.ke/handle/11295/109880>.
- Patrinós, H.A., Vegas, E., Carter-Rau, R., 2022. An Analysis of COVID-19 Student Learning Loss. No. 10033; Policy Research Working Paper. World Bank Group. Chrome-extension. <https://openknowledge.worldbank.org/bitstream/handle/10986/37400/IDU00f3f0ca808cde0497e0b88c01fa07f15bef0.pdf?sequence=1&isAllowed=y>.
- Piper, B., Zuilkowski, S.S., Kwayumba, D., Strigel, C., 2016a. Does technology improve reading outcomes? Comparing the effectiveness and cost-effectiveness of ICT interventions for early grade reading in Kenya. *Int. J. Educ. Dev.* 49, 204–214. <https://doi.org/10.1016/j.jedudev.2016.03.006>.
- Piper, B., Zuilkowski, S.S., Ong'ele, S., 2016b. Implementing mother tongue instruction in the real world: results from a medium-scale randomized controlled trial in Kenya. *Comp. Educ. Rev.* 60 (4), 776–807. <https://doi.org/10.1086/688493>.
- RTI International, 2018. About Tangerine. <https://www.tangerinecentral.org/about>.
- SAHRC, 2012. Charter of Children's Basic Education Rights. South African Human Rights Commission, p. 31. Chrome-extension. <https://www.sahrc.org.za/>.
- Searle, B., Friend, J., Suppes, P., 1976. The Radio Mathematics Project, Nicaragua, 1974–1975. USAID and the Institute for Mathematical Studies in the Social Sciences, Stanford University. https://catalog.library.vanderbilt.edu/discovery/fulldisplay/alma991009571989703276/01VAN_INST:vanui.
- Strigel, C., 2014. Evidence-based design and implementation of a mobile curriculum-based measurement system for low-resource settings. In: Kommers, P., Isaias, P. (Eds.), 12th International Conference e-Society. IADIS Press, p. 8.
- Trudell, B., 2016. Language choice and education quality in Eastern and Southern Africa: a review. *Comp. Educ.* 52 (3), 281–293.
- UNESCO, 2017. More Than One-Half of Children and Adolescents Are Not Learning Worldwide. UNESCO-UIS. Fact Sheet No. 46. <http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf>.
- UNICEF, 2020. COVID-19 and School Closures: Are Children Able to Continue Learning. UNICEF. <https://data.unicef.org/resources/remote-learning-reachability-factsheet/>.
- United Nations, 2015. Transforming Our World: The 2030 Agenda for Sustainable Development. Resolution Adopted by the General Assembly. United Nations. http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.
- U.S. Department of Education, 2021. Education in a Pandemic: The Disparate Impacts of COVID-19 on America's Students. Department of Education—Office for Civil Rights. <https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf>.
- Vegas, E., 2020. School Closures, Government Responses, and Learning Inequality Around the World during COVID-19. Brookings-CJUE. <https://www.brookings.edu/research/school-closures-government-responses-and-learning-inequality-around-the-world-during-covid-19/>.
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M., Fuso Nerini, F., 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* 11 (1), 233. <https://doi.org/10.1038/s41467-019-14108-y>.
- Virginia Dignum, M., Penagos, K.P., Vosloo, 2021. Policy Guidance on AI for Children: Recommendations for Building AI Policies and Systems that Uphold Child Rights. UNICEF-OGIP. <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>.
- Wagner, D.A., 2011. Smaller, Quicker, Cheaper: Improving Learning Assessments for Developing Countries. UNESCO-IIEP, p. 191. <https://unesdoc.unesco.org/ark:/48223/pf0000213663/PDF/213663eng.pdf.multi>.
- Wagner, D.A., 2018a. *Learning as Development: Rethinking International Education in a Changing World*, first ed. Routledge.
- Wagner, D.A., 2018b. Technology for education in low-income countries: supporting the UN sustainable development goals. In: Lubin, I.A. (Ed.), *ICT-Supported Innovations in Small Countries and Developing Regions: Perspectives and Recommendations for International Education*. Springer International Publishing, pp. 51–74. https://doi.org/10.1007/978-3-319-67657-9_3.

- Wagner, D.A., Castillo, N.M., Murphy, K.M., Crofton, M., Zahra, F.T., 2014. Mobiles for literacy in developing countries: an effectiveness framework. *Prospects* 44 (1), 119–132. <https://doi.org/10.1007/s11125-014-9298-x>.
- Wagner, D.A., Daswani, C.J., Karnati, R., 2010. Technology and mother-tongue literacy in southern India: impact studies among young children and out-of-school youth. *Inf. Technol. Int. Dev.* 6 (4), 23–43.
- Wood, S.G., Moxley, J.H., Tighe, E.L., Wagner, R.K., 2018. Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *J. Learn. Disabil.* 51 (1), 73–84. <https://doi.org/10.1177/0022219416688170>.
- Walter, S., Benson, C., 2012. Language policy and medium of instruction in formal education. In: Spolsky, B. (Ed.), *The Cambridge Handbook of Language Policy*. Cambridge University Press, Cambridge, pp. 278–300.
- World Bank, 2016. *World Development Report 2016: Digital Dividends*. World Bank.
- Wu, A.P., Gonsalves, S.-L.R., Wagner, D.A., 2021. The impact of technology on a multilingual world: problems and opportunities. In: Lubin, I.A. (Ed.), *ICT and International Learning Ecologies*, first ed. Routledge, p. 19.

On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT

Ryosuke Aoyama*, Takeshi Kajigaya, Yuya Takeda, Ryuko Kubota, and Ryan Deschambault, Department of Language and Literacy Education, University of British Columbia, Vancouver, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	825
World Englishes and its neighboring/overlapping paradigms	825
WE and pedagogy	826
Stumbling blocks	827
Ideological domain	827
Sociopolitical and structural domain	827
Practical domain	828
Scholarly domain	829
Stepping stones toward the changing paradigm	829
Awareness-raising	830
(Un)learning language ideology	830
Writing	831
Communication strategies	831
Responding to the critique of WE: pragmatic and critical utilization of boundaries	832
Conclusion	833
References	834

Introduction

In the last quarter of the 20th century, sociolinguists and scholars of English language teaching (ELT) began to pay increased attention to diverse varieties of English used in various parts of the world. This burgeoning interest led to the 1981 inauguration of the international journal *World Englishes*, co-founded by Braj B. Kachru and Larry E. Smith, and later the 1992 creation of the International Association for World Englishes (Bolton and Davis, 2018). World Englishes (WE) as a paradigm has destabilized and complexified the notion of English, recognized equal statuses among its different varieties, and questioned the hegemony enjoyed by so-called native speakers of English and English spoken by them (see Bolton, 2020; Kachru, 2013 for an overview of the development of WE).

Even though WE provoked a major paradigm shift in the conceptualization of English, it is facing dual challenges today. On the one hand, despite the large accumulation of research that has shed light on diverse ways people use Englishes and made practical suggestions that aimed to transform traditional ELT, the dominance of Inner Circle standardized English persists in many classrooms around the world. On the other hand, because of its inherently modernist conception of language based on geographical boundaries, WE has been critiqued by scholars who conceive language in a more fluid, mobile, and entangled way. In other words, WE seems to be in a position where its transformative project has not yet unleashed its potential while it is already treated by some as an outdated model.

Given this background, this chapter explores where WE is situated in today's ELT and re-examines its pedagogical utility. We begin by providing a brief overview of the theoretical development of WE and highlighting its relations to neighboring/overlapping disciplinary fields. This is followed by a summary of WE's pedagogical foci, after which we identify some of the major stumbling blocks that impede actualization of WE-informed pedagogy. Next, we highlight select classroom innovations that challenge those stumbling blocks while contextually applying WE elements to instruction. We then respond to a critique of WE while discussing a pragmatic and critical approach to WE that puts its theoretical limitations to use for pedagogical purposes.

World Englishes and its neighboring/overlapping paradigms

Since the 1980s, the diverse and complex sociolinguistic realities of English and its users in a postcolonial and globalizing world have attracted scholarly attention as signified with WE. The impetus of such exploration dates to when WE reconceptualized English use and users in international contexts, challenging the monolithic conception of English restricted to British and American

*Corresponding author

standardized varieties. One of the influential conceptual models was offered by Kachru (1985), which visually presented the global diffusion of English, drawing boundaries between three concentric circles: the Inner Circle, where English is the primary language (the USA, the UK, Canada, Australia, and New Zealand), the Outer Circle, where intranational and institutional uses of English were the result of colonization by Britain and the United States (e.g., Nigeria, India, and the Philippines), and the Expanding Circle, in which English is learned as a foreign language and mainly used for international communication (e.g., Brazil, China, and Russia).

This boundary-making resulted in the pluralization of English to Englishes and was an inevitable response to the increased complexity surrounding the roles and users of English in the world in the 1980s and 1990s. It was also an attempt to claim legitimacy and recognition of English varieties that had been traditionally marginalized. As Kachru (1992) writes, "The result of [English's global] spread is that, formally and functionally, English now has multicultural identities. The term 'English' does not capture this sociolinguistic reality; the term 'Englishes' does" (p. 357). Kachru's early articulation of Englishes became the seed for a paradigm shift. This is part of a broader politics of recognition since "claims [to recognition] often take the form of calling attention to, if not performatively creating, the putative specificity of some group, and then of affirming the value of that specificity" (Fraser, 1995, p. 74). In discussions of English, the traditional classification of English varieties predicated on native and non-native speaker dichotomy such as English as native language (ENL), English as a second language (ESL), and English as a foreign language (EFL) has thus been replaced, or at least complemented, by WE with inclusive plurality of language varieties at its core (Bolton, 2020).

As a paradigm, WE not only pluralized the English language by delineating provisional boundaries, but it also facilitated pluralization of disciplinary space in the study of English—paving the way for the development of new conceptual approaches or paradigms. English as an International Language (EIL) is one such example; Sharifian (2009) explains that EIL is "a paradigm for thinking, research and practice" in TESOL, SLA, and the applied linguistics of English, whose themes are centered on "its recognition of World Englishes, regardless of which 'circle' they belong to" (p. 2). EIL stresses diverse use of English from Kachru's three circles as international and intercultural communication, while WE studies are often associated with explorations and re-evaluations of Englishes primarily in Outer Circle contexts (Hino, 2018; Saraceni, 2015).

Another WE-related paradigm is English as a Lingua Franca (ELF), although ELF is often discussed in parallel to WE as an independent entity to represent a different theoretical approach to English varieties. ELF, defined as "any use of English among speakers of different first languages for whom English is the communicative medium of choice, and often the only option" (Seidlhofer, 2011, p. 7), explores the function and validity of different Englishes in international settings with its primary focus on Expanding Circle users of English. Jenkins (2018) explains that ELF was initially deemed "an ideological and linguistic extension of WE," (p. 12) and now as a complementary (rather than competing) paradigm, acknowledging Kachru's (1992) work as the key influence on ELF.

Critical work in sociolinguistics and applied linguistics concerned with translingual practice (Canagarajah, 2013)—such as linguistic hybridity (e.g., Otsuji and Pennycook, 2013) and translanguaging (e.g., Garcia, 2009)—also furthers inquiries across scholarly circles, challenging the notion of boundaries mobilized by the WE paradigm to pluralize English. Otsuji and Pennycook (2010) claim that "there need to be alternative ways of understanding diversity other than pluralisation" (p. 251) and argue for an approach that centers the complexity of local language practices and eschews any naïve associations between a language user and specific places, sites, backgrounds, or identities.

Responding to such diversification of disciplinary fields associated with multi/plural turn of sociolinguistics and applied linguistics research (see Kubota, 2015, 2016), there has been an effort to integrate WE, EIL, ELF, and the translingual paradigms into a comprehensive umbrella term called Global Englishes (GE). Rose and Galloway (2019) define GE as "an inclusive paradigm that embraces a broad spectrum of interrelated research that has come before it and emerged alongside it" (p. 6). Indeed, the emergence of this umbrella paradigm clearly shows the proliferation of WE-related/informed paradigms. New paradigms are often the result of critique, and WE is no exception. The three-circle model is limited in the ways it can account for the range and complexity of language practices in today's world. However, WE has long served as a catalyst for dialectic explorations, across scholarly circles, of linguistic diversity, resisting the power of monolithic linguistic normativity. As Bolton (2020) observes, this openness and flexibility to scholarly inquiries, new ideas, and innovation is the foundation of WE.

WE and pedagogy

The WE paradigm opened up disciplinary explorations pertaining to diverse use and users of English, with rigorous scholarly activities signifying postcolonial resistance to Inner Circle linguistic and cultural norms. The impact and influence of the WE paradigm on academia, however, was akin to a ripple effect: when the "stone" was thrown into the water, the primary target that was aimed for was English language pedagogy. Indeed, ELT "has always been a very central—and, possibly, the central—preoccupation" (Saraceni, 2015, p. 8) in WE studies. Teaching of WE can primarily be centered on responding to and addressing fallacies about use and users of English in the world context. These fallacies include the notions that: outside the Inner Circle, English is learned to communicate with native speakers of the language; English is learned to understand American or British cultural values; the goal of English learning and teaching is predicated on the native-speaker model which thus reduces non-native varieties of English to interlanguage as lacking or decaying; and native speakers play significant roles in the teaching of English (Kachru, 1992). Calling for this attitudinal and methodological paradigm shift in the classroom, Kachru (1992) proposed the following six foci for teaching advanced students and training professionals (pp. 360–361, emphases original):

1. Sociolinguistic profile: an overview of English in its world context with discussion of selected major varieties, their users and uses. A clear distinction to be made between the use of English in a monolingual society, as opposed to a multilingual society; and its implications (e.g., mixing, switching).
2. Variety exposure: an exposition of the repertoire of major varieties of English, native and non-native: their uses and users, specific texts related to various interactional contexts, shared and non-shared features at different linguistic levels.
3. Attitudinal neutrality: for teaching purposes, one might focus on one specific variety and at the same time emphasize *awareness* and *functional validity* of other varieties.
4. Range of uses: the functional appropriateness of the lectal range of varieties within a specific a specific variety (e.g., from educated varieties to the pidgins and basilects).
5. Contrastive pragmatics: the relationships of discoursal and stylistic innovations and their relationships to the local conventions of culture (e.g., strategies used for persuasion, phatic communion, apologies, condolences, and regrets).
6. Multidimensionality of functions: the linguistic implications of the functional range, as in, for example, the media, in literary creativity, in administration, in the legal system.

It goes without saying that this was not intended as a comprehensive list for teaching WE in the classroom; it needs to be contextually sensitive and relevant to local learning and teaching communities of practice, such as goals and objectives, students' learning and communicative needs, students' and teachers' sociolinguistic backgrounds which may or may not be shared, and so forth. Indeed, teaching from a WE perspective in ELT is challenging not only because of the conceptual and methodological innovation required (Kachru, 1992), but also because it presents such a fundamental challenge to established attitudes toward "standards" for the teaching and learning of English—despite its relevance for the changing realities of English use and users. The next section discusses existing "stumbling blocks" which stand in the way of pedagogical application of WE.

Stumbling blocks

As was already discussed, WE has the potential to transform the traditional, Inner Circle and native speaker-based ELT to a more inclusive field by emphasizing diversity among Englishes for learners. However, since such transformation requires several fundamental changes in various domains in ELT and society at large (Matsuda, 2012a; Matsuda and Friedrich, 2011; McKay, 2002), it inevitably meets challenges. In this section, we summarize these challenges as ideological, sociopolitical/structural, practical, and scholarly "stumbling blocks" that impede realization of WE-informed pedagogy. It is important to note that these four domains are not meant to be discrete and exhaustive. Rather, we use them as a workable framework to clarify the existing issues for researchers, practitioners, and policymakers to work on.

Ideological domain

Various ideologies related to language, especially native-speakerism and normativism, are extremely powerful stumbling blocks in many ways (Holliday, 2006, 2015; Kubota, 2018b). For example, native-speakerism — the belief that "native-speaker" teachers represent a 'Western culture' from which spring the ideals both of the English language and of English language teaching methodology" (Holliday, 2006, p. 385) — penetrates classrooms, employment policies, national- and supra-national education policies as well as ELT scholarship (Cook, 2007; Davidson, 2006; Hashimoto, 2013; Holliday, 2006; Kubota, 2018a; Kumaravadivelu, 2016; Phillipson, 1992). Here, "native English speakers" are often ideologically and discursively constructed as white speakers from Inner Circle communities (Kang and Rubin, 2009; Kubota, 2018b, 2021). Native-speakerism thus predominantly gives white native speakers from Inner Circle communities and Englishes spoken by them symbolic status (e.g., being recognized "authentic" or "intelligent") as well as material benefits (e.g., higher salary or chance of employment) both within and outside the English-speaking West.

In relation to realization of WE-informed pedagogy, native-speakerism leads to general assumption that the goal of learning English is to be like or to communicate with native speakers (Berns, 2006; Cook, 2007), learners' and teachers' preference for Inner Circle varieties of English over other Englishes (Matsuda, 2003; McKenzie, 2010), or teachers' association of standard English with Inner Circle standardized varieties of English, even when they think teaching different varieties of Englishes is desirable (Mimatsu, 2011). These phenomena denigrate the symbolic status of the Outer Circle and Expanding Circle Englishes, and thus are an impediment for those varieties to be accepted by teachers, learners, policy makers, and publishers.

In addition to native-speakerism, normativism, which diminishes English that is deviant from its "conventional" or "standard" form, stands as a significant stumbling block. This is especially prevalent in writing, both in classroom assignments and high-stakes academic writing such as publishing or grant proposals. Due to the strong normativism, learners often receive lower grades in writing classes and scholars lose chance to get their work published for their "incorrect" use of non-standard form of English, forcing both learners and scholars to sustain the hegemony of standardized English (Davila, 2016; Heng Hartse and Kubota, 2014; Matsuda and Matsuda, 2010; O'Regan, 2021; also see "Practical domain" and "Scholarly domain" below).

Sociopolitical and structural domain

A number of stumbling blocks elaborated in this section are sustained and reinforced due to sociopolitical and structural factors, especially neoliberal political economy, teacher education, assessment practice, and political nature of educational policy.

Current language education policy and practices as well as scholarly activities cannot ignore the influence of neoliberalism (Ricento, 2015), which is an economic and political theory that presumes that human well-being can be best advanced through individual entrepreneurial freedoms and deregulation of markets (Harvey, 2005).

In terms of WE-informed pedagogy, the neoliberal political economy functions as a stumbling block by sustaining the symbolic value of standardized English at least in three ways. First, as neoliberalism endorses borderless economy and efficiency, it valorizes English as “the” international language for global communication for business and strengthens the belief that English proficiency brings economic return (Kubota, 2011, 2018b; Ricento, 2015). In turn, these perceived values of English influence (supra-) national policies and educational practices, promoting English language education worldwide (McNamara, 2011; Sayer, 2015). Although neoliberalism does not specifically endorse “standard” English in theory—rather, it endorses the diversity of people and cultures as a catalyst for strengthening economic competitiveness (Kubota, 2016)—it inevitably sustains the hegemony of standard English because “English” in ELT practices and policies still predominantly means “standard English,” as recent studies continuously suggest (e.g., Boonsuk et al., 2021; Cushing, 2021a,b; Galloway, 2017; Rose and Galloway, 2017). Second, as neoliberalism holds scholars and institutions where they work accountable and competitive, it makes scholars complicit with sustaining the hegemony of standardized English because it is the de facto universal medium of writing and presenting in academia (see “Scholarly domain” below). Third, neoliberalism also forces people to acquire a bundle of skills to keep their employability and competitiveness (Kubota and Takeda, 2021; Urciuoli, 2008). This accelerates the need for large-scale standardized assessments like TOEFL or IELTS, whose benchmarks are largely Inner Circle-based. The issue of assessment is elaborated further below.

Teacher education plays a vital role to transform education because it can influence teachers, each of whom will then influence a number of students and potentially local school curricula (Matsuda, 2017). However, many teacher education programs are still Inner Circle- and standardized English-based (Bayyurt and Sifakis, 2017; Biric Deniz et al., 2020). Transforming teacher education is indeed a daunting task and meets a number of stumbling blocks as it is deeply embedded in the existing social structure and connected to people’s beliefs about English. For example, there is not enough guidance for teacher educators despite recent practical works that support WE-informed pedagogy (Dogancay-Aktuna and Hardman, 2017). In addition, teachers believe that they are the “custodian” (Bayyurt and Sifakis, 2017, p. 6) of standardized English, or they think that WE-informed pedagogy is desirable but not necessary (Matsuda, 2009). Furthermore, teacher candidates are reluctant to fundamentally change their traditional understanding of English to something more fluid, unstable, and heterogeneous like WE or EIL, as they might see such change as a dismissal or denial of the past effort and investment in English learning (Matsuda, 2017).

In addition to teacher education, another structural issue is assessment, which serves as a particularly strong mechanism to set, maintain, and reinforce the linguistic norms (Davidson, 2006; Hu, 2018; Shohamy, 2006, 2011). Moreover, the trend to utilize global assessment tools like TOEFL or IELTS, which is largely motivated by neoliberal pressure that holds people accountable with measurable, comparable, and globally transferable scores, has caused two problems. The first is that large-scale tests like TOEFL often use Inner Circle Englishes for test items and scoring criteria, reproducing and imposing such Englishes as a global norm. The second is that large-scale tests are often developed for specific purposes (e.g., TOEFL is to assess learners’ academic English ability to study in North American context), they may not be suitable to local learners’ needs (Chalhoub-Deville and Wigglesworth, 2005; Davidson, 2006; McNamara, 2011).

Although theoretical discussions have been conducted on linguistic inequality in assessment (e.g., Brown, 2014; Canagarajah, 2006a; Hu, 2012, 2018), transformation in practice rarely happens for three reasons. First, because test developers need to establish construct validity, they need English that is well-described, which necessarily becomes the traditional Inner Circle Englishes (Brown, 2014). Second, test-making is increasingly becoming a political act, and policy makers often decide specific benchmarks against which test-takers are evaluated, such as the Common European Framework of Reference (CEFR) (McNamara, 2011). McNamara (2011) argues that although test developers used to define benchmarks using the latest theories on language, communication, and other related areas, this “‘thinking stage’ of test development has now been removed” and replaced with “complex policy procedures” (p. 502). He shows how CEFR has pre-determined standards of language proficiency—some of which assess the proficiency of a speaker vis-à-vis that of native speakers—and how those policy-driven standards influence society to a large extent. Third, and relatedly, benchmarks, especially large-scale, influential frameworks like CEFR, are deeply embedded in existing social structure, thus changing them would be naturally challenging as multiple administrative and legislative procedures are involved (McNamara, 2011).

McNamara’s (2011) insights on sociopolitical aspects of assessment are important, as they hold true in educational policy in general, not only assessment. This is another stumbling block to realize WE-informed pedagogy. Given that calling for WE-informed pedagogy is “quite radical” (Matsuda, 2012a, p. 6), transformation of the current Inner Circle-based teaching model requires not only practitioners’ micro-level efforts, but also result-oriented activism that impacts national- and supra-national-level policy. Since policymaking is largely driven by politics, not always rationality (Edelman, 1988; Wodak, 2011), scholars’ rational critique itself may not lead to policy change. Thus, WE scholarship for transforming policies and practices needs to break the scholarly bubble and reach out to and work together with a wider audience. This leads and is connected to stumbling blocks in the scholarly domain (see “Scholarly domain”).

Practical domain

The practical domain is perhaps the biggest hindrance to realize WE-informed pedagogy. First, since the school systems still operate within standardized curricula to a great extent, teachers have almost no choice but to rely on traditional ELT. That is, while WE

scholarship appreciates and describes the diversity of Englishes in society, thus questioning such notions as norms and standards, traditional teaching still tends to rely on them because they are useful—almost necessary—to conceptualize teaching contents, set goals, create materials, and assess learning.

Second, as mentioned earlier, normative ideologies of English are especially pervasive in academic writing. This is despite the recent popularity of translanguaging, which problematizes the view that students' multilingual competence is a sum of abilities in named languages that are separated by fixed linguistic boundaries (e.g., [García and Kleyn, 2016](#)). Instead, translanguaging valorizes students' agentive maneuvering of diverse linguistic repertoires that are available to them. Yet, academic writing in English typically expects students to follow established conventions. Any deviation can risk negative consequences especially in high-stakes tasks, such as writing for tests, publications, grant proposals, job applications, and so on ([Heng Hartse and Kubota, 2014](#)). Thus, the popularized pedagogical approach to translingualism has been critiqued by second language writing professionals, who are skeptical of the academic and professional viability of this approach for students who are second language writers in English (see [Silva and Wang, 2020](#)).

Third, even if practitioners desire to practice WE-informed pedagogy, they are often not supported materially, meaning that there is a dearth of resources to guide this type of pedagogy ([Galloway and Rose, 2018](#); [Matsuda, 2012b](#)). Researchers point out that materials published by large publishers continue to feature Inner Circle Englishes and speakers ([Kumaravadivelu, 2016](#); [Matsuda, 2012b](#); [Yamada, 2015](#)), who are also often racialized as white ([Kubota, 2018b, 2021](#)). ELT scholarship has seen proliferation of theoretical critiques against traditional, Inner Circle-based ELT over the last 50 years but teaching practices and publishers of teaching materials have not kept up with the speed of such theoretical development.

Finally, even if teachers were able to include some elements of WE in their classroom (see [Stepping stones toward the changing paradigm](#) section), their practices may become simplistic and superficial because of time constraints, lack of institutional or material support, parents' demands for traditional Inner Circle-based ELT, and so forth. This reality hinders a progressive implementation of WE-informed pedagogy, because it ironically stigmatizes WE-informed practitioners—despite their good intentions—as essentializing a certain variety of English through their simplistic or superficial representations. This means that practitioners could be targets of scholarly criticisms either for continuing traditional ELT or for incorporating elements of WE in the classroom if their practices are simplistic and superficial.

Scholarly domain

Although the significance of theoretical advancement in WE scholarship is unquestionable, scholars too have their own stumbling blocks, and can be complicit in reifying the status quo in the following ways. First, some scholars critically point out that much discussion on WE-informed pedagogy remains abstract and unactionable for practitioners and policymakers ([Kubota, 2018b](#); [Matsuda, 2012a](#); [Matsuda and Friedrich, 2011](#)). Others also critique how theoretical discussions of the social necessity for WE-informed pedagogy often take place without insights and perspectives from local classrooms ([Galloway and Rose, 2018](#); [Matsuda, 2012a](#); [Tupas, 2010](#)). This is an unhealthy situation because the lack of consideration of local contexts prevents practitioners from integrating the theoretical knowledge into their everyday practice.

Second, as was already mentioned, scholars can be complicit with sustaining the hegemony of standard English due to neoliberalism ([Kubota, 2016](#); [O'Regan, 2021](#)). This is because most scholars work in institutions operating under the neoliberal regime, which champions "more and faster" as a basic principle and care about international rankings and journal impact factors for self-branding and accountability purposes. Thus, scholars are evaluated against how much they published in prestigious journals and how much grant funding they won. Here, a long-standing, unquestioned, and default medium of writing is standardized English ([Heng Hartse and Kubota, 2014](#); [O'Regan, 2021](#)), and the chances of getting accepted for publication and winning grants significantly drop when the work is written in non-standardized form of English. Therefore, ironically, even WE scholars might need to be complicit with this long-standing hegemony of standardized English to make their voice heard in the first place.

Third, scholars' critiques tend to remain within academia ([Kubota, 2018b](#)). In order to disrupt the existing hegemony sustained by Inner Circle standardized English, scholars should seek ways to reach out to the wider public and disseminate their findings, thereby informing practitioners with practical suggestions, communicating with policy makers, or making efforts to debunk popular ideologies about English language.

Fourth, scholarly activities also tend to be confined within certain conceptual orientations. As [Matsuda and Matsuda \(2018\)](#) pointed out, many scholars who are interested in the plurality of English align with their own theoretical orientation without much communication with each other. However, regardless of their conceptual differences, WE, ELF, EIL, GE, and translingual scholars all acknowledge that the messiness of English(es) cannot be captured by monolingual ideology, aim to move away from the native-speaker norm, and advocate equal rights for all languages ([Alsagoff, 2012](#); [Matsuda, 2017](#)). Resistance to the established teaching practices and policies entails a "major overhaul" ([Matsuda and Friedrich, 2012](#), p. 25) of the current system and belief, and if so, like-minded scholars should work in collaboration, not in isolation, for that same goal.

Stepping stones toward the changing paradigm

Despite these existing limitations and obstacles, ELT practitioners' efforts to contextually apply WE in the classroom continue until today, and have been documented across [Kachru's \(1985\)](#) three concentric circles: the Inner Circle (e.g., [Henderson Lee and Pandey, 2021](#); [Kubota, 2001](#); [Sharifian and Marlina, 2012](#); [Tardy et al., 2020](#)), the Outer Circle (e.g., [Ali, 2015](#); [Sung, 2015](#)), and the Expanding Circle (e.g., [Bayyurt and Altinmakas, 2012](#); [Chang, 2014](#); [D'Angelo, 2012](#); [Galloway and Rose, 2018](#); [Lee, 2012](#); [Rose and](#)

Galloway, 2017). These reports on how the WE paradigm is used in classroom settings can be seen as stepping stones that help bridge existing gaps between WE's theoretical potential and its pedagogical application. The stepping stones represent practitioners' work to bring about a WE-informed pedagogical paradigm shift, necessitating both "an entirely different set of assumptions than the teaching and learning of any other second or foreign language" (McKay, 2002, p. 1) and "a thorough re-conceptualization and a thorough re-organization of knowledge systems" (Kumaravadivelu, 2012, p. 14). In what follows, we will thematically showcase WE-informed/relevant pedagogical innovations made by teacher-scholars and highlight subsequent trends and implications for ELT.

Awareness-raising

As Matsuda (2020) argues, what both students and teachers increasingly need today is more "the metaknowledge of the linguistic diversity than the specific formal and functional rules of particular varieties" (p. 688). Indeed, the most apparent pedagogical aspect interwoven into Kachru's (1992) six proposals for classroom innovations are directly concerned with awareness-raising regarding WE. Awareness-raising instruction has often included deliberate exposure to English varieties and the validation of fluid, rather than prescriptive, uses of English.

Galloway and Rose (2018) report on a classroom-based research study in a Japanese university that investigated the instructional efficacy of a learning task on GE. The task aimed to raise awareness of GE through exploring a regional variety of English chosen from Kachru's concentric circles or investigating the use of ELF for a specific purpose. The students investigated an English variety of their personal choice using the library, internet, and other resources to conduct a ten-minute presentation using visual aids. According to Galloway and Rose, students' written reflections on the task suggest that the activity was positively received; it enabled students to discover and reflect on the linguistic features, functions, and historical development of the varieties they had selected. Reflections on language and identity, standardized English, and Indigenous languages were also observed.

Sung's (2015) study is another example that focuses on awareness-raising activities in an Asian classroom context. Sung implemented a GE-oriented component to raise students' awareness of WE variation in a university English course in Hong Kong. The classroom activities ranged from listening tasks such as accent identification of different varieties of English to discussion tasks regarding standardized English, native speakers, and ELF communication. Sung reported that students welcomed these awareness-raising activities, which prompted them to think through different varieties of English, the reality of English use in the world, issues surrounding (non)native speakers in intercultural communication, and so forth. However, the study also showed that although these activities heightened their awareness of GE, most students continued to hold negative perspectives toward non-native varieties of English, exemplifying the pedagogical conundrum of challenging the native-speaker norm within a limited instructional time.

Attempts to direct students' attention to WE in the classroom are not only made for non-native learners of English in ELT contexts but also for native speakers of English. Arguing that Inner Circle native speakers of mainstream English should also share the responsibility for successful communication, Kubota (2001) details an instructional unit that exposed American high school students to inter/intra-national varieties of English. The unit invited them to think through the global spread of English, the reality and difficulty of second language acquisition, and their responsibility and available strategies for communication with WE speakers. Instructional modes and materials included lectures, group work, whole class discussion, video clips, reflective writing, and interaction with guest WE speakers. Based on the results of the pre- and post-surveys, Kubota observed positive change in students' attitudes toward and comprehension of WE, their understanding of speech samples, as well as the difficulty of second language acquisition. However, significant improvements in perceived quality of speech samples, perceived personal traits of WE speakers, and desire to communicate with them was not confirmed. Furthermore, some students made xenophobic remarks, drew on problematic stereotypes, and showed ethnocentric biases during in-class discussions, indicating a challenge for classroom intervention. The results of this project offer pedagogical implications that highlight "the necessity of affirming linguistic diversity at all educational levels and creating a pedagogical environment conducive to developing critical consciousness on the global spread of English" (Kubota, 2001, p. 62).

(Un)learning language ideology

As the WE paradigm questions and challenges the linguistic norms established exclusively by Inner Circle varieties of English, critical discussions for decentering and (un)learning standardized language norms have been incorporated into the classroom. Such discussions are often an extension and/or application of awareness-raising instruction through exposure to varieties of WE, and challenge students to critically engage with issues surrounding language and power.

A study by Chang (2014) reports on the pedagogical impact of incorporating critical issues surrounding WE into a university-level course in Taiwan. Using activities that included lectures, discussions, student presentations, reflective writing, and a final exam, the WE course invited students to reflect on various issues such as standard language ideology, the "native" and "non-native" dichotomy, as well as the global dominance of English and its influence on their community. According to Chang, students' reflections suggested that the course had helped them redefine the purpose of language learning, while increasing their open-mindedness toward different varieties of English and prompting them to challenge the native-speaker norm. One student wrote, "You do not have to care about the red lines under it when we type the word Englishes on MS word" (p. 25), implying the course had impacted their beliefs about the legitimacy and ownership of English. The pedagogical impact elaborated in the study in the form of student reflection was overwhelmingly positive, suggesting the possibility for incorporating critical issues into the ELT classroom, such as language ideology and power.

Rose and Galloway (2017) detail a debate activity conducted in a university-level course in Japan, which used the “Speak Good English Movement” in Singapore as a topic to debate on the standardized language ideology. The movement encouraged Singaporeans to speak and write in standardized English, rather than in Singlish or a Singaporean variety of English, because it was not perceived as intelligible enough in the globalized world. The students were invited to argue for or against the movement. Pre-debate activities included a research task that encouraged students to gather evidence from newspaper articles and online videos to support their argument. After the debate activity, students shared their accounts of their positions and beliefs in the form of written reflections, which were used as a data source to gauge the impact of this activity on students’ awareness of language attitudes and ideologies. Rose and Galloway concluded that the use of the “Speak Good English Movement” for the debate successfully enabled students to increase their awareness of the sociolinguistic profile of Singaporean English, to think through its cultural significance for Singaporeans, and to challenge standardized language ideology in the process. This classroom innovation is a good example of how authentic WE-relevant issues can be introduced into ELT classrooms in which the native-speaker norm perpetuates, and can raise students’ (and teachers’) awareness “that standards are imagined benchmarks, and that it is the actual use of a language that indicates its legitimacy” (Rose and Galloway, 2017, p. 300).

Writing

Writing is one of the linguistic domains where standardized English has prevailed (e.g., Gupta, 2006). However, scholars continue to contest the pervasive preference for literacy skills predicated entirely on standardized English, advocating WE-informed pedagogical spaces not only in creative writing (e.g., Lim, 2015) but also in more conservative genres such as academic writing. Canagarajah (2006b) argues that students require a linguistic and strategic repertoire that allows them to “shuttle between communities” (p. 593) in socially situated ways, as opposed to striving for a single standard form of language. This can be achieved by a literacy practice of what Canagarajah calls code-meshing, a strategy of merging local variants with standardized English in rhetorically effective ways in academic writing. Canagarajah details the pedagogical use of code-meshing, through the use of writing models that employ African American Vernacular English in texts for multilingual students to understand how linguistic creativity can be performed in their writing. The notion of code-meshing is not simply mixing or switching back and forth different varieties of a named language or different languages. It symbolizes the ways writers shuttle between their available linguistic resources rather than moving between discrete and bounded linguistic categories (Canagarajah, 2013). This conceptualization of language, which questions linguistic competence defined by fixed linguistic boundaries, is consistent with the notion of translanguaging which has become popular in recent years (e.g., Cenoz and Gorter, 2020).

Calling for the development of global literacy responsive to the realities where students socially engage in the multiplicity of English, Matsuda and Matsuda (2010) offer guiding principles for teachers who wish to implement WE-informed writing pedagogy in the classroom. They include the teaching of (a) the dominant language forms and functions (so that students understand the privileged status certain English varieties enjoy and so that they can appropriate them when they need to); (b) the nondominant language forms and functions (so that students understand deviations are employed in the writing in socially meaningful ways; etc.); (c) the perceived boundaries between variations and errors; (d) the strategies that can be employed when using nondominant forms in socially acceptable ways; and (e) the potential risks involved (so that students become sensitive to how and when deviations from the norm can/cannot be effective especially in the dominant discourse they negotiate). ELT practitioners’ efforts to contextually apply such principles and strategies to L2 writing classrooms have been made to this date.

Henderson Lee and Pandey (2021) report on an instructional unit for multilingual writers in an American university, which incorporated WE texts as a catalyst for nurturing global literacy. Using a TED Talk called *3 Ways to Speak English* (Lyiscott, 2014) and a text from *Rotten English: A Literary Anthology* (Ahmad, 2007) as stimulus WE materials, the authors showed that students were able to raise their awareness about how and when nondominant forms can rhetorically strengthen their written texts, while positively impacting their beliefs and attitude toward language varieties. The conclusion of the study offers a list of potential resources that can navigate WE-informed writing instructions.

Tardy et al. (2020) is another recent study which took place in a course for multilingual writers in an American university. Combining content-based and language and genre awareness approaches, the course engaged students in exploring WE concepts such as standardization, nativeness, intelligibility and identity, and dominant and nondominant varieties. The article elaborates on various activities used for the course, one of which is inviting students to write an academically argumentative response to Matsuda and Matsuda’s (2010) article (mentioned above). Writing a response to scholars’ suggestions allowed students “to share their views on English language norms, validating their experiences and values” (p. 7), indicating the possibility of using scholarly articles on WE as stimulus texts in advanced literacy instruction.

Communication strategies

The teaching of pragmatic and linguistic strategies has long been explored in ELT since its early days (Dörnyei, 1995). Traditionally, such communication strategies have been regarded as verbal and non-verbal techniques to compensate for communicative breakdowns or increase communicative effectiveness (Canale, 1983). They have been discussed as part of strategic competence that constitutes communicative competence from the perspective of second language teaching (Canale and Swain, 1980). In today’s globalizing world where diverse users of English bring their linguistic repertoire into a specific communicative space, communication strategies have been reconceptualized as mediating strategies necessary for crossing linguistic and cultural boundaries. Canagarajah (2007, p. 236) lists different interaction practices where varieties are negotiated and communication strategies play a mediating role:

- Code-switching, crossing (Rampton, 1995);
- Speech accommodation (Giles, 1984);
- Interpersonal strategies: i.e. repair, rephrasing, clarification, gestures, topic change, consensus-oriented, mutually supportive (Gumperz, 1982; Seidlhofer, 2004);
- Attitudinal resources: i.e. patience, tolerance and humility to negotiate differences (see Higgins, 2003).

Stressing the relevance of communication strategies for speakers who use different varieties of English (or other languages), Canagarajah calls for a pedagogical focus on developing such strategies as opposed to (standardized) rules and conventions. Similarly, Kubota (2012) and Kubota and Takeda (2021) maintain the need to teach students communication strategies for border-crossing communication, which is “active, critical and reflective engagement in communication across diverse ethnic, racial, linguistic and socioeconomic differences” (Kubota, 2012, p. 63). Such communication contexts include both ELF situations where English serves as a common language and other situations where English does not. Furthermore, students may interact with L2 users of their L1. Border-crossing communication highlights the importance of accommodation skills, such as gauging interlocutors’ linguistic repertoires, supportive listening, paraphrasing, and adjusting speech speed. The emphasis on communication strategies rather than a sole emphasis on linguistic forms is also a response to reconceptualization of communicative competences in socially situated ways as informed by the WE and related paradigms. As Canagarajah (2007) argues, communicative proficiency in light of such paradigms requires being able to negotiate linguistic differences, where communication strategies are expected to come into play.

Responding to the critique of WE: pragmatic and critical utilization of boundaries

Having illustrated some of the pedagogical innovations and suggestions regarding WE, we would like now to turn to the core critique against WE: problematizing its reification of varieties based on geographical boundaries. In what follows, we examine the main claim of this critique and argue for pragmatic and critical utilization of WE in which even the theoretical limitations can be rendered as pedagogical opportunities.

WE contains an inherently modernist conception of language that associates certain linguistic features with geographical boundaries. Through its nation-based categorization, WE inadvertently gives an impression that those varieties are discretely codifiable and that the usage of English is somewhat uniform within those categories (Bruthiaux, 2003). With this, rather than fully decentralizing the standardized English, WE preserves the idea of centrality vis-a-vis its own boundary-making: it reproduces the idea that there are “typical” usages of the named English variety that are recognized as linguistically legitimate because they are frequently and systematically used by a certain linguistic community, while leaving “non-typical” usages unrecognized. Regardless of how extensively sociolinguistic data are collected, due to the generative nature of language use, utterances and writings that deviate from categorical norms theorized by the WE framework will continuously be produced. Thus, WE struggles to deal with deviations within a named variety. This theoretical limitation not only makes it difficult for WE to adequately account for increasing mobility and fluidity of semiotic practices in the globalizing world, but also risks reproducing the system of domination and oppression it purports to disrupt. As Pennycook (2007, p. 21) stated, “by focusing centrally on the development of new national Englishes, the world Englishes approach reproduces the very linguistics it needs to escape.”

Nonetheless, at the pedagogical level, WE continues to provide a useful framework as long as it is adopted pragmatically and critically. Pragmatic adoption of the WE framework provisionally utilizes the boundaries of varieties while recognizing the contingency and contestedness of drawing such boundaries. As we discussed earlier, strategic use of boundaries enables one to claim legitimacy and (sociopolitical) recognition of diverse ways people use English and effectively highlight those uses in classrooms. If we dismiss the existence of languages and their varieties because they are “the inventions of social, cultural and political movements” (Makoni and Pennycook, 2007, p. 2), we lose the ground for claiming their rights and ensuring equitable representations of them (MacSwan, 2020). This is a structurally similar issue created by the tensions between modernist and postmodernist forms of critique: whereas the former consolidates identity categories for emancipatory struggles, the latter dissipates those categories and rejects the totalizing tendency of the former (Pietikäinen, 2016; Takeda, 2022).

Yet, the pragmatic approach we advocate in this chapter renders these tensions into pedagogical opportunities. With the former, by tangibly identifying diverse varieties of English, we argue that teachers are better positioned to ensure the inclusivity of teaching materials and instructions. The pedagogical approaches we discussed above (e.g., Bayyurt and Altinmakas, 2012; Chang, 2014) adopt the WE framework to teach students not only about diverse varieties of English, but also about language ideologies and systemic inequalities perpetuated by Inner Circle norms. Without provisional acceptance of boundaries, linguistic differences are “flattened” (Gilyard, 2016), making it difficult to unsettle Inner Circle norms while claiming the recognition and identity of marginalized Englishes. Critical discussions of inequalities based on race, gender, nationality, and so on in association with language (Kubota, 2020) can also be made possible by such a utilization of boundaries.

While WE provides a useful framework for teachers to ensure diverse and inclusive representations of English varieties in the classroom, it must be adopted critically. Because any form of boundary work involves exclusion, uncritical utilization of the WE framework risks linguistic essentialism and delegitimization of English usages that are considered “irregular” within a given variety. In fact, WE typically claims legitimacy of a certain English variety by highlighting its systematicity and regularity or what Lee (2017) calls “frequency premise.” According to this premise, Lee states:

if a particular feature occurs frequently enough, even if it is a deviation from mainstream norms, it warrants recognition and legitimacy. The frequency premise is foundational to descriptivist approaches to language and is what drives the practice of linguistic codification (p. 28).

Although it is undeniable how this approach has successfully claimed the legitimacy of certain English varieties like Philippine English (Tupas, 2004), it simultaneously reinforces the authority of the standard—in this case, the variety of English used by educated elites in the Philippines. Because the claim of legitimacy in this way follows “the criteria of dominant apparatuses of authoritative English,” Lee (2017, p. 31) argues, “the frequency premise thus reifies the standards by focusing on the deviations from the very standards as such.”

Thus, uncritical adaptation of WE could inadvertently reinforce the idea of standards and delegitimize usages that fall outside of the boundaries. In this sense, some of the pedagogical approaches above could be potentially problematic. For instance, some of those pedagogies involved identifying accents of different varieties of English and learning about their sociolinguistic profiles. If those activities are implemented in such a way to solidify students’ idea of the varieties introduced without simultaneous destabilization of the boundaries, they risk reproducing the very ideology they aim to critique. For example, if an accent identification activity has a set of answers students are instructed to arrive at (e.g., the voice recording A is Singaporean accent) without introducing other examples that do not fall into a “typical” speech style, it is more likely to reinforce the idea of regional standards and stereotypes.

In the pragmatic and critical adoption of the WE framework, therefore, identification and profiling of English varieties should never be the end point. Rather, the boundaries should be introduced in order to be problematized. After learning about a certain variety, for example, students may engage in an activity in which they search online to find speech samples that do not correspond to the characteristics of the variety they learned (e.g., English spoken by Japanese nationals that is phonologically different from typical “Japanese English”). Such a centrifugal use of WE opens the possibility for recognizing the excess of our linguistic practices that incessantly go beyond the confines of the standard. In calling for moving from World Englishes to wayward Englishes, Lee (2017) states:

Encounters with different Englishes, even among people from the same country, are increasingly common, and the sheer heterogeneity of Englishes, even in geographic spaces traditionally understood to be culturally and linguistically homogeneous, means we are losing the ability to determine where someone is from based on their English, and what their English will be like on the basis of where they are from (p. 79).

“The sheer heterogeneity of Englishes” may never be capturable within the boundaries of WE or any other boundaries. Realization of such a theoretical limitation can be an excellent pedagogical moment. Just like a transgressive action is more visible against the background of strict rules, the waywardness of Englishes can be more tangibly encountered in the presence of boundaries. Thus, pragmatic and critical utilization of WE as a pedagogical tool introduces boundaries purposefully in order for them to be traversed. It is through this transgressive use of boundaries, WE remains a useful framework for ELT.

Conclusion

In this chapter, centering the idea of boundaries, we explored the whereabouts of WE today. For this, we first reviewed the historical development of WE and situated it in relation to the neighboring fields like EIL, ELF, and trans-/pluri-/metro-lingualism. This was followed by a discussion of the existing stumbling blocks that continue to impede actualization of pedagogy that is informed by WE and its related fields. The durability of these stumbling blocks demands pedagogical approaches that are both practical and creative. We then showcased some of the WE-informed pedagogies as stepping stones to help bridge existing gaps between WE’s theoretical potential and its pedagogical application. These pedagogical approaches make use of variational boundaries to expose students to diverse ways Englishes are used and to destabilize the idealization of the standard English. Finally, we turned to and examined the critique of WE and argued how theoretical limitations of WE can be rendered pedagogically useful by a pragmatic and critical use of linguistic boundaries. Even though some scholars regard WE as a framework that is “a 20th century construct that has outlived its usefulness” (Bruthiaux, 2003, p. 161), we posit that the pedagogical utility and potentiality of WE have not run out of steam. This is largely because our social, political, and educational systems still rely on modernist institutional conventions. Claiming linguistic rights and recognitions within those systems continues to demand strategic reifications (MacSwan, 2020). Thus, while recognizing the theoretical limitations of WE and the values of a more fluid, mobile, and entangled conception of language, we argue that the very limitations can be utilized for pedagogical purposes. In this way, WE can help to catalyze classroom explorations about linguistic diversity and its related issues, where students and teachers advance their understanding of the reality of English use and users in the globalizing world. Pragmatically and critically, we still use WE.

References

- Ahmad, D. (Ed.), 2007. *Rotten English: A Literary Anthology*. W.W. Norton, New York.
- Ali, Z., 2015. The prospect and potential challenges of teaching Englishes in Pakistan. *Asian Engl.* 17 (2), 152–169.
- Alsagoff, L., 2012. Identity and the EIL learner. In: Alsagoff, L., McKay, S.L., Hu, G., Renandya, W.A. (Eds.), *Principles and Practices for Teaching English as an International Language*. Routledge, New York, pp. 104–122.
- Bayyurt, Y., Altinmakas, D., 2012. A WE-based English communication skills course at a Turkish university. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 169–182.
- Bayyurt, Y., Sifakis, N., 2017. Foundations of an EIL-aware teacher education. In: Matsuda, A. (Ed.), *Preparing Teachers to Teach English as an International Language*. Multilingual Matters, Bristol, pp. 3–18.
- Berns, M., 2006. World Englishes and communicative competence. In: Kachru, B.B., Kachru, Y., Nelson, C.L. (Eds.), *The Handbook of World Englishes*. Blackwell, Malden, pp. 718–730.
- Biric Deniz, E., Özkan, Y., Bayyurt, Y., 2020. ELF-Awareness in Pre-service English Language Teacher Education: A Case Study From Turkey. *Hacettepe University J. Educ.* 35 (2), 270–284.
- Bolton, K., Davis, D.R., 2018. Larry Smith and world Englishes. *World Engl.* 37, 447–454.
- Bolton, K., 2020. World Englishes: current debates and future directions. In: Nelson, C.L., Proshina, Z.G., Davis, D.R. (Eds.), *The Handbook of World Englishes*, second ed. Wiley-Blackwell, New Jersey, pp. 743–760.
- Boonsuk, Y., Ambele, E.A., McKinley, J., 2021. Developing awareness of Global Englishes: moving away from “native standards” for Thai university ELT. *System* 99.
- Brown, J.D., 2014. The future of World Englishes in language testing. *Lang. Assess. Q.* 11 (1), 5–26.
- Bruthiaux, P., 2003. Squaring the circles: issues in modeling English world-wide. *Int. J. Appl. Ling.* 13 (2), 159–178.
- Canagarajah, S., 2006a. Changing communicative needs, revised assessment objectives: testing English as an international language. *Lang. Assess. Q.* 3 (3), 229–242.
- Canagarajah, A.S., 2006b. The place of World Englishes in composition: pluralization continued. *Coll. Compos. Commun.* 57 (4), 586–619.
- Canagarajah, S., 2007. After disinvention: possibilities for communication, community and competence. In: Makoni, S., Pennycook, A. (Eds.), *Disinventing and Reconstituting Languages*. Multilingual Matters, Clevedon, pp. 233–239.
- Canagarajah, S., 2013. *Translingual Practice: Global Englishes and Cosmopolitan Relations*. Routledge, London.
- Canale, M., Swain, M., 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Appl. Ling.* 1 (1), 1–47.
- Canale, M., 1983. From communicative competence to communicative language pedagogy. In: Richards, J.C., Schmidt, R.W. (Eds.), *Language and Communication*. Longman, Harlow, pp. 2–27.
- Genoz, J., Gorter, D., 2020. Teaching English through pedagogical translanguaging. *World Engl.* 39 (2), 300–311.
- Chalhoub-Deville, M., Wigglesworth, G., 2005. Rater judgment and English language speaking proficiency. *World Engl.* 24 (3), 383–391.
- Chang, Y.J., 2014. Learning English today: what can World Englishes teach college students in Taiwan? *Engl. Today* 30 (1), 21–27.
- Cook, V., 2007. The goals of ELT. In: Cummins, J., Davison, C. (Eds.), *International Handbook of English Language Teaching (Part I)*. Springer, New York, pp. 237–248.
- Cushing, I., 2021a. Policy mechanisms of the standard language ideology in England’s education system. *J. Lang. Ident. Educ.*
- Cushing, I., 2021b. ‘Say it like the Queen’: the standard language ideology and language policy making in English primary schools. *Lang. Lang. Cult. Curric.* 34 (3), 321–336.
- D’Angelo, J., 2012. WE-informed EIL curriculum at Chukyo: towards a functional, educated, multilingual outcome. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 121–139.
- Davidson, F., 2006. World Englishes and test construction. In: Kachru, B.B., Kachru, Y., Nelson, C.L. (Eds.), *The Handbook of World Englishes*. Blackwell, Malden, pp. 709–717.
- Davila, B., 2016. The inevitability of “standard” English. *Writ. Commun.* 33 (2), 127–148.
- Dogancay-Aktuna, S., Hardman, J., 2017. A framework for incorporating an English as an international language perspective into TESOL teacher education. In: Matsuda, A. (Ed.), *Preparing Teachers to Teach English as an International Language*. Multilingual Matters, Bristol, pp. 19–31.
- Dörnyei, Z., 1995. On the teachability of communication strategies. *Tesol Q.* 29 (1), 55–85.
- Edelman, M.J., 1988. *Constructing the Political Spectacle*. University of Chicago Press, Chicago.
- Fraser, N., 1995. From redistribution to recognition? Dilemmas of justice in a “postsocialist” age. *N. Left Rev.* 1 (212), 68–93.
- Galloway, N., Rose, H., 2018. Incorporating global Englishes into the ELT classroom. *ELT J.* 72 (1), 3–14.
- Galloway, N., 2017. ELF and ELT teaching materials. In: Jenkins, J., Baker, W., Dewey, M. (Eds.), *The Routledge Handbook of English as a Lingua Franca*. Routledge, London, pp. 468–480.
- García, O., Kleyn, T. (Eds.), 2016. *Translanguaging With Multilingual Students: Learning From Classroom Moments*. Routledge, New York.
- García, O., 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell, Malden.
- The dynamics of speech accommodation. In: Giles, H. (Ed.), *Int. J. Sociol. Lang.* 46 special issue.
- Gilyard, K., 2016. The rhetoric of translanguaging. *Coll. Engl.* 78 (3), 284–289.
- Gumperz, J.J., 1982. *Discourse Strategies*. Cambridge University Press, Cambridge.
- Gupta, A.F., 2006. Standard English in the world. In: Rubdy, R., Saraceni, M. (Eds.), *English in the World*. Continuum, New York, pp. 95–109.
- Harvey, D., 2005. *A Brief History of Neoliberalism*. Oxford University Press, New York; Oxford.
- Hashimoto, K., 2013. The construction of the “native speaker” in Japan’s educational policies for EFL. In: Houghton, S.A., Rivers, D.J. (Eds.), *Native-Speakerism in Japan: Intergroup Dynamics in Foreign Language Education*. Multilingual Matters, Bristol, pp. 159–168.
- Henderson Lee, S., Pandey, S., 2021. World Englishes in the first-year composition classroom: perceptions of multilingual writers. In: Hassel, H., Cole, K. (Eds.), *Transformations: Change Work Across Writing Programs, Pedagogies, and Practices*. Utah State University Press, Colorado, pp. 210–244 in press.
- Heng Hartse, J., Kubota, R., 2014. Pluralizing English? Variation in high-stakes academic texts and challenges of copyediting. *J. Sec. Lang. Writ.* 24, 71–82.
- Higgins, C., 2003. “Ownership” of English in the outer circle: an alternative to the NS-NNS dichotomy. *Tesol Q.* 37 (4), 615–644.
- Hino, N., 2018. *EIL Education for the Expanding Circle: A Japanese Model*. Routledge, London.
- Holliday, A., 2006. Native-speakerism. *ELT J.* 60 (4), 385–387.
- Holliday, A., 2015. Native-speakerism: taking the concept forward and achieving cultural belief. In: Swan, A., Aboshiha, P., Holliday, A. (Eds.), *(En)Countering Native-Speakerism: Global Perspectives*. Palgrave Macmillan, Hampshire, pp. 11–25.
- Hu, G., 2012. Assessing English as an international language. In: Alsagoff, L., McKay, S.L., Hu, G., Renandya, W.A. (Eds.), *Principles and Practices for Teaching English as an International Language*. Routledge, New York, pp. 123–143.
- Hu, G., 2018. The challenges of world Englishes for assessing English proficiency. In: Low, E.L., Pakir, A. (Eds.), *World Englishes: Rethinking Paradigms*. Routledge, New York; London, pp. 78–95.
- Jenkins, J., 2018. ELF and WE: competing or complementing paradigms? In: Low, E.L., Pakir, A. (Eds.), *World Englishes: Rethinking Paradigms*. Routledge, Oxon, pp. 12–28.
- Kachru, B.B., 1985. Standard, codification, and sociolinguistic realism: the English language in the outer circle. In: Quirk, R., Widdowson, H.G. (Eds.), *English in the World: Teaching and Learning the Language and Literatures*. Cambridge University Press, Cambridge, pp. 11–30.
- Kachru, B.B., 1992. Teaching world Englishes. In: Kachru, B.B. (Ed.), *The Other Tongue: English Across Cultures*, second ed. University of Illinois Press, Urbana, pp. 355–365.
- Kachru, B.B., 2013. History of world Englishes. In: Chapelle, C.A. (Ed.), *The Encyclopedia of Applied Linguistics*. John Wiley, West Sussex, pp. 1–12.

- Kang, O., Rubin, D.L., 2009. Reverse linguistic stereotyping: measuring the effect of listener expectations on speech evaluation. *J. Lang. Soc. Psychol.* 28, 441–456.
- Kubota, R., Takeda, Y., 2021. Language-in-education policies in Japan versus transnational workers' voices: two faces of neoliberal communication competence. *Tesol Q.* 55 (2), 458–485.
- Kubota, R., 2001. Teaching world Englishes to native speakers of English in the USA. *World Engl.* 20 (1), 47–64.
- Kubota, R., 2011. Questioning linguistic instrumentalism: English, neoliberalism, and language tests in Japan. *Ling. Educ.* 22 (3), 248–260.
- Kubota, R., 2012. The politics of EL: toward border-crossing communication in and beyond English. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, UK, pp. 55–69.
- Kubota, R., 2015. Inequalities of Englishes, English speakers, and languages: a critical perspective on pluralist approaches to English. In: Tupas, R. (Ed.), *Unequal Englishes*. Palgrave Macmillan, New York, pp. 21–41.
- Kubota, R., 2016. The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: complicities and implications for applied linguistics. *Appl. Ling.* 37 (4), 474–494.
- Kubota, R., 2018a. *Eigo kyōiku gensō* [Illusions in English Education]. Chikuma Shinsho, Tokyo.
- Kubota, R., 2018b. Unpacking research and practice in world Englishes and second language acquisition. *World Engl.* 37 (1), 93–105.
- Kubota, R., 2020. Toward a performative commitment to heterogeneity. In: Silva, T., Wang, Z. (Eds.), *Reconciling Translingualism and Second Language Writing*. Routledge, New York, pp. 163–171.
- Kubota, R., 2021. Critical antiracist pedagogy in ELT. *ELT J.* 75 (3), 237–246.
- Kumaravadivelu, B., 2012. Individual identity, cultural globalization, and teaching English as an international language: the case for an epistemic break. In: Alsagoff, L., McKay, S.L., Hu, G., Renandya, W.A. (Eds.), *Principles and Practices for Teaching English as an International Language*. Routledge, New York, pp. 9–27.
- Kumaravadivelu, B., 2016. The decolonial option in English teaching: can the subaltern act? *Tesol Q.* 50 (1), 66–85.
- Lee, H., 2012. World Englishes in a high school English class: a case from Japan. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 154–168.
- Lee, J.W., 2017. *The Politics of Translingualism: After Englishes*. Routledge, New York.
- Lim, S.G.-L., 2015. Creative writing pedagogy for world Englishes students. *World Engl.* 34 (3), 336–354.
- Lyiscott, J., 2014. 3 Ways to Speak English. Available at: https://www.youtube.com/watch?v=k9fmJ5xQ_mc.
- MacSwan, J., 2020. Translinguaging, language ontology, and civil rights. *World Engl.* 39 (2), 321–333.
- Makoni, S., Pennycook, A., 2007. Disinventing and reconstituting languages. In: Makoni, S., Pennycook, A. (Eds.), *Disinventing and Reconstituting Languages*. Multilingual Matters, Clevedon, pp. 1–41.
- Matsuda, A., Friedrich, P., 2011. English as an international language: a curriculum blueprint. *World Engl.* 30 (3), 332–344.
- Matsuda, A., Friedrich, P., 2012. Selecting an instructional variety for an EIL curriculum. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 17–27.
- Matsuda, A., Matsuda, P.K., 2010. World Englishes and the teaching of writing. *Tesol Q.* 44 (2), 369–374.
- Matsuda, A., Matsuda, P.K., 2018. Teaching English as an international language: a WE-informed paradigm for English language teaching. In: Low, E.L., Pakir, A. (Eds.), *World Englishes: Rethinking Paradigms*. Routledge, London, pp. 64–77.
- Matsuda, A., 2003. The ownership of English in Japanese secondary schools. *World Engl.* 22 (4), 483–496.
- Matsuda, A., 2009. Desirable but not necessary? The place of World Englishes and English as an international language in English teacher preparation programs in Japan. In: Sharifan, F. (Ed.), *English as an International Language: Perspectives and Pedagogical Issues*. Multilingual Matters, Bristol; Buffalo; Toronto, pp. 169–189.
- Matsuda, A., 2012a. Introduction: teaching English as an international language. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 1–14.
- Matsuda, A., 2012b. Teaching materials in EIL. In: Alsagoff, L., McKay, S.L., Hu, G., Renandya, W.A. (Eds.), *Principles and Practices for Teaching English as an International Language*. Routledge, New York, pp. 168–185.
- Matsuda, A., 2017. Introduction. In: Matsuda, A. (Ed.), *Preparing Teachers to Teach English as an International Language*. Multilingual Matters, Bristol, pp. xiii–xxi.
- Matsuda, A., 2020. World Englishes and pedagogy. In: Nelson, C.L., Proshina, Z.G., Davis, D.R. (Eds.), *The Handbook of World Englishes*, second ed. Wiley-Blackwell, New Jersey, pp. 686–702.
- McKay, S.L., 2002. *Teaching English as an International Language: Rethinking Goals and Approaches*. Oxford University Press, Oxford.
- McKenzie, R.M., 2010. *The Social Psychology of English as a Global Language: Attitudes, Awareness and Identity in the Japanese Context*. Springer, London; New York.
- McNamara, T., 2011. Managing learning: authority and language assessment. *Lang. Teach.* 44 (4), 500–515.
- Mimatsu, T., 2011. ELF versus EFL: teaching English for “international understanding” in Japan. In: Archibald, A., Cogo, A., Jenkins, J. (Eds.), *Latest Trends in ELF Research*. Cambridge Scholars Publishing, Newcastle on Tyne, pp. 251–268.
- O'Regan, J.P., 2021. *Global English and Political Economy*. Routledge, London.
- Otsuji, E., Pennycook, A., 2010. Metrolingualism: fixity, fluidity and language in flux. *Int. J. Multiling.* 7 (3), 240–254.
- Otsuji, E., Pennycook, A., 2013. Unremarkable hybridities and metrolingual Practices. In: Ruddy, R., Alsagoff, L. (Eds.), *The Global-Local Interface and Hybridity: Exploring Language and Identity*. Multilingual Matters, Bristol, pp. 83–99.
- Pennycook, A., 2007. *Global Englishes and Transcultural Flows*. Routledge, Abingdon.
- Phillipson, B., 1992. *Linguistic Imperialism*. Oxford University Press, Oxford.
- Pietikäinen, S., 2016. Critical debates: discourse, boundaries and social change. In: Coupland, N. (Ed.), *Sociolinguistics: Theoretical Debates*. Cambridge University Press, New York, pp. 263–281.
- Rampton, B., 1995. *Crossing: Language and Ethnicity Among Adolescents*. Longman, London.
- Ricento, T., 2015. Political economy and English as a “global” language. In: Ricento, T. (Ed.), *Language Policy and Political Economy: English in a Global Context*. Oxford University Press, New York, pp. 27–47.
- Rose, H., Galloway, N., 2017. Debating standard language ideology in the classroom: using the “speak good English movement” to raise awareness of global Englishes. *RELC J.* 48 (3), 294–301.
- Rose, H., Galloway, N., 2019. *Global Englishes for Language Teaching*. Cambridge University Press, Cambridge.
- Saraceni, M., 2015. *World Englishes: A Critical Analysis*. Bloomsbury, London.
- Sayer, P., 2015. “More & earlier”: neoliberalism and primary English education in Mexican public schools. *L2 J.* 7 (3), 40–56.
- Seidlhofer, B., 2004. Research perspectives on teaching English as a Lingua Franca. *Annu. Rev. Appl. Ling.* 24, 209–239.
- Seidlhofer, B., 2011. *Understanding English as a Lingua Franca*. Oxford University Press, Oxford.
- Sharifan, F., Marlina, R., 2012. English as an International Language (EIL): an innovative academic program. In: Matsuda, A. (Ed.), *Principles and Practices of Teaching English as an International Language*. Multilingual Matters, Bristol, pp. 140–153.
- Sharifan, F. (Ed.), 2009. *English as an International Language: Perspectives and Pedagogical Issues*. Multilingual Matters, Bristol.
- Shohamy, E., 2006. *Language Policy: Hidden Agendas and New Approaches*. Routledge, New York; London.
- Shohamy, E., 2011. Assessing multilingual competencies: adopting construct valid assessment policies. *Mod. Lang. J.* 95 (3), 418–429.
- Silva, T., Wang, Z. (Eds.), 2020. *Reconciling Translingualism and Second Language Writing*. Routledge, New York.
- Sung, C.C.M., 2015. Implementing a Global Englishes component in a university English course in Hong Kong. *Engl. Today* 31 (4), 42–49.
- Takeda, Y., 2022. Praxis of critical literacy: pragmatic utilization of theoretical tensions. *Crit. Educ.* 13 (1), 26–50.

- Tardy, C.M., Reed, K., Slinkard, J.R., LaMance, R., 2020. Exploring Global Englishes content and language variation in an academic writing course. *TESOL J.* 12 (1), 1–11.
- Tupas, R., 2004. The politics of Philippine English: neocolonialism, global politics, and the problem of postcolonialism. *World Engl.* 23, 47–58.
- Tupas, R., 2010. Which norms in everyday practice: and why? In: Kirkpatrick, A. (Ed.), *The Routledge Handbook of World Englishes*. Routledge, Oxon, pp. 567–579.
- Urciuoli, B., 2008. Skills and selves in the new workplace. *Am. Ethnol.* 35 (2), 211–228.
- Wodak, R., 2011. *The Discourse of Politics in Action: Politics as Usual*, second ed. Palgrave Macmillan, London.
- Yamada, M., 2015. *The Role of English Teaching in Modern Japan: Diversity and Multiculturalism Through English Language Education in a Globalized Era*. Routledge, Abingdon.

The literacy journey: from a psychological to a biological way of thinking about learning to read and write

Brian L. Cambourne^a and Debra K. Crouch^b, ^a University of Wollongong, Wollongong Heights, NSW, Australia; and ^b Educational Consultant, San Diego, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	837
How learning to read and write has been positioned and framed in the past	838
A historical overview of the different ways of framing reading and writing education	838
Framing reading and writing education from a psychological viewpoint	838
An emergence of cultural forces and individual participants	838
Implications for pedagogy from the psychological viewpoint	839
A changing paradigm	839
The emergence of constructivism	840
The emergence of a linguistic perspective	840
Extending psycholinguistic perspectives in educational research	841
The process of learning language	841
Extending the socio-psycholinguistic viewpoint for classroom practice	841
Implications for pedagogy from the linguistic perspective	841
Where the education profession finds itself today	842
A shift in thinking	842
The psychological experimentation model	842
Dramatic changes in understandings of the relationship between conceptual metaphors and discourse	843
Going forward: embracing a biological-evolutionary perspective	844
So how does biomimicry connect to educational theory?	844
How is literacy and language learning viewed from a biological-evolutionary perspective?	845
How does this shift in perspective help resolve our human challenge of learning to read and write?	845
Final thoughts on the future of educational research	846
References	846

Reading has a history.

—Robert Darnton, *The Kiss of Lamourette: Reflections in Cultural History*, 1990

History is tangled, messy, contradictory. But is where we are.

—Eamon Duffy, *Faith of Our Fathers: Reflections on Catholic Tradition*, 2006

The question is not how the language arts should be brought together in the learner's mind

but why they should ever be separated.

—Frank Smith, *Reading Like a Writer*, 1983

Introduction

Education is shaped by history, research, and politics, as is true of any domain of intellectual inquiry. In other words, education evolves. It is influenced by academic trends, cultural forces, and individual participants. This is especially true when it comes to identifying the best pedagogy for teaching literacy.

This chapter has two broad aims. First, it describes how the field of literacy education has been framed over the past hundred years or so. This chapter, however, is not an exhaustive anthology of research or researchers in the field of literacy and language development. Rather, it reflects the authors' historical synthesis of multiple threads of research. Second, it foreshadows a shift for the domain of learning, arguing that a biological-evolutionary perspective is needed.

How learning to read and write has been positioned and framed in the past

As American cultural historian [Darnton \(2014\)](#) points out, how we learn to read and write has been shaped by “cultural configurations which have varied enormously over time.” One such cultural configuration is the presence or absence of literate experiences for the majority. Another is the restriction of literacy to certain privileged classes. These influence individual lives as well as society at large in familiar as well as yet undreamed of ways. The justification of these cultural configurations has spilled over into academia and continues to shape our current educational landscape.

Australian literacy educator [Cambourne \(1988\)](#) suggests that those who are unable to access written forms of language, either as readers or writers, experience what he calls “textual alienation”; this essentially leaves them “depowered” in a society that values literacy. He concludes “this textual alienation is a function of how literacy has been taught.”

These core issues reveal that what counts as literacy, and who should have access to it, matters. And these basic parameters have colored how literacy has been researched and discussed in the professional literature for decades. Furthermore, one’s personal, often implicit or unconscious, definition of literacy determines how one approaches the teaching, and hence, the learning of literacy. These differing views of what defines literacy have influenced the relationship between reading and writing through the ages.

Historically—from the cuneiform writing system that appeared in Mesopotamia circa 3200 BC up to our present-day alphabetic system—written language recorded everyday financial transactions, governmental or legal business, or religious dogma ([Dragataki, 2021](#)). As early writing became more specialized and uniform, the profession of scribe emerged. This specialization was highly regarded and a possibility for only the most intellectually able citizens. Similarly, reading was regarded as the domain of a select few and was typically restricted to those in certain professions or from classes of people: elite professionals, wealthy citizens, or the most powerful priests of the time. But since the first uses of written language, reading and writing have always been connected. The earliest scribes’ writing, incisions on pieces of clay, created readers who would simultaneously “decipher them, recognize their meaning, give them voice” ([Manguel, 1996](#)). Just as today, writing requires a reader to bring it to life, whether the recording medium is a clay tablet or digital device.

Historically, however, composition, as we think of writing today, wasn’t encompassed within these early scribes’ definition of writing ([Manguel, 1996](#)). And reading as a critical skill necessary for a healthy, creative society and functioning democracy is also a more recent development ([Dewey, 1916](#)). Academically and culturally throughout time, reading and writing have continued to be thought of as a specialty.

In the late 1800s and early 1900s in English-dominant cultures, the tradition of vigorous, often intense debate began in education about the best pedagogy for teaching reading and writing to the masses. The academic debates typically revolved around one set of either-or questions: Is reading a perceptual skill, through which readers perceive a word in print as one unit? Or is reading about decoding and blending the sounds of the letters in a word?

A historical overview of the different ways of framing reading and writing education

As with all memories, the human brain organizes events as a linear series of episodes ([Tulving, 1983](#)). What follows in this overview is a mostly linear synthesis of two distinct historical schools of thought linking to reading and writing education. It is an attempt to make sense of a complex collection of theories. It should not be regarded as the only truth.

Framing reading and writing education from a psychological viewpoint

In the late 1800s, educators regarded the emerging field of psychology as education’s theoretical home. The alignment between these domains seemed a natural fit because learning was (and still is) considered the purpose of education; it is also a key topic for study in the field of psychology. Psychology regarded learning as “a mental activity” ([Lund, 1927](#)) devoted to “the understanding of mental activity.” An alignment of the fields flourished as educational pedagogy came to be viewed through the lens of psychology. The language (i.e., Discourse) of psychology became the platform educators could use to communicate and/or discuss theoretical issues and concerns about learning, an academic construct that continues today.

In its infancy, the field of psychology sought to apply traditional hypothesis-testing models of research, much like the long-established linked sciences of physics and chemistry. These familiar models emphasized “either-or” experimental research designs ([Tracey and Morrow, 2017](#)), and as psychology sought to be perceived as an established science, it took up this either-or paradigm. The alignment of education and psychology led education to position the early reading debates as an either-or proposition: the Method A vs. Method B dichotomy between look-say and phonics.

Education relied on psychological theories to support both sides of the reading debate. The look-say philosophy in reading instruction intersected with psychology’s Gestalt theory, which argues that people perceive objects as “wholes” ([Kohler, 1929](#)). Another more dominant psychological learning theory of the time, behaviorism, reinforced both look-say and phonics. At the core of psychology’s behaviorism ([Schneider and Morris, 1987](#)) was the belief that learning was basically habit formation; by extension, the reinforcement of correct behaviors and punishment of incorrect behaviors helped ensure that learners developed the right habits. This way of thinking about learning eventually extended to teaching of reading and writing.

An emergence of cultural forces and individual participants

In 1896, American philosopher, psychologist, and educator John Dewey founded the University of Chicago Laboratory Schools, applying and testing his progressive ideas. His views of progressive education (1916) in the United States were a precursor to today’s

constructivism. Dewey's innovative beliefs describing the role of public education in supporting a thriving democracy were revolutionary. His work represented a significant cultural shift in how public education was viewed in American society.

An example of cultural forces influencing the educational field came with World Wars I and II in the 20th century. Recruits numbering in the millions were assessed on their intelligence and other abilities to qualify them for military service and to further establish their fitness for specialized training (DuBois, 1970). Large numbers of military recruits exhibited inadequate literacy abilities, reading in particular. This recognition of low reading abilities was naturally concerning and subsequently filtered down to the school system, continuing to fuel the academic debate between the two dominant methods of teaching reading: look-say and phonics.

Several well-known individual participants played a significant role in this debate. In 1955, American educator Rudolph Flesch wrote *Why Johnny Can't Read—and What You Can Do About It* (Flesch, 1955), taking a strong stance that teaching children to sound out words using rules was the best way to teach reading. The popularity of this book within and beyond the educational world led to the phonics method of reading instruction becoming the predominant method of teaching.

In the 1960s, American educator Jeanne Chall published her groundbreaking book *Learning to Read: The Great Debate* (Chall, 1967). Chall believed that effective reading was contingent upon accurate word recognition, which was best achieved through phonic decoding. She defined the two approaches to teaching reading as code-based (phonics) and meaning-based (look-say). Although Chall used the phrase meaning-based, her definition and usage differed. She used meaning-based to refer to word recognition. Today, the phrase relates to comprehension.

In her research, Chall used concepts, theories, and the discourse of psychology to test which of the two broad approaches to teaching reading, look-say or phonics, was best. After analyzing results from empirical researchers with a variety of test scores and treatments and a reliance on a Method A and B methodology, Chall determined that phonics produced the best results. Chall's research continued a tradition of defining phonics (i.e., decoding to sound) as a basic and hierarchical set of subskills. This belief contributed to a view of reading as a phonics-first process and reinforced a behaviorist orientation in education that is still relevant today.

Implications for pedagogy from the psychological viewpoint

In the decades from the 1890s through the 1970s, writing education was linked to spelling and grammar and very little, if any, attention was paid to composition. The teaching of grammar focused on parts of speech, with students working to identify the right kind of words (e.g., nouns and adjectives, verbs and adverbs) to produce an acceptable text. The relationship between reading and writing was not considered; therefore, reading and writing were taught as separate subjects.

During these decades, comprehension, fluency, and other aspects of reading were regarded as outcomes of accurate phonic decoding. If a reader could decode the words accurately, it was assumed that reader could understand what a text was about.

Because there were two dominant theories of how reading was learned, two instructional methods emerged (Holdaway, 1979). In classroom practice, however, both methods often intertwined and overlapped. Look-say advocates argued children should learn the pattern of the whole word; therefore, instruction consisted of drills using flash cards, often followed up with worksheets. Phonics advocates argued that children should learn to analyze words and their component sounds, hence a reliance on whole class instruction in phonics. This mode of instruction involves identifying letters and corresponding sounds, segmenting and blending sounds in words, and repetitively chanting how words are spelled ("C-A-T spells CAT"), often culminating in the completion of worksheets. In both the look-say and phonics methods, worksheets were used to provide individual, repetitive practice of isolated skills.

The teaching methods that were predominant in this era followed the pattern "teach-practice-test," and they were clearly underpinned by a learning theory that "came straight out of the experimental psychologist's laboratory" (Smith, 1986). Writing instruction was taught separately and consisted of exercises designed to identify correct usage in tandem with grammar, spelling, and punctuation.

If comprehension was addressed at all, it was through questions asked by the teacher and answered by the students after reading a passage together. The teacher's questions were typically about specific details in the passage, reinforcing the idea that comprehension was an outcome of correctly decoding the words.

Assessment during these years was based on the degree to which a student's responses on a worksheet correlated to the questions posed—and whether the answers were deemed correct. The emphasis was on "getting it right." Mistakes were unacceptable and had to be ruthlessly rooted out in case they became "permanently fixed in the child's repertoire" (Cambourne, 1988).

A changing paradigm

Irish historian Eamon Duffy's memorable quote about history being "tangled, messy, contradictory" (2006) seems pertinent when reflecting on the past five decades of research into literacy and language learning. In the previous sections of this chapter, educational research was characterized by an emphasis on the design of experimental psychological research. The next section integrates research on the social aspects of language and literacy learning which promoted the use of naturalistic inquiry. This section synthesizes complex, overlapping lines of research and theory from multiple fields of study.

The emergence of constructivism

The history of constructivism is wide and varied (Tracey and Morrow, 2017). The discussion of constructivism that follows begins with Swiss biologist and psychologist Jean Piaget, “the great pioneer of constructivist theory of knowing” (Glaserfeld, 1990), who, in the early and mid-decades of the 20th century, studied cognitive development in young children. Applying some principles of naturalistic inquiry, Piaget focused on learners and the qualitative development of their understandings. His concepts of assimilation (from which learners respond to new events based on existing schema), and accommodation (whereby learners modify or create schema based on new events) articulated a model of active learning. Piaget also introduced the concept of readiness, by which children progress through stages of intellectual development.

Around the same time as Flesch’s and Chall’s writings, developmental psychologist Lev S. Vygotsky’s seminal text, *Thought and Language* (1962), was translated into English, informing and reshaping Western educational trends. Vygotsky connected speech (both silent inner speech and spoken language) to the development of mental concepts and cognitive awareness; in addition, the connection fueled developmental and educational psychology’s embrace of social constructivism. It also informed his colleague Alexander Luria’s research into the regulatory role of speech (Tupper, 1999). Vygotsky’s theories of language development and cognitive processing described the potential of human cognitive development. Vygotsky gave the field of education, particularly the area of constructivist learning, one of its most enduring and influential concepts: the zone of proximal development. The zone of proximal development is defined as the distance between what a learner can do on his or her own and what can be accomplished with the support and assistance of someone who is more knowledgeable. This concept underpins several contemporary learning theories, including the gradual release of responsibility (Pearson and Gallagher, 1983).

The emergence of a linguistic perspective

Throughout the first half of the 20th century, the dominant theory of learning was based on behavioristic modeling, pulled together into a coherent theory by B. F. Skinner and John Watson (Schneider and Morris, 1987). This theory proposed that the focus of learning research should be on observable behaviors rather than internal emotions and processes; internal processing was thought to be beyond the scope of scientific inquiry. Skinner and his colleagues and followers believed humans acquired language through direct instruction and reinforcement (Skinner, 1957; Winokur, 1976; McLaughlin, 2006). Mid-century, advances in linguistic theory challenged this thinking.

In the 1950s and early 1960s, American psychologist George Miller, the father of the psycholinguistic perspective of learning, linked language and cognition to analyze language creation and use (Miller, 1951). He also pioneered an acceptance of mental processes as proper subject matter for research. Miller’s research on working memory (1956) revealed that the brain is prewired to take in and remember up to “seven plus or minus two” bits of information at a time. This research had implications for both the size and quantity of information bits built into a learning event. His research also brought meaning into the discussion of learning, acknowledging that language is more than just words.

Noam Chomsky, an American researcher who redirected academic trends in linguistics, challenged this behaviorist view in his watershed critique, “A review of B. F. Skinner’s verbal behavior” (1959). In his response to behaviorism, Chomsky challenged the notion that internal cerebral processing could be ignored, leading him to develop his theory to explain the generative nature of all human languages. For speakers to generate sentences they had never heard before, Chomsky insisted humans possessed an innate knowledge of how any language works, a Universal Grammar. Because no learning theory at the time explained how children, who are otherwise cognitively immature, could learn something as complex as language so easily and early, Chomsky argued humans must be born with this innate knowledge. This viewpoint challenged the academic beliefs of linguistics of the time. Although multiple scholars have challenged Chomsky’s view of an innate language-acquisition device (Everett, 2012; Deacon, 1998), Chomsky offered a stark contrast to the thinking that had prevailed (Ellis and Bloch, 2021).

American researcher Ken Goodman (1965, 1967) also challenged the academic trends of the time. His psycholinguistic theory of reading redefined Chall’s term “meaning-based” from word-recognition to a process of continuous construction of meaning by learners. Understandings emerge as readers predict and then confirm or reject those predictions. Goodman defined this as a “psycholinguistic guessing game” (1967). By this, he meant that readers use multiple cues—semantic, syntactic, and graphophonic—to anticipate and hypothesize (i.e., “guess”) as they construct the author’s intent based on information found within the text and within the reader. This thinking was contemporaneous with researchers who argued that language is an integral part of literacy education. The influential researchers who fueled this trend included Rosenblatt (1978), whose transactional theory of reading emphasized the cultural experiences and language that readers bring to reading.

Other researchers in this emerging group included American sociolinguists Labov (1968), Hymes (1972), and Shuy (1976). They proposed the concept of language performance through social interactions; this idea that language has a social component contrasted with Noam Chomsky’s view of linguistic competence. These researchers believed that studying grammar rules does not help learners understand how language is used in the real world, where “meanings may be won by the sweat of the brow, and communication achieved in labor” (Hymes, 1972). Until the 1970s, competence—the internal set of rules that Chomsky’s transformational grammar identified—was considered the only legitimate area of linguistic study. But, by looking at the culture and context in which language is constructed and used, these researchers connected the purpose and production of language while expanding the field in ways that still reverberate. It is also significant that because each of these sociolinguistic researchers had background in anthropology, they applied the canons of anthropological research to the study of linguistic performance. This legitimized the movement toward naturalistic inquiry as a valid research methodology in education.

Extending psycholinguistic perspectives in educational research

By extending his research on the social and cultural dynamic of language development discussed previously, Labov (1972) was instrumental in examining the linguistic features of Black English Vernacular. His groundbreaking research of this form of English, with its various dialects, found it to be as complex as other forms and important for the culture in which it is used.

Contemporaneously with these other researchers, Canadian psycholinguist Smith (1971) introduced the notion of learning as a process. Smith's idea of reading and writing as linked processes contradicted the academic trend of the time, which viewed reading and writing as a set of sequential and hierarchical skills to be acquired by a learner. His theory influenced how educational professionals thought about literacy and also changed how researchers designed educational studies.

British linguist Halliday (1973, 1975, 1978) expanded the idea of language as a social meaning-making process while detailing how language supports various forms of communication. Halliday's research showed that function (the purpose) preceded form (the structure) in language development, challenging Chomsky's notion that form preceded function. In addition, Halliday's ideas about language development emphasized the integrated nature of learning, recognizing we "learn language, learn through language, and learn about language simultaneously" (1979, 1980, 2003).

The process of learning language

Several key researchers continued to study how young children learn the oral version of their native tongue. American psychologists and educators Carroll (1964), Brown (1973), and others explored and wrote about how children learn to talk. By observing children and documenting language development over time for individual children, their research emphasized the social dimension in literacy and language learning. American linguist and educator Carole Chomsky's research of how children learn to control the passive voice further documented how language develops over time, providing support of learning as a process (1969, 1976).

Australian educator Brian Cambourne's Conditions of Learning theory (1988) built on the existing research of native language learning. His naturalistic research explored conditions under which young children learn to talk, producing a grounded theory that explained the physical, social, emotional, and ecological aspects of first-language learning. Cambourne's framework articulated how learning occurs in a natural environment where learning is based on the interactions between a novice and more knowledgeable others with whom the novice has developed strong, bonded relationships. Cambourne proposed these same conditions could be simulated in classrooms to support language and literacy learning.

Continuing to expand the body of research into the social nature of learning, Tomasello (1999, 2003), an American developmental and comparative psychologist and linguist, interpreted social and contextual influences on language for his theory of usage-based linguistics. His social-pragmatic theory proposes that language is developed through interactions with adult speakers. During these exchanges, the child's cognitive skills enable him or her to assess the intentions of the speaker, learn to refine patterns of grammatical use over time. Language, according to Tomasello (2008), is "cognition packaged for purposes of interpersonal communication."

This research into what cognitive scientists and linguists believe is the most complex learning a human ever does had implications for learning in other symbol systems (e.g., writing, art).

Extending the socio-psycholinguistic viewpoint for classroom practice

American educators Harste and Burke (1978) introduced the term "socio-psycholinguistic" to integrate the cognitive work described by psycholinguists Miller and Goodman with the social aspects explored by Labov, Halliday, and other sociolinguist colleagues. Specifically, Harste and Burke studied the ecological contexts in which children learned to understand and respond to symbols. They observed children's initial symbolic behavior in writing, noticing how children used what they saw in their home and cultural environments to make sense of written and spoken language. In particular, Harste's "education as inquiry" (2003) theory proposes that one's ability to learn—including learning to read in all symbol systems—is social in nature.

In her research, a New Zealand developmental psychologist and educator Marie Clay continued a focus on the earliest years of literacy and language learning (1982, 1985, 1991). Her literacy processing theory echoes other socio-psycholinguistic theories. Rather than perceiving children as empty vessels to be filled, Clay's theory values the role that prior knowledge can play in pedagogy, a belief that continues to influence classroom practices in significant ways.

Implications for pedagogy from the linguistic perspective

By the 1970s and '80s, the code-based vs. meaning-based debate had morphed into another dichotomy: direct instruction vs. whole language. Direct instruction (not to be confused with explicit instruction, which means teaching with clarity) is a philosophy and series of teaching methods based on the belief that learning occurs only when the learner is told directly what is to be learned and how to accomplish the learning (Bereiter and Engelmann, 1966). Direct instruction methods of teaching broke content to be learned into hierarchical sequences, with each part being mastered without mistakes. Once the parts were mastered, they were added together to make an overall whole. Whole language proponents believe that learners must be given an opportunity to construct meaning by experiencing demonstrations of language in use. These demonstrations provide an overview of how the component parts of what is being learned fit together as a whole (Goodman, 1986).

Most classroom instruction in the 1970s and beyond reflected a psychological discourse. For example, the ideas of reinforcement being a necessary component of learning—the notion that "practice makes perfect"—and the intent to extinguish errors before they became permanent promoted an objectivist pedagogy. The Initiate-Respond-Evaluate (IRE) pattern of classroom discourse (Cazden, 1988) is a hallmark of this pedagogy. With this method of interaction, a teacher initiates an exchange with students by asking

a question. Students indicate they know the answer by raising a hand. The teacher selects a student to respond. Then the teacher evaluates whether the given answer is right or wrong.

For those who embraced a constructivist perspective through the 1970s and '80s, however, a new discourse emerged. Students were encouraged to collaborate with one another, and texts were viewed as having multiple interpretations rather than a singular (right or wrong) main idea. Terms like "mistakes" and "errors" instead became "approximations" or "miscues"—and "motivation" became "engagement." During this era, reading and writing began to be thought of as a process. For a significant number of practitioners, this change in discourse precipitated a notable shift in the pedagogy for teaching reading and writing.

Where the education profession finds itself today

From the 1980s through the first decade of the 2000s, much theorizing that emphasized a phonics-first approach occurred in educational and psychological circles (Adams, 1990; Lyon and Chhabra, 2004). American psychologist Ehri (1999) and American educators Foorman and Moats (2004) added to the body of research endorsing a skills-based approach to literacy learning. Another American cognitive and developmental psychologist and reading program developer Marilyn Jager Adams summarized the research of the time in the influential book *Beginning to Read: Thinking and Learning About Print* (1990), supporting a phonics-first view of how literacy was taught. These researchers used a psychological, hypothesis-testing paradigm that supported a phonics-first pedagogy to explore cause-effect relationships in learning to read.

Since the dawn of the 21st century, the advancement of technologies such as the MRI has made it possible to explore more deeply how the brain functions during language and literacy experiences. Early brain-to-learning researchers studied MRI scans to determine which parts of the brain are activated during certain activities; this approach resulted in a linear view of how the brain processes and responds to stimuli (Dehaene, 2010; Abadzi, 2016; Castles et al., 2018). Most research of this type continues to use an operational definition which represents reading as decoding of isolated words, rather than define reading as a constructive, meaning-making process.

Recent studies, however, provide a focus on how the brain makes meaning of stimuli it receives (Cobb, 2020a,b; Ellis and Bloch, 2021). When one defines reading as meaning making, rather than simply as decoding print, the type and design of neuroscientific research that can inform pedagogy expands significantly. Researchers, including Hirsh-Pasek et al. (2018), Romeo et al. (2018a,b), and Piazza et al. (2020), have focused on how connections that support language development take shape within the brain. Their findings show language learning to be both socially and contextually situated, reflecting the necessity of interpersonal and interactive experiences for it to occur. Other research concerning reading development details how multiple brain regions and neural pathways that connect and coordinate those regions interact (Kweldju, 2015). Neuroscientists Buzsáki (2019) and Cobb (2020a) conclude that the brain doesn't only represent information but also constructs it.

The distinction as to how reading is defined (i.e., representing information or constructing understanding) and how the brain makes meaning was explored in *Reading—The Grand Illusion: How and Why People Make Sense of Print* (2016), in which American educator Kenneth S. Goodman, American linguist Peter H. Fries, and American neuroscientist Steven L. Strauss write about the remarkable abilities of humans to produce and use language to navigate their world. As quoted earlier in this chapter, Darnton (2014) and Cambourne (1988) believe one's definition of literacy matters. Goodman, Fries, and Strauss, researchers from diverse fields, share this belief. Their research highlights how readers use their eyes, brain, and language to make meaning. It's no surprise that when professionals in different fields collaborate, they build on one another's thinking and break down walls that divide them, thereby exemplifying how meaning is constructed. This collaboration among different disciplines using a meaning-based definition of literacy has implications for moving the field of literacy education forward.

Yet, at the conclusion of this historical overview of literacy in the context of education, debates continue to rage about what distinguishes scientifically-based and evidence-based practice. American educational researcher Gabriel (2018) succinctly captures this state of affairs: "The sloppy, mudslinging nature of these debates has led to confusion, distrust and a tribe-like affiliation with single approaches among practitioners, researchers and policymakers." This present-day cauldron of competing and overlapping philosophies continues to bubble in the background of literacy and language learning. So where do we go from here?

A shift in thinking

Until this point, this chapter has presented a chronological and philosophical tour of the history of the profession. It explained philosophies and dichotomies that have influenced educational practice for the past 50 years. The section that follows explains why a shift in thinking and research about learning to read and write is necessary and why a biological-evolutionary perspective has merits as a path forward.

The psychological experimentation model

The varied psychological and linguistic perspectives that have been applied to reading and writing have existed in silos. The factor that unites many of these disparate theories is their reliance on psychological definitions of learning and psychological interpretations.

What do these psychological connections mean and why does it matter?

In psychological parlance, learning was seen as a “change in behavior” caused by external environmental factors occurring just before any such changes (Lund, 1927). This definition influenced the dominant culture of both psychological and educational research. This culture of both domains was characterized by an emphasis on controlling and testing effects of external factors such as type of motivation, duration of exposure to different external factors, or sequencing of the presentation of material to be learned. This view of learning condoned, promoted, and emphasized establishing cause-effect relationships between whatever factors were being tested.

Why is that problematic? In its earliest manifestations, psychology struggled with its place among the established sciences, including physics, chemistry, and biology. This struggle was caused by ignoring the process through which the hard sciences became credible and trustworthy. That process begins with naturalistic observation to identify and document how phenomena exist and behave in their natural environment. Mini-theories are proposed to explain observations that emerge (Kuhn, 1962; Popper, 1972; Lakatos and Feyerabend, 1999). These mini-theories are then tested in controlled experiments. If a mini-theory doesn’t hold up to scrutiny, it is discarded. Over time, a convergence toward an explanatory, overarching theory occurs that represents the best current thinking but will continue to be tested and evolve. For example, the periodic table in chemistry derived from observations to determine what the basic elements were and how to systematize them. Scientists throughout history engaged in rigorous debate as they proposed, tested, and accepted or rejected ideas and theories. They sought to find patterns to establish consistency, eventually arriving at a widely accepted consensus of understanding. Early psychologists leapfrogged this historical, observational stage. They proceeded directly into experimental design and hypothesis testing without first identifying and describing the way the phenomena they were seeking to understand occurs in an experimenter-free environment (Heft, 2001).

Ultimately, psychology—in its early development as a field of scientific study—didn’t strive to identify and apply their own set of overarching concepts as other scientific domains had. The abundance of extant (often conflicting) learning theories (Culatta and Kearsley, 2021) that have continually emerged from experimental psychology are a result of this leapfrogging. Such a lengthy list of competing theories provides a mute but powerful testament that something is seriously wrong with the way psychology (and therefore education) has conceptualized and researched learning over the last one hundred and twenty or so years.

As American psychologist Harry Heft writes: “Has there been a moment since its formal founding in the late 19th century when experimental psychology was not in a state of theoretical conflict? Select any historical point during its first 120 years and you will find psychologists embroiled in some theoretical squabble” (Heft, 2001). A closer look at the history of psychology shows that a majority of these theoretical squabbles are either directly or indirectly related to the umbrella concept of learning. If that basic overarching concept is so conflicted within education’s chosen theoretical partner, no wonder the derived theories of learning (including learning to read and write) are also in conflict.

If fact, Canadian psychologist Frank Smith claimed in an essay that “education backed the wrong horse when it backed psychology”; he argued that “psychology has never been comfortable with learning” (Smith, 1983). Psychological experiments at the time of Smith’s assertion emphasized the learning and forgetting of 3-letter nonsense syllables rather than assessing learning as a meaning-making process. The belief was that because the nonsense syllables were unknown to the test subjects, the test results could be attributed solely to the variables being tested. Learning to read with meaning and the variables that influence meaning-making were considered to be uncontrollable factors; as such, they were eliminated from the experimental design. Although Smith stated the obvious by signaling these types of experiments were not about learning, they nevertheless informed pedagogy.

Dramatic changes in understandings of the relationship between conceptual metaphors and discourse

In academia, Discourse is a theory which describes the various relationships within a particular setting. It includes the thinking and behavior group members expect others to adopt and use when engaged in the group’s “business.” Conceptual metaphors, or mental representations, are embedded in all discourse (Fauconnier and Turner, 2002; Lakoff, 2004). Conceptual metaphors guide how people think about the world and determine their language and behaviors (Lakoff and Johnson, 2003). As British historian Schaffner (2021) stated: “Metaphors matter. Far from being mere ornamental flourishes that render our speech more vivid, they create important conceptual bridges between fields The imagery we use is never neutral—it reflects deeper cultural assumptions, often about agency and personal responsibility. This has consequences.”

How one defines learning and the research designed to gauge it reflects the metaphor of learning one embraces. Holding tightly to metaphors, and the discourse that arises from their presence, imposes boundaries and can end up “limiting what and how we can think” (Cobb, 2020b). The great debate is a powerful example of how metaphors influence thinking and behavior.

And, for educators who have been immersed in psychological discourse for all their professional lives, another conceptual metaphor influencing their practice is the acquisition metaphor. In this nonconscious metaphor, learning and knowledge are thought of as some kind of reified stuff. And, like other stuff, it can “be measured and weighed as though it has mass, weight, and heft” (Crouch and Cambourne, 2020). It follows that it can also be transferred from one person to another and acquired by someone through a process of moving the stuff from one place (e.g., an expert’s mind) to another by an act of transmission (using language as the medium). It can be discussed, communicated, and thought about as the static end product of an act of learning (e.g., comprehension is seen as something that occurs after one reads rather than as a process of meaning-making).

Viewing the end product of learning as knowledge continually reinforces this underlying acquisition metaphor, one in which learning and knowledge is seen as stuff and teaching as the delivery of this stuff. For example, a teacher’s language that labels responses as right or wrong is an indicator of the heft of the ideas; signaling to “get this information into your head” reinforces

the metaphor of learning as something transmitted. Australian educator [Cambourne \(2015\)](#) identified this language and associated behaviors as the Discourse of Acquisition and argued it is the dominant discourse in education and psychology today.

An alternative to this view of learning is the Discourse of Meaning-Making ([Langer, 2011](#); [Cambourne, 2016](#)). Instead of a discourse that subconsciously implies that learning is stuff to be transmitted, a Discourse of Meaning-Making frames learning as a process of using linguistic and other symbols to engage in continuous cycles of constructing, deconstructing, and reconstructing these meanings through interactions with others. In other words, meaning is constructed using whatever medium is available to us—print, graphic, oral, signing, gestures, etc.—as we communicate with and think alongside other learners. This need to communicate is both innate and generative. Meanings are continually being adjusted and updated as we learn more. In a Discourse of Meaning-Making, learning is characterized by, rather than defined as, an ever-changing, dynamic mix of knowledge, understandings, feelings, values, and skills linking to what is being learned. In addition, a Discourse of Meaning-Making embodies an intrinsic dynamic that ensures we learn to exchange and refine meanings with others using symbols.

While parsing these different discourses may seem a trivial exercise in semantics to some, the ramifications are compelling. This reconceptualizing of learning forces a significant shift in thinking about learning, with implications for both research and pedagogy. This innate need to communicate with others using symbols positions learners as active and agentive in the learning process. The ability to use symbols combined with an innate need to communicate using those symbols are two linked human characteristics, unique to the species ([Deacon, 1998](#); [Fauconnier and Turner, 2002](#); [Tomasello, 2003](#)). Through the ages, all cultures have struggled to come up with the best way to learn this symbolic behavior. And this cultural issue is a distinctly human challenge.

Going forward: embracing a biological-evolutionary perspective

For the past century, the social sciences—especially psychology, education, and other cognate domains—have been grappling with this human challenge of learning to read and write. Interpreting this issue from a psychological perspective has hindered the ability to identify or agree on a definitive theory of learning to read and write. Biomimicry, a relatively new branch of biological-evolutionary science, represents an alternative approach, one that focuses on the science and art of emulating nature’s best biological ideas to solve human problems ([Biomimicry Institute, 2021](#), <https://biomimicry.org/>). This stance offers social scientists a fresh perspective for exploring literacy and language learning.

Biomimicry ([Benyus, 1997](#)) relies on two fundamental principles. First, evolution has already solved many of the human challenges with which science is currently grappling. Second, advocates of biomimicry insist that, to solve a human problem, it must be interpreted from a biological or natural perspective. “Biomimicry asks us to learn from and then emulate the intelligence embedded in 3.8 billion years of learning, using nature’s forms, processes, and ecosystems to create more sustainable designs” ([Fig. 1](#)) ([Biomimicry Institute](#), as cited in [Crouch and Cambourne, 2020](#)).

So how does biomimicry connect to educational theory?

At the core of biomimicry is the belief that its principles should be applied only to issues that are important for human survival. Being literate is both important to, and significant for, our survival as a species for two reasons. First, being literate is necessary for maintaining democracy as we know it. A healthy, creative society and a functioning democracy require critical thinkers and critical analysts. Given the relationship between literacy and democracy, it is imperative for schools to produce graduates who are highly effective readers and writers. Second, literacy is significant for human survival because of the vital need to address and resolve the dominant global ecological, social, economic, scientific, ethical, and moral issues related to sustaining our species on a planet that has finite resources. It follows that regular cohorts of highly literate graduates are essential for creating an ongoing international supply of talented researchers, theory builders, thinkers, problem finders, and problem solvers—and especially well-equipped

Among the multitude of designs that have evolved in nature through the ages is the pervasive ecosystem of termite mounds. Termites in the Northern Territory of Australia must keep their eggs within a certain range in the very hot climate. Over billions of years, these insects have worked out how to design their six-foot mounds to survive in the hostile conditions. By incorporating strategically placed openings and internal structures, the termites maintain temperatures that ensure their survival. Engineers have studied the termite mounds and applied similar elements of climate control in designs for human structures. This is just one example of many human quandaries that have been addressed by applying principles of biomimicry.

Fig. 1 An example of biomimicry in action.

teachers—across all domains of knowledge and expertise. Solving this challenge becomes even more urgent because of the exponential rate at which information is increasing and the skills required of readers and writers who must quickly assess the truth-value of the textual messages that bombard them.

How is literacy and language learning viewed from a biological-evolutionary perspective?

Psychological discourse tends to define and frame learning as a change in an animate organism's behavior "caused" by contiguously occurring external factors and events. Biological discourse, however, is radically different. Biological-Evolutionary discourse frames human learning as a unique form of learning, one that is available only to our species (Deacon, 1998; Iacoboni et al., 1999; Fuster et al., 2000; Donald, 2001; Ramachandran, 2004; Feldman, 2008; Tomasello, 1999, 2003, 2008). Framing human learning this way triggers the question: Why is this kind of learning and knowing so unique for our species? Addressing this question has far-reaching implications for resolving the human challenge of learning to read and write.

Evolutionary biologists argue learning and knowing using symbols is ultimately what distinguishes *Homo sapiens* from all other forms of life on Earth. Through the long process of evolution, learning and knowing symbols came to define our species; therefore, it must be a form of species-survival behavior (Fauconnier and Turner, 2002). And not only has learning and knowing using symbols made us the dominant species on the planet, but the fact that this dominance has not resulted from superior physical strength, speed, size, teeth, or claws further strengthens the argument that this ability is vital for the survival of our species.

In addition, the selection of this ability to use symbols provides our species with two unique, species-specific, cognitive capabilities. First, it allows us to create complex knowledge (or meanings) using abstract symbol systems. Without this ability, we would not have mathematics or science or poetry, for example. Second, we can apply this knowledge (or these meanings) to the everyday challenges of the survival of our species. Not only are we the only species of living organism on the planet that can construct abstract meaning using a diverse range of symbol systems, but we can also store these meanings in memory, revisit them, manipulate them, extend them, refine them, and build on them. As a consequence of these processes, we can share and communicate what we've constructed with other members of the species to solve problems, whether they involve mundane challenges or world-changing issues.

This evolutionary selection of making meaning using abstract symbols is supported by much research (Deacon, 1998; Iacoboni et al., 1999; Fuster et al., 2000; Donald, 2001; Ramachandran, 2004; Feldman, 2008; Tomasello, 1999, 2003, 2008) culminating in the same conclusion: Until our ancestors developed this way of using abstract symbols, communicating in this way to identify and solve issues of survival was a new, untried form of cognition. As American cognitive scientists and linguists Fauconnier and Turner (2002) note: "That every human being can do it, but no member of any other species can do it should tip us off immediately to the evolutionary development that was needed to make such feats possible. Only really big brains connected in special ways, and doing a lot of work as trained by the cultures, can even begin to handle these feats."

Further support for the aforementioned conclusions can be found in archeological and evolutionary records that show this previously untried form of cognition created and unleashed new forms of knowing, learning, thinking, and communicating. This not only made humans unique among all other forms of life, but has had profound impact on the planet, including in some possibly irreversible ways (Deacon, 1998; Iacoboni et al., 1999; Fuster et al., 2000; Donald, 2001; Ramachandran, 2004; Feldman, 2008; Tomasello, 1999, 2003, 2008). The ability to construct and communicate abstract meanings using symbols offered our biological ancestors a unique and powerful tool for survival in a hostile environment (Deacon, 1998; Donald, 2001). It follows that our descendants will need these same tools to address and resolve the multitudinous global ecological, social, economic, scientific, ethical, and moral issues they face.

Fauconnier and Turner (2002) summarize the process of developing this system of constructing and communicating messages this way:

Two million years ago, australopithecines, equipped with nonlinguistic ape-like mental abilities, struggled to assemble by fits and starts, an extremely crude symbolic system—fragile, difficult to learn, inefficient, slow, inflexible, and tied to ritual representation of social contracts like marriage. We would not have recognized it as language. But language then improved by two means. First, invented linguistic forms were subjected to a long process of selection. Generation after generation, the newborn brain deflected linguistic inventions it found uncongenial. The guessing abilities and intricate nonlinguistic biases of the newborn brain acted as filters on the products of linguistic invention. Today's languages are systems of linguistic forms that have survived. *The child's mind does not embody innate language structures. Rather the language has come to embody the predispositions of the child's mind* [emphasis added]

How does this shift in perspective help resolve our human challenge of learning to read and write?

The idea of language being structured by the human mind to communicate meanings with others—rather than the mind being structured by language imposed on it—is a seismic shift in cognitive science. If meaning is present for a learner, the learner will attempt to construct a form of language to express and communicate it to others. However, this language may differ from the language used in the classroom. This implies a meaning-centered pedagogy which honors children's first responses as approximations of the Discourse of Meaning-Making is fundamental for effectively teaching students to read and write. This perspective on pedagogy echoes observations from linguist Halliday (1979/1980; 2003), who said we "learn language, learn through language, and learn about language *simultaneously*" [emphasis added].

In alignment with Halliday, Biological-Evolutionary perspective equates the term “learning” and its cognates (e.g., “knowledge,” “understanding,” “comprehension,” “knowledge building”) with “meaning” and “meaning making”. Rather than use psychology’s “change in behavior” definition, a biological-evolutionary perspective explains learning as a process of making meaning using symbols. Meaning is “an unpredictable mix of personally constructed internal pictures, sounds, and feelings, which seem to be unleashed by, and closely related to, the range of symbol systems humans are constantly using and manipulating to create new topical and contextually relevant meanings that allow them make sense of the world” (Cambourne, 2015). This internal cognitive construction of the human mind helps us make sense of the world and makes communication with others possible; both are essential for the survival of the individual and the species. Biological and cultural evolution has selected the construction of meaning using this range of symbol systems as a species-survival trait for *Homo sapiens*. This is a radically different way of thinking about learning as *meanings constructed by the learner, not something done to a learner*.

Furthermore, the ability to control the dominant symbol system, the oral form of language, has been so successful as a species-survival mechanism that evolution has ensured that young members of the culture develop this ability in ways as fail-safe as possible. The principle of Occam’s razor, or “evolutionary parsimony” (Dawkins, 2004), states that the best scientific principle to explain complex issues is the simplest one. So, when applying this principle to biological-evolutionary theory, the same conditions for learning the oral form of language should also support learning to control the other symbol systems a culture values and uses, including the written form of language (i.e., reading and writing). Thus, the ecological, social, physical, and emotional conditions that support learning to talk provide a scientifically rigorous framework for a pedagogy that supports learning to read and write (Carroll, 1964; Cambourne, 1988; Brown, 1973; Tomasello, 1999, 2003, 2008).

Final thoughts on the future of educational research

In the future, to design credible and trustworthy research, educators should consider a biological-evolutionary perspective as a framework for research and pedagogy. A summary of key principles to guide educators includes:

- a focus on research that explains how systems work, rather than research that seeks to prove cause-effect relationships
- honoring the centrality of meaning-making in the processes of literacy and other learning
- respecting the natural and cultural conditions under which learning occurs in classrooms and beyond

As this chapter has detailed, our history of research in literacy development and implications for classroom instruction has been messy and complicated; nonetheless, biological-evolutionary theory has clarified the social, cultural, and physical factors that support the basis of scientifically derived learning-to-read-and-write pedagogy. This theory confronts the culturally determined disempowerment of learners through “textual alienation” referenced at the beginning of this chapter (Cambourne, 1988). With this new discourse based on biological-evolutionary theory, teaching decisions can fully support and promote learning as a meaning-making process using symbols. Simplifying the process of producing critical readers, writers, and thinkers may help ensure that democracy as we know it has a chance of surviving.

References

- Abadzi, H., 2016. Turning a molehill into a mountain? How reading curricula are failing the poor worldwide. *Prospects* 46, 319–334. <https://doi.org/10.1007/s11125-017-9394-9>.
- Adams, M.J., 1990. *Beginning to Read: Thinking and Learning About Print*. MIT Press, Cambridge, MA.
- Benyus, J.M., 1997. *Biomimicry: Innovations Inspired by Nature*. William Morrow, New York.
- Bereiter, C., Engelmann, S., 1966. *Teaching Disadvantaged Children in the Preschool*. Prentice-Hall, Englewood Cliffs, NJ.
- Biomimicry Institute, 2021. What is Biomimicry? viewed July 16, 2021. <https://biomimicry.org>.
- Brown, R., 1973. *A First Language: The Early Stages*. Harvard University Press, Cambridge, MA.
- Buzsáki, G., 2019. *The Brain From the inside Out*. Oxford University Press, Oxford, England.
- Cambourne, B.L., 1988. *The Whole Story: Natural Learning and the Acquisition of Literacy in the Classroom*. Scholastic, New York City, NY.
- Cambourne, B.L., 2015. Understanding writing and its relationship to reading. In: Turbill, J., Barton, G., Brock, C. (Eds.), *Teaching Writing in Today’s Classrooms: Looking Back to Look Forward*. Australian Literacy Educators’ Association, Sydney, AU, pp. 26–40.
- Cambourne, B.L., 2016. Reclaiming or reframing? Getting the right conceptual metaphor for thinking about early literacy learning. In: Meyer, R., Whitmore, K. (Eds.), *Reclaiming Early Childhood Literacies: Narratives of Hope, Power, and Vision*. Routledge, London, pp. 17–29.
- Carroll, J.B., 1964. Linguistics and the psychology of language. *Rev. Educ. Res.* 34, pp. 119–126.
- Castles, A., Rastle, K., Nation, K., 2018. Ending the reading wars: reading acquisition from novice to expert. *Psychol. Sci. Publ. Interest* 19, 5–51.
- Cazden, C.B., 1988. *Classroom Discourse: The Language of Teaching and Learning*. Heinemann, Portsmouth, NH.
- Chall, J.S., 1967. *Learning to Read: The Great Debate*. McGraw-Hill, New York.
- Chomsky, N., 1959. Review of Skinner’s verbal behavior. *Language* 35, 26–58.
- Chomsky, C., 1969. Linguistics and philosophy. In: Hook, S. (Ed.), *Language and Philosophy*. New York University Press, New York, pp. 17–54.
- Chomsky, C., 1976. Creativity and innovation in child language. *J. Educ.* 158, 12–24.
- Clay, M.M., 1982. *Observing Young Readers: Selected Papers*. Heinemann, Exeter, NH.
- Clay, M.M., 1985. *The Early Detection of Reading Difficulties: A Diagnostic Survey With Recovery Procedures*, third ed. Heinemann, Portsmouth, NH.
- Clay, M.M., 1991. *Becoming Literate: The Construction of Inner Control*. Heinemann, Portsmouth, NH.
- Cobb, M., 2020a. Why your brain is not a computer. *Guardian*. February 27, 2020. <https://www.theguardian.com/science/2020/feb/27/why-your-brain-is-not-a-computer-neuroscience-neural-networks-consciousness>.
- Cobb, M., 2020b. *The Idea of the Brain: The Past and Future of Neuroscience*. Hatchette Book Group, New York.
- Crouch, D., Cambourne, B., 2020. *Made for Learning: How Conditions of Learning Guide Teaching Decisions*. Richard C. Owen Publishers Katonah, NY.

- Culatta, R., Kearsley, G., 2021. Learning Theories. Viewed July 23, 2021. <https://www.instructionaldesign.org>.
- Darnton, R., 1990. *The Kiss of Lamourette: Reflections in Cultural History*. Norton, New York.
- Darnton, R., 2014. First steps toward a history of reading. *Aust. J. Fr. Stud.* 51, 152–177.
- Dawkins, R., 2004. *The Ancestor's Tale: A Pilgrimage to the Dawn of Life*. Mariner Books, Boston.
- Deacon, T.W., 1998. *The Symbolic Species: The Co-evolution of Language and the Brain*. Norton, New York.
- Dehaene, S., 2010. *Reading in the Brain: The New Science of How We Read*. Penguin Books, London.
- Dewey, J., 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan, New York, NY.
- Donald, M., 2001. *A Mind So Rare: The Evolution of Human Consciousness*. Norton, New York.
- Dragatakis, M., 2021. Cuneiform to Hieroglyphics: The Evolution of Western Alphabets. <https://www.thecollector.com/cuneiform-hieroglyphics-evolution-western-alphabets/>.
- DuBois, P.H., 1970. *A History of Psychological Testing*. Allyn & Bacon, Boston, MA.
- Duffy, E., 2006. *Faith of Our Fathers: Reflections on Catholic Tradition*. Bloomsbury, London.
- Ehri, L.C., 1999. Phases of development in learning to read words. In: Oakhill, J., Beard, R. (Eds.), *Reading Development and the Teaching of Reading: A Psychological Perspective*. Blackwell Science, Oxford, England, pp. 79–108.
- Ellis, G., Bloch, C., 2021. Neuroscience and Literacy: An Integrative View. *Transactions of the Royal Society of South Africa*. <https://doi.org/10.1080/0035919X.2021.1912848>.
- Everett, D.L., 2012. *Language: The Cultural Tool*. Pantheon Books, London.
- Fauconnier, G., Turner, M., 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. Basic Books, New York.
- Feldman, J.A., 2008. *From Molecule to Metaphor: A Neural Theory of Language*. MIT Press, Cambridge, MA.
- Flesch, R., 1955. *Why Johnny Can't Read—And What You Can Do About It*. Harper & Brothers, New York.
- Forman, B.R., Moats, L.C., 2004. Conditions for sustaining research-based practices in early reading instruction. *Remedial Special Educ.* 25, 51–60.
- Fuster, J.M., Bodner, M., Kroger, K.K., 2000. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature* 405, 347–351.
- Gabriel, R., 2018. The straw man in the new round of the reading wars. In: Valerie Strauss, *Washington Post*. September 25, 2018. <https://www.washingtonpost.com/education/2018/09/25/straw-man-new-round-reading-wars/>.
- Glaserfeld, E.V., 1990. An exposition of constructivism: why some like it radical. In: Davis, R.B., Maher, C.A., Noddings, N. (Eds.), *Monographs of the Journal for Research in Mathematics Education*, vol. 4. National Council of Teachers of Mathematics, Reston, VA, pp. 19–29.
- Goodman, K.S., Fries, P.H., Strauss, S.L., 2016. *Reading—The Grand Illusion: How and Why Readers Make Sense of Print*. Routledge, New York.
- Goodman, K.S., 1965. A linguistic study of cues and miscues in reading. *Elem. Engl.* 42, 639–643.
- Goodman, K.S., 1967. Reading: a psycholinguistic guessing game. *J. Reading Spec.* 6, 26–135.
- Goodman, K.S., 1986. What's Whole in Whole Language? A Parent/Teacher Guide to Children's Learning. Heinemann, Portsmouth, NH.
- Halliday, M.A.K., 1973. *Explorations in the Functions of Language*. Edward Arnold, London.
- Halliday, M.A.K., 1975. *Learning How to Mean: Explorations in the Development of Language*. Edward Arnold, London.
- Halliday, M.A.K., 1978. Meaning and the construction of reality in early childhood. In: Pick Jr., H.L., Saltzman, E. (Eds.), *Modes of Perceiving and Processing of Information*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 67–96.
- Halliday, M.A.K., 1979/1980. Three aspects of children's language development: learning language, learning through language, learning about language. In: Goodman, Y., Hausser, M., Strickland, D.S. (Eds.), *Oral and Written Language Development: Impact on Schools. Proceedings From the 1979 and 1980 IMPACT Conferences*. International Reading Association and National Council of Teachers of English, Urbana, IL, pp. 7–19.
- Halliday, M.A.K., 2003. On language in relation to the evolution of human consciousness. In: Webster, J. (Ed.), *On Language and Linguistics*. Bloomsbury Publishing, London, pp. 390–432.
- Harste, J.C., Burke, C.L., 1978. Toward a socio-psycholinguistic model of reading comprehension. *Viewpoints Teach. Learn.* 54, 9–34.
- Harste, J.C., 2003. What do we mean by literacy now? *Voices Middle* 3, 8–12.
- Heft, H., 2001. *Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James's Radical Empiricism*. Psychology Press, London.
- Hirsh-Pasek, K., Alper, R.M., Golinkoff, R.M., 2018. Living in Pasteur's quadrant: how conversational duets spark language at home and in the community. *Discourse Process* 55, 338–345.
- Holdaway, D., 1979. *The Foundations of Literacy*. Ashton Scholastic, Gosford, New South Wales.
- Hymes, D.H., 1972. On communicative competence. In: Pride, J.B., Holmes, J. (Eds.), *Sociolinguistics. Selected Readings*. Penguin, New York, pp. 269–293.
- Iacoboni, M., Woods, R.P., Brass, M., et al., 1999. Cortical mechanisms of human imitation. *Science* 286, 2526–2528.
- Kohler, W., 1929. *Gestalt Psychology*. Liveright, New York.
- Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Kweldju, S., 2015. Neurobiology research findings: how the brain works during reading. *PASAA A J. Lang. Teach. Learn. Thail.* 50, 125–142.
- Labov, W., 1968. The reflections of social processes in linguistic structures. In: Fishman, J.A. (Ed.), *Readings in the Sociology of Language*. Mouton & Co, The Hague, pp. 240–251.
- Labov, W., 1972. *Language in the Inner City: Studies in the Black English Vernacular*. University of Pennsylvania Press, Philadelphia.
- Lakatos, I., Feyerabend, P., 1999. In: Motterlini, M. (Ed.), *For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence*. Chicago University Press, Chicago.
- Lakoff, G., Johnson, M., 2003. *Metaphors We Live By*. University of Chicago Press, Chicago.
- Lakoff, G., 2004. *Don't Think of an Elephant!: Know Your Values and Frame the Debate—The Essential Guide for Progressives*. Chelsea Green Publishing, Hartford, VT.
- Langer, J.A., 2011. *Envisioning Knowledge: Building Literacy in the Academic Disciplines*. Teachers College Press, New York.
- Lund, F.H., 1927. *Psychology: The Science of Mental Activity*. A. G. Seiler, Reading, UK.
- Lyon, G.R., Chhabra, V., 2004. The science of reading research. *Educ. Leader* 61, 12–17.
- Manguel, A., 1996. *A History of Reading*. Penguin Books, New York.
- McLaughlin, S.F., 2006. *Introduction to Language Development*, second ed. Thomson Delmar Learning, Clifton Park, NY.
- Miller, G.A., 1951. *Language and Communication*. McGraw-Hill, New York, London.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97.
- Pearson, P.D., Gallagher, M.C., 1983. The instruction of reading comprehension. *Contemp. Educ. Psychol.* 8, 317–344.
- Piazza, E.A., Hasenratz, L., Hasson, U., Lew-Williams, C., 2020. Infant and adult brains are coupled to the dynamics of natural communication. *Psychol. Sci.* 31, 6–17.
- Popper, K.R., 1972. *Objective Knowledge: An Evolutionary Approach*. Clarendon Press, Oxford, England.
- Ramachandran, V.S., 2004. *A Brief Tour of Human Consciousness: From Impostor Poodles to Purple Numbers*. Pi Press, New York.
- Romeo, R.R., Segaran, J., Leonard, J.A., et al., 2018a. Language exposure relates to structural neural connectivity in childhood. *J. Neurosci.* 38, 7870–7877.
- Romeo, R.R., Leonard, J.A., Grotzinger, H.M., et al., 2018b. Beyond the 30-million-word gap: children's conversational exposure is associated with language-related brain function. *Psychol. Sci.* 29, 700–710.
- Rosenblatt, L.M., 1978. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Southern Illinois University Press, Carbondale, IL.
- Schaffner, A.K., 2021. You're Not a Computer. You're a Tiny Stone in a Beautiful Mosaic. *Psyche*, Aeon Media Group Ltd. <https://psyche.co/ideas/youre-not-a-computer-youre-a-tiny-stone-in-a-beautiful-mosaic>.
- Schneider, S.M., Morris, E.K., 1987. A history of the term radical behaviorism: from Watson to Skinner. *Behav. Analyst* 10, 27–39.

- Shuy, R.W., 1976. The Mismatch of Child Language and School Language: Implications for Beginning Reading Instruction. Paper Presented at the Conference on Beginning Reading Instruction. University of Pittsburgh, Pittsburgh, PA. April 1, 1976.
- Skinner, B.F., 1957. *Verbal Behavior*. Appleton-Century Crofts, New York.
- Smith, F., 1971. *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read*. Holt, Rinehart and Winston, New York.
- Smith, F., 1983. Reading like a writer. *Lang. Arts* 60, 558–567.
- Smith, F., 1986. *Insult to Intelligence: The Bureaucratic Invasion of Our Classroom*. Arbor House, New York.
- Tomasello, M., 1999. *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.
- Tomasello, M., 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, MA.
- Tomasello, M., 2008. *Origins of Human Communication*. MIT Press, Cambridge, MA.
- Tracey, D.H., Morrow, L.M., 2017. *Lenses on Reading: An Introduction to Theories and Models*, third ed. The Guilford Press, New York.
- Tulving, E., 1983. *Elements of Episodic Memory*. Clarendon Press, Oxford, England.
- Tupper, D.E., 1999. Introduction: Alexander Luria's continuing influence on worldwide neuropsychology. *Neuropsychology* 9, 1–7.
- Vygotsky, L.S., 1962. *Thought and Language*. Translation by E. Hanfmann and G. Vakur. MIT Press, Cambridge, MA.
- Winokur, S., 1976. *A Primer of Verbal Behavior: An Operant View*. Prentice-Hall, Englewood Cliffs, NJ.

Reintroducing “development” into theories of the acquisition and growth of early literacy: developmental science approaches and the cultural-historical perspective of L. S. Vygotsky

David B. Yaden, Jr. and Camille Martinez-Yaden, Department of Teaching, Learning and Sociocultural Studies, University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	849
Principles of development	849
Development is complex and nonlinear	850
A retrospective lens on interpreting developmental discontinuities in early reading and writing	850
Toward a dynamic, “historical” modeling of early literacy: from being to becoming	851
Learning to write in Spanish—managing multiple perspectives	853
The importance of showing variation in modeling literacy behavior	856
Computational vs. observation models of literacy development	856
Models of multidimensionality	857
Prospective, developmental models of learning	857
Andre’s writing sample reinterpreted	859
Bilingual children learning to write in Chinese and English	860
Conclusion	863
References	863

Introduction

L. S. Vygotsky’s essay on the *Prehistory of Written Language* (1931/1997) highlights four primary features of children’s early writing development which can also be considered fundamental to developmental science principles (cf. Gottlieb, 2003a,b; Granott and Parziale, 2002; Overton and Molenaar, 2015; Siegler and Crowley, 1991; Valsiner and Connolly, 2003; van Geert, 2017):

The history of the development of written language of the child presents great difficulties for research. To the extent that this can be judged by extant materials, the development of written language does not proceed along a single line preserving a continuity of forms. The history of the development of written language contains the most unexpected metamorphoses. We frequently encounter characteristics of spasmodic changes or breaks in the line of development. The line of development of written language in the child sometimes stops almost completely, then suddenly, as if completely out of nowhere, from outside, a new line begins and at first glance it seems that between the broken off past and the beginning new, there is absolutely no continuous connection. But only a naïve representation of development as a purely evolutionary process accomplished exclusively by gradual accumulation of separate small changes ... can conceal from our eyes the true essence of the processes that are occurring. Only someone who is inclined to imagine all the processes of development as processes of germination will deny that the history of written language in the child is fully entitled to be represented by a single line of development regardless of the breaks, dying off, and metamorphoses.

Vygotsky (1931/1997, p. 132).

Principles of development

From Vygotsky’s statement as well as from other findings within the literature of developmental science (cf. Fischer et al., 2003; Gottlieb, 2003a,b; Gottlieb et al., 2006; Overton, 2015; Valsiner and Connolly, 2003; van Geert, 2003, 2017) the following principles can be identified: Namely, *that normal growth is characterized by spurts, plateaus, and regressions, rather than gradual growth*—in other words—*development typically is not smoothly uniform and cumulative, but asynchronous and nonlinear*. As Thelen and colleagues (Thelen and Corbetta, 2002; Thelen and Smith, 1994) have pointed out in their studies of the development of infants’ early motor systems, even related subsystems involved in babies’ early kicking, walking and reaching behavior do not develop at the same pace—each subsystem seems to have its own timetable for development and ongoing coordination between all of the subsystems is dynamically driven by both internal as well as environmental factors.

Secondly, however, *increased variation often occurs after long periods of stability*. This characteristic of the development of many species is one of the fulcrum points of the theory of “punctuated equilibrium” developed by the late paleontologist, Stephen J. Gould (2002), and his colleague, Niles Eldredge, as a modification of evolutionary theory. However, similar trends in the history of scientific thought have been noted by Piaget and Garcia in their book *Psychogenesis and the History of Science* (1989). Nonetheless, similar to other revolutionary ideas in science (see Horwich, 2010; Kuhn, 1962, 2000), long-believed, well-entrenched theories are

difficult to dislodge. For example, despite empirical data to the contrary available over several 100 years, Aristotle’s ideas regarding motion survived over a millennium before seriously being questioned in physics (Piaget and García, 1989).

Thirdly, *what seems to be a phenotypic discontinuity in development is often reflective of the deeper continuity of the developmental trajectory*. Or rather, what seems to be an “absence” at times of visible development in an area, does not imply the “nonexistence” of the skill or ability in question—the latter may simply be obeying a deeper law of development in which some other level of related reorganization is taking place prior to more rapid and mature development.

The prior principle discussed above then leads to a last observation of development that *wide variation within any level of development at any one point is the rule rather than the exception*. For example, Siegler (1995) has persuasively shown that even when elementary children know a more sophisticated method for calculating mathematical sums, they will still employ lower level processes (such as counting on their fingers, or orally counting all the numbers in the equation) at least as often as the more advanced method. Similarly, Fischer and Yan (2002) have shown in an analysis of Darwin’s journals written while aboard the *Beagle* that Darwin simultaneously held many levels of beliefs about the processes of evolution both prior to and during his writing of the *Origin of the Species* (1859/2003).

It appears, then, that both children and adults at any level of development have at their command a range of abilities which they employ sometimes economically—and sometimes not—to solve a variety of problems. What *doesn’t* seem to be the case, however, under any circumstances, is that development involves the simple acquisition of a more advanced behavior or level of knowledge which then automatically becomes the normative way of operating. Rather, children and adults only “slowly abandon” old strategies which do not serve them well in the process of experimenting with newer learning strategies. Siegler (1996) has characterized this simultaneous orchestration of ability in his description of “overlapping waves theory” where, at any one time, as he has shown in mathematics learning (Siegler, 1995), and we (Yaden and Tardibuono, 2004, 2011) will illustrate for early writing later in this paper, children have several strategies at hand from which they draw in devising solutions for a variety of numerical and literacy tasks.

Development is complex and nonlinear

Another statement by Vygotsky (1931/1997) expresses the complexity involved in, particularly, child development:

But a positive picture [of child development] is possible only if we radically change our representation of child development and take into account that it is a complex dialectical process that is characterized by complex periodicity, disproportion in the development of separate functions, metamorphoses or qualitative transformation of certain forms into others, a complex merging of the process of evolution and involution, a complex crossing of external and internal factors, a complex process of overcoming difficulties and adapting (p. 99).

Vygotsky used a *geological* metaphor to characterize the above process—just as the strata of the earth is compressed, fractured and riven with heat, pressure, water and the untidy accumulation of extinct lifeforms, so are the influences of development comingled, merged, piled up and transformed. Thus, anyone who has worked closely with young children can appreciate Vygotsky’s comments particularly regarding the discontinuities, breaks, dying off, commingling and metamorphoses of which he speaks regarding children’s growth toward mastery of the written language system.

A retrospective lens on interpreting developmental discontinuities in early reading and writing

Even for researchers who have spent a majority of their professional lives studying early reading and writing behaviors, a writing sample like the one produced by 5-year-old Anthony, an emergent Spanish/English bilingual, still provokes wonder and amazement at the way a few graphic symbols can be arranged in such a dazzling display of directional exploration, combinatorial creativity, and topological range to represent both words and names, including his own (Fig. 1).

But what types of early literacy models (cf. Hanno et al., 2020; Tracey and Morrow, 2017) are there currently which can explain, in any systematic way, the range of writing performances by young emergent bilinguals (cf. Yaden et al., 2019) like Anthony’s with the obvious visual discontinuities and departures from the conventional standard? What types of models, metaphorical, statistical or otherwise (cf. Alvermann et al., 2019; Reichle, 2021; Tracey and Morrow, 2017) does the field have which can provide an explanation, for example, for how *each* graphic string represented in the sample shown is illustrative of an important feature of a unitary developmental trajectory? Or how do all of the graphic strings together speak to the typical or, perhaps, atypical trajectory being built at this moment in time?

For reasons which will be discussed shortly, the type of model needed to illuminate developmental growth is *not* what is pictured in Fig. 2 despite that model’s popularity as a model of early literacy development.

The type of model illustrated in Fig. 2 is characteristically called “Simple” (see Gough and Tunmer, 1986; Hoover and Tunmer, 2018; Petscher et al., 2020; Seidenberg, 2017) as it is based upon the equation “Decoding (D) × Language Comprehension (LC) = Reading Comprehension (RC)”. Nonetheless, Paul van Geert (2003) has also described such a model as “retrospective”—in other



Fig. 1 Anthony's writing sample at 5-years-of-age.

words, “one that looks at the developmental processes from the perspective of a static end state and views all preceding states in light of this end state” (p. 641).

In addition, as Paris (2005) has pointed out related to the modeling of early literacy skills, researchers seldom take into consideration *either* the different developmental time scales of the variables which are measured *or* the fact that strong correlations depend upon only the *partial development* of abilities—given that only partial development ensures statistical variation.

One interpretation which can be drawn from Paris' observations is that “constrained skills” or those abilities which seem to develop faster or early—such as name writing—fall out of the model quickly since most name writing scoring systems that index figural aspects reach ceiling scores very rapidly as most children learn to write their names prior to kindergarten in many cases. On the other hand, elementary outside-in skills (see Whitehurst and Lonigan, 2002) such as language development, vocabulary knowledge, conceptual and domain knowledge—which are learned across the lifespan and are harder to measure—drop out of the model later either because of wide variation within the groups being measured or minimum variation due to test difficulty.

But whether or not the variables measured statistically fall out early or late, we are faced with the decision, as Paris (2005) reminds us, of interpreting the findings of developmental “discontinuities,” at least according to the Simple Model of Reading (Gough and Tunmer, 1986), which relies upon the lack of path coefficients in the model below (see Fig. 2), as supporting the following:

1. That elementary “outside-in” knowledge (as vocabulary or concept growth) doesn't correlate with reading in grades 1 or 2,
2. That vocabulary growth becomes less important after preschool and kindergarten, and,
3. That learning how to write one's name completely captures what it means to be an early writer.

Thus, we must interpret these discontinuities as *either* being theoretically robust according to the model or, on the other hand, as *artifacts* of factors relating to both statistical conclusion validity of the measures used and of inadequate theorizing of the reading/writing acquisition processes, treating the various aspects of literacy growth as developmentally isomorphic across time—which clearly they are not.

Toward a dynamic, “historical” modeling of early literacy: from being to becoming

Thus, there are at least two reasons to be highly skeptical of “retrospective” accounts of literacy development. First, according to Siegler (2006),

Knowledge states are easier to describe than change processes ... and the prolonged emphasis on documenting static states results in a distorted description of development, one in which children appear to spend far more time *being* than *becoming*.

While retrospective or static state models may be important for sketching out possible relationships between constructs or variables, they do not actually measure change processes themselves, but rather static states of knowledge separated by arbitrary time

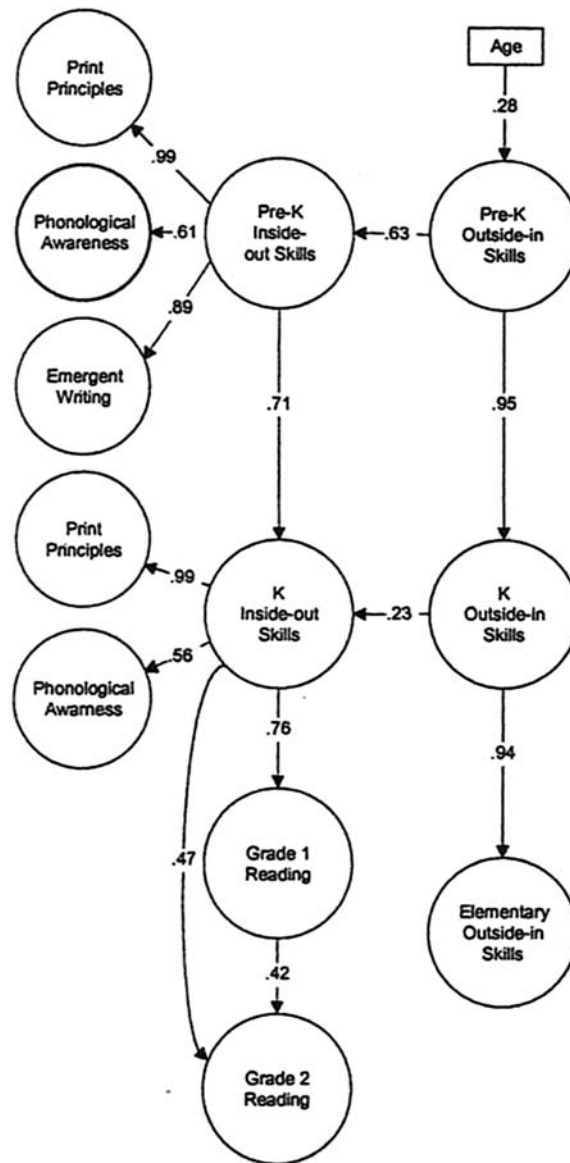


Fig. 2 A structural model of the development of inside-out and outside-in skills from pre-K through 2nd grade. From Whitehurst, G., Lonigan, C., 2002. Emergent literacy: development from prereaders to readers. In: S. B. Neuman, S.B., D. K. Dickinson, D.K. (Eds.), *Handbook of Early Literacy Research*. Guilford, New York, NY, pp. 11–29.

periods. And as [Vygotsky \(1931/1997\)](#) pointed out, it is only when the *becoming* is studied from a “historical” perspective that the richness of development shows itself.

To study something historically means to study it in motion. Precisely, this is the basic requirement of the dialectical method. To encompass in research the process of development of something in all its phases and changes—from the moment of its appearance to its death—means to reveal its nature, to know its essence, for only in movement does a body exhibit what it is.

Secondly, according to [Vygotsky \(1934/1987\)](#) and other developmentalists ([Bronfenbrenner, 1979](#); [Edelman, 1992](#); [Tomasello, 2019](#); [van der Veer and Valsiner, 1991](#); [Waddington, 1957](#)) knowledge and abilities which develop first or early do not simply fade away, but rather are incorporated into the fabric of more advanced growth and undergo change themselves on the “trailing edge” of growth, as it were, even as newer advanced concepts are added to the leading edge. Thus, as Vygotsky noted again, “The new higher concepts, in turn, transform the meaning of the lower. The adolescent who has mastered algebraic concepts has gained an advantage from which he sees the concepts of arithmetic in a broader perspective” ([Vygotsky, 1934/1987](#), p. 202). Therefore, higher order concepts illuminate and bring a broadened perspective to formerly held understandings, thus effectively transforming the entire

learning and developmental system throughout, making it even more efficient, or in Piaget’s (1977) words, a dynamic system always advancing and seeking an “optimizing equilibration.”

And we would submit that as a child embarks upon the journey of becoming a literacy learner and a user of such symbols as we associate with that learning—a journey which continues over decades—that what is learned along the way does not become less important after its initial appearance, but rather undergoes continual reorganization within the learner’s socio-emotional and cognitive networks in ways which enrich the whole learning system—making the child, adolescent and adult even more efficient in acquiring more complex forms of knowledge, in other words, nothing is wasted, nothing is lost.

Learning to write in Spanish—managing multiple perspectives

We want to now turn to some specific examples of children writing their names as well as other words to illustrate the range of variation in understanding that children exhibit during such tasks, variation which is difficult, if not impossible to demonstrate in a retrospective model such as in Fig. 2.

Depicted in the next few Figures (nos. 3–8) are several 4–5-year-old, Spanish-speaking children’s writing performances across 11 tasks (cf. Ferreiro and Teberosky, 1982) as they wrote their name, their mother’s and father’s names, several words *mama*, *papa*, *nene* (*baby*), *oso* (*bear*), *sapo* (*toad*), *mapa* (*map*), *pato* (*duck*) and a sentence—*mi nene toma sol*—*my little girl sits in the sun*.

After each word was written, the child was asked to read it back, pointing with their finger, and, in the case of the child’s name, the first and last few letters were covered, and the child was asked to read what symbols were uncovered.

In the actual testing protocol with the child, each word was written with a different marker so that it was easier to tell which string was which, although in a few cases, the child’s personal preference for certain colors overrode the researchers’ instructions.

The following are just a few examples from a corpus of 500 or so writing samples which were collected over 2 years in a study (Tardibuono, 2006; Yaden and Tardibuono, 2004, 2011) funded by the Center for the Improvement of Early Reading Achievement (CIERA). Note the differences, particularly, between how the children write their names and the other words.

As can be seen in Figs. 3–8, the children wrote using mostly a combination of uppercase letters and letter-like symbols, having a range of 3–11 symbols per word which they rearranged in different combinations when writing different words. While some were more particular about writing the words in a column, others wrote simply where there was space to write with some children (not shown here) taking as many as three pages to record all of the words. Despite the variation in how the children chose to represent each one of these words, one of the more obvious features of the samples is that the child’s name is the one word which is conventionally written. And, indeed, these examples illustrate, as many (e.g., Bloodgood, 1999; Cabell et al., 2009; Levin et al., 2005) have pointed out over the years, of the importance of the child’s name in the beginning repertoire of what they feel that they can write—at least in a situation where they are requested to write.



Fig. 3 Writing samples from 5-year-old Andres. From top left: Andres’ name, Martha (mother’s name), Alphonso (father’s name), *mama* (written vertically as “map”).



Fig. 4 Writing samples from 5-year-old Andres. From top left: oso, mapa (bp), papa (AD), pato (AE), nene (if), mi nene toma sol (PO), sapa (sp). **Figs. 3–8** from Tardibuono, J., 2006. A Psychogenetic Analysis of Spanish-Speaking Preschoolers' Emergent Writing: A Developmental Trajectory Toward the Phonetization of Speech (Publication No. 3238342). Doctoral Dissertation, University of Southern California.



Fig. 5 Bernardo's writing sample.

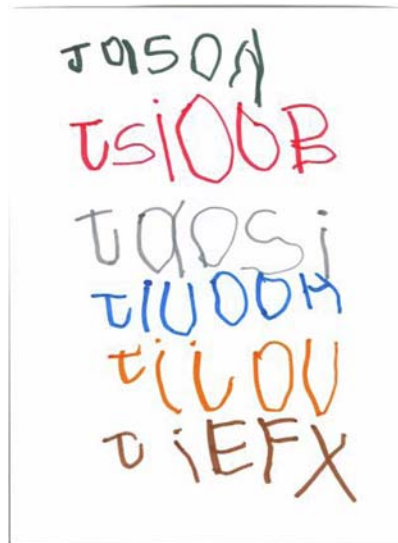


Fig. 6 Jason's writing sample.



Fig. 7 Jeanette's writing sample.



Fig. 8 Maria's writing sample.

However, in the samples illustrated, only Andres (Figs. 3 and 4), according to all that we could detect from the testing situation, understood that the symbols being written in his name and other words (e.g., mama, papa, nene, etc.) were “second-order” in that they represented smaller, linguistic segments of the name (phonemic) or word (syllabic) as it was spoken. In all the other cases, covering up portions of the child’s name resulted in the child either not being able to recognize it or changing the form to another name with which they were familiar.

Thus, for a majority of the children, the name was treated as a “logo” of sorts (cf. Tardibuono, 2006), in other words, a graphic which needed to be taken as a whole in order to be meaningful or, in Vygotsky’s terms (1931/1997), a “first-order” symbol referring directly to their identity as a person. It is also an example of what Luria (1988) described as illustrative of the “glass window theory,” indicating that children and adults too, in some cases, “look through” the form of a language directly to the meaning.

In addition, we discovered that while Andres (Figs. 3 and 4) understood that his name could be “sounded out” phonemically, this understanding did not extend to his mother’s or father’s name which were understood as intact graphic strings as was the word “oso” which he had been able to write conventionally several months earlier.

Further, while he demonstrated a clear understanding of what Ferreiro and Teberosky (1982) have termed the “syllabic principle”—that most of the two-syllable words he was asked to write needed one symbol per syllable—any individual symbol might represent a variety of sound units. In addition, he also wrote the entire sentence “my little girl sits in the sun” with two symbols as well—“PO” (see Fig. 4)—segmenting the sentence into larger clausal chunks as well as changing the meaning of the sentence to something more relevant to him, i.e., “my mother [P]... points at the sun [O].”

From what we could tell from this assessment and others we did in the classroom, Andres represented nearly the entire range of the five levels which Ferreiro and Teberosky (1982) have described—from a very global syncretism, to understanding the syllabic principle, and even to demonstrating phonemic awareness—all in the span of the 10 min it took to produce the writing sample. The point we want to underscore here is that Andres’ range of understanding of the structure and symbolic aspects of the written system cannot be adequately determined by only looking at the figural representation of his name. In fact, Andres would actually be scored lower, using some of the more common assessments of name writing such as Hildreth’s (1936) or the *Phonological Awareness Literacy Screening* (PALS) scoring system (Invernizzi et al., 2004) because of the two letter reversals, even though none of the other children who wrote their names conventionally could segment them into linguistic units.

The importance of showing variation in modeling literacy behavior

Thus, we would like to assert that name writing—in and of itself—is not completely informative of what a child is learning about written language as are other types of writing tasks the child might attempt. What can be seen from the writing samples is that children’s early writing performances are *highly varied* both between children as well as within any one child’s specific performance. A child’s accomplishment of a series of writing tasks such as ones comprising the research protocol illustrated is not seamless, but interspersed with interruptions of various sorts, hesitations, perplexities as well as certainties about how some words are to be written (cf. Rowe, 2015).

Computational vs. observation models of literacy development

If our analyses of very young children’s writing is limited to only what we see in their names, then we have hardly scratched the surface in understanding their developmental trajectory toward becoming facile with the written system. Notwithstanding the tremendous accomplishment of writing one’s name at 3- or 4-years of age, most of the intellectual effort of which is invisible in the brief moments of the act, at this period in a child’s growth, it is a “situational” performance—requiring knowledge which may or may not extend to writing other names or words even a few seconds later. Therefore, name writing must be considered within a larger context of writing tasks and purposes within the network of a child’s wider understanding of the use of written symbols before any display can be usefully interpreted.

The type of developmental complexity we have just witnessed cannot be represented in the model above (see Fig. 2). Relying on the *computation* rather than on the *observation* of change, the above depictions are “naïve,” as Vygotsky (1931/1997) stated, in the sense that, first of all, discontinuity—or nonlinear relationships between states—is assumed to be indicative of an absence or the waning of a relationship.

Secondly, as Bronfenbrenner (1977, 2005) has pointed out, standardized regression weights are a “one size fits all” computation, assuming that the trajectory of development is uniform over each individual. As we have seen (with Andres), this uniformity doesn’t even exist within individuals for many variables measured. Further, the point that all participants in an experimental group or in a randomized clinical trial *don’t* benefit from the intervention is strikingly apparent in a recent update of the Perry Preschool Study entitled *Lifetime effects: The high/scope Perry Preschool study through age 40* (Schweinhart et al., 2005) where 8 case studies are reported. Interestingly, the individual having the highest number of “success” indicators (education, employment, criminal record, and family/community life) is actually a *no-program* female—in other words, one of the control group members. Heckman (2005), a Nobel Laureate in economics, although praising the study for its many strengths, states the following:

Comparisons in this monograph are made cross-sectionally at ages 27 and 40 without analyzing the individual life cycle profiles that would be of interest for understanding human development processes. ... It would be valuable in a continuation of this study to analyze individual growth profiles rather than group means over time (p. 230).

Heckman goes on to suggest that there are several ways to be more precise, statistically, about the individual trajectories in this study which could include small sample parametric tests, event history and sensitivity analyses.

In a similar vein, but through an anthropological and sociological lens, Diana Slaughter-Defoe, also commenting upon the Perry Study, makes the following observation:

It seems that after the preschool program ended, even the program participants had evidence of serious life challenges in the form of continuing poverty, exposure to drugs in the communities in which they lived, and the multiple family and social problems that can spiral out of control without help. I wish it had been possible to ascertain how their obvious resiliencies enabled them to overcome those challenges.

Slaughter-Defoe (2005, p. 238).

The point that both of these scholars touch upon is the absence of any developmental information about just how did the preschool experience link with, as Slaughter-Defoe suggests, “the multiple family and social problems that can spiral out of control without help.” Thus, without a strong developmental and theoretical understanding of just how the early experiences linked with later ones, we are left with an unspecified, unknown alternative theory of effects that explains how a non-program participant, similar in initial conditions, but without the benefit of an intensive early childhood program, can still end up after nearly four decades as the highest achiever.

Models of multidimensionality

While methodologically it is important to do more sophisticated quantitative and qualitative research, it is perhaps more important for those of us studying the trajectory of literacy growth over time to have a more nuanced and powerful theory to explain how human beings develop (cf. Tomasello, 2019; Yaden et al., 2014).

So then, what types of models and developmental principles are suggestive in explaining complex patterns of behavior such as we see in young children learning how to write?

For some physicists, mathematicians and chemists at the [Institute of Figuring \(n.d.\)](#) in Los Angeles, these crocheted objects (Figs. 9 and 10), knitted by these same researchers, are physical models of not only of concepts in hyperbolic geometry, but also of natural shapes in the universe. It is also suggested by the group of scientists who crochet these shapes that the various curves and involutions may represent lines of development of different types of physical or conceptual development.

While we are limited in our ability to understand much beyond two dimensions, particularly in our methods of assessment, statistics and research design, we operate in the real world in a least as many dimensions as we have senses—and then some—from our personal experiences, ethnic backgrounds, family histories, school experiences, habits of thinking, and belief systems.

Thus, each one of the convolutions and involutions pictured could be seen as representing one facet of learning which is distinct in some ways, but ultimately connected to the whole system. It may very well be that these shapes are also a fairly accurate image of what children go through in learning a complex skill such as learning to read or write.

Suppose we hypothesize that each curvature in the hyperbolic plane of Fig. 11 represents a child’s growing ability to understand and eventually control the multiple aspects of the task we call learning to write. At any one time, these contours and cleavages may appear to operate somewhat in isolation, but as can be seen, they are nonetheless connected with one another in a complex, but unitary shape, as Vygotsky observed nearly a century ago.

Prospective, developmental models of learning

In contrast to retrospective theories of development, “prospective” theories such as represented by *dynamic skill theory* (Fischer and Biddell, 1998; Fischer et al., 2003), *indeterministic constraints modeling* (Valsiner, 1998), *probabilistic epigenesis* (Gottlieb, 2003a,b), and *Bronfenbrenner’s bioecological theory of human development* (Bronfenbrenner and Morris, 1998, 2007) are theoretical approaches in child and developmental psychology underscoring the “normal” processes of learning as being among other things,

(a) Characterized by spurts, plateaus, and drops rather than gradual growth, (b) exhibiting either *equifinality* (converging on the same end state from different beginnings) or (c) *multifinality* (divergent outcomes from similar beginnings), and (d) following an *epigenetic* course of development; in other words, the growth trajectory is highly susceptible to what Gottlieb (2003a,b) has termed “non-obvious experiential contingencies” as much as planned experiences such as instruction.

Further, prospective theories in developmental science suggest certain “rules” which operate to govern the progress of development in any area. Some of these include: that.



Another pseudosphere. This model was made by increasing one stitch every two stitches. Notice that it is considerably more crenellated than the previous pseudosphere. The hyperbolic radius here is about 3 centimeters. The model itself is about 10 centimeters in diameter.

Fig. 9 Crocheted hyperbolic planes. Crochet hyperbolic models by Margaret Wertheim. Photo © Institute For Figuring. <https://theiff.org/oexhibits/oe1e.html>.



Hyperbolic plane. Crocheted models of this space begin with a line of chain stitches. From there, one increases the number of single or double crochet stitches in every row, the rate of increase determining how rapidly the space will expand. This model is about 24 centimeters across and is made by increasing one stitch every 13. The structural integrity is achieved by using coarse, cheap synthetic yarn. Wool is nicer to work with (and gentler on the hands) but it does not produce the satisfying stiffness of synthetics. This model was made by Daina Taimina.

Fig. 10 Crocheted hyperbolic planes. Crochet hyperbolic models by Margaret Wertheim. Photo © Institute For Figuring. <https://theiff.org/oexhibits/oe1e.html>.

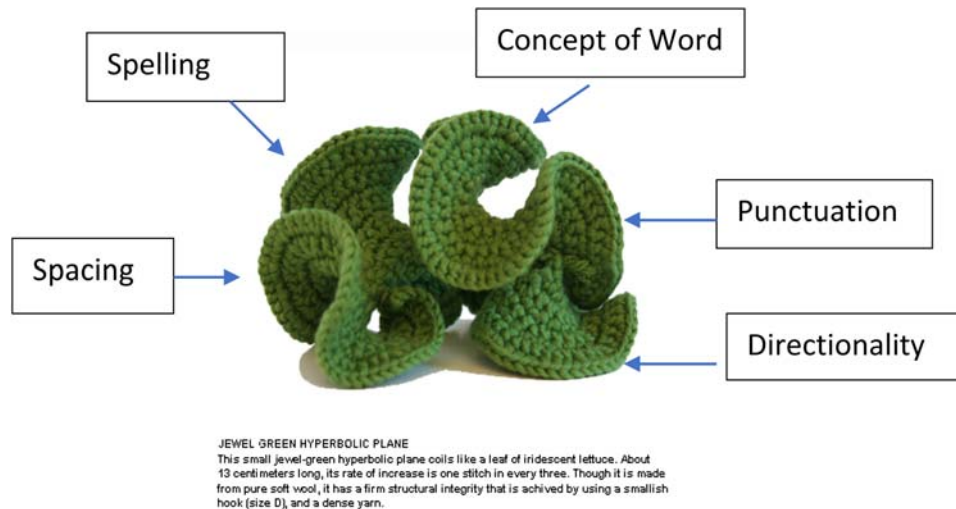


Fig. 11 Aspects of literacy behavior represented and connected by various trajectories in a crocheted hyperbolic plane. Adapted from the [Institute for Figuring](https://theiff.org/oexhibits/oe1e.html) (n.d.). Crochet hyperbolic models by Margaret Wertheim. Photo © Institute For Figuring <https://theiff.org/oexhibits/oe1e.html>.

1. *A lower-level ability is only solidified by operating at a higher level.* In other words, “mastery” is only achieved by experimenting with higher level learning—not by repeating the lower-level skill over and over again. Therefore, restricting anyone’s learning to just the task at hand hampers the developmental process. Automaticity will not develop unless the possibility exists to practice any skill at multiple levels of difficulty.
2. *Experimenting initially with a higher level of functioning usually results in the learner reverting to using lower level processes.* However, rather than “regressions,” this utilization of earlier or familiar strategies have been denoted by [Granott \(2002\)](#) as “backward transitions” that are iterative processes solidifying what has been learned so that advancement can become more rapid (See also Templeton, this volume, for similar patterns in spelling development).
3. *Learners at any time are operating on multiple levels simultaneously.* In other words, the development of competence in any skill or activity means that the learner brings a range of skill levels ([Siegler, 1996, 2006](#)) to the task at hand and uses them according to the levels of difficulty perceived to be inherent in the task to be accomplished.

Andre’s writing sample reinterpreted

Let’s return to Andres’ writing sample (Figs. 3 and 4) from the standpoint of both retrospective and prospective models. First of all, in the retrospective model briefly discussed ([Whitehurst and Lonigan, 2002](#)), only the name—the first of 11 tasks—can be used—data from the other ten tasks must be discarded. This is a tremendous waste of information. Secondly, since only the figural features of the name are scored in traditional assessments of name writing, Andres’ name—although he demonstrates phonemic awareness of the individual letters—is no more useful in the model than any of the other four names written (see Figs. 5–8) by children who did not attribute any sound values to the symbols. Thus, not only is data lost in a retrospective view, but the conceptual complexity of the data itself is reduced to and conflated with that of the description of the visual aspects alone.

While “prospective” or developmental systems theories have been typically underutilized across many fields of education, psychology, and public health ([Luke and Stamatakis, 2012](#)), in child psychology there has been a shift from a Cartesian-framed analysis of behavior to what [Overton and Molenaar \(2015\)](#) have described as a “Process-Relational and Relational-Developmental-Systems” paradigm. The advantages of using a relational-developmental-systems framework orientation in discussing Andres’ performance are that *all the data can be used* and that the conceptual range that Andres demonstrates—from the phonemic awareness of the letters in his name to the use of one symbol per syllable, or one symbol per clause, to the iconic representation of both the word “oso” and his parents’ names can be interpreted according to “overlapping waves theory” ([Siegler, 1996](#)) as well as the following principle of developmental theory that “within-child variability tends to be greatest during periods of rapid learning—though substantial variability is present in relatively stable periods as well” ([Siegler, 2006](#)).

For Andres, then, the conceptual range within which he operated at the time this sample was collected, may well have presaged a more rapid learning toward a conventional understanding of the uses of written language than perhaps other children whose writing of the 11 tasks, conceptually at least, were more uniform. However, further follow-up into the kindergarten year was not a part of the study ([Tardibueno, 2006](#)), and somewhat limits a longitudinal observation and interpretation of these children’s ultimate trajectories in learning how to write in their heritage language and in English, the language of instruction in school.

Bilingual children learning to write in Chinese and English

Finally, in addition to research with Spanish/English emergent bilingual speakers, we have conducted a series of studies with beginning bilingual/biliterate children learning languages which differ in language structure and orthography such as Persian (Mirzai, 2002), Chinese (Tsai, 2007; Yaden and Tsai, 2012) and Korean (Kwak, 2006). With young, kindergarten Chinese/English bilingual children, the following patterns have been observed (Tsai, 2007; Yaden and Tsai, 2012).

Unlike the Spanish/English emergent bilingual studies, where children were only interviewed once with several tasks, Tsai (2007) used a microgenetic research design to measure children's understanding of each written task or trial in multiple sessions over a period of 3 months. These bilingual Chinese/English children were receiving instruction in Chinese literacy after school in Los Angeles for 5 h a week while attending local kindergartens where all instruction was conducted in English. All but one of the children spoke Mandarin at home. However, all children conversed freely in English during the study sessions, although we did not test specifically for language proficiency.

Some of the patterns already discussed in this paper appear in the following displays. For example, in Figs. 12 and 13, the child shows a “backward transition” in Chinese (Fig. 12) after the first two sessions, from a syllabic or word interpretation of whole composite characters and their semantic and phonetic components to a prior interpretation of objective differences only such as the number of strokes or not recognizing the 3–4 component structure of the composite character. In interpreting English script (Fig. 13), a level 2 indicates primarily that the child interprets the string according to visual characteristics only, and not to any linguistic segment. This pattern is stable for a month, with nearly every task being scored at Level 2. Beginning in November, however, much more fluctuation appears, both up and down.

Figs. 14 and 15 offer another example of variations or “punctuations,” following Gould and Eldrege (2002), after periods of a stable understanding of interpreting both Chinese and English scripts. The child's knowledge of Chinese is higher than English

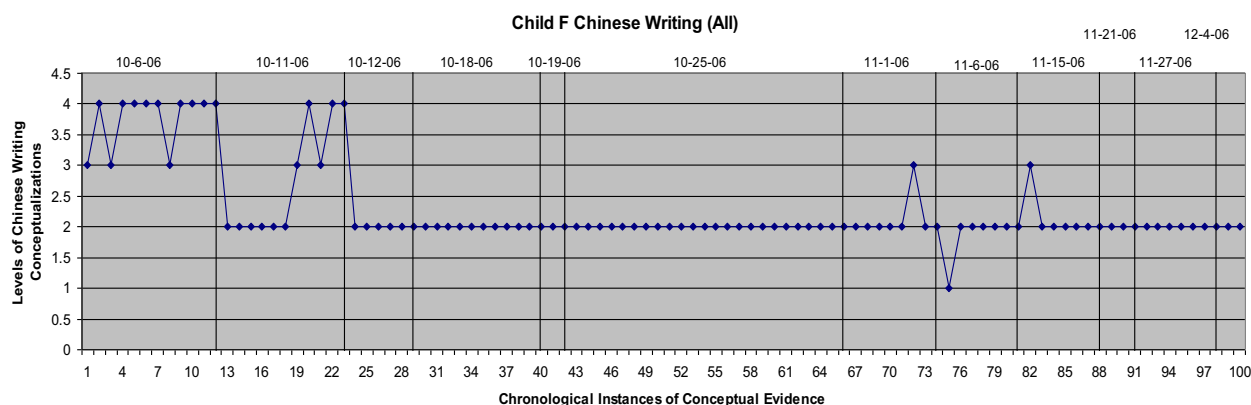


Fig. 12 A pattern of “backward transition” in the interpretation of Chinese characters in the writing of an emergent Chinese/English bilingual kindergartner. Note. Writing sample in Chinese exhibiting a “backward transition.” For all subsequent figures, the Y-axis represents increasing levels of conceptual sophistication and the scored level of each task (child's name, mother's name, father's name, etc.). The X-axis represents number of total trials. Tasks repeat with each date segment. Figs. 12–21 from Tsai, T., 2007. A Microgenetic Analysis of English/Chinese Early Writing Development (Publication No. 3291823). Doctoral dissertation, University of Southern California.

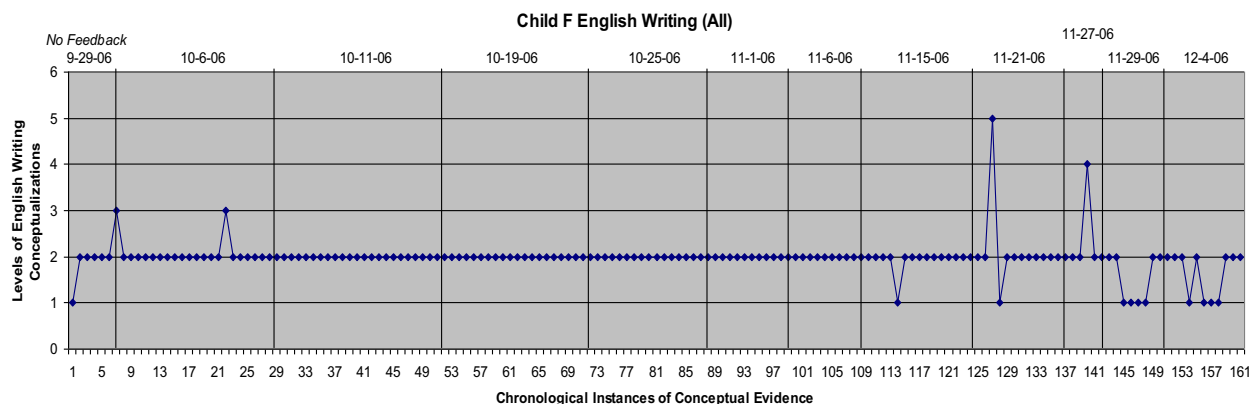


Fig. 13 A pattern of “backward transition” in the interpretation of English letters in the writing of an emergent Chinese/English bilingual kindergartner. Note. Writing sample in English representing punctuation after period of stability.

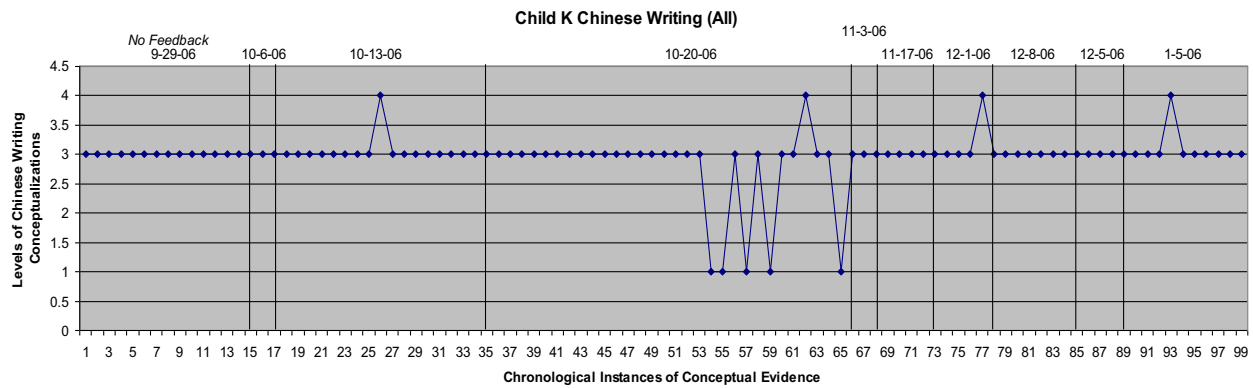


Fig. 14 A pattern of “ordered fluctuations” after a period of stability in the interpretation of Chinese characters by an emergent Chinese/English bilingual kindergartner.

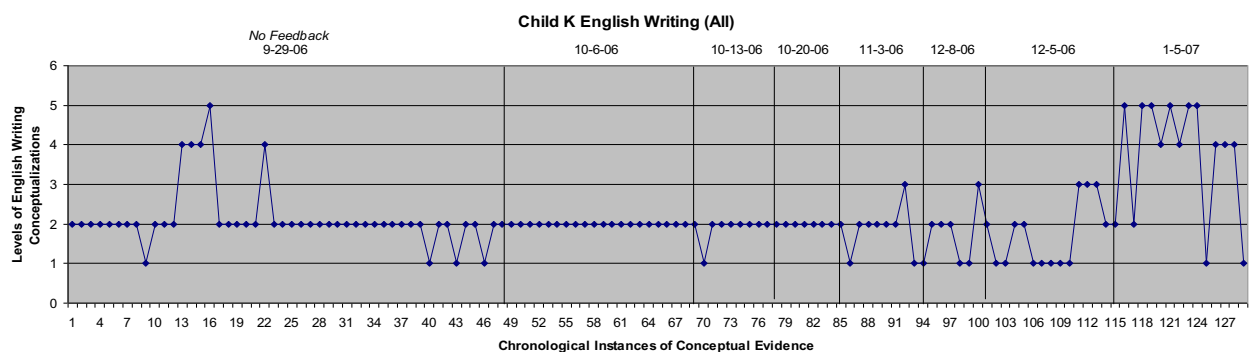


Fig. 15 A pattern of “ordered fluctuations” after a period of stability in the interpretation of English letters by an emergent Chinese/English bilingual kindergartner.

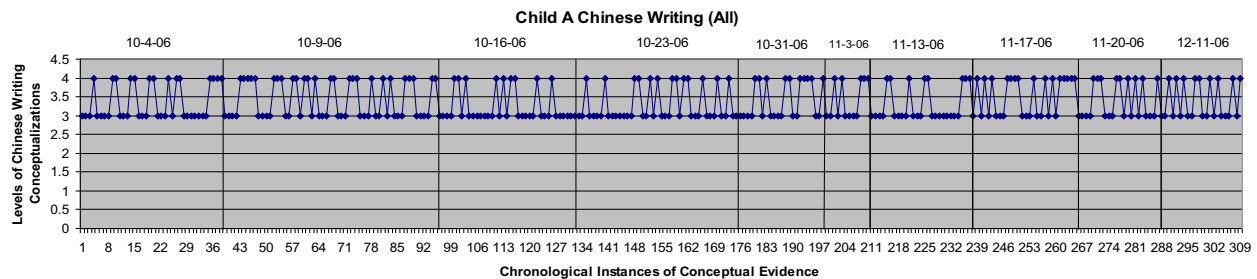


Fig. 16 A pattern of consistent “ordered fluctuations” in the interpretation of Chinese characters by an emergent Chinese/English bilingual kindergartner.

until the last session when the child begins to relate English letters and patterns to syllabic and phonemic segments. These graphs show what [Granott \(2002\)](#) has termed “ordered fluctuations” in the sense that the variations regularly vary between certain levels in both Chinese and English, until the last session in English where the variation becomes wider and the trend toward greater understanding of the English system seems probable. Again, it is clear that this child also operates on multiple levels of conceptualization in both languages.

The following displays ([Figs. 16–21](#)) show that becoming biliterate in any two languages offers varying patterns of growth. In [Figs. 16](#) and [17](#), the child demonstrates extremely ordered fluctuations in both languages.

In [Figs. 18](#) and [19](#), however, this child shows variation of understanding the tasks in the second language (English, [Fig. 18](#)), but not the first (Chinese, [Fig. 19](#)).

Finally, [Figs. 20](#) and [21](#) indicate that this child exhibits variation across several levels of conceptualization from task to task, from month to month in both the heritage language and English.

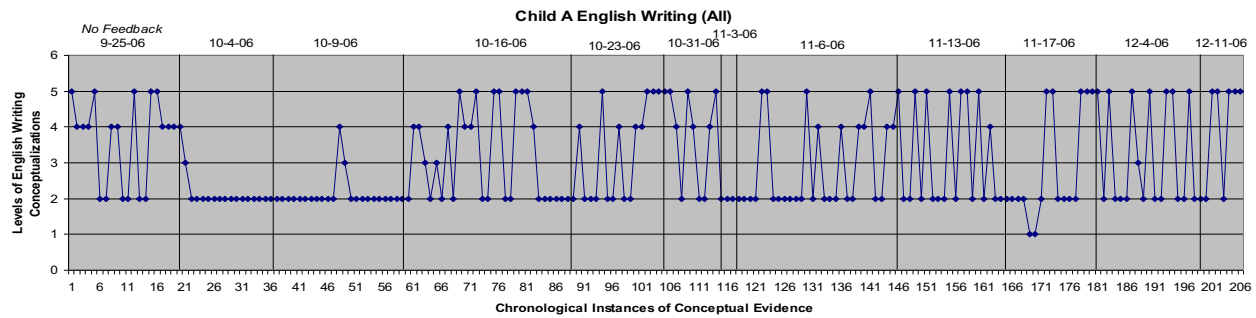


Fig. 17 A pattern of consistent “ordered fluctuations” in the interpretation of English letters by an emergent Chinese/English bilingual kindergartner.

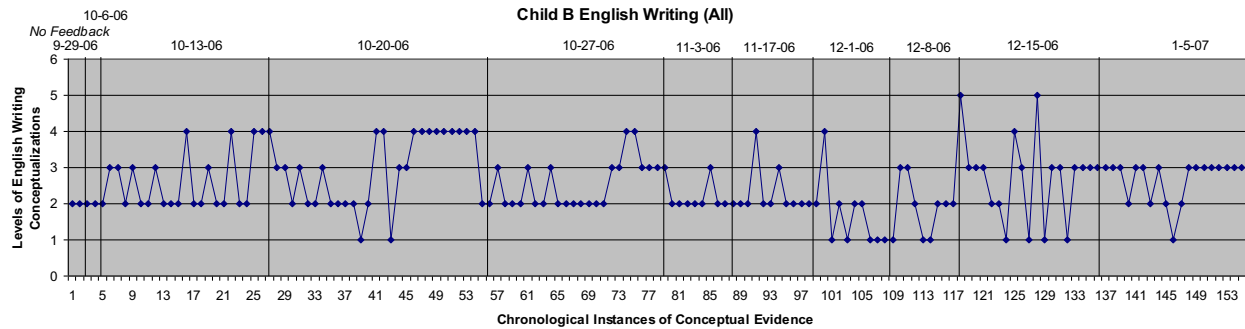


Fig. 18 Variation of conceptualizations about the function of English letters by an emergent Chinese/English bilingual kindergartner who learned English as a second language.

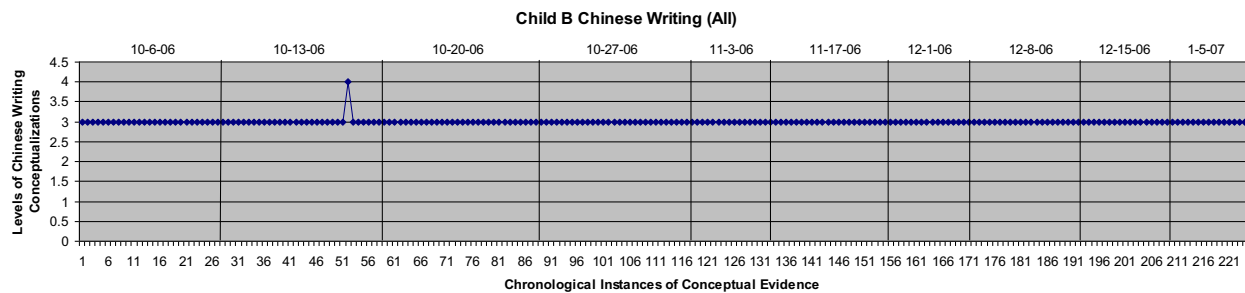


Fig. 19 A stable conceptualization in the interpretation of Chinese characters by an emergent Chinese/English bilingual kindergartner.

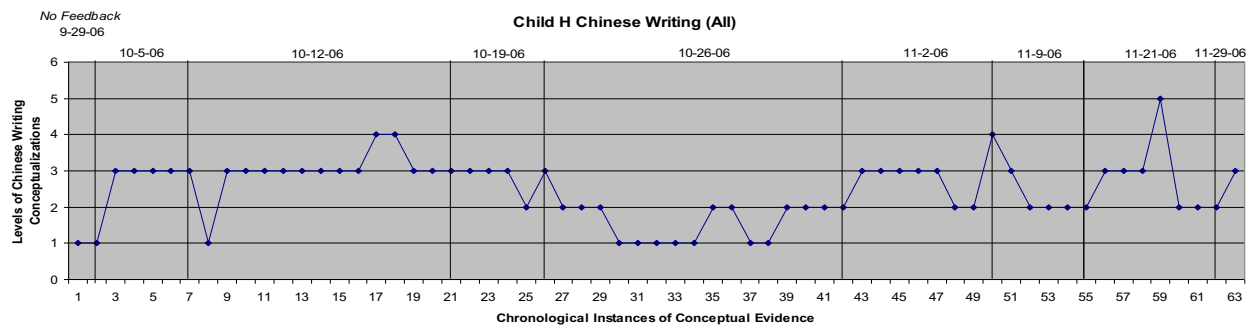


Fig. 20 Wide within-task variations of conceptualization in the interpretation of Chinese characters by an emergent Chinese/English bilingual kindergartner.

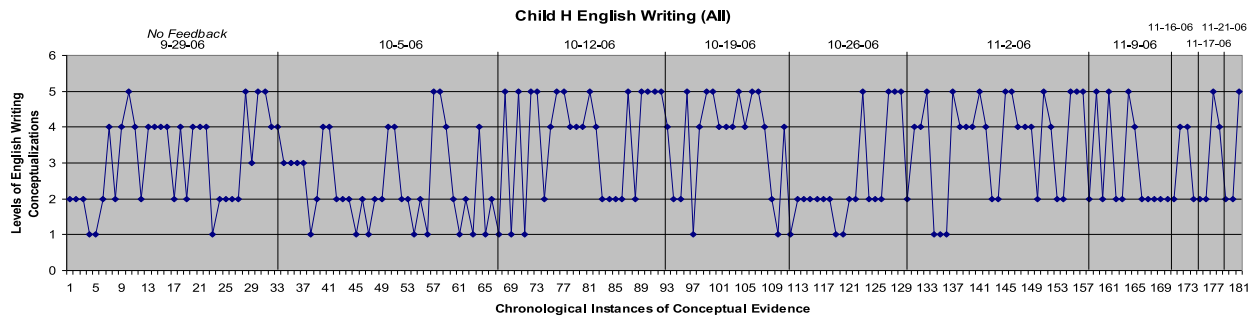


Fig. 21 Substantial within-task variations of conceptualization in the interpretation of English letters by an emergent Chinese/English bilingual kindergartner.

It is important to add that much more detailed study of the children’s home environments as well as the classrooms for learning English and Chinese would be needed to begin to understand the sources of variation or stability, for that matter. Such design features, however, were not built into this particular study.

It suffices to say, however, even without knowing where the particular sources of variability might lie, that any child’s learning of a written system of any language will exhibit, most likely, some of the types of variation demonstrated in these studies. And, if so, then the implication for models of early literacy is clearly that they become able to incorporate such diversity and variation as Paris (2005) and our studies (see Yaden and Tardibuono, 2004, 2011; Yaden and Tsai, 2012) suggest.

Conclusion

An oft-quoted passage of Vygotsky’s from the *Prehistory of Written Language* related to the teaching of writing to young children states,

If we should want to summarize the practical requirements and express them in a single position, then we could say that consideration of the problem leads us to the requirement to teach a child written language and not writing the alphabet (p. 147).

Or in the cases of scripts which index other linguistic units, not to focus so much on those forms either, but rather as Olson (1994) has suggested—teaching children about what it means to “take” meaning from any script—what are the larger purposes of being literate in any society?

Developmentally speaking, our contention is that retrospective models of early literacy such as depicted in Fig. 2 expect us, rather, to demand a slavish focus on the alphabet and figural convention. However, prospective modeling and current microgenetic, developmental theory (cf. Overton, 2015) urges us to do the opposite, as did Vygotsky, Piaget and other developmental theorists (cf. Bronfenbrenner, 2005; Edelman, 1992; Tomasello, 2019). In other words, embrace the complexity of the developmental processes and as Vygotsky further argued “bring the child to an internal understanding of writing” in whatever language that may be.

References

- Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B., 2019. *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, NY.
- Bloodgood, J.W., 1999. What’s in a name? Children’s name writing and literacy acquisition. *Read. Res. Q.* 34 (3), 342–367.
- Bronfenbrenner, U., Morris, P.A., 1998. The ecology of developmental processes. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology: Theoretical Models of Human Development*, vol. 1. John Wiley & Sons Inc, pp. 993–1028.
- Bronfenbrenner, U., Morris, P.A., 2007. The bioecological of human development. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of child psychology*. John Wiley & Sons, Hoboken, NJ, pp. 793–828.
- Bronfenbrenner, U., 1977. Toward an experimental ecology of human development. *Am. Psychol.* 32 (7), 513–531. <https://doi.org/10.1037/0003-066X.32.7.513>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press, Cambridge, MA.
- Bronfenbrenner, U. (Ed.), 2005. *Making Human Beings Human: Biological Perspectives on Human Development*. Sage, Thousand Oaks, CA.
- Cabell, S.Q., Justice, L.M., Zucker, T.A., McGinty, A.S., 2009. Emergent name writing abilities of preschool children with language impairments. *Lang. Speech Hear. Serv. Sch.* 40, 53–66.
- Darwin, C., 2003. *The Origin of the Species*. Signet, West Bengal, IN.
- Edelman, G., 1992. *Bright Air, Brilliant Fire: On the Matter of Mind*. Basic books, New York, NY.
- Ferreiro, E., Teberosky, A., 1982. *Literacy Before Schooling*. Heinemann, Portsmouth, NH.
- Fischer, K., Biddell, T.R., 1998. Dynamic development of psychological structures in action and thought. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of child psychology: Theoretical models of human development*. John Wiley & Sons, Hoboken, NJ, pp. 467–561.
- Fischer, K.W., Yan, Z., 2002. Darwin’s construction of the theory of evolution: microdevelopment of explanations of variations and change in species. In: Granott, N., Parziale, J. (Eds.), *Microdevelopment: Transition Processes in Development and Learning*. Cambridge University Press, Cambridge, UK, pp. 294–318.

- Fischer, K., Yan, Z., Stewart, J., 2003. Adult cognitive development: dynamics in the developmental web. In: Valsiner, J., Connolly, K.J. (Eds.), *Handbook of Developmental Psychology*. Sage, Thousand Oaks, CA, pp. 491–516.
- Gottlieb, G., Wahlsten, D., Lickliter, R., 2006. The significance of biology for human development: a developmental psychobiological systems view. In: Lerner, R.M. (Ed.), *Theoretical Models of Human Development*, *Handbook of Child Psychology*, sixth ed., vol. 1. John Wiley & Sons, Inc, Hoboken, NJ, pp. 210–257.
- Gottlieb, G., 2003a. On making behavioral genetics truly developmental. *Hum. Dev.* 46, 337–355.
- Gottlieb, G., 2003b. Probabilistic epigenesis of development. In: Valsiner, J., Connolly, K.J. (Eds.), *Handbook of Developmental Psychology*. Sage, Thousand Oaks, CA, pp. 3–17.
- Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7 (1), 6–10.
- Gould, S.J., Eldredge, N., 2002. Punctuated equilibrium and the validation of macroevolutionary theory. In: Gould, S.J. (Ed.), *The Structure of Evolutionary Theory*. The Belknap Press, Cambridge, MA, pp. 742–1022.
- Granott, N., Parziale, J. (Eds.), 2002. *Microdevelopment: Transition Processes in Development and Learning*. Cambridge University Press, Cambridge, UK.
- Granott, N., 2002. How microdevelopment creates macrodevelopment: reiterated sequences, backward transitions, and the zone of current development. In: Granott, N., Parziale, J. (Eds.), *Microdevelopment: Transition Processes in Development and Learning*. Cambridge University Press, Cambridge, UK, pp. 213–242.
- Hanno, E.C., Jones, S.M., McCoy, D.C., 2020. The joint development of literacy and self-regulation in early childhood. In: Moje, E.B., Afflerbach, P.P., Enciso, P., Lesaux, N.K. (Eds.), *Handbook of Reading Research*, vol. 5. Routledge, New York: NY, pp. 279–306.
- Heckman, J.J., 2005. Commentary. In: Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., Nores, M. (Eds.), *Lifetime Effects: The High Scope/Perry Preschool Study Through Age 40*. High Scope Press, Ypsilanti, MI, pp. 229–233.
- Hildreth, G., 1936. Developmental sequences in name writing. *Child Dev.* 7 (4), 291–303.
- Hoover, W.A., Tunmer, W.E., 2018. The simple view of reading: three assessments of its adequacy. *Remedial Spec. Educ.* 39 (5), 304–312.
- Horwich, P., 2010. *World Changes: Thomas Kuhn and the Nature of Science*. University of Pittsburgh Press, Pittsburgh, PA.
- Institute for Figuring, (n.d.). Crochet Model of Hyperbolic Plane by Daina Taimina. <https://theif.org/oe/exhibits/oe1e.html>.
- Invernizzi, M., Sullivan, A., Meyer, J., Swank, L., 2004. *PreK Teachers Manual*. Virginia Dept. of Education, University of Virginia, Curry School of Education, Charlottesville, VA.
- Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, IL.
- Kuhn, T.S., 2000. *The Road Since Structure*. University of Chicago Press, Chicago, IL.
- Kwak, J., 2006. *Writing Development of Korean Heritage Children* (Order No. 3237174).
- Levin, I., Both-De Vries, A., Aram, D., Bus, A., 2005. Writing starts with own name writing. From scribbling to conventional spelling in Israeli and Dutch children. *Appl. Psycholinguist.* 26, 463–477.
- Luke, D.A., Stamatakis, K.A., 2012. Systems science methods in public health: dynamics, networks, and agents. *Annu. Rev. Publ. Health* 33, 357–376. <https://doi.org/10.1146/annurev-publhealth-031210-101222>.
- Luria, A.R., 1988. The pathology of grammatical operations. In: Downing, J. (Ed.), *Advances in Psychology, Cognitive Psychology and Reading in the U.S.S.R.*, vol. 49. North Holland Publishing Company, Amsterdam, NL, pp. 95–141. [https://doi.org/10.1016/S0166-4115\(08\)60372-6](https://doi.org/10.1016/S0166-4115(08)60372-6).
- Mirzai, A., 2002. *Development of Orthographic Knowledge in a Consonantal Script: Children's Invented Spelling in Farsi* (Production No. 3073817). Doctoral dissertation. University of Southern California. <https://www.proquest.com/pqdtglobal/docview/305523542/B58B6348FF244E4BPQ/1?accountid=8360>.
- Olson, D., 1994. *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading*. Cambridge University Press, Cambridge, UK.
- Overton, W.F., Molenaar, P.C.M., 2015. Concepts, theory, and method in developmental science: a view of the issues. In: Overton, W.F., Molenaar, P.C.M. (Eds.), *Handbook of Child Psychology and Developmental Science, Theory and Method*, vol. 1. John Wiley & Sons, Inc, Hoboken, NJ, pp. 1–8.
- Overton, W.F., 2015. Processes, relations, and relational-developmental systems. In: Overton, W.F., Molenaar, P.C.M. (Eds.), *Handbook of Child Psychology and Developmental Science, Theory and Method*, vol. 1. John Wiley & Sons, Inc, Hoboken, NJ, pp. 9–62.
- Paris, S.G., 2005. Reinterpreting the development of reading skills. *Read. Res. Q.* 40 (2), 184–202.
- Petscher, Y., Cabell, S.Q., Catts, H.W., Compton, D.L., Foorman, B.R., Hart, S.A., Lonigan, C.J., Phillips, B.M., Schatschneider, C., Staacy, L.M., Nicole Patton, T., Wagner, R.K., 2020. How the science of reading informs 21st-century education. *Read. Res. Q.* 55, S267–S282.
- Piaget, J., Garcia, R., 1989. *Psychogenesis and the History of Science*. Columbia University Press, New York, NY.
- Piaget, J., 1977. *The Development of Thought: The Equilibration of Cognitive Structures*. Viking Press, New York, NY.
- Reichle, E.D., 2021. *Computational Models of Reading: A Handbook*. Oxford University Press, New York, NY.
- Rowe, D.W., 2015. The development of a descriptive measure of early childhood writing: results from the Write Start! Writing assessment. *J. Lit. Res.* 47 (2), 245–292. <https://doi.org/10.1177/1086296X15619723>.
- Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., Nores, M., 2005. *Lifetime Effects: The High Scope/Perry Preschool Study Through Age 40*. High Scope Press, Ypsilanti, MI.
- Seidenberg, M., 2017. *Language at the Speed of Sight. Why We Read, Why So Many Can't, and What Can Be Done About It*. Basic Books, New York.
- Siegler, R.S., Crowley, K., 1991. The microgenetic method: a direct means for studying cognitive development. *Am. Psychol.* 6 (5), 606–620.
- Siegler, R.S., 1995. How does change occur: a microgenetic study of number conservation. *Cognit. Psychol.* 28, 225–273.
- Siegler, R.S., 1996. *Emerging Minds: The Process of Change in Children's Thinking*. Oxford University Press, New York, NY.
- Siegler, R.S., 2006. Microgenetic analyses of learning. In: Kuhn, D., Siegler, R.S., Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology: Cognition, Perception, and Language*. John Wiley & Sons Inc, pp. 464–510.
- Slaughter-Defoe, D.T., 2005. Commentary. In: Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., Nores, M. (Eds.), *Lifetime Effects: The High Scope/Perry Preschool Study Through Age 40*. High Scope Press, Ypsilanti, MI, pp. 235–239.
- Tardibuono, J., 2006. *A Psychogenetic Analysis of Spanish-Speaking Preschoolers' Emergent Writing: A Developmental Trajectory Toward the Phonetization of Speech* (Publication No. 3238342). Doctoral dissertation. University of Southern California. <https://www.proquest.com/docview/305277930/B06DE7579E894629PQ/2?accountid=8360>.
- Thelen, E., Corbetta, D., 2002. Microdevelopment and dynamic systems: applications to infant motor development. In: Granott, N., Parziale, J. (Eds.), *Microdevelopment: Transition Processes in Development and Learning*. Cambridge University Press, Cambridge, UK, pp. 59–79.
- Thelen, E., Smith, L.B., 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press, Cambridge, MA.
- Tomasello, M., 2019. *Becoming Human: A Theory of Ontogeny*. Harvard University Press, Cambridge, MA.
- Tracey, D.H., Morrow, L.M., 2017. *Lenses on Reading: An Introduction to Theories and Models*, third ed. The Guilford Press, New York, NY.
- Tsai, T., 2007. *A Microgenetic Analysis of English/Chinese Early Writing Development* (Order No. 3291823). Available from ProQuest Dissertations & Theses Global. (304824128). Retrieved from: <https://ezproxy.library.arizona.edu/login?url=https://www.proquest.com/dissertations-theses/microgenetic-analysis-english-chinese-early/docview/304824128/se-2?accountid=8360>.
- Valsiner, J., 1998. The development of the concept of development: Historical and epistemological perspectives. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of child psychology: Theoretical models of human development*. John Wiley & Sons, Hoboken, NJ, pp. 189–232.
- Valsiner, J., Connolly, K.J., 2003. The nature of development: the continuing dialogue of processes and outcomes. In: Valsiner, J., Connolly, K. (Eds.), *Handbook of Developmental Psychology*. Sage, Thousand Oaks, CA, pp. ix–xviii.
- van der Veer, R., Valsiner, J., 1991. *Understanding Vygotsky: A Quest for Synthesis*. Blackwell, Oxford, UK.
- van Geert, P., 2003. Dynamic systems approaches and modeling of developmental processes. In: Valsiner, J., Connolly, K. (Eds.), *Handbook of Developmental Psychology*. Sage, Thousand Oaks, CA, pp. 640–672.
- van Geert, P., 2017. *The Development of Perception, Cognition and Language: A Theoretical Approach*. Routledge, London, UK.

- Vygotsky, L.S., 1987. In: Rieber, R.W., Carton, A.S., Minick, N. (Eds.), *The Collected Works of L. S. Vygotsky, Problems in General Psychology, Including the Volume Thinking and Speech*, vol. 1. Plenum Press, New York, NY.
- Vygotsky, L.S., 1997. The prehistory of written language. In: Rieber, R., Hall, M.J. (Eds.), *The Collected Works of L. S. Vygotsky, Vol. 4: The History of the Development of Higher Mental Functions*. Springer, Boston, MA, pp. 131–148. https://doi.org/10.1007/978-1-4615-5939-9_7.
- Waddington, C.H., 1957. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. George Allen & Unwin, Ltd, London, UK.
- Whitehurst, G., Lonigan, C., 2002. Emergent literacy: development from prereaders to readers. In: Neuman, S.B., Dickinson, D.K. (Eds.), *Handbook of Early Literacy Research*. Guilford, New York, NY, pp. 11–29.
- Yaden Jr., D.B., Tardibuo, J., 2004. The emergent writing development of urban Latino preschoolers: a socio-psychogenetic trajectory. *Read. Writ. Q.* 20, 1–33.
- Yaden Jr., D.B., Tardibuo, J., 2011. The emergent writing development of urban Latino preschoolers: a reanalysis of data. In: Filipo, R. (Ed.), *In Search of Common Ground*. Taylor and Francis, New York, NY.
- Yaden Jr., D.B., Tsai, T., 2012. Learning how to write in English and Chinese: young kindergarten and first grade children explore the similarities and differences in writing systems. In: Gort, M., Bauer, E. (Eds.), *Early Bilingual Development: Exploring Young Learners' Use of Their Linguistic Resources*. Routledge, New York, NY, pp. 55–84.
- Yaden Jr., D.B., Rueda, R., Tardibuo, J., Martinez, C., Mirzaei, A., Scott-Weich, B., Tsai, T., 2014. Interviewing young children: using psycho- and microgenetic design methodology to assess children's understandings of reading and writing: the promise and challenge. In: Saracho, O., Spodek, B. (Eds.), *Handbook of Research Methods in Early Childhood Education*, vol. II. Information Age Publishing, Charlotte, NC, pp. 309–348.
- Yaden Jr., D.B., Gort, M., Martinez, C., Rueda, R., 2019. The education of young emergent bilingual children: an update and call for action. In: Saracho, O. (Ed.), *Handbook for Research on the Education of Young Children*. Routledge, New York, NY, pp. 234–253.

A naturalized view of literacy education: what the neuro- and life sciences may offer language and literacy research

George G. Hruby^a and Ayan Mitra^b, ^a Collaborative Center for Literacy Development, University of Kentucky, Lexington, KY, United States; and ^b BrainLENS Laboratory (UConn-UCSF), Department of Psychological Sciences, Hoefft Laboratory for Learning Engineering and Neural Systems, University of Connecticut, Storrs, CN, United States

© 2023 Elsevier Ltd. All rights reserved.

Useful knowledge about the neurophysiology of reading	867
Technical issues	867
A narrow-gauge synthesis of neuroanatomy related to reading	868
Decoding	869
Comprehension	870
Additional neural factors relatable to reading	872
Neuroscience research related to reading difficulties and dyslexia	873
Neurobiology of reading	874
Literacy education naturalized: a wider-gauge view	875
Finding a place for new realism in literacy research	877
References	878

What can cognitive neuroscience contribute to an understanding of literacy and literacy education (in the general sense used by school and teacher educator faculty, entailing language, reading, writing, and other systems of symbolic representation for communication, learning, and thought)? To date, the answer is not as much as some may believe, particularly advocates for a brain-based education (O'Mahoney, 2021; Sousa, 2014) or a science of reading (Burkins and Yates, 2021; Seidenberg et al., 2020). To justify our restrained assessment, we review mainstream research findings and theoretical constructs from the neuro- and related life-sciences in hopes of clarifying misunderstandings about reading-relatable neuroscience, misunderstandings often shared and reiterated by teacher educators, journalists, marketers, and educational policy advocates. To better inform educational applications, we propose reliance on a complementary and comprehensive life-science narrative about literacy development.

Specifically, we propose a distinction between narrow-gauge and wide-gauge pathways for connecting neuroscience to education. Narrow-gauge approaches rely on compellingly arranged facts, claims, and findings from cognitive neuroscience to justify theoretical constructs from cognitive educational psychology (e.g., as found in brain-based education and educational neuroscience; cf. Price, 2017; Thomas et al., 2020). This foreshortened narrow-gauge view is well disseminated and seems unproblematic for non-specialists to grasp in part because the claims are very familiar, having been made by cognitive theorists in education and psychology for nearly 70 years (Hoffman and Deffenbacher, 1992; Pearson and Cervetti, 2013). Syntheses of this work, as we shall include in this review, are certainly of value. But cognitive models are not without their critics and in the hands of non-specialists can act as rhetorical springboards for interpretations requiring a broader understanding for adequate justification (e.g., for educational policy or decisions on instructional method or curricular design). Too often the result are claims about "scientific" reading or learning based on assumptions of abbreviated isomorphism suggesting a computational physiology improbably shorn of biological constraint or socio-emotional intention, and further confounded by epistemological confusions and theoretical incoherence (Busso and Pollack, 2014; Howard-Jones, 2014; Hruby, 2012).

In contrast, we hold that the neurosciences can suggest a great deal about what it means to be a neurologically endowed and symbolically capable organism when neuroscience is synthesized within a wider array of life sciences (e.g., genetics, cytology, endocrinology, systems biology) and theory from life science-adjacent fields (e.g., developmental science, cognitive ethology, evolutionary psychology, philosophy of mind or neuro-philosophy). Such syntheses make cohesive sense of neurobiological processes in context over time and how they can give rise to symbolic communication as in natural human language and literacy. A fuller neurological view would include sufficient life science background and metatheoretical support to clarify, among other things, how nervous and endocrine systems develop, how they evolved in the first place, what constraints they impose and affordances they allow, how symbolic representation as a communicative form can emerge within those constraints, and thereby how *homo sapiens* evolved to develop natural language, symbol use, and the cultural affordances of literacy. Conversely, the mere charting of the functional neuroanatomy of the cerebral cortex by the light of cognitive models of mental process is insufficient to warrant an understanding of how humans employ literacy beyond what literacy researchers relying on cognitive models understand already.

What we encourage, then, is a naturalized view of literacy grounded by conciliatory assumptions from the life sciences useful for educational applications (cf., Gee, 2021; Sigman et al., 2014). Conjoining life-science informed insights about living organisms and systems with mainstream scholarship in literacy education could provide the latter with a refreshment some have argued is needed

(Reinking and Yaden, 2020). We acknowledge that, at this early juncture, theoretical advance is more likely than empirically demonstrated instructional guidance, but for the same reasons more necessary, preliminary in both senses of the word, formative and prefatory.

Useful knowledge about the neurophysiology of reading

We begin with a focus on reading processes and their development as an example applicable to literacy education more generally. Reading processes (what goes on in the mind when a reader reads) have long been a central focus in literacy education research and have been central to psychological research since its inception (e.g., Huey, 1908). As a result, research and theory on reading processes, contexts, development, and instruction is richly diverse, as a review of the field's major handbooks will demonstrate (e.g., Alvermann et al., 2019; Christenbury et al., 2009; Israel, 2017; Moje et al., 2020; Neuman and Dickinson, 2011; Rowsell and Pahl, 2015). Reading processes have been defined statistically by behavioral factor analyses since the 1940s, parsed by models from cognitive psychology since the 1960s, and theorized by analogy with information processing systems since the 1980s (Alexander and Fox, 2019). Since the 1990s, cognitive neuroscience has focused on neurophysiological correlates to functions identified in those earlier factor analytic and cognitive models of reading (Goswami, 2015; Hruby, 2017; Hruby and Goswami, 2011, 2019).

Meanwhile, the sociocultural turn in educational research since the 1980s has broadened the study of reading instruction to include writing instruction, language and second-language development, disciplinary literacies, technology-related and popular culture literacies, and the effect of socio-emotional, societal, economic, cultural, political, and ecological contexts for literacy learning (e.g., Alvermann, 2009; Gee, 2001, 2013; Gillen, 2015; O'Brien and Rogers, 2016; as well as the panorama of topics and perspectives in this volume). Tied tightly to theoretical models of computational cognitive process, the narrow-gauge track between cognitive neuroscience and literacy instruction has appeared disengaged from if not unsympathetic to this sociocultural turn. By contrast, wider-gauge educational neuroscience informed by naturalistic understandings of human development could be more complementary to literacy's sociocultural extension.

Technical issues

The techniques used to correlate neural activity with mental processes, such as positron emission tomography (PET), magnetic resonance imaging (MRI), or electroencephalographic (EEG) techniques, date to the 1960s, but strategic combination and technological refinement, in tandem with ongoing advances in computational power, have led to more precise contemporary methods. One result has been brain imaging: the vibrant rainbow-colored illustrations of functional neuroanatomical correlates that today adorn research reports, magazine covers, and popular websites. Non-specialists can be impressed by such images without fully understanding what they represent. Nonetheless, they are persuasive. Researchers have found that scientific reports illustrated with brain images are more compelling than the same reports with data tables or bar graphs (McCabe and Castel, 2008), and that non-specialist readers find psychological reports that refer to neuroscience more believable than the same that do not, even when the neurological explanations are irrelevant, contradictory, or illogical (Weisberg et al., 2008). "Neuromyths" held by educators have been an ongoing concern for neuroscientists and educational sociologists alike (Hughes et al., 2020; MacDonald et al., 2017; Worthly et al., 2021).

Non-specialist should know that brain study imaging is only superficially akin to X-rays, ultra-sounds, or medical MRIs. Contrary to common parlance, the brain does not "light up," and brain images are not photographs of the brain in action. The illustrations are in fact color-coded statistical charts from data sets averaged across research participants usually subtracted between two conditions and the result overlain on a normative model of a human brain. There are several of these normative brain models available, and their identification of regions, gyri, and sulci somewhat differ. Moreover, just as everyone has a face, but not the same face, everyone has a brain, just not an identical one. To get the data from different subjects in brain imaging studies to align to the coordinates of a normative model, a certain amount of manipulation is required. In other words, at first blush, brain charts give a misleadingly consistent impression of what "the brain"—implying all brains—look like and do when, in fact, there is a great deal of variation in neurophysiological anatomy and its functional purposes between and within human subjects that by necessity is obscured in the final images. When distilled by popular media for wider dissemination, a view of functionally necessary human homogeneity, both false and unhelpful, is promulgated. Critical consumers of these images need to remember they are as much the product of the models, methods, and technologies employed to produce them as they are of the underlying phenomena being investigated.

The blurring between what can be discerned from subject set averages and the characteristics found in specific individuals is notable in developmental studies, as well. Developmental trajectories, at different scales of physiological, functional, and chronological analysis, differ between and within individuals over time, a fact occluded by averaging subgroups in search of robust trend-lines. Brains are generally structured and highly plastic at birth but take on their effective functionality through learned response to their environments throughout the lifespan, and both environments and initiating (and ongoing) conditions in the brain differ between individuals and diversify over the lifespan. The result is that equally functional behaviors are possible from structurally

“irregular” brains, and seemingly similar brains can be surprisingly diverse in their functionality and behaviors. This is clearly the case when reviewing the changes that can occur in the brain after successful reading instruction or intervention (Barguero et al., 2014; Chyl et al., 2018; Kukona et al., 2016; Meyler et al., 2008).

As impressive as the technical advance of imaging technologies has been, precision is an ongoing vexation (Churchland and Abbott, 2017). Magneto- and electroencephalographic techniques are more precise temporally, parsing the electrical or magnetic field signature of neural activity to 100ths of a second, but their spatial precision is poorer than other methods, and deteriorates as the signal source moves away from the surface of the cerebral cortex. Hemodynamic methods, such as fMRI and PET, are more precise spatially but are less precise temporally, capturing changes in glucose or oxygen uptake in blood flow within brain regions of interest (ROIs) over nearly a second or more (a figurative eon in the world of neural activity). Results on similar experimental conditions from these two categories of technique often mismatch as neither measure neural activity directly, and spurious (non-neural) activity in the brain can blur results further (Jorge et al., 2014; Rosa et al., 2010). However, newer methods are improving the ever-evolving state of the science, through multiple matched comparisons (e.g., Bonhage et al., 2015; Koch et al., 2016; McCormick, 2021), or with newer techniques suitable for human study such as diffusion tensor imaging of white matter tracts (Cummine et al., 2015), transcranial stimulation (Rothwell, 2018), or electrode implantation (Themistocleous et al., 2017).

Areas of activation that are equally necessary to both conditions in a subtractive design will disappear in the final image; the differential is only of areas *uniquely* active between the two conditions. This is understood by neuroanatomists, but the distinction is often lost on non-specialists who believe what is being visualized is, for instance, “where in the brain we read letters,” etc. And the more highly specific the anatomical loci of a functional dedication reported in a study, the more likely the result is a false positive. Statistically, a great deal of noise reduction is necessary to create images representing neatly modularized regions of activation, and the crisp modular boundaries typically presented in brain images are, again, more an artifact of the methods and technologies employed than they are of the underlying phenomena being explored. Setting noise thresholds involves a degree of creative approximation guided by the search for a best-match and compelling image. The percentage of participants whose data are eliminated from analysis as outliers (failing to average neatly with the majority) can run as high as 40%, drastically reducing already arguably inadequate subject set sizes (Marek et al., 2022). Most studies finish out with single- or low double-digit ns (approximately 25 on average; Szűcs and Ioannidis, 2020). Concerns over study design and statistical inadequacies have an ongoing history within the neurosciences (e.g., Button et al., 2013; Frégnac, 2017; Poldrack et al., 2017; Ramus et al., 2018), with current estimation of false positives reported in published cognitive neuroscience studies running well over 50% (Szűcs and Ioannidis, 2017, 2020). This can partially explain lack of successful study replication, the high variability of loci correlated with specific processes or functions, and why the results of metaanalysis or careful synthesis are so diverse (Marek et al., 2022).

Although the issues outlined above require a more sober assessment of neuroscience related to reading, certain *general* results are well demonstrated. Our intention in noting these issues is not to discourage belief in the potential value of neuroscience for informing fresh narratives about the nature of reading, literacy, or symbolic thought in education, but to underscore the potential, now broadly realized, for vernacular misunderstanding, misappropriation, and misuse wherein the pretense of evidence is employed to promote possibly inadequate or even harmfully inequitable literacy policies, programs, and products (Compton-Lilly et al., 2020; Hruby, 2012; Vandenbroeck, 2017). Turning educators into neuroscience experts by inundating them with a welter of titillating images and factual tidbits is an unlikely and unnecessary proposition. We propose a more narratively cohesive approach, synthesizing a theoretical view of human learning informed by life and developmental science and life science-informed philosophy of mind and of language to connect the sciences of human change (including neuroscience) with current theory and practice in literacy education.

A narrow-gauge synthesis of neuroanatomy related to reading

To provide an example of a narrow-gauge review of mainstream findings from cognitive neuroscience related to reading processes, we begin with the Simple View of Reading (Gough and Tunmer, 1986), which distinguishes decoding processes (D) from language comprehension processes (C) in the measurement of reading comprehension (R; i.e., reading as measured on a standardized reading comprehension test). Reading is thus said to be the product of students’ decoding skills and their language comprehension ability. That is, $D \times C = R$. As a result, adequate performance on R requires adequate performance in both D and C, although, as research has revealed, much else is also required (cf., Hoffman et al., 2020; Duke and Cartwright, 2021). Literacy researchers and teacher educators traditionally subdivide decoding to print concept, alphabet knowledge, phonemic awareness, phonics accuracy, sight word reading, and fluency (Gough, 1972; Shanahan, 2005; Stahl et al., 2020). Language comprehension has been subdivided into language familiarity, symbolic and strategic reasoning, prior knowledge, and sociocultural or linguistic environment (Hruby, 2020; Pearson et al., 2020). Neuroscience studies have demonstrated correlations between most of these categories, or their subdivisions, and specific cerebral loci.

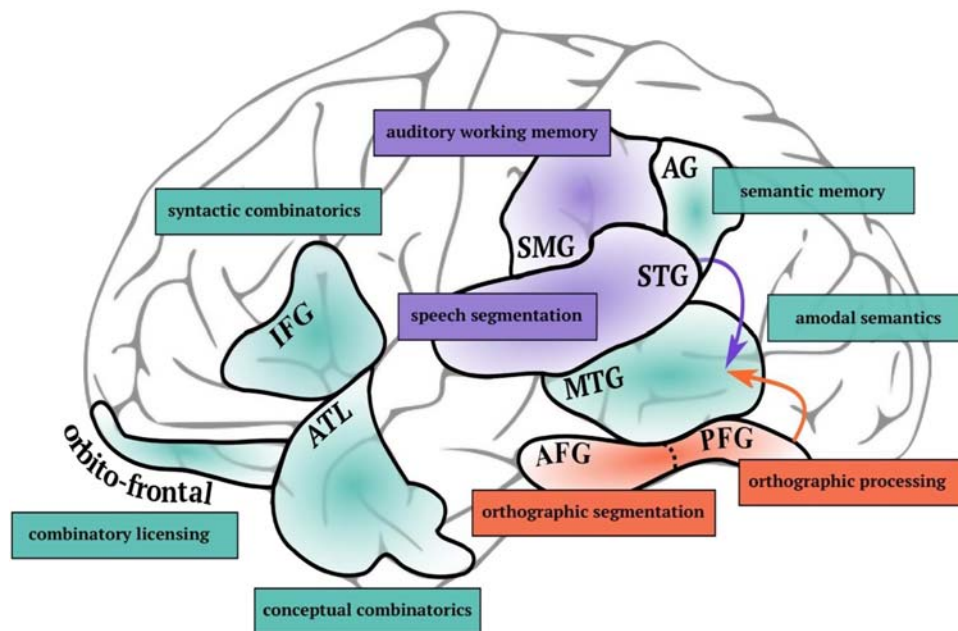


Fig. 1 Putative brain regions associated with different stages of morphological processing. Orange areas refer to modality-specific written word processes. Purple areas to modality-specific spoken word processes. Turquoise refers to putative a-modal processes. Key: PFG, posterior fusiform gyrus; AFG, anterior fusiform gyrus; STG, superior temporal gyrus; SMG, supramarginal gyrus; AG, angular gyrus; MTG, middle temporal gyrus (although here rather posterior); ATL, anterior temporal lobe; IFG, inferior frontal gyrus. From Gwilliams (2019), depicting possible language and reading loci of left-side of the cerebral cortex.

Decoding

Text decoding processes, wherein sequences of letter shapes on a page or screen are interpreted or recognized to known word forms, follow two distinct trajectories in the brain, both beginning in the occipital lobe where visual processing for letter shape identification occurs (Paz-Alonso et al., 2017; Rueckl et al., 2015). The first route traces a dorsal (upper) trajectory through the left parietal lobe in an arc toward the left angular gyrus (AG; see Fig. 1 for acronym loci) and supramarginal gyrus (SMG) at the temporal-parietal transition. The SMG is where neural activation correlates with study subjects sounding out letters or letter groupings to identify auditory word forms (Oberhuber et al., 2016) triggering recognition of a sound pattern in the left superior temporal lobe (STG) with a neural nodal nexus for word meaning into the left medial temporal gyri (MTG). Readers who demonstrate difficulty sounding out words fluently demonstrate atypical activation in the SMG and loci downstream along this trajectory (Gori et al., 2014).

The second decoding route traces a ventral (lower) trajectory, again, from the occipital lobe at the back of the brain for visual letter shape processing, along the inferior occipital-temporal sulcus, parallel to the inferior longitudinal fasciculus, through the fusiform gyrus (AFG and PFG in Fig. 1), a region often dubbed the *visual word form area* (VWFA), or *the brain's letterbox* (Dehaene, 2009). The VWFA is adjacent to areas for identifying faces and common objects (presumably by selecting for higher salience characteristics in a perceived visual pattern or learned prototypical sequence; Vogel et al., 2014). This is where orthographic patterns are recognized as "sight words," or familiar letter sequences, before routing forward into the temporal lobe (MTG, ATL), again, to activate a word's nodal nexus for meaning (Dehaene et al., 2015). Whereas emergent readers rely more on the dorsal-auditory route as they master their decoding skills, possibly facilitated by being explicitly taught to do so in school through some form of phonics instruction, fluent readers, such as literate adults, typically rely on the more efficient and rapid ventral-orthographic route. Readers thus demonstrate a shift from relying on the dorsal (phonemic processing) route to the ventral (orthographic processing) route for most word-form decoding as they mature into fluent readers and demonstrate increased reliance on sight words (Martin et al., 2015).

Average intermediate grades readers are also more likely to activate the ventral route compared to peers who demonstrate persistent decoding difficulty (Brem et al., 2020; Huettig et al., 2018; Martin et al., 2016). The absence of direct neural connection between the SMG and the VWFA suggests that fluent sight word reading is not the sounding out of letters so automatically a reader does not know they are doing it. Rather, occipital-temporal activation appears to be a form of visual pattern or sequence recognition similar to facial recognition or familiar object recognition, both of which activate adjacent and right analogous areas of the brain. Just how this works at the level of neural tissues and networks is a matter of speculation (Chen et al., 2019; Goswami, 2019; Gwilliams, 2019). It is possible the recognition of visual patterns relies on neural assemblies similar to those for sound-pattern parsing of rhythms, syllables, or morphemes (De Vos et al., 2017; Goswami, 2020; Verhoeven and Perfetti, 2011).

In the span of milliseconds, the subvocalization of sight words occurs *after* word-level comprehension is inaugurated, even as syntactic and semantic integration is engaged (although these activations seem simultaneous in conscious awareness). [Chen et al. \(2019\)](#) found that the VWFA significantly connects not only to language processing areas by way of the temporal lobe, but also directly to areas in the superior parietal cortex for quickly directing nonconscious attention, possibly facilitating pattern anticipation and recognition to most probable (i.e., coherent) perceptual pattern confirmation. Developing readers likely employ both dorsal and ventral routes redundantly for some time, using the dorsal route as “training wheels” to scaffold the ventral route until it can handle run-of-the-mill word identification needs effectively ([Martin et al., 2015](#)). That is, teaching children to sound out every letter of every word (in the SMG) is how teachers unintentionally foster children’s overlearning of the orthographic regularities (in the VWFA) for automatized recognition of print.

Activation through developed white matter tracts (bundled and myelinated neural pathways well matured by use and thus efficient in their functions; ([Wandell and Yeatman, 2013](#))) occurs within a matter of milliseconds and runs in both directions along a fascicular tract: every bit as much from frontal and superior areas of the brain for attention, interpretation, and meaning, to posterior and inferior areas for perception and sensation as in the opposite direction (from the “start” of reading) in sensation and perception forward to word meaning, constructed scenario, and interpretation ([Cartwright, 2012](#)). Possibly due to cultural traditions privileging empiricism over rationalism in anglophone epistemology, self-reflection by members of that tradition, by whom most of the research of this type has been done, enjoy the confident illusion that subvocalization of sight word texts leads to, rather than is the result of, a search for context-relevant meaning. But in coordinating a conscious response to a text, a great deal of back and forth occurs between lower-order (and fairly modular) nonconscious processes across the brain ([Liebig et al., 2017](#)). The infinitesimal timescale of these neural processes should not be confused as analogs for conscious word attack strategy or sequencing of reading instruction over weeks to years.

Comprehension

Literacy educators often refer to *comprehension* as a solitary construct, one of the so called “Five Pillars of Reading” (along with phonemic awareness, phonics, fluency, and vocabulary; [Shanahan, 2005](#); [Cassidy et al., 2010](#)). Literacy researchers have extensively explored four language comprehension subcategories: (a) *language familiarity* (e.g., knowledge of vocabulary, common grammatical structures, idioms, tropes, discourse features); (b) *symbolic and strategic reasoning ability* (e.g., lower order abilities, such as comparison, categorization, analogy, inference, determination of priority, as well as higher order abilities for strategic problem solving, such as text-level assessment, questioning, summarization, graphic analysis, reviewing, and discussion with others); (c) *knowledge base or prior knowledge* (of self, sociality, world, content, “common sense,” culturally normative beliefs, etc.); and (d) *linguistic or sociocultural environment* (from which a language user learns their language, and in which they employ and practice it to intentional, normative, and self-correcting ends) ([Hruby, 2020](#); [Pearson et al., 2020](#)).

For a first pass on how language comprehension develops (see [Fig. 2](#)), this four-component view—literate environment, language familiarity, symbolic reasoning ability, and knowledge base—may seem sufficient, although the model would be profitably extended with the incorporation of socio-emotional intention, metacognitive motivations, and agentive identity (see [Hoffman et al., 2020](#)). But as it is, the subprocesses contained within each category complicate the construction of cognitive neuroscience study designs. Parsing reality (e.g., anatomical function) “at the joints” (i.e., locating areas of activity in the cerebral cortex relevant for comprehension subprocesses) is tricky and can too easily become self-fulfilling prophecies. Studies employing general categories from linguistics has led to areas identified as relevant for lexical, syntactic, and semantic processing ([Hruby and Goswami, 2019](#)). Studies of consciousness and meta-cognition, presumed requirements of comprehension in school settings, locates multiple areas elsewhere ([Carreteiro and Figueiro, 2017](#); [Pace et al., 2020](#)) (For the theoretical quagmires lurking in the topic of consciousness see [Chalmers, 1996](#); [Koch et al., 2016](#); [Wendt, 2015](#)).

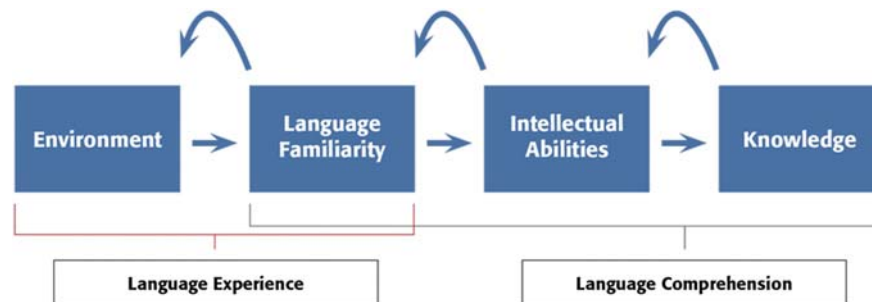


Fig. 2 With each step in the trajectory, left to right, the development of a category supports others, such that the development of language use by a child adds to the density and complexity of the linguistic environment from which it learns language; which in turn fosters the development of the child’s symbolic reasoning (language being the first and primary symbol system a child must master), thereby enriching the language familiarity and linguistic environment from which it develops; finally, symbolic reasoning, language, and environment develop the child’s knowledge base, for humans do not merely use language and reason symbolically, they do so about matters, those matters becoming knowledge. The categories are not stages, but an ongoing set of nested and recursive developmental processes ([Hruby, 2020](#)).

Because the theories, models, methods, and research designs employed in this work have as much to do with the apparent functionality of the loci as the underlying activity being studied, the current neuroscience literature on *reading comprehension* lends an unfortunate “now you see it, now you don’t” impression to which subcomponents matter, when, and where. The imprecision of localization methods is amplified by the functional and anatomical foci being too broadly bounded to indicate the neural connections and networks involved or the structure of their “processing,” but also too tightly bounded to be inclusive of all the multiple processes operating at differing scales of structure and time required for conscious understanding. Although methodological reductionism in parsing the brain’s complexity is necessary, it may inadvertently lead to unjustified theoretical reductionism.

The sharply modular representation of brain loci created by imaging technologies can be misleading. For instance, vocabulary definitions are not “stored” or “recorded” in the left medial temporal lobe, as shorthand reference may suggest. Word meaning involves networks of neural lacing across many areas of the brain including for sensation, motor activity, emotional arousal and control, executive strategy, and even sub-cerebellar areas for emotion in the limbic lobe, cerebellum, and brain-stem (e.g., Alvarez and Fiez, 2018; Damasio, 2021; Venkatraman et al., 2017). These neural nettings are insufficient in volume to appear in whole brain imaging studies, but they converge in the temporal as nodal clusters large enough to be reported as significant, hence the “localization” of word meaning in the temporal regions. Similarly, a focus on any cognitive process will image correlated activity unique for that process, while equally important factors and areas shared with the comparison condition will go missing.

Taking these warnings to heart, language comprehension processes trace through two main routes as with decoding: an upper (dorsal) pathway and a lower (ventral) pathway (see Fig. 3). The dorsal pathway traces from speech sound processing (SMG) to syntactic processing in the left inferior frontal gyrus (IFG), possibly for analyzing morphemic grammatical markers, and to a more varied dorsolateral set of areas for semantics and situation modeling. The ventral pathway traces from the medial temporal and adjacent areas (activated for lexical access or word meaning) into the IFG and adjacent dorsolateral frontal areas for syntactic and semantic integration (Catinelli et al., 2013; Walenski et al., 2019).

Neuroscientists have dependably and repeatedly correlated activation in the left temporal lobe with lexical recognition and word comprehension (Schuster et al., 2021), left inferior frontal gyrus (LIFG) with syntactic tasks (Henderson et al., 2016), and dorso-lateral and orbito-frontal, medial cingulate, and superior parietal cortex with higher-order and a-modal semantic processing (Deniz et al., 2019). Still, how neural activity incorporates these levels of language, integrating lexical meaning and syntactic structure into sentence-level comprehension, is far from established (Arana et al., 2020; Bonhage et al., 2015; Gwilliams, 2019; Hagoort, 2019; Takashima et al., 2020).

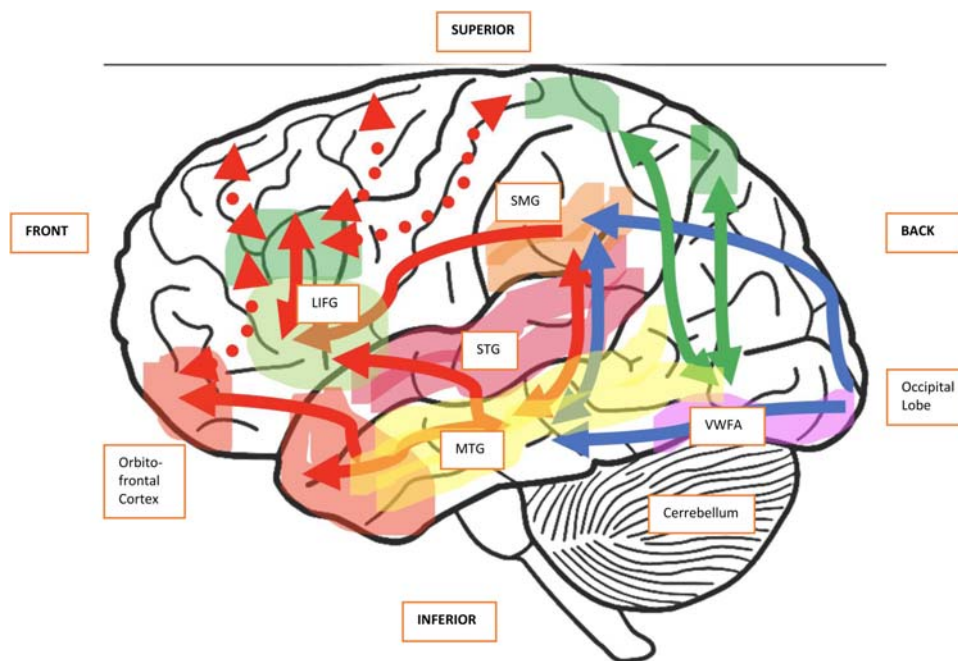


Fig. 3 Potential cerebral pathways and ROIs for reading in left cerebral cortex; Blue arrows—decoding pathways; solid red arrows—language comprehension pathways [vocabulary, grammar, clausal meaning]; dotted red arrows—a-modal comprehension pathways; green arrows—semantic and scenario pattern prediction and confirmation pathways; darker green areas: semantics and scenario; lighter green area: syntax; red-orange areas: emotional meaning associations; magenta area: word sound patterns; orange area: auditory processing; periwinkle: orthographic processing.

In comparison to the comprehension related research, the use of cognitive models to guide interpretation of neural activation during decoding seems less problematic, but this illustrates again how theoretical framing drives the interpretation of data. Because decoding is described as symbol identification and association, and because symbols map easily as artifacts of a computational processing system (computers being quantified symbol re-shufflers and modifiers), the central metaphor assumed in cognitive process modeling of the mind as an information processing system appears confirmed in decoding studies (Brette, 2019; Martin and Baggio, 2019). Word prediction and grammatical structure can be parsed similarly as pattern recognition of syntactic and semantic sequences (e.g., as demonstrated by messaging apps on smart phones) (Friederici et al., 2017; Nieuwland, 2019). However, determining the loci of conscious comprehension is more conceptually challenging. The computational metaphor implied in cognitive models presumes internal coherence as a defining characteristic, which might make sense were neural systems inherently symbolic in nature (a questionable assumption on biological grounds). But calculation of symbols or data points is an analogy of limited application because comprehension also requires non-cognitive internals such as emotional state and hormonal arousal, and externals such as situational and socio-emotional context. Of course, any of these factors may be reduced to quantitative categories (Barsalou, 2017; Boeckx and Grohmann, 2013), but rendering phenomena symbolic through quantified nominalism does not make the phenomena itself symbolic in nature (Seth, 2021).

Additional neural factors relatable to reading

Additional scales of analysis beyond sequences of linguistic indices—orthographic, lexical, syntactic, semantic, textual, or narrative—and systems for integrating them are necessary to appreciate the priority-driven nature of human attention led by emotive factors from basic associations and desires—thirst, hunger, warmth—to higher-order emotional and social associations and desires (Damasio, 2021; Grossmann, 2013; Immordino-Yang, 2015). These intentionally direct the nervous system toward functional satisfaction conditions primed through prior experience, whether perceptual, strategic, socio-emotional, or linguistic, thereby driving interpretation and anticipation on behalf of those conditions. These entailments require more than studies of neuroanatomical correlates to reading-relevant behaviors. Time and sequence bring into view the recursive bio-developmental processes that allow the predictions and confirmations guided by socially normative expectation central to conscious understanding (Grossmann, 2013; Immordino-Yang and Gottlieb, 2017).

Even in a cognitively restrained view, there is more to reading for meaning than decoding, wherein shapes are translated into recognizable sound patterns, plus language comprehension, wherein language is reduced to vocabulary, grammatical relation, and the woolly catch-all of semantics (Cervetti, et al., 2020; Snow & RAND Research Study Group, 2002). As even mid-20th century anglophone analytic philosophers were cognizant, meaning making extends across normative social relations and visceral or cultural intentions that derive signification from those interrelations (Austin, 1962; Rorty, 1978, 1990; Sellars, 1957/1997). From that view, language (and thus knowledge) evolved chiefly not for representation, but for intention and inference sharing. Hence, community-elicited desire, belonging, identity, autonomy, and other human articulations of *thumos*—ancient Greek for full engagement; “zest for life”; desire for recognition, dignity, or social position (Fukuyama, 2018)—are of a piece with the development of language, narrative, and learning to read for meaning in a literate society. We know of no cognitive or developmental literacy models in neuroscience that shadow these motivations.

However, studies confirm that sociality, emotion, cultural traditions, and the structures of societal-economic institutions present equally necessary variables for analysis (Altmann et al., 2012; Lee et al., 2021; Pace et al., 2017; Preckel et al., 2018). The potential of these factors for neuroscience research is explored in studies tracing socio-economic correlates with educational affordances and thus with brain development, or neurological correlates to degree of empathy (Brinch, and Galloway, 2012; Etchell et al., 2018; Falch and Massih, 2010; Lamm and Majdandzic, 2015; Pavlakis et al., 2015). But the weight of all that freight strains the axes of a narrow-gauge review of the literature. A wider gauge network may facilitate the inclusivity required of it all.

The development of a learner’s skills and knowledge base is driven by a pattern-detecting brain filtering experience in coordination with its sociocultural context, and this informs everything from “common sense” or normative assumptions of sequential probability to more precisely articulated theoretical, political, philosophical, or religious beliefs as frames of reference (Gee, 2013; Immordino-Yang and Yang, 2017; Kaplan et al., 2017). These together then inform and direct lexical attention with the granular detail and systemic organization afforded by linguistically mediated perception. To reduce reading to interpretively bounded “regions of interest” indicating neural “modules” is to impoverish literacy scholars’ understanding of the emotionally and culturally adaptive nature of reading development and behavior. It is not that biological elements do not matter, but that any one category of element cannot adequately represent the whole. Thus, the narrow track entices brain-based educators to the fallacy of insufficient metonymy as adequate analogy induced by the false assumption that limited data from cognitive neuroscience on decoding and language processes is a full or ultimate report. It is rather only a report on limited neurological phenomena; not on the behavioral, conscious, social, or cultural phenomena relevant to effective instructional decision making sought by educators. Moreover, the overt emphasis on how anatomical structure allows functional outcomes ignores how an impetus toward functionality over time allows emergence of adaptive structure (see Fig. 4.). And Additionally, the singularly inward gaze of cognitive neuroscience can unintentionally suggest inherent, decontextualized, and immutable causes for ability or disability.

Question Category	Definition	Example
The structure question	What is physically necessary for the behavior to occur?	What neuroanatomy, physiology, or spatial affordances are necessary for language or symbol use?
The function question	What does the behavior accomplish?	What does language or symbol use allow a human to do?
The development question	How does a propensity for the behavior express itself?	How does language ability or symbol use develop over time?
The evolution question	Under what circumstances was the behavior advantageous allowing it to be selected for as a species-specific propensity?	Why did language and symbol use emerge as unique characteristics of <i>homo sapiens</i> ?

Fig. 4 Tinbergen's Four Questions about animal behavior applied to literacy: an example. The first two categories of question—structure and function—are considered immediate or proximal causes. The second two categories—development and evolution—are considered ultimate or distal causes. (Adopted from Tinbergen, 1963). Other questions for human application might include the historical and political/economic.

Neuroscience research related to reading difficulties and dyslexia

The topic of developmental reading difficulties and disorders or disabilities is treated in other excellent entries in this volume. Here we will emphasize but one truth well understood in literacy education: students' difficulty with reading in real-time, or in the ability to develop grade-level reading skills over time, takes many forms and has many causes. Some of these may be decoding related, some language comprehension related, some related to other causes, none of them mutually exclusive and likely operating at various scales of developmentally accrued precedent with eventual higher-order impact on motivation to learn and identity as a reader (Elliott and Grigorenko, 2015). This complexity and diversity makes diagnosis and successful intervention of reading difficulties difficult, but never more so than when a blind eye is turned to the behavioral, cognitive, social, and neural particularity of the readers.

Notably, studies of functional neurophysiology between typical readers and readers with atypical difficulties present diverse results, generated by the various arbitrary boundary conditions used to distinguish contrastive categories (Chyl et al., 2021; Kuhl et al., 2020; Paulesu et al., 2014)—although in the case of neurological studies of reading development the distinctions may seem at times driven by surprisingly assertive vernacular shibboleths, such as that dyslexia is any and every kind of reading difficult (cf., Gabel et al., 2010; Peters and Ansari, 2019; Scerri and Schulte-Körne, 2010). To be generous, neuroscientists (and the technologists, data specialists, statisticians, psychologists, neurologists, graphic artists, and theorists with whom they collaborate) are not experts in the teaching of reading or its research base, and rarely seem to have consulted much applied research on the behavioral categories under investigation before studying their neuroanatomical or neurofunctional correlations.

As an example, most neuroimaging studies comparing typical readers to readers labeled *dyslexic* often cite, or at least paraphrase, a formal clinical definition in their introductory literature review defining dyslexia as a phonological processing disorder due to neurobiological irregularities (e.g., International Dyslexia Association, n.d.; National Institute of Neurological Disorders and Stroke, n.d.). Neurobiologically-based phonological difficulty is a generally accepted definition of dyslexia warranting neurobiological investigations seeking neuroanatomical or genetic atypicalities affecting the SMG, AG, and/or STG, that is, in the language sound processing areas of the cortex (Kearns et al., 2018). Yet in the methods section of said studies, it is nearly always the case that the dyslexic subjects have been selected on the basis of standardized reading assessment scores (for which, it should be remembered, $R = D \times C$; see Decoding subsection above) with no effort to exclude readers encumbered by language disparity rather than weak phonological processing, or to exclude those subjects with perceptual difficulties or disruptive comorbidities (e.g., ADHD, below average IQ) that might account for the impaired reading score, both being instances that would fail to qualify clinically as dyslexia. In other words, dyslexia is defined both as a specific phonological learning disorder and at the same time for subject selection as just any kind of difficulty with being able to read—now you see it; now you don't (cf. Norton et al., 2015; Peters and Ansari, 2019; Hendren et al., 2018; Pettigrew et al., 2016).

Depending on grade level, a varying percentage of subjects will struggle below the arbitrary cut score for reasons *other than* phonological processing difficulty, that is, for reasons other than dyslexia as formally defined (Huetting et al., 2018). Because subjects may score below the chosen cut score for reasons other than decoding issues, they are unlikely to consistently demonstrate atypical ROIs indicating neurobiological distinctions relevant to phonological processing—unless these are the result of, rather than the cause of, their reading difficulties (Kearns et al., 2018). This then presents a chicken and egg (cause and effect)

conundrum in the interpretation of noisy data that may lead, via confirmation bias and small subject sets, to the false positives of current concern. The situation calls for neuroscientists to take a closer look at their study designs, but also at mainstream reading research and its long-standing points of debate (e.g., Alvermann et al., 2019; Goodwin and Jiménez, 2020), and to consider the use of valid subskills diagnostic screening assessments to winnow subject sets more precisely before undertaking neuroimaging.

A shortcoming on the educational side of this cross-disciplinary confusion is mistaking studies of the brain while the subjects are engaged in reading-relevant behaviors as studies *on reading*. They are not; they are studies *on the brain*, that is, studies about neurological phenomena conducted by neuroscientists using neuroscientific methods, written up for publication in neuroscience journals peer-reviewed by, and intended for, other neuroscientists. Such studies are only peripherally about reading processes, and then only as neuroanatomically re-defined, but they are certainly not about reading instruction or reading development as it is understood in literacy education. Reducing reading to measures of blood flow is excessive reduction. Even if taken as a valid indicator of physiological changes in the brain, the data are not representing reading as measurable conscious understanding or instructional efficacy. Nor do such studies prove the validity of the cognitive models of mental process used to interpret the data; those are presumed true in the interpretation of the data and thus cannot be also a finding from the data. All that the successful use of such models demonstrates is that such models are useful for model-coherent data interpretation; the question of whether mental processes or minds actually exist as natural entities remains an open neuroscientific question (Churchland and Churchland, 2002).

Circular reasoning aside, to conflate levels of analysis, conscious and nonconscious, mental and physiological, anatomical and cellular, as if they were all identical or isomorphic, is indicative of category error (Han et al., 2019). In this way, the seemingly tight certainty of a narrow-gauge trek between neuroscience and reading-relevant behaviors can be misinterpreted by non-specialists suggesting it is the brain that reads, rather than the reader, or as *neuroscientific evidence for how to teach reading*, thereby misdirecting or misinforming reading instruction policies and law (Vandenbroeck, 2017). Greater attention to the dynamics of how higher-order and lower-order levels of analysis in complex systems conjointly transact, as in systems biology, is a potential conceptual corrective (e.g., Bassett and Sporns, 2017; Han et al., 2019).

The definition of dyslexia as an incurable, neurobiological disorder has been presumed for close to a century (Vellutino et al., 2004). Conflating the category of clinical (phonological) dyslexia with all “below basic” or “poor readers” (i.e., readers falling beneath a given cut score for a range of reasons) is unjustified. It does not escape notice that in most Western societies the latter group are disproportionally of color, of poverty, and often of foreign birth or immigrant parents and thus more likely second-language learners (Konishi et al., 2014). Research on phonological processing difficulty does not find for this correlation. Conflating these distinct categories of reading difficulty could perpetuate long standing cultural-historical dispositions to segregate people by disfavored characteristics presumed immutable. The call for more equitable funding for schools serving such students could be thereby undermined. Thus, to further marginalize underserved populations through misinterpretation of neuroimaging studies by thought leaders who likely understand very little about neuroscience, or about effective educational practice, or even about the social and economic impact of institutionalized segregation, could perpetuate and extend systems of structural racism and inequity.

That particular areas of the brain, such as the SMG, are the locus of phonological processing is not in much doubt (Church et al., 2011). That imaging results show significant differences in such areas when subjects demonstrate difficulty with phonological processing is not surprising (Dahhan et al., 2016). But to interpret differences reported from neuroimaging studies of those areas as evidence of inherent and immutable or incurable deficiency of already underserved populations without true or adequate longitudinal crosscheck is overreach. The design of cross-sectional studies that claim to show inherent deficiencies in *pre-readers* “at risk” for reading difficulty may someday be looked back upon as something far worse than weakly justified study design.

Neurobiology of reading

A belief in the immutability of human characteristics or potential is a reiteration of the nativist (nature) side of the nature-nurture debate, a relic of 19th and 20th-century Western psychology (Cobb, 2015; Overton, 2015). Nativism perpetuates a pre-modern Western belief in a steady-state world where nothing much ever changes (e.g., the acorn doesn’t fall very far from the tree). It was lent scientific credence through the theoretical posit of the *gene*, a unit of descent, put forth by the 19th-century Austrian monk Gregor Mendel who demonstrated that passage of traits through *genealogical descent* (i.e., as related to *kinship*, another pre-modern, aristocratic Western concern) could be quantified. After the discovery of the functional structure of DNA (Watson and Crick, 1953), a schism opened between those who held an interest in genealogy and persistence, and so relied on behavioral genetics (statistical and social) to track traits across generations, and those whose interest were biological and proteomic (empirical and developmental), and thus more focused on change processes and adaptation.

Obviously, in an overly broad sense, everything human is genetic as humans are multicellular life forms and cells rely on the proteins encoded by their genomes to function—but only in response to biochemical cues within nucleic, cytological, organismic and/or environmental contexts (Gottlieb et al., 2006). At no point in the life cycle is anyone a context-free genome. The initial functioning zygote (i.e., fertilized egg cell) formed by the fusion of genetic materials from two functioning gamete cells is a complete functioning cell responding to its intracellular context within functioning tissues within a functioning body, that body having functionally developed within presumably functional yet changing surrounds—physical, biological, nutritive, and for humans, emotional, linguistic, societal, and cultural—each level responding as necessary to maintain its capacity to function. In this, agent

and context are partners as much as are structure and function, change and stasis, or nature and nurture (Gottlieb, 2001; Overton, 2015). The preformationist idea that information exists within the genome as a context-free blueprint for the formation of structure is representationalist fallacy. In so far as genetic structure can be quantified by humans thereby giving it representational form, the information for any biological structure is as much found off-loaded in contextual specifics as in nucleic sequences, or alternatively as a matter of interrelation and transaction. This is to say, the information is not in or inherent to the organism, the protein, or the genome, but in the interrelationship of these entities with their guiding environmental surrounds to which they have evolved to effectively respond (Oyama, 2000).

There are no genome-wide assessment studies (GWAS) with sufficient ns demonstrating statistically significant genetic variants for dyslexia or propensity for reading ability. Given that approximately 3.4 million nucleotides per subject are mapped in a GWAS, approximately 20,000–30,000 subjects are required (e.g., as exists for studies of schizophrenia), with a p value threshold of 10^{-8} (see Derizotis and Fisher, 2017). The largest GWAS on dyslexia to date correlating dyslexia diagnosis with single nucleic polymorphisms (flipped nucleic pairs in the DNA sequence for a gene) failed to find variants significantly associated with dyslexia (Gialluisi et al., 2021; cf. also Derizotis and Fisher, 2017; Reader et al., 2014). Gialluisi and colleagues relied on available GWAS data from 9 different populations of European descent for a total n of 8546, of which 2274 subjects had been identified as dyslexic. Although the sample size is inadequate, it is the largest sampling to date. Other calculations, including correlations with comorbidities (e.g., ADD/ADHD, general learning disability), were more significant, but in any case rule out a clinical diagnosis of dyslexia for such subjects. (The Manhattan-plot results can be found at <https://www.nature.com/articles/s41380-020-00898-x/figures/1>).

Often cited in refutation to disappointing GWAS results, proponents of behavioral genetics (Mendelian analysis, e.g., Bishop, 2015) cast doubt on the value of GWAS insisting that mappings of actual genes and their variations reveal little about genetic effects; rather, direct observation of traits through inferential analysis of demographic data do. Asserting that identification of traits through social science survey instruments and interviews is more certain knowledge for understanding the biochemistry of cellular-level brain development than genetic assays of the sequential variants responsible for diverse developmental effects is a curious argument, one increasingly strained as the technology for mapping the genome has advanced and become commonplace (Carrion-Castillo et al., 2016; MacArthur et al., 2017).

Traits observed between siblings and twins demonstrate only heritability (characteristics found commonly within families, thus *genealogical*), not genetic variation in the strict biological sense, in other words, as distinct nucleic sequences coding for amino acids that make up proteins with potential effect on cortical development or neural function (Doust et al., 2020). Much is heritable that is not genetic in the biological sense such as first language, early income level, religion, community practices, family traditions, expressive behaviors, etc. (Kaplan et al., 2017). And then there are the effects of epigenesis (Lickliter and Honeycut, 2020; Pallarés-Domínguez, 2019). Twin adoption studies ignore the environmental consistency of shared gestation period and pre-natal environmental effects in development, and the fact that adoption agencies and governmental regulations restrict potential adoptive parents to those with certain socio-economic characteristics at variance with larger populations (e.g., a sufficient income, home stability, lack of criminal record or evidence of abusive behavior), or that most of those who adopt children likely have common reasons for doing so (Segal, 2021).

Assuming that neuro-genetic differences are the structural cause of a reading deficiency, that is, assuming still developing readers are suffering from an incurable neurobiological disorder, rather than a developmental correlate to a complex web of interacting environmental factors (Kershner, 2019; Weisberg et al., 2014), pays inadequate attention to parental education and literacy levels, socioeconomic status, inadequately funded and supported educational affordances, culturally maintained intergenerational effects, and debilitating chronic anxiety of parents from socio-economic condition and social-class stressors. Given that there are strong correlations between socioeconomic class, language development, and thus reading ability (Law et al., 2017), claiming already marginalized populations are disproportionately and incurably weak readers due to inadequate genes or brains denies other manifestly evident factors, or even perpetuates their effects as covert social subordination. If ever there was a need for a wide-gauge synthesis of available research about human adaptation and development by which to critique circular reasoning, affirmation of the consequent fallacy, egregious category error, and willful misinterpretation, student reading and learning deficiencies ought to be first among equals.

Literacy education naturalized: a wider-gauge view

We have suggested throughout this entry a distinction between narrow-gauge and wide-gauge approaches to an educational neuroscience of literacy development. We imply an analogy with rail transport, wherein narrow-gauge lines are typical for local or specialized conveyance (e.g., rapid transit, streetcars, industrial spurs, mining wagons). Wide-gauge lines transport over greater distances at higher speeds, with more powerful engines conveying more varied cargo and in greater volume (e.g., passenger and freight routes; high-speed rail).

Narrow-gauge indicates more specific and immediate connection, and brief is the distance between brain studies and their meta-analyses, or between the theoretical models that inform their designs and the data they produce (as when cognitive models of mental process suggest the functional utility of neuroanatomical correlates). Such focused narrow-gauge tracks play a crucial role in building larger systems of transdisciplinary inquiry. Yet a network of wider-gauge routes are required for longer distances

or more diverse needs and populations. In our usage, the wider gauge indicates transdisciplinary inquiry made manifest for a robust educational neuroscience.

It has been asserted before that an elongated trestle, wide gauge or not, between neuroscience and educational practice is “a bridge too far” (Bruer, 1997). The solution often suggested is to hoist the bridge across intermediary supports in a direct line: from neuroscience to cognitive neuroscience to cognitive psychology to (cognitive) educational psychology and thence to educational practice as hemmed by cognitive theory (Tommerdahl, 2010). But this offers only a single narrow route which suffers a loss of precision with each additional link, and, as noted in our introduction, it ignores literacy education’s socio-cultural turn. We instead propose routing support from the wider life sciences through three additional intermediaries to improve relevant access to education: bridges, in other words, from the life sciences to educational theory and practice via (a) developmental science, including developmental neuroscience (b) naturalized social science (drawing from life science theory), and (c) naturalized epistemology, as in current philosophy of mind, of biology, of language, and neuro-philosophy. For brevity we elaborate in 5 points.

1. *The developmental sciences.* The developmental sciences have long been highly invested in assumptions from the broader life sciences about the nature of living organisms, including human beings, and their capacity to change over time in response to environmental circumstances while ever remaining functional (e.g., Ayoub and Fischer, 2006; Boles, 2019; Fischer and Bidell, 2006; Greenberg and Partridge, 2010; Lickliter and Honeycutt, 2015; Tomasello, 2019). Developmental neuroscience and developmental biology are the switch gates between narrow gauge studies of localized brain function and wider-gauge syntheses of reading development across the lifespan in and out of school contexts (Wiberg et al., 2019).
2. *Developmental science is about change over time*, and we believe that is the unifying objective of education, at both the individual and societal levels. Nonetheless, whatever else humans can be described as—social, cultural, linguistic, ideological—they are most certainly biological and thus exhibit the adaptive developmental processes studied in the developmental sciences. The bio-ecological turn that began in developmental psychology in the 1980s (e.g., Bronfenbrenner and Ceci, 1993; Bronfenbrenner and Morris, 2007) has since carried forward into an ever more cogent and cohesive set of theoretical explanations for how developmental change occurs which we believe worthy of exploration by literacy researchers (for overviews see Bornstein and Lamb, 2015, and other editions; Lerner, 2015, and other editions). The influence of this domain has already emerged within educational theory (Nasir et al., 2021; Yaden et al., 2021).
3. *Life-science informed social science.* Life-science informed social science would include fields such as archeology, cognitive ethology, evolutionary psychology, biological anthropology, or ethno-botany. These offer empirical background for thinking about how natural language becomes developmentally possible in humans. They extend the biology of development and behavior into its transactional and ecological relationships over time. The life sciences remind social science that developmental capacities evolve to enhance sustenance, safety, and adaptive generativity (Berwick et al., 2013), an insight worth bridging to literacy research and classroom practice. Studies in neuroendocrinology on the self-regulating neuro-hormonal pathways directing emotional behaviors, social hierarchies, and trauma is a rich vein of research that is already contributing to studies on classroom community and exclusion. Postulating language as a form of animal communication may seem startling, but as a thought experiment on the meaning of meaning and its adaptations, it has its virtues (e.g., Deacon, 1997; Tomasello, 2019). And comprehensive analysis of the entwined causal relations between *structure, function, development, and evolution* (Tinbergen’s famous four questions; Beer, 2020; Nesse, 2013; see Fig. 4.) has yet to be attempted in literacy research in a way that aligns with what is today known about biological and socio-ecological dynamics, but practitioner hints have been offered (e.g., Hrubby, 2011, 2019; Immordino-Yang et al., 2019; McCormick and Scherer, 2018).
4. *Naturalized epistemology.* As particularly evident over the past 70 years in the anglophone analytic tradition, philosophy can justify and constrain what natural science inquiry may bring to an understanding of language and thought (Bredo, 2016; Cunningham and Fitzgerald, 1996; Quine, 1951). We acknowledge that other Western philosophical traditions (i.e., continental European) have held more sway in educational research over the past few decades (e.g., Foucault, 1966; Heidegger, 1927/1962; Vygotsky, 1978), and that other traditions across the globe are equally worthy of review. We turn as a starting point to the post-positivist anglophone tradition, however, because of its relative emphasis on *empirical naturalism* over rationalist cohesion, *neo-pragmatist holism* rather than reductively determined categorical foundations, and the tacit acceptance, at least as a matter of *normative convention* on matters about which most philosophers (or educators, for that matter) lack technical expertise, of natural science understandings about phenomena presumed of *the greater than human world*, as in the neuro- and life sciences.
5. *A pivot in the anglophone philosophical tradition* during the mid-20th-century (e.g., Austin, 1962; Quine, 1951; Ryle, 1949; Sellars, 1957/1997; Wittgenstein, 1953) set the stage for this bounded relational-linguistic view of meaning and its purposes constrained by the need for communicative and communitarian engagement. Echoing motifs from the American Pragmatists and phenomenology, it addressed the central ethos of philosophy, “to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term” (Sellars, 1963, p. 35), and framed subsequent philosophical investigations (e.g., Brandom, 1994; Lepore and Ludwig, 2005; McDowell, 1996; Millikan, 1984; Rorty, 1978, 1990). Furthermore, insights from the life and social sciences readily inform this tradition, as demonstrated by the diversity of views in current philosophy of mind, philosophy of biology, and neuro-philosophy (Chalmers, 1996; Cobb, 2020; Clark, 1996; Dennett,

1996, 2017; Kim, 2018; Millikan, 2005, 2017; Oyama, 2000; Seth, 2021). Given its continued attention to normative convention and science-informed analysis, we therefore hold recent anglophone epistemology indispensable for framing vigorous neuro- and life science informed literacy education theory that could be more easily translated for assimilation into the public schools and public discourse on education policy.

Finding a place for new realism in literacy research

What these three sectors share—developmental science, naturalized social science, and post-positivist epistemology is belief in, or at least an allowance for, an external, natural reality, which humans are not merely in (and in part responsible for) but are evolved, developed, and living forms of (Bronfenbrenner and Morris, 2007; Greenberg and Partridge, 2010; Lickliter and Honeycutt, 2015), and their indirect capacity for constructing understandings of that reality emergent of, and constrained by, bio-developmental, as well as social and cultural, processes for functioning in that reality (Gottlieb, 2001; Millikan, 1984, 2017). To observe that adaptive and developmental processes are best understood initially as biological is not a rationalization for yet more determinism, reductionism, and mechanism in educational theory, policy, and prescription (e.g., scripted programs, continuous assessments, leveling systems, developmental micromanagement). On the contrary, it is a call for understanding what biological systems, organisms, and human beings are: generative, dynamic, adaptively homeostatic, and opportunistically functional—change processes all that ultimately resist the deadening hand of authoritative prescription and prognostication (Fischer and Bidell, 1998).

It is not surprising that students treated to a curriculum that engages and advances their abilities through meaningful use (Alvermann, 2009; Pressley et al., 2001), seem to tap these bioecological dynamics. Naturalism can offer a complementary view to comprehensive and embodied constructs of classroom literacy practice, as it did for progressive educators a century ago who employed it to guide instructional design (Harris, 1964; Hruby et al., 2016). They recognized that people are not social, cultural, symbolic, linguistic, or ideo-political despite their biology, but because of it. Today, life science research lends further clarity to this insight: it does not follow that our nervous system is itself rationally or computationally symbolic in nature merely because it is presumed the cause of our conscious and cultural ability to be so. Cognitive computational metaphors ought to be used with caution as they are imposed categories of our own devising and not inherent to the greater-than-human world we are trying to understand. And from this it follows: we cannot transcend being emotionally embodied organisms (Marshall, 2015); there is no ghost in the machine to be taught or set free (Ryle, 1949/2009), and no mind-stuff to be uploaded to a matrix (Damasio, 1994; Seth, 2021). We are predisposed by nature to be nurtured by social belonging, approval, and meaningful identity.

Rationalist social science traditions aside, mainstream education's unease with biological idioms stems from its historical misuse to marginalize dispossess, or disenfranchise (Bazzul and Sykes, 2010; Biology and Gender Study Group, 1988; Davis, 1997; Lesko, 2012). This impetus seems still at work in educational policy efforts today, where low performers are segregated for specialized perpetuation of putative "disabilities" (Dever et al., 2016). And neurobiological rationales are still being offered up as justification (Yale Center for Dyslexia and Creativity, n.d.). But we hold this misunderstands human nature. The life sciences can be better used for more positive purposes, including appreciation for the necessary variance of individual uniqueness. We think this ought to be a strength, but, like psychology, neurobiology's emphasis on internal processes and singular distinctions can be mistaken as a validation of the individualism of modern Western culture and the troubled notion of the autonomous rational agent. Granted, Western individualism is lamented for its social atomism and anomie. The West's post-Enlightenment efforts toward democratic, literate, materialist, capitalist, and metropolitan systemics has been both its glory as well as the source of much critical discontent (Harari, 2015). Its pseudo-meritocratic classism refuses to see the elitism of the few as the source of the deprivation of the many (Piketty, 2017). And ecologically, it has become unrooted from and destructive to the earth, the natural world in whose play we remain (Beach et al., 2020).

But it is no more possible to have a vital society without agentive, intentional individuals than to have a complex living organism without agentive, intentional cells (notably neurons) or living cells without agentive, intentional proteins. The lower-order elements are organized into the higher-order entities that comprise their interrelation and are as actively crucial to the relationship within which they have emerged as the perpetuating functionality of the whole (Fischer and Bidell, 1998; Gottlieb, 2001; Lickliter and Honeycutt, 2015). Or as the self-contradicting truism would have it: the situation is not either/or but both/and: not nature or nurture but both, not structure or function but both, not agent or context but both, not any one level but all together (cf., Yaden et al., 2021). These binary distinctions are human constructs imposed to demarcate categories of aspect, not natural or observable causes. We cannot be social or cultural without also being biological. It is in our evolved nature to be nurtured by our caregivers and cultural surround (Immordino-Yang et al., 2019). And at a time when permutations of a persistent pandemic, ecological devastation, and global climate change confront us with both our intellectual fallibility and our physical mortality, accepting bioecological cohesion may serve as a first step toward responsible stewardship in education as in life. It would also be a fresh theoretical restart for literacy education research.

References

- Alexander, P.A., Fox, E., 2019. Reading research and practice over the decades: a historical analysis. In: Alvermann, D., Unrau, N., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, pp. 35–64.
- Altmann, U., Bohm, I.C., Lubrich, O., Menninghaus, W., Jacobs, A.M., 2012. The power of emotional valence—from cognitive to affective processes in reading. *Front. Hum. Neurosci.* 6 (Article 192), 1–15.
- Alvarez, T.A., Fiez, J.A., 2018. Current perspectives on the cerebellum and reading development. *Neurosci. Biobehav. Rev.* 92, 55–66.
- Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), 2019. *Theoretical Models and Processes of Literacy*, seventh ed. Routledge Taylor Francis, New York.
- Alvermann, D.E., 2009. Sociocultural constructions of adolescence and young people's literacy. In: Christenbury, L., Bomer, R., Smagorinsky, P. (Eds.), *Handbook of Adolescent Literacy Research*. Guilford Press, New York, pp. 3–28.
- Arana, S., Marquand, A., Hultén, A., Hagoort, P., Schoffelen, J., 2020. Sensory modality-independent activation of the brain network for language. *J. Neurosci.* 40, 2914–2924.
- Austin, J.L., 1962. *Sense and Sensibilia*. Oxford University Press, London.
- Ayoub, C.C., Fischer, K.W., 2006. Developmental pathways and intersections among domains of development. In: McCartney, K., Phillips, D. (Eds.), *Handbook of Early Child Development*. Blackwell, Oxford, UK, pp. 62–82.
- Barguero, L.A., Davis, N., Cutting, L.E., 2014. Neuroimaging of reading intervention: a systematic review and activation likelihood estimation meta-analysis. *PLoS One* 9, e83668.
- Barsalou, L.W., 2017. What does semantic tiling of the cortex tell us about semantics? *Neuropsychologia* 105, 18–38.
- Bassett, D.S., Sporns, O., 2017. Network neuroscience. *Nat. Neurosci.* 20, 353–364.
- Bazzul, J., Sykes, H., 2010. The secret identity of a biology textbook: straight and naturally sexed. *Cult. Stud. Sci. Educ.* 6, 265–286.
- Beach, R., Boggs, G., Castek, J., Damico, J., Panos, A., Spellman, R., Wilson, N., 2020. Fostering preservice and in-service ELA teachers' digital practices for addressing climate change. *Contemp. Issues Technol. Teach. Educ.* 20, 4–36.
- Beer, C., 2020. Niko Tinbergen and questions of instinct. *Anim. Behav.* 164, 261–265.
- Berwick, R.C., Friederici, A.D., Chomsky, N., Bolhuis, J.L., 2013. Evolution, brain, and the nature of language. *Trends Cognit. Sci.* 17, 89–98.
- Biology and Gender Study Group, 1988. The importance of feminist critique for contemporary cell biology. *Hypatia* 3, 61–76.
- Bishop, D.V.M., 2015. The interface between genetics and psychology: lessons from developmental dyslexia. *Proc. R. Soc. B* 282, 2014339. <https://doi.org/10.1098/rspb.2014.3139>.
- Boeckx, C., Grohmann, K.K. (Eds.), 2013. *The Cambridge Handbook of Biolinguistics*. Cambridge University Press, New York.
- Boles, D.B., 2019. *Cognitive Evolution*. Routledge, New York.
- Bonhage, C.E., Mueller, J.L., Friederici, A.D., Fiebach, C.J., 2015. Combined eye-tracking and fMRI reveal neural basis of linguistic predictions during sentence comprehension. *Cortex* 68, 33–47.
- Bornstein, M.H., Lamb, M.E. (Eds.), 2015. *Developmental Science: An Advanced Textbook*. Psychology Press, New York.
- Brandom, R.B., 1994. *Making It Explicit: Reasoning, Representing, & Discursive Commitment*. Harvard University Press, Cambridge, MA.
- Bredo, E., 2016. Philosophical perspectives of mind, nature, and educational psychology. In: Corno, L., Anderman, E.M. (Eds.), *Handbook of Educational Psychology*, third ed. Routledge, New York, pp. 3–15.
- Brem, S., Maurer, U., Kronbichler, M., et al., 2020. Visual word form processing deficits driven by severity of reading impairments in children with developmental dyslexia. *Sci. Rep.* 10, 18728. <https://doi.org/10.1038/s41598-020-75111-8>.
- Brette, R., 2019. Is coding a relevant metaphor for the brain? *Behav. Brain Sci.* 42, e215. <https://doi.org/10.1017/S0140525X19000049>.
- Brinch, C.N., Galloway, T.A., 2012. Schooling in adolescence raises IQ scores. *Proc. Natl. Acad. Sci. U. S. A.* 109, 425–430.
- Bronfenbrenner, U., Ceci, S.J., 1993. Heredity, environment, and the question "How?": a first approximation. In: Plomin, R., McClearn, G.E. (Eds.), *Nature, Nurture & Psychology*. American Psychological Association, pp. 313–324. <https://doi.org/10.1037/10131-015>.
- Bronfenbrenner, U., Morris, P.A., 2007. The bioecological model of human development. In: Lerner, R.M., Damon, W. (Eds.), *Handbook of Child Psychology: Vol 1. Theoretical Models of Human Development*, sixth ed. Wiley, Hoboken, NJ, pp. 793–828.
- Bruer, J., 1997. Education and the brain: abridge too far. *Educ. Res.* 26, 4–16.
- Burkins, J., Yates, K., 2021. *Shifting the Balance: 6 ways to Bring the Science of Reading into the Balanced Literacy Classroom*. Stenhouse, New York.
- Busso, D.S., Pollack, C., 2014. No brain left behind: consequences of neuroscience discourse for education. *Learn. Media Technol.* 40, 1–19. <https://doi.org/10.1080/17439884.2014.908908>.
- Button, K.S., et al., 2013. Power failure: why small sample sizes undermine the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376.
- Carretero, R.M., Figueiro, A.P.C., 2017. The importance of memory in reading processes. *Maldrige J. Neurosci.* 1, 23–31.
- Carrion-Castillo, A., van Bergen, E., Vino, A., van Zuijen, T., de Jong, P.F., Francks, C., Fisher, S.E., 2016. Evaluation of results from genome-wide studies of language and reading in a novel independent data set. *Gene Brain Behav.* 15, 531–541.
- Cartwright, K.B., 2012. Insights from cognitive neuroscience: the importance of executive function for early reading development and education. *Early Educ. Dev.* 23, 24–36.
- Cassidy, J., Valadez, C.M., Garrett, S.D., 2010. Literacy trends and issues: a look at the five pillars and the cement that supports them. *Read. Teach.* 63, 644–655.
- Catinelli, I., Borghese, N.A., Galluci, M., Paulesu, E., 2013. Reading the reading brain: a new meta-analysis of functional imaging data on reading. *J. Neurolinguist.* 26, 214–238.
- Cervetti, G.N., Pearson, D.P., Palinscar, A.S., Afflerbach, P., Kendeou, P., Biancarosa, G., Higgs, J., Fitzgerald, M., Berman, A., 2020. How the reading for understanding initiative's research complicates the simple view of reading invoked in the science of reading. *Read. Res. Q.* 55, S161–S172.
- Chalmers, D.J., 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, London, UK.
- Chen, L., Wasserman, D., Abrams, D.A., Kochalka, J., Gallardo-Diez, G., Menon, V., 2019. The visual word form area (VWFA) is part of both language and attention circuitry. *Nat. Commun.* 10, 5601. <https://doi.org/10.1038/s41467-019-13634-z>.
- Christenbury, L., Bomer, R., Smagorinsky, P. (Eds.), 2009. *Handbook of Adolescent Literacy Research*. Guilford Press, New York.
- Church, J.A., Balota, D.A., Petersen, S.E., Schlaggar, B.L., 2011. Manipulation of length and lexicality localizes the functional neuroanatomy of phonological processing in adult readers. *J. Cognit. Neurosci.* 23, 1475–1493.
- Churchland, A.K., Abbott, L.F., 2017. Conceptual and technical advances define a key moment for theoretical neuroscience. *Nat. Neurosci.* 19, 348–349.
- Churchland, P.S., Churchland, P.M., 2002. Neural worlds and real worlds. *Nat. Rev. Neurosci.* 3, 903–907.
- Chyl, K., Bartosz, K., Debska, A., Kuniewska, M., Banaszkiewicz, A., Zelechowska, A., Frost, S.J., Mencio, W.E., Wypych, M., Marchewka, A., Pugh, K.R., Jenoróg, K., 2018. Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. *JCPP* 59, 76–87.
- Chyl, K., Fraga-González, G., Brem, S., Jednoróg, K., 2021. Brain dynamics of (atypical) reading development—a review of longitudinal studies. *NPJ Sci. Learn.* 6 (4). <https://doi.org/10.1038/s41539-020-00081-5>.
- Clark, A., 1996. *Being There: Putting Brain, Body, and World Together Again*. Oxford University Press, New York.
- Cobb, M., 2015. *Life's Greatest Secret: The Race to Crack the Genetic Code*. Basic Books, New York.
- Cobb, M., 2020. *The Idea of the Brain: The Past and Future of Neuroscience*. Basic Books, New York.

- Compton-Lilly, C.F., Mitra, A., Guay, M., Spence, L.K., 2020. A confluence of complexity: intersections among reading theory, neuroscience, and observations of young readers. *Read. Res. Q.* 55, S185–S195.
- Cummine, J., Dai, W., Borowsky, J., Gould, L., Rollans, C., Boliek, C., 2015. Investigating the ventral-lexical, dorsal-sublexical model of basic reading processes using diffusion tensor imaging. *Brain Struct. Funct.* 220, 445–455.
- Cunningham, J.W., Fitzgerald, J., 1996. Epistemology and reading. *Reading Research Quarterly*, 31, 36–60.
- Dahhan, N.Z., Kirby, J.R., Munoz, D.P., 2016. Understanding reading and reading difficulties through naming speed tasks: bridging the gaps among neuroscience, cognition, and education. *AERA Open* 2 (4), 1–15.
- Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Penguin, New York.
- Damasio, A.R., 2021. *Feeling and Knowing: Making Minds Conscious*. Knopf, New York.
- Davis, K. (Ed.), 1997. *Embodied Practices: Feminist Perspectives on the Body*. SAGE, Thousand Oaks, CA.
- De Vos, A., Vanvooren, S., Vanderauwera, J., Ghesquiere, P., Wouters, J., 2017. A longitudinal study investigating neural processing of speech envelope modulation rates in children with (a family risk for) dyslexia. *Cortex* 93, 206–219.
- Deacon, T.W., 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. W.W. Norton & Co, New York.
- Dehaene, S., Cohen, L., Morais, J., et al., 2015. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. <https://doi.org/10.1038/nrn3924>.
- Dehaene, S., 2009. *Reading in the Brain: The New Science of How We Read*. Penguin, New York.
- Deniz, F., Nunez-Elizalde, A.O., Huth, A.G., Gallant, J.L., 2019. The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *J. Neurosci.* 39, 7722–7736.
- Dennett, D.C., 1996. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Simon & Schuster, New York.
- Dennett, D.C., 2017. *From Bacteria to Bach and Back: The Evolution of Minds*. Norton & Co, New York.
- Derizotis, P., Fisher, S.E., 2017. Speech and language: translating the genome. *Trends Genet.* 33, 642–656.
- Dever, B.V., Raines, T.C., Dowdy, E., Hostutler, C., 2016. Addressing disproportionality in special education using a universal screening approach. *J. Negro Educ.* 85, 59–71.
- Doust, C., Gordon, S.D., Garden, N., Fischer, S.E., Martin, N.G., Bates, T.C., Luciano, M., 2020. The association of dyslexia and developmental speech and language disorder candidate genes with reading and language abilities in adults. *Twin Res. Hum. Genet.* 1–10. <https://doi.org/10.1017/thg.2020.7>.
- Duke, N.K., Cartwright, K.B., 2021. The science of reading progresses: communicating advances beyond the simple view of reading. *Read. Res. Q.* 56 (S1), S25–S44.
- Elliott, J., Grigorenko, E., 2015. *The Dyslexia Debate*. Cambridge University Press, New York.
- Etchell, A., Adhikari, E., Weinberg, L.S., Choo, A.I., Garnett, E.O., Chow, H.M., Chang, S., 2018. A systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia* 114, 19–31.
- Falch, T., Massih, S., 2010. The effect of education on cognitive ability. *Econ. Inq.* 49, 838–856.
- Fischer, K.W., Bidell, T.R., 1998. Dynamic development of psychological structures in action and thought. In: Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology: Vol 1. Theoretical Models of Human Development*, fifth ed. Wiley, New York, pp. 467–561.
- Fischer, K. W., Bidell, T.R., 2006. Dynamic Development of Action and Thought. In: Lerner, R.M., Damon, W., (Eds.), *Handbook of child psychology: Theoretical models of human development*. John Wiley & Sons Inc, pp. 313–399.
- Foucault, M., 1966. *The Order of Things: An Archaeology of the Human Sciences*. Vintage, New York.
- Frégnac, Y., 2017. Big data and the industrialization of neuroscience: a safe roadmap for understanding the brain? *Science* 358, 470–477.
- Friederici, A.D., Chomsky, N., Berwick, R.C., Moro, A., Bolhuis, J.L., 2017. Language, mind and brain. *Nat. Human Behav.* 1, 713–722.
- Fukuyama, F., 2018. *Identity: The Demand for Dignity and the Politics of Resentment*. Farrar, Straus, & Giroux, New York.
- Gabel, L.A., Gibson, C.J., Gruen, J.R., LoTurco, J.J., 2010. Progress toward a cellular neurobiology of reading disability. *Neurobiol. Disabil.* 38, 173–180.
- Gee, J.P., 2001. A sociocultural perspective on early literacy development. In: Neuman, S.B., Dickinson, D.K. (Eds.), *Handbook of Early Literacy Research*. Guilford Press, New York, pp. 30–42.
- Gee, J.P., 2013. Reading as situated language: a sociocognitive perspective. In: Alvermann, D.E., Unrau, N.J., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Reading*, sixth ed. Newark, DE: International Reading Association, pp. 136–151.
- Gee, J.P., 2021. Sense, Sensitivity, and Language: What We Got Wrong About Human Beings and the Price We Paid. Distinguished scholar lifetime achievement award presentation for the Literacy Research Association, Atlanta, GA.
- Gialluisi, A., Andlauer, T.F.M., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., Ludwig, K.U., Czamara, D., et al., 2021. Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Mol. Psychiatr.* 26, 3004–3017.
- Gillen, J., 2015. Virtual spaces in literacy studies. In: Rowsell, J., Pahl, K., (Eds.), *The Routledge Handbook of Literacy Studies*. London, UK: Routledge, pp. 369–382.
- Goodwin, A.P., Jiménez, R.T., 2020. The science of reading: supports, critiques, and questions. *Read. Res. Q.* 55 (S1), S7–S16.
- Gori, S., Cecchini, P., Bigoni, A., Molteni, M., Facchetti, A., 2014. Magnocellular-dorsal pathway and sub-lexical route in developmental dyslexia. *Front. Hum. Neurosci.* 8 (Article 460), 1–11.
- Goswami, U., 2015. Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* 16, 43–54.
- Goswami, U., 2019. A neural oscillations perspective on phonological development and phonological processing in developmental dyslexia. *Lang. Linguist. Compass* E101111. <https://doi.org/10.1111/lnc3.12328>.
- Goswami, U., 2020. Reading acquisition and developmental dyslexia: educational neuroscience and phonological skills. In: Thomas, M.S., Mareschal, D., Dumontheil, I. (Eds.), *Educational Neuroscience: Development Across the Life Span*. Routledge, New York.
- Gottlieb, G., Wahlsten, D., Lickliter, R., 2006. The significance of biology for human development: a developmental psychobiological systems view. In: Lerner, R.M. (Ed.), *Theoretical Models of Human Development, Handbook of Child Psychology*, sixth ed., vol. 1. John Wiley & Sons, Inc, Hoboken, NJ, pp. 210–257.
- Gottlieb, G., 2001. *Individual Development and Evolution: The Genesis of Novel Behavior*. Psychology Press, New York.
- Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10.
- Gough, P.B., 1972. One second of reading. *Visible Lang.* 6, 291–320.
- Greenberg, G., Partridge, T., 2010. Biology, evolution, and psychological development. In: Overton, W.F., Learner, R.M. (Eds.), *The Handbook of Life-Span Development, Cognition, Biology, and Methods*, vol. 1. John Wiley & Sons, pp. 115–148. <https://doi.org/10.1002/9780470880166.hlsd001005>. Available from.
- Grossmann, T., 2013. The role of medial prefrontal cortex in early social cognition. *Front. Hum. Neurosci.* 7 (Article 340), 1–6.
- Gwilliams, L., 2019. How the brain composes morphemes into meaning. *Philos. Trans. R. Soc. B* 375, 20190311. <https://doi.org/10.1098/rstb.2019.0311>.
- Hagoort, P., 2019. The meaning making mechanism(s) behind the eyes and between the ears. *Philos. Trans. R. Soc. B* 375, 20190301. <https://doi.org/10.1098/rstb.2019.0301>.
- Han, H., Soylu, F., Anchan, D.M., 2019. Connecting levels of analysis in educational neuroscience: a review of multi-level structure of educational neuroscience with concrete examples. *Trends Neurosci. Educ.* 17, 100113.
- Harari, Y.N., 2015. *Sapiens: A brief history of humankind*. HarperCollins, New York.
- Harris, A., 1964. Progressive education and reading instruction. *Read. Teach.* 18, 128–138.
- Heidegger, M., 1927/1962. *Being and Time* (Trans. J. Macquarrie & E. Robinson). Harper and Row, New York.
- Henderson, J.M., Choi, W., Lowder, M.W., Ferreira, F., 2016. Language structure in the brain: a fixation-related fMRI study of syntactic surprisal in reading. *Neuroimage* 132, 293–300.

- Hendren, R.L., Haft, S.L., Black, J.M., White, N.C., Hoeft, F., 2018. Recognizing psychiatric comorbidity with reading disorders. *Front. Psychiatr.* 9 (Article 101), 1–10. <https://doi.org/10.3389/fpsyt.2018.00101>.
- Hoffman, R.R., Deffenbacher, K.A., 1992. A brief history of applied cognitive psychology. *Appl. Cognit. Psychol.* 6, 1–48.
- Hoffman, J.V., Hikida, M., Sailors, M., 2020. Contesting science that silences: amplifying equity, agency, and design research in literacy teacher preparation. *Read. Res. Q.* 55 (S1), S255–S266.
- Howard-Jones, P., 2014. Neuroscience and education: myths and messages. *Nat. Rev.* 15 (12).
- Hruby, G.G., Goswami, U., 2011. Neuroscience on reading: a review for reading education researchers. *Read. Res. Q.* 46 (2), 156–172.
- Hruby, G.G., Goswami, U., 2019. Educational neuroscience for reading education researchers. In: Alvermann, D., Unrau, N., Sailors, M., Ruddell, R. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge Taylor and Francis, New York, pp. 252–278.
- Hruby, G.G., Burns, L., Boztakis, S., Groenke, S., Hall, L., Laughter, J., Allington, R.L., Clark, G., Davis, J., 2016. The meta-theoretical assumptions of literacy engagement: a preliminary centennial history. *Rev. Res. Educ.* 40 (1), 588–643. <https://doi.org/10.3102/0091732X16664311>.
- Hruby, G.G., 2011. Minding the brain. *J. Adolesc. Adult Lit.* 54 (5), 316–321.
- Hruby, G.G., 2012. Three requirements for justifying an educational neuroscience (annual review). *Br. J. Educ. Psychol.* 82, 1–23.
- Hruby, G.G., 2017. Literacy, comprehension, and the neurosciences. In: Israel, S. (Ed.), *Handbook of Research on Reading Comprehension*, second ed. Guilford, New York, pp. 191–216.
- Hruby, G.G., 2019. The apple and the acorn: generating adolescent literacy learners. *J. Adolesc. Adult Lit.* 63 (1), 93–96.
- Hruby, G.G., 2020. Language's vanishing act in early literacy education. *Phi Delta Kappan* 101 (5), 19–24. Available at: <https://kappanonline.org/languages-vanishing-act-in-early-literacy-education/>.
- Huetting, F., Lachmann, T., Reis, A., Petersson, K.M., 2018. Distinguishing cause from effect—many deficits associated with developmental dyslexia may be a consequence of reduced and suboptimal reading experience. *Lang. Cognit. Neurosci.* 33, 333–350. <https://www.tandfonline.com/loi/plcp21>.
- Huey, E.B., 1908. *The Psychology and Pedagogy of Reading: With a Review of the History of Reading and of Writing and of Methods, Texts, and Hygiene in Reading*. MacMillan Co, New York.
- Hughes, B., Sullivan, K.A., Gilmore, L., 2020. Why do teachers believe educational neuromyths? *Trends Neurosci. Educ.* 21, 100145.
- Immordino-Yang, M.H., Gottlieb, R., 2017. Embodied brains, social minds, cultural meaning: integrating neuroscientific and educational research on social-affective development. *Am. Educ. Res. J.* 54, 344S–367S.
- Immordino-Yang, M.H., Yang, X., 2017. Cultural differences in the neural correlates of socio-emotional feelings: an interdisciplinary, developmental perspective. *Curr. Opin. Psychol.* 17, 34–40.
- Immordino-Yang, M.H., Darling-Hammond, L., Krone, C.R., 2019. Nurturing nature: how brain development is inherently social and emotional, and what this means for education. *Educ. Psychol.* 54, 185–204.
- Immordino-Yang, M.H., 2015. Emotions, Learning, and the Brain: Exploring the Emotional Implications of Affective Neuroscience.
- International Dyslexia Association, (n.d.). About Dyslexia: Definition of Dyslexia. <https://dyslexiaida.org/definition-of-dyslexia/>.
- Israel, S. (Ed.), 2017. *Handbook of Research on Reading Comprehension*, second ed. Guilford, New York.
- Jorge, J., Zwaag, W.V.D., Figueredo, P., 2014. EEG-fMRI integration for the study of human brain function. *Neuroimage* 102, 24–34.
- Kaplan, J.T., Gimbel, S.I., Dehghani, M., Immordino-Yang, M.H., Sagae, K., Wong, J.D., Tipper, C.M., Damasio, H., Gordon, A.S., Damasio, A., 2017. Processing narratives concerning protected values: a cross-cultural investigation of neural correlates. *Cerebr. Cortex* 27, 1428–1438.
- Kearns, D.M., Hancock, R., Hoeft, F., Pugh, K.R., Frost, S.J., 2018. The neurobiology of dyslexia. *Teach. Except. Child.* 51, 175–188.
- Kershner, J.R., 2019. Neurobiological systems in dyslexia. *Trends Neurosci. Educ.* 14, 11–24.
- Kim, J., 2018. *Philosophy of Mind*, third ed. Routledge, New York.
- Koch, C., Massimini, M., Boly, M., Tononi, D., 2016. Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* 17, 307–321.
- Konishi, H., Kanero, J., Freeman, M.R., Golinkoff, R.M., Hirsh-Pasek, K., 2014. Six principles of language development: implications for second language learners. *Dev. Neuropsychol.* 39, 404–420.
- Kuhl, U., Neef, N.E., Kraft, I., Schaadt, G., Dörr, L., Brauer, J., Czepezauer, I., Müller, B., Wilcke, A., Kristen, H., Emmrich, F., Boltze, J., Friederici, A., Skeide, M.A., 2020. The emergence of dyslexia in the developing brain. *Neuroimage* 211, 116633.
- Kukona, A., Braze, D., Johns, C.L., Mencl, W.E., Van Dyke, J.A., Magnuson, J.S., Pugh, K.R., Shankweiler, D.P., Tabor, W., 2016. The real-time prediction and inhibition of linguistic outcomes: effects of language and literacy skill. *Acta Psychol.* 171, 72–84.
- Lamm, C., Majdandzic, J., 2015. The role of shared neural activations, mirror neurons, and morality in empathy—a critical comment. *Neurosci. Res.* 90, 15–24.
- Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKean, C., Theakston, A., 2017. *Early Language Development: Needs Provision and Intervention for Preschool Children From Socio-economically Disadvantaged Backgrounds*. Education Endowment Foundation, London. Public Health England.
- Lee, Y., Jang, B.G., Smith, K.C., 2021. A systematic review of reading engagement research: what do we mean, what do we know, and where do we need to go? *Read. Psychol.* 42, 540–576.
- Lepore, E., Ludwig, K., 2005. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford University Press, Oxford, UK.
- Lerner, R.M. (Ed.), 2015. *Theoretical Models of Human Development*. In: *Handbook of Child Psychology*, 7th ed., 1. Wiley, New York.
- Lesko, N., 2012. *Act Your Age! A Cultural Construction of Adolescence*, second ed. Routledge, New York.
- Lickliter, R., Honeycutt, H., 2015. Biology, development, and human systems. In: Overton, W.F., Molenaar, P.C.M. (Eds.), *Handbook of Child Psychology and Developmental Science, Theory and Method*, vol. 1. John Wiley & Sons, New York, pp. 162–207.
- Lickliter, R., Honeycutt, H., 2020. Epigenesis without preformationism: taking development seriously. In: Mascolo, M.F., Bidell, T.R. (Eds.), *Handbook of Integrative Developmental Science: Essays in Honor of Kurt W. Fischer*. Routledge, New York, pp. 191–207.
- Liebig, J., Froehlich, E., Morawetz, C., Braun, M., Jacobs, A.M., Heekeren, H.R., Ziegler, J.C., 2017. Neurofunctionally dissecting the reading system in children. *Dev. Cognit. Neurosci.* 27, 45–57.
- MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Jenkins, H., et al., 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 45, D896–D901.
- MacDonald, K., Germine, L., Anderson, A., Christodoulou, J., McGrath, L.M., 2017. Dispelling the myth: training in education or neuroscience decreases but does not eliminate beliefs in neuromyths. *Front. Psychol.* 8, 1314. <https://doi.org/10.3389/fpsyg.2017.01314>.
- Marek, S., Tervo-Clemmens, B., Calabro, F.J., Montex, D.F., Kay, B.P., Haltoun, A.S., 2022. Reproducible brain-wide association studies require thousands of individuals. *Nature* 510. <https://doi.org/10.1038/s41586-022-04492-9>.
- Marshall, P.J., 2015. Neuroscience, embodiment, and development. In: Overton, W.F., Molenaar, P.C.M. (Eds.), *Handbook of Child Psychology and Developmental Science, Theory and Method*, vol. 1. John Wiley & Sons, New York, pp. 244–283.
- Martin, A.E., Baggio, G., 2019. Modeling meaning composition from formalism to mechanism. *Philos. Trans. R. Soc. B* 375, 20190298. <https://doi.org/10.1098/rstb.2019.0298>.
- Martin, A., Schurz, M., Kronbichler, M., Richlan, F., 2015. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981.
- Martin, A., Kronbichler, M., Richlan, F., 2016. Dyslexic brain activation abnormalities in deep and shallow orthographies: a meta-analysis of 28 functional neuroimaging studies. *Hum. Brain Mapp.* 37, 2676–2699.

- McCabe, D.P., Castel, A.D., 2008. Seeing is believing: the effect of brain images on judgments of scientific reasoning. *Cortex* 107, 343–352.
- McCormick, C.B., Scherer, D.G., 2018. *Child and Adolescent Development for Educators*, second ed. Guilford, New York.
- McCormick, E.M., 2021. Multi-level multi-growth models: new opportunities for addressing developmental theory using advanced longitudinal designs with planned missingness. *Dev. Cognit. Neurosci.* 51, 101001.
- McDowell, J., 1996. *Mind and World*, second ed. Harvard University Press, Cambridge, MA.
- Meyler, A., Keller, T.A., Cherkassky, V.L., Gabrieli, J.D.E., Just, M.A., 2008. Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: a longitudinal study of neuroplasticity. *Neuropsychologia* 46, 2580–2592.
- Millikan, R.G., 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. MIT Press, Cambridge, MA.
- Millikan, R.G., 2005. *Language: A Biological Model*. Oxford University Press, New York.
- Millikan, R.G., 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford University Press, New York.
- Moje, E.B., Afflerbach, P.P., Encisco, P., Lesaux, N.K. (Eds.), 2020. *Handbook of Reading Research, Volume V*, vol. 15. Routledge Taylor Francis Neuroscience, New York, pp. 817–824.
- Nasir, N.S., Lee, C.D., Pea, R., de Royston, M.M., 2021. Rethinking learning: what the interdisciplinary science tells us. *Educ. Res.* 50, 557, 555.
- National Institute of Neurological Disorders and Stroke, (n.d.). Dyslexia Information Page: What Research is Being Done? <https://www.ninds.nih.gov/Disorders/All-Disorders/Dyslexia-Information-Page>.
- Nesse, R.M., 2013. Tinbergen's four questions, organized: a response to Bateson and Leland. *Trends Ecol. Evol.* 28, 681–683.
- Neuman, S.B., Dickinson, D.K. (Eds.), 2011. *Handbook of Early Literacy Research*, vol. 3. Guilford Press, New York.
- Nieuwland, M.A., 2019. Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neurosci. Biobehav. Rev.* 96, 367–400.
- Norton, E.S., Beach, S.D., Gabrieli, J.D.E., 2015. Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* 73–78.
- Oberhuber, M., Hope, T.M.H., Seghier, M.L., Parker Jones, O., Prejawa, S., Green, D.W., Price, C.J., 2016. Four functionally distinct regions in the left supramarginal gyrus support word processing. *Cerebr. Cortex* 26, 4212–4226.
- O'Brien, D., Rogers, T., 2016. Sociocultural perspectives on literacy and learning. In: Corno, L., Anderman, E.M. (Eds.), *Handbook of Educational Psychology*, third ed. Routledge, New York, pp. 311–322.
- O'Mahoney, K., 2021. *The Brain-Based Classroom: Accessing Every Child's Potential Through Educational Neuroscience*. Routledge Taylor Francis, New York.
- Overton, W.F., 2015. Processes, relations, and relational-developmental systems. In: Overton, W.F., Molenaar, P.C.M. (Eds.), *Handbook of Child Psychology and Developmental Science, Theory and Method*, vol. 1. John Wiley & Sons, New York, pp. 9–62.
- Oyama, S., 2000. *The Ontogeny of Information: Developmental Systems and Evolution*, second ed. Duke University Press, Durham, NC.
- Pace, A., Luo, R., Hirsh-Pasek, K., Golinkoff, R.M., 2017. Identifying pathways between socioeconomic status and language development. *Annu. Rev. Linguist.* 3, 285–308.
- Pace, A., Levine, D.F., Golinkoff, R.M., Carver, L.J., Hirsh-Pasek, K., 2020. Keeping the end in mind: preliminary brain and behavioral evidence for broad attention to endpoints in pre-linguistic infants. *Infant Behav. Dev.* 58, 101425.
- Pallarés-Domínguez, D., 2019. Moral neuroeducation: proactive epigenesis and poverty. In: Calvo, P., Gracia-Calandín, J. (Eds.), *Moral Neuroeducation for a Democratic and Pluralistic Society*. Springer, Cham. https://doi.org/10.1007/978-3-030-22562-9_10.
- Paulesu, E., Danelli, L., Berlinger, M., 2014. Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Front. Hum. Neurosci.* 8 (Article 830), 1–20.
- Pavakis, A.E., Noble, K., Pavakis, S.G., Ali, N., 2015. Brain imaging and electrophysiology biomarkers: is there a role in poverty and education outcome research? *Pediatr. Neurol.* 52, 383–388.
- Paz-Alonso, P.M., Oliver, M., Lerner-Usabiaga, G., Caballero-Gaudes, C., Quiñones, I., Suárez-Coalla, P., Duñabeitia, J.A., Cuetos, F., Carreiras, M., 2017. Neural correlates of phonological, orthographic, and semantic reading processing in dyslexia. *Neuroimage* 20, 433–447.
- Pearson, P.D., Cervetti, G.N., 2013. The psychology and pedagogy of reading processes. In: Reynolds, W., Miller, G. (Eds.), *Educational Psychology, Handbook of Psychology*, second ed., vol. VII. John Wiley and Sons, New York, pp. 507–554.
- Pearson, P.D., Palinscar, A.S., Biancarosa, G., Berman, A.I. (Eds.), 2020. *Reaping the Rewards of the Reading for Understanding Initiative*. National Academy of Education, Washington, DC.
- Peters, L., Ansari, D., 2019. Are specific learning disorders truly specific, and are they disorders? *Trends Neurosci. Educ.* 17, 100115. <https://doi.org/10.1016/j.tine.2019.100115>.
- Pettigrew, K.A., Frinton, E., Nudel, R., Chan, M.T.M., Thompson, P., Hayiou-Thomas, M.E., Talcott, J.B., Stein, J., Monaco, A.P., Hulme, C., Snowling, M.J., Newbury, D.F., Paracchini, S., 2016. Further evidence for a parent-of-origin effect at the NOP9 locus on language-related phenotypes. *J. Neurodev. Disord.* 8, 1–8. <https://doi.org/10.1186/s11689-016-9157-6>.
- Piketty, T., 2017. *Capital in the 21st Century* (Trans. A. Goldhammer). Belknap Press, Cambridge, MA.
- Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.-B., Vul, E., Yarkoni, T., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18, 115–126.
- Preckel, K., Kanske, P., Singer, T., 2018. On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Curr. Opin. Behav. Sci.* 19, 1–6.
- Pressley, M., Wharton-McDonald, R., Allington, R., Block, C.C., Morrow, L., Tracey, D., Baker, K., Brooks, G., Cronin, J., Nelson, E., Woo, D., 2001. A study of effective first-grade literacy instruction. *Sci. Stud. Read.* 5, 35–58.
- Price, C.J., 2017. The evolution of cognitive models: from neuropsychology to neuroimaging and back. *Cortex* 107, 37–49.
- Quine, W.V.O., 1951. Two dogmas of empiricism. *Phil. Rev.* 60, 20–43.
- Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., Scotto di Covella, L., 2018. Neuroanatomy of developmental dyslexia: pitfalls and promise. *Neurosci. Biobehav. Rev.* 84, 434–452.
- Reader, R.H., Covill, L.E., Nudel, R., Newbury, D.F., 2014. Genome-wide studies of specific language impairment. *Curr. Behav. Neurosci.* 1, 242–250.
- Reinking, D., Yaden, D.B., 2020. Do we need more productive theorizing? A commentary. *Read. Res. Q.* 56, 383–399.
- Rorty, R., 1978. *Philosophy and the Mirror of Nature*. Princeton University Press, Princeton, NJ.
- Rorty, R., 1990. The dangers of over-philosophication—reply to Arcilla and Nicholson. *Educ. Theor.* 40, 41–44.
- Rosa, M.J., Daunizeau, J., Friston, K.J., 2010. EEG-fMRI integration: a critical review of biophysical modeling and data analysis approaches. *J. Integr. Neurosci.* 9, 453–476.
- Rothwell, J., 2018. Transcranial brain stimulation: past and future. *Brain Neurosci. Adv.* 2. <https://doi.org/10.1177/2398212818818070>.
- Rowell, J., Pahl, K., 2015. *The Routledge Handbook of Literacy Studies*. London, UK: Routledge.
- Rueckl, J.G., Paz-Alonso, P.M., Molfese, P.J., Kuo, W., Bick, A., Frost, S.J., Hancock, R., et al., 2015. Universal brain signature of proficient reading: evidence from four contrasting languages. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15510–15515.
- Ryle, G., 1949/2009. *The Concept of Mind*, 60th anniversary ed. Routledge, New York.
- Scerif, T.S., Schulte-Körne, G., 2010. Genetics of developmental dyslexia. *Eur. Child Adolesc. Psychiatr.* 19, 179–197.
- Schuster, S., Himmelstoss, N.A., Hutzler, F., Richlan, F., Kronbichler, M., Hawelka, S., 2021. Cloze enough? Hemodynamic effects of predictive processing during natural reading. *Neuroimage* 228, 117687.
- Segal, N.L., 2021. *Deliberately Divided: Inside the Controversial Study of Twins and Triplets Adopted Apart*. Rowman & Littlefield, Lanham, MD.

- Seidenberg, M.A., Borkenham, M.C., Kearns, D.M., 2020. Lost in translation? Challenges in connecting reading science and educational practice. *Read. Res. Q.* 55, S119–S130.
- Sellars, W., 1957/1997. *Empiricism and the Philosophy of Mind*. Harvard University Press, Cambridge, MA.
- Sellars, W., 1963. Philosophy and the scientific image of man. *Sci. Percept. Real.* 2, 35–78.
- Seth, A., 2021. *Being You: A New Science of Consciousness*. Penguin, New York.
- Shanahan, T., 2005. *The National Reading Panel Report: Practical Advice for Teachers*. Learning Point Associates, Naperville, IL.
- Sigman, M., Peña, M., Goldin, A.P., Ribeiro, S., 2014. Neuroscience and education: prime time to build the bridge. *Nat. Neurosci.* 17, 497–502.
- Snow, C., RAND Research Study Group, 2002. *Reading for Understanding: Toward and R&D Program in Reading Comprehension*. RAND Corporation, Santa Monica, CA. Available at: https://www.rand.org/pubs/monograph_reports/MR1465.html.
- Sousa, D.A., 2014. *How the brain learns to read*, 2nd ed. Sage, Thousand Oaks, CA.
- Stahl, K.A.D., Flanigan, K., McKenna, M.C., 2020. *Assessment for Reading Instruction*, fourth ed. Guilford, New York.
- Szűcs, D., Ioannidis, J.P.A., 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biol.* 15 (3), e2000797. <https://doi.org/10.1371/journal.pbio.2000797>.
- Szűcs, D., Ioannidis, J.P.A., 2020. Sample size evolution in neuroimaging research: an evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high impact journals. *Neuroimage* 221, 117164. <https://doi.org/10.1016/j.neuroimage.2020.117164>.
- Takashima, A., Konopke, A., Meyer, A., Hagoort, P., Weber, K., 2020. Speaking in the brain: the interaction between words and syntax in sentence production. *J. Cognit. Neurosci.* 32, 1466–1483.
- Themistocleous, M.S., Damianos, E.S., Efstathios, B., Tagaris, G., Kouyialis, A., Psachoulia, C., Stathis, P., 2017. The insertion of electrodes in the brain for electrophysiological recording or chronic stimulation is not associated with any biochemically detected neural injury. *Neuromodulation* 20, 424–428.
- Thomas, M.C., Mareschal, D., Dumontheil, I. (Eds.), 2020. *Educational Neuroscience: Development Across the Lifespan*. Routledge Taylor Francis, New York.
- Tomasello, M., 2008. *Origins of Human Communication*. MIT Press, Cambridge, MA.
- Tomasello, M., 2019. *Becoming Human: A Theory of Ontogeny*. Harvard University Press, Cambridge, MA.
- Tommerdahl, J., 2010. A model for bridging the gap between neuroscience and education. *Oxf. Rev. Educ.* 36, 97–109.
- Vandenbroeck, M. (Ed.), 2017. *Constructions of Neuroscience in Early Childhood Education*. Routledge, New York.
- Vellutino, F.R., Fletcher, J.M., Snowling, M.J., Scanlon, D.M., 2004. Specific reading disability (dyslexia): what have we learned in the past four decades? *JCPP* 45, 2–40.
- Venkatraman, A., Edlow, B.L., Immordino-Yang, M.H., 2017. The brainstem in emotion: a review. *Front. Neuroanat.* 11 (Article 15). <https://doi.org/10.3389/fnana.2017.00015>.
- Verhoeven, L., Perfetti, C.A., 2011. Morphological processing in reading acquisition: a cross-linguistic perspective. *Appl. Psycholinguist.* 32, 457–466.
- Vogel, A.C., Petersen, S.E., Schlaggar, B.L., 2014. The VWFA: it's not just for words anymore. *Front. Hum. Neurosci.* 8 (Article 88), 1–10.
- Vygotsky, L.S., 1978. *Mind in Society*. Harvard University Press, Cambridge.
- Walenski, M., Europa, E., Caplan, D., Thompson, C.K., 2019. Neural networks for sentence comprehension and production: an ALE-based meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 40, 2275, 2234.
- Wandell, B.A., Yeatman, J.D., 2013. Biological development of reading circuits. *Curr. Opin. Neurobiol.* 23, 261–268.
- Watson, J.D., Crick, F.H.C., 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- Weisberg, D.S., Keil, F.C., Goodstein, J., Rawson, E., Grey, J.R., 2008. The seductive allure of neuroscience explanations. *J. Cognit. Neurosci.* 20, 470–477.
- Weisberg, D.S., Hirsh-Pasek, K., Golinkoff, R.M., McCandliss, B.D., 2014. Mise en place. *Trends Cognit. Sci.* 18, 276–278.
- Wendt, A., 2015. *Quantum Mind and Social Science: Unifying Physical and Social Ontology*. Cambridge University Press, Cambridge, UK.
- Wiberg, A., Ng, M., Omran, Y.A., Alfaro-Almagro, F., McCarthy, P., Marchini, J., Bennet, D.L., Smith, S., Douaud, G., Fumiss, D., 2019. Handedness, language areas and neuropsychiatric diseases: insight from brain imaging and genetics. *Brain* 2019, 2938–2947.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Basil Blackwell, Oxford, UK.
- Worthy, J., Daley-Lesch, A., Tily, S., Godfrey, V., Slmeron, C., 2021. A critical evaluation of dyslexia information on the internet. *Journal of Literacy Research*, 53, 5–28.
- Yaden, D.B., Reinking, D., Smagorinsky, P., 2021. The trouble with binaries: a perspective on the science of reading. *Read. Res. Q.* 56, S119–S129.
- Yale Center for Dyslexia and Creativity, (n.d.). <http://www.dyslexia.yale.edu/>.

Literacy models and epistemology

James W. Cunningham, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Models in science	883
Issues from epistemology	884
Reductionism as an epistemological issue	884
Disciplinary reductionism	884
Explanatory reductionism	885
Emergence instead of reduction	885
Literacy models and reductionism	886
Will literacy models based on cognitive psychology be reduced to theories or models in neuroscience?	886
Are there social or cultural facts and phenomena in literacy?	886
Implications of reductionism for literacy models	887
Unobservables as an epistemological issue	887
Modeling unobservables	888
Inference	888
Abduction	888
Transcendental arguments	888
Analogy	888
Implications of unobservables for literacy models	888
Skepticism as an epistemological issue	889
Underdetermination	889
Theory-ladenness	889
Paradigm shifts	890
Postmodernism	890
An epistemological response to skepticism	890
Implications of skepticism for literacy models	891
Epistemology of two literacy models	891
Simple view of reading (SVR)	891
RAND reading comprehension heuristic	892
Lessons	893
References	893
Internet resources (uncited in paper)	894

Models in science

In both the natural and social sciences, researchers sometimes build a model of a target of inquiry. There were famous model builders in history (e.g., Pythagoras, Leonardo, Kepler, Hobbes), but models became increasingly common in fields such as physics and economics during the first half of the twentieth century. Models now play an occasional role in scholarly research using qualitative methods and an extensive role in quantitative research across a wide range of disciplines (Friedman et al., 2008).

Models are problematic to define in part because their targets vary so much. Certainly, it is difficult to say anything specific about what a three-dimensional model of a DNA molecule, a formal model in political science based on game theory, the standard model of elementary particles, the Rasch psychometric model, and a cultural model of marriage in anthropology all have in common. There are also different kinds of models as well as subtypes of each (see Fig. 1).

Less elusive is what a model does: it provides a limited representation of the character, action, or appearance of its target. Both the remove and limits of a model may seem to be liabilities, but in fact they can be strengths. Because a model is not the target but representative of it, it can contain a mixture of features previously confirmed or well established and others inferred as possibly being the case. Because a model's representation is limited, inquiry based on a model is likely to have a closer focus than would an examination of the target. A model may also highlight an aspect of the target that would otherwise be missed or ignored.

Every model's representation of its target is limited, but a few kinds of limitations seem to elicit more criticism. Some models are idealized—for example, models based on rational-choice theory in economics or political science. Some models are analogical—for example, Neils Bohr's model of the atom depicting it as being like a solar system. By contrast, scale models—for example, those designed to test the feasibility of Elon Musk's hyperloop—appeal to common sense and seem to draw less criticism.

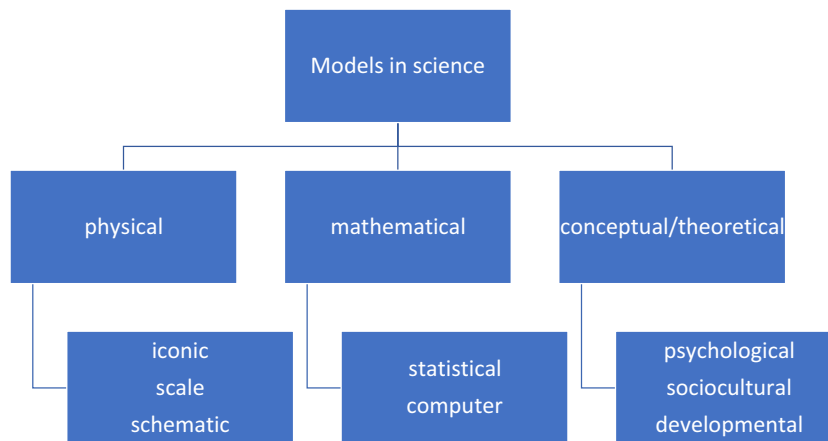


Fig. 1 Kinds of models with examples of subtypes.

A model not only represents a target of inquiry but generally has a meaningful relationship with one or more theories. The word, *model*, can be a synonym of *theory* for simple, narrow, or parsimonious theories that describe or depict a target (e.g., big bang theory/model) or when no theory yet exists that applies to a target. More often, a model is or could be part of an existing theory or has been deduced from one.

Literacy models must be built and interpreted in the context of what is known about reading and writing, how they are learned, and how they are taught. It would be a missed opportunity for the literacy field, though, if the meaning of the phrase, “what is known,” were taken as obvious or reduced to a debate about research methods. Knowledge itself has been studied for thousands of years. Epistemology is “the branch of philosophy that deals with the varieties, grounds, and validity of knowledge” (Brown, 1993, p. 838). Other entries in this encyclopedia volume appropriately address what is known about various aspects of literacy. The present chapter concentrates instead on the epistemology underlying or implicit in models of literacy.

Issues from epistemology

There are numerous issues regarding knowledge in epistemology, e.g., coherence, justification, objectivity. This article has been organized around three epistemological issues especially applicable to literacy models—reductionism, unobservables, and skepticism.

Reductionism as an epistemological issue

Disciplinary reductionism

If every object in nature consists of molecules, atoms, and subatomic particles, as almost all philosophers and researchers surely believe, why do we need any other scientific discipline besides physics? Since we have physics, why do we need chemistry or biology? Since we have the natural sciences, why do we need the social sciences? To put these questions as an epistemologist might, is what we claim to know specific to a discipline *reducible* to a more basic discipline, perhaps ultimately reducible to the definitions and laws of physics?

In the late nineteenth century, Émile Durkheim helped establish naturalistic social science, social inquiry conducted using the same research methods as the natural sciences (Durkheim, 1895/1982). At the same time, Durkheim defended the autonomy of academic sociology as a field by arguing for the existence of *social facts*, i.e., facts in society that constrain individuals, but which cannot be attributed to or explained by what individuals think, feel, or do (Durkheim, 1895/1982). Thus, Durkheim was an advocate for the unity of scientific methodology, but an opponent of the unity of science.

Starting in the 1920s, members of the philosophy of science movement known as logical positivism, later logical empiricism, likewise promoted unity in scientific methods, but also worked toward the unity of science itself. Rudolf Carnap [1891–1970] and Otto Neurath [1882–1945] argued that all sciences should adopt common principles of language. A parallel proposal for all scientific language was *operationalism*, first suggested by the physicist, Percy Bridgman (1927). Operationalism became the standard that scientific language be formulated or translated to consist only of formal and empirical statements. An empirical statement is one for which there exists a set of objective operations by which it may be observationally or experimentally tested, determining it to be confirmed (*verificationism*), rejected (*falsificationism*), or one or the other. Operationalism had its heyday during the era of behaviorism in psychology, economics, and education.

Early on, logical positivists/empiricists had been suspicious of theory, but Carl Hempel’s deductive-nomological model (e.g., Hempel and Oppenheim, 1948) was an approach to scientific theorizing that eschewed induction in favor of deduction and reduced causation to prediction. Research starts with a proposed or established law of nature that may explain the phenomenon

being investigated. It proceeds by describing the initial conditions for the phenomenon and then deducing from the law one or more hypotheses (predictions) about how the data to be collected will turn out. Hempel called his approach the “covering-law” model because for him science is the search for laws that cover (entail) the data to be collected to test them.

The deductive-nomological model became a strategy for those attempting to unify science through disciplinary reduction (e.g., Lewis, 1970; Oppenheim and Putnam, 1958). The strategy assumed that all laws (and lawlike generalizations), however well established, are hypothetical, like hypotheses before data collection in an individual study. As predictions of how data to be collected will turn out in a particular study can potentially be deduced from a law (and a statement of initial conditions), so each law of a less basic discipline was considered potentially deducible from those of a more basic discipline. Through many instances of such micro-reductions, the unity of science was eventually to be achieved (Oppenheim and Putnam, 1958).

The case made by logical positivists for having all scientific disciplines embrace common criteria of language was eventually refuted (Popper, 1965; Quine, 1951) as was operationalism (Hempel, 1966). The deductive-nomological model was substantially revised for several reasons. Despite David Hume’s argument against inductive reasoning (Hume, 1748/2007), most scientists and even most philosophers of science approve of it to some extent (Hacking, 2001; Rothchild, 2006). The covering law model cannot distinguish between generalizations that are lawlike from those that are accidentally the case and excludes kinds of scientific explanations accepted by many practicing scientists (Glennan, 2006). Much actual science does not conform itself to exceptionless laws, especially in historical sciences like biology (Rosenberg, 2016), geology, archeology, and anthropology. Finally, the unity of science failed as an epistemological project. The main reason was because each established discipline has ownership of facts more basic disciplines have been unable to explain satisfactorily (Rosenberg, 2016).

Explanatory reductionism

Complexity results from an object, process, or system having distinguishable causes or parts yet remaining a unified whole. Theoretical explanations are efforts to make the why and how of a phenomenon intelligible. Reductionism (or reductivism) is an attempt to explain the properties of a complex whole solely in terms of the properties of its causes or parts. In a regression model, for instance, a successful reduction is supported when the whole is accounted for by the weighted sum of its predictors and the model has been cross-validated.

Ockham’s razor—the many centuries old preference for simpler explanations of phenomena—has been an important motive for scientists to reduce theoretical complexity (Gauch, 2002). One advantage of parsimony (Ockham’s razor) is the methodological principle that theories (and models) with fewer hypothetical causes or components are more likely to be successful, because the more hypothetical entities a model includes, the more are available to fail an empirical test. Another advantage is that explanations with fewer assumptions simultaneously avoid the excesses of imaginative theorizing and overfitting models to the data collected.

A potential problem with explanatory reductionism, however, can arise if researchers exercise a preference for reduced explanations by ignoring facets of wholes, their complexity, or evidence supportive of more complex explanations (Dennett, 1995). Such zealous reductionism has sometimes been identified by its opponents as the “nothing-but reflex” (Davidson, 1980, p. 214), the nothing-but fallacy, or—most charmingly—nothing buttery. Wanting to defend what he calls “good reductionism” (p. 83) while acknowledging bad reductionism (nothing buttery), Dennett (1995) gives a new name, “greedy reductionism” (p. 83) to the latter. Nothing buttery (greedy reductionism) occurs when the whole is more (or less) than the sum of its parts/causes, but a reductive explanation of the whole is provided anyway.

Emergence instead of reduction

In contemporary model building, emergence is a hypothesis for why some reduced theories and models exhibit explanatory insufficiency. John Stuart Mill (1843) contrasted homopathic and heteropathic effects of a set of multiple causes. Later, G. H. Lewes (1875) renamed the same two kinds of effects *resultant* and *emergent*, respectively. Homopathic or resultant effects follow directly from the properties of their causes or equal their sum; heteropathic or emergent effects are not entirely predictable from the properties of their causes or are greater than their sum.

Emergence is manifest whenever a whole has qualities or effects that cannot be explained by the qualities or effects of its parts. For example, the behavior of an ant colony is an emergent property inexplicable from the properties of an individual ant. Some philosophers have used the term, *supervenience*, to label a condition of dependence on parts without causation.

There are two types of emergence, metaphysical and epistemological. For the British emergentists, Mill and others, and the more recent English philosopher, Roy Bhaskar (1986), many disciplines in the natural and social sciences exist as a reflection of the ontological stratification of nature. In principle, scientific disciplines can be ordered from most to least basic, consistent with the hierarchical order of the strata they study. For instance, physics is more basic than chemistry, which is more basic than biology, which is more basic than anthropology, because the layer of nature each science attends to supervenes on (depends on but is not fully explained by) more basic layers. If so, there are phenomena at each layer with characteristics beyond those resultant from causes operating at more basic strata.

Epistemological emergentists—the majority—have considered emergence to reflect the complexity of a phenomenon, to be the product of a complex web of interactions among its properties, parts, or causes. Chaos theory (Gleick, 2008) as applied across several disciplines can be seen as an example of explanation by epistemological emergence. The effects chaos theory models are assumed to result directly from their causes and initial conditions yet are extremely difficult to predict because of the complexity of their system—the quantity of causes, the quantity of their interactions of different orders (simple, 2-way, etc.)—and the complexity of the interaction of the whole system with initial conditions.

Literacy models and reductionism

Literacy is a cross-disciplinary field. Research on reading, writing, and literacy is published in literacy journals (e.g., *Reading Research Quarterly*), education journals not specific to literacy (e.g., *Journal of Educational Psychology*), journals associated with disciplines in arts and sciences (e.g., *Applied Linguistics*), journals focusing on technology or research methodology (e.g., *Behavior Research Methods*), and other academic journals (e.g., *Topics in Language Disorders*). The cross-disciplinary nature of literacy inquiry may be a key reason why models of literacy exist in such variety.

Even though the logical empiricist goal of unified science no longer animates many philosophers or scientists, there are specific disciplines today where claims to disciplinary autonomy are contested. Three of these disciplines—psychology, sociology, and cultural anthropology—have been directly or indirectly influential in literacy research. Paradoxically, psychology's integrity as a discipline is under threat of reduction today at the same time as it continues to pose a reductive threat to the disciplinary integrity of sociology and cultural anthropology. There are potential ramifications of these threats of disciplinary reduction for model building in literacy.

Will literacy models based on cognitive psychology be reduced to theories or models in neuroscience?

Psychology, including educational psychology, has been the social science within whose purview most research on literacy has been conducted. The quantity of psychologically oriented literacy research has greatly increased since the cognitive revolution. In recent decades, though it remains controversial, there seems to be wide acceptance of the conjecture that neuroscience will eventually replace psychology.

Hruby (2017) succinctly captures the challenge to psychology posed by neuroscience as it manifests itself in reading inquiry, "In approaching the neuroscience literatures on reading, most literacy scholars assume a ready connection with cognitive science. After all, is it not the case that the mind is in the brain?" (p. 192). Yet, distinctions from the philosophy of mind explain why literacy scholars should doubt an imminent transparent relationship between cognitive psychology and neuroscience.

Davidson (1970) appropriated the term, *supervenience*, to label his description of the mental being "in some sense dependent" (p. 214) on the physical without being reducible to it. A supervenient relationship has two dimensions, *dependence* and *irreducibility*. For Davidson, thoughts and actions are totally dependent on neural events in the brain, but irreducible to them.

Irreducibility of the psychological to the neurological is often argued by philosophers of mind using Putnam's (1967) *multiple realizability* thesis. This thesis asserts that two people may achieve similar perceptual, cognitive, or other psychological outcomes from different neurological states. Fodor's (1968) functionalism about mental properties also implied multiple realizability. In support of Putnam, Fodor, and their followers, Endicott (1993) reviewed findings from brain research that persons with brain injury or other conditions had developed alternative neural pathways or subprocesses to achieve specific psychological functions.

Hruby's (2017) account of the irreducibility dimension of the relationship applied to reading is an excellent summary of such insights, "perceptual and constructive abilities are possibly both predicated on a highly fluid neural architecture that allows for functional plasticity—or, in other words, that allows neurologically endowed organisms to wrap their nervous systems around their environment in highly individualized yet functional ways" (p. 208).

Whether the cognitive entirely depends on the physical, whether the cognitive is irreducible to the physical, and whether the cognitive encompasses emergent properties and powers all remain matters of contention in epistemology and the philosophy of mind. The persistence of these disagreements means models of literacy from neuroscience should be cautiously welcomed. Some will probably receive empirical support. Nevertheless, cognitive psychology still preserves its integrity and autonomy as a vibrant field despite decades of attack by disciplinary reductionists.

Are there social or cultural facts and phenomena in literacy?

Are the theories, models, and research findings of fields such as sociology and cultural anthropology (and political science and economics) reducible to psychology? Are the social and cultural dynamics operating in groups of people these fields study merely an aggregation of the psychologies of the individuals in those groups? Or do irreducible social and cultural forces and processes exist above and beyond individual psychologies that instead enable social scientists to predict or interpret the thoughts and actions of individuals? In short, are there social and cultural facts and meanings or are there only psychological facts and meanings? Competing answers to these questions come from long and continuing disputes between holists and individualists in epistemology and metaphysics.

Holists oppose disciplinary reductionism as well as explanatory reductionism by insisting that social or cultural phenomena have holistic explanations instead of or in addition to explanations based on the psychologies of the individuals who make up societies or live in cultures. As discussed earlier, Durkheim's (1895/1982) view of academic sociology included holism (the existence of social facts). The cultural anthropologist, Alfred Kroeber (1952), argued for the existence of cultural facts irreducible to facts about individuals living in a culture. The fields of sociology and cultural anthropology theorize that social and cultural wholes have social and cultural parts or causes. For many scholars in these fields, societies and cultures are social and cultural all the way down.

Individualists reject holistic explanations in favor of explanatory reductionism. They only acknowledge facts for groups as long as those facts are predictable or interpretable as the aggregated effects of the psychologies of individual people. Radical empiricists and positivists have seldom been holists because they consider emergent properties and effects to lack empirical legitimacy.

There appear to be social facts in literacy. For example, poverty has a clear impact on the social institution of schooling revealing itself in unequal outcomes in reading and writing achievement. There also appear to be cultural facts in literacy. For example, even though Whole Language conflates literacy with language, language is the product of biological evolution while literacy is the product of cultural evolution. The latter may depend on the former, but it is doubtful it can be reduced to it.

Implications of reductionism for literacy models

As of now, epistemology and the philosophy of social science leave us room to pursue literacy inquiry in ways consistent with the theories and methods of cognitive psychology, sociology, and cultural anthropology. Any disciplinary reductionism replacing these respective fields with neuroscience or psychology likely lies in the distant future if it ever comes to pass.

Explanatory reductionism is a more complicated matter. Although reductionism is a dirty word in some quarters, it performs an important function in scholarly inquiry. Even in interpretivist social science, where antireductionism is proudly owned, scholars often select a stance, lens, or theoretical framework to guide a study, presumably to avoid having to take every perspective one could take in a single inquiry. Reductionism has several possible implications for model selection and building in literacy:

- “Good reductionism” can aid builders of literacy models. Any literacy phenomenon can be modeled, whether it is a social, cultural, psychological, neurological, or other kind of phenomenon. For any literacy phenomenon, the representative nature of a model may constitute reductionism by design. The limits and remove of a literacy model can provide a literacy inquiry with needed focus and invite inferences.
- Ockham’s razor (the principle of parsimony) may serve the purposes of literacy inquiry. When two models explain a literacy phenomenon well, the one with fewer assumptions or hypothetical entities is more likely to withstand empirical scrutiny.
- Explanatory insufficiency regarding a literacy target of inquiry is a justification and motive for considering emergence, supervenience, or holism when newly modeling that target.

Unobservables as an epistemological issue

No one has ever seen a proton, let alone a quark, and physicists do not anticipate the development of an instrument that would allow either particle to be directly observed. Such *unobservables* have long presented a challenge to those with a radical empiricist epistemology.

The founders of modern empiricism, John Locke and David Hume, and their followers have largely held to the variously attributed axiom: “Nothing is in the intellect that was not first in the senses.” Locke’s (1690/1975) use of the metaphor, *tabula rasa* (blank slate), to describe the human mind in its original condition reflects his similar view that the source of all a person’s ideas is sense experience. For Locke, the senses comprise the classic five external ones plus reflection, internal awareness of one’s own mind in operation.

All empiricists share the belief that immediate sense experience constitutes the primary source and sanction for knowledge. However, it was probably Hume (1748/2007) who first formulated an empiricism so pure it merits the label, *radical*. Unlike Locke, Hume refused to accept any possibility of knowledge beyond what current observations afford. He argued against the possibility of knowing causation in science or everyday life. Later, Auguste Comte (1853/2009) coined the term *positivism* to describe his brand of radical empiricism. Comte restricted human knowledge to positive (i.e., certain) facts or laws arrived at empirically. He rejected all other claims to knowledge whether subjectively declared or the product of metaphysical argument. Radical empiricists and positivists since Hume and Comte have generally been uncomfortable with assertions about the existence or nature of anything that cannot be directly observed. The recent “constructive empiricist,” Van Fraassen (1980), is agnostic about the existence of unobservables and skeptical of all instruments used to indirectly observe them by their traces or effects.

Phenomena is Immanuel Kant’s (1787/1966) term for “object[s] or occurrence[s] perceived by the senses” (Flew, 1984, p. 266). What is experienced in hallucinations, dreams, etc. are not phenomena. In scientific realism, phenomena are different from data (Bogen and Woodward, 1988) because data are always from specific observations whereas phenomena are relatively consistent and stable across observations and observers. For example, reading comprehension data are usually answers to items on a specific test or task that purport to measure one or a set of instances of reading comprehension. Reading comprehension itself is a phenomenon that has been witnessed in many different ways through countless formal and informal observations made by innumerable observers across myriad times and places.

Some phenomena have the status of what metaphysicians call *natural kinds*. The chemical element, mercury, the animal species, moose, and the physical force, magnetism, are exemplars of natural kinds, “nature cut at its joints” as Plato [c. 428-c. 348 BCE] put it. Other phenomena remain contested in terms of whether they constitute a single category of objects or occurrences to which one may correctly refer using the name given to that category. Empiricists who believe in the existence of natural kinds await more data from more investigations before they are willing to confer that status on a phenomenon whose existence as a single category is disputed. Hacking (1999) has also argued for the existence of *social kinds*.

In the natural and social sciences today, phenomena rather than data are usually what models are constructed to explain or predict (Bogen and Woodward, 1988). Moreover, as the natural and social sciences have evolved into a search for causal mechanisms (Hedström and Ylikoski, 2010), many of the phenomena considered worthy of study are unobservable to some degree.

The existence and nature of unobservable phenomena are inferred. Their status as phenomena largely depends on the relative stability and consistency of their characteristics and facts about them supported by data from multiple investigations by multiple researchers.

Modeling unobservables

When a model of a target phenomenon is built, the model can represent both observable and unobservable objects, occurrences, processes, mechanisms, etc. Any unobservables in a model are the product of one or another kind of inferencing.

Inference

In his introduction to epistemology, Audi (1998) distinguishes between two kinds of beliefs. The first is directly based on sources of knowledge: perception, memory, consciousness, reason, or testimony. The second is based instead on other beliefs. He summarizes the value of the second kind of beliefs, "It is largely because we can inferentially build on what we already believe (or assume) that there is no limit to the richness and complexity of the ideas and theories we can construct" (p. 152).

According to Audi, inference is a kind of reasoning that enables us to acquire new beliefs and occasionally new knowledge, yet it is not a source of knowledge. If we lack knowledge, inference does us no good—there is nothing for it to build on, nothing for it to extend. When we do have knowledge, inference can increase and expand the knowledge we have. Deduction is the classic kind of logical inference that draws a necessary conclusion from one or more premises. Induction is the kind of probabilistic inference where a generalization is constructed after examining a set of particulars. Inference can also have an inventive potency based on the imagination.

Abduction

The founder of pragmatism, C. S. Peirce [1839–1914], added a third kind of reasoning to deduction and induction. *Abduction* is like induction because it is reasoning from instances (data). They differ in that, while induction makes a generalization about the data, abduction infers the best explanation (hypothesis) for the data. For example, after noticing a stack of unpaid bills on a desktop in your brother's home, you could induce that he owes a lot of money or abduce that he has been gambling excessively again.

Transcendental arguments

Transcendental arguments came to the fore in Kant's (1787/1966) epistemology. Some things are obviously the case (e.g., a particular event repeatedly occurs) and there are facts even most skeptics believe. A Kantian transcendental argument takes an obvious case or uncontroversial fact as its premise and deduces a necessary enabling condition for it from it.

During modeling in science, narrower and more contested transcendental arguments than Kant's can be made for an unobservable based on evidence, data, or prior research the model builder assumes and builds on. For instance, Brown and Smiley (1977) found that third grade participants retold 3.6 times as many of the most important quarter of the ideas than the least important quarter in unfamiliar fairy stories they had heard. A transcendental argument could be made by reasoning from this finding that the discourse structure of fairy tales (and possibly a broader range of texts) shapes listeners' (and possibly readers') sense of importance of the ideas in a text. This argument could lead a model builder to include an unobservable trait of sensitivity to such shaping in a model of text comprehension or an unobservable trait of shaping readers' sense of ideas' importance in a model of writing. Whenever a transcendental argument is used to justify the presence of a particular unobservable in a model, it is a type of inference.

Analogy

The British philosopher of science, Hesse (1963) argues that practicing scientists sometimes build a model of a target phenomenon based on an analogy with another well-understood phenomenon. She uses the familiar example of balls in motion on a billiard table as a model of gas particles in physical chemistry. For her, such an analogy is "positive" with respect to specific properties billiard balls and gas particles share (e.g., motion and impact), "negative" with respect to specific properties that distinguish them (e.g., billiard balls are hard and shiny; gas particles are not), and "neutral" with respect to any specific properties where it is currently uncertain whether billiard balls and gas particles share them. It is the "neutral" aspect of analogous models, Hesse (1963) contends, that scientists can fruitfully focus on to generate new hypotheses about a target (e.g., where another property of billiard balls in motion is theorized to also apply to gas particles). At the time of model building, it is one or more of these neutral properties with undetermined similarity to the target vis-à-vis the analog that are unobserved and inferred by analogy.

Implications of unobservables for literacy models

Many psychological, social, and cultural phenomena of interest to literacy scholars today cannot be directly observed (e.g., lexical access). Even when target literacy phenomena are observable, models of those phenomena frequently include unobservable components, processes, mechanisms, or causes. Thus, contemporary model building in literacy generally requires inferring of one kind or another. Epistemologies may be divulged by how researchers, theorists, and practitioners view such inferences. Radical empiricists resist inferring literacy unobservables or their characteristics and minimize inferences where they cannot avoid them altogether. Scientific realists work to make full use of inferences in their literacy models, limited only by the regimen of empirical testing. Social

constructionists are skeptical of inferential reasoning as an individual objective process and look instead for evidence of social or cultural reasons for the “inferences” people, including scientists, seem or claim to make. Unobservables have important implications for model selection and building in literacy:

- Unobservables constitute the vanguard of literacy inquiry. Modeling what is known or well-established in an area of literacy has summative value, but knowledge about literacy stands to grow when literacy model builders reach out into the dark, inquiring into unobservable phenomena or unobservable properties or parts of observable phenomena.
- Empirical testing of theories and models is necessary for scientific knowledge to increase. Once that is conceded, however, it must also be admitted it is not testing, but what is tested that advances understanding of the world, including the world of human experience. Inferences (deductions, inductions, abductions, transcendental arguments, and analogies) are tools by which risk takers posit the unobservables that often make empirical testing fruitful when it is. Those who desire to accelerate knowledge and understanding of literacy may do well to be wisely bold in their use of creativity and imagination to infer what can currently be seen in no other way.

Skepticism as an epistemological issue

A popular movie, *The Matrix*, showed us something out of a nightmare most of us never think about at other times. Epistemologists have long thought about such things. René Descartes [1596–1650] famously pondered the possibility that a powerful evil demon had deceived him into thinking of himself and his life as if they were real when they were an illusion. Putnam (1981) is credited with inspiring a similarly nightmarish notion in contemporary epistemology. What if I am just a brain in a vat being fed nutrients and artificial sensations from which my thought processes are constructing the world that seems real to me but isn’t? In epistemology, such scenarios have been exploited by some to defend a blanket skepticism about what the senses can tell us, but a more nuanced version of the same argument is frequently used today to support a related but lesser skepticism, one directed toward scientific theories and models. A formulation of this argument is that theories and models are *underdetermined* by data or evidence.

Underdetermination

In both the natural and social sciences, data are gained from human senses or amplifications of those senses. Data are often used to test hypotheses deduced from theories or models. What happens when data falsify a hypothesis? Philosophers and sociologists of science have noted that, rather than the hypothesis routinely being rejected, the theory, model, or methodology of testing the hypothesis is often tweaked instead to warrant the same hypothesis surviving to be tested again another day. Underdetermination of theory by data also pertains when data (including big data) are examined abductively to decide which existing theory or model best explains them. As with the evil-demon or brain-in-a-vat scenario, there is always another existing or imaginable theory or model besides the one being tested, fitted, or built that could also explain the data.

Some philosophers and researchers have been persuaded by the underdetermination of theory by data or evidence to adopt either a radically empiricist phenomenalism (all we can know are sensations, not what may lie behind them) or a postmodern relativism. The majority, however, have granted that theories and models do not achieve a consensus of acceptance, when they do, simply because the data/evidence could lead in no other direction. Rather, it has become broadly recognized that academic inquiry has other contexts and goals in addition to empiricism—historical, social, cultural, political, even esthetic and intuitive—that do and should influence which theories or models currently hold sway in a discipline.

Theory-ladenness

Another kind of skepticism is the thesis of the *theory-ladenness* of data based on Hanson’s (1958) assertion that all data and other evidence accumulated to test a theory or model are themselves theory-laden. When collecting data or other evidence to test a theory or model, the kind of data/evidence chosen, the way the observations are conducted, and the tolerances set to determine when an observation yields adequate information are themselves influenced by the theory or model being tested. How could it be otherwise? For instance, scientists would not have known to seek evidence for the curvature of space-time, or how to seek it, had they not first had Einstein’s general theory of relativity that predicted it.

Theory-ladenness has a more debatable and extreme implication, that the terminology of the hypothesis and the variable(s) observed have no meaning outside the theory or model being tested. Hence, there is no way to test theories or models without already assuming them to be correct. And might not researchers simply be finding what they expect to find? When theory-ladenness is felt to have this implication, it becomes an argument for relativism.

Theory-ladenness has become less of a damning critique in the natural and social sciences because its main thrust is to undermine the certainty of each test of a theory or model. Logical empiricism was still very much a force in 1958, but in the past six decades few philosophers of science and fewer scientists have been radical empiricists or positivists. Certainty is no longer the *sine qua non* in research to be obtained at the cost of everything else but sense data. Scientists and philosophers of science are more likely now to be scientific realists who consider science probabilistic and a process of successive approximation to knowledge, or pragmatists who are less committed to the truth of theories and models than to their usefulness.

Paradigm shifts

What has been interpreted by some as a third kind of skepticism toward scientific theories and models is Kuhn's (1970) thesis of *scientific revolutions* brought on by *paradigm shifts*. Kuhn's work was immensely important in the demise of logical empiricism. As discussed in the earlier section on reductionism, logical positivists/empiricists desired and foresaw the eventual unity of science. They understood the history of science to be (or believed it should be) a continuous evolution. During this process of incremental change, laws of nature are tested empirically, revised, and retested, and established laws in each discipline are eventually replaced by laws from a more basic discipline. Science marches on.

Kuhn's study of the history of science led him to paint a very different picture. In his account, science during a historical period adheres to a *paradigm*, a broad set of beliefs shared by scientists about the nature of things within which a set of scientific theories arises. These theories undergo empirical development and a set of scientific approaches and methods become normative. The purpose of the theories (and models, in more recent periods) and how they are evaluated is to help solve the problems presented by the ruling paradigm as important to solve. Rather than being evolutionary, scientific change as portrayed by Kuhn is revolutionary. The ruling paradigm is discarded when it no longer solves the most worrying problems of its period satisfactorily. It is replaced with a new paradigm that is incompatible and maybe incommensurable with its predecessor. As happened with Ptolemy's epicycles, the anomalies build up until a Copernicus (not always an individual) comes along to usher in a new paradigm. The subsequent change is revolutionary across the board including making what were facts during the reign of the earlier paradigm sometimes no longer be facts or no longer be relevant facts.

Kuhn's work (Horwich, 1993) still carries weight but more as a means of evaluating the current state of science in a discipline or specific area of inquiry than as a brand of skepticism. There are several reasons. The radical empiricism and positivism of the 1920s–1950s have long since ceased to matter much, so they require no opponent. More significantly, although the notion of a paradigm shift is regularly applied loosely to support relativism, Kuhn himself rejected such uses of his thesis. He acknowledged underdetermination of theories by evidence, but he also claimed to believe in objective science. And, though Kuhn drew anti-realist conclusions from the theory-ladenness of observations that was a tenet of his thesis, scientific realists also oppose the strict dichotomy between observations and theory assumed by radical empiricists/positivists.

A continuing contribution of Kuhn's view has been his account of what considerations enter into replacing a paradigm. He was careful to limit the role of social, cultural, political, and other non-scientific factors in scientific revolutions, but he gave them a place at the table, albeit an inferior one compared with empiricism.

Postmodernism

Compared with underdetermination, theory-ladenness, and paradigm shifts, *postmodernism* is a very different, much broader, and more powerful skepticism. It is not specifically directed toward theories or models or even science itself but targets the very possibility of any kind of knowledge. Postmodern (anti-)epistemology has dominated many discussions of knowledge in western Europe and North America since the late 1970s. Disenchantment with modernism had increased in the West until it reached a crescendo at that point. Writers from various fields had earlier begun discussions of knowledge that would lead them beyond modernist epistemologies. Among the more influential of these writers were Jacques Lacan (1966/1977), Louis Althusser (1965/1969), Jacques Derrida (1967/1976), Michel Foucault (1984), Jean Baudrillard (1973/1975), and Roland Barthes (1970/1974). When they and their followers were not being original, they often based their considerations of knowledge on the work of Friedrich Nietzsche [1844–1900] and Martin Heidegger [1889–1976].

Postmodernism is primarily a reaction against the claims to universality and the metanarratives dominating rationalism from Descartes through the structuralists. Postmodernists characterize modernism in a negative manner. To them modernism gave the world science, reason, Western civilization, Marxism, Freudianism, and so on. Each of these modernisms is a totalism because it organizes everything under and in subjection to whatever the universal principle is in that system. So, each of these systems tells a grand story that relates everything to everything else by using the system's universal principle as the theme. By making everything subordinate to a master narrative, each of these systems places some ideas and some groups of persons at the center and others at the periphery of the story being told. In short, according to the postmodernists, power—not truth—is the underlying principle behind all modernisms.

Modernism privileges some ideas and people; it marginalizes others. Each modernism, then, is just a perspective that forces knowledge to take a particular configuration. There is no one perspective that can be privileged above any other except by a sheer power grab.

An epistemological response to skepticism

Underdetermination, theory-ladenness, and paradigm shifts were first articulated in the philosophy of science, a branch of philosophy separate from epistemology but with direct links to it. The philosophy of science has philosophical jurisdiction over models of targets of inquiry in the natural and social sciences, but so does epistemology in a more general way that draws on a longer tradition.

In his introduction to epistemology, philosopher Audi (1998) discusses at length how one could respond to becoming aware of the possibility of a nightmare scenario like Descartes' evil demon or Putnam's brain in a vat. For Audi, it is the prospect he is hallucinating the vivid green field that seems to lie in front of him. Even though it is impossible for him not to believe it is there, he is aware the field could be a total delusion. How should a person, especially an epistemologist, think about truth and skepticism in the

light of such scenarios? For our purposes, the types of skepticism about models that follow from underdetermination, theory-ladenness, and paradigm shifts can be seen as presenting a lesser but parallel predicament to the ones confronted by Descartes, Putnam, Audi, and countless undergraduates drinking beer with friends after midnight.

For Audi, the straightforward response would simply be to doubt everything that can be doubted and then withhold judgment about the doubts. This response, however, will not do, at least not for most of us. Why not? Because Audi (1998, p. 283) thinks most people—and I would suggest most literacy researchers—have two competing “deep-seated desire[s].” One is “to believe as many significant truths as [we] can,” and the other is “to avoid believing falsehoods.” The first desire “pushes us in the direction of credulity ..., believing too much.” The second “pushes us toward ... skepticism ..., believing too little.”

On balance, each of Audi’s desires is for something good. The desire for truth motivates learning and inquiry. If we believe we know much more about nature, history, and each other compared to equally intelligent humans who lived, say, a thousand years ago, surely we can recognize desire for knowledge as a major reason for the progress. On the other hand, skepticism can motivate us not to be arrogant about what we think we know or have been taught, and not to be satisfied with explanatorily insufficient answers to important questions. In the case of underdetermination, theory-ladenness, and paradigm shifts—as we saw earlier in this section—we probably do well to take their brands of skepticism seriously without attributing to them more destructive power than they possess.

Postmodernism may be another issue altogether. Cunningham and Fitzgerald (1996) placed it in a cluster of epistemologies opposite to positivism/radical empiricism but sharing that cluster’s severe skepticism of how much can be known. Postmodernism’s thoroughgoing incredulity toward Knowledge and Truth is defended with a detailed critique of the modern project’s greatest sins, but it leaves us no prospect of any truths or knowledge except what is locally held and often contradictory across locales.

What then is to be made of postmodernism? The epistemologist, pragmatist, and critical theorist, Jürgen Habermas [1929–], has perhaps been postmodernism’s most important interlocutor and critic. For him, debate with postmodernists should focus on modern social practices and institutions instead of abstractions and theories far removed from the lives of ordinary people. Habermas has argued against the postmodernists’ totalism, their refusal to distinguish between modernism’s beneficial and oppressive practices, institutions, and ideals. With respect to empirical inquiry, Rosenberg (2016) judges that the philosophy of social science can neither take postmodernism seriously nor refute it other than by ignoring its arguments. It must simply be assumed to be false, although doing so does not confront what it may be saying to social scientists.

The postmodernist discourse most awkward to dismiss is its rehearsal of cases of rationalized oppression. No doubt it would be inconsistent with its critique for anyone to pick and choose from that critique, but it seems educators, not least literacy educators, ignore this facet of its critique at the peril of students everywhere.

Implications of skepticism for literacy models

Skepticism can be a valuable foil for serious inquiry in literacy. We need not bother doing research for those who already agree with our literacy model; if our inquiry succeeds, they stand to gain little knowledge from us. We need not do research for those so skeptical of our model or method that nothing we find could possibly change their minds. Instead, an ideal audience for any literacy inquiry is fair-minded skeptics who disagree with our model but can be persuaded. Fair-minded skeptics inspire rather than deter well-designed, conducted, and interpreted inquiry in literacy. Skepticism has possible implications for model selection and building in literacy:

- Literacy models are probably best chosen or constructed, tested, and interpreted with an audience of fair-minded skeptics in view.
- Three related brands of skepticism toward models and methods—underdetermination, theory-ladenness, paradigm shifts—merit due diligence from literacy researchers. We should rigorously look for anomalies in our data and value multiple inquiries conducted in varied ways into any unobservable.
- Literacy models for inquiry and education stand to strengthen public goods. Even so, those of us who build or employ them should be warned by history that some actions we take may have precedents in modern oppressions past, leading us to interrogate the social and cultural motives and consequences of our work.

Epistemology of two literacy models

This article ends by briefly examining the epistemology of two well-known literacy models. The models chosen contrast significantly on the three issues of reductionism, unobservables, and skepticism.

Simple view of reading (SVR)

The often-cited SVR is a “model of reading” (Gough and Tunmer, 1986, p. 6). The model’s chief and titular characteristic is simplicity: “that reading consists of only two components, decoding and linguistic comprehension” (Hoover and Gough, 1990, p. 150). Its theoretical parsimony is achieved by *nothing but* explanatory reductionism empowered by idealization of modeling.

Specifically, the SVR's explanation of reading beyond decoding is reduced to what can be measured with a listening comprehension test, an assessment requiring no reading. The SVR is an idealized model because it assumes that other empirically demonstrable aspects of reading besides decoding can safely be discounted:

The simple view does not claim that exactly the same procedures used in linguistic comprehension are employed in reading comprehension, for there are clear differences Each of these differences is likely to be reflected in processing differences. The simple view, however, argues that these differences are relatively minor in comparison to the great similarities between linguistic and reading comprehension.

Hoover and Gough (1990, p. 153).

The SVR model is epistemologically defended by directing skepticism against what are supposedly the main contrasting views: that decoding ability and therefore decoding instruction are unnecessary for reading English; and that the comprehension reading requires is different from understanding spoken language in important ways. These skepticisms are defended by citing studies finding that measures of decoding and listening comprehension together predict a measure of reading comprehension significantly and to a large extent (Hoover and Gough, 1990). Consistent with its explanatory reductionism and skepticism toward complexity, there are no unobservables inferred or modeled in the SVR. Direct measures of the three abilities modeled are assumed to provide valid and sufficient observations.

RAND reading comprehension heuristic

The RAND Reading Study Group (RRSG, 2002) presented a "formal definition of reading comprehension" (p. 4) illustrated by a "heuristic" (p. xiii). Their heuristic is a figure (p. xiv) that certainly looks like a model (see Fig. 2), but the RRSG treated it as an aid to "thinking about reading comprehension" (p. xiii) rather than a limited representation to be critiqued or tested.

The RAND heuristic could appropriately be labeled, "the complex view." It models reading comprehension as three elements—reader, text, activity—interrelating within a sociocultural context, but this is just the beginning. The thinking about reading comprehension the heuristic is to assist applies to three "domains ...: instruction, teacher preparation, and assessment" (p. xiii). The literacy research it is to foster "needs to be informed by ... multiple perspectives (including educational, cognitive, linguistic, sociolinguistic, discourse analytic, and cultural perspectives)" (p. 2), and to be research that "contributes to better theories of reading development, and produces knowledge that is useable in both classrooms and policymaking arenas" (p. xiii).

The only skepticism of consequence expressed was the RRSG's judgment that reading comprehension theories and models to date were not "adequately rich ... to provide a coherent foundation" (p. 2) for the research needed. However, some of the multiple perspectives endorsed by the RRSG (e.g., cognitive) could be consistent with inferring unobservables in literacy theories or models produced with the assistance of their heuristic, while other of the perspectives (e.g., sociolinguistic) could be suspicious of such inferences.

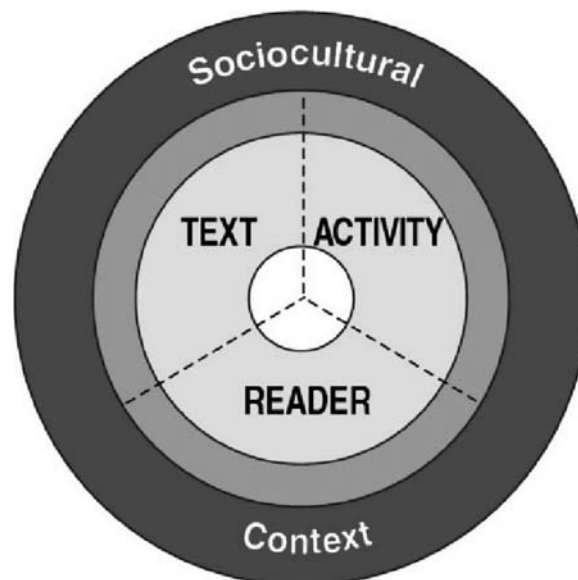


Fig. 2 RAND heuristic for thinking about reading comprehension. Reprinted from RAND Reading Study Group (2002).

Lessons

Models may sometimes embody a trade-off between explanatory reductionism and explanatory sufficiency, but neither the SVR's *nothing but* explanatory reductionism nor the RRSg's embrace of unprecedented complexity yields a sufficient explanation of reading comprehension as a phenomenon.

With its reliance on prediction and avoidance of causation, the SVR seems a vestige of logical empiricism's covering-law approach. As a result, the SVR has no answers for such practical questions as, "What cognitive mechanisms produce linguistic comprehension?" and "Does reading comprehension ability significantly influence listening comprehension ability?"

The RRSg sends us off in all directions at once to find the answers they claim are currently lacking. The SVR's model may be too simple, but it is difficult to imagine what would constitute a suitable model for the RRSg. The SVR is marked by skepticisms toward complexity, less idealized views, and unobservable constructs. The RAND heuristic includes little explicit skepticism, but the multiple perspectives the RRSg calls for in its research applications each has a core skepticism toward some or all of the other perspectives. Even research informed by multiple perspectives must choose skepticisms to endorse and ignore.

If the RAND heuristic was an attempt to bring about a paradigm shift in reading comprehension research, it appears to have failed. On the other hand, the SVR may be popular with some because it essentially undoes the paradigm shift known as the "cognitive revolution."

A model of literacy provides a limited representation of reading or writing, its components or causes, how it is learned, or how it is taught. Epistemology has relevance for understanding literacy models beyond research methods. Ultimately, literacy models are built in pursuit of knowledge and epistemology has the potential to inform all such endeavors.

References

- Althusser, L., 1969. For Marx (B. Brewster, Trans.). Allen Lane/The Penguin Press (Original Work Published 1965).
- Audi, R., 1998. Epistemology: A Contemporary Introduction to the Theory of Knowledge. Routledge.
- Barthes, R., 1974. S/Z (R. Miller, Trans.). Hill & Wang (Original Work Published 1970).
- Baudrillard, J., 1975. The Mirror of Production (M. Poster, Trans.). Telos Press (Original Work Published 1973).
- Bhaskar, R., 1986. Scientific Realism and Human Emancipation. Verso.
- Bogen, J., Woodward, J., 1988. Saving the phenomena. *Phil. Rev.* 97 (3), 303–352.
- Bridgman, P.W., 1927. The Logic of Modern Physics. Macmillan.
- Brown, A.L., Smiley, S.S., 1977. Rating the importance of structural units of prose passages: a problem of metacognitive development. *Child Dev.* 48 (1), 1–8. <https://doi.org/10.2307/1128873>.
- Brown, L. (Ed.), 1993. The New Shorter Oxford English Dictionary on Historical Principles. Clarendon Press.
- Comte, A., 2009. The Positive Philosophy of Auguste Comte (H. Martineau, Trans.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511701467> (Original Work Published 1853).
- Cunningham, J.W., Fitzgerald, J., 1996. Epistemology and reading. *Read. Res. Q.* 31 (1), 36–60.
- Davidson, D., 1970. Mental events. In: Foster, L., Swanson, J.W. (Eds.), *Experience and Theory*. University of Massachusetts Press, pp. 79–101.
- Davidson, D., 1980. *Essays on Actions and Events*. Clarendon Press.
- Dennett, D.C., 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Simon & Schuster.
- Derrida, J., 1976. Of Grammatology (G. C. Spivak, Trans.). Johns Hopkins University Press (Original Work Published 1967).
- Durkheim, E., 1982. W. D. Halls, Trans.). In: Lukes, S. (Ed.), *The Rules of Sociological Method*. The Free Press (Original Work Published 1895).
- Endicott, R.P., 1993. Species-specific properties and more narrow reductive strategies. *Erkenntnis* 38 (3), 303–321.
- Flew, A., 1984. *A Dictionary of Philosophy*, rev. second ed. St Martin's Press.
- Fodor, J.A., 1968. *Psychological Explanation: An Introduction to the Philosophy of Psychology*. Random House.
- Foucault, M., 1984. In: Rabinow, P. (Ed.), *The Foucault Reader*. Random House.
- Friedman, L.W., Friedman, H.H., Pollack, S., 2008. The role of modeling in scientific disciplines: a taxonomy. *Rev. Bus.* 29 (1), 61–67.
- Gauch Jr., H.G., 2002. *Scientific Method in Practice*. Cambridge University Press.
- Gleick, J., 2008. *Chaos: Making a New Science*. Penguin.
- Glennan, S., 2006. Explanation. In: Sarkar, S., Pfeifer, J. (Eds.), *The Philosophy of Science: An Encyclopedia*. Routledge, pp. 275–287.
- Gough, P.B., Tunmer, W.E., 1986. Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7 (1), 6–10.
- Hacking, I., 1999. *The Social Construction of What?* Harvard University Press.
- Hacking, I., 2001. *An Introduction to Probability and Inductive Logic*. Cambridge University Press.
- Hanson, N.R., 1958. *Patterns of Discovery*. Cambridge University Press.
- Hedström, P., Ylikoski, P., 2010. Causal mechanisms in the social sciences. *Annu. Rev. Sociol.* 36, 49–67.
- Hempel, C.G., Oppenheim, P., 1948. Studies in the logic of explanation. *Philos. Sci.* 15 (2), 135–175.
- Hempel, C.G., 1966. *Philosophy of Natural Science*. Prentice-Hall.
- Hesse, M.B., 1963. *Models and Analogies in Science*. Sheed and Ward.
- Hoover, W.A., Gough, P.B., 1990. The simple view of reading. *Read. Writ.* 2 (2), 127–160.
- Horwich, P., 1993. *World Changes: Thomas Kuhn and the Nature of Science*. MIT Press.
- Hruby, G.G., 2017. Literacy, comprehension, and the neurosciences. In: Israel, S. (Ed.), *Handbook of Research on Reading Comprehension*, second ed. Guilford, pp. 191–216.
- Hume, D., 2007. In: Millican, P.F. (Ed.), *An Enquiry Concerning Human Understanding*. Oxford University Press (Original Work Published 1748).
- Kant, I., 1966. *Critique of Pure Reason* (F. M. Muller, Trans.). Anchor Books (Original Work Published 1787).
- Kroeber, A.L., 1952. *The Nature of Culture*. University of Chicago Press.
- Kuhn, T.S., 1970. *The Structure of Scientific Revolutions*, second ed. University of Chicago Press.
- Lacan, J., 1977. *Écrits: A Selection [Writings: A Selection]* (A. Sheridan, Trans.). W. W. Norton (Original Work Published 1966).
- Lewes, G.H., 1875. *Problems of Life and Mind (First Series)*. Trübner.
- Lewis, D., 1970. How to define theoretical terms. *J. Philos.* 67 (13), 427–446.
- Locke, J., 1975. In: Niddich, P.H. (Ed.), *An Essay Concerning Human Understanding*. Clarendon Press (Original Work Published 1690).

- Mill, J.S., 1843. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and Methods of Scientific Investigation*. J. W. Parker.
- Oppenheim, P., Putnam, H., 1958. Unity of science as a working hypothesis. In: Feigl, H., Scriven, M., Maxwell, G. (Eds.), *Concepts, Theories, and the Mind-Body Problem*. University of Minnesota Press, pp. 3–36.
- Popper, K., 1965. *The Logic of Scientific Discovery*. Harper & Row.
- Putnam, H., 1967. Psychological predicates. In: Capitan, W.H., Merrill, D.D. (Eds.), *Art, Mind, and Religion*. University of Pittsburgh Press, pp. 37–48.
- Putnam, H., 1981. *Reason, Truth, and History*. Cambridge University Press.
- Quine, W.V.O., 1951. Two dogmas of empiricism. *Phil. Rev.* 60 (1), 20–43.
- RAND Reading Study Group, 2002. *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. RAND. https://www.rand.org/pubs/monograph_reports/MR1465.html.
- Rosenberg, A., 2016. *Philosophy of Social Science*, fifth ed. Routledge.
- Rothchild, I., 2006. Induction, Deduction, and the Scientific Method: An Eclectic Overview of the Practice of Science. Society for the Study of Reproduction. https://higherlogicdownload.s3.amazonaws.com/SSR/fbd87d69-d53f-458a-8220-829febf990b/UploadedImages/Documents/rothchild_sciemethod.pdf.
- Van Fraassen, B.C., 1980. *The Scientific Image*. Clarendon Press.

Internet resources (uncited in paper)

- Internet Encyclopedia of Philosophy, iep.utm.edu.
- Stanford Encyclopedia of Philosophy, plato.stanford.edu.

Are we there yet? Onto-epistemological qualitative language and literacy research

Marla H. Mallette^a, Jillian Mason^b, and Matthew McConn^a, ^a Department of Teaching, Learning, and Educational Leadership, College of Community and Public Affairs, Binghamton University, Vestal, NY, United States; and ^b Communications Department, Division of Liberal Arts and Education, Arizona Western College, Yuma, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Qualitative literacy and language research	895
Counter narratives	897
Personal stories	897
Stories of others	897
Composite stories	898
Intersectionality	898
Translanguaging, transliteracies and transnational	899
Conclusion	900
References	900

Qualitative literacy and language research

Qualitative research has a long and rich history in literacy and language education. As detailed by Barone (2021), early qualitative (i.e., case study) research tended to be observational in which researchers detailed language and literacy activities, typically with their own children or family members (see e.g., Butler, 1975; White, 1956). Throughout the 1970s and 1980s, interest in qualitative research among scholars grew; however, acceptance among scholars was a slow and gradual process. As Reutzel and Mohr (2015) pointed out in their sample of 200 randomly selected studies spanning the 50 years of *Reading Research Quarterly* (RRQ), from 1965 to 1994, quantitative research was published at a rate of four to one, whereas, from 1995 to 2014, it was closer with a rate of 5 to 4. However, in a more nuanced analysis of publications across nine literacy journals from 2009 to 2014, Parsons et al. (2016) found that in RRQ, 58% of the studies were quantitative, 23% were qualitative, and 8% mixed methods and in JLR 30% of the studies were quantitative, 49% were qualitative, and 8% mixed methods.

In considering a historical perspective of qualitative research publications in language and literacy, it is important to note that the depth of qualitative research resulted in lengthier manuscripts, which often exceeded journal page limits, and as such, it was not uncommon for researchers to publish their work in books, allowing them to preserve the richness of their data. Despite the overall disparity in published qualitative research, the importance of qualitative research in understanding language and literacy is unequivocal. For example, the observational data from Clay (1967) is considered the impetus for research exploring the construct of emergent literacy. Dyson's (1997) case study research and Heath's (1983) ethnography remain still today as landmark studies into understanding family literacy. Emig's (1971) study of writing among twelfth-grade students, which she described as a "unique effort to utilize case studies for eliciting data about how students behave as they write" (p. 5), is considered the foundation in understanding writing as a *process*.

Looking back, the depth and insight into understanding language and literacy advanced through qualitative research is profound. Further, as qualitative research evolves, variations of, for example, ethnography or participatory action research, although not exactly or entirely new, can be characterized as leading-edge innovations, continuously moving the field forward. As we will show, largely due to researchers' increased attention to both learner and researcher identity and intersectionality, as well as to more layers of context (see e.g., Smagorinsky, 2018), the meaning of literacy itself has become more nuanced. What researchers are able to see is expanding as the ways they look expand, as their methodologies continue to stretch beyond assessing literacy practices to include learning more about who literacy learners are. The resulting insight into the lived experience and identity of the literacy learner also widens and deepens as the identity of the researcher and scope of inquiry increasingly intertwine—ultimately revealing an expanded view of the literacy learner. The direction of literacy and language studies toward genuine inclusivity seems to be progressing in accordance with an awareness that literacy cannot be truly understood, nor truly inclusive, unless the reality of the learner, from the learner's perspective, is at the heart of the research.

This deepened and broadened understanding of the nature of literacy, and of this expanded view of the literacy learner, seems natural in hindsight, given the nature of the iterations or delineated "moments" in qualitative research as presented by Denzin and Lincoln (2000, 2008, 2017) through the five editions of their esteemed *Sage Handbook of Qualitative Research* series. Importantly, although admittedly somewhat paradoxically, they do not characterize the progression of qualitative research as a typical evolution. That is, while there is a chronological nature to the moments they name and define, they are decidedly not linear, and the more recent "advancements" are not emphasized as "improvements." As Denzin and Lincoln (2000) describe, their historical moments "overlap and simultaneously operate in the present" (p. 2). This language is important because it emphasizes a sense of qualitative

research as maturing and developing, and as continuing to build on existing and established tradition—yet without the implication that earlier iterations were immature or underdeveloped. To illuminate this somewhat counterintuitive notion of development, we offer the theoretical framework of consciousness philosopher Wilber's (2000) integral metatheory, which also helps to explain how recent trends we notice in qualitative literacy and language research seem simultaneously new and old. This lens also offers a glimpse into how we may begin to imagine a nonlinear experience of the past, present, and future of qualitative research, akin to Denzin and Lincoln's notion of historical moments simultaneously operating in the present. Wilber's integral theory emerged from his meta-analysis of over a hundred human development theories, and what is most relevant here about developmental stage theories is that they are "growth" or "actualization" hierarchies, as opposed to "dominator" hierarchies. This means that each stage "transcends but includes" or "differentiates and integrates" the previous stages. New stages are qualitatively "better" only in the sense that they are more complex and thus more capable of insight, perspective, understanding, and inclusion, and so as such, they are simultaneously able to see the limitations of previous stages *and* also deeply cherish them. In keeping with Denzin and Lincoln's contention, there is the inherent belief that lower or previous stages, offered here as parallels to the qualitative moments and their methodologies and offerings, are not "wrong," but are only partial:

The reason that earlier wisdoms need to be remembered is that in a healthy development each stage "transcends but includes" its predecessor: it goes beyond or brings forth new and emergent truths not previously found, but it also needs to include and embrace the important previous truths themselves, at pain of developmental dysfunction and disharmony.

Wilber (2017, p. 14).

To a certain extent, such a view also minimizes the relevance of distinctions between past, present, and future, while also serving to open epistemological possibilities for both prediction and hopeful aspirations for where qualitative research may continue to lead. Indeed, in their third edition (2008), Denzin and Lincoln speculate that qualitative research might be "undergoing a stunning compression of time, of moments" (p. 541). We cannot help but notice how we are living the very predictions and calls presented by Denzin and Lincoln, as well as their caveat that the revisiting of earlier moments "operates at different levels of abstractions" (2000, p. 1050)—a claim also made clearer by thinking in terms of growth hierarchies. All of the moments Denzin and Lincoln identified are still with us, and where they ended in their most recent edition is both where we are and where we still seem to be going.

Put another, more specific, way, there is a sense that each moment was born of a felt need to improve upon what preceded. For instance (though this is an oversimplification), the "traditional period" of the first half of the twentieth century blossomed into the "blurred genres" of 1970–1986, when "qualitative researchers had a full complement of paradigms, methods, and strategies to employ in their research" (Denzin and Lincoln, 2000, p. 15). The boundaries between social science and the humanities blurred, and the researcher's increased presence in the interpretive text, along with a concern that too much blurring and the narrative turn would discredit the legitimacy of qualitative research, gave way to the "crisis of representation," which was answered by the "post-modern period of experimental ethnographic writing" from the late 1980s through 1995. When trying to describe the then "future" on the horizon, which they labeled "the future (now -)," Denzin and Lincoln began using the word "sacred" rather liberally, both envisioning and calling for a future "concerned with moral discourse, with the development of sacred textualities ... sites for critical conversations about democracy, race, gender, class, nation-states, globalization, freedom, and community" (2000, p. 3). Moreover, their emphasis on qualitative research as defined by the interconnectedness of the "generic activities" of theory, method, analysis, ontology, epistemology, and epistemology that individual researchers uniquely bring to life, while another "not new" idea, was a prescient nod to the decreasing willingness of researchers to maintain boundaries between these components. As we will show, the onto-epistemological stances that we see so prevalent of late reveals an underlying assumption that ways of and beliefs about knowledge and knowing cannot be neatly parsed out from those about reality. That is, the nature of reality and of knowing cannot be considered separately, and we find this belief widely evident as insights from quantum physics have prominently arrived on the qualitative research landscape (Monk-Turner, 2020; Sailors et al., 2018). As Barad (2003) explains:

The primary epistemological unit is not independent objects with inherent boundaries and properties but rather phenomena. ... phenomena do not merely mark the epistemological inseparability of "observer" and "observed;" rather, phenomena are the ontological inseparability of agentially intra-acting "components." ... the "knower" does not stand in a relation of absolute externality to the natural world being investigated—there is no such exterior observational point. We are not outside observers of the world. Nor are we simply located at particular places in the world; rather, we are part of the world in its ongoing intra-activity (p. 815).

Barad's depiction, we suggest, resonates with Denzin and Lincoln's notions of sacredness, their imagined future of inquiry that "erases traditional distinctions among epistemology, ethics, and esthetics ... seek[ing] to ground the self in a sense of the sacred ... to illuminate the unity of the self in its relationship to the reconstructed, moral, and sacred natural world ... seek[ing] a sacred epistemology that recognizes the essential ethical unity of mind and nature" (2000, p. 1052). Together, these lines of thinking make the onto-epistemological framework that we utilize here.

Notably, in their third edition (2008), Lincoln and Denzin redubbed their "the future (now-)/out of the past" seventh moment into the "methodologically contested present" and their new imagined future became their projected eighth and ninth moments of "qualitative research in/and the fractured future." They revised an assertion from an earlier edition regarding qualitative research's coming of age, noting that it actually "had merely reached a zesty and robustly athletic late youth," and they predicted that the

genuine coming of age will be marked by a forthcoming maturation into “a new praxis that is deeply responsive and accountable to those it serves” (pp. 549–550). Similarly, and most recently, in their fifth edition (2017), they note an important shift of critical stances in interpretation and representation: that, historically, “critical perspectives were, in some cases, simply a companion, not a conjoined voice. That has changed” (p. 924). We find the final lines of their epilogue not only a compelling call for but an apt description of what we see in contemporary literacy research as situated in our current social and political landscape:

Far from being some imaginary end point, we are in fact at the edge of a new colonialism, a new era, one that we did not fully choose, and one that we must begin to understand more fully than we have to this point. The only meaningful method for that understanding is a refashioned, refunctioned, repurposed imaginary for ethnography and ethnographers. And that is yet to be invented (p. 927).

Considered in the context of their arc of their looking back and forward across moments, we interpret this new imaginary and invention in the established spirit of continuing the “re”-visiting, and at unprecedented new “levels of abstraction” (2000, p. 1050). Nestled in this interpretation, we began our inquiry with the consideration of what moments are featured in current qualitative language and literacy research. We chose to focus this analysis on qualitative research published in US based journals in the past 5 years (2015–2020) in the *Journal of Literacy Research (JLR)* and *Research in the Teaching of English (RTE)*, as these journals are highly regarded, have the highest prevalence rates of published qualitative studies, and span a range of age groups and topics. The qualitative studies published during this time frame encompassed a range of methodologies and were reflective of the many moments in qualitative research. However, what also became clear was the growing number of studies demonstrating Denzin and Lincoln’s premonition that we are arriving at where we are going. Thus, in our efforts to look forward, we turn the focus for this chapter to a consideration of how qualitative researchers seeking to understand language and literacy phenomena demonstrate clear commitments to social justice, equity and empowerment; dissolving and crossing boundaries; probing intersectionalities; and interrogating the intra-activity of people, materials, time, and place.

As anticipated, and in keeping with the onto-epistemological framework used in guiding this analysis, we discovered that the theoretical frames researchers used in guiding onto-epistemological language and literacy research overlapped, intra-acted, and intersected. The most prevalent theoretical frames we found in this body of research, which parallel Tierney’s (2018) “Dimensions of Global Meaning Making” were: (a) counter narratives, (b) intersectionality, and (c) “trans-” (e.g., translanguaging, transliteracies). There are a few caveats essential in noting to provide context to the discussion of these theoretical frames. First, and as with any theoretical shift, researchers’ portrayals of onto-epistemology varied, ranging along a continuum from identifying/labeling/defining to more integrated and holistic representations. Second, the theoretical frames, by the very nature of an onto-epistemological framework, are not discreet. Yet, for the purpose of providing a coherent discussion, we discuss each one separately.

Counter narratives

Solorzano and Yosso (2002) define counter narratives (or counter storytelling) “as a method of telling the stories of those people whose experiences are not often told (i.e., those on the margins of society)” (p. 32). These narratives challenge the status quo by countering the dominant discourse and exposing and confronting racial privilege. Solorzano and Yosso delineate three specific types of counter narrative: (a) personal stories, (b) stories of others, and (c) composite stories.

Personal stories

In Baker-Bell’s (2017) personal counter narrative, she situated her 1st years in academia within her life experiences from childhood as a young black into adulthood as a black woman. Through the complexity of the strong black woman myth, Baker-Bell’s narrative describes how she “existed, resisted, and persisted during ... on the tenure track as a Black woman and emerging language and literacy scholar with a family” (p. 540). Jang (2017) also used a personal counter narrative to describe how, in his 1st years in academia, his encounters with deficit-based racialized discourse provided insight into how he positioned himself and how he was positioned by others; recognizing how such positioning was both detrimental and potentially advantageous. In teaching literacy, he was questioned “how can you teach us this course even though you are a non-native speaker? It’s all about decoding stuff” (p. 571); however, as a quantitative researcher he was respected. These counter narrative personal stories highlight the importance of understanding the researcher as we seek to better understand language and literacy.

Stories of others

In a 3-year ethnographic study, San Pedro’s (2015) counter narrative challenged the dominant discourse through his depiction of how the use of silence with Native American students was a means for the students to enact their agency. As San Pedro noted, “storying allowed us—participants and me—space to share our developing understandings of the intentionality of silence” (p. 137). Using a life history interview approach, Vieira’s (2016) counter narrative illuminated how the life experiences among migrant families shape their literacy use upon returning to their homeland residence. Grounded in the theoretical framework of sociomateriality, Vieira captured how literacy cannot be separated from economics, politics, time, nor space.

Furthermore, in two interconnected ethnographic case studies, Ellison and Solomon (2019) used a counter-storytelling framework to analyze the ways two African American children and one parent talk about digital literacies, race, and the digital divide. By

doing so, the researchers highlighted how moving away from the deficit thinking of the digital divide provided the space for understanding, “the ways African American children are using technologies-not as passive content consumers, but as active and creative content producers whose activities support and transform their own learning” (p. 238). And finally, [Lam et al. \(2021\)](#) used a multi-modal storytelling lens to study the construction of scale. Through their research, they “contribute to literacy learning and teaching that support young people in bringing their knowledge, experiences, and narrative resources to engage with societal structures. ... In this way, we think that media production offers opportunity for students to document and mobilize their own and their communities’ language and semiotic repertoires, and to analyze and orchestrate them in relation to other community and institutional voices to extend important conversations in society” (p. 362).

Composite stories

In a composite narrative, researchers use data to create characters and then place these characters in social and political contexts. Using a version of this approach, [Tolliver \(2020\)](#) constructed a counter narrative of Avenae’J though integrating the work from her short story, fieldnotes from observations of writer’s workshop, and interviews. Tolliver included Avenae’J’s story in its entirety. In reflecting on her story, Avenae’J explained:

When I was writing this story, my problem was that people, a lot of African Americans, were being killed and then not being represented because of color. There is a lot of African American women who are not being represented even if they’re being killed just because they are African American. ... I thought that in 10 years, if it did not stop, there was going to be more killings ... In 50 years, there would be a civil war between different groups. Multiple groups fighting against multiple groups (p. 522).

Avenae’J further explained the main character in the story, “Avenae’J is one of those people ... well, she’s a lot like me. I just took myself and put me in my own story. That is why her name is my name. I like to put myself into my writing because it makes me feel better ... not feel better, but it makes me understand the story more” (p. 523). Through counter narratives, we hear the untold and unheard stories, and as Tolliver concluded, “We tell our stories to envision worlds beyond the one we currently inhabit and to imagine more just futures. Through Avenae’J story, she shared her understandings and experiences of the world, and she provided readers an opportunity to listen and respond” (p. 526).

Intersectionality

Building on Crenshaw’s original conception of intersectionality as it applied to law, [Cho et al. \(2013\)](#) delineate how this construct is now used widely across numerous disciplines. They conclude by locating research in intersectionality as follows:

By focusing on structures of power that constitute subjects in particular sociopolitical formations, we locate intersectional dynamics in social space and time. This does not mean, however, that subjects are simply structural positions. It does mean that debates in intersectional studies will circulate less around categories and identities and more around how those categories and identities and their specific content are contingent on the dynamics under study or of political interest (p. 807).

[Sciurba’s \(2017\)](#) study of male readers of color, illuminates how intersectionality provides fluidity and depth in understanding reader response and identity. The high level of engagement the participants demonstrated in two very different types of text captured the variability in the ways they connected to these texts. At times the boys connected the texts to their own lives as with the texts serving as windows and/or mirrors or the boys just found other relevant connections. “Even if the boys did not see their current worlds reflected while reading the stories, they became ‘active authors’ of their own possible worlds or found value in learning about how their current and future worlds (might) evolve. Their ‘personal and social identities’ proved to be multifaceted, too—influenced not only by race/ethnicity and gender but by popular culture” (p. 389). Similarly, [Compton-Lilly et al. \(2017\)](#) demonstrate how intersectionality offers a more nuanced understanding of literacy and identity among young children. Building on the premise that identities are fluid, contextual, and social, Compton-Lilly et al. demonstrate how intersectionality exposes how young children’s identity negotiations “are deeply situated in social contexts that bring long histories of privilege” (p. 122). In terms of early literacy

... intersectionality matters. It reveals how the teaching of reading, writing, and other literacy practices is intricately interwoven with networks of self that extend beyond classrooms. For young children, this intersectionality is not merely about the intersectional networks of race, gender, social class, and language the children bring but also about how those ways of being are taken up, valued, and negotiated by both the children and the people they encounter. As children negotiate networks of self, adults must not only support children in negotiating this multiplicity but also view children as active and invested individuals who have the capacity to contribute to their own becomings (p. 135).

Through a framework that included gender, masculinity, language, discourse, and identities, [Qin \(2018\)](#) explored the language learning experiences of Tiger, a 15-year-old student originally from Taiwan.

Tiger's masculinity performances were deeply intertwined with his participation in the classroom interactions, got entangled with the process of language learning and relations of power, and complicated his language learning experience and trajectory, leading to him being labeled as a "not serious" and "problem" student. While his masculinity performances took up and reinforced the dominant discourses of doing masculinity, his way of doing funny challenged the routinized language instruction, the censoring pedagogy, and the colonizing "good learner" discourses imposed and legitimized by the institutional power of the school (p. 449).

Translanguaging, transliteracies and transnational

Stornaiuolo et al. (2017) developed a framework to offer a conceptual understanding of transliteracy in a global world. Drawing from the work of scholars across many disciplines, the prefix trans is often thought of as across or beyond of through. In relation to literacy, Stornaiuolo, et al. appreciate the ambiguity it adds as "—it leaves unspecified who or what is moving or being moved, and how. Without assuming the relationships between what moves in advance, a trans-focus spotlights the activity of creating, maintaining, and disassembling associations across spacetimes" (p. 73). In further theorizing transliteracy, they suggest the following:

We propose that affixing trans- to the root word literacies foregrounds particular dimensions of literacies that can prove generative for research and analysis as literacies are lived, explored, described, analyzed, and enacted in critical and creative social practices, situated within and across ideological systems and characterized by the movements and interactions of people and things. This framing rests on foundational sociocultural claims that literacies are not located solely within individuals' minds (i.e., as skills or abilities) but are developed and constructed in distributed social practices constitutive of all human activity.

Duran (2017) used transliteracies to deepen our understanding of audience awareness among young writers. The students participated in a writing workshop approach modified to focus on audience awareness. Each unit incorporated different genres of writing and a range of audiences (e.g., family, peers, community member). At the start of the year, operating from a deficit model based on test-scores, except for one student, all students were positioned as at-risk learners. However, the teacher's willingness to embrace a transliteracies perspective provided a framework to shift from a deficit perspective to a gifted one. As evidenced in the findings, "students displayed rhetorical astuteness as they chose and combined language(s) in nuanced ways" (p. 107).

In a think aloud study using a transliteracies framework, Garcia and Godina (2017) explored the strategies fourth-grade, bilingual students used in interpreting and discussing Spanish and English texts. Findings indicated that the students varied in their cognitive approaches; however, they were similar in language approaches. That is, the students frequently used one language to explain their reading in the other language and they rarely used cognates, and they displayed a wider range of strategies than student who speak only one language. The findings from this study provide compelling evidence that translanguaging "appears to characterize how the fourth graders in our study approached, discussed, and interpreted texts, especially if we use Daniel and Pacheco's (2016) definition of translanguaging as "moving across languages and registers of speech to make meaning" (p. 653)" (p. 295).

In a critical ethnographic case study, Rounsaville (2017) used translanguaging to capture writing across a lifespan. Through this perspective, Rounsaville was able to demonstrate that writing is more than text-type; rather writing across one's lifetime encompasses how genres are inherited, created, and moved and are inclusive of transnational links. As Rounsaville explains:

I found that Clara built her repertoire over time, generations, and contexts. Stated attachments and assigned significances indexed her transnational aims and concerns, which accrued as patterned social actions—as defined by Clara—guiding and governing how she moved genre across time, space, generation, and text-type. ... As Clara's case has shown, tracing the actions assigned to texts from an emic perspective—through the method of genre mapping—enabled an expansive look at how "here" and "there" were coordinated. Specifically, findings revealed how orientations to literacy inherited and made in one childhood persisted over time and across contexts and were handed down to the children of other migrant families. Social actions assigned to writing, accrued through childhood and young adult writing, became a consequential bridge across difference for Clara's work as a bilingual literacy educator (p. 334).

de los Ríos and Seltzer (2017) provide additional compelling evidence that emphasize the importance of translanguaging pedagogy to improve writing instruction. de los Ríos and Seltzer center their argument on how this perspective shifts the characterization of bilingual learners from being marginalized and often characterized as at-risk to one that recognizing the strengths in their linguistic abilities, providing opportunities, and creating the space needed to situate their linguistic and cultural abilities at the core of their learning. To truly capitalize on this approach, de los Ríos and Seltzer focused the topic of the writing on challenging coloniality. Building on the importance of intersectionality, the students engaged in positive writing experiences, and were able to "imagine themselves as decolonial subjects whose futures will be on their own terms. A decolonial imaginary means moving past simply 'allowing' students to draw on their everyday language practices and instead centers instruction and curricula around who and where students are as historically colonized and racialized subjects" (p. 70).

In a two-part study focused on improving instructional practices, Stewart and Hansen-Thomas (2016) sought first to identify what transnational experiences were evident in a participant's life and literacies, and then to explore how the student responded to instructional practices that leveraged her transnational life and literacies through translanguaging. Their findings led them to conclude that, "Sanctioning a space for translanguaging as part of a 'multilingual awareness pedagogy' enables emergent bilingual

transnational students to engage in greater creativity and criticality as they engage their full linguistic repertoires. Translanguaging facilitates students' use of higher-order thinking to make decisions and evaluate all linguistic options available to them in their writing" (p. 467).

Conclusion

Using an onto-epistemological framework, qualitative researchers deepen our understanding of language and literacy through research that enacts many of Tierney's (2018) "Dimensions of Global Meaning Making." Through Counter Narratives researchers interrupted existing frames through their stories that disrupted the colonized status quo. Using an intersectionality framework, researchers interrogated truth/post-truth by challenging orthodox positionalities with multifaceted and complex depictions of identities and positionality. Through a focus on translanguaging, transliteracies and transnationalism, researchers "decolonized spaces," as they disrupted dominant political and cultural spaces.

Through their qualitative research, these scholars aimed to "disrupt the normative—an imaginary that challenges the colonized status quo" (Tierney, 2018, p. 407) as they deconstructed the researcher/researched dichotomies and afford readers the opportunity to see language and literacy through learners whose stories are often untold. Drawing on Yeats' poem *Among School Children*, in which Yeats asks, "How can we know the dancer from the dance," Bloome and Green (2015) extend the imagery of the inseparability of the dancer and the dance to broaden our understanding of literacy and literacy research. They suggest, "Literacy cannot be separated from what people are doing, how they are doing it, when, where, under what conditions and with whom they are doing it; metaphorically, there is no separation of the dancer from the dance" (p. 20).

The deconstruction of binary thinking, of course, is another not-new idea. But what is new is the evolving, subtly different onto-epistemological stance that can hold contradiction or inconclusively as not necessarily or exclusively a temporary problem to eventually figure out, or as warranting some sort of gracious open-minded conciliatory gesture or relativistic "agree to disagree" accommodation on principle, but instead as a welcome and even satisfying indication of how much remains invisible, or yet unconsidered, or even in flux (This goes beyond acknowledging gaps or limitations in the research or identifying areas for future study. This also goes beyond the deconstruction that accompanies the postmodern failings of language to represent reality, though of course both are related and relevant). This evolving onto-epistemological stance, then, is not fueled by the drive for more information or refining of analysis to eventually reveal which of the seemingly contradictory pieces can be eliminated or otherwise solved; instead, this stance is eagerly patient to wait and see how, with more information and understanding, what may emerge is a way to comprehend how all of the information can make sense as part of the larger constellation. Lack of understanding becomes, paradoxically, part of the package of understanding. Not a void to be filled, that is, but a part of the landscape. Meanwhile, it nonetheless remains as important as ever to continue pursuing knowledge and understanding.

As Barad explained:

We do not obtain knowledge by standing outside of the world; we know because "we" are of the world. We are part of the world in its differential becoming. . . . Onto-epistemology—the study of practices of knowing in being—is probably a better way to think about the kind of understandings that are needed to come to terms with how specific intra-actions matter.

Barad (2003, p. 829).

Are we there yet? Perhaps not, but it seems language and literacy qualitative researchers are heading in that direction as the studies highlighted in this piece show promising signs of scholarship that advances "an approach to our worlds that challenges authoritarian circumstances that attempt to control knowledge narrowly or position other cultures as having fake or lesser forms of narrative" (Tierney, 2018, p. 413).

References

- Baker-Bell, A., 2017. For Loretta: a Black woman literacy scholar's journey to prioritizing self-preservation and Black feminist–womanist storytelling. *J. Lit. Res.* 49 (4), 526–543. <https://doi.org/10.1177/1086296X17733092>.
- Barad, K., 2003. Posthumanist performativity: toward an understanding of how matter comes to matter. *Signs* 28 (3), 801–831. <https://doi.org/10.1086/345321>.
- Barone, D., 2021. Exploring lenses used in case study research in literacy over time. *Qual. Rep.* 26 (2), 386–411. <https://doi.org/10.46743/2160-3715/2021.4617>.
- Bloome, D., Green, J., 2015. The social and linguistic turns in studying language and literacy. In: Rowsell, J., Pahl, K. (Eds.), *The Routledge Handbook of Literacy Studies*. Routledge, pp. 19–34.
- Butler, D., 1975. *Cushla and Her Books*. The Horn Book.
- Clay, M.M., 1967. The reading behavior of five-year-old children: a research report. *N. Z. J. Educ. Stud.* 2 (1), 11–31.
- Compton-Lilly, C., Papoi, K., Venegas, P., Hamman, L., Schwabenbauer, B., 2017. Intersectional identity negotiation: the case of young immigrant children. *J. Lit. Res.* 49 (1), 115–140. <https://doi.org/10.1177/1086296X16683421>.
- Cho, S., Crenshaw, K.W., McCall, L., 2013. Toward a field of intersectionality studies: theory, applications, and praxis. *Signs* 38 (4), 785–810. <http://www.jstor.org/stable/10.1086/669608>.
- de Los Rios, C.V., Seltzer, K., 2017. Translanguaging, coloniality, and English classrooms: an exploration of two bicoastal urban classrooms. *Res. Teach. English* 55–76. <https://www.jstor.org/stable/44821287>.

- Denzin, N.K., Lincoln, Y.S. (Eds.), 2000. *Handbook of Qualitative Research*, second ed. Sage Publications.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2008. *The Landscape of Qualitative Research*, third ed. Sage Publications.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2017. *The SAGE Handbook of Qualitative Research*, fifth ed. Sage Publications.
- Durán, L., 2017. Audience and young bilingual writers: building on strengths. *J. Lit. Res.* 49 (1), 92–114. <https://doi.org.proxy.binghamton.edu/10.1177/1086296X16683420>.
- Dyson, A., 1997. *Writing Superheroes: Contemporary Childhood Popular Culture and Classroom Literacy*. Teachers College Press.
- Ellison, T.L., Solomon, M., 2019. Counter-storytelling vs. deficit thinking around African American children and families, digital literacies, race, and the digital divide. *Res. Teach. English* 53 (3), 223–244.
- Garcia, G.E., Godina, H., 2017. A window into bilingual reading: the bilingual reading practices of fourth-grade, Mexican American children who are emergent bilinguals. *J. Lit. Res.* 49 (2), 273–301. <https://doi.org/10.1177/1086296X17703727>.
- Heath, S., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press.
- Jang, B.G., 2017. Am I a qualified literacy researcher and educator? A counter-story of a professional journey of one Asian male literacy scholar in the United States. *J. Lit. Res.* 49 (4), 559–581. <https://doi.org/10.1177/1086296X17733491>.
- Lam, W.S.E., Chang, A.A., Rosario-Ramos, E.M., Smirnov, N., Easterday, M.W., Doppelt, J.C., 2021. Multimodal voicing and scale-making in a youth-produced video documentary on immigration. *Res. Teach. English* 55 (4), 340–368.
- Monk-Turner, E., 2020. Epistemology, social inquiry and quantum theory. *Qual. Res. J.* 20 (2), 228–242. <https://doi.org/10.1108/QRJ-10-2019-0085>.
- Parsons, S.A., Gallagher, M.A., George Mason University Content Analysis Team, 2016. A content analysis of nine literacy journals, 2009–2014. *J. Lit. Res.* 48 (4), 476–502. <https://doi.org/10.1177/1086296X16680053>.
- Qin, K., 2018. "Doing funny" and performing masculinity: an immigrant adolescent boy's identity negotiation and language learning in one US ESL classroom. *Res. Teach. English* 52 (4), 427–454, 130097523.
- Reutzel, R., Mohr, K.A., 2015. 50 years of reading research quarterly (1965–2014): looking back, moving forward. *Read. Res. Q.* 50 (1), 13–35. <https://doi.org/10.1002/rrq.87>.
- Rounsaville, A., 2017. Genre repertoires from below: how one writer built and moved a writing life across generations, borders, and communities. *Res. Teach. English* 51 (3), 317–340. <http://www.jstor.org/stable/44821265>.
- Sailors, M., Martinez, M., Stortz, R., Trevino, C., Davis, D.S., Jones, J.S., Goatley, V.J., Van Cura Monaco, C., 2018. Invoking quantum physics: fifty volumes of methodological complexity in literacy research. *J. Lit. Res.* 50 (1), 3–8. <https://doi.org/10.1177/1086296X18754892>.
- San Pedro, T.J., 2015. Silence as shields: agency and resistances among Native American students in the urban Southwest. *Res. Teach. English* 132–153. <https://www.jstor.org/stable/24890030>.
- Sciurba, K., 2017. Journeys toward textual relevance: male readers of color and the significance of Malcolm X and Harry Potter. *J. Lit. Res.* 49 (3), 371–392.
- Smagorinsky, P., 2018. Literacy in teacher education: "It's the context, stupid.". *J. Lit. Res.* 50 (3), 281–303. <https://doi.org/10.1177/1086296X18784692>.
- Solórzano, D.G., Yosso, T.J., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1), 23–44. <https://journals.sagepub.com/doi/pdf/10.1177/107780040200800103>.
- Stewart, M.A., Hansen-Thomas, H., 2016. Sanctioning a space for translanguaging in the secondary English classroom: a case of a transnational youth. *Res. Teach. English* 450–472. <https://www.jstor.org/stable/24889944>.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Lit. Res.* 49 (1), 68–91. <https://doi.org/10.1177/1086296X16683419>.
- Tierney, R.J., 2018. Toward a model of global meaning making. *J. Lit. Res.* 50 (4), 397–422. <https://doi.org/10.1177/1086296X18803134>.
- Toliver, S.R., 2020. Can I get a witness? Speculative fiction as testimony and counterstory. *J. Lit. Res.* 52 (4), 507–529. <https://doi.org/10.1177/1086296X20966362>.
- Vieira, K., 2016. Writing remittances: migration-driven literacy learning in a Brazilian homeland. *Res. Teach. English* 50 (4), 422–449. <http://www.jstor.org/stable/24889943>.
- White, D., 1956. *Books Before 5*. Oxford University Press.
- Wilber, K., 2000. *Integral Psychology: Consciousness, Spirit, Psychology, Therapy*. Shambhala.
- Wilber, K., 2017. *Trump and a Post-truth World*. Shambhala.

Epistemologies and methodologies in literacy research: a case study

Seth A. Parsons^a, Ashley N. Patterson^b, Jamie Colwell^c, Bong Gee Jang^d, and Betsy A. Baker^e, ^a George Mason University, Fairfax, VA, United States; ^b Penn State University, University Park, PA, United States; ^c Old Dominion University, Norfolk, VA, United States; ^d Syracuse University, Syracuse, NY, United States; and ^e University of Missouri, Columbia, MO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	902
About LRA	903
Methods	903
Data collection and analysis	903
JLR content analysis (2007–2019)	903
LR:TMP content analysis (2015–2019)	904
Self-reported epistemologies and methodologies in the conference submission system (2019 & 2020)	904
Results	904
JLR	904
LR:TMP	904
Self-reported epistemologies and methodologies	904
Discussion	905
Conclusion	908
Appendix: Options for self-reported epistemologies and methodologies for LRA conference submission	908
Epistemology	908
Theoretical perspectives derived from/rooted in...	909
Method	909
Setting	909
Acknowledgments	909
References	910

Introduction

The field of literacy is currently facing new, complex, and evolving problems (Beach & O'Brien, 2018; Moje et al., 2020). Conceptualizations of literacy are shifting, technologies are advancing, schooling is transforming to virtual and hybrid spaces, information and misinformation in the media and on social media is ever present, policies are being informed by outspoken parents and legislators or narrow conceptions of research rather than robust empirical investigation, and more. As a result—and rightfully so—the field of literacy is using new, complex, and evolving theories and methodologies to effectively address these big problems (Dillon & O'Brien, 2019; Duke and Mallette, 2021).

One concern noted in educational research generally and in literacy research specifically is the lack of theoretical articulation in published scholarship. For example, Parsons and colleagues (2016, 2020) analyzed 4305 articles published in 15 literacy journals across a 10-year period (2007–2016). They conducted a content analysis of topics studied, theoretical perspectives applied, research designs used, and data sources collected in these published articles. Pertinent to this chapter is the finding that a shocking 78% of articles did not explicitly present the theory or theories that informed the study. An important caveat is that the researchers used strict coding guidelines in identifying theoretical perspectives that required them to be explicitly presented. Nonetheless, this limited explicit presentation and application of theory in literacy research is concerning and serves as a call for researchers to clearly articulate and use theory in their research to ensure that it is informing research, theory, policy, and practice.

It must be stated that our concern with the explicit articulation of theories guiding literacy research work is not that we are using this feature as a proxy for being able to maintain a hierarchy of power that allows us to designate some work as “better” and other as “less than.” While we do believe that being grounded in theory is a prerequisite for research to be considered rigorous, we also believe that theory is continually at play in research work whether researchers acknowledge it or not. By identifying and explicating the language, assumptions, and delimitations—the theories—at work as a part of one’s inquiry, the serious researcher is laying bear for their audience what it is that undergirds the thought processes that have yielded the particular conclusions and insights they offer.

The truth of the matter is that each of us who conducts research within the field of literacy carries with us *some* set of working language, assumptions, and delimitations that are inextricably folded into our explorations. This truth exists whether we take time to make them legible for others or not and this truth exists whether or not the theory has a fancy or ‘official’ name to label it. In our estimation, the failure to explicitly identify these contributing factors, these framing theories, is evidence of another assumption: a research project that has not taken care to name the theoretical principles upon which it draws is making a (grossly erroneous) assumption that the foundational understandings of the world that it relies upon are normal and standard and universally accepted

and thus need not be clarified. We reject this notion and rather insist that researchers seeking to most justly impact individuals affected by their work must readily identify and describe the theories guiding their research endeavors.

This being said, educational research generally and literacy research particularly has experienced methodological tensions where historically quantitative and qualitative research has been perceived as conflicting approaches to inquiry. And these divides manifested in [Parsons et al.'s \(2016, 2020\)](#) content analysis of 15 literacy journals as well. That is, some journals published quantitative or qualitative research almost exclusively and seldom, if ever, published research from the other tradition. For this chapter, we present an analysis of the theoretical and methodological representation in publications of the Literacy Research Association (LRA), which, due to the LRA's sole focus on literacy research, we consider to be broadly representative of literacy research at large.

About LRA

LRA is a non-profit, international professional research organization with members "who share an interest in advancing literacy theory, research, and practice [it is] a community that engages in research and dialogue pertaining to literacy and related topics" ([Literacy Research Association, 2022](#)). A unique aspect particular to LRA is that it is decidedly focused on literacy research, unlike many literacy organizations that have a multi-lens focus on research, practice, and policy in literacy. As a stridently research-focused organization, LRA has been regarded as a leader in literacy research since its inception in the 1950s. Its annual conference and publications are viewed as *the* place for cutting edge research and theory development.

LRA publishes two journals: the *Journal of Literacy Research* (JLR) and *Literacy Research: Theory, Method, and Practice* (LR:TMP). JLR "is a peer-reviewed journal that has contributed to the advancement literacy and literacy education research for over 50 years" ([SAGE Journals, 2022](#)). It is a quarterly journal with a current impact factor of 2.918, demonstrating its significant contributions to the research literature. LR:TMP was formerly the *Yearbook of the Literacy Research Association*. All presenters at the LRA annual conference are invited to submit their presented research to the journal for blind peer review. Of the manuscripts that are submitted, only a portion (traditionally ~20%) are accepted for publication in LR:TMP. Both journals are highly selective and represent rigorous, timely, and significant scholarship in the field of literacy research.

This chapter reflects findings that emerged from a presidential charge from the Literacy Research Association to examine the range of epistemologies and methodologies represented within the organization over the previous 10-year time span. The authors were members of the ad hoc committee consisting of nine LRA members from diverse backgrounds and with diverse areas of interest and methodological expertise that was formed to meet this charge and were guided by the following questions:

1. To what extent are multiple methodologies and epistemologies represented across our conference and publications?
2. To what extent do our publications and conference address current problems in literacy via multiple methodologies and epistemologies simultaneously?

Although this charge and these questions prompted an investigation into recently published literacy research in LRA journals, we present them here simply to contextualize the catalyst for our undertaking. For this chapter, the authors considered the findings that emerged from the process, which aimed to capture literacy researchers' use of methodologies and epistemologies, to be contextualized within the current and rapidly changing literacy landscape. As such, this chapter takes a history of science ([Kuhn, 1962](#)) theoretical perspective to investigate how methodologies and epistemologies are driven by problems in literacy education and what that implies for the ways in which researchers approach research and the lenses that frame it. Additionally, we consider what [Reinking and Yaden \(2021\)](#) refer to as "productive theorizing" to speak to what we found as an absence of explicit theories in research.

Methods

This research used a mixed-method case study design. This investigation was mixed because it collected both quantitative and qualitative data. The bounded system ([Stake, 1995](#)) of this case study was the research methodologies and associated epistemologies and theories in LRA. This design is appropriate because mixed methods allowed us to collect a variety of data to examine research methodologies used within LRA from different perspectives and at various levels. Case study allows for a focused and in-depth exploration of phenomena within a bounded system ([Stake, 1995](#)).

Data collection and analysis

JLR content analysis (2007–2019)

A content analysis of the 2007–2019 volumes of *Journal of Literacy Research* was conducted. We included data from [Parsons et al. \(2020\)](#), which analyzed 2007–2016. In addition, we used the same procedures to gather data from three subsequent years: 2017, 2018, and 2019. The date range was selected to provide a view of literacy research trends over an expansive and recent time period. The peer-reviewed articles in each issue (a total of 209) were analyzed using a priori coding to document the following information for each article: (a) theoretical perspectives, (b) research design/methods, and (c) data sources.

The goal of this analysis was to explore the content of *JLR* over an extended recent period of time. Applying the a priori codes previously used by Parsons et al. (2016, 2020) allowed us to use and build upon the analysis of *JLR* already completed by he and his team in their study of 15 literacy journals. That is, *JLR* volumes from 2007 to 2016 were already coded, and our team applied the same coding to 2017–2019, thereby providing a broad analysis of *JLR*.

LR:TMP content analysis (2015–2019)

Similarly, a content analysis of the peer-reviewed articles from the 2015–2019 volumes of *Literacy Research: Theory, Methods, and Practice* were analyzed in each volume. The date range for this journal was conveniently selected because it represented the most recent five years of the journal at the time of analysis. A priori coding was used to code each article in that time frame (total of 80) the following: (a) theoretical perspectives, (b) research design/methods, and (c) data sources. We used the same a priori coding of *LR:TMP* articles as the previous *JLR* analysis to allow for consistency in coding journal articles to allow for comparison among LRA publications. The goal of this analysis was to examine trends over the previous five years in the research published in *LR:TMP*.

Self-reported epistemologies and methodologies in the conference submission system (2019 & 2020)

With the Research Committee's guidance, LRA leadership added questions about submitters' epistemologies and methodologies in the proposed sessions as a required component of the conference proposal submission system, asking them to choose from a menu of options or to select other and provide an unlisted category (see Appendix). The Research Committee collaboratively created the menu options to capture the epistemological, theoretical, and methodological approaches guiding research within LRA and the various contexts in which this research took place (e.g., grade level).

We opted not to use the same categories as those used in the journal article analysis because this was a new data source that allowed us to consolidate particular theories by the traditions under which they fall instead of documenting closely-related theories separately as was done in the analysis of journals. For example, in the journal analysis, which followed the procedures of Parsons et al. (2016, 2020), social practice theory and CHAT would have been coded separately. Researchers submitting proposals guided by social practice theory or CHAT to LRA would have categorized them as "sociocultural theories," which subsumed social practice theory and CHAT.

We collected these data in 2019 and 2020. The committee accumulated the totals for each of the codes to catalog the range of epistemologies and methodologies self-reported by researchers submitting proposals to the 2019 and 2020 conferences. Codes were analyzed by acceptances and rejections to see if particular epistemologies or methodologies were disproportionately accepted or rejected.

Results

JLR

Each of the 209 articles' theoretical frameworks was coded. The most frequently used were Not Specified (89), Multiple Theories (29), Sociocultural (24), and Other (18). The most common research designs were Qualitative (36), Case Study (34), Quantitative (30), and Nonempirical (22). The most frequently collected data sources were Observations (94), Interviews (90), Artifacts (77), Student Assessments (51), and Surveys/Questionnaires (32). The full results of this coding are presented in Table 1.

LR:TMP

The 80 articles' theoretical frameworks were coded. The most frequently used were Other (33), Multiple Theories (15), Not Specified (13), and Sociocultural (4). The most common research designs were Qualitative (15), Case Study/Collective Case Study (14), Other (12), Content Analysis (8), and Discourse Analysis (7). The most frequently collected data sources were Observations/Video Transcripts/Field notes (36), Interview Focus Groups (32), Artifacts (32), and Surveys/Questionnaires (13). The full results of this coding are presented in Table 2.

Self-reported epistemologies and methodologies

About 49% of the scholars who submitted proposals to the 2019 and 2020 LRA conferences reported that their epistemological stances fell within a constructivist paradigm. The next most commonly reported epistemology was critical/feminist/Marxist/post-structuralism (24%). Only 6% of respondents identified either post-positivist or positivist approaches as their epistemological stances. In terms of methodology, 26% of all submitted research used a qualitative case design. Interestingly, about 31% indicated that they used more than two research methods, which was coded as "mixed." A few epistemologies or methodologies appeared to be disproportionately accepted or rejected. Theoretical submissions were accepted less frequently (61%) than empirical proposals (73–79%, see Table 3) as were proposals that used Linguistics (46%) or Behaviorism (50%, see Table 5) as their theoretical perspectives. Tables 3–5 present the data from this analysis.

Table 1 Journal of literacy research theoretical perspectives, research designs, and data sources.

	Year of publication													n
Code	'07	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	'18	'9	
Theoretical perspective														
Cog./info. Processing	1	0	0	0	1	0	0	0	0	0	0	0	0	2
Constructivism	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Critical	0	0	0	0	0	1	0	1	0	1	5	3	2	13
CHAT	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Emergent literacy	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Goal theory	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Identity theories	0	0	1	0	0	1	0	0	0	0	0	2	0	4
Linguistic/sociolinguistic	0	0	1	0	1	0	1	0	1	0	0	0	0	4
Multiple theories	3	4	2	1	1	6	2	2	2	0	6	0	0	29
New literacies	0	1	0	0	2	0	3	0	0	0	2	3	1	12
Not specified	10	7	5	14	9	6	9	4	4	9	1	2	9	89
Reader response	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Simple view	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Situated learning	0	0	0	0	0	0	1	1	0	1	0	0	0	3
Social constructivist	1	0	0	0	0	0	0	0	0	0	0	1	0	2
Sociocultural	2	1	2	2	0	2	0	4	4	1	2	1	3	24
Transactional	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Bakhtinian theory	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Social practice theory	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Other	0	1	0	0	0	0	0	3	2	2	4	3	3	18
Total	18	14	14	18	14	16	16	15	15	16	20	15	18	209
Research design														
Case study	0	1	3	2	4	2	3	4	3	2	3	5	2	34
Content analysis	0	2	1	0	0	0	0	1	2	2	1	0	2	11
Discourse analysis	0	0	0	0	0	1	2	0	1	0	2	2	0	8
Ethnography	1	1	0	1	1	0	0	1	2	1	1	2	2	13
Experimental/quasi.	2	2	0	1	0	0	1	0	0	1	1	0	0	8
Formative	0	0	0	0	1	0	1	0	0	0	1	0	1	4
Grounded theory	1	0	0	0	0	0	0	1	0	0	0	0	0	2
Literature review	0	0	0	0	1	0	0	0	0	0	1	1	5	8
Meta-analysis	0	1	0	0	0	0	1	0	0	0	0	0	0	2
Mixed methods	1	2	0	2	2	4	0	0	0	0	0	0	0	11
Narrative	0	0	1	0	0	0	0	1	0	0	5	0	0	7
Non-empirical	3	0	3	5	0	1	4	2	0	2	1	0	1	22
Qualitative	4	4	2	0	1	2	2	2	4	7	4	4	0	36
Quantitative	6	1	3	4	2	5	0	3	2	1	0	0	3	30
Statistical modeling	0	0	1	3	2	1	2	0	1	0	0	1	2	13
Total	18	14	14	18	14	16	16	15	15	16	20	15	18	209
Data sources														
Artifacts	6	3	5	6	5	7	4	5	6	8	8	9	5	77
Interviews	4	6	5	4	7	5	9	9	7	8	8	9	9	90
Observations	5	3	6	6	9	10	7	8	10	7	7	7	9	94
Student assessment	8	4	4	9	5	6	3	2	4	2	2	1	1	51
Studies/articles	0	1	0	0	0	0	1	1	0	1	1	0	0	5
Surveys/questionnaire	3	2	2	5	6	5	0	3	2	2	2	0	0	32
Other	0	4	0	0	0	2	0	0	1	0	0	2	0	9
Total	26	23	22	30	32	35	24	28	30	28	28	28	24	358

Cog./info. processing = Cognitive/informational processing; CHAT = cultural historical activity theory; quasi. = quasi-experimental.

Discussion

Findings from this analysis of 209 articles from *JLR* (2007–2019), 80 articles from *LR:TMP* (2015–2019), and 1470 conference proposals (2019–2020) focused on (a) the extent of methodological and epistemological representation and (b) the extent to which LRA venues addressed current problems in literacy via those methodologies and epistemologies. We briefly consider underlying impetuses that ultimately inform the types of methodologies and epistemologies used in literacy research and their relevance in current social contexts. We also discuss what these findings might imply regarding future research.

Table 2 Literacy research: theory, method, and practice theoretical perspective, research design, and data sources by year.

	Year of publication					
Code	2015	2016	2017	2018	2019	n
Theoretical perspective						
Critical perspectives	0	1	0	0	1	2
Culturally relevant pedagogy	0	0	1	0	0	1
Identity theories	0	0	1	2	0	3
Linguistic or sociolinguistic theories	1	0	1	1	0	3
Multiple theories	1	4	0	4	6	15
New literacies	1	1	1	0	0	3
Not specified	4	6	0	3	1	13
Pragmatic perspectives	0	0	1	0	0	1
Sociocultural	2	1	1	0	0	4
Third space	1	0	0	0	0	1
Other	10	5	9	4	5	33
Total	20	18	15	14	13	80
Research design						
Case study/collective case study	0	4	2	4	4	14
Content analysis	4	1	1	1	1	8
Discourse analysis	2	2	0	2	1	7
Ethnography	1	0	1	0	0	2
Formative/design experiment	1	0	1	0	0	2
Grounded theory	1	0	0	1	1	3
Literature review	1	0	2	0	0	3
Mixed methods	1	1	0	0	1	3
Narrative	1	0	0	0	0	1
Non-empirical	0	1	1	3	1	6
Qualitative	3	4	3	1	4	15
Quantitative	1	1	1	1	0	4
Other	4	4	3	1	0	12
Total	20	18	15	14	13	80
Data sources						
Artifacts	9	7	3	4	9	32
Interviews/focus groups	7	7	5	5	8	32
Observations/video transcripts/field notes	10	7	6	6	7	36
Student assessments	3	3	2	1	1	10
Studies/articles	2	0	2	2	0	6
Surveys/questionnaires	3	1	5	2	2	13
Other	0	6	0	1	0	7
Total	34	31	23	21	27	136

Table 3 Frequencies of acceptance/rejection and percentage of acceptance of LRA conference proposals by research method 2019–2020.

Research method	Accepted (n)	Rejected (n)	Total	% Accepted
Formative experiments and design-based research	38	10	48	79.17
Literature review/research synthesis	67	25	92	72.83
Mixed methods	200	65	265	75.47
Qualitative	706	231	937	75.35
Quantitative	73	21	94	77.66
Theoretical (not data-based)	21	13	34	61.76

Methodologies used in LRA outlets included qualitative (e.g., social design methodologies, [Frankel et al., 2018](#); ethnography, [Buchholz, 2019](#)), quantitative (e.g., validation studies, [Gehsmann et al., 2017](#); correlational studies, [Lupo et al., 2017](#)) and mixed methods approaches (e.g., [O'Brien and Leighton, 2015](#)). Despite this documented diversity, most of the research presented at the LRA conference and published in LRA publications used qualitative methodologies and sociocultural and critical epistemologies. Though there were examples of both qualitative and quantitative research across our data set, qualitative studies outnumbered quantitative research at a ratio of roughly 10 to 1. The ratio of qualitative to mixed methods is approximately 10 to 3. The number of qualitative works far exceeded those with a quantitative design, creating an unequal distribution of methodologies represented across recent literacy research presented or published in LRA venues.

Table 4 Frequencies of acceptance/rejection and percentage of acceptance of LRA conference proposals by epistemology 2019–2020.

<i>Epistemology</i>	<i>Accepted (n)</i>	<i>Rejected (n)</i>	<i>Total</i>	<i>% Accepted</i>
Single category				
Constructivist	507	184	691	73.37
Critical/feminist/Marxist	162	63	225	72.00
Positivist	31	13	44	70.45
Post positivist	30	12	42	71.43
Post structural	56	17	73	76.71
Multiple categories				
Constructivist, critical/feminist/Marxist	137	37	174	78.74
Constructivist, post positivist	39	6	45	86.67
Constructivist, post structural	22	6	28	78.57
Critical/feminist/Marxist, post structural	45	12	57	78.95
Constructivist, critical/feminist/Marxist, post structural	14	4	18	77.78
Multiple	62	11	73	84.93

Combinations of epistemologies cited by 10 or more proposals were examined as their own categories. Combinations cited by fewer than 10 proposals were collapsed into the code "Multiple."

Table 5 Frequencies of acceptance/rejection and percentage of acceptance of LRA conference proposals by theoretical perspective 2019–2020.

<i>Theoretical perspective</i>	<i>Accepted (n)</i>	<i>Rejected (n)</i>	<i>Total</i>	<i>% Accepted</i>
Single category				
Behaviorism	2	2	4	50.00
Cognitive psychology	29	9	38	76.32
Critical theories	78	28	106	73.58
Linguistics	6	7	13	46.15
Psycholinguistics	2	1	3	66.67
Semiotics, multimodalities, multimedia	23	10	33	69.70
Sociocognitive theories	32	12	44	72.73
Sociocultural theories	219	73	292	75.00
Sociolinguistics	8	3	11	72.73
Other	43	16	59	72.88
Multiple categories				
Critical theories/sociocultural theories	147	44	191	76.96
Critical theories/sociocultural theories/sociolinguistics	21	2	23	91.30
Critical theories/semiotics, multimodalities, multimedia/ sociocultural theories	41	7	48	85.42
Semiotics, multimodalities, multimedia/sociocultural theories	36	15	51	70.59
Semiotics, multimodalities, multimedia/sociocultural theories/sociolinguistics	8	3	11	72.73
Sociocognitive theories/sociocultural theories	37	13	50	74.00
Sociocultural theories/sociolinguistics	35	7	42	83.33
Multiple	338	113	451	74.94

Combinations of theoretical perspectives cited by 10 or more proposals were examined as their own categories. Combinations cited by fewer than 10 proposals were collapsed into the code "Multiple."

How should we interpret this methodological imbalance? As methods are driven by the research questions being asked, we suspect that some of the methodological trends highlighted by our analysis is a reflection of the pressing questions literacy scholars are posing and investigating as called for by current social contexts. It is not that quantitative methods have fallen out of favor or have been relegated by the literacy research community to an undesirable status. Rather, questions attempting to address the needs of our ever-changing world as related to societal power structures brought to the fore in particular by the COVID-19 global pandemic that onset in early 2020 along with the racial reckoning experienced in the US and across the globe in the summer of 2020 may be better explored via qualitative methods.

We do not see trends that indicate a collective movement toward a particular worldview (either domestically or internationally) as a field, and we see that as a positive outcome. Indeed, we encourage fellow literacy researchers to be open-minded about the potentialities of a range of methodologies and to continue to use methods that are theoretically grounded in perspectives that support rigorous and robust responses to timely and important questions about the world of literacy in its many forms. In particular, we encourage the application of new or combined methodologies and epistemologies that actively disrupt methodological

and epistemological binaries (e.g., [Dillon and O'Brien, 2019](#); [Purcell-Gates et al., 2016](#)), approaches we believe make space for new considerations poised to yield new insights.

We saw evidence of literacy researchers using multiple epistemological stances. For example, when considering the self-reported data from proposal submissions, it was common for researchers to report the employment of a variety of epistemologies based on the focus of a particular study. These observations lead us to ask additional questions about how theory is woven into literacy research work: How are researchers articulating their theoretically driven understandings of the literacy world, literacy research, and methods for forming continued understandings about literacy, and how are those articulations related to the ways they are (or are not) applying theory to their research efforts? These questions become particularly important when we use the collected data in an attempt to draw conclusions about the current field of literacy research and about where the field is headed. It also signals to us that as a field we must remain diligently nimble as we continue to be responsive to the literacy needs of the world, regardless of whether this will ultimately result in a singular or dominant orientation to literacy research.

The field of literacy research is tightly connected to the social and structural factors that provide the context within which the research occurs. We posit that expanding methodologies and epistemologies is not, then, an indication of its immaturity or underdevelopment ([Kuhn, 1962](#)), but rather its ability to flexibly contend with the ever-changing context. In relationship to current social movements that are encouraging more justice-oriented and equitably-driven understandings of the world at large, it is necessary that the focal points, methods, and theories underpinning literacy research are able to evolve in ways that allow it to address the most pressing needs within the world of literacy more specifically.

At the same time, we also acknowledge the pause that an inconsistent explicit application of theory in literacy research causes us, which others have also noted ([Anders et al., 2016](#)). In our analysis of *JLR*, for example, we found that 43% of the 209 articles did not explicitly articulate a theoretical framework. In *LR:TMP* 16% of articles did not clearly state the theoretical grounding. These findings are far better than what [Parsons et al. \(2020\)](#) found across 15 journals where 78% of articles did not explicitly name a theoretical perspective. Nonetheless, we find this omission of or unclear presentation of a theoretical framework problematic. As previously mentioned, at the core of the problem is the fact that working theories are framing any and all research efforts whether they are explicitly articulated or not. The candid presentation of the theories contributing to one's research work requires the researcher to seriously grapple with the assumptions about the world, themselves, learners, literacy and more that they are—subconsciously or intentionally—folding into their inquiry explorations. As the field of literacy research continues to advance, we call for literacy researchers to dedicate concentrated efforts to both consider and articulate theoretical foundations that drive their understandings and thus their methods for gathering data and the conclusions supported by the data collected.

We particularly concur with [Reinking and Yaden \(2021\)](#) that theorizing, particularly in literacy education, may be critical to productivity in research. They note, "For example, theories *about* [emphasis in original text] literacy can reveal hidden or ignored dimensions of literacy and established valued goals or a guiding moral conscience *for* [emphasis in original text] literacy research" (p. 385). As such, identifying theories about broader literacy topics may, in turn, forge stronger connections to classroom or field-based literacy practice. Again, our operational assumption is that, even if not explicitly stated or articulated, theories are certainly and persistently already in place. Given the diversity we have seen among methodologies and epistemologies utilized in the recent literacy research data analyzed here, we assume theories exist with a level of multiplicity that will continue to fuel a research field poised to provide significant and timely observations. But if we as a field are to be able to effectively address many of the big problems we outlined at the onset of this chapter, we must take on the task of thinking seriously enough about the theoretical assumptions we are embedding into our attempts to address those problems that we become adept at describing them with clarity and directness. We encourage explicitness in theory identification to draw stronger connections across research that may support stronger practice.

Conclusion

The current field of literacy—the problems it faces and the contexts in which it faces them—is incredibly complex. At the same time, the field of literacy is growing and expanding ([Beach & O'Brien, 2018](#); [Moje et al., 2020](#)). What literacy is and what it means to be literate is evolving more rapidly than ever before. Accordingly, researchers need multiple, diverse, integrated, and innovative theories and methodologies to study and understand these complexities ([Dillon & O'Brien, 2019](#); [Duke and Mallette, 2021](#)). Therefore, we believe that researchers continually learning and growing in their epistemological, theoretical, and methodological understandings will only serve to better the effort. Such respect for and use of diverse epistemologies and methodologies and continual development as scholars will serve to unite the field of literacy research, not divide it.

Appendix: Options for self-reported epistemologies and methodologies for LRA conference submission

Epistemology

1. Positivist
2. Post Positivist
3. Constructivist/Social Constructivist
4. Critical/Feminist/Marxist
5. Post Structural

Theoretical perspectives derived from/rooted in...

1. Behaviorism
2. Linguistics
3. Psycholinguistics
4. Cognitive theories
5. Sociocognitive theories
6. Sociocultural theories
7. Semiotic, Multimodal, Multimedia theories
8. Critical theories
9. Sociolinguistics
10. New Literacies
11. Others:

Method

- a. Qualitative
- b. Ethnography
- c. Case Study
- d. Phenomenology
- e. Narrative Inquiry
- f. Content Analysis
- g. Discourse Analysis
- h. Grounded Theory
- i. Other: (specify)
- j. Quantitative
- k. Descriptive
- l. Correlational
- m. Predictive
- n. Quasi-Experimental/Experimental
- o. Survey
- p. Statistical Modeling
- q. Meta-Analysis
- r. Pre-Existing Statistical Data Using Computational Techniques
- s. Other: (specify)
- t. Mixed Methods
- u. Design-Based Research
- v. Practitioner Inquiry/Action Research/Participatory Action Research
- w. Literature Review/Research Synthesis
- x. Theoretical (not data-based)

Setting

- Early Childhood
- Elementary
- Middle School
- High School
- College
- Post-baccalaureate
- Adult
- In schools
- Out of schools
- Other: (specify)

Acknowledgments

We acknowledge the other members of the original committee who were unable to participate in this chapter: Julie Coiro, Marla Mallette, Gwen McMillon, and Alfred Tatum. We acknowledge and appreciate the data collection contributions of the following individuals: Melissa Gallagher,

Sohee Park, Samantha Ives, Deidre Cutter, Matthew Korona, Reagan Murnan, Lilly LePelch, Madelyn Wells, Sara Field, Kelly Usher, Andrew Vardas-Doane, Michelle Arnold, Lauren Loughlin, Julie Sill, Robyn Crump, and Tisha Jones.

References

- Anders, P.L., Yaden, D.B., Iddings, A.C., Katz, L., Rogers, T., 2016. Improving theory in literacy research. *J. Lit. Res.* 48 (1), 3–4. <https://doi.org/10.1177/1086296X16637169>.
- Beach, R., O'Brien, D., 2018. Significant literacy research informing the English language arts. In: Lapp, D., Fisher, D. (Eds.), *Handbook of Research on Teaching the English Language Arts*, fourth ed. Routledge, pp. 1–29.
- Buchholz, B.A., 2019. Author(iz)ing a playground game: "The arguing started once the rules were written down". *J. Lit. Res.* 51 (1), 100–122. <https://doi.org/10.1177/1086296X18821144>.
- Dillon, D.R., O'Brien, D.G., 2019. Pragmatism [not just] practicality as a theoretical framework in literacy research. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*. Routledge, pp. 582–600.
- Duke, N.K., Mallette, M.H., 2021. Conclusion. In: Mallette, M.H., Duke, N.K. (Eds.), *Literacy Research Methodologies*, third ed. Guilford, pp. 447–453.
- Frankel, K.K., Fields, S.S., Kimball-Veeder, J., Murphy, C.R., 2018. Positioning adolescents in literacy teaching and learning. *J. Lit. Res.* 50 (4), 446–477. <https://www.doi.org/10.1177/1086296X18802441>.
- Gehsmann, K., Spichtig, A., Tousley, E., 2017. Validating an online assessment of developmental spelling in grades five through eight. *Lit. Res.* 66, 279–292. <https://www.doi.org/10.1177/2381336917718834>.
- Kuhn, T., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press.
- Literacy Research Association, 2022. About LRA. <https://literacyresearchassociation.org/about/>. Accessed March 25, 2022.
- Lupo, S., Jang, B.G., McKenna, M., 2017. The relationship between reading achievement and attitudes toward print and digital texts in adolescent readers. *Lit. Res.* 66, 264–278. <https://www.doi.org/10.1177/2381336917719254>.
- Moje, E.B., Afflerbach, P.P., Enciso, P., Lesaux, N.K., 2020. Game changers in reading research. In: Moje, E.B., Enciso, P., Afflerbach, P.P., Lesaux, N.K. (Eds.), *Handbook of Reading Research*, vol. V. Routledge, pp. 3–13.
- O'Brien, L.M., Leighton, C.M., 2015. Use of increasingly complex text to advance EL's knowledge and academic language. *Lit. Res.* 64, 169–192. <https://doi.org/10.1177/2381336915617579>.
- Parsons, S.A., Gallagher, M.A., the George Mason University Content Analysis Team, 2016. A content analysis of nine literacy journals, 2009–2014. *J. Lit. Res.* 48 (4), 476–502. <https://doi.org/10.1177/1086296X16680053>.
- Parsons, S.A., Gallagher, M.A., Leggett, A.B., Ives, S.T., Lague, M., 2020. An analysis of 15 journals' literacy content, 2007–2016. *J. Lit. Res.* 52 (3), 341–367. <https://doi.org/10.1177/1086296X20939551>.
- Purcell-Gates, V., Duke, N., Stouffer, J., 2016. Teaching literacy: reading. In: Gitomer, D.H., Bell, C.A. (Eds.), *Handbook of Research on Teaching*, fifth ed. AERA, pp. 1217–1267.
- Reinking, D., Yaden Jr., D.B., 2021. Do we need more productive theorizing? A commentary. *Read. Res. Q.* 56 (3), 383–399. <https://doi.org/10.1002/rrq.318>.
- SAGE Journals, 2022. *Journal of Literacy Research*. <https://journals.sagepub.com/home/jlr>. Accessed March, 25 2022.
- Stake, R., 1995. *The Art of Case Study Research*. SAGE.

Affect theory and textual variations

Kimberly Lenters, University of Calgary, Calgary, AB, Canada

© 2023 Elsevier Ltd. All rights reserved.

Assemblage perspectives on affect	911
Affect theory	912
Expanding conceptions of affect theory	912
Affect as orienting and disorienting	913
Affect as transformative and normative	913
Recognizing affect and textual variations	914
Pikachu bonding	914
Book club instigations	914
When words are not enough	915
Openings	916
References	916

"Pikachu! Pikachu!" Crystal's voice could occasionally be heard calling out during one of my recent classroom research projects. While most of the other children participated in whole class discussions in the usual schooled way—raising hands and waiting to be called upon before responding—Crystal would occasionally jump into the flow, unbidden, squeakily inserting Pikachu from the sidelines. In these moments, Crystal was, as some would say, off-task.

What might have happened had this Pikachu-calling been recognized instead as an affective engagement with literacy with the potential to encourage and nurture Crystal's practice of literacy? In this entry, I explore this and two other moments of classroom literacy engagements that, when placed in conversation with post-structural (Deleuze and Guattari, 2004) and post-colonial (Wehliye, 2014) understandings of affect, offer a means for reconfiguring our understanding of young people's engagements with literacy.

Recognizing the importance of behavioral aspects of literacy, educational psychology has approached classroom literacy learning with constructs such as motivation and engagement for some time now (e.g., Guthrie and Klauda, 2014; Wigfield and Guthrie, 1996). Primarily interested in the relationship between reading comprehension, motivation, and engagement, the field accounts for cognitive aspects of literacy learning. However, based on a normative idea of what constitutes engagement with literacy, identifying singular measures of engagement, a cognitive-only approach offers limited scope for examining the multiple ways participation in literacy might occur. Sociocultural literacy scholars have also recognized the importance of motivation and engagement, shifting attention from the "reader-centric" focus of cognitive perspectives to an "activity-centric" approach to literacy (Pearson and Cervetti, 2015). While sociocultural literacy studies encompass a range of perspectives (e.g., Smagorinsky and O'Donnell-Allen, 1998; Snow and Verhoeven, 2001), Hruby et al. (2016) observe that sociocultural approaches to motivation and engagement commonly understand the "cause" and "effect" of engaged literacy to be located externally, situated in both historical and cultural contexts (p. 616). While both of these perspectives have provided important insights, neither of these approaches to engagement (cognitive internally focused or sociocultural externally focused) considers the role of the body and relationship—the individual as a relational member of the body collective, such as classrooms, communities, or families—in young people's engagement with literacy. The concept of affect addresses this matter, providing literacy educators and researchers with multiple avenues for recognizing and supporting young people's engagements with the textual production—a label I am using to refer to the multimodal texts (print, oral, visual, and embodied) with which young people engage.

Assemblage perspectives on affect

Affect is relational. It is aroused or sparked when bodies come into association with each other and thus affect theory needs to be understood within the larger theoretical realm of relational theories, such as affect theory, which commonly draw on the work of Gilles Deleuze and Felix Guattari. Assemblage perspectives view living beings as moving through the world as participants in heterogeneous groupings of human and more-than-human entities, otherwise referred to as assemblages (Deleuze and Guattari, 2004; DeLanda, 2006). Assemblages are constantly emerging, shifting, joining together, and breaking apart, as relations change, and participants enter and exit. They are non-hierarchical: that is, participation of humans is not automatically foregrounded in the functioning capacities of an assemblage. Thus, when an assemblage is mapped, people may be recognized as significant participants, but the participation of things such as signs, material objects, events, practices, and utterances must also be recognized.

Relations between the myriad more-than-human and human participants constitute the most important aspects of an assemblage's functioning. These relations are productive: they do things. While material things are the most easily recognizable of these productions, relations between participants also produce the immaterial, and affect is one of the most prominent of these immaterial productions. Thus, in a classroom literacy assemblage, we might easily recognize the inquiry projects, stories, notes made in

notebooks and on screens, text messages, or doodles produced by students on a page or whiteboard; however, attention to the immaterial requires us to notice the affect that is produced through students' relations with each other, their teachers, material items, and any other matter present or produced in the classroom assemblage.

Affect theory

Along with many theoretical bodies informed by assemblage perspectives, affect theory is grounded in work of 16th century Dutch philosopher, Baruch Spinoza, who is widely recognized within Western philosophy for disputing Descartes' mind-body dualism. Spinoza's writings emphasize bodily experience as an important aspect of knowing, inviting educators to consider ways people come to know, not only as epistemology (ways of knowing) but also as ontology (ways of being and doing). Spinoza famously identified affect as "the ability to affect and to be affected," highlighting affect's bi-directional force and flow but not necessarily defining it.

Over the last three decades, scholars in areas such as cultural studies, anthropology and sociology have begun to explore notions of affect, moving us toward a more precise and expansive conceptualization. [Massumi \(2004\)](#) identifies affect as "a prepersonal intensity corresponding to the passage from one experiential state of the body to another and implying an augmentation or diminution in that body's capacity to act" (p. xvi). [Gregg and Seigworth \(2010\)](#) add clarity to this notion of affect as an intensity, naming it as "those visceral forces beneath, alongside, or generally *other than* conscious knowing that can serve to drive us toward movement, thought, and ever-changing forms of relation" (p. 1). Following these scholars, I understand affect to be a form of embodied knowing, a visceral intensity that informs our immediate, practical actions and in the case of literacy, that which initiates the process of understanding and responding ([Lenters and McDermott, 2020](#)). Affect has also been described as a hunch or a visceral prompt ([Hickey-Moody, 2013](#)), an enchantment ([Bennett, 2001](#)), a feeling ([Leander and Ehret, 2019](#)), a vibration, rush, quiver, or surge ([Dutro, 2019](#)), a resonance ([Nxumalo and Villanueva, 2020](#)), and a stickiness ([Franklin-Phipps, 2020](#))—descriptive nouns that vividly animate and amplify the concept.

Affect is considered by many as distinct from emotion. While emotion is recognized as part of the meaning-making process, affect is considered to exceed consciousness ([Clough, 2007](#)), it is prepersonal ([Massumi, 2004](#)). Affect "precedes the articulations of language and thought. It is felt in the body before it is overlaid with meaning" ([Cannon, 2009](#), p. 74). In other words, affect comes before emotion, possibly only by seconds, but it is understood to be that which registers in the body before the mind processes it to form emotion. Although affect and cognition are understood, in practice, never to be fully separable ([Gregg and Seigworth, 2010](#)), this distinction between affect and emotion is useful for underscoring affect's role in social configurations, a role that can defy rational expectations for what might unfold as a body moves through its world.

Expanding conceptions of affect theory

The assemblages in which affect is produced are anything but straightforward. While much has been made of the idea that assemblages are comprised of human and more-than-human participants, critical race scholars remind us that thinking of "the human" as one generic category, belies the fact that in our social world, the experience of being human is not universal ([Weheliye, 2008](#); [Wynter, 2003](#); [Spillers, 1987](#)). This has implications for how affect is understood—in particular, the observation that not all affects are treated equally. Some people's affects count, and some do not.

Long-standing, institutionalized practices of racism, oppression and exclusion produce gendered and racialized violence that form the status quo in so many assemblages, making it difficult, if not impossible, for many to participate fully. [Weheliye \(2014\)](#) introduces the concept of *racializing assemblages* to draw to our attention these omnipresent systems. Weheliye (and others) argue that responding in racializing assemblages with confrontation only gives credence to that oppression, thereby further entrenching it ([Jocson and Dixon-Román, 2020](#); [Burton, 2017](#); [Weheliye, 2014](#)). Informed by critical race and queer theories, they advocate for ways of being and doing in assemblages that acknowledge affect "through fleshy humanness [as] a departure that survives and thrives in ... unorthodox ways" ([Burton, 2017](#), p. 35).

Classrooms and community organizations, all racializing assemblages, must allow "genres of being human" ([Wynter, 2003](#)) or in other words permit affective responses that sometimes push back on normative conceptions of what it means to "behave well" in a classroom. As [Snaza and Sandlin \(2017, p. 118\)](#) state, "classroom spaces often preserve the whiteness of educational institutions through the modulation and sanctioning of particular feelings." For instance, affective expressions of rage, despair, or defiance through spoken, embodied, and multimodal texts, particularly when associated with race or gender, are too frequently unwelcome in classrooms.

The concept of racializing assemblages thus underscores the notion that what a body does—its capacity to affect and be affected—is contoured by its relations with other bodies ([Deleuze and Guattari, 2004](#)) and what is typically considered acceptable behavior for those bodies ([Weheliye, 2014](#)). When applied to literacy learning, the concept of affect prompts educators to consider the "other than conscious" forces that work to enable, support, motivate and inspire literate engagements. Simultaneously, educators must consider what constrains the practice of literacy for some students, asking what is orienting and disorienting, transformative and normative, both in the classroom and elsewhere. In the next section, I explore some of these facets of affect in relation to literacy applications.

Affect as orienting and disorienting

Whether in school or in contexts other than school, young people make choices about the kinds of texts with which they will engage. These choices are not always rational or conscious (Jocsun and Dixon-Román, 2020; Lenters, 2016, 2018; Leander and Boldt, 2013). Something about a particular cartoon, film, comic, graphic novel, or picture book grabs their attention and holds it: perhaps the book appeals to their sense of humor; maybe it is aesthetically pleasing; possibly the story strikes a chord; or perhaps it is wildly popular with their friends and connects them to those friends in unspoken ways. Some texts invite a response in ways that seem highly organic, such as a text message or social media post in response to something posted by another, where the young person appears to barely pause before responding. The affect associated with such responses recognizable as response-driven by visceral, other than conscious knowing. In this way, affect orients young people toward literate engagement.

Affect may also be disorienting and can work in a couple of different ways. While affective feelings are so often what attract young people to a text, those feelings may just as easily be that which repels them. Even though they may not be able to name or identify, with words, exactly what is taking place, young people may form particular dislikes for certain types of texts. Thinking with affect about these disorientations directs educators to consider the assemblage in which distaste forms. Perhaps the student has formed feelings for a particular type of text (a literacy object) that relate to, for example, impossible classroom expectations around using that text, someone who was associated with that text at a particular moment, or how someone responded to a text the student produced. Often unspoken, feelings such as these can be intense and repel students from texts for reasons that have nothing to do with the text itself.

Affective disorientation can also have an upside. When texts unsettle, provoke or move students in certain ways, enlivening their senses with regard to a topic, an issue, a person, or a place, they are disoriented. As they are moved to think in new ways about things that might have been commonplace or stereotyped in their worlds, affective disorientation can open new possibilities and trajectories for them.

Affect as transformative and normative

The positive aspects of affective orientations and disorientations may prompt a tendency to focus on affect's transformative possibilities. While it is a key aspect of social and material literacy learning, a sole focus on transformation engenders a myopia that constrains attention to affect's potential in educational settings. Hemmings (2005), Dernikos et al. (2020, p. 9) state, "affect has the potential to be both transformative and normative." Hemmings contention is that "affective freedom," that is the ability for affect to radically alter an individual's subjectivity, ignores the fact that structural constraints exist and exert their influence on individuals and particular groups, and cannot be ignored in the pursuit of a research stance that is productive (as opposed to critical). When this is acknowledged, a glimpse of affect as both transformative and normative comes into view.

The call to move beyond the trap of critique for the sake of critique (Lather, 2016) that marks some applications of post-structural theory in literacy studies is valid and important. The productive, forward looking stance affect offers to those of us looking for generative practices in children's literacy education is invaluable, for it reminds us that following, for example, children's affective enchantment with particular literacy objects and practices provides important means for encouraging the transformative possibilities of alternate paths in children's literacy development. However, the cautionary note Hemmings provides is crucial in at least two ways: affect manifests in normative ways in the literacy classroom both through the educators' affective responses to children's textual productions and in the setting out of appropriate affective responses for children engaging with texts.

In classroom composing assemblages, a particular narrative around what are acceptable textual productions seems to pervade. Children who talk, sing, act, write, or draw about anything that either attends to or appears to promote violence (guns, fighting, war, anger) are often told these are inappropriate topics (e.g., Dyson, 2003). The child with productions of this sort might even be referred to a school psychologist. Not discounting the concerns that certain types of textual production engender for educators, these "inappropriate" topics are part of the fabric of many students' lives. Disallowing such productions or responding to them in particular ways can work to both disorient the student from literate engagements or worse, stigmatize the child within the classroom assemblage. Dutro (2019) provides an illustration of this as she relates the story of second-grade student, Malina. Malina's father, who had been incarcerated for several years, was a central figure in much of Malina's writing. Dutro writes,

In responding to her writing about her father, potential exists for fueling empathy, relationship, advocacy, and political commitment toward justice. But there is also the potential for pity, antipathy, fueling of deficit narratives, and political commitment toward injustice. White, heterosexual, gender normative, economically advantaged educators, like me, who benefit from privilege must be particularly vigilant against the latter response (pp. 386–387).

Educators' affective responses—pity, distaste, alarm—can function in normative ways to dictate what is appropriate, according to an unspoken ontology of privilege. Attuning to that affect is key to disrupting the classroom's racializing assemblage (Weheliye, 2014). Dutro (2019, p. 286) identifies this noticing as "a pause between sensation and language that allows us to question before we leap to certainty." By attending to and questioning the affect that precedes verbal and embodied responses to children's textual productions, educators can work toward building classroom assemblages that are transformative for all member children and not just those with whom we readily identify.

Normativity in educator affect may also be experienced by children in relation to their affective responses in literacy learning contexts. In recent years, classrooms have become more attentive to the critical need for providing literature in which all students

can see themselves; that is, diverse and culturally and linguistically responsive literature. However, particular notions of how literate engagement should be performed (e.g., [Luke, 1992](#)) can mean that diverse responses to literature are not always welcome. [Dernikos \(2020\)](#) identifies this force as *affective assimilation*. Educators may be more inclined to label a student as an engaged reader or writer when they are laser-focused on the text at hand. However, those whose impulse is to verbally or demonstratively share what excites or enrages them, as they engage with text, are often subjected to affective assimilation; that is, they are likely to receive a censorial shushing from their teacher or be labeled as disruptive. Again, by attuning to and validating such diverse or “fleshy” ([Burton, 2017](#); [Weheliye, 2014](#)) affective responses, educators can go a long way toward ensuring that affect becomes an expansive, rather than limiting, force in the classroom.

Affect should matter greatly to literacy educators—not just in effect-laden moments but also over the long-haul of a student’s educational trajectory. Just as attending to the transformative possibilities that registering affect can hold for young people’s literacy learning, research and pedagogy must be cognizant of the affective accretions that form over time. Following [Hemmings \(2005\)](#) discusses affective cycles of body-affect-emotion and argues that these cycles do not constitute singular events but rather “series of repeated moments” that form “an ongoing, incrementally altering chain—body—affect—emotion—affect—body—doubling back upon the body and influencing the individual’s capacity to act in the world” (p. 564). Put simply: affect sticks around. As [Thiel \(2015, p. 40\)](#) states, the stickiness of affect is “what gets named, remembered, embodied, and performed, serving as a catalyst in the ways we learn to become and move through the world.” For some students affect will increase their capacities for literacy learning; and for others, it will diminish it. Seen in this way, attending to affect in educational contexts is a matter of critical importance.

Recognizing affect and textual variations

In this section, I revisit encounters with three students previously introduced in other publications (I cite the publications below for those interested in a more fulsome background on these young people and their classroom and community contexts.) Affect played a key role in each of these revisited encounters in Canadian classrooms. Crystal, Ally, and Sara, in different ways, were students who appeared to be out of step, to varying degrees, with what was expected of them in the particular encounters with literacy I relate below. However, when their participation in their classroom assemblages is examined with post-colonial and post-structural concepts of affect in mind, alternative understandings of the three girls and their literacy learning become possible.

Pikachu bonding

I first return to the story that opens this article. While we rarely discussed Crystal’s ways of being in the classroom or her textual productions, her teachers did note that they interpreted things such as the Pikachu calling as maladaptive, attention-seeking behaviors that were best left unacknowledged. However, through thinking with affect, an alternative understanding becomes possible.

Crystal, a 10-year old girl, of Métis heritage, was a participant in a 2-week long series of half-day workshops with a local comedian, Ben Canon ([Lenters and Whitford, 2018](#)). In the comedy literacy workshops, as the students called them, Ben worked with Crystal’s class of 5th-grade students to immerse them in processes of composing comedy. Early on in those two weeks, Crystal, who at times struggled in her relationships with classmates and teachers, learned that Ben was fond of Pikachu. Pikachu also happened to be her favorite Pokémon character. From that time forward, she frequently sought out Pikachu conversation with Ben, sometimes interrupting discussions he was having with other students. She began bringing her stuffed Pikachu to every workshop and soon started wearing a Pikachu costume to school. Crystal clearly formed a bond with Ben, sparked by the electrifying character of Pikachu.

At some point, mid-way through the two-week workshop, Crystal began seating herself during whole-class discussions beside the whiteboard that sat on its easel to the left of the teacher’s chair. The whiteboard was primarily covered by a large pad of paper but Crystal found a small open corner and would quietly illustrate the many class stories that were collaboratively composed during the workshop. At times she selected the character of Pikachu for these illustrations and other times the figures more closely resembled those taking shape in the large group’s emerging collaborative story productions.

Not wanting to reinforce what they identified as inappropriate classroom behavior, Crystal’s textual productions—the Pikachu calling, interruptions to conversations, the stuffed Pikachu and costume, and the white-board illustrations—were not attended to by her teachers. In this way, it seemed to me that they were considered “surplus” ([Leander and Ehret, 2019](#)) (albeit, unconsciously) to the everyday business of working together well as a class. Crystal struggled with the other aspects of the workshop and I found myself wondering if the affective assimilation ([Dernikos, 2020](#)) that rendered her productions surplus might have contributed to these struggles. Crystal’s affective bonds to Pikachu and Ben, and the texts, verbal and visual, that affect produced might instead have been recognized, welcomed into the literacy circle, and could have provided openings for Crystal’s textual productions to extend into other modalities, such as writing.

Book club instigations

Ten-year old Ally had a complicated relationship to reading. She professed a love for reading when it involved certain kinds of texts, for example, graphic novels; and to roundly dislike it, in relation to the reading of other types of novels ([Lenters, 2014](#)). This meant that in school, she loved free reading time; however, participation in her classroom literature circle assemblage was fraught. In part,

this was connected to her school library's novel sets—Ally's affective disorientation to the somewhat limited offerings put her off—and the normative constraints of scripted roles in student-led literature circles. But it also had to do with who was in her literature circle group. A rivalry between Ally and Deanne, years in the making, formed a relational undercurrent of affect in Ally's experience of her lit circle group. Deanne appeared to be unaware of it, but the rivalry was felt deeply by Ally, who carried a sense of being overshadowed by Deanne's accomplishments. These aspects of her participation in literature circles created an intensity around student-led, collective reading experiences that marginalized Ally in the classroom literacy club (Smith, 1988) and led to what her teachers (and report cards) characterized as poor performance in reading.

When the opportunity to form the English Roses book club emerged in her housing co-op's shared outdoor space, things changed quite dramatically, albeit momentarily, for Ally (Lenters, 2018). The arrival of her long-awaited copy of Madonna's *English Roses: Friends for Life* (Madonna, 2007), her friends' exuberant interest in the book, the formation of the English Roses book club, and the embodied ways in which they collectively engaged with the book offered an attractive alternative to the aspects of literature clubs that Ally found so disorienting. With just one copy of the English Roses, Alli and four friends found the enthusiastic engagement with literature that her classroom literature circles (modeled on Daniels, 2002) had been unable to generate. In contrast to the disillusionment she experienced in the classroom lit circle, the book club brought meaningful discussions, electrifying embodied (re)readings, and a joyful sense of belonging, as allied Madonna lovers gathered. This palpable affect carried over into the classroom for several days, as I witnessed Ally return from recess and lunch breaks, book in hand and book club members in tow, with her head held high. This transformation stood in stark contrast to her downcast demeanor days and weeks prior, comportment that seemed to return when the English Roses book club came apart.

In the time I knew Ally (the whole of her 5th-grade year), the normative practices and the teacher-assigned group members that configured the literature circles in Ally's classroom had become highly constraining for her. In contrast, the spontaneous formation of her out-of-school book club, while precarious, were transformative and defied the constraints that threatened her literate development. Affect (re)configured Ally and her relationship to collective reading experiences. Simultaneously, even as Ally was transformed, so too was the assemblage we recognize as book club. The transformed, affectively shaped book club became a place of shared reading and embodied response. It fostered intensities that extended beyond the club meetings, affecting not only relationships to reading and texts but also interpersonal relations. As they came into a relationship with each other, the affect sparked between Alli and her book club had a mutually reconstituting effect on both assemblages. What might have happened had this intensity and the (re)configurations been noticed in Ally's classroom? Could the literature circle, as a textual production in and of itself, have thereby been transformed for all of the members of Ally's classroom assemblage?

When words are not enough

Eleven-year old Sara considered herself shy—reluctant to speak out in her 6th-grade class. As a child of South Asian descent, Sara was by no means in the minority in her culturally and linguistically diverse class (see Lenters and Whitford, *in press*). Yet, in the scene I describe next, Sara's experiences in a particular classroom improv scene animate the workings of a racializing assemblage (Weheliye, 2014) and the fleshy, affective responses to it that served to disrupt its normativizing flow. Her teacher's experiment with classroom improv had started in September with the idea that other topics would be pursued after a couple of months. However, the 30 members of the class, some of whom were familiar with improv and some not, quickly came to adore the 2 hours per week spent doing improv. Sara, like many in the class, professed to having an initial preference for being an audience member and leaving the improv performance to others. However, after a few weeks of observation she found herself compelled to take the stage.

In Sara's class, improv operated around two central rules: "Yes, and" and always make your partner look good (see Fey, 2011), rules that the class interpreted as, "No denial" and "No shining." In other words, the 6th-grade players were to endeavor to take up and build on offers (ideas, directions, suggestions, etc.) made by other players, as much as possible, and to work collaboratively to make each other look good by not personally dominating the scene. These rules and the supportive milieu they fostered most of the time, along with the teacher's gentle encouragement, soon drew even the most reluctant to Sara's classroom stage.

The classroom rules for improv were meant to guide the interactions between players on the stage, that is their embodied textual productions. However, in one particular scene, which we titled, Making Macaroni, it was affect that guided the textual production. Sara, Colton, and Liam were to improvise a scene in which they were trapped inside a computer. From the start, and throughout the scene, the boys (Colton, of Anglo-Saxon heritage, and Liam, of Asian heritage) excluded Sara (of South Asian heritage) by only speaking with each other. Sara's rather clever attempt to break into the scene by becoming Siri, the Apple IOS virtual assistant, was at first ignored by the boys. When the teacher intervened, Sara repeated her offer: "[using Siri's voice] How may I assist you?" and Colton responded with, "I'm going to disable you, Siri. You're so annoying!" Colton's retort brought uproarious laughter from the audience. As Colton continued with his "shining," Sara disrupted the scene with her own shining, moving between the boys to get to the side of the stage, saying, "I'm going to make macaroni." Seemingly unfazed, the boys continued their excluding dialogue until Sara returned, bearing a large, imaginary bowl of macaroni and dumped it in Colton's face with a peremptory, "Here, have some macaroni and cheese!", this time inviting a collective gasp, followed by silence from the audience. This type of exchange was repeated a few times during the rest of the scene.

The affects generated in this scene were multiple: exclusion, courage, frustration, laughter, shock, and perhaps embarrassment. I argue that these affects then fostered various types of textual (verbal and embodied) productions—for example, Sara-Siri cutting off the boys' excluding dialogue with movement and the introduction of offers not connected to their attempted repartee. Following

the initial Siri gambit, Sara-Siri's textual productions, as they matched the boys' exclusionary productions, technically broke with the classroom rules of improv, over and over again.

The boys' break from the improv rules drew laughter from the audience; Sara-Siri's, dismay. The scene serves as an illustration of the way affect can be used by some for normative, reinforcement of racializing assemblages in which some bodies (Sara and Sara-Siri) were excluded by a gendered and racialized status quo perpetuated by those accustomed to holding the stage (Colton and Liam). It also demonstrates how affect may be used to turn the tables on textual production. Exclusion may have been the affect generated by the boys to work against Sara; however, Sara, instead, used it to empower the character of Sara-Siri.

With her unorthodox movements, Sara-Siri established a place for herself in the improv scene (Blackburn, 2016). Her moves disrupting the excluding assemblage perpetuated by the boys. The "fleshy humanness" of her actions—the textual production sparked by affect that disrupted the male dominated narrative Colton and Liam were perpetuating—allowed Sara-Siri, paraphrasing Burton, to thrive and survive in unorthodox ways (Burton, 2017, p. 35), in both the improv scene and the larger classroom assemblage. The teacher in Sara's classroom noted the racializing assemblage the boys were creating in the storied world of the Making Macaroni improv scene. His initial intervention and invocation of the classroom rules of improv helped Sara gain entry to the assemblage, an entry she then ran with through spoken and embodied textual productions. As Sara herself broke with classroom rules of improv, he did not intervene, allowing her fleshy responses to the boys. Through these supportive moves, Sara-Siri's clever and unorthodox ways with spoken and embodied texts were both validated and allowed to do their work in disrupting the boys' attempt to dismiss and exclude.

Openings

Affect should matter greatly to educators. As Stewart (2010) reminds us, affects, as "pre-personal intensities lodge in bodies. Events, relations and impacts accumulate as the capacities to affect and be affected" (Stewart, 2010, p. 339). Put simply: affect sticks around.

I offer these encounters with Crystal, Ally, and Sara to, following Leander and Ehret (2019), "generate an alter-economy of literacy research and pedagogy wherein new modes of expression and encounter value the excess of experience by allowing that excess to become what could be" (p. 23). This alter-economy of the girls' forays provide important questions to consider. Most notably: What might happen to young peoples' engagement with classroom literacies were educators to recognize the kinds of orienting and disorienting, transformative and normalizing affects that arise in classroom assemblages?

Affect cannot be readily measured. And for this reason, it may be eschewed in schooled literacy learning contexts. However, tuning in to affect can support critical, relational pedagogies (Dutro, 2019). Attending to the various faces of affects, and the bodies in which they bloom, provides important avenues for understanding and supporting young people's textual productions. Such attention can alert educators to racializing assemblages and the normative practices that so easily, and all too commonly, marginalize certain students. Finally, attuning to young people's affective responses (the reluctances and the delights) and watching for the unsanctioned textual productions (the side-conversations, notes passed, doodles drawn, movements made), can provide educators with clues and directions for encouraging all students to feel they are recognized members of the classroom assemblage we think of as the literacy club.

References

- Bennett, J., 2001. *The Enchantment of Modern Life: Attachments, Crossings, and Ethics*. Princeton University Press, Princeton, NJ.
- Blackburn, M., 2016. Traditionally marginalized bodies making space. In: Enriquez, G., Johnson, E., Kontovourki, S., Mallozzi, C.A. (Eds.), *Literacies, Learning and the Body: Putting Theory and Research Into Pedagogical Practice*. Routledge, New York, pp. 170–181.
- Burton, J.A., 2017. *Posthuman Rap*. Oxford University Press, New York.
- Clough, P.T., 2007. Introduction. In: Clough, P.T., Halley, J.O. (Eds.), *The Affective Turn: Theorizing the Social*. Duke University Press, Durham, N.C., pp. 1–33.
- Daniels, H., 2002. *Literature Circles: Voice and Choice in Book Clubs and Reading Groups*. Stenhouse, Portland, ME.
- DeLanda, M., 2006. *A New Philosophy of Society: Assemblage Theory and Social Complexity*. Continuum, New York.
- Deleuze, G., Guattari, F., 2004. *A Thousand Plateaus: Capitalism and Schizophrenia*. Trans. In: Massumi, B. (Ed.). University of Minnesota Press, Minneapolis.
- Dernikos, B.P., Lesko, N., McCall, S.D., Niccolini, A.D., 2020. *Mapping the Affective Turn in Education: Theory, Research, and Pedagogies*. Routledge, New York.
- Dernikos, B., 2020. Tuning into "fleshy" frequencies: a posthuman mapping of affect, sound, and de/colonized literacies with/in a primary classroom. *J. Early Child. Literacy* 20 (1), 134–157.
- Dutro, E., 2019. How affect theory can support social justice in our literacy classrooms. *Lang. Arts* 96 (6), 384–399.
- Dyson, A.H., 2003. *The Brothers and Sisters Learn to Write*. Teachers College Press, New York.
- Fey, T., 2011. *Bossypants*. Little, Brown and Co, New York.
- Franklin-Phipps, A., 2020. Art encounters, racism, and teacher-education. In: Dernikos, B.P., Lesko, N., McCall, S.D., Niccolini, A.D. (Eds.), *Mapping the Affective Turn in Education*. Routledge, New York, pp. 122–140.
- Gannon, S., 2009. Difference as ethical encounter. In: Davies, B., Gannon, S. (Eds.), *Pedagogical Encounters*. Peter Lang, New York, pp. 69–88.
- Gregg, M., Seigworth, G.J. (Eds.), 2010. *The Affect Theory Reader*. Duke University Press, Durham, NC.
- Guthrie, J.T., Klauda, S.L., 2014. Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Read. Res. Q.* 49 (4), 387–416.
- Hemmings, C., 2005. Invoking affect. *Cult. Stud.* 19 (5), 548–567.
- Hickey-Moody, A., 2013. Affect as method: feelings, aesthetics, and affective pedagogy. In: Coleman, R., Ringrose, J. (Eds.), *Deleuze and Research Methodologies*. Edinburgh University Press, Edinburgh, pp. 79–95.
- Hruby, G.G., Burns, L.D., Botzakis, S., Groenke, S.L., Hall, L.A., Laughter, J., Allington, R.L., 2016. The metatheoretical assumptions of literacy engagement: a preliminary centennial history. *Rev. Res. Educ.* 40, 588–643.

- Jocson, K.M., Dixon-Román, E.J., 2020. Becoming Shuri: CTE, racializing affect, and the becoming-technologist. *Read. Res. Q.* <https://doi.org/10.1002/rq.310>.
- Lather, P., 2016. Top ten+ list: (re)thinking ontology in (post)qualitative research. *Cult. Stud.* 16 (2), 125–131.
- Leander, K., Boldt, G., 2013. Rereading “A pedagogy of multiliteracies”: bodies, texts, and emergence. *J. Literacy Res.* 45 (1), 22–46.
- Leander, K., Ehret, C., 2019. *Affect in Literacy Learning and Teaching: Pedagogies, Politics, and Coming to Know*. Routledge, New York.
- Lenters, K., McDermott, M., 2020. *Affect, Embodiment, and Place in Critical Literacy: Assembling Theory and Practice*. Routledge, New York.
- Lenters, K., Whitford, A., 2018. Failing with grace: kids, improv, and embodied literacies. *Literacy* 52 (3), 117–127.
- Lenters, K., Whitford, A., (in press). “I’m going to go make macaroni”: classroom improv as ethical encounter. *Engl. Teach. Pract. Critiq.* 19(4), 463–478.
- Lenters, K., 2014. Just doing our jobs: a case study of literacy-in-action in a fifth grade literature circle. *Lang. Literacy* 16 (1), 53–70.
- Lenters, K., 2016. Riding the lines and overwriting in the margins: affect and multimodal literacy practices. *J. Literacy Res.* 48 (2), 280–316.
- Lenters, K., 2018. Multimodal becoming: literacy in and beyond the classroom. *Read. Teach.* 71 (6), 643–649.
- Luke, A., 1992. The body literate: discourse and inscription in early literacy training. *Ling. Educ.* 4, 107–129.
- Madonna, 2007. *The English Roses—Friends for Life*. Puffin Books—Penguin Group, Toronto, Canada.
- Massumi, B., 2004. Introduction. In: Deleuze, G., Guattari, F. (Eds.), *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press, Minneapolis. B. Massumi, Trans.
- Nxumalo, F., Villanueva, M.T., 2020. (Re)storying water: decolonial pedagogies of relational affect with young children. In: Dernikos, B.P., Lesko, N., McCall, S.D., Niccolini, A.D. (Eds.), *Mapping the Affective Turn in Education*. Routledge, New York, pp. 209–228.
- Pearson, P.D., Cervetti, G.N., 2015. Fifty years of reading comprehension theory and practice. In: Pearson, P.D., Hiebert, E.H. (Eds.), *Research-Based Practices for Teaching Common Core Literacy*. Teachers College Press, New York, pp. 1–24.
- Smagorinsky, P., O’Donell-Allen, C., 1998. The depth and dynamics of context: tracing the sources and channels of engagement and disengagement in students response to literature. *J. Lit. Res.* 30 (4), 515–559.
- Smith, F., 1988. *Joining the Literacy Club: Further Essays Into Education*. Heinemann, London.
- Snaza, N., Sandlin, J., 2017. #BlackLivesMatter: racialization, the human, and critical pedagogies of race. In: Means, A.J., Ford, D.R., Slater, G.B. (Eds.), *Educational Commons in Theory and Practice: Global Pedagogy and Politics*. Palgrave-MacMillan, London, pp. 109–125.
- Snow, C.E., Verhoeven, L., 2001. *Literacy and Motivation: Reading Engagement in Individuals and Groups*. Erlbaum, Mahwah, N.J.
- Spillers, H.J., 1987. Mama’s baby, Papa’s maybe: an American grammar book. *Diacritics* 17 (2), 64–81.
- Stewart, K., 2010. *Ordinary Affects*. Duke University Press, Durham, NC.
- Thiel, J.J., 2015. “Bumblebee’s in trouble!” Embodied literacies during imaginative superhero play. *Lang. Arts* 93 (1), 38–49.
- Weheliye, A.G., 2008. After man. *Am. Lit. Hist.* 20 (1/2), 321–336.
- Weheliye, A.G., 2014. *Habeas Viscus: Racializing Assemblages, Biopolitics, and Black Feminist Theories of the Human*. Duke University Press, Durham, NC.
- Wigfield, A., Guthrie, J.T., 1996. *Reading Engagement: Motivating Readers Through Integrated Instruction*. International Reading Association, Newark, NJ.
- Wynter, S., 2003. Unsettling the coloniality of being/power/truth/freedom: towards the human, after man, its overrepresentation—an argument. *CR N. Centen. Rev.* 3 (3), 257–337.

Posthumanism and Positioning Theory: explorations of embodiment and affect from a critical perspective

Mary B. McVee^a, Kelly A. Schucker^b, and Aijuan Cun^c, ^a University at Buffalo—SUNY, Buffalo, NY, United States; ^b University at Buffalo—SUNY, Niagara Falls, NY, United States; and ^c University of New Mexico, Albuquerque, NM, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	918
Perspectives on positioning and posthumanism	918
Foucault	919
Hollway	919
Harré's Positioning Theory: a brief summary	920
Storylines and the positioning triangle	920
Posthumanism and critical positioning: examples from literacy research	921
Positioning and embodiment	922
Embodied knowledge	922
Embodied practices	922
Embodied discourses	923
Embodied being	923
Positioning and affect	923
Positioning of self and others	924
Positioning and corporeal dynamics	924
Conclusion	925
References	925

Overview

In this article, we explore several aspects of positioning and positionality in relation to posthumanism. In particular, we consider an increasingly common framework in literacy studies, that of Positioning Theory, which originated in the work of Davies and Harré (1990) and Harré and colleagues (e.g., Harré and Van Langenhove, 1991, 1999a; Harré and Moghaddam, 2003). We ask: what might Positioning Theory gain from posthuman perspectives? This question is in response to Gergen's (2015) call for theorists to move from "issues of philosophic grounding to social utility" (p. 290). That is, we aim to consider not only philosophical issues around positionality and posthumanism, but also how theories of positionality can be potentially be used for transformation and change, rather than merely to represent phenomena. With those intentions in mind, we attempt to consider theory as a utilitarian tool for literacy scholars and those interested in literacy. We define literacy broadly with emphasis on discourses that shape and define literacy processes. This includes literate processes related to traditional alphabetic systems (e.g., reading a book on paper) or processes pertaining to digital literacies (e.g., composing a multimodal text). This conception of literacy is situated within social, discursive, and multimodal perspectives. Since much has been written about these areas (e.g., Cope and Kalantzis, 2000; Gee, 2015; Jewitt and Kress, 2003; Unrau et al., 2019), we do not review these perspectives here.

To begin our exploration, we briefly define posthumanism. Next, while making some comparisons with Positioning Theory and Davies and Harré, we consider the writings of two scholars, Foucault and Hollway, who have influenced scholarship in posthumanism and positioning. Third, we provide a brief overview of Positioning Theory, primarily as explored in the work of Rom Harré and other positioning scholars. We conclude by discussing three key intersections of posthumanism and positioning: critical positioning and literacy, embodiment, and affect, and review areas of synergy with positioning.

Perspectives on positioning and posthumanism

Posthumanism is generally viewed as a reaction to, or questioning of, the aspects of Enlightenment technical, rational, objective thought and beliefs that placed humans at the center of life that gave rise to humanist philosophies. Wolfe (2010) notes that Foucault takes issue with such assertions, opting instead to consider these as tensions, and others (e.g., Taylor and Hughes, 2016) observe such a direct lineage is not entirely accurate. However, as a starting place, it can be helpful to make this comparison while noting that posthumanism does not require the rejection of all that has come before. Posthumanism is about decentering human (or anthropocentric) perspectives that neglect the role of objects, animals, technologies, materials, and biologies. The term "posthumanism" itself is somewhat misleading. Wolfe (2010) articulates that posthumanism is not a renunciation of

the human, but rather that it “actually enables us to describe the human and its characteristic modes of communication, interaction, meaning, social significations, and affective investments with *greater* specificity” (p. xxv, emphasis original). As a theoretical lens, posthumanism is additionally challenging since it spans embodiment, affect, technology, and artificial intelligence among other areas.

Vermeulen (2014) defines posthumanism as a shift “to a more affirmative engagement with the world outside the text—with bodies, animals, affects, technologies, and materialities of different kinds” (pp. 121–122). Precisely because literacy scholars do not have theories that adequately address this shift, Leander and Burris (2020) assert that posthumanist theory can be useful. In *Posthumanism in Literacy Education: Knowing/Becoming/Doing*, Kuby et al. (2019) demonstrate what it means to take up posthumanist theory and apply it to literacy. Zapata et al. (2018) articulate that posthumanism and other critical theories “seek to reimagine and reconsider understandings about the ways multifarious entities enter into, sustain, and enmesh with worldly relations” (p. 481). Positioning Theory precisely focuses on the interactional relations that are intrinsically social, networked, and embodied. Fundamentally, positioning emphasizes the nuances of relationships—relationships with group identities, with other people, and with oneself. These relations are socially-distributed and not bound by time.

Applying posthumanist ideologies to literacy education, Zapata et al. (2018), express the value of considering the feminist perspective of how each of us is able to respond, regardless of whether we desire to be included in particular relational entanglements. Parrott (2003) echoes this sentiment, asserting that the dynamism of Positioning Theory lies in how it implicates people within any interaction as active participants (p. 33). This is critical in that Positioning Theory’s idealistic function is to enlighten those working in education and other fields of the complexity of human entanglements, and how human interrelations can elicit unconscious human biases. Positioning Theory has the potential to encourage self-reflection of how we position ourselves, how we are positioned, and how we position others—whether others are people, things, or animals. As a critical theory, posthumanism can help Positioning Theory focus not only on language as discursive but also on all aspects of positions whether related to people, animals, technologies, machines as well as other humans.

Foucault

In much of the scholarship on positioning and positionality, Foucault and Smith (1969/1972) is cited as the origin for subject positions (e.g., Deppermann, 2013, 2015). While Tirado and Gálvez (2007) discuss contributions of Foucault’s work to Positioning Theory, writing about their own line of thought, Harré and collaborators point to Hollway (1984/2005) as influential. In her writing, Hollway refers to Foucault’s foundational notion of subject positions. She recognizes a gap in Foucault’s account of positioning and discourse as she writes about heterosexual gender differences and the production of subjectivity: “He [Foucault] still does not account for how people are constituted as a result of certain truths being current rather than others how is it that people take up positions in one discourse rather than another?” (p. 234). These questions, Hollway observes, demand critical examination of hierarchies and social histories so as to illuminate patterns of positioning within discourses. This “also requires a question concerning the *investment* in that position” (Hollway, 2005, p. 234, emphasis original). She acknowledges that Foucault’s usage of the term discourse and its historical nuances are crucial in understanding positions (p. 226). Discourse positions individuals within external systems because discourses are produced by power within a social order. For example, education allows access to discourse but simultaneously controls the appropriation of discourses and access to power of those discourses.

Hollway

In her writing, Hollway (1984/2005) writes of taking up positions and positioning oneself with regard to gender:

Discourses make available positions for subjects to take up. These positions are in relation to other people. Like the subject and object of a sentence ... women and men are placed in relation to each other through the meanings which a particular discourse makes available ... (2005, p. 233).

Positions are not predetermined and do not exist independently but are produced and “re-produced,” as Hollway affirms, through interaction. Hollway’s intentional use of the hyphen signals the ways in which any day-to-day practice is individually a production. These re-productions produce (and reproduce) gendered subjectivity. Hollway focuses on positioning in discourses and how different discourses “make available different positions and different powers” (2005, pp. 226–227). She emphasizes Foucault’s historical perspective on discourses and the importance of individual histories. Davies (1991) echoes this perspective in her feminist writings on positioning stating that “discourses, embodiment, and the material force of the discourses through both individuals and social structures” (p. 43) are related to subject positions and agency. Positions available to be taken up exist not only in conversational discourses but in “body/landscape relations” (Davies, 2003, p. 279). That is, discourses are embodied and shape the power and agency that individuals can access.

Harré's Positioning Theory: a brief summary

In *The Self and Others* (2003, p. 139), Moghaddam, Hanley, and Harré remark that Positioning Theory is derived from feminist developmental and social psychologists [Hollway \(1984\)](#) and [Davies \(1991\)](#). In their original article introducing Positioning Theory, [Davies and Harré \(1990\)](#) write:

Once having taken up a particular position as one's own, a person inevitably sees the world from the vantage point of that position and in terms of the particular images, metaphors, story lines and concepts which are made relevant within the particular discursive practice in which they are positioned (p. 46).

This widely quoted passage is a succinct expression of the metaphor of positioning. While the metaphor is essential, Positioning Theory is about more than merely a metaphorical sense of where one is located within an interaction. [Davies and Harré \(1990\)](#) seek alternatives to the static "concept of role in developing a social psychology of selfhood" (p. 43) and observe that positions:

permit us to think of ourselves as a choosing subject, locating ourselves in conversations according to those narrative forms with which we are familiar and bringing to those narratives our own subjective lived histories (p. 52).

Even in interactive approaches, such as Goffman's frames or footing, roles are viewed as static or fixed (see [Davies and Harré, 1990](#), pp. 52–55; [Harré and Slocum, 2003](#), pp. 126–127). By contrast, positions are dynamic, fluid, and immanent, representing "a cluster of rights and duties to perform certain actions" ([Harré and Moghaddam, 2003](#), p. 5). People understand and experience positions through lived narrative experience. This includes beliefs, emotions, and subjective histories in addition to knowledge of social structures and their accompanying rights and duties. This is not to say that "choosing subjects" are fully independent. [Parrott \(2003\)](#) indicates that positioning is "dynamic ... extending the range, flexibility, and creativity of self-definitions that are under negotiation" (pp. 32–33). However, positions are always linked to "cultural, legal, institutional, conversational and personal moral orders" ([Van Langenhove, 2017](#), pp. 5–6) which influence a person's agency or perception of agency.

Storylines and the positioning triangle

[Davies and Harré \(1990\)](#) describe the "metaphor of an unfolding narrative" (p. 53) where discourses constitute one or more positions within a storyline or storylines.

If we are to come close to understanding how it is that people actually interact in everyday life we need the metaphor of an unfolding narrative, in which we are constituted in one position or another within the course of one story, or even come to stand in multiple or contradictory positions, or to negotiate a new position by "refusing" the position that the opening rounds of a conversation have made available to us.

[Davies and Harré \(1990, p. 53\).](#)

There is a strong relationship between positioning, narrative, and conversation as ongoing, enacted processes. Storylines emphasize the "dynamics of social episodes" and how various individuals may contribute as storylines unfold ([Harré and Moghaddam, 2003](#), p. 6). The Positioning Theory triangle or triad has emerged as a key representation and analytic tool in Positioning Theory. Early theoretical works were limited in specifics of how to conduct analysis, but theorists proposed that analysis could begin from any of the triangle's vertices: storyline, position, or speech acts ([Harré and Moghaddam, 2003](#), p. 9), also sometimes referred to as "Speech and other acts" (e.g., [McVee, 2011](#), p. 6). Examples of analysis using the positioning triangle to analyze different scales of social encounters can be found in [Harré et al. \(2009\)](#). The triangle has also been adapted as a Positioning Theory Diamond to support more refined and systematic analysis and account for identity ([Slocum-Bradley, 2009](#)).

Whether conceived of as a triangle or diamond, as an analytic tool the representation can be applied to conversations. Conversations may be face-to-face dialog, but [Van Langenhove \(2019\)](#) cautions that the term "conversation" must be understood also to include other encounters such as when a person reads a sign prohibiting smoking. The essence of the social world is such "the primary social reality that can be conceived as a species-wide and history-long network of people speaking to each other in an environment that is socially meaningful" (p. 64). In prior work, [Harré and Van Langenhove \(1999b\)](#) elaborate that Positioning Theory takes as its starting point:

the idea that the constant flow of everyday life in which we all take part, is fragmented through discourse into distinct episodes that constitute the basic elements of both our biographies and the social world. The skills that people have to talk are not only based on capacities to produce words and sentences but equally on capacities to follow rules that shape the episodes of social life. Not only what we do but also what we can do is restricted by rights, duties and obligations we acquire, assume or which are imposed upon us in the concrete social contexts of everyday life (p. 4).

McVee et al. (2019b) have also refined the positioning diamond (Slocum-Bradley, 2009) to represent moral fields and the ways that storylines, speech and other acts, and positions are not bounded by linear representations of time. "The distinction between past, present, and future does not go over neatly into psychological time, partly because the social and psychological past is not fixed. The social future can influence the social past" (Van Langenhove and Harré, 1999, p. 15). Researchers have demonstrated how embodied interactions can be analyzed not only for actual spoken or written conversation but also for multimodal meaning making within moral fields and to consider identity, rights, and duties (McVee and Carse, 2016; McVee et al., 2019a) within research tied to embodiment and material artifacts (see McVee et al., 2021).

Posthumanism and critical positioning: examples from literacy research

Over the past few decades critical theories have increasingly been applied to literacy research (Anders et al., 2016; Brooks, 2019; Luke, 2019; Vasquez, 2014; Willis et al., 2008). We do not provide a review of this work, but assert that it provides a useful entry point for exploring intersections and potentials of posthumanism and Positioning Theory. Zapata et al. (2018) posit that posthumanism helps bring recognition to the ways that "seemingly disparate critical approaches (feminism, anticolonial and antiracist thought, technology studies, ecology, etc.) have a common ground in directly challenging the ways humanism has restricted politics and education" (p. 480). The definition draws attention to the importance of understanding the critical stance in the dominant discourse and academic discourse. Restrictive policies are usually designed by individuals who have power, authority, and agency (Davies, 1991) in dominant discourses. Individuals who are from different/nondominant linguistic, racial, and cultural backgrounds are usually positioned as marginalized groups who have duties and rights to perform social acts without agency in the discourses. Not only are these unequal positions historically and socially situated in the broader discourses, but these positions in relation to race or other critical issues are also consciously or subconsciously embedded within everyday teaching and learning in the classroom discourse. Zapata et al. (2018) call for a focus on "real social and physical materialities of race and class on a human" (p. 491). Several examples can illustrate this.

In 2010, the Tucson Unified School District banned the Mexican-American ethnic studies program (Delgado, 2012). This enactment dematerialized (Zapata et al., 2018) Latinx students' bodies in relation to race and ethnicity and was precipitated by those who had power and authority through a variety of discourses. Latinx students were positioned as others who had some rights (e.g., the right to go to school) but not the agency in schools to decide curriculum (i.e., Latinx students were no longer able to decide to engage with their own cultural materials). Similar events also occur in other institutional discourses. For instance, in 2019, the African American Studies program at the University of Buffalo was defunded because the number of African American faculty had decreased (The Spectrum, 2019), and the African American Studies program was merged into another department. Students were not given choices to decide if they wanted to maintain the program or to hire new faculty. These cases demonstrate positions and power related to institutional discourses. From the posthumanist perspective, however, research related to education and power structures should not just focus on dualisms (Davies, 1991) such as ethnic studies or no ethnic studies. Rather, expanding humanist perspectives by deconstructing dualisms may provide access to examine positionality within institutional discourses. Amplified attention on how various positions come into play in the institutional discourses would illuminate the influence of social hierarchies on interpersonal dynamics. Positioning Theory helps examine nuances (i.e., intersectional positions) concerning materials of critical issues such as "racism, classism, and sexism" (Zapata et al., 2018, p. 491) in the institutional discourse. Even in these cases, further analyses could analyze material or ideological positions pertaining to culture, law, conversation, and personal moral orders (Van Langenhove, 2017) surrounding institutional decisions. Such critical examination and reflection may illuminate unconscious (and perhaps unintentional) biases in decision-making that produce privilege/oppression and access/deprivation.

Not only do the posthumanist perspective and Positioning Theory help people examine power structures in societal and institutional discourses, but they also help educators and researchers examine unequal power positions and critical issues in the literacy field. Zapata et al. (2018) utilize the perspective of materiality to emphasize the materials of critical issues and ways that racism, sexism, or classism are entangled with literacy practices in the classroom. The bodies of students from linguistically and racially diverse backgrounds are often dematerialized. That is, bodies are often treated as if they are cut off from the event, are unimportant, or are inconsequential (Zapata et al., 2018). For example, as the unequal positions are sustained in school discourses, teachers experience challenges when working with the students. In the United States, 79% of public school teachers are White monolingual teachers (U.S. Department of Education, 2016). How can the teachers be prepared for working with their students? One way to achieve this purpose is to examine individual "subjectivity" (Davies, 1991, p. 47) and "become more multiple, flexible and inclusive of different points of view" (p. 51).

A number of scholars have utilized Positioning Theory to examine White teachers' subjectivities and self-critique in multiple ways, such as book-club discussions and cross-cultural experiences (Brock et al., 2006; Brock and Pennington, 2014; Glazier, 2005; McVee et al., 2004; McVee, 2014). The discussions on critical issues enabled the teachers to examine their own subjectivities in relation to the dominant discourse, learn about their students' positions from new perspectives, shape their beliefs as literacy educators, and reconsider their instruction and pedagogy for supporting all students. For example, drawing upon Positioning Theory, Brock et al. (2006) explored how six in-service teachers who enrolled in a cross-cultural summer course re-examined their own perspectives on culture and literacy instruction for students who are from diverse backgrounds. Participants had opportunities to engage in online discussions on multicultural texts and then traveled to Costa Rica to acquire cross-cultural experiences by staying with local families. The authors found that teachers had experiences with displacement when they entered a different culture and

they repositioned themselves in new storylines that were different than those they understood through their “own interpretive lenses” (p. 58). In this case, educators used Positioning Theory to re-examine their position and beliefs. Similarly, [McVee et al. \(2004\)](#) investigated White female teachers’ learning and positioning in relation to race and class when participating in book club discussions. Not only did the authors draw upon Positioning Theory as a tool, but they also developed new kinds of positioning as analytical lenses to gain insights into how the participants examined their positions and challenged their thinking in relation to critical issues and literacy. Positioning Theory has expanded peoples’ relationships and lived worlds through illuminating problematic, previously unconscious biases. Using self-reflection to crystallize positionalities creates more flexible social interactions, which has the potential to alter the influence of power in social hierarchies.

These studies demonstrate aspects of critical literacy and position concerning race, class, and gender in teacher education. Indeed, power dynamics are embedded in all discourses. Instead of viewing colorblindness as an approach to teach all children, it is significant that teachers examine their individual positions in relation to the discourses, recognize “the materialities of race,” “open up different possibilities” ([Zapata et al., 2018](#), p. 490), and explore “what it meant to be an effective and responsible teacher for all children” ([McVee et al., 2004](#), p. 13). Posthumanism and Positioning Theory can be effective theoretical lenses to identify power dynamics and to assist teachers, and others, in identifying the ways that discourses shape pedagogical decisions and actions but also in helping transcend racial, cultural binaries that are counter to democratic possibilities ([Snaza et al., 2014](#)).

Positioning and embodiment

Literacy studies that adopt posthumanism or Positioning Theory as an analytic or explanatory framework most often also draw from critical theories. Despite this critical focus, many literacy studies do not explicitly develop or explore embodiment or affect. Writing of posthumanism in the context of literacy studies, [Zapata et al. \(2018\)](#) articulate that the “philosophical perspectives on knowing (epistemology), be(com)ing (ontology), and doing (axiology)” are the basis of posthumanist scholarship (p. 478). Positioning Theory shares some of these same roots through embodiment. Although not always foregrounded, Harré’s body of work (cf. *Personal Being*, 1983) makes clear that embodiment is a fundamental tenet of Positioning Theory. We propose that the “mind extends beyond the skin” ([Wertsch, 1991](#), p. 14), in that mind is discursively produced (e.g., [Harre and Gillett, 1994](#); [Harre and Stearns, 1995](#)) and socially distributed ([Gee, 1992, 1997](#); [Lave and Wenger, 1991](#); [Rogoff and Lave, 1984](#)). The *New London Group* (1996) similarly assert that “the human mind is embodied, situated, and social” (*New London Group, 1996*, p. 82; emphasis added).

[Roelvink and Zolkos \(2015\)](#) connect *affect* with *embodiment*: “affect is often interpreted as a recognition and validation of embodied and emotive forms of knowledge production” (p. 4). They understand the importance of the body in terms of identity and intellectual awareness. The body is defined through the multiple and varied identities it creates, as well as through bodily movement. In this regard the body does not “remain fixed or static but [is] mediated by processes and practices that produce dynamic points of intersection and connection” ([Blackman, 2009](#), p. 107). The potentiality of embodiment can be observed in different forms, including embodied knowledge (i.e., spatial embodiment, and embodied practices). Thus, the role of embodiment can be understood in various ways:

1. Embodied knowledge
2. Embodied practices
3. Embodied discourses

Embodied knowledge

Embodied knowledge, or what [Bourdieu \(1986\)](#) labeled as “embodied capital,” emphasizes the value of identity-based knowledge. [Fairbanks and Ariail \(2006\)](#) explicate this idea, arguing this is knowledge that is embodied in thoughts or perceptions through which individuals engage and interact with the material world. They emphasize that embodying positional identities gives an opportunity for individuals to “resist and revise them,” (p. 316) creating new opportunities and ways to exist in the world. Embodied capital is demonstrated through a series of power relations, including race, class, ethnicity, nationality, gender and so on. Each of these facets of identity is connected with capital—some more highly valued than others—in our social world. For example, as explored in Bourdieu’s scholarship racial hierarchies have presumed that some forms of racial capital are more or less valuable than others. Historically in the US, for example, the capital of White people has been of greater perceived value than that of other races, thus giving Whites more access to discourses of power.

Embodied practices

“Embodied practices” ([Veninga, 2009](#), p. 116) are “the ways in which individuals do or perform who they are” ([Tzou and Bell, 2012](#), p. 267). This has been exemplified in literacy scholarship. In examining students doing digital video production, [Rogers et al. \(2010\)](#) observe that students embody genre practices in how “they borrow, juxtapose, hybridize, or transform genres and modes” through video composition practices to generate “new discourse narratives” (p. 301) that function both as locations of subject positions and as methods of social critique. With regard to counterstories, [Winslade \(2012\)](#) notes a similar concept arguing that this literary genre “embodies in narrative practice the impulse toward resistance in an intentional construction of an identity story. It stresses the value

of agency" (p. 29). These studies represent aspects of literacy as an embodied practice; as [McVee et al. \(2005\)](#) argue: "Language is a way of *doing* things in the world" (p. 546, emphasis added).

Embodied discourses

This leads us to the third role of embodiment, closely related to embodied practices: embodied discourse. [McVee et al. \(2005\)](#) describe how embodiment is rooted within our sociocultural worlds through language:

The discourse processes required in the moment are embedded in cultural histories not just of individuals but of kinship groups, social classes, and political ideologies. It is not merely words but social and cultural practices, opportunities, and interactions that must be provided. In addition, these experiences, opportunities, and interactions must allow for a child to engage with a range of mediational artifacts, both material and ideal, thus tying language to authentic material and cognitive practices and embodied activity (p. 546).

Embodied discourses, especially those that are multiple and contradictory ([Fitch, 2003](#)), are an active process. An example of this can be seen in an engineering literacies club ([McVee et al., 2017, 2019a](#)) wherein teachers taught third-grade students about engineering skills and disciplinary literacies and worked to instill Engineering Habits of Mind ([Royal Academy of Engineering, 2014; McVee et al., under review](#)). Teachers used engineering discourses which were, over time, adapted by the students. These discourses were more than the cultivation of scientific terminology—they represented new ways of thinking, perceiving, working collaboratively, asking questions, and using creative problem-solving skills to accomplish a shared goal involving material artifacts ([Silvestri et al., 2021](#)). The phrase *Engineering Habits of Mind* captures the interactions that took place in this engineering literacies club: using engineering discourses began as the adoption of jargon and evolved into the development of an engineering mindset through the use of artifacts and embodied interactions (e.g., physical and social positionings with, around, and through artifacts). This embodiment of discourse transformed research participants from novice students to "expert" third-grade engineers.

Embodied being

Embodiment implies a physical being that lives in a social world. Posthumanism can offer unique insights to Positioning Theory, and the field of education, in this sense. [Zapata et al. \(2018\)](#) write: "posthumanist scholarship is rooted in a relational ontology, meaning that we come to know through our being with and in the world" (p. 479). As articulated by [Gee \(2000\)](#), the field of education has been notorious for dismissing the physical bodies of students, metaphorically having students check their bodies at the door. Only in recent decades have academics begun to make the shift to perceive students in their entirety. [McVee et al. \(2005\)](#) propose this to be "a paradigm shift that involves radically transforming our perspective so that we are considering 'minds-on-bodies-on' as students engage the body-mind and researchers consider what it means to learn from a body-mind perspective" (p. 157). Zapata et al. also acknowledge the way that bodies have been dismissed—or outright ignored—specifically in terms of class, race, and gender. Positioning Theory is an attempt to subvert this idea of students leaving their bodies at the door by focusing specifically on how identities, privileges, and power are constructed, reproduced, and sustained, but also how they may be dismantled, transformed, adapted, or challenged. Posthumanism and Positioning Theory can readily be used in conjunction with each other to illuminate complex relationships across all manner of educational domains and relations.

Finally, we must note that not all embodiment happens within the body; rather, spatial embodiment is the perception of geographical spaces as embodying histories, ideologies, and value-systems ([Green and Whitsed, 2012; Hungerford-Kresser and Vetter, 2012](#)). [Tzou and Bell \(2012\)](#) agree that places embody "narratives and norms," and social interactions within a learning environment are not "neutral or random" but are entrenched in "historical and spatial contexts that invite or prohibit opportunities for learning" (p. 266) and engagement.

Positioning and affect

In "Posthuman Affect," Pieter Vermeulen addresses the connection between embodiment and affect: "a posthumanism that insists on embodiment cannot simply be reduced to one particular emotive disposition; instead, it generates a field of affective possibilities that different posthumanist practices can inhabit in different ways" ([Vermeulen, 2014](#), p. 124). One of these "affective possibilities" is the complex, nuanced role identity plays in human relationships. Posthumanism, not unlike Positioning Theory, considers affect integral to the "interplay of bodies" ([Roelvink and Zolkos, 2015](#), p. 2). Considering the fundamental aspect of humanity to be relational, posthumanism emphasizes the importance of human connections—the animation between people. Roelvink and Zolkos, in their piece "Posthumanist Perspectives on Affect," emphasize that recent scholarship on affect "has used the category of 'the animate' ... as synonymous with a certain *force of aliveness* and of a *potential for becoming*" (2015, p. 2, emphasis original). They define the primary facets of affect to be sensation, subjectivity, sociality, and politics (p. 4). Positioning Theory considers similar tenets that affect how people position each other—especially through social subjectivity, relationships, and identity. In his work [Harré](#) argues that emotions and meaning are grounded in the bodies as material things and social beings ([Harré, 1983, 1991](#)). Positioning Theory

functions on multiplicity—the position of one person, without anything in relation for comparison, is meaningless; positioning depends on interrelations.

Positioning of self and others

How people position others and are positioned by others is impacted by gender, social class, race, ethnicity, linguistic background, and other relational power dynamics. Referencing positioning scholars [Harré and Van Langenhove \(1999a\)](#), [Arkoudis \(2005\)](#) articulates that “people bring to their conversations different rights and responsibilities that reflect differences in power relations and affect their ability to position themselves and each other” (2005, p. 175). These social identities are related to how students position themselves (defined as “intentional self-positioning”), and how they position others (defined as “interactive positioning”) that impacts student learning ([De Costa, 2011](#), p. 349). We detail examples of these complex interactive possibilities below.

The following studies examined the roles of class and gender in regards to Positioning Theory. In a study focusing on student-led discussions in a classroom, [Clarke \(2006\)](#) found that gender and class affect student dialog and “discursive interactions” (p. 55). Social class, through this lens, was also examined by [Payne-Bourcy and Chandler-Olcott \(2003\)](#), who found that social class affected a research participant’s literacy learning due to differences in community language practices and academic language practices. Similarly, [Freeman \(2010\)](#), found the needs of students from working class families were handled by teachers/staff differently. She also found that working-class parents’ backgrounds influenced how they participated and also whether they did participate. It is essential to note that social class influenced perceptions and emotions tied to participation (or lack of participation), but social class alone does not determine positions. Importantly, Freeman discovered that parents were not always passive in these social interactions. In their affective and agentive responses, parents also demonstrated what they believed to be “effective countermeasures in their efforts to reposition themselves and their children in the eyes of the administrators” (2010, p. 189).

Race is also a critical factor that impacts peoples’ positioning (e.g., [Glazier, 2003](#); [Harré and Slocum, 2003](#); [Mahar, 2001](#); [McVee, 2011](#); [McVee et al., 2004](#); [Monrouxe, 2010](#); [Rickford et al., 2004](#); [Vetter, 2013](#)). [Monrouxe \(2010\)](#) examined how medical students’ perceptions of their patients were linked to race, articulating that race “affects doctors’ feelings of affiliation toward patients” (p. 44) paired with their belief regarding how likely they thought patients were to follow to the doctor’s recommendations. Positioning and race is similarly disconcerting within an academic context. [Rickford et al. \(2004\)](#) discussed historical injustice and oppression that has plagued academic institutions, specifically affecting the academic “experiences and classroom membership of youth who speak and value AAL [African American Language]” (cited in [Vetter, 2013](#), 175). This study found that students speaking African American Vernacular English (or AAL) in classrooms invoked “issues of power and status” that “affected the formation of a respectful community” (p. 200). [Vetter \(2013\)](#) also conducted a study about race in education, gathering data in a classroom with a White teacher and a racially diverse student population. As a result, she argued that it is imperative for teacher educators to create and protect opportunities for preservice teachers to reflect on the ways that power might affect teaching methodologies, and how they approach their students.

Similarly, nationality and linguistic background play a significant role in positioning. [Maringe and Jenkins \(2015\)](#) noted that international students’ self-perception and agency was inhibited by not having an established stance to voice personal concerns they had about the university (p. 622). These perceived “constraints” were influenced by students’ vulnerabilities, which included anxiety about being othered, being marked as non-native speakers, and a lack of places to voice concerns about critical issues affecting them. These concerns are important in that they reflect how the international students are positioned (or feel they are being positioned). In turn, this influences how these students position themselves. This idea of the everyday “doing of language” ([Vetter, 2013](#), p. 174) influences a person’s sense of membership within a community, which is inarguably less accessible for students with different linguistic backgrounds than students who represent a linguistic or cultural majority.

Positioning and corporeal dynamics

[Roelvink and Zolkos \(2015\)](#) considered the “affective turn,” emphasizing the increased popularity of the phrase in relation to “the renewed interest in the category of experience, especially, though not exclusively, in terms of its corporeal, non-conscious and non-verbal dimensions” (p. 5). Positioning Theory is an ideal vehicle for exploration of non-conscious or dysconscious experiences (e.g., “dysconscious racism” [[King, 1991](#), p. 133]). Very often, as is articulated specifically in many articles already cited, research participants are unaware of the misperceptions, misconceptions, and biases that impact how they perceive others, how they perceive being positioned by others, and how they self-position. [Parrott \(2003\)](#) posited that “emotions often establish a position that is defined relative to other persons’ positions” (p. 31). Positioning is, therefore, almost always a subconscious response. Posthumanism’s attention to the corporeal and non-verbal aspects of positioning is particularly useful in uncovering these implicit positions. This is a main tenet of Positioning Theory—to make visible what can easily be dismissed, or missed altogether. A significant number of studies in this field focus explicitly on how to become aware of how people position others and themselves.

As previously observed, positions are not static and unchanging. A fundamental aspect of Positioning Theory is the dynamism of human relationships. Positioning Theory attempts to capture the ways that positions are unstable and shift over time, place, and social context. Similarly, [Roelvink and Zolkos \(2015\)](#) note that scholarship on affect has disputed the concept of a human as a “singular, cohesive, and bonded unit” that could be “characterized by relative stability over time” (2015, p. 7). In one such example, [Tanghe and Park \(2016\)](#) use an example of two research participants who “transformed their preconceived notions

that affected their initial ways of positioning one another" (p. 11). This demonstrates the fluidity of positioning, where perception shifts and can be affected by a myriad of things, which alter how people perceive and position others and themselves.

Positions are not simply relational; they have significant meaning regarding actions within a context, requiring knowledge about complex social interactions, and an understanding of how a person's position will "affect what happens in particular contexts" (Rainville and Jones, 2008, p. 447). Positioning Theory, and the role affect plays, are important because metaphorical, subconscious positionings impact self-perception. For example, Hazari et al. (2015), remark that how students feel they are being positioned can affect how students themselves access particular identities. Gee (2011) argues that when people express themselves they are composing a version of reality, both about themselves and their worlds. Rainville and Jones (2008) expand this sentiment, arguing that sociality is not made in isolation "but instead is affected by the actions of others" (p. 440). Discourse reflects reality as it is perceived and felt through affect and the body.

Conclusion

The phrase, "There is nothing as practical as a good theory" is attributed to psychologist Lewin (1943, p. 118). The purpose of theories, particularly those that are analytical and exploratory rather than predictive, is to be utilized in some way. In this regard, when applied to data, posthumanism and Positioning Theory have the potential to surface positions that may go unobserved or unexcavated. Certainly, any number of articles could be written about differences or disagreements between discursive psychology and posthumanism. But, as noted in the introduction, our intent is to ask how theories may be applied in a potentially transformative manner.

Traditionally, posthumanism and Positioning Theory have originated from different disciplinary fields and scholarly perspectives. We have attempted to consider what each of these concepts might have to offer the other. Specifically, Positioning Theory can benefit from strains of posthumanism thought in encouraging or enabling scholars to more thoroughly explore embodied and nonhuman aspects of positioning, and how such dynamics function in terms of storylines, rights, duties, and identities. These embodied aspects of posthumanism and Positioning Theory are both connected to affect and the underlying felt sense that shapes understandings of ourselves and interactions with others—whether those others are animals, humans, machines, or technologies. The posthuman project to bring attention to a myriad of others from animal to artificial intelligence challenges positioning theorists to move beyond human-to-human speech interaction to other forms of positioning such as that between artifacts and people (McVee et al., 2021). Particularly as represented in studies in literacy, Positioning Theory brings an analytic specificity to identifying and uncovering often implicit positions of which people are typically unaware. Given the work already carried out across many disciplines by positioning scholars, Positioning Theory can provide analytic means to examine data and real world interactions.

Posthumanism and Positioning Theory explorations offer affordances to look beyond binaries to open up opportunities. For example, researchers can explore multiple positions around race and racism instead of only two positions: racist/anti-racist. Such investigations can allow for clearer delineations and practices to understand processes that can move people toward a desirable outcome, such as becoming anti-racist. By bringing positions to the surface, and identifying related positionings, these theories can leverage their strengths to assist researchers in analysis. With regard to education, it is especially important that, once identified, negative impacts of positioning might be mitigated or reshaped through other actions and interactions, particularly where harm has occurred or is occurring as a result of positionings. It is here that the potentially transformational power of Positioning Theory and posthumanism are revealed. Simply put, negative impacts of positions cannot be addressed until they are identified. Once identified educational practices, processes, and interventions can be considered to bring about changes. Considering these two theories for their synergistic power holds forth the possibility of offering literacy scholars new tools to transform their understanding of positions, positionality, and positioning.

References

- Anders, P.L., Yaden, D.B., Da Silva Iddings, A.C., Katz, L., Rogers, T., 2016. Entanglements, intensities, and becoming: non-representational perspectives on literacy research [Themed Issue]. *J. Lit. Res.* 48 (3). <https://doi.org/10.1177/1086296x16662393>. Available at.
- Arkoudis, S., 2005. Fusing pedagogic horizons: language and content teaching in the mainstream. *Ling. Educ.* 16 (2), 173–187. <https://doi.org/10.1016/j.linged.2006.01.006>. Available at.
- Blackman, L., 2009. *The Body: The Key Concepts*. Bloomsbury, London.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood, pp. 241–258.
- Brock, C.H., Pennington, J.L., 2014. Exploring three White American teachers' dispositional stances towards learning about racial, cultural, and linguistic diversity. *Pedagogies* 9 (4), 322–342. <https://doi.org/10.1080/1554480x.2014.955497>. Available at.
- Brock, C., Wallace, J., Herschbach, M., Johnson, C., Raikes, B., Warren, K., Nikoli, M., Poulsen, H., 2006. Negotiating displacement spaces: exploring teachers' stories about learning and diversity. *Curric. Inq.* 36 (1), 35–62. <https://doi.org/10.1111/j.1467-873x.2005.00345.x>. Available at.
- Brooks, M.D., 2019. Untapped possibilities: intersectionality theory and literacy research. In: Alvermann, D., Unrau, N., Sailors, M., Ruddell, R. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 419–429. <https://doi.org/10.4324/9781315110592-25>. Available at.
- Clarke, L.W., 2006. Power through voicing others: girls' positioning of boys in literature circle discussions. *J. Lit. Res.* 38 (1), 53–79. https://doi.org/10.1207/s15548430jlr3801_3. Available at.
- Cope, B., Kalantzis, M. (Eds.), 2000. *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge, New York.
- Davies, B., 2003. Positioning the subject in body/landscape relations. In: Harré, R., Moghaddam, F.M. (Eds.), *The self and others*. Praeger, Westport, CT, pp. 279–295.

- Davies, B., Harré, R., 1990. Positioning: the discursive production of selves. *J. Theor. Soc. Behav.* 20 (1), 43–63. <https://doi.org/10.1111/j.1468-5914.1990.tb00174.x>. Available at.
- Davies, B., 1991. The concept of agency: a feminist poststructuralist analysis. *Soc. Anal.* 30, 42–53.
- De Costa, P.I., 2011. Using language ideology and positioning to broaden the SLA learner beliefs landscape: the case of an ESL learner from China. *System* 39 (3), 347–358. <https://doi.org/10.1016/j.system.2011.07.007>. Available at.
- Delgado, R., 2012. Precious knowledge: state bans on ethnic studies, book traffickers (librotráfico), and a new type of race trial. *N. C. Law Rev.* 91, 1513–1554.
- Deppermann, A., 2013. Positioning in narrative interaction. *Narrat. Inq.* 23 (1), 1–15. <https://doi.org/10.1075/ni.23.1.01dep>. Available at.
- Deppermann, A., 2015. Positioning. In: De Fina, A., Georgakopoulou, A. (Eds.), *Handbook of Narrative Analysis*. John Wiley and Sons, Toronto, CA, pp. 369–386.
- Fairbanks, C.M., Ariail, M., 2006. The role of social and cultural resources in literacy and schooling: three contrasting cases. *Res. Teach. Engl.* 40 (3), 310–354.
- Fitch, F., 2003. Inclusion, exclusion, and ideology: special education students' changing sense of self. *Urban Rev.* 35 (3), 233–252.
- Foucault, M., 1969/1972. *The Archeology of Knowledge*. A. M. S. Smith, Trans. Pantheon Books, New York. original published 1969.
- Freeman, M., 2010. "Knowledge is acting": working-class parents' intentional acts of positioning within the discursive practice of involvement. *Int. J. Qual. Stud. Educ.* 23 (2), 181–198. <https://doi.org/10.1080/09518390903081629>. Available at.
- Gee, J.P., 1992. *The Social Mind*. Bergin & Garvey.
- Gee, J.P., 1997. Thinking, learning, and reading: the situated sociocultural mind. In: Kirshner, D., Whitson, J.A. (Eds.), *Situated Cognition: Social, Semiotic, and Psychological Perspectives*. Hampton Press, New York, pp. 235–302.
- Gee, J.P., 2000. Chapter 3: identity as an analytic lens for research in education. *Rev. Res. Educ.* 25 (1), 99–125. <https://doi.org/10.2307/1167322>. Available at.
- Gee, J.P., 2011. *Tools for Discourse Analysis*. Routledge, London, UK.
- Gee, J.P., 2015. *Social Linguistics and Literacies: Ideology in Discourses*, fifth ed. Routledge, London, UK.
- Gergen, K., 2015. From mirroring to world-making: research as future forming. *J. Theor. Soc. Behav.* 45 (3), 287–310. <https://doi.org/10.1111/jtsb.12075>. Available at.
- Glazier, J.A., 2003. Moving closer to speaking the unspeakable: white teachers talking about race. *Teach. Educ. Q.* 30 (1), 73–94.
- Glazier, J.A., 2005. Talking and teaching through a positional lens: recognizing what and who we privilege in our practice. *Teach. Educ.* 16 (3), 231–243. <https://doi.org/10.1080/10476210500204929>. Available at.
- Green, W., Whitsed, C., 2012. Reflections on an alternative approach to continuing professional learning for internationalization of the curriculum across disciplines. *J. Stud. Int. Educ.* 17 (2), 148–164. <https://doi.org/10.1177/1028315312463825>. Available at.
- Harré, R., Gillett, G., 1994. *The Discursive Mind*. Sage Publications, New York.
- Harré, R., Moghaddam, F.M. (Eds.), 2003. *The Self and Others*. Praeger, Connecticut.
- Harré, R., Slocum, N., 2003. Disputes as complex social events: on the uses of positioning theory. *Common. Knowl.* 9 (1), 100–118. <https://doi.org/10.1215/0961754x-9-1-100>. Available at.
- Harré, R., Stearns, P., 1995. *Discursive Psychology in Practice*. Sage Publications, London.
- Harré, R., Van Langenhove, L., 1991. Varieties of positioning. *J. Theor. Soc. Behav.* 21 (4), 393–407. <https://doi.org/10.1111/j.1468-5914.1991.tb00203.x>. Available at.
- Harré, R., Van Langenhove, L. (Eds.), 1999a. *Positioning Theory*. Blackwell, Oxford.
- Harré, R., Van Langenhove, L., 1999b. The dynamics of social episodes. In: Harré, R., Van Langenhove, L. (Eds.), *Positioning Theory*. Blackwell, Oxford, pp. 1–13.
- Harré, R., Moghaddam, F.M., Cairnie, T.P., Rothbart, D., Sabat, S.R., 2009. Recent advances in positioning theory. *Theor. Psychol.* 19 (1), 5–31. <https://doi.org/10.1177/0959354308101417>. Available at.
- Harré, R., 1983. *Personal Being: A Theory for Individual Psychology*. Blackwell, Oxford.
- Harré, R., 1991. *Physical Being*. Blackwell, Oxford.
- Hazari, Z., Cass, C., Beattie, C., 2015. Obscuring power structures in the physics classroom: linking teacher positioning, student engagement, and physics identity development. *J. Res. Sci. Teach.* 52 (6), 735–762. <https://doi.org/10.1002/tea.21214>. Available at.
- Hollway, W., 1984/2005. Gender difference and the production of subjectivity. In: Henriques, J., Hollway, W., Urwin, C., Venn, C., Walkerdine, V. (Eds.), *Changing the Subject*. Taylor & Francis eBook Library, pp. 223–261. Available at. <https://doi-org.gate.lib.buffalo.edu/10.4324/9780203298886>.
- Hungerford-Kresser, H., Vetter, A., 2012. Positioning and the discourses of urban education: a Latino student's university experience. *Urban Rev.* 44 (2), 219–238. <https://doi.org/10.1007/s11256-011-0193-y>. Available at.
- Jewitt, C., Kress, G.R. (Eds.), 2003. *Multimodal Literacy*. Peter Lang, New York.
- King, J.E., 1991. Dysconscious racism: ideology, identity, and the miseducation of teachers. *J. Negro Educ.* 60 (2), 133–146.
- Kuby, C.R., Spector, K., Thiel, J.J. (Eds.), 2019. *Posthumanism in Literacy Education: Knowing/Becoming/Doing*. Routledge, New York.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Leander, K.M., Burriss, S.K., 2020. Critical literacy for a posthuman world: when people read, and become, with machines. *Br. J. Educ. Technol.* 51 (4), 1262–1276. <https://doi.org/10.1111/bjet.12924>. Available at.
- Lewin, K., 1943. Psychology and the process of group living. *J. Soc. Psychol.* 17 (1), 113–131.
- Luke, A., 2019. Regrounding critical literacy: representation, facts and reality. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Framing Languages and Literacies*. Routledge, New York, pp. 349–361.
- Mahar, D., 2001. Positioning in a middle school culture: gender, race, social class, and power. *J. Adolesc. Adult Lit.* 45 (3), 200–209.
- Maringe, F., Jenkins, J., 2015. Stigma, tensions, and apprehension: the academic writing experience of international students. *Int. J. Educ. Manag.* 29 (5), 609–626. <https://doi.org/10.1108/ijem-04-2014-0049>. Available at.
- McVee, M.B., Carse, C., 2016. A multimodal analysis of storyline in "The Chinese Professor" political advertisement: narrative construction and positioning in economic hard times. *Vis. Commun.* 15 (4), 403–427. <https://doi.org/10.1177/1470357216637752>. Available at.
- McVee, M.B., Baldassarre, M., Bailey, N., 2004. Positioning theory as lens to explore teachers' beliefs about literacy and culture. In: 53rd Yearbook of the National Reading Conference. National Reading Conference, Oak Creek, WI, pp. 281–295.
- McVee, M.B., Dunsmore, K., Gavelek, J.R., 2005. Schema theory revisited. *Rev. Educ. Res.* 75 (4), 531–566. <https://doi.org/10.3102/00346543075004531>. Available at.
- McVee, M.B., Silvestri, K., Shanahan, L.E., English, K., 2017. Productive communication in an engineering afterschool club with girls who are English Language Learners. *Theor. Pract.* 56 (4), 246–254.
- McVee, M.B., Haq, K.S., Barrett, N., Silvestri, K.N., Shanahan, L.E., 2019a. The positioning theory diamond as analytic tool to examine multimodal social interaction in an engineering club. *Pap. Soc. Represent.* 28 (1), pp. 7–1. Available at. <http://psr.iscte-iul.pt/index.php/PSR/article/view/495/425>.
- McVee, M.B., Silvestri, K.N., Barrett, N., Haq, K.S., 2019b. Positioning Theory. In: *Theoretical Models and Processes of Literacy*. Routledge, pp. 397–416.
- McVee, M.B., Silvestri, K.N., Schucker, K.A., Cun, A., 2021. Positioning Theory, embodiment, and the moral orders of objects in social dynamics: how Positioning Theory has neglected the body and artifactual knowing. *J. Theor. Soc. Behav.* 51 (2), 192–214.
- McVee, M.B., Schucker, K.S., Jarmark, C.J., Shanahan, L.E. (under review). *Disciplinary Literacies and Engineering Habits of Mind*.
- McVee, M.B., 2011. Positioning Theory and Sociocultural Perspectives. *Sociocultural Positioning in Literacy: Exploring Culture, Discourse, Narrative, & Power in Diverse Educational Contexts*, pp. 1–22.
- McVee, M.B., 2014. "Some are way left, like this guy, Gloria Ladson-Billings": resistance, conflict, and perspective taking in teachers' discussions of multicultural education. *Peace Conflict* 20 (4), 536–551. <https://doi.org/10.1037/pac0000067>. Available at.

- Moghaddam, F., Hanley, E., Harré, R., 2003. Sustaining intergroup harmony: an analysis of the Kissinger papers through positioning theory. In: Harré, R., Moghaddam, F. (Eds.), *The Self and Others*. Praeger, Connecticut, pp. 137–155.
- Monrouxe, L.V., 2010. Identity, identification and medical education: why should we care? *Med. Educ.* 44 (1), 40–49. <https://doi.org/10.1111/j.1365-2923.2009.03440.x>. Available at.
- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–93. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>. Available at.
- Parrott, W.G., 2003. Positioning and the emotions. In: *The Self and Others: Positioning Individuals and Groups in Personal, Political, and Cultural Contexts*. Greenwood Publishing Group, Connecticut, pp. 29–43.
- Payne-Bourcy, L., Chandler-Olcott, K., 2003. Spotlighting social class: an exploration of one adolescent's language and literacy practices. *J. Lit. Res.* 35 (1), 551–590. https://doi.org/10.1207/s15548430jlr3501_2. Available at.
- Rainville, K.N., Jones, S., 2008. Situated identities: power and positioning in the work of a literacy coach. *Read. Teach.* 61 (6), 440–448. <https://doi.org/10.1598/rt.61.6.1>. Available at.
- Rickford, J.R., Sweetland, J., Rickford, A.E., 2004. African American English and other vernaculars in education: a topic-coded bibliography. *J. Engl. Ling.* 32 (3), 230–320. <https://doi.org/10.1177/0075424204268226>. Available at.
- Roelvink, G., Zolkos, M., 2015. Posthumanist perspectives on affect. *J. Theor. Humanit.* 20 (3), 1–20. <https://doi.org/10.1080/0969725x.2015.1065106>. Available at.
- Rogers, T., Winters, K.L., LaMonde, A.M., Perry, M., 2010. From image to ideology: analysing shifting identity positions of marginalized youth across the cultural sites of video production. *Pedagogies* 5 (4), 298–312. <https://doi.org/10.1080/1554480x.2010.509473>. Available at.
- Rogoff, B., Lave, J. (Eds.), 1984. *Everyday Cognition: Its Development in Social Context*. Harvard University Press, Cambridge, MA.
- Royal Academy of Engineering, 2014. *Thinking Like an Engineer: Implications for the Education System*. Retrieved from. <http://www.raeng.org.uk/thinkinglikeanengineer>.
- Silvestri, K.S., McVee, M.B., Jarmark, C.J., Shanahan, L.E., English, K.E., 2021. Multimodal positioning of artifacts in interaction in a collaborative elementary engineering club. *Multimodal Commun.* 10 (3), 289–309. <https://doi.org/10.1515/mc-2020-0017>.
- Stocum-Bradley, N., 2009. The positioning diamond: a trans-disciplinary framework for discourse analysis. *J. Theor. Soc. Behav.* 40 (1), 79–107. <https://doi.org/10.1111/j.1468-5914.2009.00418.x>. Available at.
- Shaza, N., Appelbaum, P., Bayne, S., Carlson, D., Morris, M., Rotas, N., et al., 2014. Toward a posthumanist education. *J. Curr. Theor.* 30 (2), 39–55.
- Spectrum, 2019. *UB Administrators Need to Hear Black Students' Cries for Help*. Retrieved from. <https://www.ubspectrum.com/article/2019/03/ub-administrators-need-to-hear-black-students-cries-for-help>.
- Tanghe, S., Park, G., 2016. "Build[ing] something which alone we could not have done": international collaborative teaching and learning in language teacher education. *System* 57, 1–13. <https://doi.org/10.1016/j.system.2016.01.002>. Available at.
- Taylor, C.A., Hughes, C. (Eds.), 2016. *Posthuman Research Practices in Education*. Palgrave Macmillan, London.
- Tirado, F., Gálvez, A., 2007. Positioning theory and discourse analysis: some tools for social interaction analysis. *Forum* 8 (2).
- Tzou, C.T., Bell, P., 2012. The role of borders in environmental education: positioning, power, and marginality. *Ethnogr. Educ.* 7 (2), 265–282. <https://doi.org/10.1080/17457823.2012.693697>. Available at.
- U.S. Department of Education, 2016. *The State of Racial Diversity in the Educator Workforce*. Retrieved from. <https://www2.ed.gov/rschstat/eval/highered/racial-diversity/state-racial-diversity-workforce.pdf>.
- Unrau, N.J., Alvermann, D.E., Sailors, M., 2019. Literacies and their investigation through theories and models. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, New York, pp. 3–34.
- Van Langenhove, L., Harré, R., 1999. Introducing positioning theory. In: Harré, R., Van Langenhove, L. (Eds.), *Positioning Theory*. Blackwell, Oxford, pp. 14–31.
- Van Langenhove, L., 2017. Varieties of moral orders and the dual structure of society: a perspective from positioning theory. *Front. Sociol.* 2 (9). <https://doi.org/10.3389/fsoc.2017.00009>.
- Van Langenhove, L., 2019. The discursive ontology of the social world. In: Christensen, B.A. (Ed.), *The Second Cognitive Revolution*. Springer, New York, pp. 63–71.
- Vasquez, V.M., 2014. *Negotiating Critical Literacies With Young Children*. Routledge, New York.
- Veninga, C., 2009. Fitting in: the embodied politics of race in Seattle's desegregated schools. *Soc. Cult. Geogr.* 10 (2), 107–129. <https://doi.org/10.1080/14649360802652103>. Available at.
- Vermeulen, P., 2014. Posthuman affect. *Eur. J. Engl. Stud.* 18 (2), 121–134. <https://doi.org/10.1080/13825577.2014.917001>. Available at.
- Vetter, A., 2013. "You need some laugh bones!": leveraging AAL in a high school English classroom. *J. Lit. Res.* 45 (2), 173–206. <https://doi.org/10.1177/1086296x12474653>. Available at.
- Wertsch, J.V., 1991. *Voices of the Mind: A Sociocultural Approach to Mediated Action*. Harvard University Press, Cambridge, MA.
- Willis, A., Montovan, M., Burke, L., Hunter, C., Herrera, A., Hall, H., 2008. *On Critically Conscious Research: Approaches to Language and Literacy Research*. Teachers College Press, New York.
- Winslade, J., 2012. Positioning among the lines of force in schooling: an issue for psychologists and counsellors in schools. *Educ. Child Psychol.* 29 (2), 18–29.
- Wolfe, C., 2010. *What is Posthumanism?* University of Minnesota Press, Chicago.
- Zapata, A., Kuby, C.R., Thiel, J.J., 2018. Encounters with writing: becoming-with posthumanist ethics. *J. Lit. Res.* 50 (4), 478–501. <https://doi.org/10.1177/1086296x18803707>. Available at.

Literacy and the spatial production of school during a pandemic

Kevin Leander and Michael Havazelet, Vanderbilt University, Nashville, TN, United States

© 2023 Elsevier Ltd. All rights reserved.

Some key contributions to literacy studies of spatial theory	929
Popular discourses of school space and Covid-19	930
Variation 1: physical, affective, and hyperreal productions of school space during Covid-19	931
The physical and affective reproduction of school space at home	931
Hyperreal school space on videoconference platforms	931
Discourse on cameras and student presence in hyperreal school space	932
Variation 2: affective shortfalls of hyperreal school spaces and remote learning	932
Anxiety and the mismatch between the presentation and experience of remote school	932
Missing connections and affective depletion in remote school	933
Techno-social discourses of student affect and mental health during Covid-19	933
Variation 3: critiquing, resisting, and transforming school space	934
TikTok critiques of the production and boundaries hyper-school	934
Resistance through play: student challenges to school space on TikTok	935
Techno-social transformation of school space	935
Final thoughts	936
References	936

In order to make a consideration of literacy and social space more accessible, memorable, and tangible, imagine we shift our imagination of a central image in theories and discussions of literacy. Imagine this shift involves seeing literacy and its relationship to the classroom differently. In this case, we picture the classroom not as an inert “container” of children, a teacher, and materials, but rather as a city. The classroom-as-city has of course all kinds of movements in it and across it: movements of trucks, motorcycles, bodies, animals, ambulances, natural gas, water, and sewage. Busses and commuter trains traverse the city’s landscape, and endless Amazon trucks skirt from building to building, along with GrubHub cars and pizzas stacked in insulated carriers on the back of scooters. The city is in constant movement and change, and yet there are parts that seem durable over time, being made again and again. Different languages are spoken and written in different regions of the city and are often juxtaposed on digital monitors and posters throughout. Every day, people flow into the city and out—people who are classed, raced, and gendered in the experiences of their homes and perhaps differently in their interactions in the city. During the day, some of these people go to the same spots at lunch or after work—parks, bars, coffee shops. These spots become familiar and feel like places of respite. At dusk the city changes again, and yet again at night.

Over a decade ago, we critiqued the “classroom-as-container” model of conceiving of learning (and schooling) and proposed instead the constructs of “learning-in-place,” “learning trajectories,” and “learning networks” (Leander et al., 2010). Re-imagining the classroom as a city moves us into considering places (e.g., the bar), trajectories (e.g., commuting) and networks (e.g., bus routes). Perhaps more importantly, such a transition in image and imagination helps us to consider the constant movements and energy of making and remaking social space, and also consider the intensive heterogeneity of the actors who contribute to this ongoing production of space—vehicles, humans, materials of all kinds, and of course, texts and our practices with them.

One of the challenges of relating literacy studies to theories of social space, and especially to those “post” or critical of modernity, is that, like any field, literacy theory does not float free from its own history and practices, including research methods. The ethnographic turn in literacy studies (e.g., Heath, 1983; Street, 1985) brought with it a powerful focus on the socially situated meanings and uses of literacy, which was powerfully marshalled, for instance, into the development of the New Literacy Studies with its critique of an “autonomous” view of literacy (Street, 1987). Still, nearly four decades after the ethnographic turn, the field of literacy studies is still uncertain how to formulate an anthropology of postmodernity. How do the assumptions of ethnography hold up for postmodernity and its entangled, flowing social spaces? How are assumptions about the distinctions between local and global, bounded views of particular “settings” and “communities,” and the work of the lonely ethnographer, embedded in her focal “research site,” constraining for an understanding of what literacy is and what it does? Can we have a literacy studies that is at once “ideological” (Street, 1987)—entirely infused with critical considerations of power in discourse—and yet also not parochial or locally mundane?

Another present challenge of more fully realizing the potentials of critical, postmodern spatial theory in literacy studies is the fascination or even fetishization of method within educational research (St. Pierre, 2011). In the case of spatial theory, the attachment to method can run two ways—either in a search to create a settled “spatial methodology” for the field, or in the tendency to not fully engage spatial theory because such a methodology is not readily evident. These prevailing modernist tendencies can envelop a more liberating engagement with the concepts and theory in their own moves to unsettle structures, codes, and singular entities. As postmodern spatial theory is relationally conceived, and as it engages the multiple and shifting relations between people,

discourses, and things, the desires to code, fix identities, stand outside of, and create a priori boundaries as embodied methodological practices in qualitative research constrain how we might embrace the complex relationalities of social space in its reproduction and transformative becoming.

Yet another challenge for a more complete appropriation of the “spatial turn” in the social sciences and humanities (Soja, 1996) has been the centrality of a discourse-based ethos within the critical corners of literacy studies. It may seem ironic to claim that discourse—with its roots in sociolinguistics and critical social theory—could be made too central within the domain of language and literacy studies. Still, for spatial theory, a key question concerns how the material, embodied world comes into contact with discourse and is shaped by discourse, without reducing the material and embodied world to texts or sign systems. A related challenge is how we might think spatially about the material-discursive world without reducing social spaces to metaphor (another discursive framing), since the field of critical social science is already rife with spatial metaphors (e.g., center/periphery, location, place, boundary, marginalization, intersection). Studies of “third space,” for instance, while drawing richly from critical and post-colonial theory (e.g., Bhabha, 1994) can sometimes tend to treat social spaces metaphorically or as entirely discursively produced, even while taking up questions of bodies, identities, and power. In contrast, Hirst (2004) offers a rich analysis of hybrid/third space, describing the power relations in a second language Australian classroom where the teacher was an Indonesian national. That material spaces, their design, and bodies participate in the production of discourse is not in question. The question concerns how bodies, the material world, and material spaces exceed our means of interpreting sign systems in their capacities to enter into socio-material relations. How is materiality not like discourse, and yet enters into relationship with it in the production of social space? For literacy scholars, how are the meanings of signs and texts, as well as their affects, entangled with material bodies of different sorts? These are pressing and important theoretical questions for spatial theory in its development within the field, some of which are being currently considered in adjacent fields such as affect theory and studies of new materialism.

Some key contributions to literacy studies of spatial theory

Despite the challenges and limitations considered thus far, as well as others, spatial theories have made important contributions to literacy research. These contributions are still changing and growing. Without offering a complete review, we would like to synthesize a few types of contributions that have influenced the field and of course our own work as part of it.

First, spatial theory has given new insights into a reconsideration of boundaries within literacy research, with importance for methodology, but also for theoretical questions such as the distributions and movements of identity and social practice. Because critical and postmodern spatial theories begin from the premise that boundaries are produced and reproduced in social-spatial life, rather than epiphenomenal to social life, then the question of boundary-making is brought to the fore. Contextual constructs such as “out of school” come into question in this modality and may shift to become considerations of how the boundaries of “school” and “out of school” are actively made, disciplined, and disrupted. In this sensibility, boundaries become perhaps more important, as the spatial shapes the social and the social shapes the spatial but are considered as actively produced rather than as a backdrop for something focal (e.g., the “literacy event”). Issues relevant to boundary (re)production, transformation, and the hybridization of overlapping boundaries within literacy research include questions about identity, about contexts for literacy, and about our broad sensibilities concerning the “virtual world” or “online world” and its relations to the “offline world” or “meat space.” A starting point for social spatial theory is not to begin with what seem to be naturalized or given boundaries, but to begin with an analysis of the forces, materials, and processes that produce boundaries and/or disrupt them.

For critical literacy studies, the spatial analysis of boundary-making can offer a significant shift of thinking beyond the individual and her experience and toward entanglements with the social world, all while not resorting to macro structural analyses. Ultimately, for critical theory, such analyses may serve to produce means of opening up and transforming boundaries of identity, literacy, and learning, when these boundaries themselves are critically appropriated as more than mere metaphor. Sheehy (2004) describes the tensions that were created in a seventh-grade classroom as students engaged in political action relevant to their proposed school closure, examining how their activity traversed boundaries and created “in-between spaces.” Dixon (2011), drawing on Foucault’s approach to disciplining space-time, demonstrated how gaze and the organization of school space were accomplished for power/knowledge in post-Apartheid South Africa. In this case, the relations between national surveillance, oppression, and spatial enclosure and the local geography and practices of the school are accomplished with a chilling synchronicity. Finally, Nespor’s (2013) incredible analysis of spatial relations in an across an urban elementary school, based on 2 years of ethnographic fieldwork, investigates, among many other issues, the school’s collaboration with business enterprises toward refashioning classrooms into business enterprises, the children’s efforts to make sense of the scheme. Nespor relentlessly opens up the boundaries of “the classroom” to examine webs of social-spatial relations involving school division, regional economics, parental concerns, popular cultures, racial politics, urban development efforts, and other forces.

Second, literacy scholars have used spatial theory broadly and theories of place more precisely to engage in research and development with liberatory, humanizing aims. This work takes seriously the qualities of different kinds of places and how places shape experiences. Sometimes inspired and informed by phenomenology, place-inspired research is also often focused on the capacities of educators and children to transform school or institutional spaces into places that matter and that offer growth. Some place-oriented research has actively engaged children in processes of textual and material place-production. For example, an ongoing project by Comber and colleagues (“Urban Renewal from the Inside Out”) has engaged children with university students from the fields of architecture, communications, and literacy students to redesign an area of the school grounds in a poor suburb of

Adelaide, South Australia (Comber and Nixon, 2008). These forms of local spatial transformation are a means of creating a critical literacy pedagogy that is more fully embodied and spatially infused. Because of its engagement in the nature of experience, place-inspired research provides a natural bridge to more recent work in effect theory and literacy (Boldt, 2021; Hollett and Ehret, 2017; Niccolini, 2019), which is concerned with the different kinds of affects produced when texts and bodies of different sorts come into contact or relation with one another.

A third key contribution we wish to highlight concerns an understanding of mobilities in literacy research as conceived through spatial theory. As discussed in the opening regarding the expansion of ethnography within the postmodern, toward something like “connected ethnography” (Hine, 2000; Leander and McKim, 2003) or another formulation, mobilities studies are given to thinking about practices, texts, bodies, and technologies in their circulations. Such circulations may include a given “situation” or context, but situatedness is often considered as a temporary arrangement or assemblage. Mobilities research in literacy has drawn variously on critical spatial theories and also on social theories and philosophy taken up by spatial theorists, including Actor Network Theory (Brandt and Clinton, 2002; Leander and Lovvorn, 2006), Deleuzian rhizome theory (Eakle, 2007; Hagood, 2004; Leander and Boldt, 2013), embodied media composition (Ehret and Hollett, 2014; Vasudevan, 2010) and rhythm-analysis (Leander and Hollett, 2017). In an important contribution to mobilities research in literacy studies, for example, Brandt and Clinton (2002) provoke: “might something be lost when we ascribe to local contexts responses to pressures that originate in distant decisions ...?” (p. 338). Drawing on Actor Network Theory, the researchers reconsider Besnier’s (1995) study of the on the geographically isolated, subsistence-based Nukulaelae people, who would wear Western-style tee-shirts “bearing risqué slogans in English” (p. 343). The slogans did not enter into the textual/meaning-making practices of the Nukulaelae, and yet the tee-shirts did. As a result of their reconsideration, Brandt and Clinton (2002) expand a mobilities mindset, invoking a “thing like status” to literacy to consider how the material dimensions of literacy travel across time and space, integrate, and endure. In this way, the global and the local become situated in movement and practice rather than being captured in representations or static backdrops.

Reading across the key contributions described above, and considering others too numerous to discuss presently, it would be difficult to describe the many possible expansions and revisions of literacy studies in their contact with spatial theories. However, one insight that we have gleaned from this effort is that the value of spatial theories will at least in part be determined by their capacity to contribute interpretive texture, “layers,” nuance and insight to the qualities that are already made present in (critical) qualitative research. Inspired by such textured, layered approaches, and in order to further illustrate an analysis of social space and schooling, we turn now to an examination of social space in the recent history of schooling during the Covid-19 pandemic.

Popular discourses of school space and Covid-19

Our approach in the following is to consider the social spaces produced in and around “schooling” during Covid as fully relational, and as produced through conceptions or intentional designs of space, perceptions or discourses of space, and lived experiences of space (Lefebvre, 1991; Soja, 1989, 1996). Within “lived experiences,” we are especially interested in the affective dimensions of the experience of living in and through the period of Covid-19, and “doing school” during this period. The media we draw on in order to record and reflect on the productions of space during Covid-19 were published popular press and the social media network TikTok from March 2020 to June 2021.

We have chosen these media because they illustrate experiences of being a teachers, students, and parents during Covid-19, and also reflect discourses around remote schooling that developed at this time. From the popular press, we searched out profiles, editorials, and interviews that depicted or addressed the experiences of remote schooling. While completing this survey, we noticed that while teachers, and parents were well represented, the perspectives of students were harder to find. When the experiences of students were depicted, they were most often mediated through an adult interviewer or journalist and published for a presumably primarily-adult audience. We therefore widened our search to include TikTok where we knew many students—as well as parents and teachers—had participated in discourses on schooling during Covid-19. We coded these media and selected telling examples of expansions the boundaries of school, spatial reproduction and simulation, embodied experience and the effects of space, and critiques, resistance, and transformations of school space.

In the following, we describe three different “variations” of social space that arise out of our analysis of these media. We intend these variations as different productions of social space that are reflected in the media we have studied, as well as partially produced through them. For each variation, we consider some of the relational dimensions of social space, including how the embodied, material world, discourses, power relations, agency and affect are interconnected and co-constructed in the production of social space. We do not consider this analysis nor these data to be a “complete” rendering of social spatial relations of schooling during Covid-19. Rather, our goal is to use them to illustrate the dynamics of social space, and the uses of spatial analysis to consider how experience, space, literacy, and discourse are co-produced. We also aim to make evident how qualitatively different experiences of school (and literacy) are spatially produced—that space not only “matters” but is also co-constituted with unique, distinct qualities of experience.

In the first variation, we detail productions of school space during Covid-19. We trace how school boundaries were expanded as classrooms were physically and affectively reproduced in homes. Next, we look at how conceptions of school, power relations, teacher decisions, and technology design interacted to simulate hyperreal classroom spaces online. We then analyze discourse around cameras to uncover ways student presence in these spaces was produced, measured, and disciplined.

The second variation traces the affective shortfalls of school space in remote learning. We examine accounts of mismatches between the hyperreal simulations of school online and embodied experiences being students, teachers, and parents during remote schooling. We consider affective impact these mismatches had on people subjected to them. We then turn to TikTok posts in this variation to illustrate experiences of remote schooling and its shortfalls, and also to trace how discourse on the platform helped shape collective understandings.

In the final variation, we focus on manifestations of student agency. We look to TikTok for collectively authored critiques of school space during Covid-19. We also examine subversions of school boundaries and collective acts of resistance on the platform. Finally, we detail accounts of students and teachers forming empathetic relations and collaboratively reshaping their remote school, and consider if these episodes point to possibility for the production of more humane school space.

Variation 1: physical, affective, and hyperreal productions of school space during Covid-19

The physical and affective reproduction of school space at home

In media we surveyed, teachers sought to replicate elements of their classroom by gathering school materials and rearranging furniture (Hadden, 2020; Hackman, 2020; Dominus, 2021). These redesigns of teacher's homes seem to have functioned not just as physical, but affective reproductions of school space. A chronicle of a day in the life of Matthew Bordello, a second-grade teacher working remotely illustrated this process (Hadden, 2020). Bordello's conception of his class had been deeply tied to physical space. "Our community as a classroom is very much on being together and being able to share things with each other and build off together in a group" he argued. The transition to remote learning could have devastated this community and Bordello's self-efficacy as a teacher. However, by rearranging his bedroom he recreated some of the routines and feelings of this in-person community. He made space for art supplies, books, and hand drawn copies of the "classroom agreement posters" that hung on the classroom walls. While he his workspace/bedroom "might seem chaotic to others" it was the "ideal environment for keeping his class focused" which let him "recreate a sense of organized chaos in his virtual classroom." For Bordello, and he believed, his class, feelings of order were restored.

Students and families reproduced school spaces at home as well. One P-12 private school, displayed this work in a series of interviews with remote students of different ages (In their own words, 2020). The interviews were punctuated with photographs (see Fig. 1) of the subjects sitting at large desks in bright bedrooms with stacks of books, notepads, and a laptop or tablet around them. While taken in different houses, they share striking physical and qualitative uniformity, seemingly curated by the same schooled expectations. The consistency of the images parallels the interviews themselves. Each student expressed some melancholy about isolation, but they mostly described their—quite similar—homeschool routines and conveyed of self-confidence and comfort. These accounts imply that through intentional redesign of home space and activity physically and affectively expanded the boundaries of school space into the home.

Hyperreal school space on videoconference platforms

The social space of schooling was not simply relocated during Covid-19. Rather, we argue, the displacement and re-placement of school qualitatively changed its space. In the media we reviewed, we can trace how simulations of school—or perhaps more accurately, of conceptions of school—produced hyperreal school spaces. Baudrillard (1994) describes the hyperreal as "the simulation of something which never really existed." It is the representation or reenactment of something which had previously existed only as an abstraction. Baudrillard points to the depiction of medieval Europe in Disney World, as an example—it doesn't simulate any place that existed in history, but a preexisting, essentialized conception. Similarly, remote schooling arguably simulated conceptions of school rather than actual school spaces. Certain "school-like" qualities, such as teacher-control and focus on attention and evaluation were more intense than they had—or ever could have—been in physical classrooms. Across the media we surveyed, conceptions of school, teacher actions, and the design of videoconferencing platforms could be see coproducing these hyperreal school spaces.



Fig. 1 Screenshots of photos in a school online publication of students working from home. Photos taken from a school publication of a twelfth and fifth grade student of working from their homes.

Bordello turned to various Zoom features to simulate his second-grade classroom (Hadden, 2020). He used breakout rooms for different ends, including sequestering students “when their distractions affect[ed] others”—a move reminiscent of timeouts. The hallway or corner where students Bordello had sent students for timeouts was thus simulated on Zoom. In some ways however, these breakout rooms were more “timeout-like” spaces than could have existed in a school. They were more isolated. A child could hear and see her classmates while sitting in a corner, but she would have no access to the Zoom meeting while in a breakout room. Teacher control was also heightened. In person, a student could resist timeout and still affect the rest of the class by making a ruckus or walking back to the group, while on Zoom, she had no recourse. The timeouts enacted in breakout rooms were in these ways hyperreal.

Breakout rooms seem to have been made into hyperreal instructional spaces as well. English professor Nguyen (2021a) argued this feature not only allowed the replication of classroom discussions, but helped make them “better than those of the in-person era.” When teaching in person, he prompted students mid-lecture to gather in small group for discussions. Teaching remotely, he put them into breakout rooms. He saw the latter structure as better suited to his goals because it was “simple to arrange and lack[ed] the cacophony” of live discussion. As with the virtual timeout, distractions that could accompany proximity were stripped away. Students no longer had to move or look around to form groups—transitions were streamlined to pushes of buttons. While buzz and bustle of other groups had been present in the classroom, on Zoom students could only see and hear their groupmates. In some ways, these breakout rooms were therefore distilled for focus on instructional activity beyond any classroom.

In a few of the cases we reviewed, teachers’ changes to videoconference settings helped produce essentialized school space. In a newspaper profile (Hackman, 2020), a sixth grade teacher described remote teaching as “all the same stuff I [did] in class ... I just needed time to figure out how to translate those skills virtually.” One way she did this was by “eliminat[ing] the class chat function, allowing to only send messages to her.” This choice constrained student discourse in a “school-like” manner, which bore striking resemblance to the Initiate-Respond-Evaluate (IRE) model that has long dominated US schools (Nystrand, 1997; Wilkinson et al., 2015). Following an IRE model, student utterances are directed, heard, and evaluated by a teacher. No matter how rigidly their classroom followed this structure though, students could circumvent expectations and surreptitiously communicate without the teacher’s notice by passing notes or sharing whispers. In this virtual classroom however, all communication was witnessed by the teacher and subject to evaluation. The videoconference was not a replication of an in-person class, but a hyperreal “school-like” simulation.

Discourse on cameras and student presence in hyperreal school space

Perhaps the most widely written about videoconference feature during the pandemic was the camera. Arguments for and against mandating students to keep their cameras came up repeatedly in our survey popular press (Hackman, 2020; Nguyen, 2021b). This discourse revealed contested understandings of what it meant to be “present” in a remote lesson. By examining this discourse, we can uncover ways in which student presence was produced and disciplined in the hyperreal school space of remote learning by technology, school policy, and student and teacher actions.

Mandates around cameras seem to have been widespread during Covid-19. In a survey of American K-12 educators, 77% stated their students “must” keep their cameras on (Will, 2020). Many of the arguments for these requirements centered on accountability and attendance. Some teachers maintained live video helped them keep the class engaged because they could hold students accountable (Hackman, 2020; Burney, 2020; Will, 2020). One assistant superintendent bemoaned that before her district’s camera mandate, students logged into class to be counted as “present,” right before “turning off the camera and walking away” (Burney, 2020). After the mandate, she claimed her district achieved “95% participation.” What she meant by “participation” was not clear, but a teacher from her district revealed in the same article that she gave participation grades based on who kept their cameras on “just like when I [took] attendance each period at school.” According to her school’s measures of accountability, grades, this teacher’s students participated in her lessons simply by visibly attending them.

The understandings of student presence in these arguments were notably thin. Discussion, collaboration, creation, and reflection were all absent. Many students undoubtedly did all of these activities in remote lessons, with or without their cameras on, but they could only be counted—or held accountable—as present if their teachers could see them. Presence was thus passive and arguably deadened. This dynamic echoed critiques forwarded by Freire (1972), Davies and Bansel (2007), Nespor (2013) and others that schooling in capitalist and neoliberal societies understand students as objects to be acted upon. In classrooms, students were often not expected to act but to receive instruction and present themselves for accountability. In remote lessons, this dynamic intensified. Physical attendance in which students had sat in classroom and regulated their bodies was simulated by videos of their faces—ignoring their embodied and affective experiences. Student subjectivity, long overlooked in school spaces, seems all but irrelevant to student presence in these hyperreal classroom simulations.

Variation 2: affective shortfalls of hyperreal school spaces and remote learning

Anxiety and the mismatch between the presentation and experience of remote school

Attempts to simulate school space during remote learning appear to have taken affective tolls on some teachers, students, and parents. Many accounts of remote learning detailed disconnect between how teachers experienced teaching during Covid and how they presented—or hoped to present—themselves. Some teachers felt anxiety and a loss self-efficacy when navigating this mismatch. In a longform article tracking the year of an AP World History class, the teacher, Ms. Everett-Kennedy, described school building was a “community-ordained safe space” and her classroom a “locus of control” that was a “near sacred space in which she

excelled" (Dominus, 2021). Extracted from these spaces, she felt these senses of order, community, and self-efficacy disappear. She attempted to regain them in part simulation. She tried to appear energetic on Zoom, as she had been in person, and to keep clutter hidden from the camera so she could project an image of order and control. But struggles over the course of the pandemic prevented her from maintaining this image. As one student described her, Everett-Kennedy often came to zoom with her "hair in that messy ponytail, tired eyes, that bottle of Diet Coke always beside her." While this student found the presentation relatable, it fell short of Everett-Kennedy's self-expectations. Frustrated by her inability to support and connect with students as she had in previous years, and regain control over her life and image Everett-Kennedy went through bouts of anxiety and disappointment, saying she expended "more and more energy to fight off her sense of failure."

Like their teachers, some students felt anxious about presenting themselves in videoconferences. They carefully considered what was displayed. Three of Everett-Kennedy's students were described angling their laptops to hide messes in their rooms (Dominus, 2021). Anxiety about self-presentation was particularly pronounced in one student, Charles. According to the profile, Charles was a straight-A student who was "used to entertaining his classmates but also impressing his teachers." As a high-achieving Black teenager he had felt pressure to live up to the high expectations of his mother while confronting a racist society and "proving these stereotypes wrong," but he felt he was on track to live up to these goals. However, when the schoolhouse closed for the pandemic, Charles began struggling with mental health and fell behind on school work. He was determined to hide these challenges. Always a "skilled performer," he became not only an actor, but a set-designer to simulate the role he had filled before the pandemic. He set up "his personal stage [in] his bedroom, where he'd hung LED lights and a tapestry with a design that evoked some graphic novel's envisioning of another planet." The journalist notes that "from that room, Charles projected the image of young man wildly amused and amusing." This projection was hyperreal, however—it showed not who Charles actual identity or experience, but how he believed he should be. The pressure to uphold this image was a drain on him. As his affect and experience worsened, this anxiety grew more intense, as did his evasions from support. He was ensnared in a devastating cycle. By the winter of 2020, Charles was convinced "failure was his destiny."

Missing connections and affective depletion in remote school

Students and teachers in the media we surveyed repeatedly described videoconferences as not-quite approximating classes in person. They often foregrounded what was missing—physical proximity to others and associated feelings of energy. This sense that remote learning, as a simulation of school space, was missing key elements it *should* have imbued accounts with effects of disconnection, longing, and depletion.

Teacher-student conversations over videoconference were often characterized as strange or awkward (Dominus, 2021; Blum, 2020; Nguyen, 2021b; Osman, 2020; Hackman, 2020). Forming relationships became more challenging. Teachers lamented talking to "a sea of black squares" (Nguyen, 2021b). Students were more reluctant to be vulnerable—"it's hard to believe someone you only met through a screen" one high schooler explained (Dominus, 2021). Physical distance weakened the affective connection between student and teacher. Both teachers and students described shifts in group affects as well. Class discussion was subdued (Nguyen, 2021b; Blum, 2020) and students had felt energized by in-person gatherings, felt drained after videoconference lessons (In their own words, 2020; Dominus, 2021).

Many of these accounts related feelings of depletion with strained connections between classmates and teachers and students and their lack of physical proximity. Anthropology professor Blum (2020) theorized this dynamic in more detail. When considering why she felt exhausted after teaching on Zoom, she concluded "it is because videoconferencing is *nearly* a replication of face-to-face interaction but not quite [that] it depletes our energy." While Zoom allowed her class to see one another's faces and hear their voices, Blum argued it did not enable "harmonized posture," "coordinated breath" or eye gazes that "communicate information." Humans, she claimed, were "delicately attuned to each other's complete presence" and so they read these nonverbal cues to support conversation in person. In many ways, videoconference made presence thinner—for example, one could not tell where another was gazing. While these changes were subtle, they were also meaningful. They left conversations halting and labored which Blum argued, which led to feelings of frustration, depletion, and a longing for more free-flowing human connection.

Techno-social discourses of student affect and mental health during Covid-19

By surveying TikTok clips with hashtags related to remote learning, we traced lines of student discourse around remote learning, affect, and mental health. Thousands of posters, including many middle school, high school, and college students uploaded clips documenting their experiences during the pandemic. The design of TikTok, where users could easily edit and repurpose audio and video from existing posts, enabled the emergence of multiple lines of discourse where thousands of students watched, commented and posted response videos to one another's clips. Though physically isolated, students were able to collaboratively publish and interact with techno-social discourses around remote learning and its effects on them.

One particularly striking motif that was seen across many TikTok pound we found was panic attacks. One TikTok which was liked over two million times showed an adolescent girl crying and breathing hard while her parents look at a laptop. Text came up over the video that read "the chem final my sister has been taking for 3 h didn't save" while pre-existing audio played of a woman repeating "just breathe." Even with her family in person, trying to help address the issue, inadequate technology and the demands of remote schooling overwhelmed this student. Many viewers could relate—the most-liked reply read "the crying and pacing while parents attempt to fix it resonates with me."

Increasing lists of assignments and problems come up in many videos. One popular post, pictured below, was representative. The camera faced an adolescent laying **their** head on a desk. Audio taken from another TikTok played where a robotic voice rapidly

listed assignments, tasks like “clean your room” and “take a shower,” invectives such as “work until it hurts” and questions such as “does he still love you?” and “what happened to you?” Words such as “due at 11:59” and “missing assignments: 35” flashed on screen. Not quite obscured, the adolescent’s head slowly sank, as though pushed down by the words above them (see Fig. 2). Clips such as these evocatively illustrated their feelings of depletion, and at times, depression. Like the teen shown in this TikTok, many adolescents appeared to fall further into hopelessness and isolation as the demands of remote learning cascaded upon them.

Notably though, the students who posted and viewed these TikToks were not entirely isolated. While physically distanced, the views, likes, replies and response videos demonstrate they saw and heard one another. Their techno-social discourse allowed remote students to share in, and, to an extent, shape their affective experience during the pandemic.

Variation 3: critiquing, resisting, and transforming school space

TikTok critiques of the production and boundaries hyper-school

From what we surveyed, TikTok posters didn’t only aim to document remote schooling, but to spoof it. Many popular videos drew humor from unbridgeable gaps between the experiences of remote schooling and the images the uploaders had hoped to present. While some of this comedy was self-deprecating, it was often pointed toward the absurdity of remote schooling and the demands it put upon them without regard to their conditions and experiences in reality. These clips may therefore shed insight not just into how students, teachers, and parents experienced the emergence of hyperreal school space of remote learning, but how they critiqued it.

A recurring theme was sights and sounds accidentally recorded during virtual lessons. In one clip, a woman sat on a rug and explained “I’m sitting on the floor, next to my children’s desks ready to mute and unmute their Zoom class,” out of view from their Chromebook cameras. She said with a wide smile that while her children knew how to unmute themselves, they often “do it at the wrong times and everyone can hear me screaming in the background.” She thus subverted and mocked the expanded boundaries of school space into the home. Remote schooling compelled her to simulate scenes of home-school life that, implicit in her telling, can only be created through extreme and surreptitious measures. What’s more, even if she succeeded in curating a scene for Zoom school, her family’s *actual* home-school would fall short of the presumed standards. If everything was going well, after all, why would she be screaming? She could not meet the expectations she assumes the teacher and/or other families have for remote learning, so she had to hide, crouching on the floor for entire lessons. The demands of simulating homeschool appeared harmful and absurd.

Many TikToks posted by students documented their struggles with staying, or appearing to stay, present during virtual lessons. One widely-viewed post panned in on a square in a grid on a Zoom screen where a cat sat alone staring at the camera while a man lectured monotonously in the background. The caption read, “I forgot to turn my camera off when I went for a walk and my friend sent me this she was there for like 15 min #catsoftiktok #zoomschool #fyp.” This text implied embarrassment not for walking away from the lesson, but for leaving the camera on to capture this humorous tableau. The poster seems to have taken it for granted that the expectation that to sit still at her computer for the entire lecture was unreasonable, or at least unimportant, but *showing* she left was a breach of social norms. Several comments, where users shared fears that this could happen to them, pointed to the same conclusion. It was impossible to meet the implied expectations of this new school space, so they instead hoped to conceal themselves. The possibility of revealing their activity and movement appears to have filled many students with anxiety and fear. By

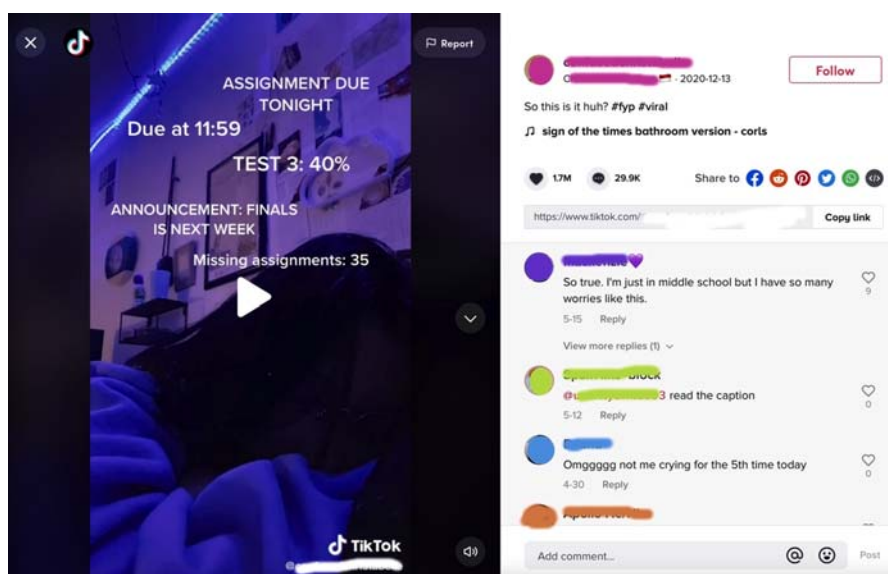


Fig. 2 Screenshot of a remote student’s TikTok post and user’s comments.

sharing these experiences and affects, these TikToks implicitly critiqued remote learning. Could such empty simulations really be worth this stress?

Resistance through play: student challenges to school space on TikTok

TikTok posts not only revealed critiques of the social space of remote school, they propagated challenges to it. Students posted hundreds of clips that documented pranks and hacks that altered remote learning spaces. Many clips appeared explicitly designed to instruct others how to replicate these tricks. They were often saturated with referential humor that evoked aspects of youth internet culture and seemed to invite viewers to respond. The techno-social discourse of remote student TikTok thus seems to have cultivated acts of resistance to remote learning that were collectively-designed by students spread across the world and physically isolated.

Some of the most-liked clips we found of this genre were pranks where students subverted expectations of their remote lesson space by collaborating and manipulating the technology. For instance, one clip, captioned, “April Fools on our professor #aprilfools #fyp #zoomschool,” showed a young woman recording and laughing at a zoom screen where a professor mumbled “what’s going on ...” as over a dozen users logged into the meeting with the same default photo of his face and name. The students disrupted their remote class by wielding the technology and their professor’s own name and image against him. Other pranks were less reliant on videoconference features. Several clips showed teachers interrupted as their class suddenly breaking into synchronized dance routines across the grid squares on the screen. These pranks were likely shocking not just because they were unexpected, but because they clearly implied the class communicated and coordinated without the teacher’s knowledge. While students were highly surveyed in many remote lessons, these pranks demonstrated they could still surreptitiously collaborate, plan, and play outside of them. Posts tagged with #schoolhacks or #collegehacks, or #studhacks were. Some of these videos offered advice on study habits and using remote learning technology. Others demonstrated ways to surreptitiously circumvent teacher’s rules and expectations. At the time of our survey, the two most liked TikToks with the tag #onlineschool showed how to use software that hacked google-form quizzes, auto-filling all of the correct answers. Similar videos demonstrated the steps to photoshop pictures of homework to send to teachers, trick mouse-tracking programs, and set looped videos as their Zoom default picture to make it appear cameras were on.

These TikTok clips revealed heteroglossic and highly participatory discourses. Comments shared stories similar to the videos. Clips referenced one another by incorporating each other’s audio or video. They often referred to other elements of online youth culture by layering sound from viral social media clips, and working in humorous images from memes (video of a bald man vigorously eating a sandwich shows up in many of these TikToks without explanation, for instance). Through this discourse, it appears that students were able to remotely form social spaces apart from remote school, where they could then—to an extent—collaborate to circumvent and resist the boundaries and expectations of school space.

Techno-social transformation of school space

While much of the accounts of remote learning focused on struggle, there were also glimpses of how school spaces might be made more humanizing and collaboratively-designed. In some instances, teachers and students appear to have developed understanding and empathy for one another as video conferences bridged their home environments and lives. One student-teacher likened this dynamic to “a throwback to a time when teachers made home visits” (Arthur, 2021). Some teachers considered equity in a new light, realizing that some of their students homelives made schoolwork more challenging—both when school was remote and in person (Osman, 2020; Will, 2020). Students too were described as reconsidering their teachers’ perspectives after seeing and hearing their struggles over zoom (Dominus, 2021; Nguyen, 2021b). These accounts suggest that as remote schooling expanded the borders of school space into the home, it also offered at least some teachers and students more insight into one another’s lives.

On TikTok, audiences seemed fairly segregated between students, teachers, and parents. Most comments on student videos appeared to made by other students and parallel patterns held with teacher and parent posts. There was, however, also discourse between these groups. Some popular teacher posts were duets with student posts. At the time of our survey, one such duet, pictured below (see Fig. 3), was most-liked post tagged with #remoteteaching. On the right, a student’s TikTok played, illustrating the overwhelming demands of remote learning. On the left, a teacher sat silently, appearing to empathetically react to the video. This post and ones like it showed meeting of space and experience. By apposing these self-authored, affectively-imbued videos the poster not only joined the environments of her and the student, but couples their perspectives. This TikTok thus showed a space co-written by a student and teacher and reflective not just of their respective roles, but their human experiences.

In the popular press there were a few examples of elements of remote school space being similarly collaboratively and humanely produced. One middle school teacher (Osman, 2020), how her student’s use of chat helped shape the discourse and content of her zoom lessons. Chat enabled multilinear conversation, as students could simultaneously talk over speaker, type to the class in the chat, and send private messages. It allowed multimodal discourse, as students used words, acronyms, emojis, and pasted images and video-links. The teachers admitted, “I am still learning this particular language they speak in, but [my students] are quite happy to translate.” Sometimes the messages were connected to curricular content, sometimes they weren’t. The teacher explained she followed deviations from her lesson plans more than she would have in person because she sensed her students longed for the human connection it provided. A few pieces described classes collaboratively establishing norms around cameras in remote lessons. Several teachers reported telling their students to turn their cameras off or on whenever they felt needed (Will, 2020; Burney, 2020; Nguyen, 2021b; Osman, 2020; Dominus, 2021). One professor (Nguyen, 2021b) detailed she opened up to her students about her struggles with teaching to blank screen, and the class then discussed balancing their needs for privacy with their classmates and teachers desire for connection as factors to consider choosing when to turn of their cameras.

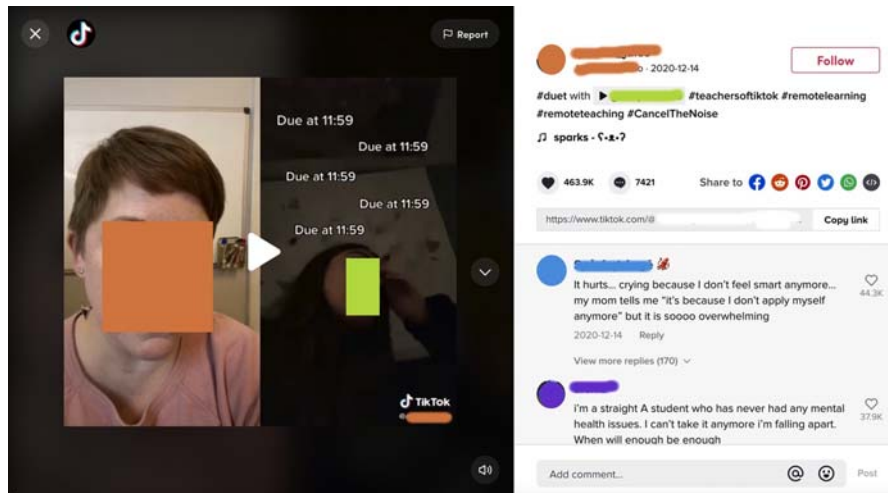


Fig. 3 A screen shot of a teacher's duet on TikTok with a student's post.

In these examples, teachers, students, and technology co-produced school spaces that were in some ways more responsive and humane than many in-person classrooms. Chat enabled new modes of communication to come into school space, many of which were more familiar to students than teachers, and closer connected to their out-of-school lives. When choosing to turn on or off their cameras, students could determine how to be present, seen, and available for in a manner unreplaceable in person. School space during Covid-19 was therefore not only reproduced, resisted, and made hyperreal. It was also transformed into spaces that could, at times, allow for collaborative and human connection.

Final thoughts

Although spatial theory to re-see approaches otherwise more familiar in qualitative literacy research, such as textual coding, observation, or discourse analysis, ultimately the values of the theory will largely depend on how it is used to bring to life the particular meanings and effects of literacy practices, texts, and discourses in their emergent relations to bodies and the material world. From the broader push of the earlier work on social space—that space matters, and that social spaces are everywhere produced—one important shift involves asking how to understand and feel the particular qualities of different types of spaces and places involving literacy and literacy learning. Our analysis of variations of social space during Covid-19 is an attempt to bring to life these qualities and affective dimensions of spatialized experience.

Among Doreen Massey's far-reaching contributions to spatial theory, she proposed three maxims for understanding space: first, as a product of interrelations, themselves constituted through interaction; secondly, as a dynamic of multiplicity in which distinct trajectories and heterogeneity co-exist; and thirdly, as always in the process of being made (Massey, 2005). Massey was fond of claiming that social space was the accumulation of "stories told so far," or an "open and ongoing production." Social space is not outside of our making of it, outside of our discourses on it, of our bodies moving through it. The openness of social space is, in the end, also much of its significant potential for social theory and for literacy studies more specifically. Spatial approaches are not merely a more robust or geographical method to "map" existing power structures; rather they open up for us how the movements of power—including those more liberating—shift and may be re-storied over time.

References

- Arthur, K., 2021. Zooming In: What It's Like to Student Teach During a Pandemic. Illinois State News. Retrieved from. www.news.illinoisstate.edu/2021/01/zooming-in-what-its-like-to-student-teach-during-a-pandemic/. (Accessed 14 April 2021).
- Baudrillard, J., 1994. In: Glaser, S. (Ed.), *Simulacra and Simulation*. University of Michigan Press.
- Besnier, N., 1995. *Literacy, Emotion, and Authority: Reading and Writing on a Polynesian Atoll*. Cambridge University Press, New York.
- Bhabha, H., 1994. *The Location of Culture*. Routledge, London.
- Blum, S.D., 2020. A Professor Explores Why Zoom Classes Deplete Her Energy (Opinion). Inside Higher ed.
- Boldt, G., 2021. Theorizing vitality in the literacy classroom. *Read. Res. Q.* 56 (2), 207–221.
- Brandt, D., Clinton, K., 2002. Limits of the local: expanding perspectives on literacy as a social practice. *J. Lit. Res.* 34 (3), 337–356.
- Burney, M., 2020. Attending Virtual School: Should Cameras Be On or Off for Remote Learners? The Philadelphia Inquirer. <https://www.inquirer.com/news/nj-education-remote-learning-cameras-virtual-classes-coronavirus-20201210.html>. (Accessed 7 July 2021).
- Comber, B., Nixon, H., 2008. Spatial literacies, design texts, and emergent pedagogies in purposeful literacy curriculum. *Pedagogies* 3 (4), 221–240.
- Davies, B., Bansel, P., 2007. Neoliberalism and education. *Int. J. Qual. Stud. Educ.* 20 (3), 247–259.
- Dixon, K., 2011. *Literacy, Power, and the Schooled Body: Learning in Time and Space*. Routledge, New York and London.

- Dominus, S., 2021. "I feel like I'm just drowning": surviving sophomore year in 2020. N. Y. Times. <https://www.nytimes.com/2021/05/13/magazine/high-school-students-coronavirus-pandemic.html>.
- Eakle, J., 2007. Literacy spaces of a Christian faith-based school. *Read. Res. Q.* 42, 472–510.
- Ehret, C., Hollett, T., 2014. Embodied composition in real virtualities: adolescents' literacy practices and felt experiences moving with digital, mobile devices in school. *Res. Teach. Engl.* 428–452.
- Freire, P., 1972. *Pedagogy of the Oppressed*. Herder and Herder, New York.
- Hackman, M., 2020. What it's like to teach sixth-grade science over zoom. *Wall St. J.* <https://www.wsj.com/articles/what-its-like-to-teach-sixth-grade-science-over-zoom-11586774461>. (Accessed 14 April 2021).
- Hadden, J., 2020. An Elementary School Teacher in San Francisco Gives an Inside Look Into Life Inside His Virtual Classroom Amid the Coronavirus Pandemic. *Business Insider*. <https://www.businessinsider.com/what-its-like-to-teach-elementary-students-remotely-amid-pandemic-2020-5>. (Accessed 14 April 2021).
- Hagood, M., 2004. A rhizomatic cartography of adolescents, popular culture, and constructions of self. In: Leander, K., Sheehy, M. (Eds.), *Spatializing Literacy Research and Practice*. Peter Lang Publishing, New York, pp. 143–160.
- Heath, S.B., 1983. *Ways With Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press, Cambridge, England.
- Hine, C., 2000. *Virtual Ethnography*. Sage, London.
- Hirst, E., 2004. Diverse social contexts of a second-language classroom. In: Leander, K., Sheehy, M. (Eds.), *Spatializing Literacy Research and Practice*. Peter Lang Publishing, New York, pp. 39–66.
- Hollett, T., Ehret, C., 2017. Civic rhythms in an informal, media-rich learning program. *Learn. Media Technol.* 42 (4), 483–499.
- In Their Own Words: What Students Say About Zoom school. 2020. Retrieved from news.coloradoacademy.org/in-their-owns-words-what-students-say-about-zoom-school.
- Leander, K., Boldt, G., 2013. Rereading "A pedagogy of multiliteracies" bodies, texts, and emergence. *J. Lit. Res.* 45 (1), 22–46.
- Leander, K.M., Hollett, T., 2017. The embodied rhythms of learning: from learning across settings to learners crossing settings. *Int. J. Educ. Res.* 84, 100–110.
- Leander, K.M., Lowm, J.F., 2006. Literacy networks: following the circulation of texts, bodies, and objects in the schooling and online gaming of one youth. *Cognit. Instr.* 24, 291–340.
- Leander, K., McKim, K., 2003. Tracing the everyday "sittings" of adolescents on the Internet: a strategic adaptation of ethnography across online and offline spaces. *Educ. Commun. Inf.* 3 (2), 211–240.
- Leander, K.M., Phillips, N.C., Taylor, K.H., 2010. The changing social spaces of learning: mapping new mobilities. *Rev. Res. Educ.* 34 (1), 329–394.
- Lefebvre, H., 1991. *The Production of Space*. Blackwell, Oxford.
- Massey, D., 2005. *For Space*. Sage, London.
- Nespor, J., 2013. *Tangled Up in School: Politics, Space, Bodies, and Signs in the Educational Process*. Routledge.
- Nguyen, B.T., 2021a. I actually like teaching on Zoom. N. Y. Times 14–15.
- Nguyen, C.T., 2021b. Zoom classes felt like teaching into a void—until I told my students why. *Wash. Post*. https://www.washingtonpost.com/outlook/zoom-classes-felt-like-teaching-into-a-void-until-i-told-my-students-why/2021/03/11/024afb2c-81e5-11eb-81db-b02f0398f49a_story.html.
- Niccolini, A., 2019. "It's something that requires passion": after-echoes of the ethnic studies ban. In: Leander, K.M., Ehret, C. (Eds.), *Affect in Literacy Teaching and Learning: Pedagogies, Politics, and Coming to Know*. Routledge, pp. 109–123.
- Nystrand, M., 1997. The big picture: language and learning in hundreds of English lessons. In: Nystrand, M. (Ed.), *Opening Dialogue: Understanding the Dynamics of Language and Learning in the English Classroom*. Teachers College Press, New York, pp. 30–74.
- Osman, S.M., 2020. What It's Like Teaching on Zoom to Students I've Never Met, and Often Never Seen or Heard. *Huffpost.com*. Retrieved from huffpost.com/entry/what-its-like-teaching-on-zoom_n_5fd7a2d9c5b689a6230cd60d. (Accessed 14 April 2021).
- Rimer, S., 2021. BU Students: Zoom vs. in-Person Classes? It's Complicated. *BU Today*. Retrieved from www.bu.edu/articles/2020/students-learn-from-anywhere/.
- Sheehy, M., 2004. Between a thick and a thin place: changing literacy practices. In: *Spatializing Literacy Research and Practice*, vol. 15. Peter Lang, New York, pp. 91–114.
- Soja, E.W., 1989. *Postmodern geographies: The reassertion of space in critical social theory*. Verso, London.
- Soja, E.W., 1996. *Thirdspace: Journeys to Los Angeles and Other Real-and-Imagined Places*. Blackwell, Oxford, England.
- St Pierre, E.A., 2011. Post qualitative research: the critique and the coming after. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, vol. 4. Sage, Thousand Oaks, CA, pp. 611–625.
- Street, B.V., 1985. Literacy in Theory and Practice. In: Burke, P., Finnegan, R. (Eds.), *Cambridge Studies in Oral and Literate Culture*. Cambridge University Press, Cambridge.
- Street, B.V., 1987. *Comparative Perspectives on Literacy Research*.
- Vasudevan, L., 2010. Literacies in a participatory, multimodal world: the arts and aesthetics of Web 2.0. *Lang. Arts* 88 (1), 43.
- Wilkinson, I.A.G., Murphy, P.K., Binici, S., 2015. Dialogue-intensive pedagogies for promoting reading comprehension: what we know, what we need to know. In: *Socializing Intelligence Through Academic Talk and Dialogue*. American Educational Research Association, pp. 37–50.
- Will, M., 2020. Most Educators Require Kids to Turn Camera on in Virtual Class, Despite Equity Concerns. *Education Week*. <https://www.edweek.org/teaching-learning/most-educators-require-kids-to-turn-cameras-on-in-virtual-class-despite-equity-concerns/2020/10>. (Accessed 10 July 2021).

Post-philosophies and the doing of inquiry in literacies studies

Candace R. Kuby, University of Missouri, Columbia, MO, United States

© 2023 Elsevier Ltd. All rights reserved.

Historic traditions and a move toward post-philosophies	940
Why post qualitative inquiry is needed: releasing our minds (and bodies) from a groove	941
What post-philosophies have produced for literacy studies	942
Invitation to think and do inquiry otherwise	944
Acknowledgments	946
References	946

Doing theory requires being open to the world's aliveness, allowing oneself to be lured by curiosity, surprise, and wonder. Theories are not mere metaphysical pronouncements of the world from some presumed position of exteriority. Theories are living and breathing reconfigurings of the world. The world theorizes as well as experiments with itself. Figuring, reconfiguring ... Thought experiments are material matters.

Barad (2012, p. 207).

Would educational researchers be willing to forgo methodology for the possibility of an unpredictable future-to-come, one that is not foreclosed by repetition of the same?...What would it be like not to know what to do next, and then next, and next when one inquires?

St. Pierre (2016d, p. 3).

In December 2015, Elizabeth St. Pierre gave an invited lecture for the Literacy Research Association (LRA) titled "Post Qualitative Inquiry". In this lecture, similarly to her chapter in the *Sage Handbook of Qualitative Research* (2011) titled "Post Qualitative Research: The Critique and the Coming After", she outlines what she calls, *conventional humanist qualitative methodology*, and offers a turn toward poststructural and posthuman inquiry. The "post" in post qualitative inquiry (PQI) is a (1) questioning of the taken-for-granted assumptions of qualitative research built on the logics of humanism and a (2) imagining, inventing, or experimenting with the not-yet-known or the unknown possibilities when thinking and inquiring with post-philosophies such as postmodernism, poststructuralism, posthumanism, and others.

Philosophies, paradigms, and theories all have a grid of intelligibility according to Foucault or dogmatic images of thought according to Deleuze that make particular inquiry practices (im)possible or (un)thinkable. The (im) and (un) before these words signals a both/and logic that something can be both possible *and* impossible simultaneously. As an example, St. Pierre's (2000) discusses at length what makes a feminist, poststructural paradigm possible for inquiry and gives examples of how one conceptualizes the subject, language, discourse, agency, freedom, power, reality, knowledge and so forth *and* how and why these conceptualizations matter in the doing of inquiry. The term post qualitative inquiry was a phrase St. Pierre created to name what she'd been doing as an inquirer since her dissertation in the 1990s (see St. Pierre, 2011, 2015, 2016b, 2016c). She conducted interviews and did all the methods deemed as appropriate qualitative research only to find out that the methods fell apart, so to speak, as she tried to write about methodology and do analysis that philosophically didn't align with the poststructural theories she was writing-thinking-with.

Prior to this invited lecture at LRA, literacy researchers were already thinking-with post-philosophies in literacy education (e.g., Leander and Rowe, 2006; Kuby, 2014), but they were few and often found it hard to fissure into a community of conferences and publications venues, that was/is still operating in a think atmosphere of logical/positivist empiricism and even phenomenological empiricism which is often found in interpretivist and conventional humanist qualitative methodology approaches (see St. Pierre, 2016c, for a discussion on three empiricisms). Thus, in order to open spaces for scholars to do inquiry inspired by post-philosophies, the broader field of literacy education has to be open to paradigms that aren't considered the normal or accepted way to do research. As St. Pierre (2016d) reminds us, we made up these paradigms and research methods at one point in time anyway. Surely, we can question them and think of new ways to do inquiry.

This shifting to post-philosophies and the doing of inquiry is sometimes called the 'the material turn' or 'ontological turn' in the broader social science and humanities communities and in the discussions among what is traditionally called qualitative research methodology. Post-philosophies often focus (primarily) on ontologies (being, reality) and do so in a mutually constitutive way with epistemology (knowing, knowledge) and axiology (ethics, relationships, aesthetics), hence the neologism, ethico-onto-epistemology (see Barad, 2007; Lenz Taguchi, 2010; Kuby, 2017a). What all this means is that how we evaluative and review research that is situated in post-philosophies can't be measured against the philosophical assumptions and methodological practices and standards of logical/positivist empiricism and even conventional humanist qualitative methodologies (phenomenological empiricism) that we are so familiar with. Nor can we take a study that operates in one of these empiricisms and logics of thought and just "add ontology and stir" (St. Pierre, 2016a, p. 2). By doing this we would retain

the descriptions of human being, language, knowledge, power, agency, resistance, rationality, materiality, and empiricism that organize the ontology, epistemology, and methodology of humanism we've been taught and have learned so well. Making this ontological turn requires, to a great extent, that we unlearn and try to forget what we've been taught about the nature of things. As [Kuhn \(1962\)](#) notes, this shift, this turn is not a tweaking or extension of what is but, rather, a rethinking from the ground up, in this instance, of an ontology in which the human/nonhuman, human/material binaries are no longer intelligible and in which empiricism is no longer an epistemological project.

[St. Pierre \(2016a, p. 2\).](#)

This might be hard for literacy researchers, and educational researchers in general to swallow, as our projects are firmly rooted in epistemology (knowing and learning benchmarks, objectives, goals, standards) and I would say largely with the absence of ontology and axiology. What is an empiricism, in educational (literacy) research, with epistemology not the center focus? And what might this empiricism make possible and produce?

Perhaps many of us did not have coursework in our graduate programs on ontology, philosophies, axiologies, and empiricisms. If we didn't, we might not know what they are and what they make (im)possible or (un)thinkable for inquiry practices. This was definitely true for me and I've spent over a decade trying to read and understand why this all matters for the doing of inquiry, pedagogies, and for literacies. As [St. Pierre \(2016d\)](#) writes, this entails having to untrain educational researchers. We have to rethink and think-anew our research practices.

I doubt that our research training prepares us for the productive confusion of not knowing what "to do" when we try to straddle disciplines ... We may not even understand the language of another intellectual tradition and demean its specificity by calling it "jargon." Its research practices may not seem "scientific," and its research process may lack the over-rated "systematicity" that seems to guarantee something is scientific. In confusion, we may cling to our training, so we know where we're going and what to do next. Hence, I have come to believe that research training too often gets in our way, prevents us from recognizing the "new" that is always already there in the world, and shuts down futures that might be—an education-to-come we might desire ... research training (repetition) can stifle invention (difference) and thereby impoverish possible futures of educational research and education.

[St. Pierre \(2016d, pp. 2 and 3\).](#)

This is hard to imagine as we are the instructors for research courses and thus in order to untrain ourselves, we have to learn alongside or with our graduate students. We have to teach the courses on philosophies, empiricisms, paradigms, axiologies and experimental ways of doing inquiry in order to learn (or rather invent, create, make) the content ourselves. We have to take a risk and allow new generations of scholars to stretch us and not become defensive or reject research when it looks and feels different. As editors of journals, conference proposals reviewers, dissertation committee members and advisors, and tenure/promotion committee members we have to be open to the otherwise.

It is important to pause upfront in this chapter to acknowledge a long history of Indigenous ways of knowing/being/doing in the world, many of which intersect with the ethico-onto-epistemological assumptions of post-philosophies. Indigenous scholars, such as [Robin Wall Kimmerer \(2013\)](#), a botanist and philosopher, writes about how the world comes to be based on long histories and traditions from the Citizen Potawatomi Nation of which she is a member. Thus, the neologism PQI came into being long after Wall Kimmerer's peoples thought about the inter-relatedness of lands, plants, animals (including human), spirits, and so forth which relates to concepts such as subject, agency, identity, and language. [Kakali Bhattacharya's \(2020\)](#) article, "Rejecting Labels and Colonization: In Exile From Post-Qualitative Approaches," speaks to her pathways to de/colonial ontoepistemic orientations of research and not needing the labels such as PQI. She describes this as a process of un/learning and self-exile. Other scholars, internationally, for example, such as [Fikile Nxumalo \(2018\)](#) and [Alison Jones and Te Kawehau Hoskins \(2016\)](#) also discuss the needs and tensions of working with/in post-inspired philosophies, Indigenous traditions, and/or de/colonizing inquiry practices.

Thus, as things are named, such as PQI, lines and boundaries are drawn, and exclusionary practices happen. People try to create what counts and what doesn't count as PQI. And as this happens, PQI expands, shifts, and becomes something that was unknown before. So while many post/decolonial and Indigenous scholars might not label themselves as PQI, I find a generative space in thinking through/with/across various traditions on knowing/being/doing, especially ones that believe the world comes to be in ethico-onto-epistemological ways. My aim is not to label or put into boxes people's work. I prefer a broader phrase "post-philosophies and the doing of inquiry" (although I realize what gets labeled as post-philosophy already draws lines) as it attempts to not limit the scope to qualitative work. As I've written about elsewhere ([Kuby and Christ, 2020](#)), inquiry is not able to be planned or created in pre-set methods, thus I use the term 'qualitative inquiry' instead of 'qualitative research' which is laden with a history of methods, which in many cases have become formulaic.

I invite you to (re)read the opening quotes by Barad and St. Pierre. Go ahead, look back at the first page, read the quotes, let them wash over you, sit with them for a little while. Are you open to the aliveness of theory? The reconfiguring of the world that theory does? Are you willing to have some thought experiments? Can you forgo well-known methodologies for the possibility of an unknown? Can you lean into not knowing what to do next and next when doing inquiry?

In this chapter, I sketch a brief history of literacy research traditions and why I found myself not able to *not* think with post-inspired philosophies once I started reading and experiencing them. Woven throughout the chapter, I discuss my research partnership, since 2010, with early childhood/elementary teacher, Tara Gutshall Rucker. We conceptualized the idea of literacy desirings based on poststructural and posthumanist philosophies. I expand on the inquiries we found ourselves thinking and experimenting about and how it shifted our research practices as well as pedagogical practices. I discuss why PQI is needed to get our (literacy

community) minds (and bodies) out of the groove and how this has unfolded for me and other literacy researchers. I offer some resources that colleagues might find helpful related to post-philosophies and the doing of inquiry and end with an invitation to readers.

Historic traditions and a move toward post-philosophies

Briefly, I discuss a history of approaches to doing literacy research; however, I do so not to present a linear, chronological story nor to set up a binary of “historic” versus “new” ways of doing inquiry. Rather, I share these to situate the chapter and the many, multiple ways literacy research happens, as many of these traditions are still alive and active in how we do research today. Historically, in literacy education we’ve operated within positivist and phenomenological logics of empiricisms (see [St. Pierre, 2016c](#), for discussion on three empiricisms). Logical empiricism or positivist empiricism assumes measurement can be precise, objective, and seeks to measure, explain, predict and then control (intervene). This manifests in most quantitative and mixed methods studies and, to some extent, in some qualitative studies (e.g., some case studies). This type of empiricism typically focuses on epistemological projects (knowing, knowledge). Phenomenological empiricism describes most conventional humanist qualitative research with claims made on people’s “lived experience” or the word-for-word transcriptions of interviews and “letting the data speak for itself.” This empiricism manifests in most qualitative studies (e.g., ethnography, narrative inquiry, phenomenology, discourse analysis) and typically focuses on epistemological projects.

As a field we’ve focused on research inspired by structuralist theories which focus on language, phonemes, phonics, vocabulary and so forth (e.g., [Adams, 1990](#)), constructivist theories of emergent literacy (e.g., [Teale and Sulzby, 1986](#)), social constructivist theories of literacy as socially and culturally mediated activities (e.g., [Vygotsky, 1934/1962](#)), transactional language development theories (e.g., [Harste, Woodward, and Burke, 1984](#)), participatory apprenticeship inspired by neo-Vygotskian and cultural-historical activity theories (e.g., [Rogoff, 1995](#)), anthropological socialization theories ([Heath, 1982](#)), dialogic identity ([Holland et al., 1998](#)), critical literacy theories ([Lewison et al., 2015](#); [Vasquez, 2014](#)), sociopolitical and discourse acquisition theories ([Street, 2000](#); [Gee, 2001](#)), social semiotic and multimodality ([Halliday, 1975](#); [Kress, 1997](#)), and new technologies and material culture ([Marsh, 2006](#), [Wohlwend, 2009, 2011](#)). For more details on these, see [Crawford \(1995\)](#) and [Kuby et al. \(2019b\)](#) who each trace various paradigmatic, methodological, and theoretical histories of (early childhood) literacy research.

For each of these philosophically and theoretically inspired bodies of research, particular research practices are (un)thinkable or (im)possible depending on the assumptions on ontology, epistemology and axiology that in turn create methodological practices of data production and analysis, and thus paradigms for doing inquiry. For example, when drawing upon social semiotic and multimodality theories, particular ways of doing inquiry are possible, such as multimodal analysis. Particular types of data are collected or produced. Particular claims about a humanist subject interacting in a world of materials, are possible. Each of the theories above all have (at least) one common thread: they are all rooted in a Cartesian, humanist logic of the subject.

An example of a shift toward post-philosophies in literacy studies, comes from my book with co-researcher/teacher Tara Gutshall Rucker ([Kuby and Gutshall Rucker, 2016](#)). We spend several chapters unpacking how post-inspired inquiry looks different and demands different ways of thinking about the (human) subject and literacy, specifically we focus on sociocultural theories and New Literacy Studies perspectives. This is not to position them as binaries or in opposition to post-philosophies, but rather to consider what post-philosophies have to offer or what they make (im)possible and (un)thinkable for doing inquiry. It is an affirming reading of one theory with/through/in another or sometimes referred to as diffractive readings ([Murrin and Bozalek, 2019](#)). One example we begin the book with is an orange-paper-frog-puppet-Riley-fly. We think with the coming to be of the puppet in relation to the second graders, to Tara, to discourses on literacy, to artistic materials and tools, to the physical space of the classroom, and so forth. Thinking-with and becoming-with this literacy desiring example produced a space for us to consider the affectual ways literacies come to be, literally there was stickiness between people and non-humans (e.g., tape sticking to desks and fingers, paper frog tongue jetting out to person/fly’s arm). Materials matter, or said another way, how materials are in relationship with each other matter in literacies coming to be. We needed post-philosophies to help us think. To help us think of *how* and *why* literacies come to be. To help us consider literacies otherwise. To help us consider our relationality as pedagogists and the ethics and justice of our work as teachers/researchers.

Poststructural and posthumanist philosophies shocked our thoughts and forced us to rethink literacies. We introduced the concept literacy desiring(s) first in a *Journal of Early Childhood Literacy* article ([Kuby et al., 2015](#)) and then in our book. We conceptualized literacy desiring as:

...literacy processes (the becoming of artifacts such as books, movies, dramas, 3-D models, wall murals, puppets, and so forth) and to emphasize the fluid, sometimes unintentional, unbounded, and rhizomatic ways multimodal artifacts come into being through intra-actions with humans and nonhumans (i.e., time, space, materials, environment). Literacy desiring is oriented toward the present (ever-changing) needs, wishes, and demands of students-with-nonhumans, but also with possible users of literacy artifacts in mind.

[Kuby and Gutshall Rucker \(2016, p. 4\).](#)

Since then, this concept has morphed, grown, and challenged us as we continued to read post-philosophies and engage with the lively relations with/in Tara’s classrooms ([Gutshall Rucker and Kuby, 2020](#)). Our inquiry processes of literacy desiring coming to be

have also changed how we think of the doing of inquiry. Or said another way, we couldn't continue to think in a Cartesian logic of humanism and explore literacy desiring pedagogies *and* post-philosophy(ies) inspired research practices. Our doings (both pedagogically and methodologically), if inspired by post-philosophies, had to be experimental, a wild empiricism as Olsson (2009) writes about. Or as St. Pierre (2016c) states, a transcendental empiricism based on post-philosophies. Transcendental empiricism research employs post-philosophies and these philosophies have specific conceptualizations for the subject, knowledge (epistemology), reality (ontology), relationships and ethics (axiology), agency and power, and so forth. Thus, analytic, inquiry practices must be invented anew in each inquiry. Instead of asking *what* questions, research in these post-philosophies traditions seek *how* and *why* questions focusing on ontological phenomena. This research is experimental, creative, and unfolds in the moment of researching rather than pre-planned methods. Or as Stephanie Springgay discusses, permeable curiosity and curational planning of the not yet known (see webinar with Stephanie Springgay https://www.youtube.com/channel/UC4P_GUK6QV2Wp_OAWEpw87Q).

Thus, I've written about how post-philosophies changes how one thinks of research questions, literature reviews, data collection and production, analytic practices, relationships with 'participants' in a lively human and more-than-human world, and ways of (re)presenting and writing up of research (see Kuby, 2017a,b, 2019a,b). For too long, methodology (and perhaps pedagogy) has overlooked ethico-onto-epistemological discussions. Why? As St. Pierre (2016d) offers, perhaps "first, because we teach too much of the former [research methods] and not enough of the latter [philosophies and theories], and second, if we do teach onto-epistemology, we teach it separately from methodology", and I'd add separate from discussions on (literacy) pedagogy(ies) (p. 4). Simply put, "...where we might go is always enabled by where we've been" (St. Pierre, 2016d, p. 6). The aim is not to perpetuate binaries of methods versus theory (or even pedagogies) or to create hierarchies among them, but rather think (do) them together.

Why post qualitative inquiry is needed: releasing our minds (and bodies) from a groove

PQI is a label that some identify with and others, who engage with post-philosophies and inquiry, do not. The focus of this chapter is not to trace the histories of PQI and try to articulate what counts (or not) as PQI and/or who identifies with that label (or not). But rather, I hope to articulate why inquiries inspired by post-philosophies are needed and how the philosophical assumptions demand different, experimental ways of doing inquiry. St. Pierre (2016d), in advocating for untraining ourselves as educational researchers, writes:

Even formalized, pre-existing methodologies in which a systemic process guarantees the truth of the product cannot save us from the unimaginable to come. One could make this more complicated by presuming that the unimaginable is not necessarily exceptional but ordinary, that our everyday, normalized thoughts and practices produce "minds in a groove" (Whitehead, 1925/1967, p. 197) that prevent us from recognizing the exceptional that is always everywhere but outside the narrow groove in which we live and think and do (p. 7).

The exceptional is always, already everywhere in a lively world of humans and more-than-humans, a world in which some humans experience the world 'more fully' than others and others in (less than) inhuman ways (see scholarship by Mel Y. Chen, Karin Murris, Julietta Singh, Nathan Snaza, Sarah Truman, Alexander Weheliye). For example, if our minds are in a groove of what counts as literacies or defines literacies, we might not 'see' the literacies and processes of literacies coming-to-be in spaces and relationships that lie outside of our definition of literacies. For example, I think of Abigail Hackett's (2021) scholarship with very young children creating literacies with pinecones and flowerpots or in museums spaces. Some might argue these encounters with language, materiality, and aspects that *escape* language and representation aren't literacies. However, Hackett writes of the ways she acted like an ethnographer but thought with posthumanism. Inspired by post-philosophies, she situates literacies as mundane politics of the everyday.

A key article I use when teaching inquiry courses is one by St. Pierre (2000) where she outlines the philosophical assumptions or grids of intelligibility of poststructural feminism. As I've written about elsewhere,

I often assign this reading in research courses I teach, not because I think all students should align with this paradigm but, rather, because I want them to be able to map out a similar paper and/or concept map for the paradigm(s) they work in as scholars. In other words, I want them to be able to articulate the assumptions on language, discourse, the subject, agency, knowledge, truth, and so forth within the paradigm(s) they align with for their research so they can articulate what is possible or thinkable for the doing of inquiry. Said another way, these aspects of a paradigm impact how one thinks of research design, data production, how one works with participants, interviews, field notes, analytical tools, and representation of insights from the study (and are all of these even possible or thinkable for all paradigms!!)

Kuby (2021, pp. 56 and 57).

Thus, as literacy researchers it matters how we think concepts such as the subject, agency, knowledge, truth, power, resistance, freedom, discourse, language, and so forth as these all produce ways of doing inquiry. It matters for how we understand our history(ies) and past as a field, it matters for the present and how we do inquiry (and evaluate it) now, and it matters for the future, not

only for us as researchers in the often siloed academy, but also for what realities, relationships, and literacies come to be for the children, families, educators, and our planet.

For example, by examining a predominant perspective in literacy education – New Literacy Studies — we can see through an affirmative and diffractive reading, that particular philosophical perspectives make some inquiry practices thinking or possible and others don't. This doesn't mean we don't learn from both perspectives, we just learn something different(ly). This shift in literacy research toward inquiry inspired by post-philosophies (e.g., poststructural and posthuman) is needed for several reasons (see [Kuby, 2017a](#)). First, most paradigmatic and theoretical perspectives in literacy education focus on logocentric and anthropocentric inquiries, meaning the philosophical ways language and human is conceptualized follows a Descartes separation of mind/body and rests on assumptions that language can represent an experience. Second, thinking with poststructural and posthuman together is productive and generative to rethink the human in relation to literacies. This thinking-together-with-poststructural-and-posthuman is an affirmative process of reading one with/through the other (diffraction). Third, post-philosophies rest on the notion of ethico-onto-epistemologies and thus for literacy educators, when we think with them, our inquiry practices and pedagogies must consider doing (relationships, ethics), being (reality), and knowing (knowledge of literacies). Ethico-onto-epistemologies thus force or demand us to consider ethics/relationships/justice *with* consideration for realities and what is produced in the world *with* discussions on literacies, knowledge, and teaching/learning. Finally, the [Kuby \(2017a\)](#) article also discusses how the 'material turn' builds on the 'linguistic turn', but also begs us to think differently about language, discourse, materiality, subjectivity, and so forth in relation to inquiry practices. Thus, research inspired by post-philosophies should feel, look, and sound different. It should produce something different in the world, it should shock our thoughts ([Massumi, 2002](#)).

In another publication ([Kuby, 2017b](#)), I expand on these four tenets with a data-example of second graders, Elza and Raisa's making of a blue-paper-pipe cleaner-purse during Writers' Studio in Tara's classroom. I illustrate, with this example, how post-philosophies shifted how and what we think is (im)possible or (un)thinkable in regard to research questions, producing data, interviews and transcripts, and analysis and writing up of research. This publication, and others, invites scholars to consider: what is (the) social? In other words, we claim to do social science research, we claim to look at social interactions as qualitative inquirers, but what do post-philosophies demand or force us to think about the concept of social? How is social conceptualized and what does our conceptualization of social do to our inquiry practices? In two additional publications, we focus on two data examples from Tara's second grade classrooms to rethink social inspired by post-philosophies: a silent puppet show or puppeting ([Kuby, 2019a](#)) and the coming to be of a to-scale solar system wall mural ([Kuby and Crawford, 2018](#)). For example, in the silent puppet show, one might see the skit as a flop or not worthy of being considered a literacy and/or even 'successful' as an example to research. There wasn't much human talk, it was a quiet skit and appeared difficult for the human bodies in the audience to follow along. Constraints of chronological time, uncertainty of a storyline, and the materials (e.g., puppets, a stool, the set) all worked in relationship with the students performing the puppet show. The social phenomenon came to be because of relationships, in the moment, but also in relation to past/present/future entanglements of time. Social could no longer be thought from a Cartesian logics of human.

As I continued to read post-philosophies, my thinking on inquiry continued to become anew. For instance, in [Kuby \(2019b\)](#), I write a section on what PQI demands of researchers and focus on the following: an understanding of new empiricism, resisting binaries and the philosophical assumptions they rest upon, understanding non-representational logic and transcendental empiricism, rethinking subjectivity, and a focus on ethico-onto-epistemologies. In an introduction to an edited book on posthumanism and literacy education ([Kuby et al., 2019a](#)), we frame the book on how posthumanisms shift our constructions of child and childhood, shifts what counts as literacy, and requires that our focus as literacy researchers is on producing (what is produced, why, and how), not solely on meaning making and literacy artifacts (epistemology).

What post-philosophies have produced for literacy studies

For myself, and I suspect for others, this shift into post-philosophies happened through reading outside of literacy education. Specifically, in qualitative inquiry, Indigenous studies, women and gender studies, feminist theories, environmental sciences, and philosophy to name a few. In addition to 'scholarly' publications, reading novels, paying close attention to mundane or everyday interactions, media, and for me, the life of classrooms and schools. For being open to something to shock my thoughts, as [Massumi \(2002\)](#) writes, and for events that wrest me from my natural stupor as [Deleuze \(1968/1994\)](#) writes. And in this thinking-reading-with-the-world I find myself asking what does this matter for pedagogies? How does this relate to the teaching/learning of literacies? What might this reading have to offer literacy education? Or what questions or ways of thinking does this reading make possible, not only for pedagogies, but also for inquiry practices? Below is a sampling of some of the scholarship in the literacy education community that has stretched us to (re)consider ethico-onto-epistemologies, the ethics and justice of this work, and to be open to literacies (and inquiry practices) as otherwise.

As Tara and I have written about elsewhere ([Kuby, in press](#); [Kuby and Gutshall Rucker, 2020b](#)), our relationship began with readings on multimodality and pedagogical resources about writing workshop. However, as Tara shares, a quote by Gail [Boldt \(2009\)](#) wrested us from our natural stupor:

To continue work on the piece at the teacher's insistence would shift his [a child, Kyle] role. Instead of using the writing for his own purposes, Kyle would have to assume a position of compliance ... space for young children to use materials, social relations, and time in the classroom for anything other than predetermined academic outcomes—in other words, the time to play with ideas, materials, and one another—has largely disappeared in today's classroom (pp. 11 and 12).

In addition to this quote, there are also a few comments or questions from me about pedagogical interactions that stuck with Tara and opened up spaces for conversations (see [Kuby and Gutshall, 2020b](#); [Kuby, in press](#)). Comments and questions that might have otherwise gone unnoticed — what to me seemed like mundane everyday interactions and questions in a researching/teaching partnership. In our first year together, we began to wonder what would happen if we opened up writing workshop with an invitation to “go be a writer” with a range of artistic and digital tools. How would the matter (artistic and digital tools) literally matter in what could be(come). This inquiry was a catalyst for all our writing-thinking thereafter. Our partnership as inquirers began to focus on time, space, materials, permission and trust (see [Kuby and Gutshall Rucker, 2016](#)). Around this same time, I had read (over and over) a piece by [Leander and Rowe \(2006\)](#) drawing on Deleuze and Guattari's concept of a rhizome with literacy examples. From this, we began to read more and more post-philosophies (e.g., Deleuze and Guattari, Derrida, Barad, Braidotti, Haraway, Bennett, Wall Kimmerer). I also began to read publications and attend conference sessions on inquiry inspired by post-philosophies. Once I began reading these perspectives, I (or we) could not *not* think-with it. It became how I saw the world coming to be as well as literacies, pedagogies, and inquiry practices. It wasn't a 'lens' for me to put on or take off or even 'apply' to data for analysis, it was a way of being *and* knowing *and* doing with/in the world.

Our questions became ethico-onto-epistemological in nature, meaning, we weren't solely focusing on epistemological questions that often dominate (literacy) education research such as: What does a piece of writing mean? What does the child know about writing or reading? What literacy skills are not evident? And so forth. Instead we were asking ontological questions about how we come to know and the ethics and justice of these inquiries and pedagogical practices. For example, what does it do or produce for children, literacies, relationships, the school, the world when children are invited to be writers with a range of artistic and digital and linguistic tools? What are the ethics of these pedagogical moves, inspired by post-philosophies, in our day-to-day, often mundane interactions? What is the ethics and justice of the now or in this moment rather than justice that may or may not come in the future?

We are now wondering: What is writing? What are literacies? And we believe how we come to conceptualize writing and literacies makes the world. Or, said another way, how we think writing and literacies also makes identities and subjectivity and relationships in literacy classrooms (see [Kuby and Gutshall Rucker, 2020a](#)). How we conceptualize writing and literacies makes some things (un)thinkable and (im)possible. For example, in a piece sub-titled “The End of the Book and the Beginning of Writing” inspired by a chapter by Derrida in his book *Of Grammatology* (1974/2016), we redefine writing and reimagine writers ([Kuby and Fontanella-Nothom, 2018](#)). We illuminate how second graders deconstructed and defer the meaning of writing and what counts as writing, with the pedagogical spaces Tara provided, relationships among humans, and other bodies (paper, yarn, tape, googly eyes, etc.). As we wrote:

[Caputo \(1997\)](#) stated, “to desire the impossible is to strain against the constraints of the foreseeable and possible, to open the horizon of possibility to what it cannot foresee or foretell” (pp. 133–134). Tara, in order to remain open to new possibilities, had to allow herself to trust that there was a future she was unable to predict. A future unforeseen and unanticipated but yet welcomed and embraced and unfolding in the present ... *the impossible is always and already the possible*. We invite you to consider how the ‘end of the book and the beginning of writing’ or the end of how we've always thought about writing might produce aporias, which are already possible for (re)imagining writing.

[Kuby and Fontanella-Nothom \(2018, p. 324, emphasis in original\)](#).

Being open to have Derrida's words wash over us, so to speak, to not rush to understanding, but rather to see what they produced for us, our thinking, and literacies was our aim. Not having a single research question or focus ahead of time, but rather, seeing what might unfold when we think with Derrida and literacies from Tara's classroom.

Another example comes from a handbook chapter where Tara and I think deeply with one article by Karen [Barad \(2003\)](#) and a video that came to be among second graders with an iPad. The handbook chapter, “Posthuman Theory as a Tool to Explore Literacy Desirings: Sara-Taylor-iPad-Book Becoming in Writers' Studio” ([Kuby and Gutshall Rucker, 2020c](#)) was a theoretical and philosophical playground for us to imagine how concepts like intra-action, entanglement, material-discursive, and agential cuts, defined by Barad, produce new ways of thinking about literacies. It forced us to consider matter (re)configuring, unpredictable relationships (among human and nonhuman bodies), bodily apparatuses and lively agency, and becoming (tripod, video) together in Writers' Studio. For us, we needed philosophical concepts about a lively, more-than-human world to help us understand *how* and *why* (not solely *what*) literacies came to be in the Sara-Taylor-iPad-book example we focus on in the chapter.

In another handbook chapter ([Kuby, 2019b](#)), I share inventive and experimental inquiry practices that I've engaged with rooted in post-philosophies to think-with animal research creations and nonfiction writing made by second graders in Tara's classroom. The chapter aims to reimagine multimodal research inspired by post qualitative inquiry practices. I outline four inquiry practices and shared examples from the animal research creation projects:

- Thinking with theory inspired by Alecia Jackson and Lisa Mazzei's writing
- To think is to experiment, a statement from Deleuze and Guattari
- Creating concepts inspired by Deleuze and Guattari's notion of philosophical concepts
- Diffracting an idea based in quantum physics and discussed by Karen Barad among others

These four ways of thinking/inquiring aren't pre-set or even known in advance of doing them, they are tools that open up spaces for thinking otherwise and doing inquiry otherwise.

Again, for me, the ethics and justice, at the core of post-philosophies is paramount for literacy educators to consider. How subjects are thought or conceptualized and how ethics and justice are conceptualized matter in *how we come to be* in the world (identities, relationships, subjectivity, and literacies) and what is possible for our scholarship to *do* in the world. In a co-authored piece (Zapata et al., 2018), we put to work several posthumanist concepts with three data encounters. We illuminate writing as intra-activity; as trans-corporeal, porous, and lively; and as translingual assemblages. We invite readers to think about literacies as the not-yet-known. We write,

Just as critical theories have worked toward an unfreezing of deficit monoliths that often subtract, and colonize and terrorize, posthuman theories work toward a more just world by exploring what phenomena—such as deficit monoliths—produce and the very real consequences they have on the worlds they create, such as the worlds of young writers.

Zapata et al. (2018, p. 497).

Thus, it is exciting to see more scholars thinking-with post-philosophies and matters of justice. Some scholars are thinking-with post philosophies in relation to translanguaging and translingual picturebook making (Zapata and Kuby, 2021); social class (Thiel and Jones, 2017); race (Franklin-Phipps and Rath, 2019; Thiel and Dernikos, 2020); language and representation (Hackett et al., 2020a); and lesbian, gay, bisexual, transgender, and queer communities (Wargo, 2019). There are edited collections in journals and books that pull together international scholars to think-with a variety of post-philosophies and the doing of inquiry for literacies and language and intentionally explore intersections with multimodal and critical theories (Hackett et al., 2020b; Kuby and Rowsell, 2017; Kuby et al., 2019a). These scholars are trying to (re)imagine language and literacies (education) and inquiry practices.

For those new and inspired to learn more about PQI, there are some resources that one might find helpful in addition to these readings. A webinar series titled, "Post Philosophies and the Doing of Inquiry", has fourteen webinars with international scholars doing inquiry with a range of post-philosophies and the recorded sessions are available on YouTube: https://www.youtube.com/channel/UC4P_GUK6QV2Wp_OAWEpw87Q. A new book series with Routledge titled, "Postqualitative, New Materialist and Critical Posthumanist Research" explores the doing of inquiry with these philosophical traditions. Two videos I also find helpful in thinking of post-philosophies, specifically posthumanism, are by Cary Wolfe (2012) and Michael Shirzadian (2018). Watch these and ask yourself questions such as:

- What does this video produce for you and your thinking about literacy?
- What are the desires and imperatives of humanism in relation to literacy? Should we question these?
- What might be the theories and philosophical perspectives of these humanist literacy projects? Are they self-defeating? Problematic?
- Why might we need a post-philosophical perspective in literacy education? What could it produce? Make possible?

In addition to these resources, there are many new books on post-philosophies and the doing of inquiry and resources such as a posthuman glossary (Braidotti and Hlavajova, 2018; Murris et al., 2022). Many post-philosophers have recorded lectures available on the internet, which I find helpful to consider alongside their writings.

Invitation to think and do inquiry otherwise

St. Pierre (2016d) writes, note my emphases added in brackets:

Methodologies seem to stand alone, separate from the interpretive [and I'd add, positivist] approaches that produce them. An educational [literacy] researcher can now "do" research by following a recipe without having to "think" much at all ... Does thought need a method to steer it? Does inquiry need a method to control it? What might educational [literacy] research become if it were free of the hegemonic pretensions of methods, if it were *methodology-free*? ... I like to imagine educational [literacy] research-to-come as difference instead of repetition when those who inquire do not know what to do next but trust that something new might come out of "experimentation in contact with the real."

Deleuze and Guattari (1980/1987, p. 12) (pp. 9 and 10, emphasis in original).

Trust that something new might come out of experimentation in contact with the real. I hope this chapter serves as an invitational space for educators, specifically in literacy education, to be open to other possibilities for the doing of inquiry. Post-philosophies demand us to (re)think what is (im)possible for doing inquiry. As Braidotti (2018) writes, "The driving

force for knowledge production is therefore not the quest for knowledge for disciplinary purity, or the inspirational force of radical dissent, but rather the modes of relations these discourses are able and willing to open up" (p. 44). What are we (you) willing to be a part of opening up? As [St. Pierre \(2016a\)](#) cautions, this turn to post-inspired inquiry with attention to ontology, is not an "add ontology and stir" approach. Rather, in regard to research practices, we must question dogmatic images of thought (inspired by Deleuze) or our minds in the groove (inspired by Whitehead). People made up what has become normalized or taken for granted ways of doing research. We can (re)think and question them *and* invent new ways of inquiring.

We need more research that operates explicitly from post-philosophies which focus not only on epistemology (literacy knowledge) but also ontology (what is produced and for whom) and axiology (how it is produced, in what relationships). We mustn't feel afraid or threatened by new ways of inquiring about/with literacies. If we continue to do research in the same ways—to repeat, repeat, repeat—our understandings about the world and our doing and being in the world will continue to produce the same. Replication and circularities in our research continue to reiterate the same ways of conceptualizing literacies, teaching literacies, and researching literacies. This matters for bodies—students, teachers, non-living bodies—and the lively world of discourses, policies, languages, animals, plants and so forth. It matters for how our world comes to be and what realities are produced. Who we are does not pre-exist our relationships with each other. We all come into being through relational fields or subjectivity (see [Kuby and Gutshall Rucker, 2020a](#)). Many trans/interdisciplinary scholars are helping us consider how humanist theories of the essentialized subject and identity(ies) (i.e., theories of intersectionality) might need to be thought-with post-philosophies if we aim to create a more just and ethical world (e.g., Jasbir Puar, Jennifer Nash, Mel Y. Chen, Karin Murris, Julietta Singh, Nathan Snaza, Sarah Truman, Alexander Weheliye). It would serve the literacy education community well to read broadly from trans/interdisciplinary communities of scholars and put the ideas to work in literacy education.

How we define something, the labels we use matter. How we define literacy is an ontological project and it matters in producing people, the world, pedagogies, assessments, relationships, literacies, and so forth. Thus, a matter of ethics and justice. We have to consider how bodies engaging in post-inspired philosophy inquiry will produce research that doesn't look like normalized, conventional humanist research that we are used to. We have to consider that how ethics and justice is thought and discussed inspired by post-philosophies might not look nor sound like traditional critical theories that we've become so familiar with. We have to consider what changes might need to happen with promotion and tenure processes, article reviews for journal, funding mechanisms for research, proposal reviews for conferences, and advising and teaching of students. As I've written about with others ([Ward et al., 2018](#)), colleagues are quick to judge and apply conventional humanist research standards on post-inspired philosophy inquiries, before pausing to consider *why* and *how* it might look and sound different. I've found an article on *humanities-oriented educational research* helpful in thinking differently about research ([AERA, 2009](#)). This article offers rationales for why *humanities-oriented educational research* is needed and ways to consider evaluating it. We need to be more generous to the otherwise and not yet known, rather than continue to enact normalized ways of doing inquiry and assumptions from the paradigm(s) one is most familiar with.

In the broader educational community, [Snaza and colleagues \(2014\)](#), among others, have begun to wrestle with the criticalness and ethics of post-philosophies inspired work in curriculum studies:

We are forwarding posthumanism not as a new "discourse" per se, but rather we are forwarding posthumanism as a way of recognizing that a wide variety of seemingly disparate critical approaches (feminism, anticolonial and antiracist thought, technology studies, ecology, etc.) have a common ground in directly challenging the ways humanism has restricted politics and education ... We problematize all human/humanist-centric theories because previous critiques of humanism's violence have functioned as what [Sedgwick \(1990, p. 85\)](#) calls "minoritizing" discourses—discourses that matter to some people (women, the formerly colonized, etc.) and can be easily dismissed or ignored by those with the privilege of not needing, wanting, or choosing to care. We believe posthumanism pushes intersectionality to the point where no one—no matter their field, interest, or position of power—can afford to ignore these critiques (p. 41).

Thus, for me, post-philosophies don't allow for Othering or binaries (me vs. them), or simply critique based in critical theories. But rather, they demand I always think about my power, position, and privilege – which comes into being with/in each relationship and encounter – in all my relationalities, that I am already, always a part of. I have a response-ability that I can't ignore ([Barad, 2007; Haraway, 2016](#)). As [Snaza and colleagues \(2014\)](#) write:

Consciously or not, we educators and educational researchers are used to looking at schools as places where humans dwell together to learn what it means to be human and to accumulate the kinds of skills and habits required to participate in human societies as adults. This occurs in spite of the fact that schools are connected with the nonhuman world in so many explicit and implicit ways ... we are not the center of the universe. Indeed, we should not be the center of conversation (pp. 39 and 40).

What is the future, that is yet to come, of literacy education? A future that is already entangled with the present and past histories and many bodies. Join in creating the not-yet-known for literacies education.

Acknowledgments

This publication would not be possible without the collaborative researching-teaching-thinking-writing partnerships with co-authors of my publications. Specifically, with Tara Gutshall Rucker, whose artfully, skillful teaching continues to inspire me. I am grateful for her continued commitment to think-with me.

References

- Adams, M.J., 1990. *Beginning to Read: Thinking and Learning about Print*. MIT Press, Cambridge, MA.
- AERA, 2009. Standards for reporting on humanities-oriented research in AERA publications. American educational research association. *Educ. Res.* 38 (6), 481–486. <https://doi.org/10.3102/0013189X09341833>.
- Barad, K., 2003. Posthumanist performativity: toward an understanding of how matter comes to matter. *Signs* 28 (3), 801–831.
- Barad, K., 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, Durham, NC.
- Barad, K., 2012. On touching: the inhuman that therefore I am. *Differences* 206–223.
- Bhattacharya, K., 2020. Rejecting labels and colonization: in exile from post-qualitative approaches. *Qual. Inq.* <https://doi.org/10.1177/1077800420941049>.
- Braidotti, R., 2018. A theoretical framework for the critical posthumanities. *Theor. Cult. Soc.* 36 (6), 31–61.
- Braidotti, R., Hlavajova, M., 2018. *Posthuman Glossary*. Bloomsbury Academic.
- Boldt, G.M., 2009. Kyle and the Basilisk: understanding children's writing as play. *Lang. Arts* 87 (1), 9–17.
- Caputo, J.D. (Ed.), 1997. *Deconstruction in a Nutshell: A Conversation with Jacques Derrida*. Fordham University Press, New York.
- Crawford, P.A., 1995. Early literacy: emerging perspectives. *J. Res. Child. Educ.* 10 (1), 71–86.
- Deleuze, G., 1968/1994. *Difference and Repetition* (P. Patton, Trans.). Columbia University Press, New York, NY.
- Deleuze, G., Guattari, F., 1980/1987. *Trans. In: Massumi, B. (Ed.), A thousand plateaus: capitalism and schizophrenia*. University of Minnesota Press, Minneapolis.
- Derrida, J., 1974/2016. *Of Grammatology* (G. C. Spivak, Trans.). Johns Hopkins Press, Baltimore, MD.
- Franklin-Phipps, A., Rath, C., 2019. Collage pedagogy: toward a posthuman racial literacy. In: Kuby, C.R., Spector, K., Thiel, J.J. (Eds.), *Posthumanism and Literacy Education: Knowing/Becoming/Doing Literacies*. Routledge, New York.
- Gee, J.P., 2001. A sociocultural perspective on early literacy development. In: Dickinson, D.K., Neuman, S.B. (Eds.), *Handbook of Early Literacy Research*. Guilford Press, New York, pp. 40–42.
- Gutshall Rucker, T., Kuby, C.R., 2020. Making and unmaking literacy desirings: pedagogical matters of concern from Writers' Studio. In: Toohey, K., Smythe, S., Dagenais, D., Forte, M. (Eds.), *Transforming Language and Literacy Education: New Materialism, Posthumanism, and Ontoethics*. Routledge, London, pp. 18–31.
- Hackett, A., 2021. *More-than-human Literacies in Early Childhood*. Bloomsbury, London.
- Hackett, A., MacLure, M., McMahon, 2020a. Reconceptualising early language development: matter, sensation, and the more-than-human. *Discourse.* <https://doi.org/10.1080/01596306.2020.1767350>.
- Hackett, A., MacLure, M., Pahl, K., 2020b. Literacy and language as material practices: Re-thinking social inequality in young children's literacies. *J. Early Child. Lit.* 20 (1), 3–12.
- Halliday, M.A.K., 1975. *Learning How to Mean: Explorations in the Development of Language*. Edward Arnold, London.
- Haraway, D., 2016. *Staying with the Trouble: Making Kin in the Chthulucene*. Duke University Press, Durham, NC.
- Harste, J.C., Woodward, V., Burke, C., 1984. *Language Stories and Literacy Lessons*. Heinemann, Portsmouth, NH.
- Heath, S.B., 1982. What no bedtime story means: narrative skills at home and school. *Lang. Soc.* 11, 49–76.
- Holland, D., Lachicotte, W., Skinner, D., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge, MA.
- Jones, A., Hoskins, T.K., 2016. A mark on paper: the matter of Indigenous-settler history. In: Taylor, C.A., Hughes, C. (Eds.), *Posthuman Research Practices in Education*. Palgrave MacMillan, pp. 75–92.
- Kress, G., 1997. *Before Writing: Rethinking the Paths to Literacy*. Routledge, London.
- Kuby, C.R. (in press). Teaching-Researching-Desiring: relational ethics and inquiries inspired by poststructural and posthumanist philosophies In A.D. Anders and G.W. Noblit (Eds.) *After Critical Ethnography*. Oxford: Oxford University Press.
- Kuby, C.R., 2021. What paradigmatic perspectives make possible: considerations for pedagogies and the doing of inquiry. In: Murriss, K., Zhao, W., Taylor, C.A., Cole, D.R., Bozalek, V., Kuby, C.R., Fullager, S., Quinn, J., Franklin-Phipps, A. (Eds.), *Navigating the Postqualitative, New Materialist and Critical Posthumanist Terrain across Disciplines: An Introductory Guide*. Routledge, New York, NY, pp. 43–61.
- Kuby, C.R., 2019a. (Re)thinking, (re)imagining, and (re)defining social with a more-than-human ontology given the limits of (re)(con)straining language. *Cult. Stud. Crit. Methodol.* 19 (2), 126–143.
- Kuby, C.R., 2019b. (Re)imagining multiliteracies research practices with post qualitative inquiry. In: Kucirkova, N., Rowsell, J., Fallon, G. (Eds.), *The Routledge International Handbook of Learning with Technology in Early Childhood*. Routledge, New York, NY, pp. 127–142.
- Kuby, C.R., 2017a. Why a paradigm shift of 'more than human ontologies' is needed: putting to work poststructural and posthumanist theories in Writers' Studio. *Int. J. Qual. Stud. Educ.* 39 (9), 877–896.
- Kuby, C.R., 2017b. Poststructural and posthumanist theories as research methodologies: tensions and possibilities. In: Zaidi, R., Rowsell, J. (Eds.), *Literacy Lives in Transcultural Times*. Routledge, New York, NY, pp. 157–174.
- Kuby, C.R., 2014. Matter matters: exploring the intra-activity of students and materials in Writers' Studio. In: Panel Presentation "Intra-activity Matters: Exploring the Material↔discursive of Literacy Learning in Multiple Contexts" with Honeyford, M. and Zapata, A. Literacy Research Association, Marco Island, FL.
- Kuby, C.R., Christ, R.C., 2020. *Speculative Pedagogies of Qualitative Inquiry*. Routledge, New York, NY.
- Kuby, C.R., Crawford, S., 2018. Intra-activity of humans and nonhumans in Writers' Studio: (Re)imagining and (re)defining 'social'. *Literacy* 52 (1), 20–30.
- Kuby, C.R., Fontanella-Nothom, O., 2018. Reimagining writers and writing: the end of the book and the beginning of writing. *Lit. Res.* 67, 310–326.
- Kuby, C.R., Gutshall Rucker, T., 2020a. (Re)Thinking children as fully (in)human and literacies as otherwise through (re)etymologizing intervene and inequality. *J. Early Child. Lit.* 20 (1), 13–43.
- Kuby, C.R., Gutshall Rucker, T., 2020b. (Non)sensical literacy, (non)sensical relationships. In: Schulte, C. (Ed.), *Ethics and Research with Young Children: 'New' Perspectives*. Bloomsbury Academic, London, pp. 211–225.
- Kuby, C.R., Gutshall Rucker, T., 2020c. Posthuman theory as a tool to explore literacy desirings: Sara-Taylor-iPad-Book becoming in Writers' Studio. In: Erstad, O., Flewitt, R., Küemmerling-Meibauer, B., Pires Pereira, Í.S. (Eds.), *The Routledge Handbook of Digital Literacies in Early Childhood*. Routledge, New York, NY, pp. 242–254.
- Kuby, C.R., Gutshall Rucker, T., 2016. *Go Be a Writer!: Expanding the Curricular Boundaries of Literacy Learning with Children*. Teachers College Press, New York, NY.
- Kuby, C.R., Gutshall, T.L., Kirchhofer, J.M., 2015. "Go Be a Writer": intra-activity with materials, time, and space in literacy learning. *J. Early Child. Lit.* 15 (3), 394–419.
- Kuby, C.R., Rowsell, J., 2017. Early literacy and the posthuman: pedagogies and methodologies. *J. Early Child. Lit.* 17 (3).
- Kuby, C.R., Spector, K., Thiel, J.J. (Eds.), 2019a. *Posthumanism and Literacy Education: Knowing/Becoming/Doing Literacies*. Routledge, New York, NY.

- Kuby, C.R., Zapata, A., Fontanella-Nothom, O., 2019b. Teaching and learning literacy in early childhood education: crossroads and intersections. In: Brown, C.P., McMullen, M.B., File, N. (Eds.), *The Wiley Handbook of Early Childhood Care and Education*. Wiley Blackwell Publishing, Hoboken, NJ, pp. 301–328.
- Kuhn, T., 1962/2012. *The Structure of Scientific Revolutions*, fourth ed. University of Chicago Press, Chicago, IL.
- Leander, K., Rowe, D.W., 2006. Mapping literacy spaces in motion: a rhizomatic analysis of a classroom literacy performance. *Read. Res. Q.* 41 (4), 428–460.
- Lenz Taguchi, H., 2010. *Going beyond the Theory/practice Divide in Early Childhood Education: Introducing an Intra-active Pedagogy*. Routledge, New York, NY.
- Lewis, M., Leland, C., Harste, J.C., 2015. *Creating Critical Classrooms: Reading and Writing with an Edge*. Routledge, New York, NY.
- Marsh, J., 2006. Popular culture in the literacy curriculum: a Bourdieuan analysis. *Read. Res. Q.* 41 (2), 160–174.
- Massumi, B. (Ed.), 2002. *A Shock to Thought: Expression after Deleuze and Guattari*. Routledge, New York, NY.
- Morris, K., Bozalek, V., Franklin-Phipps, A., Fullager, S., Kuby, C.R., Malone, L., Nxumalo, F., Taylor, C.A., Zhao, W., 2022. *A Glossary for Doing Postqualitative, New Materialist and Critical Posthumanist Research across Disciplines*. Routledge, New York, NY.
- Morris, K., Bozalek, V., 2019. Diffraction and response-able reading of texts: the relational ontologies of Barad and Deleuze. *Int. J. Qual. Stud. Educ.* 32 (7), 872–886.
- Nxumalo, F., 2018. Situating indigenous and black childhoods in the anthropocene. In: Cutter-Mackenzie, A., Malone, K., Barratt Hacking, E. (Eds.), *International Research Handbook on Childhood Nature: Assemblages of Childhood and Nature*. Springer, New York.
- Olsson, L.M., 2009. *Movement and Experimentation in Young Children's Learning: Deleuze and Guattari in Early Childhood Education*. Routledge, New York, NY.
- Rogoff, B., 1995. Observing sociocultural activity on three planes: participatory appropriation, guided participation, and apprenticeship. In: Wertsch, J.V., del Rio, P., Alvarez, A. (Eds.), *Sociocultural Studies of Mind*. Cambridge University Press, Boston, MA, pp. 139–164.
- Sedgwick, E.K., 1990. *Epistemology of the Closet*. University of California Press, Berkeley, CA.
- Shirzadian, M., 2018. There's no "I" in human: toward a posthuman ethics. https://www.youtube.com/watch?v=575cr_05kzQ&t=578s. (Accessed 5 November 2020).
- Snaza, N., Applebaum, P., Bayne, S., Carlson, D., Morris, M., Rotas, N., Sandlin, J., Wallin, J., Weaver, J., 2014. Toward a posthuman education. *J. Curric. Theor.* 30 (1), 39–55.
- St. Pierre, E.A., 2000. Poststructural feminism in education: an overview. *Int. J. Qual. Stud. Educ.* 13 (5), 477–515.
- St. Pierre, E.A., 2011. Post qualitative research: the critique and the coming after. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Sage Handbook of Qualitative Inquiry*, fourth ed. Sage, Los Angeles, CA, pp. 611–635.
- St. Pierre, E.A., 2015. Practices for the "new" in the new empiricisms, the new materialisms, and post qualitative inquiry. In: Denzin, N., Giardina, M. (Eds.), *Qualitative Inquiry and the Politics of Research*. Left Coast Press, Walnut Grove, CA, pp. 75–95.
- St. Pierre, E.A., 2016a. Curriculum for new material, new empirical inquiry. In: Snaza, N., Sonu, D., Truman, S.E., Zaliwska, Z. (Eds.), *Pedagogical Matters: New Materialisms and Curriculum Studies*. Peter Lang, New York, NY, pp. 1–12.
- St. Pierre, E.A., 2016b. Rethinking the empirical in the posthuman. In: Taylor, C.A., Hughes, C.L. (Eds.), *Posthuman Research Practices in Education*. Palgrave MacMillan, pp. 25–36.
- St. Pierre, E.A., 2016c. The empirical and the new empiricisms. *Cult. Stud. Crit. Methodol.* 16 (2), 111–124.
- St. Pierre, E.A., 2016d. Untraining educational researchers. *Res. Educ.* 96 (1), 6–11.
- Street, B.V., 2000. Literacy events and literacy practices: theory and practice in the new literacy studies. In: Martin-Jones, M. (Ed.), *Multilingual Matters*. John Benjamin's, Philadelphia, PA, pp. 17–29.
- Teale, W.H., Sulzby, E., 1986. *Emergent Literacy: Writing and Reading*. Ablex, Norwood, NJ.
- Thiel, J.J., Dernikos, B.P., 2020. Refusals, re-turns, and retheorizations of affective literacies: a thrice-told data tale. *J. Lit. Res.* <https://doi.org/10.1177/1086296X20966317>.
- Thiel, J.J., Jones, S., 2017. The literacies of things: reconfiguring the material-discursive productions of race and class in an informal learning center. *J. Early Child. Lit.* 17 (3), 315–335.
- Vasquez, V.M., 2014. *Negotiating Critical Literacies with Young Children*. Routledge, New York, NY.
- Vygotsky, L., 1934/1962. *Thought and Language* (E. Hanfman and G. Vakar, Trans.). MIT, Cambridge, MA.
- Wall Kimmerer, R., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions, Minneapolis, MN.
- Ward, A., Christ, R.C., Kuby, C.R., Shear, S., 2018. Thinking with Klosterman's razor: Diffracting "Reviewer 2" and research wrongness. *Knowl. Cult.* 6 (2), 28–50.
- Wargo, J., 2019. Lives, lines, and spacetime-mattering: an intra-active analysis of a 'once ok' adult writer. In: Kuby, C.R., Spector, K., Thiel, J.J. (Eds.), *Posthumanism and Literacy Education: Knowing/Becoming/Doing Literacies*. Routledge, New York.
- Whitehead, A.N., 1967. *Science and the Modern World* [Lowell Lectures, 1925]. The Free Press, New York, NY.
- Wolfe, C., 2012. Cary Wolfe on post-humanism and animal studies. <https://www.youtube.com/watch?v=5NN427KBZII>. (Accessed 5 November 2020).
- Wohlwend, K., 2009. Damsels in discourse: girls consuming and producing identity texts through Disney princess play. *Read. Res. Q.* 44 (1), 57–83.
- Wohlwend, K., 2011. *Playing Their Way into Literacies: Reading, Writing, and Belonging in the Early Childhood Classroom*. Teachers College Press, New York, NY.
- Zapata, A., Kuby, C.R., 2021. (Re)Making the past/present/future: experiencing the picturebook making of Duncan Tonatiuh and Yuyi Morales. In: Porter, C., Griffo, R. (Eds.), *The Matter of Practice: Exploring New Materialisms in the Research and Teaching of Languages and Literacies*. Information Age Press, Charlotte, NC, pp. 119–142.
- Zapata, A., Kuby, C.R., Thiel, J.J., 2018. Encounters with writing: becoming-with posthumanist ethics. *J. Lit. Res.* 50 (4), 478–501.

Urban humanities as experimental research praxis

Jonathan Jae-an Crisman, University of Arizona, Tucson, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: humans in the city	948
Developing urban humanities <i>in situ</i>	949
Life in the “undercommons”	950
The banking versus dialogic models of education	951
The humanities studio	951
Urbanism as collective life	952
Values of urban humanities: spatial justice and the practical future	952
Spatial justice	953
The practical future	953
Practices of urban humanities: thick mapping and filmic sensing	954
Thick mapping	954
Filmic sensing	955
Urban humanities, its future, and you	956
Acknowledgements	956
References	956

Introduction: humans in the city

We arrived at Mariachi Plaza early on a Saturday morning to set up. It was already bright out, a warm, blue, cloudless sky covering the historic plaza. Like Mexico City’s Plaza Garibaldi, Mariachi Plaza in Los Angeles is the city’s symbolic home to mariachi music, with mariachi musicians lingering in the plaza every day, waiting and hoping to be hired on the spot to perform, heritage culture meets day laborer. Located in the historic immigrant and now almost entirely Latinx neighborhood of Boyle Heights, Mariachi Plaza has at its center a kiosk (or gazebo) donated to the city by the Mexican state of Jalisco, an entrance to the LA subway, and a stage for performers. Around its periphery exist a number of family-run shops that are important to the community’s culture and livelihood. We put out tables, signage, canopy tents to provide some relief from the sun, and arranged an array of maps, art supplies, and photo booth props for our project. Over on the stage, we set up a station where artists would work with large cardboard pieces to make life-size human silhouettes.

We were setting up for an art project that we called “Boyle Heights en Movimiento.” I was the professor in a humanities studio, working with a team of students from the course, the non-profit organization People for Mobility Justice (formerly known as Multicultural Communities for Mobility) who advocates for safe, alternative transportation access in underserved communities of color in LA (2022), and the historic Latinx and Chicana art organization Self Help Graphics & Art which was based in Boyle Heights (2022). People for Mobility Justice had worked with some other community organizations to organize a group bike ride, its starting time and place coinciding with our project, and cyclists, families with children, and curious passers-by began to walk up to our booths. Some members of the student team worked with people to collect cognitive cartography, mapping out with marker and paper where cyclists were safe or unsafe, where they understood their place in the city, and what they wanted their neighborhood to become. Others conducted surveys, and others rewarded those who were willing to engage with us with a custom Polaroid portrait. We were also collecting our own photographs of bodies and bikes for our art project. Together, we were trying to imagine how we could harness the power of art and the critical capacities of the humanities to pursue advocacy for cyclists in Los Angeles.

As you read this sentence, even with all of the context of the Latinx neighborhood of Boyle Heights, the phrase, “advocacy for cyclists in Los Angeles,” comes with a pre-cooked imaginary. Typically, we can picture mostly white upper-middle class cyclists who perhaps bike to work, but are especially enthusiastic about the healthy lifestyle of cycling as a pleasurable activity. To extend the LA stereotype, let’s add in green juice, lycra shorts, acai bowls, sunshine, and other features of the health-centered lifestyle. Advocacy involves using the connections of this plugged-in segment of the population to pressure local governments to add bike lanes, and often these bike lanes lead to a large urban-scale recreational path where cyclists can enjoy their biking totally free from cars and with the benefit of local flora. While there is much to be lauded within this imaginary, it is problematic insofar as it excludes a large portion of the bike riding population. These excluded communities are typically lower income communities of color who bike not for pleasure but out of necessity, cycling to work because other means of transit are too time intensive, physically inaccessible, or expensive. These cyclists don’t see bike lane advocacy as something that happens for them — helmets and bike shorts are a luxury, while dangerous roads are a reality. These are the thorny issues that People for Mobility Justice grapples with on a daily basis, and the ones that we sought to engage with our project.



Fig. 1 Design mockup for a “ghost silhouette,” created by student team of Angélica Becerra, Peter Chesney, Paul Kurek, Lucy Seena K Lin, Jeannette Mundy, and Teo Wickland. Source: the author.

Back on that Mariachi Plaza stage, an artist from Self Help Graphics & Art took one of the photographs we were taking and re-drew it as a silhouette, projecting it to life size on a piece of cardboard and then cutting it out, fashioning a perpendicular piece as a base so that this ghost of a cyclist could stand on their own. This was the first of what would become a community of ghost silhouettes, distributed throughout Boyle Heights and Los Angeles in locations that were designated as high-risk to the cyclist community represented by People for Mobility Justice (see [Fig. 1](#)). These ghosts acted to both slow down drivers to speeds that were not lethal, and hopefully spark some tiny bit of reflection about this invisible community of bike commuters in Los Angeles.

The idea emerged from a collaborative effort where we discussed an artist that had inspired us all: Jay Lyn Gomez, the young, Mexican American artist who was taking the world by storm with her striking reimaginings of well-known works of art, advertisements, and even architectural spaces (for more information, see [Charlie James Gallery, 2022](#)). She forced spectators to recognize the people that were behind the work of art, the ones made invisible, by inserting them into our field of vision. For example, in perhaps her most well-known piece, titled *No Splash* (2014), she repainted David Hockney’s famous *A Bigger Splash* (1967), but instead of seeing the splash of a member of the leisure class jumping into a pool, we see a brown body at work, cleaning the very same pool. In another series of works titled *Cut-Outs* (2015), Gomez produced life-size cutouts of gardeners and other day laborers, inserting them into the manicured lawns of wealthy Angelenos and other locations where we might forget about these “essential workers” the moment they are outside of our field of vision. These cutouts are what inspired our own project, using cheap and easily available materials and our own creative capacities to try and spark a new consciousness around us.

Developing urban humanities *in situ*

Backing up a bit, you might be asking, how is this from a class in the humanities? What, exactly, are you even teaching? And what is a “humanities studio”? The students working together were completing the capstone of a year-long curriculum where they had closely read, analyzed, and debated texts that spanned disparate disciplines and topics but were united by being situated within the geography of the US-Mexico borderlands, reaching from Mexico City, through Tijuana, up to Los Angeles. To teach across such a wide swath of territory, we had a faculty teaching team from architecture, urban planning, and the humanities. The students, themselves, came from a range of fields, from the creative and applied fields of architecture, design, and art; from the socially and spatially oriented fields of urban studies, including urban planning, geography, and public policy; from interdisciplinary area fields such as Chicanx studies; and from the range of humanities fields, including Spanish language and literature, English, history, and philosophy. After engaging deeply with texts, films, histories, and geographies intellectually, from the classroom, we undertook site-specific and engaged field learning through trips from our home base at UCLA to Tijuana and Mexico City. All of this was financially and intellectually supported by the Mellon Foundation through its multi-year and multi-institution project called Architecture, Urbanism, and the Humanities which, to date, has distributed over \$36 m across 31 different grants ([Mellon Foundation, 2022](#)). This whirlwind of activity and resources was beyond extraordinary for a public university, but it also raised a number of questions for all of us as participants in this journey — questions that came to define the very thing that we called urban humanities (for a more in-depth exploration of urban humanities and numerous international example projects, see [Cuff et al., 2020](#)).

First, what are the humanities? While this seemingly simple question demands an answer of far more depth than I will provide here, I will at least point to a few observations relevant to the formation of urban humanities. One simple way to answer the question is by way of contrast: while both the humanities and the social sciences are categories of academic and applied practice which focus on people, where the social sciences focuses on generalizable patterns within populations, the humanities turns its gaze upon the fully humanized and embodied person. The social scientist tends toward positivist, quantitative, and macro-level approaches, while the humanist tends toward critical, interpretive, and micro-level approaches (see, for example, [Kagan, 2009](#)). Thus, the social sciences tend to be better suited to engage with some questions (e.g., “what is the economic impact of this policy on that region?”) and the humanities with others (e.g., “how does the language of this policy give insight into its authors and their intentions?”). The original question — what are the humanities? — remains, however, impossible to answer in any kind of complete and

comprehensive way, especially for a single person — and, indeed, this is precisely the conundrum that led to urban humanities. As the world becomes increasingly vast, as everything around us speeds up, it becomes increasingly difficult for any one person to hold all of the knowledge and expertise necessary to understand nearly anything in its entirety. This is true for categories as broad as the social sciences or the humanities, and it is true for the condition that we call “urban.”

The cultural object that we call “the city” could be described as a “hyperobject,” to use the philosopher Timothy Morton’s term (Morton, 2013). These are things that are so vast across spatial and temporal dimensions that our ability to see and touch them becomes disconnected from our mental capacity of understanding them in abstract terms. We can sit down on a park bench and feel as though we are “in the city” — but does that provide the necessary expertise for grasping the ubiquitous and global condition of the urban? Hyperobjects demand scores of people across multiple disciplines, working within and outside of academia, to come to any remote understanding of their nature. It’s a kind of approaching infinity: we are on a journey toward enlightenment, even if we never quite get there, nor expect to. Morton uses the term to describe environmental issues like climate change and, indeed, environmental humanities has emerged as a new scholarly practice to grapple with such a hyperobject (see, for example, Emmett and Nye, 2017; Heise et al., 2017). Other experimental humanities have similarly emerged: medical humanities, digital humanities, and so on. These cross-cutting fields demand new collaborative allegiances across fields, and with such interaction comes new languages and literacies necessary for communication, documentation, analysis, interpretation, and understanding. Urban humanities began in situ as a group of collaborators just trying to wrap their heads around the urban, and a handful of provisional values and practices emerged as we collectively struggled to move forward through darkness, engaging with a few additional questions beyond this.

Life in the “undercommons”

What of our privileged position within the university with funding from a private foundation, no less? Philanthropic giving, for all its good intentions and occasional ability to move faster or more experimentally than public funds, is noted for how it also acts as a means for the ultra-wealthy to undemocratically influence society and avoid paying taxes (see, for example, Reich, 2016). My urban humanities colleagues Jonathan Banfill, Todd Presner, and Maite Zubiaurre developed the ethic of the “porous university” as one response to this conundrum, emphasizing the need for schools to organize around “permeability, openness, interdisciplinarity, collaboration, and community engagement” (2016, p. 122). One can extend this ethos to private foundations, entities which are similarly situated to universities as non-profit, aspirational, yet powerful and often problematic. The philosopher-poets Fred Moten and Stefano Harney have written of the “undercommons,” an organizing principle that goes further than porosity, suggesting that for all the frenetic professional activity on the surface of the university, the real work happens under false pretenses, underground, in the fugitive undercommons. This teaching, scholarship, and collective life that happens off the books is both the real public good offered by the university and constitutes theft from the university (Moten and Harney, 2013). In other words, one ethical approach toward grant-funded academic work, as scholars and double agents, is to take as many of the resources in the university and from the foundation (Mellon Foundation included) and direct them toward channels of mutual aid, sending them into communities where there is the most need. This aspiration was never perfectly achieved, but guided our stance toward so-called “community-based learning and research” where we always asked if our partners were getting more out of our engagement than we were, trying in some small way to reverse the all-too-common extractive community partnerships that add to a glowing university marketing campaign, yet leave partners in the lurch at the moment that the university got what it needed.

Another thorny question: what of the problematic history of travel learning and research from the university, infamously “parachuting in” as colonizing experts who, once again, extract and exit? The UCLA urban humanities program emphasized the pressing nature of globalized urbanism and, as such, demanded collaborations and partnerships that spanned the Pacific Ocean — yet calendars and budgets only allowed for relatively short trips for faculty and students to our partners’ locales in Shanghai, Tokyo, and Mexico City (Cuff et al., 2020). Grappling with this question took several years of trial and error and was fundamentally intertwined with the nature of collaboration and partnership with community organizations under the rubric of the undercommons. One partial strategy was to spend as much time as possible “doing our homework”: studying up on the history and culture of the place that we were engaging. What was once a week of travel prep turned into a quarter-long seminar delving into historical readings and literature in translation, and this turned into two quarters of preparation before we were ever to even set foot in the place. What we learned was this: more preparation is better, but no amount is ever enough. This reality came home to roost most clearly when we undertook projects in our own backyard of Los Angeles: one astute student asked, well, why do we assume that we know LA any better than any of these other places just because we live here? Indeed, the vast, unknowable nature of the city ensured that a neighborhood just across town where we had hardly set foot was as foreign a land, filled with its own language, history, and culture, that we had no ownership over and little right to barge in, as a place on the other side of an ocean.

The organizing principle of the undercommons is, in part, an answer to this question. Inverting the “parachute” model of going into a place as an expert, extracting all the data and PR that one can out of the engagement and then heading back home to enjoy the benefits, we instead went as allies and co-conspirators with our collaborative partners, offering whatever resources we had (money, labor, information) and learning what we could about ourselves and our own struggles at home. Another key finding was the importance of duration, both for building the trust and shared language necessary for establishing meaningful collaborative partnerships, and also for the time necessary to complete impactful projects. In other words, our destination ceased to be our “site” of

intervention and instead more like service-training ground, and we returned back to our home as the real site of intervention where we could embark on projects without artificial limits to the length of time that we could commit to them. The format, over the course of a three-quarter academic year, included an introductory seminar on the “intimacies” between Los Angeles and Mexico City, a partnership-building “design seminar” with organizations on the ground in Mexico City concluding with a 10-day field practicum, and a culminating “humanities studio” where all of this learning was used as a catalyst for on-the-ground projects in Los Angeles in partnership with local organizations (see Cuff et al., 2020). Indeed, one such culminating project was the art and mobility collaboration described above. While our on-the-ground project involved issues and organizations in Los Angeles, it came out of learning from and deep engagement with texts, collaborations, and experiences in Mexico City, especially including innovative work on street safety. This inversion bears some relationship to recent ideas around decolonizing the university, foregoing the hegemonic unidirectional flow of “legitimated” knowledge from the university outward and, instead, opening up the possibility of transformation and Robinhood-like theft (see, for example, Mbembe, 2016).

The banking versus dialogic models of education

This inverted model of travel learning and research also forces us to reconsider conventional power relationships within teaching and learning. Of critical importance is educator Paulo Freire’s classic text *Pedagogy of the Oppressed* where he describes the conventional “banking model” of education — and, in turn, society — wherein the teacher, or other authority figure, merely “deposits” knowledge into the brains of students without concern for their agency, knowledge, or humanity and ultimately reproduces these oppressive power structures (Freire, 1970). Instead, he proposes a bidirectional or “dialogical” model of education wherein roles of teacher and student are dynamic and in flux, where equality brings liberation and freedom. The dramatist Augusto Boal created “theater of the oppressed” based on these principals, creating a model of performance that destabilized the roles of performer and audience in favor of a dialogical and recursive process of exploration which could yield mental and ultimately social breakthroughs for problems of injustice (Boal, 1985). And the philosopher Jacques Rancière describes the similarly disruptive notion of “intellectual emancipation” through analysis of the story of educator Joseph Jacotot who taught students through linguistic differences (Rancière, 1991). For Rancière, topical expertise is less important for a teacher than an abiding belief in the inherent intelligence and capability of human beings and the ability to create an emancipatory space of exploration, discovery, and learning. For all three of these thinkers, processes of learning mirror societal power structures: to learn based on the “banking model” whether it is in the classroom, in the theater, or in the workplace is to be told what to know and then to reproduce that process in others. Liberation and emancipation and “true learning” come through self-actualization as an agent with free will and the willingness to dialogically engage with equal peers, producing knowledge that is not dependent on social reproduction but, instead, new possibility.

These ideas are, of course, an ideal to strive toward, and never quite achieved, but they formed the basis for urban humanities work which blended practices of research, learning, and action, and positions of partner, teacher, and student. Sustaining these fuzzy definitions opened up possibility for unexpected learning, transformation, and collaboration. Design theorist Gerhard Fischer coined the term “symmetry of ignorance” for the social arrangement necessary as a precursor for equitable collaboration wherein no single actor can dominate power relations among the group (Fischer, 2000). Like Rancière’s “ignorant schoolmaster,” our shared humility and openness to each others’ contributions allowed for the formation of meaningful partnerships within our group and with organizations beyond it.

The humanities studio

All of this leads to the actual space of learning, a space that was critical to construct as the ground where dialogical and applied work could take place: the studio. The historic origins of the studio are hardly an equitable utopia: with origins in applied master-apprentice training, they are defined by “the banking model.” Nevertheless, philosopher Donald Schön sited his exploration of learning for what he called “reflective practice” within the architecture studio (Schön, 1987). The “reflective practitioner” is essentially an individual version of the liberated subject described by Freire and others, one who takes action in the world and can recursively examine one’s outcomes, learn, and then take new actions. The educational process of the architecture studio, however, demonstrates a model for collective liberation through dialogical peer learning combined with exploratory reflective practice. This studio-based learning departs from the power relationships historically embedded into it, in favor of open-ended, collaborative, and unexpected outcomes. Rather than a set of fixed topics that a student must learn, as is common in the lecture format class, or a fixed reading list that is to be discussed, as in the seminar, the studio instead poses a question or problem to solve. While the architecture studio begins with an architectural question — such as “how might you design a library for this irregular lot?” — the humanities studio begins with a humanistic question. How might the problematics of identity and representation be resolved within the film industry? How do we translate ancient texts into a format accessible to social media users? Or: how can artistic practices be repurposed to address a culturally specific urban mobility problem? Participants in the studio work collaboratively in an open-ended structure, taking initiative to learn whatever knowledge is needed on-the-fly, teaching each other as they work toward resolving a “wicked problem” (Rittel and Webber, 1973). This is the urban humanities that emerged “in situ,” built out of the questions that arose as we struggled to learn, work, and be together, striving toward an answer to a question.

Urbanism as collective life

The subject area focus of urban humanities is, ostensibly, “the urban.” But what is the urban? Beyond its condition as a hyperobject, there are a few additional qualities that are worth noting to better understand the literacies, practices, and pedagogies that come out of urban humanities. One starting gambit to this puzzle is to put forth the troubling yet illuminating claim posited by sociologist Henri Lefebvre in 1970: that *everything* is urban (Lefebvre, 1970). Lefebvre was not suggesting that every last person on the planet lived within city boundaries, nor that every object imaginable was actually a city in disguise. Rather, he was defining “the urban” as a condition of contemporary life under late capitalism, that the communicative, financial, and material flows around the globe had become so pervasive that we all, more or less, had access to urban life regardless of location. Even the rural village is filled with people who know what is happening in Washington, D.C. up to the very second through their smartphones, who can participate in a global cultural phenomenon of one dance or another on TikTok, is beholden to the same financial institutions as a co-op owner in New York city, with debt instruments held by thousands of investors around the world. We are all urban now.

This shift of the condition of urbanism away from a strictly material condition of dense living within a city’s core toward the inescapable conditions of contemporary life comes commensurate with a shift of urban studies away from, say, the study of high-rise building infrastructure (not to diminish this topic’s importance for particular questions), toward the sociotechnical and cultural manifestations of collective life under globalized and racialized capitalism. For this reason, contributing new approaches for urban studies to engage what scholar Lisa Lowe has referred to as “intimacies,” the transnational and recursive flows of culture and history, is increasingly urgent (Lowe, 2015). By focusing on these intimacies rather than more dated forms of urban studies which looked at cities in isolation, or in comparison to one another, you are able to engage with the urban condition without erasing cultural difference or, worse, reducing it to a false linear progression from places that are “more developed” or “less globalized,” to say nothing of the opposite extreme of situating urban phenomena as exotic manifestations of a “foreign” culture. Focusing on intimacies demonstrates the emptiness of terms such as “natural” or “authentic,” forcing us to reconcile our collective and interconnected dependencies and obligations with individuals on the other side of the globe. It reminds us that, if everything is urban, then to study the urban is to study everything, demanding humility and collaboration across discipline and difference.

It is no coincidence that urban humanities emerged out of the university, a place meant to capture the universal “whole”; there is something inextricable about the connection between the lofty ideals of academia and the “infinite” nature of the city, to use writer Rebecca Solnit’s term (Solnit, 2010). But now, as our work has demonstrated, the university has not turned out to be so universal, after all, and new connections must be built beyond academia, with and into the urban itself. Similarly, drawing from the theories that informed our pedagogical approach, one of Rancière’s key axioms from *The Ignorant Schoolmaster* is the idea that “everything is in everything” (Rancière, 1991). This paradoxical axiom points toward the intimacies that one can trace from anything in our interconnected and interdependent world: that to study the ring on the finger of a corrupt mayor, for example, is to quid pro quo study housing policies at the same time that one studies the international trade of blood diamonds. That urban humanities, in studying the urban, can focus upon the material histories and cultures of nearly any person, place, or thing, and spiral outward to build connections to a global web of other things, ultimately discovering new information about the realities of our contemporary conditions of collective life.

Since the urban humanities is situated within the humanities, it foregrounds a few elements that have often been overlooked through more social scientific approaches to urban study that have privileged abstraction, generalization, and quantification. One element is historical and interpretive: foregrounding the histories and cultures embedded in place highlight the importance of narrative, of the stories that come from places but also make places. These stories take the form of literal narratives, but also forms of shared language, symbols, and practices which can be interpreted and translated in dramatically different ways depending on who is participating, necessitating new literacies for this form of urban study. Local languages are, of course, incredibly important, but so are all the cultural and symbolic meanings that go beyond language, from foodways to rituals of daily life. Related to this is embodiment: the impact of positionality and the material conditions of the individual often go overlooked in the birds-eye view of quantification. The humanities force us to focus upon our human condition, dependent on one another and beholden to all of the possibilities and limitations that come with our bodies, from physical needs to mental biases, and beyond. And to extend our embodiment down this path one step further, it is especially important to recognize that interdependence: a single person does not make a society, a culture, or a city. Rather, to study the urban is to study our collectivity, to emphasize all of the ways that we help and harm one another, all of the ways that we share culture or misinterpret one another, how we are more interdependent on one another around the globe than ever before in human history. This fundamental reality is a key topic that urban humanities can uniquely mine, and one that is increasingly urgent.

Values of urban humanities: spatial justice and the practical future

To study our interdependence within the contemporary conditions of collective life is to also reflect on the values which drive such study. Academic work is often touted as purely objective and value-neutral, though any humanist will gladly dispel you of this notion: our work, regardless of field, is laden with our own values and biases, and the only way to deal with them is to be aware and intentional about them (see, for example, Haskell, 2000). To this end, two values surfaced in our work and became guiding principles to which we always returned. The first is the idea of *spatial justice*, a particular formation of justice which is particularly

attuned to the spatial orientation of urban studies. The second is the idea of *the practical future*, a critique in some ways of the academic approach which focuses on present and past without full engagement with the future, a future that we, intentionally or not, bring into being through our work.

Spatial justice

Spatial justice, especially as it connects to scholarly and pedagogy, was perhaps best described by geographer Ed Soja in his book *Seeking Spatial Justice* (Soja, 2010). In it, he applies the critiques and theories that came with the “spatial turn” of the 1970s–90s to our understanding of justice, arguing for the idea that justices and injustices can be spatially constructed, distributed, and resolved. Issues of spatial justice can play out at multiple scales. Extending his colleague David Harvey’s scholarship on the “right to the city” (2008; see also, Lefebvre, 1968), Soja notes that the drawing of boundaries for political representation can mean that some are better represented than other, which translates to differential access to power, which then turns into spatially distributed injustices such as worse pollution within an impoverished community that cannot fight the construction of a waste treatment plant, for example. But they can also operate on regional, national, and global scales: even now as I write this, there is a spatial injustice which is playing out globally where the distribution of COVID-19 vaccines has largely centered in the Global North, while almost entirely bypassing the continent of Africa. Beyond identifying spatial injustices, Soja also describes processes by which the fight for justice can also play out spatially. In one example, he reflects on the spatiality of the Los Angeles Bus Riders Union, which was a coalition of activists who fought for investment in LA’s ignored bus system, largely used by lower-income commuters who lived in areas where the flashy new urban rail system which used up public transit funds did not venture. This differential of mobility access is one which was a manifestation of numerous historical injustices which led to starkly spatial impacts. To bridge across disparate geographies, the community organizing of the Bus Riders Union drew on historical practices of labor unionism, and drew from the same legal and spatial arguments as the “separate but (un)equal” logic of segregation: spatial divisions ultimately lead to spatial injustices. The Bus Riders Union and its activism led to reinvestment in the bus system even as rail expanded, and demonstrated that justice need not be the work of a balkanized few nor lead to zero-sum outcomes.

Soja, being a distinguished professor at UCLA, then focuses his attention on the role that educational institutions can play in this fight for spatial justice. He provides a few insights that were informative for the development of urban humanities. First, the strategies and tactics used by activists are valuable skills that can be taught and practiced in the classroom for any range of projects. These skills set students up to be well rounded citizens, provide skills that are a selling point for entry into public-sector or non-profit employment, and, beyond this, allow teaching and learning to move past abstract or theoretical concerns within a classroom and instead have a real impact on making the world around us a better place. Second, many other theories and skills which have long been taught in the classroom find new life and meaning for students as they become actualized in the “real world.” In particular, queer, feminist, and critical race theories are seen in real time as spatial injustices are identified and engaged with through action in a kind of Deweyan “learning by doing” (Dewey, 1916). Third, the importance of understanding and engaging with spatiality within contemporary life is made clear. This has led students to learn an array of complementary skills, ranging from cartography and geographic information systems (GIS), to critical spatial analysis of contemporary phenomena, to connecting the relationship between spatial justice and urban and regional development and transformation. These skills are highly valued in a range of contexts and set students up for a wide range of vocations, most of all those that deal with spatial justice. In short, spatial justice has become a lens through which urban humanities work is structured and understood, but it has also become an organizing principle for what kind of learning should take place, and this learning has served students well even beyond the classroom.

The practical future

The practical future is a concept that we later identified in relation to the scholar Hayden White’s writing on “the practical past.” In his work, he has noted how the professionalization and development of the writing of history has transformed historical work into something that is primarily done for other historians, in contrast with the age-old practice of history which has emphasized the practical and public importance of learning from our past so that we can collectively make informed decisions about the future (White, 2014). We noticed, as our varied disciplines joined together, that many spheres of scholarly work had largely shied away from engagement with the future. To engage with the future was too much of a risk, it was too speculative, while the present and past provided firmer footing for scholarly engagement based on fact and evidence. Yet the future is perhaps the most critical thing of all: why else would anyone need all this knowledge if not to use it for informing how we ought to act and bring to life new, more spatially just futures? Those disciplines that did engage the future, perhaps most notably architecture and urban planning, often did so in a way that was to promote flashy new designs or plans for the future that lacked context, historical or otherwise. Philosopher Walter Benjamin’s famed description of the “angel of history” experiencing “progress” (1968) serves as a kind of warning against these two opposite yet equally dangerous practices, with one’s gaze fixed upon the past even as powerful forces drag it into an unknown future, sight unseen, any sense of agency abandoned. Urban humanities, then, sought out to define something like the inverse of White’s “practical past,” a future-oriented scholarly practice that rigorously engages past and present so that we can collectively and intentionally move into the kind of future that we want.

Two related ideas form the basis for this kind of scholarly practice. The first we describe as the “generative imperative” — a kind of ethical call to action for teachers and scholars to engage seriously with the practical future, even though it may be risky or dangerous. The greater danger still, in our view, is to ignore the future altogether, unwittingly becoming the angel of history. The creative and speculative, when introduced to collaborative scholarly work, has the potential to push beyond solitary visionings of one’s desired future, or fatalistic imagining of the horrible futures that are to come, but, rather, collective dreaming of a future that we can begin working on building together. The second we described early on as “immanent speculation,” which is the “how” for those attempting to work in this manner, to “practice the future” (Crisman, 2016). There is a wide world of speculative practices, including continental philosophical approaches, speculative fiction, critical design, urban planning, architecture, or, not least of all, financial speculation. Each has its own qualities that are foregrounded in service of thinking or making the future, from science fiction’s emphasis on playing out the logical possibilities embedded into scientific discoveries, to the tectonic and material constraints of architectural design. For immanent speculation, the most important quality is the idea that the future is hidden in plain sight, immanent within our past and present. Immanent speculation demands rigorous analysis of the present and past in order to speculate about the future: cultural studies, anthropology, history, the wide array of humanistic and social scientific knowledge is brought to bear on our thinking of the future. Thus, we introduced speculative work that brought together the historical and critical capacities of the humanities with the future-oriented design thinking of planners and architects — and would invite even wider circles of scholarly practice into this stance of looking back and forth, from past, to present, to future. The speculative work that we propose in urban humanities is not ungrounded utopian imaginings but, rather, rigorous and thoughtful engagement with the past and present such that better, more just futures can be identified, extracted, and manifested from them.

Practices of urban humanities: thick mapping and filmic sensing

What is the methodology for immanent speculation in the service of spatial justice? I am certain that there are a multiplicity of as-of-yet undiscovered means for scholarship which operates under these values, but given the specificity of our positionalities and those of our collaborators, and of our shared experiences, we developed a few practices that served us well in this direction — practices which could also be expanded and refined further. In other words, the particular capabilities, experiences, disciplines, and questions asked by any collaborative group should result in the development of a unique set of working practices which may or may not evolve into more established methodologies. In this way, we refer to our scholarly approaches in urban humanities as practices because of their nascent, evolving, and undisciplined (as in not linked to any one particular disciplinary tradition) nature. Nevertheless, important qualities for these approaches included two key elements: First, they incorporate spatiality in some capacity so that spatial justice or injustice can be documented, analyzed, and re-envisioned. Second, they incorporate a recursive process of collaborative making where the practice and the outcome become intertwined. For example, map making is both a process of scholarly work which demands critical analysis as the mapping takes place, and is a scholarly output that can then communicate these analyses. This eye toward the realities of making within creative scholarship also become important to satisfy the generative imperative, allowing for speculative work in the same format and context as documentary and analytic work. I will describe two such practices that we found especially useful in the context of urban humanities, each of which might be described as a kind of spatial ethnography. Spatial ethnography involves the combination of the long tradition of ethnography, vast and diverse in its various manifestations, from sensory and video ethnography to more traditional textual forms of embedded observation, but with the key organizing principle of space rather than time. While spatial ethnography could also be considered a kind of third practice in its own right, it is also a broader and more capacious category that could also describe the following two practices.

Thick mapping

The first practice is that of *thick mapping*. Thick mapping draws from the lineage of critical cartographers who first began critiquing maps as objects which were so often seen as objective representations of reality but, in fact, were laden with biases of the map maker, just as any other medium produces cultural objects that are anything but purely objective representations of our world (see, for example, the writings of J.B. Harley, 2002; Denis Cosgrove, 2001, 2012; Denis Wood, 2010; Jeremy Crampton, 2011; or John Pickles, 2004). These geographers went even farther to note the map’s history as an object used by the state to enforce its power, whether it be from the earliest ancient cadastral maps used for identifying territorial claims for the purposes of taxation, to the objectives of colonizing powers during from the 14th century onward who laid claim to faraway lands and peoples. Thick mapping is a practice of mapping originating in the cultural emphasis of the humanities which responds to these critiques through new forms of mapmaking. The thick map is “thick” in sense of ethnographer Clifford Geertz’s practice of “thick description”: to capture spatial phenomena in a map is also to demand that the mapmaker practice critical interpretation of multiple layers of meaning, layers which may, at times, even be at odds with each other (Geertz, 1973). A thick map reveals the hand of the mapmaker and remains open to extension and modification, rather than purport to be a finished, neutral, objective representation of reality (see Fig. 2).

Further, a thick map, as an attempt to represent the palimpsest-like nature of space and place, strives to incorporate the multiple, interconnected, and yet conflictual layers of time, voice, meaning, and history — which are typically collapsed down to flattened, singular layers which erase entire communities within conventional mapping practices. This idea is ever more important as digitally

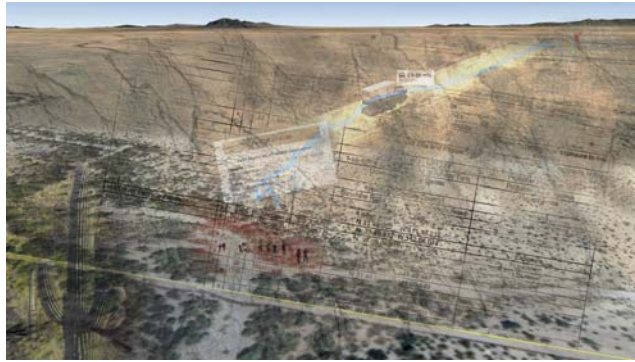


Fig. 2 Experiment in “thick mapping” on the issue of migrant death in the US-Mexico borderlands, showing the recovery and return of discovered remains from the Sonoran Desert to the Pima County Office of the Medical Examiner. Created by the author.

enabled mapping tools such as Google Maps have now become a kind of naturalized vision of our world, even as it remains as constructed as any other medium, reliant on idiosyncratic coordinate and projection systems, and drawn using largely arbitrary if widely accepted representational standards. The process of thick mapping is, itself, an experimental research praxis which uses spatial ethnography to collect data, placing that data into topological relationships within the cartographic frame of the map, self-reflexively critiquing, interpreting, and analyzing the object even as it is being made, then recursively adding, extending, iterating, and inviting others to do the same. In this way, thick mapping is both process and outcome, research and teaching, singular and collective, striving to fully capture the richness of the world within the map even as that impossibility is acknowledged at the outset. The thick map typically exists in concert with other forms of media, perhaps oral or written storytelling, or even as a series of maps extracted from a single mapping project, in order to narrate particular threads contained within the map, and to grapple with the complexity and illegibility that comes with the attempt to spatially capture the infinite nature of urban space. One global scholar whose maps demonstrate the “thickness” of the sidewalks of Ho Chi Minh City is Annette Kim, eloquently represented in her book *Sidewalk City* (2015).

Filmic sensing

The second practice is that of *filmic sensing*. Filmic sensing is something of a hybrid between the spatial ethnographic techniques of visual and sensory ethnography which use a camera as their tool and the speculative, generative, creative, and open-ended nature of the essay film. Filmic sensing involves the use of a camera to mediate, destabilize, and estrange our normalized vision of the world around us in real time. It acts as both a strategy for critical examination of space and, at the same time, as a means to visually capture these spatial and ethnographic observations for further exploration, editing, and interpretation — once again, it is both process and output. Filmic sensing, in fact, owes much to the historical origins of film as a medium: many early motion pictures were short clips of urban life. Urbanized life itself was developing as a novel and strange phenomenon just as the technology of motion picture film capture was introduced. In a way, the development of film and television has neatly tracked our global transformation into an urban society. Indeed, several remarkable early films were categorized within the newly coined “city symphony” genre of the 1920s, from Walter Ruttmann’s *Berlin* (1927) to Dziga Vertov’s *Man With A Movie Camera* (1929). These city symphony films bear key connections to the essay film genre, including their experimental formal qualities, non-linear narrative structure, and auteurist production process. Of note here is the example of Chris Marker’s *Sans Soleil* which combines speculative epistolary writing with documentary footage from disparate global locations (among other elements) to create a whole that is decidedly more than the sum of its parts, critically examining the human condition. Another contemporary global scholar whose work demonstrates filmic sensing is J.P. Sniadecki, providing an immersive ethnographic gaze into specific places throughout China, from public parks (Cohn and Sniadecki, 2012) to railroads (Sniadecki, 2014).

There are a handful of additional practices that we experimented with and found useful in different ways, from digital and multi-media installations to socially engaged art practices, and I am certain that there are myriad additional practices for working through urban humanities in ways that enable collaborative scholarly and pedagogical work that keeps a critical eye toward the future and toward spatial justice. Nevertheless, these two practices of thick mapping and filmic sensing were especially generative for our group, because of one key reason. They produced a space for collaborative creation and critical engagement (i.e., through the map or the film) that was capacious enough to incorporate both rigorous scholarship and creative or speculative impulses, yet was not methodologically predefined to such a degree that one collaborator or another could claim ownership or control over the process and outcome. This goes back to Fischer’s idea of the “symmetry of ignorance” (2000): the moment one particular methodology that is tied to one particular discipline (say, econometrics, for example), the immediately the collaborative process breaks down, the research questions become limited, and the ability to experiment and explore is curtailed. Urban humanities greatest contribution, perhaps, is the openness that comes with an experimental research praxis, allowing for deep collaboration and engagement across the infinite array of interlocutors that hold knowledge about some small piece of the city.

Urban humanities, its future, and you

The initial burst of urban humanities work and thinking owed much to the freedom afforded by the Mellon Foundation funding, which allowed us time and space to work with each other even though much of this activity did not fit neatly into institutional structures and expectations at UCLA. It provided organizational support, compensation for community partners, funding for student travel, and more. As funding has tapered off, the work has continued, yet in a way that has more constraints, tying more neatly into institutional structures, or focusing on more local travel and collaboration, for example. Mellon-funded initiatives at other universities have also, for the most part, either narrowed their scope to things such as post-doctoral research support, or are working to establish some kind of long-term institutional structure, such as the Future Histories Lab at UC Berkeley (2022). But beyond these efforts, the most promising future for urban humanities lies in the people it has engaged: over a hundred students have gone through this community-engaged research curriculum, experiencing a different way of doing scholarship from the typical solitary disciplining that occurs in conventional degree programs. These students are now doing work on their own, becoming teachers, or architects, or other kinds of practitioners who see the potential of bringing together past, present, and future in creative and critical work, in pursuit of spatial justice.

In one such example, two past urban humanities students, Jacqueline Jean Barrios and Kenny Wong, a literary scholar and architect, respectively, took the ethos of urban humanities to undertake a project which brought together experimental and creative research and pedagogy to uncover the intimacies between, of all things, philosopher Walter Benjamin's writing on the "ragpicker," artist Laura Grace Ford's zine series *Savage Messiah* (2006–9), Dickens's *Barnaby Rudge*, and the pivotal historical moment of the 1992 Los Angeles Uprising (Barrios and Wong, 2020). Their participatory method brought together researchers, students, and public to build a mental scaffolding between literary and artistic sources and the experience of the LA Uprising through the practice of "sonic scavenging," bringing new awareness to issues of race and policing and, at the same time, creating new practices of memory and commemoration. Their example demonstrates the potential not only of urban humanities, but of the humanities at large as a vital form of practice and engagement for public needs, especially in our broader historical moment of political division and social stratification where the humanities are maligned as not being "useful" enough for preparing workers in the capitalist economy. And all of this is to say nothing of the faculty, staff, and community partners who also collaborated under the auspices of urban humanities. I, myself, have moved on from UCLA to take a position in the newly established Department of Public & Applied Humanities at the University of Arizona, and have found extraordinary collaborators and potential in like-minded people working through the constellation of other experimental humanities fields, such as environmental or digital humanities.

So, I would like to conclude with an invitation: take urban humanities and make it your own. This work is unfinished, and it demands the hands and minds of participants beyond our initial group. The future of urban humanities is one that, by definition, will always be unfinished, open ended, extensible, so that it can engage the continual changes of our collective urban life. As more people bring their experience and creative capacity to collaboratively work across difference, bridging the kinds of knowledge and practice which make us human with spatial and urban studies, the better we will be able to contend with our unknown yet shared future. The practices described in this short article are an invitation for you try them out, to collaborate with people unlike yourself, with people in different social spheres, whether it is within or outside of spaces of learning, to be uncomfortable together so that you can co-create new media, analyze your surroundings, envision possible futures, address spatial injustices. If you do this, then you will be doing urban humanities. Who knows? Perhaps you already are.

Acknowledgements

Thanks to my urban humanities co-conspirators at UCLA, Dana Cuff, Anastasia Loukaitou-Sideris, Todd Presner, and Maite Zubiaurre, and to the Mellon Foundation for its generous funding of this work. Thanks also to the editors of this volume, David Yaden and Theresa Rogers, for their critical insight and feedback.

References

- Banfill, J., Presner, T., Zubiaurre, M., 2016. Urban humanities pedagogy: the classroom, education, and new humanities. *Boom A J. Calif.* 6 (3), 120–128.
- Barrios, J.J., Wong, K.H., 2020. City analog: scavenging sonic archives and urban pedagogy. *Rev. Commun.* 20 (4), 367–387.
- Benjamin, W., 1968. Theses on the philosophy of history. In: *Illuminations*. Harcourt, New York, pp. 253–264.
- Boal, A., 1985. *Theater of the Oppressed*. Theatre Communications Group, New York.
- Charlie James Gallery, 2022. Jay Lyn Gomez. <https://www.cjamesgallery.com/artists/jay-lynn-gomez>.
- Cohn, L. D. & Sniadecki, J. P. (Directors). (2012). *People's park*. [Film].
- Cosgrove, D., 2001. *Apollo's Eye: a Cartographic Genealogy of the Earth in the Western Imagination*. Johns Hopkins University Press, Baltimore.
- Cosgrove, D., 2012. *Geography and Vision: Seeing, Imagining and Representing the World*. I.B. Tauris, London.
- Crampton, J.W., 2011. *Mapping: a Critical Introduction to Cartography and GIS*. Wiley, New York.
- Crisman, J., 2016. Practicing the future: exercises in immanent speculation. *Boom A J. Calif.* 6 (3), 25–31.
- Cuff, D., Loukaitou-Sideris, A., Presner, T., Zubiaurre, M., Crisman, J.J., 2020. *Urban Humanities: New Practices for Reimagining the City*. MIT Press, Cambridge, MA.
- Dewey, J., 1916. *Democracy and Education*. Macmillan, New York.
- Emmett, R.S., Nye, D.E., 2017. *The Environmental Humanities: a Critical Introduction*. MIT Press, Cambridge, MA.
- Fischer, G., 2000. Symmetry of ignorance, social creativity, and meta-design. *Knowl. Base Syst.* 13 (7–8), 527–537.

- Freire, P., 1970. *Pedagogy of the Oppressed*. Herder and Herder, New York.
- Future Histories Lab. (2022). Homepage. <https://futurehistories.berkeley.edu/>.
- Geertz, C., 1973. *The Interpretation of Cultures*. Basic Books, New York.
- Harley, J.B., 2002. *The New Nature of Maps: Essays in the History of Cartography*. Johns Hopkins University Press, Baltimore.
- Harvey, D., 2008. The right to the city. *N. Left Rev.* 53.
- Haskell, T.L., 2000. *Objectivity is Not Neutrality: Explanatory Schemes in History*. Johns Hopkins University Press, Baltimore.
- Heise, U.K., Christensen, J., Niemann, M., 2017. *The Routledge Companion to the Environmental Humanities*. Routledge, London.
- Kagan, J., 2009. *The Three Cultures: Natural Sciences, Social Sciences, and the Humanities in the 21st Century*. Cambridge University Press, Cambridge, UK.
- Kim, A.M., 2015. *Sidewalk City: Remapping Public Space in Ho Chi Minh City*. University of Chicago Press, Chicago.
- Lefebvre, H., 1968. *Le droit à la ville*. Anthropos, Paris.
- Lefebvre, H., 1970. *La Révolution Urbaine*. Gallimard, Paris.
- Lowe, L., 2015. *The Intimacies of Four Continents*. Duke University Press, Durham.
- Mbembe, A., 2016. Decolonizing the university: new directions. *Arts Humanit. High Educ.* 15 (1), 29–45.
- Mellon Foundation, 2022. Grants Database: Architecture, Urbanism, and the Humanities. https://mellon.org/grants/grants-database/?amount-low=&amount-high=&year-start=1969&year-end=2020&city=&state=&country=&q=%22architecture%2C+urbanism%22&per_page=25.
- Morton, T., 2013. *Hyperobjects: Philosophy and Ecology after the End of the World*. University of Minnesota Press, Minneapolis.
- Moten, F., Harney, S., 2013. The university and the undercommons. In: *The Undercommons: Fugitive Planning & Black Study*. Minor Compositions, New York, pp. 22–43.
- People for Mobility Justice, 2022. Our Roots. <https://www.peopleformobilityjustice.org/about-us>.
- Pickles, J., 2004. *A History of Spaces: Cartographic Reason, Mapping, and the Geo-Coded World*. Routledge, London.
- Rancière, J., 1991. *The Ignorant Schoolmaster*. Stanford University Press, Palo Alto.
- Reich, R., 2016. Repugnant to the whole idea of democracy? On the role of foundations in democratic societies. *PS Political Sci. Polit.* 49 (3), 466–472.
- Rittel, H.W.J., Webber, M.M., 1973. Dilemmas in a general theory of planning. *Pol. Sci.* 4, 155–169.
- Schön, D., 1987. *Educating the Reflective Practitioner*. Jossey-Bass, San Francisco.
- Self Help Graphics & Art, 2022. About Us. <https://www.selfhelpgraphics.com/about-us>.
- Sniadecki, J. P. (Director). (2014). *The Iron Ministry*. [Film].
- Soja, E.W., 2010. *Seeking Spatial Justice*. University of Minnesota Press, Minneapolis.
- Solnit, R., 2010. *Infinite City: A San Francisco Atlas*. University of California Press, Oakland.
- White, H., 2014. *The Practical Past*. Northwestern University Press, Evanston.
- Wood, D., 2010. *Rethinking the Power of Maps*. Guilford Publications, New York.

Reimagining research on writing: a timely rhizomatic argument

Donna E. Alvermann and Casey Salman, University of Georgia, Athens, GA USA

© 2023 Elsevier Ltd. All rights reserved.

Chapter overview	958
Rhizomatic mapping: concept as method	958
How researchers have imagined writing	959
What writing shows	960
A counternarrative	960
What writing follows	960
No counternarrative	961
What writing helps	961
A counternarrative	961
What writing does	962
A counternarrative	962
Concluding thoughts	963
References	964

Chapter overview

Our focus in this chapter is to document the ways writing is described and researched by scholars in the field of literacy education. Specifically, we explore the writing process as viewed through scholarship that uses [Deleuze and Guattari's \(1980/1987\)](#) concepts of rhizomes, maps, and tracings to theorize potential entanglements in reading-writing relationships. In doing so, we are careful to avoid imposing these two French philosophers' thinking as if it were a mere grid-like framework in which to analyze the entanglements of reading and writing. From the start, we have heeded [Leander and Rowe's \(2006\)](#) insight that Deleuze and Guattari's ideas are best used to "spark our thinking in new directions.... [such that literacy in use involves] generating new ways of connecting texts, modalities, and performed identities" (pp. 432–434). For example, by using Deleuze and Guattari's concepts of rhizomes, maps, and tracings to rethink how writing and reading are entangled, emergent, and ever-becoming, we created an analytic method that combines traditional expository prose found in academic literature reviews with our narrativized accounts of literacy in use—that is, the *doing* of writing in relation to reading.

Rhizomatic mapping: concept as method

According to [St. Pierre \(2016\)](#), in a time of renewed interest in empirical inquiry, researchers trained in conventional humanist qualitative inquiry find themselves in need of language that decenters epistemology (i.e., *what* constitutes knowledge/known) to make room for ontology (i.e., *how* such knowing becomes 'reality'). To address that need, St. Pierre points specifically to a Deleuzian belief that language is action, a way of doing things with words. Inspired by this call for action, literacy researchers have used rhizomatic mapping as a spatial method for studying a range of literacy events involving texts. Among numerous examples, mapping has been used to look closely at adolescents' appropriation of popular cultural texts when re-negotiating their own identities ([Hagood, 2004](#)); to both situate and trouble literacy practices in Christianity's faith-based schools ([Eakle, 2007](#)); to analyze electronic poetry ([Evans, 2014](#)); to reread a pedagogy of multiliteracies ([Leander and Boldt, 2012](#)); and to conduct a non-linear analysis of a shared parent-child literacy event ([Yaden, 2004](#)).

Part of the creativity in mapping disparate phenomena at the surface level involves reading broadly in areas related to one's topic of study. For this chapter, we read in cultural studies and critical methodologies, including articles such as [St. Pierre's \(2013\)](#) explanation of how a Deleuzo-Guattarian ontology, which does not assume a subject/object binary, requires researchers to think differently about the concept of data. So differently, in fact, that Deleuze and Guattari (1987/1980) deemed it necessary to state explicitly why subject/object and human/material oppositions cannot be thought or lived:

We are not separate from the world. Being in every sense is entangled, connected, indefinite, impersonal, shifting into different multiplicities and assemblages. Our work is not "to reach the point where one no longer says I, but the point where it is no longer of any importance whether one says I." (p. 3)

We also drew from earlier research on inequitable literacy practices involving muted voices, silent texts, and cultural representations ([Alvermann, 1993, 2000](#)). The first of those two studies resulted in an in-depth literature review that became the author's

presidential address at the 42nd annual meeting of the National Reading Conference (now renamed the Literacy Research Association). Published as part of the proceedings for that meeting, it was Alvermann's first attempt at creating historical-autobiographical methods for uncovering silences, thus keeping literacy research and scholarly writing from masking more than they disclose. The second publication was a rhizomatic re-accounting of findings from her earlier research on adolescents' perceptions and negotiations of literacy practices while participating in after-school Read and Talk Clubs at a local library (Alvermann et al., 1999).

From Deleuze and Guattari's perspective, it is how texts function in relation to each other—often beyond themselves—that is of particular interest. This interest stems from their view that texts, like rhizomes, connect with other things. For instance, "a rhizome ceaselessly establishes connections between semiotic chains, organizations of power, and circumstances relative to the arts, sciences, and social struggles" (p. 7). Deleuze and Guattari's metaphoric use of the terms *maps* and *tracings* is this: a map is "a part of the rhizome ... open and connectable in all of its dimensions ... [with] multiple entryways, as opposed to the tracing, which always comes back 'to the same'" (p. 12). Maps, unlike tracings, are always becoming; they have no beginnings and endings, just middles (Deleuze, 1995). It is by looking at the middles that one begins to see how, in perspective, everything else changes. Dimitriadis and Kamberelis (1997) explained the process this way:

In drawing maps, the theorist works at the surface, "creating" possible realities by producing new articulations of disparate phenomena and connecting the exteriority of objects to whatever forces or directions seem potentially related to them. As such, maps exceed both individual and collective experiences of what seems "naturally" real. (p. 150)

Looking for middles, rather than beginnings and endings, also supports the decentering of key linkages in texts. Such decentering invites researchers to remain open to proliferations of rhizomatic ruptures and discontinuities, which in turn, endlessly create other linkages. For these reasons, Grosz (1994) argued that it is no longer appropriate to

ask what a text means, what it says, what the structure of its interiority is, how to interpret or decipher it. Instead, one must ask, what it *does*, how it connects with other things (including its reader, its author, its literary and non-literary context). (p. 199, emphasis added)

Finally, because we rely on narrative writing as part of the mapping process, it is fitting we conclude this section with a brief rationale for using that particular genre. According to Joseph Featherstone (1989), "the telling of stories can be a profound form of scholarship moving serious study close to the frontiers of art" (p. 377). Contemporaries of Featherstone (e.g., hooks, 1990; Lawrence-Lightfoot and Davis, 1997; Richardson, 1997) critiqued the notion that scholars should ever assume authority for telling other people's stories. Mindful of their critique, we offer rhizomatic accounts—or stories if you will—that are not of other people's telling; instead, they are of our own making. A major impetus for doing so is Brandt's (2015) *The Rise of Writing: Redefining Mass Literacy*. Brandt wrote it to provide an easily accessible text on her seven-year research project in which she interviewed 60 workaday adult writers. A separate set of 30 interviews involved participants whom she described as writing-intensive young people from 15 to 25 years of age who identified as authors of fiction, poetry, nonfiction, and journalistic writing. Based on findings from both sets of participants' data, Brandt claimed that writing is overtaking reading to become the more frequently applied language art in literate people's day-to-day activities. We offer such evidence not as an argument for moving reading over so that writing can take its place, but instead as support for choosing to report our rhizomatic findings as counternarratives on the ways writing has been described and researched by scholars in the field of literacy education.

How researchers have imagined writing

Smith (2019) captured various theoretical dimensions of writing since the late 1800s by using the metaphor of *waves* to characterize how these dimensions emerged, overlapped, evolved, and to some degree still linger. The product dimension is the first she noted, wherein people believed writing samples could indicate a person's cognitive abilities and developmental stages. Gradually, as thinking transitioned from product to process, researchers emphasized prewriting, writing, and rewriting as being key to one's literacy development. The practice dimension ushered in the importance of social interaction and environment, such that studies of the role writing can play in identity formation gained prominence. Building on those ideas, the pathways dimension highlighted writing as a multidimensional activity bound up with social, cultural, and environmental factors. In describing these dimensions, Smith made clear that ideas evolve over time in a nonlinear way. Thus, certain ideas endure and continue to be influential in research and practice.

In this chapter, we adapt Smith's (2019) use of the term *waves* to capture changes in how reading-writing relationships have been imagined by researchers over time. We acknowledge up front that these shifts have been neither linear in direction, nor totally separate from one another. We use the wave metaphor to denote what the action-verb of writing is. More pointedly, what writing *shows*, what writing *follows*, how writing *helps*, and what writing *does*.

What writing shows

The belief that writing can indicate a person's cognitive ability is pervasive. Smith refers to this idea as the product dimension, and although it was historically the first that she was able to locate, it is still an idea that is prevalent in 21st century literacy practices. The product dimension positions writing as a product that serves as evidence or demonstration of intelligence, learning, or more broadly, cognitive ability. This wave of thinking is the basis for much of contemporary assessment practices. Standardized testing creates writing samples that represent benchmark skills against which an individual's writing is judged. Such assessments often rest on the assumption that there are fixed cognitively related developmental stages through which a writer must pass in order to gain mastery. Consequently, it is not uncommon for assessments alone to be the driving force behind researchers' interests in designing studies that measure what writing can show. For instance, Swain et al. (2010) relied on prominent feature analysis to demonstrate how standardized writing assessments can support classroom writing instruction. By analyzing on-demand writing from state assessments, Swain et al. were able to identify 32 prominent features (e.g., sentence variety, hyperbole, and faulty punctuation) that correlated with higher scores on a state's writing assessment.

Beyond their usefulness in assessment-inspired studies, products from writing instruction also serve as indicators from which to draw conclusions about literacy related learning. In one such study (Nippold et al., 2005), persuasive writing samples representing writers from early childhood to adults in their mid-20s revealed how writers' argumentative strategies change over time. Even when researchers design studies to focus on multimodal storytelling within a New Literacies framework (and its emphasis on the social practices of literacy), options still exist for (a) relying on the logic of product analysis, and (b) drawing implications to show support for various writing standards in the Common Core (Chisholm and Trent, 2013). In short, writing that shows a connection to measurable end products retains a highly visible position in linear and ordered systems of thinking.

A counternarrative

In questioning the viability of systems that constrain writing as an end product, we offer a rhizomatic account of what a product-oriented analysis may have concealed in the Chisholm and Trent (2013) research report. Their study of a high school English class's use of digital storytelling to understand the literary element, place, centered on one focal student, Riley. Her final project served as their primary unit of analysis for identifying which of the standards (e.g., elements, visual composition) aligned with her story. Said another way, the researchers used Riley's multimodal text to create a tracing of how her work connected to the Common Core. This reverence for a "product" as well as the assumption that Riley's mastery of certain standards signified an "end" of learning prompted us to recall Deleuze and Guattari's (1987/1990) observation that "A map has multiple entryways, as opposed to the tracing, which always comes back 'to the same.'" The map has to do with performance, whereas the tracing always involves an alleged "competence" (pp. 12–13). Referencing the Common Core and Riley's so-called competence in demonstrating a degree of mastery over certain writing standards symbolizes how researchers (and yes, Casey and I as teachers) can comfortably arrive at unexamined conclusions.

Using rhizomatic mapping to reframe the Chisholm and Trent study, we envision entryways into Riley's writing as signifying middles, not beginnings and endings. Putting the tracing back on the map, as Deleuze and Guattari suggest, frees us to ask how else Riley's work connects to the real world. The last line of her digital story, placed next to a family photo and a picture of her home, jumps out at us: "[The bird's nest] reminds me of my own family, and how we used to be. I let a drawn out sigh, and realize ... I am home." Riley orients her story in a middle space—referring to memories and the past, but also descriptive of the present. A rhizomatic mapping lets us observe Riley describing her understanding of place in this excerpt of an interview from the Chisholm and T. study:

Yeah, like at the beginning, I thought that place was like, okay, this library is a place, outside is a place, you know. Everything is ... you're just there. Your place. That's where you are. But now, after [Ms. Trent] taught me about place, I realize that it is what's connected to you at the heart, not just any random spot in the world.

Riley reports that this writing project helped her understand how her connections make place significant, helping her to map her home in relation to the objects, nonhuman beings, and people that she associates with it. To summarize, a partial rhizoanalysis of Chisholm and Trent's research reveals how writing facilitated a change in Riley's perspective that was not captured in the researchers' end-product analysis. It also facilitated a closer look at how Riley was making connections between place and things that are important to her.

What writing follows

Correlational studies of reading and writing suggest a hierarchical arrangement in the cognitive realm—one in which reading leads and writing follows. However, that was not always the case. Prior to Tierney and Shanahan's (1991) landmark review of research on the relationships between reading and writing, there had been decades in which these two cognitive processes were treated as separate entities residing in different academic spheres. Indeed, it took what Gee (1990) termed 'the social turn' to disrupt the status quo.

In Alexander and Fox's (2019) historical analysis of the literacy field, the move away from research dominated by cognitive models of reading and writing to sociocultural models occurred roughly between 1986 and 1995. The conditions for such change included an increased "acceptance in the literacy community of ethnographic and qualitative modes of inquiry advocated in social and cultural anthropology" (Alexander and Fox, 2019, p. 43). While many literacy researchers in the sociocultural era focused on the interrelated nature of reading and writing, they often framed writing as an activity that follows reading. That wave of research relied heavily on correlational studies to predict reading competency based on certain types of writing skills.

Although correlational studies of literacy were most common in the 1960s through the mid-1980s, this approach to inquiry continues. For instance, researchers of two such studies (Allen et al., 2016; Allen et al., 2019) reported mixed results when studying whether or not writing flexibility in argumentative essays predicts reading comprehension. Ideologically, this kind of research supports what social anthropologist Brian V. Street (1995) labeled the autonomous model of literacy—one predicated on the assumption that reading and writing are neutral processes devoid of relations of power—whether due to socioeconomic, cultural, or political influences. In extending Street's critique of the autonomous model of literacy, 21st century activist researchers Byrd et al. (2021) call for composition studies whose purposes range from "engaging marginalized communities to taking up matters of policy or direct action" (p. 16).

No counternarrative

We refrained from attempting a rhizo-analysis of the two correlational studies mentioned above. To have done so would have been a contradiction of Deleuze and Guattari's (1987/1990) thinking. Although correlational research is capable of teasing out entangled relationships between reading and writing, the use of statistical measures to represent those relationships would produce a report that is more akin to tracing than mapping. Indeed, replications of a tracing (one after another) can lend added confidence that a relationship is "real" and worthy of implementation. Minus a mapping that keeps multiple variables in play, a researcher cannot readily think the impossible. Nor can the literacy rhizome rupture freely when it's constrained by a paradigm that prefers wrapping up messiness to letting it flow.

What writing helps

A significant question to have emerged from the Center for the Study of Reading (CSR) at the University of Illinois, Urbana-Champaign in the late 1970s through the 1980s was whether or not findings from basic research conducted on the cognitive processes underlying text-based reading comprehension had application in classroom-based instructional research, particularly as it applied to reading and writing instruction. This was the question that drove the National Reading Research Center's agenda—a collaboration between the University of Georgia and the University of Maryland. The NRRC was federally funded by a five-year (1992–1997) grant from the U. S. Department of Education Alvermann and Moje (2019). Shortly after findings from the NRRC's classroom-based instructional research were disseminated through published articles and associated instructional pamphlets aimed at teacher implementation, a relatively small but robust body of literature sprung up around the use of strategies that situated writing as a helper in promoting readers' responses to literary texts through guided journal writing (Wong et al., 2002) and literature logs (Saunders and Goldberg, 1999). On occasion, large-scale research projects, such as Graham and Hebert's (2011) meta-analysis of writing and writing instruction's impact on students' reading, have offered hope that writing would grow out of its subsidiary role as helper and occupy an entity worth studying for its own sake. Although Graham and Hebert focused their 2001 meta-analysis on writing, giving hope that writing would get the attention it deserved, the research questions still positioned writing in a helping relationship to educational outcomes. They questioned if an increase in writing improved reading comprehension, and specifically looked at whether writing about class readings improved understanding of those texts.

In that same vein, proponents of novel study—a boutique-like research domain that focuses on writing in relation to literary works—appeared. But even so, one can find traces of the helper role infiltrating writerly aspects of novel reading. For instance, sometimes specific instructional writing strategies used during literary novel study can end up being evaluated by reading comprehension assessments or adapted for use in writer's workshop (Tracy and Headley, 2013). Findings from one such study, an intervention called Reader Response Plus Granger et al. (2007) caught our attention because key participants, all of whom experienced disabilities related to reading, wrote in their journals after reading portions of a class novel to prepare for group discussion. Granger et al. used a common informal assessment, the Qualitative Reading Inventory (QRI), to measure change from pre-to post-intervention. The researchers reported that each student's comprehension level showed improvement. We were curious, however, as to what else might have been a learning outcome of Reader Response Plus.

A counternarrative

Entering that study's parameters as researchers and educators grown wary of how tracings from one curricular reform after another have returned us to the same place, we nonetheless set out to explore potential discontinuities, even ruptures, that might put us on a different pathway for dealing with the subjugation of writing by reading. In the rhizo-analysis we've temporarily tagged "The Status Quo," we question to what extent assessments (even informal ones such as the QRI), are tracings of what students at different

ages and grade levels **should** be able to do instead of maps of what they are **actually** doing when tracings aren't there to blunt creativity and the joy of breaking off to go in a different direction.

Another opportunity for pushing back on the status quo comes from the intervention's reliance on journaling and discussion—the middles rather than the typical pre- and post-assessment activities. It's as if writing has backed off its traditional role of helper to become playfully complicit in the writer's plan for novel study. Indeed, there is some empirical evidence to suggest that rhizoanalysis is not restricted to mere theorizing. For example, information in the form of observations that teachers were making about one of the case study students, David (a pseudonym), showed that initially he didn't like the class novel, but gradually began sharing more about the book in discussions with his classmates. Separately in the article, the researchers noted, "The teacher attributes the increases in personal connections made to the novel by all students as largely due to David's modeling. He began to connect personally with the book through journaling, regardless of the prompt. After hearing his personal stories, other students began to follow by example." Although Granger et al. do not mention it, we tend to give at least partial credit to writing's easing away from its helper role. In short, writing took on a quasi-facilitator's role rather than that of a lowly gatekeeper.

What writing does

In this, the last wave of our action-verb for what writing is, we envision its entanglement with reading in much the same way that [Bloome and Green \(2015\)](#) described the inseparability of literacy and the people who use it. In their words, "Literacy cannot be separated from what people are doing, how they are doing it, when, where, under what conditions and with whom they are doing it; metaphorically, there is no separation of the dancer from the dance" (p. 20). Further, because the example we use in this last wave is itself a creation derived from mapping sound, it is in part dependent on high-tech virtual interfaces for displaying what writing does. Far from being redundant, the rhizomatic analysis we use in constructing a counternarrative for this wave permits us to expand our thinking further by exploring potentially new offshoots.

That said, we remain focused on the ways in which literacy researchers are moving away from treating writing as the end product of a reading activity. In doing so, we note the considerable influence of [Stornaiuolo et al. \(2017\)](#) interest in mobilities on our own thinking. Specifically, their framework for inquiring into transliteracies pulls from Deleuze's concept of emergence: "the process by which unprecedented, unimaginable ideas come to be in the world" (as cited in Stornaiuolo et al. p. 77). Emergences, like rhizomes, are concepts that make it possible to construct counternarratives of what previously might have remained 'just findings' in reading-writing studies destined for dusty bookshelves (or worse).

Before going further into this introduction of what writing does, we need to acknowledge a branch of scholarship dedicated to understanding the intertextuality of writing and reading, particularly the way that writing becomes entangled with the textual worlds of the writer. Sometimes this research focuses on the connection between what [Dyson \(1991\)](#) calls "worlds and words" in writing and written language development. Working within these intertextual spaces, [Bloome and Egan-Robertson \(1993\)](#) examined the social nature of textual connections while [Dyson \(2018\)](#) focused on the impact of popular culture on young children's writing. As noted earlier, [Tierney and Shanahan's \(1991\)](#) review of the research on reading-writing relationships explored intertextual connections that we now credit with having foreshadowed the action term, writing that 'does'. Even earlier language and literacy researchers working with cognitive processing models for composing (e.g., [Brozick and Petrosky, 1979](#); [Flower and Hayes, 1981](#)) were careful to emphasize that writing deserved its own line of inquiry.

Finally, as evidence from [Brandt's \(2015\)](#) scholarship would suggest, writing is stepping out of reading's shadow and inserting itself into neighboring domains. For instance, one might say that writing has found a niche in the multimodal research on sound-writing ([Brownell, 2019](#); [Wargo et al., 2021](#)); imagery ([de los Rios, 2020](#)); touch ([Sanders, 2018](#)); dance ([Risner and Anderson, 2008](#)); and meme-ing ([Knobel, 2022](#); [Lankshear, 2022](#)). By entangling itself with modes that largely reject hierarchical logic and representational thinking, writing is becoming more and more an interpretive activity in its own right rather than simply evidence of that activity. We view [Wargo's \(2018\)](#) study of soundwriting as a good example of the transitioning process just described. Thus, portions of that study's abstract appear here as a way of providing context for our rhizomatic analysis of what soundwriting does.

In partnering theoretical insights from multiliteracies, queer phenomenology, and sound studies, Wargo underscores sound's capacity as a resource for adolescent multimodal writing. Refracted through a singular LGBTQ youth's practice of sonic cartography, the participant's counternarrative—comprised of notes, sounds, and neighborhood hand-drawn maps—expresses feelings of isolation in the midst of pressures to become part of a larger after-school gathering. Findings from Wargo's research suggest that sonic cartography is a form of (counter)storytelling in that it creates new ways of knowing and becoming. Wargo asks literacy researchers and educators to consider how rhizomatic stories that youth compose can turn up the volume on formal and informal literacy learning.

A counternarrative

Given the richness in multimodal writing research, why did we, as authors of a piece on trends in literacy education, settle on [Wargo's \(2018\)](#) study as our example? Mostly because Andi, a self-identified 17-year-old Latina lesbian, who delighted in calling herself a high-school dropout, was a participant in Wargo's ongoing longitudinal research project. Also, Andi was of interest because she

had access to open-source tools that she connected to her Instagram account—her composing tool of choice. Geotagging software allowed Andi to map sounds of injustice in various urban neighborhoods where she felt safe, not so safe, and/or an outsider.

We also were drawn to Andi's stance that tools for capturing video and sound provided all she needed for sonic compositions. We took Andi at her word and thus were surprised by the linguistic traces she left while mapping sounds and images at the core of her self-imposed silences. By using rhizomatic analysis to inspect Andi's sonic maps, we learned that reading and writing are inseparable literacy processes in the presence of sound and imagery. Evidence of Andi's use of three modes (sound, imagery, and language) coexist. For example, her sonic map of what she called "#MYSIDEWAYSSCHOOL" consists of what Andi heard when she audio-recorded "sounds.of.school" during her walks around the community college she attended. In her words, "you-hear=homophobia.misogyny.peoplethinkingweredumb" (p. 20) as a linguistic descriptor under an upside-down image she constructed of her college's welcoming sign. Andi attributed this positioning of the sign as a reflection of the friction she feels in attending the community college as a high school dropout" (p. 19). When we inspected her #hashtag assemblage using our literacy rhizome as an analytic, we could not detect a break or rupture in the rhizome.

Similarly, Andi's second sonic map, #SOUNDINGOUTMYSILENCE, was both an artifact and identity marker that she used in "indexing herself as a survivor of self-harm" (p. 20). In her descriptor text of the hashtag, she wrote:

you.didnt.hear.me.this.is.because.i.feel.voiceless
listen.mysilence.means.something
#selfharm.awareness.week. (p. 21)

Both assemblages viewed through a rhizomatic lens suggest that when writing assumes an agentic role, its "doing" potential is as predominant in the linguistic mode as it is in the sound mode. Still, Wargo (2018) argued persuasively, "As an audience, we hear the environment as it emerged. In short, Andi helped me discover how hearing difference is different from seeing or merely reading about it" (p. 21). Entangling these two modes, as Wargo (2018) has done, raises challenging questions—the most intriguing from our perspective being: (a) If writing is forging a new pathway in literacy circles as Brandt's (2015) seven-year research project would suggest, what added benefits will accrue to digitally literate youth, and (b), if such benefits widen the digital divide even further, what consequences might prevail?

Concluding thoughts

Although our focus in this chapter has been on the ways writing is variously described and researched by scholars in the field of literacy education, we cannot ignore the potential impact of researchers' assumptions about writing on actual classroom practice. From a pedagogical perspective, writing has historically been positioned, at least in contexts of schooled literacies (including doctoral studies), as an end product that follows reading. Writing serves as proof that reading comprehension of texts has occurred. Thus, it comes as no surprise that when reading-writing relationships are parsed—as if there exist actual beginning and ending points for the two domains—opportunities are few for experiencing movement between, among, and beyond texts.

Findings from our rhizomatic analysis of movement in reading-writing relationships are not carved in stone; quite the contrary. Accepting the unexpected (e.g., disruptions and new entanglements between writing and reading) is in keeping with Deleuze and Guattari's thinking (1987/1980). Not surprisingly, Honan and Sellers (2011) referred to rhizomatic analysis as a means of producing "(im)plausible readings" (n. p.), which was perhaps a spinoff from Spivak's (1996) earlier descriptor: "scrupulous and plausible misreadings" (p. 45). The figuration of a rhizome provided precisely what we needed to disrupt continua that signify beginnings and endings and/or cause-effect thinking in terms of reading-writing relationships. By interspersing rhizomatic maps (and the narratives they evoked) among conventionally written research reports (expositions), we were able to provide yet another example of writing's potential for 'doing' rather than simply assisting readers comprehend such reports.

In a rhizome, new shoots spring forth regularly, but these new growths do not totally abandon the rhizome, nor indicate a total rejection of shared ideas and connections. We value such emergences for their potential to erupt, rejoin, and erupt all over again. The more options the better. At the same time, we value Tierney's (2017) cautionary note regarding "meaning making in the labyrinth of multilayered text worlds" (p. 263). If meaning making manifests through internationally informed scholarship Andreotti and de Souza (2011) and draws from a broad range of human wisdom, then the rhizomatic eruptions we celebrated above are welcomed. On the other hand, if such eruptions serve to reify already established ways of thinking by dominant groups in the global western and northern hemispheres at the expense of equal representation in the eastern and southern hemispheres (Tierney et al., 2021), then we would need to reconsider the use of rhizoanalysis for learning how writing is taken up differently, say, by youth in North America (Rogers and Winters, 2010) as contrasted with youth in South Africa (Watson and Hugo, 2015). While valuing the potential for such cross-cultural differences to inform literacy research and at the same time agreeing with Tierney et al.'s call for global diversity in literacies research, we note a continuing issue that they also acknowledged in their call. Namely, as Hugo (2016) puts it, the need for radically rethinking the purposes of education more generally "in a world that is becoming increasingly unequal while global warming intensifies" (p. 11).

There is room for hope that involves scholarship in writing, though not of the kind that breeds complacency or a blind trust in writing that “does.” For is it simply that some writing “does” and some does not? Or is writing always “doing” but in ways that are not consistently recognized? We favor the latter explanation because it avoids thinking in binaries. Theory aside, writing that “does” is powerful writing, especially for those who have reliable, high-speed internet for networking; those without, regrettably, go without (Morrell and Rowsell, 2019) and pose a continuing issue for Tierney et al.’s (2021) manifesto on socioeconomic and political grounds. Based on our research, we maintain that differing perspectives, whether radical, moderate, or otherwise, can be materially dealt with when writing assumes a position of *doing* rather than *being done to*. Only then will moving away from treating writing as the end product of a reading activity achieve some degree of sustainability. Embracing writing and reading as an entangled literacy whose rhizome will continue to erupt and shoot off in different directions is movement—a condition favoring sustainable progress in the middles without the restrictions of beginnings and endings.

References

- Alexander, P.A., Fox, E., 2019. Reading research and practice over the decades: a historical analysis. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 35–64.
- Allen, L.K., Likens, A.D., McNamara, D.S., 2019. Writing flexibility in argumentative essays: multidimensional analysis. *Read. Writ.* 32 (6), 1607–1634. <https://doi.org/10.1007/s11145-018-9921-y>.
- Allen, L.K., Snow, E.L., McNamara, D.S., 2016. The narrative waltz: the role of flexibility in writing proficiency. *J. Educ. Psychol.* 108 (7), 911–924. <https://doi.org/10.1037/edu0000109>.
- Alvermann, D.E., 1993. Researching the literal: of muted voices, second texts, and cultural representations. (Presidential Address). In: Leu, D.J., Kinzer, C.K. (Eds.), *Forty-second Yearbook of the National Reading Conference* Examining Central Issues in Literacy Research, Theory, and Practice. National Reading Conference, Chicago, IL, pp. 1–10. Available at: <https://eric.ed.gov/?id=ED362855>.
- Alvermann, D.E., 2000. Researching libraries, literacies, and lives: a rhizoanalysis. In: St. Pierre, E.A., Pillow, W.S. (Eds.), *Working the Ruins: Feminist Poststructural Theory and Methods in Education*. Routledge, pp. 114–129.
- Alvermann, D.E., Moje, E.B., 2019. A relational model of adolescent literacy instruction: disrupting the discourse of “every teacher a teacher of reading.” In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 362–380.
- Alvermann, D.E., Young, J.P., Green, C., Wisenbaker, J.M., 1999. Adolescents’ perceptions and negotiations of literacy practices in after-school Read and Talk Clubs. *Am. Educ. Res. J.* 36, 221–264. <https://doi.org/10.3102/00028312036002221>.
- Andreotti, V., de Souza, L.M.T.M., 2011. *Postcolonial Perspectives on Global Citizenship Education*. Routledge.
- Bloome, D., Egan-Robertson, A., 1993. The social construction of intertextuality in classroom reading and writing lessons. *Read. Res. Q.* 28 (4), 304–333. <https://doi.org/10.2307/747928>.
- Bloome, D., Green, J., 2015. The social and linguistic turns in studying language and literacy. In: Rowsell, J., Pahl, K. (Eds.), *The Routledge Handbook of Literacy Studies*. Routledge, pp. 19–34.
- Brandt, D., 2015. *The Rise of Writing: Redefining Mass Literacy*. University of Cambridge Press.
- Brownell, C.J., 2019. Sound the alarm!: disrupting sonic resonances of an elementary English language arts classroom. *Curric. Inq.* 49 (5), 551–572.
- Brozick, J.R., Petrosky, A.R., 1979. A model for teaching writing based upon current knowledge of the composing process. *Engl. J.* 68 (1), 96–101.
- Byrd, A., Hayes, J., Turnipseed, N., 2021. Against autonomous literacies: extending the work of B. V. Street: introduction to the special issue. *Literacies in Composition Studies* 8 (2), 5–21.
- Chisholm, J.S., Trent, B., 2013. Digital storytelling in a place-based composition course. *J. Adolesc. Adult Lit.* 57 (4), 307–318.
- Deleuze, G., 1995. *Negotiations* (M. Joughin, Trans.). Columbia University Press.
- Deleuze, G., Guattari, F., 1980/1987. *A Thousand Plateaus & Schizophrenia* (B. Massumi, Trans.). University of Minnesota Press.
- de los Rios, C.V., 2020. Writing oneself into the curriculum: photovoice journaling in a secondary ethnic studies course. *Writ. Commun.* 37 (4), 487–511. <https://doi.org/10.1177/0741088320938794>.
- Dimitriadis, G., Kamberelis, G., 1997. Shifting terrains: mapping education within a global landscape. *Ann. Acad. Polit. Soc. Sci.* 551, 137–150.
- Dyson, A.H., 1991. Towards a reconceptualization of written language development. *Ling. Educ.* 3 (2), 139–161. [https://doi.org/10.1016/0898-5898\(91\)90003-2](https://doi.org/10.1016/0898-5898(91)90003-2).
- Dyson, A.H., 2018. From Superman play to singing the blues: on the trail of child writing and popular culture. *Lang. Arts* 96 (1), 37–46.
- Eakle, A.J., 2007. Literacy spaces of a Christian faith-based school. *Read. Res. Q.* 42 (4), 472–510.
- Evans, S.C., 2014. *Assemblages of Enunciation: A Rhizopoetic Analysis of Contemporary Australian Electronic Poetry*. University of Wollongong Thesis Collections.
- Featherstone, J., 1989. To make the wounded whole. *Harv. Educ. Rev.* 59, 367–378.
- Flower, L., Hayes, J.R., 1981. A cognitive process theory of writing. *Coll. Compos. Commun.* 32 (4), 365–387.
- Gee, J.P., 1990. *Social Linguistics and Literacies: Ideology in Discourses*. Taylor & Francis.
- Graham, S., Hebert, M., 2011. Writing to read: a meta-analysis of the impact of writing and writing instruction on reading. *Harv. Educ. Rev.* 81 (4), 710–744.
- Granger, N., Black, A., Miller, J., 2007. Exploring the effect of reader response plus on twelfth grade students with disabilities’ reading comprehension and attitudes toward reading. *Lang. Lit. Spectr.* 17, 14–30.
- Grosz, E., 1994. A thousand tiny sexes: feminism and rhizomatics. In: Boundas, C.V., Olkowski, D. (Eds.), *Gilles Deleuze and the Theater of Philosophy*. Routledge, pp. 187–210.
- Hagood, M., 2004. A rhizomatic cartography of adolescents, popular culture, and constructions of self. In: Leander, K.M., Sheehy, M. (Eds.), *Spatializing Literacy Research and Practice*. Peter Lang, pp. 143–160.
- Honan, E., Sellers, M., 2011. So how does it work? – Rhizomatic methodologies. <https://bit.ly/3dlrNsN>.
- hooks, b., 1990. *Yearning: race, culture and politics*. South End Press.
- Hugo, W., 2016. What education is of most worth in a world where we are consuming the future of our children? *Southern Afr. J. Environ. Educ.* 32, 11–25.
- Knobel, M., 2022. Memes and meme-ing: research and meaning. In: Sumara, D., Alvermann, D.E. (Eds.), *Ideas that Changed Literacy Practices: First-Person Accounts from Leading Voices*. Myers Education Press, pp. 193–200.
- Lankshear, C., 2022. Memes and meme-ing: rethinking internet memes for a better future. In: Sumara, D., Alvermann, D.E. (Eds.), *Ideas that Changed Literacy Practices: First-Person Accounts from Leading Voices*. Myers Education Press, pp. 211–218.
- Lawrence-Lightfoot, S., Davis, J.H., 1997. *The art and science of portraiture*. Jossey-Bass, San Francisco.
- Leander, K., Boldt, G., 2012. Rereading “A pedagogy of multiliteracies”: bodies, texts, and emergence. *J. Lit. Res.* 45 (1), 22–46.
- Leander, K.M., Rowe, D.W., 2006. Mapping literacy spaces in motion: a rhizomatic analysis of a classroom literacy performance. *Read. Res. Q.* 41 (4), 428–460.
- Morrell, E., Rowsell, J. (Eds.), 2019. *Stories from Inequity to Justice in Literacy Education: Confronting Digital Divides*. Routledge.

- Nippold, M.A., Ward-Loneragan, J.M., Fanning, J.L., 2005. Persuasive writing in children, adolescents, and adults: a study of syntactic, semantic, and pragmatic development. *Lang. Speech Hear. Serv. Sch.* 36 (2), 125–138.
- Richardson, L., 1997. *Fields of play: Constructing an academic life*. Rutgers University Press, New Brunswick, NJ.
- Risner, D., Anderson, J., 2008. Digital dance literacy: an integrated dance curriculum, pilot project. *Res. Dance Educ.* 9 (2), 113–128.
- Rogers, T., Winters, K.-L., 2010. Textual play, satire, and counter discourses of street youth zining practices. In: Alvermann, D.E. (Ed.), *Adolescents' Online Literacies: Connecting Classrooms, Digital Media, & Popular Culture*. Peter Lang, pp. 91–107.
- St. Pierre, E.A., 2013. The appearance of data. *Cult. Stud. Crit. Methodol.* 13 (4), 223–227. <https://doi.org/10.1080/00131857.2016.1151761>.
- St. Pierre, E.A., 2016. Deleuze and Guattari's language for new empirical inquiry. *Educ. Philos. Theor.* 49 (11), 1080–1089. <https://doi.org/10.1080/00131857.2016.1151761>.
- Sanders, R.K., 2018. Doctoral Student Merges Art and Literacy with GMOA Exhibition. Georgia Museum of Art, Athens, GA. <https://coe.uga.edu/news/2018/04/doctoral-student-merges-art-and-literacy-with-gmoa-exhibition>.
- Saunders, W.M., Goldenberg, C., 1999. Effects of instructional conversations and literature logs on limited- and fluent-English-proficient students' story comprehension and thematic understanding. *Elem. Sch. J.* 99 (4), 277–301.
- Smith, A., 2019. Waves of theory building in writing and its development, and their implications for instruction, assessment, and curriculum. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 65–83.
- Spivak, G., 1996. Explanation and culture: marginalia. In: Landry, D., MacLean, G. (Eds.), *The Spivak Reader*. Selected Works of Gayatri Chakravorty Spivak. Routledge.
- Stornaiuolo, A., Smith, A., Phillips, N.C., 2017. Developing a transliteracies framework for a connected world. *J. Lit. Res.* 49 (1), 68–91. <https://doi.org/10.1177/1086296X16683419>.
- Street, B.V., 1995. *Social Literacies: Critical Approaches to Literacy in Development, Ethnography, and Education*. Longman.
- Swain, S.S., Graves, R.L., Morse, D.T., 2010. Prominent feature analysis: what it means for the classroom. *Engl. J.* 99 (4), 84–89.
- Tierney, R.J., 2017. The agency and artistry of meaning makers within and across digital spaces. In: Israel, S.E. (Ed.), *Handbook of Research on Reading Comprehension*, second ed. Routledge.
- Tierney, R.J., Hingangaroa Smith, G., Kan, W., 2021. Global literacies research diversity: a manifesto for change. *J. Lit. Res.* 53 (3), 294–312. <https://doi.org/10.1177/1086296X211031262>.
- Tierney, R.J., Shanahan, T., 1991. Research on the reading-writing relationship: interactions, transactions, and outcomes. In: Barr, R., Kamil, M.L., Mosenthal, P., Pearson, P.D. (Eds.), *Handbook of Reading Research*, Vol. II. Routledge, pp. 246–280.
- Tracy, K.N., Headley, K.N., 2013. I never liked to read or write: a formative experiment on the use of a nonfiction-focused writing workshop in a fourth-grade classroom. *Lit. Res. Instr.* 52 (3), 173–191.
- Wargo, J.M., 2018. #soundingoutmysilence: reading a LGBTQ youth's sonic cartography as multimodal (counter)storytelling. *J. Adolesc. Adult Lit.* 62 (1), 13–23.
- Wargo, J.M., Brownell, C.J., Oliveira, G., 2021. Sound, sentience, and schooling: writing the field recording in educational ethnography. *Anthropol. Educ. Q.* 52 (3), 315–334. <https://doi.org/10.1111/aeq.12365>.
- Watson, A.P., Hugo, W., 2015. Adolescent literacies: chatting and learning across different planes of composition. *Educ. Change* 19 (1), 120–147.
- Wong, B.L., Kuperis, S., Jamieson, D., 2002. Effects of guided journal writing on students' story understanding. *J. Educ. Res.* 95 (3), 179–191.
- Yaden, D.B.J., 2004. Reconstructing the state space of parent-child storybook reading conversations: a brief synopsis. In: *Proceedings of the 2004 Complexity Science and Educational Research Conference*, pp. 111–119. In: <http://www.complexityandeducation.ualberta.ca/conferences/2004/proceedings>.

FOURTH EDITION

International Encyclopedia of Education

The Rise of STEM Education

VOLUME
11

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editors

Xiufeng Liu
Lei Wang



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 11

The Rise of STEM Education

VOLUME 11 EDITORS

Xiufeng Liu and Lei Wang



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

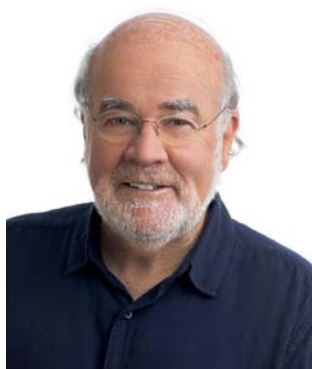
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019).

Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



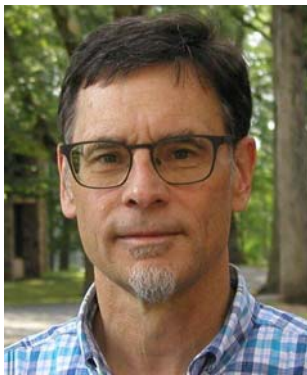
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 11: The Rise of STEM Education

Xiufeng Liu and Lei Wang

© 2023 Elsevier Ltd. All rights reserved.

Introduction

Science, Technology, Engineering, and Mathematics, or STEM, is currently a buzz word in all sectors of the society such as mass media, business, industry, government, and education in the United States and around the world. The origin of the acronym of STEM is believed to be attributed to the US National Science Foundation (NSF) that used STEM to group its program areas in education and workforce development (Hallinen, nd). Although another acronym, SMET (Science, Mathematics, Engineering, and Technology) was already in use at NSF before STEM was in common use, in 2001 the assistant director of NSF's Education and Human Resources Directorate, Judith Ramaley, suggested that STEM would be better aesthetically and conceptually (Patton, 2013). Since then, NSF has been consistently using STEM to refer to a convenient cluster of programs in its publications and budget requests to the Congress. With its inroad into public legislation and mass media, STEM workforce development with a focus on STEM's critical role for national economic development and international competitiveness became a national priority, and STEM education began to emerge as a new field of study and practice in the United States. Over the years, STEM education has become a recurring theme for education policies, workforce development, and educational research in the United States and has gradually become a world-wide movement (e.g., European Schoolnet, 2018; Office of the Chief Scientist, 2014; Center for STEM Education, 2018; Royal Academy of Engineering, 2016).

Despite its popularity and importance of STEM education, there remains no coherent conceptualization or universally agreed upon definition of STEM education (Holmlund et al., 2018). One notion accepted by many in STEM education community is the integrated nature of studies across individual disciplines of science, technology, engineering, and mathematics, given that science, technology, engineering, and mathematics are established school subjects, but interdisciplinary learning is missing. There is considerable overlap in conceptual ideas and practices between STEM education and disciplinary education of science, technology, engineering, and mathematics (Moore et al., 2020). Although STEM education emphasizes integration, STEM is not a single subject for preK-12; it does not replace individual school subjects of science, technology, engineering, and mathematics. STEM education research encompasses many areas of inquiry, such as curriculum, instruction, assessment, teacher education, educational leadership, informal education, technology integration, educational policy, career development, and relations with individual school disciplines of science, technology, engineering, and mathematics.

In education, the idea of integration is not new. Project-based learning, commonly employed in STEM teaching to promote integrated learning, has its origin in Dewey's ideas of experiential, hands-on, and student-directed learning (Maida, 2011). For Dewey, learning occurs within a social environment and knowledge is constructed through social interaction and reflection. Student learning experiences should be relevant to student everyday lives and lead to continued learning in the future (Dewey, 1938). In order for the above learning to take place, learning needs to be integrated.

The longest history of promoting integration is the integration of mathematics and science by the School Science and Mathematics Association (SSMA). SSMA was formed in 1902 with a mission to promote the integration of science and mathematics. The association's official journal, *School Science and Mathematics*, is "emphasizing research on issues, concerns, and lessons within and between the disciplines of science and mathematics in the classroom" (<https://www.ssma.org/school-science-and-mathematics-journal>). For example, *School Science and Mathematics* journal ran a feature column/department called School Science and Mathematics Integrated Lessons (SSMILes) from January 1989 all the way through the 2000s. There have been many other attempts to promote integrated learning in STEM over the past decades. In the 1970s, UNESCO initiated an international movement to promote integrated science teaching and learning in K-12 schools, with a series of publications entitled "New Trends in Integrated Science Teaching, Vols. 1-7" (Unesco, 1971, 1973, 1974, 1977, 1979). The world-wide Science-Technology-Society (STS) education movement in the 1980s and 1990s was another prominent example of integrated education (Solomon and Aikenhead, 1994). Recently, the US Next Generation Science Standards (The NGSS Lead states, 2013; NRC, 2012) promotes a strong notion of integrated science teaching and learning through a three-dimensional integration of core disciplinary ideas, science and engineering practices, and crosscutting concepts. The inclusion of engineering into science education also represents a bold vision for the future of science education in the United States. Finally, for each performance expectation, NGSS makes explicit connections to Common Core curriculum standards in which math and English language arts learning expectations are specified.

In 2014, the National Research Council published a consensus report titled "Integration in K-12 Education: Status, Prospects, and an Agenda for Research." The report presents a framework of STEM education in terms of four features: goals of integrated STEM education, outcomes of integrated STEM education, nature and scope of integrated STEM education, and implementation of integrated STEM education. In terms of goals of integrated STEM education, the following goals were identified: (a) STEM literacy, (b) 21st century competencies, (c) STEM workforce readiness, (d) interest and engagement, and (e) making connections. In terms of outcomes of integrated STEM education, the following outcomes were identified: (a) learning and achievement, 21st century competencies, (b) STEM course taken, educational persistence and graduation rates, (c) STEM related employment, (d) STEM interest, (e) Development of STEM identity, and (f) ability to make connections among STEM disciplines. In terms of nature and scope of integrated STEM education, the following aspects were identified: (a) types of STEM connections, (b) disciplinary emphasis, and (c) duration, size, and complexity of initiatives. Finally, for the implementation of integrated STEM education,

the following components were identified: (a) instructional design, (b) education supports, and (c) adjustments to the learning environment. In addition to the above features for students, the report also identified corresponding features for teachers. Specifically, goals of integrated STEM education for teachers include increased STEM content knowledge, increased pedagogical content knowledge, and changes in teaching practice. The framework provides “a vocabulary for researchers, practitioners, and others to identify, describe, and investigate specific integrated STEM initiatives in the US K-12 education system” (p. 31, NRC, 2014).

Scholarship on STEM education has been taking place since early 2000s. The *Journal of STEM Education: Innovations and Research* aims “to promote high-quality education at all levels in science, mathematics, engineering, and technology” (<https://www.jstem.org/jstem/index.php/JSTEM/about>). It emphasizes real-world case studies that focus on issues relevant and important to STEM practitioners. The *International Journal of STEM Education* was established in August 2014 (<https://stemeducationjournal.springeropen.com/>). It is a multidisciplinary journal in subject-content education that focuses on the study of teaching and learning in STEM. On the other hand, the *Journal for STEM Education Research*, launched in December 2018, “is an interdisciplinary journal in subject content education focusing on research in STEM (Science, Technology, Engineering, and Mathematics) education, which promotes the development of STEM education research as a distinct field” (<https://www.springer.com/journal/41979>). Most recently, *Disciplinary and Interdisciplinary Science Education Research* (DISER) was launched in November 2019; it promotes scholarship and best practices in education within and across science disciplines (<https://diser.springeropen.com/about>).

There have been various conceptualizations of integration (Czerniak and Johnson, 2020). There is no universally agreed definition or conceptualization of integration, and it is very difficult to derive such one. For example, for the integration of mathematics and science, a NSF-supported 3-day conference with the participation of 60 individuals representing scientists, mathematicians, science, and math educators, as well as educational psychologists and curriculum experts, was unable to reach a consensus on a definition of integrated mathematics and science (Berlin and White, 1992). Based on the conference discussion, the Berlin-White Integrated Science and Mathematics model identified six aspects that may be used to categorize various notions of integrated mathematics and science; they are (a) ways of learning, (b) ways of knowing, (c) process and thinking skills, (d) content knowledge, (e) attitudes and perceptions, and (f) teaching strategies (White, 1994).

Petrie (1992), from a philosophy of science perspective, distinguishes among disciplinary, multidisciplinary, interdisciplinary, and transdisciplinarity. He claims that interdisciplinarity cannot be understood unless we understand the concept of disciplinarity. According to him, disciplinarity is characterized as (a) a specialized knowledge, (b) unity of a common set of concepts, specialized methods, and (c) an organized group of people who study the discipline, train other practitioners, and form the social mechanism for deciding among varying truth claims within the discipline. Compared to disciplinarity, multidisciplinary is characterized as “the idea of a number of disciplines working together on a problem, an educational program, or a research study. The effect is additive rather than integrative” (p. 303). Multidisciplinary is “group work” rather than “team work.” Interdisciplinarity is characterized as “the integration of the work goes beyond the mere concatenation of disciplinary contributions” (p. 304); it is “team work.” Finally, transdisciplinarity is characterized as the integration of knowledge in meaningful ways across many disciplines. Accordingly, integration may be considered as a continuum from disciplinary to multidisciplinary to interdisciplinary to transdisciplinary (Ji et al., 2015). In this continuum, multidisciplinary is the integration in which ideas or approaches from other disciplines are drawn to solve disciplinary problems; interdisciplinary is the integration in which different disciplines form a coherent approach to solving a real-world problem while maintaining the boundaries of individual disciplines; and transdisciplinary is the integration in which problems, ideas and approaches transcend disciplines by blurring disciplinary boundaries.

Based on the above frameworks of STEM education and the conceptualization of integration, articles in this volume are organized into six sections. Section 1 includes 10 articles addressing various theoretical perspectives of STEM education. These perspectives include nature of interdisciplinary learning, characteristics of Science, Technology, Engineering, Art and Mathematics (STEAM) education, mathematics and engineering in STEM education, language, multimodalities and multiliteracies, creativity, spatial thinking, and computational thinking. Section 2 presents approaches to STEM education. It includes 16 articles on inquiry teaching, makerspace and robotics, models and modeling, problem-based learning, place-based learning, integration of socioscientific issues, play-based STEM learning, STEM schools, virtual reality, STEM curriculum development and implementation, designed-based education, technology-enhanced STEM learning, citizen science and informal science learning, and community engagement in STEM. Section 3 deals with STEM teacher education. It includes nine articles related to teacher beliefs and perceptions of STEM and STEM education, preservice- and inservice-teacher education, and scholarship of teaching for reforming university STEM teaching. Section 4 addresses the equity and inclusion in STEM Education. It includes four articles on universal design for STEM education, racial and gender equity in STEM education, and integrating indigenous knowledge in STEM education. Section 5 addresses topics on student STEM career aspirations. This section includes six articles on STEM workforce development, family influence in STEM career aspiration, student motivation, self-efficacy and career aspiration, and STEM career positioning. The final section deals with measurement and evaluation of STEM education. It includes four articles on assessment of integrated STEM learning, measurement of STEM affective learning, assessment of sensemaking, and neurocognitive and physiological measurement of STEM learning processes.

The papers in this volume follow no particular format; they are in forms of a literature review, historical account, empirical study, or summary of their own recent research; they approach a topic from diverse theoretical and methodological perspectives. In accordance with the intention of the fourth edition of the Encyclopedia to be international, this volume includes contributions from around the world to provide global conversations about the diverse conceptions, purposes, policies, and practices of STEM education. Authors are from Australia, Canada, China, Finland, Hong-Kong (China), Ireland, Israel, Sweden, and the United States. In order to reflect the latest development of STEM education, the authors were identified mainly from conference programs of the

National Association for Research in Science Teaching (NARST), a world-wide organization for improving science teaching and learning through research, American Educational Research Association (AERA), European Science Education Research Association (ESERA) from 2019 to 2021. We hope collectively the contributions provide a comprehensive introduction to STEM education world-wide.

References

- Berlin, D.F., White, A.L., 1992, October. Report from the NSF/SSMA wingspread conference: a network for integrated science and mathematics teaching and learning. *Sch. Sci. Mathematics* 92 (6), 340–342.
- Center for STEM Education, National Institute of Education Sciences, 2018. Agenda for STEM Education in China by 2029. Center for STEM Education, National Institute of Education Sciences, Beijing.
- Czerniak, C.M., Johnson, C.C., 2020. Interdisciplinary science teaching. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, vol. II. Routledge, New York, pp. 395–411. Chapter 20.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York.
- European Schoolnet, 2018. Science, Technology, Engineering and Mathematics Education Policies in Europe. Scientix Observatory Report. European Schoolnet, Brussels. http://www.scientix.eu/documents/10137/782005/Scientix_Texas-Instruments_STEM-policies-October-2018.pdf/d56db8e4-cef1-4480-a420-1107bae513d5. (Accessed 22 April 2022).
- Hallinen, J. (n.d.). STEM. *Encyclopedia Britannica*. <https://www.britannica.com/topic/STEM-education>. Accessed April 22, 2022.
- Holmlund, T., Lesseig, K., Slavt, D., 2018. Making sense of “STEM education” in K-12 contexts. *Int. J. STEM Educ.* 5, 32.
- Ji, S., Sung, S., Zhang, D., 2015. Toward an analytic framework of interdisciplinary reasoning and communication (IRC) processes in science. *Int. J. Sci. Educ.* 37 (17), 2809–2835.
- Maida, C., 2011. Project-based learning: a critical pedagogy for the twenty-first century. *Policy Fut. Educ.* 9 (6), 759–768.
- Moore, T.J., Johnston, A.C., Glancy, A.W., 2020. STEM integration: a synthesis of conceptual frameworks and definitions. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T., English, L. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 3–16.
- National Research Council [NRC], 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The National Academies Press, Washington, DC.
- National Research Council [NRC], 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. The National Academies Press, Washington, DC.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, By States*. The National Academies Press, Washington, DC.
- Office of the Chief Scientist, 2014. *Science, Technology, Engineering and Mathematics: Australia’s Future*. Commonwealth of Australia, Canberra.
- Patton, M., 2013. ATE Had Role in Naming of STEM. <https://atecentral.net/impacts/22917/ate-had-role-in-the-naming-of-stem>. (Accessed 22 April 2022).
- Petrie, H.G., 1992. Interdisciplinary education: are we faced with insurmountable opportunities? *Rev. Res. Educ.* 18, 299–333.
- Royal Academy of Engineering, 2016. *The UK STEM Education Landscape: A Report for the Lloyd’s Register Foundation From the Royal Academy of Engineering Education and Skills Committee*. <https://www.raeng.org.uk/publications/reports/uk-stem-education-landscape>. (Accessed 22 April 2022).
- Solomon, J., Aikenhead, G. (Eds.), 1994. *STS Education: International Perspectives on Reform*. Teachers College Press, New York.
- UNESCO, 1971, 1971, 1973, 1974, 1977, 1979. *New trends in integrated science teaching*, vol. 1–7. UNESCO, Paris.
- White, A.L., 1994. The Berlin-White integrated science and mathematics model. *Sch. Sci. Mathematics* 94 (1), 2–4.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 11

Brian Abramowitz

Thompson Earth Systems Institute, Department of Natural History, University of Florida, Gainesville, FL, United States

Nagarajan Akshay

University of California in San Diego, Department of Cognitive Science, La Jolla, CA, United States

Peter Appelbaum

Arcadia University, Philadelphia, PA, United States

Ryan Batkie

Department of Curriculum and Instruction, University of Wisconsin—Madison, Madison, WI, United States

Kathleen E Bordewieck

College of Education, North Carolina State University, Raleigh, NC, United States

Sheryl E Burgstahler

University of Washington, Seattle, WA, United States

Jinfa Cai

University of Delaware, Newark, DE, United States

Purvee Chauhan

Department of Mechanical and Aerospace Engineering, NYU Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY, United States

IC Chen

Create for STEM Institute, University of Michigan State, East Lansing, MI, United States

Gi Woong Choi

Instructional Design and Technology Program, School of Education, University of Cincinnati, Cincinnati, OH, United States

Heidi D Cian

STEM Transformation Institute, Department of Teaching & Learning, Florida International University, Stanford, CA, United States

Heather F Clark

UCLA School of Education & Information Studies, Los Angeles, CA, United States

SK Codere

Create for STEM Institute, University of Michigan State, East Lansing, MI, United States

Keren E Dalyot

Faculty of Education in Science and Technology, The Technion - Israeli Institute of Technology, Haifa, Israel

Emily A Dare

Department of Teaching and Learning, STEM Transformation Institute, Florida International University, Miami, FL, United States

Thomas Delahunty

Department of Education, Maynooth University, Maynooth, Co. Kildare, Ireland

Adam Devitt

Department of Teacher Education California State University, Stanislaus, Turlock, CA, United States

Meixia Ding

Temple University, Philadelphia, PA, United States

Remy Dou

STEM Transformation Institute, Department of Teaching & Learning, Florida International University, Miami, FL, United States

Daniel Douglas

Trinity College, Department of Sociology, Hartford, CT, United States

Julie Dunphy

School of Education, Arcadia University, Glenside, PA, United States

Jacob Elmore

University of Virginia, School of Education and Human Development, Charlottesville, VA, United States

Megan E Ennes

*Department of Natural History, University of Florida,
Gainesville, FL, United States*

Martha Escobar

*Department of Psychology, Oakland University,
Rochester, MI, United States*

Emily K Faulconer

*Department of STEM Education, Worldwide Campus,
Embry-Riddle Aeronautical University, Daytona Beach,
FL, United States*

Monika Fridrich

*Department of Teaching, Learning, and Educational
Leadership, Binghamton University – SUNY,
Binghamton, NY, United States*

Maureen Giannotti

*UCLA Center X School of Education & Information
Studies, Los Angeles, CA, United States*

Florence Glanfield

University of Alberta, Edmonton, Canada

Yaela N Golumbic

*Faculty of Science, The University of Sydney, NSW,
Australia*

John Griffith

*Department of STEM Education, Worldwide Campus,
Embry-Riddle Aeronautical University, Daytona Beach,
FL, United States*

Chorng-je Guo

*National Changhua University of Education, Taichung,
Taiwan*

Breonte S Guy

*Winston-Salem State University, Winston-Salem, NC,
United States*

Roxana Hadad

*UCLA Center X School of Education & Information
Studies, Los Angeles, CA, United States*

Jonas Hallström

*Technology and Science Education Research (TESER),
Department of Behavioural Sciences and Learning,
Linköping University, Norrköping, Sweden*

Kathryn Holmes

*Center for Educational Research, School of Education,
Western Sydney University, Kingswood, NSW, Australia*

Sameer Honward

University at Buffalo, NY, United States

Weiping Hu

*MOE Key Laboratory of Modern Teaching Technology,
Shaanxi Normal University, Xi'an, PR China*

Aida Hurem

*School of Education and Professional Studies, Griffith
University, Southport, QLD, Australia*

Kimberly L Ideus

*College of Education, North Carolina State University,
Raleigh, NC, United States*

David F Jackson

*Department of Mathematics, Science, and Social Studies
Education, University of Georgia, Athens, GA, United
States*

Kamini Jaipal-Jamani

*Brock University, Faculty of Education, St. Catharines,
ON, Canada*

Joseph A Johnson

*Department of Physics, Mercyhurst University, Erie, PA,
United States*

M Gail Jones

*The Friday Institute for Educational Innovation,
Department of STEM Education, North Carolina State
University, Raleigh, NC, United States*

Yong Ju Jung

*School of Library and Information Studies, University of
Oklahoma, Norman, OK, United States*

Jingoo Kang

*School of Applied Educational Science and Teacher
Education, University of Eastern Finland, Joensuu,
Finland*

Vikram Kapila

*Department of Mechanical and Aerospace Engineering,
NYU Tandon School of Engineering, Six MetroTech
Center, Brooklyn, NY, United States*

Eunjeong Kim

*Korea Foundation for the Advancement of Science &
Creativity, Seoul, Korea*

Soo Hyeon Kim

*Department of Library and Information Science, School
of Informatics and Computing, Indiana University-
Purdue University Indianapolis (IUPUI), Indianapolis,
IN, United States*

Young Ae Kim

University of Arizona, Tucson, AZ, United States

Neil A Knobloch

*Purdue University, West Lafayette, IN,
United States*

Joseph Krajcik

*Michigan State University, East Lansing, MI, United
States*

Richard Lamb

*East Carolina University, College of Education,
Neurocognition Science Laboratory, Greenville, NC,
United States*

Promail KY Leung

*Faculty of Education, The University of Hong Kong, Pok
Fu Lam, Hong Kong, China*

Siqi Li

*College of Education for the Future, Beijing Normal
University at Zhuhai, Zhuhai, Guangdong, China; and
Research Institute of Science Education, Beijing Normal
University, Beijing, China*

T Li

*Create for STEM Institute, University of Michigan State,
East Lansing, MI, United States*

Xinyi Li

*MOE Key Laboratory of Modern Teaching Technology,
Shaanxi Normal University, Xi'an, PR China*

Julie A Luft

*Mary Francis Early College of Education, University of
Georgia, Athens, GA, United States*

Tian Luo

*College of Teacher Education, Capital Normal
University, Beijing, P.R. China*

Augusto Macalalag

*School of Education, Arcadia University, Glenside, PA,
United States*

Haylee M Majewski

*Department of Psychology, Oakland University,
Rochester, MI, United States*

Mark A McDermott

University of Iowa, Iowa City, IA, United States

Mary B McVee

*University at Buffalo, SUNY, Amherst, NY, United
States*

Todd M Milford

University of Victoria, Victoria, BC, Canada

Emily Adah Miller

University of Wisconsin, Madison, WI, United States

Catherine Milne

*Department of Teaching and Learning, New York
University, New York, NY, United States*

Victor Hugo Minces

*University of California in San Diego, Department of
Cognitive Science, La Jolla, CA, United States*

Imelda Nava

*UCLA Center X School of Education & Information
Studies, Los Angeles, CA, United States*

Cynthia Nicol

University of British Columbia, Vancouver, Canada

Lisa O'Keeffe

University of South Australia, Adelaide, SA, Australia

J Steve Oliver

The University of Georgia, Athens, GA, United States

Leticia Perez

*UCLA Center X School of Education & Information
Studies, Los Angeles, CA, United States*

Chris Proctor

*Department of Learning and Instruction, University at
Buffalo, SUNY, Buffalo, NY, United States*

Senay Purzer

Purdue University, West Lafayette, IN, United States

Mohammed Qazi

*Department of Mathematics, Tuskegee University,
Tuskegee, AL, United States*

Yasmeen Rawajfih

*Department of Computer Science, Tuskegee University,
Tuskegee, AL, United States*

David Reilly

School of Applied Psychology, Southport, QLD, Australia

Kelly Riedinger

*STEM Research Center, Oregon State University,
Corvallis, OR, United States*

Elizabeth A Ring-Whalen

*Department of Education, National Center for STEM
Elementary Education, St. Catherine University, St.
Paul, MN, United States*

Audrey Conway Roberts

*School of Educational Foundations, Leadership, and
Policy, College of Education and Human Development,
Bowling Green State University, Bowling Green, OH,
United States*

Thomas Roberts

*School of Teaching and Learning, College of Education
and Human Development, Bowling Green State
University, Bowling Green, OH, United States*

Gillian H Roehrig

*Department of Curriculum and Instruction, University
of Minnesota, Learning and Environmental Sciences, St.
Paul, MN, United States*

Lilith Rüschenpöhler

Ludwigsburg University of Education, Ludwigsburg, Germany

Hal Salzman

Rutgers University, Edward J. Bloustein School of Planning & Public Policy, John J. Heldrich Center for Workforce Development, New Brunswick, NJ, United States

Konrad J Schönborn

Division of Media and Information Technology, Department of Science and Technology, Linköping University, Norrköping, Sweden

Kelly A Schucker

University at Buffalo, SUNY, Amherst, NY, United States

Samuel Severance

University of California, Santa Cruz, CA, United States

Ji Shen

Department of Teaching and Learning, University of Miami, Coral Gables, FL, United States

Katarina N Silvestri

SUNY Cortland, Cortland, NY, United States

Amber Simpson

Department of Teaching, Learning, and Educational Leadership, Binghamton University – SUNY, Binghamton, NY, United States

Wing-mui Winnie So

The Education University of Hong Kong, Tai Po, Hong Kong, P.R. China

Nancy Staus

STEM Research Center, Oregon State University, Corvallis, OR, United States

Martin Storksdieck

STEM Research Center, Oregon State University, Corvallis, OR, United States

Jennifer S Thom

University of Victoria, Victoria, Canada

Christine D Tippet

University of Ottawa, Ottawa, ON, Canada

Shawn Titus

Department of Physics, Mercyhurst University, Erie, PA, United States

Jennifer N Tripp

University at Buffalo, SUNY, Department of Learning and Instruction, Buffalo, NY, United States; and University at Buffalo, SUNY, Amherst, NY, United States

Rachel van Aswegen

University of Virginia, School of Education and Human Development, Charlottesville, VA, United States

Changzhao Wang

Department of Teaching and Learning, University of Miami, Coral Gables, FL, United States

Hui-Hui Wang

Purdue University, West Lafayette, IN, United States

Jian Wang

College of Life Sciences, Beijing Normal University, Beijing, P.R. China

Bruce White

University of South Australia, Adelaide, SA, Australia

Jeanna R Wieselmann

Department of Teaching and Learning, Southern Methodist University, Dallas, TX, United States

Valerie WY Yip

Faculty of Education, The University of Hong Kong, Pok Fu Lam, Hong Kong, China

Peter Youngs

University of Virginia, School of Education and Human Development, Charlottesville, VA, United States

Yvonne Zeegers

University of South Australia, Adelaide, SA, Australia

Xiaoming Zhai

Department of Mathematics, Science, and Social Studies Education, University of Georgia, Athens, GA, United States

CONTENTS OF VOLUME 11

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 11: The Rise of STEM Education</i>	<i>xix</i>
<i>Contributors to Volume 11</i>	<i>xxiii</i>

THEORETICAL FOUNDATIONS OF STEM EDUCATION

Interdisciplinary science learning <i>Ji Shen and Changzhao Wang</i>	1
STEAM for all: a vision for STEM and arts integration <i>Victor Hugo Minces and Nagarajan Akshay</i>	10
Mathematics in STEM education <i>Meixia Ding and Jinfa Cai</i>	19
The rise of STEM education: mathematics learning for STEM <i>Peter Appelbaum</i>	28
Engineering education <i>Senay Purzer</i>	36
The multimodal affordances of language in STEM education <i>Christine D Tippett, Mark A McDermott, and Todd M Milford</i>	42
Using multiliteracies and multimodality to explore engineering literacies and problem-scoping for multilingual learners: A perspective on multiliteracies for STEM learning <i>Mary B McVee, Kelly A Schucker, Katarina N Silvestri, and Jennifer N Tripp</i>	50
Creativity in STEM education <i>Weiping Hu and Xinyi Li</i>	64
Spatial thinking and STEM education <i>Siqi Li</i>	76
Computational thinking <i>Chris Proctor</i>	88

APPROACHES TO STEM EDUCATION

Inquiry-based teaching and STEM education <i>Yvonne Zeegers, Lisa O'Keeffe, and Bruce White</i>	96
Makerspace and robotics as/for STEM education <i>Kamini Jaipal-Jamani</i>	103
Models and modeling in STEM education: nature, roles, and implementation <i>Jonas Hallström and Konrad J Schönborn</i>	112
Aligning PBL with the NGSS: What do the NGSS offer PBL? <i>Emily Adah Miller, Samuel Severance, and Joseph Krajcik</i>	117
Imagining STEM with and as place <i>Cynthia Nicol, Jennifer S Thom, and Florence Glanfield</i>	131
Socioscientific issues and STEM learning <i>Joseph A Johnson, Ryan Batkie, Augusto Macalalag, Julie Dunphy, and Shawn Titus</i>	143
STEM curriculum development and implementation <i>Gillian H Roehrig, Emily A Dare, Jeanna R Wieselmann, and Elizabeth A Ring-Whalen</i>	153
Connecting play to STEM concepts, practices and processes: review of research on play within STEM learning environments <i>Amber Simpson and Monika Fridrich</i>	164
STEM schools: their rise, present state, and future potential <i>Jennifer N Tripp</i>	177
Virtual reality and science, technology, engineering, and mathematics education <i>Richard Lamb</i>	189
Design-based education in STEM: for learners of the 21st century <i>Yong Ju Jung, Gi Woong Choi, and Soo Hyeon Kim</i>	198
Technology-enhanced collaborative learning in STEM <i>Changzhao Wang and Ji Shen</i>	207
A pedagogical framework for mobile learning in science education <i>Xiaoming Zhai and David F Jackson</i>	215
Citizen science in STEM education: engaging students with real life science <i>Keren E Dalyot and Yaela N Golumbic</i>	224
Community engagement in STEM education <i>Wing-mui Winnie So and Chorng-jee Guo</i>	234
Informal STEM learning <i>Nancy Staus, Kelly Riedinger, and Martin Storksdieck</i>	244

STEM TEACHER EDUCATION

K-12 teachers' beliefs and practices in STEM integration <i>Hui-Hui Wang and Neil A Knobloch</i>	251
Promises and perils of STEM education: synthesizing teacher, student, & research perceptions <i>Thomas Roberts and Audrey Conway Roberts</i>	262

Pre-service STEM teacher education <i>Valerie WY Yip and Promail KY Leung</i>	270
In-service teacher education for STEM <i>Julie A Luft</i>	276
International perspectives on STEM teacher induction <i>Jacob Elmore, Rachel van Aswegen, and Peter Youngs</i>	283
Design-based pedagogy: a heuristic for developing math and science teachers' professional practice <i>Adam Devitt and Catherine Milne</i>	292
Promoting engineering research with entrepreneurship and industry experiences: a teacher professional development program <i>Purvee Chauhan and Vikram Kapila</i>	298
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program <i>Leticia Perez, Heather F Clark, Roxana Hadad, Imelda Nava, and Maureen Giannotti</i>	312
The scholarship of teaching and learning: cases from one STEM department <i>Emily K Faulconer and John Griffith</i>	321
EQUITY AND INCLUSION IN STEM EDUCATION	
Universal design in STEM education <i>Sheryl E Burgstahler</i>	326
Designing equitable STEM education: Guidelines for parents, educators and policy-makers to reduce gender/racial achievement gaps <i>David Reilly and Aida Hurem</i>	334
Gender and STEM career aspiration <i>Jingoo Kang</i>	344
Learning to balance indigenous and exogenous knowledge systems for environmental decision-making in the Kumaon Himalayas <i>Sameer Honward</i>	349
STEM CAREER ORIENTATION	
STEM education and workforce development: the history, politics, and evidence <i>Hal Salzman and Daniel Douglas</i>	358
Family influence and STEM career aspirations <i>Megan E Ennes, M Gail Jones, Heidi D Cian, Remy Dou, Brian Abramowitz, Kathleen E Bordewieck, and Kimberly L Ideus</i>	370
Secondary school students' STEM career aspirations <i>Kathryn Holmes</i>	382
Self-efficacy in STEM <i>Martha Escobar, Haylee M Majewski, Mohammed Qazi, and Yasmeen Rawajfih</i>	388

Positioning diverse students for science, technology, engineering, and mathematics (STEM) careers: reinforced learning ecosystems for STEM success	395
<i>Breonte S Guy</i>	

Student motivation in STEM: factors related to and measurement of STEM motivation	401
<i>Tian Luo, Lilith Rüschenpöhler, and Jian Wang</i>	

MEASUREMENT AND EVALUATION OF STEM EDUCATION

Integrated STEM curriculum and its assessment	409
<i>Young Ae Kim, Eunjeong Kim, and J Steve Oliver</i>	

Measurement instruments of STEM affective learning: a systematic review	421
<i>Jeanna R Wieselmann, Emily A Dare, Gillian H Roehrig, and Elizabeth A Ring-Whalen</i>	

Using flexible thinking to assess student sensemaking of phenomena in project-based learning	444
<i>EC Adah Miller, T Li, IC Chen, and SK Codere</i>	

Neurocognitive and physiological measurment of STEM learning processes	458
<i>Thomas Delahunty</i>	

Interdisciplinary science learning

Ji Shen and Changzhao Wang, Department of Teaching and Learning, University of Miami, Coral Gables, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Nature of interdisciplinary science learning	1
Interdisciplinary learning	1
Knowledge categorization and practices in interdisciplinary science learning	1
Levels of boundary crossing	2
Connecting disciplines within the natural sciences	2
Connecting science across STEM fields	3
Hierarchical mode	3
Flat mode	4
Connecting science with non-STEM fields	5
Promoting interdisciplinary science learning	6
Strategies for facilitating interdisciplinary science learning	6
Practices in interdisciplinary science learning	6
Challenges for creating and sustaining interdisciplinary science learning	7
Conclusion and recommendation	7
References	7

Nature of interdisciplinary science learning

Interdisciplinary learning

The high demand of working and thinking across disciplinary boundaries in frontier research and development mandates a rethinking of STEM education (e.g., Erduran, 2020; Shen et al., 2015). Interdisciplinary learning, in a broad sense, refers to the type of learning that cuts across disciplinary boundaries (Boix Mansilla and Duraing, 2007; Klein, 1996). It can promote deeper understanding and long-term effectiveness since it enables students to integrate and apply knowledge across disciplinary boundaries (Ivanitskaya et al., 2002; McComas and Wang, 1998; Steen, 2005). An interdisciplinary learning approach can also help students make connections to the real world and prepare them for better employment in the new century (National Academies of Sciences, Engineering, and Medicine [NASEM], 2018).

Different terms have been used to describe interdisciplinary learning approaches (National Academy of Engineering [NAE] and National Research Council [NRC], 2014). An *integrated* approach stresses blending different subject matters to form a seamless whole and as a result, no clear distinction exists among the disciplines (Lederman and Niess, 1997); A *transdisciplinary* approach focuses on addressing complex problems or issues at hand without any *a priori* disciplinary lenses (Klein, 2008); A *multidisciplinary* approach emphasizes clear disciplinary boundaries and traditions when combining ideas and methods without imposing a unified framework (Drake and Burns, 2004); A *cross-disciplinary* approach refers to the range of activities that cuts across multiple disciplines, without necessarily blending in disciplinary ideas or techniques (Cross-disciplinary Learning, 2012). These interpretations of the specific terms are not necessarily universal and may have variations across contexts (NASEM, 2018).

Acknowledging the overlaps among all these terms, we choose the term *interdisciplinarity* in this article to emphasize the space where different disciplines connect and interact. We submit that individual disciplines are the building blocks for interdisciplinary work (Tripp and Shortlidge, 2020; Walshe et al., 2017; Zhang and Shen, 2015). Therefore, it matters which disciplines are involved and which disciplines are prioritized when an interdisciplinary learning approach is adopted.

Knowledge categorization and practices in interdisciplinary science learning

We consider the nature of interdisciplinary learning from two perspectives: knowledge categorization and learner practices. During interdisciplinary learning, a learner is situated in a boundary space where two or more disciplinary fields intersect. As such, interdisciplinary learning may involve different levels of and processes in disciplinary boundary crossing. Since we are focusing on interdisciplinary learning related to science (in this article, the term *science* mainly refers to the natural sciences unless otherwise noted), boundary crossing may involve different levels of knowledge categorization with respect to how science is involved: within sciences, across STEM fields, or between science and non-STEM fields.

Despite the different types of boundary crossing, what lies in the center of interdisciplinary learning is the interdisciplinary practices: what learners are actually doing in order to connect the disciplines. In this article, we adopt the framework proposed by Shen et al. (2015) in which four interrelated interdisciplinary reasoning and communication practices are essential in interdisciplinary learning:

- Translation: Map vocabularies and terms between/among disciplines in order to understand interdisciplinary topics or effectively communicate and collaborate with others from different disciplinary backgrounds.
- Transfer: Apply concepts, theories, models, tools, and methods learned from one discipline to explain phenomena and processes in another, recognizing the similar system under study.

- Integration: Connect concepts, theories, models, tools, and methods learned from different disciplines into a coherent whole to understand natural phenomena and solve complex problems.
- Transformation: Use concepts, theories, models, tools, and methods learned from one or more disciplines to physically or conceptually change or create a system or entity typically considered in another discipline or a new context.

Fig. 1 illustrates aforementioned conceptual framing of interdisciplinary science learning. The next sections are organized around the three levels of boundary crossing (i.e., within sciences, across STEM, with non-STEM). In each section, we use different examples to illuminate insights and issues related to interdisciplinary science learning while highlighting different aspects of the interdisciplinary practices (i.e., translation, transfer, integration, and transformation). Although we introduce each of the practices in different places, the reader should keep in mind that all of these practices are closely related to each other and often occur simultaneously in an interdisciplinary learning context.

Levels of boundary crossing

Connecting disciplines within the natural sciences

Making connections across disciplines in learning has been promoted within the sciences ([American Association for the Advancement of Science \[AAAS\], 2011](#); [Czerniak and Johnson, 2014](#); [NRC, 2005](#)). The report “Vision and Change in Undergraduate Biology Education” ([AAAS, 2011](#)) lays out six core competencies and disciplinary practices in biology education that also speak to the importance of interdisciplinary thinking: (1) ability to apply the process of science, (2) ability to use quantitative reasoning, (3) ability to use modeling and simulation, (4) ability to tap into the interdisciplinary nature of science, (5) ability to communicate and collaborate with other disciplines, and (6) ability to understand the relationship between science and society.

The NEXUS/Physics ([Redish et al., 2014](#)) is an exemplar project that takes an interdisciplinary approach to reform disciplinary science instruction at the undergraduate level. Addressing the challenge that many topics in physics courses have very limited relevance to students in life sciences, the project developed a physics curriculum that connects and supports biology and chemistry courses taken by life sciences students. The curriculum emphasizes interdisciplinary examples that students may see in other

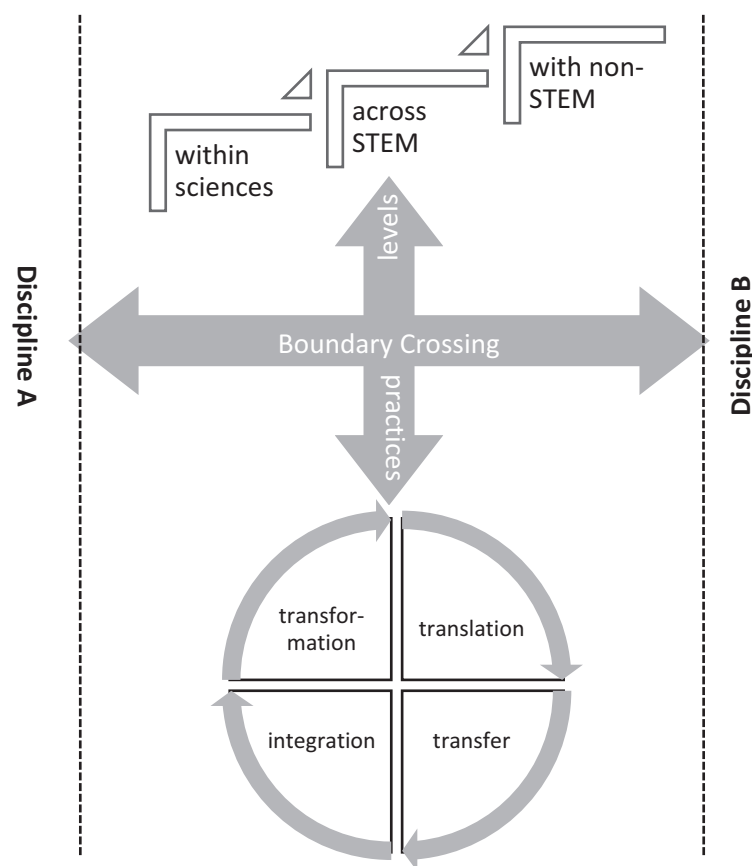


Fig. 1 Boundary crossing and interdisciplinary reasoning and communication practices in interdisciplinary science learning between two disciplines A and B.

(life) science courses. The NEXUS/Physics project highlights that developing interdisciplinary competency needs to enter as a core learning objective in an interdisciplinary curriculum. Unlike other well-defined learning objectives in educational reform, often-times we do not define or operationalize what interdisciplinary competency means.

Related to operationalizing interdisciplinary learning, one challenge that prevents effective interdisciplinary learning innovations is high quality assessment specifically designed to measure interdisciplinary learning (Gao et al., 2020). Few approaches and little empirical work existed on assessing students' interdisciplinary understanding, especially at the college level. Using an essay-based approach, Tripp and Shortlidge (2020) developed the Interdisciplinary Science Rubric (IDSR) to measure undergraduate students' interdisciplinary science thinking as demonstrated in their writing. The rubric incorporated five key aspects: objective, disciplinary grounding, integration, disciplinary humility, and collaboration. The writing assignment was administered in multiple science courses and the results indicated that the rubric produced valid data to assess students' interdisciplinary thinking in sciences. Using the topic of osmosis as a testbed, Shen et al. (2014) developed an instrument to assess college students' interdisciplinary conceptual understanding in sciences. The instrument was administered to more than 1200 students using an online platform. A Rasch partial credit analysis showed that the instrument demonstrated satisfactory psychometric properties. Student performance on the assessment suggested that the growth of college students' interdisciplinary understanding was very limited compared to the growth of their understanding of the foundational, disciplinary knowledge of the topic.

Shen et al. (2014) also revealed that the issue of translation (i.e., mapping terms across disciplines) was a key barrier for many students to solve interdisciplinary problems. A follow-up study (Sung et al., 2015) showed that osmotic pressure has been defined differently in twenty well-circulated college science textbooks. This phenomenon related to interdisciplinary translation has also been documented on other concepts common across science disciplines (e.g., for the concept of energy, see Boo, 1998; Kohn et al., 2018; Redish et al., 2014; Storey, 1992). Without accurate translation, crosscutting concepts such as energy may be interpreted differently in different science disciplines, preventing effective learning (Sung, 2012).

Likewise, educators who are engaged in designing interdisciplinary curriculum materials need to have or develop proficiency in interdisciplinary translation. The projects reported in Redish et al. (2014) and Shen et al. (2015) both illustrated the importance of having an interdisciplinary team who is willing to engage in an open dialog about the core objectives and different perspectives in an interdisciplinary educational reform. Acknowledging the different disciplinary viewpoints is a good starting point, as Redish et al. (2014) rightly pointed out:

Biology faculty and physics faculty tend to have dramatically different views about the nature and structure of the knowledge that is appropriate to teach in introductory classes. Biology students tend to have difficulty and may even actively resist seeing the value of physics and mathematics in their learning of biology.

Shen et al. (2015) described a similar situation in which a group of faculty members and graduate students from different science disciplines (physics, chemistry, biology, and veterinary medicine) struggled for a long time just to understand each other. The team was working together to create interdisciplinary assessment item that can be used in their science courses but were distracted by and stuck with terminology translation. They found that inscribing group ideas with tools such as concept map would help team members better communicate with each other.

While generally applicable in almost all different boundary-crossing scenarios, the issue of translation is often overlooked within the natural sciences—people assume that all the scientific terms have been operationalized and well defined across disciplines. In a well-designed interdisciplinary science learning environment, students need to acquire sufficient vocabulary to converse on a topic that involves multiple domains, the ability to distinguish and translate terminologies between fields, and sensitivity regarding distinctive interpretations with certain terms (Nikitina, 2005). With disciplinary knowledge becoming so narrowly specialized nowadays, it is also critical for instructors to develop the ability to effectively present and communicate knowledge among student audiences who do not speak the same scientific language.

Connecting science across STEM fields

There have been many different ways of connecting math with science (Czerniak and Johnson, 2014; Thibaut et al., 2018) or connecting fields within STEM (Kelley and Knowles, 2016). One aspect to consider is the unit of integration, e.g., in-course, within-curriculum, and co-curricular (NASEM, 2018). Here we mainly consider two modes (hierarchical and flat) that in a way are aligned with two of the IRC processes (transfer and integration). In the hierarchical (transfer) mode, one discipline serves as the knowledge base for learning the other. In other words, one discipline serves as the means and the other the ends. In comparison, a flat (or integration) mode places different disciplines at the same level with more or less equal treatment.

Hierarchical mode

The hierarchical mode is the primary mode in most college science or engineering courses when math is concerned, considering that students are expected to apply math knowledge to understand science concepts or solve engineering problems. It is often said that mathematics is the language for science (Gamow, 1988; Redish and Kuo, 2015; Wigner, 1960). Research shows, however, a common

topic may be treated very differently in a math and a science class (e.g., measurement, [Shen and Jackson, 2013](#)). Usually, math courses are taught independent of science courses at the college level, making integration much more challenging ([Gross, 2000](#); [Feser et al., 2013](#)).

Many advanced science classes require certain math courses as prerequisite or deal with certain math concepts just-in-time, as evident in NEXUS/Physics ([Redish et al., 2014](#)). [Feser et al. \(2013\)](#) described two specific approaches to integrating quantitative reasoning for biology learning: the MathBench Biology Modules at the University of Maryland rendered through an interactive web-based platform that integrates essential quantitative concepts in biology learning, and a textbook developed at Davidson College that emphasizes learning biology through the analysis and interpretation of quantitative data. In another example, [Walshe et al. \(2017\)](#) described a curriculum in which mathematics concepts were included mainly for serving inquiry science learning.

Whereas sciences may be used as an application field for math, engineering is often used as an application field for both math and science. For example, [Xie et al. \(2018\)](#) piloted the Solarize Your Home Project module with high school students in the north-east of US. In the module, students built computer models of their own houses and installed solar panels on the roof for household power supply with the aid of Energy 3D, a software that can inform engineering design with scientific simulations (e.g., visualizing the sun path, showing heat transfer through walls and windows, displaying daily or annual energy consumption curve of the house). Related science (e.g., heat transfer, energy conservation) and math knowledge (e.g., angles, balancing equations) were desired for students to create a model of an energy-efficient house.

Although in most cases, interdisciplinary transfer occurs in the direction from a more abstract or general domain to a more concrete or specific domain (e.g., from math to science), the direction of transfer could also go backwards. For instance, engineering design has been used to help students learn math or science concepts ([Becker and Park, 2011](#); [Guzey et al., 2014](#); [Xie et al., 2018](#); [Zheng et al., 2020](#)). Science and engineering have also been integrated for learning computational thinking. Including both formal and informal learning settings, [Yin et al. \(2020\)](#) reported a series of maker activities that were developed to help secondary-school students improve computational thinking dispositions and skills while working on physics and engineering activities (e.g., designing and making e-textiles by sewing electrical circuits with tiny LED lights onto clothes).

Transfer has been regarded as an important goal for learning in general. Even within individual disciplines students are expected to constantly transfer what they have learned from one situation to another, such as, from classrooms to out-of-school settings (e.g., [Bransford and Schwartz, 1999](#); [Chin and Brown, 2000](#); [Engle, 2006](#)). Successful transfer fosters deep learning by enabling students to enact the knowledge acquired previously. Interdisciplinary transfer, therefore, refers to the process whereby students apply models, concepts, and methods learned from one disciplinary context in order to understand phenomena or solve problems in another disciplinary context ([Kaminske et al., 2020](#)). In successful interdisciplinary transfer, learners should be able to recognize the core structure of the system or knowledge under study; they should be able to map out the parallel elements or corresponding parts along with their connections within the two disciplinary contexts.

Flat mode

Accompanied with the STEM movement ([Erduran, 2020](#)), many integrated STEM curricula are developed with the flat mode in mind: different disciplinary components are blended into a coherent whole ([Kelley and Knowles, 2016](#); [NAE and NRC, 2014](#)). [Gentile et al. \(2012\)](#) described the development of an intensive STEM course, Integrated Quantitative Science (IQS), offered to first-year college students that constituted half of their study schedule. The course was developed to help undergraduate students learn important concepts in and make connections among biology, chemistry, computer science, mathematics, and physics. While the study found promising results in terms of students' conceptual gains and perceptions toward interdisciplinary learning, the study also described challenges on how to balance and sequence the offering of different subject matters for first-year students.

At the precollege level, [Dierdorp et al. \(2014\)](#) described an interdisciplinary unit developed to help senior high school students build meaningful connections among natural science, mathematics, and statistics through professional practices. Their work highlighted an attempt to identify the different kinds of connections made by students, including math-science connection, that students exhibit when learning interdisciplinary topics. Their results indicated that integrating professional practices with science, mathematics, and statistics could help students build more connections among these disciplines. [Ntemngwa and Oliver \(2018\)](#) reported an integrated STEM curriculum for middle school classrooms that gave about equal attention to science, technology, and engineering. Adopting a project-based instructional approach, the students learned engineering by designing and building robots with the Lego Mindstorms robotics kits. Having robots ready, students then programmed them to fulfill assigned tasks. During the process, they needed to understand the science knowledge that was embedded as context in each activity. While teachers believed that their students improved in all three areas after the activities, they felt it was challenging for themselves to learn to construct and program robots, to select appropriate robotics activities in line with science contents, to take the role of facilitator during the activities, and to assess students' learning in the integrated STEM class. They also reported insufficient amount of time for this innovative curriculum.

Even documents of educational standards support interdisciplinary integration. For instance, the US Next Generation of Science Standards ([NGSS Lead States, 2013](#)) framed engineering and science practices as equally important in K-12 education. The standards listed eight core science and engineering practices:

- Asking questions (for science) and defining problems (for engineering)
- Developing and using models
- Planning and carrying out investigations

- Analyzing and interpreting data
- Using mathematics and computational thinking
- Constructing explanations (for science) and designing solutions (for engineering)
- Engaging in argument from evidence
- Obtaining, evaluating, and communicating information

As one can see from the list, six of the practices are presented as common across science and engineering, and two of the practices have differentiated treatment (the first one and the sixth one). These common and differentiated integration treatments provide concrete directions for connecting science with engineering in K-12 classrooms.

Similar to transfer, integration is often promoted as an important practice for both disciplinary and interdisciplinary learning. Acknowledging the fact that students bring with them a wide range of ideas, knowledge integration has been promoted as a main objective for learning within sciences (Linn, 2006; Shen and Linn, 2011). The framework emphasizes enhancing students' abilities in establishing various types of connections among scientific ideas. This framework can be expanded to interdisciplinary learning with special attention paid to integration across disciplinary boundaries (Shen et al., 2014). In interdisciplinary integration, students need to explain phenomena by incorporating multiple discipline-based concepts and tools. More importantly, they need to be more conscious about the discrepancies and mismatches in knowledge across disciplinary boundaries.

Connecting science with non-STEM fields

Interdisciplinary science learning has also been attempted in connecting science with non-STEM fields. In response to the question whether or not the disciplinary structure of higher education adequately prepares students for the 21st century, the [US National Academies of Sciences, Engineering, and Medicine \(2018\)](#) formed a committee to examine existing evidence on the effectiveness of "educational programs that mutually integrate learning experiences in the humanities and arts with science, technology, engineering, mathematics, and medicine (STEMM)" (p. 2). The committee collected a list of 218 programs and courses that integrate the humanities, arts, and STEMM fields (for the list, search "Compendium of Programs and Courses That Integrate the Humanities, Arts, and STEMM"). The committee examined a wide range of evidence and concluded that some integrative approaches are associated with positive learning outcomes, including increased critical thinking and higher-order thinking, problem solving, teamwork and communication skills, ethical decision making, and enjoyment of learning. Many of these skills are aligned with what modern employers are looking for.

Integrating science with non-STEM fields can stimulate students' motivation (Gao et al., 2020; Richardson et al., 2020). For instance, weaving arts in science learning can promote learners' inspirations and interests (Copeland et al., 2018). Burg et al. (2016) developed a curriculum in which math, science, and computer science were integrated with music and digital sound production. For another example, Bopegedera (2005) described a program in which students participated in both art workshops and science labs in order to apply scientific understanding to create artistic products. In these programs, students get to experience new learning activities that cut across disciplinary boundaries, which may engage and motivate students to learn, especially those who are fed up with rote learning.

Integrating science with non-STEM fields can promote new ways of expressing and disseminating scientific knowledge. An interesting event to consider is the Dance Your Ph.D. contest, organized by the esteemed Science Magazine. Contestants have to submit a dance performance to present their dissertation ideas. For example, the 2019 Dance Your Ph.D. winner in the physics category was about singleton electrons wandering through superconducting material (Bohannon, 2019). The 2020 winner of the overall prize was taken by Groneberg, who explored "how the motions of groups of zebrafish larvae affected each individual animal's brain development and behavior, teasing out the impact by raising some larvae in isolation and documenting any differences" (Travis, 2020). The goal of the contest was not necessarily to create new knowledge but to transform and communicate knowledge generated in Ph.D. dissertations in a new, artistic form.

Many modern societal problems, especially social scientific issues such as a global pandemic or climate change, require expertise from the STEMM fields as well as arts and humanities (NASEM, 2018). In education, this is often framed as a problem-based learning in which learning is driven by complex problems (e.g., Hmelo-Silver, 2004). For example, Braßler (2016) described an interdisciplinary problem-based course at Hamburg University to teach sustainability in which learning were driven by highly complex problems such as corruption, health inequality, and social inequity. Research has shown that disciplinary perspectives and experiences, positive or negative, play an important role in interdisciplinary problem solving in terms of conceptualizations, strategies, and attitudes (Slominski et al., 2020; Zhang and Shen, 2015). Another good example is the MIT's Terrascope program (<https://terrascope.mit.edu/overview/>). The program engages first-year college students in working on solutions addressing global problems such as clean water shortage and natural disaster with interdisciplinary research and a supportive community.

Another form of learning approach often overlapping with interdisciplinary learning is project-based learning (e.g., Solomon, 2003). In project-based learning, student learning is driven by relatively long-term projects. The STEM and Digital Literacies with Adolescents (STEM+L) project (Smith and Shen, 2017) is a good example. The STEM+L project investigated science fiction as an effective mechanism to immerse adolescents from diverse cultural backgrounds in STEM content, digital literacies, and socio-scientific issues (Ran et al., 2020). The participating students worked in small groups to design and produce multimodal science fiction in both formal and informal learning contexts. Participants chose discipline-specific roles (e.g., scientist, designer, writer) and used an array of digital tools to accomplish their collaborative projects. Culminating

activities included student presentations at a local science fiction film festival. The research revealed the distinctive patterns and trajectories about how adolescents with different backgrounds collaborated, learned, and developed distinctive disciplinary or interdisciplinary identities during the process (Jiang et al., 2019a,b). The theme of science fiction has also been explored for the interdisciplinary learning at the college level. For example, Copeland et al. (2018) reported a course, *The Science in Science Fiction*, designed by faculty members from Biology, Physics, and English departments. In the course, students read and discussed a range of science fiction texts and also conducted cross-disciplinary, collaborative research on topics of their choice. Overall, the students reported very positively about the course experience, but also acknowledged that the collaborative research across disciplines was especially challenging.

The aforementioned examples illustrate that transformation is often presented as the goal of a typical project that blends in science with non-STEM fields—generating new knowledge and/or making new products (Carlile, 2004; Carlile and Rebentisch, 2003). Intuitively, transformation is naturally linked with creation and production, suggesting an important role for arts in this endeavor. Authenticity is also a critically important consideration when including transformation in an interdisciplinary learning environment. Transformation not only involves the change in a system or entity in a different discipline in a theoretically sound way, but also results in products and solutions that are of practical importance to a local or larger audience. Having an authentic audience, which has been included as an important design feature for many of the aforementioned examples, may provide students with a sense of making real impact.

Promoting interdisciplinary science learning

Strategies for facilitating interdisciplinary science learning

The above sections presented examples of educational programs or curricula designed to facilitate interdisciplinary science learning in terms of different levels of boundary crossing. While these programs may have different goals and objectives in mind, they all promote the type of learning that requires students to connect and apply knowledge across disciplinary boundaries.

The literature suggests a wide range of pedagogies in support of an interdisciplinary science learning environment. These include generic constructivist pedagogies that encourage integrated learning (e.g., Knowledge Integration; Linn, 2006) or community building (e.g., Knowledge Building; Scardamalia and Bereiter, 2006), and pedagogies such as project-based or problem-based learning approaches that have inherent affordances to support interdisciplinary learning (Brassler and Dettmers, 2017; Lou et al., 2011, 2017).

New pedagogies and technologies should be explored and developed to better facilitate interdisciplinary science learning in the ever-changing world. For example, discipline-specific role-taking has been promoted as a strategy to facilitate interdisciplinary learning (Smith and Shen, 2017). In this strategy, students may choose and take on different roles (e.g., scientist, writer, artist, designer) when working on a collaborative project. These roles could be provided by instructor, or proposed by students themselves. They could also create hybrid roles (e.g., scientist/designer), which in a way help developing individuals' interdisciplinary identity. While working on their collaborative, interdisciplinary projects, students may see how their own interests and expertise can be connected with different areas and future careers. Research shows that this strategy may help students form core disciplinary identities (i.e., developing individual expertise and interests related to the roles they take), while gaining knowledge and expanding perspectives in other disciplines (Jiang et al., 2020).

From the technological perspective, tools should also be developed to mediate the interdisciplinary collaboration processes. For example, the iKOS platform, a web-based knowledge organization platform, has been developed to facilitate multimodal composition that is suited for interdisciplinary collaborative learning (Namdar and Shen, 2018; Shen et al., 2020). It supports different modes of knowledge representation. In addition to the distinctive knowledge entry modes, iKOS affords the creation and sharing of interactive flipbooks. This feature allows students to integrate content knowledge with multimodal writing and publish their products to the public, reaching a wider audience.

Practices in interdisciplinary science learning

Despite the various pedagogies and technologies, research on what students are actually doing when learning in an interdisciplinary context is limited. We presented the framework that includes four interrelated practices, namely, translation, integration, transfer, and transformation.

Translation is a critical, but often overlooked aspect in interdisciplinary science learning. Research shows that not only students need to develop a vocabulary base that connects different disciplinary fields, instructors or curriculum designers from different disciplinary backgrounds also need to engage in a collaborative dialog on how to develop and sustain an interdisciplinary learning program (Gentile et al., 2012; Redish et al., 2014). Developing a “disciplinary humility mindset” (Tripp and Shortlidge, 2020) with which one acknowledges and respects disciplinary differences, including their languages, is an important part of interdisciplinary learning for both students and instructors (Redish and Cooke, 2013; Slominski et al., 2020).

Integration and transfer may be construed as the two complementary practices that match the flat and the hierarchical modes in interdisciplinary science or integrated STEM learning. The flat mode refers to an interdisciplinary learning environment where different disciplinary topics and practices are juxtaposed at the same level. In this mode, integrating different disciplinary knowledge and methods to form a coherent whole is stressed. The hierarchical mode refers to an environment where one discipline serves as the base and the other as the application field. In this mode, transferring knowledge and models from one context to another needs to

happen across disciplinary boundaries. While they may occur in different settings, both the integration and transfer practices are essential in developing an interdisciplinary habit of mind.

Transformation refers to the practice when intentional changes are made in interdisciplinary learning settings. Programs that connect science learning with arts or humanities often have this practice built in their learning tasks (e.g., solving practical problems, designing new products, making artifacts for exhibit). Transformation often requires the execution of other interdisciplinary practices, as one has to be able to integrate, transfer, and translate an old system in order to transform it to a new one in a meaningful way.

These interdisciplinary practices should not be taken as purely cognitive—they also involve social processes. It matters not only how individual learners exercise these practices, but also how they interact and communicate with others in interdisciplinary learning settings. We believe the framework may provide guidance on how to better facilitate students' interdisciplinary learning.

Challenges for creating and sustaining interdisciplinary science learning

While many programs have shown promising results, the literature also indicates that interdisciplinary science learning faces many challenges. One challenge is theoretical: there is a lack of shared understanding on what interdisciplinary science learning means (Boix Mansilla and Duraising, 2007; NAE and NRC, 2014). This has practical implications in how to present learning objectives to students (Sana et al., 2020).

A closely related challenge is developing valid and reliable assessments that target interdisciplinary (science) learning (Shen et al., 2014; Tripp and Shortlidge, 2020). Reviewing how assessment was done in empirical studies in interdisciplinary STEM education, Gao et al. (2020) proposed a two-dimensional framework that includes nature of discipline (does the assessment target monodisciplinary, interdisciplinary, or transdisciplinary?) and objectives of learning (does the assessment target knowledge, skills, practices, or affective measures?). Gao et al.'s (2020) review showed that most assessments focused on monodisciplinary knowledge, monodisciplinary affective domains, and transdisciplinary affective domains.

Another constraint lies in the disciplinary structure of many institutions or organizations where learning occurs. There is a dilemma in terms of when to nurture interdisciplinary learning. On the one hand, if interdisciplinary learning needs to be built upon disciplinary foundations, we would assume interdisciplinary science learning should occur at a higher-grade level (say, secondary or post-secondary) when students have developed a substantial disciplinary knowledge base. On the other hand, the higher level the schooling goes, the more specified a course of study would be—implying more challenges and barriers for interdisciplinary science learning. For example, departmentalized structure in higher education is often difficult to penetrate.

Related to infrastructure, yet another challenge is related to the teachers or instructors who can deliver effective interdisciplinary science instruction. To implement an interdisciplinary learning approach, teachers need to have some basic level of knowledge of the disciplines involved. This has been a big hurdle for individual teachers to implement interdisciplinary curricula (e.g., Ntemngwa and Oliver, 2018). For a team approach, interdisciplinary learning implies a significant investment for team building as well as for each team member to get familiar with topics outside of his/her expertise (Gentile et al., 2012).

Conclusion and recommendation

We have described different examples that include interdisciplinary science learning in different boundary-crossing scenarios: within sciences, across STEM, and with non-STEM fields. These examples touch upon different aspects of interdisciplinary practices, including integration, translation, transfer, and transformation. In reviewing the literature, we submit that more empirical studies are still needed to design and test the effectiveness of the best strategies and tools to engender and assess these interdisciplinary learning practices in specific domains. In response to the challenges, we encourage more and deeper theoretical discussion on what interdisciplinary science learning means. We call for more empirical research on operationalizing and assessing interdisciplinary learning that goes beyond self-report. In order to successfully design and sustain an interdisciplinary science learning program or environment, stakeholders need to have an open and enduring discussion on setting the goals and objectives explicitly related to interdisciplinary learning. Acknowledging that it is often a long-term and challenging process, we recommend innovative ways for stakeholders to create platforms, harness materials, or seek resources in support of interdisciplinary collaboration.

References

- American Association for the Advancement of Science (AAAS), 2011. Vision and Change in Undergraduate Biology Education: A View for the 21st Century. Retrieved from: www.visionandchange.org.
- Becker, K., Park, K., 2011. Effects of integrative approaches among Science, Technology, Engineering, and Mathematics (STEM) subjects on students' learning: a preliminary meta-analysis. *J. STEM Educ.* 12 (5 & 6), 23–37.
- Bohannon, J., 2019. The Winner of This Year's "Dance Your Ph.D." Contest Turned Physics Into Art. Retrieved from: <https://www.sciencemag.org/news/2019/02/winner-year-s-dance-your-phd-contest-turned-physics-art>.
- Boix Mansilla, V., Duraising, E.D., 2007. Targeted assessment of students' interdisciplinary work: an empirically grounded framework proposed. *J. High Educ.* 78 (2), 215–237.
- Boo, H.K., 1998. Students' understandings of chemical bonds and the energetics of chemical reactions. *J. Res. Sci. Teach.* 35 (5), 569–581.
- Bopeggedera, A.M.R.P., 2005. The art and science of light: an interdisciplinary teaching and learning experience. *J. Chem. Educ.* 82 (1), 55. <https://doi.org/10.1021/ed082p55>.

- Bransford, J.D., Schwartz, D.L., 1999. Rethinking transfer: a simple proposal with multiple implications. In: Iran-Nejad, A., Pearson, P.D. (Eds.), *Review of Research in Education*. American Educational Research Association, Washington, DC, pp. 61–100.
- Brassler, M., Dettmers, J., 2017. How to enhance interdisciplinary competence—interdisciplinary problem-based learning versus interdisciplinary project-based learning. *Interdiscip. J. Probl. Based Learn.* 11 (2) <https://doi.org/10.7771/1541-5015.1686>.
- Braßler, M., 2016. Interdisciplinary problem-based learning—a student-centered pedagogy to teach social sustainable development in higher education. In: Filho, W.L., Pace, P. (Eds.), *Teaching Education for Sustainable Development at University Level*, World Sustainability Series. Springer, Cham. https://doi.org/10.1007/978-3-319-32928-4_17.
- Burg, J., Romney, J., Schwartz, E., 2016. *Digital Sound & Music: Concepts, Applications, and Science*. Beedle & Associates Inc, Franklin.
- Carlile, P.R., Rebentisch, E., 2003. Into the black box: the knowledge transformation cycle. *Manag. Sci.* 49, 1180–1195.
- Carlile, P.R., 2004. Transferring, translating, and transforming: an integrative framework for managing knowledge across boundaries. *Organ. Sci.* 15 (5), 555–568.
- Chin, C., Brown, D.E., 2000. Learning in science: a comparison of deep and surface approaches. *J. Res. Sci. Teach.* 37 (2), 109–138.
- Copeland, S., Furlong, M., Boroson, B., 2018. A STE[A]M approach to teaching and learning. *Int. J. Teach. Learn. High. Educ.* 30 (3), 534–548.
- Cross-Disciplinary Learning, 2012. In: Seel, N.M. (Ed.), *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1428-6_1476.
- Czerniak, C.M., Johnson, C.C., 2014. Interdisciplinary science teaching. In: Abell, S.K., Lederman, N.G. (Eds.), *Handbook of Research on Science Education*, second ed. Routledge, London, pp. 395–411.
- Dierdorff, A., Bakker, A., Van Maanen, J.A., Eijkelhof, H., 2014. Meaningful statistics in professional practices as a bridge between mathematics and science: an evaluation of a design research project. *Int. J. STEM Educ.* 1 (9), 1–15.
- Drake, S.M., Burns, R.C., 2004. *Meeting Standards Through Integrated Curriculum*. Association for Supervision and Curriculum Development (ASCD), Alexandria, VA.
- Engle, R.A., 2006. Framing interactions to foster generative learning: a situative explanation of transfer in a community of learners classroom. *J. Learn. Sci.* 15 (4), 451–498.
- Erduran, S., 2020. Nature of “STEM”? Epistemic understandings of integrated science, technology, engineering, and mathematics in education. *Sci. Educ.* 29, 781–784.
- Feser, J., Vasaly, H., Herrera, J., 2013. On the edge of mathematics and biology integration: improving quantitative skills in undergraduate biology education. *CBE Life Sci. Educ.* 12 (2), 124–128.
- Gamow, G., 1988. *One, Two, Three... Infinity: Facts and Speculations of Science*. Dover Publishing, New York.
- Gao, X., Li, P., Shen, J., Sun, H., 2020. Reviewing assessments for students’ learning in interdisciplinary STEM education. *Int. J. STEM Educ.* 7 (24) <https://doi.org/10.1186/s40594-020-00225-4>.
- Gentile, L., Caudill, L., Fetea, M., Hill, A., Hoke, K., Lawson, B., Lipan, O., Kerckhove, M., Parish, C., Stenger, K., Szajda, D., 2012. Challenging disciplinary boundaries in the first year: a new introductory integrated science course for STEM majors. *J. Coll. Sci. Teach.* 41 (5), 44–50.
- Gross, L.J., 2000. Education for a biocomplex future. *Science* 28, 807.
- Guzey, S.S., Tank, K., Wang, H., Roehrig, G., Moore, T., 2014. A high-quality professional development for teachers of grades 3–6 for implementing engineering into classrooms. *Sch. Sci. Math.* 114 (3), 139–149.
- Hmelo-Silver, C.E., 2004. Problem-based learning: what and how do students learn? *Educ. Psychol. Rev.* 16 (3), 235–266. <https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>.
- Ivanitskaya, L., Clark, D., Montgomery, G., Primeau, R., 2002. Interdisciplinary learning: process and outcomes. *Innovat. High. Educ.* 27, 95–111.
- Jiang, S., Smith, B.E., Shen, J., 2019a. Examining how different modes mediate adolescents’ interactions during their collaborative multimodal composing processes. *Interact. Learn. Environ.* <https://doi.org/10.1080/10494820.2019.1612450> online first.
- Jiang, S., Shen, J., Smith, B.E., 2019b. Designing discipline-specific roles for interdisciplinary learning: two comparative cases in an afterschool STEM+L program. *Int. J. Sci. Educ.* 41 (6), 803–826.
- Jiang, S., Shen, J., Smith, B.E., Kibler, K.W., 2020. Science identity development: how multimodal composition mediates student role-taking as scientist in a media-rich learning environment. *Educ. Technol. Res. Dev.* <https://doi.org/10.1007/s11423-020-09816-y> online first.
- Kaminske, A.N., Kuepper-Tetzl, C.E., Nebel, C.L., Sumeracki, M.A., Ryan, S.P., 2020. Transfer: a review for biology and the life sciences. *CBE Life Sci. Educ.* 19 (3), es9.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (11) <https://doi.org/10.1186/s40594-016-0046-z>.
- Klein, J.T., 1996. *Crossing Boundaries: Knowledge, Disciplinarity, and Interdisciplinarity*. University of Virginia Press, Charlottesville.
- Klein, J.T., 2008. Evaluation of interdisciplinary and transdisciplinary research: a literature review. *Am. J. Prev. Med.* 35 (2S), S116–S123. <https://doi.org/10.1016/j.amepre.2008.05.010>.
- Kohn, K.P., Underwood, S.M., Cooper, M.M., 2018. Energy connections and misconceptions across chemistry and biology. *CBE Life Sci. Educ.* 17 (1), ar3.
- Lederman, N.G., Niess, M.L., 1997. Integrated, interdisciplinary, or thematic instruction? Is this a question or is it questionable semantics? *Sch. Sci. Math.* 97 (2), 57–58.
- Linn, M.C., 2006. The knowledge integration perspective on learning and instruction. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York, pp. 243–264.
- Lou, S.J., Shih, R.C., Diez, C.R., Tseng, K.H., 2011. The impact of problem-based learning strategies on stem knowledge integration and attitudes: an exploratory study among female Taiwanese senior high school students. *Int. J. Technol. Des. Educ.* 21 (2), 195–215.
- Lou, S., Chou, Y., Shih, R., Chung, C., 2017. A study of creativity in CaC2 steamship-derived stem project-based learning. *Eurasia J. Math. Sci. Technol. Educ.* 13 (6), 2387–2404.
- McComas, W.F., Wang, H.A., 1998. Blended science: the rewards and challenges of integrating the science disciplines for instruction. *Sch. Sci. Math.* 98 (6), 340–348.
- National Academy of Engineering (NAE), National Research Council (NRC), 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*, Committee on Integrated STEM Education. The National Academies Press, Washington, DC.
- Namdar, B., Shen, J., 2018. Knowledge organization through multiple representations in a computer-supported collaborative learning environment. *Interact. Learn. Environ.* 26 (5), 638–653.
- National Academies of Sciences, Engineering, and Medicine (NASEM), 2018. *The Integration of the Humanities and Arts with Sciences, Engineering, and Medicine in Higher Education: Branches from the Same Tree*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24988>.
- NGSS Leads States, 2013. *Next Generation Science Standards: For States, by States*. The National Academies Press, Washington, DC.
- Nikitina, S., 2005. Pathways of interdisciplinary cognition. *Cognit. Instr.* 23 (3), 389–425.
- National Research Council (NRC), 2005. *Facilitating Interdisciplinary Research*. The National Academies Press, Washington, DC.
- Nternigwa, C., Oliver, S., 2018. The implementation of integrated Science Technology, Engineering and Mathematics (STEM) instruction using robotics in the middle school science classroom. *Int. J. Educ. Math. Sci. Technol.* 6 (1), 12–40. <https://doi.org/10.18404/ijemst.380617>.
- Ran, H., Shen, J., Smith, B.E., Wang, C., Kolovou, M., 2020. Storytelling for science learning: developing an assessment framework to examine adolescents’ multimodal sci-fi narratives. *Proc. Int. Conf. Learn. Sci.* 2, 817–818.
- Redish, E.F., Cooke, T.J., 2013. Learning each other’s ropes: negotiating interdisciplinary authenticity. *CBE Life Sci. Educ.* 12 (2), 175–186.
- Redish, E.F., Kuo, E., 2015. Language of physics, language of math: disciplinary culture and dynamic epistemology. *Sci. Educ.* 24, 561–590. <https://doi.org/10.1007/s11191-015-9749-7>.
- Redish, E.F., Bauer, C., Carleton, K.L., Cooke, T.J., Cooper, M., Crouch, C.H., Dreyfus, B.W., Geller, B.D., Giannini, J., Gouvea, J.S., Klymkowsky, M.W., Losert, W., Moore, K., Presson, J., Sawtelle, V., Thompson, K.V., Turpen, C., Zia, R.K.P., 2014. NEXUS/Physics: an interdisciplinary repurposing of physics for biologists. *Am. J. Phys.* 82 (5), 368–377. <https://doi.org/10.1119/1.4870386>.
- Richardson, D.S., Bledsoe, R.S., Cortez, Z., 2020. Mindset, motivation, and teaching practice: psychology applied to understanding teaching and learning in STEM disciplines. *CBE Life Sci. Educ.* 19 (3), ar46.

- Sana, F., Forrin, N.D., Sharma, M., Dubljevic, T., Ho, P., Jalil, E., Kim, J.A., 2020. Optimizing the efficacy of learning objectives through pretests. *CBE Life Sci. Educ.* 19 (3), ar43. <https://doi.org/10.1187/cbe.19-11-0257>.
- Scardamalia, M., Bereiter, C., 2006. Knowledge building: theory, pedagogy, and technology. In: Sawyer, K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York, NY, pp. 97–115.
- Shen, J., Jackson, D.F., 2013. Measure the volume of a tree: a transformative modeling lesson on measurement for prospective middle-school science teachers. *J. Sci. Teach. Educ.* 24 (2), 225–247.
- Shen, J., Linn, M.C., 2011. A technology-enhanced unit of modeling static electricity: integrating scientific explanations and everyday observations. *Int. J. Sci. Educ.* 33 (12), 1597–1623. <https://doi.org/10.1080/09500693.2010.514012>.
- Shen, J., Liu, O., Sung, S., 2014. Designing interdisciplinary assessments in science for college students: an example on osmosis. *Int. J. Sci. Educ.* 36 (11), 1773–1793.
- Shen, J., Sung, S., Zhang, D.M., 2015. Toward an analytic framework of Interdisciplinary Reasoning and Communication (IRC) processes in science. *Int. J. Sci. Educ.* 37 (17), 2809–2835.
- Shen, J., Smith, B.E., Jiang, S., 2020. Integrating Multimodal Composition Technology (MCT) in interdisciplinary learning. In: de Oliveira, L.C., Menda, A.M., Vincentini, C. (Eds.), *English Language Teaching Methods, Approaches, and Lessons*. Information Age Publishing.
- Slominski, T., Fugleberg, A., Christensen, W.M., Buncher, J.B., Momsen, J.L., 2020. Using framing as a lens to understand context effects on expert reasoning. *CBE Life Sci. Educ.* 19 (3), ar48.
- Smith, B.E., Shen, J., 2017. Scaffolding digital literacies for disciplinary learning: adolescents collaboratively composing multimodal science fictions. *J. Adolesc. Adult Literacy* 61, 85–90.
- Solomon, G., 2003. Project-based learning: a primer. *Technol. Learn.* 23 (6), 20–30.
- Steen, L.A. (Ed.), 2005. *Math & Bio 2010: Linking Undergraduate Disciplines*. Mathematical Association of America, Washington DC.
- Storey, R.D., 1992. Textbook errors & misconceptions in biology: cell energetics. *Am. Biol. Teach.* 54 (3), 161–166.
- Sung, S., Shen, J., Stanger-Hall, K.F., Wiegert, C., Li, W., Brown, S., Robertson, T., 2015. Toward interdisciplinary perspectives: using osmotic pressure as an example for analyzing textbook explanations. *J. Coll. Sci. Teach.* 44 (4), 76–87.
- Sung, S., 2012. Understanding Undergraduate Interdisciplinary Science Education from Cognitive, Curricular, and Assessment Perspectives. Doctoral dissertation. University of Georgia, Athens, GA.
- Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., et al., 2018. Integrated STEM education: a systematic review of instructional practices in secondary education. *Eur. J. STEM Educ.* 3 (1) <https://doi.org/10.20897/ejsteme/85525> article 02.
- Travis, J., 2020. Watch the Winner of This Year's "Dance Your Ph.D." Contest. Retrieved from: <https://www.sciencemag.org/news/2020/02/watch-winner-year-s-dance-your-phd-contest>.
- Tripp, B., Shortlidge, E.E., 2020. From theory to practice: gathering evidence for the validity of data collected with the Interdisciplinary Science Rubric (IDSR). *CBE Life Sci. Educ.* 19 (3), ar33.
- Walshe, G., Johnston, J., McClelland, G., 2017. Integrating mathematics into science: design, development and evaluation of a curriculum model. In: Hahl, K., Juuti, K., Lampiselkä, J., Lavonen, J., Uitto, A. (Eds.), *Cognitive and Affective Aspects in Science Education Research: Selected Papers from ESERA 2015 Conference*. Springer, Dordrecht.
- Wigner, E.P., 1960. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. Richard Courant Lecture in Mathematical Sciences Delivered at New York University, May 11, 1959.
- Xie, C., Schimpf, C., Chao, J., Nourian, S., Massicotte, J., 2018. Learning and teaching engineering design through modeling and simulation on a CAD platform. *Comput. Appl. Eng. Educ.* 26 (4), 824–840. <https://doi.org/10.1002/cae.21920>.
- Yin, Y., Hadad, R., Tang, X., Lin, Q., 2020. Improving and assessing computational thinking in maker activities: the integration with physics and engineering learning. *J. Sci. Educ. Technol.* 29 (2), 189–214. <https://doi.org/10.1007/s10956-019-09794-8>.
- Zhang, D.M., Shen, J., 2015. Disciplinary foundations for solving interdisciplinary scientific problems. *Int. J. Sci. Educ.* 37 (15), 2555–2576.
- Zheng, J., Xing, W., Huang, X., Li, S., Chen, G., Xie, C., 2020. The role of self-regulated learning on science and design knowledge gains in engineering projects. *Interact. Learn. Environ.* <https://doi.org/10.1080/10494820.2020.1761837>. Retrieved from:

STEAM for all: a vision for STEM and arts integration

Victor Hugo Minces and Nagarajan Akshay, University of California in San Diego, Department of Cognitive Science, La Jolla, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

STEM and arts	10
History and meaning of STEAM education	11
Goals of STEAM	11
STEAM's reach and barriers	13
Using STEAM to improve educational experiences within the systemic constraints	13
A STEAM-inspired vision for education	14
Conclusion	16
Acknowledgments	17
References	17
Relevant Websites	18

STEM and arts

An ancient account relates the story that Pythagoras, the mythical forefather of Greek mathematics, was preoccupied with finding a way to quantify the perception of sound, in a similar way that a scale can quantify weight, when he walked past a blacksmith's shop and heard consonant sounds (octaves, fifths, and fourth musical intervals). Excitedly, he ran into the shop and started experimenting with the sounds the blacksmiths were making, striking metal rods with varied forces, striking on different types of metal rods, and with hammers of various weights. This experience led him to carry out a series of experiments making sounds with pipes of different lengths, cups filled with different amounts of water, and strings with different lengths and tensions, thus relating the size of the vibrating objects with their frequency, and establishing the mathematical and physical underpinnings of harmonic chords and musical scales (Fig. 1).

Although not historically accurate, this tale highlights the role of science and math in understanding human sensory perception in relation to music, and the role of music in inspiring and providing a language for physics and mathematics. Throughout history, music remained an object and an inspiration for scientific research, a link that was expressed in classic education in the quadrivium, the basis of philosophical education, put forth by Plato and used through the middle ages, comprised of Arithmetic, Geometry, Music, and Astronomy (Pesic, 2014).

Perhaps the best-known personality at the intersection of art and STEM is Leonardo da Vinci. Even though he is often conceived as a polymath with diverging interests, for da Vinci art and STEM were not independent endeavors, as one of the main goals of renaissance art was creating an illusion of reality (Isaacson, 2017). This involved, for example, the use of math to create the illusion of depth through perspective drawing, understanding how objects can be modeled in three dimensions by decomposing them into simple geometrical forms, understanding how light and color are reflected on those forms and filtered through the atmosphere, understanding anatomy, and experimenting with new materials such as pigments and oils to create vivid colors.

These examples highlight that for the renaissance STEAM professionals, creating the illusion of nature involved a deep understanding of nature. And similarly, representing nature is a prerequisite for understanding it, as is also shown in da Vinci's anatomical



Fig. 1 Pythagoras researching the physical and mathematical underpinnings of musical harmony. From Franciscus Gaffurius' *Theoria Musicae* (1492).

drawings, Bouguereau's studies of clouds, and countless other illustrators whose work was fundamental for the development of the biological sciences (The Art of Innovation, BBC, See relevant websites).

It is helpful to understand the historical connections between the arts and STEM in order to see the contemporary connections and appreciate the technological aspects of art products of the past and the artistic aspects of technological products of the present. As technology evolves, what we consider "technological" changes over time. If, as computer-pioneer Alan Kay said, "technology is anything invented after you were born," it is easy for contemporary culture to overlook the technological value in the works of da Vinci and others and in the techniques described above. As a counterpart, it is sometimes easy to overlook the artistic value of technological consumer products that are integral to our culture, relegating the idea of *Art* to traditional forms found in museums, concert halls, and academic activities. The development of video games, for example, continues the renaissance's obsession with recreating reality, incorporating its technological advances, and adding newer technologies such as artificial intelligence, physics engines, and haptic interfaces to create a more immersive and flexible environment. Similarly, the Pythagorean inquiry on the nature of sound and music is extended to include the development of electronic techniques for creating music, such as the modular synthesizer (if the reader is old enough to call it technology), the development of software for sound production and editing, and the use of artificial intelligence and signal processing for autotuning the human voice (A History of Electronic Music, see relevant websites). These examples indicate that by taking a perspective that goes beyond the narrow conceptions of Art and Science typically taught in schools, the connections between STEM and the arts are easier to find than to miss.

History and meaning of STEAM education

The acronym SMET, standing for Science, Mathematics, Engineering, and Technology, was first used by the United States' National Science Foundation in the nineties and was soon scrambled into STEM to avoid vulgar connotations (Breiner et al., 2012). STEM originally designated the four field areas, without an associated educational philosophy, but with an emphasis on training and recruiting a workforce to compete in the international arena. This conception of STEM as a grouping of fields is still used frequently, particularly in policy discussions. Over time, many in the education community came to associate it with integrative project-based and inquiry learning, as is typically experienced by STEM professionals (Breiner et al., 2012), who need to bridge disciplines to solve problems. The term gained popularity in the 2000s and led to the opening of STEM schools, the designation of specialized district administrators, and professional development programs.

The first written reference to STEAM was putatively (Mejias et al., 2021) introduced by Yakman (Yakman, 2008) as an idea for greater integration of STEM and the arts, although the idea of integrating arts and STEM fields was not new (Robinson, 2013). As the recognition of the need for integration and the acronym gained popularity, many of the STEM programs added the connections to the arts and became STEAM.

The US government has promoted STEAM education through funding and conferences (Harrell and Harrell, 2010), and the US' House of Representatives released a resolution for "adding art and design to federal programs that target [STEM]" with the goal of encouraging innovation and economic growth (Langevin, 2015).

Goals of STEAM

Enough of the STEM and STEAM literature is devoted to defining, redefining, criticizing, and disentangling the meaning of the term (Breiner et al., 2012; Colucci-Gray et al., 2019; Mejias et al., 2021). We do not intend to add to that discussion but to state that the definition of STEAM varies with the stakeholder, with a general conception of project-based learning integrating the STEM fields and the arts. It should be noted that, in practice, the arts are frequently not included, for example, many self-proclaimed STEAM programs are robotic challenges (Colucci-Gray et al., 2017). The fact that experiences as such are categorized as STEAM perhaps reflects the fact that STEAM programs inherit STEM values, and that many can see the aesthetic and artistic value in the engineering processes (Harrell and Harrell, 2010). Since a large part of society's interest in STEAM is its relationship to the arts, and since this is a chapter about STEAM in a volume about STEM, here we emphasize the importance of the arts.

Just as STEAM means something different for different people, the goals for adopting STEAM education are varied, with the common assumption that children find the art-making process engaging and personally relevant. Of particular importance for their role in establishing educational policy is the view of policymakers, represented among others by the US house of representatives, whose resolution referenced above stated that:

- 1) Innovative practices of art and design play an essential role in improving Science, Technology, Engineering, and Mathematics (STEM) education and advancing STEM research.
- 2) Art and design provide real solutions for our everyday lives, distinguish United States' products in a global marketplace, and create opportunity for economic growth.
- 3) Artists and designers can effectively communicate complex data and scientific information to multiple stakeholders and broad audiences.
- 4) The tools and methods of design offer new models for creative problem solving and interdisciplinary partnerships in a changing world.

- 5) Artists and designers are playing an integral role in the development of modern technology and manufacturing.
- 6) Adding art and design to Federal STEM programs has the potential for recruiting into STEM fields children from underrepresented ethnic and economic backgrounds.

Some of these US House's statements parallel scholars' assertions that the inclusion of the arts helps develop "creativity, problem-solving skills, exploration of uncertainty, ambiguity, and fuzziness" (Colucci-Gray et al., 2017). The idea that inclusion of the arts can *infuse* STEM practice with those values will be surprising for STEM professionals, who know the importance of creativity in STEM practice, and perhaps reflects the general conception of STEM as uncreative, developed over years of engaging with science through uncreative schooling practices. Furthermore, there is no clear evidence that the creativity exercised through art practices transfers into STEM practices (Perignat and Katz-Buonincontro, 2019). It is important to note, however, that participating in a STEAM activity gives children the opportunity to *exercise* their creativity through the creation of art artifacts, with which they are often very engaged. An opportunity that they would not have otherwise.

The US House's resolution, as well as some scholarly work, indicates that STEAM is a way to engage underrepresented students in STEM fields, although the mechanism is not usually articulated. Because of the cultural relevance of the arts, and the personal relevance in terms of student interest, arts integration can achieve several dimensions that are useful in promoting equity and diversity in STEM education: They approach learning as a cultural accomplishment, allowing for science understanding to grow out of the students' lived experiences. They relate youth discourses to scientific discourses, allowing children to express their understanding in their own language. They build on prior interest and identity, as they relate to the artistic practices, tastes, and identities that the children bring. They leverage the students' cultural funds of knowledge, as they relate to the art-making practices of their communities. They make the diversity visible, highlighting the artistic (and technological) accomplishments of creators from diverse cultures and backgrounds. And they value multiple forms of expression, allowing the children's knowledge to be assessed not by standardized exams but by the creative process and the artistic artifacts that they produce (National Research Council, 2012).

Although scholarly work generally does not report on the effectiveness of art integration programs to engage underrepresented children in STEM (Colucci-Gray et al., 2017; Mejias et al., 2021; Perignat and Katz-Buonincontro, 2019), some evidence supports this claim. An example is these authors' educational program connecting music with the physics of sound and waves (Minces et al., 2021), applied with low-income and underrepresented (Latinx) eighth-grade children in a US school near the Mexican border. In this experience, children used computers to visualize and edit sound, carried out physical hands-on explorations of vibrating objects, and used their knowledge to create musical instruments. Participating in the program significantly increased the children's perception of themselves as scientists and their intention to pursue science careers, and it greatly increased their engagement with their science class.

Because the arts are manifested as projects or artistic investigations, they are naturally fit for STEM education's project-based ethos. In the example above, when children use a sound editor to create a sound composition they are making an art piece. In the process, they are visualizing the sound's waveform, which is the physical representation of sound, they are applying signal processing transformations to that waveform (for example changing the sound's speed), they are connecting those transformations to human perception (for example noticing that waveforms that evolve faster are perceived as having a higher pitch), and they are observing that superimposing sounds is equivalent to superimposing waveforms. Furthermore, children are beginning to acquire skills that are immediately useful outside of school, and that are connected to the world of work, including sound engineering but also any field that requires recording and manipulating signals. In our program, these connections are highlighted by a series of mini-documentaries profiling young underrepresented people in fields related to sound and waves. Another notable example of the power of music integration for engaging underrepresented students in STEM fields is the popular Earsketch programming platform, in which children use computing to create electronic music (Freeman et al., 2014). In these types of experiences, the arts and STEM are intertwined, as the art form is inherently technological, and participating gives children agency not only as artists but as STEM creators. These types of experiences can be qualified as *Thickly Authentic STEAM* experiences, defined as having a) personally meaningful learning experiences; b) learning that relates to the world outside of the learning context; c) learning that encourages thinking within a particular discipline (for example sound production); and d) allowing for assessment that reflects the learning process (Freeman et al., 2014; Lee and Butler, 2003; Shaffer and Resnick, 1999).

In contrast with artistic STEM projects, some experiences reported as STEAM involve art projects about STEM fields (Mejias et al., 2021), an example being interpretative dancing about planetary motion (Glinkowski and Bamford, 2009). The latter kind of experiences put arts as subservient to STEM (Robinson, 2013). It is reported that the vast majority of arts integration reports follow this subservient model (Bamford, 2006; Robinson, 2013). A problem with these subservient experiences is that they do not position the children as curious and creative participants in STEM/STEAM practices, reinforcing the popular idea that science is an established body of knowledge guarded by "others", and children's creative practices associated with science can only be making art about it. The subservient model, however, is not without advantages. It is possible that the subservient model might serve the STEAM goal of effectively communicating scientific data, although there is no guarantee that children have a deep understanding of what they are doing art about (Colucci-Gray et al., 2017), and there is no evidence that these types of experiences actually improve science communication. Another advantage is that participating in a subservient art experience might be better than the alternative, for example, dancing the planetary motion might be better than learning about it by sitting down and staring at a whiteboard. Subservient art might also offer mnemonic advantages and bring art and movement to a school day that lacks them otherwise, particularly in low-income schools (Bequette and Bequette, 2012).

The US House resolution recognizes a goal in STEAM education that is generally overlooked in scholarly articles, which is adding the value of art and design in the development of consumer products. This is better exemplified by the carefully designed consumer products of Apple –see Steve Jobs’s Stanford commencement speech for an account of the influence of art in the development of Apple products (Jobs, 2005)– which is currently the most valued company in the world, and it can be argued that Facebook and Microsoft, two of the other technology giants, owe their success not to cutting edge technologies but to their innovations in design. The goal of adding value is notable in that it sees STEAM not merely as a conduit to better or more STEM, but as worthy in itself and integral to the economy.

Given the goals typically assigned to STEAM education, it is surprising that evaluations of STEAM programs generally focus on the children’s acquisition of domain knowledge (and without control groups that would allow an understanding of STEAM’s differential contribution). The focus on domain knowledge treats STEAM as a pedagogy, in the sense of “a method for teaching something”. The goals above, however, reveal it not so much as a pedagogy, but as a policy regarding what should be taught, why, and for whom, a reflection on what education is for.

STEAM’s reach and barriers

Published reports of STEAM experiences have a large bias toward elementary school or after-school programs (Colucci-Gray et al., 2017). Published reports, however, are created within the scholar community, and are not representative of more popular STEAM practices on the ground. To understand its real reach, we carried out informal interviews with leaders in our local, and very active (Wired Magazine, 2019), STEAM ecosystem (see acknowledgments), an approach similar to Colucci-Gray et al. (2017). These leaders have created curricula, fairs, and professional development opportunities reaching more than a million students. Although there are no precise statistics, they agree that STEM and STEAM experiences decline in middle school and high school, and are more prevalent in out-of-school environments. Through our interviews and personal experience developing a popular music integration program, we have identified several barriers to implementation, which become more prominent in higher grade in-school contexts. Some of these barriers coincide with the findings of a working group of policy experts led by the *Education Commission of the States* and the *Arts Education Partnership of the US* (Dell’Erba, 2019).

(1) Siloing: Early grades are guided by one non-specialized teacher, which makes it easier for them to apply integrative subjects. As grades increase, so does teacher specialization and subject matter siloing. In such a context, teachers need to collaborate if they are to apply integrative experiences while fulfilling the standards. This approach is taken in the well-known High Tech High School System (Wagner and Dintersmith, 2015), but teacher collaboration is much harder to implement in public schools with large class sizes and overburdened teachers. Out-of-school programs, as early grades, are typically not siloed into academic subjects, which explains their prevalence in STEAM programs. (2) Curricular pressure: Teachers are often pressed to teach a rigid curriculum or a set of standards that might not be possible to align with STEAM experiences (Dell’Erba, 2019). Whereas this is true for all grades, it is particularly influential in higher grades, where curricular topics become more microscopic and abstract. For example, it is straightforward to create an engaging STEAM activity around the first grade US standard *connecting sound and vibrations*, because the standard reflects a phenomenon that children encounter regularly, for example when they stand next to a bass speaker or drum, and because it can be directly linked with the creation of musical instruments. In contrast, we have not been able to create a personally meaningful activity linked to the middle school standard addressing *wavelength*, because wavelengths cannot be felt, are hard to visualize, and are not relatable to the children’s everyday experiences. (3) Teacher preparation: an important barrier across grades is teacher preparation (Dell’Erba, 2019). This does not only involve teaching teachers how to implement specific STEAM experiences, but also requires changing rigid teaching practices that are over-reliant on memorization and testing as a way to measure the children’s performance. (4) Teacher time commitments: teachers are often overburdened, serving several classes with more than 30 students. The time commitment and the stress of classroom management leave them little time and energy for preparation, collaboration, and carrying out the type of personalized work that supervising STEAM student creations might require (Dell’Erba, 2019). (5) Research focus: because published reports of STEAM experiences are created within the scholar community, they typically involve the participation of highly trained professionals and subject domain experts, and often involve small groups of highly motivated students participating in out-of-school programs. This is problematic because the resources thus created are hard to extrapolate to less motivated students, less well-prepared facilitators, and harder curricular constraints.

STEAM is an attempt to bring to the classroom an integration of meaningful personal interests, project-based learning, performance evaluations, communication, and collaboration. These values run counter to the structure of the educational system, which is siloed, individualistic, test-oriented, understaffed, and focused on subject domain standards that children might not find a personal connection with. In this context, two strategic questions arise: (1) How can the scholar community improve the children’s educational experience given the current systemic constraints?; and (2) What radical transformations should the scholar community envision and advocate to implement in the next round of systemic reform? The next two sections explore these questions.

Using STEAM to improve educational experiences within the systemic constraints

Without denying the value of out-of-school programs, focusing on school settings has the potential to engage institutional structures serving many more children, including those that do not already have a science identity that led them to join an out-of-school

program and are therefore poised to benefit the most. The educational program being developed by these authors has been growing within, and in partnership with, actors in the formal school system, and the resources thus created are widely used in the US and are starting to be used globally (Akshay et al., 2020). Here, we draw on our experience growing the program (Minces, 2021) and our conversations with the STEAM leaders to address the implementation barriers.

A key factor for growing the program is the alignment with the curricular standards, which addresses the issue of curricular pressure mentioned in the previous section. Schools and teachers are oftentimes eager to implement resources that they perceive will engage their students, and will need a minimal measure of standards alignment to be able to do so. Therefore, the standard alignment can sometimes be rather loose, partial, or include STEAM experiences in combination with less meaningful materials. For example, our curricular resources integrate music when addressing the relationship between frequency and pitch, but are much drier when addressing the concept of *wavelength*.

Even allowing for loose alignments, some standards are so far removed from the children's everyday experiences that it cannot be expected that they will be addressed through an Authentically Thick STEAM project. In that case, a subservient art experience always offers the possibility of children making an art presentation about the standard, which might be better than the alternative.

With regard to teacher implementation, it is important that the resources help alleviate, rather than increase, the teachers' burdens. STEAM is a great opportunity to engage teachers, since they can also participate in the joy of the art-making process, which likely increases the likelihood that they will participate in teacher training. Resources should be extremely easy to start implementing, in what Resnick calls low floor (Resnick and Robinson, 2017), and allow for deeper engagement (high ceiling). A low floor includes activities requiring minimal or easily available material resources. For example, using web applications can be much more accessible than using downloadable software, which can be difficult to install on the schools' computers or on the students' devices. A low floor also means designing tools that are easy to use and that are engaging from the first moment. Involving teachers in the design process can be very helpful for fulfilling those goals (Minces, 2021). Another factor facilitating the implementation of STEAM experiences is to provide rubrics that allow teachers to guide the art-making process and evaluate the children's work with less reliance on testing, and with less time spent analyzing each artifact (for example, grading the children's compositions). Easing the implementation of a program will not only facilitate the work in the classroom but will also streamline the teachers' professional development. Regarding research, achieving broader implementation requires a shift toward policy-based research (List, 2022), that is, research that from the outset takes into consideration the possibility of scaling up the results.

A STEAM-inspired vision for education

The section above describes strategies for improving the children's educational experience within the limits of the educational system. This section discusses how STEAM can inform future rounds of systemic reform, which requires a reflection on what education is for (Biesta, 2015). The educational system does not have a clear overarching goal. Rather, it has adopted the goals of different stakeholders, which in turn might be different and contradictory (Labaree, 2012). The state might be interested in forming citizens with coherent national values and in developing the workforce to serve the economic and strategic needs of the country. Parents might value the educational system for providing a place for childcare and improving their children's economic outlook. Teachers might care about leading a classroom where students are intellectually engaged while complying with the school authorities by implementing the new set of standards and preparing children for the test. The goals for principals might be to manage the children's behavior and improve the schools' statistics by increasing the number of students pursuing higher education. Children have no choice but to participate in the system and, young children in particular, might not be so concerned about the future. They might be very interested in socializing, in not being bored, and in learning something, although not necessarily what they are expected to learn. An additional stakeholder, different from the child, is the adult that the child will be, which emerges slowly as the child approaches high school graduation and faces the need to sustain herself or himself. A goal for that adult is to have an occupation that provides life quality, which could involve economic well-being, intellectual fulfillment, or a sense of purpose. Successful reform must acknowledge and harmonize all these goals that at times appear to be in tension, in particular the state's goal for developing a specific type of workforce versus the children's personal interests. Our goal as researchers is to emphasize the interest of the children to help create a more joyful learning environment, without disregarding the interests of other stakeholders.

The policymakers' and society's interests are expressed in the curriculum and activities carried out in schools. In the United States and England, which have influenced educational systems and science education throughout the world, curricular subjects have been a matter of fiery debates and widespread reforms (Atkin & Black, 2007; DeBoer, 2014). Three main conceptual currents have informed those reforms, with elements of them alternating and taking prevalence over time. One current, dominating in the late 19 century, and referred to here as experiential, emphasized science practices and direct observation of everyday and natural phenomena. Another current, prevalent in the first half of the twentieth century in association with the progressive movement, and briefly reemerging in the sixties, emphasized the curricular subjects' value to society, value to the students, and technological applications in nascent industries. The present moment inherits the values of the education reform born during the cold war's fight for technological supremacy in the fifties, sometimes referred to as the Sputnik era of reform, and that regained strength in the eighties in the context of international industrial competition. The latter current reflected the value that society assigned to cutting edge scientific knowledge, and was shaped by the advice of prominent scientists, who advocated for including curricular subjects associated with the latest advances in their fields and the corresponding higher education courses, in what Atkin and Black

(2007) call “scientists’ science”. The latter era also saw a great increase in the value that society assigned to higher education (Baker, 2014), and in the number of people pursuing higher education, with the consequent need to sort what children could attend what institutions (Sandel, 2020). This confluence of factors produced a science and math curriculum that focuses on preparing children for pursuing careers in STEM through higher education paths, and sorting the children that can go through the hoops that such preparation requires. This ethos is captured in the popular model of STEM workforce development as a pipeline (Cannady et al., 2014; Metcalf, 2010), in which children enter at one end and emerge on the other as adult STEM professionals, with the problem that some children, those that cannot go through the hoops, drip off the pipeline, which is therefore referred to as being leaky. We note that the pipeline model is partially different in some countries in which an early tracking system sorts children putting them on a path to higher education or to technical careers through vocational schools. Such a system, which could be modeled as a bifurcating pipeline, has the advantage of producing a large pool of necessary mid-skilled labor and offering young people a safe path to economic security (Newman and Winston, 2016), but it has the disadvantage of sorting children early in life, making it difficult for someone to change and cross trajectories, perpetuating social differences, and creating resentment.

Given that the majority of children, particularly low-income children, will not finish a higher education degree, and a very large majority will not pursue a higher education STEM career, a strong focus on higher education science implies that they will spend a large portion of their time studying curricular topics that are not relevant to their lives or future occupations, that they might be resentful of being sifted-out in their science and math classes, and that by early adolescence they will lose confidence in their capacity to participate in the STEM world (Tai et al., 2006). The pipeline model does not serve the interest of society and the government either, since the exaggerated focus on preparing children for higher education in STEM fields: (a) fails to engage enough children in higher education STEM pathways (b) fails to produce the necessary mid-skill STEM practitioners that the workforce also needs and that can provide young people with a decent living wage (Newman and Winston, 2016), (c) fails to give children the tools to engage with STEM in their everyday lives, and (d) generates negative attitudes toward science, which might explain the current resentment and suspicion by a large part of the population toward the scientific community (Atkin & Black, 2007). Furthermore, it is not clear that the curricular topics being emphasized at a moment in time will be the most relevant for the workforce of the future, for example, data science and statistics are arguably more valued in the workforce today than calculus, but are not as prominently represented in the curriculum. In other words, education today prepares the workforce of the future for the jobs of the past.

The problems with the pipeline model and the intense focus on higher education topics are being recognized by the education community, as expressed by Atkin and Black: “As it is realized that scientists’ science, oriented toward the painstaking construction of conceptual frameworks that explain how the world works, is of limited use to the vast majority, new ways of enabling the citizenry to comprehend and operate in a scientific and technological culture have to be sought”. With this understanding, in the later years the education community has been striving to promote project-based learning and culturally responsive education (Ladson-Billings, 2014), conceptions from which STEM and STEAM education derive, and which are associated with the values prevalent in the experiential and progressive eras. Those efforts crash against curricular subjects that are too obtuse, for example, it is hard to think of a culturally responsive and experiential way to explore the difficult concepts of wavelength or the particle-wave duality of light, which, it can be argued, is not even understood by professional physicists (Ballentine, 2014). An additional problem of these obtuse standards is that being poorly understood by teachers and curriculum developers, they are reflected in superficial activities that misinterpret the natural phenomena that they are meant to explore, and that can only be accomplished through rote memorization (Fig. 2), which removes the children’s agency as understanders and inquirers. Those traditional curricular topics are decontextualized, in the sense that neither children nor teachers know why they are important or how they can be applied to their lives, which deprives children of a motivation to learn them other than to continue their academic life. Studying subjects without a clear motivation, especially for students that are not driven by grades and do not aspire to pursue higher education, is tedious and hard, or, as the Sputnik era discourse often calls for: rigorous (from Latin *rigorem*, meaning stiff). The emphasis on rigor as a primary value is surprising, since there is no indication that rigorous work produces more meaningful knowledge, and on the contrary, the less meaningful the knowledge the more rigorous study it requires.

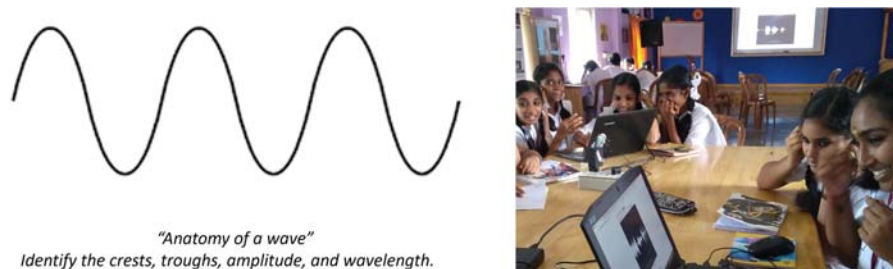


Fig. 2 STEAM versus traditional schooling practices. Left, typical test given to children to assess their knowledge of waves. Note that the axes do not have labels so it is not clear what the line is representing. Teachers can administer the test and students can comply, even without understanding the meaning of the concepts. Right, children create musical compositions using a sound editor and oscilloscope (see list of relevant websites). Through this activity children can learn to measure, visualize, and manipulate signals, which are ubiquitous practices in science and engineering, and improve their data literacy (Akshay et al., 2020).

In contrast to rigor as a primary value, STEAM education offers curiosity, passion, and the aesthetic experience (Booker and Minces, 2022; Milne, 2010), values that more faithfully reflect the experience of scientists and artists, for whom rigor is secondary to the pursuit of those passions. Through its connections with the arts, STEAM education signals a way for choosing curricular topics that are personally relevant to the child. Importantly, if the community gets past their rigid expectations of what science should look and feel like, it can be understood that Thickly Authentic STEAM topics are not any less science than other traditional curricular topics. For example, understanding how objects vibrate, how microphones work, or how to read and manipulate signals is not any less science than understanding the concept of wavelength (Fig. 2). And because Authentically Thick STEAM activities are contextualized in personal goals they are less prone to misinterpretation. In our example, a child needs to understand how to read and manipulate a signal to achieve their artistic goal of creating a sound composition. Through such activities, all children can engage with STEM as creators, develop agency and positive attitudes toward STEM, acquire knowledge that allows them to reinterpret their everyday environment—in our example, popular music—and gain skills that they can immediately apply, so they never need to ask “what are we doing this for?”. These experiences are directed at enriching the lives of all children, rather than the selected few that will follow a STEM higher education path, and these experiences are actually likely to increase the proportion of children that follow that path. Furthermore, the skills gained can be extended into other STEM fields that are highly valued in the workforce, as the Earsketch music coding platform (Freeman et al., 2014) or these authors’ data literacy work indicate (Akshay et al., 2020).

STEAM calls for rethinking the curricular topics to move them from decontextualized and impersonal to contextualized and personally relevant. Possibilities abound, as children work on woodworking projects they can learn algebra and geometry, develop their spatial intuition, explore the properties of materials and learn to design for structural stability; if they have access to laser cutters or CNC routers they can learn fabrication technologies and digital design. Importantly, they will be learning woodworking. Children can also learn electronics in the context of clothing design (Peppler, 2013), computing in the context of videogame design (Resnick and Robinson, 2017), optics in the context of photography, chemistry in the context of cooking and painting (through the creation of pigments) and, in the spirit of Leonardo, biology and geometry in the context of drawing.

The examples above point to a need to modify the academic structure to reflect project goals rather than siloed academic fields – for example introducing sound production, computer music, video game design, or woodworking classes – and to rethink the teacher workforce to include domain practitioners rather than teachers that do not have practical expertise in the fields they are teaching. Through these activities and personal interaction with practitioners, children can become aware and start exploring possible career occupations more immediately available to them after high school, such as carpentry, or that require further academic training, such as structural engineering.

This vision for education has many commonalities with vocational schooling. This is not a coincidence, since they both share the philosophy of goal-based and situated work. Mainstream schooling has a lot to learn from vocational schooling, which when implemented properly can provide mid-skill workers for the economy and has been shown to decrease school attrition and lead to good employment for young people, particularly for the many children that do not thrive in a less contextualized academic environment (Newman and Winston, 2016). As opposed to vocational education, however, this vision puts a heavier emphasis on creativity and personal meaning, and does not call for the highly specialized technical education that could limit young people’s career choices.

Thickly Authentic STEAM practices also call for shifting the way students’ capabilities are assessed, away from tests (Fig. 2) and toward performance assessments in which children demonstrate their learning through the artifacts they create (Bland and Gareis, 2018). This shift can also allow students to create a personal portfolio which can be useful in securing future employment.

The proposals above defy the role of K12 education as a sorting mechanism for higher education. This is problematic within a system that values higher education credentials in terms of social status and in terms of economics, as higher education degrees are requisite for high-paying jobs (Baker, 2014; Collins, 2019). However, these proposals might better accompany current trends in higher education and the world of work, which have been under increasing pressure – and perhaps recognize the need – to become more inclusive of social groups traditionally underrepresented. In response to that pressure, several prestigious universities have been reevaluating their admission practices to be less reliant on traditional measures of capability (Hubler, 2020; Sandel, 2020) that favor high-income applicants; and as private companies and governmental institutions have struggled to find capable employees, they have been steadily expanding hiring practices that are less reliant on the applicants’ credentials than on their skills, in what is referred to as skill-based hiring (Ark, 2021; Lohr, 2022), which can be well developed through STEAM practices and well represented through the students’ portfolios.

Conclusion

The school system appears to be in tension between an instrumentalist conception of education as a workforce factory and a humanistic conception that forefronts the children’s well-being and intellectual development. Part of this tension arises from society’s narrow views of what STEM and the arts should look like, and what are the paths necessary to get to them. The conception of STEM paths as a pipeline into higher education has deprived children of the possibility of experiencing STEM fields as fascinating and useful, alienated them from STEM, and done little to achieve its expected goal of developing the workforce. By opening the field of vision and recognizing the many ways in which children can engage creatively in STEM fields through the arts, STEAM education can harmonize that tension, engaging more diverse children in STEM paths, bringing them value and skills that they can readily use, and contributing to making their school experience more joyful.

Acknowledgments

We are grateful for the input from the following San Diego STEAM leaders: Michael Goodbody, director for STEAM Innovation at San Diego Unified School District; Leo Ulloa, STEAM coordinator at Sweetwater High School Unified School District; Victor Ciciarelli, creator of the STEAM maker space and STEAM maker fair, which draws 30,000 children annually; Karyn Warner and Glennon Stratton, associate director and director of STEAM initiatives at the CDE-Foundation, which provides STEAM professional development for teachers in California and throughout the world; and Alec Barron, director of the San Diego Science Project and former science content specialist at Escondido Union High School District. We are also grateful for the historical insights provided by Donald Peurach and the careful editing by Laleh Quinn. Minces' work is funded by the National Science Foundation's Innovative Technology Experiences for Students and Teachers (ITEST) #2048930. Increasing Students' Interest in STEM through the Science of Music.

References

- Akshay, N., Minces, V., Vazhayil, A., Gopalsamy, V., Bhavani, R.R., 2020. There's Data All Around You: Improving Data Literacy in High Schools through STEAM Based Activities. *Proceedings of Fablearn Asia 2020*.
- Ark, T.V., 2021. The Rise of Skills-Based Hiring and What It Means for Education. *Forbes*. <https://www.forbes.com/sites/tomvanderark/2021/06/29/the-rise-of-skills-based-hiring-and-what-it-means-for-education/>.
- Atkin, M., Black, p., 2007. History of science curriculum reform in the United States and the United Kingdom. In: *Handbook of Research on Science Education*. Routledge.
- Baker, D., 2014. The schooled society. In: *The Schooled Society*. Stanford University Press.
- Ballentine, L., 2014. *Quantum Mechanics: A Modern Development*. World Scientific Publishing Company.
- Bamford, A., 2006. The Wow Factor: Global Research Compendium on the Impact of the Arts in Education. Waxmann Verlag.
- Bequette, J.W., Bequette, M.B., 2012. A place for art and design education in the STEM conversation. *Art Educ.* 65 (2), 40–47.
- Biesta, G., 2015. What is education for? On good education, teacher judgement, and educational professionalism. *Eur. J. Educ.* 50 (1), 75–87.
- Bland, L.M., Gareis, C.R., 2018. Performance assessments: a review of definitions, quality characteristics, and outcomes associated with their use in K-12 schools. *Teach. Educ. J.* 11, 52–69.
- Booker, A., Minces, V., 2022. Harmonized Mutual Development through Exploring and Creating Sound. *International Conference of the Learning Sciences*, Hiroshima, Japan.
- Breiner, J.M., Harkness, S.S., Johnson, C.C., Koehler, C.M., 2012. What is STEM? A discussion about conceptions of STEM in education and partnerships: what is STEM? *Sch. Sci. Math.* 112 (1), 3–11. <https://doi.org/10.1111/j.1949-8594.2011.00109.x>.
- Cannady, M.A., Greenwald, E., Harris, K.N., 2014. Problematising the STEM pipeline metaphor: is the STEM pipeline metaphor serving our students and the STEM workforce? *Sci. Educ.* 98 (3), 443–460. <https://doi.org/10.1002/sce.21108>.
- Collins, R., 2019. *The Credential Society*. Columbia University Press.
- Colucci-Gray, L., Trowsdale, J., Cooke, C.F., Davies, R., Burnard, P., Gray, D.S., 2017. Reviewing the Potential and Challenges of Developing STEAM Education through Creative Pedagogies for 21st Learning: How Can School Curricula Be Broadened towards a More Responsive, Dynamic, and Inclusive Form of Education?
- Colucci-Gray, L., Burnard, P., Gray, D., Cooke, C., 2019. A Critical Review of STEAM (Science, Technology, Engineering, Arts, and Mathematics). *Oxford Research Encyclopedia of Education*.
- DeBoer, G.E., 2014. The history of science curriculum reform in the United States. In: *Handbook of Research on Science Education*, vol. II. Routledge, pp. 573–592.
- Dell'Erba, M., 2019. Preparing Students for Learning, Work and Life through STEAM Education. Policy Brief. Education Commission of the States.
- Freeman, J., Magerko, B., McKlin, T., Reilly, M., Permar, J., Summers, C., Fruchter, E., 2014. Engaging Underrepresented Groups in High School Introductory Computing through Computational Remixing with EarSketch, pp. 85–90.
- Glinkowski, P., Bamford, A., 2009. Insight and Exchange: An Evaluation of the Wellcome Trust's Sciart Programme. Citeseer.
- Harrell, D.F., Harrell, S.V., 2010. Strategies for Arts+ Science+ Technology Research: Executive Report on a Joint Meeting of the National Science Foundation and the National Endowment for the Arts. tech. report, Nat'l Science Foundation, 22 Sept.
- Hubler, S., May 23, 2020. Why is the SAT falling out of favor? *The New York Times*. <https://www.nytimes.com/2020/05/23/us/SAT-ACT-abolish-debate-california.html>.
- Isaacson, W., 2017. *Leonardo Da Vinci*. Simon and Schuster.
- Jobs, S., 2005. Steve Jobs' 2005 Stanford Commencement Address. <https://youtu.be/UF8uR6Z6KLc>.
- Labaree, D.F., 2012. *Someone Has to Fail: The Zero-Sum Game of Public Schooling*. Harvard University Press.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: Aka the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Langevin, J.R., 2015. Expressing the Sense of the House of Representatives that Adding Art and Design into Federal Programs that Target the Science, Technology, Engineering, and Mathematics (STEM) Fields Encourages Innovation and Economic Growth in the United States, No. 247, 114 Congress (2015). <https://www.congress.gov/bills/114/congress/house-resolution/247>.
- Lee, H.-S., Butler, N., 2003. Making authentic science accessible to students. *Int. J. Sci. Educ.* 25 (8), 923–948.
- List, J.A., 2022. The voltage effect: how to make good ideas great and great ideas scale. *Currency*.
- Lohr, S., April 8, 2022. A 4-year degree isn't quite the job requirement it used to be. *The New York Times*. <https://www.nytimes.com/2022/04/08/business/hiring-without-college-degree.html>.
- Mejias, S., Thompson, N., Sedas, R.M., Rosin, M., Soep, E., Peppler, K., Roche, J., Wong, J., Hurley, M., Bell, P., Bevan, B., 2021. The trouble with STEAM and why we use it anyway. *Sci. Educ.* 105 (2), 209–231. <https://doi.org/10.1002/sce.21605>.
- Metcalfe, H., 2010. Stuck in the pipeline: a critical review of STEM workforce literature. *InterActions* 6 (2). <https://doi.org/10.5070/D462000681>.
- Milne, I., 2010. A sense of wonder, arising from aesthetic experiences, should be the starting point for inquiry in primary science. *Sci. Educ. Int.* 21 (2), 14.
- Minces, V., 2021. Developing and popularizing STEAM online tools: the case of "listening to waves" tools for the science of music. In: *2021 ASEE Virtual Annual Conference Content Access Proceedings*, p. 36938. <https://doi.org/10.18260/1-2-36938>.
- Minces, V., Booker, A., Khalil, A., 2021. Listening to waves: engaging underrepresented students through the science of sound and music. *Connected Science Learning* 3 (4). <https://www.nsta.org/connected-science-learning/connected-science-learning-july-august-2021/listening-waves-engaging>.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- Newman, K.S., Winston, H., 2016. *Reskilling America: Learning to Labor in the Twenty-First Century*. Metropolitan Books.
- Peppler, K., 2013. STEAM-powered computing education: using e-textiles to integrate the arts and STEM. *Computer* 46.
- Perignat, E., Katz-Buonincontro, J., 2019. STEAM in practice and research: an integrative literature review. *Think. Skills Creativ.* 31, 31–43. <https://doi.org/10.1016/j.tsc.2018.10.002>.
- Pesic, P., 2014. *Music and the Making of Modern Science*. MIT Press.
- Resnick, M., Robinson, K., 2017. *Lifelong Kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play*. MIT press.

- Robinson, A.H., 2013. Arts integration and the success of disadvantaged students: a research evaluation. *Arts Educ. Pol. Rev.* 114 (4), 191–204. <https://doi.org/10.1080/10632913.2013.826050>.
- Sandel, M., 2020. *The Tyranny of Merit: What's Become of the Common Good?* Macmillan. <https://us.macmillan.com/books/9780374289980/thetyrannyofmerit>.
- Shaffer, D.W., Resnick, M., 1999. "Thick" authenticity: new media and authentic learning. *J. Interact. Learn. Res.* 10 (2), 195–216.
- Tai, R.H., Qi Liu, C., Maltese, A.V., Fan, X., 2006. Planning early for careers in science. *Science* 312 (5777), 1143–1144.
- Wagner, T., Dintersmith, T., 2015. *Most Likely to Succeed: Preparing Our Kids for the Innovation Era*. Simon and Schuster.
- Wired Magazine, 2019. How One School District Is Putting its Stamp on STEAM. <https://www.wired.com/wiredinsider/2019/08/one-school-district-putting-stamp-steam/>.
- Yakman, G., 2008. STEAM Education: An Overview of Creation a Model of Integrative Education. IEEA.

Relevant Websites

- A History of Electronic Music <https://ahoem.org>.
- Online oscilloscope: www.listeningtowaves.com/oscilloscope.
- The art of innovation, BBC podcast series on the historical connections between arts and STEM <https://www.bbc.co.uk/programmes/p07xl0v>.

Mathematics in STEM education

Meixia Ding^a and Jinfa Cai^b, ^a Temple University, Philadelphia, PA, United States; and ^b University of Delaware, Newark, DE, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	19
The relationship between mathematics and STEM	19
Mathematics in STEM	20
Theoretical frameworks	20
Mathematics as a language tool in STEM education: the content	21
Representing and modeling	21
Interpreting and explaining	22
Mathematics as a thinking tool in STEM education: the process	22
Mathematical reasoning	22
Computational thinking (CT)	23
Implications: what do we learn from the current research?	24
Setting up grade-appropriate mathematics learning goals in STEM lessons	24
Promoting explicit mathematical connection-making in STEM lessons	25
Developing a habit of mind of mathematical reasoning for STEM lessons	25
Conclusion and future directions	26
Acknowledgment	26
References	26

Introduction

With the rise of STEM education as a way to develop students' 21st century skills (Binkley et al., 2012; Johnson et al., 2020; Maass et al., 2019a,b), the role that mathematics plays in STEM teaching and learning has drawn increased attention. One of the biggest concerns is that mathematics is largely underrepresented in STEM education even though it is commonly acknowledged as a foundation for STEM activities (English, 2016; English and Kirshner, 2016; Honey et al., 2014). Often, mathematics is treated as a set of given procedures or tricks that help students complete STEM activities (Li and Schoenfeld, 2019). Such an approach to mathematics learning leads to the loss of opportunities to develop mathematical thinking (Wong and Dillon, 2019). Not surprisingly, STEM activities were found to be beneficial for other disciplines, mainly science, but rarely for mathematics (Fitzallen, 2015; Honey et al., 2014; National Research Council, 2014). Consequently, the field of mathematics education calls for making mathematics explicit in STEM activities (Larson, 2017; Maass et al., 2019b; National Council of Supervisors of Mathematics & National Council of Teachers of Mathematics [NCSM & NCTM], 2018; Shaughnessy, 2013; Silk et al., 2010). For instance, NCSM and NCTM (2018) stress the importance of teaching mathematical content and process well when tackling integrative STEM problems: "Whenever mathematics is included in a STEM activity, make sure that the mathematics addresses academic standards appropriate for the grade level and that it is taught in ways that support the development of mathematical thinking and quantitative reasoning" (p. 4). Indeed, when the relevant mathematical aspects are addressed appropriately and deeply, students' interests and dispositions toward STEM could be ignited. This is a critical component in the STEM field (President's Council of Advisors on Science and Technology [PCAST], 2010). Building on the perspectives advanced by NCSM and NCTM (2018) as well as other recent literature, we explore how critical aspects of mathematical content and process can contribute to STEM education, as well as how it can be made explicit to students. To explore these mathematical aspects within STEM education and provide context for our research focus, we start with a broad overview of the relationships between mathematics and STEM education.

The relationship between mathematics and STEM

According to NCSM and NCTM (2018), mathematics, with its unique contents and connections to non-STEM fields, does more than serve as a tool for STEM. We propose Fig. 1 to illustrate the relationship described by NCSM and NCTM (2018):

Teaching mathematics well is an important component of a comprehensive STEM program. There is more to mathematics, however, than being part of STEM. The mathematics that students learn in school includes content and thinking that can be used as tools for tackling integrative STEM problems. But it also includes content that might be considered "just math" or might be connected to non-STEM disciplines (p. 3).

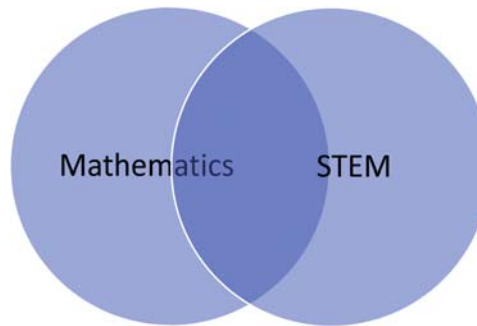


Fig. 1 Illustration of NCSM and NCTM's (2018) position about mathematics and STEM.

In Fig. 1, the three sections refer to mathematics only, STEM only, and the intersection of mathematics and STEM. On the one hand, some mathematics concepts are “pure math,” concepts that originated without a connection to STEM. For example, the axiom and propositions in Euclidean geometry were defined within its own system (Fitzpatrick, 2007). On the other hand, the STEM fields (excluding M) contain unique phenomena, concepts, and rules that may not have a relationship to mathematics. For instance, the fact that plants need sunlight, water, and air to grow is a science phenomenon that exists regardless of mathematics. Importantly, the intersection between mathematics and STEM provides an understanding of the depth and interconnectedness of the two. For instance, the STEM field provides real-world contexts in which mathematics are applied and developed, while mathematics serves as a tool to describe and explore the STEM fields.

In this article, we focus on the intersection between mathematics and STEM. Specifically, we aim to discuss mathematics in STEM education in K–12 classrooms. Ideally, mathematics learning can contribute to the success of STEM education, which in turn boosts the learning of mathematics. However, empirical studies indicate challenges with mathematics in STEM lessons. For instance, science teachers found that their students lacked important mathematics knowledge needed to successfully conduct science activities (“math blame,” Wong and Dillon, 2019). On the other hand, mathematics teachers noticed that rich opportunities to discuss mathematical concepts diminished in science courses (Wong and Dillon, 2019). In fact, researchers are concerned that mathematical concepts and reasoning are often lost in STEM activities, which does not benefit mathematics learning (Becker and Park, 2011; Honey et al., 2014). Fitzallen (2015) further argued that even though many STEM projects claimed that mathematics was involved in their projects, the ways in which “mathematics can influence and contribute to the understanding of the ideas and concepts of other STEM disciplines” (p. 241) were not addressed. In this article, we aim to narrow this gap by discussing how mathematics may contribute to STEM activities.

Mathematics in STEM

Theoretical frameworks

Fig. 2 presents two existing theoretical frameworks that together guide our study:

Fig. 2A is an illustration taken from Maass et al. (2019a), which zooms out on the role of mathematics in STEM reality. In this model, a real-world situation can be simplified into a situation model, which can be further mathematized (as illustrated by the “mathematizing” arrow at the top of the figure) to generate solutions. Conversely, mathematical solutions can be interpreted to help solve targeted STEM problems (as illustrated by the “interpreting” arrow at the bottom of the figure). The model in Fig. 2A raises the question, “What are the exact mathematical elements that play a role in STEM?” As mentioned earlier, NCSM and NCTM (2018) stress that both the content and process of mathematics should contribute to STEM learning. The Program for International Student Assessment’s (PISA) 2022 mathematics framework (Fig. 2B; Organization for Economic Co-operation and Development [OECD], 2018) offers more details about the specific elements of mathematics that play an important role in STEM education. The PISA 2022 framework identifies four targeted content areas (*quantity, uncertainty and data, change and relationship, and space and shape*) that are basically consistent with NCTM’s (2000) five content standards (*number and operations, algebra, geometry, measurement, and data analysis and probability*). The framework also highlights the cyclical process of problem-solving (or modeling): *formulate, employ, and interpret and evaluate*. The center of the problem solving cycle is mathematical reasoning.

Using these theoretical frameworks, we now discuss how mathematical content and process may contribute to STEM activities. Although mathematical content and process cannot be separated, for clarity we discuss each aspect in a separate section. Specifically, we argue that the content of mathematics can serve as a language tool for STEM activities, while the process of mathematics serves as a thinking tool for STEM experiences.

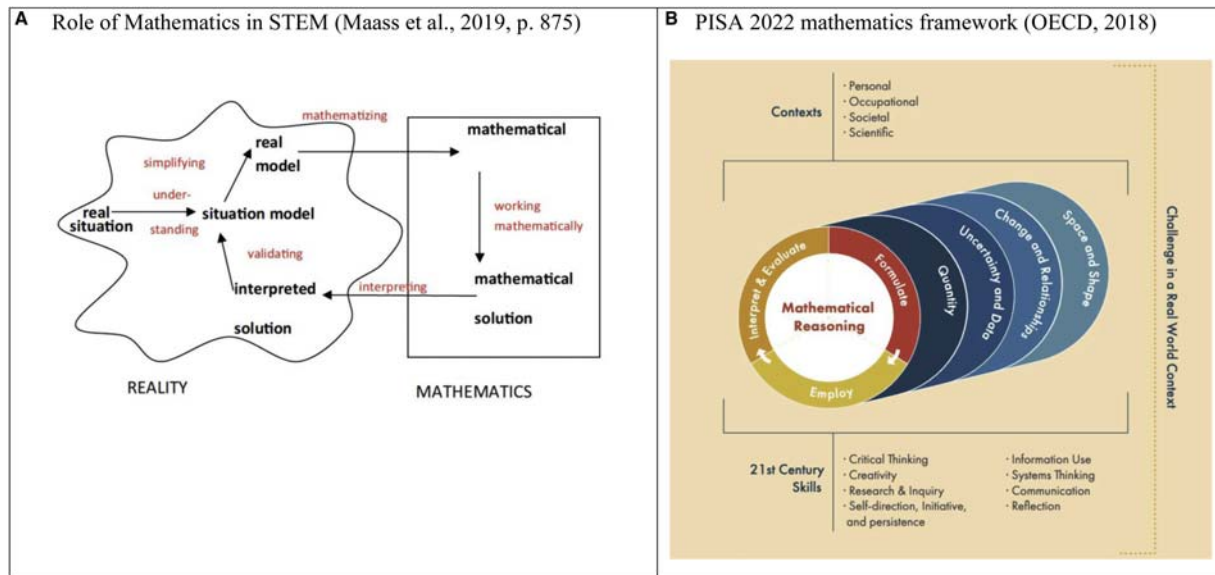


Fig. 2 Theoretical frameworks used in this study. (A) Role of Mathematics in STEM (Maass et al., 2019a,b, p. 875). (B) PISA 2022 mathematics framework (OECD, 2018).

Mathematics as a language tool in STEM education: the content

Mathematics serves as a language tool to describe, represent, and communicate about STEM activities (e.g., Dantzig and Mazur, 2007). Just like any other language that has a vocabulary, grammar, and syntax, mathematics has numbers, symbols, and concepts that operate upon certain rules. Using this mathematical language, one can represent and model real-world situations in the STEM field, producing results that can be discussed globally. In K–12 classrooms, each of the NCTM's (2000) five content standards (or PISA 2022 framework's four content areas) were involved in the reported STEM activities. For instance, NCTM's teaching journal, *Mathematical Teaching in the Middle School*, published a special issue in 2013 that contained six studies. Students in these studies participated in activities such as "building a robot, playing a battleship-like game, designing a staircase, analyzing crystalline-structure data, constructing weight-bearing tables out of paper, and designing a skate park" (Shaughnessy, 2013). Through these activities, students were able to employ a variety of mathematical content including geometry (or "space and shape"; Allen, 2013; Enderson and Grant, 2013); algebraic concepts (or "changes and relationships") such as Cartesian coordinates (Kranz et al., 2013), slope (Smith et al., 2013), ratio and proportional thinking (Stoner et al., 2013); statistical concepts (or "uncertainty and data") such as average (Magiera, 2013); and measurement (Enderson and Grant, 2013; Stoner et al., 2013). Similarly, the journal of ZDM—*Mathematics Education* published a special issue on STEM education in 2019 containing three studies conducted in elementary schools. The STEM activities (e.g., using motion detectors, computer programming-coding, and measuring plant growth) involved varied mathematical concepts such as reasoning about graphs (Duijzer et al., 2019), pattern recognition (Miller, 2019), and measurement using rulers (Geiger, 2019). Of course, in all of the previously mentioned activities, numbers and operations (or "quantity") were involved as part of the computations. However, since these examples depict a rich but broad picture of the relationship between mathematics and STEM education, the precise contributions of mathematics to STEM activities may not yet be clear. We would like to highlight two major roles that mathematics content can play in STEM education: (a) representing and modeling, and (b) interpreting and explaining.

Representing and modeling

First, mathematical content can play a role in STEM activities when concepts are used to help represent and discuss real-world situations (see Fig. 2A). Consider, for example, creating and analyzing graphs for STEM activities that involve technology. Duijzer et al. (2019) reported how primary school students (fifth graders) worked in front of a motion sensor to generate distance-time graphs. A similar activity was reported by Smith et al. (2013) with eighth graders. Using students' motions (moving, moving slowly, spinning, and standing still), a motion detector generated different graphs that represented the relationship between distance and time. In these STEM activities, mathematics (in this case, graphs) was used to describe students' body movement or the distance they walked in relation to time. As acknowledged by both studies, the underlying mathematics concept behind these graphs is related to the concept of slope and rate of change. This is an important algebraic concept in middle school and beyond (Leinhardt et al., 1990). In both studies, some of the generated graphs indicated a linear relationship while others indicated a non-linear relationship. Together, these graphs showed the functional relationship between the two related variables (time, distance), which are just as important as their algebraic representations (Leinhardt et al., 1990). STEM activities without the support of technology also demand students' use of mathematics to represent and model them. Enderson and Grant (2013) introduced an engineering design activity in

which seventh graders were asked to construct a paper table. To prepare them for their table design, students were first instructed to engage in exploring the properties of various polygons (e.g., quadrilaterals) and the connections between their strength and polygon properties. These geometrical concepts provided students with the necessary tools to successfully represent their table designs and to communicate with their peers about their design products.

Interpreting and explaining

The second role that mathematics content can perform in STEM activities is interpretation and explanation. How did a STEM phenomenon occur? Why did it occur? Will the observed pattern always be true? Why is one situation/design better than the other? How can adjustments make it better? To best understand these STEM activity questions, mathematical content/concepts serve as critical elements for reasoning, which in turn contribute to the interpretation of STEM phenomena (see Fig. 2A). Geiger (2019) shared how a science teacher, Stef, guided her first graders to explore the growth of bean sprouts under different levels of sunlight while holding other variables constant. To discuss the critical question, “What did we do/what could we have done to make the seeds grow better?” (p. 935), the class measured and compared the length of the bean sprouts under different conditions. This provided mathematics-based evidence (length determined by measurement) to help interpret and explain the science behind this phenomenon. In addition, Smith et al. (2013) reported an engineering activity in which eighth graders were instructed to explore the steepness of staircase. Students worked in small groups to measure the height and depth of two steps on a staircase. They were also asked to compare the ratio of the two measurements and discuss how they can make the staircase more or less steep. Students produced a variety of observations, yet only one group noticed that the ratio of the two measures remained the same. A follow-up discussion revealed students’ misconceptions about how to correctly use a ruler (e.g., measuring at the end of the ruler rather than the “0” marking) and why students produced inconsistent results (due to the use of two different units, inches and centimeters). After addressing the mathematical obstacles, the class reached the consensus that the ratio between height and depth would stay the same. This shed light on the underlying concept of slope that further explained the steepness of the staircase in an engineering context. In fact, students in that class also mentioned other real-world situations such as wheelchair-access ramps and how they can be explained by the same mathematical concept. In both examples above, the targeted mathematical concepts (length, slope) served as tools to help interpret and explain the science phenomenon and engineering design.

Mathematics as a thinking tool in STEM education: the process

In addition to serving as a language tool, mathematics also serves as a thinking tool for STEM activities. As Honey et al. (2014) noted, STEM activities are typically accomplished through the use of problem, project, and design-based tasks that are situated in real-world contexts. As such, problem solving is an important process that occurs in K–12 STEM classrooms, as illustrated in published empirical studies (e.g., Allen, 2013; Geiger, 2019; Magiera, 2013; Stoner et al., 2013). In the field of mathematics education, problem solving is also viewed as one of the critical mathematical processes for student learning (e.g., NCTM, 2000; National Governors Association Center for Best Practices and the Council of Chief State School Officers [NGAC & CCSO], 2010). As such, mathematical problem-solving competency, if well developed and utilized, can contribute to STEM education. Students in STEM need to make sense of the given real-world situations, formulate mathematical models based on data, and interpret and evaluate mathematical solutions to make critical decisions. As mentioned earlier, these processes in problem solving (formulate, employ, and interpret and evaluate) are clearly highlighted in PISA’s 2022 mathematics framework (see Fig. 2B). In some sense, these processes are similar to Polya’s (1945) seminal work on mathematical problem solving, which stressed four heuristic processes (understand the problem, devise a plan, carry out a plan, and look back) and general strategies (e.g., look for a pattern, draw a diagram, use a model, work backwards). However, mathematics education research shows that a focus on heuristics and general strategies does not contribute to students’ problem-solving competency (Begle, 1979; Cai, 2003; Cai and Lester, 2010; Lesh and Zawojewski, 2007; Schoenfeld, 1992; Silver, 1985). In the same vein, a focus on general mathematical problem solving processes may not contribute to STEM education. This might also be a reason why PISA’s mathematical framework has placed “mathematical reasoning” in the center of the problem solving cycle (see Fig. 2B). In addition, with the increasing use of information, computer, and technology (ICT) in current K–12 classrooms, computational thinking adds one more layer of meaning to problem solving and reasoning. PISA’s 2022 mathematics framework clearly expects that students understand computational thinking and can apply mathematical reasoning in technological contexts (OECD, 2018). Computational thinking is not only expected for the upper grade students (e.g., high school and college) but also for children in early years (Grover and Pea, 2013, 2018; Lye and Koh, 2014; Wing, 2006, 2008, 2011). Given that mathematical reasoning and computational thinking are critical (Binkley et al., 2012; Maass et al., 2019b; OECD, 2018), below we discuss how these two thinking skills play a role in K–12 STEM classrooms.

Mathematical reasoning

Mathematical reasoning can be either deductive or inductive, which is enabled by the following key understandings that undergird school mathematics (OECD, 2018):

- Understanding quantity, number systems and their algebraic properties;
- Appreciating the power of abstraction and symbolic representation;
- Seeing mathematical structures and their regularities;

- Recognizing functional relationships between quantities;
- Using mathematical modeling as a lens onto the real world; and
- Understanding variation as the heart of statistics (p. 15).

As such, instead of teaching general problem solving strategies, current mathematics education advocates teaching mathematics through problem solving (e.g., Cai, 2003; Cai and Lester, 2010; Cai et al., 2005; English and Sriraman, 2010; Lester, 1994), which lends opportunities to stress mathematical reasoning as a key process. Through solving mathematically rich problems, students will learn mathematical concepts and ideas, which in turn develops students' key understandings and reasoning skills (Cai, 2003; Cai and Lester, 2010). By "mathematically rich problems," Cai and Lester (2010) referred to tasks that are "open ended and allow for multiple correct answers and multiple solution approaches" (p. 257). However, as reviewed earlier, many STEM lessons do not actually involve mathematical reasoning. Rather, mathematics is only presented and utilized as a set of procedures or tricks (Wong and Dillion, 2019). In fact, this is also an issue for some solely mathematics lessons in K–12 classrooms (Li and Schoenfeld, 2019).

To address these issues, there is a call for "modeling" (English, 2016; Maass et al., 2019b) and "sense-making" (Li and Schoenfeld, 2019; Zhao and Schuchardt, 2021) to advance students' mathematical problem solving and reasoning. Doerr and English (2003) defined models as "systems of elements, operations, relationships, and rules that can be used to describe, explain, or predict the behavior of some other familiar system" (p. 112). English and Sriraman (2010) explicitly emphasized that "modeling problems are realistically complex situations where the problem solver engages in mathematical thinking beyond the usual school experience" (p. 273). This modeling process demands that students awaken and use their reasoning skills. In fact, English also provided examples regarding how mathematical models and reasoning are used with available data to make critical decisions for realistically complex problems such as an aerospace task (English and King, 2015). With regard to sense-making, various models were also developed to illustrate how reasoning skills are needed for both mathematics and STEM lessons (Li and Schoenfeld, 2019; Schoenfeld, 2020; Zhao and Schuchardt, 2021).

Unfortunately, these types of mathematical reasoning skills are often missing from STEM activities. Consider, for example, the aforementioned motion sensor graphs. In some of the STEM lessons (Duijzer et al., 2019; Smith et al., 2013), the technology-generated graphs were directly given to students. While the power of technology should be acknowledged, these instances also show how mathematical opportunities could have been better grasped. Leinhardt et al. (1990) pointed out that computer-generated graphs may not support student learning in subtle ways; rather, "students may develop too strong a reliance on the machine. For example, they may not understand the underlying patterns and principles that drive the production of the graphs" (p. 7). Even without technology support, some STEM teachers still simply give out graphs or needed mathematical procedures to complete science activities. Wong and Dillion (2019) reported that a secondary mathematics teacher in England made the following reflection after she observed a science lesson in her school:

And rather than talk to the students about what sort of data they were collecting and therefore what would be the best way of them representing that data and getting them to think about it [...] they were told instantly [...] "and at the end of this, you are going to draw a graph and it's going to look like this." And I just thought it was a missed opportunity [...] "Why would you want to do this graph?" [...] that discussion about why and where that comes from, from the data, those discussions didn't tend to happen (p. 793).

Interestingly, in the same study, one science teacher complained about how her secondary students lacked the ability to create graphs: "loads of them had issues drawing the axes on their graph and doing the scale properly, then plotting the graph, then drawing the line of best fit." (p. 790). Due to the lack of necessary mathematical skills, the science teacher was unable to complete her science learning objectives. The above stories indicate how mathematical reasoning should be utilized and facilitated during the process of solving STEM problems.

Computational thinking (CT)

Computational thinking is a critical way of problem solving that involves many elements of 21st century skills (e.g., systems thinking, information use, critical thinking, creativity, communication, see Fig. 2B). In her seminal work, Wing (2006) clarified what CT is and is not. CT is a way that humans solve problems rather than trying to get humans to think like computers. It is about conceptualizing rather than programming. Being able to program a computer is not CT; CT demands multiple levels of abstraction. In other words, even though CT involves the use of computers, it is different from the use of technology. Additionally, CT is not exactly the same thing as mathematical thinking (Grover and Pea, 2013) even though "its formal foundations rest on mathematics" (Wing, 2006, p. 35).

So, what is computational thinking? Although there is no consensus on the definition of CT (Cutumisu et al., 2019), Wing (2011) defined computational thinking as "the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent" (p. 1). Aho (2012) further updated this definition, making its mathematical feature more explicit: "computational thinking is the thought processes involved in formulating problems so their solutions can be represented as computational steps and algorithms" (p. 832). An algorithm "is an abstraction of a step-by-step procedure for taking input and producing some desired output" (Wing, 2008, p. 3718). This is clearly related to functional thinking in mathematics. These abstractions (computational steps and algorithms) are the essence of computational thinking (Aho, 2012; Grover and Pea, 2013). Wing (2011) explained that abstraction

is about defining patterns and generalizing from the specifics. [Grover and Pea \(2018\)](#) stated that abstraction “is a general property of symbolic systems that has been leveraged in arithmetic, algebra and other mathematical sub-domains for centuries” (p. 25). These definitions indicate how closely CT is related to mathematics.

Since [Wing \(2006\)](#) views computational thinking as “for everyone, everywhere” (p. 35), computational thinking has drawn increased attention in K–12 classrooms ([Barr et al., 2011](#); [Grover and Pea, 2013, 2018](#); [Lye and Koh, 2014](#); [NRC, 2012](#)). The [National Research Council \(2012\)](#) in its framework for K–12 science education listed “mathematics and computational thinking” as one of the eight practices for science and engineering:

Mathematics (including statistics) and computational tools are essential for data analysis, especially for large data sets. The abilities to view data from different perspectives and with different graphical representations, to test relationships between variables, and to explore the interplay of diverse external conditions all require mathematical skills that are enhanced and extended with computational skills (p. 65).

More recently, [Grover and Pea \(2018\)](#) proposed a competency framework that contained CT concepts and CT practices, which enabled further research on CT in K–12 classrooms. Among the six CT concepts, four of them—logic and logical thinking, algorithms and algorithmic thinking, patterns and pattern recognition, abstraction and generalization—are mathematically driven. Empirical studies also indicate that different tools are used in K–12 classrooms to develop students’ CT and associated STEM skills. For instance, graphical programming environments such as Scratch and Alice are widely used in classrooms to allow early experiences with CT ([Grover and Pea, 2013](#)). [Miller \(2019\)](#) reported how second graders (ages 7–8) in Australia were taught to do coding activities using Scratch or coding robots. Through these activities, students were expected to develop not only computational thinking but also mathematical understanding of patterning and structures as well as generalization skills, which are underpinnings of algebraic thinking ([Kaput et al., 2008](#)).

As grade level increases, robotic kits are used more commonly in classrooms. For instance, [Allen \(2013\)](#) reported how middle school students learned to create a robot in mathematics classrooms. Consider, for example, if students wanted their robot to “launch via a clap, roll forward 10 inches, make a 90° right turn, continue rolling until it hits a black line, and stop” (p. 343). To make their robot carry out these actions, students needed to decompose this task into smaller steps and write icon-based programs. This is where mathematics (or “abstraction”) is involved. For instance, students need to figure out how far the robot travels with one rotation of the wheels (circumference) and how many rotations are needed to complete 10 inches. They also need to use multiple sensor inputs, define variables, and reason proportionally. Furthermore, they need to use inequalities to make their robot start, turn, and stop in the exact desired locations. As [Allen \(2013\)](#) reported, through the above robotics activities, students were motivated to learn mathematics and develop their CT skills. This example addresses the challenge proposed by [Wing \(2008\)](#) on computational thinking, “we do not want the tool to get in the way of understanding the concepts. We also do not want people just to be able to use the tool but not have learned the concepts” (p. 3721). It should be noted that in this study, students’ computational and mathematical thinking also contributed to other STEM activities (e.g., the engineering component of building a robot).

Implications: what do we learn from the current research?

The above discussion of the roles of mathematics in STEM, as (a) a language tool and (b) a thinking tool, provides direction for enhancing and strengthening future STEM lessons. As mentioned earlier, the current field is largely concerned that mathematics is often diminished and lost in STEM activities ([English, 2016](#); [Honey et al., 2014](#); [Larson, 2017](#); [NCSM and NCTM, 2018](#); [Maass et al., 2019b](#); [Shaughnessy, 2013](#); [Silk et al., 2010](#)). Perhaps STEM classrooms can devote more time to explicitly addressing the above two roles that mathematics play in STEM activities. Below we have highlighted a few specific in-class actions that can help educators work toward this goal.

Setting up grade-appropriate mathematics learning goals in STEM lessons

First, there should be clear mathematical learning objectives in each STEM lesson. This is because mathematics is a foundation of STEM education ([NCTM and NCSM, 2018](#)). Furthermore, setting goals can help to ensure that mathematics learning also benefits from STEM lesson activities ([Honey et al., 2014](#)). More importantly, the mathematics involved should be targeted toward the content standards appropriate for each grade level ([Larson, 2017](#); [NCTM and NCSM, 2018](#)). To illustrate this point, we revisit the aforementioned articles by [Duijzer et al. \(2019\)](#) and [Smith et al. \(2013\)](#). Both articles used a motion detector to generate distance-time graphs. While the former study involved fifth graders (primary in Australia), the latter involved eighth graders (secondary in the US). Likely due to the grade-level differences, the former study only focused on the interpretation of the graphs while the latter brought forth the concept of slope. Although both studies provide interesting findings, we are uncertain of whether the activity in [Duijzer et al. \(2019\)](#) was grade-appropriate as there was no introduction to Australia’s mathematics standards. In fact, even though this STEM lesson spanned six lessons, mathematics content was limited to graphs without a connection to any equations. This, again, is likely due to the grade-level constraint. As such, it is important to consider grade level when choosing STEM activities.

Setting up grade-level-appropriate math learning goals is especially important when taking STEM activities from other countries. For instance, an inspection of the Common Core State Standards in the USA indicates that students are expected to understand the

relationship between dependent and independent variables (using graphs, tables, and equations) in sixth grade. In addition, it is not until eighth grade that students are expected to graph proportional relationships and interpret the unit rate as the slope of the graph. Therefore, if a teacher is interested in using activities similar to those found in [Duijzer et al. \(2019\)](#), it might be better to postpone them to a later grade. In doing so, teachers could plan deeper, more thorough learning objectives and could make sure that graphs could be linked to mathematical equations. This linkage between graphs and equations is critical for future algebraic learning, one example being proportional reasoning and functions ([Leinhardt et al., 1990](#)). The grade-level appropriate learning goals are also a challenge when developing computational thinking in K–12 classrooms, which is a quite new area that demands much attention for both research and practice ([Wing, 2008](#)).

Promoting explicit mathematical connection-making in STEM lessons

Our second suggestion is to make explicit connections between the concrete representations (e.g., real-world situations, graphs, tools) and abstract mathematical models/solutions through lesson discussion. We believe such a connection-making structure not only highlights the role of mathematics in STEM, but also promotes students' mathematical reasoning and problem-solving skills. In the aforementioned staircase example that led to the concept of slope, [Smith et al. \(2013\)](#) reported that the class discussed slope "from both geometric and algebraic standpoints." However, it is unclear whether the staircase was modeled with geometrical figures, if an algebraic equation (e.g., $y = mx$) was introduced, and whether the geometric and algebraic representations were explicitly connected for students to best learn about the concept of "slope." In fact, the eighth grade CCSS standards suggest that teachers "use similar triangles to explain why the slope m is the same between any two distinct points on a non-vertical line in the coordinate plane" ([NGAC and CCSSO, 2010](#), p. 54). In the staircase activity, students were asked to compute different ratios between the height and depth of the steps. If a teacher was to guide students and explicitly model this activity, students would generate similar triangles that could be recorded algebraically (e.g., $\frac{y^2 - y^1}{x^2 - x^1} = \frac{y^4 - y^3}{x^4 - x^3}$). This proportional reasoning can, then, be used to explain why the slope (m) remained the same. In this way, geometric and algebraic connections are made explicit.

Another example that stresses explicit mathematical discussions is related to the measurement activity that frequently occurred in science classrooms. Studies across grade levels (from elementary to secondary schools) reported that students had misconceptions with using rulers (e.g., not starting with the marking of "0") ([Geiger, 2019](#); [Wong and Dillion, 2019](#); [Zhang et al., 2015](#)). For this reason, teachers in STEM classrooms should not assume that students have these basic skills to use these mathematics tools. Rather, explicit classroom discussions need to occur either before or after the STEM activities. Depending on the grade level, a teacher may explicitly prompt students to consider the following questions: Why can we not measure an object from the end of a ruler? What can we do if a ruler does not include the marking "0"? What does a marking on a ruler actually mean? What is the relationship between the unit sizes (e.g., inches or centimeters) and the resulting measurements (the inverse relation of a bigger size yielding smaller measures)? Discussions of these seemingly simple content questions will not only ready students for STEM activities, but also prompt their mathematical reasoning skills. Likewise, in technology-related classrooms, the grade-level computational concepts (if they exist) should also be made explicit for students, which is a challenge posed by [Wing \(2008\)](#) in terms of balancing between tools and concepts.

Developing a habit of mind of mathematical reasoning for STEM lessons

To develop students' mathematical reasoning, we think that there is a need to cultivate such a habit of mind even outside the STEM classes. [English \(2016\)](#) urged the field to explore "how we might best develop in-depth understanding of core mathematics content and processes within STEM experiences, while at the same time acknowledge that not all of mathematics can or should be learned within an integrated program" (p. 4). In other words, mathematical reasoning skills (including sense-making and modeling) can be developed in regular mathematical lessons through problem solving. [Ding \(2021\)](#) reported that Chinese expert teachers prompted students to attempt to reason about a problem situation in different ways, resulting in different solutions. This was even done with routine word problems, those with only one correct answer. Such reasoning skills, once becoming a habit of mind, will contribute to STEM lessons. In fact, mathematical reasoning demands students' sense-making, which can also be developed in regular mathematics lessons. [Ding \(2021\)](#) described a Chinese fourth grade lesson in which students were formally introduced to the commutative property of addition. In this lesson, the teacher changed a "." to a "?" at the end of a written statement about the commutativity, which was derived from students' discoveries based on one real-world example. This action elicited a heated class discussion, "How can we make a generalization based on just one example?" "Will this always be true?" "Can we find more examples?" "Can we exhaust all examples?" "How can we prove it?" "If we cannot prove it, can we find a counter-example to disprove it?" These reasoning, sense-making, and critical thinking skills, as prompted by the teacher's deep questions, are vital for conducting STEM activities. For instance, in [Smith et al. \(2013\)](#), students identified that the ratio between height and depth are the same. Using the above Chinese lesson example, we believe that students could have been challenged to ask deep questions such as: "Is this always true? Why is it always true?" Such questions might lead to further investigation of the models and connection-making between the geometric and algebraic equations, as discussed in an earlier section. Indeed, such reasoning and generalization skills are also vital for computational thinking when solving problems in complex systems.

Conclusion and future directions

In this study, we have discussed, mainly based on a review of literature, how mathematics can play a role in STEM activities in K–12 classrooms. From the perspective of content, we argue that mathematics serves as a language tool. This is because mathematical concepts help describe and represent the STEM activities and can be used to explain and interpret the identified solutions. From the perspective of process, we argue that mathematics serves as a thinking tool because mathematical reasoning is at the core of problem solving. In particular with the current complex systems, computational thinking that heavily relies on mathematics is critical. To illustrate our point of view, we have analyzed empirical studies, which indicate where mathematics can play a role in STEM activities and how mathematics can be made even more explicit. Based on the reviews of the existing literature, we have proposed three in-class specific actions, which hopefully will help make mathematics explicit in K–12 STEM classrooms.

We are cognizant that to make the above suggestions happen, there is a great need for teacher development and curriculum change. For instance, do our school curricula provide a space to make in-depth STEM integration? A science teacher may be concerned about limited class time to cover required science concepts, which may hinder the action of making mathematics explicit (Wong and Dillion, 2019). One alternative is to incorporate one carefully selected STEM lesson at the end of each unit in mathematics and science classes as an application of knowledge and development of STEM skills. Another possible alternative is to offer a STEM course beyond the individual disciplinary courses. These possible curriculum changes also propose challenges for teacher development and teacher education because STEM courses demand teachers' well-rounded expertise in both mathematics and other STEM areas. We call for future studies to continue this line of research to leverage classroom practices that will contribute to developing students' 21st century skills.

Acknowledgment

This work is partially supported by the National Science Foundation in the USA (No. 1350068) at Temple university. Any opinions, findings, and conclusions in this study are those of the authors and do not necessarily reflect the views of the funding agency. The authors are grateful to Caroline Bryn-Mawr Driscoll, Monica Manfredonia, and Cassy Pressimone Beckowski for editorial assistance.

References

- Aho, A.V., 2012. Computation and computational thinking. *Comput. J.* 55, 832–835.
- Allen, K., 2013. Robots bring math-powered ideas to life. *Math. Teach. Middle Sch.* 18, 340–347.
- Barr, D., Harrison, J., Conery, L., 2011. Computational Thinking: A Digital Age. *Learning & Leading with Technology* March/April, pp. 20–23.
- Becker, K., Park, K., 2011. Effects of integrative approaches among Science, Technology, Engineering, and Mathematics (STEM) subjects on students' learning: a preliminary meta-analysis. *J. STEM Educ.* 12, 23–37.
- Begle, E.G., 1979. Critical Variables in Mathematics Education. the Mathematics Association of America and the National Council of Teachers of Mathematics, Washington D.C.
- Binkley, M., Erstad, O., Heramn, J., et al., 2012. Defining twenty-first century skills. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht, pp. 17–66.
- Cai, J., Lester, F., 2010. Why is Teaching With Problem Solving Important to Student Learning? National Council of Teachers of Mathematics, Reston, VA. Retrieved from: <https://www.washoeschools.net/cms/lib/NV01912265/Centricity/Domain/253/Math%20K-6/NCTM%20research%20brief.pdf>.
- Cai, J., Mamona-Downs, J., Weber, K., 2005. Problem solving research in mathematics education: what we have done and where we are going. *J. Math. Behav.* 24, 217–220.
- Cai, J., 2003. What research tells us about teaching mathematics through problem solving. In: Lester, F. (Ed.), *Research and Issues in Teaching Mathematics Through Problem Solving*. National Council of Teachers of Mathematics, Reston, VA, pp. 241–254.
- Cutumisu, M., Adams, C., Lu, C., 2019. A scoping review of empirical research on recent computational thinking assessments. *J. Sci. Educ. Technol.* 28, 651–676.
- Dantzig, T., Mazur, J., 2007. *Number: The Language of Science*. Penguin.
- Ding, L., 2021. Teaching Early Algebra Through Example-Based Problem Solving: Insights From Chinese and U.S. Elementary Classrooms. Routledge, New York, NY.
- Doerr, H.M., English, L.D., 2003. A modeling perspective on students' mathematical reasoning about data. *J. Res. Math. Educ.* 34 (2), 110–136.
- Duijzer, C., Van den Heuvel-Panhuizen, M., Veldhuis, M., Doorman, M., 2019. Supporting primary school students' reasoning about motion graphs through physical experiences. *ZDM* 51, 899–913.
- Enderson, M.C., Grant, M.R., 2013. Emerging engineers design a paper table. *Math. Teach. Middle Sch.* 18, 362–369.
- English, L.D., King, D.T., 2015. STEM learning through engineering design: fourth-grade students' investigations in aerospace. *Int. J. STEM Educ.* 2 (14), 1–18.
- English, L.D., Kirshner, D., 2016. Changing agendas in international research in mathematics education. In: English, L.D., Kirshner, D. (Eds.), *Handbook of International Research in Mathematics Education*, third ed. Taylor & Francis, New York, pp. 3–18.
- English, L., Sriraman, B., 2010. Problem solving for the 21st century. In: Sriraman, B., English, L. (Eds.), *Theories of Mathematics Education*. Springer, New York, NY, pp. 263–285.
- English, L.D., 2016. STEM education K-12: perspectives on integration. *Int. J. STEM Educ.* 3 (3), 1–8.
- Fitzallen, N., 2015. STEM education: what does mathematics have to offer? In: Marshman, M., Geiger, V., Bennis, A. (Eds.), *Mathematics Education in the Margins (Proceedings of the 38th Annual Conference of the Mathematics Education Research Group of Australasia)*. Sunshine Coast: MERGA, pp. 237–244.
- Fitzpatrick, R., 2007. Euclid's Elements of Geometry: The Greek Text of J.L. Heiberg (1883–1885). Retrieved from: <https://farside.ph.utexas.edu/Books/Euclid/Elements.pdf>.
- Geiger, V., 2019. Using mathematics as evidence supporting critical reasoning and enquiry in primary science classrooms. *ZDM* 51, 929–940.
- Grover, S., Pea, R., 2013. Computational thinking in K–12: a review of the state of the field. *Educ. Res.* 42 (1), 38–43.
- Grover, S., Pea, R., 2018. Computational thinking: a competency whose time has come. In: Sue Sentence, S., Barendsen, E., Schulte, C. (Eds.), *Computer Science Education: Perspectives on Teaching and Learning in School*. Bloomsbury Academic, London, pp. 19–37.
- Honey, M., Pearson, G., Schweingruber, A., 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. National Academies Press, Washington.
- Johnson, C.C., Mohr-Schroeder, M.J., Moore, T.J., English, L.D., 2020. *Handbook of Research on STEM Education*. Routledge, New York, NY.
- Kaput, J.J., Carraher, D.W., Blanton, M.L., 2008. *Algebra in the Early Grades*. Lawrence Erlbaum Associates, New York, NY.
- Kranz, S.R., Amato, C.A., Freudenthal, E.A., 2013. Coordinate an attack using the calculator. *Math. Teach. Middle Sch.* 18, 356–361.

- Larson, M., 2017. Math Education is STEM Education! NCTM President's Message. Retrieved from. <https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Matt-Larson/Math-Education-Is-STEM-Education/>.
- Leinhardt, G., Zaslavsky, O., Stein, M., 1990. Functions, graphs and graphing: tasks, learning and teaching. *Rev. Educ. Res.* 60 (1), 1–64.
- Lesh, R., Zawojewski, J.S., 2007. Problem solving and modeling. In: Lester, F. (Ed.), *The Second Handbook of Research on Mathematics Teaching and Learning*. Information Age Publishing, Charlotte, NC, pp. 763–804.
- Lester, F.K., 1994. Musings about mathematical problem-solving research: 1970–1994. *J. Res. Math. Educ.* 25, 660–675.
- Li, Y., Schoenfeld, A.H., 2019. Problematising teaching and learning mathematics as “given” in STEM education. *Int. J. STEM Educ.* 6 (44), 1–13.
- Lye, S.Y., Koh, J.H.L., 2014. Review on teaching and learning of computational thinking through programming: what is next for K-12? *Comput. Hum. Behav.* 41, 51–61.
- Maass, K., Doorman, M., Jonker, W., Wijers, M., 2019a. Promoting active citizenship in mathematics teaching. *ZDM* 51, 991–1003.
- Maass, K., Geiger, V., Ariza, M.R., Goos, M., 2019b. The role of mathematics in interdisciplinary STEM education. *ZDM* 51, 869–884.
- Magiera, M.T., 2013. Model eliciting activities: a home run. *Math. Teach. Middle Sch.* 18, 348–355.
- Miller, J., 2019. STEM Education in the primary years to support mathematical thinking: using coding to identify mathematical structures and patterns. *ZDM* 51, 915–927.
- National Council of Supervisors of Mathematics and National Council of Teachers of Mathematics (NCSM & NCTM), 2018. Building STEM Education on a Sound Mathematical Foundation. Retrieved from. <https://www.nctm.org/Standards-and-Positions/Position-Statements/Building-STEM-Education-on-a-Sound-Mathematical-Foundation/>.
- National Council of Teachers of Mathematics (NCTM), 2000. Principles and Standards for School Mathematics. National Council of Teachers of Mathematics (NCTM), Reston, VA.
- National Governors Association Center for Best Practices and the Council of Chief State School Officers (NGAC & CCSSO), 2010. Common Core State Standards for Mathematics. Retrieved from. http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf.
- National Research Council (NRC), 2012. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. The National Academies Press, Washington, DC.
- National Research Council (NRC), 2014. STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research. The National Academies Press, Washington, DC.
- Organization for Economic Co-operation and Development (OECD), 2018. PISA 2022 Mathematics Framework. Retrieved from. <https://pisa2022-maths.oecd.org/index.html>.
- Polya, G., 1945. *How to Solve It*. Princeton University Press, Princeton, NJ.
- President's Council of Advisors on Science and Technology (PCAST), 2010. Report to the President: Prepare and Inspire: K–12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future. Retrieved from. <http://www.afterschoolalliance.org/documents/pcast-stemed-report.pdf>.
- Schoenfeld, A.H., 1992. Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In: Grouws, D.A. (Ed.), *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*. Macmillan Publishing, New York, NY, pp. 334–370.
- Schoenfeld, A.H., 2020. Mathematical practices, in theory and practice. *ZDM* 52, 1163–1175.
- Shaughnessy, M., 2013. Mathematics in a STEM context. *Math. Teach. Middle Sch.* 18, 324.
- Silk, E.M., Higashi, R., Shoop, R., Schunn, C.D., 2010. Designing technology activities that teach mathematics. *Technol. Teach.* 69 (4), 21–27.
- Silver, E.A., 1985. Research on teaching mathematical problem solving: some under represented themes and needed directions. In: Silver, E.A. (Ed.), *Teaching and Learning Mathematical Problem Solving: Multiple Research Perspectives*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 247–266.
- Smith, T.M., Seshaiyer, P., Peixoto, N., et al., 2013. Exploring slope with stairs & steps. *Math. Teach. Middle Sch.* 18, 370–377.
- Stoner, M.A., Stuby, K.T., Szczepanski, S., 2013. The engineering process in construction & design. *Math. Teach. Middle Sch.* 18, 332–338.
- Wing, J., 2006. Computational thinking. *Commun. ACM* 49 (3), 33–35.
- Wing, J., 2008. Computational thinking and thinking about computing. *Phil. Trans. Math. Phys. Eng. Sci.* 366, 3717–3725.
- Wing, J., 2011. Research Notebook: Computational Thinking—What and Why? The Link: The magazine of Carnegie Mellon University's School of Computer Science. Retrieved from. <https://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why>.
- Wong, V., Dillon, J., 2019. “Voodoo maths”, asymmetric dependency and maths blame: why collaboration between school science and mathematics teachers is so rare. *Int. J. Sci. Educ.* 41, 782–802.
- Zhang, D., Orrill, C., Campbell, T., 2015. Using the mixture Rasch model to explore knowledge resources students invoke in mathematics and science assessments. *Sch. Sci. Math.* 115, 356–365.
- Zhao, F., Schuchardt, A., 2021. Development of the sci-math sensemaking framework: categorizing sensemaking of mathematical equations in science. *Int. J. STEM Educ.* 8 (10), 1–18.

The rise of STEM education: mathematics learning for STEM

Peter Appelbaum, Arcadia University, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Mathematics for STEM versus STEM for mathematics	29
Notion of learning for the sake of learning	29
Citizenship STEM or cultivation of professionals	29
STEM and its relationship with the STEM disciplines	30
Ways of working as a mathematician	30
Problem solving	30
Generalizing and specializing to generate and test conjectures	31
Problem posing	31
Different kinds of models for different sorts of purposes	32
Social justice, mathematics and STEM	32
Cross-cutting concepts and mathematizing	33
References	34
Relevant Websites	35

Glossary

Coloniality those social, institutional, economic and cultural structures that preserve and perpetuate forms of inequity endemic to colonialism after official geopolitical colonial status is no longer present

Critical Mathematics the study of mathematics as inherently political, cultural, and ideological, structuring what is understood as truth, knowledge, and forms of argument, with implications for social justice and equity

Ethnomathematics the recognition that forms of mathematization exist outside of the history of Western (school) mathematics, and that colonialism and coloniality have erased knowledges and practices, replacing them with a global assumption that mathematics is universal and neutral

Ethnomodelling the association between ethnomathematics and mathematical modeling that enables members of distinct cultural groups to perceive a different reality in relation to the nature of mathematical knowledge

Mathematization the process by which members of a distinct cultural group think and act mathematically in order to translate or interpret their own surroundings by applying mathematical knowledge, most notably universal abilities that all cultures and peoples use that of counting, measuring, classifying, patterning, gaming, quantifying, and modeling

Problem Posing a strategy for mathematizing in which one asks what-if-not about any attribute of a situation or initial question, in order to generate new problems to pursue

Mathematics is officially one of the STEM areas – science, technology, engineering, and ... mathematics. It is present in the acronym in its own right, yet perhaps also because it is understood as the “language of science” and routinely essential to the pursuit of the other associated disciplinary fields. Li and Schoenfeld (2019) refer to the special status of mathematics as a “gateway” to many scientific and technological fields, so that many future job opportunities and contributions to society are lost if and when students stop studying mathematics. The importance of mathematics is further touted when STEM skills are linked to national and global economic growth or competition (AIG 2013; OCS 2014). Parallel arguments worry about the state of mathematics given an expected increase in the demand for quantitative and computational literacy as this century unfolds, advocating shifts in the ways that mathematics is taught in order to “make mathematics a magnet” that attracts students to STEM fields and careers (Committee, 2018). It is in this latter perspective that it is often suggested that mathematics is more attractive when learned in ways that integrates its skills and concepts across disciplines, and learned through experiential, meaningful, and applied contexts. Such arguments, which see mathematics as foundational to STEM fields, seem to reduce the importance of mathematics as its own field of study, with its own value to individual self-determination and power to solve social problems independent of the other STEM fields. When mathematics is perceived only in terms of its foundational and gateway functions, it tends to be reduced to procedural skills and facts, and loses its potential for modeling, conceptual understanding and sense-making.

In simplest terms, STEM refers to a collection of separate subject areas, each with their own unique modes of inquiry and methods of discovery, problem-solving, and problem-posing. School curricula are often organized into four distinct sets of objectives, topics of study and pedagogic approaches; alternatively, STEM is potentially a discipline in its own right. STEM as its own stand-alone discipline is commonly criticized as idealistic, since it is claimed that few teachers have sufficient content knowledge

as well as pedagogical content knowledge across all four STEM disciplines to enact such a curriculum (Sanders, 2009; Fitzallen, 2015). Because of this sort of criticism, it is often advocated that STEM be treated as interdisciplinary, as a holistic approach that links the individual disciplines through solving problems that draw on concepts and procedures from the distinct fields in a context of teamwork and design methodology of engineering and suitable applications of technology. Alternatively, some propose a form of integration that does not begin with any of the specific STEM disciplines, but instead identifies knowledge development, synthesis and application in any one of the fields at one time or another, along with focused inquiry meant to support higher order learning, and other forms of rich tasks or applications to real world scenarios. This model of integration seeks to use each of the characteristics of any one or more of the STEM areas to support the development of characteristics, skills, concepts, and forms of inquiry in one or more of the other fields through ways of working that fit particular disciplines at any given time (Treacy and O'Donoghue, 2014).

Mathematics for STEM versus STEM for mathematics

The main question for mathematics in relationship to STEM, whether it is conceived as distinct or integrated disciplines, or as components of a curriculum, is, whether or not it serves mathematics or STEM fields to lump them together. On the one hand, as proponents of STEM note, integrated STEM programs have the potential to foster connectedness and meaningful applications of mathematics that might engage students, leading to greater levels of achievement or at least increased enrollment in mathematics and STEM studies beyond when it is required. At their best, integrated STEM programs make the mathematics and its real-world problem-solving potential both meaningful and accessible (Rennie et al., 2012). However, teachers often lament the loss of the standard mathematics curriculum when they experience the enactment of an integrated STEM program. The mathematics appears “watered down” to its most basic functions of data analysis, counting and measuring (Larson, 2017). The coherence of a mathematics learning program is commonly accused of being lost in the pursuit of STEM projects (Schmidt and Houang, 2007). The potential of mathematics as an aesthetic activity, as a world of inquiry in its own right independent of real-world applications, and as a unique stance to take upon the world, is typically absent, seemingly replaced by the design approaches of engineering, which otherwise would themselves struggle for a presence in the integrated activity; mathematics might in this sense be seen as sacrificed for the ways in which it would most successfully support the development of STEM as a new enterprise (Appelbaum, 2008; Bourdeau and Wood, 2019; Anderson et al., 2020). Furthermore, the inclusion of mathematics in STEM, associated with science, engineering and technology, is more likely to emphasize those aspects of mathematics that are shared with those approaches to knowledge than if it were clustered with, say, the arts, humanities, or social sciences (Sinclair, 2006; Kinniburgh and Byrd, 2008; Sahinkaya and Aladag, 2010; Appelbaum, 2012).

When mathematics is taken as a given necessity for STEM it is rarely questioned whether or not it serves STEM in general to include mathematics in the cluster. As described, the idea that mathematics is foundational to science, engineering and technology is taken for granted. Yet, mathematics is also constructed by common sense discourses as difficult and off-putting. In this way, the gateway function of mathematics is more of a problem to be solved than a contribution to STEM. Li and Schoenfeld (2019) propose a shift away from mathematics as processes and techniques of problem solving useful to science and engineering toward mathematics as a sense-making activity; that is, away from a body of content to be learned or processes in which to engage, toward the codification of experiences both of making sense and sense-making. They worry that much of the inductive parts of mathematics have been lost as it is swallowed up in STEM, but also celebrate the potential contributions of mathematics if it were to be “re-framed.” They also urge a shift in focus, away from classroom instruction, toward the student’s experience of the discipline; in other words, they wish for educators to spend less effort designing the best content sequences for mathematics, or the most optimal methods of making the content accessible to learners, and to exert more research and development work on the ways that educational environments provide opportunities for sense-making grounded in mathematical practices. Such a perspective echoes numerous calls for more attention to student experience (Appelbaum, 2008; Boaler, 2002), and mathematics as a humanistic and meaning-making endeavor (Brown, 1973, 2001), yet by placing these shifts in the context of mathematics and its support for STEM, it opens up a number of interesting, related issues.

Notion of learning for the sake of learning

To what extent is mathematics and/or STEM primarily utilitarian, and to what extent might they be concerned with capturing broader purposes of learning as contributing to a learner becoming more fully human, more capable of taking an ethical stance upon the world, better able to pursue learning for its own sake rather than its potential applications, as socially-constructed and historical forms that contribute to human endeavors and the stewardship of the planet?

Citizenship STEM or cultivation of professionals

To what extent is mathematics and/or STEM education aiming to support broad questions of public participation in and confidence in these subject areas for serving human and global needs, and on the other hand, to what extent is mathematics and/or STEM education primarily focused on the development and training of future scientific and technological professionals? A critical reading of contemporary STEM practices more generally proposes that common school experiences are highly successful in cultivating an

appreciation for the importance of STEM, science and engineering, perhaps even a reverence, and at the same time, an acceptance of others as wielding the STEM knowledge and skills, while the majority of the population remains incapable of truly grasping the details of the science (Beach, 1999; Appelbaum and Clark, 2001; Emdin, 2010). This is in some ways related to the seemingly universal embrace of mathematics as an essential component of an education, and to the equally unquestioned assumption that one or another version of mathematical literacy, numeracy, etc., is mandatory for a happy and meaningful life worth living, yet, the school mathematics curriculum has remained stuck in pre-20th Century content and methods through to contemporary life.

STEM and its relationship with the STEM disciplines

To what extent should STEM be understood as involving practices associated with mathematics, science, technology and engineering themselves, and, given the many different professional practices and philosophical approaches to each discipline, which should be represented in school experiences? Who would decide this, and, how would who is deciding make a difference for what is then constructed by the experiences as “STEM?” (Bybee, 2013) Why would the selections and decisions be chosen, and via what criteria? On the one hand, decisions might be made on the basis of assuming a neutrality of mathematics that ignores its history and continued function as supporting physical and symbolic violence (Davis and Hersh, 1981; O’Neil, 2016). On the other hand, decisions might be grounded in a desire to provide an induction into future professional practice, to a “scientific and engineering literacy” (Jarman and McClune, 2007; Hoisington, 2020; Leung et al., 2020), or a STEM-for-All introduction to content relevant to students’ current everyday life (Gitari, 2016; Cedere et al., 2020; Levrini et al., 2021). Those addressing these questions from the perspective of the T in STEM often feel caught between the academic and vocational arguments for the value of STEM in an educational program, yet propose that the weak disciplinary boundaries of the “technology field” provide a unique opportunity for learners to access a STEM curriculum (Bell et al., 2017).

These issues remain central to the relationship between mathematics and STEM, and in particular, can be the context for asking what STEM might contribute to mathematics, aside from the possibility of contexts for the application of mathematics. Larson (2017), then-President of the US National Council of Teachers of Mathematics, expressed the common litany that “we might neglect the full development of students’ mathematical understanding in order to integrate STEM “activities” into an already narrowly over-packed curriculum,” as well as the anxiety that “STEM education narrowly emphasizes learning mathematics for the workplace and for the scientific and technical communities.” He urged designers of curriculum to create STEM programs specifically tailored to developing the content and practices that characterize effective mathematics programs while maintaining the integrity of the mathematics. In his editorial, from his position as a mathematics educator, he welcomed STEM orchestrations of activity that could be understood as “vehicles to ensure students learn important mathematics at a deep level and are confident in their ability to use mathematics to be empowered in their own lives.”

One possible example of such an approach that enlists STEM activities in the pursuit of mathematics might be found in Romagnano’s (1994) story of a year in a grade 8 classroom, *Wrestling with Change*, in which STEM explorations provided models and metaphors for function concepts, which he believed at the time to be the primary conceptual power of mathematics that should be pursued and consolidated in the middle grades. Another is the story of students studying local geometric and architectural patterns on the Island of Chios in Greece (Stathopoulou et al., 2015), in which the informal knowledge of craftsmen became the basis of a dramatization created by the students of architects, historians, folklorists, and mathematicians; in this curriculum, STEM investigations led to a complex interrogation of the differences between local customs and abstract mathematical approaches. Orey and Rosa (2021) would describe the latter example as a school lesson that incorporated “ethnomodelling,” the association between ethnomathematics and mathematical modeling that enables members of distinct cultural groups to perceive a different reality in relation to the nature of mathematical knowledge. Given the dominant assumption that mathematics is the foundation for STEM, few further examples exist of STEM that serves as foundational for mathematics.

Ways of working as a mathematician

As noted above, mathematics educators often worry that the mathematics itself gets short shrift when it is absorbed into STEM. It tends to be subsumed within a set of practices that are predominantly organized by science processes of inquiry (e.g., question, observation, hypothesis, prediction, experiment, analysis of data, iterate based on results), or design processes aligned with engineering (e.g., identify need, research the problem, develop possible solutions, select the most promising solution, construct a prototype, communicate the design, redesign based on experience with the prototype). With this in mind, it can be useful to intentionally structure STEM activities to conform to processes that mathematicians commonly pursue. Those who study mathematicians notice at least three common options for such a set of processes: problem solving steps, generalizing and specializing, and problem posing.

Problem solving

Perhaps the most well-known and widely applied approach, popularized in the 1940s through Georg Polya’s (1945) *How to Solve It*, is both a summary of four steps or principles that can organize solutions to problems as well as a guide for teachers and other educational facilitators to coaching learners in these steps or principles. The first step is to “understand the problem,” which can be more likely to accomplish by asking oneself one or more of several questions: What are you asked to find or show? Can you restate the

problem in your own words? Is there enough information to enable a solution? The second step is to “devise a plan,” by using a strategy such as guess-&-check, making a systematic list, considering special cases, eliminating possibilities, or by using methods such as looking for a pattern, drawing a picture, using a formula or model, or by first solving a simpler problem. Third, one carries out the plan that was devised – Can those guess-&-check experiments or the collection of special cases be organized systematically? Can you add things to your picture and then search for relationships among the elements of that drawing? Etc. Finally, step four focuses on reviewing and extending the previous steps; potentially the most important step in the process, this involves asking if the answer makes sense, if the experience generates new questions or problems, if you would now take a different tack now that you can look back on what you did, and how you might predict a good strategy for certain problems in the future.

Using the Problem Solving approach, an educator would model the use of the processes and a creative use of the associated questions, through thinking aloud while solving problems in front of learners, and by prompting students to move through the principles when it could be helpful, until they are able to flow through all four steps on their own. Some teachers teach the steps directly, and remind the students to think about where they are in the sequence as they work.

Generalizing and specializing to generate and test conjectures

Sometimes understood as one of Polya’s heuristic techniques, the pursuit of conjectures and efforts to test them and to turn that experience into a convincing argument for the conjecture is in its own right a set of strategies and skills common to mathematical ways of working. An excellent resource for this approach is [Mason et al. \(1982/2010\)](#) *Thinking Mathematically*. This approach shares some characteristics with observational sciences such as field biology or astronomy, where one can collect data but cannot insert interventions. The mathematician collects special cases and organizes results to create categories base on attributes or outcomes. For example, one might explore a situation with small numbers, large numbers, even and odd numbers, primes, and so on; or one might try out different shapes to see how they affect outcomes, two-dimensional, three-dimensional, multi-dimensional. Once one can see patterns in the categories or outcomes, it is possible to formulate a conjecture about the situation; this generalization’ is a potential finding. Further testing of the conjecture might lead to a re-articulation of the basic requirements of the situation, if some special cases do not fit the expectations of the conjecture. Refinement of the conjecture and further testing can lead the mathematician to first convince themselves of the validity of their conjecture, and then to compose a demonstration or argument for a more skeptical audience.

This way of working would reorient some STEM activities away from the specifically utilitarian expectation that STEM solves problems and offers practical applications, toward a more conceptual interest in knowledge building and the re-framing of what is important to know about a situation. It also enables educators to harness reflection on the interaction with mathematical thinking to support learning mathematics and its STEM friends of science, technology and engineering. [Mason et al., \(1982/2010\)](#) report that skillful mathematical thinking can be improved by tacking questions conscientiously, reflecting on that experience, linking feelings with action, studying the process of resolving problems, and noticing how what is learned fits with the experience. They also note the importance of spending time on recalcitrant questions that resist resolution, and have found that answers to questions or solutions to problems are not as important as the experiences of working with the methods of thinking and organizing mathematical activity. With this in mind, it is critical to see that the most powerful mathematical experiences often should be untangled from practical applications and a solutions-fixated problem-solving or design-structured experience. One STEM curriculum strategy is to pursue this in parallel with the S- or E- grounded situation, isolating the mathematics that was used for further, independent explorations.

Problem posing

A seemingly neglected potential contribution of mathematics to STEM learning and practices is the art of problem posing that is at the heart of mathematics ([Brown and Walter 1983/2004](#)). This specifically mathematical collection of techniques for composing new and original questions could transform STEM, which tends to seek designed solutions to given problems, into the practice of re-thinking what the problems might be in the first place. Even if one prioritizes learners’ eventual ability to solve problems, Stephen [Brown \(2001\)](#) points out that the skills of problem solving and mathematical concepts and techniques that would be applied to solve problems might more readily be learned and assimilated through activities other than solving problems. This insight is akin to a sports coach who notes that there may be valuable ways to develop and refine skills and concepts other than playing the sport over and over again – exercises, isolating body awareness and so on. In the context of mathematics and STEM, problems might be compared, changed, critiqued, and so on, instead of trying to solve them. Each of these alternatives would help learners to better approach problems in the future. Certainly, the experience of critiquing given problems, seeking to understand who posed them and why, and altering problem statements is an issue of ethics: learners should not be acculturated into assuming that only others pose problems, and instead should be mentored to see themselves as contributing a critical stance on problems as citizen mathematicians and scientists.

Problem posing itself can begin with given problems. One then asks, “what if not?” about any attribute or attributes of the initial problem. Once one has done this a few times, a mathematician can see the original problem as only one very special case of a family of limitless variations on a collection of elements and relationships. To take a somewhat simple example, suppose the initial problem asks about a family with 3 adults, 1 child, and 1 1-month-old baby boarding a public transportation system. An initial question might ask how much the ride will cost, given prices of tickets. A first what-if-not transformation could ask how the total cost

would change if the ticket prices changed to other amounts; a next step in the process could ask about further changes in ticket prices. Further posing changes might ask, “why do cities charge at all for public transportation ... might the costs be different if the service were “free to ride,” reducing use of roads, bridges, and other forms of transportation? Or: why does this family need to travel all together ... where are they going, and how would their costs be different if public policies did not require them to appear in person at the government agency? Such an approach drastically alters the initial problem statement, challenges the view that STEM solves problems by making it clear that STEM can also critique social assumptions, and more.

Different kinds of models for different sorts of purposes

Independent of the approach to working as a mathematician, mathematicians often vary in the purposes of their models and representations. When STEM is guided primarily by science processes, then the models are typically used to create a representation of the relevant relationships in a system. The model represents objects, phenomena, and physical processes in a consistent and logical way, and enables the scientist to develop scientific understanding by seeing how a change in one aspect of the model would influence changes in other components. Historically important models include the billiard ball model of a gas, the Bohr model of the atom, the Lotka-Volterra model of predator-prey interaction, the Mundell-Fleming model of an open economy. Engineering models tend to be created so that the engineer can assess the advantages and disadvantages of a design architecture, using indices or criteria to compare one model with another in terms of desired outcome measures. Engineers need to consider formal versus informal models, physical versus abstract models, descriptive models, analytic models, and so on. Different kinds of models are more or less appropriate at different stages of a design process; often, the models are employed to explain to a client or audience why they should support or fund the engineer’s proposed plan of action. Mathematicians might use models for similar purposes, as they build a subfield of knowledge or attempt to convince a skeptic of the logic of a conjecture.

Yet, there is a third use of models less tied to science and engineering design, in which a mathematician creates a model for problem-posing purposes, that is, to generate questions rather than answers or solutions. What can we learn about fractions if we use an area model instead of a linear model? What can we learn about triangles if we place them on a globe instead of on a flat two-dimensional plane? How do our questions about circles and ellipses change if we look at them algebraically instead of geometrically? Learners who are lucky enough to have experiences with these kinds of activities not only develop skills of solving problems with and through the use of models, but furthermore develop a sensibility that appreciates the ways that a particular model can limit or hide alternative realities and possibly more important questions. [Brown \(2001, p. 2\)](#) asks,

Have you thought ... of designing car bumpers that make use of liquid or that are magnetized, or shaped like a football? The creation of air bags suggests yet another option ... that bumpers might be capable of inflation on impact. Or, have you thought of the possibility that they may be made of glass, the fragility of which might discourage people from driving recklessly or relying so heavily on the use of the automobile?

He further stresses a feature of “coming to know” that such playful and initially less practical uses of models can encourage: in order to move from learning as being guided by an authority toward understanding oneself as a creator of knowledge, it is important to operate upon and even modify the things we are trying to understand. Ironically, it is often only in seeing a thing as something else that we can come to appreciate and understand it. Considering how one’s choices of models can affect the questions one poses as well as being more or less accessible to an audience with whom one is attempting to communicate has yet further implications: the perspectives of different groups in society are often neglected by the inclusion or loss of information in a model.

Social justice, mathematics and STEM

Not only is mathematics often blamed for discouraging people from pursuing STEM studies and careers, it is also then held responsible for disparities in STEM interest and STEM literacy cutting across racial, gender and socioeconomic lines. It is a consistent narrative that the vast majority of students being filtered out of mathematics, and subsequently, out of STEM – i.e., those whose experiences with school mathematics have led to disengagement as they learn to believe that they “can’t do math” – are more often underprivileged, underserved, and underrepresented ([Cherkowski, 2019](#)). While it is mostly the methods and biases of typical mathematics classrooms that are at fault more than the subject matter itself, mathematics will continue to be charged with the perpetuation of social and economic stratification, divisiveness, and cultural inequity unless practices are changed. Mathematics – and, if the gatekeeping function for STEM continues to be associated with the school subject, then STEM as well, will remain linked to social justice inequities in vocational and educational opportunities, earning potential over a lifetime, emotional and social well-being, volunteerism, and job satisfaction ([OCS, 2014](#); [Committee, 2018](#); [Cherkowski, 2019](#); [Block, 2020](#)). Furthermore, cultural, ethnic and gender diversity is linked to higher innovation outcomes, a key driver of STEM progress; research further reports that ethnically diverse workforces financially outperform non-diverse workforces by 35%, and a gender diverse workforce yields a 15% increase in financial output ([Desvaux et al., 2007](#); [Block, 2020](#)). In this way, STEM, and particularly mathematics for STEM, is a social justice issue in two ways: first, in terms of how mathematics and STEM education and apprenticeship practices support increased participation, and in particular, to find creative ways to remove any barriers to opportunity; second, policy

and practice must consciously avoid “blaming” members of disadvantaged, underserved and underrepresented groups for not adequately participating in STEM programs or pursuing STEM careers – the latter would be a dangerous and reprehensible treatment of human beings as resources to be mined for economic and intellectual progress. As Block (2020, p.27) notes,

Black students’ disproportionate interest in social justice and the absence of Black STEM majors are causally related. In their courses and jobs, most STEM faculty and employers do not make social change a focus. And for many Black students, that’s a serious problem.

Black students in particular often seek university majors in disciplines associated with public service despite dim salary prospects of careers with explicit social justice orientations (Block, 2020). Yet, studies of equally prepared Black, Latino and white students from similar financial backgrounds who started out their university studies in STEM fields found that Black and Latino students were 14% points more likely to leave STEM programs, a pattern not replicated in non-STEM fields of study (Block, 2020). Although many researchers point to the lack of diversity in STEM faculty, more powerful seems to be the STEM community itself: surveys suggest that STEM programs and early career opportunities overwhelmingly focus on technical expertise and the solution of given problems. “The people who are in most STEM programs and run most STEM organizations are not from those same backgrounds and do not have those same kinds of concerns,” leading many Black students to feel lost or unwelcome, reported one interviewee in Daniel Block’s (2020) article.

Given the plethora of socially conscious and diverse STEM crises and problems needing solutions in our world, the potential for STEM productivity in such areas to reap good profits should be apparent to even the most money-focused and technique-driven STEM professionals. Opportunities for transportation not reliant on fossil fuels, new sources of energy, methods of contending with climate change, health issues related to climate, weather, and other contexts, new forms of medicine, skin care and cosmetics, ...the possibilities seem endless. Related to the ways of working, models, and the ethics of problem solving and posing, it is an important social justice issue to make sure that STEM experiences include those characteristics of mathematical practices that emphasize multiple perspectives on problems, challenging assumptions that guide work, and directly understanding the social justice implications of the STEM projects themselves. In this sense, it is critical to include forms of problem posing and modeling that critique given problems in certain ways, including:

- Who was not included in the framing of the question?
- Does this need remain a priority when compared with other possible foci of our energy, given the competing needs of a diverse community?
- What questions are generated by using an alternative model, or an alternative way of working, shifting among science processes, engineering design steps, and ways of working as a mathematician?

And,

- What are the social justice implications of this project and its potential solutions?
- What social justice questions should we be working on in this program?

Although many issues of global and local importance such as climate and associated weather change, mass migration and refugee flow, homelessness, starvation and more appear to require critical perspectives across multiple disciplines, and overlap with STEM opportunities, it is often proposed that critical mathematics education can introduce specifically mathematics problems that establish the STEM and social justice linkages (Garibay, 2018; Nicol et al., 2019). Attempts to make this work find that it is the creative dialogue across discipline experts that can support the inclusion of cultural perspectives, diverse experiences and expertise, highlighting the need for those who would want to pursue this sort of work to understand that there are more options for integration than those that maintain disciplinary boundaries (Ortiz-Revilla et al., 2020).

Cross-cutting concepts and mathematizing

Science and engineering practices provide actions by which students gather data and information, whereas cross-cutting concepts provide techniques for students to process and think about the data and information (NRC, 2012; Cooper et al., 2015). Some common cross-cutting concepts rely on mathematical thinking and action to link the different domains of science and engineering. Patterns, scale, proportion, quantity, systems and system models, structure and function, and stability and change each have this characteristic; energy and matter as a cross-cutting concept would seemingly work in the opposing direction, that is, primarily using mathematics to support a science and engineering concept. Participation in such practices helps learners to form a holistic understanding of the cross-cutting concepts and the ways that they take on specific characteristics when viewed within one discipline or another. In theory, the concepts would make students’ developing knowledge more meaningful and would embed the practices relying upon the concepts more deeply into their worldview. The M in STEM offers a rich and powerful set of additional cross-cutting concepts and practices through mathematizing: “the process [es] by which members of a distinct cultural group think and act mathematically in order to translate or interpret their own surroundings by applying mathematical knowledge, most notably universal abilities that all cultures and peoples use: that

of counting, measuring, classifying, patterning, gaming, quantifying, and modeling” (Orey and Rosa, 2021). When we pull these processes out of the science and engineering practices, to analyze in their own right, and to reflect on the choices for how to accomplish them, what they leave out, when and how they influence our questions and conclusions, and so on, we can bring them back into the STEM activities with a significant understanding of both their power and limitations.

The field of critical mathematics education has a history of exploring such power and limitations that can benefit STEM programs if the wisdom can be more fully integrated (Ernest et al., 2016; Stathopoulou and Appelbaum, 2016; Nicol et al., 2019). The important point to be gained from this perspective on mathematics is that, while traditionally assumed to be the most neutral of all disciplines, and the furthest removed from politics, culture and ideology, a more critical stance clarifies the ways in which mathematics can and has structured and filtered social and cultural conceptions of truth, fact, knowledge, forms of argument, and more. One example of the ideological and political nature of mathematics is the ways that the discipline has provided knowledge and methods for colonial powers to create hierarchical systems and institutions. Even further, the last century of ethnomathematics has established the role of mathematics and mathematical problem solving in preserving and perpetuating inequities and injustices of colonialism well after official geopolitical status is no longer present; this situation is known as “coloniality,” and is often associated with the erasure of indigenous forms of mathematizing found outside of Western historical understanding, within colonialism and coloniality itself (Stathopoulou and Appelbaum, 2016; Appelbaum and Stathopoulou, 2020). Once the ideological and cultural constraints and formatting power of mathematics is recognized, next steps include action within mathematics and mathematics education to move toward a broader sense of responsibility and social justice (Appelbaum and Stathopoulou, 2020). One can also carry out similar ethnoscience and ethnoengineering projects, as well as ethno-STEM pursuits. Any attempt to presume STEM has a neutral or mostly positive role in solving problems ignores the hidden colonialism and loss of alternative epistemologies that STEM carries with it when left unchecked.

We might say that world civilization, far beyond STEM, is anchored in mathematics as its dorsal spine (D’Ambrosio, 2016). Yet, mathematicians often claim that their subtlest thinking is beyond comprehension of those seated around the table attempting to solve the most pressing issues of survival, social justice, and human dignity. This is a failure to tackle the challenges of packing and unpacking complexity in order to render it comprehensible with loss of relationships vital to the more complex patterns. They have further ignored the potential of forms of mathematizing that are not visible if only recognized when they match existing, Western forms of mathematizing, so that such potentially essential methods of thinking and being for solving questions of survival and human dignity would be available in our world. Extending this critical stance to STEM, we see the need for STEM practices and action to incorporate an analogous perspective that asks what coloniality has made invisible. This pursuit should begin with critical ethnomathematics, simply because the mathematics typically included as the M in STEM is so commonly treated as a neutral set of tools for solving problems articulated in terms of science or engineering. However, the next steps must take techniques and principles from critical ethnomathematics into the STEM work: understanding both mathematics and the STEM fields in general as inherently political, cultural, ideological, and historically situated – the more nuanced and essential lesson to be contributed by mathematics as a discipline, rather than as a collection of tools – is the opening to a far more exciting version of STEM, perhaps termed “critical-ethno-STEM.”

References

- AIG (Australian Industry Group), 2013. Lifting Our Science, Technology, Engineering, and Maths (STEM) Skills. AIG, Sydney.
- Anderson, J., English, L., Fitzallen, N., Symons, D., 2020. The contribution of mathematics education researchers to the current STEM education agenda. In: Way, J., Attard, C., Anderson, J., Bobis, J., McMaster, H., Cartwright, K. (Eds.), *Research in Mathematics Education in Australasia 2016–2019*. Springer Singapore, Singapore, pp. 27–57. https://doi.org/10.1007/978-981-15-4269-5_3.
- Appelbaum, P., 2008. *Embracing Mathematics: On Becoming a Teacher and Changing with Mathematics*. Routledge, NY.
- Appelbaum, P., 2012. Mathematical practice as sculpture of utopia: models, ignorance, and the emancipated spectator. *Learn. Math.* 32 (2), 14–19.
- Appelbaum, P., Clark, S., 2001. Science! Fun? A critical analysis of design/content/evaluation. *J. Curric. Stud.* 33 (5), 583–600.
- Appelbaum, P., Stathopoulou, C., 2020. The taking of Western/Euro mathematics as reappropriation/repair. *Revermop* 2, e202003. <https://doi.org/10.33532/revermop.e202003>.
- Beach, D., 1999. Alienation and fetish in science education. *Scand. J. Educ. Res.* 43 (2), 157–172.
- Bell, D., Wooff, D., McLain, M., Morrison-Love, D., 2017. Analysing design and technology as an educational construct: an investigation into its curriculum position and pedagogical identity. *Curric. J.* 28 (4), 539–558.
- Block, D., 2020. Why STEM needs to focus on social justice. *Wash. Mon.* 52 (9/10), 26–29.
- Boaler, J., 2002. *Experiencing School Mathematics: Traditional and Reform Approaches to Teaching and Their Impact on Student Learning*. Routledge, NY.
- Bourdeau, D.T., Wood, B.L., 2019. What is humanistic STEM and why do we need it? *J. Humanist. Math.* 9 (1), 205–216.
- Brown, S.I., 1973. Mathematics and humanistic themes: sum considerations. *Educ. Theor.* 23 (3), 191–214.
- Brown, S.I., 2001. *Reconstructing School Mathematics: Problems With Problems and the Real World*. Peter Lang, NY.
- Brown, S.I., Walter, M., 1983/2004. *The Art of Problem Posing*. Routledge, NY.
- Bybee, R.W., 2013. *The Case for STEM Education: Challenges and Opportunities*. National Science Teachers Association, Arlington, VA.
- Cedere, D., Birzina, R., Pigozne, T., Vasilevskaya, E., 2020. Perceptions of today's young generation about meaningful learning of STEM. *Probl. Educ. 21st Century* 78 (6), 920–932.
- Cherkowski, G., 2019. Why math and STEM education is a social justice issue. *Getting Smart* 6. June 3, 2019. <https://www.gettingsmart.com/2019/06/why-math-and-stem-education-is-a-social-justice-issue/>.
- Committee on STEM Education, National Science & Technology Council, the White House, 2018. *Charting a Course for Success. America's strategy for STEM education*, Washington, DC.
- Cooper, M., Caballero, M., Ebert-May, D., Fata-Hartley, C., Jardeleza, S., Krajcik, J., Lavery, J., Matz, R., Posey, L., Underwood, S., 2015. Challenge faculty to transform STEM learning: focus on core ideas, crosscutting concepts, and scientific practices. *Science* 350 (6258), 281–282.

- D'Ambrosio, U., 2016. Ethnomathematics: a response to the changing role of mathematics in society. In: Ernest, P., Sriraman, B., Ernest, N. (Eds.), *Critical Mathematics Education: Theory, Praxis, and Reality*. Information Age Publishing, Charlotte, NC, pp. 23–34.
- Davis, P., Hersh, R., 1981. *The Mathematical Experience*. Birkhäuser, Boston.
- Desvaux, G., Devillard-Hoellinger, S., Baumgarten, P., 2007. Women Matter: Gender Diversity, a Corporate Driver. mckinsey.com. McKinsey & Company.
- Emdin, C., 2010. Affiliation and alienation: hip-hop, rap, and urban science education. *J. Curric. Stud.* 42 (1), 1–25.
- Ernest, P., Sriraman, B., Ernest, N., 2016. *Critical Mathematics Education: Theory, Praxis, and Reality*. Information Age Publishing, Charlotte, NC.
- Fitzallen, N., 2015. STEM education: what does mathematics have to offer? In: Marshman, M. (Ed.), *Proceedings of the 38th Annual Conference of the Mathematics Education Research Group of Australasia*. MERGA, Sunshine Coast, Australia, pp. 237–244.
- Garibay, J.C., 2018. Beyond traditional measures of STEM success: long-term predictors of social agency and conducting research for social change. *Res. High. Educ.* 59, 349–381.
- Gitari, W., 2016. The application of school science by urban high school youth through problem solving and everyday life. *Sci. Educ. Int.* B27 (3), 344–368.
- Hoisington, C., 2020. A scientific age. *Educ. Leader* 78 (3), 58–64.
- Jarman, R., McClune, B., 2007. *Developing Scientific Literacy*. Open University Press, Maidenhead.
- Kinniburgh, H.L., Byrd, K., 2008. Ten black dots and September 11: integrating social studies and mathematics through children's literature. *Soc. Stud.* 99 (1), 33–36.
- Larson, M., 2017. Math Education Is STEM Education. Retrieved 12 July 2021. <https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Matt-Larson/Math-Education-Is-STEM-Education/>.
- Leung, S.L., Hargrove, B., Marsh, E., Gregg, A., Thole, K., 2020. Prompting by COVID-19 to rethink the purpose of engineering laboratory education – develop practical competence to solve real-world problems. *Adv. Eng. Educ.* 8 (4), 1–7.
- Levrini, O., Fantini, P., Barelli, E., Branchetti, L., Satanassi, S., Tasquier, G., 2021. The present shock and time re-appropriation in the pandemic era missed opportunities for science education. *Sci. Educ.* 30 (1), 1–31.
- Li, Y., Schoenfeld, A., 2019. Problematising teaching and learning mathematics as “given” in STEM education. *Int. J. STEM Educ.* 6, 44. <https://doi.org/10.1186/s40594-019-0197-9>.
- Mason, J., Burton, L., Stacey, K., 1982. *Thinking Mathematically*. Pearson, Harlow, England.
- (NRC) National Research Council, 2012. A framework for K-12 science education: practices, crosscutting concepts, and core ideas. In: Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. The National Academies Press, Washington, DC.
- Nicol, C., Bragg, L., Radzinski, V., Yaro, K., Chen, A., Amoah, E., 2019. Learning to teach the M in/for STEM for social justice. *ZDM* 51 (6), 1005–1016.
- OCS (Office of the Chief Scientist), 2014. Australia's Future: STEM Launches Stars into Orbit. Australian Government, Canberra.
- O'Neil, C., 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, NY.
- Orey, D., Rosa, M., 2021. Ethnomodelling as a globalization process of mathematical practices through cultural dynamism. *Math. Enthus.* 18 (3), 439–468.
- Ortiz-Revilla, J., Adúriz-Bravo, A., Greca, I., 2020. A framework for epistemological discussion on integrated STEM education. *Sci. Educ.* 29 (4), 857–880.
- Polya, G., 1945/2010. *How to Solve It*. Princeton University Press, Princeton, NJ.
- Rennie, L., Wallace, J., Venville, G., 2012. Exploring curriculum integration: why integrate? In: Rennie, I., Venville, G., Wallace, J. (Eds.), *Integrating Science, Technology, Engineering, and Mathematics: Issues, Reflections, and Ways Forward*. Routledge, NY, pp. 1–11.
- Romagnano, L., 1994. *Wrestling with Change: The Dilemmas of Teaching Real Mathematics*. Heinemann, Portsmouth, NH.
- Sahinkaya, N., Aladag, E., 2010. The views of class teachers on the integration between social studies and mathematics. *Procedia Soc. Behav. Sci.* 2 (2), 2876–2880.
- Sanders, M., 2009. STEM, STEM education, STEMmania. *Technol. Teach.* 68 (4), 20–26.
- Schmidt, W.H., Houang, R.T., 2007. Lack of focus in the mathematics curriculum: a symptom or a cause? In: Loveless, T. (Ed.), *Lessons Learned: What International Assessments Tell us About Math Achievement*. Brookings Institution Press, Washington, pp. 65–84.
- Sinclair, N., 2006. *Mathematics and Beauty: Aesthetics Approaches to Teaching Children*. Teachers College Press, NY.
- Stathopoulou, C., Appelbaum, P., 2016. Dignity, recognition and reconciliation: forgiveness, ethnomathematics and mathematics education. *RIPEM* 6 (1), 26–44, 2016.
- Stathopoulou, C., Kotarini, P., Appelbaum, P., 2015. Ethnomathematical research and drama in education techniques: developing a dialogue in a geometry class of 10th grade students. *Rev. Latinoam. Etnomatemática RLE* 8 (2), 105–135.
- Treacy, P., O'Donoghue, J., 2014. Authentic integration: a model for integrating mathematics and science in the classroom. *Int. J. Math. Educ. Sci. Technol.* 45 (5), 703–718.

Relevant Websites

- Building STEM Education on a Sound Mathematical Foundation – A joint position statement on STEM from the National Council of Supervisors of Mathematics and the National Council of Teachers of Mathematics. <https://www.nctm.org/Standards-and-Positions/Position-Statements/Building-STEM-Education-on-a-Sound-Mathematical-Foundation/>.
- CIEAEM, The Commission for the Study and Improvement of Mathematics Teaching. <https://www.cieaem.org/index.php/en/>.
- CIEAEM Manifesto: http://www.cieaem.info/images/Documents/CIEAEM_Manifesto/CIEAEM_Manifesto.pdf.
- European Commission STEM Education website: https://ec.europa.eu/education/eat-categories/stem_mk.
- ICMI, International Commission on Mathematical Instruction. <https://www.mathunion.org/icmi>.
- ICMI Pipeline Project: <https://www.mathunion.org/icmi/activities/pipeline-project>.
- ICMI Experiencing Mathematics Travelling Exhibition: <https://www.mathunion.org/icmi/activities/experiencing-mathematics-travelling-exhibition>.
- The Mind Research Institute “The M in STEM” website: <https://blog.mindresearch.org/blog/the-m-in-stem>.
- Siemens Stiftung STEM Education Initiative in Latin America. <https://www.siemens-stiftung.org/en/projects/stem-education-for-innovation/>.
- UNESCO Cátedra Regional de Mujer, Ciencia y Tecnología website: <https://www.catunescomujer.org/>.
- UNESCO STEM Education website: <https://en.unesco.org/stemed>.
- US Department of Education STEM website: <https://www.ed.gov/stem>.
- The Youth Mathematician Laureate Project: <http://yomap.org>.

Engineering education

Senay Purzer, Purdue University, West Lafayette, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	36
Engineering in pre-college education	36
An epistemology-based framework	37
Authenticity and inclusiveness in engineering curriculum	37
Design coaching and responsive teaching	38
A focus on engineering design reasoning	39
Conclusions	40
References	40

Introduction

The early years of the 21st century yielded a growing interest in engineering education research (Sheppard et al., 2009; The Steering Committee of the National Engineering Education Research Colloquies, 2006). Worldwide, many institutions established graduate programs, research centers, and global collaboration networks in engineering education (Clark, 2011; Gumaelius and Kolmos, 2019; Yasuhara, 2021). The National Academy of Engineering in the United States published two influential reports. *The Engineer of 2020: Visions of Engineering in the New Century* envisioned the future of engineering education followed by another report which explained how educators should prepare for that future (NAE, 2004, 2005). While the early reports focused on undergraduate and graduate education, pre-college education was also targeted with recommendations to support K-12 mathematics, science, and engineering education and the need to promote public understanding of engineering. In later years, engineering education in pre-college education continued to gain momentum. This article reviews these developments and provides frameworks and strategies necessary to guide pre-college engineering education.

Engineering in pre-college education

The importance of pre-college engineering is often positioned under the umbrella of science, technology, engineering, and mathematics (STEM) education. In particular, engineering design is typically described as an approach to engage students in real-world design problems that necessitate the integration of science, mathematics, and technology. Given this definition, the STEM movement lead to a variety of curricular models. For example, other subjects are added to the STEM acronym by including the arts (STEAM) or computational thinking with STEM + CT. In other cases, robotics and maker spaces emerged as signature activities of the STEM model.

This variety can be attributed to different factors. First, the interdisciplinary nature of STEM resulted in differences in teaching models as scholars from different disciplines approach integration differently. Such variation can be explained with a disciplinary perspective that can be attributed to the epistemological stances of the scholars. For example, how scientists use and integrate technology, mathematics, and engineering in their work differs from how engineers use and integrate technology, mathematics, and science in their work. Such variation is different than the variation attributed to pedagogical differences. Second, the STEM movement emerged with a diverse set of motivations ranging from helping students gain a deeper understanding of science and mathematics concepts to growing a more diverse engineering workforce to promoting creativity. Yet, the body of literature has shown that accomplishing these benefits has not proved to be easy (Valtorta and Berland, 2015). In addition, the novelty of the STEM education movement resulted in variation of interpretations and hence variations in quality. Scholars concerned over this variation developed frameworks to guide curriculum quality (Moore et al., 2015) and comprehensively define dimensions of engineering (AE3 and ASEE, 2020). Given these factors, the vast variation in STEM education is inevitable.

In our prior work, we argued that it is beneficial to integrate engineering into formal K-12 education, but the implicit argumentative stances of the educators, who design the educational experience, must be made clear (Purzer and Quintana-Cifuentes, 2019). The value attributed to engineering design can be pedagogical, epistemological, or methodological. The pedagogical argument treats engineering and design as vehicles for motivation to learn science and mathematics concepts. For example, a lesson where students are engaged in a series of thermal transport experiments, asked to design an insulated container, and assessed on their understanding of thermal concepts is built on a pedagogical perspective. The methodological perspective puts value in skills acquisition making design the vehicle for learning collaboration, creativity, and problem solving as well as habits of mind such as comfort with failure. For example, a lesson where students are asked to follow a design process using a limited amount of materials to build a strong structure, which is tested for strength but without explicit discussion of material strengths, and assessed on their design and communication skills, is built on a methodological perspective. The epistemological argument centers on the inherently interdisciplinary nature of engineering but also positions that the pedagogical decisions must be justified with a philosophical

grounding. For example, a lesson where students are re-designing their school building to improve its energy efficiency under specified economic and structural constraints, and expected to learn about thermal concepts, earth science concepts, as well as trade-off decisions involved when designing, is built on an epistemological perspective. Understanding the specific argumentative stances helps explain variations in curriculum.

While we highlight the importance of each of these argumentative stances, we also caution educators and researchers against a monolithic approach. We argue that the epistemological perspective can most effectively guide a variety of models given the multifaceted nature of engineering. The following sections provide guidance for shaping pre-college engineering with an epistemological perspective.

An epistemology-based framework

Engineering is inherently interdisciplinary. Hence, the pedagogical role imposed on engineering as the integrator of school subjects is justified. However, simply making this connection is not sufficient. Many curricular efforts fall short of a further examination of the engineering profession, such as the necessity for making design judgments in context.

Often engineering is defined by its practice, which is design. However, to understand engineering, we must first articulate its purpose and motives. Purzer et al. (2022) defines the purpose of engineering as “the creation, monitoring, management, and improvement of technological systems through the careful negotiation of risks and benefits” (p. 8). This definition recognizes the purpose of creating new technologies but also the critical role of engineers in maintaining the existing technologies. This definition also argues that engineers engage in different types of inquiries. At the core of the practices that guide these inquiries is communication and negotiation and the need to satisfy clients or stakeholders and convince them of the quality of the design decisions. This approach resulted in the Honeycomb of Engineering framework, designed based on the principle of engineering as a socially-constructed process and the design process as the flexible, iterative, and adoptable guide for engineering (Purzer et al., 2022). The Honeycomb of Engineering takes an epistemological perspective to outline an ontological classification of six types of engineering inquiries:

1. User-centered design
2. Design-build-test
3. Engineering science
4. Engineering optimization
5. Engineering analysis
6. Reverse engineering

The flexible nature of the design process is adapted to guide different inquiries by highlighting key cells of the honeycomb. In pursuing these inquiries, relevant elements of engineering design are used while maintaining the common core practice of negotiation of risks and benefits. With a honeycomb representation, illustrated in Fig. 1, this framework positions the negotiation of risks and benefits at the center of the design process and enables iteration through communication and reflection (Purzer, 2021).

While an epistemologically-justified framework helps to guide and delineate design projects, it also helps clarify a loosely used descriptor attributed to curriculum: authenticity. A commonly used attribute of the design challenges is their authenticity—meaning the degree to which they represent real-world design projects (Garrett et al., 2016; Strobel et al., 2013). In a prior study we provided comparisons of two curricula in ways they emphasize or de-emphasize different authenticity aspects (Schmidt et al., 2021). We argue that vague references to authenticity that deviate from disciplinary epistemic practices have risks such as: student motivation will be lost; the learning outcomes will be weakened; and inaccurate perceptions of engineering will develop. These risks run counter to the aforementioned elements that bring engineering to the forefront of educational reform. Hence, there is much we can lose if authenticity is not planned into the design of curricula.

Authenticity and inclusiveness in engineering curriculum

The epistemological perspective into engineering education brings with it the elements of authenticity. For example, context authenticity defines lessons and design projects that take place in contexts similar to those in real-world contexts (Strobel et al., 2013). However, these descriptions require clarification of the terms such as “real-life” and “authentic” which are commonly used in reference to engineering curricula. Having authentic features is important for a lesson to be engaging. Yet, haphazardly labeling curricula or engineering lessons as authentic can also result in misrepresentations. Strobel and colleagues provide clear definitions of authenticity by distinguishing context authenticity, task authenticity, impact authenticity, and personal authenticity.

Projects with context authenticity embed problems in realistic contexts while promoting interdisciplinary problem solving and collaboration. Projects with task authenticity focus on ensuring specific tasks within a lesson are domain-relevant and reflects tasks that are practiced by engineers. Context and task authenticity deal mainly with the alignment of curriculum elements with the epistemic practices.

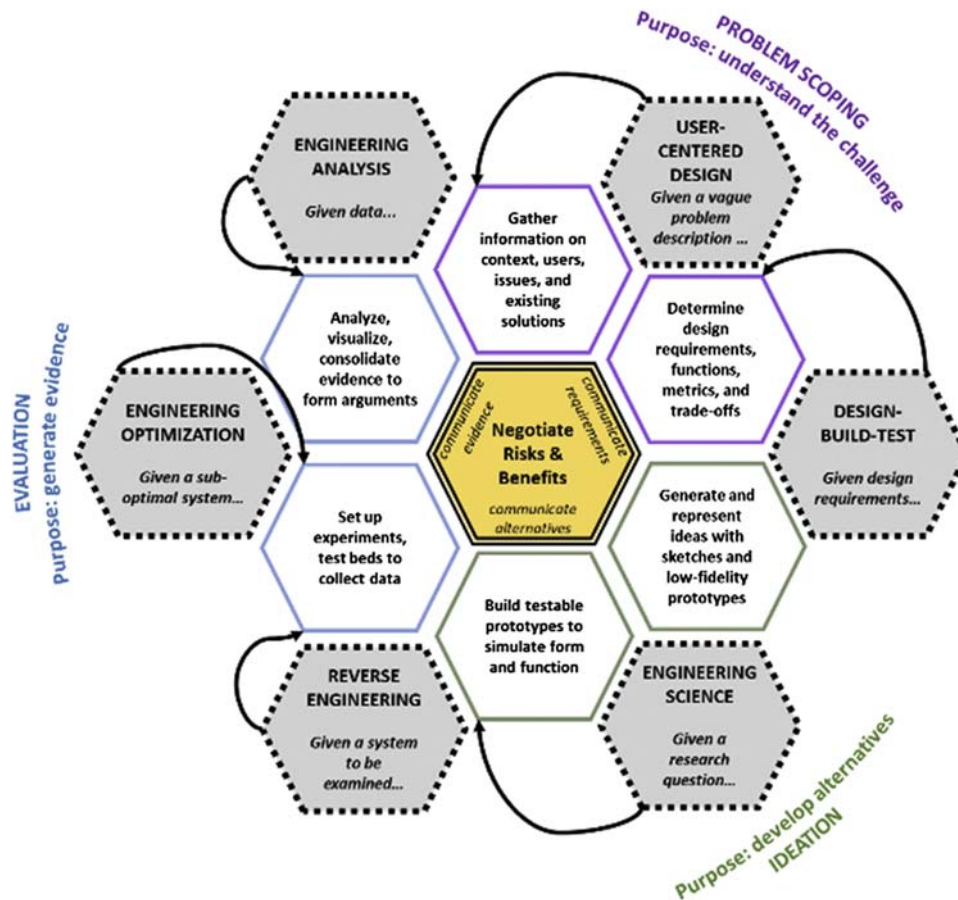


Fig. 1 The Honeycomb of Engineering (Purzer, 2021).

There are also types of authenticity that influence motivation to learn but also support inclusive education. For example, impact authenticity connects curricular elements with issues that expand beyond the classroom or school. These types of lessons can encourage decision-making and community-wide discourse around societal concerns or events. Personal authenticity, on the other hand, is designed around students and their daily life. Such lessons are designed with an understanding of students' interests, assets, and agency and aim to make explicit connections between school and aspects of students' everyday lives. Without personal authenticity, engineering lessons risk the loss of relevance for students. According to McGowan and Bell (2020) efforts to promote engineering in K-12 education continues to perpetuate stereotypical notions of engineering with old narratives that are not inclusive of women and individuals from underrepresented groups. They argue that engineering projects that connect to students' everyday experiences allow inclusive education and the opportunity for educators to build on to the students' assets.

Creating engineering curricula, especially with authenticity, requires decisions that are practical, epistemological, and pedagogical. For example, decisions must be made in terms of authenticity vs. accessibility, simplicity vs. disciplinary depth, etc. These decisions must be made explicit along with contextual constraints (such as school resources) and trade-offs that have to be made.

Design coaching and responsive teaching

Similar to how taking an epistemological perspective helps curriculum design and evaluation and ensures authenticity, such a perspective can also help frame teacher education. To establish this link, we can simply ask how design is taught in higher education. Adams et al. (2016) argue for the critical practice of coaching in design education. In their study of experienced design educators across multiple design disciplines, they identified four common coaching practices (Adams et al., 2016):

- eliciting student articulation of reasoning by directing students to share their design solutions and plans,
- helping students make conceptual connections and recognize fallacies in their design decisions,

- conveying explicit knowledge associated with design strategies and expert habits of mind, and
- encouraging students to figure out their own results by suggesting that they test the outcomes themselves and reflect on the results.

In design education, this type of teaching is called design coaching (Adams and Siddiqui, 2016). In educational research, design coaching can best be explained with the adaptive and responsive teaching model (Sherin et al., 2011; Johnson et al., 2017). Hence, through design coaching with responsive teaching, teachers can adopt strategies that facilitate student reasoning such as: noticing student reasoning, responding to student thinking, and helping student thinking evolve. Teaching engineering in such a manner that reflects authentic practices is undoubtedly a demanding task. Yet, effective teachers frequently and skillfully attend to aspects of student needs that are cognitive, emotional, and behavioral. Such attention then translates into instructional decisions such as responding to students' misconceptions, intervening to resolve social conflicts, and making revisions to classroom management (Dalvi and Wendell, 2017).

Teachers' competencies in design coaching would develop over time. To scaffold this development, educators can be provided with tools and strategies to notice and elicit student thinking, interpret student thinking, and respond to such thinking. Teachers can facilitate this through the Engineering Design Coaching Tool (Purzer et al., 2021). This instructional tool is designed to help educators elicit student reasoning through guided questioning. This questioning then helps students verbalize their thinking processes, share evidence in support of their decision, notice fallacies in their thinking and guide design decisions—hence stimulating student learning.

Another way to promote pre-college engineering education is by building on design practices in which educators already excel. According to Wendell, idea generation and feasibility analysis of potential solutions were strong assets of pre-service teachers when they engaged in engineering design (Wendell, 2014; McCormick et al., 2014; Moore et al., 2015). Conversely, these teachers were not as proficient in problem scoping. This finding helps explain results from our prior analysis with the finding that design-build-test projects were prominently common at all levels of education across primary and secondary school (Purzer et al., 2022). Design-build-test inquiries used in engineering practice emphasize ideation and feasibility analysis. Moreover, design-build-test problems entail clear design criteria and constraints and do not require in-depth problem scoping. In another study, Judson et al. (2016) analyzed middle school teachers' lesson plans in engineering and found data analysis and design comparisons were strengths of the teachers. This finding points to an opportunity suggesting that teachers may enjoy and excel in developing and teaching engineering analysis and optimization projects. Yet, these types of lessons were uncommon in our analysis, together making up only 3% of the lessons we analyzed. Hence, it is important that teacher education and professional development efforts introduce educators to the variety of inquiries used in engineering practice and ways these practices can be translated to education. In such, teachers are empowered by recognizing their assets and strengths and gradually build their competencies in coaching engineering design projects. These efforts would then translate into improved student learning and abilities in engineering design reasoning.

A focus on engineering design reasoning

Striving for responsive teaching in the context of authentic design projects requires a clear definition of what reasoning entails for the discipline of engineering. There are two important elements of engineering reasoning. One entails the ability to transition between concrete and abstract modes of thinking. The second entails the ability to think deeply within disciplinary ideas as well as broadly across multiple disciplines. At the intersection of these two spectrums of reasoning is the Quadrants of Design Reasoning with the following four elements: experiential observations, first-principles, trade-offs, and complex abstractions (See Fig. 2).

Experiential observations refer to reasoning at the concrete level with elements of procedural actions, trial and error processes, and everyday experiences. The *first-principles reasoning* refers to the applications of science and mathematics as well as principles from other disciplines such as economics, psychology, and arts. These foundational principles, which we label as first-principles, are indicators of theoretical support for design decisions. *Trade-offs* require crossing disciplinary boundaries to weigh and negotiate multiple design criteria. Reasoning with *complex abstractions* requires imagining how a solution might behave in other conditions beyond the current context or situation. Together the four modes of reasoning and the ability to fluently transition across them is essential when designing.

Some critiques of STEM education argue that the resources allocated to STEM education take away from liberal arts and civic education. However, if done properly, engineering design education would include disciplinary concepts such as ethics. Therefore, we have chosen the label, first-principles, as opposed to narrowly defining it as science and mathematics. Trade-offs are also important to consider when examining ethical issues. For example, safety is of vital importance in engineering, as well as values such as fairness and diversity. First-principles reasoning allows for deeper exploration of such disciplinary concepts. Trade-offs reasoning will also help student designers explore the intersection between the technical design requirements and the ethical design requirements. For example, designing a virus testing kit with high accuracy may limit its broad accessibility due to its high cost. Complex abstractions allow designers to imagine their products in the future so they can avoid unintended consequences. Hence, a focus on engineering reasoning in education would support learning across all school subjects and not just in STEM subjects.

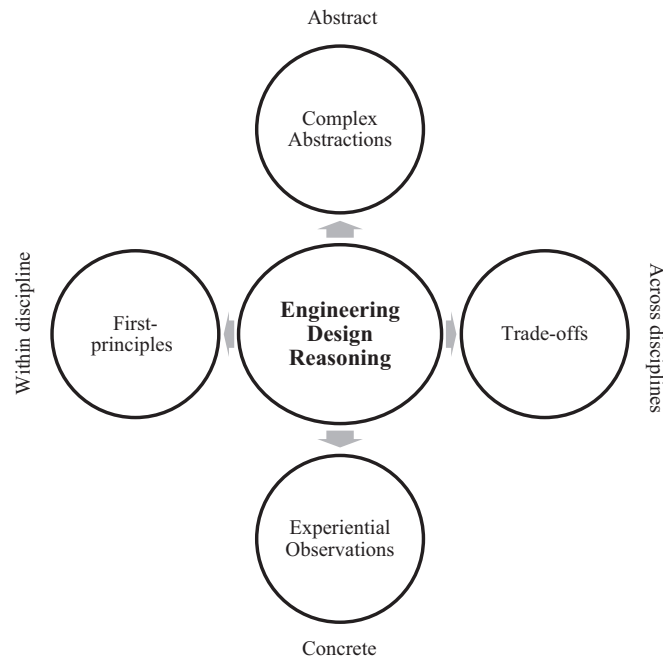


Fig. 2 Four dimensions of engineering design reasoning.

Conclusions

As education systems recognize the importance of teaching engineering in pre-college education, it is important to articulate the practices and philosophical underpinnings of the discipline of engineering. Studies of the philosophy of engineering help provide guidance. This article outlines six engineering inquiries: design-build-test, user-centered design, engineering science, engineering optimization, engineering analysis, and reverse engineering. With an understanding of different inquiries informed by the engineering philosophy, we can design curricula that enable authentic projects, engaging learning, and inclusive education.

Teaching engineering in ways that authentically reflect the engineering practice is aligned with the responsive teaching model which promotes the important practices of teaching such as guiding student ideas, responding to students' needs, recognizing creative ideas, and noticing misconceptions. Hence this article equates the role of the teacher with the role of a design coach and presents frameworks and strategies to guide pre-college engineering education. Teaching engineering in authentic ways also supports inclusive education. Design projects with impact and personal authenticity allows educators to build connections between engineering and students' everyday life and build on their assets and lived experiences. Such problem solving at the intersection of technical knowledge and lived experiences support student learning through authentic design. This article argues for a careful examination of pre-college curriculum and practices in engineering, so we can design curriculum that enable inclusive education.

Engineering education provides wide learning opportunities and will continue its momentum and expansion in pre-college education. For this movement to lead to improved student learning and engagement, authentic engineering curricula must be taught through responsive teaching and design coaching. These efforts would have the biggest impact on pre-college education when the learning outcomes goes beyond content knowledge in science, technology, engineering, and mathematics but include evidence-based and ethical decision-making as well as the ability to develop creative solutions to real-world problems.

References

- Adams, A., Forin, T., Chua, M., Radcliffe, D., 2016. Making visible the "how" and "what" of design teaching. In: *Analyzing Design Review Conversations*. Purdue University Press.
- Adams, R.S., Siddiqui, J.A. (Eds.), 2016. *Analyzing design review conversations*. Purdue University Press.
- Advancing Excellence in P-12 Engineering Education & the American Society for Engineering Education, 2020. Framework for P-12 Engineering Learning: A Defined and Cohesive Educational Foundation. <https://p12framework.asee.org>.
- Clark, R., 2011. Today's pupils, tomorrow's engineers! Pedagogy and policy: a UK perspective. *J. Eng. Des. Technol.* 9 (2), 227–241. <https://doi.org/10.1108/17260531111151087>.
- Dalvi, T., Wendell, K., 2017. Using student video cases to assess pre-service elementary teachers' engineering teaching responsiveness. *Res. Sci. Educ.* 47 (5), 1101–1125. <https://doi.org/10.1007/s11165-016-9547-5>.
- Garrett, L., Huang, L., Charleton, M.C., 2016. A framework for authenticity in the mathematics and statistics classroom. *Math. Educat.* 25 (1), 32–55. <https://tme.journals.libs.uga.edu/tme/article/view/2012/1917>.
- Gumaelius, L.B., Kolmos, A., 2019. The Nordic future engineer. In: *Proceedings of the Annual American Society for Engineering Education Conference & Exposition*, Tampa, Florida. <https://doi.org/10.18260/1-2-33411>.

- Johnson, A.W., Wendell, K.B., Watkins, J., 2017. Examining experienced teachers' noticing of and responses to students' engineering. *J. Pre-Coll. Eng. Educ. Res.* 7 (1), Article 2. <https://doi.org/10.7771/2157-9288.1162>.
- Judson, E., Erzen, J., Krause, S., Middleton, J.A., Culbertson, R.J., 2016. How engineering standards are interpreted and translated for middle school. *J. Pre-Coll. Eng. Educ. Res.* 6 (1), Article 1. <https://doi.org/10.7771/2157-9288.1121>.
- McCormick, M., Wendell, K., O'Connell, B., 2014. Student videos as a tool for elementary teacher development in teaching engineering: what do teachers notice. In: Proceedings of the Annual American Society for Engineering Education Conference & Exposition, Indianapolis, Indiana. <https://doi.org/10.18260/1-2-23060>.
- McGowan, V.C., Bell, P., 2020. Engineering education as the development of critical sociotechnical literacy. *Sci. Educ.* 29 (4), 981–1005. <https://doi.org/10.1007/s11191-020-00151-5>.
- Moore, T.J., Hynes, M.M., Purzer, S., Glancy, A.W., Siverling, E.A., Tank, K.M., et al., 2015. STEM integration: evidence of student learning in design-based curricula. In: Proceedings—Frontiers in Education Conference, FIE, vol. 2015. <https://doi.org/10.1109/FIE.2014.7044216>.
- National Academy of Engineering, 2004. *The Engineer of 2020: Visions of Engineering in the New Century*. National Academies Press, Washington, DC.
- National Academy of Engineering, 2005. *Educating the Engineer of 2020: Adapting Engineering Education to the New Century*. National Academies Press.
- Purzer, S., Quintana-Cifuentes, J.P., 2019. Integrating engineering in K-12 science education: spelling out the pedagogical, epistemological, and methodological arguments. *Discip. Interdiscip. Sci. Educ. Res.* 1 (1), 1–12. <https://doi.org/10.1186/s43031-019-0010-0>.
- Purzer, S., Quintana, J.P., Lingam, I., 2021. Engineering design coaching tool. *Sch. Eng. Educ. Open Educ. Resour.* <https://docs.lib.purdue.edu/eneoer/1>.
- Purzer, S., Quintana-Cifuentes, J.P., Menekse, M., 2022. The Honeycomb of Engineering framework: philosophy of engineering guiding pre-college engineering education. *J. Eng. Educ.* <https://doi.org/10.1002/jee.20441>.
- Purzer, S., 2021. The honeycomb of engineering. *Sch. Eng. Educ. Open Educ. Resour.* <https://docs.lib.purdue.edu/eneoer/2>.
- Schimpf, C., Purzer, S., Quintana-Cifuentes, J., Sereiviene, E., 2021. What does it mean to be authentic? Challenges and opportunities faced in creating K-12 design projects with multiple dimensions of authenticity. In: Sanzo, K., Scribner, J.P., Wheeler, J.A., Maxlow, K.W. (Eds.), *Design Thinking: Research, Innovation, and Implementation*. Information Age Publishing.
- Sheppard, S.D., Macatangay, K., Colby, A., Sullivan, W.M., 2009. *Educating Engineers: Designing for the Future of the Field*. San Francisco, CA. Retrieved from. <https://files.eric.ed.gov/fulltext/ED504076.pdf>.
- Sherin, M.G., Jacobs, V.R., Philipp, R.A., 2011. In: Sherin, M.G. (Ed.), *Mathematics teacher noticing*. Routledge, New York, NY.
- Strobel, J., Wang, J., Weber, N.R., Dyehouse, M., 2013. The role of authenticity in design-based learning environments: the case of engineering education. *Comput. Educ.* 64, 143–152. <https://doi.org/10.1016/j.compedu.2012.11.026>.
- The Steering Committee of the National Engineering Education Research Colloquies, 2006. The research agenda for the new discipline of engineering education (Special Report). *J. Eng. Educ.* 95 (4), 259–261. <https://doi-org.ezproxy.lib.purdue.edu/10.1002/j.2168-9830.2006.tb00900.x>.
- Valtorta, C.G., Berland, L.K., 2015. Math, science, and engineering integration in a high school engineering course: a qualitative study. *J. Pre Coll. Eng. Educ. Res.* 5 (1), 3. <https://doi.org/10.7771/2157-9288.1087>.
- Wendell, K.B., 2014. Design practices of preservice elementary teachers in an integrated engineering and literature experience. *J. Pre Coll. Eng. Educ.* 4 (2), 29–46. <https://doi.org/10.7771/2157-9288.1085>.
- Yasuhara, K., 2021. Engineering Education Community Resource. Retrieved from. <http://engineeringeducationlist.pbworks.com>.

The multimodal affordances of language in STEM education

Christine D. Tippet^a, Mark A. McDermott^b, and Todd M. Milford^c, ^aUniversity of Ottawa, Ottawa, ON, Canada; ^bUniversity of Iowa, Iowa City, IA, United States; and ^cUniversity of Victoria, Victoria, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	42
What is language and literacy?	42
Disciplinary literacy	43
STEM literacy	43
Science literacy	44
Technology literacy	44
Engineering literacy	45
Mathematics literacy (numeracy)	45
STEM literacy	46
What do we mean by multimodality?	46
Thinking about multimodality in STEM	46
Language-rich STEM education	48
A practical example of the affordances of language in STEM	48
Concluding remarks	48
References	49
Further reading	49

Introduction

Language has long been considered to play an essential role in science education and science literacy, with its epistemic affordances for both communicating and constructing knowledge and understanding. More recently, the role of language in science, technology, engineering, and mathematics (STEM) education and in STEM literacy development have become areas of interest for educators and researchers. STEM education has been defined by various stakeholders in different ways. For some educators, an activity focusing on any one of the four STEM disciplines would be considered a STEM learning experience; however, for most educators a true STEM learning experience involves some integration of multiple STEM disciplines. Proponents of this integrated definition range from requiring STEM education to involve a combination of any two of the disciplines, to requiring all four disciplines to be meaningfully and purposefully integrated in a single learning experience in order for true STEM education to be attained. Although debate over what constitutes the most appropriate definitions of STEM and STEM education is ongoing, our perspective is that STEM learning experiences require authentic integration between and among two or more STEM disciplines. In fact, we assert that a critical attribute of effective STEM education is disciplinary integration because of the epistemic similarities between an integrated view of STEM and the multimodal nature of the language used to communicate and construct meaning in STEM.

Here, we present our views on the multimodality of language as we explore various perspectives of literacy, including language and literacy; disciplinary literacy; individual science, technology, engineering, and mathematics literacies; and STEM literacy. We identify the multimodal affordances of language as a commonality across these perspectives and discuss three models of multimodality that can be used to consider the role of language in STEM. We describe the possibilities inherent in language-rich STEM learning experiences and propose that STEM education can be a springboard to enhanced development of language and literacy skills. Drawing on the affordances of multimodalities, learners' language and literacy skills can be developed while STEM concepts are explored, simultaneously strengthening both communication skills and STEM understanding. Language-rich integrated STEM can then be considered as an avenue for enhancing language and literacy, as well as aspects of science, technology, engineering, and mathematics literacies, rather than as a content area that draws time away from an already full curriculum.

What is language and literacy?

Literacy describes mastery of particular skills and abilities within a particular context. In the context of language, literacy once meant proficiency in reading and writing, the only two recognized strands of language. However, in response to sociopolitical changes, the conception of literacy broadened to incorporate the strands of listening and speaking. More recently, advances in communication technologies led to the complex notion of multiple literacies that encompassed static and dynamic images as well as words and sign systems. In response, the strands of viewing and representing became acknowledged as important aspects of language.

These six strands – reading, writing, listening, speaking, viewing, and representing – represent our current framework for understanding of language and literacy, although they are sometimes presented in different combinations; for example, as reading,

writing, oracy, and media literacy; or with reading and writing encompassing viewing and representing respectively. The strands of language can also be classified by function, as shown in Fig. 1. Reading, listening, and viewing can be thought of as interpretive strands, while writing, speaking, and representing can be thought of as constructive strands. These strands interact during both the construction and communication of knowledge.

The emphasis of the interpretive strands is on making sense of preestablished information, while the emphasis of the constructive strands is on using language to generate and clarify understanding and subsequently to communicate that understanding. An emphasis on inquiry based and constructivist approaches to teaching STEM means that the constructive strands are receiving growing attention. There has been a shift toward learning *with* language, rather than learning *from* language, as researchers and educators recognize the importance of all six strands – and their interactions – in communicative, constructive, and argumentative aspects of disciplinary literacy.

Disciplinary literacy

Disciplinary literacy includes using a range of representational modes (e.g., written and oral language, images, music, and gestures) to construct and communicate information, to synthesize ideas, and to formulate arguments (Moje, 2008). Just as science literacy encompasses specific ways of reading, writing, listening, speaking, viewing, and representing that are culturally mediated in the discourse of science, STEM literacy entails reading, writing, listening, speaking, viewing, and representing in ways that are particular to the discourse(s) of STEM. A useful conceptualization of disciplinary literacy differentiates among basic, intermediate, and disciplinary literacy skills and practices (Shanahan and Shanahan, 2008). Basic literacy practices are the highly generalizable practices that most students would be engaged in during the primary grades. Intermediate literacy practices, which are neither generic nor specific to a particular discipline but are more broadly useful depending upon context, are typically acquired during the intermediate grades. Disciplinary literacy practices, which tend to be highly sophisticated and not widely generalizable, reflect the discipline within which students are constructing meaning and are typically acquired during high school and university. These three clusters of literacy skills and practices are shown in Fig. 2.

In Fig. 2, each step in the progression of disciplinary literacy becomes smaller, reflecting a more limited generalizability of the practices, and the height of the steps reflects the level of specificity of the practices. Learning requires movement up and down the steps as new literacy practices are developed while other practices that have already been mastered are utilized. Also, mastery of a specific practice in one stage does not mean mastery of all practices in that stage. It is important to note that disciplinary-specific language is shaped by the specialized demands of communicating disciplinary concepts. For example, in science, the construal of scientific meaning has led to unique grammatical and textual features such as high lexical density, abstraction, and technicality, as well as the frequent use of visual representations. The presence of multiple visual representations and symbolic language means that the discourse of science is multimodal, and by association, we would argue, so is the discourse of STEM. An emphasis on the use of representations in STEM is a critical and more sophisticated aspect of disciplinary literacy that would emerge as a focus in the intermediate grades and continue into high school and university.

STEM literacy

As science education researchers, we can identify clear alignment between science and STEM, and between science literacy and STEM literacy. However, we acknowledge science is just one of the four traditional disciplines of STEM, along with technology,

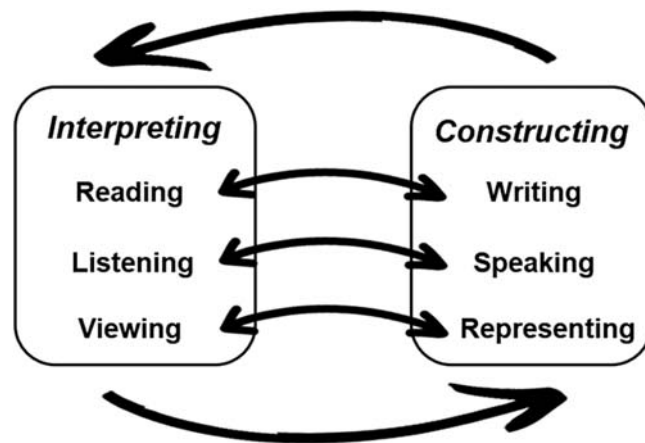


Fig. 1 The six interactive strands of language. Adapted from Tippet (2016, p. 727).

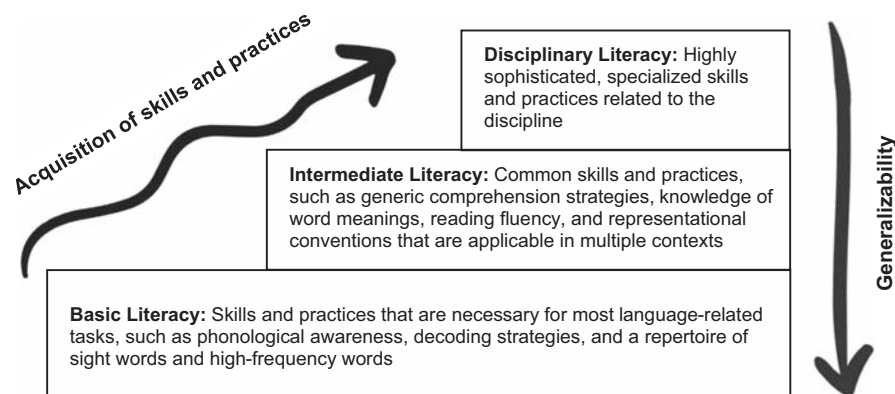


Fig. 2 Disciplinary literacy development. Adapted from [Shanahan and Shanahan \(2008\)](#).

engineering, and mathematics. We also recognize that integrated STEM experiences can provide broader and more general opportunities for the acquisition of science concepts and disciplinary skills, as well as for the development of technological and engineering literacy, numeracy, and language and literacy practices. Although there is some debate as to whether educators should be targeting an integrated STEM literacy or considering the literacies of the four individual S-T-E-M disciplines, the idea of a disciplinary literacy – particularly as proposed by [Shanahan and Shanahan \(2008\)](#) – is widely accepted ([Mohr-Schroeder et al., 2020](#); [Tang and Williams, 2019](#)). Situating STEM as a discipline in its own right, and building on the idea that the whole of STEM is greater than the sum of its parts, still calls for the simultaneous use of knowledge, skills, and practices from all four STEM disciplines. Our perspective is that even an integrated STEM literacy must consider the commonalities and differences of each of the four constituent disciplines. Here, we present an overview of conceptions of literacy within each discipline followed by a discussion of the commonalities that we, and other researchers, have identified.

Science literacy

Of the individual STEM disciplines, science has perhaps been the most studied, and current conceptions of science literacy are nuanced. [Norris and Phillips \(2003\)](#), in what may be the seminal work on the central role of language in science literacy, proposed the interacting fundamental and derived senses of science literacy, with language identified as a fundamental sense. Most subsequent conceptualizations of science literacy continue to acknowledge the critical role of language arts, with language conceptualized as a multimodal means of constructing and communication science knowledge. As an example, a recent metaphor for science literacy (see [Fig. 3](#)) reflects the growing need for a population that takes an informed role in civic actions. This metaphor situates language and literacy in the fundamental dimension of science literacy and illustrates that by engaging learners in active participatory learning with an authentic civics focus, disciplinary understanding can be enhanced by the effective use of language.

As suggested in [Fig. 3](#), true science literacy involves the integration of multiple skills, attributes, tools, and understandings. For students, and for that matter all citizens, science literacy requires an understanding of the philosophical nature of the scientific discipline and of the foundational conceptual big ideas of scientific knowledge. True science literacy would also involve an understanding of the processes and tools utilized to develop those big ideas. This conceptualization of scientific literacy obviously moves beyond rote memorization of discrete facts toward application to complex, real world problems.

Technology literacy

The discipline of technology is the least well-articulated of the four STEM disciplines, with ideas about technological literacy equally under-developed. This situation, described as a pervasive problem by [Ellis et al. \(2020\)](#), is being addressed, but misconceptions and misunderstandings about the meaning and role of technology are prevalent. There is also debate about whether technology should be considered as the T in STEM, or only in combination with engineering – part of the T/E in ST/EM. In our comparison of disciplinary literacies, we consider technology separately. Further, we take a perspective on technology and technological literacy that situates technology as tools; within STEM education, technology can be considered as tools with which to engage in authentic learning aimed at exploring STEM concepts. This perspective is a more encompassing vision of technology than is sometimes proposed because it does not focus exclusively on high-tech information and communication technologies but includes low-tech “tools” often more readily accessed by students of all ages. Consider how complex scientific concepts such as dynamic equilibrium or stoichiometric relationships can best be explored using simple tools like graduated cylinders, beakers, test tubes, and scales.

Technology does not exist in isolation. Its existence relies on systems of design, manufacture, marketing, and distribution, as well as training to operate particularly complex technologies. Technological literacy, at a minimum, includes knowledge and skills

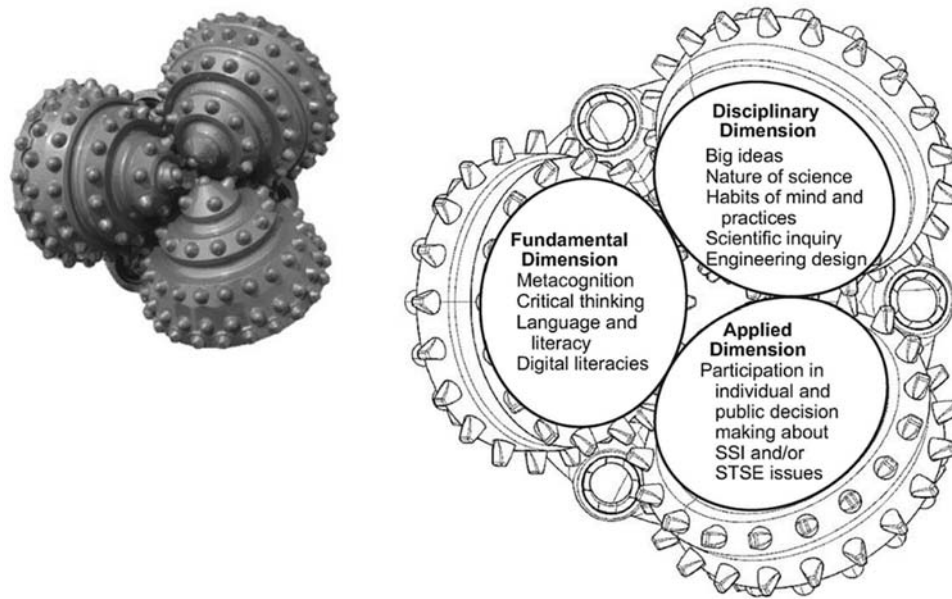


Fig. 3 A tricone bit metaphor for three dimensions of science literacy for personal and public civic action From [Tippett et al. \(2019, p. 325\)](#).

needed to work with a range of technology and to make informed decisions about the development and use of technology. Some definitions of technological literacy include communication skills, which we see as critical; consider creating or deciphering a user manual without possessing sophisticated literacy abilities! [Tang and Williams \(2019\)](#) described technological discourse as a multimodal form of language, wherein the technology community can communicate aspects of design and problem solving.

Engineering literacy

Although engineering is sometimes relegated to a complementary role with technology, we situate engineering as a separate, albeit highly interrelated, discipline with its own set of disciplinary practices. Engineering practices include identifying problems that need solving, designing solutions to problems, constructing models, applying science and mathematics concepts in problem solving, using technology in solving problems, testing and evaluating solutions, and communicating solutions. Engineering literacy can be conceptualized by articulating specific aspects of the knowledge, skills, and habits of mind that would be necessary for an engineering literate student. These specific aspects can be exemplified by the big ideas shown in [Fig. 4](#). Communication, explicitly identified as an engineering habit of mind, is also reflected in the engineering skills of engineering graphics, developing models, research and investigation, and technical writing.

Mathematics literacy (numeracy)

Mathematical literacy is the knowledge and skills needed to analyze and communicate about mathematical problems ([Organisation for Economic Cooperation and Development, 2019](#)). There are three highly interrelated components of mathematical literacy: content, competencies, and context. Fundamental mathematical content includes number sense, spatial abilities, graphing, and

Engineering Knowledge	Engineering Skills	Engineering Habits of Mind
✱ Engineering design ✱ Technology ✱ The Engineering Disciplines ✱ STEM Connections ✱ Engineering & Society ✱ Engineering Standards & Codes	✱ Designing Under Constraint ✱ Using Tools and Materials ✱ Engineering Graphics ✱ Developing Models ✱ Research & Investigation ✱ Technical Writing ✱ Mathematical Reasoning	✱ Systems Thinking ✱ Creativity ✱ Communication ✱ Optimism ✱ Optimization ✱ Iteration ✱ Collaboration ✱ Ethical Considerations ✱ Executive Functioning

Fig. 4 The big ideas in aspects of engineering literacy. Adapted from [Grubbs et al. \(2018, p. 35\)](#).

probability and statistics. Mathematical competencies include reasoning, modeling, and representing. Context refers to the idea that mathematical content and competencies are applied in a variety of real-world and authentic situations. Mathematical texts contain a unique set of symbols, drawings, proofs, and diagrams and students develop mathematical understanding through the interpretation of these specialized multimodal texts. Learners can also construct specific mathematical representations to communicate understanding and to build meaning. Therefore, a nuanced definition of mathematical literacy includes the ability to interpret, construct, and communicate with the vocabulary, symbols, and representations associated with the discourse of mathematics.

STEM literacy

Examination of different definitions of STEM literacy reveals two main perspectives: STEM literacy conceptualized as some combination of the individual disciplinary literacies – science literacy + technological literacy + engineering literacy + mathematics literacy – or STEM literacy as its own unique construct, wherein proponents typically take a fully integrated view of STEM. However, we are less concerned about possible distinctions between S-T-E-M and STEM literacy because our focus is on the commonalities across the disciplinary literacies. Consideration of the S-T-E-M literacies individually suggests that at a minimum STEM literacy would involve a combination of disciplinary literacies, each of which include, as outlined above, working with multimodal representations; understanding the conventions of specialized visual representations such as diagrams and drawings; and applying appropriate cognitive and metacognitive strategies (Tang and Williams, 2019).

The overarching commonality across disciplines is the ability to use the multimodal affordances of language to develop and to communicate understanding. Across all four disciplinary literacies, the language strands of reading, writing, speaking, listening, viewing, and representing have been identified as critical for constructing understanding and communicating that understanding in a variety of ways. Opportunities to work with discourse specific multimodal communications allow learners to most accurately portray their emerging understanding of concepts and can lead to cognitive engagement and disciplinary literacy. Students can strive to simultaneously make sense of the STEM concepts that they encounter while using the multimodal communication practices indicative of STEM discourse(s).

What do we mean by multimodality?

We focus on this idea of multimodality and how it might be conceptualized in order to expand our understanding of the possibilities inherent in language-rich STEM learning experiences. Modes can refer to sensory modes (e.g., sight, hearing, touch, smell, taste), or to specific ways of achieving complex communications (e.g., speech, static images, animated images, gestures, music, writing, three-dimensional models). Regardless, each mode has a range of affordances for communicating and constructing meaning; therefore, multimodal communication can harness the affordances of multiple senses, formats, and/or representations. Multiple representations, by definition, involve more than one representation, but those representations could all be in a single sensory mode (e.g., visual representations in the form of words, pictures, and symbols), or in multiple sensory modes (e.g., spoken words, printed images, and a hands-on demonstration). Multimodal approaches to STEM could include the use of multiple sensory modes (e.g., dramatic activities, a song containing STEM concepts, and reading a STEM information text) or the use of multiple presentation modes (e.g., verbal and nonverbal). In every STEM discipline, learners – as well as experts – demonstrate levels of representational fluency and conceptual understanding when they communicate using different modalities.

Thinking about multimodality in STEM

The idea of multimodal communication is well established in the individual STEM disciplines. Here we present three multimodal perspectives that have influenced our work in STEM education and in developing language-rich STEM learning experiences. Kozma and Russell's (2005) ideas about representational competence emerged from their work in chemistry, Tang's (2021) semiotic families are situated in science education, and the Lesh Translation Model (Glancy and Moore, 2013; Lesh and Doerr, 2003) was developed through studies of mathematics teaching and learning. Each of these three perspectives, although emerging from individual disciplines or subdisciplines, have broad applicability in STEM.

Our first perspective on multimodality is representational competence, which Kozma and Russell (2005) conceptualized as five levels of development in student decision making about representations. These levels are:

- *Representation as Depiction*, where iconic representations are based on physical features
- *Early Symbolic Skills*, where representations include physical features and some symbolic elements such as arrows to depict motion or time
- *Syntactic use of Formal Representations*, where representations are based on observed physical features and unobserved causes and the representational system focuses on syntax of use; representations are compared by surface features or syntactic rules rather than by underlying meaning
- *Semantic use of Formal Representations*, where representations are based on physical features and unobservable entities and processes, using a formal system of syntactic rules; representations are compared by shared meanings and can be transformed and used in problem solving

- *Reflective, Rhetorical use of Representations*, where multiple representations are based on physical features and unobservable entities and processes, using formal syntax and semantic rules; representations, or parts of representations, can be compared and transformed and used as rhetorical devices with appropriate selection and justification

Like disciplinary literacy, representational competence encompasses a range of levels, but movement through these levels is neither sequential nor permanent, instead influenced by context. While the lowest level of representational competence results in surface level depiction, the highest level emphasizes movement between and among multiple representations, harnessing the affordances of multimodality. With its focus on symbolic representations, the idea of representational competence aligns well with the language strands of viewing and representing, as well as with reading and writing.

Our second perspective on multimodality is the conceptualization of science discourse as consisting of semiotic families, with common families – modes – of meaning making including verbal-linguistic, visual-graphical, mathematical-symbolic, gestural-kinesthetic, and material-operational (Tang, 2021). The verbal-linguistic mode includes the reading, writing, speaking, and listening strands of language, regardless of language (e.g., English, French, Spanish). The visual-graphical mode incorporates representations such as diagrams and animations and includes the strands of viewing and representing and aligns with representational competence. The mathematical-symbolic mode includes mathematic and scientific notations based on numbers and symbols, which can be included in the strands of reading, writing, viewing, and representing. The gestural-kinesthetic mode consists of embodied movement and includes the strands of viewing and representing, with links to the verbal-linguistic mode. The material-operational mode captures hands-on processes such as practical work, demonstrations, and concrete model making and includes the viewing and representing strands of language as well as links to the verbal-linguistic and gestural-kinesthetic modes. Meaning making and communication involve movement between and within modes, as learners draw on the specific affordances of each semiotic family.

Our third perspective on multimodality, the Lesh Translation Model (e.g., Glancy and Moore, 2013; Lesh and Doerr, 2003), focuses on student meaning-making using common modes of mathematical models. Although there is a lack of consistency in published representations of the Lesh Translation Model, in every case, the critical aspect is the movement between and within modes such as visual representations, spoken symbols, and written symbols as shown in Fig. 5. The Lesh Translation Model clearly includes the six language strands as well as representational competence and shares similarities with the semiotic modes of science discourse. Although originating from the study of mathematical teaching and learning, the Lesh Translation Model has been adopted by researchers and educators in other disciplines, including integrated STEM.

Each of these three perspectives on multimodal communication differs according to context and focus; however, in each case, the suggestion is that deeper conceptual learning can occur when learners are able to move within and between different representations. This movement might be referred to as transformation (Kozmo and Russell, 2005) or translation (Glancy and Moore, 2013; Lesh and Doerr, 2003), but what is valued is working with multiple and varied representations as an opportunity to facilitate deeper learning. Each perspective showcases movement within and between multiple representations of multiple types. Deeper understanding of a science or STEM concept is enhanced by the ability to utilize different representational modes, to move within and between modes, and to connect modes. We argue that this translation, transformation, connection, or movement between and among representations is what makes the meaningful use of representations by students an epistemic tool in addition to a communication tool. Equally, the idea of movement is consistent with a perspective of integration in STEM teaching and learning serving as an epistemic tool. The commonality inherent in each perspective on multimodality and in integrated STEM is the emphasis on fluid

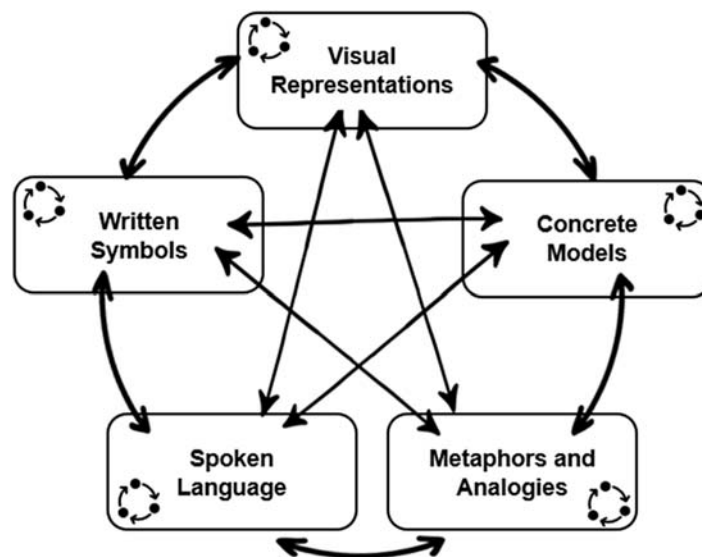


Fig. 5 The Lesh translation model. Adapted from Glancy and Moore (2013); Lesh and Doerr (2003).

movement between aspects of the construct: between multiple representational types, across semiotic families of meaning making, within and between modes, or across the STEM disciplines.

Language-rich STEM education

We have taken the position that STEM education should be integrated, because although non-integrated STEM learning experiences may still be potentially fruitful for learners, they would more appropriately be considered discipline specific. Integrated STEM learning experiences provide multiple access points for learners to think about, and construct understanding of, interrelated concepts. We have also identified language, and in particular its multimodal affordances, as a critical component of STEM education. Thus, an important goal during integrated STEM learning experiences is for students to meet curricular outcomes while learning how to use multimodal representations to create, communicate, and use knowledge within and across STEM. Advocates of integrated STEM argue that this approach is a more realistic model of what happens in the real world and can therefore more successfully prepare learners with the requisite critical thinking skills, communication abilities, entrepreneurial perspectives, and creativity to flourish in our rapidly changing society. The key benefit of an integrated approach to STEM education is the fact that the integration itself becomes an epistemic tool for students. In the act of working through multidisciplinary problems, by using and combining information and practices from STEM disciplines in an attempt to solve these problems, students are expanding their conceptual and procedural understanding of highly interrelated concepts, skills, and practices.

Instruction encouraging the effective use of multiple modes in writing-to-learn tasks is an area of study in science education. Multimodal writing-to-learn tasks have been shown to be effective instructional approaches for helping learners develop scientific understanding and these types of tasks are likely to be effective for teaching STEM. Learning benefits have been most evident when multiple modes are purposefully linked within communication tasks, regardless of context. Research suggests that students who can integrate multiple modes in science are likely to experience greater learning, in part because they are translating and transforming information across different modalities (McDermott and Hand, 2013). STEM learning environments afford numerous opportunities for translating across disciplines as well as for incorporating multiple modes to improve student understanding. A language-rich STEM learning environment allows learners to make connections between disciplines, modes, and strands of language.

A practical example of the affordances of language in STEM

A practical example of a language-rich STEM learning experience would involve multiple access points for learners to think deeply about related concepts. A STEM activity can be infused with specific multimodal tasks in order to develop an effective, integrated, language-rich learning experience. Here we describe a fairly common STEM activity – building a marble roller coaster – presented in a way that emphasizes language and its multimodal affordances. The basic materials and equipment for this activity are easily obtained: foam pipe insulation, marbles, paper cups, and masking tape. Optional materials could include markers, chart paper, graph paper, recording devices (e.g., smart phones), and tablets. As learners follow an engineering design process to plan and construct their marble roller coaster, there are multiple opportunities for educators to emphasize the roles and functions of language. For example, learners might document their plans using scaled drawings on graph paper; diagrams, lists, and sketches on paper; or mind maps in digital notepads and could document their construction processes using sketches, photographs, or videos.

During the learning experience, big ideas of science such as kinetic and potential energy, momentum, and friction emerge naturally. The technology used to design and construct the marble roller coaster can range from simple to complex, depending on the tools available and the age of the students; the end result itself can be considered a piece of technology. Aspects of mathematical literacy such as number sense, spatial abilities, and modeling and representing are central to the successful construction of a roller coaster. At any phase of the learning experience, learners utilize a variety of engineering knowledge, skills, and habits of mind, including designing under constraint, using tools and materials, creativity, and communication. Learners can communicate their growing understanding of STEM concepts by sharing a variety of representations with classmates and with their instructor. The multimodal affordances of language allow students to move between and among representations as they move between and among STEM disciplines to solve an engaging problem. In turn, the STEM learning experience provides an authentic context for developing language and literacy.

Concluding remarks

With a growing emphasis on STEM education, the role of language in STEM is emerging as a related area of interest. The multimodal nature of language is a commonality across individual science, technology, engineering, and mathematics literacies and STEM literacy. Each STEM discipline is a multimodal discourse with unique symbols and representations such as mathematical formulas, graphs, photographs, tables, animations, and diagrams, all of which are highly specialized representations and essential tools for conceptualizing STEM ideas.

STEM education can provide an authentic context in which to incorporate diverse language modes and functions. In fact, the epistemic affordances of the multimodality inherent in STEM learning environments can be leveraged to strengthen general literacy

achievement. Exploring the functions of language (e.g., interpretation and communication) and the relationship of language pairs of reading/writing, listening/speaking, and viewing/representing can provide opportunities for deeper learning.

An understanding of the affordances and limitations of different modes of communication is a crosscutting factor in science, technology, engineering, and mathematics literacies and is a key characteristic in STEM literacy. The affordances of multimodality can be more widely emphasized through purposeful STEM learning experiences such as the one described here. It is impossible to teach STEM without using language and a language-rich integrated STEM approach can help learners to simultaneously strengthen discipline specific content knowledge and literacy skills.

References

- Ellis, J., Wieselmann, J., Sivaraj, R., Roehrig, G., Dare, E., Ring-Whalen, E., 2020. Toward a productive definition of technology in science and STEM education. *Contemp. Issues Technol. Teach. Educ.* 20 (3), 472–496.
- Glancy, A.W., Moore, T.J., 2013. Theoretical Foundations for Effective STEM Learning Environments. School of Engineering Education. Working Papers, Paper 1. <http://docs.lib.purdue.edu/enevp/1>.
- Grubbs, M.E., Strimel, G.J., Huffman, T., 2018. Engineering education: a clear content base for standards. *Technol. Eng. Teac.* 77 (7), 32–38.
- Kozma, R., Russell, J., 2005. Students becoming chemists: developing representational competence. In: Gilbert, J.K. (Ed.), *Visualization in Science Education*. Springer, pp. 121–145.
- Lesh, R.A., Doerr, H.M., 2003. Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In: Lesh, R.A., Doerr, H.M. (Eds.), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning, and Teaching*. Lawrence Erlbaum Associates, pp. 3–33.
- McDermott, M., Hand, B., 2013. The impact of embedding multiple modes of representation within writing tasks on high school students' chemistry understanding. *Instr. Sci.* 41 (1), 217–246.
- Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Nickels, M., 2020. Moving toward an equity-based approach for STEM literacy. In: Johnson, C.C., Mohr-Schroeder, M.J., Moore, T.J., English, L.D. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 29–38. <https://www.routledgehandbooks.com/doi/10.4324/9780429021381-4>.
- Moje, E.B., 2008. Foregrounding the disciplines in secondary literacy teaching and learning: a call for change. *J. Adolesc. Adult Liter.* 52, 96–107.
- Norris, S.P., Phillips, L.M., 2003. How literacy in its fundamental sense is central to scientific literacy. *Sci. Educ.* 87 (2), 224–240. <https://doi.org/10.1002/sce.10066>.
- Organisation for Economic Cooperation and Development, 2019. PISA 2018 Assessment and Analytical Framework. Author. <https://doi.org/10.1787/b25efab8-en>.
- Shanahan, T., Shanahan, C., 2008. Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harv. Educ. Rev.* 78, 40–59.
- Tang, K.-S., Williams, J.P., 2019. STEM literacy or literacies? Examining the empirical basis of these constructs. *Rev. Educ.* 7 (3), 675–697. <https://doi.org/10.1002/rev3.3162>.
- Tang, K.-S., 2021. *Discourse Strategies for Science Teaching and Learning: Research and Practice*. Routledge.
- Tippett, C.D., Milford, T.M., Yore, L.D., 2019. Epilogue: the current context of Canadian science education and issues for further consideration. In: Tippett, C.D., Milford, T.M. (Eds.), *Science Education in Canada: Consistencies, Commonalities, and Distinctions*. Springer, pp. 311–337. <https://doi.org/10.1007/978-3-030-06191-3>.
- Tippett, C.D., 2016. What recent research on diagrams suggests about learning with rather than learning from visual representations in science. *Int. J. Sci. Educ.* 38 (5), 1–22. <https://doi.org/10.1080/09500693.2016.1158435>.

Further reading

- Alvermann, D.E., Wilson, A.A., 2011. Comprehension strategy instruction for multimodal texts in science. *Theory Into Pract.* 50, 116–124.
- Anderson, J., Li, Y., 2020. *Integrated Approaches to STEM Education: An International Perspective*. Springer. <https://doi.org/10.1007/978-3-030-52229-2>.
- Gillies, R.M., Baffour, B., 2017. The effects of teacher-introduced multimodal representations and discourse on students' task engagement and scientific language during cooperative, inquiry-based science. *Instr. Sci.* 45, 493–513. <https://doi.org/10.1007/s11251-017-9414-4>.
- Hand, B., McDermott, M., Prain, V., 2016. Using Multimodal Representations to Support Learning in the Science Classroom. Springer. <https://doi.org/10.1007/978-3-319-16450-2>.
- Hand, B., 2017. Exploring the role of writing in science: a 25-year journey. *Liter. Learn.* 25 (3), 16–23.
- Jewitt, C., 2008. Multimodality and literacy in school classrooms. *Rev. Res. Educ.* 32, 241–267.
- Li, Y., Schoenfeld, A.H., diSessa, A.A., Graesser, A.C., Benson, L.C., English, L., Duschl, R., 2019. On thinking and STEM education. *J. STEM Educ. Res.* 2 (1), 1–13. <https://doi.org/10.1007/s41979-019-00014-x>.
- Roehrig, G.H., Dare, E.A., Ring-Whalen, E., Wieselmann, J.R., 2021. Understanding coherence and integration in integrated STEM curriculum. *Int. J. STEM Educ.* 8 (2), 1–21. <https://doi.org/10.1186/s40594-020-00259-8>.
- Russell, T., 2014. Multimodal Representations and Science Learning. *Encyclopedia of Science Education*. https://doi.org/10.1007/978-94-007-6165-0_124-2.
- Walsh, M., 2010. Multimodal literacy: what does it mean for classroom practice? *Aust. J. Lang. Liter.* 33 (3), 211–239.
- Yore, L.D., Hand, B.M., 2010. Epilogue: plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Res. Sci. Educ.* 40 (1), 93–101. <https://doi.org/10.1007/s11165-009-9160-y>.

Using multiliteracies and multimodality to explore engineering literacies and problem-scoping for multilingual learners: A perspective on multiliteracies for STEM learning

Mary B. McVee^a, Kelly A. Schucker^a, Katarina N. Silvestri^b, and Jennifer N. Tripp^a, ^aUniversity at Buffalo, SUNY, Amherst, NY, United States; and ^bSUNY Cortland, Cortland, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Science, technology, engineering, and mathematics (STEM)	50
Multiliteracies	51
History and principles of multiliteracies	51
Multiliteracies and metalanguage to describe and interpret design elements and modes	52
Clarification of terms often used interchangeably or conflated with multiliteracies	52
New Literacy Studies (NLS)	52
Multimodality	54
new literacies	54
New Literacies	54
Multimodal literacies and engineering design	54
Context and participants	55
An exploration of multiliteracies in engineering literacies club (ELC)	55
Phase 1: Aameena and Emma begin to sketch bridge designs	56
Phase 2: Emma and Aameena shift their planning to include the template	57
Phase 3: Emma and Aameena discuss and plan which materials to use	59
Phase 4 : Aameena and Emma gather materials and organize guides	59
Discussion	59
Conclusion	60
Acknowledgement	61
References	61

The purpose of this article is to broadly consider the relationship between multiliteracies and Science, Technology, Engineering, and Mathematics (STEM) education. After providing a brief introductory overview on STEM and multiliteracies, we foreground one aspect of multiliteracies (i.e., multimodality) and one aspect of STEM (i.e., engineering). We first consider the meaning of the terms *STEM* and *multiliteracies*. We then provide an illustrative example that foregrounds multimodality and engineering design. The example derives from an afterschool Engineering Literacies Club (ELC) with children ages 8–10 learning about engineering, engineering design, and disciplinary literacies (Brock et al., 2014).

Science, technology, engineering, and mathematics (STEM)

The term “STEM” originated in the 1990s at the National Science Foundation to signify science, technology, engineering, and mathematics to efficiently identify and communicate the four disciplines (Bybee, 2013). While STEM recently became a popular buzzword, the impetus of STEM education in the United States originated during the Sputnik era when science and mathematics were foregrounded in educational reform efforts. Science education documents, such as *Science for All Americans* (AAAS, 1989), *Benchmarks for Science Literacy* (AAAS, 1993), and the *National Science Education Standards* (NRC, 1996) have called for incorporation of technology and engineering into science education. More recently, the National Research Council introduced a framework for science education that highlighted the centrality of teaching science, engineering, and technology for all:

The overarching goal of our framework for K-12 science education is to ensure that by the end of 12th grade, *all* students have some appreciation of the beauty and wonder of science; possess sufficient knowledge of science and engineering to engage in public discussions on related issues; are careful consumers to scientific and technological information related to their everyday lives; are able to continue to learn about science outside school; and have the skills to enter careers of their choice, including (but not limited to) careers in science, engineering, and technology.

NRC (2012), p. 1.

This framework informed the development of the *Next Generation Science Standards* (NGSS) where engineering is situated within three distinct dimensions of science learning: disciplinary core ideas, crosscutting concepts, and science and engineering

practices. In NGSS, crosscutting concepts and relationships and convergences with literacy and math practices demonstrate the interdisciplinary nature of science and STEM. That is, crosscutting concepts such as patterns, cause and effect, energy and matter, and structure and function cut across various disciplines and unify scientific thinking about the natural world. Practices such as constructing viable arguments from evidence and critiquing reasoning from others are shared among science and engineering, math, and English Language Arts.

Despite reform documents spanning decades and the increased global interest in STEM education, there often remains little clarity and consensus as to what STEM means or what STEM education entails (Bybee, 2013). In reviewing STEM education literature, Martín-Páez and colleagues (2019) uncovered multiple interpretations of STEM education, including varying perspectives toward disciplinary integration, such as disciplinary, multidisciplinary, interdisciplinary, and transdisciplinary approaches. While they suggest that “engineering could be a discipline with a high integrating value” (Martín-Páez et al., 2019, p. 816), others have disapproved of the dominance of one discipline (Honey et al., 2014) and have instead proposed the integration of two or more disciplines (Kelley and Knowles, 2016). Certainly, from policymakers, to researchers, administrators, teachers, and the public more broadly, STEM education is conceptualized in a multitude of ways, and integration among the disciplines is still in its infancy (Dare et al., 2019).

When applied in STEM educational contexts to design, implement, and assess various initiatives, programs, and practices, this ambiguity poses genuine concerns. Given the diverse and extensive nature of work related to defining, integrating, and researching STEM disciplines, “STEM” as a domain far exceeds the parameters of what can fully be explored in this entry. For this reason, we choose to focus on only one aspect of STEM—engineering. In so doing, we recognize that engineering itself is a complex discipline which integrates science, math, and technology but is also a domain that has distinct differences and epistemic practices (Cunningham and Kelly, 2017).

In this chapter, we will focus on multiliteracies, particularly multimodality, from a pre-college engineering (i.e., K-12) education perspective. Before introducing examples of engineering, a definition and brief history of multiliteracies is likely to be helpful. To further assist readers, we also consider common terms that are often conflated with, or used interchangeably with multiliteracies.

Multiliteracies

History and principles of multiliteracies

The term “multiliteracies” originates from an article in *Harvard Education Review* by the New London Group (NLG, 1996), later reprinted in *Multiliteracies: Literacy Learning and the Design of Social Futures*, edited by Cope and Kalantzis (2000). NLG scholars have continued to elaborate and expand the framework of multiliteracies as it was conceived in 1996. Their research includes multimodality (Kress, 2010), critical discourse analysis (Fairclough, 2010; Gee, 2014), critical literacies (Luke, 2012), curriculum and learning design, (Cope and Kalantzis, 2015), multiliteracies and Indigenous perspectives (Nakata, 2001), and classroom discourse (Cazden, 2001). The NLG constituted 10 scholars who sought to expand understandings of literacy and literacy pedagogy to more equitably account for culturally and linguistically diverse societies (1996, p. 64). Although scholars’ research foci varied (see Table 1), they convened with a shared goal of reframing how educators and researchers approach literacy teaching and learning. They agreed “the fundamental problem” was “that the disparities in educational outcomes did not seem to be improving” (1996, p. 63). They collectively outlined an agenda for what they referred to as a “pedagogy of multiliteracies” (1996, p. 60; Cope and Kalantzis, 2009). This perspective was grounded equity and social justice, and confidence that new media might redefine exclusionary practices “from and by the institutions of print literacy” (García et al., 2018, p. 73).

Framing traditional literacy pedagogy as “a carefully restricted project,” the NLG posited that privileging sanctioned literacies dismissed the realities of students’ lived social worlds, and the realized realities of future occupations in a rapidly globalizing world. Developing multiliteracies was necessary because “dramatic changes in everyday life in the realms of work, citizenship, and identity” rendered traditional literacy pedagogy practices “anachronistic” (Cope and Kalantzis, 2015, p. 3). The NLG proposed that

Table 1 New London Group scholars described in Cope and Kalantzis (2000).

Scholar	Location	Research focus
Courtney Cazden	United States	Classroom discourse, literacy pedagogy
Bill Cope	Australia	Cultural diversity in education literacy pedagogy
Norman Fairclough	Great Britain	Language changes represent of sociocultural changes
James Gee	United States	Discourse and psycholinguistics in capitalist jobs
Mary Kalantzis	Australia	Social education, literacy, citizenship education
Gunther Kress	Great Britain	Language, semiotics, visual and multimodal literacies
Allen Luke	Australia	Sociological analysis of teaching reading and writing
Carmen Luke	Australia	Feminist pedagogy
Sarah Michaels	United States	Learning in diverse urban classrooms
Martin Nakata	Australia	Literacy in indigenous communities

developing a multiliteracies approach was imperative in a diverse and globalized world, especially to address the gap in educational access and equity among diverse students and to “ensure that differences of culture, language, and gender are not barriers to educational success” (NLG, 1996, p. 61). Responding to this challenge, NLG prioritized “creating the learning conditions for full social participation,” where “the issue of differences becomes critically important” (1996, p. 61). NLG’s primary concerns included “pedagogical tension” between teaching methods, response to increasing diversity, the development of new modes and technologies, and “changing text usage in restructured workplaces” (1996, p. 62).

In 2022, these are even greater challenges. In the United States, for example, classrooms are increasingly linguistically, racially, and ethnically diverse, and repeatedly concerns are raised about test scores. For example, students in high-poverty schools are more likely to have lower math and reading scores across all racial and ethnic groups (National Center for Education Statistics, 2019). Persistently, such standardized measures have presented deficit-oriented portraits of students of color and multilingual learners as these assessments are not valid indicators of students’ intellect, holistic knowledge, or potential and have thus contributed to furthering racial inequality in education (Au, 2016; Hagopian, 2014; Menken, 2008; Milner, 2012; NASEM, 2018). Less scholarship has considered how inequitable opportunity structures and opportunities for learning contribute to the systemic inequities undergirding student performance. From its genesis, the multiliteracies framework sought to address inequities and consider how to transform pedagogy and schools to create opportunities for all learners (Cope et al., 2018). This focus on equity is foundational to multiliteracies perspectives.

Multiliteracies and metalanguage to describe and interpret design elements and modes

Prompted by global shifts in communication, the NLG (1996) considered literacy and learning in relation to multiple dimensions pertaining to multiliteracies, multimodality, and multilingualism. The term *literacies* plural provides reference not only to skills-based decoding of words or print, but also social meanings, ideologies, and emerging forms of literacy practices. From the outset, multiliteracies were linked with active design and transformation of teaching, curriculum, and educational contexts. Multiliteracies included multiple modes for meaning-making while recognizing the continued need for overt instruction. Underlying concerns about shifts in communication, labor, and society were considerations of equity in schooling and how futures are shaped by schooling. Ideas pertaining to design and pedagogy also included *situated practice*, *overt instruction*, *critical framing*, and *transformed practice* (NLG, 1996, p. 88). Some design elements were expressed in a well-known figure: “Multiliteracies: Metalanguages to Describe and Interpret Different Design Elements of Different Modes of Meaning” (see Fig. 1). Scholars have continued to define particular aspects of the figure such as mode and multimodality (e.g., Kress, 2009, 2010; Kress et al., 2001). Likewise, others have continued to refine additional key ideas such as situated practice, overt instruction, critical framing, and transformed practice as these pertain to multiliteracies (see Cope and Kalantzis, 2009, 2015; Kalantzis et al., 2005). Scholars continue to explore multiliteracies as evolving and contemporary (Serafini and Gee, 2017) and in contexts of diversity (Boyd and Brock, 2015), critical media literacy (Mirra et al., 2018), technologies (Lewis Ellison & Kirkland, 2014), transnationalism (Skerrett, 2015), and other topics.

The original NLG “call to understand knowledge and pedagogy as multimodal was radical” (Jewitt, 2009, p. 19) although in the past decade, many disciplines have turned toward multimodality. In traditional schooling, teaching and learning and academic curricula have centered primarily, and often exclusively, on linguistic design of meaning-making (i.e., reading words, writing words, speaking words). In contrast, the NLG proposed that audio, spatial, gestural, and visual design should also be considered in both teaching and learning. While not fully elaborated, multimodality was recognized as the embodied processes of literacy learning and teaching (McVee et al., 2017a). As such, literacy was not viewed as limited to skills of reading words or writing words but was situated in cultural contexts and enacted through embodied practices. The NLG is a keystone in many explorations of multimodality because of the ways that the original NLG manifesto (1996) pressed literacy scholars to consider and investigate multiliteracies and multimodality.

Clarification of terms often used interchangeably or conflated with multiliteracies

In the literature, the term “multiliteracies” is often used interchangeably or conflated with other terms. Upon encountering these varied terms, those new to multiliteracies might wonder: What are the distinctions between *multiliteracies*, *multimodality*, *new literacies*, *New Literacies*, and *New Literacy Studies*? Do these differences matter? In this section, we take a small step toward answering these questions. The explanations here are, by necessity, brief and partial. The goal is to provide readers with a rudimentary definition of terms while including sign posts for further reading and exploration, since the complexity of these particular terms and their history cannot be thoroughly presented here. We begin with a discussion of New Literacy Studies (NLS), a movement that emerged prior to the drafting of the NLG multiliteracies publication and whose scholars had a profound impact on the development and legacy of multiliteracies.

New Literacy Studies (NLS)

Grounded in the ethnography of communication, NLS is a foundational framework that has foregrounded ethnographic explorations of local contexts whereas multiliteracies has primarily foregrounded pedagogical considerations grounded in global and historical contexts (Cope and Kalantzis, 2009). There is, however, significant overlap between NLS and multiliteracies. Street (2012) noted that NLS is an approach that “entails the recognition of ‘multiple literacies’” (p. 27). NLS is founded upon

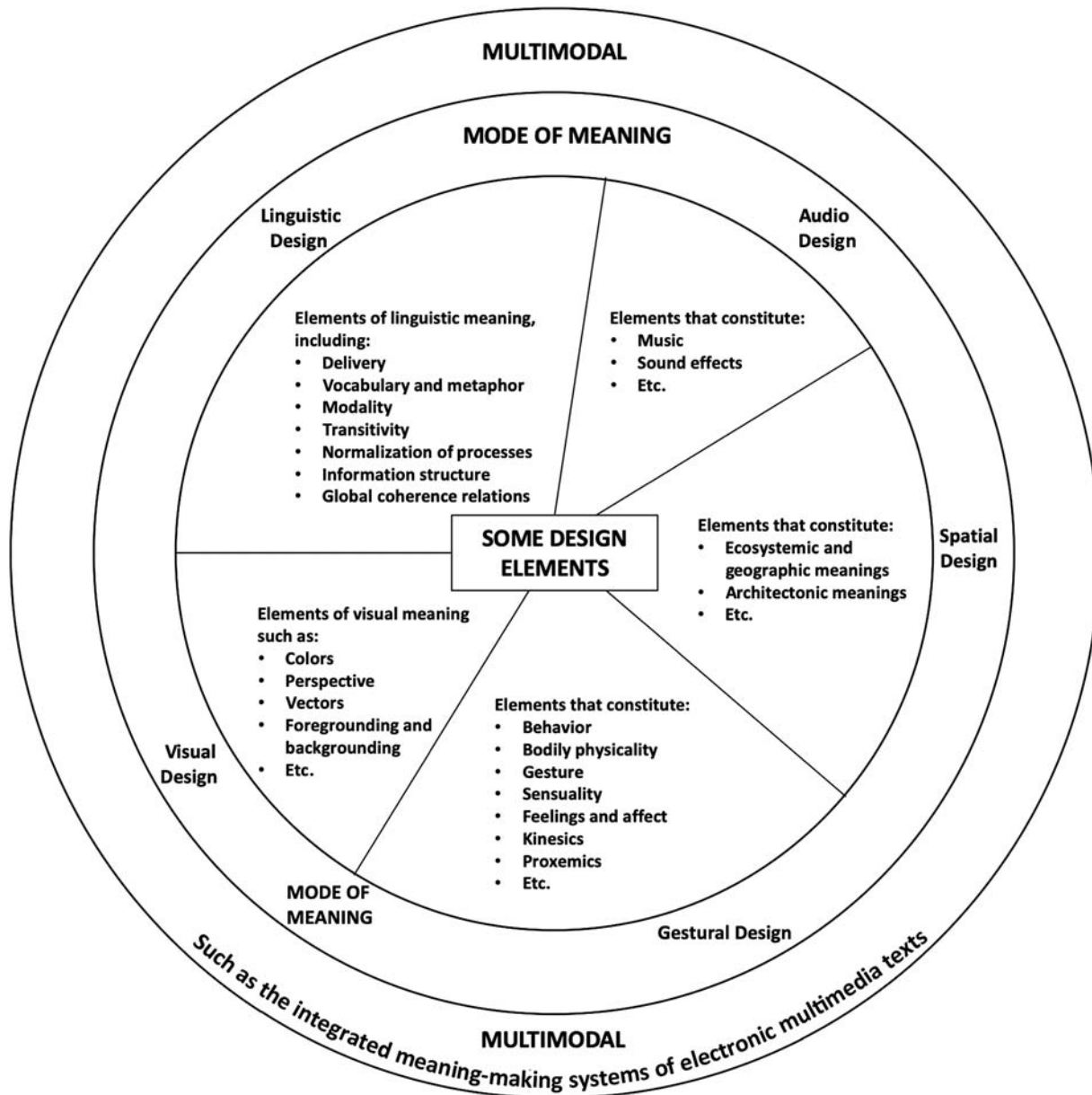


Fig. 1 Multiliteracies: metalanguages to describe and interpret different design elements of different modes of meaning (NLG, 1996, p. 83). Reprinted with permission from The New London Group, A pedagogy of multiliteracies: designing social futures, Harvard Educational Review, Volume 66:1 (Spring 1996), pp. 60–92. Copyright (c) by the President and Fellows of Harvard College.

scholarship from varied disciplines and the ethnographic or contextualized study of discourse (Gee, 2015). In particular, as postulated by NLS scholars (e.g., Barton, 1994; Gee, 1990; Heath, 1983; Street, 2003), multiliteracies and NLS share an emphasis on literacy as socially and culturally situated. A key distinction for many NLS scholars is that of literacy events and literacy practices. Grounded in a landmark ethnographic study, Heath (1982) identified a *literacy event* as “any occasion in which a piece of writing is integral to the nature of the participants’ interactions and their interpretative processes” (p. 93). Based in part on his ethnographic work in Iran, Street (1984) argued that literacy was more than mere acquisition of technical proficiency or skills in reading or writing—what he referred to as the “autonomous model of literacy” (p. 19). Street (1984) posited that “literacy practices” (p. 1) were embedded in social practices that depend “upon the context and the nature of the social formation” (Street, 1984, p. 1). Street (1988) further observed that literacy practices are broader than literacy events because they incorporate literacy events and the “ideological preconceptions that underpin them” (p. 61). Other scholars have continued to explore these distinctions (e.g., Barton et al., 2000; Burnett and Merchant, 2020) and have also considered NLS and multimodality (e.g., Rowsell et al., 2019).

Multimodality

In their original publication, the NLG (1996) posited that “One of the key ideas informing the notion of multiliteracies is the increasing complexity and inter-relationship of different modes of meaning” (p. 78). They indicated modes pertaining to linguistic, visual, audio, gestural, spatial, and multimodal design (see Fig. 1). The NLG differentiated multimodal design from other modes noting: “Multimodal Design is of a different order to the other five modes of meaning; it represents the patterns of interconnection among the other modes” (NLG, 1996, p. 78). Multiliteracies are thus inclusive of, but not the same as, multimodality. Researchers who explore multimodality often draw upon social semiotic theories of literacy (e.g., Lemke, 1989, 1998; Kress, 2010; Kress et al., 2001) and may frequently draw from or combine these with systemic functional linguistics (e.g., Jewitt et al., 2016; Unsworth, 2008). Regardless of the theoretical frameworks used, multimodality considers the ways in which people draw upon all modes to make meanings (e.g., combining images with print, using a gesture, using a gesture with a word and an image and so on). Body position, colors, movement, and sound are also explored in multimodal studies (Jewitt et al., 2016). Multimodality is frequently linked to multiliteracies (Mills, 2009), new literacies (Albers and Harste, 2007), and New Literacy Studies (NLS) (Rowse et al., 2019).

new literacies

The lowercase phrase, “new literacies,” typically encompasses the varied, emerging, and rapidly changing ways that literacy is being transformed beyond reading and writing of alphabetic text. Street (2003) has focused “what it means to think of literacy as a social practice,” (p. 77) distinguishing between literacy as a technical, and neutral skill, and literacy as an ideologically and culturally situated practice. Some authors foreground the affordances of multimodality in making literacies new, whether through traditional means such as art (Albers and Harste, 2007) or digital means for multimodal composing (Miller and McVee, 2012). Scholars also explore the ways in which digital texts and online contexts shift the reading strategies and positions of readers (Leu et al., 2008) with emphasis on social aspects of literacy (Lankshear and Knobel, 2011). The impact of new literacies on learning and design—what Cope and Kalantzis refer to as “learning by design” (2015)—is another thread of new literacies. Writings related to new literacies wend their way through landscapes of political, social, cultural, and technological change and share a common goal of transforming education and learning for teachers and students.

New Literacies

New Literacies foregrounds the Internet and the ways in which digital literacy, particularly the Internet and all it affords, is transforming literacy. As the editors of *The Handbook of Research in New Literacies* put it, “How do the Internet and other information and communication technologies (ICTs) alter the nature of literacy?” (Coiro et al., 2008). Leu et al. (2019) distinguish between a theory of New Literacies with uppercase N and L and new literacies, all in lowercase letters in what they refer to as a “dual level theory” (p. 324). Specific types of literacy changes or new forms of social communication such as text messages are “new literacies” (see pp. 325–326). As a broader theory, *New Literacies* conveys a sense of what happens across various domains where new literacies are emerging, and common patterns that emerge are “included in a broader, common, New Literacies theory” (p. 326). Gee (2015) refers to the New Literacies, as a “natural offshoot” (p. 44) of New Literacy Studies (NLS). It should be noted that while Leu and colleagues, Gee, and others have at times made careful distinctions between new literacies, New Literacies, and New Literacies Studies, not all authors adhere to these distinctions, particularly in earlier writings.

Having briefly introduced multiliteracies and terminology, the next section provides an illustration of one aspect of multiliteracy (i.e., multimodality) one aspect of STEM (i.e., engineering). Through this illustration, we explore multimodality and engineering literacies while foregrounding considerations of equity.

Multimodal literacies and engineering design

In a recent state-of-the-art literature review studying engineering, disciplinary literacies, and communication at the pre-college (K-12) level, a strand of research which focuses on studies using the perspectives of multimodality and multiliteracies has emerged to explain interactions within student engineering design teams (Silvestri et al., 2021). These studies purposefully foreground the ways in which engineers communicate their ideas, designs, arguments, and other meanings using a plethora of tools as well as multiple modalities for communication beyond speech and writing (e.g., demonstrations, gestures, inscriptions, etc.). Some publications foreground the multimodal nature of communication within design teams (Jordan, 2010; McVee et al., 2017b; Shanahan et al., 2016) whereas others foreground the types of multimodal tools (including digital notebooking) used to convey design ideas (Shanahan et al., 2017) or disciplinary literacy with multilingual learners (Paugh et al., 2018).

Within this strand of engineering and literacies research, special attention has focused on how foregrounding multimodal communication practices within engineering design teams can be of particular benefit for multilingual learners (McVee et al., 2017b; Shanahan et al., 2016; Shanahan et al., 2022). This intersection of engineering education and literacy research revealed “researchers attending to the power of multiple modes of communication, of ways students can be encouraged to interact with each other beyond linguistic means, and of the beneficial consequences of such practices for inclusion of English learners in STEM activities” (Silvestri et al., 2021, p. 10). It should be noted that multimodality is not limited only to engineering as increasing numbers of scholars have considered the role of multimodality in other STEM areas, for example, in science

(e.g., Grapin, 2019; Lee et al., 2019; Lemke, 1998; Tucker-Raymond et al., 2007; Varelas and Pappas, 2013; Wilmes and Siry, 2020; Williams and Tang, 2020).

Engineering and engineering design literacies are explored in the following example. After a brief description of participants and context, we explore the “communicative literacies” (Silvestri et al., 2021, p. 2) of students engaged in an engineering design task. Students engage in a broad range of multiliteracies practices. Primary analysis focuses on multimodality through analysis of gesturing, sketching (i.e., images), touching and handling materials, and speaking. The episodes illustrate how young, novice designers engage in *problem-scoping*, a process essential in engineering design. Watkins et al. (2014) define problem-scoping as: “the process of discovering and understanding aspects of the problem that need consideration” (p. 44). We explore the questions: What types of multimodal communication are used by children during the engineering design process? What do these multimodal communications reveal about problem-scoping?

Context and participants

The examples in this article derive from an elementary ELC in the northeastern United States. The ELC met after school twice a week for 1 h per session over seven weeks. The ELC was developed by elementary school teachers, university literacy faculty, and university engineering faculty and designed to serve children in grades three through five (i.e., 7–11 years of age). The purpose was to introduce engineering, teach basic engineering design principles, and stimulate interest in engineering, especially for groups who are minoritized in engineering as a college major and profession (e.g., students of color, girls). In particular, school leaders, teachers, and researchers agreed that students who were classified as English as a New Language (ENL) students (i.e., the state’s term for multilingual learners) should be specifically encouraged to participate. Additionally, students were accepted in the ELC regardless of academic standing.

The ELC used project-based learning (Krajcik and Shin, 2014) and engineering design process models (e.g., NASA & Design Squad PBS Kids, 2021) as a basis for minilessons, interactive discussions, and design challenges in small groups that would support the foundations of engineering knowledge and practices. While most sessions began with minilessons and interactive discussions, most of ELC was spent with students in design pairs or teams of up to five students, providing children the opportunity to design and build a bridge model, construct a rubber-band powered car, and create an amplifier that could be used with mobile devices to play music.

Data was collected throughout the ELC using audio recording, video recording, and fieldnotes. Student artifacts (e.g., drawn designs, photos of built designs, writing) were also collected. Audio and video recordings were transcribed; analytic memos written during transcription were used in analysis. Episodes in this article focus on fourth-grade girls, Ameena and Emma, as they designed a model bridge. Ameena, age 10, is an Arabic-English speaker. Emma, age 9, is a Chinese (Mandarin)-English speaker.

An exploration of multiliteracies in engineering literacies club (ELC)

The third- and fourth-grade students in ELC huddle together at the carpet watching a historical video clip (British Pathé, 2013) of the Tacoma Narrows Bridge collapse (<https://www.youtube.com/watch?v=XggxeuFDaDU>). After the dramatic finish, the lead teacher, Mrs. Stasik, prompts:

What do you notice? What do you wonder? You are going to go and design and build a bridge today with your group. What can you learn from this to help you with your design?

Within the whole-group setting, most children turn and talk to one another animatedly. Mrs. Stasik asks, “Did anyone’s conversation have to do with shapes? I’m not saying you are right or wrong; I’m just curious.” Ameena says, “I think there was something about designing with triangles.”

This discussion is situated in knowledge built across three previous sessions where students had been introduced to engineering, learning through failure, and engineering design process models (NASA & Design Squad PBS Kids, 2021). Children were introduced to engineers as people who solve problems and engage in productive communication (Shanahan et al., 2016). Children explored how technology could solve problems, pondering whether bridges were technology. Children wrote and discussed responses to questions such as: *What do bridges do for humans? What problems do bridges help with? Who builds bridges? What are bridges made of? What are some different kinds of bridges?* Ameena’s mention of triangles is a specific reference to a prior ELC session on strong shapes from TeachEngineering.org (2021). Although designed for older learners, ELC planners adapted the lesson for younger learners, for example, adding an activity using magnetized plastic rods to enable students to experiment with how shapes help create sturdy structures.

On this day, Ameena’s comment about triangles launches a conversation about how triangles can provide support for weight-bearing structures and provides an authentic and interactive review led by Mrs. Stasik. Mrs. Stasik then introduces and discusses the Client Requirements on a PowerPoint slide (Fig. 2). She reminds the children that individual sketches from the previous session were a type of brainstorming. She encourages them to communicate and plan before they build, stating “Communicating can look like sketching, talking, sketching, a mix of that, but before you get materials you need to have a plan. Mrs. Stasik uses the



Fig. 2 Engineering literacies club bridge project client requirements.

PPT to introduce the Design Review Checklist for planning. Building on the previous sessions, she reviews images of bridges (e.g., truss, suspension, beam, arch) stopping for discussion of ideas on how triangles are strong shapes. She reviews the engineering problem (i.e., designing a bridge that will reduce traffic backups and pollution, span the river, see template Fig. 4, and allow boats to go under it). A student asks if they will get a copy of the design requirements. With typical enthusiasm, Mrs. Stasik responds: “Yes, it’s in your baskets! What a good question. You are thinking like an engineer!” She adds that the Engineering Problem, Client Requirements and copy of the engineering design process are on the tables which are covered with paper for sketching. She reminds students of the requirement to use a variety of materials and projects a slide that shows images and labels of all materials, adding that this is also on the table. A student responds, “If we didn’t have that paper, then we would’ve forgot what ... materials there were.” After a few additional reminders, students move into their design teams.

Phase 1: Aameena and Emma begin to sketch bridge designs

Aameena and Emma settle down to work on their bridge design. Aameena holds her sketches made the previous session out toward Katie (second author on this article). She asks Katie, “Can I redraw? I want to make changes.” Katie affirms that they can certainly change their designs. Aameena then turns to Emma and says, “Let’s draw together, so we both agree with it.” Aameena starts sketching. Emma watches her and then begins drawing. When they have sketched for a while, Aameena looks away from her own sketch, which she has just erased. Aameena asks Emma “What kind of bridge?” In response, Emma says, “We should make the top ...” and gestures over the bridge in an arch (Fig. 3). Her full meaning is made apparent only in the gesture over the arch. Meanwhile, Aameena leans in, gesturing over Emma’s sketch and asks several questions. Then, Aameena begins a new sketch. As they sketch, the girls make comments about their designs.

This interaction between Aameena and Emma demonstrates multimodality as essential in communication and learning. The girls use spoken words (i.e., a linguistic mode) to communicate, but meanings do not reside in the words alone. Meanings are situated within the context and represented in artifacts such as the paper drawing (i.e., design images from the previous day) and the emerging sketch (i.e., the emerging image they are co-creating on this day). Additional affordances for meaning are present in Emma’s gesture—the movement of her hand across the span of the bridge arch represented on paper which, as noted, is not merely complementary to her speech but essential because words alone do not convey the full meaning. Her gesture brings a unique



Fig. 3 Emma using an arch-like gesture to communicate with Aameena. Note. This shows the beginning of Emma’s sketch. She later augmented the arch with more lines (i.e., triangles), showing the trusses.

contribution to the conversation (Goldin-Meadow and Alibali, 2013), referring to the visual-spatial form of the arch. The gesture, known as an “iconic gesture” (Roth, 2001, p. 370) is similar to the arch itself. Emma combines gesture with speech and the sketch (i.e., a representation) to complete her meaning.

When considering the practices of real-world engineers, sketching is an important mode of representation, and representations are essential in engineering design (Johri et al., 2013). Ferguson (1992, cited in Brun et al., 2016, p. 3) describes three types of sketching: “‘thinking sketches’ made by an engineering looking for new ideas, ‘talking sketches’ made when two engineers communicate, and ‘prescriptive sketches’ that are meant to please and convince people outside the design process” (Brun et al., 2016, p. 3). As the girls begin their collaboration, Ameena positions the sketch specifically as a means to communicate; she says “Let’s draw together, so we both agree with it.” Ameena and Emma do in fact use the drawing this way, exchanging different ideas while gesturing to and adding onto the drawing. The drawing mediates their design planning. In contrast, Ameena’s previous design can be viewed as a thinking sketch made for herself. On this day, the girls develop seven different sketches. (Sketches 4–7 will be introduced in the following section). The sketches function to help them communicate and clarify their ideas and reveal valuable insights into how budding engineers may engage in approaching collaborative problem-solving and design. Their approach to the sketches can also be seen as an example of “establishing collaborative group work,” a type of problem-scoping (English and King, 2017, p. 348).

Phase 2: Emma and Ameena shift their planning to include the template

At this point, Mrs. Stasik gives the group a river template (i.e., cardboard with blue lines and ripples) that will help students conceptualize the required length of the bridge to span the river (Fig. 4). Ameena promptly asks if they are able to draw on it, and Mrs. Stasik says yes but to use a pencil. The river template jumpstarts a new cycle of brainstorming and designing efforts mediated by rich intersections and affordances of the modalities of gesture, drawing, and speech. The physical artifact (i.e., river template) provided by Mrs. Stasik prompts a conversation about their design planning (i.e., problem-scoping (Watkins et al., 2014)).

Ameena:	So, um, let’s think of how we’re gonna do it before we draw.
Emma:	Yeah.
Ameena:	So, this this is it, and this is the river. So, if this is, like, water then land over here,
Emma:	Yeah.
Ameena:	The bridge would start right here.
Emma:	Right.
Ameena:	Oh the bridge would start here and then would like ... there’s room for people to walk on ... like that ...
Emma:	Yeah.

Ameena’s comment that they should “think” before they draw, and Emma’s agreement, indicates that they do not merely jump to solutions for the problem (i.e., bridge-building) without forethought but seek to further understand the problem. The girls focus their gaze on the template, noting which parts represent the land and which parts represent the water. Similar to students studied in Watkins et al. (2014), a nuanced, contextualized exploration of the girls’ interactions reveals an on-going process of problem-scoping (i.e., exploring and describing the boundaries of the problem they are to address). This consideration is key because most studies of expert and novice designers, often in an undergraduate engineering context, indicate that novices treat design tasks

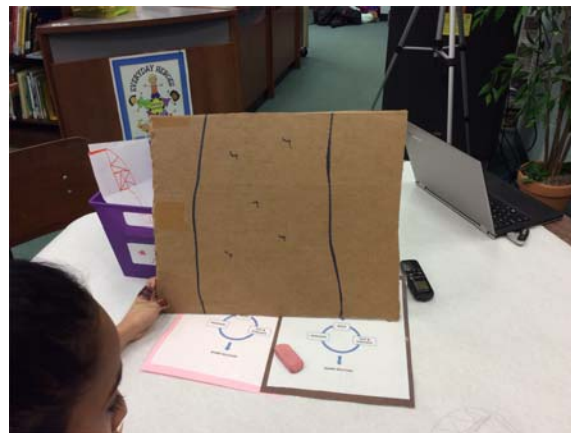


Fig. 4 The river template provided by Mrs. Stasik prompts a conversation about design plans.

as “a well-defined straightforward problem they prematurely attempt to solve” (Crismond and Adams, 2012, p. 748). In contrast, others (e.g., Ehsan and Cardella, 2020; Watkins et al., 2014) argue that nuanced, contextualized portraits reveal complex processes of problem-scoping even with children.

The girls continue to consider how to address their bridge design, specifically, engaging in what Ehsan and Cardella (2020) refer to as “problem-framing,” which is “adding/considering meaningful context” (p. 15). This meaningful context includes the physical artifact (i.e., the cardboard template) and the representative elements of the template (i.e., how the template represents land and water). The template serves as a mediational artifact (Roth et al., 1999; Pahl and Rowsell, 2010) that helps to support and refine their design process and requires them to grapple with issues of scale, engineering criteria, and constraints.

The girls begin to consider how their bridge design, which included beams and an arch with triangular trusses, might be represented on the template. As they talk, Ameena begins a smaller fourth sketch just below the template on the paper covering the table. She says “we should make the triangles under” but does not finish her thought. Emma asks, “Under what? Like, on the side? I mean, like, make our arch and so we can put it [triangles] like, through it?” Emma is quietly explaining her idea, gesturing to the paper. Ameena looks perplexed. “Can you draw it?” she asks. Emma makes a few gestures in the air and then begins a fifth sketch on the paper in front of her, pausing occasionally to gesture. She continues to explain what she is doing (i.e., adding the triangles through the arch). “Oh! Yeah, yeah, I get what you’re saying. So it’s like ...” Here Ameena rapidly draws an arc and draws triangles underneath it—the sixth sketch. Then she crosses out the drawing. She turns back to Emma pointing a pencil at Emma’s sketch asking: “Is there, like, gonna be some [something], like, here, too, for the cars?” Emma corrects her “No, like this” indicating where cars will traverse the bridge. “I get it,” Ameena says. “Let’s start drawing.”

Using pencil Ameena lightly draws the trusses (i.e., arches and triangle supports discussed in previous sessions) and columns supporting the platform—their seventh sketch—onto the template, constantly looking back and forth to the drawing that Emma had drawn on the table and talking with Emma. While drawing sidewalks on the side of the bridge as Emma suggested earlier, Ameena comments “And then here is where the, um,” and where she pauses, Emma interjects “the people part” goes. In this process, the girls work to combine their first drawings with clarifications they had recently sketched, so they can represent how their design will fit onto the cardboard template. They communicate through visual and gestural modes with little speech. Even when explaining verbally, few words are used. Ameena’s and Emma’s communication is indicative of what Schon (1992) observes about individuals engaged in design processes: “Designers know more than they can say, tend to give inaccurate descriptions of what they know, and can best (or only) gain access to their knowing-in-action by putting themselves into the mode of doing” (p. 131). Their use of the cardboard template represents how design is both social and material (Roth, 2014, 2017; Schon and Wiggins, 1992).

Across the phases of the project on this day, Emma and Ameena develop seven sketches. The first three sketches are individual designs developed while exchanging comments, questions (i.e., talking sketches). Ameena abandons her fourth sketch; not understanding what Emma is telling her, she asks Emma to draw a fifth sketch which Emma does. Ameena listens to and watches Emma sketch, but quickly turns and produces a sixth sketch copying Emma’s which she quickly scribbles out (i.e., clarifying her own understanding). Ultimately, Ameena draws the seventh and final sketch but does while referring back to Emma’s fifth sketch.

The final seventh sketch (see Fig. 5) is a very light pencil drawn on the cardboard. Throughout the episode, the girls communicate nonverbally through the sketch (i.e., image) but also through gestures (e.g., pointing to or making gestures over the images to demonstrate shapes or relevant aspects of the design). Their nonverbal communication is layered with verbal interactions. While their communication and disciplinary knowledge lacks the sophistication of practiced engineers, their use of drawing mirrors adult designers who use sketching for multiple purposes (Brun et al., 2016). Their sketching also represents characteristics described by de Vere et al. (2011) who note that “Sketching is a fluent and flexible process, producing implicit, abstract and inexact representations of possibilities, rather than well-defined solutions” (Section 4.1, n.p.).



Fig. 5 Ameena gestures across the triangles sketched on the template.

Phase 3: Emma and Ameena discuss and plan which materials to use

After completing their drawing on the cardboard template, Ameena prompts yet another round of pre-planning. She says:

So, before we start making this [the bridge], let's see what| how we're gonna make it first. So, do you think we should use the clay to make this part? (She gestures indicating the platform of the bridge). And then ... and then we use| bamboo sticks to make this. (She gestures across the triangle shapes drawn on the template. [See Fig. 5]).

Emma suggests instead that they use popsicle sticks and that they could go under the bridge to support it. When Ameena suggests using glue, Emma says to glue it and then put the clay on the popsicle sticks under the platform. They decide to use wire to make their arch.

Here, the girls engage in a number of problem-scoping activities that Ehsan and Cardella (2020) label “information gathering” (p. 15) by “exploring material” (p. 15). The girls first do this conceptually and verbally, referencing the drawn design and available materials. They discuss their materials with reference to their drawing on the template, frequently gesturing to the template (Fig. 5). As the girls complete their design and materials discussion, researcher Katie, who is taking notes nearby, asks them to explain their drawing. The girls explain how their built design will take shape and the materials they will use. Mrs. Stasik then checks in, and after a quick update, Ameena and Emma go to the materials table.

Phase 4 : Ameena and Emma gather materials and organize guides

After returning from the materials table, the girls engage in organizational planning practices. First, they prop up the cardboard template so they can refer to the image drawn lightly in pencil. On the table directly in front of them, they carefully arrange left to right: *Client Requirements* (see Fig. 2), *Checklist, Your Engineering Problem*. Only after setting out all the guide sheets do they begin working with the modeling clay that will serve a base for each pillar (see Fig. 6).

These organizational practices are similar to the ways practicing engineers might review client requirements and the problem they are tasked to solve. The girls situate printed client requirements, checklist, and engineering problem centrally. They do not explicitly review the guide sheets verbally, but place them where they can use them authentically as visual and mnemonic devices. Here, the girls are engaging the use of print text as a disciplinary literacy (i.e., using text in a specific disciplinary manner). They then begin working with the materials (e.g., modeling clay) and continue building in a collaborative manner.

Discussion

Problem-scoping is often included as one step in engineering design process models. Similar to others who have studied problem-scoping (e.g., Ehsan and Cardella, 2020; Watkins et al., 2014), in Emma and Ameena's group, problem-scoping was not limited to one step. Examples of problem-scoping exist on earlier days or whole-group settings as the problem was being introduced to the children. Later, after building their bridge, the girls also engaged in problem-scoping when they arrived at ELC and found their structure had collapsed. Problem-scoping was essential in fixing the problem and redesigning their structure. Comparable to Ehsan and Cardella (2020) and Watkins et al. (2014), the examples from Emma and Ameena show that the problem-scoping of young learners is more nuanced and sophisticated than has been portrayed by some (e.g., Crismond and Adams, 2012). When considered from a multimodal and multiliteracies perspective, the episodes presented here constitute several considerations and possible extensions for scholars of communicative literacies and engineering education, particularly when considering engineering design.

First, we explicitly attended to multiple modalities (e.g., gaze, gesture, touch, object and materials handling) in addition to spoken interactions, noting the ways in which meaning resides in these embodied processes and beyond the words spoken. We explored multiliteracies of engineering through “communicative literacies” (Silvestri et al., 2021, p. 2). In their review of the



Fig. 6 Template and guide sheets with materials chosen for bridge construction.

communicative literacies of engineering design, Silvestri et al. advocated for additional research on communicative or disciplinary literacies and posited that the multiliteracies framework could be applied to interdisciplinary studies of communication in K-12 engineering design teams. Specifically, Silvestri et al. proposed use of social semiotics (e.g., Kress, 2010; Norris, 2004) which can “provide a nuanced theoretical framing and vocabulary through which to consider within-group interactions” (Silvestri et al., 2021, p. 17). While numerous explorations of science have considered the registers and repertoires of language in science using social semiotics (e.g., Lemke, 1989, 1990, 1998; Halliday and Martin, 1993), there are, as yet, limited investigations of engineering design processes, especially for younger learners, using a multimodal lens. As part of the framework of multiliteracies, multimodality underscores the varied and rich resources for making meaning. In the episodes presented here, the process of sketching representations (i.e., images) are a nexus through which Emma and Ameena engage with ideas and problem-scoping. In these instances, the sketches take on the superordinate position followed by other modes such as gesture and speech.

Second, we explored how images, represented through sketches and the sketching process, play a role in the girls’ problem-scoping. Goldschmidt (2003) posits that “sketching assists in generating ideas and strengthening them by interpreting the ‘backtalk’ of a sketch in progress” (p. 72). She describes how children “read” drawings and that “designers habitually practice a similar process in the early idea-generation phase of the design process” (p. 72). Interestingly, while many engineering educators begin by emphasizing what experts do (Crismond and Adams, 2012), Goldschmidt uses the process of child development as a starting point and then argues that what expert designers do when sketching is actually an extension of what most children learn to do as they develop.

Few studies examine young children’s drawings during the engineering design process, but those that do typically explore the connection between children’s sketches and the outcome (i.e., built object) (Portsmore et al., 2012). Some researchers have found little or no correlation between sketching and outcomes (Rogers and Wallace, 2000), while others identified varying relationships (Fleer, 2000; Portsmore, 2009). Although Emma and Ameena ran out of time to build their arch, the built artifact (i.e., bridge) bears a strong resemblance to their sketches with pillars, trusses, and platform. Goldschmidt (2003) posits that the connection between the drawing and the built artifact is not the most crucial point. Goldschmidt proposes that sketching is important because: (1) sketching promotes idea generation; (2) sketching can be individually or collaboratively generative, and (3) sketching provides opportunities for reflection (Schon, 1983). The processes of drawing and multiple iterations as represented in Emma’s and Ameena’s designs and their multimodal engagement supports multiple types of problem-scoping. One paramount outcome is the disciplinary communication and meaning-making that takes place through, around, because of, and in response to the sketches and the utilization of other modalities.

The sketch affords opportunities for partners to seek clarification, work toward collaborative consensus, correct a partner’s misunderstandings or questions, and consider materials. These outcomes/processes/affordances of sketching are pertinent because they reveal the power of multiliteracies: Multiple semiotic signs (i.e., gestures, drawings, words) work together (i.e., inter-semiotically) to position learners. This meaning is not expressed in any one outcome—whether that outcome is a sentence or a built artifact. Meaning resides in the situated context and semiotic modes unfolding across time. Multimodality thus has tremendous power and potential to position young learners with multiple opportunities to learn. While Ameena and Emma’s final artifact (i.e., bridge) did, in fact, resemble their sketches, the resemblance is less significant than the processes of disciplinary communication and problem-scoping that they engaged in. From this vantage point, their position is not that of the deficit, a position often ascribed to students who are not yet fully bilingual. Rather, Emma and Ameena are positioned within the ELC community as competent problem-solvers and designers who can communicate using *multiple* semiotic signs. These signs may not yet include spoken words that indicate “precision as exactness in communicating disciplinary meaning” (Grapin et al., 2019, p.71), but this does not limit their opportunities for problem-solving or learning. Furthermore, as Roth (2017) observes “when students have the opportunity to design and produce engineering artifacts, their communicative competencies initially tend to be tied to work-related movements and materials for doing and exhibiting what is required to get the job done” (p. 257). Bodily movements thus ground abstract discourses with which students later engage (Roth, 2015).

Conclusion

Reflecting on multimodality and multiliteracies perspectives in engineering and STEM, several factors come to the fore. First, teachers are crucial in designing environments where multiliteracies can flourish. While the teachers do not figure prominently in the episodes described in this article, the teachers were responsible for creating a community where the multiple modes and language about design (Fig. 1) were used for learning and teaching. The teachers provided “overt instruction” (NLG, 1996, p. 88) or what Kalantzis et al. (2005) later identified as “conceptualizing” (p. 74) as needed to assist young learners. With regard to engineering and engineering design, teachers supported “defining and applying concepts” and “putting the concepts together that make discipline knowledge” (Kalantzis et al., 2005, p. 74). The ELC was created to engage in “situated practice” (NLG, p. 88) related to engineering and engineering design. Kalantzis et al. (2005) later labeled this “experiencing,” drawing upon “personal knowledge, evidence from learners’ everyday lives,” and “immersion in new information and experiences” (p. 73). This notion of engineering is particularly integral because engineering is a situated and material-centric practice (Johri et al., 2011).

Second, the context, where all learners engaged with engineering activities and the engineering design process regardless of their language background, academic standing, gender, or other characteristics, was essential to the success of the ELC. Following the NLG’s principle of equity for all students, the teachers, school leaders, and researchers agreed that no child would be excluded from ELC, particularly students who are minoritized in engineering careers. Multilingual students were encouraged to participate

regardless of their classifications in state hierarchies of language proficiency, and no students were excluded due to academic standing. The educators who established the ELC assumed that if multimodality was leveraged as a resource in a materials-rich engineering environment, students, especially multilingual learners, would thrive. In this way, the founders of the ELC took a step toward a more critical engineering design process pedagogy.

As always, there is room for improvement in teaching and learning for raising critical consciousness. In looking back on these episodes and others in the ELC, there could have been opportunities to directly link to students' and families' "funds of knowledge" (Gonzalez et al., 2005) to solve engineering problems in the school or engage in community-centered engineering projects (Jordan et al., 2021). While students learned about engineers as problem-solvers and some problems did address local community inequities, ELC could have included more direct exploration of critical literacies, justice, and equity pertaining to the "health, safety, welfare of the public" (Wilson-Lopez et al., 2017, p. 242) to help children foster deeper connections and critical perspectives of engineering (McGowan and Bell, 2020). While students were positioned in a non-deficit manner, more could have been done to explicitly create linkages that were asset-based approaches to engineering (Martin and Wendell, 2021) to further enhance opportunities to learn. Young engineers could also be introduced to multimodality as a resource for "futuremaking" (Turner and Girffin, 2020, p. 109) to envision themselves in STEM-related careers.

Researchers of engineering education can also learn much from the NLG (1996) multiliteracies framework and the outgrowth of multimodal research since the first publication of the multiliteracies framework. Writing in the *Journal of Engineering Education*, Case and Light (2011) noted that more diverse methods are needed in the study of engineering education and more explicit discussion of methods is needed. Multimodal methods (e.g., Norris, 2004; O'Hallorin, 2013) and multiliteracies bring methodological pluralism needed in studies of engineering education. Multimodality and multiliteracies can readily ground asset-based research approaches to engineering education, thus contributing to both classroom practice and research. In particular, multimodal research methods can surface non-linguistic communication that occurs frequently during the engineering design process, thus avoiding over-emphasis on verbocentric processes and outcomes that too often characterize research, teaching, and learning.

Acknowledgement

This research was supported in part by funding from The International Literacy Association Elva Knight Grant; University at Buffalo, SUNY IMPACT Grant, and University at Buffalo Center for Literacy and Reading Instruction (CLaRI).

References

- Albers, P., Harste, J.C., 2007. The arts, new literacies, and multimodality. *Engl. Educ.* 40 (1), 6–20.
- American Association for the Advancement of Science (AAAS), 1989. *Science for All Americans: A Project 2061 Report on Goals in Science, Mathematics, and Technology*. AAAS, Washington, DC.
- American Association for the Advancement of Science (AAAS), 1993. *Benchmarks for Science Literacy*. Oxford University Press.
- Au, W., 2016. Meritocracy 2.0: high-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educ. Pol.* 30 (1), 39–62. <https://doi.org/10.1177/0895904815614916>.
- Barton, D., 1994. *Literacy: An Introduction to the Ecology of Written Language*. Blackwell.
- Barton, D., Hamilton, M., Ivancic, R., 2000. *Situated Literacies: Reading and Writing in Context*. Routledge.
- Boyd, F.B., Brock, C.H. (Eds.), 2015. *Social Diversities within Multiliteracies: Complexity in Teaching and Learning*. Routledge.
- British Pathé, June 18, 2013. Tacoma Bridge Collapse: The Wobbliest Bridge in the World?, 1940 | British Pathé [Video]. YouTube. <https://www.youtube.com/watch?v=XgxeuFDaDU>.
- Brock, C.H., Goatley, V.J., Raphael, T.E., Trost-Shahata, E., Weber, C.M., 2014. *Engaging Students in Disciplinary Literacy, K-6: Reading, Writing, and Teaching Tools for the Classroom*. Teachers College Press.
- Brun, J., Le Masson, P., Weil, B., 2016. Designing with sketches: the generative effects of knowledge preordering. *Des. Sci.* 2 (13). <https://doi.org/10.1017/dsj.2016.13>.
- Burnett, C., Merchant, G., 2020. Literacy-as-event: accounting for relationality in literacy research. *Discour. Stud. Cult. Polit. Educ.* 41 (1), 45–56.
- Bybee, R.W., 2013. *A Case for STEM Education*. NSTA Press.
- Case, J.M., Light, G., 2011. Emerging methodologies in engineering education research. *J. Eng. Educ.* 100 (1), 186–210.
- Cazden, C., 2001. *Classroom Discourse: The Language of Teaching and Learning*, second ed. Heinemann.
- Coiro, J., Knobel, M., Lankshear, C., Leu, D.J. (Eds.), 2008. *Handbook of Research on New Literacies*. Taylor & Francis.
- Cope, B., Kalantzis, M., 2009. 'Multiliteracies': new literacies, new learning. *Pedagogies: Int. J.* 4, 164–195. <https://doi.org/10.1080/15544800903076044>.
- Cope, B., Kalantzis, M. (Eds.), 2015. *A Pedagogy of Multiliteracies: Learning by Design*. Palgrave Macmillan.
- Cope, B., Kalantzis, M. (Eds.), 2000. *Multiliteracies: Literacy Learning and the Design of Social Futures*. Routledge.
- Cope, B., Kalantzis, M., Smith, A., 2018. Pedagogies and literacies, disentangling the historical threads: an interview with Bill Cope and Mary Kalantzis. *Theory Into Pract.* 57 (1), 5–11. <https://doi.org/10.1080/00405841.2017.1390332>.
- Crismond, D.P., Adams, R.S., 2012. The informed design teaching and learning matrix. *J. Eng. Educ.* 101 (4), 738–797.
- Cunningham, C.M., Kelly, G.J., 2017. Epistemic practices of engineering for education. *Sci. Educ.* 101 (3), 486–505. <https://doi.org/10.1002/sce.21271>.
- Dare, E.A., Ring-Whalen, E.A., Roehrig, G.H., 2019. Creating a continuum of STEM models: exploring how K-12 science teachers conceptualize STEM education. *Int. J. Sci. Educ.* 41 (12), 1701–1720. <https://doi.org/10.1080/09500693.2019.1638531>.
- de Vere, I., Melles, G., Kapoor, A., 2011, August 15–18, 2011. Developing a drawing culture: new directions in engineering education. ICED 11, August 15–18, 2011. In: Paper Presented at the International Conference on Engineering Design. Lyngby, Copenhagen, Denmark.
- Ehsan, H., Cardella, M.E., 2020. Capturing children with Autism's engagement in engineering practices: a focus on problem scoping. *J. Pre-College Eng. Edu. Res.* 10 (1), 12–22. <https://doi.org/10.7771/2157-9288.1262>.
- English, L., King, D., 2017. Engineering education with fourth-grade students: introducing design-based problem solving. *Int. J. Eng. Educ.* 1 (33), 346–360.
- Fairclough, N., 2010. *Critical Discourse Analysis: The Critical Study of Language*, second ed. Routledge.

- Ferguson, E.S., 1992. *Engineering and the Mind's Eye*. MIT Press.
- Fleer, M., 2000. Working technologically: investigations into how young children design and make during technology education. *Int. J. Technol. Des. Educ.* 10, 43–59.
- Garcia, A., Luke, A., Seglem, R., 2018. Looking at the next 20 years of multiliteracies: a discussion with Allan Luke. *Theory Into Pract.* 57 (1), 72–78. <https://doi.org/10.1080/00405841.2017.1390330>.
- Gee, J., 1990. *Social Linguistics and Literacies: Ideology in Discourses*. The Falmer Press.
- Gee, J.P., 2014. *An Introduction to Discourse Analysis: Theory and Method*, fourth ed. Routledge.
- Gee, J.P., 2015. The new literacy studies. In: Rowsell, J., Pahl, K. (Eds.), *The Routledge Handbook of Literacy Studies*. Routledge, pp. 35–48.
- Goldin-Meadow, S., Alibali, M.W., 2013. Gesture's role in learning and development. *Annu. Rev. Psychol.* 64, 257–283.
- Goldschmidt, G., 2003. The backtalk of self-generated sketches. *Des. Issues* 19 (1), 72–88.
- Gonzalez, N., Moll, L.C., Amanti, C., 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities and Classrooms*. Erlbaum. <https://doi.org/10.4324/9781410613462>.
- Grapin, S., 2019. Multimodality in new content standards era: implications for english learners. *Tesol Q.* 53 (1), 30–55.
- Grapin, S.E., Llosa, L., Haas, A., Goggins, M., Lee, O., 2019. Precision: toward a meaning-centered view of language use with english learners in the content areas. *Ling. Educ.* 50, 71–83.
- Heath, S.B., 1982. Protean shapes in literacy events: ever-shifting oral and literate traditions. In: Tannen, D. (Ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Ablex Publishing, pp. 91–118.
- Hagopian, J., 2014. In: Hagopian, J. (Ed.), *More than a Score : The New Uprising against High-Stakes Testing/*. Haymarket Books.
- Halliday, M.A.K., Martin, J.R., 1993. *Writing Science: Literacy and Discursive Power*. The Falmer Press, Bristol, PA.
- Heath, S.B., 1983. *Ways with Words: Language, Life, and Work in Communities and Classrooms*. Cambridge University Press.
- Honey, M., Pearson, G., Schweingruber, H. (Eds.), 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. National Academies Press.
- Jewitt, C., 2009. An introduction to multimodality. In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*. Routledge, New York, pp. 14–27.
- Jewitt, C., Bezemer, J., O'Halloran, 2016. *Introducing Multimodality*. Routledge.
- Johri, A., Olds, B.M., Esmonde, I., Madhvan, K., Roth, W.-M., Schwartz, D.L., Tabak, I., 2011. Situated engineering learning: bridging engineering education research and the learning sciences. *J. Eng. Educ.* 100 (1), 151–185.
- Johri, A., Roth, W.-M., Olds, B.M., 2013. The role of representations in engineering practices: taking a turn towards inscriptions. *J. Eng. Educ.* 102 (1), 2–19.
- Jordan, M.E., 2010. Collaborative robotics engineering projects: managing uncertainty in multimodal literacy practice in a fifth-grade class. *Yearbk. Natl. Read. Conf.* 59, 260–275.
- Jordan, M.E., Zuiker, S., Wakefield, W., DeLaRosa, M., 2021. Real work with real consequences: enlisting community energy engineering as an approach to envisioning engineering in context. *J.Pre-College Eng. Edu. Res.* 11 (1), 230–255. <https://doi.org/10.7771/2157-9288.1294>.
- Kalantzis, M., Cope, B., Learning By Design Project Group, 2005. *Learning by Design*. Common Ground.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (11), 2–11.
- Krajcik, J.S., Shin, N., 2014. Project-based learning. In: Sawyer, R.K. (Ed.), *Handbook of the Learning Sciences*, second ed. Cambridge University Press, New York, pp. 275–297.
- Kress, G., 2009. What is mode? In: Jewitt, C. (Ed.), *The Routledge Handbook of Multimodal Analysis*. Routledge, pp. 54–67.
- Kress, G., 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge.
- Kress, G., Jewitt, C., Ogborn, J., Tsatsarelis, C., 2001. *Multimodal Teaching and Learning: Rhetorics of the Science Classroom*. Continuum.
- Lankshear, C., Knobel, M., 2011. *New Literacies: Everyday Practices and Social Learning*, third ed. Open University Press.
- Lee, O., Llosa, L., Grapin, S., Haas, A., Goggins, M., 2019. Science and language integration with English learners: a conceptual framework guiding instructional materials development. *Sci. Educ.* 103 (2), 317–337.
- Lemke, J., 1989. Social semiotics: a new model for literacy education. In: Bloome, D. (Ed.), *Classrooms and Literacy*. Ablex Publishing, pp. 289–309.
- Lemke, J.L., 1990. *Talking Science: Language, Learning, and Values*. Ablex Publishing Corporation.
- Lemke, J., 1998. Multiplying meaning: visual and verbal semiotics in scientific text. In: Martin, J.R., Veal, R. (Eds.), *Reading Science: Critical and Functional Perspectives on Discourses of Science*. Routledge, pp. 87–113.
- Lewis Ellison, T., Kirkland, D.E., 2014. Motherboards, microphones and metaphors: re-examining new literacies and black feminist thought through technologies of self. *E-learn. Digit. Media* 11 (4), 390–405. <https://doi.org/10.2304/elea.2014.11.4.390>.
- Leu, D.J., Coiro, J., Castek, J., Hartman, D.K., Henry, L.A., Reinking, D., 2008. Research on instruction and assessment in the new literacies of online comprehension. In: Block, C.C., Parris, S.R. (Eds.), *Comprehension Instruction: Research-Based Best Practices*. Guilford, pp. 321–346.
- Leu, D.J., Kinzer, C.K., Coiro, J., Castek, J., Henry, L.A., 2019. A dual level theory of the changing nature of literacy, instruction, and assessment. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R.B. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 319–346.
- Luke, A., 2012. Critical literacy: foundational notes. *Theory Into Pract.* 51 (1), 4–11.
- Martin-Páez, T., Aguilera, D., Perales-Palacios, F.J., Vilchez-González, J.M., 2019. What are we talking about when we talk about STEM education? A review of literature. *Sci. Educ.* 103 (4), 799–822. <https://doi.org/10.1002/sce.21522>.
- Martin, L., Wendell, K.B., 2021. Reflections on asset-based pre-college engineering education to promote equity: an introduction to the special issue. *J. Pre-College Eng. Edu. Res.* 11 (1), 42–45. <https://doi.org/10.7771/2157-9288.1325>.
- McGowan, V.C., Bell, P., 2020. Engineering education as the development of critical sociotechnical literacy. *Sci. Educ.* 29, 981–1005. <https://doi.org/10.1007/s11191-020-00151-5>.
- Menken, K., 2008. *English Learners Left behind: Standardized Testing as Language Policy*. Multilingual Matters, Buffalo, NY.
- Milner, H.R., 2012. Beyond a test score: explaining opportunity gaps in educational practice. *J. Black Stud.* 43 (6), 693–718. <https://doi.org/10.1177/0021934712442539>.
- McVee, M.B., Gavelek, J., Shanahan, L.E., 2017a. Beyond multimodality and toward embodiment in multiliteracies: what literacy scholars can learn from embodied cognition. In: Serafini, F., Gee, E. (Eds.), *Remixing Multiliteracies: Theory and Practice from New London to New Times*. Teachers College Press, pp. 148–161.
- McVee, M.B., Silvestri, K., Shanahan, L.E., English, K., 2017b. Productive Communication in an Engineering afterschool club with girls who are English Language Learners. *Theory Into Pract.* 56 (4), 246–254.
- Miller, S.M., McVee, M.B., 2012. *Multimodal Composing: Learning and Teaching for the Digital World*. Routledge.
- Mills, K.A., 2009. Multiliter. *Interrog. Compet. Discour. Lang. Educ.* 23 (2), 103–116.
- Mirra, N., Morrell, E., Filipiak, D., 2018. From digital consumption to digital invention: toward a new critical theory and practice of multiliteracies. *Theory Into Pract.* 57 (1), 12–19.
- Nakata, M., 2001. Multiliteracies and indigenous standpoints. In: Kalantzis, M., Cope, B. (Eds.), *Transformations in Language and Learning: Perspectives on Multiliteracies*. Common Ground Publishing, pp. 113–120.
- NASA, Design Squad PBS KIDS, 2021. The design process in action. In: *Parents and Educators: Online Workshop*. Retrieved from. <http://pbskids.org/designsquad/parentseducators/workshop/process.html>.
- National Center for Education Statistics (NCES), 2019. *Status and Trends in the Education of Racial and Ethnic Groups 2018*. U.S. Department of Education. Retrieved from. <https://nces.ed.gov/pubs2019/2019038.pdf>.
- National Research Council (NRC), 1996. *National Science Education Standards*. National Academies Press.
- National Research Council (NRC), 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- National Academies of Sciences, Engineering, and Medicine [NASEM], 2018. *English Learners in STEM Subjects. Transforming Classrooms, Schools, and Lives*. National Academies Press, Washington, DC.

- New London Group, 1996. A pedagogy of multiliteracies: designing social futures. *Harv. Educ. Rev.* 66 (1), 60–92.
- Norris, S., 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. Routledge.
- O'Halloran, K.L., 2013. *A Multimodal Approach to Classroom Discourse Analysis*. Equinox Publishing.
- Pahl, K., Rowsell, J., 2010. *Artifactual Literacies: Every Object Tells a Story*. Teachers College Press, New York.
- Paugh, P., Wendell, K., Wright, C., 2018. Elementary engineering as a synergistic site for disciplinary and linguistic learning in an urban classroom. *Liter. Res. Theory Method Pract.* 67 (1), 261–278.
- Portsmore, M.D., 2009. *Exploring How Experience with Planning Impacts First Grade Students' Planning and Solutions to Engineering Design Problems*. Unpublished doctoral dissertation. Tufts University.
- Portsmore, M., Watkins, J., McCormick, M., 2012. Planning, drawing and elementary school students in an integrated engineering design and literacy activity. Washington D.C. April 26–28, 2012. In: Paper Presented at the 2nd P-12 Engineering and Design Education Research Summit.
- Rogers, G., Wallace, J., 2000. The wheels of the bus: children designing in an early years classroom. *Res. Sci. Technol. Educ.* 18 (1), 127–136.
- Roth, W.-M., 2017. The thinking body in/of multimodal engineering literacy. *Theory Pract.* 56, 255–262. <https://doi.org/10.3102/00346543071003365>.
- Roth, W.M., 2014. The social nature of representational engineering knowledge. In: Aditya, J., Olds, B. (Eds.), *Cambridge Handbook of Engineering Education Research*. Cambridge University Press, New York, pp. 67–82.
- Roth, W.-M., 2015. The emergence of signs in hands-on science. In: Trifonas, P. (Ed.), *International Handbook of Semiotics*. Springer, pp. 1271–1289.
- Roth, W.-M., 2017. The thinking body in/of multimodal engineering literacy. *Theory Pract.* 56, 255–262. <https://doi.org/10.1080/00405841.2017.1389218>.
- Roth, W.-M., McGinn, M.K., Woszczyzna, C., Boutonne, S., 1999. Differential Participation during science conversations: the interaction of focal artifacts, social configurations, and physical arrangements. *J. Learn. Sci.* 8 (3–4), 293–347.
- Rowsell, J., Kress, G., Pahl, K., Street, B., 2019. The social practice of multimodal reading: a new literacy studies-multimodal perspective on reading. In: Alvermann, D.E., Unrau, N.J., Sailors, M., Ruddell, R. (Eds.), *Theoretical Models and Processes of Literacy*, seventh ed. Routledge, pp. 514–532.
- Schon, D.A., 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.
- Schon, D.A., 1992. Designing as reflective conversation with the materials of a design situation. *Res. Eng. Des.* 3, 131–147. <https://doi.org/10.1007/BF01580516>.
- Schon, D.A., Wiggins, G., 1992. Kinds of seeing and their function in designing. *Des. Stud.* 13 (2), 135–156.
- Serafini, F., Gee, E. (Eds.), 2017. *Remixing Multiliteracies: Theory and Practice from New London to New Times*. Teachers College Press.
- Shanahan, L.E., McVee, M.B., Silvestri, K.N., Haq, K., 2016. Disciplinary Literacies in an engineering club: exploring productive communication and the engineering design process. *Liter. Res. Theory Method Pract.* 65 (1), 404–420.
- Shanahan, L.E., McVee, M., Silvestri, K., 2022. Elementary students' communicative practices in the engineering design process across materials and modes. In: Wilson-Lopez, A., Tucker-Raymond, E., Esquinca, A., Mejia, J.A. (Eds.), *The Literacies of Design: Studies of Equity and Imagination in Engineering and Making*. Purdue University Press, West Lafayette, pp. 42–67.
- Shanahan, L.E., Silvestri, K., McVee, M.B., 2017. Engineering design team journals: providing multimodal opportunities for English Language Learners to explain design choices. *J. Adolesc. Adult Literacy* 61 (4), 445–451.
- Silvestri, K., Jordan, M., Paugh, P., McVee, M.B., Schallert, D., 2021. Intersecting engineering and literacies: a review of the literature on communicative literacies in K-12 engineering education. *J. Pre-College Eng. Edu. Res.* 11 (1), 1–25. <https://docs.lib.purdue.edu/jpeer/vol11/iss1/1/>.
- Skerrett, A., 2015. *Teaching Transnational Youth: Literacy Education in a Changing World*. Teachers College Press.
- Street, B., 1984. *Literacy in Theory and Practice*. Routledge.
- Street, B., 1988. Literacy practices and literacy myths. In: Saljo, R. (Ed.), *The Written World: Studies in Literate Thought and Action*. Springer-Verlag Press, pp. 59–72.
- Street, B., 2003. What's "new" in new literacy studies? Critical approaches to literacy in theory and practice. *Curr. Issues Comp. Educ.* 5 (2), 1–14.
- Street, B., 2012. New literacy studies. In: Grenfell, M., Bloome, D., Hardy, C., Pahl, K., Rowsell, J., Street, B. (Eds.), *Language, Ethnography, and Education: Bridging New Literacy Studies and Bourdieu*. Routledge, pp. 27–49.
- TeachEngineering, 2021. *Triangles and Trusses [Lesson Plan]*. Retrieved from. https://www.teachengineering.org/lessons/view/cub_trusses_lesson01.
- Tucker-Raymond, E., Varelas, M., Pappas, C.C., Wentland, A., 2007. "They probably aren't named Rachel": young children's scientist identities as emergent multimodal narratives. *Cult. Stud. Sci. Educ.* 1, 559–592. <https://doi.org/10.1007/s11422-006-9017-x>.
- Turner, J.D., Griffin, A.A., 2020. Brown girls dreaming: adolescent Black girls' futuremaking through multimodal representations of race, gender, and career aspirations. *Res. Teach. Engl.* 55 (2), 109–133.
- Unsworth, L. (Ed.), 2008. *Multimodal Semiotics: Functional Analysis in Contexts of Education*. Continuum International Publishing.
- Varelas, M., Pappas, C.C., 2013. *Children's Ways with Science and Literacy: Integrated Multimodal Enactments in Urban Elementary Classrooms*. Routledge.
- Watkins, J., Spencer, K., Hammer, D., 2014. Examining young students' problem scoping in engineering design. *J. Pre-College Eng. Edu. Res.* 4 (1), 5.
- Williams, M., Tang, K.S., 2020. The implications of the non-linguistic modes of meaning for language learners in science: a review. *Int. J. Sci. Educ.* 42 (7), 1041–1067. <https://doi.org/10.1080/09500693.2020.1748249>.
- Wilmes, S.E., Siry, C., 2020. Science notebooks as interactional spaces in a multilingual classroom: not just ideas on paper. *J. Res. Sci. Teach.* 57 (7), 999–1027. <https://doi.org/10.1002/tea.21615>.
- Wilson-Lopez, A., Strong, K., Sias, C., 2017. Critical literacy, disciplinary literacy: reading the engineering-designed world. *Theory Into Pract.* 57 (4), 238–245.

Creativity in STEM education

Weiping Hu and Xinyi Li, MOE Key Laboratory of Modern Teaching Technology, Shaanxi Normal University, Xi'an, PR China

© 2023 Elsevier Ltd. All rights reserved.

Creativity	64
Creativity in STEM education: scientific and technological creativity	65
Development of adolescents' creativity in STEM education	65
Factors affecting creativity in the educational environment	66
Educational environment	67
Teaching method	67
Peer interaction	67
Emotion	67
Motivation	68
Knowledge	68
Personality	68
Cognitive inhibition	68
Information processing speed	69
Cultivation of creativity in STEM education	69
Cultivation of creativity through projects	69
Cultivation of creativity in STEM education	70
Individual	70
Environment	70
Process	70
Strategies	71
Future research prospects	71
Developing effective STEM curriculums	71
Exploring effective STEM classroom teaching	71
Constructing innovative and operational evaluation system for STEM education	72
References	73

The 21st century is an era of knowledge and economic globalization, and more and more attention is paid to scientific and technological innovation. With the increasing demand for science, technology and engineering talents, STEM education has become an educational mode and trend for cultivating comprehensive talents (Bressoud, 2015; Curtis, 2014). There has been a wave of STEM educational research in the world. Moreover, many schools begin to emphasize students' experience of science, and focus on cultivating students' scientific literacy and innovation capacity. STEM education is about the process of producing products and solving real-world problems through activities of engineering design (Kelley and Knowles, 2016). STEM courses integrate the knowledge of multiple disciplines, which enable students to understand science, technology, engineering and mathematics in the process of project-based learning and exploration. Specifically, STEM education emphasizes the use of creativity in the production of products (Hathcock et al., 2015). In this process, students use divergent thinking, convergent thinking and critical thinking to solve problems. They also collaborate to explore and analyze data to produce products. Therefore, one main goal of STEM education is to cultivate students' creativity (Lou et al., 2017), and development of innovative literacy is an important indicator of STEM education (Judson, 2014).

Creativity

Creativity is a primary driving force for individual and social development (Guilford, 1950). Creative thinking is the core of creativity. The cultivation of creative thinking is one primary goal for education (Hu et al., 2010a,b; Lin et al., 2003). Originality and usefulness are two key characteristics of creative thinking (Stein, 1953; Hennessey et al., 2010). In the process of creative thinking, individuals must give up the general method, break the original thinking sets, and think in a unique way (Stein, 1953; Runco and Jaeger, 2012). Creative thinking consists of divergent thinking and convergent thinking (Guilford, 1950, 1967). Divergent thinking is the process of generating multiple answers from different directions to an open question (Sternberg, 1999). Convergent thinking is the process of creatively generating correct answers from given information.

The measurement of creativity is the foundation of research on creativity. Most creative tests are domain-general. According to categories of creativity, the related tests can be divided into divergent thinking tests and convergent thinking tests (Sternberg, 1999). Divergent thinking tests are generally presented in the form of open questions, requiring participants to list as many answers as possible. Divergent thinking tests include Torrance Tests of Creative Thinking (TTCT; Torrance, 1965), Alternate Use Task (AUT),

the Instance Task, and the Consequence Task. The AUT task is to come up with as many novel uses for conventional objects (such as bricks) as possible (Fink et al., 2009; Fink and Benedek, 2014). The Instance Task is to list novel items in common categories (Silvia, 2011). The Consequence Task is to imagine a novel result that might occur in a virtual situation (Silvia, 2011). A convergence thinking test requires individuals to integrate and process target-related information through cognitive processes such as remote association, thinking reorganization, and prototype inspiration (Mednick, 1962). The convergent thinking tests include Remote Association Test (RAT) and Insight Problem Solving. The RAT is to come up with a word (e.g., candle) that can be combined with three unrelated words (e.g., light, birthday, wax) or words that have semantic associations (Mednick, 1968; Mednick and Mednick, 1967). Insight Problem Solving involves an experience of “Aha” feeling at the moment of problem solving. The classic insight problems include the candle problem (Duncker and Lees, 1945) and the Matchstick problem. Tests related to creativity also include self-evaluation or evaluation of personality traits or behaviors such as Creativity Achievement Questionnaire (Carson et al., 2005), Williams Creativity Scale (Williams, 1979). The measurement of creativity has contributed to the development of creative research.

Creativity in STEM education: scientific and technological creativity

Creativity is one of important educational goals (Bermejo et al., 2016). Education is an important way to cultivate students' creativity, especially the STEM education because STEM education can stimulate students' curiosity and focus on inquiry and the cultivation of critical thinking (Henriksen, 2014). STEM education encourages students to use interdisciplinary knowledge and methods to solve real situation problems and unify scientific inquires, technical productions, engineering designs and mathematical methods, aiming at improving students' scientific literacy and innovation ability.

Creativity is not only domain-general, but also domain-specific. Knowledge and skills in a specific field are main components of creativity. Many studies have emphasized that creative thinking requires specific domain or subject-based knowledge and skills (Amabile and Gryskiewicz, 1987). Especially in the fields of STEM (Science, Technology, Engineering, Mathematics), creativity involves the process of using information to produce novel products during solving scientific problems (Amabile, 1996; Hu and Adey, 2002). This process is called scientific creativity (Hu and Adey, 2002; Liang, 2002). Creative products of science and technology generated in STEM or science courses generally have less social value but more personal development value. During STEM learning, the creativity of science and technology is embodied in putting forward novel scientific problems, mastering scientific knowledge profoundly, transferring the mastered scientific knowledge to new situations, discovering unique methods to solve a certain problem, improving the experiment in textbooks, and producing a novel product in activities (Koray, 2003).

Hu and Adey (2002) proposed a three-dimensional Scientific Creativity Structure Model (SCSM), which emphasizes that the measurement of scientific creativity should consider the creative process, trait and product. The process involves designing products through divergent and convergent thinking, and the evaluation of scientific creative products according to the fluency, originality and flexibility of creativity. The products include not only creative outcomes and technical products, but also scientific knowledge gained by individuals and the findings and solutions of scientific phenomena and problems. According to this model, scientific and technological creativity can be demonstrated through observation of scientific phenomena, teaching of scientific knowledge, scientific problem solving, scientific inquiries and creative activities. Individuals can improve scientific and technological creativity by developing creative thinking and creative imagination.

There are several aspects to consider when studying scientific creativity. Different from artistic creativity, scientific creativity involves creative scientific inquiry, creative scientific problem finding and solving, and creative scientific activities. Scientific creativity does not include non-intellectual aspects, although non-intellectual aspects may affect scientific creativity. Scientific creativity relies on scientific knowledge and skills. The structure of scientific creativity reflects the main aspects of scientific creativity, which can guide the measurement, development and research of scientific creativity and the cultivation of young people's scientific creativity.

Researchers have studied ways and techniques to improve creativity (Hu and Adey, 2002; Liang, 2002; Ayas and Sak, 2014). Based on the structural model of scientific creativity and the Torrance Tests of Creative Thinking, the Scientific Creativity Test was developed (Hu and Adey, 2002). The test includes seven open-ended questions to assess students' scientific creativity. For example, one question is “Please write down as many scientific uses of a piece of glass as possible” (Hu and Adey, 2002). Students need to combine the scientific characteristics of glass to answer this question through divergent thinking. The test includes seven abilities or dimensions: creative object application ability, problem solving ability, product improvement ability, imagination ability, problem solving ability, experimental design ability, and technical product design ability. All the questions are open-ended, affording the fluency, flexibility and originality of the subjects' thinking and imagination in the process of answering the questions. Combining scientific creativity with problem finding, the Creative Science Problem Finding Test was also developed (Hu et al., 2010a,b). This test has two questions. One is open-ended, requiring subjects to find questions based on their daily life experience. The other is close-ended, asking subjects to find questions based on a picture of an astronaut standing on the moon. The Technical Creativity Test was also developed to explore the use of technical knowledge and skills by individuals to produce innovative products (Hu et al., 2011).

Development of adolescents' creativity in STEM education

The development of individuals' creativity is closely related to the development of individuals' cognition. In general, creativity increases with the increase of cognitive abilities. Many studies have explored the cultivation of individuals' creativity (Hu et al.,

2004; Kleibeuker et al., 2013a,b). The age trend of creativity, especially the development of creativity in adolescence, has attracted the interest of many researchers. The peaks and valleys in the development of creativity have also been the focus of research.

The development of individuals' scientific creativity of adolescence is a complex dynamic process, which is influenced by various factors. One study compared the developmental trends of scientific creativity among scientific creativity of Chinese and English adolescents (Hu et al., 2004). Chinese and English adolescents are exposed to different cultural backgrounds and educational systems. Culture, social environment, family education, science courses, science teaching, science teachers, and examination methods all have an impact on the development of young people's scientific creativity, which may lead to differences in the development of young people's scientific creativity in China and England. Specifically, teenagers in China and England were assessed using the Scientific Creativity Test. The results showed that the developmental trends of scientific creativity among Chinese and English teenagers was similar. Eleven to 13 and 14–16 years old are the critical periods for the rapid development of adolescents' scientific creativity. Adolescents' scientific creativity in these two stages show a continuously rising trend. At the age of 14, however, other dimensions of scientific creativity performance of adolescents show a decreased trend. At the age of 15, the scientific creativity of adolescents increases again. Adolescents' scientific creativity stabilizes at the age of 17.

With the growth of age, as adolescents master more technical knowledge and accumulate creative experience, their thinking ability and imagination ability gradually develop. Studies have suggested that domain knowledge and skills are a major factor affecting creativity (Amabile, 1983; Feldhusen, 1995). Therefore, with the accumulation of knowledge and experience, the creativity of teenagers is on the rise from the ages of 11 to 13 and 14 to 16 (Smith and Carlsson, 1983, 1985, 1990). This finding is consistent with other research studies (Claxton et al., 2005; Kleibeuker et al., 2013a,b). Claxton et al. (2005) examined the development of creativity of students in grades 4–9 through a longitudinal study. The results indicate that the students' graphic creativity in grade 9 is better than that in grade 6. Kleibeuker et al. (2013a,b) examined the development of the divergent thinking and convergent thinking abilities of individuals in four age groups (12–13, 15–16, 18–19, and 25–30). The results indicate that the performances of 18–19 and 25–30 years old groups are better than that of 12–13 and 15–16 years old on the insight problem-solving task. The 25–30 years old group performs better than 12–13 and 15–16 years old groups on verbal tasks of divergent thinking. The age of 15 is the peak time of development in visual space of divergent thinking.

There is a "low tide" in creative development at the age of 14. This is consistent with previous research (Cheung et al., 2004; Torrance, 1962; Lau and Cheung, 2010) which has indicated that the development of creative thinking from children to adults does not rise smoothly, but increases in a curved manner. There are a total of four lows, which are at 5 years old, 9 years old, 13 years old and 17 years old. Adolescents tend to be sensitive to pressures of social conventions during the transition period of their physical and mental development (Evans, 2006). The age of 14 is a time of transition from adolescents to youth, during which their logical thinking skills transition from experiential to theoretical (Larson, 2000). During this time, adolescents are easily affected by pressure from society, schools, teachers, and peers and feel unsafe and unreliable. This feeling may inhibit creativity by limiting divergent thinking. Another reason is that teachers do not pay much attention to the training of divergent thinking ability (Lau and Cheung, 2010). While logical thinking develops with age, divergent thinking, an important component of creative thinking, has not improved.

The scientific creative ability of 17 years old tends to stabilize. At this age, most of the thinking abilities and mental characteristics of adolescents have matured. In addition, 17 years old adolescents have clear learning motivation, stable learning strategies and personality characteristics. Therefore, the strong cognitive abilities of 17 years old lay the foundation for a steady stream of scientific creativity. The stable cognitive abilities of 17 years old adolescents provide the impetus for the development of their scientific creativity (Hu et al., 2004).

There is also a gender differences in scientific creativity. The performance of males is better than females in scientific creativity, especially in scientific problem-solving, spatial visualization and technical design. This finding is consistent with previous research results (Lau and Cheung, 2010; Alfonso-Benlliure and Santos, 2016; Lau and Cheung, 2015). Lau and Cheung (2015) examined gender differences in the variability of creativity distribution. The results indicate that the variability of the graphic creativity of boys in grades 1–9 is greater than that of girls, but the difference in verbal creativity is not significant. In the process of dealing with problems, adolescent females are more likely to be constrained by mindset than young men (Lau and Cheung, 2010). Adolescent males tend to focus on understanding and exploring problems and adopt unique strategies when facing difficult problems.

In short, the adolescent stage is the peak period for the development of scientific and technological creativity. Science and STEM education can significantly improve the scientific and technological creativity of adolescents. At the adolescent stage, students learn and master subject knowledge, problem solving skills and cognitive strategies. In addition, the adolescence stage is a critical period for the formation and development of self-identity (Arnold, 2017). The practical need for adolescents to apply STEM knowledge stimulate the maintenance of their self-identity. Therefore, in the adolescent stage, the rapid development of knowledge and cognitive abilities promote the development of creativity, leading to high performance in STEM learning for students.

Factors affecting creativity in the educational environment

According to the development trend of adolescents' creativity, the level of creativity is not fixed, but develops with the growth of the individual. Many cognitive factors such as domain-related skills, metacognition and motivation can promote the development of creativity through the influence of education and environment. Current studies focus on the factors that influence creativity in the educational environment. Many studies used the methods of behavioral science, cognitive neuroscience and educational

experiments to investigate the influence of students' internal factors and external factors on the cultivation of creativity in the educational environment (Lin et al., 2003; Hu and Han, 2006). In STEM education, the external environmental factors that affect the development of students' creativity include educational environment, teaching style, peer interaction, emotion and motivation. The internal individual factors that affect the development of students' creativity include knowledge, personality, cognitive inhibition and information processing speed.

Educational environment

Environmental factors that affect creativity can be divided into family environment, school environment and teaching environment. Family plays an important role in the development of the individual (Conger et al., 2010). Feldman (1999) has pointed out that parents' occupation, income, children's birth order and family religion are all factors that lead to differences on the performance of creativity. MacKinnon (1978) studied creative architects and found that their parents have some common characteristics. For example, parents give their children sufficient understanding and respect to allow them to live independently. At the same time, children feel supported from their parents. Baumrind (1991) classified family parenting styles into four types: authoritative, authoritarian, indulgent, neglecting. When parents adopt the authoritative parenting style, their children's creative performance will be better than those who adopt the authoritarian parenting style (Snowden and Christian, 1999).

The school environment affects the ability of creative scientific problem-finding by affecting the creative personality trait (Li et al., 2010). Openness is positively correlated with students' creative scientific problem-finding ability. Especially in the school environment, a good relationship between teachers and students can enhance the positive relation between openness trait and creative scientific problem-finding ability. In the class environment, the teaching environment that provides free discussion can promote the development of creativity. While the teaching environment where the teachers are in control and have too many institutional constraints can inhibit the development of students' creativity (de Souza Fleith, 2000). Besançon and Lubart (2008) have found that an open and democratic learning environment are more conducive to cultivating students' creative thinking than the traditional teaching learning environment. Research has also found that creativity can only be inspired in an atmosphere of tolerance and freedom (Kim, 2005). Moreover, STEM teaching in the laboratory is more conducive to the cultivation of creativity than in the classroom (Chanthala et al., 2018). In STEM teaching environment, the laboratory provides more real learning resources than the ordinary classroom, and promotes students' problem awareness and operation ability, so as to obtain more creative outcomes.

Teaching method

Teaching methods can also affect the development of students' creativity (Hu, 2015). The teaching methods usually include traditional Lecture-based teaching (LBT) and Inquiry-based teaching (IBT). LBT usually follows a subject curriculum. Each subject has a fixed structure and sequence, and uses standard textbooks (Jayawickramarajah, 1996). IBT is a teaching method that combines advanced questioning, student-centered discussion and discovery of central concepts through laboratory activities (Damjanovic, 1999; NRC, 1996). IBT focuses on students' critical thinking, practical ability and problem-solving ability (Kitot et al., 2010). Teaching research has suggested that the combination of LBT and IBT can enhance the teaching effect and improve the academic performance of students in chemistry (Shen, 2009) and politics (Zhou, 2012) subjects. In a study comparing the effects of different teaching methods on students' performance of creative scientific problem finding (Jia et al., 2017), it was found that the mixed teaching method combining LBT and IBT is better than the single LBT or IBT, especially in the dimension of flexibility and originality.

Peer interaction

In STEM teaching, peer interaction plays an important role in group creativity. The establishment of appropriate peer interaction groups can enable peers to share high-level thinking and learning outcomes and promote the development of students' critical thinking (Guiller et al., 2008; Han et al., 2013). Han et al. (2013) examined the effect of peer group types on the scientific creative problems finding at different levels of material learning. Interaction groups were divided into homogeneous group, heterogeneous group, and voluntary group. The results showed that after learning the low-difficulty materials, the students in the homogeneous group performed better on the scientific creative problem finding task than those in the voluntary and heterogeneous group. After learning the high-difficulty materials, the students in the voluntary group performed better than those in the homogeneous and the heterogeneous groups on the scientific creative problem finding task. This may be because the students in the voluntary group have a positive emotional state and a high willingness to cooperate. In addition, there may be an impact of peer ideas and openness on creativity in online collaborative teamwork (Pi et al., 2019a,b). Students with higher openness would pay more attention to novel ideas and be more creative when they were exposed to high percentage of original ideas from their peers. Above findings suggest that teachers should encourage students to share more of their own original ideas and pay more attention to the ideas of peers in online collaboration groups.

Emotion

Emotion is also an important factor that can affect creativity. Studies have found that positive emotions are conducive to scientific creative performance, and negative emotions hinder scientific creative performance (Abele-Brehm, 1992; Chen et al., 2016; Lyubomirsky et al., 2005; Isen, 2008). Positive emotions can improve individuals' cognitive flexibility and avoid the input of irrelevant or

extra information (De Dreu et al., 2008). However, research on convergent thinking has found the opposite result. Negative emotions can promote insight problem solving through cognitive persistence compared with positive emotions (Li et al., 2013). It was further found that current processing types regulate the influence of emotion on the insight problem solving (Li et al., 2020). The primed global processing expands individuals' attention scope. In this case, the negative emotion enables the individual to focus on the current problem by inhibiting the interference of irrelevant stimuli, thus promoting the insight problem solving. Primed local processing reduces the scope of attention. In this case, positive emotions promote insight problem solving by promoting defocused attention.

Motivation

Motivation is an important source of students' learning behaviors and can affect the creative performance (Feldhusen, 1995; Eisenberger and Shanock, 2003). Motivation can be divided into internal motivation and external motivation (Plant and Devine, 1998). Internal motivation focuses on the individual's attention and interest to the activity itself, while external motivation focuses on external rewards and guidance (Collins and Amabile, 1999; Levesque et al., 2010). Research investigated the influence of internal and external motivation on the finding of creative scientific problems (Hu and Zhou, 2010). Results indicate that internal motivation promotes the finding of creative scientific problems through positive incentives, which impacts the fluency and originality. There is an interaction between external motivation and the ability of creative scientific problem finding. The external motivation to expect the positive evaluation has no influence on the performance of creative scientific problem finding, but the external motivation to avoid the negative evaluation has a significant inhibiting effect on the performance of creative scientific problem finding.

Knowledge

Internal factors that affect the development of creativity include knowledge, personality, cognitive inhibition, information processing speed. Domain knowledge is one of the main factors affecting creativity. Without domain knowledge, it is impossible to achieve creative outputs (Amabile, 1983; Feldhusen, 1995). Mastery of a great deal of knowledge is a basic condition for creation, but creation is also affected by the quality of knowledge mastery (Weisberg, 1999). Research in education usually uses the knowledge visualization method (concept map) to explore the knowledge structure of students (Brinkmann, 2003). Flower and Hayes (1984) found that when organizing information and materials, many writers used concept mapping strategies to carry out "brainstorming", which helped them to be highly active during thinking. Riley and Åhlberg (2004) found that concept mapping tasks promote students' creativity.

Differences in scientific creativity performance of individuals with different knowledge structures was investigated (Hu and Zhang, 2007). It was found that there is a significant positive correlation between concept mapping performance and scientific creativity. The nodes of the concept map are significantly correlated with the fluency, originality and flexibility of scientific creativity. This study has also examined the process of students using knowledge to create. First, learners retrieve theme-related information in long-term memory in order to come up with as many concepts as possible. Then, learners need to detect the connection patterns among the retrieved information. Finally, they establish reasonable and novel connections between concepts. In addition, the type of knowledge is also important. Curriculum type also has an influence on students' scientific creativity (Cotabish et al., 2013). Innovative science and technology courses can promote the cultivation of creativity (Cotabish et al., 2013). Innovative science and technology courses focus on design, production and creative applications. Such courses can promote students' advanced cognition and innovation through the application of STEM knowledge.

Personality

Both individual factors and situational factors affect the generation of creative ideas (Amabile, 1996). Therefore, individuals' different personality characteristics cannot be ignored (Barron and Harrington, 1981; Mumford and Gustafson, 1988). Studies on the relationship between personality and creativity mainly use the personality theory to explain creativity, examine personality characteristics of famous people, and investigate the special personality characteristics that may be related to creativity. One study analyzed seven personalities that are highly related to creativity out of 38 personality structures (Martinsen, 2011). The seven personality types are: need for originality, flexibility, instability, ambition, agreeableness, associative orientation and motivation. Moreover, a study explored the relationship between school environment, students' personality traits and the ability of creative scientific problem finding. The results indicate that the openness personality significantly predicts the ability of creative scientific problem finding. High openness individuals tend to be curious, imaginative and receptive to new ideas, thus they perform better in scientific creativity (Li et al., 2010).

Cognitive inhibition

A large number of studies have found that individuals with high creativity show variable attention and flexible cognitive inhibition (Vartanian et al., 2007; Zabelina and Robinson, 2010). A study explored the differences in cognitive inhibition between individuals with high and low levels of ability on scientific creative problem solving (Bai et al., 2014). The result indicates that individuals with high creativity could freely switch between focused attention and defocused attention to realize the flexible allocation of attention

resources. When interference information is presented, high cognitive inhibition ability of individuals with high scientific creativity can flexibly shift their focus of attention at different stages of problem solving. Furthermore, there is a relationship among cognitive style, cognitive inhibition, and scientific creativity (Cheng et al., 2016). The results suggest that cognitive style plays a mediating role in the relationship between cognitive inhibition and creative scientific creativity. The influence of cognitive style on creative scientific problem finding is modulated by external cues. Individuals with strong cognitive inhibition have a strong tendency of field independence in information processing and a weaker dependence on external heuristic information, so that they can more effectively use heuristic information to reduce processing load. On the contrary, individuals with weaker cognitive inhibition show a strong tendency of field dependence, and are more susceptible to the influence of irrelevant stimuli in information processing, which can lead to poor performance on scientific creative problem finding.

Information processing speed

Information processing speed is an important indicator of an individual's mental ability (Luciano et al., 2001). It refers to the number of tasks completed within a specified time. The information processing speed also affects creativity. During cognitive processing, learners with fast information processing speed are more actively thinking and are more likely to solve problems through insight (Hu et al., 2010). Compared with students with low creativity, students with high creativity process information more quickly (Silvia et al., 2013). In the process of thinking about new scientific ideas, interference information plays a moderating role in the relationship between creativity and processing speed (Dorfman et al., 2008; Vartanian et al., 2007). When interference information needs to be suppressed, scientific creativity is significantly positively correlated with processing speed (Hu et al., 2010). In the absence of interference information, there is a significant negative correlation between scientific creativity and processing speed.

Cultivation of creativity in STEM education

In STEM education, students' innovation literacy is usually cultivated in the following ways: project, curriculum and classroom teaching. Project training refers to project activities can be carried out at the national or school level to promote students' scientific inquiry and innovative literacy. Students can experience the fun of scientific creation, improve their curiosity and scientific motivation, and develop divergent and critical thinking. Curriculums can support the development of students' creativity (Hu, 2015). Some studies have indicated that the activity curriculum aims at cultivating students' thinking ability is effective (Hu, 2015; Cornelius-White, 2007). On the one hand, students can learn thinking methods through activity courses. On the other hand, thinking training integrated with curriculum knowledge helps students to acquire knowledge and learn to apply and transfer knowledge. Innovative classroom teaching is also an important way to cultivate students' creativity. Classroom teaching that promotes creativity has four key characteristics: learning motivation, cognitive conflict, social construction, self-monitoring, application. In the process of teaching, teachers need to set up cognitive conflicts to arouse students' motivation. Students explore, summarize and reflect through interaction, gain knowledge through operation and learn to transfer knowledge to solve practical problems.

Cultivation of creativity through projects

STEM teaching can be integrated into basic education courses, or combined with students' life to develop separate STEM curriculums. These STEM teaching strategies promote students' single-subject knowledge, creative motivation, and practical innovation ability in complex learning situations. Specifically, a large number of STEM schools have been established. These schools pay particular attention to creation of new courses and cultivation of high-level thinking skills by offering research-based learning activities for students (Atkinson and Mayo, 2010; Gonzalez and Kuenzi, 2012). PLTW (Project Lead The Way) has been developed (Project Lead The Way, 2018) in order to promote the training of students in STEM fields. PLTW has created a continuous pathway of education from secondary education to college in STEM fields. PLTW is a question-based course designed to increase interest and knowledge in mathematics, science and engineering. The MSIP (Mars Student Imaging Project) aims to encourage educators and students to participate in and experience professional NASA scientific research and engineering technology activities in real or simulated environments (Klug and Christensen, 2001). MSIP provides an authentic platform for students to practice science. Students can propose hypotheses, identify research topics, conduct surveys, collect and analyze data, solve problems, and communicate with Martian scientists. MSIP not only promotes the teaching of traditional subject knowledge, but also provides students with the opportunity to participate in the creative process that is closely related to reality.

"Spring shoot project" aims at cultivating creative talents jointly between universities and middle schools (Hu, 2015). The plan includes selecting middle school students with creative potential to go to the university laboratory to participate in research projects. So far, the selected students have carried out 33 projects in engineering applications, transportation, electronic technology, aerospace, nanotechnology, environmental engineering, network and so on, and have won many awards. It has been found that these students' creativity has significantly improved. This project has not only cultivated students' innovation ability, but also developed a group of research-oriented middle school teachers, who have begun to pay attention to and encourage students who have unique ideas in class. Schools have also begun to emphasize the importance of practice, encouraged interaction and cooperation between the students.

Cultivation of creativity in STEM education

Students' scientific and technological creativity can be cultivated through education. One goal of STEM education is to cultivate students' innovative literacy, so that it can promote the development of creativity beyond conventional education. The cultivation of students' creativity in STEM education can be carried out from four aspects: individual, environment, process and strategy.

Individual

STEM education can cultivate students' personality traits, such as curiosity, inspiration, independence and critical thinking. Studies have shown that personality traits have a significant predictive effect on the generation of creative ideas (Eysenck, 1993; Prabhu et al., 2008). The interdisciplinary and situational characteristics of STEM education play an important role in supporting the cultivation of students' scientific and technological creativity. STEM education is based on real world problem solving and emphasizes the integration of knowledge and skills in science, technology, engineering and mathematics. This can help students break the knowledge barrier caused by traditional subject-based education and learn to understand the world from a systematic perspective. The learning of interdisciplinary knowledge can promote the accumulation of students' experience, thus improving the thinking flexibility, enhancing the association between different ideas, and overcoming the functional fixation (Moirano et al., 2020). STEM education is based on problem or project-based learning. Therefore, STEM activities stimulate students' motivation and help them solve problems in creative ways by designing activities that promote knowledge integration and transfer.

Environment

STEM teaching environment mainly includes learning atmosphere, teacher-student relationship and teaching resources. The studies have indicated that creative classroom atmosphere, democratic teacher-student relationship, harmonious collaboration between the students, and abundant educational resources play an important role in supporting the development of creativity (de Souza Fleith, 2000; Cornelius-White, 2007). On the one hand, STEM teaching emphasizes that teachers should organize and inspire students strategically to build exploratory and challenging teaching content for students and encourage a creative classroom atmosphere. On the other hand, collaborative inquiry learning in STEM education stimulates students' engagement, which promotes the exchange of ideas among students. This process is conducive to the construction of collective knowledge and the development of creative thinking. In addition, STEM education usually relies on experimental environments equipped with advanced teaching tools, abundant resources and creative atmosphere. In such an environment, students are more likely to creatively solve problems in multiple ways and share their creative ideas. Therefore, this environment also promotes the development of students' scientific and technological creativity.

Process

STEM teaching process mainly refers to activity design, teaching method and evaluation method. Creative activity design and teaching methods can provide an effective path for the development of creativity (Kolodner, 2002). Most of the activity designs of STEM teaching are project courses based on problem solving (Capraro et al., 2013). Problem-oriented STEM learning can help students connect knowledge with real-world problem situations and learn knowledge in the process of solving real-world problems (Parker and Lazaros, 2014). Moreover, another characteristic of project-based STEM courses is the integration of interdisciplinary knowledge. An important goal of STEM learning is the construction and learning of interdisciplinary concepts. Students should not only acquire the core knowledge of the subject, but also learn the common concepts that connect different subject areas. Research has indicated that interdisciplinary STEM activities can help students better understand subject concepts (Fulton and Britton, 2011). Teaching methods are also an important factor in cultivating students' scientific and technological creativity. Rogers et al. (2015) have proposed a teaching model that uses continuous themes and a multidisciplinary approach for interdisciplinary integration. This teaching model not only retains the advantages and contents of traditional courses, but also helps students solve real problems in real life. This teaching method that combines traditional curriculum content with real problems promotes students' creative problem solving.

In addition, studies have emphasized the role of teachers in project-based STEM learning (Fosnot and Dolk, 2001; Mehalik et al., 2008). STEM activities require teachers to design activities from concept to behavior based on course content. In the teaching process, teachers encourage students to work together as groups (Puntambekar and Kolodner, 2005). This may increase students' creative activities and promote the development of STEM skills for problem solving. Saxton et al. (2014) have designed a STEM common measurement system to evaluate the development of STEM education from four aspects: student learning, teaching practice, teacher professional development and school support. Studies have explored the impact of learning goals and performance standards set by teachers in classroom learning on diagnostic teaching (Webb and Jones, 2009).

With more diversified methods, evaluation of students' innovation ability is also included in the evaluation system (Juliantari et al., 2018; Keinänen et al., 2018). STEM education connects the learning and application of scientific, technological, engineering and mathematical concepts, emphasizes the use of scientific inquiry to explore the unknown things and phenomena, and promotes the creative design (Bequette and Bequette, 2012; Honey and Kanter, 2013). This process is not only the process of solving problems, but also the process of generating creative ideas (Guilford, 1959).

Strategies

The strategy part mainly refers to creativity training methods and the cultivation strategies. Creativity training includes brainstorming and thinking training methods. Creativity cultivation strategies include creating problem situations, stimulating students' interest, setting cognitive conflicts and assigning open tasks (Hu, 2015). The characteristic of STEM teaching is to emphasize students' design ability and problem solving ability in complex situations. Many studies have explored the learning strategies in STEM teaching (Mayo, 2009; Smith et al., 2009). Apedoe et al. (2008) used design-based learning to integrate engineering design and scientific inquiry to help students learn chemical concepts. Mayo (2009) analyzed the impact of game-based learning on creativity. Although researchers try to use different teaching methods to study STEM teaching strategies, problem-based learning and project-based learning are currently two most commonly used teaching methods in STEM education (Barell, 2006; Capraro et al., 2013). Project-based learning in STEM education focuses on the situational characteristics of the problem and stimulates students' design orientation (Kennedy and Odell, 2014). Research has indicated that this teaching strategy can promote students' cooperation and promote the implementation of creativity training and cultivation strategies (Kennedy and Odell, 2014). Specifically, project-based learning strategies can guide students to generate a variety of ideas or inspirations through accessing original experiences to solve problems and promote the production of creative outcomes (Parnes, 1987). Meanwhile, students gain experience in the process of practice, along with the review, thinking and insight of knowledge and experience. Research has found that in the process of engineering design, brainstorming and collaboration among team members, cognitive differences among members are conducive to the generation of creativity (Wang et al., 2011).

Future research prospects

Developing effective STEM curriculums

The focus of STEM courses is to promote the application of knowledge and the transfer of skills. The engineering and technical disciplines emphasize the application of skills by students. However, STEM courses and teaching that emphasize the use of technical tools may limit the depth and multidisciplinary integration of STEM content, making students only have shallow learning and fail to achieve deep learning. STEM education is not simply a combination of science, technology, engineering, and mathematics, but an effective integration of them. In real-world problem-solving, students gain knowledge through practice and application (Fortus et al., 2005). Studies have shown that due to the interdisciplinary characteristic of real-world problems, adopting this interdisciplinary learning approach in solving real problems can promote the understanding of STEM concepts. Therefore, future research should develop STEM courses that promote knowledge integration and break the disconnection between STEM disciplines and with other non-STEM disciplines. In addition, future research should integrate interdisciplinary learning methods into basic education courses to effectively conduct interdisciplinary STEM learning activities.

Specifically, the integrated curriculum should emphasize the application of knowledge and focus on the relationship between disciplines. The construction of STEM course resources should emphasize the "problem context" and the authenticity of the context. Context and problem are the key to the integration and connection of interdisciplinary knowledge. Real context problems help develop the ability of knowledge transfer for students.

Importantly, introducing real-world problems into the STEM curriculums and enabling students to explore knowledge in practical activities to solve real-world problems can promote students' creative problem-solving ability (Hu, 2015; English and Mousoulides, 2015). Moreover, STEM courses should set up situational activities that promote the cooperation and interaction among students and encourage students to be active, which can provide a positive environment for the construction of knowledge. Studies in STEM education should develop feasible courses based on the local economic and cultural characteristics. Such STEM courses will give students a sense of achievement, make them realize the close connection between STEM education and life, and enhance their learning motivation. In addition, it is necessary for future research to construct STEM courses with the core teaching goal of cultivating creativity. In terms of teaching goals, more attention should be paid to basic knowledge and creativity. Basic knowledge goals are related to the STEM project theme. Creative performance can be measured from three aspects: creative consciousness, creative thinking and creative products. Creative performance can be promoted through the learning of interdisciplinary knowledge in STEM courses.

Exploring effective STEM classroom teaching

Classroom teaching is the basic organizational form of school teaching and the main way of quality-oriented education and cultivating students' creativity. Herschbach (2011) has indicated that activities are an effective form of organization for STEM courses. STEM courses generally focus on projects, tasks or problems. During STEM learning, students need to solve problems or complete project tasks through learning activities one by one. These activities can connect the learning content with the real world. Students complete these activities through cooperative and practical learning. The completion of these activities can promote deep learning of students and cultivate their awareness and ability of learning.

However, in STEM teaching, the roles of students and teachers may be ambiguous, and the rules of teaching activities may not be clear, which could lead to inappropriate intervention of teachers in the process of teaching, as well as disorderly teaching patterns in the classroom. This is not conducive to the achievement of students' learning goals, and thus unable to effectively achieve the teaching goals. Future research should explore more effective forms of organization of teaching activities. In addition, future

research should develop more effective teaching activities, and analyze behaviors of teachers and students in the process of activities, as well as the interaction and communication between students. Future research should also explore the influence of teaching activity organization on students' creative problem-solving ability by focusing on groups and individuals.

Specifically, future research should explore the impact of teaching methods in STEM education on students' scientific and technological creativity. STEM education mainly adopts the following teaching methods: project learning, lecturing and discussion. Project learning is a method for students to build knowledge through teamwork. Factors such as project-based learning materials, relevance of the project to students, complexity of the project, provision of appropriate support, prior knowledge of students and teamwork ability will all affect the learning of the project. In studying STEM subject contents, students not only need to collect relevant materials for learning, but also need to be guided by teachers through lecturing. Providing instruction through lecturing is particularly important for students in lower grades. Future research should explore the use of various teaching methods and explore the optimal combination of teaching methods for cultivating students' creativity. In addition, future research should explore the relationship between teachers and students in the teaching process. Traditional teaching activities are mainly led by teachers, which is not conducive to students' creative thinking. In the process of solving creative problems, an individual's self-construction process is particularly important. Therefore, future research should explore the influence of teachers' intervention degree on students' self-construction in the teaching process.

Future research should explore suitable teaching contents for students based on STEM courses, such as the balance between practical knowledge and theoretical knowledge. Previous STEM curriculum activities mainly focused on practice. Although such activities could promote the development of students' problem-solving ability, they ignored the underlying principles of behaviors. Therefore, future research should focus on the balance between practical knowledge and theoretical knowledge in classroom teaching. In addition, future research should develop interesting and risky classroom activities to improve students' learning motivation. Scientific learning motivation has the functions of arousal, orientation, selection and reinforcement, and it affects students' psychological processes such as observation, memory, thinking, imagination and problem solving. It also affects the development of students' scientific and technological creativity (Bhakti et al., 2018). In science teaching, students' motivation to learn science should be stimulated by setting up problems and connecting science teaching with real world.

Constructing innovative and operational evaluation system for STEM education

There are two research directions for STEM educational evaluation. The first one is to conduct STEM education evaluation research through specific thinking ability cultivation (Lin et al., 2015). The second one is to construct the educational evaluation system according to the subject curriculum (English and King, 2015). The evaluation of STEM education that adopts the relevant curriculum model is mostly based on the curriculum standards of science, mathematics, and technology. However, the STEM teaching evaluation model that aims at the cultivation of thinking ability does not have a unified standard for reference. Moreover, the current evaluation of STEM education is based on a single measurement standard; it does not have a clear measurement goal for learning outcomes, which may not only waste teaching and learning time, but also reduce classroom creativity. It may also limit the development of students with high creative thinking. Educational evaluation should focus on the scientific measurement of students' continuous learning ability, personal literacy, and academic inquiry ability. Therefore, future studies could improve measurement methods to integrate STEM educational evaluation criteria. It is necessary to measure students' mastery of various skills and knowledge in the learning process and to increase the evaluation of students' innovation ability.

Real-time evaluation can help teachers identify meaningful behavioral patterns and thus enable teachers to adjust teaching method according to learning measurement results. The convergence of technology and education could provide new measures for STEM education assessment. Educational games, real-time evaluation and cognitive neuroscience technology are innovations in learning measurement. Future research should develop innovative and operational learning assessment methods and adopt technical measures to support the process of personalized assessment. For example, future research could develop software or programs to measure learning of students, and provide personalized and timely support. For teachers, these software or programs can provide teachers with real-time data on students' learning, and teachers can observe and record students' skills and norms of behavior, thus being able to evaluate students' learning behaviors in specific situations. At the same time, this type of assessment can provide immediate feedback to the teacher, allowing the teacher to provide other support to help students find key points in the learning process. Educational games can introduce students to scientific concepts through a variety of representations and provides real-time assistance (Filament Games, 2009; Herro and Clark, 2016). Educational games such as "Ripple Effects" (Olberding and Olberding, 2010) and "The Social Express" (The Language Express, 2013), which provide virtual environments, are new evaluation means to measure learners' Social ability.

With the development of educational neuroscience, exploring the effects of STEM teaching through brain imaging technology may become a future research trend. Brain imaging techniques such as functional MRI and event-related potentials can be used to explore appropriate learning resources and environments in a scientific way. Previous studies on the evaluation of teaching effect mostly focused on knowledge mastery and lacked systematic evaluation of learning ability, learning process and learning efficiency. Moreover, online learning is becoming more and more popular and has become the preferred way of learning for some students. Through brain imaging technology, the teaching design and learning effects of online learning can be monitored and guided. Educational neuroscience can measure students' learning state, ability, attention and emotion, and reveal the underlying cognitive neural mechanism of teaching effect, which can help teachers understand the internal cognitive state of students in the learning process, thus contributing to the effective control of classroom teaching (Gallistel and Matzel, 2013). In addition, brain imaging technology

can also measure the difference in teaching effect under different teaching methods, which is helpful for teachers to choose appropriate teaching methods in teaching and improve the teaching evaluation system (Bowers, 2016). For example, studies have explored differences in brain activity between high and low academic achievers (Chrysikou et al., 2020). The results showed that there are differences in brain function among students with different academic achievement levels. Therefore, future research can develop teaching evaluation system based on cognitive ability and brain network through brain imaging technology.

References

- Abele-Brehm, A., 1992. Positive and negative mood influences on creativity: evidence for asymmetrical effects. *Pol. Psychol. Bull.* 23 (3), 203–221.
- Alfonso-Benlliure, V., Santos, M.R., 2016. Creativity development trajectories in elementary education: differences in divergent and evaluative skills. *Think. Skills Creativ.* 19, 160–174.
- Amabile, T., Grysiewicz, S.S., 1987. Creativity in the R&D Laboratory. Center for Creative Leadership.
- Amabile, T.M., 1983. The social psychology of creativity: a componential conceptualization. *J. Pers. Soc. Psychol.* 45 (2), 357.
- Amabile, T.M., 1996. Creativity and Innovation in Organizations.
- Apedoe, X.S., Reynolds, B., Ellefson, M.R., Schunn, C.D., 2008. Bringing engineering design into high school science classrooms: the heating/cooling unit. *J. Sci. Educ. Technol.* 17 (5), 454–465.
- Arnold, M.E., 2017. Supporting adolescent exploration and commitment: identity formation, thriving, and positive youth development. *J. Youth Dev.* 12 (4), 1–15.
- Atkinson, R.D., Mayo, M.J., 2010. Refueling the US Innovation Economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education. The Information Technology & Innovation Foundation. Forthcoming.
- Ayas, M.B., Sak, U., 2014. Objective measure of scientific creativity: psychometric validity of the creative scientific ability test. *Think. Skills Creativ.* 13, 195–205.
- Bai, X., Gong, Y., Hu, W., Han, Q., Yao, H., 2014. The inhibitory mechanism of individuals with different scientific creativity. *Stud. Psychol. Behav.* 12 (2), 151–155.
- Barell, J.F., 2006. Problem-Based Learning: An Inquiry Approach. Corwin Press.
- Barron, F., Harrington, D.M., 1981. Creativity, intelligence, and personality. *Annu. Rev. Psychol.* 32 (1), 439–476.
- Baumrind, D., 1991. The influence of parenting style on adolescent competence and substance use. *J. Early Adolesc.* 11 (1), 56–95.
- Bequette, J.W., Bequette, M.B., 2012. A place for art and design education in the STEM conversation. *Art Educ.* 65 (2), 40–47.
- Bermejo, M.R., Ruiz-Melero, M.J., Esparza, J., Ferrando, M., Pons, R., 2016. A new measurement of scientific creativity: the study of its psychometric properties. *An. Psicolog.* 32 (3), 652–661.
- Besançon, M., Lubart, T., 2008. Differences in the development of creative competencies in children schooled in diverse learning environments. *Learn. Individ. Differ.* 18 (4), 381–389.
- Bhakti, Y.B., Astuti, D., Agustina, I., 2018. The influence process of science skill and motivation learning with creativity learn. *J. Educ. Learn.* 12 (1), 30–35.
- Bowers, J.S., 2016. The practical and principled problems with educational neuroscience. *Psychol. Rev.* 123 (5), 600.
- Bressoud, D., 2015. Insights from the MAA national study of college calculus. *Math. Teach.* 109 (3), 179–185.
- Brinkmann, A., 2003. Graphical knowledge display–mind mapping and concept mapping as efficient tools in mathematics education. *Math. Educ. Rev.* 16 (4), 35–48.
- Capraro, R.M., Capraro, M.M., Morgan, J.R. (Eds.), 2013. STEM Project-Based Learning: An Integrated Science, Technology, Engineering, and Mathematics (STEM) Approach. Springer Science & Business Media.
- Carson, S.H., Peterson, J.B., Higgins, D.M., 2005. Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativ. Res. J.* 17 (1), 37–50.
- Chanthala, C., Santiboon, T., Ponkham, K., 2018. January). Instructional designing the STEM education model for fostering creative thinking abilities in physics laboratory environment classes. In: AIP Conference Proceedings, vol. 1923 (1), p. 030010.
- Chen, B., Hu, W., Plucker, J.A., 2016. The effect of mood on problem finding in scientific creativity. *J. Creativ. Behav.* 50 (4), 308–320.
- Cheng, L., Hu, W., Jia, X., Runco, M.A., 2016. The different role of cognitive inhibition in early versus late creative problem finding. *Psychol. Aesth. Creat. Arts* 10 (1), 32.
- Cheung, P.C., Lau, S., Chan, D.W., Wu, W.Y.H., 2004. Creative potential of school children in Hong Kong: norms of the wallach-kogan creativity tests and their implications. *Creativ. Res. J.* 16 (1), 69–78.
- Chrysikou, E.G., Wertz, C., Yaden, D.B., Kaufman, S.B., Bacon, D., Wintering, N.A., et al., 2020. Differences in brain morphometry associated with creative performance in high- and average-creative achievers. *NeuroImage* 116921.
- Claxton, A.F., Pannells, T.C., Rhoads, P.A., 2005. Developmental trends in the creativity of school-age children. *Creativ. Res. J.* 17 (4), 327–335.
- Collins, M.A., Amabile, T.M., 1999. IS motivation and creativity. In: *Handbook of Creativity*, vol. 297, pp. 1051–1057.
- Conger, R.D., Conger, K.J., Martin, M.J., 2010. Socioeconomic status, family processes, and individual development. *J. Marriage Fam.* 72 (3), 685–704.
- Cornelius-White, J., 2007. Learner-centered teacher-student relationships are effective: a meta-analysis. *Rev. Educ. Res.* 77 (1), 113–143.
- Cotabish, A., Dailey, D., Robinson, A., Hughes, G., 2013. The effects of a STEM intervention on elementary students' science knowledge and skills. *Sch. Sci. Math.* 113 (5), 215–226.
- Curtis, T. (Ed.), 2014. Science, Technology, Engineering, and Mathematics Education: Trends and Alignment with Workforce Needs. Nova Science Publishers, Incorporated.
- Damnjanovic, A., 1999. Attitudes toward inquiry-based teaching: differences between preservice and in-service teachers. *Sch. Sci. Math.* 99 (2), 71–76.
- De Dreu, C.K., Baas, M., Nijstad, B.A., 2008. Hedonic tone and activation level in the mood-creativity link: toward a dual pathway to creativity model. *J. Pers. Soc. Psychol.* 94 (5), 739.
- de Souza Fleith, D., 2000. Teacher and student perceptions of creativity in the classroom environment. *Roeper Rev.* 22 (3), 148–153.
- Dorfman, L., Martindale, C., Gassimova, V., Vartanian, O., 2008. Creativity and speed of information processing: a double dissociation involving elementary versus inhibitory cognitive tasks. *Pers. Individ. Differ.* 44 (6), 1382–1390.
- Duncker, K., Lees, L.S., 1945. On problem-solving. *Psychol. Monogr.* 58 (5), i.
- Eisenberger, R., Shanock, L., 2003. Rewards, intrinsic motivation, and creativity: a case study of conceptual and methodological isolation. *Creativ. Res. J.* 15 (2–3), 121–130.
- English, L.D., King, D.T., 2015. STEM learning through engineering design: fourth-grade students' investigations in aerospace. *Int. J. STEM Educ.* 2 (1), 14.
- English, L.D., Mousoulides, N.G., 2015. Bridging STEM in a real-world problem. *Math. Teach. Middle Sch.* 20 (9), 532–539.
- Evans, G.W., 2006. Child development and the physical environment. *Annu. Rev. Psychol.* 57, 423–451.
- Eysenck, H.J., 1993. Creativity and personality: suggestions for a theory. *Psychol. Inq.* 4 (3), 147–178.
- Feldhusen, J.F., 1995. Creativity: a knowledge base, metacognitive skills, and personality factors. *J. Creativ. Behav.* 29 (4), 255–268.
- Feldman, D.H., 1999. The development of creativity. In: *Handbook of Creativity*, vol. 169.
- Filament Games, 2009. iCivics. iCivics, Washington, DC.
- Fink, A., Benedek, M., 2014. EEG alpha power and creative ideation. *Neurosci. Biobehav. Rev.* 44, 111–123.
- Fink, A., Graif, B., Neubauer, A.C., 2009. Brain correlates underlying creative thinking: EEG alpha activity in professional vs. novice dancers. *NeuroImage* 46 (3), 854–862.
- Flower, L., Hayes, J.R., 1984. Images, plans, and prose: the representation of meaning in writing. *Writ. Commun.* 1 (1), 120–160.
- Fortus, D., Krajcik, J., Dershimer, R.C., Marx, R.W., Mamlok-Naaman, R., 2005. Design-based science and real-world problem-solving. *Int. J. Sci. Educ.* 27 (7), 855–879.

- Fosnot, C.T., Dolk, M.L.A.M., 2001. *Young Mathematicians at Work*. Heinemann, Portsmouth, NH.
- Fulton, K., Britton, T., 2011. *STEM Teachers in Professional Learning Communities: From Good Teachers to Great Teaching*. National Commission on Teaching and America's Future.
- Gallistel, C.R., Matzel, L.D., 2013. The neuroscience of learning: beyond the Hebbian synapse. *Annu. Rev. Psychol.* 64, 169–200.
- Gonzalez, H.B., Kuenzi, J.J., 2012, August. *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*. Congressional Research Service, Library of Congress, Washington, DC.
- Guildford, J., 1959. Trends in creativity. In: Anderson, H. (Ed.), *Creativity and Its Cultivation*. Wiley, New York.
- Guilford, J.P., 1950. Creativity. *Am. Psychol.* 5 (9), 444.
- Guilford, J.P., 1967. Creativity: yesterday, today and tomorrow. *J. Creativ. Behav.* 1 (1), 3–14.
- Guiller, J., Durndell, A., Ross, A., 2008. Peer interaction and critical thinking: face-to-face or online discussion? *Learn. Instruct.* 18 (2), 187–200.
- Han, Q., Hu, W., Liu, J., Jia, X., Adey, P., 2013. The influence of peer interaction on students' creative problem-finding ability. *Creativ. Res. J.* 25 (3), 248–258.
- Hathcock, S.J., Dickerson, D.L., Eckhoff, A., Katsioloudis, P., 2015. Scaffolding for creative product possibilities in a design-based STEM activity. *Res. Sci. Educ.* 45 (5), 727–748.
- Hennessey, B.A., Amabile, T.M., Mueller, J.S., 2010. *Consensual Assessment*, Encyclopedia of Creativity.
- Henriksen, D., 2014. Full STEAM ahead: creativity in excellent STEM teaching practices. *STEAM J.* 1 (2), 15.
- Herro, D., Clark, R., 2016. An academic home for play: games as unifying influences in higher education. *Horizon* 24 (1), 17–28.
- Herschbach, D.R., 2011. The STEM initiative: constraints and challenges. *J. STEM Teach. Educ.* 48 (1), 96–122.
- Honey, M., Kanter, D.E. (Eds.), 2013. *Design, Make, Play: Growing the Next Generation of STEM Innovators*. Routledge.
- Hu, W., Adey, P., 2002. A scientific creativity test for secondary school students. *Int. J. Sci. Educ.* 24 (4), 389–403.
- Hu, W., Han, Q., 2006. The development of elementary school students' creative science question-asking ability. *Psychol. Sci.* 29 (4), 944–946.
- Hu, W., Zhang, C., 2007. The relationship between the abilities of interdisciplinary concept mapping and scientific creativity. *Acta Psychol. Sin.* 39 (4), 697–705.
- Hu, W., Zhou, B., 2010. The influences of motivation on creative scientific problem finding ability of first year students in high school. *Psychol. Dev. Educ.* (01), 31–36.
- Hu, W., Philip, A., Shen, J., Lin, C., 2004. The comparisons of the development of scientific creativity between English and Chinese adolescents. *Acta Psychol. Sin.* 36 (06), 718–731.
- Hu, W., Shi, Q.Z., Han, Q., Wang, X., Adey, P., 2010a. Creative scientific problem finding and its developmental trend. *Creativ. Res. J.* 22 (1), 46–52.
- Hu, W., Liu, X., Jia, X., 2010b. The relationship among the speed of information processing and creativity, general intelligence of middle school students. *Psychol. Sci.* 33 (6), 1417–1421.
- Hu, W., Wan, X., Yu, L., 2011. The development of technological creativity of children and adolescents. *Psychol. Res.* 4 (2), 24–28.
- Hu, W., 2015. Thinking-based classroom teaching theory and practice in China. In: *The Routledge International Handbook of Research on Teaching Thinking*, pp. 92–102.
- Isen, A.M., 2008. Some ways in which positive affect influences decision making and problem solving. In: *Handbook of Emotions*, vol. 3, pp. 548–573.
- Jayawickramarajah, P.T., 1996. Problems for problem-based learning: a comparative study of documents. *Med. Educ.* 30 (4), 272–282.
- Jia, X., Hu, W., Cai, F., Wang, H., Li, J., Runco, M.A., Chen, Y., 2017. The influence of teaching methods on creative problem finding. *Think. Skills Creativ.* 24, 86–94.
- Judson, E., 2014. Effects of transferring to STEM-focused charter and magnet schools on student achievement. *J. Educ. Res.* 107 (4), 255–266.
- Juliantari, N.K., Sudarsana, I.K., Sutriyanti, N.K., Astawa, I.N.T., Putri, I.D.A.H., Saddhono, K., 2018. Educational games based in information technology as innovation evaluation activity in learning. *J. Phys. Conf.* 1114 (1), 012041. IOP Publishing.
- Keinänen, M., Ursin, J., Nissinen, K., 2018. How to measure students' innovation competences in higher education: evaluation of an assessment tool in authentic learning environments. *Stud. Educ. Eval.* 58, 30–36.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (1), 11.
- Kennedy, T.J., Odell, M.R.L., 2014. Engaging students in STEM education. *Sci. Educ. Int.* 25 (3), 246–258.
- Kim, K.H., 2005. Can only intelligent people be creative? A meta-analysis. *J. Second. Gift. Educ.* 16 (2–3), 57–66.
- Kitot, A.K.A., Ahmad, A.R., Seman, A.A., 2010. The effectiveness of inquiry teaching in enhancing students' critical thinking. *Procedia-Soc. Behav. Sci.* 7, 264–273.
- Kleibueker, S.W., De Dreu, C.K.W., Crone, E.A., 2013a. The development of creative cognition across adolescence: distinct trajectories for insight and divergent thinking. *Dev. Sci.* 16 (1), 2–12.
- Kleibueker, S., Koolschijn, P.C., Jolles, D., De Dreu, C., Crone, E.A., 2013b. The neural coding of creative idea generation across adolescence and early adulthood. *Front. Hum. Neurosci.* 7, 905.
- Klug, S.L., Christensen, P.R., 2001. Involving Students in Active Planetary Research During the 2001 Mars Odyssey Mission: The THEMIS Student Imaging Project. LPI, p. 1965.
- Kolodner, J.L., 2002. Facilitating the learning of design practices: lessons learned from an inquiry into science education. *J. Ind. Teach. Educ.* 39 (3), 1–31.
- Koray, Ö., 2003. Unpublished Doctoral Dissertation. *The Influence of Science Education Based on Creative Thinking on Learning Products*, vol. 40. Gazi University Institute of Social Sciences, Ankara, pp. 580–599.
- Larson, R.W., 2000. Toward a psychology of positive youth development. *Am. Psychol.* 55 (1), 170.
- Lau, S., Cheung, P.C., 2010. Developmental trends of creativity: what twists of turn do boys and girls take at different grades? *Creativ. Res. J.* 22 (3), 329–336.
- Lau, S., Cheung, P.C., 2015. A gender-fair look at variability in creativity: growth in variability over a period versus gender comparison at a time point. *Creativ. Res. J.* 27 (1), 87–95.
- Levesque, C., Copeland, K.J., Pattie, M.D., Deci, E.L., 2010. Intrinsic and Extrinsic Motivation.
- Li, H., Hu, W., Shen, J., 2010. School environment effects on the relations between adolescents' personality and creative scientific problem finding. *Psychol. Sci.* 33 (5), 1154–1158.
- Li, Y., Xiao, X., Ma, W., Jiang, J., Qiu, J., Zhang, Q., 2013. Electrophysiological evidence for emotional valence and competitive arousal effects on insight problem solving. *Brain Res.* 1538, 61–72.
- Li, X., Li, Y., Wang, X., Fan, X., Tong, W., Hu, W., 2020. The effects of emotional valence on insight problem solving in global-local processing: an ERP study. *Int. J. Psychophysiol.* 155, 194–203.
- Liang, J.C., 2002. *Exploring Scientific Creativity of Eleventh Grade Students in Taiwan* (Doctoral dissertation).
- Lin, C., Hu, W., Adey, P., Shen, J., 2003. The influence of CASE on scientific creativity. *Res. Sci. Educ.* 33 (2), 143–162.
- Lin, K.Y., Yu, K.C., Hsiao, H.S., Chu, Y.H., Chang, Y.S., Chien, Y.H., 2015. Design of an assessment system for collaborative problem solving in STEM education. *J. Comp. Educ.* 2 (3), 301–322.
- Lou, S.J., Chou, Y.C., Shih, R.C., Chung, C.C., 2017. A study of creativity in CaC2 steamship-derived STEM project-based learning. *Eurasia J. Math. Sci. Technol. Educ.* 13 (6), 2387–2404.
- Luciano, M., Wright, M.J., Smith, G.A., Geffen, G.M., Geffen, L.B., Martin, N.G., 2001. Genetic covariance among measures of information processing speed, working memory, and IQ. *Behav. Genet.* 31 (6), 581–592.
- Lyubomirsky, S., King, L., Diener, E., 2005. The benefits of frequent positive affect: does happiness lead to success? *Psychol. Bull.* 131 (6), 803.
- MacKinnon, D.W., 1978. *In Search of Human Effectiveness*. Creative Education Foundation.
- Martinsen, Ø.L., 2011. The creative personality: a synthesis and development of the creative person profile. *Creativ. Res. J.* 23 (3), 185–202.
- Mayo, M.J., 2009. Video games: a route to large-scale STEM education? *Science* 323 (5910), 79–82.
- Mednick, S.A., Mednick, M.T.S., 1967. *Examiner's Manual: Remote Associates Test*. Houghton Mifflin.
- Mednick, S., 1962. The associative basis of the creative process. *Psychol. Rev.* 69 (3), 220.
- Mednick, S.A., 1968. The remote associates test. *J. Creativ. Behav.*

- Mehalik, M.M., Doppelt, Y., Schuun, C.D., 2008. Middle-school science through design-based learning versus scripted inquiry: better overall science concept learning and equity gap reduction. *J. Eng. Educ.* 97 (1), 71–85.
- Moirano, R., Sánchez, M.A., Štěpánek, L., 2020. Creative interdisciplinary collaboration: a systematic literature review. *Think. Skills Creativ.* 35, 100626.
- Mumford, M.D., Gustafson, S.B., 1988. Creativity syndrome: integration, application, and innovation. *Psychol. Bull.* 103 (1), 27.
- National Research Council NRC, 1996. National Science Education Standards. National Academy of Sciences.
- Olberding, J.C., Olberding, D.J., 2010. "Ripple effects" in youth peacebuilding and exchange programs: measuring impacts beyond direct participants. *Int. Stud. Perspect.* 11 (1), 75–91.
- Parker, J., Lazaros, E.J., 2014. Teaching 21st century skills and STEM concepts in the elementary classroom. *Child. Technol. Eng.* 18 (4), 24–27.
- Parnes, S.J., 1987. The creative studies project. *Front. Creat. Res.* 156–188.
- Pi, Z., Yang, J., Hu, W., Hong, J., 2019a. The relation between openness and creativity is moderated by attention to peers' ideas in electronic brainstorming. *Interact. Learn. Environ.* 1–9.
- Pi, Z., Hong, J., Hu, W., 2019b. Interaction of the originality of peers' ideas and students' openness to experience in predicting creativity in online collaborative groups. *Br. J. Educ. Technol.* 50 (4), 1801–1814.
- Plant, E.A., Devine, P.G., 1998. Internal and external motivation to respond without prejudice. *J. Pers. Soc. Psychol.* 75 (3), 811.
- Prabhu, V., Sutton, C., Sauser, W., 2008. Creativity and certain personality traits: understanding the mediating effect of intrinsic motivation. *Creativ. Res. J.* 20 (1), 53–66.
- Project Lead The Way, 2018. Our Approach. Indianapolis. Author, IN. <https://www.pltw.org/about-us/our-approach>. Retrieved 25 Nov 2018.
- Puntambekar, S., Kolodner, J.L., 2005. Toward implementing distributed scaffolding: helping students learn science from design. *J. Res. Sci. Teach.* 42 (2), 185–217.
- Riley, N.R., Åhlberg, M., 2004. Investigating the use of ICT-based concept mapping techniques on creativity in literacy tasks. *J. Comput. Assist. Learn.* 20 (4), 244–256.
- Rogers, M., Pfaff, T., Hamilton, J., Erkan, A., 2015. Using sustainability themes and multidisciplinary approaches to enhance STEM education. *Int. J. Sustain. High Educ.* 16 (4), 523–536.
- Runco, M.A., Jaeger, G.J., 2012. The standard definition of creativity. *Creativ. Res. J.* 24 (1), 92–96.
- Saxton, E., Burns, R., Holveck, S., Kelley, S., Prince, D., Rigelman, N., Skinner, E.A., 2014. A common measurement system for K-12 STEM education: adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking. *Stud. Educ. Eval.* 40, 18–35.
- Shen, Y., 2009. The Research on the Effective Conformity of Inquiry Teaching and Traditional Teaching in Chemistry Lesson. Unpublished Master dissertation. Inner Mongolia Normal University, Hohhot, Inner Mongolia, China.
- Silvia, P.J., Beaty, R.E., Nusbaum, E.C., 2013. Verbal fluency and creativity: general and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence* 41 (5), 328–340.
- Silvia, P.J., 2011. Subjective scoring of divergent thinking: examining the reliability of unusual uses, instances, and consequences tasks. *Think. Skills Creativ.* 6 (1), 24–30.
- Smith, G.J., Carlsson, I., 1983. Creativity in early and middle school years. *Int. J. Behav. Dev.* 6 (2), 167–195.
- Smith, G., Carlsson, I., 1985. Creativity in middle and late school years. *Int. J. Behav. Dev.* 8 (3), 329–343.
- Smith, G.J., Carlsson, I.M., 1990. The creative process: a functional model based on empirical studies from early childhood to middle age. *Psychol. Issues* (57), 1.
- Smith, K.A., Douglas, T.C., Cox, M.F., 2009. Supportive teaching and learning strategies in STEM education. *N. Dir. Teach. Learn.* 2009 (117), 19–32.
- Snowden, P.L., Christian, L.G., 1999. Parenting the young gifted child: supportive behaviors. *Roeper Rev.* 21 (3), 215–221.
- Stein, M.I., 1953. Creativity and culture. *J. Psychol.* 36 (2), 311–322.
- Sternberg, R.J. (Ed.), 1999. *Handbook of Creativity*. Cambridge University Press.
- The Language Express, 2013. *The Social Express Teaching Manual*, first ed. Author, San Diego.
- Torrance, E.P., 1962. *Developing Creative Thinking through School Experiences. A Source Book for Creative Thinking*. Scribner's, New York, pp. 31–47.
- Torrance, E.P., 1965. *Rewarding Creative Behavior; Experiments in Classroom Creativity*.
- Vartanian, O., Martindale, C., Kwiatkowski, J., 2007. Creative potential, attention, and speed of information processing. *Pers. Individ. Differ.* 43 (6), 1470–1480.
- Wang, H.C., Rose, C.P., Chang, C.Y., 2011. Agent-based dynamic support for learning from collaborative brainstorming in scientific inquiry. *Int. J. Comp. Supp. Collab. Learn.* 6 (3), 371.
- Webb, M., Jones, J., 2009. Exploring tensions in developing assessment for learning. *Assess. Educ. Princ. Pol. Pract.* 16 (2), 165–184.
- Weisberg, R.W., 1999. I2 creativity and knowledge: a challenge to theories. In: *Handbook of Creativity*, vol. 226.
- Williams, F.E., 1979. Assessing creativity across Williams' cube" model. *Gift. Child. Q.* 23 (4), 748–756.
- Zabelina, D.L., Robinson, M.D., 2010. Creativity as flexible cognitive control. *Psychol. Aesth. Creat. Arts* 4 (3), 136.
- Zhou, Y., 2012. The Research on the Conformity of Inquiry Teaching and Traditional Teaching in Ideological and Political Lesson. Unpublished Master dissertation. Liaoning Normal University, Dalian, Liaoning Province, China.

Spatial thinking and STEM education

Siqi Li^{a,b}, ^a College of Education for the Future, Beijing Normal University at Zhuhai, Zhuhai, Guangdong, China; and ^b Research Institute of Science Education, Beijing Normal University, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	76
What is spatial thinking?	76
Existing concepts of spatial thinking	76
Spatial thinking framework introduced by National Research Council	77
How to measure spatial thinking?	77
Why should we consider spatial thinking in STEM education?	78
High demand versus low performance in students' spatial thinking	80
Low performance in overall spatial thinking	80
Low performance in spatial components that highly relate to STEM	80
Gender gap in spatial thinking	81
Facilitating the growth of spatial thinking	81
How to cultivate spatial thinking in STEM education?	81
Geo-technology and spatial thinking	82
The effectiveness of GIS integration in STEM classes	83
Approaching "best practice" in facilitating students' spatial thinking	83
Ensure enough frequency in taking GIS integrated STEM class	84
Emphasize the role of STEM teacher	85
Summary	86
References	86

Introduction

Researchers (i.e., geographers, astronomers, psychologists and educators) in National Research Council (NRC) argue that spatial thinking is universal and useful in a wide variety of academic disciplines and everyday problem-solving situations (NRC, 2006; NRC, 2013). The researchers also claim that spatial thinking can, and should, be taught at all levels in the educational system (NRC, 2006). The crucial role of spatial thinking has been widely acknowledged in the education field, especially for STEM disciplinary and interdisciplinary teaching and learning - in recent years; a growing number of STEM education research are conducted regarding spatial thinking (e.g., Bodzin, 2011; Jo and Bednarz, 2014; Shin et al., 2016; Sorby, 2009). Supporting by many research findings, a picture of the relationship between spatial thinking and STEM education is gradually emerging.

Past literatures of spatial thinking and STEM education suggest that the relationship between them is mainly two-fold: (1) Spatial thinking is one factor that may influence students' capacity in STEM disciplinary and interdisciplinary learning (Hegarty, 2010; Uttal et al., 2013; Uttal and Cohen, 2012); and (2) Taking well-designed STEM lessons, especially the ones that integrate geo-spatial technologies, could facilitate students' development of spatial thinking. In other words, the relationship between spatial thinking and STEM education is mutual; students' STEM learning and practice require a certain level of spatial thinking ability, meanwhile, doing STEM activities cultivates students' spatial thinking ability.

This article unpacks the above two major relationships between spatial thinking and STEM education by introducing the existing concepts and measurements of spatial thinking (i.e., what is spatial thinking and how to measure it), importance of spatial thinking in STEM learning (i.e., why educators need to draw attention to spatial thinking in STEM education), and the ways of cultivating students' spatial thinking in STEM classes (i.e., how to well-prepare students' spatial thinking development through STEM learning). Through illustrating what, why, and how in spatial thinking and STEM education, I will present several highlights in this research area at the end of the article.

What is spatial thinking?

Existing concepts of spatial thinking

The term "spatial thinking" has been used in both academic and non-academic settings extensively. A variety of definitions of spatial thinking exist in different disciplines. For example, Lohman (1979) defines spatial thinking ability as "the ability to generate, retain, and manipulate abstract visual images" (p. 126), while Schneider and McGrew (2012) define spatial thinking ability as the ability to make use of simulated mental imagery (often in conjunction with currently perceived images) to solve problems (2012, p. 129). Additionally, in geography education, the concept of spatial thinking has a somewhat nebulous

meaning, as various terms, such as “thinking spatially”, “geospatial thinking”, “spatial intelligence”, and “spatial ability”, are used synonymously (Huynh and Sharpe, 2013).

However, the existing concepts and terms either only refer to spatial thinking in certain field which is not a general and representative definition, such as in cognitive psychology and geography fields, or only represent a narrow concept which in fact is one of the many components in spatial thinking (Lee and Bednarz, 2011), such as spatial visualization (Carroll, 1993). Many researchers in this field have reached an agreement that providing a good definition of spatial thinking is not as easy as one would expect (Lee and Bednarz, 2011; Gersmehl and Gersmehl, 2007; Uttal and Cohen, 2012), which also raises a series of questions such as: is there one spatial ability, or are there many? How the different terms regarding to spatial thinking relate to one another? (Gersmehl and Gersmehl, 2007; Uttal and Cohen, 2012). Apparently, a widely accepted, broadly covered, and specific framework of spatial thinking is needed to clearly define a research field.

Spatial thinking framework introduced by National Research Council

Through a multi-year collaboration among experts in different fields, including education, geography, earth science, planetary science, psychology, and cognitive science, a NRC report on Learning to Think Spatially (NRC, 2006) clarifies definitions, concepts, and categories of spatial thinking, and provides a framework of spatial thinking for today’s educational milieu (NRC, 2006). In addition, it examines how spatial thinking could be incorporated into existing standards-based curricula, and emphasizes that spatial thinking must be recognized as a fundamental part of K-12 education, and as an integrator and a facilitator for problem solving across curricula (NRC, 2006). The framework has been widely recognized and applied by STEM educators (e.g., Bodzin, 2011; Lee and Bednarz, 2009; Tomaszewski et al., 2015).

The NRC (2006) defines spatial thinking as the knowledge, skills, and habits of mind to use concepts of space and representations, and processes of reasoning to structure problems, find answers, and express solutions to these problems. They introduce three spatial thinking contexts: life space (cognition in space), physical space (cognition about space), and intellectual space (cognition with space). The report also introduces several ways in which spatial thinking may be characterized (NRC, 2006). For example, spatial thinking can be broken down into different components, such as spatial knowledge, spatial ways of thinking and acting, and spatial capabilities; or concepts of space, tools of representation, and processes of reasoning. It also can be characterized through the differentiation of spatial component tasks, such as extracting spatial structures, performing spatial transformations, and drawing functional inferences (NRC, 2006). The interpretations of the above spatial thinking contexts and categories are presented in Table 1. Thus, it is important to notice that within the framework of spatial thinking provided by the NRC (2006), there is no one way to characterize spatial thinking, and the various characterizations of spatial thinking are all acceptable.

Within the contexts and characteristics of spatial thinking introduced by the NRC (2006), several theories of spatial thinking components have been developed following NRC’s (2006) framework (e.g., Golledge, 2008; Lee and Bednarz, 2011; Janelle et al., 2009). Most of the updated theories and framework are highly related to STEM teaching and learning, which can be regarded as reference for STEM educators. For example, Lee and Bednarz (2011) proposed eight specific spatial thinking components:

- (1) Visually navigating road maps using verbal information, including a participant’s current location, directions to destinations, and street information;
- (2) Recognizing map patterns and representing them in a graphic form;
- (3) Selecting an ideal location for a fictitious facility, based on multiple pieces of spatial information;
- (4) Creating a profile of topography along a proposed line on a contour map and properly orienting it;
- (5) Identifying spatial correlations between sets of maps, and displaying the identified spatial relationship in a graphic form;
- (6) Mentally visualizing a 3D image based on 2D information;
- (7) Visually verifying a map overlay process, and selecting the appropriate maps layers involved in the overlay;
- (8) Visually extracting types of spatial data from verbally expressed spatial information.

In summary, the NRC report (2006) provided an up-to-date definition of spatial thinking, reflecting a wider variety of academic disciplines and everyday problem-solving situations, and aligning with 21st century skills (NRC, 2006; Lee and Bednarz, 2004). NRC’s framework allows STEM educators and researchers accurately locate their position in studying spatial thinking and clarify the relationship between spatial thinking in STEM field and in others.

How to measure spatial thinking?

Assessing students’ and teachers’ spatial thinking ability in STEM field is needed in conducting relevant research. Spatial thinking is complex, and involves elements of interest to both psychologists and geographers (Huynh and Sharpe, 2013). In the field of psychology, a variety of tests have been widely used to measure individuals’ spatial abilities, but the tests in psychology are content general (Lee and Bednarz, 2011). Furthermore, spatial ability is a different concept from spatial thinking and it is a narrower concept than spatial thinking (NRC, 2006). In order to measure spatial thinking as defined by the NRC report (2006), researchers need to go beyond the limits of existing content general spatial tests (Eliot and Czarnolewski, 2007). Thus, only two measurement instruments that align with the NRC’s (2006) framework are discussed in this section, which might be used for STEM educators and researchers to obtain students and teachers’ spatial thinking ability.

Table 1 Interpretations of spatial thinking contexts and categories in [NRC \(2006\)](#).

<i>Categories and terms</i>			<i>Interpretations</i>
Spatial thinking contexts	Life space (cognition in space)		Building on the four-dimensional world of space-time to grip the spatial relations between self and objects in the physical environment.
	Physical space (cognition about space)		Building on the four-dimensional world of space-time with a focus on understanding the nature, structure, and function of phenomena.
	Intellectual space (cognition with space)		Understanding relationships among concepts and objects that are not in and of themselves necessarily spatial, but rather the nature of the space that is defined by the particular problem.
Spatial thinking category examples	Example 1	Spatial knowledge	Concepts such as symmetry, orientation, rotation, scale, relative versus absolute distance, and distance decay.
		Spatial ways of thinking and acting	Recognizing and using the similarity distance metaphor, using diagramming or graphing in problem-solving, recognizing clusters and patterns in data, separating change over space from change over time, knowing how, where, and when to use the various spatial-thinking strategies, and etc.
		Spatial capabilities	The ability to use supporting tools and technologies such as spreadsheet, graphical, statistical and GIS software to aid in problem-solving.
	Example 2	Concepts of space	Such as the relationships among units of measurement, different ways of calculating distance.
		Tools of representation	Such as the understanding of the relationships among views, and the effects of geographic projections.
		Processes of reasoning	Such as the ability to think about the shortest distances in different ways, and to make decisions based on spatial information.

The first relevant instrument is the Spatial Thinking Ability Test (STAT). [Lee and Bednarz \(2011\)](#) conducted a study focusing on the components of spatial thinking in the geoscience domain. In their study, they first reviewed the concepts of spatial thinking introduced in previous studies ([Gersmehl, 2005](#); [Janelle et al., 2009](#)). By reorganizing the core concepts of spatial thinking as introduced in prior studies, they developed the STAT, which contained 16 questions measuring 8 spatial thinking components relevant to geoscience. Examples of STAT questions and spatial thinking ability components present in [Fig. 1](#). The STAT covers spatial thinking components over a wide range ([Lee and Bednarz, 2011](#)). There are two forms of the STAT; they have the same types of question, and measure the same spatial thinking components, but the questions in the two forms are different. Since the two forms of the test are of equal difficulty, they can be used as pre- and post-test in measuring changes in students' spatial thinking ability ([Lee and Bednarz, 2011](#)).

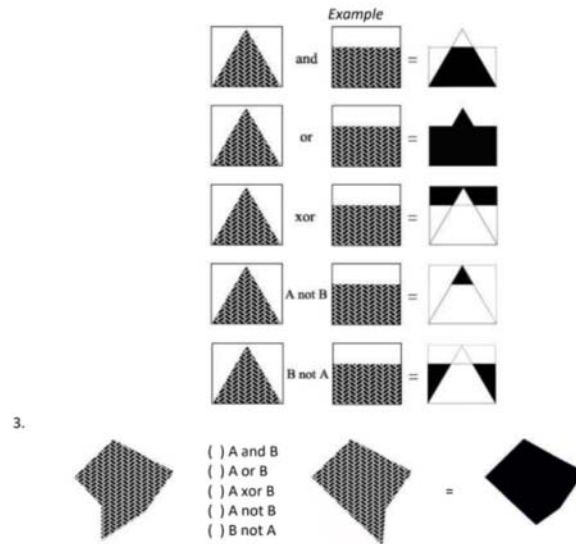
Besides the STAT, [Cohen and Hegarty \(2012\)](#) also designed a spatial thinking test called the Santa Barbara Solids Test (SBST). Rather than focusing on all spatial thinking concepts, the SBST only explores students' spatial skill of cross-section ([Cohen and Hegarty, 2012](#)). The test consists of 30 multiple choice questions; three levels of geometric structure and two types of cutting planes are distributed evenly across the 30 questions ([Cohen and Hegarty, 2012](#)). The results also indicated that "complexity of geometric object being sliced, and orientation of the cutting plane" are two difficult items for students.

A few other questionnaires on spatial thinking exist. However, they are not within the framework that [NRC \(2006\)](#) provides, such as [Dean and Morris \(2003\)](#), and [Hall-Wallace and McAuliffe \(2002\)](#). Other measurement instruments such as [Hegarty et al. \(2002\)](#) are more appropriate for classifying types of spatial behavior, rather than determining levels of student spatial ability ([Lee and Bednarz, 2011](#)).

Why should we consider spatial thinking in STEM education?

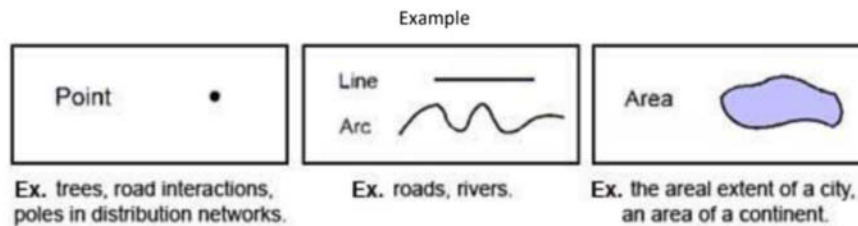
Prior studies provided evidence for the idea that spatial thinking is fundamental to the theoretical and practical underpinnings of geography (e.g., [Huynh and Sharpe, 2013](#); [Kim and Bednarz, 2013](#); [Favier et al., 2014](#); [Xiang and Liu, 2017](#)). The nature of the study of geography is to explore the patterns and processes of phenomena on the Earth's surface, at a variety of scales ([Huynh and Sharpe, 2013](#)). Due to the nature of geography, the exploration of the concept of earthly space, spatial thinking is the core ability at play during this process ([Huynh and Sharpe, 2013](#)). However, there is a common misunderstanding that spatial thinking is only important for the study of geography, when in actuality, it is equally important for other STEM subjects, such as science and technology

*DIRECTIONS: Solve the following questions based on the example below. Please mark (v) an answer.
Question #9 and #10 are adapted from Albert and Gollege (1999)*



Spatial Thinking Ability: visually verify a map overlay process and select the appropriate maps layers involved in the overlay (Lee & Bednarz, 2011)

*DIRECTIONS: Real world objects can be represented explicitly by point, line (arc), and area (polygon).
Based on the examples below, classify the followings spatial data.*



5. Locations of weather stations in Washington County _____.

- (A) Lines
- (B) Area
- (C) Points and Lines
- (D) Points and Area

6. Mississippi River channels and their basins _____.

- (A) Lines
- (B) Area
- (C) Points and Lines
- (D) Lines and Area

Spatial Thinking Ability: to be able to visually extract types of spatial data from verbally expressed spatial information (Lee & Bednarz, 2011)

Fig. 1 Examples of STAT questions and spatial thinking ability components.

(Uttal and Cohen, 2012; Wu and Shah, 2004). A comprehensive understanding of students' and teachers' spatial thinking ability, and suggestions on how to improve it, keeping in perspective a variety of STEM disciplines, is needed.

Specifically, research indicates that proficiency in spatial thinking is associated with success in engineering, architecture, medicine, mathematics, and the sciences, and students' spatial thinking ability is correlated with STEM achievement and attainment (e.g., Hegarty, 2010; Uttal et al., 2013; Uttal and Cohen, 2012; Wu and Shah, 2004; Stieff and Uttal, 2015). For example, Uttal et al. (2013) reviewed the relationship between STEM achievement and spatial thinking, and the results indicate that spatial training could improve STEM learning. Wai and Benbow (2010) conducted a 25-years longitudinal study and found that psychometrically-assessed spatial skills are a strong predictor of STEM attainment (Wai et al., 2009).

Moreover, in STEM disciplinary learning, a number of studies indicate the close relationship between STEM content and spatial thinking components (e.g., Hegarty et al., 2009; Uttal and Cohen, 2012; Sorby, 2011). For example, in chemistry, the behavior of isomers has a close relationship with spatial configurations (Uttal and Cohen, 2012); in biology, the understanding of cell structure can be enhanced through spatial visualization (Wu and Shah, 2004); and in physics, solving kinematics problems is also in relation to spatial visualization (Kozhevnikov et al., 2007).

Despite the importance of spatial thinking ability in teaching and learning within a multitude of fields of study, most studies of spatial thinking ability are confined to the subject of geography (Jo and Bednarz, 2014; Shin et al., 2016; Lee and Bednarz, 2009); limit research explored spatial thinking in other disciplines, such as science and technology (Bodzin, 2011; Sorby, 2009). Thus, more research regarding spatial thinking in STEM education field needs to be done in a broader set of disciplines.

High demand versus low performance in students' spatial thinking

The above literatures documented the close relationships between the need for spatial thinking ability and STEM education. Research also indicates that the STEM job market is facing a shortage of geospatially-trained workers (NRC, 2013). Therefore, it is also important that students' spatial thinking abilities are improved so that they are qualified for relevant future work. However, comparing the high demand of students' spatial thinking ability, their actual performance is relatively low, especially for students in K-12 settings. The status of students' current spatial thinking does not allow STEM educators to be optimistic.

Low performance in overall spatial thinking

Previous empirical studies evidenced that students' spatial thinking ability is still relatively low. For example, Bodzin (2011) examined secondary school students' spatial thinking abilities, by testing the three elements of spatial thinking addressed by the NRC (2006): concepts of space, tools of representation, and processes of reasoning. On average, students correctly answered 8 questions out of 18 total questions, amounting to less than a half percent rate of correctness. In addition, using the STAT, Tomaszewski et al. (2015) found that, on average, secondary school students only correctly answered 4 to 5 questions out of 16 total questions; identically, in Li (2020), students also correctly answered 4 to 5 questions with 28% accuracy. It is worth to notice that student participants from Li (2020) were only from urban public secondary schools, which may not represent students from all school settings, such as private schools, or suburban and rural schools. Tomaszewski et al. (2015) found that tested urban school students outperformed tested rural school students. In other words, rural school students might have even lower spatial thinking abilities than do urban school students. Thus, students' overall spatial thinking abilities must be improved, and further research on the status of students' spatial thinking abilities - across various school settings - is needed.

Low performance in spatial components that highly relate to STEM

In order to provide more precise and effective advice on improving students' spatial thinking ability, researchers suggest the importance of studying students' spatial thinking in terms of the different components (Lee and Bednarz, 2004; Li, 2020; Tomaszewski et al., 2015). For example, Tomaszewski and colleagues studied students' spatial thinking abilities according to different components, and their abilities varied. However, some students' low performed spatial thinking ability components are highly in relation to their STEM learning.

For example, Li (2020) found that secondary school students still lacked skills in comprehending overlay and dissolve, mentally visualizing 3-D images based on 2-D information, comprehending spatial shapes and patterns, and solving complex spatial questions with various components; they had better abilities in comprehending orientation and direction, discerning spatial patterns, and comprehending spatial association. In taking a closer look at spatial thinking ability components, it becomes clear that the components for which students exhibited a low ability are actually some of the principal components required in STEM interdisciplinary learning at the secondary level. For instance, transforming perceptions, representations, and images from one dimension to another, and the reverse, is one of the important spatial thinking abilities that is frequently used in students' science, math, and technology classes. To understand a certain concept or features of a subject, students are often required to transfer a 2D graph or picture to 3D, such as in learning about cellular structure in biology, molecular structure in chemistry, the earth's rotation and revolution in earth science, and solid geometry.

Furthermore, STEM researchers and educators also need to notice that during STEM disciplinary learning and everyday problem-solving situations, students usually encounter questions which require more than one spatial thinking ability component to solve. For example, in the STAT, students were asked to produce the image of a slope profile based on a topographic map, which requires the spatial thinking ability of recognizing spatial form, being able to transform perceptions, representations and images from one dimension to another and the reverse, and graphing a spatial transition. However, working on complex spatial thinking problems

requires strong spatial thinking abilities on higher spatial thinking components, which may be difficult for students with low spatial thinking abilities. Thus, STEM researchers should attend to these low ability components when studying students' spatial thinking abilities, and STEM teachers should take into account students' strengths and weaknesses in spatial thinking when designing and planning relevant lessons, especially when considering complex spatial problems.

Gender gap in spatial thinking

A number of empirical studies indicate that gender differences in spatial thinking ability exist - that male students had better spatial thinking ability than female students. Specifically, some researchers found that male students outperformed females in spatial thinking ability in secondary schools (Li, 2020; Tomaszewski et al., 2015), and others found a gender gap in spatial thinking ability among college students (Shin et al., 2016). The existing gender gap in spatial thinking might influence students' STEM learning, and that female students might lack the spatial skills to succeed in STEM (Li, 2020; Newcombe and Stieff, 2012). Thus, in order to provide students better STEM learning experience, STEM education researchers might investigate ways to narrow the spatial thinking gender gap, and teachers need to consider that some STEM instructional design efforts for imparting spatial ability skills may not be as applicable for female students.

In addition, STEM educators should also consider that a gender difference might exist when conducting an intervention (i.e., STEM activities) in developing students' spatial thinking abilities. For example, using geospatial technologies as a spatial thinking ability intervention, Crews (2008) found that spatial thinking abilities were improved for male students, but not for female students. It might be due to that a gender gap exists in the effectiveness of STEM activities. Thus, if we do not consider spatial thinking ability differences according to gender when conducting STEM activities, differences in spatial thinking ability and gender gap in STEM may be exacerbated. Future research should address gender differences when designing STEM activities centered around spatial thinking abilities.

Facilitating the growth of spatial thinking

Ideally, students' spatial thinking ability should increase along with their age to support their STEM learning, and relevant empirical studies reveal that age was a significant factor in predicting students' spatial thinking abilities (Newcombe and Stieff, 2012; Lee and Bednarz, 2004). For example, university students' spatial thinking abilities are greater than those of secondary school students (Lee and Bednarz, 2004). However, if we take a closer look at secondary school students' spatial thinking ability, researchers found the relationship between grade level and spatial thinking ability was not linear (Li, 2020; Tomaszewski et al., 2015). In other words, students' spatial thinking abilities did not always increase along with their grade levels. This might be a problem in students' STEM learning, since STEM learning in higher grades might require higher order spatial thinking, such as in taking STEM AP classes (Li, 2020).

As stated above, spatial thinking ability is an important and essential ability, not only for students' STEM learning in college level studies, but actually it is more important in students' secondary school study. Thus, we should ask whether middle and high school students are well-prepared for the spatial thinking abilities required for their college studies, and how to best support students' growth of spatial thinking through highly relevant STEM learning.

More importantly, very little attention has been drawn on teachers' growth of spatial thinking ability. In the most recent research, Li (2020) examined middle and high school STEM teachers' level of spatial thinking ability using the STAT. The results of students' and teachers' average percentage of correct answers per question on the STAT can be found in Fig. 2. The results indicate that teachers' and students' patterns of higher and lower spatial thinking ability components were similar. In other words, the ability components that are hard for students are also hard for teachers, such as "comprehending overlay and dissolve," "being able to transform perceptions, representations and images from one dimension to another and the reverse," and "overlaying and dissolving maps" (Gollidge, 2002; Gersmehl, 2005), which are the abilities that demand in STEM teaching and learning. Additionally, in Li's (2020) study, teacher participants were experienced STEM teachers; most had more than ten years of teaching experience. It is important to note that if experienced teachers are not adept at working on spatial problems that require these various ability components, it will be difficult for them to facilitate students' development of spatial thinking ability components. Thus, STEM teachers' spatial thinking abilities growth also need to be paid attention in order to better support the development of students' spatial thinking abilities.

How to cultivate spatial thinking in STEM education?

Although it is tempting to believe that cognitive abilities are fixed, there is copious evidence to support the supposition that spatial thinking ability can be improved through various learning activities (e.g., Terlicki et al., 2008; Wright et al., 2008; Uttal et al., 2013; Gold et al., 2018). Existing evidences in previously studies reveal that there are three effective ways to cultivate students' spatial thinking: (1) Through direct spatial thinking ability training (e.g., Cheng and Mix, 2013; Ehrlich et al., 2006; Wright et al., 2008); (2) Through learning the use of geo-spatial technologies (e.g., Kim and Bednarz, 2013; Milson and Earle, 2008; Madsen and Rump, 2012); (3) Through taking geo-spatial technology integrated disciplinary and interdisciplinary lessons (e.g., Bodzin, 2011; Gold et al., 2018). Newcombe and Stieff (2012) found that the differences among people's spatial thinking abilities may

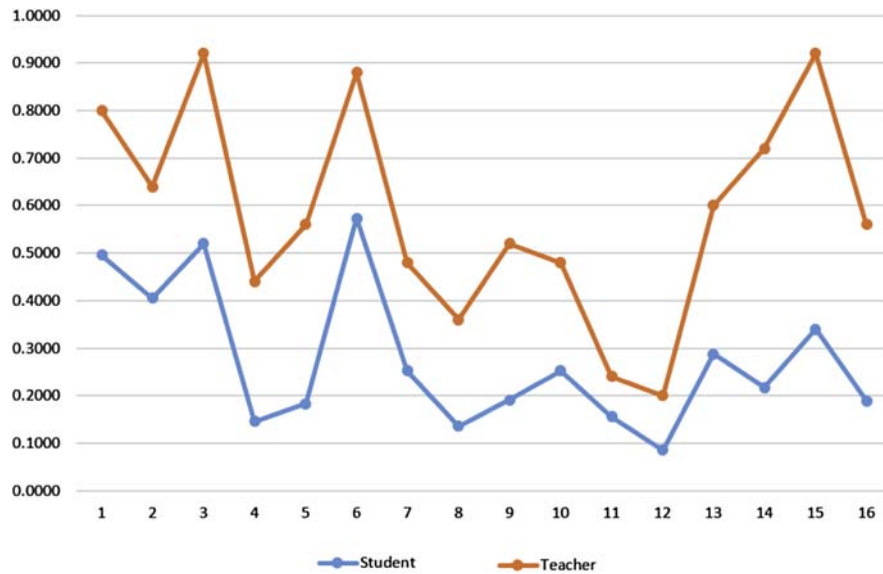


Fig. 2 Result from Li (2020), students' and teachers' average percentage of correct answers per question on the STAT. Source: Li, S., 2020. Effects of GIS Integration in Secondary School Classrooms: A Mixed-Method Study on Student Spatial Thinking Ability (Doctoral dissertation, State University of New York at Buffalo).

be due to differences in length and frequency of doing spatial thinking related activities, or the use of training techniques that are more efficacious, or both. In other words, it is important to have appropriate techniques, and intervention design, in order to improve spatial thinking ability. Since direct spatial thinking ability training is separate from STEM curriculum, only the use of geo-spatial technologies and taking geo-technology integrated STEM lessons are discussed in this section as effective ways to cultivate students' spatial thinking in STEM learning.

Geo-technology and spatial thinking

Geospatial technologies are technologies for the collecting or processing of data related to space and location; they include geographic information systems (GIS), remotely sensing satellites and aerial images, ArcExplorer, and etc. Those geospatial technologies allow individuals to examine the world through multiple layering of data within a spatial environment (Doering and Veletsianos, 2008). Geospatial technologies have become readily accessible, widely available, and more apparent in our daily lives than ever before (Bodzin and Cirucci, 2009), and they are starting to be used for teaching and learning within STEM classrooms.

Among numerous advanced technologies, Geographic Information System (GIS) is one of the most significant tools to have been called to use in educational reforms and policies. GIS is an integrated system of hardware, software, and procedures designed to support the collection, management, manipulation, analysis, modeling, and display of spatially referenced data about the Earth's surface, in order to solve complex planning and management problems (NRC, 2006). Most importantly, GIS is identified as a potential system for facilitating the development of students' spatial thinking ability (Jo et al., 2016; Lee and Bednarz, 2009; Bodzin, 2011). For the purpose of developing spatial thinking in STEM teaching and learning, NRC (2006) emphasized the significance of using GIS. NRC (2006) stated that "GIS has a clearly demonstrated potential as a support system for spatial thinking" (p. 221) and "must be brought to bear in fostering an understanding for and appreciation of spatial thinking" (p. 241). All students deserve, and need, the opportunity to be challenged, to be supported, and to become critical spatial thinkers, especially in STEM field.

Table 2 shows some sample GIS tools and related spatial thinking components identified by geo-spatial technology experts (Li, 2020). For example, Sketch Up is one of the GIS tools for a wide range of 3D modeling applications such as architectural, interior design, landscape architecture, civil and mechanical engineering. The use of Sketch Up might require the spatial thinking ability component of "being able to transform perceptions, representations and images from one dimension to another and the reverse" (Lee and Bednarz, 2011). Information in Table 2 helps us to figure out that although the GIS tools are not able to cover all the spatial thinking components identified in NRC's (2006) spatial thinking framework, the GIS tools are highly related to several important spatial thinking components. Thus, in order to cultivate students' spatial thinking ability, there is growing interest in studying the use of GIS in STEM disciplinary and interdisciplinary classroom (Bodzin, 2011; Lee and Bednarz, 2009; Li, 2020; Jo et al., 2016).

Table 2 Sample GIS tools and related spatial thinking components (Li, 2020).

Sample GIS tools	Spatial thinking components (Lee and Bednarz, 2011)
Google earth	• Comprehending orientation and direction • Discerning spatial patterns and graphing a spatial transition • Overlaying and dissolving maps • Representations and images from one dimension to another and the reverse • Comprehending integration of geographic features represented as points, networks, and regions • Comprehending spatial shapes and patterns
ArcGIS online and collector	• Discerning spatial patterns and graphing a spatial transition • Overlaying and dissolving maps • Comprehending integration of geographic features represented as points, networks, and regions • Comprehending spatial shapes and patterns
Esri story maps	• Comprehending orientation and direction • Comprehending integration of geographic features represented as points, networks, and regions • Comprehending spatial shapes and patterns
Sketch up	• Being able to transform perceptions, representations and images from one dimension to another and the reverse
Drones	• Recognizing spatial form • Comprehending orientation and direction • Representations and images from one dimension to another and the reverse

Source: Li, S., 2020. Effects of GIS Integration in Secondary School Classrooms: A Mixed-Method Study on Student Spatial Thinking Ability (Doctoral dissertation, State University of New York at Buffalo).

The effectiveness of GIS integration in STEM classes

A number of empirical studies have tested the idea of the NRC (2006) that learning to use GIS can positively impact students' spatial thinking ability (e.g., Lee and Bednarz, 2009; Madsen and Rump, 2012). For example, Lee and Bednarz (2009) studied GIS's effect on university students' spatial thinking ability. They conducted pre- and post-tests with 80 university students before and after taking GIS related courses, such as GIS Introduction and Computer Cartography. Non-GIS users served as the control group in the study. They found that hands-on lab- or project-based experience could improve students' spatial thinking, and greater improvement occurred in skills closely connected to, or supported by, students' experience with spatial representations through their completion of geo-technology course work (Lee and Bednarz, 2009).

Madsen and Rump (2012) also studied the use of GIS as an instrument for developing spatial thinking. The participants for this study were first-year geography major students who took GIS and Cartography courses, and the study focused on students' learning processes in using GIS (Madsen and Rump, 2012). The results of the study indicated that in the process of using GIS, students have different strategies for creating their personal instrument for spatial thinking; it is important to support students in the process of using GIS, to make GIS a personal instrument for the development of their spatial thinking (Madsen and Rump, 2012).

Kim and Bednarz (2013) focused on developing students' critical spatial thinking through the use of GIS. Similarly, they hypothesized that GIS would have a positive effect on students' spatial thinking (Kim and Bednarz, 2013). The participants of their study were upper-level undergraduate students, enrolled in introductory GIS and geography courses. The results indicate that GIS learning is beneficial in enhancing students' critical spatial thinking, identified as the ability to assess data reliability, to use sound spatial reasoning, and to evaluate problem-solving validity.

The above evidences on the effectiveness of using GIS in facilitating students' development of spatial thinking ability are in post-secondary school. Previously research also indicated the use of GIS can also cultivate middle and high school students' spatial thinking (Bodzin, 2011; Li, 2020). For example, Li (2020) examined whether taking GIS integration science, technology, and social study lessons facilitates secondary school students' development of spatial thinking ability. The results indicate that when controlling for students' background, students who used GIS during the academic year had significantly higher spatial thinking abilities than students who did not use GIS (Li, 2020). If we take a closer look at the results, for their control group students who did not use GIS, the difference in their spatial thinking ability between post- and pre-tests was very small. In other words, secondary school students might not improve in their spatial thinking abilities without intervention during the academic year. Thus, the study result not only adds evidence on the effectiveness of GIS use in facilitating students' development of spatial thinking, it also reminds STEM educators that it is important to support students' spatial thinking abilities at the stage of secondary learning, and GIS is a feasible choice when considering early exposure in improving students' spatial thinking abilities.

Approaching "best practice" in facilitating students' spatial thinking

NRC (2006) identifies a great challenge to the development of students' spatial thinking: How might current versions of GIS be incorporated into existing standards-based instruction in all knowledge domains across school curricula? This challenge has not been fully addressed. Most existing studies in GIS intervention were of direct GIS trainings, rather than integrating GIS within

the existing curriculum, which made the use of GIS limited to the subject of geography as opposed to in other STEM disciplines, such as science and technology. Only a few empirical studies support that it is possible to integrate GIS into existing standards-based secondary science, technology, and social studies curriculum, and it was effective in the improvement of students' spatial thinking abilities (Li, 2020). Therefore, when bringing GIS into students' learning, GIS curriculum integration is an effective intervention method that can broaden GIS's applications into wider STEM discipline use, which could better respond to NRC's (2006) call to incorporate GIS in all knowledge domains across the school curriculum. GIS STEM curriculum integration might be the "best practice" in facilitating students' spatial thinking ability.

Ensure enough frequency in taking GIS integrated STEM class

To teach effective GIS integrated STEM lessons which could facilitate students' spatial thinking ability, educators need to ensure sufficient frequency in students' GIS use.

Li (2020) conducted a quasi-experimental study testing the effectiveness of GIS STEM integration in improving students' spatial thinking ability. The results indicate that students who took more lessons with GIS integration tended to have higher gains in spatial thinking abilities, as compared with other students who took fewer lessons with GIS integration. In other words, frequency of taking GIS integration lessons could influence the effectiveness of GIS integration interventions on students' spatial thinking abilities.

For the most part, previous studies found that the outcomes of GIS use in facilitating students' learning were positive. That is, GIS use was positively correlated with gains in students' spatial thinking abilities, or their learning achievement (Bednarz and Lee, 2019; Bodzin, 2011; Kulo and Bodzin, 2013; Gold et al., 2018). For example, Hedley and colleagues (2013) indicated that the use of GIS had a positive effect on students' understanding of atmospheric concept knowledge, and Kulo and Bodzin (2013) reported a significant increase in the energy content knowledge through the use of geospatial technologies. However, other studies indicated an insignificant gain in students' learning outcomes or spatial thinking abilities based on GIS intervention. For example, Crews (2008) indicated that students did not undergo significant changes in their spatial thinking abilities when their teachers participated in geospatial technologies for a professional development project. Li's (2020) findings help with explaining the ineffectiveness of GIS interventions. That is, a longer intervention, with higher frequency of GIS use, may be required to make significant changes in spatial thinking outcomes.

The requirement of a higher frequency of GIS integration intervention lessons may be due to the fact that both students and teachers are inexperienced in using GIS. It could take time for students and teachers to adapt GIS use, especially when integrating GIS in STEM disciplinary teaching and learning. Once students and teachers become familiar and more skilled in GIS use, the experience of taking GIS integrated STEM lessons might then effectively facilitate students' spatial thinking abilities and learning outcomes. Thus, a certain amount of GIS use is important in conducting GIS intervention research, and requires further investigation.

Li (2020) also explored the threshold of GIS integration frequency in STEM class, which might be effective in improving students' spatial thinking abilities. The results present in Fig. 3, which shows students' average gain in spatial thinking ability by different GIS integration frequencies in STEM class. The results indicated that only taking three GIS integration lessons was not enough to make significant change in students' spatial thinking ability; taking 8 to 13 GIS integration STEM lessons per academic year might be necessary. However, the result is still tentative; the exact frequency number deserves future studies. So far, no other studies have investigated the effective frequency and added evidence on this important topic, STEM education researchers are suggested to examine the GIS intervention frequency that could make the most significant change in students' learning and cognitive development (Li, 2020).

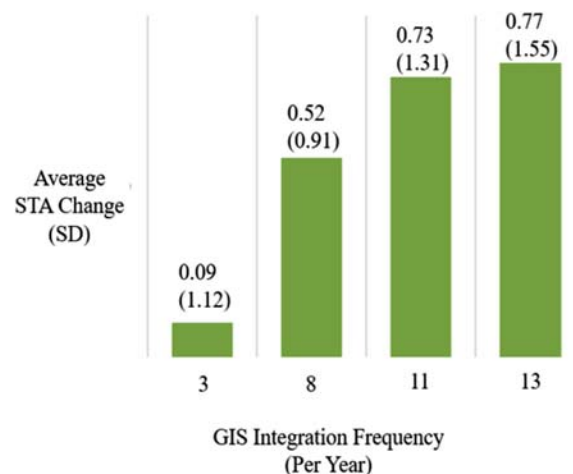


Fig. 3 Results from Li (2020), students' average gain in spatial thinking ability by different GIS integration frequencies in STEM class. Source: Li, S., 2021. Effects of GIS Integration in Secondary School Classrooms: A Mixed-Method Study on Student Spatial Thinking Ability (Doctoral dissertation, State University of New York at Buffalo).

Emphasize the role of STEM teacher

To make the use of GIS in STEM classroom more sustainable and realistic in order to facilitate students' development of spatial thinking ability, STEM teachers' role should be clarified and emphasized. In previous studies regarding GIS integration in STEM class, as teachers tend to be familiar with their curriculum but have very limited knowledge and skills of GIS use, STEM teachers' role were weakened in designing the integration lessons. For example, experts in GIS, instructional design and development, science education, environmental education, or educators in other disciplines have usually been involved in designing GIS integration lessons (Bodzin, 2011; Kulo and Bodzin, 2013; Gold et al., 2018); while the majority of teachers spent their time learning how to use specific GIS tools in order to teach lessons to their students. However, this mode of conducting GIS integration lacks sustainability, given that only a single STEM lesson, or set of lessons, within one curriculum topic were designed and taught to students (Bodzin, 2011; Kulo and Bodzin, 2013; Gold et al., 2018); teachers on their own may not be able to further extend the use of GIS in their STEM disciplinary teaching.

Based on the practice and design from prior research regarding GIS curriculum integration in STEM classes and by reviewing the literatures regarding teacher GIS professional development (PD), Li (2020) developed an updated mode of GIS application in STEM class. The mode in Li (2020) allows teacher to have more authority and freedom in designing and teaching GIS integration STEM lessons as compared with that found in previous studies, which made GIS use in the classroom more sustainable.

The study design of Li (2020) is presented in Fig. 4. In particular, a comprehensive GIS training with various GIS tools were conducted with teachers first during 3-week summer camp, which made teachers qualified in using different GIS tools in their classroom. Then, instead of relying on guidance from the researchers, as teachers had the knowledge and skills to use GIS, teachers were the leaders in designing GIS integration classes, with the freedom to choose their GIS tools and STEM class content. In contrast, researchers only facilitated teachers' work throughout the process. Under this design, each teacher designed several different GIS integration lessons, and the lessons covered various STEM disciplines and utilized different GIS tools. With GIS use abilities, teachers can design and teach more GIS integration classes without researchers', or other experts' support. In this way, GIS integration becomes sustainable, more integration lessons can be designed in various disciplines with class content, and more students are able to experience GIS use in their learning. Thus, conducting GIS teacher training first allows teachers to assume a central role in designing and teaching GIS integration lessons that are feasible in GIS interventions in STEM class.

The theory behind the above sustainable GIS STEM curriculum integration mode might also provide insight for STEM researchers in conducting future study. Specifically, Desimone (2009) proposed a core conceptual framework for studying the effects of PD's on teachers and students. This conceptual framework illustrated the path of how effective PD influence students' learning through changing teachers' knowledge, perceptions, and practice, which is presented at the top part of Fig. 5. Guided by Desimone's (2009) model, Li (2020) developed a framework on how GIS integration PD influences students' spatial thinking ability through teacher GIS practice as presented at the bottom part of Fig. 5. It shows how GIS integration PD would increase

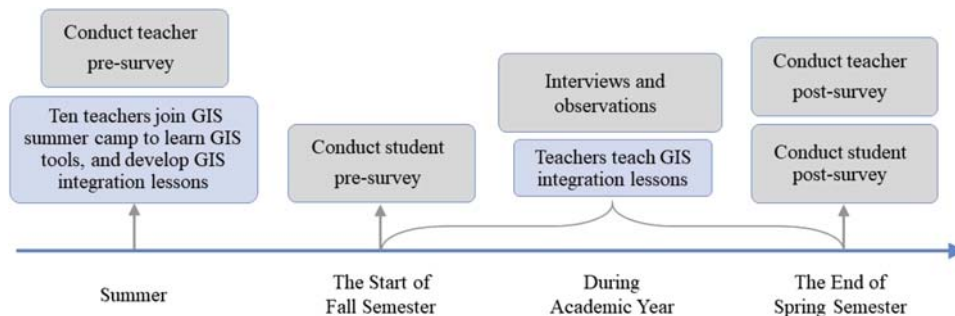


Fig. 4 Research design in Li (2020), the mode of teacher lead GIS curriculum integration STEM lesson design. Source: Li, S., 2022. Effects of GIS Integration in Secondary School Classrooms: A Mixed-Method Study on Student Spatial Thinking Ability (Doctoral dissertation, State University of New York at Buffalo).

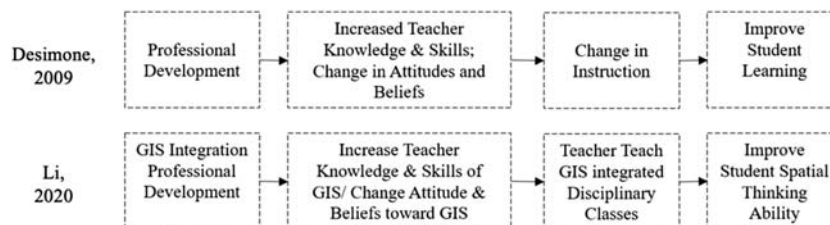


Fig. 5 Pd theory (Desimone, 2009) and the path of effective GIS PD in facilitating student spatial thinking ability (Li, 2020). Source: Li, S., 2023. Effects of GIS Integration in Secondary School Classrooms: A Mixed-Method Study on Student Spatial Thinking Ability (Doctoral dissertation, State University of New York at Buffalo).

teacher knowledge and skills of GIS use, and change their attitude and beliefs toward GIS, which in turn would lead to changes in their instruction, from teaching through traditional methods to teaching GIS integrated STEM lessons. Subsequently, after students taking GIS integrated STEM lessons and experience GIS use in the classroom, their spatial thinking ability could be improved (Li, 2020). Supported by PD theory, the effective paths from conducting STEM teacher GIS PD to the improvement of students' spatial thinking ability were identified in Li (2020). This mode is suggested for STEM researchers to further study in order to continually facilitate students' spatial thinking ability and upgrade the practice of GIS in STEM education.

Summary

Previously literatures have suggested a close relationship between spatial thinking and STEM education (e.g., Uttal and Cohen, 2012; Stieff and Uttal, 2015). Building on prior literatures, the current article elaborates the NRC's (2006) spatial thinking framework and provides several practical suggestions in facilitating students' spatial thinking ability by conducting geo-technology integrated STEM classes. STEM educators and researchers are encouraged to take active steps to ensure that students are well-prepared in their spatial thinking ability - for cultivating future STEM talents and for the improvement of STEM achievement and attainment among all students.

Several highlights of this article are listed below:

- (1) the NRC's (2006) spatial thinking definition reflects a wider variety of academic disciplines, which could be the major framework for relevant researchers;
- (2) Secondary school students' spatial thinking ability are still very low, and may influence students' capacity in STEM learning;
- (3) To have a deep understanding of students' spatial thinking and improve it, researchers need to study spatial thinking by different components and clarify its' relationships;
- (4) STEM teachers' have low spatial thinking ability on some components that are important in their STEM teaching; improving teachers' spatial thinking should also be paid attention to;
- (5) Integrating GIS in students' STEM disciplinary and interdisciplinary learning is evidenced to be an effective and practical way to facilitate students' spatial thinking ability;
- (6) In order to have effective GIS intervention in STEM learning to facilitate students' spatial thinking ability, sufficient frequency in GIS use needs to be ensured;
- (7) STEM teachers should have more autonomy in GIS curriculum integration to make GIS intervention sustainable and practical.

References

- Bednarz, R., Lee, J., 2019. What improves spatial thinking? evidence from the spatial thinking abilities test. *Int. Res. Geogr. Environ. Educ.* 28 (4), 262–280.
- Bodzin, A.M., Cirucci, L., 2009. Integrating geospatial technologies to examine urban land use change: a design partnership. *J. Geogr.* 108 (4–5).
- Bodzin, A.M., 2011. The implementation of a geospatial information technology (GIT) supported land use change curriculum with urban middle school learners to promote spatial thinking. *J. Res. Sci. Teach.* 48 (3), 281–300.
- Carroll, M., 1993. Deictic and intrinsic orientation in spatial descriptions: a comparison between English and German. In *Adv. Psychol.* 103, 23–44. North-Holland.
- Cheng, Y.-L., Mix, K.S., 2013. Spatial training improves children's mathematics ability. *J. Cogn. Dev.* 15 (1), 2–11.
- Cohen, C.A., Hegarty, M., 2012. Inferring cross sections of 3D objects: a new spatial thinking test. *Learn. Indiv. Differ.* 22 (6), 868–874.
- Crews, J.W., 2008. Impacts of a teacher geospatial technologies professional development project on student spatial literacy skills and interests in science and technology in grade 5–12 classrooms across Montana. University of Montana.
- Dean, G., Morris, P., 2003. The relationship between self-reports of imagery and spatial ability. *Br. J. Psychol.* 94, 245–273. <https://doi.org/10.1348/000712603321661912>.
- Desimone, L.M., 2009. Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educ. Res.* 38 (3), 181–199.
- Doering, A., Veletsianos, G., 2008. An investigation of the use of real-time, authentic geospatial data in the K-12 classroom. *J. Geogr.* 106 (6), 217–225.
- Ehrlich, S., Levine, S., Goldin-Meadow, S., 2006. The importance of gestures in children's spatial reasoning. *Dev. Psychol.* 42, 1259–1268.
- Eliot, J., Czarnolewski, M., 2007. Development of an everyday spatial behavioral questionnaire. *J. Gen. Psychol.* 134, 361–381. <https://doi.org/10.3200/GENP.134.3.361-381>.
- Favier, T.T., van der Schee, Joop, A., 2014. The effects of geography lessons with geospatial technologies on the development of high school students' relational thinking. *Comput. Educ.* 76, 225–236.
- Gersmehl, P.J., Gersmehl, C.A., 2007. Spatial thinking by young children: neurologic evidence for early development and "educability". *J. Geogr.* 106 (5), 181–191.
- Gersmehl, P., 2005. *Teaching Geography*. The Guilford Press, New York.
- Gold, A.U., Pendergast, P.M., Ormand, C.J., Budd, D.A., Mueller, K., 2018. Improving spatial thinking skills among undergraduate geology students through short online training exercises. *Int. J. Sci. Educ.* 40 (18), 2205–2225.
- Golledge, R.G., 2002. The nature of geographic knowledge. *Ann. Assoc. Am. Geogr.* 92 (1), 1–14.
- Golledge, R.G., 2008. Behavioral geography and the Theoretical/Quantitative revolution. *Geogr. Anal.* 40 (3), 23.
- Hall-Wallace, M.K., McAuliffe, C.M., 2002. Design, implementation, and evaluation of GIS-based learning materials in an introductory geoscience course. *J. Geosci. Educ.* 50 (1), 5–14.
- Hegarty, M., Keehner, M., Khooshabeh, P., Montello, D.R., 2009. How spatial ability enhances, and is enhanced by, dental education. *Learn. Indiv. Differ.* 19, 61–70.
- Hegarty, M., Richardson, A.E., Montello, D.R., Lovelace, K., Subbiah, I., 2002. Development of a self-report measure of environmental spatial ability. *Intelligence* 30 (5), 425–447.
- Hegarty, M., 2010. Components of spatial intelligence. In *Psychol. Learn. Motiv.* 52, 265–297. Academic Press.
- Huynh, N.T., Sharpe, B., 2013. An assessment instrument to measure geospatial thinking expertise. *J. Geogr.* 112 (1), 3–17.
- Janelle, D.G., Rebelo Hespanha, S., Goodchild, F., Goodchild, M.F., 2009. Workshops and national dissemination of geographic analysis in the social sciences: the CSISS experience in the USA. *J. Geogr. High Educ.* 33 (Suppl. 1), S88–S103.
- Jo, I., Bednarz, S.W., 2014. Dispositions toward teaching spatial thinking through geography: conceptualization and an exemplar assessment. *J. Geogr.* 113 (5), 198–207.
- Jo, I., Hong, J.E., Verma, K., 2016. Facilitating spatial thinking in world geography using web-based GIS. *J. Geogr. High Educ.* 40 (3), 442–459.

- Kim, M., Bednarz, R., 2013. Development of critical spatial thinking through GIS learning. *J. Geogr. High Educ.* 37 (3), 350–366.
- Kozhevnikov, M., Motes, M.A., Hegarty, M., 2007. Spatial visualization in physics problem solving. *Cogn. Sci.* 31, 549–579.
- Kulo, V., Bodzin, A., 2013. The impact of a geospatial technology-supported energy curriculum on middle school students' science achievement. *J. Sci. Technol. Educ. Res.* 22 (1), 25–36.
- Lee, J., Bednarz, R., 2004. GIS learning and spatial abilities: how GIS learning affects students' spatial abilities. In: Paper Presented at AAG Annual Conference, Philadelphia.
- Lee, J., Bednarz, R., 2009. Effect of GIS learning on spatial thinking. *J. Geogr. High Educ.* 33 (2), 183–198.
- Lee, J., Bednarz, R., 2011. Components of spatial thinking: evidence from a spatial thinking ability test. *J. Geogr.* 111 (1), 15–26.
- Li, S., 2020. Effects of GIS integration in secondary school classrooms: a mixed-method study on student spatial thinking ability. (Doctoral dissertation, State University of New York at Buffalo).
- Lohman, D., 1979. Spatial Ability: A Review and Reanalysis of the Correlational Literature. Office of Naval Research.
- Madsen, L.M., Rump, C., 2012. Considerations of how to study learning processes when students use GIS as an instrument for developing spatial thinking skills. *J. Geogr. High Educ.* 36 (1), 97–116.
- Milson, A.J., Earle, B.D., 2008. Internet-based GIS in an inductive learning environment: a case study of ninth-grade geography students. *J. Geogr.* 106 (6), 227–237.
- National Research Council of the National Academies, 2013. Future U.S. Workforce for Geospatial Intelligence. The National Academies Press.
- National Research Council, 2006. Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum. The National Academies Press, Washington, DC.
- Newcombe, N.S., Stieff, M., 2012. Six myths about spatial thinking. *Int. J. Sci. Educ.* 34 (6), 955–971.
- Schneider, J., McGrew, K., 2012. The Cattell–Horn–Carroll model of intelligence. In: Flanagan, D., Harrison, P. (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, third ed. Guilford Press, New York, pp. 99–144.
- Shin, E.E., Milson, A.J., Smith, T.J., 2016. Future teachers' spatial thinking skills and attitudes. *J. Geogr.* 115 (4), 139–146.
- Sorby, S.A., 2009. Educational research in developing 3-D spatial skills for engineering students. *Int. J. Sci. Educ.* 31 (3), 459–480.
- Sorby, S.A., 2011. *Developing Spatial Thinking*. Independence: Cengage.
- Stieff, M., Uttal, D., 2015. How much can spatial training improve STEM achievement? *Educ. Psychol. Rev.* 27 (4), 607–615.
- Terlicki, M., Newcombe, N.S., Little, M., 2008. Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Appl. Cogn. Psychol.* 22, 996–1013.
- Tomaszewski, B., Vodacek, A., Parody, R., Holt, N., 2015. Spatial thinking ability assessment in rwandan secondary schools: baseline results. *J. Geogr.* 114 (2), 39–48.
- Uttal, D.H., Cohen, C.A., 2012. 4 spatial thinking and STEM education: when, why, and how? *Psychol. Learn. Mot. Adv. Res. Theory* 57, 147.
- Uttal, D.H., Miller, D.I., Newcombe, N.S., 2013. Exploring and enhancing spatial thinking: links to achievement in science, technology, engineering, and mathematics? *Curr. Dir. Psychol. Sci.* 22 (5), 367–373.
- Wai, J., Lubinski, D., Benbow, C.P., 2009. Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101 (4), 817.
- Wai, L., Benbow, 2010. Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM Educational Dose: a 25-year longitudinal study. *J. Educ. Psychol.* 102, 860–871.
- Wright, R., Thompson, W.L., Ganis, G., Newcombe, N.S., Kosslyn, S.M., 2008. Training generalized spatial skills. *Psychon. Bull. Rev.* 15 (4), 763–771.
- Wu, H.K., Shah, P., 2004. Exploring visuospatial thinking in chemistry learning. *Sci. Educ.* 88 (3), 465–492.
- Xiang, X., Liu, Y., 2017. Understanding 'change' through spatial thinking using google earth in secondary geography: understanding 'change' using google earth. *J. Comput. Assist. Learn.* 33 (1), 65–78. <https://doi.org/10.1111/jcal.12166>.

Computational thinking

Chris Proctor, Department of Learning and Instruction, University at Buffalo, SUNY, Buffalo, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	88
History of computational thinking	88
Computational thinking in educational policy and practice	89
Assessment of computational thinking	90
Computational thinking as a research construct	91
How does computational thinking define the heart of computer science?	91
How does computational thinking describe the borderlands?	92
Three framings of computational thinking	93
References	93

Introduction

As the individual and social importance of computing has grown since the end of the second world war, computer science emerged as an academic field in higher education and more recently in primary and secondary education as well. Modern educational institutions require operational definitions of content knowledge as the basis for curriculum standards, graduation requirements, standardized assessments, teacher certification standards, and other forms of setting goals and measuring progress. Computational thinking has emerged as *boundary object* (Star and Griesemer, 1989) between academic computer science and K12 education, providing a mutually-intelligible construct to practitioners with different needs, different disciplinary practices, and different contexts of use. Computer scientists have tended to approach computational thinking from the inside, introspectively articulating the kind of thinking made possible by thinking-with-computers, which is uniquely characteristic of the field of computer science. Educators and education researchers often approach computational thinking from the outside, asking how these new skills and knowledge should be defined and assessed, and how they might fit into the broader scope of K12 learning goals. The ongoing controversy over how to define computational thinking emerges partly from issues of translation across these contexts and perspectives, but it can also be understood as a proxy for a political debate about the role computers ought to play in society and what children ought to learn in school.

This article provides a history of computational thinking, a summary of computational thinking's role in educational policy and practice, role as a research construct, and concludes by considering the future of computational thinking. Throughout, computational thinking is considered in terms of two opposing questions:

1. How does computational thinking define the heart of computer science?
2. How does computational thinking describe the borderlands?

History of computational thinking

The precursors to computational thinking predate the creation of functional computers. Ada Lovelace, credited with writing the first computer program in 1843, recognized that planning processes of symbolic manipulations (algorithms) represented a form of human thought distinct from the machinery which would carry them out. The machine, she wrote, "can do whatever we know how to order it to perform" (1843, p. 722), and imagined a future in which computers would be able to process symbolic representations of mathematics, language, and music. Although no computer capable of running symbolic programs would be built during her lifetime, Lovelace imagined a future in which new forms of thinking would be brought about by "more intimate practical acquaintance with the powers of the engine" (as cited in Fuegi and Francis (2003), p. 24). Lovelace's famous correspondence with Charles Babbage, who was more concerned with the mechanical details of building a computer, foreshadows the later distinction between computational thinking as a set of abstract knowledge, skills, strategies, and dispositions, and programming as an activity which concretely puts computational thinking to use.

The successful military use of computers in World War II led to speculation about how computers might also transform civilian life. Vannevar Bush, who led the United States's wartime research and development, anticipated that computers would help people find, remember, and organize the exponentially-growing amounts of information generated by society (1945). Perlis argued in the early 1960s that programming ought to be a standard part of a liberal education, enabling everyone to frame problems in computational terms (Guzdial, 2008). Engelbart's (1962) goal of using computers to augment human intellect led to the invention of personal computing and ultimately today's consumer computer ecosystem. As computers became smaller and less expensive,

educator-technologists such as Papert (1980) and Turkle (1984) explored how computers might offer powerful learning experiences to children.

Papert, the first to use the phrase “computational thinking” (1980, p. 182), was interested in using computers not “to put children through their paces” (1980, p. 19) but to serve as partners in thought providing contexts and representations through which children might form personal relationships with powerful ideas. He was dismayed at the tendency of school curricula to focus on the delivery of denatured and decontextualized “content,” similar to Freire’s (1968) critique of the “banking concept of education.” Writing about probability, for example, Papert argued that “one of the most powerful ideas in our intellectual heritage is (not untypically) disempowered in its school presentation, where it is reduced to shallow manipulations that seldom connect to anything the student experiences as important” (1996, p. 4). Papert argued that computers could serve as “micro-worlds” in which children might learn powerful ideas as easily and naturally as children learn French by growing up in France. Building off the work of Piaget and Vygotsky, Papert and later Constructionists hypothesized that computer cultures might produce new developmental trajectories for children and new futures for society.

The current era of computational thinking begins in 2006, when Wing’s argument for computational thinking as a “universally applicable attitude and skill set everyone, not just computer scientists, would be eager to learn and use” (p. 33) catalyzed legislatures, policymakers, researchers, and businesses around the goal of *CS for All*, or adding computer science to the fundamental competencies every child should develop throughout primary and secondary education. Wing defines computational thinking as involving “solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science” (p. 33). The article offers a long list of examples of computational thinking, starting with concepts such as recursion, parallel processing, type checking, and indirect addressing, which have specific technical meanings in computer science but which could be applied more generally if one frames a problem in computational terms. She encourages such abstraction, providing several examples of how concepts from computer science can be used to reinterpret situations from everyday life and envisioning a future in which computational ways of thinking about the world have become widespread and privileged: “Computational thinking will have become ingrained in everyone’s lives ... when trees are drawn upside down,” suggesting that tree-as-data-structure will have eclipsed the ecological concept of a tree (p. 34).

Wing offers a number of criteria for locating the boundaries of computational thinking which have been broadly influential in shaping the development of the concept in policy and research, as well as presciently indexing some of its key definitional tensions. First, she argues that computational thinking is “conceptualizing, not programming” (p. 35), an idea later interpreted by numerous standards documents to mean that computational thinking and programming are disjoint—a starker distinction than Wing may have intended. Furthermore, computational thinking is “fundamental, not a rote skill” (p. 35), anticipating the argument that computational thinking ought to be the emphasis in general and interdisciplinary education, leaving programming to the subset of students preparing for specialized careers in computer science and engineering. Wing suggests that computational thinking “complements and combines mathematical and engineering thinking” (p. 35), an assertion explored in subsequent literature on what is distinct about computational thinking and how it relates to other forms of disciplinary thinking. Finally, Wing argues that computational thinking is about “ideas, not artifacts” (p. 35), echoing prior arguments (e.g., Dijkstra, 1989) that computer programs ought to be taught as formal mathematical systems and not in terms of their behaviors.

Wing’s definition of computational thinking would form the basis of a rapid adoption of computational thinking as a primary and secondary learning goal around the world. As the concept was put into practice in educational policy and research, tensions and contradictions emerged around how computational thinking should be defined and how it should be taught. Computational thinking found its way into schools through standalone computer science courses as well as through interdisciplinary implementations across the curriculum, leaving researchers to reconcile narrow and broad conceptualizations.

Computational thinking in educational policy and practice

The goal of universal K12 computer science education was taken up by a coalition of public- and private-sector advocates of computing education such as the Association for Computing Machinery, Code.org, the Computer Science Teachers association (computer scienceTA), the International Society for Technology in Education (ISTE), the UK’s Computing at School, and the Australian Curriculum, Assessment, and Reporting Authority. These groups pursued a strategy of uniting behind a clear, uniform definition of the essential skills and knowledge of computer science, and they selected computational thinking as this definition despite unresolved details about the meaning and scope of the term, noted above (Barr and Stephenson, 2011). Computer science was introduced as a standalone course (particularly at the high school level) as well as an interdisciplinary literacy which could be taught across the curriculum; as a result, computational thinking was tasked with articulating both the essential heart of computer science as well as its interdisciplinary potential.

Major reports commissioned in the United States (Wilson et al., 2010) and the United Kingdom (The Royal Society, 2012) found that schools were far from preparing students with the computing skills they would need in the anticipated digital economy, and called on governments to create K12 computer science standards organized around computational thinking. Following initial standards developed by computer scienceTA (2011) and ISTE (2011) identifying computational thinking as a core strand of computer science, a broad coalition of stakeholders developed the K12 Computer Science Framework (2016) to guide individual states in defining their computer science content standards. This framework proposed seven core practices at the heart of computer science, and identified four of them as comprising computational thinking: recognizing and defining computational problems; developing

and using abstractions; creating computational artifacts; and testing and refining computational artifacts (p. 68). The framework largely followed Wing's earlier definition of computational thinking, emphasizing that computational thinking is distinct from computer science, that computer science is the best context for learning computational thinking, and that computational thinking should be included in standards for other subjects as well.

These efforts to define computational thinking laid the groundwork for computational thinking educational policy initiatives around the world. A 2019 review of global computational thinking educational policy initiatives (Hsu et al., 2019) found that five countries had implemented national computational thinking curricula: the United Kingdom, Australia, the United Arab Emirates, Finland, and Russia. Others, including India, Chile, Japan, China, Korea, Taiwan, and Hong Kong, have processes underway (Hsu et al., 2019; So et al., 2020). Hsu et al. (2019) found wide conceptual variety in how computational thinking is defined, as well as how computational thinking is understood to relate to computer science and to programming. National curricular reforms vary in the extent to which they implement computational thinking by adding computer science as a new subject area or embed interdisciplinary computational thinking across existing subject areas.

In addition to interdisciplinary applications, computational thinking has been used as a bridge between formal academic computer science and other contexts. Although many authors caution that the mere use of a computer (in or out of school) does not imply the presence of computational thinking, informal contexts such as clubs and online communities have been particularly important places for learners who have been marginalized by formal computer science to develop interest and experience which may serve as the foundations of computational thinking learning (National Academies of Sciences, Engineering, and Medicine, 2021). Informal learning environments have also been particularly welcoming to sociocultural reframings of computational thinking (Kafai and Burke, 2013; Reppenning et al., 2010), discussed in more detail below.

A lack of access to early computing experiences also plays a role in inequitable computer science participation (Margolis and Fisher, 2003), leading to increased emphasis on computational thinking learning opportunities earlier in high school (e.g., introductory courses such as Exploring Computer Science (Goode and Margolis, 2011)), the middle and primary grades, and even early childhood education. Wang and Proctor's (in press) review of computational thinking in early childhood education finds significant tensions between standard framings of computational thinking which prioritize interpreting phenomena in terms of computational models and the holistic developmental priorities of early childhood education. Instead, approaches grounded in Papert's understanding of computational thinking as a relationship between learners and computers has found more traction (Bers, 2017).

Assessment of computational thinking

Implementation of computational thinking standards and curricula is still quite recent so standardized assessments of computational thinking are generally not yet in place. As of 2010, "assessments for computer science education [were] virtually non-existent" (Wilson et al., 2010, p. 14). Assessment tends to drive implementation in K12 education (Eisenhart et al., 1996; Grover and Pea, 2013), so it is not surprising that the development of computational thinking assessments has been a policy priority as well as an active area of research. The development of assessments has also helped surface tensions within definitions of computational thinking, as an assessment requires precisely operationalizing a construct and resolving ambiguities in standards.

Brennan and Resnick's (2012) exploration of how computational thinking might be assessed remains by far the most cited framework for approaching the assessment of computational thinking. Brennan and Resnick identify three dimensions of computational thinking—concepts, practices, and perspectives—and consider several process-based strategies for assessing them, including studying students' design processes, artifact-based interviews, and design scenario tasks. Brennan and Resnick's focus on the process of artifact creation rather than the product is motivated by their primary interest in supporting youth in creating personally-meaningful artifacts, but their argument aligns with other arguments for attention to process. An artifact created with code does not constitute evidence of computational thinking without also considering the process with which it was made ("K-12 Computer Science Framework," 2016; Salac and Franklin, 2020). Similarly, Denning (2017) argues for understanding computational thinking as a competency characterized by how people think about problems (specifically, employing abstractions which represent a problem in terms of a computational model), not just what they know or what they have produced. This view presents a challenge to the decontextualized, content-based standardized testing characteristic of national policy initiatives.

Another difficulty in assessing computational thinking is that computational thinking appears to be tightly-bound to the modality in which computational thinking has been learned (Weintrop and Wilensky, 2015), reinforcing the importance of the relationship between cognition and the medium of representation (Wilensky, 2010). Despite the argument for the broad applicability of computational thinking, students' skills do not transfer easily between programming languages or representations, nor is there evidence that computational thinking improves intelligence in general (Guzdial, 2015; Pea and Kurland, 1984). To address this dependence on modality, Werner et al. (2012) developed a performance assessment embedded in the programming environment and task structure students had been using to learn computational thinking.

Several recent reviews conclude that substantial difficulties remain in the assessment of computational thinking. In terms of Brennan and Resnick's framework of concepts, practices, and perspectives, the vast majority of computational thinking assessments around the world focus on concepts, with little consensus on which concepts are included and how they are defined (Cutumisu et al., 2019). Most assessments are designed to test content knowledge without attending to process, and very few have been validated (Lye and Koh, 2014; McGill et al., 2019). Computational thinking assessment has made substantial recent progress in early childhood education, with efforts to stabilize the construct (Bers, 2017) and a validated assessment (Relkin et al., 2020).

Computational thinking as a research construct

As suggested by the uneven state of computational thinking assessment, the uptake of computational thinking in educational policy has outpaced its development as a research construct (Wilson and Guzdial, 2010). Since Wing's 2006 article, researchers have engaged in a spirited debate about what computational thinking ought to mean and how it ought to be defined, measured, and taught. However, based on the distribution of citations, this debate has had more urgency within the computer science education community than in the broader literature on education (Saqr et al., 2021). Empirical research on computational thinking often begins by recounting the history of these disagreements before proceeding with an ad-hoc measure of computational thinking. The two questions organizing this article, *How does computational thinking define the knowledge, skills, and practices at the heart of computer science?* And *How does computational thinking describe the borderlands?*, represent proximal and distal poles in a fundamental question underlying much of the controversy around computational thinking: how broad is computational thinking? (Proctor and Blikstein, 2018; Voogt et al., 2015).

How does computational thinking define the heart of computer science?

Between roughly 2006–2013, the academic conversation around computational thinking focused largely on how the construct articulated the heart of computer science, itself a dynamic and loosely-bounded discipline. In 2009 the US National Research Council convened a workshop aimed at defining the scope and nature of computational thinking. While the participants did not reach consensus on either topic, they did largely agree on the centrality of computational thinking in computer science, and that computational thinking is fundamentally a process of abstraction. In a subsequent article, Wing (2011) developed this idea into a concise definition of computational thinking:

Computational thinking is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can effectively be carried out by an information-processing agent (p. 1).

The workshop also considered the relationship between computational thinking and programming. Although examples of unplugged pedagogies (by which computational thinking is taught without the use of computers (Bell et al., 2009)) were shared, the workshop appeared to generally agree that while computational thinking is distinct from programming, code plays an important role in computational thinking. This is embedded in Wing (2011) definition, as programming is a primary (though not the only) mode of representing solutions which can be interpreted by a computer. In the words of Mitchel Resnick, “computational thinking is more than programming, but only in the same way that language literacy is more than writing ... programming, like writing, is a means of expression and an entry point for developing new ways of thinking” (p. 13).

These two findings—that computational thinking describes a thought process of applying abstractions to problems, and that programming plays an important role in computational thinking—are notable in retrospect for the extent to which they have been minimized in the implementation of computational thinking in policy and practice. In the decade following Wing's 2006 article, numerous frameworks proposed taxonomies of what should be included within computational thinking, with substantial variation (Angeli et al., 2016; Barr and Stephenson, 2011; Brennan and Resnick, 2012; Grover and Pea, 2013; Kalelioglu et al., 2016; Repenning et al., 2016; Seiter and Foreman, 2013; Selby, 2012). Abstraction (the “undisputed” keystone of computational thinking (Grover and Pea, 2013, p. 39)) was always included as a core element of computational thinking; algorithms, data, decomposition, parallelization, testing and debugging, and control structures were often included as well (Zhang and Nouri, 2019). A recent review proposed six facets of computational thinking: decomposition, abstraction, algorithm design, debugging, iteration, and generalization (Shute et al., 2017). Nevertheless, K12 computational thinking curricula and assessments disproportionately address computational thinking concepts, suggesting that computational thinking is taught as a body of knowledge rather than a process (Lye and Koh, 2014).

Recent implementations of computational thinking have also de-emphasized programming for principled and for practical reasons. In contrast to Resnick's quotation above suggesting programming is like writing, Lu and Fletcher (2009) view programming as a specialized practice, unnecessary and uninviting for beginners: “programming is to Computer Science what proof construction is to mathematics, and what literary analysis is to English” (p. 260). Others have suggested that teaching programming focuses students' attention on low-level details which are becoming less important as interaction with computers progresses to higher-level abstractions (Pears et al., 2021). In primary and secondary education, teaching computational thinking without programming has had an easier time finding purchase in interdisciplinary contexts and as the basis for pedagogical approaches which are manageable for teachers without a background in computer science.

As computational thinking gained widespread adoption in K12 education globally, concerns emerged from within the academic computer science community that the desire for broad applicability had led to an overly-vague definition of computational thinking which had lost its focus on the essential process of framing problems in computational terms (Tedre and Denning, 2022). Denning (2017) envisions computational thinking defined in terms of competencies—skills and sensibilities demonstrated in practice—while noting that assessments of computational thinking are largely focused on static content knowledge. He takes issue with curricula based on these assessments structured as “progressions of increasingly sophisticated learning objectives” rather than sequences

of skill acquisition (p. 36). Even if programming is not necessary in computational thinking, Aho (2012) emphasizes the centrality of computational models (formal systems which can run automatically) in computational thinking, arguing that interdisciplinary computational thinking requires a discipline's constructs discipline to be represented in terms of a computational model. Computational models are widely-accepted in adjacent STEM fields such as the life sciences, but in the social sciences and particularly the humanities, whose questions more often address interpretation, positionality, and subjectivity, the role of computational models is less clear.

Grover and Pea's (2013) review of computational thinking in *Educational Researcher* marks a milestone in the research trajectory of computational thinking, taking stock of progress and projecting the research agenda ahead, while also mediating the perspectives of the computer science education community and the broader community of educational research. After summarizing the prior research on computational thinking (much of which, they note, was focused on defining computational thinking), Grover and Pea proposed broadening the scope of computational thinking research by focusing on interest, identity, and stereotypes contributing to a lack of diversity in computational thinking enrollment, computational thinking research grounded in sociocultural and situated theories of learning, and interdisciplinary computational thinking. At the same time, they called for more empirical research on K12 computational thinking practice, including the development of evidence-based curricula, pedagogies, and tools for K12 computer science and models for integrating computational thinking into professional development and teacher preparation. While Grover and Pea's review does not demarcate a clean division, it marked a shift in emphasis from computational thinking as it relates to computer science to computational thinking as it relates to K12 education more broadly.

How does computational thinking describe the borderlands?

The term *borderlands*, referencing Anzaldúa (1987), points to liminal spaces and the possibilities of hybridity, crossing, or syncretism drawing on multiple sources. The borderlands of computational thinking are at the edge of what counts from the perspective of computer science, sometimes crossing into other disciplines or communities of practice with their own perspectives and priorities. The borderlands include teachers and students in schools and informal contexts (who may be more interested in creating meaningful experiences than fidelity of implementation), those oppressed by race, gender, or other categories (who may be indifferent or hostile to their powerful computing practices being identified as computational thinking), computational thinking in interdisciplinary contexts (where computational thinking may play an instrumental role in the service). As computational thinking has made its way into K12 practice over the last decade, the community of researchers interested in computational thinking has expanded to include those whose primary concern is these spaces and peoples.

Following the 2009 National Research Council workshop on the scope and nature of computational thinking, another workshop was convened to consider the pedagogical aspects of computational thinking (National Research Council, 2011). Although K12 teachers were not represented, the workshop affirmed the importance of teachers and pedagogical skill, and recognized the potential for computational thinking in formal primary and secondary education, as well as informal spaces. As the visibility of computer science practitioners grew within the computing education research world (led by the Computer Science Teachers Association), an alternative emerged to the theoretical debate over how to define computational thinking. The Use-Modify-Create framework (Lee et al., 2011), in which a group of researchers based in K12 schools and informal contexts chose to focus not on what computational thinking means in theory, but what it looks like in practice. The report aimed to help educators understand and teach computational thinking in the context of children's socioeconomic and cultural contexts and their developmental trajectories. In the framework, students first *use* computational artifacts created by others in order to understand their external functionality, then *modify* artifacts to understand their inner workings, and finally *create* their own artifacts. This strategy grounds computational thinking in a sociocultural account of learning, aiming to support students' inbound trajectories of participation and developing a sense of authorship and agency with the artifacts and the code in which they are implemented. The epistemic importance of teachers has only grown, with recent calls to substitute the top-down, "abstract and agenda-driven" processes for defining computational thinking and designing pedagogical approach with emergent definitions grounded in local needs and realities (Wilkerson et al., 2020, p. 265).

A substantial body of literature has documented the persistent inequities present in K12 computer science education. Equity-oriented researchers studying how mainstream computer science learning environments tend to support dominant-culture youth at the expense of their marginalized peers have found it useful to frame learning in terms of sociocultural constructs such as identity-building (Shaw and Kafai, 2020; Vakil, 2020) and participation in communities of practice (Scott and White, 2013), and to examine the relationship between computing and structures such as race and gender (Benjamin, 2019; Margolis and Fisher, 2003). However, these sociocultural and critical perspectives are substantially less present in the literature on computational thinking. The K12 Computer Science Framework (2016) recognizes the importance of practices such as collaboration, communicating, and fostering an inclusive computing culture, but clearly demarcates them from computational thinking. Similarly, the AP Computer Science Principles course, which was designed around computational thinking, views creativity and collaboration as "pedagogical strategies to be used to develop a diverse, appealing, and inclusive classroom environment" rather than as part of computational thinking (College Board, 2020). In empirical research, race, gender, and other categories are more often treated as contextual factors or moderating variables, rather than as formations shaping the nature of activity.

If computational thinking is essentially about the process of formulating problems so that their solutions can be carried out by an information-processing agent, sociocultural and critical constructs are relevant to who is formulating which problems, and to what end. Vakil's (2018) agenda for a justice-centered approach to equity in computer science education would emphasize the

social and political implications of computational thinking, and the disciplinary, civic, and political identities which authorize students and legitimize problems as suitable for computational thinking. Similarly, ethnocomputing applies the processes of computational thinking to the “knowledges, practices, and designs of Indigenous or vernacular cultural systems” (Lachney et al., 2022, p. 112). Several alternatives to computational thinking have also been proposed, such as *computational participation* (Kafai and Burke, 2013) *computational action* (Tissenbaum et al., 2019), and *computational literacies* (Kafai and Proctor, 2021) which view constructs such as identity, interest, power, and formations (Omi and Winant, 2001) such as race and gender as integral to the practices of computational thinking.

Finally, interdisciplinary borderlands of computational thinking have been receiving increasing attention at the K12 level, following the increasing importance of computing in research across disciplines (National Research Council, 2010). Computational thinking has been deployed in an effort to “highlight the integral relationship between computing and other fields” (Wilkerson et al., 2020, p. 264). K12 STEM disciplines have eagerly adopted computational thinking. For example, the Next Generation Science Standards include computational thinking among the core practices of science (Weintrop et al., 2016). Domain-specific “teaspoon languages” have recently been developed which “add a teaspoon of computing to other subjects.” The goal is that by the time that US students take a computer science class (typically, in high school or undergraduate), they will have had many programming experiences, have seen a variety of types of programming languages, and have a sense that “programming isn’t hard” (Guzdial and Shreiner, 2022).

Three framings of computational thinking

The two questions organizing this article are positioned respectively within computer science and within K12 education. They represent different stakeholders in the research community, and are oriented toward different future research trajectories, raising the prospect of fragmentation of the computing education research community. In response to a proposal that computing education ought to avoid *theory bias* by limiting its engagement with educational theory (Nelson and Ko, 2018), Kafai et al. (2020) argued that all accounts of learning are grounded in theoretical frames and instead proposed a commitment to *theory dialog*. Their goal was to avoid the so-called “literacy wars” and “math wars” in which researchers committed to different paradigms were unable to find common ground with respect to disciplinary learning goals, what would constitute evidence of success, or sometimes even legitimate methods of research.

Toward this end, Kafai, Proctor, and Lui identified three framings of computational thinking: cognitive, situated, and critical, which distinguish the theoretical underpinnings of various projects within the computer science education research community. Cognitive computational thinking seeks to provide students with an understanding of key computational concepts and competencies. Situated computational thinking understands learning in terms of creative expression, interest development, and identity authorship and participation. Critical computational thinking recognizes that computing is not an unequivocal social good, and proposes an analytical approach to the values, practices, and infrastructure underlying computation as part of a broader goal of education for justice and participation in shaping our shared future.

Kafai, Proctor, and Lui represented these framings as three concentric circles, suggesting units of analysis at the scale of the individual, the collaborative group, and the city or society. These three framings are loosely parallel to the concepts, practices, and perspectives proposed by Brennan and Resnick (2012). Just as the vast majority of computational thinking assessments currently focus exclusively on concepts, cognitive framings of computational thinking remain dominant. As the social importance of computing continues to grow and educational stakeholders from outside of computing education participate in the academic conversation around computational thinking, the authors hope that situated and critical research will be legible to researchers whose work is largely cognitive and dialog across framings will be possible.

The first decades of the 21st century have seen widespread adoption of computational thinking as an educational priority around the world, accompanied by a research base whose growth continues to accelerate (Saqr et al., 2021). As the social importance of computing continues to grow, computational thinking appears to be well on its way toward becoming a fundamental literacy at the heart of K12 education. With this integration the two questions anchoring this article, asking how computational thinking defines the heart of computer science and how computational thinking describes the borderlands, are likely to be increasingly in dialog with each other.

References

- Aho, A.V., 2012. Computation and computational thinking. *Comput. J.* 55, 832–835. <https://doi.org/10.1093/comjnl/bxs074>.
- Angeli, C., Voogt, J., Fluck, A., Webb, M., Cox, M., Malyn-Smith, J., Zagami, J., 2016. A K-6 computational thinking curriculum framework: implications for teacher knowledge. *J. Educ. Technol. Soc.* 19, 12.
- Anzaldúa, G., 1987. *Borderlands: La frontera*. Aunt Lute, San Francisco.
- Barr, V., Stephenson, C., 2011. Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *Acm Inroads* 2, 48–54.
- Bell, T., Alexander, J., Freeman, I., Grimley, M., 2009. *Computer Science Unplugged: School Students Doing Real Computing Without Computers*, vol. 10.
- Benjamin, R., 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Social Forces.
- Bers, M.U., 2017. *Coding as a Playground: Programming and Computational Thinking in the Early Childhood Classroom*. Routledge.

- Brennan, K., Resnick, M., 2012. New frameworks for studying and assessing the development of computational thinking. In: Proceedings of the 2012 Annual Meeting of the American Educational Research Association. Vancouver, Canada, p. 25.
- Bush, V., 1945. As we may think. In: *The Atlantic Monthly*, vol. 176, pp. 101–108.
- College Board, 2020. AP Computer Science Principles Course and Exam Description.
- Cutumisu, M., Adams, C., Lu, C., 2019. A scoping review of empirical research on recent computational thinking assessments. *J. Sci. Educ. Technol.* 28, 651–676. <https://doi.org/10.1007/s10956-019-09799-3>.
- Denning, P.J., 2017. Remaining trouble spots with computational thinking. *Commun. ACM* 60, 33–39. <https://doi.org/10.1145/2998438>.
- Dijkstra, E.W., 1989. On the cruelty of really teaching computing science. *Commun. ACM* 32, 1398–1404.
- Eisenhart, M., Finkel, E., Marion, S.F., 1996. Creating the conditions for scientific literacy: a re-examination. *Am. Educ. Res. J.* 33, 261–295.
- Engelbart, D.C., 1962. Augmenting Human Intellect: A Conceptual Framework, pp. 64–90.
- Freire, P., 1968. *Pedagogy of the Oppressed*.
- Freige, J., Francis, J., 2003. Lovelace & Babbage and the creation of the 1843 “Notes”. *IEEE Ann. Hist. Comput.* 16–26.
- Goode, J., Margolis, J., 2011. Exploring computer science: a case study of school reform. *J. Educ. Resour. Comput.* 11, 1–16. <https://doi.org/10.1145/1993069.1993076>.
- Grover, S., Pea, R., 2013. Computational thinking in K–12: a review of the state of the field. *Educ. Res.* 42, 38–43. <https://doi.org/10.3102/0013189X12463051>.
- Guzdial, M., Shreiner, T., 2022. Integrating computing through task-specific programming for disciplinary relevance: considerations and examples. In: Yadav, A., Berthelsen, U.D. (Eds.), *Computational Thinking in Education: A Pedagogical Perspective*, pp. 172–190.
- Guzdial, M., 2008. Education paving the way for computational thinking. *Commun. ACM* 51, 25. <https://doi.org/10.1145/1378704.1378713>.
- Guzdial, M., 2015. Learner-centered design of computing education: research on computing for everyone. *Synth. Lect. Hum. Centered Inf.* 8, 1–165. <https://doi.org/10.2200/S00684ED1V01Y201511HCI033>.
- Hsu, Y.-C., Irie, N.R., Ching, Y.-H., 2019. Computational Thinking Educational Policy Initiatives (CTEPI) Across the Globe. *TechTrends*. <https://doi.org/10.1007/s11528-019-00384-4>.
- ISTE, 2011. *ISTE Standards for Computer Science Educators*.
- K-12 Computer Science Framework, 2016.
- Kafai, Y.B., Burke, Q., 2013. The social turn in K-12 programming: moving from computational thinking to computational participation. In: *Proceeding of the 44th ACM Technical Symposium on Computer Science Education—SIGCSE '13*. ACM Press, Denver, Colorado, USA, p. 603. <https://doi.org/10.1145/2445196.2445373>.
- Kafai, Y.B., Proctor, C., 2021. A reevaluation of computational thinking in K-12 education: moving towards computational literacies. *Educ. Res.* <https://doi.org/10.3102/0013189X211057904>.
- Kafai, Y.B., Proctor, C., Lui, D., 2020. From theory bias to theory dialogue: embracing cognitive, situated, and critical framings of computational thinking in K-12 CS education. *ACM Inroads* 11, 44–53.
- Kalelioglu, F., Gülbahar, Y., Kukul, V., 2016. A framework for computational thinking based on a systematic research review. *Baltic J. Modern Comput.* 4, 14.
- Lachney, M., Green, B., Allen, M.C., Foy, L., 2022. Ethnocomputing and computational thinking. In: Yadav, A., Berthelsen, U.D. (Eds.), *Computational Thinking: A Pedagogical Perspective*. Routledge, pp. 112–135.
- Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., Werner, L., 2011. Computational thinking for youth in practice. *ACM Inroads* 2, 32–37. <https://doi.org/10.1145/1929887.1929902>.
- Lovelace, A., 1843. Notes on I. Menabrea's “Sketch of the analytical engine invented by Charles Babbage, Esq”. *Taylor Sci. Memoirs* 3, 1843.
- Lu, J.J., Fletcher, G.H., 2009. Thinking about computational thinking. In: *Proceedings of the 40th ACM Technical Symposium on Computer Science Education*, pp. 260–264.
- Lye, S.Y., Koh, J.H.L., 2014. Review on teaching and learning of computational thinking through programming: what is next for K-12? *Comput. Hum. Behav.* 41, 51–61. <https://doi.org/10.1016/j.chb.2014.09.012>.
- Margolis, J., Fisher, A., 2003. *Unlocking the Clubhouse: Women in Computing*. MIT Press.
- McGill, M.M., Decker, A., McKlin, T., Haynie, K., 2019. A gap analysis of noncognitive constructs in evaluation instruments designed for computing education. In: *Proceedings of the 50th ACM Technical Symposium on Computer Science Education—SIGCSE '19*. ACM Press, Minneapolis, MN, USA, pp. 706–712. <https://doi.org/10.1145/3287324.3287362>.
- National Academies of Sciences, Engineering, and Medicine, 2021. *Cultivating Interest and Competencies in Computing: Authentic Experiences and Design Factors*. National Academies Press, Washington, D.C. <https://doi.org/10.17226/25912>.
- National Research Council, 2010. *Report of a Workshop on the Scope and Nature of Computational Thinking*. National Academies Press.
- National Research Council, 2011. *Report of a Workshop on the Pedagogical Aspects of Computational Thinking*. National Academies Press.
- Nelson, G.L., Ko, A., 2018. On use of theory in computing education research. In: *Proceedings of the 2018 ACM Conference on International Computing Education Research—ICER '18*. ACM Press, Espoo, Finland, pp. 31–39. <https://doi.org/10.1145/3230977.3230992>.
- Omi, M., Winant, H., 2001. Racial formation. In: Essed, P., Goldberg, D.T. (Eds.), *Race Critical Theories: Text and Context*. Blackwell, Malden, MA, pp. 123–145.
- Papert, S., 1980. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, Inc.
- Papert, S., 1996. An exploration in the space of mathematics educations. *Int. J. Comput. Math. Learn.* 1. <https://doi.org/10.1007/BF00191473>.
- Pea, R.D., Kurland, D.M., 1984. On the cognitive effects of learning computer programming. *New Ideas Psychol.* 2, 137–168. [https://doi.org/10.1016/0732-118X\(84\)90018-7](https://doi.org/10.1016/0732-118X(84)90018-7).
- Pears, A., Tedre, M., Valtonen, T., Vartiainen, H., 2021. What Makes Computational Thinking So Troublesome? <https://doi.org/10.13140/RG.2.2.20480.35842>.
- Proctor, C., Blikstein, P., 2018. How broad is computational thinking? A longitudinal study of practices shaping learning in computer science. In: Kay, J., Luckin, R. (Eds.), *Rethinking Learning in the Digital Age. Making the Learning Sciences Count*. International Society of the Learning Sciences, London, UK, pp. 544–551.
- Relkin, E., de Ruiter, L., Bers, M.U., 2020. TechCheck: development and validation of an unplugged assessment of computational thinking in early childhood education. *J. Sci. Educ. Technol.* 29, 482–498. <https://doi.org/10.1007/s10956-020-09831-x>.
- Repenning, A., Webb, D., Ioannidou, A., 2010. Scalable game design and the development of a checklist for getting computational thinking into public schools. In: *Proceedings of the 41st ACM Technical Symposium on Computer Science Education, SIGCSE '10*. Association for Computing Machinery, New York, NY, USA, pp. 265–269. <https://doi.org/10.1145/1734263.1734357>.
- Repenning, A., Basawapatna, A., Escherle, N., 2016. Computational thinking tools. In: *2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, Cambridge, United Kingdom, pp. 218–222. <https://doi.org/10.1109/VLHCC.2016.7739688>.
- Salac, J., Franklin, D., 2020. If they build it, will they understand it? Exploring the relationship between student code and performance. In: *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*. ACM, Trondheim Norway, pp. 473–479. <https://doi.org/10.1145/3341525.3387379>.
- Saqr, M., Ng, K., Oyeler, S., Tedre, M., 2021. People, ideas, milestones: a scientometric study of computational thinking. *J. Educ. Resour. Comput.* 21, 1–17. <https://doi.org/10.1145/3445984>.
- Scott, K.A., White, M.A., 2013. COMPUGIRLS' standpoint: culturally responsive computing and its effect on girls of color. *Urban Educ.* 48, 657–681. <https://doi.org/10.1177/0042085913491219>.
- Seiter, L., Foreman, B., 2013. Modeling the Learning Progressions of Computational Thinking of Primary Grade Students. ACM Press, p. 59. <https://doi.org/10.1145/2493394.2493403>.
- Selby, C.C., 2012. Promoting computational thinking with programming. In: *Proceedings of the 7th Workshop in Primary and Secondary Computing Education on—WiPSC '12*. ACM Press, Hamburg, Germany, p. 74. <https://doi.org/10.1145/2481449.2481466>.
- Shaw, M., Kafai, Y., 2020. Charting the identity turn in K-12 computer science education: developing more inclusive learning pathways for identity. In: *International Conference of the Learning Sciences*. Nashville, US (Conference Cancelled).

- Shute, V.J., Sun, C., Asbell-Clarke, J., 2017. Demystifying computational thinking. *Educ. Res. Rev.* 22, 142–158. <https://doi.org/10.1016/j.edurev.2017.09.003>.
- So, H.-J., Jong, M.S.-Y., Liu, C.-C., 2020. Computational thinking education in the Asian Pacific region. *Asia Pac. Educ. Res.* 29, 1–8. <https://doi.org/10.1007/s40299-019-00494-w>.
- Star, S.L., Griesemer, J.R., 1989. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's museum of Vertebrate Zoology, 1907–39. *Soc. Stud. Sci.* 19, 387–420.
- Tedre, M., Denning, P., 2022. Computational thinking: a professional and historical perspective. In: Yadav, A., Berthelsen, U.D. (Eds.), *Computational Thinking: A Pedagogical Perspective*. Routledge, pp. 1–17.
- The Royal Society, 2012. Shut Down or Restart? The Way Forward for Computing in UK Schools. The Royal Society.
- Tissenbaum, M., Sheldon, J., Abelson, H., 2019. From computational thinking to computational action. *Commun. ACM* 62, 34–36. <https://doi.org/10.1145/3265747>.
- Turkle, S., 1984. *The Second Self: Computers and the Human Spirit*, 20th anniversary ed., first MIT Press ed. MIT Press, Cambridge, Mass.
- Vakil, S., 2018. Ethics, identity, and political vision: toward a justice-centered approach to equity in computer science education. *Harv. Educ. Rev.* 88, 26–52. <https://doi.org/10.17763/1943-5045-88.1.26>.
- Vakil, S., 2020. "I've always been scared that someday I'm going to sell out": exploring the relationship between political identity and learning in computer science education. *Cognit. Instruct.* 38, 87–115. <https://doi.org/10.1080/07370008.2020.1730374>.
- Voogt, J., Fisser, P., Good, J., Mishra, P., Yadav, A., 2015. Computational thinking in compulsory education: towards an agenda for research and practice. *Educ. Inf. Technol.* 20, 715–728. <https://doi.org/10.1007/s10639-015-9412-6>.
- Wang, X.C., Proctor, C., (in press). Computational thinking (CT) meets young children: critical review of conceptualization and implementation of CT in early childhood. *Early Child. Res. Q.*
- Weintrop, D., Wilensky, U., 2015. Using commutative assessments to compare conceptual understanding in blocks-based and text-based programs. In: *Proceedings of the Eleventh Annual International Conference on International Computing Education Research, ICER '15*. ACM, New York, NY, USA, pp. 101–110. <https://doi.org/10.1145/2787622.2787721>.
- Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., Wilensky, U., 2016. Defining computational thinking for mathematics and science classrooms. *J. Sci. Educ. Technol.* 25, 127–147. <https://doi.org/10.1007/s10956-015-9581-5>.
- Werner, L., Denner, J., Campe, S., Kawamoto, D.C., 2012. The fairy performance assessment: measuring computational thinking in middle school. In: *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*. ACM, pp. 215–220.
- Wilensky, U., 2010. *Restructurations: Reformulating Knowledge Disciplines Through New Representational Forms*, vol. 15.
- Wilkerson, M.H., D'Angelo, C.M., Litts, B.K., 2020. Stories from the field: locating and cultivating computational thinking in spaces of learning. *Interact. Learn. Environ.* 28, 264–271. <https://doi.org/10.1080/10494820.2020.1711326>.
- Wilson, C., Guzdial, M., 2010. How to make progress in computing education. *Commun. ACM* 53, 35–37. <https://doi.org/10.1145/1735223.1735235>.
- Wilson, C., Sudol, L.A., Stephenson, C., Stehlik, M., 2010. Running on Empty: The Failure to Teach K–12 Computer Science in the Digital Age. ACM, New York, NY, USA. <https://doi.org/10.1145/3414583>.
- Wing, J., 2011. Research notebook: computational thinking—what and why. *Link Mag.* 6, 20–23.
- Zhang, L., Nouri, J., 2019. A systematic review of learning computational thinking through scratch in K-9. *Comput. Educ.* 141, 103607. <https://doi.org/10.1016/j.compedu.2019.103607>.

Inquiry-based teaching and STEM education

Yvonne Zeegers, Lisa O’Keeffe, and Bruce White, University of South Australia, Adelaide, SA, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	96
Inquiry-focused questioning for STEM	97
Mathematics and STEM	100
Summary	102
References	102

Introduction

Science, Technology, Engineering and Mathematics (STEM) education has been identified by the Australian Government as being of critical importance to the economic future and development of Australia (Department of Education, Skills and Employment, 2019). This in turn has led to a STEM policy focus that prioritizes STEM for future careers which, as Panizzon et al. (2018) forewarn, could promote a narrow interpretation of STEM. However, the Australian Government’s focus on careers related to STEM has led to an increased emphasis on STEM in primary and secondary schools. The implementation of STEM in schools has taken many forms and the extent of its success has varied. While there are many positive examples of STEM practices in schools, there are also elements such as inflexible timetable structures, particularly in Secondary Schools, that can be a barrier to its implementation. This article is based on our work with teachers and students across a range of STEM in Schools projects. For example, O’Keeffe et al. (2018, 2021), Panizzon et al. (2021), Scott et al. (2021).

To support teachers in planning specific conceptual and skill development through STEM inquiry, there needs to be a focus on planning for intentional learning. While students are often able to articulate *what* they have produced as part of an inquiry project they do not always recognize, and are not always able to articulate, the scientific/mathematical/technological concepts involved. Besides reframing their practices in planning for inquiry-based learning, the formulation of sound questions, be they the teachers or their students, is critical to the success of an inquiry. Moreover, some teachers have found it difficult to make links to the mathematics curriculum and have felt that what had been delivered was “tokenistic” and lacked the conceptual depth for student learning in this area. In this article we discuss the importance of inquiry based, question oriented pedagogical approaches as well as why mathematics is essential in STEM focused inquiry-based learning.

Inquiry-based learning is closely connected with STEM teaching learning and many descriptions of STEM education include inquiry in some form. For example, the “Guiding principles for schools to support STEM education” in the Australian “National STEM school education strategy 2016–2026” which has as one of its principles

Build on students’ curiosity and connect STEM learning to solving real world problems, including through collaborative and individual learning experiences that are hands-on and inquiry-based and support the achievement of deep knowledge.

Education Council (2015, p. 11).

While Inquiry-based learning is not restricted to STEM alone, researchers such as Bybee (2013), have included inquiry as an aspect of STEM literacy while Kennedy and Odell (2014) identified inquiry as a component of high-quality STEM programs. Inquiry can also be viewed as an approach to teaching (Metz, 1995) which allow students to “engage in authentic and meaningful activities that are connected to the real world” (p. 538). For example, Paige et al. (2020) discuss the ways in which inquiry-based approaches to STEM can connect students/young children to their local areas by authentically engaging them in local issues via an eco-justice or sustainability lens. As Scott et al. (2021) found,

Inquiry-based learning (IBL) approaches have the capacity to transform practice from a passive transmission type teaching approach to something more engaging, cognitively stimulating, and ultimately more effective for learning.

Teaching Science (2021).

The extent to which STEM based inquiry projects align to what might be considered the traditional school curriculum can sometimes be unclear to teachers. Addressing this requires strategies such as mapping the project against each of the STEM components. This planning process incorporates the conceptual requirements, pedagogical approach, and assessment of learning. Assessing student learning can be problematic in STEM inquiries because the range of curriculum requirements, the

additional skills required for inquiry, and the diversity of ways for students to communicate project outcomes are not always part of a system's customary assessment and reporting process. Thus, teachers need to understand that while learning design principles for STEM inquiries may be qualitatively and mechanistically different from learning through other pedagogical strategies, they can assist their planning for learning. Such an approach has the potential for some teachers to minimize planning for full engagement in a STEM inquiry project, and could initially result in a project being "trialled" as an addition to the regular curriculum requirements.

The structure and content of inquiry-based projects are by their nature diverse; however, some common features can be found. For example, from our experiences in working on several STEM focused professional learning projects with schools in both metropolitan and regional areas, and from private and public systems, inquiry-based projects had either a question or a problem to be "solved" by groups of students. In reporting their experiences of using this approach to address STEM, teachers said that it had increased the collaborative opportunities for students to work together and had developed skills such as clear communication. The projects also fostered skills relevant to STEM, such as critical thinking, collecting data, using information to generate ideas, and being able to gather additional information to expand their ideas.

When implementing the STEM inquiry projects in schools, we found that teachers needed time to develop the classroom practices and processes required. Initially teachers worked through an "exploration" phase where they trialled their approach. In this initial stage the STEM inquiry projects often missed key elements in planning such as deep conceptual connections for each curriculum area, or clear assessment tasks. However, this form of productive struggle, or of being in the "Learning Pit" (Nottingham, 2016, p. 109), was deemed a critical component in teacher change. This struggle became even more evident in situations where new teachers joined a project in its second year (due to teacher turnover after the first year). While these teachers had the advantage of learning from their colleagues already in the project, in most cases they found the approaches now being implemented by their colleagues too daunting, for example in managing student agency. Instead, they opted to work through the initial "exploration" phase with its narrower, more focused approach. With time and practice, these teachers then worked through the second phase and "refined" their approach, ensuring a greater focus on student inquiry, clear curriculum connections and focused assessment (Scott et al., 2021). These teachers also spoke of "next year" demonstrating that they had already reflected on aspects of their approach that could be improved to enhance student learning. Interviews with the teachers, particularly those who worked on collaborative projects for two or more years, reported significant increases in their confidence to organize inquiry-based learning in their classrooms, about their knowledge of enacting an inquiry-based learning approach and, about how this might be more fully implemented in their classrooms.

The teachers with whom we worked across a range of different teacher professional learning projects often wanted to be able to extend opportunities for their students to experience authentic STEM-related activities. This goal motivated them to actively engage their students with local industries and community-based projects, enabling students to realize the application of STEM outside of the school environment. Connection to community was established in a variety of ways, for example, with the local council or with a local industry. Overall, the breadth and depth of inquiry-based projects in which teachers engaged their students ranged from projects that teachers developed using structured inquiries, to those that were wholly student designed and implemented. We found that the more student agency allowed for within an inquiry-based project the more students were able to provide detailed articulations of their learning across different disciplines and/or skill capabilities.

Across the STEM projects we consistently found that, where teachers had autonomy, strong leadership support, an opportunity for collaboration with their colleagues, and access to external professional development, or industry connections, then authentic, student led inquiry was well integrated into their teaching. This in turn has the potential for a positive impact on their general pedagogical practices, as well as on student dispositions and attitudes toward the individual STEM disciplines.

Inquiry-focused questioning for STEM

As Oliveira has reported (2010, p. 422), there is growing evidence in the science education literature that teachers commonly have difficulties interacting with students in inquiry-based settings. In our work with teachers within and across schools, when investigating what changes in curriculum, pedagogies and/or practices were evident from teachers' involvement in inquiry-based STEM projects, we found that three elements stood out as being particularly influential to teachers:

- The ongoing support at multiple levels (School, Sector, University) through professional learning workshops, and the affordance of the project to take a chance and try something new or something the teachers have wanted to try out,
- The opportunity for these teachers to plan and implement STEM curriculum as a group of teachers (within school or across school collaboration), and
- The ongoing refinement of the teachers' understanding of STEM and its connection to the Australian Curriculum.

A focus of the STEM-based projects that we developed was on pedagogical change through an inquiry-based teaching approach. This generally involved professional learning through a series of workshops with all the teachers. The workshops were designed to build teacher confidence and expertise in the planning, assessment, and teaching of STEM discipline knowledge and skills through experiencing this themselves. The workshops were also typically supported by small group sharing and mentoring of the teachers during half day, school-based meetings.

Teacher and student engagement with STEM can be fostered through an inquiry-based approach that also involves reasoning and critical and creative thinking through the generation and utilization of questions that connect student interest to meaningful learning. Fundamental to this is the need for teachers to pose intellectually challenging questions that foster a curiosity for intentional learning. As Toshalis and Nakkula (2012, p. 2) commented, “research has shown that the more educators give their students’ choice, control, challenge, and opportunities for collaboration, the more their motivation and engagement are likely to rise.” To be able to do this well, teachers themselves need both the confidence and preparedness to pose questions that they themselves might not know to answers to. They need to take the time to explore and extend their own questioning skills to first foster student interest in STEM based inquiries and second, to encourage their students to ask intellectually significant questions that lead to inquiry-based STEM investigations. As Sedova et al. (2016, p. 15) found, “It makes sense to teach teachers the specific skills of dialogic teaching.” Hence, in our work with teachers we typically included a focus on teachers practicing, developing, and extending their STEM focused questioning skills. For example, in one project we challenged participants to identify key differences between utilizing open and closed questions while fostering inquiry-based student learning. This was enacted through introducing an unusual object to them (e.g., an early 1900s coal fueled iron) and posing questions that ranged from “what is this object used for/made of” to “what else could it be made of/used for/improved by?” In small groups the participants then examined a range of arthropods and either asked or answered closed questions to be able to identify them in a “line-up.” This was intended to foster animated discussion about the questioning process (closed vs. open), the difficulty in asking students focused yet inquiry generating questions and, the necessity for clarity of language when questioning with purpose. A final phase was to develop and apply questioning skills. Our example used pictures of shoes through the ages, where participants were tasked with generating open questions and interrogating the STEM related connections that included mathematical, scientific, and technological elements. For example, “How might the construction of a heel influence it’s balance?” As Zeegers and Elliott (2018, p. 3) found, “Implicit in these approaches is the need for teachers to pose intellectually challenging and engaging questions in the classroom.”

In the following section, a case study of two teachers from one of the project’s primary schools has built on the previous provocations around questioning to utilize inquiry-based questioning as the launch pad for engaging students in a STEM related learning theme lasting one school term.

This school in which this case study is located is a small (~350 students) co-education Catholic school located in the metropolitan areas of Adelaide, South Australia. Teachers, Cara, and Mary (pseudonyms used) had opted to take part in our 2-year STEM inquiry project that was centered on developing STEM inquiry-based projects that were either within, or across school collaborations.

Cara and Mary’s aim for the 2-year project was to “develop a school-based STEM philosophy to inform the teaching and learning practices at their School using a focus on the design of a new learning space linked to the School’s Building Project.” In 2018 the Year 4/5 students had explored the features of classroom design that would make comfortable and environmentally friendly learning spaces. The emphasis was on designs for breakout spaces inside the building.

In 2019, Cara and Mary’s first steps were to identify an authentic inquiry for the junior primary aged students. As such, their project remained focused on the funded 2-year school building project that included a dedicated STEM room with 2 breakout spaces and extended this to having students design some outdoor learning areas. While the inquiry’s focus was on STEM and how teachers could incorporate student input to help design and create engaging learning spaces within the build, the project also provided multiple opportunities for addressing Australian Curriculum General Capabilities including Literacy; Numeracy; Information and Communication Technology Capability, Critical and Creative Thinking and, Personal and Social Capability.

The teachers involved their forty-four 6-year-old students acting as sound engineers to explore elements that could be included to create an engaging outdoor area for other students. The students’ task was to design and create a model of a sound-making element that could be added to the outdoor space. To build on student curiosity, the teachers discussed the processes used by engineers and designers to solve a problem by creating new products, systems, or environments. The TASC (Thinking Actively in a Social Context) Framework (Wallace, 2019), was used to record student ideas throughout the process. This visual reminder was on constant display in the classroom and enabled students to refer to these ideas when creating their designs. Cara and Mary also used this framework to plan out learning experiences that included gathering students’ prior knowledge and questions about sound, to having them then design, create and test a sound-making element prototype. As part of their teaching approach Cara and Mary sent a note home asking for recycled materials so that students could create their own instruments and explore how sounds were made. They encouraged students to go outside and investigate elements of sound. They also set a “School Holiday Playground Challenge” wherein students returned to school with photos of places they had visited and things they had discovered about sound. Students then participated in an African drumming incursion where they explored beat and rhythm and discovered the materials used to make traditional instruments. The students also visited a local park to investigate and experience sound elements during play.

Part of their planning included some Explicit Direct Instruction because Cara and Mary wanted to ensure the students had an accurate knowledge base, they could use prior to creating their prototype. With a focus on “*the Science in Sound*,” students investigated how they could make sound by singing a nursery rhyme with their hand over their throat and explaining what they noticed; completed a pitch and volume investigation; discovered how they could make an instrument play a soft sound or a loud sound; listened to sound effects and discussed how the sound is made and heard.

In tune with the STEM focus, in *Mathematics* the emphasis was on measurement, with students investigating how length and volume affect the pitch; looking at the length of the bars on a xylophone; creating musical water jars and comparing the sizes of

a range of instruments; In *Technologies & Arts* the students investigated making sound, identifying how sound can be manipulated (changing the pitch) and, investigating which materials were suited to making instruments. Recycled materials were used to make a range of different musical instruments including drums and maracas. The unit connected with National Simultaneous Story-time which featured the book “Alpacas with Maracas.” In their specialist music lessons, students explored beat and rhythm using songs including the Queen song “We Will Rock You.” In *English, Humanities, and Health*, students were introduced to exposition focused writing by exploring language that could be used to present a point of view. They investigated distinctive features of places such as play spaces, comparing their experiences to the experiences of their parents or grandparents. They also explored the features of traditional indigenous instruments including rain sticks, clap sticks and the didgeridoo. In addition, because two students had cochlear implants, Nicola from “Can: Do 4 Kids” talked to all the students about how their ears function and how cochlear implants work.

Equipped with this background knowledge and having a deeper understanding about sound, the formal inquiry commenced. The principal visited the class to explain the students’ role as “sound engineers” and some of the elements that she would like to have designed for the outside space. The design of such sound elements involved students in looking at musical instruments and brochures before discussing and sketching ideas with their partners. Once approved by their class teacher, students constructed their prototypes using several of the recycled materials including offcuts donated from the actual building site such as piping, wood, and guttering. Students tested their prototype before making any final changes and preparing to share at the community expo. Design ideas from sketching an initial idea to producing the final product were presented in their portfolio for display. Parents and the school community attended the expo wherein the students wrote expositions, shared their portfolios, and filmed short statements to convince the principal to choose their design. Following the expo, students had the opportunity to reflect on their learning experience. Their comments included enjoyment in working collaboratively; having fun; being creative; having the opportunity to share their learning; and being proud of their efforts. A photo display was mounted in the front office, and the students presented their models during a whole school assembly.

This was followed by the students sharing and communicating their ideas with the wider school community and then reflecting on their learning. Throughout the process the students created a portfolio containing their questions, design ideas and reflections on their learning. In addition to the student reflections, staff reflections on their learning experiences and the visitor feedback from the “community expo” helped Cara and Mary to reflect on the entire teaching and learning experience. As the School Principal stated,

A learning space designed by students acknowledges their agency. In 2018 our Year 4/5 students designed an internal space. In 2019 the Year 1s created an outdoor musical space. The musical instruments they designed influenced the school’s selection of large outdoor instruments which will be part of the outdoor learning space and musical garden. These students have shown the school community that they love a challenge!

Principal Interview.

And, as Cara commented

The enthusiasm of the dedicated staff and students, and the opportunity for students to share their learning with the community at the learning expo with enthused parents, staff and visitors was inspiring.

In 2020 the STEM focus continued with inquiry-based learning and a continuing focus on real world/authentic problems that students could connect with. For example, after investigating in the importance of being more sustainable with water use, the year 2 students created a water filtering device that could clean water found in the local area, and the year 6 students designed, made, and marketed a working prototype of a solar powered device that could be used in the school.

As Cara stated,

We are looking forward to continuing our STEM journey and have enjoyed the opportunity to be involved in a project that has both strengthened and reinforced our knowledge of STEM. The opportunity to network with teachers and mentors both through the cluster workshops and the mentor meetings gave us the opportunity to share our progress, discover other STEM ideas developed by colleagues and seek feedback to support our inquiry. Engaging in the project also helped us as a school to develop a clear STEM philosophy with a focus on students engaging in learning activities that investigate the world.

Through our work with teachers on developing STEM focused inquiry-based projects, we have found that teachers genuinely attempted to adopt a problem-solving/inquiry-based approach in their STEM focused projects, and that during this their students’ team building and collaboration skills also developed. Additionally, like the findings of [Toshalis and Nakkula \(2012\)](#), students’ communication processes, along with their motivation and engagement were also seen to improve when an inquiry process that required the formulation and investigation of their own questions was prioritized through fostering student agency and the critical analysis of what they had done.

Mathematics and STEM

Inquiry based approaches to STEM create opportunities for educators to create interdisciplinary programs of learning which include components of two or more of the different disciplines. The [National Council of Teachers of Mathematics \(NCTM\) \(2018\)](#) reminds us that in STEM Inquiry programs the curriculum subjects, which in an Australian context are science, mathematics, and technology, often play a different role to the discipline of engineering. Therefore, it is not surprising that the drivers of STEM inquiry in a primary or middle school context are predominantly science, technology, and mathematics. However, these disciplines are generally not prioritized equally due to factors ranging from teacher preferences to teacher confidence in cross-curriculum alignment. As [English \(2016\)](#) highlighted, many STEM initiatives foreground one of the disciplines, typically science or technology and even engineering (not a curriculum subject in Australia), yet mathematics often has limited emphasis or inclusion.

Some argue that mathematics is a core part of STEM because STEM provides a context for engaging in mathematic ideas and concepts. Such a restricted view of mathematics in STEM might in fact limit the opportunity for deep mathematics learning. [Johnson \(2012\)](#) however, argues that the STEM disciplines of science, technology and engineering are inherently connected by the way they require application of mathematics to solve problems. Thus, implying that rather than being a “forgotten” aspect in STEM, mathematics is the lynchpin that holds it all together. [Fitzallen \(2015\)](#), in her exploration of the way mathematics is integrated into STEM programs of learning, discusses the variance regarding the way in which mathematics is included in STEM and notes that in some studies (such as [Hefty, 2015](#)) the effect size for mathematics is low in integrated STEM learning whereas in others there is significant learning in mathematics. From our work with teachers [O’Keeffe et al. \(2021, 2018\)](#), [Panizzon et al. \(2021\)](#), [Scott et al. \(2021\)](#), we found that inquiry-based approaches to STEM resulted in incremental changes in teaching approaches, with students indicating that there was a small but significant (p. 006) decrease in their experiences of the traditional approaches to mathematics teaching and learning ([O’Keeffe et al., 2021](#)). Additionally, students typically reported small, but in some cases significant increase in opportunities to design their own investigations in Mathematics.

As part of one of our STEM inquiry-based projects, student pre-surveys across 30 schools indicated that they were not confident in their Mathematics and Science ability and were even less confident in being able to apply their knowledge of Mathematics and Science ([O’Keeffe et al., 2021](#)). Over 40% of these students surveyed ($n = 644$) were either not confident or were unsure about their ability in Mathematics and Science. However, analysis of the “post” survey data indicated that their involvement in the inquiry-based approach to STEM had had a positive impact on their confidence in their ability in Mathematics and Science. Interestingly, this increase in confidence was more likely among primary school students than secondary school students. Even so, students reported not enjoying mathematics and science, with mathematics presenting as more of an issue than science. Teachers who were involved in these inquiry-based approaches to STEM themselves reported the challenge of integrating mathematics, and with the assessment of mathematics as part of STEM. One factor mentioned across a range of our projects with teachers, both primary and secondary, has been their recognition of the importance of making a sustained and explicit commitment to embracing inquiry-based learning more holistically. This has meant incorporating a range of assessment strategies, such as detailed rubrics; student presentations; and formative interim reports during the inquiry. It has also required a shift from an end-product focus (summative assessment) to continuous assessment of both process and conceptual learning from each of the STEM related disciplines. For example, students might be assessed on their practical work, their reports, how focused their notes were, on group generated posters promoting their designed and constructed product, and so on. Instigating a formative assessment approach allowed a teacher flexibility to address elements of the different disciplines across the range of assessments, and at different times.

Across the STEM inquiry projects [O’Keeffe et al. \(2021, 2018\)](#), [Paige et al. \(2018\)](#), [Panizzon et al. \(2021\)](#), [Scott et al. \(2021\)](#), teachers reported that using a more flexible approach to assessment required investing time in the initial mapping process mentioned earlier, in the development of different assessment strategies, in collaborating with colleagues and collaborating with community/industry personnel. They also found that inquiry-based STEM projects that were authentic for students, that required critical thinking and a need to communicate and collaborate created more opportunities for student centered approaches to teaching and learning. This in turn impacted positively on their pedagogical practices in STEM. As a result, students identified a greater appreciation for the role of mathematics in STEM Inquiry. As one teacher reported,

[we understand] the role of Mathematics within Science a bit better. We’ve gone to lengths to ensure the maths we’re using, it’s not just volumes, it’s not just area, it’s not just taking averages, there’s some actual real maths. That one [exercise] with the pyramids, they (also) need to use trigonometric functions’.

Secondary Teacher.

Like [Shaughnessy’s \(2013\)](#) warning that the “M” in STEM must be made more transparent, our research also found that inquiry projects need to ensure that the connections between the project content/problem and the specific discipline knowledge required, needs to be made explicit for students. This is of particular importance in Mathematics. While teachers may feel that the mathematics embedded in an Inquiry Project is obvious, that is not necessarily so for students, who tend to find it easier to identify and explain the Science and Technology in an inquiry. For example, students in one focus group were clearly able to articulate their learning in science, making note of aspects such as “Rate of decomposition” and “Molecular structure and properties,” and their

learning in Technology, noting examples such as design, 3D printing, using TinkerCad. However, the detail they provided about their mathematics learning was limited. There were however several examples where students were able to explain various aspects of a given process in scientific or mathematical terminology. As one teacher stated,

An upper primary group had made stress balls using cornflour and water. The students were able to explain how the various elements of water and cornflour, when combined and placed in a balloon, could make a stress ball. Non-Newtonian fluids like cornflour are good at absorbing and dissipating energy. Viscosity increases with applied pressure.

Primary Teacher.

Planning a unit of work that focused on mathematics was least evident in all the school STEM inquiry projects that we have engaged with. As many STEM inquiry projects tend to focus on science, it is not surprising that students find it more difficult to articulate their learning in mathematics. When mathematics was included, the mathematics content tended to be focused on measurements, ratios, or data collection. Such limited mathematical inclusions run the risk of reinforcing students' negative views of the potential and power of mathematics to enhance STEM inquiries, with some indicating they didn't learn any mathematics, and some only providing superficial examples, such as "it was kind of basic maths," "didn't use much maths except for measurements." When probed on their mathematical learning during interviews, the students generally struggled to add any more detail and explain what they were learning in undertaking the inquiry, something the teachers also noted in post project reflections. It is also important to note that this was an area of challenge for some teachers. As with any professional learning initiative, teachers were keen that students were engaged and understood the relevance of what they were undertaking. However, a common concern raised by secondary teachers, was the challenge they faced in planning an appropriate level of content for STEM curricula in their "non" disciplines. Some indicated that ideas for "getting started" initially came from a "fun activity" undertaken during current or previous project/interests (e.g., flying drones, coding a game, making a robotic arm), rather than being driven by intended content learning outcomes, and most particularly with the mathematics curriculum. As such, they often felt that what had been addressed was "tokenistic" and were concerned about the depth of learning by students in this discipline. Some examples provided were the inclusion of budgeting elements, business planning, or a survey component without these being relevant to the context of the inquiry. Such superficial inclusions of mathematics have the potential to reinforce [Shaughnessy's \(2013\)](#) concern that "the M will become silent" in STEM.

However, this does not mean that mathematics cannot be effectively learned in integrated STEM inquiry. For example, in several cases there were missed opportunities to extend students' conceptual understanding in mathematics, such as by encouraging them to explain why something had happened, or why they chose a particular approach. An example that clearly reflects this was a group of Year 7 and 8 students who were working on a handheld water filter ([Fig. 1](#)). The students could clearly articulate the scientific principles of their filter and the how the different properties of the materials they had chosen would contribute to filtering the water. In their initial investigations the students were "trying different ratios of materials" however when asked what ratios they had tried and how the different ratios impacted on the pH of the water the students couldn't answer. This was because the task had not required them to record what they were doing and why, and hence an opportunity to extend their mathematical thinking and understanding of ratio was missed. Another example involved prototypes and models that weren't built to scale. It should be noted however, that examples such as these were more frequent in a teacher's first year of a project when they were still grappling with new planning and pedagogical approaches, let alone the complexities of managing space and equipment across multiple sites. This highlights the importance of initiating long-term professional learning projects in schools.



Fig. 1 Example of handheld water filter developed as part of STEM Inquiry project.

Summary

Choosing or negotiating authentic inquiry-based tasks with learners is essential in ensuring that students have an effective STEM focused learning experience. This isn't to say that there are "good" or "bad" tasks. Rather, inquiry-based tasks must be meaningful enough to allow for purposeful investigation and exploration. When STEM inquiries allow student agency, both teachers and students can provide more detailed explanations of the learning across the STEM disciplines.

Our experiences with STEM focused projects within South Australian schools, both primary and secondary, regional, and metropolitan, public, and private, have demonstrated that STEM focused inquiry-based projects create a unique opportunity for teachers to develop topics/projects such as those based on concepts associated with sustainability, eco-justice, critical numeracy and so on. Critically, and foremost, teachers need to be supported in their STEM journey by school leadership and their education systems, as was done for each of the projects that we designed and led. Such support needs to assist teachers to be able to collaboratively plan, share, learn, apply, and reflect on the content and intended learning outcomes, that incorporate a range of assessment strategies, in *each* of the STEM disciplines. Additionally, and fundamental to its success is that teachers clearly articulate the purpose and conceptual relevance of each inquiry to their students.

References

- Bybee, R.W., 2013. The Case for Education Challenges and Opportunities. National Science Teachers Association. Retrieved from. <http://static.nsta.org/files/PB337Xweb.pdf>.
Department of Education, Skills and Employment, 2019, <https://www.dese.gov.au/australian-curriculum/support-science-technology-engineering-and-mathematics-stem>, Accessed 11th November 2021.
- Education Council, 2015. National STEM School Education Strategy, 2016–2026: A Comprehensive Plan for Science, Technology, Engineering, and Mathematics Education in Australia. Author, Melbourne.
- English, L., 2016. STEM education K-12: perspectives on integration. *Int. J. STEM Educ.* 3 (3), 1–8.
- Fitzallen, N., 2015. STEM education: what does mathematics have to offer? In: Marshman, M., Geiger, V., Bennison, A. (Eds.), *Mathematics Education in the Margins* (Proceedings of the 38th Annual Conference of the Mathematics Education Research Group of Australasia). Sunshine Coast: MERGA, pp. 237–244.
- Hefty, L.J., 2015. Applying mathematics during engineering design challenges can help children develop critical thinking, problem solving, and communication skills. *Teach. Child. Math.* 21 (7), 422–429.
- Johnson, C.C., 2012. Four key premises of STEM (editorial). *Sch. Sci. Math.* 112 (1), 1–2.
- Kennedy, T., Odell, M., 2014. Engaging students in STEM education. *Sci. Educ. Int.* 25 (3), 246–258.
- Metz, K.E., 1995. Reassessment of developmental constraints on children's science instruction. *Rev. Educ. Res.* 65 (2), 93–127. <https://doi.org/10.3102/00346543065002093>.
- National Council of Teachers of Mathematics, 2018. Joint Position Statement: Building STEM Education on a Sound Mathematical Foundation. <https://www.mathedleadership.org/docs/resources/positionpapers/NCSMPositionPaper17.pdf>.
- Nottingham, J., 2016. *Challenging Learning Theory, Effective Practice, and Lesson Ideas to Create Optimal Learning in the Classroom*, second ed. Pub Routledge, New York.
- O'Keeffe, L., White, B., Panizzon, D., Elliott, K., Semmens, A., 2018. Mathematics anxiety: year 7 and 8 student perceptions. In: *Proceedings of the 41st Annual Conference of the Mathematics Education Research Group of Australasia*. MERGA, Adelaide, pp. 607–614.
- O'Keeffe, L., White, B., Zeegers, Y., Leonard, S., 2021. CESA-UniSA STEM, Final Report. University of South Australia, Australia.
- Oliveira, A., 2010. Improving teacher questioning in science inquiry discussions through professional development. *J. Res. Sci. Teach.* 47 (4), 422–453.
- Paige, K., Caldwell, D., Elliott, K., O'Keeffe, L., Osborne, S., Roetman, P., et al., 2018. *Fresh Water Literacies Transdisciplinary Learning for Place and Eco Justice*, Final Report. University of South Australia, Australia.
- Paige, K., O'Keeffe, L., Lloyd, D., 2020. More than STEM connecting students' learning to community through eco-justice. In: Fitzgerald, A., Haeusler, C., Pfeiffer, L. (Eds.), *Stem Education in Primary Classrooms: Unravelling Contemporary Approaches in Australia and New Zealand*. Routledge, pp. 151–167.
- Panizzon, D., Geer, R., Paige, K., O'Keeffe, L., Schultz, L., Zeegers, Y., Brown, L., 2018. Exploring the "hard facts" around STEM in Australia: females, low socioeconomic status and absenteeism. *Int. J. Innovat. Sci. Mod. Eng.* 26 (8), 30–44.
- Panizzon, D., White, B., Elliott, K., Semmens, A., 2021. Using collaborative inquiry projects to enhance students' self-efficacy and self-concept in science: patterns and surprises in the data. In: Geelan, D., Nichols, K., McDonald, C. (Eds.), *Complexity and Simplicity in Science Education* Springer Nature Switzerland AG.
- Scott, L., Pilla, L., O'Keeffe, L., White, B., 2021. STEM through inquiry projects for students: a teacher's perspective. *Teach. Sci. J. Aust. Sci. Teachers Assoc.* 67 (3), 26–37.
- Sedova, K., Sedlacek, M., Svaricek, R., 2016. Teacher professional development as a means of transforming student classroom talk. *Teach. Teach. Educ.* 57, 14–25.
- Shaughnessy, M., 2013. By way of introduction: mathematics in a STEM context. In: *Mathematics Teaching in the Middle School*.
- Toshalis, E., Nakkula, M., 2012. Motivation, Engagement, and Student Voice. The Students at the Center Series. Executive Summary. April 2012. Retrieved from. http://www.studentsatthecenter.org/sites/sci.dl-dev.com/files/field_attach_file/Exec_Toshalis%26Nakkula_032312.pdf.
- Wallace, B., 2019. Thinking Actively in a Social Context (TASC). <https://www.tascwheel.co.uk/>. (Accessed 18 October 2021).
- Zeegers, Y., Elliott, K., 2018. Who's asking the questions in classrooms? Exploring teacher practice and student engagement in generating engaging and intellectually challenging questions. *Pedagogies* 14 (1), 17–32. <https://doi.org/10.1080/1554480X.2018.1537186>.

Makerspace and robotics as/for STEM education

Kamini Jaipal-Jamani, Brock University, Faculty of Education, St. Catharines, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

STEM and STEM education	103
Robotics as/for STEM education	105
Robotics in Informal Science Education settings	105
Robotics in formal educational settings	105
Implementing robotics in K-12 settings	106
Makerspace as/for STEM education	107
Makerspace in informal educational settings	107
Makerspace in formal educational settings	108
Implementing makerspace in K-12 settings	108
Conclusion	109
Definitions	109
References	110

STEM and STEM education

The acronym STEM refers to the disciplines of science, technology, engineering, and mathematics. Since diverse economic sectors are dependent on the technological innovations occurring in the STEM fields (Orpwood et al., 2012), many countries are prioritizing STEM in education as a strategy to promote sustainable economic development, job growth and competitiveness in the global economy (Carnevale et al., 2011; National Research Council, 2014; Warde and Sah, 2019). While the impetus for STEM in education emerged from economic and political contexts, STEM Education is being repositioned as knowledge that promotes intellectual and social competencies and humanistic goals such as democracy and social justice to benefit society (Bybee, 2013; Park et al., 2020; Zeidler, 2016). Three broad goals for STEM education have been identified as STEM literacy, a STEM workforce equipped with 21st century competencies (e.g., critical and creative thinking, collaboration and communication) and STEM researchers and innovators (Bybee, 2013; National Research Council, 2014). In educational settings, the main goal of STEM education is STEM literacy (Vasquez et al., 2013). Bybee (2013) conceptualizes STEM literacy as involving three aspects: (1) having knowledge, attitudes, and skills to ask questions, problem solve, explain phenomena, and draw conclusions from evidence about STEM-related issues, (2) an understanding of how to use inquiry and design, and (3) as using ideas of STEM to reflect on issues as concerned citizens. Tang and Williams (2019) initiated an empirical analysis of the characteristics of each STEM discipline and demonstrated that STEM literacy is a useful educational construct. Their analysis showed that STEM literacy captures some common practices across the STEM disciplines (e.g., the use of diagrams such as charts and schematics) that can be applied by the general public to understand and engage in civic discussions of STEM issues. The authors, however, argue there is a need to distinguish STEM literacy from the unique skills and knowledge of a discipline such as disciplinary language and ways of interpreting, critiquing, and creating that are not applicable to other STEM disciplines. The unique skills associated with a specific discipline, they refer to as S.T.E.M. literacies. Tang and Williams purport that literacy development in a person integrates the skills and knowledge associated with STEM literacy and individual S.T.E.M. literacies. The challenges arising around discipline specific S.T.E.M skills and common STEM skills, as will be discussed below, has to some extent influenced how STEM Education is implemented in educational settings.

While there is a degree of consensus on what the goals of STEM Education are, there is a lack of consensus on what STEM Education is and what a STEM Education program entails (Holmlund et al., 2018; McLoughlin et al., 2020; Siegel and Giamellaro, 2020). In many countries, such as in Canada and the US, STEM in schools has been operationalized through the disciplinary subjects of science and mathematics with strategies to intentionally integrate concepts and practices of science and/or mathematics with the concepts and practices of technology and engineering (Vasquez et al., 2013). According to Vasquez et al. (2013), the different approaches to STEM education can be represented on a continuum of levels of STEM integration. These authors describe a continuum of STEM integration from multidisciplinary (learning disciplinary knowledge separately but around a common theme), interdisciplinary (applying relevant disciplinary knowledge from two or more disciplines at the same time to understand a concept/topic/issue) to transdisciplinary (applying knowledge from two or more disciplines to a project or to solve a problem). The latter two levels of integration align closely with an integrated STEM approach. An interdisciplinary approach begins with the discipline and integrates other STEM disciplinary knowledge and skills to address issues and a transdisciplinary approach begins with a social issue or real-world problem and draws on knowledge of more than one STEM discipline during the problem solving process to find solutions (Fensham, 2009; Quigley and Herro, 2016). The adoption of an integrated STEM approach does not come without challenges in practice.

Within both the interdisciplinary and transdisciplinary approaches, is an implicit expectation that practitioners have a good knowledge of content knowledge and pedagogical skills of two or more disciplines in STEM. Hence the implementation of an

integrated STEM approach is a challenge for many practitioners at K-12 schools and higher education institutions. Many educators lack knowledge of the content, practices and pedagogy associated with the different STEM disciplines. Studies show that K-12 math and science teachers lacked confidence to teach engineering concepts, and K-12 technology teachers identified limited content knowledge in science and math as barriers to teaching engineering design (Ames, 2014; Kelley and Wicklein, 2009). Teachers are either experts in a single STEM discipline (e.g., math or science) in high schools or may be generalists and teach all subject areas (e.g., elementary school teachers; Park and Ertmer, 2008). Quigley and Herro (2016) found that middle school math and science teachers had difficulties with transdisciplinary STEM teaching because of their desire to place their content specializations at the forefront. The latter may be motivated by the need to meet disciplinary knowledge expectations while trying to make connections across disciplines (NRC, 2014). Research has shown that another limitation of interdisciplinary STEM tasks that included mathematics, was the limited development of foundational math thinking and concepts as assessed by curriculum standards (Tytler, 2020). While interdisciplinary tasks did enhance student interest and enthusiasm for mathematics (Tytler et al., 2019), the smallest gains occurred in mathematics achievement (Williams et al., 2016). Another challenge for teachers is assessing practices and applications of other disciplines which they may have limited knowledge about – for example how argumentation practices in science and math are different (NRC, 2014). Hence, as reported by Siegel and Giamellaro (2020, p. 744),

The adoption of STEM relies heavily on the stakeholders, within their unique context, actively defining an innovation and that practitioners made sense of and connected to STEM based on existing context and practices.

Siegel and Giamellaro (2020) concluded that understanding and implementing STEM is a process that is active, context-based, negotiated, co-constructed and evolving. The implication for practice is that the implementation of STEM is dependent on the different ways teachers implement STEM and this implementation affects student learning of STEM. As reported by the National Research Council (2014, p. 7),

The expertise of educators working in classrooms and in after-/out-of school settings is a key factor – some would say *the* key factor – in determining whether integrated STEM education can be done in ways that produce positive outcomes for students.

Therefore, an important question posed by the National Research Council (p. 10) was: “What instructional approaches or contexts are most likely to lead to student outcomes related to making connections between and among the STEM disciplines?” The NRC conducted a study over 2 years on Integrated STEM Education in K-12 in the US and found that STEM integration that was characterized by explicit support to learn disciplinary knowledge and skill showed promise to increase learning of concepts within the disciplines, especially in engineering and technology. Targeted supports included strategies such as explicit prompts and teacher demonstrations that focused on disciplinary concepts (Schnittka and Bell, 2011).

Given the challenges many teachers face related to lack of disciplinary knowledge in STEM fields, Kelly and Knowles (2016) suggest using “an integrated approach [that] seeks to locate connections between STEM subjects and provide a relevant context for learning the content” and they propose engineering design as an ideal context for integrating STEM because of the “systematic approach to solving problems that often occur naturally in all of the STEM fields” (p. 3). Consistent with the latter, Moore and Smith (2014, p. 5) conceptualize STEM integration as student participation in engineering design in K-12 education and explain:

Integrated STEM education is an effort to combine the four disciplines of science, technology, engineering, and mathematics into one class, unit, or lesson that is based on connections among these disciplines and real-world problems ... with students participating in engineering design as a means to develop relevant technologies that require meaningful learning through integration and application of mathematics and/or science.

The National Research Council (2014) explains further that engineering design, scientific inquiry and problem based learning (PBL) share features that provide relevant contexts to apply STEM concepts. For example, they explain that all three learning strategies are characterized by features such as experiential learning, student centeredness, problems acting as both the focus and stimulus for learning, and self-directed learning.

Robotics and Makerspace activities, which incorporate engineering design, scientific inquiry practices, and PBL have potential to promote interdisciplinary and transdisciplinary approaches to integrated STEM education and provide the context to address practitioner challenges and ease the way to STEM adoption within the curriculum (Bevan et al., 2015; Sullivan, 2008). Robotics and makerspace act as the context to apply knowledge from the four STEM disciplines as students design and construct artifacts/products or create solutions to solve real-world problems. The focus in these two contexts is on STEM: engineering design, technological construction process and discipline specific knowledge and inquiry skills from science and math applied during the activity. The next two sections discuss how robotics and makerspace respectively act as the context for and in support of integrated STEM Education in informal and K-12 education.

Robotics as/for STEM education

Robotics presents an opportunity to learn about the four disciplinary areas of STEM while promoting the learning of competencies and skills such as critical thinking, problem-solving, computational thinking and inquiry (Nugent et al., 2014; Ma and Williams, 2013). Research studies show that robotics supports STEM Education by enhancing STEM content learning and skills, improving interest and motivation in STEM subjects, and broadening diversity and participation of underrepresented groups (Anwar et al., 2019; Benitti, 2012). Specific benefits of educational robotics include enhancing understanding of specific STEM concepts in science and math respectively (Ardito et al., 2014; Park, 2015; Williams et al., 2012), developing science process skills (e.g., Sullivan, 2008) and increasing student engagement, interest, creativity and motivation in STEM courses and careers (Hudson et al., 2020; Miller et al., 2018; Master et al., 2017). Interest in STEM is important as it has been shown to positively influence school achievement (Renninger and Hidi, 2011; Wigfield and Cambria, 2010) and science degree attainment (e.g., Tai et al., 2006). Studies also show that robotics supports the progressive development of skills such as computational thinking (Kazakoff et al., 2013; Sullivan and Heffernan, 2016), metacognition (e.g., La Paglia et al., 2011) and problem solving (e.g., Atmatzidou et al., 2018). The aforementioned benefits of robotics reinforce the use of robotics for and as STEM Education.

Robotics in Informal Science Education settings

Robotics offers a learning context to engage youth in authentic, hands-on, design, creation, and programming using knowledge and skills that engineers and programmers apply. Such programs help students learn scientific and engineering skills that are not easily learned in traditional formal educational settings (Gura, 2012). Robotics programs are commonly implemented by informal organizations such as science centers, scientific organizations and faculties of engineering at colleges and universities as after school programs (Nugent et al., 2010; Williams et al., 2007) and as robotics competitions (Chung et al., 2014; Karp and Maloney, 2013).

Informal Science Education (ISE) programs play an important role in supporting STEM Education. ISE has long been used as the context to promote interest in science and sustain participation in STEM careers among youth and underrepresented groups (Fadigan and Hammrich, 2004; Tytler and Osborne, 2012). A recent longitudinal study (Habig et al., 2020) investigating how an ISE program, conducted at a museum, impacted the STEM trajectories of youth, found that 82.3 % of youth alumni enrolled in a STEM major and 63.1% engaged in a STEM career. The study also reported that the majority alumni in these STEM areas were females and/or members of historically underrepresented racial and ethnic groups. Findings of this study also revealed that the most important ISE program features that contributed to sustained participation in STEM were the provision of multiple opportunities to practice and do hands-on STEM, building social networks and relationships with STEM practitioners such as scientists, and exposure to a variety of STEM areas and careers. However, all students, especially students from economically disadvantaged communities, will not have access to such ISE programs as many programs are scheduled outside of formal school hours, require extra time and may involve additional costs. Hence, including robotics in formal educational settings as part of the curriculum is also needed to promote equitable access so all students can benefit from this STEM approach.

Robotics in formal educational settings

According to Benitti and Spaloor (2017), robotics, which combines technology and engineering principles in real world applications, supports teachers to concretely demonstrate science and mathematics concepts in real-world applications. Scientific inquiry as investigating phenomena to find explanations and engineering design as devising solutions to problems parallels how scientists and engineers engage in knowledge creation and knowledge application respectively (NRC, 2012). In many countries, engineering design is incorporated into the K-12 science curriculum. For example, since 1998, engineering design has been included in the Canadian elementary school science curriculum. In the US, the practices that characterize scientific inquiry and engineering design for school curricular are outlined in the *Framework for K-12 Science Education* (National Research Council, 2012). According to this Framework, there are eight practices considered essential elements of the K-12 science and engineering curriculum: “asking questions (for science) and defining problems (for engineering), developing and using models, planning and carrying out investigations, analyzing and interpreting data, using mathematics and computational thinking, constructing explanations (for science) and designing solutions (for engineering), engaging in argument from evidence, and obtaining, evaluating, and communicating information” (p. 49).

Research (Ziaeeefard et al., 2017) shows that robotics engages students in engineering design practices such as defining the engineering problem (e.g., how does the robot work to solve the problem), proposing the solution to the problem (e.g., how to build the robot), and optimization (e.g., how to improve the efficiency of the robot to complete the task). While there are many practices common to scientific inquiry and engineering design such as stating a hypothesis, designing investigations, repeating experiments, analyzing and interpreting data to confirm or reject the hypothesis, using evidence and evaluating, and communicating results or information, there are differences. For example, how scientists and engineers develop and use models, plan and carry out investigations, and analyze and interpret data are different. Robotics can be used as a real world context to show the different ways scientists and engineers work and is a promising strategy to promote knowledge of technology and engineering in K-8 schools where these disciplines are most underrepresented (Bybee, 2010).

Research also indicates that robotics can provide opportunities for even elementary school students to learn science and engineering concepts and practices earlier than expected and these robotics learning experiences can develop students' confidence and

interest in math, science, and engineering (Park, 2015; Rogers and Portsmore, 2004). An analysis of research studies conducted over 15 years (Bascou and Menekse, 2016) found that “active learning based pedagogies, which were made possible through educational robots, were considerably more effective than pedagogies that passively transmitted information from teacher to student” (p. 6). Active learning pedagogies are student centered pedagogies grounded in constructivist theories of learning that posit that knowledge acquisition or learning involves the personal cognitive construction of knowledge mediated by social interactions (Duit, 1996). The well known 5E Inquiry-Based instructional model is based upon the constructivist theory of learning (Bybee and Landes, 1990) and consists of five phases: engage, explore, explain, elaborate, and evaluate. This model considers students’ prior knowledge with instructor monitoring and feedback and opportunities for student reflection. Robotics activities incorporate many components of the 5E inquiry model. For example, robotics engages students in open inquiry through the use of concrete tools (Sullivan, 2008), has in-built scaffolds such as structured prompts and visuals to guide students through activities (Mitnik et al., 2008) and provides just in time feedback (Sullivan, 2008). Robotics activities, with active hands-on and minds-on learning, can act as a context for inquiry in the STEM disciplines such as science and math. The inclusion of robotics in science or math classes is not consistent and widespread and may be due to a variety of factors. For example, robotics is not an expectation of many K-12 science or math curricula even though robotics lends itself to inquiry-based learning and engineering design and, there are challenges associated with its implementation in K-12 settings.

Implementing robotics in K-12 settings

While robotics has potential to engage students in STEM learning, the gains in learning are not guaranteed just by the simple application of robotics. As reported in a meta analysis of robotics studies by Benitti (2012), most studies indicated gains in student learning of STEM concepts and skills, however, four studies reported non-significant gains in student learning. It is therefore important to recognize that the implementation of a robotics program does not necessarily enhance student learning of STEM concepts (Benitti, 2012; Park, 2015). A number of factors contribute to knowledge gains for students during robotics implementation. For example, one of the factors cited for the low gains in geospatial concepts in the study by Nugent et al. (2008) was the lack of alignment of concepts with topics in the traditional middle school science curriculum. Another factor that appeared to affect student learning from robotics interventions was the length of time for the intervention. Students, participating in a longer robotics program demonstrated significantly greater understanding of STEM concepts and showed higher self-efficacy than students participating in a shorter intervention (Nugent et al., 2010). It was also noted that the shorter intervention did have a positive effect on student attitude—students’ attitudes and interest toward robotics increased. The latter result suggests that shorter robotics interventions can play a role in motivating students and increasing interest in STEM subjects which creates a high degree of student interest and engagement in STEM careers (Nugent et al., 2010; Robinson, 2005; Rogers and Portsmore, 2004).

In a recent meta analysis by Anwar et al. (2019), the authors reviewed 32 studies that involved the use of robotics to enhance construction of knowledge and application in STEM through an active learning process. These studies revealed that robotics did provide experiences that helped students increase understandings required for constructing knowledge and enhanced students’ ability to transfer knowledge. However, the authors also draw attention to the short-term nature of many of the comparative studies and call for more long-term studies to reinforce findings of knowledge and transfer gains. Another meta-analysis of 144 experimental studies investigating the effectiveness of computer based scaffolding on problem centered approaches in STEM education such as inquiry, design, and problem based learning showed that students who received scaffolds did better on cognitive tests than students who did not receive scaffolds and all scaffolding intervention types improved cognitive outcomes (Belland et al., 2017). Providing scaffolding (e.g., just-in time tutorials, lessons, immediate feedback and debriefings) during the problem-solving robotics activities and providing different activities to differentiate for different levels of content knowledge and problem solving skills, helps students make connections to scientific and engineering concepts (Williams et al., 2007). Other factors that enhanced STEM learning during robotics activities were large physical spaces for students to engage in experimentation and testing, small student groups of 2–3 pupils per robotics kit, and accommodating for diverse student interests and learning styles – e.g., providing multiple entry points into robotics (Lindh and Holgersson, 2007; Rusk et al., 2008).

The literature also suggests that the extent and quality of robotics learning is dependent, to a larger degree, on the teacher. Studies have identified the following characteristics of a teacher that impact the successful implementation of robotics:

- Comfort and skill level using robotics software for programming (Vollstedt et al., 2007),
- Attitude of the teacher – how to motivate students and promote positive attitudes toward the use of robotics (Lindh and Holgersson, 2007)
- Knowledge of appropriate pedagogical approaches to teach STEM concepts using robotics such as robotics enhanced inquiry-based learning (REIBL; Park, 2015) or engineering design (Rogers and Portsmore, 2004).
- Knowledge of specific scaffolding to show how robotics activities relate to science and engineering processes and to curriculum learning goals and subjects (Nugent et al., 2010).

Successful robotics implementation as and for STEM in K-12 settings depends on the teacher’s knowledge of robotics, STEM content, and student centered pedagogy and their comfort level to implement robotics based activities within classrooms. Teachers must be provided with professional development support that is effective at building content and pedagogical knowledge, and the technical skills for using robotics and other technology to promote student learning of STEM. Professional development to enhance teacher self-efficacy to integrate robotics in instruction and engage in curriculum design (Anwar et al., 2019),

as well as technical support during implementation (Rogers and Portsmore, 2004) are prerequisites for effective robotics implementation in K-12 settings.

Makerspace as/for STEM education

The maker movement arose from the “Do-it-Yourself” culture comprised of crafters, hackers, tinkerers, quilters and other makers who collectively became known as makers (Bevan, 2017). Dougherty (2013), the founder of Make Magazine describes makers as enthusiasts “who played with technology to learn about it” (p. 7). The distinction of the maker movement from the prior activities of crafters, tinkerers and inventors is the prevalent use of digital technologies such as 3D printers and microcontrollers, and the online sharing and collaboration that characterizes current maker communities (Halverson and Sheridan, 2014). Martin (2015), drawing from various definitions, defines making “as a class of activities focused on designing, building, modifying, and/or repurposing material objects, for playful or useful ends, oriented toward making a ‘product’ of some sort that can be used, interacted with, or demonstrated” (p. 31). Although there is not one unified definition for makerspace, a common feature is the idea of it being a shared space or “gathering point” where people can come together and work on projects, alone or with help from persons who may be knowledgeable in the area of exploration, with access to various tools and new technologies (Maker Media, 2013, p. 1). The nature of making therefore resonates with the T and E in STEM. Engineers design, test, modify, develop and conduct failure analysis of a variety of products where technology is used in the manufacturing and construction process to create technological innovations (Lucas, 2014). Similarly, makers create and problem solve in spaces called makerspaces, hackerspaces, tinkering studios and playlabs where they come together to create and design technological products (Anderson, 2012). The focus on design activities in makerspaces align with many science education goals (e.g., Next Generation Science Standards, NGSS Lead States, 2013; The Ontario Curriculum Grades 1–8, Science and Technology, Ministry of Education, 2022) that incorporate engineering design and technology. Halverson and Sheridan (2014), however, caution against limiting maker to K-12 learning activities as this would limit the democratizing potential of maker as a way to provide access to the power discourses usually employed by producers of artifacts using 21st century technologies.

Makerspace in informal educational settings

Making in informal spaces is a spontaneous activity and stems from people’s enthusiasm, passion and interests and involves making things, taking them apart and remaking to create new innovations and products (Dougherty, 2012). Many community spaces such as libraries, museums, and science centers have created makerspaces to promote active participation and hands-on learning in exhibits related to STEM and art (Bevan et al., 2015; Gutwill et al., 2015; Kafai et al., 2014). According to Educause (2013),

Makerspaces are zones of self-directed learning. Their hands-on character, coupled with the tools and raw materials that support invention, provide the ultimate workshop for the tinkerer and the perfect educational space for individuals who learn best by doing ... they promote multidisciplinary thinking and learning, enriching the projects that are built there and the value of the makerspace as an educational venue (p. 2).

The unique nature of makerspaces derives from addressing what the community needs (Hughes, 2017) and therefore building or creating a space that suits the interest of the community of makers using the makerspace is necessary. According to Bullock and Sator (2015), making is:

A participatory activity that draws the whole body of the maker into the creation of the project and the resulting artefact(s) represent the creative process of design, social interaction, and ethos of sharing inherent in the personal and collective learning process. (p. 70)

Informal makerspaces are self-directed and involve makers across age groups. Sheridan et al. (2014) examined three makerspaces in informal settings comprised of makers who were adults, families and youth. Their study identified the following characteristics common across the three makerspaces: diverse tools, materials and processes; makers worked across disciplines in terms of the processes engaged in and the products made; formal pedagogical structures such as computer workshops or skill demonstrations were incorporated with voluntary attendance; mentoring and leadership roles were assumed as needed; learning was embedded in the process and product; members found projects and products to work on and engaged in the design process; and creations and products were shared.

A few studies have investigated STEM related, informal makerspaces. Bonnette et al. (2019) conducted a study with 983, sixth grade students to explore the relationship between participation in informal science activities and children’s fascination/interest in science. Their findings indicated that “informal experiences are positive predictors of fascination. Children who get involved in informal science before middle school and stay involved have a higher fascination, which in turn predicts greater involvement in informal science during the school” (p. 1636). Other research also suggests that engaging in maker-centered learning activities may encourage more students to pursue learning and career pathways related to STEM content areas (Bevan et al., 2014; Hsu

et al., 2017). Fostering student interest and participation in STEM is one of the goals of STEM educators in K-12 formal learning environments. However, tensions exist between making practices and formal educational practices. The question to ask is: how can the features of informal makerspaces be translated into formal educational environments in ways that reflect the unstructured, creative, and collaborative nature of learning in makerspaces compared to the structured, curriculum aligned learning in formal K-12 educational environments?

Makerspace in formal educational settings

Making in formal educational settings is associated with the theory of constructionism (Papert, 1991) where the learner builds personal, concrete knowledge through construction of artifacts chosen by the learner. The four elements of a maker mindset outlined by Martin (2015) provide a workable framework for designing makerspaces in education that align with constructionism. Martin proposed that makerspaces be characterized by the following: (1) being playful – activities selected are those that learners find fun and playful; (2) being growth oriented – provide free choice of activities to enhance skills consistent with strengths of learners; (3) failure-positive – make failure an option and encourage students to use it to adjust, learn and create; and (4) collaborating and sharing – learners work together to discuss ideas and solutions to generate new knowledge and products. These makerspaces therefore provide a space and an opportunity for students to work collaboratively with others to design and create and use trial and error to solve problems through several iterations. Choice in activities provides learners with a degree of agency, similar to informal makerspaces, and promotes access to all learners by offering multiple paths to learning (Martinez and Stager, 2013). Studies show that well designed makerspaces in formal educational environments allow students to experience learning in new or different ways and cater to a range of different learning styles (Blackley et al., 2017; Sheffield et al., 2017) and contribute to an equitable learning space (Martin et al., 2018). A STEM makerspace is therefore another viable strategy to add to the repertoire of hands-on, concrete, inquiry based activities already being implemented in formal science and mathematics classes.

Implementing makerspace in K-12 settings

The formal implementation of a makerspace approach as a pedagogical strategy in K-12 schools is in its infancy. While Martin's (2015) framework provides a guide for creating a maker experience in formal educational settings, it does not address how the goals of formal curriculum are incorporated into makerspace experiences and what skills educators need to be able to implement makerspaces for effective student learning. With respect to the formal school curriculum, in integrated STEM education, learners draw on knowledge and skills from the different STEM disciplines during learning. The latter is congruent with a makerspace approach that deliberately sets student learning in contexts that draw on skills and knowledge from the STEM areas "to create, construct and critique a product that has one or more engineering design elements" (Sheffield et al., 2017, p. 152). Sheffield et al. (2017) argue that the makerspace approach can be a robust pedagogical strategy when strong and explicit connections are made between the making activities and resultant product and curriculum standards of the different STEM areas. Ensuring that relevant curriculum connections are made during the makerspace activity for students is dependent on the pedagogical knowledge and skills of the classroom teacher or educator. Some teacher education programs are introducing makerspace as an approach to teach STEM. Blackley et al. (2017) examined how preservice teachers learn to teach STEM through a makerspace type activity. A small group of preservice teachers (PSTs) participated in a STEM makerspace workshop where they were mentored by engineering students and learned how to design and create an engineering artifact. An important component of PSTs preparation was also learning pedagogical strategies to manage groups, ask questions and make connections to the science concepts related to the artifact. PSTs then worked with 3–6 elementary students in a makerspace activity at a school and provided students with the design task and materials to plan and create the product. The findings of this study with PSTs indicated that maker activities, when introduced into formal educational settings such as classrooms, which are constrained by curriculum standards and formal assessments, require pedagogical scaffolding from the teacher for effective student learning. Scaffolding during maker activities such as teacher questions, is necessary to show connections between the inquiry or design process and knowledge to be learned (Blackley et al., 2017; Darling-Hammond, 2008). Studies (e.g., O'Brien et al., 2016) also suggest that effective implementation of makerspaces requires teachers to be knowledgeable at facilitating less structured instructional strategies and being responsive to the spontaneous nature of makerspaces.

In relation to effectiveness for student learning, Halverson and Sheridan (2014) questioned whether making really improves STEM learning. Recent research shows that STEM makerspaces in the school curriculum have potential as an instructional strategy to support integrated STEM education and the acquisition of 21st century learning skills. In a meta analysis of STEM makerspaces in K-12 education (Schad and Jones, 2020, p. 69–70), eight of 14 studies "suggested benefits for maker-centered learning for STEM content areas" that included "improved motivation, positive attitude toward STEM subjects, and improved learning outcomes in specific content areas." Other outcomes of maker-centered learning revealed by the findings of the meta-analysis were in the areas of equity, creativity, and independent thinking. There are, however, very few studies that have quantitatively investigated the effect of STEM makerspaces on the learning of STEM concepts. Litts et al. (2017) and Jaipal-Jamani (2021) found that STEM makerspaces did result in statistically significant improvement in learning of targeted STEM concepts (e.g., knowledge and design of circuits in Litts et al. study; knowledge of gears in Jaipal-Jamani study). In another recent study, Sheffield et al. (2017) investigated how a makerspace engaged females ($n = 71$) in STEM education at an elementary girl's school. The results showed the potential for a makerspace approach to shift the focus of STEM as a cognitive endeavor for conceptual learning to STEM as inquiry leaning with intrinsic affective and motivational factors.

Successful makerspace implementation as and for STEM depends on the pedagogical skill and comfort level of the teacher to take on the role of facilitator of learning and embrace the flexible nature of the role of teacher and students. Teachers should be provided with exemplars of effective STEM makerspaces and the pedagogical strategies to facilitate learning that suits the needs of individual learners, give students choice, and enables student ownership of learning, yet at the same time, can be assessed for integrated STEM curriculum outcomes. As Fleming (2015, p. 13) succinctly advises, “Talk to your students. They are your stakeholders. Invite students to discuss their interests and share their opinions. This model of integrated decision making, focused on the goal of maximizing the impact of your makerspace, will help ensure a student-centered learning environment”.

Conclusion

Robotics and makerspace are robust instructional approaches that support STEM in different ways such as motivating learners to engage in STEM learning and pursue STEM careers, to improving understanding of concepts and practices in the different STEM disciplines. The literature shows that both instructional approaches involve the learner actively participating in knowledge creation, design, and hands-on construction of artifacts through inquiry, PBL, and engineering design, while drawing on the knowledge and skills from the disciplinary areas of STEM. These are viable instructional strategies to incorporate into disciplinary subjects such as math and science to promote integrated STEM education through lessons or an integrated unit topic. These two approaches can also be used as independent contexts to promote integrated STEM curriculum outcomes. Robotics and makerspace approaches are also grounded in constructivism and constructionism learning theories and can be readily incorporated into preservice teacher education and inservice teacher professional development. Teachers who have been introduced to these approaches enhance the degree of scalability to integrate it across K-12 education.

Widespread implementation of robotics and makerspace in many countries may be limited as it is contingent on funding. Another challenge to implementation arises when engineering design is the central activity in robotics and makerspace experiences. Students may be subject to some of the challenges experienced by students in engineering design. Three challenges that students experience during engineering design that will also apply to integrated STEM learning with robotics and makerspace are cognitive limits of attention and memory, learning from real world solutions and making connections between multiple representations (National Academies of Engineering and Medicine, 2020). The National Academies of Sciences Engineering and Medicine (2020) went on to state that “repeated experience with ideas that cross STEM boundaries in the context of engineering design activities can reduce challenges related to cognitive limits, real-world problem solving, and multiple representations” (p. 61). They suggest the following strategies for how educators can alleviate these three challenges: scaffolding design tasks from well-defined to moderately ill defined to minimize cognitive load, providing just-in time knowledge of science and math concepts to show connections to the real world, and highlighting how models used as representations in the different STEM field are similar and different.

Current approaches to STEM Education increasingly conceptualize STEM as an integrated field where the knowledge and practices from the different STEM disciplines are applied to solve real world problems which parallels how professional engineers and scientists draw on multiple STEM disciplines to find solutions (Czerniak and Johnson, 2014; National Research Council, 2014). While educators implement integrated STEM programs within the parameters outlined in country, state or provincial curriculum standards, the emergence of new careers and multidisciplinary, innovative fields that require knowledge of crosscutting STEM concepts and its applications (e.g., bio-engineering, artificial intelligence, and gene editing), point to the potential for STEM to be enacted as a distinct field. The literature suggests that robotics and makerspaces can act as and for STEM education. When the two approaches are used in disciplines such as science and math they are support for STEM education (interdisciplinary and transdisciplinary STEM approaches). When robotics and makerspace act as the context for STEM outcomes such as bioengineering, they can be conceptualized as STEM education.

Definitions

STEM Education: is applying the disciplines of science, technology, engineering and mathematics through multidisciplinary, interdisciplinary and transdisciplinary learning approaches.

Integrated STEM Education: is a learning approach that involves understanding a concept, solving a real world problem or designing and creating an artifact by applying the knowledge and practices/skills from the different STEM disciplines.

Multidisciplinary STEM Learning: is when disciplinary knowledge and skills are learned separately but around a common theme/topic.

Interdisciplinary STEM Learning: is when knowledge and skills from two or more disciplines are applied at the same time to gain understanding of a concept or topic.

Transdisciplinary STEM Learning: begins with a real world problem or project and knowledge and skills from two or more of the STEM disciplines are applied to solve the problem/project.

STEM Literacy: is having knowledge, attitudes, and skills to ask questions, problem solve, explain phenomena, and draw conclusions from evidence to reflect and make decisions on STEM-related issues.

Makerspaces: are physical spaces where community members create, design, make and repurpose artifacts or engage in projects using physical and digital technologies.

References

- Ames, R.T., 2014. A Survey of Utah's Public Secondary Education Science Teachers to Determine Their Feelings of Preparedness to Teach Engineering Design. Utah State University, Logan, Utah.
- Anderson, C., 2012. *Makers: The New Industrial Revolution*. Crown, New York.
- Anwar, S., Bascou, N.A., Menekse, M., Kardgar, A., 2019. A systematic review of studies on educational robotics. *J. Pre-Coll. Eng. Educ. Res.* 9 (2), 2.
- Ardito, G., Mosley, P., Scollins, L., 2014. We, robot: using robotics to promote collaborative and mathematics learning in a middle school classroom. *Middle Grades Res. J.* 9 (3).
- Atmatzidou, S., Demetriadis, S., Nika, P., 2018. How does the degree of guidance support students' metacognitive and problem solving skills in educational robotics? *J. Sci. Educ. Technol.* 27 (1), 70–85.
- Bascou, N.A., Menekse, M., 2016. Robotics in K-12 formal and informal learning environments: a review of literature. In: Paper Presented at 2016 ASEE Annual Conference and Exposition, New Orleans, Louisiana.
- Belland, B.R., Walker, A.E., Kim, N.J., Lefler, M., 2017. Synthesizing results from empirical research on computer-based scaffolding in STEM education: a meta-analysis. *Rev. Educ. Res.* 87 (2), 309–344.
- Benitti, F.B.V., Spolaôr, N., 2017. How have robots supported STEM teaching? In: Khine, M. (Ed.), *Robotics in STEM Education*. Springer, Cham.
- Benitti, F.B.V., 2012. Exploring the educational potential of robotics in schools: a systematic review. *Comput. Educ.* 58 (3), 978–988.
- Bevan, B., Petrich, M., Wilkinson, K., 2014. Tinkering is serious play. *Educ. Leader.* 72 (4), 28–33.
- Bevan, B., Gutwill, J.P., Petrich, M., Wilkinson, K., 2015. Learning through STEM-rich tinkering: findings from a jointly negotiated research project taken up in practice. *Sci. Educ.* 99 (1), 98–120.
- Bevan, B., 2017. The promise and the promises of Making in science education. *Stud. Sci. Educ.* 53 (1), 75–103.
- Blackley, S., Sheffield, R., Maynard, N., Koul, R., Walker, R., 2017. Makerspace and reflective practice: advancing pre-service teachers in STEM education. *Aust. J. Teach. Educ.* 42 (3), 22 (Online).
- Bonnette, R.N., Crowley, K., Schunn, C.D., 2019. Falling in love and staying in love with science: ongoing informal science experiences support fascination for all children. *Int. J. Sci. Educ.* 41 (12), 1626–1643.
- Bullock, S.M., Sator, A.J., 2015. Maker pedagogy and science teacher education. *J. Can. Assoc. Curric. Stud.* 13 (1), 60–87.
- Bybee, R., Landes, N.M., 1990. Science for life and living: an elementary school science program from Biological Sciences Improvement Study (BSCS). *Am. Biol. Teach.* 52 (2), 92–98.
- Bybee, R., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teach.* 70 (1), 30–35.
- Bybee, R.W., 2013. *The Case for STEM Education: Challenges and Opportunities*. NSTA press, Arlington, VA.
- Carnevale, A.P., Smith, N., Melton, M., 2011. *STEM: Science Technology Engineering Mathematics*. Georgetown University Center on Education and the Workforce.
- Chung, C.C., Cartwright, C., Cole, M., 2014. Assessing the impact of an autonomous robotics competition for STEM education. *J. STEM Educ. Innovations Res.* 15 (2), 24.
- Czerniak, C.M., Johnson, C.C., 2014. Interdisciplinary science teaching. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, first ed., vol. II. Routledge, New York, pp. 409–425.
- Darling-Hammond, L., 2008. Teacher learning that supports student learning. In: *Teaching for Intelligence*, vol. 2, pp. 91–100, 1.
- Dougherty, D., 2012. The maker movement. *Innovations* 7 (3), 11–14.
- Dougherty, D., 2013. The maker mindset. In: Honey, M., Kanter, D.E. (Eds.), *Design. Make. Play. Growing the Next Generation of STEM Innovators*. Routledge, New York, NY, pp. 7–16.
- Duit, R., 1996. The constructivist view in science education—what it has to offer and what should not be expected from it. *Investigações em ensino de ciências* 1 (1), 40–75.
- Educause, 2013. ELI: 7 things you should know about makerspaces. *Educause Learn. Init.* 1–2.
- Fadigan, K.A., Hammrich, P.L., 2004. A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program. *J. Res. Sci. Teach.* 41 (8), 835–860.
- Fensham, P.J., 2009. Real world contexts in PISA science: implications for context-based science education. *J. Res. Sci. Teach.* 46 (8), 884–896.
- Fleming, L., 2015. *Worlds of Making: Best Practices for Establishing a Maker Space at Your School*. Corwin, Thousand Oaks, CA.
- Gura, M., 2012. Lego robotics: STEM sport of the mind. *Learn. Lead. Technol.* 40 (1), 12–16.
- Gutwill, J.P., Hido, N., Sindorf, L., 2015. Research to practice: observing learning in tinkering activities. *Curator* 58 (2), 151–168.
- Habig, B., Gupta, P., Levine, B., Adams, J., 2020. An informal science education program's impact on STEM major and STEM career outcomes. *Res. Sci. Educ.* 50 (3), 1051–1074.
- Halverson, E., Sheridan, K., 2014. The maker movement in education. *Harv. Educ. Rev.* 84 (4), 495–505.
- Holmlund, T.D., Lesseig, K., Slavitt, D., 2018. Making sense of “STEM education” in K-12 contexts. *Int. J. STEM Educ.* 5 (1), 1–18.
- Hsu, Y.C., Baldwin, S., Ching, Y.H., 2017. Learning through making and maker education. *TechTrends* 61 (6), 589–594.
- Hudson, M.A., Baek, Y., Ching, Y.H., Rice, K., 2020. Using a multifaceted robotics-based intervention to increase student interest in STEM subjects and careers. *J. STEM Educ. Res.* 3 (3), 295–316.
- Hughes, J.M., 2017. Digital making with “At-Risk” youth. *Int. J. Inf. Learn. Technol.* 34 (2), 102–113.
- Jaipal-Jamani, K., 2021. Comparing Preservice Teachers Learning of Science in a Maker and a Scaffolded Robotics Intervention. In: Langran, E., Rutledge, D. (Eds.), *Proceedings of SITE Interactive Conference. Association for the Advancement of Computing in Education (AACE)*, pp. 297–300. Online, United States.
- Kafai, Y., Fields, D., Searle, K., 2014. Electronic textiles as disruptive designs: supporting and challenging maker activities in schools. *Harv. Educ. Rev.* 84 (4), 532–556.
- Karp, T., Maloney, P., 2013. Exciting young students in grades K-8 about STEM through an afterschool robotics challenge. *Am. J. Eng. Educ.* 4 (1), 39–54.
- Kazakoff, E.R., Sullivan, A., Bers, M.U., 2013. The effect of a classroom-based intensive robotics and programming workshop on sequencing ability in early childhood. *Early Child. Educ. J.* 41 (4), 245–255.
- Kelley, T.R., Wicklein, R.C., 2009. Teacher challenges to implement engineering design in secondary technology education. *J. Ind. Teach. Educ.* 46 (3), 34–50.
- Kelly, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (11).
- La Paglia, F., Rizzo, R., La Barbera, D., 2011. Use of robotics kits for the enhancement of metacognitive skills of mathematics: a possible approach. *Stud. Health Technol. Inf.* 167, 26–30.
- Lindh, J., Holgersson, T., 2007. Does lego training stimulate pupils' ability to solve logical problems? *Comput. Educ.* 49 (4), 1097–1111.
- Litts, B.K., Kafai, Y.B., Lui, D.A., Walker, J.T., Widman, S.A., 2017. Stitching codeable circuits: high school students' learning about circuitry and coding with electronic textiles. *J. Sci. Educ. Technol.* 26 (5), 494–507.
- Lucas, J., 2014. What is Engineering? Types of Engineering. *LIVESCIENCE*. <https://www.livescience.com/>.
- Ma, Y., Williams, D.C., 2013. The potential of a first LEGO league robotics program in teaching 21st century skills: an exploratory study. *J. Educ. Technol. Dev. Exc.* 6 (2), Art 2.
- Maker Media, 2013. *The Makerspace Playbook. School Edition*. <https://makered.org/resources/spaces-places/>.
- Martin, L., Dixon, C., Betser, S., 2018. Iterative design toward equity: youth repertoires of practice in a high school maker space. *Equity and Excell. Educ.* 51 (1), 36–47.
- Martin, L., 2015. The promise of the maker movement for education. *J. Pre-Coll. Eng. Educ. Res.* 5 (1), 4.
- Martinez, S.L., Stager, G., 2013. *Invent to Learn. Making, Tinkering, and Engineering in the Classroom*. Torrance, Canada. Constructing Modern Knowledge Press.

- Master, A., Cheryan, S., Moscatelli, A., Meltzoff, A.N., 2017. Programming experience promotes higher STEM motivation among first-grade girls. *J. Exp. Child Psychol.* 160, 92–106.
- McLoughlin, E., Butler, D., Kaya, S., Costello, E., 2020. STEM Education in Schools: What Can We Learn from the Research? Dublin City University, Ireland. ATS STEM Report #1.
- Miller, K., Sonnert, G., Sadler, P., 2018. The influence of students' participation in STEM competitions on their interest in STEM careers. *Int. J. Sci. Educ.* 8 (2), 95–114.
- Ministry of Education, 2022. Ontario Curriculum, Grades 1–8: Science and Technology. Online at <https://www.dcp.edu.gov.on.ca/en/curriculum/science-technology>
- Mitnik, R., Nussbaum, M., Soto, A., 2008. An autonomous educational mobile robot mediator. *Aut. Robots* 25 (4), 367–382.
- Moore, T.J., Smith, K.A., 2014. Advancing the state of the art of STEM integration. *J. STEM Educ. Innovations Res.* 15 (1), 5.
- National Academies of Sciences, Engineering, and Medicine, 2020. Building Capacity for Teaching Engineering in K-12 Education. The National Academies Press, Washington, DC. <https://doi.org/10.17226/25612>.
- National Research Council, 2012. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. The National Academies Press, Washington, DC.
- National Research Council, 2014. STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research. The National Academies Press, Washington, DC.
- NGSS Lead States, 2013. Next Generation Science Standards: For States, by States. The National Academies Press, Washington, DC.
- Nugent, G., Barker, B., Grandgenett, N., 2008. The effect of 4-H robotics and geospatial technologies on science, technology, engineering, and mathematics learning and attitudes. In: Luca, J., Weippl, E. (Eds.), *Proceedings of ED-MEDIA 2008—World Conference on Educational Multimedia, Hypermedia & Telecommunications*. Association for the Advancement of Computing in Education (AACE), Vienna, Austria, pp. 447–452.
- Nugent, G., Barker, B., Grandgenett, N., Adamchuk, V.I., 2010. Impact of robotics and geospatial technology interventions on youth STEM learning and attitudes. *J. Res. Technol. Educ.* 42 (4), 391–408.
- Nugent, G.C., Barker, B.S., Grandgenett, N., 2014. The impact of educational robotics on student STEM learning, attitudes, and workplace skills. In: Khosrow-Pour, M. (Ed.), *Robotics: Concepts, Methodologies, Tools, and Applications*. GI Global, Hershey, Pennsylvania, pp. 1442–1459.
- O'Brien, S., Hansen, A.K., Harlow, D.B., 2016. Educating teachers for the maker movement: pre-service teachers' experiences facilitating maker activities. In: *Proceedings of the 6th Annual Conference on Creativity and Fabrication in Education*. Association for Computing Machinery, New York, NY, pp. 99–102.
- Orpwood, G.W., Schmidt, B.A., Hu, J., 2012. Competing in the 21st Century Skills Race. Canadian Council of Chief Executives, Ottawa, Ca.
- Papert, S., 1991. Situating constructionism. In: Harel, I., Papert, S. (Eds.), *Constructionism*. Ablex, Norwood, NJ, pp. 1–11.
- Park, S.H., Ertmer, P.A., 2008. Examining barriers in technology-enhanced problem-based learning: using a performance support systems approach. *Br. J. Educ. Technol.* 39 (4), 631–643.
- Park, W., Wu, J.Y., Erduran, S., 2020. The nature of STEM disciplines in the science education standards documents from the USA, Korea and Taiwan. *Sci. Educ.* 29 (4), 899–927.
- Park, J., 2015. Effect of Robotics enhanced inquiry based learning in elementary Science education in South Korea. *J. Comput. Math. Sci. Teach.* 34 (1), 71–95.
- Quigley, C.F., Herro, D., 2016. Finding the joy in the unknown: implementation of STEAM teaching practices in middle school science and math classrooms. *J. Sci. Educ. Technol.* 25 (3), 410–426.
- Renninger, K.A., Hidi, S., 2011. Revisiting the conceptualization, measurement, and generation of interest. *Educ. Psychol.* 46 (3), 168–184.
- Robinson, M., 2005. Robotics-driven activities: can they improve middle school science learning. *Bull. Sci. Technol. Soc.* 25 (1), 73–84.
- Rogers, C., Portsmore, M.D., 2004. Bringing engineering to elementary school. *J. STEM Educ.* 5 (3 & 4), 17–28.
- Rusk, N., Resnick, M., Berg, R., Pezalla-Granlund, M., 2008. New pathways into robotics: strategies for broadening participation. *J. Sci. Educ. Technol.* 17 (1), 59–69.
- Schad, M., Jones, W.M., 2020. The maker movement and education: a systematic review of the literature. *J. Res. Technol. Educ.* 52 (1), 65–78.
- Schnittka, C., Bell, R., 2011. Engineering design and conceptual change in science: addressing thermal energy and heat transfer in eighth grade. *Int. J. Sci. Educ.* 33 (13), 1861–1887.
- Sheffield, R., Koul, R., Blackley, S., Maynard, N., 2017. Makerspace in STEM for girls: a physical space to develop twenty-first-century skills. *Educ. Media Int.* 54 (2), 148–164.
- Sheridan, K., Halverson, E.R., Litts, B., Brahms, L., Jacobs-Priebe, L., Owens, T., 2014. Learning in the making: a comparative case study of three makerspaces. *Harv. Educ. Rev.* 84 (4), 505–531.
- Siegel, D., Giamellaro, M., 2020. Defining STEM within a school district: a co-constructed and evolving process. *Cult. Stud. Sci. Educ.* 15 (3), 739–773.
- Sullivan, F.R., Heffernan, J., 2016. Robotic construction kits as computational manipulatives for learning in the STEM disciplines. *J. Res. Technol. Educ.* 48 (2), 105–128.
- Sullivan, F.R., 2008. Robotics and science literacy: thinking skills, science process skills and systems understanding. *J. Res. Sci. Teach.* 45 (3), 373–394.
- Tai, R.H., Liu, C.Q., Maltese, A.V., Fan, X., 2006. Planning early for careers in science. *Life Sci.* 1, 0–2.
- Tang, K.S., Williams, P.J., 2019. STEM literacy or literacies? Examining the empirical basis of these constructs. *Rev. Educ.* 7 (3), 675–697.
- Tytler, R., Osborne, J., 2012. Student attitudes and aspirations towards science. In: Fraser, B., Tobin, K. (Eds.), *Second International Handbook of Science Education*. Springer, Dordrecht, pp. 597–625.
- Tytler, R., Williams, G., Hobbs, L., Anderson, J., 2019. Challenges and opportunities for a STEM interdisciplinary agenda. In: Doig, B., Williams, J., Swanson, D., Borromeo Ferri, R., Drake, P. (Eds.), *Interdisciplinary Mathematics Education: The State of the Art and beyond*. Springer, Cham, pp. 51–81.
- Tytler, R., 2020. STEM Education for the twenty-first century. In: Anderson, J., Li, Y. (Eds.), *Integrated Approaches to STEM Education: An International Perspective*. Springer, Cham, Switzerland, pp. 21–43.
- Vasquez, J.A., Sneider, C.I., Comer, M.W., 2013. *STEM Lesson Essentials, Grades 3–8: Integrating Science, Technology, Engineering, and Mathematics*. Heinemann, Portsmouth, NH.
- Vollstedt, A.M., Robinson, M., Wang, E., 2007. Using robotics to enhance science, technology, engineering, and mathematics curricula. In: *Proceedings of American Society for Engineering Education Pacific Southwest Annual Conference*. Hawaii, Honolulu.
- Warde, C., Sah, D., 2019. Creating a STEM-based economic pillar for the Caribbean: a blueprint. *J. Educ. Dev. Caribb.* 18 (1), 234–270.
- Wigfield, A., Cambria, J., 2010. Students' achievement values, goal orientations, and interest: definitions, development, and relations to achievement outcomes. *Dev. Rev.* 30, 1–35.
- Williams, D.C., Ma, Y., Prejean, L., Ford, M.J., Lai, G., 2007. Acquisition of physics content knowledge and scientific inquiry skills in a robotics summer camp. *J. Res. Technol. Educ.* 40 (2), 201–216.
- Williams, K., Igel, I., Poveda, R., Kapila, V., Iskander, M., 2012. Enriching K–12 science and mathematics education using LEGOs. *Adv. Eng. Educ.* 3 (2).
- Williams, J., Roth, W.M., Swanson, D., Doig, B., Groves, S., et al., 2016. *Interdisciplinary Mathematics Education: A State of the Art*. Springer Open, Hamburg, Germany.
- Zeidler, D.L., 2016. STEM education: a deficit framework for the twenty first century? A sociocultural socioscientific response. *Cult. Stud. Sci. Educ.* 11 (1), 11–26.
- Ziaefard, S., Miller, M.H., Rastgaar, M., Mahmoudian, N., 2017. Co-robotics hands-on activities: a gateway to engineering design and STEM learning. *Robot. Autonom. Syst.* 97, 40–50.

Models and modeling in STEM education: nature, roles, and implementation

Jonas Hallström^a and Konrad J. Schönborn^b, ^a Technology and Science Education Research (TESER), Department of Behavioural Sciences and Learning, Linköping University, Norrköping, Sweden; and ^b Division of Media and Information Technology, Department of Science and Technology, Linköping University, Norrköping, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Introduction	112
The nature of models and their representation	112
Roles of models in relation to specific modeling processes, knowledge and skills	113
Implementation of models and modeling in STEM education	114
Strategies for teaching models and modeling in STEM education	114
Integration of models and modeling in STEM curricula and practice	114
Conclusions and implications for STEM education	115
References	115

Introduction

In a 2019 commentary, we reinforced the argument that models and modeling could serve as a bridge between the STEM subjects in educational practice (Hallström and Schönborn, 2019). This synthesis emphasized that models and modeling are central to fostering authentic STEM education. The subjects of science, technology, engineering and mathematics are often subsumed into the notion of “STEM”, an acronym used particularly in the last two decades to define content, craft curricular goals, define educational policy and/or motivate political agendas. However, a precise interpretation and operational definition of STEM remains elusive (Breiner et al., 2012) with various camps of scholars suggesting that it is often used in far too general or in far too specific terms (e.g., Henderson et al., 2017; Tang and Williams, 2019). Another saliently discussed attribute is the relative contribution of each “S”, “T”, “E” and “M” subject to the overall STEM construct, with contemporary developments also advocating the inclusion of an “A” for artistic subjects (STEAM), and “i” for information technologies (iSTEAM). Such discussion calls for further and clearer considerations of the nature, roles and implementation of STEM, and how the subjects can be meaningfully linked. In this regard, models and modeling act as point of departure for linking the STEM subjects, thereby contributing to the discussion of the meaning, integration and application of STEM education.

Models and modeling are crucial for solving problems, making predictions, and communicating concepts and constructs (e.g., Müller, 2009; Vincenti, 1990). Models clarify aspects of reality and range from simple conceptual diagrams and early prototypes to advanced mathematical models and machine learning algorithms. Therefore, the competencies needed to create, use, apply, evaluate and revise models are necessary for gaining an in-depth understanding of scientific practice, technological and engineering design, and mathematical tools (Schwarz et al., 2009).

In this article we elaborate upon the nature, roles, and implementation of models and modeling in STEM education, by posing the following questions:

- What is the nature of models and how are they represented?
- What are the roles of models in relation to specific modeling processes, knowledge and skills?
- How can models and modeling be implemented in STEM education?

In pursuit of implementing models and modeling across the four STEM disciplines, the relationship between the nature, roles and implementation of models should be seen as being both intertwined as well as informing one another. The following sections unpack these three key aspects of models and modeling in STEM education by responding to each of the questions that frame the article.

The nature of models and their representation

According to Gilbert et al. (2000), models are *ubiquitous representations of phenomena* such as objects, ideas, concepts, systems, events, and processes. Models describe, explain, construct, and predict phenomena by organizing relevant information, generating hypotheses, and explaining how information may be related (Brady et al., 2015; Lesh et al., 1997, 2013). In doing so, models communicate a simplified version of reality through concrete, conceptual, and formal/symbolic depictions (de Vries, 2013). In this regard, models portray entities in smaller (e.g., a spaceship blueprint) or larger (e.g., a coronavirus diagram) visual forms than the actual represented phenomenon, as abstractions (e.g., force field arrows), or as depictions of concrete and abstract entities (e.g., augmented predicted visual path of a photographed projectile) (Gilbert et al., 2000). Furthermore, models represent dynamic systems such as cause-and-effect relationships as well as in coordinated depictions where the system is communicated as more than the “sum of its parts” (Zawojewski et al., 2008; Diefes-Dux et al., 2008).

Different *modes of representation* are used to communicate a model (Schönborn and Anderson, 2009). Herein, a model can be expressed in at least five different modes of representation, namely concrete (e.g., 3D physical models), verbal (e.g., spoken or written description of model entities); symbolic (e.g., chemical formulae and algorithms), visual (e.g., diagrams and animations), or gestural (e.g., bodily representation of model entities) (Gilbert, 2004). While in science education models are often communicated through concrete, visual, verbal, mathematical, or gestural representations, in technology education models often include iconic (e.g., a sketch), analogue (e.g., simulation), and symbolic (e.g., mathematical) models (Davies and Gilbert, 2003).

Models in science are often central to knowledge building, and fundamental for providing explanations and predictions. Scientific models can comprise consensus models (used in research), which in turn serve to illustrate the nature of science, or historical models (replaced by revised models over time) (Justi and Gilbert, 2002). Nia and de Vries (2017) consider models as “techno-scientific artifacts”. Herein, they have an intrinsic nature, intentional nature, as well as an intrinsic-intentional interconnection. In this regard, Nia and de Vries (2017) have proposed a framework for the “dual nature” (cf. de Vries and Meijers, 2013) of models which may be applied across the STEM subjects. Nia and de Vries (2017) describe the “intrinsic” nature of models as concerning the material structure and form of models, whereas the “intentional” nature of models concerns their purposes or functions, that is, whether they are used for exploration, design or communication. In technology and engineering, knowledge can be developed by constructing and manipulating models, wherein models are used to comprehend design concepts and optimize prototypes (Citrohn and Svensson, 2020; France et al., 2011; France, 2018). In mathematics, a model can be defined as the combination of an extra-mathematical domain, a mathematical domain, and the translation between the two (Niss, 2012).

Roles of models in relation to specific modeling processes, knowledge and skills

Justi and Gilbert (2002) have advocated that models and modeling play a central role both in *learning science* and *learning how to do science*. The modeling process includes determining the purpose of the model, which often includes formulating a mental model of the phenomenon, and ascertaining in what mode of representation (e.g., visual, concrete or symbolic) to express the model. Subsequently, while “testing” whether a model satisfies its foreseen purpose, the modeller infers the scope and limitations of the model.

Regarding the specific roles of models for learning science, Gilbert (2004) states that students should understand what a model is; understand the entities that a particular model represents and how these entities interact with each other; mentally visualize models; display visual literacy skills associated with interpreting models; understand analogy and metaphor in relation to describing model components; and, understand how a model can be used. According to Gilbert et al. (2000) modeling can include several different model types with different roles: mental models (cognitive representations); expressed models (available for others to interpret); consensus models (expressed models that gain acceptance); scientific models (tested expressed models that become predictive tools); historical models (exist in a context and perhaps later displaced); curricular models (historical models in curricula); teaching models (aid interpreting historical and curricular models); and, hybrid models (coordinate scientific, historical or curricular models).

It follows that two roles of modeling in science and design & technology education are modeling ideas in the mind – communicating with oneself – and modeling ideas in the world – communicating with others. The modeling process includes having experience of the phenomenon or problem; formulating suitable metaphors and analogies to express the model; visualizing the outcome of the modeling process; producing a representation of the model, and, evaluating the scope and limitations of the produced model (Davies and Gilbert, 2003).

In technology and engineering education, models support development of theories and artifacts through manipulation (e.g., concrete models) and mental exploration (e.g., conceptual models, sketches). Furthermore, the intrinsic and intentional nature of models in technology and engineering can support building, revising, and communicating knowledge and artifacts related to pedagogical use (e.g., educational models); procedural use (e.g., computer-aided design); and, decisional use (e.g., climate change models). The intrinsic-intentional interrelation of models informs design (designers’ and users’ points of view); simplification (abstraction and idealization); iterativity (trial and error); and, adequacy (judging appropriateness and effectiveness) (Nia and de Vries, 2017; de Vries, 2013). Another aspect of modeling in technology and engineering education is functional modeling, which concerns the development of a design concept and prototyping of the realized outcome, often as an artifact/system. Technological modeling thus provides epistemic strategies for reinforcing that the technological outcome is “fit for purpose”; the purpose is a designed intervention where the outcome is judged by a successful function (France et al., 2011; France, 2018).

In engineering education modeling processes are often similar to that of technology education, but principles from mathematics education for so-called model-eliciting activities (MEAs) may also include model construction, reality, self-assessment, model documentation, model shareability and re-usability, and production of an effective prototype (Zawojewski et al., 2008; Diefes-Dux et al., 2008). Model-eliciting activities (MEAs) in mathematics and engineering education thus include posing the following questions: is the situation authentic? (the reality principle); is the construction or modification of a model required? (the model construction principle); are there clear criteria for assessing the usefulness of the model? (the self-evaluation principle); does the model apply to multiple situations? (the model generalization principle); and, will the solution provide a useful prototype for interpreting other similar situations? (the simple prototype principle) (Brady et al., 2015; Lesh et al., 1997, 2013).

Implementation of models and modeling in STEM education

Strategies for teaching models and modeling in STEM education

To promote a model-based teaching approach, science and technology educators need to communicate to learners what representational components make up a model, demonstrate the scope and limitations of different models, adapt model usage depending on the content taught, and, design meaningful activities that include learners' active model construction (Gilbert, 2004). Justi and Gilbert (2002) suggest the following modeling strategies during teaching: communicate the purpose of a particular modeling activity; provide an authentic experience of the phenomenon being modeled during any practical work; specify the model source; support the mental imagery of a particular model; and, show relationships between different modes of representation of the model (cf. Davies and Gilbert, 2003). It follows that modeling in science and technology involves both a developmental cycle where iterative changes are associated with the produced model, a "fitness for purpose" where a certain specification is envisioned, as well as, a visualization of the intended outcome of the process. For Gilbert et al. (2000), teaching models and modeling contributes to learning because mental modeling is central to understanding. Moreover, expressing and testing models also reflect the "doing" of science, and understanding science relies on interpreting scientific and historical models.

According to de Vries (2013), fusing modeling and design activities enhances STEM learning since design connects scientific, technological, engineering and mathematical components. In addition, modeling activities provide a bridge between a practical scenario and required mathematical analytical tools to model different aspects of reality. In particular, this is the case in problems when understanding reality (science) and manipulating reality (technology and engineering) are foreseen. Developing students' conceptual understanding of modeling and their intrinsic-intentional perspective of models, can foster an enhanced individual modeling ability. In this way modeling can be used as a pedagogical strategy to support STEM learning (France et al., 2011; France, 2018; Nia and de Vries, 2017).

Effective teaching of models and modeling requires paying concerted attention to the design of activities, with adequate time for the activities to be realized (Niss, 2012). In mathematics, students should be provided with an opportunity to experience how mathematical models come to be, and interrogate the trade-offs involved in developing a mathematical model, including assessing the limitations and strengths of different models. When students engage in model-eliciting activities to assess and monitor their own work using authentic tools, they induce the construction, modification, and refinement of powerful conceptual models. Overall, modeling abilities in STEM can be developed along various dimensions that include: from concrete to abstract, from specific to general, from local/refined to global, or from intuitions to formalisations (Brady et al., 2015; Lesh et al., 1997, 2013; Zawojewski et al., 2008; Diefes-Dux et al., 2008).

Integration of models and modeling in STEM curricula and practice

According to Gilbert (2004), integrating models and modeling can increase the authenticity of STEM curricula by explicitly training science teachers in the nature and role of models. Integrating models and modeling in STEM curricula can be promoted through pupils learning to use models, revise models, reconstruct models, and construct models *de novo*. Pre- (and in-) service STEM education should also focus heavily on unpacking the nature of a "model" and how to use different models in different contexts (transfer). This can be promoted by following the historical sequence of model development in a certain topic area as a means of cognitive reconstruction in modeling, and, providing skills for evaluating the strengths and limitations of models (Justi and Gilbert, 2002).

Furthermore, implementing models and modeling to embrace STEM education requires considering the importance of the learning context - curricula need to be enterprising, effective, of good quality and personally relevant (Margot and Kettler, 2019). Such curricula will allow learners to perceive relationships between science and technology content while applying knowledge in real-world problem solving (Davies and Gilbert, 2003). Models and modeling should therefore be viewed by educators and curriculum developers as a valuable bridge between science, technology, engineering and mathematics, which in turn, may promote authentic STEM education. Furthermore, there must be opportunities for students to develop activities to find solutions to, and make informed decisions about real-world issues by integrating both science and technology concepts (Gilbert et al., 2000).

For de Vries (2013), the encompassing facets of modeling allow it to be integrated at various points of student development in STEM curricula. Primary students could, for example, be provided with early experiences of modeling through concrete models, whereas secondary students could engage with formal aspects of modeling, which include nature, types and functions of models and modeling. The intrinsic, intentional and intrinsic-intentional perspective of models can, according to Nia and de Vries (2017), be used to analyze curricula and policy documents on the integration of models and modeling. These perspectives are a potential benchmark of what a curriculum should contain with respect to models and modeling in the STEM subjects. Moreover, modeling can support the development of an understanding of the nature of technology and the nature of science both as separate domains, but also in relationship with one other as well as engineering and mathematics (France et al., 2011; France, 2018).

According to Niss (2012), teaching about modeling should be used to support students' concept formation and sense-making in mathematics, especially when it comes to the transfer to authentic problems. Thus, modeling requires students to actively engage multiple skills (Niss, 2012). Implementing models and modeling in authentic learning programmes can help students construct, modify and refine conceptual models that are applicable not only to mathematics but also to other modeling adaptation activities in engineering, technology and science (Brady et al., 2015; Lesh et al., 1997, 2013). In this way, students are potentially better positioned to understand the strengths and weaknesses of a conventional model, and better prepared to apply, adapt, and even create new models for novel and similar situations, across the STEM subjects (Zawojewski et al., 2008; Diefes-Dux et al., 2008; Kertil and Gurel, 2016).

Conclusions and implications for STEM education

This article has elaborated upon the nature, roles, and implementation of models and modeling in STEM education. Regarding the nature of modeling, although models and modeling differ slightly between the STEM subjects there are some clear similarities, for example, concerning visual models and representations (e.g., Tang and Williams, 2019). Exploring these similarities and bringing differences to light could strengthen STEM education and STEM literacy. Modeling is often about representing simplified versions of reality that take on concrete/physical, conceptual, verbal, gestural or symbolic/mathematical forms (Gilbert, 2004). Models are therefore representations of ideas, objects, systems, events, or processes which are central in the STEM disciplines. At a conceptual level, models are even systems of description in themselves; for explaining, constructing, modifying, manipulating and/or predicting a complex series of experiences. Models thereby help to organize relevant information so as to generate or (re)interpret hypotheses about given situations, designs or processes, or explain how information is related, something which is at the core of all the STEM subjects.

Given the above, we must also be cognisant of the influence of the rapidly developing digital landscape on models and modeling in STEM education. Notably, artificial intelligence (AI), neural networks and machine learning are changing the way models are designed and how modeling is carried out, because some of the modeling is now done not by humans but by computers (Campbell, 2020; Carr, 2015; Wiberg et al., 2019), which are also “learning” to be more proficient modellers. Furthermore, models are increasingly becoming digital and mathematical in conjunction with more complex technological systems, which will also have repercussions on the field of STEM education; if machines learn to model increasingly complex systems, humans may become less involved in, or even lose some of the control over, STEM projects (cf. Hallström, 2022).

When it comes to the roles of models and modeling in STEM education, models support design of artifacts and theory development through manipulation (e.g., concrete models) or mental exploration (e.g., conceptual models, sketches) (de Vries, 2013), and, in the latter case, modeling ideas in the mind for communicating with oneself and modeling ideas in the world for communicating with others (Davies and Gilbert, 2003). Some of the primary skills associated with modeling include: understanding what a model is and how to use it; carefully defining the context of the modeling process; mentally visualizing a model outcome; deciding what mode of representation to express the model; and understanding how a model can be constructed, interpreted, evaluated and revised. A critical ability is also being able to assess the strengths and limitations of a particular model (Schönborn and Anderson, 2009).

Concerning the implementation of models and modeling in STEM education, a lot of evidence points to the relevance of including modeling in authentic engineering design projects. In such engineering projects students will formulate scientific and mathematical models and algorithms to optimize their designs, which are tested with engineering prototypes and functional models. Hence, new knowledge is developed not only about science, mathematics and technology but also about the engineering design process itself; students are required to apply existing knowledge about, for instance, modeling previously learnt in science, technology and/or mathematics (Ammon, 2017; de Vries, 2013; Hallström and Ankiewicz, 2019; Kertil and Gurel, 2016).

Modeling is an indispensable feature of STEM education (see e.g., Banks and Barlex, 2014; Gilbert, 2004; Williams, 2017). Recent studies show that there are advantages in teaching and learning modeling as an integrated STEM literacy. For instance, Krell and Krüger (2017) have shown that STEM students have more advanced “meta-modeling knowledge” than students of other, single disciplines at the tertiary level. Additionally, models and modeling are often seen as a “language” of pursuing an integrative STEM literacy (e.g., Kertil and Gurel, 2016). Albeit so, further research is needed to ascertain how modeling can be developed to purposefully and systematically interconnect the STEM subjects.

In conclusion, this article emphasizes that modeling activities can serve as an important path toward authentic STEM education (e.g., France, 2018; Gilbert et al., 2000). In addition, models and modeling can be used as a vehicle to foster STEM literacy and the transfer of knowledge and skills between STEM subjects (e.g., Niss, 2012). To accomplish this, teaching must take into account modeling frameworks that are based on authentic STEM practices (Breiner et al., 2012). In pursuing this vision, it must be realized that integrating science, technology, engineering and mathematics remains a complex challenge that calls for what Kelley and Knowles (2016) term “a new generation of STEM experts” (p. 1) that can navigate the future expected transitions between the STEM disciplines. Finally, in order to implement and fulfill an integrated STEM literacy, it is crucial that model-based pedagogies intended for STEM education classrooms are empirically investigated and evaluated in real education contexts (Margot and Kettler, 2019).

References

- Ammon, S., 2017. Why designing is not experimenting: design methods, epistemic praxis and strategies of knowledge acquisition in architecture. *Philos. Technol.* 30, 495–520.
- Banks, F., Barlex, D., 2014. *Teaching STEM in the Secondary School: Helping Teachers Meet the Challenge*. Routledge, London.
- Brady, C., Lesh, R., Sevis, S., 2015. Extending the reach of the models and modelling perspective: a course-sized research site. In: Stillman, G.A., et al. (Eds.), *Mathematical Modelling in Education Research and Practice*. Springer International Publishing, Dordrecht, pp. 55–66.
- Breiner, J.M., Harkness, S.S., Johnson, C.C., Koehler, C.M., 2012. What is STEM? A discussion about conceptions of STEM in education and partnerships. *Sch. Sci. Math.* 112 (1), 3–11.
- Campbell, M., 2020. Automated coding: the quest to develop programs that write programs. *Computer* 53 (2), 80–82. <https://doi.org/10.1109/MC.2019.2957958>.
- Carr, N., 2015. *The Glass Cage: How Our Computers are Changing Us*. Norton, New York.
- Citrohn, B., Svensson, M., 2020. Technology teacher's perceptions of model functions in technology education. *Int. J. Technol. Des. Educ.* 1–19. <https://doi.org/10.1007/s10798-020-09632-8>.

- Davies, T., Gilbert, J., 2003. Modelling: promoting creativity while forging links between science education and design and technology education. *Can. J. Sci. Math. Technol. Educ.* 3 (1), 67–82.
- de Vries, M.J., 2013. Modeling in technology and engineering education. In: Williams, P.J., Gedera, D. (Eds.), *PATT 27, Technology Education for the Future: A Play on Sustainability*, Christchurch, New Zealand, 2–6 December, 2013. University of Waikato, Waikato, NZ.
- de Vries, M.J., Meijers, A.W.M., 2013. Beliefs, acceptances and technological knowledge. In: de Vries, M.J., Hansson, S.O., Meijers, A.W.M. (Eds.), *Norms in Technology*. Springer, Dordrecht, pp. 55–65.
- Diefes-Dux, H.A., Hjalmarson, M.A., Miller, T.K., Lesh, R., 2008. Model-eliciting activities for engineering education. In: Zawojewski, J.S., Diefes-Dux, H., Bowman, K. (Eds.), *Models and Modeling in Engineering Education: Designing Experiences for All Students*. Sense Publishers, Rotterdam.
- France, B., 2018. Modeling in technology education: a route to technological literacy. In: de Vries, M.J. (Ed.), *Handbook of Technology Education*. Springer, Dordrecht.
- France, B., Compton, V.J., Gilbert, J.K., 2011. Understanding modelling in technology and science: the potential of stories from the field. *Int. J. Technol. Des. Educ.* 21 (3), 381–394.
- Gilbert, J.K., Boulter, C.J., Elmer, R., 2000. Positioning models in science education and in design and technology education. In: Gilbert, J.K., Boulter, C.J. (Eds.), *Developing Models in Science Education*. Kluwer Academic Publishers, Dordrecht.
- Gilbert, J., 2004. Models and modelling: routes to more authentic science education. *Int. J. Sci. Math. Educ.* 2 (2), 115–130.
- Hallström, J., 2022. Embodying the past, designing the future: technological determinism reconsidered in technology education. *Int. J. Technol. Des. Educ.* 32, 17–31. <https://doi.org/10.1007/s10798-020-09600-2>.
- Hallström, J., Ankiewicz, P., 2019. Laying down the “T” and “E” in STEM education: design as the basis of an integrated STEM philosophy. In: Pulé, S., de Vries, M.J. (Eds.), *Proceedings PATT 37: Developing a Knowledge Economy Through Technology and Engineering Education*, 3–6 June 2019 University of Malta, Msida Campus. University of Malta, Msida, pp. 187–194.
- Hallström, J., Schönborn, K.J., 2019. Models and modelling for authentic STEM education: reinforcing the argument. *Int. J. STEM Edu.* 6, 22. <https://doi.org/10.1186/s40594-019-0178-z>.
- Henderson, C., Connolly, M., Dolan, E.L., Finkelstein, N., Franklin, S., Malcom, S., Rasmussen, C., Redd, K., St John, K., 2017. Towards the STEM DBER Alliance: why we need a discipline-based STEM education research community. *Int. J. STEM Edu.* 4 (14), 1–4.
- Justi, R.S., Gilbert, J.K., 2002. Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *Int. J. Sci. Educ.* 24 (4), 369–387.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Edu.* 3 (1), 11.
- Kertil, M., Gurel, C., 2016. Mathematical modeling: a bridge to STEM education. *Int. J. Educ. Math. Sci. Technol.* 4 (1), 44–55.
- Krell, M., Krüger, D., 2017. University students' meta-modelling knowledge. *Res. Sci. Technol. Educ.* 35 (3), 261–273.
- Lesh, R., Amit, M., Schorr, R.Y., 1997. Using “real-life” problems to prompt students to construct conceptual models for statistical reasoning. In: *The Assessment Challenge in Statistics Education*. IOS Press, pp. 65–83.
- Lesh, R., English, L., Sevis, S., Riggs, C., 2013. Modeling as a means for making powerful ideas accessible to children at an early age. In: Hegedus, S.J., Roschelle, J. (Eds.), *The SimCalc Vision and Contributions*. Springer, Dordrecht, pp. 419–436.
- Margot, K.C., Kettler, T., 2019. Teachers' perception of STEM integration and education: a systematic literature review. *Int. J. STEM Edu.* 6 (2), 1–16.
- Müller, R., 2009. The notion of a model: a historical overview. In: Meijers, A. (Ed.), *Philosophy of Technology and Engineering Sciences*. North Holland, Burlington, MA, pp. 637–664.
- Nia, M.G., de Vries, M.J., 2017. Models as artefacts of a dual nature: a philosophical contribution to teaching about models designed and used in engineering practice. *Int. J. Technol. Des. Educ.* 27 (4), 627–653.
- Niss, M., 2012. Models and modelling in mathematics education. *EMS Newsl.* 86, 49–52.
- Schönborn, K.J., Anderson, T.R., 2009. A model of factors determining students' ability to interpret external representations in biochemistry. *Int. J. Sci. Educ.* 31 (2), 193–232.
- Schwarz, C.V., et al., 2009. Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *J. Res. Sci. Teach.* 46 (6), 632–654.
- Tang, K.S., Williams, P.J., 2019. STEM literacy or literacies? Examining the empirical basis of these constructs. *Rev. Educ.* 7 (3), 675–697.
- Vincenti, W.G., 1990. *What Engineers Know and How They Know It*. Johns Hopkins University Press, Baltimore, Maryland.
- Wiberg, A., Persson, J., Ölvander, J., 2019. Design for additive manufacturing—a review of available design methods and software. *Rapid Prototyp. J.* 25 (6), 1080–1094. <https://doi.org/10.1108/RPJ-10-2018-0262>.
- Williams, J., 2017. The Vocational Goals of STEM Education: Is That Enough? Paper presented at PATT 2017 Technology & Engineering Education – Fostering the Creativity of Youth Around The Globe, Philadelphia, PA, pp. 10–14. July 2017.
- Zawojewski, J.S., Hjalmarson, M.A., Bowman, K.J., Lesh, R., 2008. A modeling perspective on learning and teaching in engineering education. In: Zawojewski, J.S., Diefes-Dux, H., Bowman, K. (Eds.), *Models and Modeling in Engineering Education: Designing Experiences for All Students*. Sense Publishers, Rotterdam.

Aligning PBL with the NGSS: What do the NGSS offer PBL?

Emily Adah Miller^a, Samuel Severance^b, and Joseph Krajcik^c, ^a University of Wisconsin, Madison, WI, United States; ^b University of California, Santa Cruz, CA, United States; and ^c Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	117
Related research and theoretical foundation	118
Promise of project-based learning for supporting student learning	118
Challenge presented by the Next Generation Science Standards	118
Research questions	119
Methodology	119
Context	119
Multiple Literacies in Project-based Learning	119
Context of implementation	120
Approach to design	120
Employing design-based research methods	120
Logic model	121
Data collection	122
Studying implementation of materials	122
Data analysis	122
Design cycles	123
Iterative cycles of design, testing, and revising	123
First cycle teaching experiments	123
Design	123
Testing and analysis	124
Outcomes and reflections from the first design cycle	124
Second cycle pilot studies	124
Design	125
Testing and analysis	125
Outcomes and reflections from the second design cycle	125
Third cycle field experiments	126
Design	126
Testing and analysis	126
Outcomes and reflections toward fourth cycle	127
Discussion	127
Conclusion	129
Acknowledgment	129
References	129

Introduction

The *Framework for K-12 Science Education* (Framework; NRC, 2012) presents a new vision of science teaching and learning in which students are to develop knowledge to make sense of phenomena or design solutions to engineering problems. Because of the uniqueness of what the Framework and the resulting Next Generations of Science Standards (NGSS; NGSS Lead States, 2013) proposed, districts are currently grappling with finding ways to address the NGSS at the elementary level. In this paper, we present one promising solution.

One of the most significant—and challenging—innovations put forth in the Framework and NGSS is a call for “three-dimensional science learning.” Three-dimensional science learning means that students use science and engineering practices, core scientific ideas, and crosscutting concepts that span science disciplines together in an integrated manner to explain phenomena or solve problems. For example, to figure out why some bird species prefer wetlands over prairies, a student would need to use science and engineering practices like carrying out an investigation and analyzing data of a bird’s habitat, while thinking about the phenomenon through the lens of crosscutting concepts like systems and cause and effect, and also applying core ideas about ecology such as interdependent relationships.

Project-based learning (PBL) is a curriculum approach that shows promise in creating environments for meaningful student learning, including three-dimensional science learning (Miller and Krajcik, 2019; Miller et al., 2021b). Researchers describe PBL in different ways, but they always include the following features: (1) An artifact that drives and culminates a cycle of investigation,

(2) Driving Questions that motivate practices and are answered by the artifact, and (3) Student learning has authentic and student-centered application (Barron et al., 1998; Krajcik and Blumenfeld, 2006).

As part of a large design-based research project, our team designed upper elementary project-based learning environments that extended our vision of PBL to three-dimensional learning. Our task was to develop and refine units of instruction that would provide evidence of the value of PBL in promoting three-dimensional science learning and community. This paper reports on the design cycles that led to our final third and fourth grade products, prior to an efficacy study and the testing of fifth grade materials.

Related research and theoretical foundation

Promise of project-based learning for supporting student learning

Project-based learning (PBL) has its roots in the progressive era with Dewey (Dewey, 1902), who advocated a student-centered education approach. Instead of lecture, Dewey believed learning should have practical meaning and direct application. He moved away from teacher-centered to student-centered teaching and learning. Dewey advocated for an educational approach that would engage students in experience as active participants in learning, an approach that he argued promoted their interest and motivation (Weiss et al., 2005).

Despite evidence of the promise of PBL as an effective curricular approach for students, PBL is under researched, especially in elementary education (Duke et al., 2021). Studies support PBL as motivating for students (Harris et al., 2015; Miller et al., 2021a), and that PBL enhances subject matter learning (Condliffe et al., 2017; Geier et al., 2008; Miller et al., 2020b). PBL is defined differently across scholars and program designers (Barron et al., 1998; Blumenfeld et al., 1991), yet all iterations draw on situated cognition (Brown et al., 1989) and constructivism (Bransford and Schwartz, 1999). The lack of consistency in definitions of PBL has caused some confusion in the field (Thomas, 2000). Current descriptions of PBL have the following salient features: (1) The project motivates learning and includes a culminating and “concrete” artifact (Helle et al., 2006); (2) PBL features a driving question that is student-centered and promotes sustained engagement toward the project development; (3) Engagement in the project is collaborative and demands authentic application (Blumenfeld et al., 1994); and, (4) The teacher scaffolds interactions between students and students’ cognitive processes and skills. Divergences in PBL across program approaches appear at the degree of integration of computer technology (Moursund, 1999), the degree of reliance on learning goals versus place-based authenticity (Blumenfeld et al., 1994; Lee and Miller, 2016), and level of individualization (Krajcik et al., 1998).

Much of the research on PBL focuses on the middle school and high school levels with a paucity of research in the elementary grades. The majority of the research that is available suggests a positive correlation between PBL and various student outcomes (Charara et al., 2021; Kokotsaki et al., 2016) and tangible benefits from PBL for student learning of content, motivation, or problem solving (Kokotsaki et al., 2016).

One study conducted by Harris and colleagues (2015), and using a randomized-controlled trial approach, indicates that sixth grade students who used materials designed with a PBL perspective outperformed students in the comparison curriculum on a robust outcome measure that was developed by the researchers aligned to ideas in the *Framework*. Specifically, the units aligned with several core science ideas and two science practices in the *Framework*—developing and using models and planning and conducting investigations. The results from the first-year implementation of a 2 year study demonstrated that project-based curriculum materials help students achieve specific inquiry learning outcomes when treatment teachers use materials and pacing guides with school support, including extensive professional development aligned to the materials.

Further results supporting PBL were reported in a high school math study by Boaler (2002), which did not employ random selection but matched schools for SES and race. It compared 300 matched students in one grade across three years in two schools, approximately half engaged either with the PBL approach while the other half had traditional instruction. The number of students in the PBL group who passed the national exam was three times that in the traditional group. The students in PBL reported less anxiety and exhibited increased facility with problem solving. Importantly, Boaler’s (2002) study suggests that PBL mitigates the effects of economic difference among students, whereas the traditional teaching approach exacerbated correlations between socioeconomic status and test results.

Historically, as a curriculum approach, PBL has had considerable purchase as being student-driven and individualized, with teachers enacting it in local contexts (Thomas, 2000). The localized design of PBL has promoted student choice and interest through deeply authentic questions, flexibility in process, and engagement in, and investment from, the local community (Larmer and Merendoller, 2010). In addition, individualized PBL experiences led to investment of teachers and students because of an opportunity to balance expertise on the part of the teacher while also drawing from students’ strengths, questions and skills. Studies also report that PBL enhanced teacher competency through soliciting and accessing available tools, community experts, and technology (Halvorsen et al., 2018; Miller and Berland, 2020).

Challenge presented by the Next Generation Science Standards

The Next Generation Science Standards (NGSS Lead States, 2013) are based on current research and their enactment is intended to scale throughout the United States. They describe sets of performance expectations for science understanding that students build

toward meetings by the end of each grade level or grade band across grades K-12. Based on the *Framework*, NGSS draw from research that very young students can engage in science practices, that practices are integral to developing understanding of science ideas, and that science proficiency builds over time with increasing sophistication (NRC, 1999, 2007, 2009, 2012). The *Framework* calls for learners to use disciplinary core ideas, scientific and engineering practices, and crosscutting concepts together to make sense of phenomena or solve problems—to engage in “three-dimensional science learning.” As such the *Framework* presents a vision that mirrors the conceptual work of science practitioners. The NGSS concretizes the *Framework*’s call for integrating different dimensions of science learning within three-dimensional performance expectations, envisioning what students should be able to do with knowledge and not just simply *know*. With developmental theory (NRC, 2007; Smith and Nadelson, 2017) and situated learning grounding its development (Bransford and Schwartz, 1999), the *Framework* and NGSS aim to address inequities in science learning experienced by underserved student populations in the US and provide more rigorous levels of teaching and learning of science understandings (Miller et al., 2015; Reigh and Miller, 2020).

Enactment of the three dimensions called for in the *Framework* and NGSS has presented a considerable challenge at the individual classroom level and at scale within districts. Proponents of the NGSS acknowledge a gap persists between what is promoted by the science research and policy community and what is taken up by teachers, school districts, and local communities (Duschl and Bybee, 2014). In practice, the implementation of the NGSS seems to require major investment in science teachers’ professional learning, an investment underfunded districts may not be able to make, or may be unable to prioritize.

Building on positions from *Taking Science to School* (NRC, 2007) and *How People Learn* (Bransford and Schwartz, 1999), the NGSS provide a tangible series of performance expectations that codify science learning as increasingly sophisticated demonstrations of science understanding where learners integrate disciplinary core ideas, seven conceptual understandings that cross disciplines, and eight science practices including argumentation, modeling, and mathematical and computational thinking. Unlike many state standards, which dictate the reproduction of isolated facts and ossified routines like the “scientific method,” the NGSS move away from discrete science ideas and memorization of steps to express what students should know and be able to do to meet performance expectations.

Merging the NGSS, with its focus on engaging all students across the US in three-dimensional science learning, with PBL and its focus on framing learning and engaging all learners with motivating and authentic contexts, we argue presents a powerful learning opportunity. Our goal is to productively merge these two approaches to support students in deeper learning, particularly at the elementary level where the need is great.

Research questions

In order to develop materials that use a PBL approach to engage students in three-dimensional learning, we ask.

1. How can we iteratively develop a comprehensive implementation system (i.e., teacher and student-facing materials, assessments, and professional learning) of PBL curriculum materials over three years of design cycles?
2. What would be a productive design approach to developing PBL materials that effectively address the NGSS?
3. What tangible characteristics would a final product feature that uses both PBL and the NGSS to promote student learning?

Methodology

We used design-based research (DBR; Barab and Squire, 2004) approach to iteratively design, develop, test and revise third and fourth grade PBL curriculum aligned with the vision of the *Framework* and the NGSS. Our Grade 3 work began with conducting exploratory teaching experiments with select teachers, moved to a pilot study of more refined materials with a larger group of teachers, a field study of polished materials with yet more teachers and multiple districts, and culminated in sharing the materials on an open-source online platform. The same development and revision cycle was used for the creation of fourth grade materials. In essence, we designed materials that could be implemented at a large scale to support all students in meeting the NGSS across entire districts, and we propose a more effective pedagogical model that combines the full scope of the NGSS with PBL. Using DBR, we have engaged in tweaking on-the-ground elements of this marriage to best address challenges with implementation creatively with iterative design.

Context

Multiple Literacies in Project-based Learning

The Multiple Literacies in Project-Based Learning (ML-PBL; Krajcik et al., 2015) project is a collaboration between Michigan State University and the University of Michigan. The intent of the project is to develop, test and iteratively refine Grades 3–5 project-based science curriculum materials aligned with the NGSS and PBL. To support students’ learning of science as envisioned in the NGSS, the design of PBL materials also focused on integrating additional literacies, primarily English language arts and math, and that also focused on cultural relevance, Social and Emotional Learning (SEL), and classroom discourse. Importantly, our approach goes beyond simply providing curriculum materials, and seeks instead to deeply attend to issues of implementation (Fishman et al., 2013; Penuel and Fishman, 2012). We seek to create a comprehensive implementation system where teachers are supported by

Table 1 Development and enactment cycle for third grade.

<i>Cycle of development</i>	<i>Types of enactment</i>	<i>Numbers of teachers^a</i>	<i>Levels of support</i>	<i>Project data sources^b</i>
First cycle	Teaching experiments (Gd 3 15–16)	6	Researcher(s) working with teacher Instructional support	Field notes/Classroom observations (~daily) Interviews (T and S) S work artifacts
Second cycle	Pilot studies (Gd 3 16–17)	12	Professional learning F2F PL - 6 days CheckIn 2–3/unit PL surveys and discussion	T surveys - PL discussions/feedback - Lesson enactment Unit assessments Classroom observations (2–4/unit; observation tool) SEL survey Interviews (T) Classroom observations (~daily) Interviews (T and S) S work artifacts
	Teachers from first year enactment.	4	Researcher(s) working with teacher	
Third cycle	Field studies (Gd 3 17–18)	20 (21+ classes) + Comparison	Professional learning F2F PL - 6 days CheckIn 2–3/unit PL surveys and discussion	T surveys - PL discussions/feedback - Lesson enactment Unit assessments Classroom observations (2–4/unit; observation tool) SEL survey (Tx and Co) Interviews (T) (case studies and PD follow-up) Summative assessment (Tx and Com)
	Teaching experiments	4	Researcher(s) working with teacher	Classroom observations (~daily) Interviews (T and S) S work artifacts

Abbreviations: T, Teacher; S, Student; PL, Professional Learning; F2F, Face-to-Face PL; Tx, Treatment; Com, Comparison; Co - Control.

^aNumbers of T and Classes taught varies by Grade level study; entries represent grade 3 across development cycles.

^bAdditional data not generated within the project is also collected (district reading/math benchmark assessments and demographic data; state end-of-year assessment data).

ongoing, sustainable forms of professional learning and have access to tailored three-dimensional assessments that can support enactment and instruction of ambitious PBL curriculum materials that support enactment of the NGSS.

Context of implementation

This ongoing study has taken place across multiple states in the US, involving more than 100 elementary teachers from 12 districts. (Insert information on teachers, students, districts in study).

Approach to design

The project has involved multiple cycles of curriculum development and revision. The study began with Grade 3 development in year one and began a new cycle of development for Grade 4 in years two and three. With each year, more teachers were added to the design project (See **Table 1**). Our first cycle of design and implementation culminated in what we refer to as *teaching experiments*. In this phase, we seek to observe the initial challenges the curriculum materials, professional development, and assessments present to a small cadre of teachers working closely with researchers, usually within one of two classrooms that cross districts. Data from the first phase is then used to inform the re-development of materials for a larger *pilot* phase of implementation in the second phase, involving significantly more teachers—in addition to working closely with a small cadre through teaching experiments—and where we begin to define larger problems of scaling across entire education systems. In an iterative fashion, data from the second phase is then fed into further refinements for use in *field experiments* in the third phase. Final refinement to our curricular materials, professional development and assessments are then made, and the materials are placed on Sprocket, an open-source, on-line medium, accessible across the globe.

Table 1 presents a summary of the cycles of development of the third-grade materials.

Employing design-based research methods

This project uses a framework that rests on the assertion that tenets of PBL and the NGSS can be productively woven together in materials to support enactment at large scales that shows integrity to PBL and motivates students to develop three-dimensional science understanding. We employed design-based research (DBR) methods to undergo iterative cycles of development to build

on theory of how learning occurs through the use of pedagogical and curriculum materials, professional learning, and assessment. This project reflects key characteristics of DBR: (1) a dual goal of developing pragmatic interventions for real-world use and the development of theory on how learning occurs through the use of the interventions, (2) iterative cycles in the research process, and (3) an appropriately robust approach to data collection on student learning (Wang and Hannafin, 2005). We created and revised materials in iterative cycles of design, implementation, as well as synergistically reviewed and examined opportunities for teacher professional learning. Importantly, we developed a set of design principles to support the implementation of these practices in the classroom (Miller et al., 2021b).

Barab and Squire (2004) suggest the use of DBR to examine the effects and promise of a learning design by applying instructional practices to real teaching contexts (Barab and Squire, 2004). Through the implementation of instructional practices, we seek to understand both students' learning processes and the impact of the designed intervention on student learning. DBR takes into account the underlying theory of the interventions, the features of the instructional activities or materials, and the contextual factors (Brown, 1992). Our theory within the ML-PBL project is that through thoughtful refinement of our PBL approach to incorporate the NGSS for learning and a coherent lesson-level design, we could develop a curriculum that has broad scalability as well as local potential to motivate and sustain learning for diverse contexts. In other words, we explored how to productively merge PBL features with principles of the NGSS. Teachers' understanding of the curriculum design approach used in ML-PBL and their capacity to implement these materials with integrity to designers' intentions was supported with recurring professional learning.

Logic model

Our overarching research questions explore how to support all students—particularly students from diverse backgrounds, including English language learners and students at-risk for underachieving or disengaging from science and school—in learning science through the design, development, and testing of PBL materials that integrate tenets of the NGSS.

The logic model (Fig. 1) shows how the features of PBL build from underlying theoretical ideas. To meet the challenges of diversity, universal design for learning is critical to provide differential scaffolds. We can accomplish this by providing various types of scaffolds and using the principles of universal design for learning—multiple representations, multiple means of expression and multiple means of engagement (Jimenez et al., 2007). Our PBL interventions have three major, complementary components:

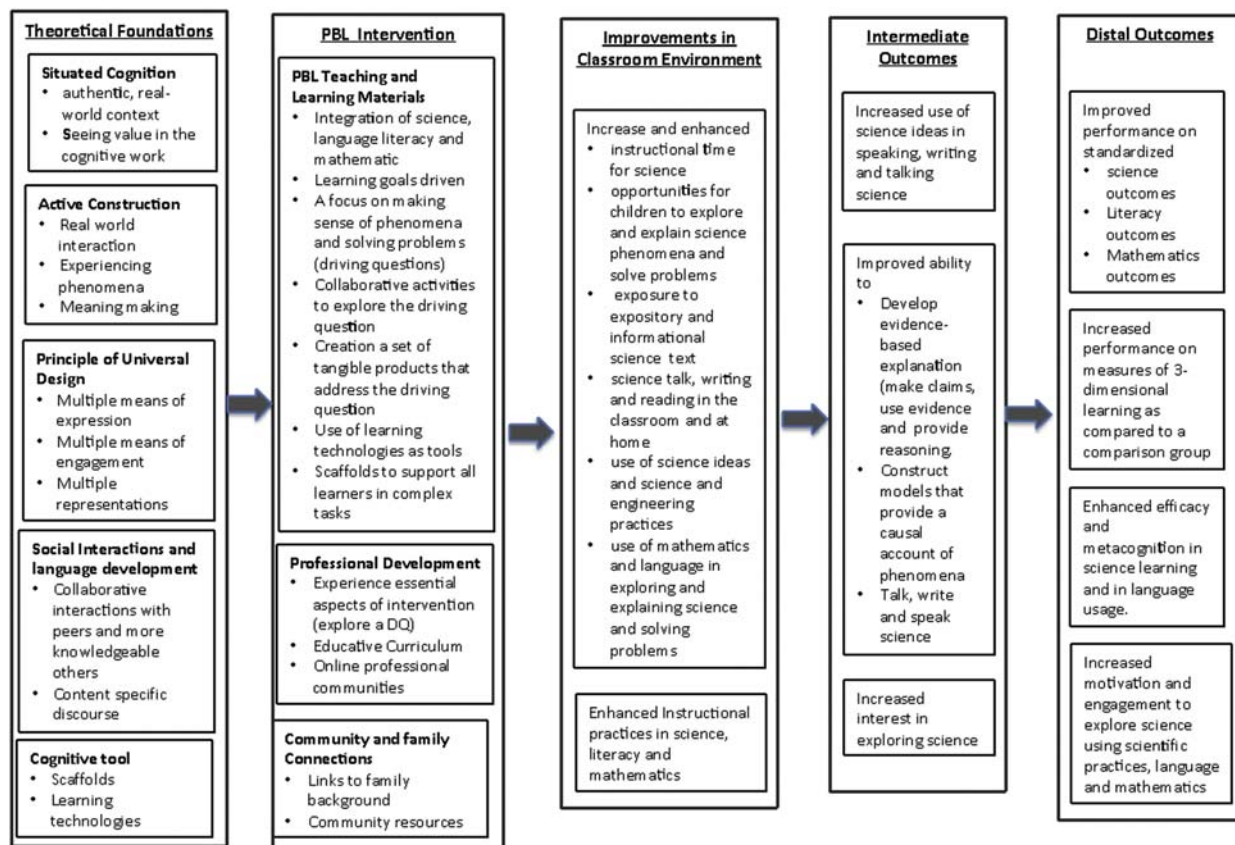


Fig. 1 Logic model for ML-PBL design.

(1) PBL teaching and learning materials, (2) professional development, and (3) connections to community and family resources. These components, working together, should result in both classroom transformation and intermediate and distal outcomes. **Fig. 1** presents a visual of the logic model that explains how elementary students' develop increased engagement and performances in science.

Data collection

According to [Cobb et al. \(2003\)](#), "*design experiments create the conditions for developing theories yet must place these theories in harm's way*" (p.10). We expected to observe the need to readjust the union of PBL and NGSS as our project unfolded through the collection of (1) field notes of ethnographic close classroom observations, (2) interviews of students and teachers, and (3). Collection of artifacts and assessments from the classrooms.

Studying implementation of materials

In the first 3 years of the research study, we collected careful and detailed field notes from multiple school sites within districts in the Midwest US. For the research reported in this manuscript, data was collected in schools where the team developed field notes and collected video data one to three times a week (~40 min) across classrooms. Our protocol for studying classroom enactment includes an in-person observation by an observer who is familiar with the intent and flow of the unit, and has learned to look for the features of PBL that frame the lessons using an Observation Tool designed specifically for the ML-PBL project. Using the tool, we scored teacher enactment of specified components of the lesson, as well as various features that are consistent with PBL, such as use of the driving question, engaging students in figuring out, and sense making, discourse, and development of artifacts.

The focus of the collection used a wide lens to capture whole class implementation of the teacher and whole group discussion in order to broadly ascertain fidelity to the curriculum as well as a more narrow lens employing interviews and following specific students over time. During unit implementation in the second-year pilot and third year field tests, we conducted classroom observations of each teacher at least once. We averaged three to five observations per teacher per unit.

The research team met regularly to compare field notes, and observations of lesson enactment, and professional learning enactment. Questions the team asked during this phase included, "Does the design of PBL and NGSS work to promote engagement and learning?" "What are the processes for student learning, including socio-cognitive aspects?" "How can we improve opportunities for students to engage in the NGSS practices and use their science learning in projects?" "What challenges do teachers experience in enacting the curriculum?"

We found that the most profound outcomes were formulated after the first and second years of the enactments. As we concluded the first round of completion of the units, our work revolved around the design work of ML-PBL units: (1) organizing learning to build coherently over time, (2). Organize learning to culminate in artifacts, and (3). Support shifts in teacher practices through use of assessment, supportive features, and professional learning. During each cycle, we integrated on development of units and continued to inform a final system of assessment. In the final cycle, we focused most intently on professional learning that emphasizes integrative design. Together curriculum, assessment and professional learning worked as a system that would be tested and redesigned to derive increasingly scalable and sustainable materials.

Data analysis

To analyze field notes and observations, we followed a systematic data analysis process modified from [Strauss et al. \(1996\)](#).

Step 1: Researchers read through fieldnotes and observation data several times to create tentative proposed themes that were emphasized or repeated by a teacher. We developed themes related to the challenges and opportunities related to features of PBL and the three dimensions of the NGSS, adding more themes in subsequent reads (i.e., student voice, adaptation of feature, community involvement). We used the following questions to search for tentative themes of challenges and opportunities for design and revision:

1. How do features of PBL align with meaningful learning of the NGSS?
2. How can we describe challenges that the features of PBL present for teacher enactment?
3. How does professional learning support teachers in enacting the materials to use PBL for meaningful three-dimensional learning?
4. How are scaffolds in the materials taken up to support PBL for deep learning of the NGSS?

Across teachers and sites, we sought patterns indexing themes related to these questions. We discussed, took extensive meeting notes, and wrote summaries of these themes across teachers, highlighting specific responses to our four questions. The proposed themes were tentative themes that emerged from analyzing the data sources. After this initial analysis, we would address an opportunity or challenge identified in the classrooms by adjusting the curricular intervention (e.g., teachers adding to the lesson, adapting for students with IEPs, skipping lessons, or parts of lessons), gather more data, and follow the same processes looking for new proposed themes or evidence to support a previously proposed theme. In each step, we returned to our theory of PBL (i.e., meaningful and engaging question contexts, and useable supportive design for teachers) and theory underlying the NGSS (i.e., coherent, rigorous, and evidence-based learning progressions) and its enactment.

- Step 2: Across units, we continued to analyze additional data as we collected it, and as we introduced new designs. Seeking to further refine themes, we looked for evidence to support themes informing iterations that address challenges and opportunities.
- Step 3: In Step 3, we focused on responses to iterations. We sought additional support for the themes of challenges and opportunities of the NGSS and PBL by reviewing recent data to find possible evidence of responses to iterations.
- Step 4: In Step 4, we searched for patterns and shifts in the themes to create an overall summary of the themes that occurred across the year. In this stage we summarized solutions to challenges, and critically analyzed the four questions to inform further revision of the design. We again kept focal theories of PBL and those informing the NGSS (i.e., three-dimensional learning) and examined theory for scaling. This step completed a cycle.

Design cycles

For development, we engaged in multiple design cycles (Brown, 1992; Collins et al., 2004; Sandoval, 2014). Here we report on three of our yearly design cycles (See Table 1), reporting on the design cycles that occurred from school years spanning 2015 through January 2018. Each cycle included design, development, enactment, critical review and analysis of collected data, and revision of the curricular tools (Barab et al., 2007), implementation structures (Penuel et al., 2011), and professional learning environment design (Rosebery et al., 2016). Cycles were ongoing from September to August of each school year and culminated with new broader and revised theoretical principles as well as refined and further developed curriculum materials, professional learning strategies and assessments in the late summer of each school year (Fig. 2).

Iterative cycles of design, testing, and revising

During the first design cycle, we collected and analyzed field notes and observation notes. In the second cycle, we created assessments, or performance tasks and revised the curricular format based on the review of the data, and again conducted ethnographic and interview data collection. In the third cycle, more revisions were made leading to the development of the final curriculum and the design of professional learning supports based on current theoretical principles. We will reflect on the need for more changes as the current design cycle comes to a close.

First cycle teaching experiments

Design

Extensive planning and preparation in the first cycle provided overarching plans for unit development, including anchoring phenomena of units, assignment of standards to be addressed, development of lesson-level learning goals or Learning Performance statements (LPs), as well as unit and lesson templates. In this cycle of design, we solidified and instantiated principles for incorporating English language arts, mathematics, cultural relevance, SEL, and classroom discourse.

The process of resource design and development reinforced design team members' understanding that the development of rich PBL resources to support 3-D learning and meet the NGSS and other interrelated goals (e.g., SEL) required a purposeful collaborative and iterative process. Design team members collaborated on developing and revising evidence statements, SEL learning objectives, and cultural relevance and discourse objectives, making these explicit and a central part of the lesson design and unit progressions. Identifying stages and responsibilities within the iterative process, and continually revising plans based on desired versus actual experiences and outcomes and balancing the need for rigorous learning goals with providing adequate supports to help students meet them, provided extensive opportunities for rich inter-team discussion and for thinking of ways to identify, organize, prioritize, and eventually streamline steps in the process.

The team worked to maintain focus on the goals of PBL and 3-dimensional learning from the NGSS. Weaving together PBL and 3-dimensional learning offered a new vision for classrooms in which all children, regardless of their backgrounds, experience rigorous learning that pushes their thinking forward, so they have the intellectual resources to develop agency in science and in making sense of the world in which they live. At times, because of time and capacity pressures and differences in our experiences within the educational system we regressed into traditional approaches and lenses in the curriculum and theoretical framing. The

Cycle 1 Focus on initial Development SEL & Equity	→	Cycle 2 Revise material from Cycle 1 Focus on Assessment	→	Cycle 3 Revise material from Cycles 1 & 2 Focus on PL	<i>Place units on Sprocket (on- line Portal).</i> <i>Conduct Efficacy Study</i>
---	---	---	---	--	--

Fig. 2 Foci of design cycles.

need to continually recommit to an innovative perspective—and break away from traditional approaches—is a tension that all designers face. To push the field forward and learn what is possible, we needed to stay true to the premises of PBL and NGSS, keep our learning goals rigorous, and believe in what is possible, not what is expected.

Testing and analysis

In the first cycle, we developed and implemented the first curriculum design and associated professional learning opportunities. Teaching experiments using unit resources occurred with a small cohort of teachers in multiple schools. As teachers received professional development on using the draft resources in teaching experiments, researchers engaged in the iterative process inherent in the development of PBL curriculum resources. Purposeful iterative planning, and retrospective analysis of observations and field notes identified needs that served as foci within the development process.

In the testing and analysis phase, we were cognizant of the importance of interdisciplinary teams: The various team members—researchers and practitioners with expertise spanning science education, math and language literacies, and SEL among other areas—brought different strengths and perspectives to the project work. This assemblage of expertise enhanced the materials and what could be accomplished.

The testing, analysis, and revision work proved extremely iterative in nature. Throughout the process, the learning goals for lessons (i.e., 3-dimensional learning in science, social emotional, and discourse aims) remained at the forefront. Unit-level learning goals (expressed as performance-based evidence statements) were expected to frame the vision of each unit. Our work made use of data from the small cohort of teachers enacting the lessons to drive revisions of lessons and the revisions of progressions and scaffolds. As the team collected a first round of rich data from the first implementation, we needed to elaborate a set of collaborative practices around the revision process, expressed as a set of shared design principles. We designed and redesigned prior and subsequent units using these design principles, leading to student materials used in the second cycle.

Outcomes and reflections from the first design cycle

Major lessons learned with respect to the design principles from the first cycle included the following:

- (1) **Units needed to better attend to issues of coherence:** The simultaneous development and redesign of units, and rapid development of additional units challenged the team to create tight connections between and across units with respect to promoting coherence, where learning becomes progressively more sophisticated within the year and across years (Fortus and Krajcik, 2012). Specifically, the team experienced challenges in developing lesson-level learning objectives (i.e., learning performances) that increased in sophistication over time and pushed students to develop a deeper understanding of disciplinary core ideas while increasing their engagement with practices, and furthering their capacity for productive discourse and opportunities for SEL. The group confronted this tension, moving forward, to be resolved in subsequent cycles.
- (2) **Current level of scaffolding proved insufficient:** This item is related to staying focused on PBL and 3-dimensional learning. The group needed to design materials to support students to take part in classroom tasks that require planning, synthesizing, sensemaking and using ideas to understand their world. All students can take part in these synthesizing and integrating tasks. Our challenge was to develop and include the structures and scaffolds to support rich academic practices in which all learners can productively and collaboratively engage. Finding the right level of scaffolding that supported learners in meeting the multifaceted aims of lessons (i.e., 3-dimensional learning, SEL, discourse) without overcomplicating lessons and diminishing their usability required attention in the revision process.
- (3) **Designing supportive features for teachers:** Providing the right level of support for teachers was a challenge for design. One of our goals was to provide highly specified development materials for students and teachers. However, too much support (as text in the lessons), created situations where teachers did not read the materials. Conversely, there was a danger of providing too little support. The team returned to this tension repeatedly. We tried providing links to supports from a lean document that provided a more succinct overview of lesson structure. In addition, taking an asset-based approach to teacher support, which aligns to evidence that teachers will develop skills in a just in time-based fashion (Miller and Shelton, 2021), meant taking into account the reciprocal nature of practices, goals and supports and context (including community), and how they work together to enhance teacher practice and reflection on practice.

Second cycle pilot studies

In the second cycle, the ML-PBL Team focused on completing teaching experiments with Grade 3 units, reviewing and revising Grade 3 units, and designing and developing initial drafts of Grade 4 unit resources. Year 2 early teaching experiments with Grade 4, and pilots of Grade 3 took place in seven schools. The team continued to design to address issues of coherence identified in the previous cycle, as well as teacher and student support for teaching and learning. This cycle was laser-focused on supporting teachers by providing a variety of assessment supports. Teachers could address how they were supporting student learning and reflect on their own teaching practices through multiple informal, formative, and summative assessments.

Design

This cycle saw enhancements in our design-based research process for unit development, particularly as it relates to assessment and evidence of 3-dimensional learning. Although the process was iterative, the design process began by developing learning goals—specifically, lesson-level 3-dimensional Learning Performances—aligned to PEs. These learning goals were enhanced and revised throughout unit development and subsequent iterating in response to ongoing data collection.

Testing and analysis

The development team made several adjustments and additions to the testing and analysis process. Development and revision of design for materials resulted from feedback from teachers during professional learning sessions and from talking with teachers during the enactments, review of videos taken during the first-year enactments, field notes taken during the enactments, and review and feedback regarding the units given by project team members. We were interested in how teachers were using the formative assessments for understanding goals of the lessons and planned to revise assessments based on feedback in the third cycle. We found the feedback provided by the teachers extremely valuable in revising the units and lessons and reexamination of documents.

Outcomes and reflections from the second design cycle

Unit Development revisions. An important modification made to the units in this cycle was using a theme that cuts through each of the units. For example, because energy and energy transfer is an important disciplinary core idea and crosscutting concept, we decided that each of the fourth grade units would build and reinforce this important idea. We also more systematically focused our efforts on making the units more interdisciplinary. This can be seen, for instance, in the first fourth Grade unit, *What causes the land in our community and world to be shaped in different ways?* In this unit, we focus on students using and revising their developing understanding of energy and energy transfer to figure out how water and wind change the earth. Another revision was to make explicit how we engage children in the engineering design process and in various design elements throughout all of the units. We focused on integrating engineering, so children have opportunities to engage in design processes. Our revision process also involved focusing more explicitly on discourse, equity, and SEL. Our goal was to include discourse prompts and “look fors” (i.e., tangibles that teachers can look for to see if students are engaging in the lesson as intended) in every lesson and either an SEL or equity lesson in each learning set.

Assessment Development. During the second design cycle, team members collaborated on refining design principles as well as designing and revising components of an assessment system, including: (a) formative and summative assessments based on learning performance statements, (b) discourse objectives, (c) formative assessment, “look fors,” and (d) evidence statements, all integrated in lesson design and unit progressions. We also developed a plan for an end-of-year series of assessment tasks to be used during field testing in Year 3.

The ML-PBL project aimed to provide meaningful and actionable information about the progress and capacities of individual students and groups of students toward meeting learning goals, as well as students’ ability to demonstrate and apply understandings in new contexts. Assessments informed teachers’ instructional practices and interactions with students and for describing a bigger portrait of student learning. As designers, we found the assessments valuable in providing information with respect to the design of units.

To accomplish these aims, the assessment strategy included specification of learning performances (LPs; the lesson-level 3-dimensional learning goals), as well as end-of-unit learning goals (which students’ progress toward served as evidence to make “claims” of students’ enduring understanding). This variety of performance learning goals necessitated multiple types of performance-based, student-generated artifacts. Adherence to the project-based context helped to ensure coherence among artifacts. Throughout each unit, learning goals and corresponding artifacts build on each other and offer opportunities for students to reflect on and revise science understandings and practices.

Formative assessment opportunities within each unit continuously gather evidence about student learning. Assessments included a variety of strategies—including teacher observation of student interactions and work (“look fors”), planned-for interactions, and curriculum-embedded assessments—all designed to support teachers in identifying each student’s place along a progression of learning, provide feedback, guide plans for adjusting instruction and providing scaffolds, and for developing students’ self-regulation strategies.

Units also include end-of-unit culminating artifacts that provide tangible means for students to demonstrate their ability to connect learning experiences in response to the unit driving question. A performance-based assessment at the end of each unit provides an opportunity for students to extend and apply what they learned in a related, new context, as evidence of enduring understanding.

Along with culminating artifacts, to help triangulate whether students have developed the enduring understanding goals of the unit, our assessment system includes Unit-Specific Pre- and Post- Assessments of Enduring Understanding in which students utilize disciplinary core ideas, science and engineering practices, and crosscutting concepts while interpreting text and images and determining evidence to support two claims (i.e., generalizations abstracted from the enduring understanding).

The ML-PBL team also developed a series of summative assessment performance tasks (3–4 tasks) to assess student learning. These 3-dimensional tasks are administered at the end of the year, assess the scope of grade-level PEs, employ contexts other than those used in ML-PBL units, and provide a means of comparing the performance of students in ML-PBL classrooms with those from non-ML-PBL classrooms.

Third cycle field experiments

The third cycle continued to provide rich data to support the redesign of the units and the assessments. During this cycle, we focused on revision, including the testing and analysis of Professional Learning materials. We applied an iterative process informed by design-based research to develop and redesign professional learning experiences, which proved key to supporting enactment with higher fidelity to our design principles. We gathered and analyzed data from enactment surveys and observations, whole group reflection during professional learning sessions, and teacher interviews to modify and improve professional learning for teachers.

Design

With each new cycle we enhanced our support for teacher sensemaking about PBL practices. During this cycle we incorporated small-group and whole-group discussion prompts that engaged the teachers in engaging in argumentation about teaching and learning within the units. We developed a Professional Learning Framework that facilitated sensemaking about lesson design, and strategically involved our returning teachers.

The professional learning framework focused on the following: (1) Targeting experiences of one (or more) key lesson(s) in each learning set that engage teachers in sensemaking about how to support students in the “figuring out” processes that occur throughout the unit to move toward the final artifact and building proficiency in meeting bundles of PEs; (2) Soliciting teacher-driven discussions on the affordances of the features of ML-PBL lesson design; (3) Fostering conversation among returning and new ML-PBL teachers about enacting PBL; (4) Highlighting the coherence across learning performances; and (5) Using video of classroom discussion, artifacts, questions, and examples of student work to support teachers in critically reflecting on the ways that PBL supports equity, engagement, and community building.

Further design for professional learning focused on coherence across experiences. Each professional learning opportunity built on earlier opportunities, deepening the teacher-facing professional learning DQ (How can we support students in sensemaking in PBL?), artifacts, and engagement with the lesson features. Although the Professional Learning structural framework remained the same, the Professional Learning outcomes were designed to build one upon the other. We began the face-to-face Professional Learning by recording some of the key take-aways of the last unit, including use of crosscutting concepts, science and engineering practices, and disciplinary core ideas. Then we asked and recorded what the take-away would look like if students had made greater gains.

Each video conference meeting focused the teachers to collaboratively reflect on the feature(s) of the student-facing materials with respect to the lesson-level 3-dimensional learning performances. They discussed the alignment of the feature with the learning performance, and the evidence statement. They were tasked to work together to describe how the lesson feature provides direction for the informal and formal assessment. In the face-to-face Professional Learning experiences, we repeatedly returned to the lesson structures. We analyzed how the discourse supports were used by the presenter with teachers, how the learning performance was met in the lesson, or how the evidence statement is instantiated in the artifact. We returned to the unit Driving Question and three-dimensional learning performances of student-facing materials to support teachers in explaining the intentionality of each learning set, how learning sets build on one another, and how coherence is met.

Testing and analysis

In this cycle, we solicited teacher-driven discussions about the affordances of the features of the current iteration of the ML-PBL lesson design. One aim for the professional learning was to provide support for how teachers can use and interpret the lesson structure. Based on teacher feedback, our lesson structure was initially viewed as “lengthy” and difficult to navigate. However, we saw consistent evidence (as confirmed through classroom observations) that the lesson format promotes 3-dimensional teaching in PBL when teachers engaged with provided supports: teachers enacted lessons efficiently with positive results.

During interviews, teachers described that they appreciated the new professional learning design. Multiple teachers have expressed that they now understand the whole picture of each unit and know how the learning progresses, indicating a deeper grasp of the coherence intended within and across units. They found the professional learning “interesting” and “purposeful” and explained (during interviews) the “purpose of each learning set and how 3D learning develops throughout the unit.”

Analysis of the revised design of the video conferences and face-to-face professional learning experiences by the ML-PBL team drove sensemaking about the relationship between lesson structure and features and enhanced three-dimensional instruction. The third cycle of design focused on increasing awareness of the need for changes in practice on the part of the teacher. The video conference discussions from professional learning sessions were analyzed, providing more strategic guidance on what proved productive for teachers’ learning. For example, when the PL facilitator encouraged examination of classroom practice for meeting the 3-D learning performance and/or identified how that practice hindered or supported student learning, agency, and SEL, this proved a high-leverage move within the professional learning series.

Our iterative design of the professional learning opportunities focused on how teachers’ learning progressed, and highlighted changes in students’ practice as evidence of learning and growth in PBL and the NGSS. For example:

- After completing the first third grade unit, teachers stated that their students had started to listen to their peers’ ideas in small group discussions. They agreed that the students would need to work next on building students’ capacity to listen and productively use peers’ ideas.

- After the second third grade unit, the teachers agreed that their students had created models featuring cause and effect. They thought that the next step would be to model interactive relationships where one component is changing. (This is a goal for practice in the third unit, Birds.)
- Using this design change, some teachers reported that they saw themselves as agents in sensemaking about Teacher Practice Knowledge. This change indicates the professional learning promoted (1) enhanced reflective practice on the part of the teaching community, and (2) more thoughtful engagement with the educative features of units.

Notably, rubrics developed for key formative assessment opportunities supported teachers in providing students with timely feedback to guide their next steps in learning. Assessment data from this cycle will be analyzed to provide evidence of resource effectiveness and future design directions.

Outcomes and reflections toward fourth cycle

The new professional learning design was intended to emphasize teacher sensemaking about the progression of students' learning as well as the purpose of the DQ and the final artifact in the units. The professional learning highlighted the features of PBL by having the format teachers' experience employ key features of PBL, including a teaching-focused Driving Question (i.e., How can we support students in sensemaking in PBL?) and a culminating artifact. The final artifact for professional learning was to develop a model of the progression of students' 3-dimensional learning. Coherence, being a key feature of PBL and the NGSS, also featured prominently in professional learning. We see our role as facilitators to support teachers in employing PBL practices effectively through intentional reflection of evidence of what is happening in their classroom (i.e., engagement, informal assessments, and "look fors") and continued engagement with how lesson-level learning performances coherently build into unit-level performances (enduring understandings and unit-based and summative assessments).

Lesson-level features (i.e., learning performances, figuring out statements, "look fors", evidence statements) guided teachers in 3-dimensional instruction where students actually use disciplinary core ideas, science and engineering practices, and crosscutting concepts to make sense of the world. Lesson-level features supported teachers in conducting formative assessments throughout each unit. For most teachers, 3-dimensional instruction represents a major change in practice. With its focus on teachers understanding three-dimensional learning performances, our professional learning design in tandem with associated curricular materials provided robust support for helping teachers move toward 3-dimensional instruction. Evidence we have collected to date indicates that teachers use lesson-level features with increasing sophistication over time.

From the outset of the ML-PBL Project, our team has taken steps to more clearly articulate the vision of the *Framework* and the NGSS—engaging students in using multiple science and engineering practices, disciplinary core ideas, and crosscutting concepts as tools to make sense of complex phenomena. We have designed units of instruction that provide a coherent progression of rich opportunities for students to develop sophistication in each of the three dimensions of science proficiency—opportunities for filling a toolbox with practices, ideas, and conceptual tools that can be used to explain phenomena and solve problems relevant to learners.

In terms of assessment, our continued goal is to support teachers in supporting students to access this toolbox of ideas, ways of knowing, and negotiate, reorder, or otherwise revise these resources for immediate use in an assessment environment. We strive to design assessment tasks that will allow students to consider how to use multiple scientific practices, and various elements of disciplinary core ideas and crosscutting concepts to explain phenomena or solve problems presented. We believe that assessing in this manner will promote a deeper understanding on the part of students and teachers of what it means to "do science" and provide opportunities for students to develop persistence and employ creativity in science.

Discussion

This multi-cycle design-based research project offers a promising example of how 3-dimensional project-based learning units, when tightly coordinated with assessments and professional learning in a comprehensive implementation system, can provide rich learning opportunities that productively blend elements of PBL with those of the *Framework* and NGSS, namely 3-dimensional learning. We began with an initial understanding of PBL that consolidated and expanded the features of PBL in the literature, based primarily on PBL science research at the middle school (Blumenfeld et al., 1991). We showed that PBL environments, when thoughtfully integrating elements of the NGSS, can support instruction that is well-scoped and coherent, and implementable at large scales. This work also points to how the three-dimensional learning of the NGSS provides a vehicle for not only important correlated goals (e.g., SEL, discourse), but can support the joining together of learning as a participatory performance in a community with using science to explain phenomena and solve problems. Tangibly, we extended and modified the 6 features of PBL to include the NGSS (in italics) below.

1. They start with a *driving question about a phenomenon to be explained or an engineering problem* to be solved. The NGSS adds parameters to the scope and nature of the Driving Question in PBL, specifically addressing the nature of the question at hand, and infusing the disciplinary context.
2. They focus on *3-dimensional learning goals* that align with key NGSS science standards and frame demonstration of student mastery on assessments.

3. They require *participation in NGSS scientific practices as described by the Framework and NGSS*—processes of problem solving that are central to expert performance in the discipline *and the integration of literacy as tools to support students in sensemaking*.
4. As students explore the driving question, they engage in *collaborative sensemaking activities* to build shared knowledge—the solutions to the driving question.
5. While engaged in the *practices of science*, students are *scaffolded with tools* that help them participate in activities considered beyond their ability.
6. Over time, students iteratively *create a set of tangible products* that address the driving question using *science and/or engineering with increasing sophistication*. These are shared artifacts with local impact, publicly accessible external representations of the class's learning.

The NGSS specify the learning goals as three-dimensional and as collaborative participatory practices. In this way, the NGSS condense features 2, 3, and 4. Students must show mastery through engaging in the practices, which cannot be developed in non-collaborative contexts. This condensing also offers new conundrums about assessment, which we also address in the project.

Driving Questions. This work highlights how driving questions can serve as an integration point for PBL and the NGSS. We found **driving questions** needed to be meaningful and focus on a phenomenon or problem at the unit level. For example, “How can I help the birds in my community, so they grow up and thrive?” These driving questions were meant to motivate the students, be answerable with different levels of sophistication over five weeks and leverage a bundle of NGSS Performance Expectations. This driving question encompassed a phenomenon that required understanding of life cycles, genetics, behavioral characteristics, and climate and weather. The questions highlighted the need to use crosscutting concepts—patterns, scale, stability and change, and cause and effect, and use all of the practices, including developing models, constructing explanations and designing solutions. For example, in the bird unit, the students need to figure out where their bird might be in its life cycle and how many were sighted locally. They used this information to consider what sort of food, and what quantity, the bird might need.

3-Dimensional Performance-based learning goals and Collaborative Practices. We use **3-dimensional learning performances (LPs)** for each learning set and lesson in the learning set, which act as learning goals at a smaller grain size using the same 3-dimensional structure of performance expectations. A learning set is a coherent set of lessons that build to answer a sub-driving question necessary to answer the driving question of the unit. For each unit, we have roughly 25 learning performances. Since the learning sets were related, the LPs often changed as the idea deepened or the practices changed. The LPs in each learning set, a coherent set of lessons that build to answer a sub-question, focused on addressing different aspects of the same element of the disciplinary core idea. We provided an organization showing how each lesson-level learning performance aligned with an evidence statement that clearly explained what data the teacher would observe as evidence that the three dimensions of the learning performance were met by students. In a unit, the learning performances encapsulated all of the dimensions in the PE bundle, usually about 4–5 cross-disciplinary PEs. To complete the project, the LPs in one unit called for all of the NGSS science and engineering practices and several crosscutting concepts. All of the project work was collaborative for the students and required the students to use sensemaking to exchange and creatively build on one another's ideas. In our units, the practices were incorporated into the learning performances, and they were enacted by the entire class at the same time, with direction and support from the teacher.

Artifacts. This work showcased how our team conceptualized the culminating artifacts in PBL in relation to the NGSS. We began our work with a theoretical design for artifacts that reflected the coherent building of knowledge over time intended in the units, directly related to answering the driving questions, and aligned with focal science and engineering practices. Notably, artifacts were also meant to remove barriers between the school and the community with a presentation of new understanding with a tangible artifact.

Discourse Moves. This work highlighted the central role of discourse moves to integrate PBL and NGSS. Developed by the Wisconsin Center for Educational Research (WIDA), our discourse moves were purposefully integrated into units and their use was made explicit to teachers in professional learning and in the lesson plans developed. The discourse moves respond to the call in the English Language Proficiency Development (ELPD) Framework for supporting the Language of doing science (van Lier, 2006; MacDonald et al., 2014). Our approach shows the tight integration of elements with one another, including discourse with SEL practices. The Discourse moves support all of the purposes for engaging in a practice through discourse.

For example, after a student expresses an idea while engaging in modeling, the move, “help clarify an idea” means that the teacher would ask questions seeking to make the language more specific and clarify the idea; and so both linguistic and cognitive demand would be supported—and the purposes of expressing the idea are accomplished.

Equity Goals and SEL Goals. We also showed how our approach productively addresses SEL as an essential correlated goal that supports the integration of ideas from PBL and NGSS. Our first design included Equity and SEL goals. In this iteration, we developed these goals where they made sense in terms of the narrative of the unit (i.e., we did not create SEL lessons for the sake of “covering” SEL). For example, the SEL goal is the following: “Identity development during challenges: We can learn to be comfortable with uncertainty and take on the challenge of collaborating with others and new ways of engaging with content.” In our first design, this goal was written as a corroborating learning goal that manifested as the student participated in the activity.

Through the curriculum, we promoted the view that SEL builds from opportunities within PBL, but at the same time, that it is an essential component for success in PBL. The co-development of SEL and 3-dimensional learning was intended to form a self-sustaining system that fosters an environment of strategic reasoning, self-monitoring and reflection, risk taking, collaborative skills, and problem solving. In other words, we hold that the development of SEL and 3-dimensional learning must be reciprocally supported to achieve success in project-based learning.

Assessment Design and Approach. Lastly, we specified an assessment framework that articulates how learning performances (lesson-level learning goals) can serve as a practical means to describe how students are making coherent (i.e., building over time) progress toward the end-of-unit or end-of-year NGSS performance expectations. In addition, we showed how we can support teachers' use of learning performances through tangible supports like evidence statements that describe what students must say/doing to demonstrate proficiency.

Conclusion

The *Framework for K-12 Science Education* and the subsequent NGSS present a new vision for science teaching learning. The *Framework's* vision is for all learners throughout the United States to develop the intellectual tools, creativity and personal disposition so that they can make sense of phenomena and solve problems that will make a difference in their lives or the lives of members of their community. But for the vision of the *Framework and NGSS* to be realized, school districts and teachers will need curriculum materials, assessments and professional learning that embody that vision—they will need a coordinated comprehensive implementation system. PBL with its focus on students finding solutions or building understandings in a collaborative manner represents a promising approach to use with the NGSS. As our research shows through several cycles of enactments, however, modifications to the design features of PBL need to be made to properly address challenging learning goals aligned to the NGSS and ensure the engagement of learners.

Our design-based research project that we have conducted over multiple cycles has focused our redesign on several key features of PBL. We encourage other researchers and designers to engage with these features and test their efficacy in further settings.

- Professional learning and community building that puts teachers in the role of “figuring out” that students need to grapple with making sense of complex and compelling phenomena or meaningful problems and that such efforts take time. Learners build knowledge that they can use over time as they grapple with situations that are important to them.
- Materials need to focus on supporting teachers' understanding of what the learning goal is and how the artifact helps learners build that understanding as well as social emotional goals of sticking to tasks.
- Scaffolds, including various forms of embedded assessment, need to focus on using student ideas as conceptual tools to further make sense of phenomena or design solutions to problems. In such an approach, not only will students' science proficiencies increase over time, but how they orient toward science will change as well.

The ML-PBL project focuses on integrating the vision of student engagement of PBL environments to support the 3-dimensional learning vision of the *Framework and NGSS*. Although the coordination of materials for teachers and students, professional learning and assessments show promise in promoting learning and social emotional outcomes, more work needs to be done, particularly around how to maintain such a capacity for supporting teachers' implementation over time.

Project-based learning that merges with the vision of 3-dimensional learning from the NGSS promises to promote increased engagement and learning. The work reported in this paper provides insights for how to iteratively design teacher and student project-based materials to focus on students meeting learning goals yet maintain a high level of student cognitive engagement. More broadly, this work presents one approach, a promising avenue where a comprehensive implementation system supports a generation of learners to build their capacity to work together to innovate together, and to solve the complex problems of our times.

Acknowledgment

Funding for research was provided by George Lucas Educational Research.

References

- Barab, S., Squire, K., 2004. Design-based research: putting a stake in the ground. *J. Learn. Sci.* 13 (1), 1–14.
- Barab, S.A., Sadler, T.D., Heiselt, C., Hickey, D., Zuiker, S., 2007. Relating narrative, inquiry, and inscriptions: supporting consequential play. *J. Sci. Educ. Technol.* 16 (1), 59–82.
- Barron, B.J., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., Bransford, J.D., 1998. Doing with understanding: lessons from research on problem- and project-based learning. *J. Learn. Sci.* 7 (3–4), 271–311.
- Blumenfeld, P.C., Soloway, E., Marx, R.W., Krajcik, J.S., Guzdial, M., Palincsar, A., 1991. Motivating project-based learning: sustaining the doing, supporting the learning. *Educ. Psychol.* 26 (3–4), 369–398.
- Blumenfeld, P.C., Krajcik, J.S., Marx, R.W., Soloway, E., 1994. Lessons learned: how collaboration helped middle grade science teachers learn project-based instruction. *Elem. Sch. J.* 94 (5), 539–551.
- Boaler, J., 2002. The development of disciplinary relationships: knowledge, practice and identity in mathematics classrooms. *Learn. Math.* 22 (1), 42–47.
- Bransford, J.D., Schwartz, D.L., 1999. Chapter 3: rethinking transfer: a simple proposal with multiple implications. *Rev. Res. Educ.* 24 (1), 61–100.
- Brown, A.L., 1992. Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2 (2), 141–178.
- Brown, J.S., Collins, A., Duguid, P., 1989. Situated cognition and the culture of learning. *Educ. Res.* 18 (1), 32–42.
- Charara, J., Miller, E.A., Krajcik, J., 2021. Knowledge in use: designing for play in kindergarten science contexts. *J. Leadership Equity Res.* 7 (1), n1.
- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., Schauble, L., 2003. Design experiments in educational research. *Educ. Res.* 32 (1), 9–13.

- Collins, A., Joseph, D., Bielaczyc, K., 2004. Design research: theoretical and methodological issues. *J. Learn. Sci.* 13 (1), 15–42.
- Condliffe, B., Quint, J., Visher, M.G., Bangser, M.R., Drohojowska, S., Saco, L., Nelson, E., 2017. *Project-based Learning. A Literature Review* (New York).
- Dewey, J., 1902. *The Child and the Curriculum*. The University of Chicago Press, Chicago, IL.
- Duke, N.K., Halvorsen, A.L., Strachan, S.L., Kim, J., Konstantopoulos, S., 2021. Putting PjBL to the test: the impact of project-based learning on second graders' social studies and literacy learning and motivation in low-SES school settings. *Am. Educ. Res. J.* 58 (1), 160–200.
- Duschl, R.A., Bybee, R.W., 2014. Planning and carrying out investigations: an entry to learning and to teacher professional development around NGSS science and engineering practices. *Int. J. STEM Educ.* 1 (1), 1–9.
- Fishman, B., Penuel, W., Allen, A.R., Cheng, B., Sabelli, N.O.R.A., 2013. Design-based implementation research: an emerging model for transforming the relationship of research and practice. *Teach. Coll. Rec.* 115 (14), 136–156.
- Fortus, D., Krajcik, J., 2012. Curriculum coherence and learning progressions. In: *Second International Handbook of Science Education*. Springer, Dordrecht, pp. 783–798.
- Geier, R., Blumenfeld, P.C., Marx, R.W., Krajcik, J.S., Fishman, B., Soloway, E., Clay-Chambers, J., 2008. Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *J. Res. Sci. Teach.* 45 (8), 922–939.
- Halvorsen, A.L., Duke, N.K., Strachan, S.L., Johnson, C.M., 2018. Engaging the community with a project-based approach. *Soc. Educ.* 82 (1), 24–29.
- Harris, C.J., Penuel, W.R., D'Angelo, C.M., DeBarger, A.H., Gallagher, L.P., Kennedy, C.A., Cheng, B.H., Krajcik, J.S., 2015. Impact of project-based curriculum materials on student learning in science: results of a randomized controlled trial. *J. Res. Sci. Teach.* 52 (10), 1362–1385.
- Helle, L., Tynjälä, P., Olkinuora, E., 2006. Project-based learning in post-secondary education—theory, practice and rubber sling shots. *High Educ.* 51 (2), 287–314.
- Jimenez, T.C., Graf, V.L., Rose, E., 2007. Gaining access to general education: the promise of universal design for learning. *Issues Teach. Educ.* 16 (2), 41–54.
- Kokotsaki, D., Menzies, V., Wiggins, A., 2016. Project-based learning: a review of the literature. *Improv. Sch.* 19 (3), 267–277.
- Krajcik, J.S., Blumenfeld, P.C., 2006. Project-based learning (pp. 317–34). In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge, New York, pp. 317–334.
- Krajcik, J.S., Blumenfeld, P.C., Marx, R.W., Bass, K.M., Fredricks, J., Soloway, E., 1998. Inquiry in project-based science classrooms: initial attempts by middle school students. *J. Learn. Sci.* 7, 313–350.
- Krajcik, J.S., Palincsar, A., Miller, E., 2015. *Multiple Literacies in Project-Based Learning*. Michigan State University, East Lansing MI.
- Larmer, J., Mergendoller, J.R., 2010. Seven essentials for project-based learning. *Educ. Leader* 68 (1), 34–37.
- Lee, O., Miller, E., 2016. Engaging in phenomena from project-based learning in a place-based context in science. In: Oliveira, L.C. (Ed.), *The Common Core State Standards in literacy in history/social studies, science, and technical subjects for English language learners: Grades 6–12*. Teaching English to Speakers of Other Languages, Alexandria, VA, pp. 59–73.
- MacDonald, R., Cook, H.G., Miller, E.C., 2014. *Doing and Talking Science: A Teacher's Guide to Meaning-Making with English Learners*. Board of Regents of the University of Wisconsin System, Madison, WI.
- Miller, E., Berland, L., 2020. In what contexts do teachers experience changing their science teaching as satisfying?. In: *International Society of the Learning Sciences (ISLS) ICLS 2020 Proceedings*.
- Miller, E.C., Krajcik, J.S., 2019. Promoting deep learning through project-based learning: a design problem. *Disciplin. Interdisciplin. Sci. Educ. Res.* 1 (1), 1–10.
- Miller, Shelton, 2021. The role of crosscutting concepts in teacher sensemaking and empowerment. In: Nordine, J., Lee, O. (Eds.), *Crosscutting Concepts Strengthening Science and Engineering Learning*. NSTA Press, Arlington, VA, pp. 357–377.
- Miller, E., Januszky, R., Lee, O., 2015. NGSS in action. *Sci. Child.* 53 (2), 64.
- Miller, E.C., Severance, S., Krajcik, J., 2020. Connecting computational thinking and science in a US elementary classroom. In: *Integrated Approaches to STEM Education*. Springer, Cham, pp. 185–204.
- Miller, E.C., Reigh, E., Berland, L., Krajcik, J., 2021a. Supporting equity in virtual science instruction through project-based learning: opportunities and challenges in the era of COVID-19. *J. Sci. Teach. Educ.* 1–22.
- Miller, E.C., Severance, S., Krajcik, J., 2021b. Motivating teaching, sustaining change in practice: design principles for teacher learning in project-based learning contexts. *J. Sci. Teach. Educ.* 32 (7), 757–779.
- Moursund, D.G., 1999. *Project-based Learning Using Information Technology*. International society for technology in education, Eugene, OR.
- National Research Council, 1999. *How People learn: brain, mind, experience, and school*. In: Bransford, J.D., Brown, A.L., Cocking, R.R. (Eds.), *Committee on Developments in the Science of Learning*. National Academy Press, Washington, DC.
- National Research Council, 2007. *Taking Science to School: Learning and Teaching Science in Grades K–8*. The National Academies Press, Washington, DC.
- National Research Council, 2009. *Learning science in informal environments: people, places, and pursuits*. In: Bell, P., Lewenstein, B., A.W. (Eds.), *Committee on Learning Science in Informal Environments*.
- National Research Council, 2012. *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, by States*. The National Academies Press, Washington, DC.
- Penuel, W.R., Fishman, B.J., 2012. Large-scale science education intervention research we can use. *J. Res. Sci. Teach.* 49 (3), 281–304.
- Penuel, W.R., Fishman, B.J., Haugan Cheng, B., Sabelli, N., 2011. Organizing research and development at the intersection of learning, implementation, and design. *Educ. Res.* 40 (7), 331–337.
- Reigh, E., Miller, E., 2020. Viewing science as practice: implications for the language development and inclusion of english learners. In: Gresalfi, M., Horn, I.S. (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, 4*. International Society of the Learning Sciences, Nashville, Tennessee, pp. 2269–2272. <https://repository.isls.org/handle/1/6529>.
- Rosebery, A.S., Warren, B., Tucker-Raymond, E., 2016. Developing interpretive power in science teaching. *J. Res. Sci. Teach.* 53 (10), 1571–1600.
- Sandoval, W., 2014. Conjecture mapping: an approach to systematic educational design research. *J. Learn. Sci.* 23 (1), 18–36.
- Smith, J., Nadelson, L., 2017. Finding alignment: the perceptions and integration of the next generation science standards practices by elementary teachers. *Sch. Sci. Math.* 117 (5), 194–203.
- Strauss, A.L., Corbin, J.M., Niewiarra, S., Legewie, H., 1996. *Grounded Theory: Grundlagen Qualitativer Sozialforschung*. Beltz, Psychologie-Verlag-Union, Weinheim.
- Thomas, J.W., 2000. *A Review of Research on Project-Based Learning*. Autodesk Foundation, San Rafael, CA.
- Van Lier, L., 2006. *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective*, vol. 3. Springer Science & Business Media.
- Wang, F., Hannafin, M.J., 2005. Design-based research and technology-enhanced learning environments. *Educ. Technol. Res. Dev.* 53 (4), 5–23.
- Weiss, S.G., DeFalco, A.A., Weiss, E.M., 2005. Progressive= permissive? Not according to John Dewey... subjects matter! *Essays Educ.* 14 (1), 7.

Imagining STEM *with* and *as* place

Cynthia Nicol^a, Jennifer S. Thom^b, and Florence Glanfield^c, ^a University of British Columbia, Vancouver, Canada; ^b University of Victoria, Victoria, Canada; and ^c University of Alberta, Edmonton, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introductory words	131
STEM as integration	131
Turning to river, story, and Indigenous thought	132
Styres' framework for Indigenous thought: philosophies of land	132
Vision-(re)centering	133
Returning to the river	134
Relationships-(re)membering	134
Returning to the river	135
Knowledge-(re)cognizing	135
Returning to the river	137
Action-(re)generating	137
Returning to the river	137
Concluding words: (re)actualizing STEM <i>as</i> place through more-than-human ancestor stories	138
Teaching 1: courage to respect lessons of more-than-human ancestors	138
Teaching 2: being humble to more-than-human intelligences brings humans back down to earth	138
Teaching 3: honoring the truth of interdependence among humans and more-than-humans	139
Teaching 4: wisdom of land for sustainability	140
References	141

Introductory words

Nêhiyaw and Saulteaux scholar Margaret Kovach (2009) states, “we know what we know from where we stand” (p. 7). We stand on, live with, and learn from the traditional lands of the Indigenous peoples of Musqueam, Squamish and Tsleil-Waututh Nations (Nicol, Vancouver, Canada), Songhees, Esquimalt and WSÁNEĆ Nations (Thom, Victoria, Canada), and Cree, Blackfoot, Dene, Métis, Nakota Sioux, Iroquois, and Saulteaux/Ojibway/Anishinaabe Nations (Glanfield, Edmonton, Canada). As we write this paper, we welcome the recent rains that suppress nearby forest fires both to our south and within our provinces. The rains, clearing air of ash and haze, nourish thirsty great cedar trees as salmon make their return from the Pacific Ocean to where their journey begins in the streams and rivers, high in the Canadian Rocky Mountain headwaters.

We open this paper by acknowledging our relationship to and with land and the ways these lands, where we live and work, nourish us in multiple ways. We use place/land in these introductory words to recognize these lands as the territories of the First Nations. Through this acknowledgment, and our subsequent writing, we invite you into Cree scholar Willie Ermine's (2007) notion of an ethical space, a space that grows when different positions, worldviews, philosophies, and intentions meet, confront, and engage with each other. We recognize that different worldviews frame how place and land are conceptualized in the literature. In this paper we use place and land interchangeably.

In the next sections we consider how STEM integration is typically conceived and the ways such perspectives frame relationships between humans and land/place. What results is a view of STEM (science, technology, engineering, and mathematics in K-12 and higher education) discourses as not currently connected *with* place. More specifically, we observe and experience STEM as positioned *on* land and that which *uses* land as a resource to be manipulated for human-centred economic desires. Through this writing, we work to open an ethical space in which to explore how Indigenous and ecological worldviews can enable understanding humans and more-than-humans as connected *with* place and *as* place, and in turn, occasion radically different perspectives and possibilities for STEM.

STEM as integration

Growing international interest in STEM integration focuses on the respective disciplines providing interdisciplinary settings for engagement (e.g., National Science and Technological Council, 2018; Parkin, 2018). For many, such as the UK Royal Academy of Engineers, STEM integration grounded in engineering principles supports students in “thinking like an engineer” in order to “develop more and better engineers” and to meet the growing economic demand of STEM skills (Lucas et al., 2014, p. 1). Research in this area demonstrates how even young students build disciplinary knowledge through engineering and modeling processes by

designing earthquake resistant buildings (English et al., 2017), building and modeling bridges (English, 2017), creating periscope optical instruments (King and English, 2016), and designing gears (Deis and Julius, 2017) including windmills and circuits (Wendell et al., 2017). Studies and calls for STEM integration tend to focus on developing and applying STEM disciplinary knowledge and skills. Less common but increasing are examples and calls for STEM education that critically question cultural values and environmental implications of STEM activities, such as: Whose interests (human and more-than-human) do these STEM activities serve? How do these particular STEM activities create benefits for some and barriers for others including individuals, communities, corporations, and living systems?

In the area of mathematics education Richard Barwell (2018), drawing upon mathematics educator Skovsmose (1994) approach to critical mathematics education, responds to the global climate change challenges by examining how mathematics can format our actions and influence “what we experience and how we act” (Barwell, 2018, p. 152). In the case of the current pandemic, mathematics or STEM formats how we act and interact with each other in terms of government recommendations (or requirements depending on where you live) for 2 m physical distancing and the wearing of masks, along with restrictions in mobility, employment, business operations, and receipt of economic stimulus plans. Our human or societal response to the pandemic is formatted by the predictions of mathematical models and machine learning algorithms. And, just as Barwell notes in referencing climate change, it is our response to the crisis that changes the structure of society through implementing (or not) various policies, laws and practices. The global response to COVID-19 is changing how we, as humans, interact with each other, how we treat each other, and how we interact within the world we live as human and more-than-human worlds (Abram, 1996).

Turning to river, story, and Indigenous thought

In British Columbia (BC), glaciers shape river headwaters into mountains and valleys running north to south. Here the Columbia River begins its journey in Canada’s Rocky Mountains, draining the snow melt from high mountains and interior rainforests through flowing streams and tributaries (see Fig. 1). As it flows, the river constructs a pathway, carrying minerals from mountains to the lower valleys, tumbling boulders like pebbles while cleaning the gravel bed, transporting nutrients to the flood plains, and washing fallen trees clearing debris along the banks. An international boundary cuts through the basin’s natural landscape—rivers and mountains sliced and severed by an invisible border that divides the river’s geography. South of the boundary the river flows through glacial-formed flood plateaus before twisting and turning to meet the Pacific Ocean. As a life force, a source of energy and power, the river once welcomed tens of millions of salmon at its mouth for their annual swim 2000 km upstream and hundreds of kilometers inland to spawn in the upper basin. For thousands of years the river has fed many: salmon and trout, grizzly and black bears, eagles and herons, sea lions and great killer whales, old growth forests, mosses, and wetlands, as well as Indigenous peoples and later settler colonizers.

Since the late 1800s the river has served as a resource for human use; a source of trade currency, and something to be controlled. Flowing through mountain gorges to great plateaus meant the river’s water levels fluctuate throughout the year, flooding properties and towns. Humans had the mind-set and tools through science, technology, engineering and mathematics (STEM), to construct massive dams along the river. And while dams offered solutions to managing the river’s unpredictable floods and producing electric power, these solutions also created problems, including severe destruction of the river’s spawning ocean salmon. In response, humans offered more solutions. Through STEM fish were managed by building ladders for them to move up the dammed river, fish hatcheries that raised and released juvenile fish to increase salmon stock production, and fish cultivation to satisfy increased demand. The solutions, perceived from a particular stance, positioned STEM *over* place where humans took control *over* the river, the salmon, and the entire ecosystem.

Styres’ framework for Indigenous thought: philosophies of land

Kanien’kehá:ka Mohawk scholar Sandra Styres (2017) offers a storied circular framework “about the ways Indigenous philosophies are formed and informed by connections to land and understandings of self-in-relationship” (p. 19) in her book *Pathways for Remembering and Recognizing Indigenous Thought in Education: Philosophies of Iethi’nihsténha Ohwentsia’kékha (Land)*. Concepts of circularity through what Styres refers to as the elements Vision, Relationships, Knowledge, and Action, animate how Indigenous philosophies and pedagogies come from and are informed by relationships with land, language use, and self. The framework offers us opportunities for lively discussions that embody circularity with which to (re)think connections to land, opportunities for what Stó:lō scholar Q’um Q’um Xiiem Jo-ann Archibald (2008) refers to as Indigenous storywork, and reflections on our own (re)forming relationships with land including with human and more-than-human (Abram, 1996) worlds. Styres’ dynamic, generative, and fluid framework enables each element to be in relation to the collective whole. The unifying concepts of circularity and recursivity emphasize Archibald’s (2008) emergent Indigenous storywork principles of synergy, inter-relatedness, and holism. In this section of the paper we introduce, and then use, Styres’ framework to examine and articulate examples of imagining STEM *with* and *as* place/land.



Fig. 1 Columbia river water basin. Source: Columbia River Water Basin, Kmusser, CC BY-SA 3.0, <https://creativecommons.org/licenses/by-sa/3.0>, Via Wikimedia Commons: <https://commons.wikimedia.org/wiki/File:Columbiarivermap.png>.

Vision-(re)centering

The smoke is the place of visioning—where we imagine what is or could be possible, where we identify and articulate the importance of doing this work, why now—here—in this time. Visioning opens opportunities for deep and profound insight and introspection as we explore and examine the ways we engage with the natural, spiritual, and built worlds.

Styres (2017, p. 5).

Land, in Styres' framework is more than a physical and geographic place, land is "mother earth", a sentient being that both cares for us and requires care from us. Land in this sense includes waterways, oceans, sky, urban and rural areas, the animate and inanimate, and the human and more-than-human. Okanagan poet and scholar Jeannette Armstrong (1998) writes of the life force of land that "holds all knowledge of life and death and is a constant teacher ... the land constantly speaks. It is constantly communicating" (p. 176). According to Armstrong it is this understanding of land that brings forth languages where, not only do the vocal sounds parallel the rhythm and movement of mother earth, they also engender respect and recognition of co-dependency and reciprocity for survival. This connectedness is dynamic and reinforced through the spoken word of the language and stories. In this way, the land speaks and thus has (educational) agency.

Other worldviews, such as settler colonial perspectives, envision nature, place/land as a resource to be named, used, and exploited (Bowers, 2016). These colonial spaces are, note Larsen and Johnson (2017), created by displacing and possessing in acts of oppression that reject and ignore peoples' (human and non) relationships with land. Colonial conceptions of place and land, privilege land as serving pathways for human agency, where land is understood as a backdrop to, even a stage for, human activity such as STEM.

Land as a backdrop for STEM activities involves what Skovsmose (1994) calls mathematical formatting--where mathematics participates in creating our world. For instance, mathematical modeling formats our responses and decisions through what can be measured, predicted and controlled. A measurable way of knowing the world, indeed a STEM literacy way of knowing the world, does not typically include our ancestors' stories, or what Styres (2017) calls very old pedagogies, or what ecophilosopher David Abram (2011) calls the "sensory craft of listening and the sensuous art of storytelling" about local and direct experiences in a particular place (p. 16). Consider again the Columbia River. Viewing STEM from the perspective as using place, the consequences of damming the Columbia River, including displacing lives and disrupting relationships *with* place/land, are seen as minimizing costs for maximizing human and economic gains.

Visioning and re(centering) STEM education to include an eco-critical focus is also found in the works of Wolfmeyer, Lupinacci and colleagues (e.g., Wolfmeyer et al., 2018). Drawing upon ecojustice education frameworks, these authors examine possibilities for STEM education that emphasize the inseparability of environmental and social justice issues. Like ecological philosopher Plumwood (2014) and ecojustice theorist Bowers (2016), Wolfmeyer and colleagues challenge the human/nature dualism that assumes humans as possessing higher consciousness and thus agency over nature. The authors argue that such assumptions enable the further rationalizing of humans as exceptional, separating the rest of the world as non-human (animals, forests, rivers, oceans, air, etc), and reducing them as mere resources, tools or slaves with which to serve human needs. As an example, Wolfmeyer et al. (2018) offer the story of incarceration as a possible critical STEM project to engage students in questioning the role and consequences of this human/nonhuman divide. Their project highlights the inhumane practices that human confinement embed through the physical imprisonment and incarceration (slavery) of nonhuman animals through agriculture (e.g., dairy farms), experimentation (e.g., research on animals), and human entertainment (e.g., zoos and aquariums). This leads us to ask: In what ways can we (re) vision and (re)center STEM education so we may see ourselves as part of ecosystems rather than as exceptional beings separate from nature? And in doing so, how do these ways then allow us to attend to the interests of humans and the more-than-human worlds (Kimmerer, 2013)?

Returning to the river

Consider the harnessing power of the Columbia River running through BC, Canada and the two American states of Washington and Oregon. Recognized as the fourth largest North American river (by volume), the Columbia runs across more than 2000 km of land, dropping through rocky canyons and high mountain ranges before meeting the Pacific Ocean in Oregon (Northwest Power and Conservation Council, 2020a). In one year, the river moves an average of 192 million acre-feet of water to its mouth (an acre-foot is an engineering term to describe the volume of water covering one acre one-foot deep). The river's power, length, and flow, create a home for fish and other plants and animals/wildlife. For Indigenous peoples the river sustained a way of life, language and culture. Salmon and salmon fishing provided economic, ceremonial, and nutritional benefits as well as forms of cultural practices. For colonial settlers, the river, predominantly held potential for agriculture irrigation and hydroelectric power, resources necessary to respond to the growing industrial and consumer demands of the two countries.

Between 1933 and 1941 the Grand Coulee Dam was built on the Columbia River in Washington state to address flooding issues along the river, as well as the need for irrigation, less expensive power generation, and job creation during the Great Depression. The Grand Coulee Dam is a massive structure, considered at the time as one of the largest concrete structures in the world, and known by some as the world's Eighth Wonder (McClung, 2009). Forming a lake 243 km long behind it, the Grand Coulee created a spillway of 393 m discharging about 1,000,000 L of water per second (Northwest Power and Conservation Council, 2020a). The dam represents a human STEM triumph designed to transform the river's current, gravity flow, and power into electricity and irrigation. But at what cost to place/land?

Relationships-(re)membering

To envision other possibilities for STEM, possibilities that decenter human's asymmetrical use of place for STEM, we return to Styres' circular framework and the next element of Relationships and (Re)membering. For Styres (2017) (re)membering involves the "ways one is able to bring forward into the present a conscious awareness of past events, individuals, and knowledges ..." (p. 43). A new calling to mind of "new dynamic ways to (re)member our relationships to Land" (p. 43).

Indigenous perspectives emphasize the significance of place as acknowledged in its power to shape experiences, language and culture (Deloria and Wildcat, 2001). Place shapes us and defines us. Relational and embodied, place imbues the complex interconnections that make humans, "as much a part of the land as any other element" (Coulthard, 2010, p. 80). This view aligns with what Tewa biologist Greg Cajete (2004) calls natural democracy--an interdependence of humans "with plants, animals, stones, water, and everything else" (p. 46). Moreover, that land is alive with stories and our teacher (Abram, 2011; Simpson, 2017; Styres, 2017). Contrast this with a more colonial understanding and experiencing of place. For example, place known and lived as an abstract geometric position rather than an intimate experience of dwelling (Abram, 2011) enables objectifying, separating, and privileging

humans over all other life forms. Such logic which is rooted in understanding place and land as objects, can give rise to systemic forms of domination by humans such as ownership and exploitation, violence and oppression, not only over land but also over people, language, and cultures (Calderon, 2014; Seawright, 2014).

The importance of place in education is not new and over the past few decades has taken a number of directions. American teacher-scholar Gardner Seawright (2014) examines three directions in his analysis of place, education and settler colonialism: (1) place education grounded in the local that begins with community and environmental issues important to these contexts (e.g., Sobel, 2004); (2) critical consciousness as a philosophy that encourages re-construction or re-learning to live in more relational and holistic ways in place (e.g., Greenwood, 2013); and (3) Indigenous perspectives in which place is considered as a significant relationship between human and more-than-human worlds (e.g., Barnhardt and Kawagley, 2005; Cajete, 1994; Onwu and Mosimege, 2004). A view of place education informed by Indigenous and ecological perspectives focuses on a more “harmonious relationship” between humans and the natural world compared to perspectives that support colonial systems of domination (Seawright, 2014, p. 561). Following Seawright, American scholar Calderon (2014) examines how settler notions of territory or land not only privilege settler access to land but also eliminate and dis-place Indigenous peoples on those lands, severing their deep connections with land, language and culture developed over thousands of years. Land education therefore requires, as Calderon (2014) argues, engagement in studies of place and land beyond a focus on environmental education.

Similarly, Bowers (2017) critiques critical perspectives on place stating that such views focus on power relationships between humans while neglecting the complex power relationship of humans and the environment. Bowers (2008) argues much of the current environmental/cultural ecosystem degradation comes from human domination. He contends that a critical pedagogy of place education does not address these issues as it perpetuates the very philosophical assumptions that contribute to technological capitalism. This, Bowers (2011) asserts, produces double binds (Bateson, 1972) wherein attempts to address a problem only serve to exacerbate it and thereby (re)create the same dilemma.

Understanding place relationally through Indigenous and ecological perspectives enables a more fluid sense of place/land. Rather than a Cartesian or geographically positioned view of place, a more relational perspective allows for place to flow with other places and with human and more-than-human worlds (Rubel and Nicol, 2020). It is important to clarify that we do not advocate for views that essentialize people to a specific type of education (e.g., academic or non-academic/vocational) designed specifically for that place thereby fixing people to that place. Instead, we support educational perspectives that emphasize historical, cultural and environmental distinctiveness of a particular place. However, we also recognize, following Bowers (2015), that places are also made, commodified, named, claimed, and even stolen.

Returning to the river

The Grand Coulee Dam, built with human needs in mind, rose 116 m above the river’s bed. Consequently, the dam was far too massive to accommodate any type of constructed fish passage and as such, severed fish pathways. Salmon were unable to return from the Pacific Ocean to spawn further upstream in BC waters. Canadian governments, when consulted, approved the dam, citing there were no economic losses for BC—salmon spawning above the dam were not considered part of the economy’s commercial fishery. However, Indigenous peoples living upstream and across the international boundary in Canada found they no longer had, as Sylix Elder Hazel Squakin, states, “access to culture ... [for] when you lose fish, you lose culture” (Columbia Basin Trust, 2020).

In 1964 Canada and the United States signed the Columbia River Treaty agreeing on the co-development, regulation and management of the river (Columbia Basin Trust, 2020). The Treaty required Canada to build three dams along the Columbia River Basin from 1967 to 1973. Currently, the dams bring \$120 million (US) to BC each year in power benefits, helping to: control floods along the river’s communities; provide further dam development projects and generator stations; increase economic opportunities; and create low-cost hydroelectricity for BC industries and communities (Columbia Basin Trust, 2020).

Knowledge-(re)cognizing

Styres (2017) offers Knowledge-(Re)cognizing as a third element in her framework of Indigenous philosophies of land. What diverse knowledges support visioning STEM *with place*? Here we explore knowledge sources and perspectives within geographical, linguistic, and culturally diverse settings. This aspect of the framework includes Indigenous teaching practices formed and informed by relationships with land that is, “contemporary (re)memberings of ancient knowledges and very old pedagogies,” (Styres, 2017, p. 81). Examining and increasing our awareness of these relationships and ourselves-in-relation to land provide new opportunities for collective knowing and enacting. “Place is a way of knowing, experiencing and relating with the world” (Coulthard, 2010, p. 79) and thus embraces a broad view including animals and humans, oceans and waterways, and mountains and trees, and ancestors. We acknowledge there are multiple ways of being in/*with* place and thus multiple ways of knowing.

Armstrong (2017) connects language with land, stating “All my elders say that it is land that holds all knowledge of life and death and is a constant teacher ... that the land constantly speaks. It is constantly communicating. Not to learn its language is to die.” (p. 142). What does it mean to speak the language of place, of land? For many it involves observing, listening to the places, attending to experiences of the places we inhabit. Like Glanfield et al. (2020) who offer the term “land-guaging,” Anishinaabe lawyer John Borrows (2018) writes that Anishinaabeg speakers refer to a “langscape” in which language and land form a place “where physical space interacts with human observation to give meaning to the natural and human worlds” (Borrows, 2018, p. 51). Language and meaning therefore are grounded in, or further still, inherently borne of the land. Indigenous languages reflect

this birth, this living, in the language itself. Learning to speak a dialect of Anishinaabemowin, Potawatomi plant ecologist [Kimmerer \(2013, 2017\)](#), writes how the land (the world) in her language is understood as constantly in motion with agency. Similar to other Indigenous languages ([Cruikshank, 2005](#)), and unlike the noun-based English language, Anishinaabemowin is mostly verb-based. It is a language of animacy. [Kimmerer \(2017\)](#) explains:

Rocks are animate, as are mountains and water and fire and places. Beings that are imbued with spirit, our sacred medicines, our songs, drums and even stories, are all animate. The list of inanimate seems to be smaller, filled with objects that are made by people. Of an inanimate being, like a table, we say "What is it?" And we answer Dopwen yewe. Table it is. But of apple we must say, "Who is that being?" And reply Mshimin yawe. Apple that being is. (p. 132).

The English language, on the other hand, refers only to humans and perhaps pets, as being other than "its." Nonhumans are considered distinct and often singular things and objects, even in their absence, as no-thing. Thinking and writing in an inanimate language, emphasizes human thought and knowledge as separate from, and in many cases, superior to the more-than-human world. As such, opportunities to deepen relationships with and acknowledge respect for connections to other forms of intelligence, other ways of being with the world are limited. Things with and in animate languages elucidates human and nonhuman entanglements, and discloses a world alive with agency.

For example, understanding rivers, beavers, trees, and salmon as endowed with agency and wisdom creates a very different kind of coexistence with them—one of respect and humility, one of openness to listen to others. "Imagine," writes [Kimmerer \(2013\)](#) "the access we would have to different perspectives, the things we might see through other eyes, the wisdom that surrounds us" (p. 133). [Kimmerer](#), noting how language limits our ability to show respect for the more-than-human beings of the world suggests that a change in grammar could help us (re)cognize our interconnectedness to all beings. Similarly, anthropologist [Cruikshank \(2005\)](#), in her conversations with Athapaskan and Tlingit Elders (northern Alaska, USA), points to how dominant languages such as English make it more challenging to consider landscapes, mountains, rivers and glaciers as animate and having agency. Language defines our relationships and how we (re)cognize our relationships with others. By (re)cognizing humans as superior beings we diminish and obscure our responsibility to care for the more-than-human world.

Distinguishing the world as animate and inanimate posits one (animate) as more powerful and deserving than the other (inanimate). Métis scholar Elmer Ghostkeeper, says "[f]rom my perspective, I am the land, I am the environment. There is no separation between us and the land. We are 70% air, water and heat—the same as the land" (Ghostkeeper as quoted in [Surkan, 2018](#), para. 5). Scientist and environmental activist David Suzuki expands on this point arguing that viewing our bodies as separate from our world leads to thinking that humans are above and more deserving of the earth's resources (Suzuki in [Films for Action, 2010](#)). Suzuki contends there is no divide between us and the more-than-human world. There is no environment out there, outside of our bodies, that we watch and interact with, "we are the environment ... whatever we do to our surroundings we do to ourselves ... we can't draw a line and say air ends here and I begin there. There is no line. It's in us. It's stuck to us. We are air" (Suzuki as quoted in [Films for Action, 2010](#)). The animate/inanimate dichotomy does not exist.

Forest ecologist Suzanne Simard challenges such a dichotomy as she documents how trees communicate with each other ([Bingham and Simard, 2012; Simard, 1990](#)). Using radioactive isotopes of carbon, Simard found that Douglas fir and paper birch trees communicated using an underground network of fungi, known as mycelium. All trees, Simard contends, form communicative relationships through mycelium, facilitating nutrient, carbon and water exchange. Communication extends beyond resource transfers and to kin recognition—what Simard refers to as defense signaling. Trees respond when a neighboring tree needs support. For example, when Douglas fir were shaded the birch transferred its excess carbon to the fir, and when the birch lost its leaves in the fall while the Douglas fir still photosynthesizing, transferred its excess carbon to the birch. The mycelium threads itself underground. By ensuring these transfers occur, the mycelium sustains itself codependently with trees. Furthermore, Simard and her research team found that older trees with larger root systems have associated larger mycelium networks and therefore more connections to other trees in the network. The research team even found that these older trees recognize their own seedlings, their kin. Named by Simard's team as "mother trees," these old trees when injured also communicated or signaled this through carbon release into the network preparing other trees for a possible new role in the ecosystem.

These examples of STEM with place provide insight into Styres' notion of (re)cognizing human and more-than-human intelligences. The trees, fish, rivers, are considered as kin, part of a neighborhood, a community of living in co-dependency, a family communicating with each other. Conceptualizing STEM in these ways requires awareness of how language and culture frame what is possible and impossible regarding our relationship with one another. How can we (re)cognize the unique intelligence of the more-than-human world and, as [Abram \(2011\)](#) writes, "hear those many voices that do not speak in words"? (p. 16). [Abram \(2011\)](#) makes a strong argument for reconnecting with oral culture, what [Styres \(2017\)](#) might call the act of (re)cognizing very old pedagogies. [Abram \(2011\)](#) writes:

to live in a storied world is to know that intelligence is not an exclusively human faculty located somewhere inside our skulls, but is rather a power of the animate earth itself, in which we humans, along with the hawks and the thrumming frogs, all participate. It is further to know that each land, each watershed, each community of plants and animals and soils, has its particular style of intelligence, its unique mind or imagination evident in the particular patterns that play out there, in the living stories that unfold in that valley, and that are told and retold by the people of that place" (p. 12–13).

These examples of STEM *with place* also challenge frameworks that privilege human-centeredness, and confront what environmental philosopher Plumwood (2014) refers to as the “hyperseparation” of human and nonhumans that “is expressed in denying both the mind-like aspects of nature and the nature-like aspects of the human” (p. 444). As Kimmerer (2013), addresses, even our language separates humans from more-than-human, and in so doing reinforces a false sense of human superiority, agency, and autonomy which in turn blinds us from understanding how interconnected and inter-dependent we are with nature’s ecosystems, Simard’s research questions whether mind is solely a human phenomenon while Ghostkeeper’s “I am the environment” and Suzuki’s “we are air” philosophies deny separation between humans and world. Visioning STEM *with place*, could lead us to be, as theoretical physicist and biologist Stuart Kauffman says “of the earth rather than above the earth” (Kauffman as quoted in Price, 2020). More-than-human worlds are active and intelligent.

Returning to the river

The Columbia Basin dams change every aspect of the river: its speed, depth, temperature, volume, water quality, and what it carries as well as the life it sustains. Flooding mountain valleys displaced thousands of people, flooded communities, submerged Indigenous cultural and burial sites, devastated fish and wildlife populations, and wiped-out access to fertile soil for animal and plant life. Once one of the highest salmon producing rivers in the world, reports now count 106 runs of the river’s salmon as extinct (Northwest Power and Conservation Council, 2020b). However, dams are not the sole cause of salmon depletion or what environmental historian Joseph Taylor (1999) calls the “salmon crisis” (p. 3). The Columbia River receives toxic waste from thousands of pipes along the many tributaries. Factories, industries, wastewater treatment plants, as well as a nuclear reservation, all pump massive toxic chemicals into the river. But such pollution is not new and, in fact, treating the river as a waste deposit goes back as far as the late 1800s (Taylor, 1999).

This story of the Columbia River, of conflict between settler industry and salmon runs, depicts a change in the relationship between humans and fish. What Blackfoot scholar LeRoy Little Bear (2000) refers to as “jagged worlds colliding.” The Columbia River and tributaries flow without conforming to international boundaries. The US–Canada border on the 49th parallel slices hundreds of waterways, but for thousands of years no such border existed. Salmon made their way through cascading rock waterfalls such as those at Kettle Falls, a gathering place for Indigenous Salish groups to trade, harvest, and engage in cultural exchange.

Action-(re)generating

The fourth aspect of Styres (2017) circular framework is Action, a process of (re)generating and working toward actualizing the Vision; in our case, imagining STEM *with place*. For Styres (2017) “action and (re)generation represent resistance to and (de)centering of colonially systemic relations of power and a (re)centering of Indigenous thought” (p. 123). Here an “outcome of action leads to (re)generated and (re)actualized vision through a process that is dynamic, circular, and iterative” (Styres, 2017, p. 123). This meaning should not be taken as synonymous with the Latin meaning of regeneration as “being born again.” Styres problematizes the thinking which such language engenders by describing industrialized forest reforestation practices primarily intended to replenish trees cut down by commercial logging of entire forests.

Here, as with the Columbia River, perpetuating land exploitation and commodification of trees suggests regeneration through reforestation and “could be seen as another form of slavery” (Styres, 2017, p. 124); that is, ultimately, culturally and environmentally destructive practices remain directed toward promoting human material wealth and consumer needs. In considering these choices as ethical rather than preferential choices, Suzuki urges us to think carefully about our relations with land asking: “is the river the veins and capillaries of the land or an opportunity for irrigation and energy? Is the forest a sacred grove or just an opportunity for lumber and pulp?” (Zuckerbrot and Zuckerbrot, 2016). Like Suzuki’s description of a river, Indigenous meanings of (re)generating imbue self-in-relation to land as being “born of the land” (Styres, 2017, p. 124). Such meanings land-guage circularity and connections to the more-than-human world as sentient, (re)membering humans as earth through deep and spiritual relationships *with place* (Styres, 2017). Our paper contributes to conversations about how we, as humans, can bring forth a more-than-human world. STEM *with place* draws attention to imagining goals and purposes for STEM education beyond economic ones whereby pedagogies as more-than-human engender understanding places as teacher.

Returning to the river

We continue to follow the Columbia River from high in the Canadian Rocky Mountain headwaters, across the terrain, through forests, down to the watershed and now out to the ocean. Throughout this paper, the river introduced us to more-than-human ancestors, including forests, rivers, mountains, mosses, fish, and even air. These ancestors, as teachers, invite us to imagine STEM as an ecology, offering lessons and stories of coexistence or what Kauffman calls, “a well-lived life” (Kauffman as quoted in Price, 2020). However, how do we open ourselves to receive these kinds of teachings when the many voices, languages, and ways of communicating “do not speak in words” (Abram, 2011, p. 16)?

(Re)awakening ourselves to the power of our ancestors’ stories requires “synergistic interaction” between storyteller (e.g., forests and their underground fungal networks) and listener (e.g., humans) (Archibald, 2008, p. 33). This involves a kind of listening that requires all our senses to be aware of and attune to the living forms and rhythms of our more-than-human ancestors. Such listening makes it possible to recognize not only the intelligences of our ancestors but also the varied patterns generated and (re)generated by

them. Returning to the river, we realize there are more and different stories if we are reading to listen. Following the river to the ocean waters, we gaze beneath the surface and to the coral reefs. Awake, we turn to these more-than-human ancestors, to further explore STEM *with* place and importantly, to (re)actualize STEM *as* place.

Concluding words: (re)actualizing STEM *as* place through more-than-human ancestor stories

Here as the result of a recursively circular process of (re)cognizing, (re)membering, and learning from corals as one of our more-than-human ancestors, we (re)generate and (re)actualize our Vision of seeing and understanding STEM not only *with* place but *as* place. We wonder what stories the corals, as one more-than-human ancestor, might share with us so that we may (re)generate STEM as land-gauge which (re)cognizes and (re)members humans as Earth? Said another way, and described earlier by [Borrows \(2018\)](#), how might language and land form a place “where physical space interacts with human observation to give meaning to the natural and human worlds[?]” (p. 51). Here we open ourselves to learn from the stories of coral reefs toward imagining STEM *as* place. In turn, we offer four teachings we see as vital to (re)actualizing STEM through renewing relationships with place.

Teaching 1: courage to respect lessons of more-than-human ancestors

Whether arctic or tropical seas, corals have been on Earth, according to scientists, for as long as 400 million years. Today’s coral reefs, which span five to ten thousand years old, embody and recursively carry forth 25 million years’ worth of evolution. Corals are ancestors. Their reefs are the largest biological structures on Earth. Next to humans, corals are the only organisms capable of transforming the planet’s landscape to such an extent. Not only are corals ancestors, from an engineering and technological perspective, they are both the architects and the architectures—i.e., architect(ure)s. Action as (re)generation reveals corals not as things but inherently “[n]ature as a dynamic [life]force” ([Styres, 2017](#), p. 129). Doing so (re)members, (re)cognizes, and (re)centers corals as creators of Earth’s massive land structures as well as the shapers and protectors of the shorelines of continents. For thousands and millions of years, borne of the land, corals are land. Similar to other natural fractal forms such as clouds, mountains, trees, and our circulatory system ([Mandelbrot, 1982](#)), corals deny spatial and temporal linearities ([Purkis and Riegl, 2006](#)). Irreducible, they cannot be simplified. The closer or farther away you look, the more detail you see. The Great Barrier Reef, whether viewed from outer space at 2500 km or at a closer range of 250 m, 25 cm, 2.5 mm, or even through an electron microscope at 25 μ m, demonstrate self-similarity and increasing complexity ([Muir, 2015](#)). As such, scientists ask “whether the fractal patterns observed in the small-scale structure of corals (e.g., [Bradbury and Reichelt, 1983](#); [Basillais, 1997](#)) are related to the scale-invariance observed at reef-scale?” ([Purkis and Riegl, 2006](#), p. 1760).

Teaching 2: being humble to more-than-human intelligences brings humans back down to earth

Now consider *Hexacorallia* (see [Fig. 2](#)) by artist, philosopher, physician, and scientist who coined the term “oecologie” (ecology) Ernst [Haeckel \(1904\)](#). Stunningly detailed, Haeckel’s print provides a glimpse into magnificence of corals as still objects. Yet, corals as dynamically emergent, reveal Nature’s ancient and living architect(ure)s as capable of even greater mathematical, engineering, technological, and scientific eloquence.

Contemporary science writer and artist Margaret [Wertheim \(2017\)](#) explains: “Many species of corals, then [400 million years ago] and now, also have hyperbolic structures, which we immediately recognize by the frills and crenulations of their forms” where “[a] head of coral is actually a colonial organism made up of thousands of individual polyps growing together; collectively organize [themselves] into a mathematical surface disallowed by Euclid’s axiom about parallel lines” (para. 3). We see corals resisting and de-centering human colonial intentions that attempt to fix, limit, simplify, and ultimately abstract. Here Indigenous thought calls heed to corals—“that it is land that holds all knowledge of life and death and is a constant teacher ... that land constantly speaks” and that “[l]and is constantly communicating” ([Armstrong, 2017](#), p. 142). Emergent is an ecological interweaving of intelligence, not solely as a human quality but of Nature as a whole in which the more-than-human world ([Abram, 1996](#)) enacts intelligently “in ways that are often beyond what humans can understand or replicate.” ([Bowers, 2011](#), p. 315) Corals elucidate what “dominant Western epistemology fails to understand ... the natural environment is not reducible to matter and blind forces” ([Bowers, 2011](#), p. 315).

Intelligently so, for hundreds of millions of years, corals exceeded human mathematics. Here, “[h]yperbolic geometry was only discovered in the 19th century, revolutionizing the field of mathematics and eventually paving the way for Einstein’s general theory of relativity.” (Margaret Wertheim as quoted in [Barnett, 2017](#), para 1.) Yet in being coral, corals are wondrously beautiful and eloquent geometer(ie)s. They live “alternative[s] to the Euclidean geometry that we learn about in school, and a form that, over hundreds of years, many great mathematical minds tried to prove impossible.” ([Wertheim, 2017](#), para. 1). Impossible only because the geometry as “measurement of earth or land” ([Online Etymology Dictionary, 2001–2020](#)) was humanly unimaginable. As 19th century mathematician Wolfgang Bolyai tried to convince his son János Bolyai of the futility regarding hyperbolic geometry and to abandon the work, he pleaded—“For God’s sake, please give it up. Fear it no less than the sensual passion, because it, too, may take up all your time and deprive you of your health, peace of mind and happiness in life” (Bolyai as quoted in [Henderson and Taimina,](#)

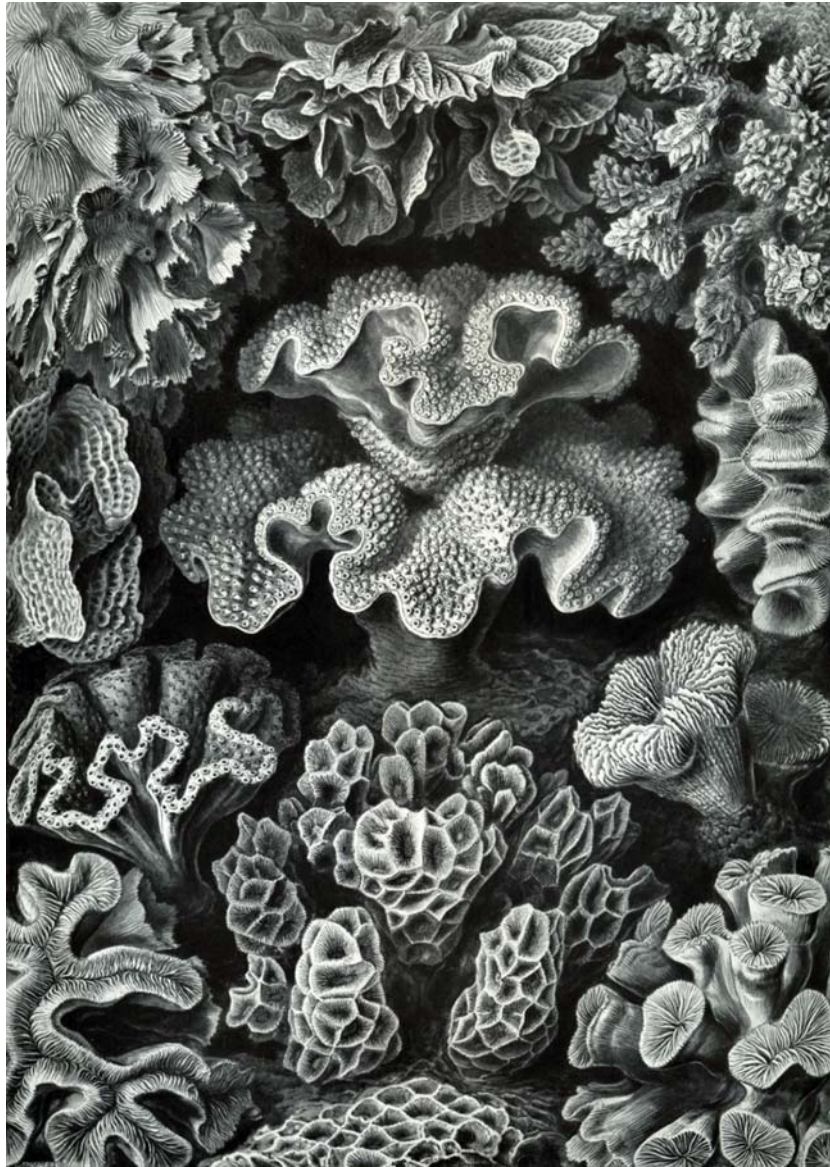


Fig. 2 Hexacoralla print by Ernst Haeckel. Source: <https://www.sciencesource.com/archive/Ernst-Haeckel-Hexacorallia-Corals-SS21690898.html>.

2001, p. 17). It was also not until 1997 when mathematician Daina Taimina explored through knitting and then discovered through crocheting that tangible models of hyperbolic planes could be made (see [Henderson and Taimina, 2001](#)). The story of mathematician Benoit Mandelbrot's discovery of fractal geometry, like hyperbolic geometry, plays out in a strikingly similar manner (see [Lesmoir-Gordon, 1994, 2016](#)). Humbled by the intelligence of corals, they bring humans back down to earth and connect humans with land.

Teaching 3: honoring the truth of interdependence among humans and more-than-humans

Consider humans with corals again. Humans have evolved to be one species and corals persist as 6000 species ([National Geographic Society, 2019](#)). Averaging not much more than 1.5 cm in diameter and no brain but a simple nerve net extending from mouth to tentacles, corals build among the most diverse ecosystems on the planet. Coral "colonies" "occupy" a mere 0.2% of the total ocean area yet the habitat enabled by these architect(ure)s supports at least one-quarter of all marine animals. These include more than 4000 species of fish, sponges, anemones, sponges, sea stars, clams, crabs as well as other mollusks and crustaceans ([United States Environmental Protection Agency, 2018](#)). The Great Barrier Reef is an ecosystem of more than 350 species of stony corals and millions of marine creatures including 1500 species of fish, 5000 species of mollusks, 215 species of birds, as well as crocodiles, varieties of dolphins, green turtles, whales, and dugongs ([Zimmerman, 2012](#)). Indeed, it is hard to imagine how this is

possible. Next, we describe three land-guage acts of (re)generating wherein we identify insights as they arise while learning about interrelationships among corals and the different animals that make up ocean “colonies.” We consider these land-guaging acts as a way that (re)generatively (re)stor(i)es interdependence as collaborative and mutually beneficial; lived and living; and, dwelling and (re)building.

First, within coral colonies, algae that inhabit the stony coral polyps photosynthesize and produce almost all nourishment which the coral polyps need to live. The corals in return make proteins, fats, carbohydrates, calcium carbonate, secreting an exoskeleton which provide shelter, minerals, and carbon dioxide which the algae must have to survive. These actions contrast starkly with human colonial practices of dominance that separate the natural world including humans, as “lesser” and “other” to gain cooperation from them with the sole purpose of controlling and exploiting them. Cooperation between algae and coral is not only mutually beneficial but absolutely essential to the survival and flourishing of both creatures.

Second, while alive, stony corals continuously secrete calcium carbonate which extends the reef’s foundation by enlarging their individual skeletons and adding chambers underneath. Even when the polyps die, their skeletons remain and become the foundation for new coral polyps—billions—to attach and continue the building process of the reef through further layering on from the previous skeletons. In this way, (re)generating of the reef necessitates both those corals which are living and those which have lived.

Third, creatures other than corals and algae also play integral roles in creating the complex architectures of reefs. Red algae, for example, produce their own encrusting structures that preserve and keep fragments or loose sediment attached to the reef which would otherwise wash away. And, marine animals such as clams, snails, and sponges that live out their lives in the crevices and holes of reefs eventually die and become the land on which new coral polyps grow. Through the reciprocal process of life and death, remains left behind by these dwellers continuously and regeneratively rebuild the reef.

Teaching 4: wisdom of land for sustainability

Following the Columbia River out to the ocean and turning our gaze beneath the water’s surface teaches us much. What we learn from corals is that (re)generative action both ecological and consistent with Indigenous thinking is inherently “dynamic, circular, and iterative” (Styres, 2017, p. 123). That the survival of the reefs is not simply and singularly driven by one species of coral or even the six thousand species of them. Rather, survival, strength, and innovation necessitate imagining sustainability as something very different than “conserving an ecological balance by avoiding depletion of natural resources” (Sustainability, 2020). Sustainability here implies living life forth as complex, sacred, and regenerative interdependence. In keeping with Styres’ framework (2017), sustaining coral reefs *as place*, *as land*, *as Earth*, does not imply regeneration as “being born again” or merely preventing the depletion of natural resources—e.g., engineering temperature tolerant corals. Instead, magnificently as life and death, corals and their co-inhabitants are always selves-in-relation to Land, born of the land, *as Land* (Coulthard, 2010; Ghostkeeper in Surkan, 2018; Styres, 2017). Corals enable us to understand how (re)generation can serve to further STEM, now imagined *as place*. Here arises land-guage that circles back and (re)members STEM *as place*, deeply and spiritually connected “to ancestors, lands, waterways, and mountains” (Styres, 2017, p. 124).

We see opportunities for (re)centering place/land education with STEM for (re)membering very old pedagogies of the river, forests, and coral that teach us to (re)cognize intelligences in a more-than-human world, and (re)generate STEM *with* place and *as place*. STEM *with* place emphasizes our interdependent relations and the role of place as teacher and mentor. STEM *as place* inspires collaboration beyond human worlds, moving us toward being with the world in new ways, a kind of mutual flourishing and recognition of interdependence. Ramámuri scholar Enrique Salmón introduces the concept of kincentricity to express this relationship in which humans are directly related to all of the natural world including animals, plants, and even stones (Salmón, 2017). A Ramámuri worldview, writes Salmón (2017):

[d]oes not differentiate or separate ontological spaces beyond and between the human and nonhuman worlds. We feel that we are directly related to everything around us. The trees are us; we are the trees. I am rain; rain is me. The rain is all around me; it aligns inside me. (p. 25)

Kincentricity describes STEM *with* place and *as place*. It centers kinship and with it a responsibility to honor and recognize not only the rights of humans but also more-than-humans in interdependent relations. Kincentricity acknowledges the rights of codependent systems to be generative and sustaining including the rights of rivers to flow, of coral to be architects, and forests to share wisdom. Kincentricity situates humans and STEM in relation within a more-than-human world, one where well-being reaches beyond simply a single person, peoples, or species. In doing so, we appreciate the challenge that moving beyond human contexts poses, especially now where silencing and privileging certain groups over others diminishes opportunity for all humans to live well. At the same time, understanding STEM *as place* prompts dialog and caring for the world’s current challenges: species degradation, climate change, sea levels rising, human population increasing, and new pandemic viruses like COVID-19.

“Land is a living entity, and it is constantly in conversation with those of us who know how to listen to it” (Salmón in conversation with Young, 2021). Rivers, forests, corals, and other more-than-human relations and ancestors continuously teach by enacting courage, humility, wisdom, truth, life and death on Earth. Hearing their stories, both ancient and current, brings awareness that humans are in an ethical space with more-than-human relations. This awareness opens us to not only imagine STEM *with* place, but also bring hope that humans might listen deeply and radically live STEM *as place*.

References

- Abram, D., 1996. *The Spell of the Sensuous: Perception and Language in a More-Than-Human World*. Vintage Books.
- Abram, D., 2011. Storytelling and wonder: on the rejuvenation of oral culture as an ecological imperative. In: Urbain, O., Temple, D. (Eds.), *Ethical Transformations for a Sustainable Future*. Routledge, pp. 9–17. <https://doi.org/10.4324/9781351311403>.
- Archibald, J. Q'um Q'um Xiem, 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. UBC Press.
- Armstrong, J., 1998. Land speaking. In: Ortiz, S. (Ed.), *Speaking for the Generations: Native Writers on Writing*. University of Arizona Press, pp. 174–194.
- Armstrong, J., 2017. Land speaking. In: McCall, S., Reder, D., Gaertner, D., L'Hirondelle Hill, G. (Eds.), *Read, Listen, Tell: Indigenous Stories from Turtle Island*. Wilfrid Laurier University Press, pp. 141–155.
- Barnett, R., 2017. What Happens When You Mix Math, Art, and Crochet? It's Mind-Blowing. <https://ideas.ted.com/gallery-what-happens-when-you-mix-math-coral-and-crochet-its-mind-blowing/#~:text=Hyperbolic%20geometry%20was%20discovered%20in,Einstein's%20general%20theory%20of%20relativity>.
- Barnhardt, R., Kawagley, O., 2005. Indigenous knowledge systems and Alaska native ways of knowing. *Anthropol. Educ. Q.* 36 (1), 8–23. <https://doi.org/10.1525/aeq.2005.36.1.008>.
- Barwell, R., 2018. Some thoughts on a mathematics education for environmental sustainability. In: Ernest, P. (Ed.), *The Philosophy of Mathematics Education Today*. Springer, pp. 145–160.
- Basillais, É., 1997. Coral surfaces and fractal dimensions: a new method. *Compt. Rendus Acad. Sci. III Sci. Vie.* 320 (8), 653–657.
- Bateson, G., 1972. *Steps to an Ecology of Mind*. University of Chicago Press.
- Bingham, M.A., Simard, S.W., 2012. Mycorrhizal networks affect ectomycorrhizal fungal community similarity between conspecific trees and seedlings. *Mycorrhiza* 22 (4), 317–326. <https://doi.org/10.1007/s00572-011-0406-y>.
- Borrows, J., 2018. Earth-bound: indigenous resurgence and environmental reconciliation. In: Asch, M., Borrows, J., Tully, J. (Eds.), *Resurgence and Reconciliation: Indigenous-Settler Relations and Earth Teachings*. University of Toronto Press, pp. 49–81.
- Bowers, C.A., 2008. Why a critical pedagogy of place is an oxymoron. *Environ. Educ. Res.* 14 (3), 325–335. <https://doi.org/10.1080/13504620802156470>.
- Bowers, C.A., 2011. Perspectives on the Ideas of Gregory Bateson, Ecological Intelligence, and Educational Reform. Eco-Justice Press, LLC.
- Bowers, C.A., 2015. How the Worlds of Data Destroy Ecological Intelligence. YouTube. <https://www.youtube.com/watch?v=PgM-kGB1fol>.
- Bowers, C.A., 2016. A Critical Examination of STEM: Issues and Challenges. Routledge.
- Bowers, C.A., 2017. An ecojustice approach to educational reform in adult education. *N. Dir. Adult Cont. Educ.* 2017 (153), 53–64. <https://doi.org/10.1002/ace.2017.2017.issue-153>.
- Bradbury, R.R., Reichelt, R.R., 1983. Fractal dimension of a coral reef at ecological scales. *Mar. Ecol. Prog.* 10, 169–171.
- Cajete, G., 1994. Look to the Mountain: An Ecology of Indigenous Education. Kivaki Press.
- Cajete, G., 2004. Philosophy of native science. In: Waters, A. (Ed.), *American Indian Thought*. Blackwell Publishing, pp. 45–57.
- Calderon, D., 2014. Speaking back to manifest destinies: a land education-based approach to critical curriculum inquiry. *Environ. Educ. Res.* 20 (1), 24–36. <https://doi.org/10.1080/13504622.2013.865114>.
- Columbia Basin Trust, 2020. Basin stories: first nations. In: *The Basin: A Resource of Columbia Basin Trust*. <https://thebasin.ourtrust.org/resource/basin-stories/>.
- Coulthard, G., 2010. Place against empire: understanding indigenous anti-colonialism. *Affinities* 4 (2), 79–83.
- Cruikshank, J., 2005. *Do Glaciers Listen? Local Knowledge, Colonial Encounters, & Social Imagination*. UBC Press.
- Deis, T., Julius, J., 2017. A STEM exploration with gears. *Math. Teach. Middle Sch.* 22 (8), 495–499.
- Deloria, V., Wildcat, D., 2001. *Power and Place: Indian Education in America*. Fulcrum.
- English, L., King, D., Smeed, J., 2017. Advancing integrated STEM learning through engineering design: sixth-grade students' design and construction of earthquake resistant buildings. *J. Educ. Res.* 110 (3), 255–271.
- English, L., 2017. Advancing elementary and middle school STEM education. *Int. J. Sci. Math. Educ.* 15 (S1), 5–24.
- Ermine, W., 2007. The ethical space of engagement. *Indig. Law J.* 6 (1), 193–203. <https://jps.library.utoronto.ca/index.php/ilj/article/view/27669/20400>.
- Films for Action, 2010. David Suzuki: An Elder's Vision for Our Sustainable Future. <https://www.filmsforaction.org/watch/david-suzuki-an-elders-vision-for-our-sustainable-future/>.
- Glanfield, F., Thom, J.S., Ghostkeeper, E., 2020. Living landscapes, archi-text-ures, and land-guaging algo-rhythms. *Can. J. Sci. Math. Technol. Educ.* 20 (2), 246–263.
- Greenwood, D., 2013. A critical theory of place-conscious education. In: Stevenson, R., Brody, M., Dillon, J., Wals, A. (Eds.), *International Handbook of Research on Environmental Education*. Routledge, New York, NY, pp. 93–100.
- Haeckel, E., 1904. *Hexacorallia*. Public Domain, Via Wikimedia Commons. https://upload.wikimedia.org/wikipedia/commons/9/9f/Haeckel_Hexacoralla_69.jpg.
- Henderson, D.W., Taimina, D., 2001. Crocheting the hyperbolic plane. *Math. Intel.* 23 (2), 17–28.
- Kimmerer, R.W., 2013. *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants*. Milkweed Editions.
- Kimmerer, R.W., 2017. Learning the grammar of Animacy. *Anthropology of Consciousness* 28 (2), 128–134.
- King, D., English, L., 2016. Engineering design in the primary school: applying STEM concepts to build an optical instrument. *Int. J. Sci. Educ.* 38 (18), 2762–2794.
- Kovach, M., 2009. *Indigenous Methodologies: Characteristics, Conversations, and Contexts*. University of Toronto Press.
- Larsen, S.C., Johnson, J.T., 2017. *Being Together in Place: Indigenous Coexistence in a More than Human World*. University of Minnesota Press.
- Lesmoir-Gordon, N., 1994. *The Colours of Infinity* [Documentary]. New Moon Productions.
- Lesmoir-Gordon, N., 2016. *The Colours of Infinity: The Beauty and Power of Fractals*. Springer.
- Little Bear, L., 2000. Jagged worldviews colliding. In: Battiste, M. (Ed.), *Reclaiming Indigenous Voice and Vision*. UBC Press, pp. 77–85.
- Lucas, B., Claxton, G., Hanson, J., 2014. *Thinking Like an Engineer: Implications for the Education System*. Royal Academy of Engineers. www.raeng.org.uk/thinkinglikeanengineer.
- Mandelbrot, B.B., 1982. *The Fractal Nature of Geometry*. W. H. Freeman and Co.
- McClung, C., 2009. Grand Coulee Dam: Leaving a Legacy. The Great Depression in Washington State. https://depts.washington.edu/depress/grand_coulee.shtml.
- Muir, P., October 26, 2015. Coral Polyps: Microscopic Builders of our Great Barrier Reef. <https://www.pechakucha.com/presentations/reef-corals-and-fractals>.
- National Geographic Society, 2019. Coral. www.nationalgeographic.org/encyclopedia/coral/#~:text=There%20are%20about%20six%20thousand,seafloor%20of%20the%20open%20ocean.
- National Science & Technology, 2018. Charting a Course for Success: America's Strategy for STEM Education. <https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf>.
- Northwest Power and Conservation Council, 2020a. Columbia River: Description, Creation, and Discovery. <https://www.nwcouncil.org/reports/columbia-river-history/columbiariver>.
- Northwest Power and Conservation Council, 2020b. Grand Coulee Dam: Impacts on Fish. <https://www.nwcouncil.org/reports/columbia-river-history/grandcouleeimpactsnfish>.
- Online Etymology Dictionary, 2001–2020. Geometry. [Definition] (2001–2020). In: *Online Etymology Dictionary*. October 12, 2020. <https://www.etymonline.com/word/geometry>.
- Onwu, G., Mosimege, M., 2004. Indigenous knowledge systems and science and technology education: a dialogue. *Afr. J. Res. Math. Sci. Technol. Educ.* 8 (1), 1–12.
- Parkin, A., 2018. Supporting Education Transformation in Canada. https://canada2067.ca/app/uploads/2018/11/Canada-2067-Supporting-Education-Transformation-in-Canada_Final.pdf.
- Plumwood, V., 2014. Nature in the active voice. In: Harvey, G. (Ed.), *The Handbook of Contemporary Animism*. Taylor & Francis, pp. 441–453.
- Price, J., 2020. COVID: civilizational, global, & existential. A conversation with Stuart Kauffman. Episode 55 (audio podcast episode). In: *The Sacred Speaks*.

- Purkis, S.J., Riegl, B., 2006. Fractal patterns of coral communities: evidence from remote sensing (Arabian Gulf, Dubai, UAE). In: Dodge, R.E. (Ed.), Proceedings of 10th International Coral Reef Symposium, Okinawa, Japan. https://nsuworks.nova.edu/occ_facpresentations/13/.
- Rubel, L.H., Nicol, C., 2020. The power of place: spatializing critical mathematics education. *Math. Think. Learn.* 22 (3), 1–22.
- Salmón, E., 2017. No word. In: van Horn, G., Hausdoerffer, J. (Eds.), *Wildness: Relations of People and Place*. University of Chicago Press, pp. 24–32.
- Seawright, G., 2014. Settler traditions of place: making explicit the epistemological legacy of white supremacy and settler colonialism for place-based education. *Educ. Stud.* 50 (6), 554–572. <https://doi.org/10.1080/00131946.2014.965938>.
- Simard, S., 1990. Competition Between Sitka Alder and Lodgepole Pine in the Montane Spruce Zone in the Southern Interior of British Columbia. FDRA Report. BC Government. <https://www.for.gov.bc.ca/hfd/pubs/docs/Frr/Frr150.htm>.
- Simpson, L., 2017. *As We Have Always Done: Indigenous Freedom Through Radical Resistance*. University of Minnesota Press.
- Skovsmose, O., 1994. *Towards a Philosophy of Critical Mathematics Education*. Kluwer Academics Publishers.
- Sobel, D., 2004. *Place-based Education: Connecting Classrooms and Communities*. Orion Society.
- Styres, S.D., 2017. *Pathways for Remembering and Recognizing Indigenous Thought in Education: Philosophies of Iethi'nihténha Ohwentsia'kékha (Land)*. University of Toronto Press.
- Surkan, J., 2018. The Métis Architecture: Conversations: Elmer Ghostkeeper. <https://metisarchitect.com/2018/06/04/conversations-elmerghostkeeper/>.
- Sustainability [Def 1.1], 2020. Oxford Living Dictionaries. Retrieved, October 12, 2020. <https://en.oxforddictionaries.com/definition/sustainability>.
- Taylor, J.E., 1999. *Making Salmon: An Environmental History of the Northwest Fisheries Crisis*. University of Washington Press.
- United States Environmental Protection Agency, 2018. Basic Information About Coral Reefs. www.epa.gov/coral-reefs/basic-information-about-coral-reefs#:~:text=Coral%20reefs%20are%20among%20the,point%20in%20their%20life%20cycle.
- Wendell, K., Wright, C., Paugh, P., 2017. Reflective decision-making in elementary students' engineering design. *J. Eng. Educ.* 106 (3), 356–397.
- Wertheim, M., 2017. How to play mathematics. Aeon. <https://aeon.co/essays/theres-more-maths-in-slugs-and-corals-than-we-can-think-of>.
- Wolfmeyer, M., Lupinacci, J., Chesky, N., 2018. Critical education. *Crit. Educ.* 8 (15), 37–39. https://doi.org/10.1007/978-1-137-42617-8_15.
- Young, A., 2021. Enrique Salmón on moral landscapes amidst changing ecologies. Episode 255 (audio podcast episode). In: *For the Wild*.
- Zimmerman, K., 2012. Great Barrier Reef: Facts, Location and Animals. Live Science. Retrieved, October 12, 2020. <https://www.livescience.com/6290-great-barrier-reef.html>.
- Zuckerbrot, D., Zuckerbrot, D., 2016. Suzuki@80 (season 55, episode 18). In: *The Nature of Things, Reel Time Images*.

Socioscientific issues and STEM learning

Joseph A. Johnson^a, Ryan Batkie^b, Augusto Macalalag^c, Julie Dunphy^c, and Shawn Titus^a, ^a Department of Physics, Mercyhurst University, Erie, PA, United States; ^b Department of Curriculum and Instruction, University of Wisconsin–Madison, Madison, WI, United States; and ^c School of Education, Arcadia University, Glenside, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Socioscientific issues in the context of STEM education	143
The goals of SSI in STEM education	144
Teachers' sense of purpose in teaching STEM	145
Teachers' conceptions of SSI	146
Building new curriculum and adopting new repertoires of instruction	146
Hospitality of school contexts	147
Developing teachers' pedagogical content knowledge in SSI	148
Conclusions	150
References	150

Socioscientific issues in the context of STEM education

Socioscientific issues (SSI) are debatable and ill-defined problems that have a basis in science but necessarily include moral and ethical choices and cannot be resolved through science alone (Ratcliffe and Grace, 2003). As such SSI can provide meaningful contexts for students to learn concepts and practices in science, technology, engineering, and mathematics (STEM) disciplines (Zeidler et al., 2005). As an educational reform, SSI parallels other reform efforts, such as teaching through science inquiry, the engineering design processes, technological literacy, mathematics and computational thinking, and environmental education, while remaining distinct in focus and application. This article explores SSI in the landscape of parallel educational reform movements, the goals of STEM education, and how SSI can provide an avenue for attaining what are often viewed as disparate goals. The article concludes by highlighting challenges to SSI implementation found in the literature. These challenges are followed by potential avenues to foster and support SSI implementation in STEM through teacher education and professional development.

Similar to science inquiry that promotes development and testing of scientific models through evidence-based argumentation (Abd-El-Khalick et al., 2004), SSI allows students to identify arguments and counter arguments based on models (Albe and Gombert, 2012). As science inquiry promotes understanding of and experience with the nature of science, that scientific knowledge of the natural world is tentative and relies on iterative science inquiry processes of evidence-based explanation and argumentation from people of all cultures (Flick and Lederman, 2006). SSI contexts also provide ways for students to enhance science content knowledge, the ability to study the empirical and culturally embedded nature of science, and ability to evaluate different assumptions and claims made during the discussion of SSI (Akerson et al., 2009). However, unlike science inquiry that focuses on the systematic and iterative process to understand the natural world (National Research Council, 2012), SSI cases can go beyond the study of nature extending to areas such as citizenship education (Barrue and Albe, 2013) and character and values (Lee et al., 2013).

The engineering design process and SSI also share common features. Both allow for the exploration and development of knowledge, skills, and technologies to solve problems, taking into account multiple factors such as safety risks, cost analyses, alternative routes, etc. (Guzey et al., 2016). According to the Standards for Technological Literacy, students need to develop understanding of the attributes of design, the engineering design process, and the role of research and development as well as experimentation in problem solving (International Technology Education Association, 2007). The engineering design process includes defining the problem, imagining possible solutions, planning investigations, and creating, testing and redesigning prototypes (Cunningham, 2018). In addition to the engineering design process, technological literacy engages students to develop an understanding of: (a) the cultural, social, economic, and political effects of technology, (b) the effects of technology on history, society and the environment, and (c) the designed world such as medical and biotechnologies, energy and power technologies, and information technologies (ITEA, 2007). Depending on the SSI and how it is incorporated, there can be significant overlap between SSI and engineering design. For instance, learners may compare and contrast alternative energy sources (solar, wind and water) in terms of efficiency and cost, use the engineering design process to create and test model windmills, and examine the impact of cultural practices on energy consumption and carbon footprint (Macalalag et al., 2017). However, while engineering design revolves around systematic and iterative problem solving, it does not always consider the political, moral, ethical, and cultural aspects of the problem, components central to SSI (Zeidler, 2016).

Mathematics and computational thinking play an important role in STEM education. In science, mathematics and computation allow students to collect, analyze and represent data from their investigations, investigate and interpret models from computer simulations, and study patterns and predict behavior of physical phenomena or systems. While in engineering, computer simulations allow students to test, collect data, retest, and redesign their prototypes. Mathematics can also communicate aspects of design features and characteristics (National Research Council, 2012) through real-world connections for mathematics concepts. For

instance, mathematical modeling is a process of interpreting, testing, and revising general real life situations or complex systems (Kertil and Gurel, 2016). According to Gainsburg (2008), real-world connections in teaching mathematics can potentially enhance students' understanding of mathematical concepts, motivate learning of mathematics, and apply concepts to solve real world problems. However, according to her study, most real world connections were brief and many were made by teachers and not by students. SSI can offer a means of addressing this concern by providing an engaging context for applying mathematical concepts to real world scenarios that are culturally relevant to the student. The moral and ethical decision making embedded in SSI requires students, not just the teacher, to connect content to the real world scenarios.

The study of environmental education has been influential in STEM education in recent decades, providing beneficial education on the outdoors, land and wildlife conservation, urban studies, and sustainability, to educate society on how to improve the quality of the environment (Tilbury, 1995) and educate the community about the interdependence of sustainable development and conservation (Gough, 2014). With this focus on modifying our behavior to encourage the use of less damaging alternatives to the environment, environmental education has brought forth the need to bridge the gap between economic development and environmental conservation, and develop a greater understanding of environmental concerns in socioeconomic and political contexts (Tilbury, 1995). Like environmental education, SSI has provided new ways for students to learn about renewable energy sources (e.g., nuclear power) and methods to reduce carbon emission and its effects on the climate (Lee et al., 2012). SSI has also influenced the examination of cultures, politics, and economic factors in the consumption of sea resources and water pollution (Tal and Kedmi, 2006). However, SSI provides focus on the moral and ethical aspects situated within the underlying social and political contexts of environmental issues (Macalalag et al., 2019) such as examining the proponents and consequences of fluoridation of public water supplies, fast food health risks, and stem cell research (Eastwood et al., 2012).

The goals of SSI in STEM education

SSI is firmly situated in the landscape of educational reform and has been shown to have the potential to address the concerns and sometimes apparently disparate goals of STEM education. According to the landmark report *Science for all Americans* (Rutherford and Ahlgren, 1991), STEM courses should "equip [students] ... to participate thoughtfully with fellow citizens" (p. xiii). Presently the *Framework for K-12 science education* (NRC, 2012) holds a prominent place of influence for science teachers while also fostering connections to engineering and environmental topics. That document conveys that learning outcomes from STEM courses should improve the "civic decision making [that] is critical to good decisions about the nation's future" (p. x). It goes on to say that students need to learn about science and technology so they can "engage with the major public policy issues of today" (NRC, 2012, p. 7). However, while there is a clear consensus that STEM education must contribute to the students' ability to deal with scientific issues as they appear in their lives, there is no similar level of agreement as to what kind of classroom experience is best suited to that outcome. SSI researchers have argued that presenting technical knowledge as isolated and divorced from the context of its use is insufficient, given that the goal is to educate people to be able to solve problems that are not entirely technical (e.g., Levine Rose and Calabrese Barton, 2012; Ryder, 2001; Zeidler, 2016).

Feinstein (2011) claimed that the aim should be to prepare "competent outsiders", whom he describes as those who see that "scientific knowledge and knowledge about science are both relevant and accessible—that someone starting on the outside of a problem, without much background, can plunge into the deep water of conflicting expertise and emerge with something resembling an answer" (p. 182). If STEM schooling is geared toward this, it can attend to a longstanding, recognized need for students to recognize what they learn about in school as a relevant and necessary part of navigating their own life and sustaining their cultures (Ladson-Billings, 2014). SSI meets this need by centering on moral and ethical issues, thus rendering the STEM disciplines as an important and relevant contributory means of addressing these issues (Feinstein et al. 2013; Hodson, 2011; Ratcliffe and Grace, 2003). Through addressing moral and ethical issues founded in science-based content, but often left out of traditional STEM education, SSI can offer an authentic immersion into why science matters to all kinds of people in their lives beyond school (Layton et al., 1993). For teachers, this means organizing their curriculum and instruction in ways in which the clarification and resolution of SSI themselves can be the driver of student learning (Sadler, 2009).

A primary reason to pursue SSI is that it can make STEM thoroughly relevant, provided teachers can be responsive and inclusive of their students' interests in and connections to STEM (Basu and Barton, 2007). The potential for SSI to achieve relevance through context is also indicative of its key challenges. As the term "socioscientific" implies, it provides an avenue to marry "the social" with "the science", yet many teachers may not see this as a virtue. STEM teachers tend to identify as authoritative explainers of their discipline's knowledge and instruct accordingly (Levinson, 2004). When teachers do pursue SSI, they are often ill-prepared to implement appropriate forms of instruction (Owens et al., 2019). The main goal of SSI teaching is to achieve what Ryder terms functional scientific literacy (2001), a construct of literacy that is active and seeks to bring the discourse that the public actually engages with science into the classroom. Discourse involves the ways of speaking, thinking, and acting that separate and identify a specific group or network (Gee, 1988) and the scientific discourse in functional scientific literacy must involve the tools needed to thoughtfully negotiate SSI (Sadler et al., 2006). Teachers implementing SSI to foster functional scientific literacy necessarily must model the desired discourse practices (Johnson et al., 2020b). Thus key questions remain that will determine if SSI reforms are successfully implemented: Do teachers have the motivation and disposition to incorporate the extra-scientific components within a STEM content course? Are teachers clear about the pedagogical and content knowledge required for SSI implementation, even as they seem to agree with its basic premise? Will they find their school context to be supportive of enacting SSI? The remainder of this article will

highlight the challenges to reform found in the literature on SSI. Following these descriptions, we will present potential avenues to foster and support SSI implementation in STEM through teacher education and professional development.

Teachers' sense of purpose in teaching STEM

The success of any classroom reform effort is largely dependent on the commitments of the teacher meant to enact it (Haney et al., 1996). Sadler et al. (2006) explored whether and how teachers incorporated ethics, a core element of SSI, into their science teaching. Their findings indicated that many did not because they felt that science classrooms should be value-free. Levinson (2004), in a large survey of British teachers across scientific and non-scientific disciplines, came to essentially the same conclusion. These findings speak to the typical sense of purpose that STEM teachers have regarding their work: successful preparation in STEM evidenced by students' greater acquisition of content knowledge. However from an SSI perspective, this notion of a value-free STEM classroom that does not intentionally make connections to students' cultures is problematic for two reasons.

First, representing the STEM disciplines either implicitly or explicitly as acultural and value-free is a misrepresentation of those fields. Science teachers and science education reformers often hold a misconceived notion of scientific objectivity as a product of its removal and protection from the influences of culture and values (Johnson, 2011). Aikenhead (1996) stated that, "Science curriculum, more often than not, provides students with a stereotype image of science: socially sterile, authoritarian, non-humanistic, positivistic, and absolute truth ... To suggest that school science can present science as culture-free is simply to characterize the subculture of science by its stereotype features" (p. 10). This version of science does a disservice to students who need to engage with science in life by leaving them unprepared to deal with complex issues that involve both scientific and non-scientific epistemologies (Aikenhead, 1985). Chinn and Malhotra (2002) wrote, "When students learn an oversimplified, algorithmic form of scientific reasoning in school, they are likely to reject scientific reasoning as irrelevant to any real-world decision making" (pp. 213–214). Across STEM fields, students may be needlessly marginalized by instruction that portrays ways of knowing as above the influence of context and free from value, putting it at odds with their personal cultural self-identity (Malcom and Feder, 2016). Textbook STEM education, free of discourse and culture, gives students a false impression of actual STEM practice.

In contrast to the culture-free myth, ethnographies examining practices in STEM fields, observed in the context of labs, field work, and conference activities, reveal a strong culture with specific language, norms, and practices held by an elite class of knowledge producers (e.g., Woolgar and Latour, 1986; Traweek, 1990). Uncovering and exploring the cultural dimensions of STEM is important for teachers and students because they foster students' connections to STEM and how they view themselves as learners in those fields. This in turn affects their efforts and academic achievement (Malcom and Feder, 2016). STEM teachers, in an attempt to transmit an objective, culture-free positivist notion of science actually promote a culture foreign to the students that is socially and intellectually distant from the students it is intended to help (Brickhouse, 1990; Lederman, 1992). Barrett and Nieswandt (2010) found that pre-service teachers self-identified strongly within their respective disciplinary group. As such, teachers in this group demonstrated resistance to take up "non-scientific" issues, a characteristic that did not alter throughout their teacher preparation program. It is evident that when faced with teaching SSI, new teachers may struggle to reconcile the disparate views of traditional classroom STEM learning and the SSI framework, which may result in open and often persistent resistance to SSI integration.

The second way that a presumption of appropriateness of value-free science is a problem is that it runs counter to the very essence of SSI as issues-based education. The goal of SSI teaching is to prepare students to deal thoughtfully with scientific issues as they exist in context. Furthermore, SSI are defined as issues for which science alone cannot impose resolution; ultimately only a value judgment can bring closure. Thus SSI classrooms are ideally places where the limits of scientific knowledge are intentionally examined (a goal explicitly supported in the Next Generation Science Standards). These classrooms are learning spaces where boundaries between disciplines are meant to be crossed and explored rather than avoided. Gieryn (1999), in presenting his seminal case studies on the "cultural boundaries" that demarcate science from non-science, wrote that

[p]utatively factual explanations or predictions about nature do not move naked from lab or scientific journal into courtrooms, boardrooms, newsrooms, or living rooms. Rather, they are clothed in sometimes elaborate representation of science—compelling arguments for why science is uniquely best as a provider of trustworthy knowledge, and compelling narrations for why my science (but not theirs) is bona fide" (p. 4).

In other words, any such borders are cultural accomplishments. For the lay person needing to engage with science in order to solve a problem that resides in large part *outside* of science proper, it makes little sense to exclude instances of "elaborate representation" from the curriculum. Effective SSI teachers do not presume that students will learn how to put "ready made" science (Latour, 1987) back into the contexts that have been stripped away by an education that has adhered to a firmly bordered science curriculum.

Another issue stems from resistance to ideological change or resistance to teaching for diversity. This is defined as the resistance to changing one's beliefs and value system (Rodriguez, 2005). Prospective teachers who show resistance to ideological change might believe that all students need only to work hard enough to be successful in science or mathematics or in life in general. They may also believe that students' gender, ethnic backgrounds, socioeconomic status, and language abilities, do not matter with regard to academic success. This "boot straps" mentality can present a major stumbling block in attempting to prepare STEM teachers to teach for diversity using more inclusive or culturally relevant strategies, like integration of the SSI framework, because it undermines the

need for teachers to be culturally responsive. Teachers may demonstrate this resistance not because they disagree with this orientation, but instead because they may lack the awareness, support, confidence, or knowledge and skills to implement socially relevant curriculum and culturally responsive strategies (Rodriguez, 2005).

Teachers' conceptions of SSI

According to Nuangchalerms (2009), it is important that pre-service teachers develop their own conception of SSI-based pedagogy to help their students understand the criteria of real science. Sadler et al. (2006) indicated that, while a great deal of diversity exists among teacher perspectives regarding ethics in science and science teaching, shared views exist on a variety of topics. Some educators have argued for the explicit consideration of the social contexts of the nature of science, even if such consideration reduces focus on some content, to provide an opportunity to "view science as a social enterprise embedded in a community of scientists" (Zeidler et al., 2002, p. 361).

Implementing SSI requires teachers to consider many different factors including how social values influence content, the epistemic values of reliable and productive scientific practice, and ethical questions raised as a result of scientific progress. The acknowledgment and understanding of how science and ethics are interdependent necessarily influences how teachers conceptualize SSI in their classrooms (Sadler et al., 2006). Specifically, research has shown that teachers implementing SSI recognize that "modern, democratic societies demand an informed populace capable of making decisions on controversial, value-laden issues" (Sadler et al., 2006, p. 362) and "SSI serve as effective means of relating science to students' lives" (p. 362). SSI are recognized for their potential to promote scientific literacy and to create an authentic, humane image of the processes of science, important tools for an informed society regarding critical decision-making regarding socioscientific issues (Ries and Galvão, 2009). According to Sadler et al. (2006) teachers tend to believe that the primary goal of integrating SSI and ethics into science was to promote student consideration of their own values and thus they should not impose their own values on students. A majority of the teachers participating in their study stated the belief that both sides of an issue were necessary when dealing with controversial topics, and that they avoided sharing personal beliefs for a variety of reasons (i.e., to avoid being perceived as trying to advance a personal agenda or to avoid legal ramifications). Teachers' personal knowledge of SSI being incorporated into the classroom enabled them to provide foundational support of social issues being developed and debated and clarity on the ethical values attached to SSI. Introducing SSI teaching to the STEM classroom can be difficult as it requires that teachers advance their own scientific literacy to better understand the issues that they incorporate to provide a well-rounded teaching environment (Nuangchalerms, 2009).

Institutionally speaking, many factors affect what teachers choose to do, including teachers' conceptions regarding accepted curriculum being the driving force regarding the concepts taught and means of assessment (Yerrick et al., 1997). Teachers may not recognize how SSI align with accepted curriculum and so constrain their choices to more traditional pedagogies, avoiding depth of coverage for breadth in topics covered. Teachers may avoid discussing controversial issues out of fear of complaints from students' parents or the sense of lack of control during classroom discussions (Ries and Galvão, 2009). This latter constraint reflects a lack of management skills related to classroom discussions and a lack of knowledge on the SSI. For teachers to be proficient at incorporating controversial SSI topics to address the nature of science and to improve reasoning skills, the teachers must be knowledgeable and experienced in this area (Zeidler et al., 2002). Other constraints to SSI implementation include time limitations, lack of administrative support, and pressures of standardized testing (Sadler et al., 2006; Yerrick et al., 1997), yet teachers who successfully incorporate SSI treat these as minor inconveniences rather than serious impediments. Others are more resistant to classroom implementation of SSI. Such resistance stems from numerous sources and is discussed in detail in upcoming sections.

Building new curriculum and adopting new repertoires of instruction

It is important that, when considering the purposes of STEM education, the goals, form, and function are aligned. Rudolph and Horibe (2016) sought to make explicit what is meant by teaching science "for civic engagement" and found that there are two basic ways in which the public interacts with science: how science is used to form policy and to a lesser extent how the conduct of science itself is to be supported, both through allocation of public resources and its accountability through regulation. They further pointed out that any notion of how science education should be done must reference the actual "avenues of engagement" that people actually have access to. To that point they reference Ryder's (2001) notion of functional scientific literacy. Ryder analyzed case studies of non-scientists engaging with institutional science, either directly with scientists or with representations of science such as media reports. He found that what really constitutes "functional literacy" is an ability to make sense of how science is represented in mass communication, an understanding of the epistemology of science (Ryder, 2001), and to a lesser but non-trivial extent the sociology of science itself or the notion that scientists work in a community of colleagues where they argue with each other toward establishing stable claims. It is important to note what he finds is *not* so crucial: an encyclopedic knowledge of scientific facts. Layton et al. (1993) performed a study with similar results. Lemke (1998) contended that learning about natural and technological phenomena includes engaging students in the kinds of patterns of action that scientists use in their work in meaningful ways. Duschl et al. (2007) identified the scientist's central activity as "argument in communities of practice for the purpose of persuading colleagues of the validity of one's own ideas and the ideas of others," (p. 187).

Researchers have argued that if scientific literacy is indeed a central and important goal of STEM education, “then scientific literacy must entail, at least in part, the ability to thoughtfully negotiate SSI and contribute to discourse regarding these topics” (Sadler et al., 2006, p. 354). For students, aspects of science discourse include the ways they construct scientific knowledge and make meaning from observation and participation in scientific processes (Johnson, 2011). However the most effective means of doing this are still being explored. According to Zeidler et al. (2005), the framework for teaching SSI includes four important pedagogical areas; (1) nature of science issues, (2) classroom discourse issues, (3) cultural issues, and (4) case-based issues (p. 361), that come together to inform teaching and promote the development of functional scientific literacy among students. While this framework points to four dimensions of curriculum-building work and classroom instruction that SSI teachers must attend to, they in turn can also be seen through the lens of ideological commitment. It is important to be specific about what SSI teaching entails when questioning whether teachers see its value because while “science literacy for all” is an easy mantra to profess; this consensus that “literacy” should be pursued tends to break down once we consider the kinds of pedagogy that would entail. Genuine scientific argument is rarely seen in classrooms. Instead, tacit turn-taking is often the primary means of communication in the classroom. Several factors influence the classroom discourse and limit teachers’ use of scientific argument. Many teachers are uncomfortable with argument and both teachers and students are often unable to distinguish between scientific argument and everyday argument (Duschl et al., 2007). Teachers may find difficulty in connecting claims to evidence and reasoning in discussing SSI cases. This has major implications in that teachers implementing SSI to foster functional scientific literacy must model the desired discourse practices (Johnson et al., 2020b). Nevertheless, participation is important. When curriculum presents students with challenging, socially relevant new knowledge, it provides the opportunity to become active participants in the community and engage in meaningful dialog (Rodriguez and Berryman, 2002). Fortunately, directed coursework and professional development can provide an avenue for perspective teachers with the resources and experience to develop the necessary skills for effective SSI implementation (Macalalag et al., 2019).

Providing teachers with firsthand experience in culturally relevant cases and familiar content allows for deep engagement in argumentation. This can build a foundation for application to teachers’ future instruction, allowing them to better align selected cases to student interests and backgrounds (Johnson et al., 2020b). One way that the SSI movement differs from prior efforts to infuse real-world applications into curricula and instruction is the “explicit attention to the ethical aspects of social issues with conceptual, methodological, and/or technological ties to science (Sadler et al., 2006, p. 354). SSI embeds STEM content into widely recognizable social and political contexts. The context is central in that when the ethical aspects of selected social issues align with students’ background and interest, the content is never irrelevant in the first place. Content is positioned as a necessary part of understanding an issue rather than an end in and of itself. This empowers learners to access scientific content in a completely different way and makes the primary goal of learning to make sense of or solve an authentic, personal problem that is directly faced by students in their own communities. The value of SSI is that it can teach not just what science is, but also how science is done, why science matters, and how science impacts everyone.

Hospitality of school contexts

The decision to implement SSI as an approach to STEM instruction does not happen in isolated spaces. Teachers must navigate numerous factors within their school context that both support and discourage teachers’ use of SSI in the classroom. These factors impact whether or not a teacher uses SSI as a context for instruction, and if so, influences the cases which teachers select to include in their curriculum (Hancock et al., 2019). Factors such as climate, schedules, high stakes testing, mandated curriculum standards and various school stakeholders’ philosophical ideals about the purpose of education all converge to impact teachers’ use of SSI at various levels including the classroom, school, and community.

One major factor cited by teachers as a barrier to implementing SSI is their perception that students will develop misconceptions of STEM content, thus negatively impacting learning outcomes (Aivelo and Uitto, 2019). Often because of this fear, teachers tend to shy away from SSI altogether or relegate their use of SSI to the end of a unit or article if time allows (Aivel and Uitto, 2019; Wong et al., 2011). This notion that “content” must be delivered is hard, or maybe impossible, to get away from when courses are tied to a high stakes test. Testing can become a bludgeon that enforces that the “delivery” task must be accomplished before SSI can happen.

Participants in a study conducted by Hancock et al. (2019) cited a focus on high stakes testing within their school as a barrier to their implementation of SSI even after a workshop guiding them to develop standards aligned lessons. Multiple studies with SSI case-based instruction have found that students tend to make progress on standards-based and standardized science assessments through engagement in SSI (Klosterman and Sadler, 2010). Furthermore, research supports the use of SSI cases to improve students’ abilities in evidence based arguments and critical thinking skills, as well as their understanding of both science concepts and the nature of science (Pinzino, 2012).

Directed course work and professional development can influence teachers intentions to implement SSI (Macalalag et al., 2019), but even when teachers do intend to use SSI cases in their classrooms they encounter challenges such as the need to use common curriculum materials and assessments, and the need to follow the same curriculum map and pacing guide as other teachers with whom they work. Hancock et al. (2019) found, however, that when working as part of a professional learning community team, these issues could be navigated and teachers on the team were able to find ways to include SSI within the confines their school context presented.

Firsthand experience with SSI cases is also an important factor in teachers' implementation of SSI. Through engaging directly in SSI, teachers gain knowledge and experience in the process of SSI implementation, yet they will engage most deeply in issues that they are interested in or have experience with (Johnson et al., 2020b). In a study conducted by McGinnis and Simmons (1999) inservice teachers engaged in a variety of SSI topics themselves before selecting and planning how they might implement them in their own classrooms. Findings revealed that teachers identified non-taboo issues as those that they felt comfortable teaching and cited their job security being threatened as a reason to stay away from more personally relevant cases, such as clean water in their building. Teachers also identified themselves as outsiders of the local school community, thus decreasing their willingness to teach topics they view as controversial to the community, despite seeing the value in such topics. Teachers often steered away from teaching that would result in some action being taken by students in response to controversial topics (McGinnis and Simmons, 1999).

Yet scholars such as McAvoy and Hess (2013) argue that avoidance of politics constitutes a missed opportunity to learn the necessary civic skills of deliberation on polarizing topics. Therefore controversial socioscientific issues are precisely the cases that teachers should include in their courses. Schools are important venue for learning to make decisions across different perspectives (Parker, 2003) because they can bring together diverse stakeholders. STEM classrooms, rather than being sites of uniform purpose meant only to prepare future STEM professionals, are likely to be some of the only places in which students can be exposed to more diverse beliefs and views than their immediate social circles (McAvoy and Hess, 2013). In light of this factor, and the likelihood that students are still developing their own political identity and therefore less likely to be politically engrained in ideas, the classroom is the ideal place for students to practice discussing and deliberating on controversial issues. On the contrary, teachers are more likely to be resistant to engage their students in such topics especially in politically polarized times (McAvoy and Hess, 2013).

Developing teachers' pedagogical content knowledge in SSI

Despite the challenges outlined in the preceding sections, SSI implementation can be fostered and maintained. Directed course work and professional development can develop teacher knowledge necessary for effective SSI implementation. STEM teachers who teach using SSI should include the following components in their lessons: (a) identify the issue, (b) explore and explain the underlying scientific phenomena, (c) engage in scientific modeling, (d) consider system dynamics, (e) employ reflective skepticism, (f) compare and contrast multiple perspectives, and (g) elucidate their own position or solution (Sadler et al., 2019). The pedagogical content knowledge (PCK) framework can provide a lens for understanding the requirements for the development of teachers' in planning, teaching, and reflecting on SSI (Bayram-Jacobs et al., 2018) as contexts for learning STEM concepts like arguments on global warming (Albe and Gombert, 2012), scientific literacy (Tal and Kedmi, 2006), and citizenship education (Barrue and Albe, 2013). According to Shulman (1986), PCK is the knowledge teachers must have within a discipline that is essential for them to effectively plan and implement teaching methods in order to help learners of various levels and backgrounds learn concepts and skills during instruction. PCK in general, and PCK for teaching SSI, includes several subdomains. Knowledge about the curriculum includes a teacher's awareness of the curriculum goals, objectives, and the vertical alignment and progressions of students' learning (Magnusson et al., 1999; Bayram-Jacobs, 2019). A teacher's knowledge about curriculum may come from state and national standards documents and curricular programs being used (Magnusson et al., 1999). Bayram-Jacobs et al. (2019) included the teacher's thinking about the importance for students and the learning objectives within lessons in this part of a teacher's PCK for SSI. In a study with inservice teachers, teacher PCK for SSI curriculum was demonstrated through the selection of controversial case-based issues (Macalalag et al., 2019). Teachers cited raising environmental awareness and social responsibility as their purpose in using the SSI cases which they selected.

In addition to knowledge of the curriculum, a teacher's PCK for SSI must also include knowledge of students' understanding. This includes an awareness of students' background knowledge and how students generally understand specific science concepts, the common misconceptions and preconceptions they may hold, and strategies to help students' learning move forward (Magnusson et al., 1999). Learning about students' backgrounds is especially important in regards to teachers' selection of SSI cases, so that teachers can align SSI cases to students' lives and make them personally relevant (Saunders and Rennie, 2013; Yerrick and Johnson, 2011). The moral and ethical reasoning involved in SSI teaching is what sets it apart from other orientations to teaching science (Zeidler, 2016). Saunders and Rennie (2013) suggest the use of an ethical framework to help support students' thinking throughout this process. Tal and Kedmi (2006) suggested continuous feedback between the teacher and students was important in SSI cases. Bayram-Jacobs et al. (2019) found that teachers with strong PCK tend to have a strong connection between knowledge of students' understanding and knowledge of instructional strategies.

In terms of knowledge of assessments, the teacher of SSI must know what, when and how to use suitable assessments in order to support learners during their exploration of SSI. For instance, in a study describing an Earth science SSI unit on beach sand replacement (Dolan et al., 2009), assessment was based on how students were able to engage in argumentation in order to provide both their scientific knowledge and personal beliefs to protect and restore the beach. Teachers may choose to assess individual or small groups of students' arguments using criteria such as the number of justifications, evidence provided, and counterclaim and rebuttals (Tal and Kedmi, 2006). Teachers must develop and apply their PCK on assessments in order to support students' learning of SSI. Teachers must be able to create a learning environment and assessment strategies in order to monitor and to support students' knowledge development in STEM, their ability to work with and to effectively engage in argumentation with other students, and

their ability to apply concepts to solve real world problems. This learning environment may include a continuous cycle of thinking and discussing, feedback, and reflection (Tal and Kedmi, 2006).

Finally, a teacher's knowledge of instructional strategies includes a teacher's ability to make appropriate choices about teaching and learning strategies in incorporating SSI (Magnusson et al., 1999). This is strongly connected to the teacher's knowledge of student understanding and assessment. A teacher of SSI must be able to incorporate a variety of teaching strategies that allow students to explore and explain the underlying scientific phenomena, engage in scientific modeling, employ reflective skepticism, compare and contrast multiple perspectives, and elucidate their own position or solution, which are several essential components of SSI (Sadler et al., 2019). As the teacher assesses students' evidence based arguments, they gain a sense of students' understanding and difficulties and use this to plan appropriate strategies to move students forward. The use of case-based strategies can provide a context for development of informal reasoning, such as Topcu et al.'s (2011) examples with human cloning, gene therapy and global warming. Minken et al. (2020) studied inservice teachers' use of multiple perspectives as an instructional strategy and found that although this is an important strategy for argumentation in SSI, teachers infrequently used it in their lesson plans.

Studies of course work and professional development models have suggested ways to educate and prepare both preservice and inservice teachers to implement SSI (Macalalag et al., 2019; Zeidler, 2016; Yerrick et al., 1997; Rodriguez, 2005). In several of these studies, teachers acted as learners to engage in scientific practices, such as evidence-based reasoning, scientific skepticism, and exploring connections between science and culture, through participation in selected case-based issues. Through coursework and professional development opportunities, teachers can engage in scientific practices themselves. In order to teach SSI, experiences in developing evidence-based reasoning are critical. Often, when teachers engage in evidence based reasoning, they cite vague evidence and lack a connection between their claim and the evidence suggested (Macalalag et al., 2019). Experiences in developing arguments centered on SSI cases can be a means to develop teachers' reasoning (Macalalag et al., 2019). Johnson et al. (2019) also found that teachers lack cohesion between claims and evidence in scientific arguments and suggest that this may be due to lack of experience in this type of communicating. As Michaels and O'Connor (1990) point out, if strategies and norms for participating in science discourse like argumentation are not actively pursued and promoted in traditional classrooms, they rarely develop on their own. Teachers who have never experienced this type of argumentation themselves will be unfamiliar and uncomfortable implementing it in their classrooms. Yet, course work and professional development can help. Coursework can provide avenues for teachers to evaluate the sources of evidence and information. In the works of Macalalag et al. (2019) and Kolsto (2006), teachers were skeptical of sources provided as they searched for evidence and examined the motivations and possible biases of the authors as part of their discussions of SSI. Opportunities for evidence based reasoning with scaffolded support are necessary for teachers to develop these abilities (Johnson et al., 2019). A clear understanding of teachers' abilities in these areas can also assist in the development of targeted experiences for professional development and coursework (Johnson et al., 2019).

As we have illustrated, resistance to change in teaching practice may stem from many sources. SSI integration can provide a powerful tool for teaching in culturally relevant ways as it allows teachers to gain a deeper understanding of students' conceptual frameworks during discussions of socioscientific issues (Zeidler et al., 2002). However, prospective teachers may purposely refuse to teach for diversity and understanding for a variety of reasons (Rodriguez, 2005). School policies, preservice training, strict accountability, rigorous state assessments, teacher socialization, and mandated curriculum and assessment may influence teachers so much that they resist thinking about content and teaching in any other way (Yerrick et al., 1997). This type of resistance to pedagogical change or resistance to learning to teach for understanding results from teachers' avoiding, refusing, or feeling unable to change their perceptions of what it means to be an effective teacher (Rodriguez, 2005), and is an ongoing challenge in preparing STEM educators. Educators who demonstrate this form of resistance may prefer lecture and individual student work, providing the students the content first as a de facto prerequisite before they can be allowed to do any kind of application or experiment. These teachers may view science teaching as the transmission of lists of concepts and students as passive recipients (Yerrick et al., 1997). There is also the possibility that teachers may show resistance to pedagogical change because they lack the knowledge, skills, and confidence to move away from transmissive models of teaching which they themselves have experienced throughout their education (Lortie, 1975). As Yerrick et al. (1997) point out, the ubiquitous adoption of transmissional modes of teaching is especially troubling since it conflicts so strongly with methods shown to be effective for teaching for the ever increasing diversity within schools.

Numerous studies have shown that, while the forms of resistance described above are all too common, they can be overcome in preparing teachers to teach for diversity in inclusive and culturally responsive ways (Rodriguez and Kitchen, 2005). In particular, SSI provides a strong framework for engaging students and teachers in meaningful and relevant scientific discourse in the development of functional scientific literacy (Zeidler et al., 2005) and although it can be difficult for teachers to engage in arguments around SSI, coursework and professional development can provide a means for them to do so (Macalalag et al., 2019; Nuangchalerms, 2009). Varelas et al. (2018) examined an intervention intended to raise awareness among pre-service teachers regarding environmental racism. Their work showed that pre-service teachers began to see their role as more than just a science teacher as they expanded their views of the ways science is relevant in peoples' lives. Macalalag et al. (2019) indicated positive views among in-service and pre-service teachers regarding teaching controversial and complex case-based SSI lessons following completion of a STEM certification endorsement program founded in the SSI framework. Participating teachers indicated raising environmental awareness, teaching social responsibility, and teachers' interest on SSI topics as reasons for their views. This study attributed the development of the cultural issues component of SSI to teachers' field experiences in Pennsylvania, USA and in Sicily, Italy embedded within the STEM endorsement program, which included shifts on teachers' perceived cultural practices, reducing carbon dioxide, and teaching of SSI in their classrooms. In another study by Macalalag et al. (2017), following the methods course within the STEM endorsement program, more teachers identified connections between science and engineering and demonstrated more detailed and specific

understandings of STEM concepts and processes based on their experiences with SSI. Their experiences and discussion of SSI within the course and field experiences provided them a template for integrating SSI into their own teaching. Following the course their focus shifted to actual implementation as opposed to simply identifying the need.

Conclusions

Incorporating SSI into STEM classrooms provides a way to teach all students about STEM content because it situates STEM content in contexts that are meaningful to them, however this asks STEM teachers to teach in ways that they are often ill-prepared for, unsupported in, and/or personally resistant to on ideological grounds. STEM teacher education and professional development has been shown to be effective and necessary in fostering SSI implementation, particularly at sites where STEM teachers are supported to work through the problems they are likely to face when enacting SSI teaching.

A first necessary step toward reform is to raise awareness of the ways that STEM impacts people. Varelas et al. (2018) examined an intervention intended to develop a consciousness among pre-service teachers about environmental racism. Their work showed that pre-service teachers began to see their role as more than just a science teacher as they expanded their views on the ways science is relevant in peoples' lives, sometimes in adverse ways.

To facilitate and maintain integration of SSI among pre-service and in-service teachers, teacher educators must provide a variety of experiences in which developing teachers directly participate in SSI from the students' point of view. In doing so, teachers will develop an extensive repertoire of SSI strategies and contexts, while also learning how to learn, develop, and implement their own SSI lessons focusing on contexts specific to their classroom (Johnson et al., 2020a). Experience with multiple SSI cases within a teacher preparation program will better prepare teachers to provide culturally relevant and engaging opportunities for their students to build understanding. Culturally relevant cases with familiar content allow for deep engagement in argumentation and application to teachers' future instruction, allowing them to better align selected cases to student interests and backgrounds.

References

- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N.G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., Tuan, H.L., 2004. Inquiry in science education: international perspectives. *Sci. Educ.* 88 (3), 397–419.
- Aikenhead, G.S., 1985. Collective decision making in the social context of science. *Sci. Educ.* 69 (4), 453–475.
- Aikenhead, G.S., 1996. Science education: border crossing into the subculture of science. *Stud. Sci. Educ.* 27, 1–52.
- Avelo, T., Uitto, A., 2019. Teachers' choice of content and consideration of controversial and sensitive issues in teaching of secondary school genetics. *Int. J. Sci. Educ.* 41 (18), 2716–2735.
- Akerson, V.L., Cullen, T.A., Hanson, D.L., 2009. Fostering a community of practice through a professional development program to improve elementary teachers' views of nature of science and teaching practice. *J. Res. Sci. Teach.* 46 (10), 1090–1113.
- Albe, V., Gombert, M.J., 2012. Students' communication, argumentation and knowledge in a citizens' conference on global warming. *Cult. Stud. Sci. Educ.* 7 (3), 683–691.
- Barrett, S.E., Nieswandt, M., 2010. Teaching about ethics through socioscientific issues in physics and chemistry: teacher candidates' beliefs. *J. Res. Sci. Teach.* 47 (4), 380–401.
- Barrie, C., Albe, V., 2013. Citizenship education and socioscientific issues: implicit concept of citizenship in the curriculum, views of French middle school teachers. *Sci. Educ.* 22 (5), 1089–1114.
- Basu, S.J., Barton, A.C., 2007. Developing a sustained interest in science among urban minority youth. *J. Res. Sci. Teach.* 44 (3), 466–489.
- Bayram-Jacobs, D., Henze, I., Barendsen, E., 2018. The influence of ENGAGE materials on students' learning about socio-scientific issues. In: Finlayson, O. (Ed.), *Proceedings of the 12th Conference of the European Science Education Research Association. ESERA*, Dublin, Ireland.
- Bayram-Jacobs, D., Henze, I., Evagorou, M., Schwartz, Y., Aschim, E.L., Alcaraz-Dominguez, S., Barajas, M., Dagan, E., 2019. Science teachers' pedagogical content knowledge development during enactment of socioscientific curriculum materials. *J. Res. Sci. Teach.* 56 (9), 1207–1233.
- Brickhouse, N.W., 1990. Teachers' beliefs about the nature of science and their relationship to classroom practice. *J. Teach. Educ.* 41 (3), 53–62.
- Chinn, C.A., Malhotra, B.A., 2002. Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Sci. Educ.* 86 (2), 175–218.
- Cunningham, C., 2018. *Engineering in Elementary STEM Education: Curriculum, Design, Instruction, Learning, and Assessment*. Teachers College Press, New York, New York.
- Dolan, T.J., Nicholas, B.H., Zeidler, D.L., 2009. Using socioscientific issues in primary classrooms. *J. Elem. Sci. Educ.* 21 (3), 1–12.
- Duschl, R.A., Schweingruber, H.A., Shouse, A.W. (Eds.), 2007. *Taking Science to School: Learning and Teaching Science in Grades K-8*, vol. 500. National Academies Press, Washington, DC.
- Eastwood, J.L., Sadler, T.D., Zeidler, D.L., Lewis, A., Amiri, L., Applebaum, S., 2012. Contextualizing nature of science instruction in socioscientific issues. *Int. J. Sci. Educ.* 34 (15), 2289–2315.
- Feinstein, N.W., Allen, S., Jenkins, E., 2013. Outside the pipeline: reimagining science education for nonscientists. *Science* 340 (6130), 314–317.
- Feinstein, N., 2011. Salvaging science literacy. *Sci. Educ.* 95 (1), 168–185.
- Flick, L.B., Lederman, N.G., 2006. *Scientific Inquiry and Nature of Science*. Springer, Netherlands.
- Gainsburg, J., 2008. Real-world connections in secondary mathematics teaching. *J. Math. Teach. Educ.* 11, 199–219.
- Gee, J.P., 1988. The legacies of literacy: from Plato to Freire through Harvey Graff. *Harv. Educ. Rev.* 58 (2), 195–213.
- Gieryn, T.F., 1999. *Cultural Boundaries of Science: Credibility on the Line*. University of Chicago Press, Chicago, IL.
- Gough, A., 2014. A long, winding (and rocky) road to environmental education for sustainability in 2006. *Aust. J. Environ. Educ.* 30 (1), 96–101.
- Guzey, S.S., Moore, T.J., Harwell, M., 2016. Building up STEM: an analysis of teacher-developed engineering design-based STEM integration curricular materials. *J. Pre-Coll. Eng. Educ. Res.* 6 (1), 2.
- Hancock, T.S., Friedrichsen, P.J., Kinslow, A.T., Sadler, T.D., 2019. Selecting socio-scientific issues for teaching: a grounded theory study of how science teachers collaboratively design ssi-based curricula. *Sci. Educ.* 28, 639–667.
- Haney, J.J., Czerniak, C.M., Lumpe, A.T., 1996. Teacher beliefs and intentions regarding the implementation of science education reform strands. *J. Res. Sci. Teach.* 33 (9), 971–993.
- Hodson, D., 2011. *Looking to the Future: Building a Curriculum for Social Activism*. Sense, Rotterdam, Netherlands.
- International Technology Education Association, 2007. *Standards for Technological Literacy: Content for the Study of Technology*. ITEA Press, Reston, Virginia.

- Johnson, J., Dunphy, J., Macalalag, A., 2019. Teacher understandings of claims, evidence, and reasoning within socioscientific issue implementation. In: Proceedings of the American Educational Research Association Annual Meeting. Canada, Toronto.
- Johnson, J., Macalalag, A., Dunphy, J., 2020. Encouraging teacher use of SSI: developing scientific argumentation and classroom implementation. In: Proceedings of the Conference for Academic Research in Education. UNLV, Las Vegas, NV.
- Johnson, J., Macalalag, A.Z., Dunphy, J., 2020. Incorporating socioscientific issues into a STEM education course: exploring teacher use of argumentation in SSI and plans for classroom implementation. *Discip. Interdiscip. Sci. Educ. Res.* 2 (1), 1–12.
- Johnson, J.A., 2011. Science is for Me: Meeting the Needs of English Language Learners in an Urban, Middle School Science Classroom through an Instructional Intervention. State University of New York at Buffalo, Buffalo, NY.
- Kertl, M., Gurel, C., 2016. Mathematical modeling: a bridge to STEM education. *Int. J. Educ. Math. Sci. Technol.* 4 (1), 44–55.
- Klosterman, M.L., Sadler, T.D., 2010. Multi-level assessment of scientific content knowledge gains associated with socioscientific issues-based instruction. *Int. J. Sci. Educ.* 32 (8), 1017–1043.
- Kolstø, S.D., 2006. Patterns in students' argumentation confronted with a risk-focused socio-scientific issue. *Int. J. Sci. Educ.* 28 (14), 1689–1716.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: aka the remix. *Harv. Educ. Rev.* 84 (1), 74–84.
- Latour, B., 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Harvard university press, Cambridge, MA.
- Layton, D., Jenkins, E., Macgill, S., Davey, A., 1993. Inarticulate science? Perspectives on the public understanding of science and some implications for science education. *Studies in Education*, Driffield, East Yorkshire.
- Lederman, N.G., 1992. Students' and teachers' conceptions of the nature of science: a review of the research. *J. Res. Sci. Teach.* 29 (4), 331–359.
- Lee, H., Chang, H., Choi, K., Kim, S.W., Zeidler, D.L., 2012. Developing character and values for global citizens: analysis of pre-service science teachers' moral reasoning on socioscientific issues. *Int. J. Sci. Educ.* 34 (6), 925–953.
- Lee, H., Yoo, J., Choi, K., Kim, S.W., Krajcik, J., Herman, B.C., Zeidler, D.L., 2013. Socioscientific issues as a vehicle for promoting character and values for global citizens. *Int. J. Sci. Educ.* 35 (12), 2079–2113.
- Lemke, J.L., 1998. Teaching all the languages of science: words, symbols, images, and actions. In: Proceedings of the Conference on Science Education in Barcelona. Barcelona, Spain.
- Levine Rose, S., Calabrese Barton, A., 2012. Should great lakes city build a new power plant? How youth navigate socioscientific issues. *J. Res. Sci. Teach.* 49 (5), 541–567.
- Levinson, R., 2004. Teaching bioethics in science: crossing a bridge too far? *Can. J. Sci. Math. Technol. Educ.* 4 (3), 353–369.
- Lortie, D., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, London, UK.
- Macalalag Jr., A.Z., Johnson, J., Lai, M., 2017. STEM education course: enhancing K-12 teachers' cultural awareness through reflections of socioscientific issues. *Pennsyl. Teacher Educ.* 16, 76–92.
- Macalalag, A.Z., Johnson, J., Lai, M., 2019. How do we do this: learning how to teach socioscientific issues. *Cult. Stud. Sci. Educ.* 3 (8), 1–25.
- Magnusson, S., Krajcik, J., Borko, H., 1999. Nature, sources, and development of pedagogical content knowledge for science teaching. In: Gess-Newsome, J., Lederman, N.G. (Eds.), *Examining Pedagogical Content Knowledge*. Kluwer, The Netherlands.
- Malcom, S., Feder, M., 2016. The culture of undergraduate STEM education. In: *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. National Academies Press (US).
- McAvoy, P., Hess, D., 2013. Classroom deliberation in an era of political polarization. *Curric. Inq.* 43 (1), 14–47.
- McGinnis, J.R., Simmons, P., 1999. Teachers' perspectives of teaching science-technology-society in local cultures: a sociocultural analysis. *Sci. Educ.* 83, 179–211.
- Michaels, S., O'Connor, M.C., 1990. Literacy as reasoning within multiple discourses: implications for policy and educational reform. In: Proceedings the Meeting of the Council of Chief State School Officers 1990 Summer Institute.
- Minken, M.Z., Macalalag Jr., A.Z., Naylor, N., 2020. What will you do to help elementary students who struggle in the engineering design process? Analysis of teachers' reflections. In: Proceedings of the American Society for Engineering Education Virtual Conference. ASEE.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The National Academies Press, Washington, DC.
- Nuangchalem, P., 2009. Development of socioscientific issues-based teaching for preservice science teachers. *J. Soc. Sci.* 5 (3), 239–243.
- Owens, D.C., Sadler, T.D., Friedrichsen, P., 2019. Teaching practices for enactment of socio-scientific issues instruction: an instrumental case study of an experienced biology teacher. *Res. Sci. Educ.* 1–24.
- Parker, W.C., 2003. *Teaching Democracy: Unity and Diversity in Public Life*. Teachers College Press.
- Pinzino, D.W., 2012. Socioscientific Issues: A Path towards Advanced Scientific Literacy and Improved Conceptual Understanding of Socially Controversial Scientific Theories. Graduate Theses and Dissertations. <http://scholarcommons.usf.edu/etd/4387>.
- Ratcliffe, M., Grace, M., 2003. *Science Education for Citizenship: Teaching Socio-Scientific Issues*. McGraw-Hill Education, UK.
- Reis, P., Galvão, C., 2009. Teaching controversial socio-scientific issues in biology and geology classes: a case study. *Electron. J. Sci. Educ.* 13.
- Rodríguez, A.J., Beryman, C., 2002. Using sociotransformative constructivism to teach for understanding in diverse classrooms: a beginning teacher's journey. *Am. Educ. Res. J.* 39 (4), 1017–1045.
- Rodríguez, A.J., Kitchen, R.S. (Eds.), 2005. *Preparing Mathematics and Science Teachers for Diverse Classrooms: Promising Strategies for Transformative Pedagogy*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Rodríguez, A.J., 2005. Teachers' resistance to ideological and pedagogical change: definitions, theoretical framework, and significance. In: Rodríguez, A.J., Kitchen, R.S. (Eds.), *Preparing Mathematics and Science Teachers for Diverse Classrooms: Promising Strategies for Transformative Pedagogy*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 17–32.
- Rudolph, J.L., Horibe, S., 2016. What do we mean by science education for civic engagement? *J. Res. Sci. Teach.* 53 (6), 805–820.
- Rutherford, F.J., Ahlgren, A., 1991. *Science for All Americans*. Oxford university press.
- Ryder, J., 2001. Identifying science understanding for functional scientific literacy. *Stud. Sci. Educ.* 36, 1.
- Sadler, T.D., Amirshokoobi, A., Kazempour, M., Allspaw, K.M., 2006. Socioscience and ethics in science classrooms: teacher perspectives and strategies. *J. Res. Sci. Teach.* 43 (4), 353–376.
- Sadler, T.D., Friedrichsen, P., Zangori, L., 2019. A Framework for teaching for socio-scientific issue and model based learning (SIMBL). *J. Educação e Fronteiras* 9 (25), 8–26.
- Sadler, T.D., 2009. Situated learning in science education: socio-scientific issues as contexts for practice. *Stud. Sci. Educ.* 45 (1), 1–42.
- Saunders, K.J., Rennie, L.J., 2013. A pedagogical model for ethical inquiry into socioscientific issues in science. *Res. Sci. Educ.* 43 (1), 253–274.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 67, 4–14.
- Tal, T., Kedmi, Y., 2006. Teaching socioscientific issues: classroom culture and students' performances. *Cult. Stud. Sci. Educ.* 1 (4), 615–644.
- Tilbury, D., 1995. Environmental education for sustainability: defining the new focus of environmental education in the 1990s. *Environ. Educ. Res.* 1 (2), 195–212.
- Topçu, M.S., Yılmaz-Tüzün, Ö., Sadler, T.D., 2011. Turkish preservice science teachers' informal reasoning regarding socioscientific issues and the factors influencing their informal reasoning. *J. Sci. Teach. Educ.* 22 (4), 313–332.
- Traweeek, S., 1990. *Beamtimes and Lifetimes: The World of High Energy Physicists*. Harvard University Press, Cambridge, MA.
- Varelas, M., Morales-Doyle, D., Raza, S., Segura, D., Canales, K., Mitchener, C., 2018. Community organizations' programming and the development of community science teachers. *Sci. Educ.* 102 (1), 60–84.
- Wong, S.L., Wan, Z., Cheng, M.M.W., 2011. Learning nature of science through socioscientific issues. In: Sadler, T.D. (Ed.), *Socio-scientific Issues in the Classroom: Teaching, Learning, and Research*. Springer, Dordrecht, pp. 245–269.

- Woolgar, S., Latour, B., 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton, NJ.
- Yerrick, R., Johnson, J., 2011. Negotiating White science in rural Black America: a case for navigating the landscape of teacher knowledge domains. *Cult. Stud. Sci. Educ.* 6 (4), 915.
- Yerrick, R., Parke, H., Nugent, J., 1997. Struggling to promote deeply rooted change: the "filtering effect" of teachers' beliefs on understanding transformational views of teaching science. *Sci. Educ.* 81 (2), 137–159.
- Zeidler, D.L., Walker, K.A., Ackett, W.A., Simmons, M.L., 2002. Tangled up in views: beliefs in the nature of science and responses to socioscientific dilemmas. *Sci. Educ.* 86 (3), 343–367.
- Zeidler, D.L., Sadler, T.D., Simmons, M.L., Howes, E.V., 2005. Beyond STS: a research-based framework for socioscientific issues education. *Sci. Educ.* 89 (3), 357–377.
- Zeidler, D.L., 2016. STEM education: a deficit framework for the twenty first century? A sociocultural socioscientific response. *Cult. Stud. Sci. Educ.* 11, 11–26.

STEM curriculum development and implementation

Gillian H. Roehrig^a, Emily A. Dare^b, Jeanna R. Wieselmann^c, and Elizabeth A. Ring-Whalen^d, ^a Department of Curriculum and Instruction, University of Minnesota, St. Paul, MN, United States; ^b Department of Teaching and Learning, STEM Transformation Institute, Florida International University, Miami, FL, United States; ^c Department of Teaching and Learning, Southern Methodist University, Dallas, TX, United States; and ^d National Center for STEM Elementary Education, St. Catherine University, St. Paul, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	153
Brief history of integrated STEM curriculum development	154
Integrated STEM curriculum development and implementation: the case of <i>EngrTEAMS</i>	154
Need for an integrated STEM curriculum framework	154
Broad approaches to engineering integration	156
STEM integration curriculum assessment	156
Conceptual flow graphics	157
Considerations for selecting an engaging context	159
Implementing integrated STEM curricula in K-12 classrooms	159
Understanding integrated STEM implementation through student discourse	160
Future research in integrated STEM curriculum implementation	160
Acknowledgment	161
References	161

Introduction

The rise of STEM education is highly influenced by political and business leaders related to concerns about shortages in the STEM workforce. For example, the National Governors Association (NGA) (Toulmin and Groome, 2007) argued,

America's economic growth in the 21st century will be driven by our nation's ability to generate ideas and translate them into innovative products and services. A strong consensus is emerging among scientific, business, and education leaders that America's ability to innovate and compete in the global marketplace is directly tied to the ability of our public schools to adequately prepare all of our children in STEM (p. 4).

Arguments related to the STEM workforce have similarly propelled countries around the world to establish STEM policies and initiatives designed to improve the quality of STEM education (Freeman et al., 2014). Other equally important arguments for global STEM initiatives are more equitable education, poverty reduction, and increased STEM literacy and awareness (Freeman et al., 2014; National Research Council [NRC], 2014).

Concurrently, professional societies, such as the American Society for Engineering Education (ASEE) and the National Academy of Engineering (NAE), promoted K-12 engineering and set the stage for a more integrated approach to STEM education. In *Engineering in K-12 Education: Understanding the Status and Improving the Prospects*, the NAE and NRC argued that, "in contrast to science, mathematics, and even technology education, all of which have established learning standards and a long history in the K-12 curriculum, the teaching of engineering in elementary and secondary schools is still very much a work in progress (NAE and NRC, 2009, p. 2). Rationales for the inclusion of engineering into the K-12 curriculum include: (a) engineering as a real-world context to motivate science and mathematics learning, and (b) engineering design tasks as a context to promote the development of 21st century skills (Brophy et al., 2008; Hirsch et al., 2007; Koszalka et al., 2007). Ultimately, engineering is viewed as "a catalyst for integrated STEM education" given the "considerable potential value, related to student motivation and achievement, in increasing the presence of technology and, especially, engineering in STEM education ... in ways that address the current lack of integration in STEM teaching and learning (NAE and NRC, 2009, p. 150)."

The subsequent release of the *Framework for K-12 Science Education* (NRC, 2012) and *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013), as well as parallel international policies (e.g. Commonwealth of Australia, 2015; Department of Education, 2009), created demand for the rapid development of professional learning opportunities and quality STEM curricula for K-12 teachers. One challenge has been in simultaneously developing curricular components for use in K-12 classrooms while also conceptualizing and actualizing what makes for high-quality integrated STEM instruction. This paper addresses this challenge in reform efforts by describing the development and implementation of integrated STEM curricula in K-12 classrooms.

Brief history of integrated STEM curriculum development

New curricula are considered central to education reforms (Powell and Anderson, 2002), and these curricula may be developed as a mandated national curriculum or as in the United States through state, school, and teacher-level decisions. In the case of integrated STEM education, it is important to acknowledge that interdisciplinary approaches to curriculum development is not a new idea. Researchers have argued for decades that meaningful learning requires that students have opportunities to learn through the development and application of knowledge and skills to real-world problems (e.g. Beane, 1995; Brooks and Brooks, 1993; Czerniak et al., 1999; Hirst, 1974; Jacobs, 1989). Given that real-world problems are inherently interdisciplinary in nature, teaching approaches anchored by real-world problems require integrated approaches. For example, Problem-Based Learning (PBL), which originated in medical education, uses complex real-world problems to promote student learning of concepts and principles by proposing possible solutions to the problem (Hung et al., 2008). Although PBL can be used across all disciplines, its approach to teaching and learning is well suited to contexts that require students to draw upon STEM knowledge to solve problems.

Integration within the STEM disciplines has historically focused on the integration of science and mathematics (e.g., Berlin and White, 1995; Davison et al., 1995; Huntley, 1998), with little attention to technology and engineering (Bybee, 2010; Hoachlander and Yanofsky, 2011). However, the recent addition of engineering into national science standards (NRC, 2012; NGSS Lead States, 2013) has led to a renewed focus on integration. Engineering by definition requires the application of knowledge from science, mathematics, and technology to solve real-world problems through engaging in an engineering design process (Brophy et al., 2008; Douglas et al., 2004). Prior to calls from educational reforms to explicitly include engineering into science standards, a review of K-12 engineering curricula (NAE and NRC, 2009) found that while there are implicit connections between engineering and the disciplines of mathematics, science, and technology, existing K-12 engineering curricula do not clearly and explicitly integrate across the STEM disciplines. The review found a wide variety of curricular approaches to K-12 engineering with the only strong commonality being the engagement of students in the engineering design process; “Virtually all of the curricula present a paradigm for designing solutions to problems that include a cyclical pattern of steps” (NAE and NRC, 2009, p. 82). While engineering was often used to contextualize science learning with a real-world problem, rarely were students asked to engage in scientific inquiry “to generate data that could then be used to inform engineering design decisions” (NAE and NRC, 2009, p. 157). Connections to mathematics were found to be weak and limited to gathering, organizing, analyzing, interpreting, and presenting data, additionally, technology was most often represented simply as the product of engineering. The release of the *Framework for K-12 Science Education* (NRC, 2012) and the NGSS (NGSS Lead States, 2013) led to a critical need for both quality integrated STEM curriculum and professional development for K-12 teachers.

Integrated STEM curriculum development and implementation: the case of *EngrTEAMS*

One response to this demand was the National Science Foundation (NSF) funded Mathematics and Science Partnership grant, *EngrTEAMS: Engineering to Transform the Education of Analysis, Measurement, and Science* (DRL #1238140). *EngrTEAMS* aimed to increase grade 4–8 student learning of science concepts, as well as the mathematics concepts related to data analysis and measurement, by using an engineering design-based approach to curricular development and teacher professional development. Each summer of the *EngrTEAMS* grants, teachers from five partner school districts attended a three-week summer professional development (PD). The PD included learning opportunities for teachers related to STEM teaching and learning, including integration work related to a specific science content area (earth science, life science or physical science topics) relevant to their teaching assignment. During the PD, small teams of teachers developed new integrated STEM units with the support of a graduate student coach. Teachers piloted their curriculum in an informal summer camp on a university campus and used this experience to revise their curriculum before implementing the unit in their respective classrooms during the academic year. Final revisions to the integrated STEM units were made at the end of each academic year. By the end of the project, a total of 87 teachers participated in writing 50 unique integrated STEM curricula.

Need for an integrated STEM curriculum framework

Even after the implementation of the NGSS in the United States and related global policies, there remains disagreement on focus, models, and effective approaches for integrated STEM instruction (Bybee, 2010; Moore et al., 2020). In spite of these global policy calls for STEM education, there is no consensus on a theoretical or practical definition of integrated STEM to guide curriculum development. Based on common implementations of STEM education, Bybee (2010) outlined a continuum of models ranging from STEM disciplines as completely separate and disconnected to full integration of all four STEM disciplines, more akin to interdisciplinary and transdisciplinary models described by Vasquez et al. (2013). While disagreements still exist, there is a developing consensus around features indicative of quality integrated STEM that can be used to guide curriculum development (Moore et al., 2020). Consensus features of quality integrated STEM include: (a) using real-world contexts to engage students in authentic and meaningful learning (Kelley and Knowles, 2016; Moore et al., 2014a; Sanders, 2009), (b) using student-centered pedagogies (Kelley and Knowles, 2016; Moore et al., 2014a), (c) supporting the development of 21st century skills such as creativity, collaboration, communication, and critical thinking (Honey et al., 2014; Moore et al., 2014b), and (d) making the connections between STEM disciplines explicit to students (English, 2016; Herschbach, 2011; Honey et al., 2014; Kelley and Knowles, 2016; Moore et al.,

Table 1 Six tenets of the Framework for Integrated STEM.

Tenet	Tenet description	Supporting literature
1	Quality STEM integration contextualizes learning for students by engaging students in motivating and authentic problems	Breiner et al. (2012); Brophy et al. (2008) Brown et al. (2011); Carlson and Sullivan (2004); Frykholm and Glasson (2005); Kelley and Knowles (2016); Kennedy and Odell (2014); Labov et al. (2010); Rinke et al. (2016); Sanders (2009)
2	Integrated STEM curricula must include mathematics and/or science content as main objectives of the activity.	Harris and Felix (2010); Fortus et al. (2004); Mehalik et al. (2008)
3	Students develop and apply mathematics and science content by engaging in the engineering design process .	Brown et al. (2011); Dym et al. (2005); English (2016); Herschbach (2011); Honey et al. (2014); Kelley and Knowles (2016); Kennedy and Odell (2014); Labov et al. (2010); Rinke et al. (2016); Sanders (2009)
4	Engineering problems have multiple solution pathways and engineers use evidence-based reasoning to iteratively improve their design solutions providing opportunities for students to learn from failure through redesign	Dym et al. (2005); Morrison (2006); NRC and NAE (2009), NRC (2012)
5	The implementation of integrated STEM requires the use of student-centered pedagogies such as inquiry, discovery, and engineering design that help students develop and apply knowledge in a manner that deepens conceptual understanding	Breiner et al. (2012); Furner and Kumar (2007); Kelley and Knowles (2016); Kennedy and Odell (2014); Labov et al. (2010); Rinke et al. (2016); Sanders (2009); Smith et al. (2005); Stinson et al. (2009)
6	Quality integrated STEM curricula emphasizes teamwork and communication .	Brown et al. (2011); Carlson and Sullivan (2004); Dym et al. (2005); Honey et al. (2014); Kennedy and Odell (2014); NRC and NAE (2009); NGA (2010); Rinke et al. (2016); Smith et al. (2005)

2014a, 2014b). Given the prominence of engineering within STEM policy documents (e.g. NRC, 2011, 2012; NGSS Lead States, 2013), an engineering context or problem is often central to integrated STEM curriculum (Bryan et al., 2016; Hmelo et al., 2000; Mehalik et al., 2008; Moore et al., 2014a; Sadler et al., 2000).

Specifically, the teachers' curriculum writing within *EngrTEAMS* was guided by two frameworks for STEM integration (Moore et al., 2014a, 2014b), as well as by state science standards, which included engineering at the time of the project. The framework for integrated STEM instruction (Moore et al., 2014a) includes six tenets to guide integrated STEM curriculum development (Table 1). The *Framework for Quality K-12 Engineering Education* (Moore et al., 2014b) describes nine key indicators, which expand upon the framework for STEM integration and provide more specific descriptions for the integration of engineering content and practices (Table 2). The indicators in the *Framework for Quality K-12 Engineering Education* articulate quality features for the full

Table 2 Framework for quality K-12 engineering education.

Key indicator	Description
Process of design (POD)	Engineering design process
Problem and background (POD-PB)	Identifying an engineering problem in need of a solution—includes researching the problem, participating in learning activities to gain background knowledge, and identifying constraints
Plan and implementation (POD-PI)	Developing and implementing a plan for a design solution to an engineering problem—includes brainstorming, developing multiple solutions, and evaluating pros and cons of multiple solutions
Test and evaluate (POD-TE)	Generating a testable hypothesis/question and designing experiments to evaluate them—includes data collection and analysis for use in redesign
Apply science, engineering, and mathematics (SEM)	Applying mathematics or science in the context of solving engineering problems
Engineering thinking (EThink)	Applying independent, reflective, and metacognitive thinking to improve solutions and learn from failure—includes systems thinking, creativity, optimism, perseverance, and innovation
Conceptions of engineers and engineering (CEE)	Developing an understanding of what engineers do—includes work driven by client needs, design under constraints, and learning about various engineering disciplines
Engineering tools (Etool)	Developing proficiency using techniques, skills, processes, and tools used in engineering
Issues, solutions, and impacts (ISI)	Understanding the impact of solutions in a global, economic, environmental, and societal context
Ethics	Understanding ethical considerations inherent in engineering - includes responsibility to natural and client resources; effects of a design on public health and safety; governmental regulations and professional standards; and integrity
Teamwork	Developing teamwork and interpersonal skills—includes participating in collaborative groups in a variety of roles
Communication related to engineering (Comm-Engr)	Developing technical writing skills and ability to communicate technical ideas in common language—includes symbolic and pictorial representations

Table 3 Alignment of Diggin' for Fools Gold curriculum with STEM curriculum frameworks.

	<i>Lesson description</i>	<i>Alignment to STEM frameworks</i>
1	Students are introduced to the engineering design challenge. Students review the engineering design process and engage in problem scoping as they discuss and identify relevant criteria and constraints.	Tenets 1, 5, and 6 CEE, POD-PB, teamwork
2	Students explore properties of magnets and magnetism through a series of laboratory stations.	Tenets 2,5, and 6 POD-PB, teamwork
3	Students explore how they can use an electromagnet to pick up magnetic materials and generate a list of variables that might impact the strength or performance of the electromagnet.	Tenet 2, 5, and 6 POD-PB, teamwork
4	Students carry out a guided inquiry activity to test the impact of the number of coils of the strength of the electromagnet. Students represent their data in graphical form and writing claims supported with evidence about the effect of coils on electromagnets.	Tenets 2, 5, and 6 POD-PB, teamwork
5	Teams choose another variable to test, create tables/visual displays to look for patterns in the data, and share their claims and evidence with the class.	Tenets 2, 5, and 6 POD-PB, teamwork
6	Using data from lessons 4 and 5, students design, build, and test an electromagnetic claw arm prototype, using their test data to justify their design decisions.	Tenets 3, 4, 5, and 6 POD-PI, SEM, Ethink, Comm-Engr, teamwork
7	Students redesign and retest their new designs with several different toys of different sizes. The students present their best design to the client describing the results of the tests, as well as the reasoning behind their design choices.	Tenet 3, 4, 5, and 6 POD-TE, SEM, Ethink, Comm-Engr

K-12 scope and sequence; it is not intended that all nine indicators be present in every unit or lesson, rather that students experience all of the indicators at some point in their K-12 education. Three of the indicators, Process of Design (POD), Science, Engineering, and Mathematics (SEM), and Engineering Thinking (ETthink) are considered the most central and should be included in all curricular units (Moore et al., 2014b).

An example of an integrated STEM curriculum that aligns with the above frameworks is *Electromagnetic Crane Arm*. This curriculum was developed by a team of three elementary science specialists as part of the *EngrTEAMS* project to be used in their fourth grade classrooms. In this curriculum, students are introduced to a fictitious client, Galactic Games, that is concerned that people are no longer playing their claw arcade game due to reports that these games are rigged. The students are asked to develop a prototype of an electromagnetic attachment for the arm that will be used in the game instead of the mechanical claw. Students learn about magnetism and electromagnetism through a series of guided inquiry science lessons and then apply their knowledge of science to the iterative design of a new electromagnetic attachment that meets the criteria and constraints set by the client. Mathematics is integrated as measurement, data analysis, and graphing. Technology is integrated both as the product of engineering and through the integration of instructional technology to support data collection and analysis. Details for each of the seven lessons and their alignment to the STEM frameworks are shown in Table 3.

Broad approaches to engineering integration

Crotty et al. (2017) analyzed the 20 integrated STEM curricula developed during the first year of *EngrTEAMS* to understand the curricular approaches to the integration of an engineering design challenge into science units. Three approaches to engineering integration were identified: (a) a culminating project where engineering was introduced only at the end of a science unit and aimed to require students to use their science knowledge, (b) implicit integration where the engineering design challenge provided context for the unit and was revisited as a culminating project at the end of the unit, and (c) explicit integration where the EDC not only provided context for the unit, but was purposefully revisited during each lesson leading up to the culminating EDC. Simply including an engineering design challenge as a culminating project produced significantly lower student learning related to engineering than when the engineering design challenge was also included at the beginning of the unit to frame the learning and revisited throughout the unit. Interestingly, no differences in student learning were found between implicit and explicit approaches, which suggests a more nuanced assessment of curriculum quality is necessary. While valuable in understanding the use of an EDC within a STEM curricular unit, this work did not assess the quality of the integration or coherence of conceptual development, which will be addressed in the following two sections.

STEM integration curriculum assessment

The STEM Integration Curriculum Assessment (STEM-ICA) (Guzey et al., 2016) was developed from the framework for integrated STEM instruction (Moore et al., 2014a) as a mechanism to evaluate integrated STEM curricula. The STEM-ICA is composed of nine items: (a) motivating and engaging context, (b) engineering design, (c) integration of science content, (d) integration of mathematics content, (e) instructional strategies, (f) teamwork, (g) communication, (h) assessment, and (i) organization. Each item in the STEM-ICA is rated on a 5-point scale from 0 to 4 (0: not present, 1: weak, 2: adequate, 3: good, 4: excellent). Table 4 provides some examples of the descriptive elements within each of the nine items, and the full version of the STEM-ICA can be found in the published paper (Guzey et al., 2016).

Table 4 Description of the STEM-ICA items.

<i>STEM-ICA item</i>	<i>Sample descriptive elements</i>
To what extent does the curriculum unit use a motivating and engaging context?	- allow students to draw on their personal knowledge and experiences? - engage and motivate students from different backgrounds? - provide a context with a compelling purpose.
To what extent does the curriculum unit allow students to learn engineering design by integrating an engineering design challenge?	- provide opportunities to apply engineering process in authentic situations? - require students to use engineering design processes? - allow students opportunities to learn from failure? - allow students to redesign? - require students to consider constraints, safety, reliability, risks, alternatives, trade-offs, and/or ethical considerations? - promote engineering habits of mind?
To what extent does the curriculum unit integrate science content needed to solve the engineering challenge and support in-depth understanding?	- integrate science concepts that are grade level appropriate? - require students to learn, understand, and use fundamental science concepts necessary to solve the engineering challenge? - promote coherent conceptual understanding of science?
To what extent does the curriculum unit integrate mathematics content needed to solve the engineering challenge and support in-depth understanding?	- integrate mathematics concepts that are grade level appropriate? Require students to learn, understand, and use fundamental mathematics concepts necessary to solve the engineering challenge? - promote coherent understanding of mathematical thinking?
To what extent does the curriculum unit support student centered teaching strategies?	- contains lessons and activities that are student-centered? - contain activities that require students to collect and analyze information or data before arriving at a solution? - engage students in argumentation and evidence-based reasoning?
How well does the curriculum unit enable students to develop teamwork skills?	- require students to collaborate with others? - build in instructional strategies that encourage positive team interactions?
How well does the curriculum unit enable students to develop communication	- require students to communicate science concepts and engineering thinking/engineering solutions/products? - include a requirement for using evidence to support claims and design decisions?
To what extent do the assessments and required assignments in the curriculum unit measure students' knowledge and skills?	- are closely aligned with the learning objectives and goals and content from the multiple disciplines of STEM? - provide students opportunities to produce evidence of understanding and abilities in different ways through performance tasks?
How well is the curriculum unit organized?	include activities/lessons that flow in a logical and sequential order so they build on each other?

The STEM-ICA was used to evaluate the integrated STEM curricula developed during the first year of *EngrTEAMS* (see [Guzey et al., 2016](#) for details). Two items, mathematics integration and communicating mathematics, science, and engineering thinking, were found to only weakly contribute to the overall quality of the STEM units. Additionally, none of the curricula scored highly for science integration. The STEM-ICA provides valuable insight regarding the level of integration of the STEM disciplines, but similar to challenges with [Crotty et al.'s \(2017\)](#) broad approach to assessing curriculum quality, it does not provide information about the coherence of the curriculum.

Conceptual flow graphics

Given the struggle to produce integrated STEM curricula with strong integration of science and mathematics ([Guzey et al., 2016](#)), new mechanisms are needed for both researchers and teachers to understand the nature of integration within curricula. A Conceptual Flow Graphics (CFG) offers a visual means to evaluate integration across the disciplines within a given curriculum. Further, these visual depictions offer a way to identify and assess curricular coherence. CFGs are developed by identifying the main content learning for the unit and each individual lesson, then mapping the degree of connectivity between the lessons and between each lesson and the main unit goal (see [Roehrig et al., 2021](#) for detail). [Fig. 1](#) provides an example of a CFG generated for the *Electromagnetic Crane Arm* curriculum (see [Table 3](#)). A bold black border around the content goals for a lesson indicate that there is an explicit connection to the EDC through a letter or memo from the client and a solid arrow show strong connectivity between lessons. Each lesson in the *Electromagnetic Crane Arm* unit is coherently connected, for example students cannot plan a prototype design for their electromagnetic arm attachment (Lesson 5) without understanding the impact of variables such as the number of coils, number and type of batteries, and wire gauge on the strength of an electromagnet. In contrast, [Fig. 2](#) shows a CFG for a curriculum where students are asked to design a means to prevent cross-pollination between fields of non-GMO and GMO corn. Pedagogical connections through teacher guidelines within the written curriculum that direct the teacher to remind students about the engineering design challenge and the client's needs are designated with a dashed black box around the lesson concepts. A dashed arrow represents a weak connection between lessons, and the absence of an arrow represents that there is no conceptual connection between lessons. In this curriculum, while the science lessons build conceptually on each other, there are minimal

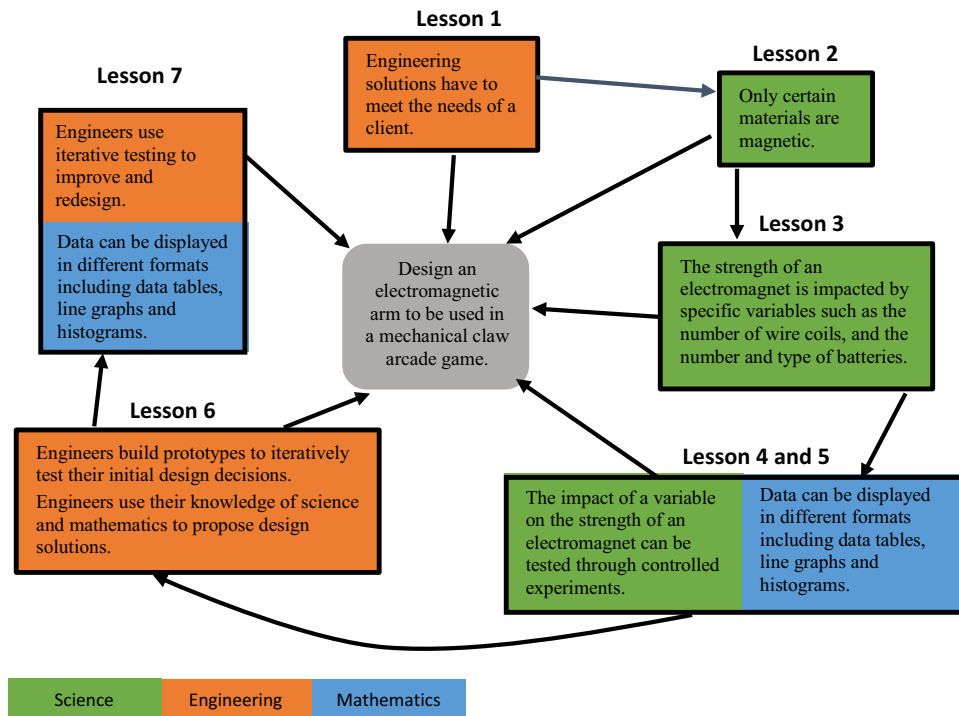


Fig. 1 Conceptual flow graphic for the *electromagnetic crane arm* unit. A *solid arrow* represents strong conceptual connection between lessons or a lesson and the Engineering Design challenge. A *solid box* around a lesson represents a contextual connection to the client through curricular materials such as a client letter.

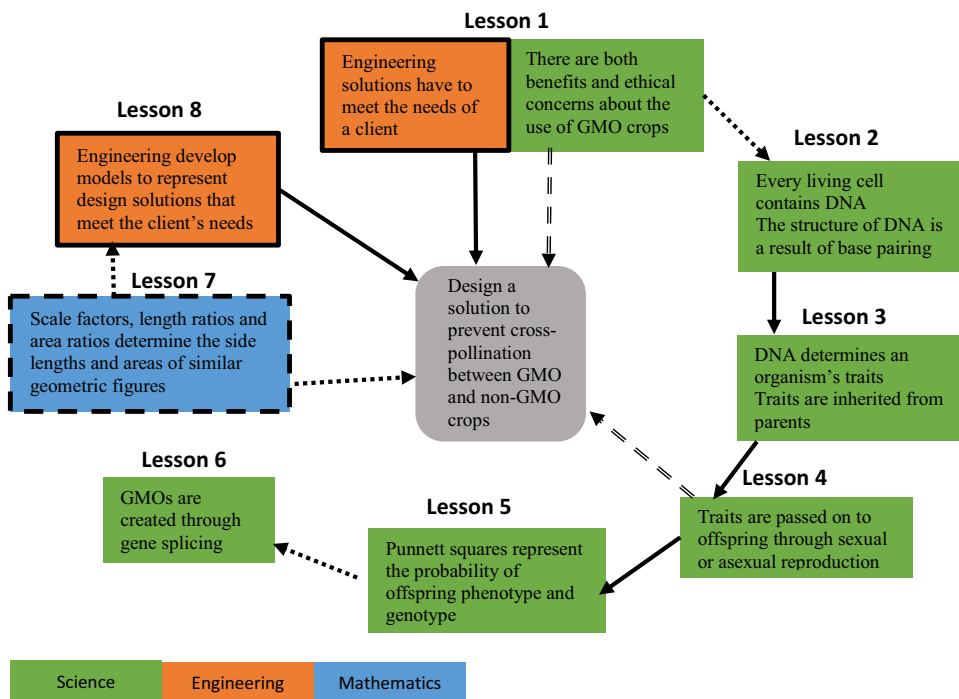


Fig. 2 Conceptual flow graphic for the *GMO corn* unit. A *solid arrow* represents strong conceptual connection between lessons or a lesson and the Engineering Design challenge. A *dotted arrow* represents weak conceptual connection between lessons or a lesson and the Engineering Design challenge. A *dashed arrow* represents a contextual connection between a science lesson and the Engineering Design challenge. A *solid box* around a lesson represents a contextual connection to the client through curricular materials such as a client letter. A *dashed box* around a lesson represents a contextual connection to the client to be made verbally by the teacher as suggested in teacher guidelines accompanying the curriculum.

connections between science lessons and the EDC. In other words, the students can propose solutions to the EDC without using any scientific content learned in the unit.

Roehrig et al. (2021) analyzed all 50 *EngrTEAMS* curricula (16 earth science, 15 life science, 19 physical science), through the use of CFGs. Of the 21 curricula that included strongly integrated and coherent lessons, as illustrated in Fig. 1, 18 were physical science curricula. Conversely, all of the life science curricula represented weak integration with coherence of the EDC with science lessons limited to providing information about the engineering context rather than content knowledge needed to develop design solutions. Similarly, all but one of the earth science curricula included engineering design lessons with, at best, weak integration between science and engineering lessons. As with other studies (e.g. Guzey et al., 2016; NAE and NRC, 2009), the CFG analysis revealed that mathematics concepts represented less than 10% of the main concepts to be learned in each lesson across all integrated STEM curricula. Aligning motivating and engaging engineering contexts with grade-level science and mathematics standards is not easy. Researchers have outlined a backward design approach (Wiggins and McTighe, 1998) where relevant STEM concepts are identified from analysis of an EDC (e.g. Chalmers et al., 2017) to generate curricular activities. However, if the relevant STEM concepts are not aligned with grade-level standards, teachers will be hesitant to adopt a curriculum that takes time from required content standards. For example, genetic engineering is a field that has myriad engaging topics, and while topics such as the polymerase chain reaction and gel electrophoresis are common applications of genetics knowledge in undergraduate courses (Phillips et al., 2008), they are beyond the expectations of a middle school or even high school science class.

Other important findings from the CFG analysis of integrated STEM curricula indicate the importance of maintaining connections to the EDC as a means to contextualize learning throughout the unit and explicitly connect science and mathematics learning to the engineering context. Two fruitful mechanisms to make the integration amongst the disciplines explicit are the use of client letters or memos and evidence-based reasoning; these mechanisms are often intertwined, as demonstrated below. Professional engineers produce design solutions to meet the needs of a client; thus, it is recommended that K-12 engineering lessons also focus on generating solutions for a client, mirroring the work of professional engineers (Brophy et al., 2008; Cunningham and Carlsen, 2014). In almost all *EngrTEAMS* curricula, a client letter was used to introduce the engineering design challenge and establish relevant criteria and constraints. In the strongly coherent and integrated curricula, there were frequent interactions with the client, including: (a) follow-up memos from the client, (b) sharing of prototypes, test data, and final designs with the client, and (c) opportunities to ask clarifying questions or get feedback from the client. Evidence-based reasoning, which will be discussed in more detail in the implementation section, was used as a specific form of communication with the client where students were expected to justify their design decisions with scientific and/or mathematical evidence (Siverling et al., 2019).

Considerations for selecting an engaging context

As already discussed, there is consensus that integrated STEM curricula should utilize authentic problems to contextualize students' learning (Kelley and Knowles, 2016; Moore et al., 2014a; Sanders, 2009). However, given the need to diversify the STEM fields (NSF, 2019), there needs to be careful consideration of the selection and representation of real world problems. For example, research suggests girls and students of color are more motivated by real-world problems that are people-oriented and focused on societal issues such as health, the environment, and social justice as opposed to more gendered engineering projects such as designing cars and rockets (Dare and Roehrig, 2020; Diekman et al. 2010; Leammukda and Roehrig, 2020).

Many engineering and integrated STEM curricula use global contexts in developing engineering design tasks. However, the use of contexts not centralized on the communities or lived experiences of the students can be problematic (Gruenewald, 2003). Keratithamkul et al. (2020) found that the use of contexts that ask students to design for a non-dominant culture can be problematic and lead to students developing deficit-based views of the people and places presented in the lesson, rather than the desired culturally competent views. Researchers have suggested a place-based (Semken and Freeman, 2008) or community-based approach (Melaville et al., 2006) to STEM curriculum development that situate learning locally and provide an opportunity for students to make a difference in their communities (Keratithamkul et al., 2020; Leammukda and Roehrig, 2020; Nadelson and Seifert, 2013).

Implementing integrated STEM curricula in K-12 classrooms

To date, little research has been published surrounding teacher and student learning related to integrated STEM, although research related to policy, curriculum, evaluation, and assessment has accounted for nearly half of STEM education-related publications (Li et al., 2020). As part of the *EngrTEAMS* project, several research studies have focused on the implementation of integrated STEM curricula. For instance, Dare et al. (2018) found that sixth grade physical science teachers struggled to make connections between the STEM disciplines. Teachers who used an implicit approach to integration noted that their students did not make connections between the science content and developing solutions to the associated EDC, and when students did not access and apply relevant content knowledge, they resorted to design by tinkering. The teachers with the highest quality lessons explicitly connected science content to engineering design decisions. This required students to develop a deep conceptual understanding of the science content; for example, simply knowing the definition of an insulator was not enough to purposefully select materials to effectively insulate a cooler. Successful teachers also provided multiple opportunities for students to explain how the science content informed their design solutions, both verbally and in writing. Successful implementation of integrated STEM curricula takes time and requires strong pedagogical skill in facilitating and supporting learning. Teachers in this study shared the constant dilemma of having to

decide how much support and guidance to provide students in their design work, balancing frustration and opportunities to problem solve and engage in creative thinking.

Dare et al. (2018) also found that teachers focused primarily on science and engineering, finding little time to explicitly include mathematics into their instruction. This minimization of including mathematics reflects teachers' conceptions of STEM as reported by Ring-Whalen et al. (2017) and Ring-Whalen et al. (2018), wherein mathematics was often treated as a means to an end or a tool only used to analyze data for science and engineering learning. Ring-Whalen et al. (2018) noted that the ways in which teachers conceptualized STEM education played a role in how curricular units were designed and implemented. Similarly, other researchers have reported that teachers' attitudes to STEM and associated reform-based teaching practices influence their implementation of integrated STEM (Thibaut et al., 2018).

Understanding integrated STEM implementation through student discourse

Given the focus of integrated STEM curricula on developing teamwork and collaboration skills, quality curricula call for students to engage in science and engineering practices in small groups. The prevalence of small group work in STEM education means that there is a critical need for implementation research to examine how students interact to construct knowledge and design solutions with their small groups, in addition to implementation research focused on the role of the teacher (e.g. Dare et al., 2018).

Several studies have explored the ways in which students use STEM disciplinary knowledge to justify their design decisions during the implementation of *EngrTEAMS* curricula (Guzey and Aranda, 2017; Mathis et al., 2016; Mathis et al., 2018; Siverling et al., 2019). Using audio-recordings of one small group in each of seven integrated STEM curriculum implementations, Siverling et al. (2019) identified 15 categories that describe how students generate and justify their design decisions throughout their small group work. Students used science content aligned with the curriculum under investigation, but also invoked science content not taught in the curriculum. Similarly, students used mathematics content both from within and outside the curriculum. Categories for justifying design decisions based on engineering principles included: (a) overall design, (b) material type, (c) material amount, (d) functionality (e) structure, (f) execution of design, (g) ease of use, (h) cost, (i) procedural step, and (j) aesthetics. Finally, students used their knowledge of and familiarity with existing technologies to justify design decisions. Results are mixed in terms of how students engaged in evidence-based reasoning during the engineering design process. Siverling et al. (2019) reported that all seven small groups used science from the curriculum and five small groups used mathematics from the curriculum in their design work. However, Guzey and Aranda (2017) found that small groups rarely used science in their justifications; rather they focused on cost as the primary factor for making design decisions. These mixed findings suggest that more work is needed to explore facets of curriculum and task design that afford opportunities for students to apply of science and mathematics knowledge during engineering design lessons. Other researchers have also found that students place an emphasis on cost and material limitations when engaging in engineering design tasks (e.g., English et al., 2013), suggesting that more detailed analysis of how constraints are integrated into integrated STEM curricula is needed. It is also noteworthy that Guzey and Aranda (2017) reported that the group that most often used science to justify their designs was explicitly prompted to do so by the teacher.

Given concerns about the under-representation of women and racial minorities in the STEM fields (NSF, 2019), it is also important to consider small group composition during integrated STEM lessons. While benefits of group work are well established (e.g. Baines et al., 2003; Fredricks et al., 2018; Fung et al., 2018), there are also well documented issues related to equity in small group participation (e.g., Jovanovic and King, 1998; Mewborn, 1999; Schnittka and Schnittka, 2016). Thus, the frequency of group work within STEM could promote inequitable participation and gender and racial stereotypes. One such strand of research has explored gender differences in participation in small group STEM activities through multimodal discourse analysis of video and transcripts of all boy, all girl, and mixed gender groups (Wieselmann et al., 2020a). This work suggests that boys and girls participate in different ways during small group STEM activities, with girls taking on more managerial roles (e.g., note-taking, keeping the group on task) and boys more inclined to participate when materials are physically manipulated (Wieselmann et al., 2019, 2020b). Interestingly, patterns of interactions shift as students transition from science-focused activities to engineering-focused activities (Wieselmann et al., 2019, 2020b). While girls did not appear to struggle in gaining access to or completing science tasks, they were more hesitant when faced with an open-ended engineering task that did not have a clear set of steps to follow to reach a pre-determined conclusion. Simultaneously, boys' participation in engineering increased when compared to the more-procedural science tasks, such that maintaining equitable participation became problematic within mixed-gendered groups. Assigned roles within these tasks that mirror cooperative learning rather than collaborative learning may be beneficial to minimize the challenges in navigating small group work in integrated STEM tasks (Dare and Roehrig, 2020). From these studies, it becomes clear that students need guidance with respect to working in small groups, which demands teacher effort to explicitly teach collaborative skills. Research has suggested that integrated STEM is a mechanism to promote interest in pursuing STEM careers and increase diversity within the field; however, this body of research suggests that the field should proceed with caution. More studies are needed to understand how the nature of the STEM tasks and the teachers' implementation of STEM curricula affect participation of groups underrepresented in STEM.

Future research in integrated STEM curriculum implementation

In order to continue to improve integrated STEM education, especially with respect to teachers' professional learning and curriculum implementation, more research needs to take place. Much of the research related to implementation relies on teacher self-report

data (e.g., Thibaut et al., 2018), and there is a need for more observational research. One of the limitations is the lack of an appropriate observation instrument that attends to the nuances of integrated STEM instruction and curriculum implementation. Existing instruments most commonly used by researchers, such as the Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002), were designed prior to the development of the NGSS and were intended for use in reform-based science classrooms rather than integrated STEM learning environments. In order to facilitate the implementation of K-12 integrated STEM curricula and the development of the nascent integrated STEM education literature, new integrated STEM observation protocols are needed for use in K-12 science and engineering classrooms. Building on the experiences of *EngrTEAMS*, a new STEM Observation Protocol (STEM-OP) has been developed (Dare et al., 2021; Roehrig et al., 2019) which provides a common metric for future research related to the implementation of integrated STEM curricula in K-12 classrooms. With additional research on curriculum development and implementation, as well as teacher professional development opportunities, the field of STEM education can become better poised to meet the challenges of preparing all students to be proficient in STEM practices and habits of mind, regardless of whether they enter a STEM career.

Acknowledgment

This work was supported by the National Science Foundation under Award Numbers 185401, 1812794, 1813342, and 1238140. The findings, conclusions, and opinions herein represent the views of the authors and do not necessarily represent the view of personnel affiliated with the National Science Foundation.

References

- Baines, E., Blatchford, P., Kutnick, P., 2003. Changes in grouping practices over primary and secondary school. *Int. J. Educ. Res.* 39, 9–34.
- Beane, J., 1995. Curriculum integration and the disciplines of knowledge. *Phi Delta Kappan* 76, 616–622.
- Berlin, D.F., White, A.L., 1995. Connecting school science and mathematics. In: House, P.A., Coxford, A.F. (Eds.), *Connecting Mathematics across the Curriculum*. 1995 National Council of Teachers of Mathematics Yearbook. National Council of Teachers of Mathematics, Reston, VA, pp. 22–33.
- Breiner, J.M., Harkness, S.S., Johnson, C.C., Koehler, C.M., 2012. What is STEM? A discussion about conceptions of STEM in education and partnerships. *Sch. Sci. Math.* 11, 3–11.
- Brooks, G.J., Brooks, G.M., 1993. In Search of Understanding: The Case for Constructivism Classroom. Association for Supervision and Curriculum Development, Alexandria, VA.
- Brophy, S., Klein, S., Portsmore, M., Rogers, C., 2008. Advancing engineering education in P-12 classroom. *J. Eng. Educ.* 97, 369–387.
- Brown, R., Brown, J., Reardon, K., Merrill, C., 2011. Understanding STEM: current perceptions. *Technol. Eng. Teacher* 70, 5–9.
- Bryan, L.A., Moore, T.J., Johnson, C.C., Roehrig, G.H., 2016. Integrated STEM education. In: Johnson, C.C., Peters-Burton, E.E., Moore, T.J. (Eds.), *STEM Road Map: A Framework for Integrated STEM Education*, first ed. Routledge, New York, NY, pp. 23–37.
- Bybee, R.W., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teacher* 70, 30–35.
- Carlson, L., Sullivan, J., 2004. Exploiting design to inspire interest in engineering across the K–16 engineering curriculum. *Int. J. Eng. Educ.* 20, 372–380.
- Chalmers, C., Cater, M.L., Cooper, T., Nason, R., 2017. Implementing big ideas to advance the teaching and learning of science, technology, engineering, and mathematics (STEM). *Int. J. Sci. Math. Educ.* 15, S25–S43.
- Commonwealth of Australia, 2015. National Innovation and Science Agenda. Commonwealth of Australia, Canberra. <https://www.industry.gov.au/strategies-for-the-future/boosting-innovation-and-science>.
- Crotty, E.A., Guzey, S., Roehrig, G.H., Glancy, A.W., Ring-Whalen, E.A., Moore, T.J., 2017. Approaches to integrating engineering in STEM units and student achievement gains. *J. Precollege Eng Educ Res* 7, 1–14.
- Cunningham, C.M., Carlsen, W.S., 2014. Teaching engineering practices. *J. Sci. Teach. Educ.* 25, 197–210.
- Czerniak, C.M., Weber, W.B., Sandmann Jr., A., Ahern, J., 1999. Literature review of science and mathematics integration. *Sch. Sci. Math.* 99, 421–430.
- Dare, E.A., Roehrig, G.H., 2020. Beyond Einstein and explosions: understanding 6th grade girls' and boys' perceptions of physics, school science, and STEM careers. *J. Women Minor. Sci. Eng.* 26 (6), 541–577.
- Dare, E.A., Ellis, J.A., Roehrig, G.H., 2018. Understanding science teachers' implementations of integrated STEM curricular units through a phenomenological multiple case study. *Int. J. STEM Educ.* 5.
- Dare, E.A., Hiwatig, B., Keratithamkul, K., Ellis, J.A., Roehrig, G.H., Ring-Whalen, E.A., ... Crotty, E. A., 2021. In Improving integrated STEM education: The design and development of a K-12 STEM observation protocol (STEM-OP) (RTP) (Ed.), *Proceedings of the 2021 ASEE Annual Conference & Exposition*. <https://peer.asee.org/improving-integrated-stem-education-the-design-and-development-of-a-k-12-stem-observation-protocol-stem-op-rtp.pdf>.
- Davison, D.M., Miller, K.W., Metheny, D.L., 1995. What does integration of science and mathematics really mean? *Sch. Sci. Math.* 95, 226–230.
- Department of Education, 2009. Report of the STEM Review. <https://www.education-ni.gov.uk/publications/report-stem-review>.
- Diekmann, A.B., Brown, E.R., Johnston, A.M., Clark, E.K., 2010. Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21, 1051–1057.
- Douglas, J., Iversen, E., Kalyandurg, C., 2004. *Engineering in the K-12 Classroom: An Analysis of Current Practices & Guidelines for the Future*. American Society for Engineering Education, Washington, DC.
- Dym, C.L., Agogino, A., Eris, O., Frey, D.D., Leifer, L.J., 2005. Engineering design thinking, teaching, and learning. *J. Eng. Educ.* 103–120.
- English, L.D., 2016. STEM education K–12: perspectives on integration. *Int. J. STEM Educ.* 3, 3.
- English, L.D., Hudson, P., Dawes, L., 2013. Engineering-based problem solving in the middle school: design and construction with simple machines construction with simple machines. *J. Precollege Eng. Educ. Res.* 3, 43–55.
- Fortus, D., Dershimer, R.C., Krajcik, J., Marx, R.W., Mamlok-Naaman, R., 2004. Design-based science and student learning. *J. Res. Sci. Teach.* 41, 1081–1110.
- Fredricks, J.A., Hofkens, T., Wang, M.-T., Mortenson, E., Scott, P., 2018. Supporting girls' and boys' engagement in math and science learning: a mixed methods study. *J. Res. Sci. Teach.* 55, 271–298.
- Freeman, B., Marginson, S., Tytler, R., 2014. *The Age of STEM: Educational Policy and Practice across the World in Science, Technology, Engineering and Mathematics*. Routledge, New York, NY.
- Frykholm, J., Glasson, G., 2005. Connecting science and mathematics instruction: pedagogical context knowledge for teachers. *Sch. Sci. Math.* 105, 127–141.
- Fung, D., Hung, V., Lui, W., 2018. Enhancing science learning through the introduction of effective group work in Hong Kong secondary classrooms. *Int. J. Sci. Math. Educ.* 16, 1291–1314.

- Furner, J.M., Kumar, D.D., 2007. The mathematics and science integration argument: a stand for teacher education. *Eurasia J. Math. Sci. Technol. Educ.* 3, 185–189.
- Gruenewald, D.A., 2003. Foundations of place: a multidisciplinary framework for place-conscious education. *Am. Educ. Res. J.* 40, 619–654.
- Guzey, S.S., Aranda, M., 2017. Student participation in engineering practices and discourse: an exploratory case study. *J. Eng. Educ.* 106, 585–606.
- Guzey, S.S., Moore, T.J., Harwell, M., 2016. Building up STEM: an analysis of teacher-developed engineering design-based STEM integration curricular materials. *J. Precollege Eng. Educ. Res.* 6, Article 2.
- Harris, J., Felix, A., 2010. A project-based, STEM-integrated team challenge for elementary and middle school teachers in alternative energy. In: Crawford, C. (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010*. AACE, Chesapeake, VA, pp. 3566–3573.
- Herschbach, D.R., 2011. The STEM initiative: constraints and challenges. *J. STEM Teacher Educ.* 48, 96–112.
- Hirsch, L.S., Carpinelli, J.D., Kimmel, H., Rockland, R., Bloom, J., 2007. The differential effects of pre-engineering curricula on middle school students' attitudes to and knowledge of engineering careers. In: *Proceedings of 2007 Frontiers in Education Conference*, Milwaukee, WI.
- Hirst, P.H., 1974. *Knowledge and the Curriculum: A Collection of Philosophical Papers*. Routledge and Kegan Paul, London.
- Hmelo, C., Douglas, H., Kolodner, J., 2000. Designing to learn complex systems. *J. Learn. Sci.* 9, 247–298.
- Hoachlander, G., Yanofsky, D., 2011. Making STEM real: by infusing core academics with rigorous real-world work, linked learning pathways prepare students for both college and career. *Educ. Leader* 68, 60–65.
- Honey, M., Pearson, G., Schweingruber, A., 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. National Academies Press, Washington.
- Hung, W., Jonassen, D.H., Liu, R., 2008. Problem-based learning. In: Spector, J.M., van Merriënboer, J.G., Merrill, M.D., Driscoll, M. (Eds.), *Handbook of Research on Educational Communications and Technology*, third ed. Erlbaum, Mahwah, NJ, pp. 485–506.
- Huntley, M.A., 1998. Design and implementation of a framework for defining integrated mathematics and science education. *Sch. Sci. Math.* 98, 320–327.
- Jacobs, H.H., 1989. *Interdisciplinary Curriculum: Design and Implementation*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Jovanovic, J., King, S.S., 1998. Boys and girls in the performance-based science classroom: who's doing the performing? *Am. Educ. Res. J.* 35, 477–496.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3, 1–11.
- Kennedy, T.J., Odell, M.R.L., 2014. Engaging students in STEM education. *Sci. Educ. Int.* 25, 246–258.
- Keratithamkul, K., Kim, J.N., Roehrig, G.H., 2020. Cultural competence or deficit-based view? A qualitative approach to understanding middle school students' experience with culturally framed engineering. *Int. J. STEM Educ.* 7.
- Koszalka, T., Wu, Y., Davidson, B., 2007. Instructional design issues in a cross-institutional collaboration within a distributed engineering educational environment. In: Bastiaens, T., Carliner, S. (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007*. AACE, Chesapeake, VA, pp. 1650–1657.
- Labov, J.B., Reid, A.H., Yamamoto, K.R., 2010. Integrated biology and undergraduate science education: a new biology education for the twenty first century? *CBE Life Sci. Educ.* 9, 10–16.
- Leammukda, F.D., Roehrig, G.H., 2020. Community-based conceptual framework for STEM integration. Paper presented at the annual meeting of the association for science teacher education. San Antonio, TX.
- Li, Y., Wang, K., Xiao, Y., Froyd, J.E., 2020. Research and trends in STEM education: a systematic review of journal publications. *Int. J. STEM Educ.* 7, 1–16.
- Mathis, C.A., Siverling, E.A., Glancy, A., Guzey, S.S., Moore, T.J., 2016. Students' use of evidence-based reasoning in K-12 engineering: a case study (Fundamental). In: *Proceedings of ASEE Annual Conference and Exposition, Conference Proceedings*. ASEE, New Orleans, LA.
- Mathis, C.A., Siverling, E.A., Moore, T.J., Douglas, K.A., Guzey, S.S., 2018. Supporting engineering design ideas with science and mathematics: a case study of middle school life science students. *Int. J. Educ. Math. Sci. Technol.* 6, 424–442.
- Mehalik, M.M., Doppelt, Y., Schunn, C.D., 2008. Middle-school science through design-based learning versus scripted inquiry: better overall science concept learning and equity gap reduction. *J. Eng. Educ.* 97, 71–85.
- Melaville, A., Berg, A.C., Blank, M.J., 2006. Community-based learning: engaging students for success and citizenship. *Partnerships/Community* 40.
- Mewborn, D.S., 1999. Creating a gender equitable school environment. *Int. J. Leader. Educ.* 2, 103–115.
- Moore, T.J., Johnston, A.C., Glancy, A.W., 2020. STEM integration: a synthesis of conceptual frameworks and definitions. In: Johnson, C.C., Moore, T.J., Mohr-Schroeder, M., English, L. (Eds.), *Handbook of Research in STEM Education*, first ed. Routledge, New York, NY, pp. 3–16.
- Moore, T.J., Glancy, A.W., Tank, K.M., Kersten, J.A., Smith, K.A., 2014a. A framework for quality K-12 engineering education: research and development. *J. Precollege Eng. Educ. Res.* 4, 1–13.
- Moore, T.J., Stohlmann, M.S., Wang, H.-H., Tank, K.M., Roehrig, G.H., 2014b. Implementation and integration of engineering in K-12 STEM education. In: Strobel, J., Purzer, S., Cardella, M. (Eds.), *Engineering in Pre-college Settings: Research into Practice*, first ed. Sense Publishers, Rotterdam, Netherlands, pp. 35–60.
- Morrison, J., 2006. *TIES STEM Education Monograph Series, Attributes of STEM Education*. TIES, Baltimore, MD.
- Nadelson, L.S., Seifert, A.E., 2013. Perceptions, engagement, and practices of teachers seeking professional development in place-based integrated STEM. *Teach. Educ. Pract.* 26, 242–265.
- National Academy of Engineering, National Research Council, 2009. *Engineering in K-12 Education: Understanding the Status and Improving the Prospects*. The National Academies Press, Washington, DC.
- National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010. *Common Core State Standards*. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington, DC.
- National Research Council, 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. The National Academies Press, Washington, DC.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press, Washington, DC.
- National Research Council, 2011. *Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics*. National Academies Press, Washington, DC.
- National Science Foundation, 2019. *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019*. (Special Report NSF 19-304). National Center for Science and Engineering Statistics, Arlington, VA.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, by States*. National Academies Press, Washington, DC.
- Phillips, A.R., Robertson, A.L., Batzil, J., Harris, M., Miller, S., 2008. Aligning goals, assessments, and activities: an approach to teaching PCR and gel electrophoresis. *CBE Life Sci. Educ.* 7, 96–106.
- Powell, J.C., Anderson, R., 2002. Changing teachers' practice: curriculum materials and science education reform in the USA. *Stud. Sci. Educ.* 37, 107–135.
- Rinke, C.R., Gladstone-Brown, W., Kinlaw, C.R., Cappiello, J., 2016. Characterizing STEM teacher education: affordances and constraints of explicit STEM preparation for elementary teachers. *Sch. Sci. Math.* 116, 300–309.
- Ring-Whalen, E.A., Dare, E.A., Roehrig, G.H., Titu, P., Crotty, E., 2018. How teachers' conceptions of integrated STEM education are reflected in curriculum writing. *Int. J. Educ. Math. Sci. Technol.* 6, 343–362.
- Ring, E.A., Dare, E.A., Crotty, E.A., Roehrig, G.H., 2017. The evolution of teacher conceptions of STEM education throughout an intensive professional development experience. *J. Sci. Teach. Educ.* 28, 444–467.
- Roehrig, G.H., Dare, E.A., Ring-Whalen, E.A., Wieselmann, J.R., 2021. Understanding coherence and integration in integrated STEM curriculum. *Int. J. STEM Educ.* 8 (2) <https://doi.org/10.1186/s40594-020-00259-8>.
- Roehrig, G.H., Ring-Whalen, E.A., Dare, E.A., Ellis, J.A., Wieselmann, J.R., 2019. The development of a K-12 integrated STEM observation protocol. In: *Proceedings of the 2019 ASEE Annual Conference and Exposition*. ASEE, Tampa, FL.
- Sadler, P., Coyle, H., Schwartz, M., 2000. Engineering competitions in the middle school classroom: key elements in developing effective design challenges. *J. Learn. Sci.* 9, 299–327.

- Sanders, M., 2009. STEM, STEM education, STEM mania. *Technol. Teach.* 68, 20–26.
- Sawada, D., Piburn, M.D., Judson, E., Turley, J., Falconer, K., Benford, R., Bloom, I., 2002. Measuring reform practices in science and mathematics classrooms: the reformed teaching observation protocol. *Sch. Sci. Math.* 102, 245–253.
- Schnittka, J., Schnittka, C., 2016. “Can I drop it this time?” gender and collaborative group dynamics in an engineering design-based afterschool program. *J. Precollege Eng. Educ. Res.* 6, 1–24.
- Semken, S., Freeman, C.B., 2008. Sense of place in the practice and assessment of place-based science teaching. *Sci. Educ.* 92, 1042–1057.
- Siverling, E.A., Suazo-Flores, A., Mathis, C.A., Moore, T.J., 2019. Students’ use of STEM content in design justifications during engineering design-based STEM integration. *Sch. Sci. Math.* 119, 457–474.
- Smith, K.A., Sheppard, S.D., Johnson, D.W., Johnson, R.T., 2005. Pedagogies of engagement. Classroom-based practices. *J. Eng. Educ.* 94, 87–101.
- Stinson, K., Harkness, S., Meyer, H., Stallworth, J., 2009. Mathematics and science integration: models and characterizations. *Sch. Sci. Math.* 109, 153–161.
- Thibaut, L.H., Knipprath, H., Dehaene, W., Depaepe, F., 2018. The influence of teachers’ attitudes and school context on instructional practices in integrated STEM education. *Teach. Teach. Educ.* 71, 190–205.
- Toulmin, C., Groome, M., 2007. Building a Science, Technology, Engineering and Math Agenda. National Governors Association, Washington, DC.
- Vasquez, J., Sneider, C., Comer, M., 2013. *STEM Lesson Essentials, Grades 3–8: Integrating Science, Technology, Engineering, and Mathematics*. Heinemann, Portsmouth, NH.
- Wieselmann, J.R., Keratithamkul, K., Dare, E.A., Ring-Whalen, E.A., Roehrig, G.H., 2020a. Discourse analysis in integrated STEM activities: methods for exploring power dynamics in small group interactions. *Res. Sci. Educ.* 51, 113–133.
- Wieselmann, J.R., Dare, E.A., Ring-Whalen, E.A., Roehrig, G.H., 2020b. “I just do what the boys tell me”: exploring small group student interactions in an integrated STEM unit. *J. Res. Sci. Teach.* 57, 112–144.
- Wieselmann, J.R., Dare, E.A., Roehrig, G.H., Ring-Whalen, E.A., 2019. Participation in small group engineering design activities at the elementary and middle school level: an investigation of gender differences. In: *Proceedings from the 126th ASEE Annual Conference & Exposition*. American Society for Engineering Education, Tampa, FL.
- Wiggins, G., McTighe, J., 1998. *Understanding by Design*. ASCD, Alexandria, VA.

Connecting play to STEM concepts, practices and processes: review of research on play within STEM learning environments

Amber Simpson and Monika Fridrich, Department of Teaching, Learning, and Educational Leadership, Binghamton University – SUNY, Binghamton, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	164
Methods	165
Inclusion and exclusion criteria	165
Search and data collection process	165
Data analysis	166
Research question one	166
Research question two	167
Limitations	168
Results	168
Research question one	168
Research question two	170
Free play	170
Guided play	171
Game play	171
Other play types	171
Discussion	172
Research question one	172
Research question two	173
Conclusion	173
References	174
Relevant website	176

Introduction

Play is a leading activity that is fundamental to a child's learning and development (Piaget, 1952; Vygotsky, 2016). Play naturally provides high motivational forces, a zone of proximal development, and social engagement (Vygotsky, 2016). Through play, children learn about the world around them and develop self-regulation, oral language, cognitive, and social competence skills (Coople and Bredekamp, 2009; Elkind, 2007; Piaget, 1952; Vygotsky, 2016).

Unfortunately, play has been rapidly disappearing from our children's lives (Brown and Vaughan, 2009; Gray, 2013; Hofferth and Sandberg, 2009; Louv, 2005; Miller and Almon, 2009). For example, a comparison study of how US children spent their time showed that children age six to eight spent 25% less time in indoor/outdoor play in 1997 than they did in 1981 (Hofferth and Sandberg, 2009). Among the many reported reasons for this decline of play were increase demands of academic standards in younger age (Bassok et al., 2016; Carlsson-Paige et al., 2015; Miller and Almon, 2009), and social pressure of extra-curriculum activities (Brown and Vaughan, 2009; Gray, 2013; Louv, 2005). While these and other factors push play out of the children's lives, at the same time more reports of declining physical and mental health of children, such as child's obesity, hyperactivity, and general anxiety, circulate through public media and academic communities (Burstein et al., 2019; Fryar et al., 2014). Recognizing play as a leading activity of learning and increased academic demands on children from a very young age has prompted educators and researchers to look for ways to integrate play within and across all learning environments (Zosh et al., 2018).

In addition to framing play as fundamental to children's development, science, technology, engineering, and mathematics (STEM) is an important component for education – economic and national security in terms of workforce development (Kuenzi, 2008; Jang, 2016; National Academy of Science [NAS], 2011), diversifying the field to a more diverse population (van den Hurk et al., 2019; Watson and Froyd, 2007), and development and use of STEM-literacy skills, practices, and knowledges by citizens (Bybee, 2010; Maass et al., 2019; McDonald, 2016; Zollman, 2012). However, research has shown how K-12 schools and universities are not preparing students to be agents of social change (Beckwith and Huang, 2005; Garibay, 2015), to develop the skills, knowledges, and work activities needed to enter the STEM workforce (Jang, 2016; Maass et al., 2019; Organization for Economic Cooperation and Development, 2016), or to view STEM as a meta-discipline as opposed to independent disciplines (Zollman, 2012). Changes to the STEM learning environment has been proposed as one factor as individuals are in a continuous dialog with their physical and socio-cultural environment (Eshach, 2007; Falk and Dierking, 2000). Specific changes include use of

authentic, student-generated, and/or interdisciplinary real-world problems, provide opportunities for students to apply previously learned skills and knowledge through multiple forms of representations, encourage collaboration, and teacher guidance (Glancy and Moore, 2013; Hatisaru et al., 2019; Struyf et al., 2019).

We argue that another approach to addressing how K-12 schools and universities are preparing students for global, national, and societal issues in STEM is through a play perspective. Conversely, we also contend that the STEM learning environment is a potential context for addressing the decline in play. Therefore, we conducted a literature review specific to play in STEM learning environments to address the following two research questions: (1) How is play conceptualized and operationalized in research within STEM learning environments? How might this differ by STEM field? (2) Utilizing play as a research lens, what characteristics of the STEM learning environment lead to positive developmental outcomes in children and young adolescents? The significance of this study is that it examined research literature across two fields, play research and STEM learning environments. This study also revealed gaps in the research literature related to play within the STEM learning environments including lack of research on the play for upper elementary and middle grades (9–14 years) and favoring math over other STEM fields. Additionally, this study highlighted the need for further research in embedding play within academic standards for all learners and exploring the social nature of play within the STEM learning environments.

Methods

Below we outline the process used to conduct the systematic literature review, which was informed by the PRISMA guidelines for systematic literature reviews (Moher et al., 2009). We also detail our analysis of the included articles to gain an understanding of play scholarship across formal, informal, and non-formal STEM learning environments.

Inclusion and exclusion criteria

Articles were initially screened based on their abstract and title, and their potential to meet the following inclusion criteria: (1) full-text empirical research utilizing methods and procedures of reporting standards for research studies as described by the American Psychological Association (APA) Publications and Communications Board Task Force (see Appelbaum et al., 2018; Levitt et al., 2018); (2) published between 1989–April 2020 in peer-reviewed journals, book articles, conference proposals, or conference proceedings; (3) published or translated into English; (4) included birth to 18 age group as a participant group or individual; (5) defined and/or situated within a theoretical or conceptual framework specific to play (e.g., Fleer, 2009; Vygotsky, 1978); and (6) conducted in a STEM learning environment defined as any physical space in which children and/or adolescents are provided with opportunities to engage with peers, adults (e.g., parents, educators), and objects/materials that promote an understanding of science, engineering, technology, and/or mathematics knowledges, literacies, and practices. This included objects and material that may not initially be viewed as related to STEM, but were transformed as such through children and/or adult's play (e.g., barrels of toy food, shopping baskets, and cash register used to engage in spontaneous numerical moments; Braham et al., 2018).

Search and data collection process

Articles were identified through the following databases: Academic Search Complete, Education Source, Environment Complete, ERIC, JSTOR, Library, Informational Science, & Technology, and PsychINFO. We used a combination of search terms, with the term “play” in the subject line being used in every search as to narrow and focus on the most relevant publications. This was used with a combination of search terms focused on the environment (i.e., learning environment, formal learning or non-formal learning or informal learning, out-of-school, summer programs, after school programs, library, museum, education environment) and STEM (e.g., STEM Education or science or technology or engineering or mathematics, science, technology, engineering, and mathematics). An asterisk was used to find relevant terms. For example, math* should capture similar terms such as maths and mathematics.

Based on the title and abstract, we maintained 87 full text articles to assess for eligibility based on the inclusion criteria. Next, we extracted the following information in the form of direct quotes: (a) research aim/questions, (b) how play was defined and situated in relevant literature, (c) context (e.g., preschool, playground), (d) STEM learning environment, and (e) results. Through this process, we discussed and removed 39 articles for four reasons: (1) did not clearly define play (e.g., Nel et al., 2017), (2) was not situated with a STEM learning environment (e.g., Hoskins and Smedley, 2019), (3) not in the targeted age group (e.g., Savin-Baden, 2010), and (4) not an empirical study (e.g., Castell et al., 2014). Forty-eight articles remained that fit our inclusion criteria. To ensure that we captured the relevant empirical research for this study, we used Web of Science to search for scholarship that cited the articles maintained (i.e., forward search) and fit our inclusion criteria. This yielded an additional 14 articles. Next, we conducted a backward search, which involved examining the references of the 48 articles for relevant articles. This produced an additional 14 articles. This process accumulated 76 articles, book articles, conferences, or proceedings to address the two research questions (see Fig. 1).

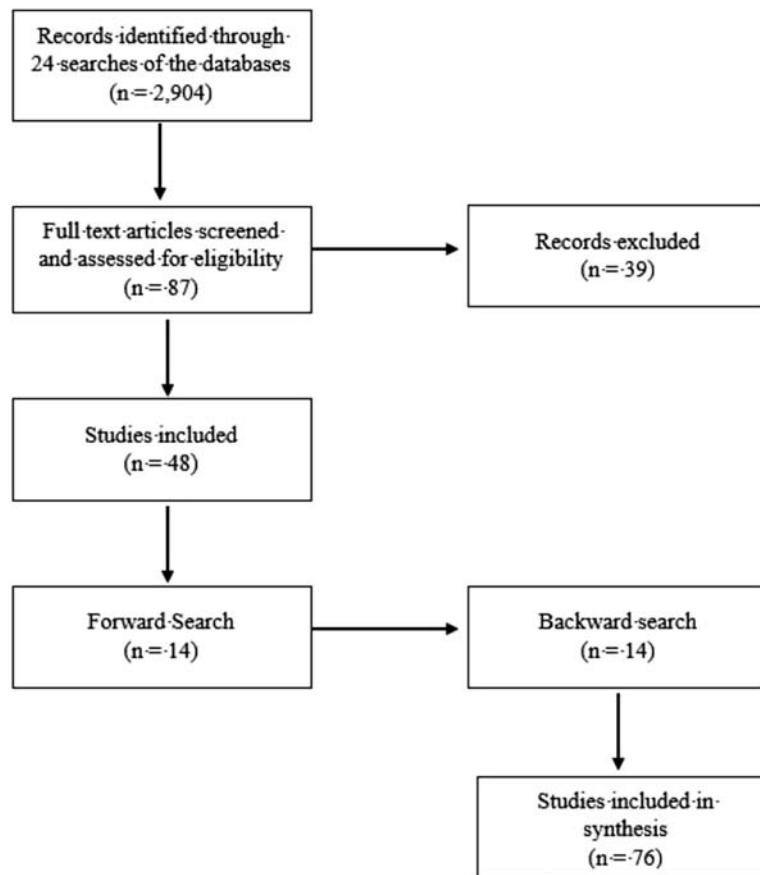


Fig. 1 Flow diagram illustrating the search process. Adapted from Moher et al. (2009).

Data analysis

Research question one

Play was operationalized according to how the play was initiated, who the play was directed by, and whether there was an explicit learning goal (Zosh et al., 2018). The context of the studies was denoted as free play, guided play, games, co-op play, playful instruction, direct instruction, or a combination of play. These were identified by each study's description of the activities located in the study's aims/purpose and/or methods section. Free play was characterized as an activity during which a child initiates and directs play without a pre-determined learning goal. To fully embrace the scope of the free play within the study, subcategories identifying the setting of the free play were considered (indoor, outdoor). Guided play was characterized as initiated and structured by an adult with an explicit learning goal in mind and directed by a child. Upon examination, two subcategories were added to further specify the characteristics of the guided play based on whether the initial input was an adult setting up the physical environment or whether the initial input was an adult inviting a child in social learning (Vygotsky, 1978). Game play was characterized as an activity that was adult-initiated and child-directed with an explicit learning goal using a game with rules. Two subcategories, didactical (i.e., non-technical) and virtual/digital were added to specify the setting of the play environment. Co-opted play was characterized as an activity that was initiated by a child and then directed by adults to achieve an explicit learning goal. Zosh et al. (2018) described the co-opted game as the adult intervening, redirecting, and changing the child's initial task to match the adult's learning goal. We extended the co-opted play characterization to include play initiated by a child during which an adult intervened by matching a specific learning goal to the child's original play. Playful instructions were characterized as adult initiated and directed with an explicit learning goal. Direct instructions were characterized as adult initiated and directed with an explicit learning goal. Articles that reported studies with exclusively direct instructions were not included in our study since they did not satisfy our inclusion criteria. However, there were several articles reporting comparison studies of free play, guided play and direct instructions (e.g., Bulunuz, 2013; Eason and Ramani, 2020). These are identified in our analysis as comparison.

Table 1 summarizes the types of play utilized to address the first research question, as well as the type of learning environment (i.e., formal, non-formal, or informal; Eshach, 2007) that was associated with each play type. Eshach's characterizations are not restricted to a physical space (e.g., all classrooms as formal environments), but how the learning environment is structured. Formal

Table 1 Types of play as associated with learning environments.

<i>Codes</i>	<i>Play definitions (Zosh et al., 2018)</i>	<i>Settings (Eshach, 2007)</i>	<i>Examples</i>
Free play	Child-initiated	Informal	Edens and Potter (2013)
• Indoor	Child-directed		Hallström et al. (2015)
• Outdoor	No explicit learning goal		Worthington and van Oers (2016)
Guided play	Adult-initiated	Formal	Adbo and Carulla (2019)
• Environment setup	Child-directed	Non-formal	Fleer (2020)
• Social learning setup	Explicit learning goal	Informal	Tian and Luo (2019)
Games play	Adult-initiated	Formal	Guberman et al. (1998)
• Didactical	Child-directed	Non-formal	Danby et al. (2018)
• Virtual/Digital	Explicit learning goal	Informal	
	Games with rules		
Co-opted play	Child-initiated	Informal	Hao and Fleer (2017)
	Adult-directed		Magnusson and Pramling (2018)
	Explicit learning goal		
Playful instruction	Adult-initiated	Informal	Watanabe (2017)
	Adult-directed		
	Explicit learning goal		
Direct instruction (non-play)	Adult-initiated	Formal	Fisher et al. (2013)
	Adult-directed		
	Explicit learning goal		

environments were characterized as structured, compulsory, teacher-led, potentially oppressive, and sequential. Non-formal environments were characterized as structured, voluntary, guided or teacher-led, usually supportive, and non-sequential. Informal environments were described as unstructured, voluntary, spontaneous, learner-led, and supportive. As an example, Bairaktarova et al. (2011) conducted a study in two mixed age preschools in a university-affiliated developmental center. This study was categorized as informal because the environment was created as free playtime - learner-centered, unstructured, and spontaneous (Eshach, 2007). As we coded the STEM environment for each study, we noticed how the learning environment was aligned with the play definitions by Zosh et al. (2018). This association is represented in Table 1.

Similar to Radovic et al. (2018), conceptualization of play was completed through highlighting central words and concepts from each article's definition of play. These central words and concepts were coded along three dimensions of play as play is a complex, in-flux activity (Hännikäinen and Munter, 2018; Reikerås, 2020). For the first dimension, we coded as to whether play was defined as a solitary activity or as a social activity in which play was done with others including peers, adults, and siblings. The second dimension we identified was concepts or practices/processes. Concepts were framed around play for cognitive development, such as scientific and mathematical concepts, while practices/processes were focused on ways of talking, acting, and behaving like an individual in a STEM field. The last dimension was coded along imaginary (e.g., role-play, hand is a telephone), manipulative (e.g., blocks, puzzles), learning (e.g., games), and creative (e.g., creating videos) (Bronson, 1995). To illustrate, we included a short excerpt from Schmitt et al. (2018, pp. 181–182) in which we italicized defining play features.

Block play in preschool provides a context where children can learn various skills (e.g., problem solving, perspective taking) through their interactions with both blocks and peers (Ginsburg, 2006). Engaging with blocks provides children with opportunities to reason and understand concepts of space and physical properties of objects (Verdine et al., 2017). Further, block play is social, verbal, and facilitates shared play, thereby giving children an opportunity to exchange ideas with peers and work together.

We coded the conceptualization of this study as social (e.g., “facilitates shared play”), manipulative (e.g., “physical properties of objects”), and practices/processes (e.g., “learn various skills (e.g., problem solving, perspective taking)”). These dimensional codes were not mutually exclusive as some studies may have been identified as imaginary and manipulative for example.

Lastly, the STEM field or learning environment was characterized as science, technology, engineering, and/or mathematics as identified in each study's aims/purpose. As an example, Wolfgang et al. (2001) was characterized as a mathematics learning environment because the researchers examined how block play performance among preschoolers was a predictor (or not) of later mathematics achievement.

Research question two

For the second research question, articles were grouped by type of play as researchers continue to consider the types of play experiences that best support young children development (e.g., Eason and Ramani, 2020; Ferrara et al., 2011). Next, positive development outcomes were synthesized across studies within a type of play while also considering components of the learning environment that researchers highlighted as shaping and impacting youths' and adolescents' positive development.

Limitations

Several limitations must be acknowledged as part of this systematic review of the literature on play-STEM learning environment. One, we do not claim to have included every empirical study that met the inclusion criteria. As such, we limited our search to the six databases listed above and searched for the term “play” within the subject line only. Yet, as argued by Doyle (2003), it may not be necessary to find every research study as the purpose of this literature review is interpretation as opposed to prediction. Two, the search terms and use of Booleans (i.e., and, or) potentially restricted the number of included studies. Three, we recognize the threat of publication bias in that scholarship with results that are unfavorable are not published (Petticrew et al., 2008). Similarly, the second research question is focused on positive developmental outcomes for children and young adolescents, which we acknowledge does provide a complete picture of characteristics of the STEM learning environment that may not be fruitful. There were only two studies included in this review that did not conclude some positive outcome specific to STEM (Edens and Potter, 2013; Hanline et al., 2010).

Results

The majority of included studies in this literature review (87%) included participants that were between the ages of birth to eight while the other 13% were with participants aged 9–14 years. Additionally, a little more than half of the studies (52.6%) focused on mathematics, 21.1% focused on science, 14.5% on technology, 10.5% on engineering, and 3.9% on STEM broadly. Additional information (e.g., context, age/grade of participants) regarding the included studies can be found online at https://orb.binghamton.edu/mathematics_fac/10.

Below, we first provide information related to how play was conceptualized and operationalized in research within STEM learning environments. Next, we synthesized positive developmental outcomes for child participants by play type, as well as consider the nature of the learning environment. In the results, not every included study was cited but accounted for within the synthesis of the findings. These studies are identified with an asterisk (*) in the reference section.

Research question one

Fig. 2 illustrates how play was conceptualized in research within STEM learning environments, as well as any differences that may exist by STEM field. The majority of research emphasized the social nature of play (e.g., Danby et al., 2018). Moreover, about two-thirds of the studies emphasized the application and understanding of STEM skills and concepts (e.g., Bulunuz, 2013; Worthington and van Oers, 2016). As an example, Gold et al. (2020) examined the association between engineering play with wooden blocks

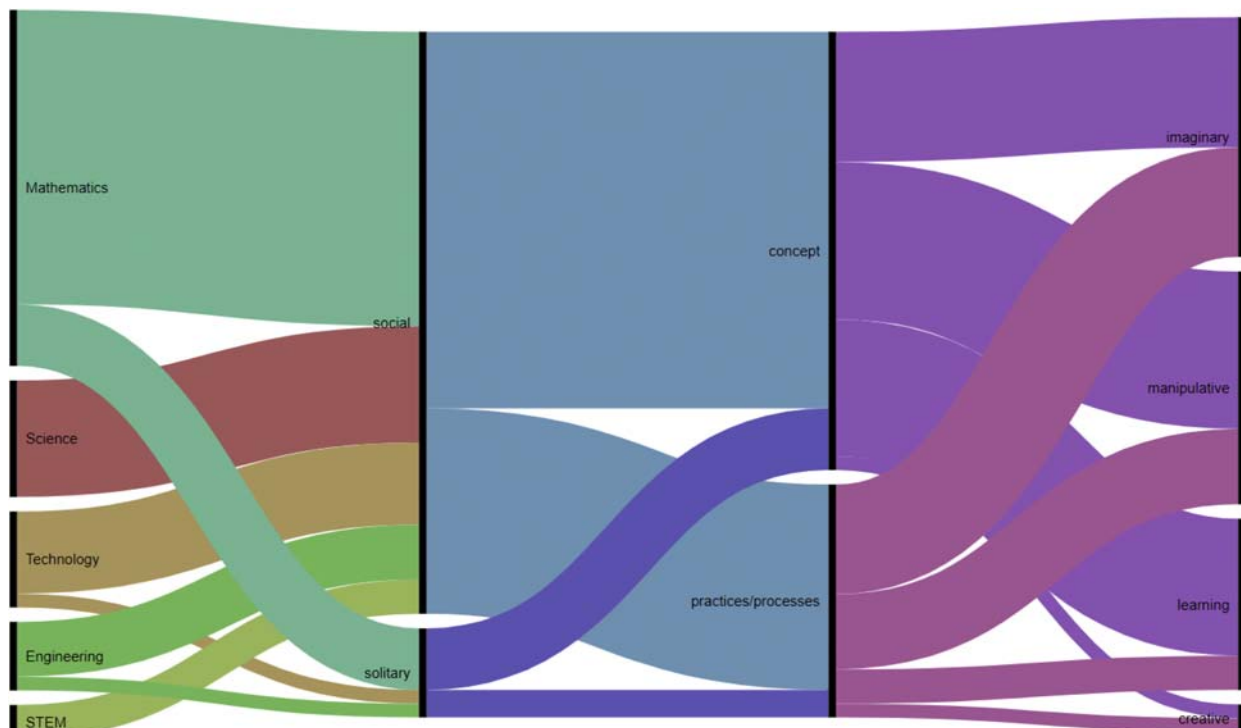


Fig. 2 Alluvial representation of the three dimensions of conceptualizations of play by STEM field.



Fig. 3 Dimensions for conceptualizing play, concepts and/or practices/processes and play perspectives, by STEM field.

and 110 preschoolers' knowledge of numeracy and geometry, as well as the preschoolers' spatial skills. Additionally, with the exception of creative play, the majority of studies in this literature review are conceptualizing play as imaginative (e.g., Bird, 2020), manipulative (e.g., Tian and Luo, 2019), and learning (e.g., Bachman and Dierking, 2011).

Fig. 3 highlights any differences by STEM field based on two of the dimensions for conceptualizing play, concepts and/or practices/processes and play perspectives, respectively. Research specific to the fields of science and mathematics more often viewed play in terms of conceptual development (e.g., Solis et al., 2017; Wolfgang et al., 2001) while research situated in the technology field viewed play in terms of the development of children's ways of talking, acting, and behaving like a STEM individual (i.e., practices/processes) (e.g., Lawrence, 2018). With respect to the other dimension, research in mathematics and engineering learning environments more often conceptualized play as manipulative, objects and materials that can be physically explored and handled (e.g., Bairaktarova et al., 2011; Trawick-Smith et al., 2017). On the other hand, research in science and technology learning environments included a higher percentage of studies that conceptualized play as imaginative (Enyedy et al., 2012; Han et al., 2015).

Fig. 4 displays how play was operationalized in research within STEM learning environments. The most common type of play was guided play (47%), social (28%) and environmental (19%), respectively. Therefore, a little less than half of the studies in this literature review are environments in which play is initiated and structured by an adult, but directed by the child (e.g., Andree and Lager-Nyqvist, 2013; Hassinger-Das et al., 2020; Howe et al., 2016). Free play accounted for 25% of the

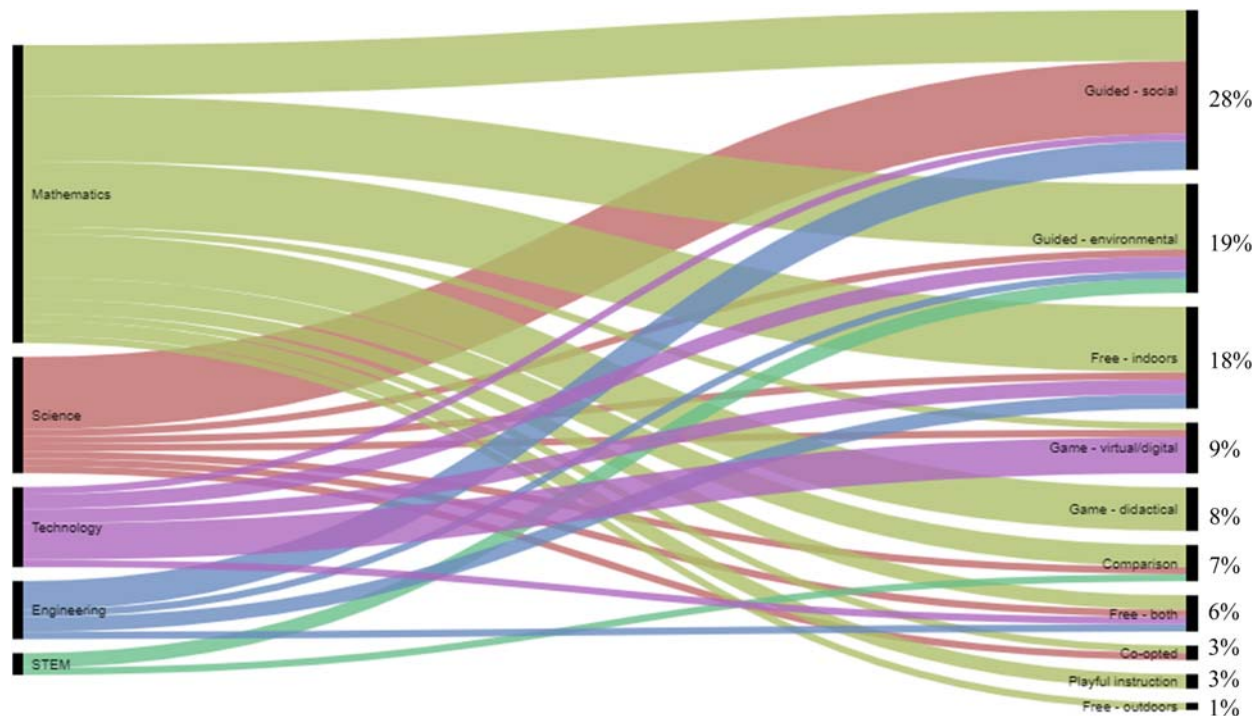


Fig. 4 Alluvial representation of types of play by STEM field. Percentages do not add up to 100% due to rounding to the nearest whole percentage.

studies (e.g., [Bagiati and Evangelou, 2016](#); [Reikerås, 2020](#)) and game play accounted for 13% (e.g., [Siyahhan et al., 2010](#); [Stebler et al., 2013](#)). In addition, there are a few things to note about types of play by STEM discipline. One, in considering the spread of articles in the science field, there is a higher concentrate that operationalize play through a guided play – social lens (e.g., [Carulla and Abdo, 2019](#)). Two, the field of mathematics is the only discipline that has researched play through games that do not include a technology component (i.e., didactical), but were more often board games (e.g., [Guberman et al., 1998](#)) and/or card games ([Vogt et al., 2018](#)).

Research question two

Free play

In free play STEM environments, children are unrestricted to explore the objects within their environment (e.g., blocks, mirror, plastic food) and engage in play either independently or with others such as their peers. Research characterized as free play has shown several positive outcomes. One, children experienced and began developing STEM related concepts such as numeracy (e.g., [Emfinger, 2009](#)), geometrical shapes (e.g., [Gejard and Melander, 2018](#)), spatial transformations (e.g., [Levine et al., 2012](#)), measurement concepts such as money and time (e.g., [Worthington and van Oers, 2016](#)), and physical concepts such as magnetism (e.g., [Solis et al., 2017](#)). Two, there is a positive and significant correlation between early block play, as well as LEGO play, and children's mathematical achievement at the middle and high school levels ([Wolfgang et al., 2001, 2003](#)). Three, free play afforded children the opportunity to explore physical and digital objects and make connections to their everyday experiences (e.g., [Worthington and van Oers, 2016](#)), as well as transform these objects to be closer to their own interests (e.g., [Kjällander and Moinian, 2014](#)). Lastly, children engaged in practices common to engineers and mathematicians such as problem solving, innovation, argumentation, and goal setting (e.g., [Bairaktarova et al., 2011](#); [Sumpter and Hedefalk, 2015](#)).

Beyond access to physical and digital material in the environment, researchers examined other components of a learning environment that had a positive outcome for children engaged in free play. In general, researchers noted a positive shift when children collaborated with one another (e.g., [Trawick-Smith et al., 2017](#); [Reikerås, 2020](#)), as well as when children were engaged in independent play ([Reikerås, 2020](#)). Additionally, [Gold et al. \(2015\)](#) found that different free play environments elicited differences in the ways that children participated as engineers. For example, children were found to use more technical language per hour in the dramatic play area than in a traditional playground or Imagination Playground™ environment.

However, some researchers that examined the impact of free play concluded that teachers are a vital part of the learning environment as they are able to bridge or connect children's play experiences and development to scientific ([Fleer, 2009](#)) and mathematical concepts ([Trawick-Smith et al., 2016](#)). Time engaged in play, or even the number of structures built, does not necessarily have an impact on children's cognitive development ([Trawick-Smith et al., 2016](#)) as other researchers have found (e.g., [Levine et al., 2012](#)). These conclusions highlight a tension in the educational field ([Hedges, 2014](#)).

Guided play

Guided play studies (adult-initiated and child-directed) in which changes were made to the physical environment were done so in four ways – signage or caregiver prompt cards (e.g., [Hassinger-Das et al., 2018](#)), access to and structure around materials, objects and models (e.g., [Simoncini et al., 2020](#)), stations or exhibits (e.g., [Reinking et al., 2017](#)), and less whole-group teacher instruction ([Andrée and Lager-Nyqvist, 2013](#)). The studies that provided signage and/or caregiver prompt cards, as well as stations or exhibits, was to increase caregiver-child STEM-related interactions in everyday contexts such as supermarkets, bus stops, playgrounds, and museums, and to indirectly impact child outcomes (e.g., [Hassinger-Das et al., 2018](#); [Reinking et al., 2017](#)). Researchers concluded that providing such support and opportunities for caregivers led to positive impact on children's spontaneous focus on number ([Braham et al., 2018](#)) and some math language (e.g., fractions, measurement, whole numbers, spatial; [Hassinger-Das et al., 2018, 2020](#)). Studies in which novel materials, objects and models were provided and supported through prompts (e.g., include a ramp in your design; [Schmitt et al., 2018](#)) and educator provisions (e.g., [Bird, 2020](#)), positive child outcomes included shape recognition ([Schmitt et al., 2018](#)), early arithmetic strategies ([Simoncini et al., 2020](#)), and an understanding of how materials function within a meaningful and personal context ([Parekh and Gee, 2018](#)). Lastly, the study conducted by [Andrée and Lager-Nyqvist \(2013\)](#) argued that when teacher instruction mirrors a more inquiry-based approach, students are able to make sense of science concepts – the variety of nutrients that humans need – through spontaneous play that some may view as off-task behavior.

Studies in which adults created and provided social support within the STEM learning environment were more often planned and embedded play-based learning activities in non-formal (e.g., [Leung et al., 2020](#)) and formal learning environments (e.g., [Enyedy et al., 2012](#); [Sliogeris and Almeida, 2019](#)), including formal environments that create out-of-school contexts (e.g., bakery; [Edo et al., 2009](#)) as opposed to more natural and everyday interactions and experiences with family members (e.g., [Sikder and Fler, 2015](#)) and peers ([Ramani et al., 2014](#)). For example, three preschool educators in a study by [Gomes and Fler \(2019\)](#) planned and implemented play-based activities around Jack and the Beanstalk. These activities were embedded in children's regular activities and routines, and educators engaged and guided children in science concepts such as sound, vibration, and growth of plants through imaginary play. Findings from these studies concluded that children developed an understanding of STEM concepts such as complex numbers, operations and arithmetical situations ([Edo et al., 2009](#)), force and motion concepts ([Enyedy et al., 2012](#)), atoms and molecules ([Carulla and Adbo, 2019](#)), and the relationship between animals, their habitats and needs ([Sliogeris and Almeida, 2019](#)). Additionally, children engaged in age-appropriate practices similar to scientists, engineers, and technologists/artists (e.g., [Fler, 2020](#); [Leung et al., 2020](#); [Simpson et al., 2020](#)). For example, [Adbo and Carulla \(2020\)](#) described how children became familiar with how to use scientific tools such as magnifying glasses and microscopes.

As such, supporting young children's development within institutional practices and expectations can be challenging for teachers as play is not viewed as a legitimate way to meet learning expectations (e.g., [Andree and Lager-Nyqvist, 2013](#); [Sliogeris and Almeida, 2019](#)). So, what did the researchers in this literature review highlight as effective pedagogical strategies for play-based (or guided-play) instruction in STEM learning environments? These instructional moves are conscious and intentional; they see opportunities during play and make use of them in an appropriate way (e.g., [Carulla and Adbo, 2019](#); [van Oers, 1996](#)). One, educators posed questions that encouraged young children to describe their observations, make connections to STEM concepts, and test their certainty about assertions and solutions (e.g., [Adbo and Carulla, 2019](#); [DeLiema et al., 2019](#); [van Oers, 1996](#)). Two, educators utilized young children's everyday experiences and concepts within play to bridge connections to STEM concepts (e.g., [Gomes and Fler, 2019](#); [Sliogeris and Almeida, 2019](#)). [Adbo and Carulla \(2020\)](#) referred to this as sustained shared thinking, which implies the social nature of the interaction between teachers and children. Three, educators introduced and used mental and physical models and tools to conceptually orient children's play (e.g., [Carulla and Adbo, 2019](#); [DeLiema et al., 2019](#)). Additional strategies included wondering discourse, creation of imaginary situations ([Fler, 2019](#)), and use of gestures to support toddlers ([Lee et al., 2015](#)).

Game play

Game play, regardless if virtual or didactic, seemed to initiate and engage children in spontaneous mathematical (e.g., comparison of different objects; [Avraamidou et al., 2012](#)) and scientific understandings (e.g., how particles move based on its form; [DeLiema et al., 2019](#)). For example, 6-year old children engaged in an adapted game called Shut the Box (i.e., sum of two dice to shut a wooden tile) and were observed using and expanding upon a variety of counting strategies (or none at all) based on their individual competency levels ([Stebler et al., 2013](#)). Children were also observed engaging in problem solving strategies ([Guberman et al., 1998](#)). These spontaneous moments were not only a function of the game's rules and artifacts, but also supported through interactions with other such as their peers (e.g., [Danby et al., 2018](#)) and parents ([Siyahhan et al., 2010](#)). As expressed by [Danby et al. \(2018\)](#), collaboration and social interactions between novice and expert children were necessary for the novice to gain access to and be successful within the game play virtual environment. Providing such play-based games may also have more of an influence on children's positive dispositions as they expressed a sense of enjoyment and fun, described as less school like ([Vogt et al., 2018](#)).

Other play types

In looking across play types that were less common, the research continued to highlight children's positive development. For example, the two studies that operationalized play as playful instruction observed an increase in a child's mathematical ability (e.g., conservation concepts of numbers, liquids, and lengths; [Watanabe, 2017](#)) when the adult, in this case a parent, initiated

and directed a math quiz game and Piaget tasks during conversations and play in home environments (Watanabe, 2017, 2019). The co-opted STEM play environments illustrated how adults engaged in joint creation and elaboration of a child's imagined or fantasy world(s) grounded in STEM topics may lead to an understanding of how numbers are sequenced and how they can be combined (Magnusson and Pramling, 2018) or learning about Earth and space (Hao and Fleer, 2017). As such, the adult supports the child's experience within their playful exploration and as-if world through becoming a part of that world as opposed to situating the child's interest and engagement in STEM through transitioning from an imaginary play environment to reality.

Lastly, the comparison studies showed differing results in terms of which play environment was more beneficial for children's cognitive development. Eason and Ramani (2020) found that parents and children in a didactic learning environment engaged in more math talk with one another than parents and children in a comparable guided play learning environment. However, parents in both the formal and guided environments posed about the same amount of questions that encouraged their child to explain their thinking (e.g., "why"). Parents and children in the free or unguided play environment was found to use the least amount of math talk. Conversely, Ferrera et al. (2011) noted how the guided play environment with blocks elicited more spatial language from parents and children than parents and children in the preassembled block or formal play environment. Both of these environments were found superior to the amount of spatial language spoken by parents and children in the free block play environment. Moreover, Ferrera et al. (2011) also noted that these three play contexts – free, guided, and preassembled – elicited more spatial language than other play contexts that do not involve construction.

Fisher et al. (2013) and Bulunz (2013) also found favor with guided play environments than didactic or formal environments, as well as free play environments. For example, children in guided play demonstrated a stronger conceptual understanding of shapes (Fisher et al., 2013) and science concepts (e.g., make predictions and explanations, build cause and effect relationships; Bulunz, 2013) than children in comparable play environments. These results raise the question, what is it about the free play experience that did not lead to more math talk or a stronger conceptual base than guided play or didactic instruction? As expressed, free play environments do not necessarily provide children with the supports to notice particular features or objects as they are irrelevant to the children's play context (e.g., Ferrera et al., 2011; Fisher et al., 2013).

Discussion

Descriptives from the included studies situate the majority of research on play in STEM learning environments within early childhood contexts (i.e., birth to age 8) across a variety of settings such as school, home, playgrounds, and museums. This raises the question, Why? Why is this case when the research included here showed positive outcomes in terms of children's STEM development? Why is this case when research has indicated a decline in factors that strongly influence persistence in STEM (e.g., interest; Maltese and Tai, 2011) around middle grades (ages 11–14) (Osborne et al., 2003)? We argue that additional research should consider the characteristics of the STEM play learning environment that is most conducive to learning and development for learners of all ages. How might play be conceptualized and operationalized similarly and/or differently from that in early childhood settings? As noted by Leontev (1981) play is intrinsically motivating and unique in that the motive of play "does not lie in its result but in the content of the action itself" (p. 368); thus, emphasizing the process rather than the product. Additionally, the majority of scholarship included in this literature review was in the field of mathematics and less often in engineering and technology. With the increase in the use of technology (e.g., robotics, coding apps, and virtual reality) and engineering education across PK-12, there is an opportunity to further our understanding of play within STEM fields. We also argue a need for research on play across the STEM fields as K-12 and postsecondary learners are not being prepared to enter the STEM workforce viewing STEM as a meta-discipline (Zollman, 2012).

Research question one

In looking across the STEM play learning environment scholarship in this literature review, the majority conceptualized play as a social as opposed to a solitary endeavor, which support young children as social actors with social competencies (Engdahl, 2011) and the development of communication and collaboration skills vital to the STEM workforce (e.g., Jang, 2016). Additionally, the studies rarely viewed play as creative, but more often as imaginary, manipulative, and learning. We argue a need for research utilizing creative play within STEM learning environments as thinking creatively while solving ill-defined problems, as well as an esthetic or artistic sensibility, is considered an important work activity for individuals who work in STEM disciplines (Jang, 2016; Root-Berstein, 2015). There were also differences in terms of STEM fields as mathematics and engineering were more likely to conceptualize play through use of manipulatives (e.g., blocks, puzzles) while technology were more likely to conceptualize play as imaginary (e.g., creating fantasy worlds). Research grounded in science and mathematics also emphasized conceptual development more often than the development of practices and process. We hypothesize that this may be due to the emphasis of these two subject areas in formal schooling contexts and teacher preparation programs, as well as the number of standardized tests as markers of "success" in mathematics and science fields (see Chandler et al., 2011 and Roehrig et al., 2012 for a similar argument).

Research question two

As a body of scholarship, the included studies in this literature highlight the value of all types of play for young children and adolescents' development within STEM across a range of learning environments. Children were provided opportunities to develop an understanding of a range of science and math concepts (e.g., DeLiema et al., 2019; Enyedy et al., 2012; Fisher et al., 2013; Gejard and Melander, 2018), use mathematical language (e.g., Ferrera et al., 2011; Hassinger-Das et al., 2020), and engage in practices common to STEM professionals (e.g., Simpson et al., 2020; Sumpter and Hedefalk, 2015). Thus, why do we see a decrease in play occurring both in and outside what can be considered schooling hours (e.g., Hofferth and Sandberg, 2009)? Why do we see a decrease in play when there is potential for long-term STEM impacts (Wolfgang et al., 2001, 2003)? How do we embed play within the demands of the academic standards for all learners and not only learners in early childhood settings? The findings from this analysis provides some insights into this question.

Across the play types, there are at least two commonalities that we can utilize when developing play in STEM learning environments for all learners. One is access to objects and materials for their own right (e.g., blocks, microscopes) and for their potential to transform reality into an imaginary situation (e.g., plastic banana becomes a phone; Bird, 2020). We highlight access as Simoncini et al. (2020) described how access to resources such as Duplo Lego is a barrier in some countries. But we also see potential in transforming household objects and recyclable materials (e.g., cardboard box), as well as items in nature (e.g., tree branch), into opportunities for play with STEM learning environments (e.g., Parekh and Gee, 2018). Two is the connection to learners' everyday experiences, everyday experiences that lead to the development of STEM concepts (e.g., Hao and Fleer, 2017; Vogel, 2013). However, as argued by some, the teacher is a vital part of making these connections between children's everyday experiences and understanding of concepts (e.g., Gomes and Fleer, 2019; Sliogeris and Almeida, 2019). An approach to this may be grounded in the funds of knowledge framework, a framework for conceptualizing students' homes, cultural, and everyday experiences as sources of knowledge to use in the learning environment (González et al., 2005). Utilization of this knowledge creates a "third space" or an "integration of knowledges and Discourses drawn from different spaces ... that merges the "first space" of people's home, community, and peer networks with the "second space" of the Discourses they encounter in more formalized institutions such as work, school, or church" (Moje et al., 2004, p. 41). We contend that this extends to STEM learning environments whether non-formal, informal, or formal setting and with caregivers as they have a foundational knowledge of their child's everyday experiences and have the ability to build upon and make connections between these experiences and STEM concepts, practices and processes within play.

As alluded to above, the social nature of play is an important component of the STEM learning environment, but there seems to be a difference in outcomes based on the role and interaction with others – educators, caregivers, siblings, and/or peers. The comparative studies included in this literature review highlight this as the majority confirmed that guided play (adult-initiated, child-directed) led to more mathematical language from parents and children, for example, than mathematical language from parents and children in formal play and free play STEM learning environments (Ferrera et al., 2011). It is argued that free play environments do not afford children with the necessary supports to make connections to STEM concepts, practices and processes (e.g., Fisher et al., 2013). We also see this argument by others scholars such as Fleer (2009), Trawick-Smith et al. (2016), van Oers (1996), Bulunz (2013), and Abdo and Carulla (2020) to name a few. We argue that additional research be conducted that compares developmental outcomes within different STEM play learning environments, including co-opted play environments as the two studies in this literature review illustrated how the adult entered the child's playful world to co-create mathematical and scientific concepts specific to the child's interest (Hao and Fleer, 2017; Magnusson and Pramling, 2018). Additionally, as these studies argued, the instructional practices here are conscious and intentional (e.g., van Oers, 1996) as there is a fine line between supporting the child's play and directing the child's play to an experience that does not resemble play. Simpson et al. (2020) referred to this as missed opportunities, opportunities to engage children in further play around STEM concepts, practices and processes mirror that of instruction and not that of play.

Conclusion

The goal of this systematic literature review was to examine play in STEM learning environments as a means to consider play as a potential for addressing global, national, and societal issues in STEM (e.g., Kuenzi, 2008; van den Hurk et al., 2019) and conversely, how STEM learning environments can be framed as a context for human (not only child) development through play (e.g., Bassok et al., 2016; Coople and Bredekamp, 2009). The positive children and adolescent outcomes outlined in this study support the notion of play in STEM learning environments as establishing the seeds or foundation to address concerns around economic and national security, a diverse STEM workforce, and a STEM-literate society; yet, our understanding does not extend beyond the few studies that included a population beyond the age of 14. Reframing STEM research through a play lens with a wider age range of learners may shed more light on this, including individuals in the STEM workforce and their opportunities to "play" with ideas and objects. This literature review also highlighted how STEM play learning environments can be utilized to alleviate some of the anxiety and pressures around achievement and academic standards of educators and caregivers (Bassok et al., 2016; Miller and Almon, 2009) as young children and adolescents were gaining an understanding of STEM concepts, practices and processes through play.

References

- Adbo, K., Carulla, C.V., 2019. Designing play-based learning chemistry activities in the preschool environment. *Chem. Educ. Res. Pract.* 20 (3), 542–553. <https://doi.org/10.1039/c8rp00306h>.
- Adbo, K., Carulla, C.V., 2020. Learning about science in preschool: play-based activities to support children's understanding of chemistry concepts. *Int. J. Early Child.* 52, 17–35. <https://doi.org/10.1007/s13158-020-259-3>.
- Andrée, M., Lager-Nyqvist, L., 2013. Spontaneous play and imagination in everyday science classroom practice. *Res. Sci. Educ.* 43 (5), 1735–1750. <https://doi.org/10.1007/s11165-012-9333-y>.
- Appelbaum, M., Cooper, H., Kline, R.B., Mayo-Wilson, E., Nezu, A.M., Rao, S.M., 2018. Journal article reporting standards for quantitative research in psychology: the APA publications and communications board task force report. *Am. Psychol.* 73 (1), 3–25. <https://doi.org/10.1037/amp0000191>.
- Avraamidou, A., Monaghan, J., Walker, A., 2012. Abstraction through game play. *Technol. Knowl. Learn.* 17 (1–2), 1–21. <https://doi.org/10.1007/s10758-012-9189-2>.
- Bachman, J., Dierking, L., 2011. Co-creating playful environments that support children's science and mathematics learning as cultural activity: insights from home-educating families. *Child. Youth Environ.* 21 (2), 294–311.
- Bagiati, A., Evangelou, D., 2016. Practicing engineering while building with blocks: identifying engineering thinking. *Eur. Early Child. Educ. Res. J.* 24 (1), 67–85. <https://doi.org/10.1080/1350293X.2015.1120521>.
- Bairaktarova, D., Evangelou, D., Bagiati, A., Brophy, S., 2011. Early engineering in young children's exploratory play with tangible materials. *Child. Youth Environ.* 21 (2), 212–235.
- Bassok, D., Latham, S., Rorem, A., 2016. Is kindergarten the new first grade? *AERA Open* 2 (1), 1–31. <https://doi.org/10.1177/2332858415616358>.
- Beckwith, J., Huang, F., 2005. Should we make a fuss? A case for social responsibility in science. *Nat. Biotechnol.* 23 (12), 1479–1480.
- Bird, J., 2020. "You need a phone and camera in your bag before you go out": children's play with imaginative technologies. *Br. J. Educ. Technol.* 51 (1), 166–176. <https://doi.org/10.1111/bjet.12791>.
- Braham, E.J., Libertus, M.E., McCrink, K., 2018. Children's spontaneous focus on number before and after guided parent-child interactions in a children's museum. *Dev. Psychol.* 54 (8), 1492–1498. <https://doi.org/10.1037/dev0000534>.
- Bronson, M.B., 1995. *The Right Stuff for Children Birth to Eight: Selecting Play Materials to Support Development*. National Association for the Education of Young Children, Washington D.C.
- Brown, S., Vaughan, C.C., 2009. *Play: How It Shapes the Brain, Opens the Imagination, and Invigorates the Soul*. Avery, New York.
- Bulunuz, M., 2013. Teaching science through play in kindergarten: does integrated play and science instruction build understanding? *Eur. Early Child. Educ. Res. J.* 21 (2), 226–249. <https://doi.org/10.1080/1350293X.2013.789195>.
- Burstein, B., Agostino, H., Greenfield, B., 2019. Suicidal attempts and ideation among children and adolescents in U.S. Emergency departments, 2007–2015. *JAMA Pediatr.* 173 (6), 598–600. <https://doi.org/10.1001/jamapediatrics.2019.0464>.
- Bybee, R.W., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teacher* 70, 30–35.
- Carlsson-Paige, N., McLaughlin, G.B., Almon, J.W., 2015. *Reading Instruction in Kindergarten: Little to Gain and Much to Lose*. Retrieved from. <https://allianceforchildhood.org/publications-and-reports>.
- Carulla, C.V., Adbo, K., 2019. Using cultural-historical theory to design and assess a chemistry play-based learning intervention. *Cult. Hist. Psychol.* 15 (4), 35–43. <https://doi.org/10.17759/chp.2019150404>.
- Chandler, J., Fontenot, A.D., Tate, D., 2011. Problems associated with a lack of cohesive policy in K-12 pre-college engineering. *J. Pre-Coll. Eng. Edu. Res.* 1 (1), 40–48. <https://doi.org/10.7771/2157-9288.1029>.
- Copple, C., Bredekamp, S., 2009. *Developmentally Appropriate Practice in Early Childhood Programs Serving Children From Birth Through Age 8*. National Association for the Education of Young Children, Washington, DC.
- Danby, S., Evaldsson, A.C., Melander, H., Aarsand, P., 2018. Situated collaboration and problem solving in young children's digital gameplay. *Br. J. Educ. Technol.* 49 (5), 959–972. <https://doi.org/10.1111/bjet.12636>.
- de Castell, S., Jensen, J., Thumler, K., 2014. From simulation to imitation: controllers, corporeality, and mimetic play. *Simulat. Gaming* 45 (3), 332–355. <https://doi.org/10.1177/1046878114542316>.
- DeLiema, D., Enyedy, N., Danish, J.A., 2019. Roles, rules, and keys: how different play configurations shape collaborative science inquiry. *J. Learn. Sci.* 28 (4–5), 513–555. <https://doi.org/10.1080/10508406.2019.1675071>.
- Doyle, L.H., 2003. Synthesis through meta-ethnography: paradoxes, enhancements, and possibilities. *Qual. Res.* 3 (3), 321–344.
- Eason, S.H., Ramani, G.B., 2020. Parent-child math talk about fractions during formal learning and guided play activities. *Child Dev.* 91 (2), 546–562. <https://doi.org/10.1111/cdev.13199>.
- Edens, K.M., Potter, E.F., 2013. An exploratory look at the relationships among math skills, motivational factors and activity choice. *Early Child. Educ. J.* 41 (3), 235–243. <https://doi.org/10.1007/s10643-012-0540-y>.
- Edo, M., Planas, N., Badillo, E., 2009. Mathematical learning in a context of play. *Eur. Early Child. Educ. Res. J.* 17 (3), 325–341. <https://doi.org/10.1080/13502930903101537>.
- Elkind, D., 2007. *The Power of Play: How Spontaneous, Imaginative Activities Lead to Happier, Healthier Children*. Da Capo Lifelong Books, Boston, MA.
- Emfinger, K., 2009. Numerical conceptions reflected during multiage child-initiated pretend play. *J. Instr. Psychol.* 36 (4), 326–334.
- Engdahl, I., 2011. Toddler interaction during play in the Swedish pre-school. *Early Child. Dev. Care* 181 (10), 1421–1439.
- Enyedy, N., Danish, J.A., Delacruz, G., Kumar, M., 2012. Learning physics through play in an augmented reality environment. *Int. J. Comp. Support. Collabor. Learn.* 7 (3), 347–378. <https://doi.org/10.1007/s11412-012-9150-3>.
- Eshach, H., 2007. Bridging in-school and out-of-school learning: formal, non-formal, and informal education. *J. Sci. Educ. Technol.* 16 (2), 171–190. <https://doi.org/10.1007/s10956-006-9027-1>.
- Falk, J.H., Dierking, L.D., 2000. *Learning From Museums: Visitors Experiences and Their Making of Meaning*. Altamira Press, Walnut Creek, CA.
- Ferrara, K., Hirsh-Pasek, K., Newcombe, N.S., Golinkoff, R.M., Lam, W.S., 2011. Block talk: spatial language during block play. *Mind Brain Educ.* 5 (3), 143–151.
- Fisher, K.R., Hirsh-Pasek, K., Newcombe, N., Golinkoff, R.M., 2013. Taking shape: supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Dev.* 84 (6), 1872–1878. <https://doi.org/10.1111/cdev.12091>.
- Fleer, M., 2009. Understanding the dialectical relations between everyday concepts and scientific concepts within play-based programs. *Res. Sci. Educ.* 39 (2), 281–306. <https://doi.org/10.1007/s11165-008-9085-x>.
- Fleer, M., 2019. Scientific playworlds: a model of teaching science in play-based settings. *Res. Sci. Edu.* 49 (5), 1257–1278. <https://doi.org/10.1007/s11165-017-9653-z>.
- Fleer, M., 2020. Studying the relations between motives and motivation—How young children develop a motive orientation for collective engineering play. *Learn. Cul. Soc. Inter.* 24, 100355. <https://doi.org/10.1016/j.lcsi.2019.100355>.
- Fryar, C.D., Carroll, M.D., Ogden, C.L., 2014. Prevalence of Overweight and Obesity Among Children and Adolescents: United States, 1963–1965 Through 2011–2012. *Health E-Stats*. Retrieved December 21, 2017, from. https://www.cdc.gov/nchs/data/hestat/obesity_child_11_12/obesity_child_11_12.htm.
- Garibay, J.C., 2015. STEM students' social agency and views on working for social change: are STEM disciplines developing socially and civically responsible students? *J. Res. Sci. Teach.* 52 (5), 610–632. <https://doi.org/10.1002/tea.21203>.
- Gejard, G., Melander, H., 2018. Mathematizing in preschool: children's participation in geometrical discourse. *Eur. Early Child. Educ. Res. J.* 26 (4), 495–511. <https://doi.org/10.1080/1350293X.2018.1487143>.

- Ginsburg, H.P., 2006. Mathematical play and playful mathematics: a guide for early education. In: Singer, D.G., Golinkoff, R.M., Hirsh-Pasek, K. (Eds.), *Play = Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*. Oxford University Press, New York, pp. 145–165.
- Glancy, A.W., Moore, T.J., 2013. Theoretical foundations for effective STEM learning environments. In: *School of Engineering Education Working Papers*. Retrieved from: <http://docs.lib.purdue.edu/enewp/1>.
- Gold, Z.S., Elicker, J., Choi, J.Y., Anderson, T., Brophy, S.P., 2015. Preschoolers' engineering play behaviors: differences in gender and play context. *Child. Youth Environ.* 25 (3), 1–21.
- Gold, Z.S., Elicker, J., Kellerman, A.M., Christ, S., Mishra, A.A., Howe, N., 2020. Engineering play, mathematics, and spatial skills in children with and without disabilities. *Early Educ. Dev.* <https://doi.org/10.1080/10409289.2019.1709382>. Advanced online publication.
- Gomes, J., Fleer, M., 2019. The development of a scientific motive: how preschool science and home play reciprocally contribute to science learning. *Res. Sci. Educ.* 49 (2), 613–634. <https://doi.org/10.1007/s11165-017-9631-5>.
- González, N., Moll, L.C., Amanti, C. (Eds.), 2005. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Routledge.
- Gray, P., 2013. *Free to Learn: Why Unleashing the Instinct to Play Will Make Our Children Happier, More Self-Reliant, and Better Students for Life*. Basic Books, New York, NY.
- Guberman, S.R., Rahm, J., Menk, D.W., 1998. Transforming cultural practices: illustrations from children's game play. *Anthropol. Educ. Q.* 29 (4), 419–445.
- Hallström, J., Elvstrand, H., Hellberg, K., 2015. Gender and technology in free play in Swedish early childhood education. *Int. J. Technol. Des. Educ.* 25 (2), 137–149. <https://doi.org/10.1007/s10798-014-9274-z>.
- Han, J., Jo, M., Hyun, E., So, H.J., 2015. Examining young children's perception toward augmented reality-infused dramatic play. *Educ. Technol. Res. Dev.* 63 (3), 455–474. <https://doi.org/10.1007/s11423-015-9374-9>.
- Hanline, M.F., Milton, S., Phelps, P.C., 2010. The relationship between preschool block play and reading and maths abilities in early elementary school: a longitudinal study of children with and without disabilities. *Early Child. Dev. Care* 180 (8), 1005–1017. <https://doi.org/10.1080/03004430802671171>.
- Hännikäinen, M., Munter, H., 2018. Toddlers' play in early childhood education settings. In: Smith, P.K., Roopnarine, J.L. (Eds.), *The Cambridge Handbook of Play: Developmental and Disciplinary Perspectives*. Cambridge University Press, Cambridge, pp. 491–510.
- Hao, Y., Fleer, M., 2017. Collective interpretations of early science learning about earth and space: a cultural-historical study of family settings for scientific imagination. *Pedagogies* 12 (4), 354–373. <https://doi.org/10.1080/1554480X.2017.1375411>.
- Hassinger-Das, B., Bustamante, A.S., Hirsh-Pasek, K., Golinkoff, R.M., 2018. Learning landscapes: playing the way to learning and engagement in public spaces. *Educ. Sci.* 8 (2), 74–95. <https://doi.org/10.3390/educsci8020074>.
- Hassinger-Das, B., Palti, I., Golinkoff, R.M., Hirsh-Pasek, K., 2020. Urban thinkscape: infusing public spaces with STEM conversation and interaction opportunities. *J. Cogn. Dev.* 21 (1), 125–147. <https://doi.org/10.1080/15248372.2019.1673753>.
- Hatisaru, V., Beswick, K., Fraser, S., 2019. STEM learning environments: perceptions of STEM education researchers. In: Hine, G., Blackley, S., Cooke, A. (Eds.), *Proceedings of the 42nd Annual Conference of the Mathematics Education Research Group of Australasia. MERGA, Perth*, pp. 340–347.
- Hedges, H., 2014. Children's content learning in play provision: competing tensions and future possibilities. In: Brooker, L., Blaise, M., Edwards, S. (Eds.), *The Sage Handbook of Play and Learning in Early Childhood*. Sage Publications, London, pp. 192–203.
- Hofferth, S.L., Sandberg, J.F., 2009. How American children spend their time. *J. Marriage Fam.* 63, 295–308. <https://doi.org/10.1111/j.17413737.2001.00295.x>.
- Hoskins, K., Smedley, S., 2019. Protecting and extending Froebelian principles in practice: exploring the importance of learning through play. *J. Early Child. Res.* 17 (2), 73–87. <https://doi.org/10.1177/1476718X18809114>.
- Howe, N., Adrien, E., Della Porta, S., Peccia, S., Recchia, H., Osana, H.P., Ross, H., 2016. 'Infinity means it goes on forever': siblings' informal teaching of mathematics. *Infant Child Dev.* 25 (2), 137–157. <https://doi.org/10.1002/icd.1928>.
- Jang, H., 2016. Identifying 21st century STEM competencies using workplace data. *J. Sci. Educ. Technol.* 25 (2), 284–301. <https://doi.org/10.1007/s10956-015-9593-1>.
- Kjällander, S., Moilanen, F., 2014. Digital tablets and applications in preschool-preschoolers' creative transformation of didactic design. *Des. Learn.* 7 (1), 10–33. <https://doi.org/10.2478/dlf-2014-0009>.
- Kuenzi, J.J., 2008. *Science, Technology, Engineering, and Mathematics (STEM) Education: Background, Federal Policy, and Legislative Action*. Congressional Research Service, Washington, DC.
- Lawrence, S.M., 2018. Preschool children and iPads: observations of social interactions during digital play. *Early Educ. Dev.* 29 (2), 207–228. <https://doi.org/10.1080/10409289.2017.1379303>.
- Lee, J., Kotsopoulos, D., Tumber, A., Makosz, S., 2015. Gesturing about number sense. *J. Early Child. Res.* 13 (3), 263–279. <https://doi.org/10.1177/1476718X13510914>.
- Leont'ev, A.N., 1981. *Problems of the Development of the Mind*. Progress, Moscow.
- Leung, S.K., Choi, K.W., Yuen, M., 2020. Video art as digital play for young children. *Br. J. Educ. Technol.* 51 (2), 531–554. <https://doi.org/10.1111/bjet.12877>.
- Levine, S.C., Ratliff, K.R., Huttenlocher, J., Cannon, J., 2012. Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Dev. Psychol.* 48 (2), 530–542. <https://doi.org/10.1037/a0025913>.
- Levitt, H.M., Bamberg, M., Creswell, J.W., Frost, D.M., Josselson, R., Suárez-Orozco, C., 2018. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: the APA publications and communications board task force report. *Am. Psychol.* 73 (1), 26–46. <https://doi.org/10.1037/amp0000151>.
- Louv, R., 2005. *Last Child in the Woods: Saving Our Children from Nature-Deficit Disorder*. Chapel Hill. Algonquin Books of Chapel Hill, NC.
- Maass, K., Geiger, V., Ariza, M.R., Goos, M., 2019. The role of mathematics in interdisciplinary STEM education. *ZDM* 51, 869–884. <https://doi.org/10.1007/s11858-019-01100-5>.
- Magnusson, M., Pramling, N., 2018. In 'Numberland': play-based pedagogy in response to imaginative numeracy. *Int. J. Early Years Educ.* 26 (1), 24–41. <https://doi.org/10.1080/09669760.2017.1368369>.
- Maltese, A.V., Tai, R.H., 2011. Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among U.S. students. *Sci. Educ.* 95 (5), 877–907.
- McDonald, C.V., 2016. STEM education: a review of the contribution of the disciplines of science, technology, engineering and mathematics. *Sci. Educ. Int.* 27 (4), 530–569.
- Miller, E., Almon, J., 2009. *Crisis in the Kindergarten: Why Children Need to Play in School*. Alliance for Childhood, College Park, MD.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6 (7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Moje, E.B., Ciechanowski, K.M., Kramer, K., Ellis, L., Carrillo, R., Collazo, T., 2004. Working toward third space in content area literacy: an examination of everyday funds of knowledge and discourse. *Read. Res. Q.* 39 (1), 38–70.
- National Academy of Sciences (NAS), 2011. Expanding underrepresented minority participation: America's science and technology talent at the crossroads. In: *Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline*. The National Academies Press, Washington, DC.
- Nel, A., Joubert, I., Hartell, C., 2017. Teachers' perceptions on the design and use of an outdoor learning environment for sensory and motor stimulation. *S. Afr. J. Chem. Eng.* 7 (1), 1–11.
- Organisation for Economic Cooperation and Development (OECD), 2016. *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD Publishing, Paris.
- Osborne, J., Simon, S., Collins, S., 2003. Attitudes towards science: a review of the literature and its implications. *Int. J. Sci. Educ.* 25, 1049–1079. <https://doi.org/10.1080/0950069032000032199>.
- Parekh, P., Gee, E.R., 2018. Zooming into a tinkering project: the progression of learning through transitional objects. *Interdisc. J. Prob. Based Learn.* 12 (2) <https://doi.org/10.7771/1541-5015.1751>.

- Petticrew, M., Egan, M., Thomson, H., Hamilton, V., Kunkler, R., Roberts, H., 2008. Publication bias in qualitative research: what becomes of qualitative research presented at conferences? *J. Epidemiol. Comm. Health* 62 (6), 552–554. <https://doi.org/10.1136/jech.2006.059394>.
- Piaget, J., 1952. *Plays, Dreams, Imitations*. WW Norton.
- Radovic, D., Black, L., Williams, J., Salas, C.E., 2018. Towards conceptual coherence in the research on mathematics learner identity: a systematic review of the literature. *Educ. Stud. Math.* 99 (1), 21–42. <https://doi.org/10.1007/s10649-018-9819-2>.
- Ramani, G.B., Zippert, E., Schweitzer, S., Pan, S., 2014. Preschool children's joint block building during a guided play activity. *J. Appl. Dev. Psychol.* 35 (4), 326–336. <https://doi.org/10.1016/j.appdev.2014.05.005>.
- Reikerås, E., 2020. Relations between play skills and mathematical skills in toddlers. *ZDM*. <https://doi.org/10.1007/s11858-020-01141-1>. Advanced online publication.
- Reinking, A.K., Vetere III, M.J., Percell, J.C., 2017. Collaborating with theatre, nature, and STEM: a multigenerational family event. *New Waves Educ. Res. Dev.* 20 (1), 23–37.
- Roehrig, G.H., Moore, T.J., Wang, H.H., Park, M.S., 2012. Is adding the E enough? Investigating the impact of K-12 engineering standards on the implementation of STEM integration. *Sch. Sci. Math.* 112 (1), 31–44.
- Root-Bernstein, R., 2015. Arts and crafts as adjuncts to STEM education to foster creativity in gifted and talented students. *Asia Pac. Educ. Rev.* 16 (2), 203–212. <https://doi.org/10.1007/s12564-015-9362-0>.
- Savin-Baden, M., 2010. Changelings and shape shifters? Identity play and pedagogical positioning of staff in immersive virtual worlds. *Lond. Rev. Educ.* 8 (1), 25–38.
- Schmitt, S.A., Korucu, I., Napoli, A.R., Bryant, L.M., Purpura, D.J., 2018. Using block play to enhance preschool children's mathematics and executive functioning: a randomized controlled trial. *Early Child. Res. Q.* 44, 181–191. <https://doi.org/10.1016/j.ecresq.2018.04.006>.
- Sikder, S., Fleer, M., 2015. *Small Science*: infants and toddlers experiencing science in everyday family life. *Res. Sci. Educ.* 45 (3), 445–464. <https://doi.org/10.1007/s11165-014-9431-0>.
- Simoncini, K., Forndran, A., Manson, E., Sawi, J., Philip, M., Kokinai, C., 2020. The impact of block play on children's early mathematics skills in rural Papua New Guinea. *Int. J. Early Child.* 52, 1–17. <https://doi.org/10.1007/s13158-020-00261-9>.
- Simpson, A., Burris, A., Maltese, A., 2020. Youth's engagement as scientists and engineers in an afterschool making and tinkering program. *Res. Sci. Educ.* 50, 1–22. <https://doi.org/10.1007/s11165-017-9678-3>.
- Siyahhan, S., Barab, S.A., Downton, M.P., 2010. Using activity theory to understand intergenerational play: the case of Family Quest. *Int. J. Comp. Support. Collabor. Learn.* 5 (4), 415–432. <https://doi.org/10.1007/s11412-010-9097-1>.
- Sligeris, M., Almeida, S.C., 2019. Young children's development of scientific knowledge through the combination of teacher-guided play and child-guided play. *Res. Sci. Educ.* 49 (6), 1569–1593. <https://doi.org/10.1007/s11165-017-9667-6>.
- Solis, S.L., Curtis, K.N., Hayes-Messinger, A., 2017. Children's exploration of physical phenomena during object play. *J. Res. Child. Educ.* 31 (1), 122–140. <https://doi.org/10.1080/02568543.2016.1244583>.
- Stebler, R., Vogt, F., Wolf, I., Hauser, B., Rechsteiner, K., 2013. Play-based mathematics in kindergarten. A video analysis of children's mathematical behaviour while playing a board game in small groups. *J. für Mathematik-Didaktik* 34 (2), 149–175. <https://doi.org/10.1007/s13138-013-0051-4>.
- Struyf, A., De Loof, H., Boeve-de Pauw, J., Van Petegem, P., 2019. Students' engagement in different STEM learning environments: integrated STEM education as promising practice? *Int. J. Sci. Educ.* 41 (10), 1387–1407.
- Sumpter, L., Hedefalk, M., 2015. Preschool children's collective mathematical reasoning during free outdoor play. *J. Math. Behav.* 39, 1–10. <https://doi.org/10.1016/j.jmathb.2015.03.006>.
- Tian, M., Luo, T., 2019. The development of 3-to 6-year-olds' symbolic representation ability in a block construction activity. *Early Child. Dev. Care* 1–19. <https://doi.org/10.1080/03004430.2019.1656619>.
- Trawick-Smith, J., Swaminathan, S., Liu, X., 2016. The relationship of teacher–child play interactions to mathematics learning in preschool. *Early Child. Dev. Care* 186 (5), 716–733. <https://doi.org/10.1080/03004430.2015.1054818>.
- Trawick-Smith, J., Swaminathan, S., Bato, B., Danieluk, C., Marsh, S., Szarwacki, M., 2017. Block play and mathematics learning in preschool: the effects of building complexity, peer and teacher interactions in the block area, and replica play materials. *J. Early Child. Res.* 15 (4), 433–448. <https://doi.org/10.1177/1476718X16664557>.
- van den Hurk, A., Meelissen, M., van Langen, A., 2019. Interventions in education to prevent STEM pipeline leakage. *Int. J. Sci. Educ.* 41 (2), 150–164. <https://doi.org/10.1080/09500693.2018.1540897>.
- van Oers, B., 1996. Are you sure? Stimulating mathematical thinking during young children's play. *Eur. Early Child. Educ. Res. J.* 4 (1), 71–87.
- Verdine, B.N., Golinkoff, R.M., Hirsh-Pasek, K., Newcombe, N.S., 2017. Spatial skills, their development, and their links to mathematics. *Monographs of the Society for Research in Child Development* 82, 7–30. <https://doi.org/10.1111/mono.12280>.
- Vogel, R., 2013. Mathematical situations of play and exploration. *Educ. Stud. Math.* 84 (2), 209–225. <https://doi.org/10.1007/s10649-013-9504-4>.
- Vogt, F., Hauser, B., Stebler, R., Rechsteiner, K., Urech, C., 2018. Learning through play–pedagogy and learning outcomes in early childhood mathematics. *Eur. Early Child. Educ. Res. J.* 26 (4), 589–603. <https://doi.org/10.1080/1350293X.2018.1487160>.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Vygotsky, L.S., 2016. Play and its role in the mental development of the child. *Int. Res. Early Child. Edu.* 7 (2), 3–25.
- Watanabe, N., 2017. Acquiring Piaget's conservation concept of numbers, lengths, and liquids as ordinary play. *J. Educ. Dev. Psychol.* 7 (1), 210–217. <https://doi.org/10.5539/jedp.v7n1p210>.
- Watanabe, N., 2019. Effective simple mathematics play at home in early childhood: promoting both non-cognitive and cognitive skills in early childhood. *Int. Electron. J. Math. Educ.* 14 (2), 401–417. <https://doi.org/10.29333/iejme/5739>.
- Watson, K., Froyd, J., 2007. Diversifying the US engineering workforce: a new model. *J. Eng. Educ.* 96 (1), 19–32.
- Wolfgang, C.H., Stannard, L.L., Jones, I., 2001. Block play performance among preschoolers as a predictor of later school achievement in mathematics. *J. Res. Child. Educ.* 15 (2), 173–180.
- Wolfgang, C., Stannard, L., Jones, I., 2003. Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child. Dev. Care* 173 (5), 467–475. <https://doi.org/10.1080/0300443032000088212>.
- Worthington, M., van Oers, B., 2016. Pretend play and the cultural foundations of mathematics. *Eur. Early Child. Educ. Res. J.* 24 (1), 51–66. <https://doi.org/10.1080/1350293X.2015.1120520>.
- Zollman, A., 2012. Learning for STEM literacy: STEM literacy for learning. *Sch. Sci. Math.* 112 (1), 12–19.
- Zosh, J.M., Hirsh-Pasek, K., Hopkins, E.J., Jensen, H., Liu, C., Neale, D., Solis, S.L., Whitebread, D., 2018. Accessing the inaccessible: redefining play as a spectrum. *Front. Psychol.* 9, 1124. <https://doi.org/10.3389/fpsyg.2018.01124>.

Relevant website

Additional information (e.g., context, age of participants) from each included study can be found online at https://orb.binghamton.edu/mathematics_fac/10.

STEM schools: their rise, present state, and future potential

Jennifer N. Tripp, University at Buffalo, SUNY, Department of Learning and Instruction, Buffalo, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	177
Rationale for the rise of STEM education	178
Varieties of STEM schools	178
Selective STEM schools	179
Inclusive STEM schools	179
STEM Career and Technical Education (CTE) schools	180
STEM school models	180
Characteristics of STEM schools	181
Critical components explained	182
Review of the STEM school literature	182
Selective STEM school outcomes	182
Inclusive STEM school outcomes	182
Rigorous course-taking patterns and experiences	182
Achievement	183
STEM interests and aspirations	183
Racial/ethnic, economic, and gender disparities	184
Summary of findings	184
Going forward: a vision for STEM schools of tomorrow	185
Equity considerations	185
References	186

Introduction

STEM is ubiquitous. Toys, coloring books, and magazines for children. After school programs, summer camps, and podcasts. Teachers and professional development programs. Courses, internships, and schools. International journals, special issues, and conferences. Occasionally there is a merging of the arts and design (STEAM or STEM-D), reading or religion (STREAM), or medicine (STEMM). Nevertheless, STEM has remained most popular and become a favored marketing maneuver for products, events, practices, programs, and policies.

Indeed, STEM has proliferated in popularity and become a natural part of the vernacular since the acronym was coined at the National Science Foundation (NSF) in the 1990s as a way to group science, technology, engineering, and mathematics (Bybee, 2013). Yet, the meaning of the term, STEM, remains an enigma and varies based on individual interpretation; from teachers to administrators, policymakers, and the general public, a universally accepted definition does not exist. However, an integrated definition of STEM, in which two or more STEM domains are “bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning” (Kelley and Knowles, 2016, p. 3), has been more widely embraced (Koehler et al., 2015; Vaval et al., 2019).

This integrative nature of STEM is recognized by the *Next Generation Science Standards* (NGSS), from unifying crosscutting concepts to math and literacy connections:

What is different in Next Generation Science Standards (NGSS) is a commitment to fully integrate engineering design, technology, and mathematics into the structure of science education by raising engineering design to the same level as scientific inquiry when teaching science disciplines at all levels from kindergarten to grade 12. This new integrated approach to science education is sometimes referred to by the acronym STEM.

NGSS Lead States (2013).

In effect, STEM is a recognition that the disciplines are not siloed. Everyday challenges are “multidisciplinary, and many require integration of multiple STEM concepts to solve them” (Roehrig et al., 2012, p. 31).

Before a thorough consideration of what STEM is and should be, particularly in practice, there has been an international race to jump on the STEM bandwagon (Lesseig et al., 2017). The rationale is that STEM propels innovation, economic growth, and a nation’s standing and prestige (Sahin et al., 2015). It is in this context that STEM schools have been established, which is the focus of this entry. First, I contextualize the propagation of STEM schools by exploring the rise of STEM. What follows is an overview of the varieties of STEM schools and models, along with their common characteristics. I then offer an overview of the extant empirical

literature on STEM schools before closing with recommendations for future research endeavors and thoughts on the way forward, with equity considerations in mind.

Rationale for the rise of STEM education

Rallying cries for STEM education are not new. In fact, much of the STEM hype can be traced back to the late 1950s when the Soviet Union launched the satellite, Sputnik 1, on October 4, 1957 (Bybee, 2013). This event sparked a fundamental reckoning about the United States' global competitiveness. In particular, Sputnik stimulated societal calls and support for major science and mathematics curriculum reform projects funded by the National Science Foundation (NSF), such as the Biological Sciences Curriculum Study (BSCS) and the School Mathematics Study Group (MSG), that spanned decades afterward (Koehler et al., 2015). Without accompanying, ongoing professional development to support teachers with a new paradigm of inquiry-based teaching, these changes did not make a widespread dent in the fabric of daily classroom life; teachers relapsed into teaching the way they had taught previously, which was at odds with reform-based teaching and learning.

In 1983, *A Nation at Risk* was released, which painted a grave portrait of the United States' education system (NCEE, 1983). This report stated that U.S. students were falling markedly behind their counterparts in other developed nations and that students had abysmal deficits in basic academic skills and knowledge, let alone critical thinking. Frameworks and learning standards followed, such as *Project 2061: Science for All Americans* (AAAS, 1989), *Benchmarks for Science Literacy* (AAAS, 1993), and the *National Science Education Standards* (NRC, 1996). These documents communicated the importance of scientific literacy for all Americans, not just those deemed gifted and talented as had been the case in the 1960s (Bybee, 2013). Furthermore, *Rising Above the Gathering Storm* (NAS, 2007) was another report that issued a clarion call for the U.S. to revive its global status through major overhauls in STEM education. Some of the recommendations included vast improvements to K-12 science and mathematics education, an intensified approach to research and innovation, and the creation of high-paying jobs, accompanied by innovation, manufacturing, and marketing enhancements.

Indeed, countries the world over have substantially invested in such plans to cultivate STEM professionals capable of addressing complex, global challenges. These initiatives have been interpreted as critical infrastructure for bolstering a nation's well-being and include STEM education, teacher preparation, and professional development programs (Sharma and Yarlagaadda, 2018). For example, researchers of STEM education in the African countries of Liberia and Ethiopia have embraced STEM training to guard against the next pandemic and tackle road and housing construction concerns (Tardy Hackman et al., 2021; Melak and Singh, 2021). Similarly, others have acknowledged localized challenges in Vietnam due to climate change—including rising sea levels, increased flooding, landslides, and droughts—that could be further understood and addressed with integrated STEM expertise (Nguyen et al., 2020). China, India, Japan, The Philippines, Qatar, Singapore, South Korea, Thailand, Turkey, and Vietnam are countries in Asia documented in the literature with increased STEM interest and engagement (Ali et al., 2021; Changtong et al., 2020; Choi, 2014; Kurt and Benzer, 2020; Nguyen et al., 2020; Sharma and Yarlagaadda, 2018; Teo and Osborne, 2014).

Central and South American countries pursuing STEM include Ecuador and Chile (Bascope and Reiss, 2021; Ching-Chiang and Fernández-Cárdenas, 2020). In particular, Bascope and Reiss (2021) have advocated for sustainable STEM education in Chile that supports "habits of using scientific evidence to justify sustainable practices" and "the value of STEM education for society as a whole" (p. 2). Australia and New Zealand have invested heavily in training math and science teachers (Gough, 2021). Germany, Serbia, Spain, Sweden, and the United Kingdom are European countries that have been documented with increased support in STEM education as well (Andrée and Hansson, 2013; Hernández-Serrano and Muñoz-Rodríguez, 2020; Maksimović et al., 2020; Moote et al., 2020; Stoeger et al., 2021).

Undoubtedly, countries around the world are participating in this development of STEM infrastructure. At the same time, there is growing recognition that "STEM education should be inclusive and available to every student" (Nguyen et al., 2020, p. 12), given enduring gender gaps in physics, engineering, and computer science, which have continued to be male-dominated (Gough, 2021; Stoeger et al., 2021). Discrepancies in STEM education access and opportunities for ethnic minorities and individuals in rural areas are also worldwide injustices (Wang and Yang, 2021; Tardy Hackman et al., 2021; Melak and Singh, 2021). As Nguyen et al. (2020) put it, in the context of Vietnam, STEM is still largely often only for the elite given that "STEM education has become commercialised and dominated by private high-tech companies and schools in the large cities" (p. 12). Thus, STEM education has the potential to close opportunity gaps as well as exacerbate them.

Varieties of STEM schools

STEM schools are a manifestation of these global calls for STEM education and have recently carved out a unique area of educational research. While Teo and Osborne (2014) noted that STEM schools exist in the Philippines, Singapore, South Korea, and Australia, these schools have not been examined in the empirical research base. To date, the research literature has only examined STEM schools in the United States, and as such, this entry will focus on STEM schools in this context. In this section, I will provide an overview of the varieties of STEM schools, outlined in Fig. 1.

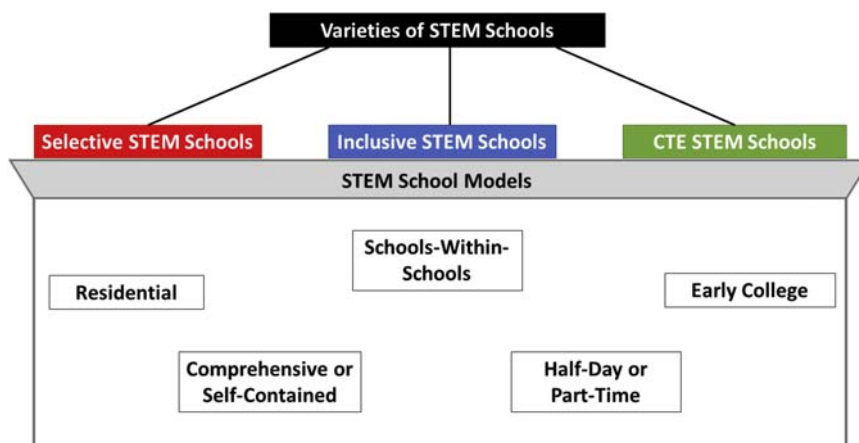


Fig. 1 Varieties of STEM schools.

Selective STEM schools

Specialized STEM schools are not new; they have been around since the turn of the 20th century, although their name has changed over time. Initially referred to as “manual training schools,” they became specialized science, mathematics, and technology (SMT) schools, and most recently have been referred to as “selective” STEM schools (Erdogan and Stuessy, 2015). In fact, Stuyvesant High School in New York City, which opened in 1904, is considered the first selective science and math high school in the U.S., originally only admitting boys. Others such as Brooklyn Technical High School and Bronx High School of Science in New York City opened in 1922 and 1938, respectively, with the goal of preparing students to attend college or begin technical careers. Like Stuyvesant High School, initially only male students were admitted.

Targeting the development of gifted scientific researchers and engineers, the National Defense Education Act spawned the creation of more specialized science high schools in 1964 (Subotnik et al., 2019). These STEM schools are known as selective STEM schools because they have been discerning in their entrance requirements and therefore admit only a select few students. They have narrow admissions requirements, often based on prior academic achievement and entrance exams; in some cases, they also require financial means to attend (Johnson and Sondergeld, 2020). Selective STEM schools have been lauded as “incubators for talent development” (Subotnik et al., 2016, p. 236), and the National Consortium for Specialized Secondary Schools of Mathematics, Science, and Technology (NCSSSMST), founded in 1988, has also played a significant role in the advancement of these schools (Erdogan and Stuessy, 2015).

Yet, selective STEM schools tend to serve students from traditionally privileged backgrounds and are less diverse. For example, in the Almarode et al. (2016) study, which explored the educational and career outcomes of selective STEM high schools, 70.1% of participants were Caucasian, 78.8% had English as their primary language, 81.2% had a parent/guardian with postsecondary education, and 56.7% had a parent/guardian in a STEM-related career. Additionally, with entrance exam performance as the sole criterion for admissions in a selective New York City STEM high school, only 4% of the student population was Black and Hispanic (Pride, 2014). In essence, selective STEM schools often lack racial/ethnic and economic diversity and tend to serve students from privileged backgrounds, typically overrepresented in STEM. As such, they operate with a tacit view that talent and giftedness are innate and that STEM education should be reserved for a small group of individuals.

Inclusive STEM schools

In contrast, more recent STEM education initiatives and trends aligned with education reform documents have sought to address disproportionate STEM participation through more intentional efforts at ensuring STEM access and opportunities for minoritized students (Johnson and Sondergeld, 2020). As Vaval et al. (2019) put it, the goal has been to “increase the number of students pursuing advanced degrees and careers in STEM, STEM literacy for all students, and the participation of women and minorities in the STEM workforce” (p. 1153). That is, while the STEM movement in the 1960s was primarily concerned with high-performing students and creating career scientists and mathematicians, today’s movement is more inclusive, centering STEM literacy for all students, regardless of their career intentions. All individuals should be able to acquire STEM knowledge and apply it, recognize how STEM disciplines shape society and vice versa, understand STEM disciplines as human endeavors, and engage in critical issues as concerned, informed individuals (Bybee, 2007).

It is with this inclusive mindset that in 2010, under former President Barack Obama, a report from the President’s Council of Advisors on Science and Technology (PCAST) clamored for the recruitment and training of 100,000 new STEM teachers with deep content knowledge and pedagogical expertise to effectively inspire and educate the nation’s youth. The report also called for the creation of at least 200 new STEM-focused high schools and 800 STEM-focused elementary and middle schools within the next decade. Further, the report stated that many of these schools should serve “minority and high-poverty communities”

and “take steps to ensure that all schools and school systems have access to relevant STEM-expertise” (PCAST, 2010, p. 8). STEM schools were lauded as unique national resources and laboratories for innovative STEM education approaches.

Since the PCAST report, the number of inclusive STEM-focused schools has increased in the United States (Johnson and Sondergeld, 2020). Rogers-Chapman (2014) had a list of 221 inclusive STEM high schools in the U.S., up from 100 in 2008 as documented in the Means et al. (2008) study. Statewide initiatives in North Carolina, Ohio, and Texas also catalyzed the creation of more ISHSs. In fact, the first inclusive STEM schools in Texas, known as T-STEM schools, were funded in 2006. A funding stipulation required that schools could not select students based on prior academic performance and had to have a student population that was more than 50% economically disadvantaged or more than 50% from racially/ethnically minoritized groups. From 2009–2010, Texas invested approximately \$120 million in 51 academies and 7 T-STEM technical assistance centers, a majority of which were established in major metropolitan contexts (Young et al., 2016).

Unlike selective STEM schools, inclusive STEM schools admit students who wish to attend based on interest, without exclusive measures of aptitude and prior achievement. As such, inclusive STEM schools actively recruit students interested in STEM who are traditionally underrepresented based on race/ethnicity, family income, and gender. Inclusive STEM schools consider “instructional and cultural supports to meet students where they are, providing relevance to their background, life experiences, and cultural values/norms to facilitate learning” (Johnson and Sondergeld, 2020, p. 2). Inclusive STEM schools are built on the idea that talent resides within every child and that STEM competencies can be developed and cultivated with appropriate access, opportunities, and scaffolds.

STEM Career and Technical Education (CTE) schools

A third broad category of STEM schools include STEM-focused Career and Technical Education (CTE) schools, often located in educational centers, comprehensive high schools, or career academies (NRC, 2011). STEM CTE schools aim to prepare students for college with a primary focus on practical, real-world applications through STEM-related career preparation, such as health services and information technology (Erdogan and Stuessy, 2015; Vaval et al., 2019). These schools have a broader purpose of preventing students from dropping out of school. In particular, students with learning disabilities are traditionally underrepresented in high school STEM courses, as they accrue fewer STEM courses in high school and graduate college less ready in STEM areas than students without learning disabilities (Freeman et al., 2021). The recent reauthorizations of the Carl D. Perkins Career and Technical Education Act in 2006 and 2018 called for increased access to a CTE curriculum as well as increased STEM supports for students with disabilities. As a result, there has been a greater focus on applied STEM CTE coursework—such as engineering technology and information technology—that center on application of science and math concepts for real-world STEM tasks (Plasman and Gottfried, 2018). Applied STEM coursework is one potential path by which STEM material can become more accessible for students (Freeman et al., 2021).

STEM school models

Regardless of category, STEM schools come in a variety of models including residential, comprehensive or self-contained, schools-within-schools, half-day or part-time, and early college (Almarode et al., 2016). In residential STEM schools, students live and study in a school environment that focuses on STEM learning. Residential STEM schools tend to serve students from across a state. Comprehensive or self-contained schools are similar to traditional high schools; they are independently administered and fully functional, stand-alone schools, with a concentrated focus on advanced STEM topics.

In the school-within-a-school model, independent STEM programs exist within larger, comprehensive schools (Tofel-Grehl and Callahan, 2014). Students can study in specific STEM classes while still being members of the broader school community and also participate in classes with peers outside of the specialized school curriculum. Likewise, students not enrolled in the STEM program can take advanced courses in the program.

In half-day or part-time programs, students enroll in advanced STEM courses outside of their home school but attend non-STEM classes with peers in their home school. While the comprehensive and school-within-a-school programs tend to be located in major urban areas, the half-day or part-time programs often operate in poorly resourced, rural counties that serve entire geographical regions. Students are bused to the school for part of the day from a number of secondary schools to participate in science-related disciplines for a portion of the day, while the rest of the day is spent at their home high school. Sometimes these programs occur at local colleges and universities. Early college STEM programs allow students to skip the last 2 years of high school and enroll in a local college early (Almarode et al., 2016). Due to partnerships between a high school and local college or university, students in early college programs can graduate from high school with a diploma and up to 2 years of college credit or an associate’s degree in a STEM-related field (Tofel-Grehl and Callahan, 2014).

Overall, STEM schools exist in a variety of geographical contexts. Although most are located in large metropolitan areas, there are STEM schools that serve rural populations. STEM schools can be private or public, including charter and magnet schools (Kloser et al., 2018; Judson, 2014; Means et al., 2008), and exist at elementary, middle, and high school levels. STEM schools tend to be small schools of choice with a STEM-related focus or theme such as health sciences or engineering and technology (Edmunds et al., 2017).

Characteristics of STEM schools

As evident from the previous section, no single model for STEM schools exists. However, STEM schools share a number of characteristics, or critical components, including a mission to prepare students for STEM pathways. A focus on one or more STEM disciplines, well-trained teachers, advanced coursework, innovative technology, and partnerships with community members for authentic research experiences, mentorships, and internships are others (Subotnik et al., 2016). While inclusive and selective STEM schools both share these features, selective STEM schools do not share an inclusive vision of actively recruiting and selecting students from underrepresented backgrounds without record of achievement based on traditional measures.

A body of research details critical components of inclusive STEM-focused high schools (ISHSs) based on literature reviews (Peters-Burton et al., 2014a) as well as document analyses according to school websites and national statistics databases, standardized test scores, interviews, and other publicly available data (LaForce et al., 2016; Scott, 2012). Scott (2012) noted that ISHSs have rigorous course requirements with electives centered on STEM content and real-world problem-solving applications, internships, and capstone projects. From an examination of 20 ISHSs across the U.S., LaForce et al. (2016) identified eight essential elements: personalization of learning; problem-based learning; rigorous learning; career, technology, and life skills; school community and belonging; external community; staff foundations; and external factors. Another study explored the cultural dimensions that influence curriculum and instruction for inclusive STEM schools at elementary, middle, and high school levels (Lesseig et al., 2018). Unique student learning cultural dimensions across the grade levels include the school's vision, community partnerships, course scheduling, and testing pressures along with instructional practices, interdisciplinary collaboration, and teacher content knowledge.

Perhaps most encompassing of ISHS critical components is the work of Peters-Burton et al. (2014a) and Lynch et al. (2018). Peters-Burton et al. (2014a) outlined ten characteristics of successful ISHSs from the literature. These components include: STEM-focused curriculum; reform instructional strategies and project-based learning; integrated, innovative technology use; blended formal/informal learning beyond the typical school day, week, or year; real-world STEM partnerships; early college-level coursework; well-prepared STEM teaching staff; inclusive STEM mission; administrative structure; and supports for underrepresented students. Combined, these critical components comprise opportunity structures that advantageously position students on STEM pathways and careers.

Lynch et al. (2018) expanded on this work, introducing four critical components developed inductively through their cross-case analysis of eight ISHSs. These critical components include data-driven decision making for continuous improvement; innovative and responsive leadership; positive school community and culture of high expectations for all; and agency and choice. Although the substance remains the same, Lynch and colleagues also slightly modified the terminology of several critical components from the Peters-Burton et al. (2014a) study. Fig. 2 offers a visual overview of these ISHS critical components and their link to STEM school outcomes.

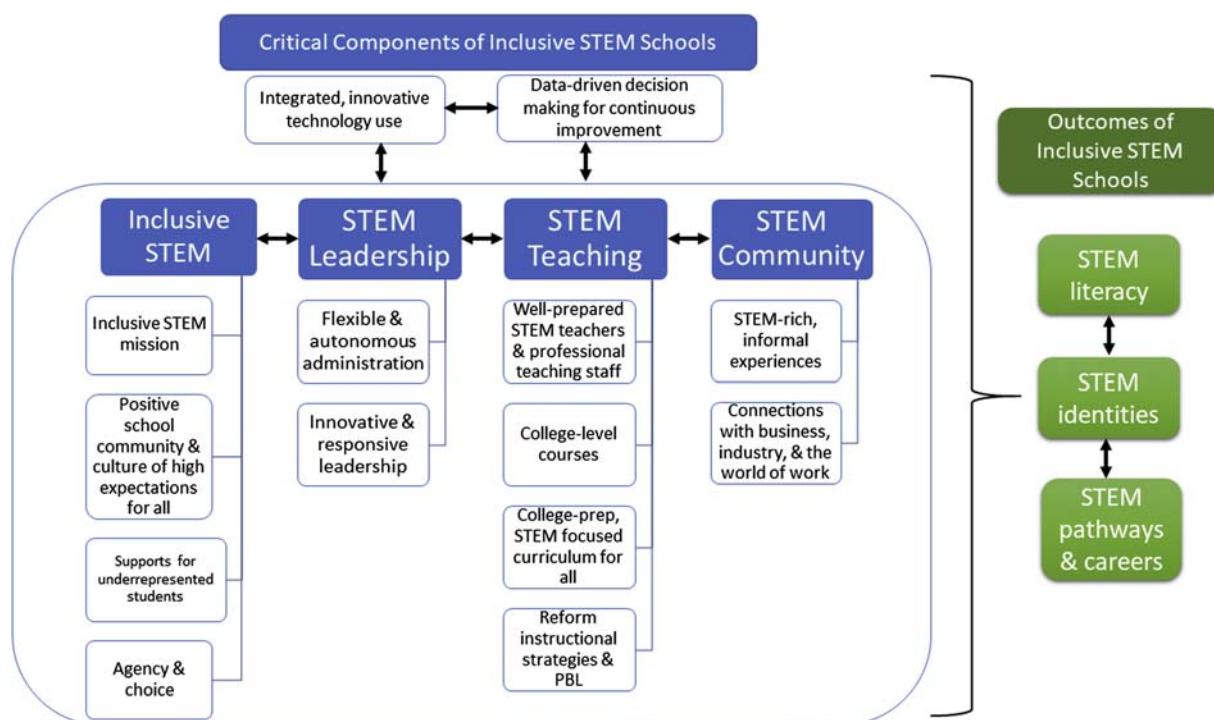


Fig. 2 STEM school critical components and their links to STEM school outcomes.

Critical components explained

In this section, I offer a short description of the critical components of inclusive STEM-focused high schools. Given that students choose to attend these schools, they harbor a sense of purpose that aligns with the school's STEM focus and recognize the challenges involved with a rigorous curriculum. The critical component of a college-prep, STEM-focused curriculum for all (Lynch et al., 2018; Peters-Burton et al., 2014a) entails rigorous courses in all four STEM disciplines that support STEM pathways and careers. The extent to which ISHSs practice integrated STEM, however, is unclear, although they often have project- or problem-based learning approaches that support the interdisciplinary connections and development of 21st century skills such as collaboration, critical thinking, creativity, and leadership (Holmlund et al., 2018). Reform instructional strategies include project- and problem-based learning and performance-based assessments that encourage students to take ownership of their learning.

ISHSs are well-resourced, including technology that supports students in research, modeling, and STEM-related social networking (Vaval et al., 2019). In order to support students in meeting rigorous expectations, ISHSs offer tutoring, advisories, and special classes such as college and career readiness. Learning extends beyond the school day, with rich, informal STEM learning opportunities, including mentoring, apprenticeships, internships, after school programs, and field trips, furthered by community, business, and industry partnerships. Schools may participate in Science Olympiads and science fairs.

STEM teachers and professionalized teaching staff are considered highly qualified due to their certifications, advanced STEM content knowledge, and/or practical experience in STEM careers. Teachers are supported with professional development, collaboration, and interactions with STEM professionals and have high expectations for their students (Means et al., 2017). Moreover, due to liberties from the district to address innovative goals and partnerships with organizations that offer expertise, leadership, and resources, administration is flexible and autonomous. Leadership is also responsive and innovative and proactively addresses the needs of teachers, students, and the larger community, through open communication and data-based decision making. Overall, ISHSs possess a culture of continuous improvement and high expectations for students and staff who "feel a sense of personal, intellectual, and socio-emotional safety" (Lynch et al., 2018, p. 718).

Review of the STEM school literature

In this section, I review the current peer-reviewed articles on STEM schools. Apart from the characteristics of STEM schools detailed in the previous section, the existing body of STEM school studies primarily centers on STEM school effectiveness, measured in terms of outcomes. While there is still much to be learned about their effectiveness, extant research offers initial results. In the following sections, I discuss the literature on selective and inclusive STEM school outcomes.

Selective STEM school outcomes

In a study of selective STEM schools that examined students' post-graduation plans, 75% of participants expressed a desire to go to college and 40% planned to obtain doctorate degrees (Thomas, 2000). Slightly over half of the students who graduated from these schools majored in science and about 10% majored in mathematics in college. Likewise, according to a national survey of alumni of 25 selective STEM schools, almost 50% of students graduated with a STEM degree (Subotnik et al., 2013). Students also demonstrated positive outcomes on state assessments, college entrance exams, and end-of-course assessments. In another study, students in selective STEM schools were more likely to report earning a STEM-related degree if their teachers asked questions with pre-determined answers, made connections with other content areas, focused on deep understanding of complex content, accepted criticism of their ideas, and reinforced self-confidence in STEM activities (Almarode et al., 2016). Moreover, participation in an internship or mentorship experience were additional characteristics positively associated with earning a STEM-related degree. Overall, studies of selective STEM school outcomes have been positive with respect to post-secondary STEM achievements and aspirations.

Inclusive STEM school outcomes

It is argued that ISHSs support minoritized students through curricular, instructional, and organizational opportunity structures that promote STEM knowledge and skill development and positively position students on STEM pathways (Lynch et al., 2018). An emerging and growing area of research on ISHSs offers some insights into student outcomes. In the sections that follow, I detail the ISHS research on student course-taking patterns and experiences, student achievement, student aspirations and interests, and effectiveness of addressing disparities.

Rigorous course-taking patterns and experiences

Using survey data, Means et al. (2016) compared ISHS students' high school experiences and achievement with that of students in non-STEM schools in North Carolina. Students who attended an ISHS were more likely to complete pre-calculus and chemistry in high school. Similarly, Saw (2019) found positive associations between ISHS enrollment and student completion of advanced level math courses in high school. Another study that examined 94 ISHSs and almost 10,000 students found positive effects on students' completion of key STEM courses and participation in STEM activities as well (Means et al., 2021).

Likewise, instrumental case studies have confirmed reform-oriented instructional approaches in ISHSs (Lynch et al., 2017; Morrison et al., 2015; Slavit et al., 2016). For example, graduates from T-STEM academies in Texas described their teachers as having extensive interdisciplinary knowledge. They also mentioned inquiry and project- and problem-based learning experiences that allowed for group work and real-life, hands-on learning applications (Bicer et al., 2020). Students felt supported by tutoring programs to meet the rigorous expectations of Advanced Placement courses and dual credit opportunities. Authentic research collaborations with university professors and internships in key industries were also highlights of their experiences.

Moreover, a cross-case analysis that explored students' engineering education experiences indicated that students learned key components of engineering such as design thinking, engineering habits of mind, and communication skills, although other elements such as the tension between optimization and constraints, systems thinking, and mathematical modeling were lacking (Peters-Burton and Johnson, 2018). Despite limited resources that were budgetary, technological, and logistical in nature, Wayne School of Engineering (WSE), a rural high school in North Carolina, maintained increased rigor of courses and connections with outside partners (Peters-Burton et al., 2014b). In these ways, a majority of the ISHS critical components that address rigorous, college-level STEM experiences and supports were present.

In contrast, some ISHS studies have detailed an erosion of opportunity structures as schools have attempted to meet accountability demands, including raising state exam scores and graduation rates while facing multiple challenges simultaneously (Eisenhart et al., 2015; Weis et al., 2015). For example, gatekeeping courses such as chemistry, physics, and pre-calculus—that would positively position students on STEM pathways—were no longer offered at an ISHS in the Weis et al. (2015) study by the 3rd year. Instead, only repeater and remedial math and science courses that focused heavily on standardized test preparation were provided. In an ISHS where these courses were still offered, they were found to exist in name only; that is, only Algebra I and II content were covered in a pre-calculus course. Moreover, Stehle and Peters-Burton (2019) conducted a mixed methods study that examined the extent to which student work samples and teacher lesson plans from seven exemplary ISHSs engaged in 21st century skills. Results indicated that out of 60 lessons, 50 included instruction on 21st century skills, but most were at an introductory level. Lesson plans for grades 11 and 12 tended to involve higher-order skills. Overall, there is much to be explored with regard to in-depth, instructional practices.

Achievement

Another area of the ISHS literature has focused on student outcomes such as achievement on standardized assessments. Scott's (2012) multiple case study of ten high schools revealed that students were more likely to score higher than the state average in assessment areas of math, science, and English. Also promising, ninth grade students at T-STEM academies performed slightly better on mathematics state tests, and 10th grade students performed slightly better on both mathematics and science state tests (Young et al., 2011). In a study in which academic and attendance outcomes of the first inclusive, integrated STEM high school in the United States were explored, not only were student attendance rates higher, but more than two-thirds of students were deemed college-ready on the ACT as compared with about 40% in the district (Johnson and Sondergeld, 2020).

There have also been mixed or negligible achievement results. For instance, findings from Means et al. (2021) revealed a small positive impact on science test scores for students attending ISHSs, particularly economically disadvantaged students, although there were no discernible impacts on math test scores. In another study, no significant student gains were found in reading, mathematics, and science from attending the inclusive Texas STEM schools; however, the study only covered the first 4 years of the academies (Sahin et al., 2015). Results from Saw (2019) indicate that Texas ISHSs had no impact on student test scores in math and science. Additionally, Gnagey and Lavertu (2016) found that attendance during the first 2 years of high school in Ohio ISHSs can have a negligible or negative effect on academic achievement. In some ISHSs, there were achievement gains in science courses, but these gains were accompanied by lower performance in non-STEM courses.

STEM interests and aspirations

Research also exists with respect to students' STEM interests and college and career aspirations. Means et al. (2016) found students to be more interested in science careers and involved in STEM extracurricular activities, which were positively associated with students' attendance and GPA. These findings are consistent with that of the LaForce et al. (2019) study in which students had positive attitudes toward science and STEM careers, and those with higher ratings of supportive relationships with the school community also reported significantly higher ratings on science ability beliefs (LaForce et al., 2019). However, as students progressed through the grade levels, their science attitudes declined.

Alternatively, results of another study revealed that students who participated in key STEM courses and self-selected STEM activities identified more strongly with math and science as seniors, and they were more likely to be interested in one or more STEM careers (Means et al., 2021). Furthermore, all participants in the Bicer et al. (2020) study who had completed 1–3 years of undergraduate education when the study was conducted had passed their college courses. They also majored in STEM areas of expertise including engineering, mathematics, computer science, molecular biology, animal science, and genetics. Overall, ISHSs “do typically enhance STEM coursetaking and career interest across a range of different state contexts” (Means et al., 2021, p. 17). In these ways, the possibilities of ISHSs extend beyond entry into STEM majors and careers and also potentially promote science dispositions and affective factors that could facilitate lifelong, sustained engagement in STEM.

Racial/ethnic, economic, and gender disparities

Means et al. (2021) found that ISHSs attracted students from groups underrepresented in STEM, with 63% of students from low-income homes and 69% classified as underrepresented minorities. Provided that ISHSs are designed with a particular mission to support underrepresented students in STEM, studies have also examined whether these schools have been effective at reducing persistent racial/ethnic, economic, and gender disparities. In one study, high school graduation rates for racialized minorities and low-income students improved, given that the math performance of low-income students in ninth grade increased, and female students were more likely to complete more AP or IB credits in math (Saw, 2019). In another study, female Hispanic and African American students in North Carolina and Texas ISHSs were more likely to take advanced STEM courses as compared with students in traditional schools (Means et al., 2017). In both states, 12th grade females also expressed greater interest in STEM careers if they attended an ISHS.

In contrast, LaForce et al. (2019) found that traditional race and gender gaps persisted, and Black and Hispanic students fell below White students in attitudes and achievement. While females consistently demonstrated higher GPA than males, they had less positive attitudes toward science and STEM careers. As the researchers posited, “While STEM schools seem to be successful at providing a diverse body of students with access to an environment that promotes STEM success, access alone may not be sufficient to eliminate gaps in science attitudes and GPA across race/ethnicity and gender identity” (LaForce et al., 2019, p. 12, emphasis in original).

Despite claims to equity and fairness, STEM schools in Chicago have given cover for school closures, the displacement of Black and Latinx communities, and the reification of discriminatory tracking, while White middle-class families and corporate interests have been prioritized (Morales-Doyle and Gutstein, 2019). Black students have been drastically underrepresented in the most desirable, selective-enrollment Chicago STEM schools and overrepresented in schools that have steered students toward working-class jobs needed for Chicago labor markets. In fact, predominantly Black schools in Chicago offered zero science or mathematics dual-credit options and emphasized workplace skills over science. This reality contradicts the touted “college-prep STEM curriculum for all” and pathways for lucrative STEM careers (Morales-Doyle and Gutstein, 2019). As documented in the Eisenhart et al. (2015) and Weis et al. (2015) studies, the STEM school movement has also been a driver of inequality.

Summary of findings

Overall, research into STEM school outcomes is in its early stages. Given that STEM schools are relatively new, reform effects can take 5 or more years (Borman et al., 2003). Apart from identifying distinguishing characteristics of these schools, existing literature on STEM schools consider their effectiveness, based on factors such as course taking patterns and experiences, achievement on standardized tests, STEM interests and postsecondary aspirations, and reducing educational disparities. Though there are mixed results on the effectiveness of STEM schools, for the most part, studies signal slight positive outcomes on science and math, attendance, high school graduation rates, and STEM college and career pathway interest (Bottia et al., 2018; Erdogan and Stuessy, 2015; Means et al., 2017, 2016; Wiswall et al., 2014). As Means et al. (2017) concluded, “The ISHS data suggest that regardless of their demographic background, students who have an interest in STEM can benefit from a rigorous STEM-focused curriculum if provided with the kinds of instruction and supports emphasized in the ISHS model” (p. 17). There is some promise, although scrutiny and vigilance are necessary, given the warnings from other cases (Bullock, 2017; Eisenhart et al., 2015; Weis et al., 2015; Morales-Doyle and Gutstein, 2019).

When it comes to elementary, middle, charter, magnet, and private STEM schools, there is a paucity of research, and their effectiveness is less promising. In a study of nine STEM-focused charter schools and two STEM-focused magnet schools serving elementary-aged students, Judson (2014) found no effect on achievement in math, language arts, and reading in the magnet schools and some positive effects in the STEM-focused charter schools. Additionally, an examination of Catholic kindergarten to eighth grade STEM-focused schools revealed tensions in navigating dual Catholic and STEM-focused identities (Kloser et al., 2018).

What is more, the impact of STEM-focused CTE programs remains largely unexplored. While there are no specific studies of STEM-focused CTE schools, there are studies of STEM CTE course-taking outcomes, which have produced negligible or mixed results. For example, Gottfried and Sublett (2018) found that students with learning disabilities were less likely to take applied STEM courses in high school as compared with the general population, while enrolling in applied STEM courses was not associated with taking advanced math and science coursework later in high school.

Upon examination of CTE-STEM course enrollment in Texas, another study found a larger gender disparity in advanced science courses than advanced math courses, and racial/ethnic and gender disparities worsened over the years (Yoon and Strobel, 2017). Moreover, Freeman et al. (2021) found no evidence that earning more Applied STEM Career and Technical Education (AS-CTE) credits in high school influenced college enrollment patterns or major selection for students with learning disabilities. However, students with learning disabilities who earned more AS-CTE credits in high school were more likely to seriously consider and declare AS-related majors in college. For low-income students as compared with their non-STEM-CTE counterparts, taking STEM-CTE courses has been associated with higher school engagement (Plasman et al., 2021). Overall, these varied results warrant additional exploration.

Going forward: a vision for STEM schools of tomorrow

STEM schools are considered full of potential and promise for STEM reform efforts. As [Morales-Doyle and Gutstein \(2019\)](#) noted, “The cachet of STEM education is related to the high status and pay of the associated careers and is also connected with the prestige of canonical scientific and mathematical knowledge” (p. 533). At this juncture, a majority of STEM schools have been around for less than 10 years. Accordingly, STEM school educational research is still in its infancy. While much has been learned about their common features and effectiveness, gaping unknowns remain. In this section, I offer some suggestions for further research and a vision for STEM schools of the future.

The current empirical literature base on STEM schools is dominated by United States contexts and as such, further examination into STEM schools in other international contexts would be fruitful, especially given the global rise of STEM. A lot can also be learned from studies of STEM schools across grade levels, including elementary and middle schools. In fact, the 2010 PCAST report called for 80% of the 1,000 new STEM-focused schools to be elementary and middle schools. A range of other school contexts should be explored as well, from rural to private and CTE, among various models including residential and early college programs.

Existing ISHS studies have offered snapshots of classroom life due to participant observation spanning two-to four-day time frames and across courses in different disciplines and grade levels ([Bruce-Davis et al., 2014](#); [Eisenhart et al., 2015](#)). As such, in-depth, qualitative studies that engage continuous observations that can further understandings into organizational, instructional, and cultural practices are imperative ([Tripp, 2021](#)). This nuanced contextualization will be key to understanding distinctive educational practices, the fidelity and effectiveness of implementation, and navigation of tensions involved with accountability and other pressures that stakeholders might face ([Nasir and Vakil, 2017](#)). Deeper interrogation into critical components such as STEM school leadership and innovative, integrative technology use are also valuable ([Lynch et al., 2018](#)).

Another arena for prospective research involves an expanded view of STEM school outcomes, based on multiple, varied measures of assessment that more authentically evaluate what students know and are able to do and are informed by culturally responsive, relevant, and sustaining approaches ([Gay, 2010](#); [Ladson-Billings, 2014](#); [Paris and Alim, 2017](#)). As the [NRC \(2011\)](#) contends, “Test scores, however, do not tell the whole story of success” (p. 6). Outcomes such as shifting STEM identities, motivations, feelings of belonging, and attitudes toward STEM among the STEM school community—along with the effectiveness of these schools in redressing opportunity gaps ([Milner, 2012](#))—should be considered as well. In what ways might joy, community, criticality, and commitments to morality and humanity be fostered and measured ([Morales-Doyle, 2019](#); [Szostkowski and Upadhyay, 2019](#))? In what ways might the longstanding stereotypes about science and STEM be challenged and overturned in STEM schools ([King and Pringle, 2019](#))? Do students and communities feel equipped and empowered to tackle authentic, socioscientific issues relevant to their everyday lives? Perspectives from multiple stakeholders, including school leaders and family and community members, to multilingual students and students with disabilities should be explored.

Efforts to understand school- and student-level contextual conditions surrounding outcomes are essential as well. Longitudinal investigations that follow STEM school students across elementary, middle, and high school—as well as post-graduation—would be ideal. Such studies would offer more insights into the supports that all stakeholders in STEM school communities need to succeed—in addition to the assets that they bring. This research agenda would inform a more concrete direction for improving STEM instruction and learning in the areas of design, implementation, and evaluation of STEM schools.

Equity considerations

It is well known that there are persistent racial/ethnic, gender, socioeconomic, and disability-based inequalities with respect to participation in STEM fields. Disparities also exist in rural schools with regard to advanced opportunities such as college-level coursework ([Assouline et al., 2017](#)). In fact, [Saw and Agger \(2021\)](#) found that there were fewer STEM opportunities for rural and small-town students in the United States, which adversely affected STEM career aspirations, academic preparation, and postsecondary STEM participation.

While parity in participation is important, equity in STEM is more than “simply access or exposure;” it is also about “issues of knowledge production, epistemology, culture, and power” ([Nasir and Vakil, 2017](#), p. 378) which have fueled these disparities. Indeed, the culture of STEM and the spaces in which STEM opportunities have existed often mirror the interests, values, and experiences of the dominant culture and are often unwelcoming to minoritized groups as a result. Instead of reinscribing traditional White male cultural norms and Eurocentric, positivist teachings of STEM, it is important for all stakeholders to recognize and seek “multiple perspectives, knowledge frames, and approaches” ([Mensah and Jackson, 2018](#), p. 32). STEM should be considered with “local knowledge, stories, narratives, traditional practices, wisdom, and collective intuitions” ([Bascope and Reiss, 2021](#), p. 2). Moreover, STEM alone cannot be a panacea to all of society’s problems. Rather than portray STEM as “apolitical and inherently good” ([Hyttén and Stenmagen, 2020](#), p. 28), STEM education should afford students opportunities to unpack the ways in which STEM can and has been used to perpetuate both benefit and harm.

As [Bullock \(2017\)](#) noted, “it is not possible to achieve an equitable future without correcting the discriminatory policies and practices that make up the fabric of schooling” (p. 637). Without critically interrogating embedded assumptions and power dynamics in the design and implementation of STEM schools, without intentionally integrating culturally and linguistically

sustaining, anti-racist, and decolonizing approaches, inequities have and will continue to endure (hooks, 1994; Paris and Alim, 2017). STEM education should prepare and empower marginalized students to critically contend with issues relevant to their lives and focus on ways to address them in solidarity with one another (Morales-Doyle and Gutstein, 2019). Inclusive STEM school models have promising approaches, yet beyond adhering to their design with fidelity, opportunity structures of all dimensions must be reexamined and monitored such that they become centered around students' identities, nurture competence, raise critical consciousness, and cultivate joy, empowerment, and community (Tripp, 2021). In these ways, with an aim of continuous improvement and in the spirit of collective interdependence, STEM schools of the future can be welcoming, stimulating, and nurturing hubs of excellence for everyone.

References

- Ali, R., Bhadra, J., Siby, N., Ahmad, Z., Al-Thani, N.J., 2021. A STEM model to engage students in sustainable science education through sports: a case study in Qatar. *Sustainability* 13, 1–26.
- Almarode, J., Subotnik, R., Lee, M., 2016. What works: lessons from a research study of specialized science high school graduates. *Gift. Child Today* 39, 185–190.
- American Association for the Advancement of Science (AAAS), 1989. Project 2061: Science for All Americans. Oxford University Press, New York.
- American Association for the Advancement of Science (AAAS), 1993. Benchmarks for Science Literacy. Oxford University Press, New York.
- Andrée, M., Hansson, L., 2021. Industry, science education, and teacher agency: a discourse analysis of teachers' evaluations of industry-produced teaching resources. *Sci. Educ.* 105, 353–383.
- Assouline, S.G., Ihrig, L.M., Mahatmya, D., 2017. Closing the excellence gap: investigation of an expanded talent search model for student selection into an extracurricular STEM program in rural middle schools. *Gift. Child. Q.* 61, 250–261.
- Bascope, M., Reiss, K., 2021. Place-based STEM education for sustainability: a path towards socioecological resilience. *Sustainability* 13, 1–16.
- Bicer, A., Lee, Y., Perihan, C., 2020. Inclusive STEM high school factors influencing ethnic minority students' STEM preparation. *J. Ethn. Cult. Stud.* 7, 147–172.
- Borman, G.D., Hewes, G.M., Overman, L.T., Brown, S., 2003. Comprehensive school reform and achievement: a meta-analysis. *Rev. Educ. Res.* 73, 125–230.
- Bottia, M.C., Stearns, E., Mickelson, R.A., Moller, S., 2018. Boosting the numbers of STEM majors? The role of high schools with a STEM program. *Sci. Educ.* 102, 85–107. <https://doi.org/10.1002/sce.2131>.
- Bruce-Davis, M.N., Gubbins, E.J., Gilson, C.M., Villanueva, M., Foreman, J.L., Rubenstein, L.D., 2014. STEM high school administrators', teachers', and students' perceptions of curricular and instructional strategies and practices. *J. Adv. Acad.* 25, 272–306. <https://doi.org/10.1177/1932202X14527952>.
- Bullock, E.C., 2017. Only STEM can save us? Examining race, place, and STEM education as property. *Educ. Stud.* 53, 628–641.
- Bybee, R.W., 2007. Do we need another Sputnik? *Am. Biol. Teach.* 69, 454–457.
- Bybee, R.W., 2013. Case for STEM Education: Challenges and Opportunities. National Science Teachers Association, Arlington.
- Changtong, N., Maneajak, N., Yasri, P., 2020. Approaches for implementing STEM (Science, Technology, Engineering & Mathematics) activities among middle school students in Thailand. *Int. J. Educ. Methodol.* 6, 185–198. <https://doi.org/10.12973/ijem.6.1.185>.
- Ching-Chiang, L.-W.C., Fernández-Cárdenas, J.M., 2020. Analysing dialogue in STEM classrooms in Ecuador: a dual socioeconomic context in a high school. *J. N. Approaches Educ. Res.* 9, 194–215.
- Choi, K.M., 2014. Opportunities to explore for gifted STEM students in Korea: from admissions criteria to curriculum. *Theory Pract.* 53, 25–32.
- Edmunds, J., Arshavsky, N., Glennie, E., Charles, K., Rice, O., 2017. The relationship between project-based learning and rigor in STEM-focused high schools. *Interdiscip. J. Probl. Based Learn.* 11. <https://doi.org/10.7771/1541-5015.1618>.
- Eisenhart, M., Weis, L., Allen Carrie, D., Cipollone, K., Stich, A., Dominguez, R., 2015. High school opportunities for STEM: comparing inclusive STEM-focused and comprehensive high schools in two U.S. cities. *J. Res. Sci. Teach.* 52, 763–789. <https://doi.org/10.1002/tea.21213>.
- Erdogan, N., Stuessy, C.L., 2015. Modeling successful STEM high schools in the United States: an ecology framework. *Int. J. Educ. Math. Sci. Technol.* 3, 77–92.
- Freeman, J.A., Gottfried, M.A., Plasman, J.S., 2021. STEM-focused career courses and college pipeline for students with learning disabilities. *Educ. Pol.* 0, 1–31.
- Gay, G., 2010. Culturally Responsive Teaching: Theory, Research, and Practice, second ed. Teachers College Press, New York.
- Gnagye, J., Lavertu, S., 2016. The impact of inclusive STEM high schools on student achievement. *AERA Open* 2, 1–21.
- Gottfried, M.A., Sublett, C., 2018. Does applied STEM course taking link to STEM outcomes for high school students with learning disabilities? *J. Learn. Disabil.* 51, 250–267.
- Gough, A., 2021. All STEM-Educ: gaps and silences around ecological education in Australia. *Sustainability* 13, 1–19. <https://doi.org/10.3390/su13073801>.
- Hernández-Serrano, M.J., Muñoz-Rodríguez, J.M., 2020. Interest in STEM disciplines and teaching methodologies: perception of secondary school students and preservice teachers. *Educator* 56, 369–386.
- Holmlund, T.D., Lesseig, K., Slavitt, D., 2018. Making sense of “STEM education” in K-12 contexts. *Int. J. STEM Educ.* 5, 1–18.
- hooks, B., 1994. Teaching to Transgress: Education as the Practice of Freedom. Routledge, New York, NY.
- Hyttén, K., Sternhagen, K., 2020. When STEM and STEAM really mean ABC: a democratic critique of “anything but civics” schools. *Educ. Stud.* 56, 18–36.
- Johnson, C.C., Sondergeld, T.A., 2020. Outcomes of an integrated STEM high school: enabling access and achievement for all students. *Urban Educ.* 0, 1–27. <https://doi.org/10.1177/0042085920914368>.
- Judson, E., 2014. Effects of transferring to STEM-focused charter and magnet schools on student achievement. *J. Educ. Res.* 107, 255–266.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3, 1–11. <https://doi.org/10.1186/s40594-016-0046-z>.
- King, N.S., Pringle, R.M., 2019. Black girls speak STEM: counterstories of informal and formal learning experiences. *J. Res. Sci. Teach.* 56, 539–569.
- Kloser, M., Wilsey, M., Hopkins, D.W., Dallavis, J.W., Lavin, E., Comuniello, M., 2018. Dual identities: organizational negotiation in STEM-focused Catholic schools. *Cult. Stud. Sci. Educ.* 13, 549–579.
- Koehler, C., Binns, I.C., Bloom, M.A., 2015. The emergence of STEM. In: Johnson, C.C., Moore, T., Peters-Burton, E.E. (Eds.), *STEM Road Map: A Framework for Integrated STEM Education*. Taylor & Francis, pp. 13–22.
- Kurt, M., Benzer, S., 2020. An investigation on the effect of STEM practices on sixth grade students' academic achievement, problem solving skills, and attitudes towards STEM. *J. Sci. Learn.* 3, 79–88.
- Ladson-Billings, G., 2014. Culturally relevant pedagogy 2.0: A.K.A. the remix. *Harv. Educ. Rev.* 84, 74–84.
- LaForce, M., Noble, E., King, H., Century, J., Blackwell, C., Holt, S., Ibrahim, A., Loo, S., 2016. The eight essential elements of inclusive STEM high schools. *Int. J. STEM Educ.* 3, 1–11. <https://doi.org/10.1186/s40594-016-0054-z>.
- LaForce, M., Zuo, H., Ferris, K., Noble, E., 2019. Revisiting race and gender differences in STEM: can inclusive STEM high schools reduce gaps? *Eur. J. STEM Educ.* 4, 1–15. <https://doi.org/10.20897/ejsteme/5840>.
- Lesseig, K., Slavitt, D., Holmlund Nelson, T., 2017. Jumping on the STEM bandwagon: how middle grades students and teachers can benefit from STEM experiences. *Middle Sch. J.* 48, 15–24.

- Lesseig, K., Firestone, J., Morrison, J., Slavit, D., Holmlund, T., 2018. An analysis of cultural influences on STEM schools: similarities and differences across K-12 contexts. *Int. J. Sci. Math. Educ.* 17, 449–466.
- Lynch, S.J., Spillane, N., House, A., Peters-Burton, E., Behrend, T., Ross, K.M., Han, E.M., 2017. A policy-relevant instrumental case study of an inclusive STEM-focused high school: Manor New Tech High. *Int. J. Educ. Math. Sci. Technol.* 5, 1–20. <https://doi.org/10.18404/ijemst.75656>.
- Lynch, S.J., Peters-Burton, E., Behrend, T., House, A., Ford, M., Spillane, N., Matray, S., Han, E., Means, B., 2018. Understanding inclusive STEM high schools as opportunity structures for underrepresented students: critical components. *J. Res. Sci. Teach.* 55, 712–748. <https://doi.org/10.1002/tea.21437>.
- Maksimović, J.Z., Osmanović, J.S., Mamutović, A.S., 2020. Perspectives of STEM education regarding Serbian secondary school students' motivation for career choice. *J. Baltic Sci. Educ.* 19, 989–1007. <https://doi.org/10.33225/jbse/20.19.989>.
- Means, B., Confrey, J., House, A., Bhanot, R., 2008. STEM High Schools: Specialized Science, Technology, Engineering, and Mathematics Secondary Schools in the U.S. Report Prepared for the Bill & Melinda Gates Foundation. SRI International, Menlo Park, CA. Retrieved from. <http://ctl.sri.com/publications/displayPublicationResults.jsp>.
- Means, B., Wang, H., Young, V., Peters, V.L., Lynch, S.J., 2016. STEM-focused high schools as a strategy for enhancing readiness for postsecondary STEM programs. *J. Res. Sci. Teach.* 53, 709–736. <https://doi.org/10.1002/tea.21313>.
- Means, B., Wang, H., Wei, X., Lynch, S., Peters, V., Young, V., Allen, C., 2017. Expanding STEM opportunities through inclusive STEM-focused high schools. *Sci. Educ.* 101, 681–715. <https://doi.org/10.1002/sce.21281>.
- Means, B., Wang, H., Wei, X., Young, V., Iwatani, E., 2021. Impacts of attending an inclusive STEM high school: meta-analytic estimates from five studies. *Int. J. STEM Educ.* 8, 1–19.
- Melak, A., Singh, S., 2021. Women's participation and factors affecting their academic performance in engineering and technology education: a study of Ethiopia. *Sustainability* 13, 1–22.
- Mensah, F.M., Jackson, I., 2018. Whiteness as property in science teacher education. *Teach. Coll. Rec.* 120, 1–38.
- Milner, H.R., 2012. Beyond a test score: explaining opportunity gaps in educational practice. *J. Black Stud.* 43, 693–718.
- Moote, J., Archer, L., DeWitt, J., MacLeod, E., 2020. Science capital or STEM capital? Exploring relationships between science capital and technology, engineering, and math aspirations and attitudes among young people aged 17/18. *J. Res. Sci. Teach.* 57, 1228–1249.
- Morales-Doyle, D., Gutstein, E., 2019. Racial capitalism and STEM education in Chicago public schools. *Race Ethn. Educ.* 22, 525–544.
- Morales-Doyle, D., 2019. There is no equity in a vacuum: on the importance of historical, political, and moral considerations in science education. *Cult. Stud. Sci. Educ.* 14, 485–491. <https://doi.org/10.1007/s11422-019-09925-y>.
- Morrison, J., Roth McDuffie, A., French, B., 2015. Identifying key components of teaching and learning in a STEM school. *Sch. Sci. Math.* 115, 244–255. <https://doi.org/10.1111/ssm.12126>.
- Nasir, N.S., Vakili, S., 2017. STEM-focused academies in urban schools: tensions and possibilities. *J. Learn. Sci.* 26, 376–406. <https://doi.org/10.1080/10580406.2017.1314215>.
- National Academy of Science (NAS): Committee on Science, Engineering, Public Policy, 2007. *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. National Academies Press, Washington, D.C. www.nap.edu/books/0309100399/html
- National Commission on Excellence in Education (NCEE), 1983. *A Nation at Risk*. US Department of Education, Washington, D.C.
- National Research Council (NRC), 1996. *National Science Education Standards*. National Academies Press, Washington, D.C.
- National Research Council, 2011. *Successful K-12 STEM education: identifying effective approaches in science, technology, engineering, and mathematics*. In: Committee on Highly Successful Science Programs for K-12 Science Education. Board on Science Education and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. The National Academies Press, Washington, D.C.
- Next Generation Science Standards Lead States, 2013. *Next Generation Science Standards: For States, by States*. National Academies Press, Washington, DC. Retrieved from. <https://www.nextgenscience.org/>.
- Nguyen, T.P.L., Nguyen, T.H., Tran, T.K., 2020. STEM education in secondary schools: teachers' perspective towards sustainable development. *Sustainability* 12, 1–16.
- Paris, D., Alim, H.S. (Eds.), 2017. *Culturally Sustaining Pedagogies: Teaching and Learning for Justice in a Changing World*. Teachers College Press, New York.
- Peters Burton, E., Kaminsky, S.E., Lynch, S., Behrend, T., Han, E., Ross, K., House, A., 2014b. Wayne school of engineering: case study of a rural inclusive STEM-focused high school. *Sch. Sci. Math.* 114, 280–290. <https://doi.org/10.1111/ssm.12080>.
- Peters-Burton, E.E., Johnson, T., 2018. Cross-case analysis of engineering education experiences in inclusive STEM-focused high schools in the United States. *Int. J. Educ. Math. Sci. Technol.* 6, 320–342.
- Peters-Burton, E.E., Lynch, S.J., Behrend, T.S., Means, B.B., 2014a. Inclusive STEM high school design: 10 critical components. *Theory Pract.* 53, 64–71. <https://doi.org/10.1080/00405841.2014.862125>.
- Plasman, J.S., Gottfried, M.A., 2018. Applied STEM coursework, high school dropout rates, and students with learning disabilities. *Educ. Pol.* 32, 664–696.
- Plasman, J.S., Gottfried, M.A., Klasik, D.J., 2021. Do career-engaging courses engage low-income students? *AERA Open* 7, 1–17.
- President's Council for Science and Technology, 2010. *Prepare and Inspire: K-12 Science, Technology, Engineering, and Math (STEM) Education for America's Future*. PCAST, Washington, D.C.
- Pride, L.D., 2014. Using learning stories to capture "gifted" and "hard worker" mindsets within a NYC specialized high school for the sciences. *Theory Pract.* 53, 41–47. <https://doi.org/10.1080/00405841.2014.862121>.
- Roehrig, G.H., Moore, T.J., Wang, H.-H., Park, M.S., 2012. Is adding the E enough? Investigating the impact of K-12 engineering standards on the implementation of STEM integration. *Sch. Sci. Math.* 119, 31–44.
- Rogers-Chapman, M.F., 2014. Accessing STEM-focused education: factors that contribute to the opportunity to attend STEM high schools across the United States. *Educ. Urban Soc.* 46, 716–737.
- Sahin, A., Oren, M., Willson, V., Hubert, T., Capraro, R.M., 2015. Longitudinal analysis of T-STEM Academies: how do Texas inclusive STEM academies (T-STEM) perform in mathematics, science, and reading? *Int. Online J. Educ. Sci.* 7, 11–21.
- Saw, G.K., Agger, C.A., 2021. STEM pathways of rural and small-town students: opportunities to learn, aspirations, preparation, and college enrollment. *Educ. Res.* 50, 595–606.
- Saw, G., 2019. The impact of inclusive STEM high schools on student outcomes: a statewide longitudinal evaluation of Texas STEM academies. *Int. J. Sci. Math. Educ.* 17, 1445–1457.
- Scott, C., 2012. An investigation of Science, Technology, Engineering and Mathematics (STEM) focused high schools in the US. *J. STEM Educ.* 13, 30–39.
- Sharma, J., Yarlagadda, P.K.D.V., 2018. Perspectives of "STEM education and policies" for the development of a skilled workforce in Australia and India. *Int. J. Sci. Educ.* 40, 1999–2022.
- Slavit, D., Nelson, T.H., Lesseig, K., 2016. The teachers' role in developing, opening, and nurturing an inclusive STEM-focused school. *Int. J. STEM Educ.* 3, 1–17. <https://doi.org/10.1186/s40594-016-0040-5>.
- Stehle, S.M., Peters-Burton, E.E., 2019. Developing student 21st century skills in selected exemplary inclusive STEM high schools. *Int. J. STEM Educ.* 6, 1–15.
- Stoeger, H., Heilemann, M., Debatin, T., Hopp, M.D.S., Schirmer, S., Ziegler, A., 2021. Nine years of online mentoring for secondary school girls in STEM: an empirical comparison of three mentoring formats. *Ann. N. Y. Acad. Sci.* 1483, 153–173.
- Subotnik, R.F., Tai, R.H., Almarode, J., Crowe, E., 2013. What are the value-added contributions of selective secondary schools of mathematics, science and technology? Preliminary analyses from a U.S. national research study. *Talent Dev. Excell.* 5, 87–97.
- Subotnik, R.F., Almarode, J., Lee, G.M., 2016. STEM schools as incubators of talent development. *Gift. Child Today* 39, 236–241.

- Subotnik, R.F., Stoeger, H., Luo, L., 2019. Exploring compensations for demographic disadvantage in science talent development. In: Subotnik, R.F., Assouline, S.G., Olszewski-Kubilius, P., Stoeger, H., Ziegler, A. (Eds.), *The Future of Research in Talent Development: Promising Trends, Evidence, and Implications of Innovative Scholarship for Policy and Practice*, New Directions for Child and Adolescent Development, vol. 168, pp. 101–130.
- Szostkowski, A., Upadhyay, B., 2019. Looking forward by looking back: equity in science education as socially just and morally healing action. *Cult. Stud. Sci. Educ.* 14, 335–353. <https://doi.org/10.1007/s11422-019-09916-z>.
- Tardy Hackman, S., Zhang, D., He, J., 2021. Secondary school science teachers' attitudes towards STEM education in Liberia. *Int. J. Sci. Educ.* 43, 223–246. <https://doi.org/10.1080/09500693.2020.1864837>.
- Teo, T.W., Osborne, M., 2014. Understanding accountability from a microanalysis of power dynamics in a specialized STEM school. *Crit. Stud. Educ.* 55, 229–245.
- Thomas, J., 2000. First year findings: NCSSMST longitudinal study of gifted students. *NCSSMST J.* 5, 4–6.
- Tofel-Grehl, C., Callahan, C.M., 2014. STEM high school communities: common and differing features. *J. Adv. Acad.* 25, 237–271.
- Tripp, J.N., 2021. *Towards a Framework of Liberatory Science Education Opportunities for Minoritized Students in an Urban, Inclusive STEM-Focused High School (ISHS)*. Ph.D. thesis. University at Buffalo, SUNY.
- Vaval, L., Bowers, A.J., Snodgrass Rangel, V., 2019. Identifying a typology of high schools based on their orientation toward STEM: a latent class analysis of HLS:09. *Sci. Educ.* 103, 1151–1175.
- Wang, J., Yang, M., 2021. A study on the relationship between preservice STEM teachers' beliefs about migrant students and teachers' roles in Chinese urban schools. *Educ. Urban Soc.* 53, 206–230.
- Weis, L., Eisenhart, M., Cipollone, K., Stich, A.E., Nikischer, A.B., Hanson, J., Ohle Leibbrandt, S., Allen, C.D., Dominguez, R., 2015. In the guise of STEM education reform. *Am. Educ. Res. J.* 52, 1024–1059. <https://doi.org/10.3102/0002831215604045>.
- Wiswall, M., Stiefel, L., Schwartz, A.E., Boccardo, J., 2014. Does attending a STEM high school improve student performance? Evidence from New York city. *Econ. Educ. Rev.* 40, 93–105. <https://doi.org/10.1016/j.econedurev.2014.01.005>.
- Yoon, S.Y., Strobel, J., 2017. Trends in Texas high school student enrollment in mathematics, science, and CTE-STEM courses. *Int. J. STEM Educ.* 4, 9.
- Young, V.M., House, A., Wang, H., Singleton, C., Klopfenstein, K., 2011. *Inclusive STEM Schools: Early Promise in Texas and Unanswered Questions*. Paper presented at the National Research Council Workshop on Successful STEM Education in K-12 Schools. Available at: http://www7.nationalacademies.org/bose/STEMSchools_Workshop_Paper_Young.pdf.
- Young, V.M., House, A., Sherer, D., Singleton, C., Wang, H., Klopfenstein, K., 2016. Scaling up STEM academies statewide: implementation, network supports, and early outcomes. *Teach. Coll. Rec.* 118, 1–26.

Virtual reality and science, technology, engineering, and mathematics education

Richard Lamb, East Carolina University, College of Education, Neurocognition Science Laboratory, Greenville, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	189
Virtual reality in STEM education	190
Immersion and interactivity two key factors	191
Use case	192
VR laboratory environment	192
VR and knowledge generation	193
Methods	194
Results	194
Theme 1, technological savvy	195
Theme 2, life like details in the VR simulations	195
Theme 3, content appropriateness	195
Theme 4, STEM practice development and difficulty appropriateness	195
Discussion	196
Conclusion	196
References	197

Introduction

Virtual reality is a part of a family of technologies designed to create a synthetic or artificial environment. Technologies related to virtual reality include augmented reality, mixed reality, and “flat” two-dimensional environments. Virtual Reality (VR) is defined as a three-dimensional synthetic digital environment which allows multiple degrees of freedom for the user to interact with the environment and engage in immersive interactions within these environments. While there are several types of digital environments which are related to VR e.g., Augmented Reality, Mixed Reality, etc., VR is the only digital environment which is fully synthetic and completely immersive.

Immersion is one of the most critical characteristics of VR and related technology. It is immersion or the ability to become emotionally, cognitively, and behaviorally engrossed in the environment which creates the specific affordances of VR that promote the learning outcomes seen when using VR. Environments within VR may be ultra-realistic and almost indistinguishable from real world interactions. Lamb among others in multiple studies has illustrated that psychophysiological and neuro-cognitive responses of persons using VR environments very closely mirror the responses of the same person in similar real-world environments (Lamb et al., 2018a,b,c, 2021; Minderer et al., 2016). It is the similarity of the cognitive and sensory responses that helps to facilitate STEM learning using VR by triggering the same cognitive and affective responses as if they were in the real-world.

While there is tremendous value in using VR to explore realistic environments which may otherwise be unreachable by the user, VR has even greater value in the exploration of environments which are not typically encountered in the real-world. These environments could include visualizations and interactions within microscopic worlds e.g., at the atomic scale, interactions within environments which would be dangerous in the real-world e.g., walking at the bottom of the ocean, or interactions with environments which may be otherwise unreachable such as the surface of another planet. In many ways, VR is an important tool to provide experiences and access to almost any environment and interaction imaginable. The broad applicability of VR, coupled with the experiential nature of VR, makes VR a potentially transformative tool in STEM Education and a key area of exploration for educational researchers. The immersion, experiential nature, and ability to manipulate the environment drives much of the successes seen with VR and similar tools.

Current foci in STEM Education disciplines require students to develop sophisticated abstract understandings, manipulate multiple ideas simultaneously, integrate complex and abstract ideas, and align observations in order to learn content and practices in STEM disciplines. However, some of the opportunities to engage in environments which promote these activities are impossible to replicate in educational settings. To address the teaching and learning of complex and abstract ideas, educators focus on pedagogies which stress sense-making through connections to prior experiences and extrapolation to observed phenomena through mental models. These mental models often use representations and simplifications to help ground understandings in concrete and discrete terms particularly for younger students. The use of concrete and discrete terms and ideas links the abstract concepts in STEM content to student everyday experiences. The everyday experiences become the

building blocks of the mental-models students use to understand the world around them. From these mental models students then explain the world around them and generate understanding (Johnson-Laird and Khemlani, 2017). To this end, students in all disciplines, but in STEM disciplines specifically, learn from experiences by interacting with practices and content through their senses. The technologies used to create VR environments replace the stimulus from the real-world with digital stimulus which is more able to be personalized by the content creators and even the student. The personalization of the stimulus allows the student to individualize their experiences and promotes development of mental models which are more easily accommodated by the student. The mechanism by which this is thought to occur is cognitive priming and affective priming for learning (Lamb et al., 2015a). In other words, VR sets the conditions for the priming of specific neurocognitive systems by allowing the student a greater degree of control and immersion which improves learning opportunities.

Virtual reality in STEM education

The approaches and reasons to use VR in STEM education are often similar to the reasons an educator may use a two-dimensional, computer-based simulation. These reasons include the presence of a soft-failure environment, reduction in the cost of materials, commonality of experiences across students, and the ability to provide experiences not otherwise available (Protopsaltis and Baum, 2019). However, in addition to the reasons identified for using a computer-based instructional approach, VR offers additional affordances beyond these. VR has the potential to create individualized and custom learning environments which can accommodate the distinctive needs of the learning. Through this immersion and individualization exploration and discovery are promoted, ultimately this will motivate, encourage, and excite students about STEM discipline content increasing engagement with the content. The learner in the VR context can take part in the learning environment with a sense of embodiment which is not available in flat screen interactions. Conceptually, VR provides the following affordances: (1) VR is a first-person concrete representation of STEM discipline phenomena and ideas that can align to authentic content and practices within STEM disciplines. When combined with specific pedagogies VR helps students to learn complex and abstract material. (2) VR provides opportunities for experiences that may not be obtained by learners in other ways due to cost, complexity, or risk. (3) VR makes use of first-person embodied experiences, and this form of experience tends to make up the majority of our non-classroom-based interactions related to our environment. In contrast STEM classrooms, particularly for older students, tend to focus on third-person abstract and representational experiences reducing the construction of understanding and sense-making.

From an understanding construction and sense making perspective, VR and related technologies promote learning through the ability of the user to manipulate the scale of the experiences and move from the microscale to the macroscale. Through transformations at each of these scales, new sources of information and understanding may be experienced. VR is one of the only tools (outside of actual experiences) that allow students to learn from non-symbolic, immersive, first-person experiences.

VR as a tool for learning allows students to access the symbol systems associated with the content and practices of STEM disciplines in ways that work the best for them. VR promotes success in learning for students who may need to make use of alternative means of learning because written content does not allow them to access the content in a meaningful way. Pedagogically VR provides new forms and methods of content visualization and sensory stimulation. This occurs by drawing on not just the strengths of visual representations but the ability of VR to integrate multi-modal and cross modal cues related to the content reducing cognitive demand (Lamb et al., 2015a,b). VR also provides multiple alternate methods for content and practice presentation alongside the ability to manipulate the content in three-dimensional space. For example, a student may hold the planet Jupiter and Earth in their hand and directly compare the sizes of the planets. This allows the student to conceptualize the true proportions of the planets at a scale that makes sense to them. VR can also be used to clarify features, processes, and activities using novel means such as data rotations or other “physical” manipulations in VR environments. These manipulations allow close-up examination of an object to understand subsidiary structures, observation from a great distance to see overall patterns, and observation and examination of areas and events across time or that are unavailable by other means.

From an affective perspective, VR motivates and engages students by providing perspectives which are engaged with in novel ways. The first-person nature of VR requires interaction and encourages active participation in learning rather than passive participation as students engage with the experiences. Some types of virtual reality such as multi-user VR encourage and require collaboration providing socialization and interaction with others in ways that are not available in the real-world or on flat screens. Differentiation of experience and customization of learning opportunities in VR allows the learner to proceed through an experience at their own pace and in locations outside of the classroom. Importantly these features of VR allow students with differences to participate in learning environments which if they were in the real-world may limit their interactions. Because of the symbolic and representational nature of VR it can also transcend language barriers allowing students with English as an L2 to make sense of phenomena in their own way. The pedagogical affordances of VR in summary are: (1) visualization and reflection; (2) an alternate method for presentation of material; (3) learning in contexts impossible or difficult to experience in real life; (4) motivation, collaboration, adaptability, and differentiation; and (5) evaluation and assessment occurring in near authentic environments and with near authentic activities.

Immersion and interactivity two key factors

To fully take advantage of the affordances of VR two forms of immersion must be present: sensory immersion and psychological immersion. These two distinct forms of immersion help to create successful learning experiences within VR environments by helping to create physical sensations and psychological connections within the VR environment (Lamb and Etopio, 2019a,b). Sensory immersion occurs when students move or interact in some “physical” way within the VR environment. The visual, auditory, and kinesthetic feedback from the digital environment to the student’s sensory memory systems creates the immersive illusion that the student experiences as being bodily present in the environment. To accomplish sensory immersion visual, auditory, or haptic devices related to the environment must respond in a realistic, timely, and meaningful way, making use of inputs from the user. Ideally the sensory experiences are physiologically the same as those found in real-life thus triggering the same neurocognitive systems (Ferreira and Mavroidis, 2006). When VR environments are developed with both sensory immersion and pedagogical approaches in mind, users can interpret visual, auditory, and haptic cues to gather information and move toward specific learning outcomes (Lamb et al., 2018a,b,c; Lamb, 2016; Liu et al., 2020).

While sensory immersion results from engagement with the physical senses, mental immersion refers to the “state of being deeply engaged” psychologically within a VR environment (Huang et al., 2010). Mental immersion is critical to the success of VR as with on sensory immersion there is still the understanding that the environment is not “real,” and this may reduce the experiences and not trigger the appropriate neurocognitive systems. While sensory immersion is used to develop the mental immersion, mental immersion is how a VR system engages the students in learning about STEM discipline content. Essentially, the user “forgets” they are in a digital virtual environment and the interaction then becomes embodied. For VR environments designed for learning related to STEM disciplines, a significant marker of the success of the environments is predicated upon how involved the user becomes within the learning environments. One potential way to measure this marker in real-time is through psychophysiological and neurocognitive measurements which when combined with machine learning classification can adapt the environment in real-time (Lamb et al., 2022).

A second key feature of VR that helps to distinguish it from other educational technologies, particularly flat screens, is the real-time interactivity and responsiveness in digital stereoscopic three-dimensional environments. Real-time refers to the ability to make use of data with little (millisecond) or no delay. The fluidity of the interactivity with the digital environment helps to create the immersive experiences which are the fundamental attraction of VR for educators. VR as a system can interpret gestures and other actions and respond as if the digital environment is a real-life physical environment. The responsiveness occurs with relatively little lag resulting in a more authentic, true-to-life experience for the student. Interactivity coupled with the ability of VR to respond realistically promotes the sensation of sensory immersion by responding to user behaviors with actions on the screen. Auditory, haptic, and tactile inputs (Klatzky et al., 2014) allow users to not only interact visually with the environment but to manipulate (i.e., touch and feel) graphic objects on the screen. The ability to manipulate items sets VR apart. In short, VR technology is well suited to convey difficult abstract concepts, practice skills, and interact in complex environments due to the visualization, fluidity, interactivity, and immersion of the environment (Psotka, 2013; Lamb and Firestone, 2018). Importantly because the VR environments are fully digital, they can be programmed and shaped in any way necessary to promote learning. When educators design an environment for delivery of complex concepts, it is necessary to ensure the presence of the three features of (1) interaction, (2) immersion, and (3) authenticity in the learning environment. Fig. 1 illustrates the continuum of the relationship between interactivity and immersion as it relates to VR and related technologies. Authenticity is simply the need to make the interaction and environment as realistic as is needed to trigger the neurocognitive and sensory systems to promote learning (Fig. 2).

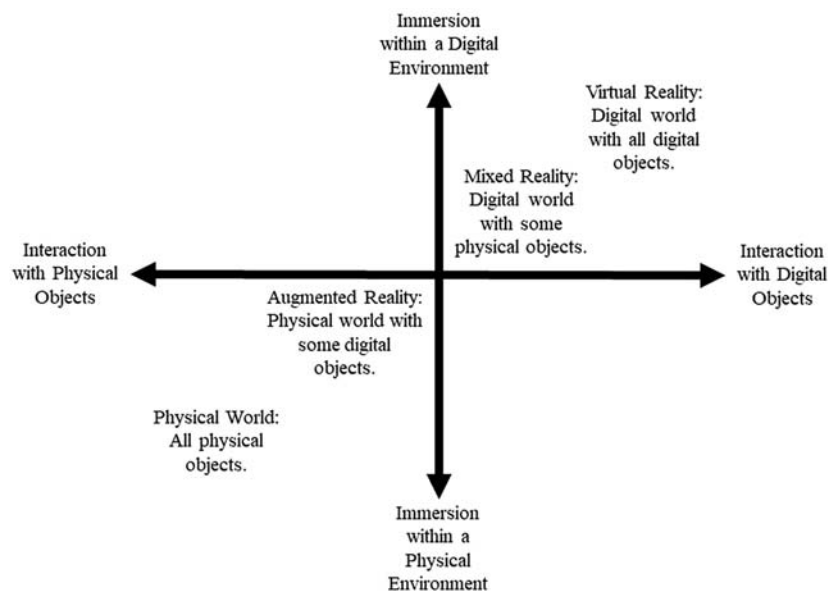


Fig. 1 Overview of VR in relation to other similar technologies.



Fig. 2 Example VR laboratory views available to the student.

Use case

This use case seeks to explain barriers and affordances related to content learning and immersion when using VR designed to replicate a hands-on laboratory environment for educational purposes. The primary means for data collection associated with this use case is a combination of multiple-choice survey combined with open-ended interviews.

VR laboratory environment

The VR laboratory used in this study was developed as a means to supplement and replace existing laboratories in the biology and chemistry. The VR laboratories are fully digital and do not make use of 360-degree camera footage. The VR classroom was experienced using a standard HTC Vive Virtual Reality headset. The HTC Vive was not specifically modified for the interactions apart from the use of noise canceling headphones. The VR headset allows a room scale of $12' \times 12'$ virtual environment with 6 degrees of freedom. The screen resolution was 2160×1200 and uses a refresh rate of 90 Hz. The headset equipment consists of the headset, two wall-mounted sensors for bodily positioning, and handheld motion controllers to position the hands. The computer used to generate the environment was a Dell G5 Gaming Desktop with a virtual reality ready video card.

Two primary views of the laboratory were available to the students. The first view is an overall view of the laboratory used to access various pieces of equipment within the laboratory. Students interacted with this view by using the thumb controller to move an avatar around the laboratory and to position the avatar at a specific bench for work. The second view available to the student once they reached a specific bench area was a first person view with two hands. The two hands were controlled by the handheld controls. The digital hands could pick-up items, move items, pour items, and otherwise are used to interact with the environment. In addition to using the controllers to move the hands students could also walk short distances while at the bench to reach other items. Sounds associated with the use of various laboratory equipment were also present in the environment. Overall, the environment is considered an open-world sandbox type environment. Meaning that the students could move around and complete tasks in any order. It should be noted as with real-life laboratories if the student did not appropriately follow the protocols the laboratory did not produce results. [Fig. 1](#) provides a visual depiction of the laboratory.

For this use case, five laboratories were selected for study: three chemistry and two biology. The specific five laboratories were selected by classroom instructors as they felt these topics were most aligned with the needs of the course. The three chemistry laboratories covered the topics of acids and bases (26 min), atomic structures (34 min), and basic chemistry thermodynamics (42 min). The biology laboratory topics consisted of animal genetics (53 min) and antibodies and blood type (37 min). [Table 1](#) provides a summary of the five laboratories and the specific concepts covered. During the completion of each laboratory the students were able to keep notes about the results, redo experiments as often as they like, and upon completion were assessed using a multiple-choice assessment.

Table 1 Overview of topics covered in this use case.

<i>Laboratory</i>	<i>Concepts covered</i>
Acids and bases	Measure of pH and evaluation of mixtures of acids and bases.
Atomic structure	Properties of element samples, atoms, and ions. Prediction of properties of elements based upon periodic table position.
Basic chemistry thermodynamics	Bomb calorimetry, concepts of thermodynamics, renewable energy.
Animal genetics	Mendelian inheritances, gene linkage analysis, mutations, and gene expression.
Antibodies and blood type	Antibodies antigen interactions, ABO and Rhesus blood groupings, risk for Rhesus disease, blood transfusions.

VR and knowledge generation

VR in the context of this study is thought to construct knowledge through experiences and provide the ability to examine phenomena at multiple scales. The selected laboratories were developed with a knowledge construction approach using a user experience construction lens. By providing a fully immersive environment with the sights, sounds, and tactile feedback the intent was to fully simulate a laboratory experience. To accomplish knowledge construction through a VR environment, VR designers must create activities which allow learners to take an active role in their learning through fluid and immersive environmental interactions (Vygotsky, 1967). The intention of creating a fluid and immersive environment is to sufficiently simulate a laboratory experience such that students can connect current information with prior knowledge (Vygotsky, 1967). To accomplish this connection the environment and specific design elements such as the interactions with the laboratory equipment, the concepts explored, and the practices associated with the content must be sufficiently realistic. Sufficiently realistic means that the environment is sufficiently real enough to trigger the same neurocognitive response as would be triggered by the real-world equivalent action or environment. The realism and responsiveness of the environment plays a critically important role in learning (Kintu and Zhu, 2016). When responsiveness and sufficient realism are not present this can lead to affective priming (e.g., Lamb et al., 2015a,b) and ultimately to frustration and lack of interest in completing the laboratory. By providing realism and responsiveness through immersion and interactivity, students have the opportunity to directly apply their practices and content to a context which provides meaningful feedback and practice. To successfully learn, students must be allowed to directly apply their knowledge to realistic and authentic contexts, engage in failure, and explore their successes and failures in a safe low consequence environment without the structures of the experience e.g., the VR system complicating the interaction (Stefaniak et al., 2020).

Technologies such as VR allows learners to engage in experiential and exploratory learning within low consequence environments. In the case of these laboratories the students could redo the laboratories multiple times. VR fills a core need within STEM education; that learning should be experimental and experiential (Long et al., 2020). Using VR, the educator's role is to help shape and facilitate the learners' experiences and understand how the surrounding digital environment promotes or impedes necessary learning (Ebenezer et al., 2020). Considering this responsibility, it is important that educators consider the user experience when implementing VR in the classroom. Within the user experience the educator should focus on how the content and practices are developed in VR, and how the learner's control of the learning process impedes or assists their learning and task completion. Therefore, VR designs for education must assume that not all learners can formulate their current knowledge as discrete concepts and tie it to experiences and authentic tasks. Designers must scaffold and provide other adaptive supports which meet the needs of varying levels of learners. A user experiences centered approach to VR provides learners more freedom to select and coordinate their learning processes and to construct knowledge through these means. User experience centered constructivists emphasize the design of learning environments rather than instructional sequences and focus on the individual differences of the learner. The learning environments should provide real-world, case-based environments for meaningful and authentic knowledge construction with minimum system friction.

In the case of VR for STEM education, it is important to not only focus on the user experience but also provide educational means and opportunities to examine questions, claims, and evidence. The inclusion of the pedagogical component is a key differentiator between user experience researchers outside of education and user experience researchers within educational contexts. Current research related to educational technology suggests that user experience centered constructivist principles fundamentally underlie learning in an VR environment (Levin and Tsybulsky, 2017; Lamb and Etopio, 2019a,b; Lamb et al., 2018a,b,c). User experience centered constructivist learning in VR has multiple characteristics: (a) the learning involves the exploration, internalization, and discovery within the prebuilt, interactive, immersive representation of the real-world, (b) engagement with prior knowledge, (c) learning processes that allow educators to examine VR features which support learning, and (d) deep consideration of how the user interface and related software and hardware influences learning outcomes and success.

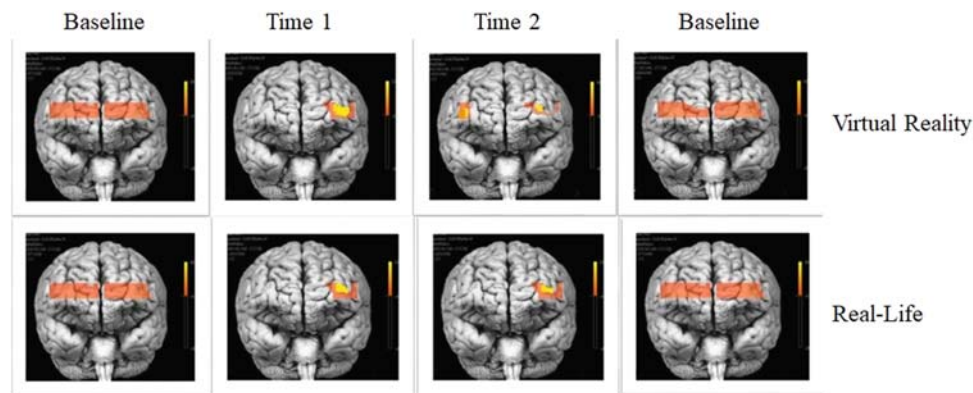


Fig. 3 Neurocognitive comparison of VR and real-life activities. Images supplied by the East Carolina University Neurocognition Science Laboratory.

VR allows learners to interact with simulated environments in real time and engage with soft failure (Nelson, 2016). VR also offers greater sensory cueing and multimodal feedback to enable the transfer of concepts learned in VR into real-world understandings related to the VR topics (Smith et al., 2020; Lamb, 2016). For example, learners can view 3D objects from multiple representations, viewpoints, and scales in addition to examining and exploring interactions and relationships between the representations and concepts. VR used in more traditional classroom learning environments allows educators to provide experiences which otherwise would not be possible in STEM classrooms (e.g., walking inside of a volcano or seeing a virus infect a cell). Further, immersive environments create a strong sense of presence in the environment, which in turn motivates and thereby causes the learner to cognitively process the learning material more deeply (Shu et al., 2019). A neuroimaging study conducted by Lamb (2019) reveals that when learners interact with VR environments the learners' cognitive systems process the VR immersive environment in the same way that real-world environmental cues are processed (See Fig. 3).

Methods

This use case was based upon focus group discussions and user responses on surveys collected during and after the use of the VR environment in multiple STEM classrooms. One-hundred faculty, representing 300 classrooms, and approximately 1000 students took part in this use case. A subsample of three instructors and three students were randomly selected for the focus groups. The five modules were selected by the instructors based upon alignment to current topics covered in class. VR outcomes were presented as relationships between characteristics of performance, which either impeded or promoted learning for the students in the classroom. Using thematic inquiry, the researchers were able to generate several themes. Data was summarized and used to identify potential relationships between the emergent themes. The VR environment is an immersive 3D problem-based learning virtual simulation of five laboratories. The simulation makes use of realistic 3D animations and interactions to address the hands-on learning goals the instructors have for their class. The VR animations and simulations allowed students, for example, to examine molecular scale representations as learners engage in the performance of experiments to reinforce content learned within non-laboratory components of the class. Upon completion of the VR tasks, instructors and students were asked to attend a focus group and complete a survey.

The areas covered in the focus groups and survey were intended to examine components of immersion and fluidity among the students and instructors through a user experience lens. The areas of user experience are Presentation of Technical information (Immersion), Laboratory Work Mechanics (Fluidity), Learning Outcomes (Immersion), Student Engagement (Immersion and Fluidity), and Ease of Navigation (Fluidity). Instructors and students rated these areas on a 1 (Poor) through 7 (Excellent) scale. Responses to the focus group questions were noted and triangulated with survey responses to develop the themes more fully.

Results

The most common theme emerging from the focus group was a relationship between the technological skill level of the student with respect to the use of VR and the student's history with online learning management software. Results also illustrate that a significant technological barrier existed. This technological barrier increased the difficulty and utility of VR when considered for the full implementation in the replacement and supplementing of hands-on laboratory work. This is as opposed to conceptual barriers meaning that the ideas associated with the presentation of content and practices in this VR environment met the needs of the students, but the

user experience in terms of the immersion and fluidity are not fully developed and require a technological solution e.g., less lag in the response, more articulation with the controls, etc. User responses show that instructors and students identified Student Engagement (Immersion and Fluidity) and Ease of Navigation (Fluidity) as the most impactful and positive aspects of the VR experience. The inclusion of the two factors “Ease of Content Navigation” and “Student Engagement” is substantiated because inclusion and specific focus on these two factors of VR experience are shown to be very influential in learning outcomes and the likelihood of persisting through difficulties. Survey and focus group outcomes were integrated to generate themes.

Theme 1, technological savvy

Participants using the VR environment suggested that their savvy with technology played a major role in their acceptance of VR as a means for teaching and learning in STEM disciplines. Particularly because technological savvy allows students to engage with the material in a more fluid and immersive way while ignoring or correcting items which detract from the fluidity or immersion. Technologically savvy students were able to continue through the content despite having to do “on-the-fly” corrections due to glitches within the laboratory. Those with a lack of technological savvy, meaning they had difficulty using input devices, difficulty working with computers, lack of basic knowledge of computer troubleshooting, and did not have high levels of content skills, found working with the VR software extremely difficult. This was because they said it [*the user interface*] when combined with trying to learn new material overwhelmed them. The students could not focus on learning the STEM discipline content because they were too focused on trying to figure out how to correct aspects of the VR environment which did not operate properly. From a neurocognitive perspective, this aligns with previous studies illustrating that there are limited cognitive resources and when threshold levels are exceeded students become frustrated and disengage (Lamb et al., 2018a,b,c). Exceeding cognitive demand levels also reduces a student’s level of motivation to engage with the software due to frustration with the platform.

Theme 2, life like details in the VR simulations

There are two major aspects of the VR experiences which reduced learning outcomes in the VR experience. The first item was the lack of a fluidity and the loss of psychological immersion. Experiences in VR must be mediated through a responsive computer interface and when the interface does not respond in an expected way it creates difficulties in learning. This occurs because the “incorrect” response by the environment must be accommodated. To accommodate the response, the primary cognition used for the process of learning must be reallocated to incorporating the incorrect response breaking the psychological immersion. The second factor was lack of sufficient realism in the VR experience which breaks the sensory immersion. Without sufficient realism, sensory stimuli from the VR system are misaligned creating cognitive dissonances which must be accommodated. Some of the difficulties for the students resulted from a lack of normally anticipated things found in an environment e.g., flipping a switch and nothing happening, or attempting to pick-up an object and not having the object behave as expected. The software also seemed to lack basic science equipment found in traditional life science classrooms such as a microscope, though a microscope is visible in the comprehensive view of the laboratory. However, there was an appreciation for the way safety procedures are represented and incorporated into the software. Students even within the VR environment still had to don and doff personal protective equipment helping to improve sensory and psychological immersion. However, despite these positives, the limitations associated with the simulation equipment, (i.e., the program crashing, lack of tactile feedback, reduced interactivity, and lack of objects to engage in practices) created difficulty for the immersive and fluid aspects of the experience, leading the students and instructors to see the VR experiences as less than life-like and therefore not as good as the physical real-life experiences.

Theme 3, content appropriateness

Content at the appropriate instructional level and that also aligned specifically with learning objectives was a challenge. The lack of alignment reduced the psychological immersion of the VR simulation. Students often became “lost” in attempting to link the presented content to their existing knowledge and expectations. While a positive attribute of the VR experiences is that there was a reduced need for physical space and increased opportunities to engage with the content multiple times, this did not seem to make up for the other areas of concern noted by the instructors and students. Meaning the lack of alignment between the presented content and the expected content reduced the immersive qualities of the VR environment. Thus, the benefit of VR from a cost perspective did not outweigh the lack of sufficient realism and appropriate content. Specifically, the educators felt that practice skill development was deeply hindered by the lack of virtual reality interactivity and sufficient realism.

Theme 4, STEM practice development and difficulty appropriateness

The instructors felt that the VR did not successfully replicate physical world experiences sufficiently. However, the students while agreeing with the instructors, they did illustrate more optimism about the potential of VR. The instructors suggest that may lead to the idea that the inclusion of VR as a supplemental form of instruction. While they see some potential the instructors suggested they need to deeply consider how to make use of VR and specifically they suggested that as it stands the VR environments within this study need improved fluidity and immersion to more fully replace hands-on laboratories.

Discussion

The general assessment of VR for the teaching and learning of STEM content and practices is, that without significant technological improvements VR fails to meet the student's needs associated with studies in STEM disciplines. Specifically, the technological improvements should be focused on increasing the levels of immersion and fluidity, particularly around more sufficiently realistic responses within the environment. The improvements as suggested by students and instructors results from difficulties associated with navigation of the environment (fluidity) and engagement (immersion and fluidity). In the case of this study the VR environment, the learning activities, tasks, and assessments were designed without specific pedagogical approaches embedded into the VR as suggested by [Lamb et al. \(2018\)](#). The lack of alignment of the user experience to specific and explicit pedagogies and learning outcomes led to a disjointed and unstructured learning environment in which outcomes were not able to be anticipated by the users. As instructional designers or educators develop and deploy specific features of VR as a replacement for face-to-face instruction there is a need to consider the student user experience. Importantly they must consider what factors play a role in successful use of the VR environment for learning and how to build immersion and fluidity into the environment to reduce frustration, student training time, instructor training time, and have sufficient infrastructure to support the VR laboratories.

One of the most promising aspects of the use of VR in the classroom is the ability to develop interactive, highly controlled, ultra-realistic, learning environments and experiences which are specifically tailored to those using VR. Appropriate pedagogical approaches involving construction of knowledge in a virtual environment requires that interactions with the VR environment have appropriate levels of fluidity, immersion, and authenticity allowing realistic interactions which trigger the necessary cognitive systems for learning. Without this critical level of fluidity of interaction with the environment the learning process is too difficult to sustain, and the user may not persist. If developed with appropriate levels of fluidity and immersion VR environments can be a reasonable and valuable substitute or supplement for real life experiences ([Lamb and Etopio, 2019a,b](#)). However, as in the case of this study, affective aspects of the interaction, such as frustration, will quickly impede the learning process. Unfortunately, these barriers and impediments develop quickly when working with difficult user interfaces such as the one in this study.

During the VR interactions learners attempted to undertake actions allowing them to put new understandings and new skills into practice, however the attempts to make use of the learned skills in this VR environment were unsuccessful due to the interface and its lack of immersion and fluidity. This lack of success created a negative view of the VR application and reduces the student desire to participate in the environment. In considering future designs and interactions, these interactions must occur in a life-like and measured way. While VR environments may allow learners to acquire knowledge with less difficulty than that of traditional learning processes, poor organization of the environment will have the opposite outcome.

The transfer of skills from the VR environment to real world environments is of critical concern to educators. To accomplish this transition from one environment to the other, VR environments require immersion and realism along with the ability to construct knowledge through interactions. Static VR environments are not of much use and will only act in much the same way a photograph will for learning purposes. This is because of the lack of immersion within the environment.

Immersion and fluidity allow VR tools to be used for similar tasks and reasoning as the tasks and tools found in the real world ([Lamb, 2016](#); [Lamb et al., 2018a,b,c](#)). As a result, VR environments provide rich teaching opportunities and help to improve learners' ability to analyze problems and explore new concepts associated with the environment. The multisensory aspect of the VR technology promotes greater learner immersion by prompting attention and stimulating curiosity. In the case of this VR environment, the multisensory aspect of the environment impeded learning due to the lack of high-quality visuals, poor performance of the application, poor assessment, and poor feedback from the environment e.g., lack of fluidity. That is, features of interaction and immersion will only take a student so far into the environment, if the environment is not fluid and highly responsive. Fluidity and responsiveness are the main characteristics, which maintain engagement with the VR environment ([Lamb et al., 2016](#)).

Conclusion

While this VR environment was missing instructionally important characteristics such as feedback and successive skill development, results from this study provide insight for the exploration of needed characteristics for future iterations of the VR environment. In particular, this study illustrates how to consider inclusion of fluidity and immersion as central themes to the user experience during development. As more students and instructors focus on the use of VR technology and VR applications for education, content will become easier to use and incorporate *a priori* pedagogical approaches with fluid and immersive environments. To promote the use of VR for learning, educators need to understand the challenges students face when using VR technology for instruction for the first time and understand the limitations of the environments rather than counting on the novelty to maintain and promote learning outcomes in STEM disciplines. It is imperative that an instructor making use of VR, keep in mind, that VR is another tool to promote learning and not meant to replace the instructor.

References

- Ebenezer, J., Kaya, O.N., Kassab, D., 2020. High school students' reasons for their science dispositions: community-based innovative technology-embedded environmental research projects. *Res. Sci. Educ.* 50 (4), 1341–1365.
- Ferreira, A., Mayrroidis, C., 2006. Virtual reality and haptics for nanorobotics. *IEEE Robot Autom. Mag.* 13 (3), 78–92.
- Huang, H.M., Rauch, U., Liaw, S.S., 2010. Investigating learners' attitudes toward virtual reality learning environments: based on a constructivist approach. *Comput. Educ.* 55 (3), 1171–1182.
- Johnson-Laird, P.N., Khemlani, S., 2017. Mental models and causation. In: *Oxford Handbook of Causal Reasoning*, pp. 1–42.
- Kintu, M.J., Zhu, C., 2016. Student characteristics and learning outcomes in a blended learning environment intervention in a Ugandan University. *Electron. J. e-Learn.* 14 (3), pp181–195.
- Klatzky, R.L., Giudice, N.A., Bennett, C.R., Loomis, J.M., 2014. Touch-screen technology for the dynamic display of 2D spatial information without vision: promise and progress. *Multisensory Res.* 27 (5–6), 359–378.
- Lamb, R.L., 2016. Examination of the effects of dimensionality on cognitive processing in science: a computational modeling experiment comparing online laboratory simulations and serious educational games. *J. Sci. Educ. Technol.* 25 (1), 1–15.
- Lamb, R., 2019. Successful Use of a Novel Artificial Neural Network to Computationally Model Cognitive Processes in High School Students Learning Science. *Comput. rev. J.* 3, 425–435.
- Lamb, R., Etopio, E., 2019a. Preservice science teacher preparation using virtual reality. In: *Society for Information Technology & Teacher Education International Conference*, pp. 162–167. Association for the Advancement of Computing in Education (AACE).
- Lamb, R.L., Etopio, E., 2019b. Virtual reality simulations and writing: a neuroimaging study in science education. *J. Sci. Educ. Technol.* 28 (5), 542–552.
- Lamb, R., Firestone, J.B., 2018. Science Teacher Education as a Way Forward for Medical Schools: A Case for Medical Pedagogical Content Knowledge.
- Lamb, R., et al., 2018. Examining human behavior in video games: the development of a computational model to measure aggression. *Soc. Neurosci.* 13 (3), 301–317.
- Lamb, R., et al., 2019. A computational model of student cognitive processes while solving a critical thinking problem in science. *The Journal of Educational Research* 112 (2), 243–254.
- Lamb, R., Akmal, T., Petrie, K., 2015a. Development of a cognition-priming model describing learning in a STEM classroom. *J. Res. Sci. Teach.* 52 (3), 410–437.
- Lamb, R., Annetta, L., Vallett, D., 2015b. La interrelación de la creatividad, fluidez, pensamiento lateral y tecnología a la hora de diseñar Juegos Educativos Serios en un aula de ciencias naturales. *Electron. J. Res. Educ. Psychol.* 13 (2), 219–242.
- Lamb, R., Cavagnetto, A., Akmal, T., 2016. Examination of the nonlinear dynamic systems associated with science student cognition while engaging in science information processing. *Int. J. Sci. Math. Educ.* 14 (1), 187–205.
- Lamb, R.L., Annetta, L., Firestone, J., Etopio, E., 2018a. A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Comput. Hum. Behav.* 80, 158–167.
- Lamb, R., Annetta, L., Vallett, D., Firestone, J., Schmitter-Edgecombe, M., Walker, H., et al., 2018b. Psychosocial factors impacting STEM career selection. *J. Educ. Res.* 111 (4), 446–458.
- Lamb, R., Antonenko, P., Etopio, E., Seccia, A., 2018c. Comparison of virtual reality and hands on activities in science education via functional near infrared spectroscopy. *Comput. Educ.* 124, 14–26.
- Lamb, R., Fortus, D., Sadler, T., Neumann, K., Kavner, A., Annetta, L., 2021. Exploration of Teacher-Student Neural Coupling Occurring During the Teaching and Learning of Science.
- Lamb, R., Neumann, K., Linder, K.A., 2022. Real-time prediction of science student learning outcomes using machine learning classification of hemodynamics during virtual reality and online learning sessions. *Comput. Educ.* 100078.
- Levin, I., Tsybulsky, D., 2017. The constructionist learning approach in the digital age. *Creativ. Educ.* 8 (15), 2463.
- Liu, R., et al., 2020. Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons. *Br. J. Educ. Technol.* 51 (6), 2034–2049.
- Long, N.T., Yen, N.T.H., Van Hanh, N., 2020. The role of experiential learning and engineering design process in K-12 STEM education. *Int. J. Educ. Pract.* 8 (4), 720–732.
- Minderer, M., Harvey, C.D., Donato, F., Moser, E.I., 2016. Virtual reality explored. *Nature* 533 (7603), 324–325.
- Nelson, D., 2016. *Creating Disruptive Innovators: Serious Educational Game Design on the Technology and Engineering Spectrum*. Connecting Science and Engineering Education Practices in Meaningful Ways. Springer, Cham.
- Protopsaltis, S., Baum, S., 2019. Does Online Education Live Up to Its Promise? A Look at the Evidence and Implications for Federal Policy. Center for Educational Policy Evaluation.
- Psotka, J., 2013. Educational games and virtual reality as disruptive technologies. *J. Educ. Technol. Soc.* 16 (2), 69–80.
- Shu, Y., Huang, Y.Z., Chang, S.H., Chen, M.Y., 2019. Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Real.* 23 (4), 437–446.
- Smith, C., Pezent, E., O'Malley, M.K., 2020. Spatially separated cutaneous haptic guidance for training of a virtual sensorimotor task. In: *2020 IEEE Haptics Symposium (HAPTICS)*. IEEE, pp. 974–979.
- Stefaniak, J.E., Reese, R.M., McDonald, J.K., 2020. Design considerations for bridging the gap between instructional design pedagogy and practice. *J. Appl. Instr. Des.* 9 (3).
- Vygotsky, L.S., 1967. Play and its role in the mental development of the child. *Sov. Psychol.* 5 (3), 6–18.

Design-based education in STEM: for learners of the 21st century

Yong Ju Jung^a, Gi Woong Choi^b, and Soo Hyeon Kim^c, ^a School of Library and Information Studies, University of Oklahoma, Norman, OK, United States; ^b Instructional Design and Technology Program, School of Education, University of Cincinnati, Cincinnati, OH, United States; and ^c Department of Library and Information Science, School of Informatics and Computing, Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: learner as the designer	198
Foundation and rationale of design-based education	198
Theoretical foundation of design-based education	198
Rationale of design-based education: for the 21st century	199
Current applications of design-based education in STEM	200
DBE with engineering design and maker education	200
K-12 settings	200
Out-of-school settings	201
Higher education settings	201
DBE with computational thinking and STEM+C	201
K-12 settings (both school and out-of-school)	202
Higher education settings	202
DBE with digital literacy and new media creation	202
K-12 settings (both school and out-of-school)	203
Higher education settings	203
Interdisciplinarity of the current applications	204
Designing design-based education for STEM learning: principles and gaps	204
Conclusion	204
References	205

Introduction: learner as the designer

Design is the process of manufacturing products—tangible and intangible—through constant iterations of analysis, synthesis, and evaluation (Gero, 1999). Design, therefore, includes actions ranging from framing problems, setting goals, and making plans to creating (new) products and assessing them iteratively (Koh et al., 2015; Li et al., 2019; Purao et al., 2008; Razzouk and Shute, 2012). In the past, design was regarded as a specialized skill of a few with specific vocations and professions (e.g., fashion designer, game designer, instructional designer). Similarly, in educational contexts, the term design mainly belonged to the educators who designed curricula and learning activities. However, with the recent acknowledgment that suggest “everyone designs and everyone can design” (Li et al., 2019, p. 95), any learner could become agents who design (i.e., devise, create, assess) rather than users of designed products. Accordingly, design thinking—a way of thinking through the process of design—and design practices have been widely used for and by learners to empower their learning experiences, especially in science, technology, engineering, and mathematics (STEM) educational settings (Dym et al., 2005; Li et al., 2019).

We define design-based education (DBE) as educational theory and practice that integrate and transform design processes and design thinking into *learning* experiences, especially for STEM education. DBE has been called with different names such as learning-by-design (e.g., Puntambekar and Kolodner, 2005), design-based learning (e.g., Bekker et al., 2015; Won et al., 2015), engineering design (e.g., Donna, 2012), computational design process (e.g., Kafai et al., 2014), embedded design (e.g., Borge et al., 2020), or making (e.g., Sheridan et al., 2014) and embraces a wide range of curricula and educational programming through which learners participate in hands-on designing for and as part of their STEM learning. Despite its wide range of the activities, DBE often includes iterative, non-linear cycles of analyzing/interpreting, planning, creating, and assessing by the learners. In this article, we demonstrate the concept of DBE and its benefits that align with the 21st century education, review current examples demonstrating innovative movements in STEM education, and discuss suggestions for better applications of DBE.

Foundation and rationale of design-based education

Theoretical foundation of design-based education

While our definition of design-based education (DBE) is centered on the use of design, DBE in essence is based on various theories on learning. DBE is grounded upon socio-constructive and sociocultural perspectives rooted from Vygotsky (Vygotsky, 1978) that consider learning to occur through participation and active construction of meaning when learners interact with their surrounding

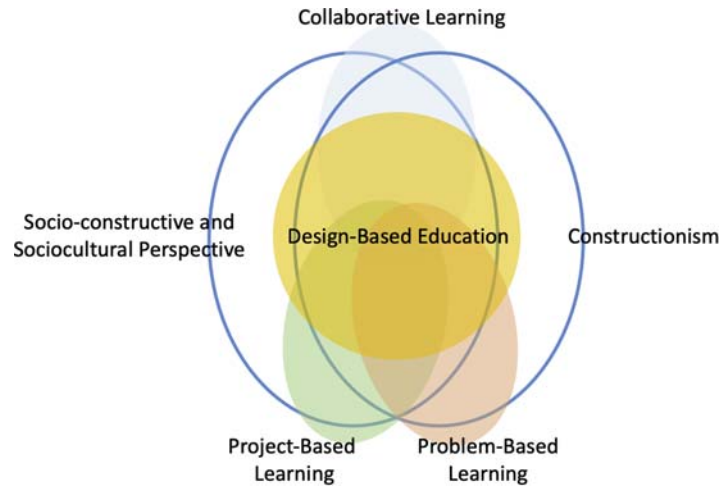


Fig. 1 Theoretical foundation of DBE: at the nexus of various perspectives and theories on learning.

contexts (Kolodner, 2004; Rogoff, 1994). Design process can be regarded as mediational means for learners to construct meaning. Incorporating design process into the core learning experience is also well-aligned with Papert's constructionism (Papert and Harel, 1991), which posits that learning is "building relationships between old and new knowledge, in interactions with others, while creating artifacts of social relevance" (Kafai, 2006, p. 35).

DBE also aligns with the theories of collaborative learning. Collaboration and social interactions are particularly relevant to design, as design in the real world often involves multiple stakeholders and experts. For instance, designing a smartphone requires collaboration between user experience designers, mobile application programmers, hardware engineers, and other stakeholders. DBE is built upon diverse approaches of collaboration, such as knowledge co-construction (Borge et al., 2020), negotiations (Jung et al., 2017), and peer supports (Melander Bowden, 2019) to empower people to learn with and from each other.

In addition, project-based learning and problem-based learning are connected with DBE. Project-based learning supports learners to be situated in their authentic projects, through which they can apply ideas and solve authentic problems (Krajcik and Shin, 2014). Similar to project-based learning, problem-based learning enables learners to be active agents who analyze and solve ill-structured and complex problems (Lu et al., 2014). The features of DBE overlap with both approaches because they are centered on learning-by-doing and inquiry-based procedure. DBE differs from either approach as it focuses on creating a product (either tangible or intangible), and iterative, numerous problem-solving occur during the design process. Acknowledging these commonalities and features of DBE, Barak and Assal (2018) developed the P3 framework combining practice, problem-solving, and project-based learning to support STEM DBE with designing robotics. Building upon these studies, as shown in Fig. 1, our framework illustrates that research and practice of DBE are at the nexus of various perspectives and theories on learning.

Rationale of design-based education: for the 21st century

The 21st century education is distinct in many ways from the traditional (20th century) education; it focuses on learner-centered pedagogy, critical thinking, collaboration, interest-driven learning, and interdisciplinarity (Thomas and Brown, 2011). The practical and theoretical foundations of DBE align well with these foci.

One of the main characteristics of the 21st century education is its emphasis on learner-centeredness (Thomas and Brown, 2011). The 21st century education centers around learners' active roles that entail what and how they think and do, which is different from instruction-based education that focuses on delivering content knowledge and positions learners as passive receivers. Hands-on design experiences, through which learners become creators, have been integrated in the form of project-based learning or constructive learning to facilitate learner-oriented activities in STEM education (Dym et al., 2005; Keune et al., 2015; Kim and Zimmerman, 2019).

The 21st century education also emphasizes higher-order cognitive skills such as critical thinking and scientific reasoning, which have been highlighted to be important skills for STEM learning. As learners apply these higher-order cognitive skills in the design process (e.g., analyze, plan, assess), DBE for STEM education can help learners develop essential 21st century skills (Razzouk and Shute, 2012). Furthermore, DBE can bridge between real-life cases and what they learn in classrooms (or other educational settings) (Koh et al., 2015) by situating learners within authentic STEM problems and providing a learning space to apply and train higher-order cognitive skills.

Moreover, collaborative skills, which is another important competence for the 21st century learners, can be empowered through DBE (Koh et al., 2015). The design process involves analyzing and coordinating diverse perspectives and solutions that may come

from different stakeholders, and this requires effective communication as well as collaboration. Similarly, for many professions in STEM fields, the ability to communicate and collaborate are essential skills for success. This means that DBE can help learners improve communication skills and help them practice collaborative problem-solving with other learners.

In addition to cognitive skills and competencies, DBE can also support emotional and affective aspects (e.g., interest and agency) of STEM learning. Developing interest and supporting interest-driven learning have been recognized as a key strategy for the 21st century education because interest can lead to deeper and voluntary engagement (Ito et al., 2013). Also, hands-on designing experiences have been shown to maintain learners' interest and develop epistemic agency (Jung et al., 2020). Therefore, DBE can support both cognitive and emotional growth in STEM of the 21st century learners.

Lastly, unlike the traditional education that was strict with the boundaries between disciplines, the 21st century education has focused on crossing diverse disciplines (i.e., inter- and multi-disciplinarity). Such shift led to emphasis on learners' ability to integrate knowledge from diverse areas. This is particularly true for STEM education because design process entails incorporating diverse perspectives that are derived from different people and fields. Therefore, DBE can bridge across the disciplines of science, technology, engineering, and mathematics for the integrative STEM education (Donna, 2012; Li et al., 2019).

Current applications of design-based education in STEM

Design-based education (DBE) has been applied to diverse educational settings in and out of schools across K-16, parallel to innovative movements in STEM education. In this article, we focus on the application of DBE in three educational settings: (a) engineering design and maker education, (b) computational thinking and STEM+C (computer science) education, and (c) digital literacy and new media creation. While we divide educational settings into three categories for the purpose of reviewing, we state that these categories are not mutually exclusive from each other. Rather, they overlap with one another because of the interdisciplinary and connective nature of DBE.

DBE with engineering design and maker education

Among different fields within STEM, engineering has been historically known to incorporate design thinking approach and design processes at their core. Accordingly, DBE for STEM has utilized the pedagogical approach of engineering design and making. Engineering design is a process practiced by engineers to identify and solve real-world problems. A framework for K-12 science education in USA (National Research Council, 2012) lists the following practices as part of engineering design process: (1) defining problems, (2) developing and using models, (3) planning and carrying out investigations, (4) analyzing and interpreting data, (5) using mathematics and computational thinking, (6) designing solutions, (7) engaging in argument from evidence, and (8) obtaining, evaluating, and communicating information. While engineering design continues to be integrated in STEM curricula within formal settings (Lachapelle and Cunningham, 2019; Moore et al., 2014), "making" has gained attention from educators within out-of-school environments as a way to support STEM learning in personally-relevant, playful ways with the rise of the Maker Movement.

The Maker Movement can be defined as participation in a community of practice that promotes creative production in art, science, and engineering with the use of digital and physical technologies in physical and/or virtual settings (Dougherty, 2012; Sheridan et al., 2014). With its origin in DIY and hacker culture, the Maker Movement is characterized by open exploration that values collaboration and distributed expertise. Research has shown that making can be one of the ways to engage learners to imagine STEM-related identity and future trajectories (Vossoughi and Bevan, 2014) as learners move toward full participation in the socio-cultural practices of the maker community. While comprehensive reviews of literature and best practices of the Maker Movement can provide more holistic overview of the learning outcomes (e.g., Association of Science-Technology Centers and Urban Libraries Council, 2014; Vossoughi and Bevan, 2014; Whyte, 2016), we particularly interpret it as part of the DBE pedagogy and present examples from K-12, out-of-school, and higher-education settings.

K-12 settings

The promising results of the makerspaces has led formal learning institutions to adopt them. An example of early efforts that introduced engineering design and making to formal STEM education settings is the development of FabLearn Labs. FabLearn Labs provides digital fabrication and high-tech equipment for construction to K-12 schools and university partners in 24 countries across six countries (Blikstein and Krannich, 2013). A case study of two FabLearn Labs in middle schools illustrated that problem-based learning approach in digital fabrication makerspaces, when positioned as part of the core curriculum rather than electives, provided a huge educational potential (Chan and Blikstein, 2018). However, the study also highlighted that self-directed and inquiry-driven STEM learning was dependent upon vigilant teaching and facilitation and carefully crafted design challenges and pedagogy.

When the provision of high-tech materials and a designated makerspace is not viable, public schools promote the tinkering disposition through creating design challenges that align with project-based learning principles. Purdue Polytechnique High School located in downtown Indianapolis is an example of how maker ethos is nurtured within the entire curriculum (Purdue University, 2019). It is a free public charter school that encourages project-based learning, in which students team up to develop a solution for real-world challenges that are initiated by industry partners and experts. Their curriculum follows the five-step iterative design process that are similar to engineering design process: initiate (discovering the challenge), empathize (understand users' needs),

analyze (define and ideate), execute (prototype and implement solutions), and reflect. Some of their design challenge examples are from conservation efforts in partnership with diverse community partners: how to enhance fan and player experience on the rink and ticket booth in partnership with Indy Fuel Hockey or how to improve transportation network in Indianapolis in collaboration with IndyGo.

The inclusion of engineering experiences in STEM curricula is also underway through developing frameworks (Moore et al., 2014), articulating the engineering practices in education (Cunningham and Kelly, 2017), and documenting how engineering design approaches help to contextualize mathematics and science principles while enhancing motivation and STEM interest to solve complex, real-world problems (Barrett et al., 2014; Bowker, 2007). These curricula activities, for example, asked students to design, construct, test, and evaluate a trebuchet catapult, a 3D model of a paper airplane that would be airborne for as long as possible, and a wind tunnel by considering basic tornado mechanics and structural engineering principles. In addition, public schools have been taking initiatives to use school librarians or media specialists to work closely with science teachers to develop design-based STEM curricula (Philips et al., 2018).

Out-of-school settings

Emerging research of makerspaces and making programs in out-of-school settings demonstrates that making can instill learners' STEM interest and position learners to pursue future STEM activities and identity. For instance, the Exploratorium created an after-school tinkering program—in partnership with Lighthouse community charter school, the University of Washington, and the Boys and Girls club—to expand tinkering dispositions beyond one-time visit to an informal learning institution and develop interdisciplinary forms of STEM learning. Their curricula included designing and making scribble machines, stop-motion animation films, shadow plays, wooden pinball machines, and musical instruments, which focused on connecting the scientific concepts with personal interests and cultural histories to support meaning-making, while promoting practices of iterative problem-solving and play. Making a scribbling machine, for example, utilized markers, craft materials, and battery-powered motors to move and create scribbling lines, thus it required learners' engagement with various STEM concepts such as circuitry, vibration, kinetic motion, and symmetry. This tinkering activity supported learners to experiment with different materials, change different parameters within the design (i.e., place a material to a different position), and engage in engineering practices through the iterative cycles of trial-and-error. As such, the afterschool tinkering program supported learners, particularly those who might be positioned as less successful according to the school standards, to make connections to ideas and practices from home or school into the afterschool setting and vice versa (Vossoughi et al., 2013).

In rural communities in which permanent makerspaces are not a viable option, pop-up makerspaces or making programs are being introduced to share the STEM resources and expertise with underserved and marginalized populations. For example, mobile makerspaces (Moorefield-Lang, 2015), circulation making kits (Simpson and Maltese, 2019), and pop-up STEM programs (Kim and Zimmerman, 2019; Philips et al., 2018) underscored engineering design process in the making activities. Also, the Chicago Public library opened Mini MakerLabs, utilizing a van to take equipment such as 3D printers, laser cutters, milling machine, to provide resources for learners from marginalized populations to engage in engineering design.

Higher education settings

Similarly, higher education settings have been increasingly implementing makerspaces to support academic, extracurricular, and personally-driven projects for university students, faculty, and staff. Many of the makerspaces are either located within engineering-related departments—to address particular needs for engineering students—or university libraries—to help non-engineering students apply engineering design thinking and process. The rise of academic makerspaces has accelerated the culture of promoting engineering design. While approaches differ, a common theme in academic makerspaces is to provide a space to develop skills to design, manufacture, test, and improve solutions to real-world problems (Wilczynski, 2015). Recently, efforts have taken place to connect hands-on making activities into university curriculum across different disciplines to cultivate the culture of innovation and creativity. For instance, the University of Texas at Arlington—in partnership with University of Nevada-Reno—implemented cross-disciplinary maker-based competencies into multiple undergraduate curricula (i.e., art, biology, education, mathematics). The project created newly adapted assignments that incorporated making into the undergraduate curriculum, while still addressing the desired course learning outcomes (Wallace et al., 2017). As such, academic makerspaces contribute to expanding the design-centered approach of engineering design to university students, faculty, and staff who are less familiar with engineering, design, and making.

DBE with computational thinking and STEM+C

Computational thinking is another skill set that can benefit from DBE. With growing interests toward the role of information technology in society, computational thinking has become one of the key agendas in the realm of STEM education. Of varying components of design, computational thinking is a critical one in fostering problem-solving skills (Brooks and Sjöberg, 2020; Wu et al., 2019) that are required for the 21st century education. Despite various definitions of computational thinking, we use Rode et al.'s list of characteristics (2015, p. 240) that was adapted from Barr and Stephenson's (2011):

- Analyzing and organizing data logically
- Visualizing data through abstractions, modeling, and simulations

- Formulating problems in a way a computer can help in solving them
- Identifying, analyzing, and implementing solutions effectively and efficiently
- Using algorithmic thinking to automate solutions
- Generalizing and applying this problem-solving process to other kinds of problems

The examples of DBE with computational thinking are derived from STEM and/or STEM+C (STEM+Computer Science) education settings where learners engage with design process to address interdisciplinary, STEM-related problems by utilizing and developing computational thinking. Learners in these settings are supported to design games, robotics, and/or coding (programming) by using computers and other technological tools.

K-12 settings (both school and out-of-school)

DBE with computational thinking has been explored extensively in K-12 STEM education settings both in and outside of classrooms. Due to computational thinking's close link with programming and electrical circuitry, most of the DBE activities revolve around engaging learners to perform fun computational projects. Oftentimes, these projects involve block/visual programming, game design, and maker projects with small hardware. For instance, [Rode et al. \(2015\)](#) conducted participatory action research about how children in a computer club at a German primary school demonstrated skill sets related to computational thinking during their engagement with e-textiles and the LilyPad Arduino. Children participated in two activities: *Bunny Bright* and *Monster Making*. During the first activity, children used circuit board, LED lights, and textiles to become familiar with basic electrical concepts. The second activity integrated block programming component. The study illustrated that children were able to engage in computational skill sets and practice new core skills such as *aesthetics, creativity, constructing, visualizing multiple representations, and understanding materials* (p. 241), which are closely related to computational making and design thinking. Through their study, they asserted the importance of expanding the notion of computational thinking to computational making because computing has extended beyond desktop usage. Similarly, [Richard and Giri \(2019\)](#) observed an elective mini-course on making at a public magnet high school where students engaged in bidirectional design activities by using between digital and physical design kits. During the workshop, students built models through collaborative projects with digital and tangible maker toolkits, such as Makey Makey, LilyPad Arduino, ModKit, and Scratch. Through multimodal analysis of the students' activities, the researchers found evidence for different practices of computational thinking among learners, designing circuits, coding, and understanding complex systems. [Mel-ander Bowden \(2019\)](#) used Scratch to engage students in problem solving through game design activity. During the activity, students were prompted to redesign a Scratch game by modifying visual and auditory components. The author focused on exploring the social nature of problem solving throughout design process. [Kynigos and Grizioti \(2020\)](#) saw game modding as an effective tool for developing computational thinking skills.

DBE with computational thinking has also addressed the needs of learners in out-of-school settings. [Brooks and Sjöberg \(2020\)](#) explored elementary-aged children in creative workshops where children in groups designed games using creative materials and digital technology. They were asked to create a short stop-motion film for their game design prototype. Through the interaction analysis and quasi-content analysis approaches, the authors discovered that children developed specific kinds of computational skills such as step-by-step procedural skills, design and arrangement skills, and computational perspectives (p. 93). The authors concluded that reverse engineering approach that involve the designerly activity can help learners understand computational thinking skills. [Wagh et al. \(2017\)](#) reported on an ethnographic study involving a group of high school students working on a project in a paid summer program, the Youth Maker-Teacher Program. The group of students worked on a project called WaterWorks project, which was described as "a musical water fountain that changes color when different Jumbo piano keys are pressed" (p. 3). By analyzing video and audio recordings, field notes, photographs, artifacts, and group's blog posts, the researchers discovered that the group mainly practiced problem decomposition and simulation as well as troubleshooting and debugging. Both studies involved workshop activities where students would work on design activities as a group. They would participate in problem solving activities, which, in turn, could facilitate computational thinking practices.

Higher education settings

Game-related activities have been found to be useful in higher education settings as well. Interests toward computational thinking have been expanded to not only STEM majors but also diverse disciplines. For instance, [Marcelino et al. \(2018\)](#) used Scratch to train elementary school teachers in distance higher-education settings to engage them in computational thinking related activities. The trainees went through computational thinking-related themes with various activities and completed a final project during the course. During the course, the trainees learned about computational thinking concepts, explored the usage of Scratch in school settings, learned to program and build an application using Scratch. To teach introductory programming within the STEM discipline, [Kazimoglu et al. \(2012\)](#) used game prototype and helped students develop computational thinking skills that were expressed through practices, such as problem solving, building algorithms debugging, simulation, and socializing.

DBE with digital literacy and new media creation

Digital literacy (i.e., media literacy, multimedia literacy) is the ability to understand, manipulate, transform, and distribute digital media ([Burdick and Willis, 2011](#)). Digital literacy is critical in the 21st century education because such ability of utilizing digital

media and resources can impact learning performances and outcomes (Banerjee and Gibbs, 2016; Lemke et al., 2014). In other words, digital literacy is the core competence to understand, collect, analyze, and create information in any discipline of STEM. As a result, it is gaining more attention as a bridge to connect different educational practices from multiple STEM disciplines. Studies have also shown that digital literacy matters to learners' STEM-related interests and future careers (Bekker et al., 2015). Thus, understanding and promoting digital literacy could enhance the quality of integrative STEM education.

New media creation, which is tightly related to digital literacy, is another area of design. In addition to tangible and physical products, creating web-based or hybrid digital media also requires design thinking and process. DBE with new media can include a wide range of creative production, including digital storytelling using videos, websites, or blogs (e.g., Chan et al., 2017; Hill and Grinnell, 2015; Rahm and Moore, 2015) and creative visualization utilizing digital media and data (e.g., Stornaiuolo, 2020). New media creation needs skills for research, problem-solving, assessing, technology, interpretation, and reflection, which are all critical for STEM learning (Chan et al., 2017; Hill and Grinnell, 2015). In this section, we review the cases of DBE with (and for) digital literacy and/or new media creation in STEM education settings.

K-12 settings (both school and out-of-school)

DBE with multimedia design has been integrated into diverse curricula and educational activities in K-12 settings. Liu (2003) designed a three-phase curriculum called Learner-As-Multimedia-Designer, through which students in elementary, middle, and high schools designed multimedia products. In the first phase, students learned about required tools for media creation (e.g., software and hardware), in the second phase, students designed products for clients (e.g., a multimedia presentation for teachers' job fair) through planning, designing, producing, and evaluating and revising, and lastly in the third phase, they repeated the cycle of the second phase for another project. Groups of students collaborated on attaining graphics, creating digital video and audio clips and animations, and merging the digital resources for their final media product. This multimedia design activity was shown to enhance the students' cognitive and collaborative skills.

Diverse DBE programs with new media were implemented in out-of-school settings as well. In a weekly afterschool club from Rahm's study (2016), nine high school students participated in video production activities over 12 weeks. The DBE activities included brainstorming, fieldtrips to community science centers (e.g., botanical garden), topic development for collaborative video production, as well as storyboard making, filming, and editing videos. Students took the photos and videos by themselves from the fieldtrip and utilized them as sources of their video design. Another notable example is YOUMedia. YOUMedia supported interest-driven and collaborative learning through digital media creation across diverse environments, including libraries, museums, and afterschool clubs (Ito et al., 2013). YOUMedia's programs utilized digital media creation to hang out, mess around, and geek out on areas related to STEM (e.g., electronic gaming, design) and the arts (e.g., music, writing). Teenagers in this program chose their own interest, by communicating with their mentors and peers in person as well as through social media (e.g., iRemix), and participated in a range of design projects with new media. They were encouraged to design and create new media production about their own interests, such as podcasts, blog posts, and journal articles, while their content knowledge and digital literacy were also empowered.

To enhance digital literacy, some incorporated both digital and physical technology into learners' design activities. In Stornaiuolo's study (2020), students in a media makerspace at a public high school learned how to interpret and analyze web-based data using Google Trends and DataBasic web tools. This DBE curriculum focused on digital literacy related to data (i.e., data literacy). Students in this curriculum collected personal data and designed their own visualization of their data story, with which they created a key and printed onto T-shirts. Through designing their own visualization with data and data stories, students learned how to understand and use data in creative ways. They also developed agency in terms of utilizing digital resources and data. Bekker et al. (2015) also focused on enhancing digital literacy of elementary and secondary school students by utilizing digital and tangible toolkits for their design-based learning efforts. Students were asked to design innovative products, such as new interactive and sustainable lamp. For their design projects, students followed the steps of imagine/discover, make/prototype, test/present, and reflect, which are essential components of DBE. Throughout their design processes, students were able to engage with STEM-related contents as well as enhance digital literacy.

Higher education settings

DBE with new media creation, especially in a format of digital storytelling, has been widely used in higher education contexts. Hill and Grinnell (2015) adopted digital storytelling with infographics for STEM major students. Students were asked to create "story resumes"—a form of digital resume through storytelling—including infographics as well as videos. They created movies using their smartphones, PowerPoint, as well as sound and video functions. By designing their story resumes, these students were able to reflect on their competencies and attained digital literacy skills to create the infographics and videos. In a course about technology at a community college in Hong Kong, Chan et al. (2017) employed digital storytelling to enhance students' deep learning. Students designed their own storytelling videos through brainstorming, planning and storyboard sketching, and filming and producing videos about the themes of their choice. While this process improved the students' skills and competence related to digital literacy as they had to learn about multimedia concepts, they also developed skill sets for design by using software for video editing (e.g., Adobe Premiere Pro) and hardware devices (e.g., video cameras and microphones) to create their storytelling videos. This process also helped them deeply engage with the concept knowledge of their choice. The d.School from the Stanford Hasso Plattner

Institute of Design is another example that employed design thinking in higher education, especially across diverse disciplines within and beyond STEM. Professors from different departments taught courses together, and students could enhance design thinking through interdisciplinary opportunities and projects. To pursue innovation, d.school emphasized enhancing students' ability to use and leverage technology and knowledge (Banerjee and Gibbs, 2016). In other words, digital literacy, as well as design process, acted as the core ability and the bridge across various disciplines.

Interdisciplinarity of the current applications

Although we divide the study cases into STEM (engineering design and maker education, computational thinking and STEM+C education, and digital literacy and digital media creation) categories, some of the examples fall into all three categories. For instance, Bekker et al. (2015) DBE program focused on enhancing digital literacy but through designing activities with both digital and physical toolkits, and these are also tightly related with maker education and computational thinking. Also, d.school represents a large-scale educational effort that supported diverse types of innovative thinking and behaviors rooted from interdisciplinary knowledge. As such, DBE has potential to support digital literacy engineering thinking, maker activities, computational thinking, as well as conceptual learning in STEM, on the basis of design thinking and process.

Designing design-based education for STEM learning: principles and gaps

How to design learner experience in design-based education (DBE) in STEM is a critical question for researchers and practitioners to better serve learners, especially because DBE entails complex problem-solving and higher-order cognitive activities. Puntambekar and Kolodner (2005) employed distributed scaffolding to support middle school students' design-based learning activities that are related to a scientific phenomenon, erosion. The authors implemented principles of scaffolding for DBE: (a) creating common goals, (b) providing ongoing diagnosis, (c) providing dynamic and adaptive support, (d) including dialogs and interactions, and (e) positing fading and transfer of responsibility. In particular, they emphasized "understanding of the processes involved in learning science by design and what [educators] knew about student difficulties" (p. 190). In other words, rather than product-oriented assessments and facilitation, process-oriented supports are essential for better learning experiences in DBE.

However, training teachers, as well as informal educators and parents, to become facilitators of DBE is still a remaining task for the future of DBE. The cases we reviewed barely demonstrated how teachers were trained to design and implement DBE. Moreover, some studies (e.g., Chan and Blikstein, 2018; Lee et al., 2017) even discussed the necessity of proper preparation for this new teaching method. DBE requires immense efforts of educators to accommodate differentiated needs depending on different designs. Without proper training, educators and parents in DBE may impose learners on pre-determined answers despite the fact that it is important to support learners' own tinkering and making throughout the design process (Jung et al., 2020). Therefore, we point out the necessity to come up with systemic and concrete plans for supporting educators to empower learners in DBE and better design DBE activities for STEM learning.

We also identify the gap between the interdisciplinarity that many DBE programs pursued and the integrative STEM education. Despite the meaningful insights of DBE from current examples, we barely found the examples of DBE that explicitly connect with mathematics education among STEM areas. In fact, most of them were related to technology and engineering. Although DBE ideally pursues various, individual disciplines in STEM as well as integrative STEM education (Donna, 2012; Li et al., 2019), current DBE practices only allowed limited integration of STEM disciplines. For integrative STEM education, establishing intentional connections among concepts and practices of S, T, E, and M disciplines is essential (Sanders, 2012). Therefore, efforts to make precise, clear, and explicit connections between each of the STEM areas is another remaining task when planning and implementing DBE in STEM.

Conclusion

In this article, we conceptualize design-based education (DBE) in STEM educational settings by overviewing related theories, including socio-constructive and sociocultural perspectives, constructionism, project-based and problem-based learning, and collaborative learning. We also introduce the rationale and benefits of DBE, especially for the 21st century education. We also review the current examples of DBE practices, which provide insights for what types of DBE exist and how researchers can apply DBE to certain STEM education settings. While our review of the DBE practices categorized three innovative STEM education movements (engineering design and maker education, computational thinking and STEM+C education, and digital literacy and digital media creation), we also highlight the potential of DBE in connecting all of these categories. If the lack of explicit efforts for integrative STEM education and the systemic supports for training educators can be addressed, the future DBE will better support children and youths to grow as critical and creative leaders for the 21st century.

References

- Association of Science-Technology Centers, Urban Libraries Council, 2014. Learning Labs in Libraries and Museums: Transformative Spaces for Teens. Association of Science-Technology Centers and Urban Libraries Council, Washington, DC.
- Banerjee, B., Gibbs, T., 2016. Teaching the innovation methodology at the Stanford d.school. In: Banerjee, B., Geri, S. (Eds.), *Creating Innovation Leaders: A Global Perspective*. Springer, pp. 163–174.
- Barak, M., Assal, M., 2018. Robotics and STEM learning: students' achievements in assignments according to the P3 task taxonomy—practice, problem solving, and projects. *Int. J. Technol. Des. Educ.* 28, 121–144. <https://doi.org/10.1007/s10798-016-9385-9>.
- Barr, V., Stephenson, C., 2011. Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *ACM Inroads* 2, 48–54.
- Barrett, B.S., Moran, A.L., Woods, J.E., 2014. Meteorology meets engineering: an interdisciplinary STEM module for middle and early secondary school students. *Int. J. STEM Educ.* 1, 1–7. <https://doi.org/10.1186/2196-7822-1-6>.
- Bekker, T., Bakker, S., Douma, I., van der Poel, J., Scheltenaar, K., 2015. Teaching children digital literacy through design-based learning with digital toolkits in schools. *Int. J. Child Comput. Interact.* 5, 29–38. <https://doi.org/10.1016/j.ijcci.2015.12.001>.
- Blikstein, P., Krannich, D., 2013. The Makers' Movement and FabLabs in Education, p. 613. <https://doi.org/10.1145/2485760.2485884>.
- Borge, M., Toprani, D., Yan, S., Xia, Y., 2020. Embedded design: engaging students as active participants in the learning of human-centered design practices. *Comput. Sci. Educ.* 30, 47–71. <https://doi.org/10.1080/08993408.2019.1688592>.
- Bowker, R., 2007. Children's perceptions and learning about tropical rainforests: an analysis of their drawings. *Environ. Educ. Res.* 13, 75–96. <https://doi.org/10.1080/13504620601122731>.
- Brooks, E., Sjöberg, J., 2020. A Designerly Approach as a Foundation for School Children's Computational Thinking Skills While Developing Digital Games. *Proceedings of the Interaction Design and Children Conference IDC 2020*, pp. 87–95. <https://doi.org/10.1145/3392063.3394402>.
- Burdick, A., Willis, H., 2011. Digital learning, digital scholarship and design thinking. *Des. Stud.* 32, 546–556. <https://doi.org/10.1016/j.destud.2011.07.005>.
- Chan, M.M., Blikstein, P., 2018. Exploring problem-based learning for middle school design and engineering education in digital fabrication laboratories. *Interdiscip. J. Probl. Learn.* 12, 9–10. <https://doi.org/10.7771/1541-5015.1746>.
- Chan, B.S.K., Churchill, D., Chiu, T.K.F., 2017. Digital literacy learning in higher education through digital storytelling approach. *J. Int. Educ. Res.* 13, 1–16. <https://doi.org/10.19030/ijer.v13i1.9907>.
- Cunningham, C.M., Kelly, G.Y.J., 2017. Epistemic practices of engineering for education. *Sci. Educ.* 101, 486–505. <https://doi.org/10.1002/sce.21271>.
- Donna, J., 2012. A model for professional development to promote engineering design as an integrative pedagogy within STEM education. *J. Pre-Coll. Eng. Educ. Res.* 2, 1–8. <https://doi.org/10.5703/1288284314866>.
- Dougherty, D., 2012. The maker movement. *Innovations* 7, 11–14.
- Dym, C., Agogino, A., Eris, O., Frey, D.D., Leifer, L.J., 2005. Engineering design thinking, teaching, and learning. *J. Eng. Educ.* 94, 103–120. <https://doi.org/10.1002/j.2168-9830.2005.tb00832.x>.
- Gero, J.S., 1999. Constructive memory in design thinking. *Archit. Sci. Rev.* 42, 29–35.
- Hill, S., Grinnell, C., 2015. Using Digital Storytelling With Infographics in Stem Professional Writing Pedagogy, 2014 IEEE International Professional Communication Conference (IPCC). <https://doi.org/10.1109/IPCC.2014.7020367>.
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salor, K., Schor, J., Sefton-Green, J., Watkins, S.C., 2013. *Connected Learning: An Agenda for Research and Design*. Connected Learning, Irvine, CA.
- Jung, Y.J., Toprani, D., Yan, S., Borge, M., 2017. Children's participation in rulemaking to mitigate process problems in CSCL. In: Smith, B.K., Borge, M., Mercie, E., Lim, K.Y. (Eds.), *Making a Difference*. International Society of the Learning Sciences, Philadelphia, PA, pp. 652–655.
- Jung, Y.J., Whalen, D.P., Zimmerman, H.T., 2020. Epistemic agency shifts between children and parents during inventing with robotics at museum-based makerspace. In: *ICLS 2020*, pp. 851–852.
- Kafai, Y.B., Searle, K., Martinez, C., Brayboy, B., 2014. Ethnocomputing With Electronic Textiles: Culturally Responsive Open Design to Broaden Participation in Computing in American Indian Youth and Communities. *SIGCSE 2014—Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pp. 241–246. <https://doi.org/10.1145/2538862.2538903>.
- Kafai, Y.B., 2006. Constructionism. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, pp. 35–46.
- Kazimoglu, C., Kiernan, M., Bacon, L., MacKinnon, L., 2012. Learning programming at the computational thinking level via digital game-play. *Proc. Comput. Sci.* 9, 522–531. <https://doi.org/10.1016/j.procs.2012.04.056>.
- Keune, A., Gomoll, A., Peppeler, K., 2015. Flexibility to learn. *FabLearn* 15.
- Kim, S.H., Zimmerman, H.T., 2019. Understanding the practices and the products of creativity: making and tinkering family program at informal learning environments. In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. ACM, pp. 246–252. <https://doi.org/10.1145/3311927.3323117>.
- Koh, J.H.L., Chai, C.S., Wong, B., Hong, H.Y., 2015. *Design Thinking for Education: Conceptions and Applications in Teaching and Learning*. Springer Science+Business Media Singapore, Singapore. <https://doi.org/10.1007/978-981-287-444-3>.
- Kolodner, J.L., 2004. The learning sciences: past, present, future. *Educ. Technol.* 44, 34–40.
- Krajcik, J.S., Shin, N., 2014. Project-based learning. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, pp. 275–297.
- Kynigos, C., Griziotti, M., 2020. Modifying games with ChoiCo: integrated affordances and engineered bugs for computational thinking. *Br. J. Educ. Technol.* 0, 1–16. <https://doi.org/10.1111/bjet.12898>.
- Lachapelle, C., Cunningham, C.M., 2019. Engineering in elementary schools. In: *Engineering in Pre-college Settings*, pp. 61–88. <https://doi.org/10.2307/j.ctt6wq7bh.8>.
- Lee, V.R., Lewis, W., Searle, K.A., Recker, M., Hansen, J., Phillips, A.L., 2017. Supporting Interactive Youth Maker Programs in Public and School Libraries: Design Hypotheses and First Implementations. *IDC 2017—Proceedings of the 2017 Conference on Interaction Design and Children*, pp. 310–315. <https://doi.org/10.1145/3078072.3079741>.
- Lemke, J., Lecusay, R., Cole, M., Michalchik, V., 2014. Documenting and Assessing Learning in Informal and Media-Rich Environments. The MIT press, Cambridge, MA. <https://doi.org/10.1007/s13398-014-0173-7.2>.
- Li, Y., Schoenfeld, A.H., DiSessa, A.A., Graesser, A.C., Benson, L.C., English, L.D., Duschl, R.A., 2019. Design and design thinking in STEM education. *J. STEM Educ. Res.* 2, 93–104. <https://doi.org/10.1007/s41979-019-00020-z>.
- Liu, M., 2003. Enhancing learners' cognitive skills through multimedia design. *Interact. Learn. Environ.* 11, 23–39. <https://doi.org/10.1076/1lee.11.1.23.13686>.
- Lu, J., Bridges, S., Hmelo-Silver, C.E., 2014. Problem-based learning. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, pp. 298–318.
- Marcelino, M.J., Pessoa, T., Vieira, C., Salvador, T., Mendes, A.J., 2018. Learning computational thinking and scratch at distance. *Comput. Hum. Behav.* 80, 470–477. <https://doi.org/10.1016/j.chb.2017.09.025>.
- Melander Bowden, H., 2019. Problem-solving in collaborative game design practices: epistemic stance, affect, and engagement. *Learn. Media Technol.* 44, 124–143. <https://doi.org/10.1080/17439884.2018.1563106>.

- Moore, J.T., Glancy, A.W., Kristina, M.T., Jennifer, A.K., Karl, A.S., 2014. A framework for quality K-12 engineering education: research and development. *J. Pre-Coll. Eng. Educ. Res.* 4, 2.
- Moorefield-Lang, H.M., 2015. When makerspaces go mobile: case studies of transportable maker locations. *Libr. Hi Technol.* 33, 462–471. <https://doi.org/10.1108/LHT-06-2015-0061>.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press, Washington, D.C.
- Papert, S., Harel, I., 1991. Situating constructionism. In: *Constructionism*. Ablex Publishing Corporation, pp. 1–11.
- Philips, A.L., Lee, V., Recker, M., 2018. Small-town libraries as experience engineers. In: *Reconceptualizing Libraries: Perspectives from the Information and Learning Sciences*. Routledge, New York, NY.
- Puntambekar, S., Kolodner, J.L., 2005. Toward implementing distributed scaffolding: helping students learn science from design. *J. Res. Sci. Teach.* 42, 185–217. <https://doi.org/10.1002/tea.20048>.
- Purao, S., Baldwin, C.Y., Hevner, A., Storey, V.C., Pries-Heje, J., Smith, B., Zhu, Y., 2008. The sciences of design: observations on an emerging field. *Commun. Assoc. Inf. Syst.* 23, 523–546. <https://doi.org/10.2139/ssrn.1281643>.
- Purdue University, 2019. Purdue Polytechnic High School.
- Rahm, J., Moore, J., 2015. A case study of long-term engagement and identity-in-practice: insights into the STEM pathways of four underrepresented youths. *J. Res. Sci. Teach.* <https://doi.org/10.1002/tea.21268>.
- Razzouk, R., Shute, V., 2012. What is design thinking and why is it important? *Rev. Educ. Res.* 82, 330–348. <https://doi.org/10.3102/0034654312457429>.
- Richard, G.T., Giri, S., 2019. Digital and physical fabrication as multimodal learning. *ACM Trans. Comput. Educ.* 19, 1–35. <https://doi.org/10.1145/3243138>.
- Rode, J.A., Marshall, A., Weibert, A., Aal, K., Von Rekowski, T., El Mimouni, H., Sharma, A., Jobs, J., Schleeter, A., Booker, J., 2015. From Computational Thinking to Computational Making. *UbiComp ISWC 2015—Proceedings of 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, pp. 401–402. <https://doi.org/10.1145/2800835.2800926>.
- Rogoff, B., 1994. Developing understanding of the idea of communities of learners. *Mind Cult. Activ.* 1, 209–229. <https://doi.org/10.1080/10749039409524673>.
- Sanders, M.E., 2012. Integrative STEM education as “best practice”. In: Middleton, H. (Ed.), *7th Biennial International Technology Education Research Conference*. Griffith Institute of Educational Research, Queensland, Australia, pp. 103–117.
- Sheridan, K.M.K., Halverson, E.R., Litts, B.K., Brahms, L., Jacobs-Priebe, L., Owens, T., 2014. Learning in the making: a comparative case study of three makerspaces. *Harv. Educ. Rev.* 84, 505–565. <https://doi.org/10.17763/haer.84.4.brr34733723j648u>.
- Simpson, A., Maltese, A., 2019. Developing and using low-cost MAKEngineering kits for youth engagement. *Teach. Libr.* 46, 14–17.
- Stornaiuolo, A., 2020. Authoring data stories in a media makerspace: adolescents developing critical data literacies. *J. Learn. Sci.* 29, 81–103. <https://doi.org/10.1080/10508406.2019.1689365>.
- Thomas, D., Brown, J.S., 2011. *A New Culture of Learning: Cultivating the Imagination for a World of Constant Change*. CreateSpace, Lexington, KY.
- Vossoughi, S., Bevan, B., 2014. Making and tinkering: a review of the literature. *Natl. Res. Counc. Comm. Sch. Time STEM* 1–55.
- Vossoughi, S., Escudé, M., Kong, F., 2013. *Tinkering, Learning & Equity in the After-School Setting*.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes, Memory*. Harvard University, Cambridge, MA.
- Wagh, A., Gravel, B., Tucker-Raymond, E., 2017. The Role of Computational Thinking Practices in Making: How Beginning Youth Makers Encounter & Appropriate CT Practices in Making. *Proceedings of FabLearn 2017 Conference, Palo Alto, California USA, October 2017 (FABLEARN 2017)*. <https://doi.org/10.1145/3141798.3141808>.
- Wallace, M.K., Trkay, G., Musick Peery, K., Chivers, M., Radniecki, T., 2017. Maker competencies and the undergraduate curriculum. In: *Proceedings of the 2017 International Symposium on Academic Makerspaces*. Cleveland, OH.
- Whyte, J., 2016. Makerspaces in libraries. *iJ. Grad. Student J. Fac. Inf.* 2. <https://doi.org/10.4324/9781315143422-3>.
- Wilczynski, V., 2015. Academic makerspaces and engineering design. In: *ASEE Annual Conference and Exposition*. American Society for Engineering Education, Seattle, WA.
- Won, S.G.L., Evans, M.A., Carey, C., Schnittka, C.G., 2015. Youth appropriation of social media for collaborative and facilitated design-based learning. *Comput. Hum. Behav.* 50, 385–391. <https://doi.org/10.1016/j.chb.2015.04.017>.
- Wu, B., Hu, Y., Ruis, A.R., Wang, M., 2019. Analyzing computational thinking in collaborative programming: a quantitative ethnography approach. *J. Comput. Assist. Learn.* 35, 421–434. <https://doi.org/10.1111/jcal.12348>.

Technology-enhanced collaborative learning in STEM

Changzhao Wang and Ji Shen, Department of Teaching and Learning, University of Miami, Coral Gables, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	207
Parameters and forces of collaborative learning	207
Benefits of collaborative learning in STEM education	208
TECL in STEM education	208
Connecting learners across space	208
Remote collaborative learning environments	208
Multi-user virtual environments (MUEs)	209
Facilitating group written discourse	209
Synchronous discussion forum	209
Asynchronous discussion tool/forum	210
Guiding collaborative process	210
Generate, deposit, and visualize collective knowledge	211
Assess and analyze TECL in STEM	211
Conclusion	212
References	212

Introduction

Collaborative learning is an ancient wisdom. In one of the Confucian classics *Li Ji* (礼记), there is a famous saying—“one becomes narrow-minded if learning without peers.” In contrast to individual learning, collaborative learning provides learners with the opportunity to interact with others, and hence, increase their opportunities to learn.

Many complex problems in the science, technology, engineering, and mathematics (STEM) fields such as climate change and infectious disease control require collaborative problem solving. For example, to fight against the new corona virus pandemic, policy makers need to work with different sectors—including scientists, health care workers, and the media—to make and implement effective policies and strategies. It is necessary to prepare younger generations with a collaborative mindset to be able to tackle real life problems and challenges.

There is also an increasing demand of new forms of collaboration due to the widespread application of modern digital tools and social-network platforms. Consider some of the common scenarios today in- and out-of-school environments: a small team works in an online chat room to solve a math problem together; a class is using a discussion board asynchronously to brainstorm ideas on a science topic; a community of game players discuss mechanisms to solve embedded technical bugs. To reflect the trend of integrating technologies in education, this article focuses on Technology-Enhanced Collaborative Learning (TECL) in STEM.

Parameters and forces of collaborative learning

In this article, we adopt the broad definition of collaborative learning by [Dillenbourg \(1999\)](#): “a situation in which two or more people learn or attempt to learn something together” (p. 1). In theory, collaborative learning is built upon social constructivism ([Adams, 2006](#)), which submits that learning happens during the process of social interaction ([Vygotsky, 1978](#)). In practice, the key to defining a specific form of collaborative learning is through specifying its various pragmatic parameters: The scale of collaborative learning can be as small as a pair on one end, and as big as a large community on the other; the timing can vary, such as in a year-long course, or in a one-shot activity (e.g., a team-building activity at a firm); the setting can change from physical collaborative learning to virtual ones; the medium may include various technological platforms, and so forth.

Besides these pragmatic parameters, the dynamics of collaborative learning are shaped by different forces at play. One such force is the preference to individualism or collectivism in terms of the learning objective. If the learning objective is set for the individual, collaborative learning will foreground the improvement of individual knowledge or skills. For instance, to assess students’ collaborative problem-solving competency in the PISA test, [OECD \(2017, p. 6\)](#) defined it as “the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution.”

Alternatively, when the learning objective is for the collective, community building and collective knowledge construction are prioritized. For example, the notion “team-based learning” focuses on small-group interaction in a class setting and requires that assignments promote team development in addition to individual learning ([Michaelsen, 1992](#); [Michaelsen and Sweet, 2008](#)). In a similar vein, those who advocate “communities of practice” ([Wenger, 2011](#)) or Knowledge Building ([Scardamalia and Bereiter,](#)

2006, 2014) emphasize the importance of building a community for learners. Communities of practice are “groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly” (Wenger, 2011, p. 1). The Knowledge Building theory stresses more on the creation and innovation of collective knowledge, and prioritizes community knowledge advancement instead of individual accomplishment (Scardamalia and Bereiter, 2006).

Another force that influences collaborative learning is division of labor. In many learning occasions labor division is set clearly by instructors or supervisors (i.e., the work is divided into independent sub-tasks and then the group assembles different parts together; Nugraha et al., 2019). “Cooperative learning” (Johnson and Johnson, 1994) is often used in these cases. In other collaborative learning occasions, work is not split, but negotiated and completed jointly by the individuals involved (Micari and Pazos, 2020).

Often conceived as resistance to collaboration, competition refers to the action that two or more parties battle for a common benefit or recognition that cannot be shared. In the collaborative learning environment, all group members contribute to the same learning task. In contrast, in a competitive environment, learners work independently and try to outperform others. However, collaboration and competition are not necessarily incompatible to each other. Contradiction may serve as a driving force in advancing an activity system (Roth et al., 2004). Likewise, competition may serve as a positive ingredient to collaborative learning (Mayer, 2009; Ziegler and Hamker, 2011). It is often coexisting with collaboration within or across group learning settings (Shen et al., 2020a). Therefore, a competitive mindset could be a double-edged sword: When carefully scaffolded and well-harnessed, it can facilitate the collaborative problem-solving process; if not, competition may undermine collaboration.

Benefits of collaborative learning in STEM education

Collaborative learning has been well embraced in STEM education (see a systematic literature review by Thibaut et al., 2018). Building student learning communities is recommended as a strategy to increase students’ interest and persistence in college STEM programs (Sithole et al., 2017). Several meta-analysis studies have demonstrated that collaborative learning promotes student learning in general (Chen et al., 2018; Kyndt et al., 2013; Roseth et al., 2008). Moreover, according to a more recent meta-analysis, computer-supported collaborative learning has shown notable effectiveness in STEM education, including the process, knowledge, and affective outcomes (Jeong et al., 2019).

More importantly, research has shown strong evidence that collaborative learning is contributing to addressing critical issues in STEM education (Laal and Ghodsi, 2012; Panitz, 1999). One such issue is low enrollment and high attrition (Chen and Soldner, 2013; HESA, 2021; Lichtenberger and George-Jackson, 2013; Romash, 2019; Sithole et al., 2017). Collaborative learning may enhance students’ self-esteem and sense of achievement by, e.g., allowing them to help others and having them involved in the design of class activities (Johnson and Johnson, 1990). Students can develop more higher-order thinking skills and learn better on STEM content as a result of collaborative learning than individual learning (Gokhale, 1995; Johnson and Johnson, 2009; Webb, 1982). Collaborative learning may also foster students’ communication skills and abilities to resolve conflict (Johnson and Johnson, 1990; Tolmie et al., 2010).

Another issue is equity in STEM education. For example, female students generally have a lower enrollment rate yet higher attrition rate in postsecondary STEM majors than male students. Likewise, students from low-income families exhibit similar patterns when compared with students from high-income families (Chen and Soldner, 2013; English, 2017; HESA, 2021; Lichtenberger and George-Jackson, 2013). Research has shown that students’ satisfaction toward the positive experience of collaborative learning makes them more intrinsically interested in learning and therefore, they become more willing to attend school and to persist in STEM learning. Additionally, collaborative learning promotes communication and interdependence among all group members, regardless of race, gender, and level of achievement, leading to a more equitable learning environment (Esmonde, 2009; Panitz, 1999).

TECL in STEM education

Technology plays an important role in facilitating collaborative learning. It may provide visual representations of learning tasks, guide collaboration processes, and serve as scaffolding to building collective knowledge (Chen et al., 2018; Goodyear et al., 2014). In the following, we focus on examples of technologies that directly support collaboration during the learning process. Some technologies may be used for collaborative learning, but do not play a central role in collaboration. For example, Lee et al. (2016) documented a study in which students collected their own activity data (e.g., step counts with time) by wearing sport watches and learn statistics by exploring the data in groups. The sport watch was used for data collection, rather than to facilitate collaboration. This type of studies is beyond the scope of this article.

Connecting learners across space

Remote collaborative learning environments

Audio and video conferencing tools (e.g., Zoom, Skype) are widely used nowadays in synchronous, distance learning settings. These tools provide a mechanism to connect learners across space. Many of them have specific features that support collaborative learning such as screen sharing, chat room, and annotation. Babaian and Schiano (2020) listed several practical tips for facilitating small group work in online learning. They encouraged both instructors and students be familiar with the features that support small group

work, including managing display names and changing screen name to group ID/name, navigating between breakout rooms and the main room, sending messages to all groups or a specific group in the breakout room, and being familiar with sharing or annotation tools. Research has also been conducted to examine these features. For instance, Singhal (2020) showed how breakout rooms in Zoom could be utilized in a virtual pharmacotherapeutics course during COVID-19. In the class, every five to six students were grouped and assigned to a breakout room, and the instructor was able to move around the rooms to facilitate discussions. The author argued that using breakout room could aid students' small team assignment in promoting active learning and therefore, leading to more participation.

Multi-user virtual environments (MUEs)

Multi-user virtual environments (MUEs) have been used for facilitating collaborative learning. In these environments, learners usually take on avatars to interact with one another in a virtual world, similar to popular (massive) multiplayer online role-playing games (Barab et al., 2005). It has been found that these learning environments may bring learning benefits to students with autism spectrum conditions (e.g., Cheng and Ye, 2010).

A pioneering example in science education is River City, an MUE learning environment designed to engage middle grade students in standards-based scientific inquiry practices (<https://muve.gse.harvard.edu/>). The curriculum improved over multiple versions of development (Ketelhut et al., 2010). For instance, the very first unit focused on biology and ecology. Joining via avatars, students could investigate authentic problems (e.g., diseases in the town) in a virtual town with a river running through it. To solve the problems, students must collaborate to share information and communicate with team members on what data they have gathered via a chat-in-text tool. In retrospection, Ketelhut et al. (2010) concluded that "MUEs are viable tools allowing learner-centered, collaborative inquiry learning that can provide an alternative to more conventional kinds of inquiry-based science instruction" (p. 108).

Ibanez et al. (2013) proposed a 3D virtual collaborative learning model that guides learners' collaborative activities in MUEs. The model specifies synchronization learning points in order to achieve the desired learning objectives of individuals or groups, as well as division of labor to promote positive interdependence. Based on the model, they conducted a case study in which participants (via avatars) walking up and down a street in the 3D environment and collecting information related to available theater shows, with the goal of purchasing ticket for the show they would like to attend. The participants needed to carry out several collaborative activities including sharing information and group assessments. The researchers were able to identify factors that influenced interactive collaboration in the first iteration and improved the design that resulted better collaboration in the second one.

Different from the examples above that have game-like settings, Hwang and Hu (2013) described a unit on geometry for fifth graders in which students worked in a virtual classroom setting. In this environment, students could manipulate virtual 3D geometric objects, observe peers' problem-solving processes, and send critiques and comments to each other. Their study showed that the experiment group performed significantly better than the control group on learning geometry. According to their study, manipulating the virtual objects and collaborating with peers (e.g., observing and criticizing peers' solutions) were key processes for students to learn abstract geometry concepts (Hwang and Hu, 2013).

Note that manipulations of virtual objects could also facilitate face-to-face peer interactions. Jackson et al. (2013) developed a tabletop-based math game activity to supplement and reinforce what students had learned in the fourth-grade math class. The math problem was displayed in the center of the tabletop screen (which allows multiple users to manipulate on the screen simultaneously) and required students to fill in blanks with elements from their resource pools. Four students worked in a group and each of them controlled a resource pool that may contain necessary elements in the solution, so they must collaborate to complete the problem. Students demonstrated favor to the collaborative game learning activity and boys significantly improved their math grades after the program.

In sum, both audio/video conferencing tools and MUEs are mechanisms that connect learners across space. While audio/video conferencing tools have been used in our daily life, more empirical studies on specific features of these tools that can best facilitate collaborative learning are still needed. Unlike audio/video conferencing tools, MUEs not only provide a common, alternative space for learners, but also simulate task scenarios. Through exploring virtual environments and manipulating virtual objects, learners could experience the problem-solving process in a more authentic way.

Facilitating group written discourse

While audio or video conferencing tools makes remote communication convenient and efficient, they require high speed Internet and extra devices (i.e., camera and microphone). In comparison, written discourse can satisfy the need of collaborative learning without these constraints.

Synchronous discussion forum

A synchronous discussion forum provides the opportunity of real time interaction with written texts or symbols (Kuyath, 2008; Rosé and Ferschke, 2016). In addition to enabling communication, some advanced synchronous discussion forums for educational purposes even include virtual intelligent agents to provide in-time feedback on the contents of learners' discussion. Wang et al. (2011) conducted a study in which high school students worked in pairs on geology-related brainstorming tasks. They used VIBRANT, a synchronous chat software with an embedded virtual agent that was able to provide real-time feedback. Tegos et al. (2016) showed a similar synchronous conversational platform MentorChat with a virtual agent that facilitated dyad collaborative

learning of 96 university students on the topic of human-computer interaction. In both learning environments, the virtual agent appeared as the third person to participate in the team dialog, commenting and asking guiding questions based on their discourse texts. Both studies indicated positive effect of using the virtual agent in enhancing students' collaborative learning.

Asynchronous discussion tool/forum

Whereas synchronous discussion tools have the advantage of satisfying learner's desire of having instant interactions and feedback, asynchronous discussion tools allow learners adequate time to review peers' posts and to provide more contributive comments. In one study that investigated collaborative learning in an undergraduate engineering course, the instructor asked group members to discuss the team task through email and to turn in a group report based on the email communication (Liu et al., 2018).

A shared web-based space (e.g., blogs) can serve as an asynchronous discussion forum for viewpoint exchange. For example, Jimoyiannis and Angelaina (2012) engaged 21 middle school students in the investigation of the acid rain problem using educational blogs. All the students who participated in the blog formed a learning community. They posted their individual findings and commented on each other's blogs. The study found that students presented increased interest in the activity, demonstrated mutual social and emotional support, and enhanced knowledge construction through their interaction on the blog.

In a study that investigated learners' collaboration mechanism, a total of 25 pre-service teachers interacted through an asynchronous online whiteboard (Gressick and Derry, 2010). They worked in small groups of five to design an interdisciplinary instructional unit that integrated math and science. Any member in the group could post his/her design proposal and others could comment, suggest edits, and rate. They found although there was no initial labor division by the instructor, participants took on different responsibilities and leadership roles (e.g., seeking input, knowledge contribution, topic control) were highly distributed among team members.

Synchronous or asynchronous discussion forums have been around for quite some time, but they keep evolving with the development of new technology. In recent years, multimodal inputs such as audio or video have been incorporated in these online discourse platforms (e.g., Flipgrid; Guffy et al., 2021), and machine learning algorithms have been embedded to support the automated and customized feedback to learners (Rosé and Fersckhe, 2016).

Guiding collaborative process

The aforementioned technologies create more collaborative opportunities by connecting learners across time and space, yet collaboration among students is not always successful as a result of ineffective communication, lack of interaction, improper labor division, to name a few (Le et al., 2018). Researchers have explored the affordances of collaborative scripts in computer-based learning environments to guide collaborative processes. For example, Rummel et al. (2012) conducted a study of scripted collaborative math learning with 106 high school students. They added collaborative scripts to the Cognitive Tutor Algebra software, a computer-based tutoring system. Two students first worked independently on different problems, then joined on one computer to collaborate on the problem provided by the script, which combined the two individual problems they had previously. During their collaboration, more scripts popped up to prompt certain collaborative behaviors at specific time points (e.g., making contributions based on the independent problem-solving experience, listening to peers, and asking questions if not understanding anything). The results showed that the students in the scripted group presented higher quality of collaboration and problem solving than those in the unscripted group.

Apart from guiding collaborative behaviors, collaborative scripts have also been used for peer monitoring. Bouyias and Deme-triadis (2012) compared the effects of compulsory peer-monitoring prompts and the fading-out scaffolding scripts for collaborative argumentation activities in an undergraduate computer science course. Students used iArgue, a web-based argumentation learning environment. Thirty-four students were randomly assigned into groups of two. In the peer-monitoring groups, when student A submitted a task, his/her partner B would receive the submission along with a peer-monitoring prompt that requested student B to check whether A's argumentation followed the model or not, and if not, to point out where it should be improved. The monitoring task has to be completed before student B can proceed with his/her own work. Apart from peer-monitoring prompts, they also received consistent scaffolding scripts (e.g., guidance on argument construction). In the fading group, students received no peer-monitoring prompts but fading-out scaffolding scripts (i.e., less scaffolding after students' second post). The study found the peer-monitoring scripts significantly promoted learning outcomes while the fading-out scaffolding scripts did not.

In another study that investigated the function of collaborative scripts in argumentative learning, the collaborative scripts not only served as peer-monitoring prompts, but also guided the learner to systematically analyze his/her partner's argumentation and foster his/her own argumentative construction—to paraphrase, criticize, ask questions, give counterarguments, and propose new arguments (Noroozi et al., 2013). Sixty university students participated in the study and collaborated in pairs on tasks that involved both of their disciplinary backgrounds; each pair of students consisted of one from water management and the other from international development studies. The study showed that the scripts were effective in facilitating argumentative knowledge construction (Noroozi et al., 2013).

These findings suggest that collaborative scripts may guide and improve collaborative processes through prompting collaborative behaviors, providing peer monitoring, and facilitating knowledge construction, especially for small-group collaboration. Therefore, we call for more integration of collaborative scripts into teaching practices and more research on strategies to facilitate the collaboration within larger groups.

Generate, deposit, and visualize collective knowledge

Apart from increasing collaborative opportunities and guiding collaborative processes, technological tools or platforms have also been developed to facilitate collective knowledge construction and sharing. After twenty years of development, Wikipedia, an online encyclopedia, now has a repertoire of more than 56 million articles in more than 300 different language versions (https://en.wikipedia.org/wiki/List_of_Wikipedias). It has become one of the most visited websites. The idea that anyone can contribute to the most comprehensive encyclopedia, representing the state of the art of collective human knowledge, has thrived.

Wikipedia or technologies alike have been applied in classrooms at different levels. For instance, Pifarré and Kleine Staarman (2011) examined the collaborative processes of primary students working together to create a science text in a wiki environment. Using a dialogic perspective that attends to, the researchers examined student-generated texts in terms of the number of paragraphs, the number of words, the number of sentences, and the number of reasoning connectors. They concluded that a dialogic approach should be used to analyze the collaborative interaction processes and can be used to help design more effective pedagogies related to the use of wikis in classrooms. At the college level, WIKled Biology is a National Science Foundation funded project in which the researchers developed and implemented a student-centered inquiry learning model for biology classes (Frisch et al., 2013). In the project, students used Web2.0 tools to collaborate with each other, managed and tagged learning materials found on the Internet, and worked in small groups to design and publish website pages based on their scientific inquiry projects. Results showed that the students increased their understanding of certain biology topics, critical thinking skills, as well as the significance of scientific communication.

Like Wikipedia, citizen science can be considered as crowdsourcing science activities to the general public. Citizen science refers to the practice of science endeavor (e.g., data collection) by nonprofessional scientists that contributes to generating professional scientific knowledge (Bonney, 1996). For example, Digital Earth is an effort where people help collect and convert information related to Earth into its digital forms (Brovelli et al., 2020). One example of Digital Earth is OpenStreetMap (www.openstreetmap.org), a collaboration effort to create free, editable map of the world infused with geographic data collected by individual users. Its college student article, YouthMapper (www.youthmappers.org), is a global network of students who actively engage in collaborative mapping using OpenStreetMap. Coetzee et al. (2018) examined student collaborative mapping experiences in three universities in Europe, North America, and Africa. They found that these mapathon (i.e., marathon of mapping) activities could enhance students' learning not only in technical skills and subject matter knowledge, but also in terms of becoming socially engaged citizens.

A pioneer platform of collective knowledge building in K12 classrooms is Knowledge Forum, led by Marlene Scardamalia and Carl Bereiter over three decades of research and development (Scardamalia and Bereiter, 2014). Knowledge Forum is used to help and support knowledge building communities. Students create knowledge notes related to a specific topic. These posts can be organized and connected into different layers and form a part of "World 3—the body of human knowledge expressed in its manifold forms" (Chen and Hong, 2016, p. 269). One strength of Knowledge Forum is the many scaffoldings provided for students to generate their entries. For instance, prescribed prompts such as "My Theory", "My Problem of Understanding" can be selected to start a statement; rise-above notes can be added to provide high-order overviews. Key to these supports is that they are informed by and reinforce the Knowledge Building design principles. Abundant empirical research has shown that students benefit from using Knowledge Forum in improving their basic, domain-specific, and epistemic literacies (Chen and Hong, 2016).

The iKOS platform, a web-based knowledge organization platform, has been developed to facilitate interdisciplinary, collaborative learning in a classroom setting (Namdar and Shen, 2016). The platform allows users to construct, share, and organize knowledge and has the advantage of embedding multimodal elements (Jiang et al., 2019; Shen et al., 2020b). It supports multiple modes of knowledge representation: In the Wiki mode, users can create knowledge entries similar to Wikipedia; In the PicTag mode, users can tag and annotate pictures or photos; In the Concept Map mode, users can construct concept maps to visualize the connections among big ideas; In the Flipbook mode, users can publish their multimodal pages and interactive flipbooks to the general public. Research has shown that preservice science teachers are able to generate a relatively dense knowledge network after taking a unit on nuclear energy (Namdar and Shen, 2018).

Another type of tools we would like to mention is personal response systems (e.g., clickers) or polling apps (e.g., Top Hat). These tools typically go well with the instructional practice called Peer Instruction (Crouch and Mazur, 2001)—a pedagogy developed to engage students in discussion in typically lecture-based classes. A key process is to project and visualize the distribution of students' answers and responses on the screen using the clicker or the polling app. This projection of the status of the collective knowledge of a class drives and engages students in discussing with their neighbors about their choices or reasoning. This method has shown success in improving college students' learning and retention in STEM fields (Deslauriers et al., 2011; Watkins and Mazur, 2013).

The aforementioned platforms provide the space for learners from a large community to deposit, construct, and share collective knowledge. From contributing to the final product of collective knowledge, learners may not only advance community and knowledge building, but also benefit in their individual knowledge construction and other personal skills in the process.

Assess and analyze TECL in STEM

Despite the popularity, relatively little empirical work has been done in the area of assessment of TECL. One may approach assessment of collaborative learning from two aspects: collaborative products and processes. The former refers to what a collaborative group generates during the learning process, i.e., the group-level knowledge/artifacts or group performance of a learning task. These

collective outcomes can be examined to infer the collaboration processes or the team members' collaboration knowledge. The latter focuses on the messy dynamics (e.g., social interactions among group members) that bridge the individual and the collective.

For products, researchers have exploited the discourse or artifacts logged within the technological platforms during collaborative learning. For example, a synchronous or asynchronous discussion forum may record the discourses among members of the learning community, providing rich data source for assessing and analyzing their collaboration as well as knowledge construction. Jimoyianis and Angelaina (2012) conducted Community of Inquiry (CoI) analysis and Social Network Analysis (SNA) on the total 131 blog publications about the acid rain problem. Through CoI analysis, they explored the community characteristics by categorizing the publications into social presence (i.e., involving no domain knowledge but for emotional communication and group cohesion), cognitive presence (i.e., involving domain knowledge), and teacher presence (i.e., including instructions or scaffolding). Their SNA analysis quantified individual participation and the social relations among the individuals. For instance, the number of direct connections each individual had with others may indicate power distribution among students.

For processes, utilizing external tools such as video/audio recording is a common strategy to investigate collaboration processes. For example, Birchfield and Megowan-Romanowicz (2009) examined the use of a mixed-reality environment in facilitating high school students' Earth science learning. Video and audio recordings were coded based on the discourse between the students and the teacher (e.g., questions, responses) and that among students (e.g., comments, questions, or responses between individual students). The findings showed that the utterances were significantly increased in terms of student-driven exchanges.

It is common that a computer-based learning platform logs student behaviors on the platform (e.g., time stamps, clicking a button or page, edit texts). Recently, this type of log data is examined by researchers for understanding group collaborative processes. Olsen et al. (2019) conducted a study to compare how elementary students solved math problems through an online platform with a virtual intelligent tutor in different learning conditions: collaborative, individual, and combined. Researchers analyzed the log data of students' interactions with the virtual tutor—including attempts, errors, and hint requests—as an evaluation of each student's problem-solving efficiency under the assigned learning condition.

Some studies attended to both collaborative processes and products. For example, Altebarmakian and Alterman (2019) examined both the system log data and the contents of students' posts to probe 29 college students' engagement in collaboration in a course on computer science and psychology. Students collaborated on homework in small groups of 3–5 through an educational blog platform. They needed to draft and upload their independent responses to an initial problem, comment on others' posts, reply to questions raised by others, modify their own responses based on group discussions, and submit final responses. Each student's engagement in the collaboration was analyzed in terms of behavioral, social, and cognitive aspects. Behavioral engagement was measured by quantifying the student's reading, editing, and commenting behaviors logged by the system. Interactivity was based on whether or not the student referenced previous comments, which also came from the log data. Cognitive engagement was rated based on the contents of student's posts.

One observation of the assessment aspect of the collaborative learning literature is that most studies used assessments to evaluate the impact of collaboration on individual students' academic or affective outcomes. These assessments do not capture the growth of students' collaboration knowledge and/or collaborative practices. Moreover, the formative assessment during the process receives much less attention in TECL than outcome measures. When students carry out collaborative learning, it is critical to provide timely and effective feedback to individuals and groups. On-going assessment of collaborative learning enables meaningful communication between collaborative group members and between students and their teacher, resulting in a better collaborative learning experience for students. Therefore, more research is needed in the area of assessment in TECL in STEM education.

Conclusion

This article identifies the parameters, forces, and benefits of collaborative learning in STEM education. The article also provides examples of technologies that facilitate collaborative learning in different ways: connecting learners across space and time, fostering written (and multimodal) discourses among learners, guiding collaborative processes, and promoting collective knowledge. Whereas assessments of TECL have been mainly focused on collaborative products and processes to evaluate the impact of collaborative learning on (individual) students' academic and affective outcomes, formative assessment of students' collaboration knowledge and practices needs to be further explored.

References

- Adams, P., 2006. Exploring social constructivism: theories and practicalities. *Educ. 3–13* 34 (3), 243–257.
- Altebarmakian, M., Alterman, R., 2019. Cohesion in online environments. *Int. J. Comput. Support. Collab. Learn.* 14 (4), 443–465.
- Babaian, T., Schiano, B., 2020. Energize Your Online Course with Group Work. A How-to Guide for Making the Most of Digital Breakout Rooms. Accessed on Mar 21st, from. <https://hsbp.harvard.edu/inspiring-minds/energize-your-online-course-with-group-work>.
- Barab, S., Thomas, M., Dodge, T., Carteaux, R., Tuzan, H., 2005. Making learning fun: quest Atlantis, a game without guns. *Educ. Technol. Res. Dev.* 53 (1), 86–107.
- Birchfield, D., Megowan-Romanowicz, C., 2009. Earth science learning in SMAL Lab: a design experiment for mixed reality. *Int. J. Comput. Support. Collab. Learn.* 4 (4), 403–421.
- Bonney, R., 1996. Citizen science: a lab tradition. *Living Bird.* 15 (4), 7–15.
- Bouyias, Y., Demetriadis, S., 2012. Peer-monitoring vs. micro-script fading for enhancing knowledge acquisition when learning in computer-supported argumentation environments. *Comput. Educ.* 59 (2), 236–249.

- Brovelli, M.A., Ponti, M., Schade, S., Solís, P., 2020. Citizen science in support of digital earth. In: Guo, H., Goodchild, M.F., Annoni, A. (Eds.), *Manual of Digital Earth*. Springer, Singapore. https://doi.org/10.1007/978-981-32-9915-3_18.
- Chen, B., Hong, H.-Y., 2016. Schools as knowledge-building organizations: thirty years of design research. *Educ. Psychol.* 51 (2), 266–288.
- Chen, X., Soldner, M., 2013. STEM Attrition: College Students' Paths Into and Out of STEM Fields—Statistical Analysis Report. National Center for Education Statistics, Washington. Retrieved from: <http://nces.ed.gov/pubs2014/2014001rev.pdf>.
- Chen, J., Wang, M., Kirschner, P.A., Tsai, C.C., 2018. The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: a meta-analysis. *Rev. Educ. Res.* 88 (6), 799–843. <https://doi.org/10.3102/0034654318791584>.
- Cheng, Y., Ye, J., 2010. Exploring the social competence of students with autism spectrum conditions in a collaborative virtual learning environment—the pilot study. *Comput. Educ.* 54, 1068–1077.
- Coetzee, S., Minghini, M., Solís, P., Rautenbach, V., Green, C., 2018. Towards understanding the impact of Mapathons—reflecting on youthmappers experience. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-4/W8, 29–31, 2018 FOSS4G 2018—Academic Track.
- Crouch, C., Mazur, E., 2001. Peer instruction: ten years of experience and results. *Am. J. Phys.* 69 (9), 970–977.
- Deslauriers, L., Schelew, E., Wieman, C., 2011. Improved learning in a large-enrollment physics class. *Science* 332 (6031), 862–864.
- Dillenbourg, P., 1999. What do you mean by collaborative learning? In: Dillenbourg, P. (Ed.), *Collaborative-Learning: Cognitive and Computational Approaches*. Elsevier, Oxford, pp. 1–19.
- English, L.D., 2017. Advancing elementary and middle school STEM education. *Int. J. Sci. Math. Educ.* 15 (1), 5–24.
- Esmonde, I., 2009. Ideas and identities: supporting equity in cooperative mathematics learning. *Rev. Educ. Res.* 79 (2), 1008–1043.
- Frisch, J.K., Jackson, P.C., Murray, M.C., 2013. WikiED: using Web2.0 tools to teach content and critical thinking. *J. Coll. Sci. Teach.* 43 (1), 71–80.
- Gokhale, A.A., 1995. Collaborative learning enhances critical thinking. *J. Technol. Educ.* 7 (1), 22.
- Goodyear, P., Jones, C., Thompson, K., 2014. Computer-supported collaborative learning: instructional approaches, group processes and educational designs. In: *Handbook of Research on Educational Communications and Technology*. Springer, New York, pp. 439–451.
- Gressick, J., Derry, S.J., 2010. Distributed leadership in online groups. *Int. J. Comput. Support. Collab. Learn.* 5 (2), 211–236.
- Guffy, S., Gaston, J., Rand, A., 2021. Using Flipgrid as a Reflection Tool to Capture Students' Design Thinking in a Second-Grade Science Classroom. Paper Presented at the 2021 Annual International Conference of National Association for Research in Science Teaching (NARST).
- HESA, 2021. Higher Education Student Statistics: UK, 2019/20—Subjects Studied. <https://www.hesa.ac.uk/news/27-01-2021/sb258-higher-education-student-statistics/subjects>. (Accessed 28 January 2021).
- Hwang, W.Y., Hu, S.S., 2013. Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving. *Comput. Educ.* 62, 308–319.
- Ibanez, M.B., Rueda, J.J.G., Maroto, D., Kloos, C.D., 2013. Collaborative learning in multi-user virtual environments. *J. Netw. Comput. Appl.* 36 (6), 1566–1576.
- Jackson, A.T., Brummel, B.J., Pollet, C.L., Greer, D.D., 2013. An evaluation of interactive tabletops in elementary mathematics education. *Educ. Technol. Res. Dev.* 61 (2), 311–332.
- Jeong, H., Hmelo-Silver, C.E., Jo, K., 2019. Ten years of computer-supported collaborative learning: a meta-analysis of CSCL in STEM education during 2005–2014. *Educ. Res. Rev.* 28, 100284.
- Jiang, S., Smith, B.E., Shen, J., 2019. Examining how different modes mediate adolescents' interactions during their collaborative multimodal composing processes. *Interact. Learn. Environ.* 1–14.
- Jimoyiannis, A., Angelaina, S., 2012. Towards an analysis framework for investigating students' engagement and learning in educational blogs. *J. Comput. Assist. Learn.* 28 (3), 222–234.
- Johnson, D.W., Johnson, R.T., 1990. Using cooperative learning in math. In: Davidson, N. (Ed.), *Cooperative Learning in Mathematics: A Handbook for Teachers*. Addison-Wesley Pub. Co, Menlo Park, Calif, pp. 103–125.
- Johnson, R.T., Johnson, D.W., 1994. An overview of cooperative learning. In: Thousand, J.S., Villa, R.A., Nevin, A.I. (Eds.), *Creativity and Collaborative Learning: The Practical Guide to Empowering Students, Teachers, and Families*, second ed. Brookes Publishing, Baltimore, Maryland, USA.
- Johnson, D.W., Johnson, R.T., 2009. An educational psychology success story: social interdependence theory and cooperative learning. *Educ. Res.* 38 (5), 365–379.
- Ketelhut, D.J., Clarke, J., Nelson, B., 2010. The development of River City, a multi-user virtual environment-based scientific inquiry curriculum: historical and design evolutions. In: Jacobsen, M. (Ed.), *Designs for Learning Environments of the Future: International Learning Sciences Theory and Research Perspectives*. Springer, New York.
- Kuyath, S.J., 2008. The Social Presence of Instant Messaging: Effects on Student Satisfaction, Perceived Learning, and Performance in Distance Education (Order No. 3320887). Available from: ProQuest Dissertations & Theses Global. (304351136). Retrieved from: <http://access.library.miami.edu/login?url=https://www.proquest.com/dissertations-theses/social-presence-instant-messaging-effects-on/docview/304351136/se-2?accountid=14585>.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., Dochy, F., 2013. A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educ. Res. Rev.* 10, 133–149.
- Laal, M., Ghodsi, S.M., 2012. Benefits of collaborative learning. *Procedia Soc. Behav. Sci.* 31, 486–490.
- Le, H., Janssen, J., Wubbels, T., 2018. Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration. *Camb. J. Educ.* 48 (1), 103–122.
- Lee, V.R., Drake, J.R., Thayne, J.L., 2016. Appropriating quantified self technologies to support elementary statistical teaching and learning. *IEEE Trans. Learn. Technol.* 9 (4), 354–365.
- Lichtenberger, E., George-Jackson, C., 2013. Predicting high school students' interest in majoring in a STEM field: insight into high school students' postsecondary plans. *J. Career Tech. Educ.* 28 (1), 19–38.
- Liu, H., Spector, J.M., Ikle, M., 2018. Computer technologies for model-based collaborative learning: a research-based approach with initial findings. *Comput. Appl. Eng. Educ.* 26 (5), 1383–1392.
- Mayer, K.C., 2009. *Collaborative Competition: A Women's Guide to Succeeding by Competing*. Collaborative Competition Press.
- Micari, M., Pazos, P., 2020. Beyond grades: improving college students' social-cognitive outcomes in STEM through a collaborative learning environment. *Learn. Environ. Res.* <https://doi.org/10.1007/s10984-020-09325-y>.
- Michaelsen, L.K., Sweet, M., 2008. The essential elements of team-based learning. *N. Dir. Teach. Learn.* 2008 (116), 7–27.
- Michaelsen, L.K., 1992. Team learning: a comprehensive approach for harnessing the power of small groups in higher education. *To Impr. Acad.* 11 (1), 107–122.
- Namdar, B., Shen, J., 2016. Intersection of argumentation and the use of multiple representations in the context of socioscientific issues. *Int. J. Sci. Educ.* 38 (7), 1100–1132.
- Namdar, B., Shen, J., 2018. Knowledge organization through multiple representations in a computer-supported collaborative learning environment. *Interact. Learn. Environ.* 26 (5), 638–653.
- Noroozi, O., Weinberger, A., Biemans, H.J.A., Mulder, M., Chizari, M., 2013. Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Comput. Educ.* 61, 59–76.
- Nugraha, F., Siahaan, P., Chandra, D.T., 2019. The effect of structured individual responsibility on students' achievement increase in cooperative learning science class. *J. Phys. Conf.* 1157, 032066.
- OECD, 2017. PISA 2015 Collaborative Problem-Solving Framework. <https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>.

- Olsen, J.K., Rummel, N., Alevan, V., 2019. It is not either or: an initial investigation into combining collaborative and individual learning using an ITS. *Int. J. Comput. Support. Collab. Learn.* 14 (3), 353–381.
- Panitz, T., 1999. The motivational benefits of cooperative learning. *N. Dir. Teach. Learn.* 1999 (78), 59–67.
- Pifarré, M., Kleine Staarman, J., 2011. Wiki-supported collaborative learning in primary education: how a dialogic space is created for thinking together. *Int. J. Comput. Support. Collab. Learn.* 6 (2), 187–205.
- Romash, Z.M., 2019. Leaving STEM: An Examination of the Stem to Non-stem Major Change and How the Stem Curriculum Relates to Academic Achievement in Non-stem Fields. Seton Hall University Dissertations and Theses (ETDs), p. 2675. <https://scholarship.shu.edu/dissertations/2675>.
- Rosé, C.P., Ferschke, O., 2016. Technology support for discussion based learning: from computer supported collaborative learning to the future of massive open online courses. *Int. J. Artif. Intell. Educ.* 26 (2), 660–678.
- Roseth, C.J., Johnson, D.W., Johnson, R.T., 2008. Promoting early adolescents' achievement and peer relationships: the effects of cooperative, competitive, and individualistic goal structures. *Psychol. Bull.* 134 (2), 223–246.
- Roth, W.M., Tobin, K., Elmesky, R., Carambo, C., McKnight, Y., Beers, J., 2004. Re/making identities in the praxis of urban schooling: a cultural historical perspective. *Mind Cult. Activ.* 11, 48–69.
- Rummel, N., Mullins, D., Spada, H., 2012. Scripted collaborative learning with the cognitive tutor algebra. *Int. J. Comput. Support. Collab. Learn.* 7 (2), 307–339.
- Scardamalia, M., Bereiter, C., 2006. Knowledge building: theory, pedagogy, and technology. In: *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York, NY, US, pp. 97–115.
- Scardamalia, M., Bereiter, C., 2014. Knowledge building and knowledge creation: theory, pedagogy, and technology. In: Sawyer, K.R. (Ed.), *The Cambridge Handbook of the Learning Sciences*, second ed. Cambridge University Press, Cambridge, UK, pp. 397–417.
- Shen, J., Smith, B.E., Jiang, S., 2020a. Examining Disciplinary Role-Taking in a Collaborative, Multimodal Learning (CML) Environment.
- Shen, J., Smith, B.E., Jiang, S., 2020b. Integrating multimodal composition technology (MCT) in interdisciplinary learning. In: de Oliveira, L.C., Menda, A.M., Vincentini, C. (Eds.), *English Language Teaching Methods, Approaches, and Lessons*. Information Age Publishing.
- Singhal, M.K., 2020. Facilitating virtual medicinal chemistry active learning assignments using advanced Zoom features during COVID-19 campus closure. *J. Chem. Educ.* 97, 2711–2714.
- Sithole, A., Chiyaka, E.T., McCarthy, P., Mupinga, D.M., Bucklein, B.K., Kibirige, J., 2017. Student attraction, persistence and retention in STEM programs: successes and continuing challenges. *High Educ. Stud.* 7 (1), 46. <https://doi.org/10.5539/hes.v7n1p46>.
- Tegos, S., Demetriadis, S., Papadopoulos, P.M., Weinberger, A., 2016. Conversational agents for academically productive talk: a comparison of directed and undirected agent interventions. *Int. J. Comput. Support. Collab. Learn.* 11 (4), 417–440.
- Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., et al., 2018. Integrated STEM education: a systematic review of instructional practices in secondary education. *Eur. J. STEM Educ.* 3 (1), 02.
- Tolmie, A.K., Topping, K.J., Christie, D., Donaldson, C., Howe, C., Jessiman, E., et al., 2010. Social effects of collaborative learning in primary schools. *Learn. InStruct.* 20 (3), 177–191.
- Vygotsky, L., 1978. *Mind in Society: The Development of Higher Psychological Process*. Harvard University Press, Cambridge, MA.
- Wang, H.C., Rosé, C.P., Chang, C.Y., 2011. Agent-based dynamic support for learning from collaborative brainstorming in scientific inquiry. *Int. J. Comput. Support. Collab. Learn.* 6 (3), 371–395.
- Watkins, J., Mazur, E., 2013. Retaining students in science, technology, engineering, and mathematics (STEM) majors. *J. Coll. Sci. Teach.* 42, 36–41.
- Webb, N.M., 1982. Group composition, group interaction, and achievement in cooperative small groups. *J. Educ. Psychol.* 74 (4), 475–484.
- Wenger, E., 2011. Communities of Practice: A Brief Introduction. Paper Presented at the STEP Leadership Workshop, University of Oregon, Oregon, USA. Retrieved Jan 24, 2021 from <https://scholarsbank.uoregon.edu/xmlui/handle/1794/11736>.
- Ziegler, R., Hamker, N., 2011. The idea of a collaborative competition—a case study. *Futures* 43 (4), 441–449.

A pedagogical framework for mobile learning in science education

Xiaoming Zhai and David F. Jackson, Department of Mathematics, Science, and Social Studies Education, University of Georgia, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	215
Mobile learning in science education	216
Portability	216
Accessibility and affordability	216
Multifunctionality	217
Compatibility	217
A three-dimensional pedagogical framework	217
The degree to which technology is integrated with learning activities	217
Substitution	218
Augmentation	218
Learning outcomes varied at different levels of mobile integration	219
Pedagogical roles that students and teachers play	219
Student-led technology	219
Teacher-led technology	220
Collaborative technology	220
Student-led and collaborative functions outperform teacher-led functions on student achievements	221
Student perceptions	221
Remarks	221
Ethical approval	222
References	223

Introduction

The 21st century has witnessed significant progress in science and technology, and people's lives are being dramatically changed by technological advances (Waight and Neumann, 2020). For example, more and more people are shopping online and sometimes having a better experience than with physical shopping; social media giants such as Facebook and Twitter have announced policies of permanent working from home for many of their employees, which would not be possible without highly-developed communication technology (Berliner, 2020). Nevertheless, compared to other aspects of life that science and technology have affected in profoundly revolutionary ways, education continuously remains the area that is least impacted by these advantages (Zhai et al., 2020). Partially, we argue that a reason for this dilemma is due to the very nature of education: A complex process of facilitating the acquisition of knowledge, skills, beliefs, and habits through utilizing media, communicating, and managing. It is not the technology itself, but also the roles of users, the integration of technology and learning, as well as the users' perception of technology, that could facilitate effective learning (Koehler and Mishra, 2009; Daughtery and Berge, 2017; Rosenberg and Koehler, 2015). In light of this fact, even in the area of science education—where science and technology are supposed to have been acknowledged as advantageous (Waight and Abd-El-Khalick, 2012), technology is mostly serving as a supplement, scaffold, or a secondhand substitute for more traditional modes of education (Zhai et al., 2019b). Science education is still suffering from a lack of robust programs to underpin innovations driving workforce development in the future job market (Lobanova and Glushko, 2016).

To cope with such an educational dilemma, the authors of this paper argue that technological innovation could make sense for learning only if embedded in appropriate pedagogy. A similar argument has been made by Crompton and Burke (2018) after reviewing the use of the most updated mobile technology in higher education. While Crompton and Burke identified such a gap between potential and current reality, they did not specify what they saw as the most critical pedagogical aspects of mobile learning. In this paper, we introduced a pedagogical framework for mobile learning developed in the course of our multiple-year project by considering three critical aspects:

- The degree to which technology is integrated with learning activities
- The roles that students and teachers play
- Learners' perception of the technology

We elaborate this argument drawing on data and evidence from literature that particularly focused on the use of mobile technology in science classrooms, including our publications (e.g., Zhai et al., 2018, 2019a,b; Zhai and Shi, 2020). In the following sections, we first review mobile learning in science education. We then articulate a three-dimensional pedagogical framework including the three critical aspects mentioned above. To support this framework, we synthesize evidence for each of the three

dimensions from literature. We conclude with four implications for science educators seeking to better implement mobile technology to facilitate science learning.

Mobile learning in science education

Mobile learning has been a new paradigm in science education. Grant (2019) examined existing definitions of mobile learning and summarized four major categories: relationship to distance education and e-learning, exploitation of devices and technologies, mediation with technology, and the nomadic nature of the learner and learning. These definition categories reflect significant variability in mobile learning which should be grounded in the characteristics of specific disciplinary and research purposes. In science education, students may utilize mobile technology such as tablets or cellphones to solve problems, design solutions, and demonstrate thinking. Such learning may be remote or in-person and can be in formal or informal learning environments (Delen and Krajcik, 2017). Mobile applications in science education have the potential to advance audiovisual representations, increase location awareness, and make virtual learning possible almost anytime and anywhere (Zydney and Warner, 2016; Zhang et al., 2010). Most of these mobile applications, as reviewed by Crompton et al. (2016), yielded positive learning outcomes, drawing great interest in science education. Mobile applications have demonstrated multiple advantages over traditional educational technology in science education, such as portability, accessibility and affordability, multifunctionality, ease of use, and compatibility.

Portability

A fundamental goal of science educators is to engage students in “answering questions about the world or understanding an observed pattern” (National Research Council, 2012). Achieving such a learning goal largely could depend on how we can guide students directly to observe and investigate science phenomena outside of the classroom. Traditional science learning has been constrained in classrooms and labs because most experimental equipment for data collection and analysis is difficult to move to outside of these locations. Therefore, students must primarily rely on experiences of memorization, imagination, and a limited number of lab demonstrations in the classroom to participate in science learning. This learning structure has changed significantly since mobile technology became more widely accessible and is now being applied to science education (Delen and Krajcik, 2017). Mobile technology such as the recently released iPhone 11 Pro, which only weighs 188 g and can provide up to 20 h of video playback, allows for portable, durable, and high-capacity use in science learning. Mobile technology allows students to experience science learning almost everywhere, including in field settings. In their study, Wallace and Witus (2013) incorporated iPad tablet devices into graduate-level K-12 in-service teacher field-based courses. Students used *AMS Sand Sludge Sediment Probe* installed in iPads to collect field data (e.g., shallow sediment cores in different locations), which was difficult to achieve without mobile technology, and analyzed it later in the lab (see Fig. 1 left). Due to the portability of the iPad, they found that teachers increased their confidence in technology use both as students in their outreach class and in their own teaching.

Accessibility and affordability

Accessibility and affordability have been regarded as two of the fundamental factors in engaging students and teachers in using educational technology (Cuban et al., 2001; Ally and Tsinakos, 2014). Mobile technology has been reported to be the most affordable educational technology ever (Tsinakos, 2013). The Pew Research Center (2019) reported that more than 96% of US adults own a mobile device, such as a cellphone or tablet, that can be used for learning. “Digital natives” at the age of eight or less who are immersed in technology have been reported to spend an average of 2.3 h per day using mobile devices (Bernacki et al., 2020).

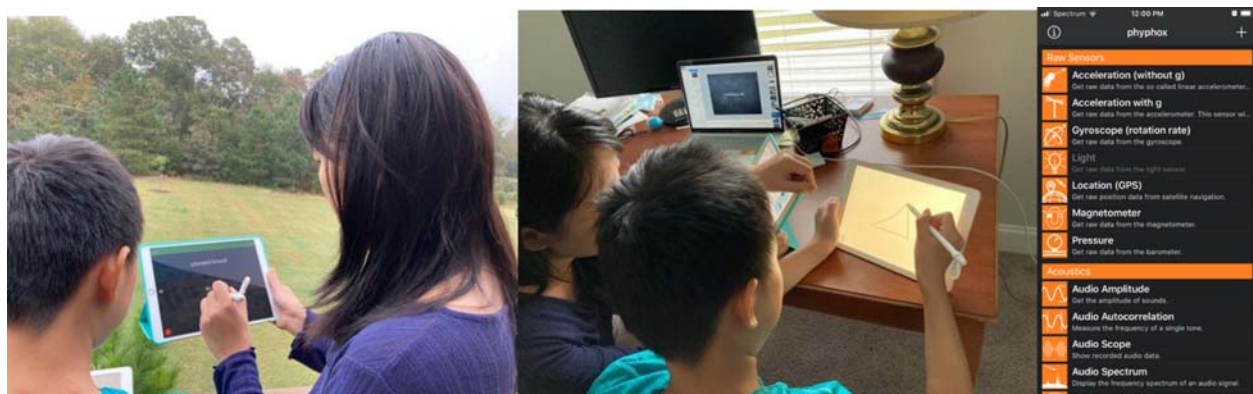


Fig. 1 Field learning using iPad (left) and sensors in Phyphox (right).

Even infants and toddlers, as reported by Levine et al. (2019), are frequently using mobile devices for learning. Given such great accessibility and affordability, it is not surprising that mobile devices have been broadly utilized in formal and informal science learning (Lai, 2020).

Multifunctionality

Science investigators need to collect and analyze data to better understand relationships between variables and to construct explanations for phenomena. The data are not always observable to the naked eye and often need technological devices such as sensors to measure relevant variables. Mobile devices can help in this regard as they are usually equipped with multiple sensors that can be used to detect phenomena and convert them to measurable data. For example, Fig. 1 right shows a free mobile phone application, *Phyphox*, that integrates multiple embedded sensors to collect up to 29 types of data in physics. More importantly, the data collected can be represented in various forms. In their study, Becker et al. (2020) engaged students in using mobile applications that could provide multiple representations of measurement data and found this approach reduced students' extraneous cognitive load and increased their conceptual knowledge, compared to traditional physical experiments.

Compatibility

Compatibility denotes the degree to which the technology is consistent with existing uses, need, and values of users. After surveying 200 college students' use of mobile technology for an online course, Chung et al. (2015) found that compatibility is one of the most prominent predictors of students' intention to use mobile technology in college learning. In science education, high compatibility indicates that students would be able to incorporate mobile technology in aspects of scientific inquiry already commonly acknowledged as important. For example, Liu et al. (2017) developed a mobile application, *InduLab*, for scientific modeling. In their study, students used a smartphone to record the changing angles of a pendulum to collect and store data. The data were plotted within *InduLab* for students to model harmonic motion. They found students were positively engaged in the practice of scientific modeling and 90% of them developed a meaningful explanatory model.

Existing technology such as animations and simulations had been broadly adopted in science learning to support students' scientific inquiry, but such applications were constrained to the computer-based learning environment and thus limited to the science classroom environment. Using mobile technology, Al-Balushi et al. (2017) embedded a package with animations and simulations in a mobile app to help students learn chemistry. Students used the mobile delivered simulations and animations to manipulate variable, collect data, and made predictions. After using the mobile application for 9 weeks, researchers found significant differences in spatial ability between students who used the mobile application and those who did not.

Though mobile technology has all of the potential advantages mentioned above, its ultimate effectiveness depends on how students and teachers use the technology. In the following section, we present a three-dimensional pedagogical framework for analyzing such use.

A three-dimensional pedagogical framework

Though mobile learning has been incorporated in science education to a limited degree for decades (Seppälä and Alamäki, 2003; Chen et al., 2003), there has been a lack of a comprehensive pedagogical framework to guide the use of mobile technologies in science classrooms (Li and Wang, 2018; Daughtery and Berge, 2017). In her review study, Bai (2019) categorized mobile learning into nine categories of pedagogical practices, including situated learning, communication and collaborative learning, game-based learning, inquiry-based learning, informal learning, personalized learning, behaviorist learning activities, learning content access and review, student engagement and motivation. However, she did not abstract the major pedagogical theories implicitly underlying the practices. A review study conducted by Zydney and Warner (2016) suggests that much research on the application of mobile learning is grounded in learning theories, most prominently situated learning. Situated learning theorists argue that science learning needs authentic scenarios to allow students to physically experience the phenomena being investigated, and this perspective has assisted many application designers to identify the critical features of mobile technology to better serve science learning. However, this and other learning theories have not yet widely informed teachers and students in regard to how they should apply the technology in learning activities to better facilitate knowledge construction by specifying what kinds of roles students and teachers should play. More importantly, we don't know how students perceive the usefulness of mobile functions and how their perceptions are associated with their achievement (Crompton and Burke, 2018). Therefore, this article aims to develop a three-dimensional framework to articulate how the critical aspects of pedagogy contribute to mobile learning in science education.

The degree to which technology is integrated with learning activities

The first dimension of the framework looks at the level of integration between mobile technology and learning activities. In their seminal research, Koehler and Mishra (2009) suggested that three forms of knowledge (i.e., of content, pedagogy, and technology) which combine to form technological pedagogical content knowledge (TPACK), are essential for teachers to integrate technology in teaching. This model has been broadly acknowledged and cited in educational research partially because it reveals the fundamental

relationships between technology, content, and pedagogy. Effective instruction has to consider how to integrate the technology with the content and learning activities in a form that could engage students in learning (Kopcha et al., 2020). In a recent study, Jeno et al. (2020) specified the nature of the learning goals (i.e., intrinsic and extrinsic) as an anchor to frame their inquiry tasks for biology students. They randomly assigned students into mobile learning (i.e., a mobile app) and traditional learning (i.e., textbook) groups. Their results suggest an important interaction between mobile learning and intrinsic goal-framing for students' intentions and regulation—how teachers frame learning tasks using technology is critical to the students' success in science learning.

Puentedura (2006) employed Koehler and Mishra's (2009) lens and developed a four-level model (i.e., Substitution, Augmentation, Modification, and Redefinition [SAMR]) to describe how technology can be integrated in learning. In a prior study, Zhai et al. (2018) adopted Puentedura's SAMR model (2006) to categorize 20 specific uses of mobile technology into two of the four levels of integration: the Substitution and Augmentation levels (see Fig. 2). They did not find uses in their study at the Modification and Redefinition levels.

Substitution

At the substitution level, mobile technology is used to directly substitute for conventional learning activities without functional improvement. In their study, Zhai et al. (2018) found that functions such as *My Textbook* was typically used at the substitution level. According to their interview with the lead physics teacher and the authors' close examination of its function, *My Textbook* was simply a digital media representation of the hardcopy of the textbook. Students were able to review the textbook, search content, or make notes on the textbook, but the digital textbook did not include new functionalities. Similar substitution functions included Discussion Aid and Picture Uploading: The former was used to provide resources for discussion which were primarily presented in other forms (e.g., teacher-provided slides) in traditional classrooms; the latter was used to upload students' drawings or pictures to share with others, traditionally achieved by asking students to share their work on the blackboard. These uses have not contributed to functional improvement in conventional science learning.

Augmentation

Mobile technology at the Augmentation level is used to substitute for conventional learning activities with significant functional improvement. For example, there was a mobile function described in Zhai et al. (2018) called *Class Note*, a substitute for traditional note taking. Moreover, it provided functional improvement in allowing students to reorganize, tag, and modify notes much more



Fig. 2 Technology integration model. Modified from Zhai et al. (2018), Puentedura (2006).

efficiently. In Zhai et al. (2018) survey, seventeen specific uses of mobile technology in physics classrooms were classified at this level. These functions represent a higher level of integration with student learning activities than simple substitution.

Learning outcomes varied at different levels of mobile integration

A higher level of mobile integration indicates that the learning activities are better accommodated with mobile technology, indicating greater opportunities for learners to achieve learning goals. Empirical research has shown such potential in mobile learning. For example, to examine how the different levels of integration were associated with students' physics learning achievement, Zhai et al. (2019a) collected students' frequency of use for each function, as well as other information (e.g., demographic, learning achievement and interest), from 803 high-school physics students. According to a multi-level covariate analysis, they found that these mobile technology uses (e.g., *substitution* and *augmentation* uses) could explain 5% of the total variance of students' physics learning achievement. This finding is aligned with assumptions in prior research that the integration levels matter to learning efficacy (Puentedura, 2006). Further analysis found that the frequency of *augmentation* uses had a significant positive correlation with student achievement. However, Zhai, Li and Chen did not find a significant correlation of the frequency of *substitution* uses with student achievement. This result suggested that a higher level of mobile technology integration yields greater physics achievement (for details see Zhai et al., 2019b).

Even though mobile integration levels are critical to the efficacy of using mobile technologies, it is unlikely that teachers would be able to develop high-level of mobile integration learning activities without professional learning. This is because a higher level of integration design would require teachers to have TPACK to transform content knowledge to practice knowledge in designing learning activities that utilizes the advantage of mobile technology. Such discipline-dependent practice knowledge is not likely to be inherent in teachers' general professional preparation. For example, after examining college faculty use of one-on-one iPads in their classrooms, Cavanaugh et al. (2013) found no statistical difference in terms of numbers of teachers between those who designed the Substitution-level learning activities and those who designed higher-level learning activities. Teachers at all levels may only use mobile technology as a substitute rather than to transform learning. Therefore, efforts to support teachers to develop TPACK in specific disciplines are needed.

Pedagogical roles that students and teachers play

Much recent research on pedagogy has highly focused on the roles of students and teachers in processes of technology use (Li and Wang, 2018). This is partially because learning is deemed as not only a psychological but also a social process (Richardson, 2003). While the psychological aspect looks at the developmental characteristics of learning and how students' learning progresses with the support of teachers and technology, the latter aspect focuses on the authority and identity of students and teachers and how their identities may interact with knowledge acquisition and construction. We value both aspects but focus on the latter aspect for this mobile learning research because we regard the primary function of technology as assisting in building a social learning environment. Students have to interact with the environment in order to develop their knowledge and achieve learning goals. Teachers, though, can also use technology to shape that environment, playing a role in facilitating students' learning and help students achieve their learning goals. To have a meaningful interaction with the environment, it is critical to examine the roles of students and teachers play in terms of how they use the technology (see Fig. 3).

In a prior study, Zhai et al. (2019a) classified students' technology uses into three types according to "who leads the use." They acknowledge that the uses may also depend on the specific learning content and situation. One given function may be led by students in one activity but by teachers in another one. Therefore, the classification was based on teachers' typical uses in their specific classrooms.

Student-led technology

In this mode, students have authority to determine when and how to use the technology in specific learning activities, self-regulating their own learning pace as needed. Students are active technology users rather than users passively engaged according to teachers'

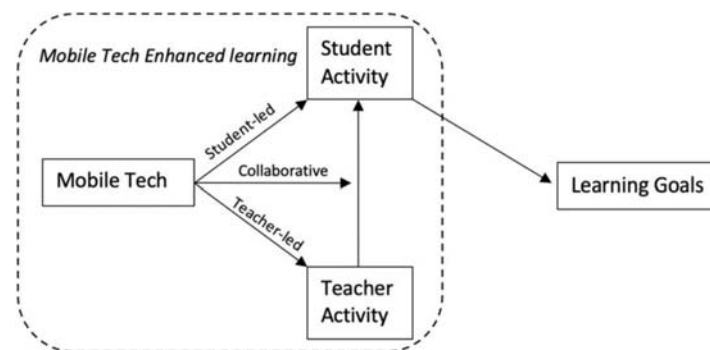


Fig. 3 The STC model. Adopted from Zhai et al. (2019a).

instruction. Note, however, that active technology use does not equal active learning—it is possible that students actively engage in technology use but not knowledge building, even though this type of uses is a trend in mobile learning.

Wang et al. (2020) developed a context-aware plant ecology learning system (CAPELS) to support students' field learning of plant ecology. CAPELS adapts machine learning algorithms to automatically identify plant leaves. While conducting field learning, students could utilize the CAPELS to identify leaves and collect information whenever they felt confused. This implementation is led by students and teachers had limited input in this process, and the study found CAPELS effectively improved students' motivation and learning performance.

In their study, Zhai et al. (2019a) classified student use of the mobile app *Mindmap* as a student-led function. Their research suggests that students primarily used *Mindmap* to construct concept maps for topics and to establish relationships between topics. When learning new concepts, students would revise their concept maps to connect new concepts with their old knowledge structure. Though students did not use *Mindmap* in every class, they constructed or revised concept maps after completing a unit or revisited the concept maps whenever they felt confused by some concepts. Their physics teacher usually spent time to help students understand what concept maps are and how to use *Mindmap*. Teachers also suggested students develop concept maps for each topic, or graded student concept maps, but they seldom explicitly indicated to students when to do so. It was up to the students to decide when and how to use *Mindmap* in their physics learning.

Learning with *Mindmap* was primarily self-regulated by students. Teachers might have included instruction at the beginning of a learning unit or monitor students' uses of the application at its end, but they would not intervene to influence the frequency or timing of students' use during the course. Students had full authority to determine when, where, and how to use the technology. This type of freedom allows students at different learning levels to personalize their learning. For example, lower-level students may want to spend more time on revisiting *Mindmap* than higher-level students when not constrained by any rules.

Note that "student-led uses" of technologies may also be embedded in structured learning activities, in which students have very limited flexibility in leading mobile learning activities. For example, Ahmed and Parsons (2013) described the mobile app *ThinkLearn* and how they implemented this app in abductive science inquiry. Students used the app during four learning phases: exploration, examination, selection, and explanation. However, all the learning activities were ordered by the teacher, and students were involved in a series of linear activities. In this situation, the effectiveness of mobile learning will largely depend on the design of the learning activities. Therefore, we should consider students' authority in not only the use of mobile devices, but also in the larger context of the learning activities.

Teacher-led technology

In this mode, technology is primarily developed and used to support teachers' instructional activities. Often, teachers initiate and direct use of the technology in the technology-enhanced activities, leaving little or no opportunity for students to direct the use of the technology. In their prior study, Zhai et al. (2019a) described a mobile function named *Doodle* used by students and teachers in physics learning. By using *Doodle*, students could draw models, take screenshots and submit them to teachers. For example, in teaching the concept of acceleration to high school students, the teacher broadcast a video on student tablets showing how a motor and a plane accelerate when they start to move. He then asked students to make screenshots at time 2 s, 4 s, 6 s, 8 s, and 10 s using the *Doodle* tool, and then draw arrows to represent the speed of both the motor and the plane. Students watched video, made screenshots, drew arrows, and submitted their "doodles" to teachers following teachers' instructions. In this example of using *Doodle*, the teacher initiated the use and instructed students to do exactly what he wanted students to do. Students had almost no authority to determine when and how to use the function, even though they used the function to actively add in their own drawing. This use had constrained students' freedom of use and seemed primarily to serve for the teacher to organize the instruction.

Some studies suggest that this teacher-led mode may have less impact on students' learning performance. For example, Lee et al. (2019) investigate the use of an instant response system *Kahoot* in middle school earth science. They found no significant difference in terms of learning motivation, value of earth science, expectation of earth science, and learning outcomes between students who used *Kahoot* in the experimental group and those in the control group. This is not surprise as students were purely passively involved in the mobile learning and there is little functional improvement provided by the mobile function compared to traditional learning.

Collaborative technology

When both students and teachers actively participate in the mobile learning, technology use becomes a medium to connect students and teachers in learning activities. In this mode of technology use, either students or teachers may initiate the use while the other may follow and give active input on their use of the technology. The major characteristic of collaborative technology use, as compared to student- and teacher-led functions, is that both students and teachers play active and significant roles. Note that collaborative technology use is different from collaborative learning, which emphasizes collaboration between students. In their research, Zhai et al. (2019a) found the digital *Learning Guide* was frequently used in a collaborative manner, with teachers and students both actively involved. Sometimes teachers required students to preview *Lesson Guide* before class—teachers played a lead role in initiating the use. However, students also spontaneously referred to *Lesson Guide* because they found it useful in many ways. For example, students might raise questions when reviewing the *Learning Guide*, and they would make notes on it or send a message to their teacher. Teachers could reply responsively to their questions because of the high accessibility of mobile technology. Mobile apps like the *Learning Guide* plays a significant role in connecting teachers and students.

Student-led and collaborative functions outperform teacher-led functions on student achievements

Consistent with constructivist views of teaching and learning, student-led functions and collaborative functions both actively involve students in mobile learning, though to a varying degree. Student-led mobile functions require students to be most actively involved in science learning, compared to collaborative and teacher-led functions. Therefore, students who frequently used this type of functions are expected to outperform other students, and empirical evidence supports this. For example, Zhai et al. (2019a) examines the effectiveness of the three types of mobile technology uses. In their study, Zhai et al. (2019a) classified 15 specific mobile functions into three groups according to “who led the use.” They conducted an analysis of variance on achievement scores using the class as a fixed independent variable and controlling for the pretest score and found that there is a significant main effect of class on student post-test achievement. After controlling for students’ gender and the pre-score, student-led functions and collaborative functions both significantly predicted students’ physics achievement, but teacher-led functions failed to offer that prediction. Their results suggest that for one standard deviation increase of student-led functions of mobile technology use, student achievement increased 0.20 standard deviations, higher than for collaborative function use (0.08 standard deviations). This finding explicitly supported their prediction.

Student perceptions

Student perception of technology has been regarded as one of the most crucial factors that may impact how technology is used in the classroom (Ferguson and Bargh, 2004; Karahanna and Straub, 1999). This insight also applies to mobile technology in science education. According to their review, Crompton and Burke (2018) suggested that 29% of mobile learning studies they reviewed focused on students’ perceptions of mobile learning. This finding is not surprising as student-perceived usefulness has been found to be a factor that may influence student behavior. The question remains how students’ perception is associated with frequency of technology use and student learning outcomes.

The learner’s perception of technology is critical and may be moderated by factors such as the roles that students and teachers play in mobile learning. In their study, Zhai and Shi (2020) developed a path model accounting for this mechanism. In the model, they used student physics learning interest and prior physics learning scores as covariates and examined how the frequency of use of 15 mobile functions directly and indirectly predict physics learning outcomes (see Fig. 4). Given that male and female students may have different perceptions of mobile technology (Wang and Wang, 2008; Wang et al., 2009), they applied a multi-group structural equation model to analyze the data collected with a total of 803 high school students. Their findings suggest that perceived usefulness was a significant predictor for use frequency by pedagogical category for both male and female students. They found that the student-led and collaborative functions had smaller coefficients than that of teacher-led functions for both groups—when teachers, as compared to students, have greater authority in classroom in terms of directing the frequency of technology use, student perception of the technology matters less.

Student gender may moderate their perceptions of the mobile usefulness. Zhai and Shi (2020) suggest a significant total effect on student physics achievement of student perception of collaborative functions for male students, while student perception of student-led functions had a significant total effect for female students. The findings suggest that male and female students’ perception may result in different learning outcomes due to their roles in the uses of mobile technology. For both male and female students, the perception of teacher-led mobile functions had a total negative effect on student physics achievement.

Remarks

While earlier research was concerned with the paradox of “high access and low use” (Cuban et al., 2001), mobile technology is now advancing to the stage of “high use and high effectiveness.” Recent research (Zhai et al., 2018) has indicated that mobile technology, differentiated from other electronic technologies, is now frequently used in science learning. “Low use” has not yet been a concern for mobile technology due to its multi-functionality, portability, and accessibility. The concern became how to make use of mobile technology to facilitate effective science learning. This article provides a comprehensive pedagogical framework for us to better understand the uses of mobile technology in science learning.

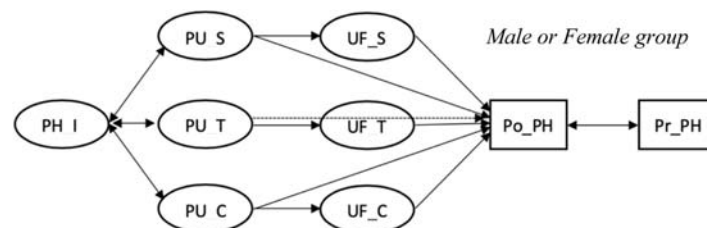


Fig. 4 Theoretical path model. *Note.* PH_I = physics interest, PU_S = perceived usefulness_student-led, PU_T = perceived usefulness_teacher-led, PU_C = perceived usefulness_collaborative, UF_S = use frequency_student-led, UF_T = use frequency_teacher-led, UF_C = use frequency_collaborative, Po_PH = post physics achievement, Pr_PH = pre physics achievement. Adapted from Zhai and Shi (2020).

We acknowledge that with the advent of wide availability of mobile technology to nearly all teachers and students, it becomes increasingly realistic for science teachers to broaden their vision of settings appropriate for formal science learning to incorporate field settings, including not only the types of locations where professional scientists might gather research data, but also the homes and local communities of students. However, understanding how to better implement mobile learning in science education is critical (McMahon et al., 2016). Prior research has highlighted the importance of pedagogical context in mobile learning (Dennen and Hao, 2014; Aguayo et al., 2017; Li and Wang, 2018); yet no research has explicitly developed a comprehensive framework to help researchers and educators in science to understand this novel learning opportunity. Such a framework is especially valuable since the COVID-19 pandemic has significantly impacted schooling and many students have had to rely on remote learning (Bond et al., 2020). Without such a framework, mobile technology may not be used appropriately or perceived as useful in science learning, especially when more students are learning remotely than ever before during this critical pandemic time. The framework provides such a vision for understanding mobile learning and has implications for the following aspects of it.

First, mobile technology, with all of its advantages (over full-size computers) in portability, accessibility, affordability, and multifunctionality, is unlikely to improve science learning unless its use substantively and qualitatively enhances the capabilities of students to engage with scientific ideas. Providing a more convenient substitute for functions normally or previously carried out by teachers and/or students (either with full-size computers or in “low-tech” fashion) without its assistance is not sufficient for measurable improvements in achievement. This knowledge should help science educators to *develop learning materials and activities that are deeply integrated with mobile technology*. Such uses should prove especially valuable for engaging students in scientific practices such as modeling, argumentation, and computational thinking using mobile technology, in classroom, field, or remote learning spaces (Liu et al., 2017; Hwang et al., 2013).

Second, student-led functions of mobile technology—that is, those for which students, rather than their teacher, are in control of at least the timing and frequency of their use in science learning (if not, in practice, often the nature or purpose of their use)—are more likely to lead to increased learning, if the mobile devices are used often, than teacher-led functions—those based on prescriptive, inflexible instructions from the teacher—even if the students are operating the devices rather than merely observing them used by their teacher. Collaborative use of mobile technology, in which the devices provide a means for students and teachers to initiate (from either side of the dyad) and maintain interactive communication, were associated with more modest learning gains in these studies, but their relatively infrequent use suggests that examining classrooms in which they are more common would be a promising future direction for research. Findings in this aspect highlight the roles that students and teachers can potentially play in mobile learning. This dimension of the framework provides a lens for educators to *design their learning activities by coordinating students’ and teachers’ roles in a way that is effective for science learning*.

Third, the framework highlights learners’ perceptions about mobile technology. Though perception has been used before to account for why users accept technology (Karahanna and Straub, 1999), there remains a need to understand how this perception impacts the actual frequency of use and student learning outcomes. The mechanism tends to be complex—that is, high perceived usefulness suggests high use frequency but does not necessarily indicate high effectiveness. This is because students’ perception is associated with not only their use frequency but also their roles, gender, etc. The potential for more profound, truly transformative effects on science learning (major modifications of the design of learning tasks or, eventually, the facilitation of entirely new tasks practically impossible without mobile technology) would require careful simultaneous attention to a large number of factors. No instances of such outcomes were observed in these studies, despite quantitatively demonstrable incremental gains in learning associated with the several promising theoretical frames we employed. Simply assuming a relationship between perceived usefulness and learning outcomes may not work in authentic and complicated science learning environments. This suggests that *multivariate quantitative research studies of affective factors contributing to student achievement* in such exemplary classrooms would probably be premature, and exploratory, descriptive, largely or wholly qualitative studies may be needed to generate new theoretical constructs and/or hypotheses about quantitative effects in such learning environments.

Last but not least, gender differences in science learning have long been of great interest to science educators and have often been posited (and sometimes empirically identified) to be most prominent in physics in particular (Brewer and Sawtelle, 2016). The studies summarized here suggest that further research on the effects of student perceptions of the use of mobile technologies may be a particularly fruitful new angle from which to view the role of gender in the physics classroom, and that emphasizing student-led functions of these technologies may improve the achievement of female students in particular. Specifically, male students’ perceptions of technologies with which students play a leading role in the uses may positively affect their learning outcomes, while female students’ perceptions of collaborative use of technology is a positive predictor of theirs. Thus, educators are strongly encouraged to *implement mobile learning in science in light of the effect of student gender differences in their perceptions, frequency of use, as well as the nature of their roles in technology use*.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed Consent: Not applicable.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

- Aguiayo, C., Cochrane, T., Narayan, V., 2017. Key themes in mobile learning: prospects for learner-generated learning through AR and VR. *Australas. J. Educ. Technol.* 33.
- Ahmed, S., Parsons, D., 2013. Abductive science inquiry using mobile devices in the classroom. *Comput. Educ.* 63, 62–72.
- Al-Balushi, S.M., AL-Musawi, A.S., Ambusaidi, A.K., AL-Hajri, F.H., 2017. The effectiveness of interacting with scientific animations in chemistry using mobile devices on grade 12 students' spatial ability and scientific reasoning skills. *J. Sci. Educ. Technol.* 26, 70–81.
- Ally, M., Tsinakos, A., 2014. Increasing Access Through Mobile Learning. *Commonwealth of Learning (COL)*.
- Bai, H., 2019. Pedagogical practices of mobile learning in K-12 and higher education settings. *Tech Trends* 63 (5), 611–620.
- Becker, S., Klein, P., Gösslin, A., Kuhn, J., 2020. Using mobile devices to enhance inquiry-based learning processes. *Learn. InStruct.* 69, 101350.
- Berliner, U., 2020. Get a Comfortable Chair: Permanent Work from Home is Coming. Available: <https://www.npr.org/2020/06/22/870029658/get-a-comfortable-chair-permanent-work-from-home-is-coming>. (Accessed 12 September 2020).
- Bernacki, M.L., Greene, J.A., Crompton, H., 2020. Mobile technology, learning, and achievement: advances in understanding and measuring the role of mobile technology in education. *Contemp. Educ. Psychol.* 60, 101827.
- Bond, E.C., Dibner, K., Schweingruber, H., 2020. Reopening K-12 Schools During the COVID-19 Pandemic.
- Brewe, E., Sawtelle, V., 2016. Focused collection: gender in physics. *Phys. Rev. Phys. Educ. Res.* 12, 020001.
- Cavanaugh, C., Hargis, J., Kamali, T., Soto, M., 2013. Substitution to Augmentation: Faculty Adoption of iPad Mobile Learning in Higher Education. *Interactive Technology and Smart Education*.
- Chen, Y.-S., Kao, T.-C., Sheu, J.-P., 2003. A mobile learning system for scaffolding bird watching learning. *J. Comput. Assist. Learn.* 19, 347–359.
- Chung, H.-H., Chen, S.-C., Kuo, M.-H., 2015. A study of EFL college students' acceptance of mobile learning. *Proc. Soc. Behav. Sci.* 176, 333–339.
- Crompton, H., Burke, D., 2018. The use of mobile learning in higher education: a systematic review. *Comput. Educ.* 123, 53–64.
- Crompton, H., Burke, D., Gregory, K.H., Gräbe, C., 2016. The use of mobile learning in science: a systematic review. *J. Sci. Educ. Technol.* 25, 149–160.
- Cuban, L., Kirkpatrick, H., Peck, C., 2001. High access and low use of technologies in high school classrooms: explaining an apparent paradox. *Am. Educ. Res. J.* 38, 813–834.
- Daughtery, C., Berge, Z.L., 2017. *Mobile Learning Pedagogy*. UMBC Faculty Collection.
- Delen, I., Krajcik, J., 2017. Using mobile devices to connect teachers and museum educators. *Res. Sci. Educ.* 47, 473–496.
- Dennen, V.P., Hao, S., 2014. Intentionally mobile pedagogy: the M-COPE framework for mobile learning in higher education. *Technol. Pedagog. Educ.* 23, 397–419.
- Ferguson, M.J., Bargh, J.A., 2004. How social perception can automatically influence behavior. *Trends Cogn. Sci.* 8, 33–39.
- Grant, M.M., 2019. Difficulties in defining mobile learning: analysis, design characteristics, and implications. *Educ. Technol. Res. Dev.* 67 (2), 361–388.
- Hwang, G.-J., Wu, P.H., Zhuang, Y.Y., Huang, Y.M., 2013. Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interact. Learn. Environ.* 21, 338–354.
- Jeno, L.M., Dettweiler, U., Gytmes, J.-A., 2020. The effects of a goal-framing and need-supportive app on undergraduates' intentions, effort, and achievement in mobile science learning. *Comput. Educ.* 159, 104022.
- Karahanna, E., Straub, D.W., 1999. The psychological origins of perceived usefulness and ease-of-use. *Inf. Manag.* 35, 237–250.
- Koehler, M., Mishra, P., 2009. What is technological pedagogical content knowledge (TPACK)? *Contemp. Issues Technol. Teach. Educ.* 9, 60–70.
- Kopcha, T.J., Neumann, K.L., Ottenbreit-Leftwich, A., Pitman, E., 2020. Process over product: the next evolution of our quest for technology integration. *Educ. Technol. Res. Dev.* 68, 729–749.
- Lai, C.L., 2020. Trends of mobile learning: a review of the top 100 highly cited papers. *Br. J. Educ. Technol.* 51, 721–742.
- Lee, C.-C., Hao, Y., Lee, K.S., Sim, S.C., Huang, C.-C., 2019. Investigation of the effects of an online instant response system on students in a middle school of a rural area. *Comput. Hum. Behav.* 95, 217–223.
- Levine, L.E., Waite, B.M., Bowman, L.L., Kachinsky, K., 2019. Mobile media use by infants and toddlers. *Comput. Hum. Behav.* 94, 92–99.
- Li, Y., Wang, L., 2018. Using iPad-based mobile learning to teach creative engineering within a problem-based learning pedagogy. *Educ. Inf. Technol.* 23, 555–568.
- Liu, C.-Y., Wu, C.-J., Wong, W.-K., Lien, Y.-W., Chao, T.-K., 2017. Scientific modeling with mobile devices in high school physics labs. *Comput. Educ.* 105, 44–56.
- Lobanova, Y.I., Glushko, K.V., 2016. About assessment criteria of driver's accidental abilities. *Int. J. Environ. Sci. Educ.* 11, 9969–9979.
- McMahon, D., Wright, R., Cihak, D.F., Moore, T.C., Lamb, R., 2016. Podcasts on mobile devices as a read-aloud testing accommodation in middle school science assessment. *J. Sci. Educ. Technol.* 25, 263–273.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- Pew Research Center, 2019. *Mobile Fact Sheet—Mobile Phone Ownership Over Time*.
- Puentedura, R.R., 2006. SAMR: A Brief Introduction. unpublished. Retrieved from: http://www.hippasus.com/rpweblog/archives/2013/10/02/SAMR_ABriefIntroduction.pdf.
- Richardson, V., 2003. Constructivist pedagogy. *Teach. Coll. Rec.* 105, 1623–1640.
- Rosenberg, J.M., Koehler, M.J., 2015. Context and technological pedagogical content knowledge (TPACK): a systematic review. *J. Res. Technol. Educ.* 47, 186–210.
- Seppälä, P., Alamäki, H., 2003. Mobile learning in teacher training. *J. Comput. Assist. Learn.* 19, 330–335.
- Tsinakos, A., 2013. State of mobile learning around the world. In: *Global Mobile Learning Implementations and Trends*, pp. 4–44.
- Waight, N., Abd-El-Khalick, F., 2012. Nature of technology: implications for design, development, and enactment of technological tools in school science classrooms. *Int. J. Sci. Educ.* 34, 2875–2905.
- Waight, N., Neumann, K., 2020. 21st-Century science education digital ecologies: technology, technique, shoelaces, promise, and pitfalls? *J. Res. Sci. Teach.* 57, 1313–1321.
- Wallace, D.J., Witus, A.E., 2013. Integrating iPad technology in earth science K-12 outreach courses: field and classroom applications. *J. Geosci. Educ.* 61, 385–395.
- Wang, H.Y., Wang, Y.S., 2008. Gender differences in the perception and acceptance of online games. *Br. J. Educ. Technol.* 39, 787–806.
- Wang, Y.S., Wu, M.C., Wang, H.Y., 2009. Investigating the determinants and age and gender differences in the acceptance of mobile learning. *Br. J. Educ. Technol.* 40, 92–118.
- Wang, C.-C., Lo, C.-L., Hsu, M.-C., Tsai, C.-Y., Tsai, C.-M., 2020. Implementation a context-aware plant ecology mobile learning system. *SAGE Open* 10, 2158244020920701.
- Zhai, X., Shi, L., 2020. Understanding how the perceived usefulness of mobile technology impacts physics learning achievement: a pedagogical perspective. *J. Sci. Educ. Technol.* 29, 743–757.
- Zhai, X., Zhang, M., Li, M., 2018. One-to-one mobile technology in high school physics classrooms: understanding its use and outcome. *Br. J. Educ. Technol.* 49, 516–532.
- Zhai, X., Li, M., Chen, S., 2019a. Examining the uses of student-led, teacher-led, and collaborative functions of mobile technology and their impacts on physics achievement and interest. *J. Sci. Educ. Technol.* 28, 310–320.
- Zhai, X., Zhang, M., Li, M., Zhang, X., 2019b. Understanding the relationship between levels of mobile technology use in high school physics classrooms and the learning outcome. *Br. J. Educ. Technol.* 50, 750–766.
- Zhai, X., Haudek, K.C., Shi, L., Nehm, R., Urban-Lurain, M., 2020. From substitution to redefinition: a framework of machine learning-based science assessment. *J. Res. Sci. Teach.* 57, 1430–1459.
- Zhang, B., Looi, C.-K., Seow, P., Chia, G., Wong, L.-H., Chen, W., So, H.-J., Soloway, E., Norris, C., 2010. Deconstructing and reconstructing: transforming primary science learning via a mobilized curriculum. *Comput. Educ.* 55, 1504–1523.
- Zydney, J.M., Warner, Z., 2016. Mobile apps for science learning: review of research. *Comput. Educ.* 94, 1–17.

Citizen science in STEM education: engaging students with real life science

Keren E. Dalyot^a and Yaela N. Golumbic^b, ^a Faculty of Education in Science and Technology, The Technion - Israeli Institute of Technology, Haifa, Israel; and ^b Faculty of Science, The University of Sydney, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	224
Learning through citizen science	225
Citizen science in schools—exploring a broad potential	226
Societal aspects of science	226
Awareness to science and the environment	227
Civic action	227
Case studies	228
Enhancing learning and environmental stewardship through citizen science: the case of the radon home survey	228
Learning about the complexity of science through citizen science: the case of breaking good	229
The classroom as a platform for community engagement: the case of Food Science Class	229
Integration	230
Summary	231
Acknowledgment	232
References	232
Further reading	233
Relevant websites	233

Introduction

In recent years, there has been widening interest in and opportunities for public engagement with science through citizen science initiatives, worldwide (Bonney et al., 2014). As a collaborative effort in nature, citizen science promotes public participation in scientific endeavors and can bring communities and scientists together around scientific projects that are local to global in scope. Through this collaboration, great potential exists for the advancement of science alongside transformative effects, such as development of individual skills, learning processes and creating a dialog between researchers and participants (see Fig. 1) (Shirk et al., 2012; Turrini et al., 2018).

Citizen science has many origins, from 19th century weather observations to 20th century environmental monitoring, to 21st century big data opportunities. Today citizen science is used to advance research in biology, astronomy, environmental sciences, physics, and other fields. Each project represents a different vision of what non-experts contribute to the production of reliable knowledge about the natural, social, and cultural worlds as well as a vision for building students' capacity for engaging with real life socio-scientific dilemmas.

Organizations that support citizen science initiatives are mushrooming in many places around the world. Here we note for example the USA based Citizen Science Association (CSA), the European Citizen Science Association (ECSA), the Australia Citizen

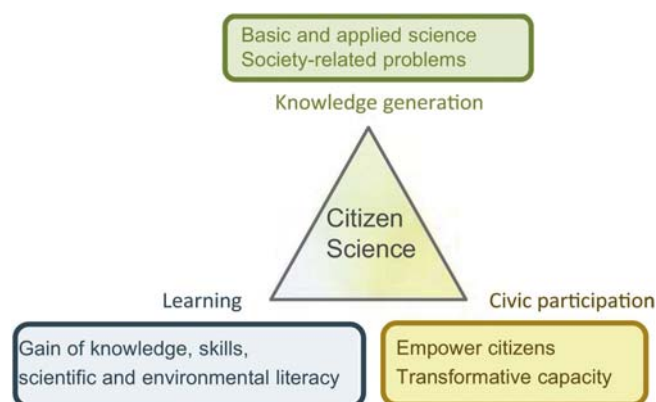


Fig. 1 The threefold potential of citizen science for generating new knowledge, learning outcomes and civic participation. Source: Turrini, T., Dörler, D., Richter, A., Heigl, F., Bonn, A., 2018. The threefold potential of environmental citizen science—generating knowledge, creating learning opportunities and enabling civic participation. *Biol. Conserv.* 225, 176–186. <https://doi.org/10.1016/j.biocon.2018.03.024>, p. 177.

Science Association (ACSA), the CitizenScience.Asia association and others (websites of all of these organization appear at the relevant websites section at the end of this article). The overarching goal of these organizations is to support the theory and practice of citizen science, as they play a critical role in facilitating and establishing citizen science as a field of practice and area of scholarship, and cultivate a new citizen science culture (Storksdiack et al., 2016). The institutionalization of citizen science is also broadening with recognition by policymakers and funding agencies, the launch of a specified citizen science journal and several professional and practitioner conferences.

The book *"The Science of Citizen Science"* (Vohland et al., 2021) published by the citizen science COST action, outlines the various aspects of citizen science and reflects on the contribution of citizen science to societal development, education, and innovation. Using a holistic approach and viewing citizen science as a social and civic enterprise, multiple stakeholders of citizen science are outlined, from scientists to citizens and from policy makers to environmental organizations. In addition to the contribution of citizen science, the book deals with ways to improve citizen science practice and discusses its scientific quality, data standardization, use of technologies, participant motivations and developing synergies with education systems.

To this end, Kloetzer et al. (2021) discuss the diverse ways in which participants of citizen science learn and the many settings in which such learning can take place. These include online settings, museums, community engagement, families and formal and out-of-school education settings. Because citizen science opens new opportunities for science learning, especially when conducting scientific inquiry with students (Bonney et al., 2014), these projects are becoming more and more popular in the field of science education. Schools, students, and teachers are increasingly seen as important stakeholders of citizen science and dedicated programs are created for incorporating citizen science in schools (Shah and Martinez, 2016; Harlin et al., 2018).

The goal for this article is to discuss the potential for incorporating citizen science in schools for promoting science education, active citizenship, and a vision of science as relevant and important to students' daily lives. As we sketch out the web of potential educational benefits citizen science has to offer, we illustrate a lively science education arena, in which students and schools actively engage with science through citizen science, leading to diverse learning outcomes.

Our article is designed to elicit a discussion about the different ways citizen science in schools can promote meaningful and broad learning outcomes, and provide students the opportunities to develop knowledge and skills they need to live in today's scientifically and technologically enhanced society (Ryder, 2001; National Academies of Sciences Engineering and Medicine, 2016).

Through presenting three case studies, representing the varied landscape of citizen science projects in schools, we will expand upon the many different ways learning can take place and discuss ways in which citizen science can sustain active citizenship and increase student awareness and relatedness to science.

Learning through citizen science

Research into learning opportunities through citizen science reveal diverse learning outcomes from student participation. These may include cognitive dimensions (what people know), affective dimensions (what people feel) and behavioral dimensions (what people do) (Phillips et al., 2014). Citizen science enables students to engage in hands-on, authentic research and learn about science from broad perspectives. A variety of individual learning outcomes from citizen-science have been documented. These include interest in science and the environment; self-efficacy for science and the environment; motivation for science and the environment; increased scientific knowledge; skills of science inquiry; and environmental stewardship (Phillips et al., 2018). At the same time, how this learning takes place and what factors are involved in these processes are not well understood. Furthermore, while some participants in citizen science projects have demonstrated increased content knowledge and scientific skills, higher level learning and increased understanding of the Nature of Science (NOS), its relevance to day-to-day life has not been documented in school context.

In 2018 the National Academies of Sciences, Mathematics and Engineering published a consensus report: *Learning Through Citizen Science: Enhancing Opportunities by Design*, that set out to examine the complex relationship between the potential of citizen science for science education and the way this potential is fulfilled as well as to set out the evidence supporting this claim. The report highlights a list of conclusions and recommendations, three of which are relevant to our discussion of learning outcomes. First, the writers conclude that "Citizen science supports learning outcomes related to scientific practices, content, identity, agency, data, and reasoning" (p. 147), however they emphasize that to realize these outcomes the design of the projects needs to be focused on them and some type of support for learning needs to be provided. They stress: "Without intentional design, it can be hard to anticipate what learning outcomes will be achieved" (p. 150). Moreover, the report ends with a recommendation that citizen science projects who wish to enhance learning opportunities need to work toward designing the project in a premeditated way, working with stakeholders and especially those with expertise in education to design the project. Similar findings have been discussed by Roche et al. (2020), who highlight the competing goals of science, learning and social advocacy and call for the alignment of goals through deliberation and co-creation.

In an attempt to support and direct learning outcomes from citizen science, multiple initiatives for incorporating citizen science in schools have been established. Through engaging students in real-world authentic investigations, these initiatives aim to increase student's interest and eventually, learning. Some examples include:

- Taking Citizen Science to Schools (TCSS) center in Israel. A collaboration between two academic institutions linking education researchers, scientists and schools, in an effort to facilitate implementation of a variety of citizen science projects in schools and increase student learning outcomes. TCSS focuses on both the practice of implementing citizen science in schools and the practical knowledge needed for sustaining and enhancing such endeavors. They emphasize on developing research-practice partnerships for ongoing evaluation making projects relevant and beneficial for multiple stakeholders, especially schools.

- The citizen science platform in the New Zealand science learning hub. This is a digital platform designed for teachers who are interested in incorporating citizen science in their classroom, offering teachers relevant projects, proposed classroom activities and dedicated lesson plans. The platform also offers webinars on citizen science and allows teachers to tailor the educational work to specific scientific skills they are interested in.
- Cornell Lab of Ornithology- K-12 Education Team. The Cornell lab is one of the pioneers of citizen science in North America. Their education team is devoted to creating resources for both formal and informal education teams, using an inquiry-based approach to connect students to their local habitats and build their science skills. The Cornell Lab K-12 Education team offers teachers professional development opportunities and educator's guides as well as curriculum-based activities for students and activities for children out of school.

These initiatives, among others, are focused around learning outcomes and developing scientific skills. They demonstrate the growing interest in incorporating citizen science into formal education systems as mentioned in the report above, and the vast potential for increasing long term learning outcomes. But can citizen science go beyond conventional learning outcomes, promoting critical thinking and urging students to realize the complexity of science as a real-world endeavor? This will be the focus in the next section of this article where we further explore the potential of citizen science for promoting a holistic view of science, civic action and democratic participation.

Citizen science in schools—exploring a broad potential

In a paper published a few years ago in *Science Magazine*, citizen science was suggested as a form of collaborative research among scientists, educators, and the public (Wals et al., 2014). Wals et al. (2014) argue that citizen science might lead to more effective processes of public engagement and learning that could stimulate meaningful socio-ecological outcomes. In the “post-truth” era where trust in science seems to be diminishing, formal education can benefit from becoming part of new hybrid forms of knowledge co-creation, that seek to contribute to a more sustainable world. Furthermore, with such ill-defined, highly contextual and ambiguous key issues of our time (also described as wicked problems that are difficult to solve because of incomplete, contradictory, and often changing and undefined requirements), alternative forms of research and inquiry are required. Citizens, including younger generations, need to have agency; access to multiple knowledges—including scientific knowledge, indigenous knowledge and local knowledge—and the opportunity to reflexively act to improve a situation jointly with the multiple stakeholders affected by an issue (Wals and Peters, 2017).

However, to realize this vision, a more holistic approach is needed for ways to engage learners for optimal individual and community outcomes. Learning which goes beyond scientific knowledge and incorporates social dimensions, cultural awareness and individual responsibility. In this section we will focus on three important aspects that contribute to meaningful impact of citizen science projects in schools and to broadening the learning outcomes. These include (a) learning about Societal Aspects of science, (b) raising Awareness to science and the environment and (c) engaging in Action on the local and global level (SAAA). We will then introduce three case studies that demonstrate how these aspects play out in three citizen science projects of distinct nature.

Societal aspects of science

Science, or rather the scientific method, is comprised of diverse practices and concepts that collectively aim to minimize error and provide reliability and reportability to a study (Bauer, 1994). The value of this practice is in developing a culture or mindset of thinking both analytically and critically about complex situations. It provides the ability to make logically deductive arguments from simple premises, and identify salient variables, patterns in data and numerical fluency (Claussen and Osborne, 2013). These scientific practices take into consideration a number of core ideas and concepts, which are fundamental for scientific inquiry and research, as well as application of science in everyday life.

However, formal science education has a tradition of disengaging science from the real life of students (Feinstein, 2011) and not proving as useful for decision making in later life (Crowell and Schunn, 2016). Many science education researchers offer ways in which teachers can make science more relevant, often by using what has been termed Socio-Scientific Issues (SSI) (Tal and Dierking, 2014; Kinslow and Sadler, 2018). SSI are science-related topics which have a strong social dimension and are often controversial, such as environmental and health related issues. Introducing such issues in a classroom environment is suggested as an effective way to engage students with science while considering its social, cultural and personal relevance.

This is of great importance, since both adults and children encounter “science” in many daily situations and are often expected to make related decisions. Since “[s]cience is never separate from the values of those who have a stake in it” (Mueller et al., 2012, p. 7), it is essential to discuss and teach science from the preceptive public encounters with science and consider their context and culture. Moreover, studies have consistently found that science knowledge is best acquired when it is personally or socially relevant (Shauli and Baram-Tsabari, 2019).

In this context, citizen science offers a platform where science is not disengaged from society and non-experts can be meaningfully engaged in a real scientific endeavor (Mueller et al., 2012). Through such hands-on experiences, participants can make connections between science and daily lives, providing an authentic scientific experience in a day-to-day context and contributing to life-long

learning processes. For example, [Koomen et al. \(2018\)](#) studied middle school students' engagement with a science fair project. One of their key findings was that participating in an authentic context provided at a summer citizen science program, served as a crucial starting point for their future participation in school science programs. Similar findings have been reported following student participation in a local air quality monitoring project with increased contextualized understanding of local air quality hazards and their relevance to students' personal exposure and health ([Golumbic et al., 2016](#)).

Awareness to science and the environment

Research over the past few decades has demonstrated that "raising levels of knowledge does not directly lead to a more positive attitude toward science" ([Edwards, 2004](#), p. 261). Thus, educational programs wishing to truly engage and impact student attitudes, need to adopt an approach that takes this into consideration and develop engagement and skills that go beyond knowledge. One of the processes that enable this is raising awareness of students to the interconnectedness of science to their lives, an important outcome, even if students do not foresee a science related career. Citizen science projects in schools are often designed to elicit exactly that—in depth engagement of students with science relevant to student's lives. Raising awareness and stewardship as learning outcomes of citizen science projects have been suggested widely in the literature ([Jordan et al., 2012](#)).

Public awareness of science and technology is a term coined by [Gilbert et al. \(1999\)](#) to be used in the context of science centers and museums. It is defined as "a set of attitudes, a predisposition toward science and technology, which are based on beliefs and feelings and which are manifest in a series of skills and behavioral intentions" (as cited in [Stockmayer and Gilbert \(2002\)](#), p. 836).

Raising awareness of student to one kind of relevant scientific issues: environment issues, is a long tradition within Environmental Education programs, although it has been acknowledged that this awareness is also dependent upon cultural and social characteristics and context ([Tal and Alkahr, 2010](#)). Two examples below demonstrate how contextualizing science contributes to raising awareness—both as a learning goal and as a platform for further action. First, in a study about community based climate change awareness in the Pacific, [McNamara \(2013\)](#) emphasizes that in a reality where knowledge is changing and expanding constantly, raising awareness needs to consider both who the participants of the project are, as well as their cultural and social background. In a different context, [Braschler et al. \(2010\)](#) who studied an ant monitoring biodiversity project in South Africa, found that students were empowered by realizing the local manifestation of larger global issues, as powerfully described:

Perhaps one of the most significant forms of awareness generated is the realization that many areas close to schools are in a poor condition. Scholars see for the first time the refuse and disturbance in their local environments and how this has an effect on biodiversity. In this way they rapidly come to understand that the global problems they read and hear about more generally, and in environmental lessons, are problems that have their basis at a local level and can, at least to some extent, be resolved at a local level. (p. 25)

In a US based study, [Ballard et al. \(2017\)](#) use the concept of environmental science agency (ESA) to identify ways youth develop and identify their own expertise in environmental science as well as a sense of agency with respect to environmental and community actions. They identified several structures and strategies that supported the development of agency and are unique to community and citizen science contexts, for example: using environmental science as a foundation for change.

Indeed, participation in citizen science projects can and has contributed to heightened awareness of science both on the individual and community levels ([Mueller et al., 2012](#)). By developing students awareness to the relevance of science we empower them to "engage in ongoing debates about local and global environmental and sustainability issues" ([Dillon et al., 2016](#), p. 450). This engagement, or rather call for action, will be discussed in the next section.

Civic action

Citizen science has been discussed in the literature as a gateway to civic action. For example, [Ottinger \(2016\)](#) describes a citizen science project monitoring air-toxics to demonstrate the potential of citizen science for social movements. However, within educational systems it is still not widely acceptable that there is a close link between science and citizenship ([Cooper, 2012](#); [Mueller et al., 2012](#)). Although citizen science projects in schools enable students to take an active role in the learning process, and science is often authentic and contextualized ([Koomen et al., 2018](#)), still promoting action that is beyond the learning process is often limited. [Phillips et al. \(2018\)](#) define an important outcome of citizen science in this sense—one that looks at "behavior and stewardship as measurable actions resulting from engagement in citizen science, but external to the protocol activities and the specific project-based skills of the citizen science project". (p. 10)

Stewardship, as suggested by them refers to action taken by citizen science participants, at both local and global scale, as they become more civically engaged. [Turrini et al. \(2018\)](#) reiterate that "citizen science can empower members of the public to get involved in policy-relevant citizen processes and to gain the knowledge that is needed to do so, thereby fostering deliberative decision-making in societally relevant processes" (p. 177).

Some examples of citizen science projects with a more civic participatory approach can be found in the field of crowdsourcing mapping projects, or volunteered geographic information (VGI) ([Haklay, 2013](#)). In these projects the common protocol is of "a collection of people working independently while responding to the needs of the local communities ..." ([Dalyot and Dalyot, 2018](#)).

In these types of projects participants tend to contribute data that is personally relevant, to solve (mapping and access) problems that are appropriate for them, and often the line between them and the scientist is left blurred (Haklay, 2013).

Despite the fact that more and more researchers see this great potential for citizen science projects to encourage and even facilitate social or civic action, research on the pathways to better civic engagement is quite mixed and often limited (Phillips et al., 2018).

Case studies

As demonstrated above, incorporating citizen science into education systems has the potential to lead to effective processes of student learning as well as promoting active citizenship and increasing student relatedness to science as a day-to-day encounter. Such outcomes have been successfully demonstrated in several citizen science projects and are further articulated in the next section, using examples from three distinct citizen science projects.

Enhancing learning and environmental stewardship through citizen science: the case of the radon home survey

Student's learning trajectories and development of social and environmental stewardship has been documented in the environmental citizen science project the Radon home survey. Initiated jointly by environmental and educational researchers, the project based in Israel engages students and communities in measuring and mapping Radon concentrations in buildings, and aims to raise awareness and create mitigation opportunities for lowering indoor Radon concentrations (Tsapalov et al., 2020).

Radon is a dangerous radioactive gas that, in high concentrations, can cause lung cancer. While early detection enables mitigation, many citizens are not aware of its presence, since Radon is both colorless and odorless. Using a simple and cheap technology for passive air absorption, the Radon home survey involved school students in measuring Radon in their homes. Following a series curriculum-based lessons, students learned about Radon gas and its dangers, engaged in class discussions, and analyzed the data received from their measurements.

Research examining the learning processes of students involved in the project and their contributing factors, revealed six common themes which were expressed at different stages of student participation (Columbic et al., 2022) (see Fig. 2). Together, the six themes illustrate a process students experienced throughout their participation: While initially unaware of Radon and its dangers, the students have recognized its relevance to their life, become curious of the topic, motivated to learn about it, and expressed interest in promoting awareness of Radon in their local environment.

The first two themes identified—relevance and identifying unawareness—demonstrate students initial exposure to the project topic and their recognition of its importance and lack of awareness in their close environment. This new understanding inspires the next two themes—interests and experience—as students spark new interest in learning about the subject, and do so in a nontraditional way by engaging in citizen science. This unique experience of learning through hands-on involvement, increases again the desire for learning and stimulates the ongoing interest. Collectively these processes lead to the fifth theme—learning—which is achieved by studying the subject in the classroom, performing Radon measurement and data analysis, and engaging in class

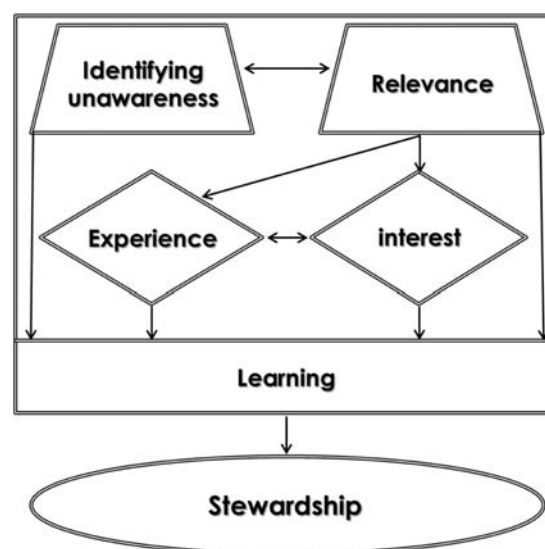


Fig. 2 Six themes which emerged from the data analysis, illustrating the typical learning process in the radon survey. Modified from Columbic et al. (2022).

discussions. Students build their knowledge base about Radon gas, develop an understanding of the importance of its measurement and learn to think critically about socio-scientific issues.

Lastly, after developing an understanding and awareness to Radon, students developed an organic interest in promoting awareness to Radon in their local communities and in becoming agents of change in their close environment. For example, students from one school initiated a community event in order to present the results of their research and raise awareness to Radon in the broader community. Some students suggested that they have a personal responsibility toward their family and community in addressing this matter. Finally, many of the students raised questions about social norms and democratic processes, investigating the role of governments and decision makers in maintaining public health and well-being.

The findings of this study emphasize the potential of citizen science in raising awareness to environmental hazards, empowering students and publics, and promoting their environmental stewardship and responsibility. It eludes to the potential of citizen science in supporting learning outcomes that are beyond formal curriculum requirements and develop student's accountability, forming the next generation of responsible, caring citizens.

Learning about the complexity of science through citizen science: the case of breaking good

Student's learning about the complexity of science and its multifaceted nature, is a key goal of the open source drug discovery project Breaking Good. The project, operating in Australia, the UK and the US, engages students and citizens in synthesis processes of new and existing molecules for human health, and in the investigation of access to important medicines for people around the world (Motion (nee Williamson), 2017).

Such essential medicines have been defined by the World Health Organization as medicines that satisfy the priority health care needs of the population, and should be available at all times, in adequate amounts and at affordable prices. However, there are several barriers which prevent the accessibility of such medicines to people around the world. These include low market incentive, competition and profit considerations, supply shortages, in addition to a range of societal and political challenges.

This complex system provides the Breaking Good project, good grounds for student learning about science and its broader societal context. It inspires a deep conversation about science and society, challenging students to think about essential medicines in broad public health context and consider its relevance to their own lives and communities. Building on this understanding, students engage in drug discovery by learning organic chemistry and following pharmaceutical procedures for molecules synthesis.

Over the past years, high school and undergraduate university students have engaged in Breaking Good, through hands on synthesis of molecules, as part of their teaching laboratories. Rather than throwing away their samples at the end of class, students worked on the synthesis of novel antimalarials drugs, sending their samples for activity screening against malaria. This process resulted in a number of novel molecules which showed promising activity at killing the malaria parasite, contributing to the Open Source Malaria (OSM) consortium which works to speed up and reduce costs of malaria drug identification (Motion (nee Williamson) et al., 2016).

In another case, high school students from Australia reproduced an existing medicine for treating toxoplasmosis (a parasitic infection), that has been price-hiked earlier that year, making it unaffordable for many people around the world. This project is a proof of concept and exemplifies how much can be done to promote science and public health by young and inexperienced scientists, given the right facilities, instruction and personal incentive.

Students have also discussed global production of medicines, equity and resource distribution in medicine and health care and learned about the social, political and financial variables which affect access to some medicine. Such learning opportunities provide a holistic understanding of the complexity of science and the many factor which may influence its development and acceptance.

The classroom as a platform for community engagement: the case of Food Science Class

Citizen science, as a collaborative social and scientific effort, has the power to bring together scientists and communities to work together on relevant issues. This is an especially crucial point when it comes to food science and food literacy, however "citizen science" is rarely discussed in this context (Ryan et al., 2018). The FoodScienceClass project brings together citizen science principles to demonstrate the multidisciplinary nature of [food] science.

This is a collaborative multi-disciplinary project-based learning program that is centered around student led research into the origins of their commonly consumed industrialized food items (e.g., breakfast cereals, yogurts, ketchup, etc.). The project targets young teenagers (ages 13–16 in middle school) at a stage when they start making independent food choices. This is also an age where interest in science is slowly decreasing (King and Henderson, 2018), hence one of the goals of this project is to demonstrate the relevance of science to every day decision and issues—especially on a societal level.

In more detail—the project seeks to engage teenagers with food science processes, social dilemmas, and policy aspects, but at the same time teach them science communication skills. In the context of civic engagement, the component of science communication skills is extremely important as they enable teenagers (now and later in life) to engage with their communities on relevant scientific issues. Thus, students are perceived and addressed as empowered citizens, capable of communicating societally significant scientific issues. Since the programs runs in parallel both in Israel and in Poland, it presents additional opportunities to discuss global vs. local issues within a formal class setting. This poses a rare opportunity in education systems everywhere.

FoodScienceClass project introduces to students the inter and multidisciplinary aspects of food science (and science in general) and strives to make both the subject matter and skills gained relevant to their lives. Throughout the project a strong emphasis was put on the question: “how is this relevant to my daily life?” and students were encouraged to critically consider and discuss this question. For example, when conducting a lab on yeast fermentation (See [Box 1](#))—a chemistry related activity—students discussed industrial uses of yeast and how their conclusions from the experiment are related to cooking at home.

Box 1 Excerpt from info-sheet students receive as part of their yeast experiments

The importance of yeasts in the food industry

Yeasts are vital microorganisms in the food industry. The Bakery yeast (which is also known as the Brewers' Yeast), *Saccharomyces cerevisiae*, and its different species is the most common in the making of food products. Without yeasts, the dough will not rise (sour) and baking goods that we all know and love, like bread, cakes etc. will not exist. When the dough is in its rising stage, which is mostly an anerobic stage, the yeasts ferment the sugar in the dough (which is the flour itself, when no sugar is added) and emit ethanol and CO₂, which is carbon dioxide. The gas emitted fills the dough with air and makes it rise, and the ethanol, which evaporates in high temperatures, gives the dough its special flavor of baking goods.

The use of yeast in baking and the making of alcoholic beverages is believed to be earlier than the invention of script. An ancient clay tablet from Babylon dated 8000 years ago and Egyptian hieroglyphs from ancient Egypt dated 5000 years ago, indicate the existence of these processes, but without the understanding of how they occur. It is believed that dough fermentation began with a coincidence contamination of flour by the bakery yeast which fermented it, giving it its unique qualities.

Although not a traditional and scientist led citizen science project, this project is conceptualized around “problem-based learning” experiences that are known as a useful and important tool in supporting students in gaining knowledge and skills, especially in the sciences ([Willard and Duffrin, 2003](#)). The problem being industrial food consumption in their communities and its effects. They start by understanding what industrial and processed food is, continue with developing a survey to assess the consumption and habits in their families and community at large, and then distribute and analyze the survey. While discussing the survey findings they come up with ideas for further activities and questions, compare results with classes in another country and learn more about the most popular items in their community. Finally, they receive expert science communication training and are then asked to prepare their own communication or outreach activity to disseminate what they have learned in the project.

Since this project is not based on lectures but rather on self-learning and group work, the transition of education systems to remote online learning during the COVID-19 pandemic has not halted the implementation of the project. The fact that project staff were flexible and worked with schools to accommodate needs and logistics demonstrates an important design principle that contributes to success of such projects—working with stakeholder to adjust and fit the program to their need is instrumental.

Integration

The three case-studies presented above, represent a range of citizen science-based projects in various scientific fields, implemented within school environments. While each aims to enhance student knowledge in its respected field, they all have additional learning outcomes they plan and strive for, which include improving student skills, critical thinking and stewardship. Together, these case-studies exemplify how student participation in citizen science can go beyond content knowledge generation to include learning about Societal Aspects of science, raise Awareness to science and the environment and engage in Action on the local and global level (SAAA) (See [Fig. 3](#)).

The Radon Home survey discusses science in the context of daily life, encouraging students to recognize the relevance of the topic studied to their own environments (Societal Aspects). By doing so, it raises students' awareness and inspires their continued interest in learning about the topic and conducting relevant measurements (Awareness). Lastly, and interrelated to the two processes above, students are inspired to take action and positively impact their communities and families (Action).

Breaking Good has a strong emphasis on public health and the social, financial and cultural influences on health care, revealing science research and practice are very much embedded in society (Societal Aspects). It confronts students with the complexity of health systems and science more generally, broadening their perceptions and perspective of “how science works” (Awareness). Finally, it calls for students to assist in the creation of new, open sourced medications and become advocates of such collaborative science (Action).

FoodScienceClass connects students to science through one of the most common activities we all do—eat. It exemplified how interconnected this experience is with food science processes, social dilemmas, and policy aspects (Societal Aspects). By raising students' attention to different ways of food production, marketing and advertising, the project helps student make informed

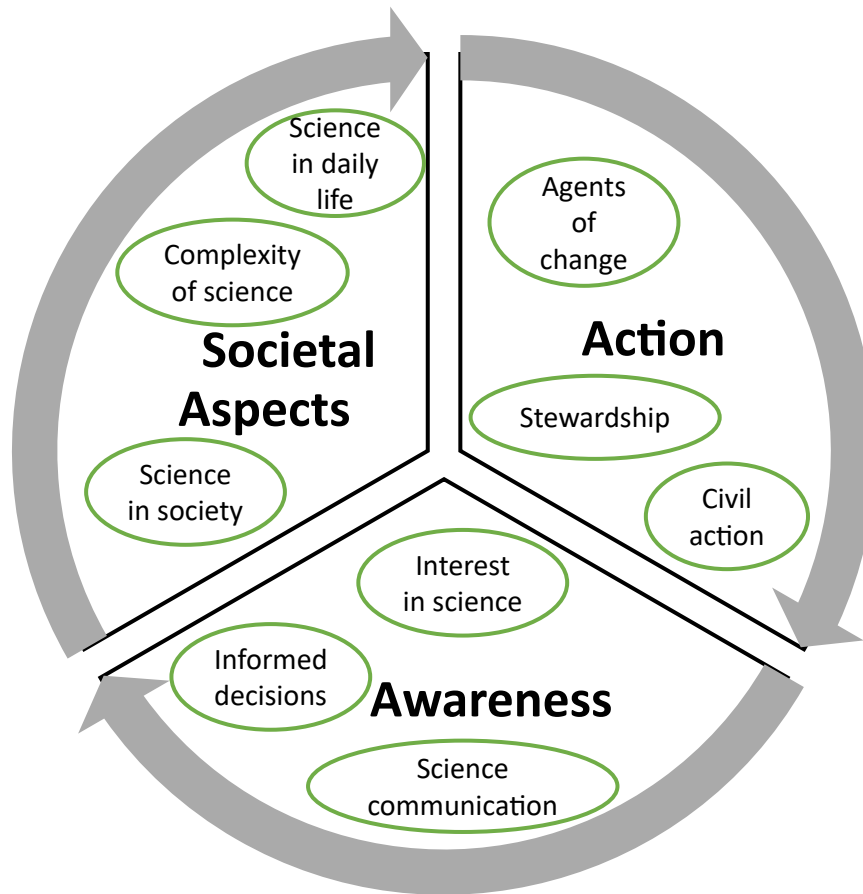


Fig. 3 Framework for potential of citizen science for engaging students.

decisions and think critically (Awareness). The project also teaches important science communication skills, encouraging students to share this understanding within the community to create change (Action).

Summary

This article discussed incorporating citizen science into education systems for encouraging more effective processes of learning science in the context of real life experiences and broadening students learning outcomes. Using three case studies of distinct nature, we demonstrate the potential for learning outcome in three important aspects of science education; Societal Aspects, Awareness, and Action (SAAA), which go beyond content knowledge generation. We further describe the pathways for such learning to occur, rooted in the continued effort of educators and project facilitators to design for these causes.

While broad in scope and expressed in a variety of contexts, SAAA demonstrates that “by providing both detail of specific science knowledge and also examples of contexts in which such knowledge is important” (Ryder, 2001, p. 2) comprehensive and broad learning can occur. Eluding to the vast potential of citizen science for effective and meaningful learning processes.

However, to fulfill this potential of citizen science several prerequisites must occur. First and foremost, projects need to plan for the educational goals of SAAA. Without such intentional design and consideration of learning goals and processes, it is unlikely to achieve many of these goals (National Academies of Sciences, Engineering and Medicine, 2018). Second, projects must make explicit to participants the scientific, social and cultural issues that they are encountering. This goes beyond explaining scientific processes of data collection to an in-depth justification of the project rationale, values and end goals (Brossard et al., 2005). Lastly, students’ interaction with scientists as well as with the entire scientific protocol needs to be considered in order to allow students to fully benefit from their participation and connect on a personal level. This is increasingly important as many citizen science projects are used as a pathway for data collection with little or no direct interaction with the researcher behind the project or with the complete scientific process (Mueller et al., 2012). Ultimately, in order for citizen science to become meaningful in science education, it has to offer a variety of avenues in which open, collaborative and complex initiatives can emerge. We call for further advancement of citizen science practices, which design for, with and toward the progression of learning on all its complexities.

Acknowledgment

The ideas for this article came from part of our work on a symposium for NARST annual conference 2019. We are thus grateful for the contributions of—Ayelet Baram-Tsabari, Tali Tal, Bruce Lewenstein, Heidi Ballard, Caren Cooper, Arjen Wals and Anne Bowser. We would also like to thank Yael Barel Ben-David for providing useful comments on this article. KD would like to acknowledge the generous support of EIT FOOD in the development and implementation of the *The classroom as a platform for community engagement with food production and science* project.

References

- Ballard, H.L., Dixon, C.G.H., Harris, E.M., 2017. Youth-focused citizen science: examining the role of environmental science learning and agency for conservation. *Biol. Conserv.* 208, 65–75. <https://doi.org/10.1016/j.biocon.2016.05.024>.
- Bauer, H.H., 1994. *Scientific Literacy and the Myth of the Scientific Method*. University of Illinois Press, Chicago.
- Bonney, R., Phillips, T., Enck, J., Shirk, J., Trautmann, N., 2014. Citizen science: next steps for citizen science. *Science* 343 (6178), 1436–1437.
- Braschler, B., Mahood, K., Karenyi, N., Gaston, K.J., Chown, S.L., 2010. Realizing a synergy between research and education: how participation in ant monitoring helps raise biodiversity awareness in a resource-poor country. *J. Insect Conserv.* 14 (1), 19–30.
- Brossard, D., Lewenstein, B., Bonney, R., 2005. Scientific knowledge and attitude change: the impact of a citizen science project. *Int. J. Sci. Educ.* 27 (9), 1099–1121.
- Claussen, S., Osborne, J., 2013. Bourdieu's notion of cultural capital and its implications for the science curriculum. *Sci. Educ.* 97 (1), 58–79.
- Cooper, C.B., 2012. Links and distinctions among citizenship, science, and citizen science. A response to “the future of citizen science”. *Democr. Educ.* 20 (2), 13. Available at: <http://democracyeducationjournal.org/home/vol20/iss1/2/>. (Accessed 3 August 2020).
- Crowell, A., Schunn, C., 2016. Unpacking the relationship between science education and applied scientific literacy. *Res. Sci. Educ.* 46, 129–140.
- Dalyot, K., Dalyot, S., 2018. Towards the use of crowdsourced GIS data to georeference child well-being globally. *Soc. Indic. Res.* 139 (1), 185–204.
- Dillon, J., Stevenson, R.B., Wals, A.E.J., 2016. Introduction to the special section: moving from citizen to civic science to address wicked conservation problems. *Conserv. Biol.* 30 (3), 450–455.
- Edwards, C., 2004. Evaluating European public awareness of science initiatives: a review of the literature. *Sci. Commun.* 25 (3), 260–271.
- Feinstein, N., 2011. Salvaging science literacy. *Sci. Educ.* 95, 168–185.
- Gilbert, J.K., Stockmayer, S., Garnett, R., 1999. Mental modelling in science and technology centres: what are visitors really doing? In: Stockmayer, S., Hardy, T. (Eds.), *Learning Science in Informal Contexts*. Questacon, Canberra, pp. 7–15. Australia.
- Golumbic, Y.N., Baram-Tsabari, A., Fishbain, B., 2016. Increased knowledge and scientific thinking following participation of school students in air-quality research. In: *Proceedings of Indoor Air 2016*. Ghent Belgium. Available at: <https://www.isiaq.org/docs/Papers/Paper832.pdf>. (Accessed 29 October 2020).
- Golumbic, Y.N., Peri, A., et al., 2022. The radon home survey: a citizen science project for involving the public in authentic research combining science and society. *Israeli Sociol. [HEBREW]*.
- Haklay, M., 2013. Citizen science and volunteered geographic information: overview and typology of participation. In: Sui, D.Z., Elwood, S., Goodchild, M.F. (Eds.), *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. Springer, Netherlands, pp. 105–122.
- Harlin, J., Kloetzer, L., Patton, D., Leonhard, C., Leysin American School High School Students, 2018. Turning students into citizen scientists. In: Hecker, S., Haklay, M., Bowser, A., et al. (Eds.), *Citizen Science: Innovation in Open Science, Society and Policy*. UCL Press, pp. 410–428.
- Jordan, R.C., Ballard, H.L., Phillips, T.B., 2012. Key issues and new approaches for evaluating citizen-science learning outcomes. *Front. Ecol. Environ.* 10, 307–309.
- King, D., Henderson, S., 2018. Context-based learning in the middle years: achieving resonance between the real-world field and environmental science concepts. *Int. J. Sci. Educ.* 40 (10), 1221–1238. <https://doi.org/10.1080/09500693.2018.1470352>.
- Kinslow, A.T., Sadler, T.D., 2018. Making science relevant: using socio-scientific issues to foster critical thinking. *Sci. Teach.* 86 (1), 40–45.
- Kloetzer, L., Lorke, J., Roche, J., et al., 2021. Learning in citizen science. In: Vohland, K., Land, A., Ceccaroni, L., et al. (Eds.), *The Science of Citizen Science*. Springer International Publishing.
- Koomen, M.H., Rodriguez, E., Hoffman, A., Petersen, C., Oberhauser, K., 2018. Authentic science with citizen science and student-driven science fair projects. *Sci. Educ.* 102 (3), 593–644.
- Mcnamara, K.E., 2013. Raising awareness about climate change in Pacific communities. *Environ. Educ. Res.* 19 (6), 864–871.
- Motion (nee Williamson), A.E., Ylloja, P.M., Robertson, M.N., et al., 2016. Open source drug discovery: highly potent antimalarial compounds derived from the tres cantos arylpyrroles. *ACS Cent. Sci.* 2 (10), 687–701.
- Motion (nee Williamson), A.E., 2017. Open source drug discovery: global solutions to global problems. *Aust. Q.* 88 (2), 3–8.
- Mueller, M., Tippins, D., Bryan, L., 2012. The future of citizen science. *Democr. Educ.* 20 (1), 1–12. Available at: <http://democracyeducationjournal.org/home/vol20/iss1/6http://democracyeducationjournal.org/home/vol20/iss1/12http://democracyeducationjournal.org/home/vol20/iss1/13http://democracyeducationjournal.org/home/vol20/iss2/14>. (Accessed 3 August 2020).
- National Academies of Sciences, Engineering and Medicine, 2016. *Science Literacy: Concepts, Contexts, and Consequences*. The National Academies Press, Washington, DC.
- National Academies of Sciences, Engineering and Medicine, 2018. *Learning Through Citizen Science: Enhancing Opportunities by Design*. The National Academies Press, Washington, DC.
- Ottinger, G., 2016. Social movement-based citizen science. In: Cavalier, D., Kennedy, E.B. (Eds.), *The Rightful Place of Science: Citizen Science*. Consortium for Science, Policy & Outcomes, Tempe, AZ, pp. 89–103.
- Phillips, T., Ferguson, M., Minarchek, M., Porticella, N., Bonney, R., 2014. User's Guide for Evaluating Learning Outcomes in Citizen Science. Cornell Lab of Ornithology, Ithaca, NY.
- Phillips, T., Porticella, N., Constat, M., Bonney, R., 2018. A framework for articulating and measuring individual learning outcomes from participation in citizen science. *Citiz. Sci. Theory Pract.* 3 (2), 1–19.
- Roche, J., Bell, L., Galvão, C., Golumbic, Y.N., Kloetzer, L., Knoben, N., Laakso, M., Lorke, J., Mannion, G., Massetti, L., Mauchline, A., Pata, K., Ruck, A., Taraba, P., Winter, S., 2020. Citizen science, education, and learning: challenges and opportunities. *Front. sociol.* 5, 110. <https://doi.org/10.3389/fsoc.2020.613814>.
- Ryan, S.F., Adamson, N.L., Aktipis, A., et al., 2018. The role of citizen science in addressing grand challenges in food and agriculture research. *Proc. Biol. Sci.* 285 (1891), 20181977.
- Ryder, J., 2001. Identifying science understanding for functional scientific literacy. *Stud. Sci. Educ.* 36 (1), 1–44.
- Shah, H.R., Martinez, L.R., 2016. Current approaches in implementing citizen science in the classroom. *J. Microbiol. Biol. Educ.* 17 (1), 17–22.
- Shauli, S., Baram-Tsabari, A., 2019. The usefulness of science knowledge for parents of hearing-impaired children. *Publ. Understand. Sci.* 28 (1), 19–37.
- Shirk, J.L., Ballard, H.L., Wilderman, C., et al., 2012. Public participation in scientific research: a framework for deliberate design. *Ecol. Soc.* 17 (2), 29.
- Stockmayer, S., Gilbert, J.K., 2002. New experiences and old knowledge: towards a model for the personal awareness of science and technology. *Int. J. Sci. Educ.* 24 (8), 835–858.
- Storksdieck, M., Shirk, J.L., Cappadonna, J., et al., 2016. Associations for citizen science: regional knowledge, global collaboration. *Citiz. Sci. Theory Pract.* 1 (2), 10.
- Tal, T., Alkaher, I., 2010. Collaborative environmental projects in a multicultural society: working from within separate or mutual landscapes? *Cult. Stud. Sci. Educ.* 5 (2), 325–349.

- Tal, T., Dierking, L.D., 2014. Editorial: learning science in everyday life. *J. Res. Sci. Teach.* 51 (3), 251–259.
- Tsapalov, A., Kovler, K., Shpak, M., et al., 2020. Involving schoolchildren in radon surveys by means of the "RadonTest" online system. *J. Environ. Radioact.* 217, 106215.
- Turrini, T., Dörler, D., Richter, A., Heigl, F., Bonn, A., 2018. The threefold potential of environmental citizen science: generating knowledge, creating learning opportunities and enabling civic participation. *Biol. Conserv.* 225, 176–186.
- Vohland, K., Land, A., Ceccaroni, L., et al. (Eds.), 2021. *The Science of Citizen Science*. Springer International Publishing.
- Wals, A.E.J., Peters, M.A., 2017. Flowers of resistance: citizen science, ecological democracy and the transgressive education paradigm. In: König, A., Ravetz, J. (Eds.), *Sustainability Science: Key Issues*. Earthscan/Routledge, London.
- Wals, A.E.J., Brody, M., Dillon, J., Stevenson, R.B., 2014. Convergence between science and environmental education. *Science* 344, 583–584.
- Willard, K., Duffrin, M.W., 2003. Utilizing project-based learning and competition to develop student skills and interest in producing quality food items. *J. Food Sci. Educ.* 2 (4), 69–73. <https://doi.org/10.1111/j.1541-4329.2003.tb00031.x>.

Further reading

- Bonney, R., Ballard, H., Jordan, R., et al., 2009. Public Participation in Scientific Research: Defining the Field and Assessing its Potential for Informal Science Education. A CAISE Inquiry Group Report, pp. 1–58.
- Dalyot, K., Golumbic, Y.N., Ballard, H., Baram-Tsabari, A., Bowser, A., Cooper, C., Lewenstein, B., Tal, T., Wals, A., 2019. Citizen science in STEM education: linking society, scientists and education systems. In: Symposium at the 92nd NARST Annual International Conference, Baltimore, MD, USA (March 31–April 3).
- Golumbic, Y.N., Fishbain, B., Baram-Tsabari, A., 2020. Engagement styles in an environmental citizen science project. *J. Sci. Commun.* 19 (6), A03.
- Sagy, O., Golumbic, Y.N., Abramsky, H., et al., 2019. Citizen science: an opportunity for learning in the networked society. In: Kali, Y., Baram-Tsabari, A., Schejter, A.M. (Eds.), *Learning in a Networked Society (LINKS)*. Springer, Cham, pp. 97–115.

Relevant websites

<https://ayeletlab.net.technion.ac.il/the-classroom-as-a-platform-for-community-engagement-with-food-production-and-science/>.

<https://www.birds.cornell.edu/k12>.

<https://www.breakinggoodproject.com>.

<https://citizenscience.org.au>.

<https://www.citizenscience.org>.

<https://cs-eu.net>.

<https://ecsa.citizen-science.net>.

<https://medium.com/citizenscience-asia>.

<https://nbri.net.technion.ac.il/en/radon-project/>.

https://www.sciencelearn.org.nz/citizen_science.

<https://scistarter.org/>.

<https://www.tcscs.center/home-en>.

<https://www.zooniverse.org/>.

Community engagement in STEM education

Wing-mui Winnie So^a and Chong-jee Guo^b, ^aThe Education University of Hong Kong, Tai Po, Hong Kong, P.R. China; and
^bNational Changhua University of Education, Taichung, Taiwan

© 2023 Elsevier Ltd. All rights reserved.

Introduction	234
Background	234
STEM learning ecosystem	235
The role of individuals in STEM learning	235
The role of family and peers in STEM learning	235
The role of school in STEM learning	236
The role of society in STEM learning	236
Students' STEM aspirations	237
Principles to support STEM education across various settings	237
Various practices with community engagement in STEM education	238
STEM service learning	238
STEM fair/competition	238
STEM information website	238
STEM network website	238
Science museum visit	239
STEM outreach activities	239
Family engagement in STEM	239
Recommendations	240
Establishing STEM professional-school collaboration	240
Developing STEM communities of practice (STEM CoPs)	240
Highlighting STEM professional contribution to the society	240
Concluding remarks - a look into the future	241
Acknowledgments	242
References	242
Relevant Websites	243

Introduction

The advancing of STEM education goes beyond reform of the curriculum, teaching and learning in the school setting. There have been calls for more community engagement in STEM education with contributions from families and society. This article discusses the various conceptions and practices related to community engagement in STEM education, supported with research evidence from the literature. This is to engage the community through more public awareness, and to involve the STEM industries/professionals in achieving corporate social responsibilities, to bring real world STEM to STEM education, and to empower teacher professional development to provide quality STEM learning, so as to inspire the next generation with STEM aspirations in their study and career choice.

This article will provide some background information of the STEM education initiatives in the school curriculum and the STEM learning opportunities outside school. It is followed by a discussion of the STEM learning ecosystem with the different individual, family, school and society factors influencing students' STEM learning and STEM aspirations. The various studies on STEM aspirations provide insights into the importance of learning from the community. Some successful cases of community engagement in STEM education will be outlined. The conceptual frameworks of "Establishing STEM Professional-School Collaboration", "Developing STEM Communities of Practice" and "Highlighting STEM professional contribution to the society" are suggested to engage the community more effectively in STEM education.

It is hoped that the suggested frameworks can provide some guidelines for educators and community stakeholders in designing and implementing STEM education. Likewise, it is also timely to have more evidence from researchers with studies on student learning outcomes during the various community engagement opportunities. The evidence will be insightful for educators and community stakeholders and will add knowledge to the related field of study.

Background

In order to prepare our students to engage in innovation and technology, which is one of the main driving forces for the nation's economic growth and productivity, the call for STEM education has become one of the global key educational initiatives. This

development has given rise to active discussions on the forms, strategies and effectiveness of STEM education, not only in the education field, but also among the various stakeholders in the community involved in STEM education.

The substantial energy to improve STEM education is mainly focused on the reform of the existing disciplinary curriculum, teaching and learning, and assessment. The focus on teacher professional development is becoming more intense in its call for a more competent teaching force with relevant knowledge and skills for providing STEM education to students. However, school curricula usually reveal what teachers consider as essential rather than the knowledge and skills that are actually required in the community. Moreover, what should be included in school STEM education also relies on teachers' understanding of what STEM and STEM education are. Unfortunately, STEM is also new to most teachers as they were mostly educated and trained under the disciplinary learning approach.

It is also believed that students do not acquire STEM knowledge or develop STEM skills in schools only, but on many occasions and in different ways after school. For instance, there are STEM learning opportunities during afterschool programs/workshops; when visiting museums and exhibitions; reading (online information or books); watching TV or educational programs; during field trips and outings, or even at home while interacting with peers and family members. These various STEM learning opportunities may interact with and reinforce students' STEM learning in school. Hence, linking education and community is necessary for community economic growth, the advancing quality of life, and promoting citizen's wellbeing.

Gaps between STEM education and required workplace skills have been identified in industry, academia and the government (Jang, 2016). There is research evidence that, although STEM professionals are not educators, their roles in STEM education are indispensable as they have first-hand experience of what knowledge, skills, attitudes and work activities, which can be understood as STEM competence, are performed in their daily tasks. Individual learning is surrounded by layers of influences that have an effect on learning (Olson and Labov, 2014). All of the above diverse influences within and beyond the school setting where STEM learning takes place is conceptualized as the STEM Learning Ecosystem.

STEM learning ecosystem

As STEM learning ecosystems are emerging (Olson and Labov, 2014), there have been discussions of such ecosystems (Traphagen and Traill, 2014), and the ecological framework of factors are influencing students' participation, achievement and progression in STEM studies (Chavatzia, 2017, p. 40).

As suggested in Chavatzia's report for UNESCO (2017), there are multiple and overlapping factors which influence the participation, achievement and progression in STEM studies and careers, all of which interact in complex ways. An ecological framework which encompasses factors at the individual, family and peer, school and societal levels is suggested to explain these factors and understand the interrelations among them.

Olson and Labov (2014) described the concentric circles of influences surrounding the individual learner. Closest to the learner are family, friends, role models, caregivers, mentors, and peers. Beyond these influences, a set of sectors seeks to teach or otherwise provide experiences to learners, including schools, museums, media, libraries, and afterschool clubs. Finally, infrastructural elements such as policies, culture, communities, and values create the rules of the game for sectors as well as individuals.

Similarly, the STEM learning ecosystem suggested by Traphagen and Traill (2014) comprises schools, community settings such as after-school and summer programs, science centers and museums, and informal experiences at home and in a variety of environments that together constitute a rich collection of learning opportunities for students.

The reasons to pursue STEM study and careers may vary among students, ranging from individual reasons such as achieving personal goals, to a greater societal vision such as contributing to the world's advancement (So and Chiu, 2020). Following are discussions of the roles of individuals, family and peers, schools and society in students' STEM learning.

The role of individuals in STEM learning

The role of individuals consists of biological factors that may influence individuals' abilities, skills, and behavior such as brain structure and function, hormones, genetics, and cognitive traits including spatial and linguistic skills. It also considers psychological factors, including self-efficacy, interest and motivation (Chavatzia, 2017). Yet, individual STEM identities have been found to be impacted by both internal and external factors (Hansen, 2018).

Research results have shown that educators, peers, and family influenced youth STEM interest, which in turn predict their STEM self-efficacy and career outcome expectancy. STEM career orientation is fostered by youth-expected outcomes for such careers (Nugent et al., 2015). It was also found in Han et al., (2015) study of 3 years of STEM PBL instruction that low performing students showed statistically significantly higher growth rates on mathematics scores than high and middle performing students. Moreover, high levels of self-efficacy in secondary students have been positively correlated to students' pursuit of post-secondary education (Chachashvili-Bolotin et al., 2016).

The role of family and peers in STEM learning

Family and peer factors comprise parental beliefs and expectations, parental education and socioeconomic status, and other household factors, as well as peer influences (Chavatzia, 2017). Han et al., (2015) noticed that students' ethnicity and economic status

were good predictors of academic achievement. Archer et al. (2013) also found that family exert a considerable influence on students' aspiration.

In Lloyd et al. (2018)'s study, students who expressed an interest in STEM were typically high achieving, and just over 90% of their parents had aspirations for them to attend university—although this was proportionally higher for sons than for daughters. Even when parents created a supportive environment, there was little evidence indicating that girls were encouraged to pursue STEM. The research also highlighted the complexity and importance of parental influences on student aspirations.

For the influence of peers, the result of Micari et al. (2016) study indicated that academic-preparedness diversity was mostly associated with positive learning outcomes. They found that academically less-prepared students got greater benefits, and those less-prepared students performed better when they were not alone in a group of highly prepared students.

The role of school in STEM learning

The factors within the school learning environment include teachers' profile, experience, beliefs and expectations, curricula, learning materials and resources, teaching strategies and student teacher interactions, assessment practices and the overall school environment (Chavatzia, 2017). Among the offering of definitions and conceptions of STEM, putting STEM education into practice requires teachers to have a different knowledge base than that of subject teachers, to enable them to play an important role in providing students with the integrated learning opportunities with STEM. Participation in advanced STEM coursework at the secondary level increases female students' likelihood of pursuing a STEM pathway in postsecondary education (Hinojosa et al., 2016).

The role of society in STEM learning

STEM permeates the modern world. The jobs people do, the foods they eat, the vehicles in which they travel, the information they receive, the medicines they take, and many other facets of modern life are constantly changing as STEM knowledge steadily accumulates (Olson and Labov, 2014). Social and cultural norms are related to gender equality, and gender stereotypes in the media (Chavatzia, 2017). Lee et al. (2018) identified students' social perspectives for their pursuit of a STEM career, to help people, to interact with others with 21st century ideas, and to impact the world.

Fig. 1 near here shows the recommendation by So and Chiu's (2020) in their policy research at different levels of individual, family, school and society of the STEM learning ecosystem. For the individual level, there is need to make available learning opportunities and narrow the gender gap. For the family level, it is recommended to enhance parent education on students' study and career choices. For the school level, it is essential to articulate STEM learning progression across key stages and integrate STEM career understanding into school curriculum. For the society level, it is indispensable to improve quality of STEM education and highlight STEM contribution to the contemporary society.

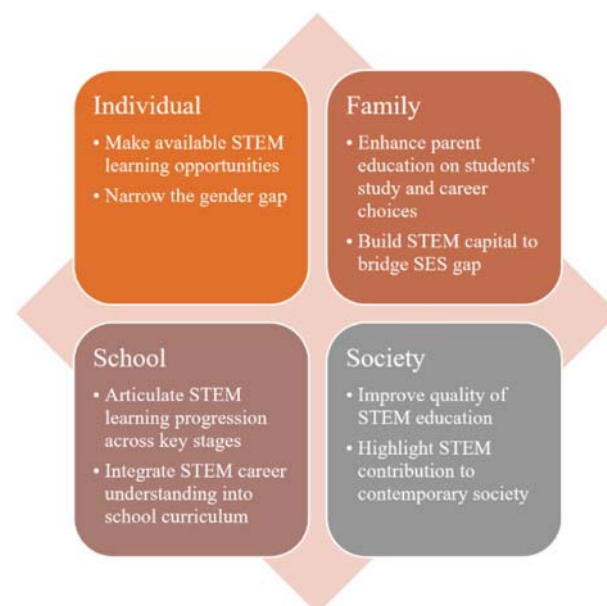


Fig. 1 The recommendation at the various levels of the STEM Learning Ecosystem. Source: So and Chiu, 2020, p. 63.

Students' STEM aspirations

A learning ecosystem drives the unique contributions of each different setting to offer STEM learning for children (Traphagen and Traill, 2014). Unfortunately, these various settings are in different stages of evolution, and people from one setting often do not know the people or organizations from other sectors. How to motivate and engage the people from different settings and work with others as part of a unified endeavor is not an easy task; it needs some mutual understanding of the roles of each sector in the STEM learning ecosystem. So and Chiu (2020) argued that understanding students' intention to pursue a STEM career is required for strategy formulation to nurture young people with the necessary skills, education and experience to engage as STEM professionals.

Archer et al. (2014) explained why aspirations are an interesting and important focus for research, and young people's aspirations remain an enduring focus of education policy interest and concern. Their research analysis indicated patterns of association between particular sources of influence and types of career aspiration. For example, aspirations for sports and the arts were overwhelmingly related to children's out-of-school activities. Middle-class children were more likely to cite family capital in relation to professional careers in the fields of medicine and teaching. Working-class students were much less likely to aspire to follow in a family member's footsteps or to develop their aspirations through an out-of-school hobby/activity. The researcher criticized that this is an unfair social enterprise that only the privileged few are likely to attain.

However, the reasons why students not choosing STEM careers may due to common misunderstandings that STEM is a difficult, uncreative, passionless and socially isolated enterprise (Masnick et al., 2010). Students' lack of interest or aspirations in STEM careers is a concern worldwide (The White House, 2018).

Lee et al. (2018) indicated that for those who selected STEM careers for their future employment, the perspectives they held that motivated their STEM career aspirations were mainly from personal and societal perspectives. Students with a personal context usually have positive emotions, personal development, and developing tools for the job. Students whose perspectives best reflected a societal context were helping people, interacting with others, and impacting the world.

On top of Mau and Li's (2018) research result that race, gender, socioeconomic status, mathematics interest, and science self-efficacy were the most important predictors of STEM career aspirations, So and Chiu (2020) suggested building STEM capital to bridge the socio-economic status gap as more informal STEM activities are one of the significant predictors of students' STEM aspiration.

Principles to support STEM education across various settings

Olson and Labov (2014) indicated that the STEM learning system is remarkably diverse, and all parts of the system have something to contribute to STEM learning. The challenge is to mesh these contributions synergistically while adapting models that have worked well in one place to the culture, governance, and idiosyncrasies in other settings. In other words, the goal is to create an actual integrated system for teaching and learning wherever learning occurs, both in and outside of school, rather than a set of uncoordinated activities.

Penuel et al. (2014) reviewed the literature on cross-setting learning, and reported on several design principles they identified which are useful for STEM learning: (1) Draw on values and practices from the various settings and identify practices in one setting that can be used as a resource to support learning in others. (2) Structure partnerships to embrace the goals of all stakeholders. (3) Engage participants in building stories, imaginative worlds, and artifacts that span contexts and that facilitate meaning-making across contexts. (4) Help students identify learning opportunities to contribute to authentic endeavors. (5) Use brokering purposefully, for example, preparing educators and parents to play roles as brokers to facilitate movement across settings.

The guide to STEM Community Engagement provided schools, districts, and local community organizations with self-guided tools to engage a broad community of stakeholders in sustainable education innovation. It suggested five design principles to drive the work, namely: (1) Equitable; (2) Scalable & sustainable; (3) Innovative, (4) STEM-focused; and (5) collaborative.

These principles aim to make students the center of our communal efforts across settings. Although Tytler et al.'s (2018) discussion of the community partnerships is focusing on sustainability education, the benefits they mentioned are useful and can be translated to STEM education, including curriculum links, community network formation, volunteering in STEM work, students' career and attitude development.

The above three principles show there are different ways to support STEM learning. Based on the work by Morrison and Fisher (2019) on measurement in STEM learning ecosystems that build on expectations have already been articulated concerning progressions in the quality of partnerships, collaborations, and outcomes. We propose a summary of the principles to support STEM learning across sectors as follows:

1. Connect parents and/or teachers with links to the curriculum and daily life for a sustainable effect
2. Address contributions across collaborators to achieve one's goals and to share best practices
3. Facilitate meaningful learning opportunities with innovative and authentic tasks that are transferrable across contexts
4. Focus on students' attitudes and career development as outcomes

Various practices with community engagement in STEM education

There are different practices of community engagement in STEM education based on the various principles. Most of them are also practices used in other areas but with special focuses on STEM learning. They are service learning, fair/competition, information websites, network websites, science museums, outreach programs and family engagement in STEM education.

STEM service learning

Preece (2017) claimed that service is a largely philanthropic endeavor. The conceptual movement is from one of “service” to the community to one of “engaged scholarship” as an integral element of teaching, research and knowledge production, embracing the interdisciplinary, problem-based and experiential concepts. Service-learning is also a form of teaching and learning that combines subject learning with civic engagement. In the “Taking Social Responsibilities” project, learners put knowledge and skills learned in science class to practical use while making an active contribution to their communities.

The study conducted by Jensen and Burr (2006) found that service-learning activities can serve as a tool for engaging some students who would not normally respond to traditional methods of teaching. Moreover, results from Newman et al. (2015) study showed gains in three areas of measures of academic achievement and academic engagement, civic responsibility and resiliency.

Hayford et al. (2014) found that improving STEM literacy is also a focal point of STEM education and has been explicitly linked to service learning activities. Hayford et al. (2015) postulated that components of service-learning, including the deliverables and reflections, are examples of fundamental STEM literacy which can be assessed for STEM learning outcomes.

Moreover, the study (Daniel and Mishra, 2017) measured student affective, behavior, and content (ABC), as well as global awareness outcomes after participating in a STEM-based international service-learning course indicated positive improvements in affective outcomes and significant gains in civic awareness. The differences in behavior based on class of student, the specific content gains related to the service-learning activities. There were global awareness gains for all students and differential impacts on retention in STEM-related fields.

STEM fair/competition

There has been increasing student participation in science fairs, robotics competitions, computing contests and other STEM competitions every year. And research conducted on the relationship between student competition participation and their learning are reporting positive effect. For instance, the research on students participated in the RoboCupJunior, one of the World Championship, reported the effectiveness of its practice for enhancing learning of STEM content and skills for innovation and creativity among participating students (Eguchi, 2016). The result from 4th to 12th graders by a STEM assessment before and after the Robofest competition showed significant improvement in mathematics scores (Chung et al., 2014).

Miller et al. (2017) study using data from the “Outreach Programs and Science Career Intentions” survey reported that competitions are an effective way to foster career interest in specific STEM careers. Their result revealed that, first, students who participated in STEM competitions are more likely to express interest in a STEM-related career. Second, the relationship between competition participation and interest in a STEM career appears to be domain specific. Third, the impact of competition participation on pursuit of a STEM career is three times stronger when students compete in more than one competition.

So et al. (2017) adopted a researcher-generated framework to analyze STEM activities demonstrated in primary students’ projects in an annual STEM event, found that the framework does not only provided better understanding of assessment of students’ performance of STEM integration in projects, it also offered baseline information and suggestions for further research and instructional efforts in the planning and implementation of the future development of quality STEM education at the primary school level.

STEM information website

Nakamura et al. (2018) considered that knowledge structure of STEM is huge and complicated, that visualization of knowledge structure is very useful for learning STEM. The STEM@HKUST not only seeks to be a one-stop shop of information on STEM-related events and competitions, but also showcases original videos produced by the university faculty, staff and students which explain everyday conundrums in a lively manner. Students can suggest topics or experiments they want to watch, while secondary school teachers with queries on STEM-related topics are also welcomed to raise questions through the website. Wabisabi learning also provided a list of useful websites with contributions for various STEM professionals to facilitate STEM learning including how to begin preparing students for these swiftly approaching occupations. Creating STEM classroom opportunities that excite and inspire our children is a great place to start.

STEM network website

National Lab Network is an initiative that matches teachers and students with members of the STEM community. It began as “National Lab Day” by connecting schools with educational, scientific, and engineering organizations nationwide to work together on projects related to STEM. The original one-day event has evolved into an ongoing, collaborative network of schools and community leaders. Teachers can request community support by posting information about their STEM classroom projects. They were

matched with local STEM volunteers including university STEM students and faculty, scientists, engineers, professionals, and other community members involved in STEM. Projects ranged from helping students with science fairs to coordinating or hosting field trips to mentoring a student with STEM work.

Science museum visit

Given the growing interest in supporting STEM learning opportunities for youth, museums and other designed settings offer extensive experience and expertise as resources to support youth and communities in deepening and expanding such opportunities (Chi et al., 2015). The informal science learning environments such as science museums might provide valuable spaces within which to engage visitors (Archer et al., 2016). They explored the experiences of socially disadvantaged families in a science museum visit and found that families' experiences could be characterised into three discourses, as "disorientating", "fun" or "meaningful" visits. Analysis identified the families' experiences, and the likelihood of deriving science learning from the visit were shaped through interactions of habitus and capital.

STEM-rich learning environments like science museums which expose children to STEM can spark and nurture children's interest in and enthusiasm for STEM (Traphagen and Traill, 2014). The study conducted by Haden et al. (2014) facilitated an educational program in a museum. Prior to the exhibition, the families were randomly assigned to conditions that varied in terms of the instructions about a key engineering principle and elaborative question-asking they received. The participants were then observed during the exhibition. Conversation instruction resulted in adults asking double the number of Wh-questions compared to families who did not receive the instruction. The study found that the instruction was effective in terms of promoting family conversations and children's learning about STEM.

STEM outreach activities

Vennix et al. (2018) investigated outreach activities developed by STEM-based companies or universities with an aim to inform students about and motivate them for a career in STEM by connecting the work-context with school-science. Their research investigated students' perceptions with the outreach learning environment, perceived need-fulfilment, self-reported motivation and attitudes toward STEM. The findings show that outreach learning environments created opportunities to increase students' motivation in STEM and attitude toward STEM, but the impact varied according to the particular characteristics of the activities. Moreover, the effectiveness of STEM outreach program studied by Dubetz and Wilson (2013) noted the increase in scientific familiarity that encouraged girls to consider related careers and helped develop better informed scientific citizens.

Another study on paired peer learning through engineering education outreach by Fogg-Rogers et al. (2016) observed that the activities are having multiple benefits to both students, teachers and engineers. First, increased school children's interest and engagement with science and engineering. Second, provided teachers with expert contributions to engineering subject knowledge. And third, developing engineers with professional generic skills of communication and teamwork.

The scientific and technological community has a vital role to play in reaching out to make education reform efforts and encourage young people to study and pursue careers in STEM (Kesidou and Koppal, 2004). Yet, it would be better to focus on identifying resources and strategies that can be used to enrich outreach efforts that aim to supplement or enhance STEM content for students. The three recommendations are: (1) align outreach efforts to relevant content standards, that is, close to the curriculum; (2) pay attention to what students are thinking, that is to have continuous evaluation of students' achievement; and (3) take advantage of instructional strategies that work, that is to have association with teaching and learning in classrooms.

Family engagement in STEM

Jackson and Harbison (2014) showed that family involvement in their child's education has a greater impact on the child's learning than variables such as social class, level of parental education or parental income. In a study about young children and science (Silander et al., 2018), it was found that 99% of parents wanted to be involved in their children's education, but they did not understand the role they can play. For instance, they did not know that there was no need to have the right answer nor communicate the information appropriately; rather, parents were there to support their children's exploration.

It is suggested by the family engagement, A STEM NEXT Project Opportunity fund initiative, to empower parents to support their children's participation and persistence in STEM. Family engagement can start with listening, learning, and building relationships with families. Parents should be empowered with research and resources. They do not need to be the experts or to have the answers, particularly in relation to STEM; rather, it is their encouragement that matters. Providing professional development to team members for impactful family engagement to better engage parents and sustain further exploration at home is also one key practice for successful family engagement. Last but not least, evaluating the impact with a brief survey instrument using social media and through online platforms is recommended. During the whole process, it is also important to prioritize access and inclusion to consider who benefits and who is left out and left behind.

The effect of parent engagement in children's STEM learning is still under investigation. The meta-analysis of the literature on parental engagement with children's STEM education (Milner-Bolotin and Marotto, 2018) reported the followings themes: (1) The challenges of supporting parents with children's STEM education; (2) STEM education as a bridge between school and family;

(3) STEM education as a gateway for children's future economic success; (4) STEM education as a vehicle for promoting student communication skills; and, (5) the role of hands-on inquiry-based activities in enhancing student engagement.

Recommendations

Based on the above discussions of STEM learning ecosystems with the roles of individuals, families, schools and the society in STEM learning; the importance of students' STEM aspiration to bridge the gap between STEM education and the required work force for the growth of innovation and technology in the contemporary world; and the significance of engaging communities in STEM education and the various practices with community engagement in STEM education, the following recommendations are proposed for continuous and future development to engage community stakeholders in students' STEM learning. These include establishing STEM professional-school collaboration, developing STEM communities of practice, and highlighting STEM professional contributions to society.

Establishing STEM professional-school collaboration

There are limited practical ways to help students understand the nature of science careers. Wyss, Heulskamp, and Siebert (2012) suggested providing school students with accurate information about STEM careers to enable them to make more knowledgeable choices about courses of study and career paths. They investigated the use of video interviews with STEM professionals to better inform students about STEM career possibilities, and got compelling evidence of implementing video interviews to arouse students' interest in pursuing STEM careers.

Watters and Diezmann's (2013) study established and implemented partnerships with industry experts in an attempt to stimulate youth's interest in the contemporary world of science. The formation of relationships and partnerships induced students' high degree of autonomy and sense of responsibility, which are important attitudes toward STEM.

El-Deghaidy and Mansour (2015) conducted a STEM partnership-based Professional Development program, in which participating teachers suggested that a direct dialogue between science teachers, mathematics teachers, scientists, and engineers about STEM applications and activities is essential for promoting STEM education in schools. This provides insights for teachers and scientists to work together to assist students in making learning more meaningful.

Cheng and So (2020) reported an example of the total integration of STEM learning which provided opportunities for students to engage in simulation activities as water engineers. It began with some basic information about the daily tasks of water engineers, followed by students' mimicking water engineering to solve the problem of a sustainable supply of energy for reservoirs by designing a mini floating solar power system to capture renewable energy from the sun. At the same time, students could work on investigating problems, such as the effect on algal growth arising from the use of the solar panels on reservoirs. The design of the STEM learning scaled up from simulation to real situations to solve an authentic problem encountered by the STEM professional.

Developing STEM communities of practice (STEM CoPs)

STEM education for our young generation is more than just about knowledge, skills and thinking processes; it is a matter of literacy that all school leavers should acquire (So and Chiu, 2020). Scientists helped us better understand the behaviors of the natural and physical world, based on scientific principles, engineers solved problems and created things that had never existed before. Mackie (2019) proposed a high-functioning STEM community which is youth-focused, action-oriented, hands-on, and adult-coached. It is a means for the STEM community to not only transfer relevant knowledge, best practices, successful practicalities, and expected social norms, but also to share, embed, and make easily reachable best practices and lessons learned from the STEM stakeholders in the community. An example is the community of STEM NOLA which leveraged what is believed to be the best from that classic model, with one group of paid college students to help with the group of volunteer STEM professionals, retired professionals and community elders.

The Community of Practice (CoP) was defined by Wenger et al. (2002) as "Groups of people who share a concern, a set of problems, a passion about a topic and who deepen their knowledge and expertise in that area by interacting on an ongoing basis" (p. 4). STEM education through a Community of Practice (STEM CoP) can also provide opportunities for teachers and students to learn the common and unique practices of STEM professionals in the community and to identify key learning outcomes necessary to achieve what are grounded in common shared practices in real-world contexts. An example of a STEM CoP is the community effort in plastic waste recycling (Chow et al., 2016) focusing on the plastic waste problem in a city. The common practices are the practical experiences of developing, testing, using and maintenance of bins for plastic recycling. This is followed by interaction with the community for innovative ideas and strategies for recycling plastic "waste".

Highlighting STEM professional contribution to the society

Due to the rapid growth in innovation and technology, the future jobs and occupations taken by today's learners may not exist as of today; furthermore, many of these future jobs are probably STEM-related. Yet, without even mentioning future occupations,

students nowadays do not even know the existing STEM occupations serving our community. How to excite and inspire our children and prepare students for the demand of the innovation and technology growth and swiftly approaching STEM occupations in our community is a timely question.

One of the directions of STEM education should be closely related to the daily operation and development of the local community with various government services to serve the people living in its jurisdiction. To strengthen this linkage, we should consider stepping up and pitching the public education programs from various STEM-related Government departments toward our young generation (So and Chiu, 2020). The services, ranging from food packaging and safety, manufacturing of disinfecting cleaners during pandemics, architecture for city planning and road construction, drainage systems for sewerage treatment and flood prevention, firefighting and calamities protection, water and energy supply, could be made available for students to relate their STEM learning to everyday life, and to spark their STEM interest.

The mass media are also a key channel to promote STEM. Any promotion messages and programs can benefit students and their parents, to have some impact on steering them toward the study and career choice of a STEM profession. This will eventually contribute to the community goal of economic development through science and technology. A television program “Our Scientist” produced by Radio Television Hong Kong not only aroused public concern for and awareness of the importance of STEM contributions to society, but was also successful in revealing scientists’ non-nerdy images as a person (So and Chiu, 2020). The program was successful in demonstrating that, other than scientific work, the male environmental hydraulics professional being interviewed also excels in ballet and table-tennis.

Concluding remarks - a look into the future

Reflecting on the development of STEM education in the United States during the past few decades, the NSTA Board of Directors adopted in February 2020 a new Position Statement on STEM Education Teaching and Learning indicating that STEM education has been evolving from a convenient clustering of four overlapping disciplines toward a more cohesive knowledge base and skill set critical for the 21st century (<https://www.nsta.org/about/positions/stem.aspx>). New goals for a Modern STEM education that will promote not only skills such as critical thinking, problem solving, higher-order thinking, design, and inference, but also behavioral competencies such as perseverance, adaptability, cooperation, organization, and responsibility are advocated. A strong foundation in language and creative arts, together with the ability to fully comprehend global issues were also emphasized. With a number of recommendations, STEM education leaders are advised to continue to broaden and deepen the scope of STEM education and to involve the entire community to play a critical role in shaping the STEM vision and supporting its implementation through active participation. It is worth noting that the emphasis on new goals of STEM education and some additional competencies as mentioned above is in line with a recent global trend concerning education for the future and the development of 21st Century literacy/competencies (Howard et al., 2019; OECD, 2018).

Increasingly, in the past two decades, educators and policy makers around the world have come to realize the importance and potential impacts of the volatile, uncertain, complex, and ambiguous nature of the 21st Century on the future development of citizens, the human race, and our entire planet. Taking a more holistic and systemic way of thinking, a number of educators, research projects, and international organizations have reflected on the shortcomings and limitations of our past and current education systems and practices, while offering a range of suggestions for the transformation to a better global future education. Some earlier works in this direction have been extensively reviewed by a report supported by the World Innovation Summit for Education (Liu, 2015), while results of more recent studies, conference reports and so on continue to appear (GELP, 2017, 2019; Luksha et al., 2018; OECD, 2018, 2019).

For instance, based on a series of previous studies, the OECD proposed the Learning Framework 2030 which offers a vision and some principles for the future of education (OECD, 2018). This framework emphasizes a commitment to helping every learner develop as a whole person, fulfill his or her potential and help shape a shared future built on the well-being of individuals, communities and the planet. It also outlines a number of underpinning principles for the future of education systems, including the need for learner agency in terms of navigating through a complex and uncertain world, the need for a broad set of knowledge, skills, attitudes and values in action, and especially a new set of what they called “Transformative Competencies” that together address the growing need for young people to be able to create new value, to reconcile tensions and dilemmas, and to take responsibility. To ensure that the new learning framework is actionable, the OECD Learning Framework 2030 also includes some design principles for moving toward an eco-systemic change.

Besides the OECD (2018), GELP (2017, 2019), Schleicher (2018), Cook (2019), and many others have made efforts along similar lines. Without going into further details, suffice to say that as an integral part of the education system, STEM education, especially with community engagement, has significant new roles to play in the future. The concepts and practices briefly introduced in this article can be integrated into efforts to reform education to better equip future generations to meet the challenges posed not only by issues of technological and economic advancement, but also equity, social justice, environmental degradation, climate change and so on, and to help shape a shared future built on the well-being of individuals, communities and the planet. There are a significant number of partnerships and programs in place that bring together schools, industry and universities to support STEM engagement. However, more needs to be done to improve the coordination and reach of these arrangements, particularly for under-represented students, and to identify best practices (Education Council, 2015). Of course, developing a STEM community is time-intensive, challenging, and demands dedication, but it can be done.

Acknowledgments

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [EDUHK 18607717].

References

- Archer, L., Dawson, E., Seakins, A., Wong, B., 2016. Disorientating, fun or meaningful? Disadvantaged families' experiences of a science museum visit. *Cult. Stud. Sci. Educ.* 11, 917–939.
- Archer, L., DeWitt, J., Wong, B., 2014. Spheres of influence: what shapes young people's aspirations at age 12/13 and what are the implications for education policy? *J. Educ. Pol.* 29, 58–85.
- Archer, L., Osborne, J., DeWitt, J., Dillon, J., Wong, B., Willis, B., 2013. *ASPIRES: Young People's Science and Career Aspirations, Age 10–14*. King's College, London.
- Chachashvili-Bolotin, S., Milner-Bolotin, M., Lissitsa, S., 2016. Examination of factors predicting secondary students' interest in tertiary STEM education. *Int. J. Sci. Educ.* 38, 366–390.
- Chavatzia, T., 2017. *Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM)*. United Nations Educational, Scientific and Cultural Organization (UNESCO), Paris, France.
- Cheng, Y.C., So, W.W.M., 2020. Managing STEM learning: A typology and four models of integration. *Int. J. Educ. Manag.* 34 (6), 1063–1078. <https://doi.org/10.1108/IJEM-01-2020-0035>.
- Chi, B., Dorph, R., Reisman, L., 2015. *Evidence & Impact: Museum-Managed STEM Programs in Out-of-School Settings* (University of California, Berkeley: The Lawrence Hall of Science).
- Chow, C.F., So, W.W.M., Cheung, T.Y., 2016. Research and development of a new waste collection bin to facilitate education in plastic recycling. *Appl. Environ. Educ. Commun. Int. J.* 15, 45–57.
- Chung, C.J.C.J., Cartwright, C., Cole, M., 2014. Assessing the impact of an autonomous robotics competition for STEM education. *J. STEM Educ.* 15, 24–34.
- Cook, J.W., 2019. In: Cook, J.W. (Ed.), *Sustainability, Human Well-Being, and the Future of Education*, first ed. Springer Nature, Cham, Switzerland.
- Daniel, K.L., Mishra, C., 2017. Student outcomes from participating in an international STEM service-learning course. *SAGE Open* 7, 1–11.
- Dubetz, T.A., Wilson, J.A., 2013. Girls in engineering, mathematics and science, GEMS: a science outreach program for middle-school female students. *J. STEM Educ.* 14, 41–47.
- Education Council, 2015. *National STEM School Education Strategy. A Comprehensive Plan for STEM Education in Australia*. Available at: <http://www.educationcouncil.edu.au/site/DefaultSite/filesystem/documents/National%20STEM%20School%20Education%20Strategy.pdf>.
- Eguchi, A., 2016. RoboCupJunior for promoting STEM education, 21st century skills, and technological advancement through robotics competition. *Robot. Autonom. Syst.* 75, 692–699.
- EL-Deghaidy, H., Mansour, N., 2015. Science teachers' perceptions towards STEM education in Saudi Arabia: possibilities and challenges. *Int. J. Learn. Teach.* 1 (1), 51–54. <https://doi.org/10.18178/ijlt.1.1.51-54>.
- Fogg-Rogers, L., Lewis, F., Edmonds, J., 2016. Paired peer learning through engineering education outreach. *Eur. J. Eng. Educ.* 42, 75–90.
- GELP, 2017. *Transforming Education for Complexity: Why, what, and How*. Report on GELP Moscow 2017 Conference Results [pdf] Available at: https://www.gelponline.org/wp-content/uploads/2019/03/GELP17_feb27_2.pdf.
- GELP, 2019. *Anticipating the Future: Leading Education Transformation*. Report on GELP Israel 2019 Conference Results [pdf] Available at: https://www.gelponline.org/wp-content/uploads/2019/10/GELP_Israel_report_Final%20Sep9.pdf.
- Han, S., Capraro, R., Capraro, M., 2015. How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: the impact of student factors on achievement. *Int. J. Sci. Math. Educ.* 13, 1089–1113.
- Haden, C.A., Jant, E.A., Hoffmana, P.C., Marcus, M., Geddes, J.R., Gaskins, S., 2014. Supporting family conversations and children's STEM learning in a children's museum. *Early Child. Res. Q.* 29, 333–344.
- Hansen, L.M., 2018. Examining ways to meaningfully support students in STEM. *Int. J. STEM Edu.* 5, 1–6.
- Hayford, B., Blomstrom, S., Deboer, B., 2014. STEM and service-learning: does service-learning increase STEM literacy? *IJRSLE* 2, 32–43.
- Hayford, B., Blomstrom, S., Mumpower, L., 2015. Formation and assessment of a tool to evaluate STEM literacy in service-learning projects. *J. Excel. Coll. Teach.* 26, 73–101.
- Hinojosa, T., Rapaport, A., Jaciw, A., LiCalsi, C., Zacamy, J., 2016. Exploring the foundations of the future STEM workforce: K–12 indicators of postsecondary STEM success (REL 2016–122). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest, Washington, DC. <http://ies.ed.gov/ncee/edlabs>.
- Howard, P., O'Brien, C., Kay, B., O'Rourke, K., 2019. Leading educational change in the 21st Century: creating living schools through shared vision and transformative governance. *Sustainability* 11, 4109.
- Jackson, J., Harbison, L., 2014. An evaluation of the utility of homework in Irish primary school classrooms. *Irish Teach. J.* 2 (1), 47–62.
- Jang, H., 2016. Identifying 21st Century STEM competencies using workplace data. *J. Sci. Educ. Technol.* 25, 284–301.
- Jensen, S., Burr, K.L., 2006. Participation and learning relationships: a service-learning case study. *JSTE* 43, 6–28.
- Kesidou, S., Koppal, M., 2004. Supporting goals-based learning with STEM outreach. *J. STEM Edu.* 5, 5–16.
- Liu, J., 2015. *Education for the Future: The Global Experience of Developing 21st Century Skills and Competencies*. World Innovation Summit for Education Report. China Education Innovation Institute, Beijing Normal University [pdf] Available at: https://www.wise-qatar.org/app/uploads/2019/04/education_for_the_future.pdf.
- Lloyd, A., Gore, J., Holmes, K., Smith, M., Fray, L., 2018. Parental influences on those seeking a career in STEM: the primacy of gender. *Int. J. Gend. Sci. Technol.* 10, 308–328.
- Lee, Y., Carpraro, M.M., Viruru, R., 2018. The factors motivating students' STEM career aspirations: personal and societal contexts. *Int. J. Innovat. Sci. Mod. Eng.* 26, 36–48.
- Luksha, P., Cubista, J., Laszlo, A., Popovich, M., Ninenko, I., 2018. *Global education futures report. Educational ecosystems for societal transformation* [pdf] Available at: http://www.globaledufutures.org/images/people/GEF_april26-min.pdf.
- Mackie, C., 2019. Developing A high functioning STEM community. *Forbes*, 28 August. Available at: <https://www.forbes.com/sites/calvinmackie/2019/08/28/how-to-develop-a-high-functioning-stem-community/#ed9b7f23c444>.
- Masnick, A.M., Valenti, S.S., Cox, B.D., Osman, C.J., 2010. A multidimensional scaling analysis of students' attitudes about science careers. *Int. J. Sci. Educ.* 32 (5), 653–667.
- Mau, W.C.J., Li, J., 2018. Factors influencing STEM career aspirations of underrepresented high school students. *Career Dev. Q.* 66, 246–258.
- Micari, M., Winkle, Z.V., Pazos, P., 2016. Among friends: the role of academic-preparedness diversity in individual performance within a small group STEM learning environment. *Int. J. Sci. Educ.* 38, 1904–1922.
- Miller, K., Sonner, G., Sadler, P., 2017. The influence of students' participation in STEM competitions on their interest in STEM careers. *Int. J. Sci. Educ.* 8, 95–114.
- Milner-Bolotin, M., Marotto, C.C.F., 2018. Parental engagement in children's STEM education. Part I: meta-analysis of the literature. *Int. J. Math. Sci. Technol. Edu.* 6, 41–59.
- Morrison, J., Fisher, W.P., 2019. Measuring for management in science, technology, engineering, and mathematics learning ecosystems. *J. Phys. Conf. Ser.* 1379, 012042.

- Nakamura, A., Kudo, T., Nishioka, K., 2018. Development of the visualizing system of knowledge structure based on STEM eLearning website. In: Proceedings of the 9th International Conference on Language, Innovation, Culture and Education. Bangkok, 24–25 February 2018, pp. 55–61. Available at: <http://icsai.org/procarch/9iclice/9iclice-036.pdf>.
- Newman, J.L., Dantzier, J., Coleman, A.N., 2015. Science in action: how middle school students are changing their world through STEM service-learning projects. *Theory Into Pract.* 54, 47–54.
- Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C.R., Nelson, C., 2015. A model of factors contributing to STEM learning and career orientation. *Int. J. Sci. Educ.* 37, 1067–1088.
- OECD, 2018. The Future of Education and Skills: Education 2030. Progress Report of the OECD Education 2030 8th International Working Group. OECD, Paris. Available at: [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf).
- OECD, 2019. Trends Shaping Education 2019. OECD Publishing, Paris. Available at: <http://www.oecd.org/education/trends-shaping-education-22187049.htm>.
- Olson, S., Labov, J.B., 2014. STEM Learning is Everywhere: Summary of a Convocation on Building Learning Systems. The National Academies Press, Washington, DC.
- Penuel, W.P., Lee, T., Bevan, B., 2014. Designing and Building Infrastructures to Support Equitable STEM Learning across Settings. The Research+Practice Collaboratory, San Francisco, CA.
- Preece, J., 2017. Community engagement and its evolving terminology. In: Preece, J. (Ed.), *University Community Engagement and Lifelong Learning*. Springer Nature, Druban, South Africa, pp. 49–73.
- Schleicher, A., 2018. World Class: How to Build a 21st-Century School System. OECD Publishing, Paris. Available at: <https://www.oecd.org/education/world-class-9789264300002-en.htm>.
- Silander, M., Grindal, T., Hupert, N., Garcia, E., Anderson, K., Vahey, P., Pasnik, S., 2018. What Parents Talk About When They Talk About Learning: A National Survey About Young Children and Science. Education Development Center, Inc., & SRI International, New York, NY & Menlo Park, CA.
- So, W.M.W., Chiu, W.K.S., 2020. Challenges and Opportunities With Hong Kong Students' Science, Technology, Engineering and Mathematics Aspirations. Public Policy Research Funding Scheme Report. Policy Innovation and Co-ordination Office of the Government of the HKSAR.
- So, W.W.M., Zhan, Y., Chow, S.C.F., Leung, C.F., 2017. Analysis of STEM activities in primary students' science projects in an informal learning environment. *Int. J. Sci. Math. Educ.* 16, 1003–1023.
- STEM Next, 2020. A STEM NEXT Project Opportunity Fund Initiative. Available at: <https://stemnext.org/engaging-families/>.
- STEM NOLA, 2020. STEM NOLA. Available at: <http://stemnola.com/>.
- The Hong Kong University of Science and Technology, 2020. STEM@HKUST. Available at: <https://stem.ust.hk/>.
- The White House, 2018. Summary of the 2018 State-Federal STEM Education Summit. <https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/06/Summary-of-the-2018-White-House-State-Federal-STEM-Education-Summit.pdf>.
- Traphagen, K., Traill, S., 2014. How Cross-Sector Collaborations are Advancing SETM Learning. The Noyce Foundation, Los Altos, CA.
- Tytler, R., Symington, D., Williams, G., White, P., 2018. Enlivening STEM education through school-community partnerships. In: Jorgensen, R., Larkin, K. (Eds.), *STEM Education in the Junior Secondary*. Springer, Singapore, pp. 249–272.
- Vennix, J., Brok, P.D., Taconis, R., 2018. Do outreach activities in secondary STEM education motivate students and improve their attitude towards STEM? *Int. J. Sci. Educ.* 40, 1263–1283.
- Wabisabi learning. A STEM Resource List of More Than 40 Useful Websites. Available at: <https://wabisabilearning.com/blogs/stem/stem-resource-list-40-useful-websites>.
- Watters, J., Diezmann, C., 2013. Community partnerships for fostering student interest and engagement in STEM. *J. STEM Educ.* 14 (2), 47–55.
- Wenger, E., McDermott, R., Snyder, W.M., 2002. *Cultivating Communities of Practice*. Harvard Business School Press, Boston, MA.
- Wyss, V.L., Heulskamp, D., Siebert, C.J., 2012. Increasing middle school student interest in STEM careers with videos of scientists. *Int. J. Environ. Sci. Educ.* 7 (4), 501–522.

Relevant Websites

- A STEM NEXT Project Opportunity Fund Initiative ' <https://stemnext.org/engaging-families/>.
- Community Involvement in STEM Education' <http://occhildrenandfamilies.com/community-involvement-in-stem-education/>.
- Developing A High Functioning STEM Community' <https://www.forbes.com/sites/calvinmackie/2019/08/28/how-to-develop-a-high-functioning-stem-community/#ed9b7f23c444>.
- RTHK.HK' https://www.rthk.hk/tv/dtt31/programme/ourscientists_eng?lang=en.
- STEM NOLA' <http://stemnola.com/>.
- STEM@HKUST' <https://stem.ust.hk/>.
- Taking Social Responsibilities' <https://www.siemens-stiftung.org/projects/service-learning-in-stem/>.
- Wabisabi learning' <https://wabisabilearning.com/blogs/stem/stem-resource-list-40-useful-websites>.

Informal STEM learning

Nancy Staus, Kelly Riedinger, and Martin Storksdieck, STEM Research Center, Oregon State University, Corvallis, OR, United States

© 2023 Elsevier Ltd. All rights reserved.

What is informal STEM learning?	244
Learning in informal STEM environments	245
The role of informal STEM learning in the larger STEM ecosystem	245
Broadening participation in informal STEM learning	246
Conclusions and future directions	247
Supporting equity and inclusion for all learners	247
Addressing adult life-long informal STEM learning	248
Documenting long-term informal STEM learning pathways	248
References	248

What is informal STEM learning?

Individuals learn science, technology, engineering and mathematics (STEM) across a wide range of contexts, settings and experiences. While some STEM learning takes place within the formal classroom, much of it happens outside of school through the process of informal STEM learning (National Research Council [NRC], 2009; National Academies of Sciences, Engineering and Medicine, 2018; Rennie, 2007). Although the term “informal” may suggest a lack of seriousness or rigor, in reality it refers to the context in which these learning experiences take place. Informal STEM learning represents a range of learning opportunities that happen beyond the classroom or other formal schooling environments such as through printed media, the internet, at museums and in afterschool programs (Crane, 1994; NRC, 2009, 2015; Rennie, 2014; Storksdieck et al., 2016). Because it takes place outside of the formal classroom, informal STEM learning is usually voluntary, self-motivated, and guided by the learners’ needs and interests rather than by curriculum or standards (NRC, 2009). In addition, whereas learning in the classroom is usually an individual endeavor, learning in informal STEM settings is often socially constructed as families or groups of friends make sense of informal STEM experiences at museums or related institutions and during everyday activities at home (Rennie et al., 2003).

Although informal STEM learning and formal school learning are often presented as dichotomous entities, they can be conceptualized as a continuum of learning opportunities defined by a range of characteristics including choice, facilitation, purpose and assessment. Formal schooling, for example, falls at one end of the spectrum and is characterized as highly structured, designed to address a specific curriculum, provides limited choice and incorporates assessments that determine whether or not a learner has adequately met a particular objective. Opportunities to learn STEM in everyday settings such as watching a science documentary at home fall at the other end of the spectrum in that they are not aligned with any particular curriculum or learning outcomes, offer high levels of choice and have limited structure and facilitation strategies. However, many other informal STEM learning experiences fall somewhere in between these extremes. Field trips, for instance, often take place in informal settings such as museums or zoos, and may offer varied levels of choice for learners, but they also can be highly structured and facilitated, may incorporate assessments that count toward a student’s grade and typically are aligned with a curriculum or have an intended and specific learning outcome (Behrendt and Franklin, 2014; DeWitt and Storksdieck, 2008; Storksdieck, 2006).

In light of the inherent complexity of categorizing informal STEM learning experiences based on the above range of characteristics, the Committee on Learning Science in Informal Environments (NRC, 2009) chose to organize its analysis based on the places in which informal STEM learning typically occurs, focusing on the following three categories of venues: everyday learning, designed environments and programs.

Everyday learning represents the day-to-day types of activities where individuals engage with STEM, either intentionally or incidentally (NRC, 2009, 2015; Tal and Dierking, 2014). In these settings, STEM learning is not necessarily an explicit or intended outcome but may result as individuals encounter STEM content, practices or ways of thinking. Generally, everyday environments have a low degree of structure and facilitation. Examples of everyday environments include (but are not limited to) engaging with STEM oriented media or the internet, playing at the park and family activities where science is encountered (e.g., cooking, gardening, family discussions) (NRC, 2009, 2015; Storksdieck and Falk, 2020).

Designed environments are intentionally created for specific STEM learning outcomes and commonly have collections of artifacts, materials or animals. As compared to everyday settings for learning, these environments may have a higher degree of structure and facilitation but are more sporadic in nature than programs. Learners usually have significant choice in designed settings to select what exhibits and experiences to attend and for how long they want to engage. Examples of designed settings including museums, science centers, nature centers, zoos and aquariums (NRC, 2009, 2015; Falk and Storksdieck, 2005).

Programs are also designed for a specific outcome and like designed environments, tend to have a higher degree of structure and facilitation. However, they are generally a more consistent (likely to occur over time) learning opportunity as compared to designed settings as well as more aligned with a specific curricular goal. In some cases, programs are offered through a designed environment such as a summer science camp program offered at a local zoo. Generally, programs tend to focus on a prescribed audience or group of learners such as an afterschool program targeting young learners, or a volunteer group that caters to adult learners. Examples of STEM programs for youth include field trips, science clubs, afterschool programs, and robotics programs, while adults may participate in hobby clubs, citizen science programs or learning vacations. In addition there are specialized STEM programs such as teacher professional development opportunities in non-school settings (e.g., museums) (NRC, 2009, 2015).

Learning in informal STEM environments

Significant STEM learning can and does take place in each of the venues described above; however, it can be challenging to document and measure the extent of learning outcomes due to the unique, self-directed and highly personal nature of such experiences which make them difficult to evaluate with standard measures (Allen and Peterman, 2019). Consequently, most informal STEM learning research to date has been limited in both duration and scope, designed to document the results of short-term experiences such as the learning outcomes from a single designed environment visit (e.g., science center) or participation in a program (e.g., field trip), while comparatively little research has addressed the learning that takes place in unstructured everyday environments or the long-term outcomes of participation (Allen et al., 2019; Dabney et al., 2012; NRC, 2015; Young et al., 2017).

Substantial research has examined informal STEM learning in designed spaces such as museums, science centers, zoos and aquariums, and these studies have provided considerable evidence of learning, even over the span of just a single visit (NRC, 2009). Visitors of all ages often learn STEM content and increase their scientific understandings based on the exhibits and activities they engage in, as evidenced by significant factual recall even several months after the experience (Anderson et al., 2007; Falk and Storksdieck, 2005, 2010; Jones, 2005; Korn, 2006; Medved and Oatley, 2000). In addition, researchers have demonstrated that many visitors engage in scientific reasoning behaviors such as meaning making and explanation of scientific principles as they explore and make sense of experiences, particularly interactive exhibits, in designed settings (e.g., Allen, 2007; Goldowsky, 2002; Randol, 2005; Zimmerman et al., 2010). Others have found that these learning outcomes are not universal, but in fact are strongly influenced by visitor goals and motivations. For example, Packer and Ballantyne (2002) identified five categories of visitor goals reflecting peoples' stated reasons for visiting educational leisure institutions such as museums and interpretive sites: (1) passive enjoyment; (2) learning and discovery; (3) personal self-fulfillment; (4) restoration; and (5) social contact. Their research showed that only visitors with learning and discovery goals exhibited significant learning and interest development outcomes. Using a similar visitor motivation taxonomy, Falk et al. (2008) found that motivations accounted for about a quarter of the conservation-related attitudes reported by zoo and aquarium visitors, further suggesting that visitor goals and motivations may be strong drivers for what participants ultimately learn in designed settings.

Out-of-school STEM programs (including afterschool, summer camps, clubs, and field trips) also have been extensively studied to better understand how they may benefit STEM learning for participating youth. Such programs provide expanded content-rich learning opportunities, often engaging students in rigorous, purposeful activities that feature hands-on engagement that inspire inquiry, reasoning, and problem-solving (Afterschool Alliance, 2014; NRC, 2015). In addition, out-of-school STEM activities may allow students to meet STEM professionals and learn about STEM careers (Bevan and Michalchik, 2013; Fadigan and Hamrich, 2004), and can help learners to expand their identities as achievers in the context of STEM as they are actively involved in producing scientific knowledge and understanding (Barton and Tan, 2010). Despite the strong potential for out-of-school STEM programs to provide positive benefits to participants, it has been difficult for researchers and evaluators to document significant changes in outcomes for youth as a result of participation in these programs (NRC, 2009). This may be due in part to a measurement limitation called the ceiling effect which can occur when a large proportion of subjects begin a study with very high scores on the measured variable(s), such that participation in an educational experience cannot yield significant gains among these learners (Judson, 2012; Staus et al., 2021b). This effect is often attributed to the biased nature of participation in informal STEM programs; that is, youth generally choose to participate because they are already interested in and potentially knowledgeable about the content area. Despite this limitation, several recent large-scale research and evaluation projects suggest that in the short term such programs can lead to increases in STEM interest, engagement, identity, career interest, career knowledge, and critical thinking (Allen et al., 2019; Riedinger and McGinnis, 2017; Young et al., 2017).

The role of informal STEM learning in the larger STEM ecosystem

Although short-term learning outcomes are important and worthwhile, evidence suggests that the principal benefits of informal STEM experiences are long-term transformations in learners as they pursue a "cascade" of experiences cumulatively across a diversity of physical and social contexts including school, home, museums, media, and science programs (Bronfenbrenner, 1979; National Academies of Sciences Engineering and Medicine, 2018a; NRC, 2009, 2010). For example, an individual's understanding of astronomy might represent the cumulative experiences of completing a classroom assignment about the planets, reading a book about cosmology by Stephen Hawking, visiting an air and space museum, watching relevant television programs, stargazing

through a telescope, and joining an astronomy club. For the individual, each of these experiences, most of them in the realm of informal STEM learning, combine to create a personal understanding of astronomy.

While fictitious, the above example of astronomy learning appears to be an accurate representation of the importance of informal STEM learning experiences in the learning of STEM more broadly. Given that most people spend only a small fraction of their life in school, informal STEM learning experiences such as those described above appear to be particularly important in contributing to people's STEM understandings over time (Falk and Dierking, 2010; Sosniak, 2001). Because virtually everyone engages in informal STEM learning activities nearly every day, from reading a news article about science to tending a garden, people accumulate STEM knowledge throughout their lifetimes, not just the years they are in school.

Accordingly, scholars have begun examining STEM learning from an ecological perspective, in which the educational "ecosystem" is conceptualized as a set of environments, both physical and virtual, in- and out-of-school, that provide people of all ages and backgrounds with opportunities to learn (Barron, 2010; Jackson, 2013; Traphagen and Traill, 2014). Within the larger ecosystem, learners construct unique STEM learning pathways guided by personal goals and interests and supported by a variety of sociocultural and physical factors including activities, resources, and relationships that the learner can engage with to support their personal interests and learning goals (Barron, 2006, 2010; Corin et al., 2017; Crowley et al., 2015; Shaby et al., 2021).

These studies have revealed that informal STEM learning experiences, particularly during adolescence, are important and may be critical for the development of enduring STEM interest over time. For example, Crowley et al. (2015) conducted retrospective life-history interviews with adult scientists and engineers inquiring about their early disciplinary interests and the ways those interests were developed and supported throughout their lives. They found that individual interest in science often emerged before high school and was supported by ample opportunities to learn in informal STEM learning settings such as enrolling in STEM programs, visiting informal learning settings, and/or pursuing information in books or online. Similarly, Hecht et al. (2019) explored retrospective life histories of naturalists to examine experiences that they recalled supporting their interest development in the natural world and found that persistent exposure to nature throughout their childhoods was critical for their long-term commitment to engagement with nature. In a large-scale survey of nearly 7000 nationally representative university students in the US, Dabney et al. (2012) found that students who participated in either structured or unstructured informal STEM learning activities were up to 1.5 times more likely to report a career interest in a STEM discipline. More recently, a longitudinal study of adolescents aged 11 to 13 indicated that more frequent participation in out-of-school STEM activities was a strong predictor of persistent STEM interest over time (Staus et al., 2020).

In addition to the importance of informal STEM learning early in life, there is growing evidence that suggests these experiences support the life-long learning trajectories of adults. A recent study of astronomy and birding hobbyists indicated that informal STEM learning organizations such as museums, universities and clubs were instrumental in connecting adult learners with relevant opportunities within the STEM learning ecosystem (Corin et al., 2017; Storksdieck et al., 2012; Yocco et al., 2012). As the hobbyists became more skilled and knowledgeable, such organizations then acted as catalysts to facilitate access to other elements of the STEM learning ecosystem such as hobby-related conferences or formal learning institutions to further develop their hobby. Similarly, citizen science programs aimed at adult participants are emerging as important informal STEM learning activities that facilitate both increased science literacy and the collection of highly rigorous, longitudinal data for researchers (He et al., 2019). However, although citizen science projects are rapidly increasing in number, the effect of such projects on adults' STEM learning outcomes and trajectories are not yet well documented (Brossard et al., 2005; National Academies of Sciences, Engineering, and Medicine, 2018b).

Broadening participation in informal STEM learning

Given the importance of informal STEM learning experiences for triggering and sustaining interest in STEM and providing valuable opportunities for individuals of any age to engage with and learn about STEM, there is growing interest in broadening participation in informal STEM learning, particularly for groups who have been historically under-represented or marginalized in STEM fields (DeWitt and Archer, 2017). In general, learning is maximized when it is supported and reinforced through participation in multiple experiences within the STEM learning ecosystem (DeWitt and Storksdieck, 2008; NRC, 2015). However, while a STEM learning ecosystem may contain a large variety of informal STEM learning experiences and resources, such opportunities may not be equally accessible to all learners. In fact, informal STEM learning resources are often disconnected and uncoordinated in such a way that learners may have difficulty finding "the next thing" that supports their particular interest (Bell, 2013). The disjointed nature of resources within learning ecosystems is particularly problematic for individuals from under-represented or disempowered groups (DeWitt and Archer, 2017). For example, a recent study of STEM interest and participation of adolescents in a diverse low-income neighborhood in Portland, Oregon, USA, found that the disconnected nature of the informal STEM opportunities in the learning ecosystem proved to be a significant constraint to persistence in STEM interests and activities for some youth whose families had difficulty identifying and connecting to relevant resources (Shaby et al., 2021). A major goal in the future will be to develop effective supports to enhance navigation of STEM ecosystems, perhaps by utilizing brokers who can help learners connect to experiences of interest (Allen et al., 2020; Ching et al., 2016).

The challenge of connecting learning settings appears to be widespread across learning ecosystems worldwide. A recent landscape study of UK science learning resources (Falk et al., 2015), revealed that the science education community was highly interconnected within sectors, but only moderately interconnected and collaborative between sectors (e.g., science centers, schools, hobby clubs). In particular, the school sector emerged as a major outlier that rarely connected or collaborated with informal science

learning sectors or even other schools, making it difficult for learners to identify and connect with additional learning opportunities related to their science interests. [Shaby et al. \(2021\)](#) also noted a lack of collaboration between formal and informal STEM providers in the learning ecosystem and suggested that successful ecosystem coordination requires that all STEM providers in an ecosystem commit to on-going and continuous communication among and between themselves, ensuring that opportunities and options are clearly signposted for youth and their families.

Recently, the *Connected Science Learning: Linking In-School and Out-of-School STEM Learning* (CSL) project – a collaborative effort between the National Science Teaching Association (NSTA) and the Association of Science and Technology Centers (ASTC) – addressed this issue by creating a scholarly journal designed to connect the work of STEM education practitioners with each other and with those conducting research in the area of informal STEM learning. An evaluation of the project showed that the journal provided great value to the field by featuring successful partnerships and models that create connections across the learning ecosystem. Additionally, STEM education professionals noted that the CSL journal translated research to practical applications, offered a mechanism for connecting STEM education practitioners, and provided an outlet where out-of-school education practitioners could write about and share details of their work ([Riedinger, 2017](#)). Efforts such as this have the potential to create more connected, functional STEM ecosystems for more learners.

But helping more people identify and connect with informal STEM learning resources does not address an even more pressing issue, which is the fact that some opportunities are not equally available to all learners due to social or cultural barriers to participation. Despite the perception that informal STEM learning by its nature is equally available to all, there is some evidence that institutions such as museums and science centers, rather than being easily accessible STEM resources for their communities, may actually exclude low-income and/or minoritized ethnic groups through their culture and practices ([Dawson, 2014a, 2014b, 2019](#)). For example, many institutions provide interpretive materials in only the dominant language, making it difficult for learners who speak different languages to understand and fully participate in the STEM activities ([Rahm and Ash, 2008](#)). In addition, many exhibits require an existing understanding of scientific terms and concepts that may exclude some visitors. The high financial cost of admission also may deter many potential visitors with limited incomes. These and other institutional practices reinforce social disadvantages such that minoritized visitors experience informal STEM learning institutions as unwelcoming places where their knowledge and practices are not valued or respected ([Young, 2000](#)). Clearly there is work to be done toward making informal STEM learning experiences inclusive to meet the needs of a diverse range of STEM learners.

Conclusions and future directions

Informal environments can provide rich opportunities for learning about STEM. Because virtually everyone participates in informal STEM learning in the course of daily life, from looking up health information on the internet to caring for a pet, informal STEM learning environments have the potential to support a broad range of learning experiences for any individual, of any age or background, with an interest in STEM topics or activities. Informal STEM learning experiences can stimulate interest in STEM, build STEM knowledge and skills, and support pathways of life-long STEM learning. However, there remain a number of areas for further exploration to provide more detailed understandings of learning processes and outcomes, as well as how to better support more learners in their STEM investigations.

Supporting equity and inclusion for all learners

Due to the often significant barriers to identifying and accessing informal STEM learning resources for some learners, especially members of low-income and minoritized communities, in the future informal STEM learning providers must better account for the fact that learners are not homogeneous groups but are rather comprised of members of numerous different socio-cultural groups whose science explorations are shaped by their unique backgrounds and lived experiences reflecting a diversity of science perspectives ([NRC, 2009](#)). In particular, there must be a rethinking of designed spaces and programs in order to provide more equitable and inclusive opportunities for all to participate in informal STEM learning experiences and environments ([Dawson, 2014a](#)). For example, providing multilingual exhibit texts and program delivery, training staff to work with a diverse range of visitors, incorporating alternative interpretations and different perspectives, and designing opportunities for cross-cultural meaning making within informal STEM learning institutions may be effective ways to create more inclusive experiences for more visitors and participants ([Dawson, 2014b; Rahm and Ash, 2008](#)). Another potential model for reaching a broader audience is illustrated by the Knowledge rooms project in Vienna, Austria in which science center activities were temporarily offered in empty shops in under-resourced urban districts whose residents did not have easy access to other designed spaces ([Streicher et al., 2014](#)). These “pop-up science centers” were well accepted and utilized by a socially mixed audience because they were welcoming and comfortable for the visitors, easy to access, and were seen as trustworthy actors inside the communities they served. This project showed that by meeting people where they are, we can potentially increase social inclusion in STEM and empower more communities to develop and generate STEM awareness and knowledge. A similar approach has been taken by Guerrilla Science, an approach to informal STEM learning that aims at reaching audiences who may not ordinarily engage with science ([Rosin et al., 2019](#)). Empirical studies ([O’Connell et al., 2020](#)) have shown that Guerrilla Science removes barriers to participation by embedding science into public places where science might be unexpected (access), and designs activities that speak to the identities of participants who might not ordinarily engage with science voluntarily (inclusion).

Addressing adult life-long informal STEM learning

As discussed earlier, research about the effectiveness of informal STEM experiences such as afterschool programs, summer camps, and field trips on youth outcomes is much more prevalent than research about how adults engage in informal STEM activities to support their STEM learning goals. However, several recent studies suggest that informal STEM learning experiences do contribute to adults' knowledge of STEM in meaningful ways. For example, a telephone survey of a thousand adult residents in greater Los Angeles, California, USA, showed that adult informal STEM learning experiences such as reading books and magazines about science and technology, using the internet, and watching science related documentaries and videos were significant predictors of self-reported knowledge of science and technology (Falk and Needham, 2013). Similarly, an international investigation of over 6000 adults living in 13 countries revealed that individuals who used science centers had significantly higher understanding, interest and curiosity, and identity relative to science and technology than did individuals who did not visit (Falk et al., 2016). And an online study amongst amateur astronomers revealed that basic astronomy knowledge was similar to or even higher than that of university instructors if the amateur astronomers were members of a club and had engaged in at least three years of public outreach (Berendsen and Storksdieck 2007; Storksdieck and Berendsen 2007). While these studies begin to build knowledge about the role of informal STEM in adult learning, given the widespread interest in better supporting public understanding of STEM in order to promote more scientifically literate societies, more research is needed to better understand the role of informal STEM resources and experiences in contributing to the STEM learning of adults (American Association for the Advancement of Science, 1994; Brown et al., 2005; Stockmayer et al., 2010).

Documenting long-term informal STEM learning pathways

Despite the long-term, cumulative nature of informal STEM learning, most research to date has been limited in both duration and scope (Anderson et al., 2007). In order to advance the informal STEM learning field, longitudinal studies and innovations in assessment will be essential to better document the construction of STEM learning pathways in order to illuminate how STEM interests and understandings originate and evolve over time across a variety of informal STEM learning contexts and settings (NRC, 2015; Staus et al., 2021a). While there are challenges to investigating long-term learning outcomes in any context, it is particularly the case within informal spaces in which learning pathways are unique to each individual, making it difficult to attribute learning to certain events or experiences (Falk et al., 2018). Similarly, because learners choose what, when and how long to participate in informal STEM learning programs and other activities, it is difficult to follow individuals over time, increasing the chance of potential measurement issues due to attrition (Taplan, 2005). Clearly, researchers will need to develop innovative methods to capture the accumulation of knowledge and understandings derived from multiple sources over longer time frames to address these challenges in future longitudinal informal STEM learning studies.

References

- Afterschool Alliance, 2014. *America After 3 pm: Afterschool Programs in Demand*. Washington, D.C. Retrieved from: <http://www.afterschoolalliance.org/AA3PM/>.
- Allen, S., Peterman, K., 2019. Evaluating informal STEM education: issues and challenges in context. In: Fu, A.C., Kannan, A., Shavelson, R.J. (Eds.), *Evaluation in Informal Science, Technology, Engineering, and Mathematics Education*, vol. 161. *New Directions for Evaluation*, pp. 17–33.
- Allen, P.J., Chang, R., Gorral, B.K., Waggenspack, L., Fukuda, E., Little, R.D., Noam, G.G., 2019. From quality to outcomes: a national study of afterschool STEM programming. *Int. J. STEM Educ.* 6 (37), 1–21. <https://doi.org/10.1186/s40594-019-0191-2>.
- Allen, S., Kastelein, K., Mokros, J., Atkinson, J., Byrd, S., 2020. STEM Guides: professional brokers in rural STEM ecosystems. *Int. J. Sci. Educ.* 10 (1), 17–35.
- Allen, S., 2007. *Secrets of Circles Summative Evaluation Report*. Report prepared for the Children's Discovery Museum of San Jose. Available: <https://www.informalscience.org/secrets-circles-summative-evaluation-report>. Accessed April 2021.
- American Association for the Advancement of Science, 1994. *Benchmarks for Science Literacy*. Oxford University Press, New York.
- Anderson, D., Storksdieck, M., Spock, M., 2007. The long-term impacts of museum experiences. In: Falk, J., Dierking, L., Foutz, S. (Eds.), *In Principle, in Practice: New Perspectives on Museums as Learning Institutions*. AltaMira Press, Walnut Creek, CA, pp. 197–215.
- Barron, B., 2006. Interest and self sustained learning as catalysts of development: a learning ecology perspective. *Hum. Dev.* 49, 193–224.
- Barron, B., 2010. Conceptualizing and Tracing Learning Pathways over Time and Setting, vol. 109. *National Society for the Study of Education*, pp. 113–127.
- Barton, A.C., Tan, E., 2010. "We be burnin!": agency, identity, and science learning. *J. Learn. Sci.* 19 (2), 187–229.
- Behrendt, M., Franklin, T., 2014. A review of research on school field trips and their value in education. *Int. J. Environ. Sci. Educ.* 9 (3), 235–245.
- Bell, P., 2013. Introduction: understanding how and why people learn across settings as an educational equity strategy. In: Bevan, B., Bell, P., Stevens, R., Razfar, A. (Eds.), *LOST Opportunities: Learning in Out-of-School Time*. Springer, New York, NY, pp. 95–98.
- Berendsen, M., Storksdieck, M., 2007. Knowledge of astronomy among amateur astronomers. In: Gibbs, M.G., Berendsen, M., Storksdieck, M. (Eds.), *Science educators under the stars: Amateur astronomers engaged in education and public outreach*. Astronomical Society of the Pacific, San Francisco, pp. 43–56.
- Bevan, B., Michalchik, V., 2013. Where it Gets Interesting: Competing Models of Stem Learning After School. *Afterschool Matters*. Retrieved from: <http://eric.ed.gov/?id=EJ1003837>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press.
- Brossard, D., Lewenstein, B., Bonney, R., 2005. Scientific knowledge and attitude change: the impact of a citizen science project. *Int. J. Sci. Educ.* 27 (9), 1099–1121.
- Brown, B.A., Reveles, J.M., Kelly, G.J., 2005. Scientific literacy and discursive identity: a theoretical framework for understanding science learning. *Sci. Educ.* 89, 779–802.
- Ching, D., Santo, R., Hoadley, C., Peppler, K., 2016. Not just a blip in someone's life: integrating brokering practices into out-of-school programming as a means of supporting and expanding youth futures. *Horizon* 24 (3), 296–312.
- Corin, E.N., Jones, M.G., Andre, T., Childers, G.M., Stevens, V., 2017. Science hobbyists: active users of the science-learning ecosystem. *Int. J. Sci. Educ.* 7 (2), 161–180.
- Crane, V., 1994. An introduction to informal science learning and research. In: Crane, V., Nicholson, H., Chen, M., Bitgood, S. (Eds.), *Informal Science Learning: What the Research Says about Television, Science Museums and Community-Based Projects*. Research Communications Ltd, Deham, MA.

- Crowley, K., Barron, B.J., Knutson, K., Martin, C., 2015. Interest and the development of pathways to science. In: Renninger, K.A., Nieswandt, M., Hidi, S. (Eds.), *Interest in Mathematics and Science Learning*. American Educational Research Association.
- Dabney, K.P., Tai, R.H., Almarode, J.T., Miller-Friedmann, J.L., Sonnet, G., Sadler, P.M., Hazari, Z., 2012. Out-of-school time science activities and their association with career interest in STEM. *Int. J. Sci. Educ.* 2 (1), 63–79.
- Dawson, E., 2014a. Equity in informal science education: developing an access and equity framework for science museums and science centres. *Stud. Sci. Educ.* 50 (2), 209–247.
- Dawson, E., 2014b. "Not designed for us": how science museums and science centers socially exclude low-income, minority ethnic groups. *Sci. Educ.* 98 (6), 981–1008.
- Dawson, E., 2019. *Equity, Exclusion and Everyday Science Learning: The Experiences of Minoritised Groups*, 1st ed. Routledge. <https://doi.org/10.4324/9781315266763>.
- DeWitt, J., Archer, L., 2017. Participation in informal science learning experiences: the rich get richer? *Int. J. Sci. Educ.* 7 (4), 356–373.
- DeWitt, J., Storksdieck, M., 2008. A short review of school field trips: key findings from the past and implications for the future. *Visitor Stud.* 11 (2), 181–197.
- Fadigan, K.A., Hammrich, P.L., 2004. A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program. *J. Res. Sci. Teach.* 41 (8), 835–860.
- Falk, J.H., Dierking, L., 2010. The 95 percent solution. *Sci. Am.* 98, 486–493.
- Falk, J.H., Needham, M.D., 2013. Factors contributing to adult knowledge of science and technology. *J. Res. Sci. Teach.* 50 (4), 431–452.
- Falk, J.H., Storksdieck, M., 2005. Using the contextual model of learning to understand visitor learning from a science center exhibition. *Sci. Educ.* 89, 744–778.
- Falk, J.H., Storksdieck, M., 2010. Science learning in a leisure setting. *J. Res. Sci. Teach.* 47 (2), 192–212.
- Falk, J.H., Heimlich, J., Bronnenkant, K., 2008. Using identity-related visit motivations as a tool for understanding adult zoo and aquarium visitors' meaning making. *Curator* 51 (1), 55–79.
- Falk, J.H., Dierking, L., Osborne, J., Wenger, M., Dawson, E., Wong, B., 2015. Analyzing science education in the United Kingdom: taking a system-wide approach. *Sci. Educ.* 99 (1), 145–173.
- Falk, J.H., Dierking, L.D., Swanger, L., Staus, N., Back, M., Barriault, C., Catalao, C., Chambers, C., Chew, L.-L., Dahl, S.A., Falla, S., Gorecki, B., Lau, T.C., Lloyd, A., Martin, J., Santer, J., Singer, S., Solli, A., Trepanier, G., Tytstjärvi, K., Verheyden, P., 2016. Correlating science center use with adult science literacy: an international, cross-institutional study. *Sci. Educ.* 100 (5), 849–876.
- Falk, J.H., Koke, J., Price, C.A., Pattison, S., 2018. *Investigating the Cascading, Long Term Effects of Informal Science Education Experiences Report*. Institute for Learning Innovation. Available: <https://www.informalscience.org/investigating-cascading-long-term-effects-informal-science-education-experiences-report>.
- Goldowsky, N., 2002. *Lessons from Life: Learning from Exhibits, Animals, and Interaction in a Museum*. UMI#3055856. Unpublished doctoral dissertation. Harvard University.
- He, Y., Parrish, J.K., Rowe, S., Jones, T., 2019. Evolving interest and sense of self in an environmental citizen science program. *Ecol. Soc.* 24 (2), 33. <https://doi.org/10.5751/ES-10956-240233>.
- Hecht, M., Knutson, K., Crowley, K., 2019. Becoming a naturalist: interest development across the learning ecology. *Sci. Educ.* 103 (3), 1–23.
- Jackson, N.J., 2013. *Learning ecology narratives*. In: *Lifewide Learning, Education and Personal Development E-Book*.
- Jones, J., 2005. *Huntington Botanical Gardens Summative Evaluation Conservatory for Botanical Science*. Report for Huntington Botanical Gardens. Available: <https://www.informalscience.org/huntington-botanical-gardens-summative-evaluation-conservatory-botanical-science>. accessed April 2021.
- Judson, E., 2012. Learning about bones at a science museum: examining the alternate hypotheses of ceiling effect and prior knowledge. *Instr. Sci.* 40, 957–973.
- Korn, R., 2006. *Search for Life: Summative Evaluation*. New York Hall of Science, New York. Available: <https://www.informalscience.org/summative-evaluation-search-life>. accessed April 2021.
- Medved, M.I., Oatley, K., 2000. Memories and scientific literacy: remembering exhibits from a science centre. *Int. J. Sci. Educ.* 22 (10), 1117–1132.
- National Academies of Sciences, Engineering, and Medicine, 2018a. *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24783>.
- National Academies of Sciences, Engineering, and Medicine, 2018b. *Learning Through Citizen Science: Enhancing Opportunities by Design*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/25183>.
- National Research Council, 2009. *Learning Science in Informal Environments: People, Places, and Pursuits*. National Academies Press, Washington, DC. <https://doi.org/10.17226/12190>.
- National Research Council, 2010. *Surrounded by Science: Learning Science in Informal Environments*. National Academies Press, Washington, DC. <https://doi.org/10.17226/12614>.
- National Research Council, 2015. *Identifying and Supporting Productive STEM Programs in Out-of-School Settings*. National Academies Press, Washington, DC. <https://doi.org/10.17226/21740>.
- O'Connell, K.B., Keys, B., Storksdieck, M., Rosin, M., 2020. Context matters: using art-based science experiences to broaden participation beyond the choir. *Int. J. Sci. Educ.* B 10 (2), 166–185. <https://doi.org/10.1080/21548455.2020.1727587>.
- Packer, J., Ballantyne, R., 2002. Motivational factors and the visitor experience: a comparison of three sites. *Curator* 45 (3), 183–198.
- Rahm, J., Ash, D., 2008. Learning environments at the margin: case studies of disenfranchised youth doing science in an aquarium and an after-school program. *Learn. Environ. Res.* 11 (1), 49–62.
- Randol, S.M., 2005. *The Nature of Inquiry in Science Centers: Describing and Assessing Inquiry at Exhibits*. Unpublished doctoral dissertation. University of California, Berkeley.
- Rennie, L.J., Feher, E., Dierking, L.D., Falk, J.H., 2003. Toward an agenda for advancing research on science learning in out-of-school settings. *J. Res. Sci. Teach.* 40 (2), 112–120.
- Rennie, L., 2007. Learning science outside of school. In: Abell, S.K., Lederman, N.G. (Eds.), *Handbook of Research in Science Education*. Lawrence Erlbaum, pp. 125–167.
- Rennie, L., 2014. Learning science outside of school. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research in Science Education*, second ed. Routledge, pp. 120–144.
- Riedinger, K., McGinnis, J.R., 2017. An investigation of the role of learning conversations in youth's authoring of science identities during an informal science camp. *Int. J. Sci. Educ.* 7 (1), 76–102.
- Riedinger, K., 2017. *Connected Science Learning: Linking In-School and Out-Of-School STEM Educators: Final Evaluation Report*. Oregon State University. <https://doi.org/10.5399/osu/1135>.
- Rosin, M., Wong, J., O'Connell, K., Storksdieck, M., Keys, B., 2019. Guerilla science: mixing science with art, music and play in unusual settings. *Leonardo*. https://doi.org/10.1162/leon_a_01793.
- Shaby, N., Staus, N.L., Dierking, L., Falk, J., 2021. Pathways of interest and participation: how STEM-interested youth navigate a learning ecosystem. *Sci. Educ.* 1–25. <https://doi.org/10.1002/sce21621>.
- Sosniak, L., 2001. The 9% challenge: education in school and society. *Teach. Coll. Rec.* 103, 15.
- Staus, N.L., Falk, J.H., Penuel, W., Dierking, L., Wyld, J., Bailey, D., 2020. Interested, disinterested, or neutral: exploring STEM interest profiles and pathways in a low-income urban community. *Eurasia J. Math. Sci. Technol. Educ.* 16 (6), em1853.
- Staus, N.L., Falk, J.H., Price, A., Tai, R., Dierking, L., 2021a. Measuring the long-term effects of informal science education experiences: challenges and potential solutions. *Discipl. Interdiscipl. Sci. Educ. Res.* 3, 3. <https://doi.org/10.1186/s43031-021-00031-0>.
- Staus, N.L., O'Connell, K., Storksdieck, M., 2021b. Addressing the ceiling effect when assessing STEM out-of-school time experiences. *Front. Educ.* <https://doi.org/10.3389/educ.2021.690431>.
- Stockmayer, S.M., Rennie, L.J., Gilbert, J.K., 2010. The roles of the formal and informal sectors in the provision of effective science education. *Stud. Sci. Educ.* 46, 1–44.
- Storksdieck, M., Falk, J.H., 2020. Valuing free-choice learning in national parks. *Parks Stewardship Forum* 36 (2), 271–280. <https://escholarship.org/uc/item/2z94016m>.
- Storksdieck, M., Stein, J.K., Jones, E.C., 2012. Hobbyists in the role of environmental educator: the case of amateur astronomy clubs. In: Kopnina, H. (Ed.), *Anthropology Environ. Educ.* Nova Science Publishers, New York, pp. 91–116.

- Storksdieck, M., Stylinski, C., Bailey, D., 2016. Typology for public engagement with science: a conceptual framework for public engagement involving scientists. Center for Research on Lifelong STEM Learning, Corvallis, OR. https://www.aaas.org/sites/default/files/content_files/AAAS_Typology.pdf.
- Storksdieck, M., 2006. Field trips in environmental education. Berliner Wissenschafts-Verlag, Berlin, Germany.
- Storksdieck, M., Berendsen, M., 2007. Attributes and practices of amateur astronomers who engage in education and public outreach. In: Gibbs, M.G., Berendsen, M., Storksdieck, M. (Eds.), *Science educators under the stars: Amateur astronomers engaged in education and public outreach*. Astronomical Society of the Pacific, San Francisco, pp. 31–42.
- Streicher, B., Unterleitner, K., Shulze, H., 2014. Knowledge rooms—science communication in local, welcoming spaces to foster social inclusion. *J. Sci. Commun.* 13 (2), C03.
- Tal, T., Dierking, L., 2014. Learning science in everyday life. *J. Res. Sci. Teach.* 51 (3), 251–259.
- Taplan, S., 2005. Methodological Design Issues in Longitudinal Studies of Children and Young People in Out-Of-Home Care: A Literature Review. Technical Report. NSW Centre for Parenting & Research. NSW Department of Community Services.
- Traphagen, K., Traill, S., 2014. How Cross-Sector Collaborations Are Advancing STEM Learning. Noyce Foundation.
- Yocco, V., Jones, E.C., Storksdieck, M., 2012. Factors contributing to amateur astronomers' involvement in education and public outreach. *Astr. Educ. Rev.* 11 (1). <https://doi.org/10.3847/AER2011040>.
- Young, J.R., Ortiz, N.A., Young, J.L., 2017. STEMulating interest: a meta-analysis of the effects of out-of-school time on student STEM interest. *Int. J. Educ. Math. Sci. Technol.* 5 (1), 62–74.
- Young, I.M., 2000. *Inclusion and Democracy*. Oxford University Press, Oxford, England.
- Zimmerman, H.T., Reeve, S., Bell, P., 2010. Family sense-making practices in science center conversations. *Sci. Educ.* 94 (3), 478–505.

K-12 teachers' beliefs and practices in STEM integration

Hui-Hui Wang and Neil A. Knobloch, Purdue University, West Lafayette, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	251
Definition of STEM integration	252
Teachers' beliefs and practices	253
Teachers' beliefs in STEM integration	253
Disciplinary ways of thinking	253
Integrated STEM examples based on teacher's beliefs	255
Single CTE discipline	255
Single S/M discipline	256
Multidisciplinary	256
Interdisciplinary	257
Transdisciplinary	257
Conclusions	257
Teachers' disciplinary ways of thinking	257
A broader and developmental approach to STEM integration	258
Challenges to implementing a broader approach to integrated STEM learning	259
Summary	259
References	259

Introduction

Worldwide, science, technology, engineering, and mathematics (STEM) education is experiencing reform and opening to try to keep pace with the rapidly changing world. Drawing from national and international student achievement data, such as Programme for International Student Assessment (PISA), and statistics from STEM occupations and employment, professional societies call for new educational approaches that simulate real-world complex problems, and focus on the hands-on, interdisciplinary and socially relevant aspects of STEM to improve K-12 STEM education and to build future STEM workforce pipelines.

STEM integration represents a way to think about curriculum change. It is a concept of how to restructure what we teach and what students learn (Herschbach, 2011). Integrated STEM learning could lead to important educational outcomes that include: (1) recognizing system dynamics (Repko, 2008); (2) closing gaps between STEM education from high school to college and the workforce (Tripp et al., 2017); (3) developing workforce skills to cope with complexity (USDA, 2015); and, (4) making scientifically and mathematically sound decisions when students solve problems in a complex designed system (Dym et al., 2005).

Although the theoretical basis for integrated STEM teaching is familiar to many STEM educators, the current educational reform movement no doubt creates certain barriers in the K-12 STEM education system (Herschbach, 2011). In the secondary education level, STEM subjects are typically taught separately, and tend to have disciplinary-specific curriculum models (Herschbach, 2011). Science and mathematics are considered as foundational and formal knowledge and are aligned with standardized tests and discipline-based structure and rules, which could be argued that science and mathematics are focused on preparing high school students for college. In contrast, technology, engineering, and agriculture, are recognized as career and technical education (CTE) career clusters and are considered to be practical subjects that apply some science and mathematics concepts. CTE courses (aka, vocational-oriented courses) are aligned with teaching high school students how to apply knowledge in real-world settings and develop 21st century skills, which could be argued that they are focused on preparing high school students for careers and the workforce. Because each STEM subject has its unique characteristics, K-12 science, mathematics, and CTE educators are overwhelmed about the changes and are unclear on how to navigate these challenges (Wang et al., 2020). In order to successfully integrate STEM subjects in K-12 classrooms, understanding educator's beliefs is the first step and at the heart of effective educational practices of STEM integration (Wang et al., 2020).

By synthesizing current related literature, we start this article by defining STEM integration and describing teachers' beliefs and practices. We provide a new definition of STEM integration that has a broader and development approach to integrated STEM learning after comparing STEM subjects in K-12 education. Next, we share examples of different levels of integrated STEM learning and how teachers' beliefs and practices STEM integration support the role of disciplinary ways of thinking and a broader and developmental approach to integrated STEM learning. At the end of the article, we shared a couple conclusions and identified some challenges teachers will likely navigate when using a broader and developmental approach to integrated STEM teaching and learning.

Definition of STEM integration

Integrated STEM has been ill-defined, and it has become a slogan that yields a myriad of operational definitions and implementation models (Bybee, 2013; English, 2016; NRC, 2014; Scherer et al., 2019). Developing a precise definition of integrated STEM education has been a challenge. Smith and Karr-Kidwell (2000) pointed out integrated STEM is “a holistic approach that links the disciplines, so the learning becomes connected, focused, meaningful, and relevant to learners” (p. 22). STEM integration could signify teaching science and mathematics by integrating engineering and technology into the K-12 curricula (Bryan et al., 2016; Bybee, 2010; Ramaley, 2007), and use real-world problems that relate to learners' everyday lives (Bybee, 2010; Moore et al., 2014; NRC, 2014).

In framing STEM integration, Herschbach (2011) argued there are two basic ways STEM curricula are organized: (1) correlated curriculum; and, (2) broad fields curriculum. Table 1 illustrates our understanding of Herschbach's description of these two curriculum models. The *correlated curriculum* represents a multidisciplinary integration approach where existing content areas retain their formal knowledge structure and teachers coordinate to illustrate interrelationships between or among related content areas. The level of adaptation is limited, and the existing structure of formal knowledge frames the curriculum design and instructional process. The *broad fields curriculum* is an interdisciplinary integration approach where related content areas are reorganized around an activity or sequence of activities. As such, individual subject matters lose their individual identities and the level of adaptation is much higher in coordinating instruction using a new framework that is activity or project-driven.

Moreover, Herschbach (2011) argued that integrating science and mathematics with vocational-oriented subjects (aka, CTE), such as technology, engineering, and agriculture, would help students have an enhanced understanding of knowledge, make learning more relevant, and be more college and career and workforce ready. For example, some scholars agreed that engineering design problems have the potential to help students practice problem solving by using science and/or mathematics knowledge through a systematic and iterative approach (Bryan et al., 2016; NRC, 2014). Table 2 summarizes our analysis of Herschbach's (2011) comparison of science and math to vocational-oriented subjects, such as technology, engineering and agriculture.

Table 1 Comparison of two basic STEM curricula.

Considerations	Correlated	Broad fields
Disciplinary Content areas	Multidisciplinary	Interdisciplinary
Organization of courses	Retains identity of each subject	Individual subjects may lose their separate identities
	Taught as separate subject matter courses	Cluster of related subjects are organized into a single area of study or sequence of related courses
Coordination and planning	Among different stand-alone content areas	Commonly done through a sequence of activities rather than structure of content
Challenges	High level of communication among teachers to reorganize existing subject matter to coordinate relationships among subjects	Clear organizing framework is needed to integrate related content areas into a single course, and structure inherent in existing content areas is lost
Level of adaptation	Very limited	Moderate or substantial
Instructional focus	Formal knowledge taught deductively	Applied knowledge taught inductively through activities (e.g., design process)
Organization of knowledge	Linear and sequential acquisition of knowledge	Non-linear and driven by solving problems

Table 2 Comparison of science and math to career and technical education subjects.

Considerations	Science & mathematics	Technology, engineering, & agriculture
Structure of knowledge	More formal structure “Formal knowledge”	Less formal structure “Applied knowledge”
Kind of knowledge	Conceptual & theoretical	Technical & contextual
Uses of knowledge	Formal knowledge & formal knowledge applied to specific activities	Formal knowledge applied to specific activities & knowledge specific to tasks
Knowledge selection	Non-selective; instruction organized around content	Selective; instruction organized around process and content is selected and applied around application
Purpose of knowledge	Understanding of intellectual content	Utilitarian; skill-training use
Problem-solving	Deductive, theoretical	Inductive, applied
Disciplinary	Single academic subjects of formal disciplinary knowledge	Interdisciplinary and integrated subjects that apply formal knowledge
Epistemological characteristics	S: Discovery and advance knowledge M: Analytical tool and analyses	T: Production and use of abstract and concrete artifacts E: Design of artifacts by selectively using knowledge from STM
Instruction	Independence (stand-alone)	Integrative

A common definition of integrated STEM teaching approaches that is accepted by many scholars is to mirror solving a real-world problem in a complex designed system, where students use knowledge and skills from multiple disciplines that relate to their everyday lives (Bryan et al., 2016; Bybee, 2010, 2013; English, 2016; NRC, 2014; Scherer et al., 2019). We argue that there needs to be a more disciplinary inclusive definition of integrated STEM with CTE career clusters to help bridge cross-disciplinary learning to increase the potential of reaching transdisciplinary learning. *Integrated STEM education is intentionally and purposively blending multiple disciplines (i.e., academic and vocational) to help students meaningfully learn and apply academic content through real-world problems framed in designed complex systems and grounded in career and technical contexts that facilitate multidisciplinary, interdisciplinary, or transdisciplinary learning for the development of life-long and workforce development connections and skills.*

Teachers' beliefs and practices

Educators' beliefs shape who they are as individuals and are influential in their decision-making, thinking, and classroom practices (Buehl and Beck, 2014; English, 2016; Margot and Kettler, 2019; Wang et al., 2020). Teachers' beliefs are a fundamental type of the through process or cognition, such as judgment of the truth or falsity of a proposition or assumption (Nespor, 1987; Pajares, 1992), that occur inside teachers' heads and are unobservable (Fang, 1996). Although teachers' beliefs, compared with their knowledge, are stronger predictors of their actions (Fang, 1996; Pajares, 1992), researchers discovered teachers' beliefs and their classroom practices are not always correlated. For example, teachers might hold beliefs in learner-centered teaching, but their instructions may be didactic in nature (Lim and Chai, 2008; Liu, 2011). Many factors, internal and external (Buehl and Beck, 2014), have impact on teachers' beliefs and influence their practices. These factors include administrators and colleagues' attitudes (Fang, 1996), content knowledge and abilities (Buehl and Alexander, 2006; Buehl and Beck, 2014), teacher preparation programs (Rice and Kitchel, 2017), personal interests and disciplinary identities (Ryu et al., 2019), how they were taught as a student (Balschweid et al., 2014), time and resources (Knobloch et al., 2007; Raymond, 1997), curriculum and standards (Buehl and Beck, 2014), years of teaching experiences (Lumpe et al., 2000), school and classroom environments, and school and education-related policies (Lumpe et al., 2000; Pajares, 1992). It is critical to understand the possible internal and external factors that may support or hinder the connections between beliefs and practices (Buehl and Beck, 2014).

Teachers' beliefs in STEM integration

When integrating STEM, often teachers tend to start with the to-be-learned STEM content knowledge, then, using engineering design challenges to engage students' STEM content learning by helping students see the connections between what they have learned with real-world applications (Crotty et al., 2017; Guzey et al., 2017; Wang et al., 2020). Teachers believed that engineering design activities with the hands-on aspect of design challenges increases students' engagement and motivation to learn STEM contents and practicing their problem-solving abilities (Kendall and Wendell, 2012; Park et al., 2017; Van Haneghan et al., 2015). Teachers also believed mastering STEM content knowledge that is relevant to the design challenges is essential before students can solve the problems. Although teachers believed that science, mathematics, and engineering are commonly used together in solving real-world problems (Wang et al., 2011), and they value integrated STEM teaching strategies, teachers feel low self-efficacy and discomfort to employ integrated STEM instruction, if they do not feel they have enough STEM content knowledge (Nadelson et al., 2013). Margot and Kettler (2019) analyzed 25 articles to picture a holistic view of teachers' perceptions of STEM integration. They found five pedagogical challenges: (1) curriculum challenges, (2) structural challenges, (3) concerns about students, (4) concerns about assessment, and (5) lack of teacher support might hinder teachers' willingness to adapt integrated STEM teaching in their classrooms. Ryu et al. (2019) found preservice teachers' competence to implement integrated STEM practices was "limited in terms of content knowledge, practices, nature of reasoning in disciplines other than their own as well as relations among STEM disciplines" (p. 13). Moreover, preservice teachers had limited and simplistic understandings of the disciplines (i.e., mathematics), sub-disciplines (i.e., science) and related disciplinary connections (i.e., agriculture). Limited understanding of disciplinary relations among STEM content areas limited the breadth and depth of STEM integration. Preservice teachers viewed STEM as a checklist approach to integration as they mentioned they needed to include all four content areas. Wang et al. (2020) conducted a case study with six high school STEM and agriculture teachers to explore their beliefs and practices through interdisciplinary collaboration. Their results echoed Buehl and Beck's findings. Several internal (e.g., teachers' personalities and attitudes to incorporate less familiar content) and external (e.g., administration supports) hindrances challenged teachers to implement STEM integration. The internal and external challenges also resulted in two interdisciplinary collaboration models, multi-classroom and extracurricular activity models. In conclusion, teachers' beliefs have impacts on how they perceive and practices STEM integration and interdisciplinary collaboration.

Disciplinary ways of thinking

Disciplines consist of organized knowledge, including core concepts and professional practices, which is supported by foundational beliefs and assumptions that represent a way of thinking in the field. In Table 3, we outlined how different disciplines have different ways of thinking. For example, the nature of science represents how scientists think and solve problems. Engineering and technological design represents how engineers and technologists think and solve problems. Applied and agroecosystem thinking

Table 3 Ways of thinking and problem-solving process of disciplinary.

<i>Disciplinary</i>	<i>Way of thinking</i>	<i>Problem-solving process</i>
Science	Nature of science (scientific method)	Scientific inquiry
Technology & engineering	Design process	Engineering design process
Mathematics	Nature of mathematics (science of order, logic, and representation; mathematizing and structuring)	Reasoning and representation (e.g., patterns, numbers, shapes, and modeling)
Agriculture, food & natural resources	Applied thinking & agroecosystem thinking	Pragmatic experiences (e.g., learning by doing, observations, and conversations) and systems thinking

represents how agriculturalists think and solve problems. Next, we explain these disciplinary ways of thinking and how they can play a role in integrated STEM education.

The STEM disciplines have unique epistemological characteristics that should be understood to maintain the integrity of each discipline and know how they each contribute to integrated learning (Herschbach, 2011). Disciplines accumulate and articulate sophisticated knowledge and understandings through scholarly endeavors (Rogers, 1999), and it is important to learn through different disciplines because students can learn “significant ways of knowing that have been built by thinkers over time and that operate in and shape our society today” (p. 35). These unique epistemological characteristics are considered “ways of thinking,” because this situates these epistemological characteristics within each disciplinary field, especially from a professional and expert’s perspective in the discipline and workforce, and it easier to communicate the concept of how scientists or engineers think. We summarized the ways of thinking for four disciplines in Table 3. A fourth discipline was included because agriculture, food, and nature resources (AFNR) brings an additional and unique way of thinking as a CTE subject area. Solving-problems is a unifying process across all disciplines, yet each discipline has unique epistemological characteristics that represent how a professional in each discipline would be informed by a way of thinking to solve a problem. Although each discipline has defined epistemological characteristics and ways of thinking, some similarities exist across the different disciplinary ways of thinking.

Nature of science is the way of thinking for scientists. Science involves asking questions about the world and then developing scientific investigations to answer the questions (Abd-El-Khalick, 2012). The nature of science is guided by empiricism, inference, creativity, theory, tentatively, and deductive/inductive/hypothetico-deductive inquiry (Abd-El-Khalick, 2012). Teachers use inquiry-based learning as a pedagogy to simulate the scientific inquiry process that scientists use to solve problems. It is important to note that the scientific inquiry process can be conducted in multiple ways to ask questions, collect and analyze evidence, and community understandings of the phenomenon, which also is known as scientific method (NRC, 1996).

Design process is the way of thinking for technologists and engineers (Moore et al., 2014). The argument could be made that engineers are more theoretical and design tools, whereas technologists are more applied and make and/or maintain tools. We see technology and engineering as being complementary in using the engineering design process. The engineering design process is an iterative process that involves formulating or identifying the engineering problem (including researching the problem and identifying constraints), planning a design solution (including brainstorming different solutions and evaluating competing solutions), and testing and evaluating a prototype or model (including data collection and analyses for feedback in redesign). This design process is repeated until a desirable design solution is created (Moore et al., 2014).

Nature of mathematics is the way of thinking for mathematics (AAAS, 1990). Mathematics is a science of order, logic, reasoning, and representation. Mathematicians use a process known as structuring to solve problems. Structuring can be accomplished through mathematizing and modeling. Mathematizing to structure real-world and mathematical problems known as creative structuring processes. Structuring can be notation, diagrams, definitions, analyses, or models to a class of problems that lead to solutions (Selden and Selden, 2005). Selden and Selden explained mathematizing is “a process of consideration and reflection that alternates between two forms—horizontal, during which the mathematics at hand becomes more familiar and is broadened, and vertical, during which new mathematics (notation, algorithms, definitions, etc.) is created” (p. 9). Next, modeling is another form of structuring, which may involve adding structure (e.g., variables, diagrams, equations, functions) to a familiar real-world or mathematical problem “with a view to converting the given problem into a more tractable one in a (or another) familiar mathematical domain” (Selden and Selden, 2005, p. 9). Finally, mathematicians use another form of complex thinking and mathematizing called validation or validating proofs. Teachers facilitate mathematizing for students using inquiry-oriented pedagogy to help students explain and justify their thinking (Selden and Selden, 2005). Mathematics can be used to express ideas or solve problems through analogies of things or natural phenomena abstractly, manipulating the abstractions to illustrate new relationships, and communicating the abstraction to communicate meaning about the thing or natural phenomenon (AAAS, 1990).

Applied thinking and agroecosystem thinking is the way of thinking for agriculturalists. Agriculture as a discipline was informed by pragmatic thought (Mars and Ball, 2016; Talbert et al., 2005), and applied learning in agriculture is rooted in experiential learning and authentic learning (Knobloch, 2003). In agriculture, a common thread was based on inductively solving problems in a real-world context based on four tenets: (1) John Dewey’s learning in context; (2) Seamen Knapp’s learning by doing; (3) Rufus Stimson’s learning through projects; and, (4) William Henry Lancelot’s learning by using knowledge through solving problems and explaining knowledge (Knobloch, 2003). Learning in agriculture has used an integrated model of experiential learning (Croom, 2008) that promotes application of academic subjects through an integrated curriculum (Talbert et al., 2005); (2) connects learning

to thinking and applying knowledge to careers and life-long learning (Knobloch, 2003); and, (3) holistic student development of 21st century skills to become contributing citizens in a global society. Ball and Knobloch (2005) shared, "The pedagogical roles of the agriculture teacher are complex and varied" (p. 55), and it is important that agriculture teachers understand the ill-structured nature of how to teach concepts in the agriculture and food system. Further, Mars and Ball (2016) found agriculturalists' ways of thinking were varied and informed by cultural and generational transmission, individual and community engagement to address local food system needs, and production efficiencies and economic sustainability.

As such, Knobloch (2003) stated that agricultural educators and the development of future agriculturalists was situated in the context of the social, economic, political, and agricultural systems. In the 1970s, system thinking became more prevalent among agricultural scientists and became known as agroecology. Agroecosystem thinking makes food systems decisions that are productive and efficient while being economical and cost-effective, environmentally sustainable, and socially responsible and acceptable to consumers (Aguna et al., 2005; Raven, 2020). Systems thinking considers a system as a whole and how elements of the system interact with each other (Bartlett, 2001; Kasser, 2018; Ponto and Linder, 2011). System thinking considers how different elements of the system affect each other and the intended and unintended consequences of making various decisions (Ponto and Linder, 2011). It is complex to understand how to integrate social, economic, and behavioral sciences into science and engineering (NRC, 2012). Agriculture, food and natural resources helps facilitate the integration of social, economic, and behavioral sciences because the production system is situated in the environmental system to serve consumers in exchange for economic and social value (Raven, 2020). Charoenmuang (2020) argued food systems education is an interdisciplinary education the "encourages reflective thinking and judgment by helping students use analysis, synthesis, and evaluation in problem-solving oriented approaches; and representing a complex view of knowledge to build the abilities to recognize, evaluate, choose among, and integrate multiple perspectives (Hilimire, 2016)" (p. 36–37).

These different ways of thinking can contribute and flavor the integrated learning experience for students. Interdisciplinarity and integrated STEM education is blending different disciplines together to create new meanings or new solutions to real-world problems (Raven, 2020). Because disciplines are defined by a corpus of knowledge and epistemic characteristics, ways of thinking provide a representation of different approaches to how professionals in specific disciplines solve problems. For example, a scientist would solve a problem using the nature of science, whereas an agriculturalist would solve a problem by inductively testing possible solutions based on feedback from the system components. The different approaches to solving problems provide teacher educators more common language to seeing how different disciplines see and approach solving problems differently. This helps facilitate solving complex problems by facilitating different disciplines to blend together and facilitate interdisciplinary learning experiences for students. Integrated learning is more about using a way of thinking that is representative of a discipline and how the ways of thinking inform the learning process when mixing content and context from different disciplines. As such, we provide five examples of how different disciplinary perspectives informed the development of integrated STEM education units.

Integrated STEM examples based on teacher's beliefs

The examples that we provided in this section were developed in two STEM teaching methods courses (and teacher professional development workshop modeled after one of the courses) at a large public university in the US. Predominately, the 3-credit graduate courses aimed to equip graduate students and upper level undergraduate students, as preservice educators, with the knowledge and skills to develop integrated STEM lessons (Wang and Knobloch, 2018; Ryu et al., 2019). As an example, we will describe the instructional process for one of the courses. From week one to week four, preservice educators experienced the nature of S, T, E, M as single disciplines. In addition, they explored their views of STEM integration, both individually and as groups by drawing concept maps. During week five and six, integrating AFNR into STEM was introduced to preservice educators. Different examples of STEM integration were shared in the course to generate discussion of the different characteristics of STEM integration through AFNR instruction. After sharing examples, preservice teachers started to develop their big idea of integrated STEM through AFNR. They were instructed that no one existing integrated model or teaching method is the best approach to teach integrated STEM through AFNR, which supported an openness to facilitate creative thinking using experiential learning strategies (Roberts et al., 2016). Preservice teachers had freedom to develop their STEM integrated lesson plans by using what they believed would be the best integrated model and practices for teaching STEM through AFNR. Then, from week 10 to week 13, preservice teachers conducted a microteaching of their integrated STEM through AFNR mini-units that they developed, and then delivered the mini-units through field teaching. At the end, preservice educators reflected on their teaching experiences to modify their mini-units. The final product of each mini-unit contained three to five lessons, and each lesson was about 45–60 min. These mini-units demonstrated the different structure of integrated STEM teaching when teachers hold different beliefs (Table 4). These examples showed that preservice educators tended to start teaching by being content-focused. Then, they transition to becoming learner-focused by shifting to activities (hands-on learning), problem-based learning, and/or inquiry-based learning.

Single CTE discipline

The *Where'd You Get That?! By-products of Animals in Agricultural Settings* mini-unit was developed by a non-formal educator, who was a licensed high school agriculture teacher. The mini-unit had five lessons that focused on introducing animal by-products (e.g., ice cream, bee wax, and wool) to fifth grade students. The overarching learning objective was "student understanding and knowledge through STEM-related activities by raising awareness of where by-products originate" Each lesson in the mini-unit followed the same instructional structure. A lesson started with reviewing what students had learned from the previous lesson. Then, the

Table 4 Integrated STEM examples.

<i>Disciplinary focus</i>	<i>Disciplinary process/thinking</i>	<i>STEM content knowledge</i>	<i>Contextual knowledge</i>
Single CTE discipline	Applied thinking	None/limited	Understand byproduct from animals, such as wool from sheep Design a hydroponics
Single S/M discipline	Science inquiry & scientific method	S: Force & motion M: Algebra and data interpretation	
Multidisciplinary	Applied thinking & science inquiry	S: Priorities M: Algebra	Understand and connect wool priorities and production
Interdisciplinary	Scientific inquiry, scientific method, & engineering design	S: Priorities, density, and buoyancy M: Algebra and data interpretation T/E: Engineering design	Design a sturdy chair that can pass tensile strength test.
Transdisciplinary	Nature of mathematical (modeling) & agroecosystem thinking	S: Food desert M: Function, logic, and modeling T/E: Google maps & food desert index	Design a food desert index that could be used to weight factors that might contribute to food distribution and economics problems in a local community

educators taught background knowledge about the by-products, such as how people sheared sheep to harvest wool, and how wool was spun into yarn. Then, students did hands-on activities, such as “shaving” cream off of a balloon using a craft stick, which the balloon coated with shaving cream represented a sheep’s wool. Then, the educator asked students about the purpose of shearing a sheep, and usage of wool at the end of the class. When asked how integrated STEM plays a role in the lessons, the educator, who developed the lessons answered, “In my opinion, I feel that STEM [integration] obviously ... like a tool that’s utilized to teach the content ...” As for the STEM content, the educator said, “I wanted them [students] to understand that everyday household items that they see actually come from animals. They [students] did not see these [STEM and AFNR] connections.” The educator’s perspective implied he wanted students to understand applied thinking using agriculture and food activities.

Single S/M discipline

The *hydroponics lab* mini-unit was developed for ninth grade integrated chemistry and physics students by a licensed high school science teacher. The mini-unit included five lessons that focused on conducting experiments to test different variables that might affect plant growth in a hydroponics system. For example, determining the power needed to pump water in a certain height was the learning objective of the *Power vs. Height Pumped Lab* lesson in the mini-unit. Each lesson of the *hydroponics lab mini-unit* followed the same instructional structure. In the beginning of the lesson, the educator proposed a problem, such as how high the water can be lifted by the pump? Then, students used steps of the scientific method to conduct scientific experiments. Students predicted an outcome, collected data, graphed the data, and interpreted data. At the end of each lesson, students reported and explained the relationship between the variables by using their data. The purpose of integrated STEM instruction was to teach science content to be used to solve a problem. The educator, who designed the mini-unit, said, “Students need to solve a real problem at the end of an integrated STEM lesson. So, they’re going to use science concepts whatever it is to solve problems.” As for the STEM content in the mini-unit, he indicated that “Students need to understand and tell me how high the pump can pump the water. So, I gave both variables [time/height]. By doing scientific experiments, they learned the concept of force and motion, and could apply that to modify the water pump in a hydroponics system.” The educator’s perspective was that he wanted the students to learn science content through a deductive approach of scientific inquiry. He wanted students to learn content by working through the scientific process.

Multidisciplinary

The *Sheep Byproducts: Properties of Wool* mini-unit was developed by a non-formal educator, who was a licensed high school agriculture teacher. It had three lessons to educate third graders about sheep and their byproducts from start to finish. In the first lesson, students explored how much wool is needed for everyday products. Students calculated how many sheep it takes to create a ball of wool. The lesson focused on adding and subtracting multi-digit whole numbers and recognizing the relationships between addition and multiplication. The second lesson focused on the properties of wool like its ability to retard fire, and why wool is a desirable fiber that used to design clothes for certain purposes. Lesson two guided students to practice scientific method through science inquiry. Lesson three asked students to use a simple loom to weave a small piece of fabric. Lessons attempted to build on one another but did not have strong connections. Each lesson focused on one STEM subject area, such as mathematics was the main concept that students learned and practiced in lesson one, and lesson two focused on conducting scientific experiments through science inquiry. Although the mini-unit attempted to incorporate engineering design in lesson three, students did not need to use the content knowledge that they learned from the previous lessons to wave a fabric by using a craft loom machine. The educator, who developed the mini-unit, acknowledged the mini-unit did not have every piece of STEM integrated at the same time. Purposeful STEM integration was what the educator believed when she developed her lessons. She indicated, “I don’t want the lessons to feel like they were doing math or science just for the sake of doing it, but for them [students] to understand that math and science are real-world experiences in agriculture that people in the past and today use.” When it comes to the role of STEM concepts, the

educator said, "I feel this is where I put the most effort into, making sure what they [students] were learning was transferable skills like the scientific process and mathematical equations." As such, the educator's perspective reflected she wanted students to learn mathematical representation and applied thinking in a complementary and relevant manner.

Interdisciplinary

The *Ultimate Building Material* mini-unit was developed by a non-formal educator, who worked as technical staff in a wood lab. The mini-unit had three lessons that focused using science concepts, such as buoyancy and density to determine properties of different wood species, and design and make quarter scale chairs. The mini-unit was designed for upper elementary and middle school levels. Students conducted several experiments in the first two lessons to calculate volume, and density of different wood species. For example, students made a prediction if a wood material will sink or float in water and used the data that they collected and analyzed to synthesize how density and buoyancy are related. The first two lessons helped students learn the relative strength of different wood species given their densities. Students followed scientific methods to conduct experiments in lesson one and two. In the last lesson, students used engineering design process, and what they learned about the density and strength of two different wood species in the first two lessons to design and make a chair. Students learned that there are often trade-offs between weight and strength when designing a chair. After students designed and built the chair, they break their chairs on a test apparatus. At the end of lesson three, students used mathematics and science knowledge and skills to analyze the test results and presented them in the class. Lessons built on one another and had strong connections. Students needed to apply what they used in the previous lessons to help them design their final products. The educator, who developed the mini-unit, believed that STEM integration is "using knowledge gained in math and science to create unique solutions to solve problems and for innovation." As for the STEM content, he said, "The lowest level [STEM integration] includes some math calculation, science fact, or using technology such as the internet. This helped students create an awareness of connections between a STEM discipline. The highest level of integration is where the STEM knowledge is applied to solve problems. This often includes designing something using math skills, science knowledge, and technology in a giving context." As such, the educator's perspective indicated he wanted students to learn the engineering design process and test wood product variables using scientific inquiry to solve the problem. Engineering design and scientific method were integrally interwoven and were simultaneously used to solve the design challenge.

Transdisciplinary

The lesson *Barriers to Food Access in Your Community* was the first lesson of the *Design a Greenhouse* mini-unit that focuses on building context for food insecurity in urban areas. The lesson was developed by a per-service technology education teacher. The lesson asked high school students to define "food desert" by weighing various factors, synthesizing statistics and information in order to draw conclusions to create a mathematical model (i.e., Food Desert Index) that can be used to explain their view of food desert. The educator first presented an existing definition of a food desert and asked students to identify the areas of the definition that seemed vague or unclear. Then, she proposed the question "what do food deserts look like in my community?" and asked students to investigate the problem. Students used Google Maps to investigate the number of grocery stores (aka, food markets) where they can buy fresh food in a community, and measure distances and time, such as by walking, taking public buses, driving, when people go to grocery stores. Students also use city data from the local government, such as average income, residents below the poverty line, percentage of residences with vehicles, as other critical factors that could contribute to food deserts in the community where students lived. Students weighed these factors by using a scale from 0 to 40 to describe the severity of a food desert in the community. At the end of the lesson, students needed to create a Food Desert Index equation that could be used to determine the food desert scores in their community, and also can be used to standardize food desert problems in other urban areas. The educator, who developed the lesson, believed critical thinking and systems thinking are the core STEM knowledge that she wanted her students to practice in this lesson. She said, "My view of integrated STEM is not about teaching specific content. Instead of building subject-specific content knowledge, I think integrated STEM should focus on helping students use critical thinking to solve problems," and "The problem needs to be meaningful to students. The boundary of each subject [STEM] is blurry, but each subject plays an important role to help students understand the problem and they are able to solve the problem by using [STEM] in a meaningful way." As such, the educator's perspective indicated she wanted students to learn transdisciplinary by solving a complex community-based problem (aka, food desert) using two disciplinary ways of thinking—mathematically modeling and systems thinking that were integrally interwoven and were simultaneously used to solve the complex problem.

Conclusions

Teachers' disciplinary ways of thinking

Teachers' disciplinary ways of thinking play an important role in integrated STEM learning. Teachers are informed by their previous experiences, funds of knowledge, disciplinary identities, interests, and professional training (Ryu et al., 2019). These sources shape teachers' beliefs and thinking (Pajares, 1992). Because beliefs play a critical role in teacher education to support both preservice and practicing teachers' professional development, such as forming personalized teaching philosophy, motivating student learning, and developing effective pedagogical approaches (Pajares, 1992; Turner et al., 2009), it is important to consider how teachers' beliefs and disciplinary ways of thinking will shape how they design and implement integrated STEM learning experiences. Teacher educators should recognize the role beliefs play in the acquisition of knowledge as well as encourage the reflection that brings beliefs to

the level of conscious awareness to help improve teachers' classroom practices (Caudle and Moran, 2012; Lumpe et al., 2000). As such, teachers' ways of thinking can help frame and focus integrated STEM curricula. For example, if the teacher uses their primary way of thinking to structure and facilitate a lesson, the learning experience will be more strongly anchored in a discipline and make connections with other disciplines. This experience is more likely to be a single discipline or multidisciplinary learning experience. In contrast, if a teacher uses multiple ways of thinking and purposefully structures and facilitates a lesson that engages students to use two or more ways of thinking in complementary or blended ways, then the lesson is likely to be an interdisciplinary or transdisciplinary lesson. The students' ways of thinking that are engaged and expressed in a learning experience is a criterion to determine the level of integrated learning.

A broader and developmental approach to STEM integration

Although content plays an important role in integrated STEM learning experiences, a more inclusive curricular and cross-disciplinary approach of academic and CTE subjects is needed to reach interdisciplinary and transdisciplinary levels of learning. It is important to consider several factors, not just the role of STEM content, to design and facilitate integrated STEM lessons, including: (1) the role of integration in learning objectives (2) the role (i.e., presence and usage); of STEM concepts, content knowledge, and skills; (3) role of learning outcomes; (4) role of the instructor and type of instruction; (5) role of CTE content knowledge (e.g., AFNR); and, (6) role of students' thinking (Wang and Knobloch, 2018). We emphasize the role of students' thinking should play an important part of planning and facilitating integrated STEM learning experiences. Teachers should design integrated STEM lessons to engage students to think in disciplinary ways, whether it be predominately from a single discipline way of thinking or from a blending of two complementary disciplinary ways of thinking. Helping students see how disciplinary thinking helps them solve problems would encourage them to be more creative, critical, and multi-dimensional in finding solutions. Moreover, intersectionality of disciplinary ways of knowing will help students understand the nature of reasoning in disciplines other than their own as well as relations among STEM and CTE disciplines (Ryu et al., 2019), which would help developing multidisciplinary, interdisciplinary and transdisciplinary understandings.

Cross-disciplinary learning experiences should not be superficially regarded as content from different disciplines (Ryu et al., 2019), but as informed by different disciplinary thinking and epistemic characteristics that are exemplified by experts and professionals who were trained in the respective disciplines. A single disciplinary or multidisciplinary approach is a good starting point for teachers who wish to begin and develop their skills to design and facilitate integrated STEM learning experiences. By doing so, students can explore how specific disciplines have a unique way of thinking and how they connect to other disciplines. We embrace that all disciplines can be a starting point for creating integrated STEM learning experiences, and it is not necessary for integrated STEM education to be inclusive of all S, T/E and M content areas. Developmentally, it may be better to transition multidisciplinary learning to interdisciplinary or transdisciplinary learning by focusing on two disciplines that have complementary ways of thinking or the ways of thinking can be blended to solve a complex real-world problem. Contextualizing integrated STEM learning using real-world problems would help students understand how to solve problems using different disciplinary perspectives. We believe that an academic subject and a CTE subject can be used in complementary and blended ways thinking to help facilitate interdisciplinary or transdisciplinary learning experiences for students to experience the academic richness and rigor that comes when solving real-world problems that are complex, authentic, and relevant to students (Knobloch, 2003). Interdisciplinary and transdisciplinary learning experiences may be desirable among teachers who are interested in designing and facilitating integrated STEM learning experiences. However, we argue a continuum of different levels of integrated STEM learning can be used to scaffold integrated learning experiences for both teachers and students to develop their understanding and experiences of meaningful and integrated STEM learning.

The role of students' thinking can be reflected as a continuum of three different levels of STEM integration (Wang and Knobloch, 2018). First, *Exploring STEM* (Level 1) is when a teacher exposes students to one disciplinary way of thinking, but may see the possibility of another way of knowing upon completion of solving a complex problem. For example, the single CTE discipline and single S/M discipline examples showed the educators only used either applied thinking in the mini-unit design by showing where animal by-products come from and how they are used by consumers (e.g., *Where'd You Get That?! By-products of Animals in Agricultural Settings* mini-unit). Next, the educator demonstrated the scientific method by having students test the variables of height and time using a water pump for a closed-loop hydroponics system (e.g., *Hydroponics Lab* mini-unit). Second, *Developing STEM* (Level 2) is when a teacher has a predominant way of thinking that drives the instruction or problem-solving process, but engages students to use a secondary disciplinary way of knowing to introduce students to additional tasks that were part of completing the integrated STEM project. For example, the multidisciplinary example demonstrated that students could use mathematics in a pragmatic way, which they calculated how much wool is needed to weave a sweater. However, there was a weak connection between knowing how many wools are needed to wave a sweater and how to conduct a scientific experiment to test the properties of wool in the mini-unit. Finally, *Advancing STEM* (Level 3) is when a teacher intentionally and purposefully teaches students two ways of thinking that complement each other in meaningful ways to holistically solve a complex problem. In doing so, students demonstrate higher-order thinking such as systems thinking, critical thinking, creative thinking, and/or complex problem-solving. For example, in the interdisciplinary example, students needed to understand the science concepts (density and priorities) and used engineering design process to design a sturdy chair. Students also needed to use scientific method to conduct experiments to test tensile strength of their chairs. As for the transdisciplinary example, in order for students to design a mathematical model (i.e., Food Desert Index) to standardize various factors that would contribute to the food desert problem in a community, students needed to rely on systems thinking to evaluate data and make evidence-based decisions to solve the complex problem.

Challenges to implementing a broader approach to integrated STEM learning

Integrated STEM education should be intentional and purposeful in blending multiple disciplines to help students meaningfully learn and apply academic content through real-world problems. These problems should be framed in designed complex systems and grounded in career and technical contexts that facilitate multidisciplinary, interdisciplinary, or transdisciplinary learning for the development of life-long and workforce development connections and skills. By implementing integrated STEM with CTE career clusters, students will be engaged to solve complex real-world problems and help bridge cross-disciplinary learning to increase the potential of reaching transdisciplinary learning.

We acknowledge that teachers and integrated STEM education specialists will face challenges when implementing this more holistic and broader STEM integration approach in the current K-12 education system and teacher educators, curriculum developers, school administrators, and professional development specialists can play important roles in helping teachers to navigate these challenges. First, K-12 teachers are traditionally trained to teach domain-specific content. Domain-specific teacher preparation programs impact teachers' beliefs (e.g., teaching philosophy) and classroom practices (Taconis et al., 2001; Wang et al., 2020). Teachers feel neither they have enough expertise to incorporate subject outsiders of their disciplinary (Baker et al., 2015; Graves et al., 2016), nor do they have abilities to align their domain-specific content with other subjects through the lens of interdisciplinary collaboration (Frykholm and Glasson, 2005). Consequently, how teachers trained in one of the STEM domains are not equipped to incorporate other STEM subjects into their teaching becomes a growing concern to promote integrated STEM instructions. Second, STEM integration needs a high level of on-going planning and coordination among different subjects (Herschbach, 2011; Capraro and Jones, 2013). However, inflexible class schedules, overloaded daily work, and stringent timelines for implementing curricula discourage the interdisciplinary nature of integrated STEM curricula and instructions (Lesseig et al., 2017; Wang et al., 2020). Finally, the STEM integration reform movement is widely recognized in science education (Bybee, 2010, 2013; NRC, 2014), but not in other subject areas. For instance, in the US, secondary science teachers are under stress to adapt their teaching from disciplinary silos to become more integrated styles, but not in other subject areas. Additionally, current secondary education, such as high schools, are structured in ways that continue to encourage teachers to stay in their teaching disciplinary silos (Boyd, 2017). The culture and structure of secondary education might hinder interdisciplinary collaboration to develop and implement integrated STEM curricula and instructions.

Summary

Integrated STEM learning has been growing and developing across many disciplines in the last 20 years. Because of this growth and interest among many disciplines, we believe the future of integrated STEM is exciting as educational professionals align the K-12 curricula and teacher professional development and education programs with complex real-world problems that drive what professionals in different disciplines are working on every day, and that students will work on in their future careers someday. By preparing students to learn how to solve complex problems in new interdisciplinary and transdisciplinary ways, we can hope there are new solutions to improve the way teachers teach, students learn, and professionals collectively solve complex challenges in the 21st century. Bringing academic and vocational disciplines together in education will model how interdisciplinary and integrated teams in the workforce solve complex real-world problems.

References

- Abd-El-Khalick, F., 2012. Examining the sources for our understandings about science: enduring conceptions and critical issues in research on nature of science in science education. *Int. J. Sci. Educ.* 34 (3), 353–374.
- Agunda, R., Connors, J.J., Chen, H.Y., 2005. A study of the Ohio State University College of food, agricultural, and environmental sciences' ecological-paradigm model. *Nacta J.* 2–10.
- American Association for the Advancement of Science, 1990. Science for All Americans: A Project 2061 Report on Literacy Goals in Science, Mathematics, and Technology. AAAS, Washington, DC. <http://www.project2061.org/publications/sfaa/online/sfaatoc.htm>.
- Baker, M.A., Bunch, J.C., Kelsey, K.D., 2015. An instrumental case study of effective science integration in a traditional agricultural education program. *J. Agric. Educ.* 56 (1), 221–236. <https://doi.org/10.5032/jae.2015.01221>.
- Ball, A.L., Knobloch, N.A., 2005. A document analysis of the pedagogical knowledge espoused in agriculture teaching methods courses. *J. Agric. Educ.* 46 (2), 47–57. <https://doi.org/10.5032/jae.2005.02047>.
- Balschweid, M., Knobloch, N.A., Hains, B.J., 2014. Teaching introductory life science courses in colleges of agriculture: faculty experiences. *J. Agric. Educ.* 55 (4), 162–175. <https://doi.org/10.5032/jae.2014.04162>.
- Bartlett, G., 2001. Systemic Thinking: A Simple Thinking Technique for Gaining Systemic (System-Wide) Focus. Paper presented at the International Conference on Thinking, Auckland, New Zealand. Retrieved from: http://www.probsolv.com/systemic_thinking/Systemic%20Thinking.pdf.
- Boyd, A.S., 2017. Social Justice Literacies in the English Classroom: Teaching Practice in Action. Teachers College Press, New York, NY.
- Bryan, L.A., Moore, T.J., Johnson, C.C., Roehrig, G.H., 2016. Integrated STEM education. In: Johnson, C.C., Peters-Burton, E.E., Moore, T.J. (Eds.), *STEM Road Map: A Framework for Integrated STEM Education*. Routledge, New York, pp. 23–37.
- Buehl, M.M., Alexander, P.A., 2006. Examining the dual nature of epistemological beliefs. *Int. J. Educ. Res.* 45, 28–42. <https://doi.org/10.1016/j.ijer.2006.08.007>.
- Buehl, M.M., Beck, J.S., 2014. The relationship between teacher's beliefs and teachers' practices. In: Fives, H., Gill, M.G. (Eds.), *International Handbook of Research on Teachers' Beliefs*. Routledge, New York, NY.
- Bybee, R., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teach.* 70 (1), 30–35.

- Bybee, R.W., 2013. A Case of STEM Education. NSTA press, Arlington, VA.
- Capraro, M.M., Jones, M., 2013. Interdisciplinary STEM project-based learning. In: Capraro, R.M., Capraro, M.M., Morgan, J.R. (Eds.), *STEM Project-Based Learning: An Integrated Science, Technology, Engineering, and Mathematics (STEM) Approach*. Sense Publishers, Rotterdam, The Netherlands, pp. 51–58.
- Caudle, L.A., Moran, M.J., 2012. Changes in understandings of three teachers' beliefs and practice across time: moving from teacher preparation to in-service teaching. *J. Early Child. Teach. Educ.* 33 (1), 38–53. <https://doi.org/10.1080/10901027.2011.650784>.
- Charoenmuang, M., 2020. High School Students' Systems Thinking in the Context of Sustainable Food Systems. Doctoral dissertation. Purdue University Graduate School.
- Croom, D.B., 2008. The development of the integrated three-component model of agricultural education. *J. Agric. Educ.* 49 (1), 110–120. <https://doi.org/10.5032/jae.2008.01110>.
- Crotty, E.A., Guzey, S.S., Roehrig, G.H., Glancy, A.W., Ring-Whalen, E.A., Moore, T.J., 2017. Approaches to integrating engineering in STEM units and student achievement gains. *J. Pre-Coll. Eng. Educ. Res.* 7 (2), 1–14.
- Dym, C.L., Agogino, A.M., Eris, O., Frey, D.D., Leifer, L.J., 2005. Engineering design thinking, teaching, and learning. *J. Eng. Educ.* 94 (1), 103–120.
- English, L.D., 2016. STEM education K-12: perspective on integration. *Int. J. STEM Educ.* 3 (1), 3. <https://doi.org/10.1186/s40594-016-0036-1>.
- Fang, Z., 1996. A review of research on teacher beliefs and practices. *Educ. Res.* 38 (1), 47–65. <https://doi.org/10.1080/0013188960380104>.
- Frykholm, J., Glasson, G., 2005. Connecting science and mathematics instruction: pedagogical content knowledge for teachers. *Sch. Sci. Math.* 105 (3), 127–141. <https://doi.org/10.1111/j.1949-8594.2005.tb18047.x>.
- Graves, L.A., Hughes, H., Balgopal, M.M., 2016. Teaching STEM through horticulture: implementing an edible plant curriculum at a STEM-centric elementary school. *J. Agric. Educ.* 57 (3), 192–207. <https://doi.org/10.5032/jae.2016.03192>.
- Guzey, S.S., Ring-Whalen, E.A., Harwell, M., Peralta, Y., 2017. Life STEM: a case study of life science learning through engineering design. *Int. J. Sci. Math.* <https://doi.org/10.1007/s10763-017-9860-0>.
- Herschbach, D.R., 2011. The STEM initiative: constraints and challenges. *J. STEM Teach. Educ.* 48 (1), 96–122. <https://doi.org/10.30707/JSTE48.1Herschbach>.
- Hilimire, K., 2016. Theory and practice of an interdisciplinary food systems curriculum. *Nacta J.* 60 (2), 227–233.
- Kasser, J.E., 2018. *Systems Thinker's Toolbox: Tools for Managing Complexity*. CRC Press.
- Kendall, A., Wendell, K.B., 2012. Understanding the Beliefs and Perceptions of Teachers Who Chose to Implement Engineering-Based Science Instruction. Paper presented at the American Society for Engineering Education Annual Conference and Exposition, San Antonio, Texas, June 2012.
- Knobloch, N.A., Ball, A.L., Allen, C.A., 2007. The benefits of teaching and learning about agriculture in elementary and junior high schools. *J. Agric. Educ.* 48 (3), 25–36. <https://doi.org/10.5032/jae.2007.03025>.
- Knobloch, N.A., 2003. Is experiential learning authentic? *J. Agric. Educ.* 44 (4), 22–34. <https://doi.org/10.5032/jae.2003.04022>.
- Lesseig, K., Elliott, R., Kazemi, E., Kelly-Petersen, M., Campbell, M., Mumme, J., Carroll, C., 2017. Leader noticing of facilitation in videocases of mathematics professional development. *J. Math. Teach. Educ.* 20 (6), 591–619. <https://doi.org/10.1007/s10857-016-9346-y>.
- Lim, C.P., Chai, C.S., 2008. Teachers' pedagogical beliefs and their planning and conduct of computer-mediated classroom lessons. *Br. J. Educ. Technol.* 39 (5), 807–828. <https://doi.org/10.1111/j.1467-8535.2007.00774.x>.
- Liu, S.-H., 2011. Factors related to pedagogical beliefs of teachers and technology integration. *Comput. Educ.* 56 (4), 1012–1022. <https://doi.org/10.1016/j.compedu.2010.12.001>.
- Lumpe, A.T., Haney, J.J., Czerniak, C.M., 2000. Assessing teachers' beliefs about their science teaching context. *J. Res. Sci. Teach.* 37 (3), 275–292. [https://doi.org/10.1002/\(sici\)1098-2736\(200003\)37:3<275::aid-tea4>3.0.co;2-2](https://doi.org/10.1002/(sici)1098-2736(200003)37:3<275::aid-tea4>3.0.co;2-2).
- Margot, K.C., Kettler, T., 2019. Teachers' perception of STEM integration and education: a systematic literature review. *Int. J. STEM Educ.* 6, 2. <https://doi.org/10.1186/s40594-018-0151-2>.
- Mars, M.M., Ball, A.L., 2016. Ways of knowing, sharing, and translating agricultural knowledge and perspectives: alternative epistemologies across non-formal and informal settings. *J. Agric. Educ.* 57 (1), 56–72. <https://doi.org/10.5032/jae.2016.01056>.
- Moore, T.J., Stohlmann, M.S., Wang, H.H., Tank, K.M., Glancy, A.W., Roehrig, G.H., 2014. Implementation and integration of engineering in K-12 STEM education. In: Purzer, S., Strobel, J., Cardella, M.E. (Eds.), *Engineering in Pre-college Settings: Synthesizing Research, Policy, and Practices*. Purdue University Press, Indiana: West Lafayette.
- Nadelson, L.S., Callahan, J., Pyke, P., Hay, A., Dance, M., Pfister, J., 2013. Teacher STEM perception and preparation: inquiry-based STEM professional development for elementary teachers. *J. Educ. Res.* 106 (2), 157–168. <https://doi.org/10.1080/00220671.2012.667014>.
- National Research Council, 1996. *National Science Education Standards*. National Academies Press.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Science and Education. The National Academies Press, Washington, DC.
- Nespor, J., 1987. The role of beliefs in the practice of teaching. *J. Curric. Stud.* 19 (4), 317–328. <https://doi.org/10.1080/0022027870190403>.
- NRC (National Research Council), 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. National Academies Press, Washington, DC.
- Pajares, M.F., 1992. Teachers' beliefs and educational research: cleaning up a messy construct. *Rev. Educ.* 62 (3), 307–332. <https://doi.org/10.2307/1170741>.
- Park, M.H., Dimitrov, D.M., Patterson, L.G., Park, D.Y., 2017. Early childhood teachers' beliefs about readiness for teaching science, technology, engineering, and mathematics. *J. Early Child. Res.* 15 (3), 275–291. <https://doi.org/10.1177/1476718X15614040>.
- Ponto, C.F., Linder, N.P., 2011. *Sustainable Tomorrow: A Teachers' Guidebook for Applying Systems Thinking to Environmental Education Curricula*. Retrieved from <https://www.fishwildlife.org/application/files/1715/1373/1187/ConEd-Sustainable-Tomorrow-Systems-Thinking-Guidebook.pdf>.
- Ramaley, J.A., 2007. *Facilitating Change: Experience with the Reform of STEM Education*. Retrieved from <http://www.wmich.edu/science/facilitating-change/Products/RamaleyPresentation.pdf>.
- Raven, M.R., 2020. Regenerative agriculture and implications for agriculture, food, and natural resources education. *J. Agric. Educ.* 61 (1), 1–12. <https://doi.org/10.5032/jae.2020.01001>.
- Raymond, A.M., 1997. Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *J. Res. Math. Educ.* 28 (5), 550–576. <https://doi.org/10.2307/749691>.
- Repko, A.F., 2008. Assessing interdisciplinary learning outcomes. *Acad. Exchange Q.* 12 (3), 171–178.
- Rice, A.H., Kitchel, T., 2017. Agriculture teachers' integrated belief systems and its influence on their pedagogical content knowledge. *J. Agric. Educ.* 59 (1), 51–69. <https://doi.org/10.5032/jae.2018.01059>.
- Roberts, R., Baker, M.A., Goossen, C.E., 2016. The chasm between beliefs and practice: a case study of the epistemological positions of pre-service agricultural education teachers. *J. Agric. Educ.* 57 (2), 172–186. <https://doi.org/10.5032/jae.2016.02172>.
- Rogers, B., 1999. Conflicting approaches to curriculum: recognizing how fundamental beliefs can sustain or sabotage school reform. *Peabody J. Educ.* 74 (1), 29–67.
- Ryu, M., Mentzer, N., Knobloch, N., 2019. Preservice teachers' experiences of STEM integration: challenges and implications for integrated STEM teacher preparation. *Int. J. Technol. Des. Educ.* 29 (3), 493–512. <https://doi.org/10.1007/s10798-018-9440-9>.
- Scherer, H.H., McKim, A.J., Wang, H.H., DiBenedetto, C.A., Robinson, K., 2019. Making sense of the buzz: providing a taxonomy of "STEM" in agriculture, food, and natural resources education. *J. Agric. Educ.* 60 (2), 28–53. <https://doi.org/10.5032/jae.2019.02028>.
- Selden, A., Selden, J., 2005. Perspectives on advance mathematical thinking. *Math. Think. Learn.* 7 (1), 1–13. https://doi.org/10.1207/s15327833mtl0701_1.
- Smith, J., Karr-Kidwell, P., 2000. *The Interdisciplinary Curriculum: A Literary Review and a Manual for Administrators and Teachers*. Retrieved from <https://eric.ed.gov/?id=ED443172>.

- Taconis, R., Ferguson-Hessler, M., Broekkamp, H., 2001. Teaching science problem solving: an overview of experimental work. *J. Res. Sci. Teach.* 38 (4), 442–468. <https://doi.org/10.1002/tea.1013>.
- Talbert, B.A., Vaughn, R.C., Croom, D.B., 2005. *Foundations of Agricultural Education*. Professional Educators Publications.
- Tripp, S., Helwig, R., Yetter, D., 2017. *The Importance of Research Universities: With Examples of Their Functional Role and Impacts Within the State of Indiana*. Biocrossroads, Indianapolis, IN.
- Turner, J.C., Christensen, A., Meyer, D.K., 2009. Teachers' beliefs about student learning and motivation. In: Saha, L.J., Dworkin, A.G. (Eds.), *International Handbook of Research on Teachers and Teaching*, Springer International Handbooks of Education, vol. 21. Springer, Boston, MA.
- United States Department of Agriculture (USDA), 2015. Employment Opportunities for College Graduates in Food, Agriculture, Renewable Natural Resources, and the Environment. Retrieved from: <https://www.purdue.edu/usda/employment/wp-content/uploads/2015/04/2-Page-USDA-Employ.pdf?F>.
- Van Haneghan, J.P., Pruet, S.A., Neal-Waltman, R., Harlan, J.M., 2015. Teacher beliefs about motivating and teaching students to carry out engineering design challenges: some initial data. *J. Pre-Coll. Eng. Educ. Res.* 5 (2). <https://doi.org/10.7771/2157-9288.1097>. Article 1.
- Wang, H.H., Knobloch, N.A., 2018. Levels of STEM integration through agriculture, food, and natural resources. *J. Agric. Educ.* 59 (3), 258–277. <https://doi.org/10.5032/jae.2018.03258>.
- Wang, H.H., Moore, T., Roehrig, G.H., Park, M.S., 2011. STEM integration: teacher perception and practice. *J. Pre-Coll. Eng. Educ. Res.* 1 (2), 1–13. <https://doi.org/10.5703/1288284314636>.
- Wang, H.H., Charoenmuang, M., Knobloch, N.A., Tormoehlen, R.L., 2020. Defining interdisciplinary collaboration at high school settings through teachers' beliefs and practices of STEM integration by using a complex designed system. *Int. J. STEM Educ.* 7 (3). <https://doi.org/10.1186/s40594-019-0201-4>.

Promises and perils of STEM education: synthesizing teacher, student, & research perceptions

Thomas Roberts^a and Audrey Conway Roberts^b, ^aSchool of Teaching and Learning, College of Education and Human Development, Bowling Green State University, Bowling Green, OH, United States; and ^bSchool of Educational Foundations, Leadership, and Policy, College of Education and Human Development, Bowling Green State University, Bowling Green, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	262
Teacher and student perceptions of STEM	262
Teachers' perceptions of STEM	263
Students' perceptions of STEM	263
Synthesis of students' and teachers' perceptions	265
Methodological perceptions, considerations, and approaches in STEM education	265
Operationalizing STEM	265
Integration of STEM	266
Development of STEM integration	266
Implications and future directions	266
References	267

Introduction

Science, technology, engineering, and mathematics (STEM) education has increased in popularity ([Moore et al., 2020](#); [Ehlert and Roberts, 2021](#)) around the world. Research and policy positions focused on increasing access to STEM education exist in many countries including the United States (e.g., [Mohr-Schroeder et al., 2014](#); [National Research Council, 2014](#)), the United Kingdom ([Laconte, 2020](#)), Ethiopia (e.g., [Melak and Singh, 2021](#)), South Korea (e.g., [Kang, 2019](#); [Korean Ministry of Education Science and Technology, 2011](#)), Australia (e.g., [Office of the Chief Scientist, 2014](#)), Indonesia (e.g., [Arivina and Jailani, 2020](#)), and Pakistan (e.g., [Ramzan et al., 2021](#)). STEM education has been positioned as a solution for societal and economic needs. For example, some argue STEM education is necessary for people to function in a rapidly changing world increasingly driven by science, technology, engineering, and/or mathematics ([International Technology and Engineering Educators Association \[ITEEA\], 2020](#); [Mohr-Schroeder et al., 2020](#)). Others argue that STEM education should be emphasized to meet national security and workforce needs ([National Science Board, 2015](#); [Maiorca et al., 2021](#)). General fears about falling behind on international test scores also drive an emphasis on STEM education ([Fleischman et al., 2010](#)). These competing rationales for STEM education preview the diverse perceptions found in the current literature.

[Moore et al. \(2020\)](#) synthesized conceptual frameworks and definitions of STEM education. Ultimately, four common themes were identified as most common in STEM education research: STEM education is based on real-world problem solving, individual STEM disciplines share large conceptual ideas and practices, integration strategies vary, and integration relies on structures based in individual disciplines and/or pedagogies ([Moore et al., 2020](#)). Even with these common themes, STEM education research has significant variation in theoretical frameworks and how these practices and policies are enacted. Moreover, the way these STEM practices and policies are enacted often exclude the needs of historically excluded learners ([Berry, 2021](#)). Given this landscape in STEM education, this article has three purposes: (1) synthesize the literature around teacher and student perceptions on STEM education; (2) share perceptions about STEM education research methodologies; and (3) based on the synthesis of these diverse perceptions on STEM education, share implications and recommendations for future research directions.

Teacher and student perceptions of STEM

The emphasis on STEM education directly impacts classrooms as countries, states, districts, schools, and teachers begin to implement STEM learning experiences. The success of these approaches depends upon the success of teachers. Successfully teaching STEM not only requires teacher knowledge ([Nesmith and Cooper, 2020](#)) and affective characteristics, such as dispositions, attitudes, and beliefs ([Thompson, 1992](#); [Philipp, 2007](#)), but also relies on teachers' prior experiences ([Margot and Kettler, 2019](#)). Therefore, it is imperative to understand teachers' perceptions about STEM education.

Teachers' perceptions of STEM

Teachers' perceptions of STEM offer both hope and insight into needed structural changes to effectively implement STEM education in formal settings. Overall, teachers perceive STEM education to be important for their students (Bruce-Davis et al., 2014; Margot and Kettler, 2019; Park et al., 2017; Nesmith and Cooper, 2020). Similar to their students, teachers perceive STEM education to be a great way to apply discipline specific content (Asghar et al., 2012). Teachers also perceive STEM education to be relevant to students' future careers in that real-world problem solving is often interdisciplinary (Bruce-Davis et al., 2014). Within the classroom, teachers believe STEM education can positively impact student learning and learning outcomes (Park et al., 2017) and increase student engagement, interest, and confidence (Dare et al., 2014; Herro and Quigley, 2017; Stubbs and Myers, 2016; Wang et al., 2011). There is also a belief that STEM helps students learn how to deal with failure and how to learn from mistakes (Dare et al., 2014; Roberts and Schnepf, 2020). With such positive perceptions of STEM education, the question must be asked: Why is STEM education valued but not done well in formal school settings? Teachers' perceptions to STEM education suggest the answer lies in structural dysfunctions within the educational system.

One major barrier to effective STEM education in formal school settings is a lack of teacher knowledge (McMullin and Reeve, 2014) from a lack of teacher preparation (Nesmith and Cooper, 2020; Radloff and Guzey, 2016; Stohlmann et al., 2011). Teachers may not know or recognize effective strategies to integrate STEM disciplines (Shernoff et al., 2017) or to implement integrated programs (Nadelson and Seifert, 2013). Other times this focuses on content knowledge and teacher training. For example, typically elementary teachers are not trained for STEM integration in their teacher education programs (Nesmith and Cooper, 2020). Likewise, secondary teachers are often trained in single disciplines (Stohlmann et al., 2011). The lack of effective preparation also leads to teachers being concerned about the role of each discipline in integrated STEM learning experiences (Herro and Quigley, 2017). With this lack of preparation, teachers cannot be expected to effectively implement STEM learning experiences. Moreover, because teachers have less knowledge in STEM, they are less likely to teach it (Nadelson and Seifert, 2013).

In addition to the lack of preparation, teachers also perceive dramatic shifts in classroom functioning to implement STEM learning experiences. For example, teachers regularly highlight the lack of coherent curricula in STEM (Ehlert and Roberts, 2021; Margot and Kettler, 2019; Wang et al., 2011). Without having curricula, teachers need time to collaborate and create intentional STEM learning experiences for their students. Unfortunately, teachers often are not given the time and cite the lack of collaboration and increased amount of time to prepare and implement STEM as barriers to doing more STEM in their classrooms (Margot and Kettler, 2019; Al Salami et al., 2017; Park et al., 2017). Providing STEM experiences necessitates a shift from teacher-led instruction to student-led learning which can be difficult for teachers to implement (Park et al., 2017). Finally, teachers perceive STEM assessments to be lacking in formative assessments (Asghar et al., 2012), in their ability to assess individuals' knowledge of content (Herro and Quigley, 2017), and in alignment to standards and accountability measures such as standardized tests (Nadelson and Seifert, 2013; Ehlert and Roberts, 2021). Although teachers readily identify benefits of STEM, teachers are often evaluated on their instruction and on student learning. If teachers do not perceive appropriate curricular, pedagogical, and assessment tools are available for any subject area, it is logical that teachers would not actively choose to develop and implement these items when they have already identified lack of planning and collaboration time as a barrier to implementation (Margot and Kettler, 2019).

The ways in which schools are structured and teachers are trained creates a system that is not conducive to STEM learning experiences. Student schedules are not designed for cross-curricular or integrated learning time (Margot and Kettler, 2019). Because they are trained in individual disciplines and evaluated on student performance in individual disciplines, different content teachers have different foci when developing STEM learning experiences (Wang et al., 2011). Even then, teachers underestimate students' abilities to solve STEM problems (Asghar et al., 2012) and often think students need specific disciplinary knowledge to participate in integrated STEM learning activities (Wang et al., 2011). Teachers are placed in an impossible position where society emphasizes STEM, teachers see the value of STEM (Margot and Kettler, 2019; Park et al., 2017), but the system is not designed to support teachers or students to do STEM in formal school settings. Expecting teachers to provide high-quality STEM learning experiences without appropriate time and support is irrational. Moreover, as STEM education is still relatively new, people have limited prior experiences with STEM education, particularly in formal school settings.

Research (e.g., Densen et al., 2015; Roberts et al., 2018; Kwon et al., 2021; Maiorca et al., 2021) that noted the structural constraints placed on formal learning environments do not allow sufficient time, integration, and/or access to expertise to develop students' interest in authentic STEM fields. Gupta et al. (2020) studied how the United States public perceived their interactions with STEM. Most participants identified informal learning spaces such as science centers, zoos, and aquariums as places where they encountered STEM, suggesting many STEM education interactions occurring within the informal learning space (Gupta et al., 2020). Informal learning environments promote learning through real-world modeling and authentic examples of STEM work (Nugent et al., 2010; Popovic and Lederman, 2015; Roberts et al., 2018). They occur outside of formal school environments and are often less structured and driven by student interest (NRC, 2009). Given the prevalence of interactions with STEM outside of school settings and the lack of consistent STEM offerings in formal school settings (Roberts et al., 2018), much of the research on students' perceptions of STEM occurs in informal learning environments.

Students' perceptions of STEM

Research on students' perceptions of STEM primarily focus on informal learning opportunities in the form of summer camps or after school programs. This body of research does not consistently define what is meant by perceptions or what is meant by STEM. Often,

constructs like attitudes, beliefs, and motivation are incorporated with the focus on perceptions. Similarly, STEM is not clearly defined with conceptualizations ranging from a math and/or science focused camp to an integrated STEM learning experience. With these limitations acknowledged, four primary themes are evident in research around students' perceptions of STEM learning experiences: (1) STEM learning experiences help students make real-world connections for applications of content; (2) STEM learning experiences often are grounded in the real world with career connections; (3) STEM learning experiences are typically hands-on and practice focused; and (4) STEM learning experiences provide students with access and opportunities to authentic STEM learning.

Students often perceive STEM learning experiences as places where real-world connections and applications of content are provided. An emphasis on real-world learning is central to STEM education broadly (Moore et al., 2020) and students express the benefits of having the real-world connections because it allows them to see how content is applied (Roberts et al., 2018; Densen et al., 2015; Jackson et al., 2021a,b). This is particularly true for students seeing applications of mathematics and science (Densen et al., 2015; Roberts et al., 2018). Robotics is one common way students make connections to mathematics content (Coxon et al., 2018; Wang and Frye, 2019), particularly around ideas of logic, measurement, and decimals (Roberts et al., 2019). Students also made connections to specific topics they learned in school, such as learning about DNA through a DNA extraction activity (Roberts et al., 2018), understanding how the scientific method was applied when designing a prosthetic leg (Bush et al., 2020), and learning about anatomy by dissecting a pig instead of reading about anatomy in a textbook (Jackson et al., 2021a,b). By connecting activities to real-world STEM practices, students see the applicability of the content (Jackson et al., 2021a,b; Christensen et al., 2014; Densen et al., 2015). Importantly, a focus on applied content and curricula is important to the success of historically excluded learners in STEM (Jong et al., 2020). Informal STEM learning environments provide ideal settings for these types of learning experiences for historically excluded learners (Calabrese Barton and Tan, 2018) as formal schooling often reinforces structures (e.g., tracking, siloed-instruction) that negatively impact historically excluded learners (Berry, 2021; Crisp et al., 2009; Martin et al., 2010).

Students can make real-world connections through informal STEM learning experiences that emphasize STEM careers and mentoring. Many informal STEM learning experiences connect the content to the real world by partnering with STEM professionals (e.g., Mohr-Schroeder, 2015). This provides students with authentic learning experiences (Roberts and Chapman, 2017) and increases interest in STEM careers (Maiorca et al., 2021). Informal STEM learning is particularly effective at this because students can see how content is integrated, which they often do not experience in formal schooling where subjects are frequently instructed as siloed-disciplines (Maiorca et al., 2021). Students who participated in informal STEM learning experiences grounded in real-world applications of STEM were 1.8 times more likely to pursue a STEM career than other students (Kitchen et al., 2018). Informal STEM learning opportunities can also increase the chances of students interacting with diverse STEM professionals. This is imperative as non-White and non-male persons are often invisible in STEM professions, which makes it difficult for historically excluded learners to see themselves as members of the STEM community (Jong et al., 2020; Morton and Smith-Mutegi, 2018). Having access to diverse STEM role models helps students see themselves as a member of the STEM community (Coxon et al., 2018) and improve self-perception as a member of STEM (Benavent et al., 2020).

Students also perceive STEM learning experiences to be hands-on and engaged in the "doing" of STEM. Hands-on learning experiences made learning engaging and fun to students (Anand and Dogan, 2021; Bush et al., 2020; Densen et al., 2015; Kwon et al., 2021). Students perceived the informal STEM learning experiences as more engaging because of the constraints they and their teachers have in the formal school setting (Roberts et al., 2018). The way students characterized the engaging, hands-on STEM learning experiences emphasized disciplinary practices (e.g., Standards for Mathematical Practice, Science and Engineering Practices, Technology and Engineering Practices) and doing and making (International Technology and Engineering Educators Association [ITEEA], 2020) as students noted the importance of trying ideas, learning from mistakes, and trying again (Roberts and Schnepf, 2020). Students noted the importance of collaborating (ITEEA, 2020) in completing hands-on projects and 3-D printing (Anand and Dogan, 2021). When doing robotics, students noted the importance of being precise (NGA Center for Best Practices and CCSO, 2010) in their programming and the mathematical implications of their trial and error (NGSS Lead States, 2013). Thus, as students perceived the importance of real-world applications of STEM content knowledge, they also perceived important habits of mind and dispositions characterized by disciplinary practice standards.

Informal STEM learning experiences frequently provide students with access and opportunities that they otherwise would not have had. Students typically perceived informal STEM learning environments to be more interesting than formal school environments (Roberts et al., 2018; Densen et al., 2015). For example, students noted the lack of technology in their schools, lack of access to rigorous core content, and lack of access to integrated STEM courses or electives (Roberts et al., 2018). The implication from this finding is the need for each and every student to have access to quality STEM learning experiences because the way in which students have the opportunity to interact with STEM content impacts their desire to persevere in STEM subjects (Jackson et al., 2021a,b). In the United States, many students decide by the eighth grade that they do not want to pursue more advanced STEM (The President's Council of Advisors on Science and Technology [PCAST], 2010) which implies they perceive STEM as not for them because of how the content has been presented to them (Bell et al., 2018). This has long term implications as access to advanced STEM courses in high school is a primary factor in students' decisions to major in a STEM discipline (Jong et al., 2020). Parents also perceive this structural problem with formal schooling in the United States as it is a reason they enroll their children in informal STEM learning opportunities (Anand and Dogan, 2021).

Students typically have positive perceptions of informal STEM learning experiences and recognize the limitations of formal STEM learning experiences in school settings. Parents recognize it, too (Anand and Dogan, 2021). The general public admit that their interactions with STEM happen primarily in informal learning spaces (Gupta et al., 2020). Even so, there has been significant

work within the United States to increase STEM education in formal school settings. The Early Childhood STEM Working Group (2017) called for earlier STEM experiences for students but acknowledged the lack of attention early STEM education received. Pressures related to standardized testing pose an obstacle to integrated STEM (Shernoff et al., 2017; Sikma and Osborne, 2014). STEM specific schools have failed because of systemic constraints that prevent them from being successful (Weis et al., 2015). In addition to the systemic issues, typically in the form of district mandates and federal and state accountability measures, some teachers may not feel comfortable or prepared to teach STEM in an integrated way (Shernoff et al., 2017). It is important to understand why, given the emphasis on STEM education, structures have not been enacted to achieve this goal.

Synthesis of students' and teachers' perceptions

Students and teachers both have great hope for the promise of STEM education. Students see STEM learning opportunities as engaging ways to explore disciplinary content ideas applied to real-world problems (Roberts et al., 2018, 2019; Densen et al., 2015; Jackson et al., 2021a,b). Teachers also recognize the benefits of STEM for connecting content to the real-world and to increasing student engagement and motivation (Dare et al., 2014; Herro and Quigley, 2017; Wang et al., 2011). Unfortunately, systemic barriers at many levels prevent widespread implementation of high quality STEM education. In South Korea, STEM is less popular than other curriculum acronyms such as STEAM (STEM infused with Art) or STREAM (STEM plus Art, Robotics or Reading) as students express high levels of performance in science and mathematics, but very low interest in these activities (Korea Institute of Curriculum and Evaluation, 2014).

In many Latin American countries, Bascopé et al. (2020) suggest there are positive steps toward incorporation of more integrated learning, however, STEM education as a concept is not explicitly referred to as such at the policy level. Students need access and opportunity to high quality STEM education learning experiences in both informal and formal settings. In formal settings, teachers need to be prepared to teach STEM (Nesmith and Cooper, 2020; Park et al., 2017) and have the curriculum and support (e.g., collaborative planning time, flexible student scheduling, adequate assessments) to teach STEM (Margot and Kettler, 2019). Both bodies of literature suffer from a lack of operationalized constructs. STEM is difficult to implement and measure because it is poorly and inconsistently defined (Moore et al., 2020; Pringle et al., 2020). Therefore, it is important to consider basic methodological perceptions of STEM education.

Methodological perceptions, considerations, and approaches in STEM education

Operationalizing STEM

Currently, there is no agreed upon definition for STEM education (English, 2016; Moore et al., 2020; Pringle et al., 2020). For example, Stohlmann et al. (2012) emphasized integrated STEM uses the engineering design process to create a real-world context for learning content within discipline specific courses. Others have suggested integrated STEM engages students in big ideas and skills for at least two individual STEM disciplines through real-world contexts (Roehrig et al., 2012). Roberts et al. (2022) defined integrated STEM as "using multiple STEM disciplines to collaboratively solve real-world problems with the ultimate goal of increasing STEM literacy" (p. 2). The last definition is one of the few that not only defines what STEM education is but what the goal of STEM education should be. Not only do these competing definitions exist, there are even different names (e.g., integrated STEM, STEM education, integrative STEM) and related fields (e.g., STEAM and STREAM). Part of this could be because of misalignment between STEM initiatives and the siloed nature of educational standards (Bybee, 2018). Therefore, STEM research is inherently complex.

Compounding the problems associated with a lack of clear definitions, STEM education also has competing disciplines and epistemological considerations that are foundational to the methods of research undertaken. Even at the federal level in the United States, different agencies define what is under the umbrella term of STEM, adding further ambiguity (Gonzalez and Kuenzi, 2012). Educational researchers in STEM are influenced by the "traditions, languages, and cultures ... [that] define and limit how they understand and explore this world ... [to] draw upon a variety of theoretical frameworks" (Bussey et al., 2020, p. 52). After a critical review of available discipline based educational research in mathematics, engineering, physics, biology, and chemical education, Bussey et al. (2020) concluded that each discipline and related fields of study within disciplines prioritizes different theoretical frameworks, which could prove to be a rich source of interdisciplinary growth and cohesion, or negatively impact the potential for true integration of STEM concepts. Due to the lack of an operationalized definition and varied disciplinary, interdisciplinary, and transdisciplinary approaches to STEM education research, there are significant gaps in foundational methodological considerations in STEM education research (Li et al., 2020).

There is not one preferred methodological paradigm in STEM education research, rather there are diverse study designs, data collection, and analysis techniques used. Li et al. (2020) in a critical review of the state of STEM education reviewed approximately 800 articles published between 2000 and 2018 in 36 academic journals to understand trends in STEM scholarship. The authors found a dramatic increase in published STEM research beginning in 2015, in research encompassing multiple disciplines concurrently of STEM, with global contributions (e.g., USA, Canada, Australia, UK, Israel, Turkey), and a focus on multi-author publications. When considering methodological approaches, a quantitative focus was the most predominant, however, qualitative, mixed methods, and non-empirical studies were well-represented (Li et al., 2020). The rapid evolution of and interest in STEM education research being published in academic journals over the last two decades suggest that the field will continue to shift to meet changing needs.

Integration of STEM

For many, STEM represents any curriculum, instruction, and/or activities pertaining to any of the four disciplines (English, 2016; Li et al., 2020). Developing a culturally relevant geometry curriculum, primary school children building wind turbines, or research on the perceptions of adolescents' robotics designs could all fall under the umbrella term of STEM education. In addition, Krapp (2002) and Krapp and Prenzel (2011) noted that STEM interest is not always domain specific. Rather, students may have a stronger interest in a particular topic within domains but report less interest in the domain itself (e.g., computer programming but not technology). This practice of mono-discipline instruction in STEM has carried over into the instruments used to assess STEM perceptions, attitudes, interests, and content knowledge from both teachers and students, with existing instruments for measuring STEM perceptions commonly focused on collectively assessing individual attitudes, perceptions, or knowledge individually (Thibaut et al., 2018; Roller et al., 2020).

Yet, there is a growing perception in the research and practitioner field that STEM should be treated as an integrated or interdisciplinary model that unites all four fields into something unique (Honey et al., 2014; Holmlund et al., 2018; Gao et al., 2020). Indeed, education standards like the Next Generation Science Standards (NGSS) in the USA, Australia (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015), and the United Kingdom (UK Department for Education, 2015) suggest an interdisciplinary instructional approach. However in practice, and as evidenced by research, STEM disciplines frequently remain siloed in their respective content areas present challenges when designing appropriate validated and reliable survey instruments (Staus et al., 2020).

Development of STEM integration

The development of STEM survey instruments is ongoing, as the field of STEM Education as an integrated discipline has risen to prominence in the past decade (Thibaut et al., 2018). As researchers apply a variety of theoretical frameworks to understand the perceptions, attitudes, etc. of students and teachers, using qualitative, quantitative, and mixed paradigms, there is space for contributions to the field. Integrated or interdisciplinary STEM approaches have been under-researched, with gaps in understanding instruction, professional development and training, and student outcomes (Griese et al., 2015; Herro et al., 2017; You et al., 2018). Therefore, it becomes necessary from an assessment perspective to be able to measure interest, attitudes, and knowledge in the four elements comprising STEM, while using precise contexts and content area terms, as well as integrated STEM as its own "meta-discipline" to understand how everything works together (Morrison, 2006).

As there is no consensus with an operationalization of integrated or interdisciplinary STEM education, and a documented pattern of assessments that are mono-discipline specific (Gao et al., 2020), there is a critical need for the creation of valid and reliable assessments that encompass the integration of science, technology, engineering, and mathematics. Some of this work has begun, however, the lack of consensus for an inclusive theoretical framework prevents any one singular instrument from being appropriate. Gao et al. (2020) conducted a literature review of available research of integrated STEM educational programs for Grade 6-university students, finding approximately 49 appropriate studies, 40 of which included science with engineering as the dominant disciplines. The authors concluded that although all of the studies were presented as using an integrated STEM approach, the assessments used to measure students' content knowledge were typically specific to a singular discipline. Moreover, the assessments used were chosen or adapted from existing instruments. Less than 15% of the articles chosen included assessments that incorporated an interdisciplinary approach to content knowledge (Gao et al., 2020).

Implications and future directions

STEM education is positioned as important for workforce, national security, and citizens' needs to function in a world increasingly driven by STEM (ITEEA, 2020; Mohr-Schroeder et al., 2020; Maiorca et al., 2021). In education systems, teachers and students perceive significant benefits of STEM education, particularly for increasing student engagement and understanding when content is applied in real-world contexts. However, like many educational initiatives in the past, the enactment of reforms and policies do not typically meet the needs of historically excluded learners (Berry, 2021). As a field, it is imperative to undertake the conscientious work needed to realize the hope many have for STEM education.

Currently, STEM education is not well defined (Moore et al., 2020; Radloff and Guzey, 2016). Clearly operationalized definitions of STEM education must be paramount for moving the field forward. Without operationalized definitions, rigorous measurement of important constructs and concepts cannot occur. This impedes the progress of truly understanding the impact of high-quality STEM learning experiences on teaching and learning STEM. While discipline-specific approaches to STEM continue, other research (e.g., Roberts et al., 2022) encourages an interdisciplinary or transdisciplinary approach to STEM education in both research and practice. Frameworks for integrated STEM practices (e.g., Lowrie et al., 2018; Roberts et al., 2022) and approaches that are grounded in equitable approaches to STEM teaching and learning do exist (e.g., Jackson et al., 2021a,b). The field of STEM education should use these interdisciplinary perspectives to develop and evaluate systems and practices to support the teaching and learning of STEM so that students' and teachers' hopes for STEM education can be realized.

References

- Al Salami, M.K., Makela, C.J., de Miranda, M.A., 2017. Assessing changes in teachers' attitudes toward interdisciplinary STEM teaching. *Int. J. Technol. Des. Educ.* 27 (1), 63–88. <https://doi.org/10.1007/s10798-015-9341-0>.
- Anand, N., Dogan, B., 2021. Impact of informal learning environments on STEM education—views of elementary students and their parents. *Sch. Sci. Math.* 121 (6), 369–377. <https://doi.org/10.1111/ssm.12490>.
- Arivina, A., Jailani, J., 2020. Development of trigonometry learning kit with a STEM approach to improve problem solving skills and learning achievement. *Jurnal Riset Pendidikan Matematika* 7 (2), 178–194. <https://doi.org/10.21831/jrpm.v7i2.35063>.
- Asghar, A., Ellington, R., Rice, E., Johnson, F., Prime, G.M., 2012. Supporting STEM education in secondary science contexts. *Interdiscip. J. Probl. Based Learn.* 6 (2), 85–125. <https://doi.org/10.7771/1541-5015.1349>.
- Bascope, M., Reiss, K., Morales, M., Robles, C., Reyes, P., Duque, M.I., Andrade, J.C., 2020. A review of recent initiatives on STEM education in four Latin American countries. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T.J., English, L.D. (Eds.), *Handbook of Research on STEM Education*, first ed. Routledge, New York, pp. 51–62.
- Bell, D., Morrison-Love, D., Wooff, D., McLain, M., 2018. STEM education in the twenty-first century: learning at work—an exploration of design and technology teacher perceptions and practices. *Int. J. Technol. Des. Educ.* 28, 721–737. <https://doi.org/10.1007/s10798-017-9414-3>.
- Benavent, X., de Ves, E., Forte, A., Botella-Masarell, C., López-Iñesta, E., Rueda, S., Roger, S., Perez, J., Portalés, C., Dura, E., Garcia-Costa, D., Marzal, P., 2020. Girls4STEM: gender diversity in STEM for a sustainable future. *Sustainability* 12 (15), 6051. <https://doi.org/10.3390/su12156051>.
- Berry III, R.Q., 2021. 2021 Founders Lecture: examining mathematics education reforms' impact on historically excluded learners. *Investig. Math. Learn.* 13 (3), 153–166. <https://doi.org/10.1080/19477503.2021.1938870>.
- Bruce-Davis, M.N., Gubbins, E.J., Gilson, C.M., Villanueva, M., Foreman, J.L., Rubenstein, L.D., 2014. STEM high school administrators', teachers', and students' perceptions of curricular and instructional strategies and practices. *J. Adv. Acad.* 25 (3), 272–306. <https://doi.org/10.1177/1932202X14527952>.
- Bush, S.B., Cook, K.L., Edelen, D., Cox, R., 2020. Elementary students' STEAM perceptions: extending frames of reference through transformative learning experiences. *Elem. Sch. J.* 120 (4), 692–714. <https://doi.org/10.1086/708642>.
- Bussey, T.J., Lo, S.M., Rasmussen, C., 2020. Theoretical frameworks for STEM education research. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T.J., English, L.D. (Eds.), *Handbook of Research on STEM Education*, first ed. Routledge, New York, pp. 51–62.
- Bybee, R.W., 2018. STEM Education Now More Than Ever. National Science Teachers Association, Arlington, VA.
- Calabrese Barton, A., Tan, E., 2018. A longitudinal study of equity-oriented STEM-rich making among youth from historically marginalized communities. *Am. Educ. Res. J.* 55 (4), 761–800. <https://doi.org/10.3102/0002831218758668>.
- Christensen, R., Knezek, G., Tyler-Wood, T., 2014. Student perceptions of science, technology, engineering and mathematics (STEM) content and careers. *Comput. Hum. Behav.* 34, 173–186. <https://doi.org/10.1016/j.chb.2014.01.046>.
- Coxon, S.V., Dohrman, R.L., Nadler, D.R., 2018. Children using robotics for engineering, science, technology, and math (CREST-M): the development and evaluation of an engaging math curriculum. *Roeper Rev.* 40 (2), 86–96. <https://doi.org/10.1080/02783193.2018.1434711>.
- Crisp, G., Nora, A., Taggart, A., 2009. Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: an analysis of students attending a Hispanic serving institution. *Am. Educ. Res. J.* 46 (4), 924–942. <https://doi.org/10.3102/0002831209349460>.
- Dare, E.A., Ellis, J.A., Roehrig, G.H., 2014. Driven by beliefs: understanding challenges physical science teachers face when integrating engineering and physics. *J. Pre-Coll. Eng. Educ. Res.* 4 (2), 47–61. <https://doi.org/10.7771/2157-9288.1098>.
- Densen, C.D., Stallworth, C.A., Hailey, C., Householder, D.L., 2015. Benefits of informal learning environments: a focused examination of STEM-based program environments. *J. STEM Educ.* 16, 11–15.
- Early Childhood STEM Working Group, 2017. Early STEM Matters: Providing High-Quality STEM Experiences for All Young Learners: A Policy Report by the Early Childhood STEM Working Group. UChicago STEM Education, Chicago, IL.
- Ehlert, J., Roberts, T., 2021. Exploring public schools' approaches to STEM: a Pilot Study. In: Herron, K. (Ed.), *Proceedings of the 120th annual convention of the School Science and Mathematics Association*, Vol. 8. SSMA, Virtual.
- English, L.D., 2016. STEM education K-12: perspectives on integration. *Int. J. STEM Educ.* 3, 3. <https://doi.org/10.1186/s40594-016-0036-1>.
- Fleischman, H.L., Hopstock, P.J., Pelczar, M.P., Shelley, B.E., 2010. Highlights From PISA 2009: Performance of US 15-Year-Old Students in Reading, Mathematics, and Science Literacy in an International Context. NCES 2011-004. National Center for Education Statistics.
- Gao, X., Li, P., Shen, J., Sun, H., 2020. Reviewing assessment of student learning in interdisciplinary STEM education. *Int. J. STEM Educ.* 7 (24), 1–14. <https://doi.org/10.1186/s40594-020-00225-4>.
- Gonzalez, H.B., Kuenzi, J.J., 2012. Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer. Congressional Research Service: Library of Congress.
- Griese, B., Lehmann, M., Roesken-Winter, B., 2015. Refining questionnaire-based assessment of STEM students' learning strategies. *Int. J. STEM Educ.* 2 (1), 12. <https://doi.org/10.1186/s40594-015-0025-9>.
- Gupta, R., Voiklis, J., Rank, S.J., Dwyer, J.d.T., Fraser, J., Flinner, K., Nock, K., 2020. Public perceptions of the STEM learning ecology—perspectives from a national sample in the US. *Int. J. Sci. Educ.* B 10 (2), 112–126. <https://doi.org/10.1080/21548455.2020.1719291>.
- Herro, D., Quigley, C., 2017. Exploring teachers' perceptions of STEAM teaching through professional development: implications for teacher educators. *Prof. Dev. Educ.* 43, 416–438. <https://doi.org/10.1080/19415257.2016.1205507>.
- Herro, D., Quigley, C., Andrews, J., Delacruz, G., 2017. Co-measure: developing an assessment for student collaboration in steam activities. *Int. J. STEM Educ.* 4 (1), 26. <https://doi.org/10.1186/s40594-017-0094-z>.
- Holmlund, T.D., Lesseig, K., Slavik, D., 2018. Making sense of “STEM education” in K-12 contexts. *Int. J. STEM Educ.* 5 (32). <https://doi.org/10.1186/s40594-018-0127-2>.
- Honey, M., Pearson, G., Schweingruber, A., 2014. STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research. National Academies Press, Washington.
- International Technology and Engineering Educators Association, 2020. Standards for Technological and Engineering Literacy. ITEEA, Reston.
- Jackson, C., Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Roberts, T., Yost, C., Fowler, A., 2021a. Equity-oriented conceptual framework for K-12 STEM literacy. *Int. J. STEM Educ.* 8 (38). <https://doi.org/10.1186/s40594-021-00294-z>.
- Jackson, C., Mohr-Schroeder, M.J., Roberts, T., Maiorca, C., 2021b. Students' Perception of Mathematics and Science. Paper Presented at the American Educational Research Association Annual Conference, Accepting Educational Responsibility, Virtual, 8 April–12 April 2021.
- Jong, C., Priddle, C., Roberts, T., Museus, S., 2020. Race-related factors in STEM: a review of research on educational experiences and outcomes for racial and ethnic minorities. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T.J., English, L.D. (Eds.), *Handbook of Research on STEM Education*, first ed. Routledge, New York, pp. 51–62.
- Kang, N.H., 2019. A review of the effect of integrated STEM or STEAM (science, technology, engineering, arts, and mathematics) education in South Korea. *Asia-Pac. Sci. Educ.* 5 (6). <https://doi.org/10.1186/s41029-019-0034-y>.
- Kitchen, J.A., Sonnet, G., Sadler, P.M., 2018. The impact of college- and university-run high school summer programs on students' end of high school STEM career aspirations. *Sci. Stud. Sci. Educ.* 102 (3), 529–547. <https://doi.org/10.1002/sce.21332>.
- Korea Institute of Curriculum and Evaluation, 2014. PISA 2012 Main Outcomes: Mathematics, Reading, Science and Problem Solving (Vol. PIM 2014-12). KICE, Seoul.
- Korean Ministry of Education, Science and Technology, 2011. The Second Basic Plan to Foster and Support the Human Resources in Science and Technology (2011–2015). MEST, Seoul.
- Krapp, A., Prenzel, M., 2011. Research on interest in science: theories, methods, and findings. *Int. J. Sci. Educ.* 33 (1), 27–50. <https://doi.org/10.1080/09500693.2010.518645>.

- Krapp, A., 2002. An educational-psychological theory of interest and its relation to self-determination theory. In: Deci, E., Ryan, R. (Eds.), *The Handbook of Self-Determination Research*. University of Rochester Press, Rochester, NY, pp. 405–427.
- Kwon, H., Capraro, R.M., Capraro, M.M., 2021. When I believe, I can: success STEMs from my perceptions. *Can. J. Sci. Math. Technol. Educ.* 2021, 1–19. <https://doi.org/10.1007/s42330-020-00132-4>.
- Lacoste, K., 2020. Creativity and active learning in libraries: variations on “STEM” in the United Kingdom and Ireland. *Int. Leads* 35 (2), 17–19.
- Li, Y., Wang, K., Xiao, Y., Froyd, J.E., 2020. Research and trends in STEM education: a systematic review of journal publications. *Int. J. STEM Educ.* 7, 11. <https://doi.org/10.1186/s40594-020-00207-6>.
- Lowrie, T., Leonard, S., Fitzgerald, R., 2018. STEM Practices: a translational framework for large-scale STEM education design. *Educ. Des. Res.* 2 (1). <https://doi.org/10.15460/eder.2.1.1243>.
- Maiorca, C., Roberts, T., Jackson, C., Bush, S., Delaney, A., Mohr-Schroeder, M.J., Soledad, S.Y., 2021. Informal learning environments and impact on interest in STEM careers. *Int. J. Sci. Math. Educ.* 19, 45–64. <https://doi.org/10.1007/s10763-019-10038-9>.
- Margot, K.C., Kettler, T., 2019. Teachers' perception of STEM integration and education: a systematic literature review. *Int. J. STEM Educ.* 6, 2. <https://doi.org/10.1186/s40594-018-0151-2>.
- Martin, D.B., Gholson, M.L., Leonard, J., 2010. Mathematics as gatekeeper: power and privilege in the production of knowledge. *J. Urban Math. Educ.* 3 (2), 12–24. <https://doi.org/10.21423/jume-v3i2a95>.
- McMullin, K., Reeve, E., 2014. Identifying perceptions that contribute to the development of successful project lead the way pre-engineering programs in Utah. *J. Technol. Educ.* 26 (1), 22–46. <https://doi.org/10.21061/jte.v26i1.a.2>.
- Melak, A., Singh, S., 2021. Women's participation and factors affecting their academic performance in engineering and technology education: a study of Ethiopia. *Sustainability* 13, 2246. <https://doi.org/10.3390/su13042246>.
- Mohr-Schroeder, M.J., Jackson, C., Miller, M., Walcott, B., Little, D.L., Speler, L., Schooler, W., Schroeder, D.C., 2014. Developing middle school students' interests in STEM via summer learning experiences: see Blue STEM camp. *Sch. Sci. Math.* 114 (6), 291–301. <https://doi.org/10.1111/ssm.12079>.
- Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Nickels, M., 2020. Moving toward an equity-based approach for STEM literacy. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T., English, L. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 29–38.
- Mohr-Schroeder, M., 2015. Track 3 panel session: national models for broadening participation. In: Invited Panelist Speaker at the 24th National EPSCoR National Conference, Portsmouth, NH.
- Moore, T.J., Johnston, A.C., Glancy, A.W., 2020. STEM integration: a synthesis of conceptual frameworks and definitions. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T., English, L. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 3–16.
- Morrison, J., 2006. *TIES STEM Education Monograph Series, Attributes of STEM Education*. TIES, Baltimore, MD.
- Morton, C.H., Smith-Mutegi, D., 2018. Girls STEM institute: transforming and empowering Black girls in mathematics through STEM. *Annu. Perspect. Math. Educ.* 23–38.
- Nadelson, L.S., Seifert, A., 2013. Perceptions, engagement, and practices of teachers seeking professional development in place-based integrated STEM. *Teach. Educ. Pract.* 26 (2), 242.
- National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010. *Common Core State Standards for Mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C.
- National Research Council, 2009. *Learning Science in Informal Environments: People, Places, and Pursuits*. National Academies Press.
- National Research Council [NRC], 2014. *Developing Assessments for the Next Generation Science Standards*. The National Academies Press, Washington, DC.
- National Science Board, 2015. *Revisiting the STEM Workforce: A Companion to Science and Engineering Indicators 2014*. National Science Board.
- Nesmith, S.N., Cooper, S., 2020. Elementary STEM learning. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T., English, L. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 101–114.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, by States*. The National Academies Press, Washington, DC.
- Nugent, G., Barker, B., Grandgenett, N., Adamchuk, V.I., 2010. Impact of robotics and geospatial technology interventions on youth STEM learning and attitudes. *J. Res. Technol. Educ.* 42 (4), 391–408. <https://doi.org/10.1080/15391523.2010.10782557>.
- Office of the Chief Scientist, 2014. *Science, Technology, Engineering and Mathematics: Australia's Future*. Commonwealth of Australia, Canberra.
- Park, M., Dimitrov, D.M., Patterson, L.G., Park, D., 2017. “Early childhood teachers’ beliefs about readiness for teaching science, technology, engineering, and mathematics. *J. Early Child. Res.* 15, 275–291. <https://doi.org/10.1177/1476718X15614040>.
- Phillip, R.A., 2007. Mathematics teachers’ beliefs and affect. In: Lester, F.K. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, vol. 1, pp. 257–315.
- Popovic, G., Lederman, J.S., 2015. Implications of informal education experiences for mathematics teachers’ ability to make connections beyond the formal classroom. *Sch. Sci. Math.* 115 (3), 129–140. <https://doi.org/10.1111/ssm.12114>.
- President's Council of Advisors on Science and Technology (PCAST), 2010. *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future*. Author, Washington.
- Pringle, R.M., Lord, C.G., Sheppard, T.D., 2020. Secondary STEM learning. In: Johnson, C., Mohr-Schroeder, M.J., Moore, T., English, L. (Eds.), *Handbook of Research on STEM Education*. Routledge, pp. 115–127.
- Radloff, J., Guzey, S., 2016. Investigating preservice STEM teacher conceptions of STEM education. *J. Sci. Educ. Technol.* 25 (5), 759–774. <https://www.learnedtechlib.org/p/176204/>.
- Ramzan, M., Kashif, N.U., Ali, A., 2021. The significance of STEM in effective teaching-learning process at secondary schools. *S. J. E. S. R.* 4 (3), 37–42.
- Roberts, T., Chapman, P., 2017. Authentically engaging elementary students in the designed world. *Child. Technol. Eng.* 21 (4), 15–17.
- Roberts, T., Schnepf, J., 2020. Building problem-solving through STEAM. *Technol. Eng. Teach.* 79 (8).
- Roberts, T., Jackson, C., Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Cavalcanti, M., Schroeder, D.C., Delaney, A., Putnam, L., Creameans, C., 2018. Students' perceptions of STEM learning after participating in a summer informal learning experience. *Int. J. STEM Educ.* 5 (35). <https://doi.org/10.1186/s40594-018-0133-4>.
- Roberts, T., Jackson, C., Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Delaney, A., 2019. Exploring applications of school mathematics: students' perceptions of informal learning experiences. In: Otten, S., Candela, A.G., de Araujo, Z., Haines, C., Munter, C. (Eds.), *Proceedings of the Forty-First Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. University of Missouri, St. Louis, MO, pp. 1515–1520.
- Roberts, T., Maiorca, C., Jackson, C., Mohr-Schroeder, M.J., 2022. Integrated STEM as problem-solving practices. *Investig. Math. Learn.* 14 (1). <https://doi.org/10.1080/19477503.2021.2024721>.
- Roehrig, G.H., Moore, T.J., Wang, H.H., Park, M.S., 2012. Is adding the E enough? Investigating the impact of K-12 engineering standards on the implementation of STEM integration. *Sch. Sci. Math.* 112 (1), 31–44. <https://doi.org/10.1111/j.1949-8594.2011.00112.x>.
- Roller, S.A., Lampley, S.A., Dillihunt, M.L., Benfield, M.P.J., Gholston, S.E., Turner, M.W., Davis, A.M., 2020. Development and initial validation of the student interest and choice in STEM (SIC-STEM) survey 2.0 instrument for assessment of the social cognitive career theory constructs. *J. Sci. Educ. Technol.* 29, 646–657. <https://doi.org/10.1007/s10956-020-09843-7>.
- Shermoff, D.J., Sinha, S., Bressler, D.M., Ginsburg, L., 2017. Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *Int. J. STEM Educ.* 4 (13). <https://doi.org/10.1186/s40594-017-0068-1>.
- Sikma, L., Osborne, M., 2014. Conflicts in developing an elementary STEM magnet school. *Theory Pract.* 53 (1), 4–10. <https://doi.org/10.1080/00405841.2014.862112>.
- Staus, N.L., Lesseig, K., Lamb, R., Falk, J., Dierking, L., 2020. Validation of a measure of STEM interest for adolescents. *Int. J. Sci. Math. Educ.* 18, 279–293. <https://doi.org/10.1007/s10763-019-09970-7>.

- Stohlmann, M., Moore, T.J., McClelland, J., Roehrig, G.H., 2011. Impressions of a middle grades STEM integration program: educators share lessons learned from the implementation of a middle grades STEM curriculum model. *Middle Sch. J.* 43 (1), 32–40. <https://doi.org/10.1080/00940771.2011.11461791>.
- Stohlmann, M., Moore, T.J., Roehrig, G.H., 2012. Considerations for teaching integrated STEM education. *J. Pre-Coll. Eng. Educ. Res.* 2 (1). <https://doi.org/10.5703/1288284314653>.
- Stubbs, E.A., Myers, B.E., 2016. Part of what we do: teacher perceptions of STEM integration. *J. Agric. Educ.* 57 (3), 87–100. <https://doi.org/10.5032/jae.2016.03087>.
- Thibaut, L., Knipprath, H., Dehaene, W., Depaepe, F., 2018. The influence of teachers' attitudes and school context on instructional practices in integrated STEM education. *Teach. Educ.* 71, 190–205. <https://doi.org/10.1016/j.tate.2017.12.014>.
- Thompson, A.G., 1992. Teachers' beliefs and conceptions: a synthesis of the research. In: Grouws, D.A. (Ed.), *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*. Macmillan Publishing Co, Inc, pp. 127–146.
- Wang, C., Frye, M., 2019. miniGEMS 2018: a mixed methods study exploring the impact of a STEAM and programming camp on middle school girls' STEM attitudes. In: Paper Presented at 2019 American Society for Engineering Educational Annual Conference & Exposition, Tampa, Florida. <https://doi.org/10.18260/1-2-33113>.
- Wang, H.H., Moore, T.J., Roehrig, G.H., Park, M.S., 2011. STEM integration: teacher perceptions and practice. *J. Pre-Coll. Eng. Educ. Res.* 1 (2). <https://doi.org/10.5703/1288284314636>.
- Weis, L., Eisenhart, M., Cipollone, K., Stich, A.E., Nikischer, A.B., Hanson, J., Leibrandt, S.O., Allen, C.D., Dominguez, R., 2015. In the guise of STEM education reform: opportunity structures and outcomes in inclusive STEM focused high schools. *Am. Educ. Res. J.* 52 (6), 1024–1059. <https://doi.org/10.3102/0002831215604045>.
- You, H.S., Marshall, J.A., Delgado, C., 2018. Assessing students' disciplinary and interdisciplinary understanding of carbon cycling. *J. Res. Sci. Teach.* 55 (3), 377–398. <https://doi.org/10.1002/tea.21423>.

Pre-service STEM teacher education

Valerie W.Y. Yip and Promail K.Y. Leung, Faculty of Education, The University of Hong Kong, Pok Fu Lam, Hong Kong, China

© 2023 Elsevier Ltd. All rights reserved.

Challenges of pre-service STEM teacher education	270
Models of STEM pre-service programmes	270
Technology in STEM education	271
Support from stakeholders	271
The way forward	272
Attributes of STEM teachers	272
Professional development models for pre-service STEM teachers	272
Strategies of pre-service STEM teacher education	272
Situated learning in the course design	272
Project-based learning	273
Place-based education	273
Video-based education	273
Support from stakeholders	274
Future research initiatives	274
References	274

The acronym STEM (Science, Technology, Engineering and Mathematics) represents an integration of the four fields of studies that can remain independent on their own. Since STEM has been introduced in the 1990s by the National Science Foundation of US, it has gained a lot of attention from the education community and it has been a vibrant field of the research area in the past 10 years. From 2000 to 2018, the number of studies on STEM education grew enormously in 36 selected journals. STEM education research has been increasingly recognized and published across the world (Li et al., 2020).

The interdisciplinary nature of STEM education advocates an integrative instructional approach (Sanders, 2009) that has been examined by many studies (e.g., Berlin and White, 1995; Kennedy and Odell, 2014; Stohlmann et al., 2012). Given the complex, ever-changing nature of real-world problems, STEM teachers need to continuously understand and identify the complexity of problems, learn how to integrate various disciplinary knowledge in their curriculum design, and evaluate their instructions. They need to cater for the cultural needs of the students, create multi-facet learning scenarios, and promote higher order thinking skills such as critical thinking to solve the real-world problems (Stillman and Anderson, 2014). Another challenge faced by STEM educators is a lack of common understandings of the comprehensiveness of STEM education (Bybee, 2010; Keefe, 2009). It has been difficult to define what STEM education is, how school pupils' STEM literacy can be developed and consequently to identify effective ways to enhance teachers' proficiencies in teaching STEM. Pre-service teachers particularly experience difficulties to adopt integrative STEM instruction for their inadequate subject content knowledge in different fields (Blackley and Howell, 2015), limited authentic teaching experience in general (Verloop et al., 2001), and the absence of integrative STEM learning environment in the historical, discipline-specific teacher education they have been engaged in. Without clear STEM-related conceptions, pre-service STEM teachers will misunderstand the whole exertion of interdisciplinary teaching and learning, and possibly ended up with frustrations and withdrawal from STEM-related instruction (Atkinson, 2012; Atkinson and Mayo, 2010; Golding, 2009).

Challenges of pre-service STEM teacher education

The challenges of initial STEM teacher training not only stem from the personal factors of the novice teachers, but also from multiple factors such as a lack of evidence-based professional development models, the under-researched teachers' STEM-related technological pedagogical content knowledge and strong support from various stakeholders.

Models of STEM pre-service programmes

STEM teachers require the ability to provide STEM-related experiences to promote pupils' STEM literacy, for which the latter can be defined as acquiring "the knowledge and understanding of scientific and mathematical concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity" (NRC, 2011, p. 5). The call for integrative STEM instruction is coherent with the notion that an integrated mode of teacher preparation (Bragg, 1998) with clear connections between K-12 and higher education (Kalantzis and Cope, 2014; Ledbetter, 2012).

An important aspect of teacher education is to examine teachers' thought processes in making pedagogical decisions. Richardson (1996) claims that attitudes and beliefs are important factors affecting pre-service teachers to decide what to teach and how to

achieve the learning objectives. Due to limited practical experience, pre-service teachers tend to teach in the same way they learned from previous experience (Deemer, 2004; Koballa et al., 2000; Moore-Russo and Waight, 2019). Since STEM pre-service teachers are subject-based trained and neither have interdisciplinary knowledge nor prior experiences and practice to reproduce their teacher identity, pre-service teacher education program and the concurrent STEM-related teacher practicum experience play a vital role in developing their perception of STEM education (Radloff and Guzey, 2017).

Nevertheless there are conflicts between the ideal and realistic initial STEM teacher education. First, the corresponding pre-service teacher preparation should view STEM teaching as boundary-crossing and integrative in nature (Stillman and Anderson, 2014); while on the other it has been difficult to articulate the notion of interdisciplinarity into epistemological and pedagogical approaches (Jenlink and Jenlink, 2019) for the fluidity of STEM and STEM education over the years. Second, an innovative, adaptive and holistic curriculum is uncommon in K-12 STEM education and hence, it is less likely to have favorable impacts on teacher preparation programs. Although STEM education instructional units with exemplars (e.g., Fulmer et al., 2018) can provide clear standards, guidance and tools for novice teachers to reflect on their own instructional design, a research-based curriculum framework would be desirable for educators to propose their customized rather than standardized STEM education curriculum for everyday classroom teaching as well as pre-service teacher preparation. Third, STEM teachers' identity, role and understanding of STEM education have been continuously studied in the field but few have implications on the STEM teachers' initial education. Only a few reported such programs (Babb et al., 2016; Lee and Nason, 2012) which leaves a gap in STEM teachers' initial preparation. All these point toward a need to have effective models of integrated teacher preparation that are supported by empirical evidence for developing future pre-service teachers' subject content knowledge, pedagogical knowledge and attitudes and beliefs.

Technology in STEM education

"Teachers are the single most important factor in the K-12 education system ... Great STEM teachers have at least two attributes: deep content knowledge in STEM, and strong pedagogical skills for teaching their students STEM" (Holdren et al., 2010, p. 63). STEM pedagogy is a collection of the qualities derived from STEM disciplines, including higher order thinking skills, computational thinking, creative designs, esthetic appreciations, and social interactions. These are necessary to advance a new paradigm of STEM teaching.

Eckman et al. (2016) reported that successful STEM teaching demands pedagogical content knowledge (PCK) in inquiry and activity-based STEM teaching practices for teachers to communicate STEM concepts effectively in a way that students can understand. The UNESCO (2008) ICT-CST framework and standards were also designed to assist countries in preparing innovative and self-directed teachers and pre-service teachers who would be able to integrate technology with their interdisciplinary knowledge dynamically in making pedagogical decisions to teach socio-cultural and political issues (Kurup et al., 2019). Using technology in STEM education seems promising to cope with the ever-changing society and align with 21st-century learning.

As STEM teachers should use technology in their instruction, it would be desirable for them to possess STEM-related technological pedagogical content knowledge (i.e., STEM-TPCK) so that they would be proficient in identifying the needs of integrating different disciplines and learning how to synthesize different perspectives to advocate integrative education. STEM-TPCK is not unique to technology teachers who always have to adopt and teach technology. As Sanders (2009) described these teachers:

Many technology teachers are fond of saying they teach science and math[ematics] in their technology education programs. In truth, it is exceedingly rare for a technology teacher to explicitly identify a specific science or mathematics concept or process as a desired learning outcome and even rare for technology teachers to assess a science or mathematics learning outcomes. (p. 21)

The discrepancy between the claim and practices of teachers might not be only accounted for by their beliefs and attitudes, but also due to a lack of STEM-TPCK. The possible ways to develop teachers' STEM-TPCK deserve further investigation.

Support from stakeholders

In the four dimensions framework model, Bybee (2013) reported that the purposes and policy of STEM education were fundamental to guide the policy-maker and educators to develop useful society-based STEM curriculum and instructional materials to inform STEM literate citizens (Kurup et al., 2019; Zeidler, 2016; Zeidler et al., 2005). Moreover, it is desirable to establish clear guidelines to make a balance between the integrative STEM instruction and teaching individual disciplines in highly structured schooling systems. These can be exemplified by: having guidance in disciplinary-based lesson and curriculum to help teachers to experience teaching environments that resemble K-12 school settings (Corlu et al., 2014; Shernoff et al., 2017), pre-service teacher education framework (e.g., framework for reforming preparation of STEM teachers in the US) (Lee and Nason, 2012), or pioneering STEM Education Course (e.g., the five dimensions of engaging learning that conceptualized the teaching preparation program in Canada) (Babb et al., 2016) to train the STEM teachers. By date, training interdisciplinary STEM teachers is usually not an intentional part of the teacher education program and is regarded as a subsidiary (Radloff and Guzey, 2017; Wilson's report as cited in Rogers et al., 2016). STEM teachers often receive training from part of a major methods program or STEM elective courses. This would consequently hinder the development of both STEM education and STEM career manpower pipeline (e.g., Adams et al., 2014; Babb et al., 2016; Bartels et al., 2019; Radloff and Guzey, 2017).

The way forward

Pre-service teacher education is fundamental to prepare and sustain the workforce in the STEM education community and ensure quality teaching. To address the challenges of initial STEM teacher training, it is essential to identify the attributes of STEM teachers, develop models for teacher training, identify useful strategies to support student teachers in building their knowledge base, skills and attitudes and beliefs. From the sociological perspective, support from different stakeholders is crucial for systemic changes.

Attributes of STEM teachers

Attributes or qualities of STEM teachers should be identified when we develop useful pre-service teacher education programs. In other words, the goals (i.e., attributes of STEM teachers) would determine the training models adopted in different social contexts. There has been research on the STEM literacy K-12 students should acquire, but much less on what is expected of the novice teachers after they have received proper training. In view of the interdisciplinary nature of STEM, the qualities expected on the novices may range from generic to specific and should be comprehensive. In general, teachers should be capable to reflect critically (e.g., [Loughran, 2002](#)) on educational practices. STEM teachers should be particularly reflective so that they can respond promptly to the changes in technology, to the characteristics and learning needs of K-12 students such as their attitudes toward STEM and social demands. STEM teachers who are prepared for self-directed learning may have better skills to analyze their practices continuously and solve problems effectively ([Gencel and Saracaloğlu, 2018](#)). This attribute echoes with the purpose of STEM to strengthen individuals' capabilities in dealing with complex and evolving problems. In addition, creativity and STEM education have a tight connection. Teachers with limited learning experience from a creative and divergent environment would consequently confine their motivation and practices to nurture creativity of their students ([Harris and Bruin, 2018](#)). Before STEM teachers would acquire STEM-TPCK, their digital literacy such as managing technological tools, coding and programming, computational thinking, marshalling evidence and data presentation has to be developed. Afterward they should have the opportunity to apply technology in STEM teaching.

Professional development models for pre-service STEM teachers

There have been a number of teacher preparation programs aiming to develop the competencies of pre-service teachers. Despite the successful and failing attempts, there has been a lack of effective professional development models supported by empirical evidence that focuses on developing pre-service teachers' subject content knowledge, pedagogical knowledge and attitudes and beliefs.

[Rogers et al. \(2016\)](#) proposed their "authentic and sustainable 4-pillar model" based on the Project Engage. The four pillars include: facilitating student teachers to gain insights into STEM career by interacting with STEM specialists in career events, bringing them to experience real-world STEM applications by exploring STEM facilities and workplaces, supporting them to gauge STEM knowledge and skills through mentoring programs, and gaining experience about STEM teaching by supporting K-12 students in STEM competitions. The four pillars were built upon the authenticity model proposed by [Strobel et al. \(2013\)](#) that emphasizes K-12 STEM learning should be context authentic, task authentic, impact authentic and personal/value authentic. The project focused on developing pre-service teachers' understanding of subject matter knowledge and pedagogy to foster collaborative inquiry, problem-solving, design thinking, model building and reasoning. The social interactions between the student teachers and STEM-related stakeholders were essential to support student teachers' in their learning. Being one of the very few STEM initial teacher development models, the researchers regarded their suggestion as a "side product" of the project that required further empirical studies to support its usefulness.

The importance of engaging student teachers actively in STEM and STEM education related activities proposed by [Rogers et al.'s \(2016\)](#) echoes well with the first two principles of another framework for reforming STEM teacher preparation ([Lee and Nason, 2012](#)). There are three principles suggested in the reform framework. The first focuses on developing the novices' disciplinary and trans-disciplinary subject content knowledge and pedagogical knowledge, and the second principle emphasizes the development of "situated theoretical knowledge" that would allow the pre-service teachers to reflect on the theoretical underpinnings of the educational practices such as the classroom learning activities and STEM curriculum ([Custer and Daugherty, 2009](#)). A unique dimension suggested by [Lee and Nason \(2012\)](#) is to support the novices to overcome their negative dispositions ([Sherman and MacDonald, 2007](#)) and fears toward teaching "new" STEM. Appropriate beliefs and dispositions are assumed to be built upon better understandings of STEM knowledge and pragmatic experience; while the affective and emotional aspects of student teachers can be taken care of through continuous intervention and reflective practices.

Strategies of pre-service STEM teacher education

The useful strategies of integrative STEM teacher preparation programs have been reported according to the theoretical underpinnings of individual projects. Nevertheless, these programs "implicitly" followed the models suggested in the previous section.

Situated learning in the course design

Preservice teacher education programs usually emphasize on developing the novices' subject matter knowledge and pedagogical knowledge to integrate the four fields in their teaching. [Rinke et al. \(2016\)](#) implemented a STEM block that focused on the process

of technology and engineering to a group of elementary pre-service teachers. At the end of the course, the student teachers showed a significant increase in the content integration of their lesson planning, made better use of engineering and design concepts, and were more confident to teach STEM. Similarly, [Berlin and White \(2010\)](#) introduced the integrative STEM pedagogies to the master program pre-service teachers, involved them to produce teaching packages and carried out action research. Upon the completion of the program, the beginning teachers valued the integrative approach of STEM. Indeed a holistic approach incorporated with situated learning that engages the novices in planning and implementing lesson plans is common in teacher preparation programs (e.g., [Bartels et al., 2019](#)).

Project-based learning

In theory, project-based learning or problem-based learning in teacher education may simulate the ways K-12 students to solve daily life problems in the context of STEM. Pre-service teachers need to learn connecting subject matter knowledge and educational practices with technology so that technological applications can be used effectively in teaching and learning ([Niess, 2005](#)). In [Awad and Barak's \(2018\)](#) study, student teachers carried out scientific inquiries and solved problems on the topic of sound waves and communication systems by making use of information and communications technology (ICT). The integration of technology and content, adoption of blended learning and the use of ICT were found to be useful for improving student teachers' subject matter knowledge and interests in ICT. Consequently their STEM-TPCK would have been developed through the use of technology. On the other hand, the impacts of project-based learning on increasing their self-efficacy were inconclusive. The student teachers in [Awad and Barak's \(2018\)](#) project reported that working independently in an ICT environment distanced them from STEM. As such, more empirical studies would be useful to provide evidence on whether collaborative learning would be necessary for student teachers to develop their STEM teaching competence in an online mode ([Lee and Nason, 2012](#)).

Place-based education

Place-based education provides an authentic, situated learning experience for pre-service STEM teachers. Teacher educators need to model and provide authentic and meaningful experiences that facilitate community building. This would increase the confidence of pre-service teachers as they can find resources readily available. Enthusiastic community partners are also ready to guide them in bridging the theories and practices and overcoming the challenges of teaching through apprenticeship. There are two aspects of field-based STEM teacher training: preparing pre-service teachers with adequate understandings of different fields in STEM and developing them to become competent in teaching.

Teacher preparation should provide more opportunities for STEM student teachers to work in the related industries, particularly at their early stage of training. The government can collaborate with teacher education institutions (TEIs) and industries to introduce student internships, fellowships or summer jobs in the STEM-related industries ([Atkinson and Majo, 2010](#)). The government can increase the incentives for such collaboration through providing subsidies, tax reduction or free time for the mentors to guide their mentees. This would encourage more students to stay in the STEM-related career pipeline.

Another aspect of field-based training can be exemplified in teaching practicum. In [Adams et al. \(2014\)](#) project, the elementary school teachers stayed in a residential field campus for collaborative scientific inquiry about the local natural environment. They explored the STEM content and carried out research and presentation that would be similar to what their future students would do. [Eckman et al. \(2016\)](#) adopted a cooperative education model that is usually adopted in engineering education. It combined academic study with practical work experience to prepare the STEM pre-service teachers in 12 to 16 months. The pre-service "engineering" teachers were placed in urban middle and high schools and regularly met the university instructors. Results indicated that STEM teachers prepared through this education model would exceed their counterparts learning by traditional mode in a few areas: being more reflective in teaching practices, better mastery of central concepts of STEM fields, delivering appropriate instructions to all students, providing differentiated instructions to meet the diverse needs of students, and overcoming the structural barriers to student learning. In conclusion, field-based experience allows the novice teachers to build meaningful connections between theories learned from the coursework and the classroom practices.

Video-based education

Making use of videos showcasing integrative STEM instruction is another possible strategy to open a "window" for the pre-service teachers to enhance their noticing abilities ([Sherin and van Es, 2005](#)) and establish better connections between theories and practices without entering schools. [Radloff and Guzey \(2017\)](#) carried out a video-based intervention – which is relatively rare in STEM teacher preparation – for a group of elementary pre-service teachers in a semester. The participants were guided to analyze and reflect over videos and create lesson plans. The pre-service teachers demonstrated more alignment with the principles of integrative STEM education. Although videos are not a new pedagogical tool in teacher education, this technology has gained more importance in the period of Covid-19 – a time at which pre-service teachers might not have place-based learning (professional practicum) when many schools were suspended. There are two possible ways of using videos in this situation. Similar to [Radloff and Guzey's \(2017\)](#) study, teacher educators can use pre-recorded STEM lesson videos in their programs to develop pre-service teachers' observation, analytical abilities and reflection as a supplement to the missing face-to-face teaching experience in the practicum. In addition, the pre-service teachers may have online professional practicum during school suspension at which they have to observe and/or co-teach the online lessons with their mentor teachers, or they may conduct independent online lessons. Videos of the online lessons taken during school suspension, together with the live online experience of the same lessons, can serve as a catalyst for the pre-service teachers

to develop understandings of what “authenticity” means in a new era of teaching. Such a hybrid mode of learning closely resembles what has been experienced by their future K-12 students in terms of “authenticity”. Pre-service STEM teachers would also be able to develop their STEM-TPCK with the evidence shown in the videos.

Support from stakeholders

The establishment of professional training models for pre-service teachers and relevant practices is closely related to the policies for STEM education. Bybee (2013) made use of an analogy that “the policies are like blueprints for a house” to describe how they are related to the purposes, instructional practices and material development in STEM education. The policies also inevitably affect the teacher professional development, for instance how the educators would develop the novice teachers in interpreting the national/central curriculum, and hence their own beliefs and practices in STEM learning and teaching. Issues such as the extent to which the four fields of STEM can integrate, how they can integrate while maintaining the uniqueness of individual disciplines are also guided by the policies. Frequent changes in and highly structured school curriculum without research evidence would hamper the development of meaningful teacher education programs to prepare the novice STEM teachers.

To achieve a fruitful STEM-focused teacher training program, it requires innovations, collaborations, and decisions from all levels of individuals and governments (Fensham, 2016). Governments can manage the teacher training programs by providing resources to and endorsing the programs offered by teacher education institutions (TEIs) and issuing teacher certificates (Shernoff et al., 2017). For instance, a unique and innovative double major STEM education teaching certificate program was offered by the University of Arizona in response to the call of government (Mohr-Schroeder et al., 2019). This program was the first one in US resulting in a degree and certification that truly prepares teacher candidates as secondary (grades 8 to 12) STEM specialists. The program encouraged future STEM pre-service teachers to engage in a more interdisciplinary and transdisciplinary study through a STEM perspective.

Another example to illustrate the importance of clear government policy was Virginia Tech’s program, the first of its kind in US that was launched in 2006. It served as an exemplar of a STEM-integration paradigm for preparing 21st century educators, leaders, scholars, and researchers to investigate, teach, and disseminate new integrative approaches to STEM education. The integrative STEM education program provided the theoretical and pedagogical premise for technological/engineering (T/E) design-based teaching and learning practices (Mohr-Schroeder et al., 2019) which required various stakeholders to collaborate.

The experience of pre-service teachers in field-based learning as an important part of teacher training is affected by the ethos of the schools and/or industries at which they would spend a lot of time in being guided to observe, practice and reflect (Hardy et al., 2008). A clinically based preparation program of Kentucky, US (O’Brien et al., 2014) that aligned with the National Council for the Accreditation of Teacher Education (NCATE, 2010) provided such an in-depth learning experience. The program allowed the novice teachers to visit various workspaces and have at least 255 h of placement, giving them hands-on experience with adolescents in classrooms and non-traditional settings to foster their STEM education experience in an integrated manner.

In echoing the blueprint analogy (Bybee, 2013), government policy plays a key role in building the STEM teachers’ capacity in dealing with STEM education promotion. The government can further regulate STEM teacher employment, their relocation, exercise power over classroom practices through quality assurance of schools (Corlu et al., 2014), and propose competency frameworks with sufficient professional development support for STEM teachers at different stages of their career.

Future research initiatives

All policies, curriculum and instructional practices and STEM pre-service program models require the support of empirical studies. Future initiatives on pre-service STEM teacher preparation can focus on investigating teaching STEM from K-12 levels, on the effects of particular learning activities on novice teachers’ subject matter knowledge, technological knowledge, pedagogical knowledge, and dispositions and beliefs. Longitudinal studies to reveal the long-term effects of pre-service training, as well as more systematic, comparative reviews of initial teacher education programs would be useful to develop effective models of teacher professional development.

References

- Adams, A.E., Miller, B.G., Saul, M., Pegg, J., 2014. Supporting elementary pre-service teachers to teach stem through place-based teaching and learning experiences. *Electr. J. Res. Sci. Mathe. Educ.* 18 (5).
- Atkinson, R.D., 2012. Why the current education reform strategy won't work. *Issues Sci. Technol.* 28 (3), 29–36.
- Atkinson, R.D., Mayo, M.J., 2010. Refueling the US Innovation Economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education. The Information Technology & Innovation Foundation. Forthcoming.
- Awad, N., Barak, M., 2018. Pre-service science teachers learn a science, technology, engineering and mathematics (STEM)-oriented program: the case of sound, waves and communication systems. *Eurasia J. Math. Sci. Technol. Educ.* 14 (4), 1431–1451.
- Babb, A.P., Takeuchi, M., Yáñez, G.A., Francis, K., Gereluk, D., Friesen, S., 2016. Pioneering stem education for pre-service teachers. *Int. J. Eng. Pedag.* 6 (4), 4–11.
- Bartels, S.L., Rupe, K.M., Lederman, J.S., 2019. Shaping preservice teachers' understandings of stem: a collaborative math and science methods approach. *J. Sci. Teach. Educ.* 30 (6), 666–680.
- Berlin, D.F., White, A.L., 1995. Connecting school science and mathematics. In: Coxford, A.F., House, P. (Eds.), *Connecting Mathematics across the Curriculum*. National Council of Teachers of Mathematics, Reston, VA, pp. 22–23.

- Berlin, D.F., White, A.L., 2010. Preservice mathematics and science teachers in an integrated teacher preparation program for grades 7–12: a 3-year study of attitudes and perceptions related to integration. *Int. J. Sci. Math. Educ.* 8 (1), 97–115.
- Blackley, S., Howell, J., 2015. A STEM narrative: 15 years in the making. *Austr. J. Teach. Educ.* 40 (7) <https://doi.org/10.14221/ajte.2015v40n7.8>.
- Bragg, S., 1998. Investing in Tomorrow's Teachers: The Integral Role of Two-Year Colleges in the Science and Mathematics Preparation of Prospective Teachers. A Report from a National Science Foundation Workshop.
- Bybee, R.W., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teach.* 70 (1), 30–35.
- Bybee, R.W., 2013. The case for STEM education: challenges and opportunities. NSTA press.
- Corlu, M.S., Capraro, R.M., Capraro, M.M., 2014. Introducing stem education: implications for educating our teachers in the age of innovation. *Eğitim ve Bilim* 39 (171), 74–85.
- Custer, R.L., Daugherty, J., 2009. Professional development for teachers of engineering: research and related activities. *Bridge* 39 (3), 18.
- Deemer, S., 2004. Classroom goal orientation in high school classrooms: revealing links between teacher beliefs and classroom environments. *Educ. Res.* 46 (1), 73–90.
- Eckman, E.W., Williams, M.A., Silver-Thorn, M.B., 2016. An integrated model for stem teacher preparation: the value of a teaching cooperative educational experience. *J. STEM Teach. Educ.* 51 (1), 8.
- Fensham, P.J., 2016. The future curriculum for school science: what can be learnt from the past? *Res. Sci. Educ.* 46 (2), 165–185.
- Fulmer, G.W., Tanas, J., Weiss, K.A., 2018. The challenges of alignment for the next generation science standards. *J. Res. Sci. Teach.* 55 (7), 1076–1100.
- Gencel, I.E., Saracaloğlu, A.S., 2018. The effect of layered curriculum on reflective thinking and on self-directed learning readiness of prospective teachers. *Int. J. Prog. Educ.* 14 (1), 8–20.
- Golding, C., 2009. Integrating the Disciplines: Successful Interdisciplinary Subjects. Centre for the Study of Higher Education. The University of Melbourne, Melbourne, Australia.
- Hardy, G., Howes, A., Spendlove, D., Wake, G., 2008. The STEM Agenda in the Context of Initial Teacher Education: Challenges and Potential Ways Forward. Doctoral dissertation. Macmillan Publishers India Limited.
- Harris, A., de Bruin, L.R., 2018. Secondary school creativity, teacher practice and STEAM education: an international study. *J. Educ. Change* 19 (2), 153–179.
- Holdren, J.P., Lander, E.S., Varmus, H., 2010. Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future. Executive Report. President's Council of Advisors on Science and Technology, Washington, DC.
- Jenlink, P.M., Jenlink, K.E., 2019. The Next Generation of STEM Teachers: An Interdisciplinary Approach to Meet the Needs of the Future. Rowman & Litterfield, New York.
- Kalantzis, M., Cope, B., 2014. 'Education is the new philosophy', to make a metadisciplinary claim for the learning sciences. In: Reid, A.D., Hart, E.P., Peters, M.A. (Eds.), *A Companion to Research in Education*. Springer, Netherlands.
- Keefe, B., 2009. The Perception of STEM: Analysis, Issues, and Future Directions. Entertainment & Media Communication Institute, Division of Entertainment Industries Council, Burbank, CA.
- Kennedy, T.J., Odell, M.R.L., 2014. Engaging students in STEM education. *Sci. Educ. Int.* 25 (3), 246–258.
- Koballa Jr., T., Graber, W., Coleman, D.C., Kemp, A.C., 2000. Prospective gymnasium teachers' conceptions of chemistry learning and teaching. *Int. J. Sci. Educ.* 22 (2), 209–224.
- Kurup, P.M., Li, X., Powell, G., Brown, M., 2019. Building future primary teachers' capacity in stem: based on a platform of beliefs, understandings and intentions. *Int. J. STEM Educ.* 6 (1), 10.
- Ledbetter, M.L.S., 2012. Teacher preparation: one key to unlocking the gate to stem literacy. *CBE-Life Sci. Educ.* 11 (3), 216–220.
- Lee, K.T., Nason, R.A., 2012. Reforming the preparation of future stem teachers. In: *Proceedings of the 2nd International STEM in Education Conference*. Beijing Normal University, Beijing, pp. 33–39.
- Li, Y., Wang, K., Xiao, Y., Froyd, J.E., 2020. Research and trends in STEM education: a systematic review of journal publications. *Int. J. STEM Educ.* 7, 11.
- Loughran, J.J., 2002. Effective reflective practice: in search of meaning in learning about teaching. *J. Teach. Educ.* 53 (1), 33–43.
- Mohr-Schroeder, M., Jackson, C., Schroeder, D.C., Wilhelm, J., 2019. Developing a stem education teacher preparation program to help increase stem literacy among preservice teachers. In: Jenlink, P.M., Jenlink, K.E. (Eds.), *The Next Generation of STEM Teachers: An Interdisciplinary Approach to Meet the Needs of the Future*. Rowman & Litterfield, New York.
- Moore-Russo, D., Waight, N., 2019. Mathematics, science, and technology reform in teacher education. In: *The Next Generation of STEM Teachers: An Interdisciplinary Approach to Meet the Needs of the Future*, p. 69.
- National Research Council, 2011. Successful K-12 Stem Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics. National Academies Press.
- NCATE, 2010. Transforming teacher education through clinical practice: a national strategy to prepare effective teachers. In: *Report of the Blue Ribbon Panel on Clinical Preparation and Partnerships for Improved Student Learning*. ERIC Clearinghouse.
- Niess, M.L., 2005. Preparing teachers to teach science and mathematics with technology: developing a technology pedagogical content knowledge. *Teach. Teach. Educ.* 21 (5), 509–523.
- O'Brien, S., Karsnitz, J., Sandt, S., Bottomley, L., Parry, E., 2014. Engineering in preservice teacher education. In: Purzer, S., Strobel, J., Cardella, M.E. (Eds.), *Engineering in Pre-college Settings: Synthesizing Research, Policy, and Practices*. Purdue University Press, IN, USA.
- Radloff, J., Guzey, S., 2017. Investigating changes in preservice teachers' conceptions of stem education following video analysis and reflection. *Sch. Sci. Math.* 117 (3–4), 158–167.
- Richardson, V., 1996. The role of attitudes and beliefs in learning to teach. In: *Handbook of Research on Teacher Education*, vol. 2, pp. 102–119.
- Rinke, C.R., Gladstone-Brown, W., Kinlaw, C.R., Cappiello, J., 2016. Characterizing STEM teacher education: affordances and constraints of explicit STEM preparation for elementary teachers. *Sch. Sci. Math.* 116 (6), 300–309.
- Rogers, R.R., Winship, J., Sun, Y., 2016. Systematic support for stem pre-service teachers: an authentic and sustainable four-pillar professional development model. In: Dikiitis, K. (Ed.), *Innovative Professional Development Methods and Strategies for STEM Education*. IGI Global, PA, USA.
- Sanders, M., 2009. STEM education, STEMmania. *Technol. Teach.* 20–26. December/January.
- Sherin, M., van Es, E., 2005. Using video to support teachers' ability to notice classroom interactions. *J. Technol. Teach. Educ.* 13 (3), 475–491.
- Sherman, A., MacDonald, L., 2007. Pre-service teachers' experiences with a science education module. *J. Sci. Teach. Educ.* 18 (4), 525–541.
- Shermoff, D.J., Sinha, S., Bressler, D.M., Ginsburg, L., 2017. Assessing teacher education and professional development needs for the implementation of integrated approaches to stem education. *Int. J. STEM Educ.* 4 (1), 13.
- Stillman, J., Anderson, L., 2014. Preparing the next generation of teacher educators. *Teach. Coll. Rec.* 17581. <http://www.tcrecord.org>.
- Stohlmann, M., Moore, T.J., Roehrig, G.H., 2012. Considerations for teaching integrated STEM education. *J. Pre-Coll. Eng. Educ. Res.* 2 (1) <https://doi.org/10.5703/1288284314653>.
- Strobel, J., Wang, J., Weber, N.R., Dyehouse, M., 2013. The role of authenticity in design-based learning environments: the case of engineering education. *Comput. Educ.* 64, 143–152.
- UNESCO, 2008. Ict Competency Standards for Teachers: Competency Standards Modules. ERIC Clearinghouse.
- Verloop, N., van Driel, J.H., Meijer, P., 2001. Teacher knowledge and the knowledge base of teaching. *Int. J. Educ. Res.* 35 (5), 441–461.
- Zeidler, D.L., 2016. Stem education: a deficit framework for the twenty first century? A sociocultural socioscientific response. *Cult. Stud. Sci. Educ.* 11 (1), 11–26.
- Zeidler, D.L., Sadler, T.D., Simmons, M.L., Howes, E.V., 2005. Beyond sts: a research-based framework for socioscientific issues education. *Sci. Educ.* 89 (3), 357–377.

In-service teacher education for STEM

Julie A. Luft, Mary Francis Early College of Education, University of Georgia, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

In-service teacher education	276
STEM disciplines within in-service teacher education	277
Teacher participation in in-service education	277
STEM in-service teacher education	277
High quality STEM in-service teacher education	278
Knowledge focus	278
Collaborative opportunities	278
Appropriate time and resources	278
In-service program and school alignment	278
Ongoing program evaluation	279
Design decisions associated with STEM in-service teacher education	279
Emerging developments in STEM in-service teacher education	280
Studying STEM in-service teacher education	281
References	281
Further reading	282

In-service teacher education

Teachers are consistently improving their classroom practice to improve student learning. As preservice teachers, they build an instructional foundation through coursework and curated experiences with students. As in-service teachers, the learning opportunities become more tailored to ensure their ongoing professional growth. These learning opportunities vary significantly and are referred to as in-service teacher education, staff development, professional development, continuous professional development, or professional learning.

As learning opportunities, in-service teacher education programs are often developed with attention to local or national standards or reform documents. They have the goal of assisting teachers in attaining the vision of these documents. In-service teacher education programs can also be structured to help teachers advance professionally, which includes new certifications to be a master teacher, learning how to mentor, or becoming a teacher leader. The different pathways that in-service teachers embark upon are often a result of their instructional needs or their professional goals.

In-service teacher education programs vary significantly in format and by provider. Short-term workshops can allow teachers to learn new skills, while demonstration lessons can highlight an instructional approach. Year-long programs can provide teachers with opportunities to develop, implement and revise curriculum. Faculty in higher education can offer in-service education programs that share new pedagogical or content knowledge, or new instructional approaches. Additionally, school and district educators can provide programs to support teachers in their instructional compliance. Teachers can also form their own learning communities to enhance their classroom practice. Ultimately, the different providers and formats of in-service programs are dependent upon the information that will be shared with the teacher.

It is through these designed programs that teachers are provided with opportunities to enhance their knowledge, instructional practices, skills, or cognitive schema (e.g., beliefs, mindset, attitudes). As teachers engage in in-service education programs, they acquire new knowledge, skills, practices, or attributes that better support the learning of their students. Even though the programs are designed to support learning in specific areas, teachers learn differently. The prior knowledge and experiences of teachers influences how they understand and use presented material. Ideally, in-service programs are designed to meet the learning variations of teachers, monitor the learning of teachers, and adjust to improve the learning of teachers. The ultimate outcome of in-service education is improved student learning (see Fig. 1).

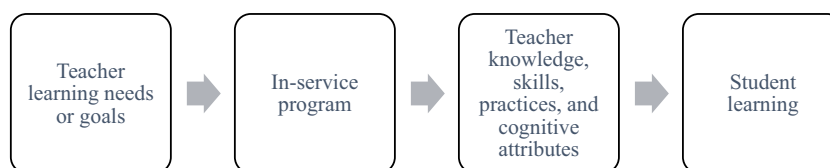


Fig. 1 Framework for in-service education teacher program.

STEM disciplines within in-service teacher education

The discipline associated with an in-service program influences the composition of the activities and the approach of the instructor. Design decisions can involve how to support the subject matter knowledge development of a teacher. A disciplinary approach can emphasize the logical structures of a discipline (Schwab, 1978, p. 246). This syntactical approach varies between different disciplines. In investigative science, for instance, there would be an emphasis on the methods of verification and justification, the use of inquiry, the scientific practices, or the process skills. A disciplinary approach can also be conceptual (Schwab, 1978, p. 246). This substantive approach emphasizes the conceptual connections between and within disciplines. An example would be when the concept of energy is explored or described within biology and physics in a way that is unique to these disciplines, yet highlights the capacity to do work.

Design decisions can also involve the subject matter knowledge that is needed for teaching. Shulman (1986, 1987) first characterized this knowledge as pedagogical content knowledge (PCK), which is unique to teachers and important in transforming subject matter knowledge into learning experiences for students. This teacher knowledge construct has been modified to account for different disciplines. In the sciences and engineering, PCK is considered to be topic-specific, and can include knowledge of students, curriculum, instructional strategies, or assessments (Neumann et al., 2019). In mathematics, the mathematical knowledge for teaching (MKT) construct focuses on common content knowledge, knowledge at the mathematical horizon, specialized content knowledge, knowledge of content and students, knowledge of content and teaching, and knowledge of the curriculum (Ball et al., 2008). In technology, teacher knowledge frameworks are grounded in technological pedagogical content knowledge (TPCK) (Koehler and Mishra, 2009). Technology is viewed as a nimble construct that is responsive to the subject and classroom.

The disciplines of science, mathematics, engineering, and technology (STEM) add to the complexity of the program design. A review of STEM in-service programs found that teacher knowledge and instructional approaches were supported in different ways (Luft et al., 2020). One type of STEM program involved a focus on one disciplinary area and then included another disciplinary area. An in-service program with a focus on science, but opportunities to learn mathematics to explain the phenomenon would be an example. Another type of STEM in-service education program emphasized the different STEM disciplines through a problem or project-based instructional approach. Finally, a few STEM in-service programs blended the different disciplines. In-service education programs on biotechnology or environmental science are examples. Technology had the greatest variation in use across the different configurations of STEM in-service programming.

Teacher participation in in-service education

Data from the Teaching and Learning International Survey (TALIS), which is sponsored by the Organization for Economic Cooperation and Development (OECD), provides information about teacher in-service education participation. TALIS is an important source of data because it is a comparative international education study that collects data on a nationally representative sample of lower grade secondary teachers. By 2018, over 260,000 teachers from 49 countries responded to the survey.

The 2018 TALIS data revealed that teachers participated differently in-service education activities. Among the entire group of teachers, only 49% participated in any formal or informal program, and among new teachers only 22% had an assigned mentor. Schleicher's (2020) report on TALIS also revealed, by country, the number of in-service education activities teachers participated in ranged from 7 to 2.5 in-service activities over a 12-month period. Singapore was at the high end of activity participation by teachers, while France was at the low end. The activity with the greatest amount of participation by teachers (at approximately 75%) was attending courses or seminars and reading professional literature. The in-service activity with the least amount of participation by teachers involved training on peer/self-observation and coaching, and networking.

Scheicher (2020) also pointed out important areas of benefit and need in terms of in-service education. According to the teachers, the most useful activities were those that connected directly to the content they were teaching and involved collaborative approaches to instruction. Teachers also reported that they wanted more support in working with students with special needs, using information and communication technologies, and teaching diverse groups of students.

STEM in-service teacher education

There has been a growing emphasis across the globe in STEM, which is closely related to building economies that are innovative and have the potential of growth (OECD, 2016). A result of this vision are changes in national standards and curriculum that advocate for more individual and integrated STEM coursework. In the United States, for example, the national science standards point to the importance of mathematics and engineering (National Research Council, 2013). Enacting the vision of national/regional standards or curriculum requires that teachers have opportunities to improve their knowledge and instructional practices through different STEM in-service education programs.

At the school level, these standards and curricula are translated into courses that vary in the integration of different STEM disciplines. Some courses are still primarily one disciplinary area with occasional connections to other disciplinary areas. New courses

are emerging that involve more opportunities for integration, which range from one area leading the integration effort to blended ways of integration. Teachers who are faced with teaching these courses will need STEM in-service education programs that support their knowledge and instruction of these integrated courses.

Another challenge in the educational terrain is a shortage of STEM teachers. Some countries are experiencing a shortage of STEM teachers because fewer students are pursuing the teaching profession, or because STEM teachers are leaving the profession in high numbers. With this shortage, STEM in-service teachers are finding themselves assigned to the instruction of STEM courses in which they do not have an adequate background. This phenomenon is referred to as teaching out-of-field and is noted as a global problem (Hobbs and Törner, 2019). Among STEM teachers, out-of-field teaching can range from a few classes to all of their classes. When teachers are out-of-field, they can be completely out-of-field (e.g., mathematics teacher assigned to teach biology), or they can be assigned to teach a discipline in which they have some background (e.g., mathematics teacher assigned to teach physics).

Teaching out-of-field can happen to secondary teachers and elementary teachers. It is more noticeable among secondary teachers because of their qualifications in a subject specific area. Among elementary teachers, out-of-field teaching occurs when the teacher is reassigned to new grade level. The shift in grade level results in an elementary teacher being unfamiliar with the standards, content or curriculum. When STEM teachers are out-of-field, in-service education becomes a necessary way to improve their content background and instruction.

High quality STEM in-service teacher education

STEM education researchers have studied data from in-service education programs to determine the impact of the program on teachers and ultimately students (e.g., Garet et al., 2001; Heck et al., 2008; Johnson et al., 2017). Many of these studies are the result of an analysis of in-service program data associated with the teachers and students. The results of these studies point to design features that should be considered when developing in-service education programs, which support teacher and student learning. The different features have been described previously in Luft et al. (2020). The following paragraphs are a revised version of Luft et al. (2020).

Knowledge focus

STEM teachers benefit significantly when the in-service program has a subject matter knowledge or PCK focus. STEM teachers can enhance their own subject matter knowledge to support their classroom instruction. By improving their subject matter knowledge, STEM teachers improve their depth of understanding of the content area and increase their flexibility with the content they are teaching. When the knowledge focus pertains to student learning within a subject matter area, STEM teachers are better able to respond and support the developing knowledge of students.

Collaborative opportunities

When teachers work together to learn more about a content-based instructional approach or conceptual area, they have opportunities to challenge and refine their own understandings of these areas. Teachers need to collaborate when they are learning new knowledge and instructional approaches, which will enable them to elicit, challenge, contemplate, and reformulate their own understandings about the focal areas of the in-service program.

Appropriate time and resources

Not all content areas take an equivalent amount of time to learn. Some topics are skill based and easy to learn for a teacher. Learning to use a laboratory device, for instance, in the teacher's content area of expertise will not require extensive time. New content areas or new forms of instruction can require more time to learn. STEM teachers need time to experience these areas as learners, undertake them in their classrooms, and then assess the impact of their new knowledge or instructional approaches with students. Ample time ensures that the teachers and students benefit from the in-service education program. In addition, it is important that STEM teachers have adequate resources to use when they return to their classrooms. These resources can include instructional materials, the support of colleagues, or the knowledge or skills of the teacher. In the absence of adequate resources the impact of the in-service education program on the STEM teacher will be diminished over time.

In-service program and school alignment

The content of the in-service program needs to be coherent with the teachers' workplace. This type of coherence exists between the focus of the in-service education program and the instructional content area of the teacher, the curricular materials that the teachers are using in their classrooms, and the standards of the school, district, or region. When there is coherence, teachers are more likely to use the content presented in the in-service education programming. Without coherence, it can be challenging for teachers to implement the in-service education material in their classrooms.

Ongoing program evaluation

The teachers' learning process should be constantly monitored during an in-service education program. The ongoing evaluation of teacher learning is essential in ensuring the in-service program is meeting the learning needs of the teachers. This evaluation involves a pre-assessment of the teachers in terms of their prior knowledge and specific learning needs in terms of the schools' standards. It also involves monitoring the STEM teachers throughout the program to determine their acquisition of the presented material. This can include an evaluation of the teachers using the in-service program material in their classrooms. As this data is collected and examined, there should be changes in the in-service program. Finally, there should be a summative evaluation to determine the attainment of the in-service education program goals.

In summary, high quality STEM in-service education programs are purposefully crafted to support teacher learning and to align with their working conditions. Throughout the program there is ongoing attention to the learning of the teacher, which may result in adjustments in the in-service program.

Design decisions associated with STEM in-service teacher education

STEM in-service teacher education is a multi-dimensional construct. It begins with the professional needs of the teachers and their goals. These are often determined by the teachers or in consultation with administrators. Identified needs or goals should ensure that teachers continue to grow professionally as classroom educators. Ideally, these needs or goals are shared with in-service education providers before they design the STEM in-service teacher education program.

There are three essential dimensions to consider when contemplating the design of a STEM in-service teacher education program. The first decision pertains to the focus on STEM subject matter; specifically, how will the STEM disciplines be integrated? Different conceptualizations of STEM have been suggested by scholars in the field. Honey et al. (2014) suggested that integration involves concepts and practices. Moore et al. (2020) concluded that STEM integration connects to real world problems, involves connecting disciplines by ideas and skills, and has varying degrees of integration. Dare et al. (2019) reported that STEM teachers conceptualized STEM as separate disciplines, used STEM when solving real-world problems solving, or integrated STEM for problem solving.

In one of the few examinations of STEM in-service programs, Luft et al. (2020) found different representations of disciplinary integration (see Fig. 2). The disciplines could be siloed with each discipline emphasized during the program. This type of program often provided specific content instruction for teachers in different disciplinary areas. When STEM in-service programs had two or more disciplines, the disciplines could be: sequential (i.e., one taught after the other); embedded (i.e., one in service of the other); or connected (i.e., two or more integrated to represent a new discipline). Sequential programs involve teachers learning to use problem based or project-based learning and emphasize the individual STEM disciplines. Embedded programs use one or two of the disciplines (e.g., science or engineering) as a setting for additional disciplinary knowledge. Connected programs involve learning new approaches or knowledge that are a result of the disciplinary integration of specific STEM areas (e.g., biochemistry, biotechnology).

There are also variations in the origin of a STEM in-service education program, which pertains to the people offering the in-service program. Some programs can originate outside of a school setting, but still align with the school curriculum or standards. These programs can be developed by individuals in higher education who are testing or promoting new knowledge or instructional approaches, those in informal learning environments, or industry professionals. National STEM organizations can create programs (e.g., conferences, workshops) that support the learning of teachers. Programs can also originate from school administrators who are interested in improving the instruction of their teachers and the learning of their students. Finally, programs can originate from teachers who want to improve their own educational environment. These types of programs allow teachers to support their own learning by examining, for instance, the learning of their students during different activities. Often, the origin of each of these programs contributes differently to the learning of STEM teachers.

There are also variations in the professional learning objectives of STEM in-service education programs, which relates to what the teacher will learn. Programs are often designed to help teachers improve their content knowledge, PCK, or instruction. These programs can include learning additional subject matter knowledge, new instructional approaches, or about curricular changes. There are also programs that target the different instructional needs of STEM teachers. For instance, there are programs for teachers who are teaching out-of-field, or teachers who are new to the field. There are also programs that encourage the professional advancement of STEM teachers, by allowing teachers to build their abilities to mentor or supervise other STEM teachers, or cultivate their leadership capacity in STEM.

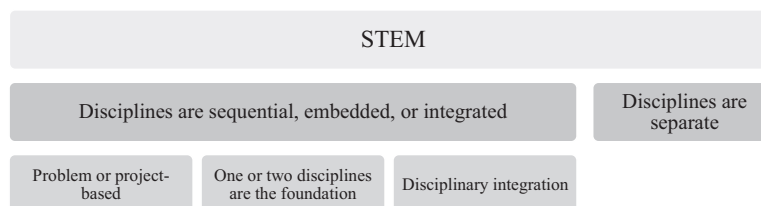


Fig. 2 The representation of STEM in in-service education programs.

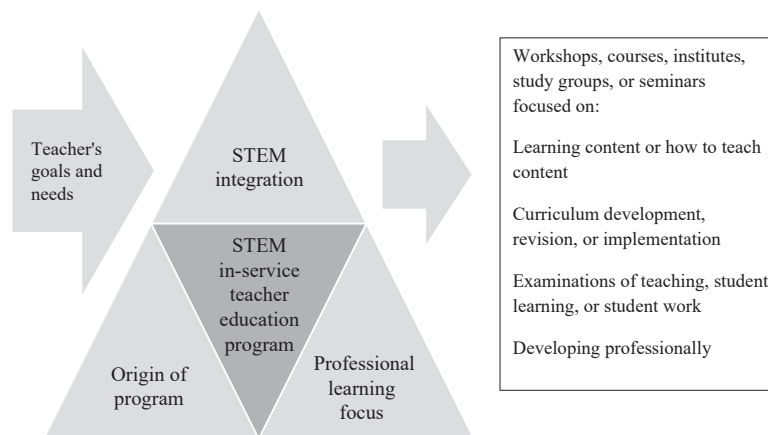


Fig. 3 Dimensions of STEM in-service teacher education.

Fig. 3 shows how the STEM in-service education program format and duration is a result of how STEM is integrated; who is supporting/offering the program; what is the learning focus. There are different program formats that range from short-term workshops, to extended course work, to yearlong professional learning communities/study groups (see [Loucks-Horsley et al., 2009](#)). Within these different formats, STEM teachers can grow professionally in different ways. They can improve their disciplinary knowledge and instruction, they can work to align or revise the curriculum within their school, or they can study student work to support their own instruction. They can also grow professionally as a mentor of other STEM teachers or learn how to guide the implementation of STEM curriculum.

Emerging developments in STEM in-service teacher education

STEM in-service teacher education is expanding as more individuals become involved in the development and implementation of these programs. Some of the emerging developments offer teachers new ways to contemplate the enactment of STEM in their classrooms. Other developments pertain to the design of the in-service education program. Some of these emerging areas already involve STEM teachers and some are well-positioned to be adapted to STEM in-service education. Ultimately, these are new and emerging areas that provide new ways to enhance the learning of STEM teachers.

An area that is expanding in STEM in-service teacher education pertains to culturally responsive or equitable instruction. These programs have orientations that expand the participation of students in STEM. The culturally responsive approach recognizes that groups of students have been historically marginalized in schools but offers a vision that encourages the success of all students, while affirming their experiences, identities, and histories ([Brown, 2017](#)). An equitable orientation improves educators' actions in order to improve students access to STEM, which has been hindered because of factors such as, race, class, ability, religion, ethnicity, disability, appearance, gender, language, or immigration status ([Bancroft and Nyirenda, 2020](#)). Infusing STEM in-service education programs with culturally responsive or equitable principles enhances how STEM is presented to teachers and students. For example, programs can highlight how different STEM advancements are a result of different cultures, or programs can emphasize ways to make STEM instruction more gender inclusive. The different ways that programs are configured to be more responsive or equitable are often a result of the theoretical framing associated with the STEM in-service education program.

Within the STEM online in-service education environment there are different ways in which teachers can advance their instruction. Teachers can receive information directly from online programs that are designed to improve their instruction. These formal programs are designed with specific learning objectives. There are also informal online programs that consist of networks, sharing platforms, or other collaborative communities. Within these communities, STEM teachers may share information, create different media (e.g., podcasts, blogs), or co-create material (e.g., wikis, collaborative media) ([Dede and Eisenkraft, 2016](#)). The platforms that support these different programs range from computers, to tablets, to mobile devices. Within these different platforms and program orientations, and an adequate internet connection, the potential for teacher learning is unlimited.

Personalized learning is another emerging area within STEM in-service education. Within a personalized environment, the pace of learning and instruction are optimized through the monitoring of teacher learning and the differentiation of the program material ([Walkington and Bernacki, 2020](#)). In these environments, teachers engage in just the material they need to achieve a specific outcome. In-service education programs are moving toward programming that can be readily adapted to the learning of teachers. The online environment will be important in developing these personalized learning environments. For STEM teachers who are tackling new content and instructional approaches, these types of programs will be essential. In a hybrid or online format, teachers can be easily directed to different activities, which can better support their learning needs. Ultimately, this may result in reduced professional development time, or supplemental learning opportunities for teachers. For in-service teacher education providers, the proliferation of educational material on the internet is a resource that can support personalized programming.

Another area that is emerging in STEM in-service education pertains to the development of the teacher as a person. Bell and Gilbert (1996) initiated this discussion in the science education community when they suggested that teacher development involved social, personal and professional development. The social component involves reframing one's view of being a teacher, which occurs as a teacher interacts with colleagues. The personal component involves navigating the emotions and knowledge that come with being a science teacher. The professional development component involves how teachers draw upon, enact, and support the different activities they use in the classroom. Korthagen (2017) called for more attention to the person in the in-service education experience. This involves planning for the cognitive, emotional, and motivational dimensions of learning during the in-service education program.

Studying STEM in-service teacher education

The global expansion of STEM in-service teacher education is underway. As teachers engage in these programs, there will be opportunities for research that will shed light on the learning of teachers and students. Within the variation of studies that are planned and carried out, some of the most important studies will be conducted to examine the design of different STEM in-service teacher education programs, study STEM teacher and student learning, and consider the role of context during and after STEM in-service education programming. Within these broad areas, there are some priority research areas.

One of the more pressing questions in STEM in-service education pertains to how teachers learn and enact the STEM disciplines. Most in-service education programs are often configured to build a teacher's knowledge and practices in one discipline. The addition of STEM adds a dimension of complexity to an in-service education program but is essential in supporting STEM teachers. One type of question that is important pertains to understanding how an integrated STEM in-service education program improves the disciplinary knowledge of a STEM teacher. It may be that one disciplinary area acts as an anchor for the teacher's learning of other disciplinary areas, or that disciplinary areas are constructed independently of each other. Additional questions that are focused on teachers can involve their emotions or identities when learning about STEM.

Another pressing question pertains to the representation of STEM knowledge in the classroom as a result of the in-service teacher education program. Understanding the supports in a program that can help a teacher enact STEM lessons will be important. For instance, it may be that there are high leverage professional development activities, which can support STEM teachers in enacting STEM lessons. For instance, more change in the instruction of a teacher may result from studying instructional materials than does viewing and analyzing videos of successful instruction, or that building one's interdisciplinary knowledge is essential before a lesson can be enacted with students.

Attending to context is important in the enactment of STEM in-service education programs. School policies, students, or the school community can impact how teachers implement STEM instruction in their classrooms. For instance, it is important to understand how teachers enact novel STEM instructional approaches in their classrooms after an in-service program. It may be that student enthusiasm can support the use of novel instructional approaches. It is widely acknowledged that context is important, but an increased student interest in STEM may be a new dimension of context.

Within these areas of study, there should always be a focus on the teacher, as this is what will help inform STEM in-service teacher education. As these studies are designed, they can be exploratory and make initial statements. This type of study allows for new ideas or approaches to be investigated and conceptualized for potential examination. With some evidence from exploratory studies, impact studies can be conducted that elucidate specific conditions, events, or attributes. Impact studies that accumulate additional evidence can eventually result in effectiveness studies. These studies can compare and contrast specific conditions, events or attributes at a larger scale and demonstrate an impact on teacher learning. These studies do contribute to our expanding knowledge base and are important in future work with teachers in the area of STEM.

References

- Ball, D.L., Thames, M.H., Phelps, G., 2008. Content knowledge for teaching: what makes it special. *J. Teach. Educ.* 59 (5), 389–407.
- Bancroft, S.F., Nyirenda, E.M., 2020. Equity-focused K-12 science teacher professional development: a review of the literature 2001–2017. *J. Sci. Teach. Educ.* 31 (2), 151–207.
- Bell, B., Gilbert, J.K., 1996. *Teacher Development: A Model from Science Education*. Falmer Press, Bristol, PA.
- Brown, J.C., 2017. A meta synthesis of the complementarity of culturally responsive and inquiry-based science education in K-12 settings: implications for advancing equitable science teaching and learning. *J. Res. Sci. Teach.* 54 (9), 1143–1173.
- Dare, E.A., Ring-Whalen, E.A., Roehrig, G.H., 2019. Creating a continuum of STEM models: exploring how K-12 science teachers conceptualize STEM education. *Int. J. Sci. Educ.* 41, 1701–1720.
- Dede, C., Eisenkraft, A., 2016. Online and blended teacher learning and professional development. In: Dede, C., Eisenkraft, A., Frumin, K., Hartley, A. (Eds.), *Teacher Learning in the Digital Age: Online Professional Development in STEM Education*. Harvard Education Press, Cambridge, MA.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., Yoon, K.S., 2001. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38 (4), 915–945.
- Heck, D.J., Banilower, E.R., Weiss, I.R., Rosenberg, S.L., 2008. Studying the effects of professional development: the case of the NSF's local systemic change through teacher enhancement initiative. *J. Res. Math. Educ.* 39 (2), 113–152.
- Hobbs, L., Törner, G. (Eds.), 2019. *Examining the Phenomenon of "Teaching Out-of-Field": International Perspectives on Teaching as a Non-specialist*. Springer, Singapore.
- Honey, M., Pearson, G., Schweingruber, H.A. (Eds.), 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*, vol. 500. National Academies Press, Washington, DC.

- Johnson, C.C., Sondergeld, T.A., Walton, J., 2017. A statewide implementation of the critical features of professional development: impact on teacher outcomes. *Sch. Sci. Math.* 117 (8), 341–349.
- Koehler, M., Mishra, P., 2009. What is technological pedagogical content knowledge (TPACK)? *Contemp. Issues Technol. Teach. Educ.* 9 (1), 60–70.
- Korthagen, F., 2017. Inconvenient truths about teacher learning: towards professional development 3.0. *Teach. Teaching* 23, 387–405.
- Loucks-Horsley, S., Stiles, K.E., Mundry, S., Love, N., Hewson, P.W., 2009. *Designing Professional Development for Teachers of Science and Mathematics*, third ed. Corwin press, Thousand Oaks, CA.
- Luft, J.A., Diamond, J., Zhang, C., White, D.Y., 2020. K-12 STEM professional development programs: building teacher knowledge and practice through the lens of transfer. In: Johnson, C., Schoeder, M., Moore, T. (Eds.), *Handbook of Research on STEM Education*. Routledge, London, pp. 361–374.
- Moore, T.J., Johnston, A.C., Glancy, A.W., 2020. STEM integration: a synthesis of conceptual frameworks and definitions. In: Johnson, C., Schoeder, M., Moore, T. (Eds.), *Handbook of Research on STEM Education*. Routledge, London, pp. 3–16.
- National Research Council, 2013. *Next Generation Science Standards: For States, by States (NGSS)*. The National Academies Press, Washington, DC.
- Neumann, K., Kind, V., Harms, U., 2019. Probing the amalgam: the relationship between science teachers' content, pedagogical and pedagogical content knowledge. *Int. J. Sci. Educ.* 41 (7), 847–861.
- Organization for Economic Cooperation and Development (OECD), 2016. *Trends Spotlight #15: A Brave New World: Technology and Education*. Trends Shaping Education 2016. OECD Publishing, Paris.
- Schleicher, A., 2020. *Teaching and Learning International Survey TALIS 2018: Insights and Interpretations*. OECD Publishing, Paris.
- Schwab, J.J., 1978. Education and the structure of the disciplines. In: Westbury, I., Wilkoff, N.J. (Eds.), *Science Curriculum and Liberal Education: Selected Essays of Joseph J. Schwab*. University of Chicago Press, Chicago, IL, pp. 229–272.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educ. Res.* 15 (2), 4–14.
- Shulman, L.S., 1987. Knowledge and teaching: foundations of the new reform. *Harv. Educ. Rev.* 57, 1–22.
- Walkington, C., Bernacki, M.L., 2020. Appraising research on personalized learning: definitions, theoretical alignment, advancements, and future directions. *J. Res. Technol. Educ.* 52, 235–252.

Further reading

- Borko, H., 2004. Professional development and teacher learning: mapping the terrain. *Educ. Res.* 33 (8), 3–15.
- Desimone, L.M., 2011. A primer on effective professional development. *Phi Delta Kappan* 92 (6), 68–71.
- Kennedy, M.M., 2016. How does professional development improve teaching? *Rev. Educ. Res.* 86, 945–980.

International perspectives on STEM teacher induction

Jacob Elmore, Rachel van Aswegen, and Peter Youngs, University of Virginia, School of Education and Human Development, Charlottesville, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Conceptual framework	284
Induction programs in Singapore, Japan, Shanghai, Australia, and France	284
Singapore	284
Japan	285
Shanghai (China)	285
Australia	286
France	286
Research on mentoring and induction programs that promote culturally relevant teaching	286
Mentoring and induction programs and beginning STEM teacher effectiveness	288
Conclusion	289
Appendix	290
References	290

In the 2015 Trends in International Mathematics and Science Study (TIMSS), 4th- and 8th-grade students in several Asian countries performed considerably higher on average on assessments of mathematics and science knowledge than their counterparts in the United States and a number of Western European nations (Provasnik et al., 2016). In addition, on the 2018 Program for International Student Assessment (PISA), 15-year-olds from Asian nations scored much higher on average on assessments of mathematics and science knowledge than 15-year-olds in the US and most Western European countries (Organization for Economic Co-operation and Development [OECD], 2019b). While numerous factors likely contribute to these differences in student performance, in this article we examine the nature of mentoring and induction programs for beginning STEM teachers in three Asian countries (i.e., Singapore, Japan, and Shanghai [China]) and compare them to such programs for novice STEM teachers in Australia and France. We also draw on research from the US and New Zealand to investigate ways that mentoring and induction programs can help prepare beginning teachers to work effectively with culturally diverse students. Finally, we review research from the US on characteristics of mentoring and induction programs that are associated with novice STEM teacher effectiveness.

In this article, *STEM teachers* are defined as (a) science, technology, engineering, and/or mathematics teachers who work at the secondary level or (b) teachers who teach mathematics and/or science at the elementary level, including multiple-subject teachers who also teach reading/language arts and/or social studies. We use the term *induction programs* to refer to formal support programs for 1st-year teachers and sometimes 2nd- and 3rd-year teachers that are provided by school districts, schools, or other organizations. Such programs typically feature formal mentoring, orientations prior to the start of the school year, and regular professional development activities. Induction programs sometimes include mentor training, release time for 1st-year teachers and mentors, and formative or summative assessments of beginning teachers. In contrast to *induction programs*, the term *induction* refers to the career stage that usually comes after pre-service teacher education. This period often lasts about two to four years or until the beginning teacher achieves more permanent employment status and/or a regular/professional teaching certificate or license. Individuals who enter teaching through alternative certification pathways and become full-time teachers of record while they are completing education coursework are often in the preparation and induction stages simultaneously (Youngs et al., 2019).

In the first section of this article, we present a conceptual framework that outlines ways in which mentoring and induction programs for beginning STEM teachers across national contexts can promote instructional quality, culturally relevant instruction, and teacher effectiveness. The second section describes mentoring and induction programs in Singapore, Japan, and Shanghai, and compares them to such programs in Australia and France. We argue in this section that mentoring and induction in Asian countries tend to focus on centralized standards for teaching practice; collective responsibility for improvement; professionalizing the career of teaching; connecting theory to practice; and coordinating teacher preparation, induction, and evaluation in relation to national standards. In contrast, such programs in Western nations are typically decentralized and place less emphasis on centralized standards for instruction, professionalizing teaching, or collective efforts to improve instruction.

In the third section, we review research on mentoring and induction programs focused on supporting effective instruction for diverse learning. In this section, we contend that in order to support novice teachers' work with diverse learners, mentors themselves need to have both knowledge about and experience working effectively with diverse students. The fourth section describes characteristics of mentoring and induction programs that promote beginning teacher effectiveness. Here we present evidence from US based research that two years of comprehensive induction has a stronger effect on beginning elementary teacher effectiveness than less intensive approaches to mentoring and induction (Glazerman et al., 2010; Schmidt et al., 2017).

Conceptual framework

Teaching is viewed in different ways throughout the world, and there are multitudinous ways in which teachers enter and are inducted into the profession. Examining different nations' approaches to beginning and continuing teaching yields many insights about the roles that teachers play internationally; such analysis can also turn inward and provide perspective on one's own system of teacher education, induction, and professional development.

It can be useful to use global-comparative exams such as TIMSS and PISA to examine educational practices in countries that perform at the highest levels on these exams. Both of these exams include components specific to mathematics and science achievement. On both tests, in both mathematics and science, for all grades tested, Singapore, Japan, and China are all among the top five countries while Australia and France are mid-level performers (Mullis et al., 2016; OECD, 2019b; Provasnik et al., 2016). There are many factors that affect a country's score on these international exams, including curricula, teacher compensation, and the mindsets of teachers and students. For example, Australian students demonstrate academic resilience across socio-economic brackets, which OECD, the managing body for PISA, attributes to positive school climates and high levels of growth mindset (OECD, 2019a).

While directly connecting specific teacher induction practices to student achievement has posed a major challenge to researchers (for exceptions, see Glazerman et al., 2010; Schmidt et al., 2017), it is valuable to examine themes that emerge related to teacher induction in regions and countries achieving progress and success on international exams. For the purposes of this article, we will examine how teachers in Singapore, Japan, China (Shanghai), Australia, and France are supported as they enter the classroom. Broadly speaking, there is consistency in the three Asian countries when it comes to emphasizing collective responsibility for improvement; professionalizing the career of teaching; connecting theory to practice; and coordinating teacher preparation, induction, and evaluation in relation to national standards. The second and third sections of our article consider several core questions outlined by Britton et al. (2003) as they relate to these themes. First, they ask about the focus of induction systems. Second, they examine the curriculum of induction systems and whose knowledge is being shared with beginning teachers. Third, they ask about the structure of the overall system of induction (Britton et al., 2003). In the fourth section, we address a different question raised by Britton and colleagues: whether or how teacher induction contributes to student achievement.

In a broad study of induction in OECD countries, teacher participation in induction and in-service professional development programs contributed to the mean overall country score for teacher professionalism (Schleicher, 2016). In correlating these scores with PISA results, ultimately, there was a weak association ($r = 0.285$), probably because isolating information about any one component of teacher training, professionalism, and induction on a national level is difficult to do. Even comparing the rates of access to and participation in formal induction programs with a country's PISA scores does not yield a conclusion about the association between teacher induction and student achievement using PISA data. This more recent data suggests that Britton et al.'s conclusions about the connections between teacher induction and student achievement still being tenuous, complicated, and lacking sufficient variability to make distinctions among programs have remained true over the past two decades in most countries (Britton et al., 2003; OECD, 2019a).

Table 1 summarizes the structure of induction programs in each of the countries examined in this article. Structure refers to the overall format and organizing body of induction programs and standards; level of experience targeted refers to when in a teacher's career induction takes place; focus and curriculum refer to the content of induction programs; and purpose refers to the reasoning provided by governments and officials for providing induction programs.

Induction programs in Singapore, Japan, Shanghai, Australia, and France

Singapore

Though a direct link between teacher induction practices and student performance is hard to demonstrate, Singapore has had a great deal of success on international testing for the past two decades, showing consistently high and increasing scores from 1999 to 2015 for both mathematics and science in grades 4 and 8 (Mullis et al., 2016; Schleicher, 2019; Wang et al., 2003). In the system Singapore has developed since its independence in 1965, teachers are generously supported both financially and socially (Darling-Hammond, 2017) and nearly all beginning teachers in Singapore have access to and participate in induction programs when they begin teaching (Schleicher, 2016). Initially, novice teachers spend their first three days on the job at a nationally-organized induction program called the Beginning Teachers' Orientation Program (Schleicher, 2016).

At this training, novice teachers are inducted formally into the professional "teaching fraternity" (Schleicher, 2016, p. 50). Further induction programs continue during the teachers' first two years that feature nationally-trained mentors, in-service training, and a peer-buddy system, which teachers can take advantage of due to a reduced teaching schedule. These courses and peer and mentor reflections focus on equipping beginning teachers with both the content knowledge and practical skills to handle the pressures of classroom management, instruction, and assessment needed to be a successful teacher (Darling-Hammond, 2017). Throughout a teacher's career, Singapore's Ministry of Education constantly monitors their performance in the areas of knowledge, skills, and attitudes; and provides feedback for improvement and opportunities for advancement based on their performance. These reviews hold greater weight as teachers advance in their careers (Schleicher, 2016). In this way, Singapore's teachers are encouraged to be teachers of high caliber at every career stage.

From initial teacher recruitment to curricula designed by teachers and continuous professional development throughout their careers, Singapore's system is focused on the national ethos of establishing and maintaining excellence in their classrooms

(Darling-Hammond, 2017; Schleicher, 2016). It should also be noted that these programs are coordinated and financially supported at the national level (Darling-Hammond, 2017). There is, simultaneously, a great deal of autonomy and professional expectation for teachers, as schools and mentors are left to direct their own teacher induction programs by defining problems and new practices that will best fit their school environment to address the identified problems. This culture of professional excellence is fostered by teachers and further promotes collaboration within the field toward nationally-defined goals (Schleicher, 2016).

Japan

The overall structure for teacher induction in Japan includes Ministry of Education, Culture, Sports, Science, and Technology (MEXT) guidelines for a one-year induction period for beginning teachers, involving 120 h of in-school professional development and 180 h (one day per week) of out-of-school training (OECD, 2019a; Schleicher, 2019). Professional development in the school consists of working with mentor teachers, subject specialists, and the school principal to address issues related to academic instruction, student guidance, health and safety, and homeroom paperwork. Off-site professional development addresses pedagogical theory, visits to social service agencies and other institutions connected to schooling and student well-being, and volunteering (OECD, 2018).

Japanese education boards prioritize professional development for novice teachers to train them within their school contexts and prepare them for persistent challenges that they face in schools (Fujimara and Mistilina, 2020; OECD, 2019a). The strengths-weaknesses-opportunities-threat analysis of the program, conducted by the OECD, revealed that induction programs following the MEXT guidelines face threats of demographic change and the need to balance autonomy and control with regard to teaching and learning. These threats are managed successfully through reforms focused on excellence, strong subject knowledge, and professionalism of teachers.

These structural shifts have happened during a time when Japan has shown major growth in mathematics and science achievement. Japan's 4th and 8th graders demonstrated a steady increase from the mid-550s up to the 600 range on TIMSS science assessments between 1995 and 2015 (Mullis et al., 2016). During the same period, grade 4 mathematics scores increased in the same range, and grade 8 students mathematics scores remained consistently in the 500s, meaning Japan has been, since inception of the TIMSS testing, a top-scoring nation. Further, when teachers were asked about their participation in professional development, they reported above-average rates of participation in learning about mathematics and science content and pedagogy.

Beyond the training that beginning teachers participate in during their first year, Japanese teachers of all experience levels engage in a "collective enquiry process" referred to as lesson study (Schleicher, 2016, p. 65). Lesson study is designed to support a collaborative culture toward self-reflection and improvement (OECD, 2018). In this process, teachers at all levels of teaching experience work together to build their reflective capacity through research, planning, teaching, and observation. Lessons observed by a team of teachers are used to examine ways for the collective group to improve their practice, supporting collaboration among team members. In this way, the team of teachers is setting the agenda for their own learning, and building on one another's ideas, skills, teaching, and research.

Shanghai (China)

In Shanghai, China, novice teachers' access to and participation in induction programs hovers right around 100% (Schleicher, 2016). The regions of Shanghai, Beijing, Jiangsu, and Zhejiang are separated from the rest of China in the research literature because their 15-year-olds have outperformed all of the other 78 countries participating in PISA on their own. While this group of students is not representative of China as a whole, in an analysis of the 2018 administration of PISA, Schleicher (2019) noted that the population of each of these four regions is similar to that of the other countries participating in the exams. For these reasons, researchers have singled out each region in research, and Shanghai demonstrates particularly compelling teacher induction practices.

Overall, the induction programs in Shanghai are centrally coordinated, requiring novice teachers to participate in 240 h of training in their first five years of teaching, and rewarding more senior teachers with leadership positions in their schools when they participate in 540 h of training over five years at any point in their career. These opportunities to learn are coupled with a lower average number of hours teaching (i.e., 10–12 h per week) to allow teachers to participate in these programs effectively (Schleicher, 2016). This time is spent with teaching and research groups, focusing on content areas and student learning; and classroom groups, focused on serving the same group of students across disciplines to ensure an equitable experience for all students (Zhang and Kong, 2012).

These regionally-coordinated training sessions work in partnership with teacher preparation colleges to ensure access to quality professional development. Induction focuses on creating a culture of collaboration and learning in schools, through mentoring and learning communities, to develop a robust faculty to take on distributed school leadership roles (Schleicher, 2016). When more senior teachers excel in professional development, they have incentives to take on leadership positions within the programs, which develops a shared sense of purpose among faculty to continuously improve their teaching practice as a collective. Thus training is coordinated at a national level, with local delivery by experts who are also embedded in the schools, ensuring that professional development is highly relevant to the context of a given school.

Shanghai teacher professional development features an improvement cycle focused on student learning using data and evidence through a whole system approach, where all members of the school community work to develop a culture of improvement, rather than isolated policy changes (Schleicher, 2016). Through this professional development, teachers are equipped with teaching

methodologies from throughout the world, such as classroom observation and evaluation structures similar to those in Japan, and improving learning through the Montessori teaching method or the mindset of mastery teaching from Bloom in the US (Zhang and Kong, 2012). This openness to best practices, coupled with a collective mindset to implement these changes, were a driving factor in improvement of students' PISA scores in 2009 (Zhang and Kong, 2012).

Australia

Australian teachers participate at higher-than-average rates in professional development focused on mathematics and science content and pedagogy (Mullis et al., 2016). Their students have consistently scored in the low 500s on TIMSS in science, and have seen slight increases in their mathematics scores from 1999 to 2015 (Mullis et al., 2016). In contrast to Singapore, Japan, and Shanghai, there is no nationally-organized induction program for beginning teachers Australia. At the same time, all novice teachers are required to complete a probationary period of five years; during this time, they must demonstrate pedagogical content knowledge and promote student achievement in the classroom (OECD, 2019a).

Each school or district authority in Australia can determine how induction will be structured for their new teachers. Across the country, there are a handful of schools and jurisdictions working to connect the content of teacher education programs to the challenges of full-time teaching, and address teacher shortages (OECD, 2019a). Because teacher accreditation standards are determined at local levels, this has left room for teachers and local administrators to establish them. This autonomy to set the direction for the profession of teaching increases teachers' levels of professionalism and self-regulation overall, a theme consistent with other high-performing countries. However, because the standards are not centralized, there is often weak feedback for teacher preparation institutions, and a lack of national focus overall (OECD, 2019a). While there are many opportunities for training and professional development for in-service teachers in Australia, there are elements missing from its approach to induction which could propel the country further.

France

Hovering around the OECD country average on PISA tests, teacher preparation in France has undergone major shifts from a university-centric model to a professional-centric model (Cornu, 2015; OECD, 2019a). Fourth-grade student achievement in both mathematics and science on the 2015 administration of TIMSS also showed that French students are close to the international average (Mullis et al., 2016). As this was the first time French students participated in TIMSS, trend data for mathematics and science is not available; PISA data for reading has been flat since 2000 (Schleicher, 2019).

The French induction system for novice teachers has undergone major reform in the past ten years, moving in a direction contrary to the three Asian countries. Until 2010, beginning teachers in France participated in coursework at teacher preparation institutes, met with mentor teachers, and had a reduced teaching load that enabled them to engage in planning and assessing student work. Between 2005 and 2010, though, French teacher preparation institutes became more professionalized (Cornu, 2015; Howe, 2006). Changes included shifting from a specifically induction-oriented structure in the second year after receiving a Bachelor's degree, to a Master's degree program, where future teachers teach a half-load, write their theses, and prepare for qualifying national teaching exams. This meant that France's model of preparation no longer included induction during a teacher's first year of teaching, but rather shorter-term internships in schools, intended to complement coursework (Cornu, 2015).

As this shift was occurring in initial teacher preparation, responsibility for professional development and induction was shifted from the centralized university system to the more decentralized secondary school system (Cornu, 2015). While goals and standards for teacher education are still set at the national level, the requirement that beginning teachers earn Master's degrees has left induction programs greatly reduced, and they are now integrated into overall in-service professional development for all teachers (Schleicher, 2016). These decentralized and unmonitored activities tend to focus on classroom management, whereas previously, centralized induction activities at universities covered a wide range of topics to support "new teachers in managing a full-time job" (Schleicher, 2016, p. 50).

Research on mentoring and induction programs that promote culturally relevant teaching

There has been relatively little research in the area of induction programs designed to help beginning secondary STEM teachers enact culturally responsive practices and work with diverse student populations. At the same time, there are a few studies of novice elementary teachers, who often teach multiple subjects including mathematics and science, and their experiences with district- or state-sponsored induction programs designed to help them work effectively with diverse students. These studies often include mentors' perspectives with regard to supporting beginning teachers to work effectively with diverse student populations. Most of the studies featured in this section were located in urban school districts in the US while one study took place in New Zealand; none of the induction programs featured in this section are necessarily representative of all induction programs in these countries.

This section focuses on three core questions from Britton et al. (2003): What is the focus of a given induction system? What are the curricula of induction systems and whose knowledge is being learned by novice teachers? And, what is the structure of the overall system of induction in a given jurisdiction? While the studies in this section represent different foci and settings, they converge on two particular themes: (a) in order to effectively mentor beginning teachers to work with diverse students, mentors themselves need to both have the experience and deep knowledge base for working effectively with diverse students; and (b) mentors face multiple obstacles that make it challenging to support novice teachers' development to effectively work with diverse students.

To start, Achinstein and Barrett (2004) conducted a two-year qualitative case study of beginning teacher-mentor pairs in northern California in the US. The study featured 12 schools, 15 novice teachers, and 11 mentors. The schools were described as medium-sized and were located in low-income communities with high percentages of English language learners and low-performing students. The induction program represented a collaboration among a university, county office, and district consortium of educators. The program included two years of on-site weekly meetings with mentors and monthly seminars for beginning teachers. The induction program focused on the California Standards for the Teaching Profession (CSTP) and featured goal setting, sharing resources, mentor and novice conversations about lesson planning, mentor classroom observations, post-observation conferences, modeling lessons, and analysis of student work.

Achinstein and Barrett (2004) focused on what frames (beliefs and attitudes) mentors and beginning teachers had toward diverse students; what pattern frames were used in mentor-novice reflection conversations; how frames were used differently by mentors and novices; and how mentors supported them to reframe their views. In their study, the authors drew on audiotapes and transcripts of mentoring conversations, interviews with beginning teachers and mentors, and videotaped classroom observations. The researchers found that novice teachers were generally situated in the managerial frame (classroom control and student compliance). Mentors had to use reframing conversations to help shift their perceptions of how to work effectively with diverse students—which included political (social change) and human relation (relationships between individuals) frames. The mentors who reframed novices' perceptions of students enabled them to see and learn about their students and were able to focus on areas such as curriculum and instruction—which also bolstered novices' classroom management. However, some mentors attended only to classroom management and this limited beginning teachers' practice to support diverse learners. The researchers made it clear that mentors need to be aware of what types of conversations (or frames) predominate mentor-novice conversations and they need opportunities for professional development to develop skills in reflection conversations to reframe beginning teachers' beliefs and classroom instruction.

While mentor teachers need to be able to reframe conversations for beginning teachers, they must also have deep knowledge about working with diverse students. Achinstein and Athanases (2005) investigated the knowledge base that mentors should have with regard to diversity and equity, and what it looks like in practice. They used questionnaire responses of 37 veteran mentors and new teacher program directors in the Leadership Network for Teacher Induction (LNTI) to define this knowledge base. LNTI is a reform network in northern California of induction leaders who support approximately 2750 new teachers in 60 school districts, many of which served high percentages of English Learners. The focus of LNTI included developing members' knowledge of practice through inquiry projects. These projects mainly focused on critical educational issues through learning and collaboration. In particular, LNTI focused on equity, diversity, and anti-racist practices. In analyzing survey data from the induction leaders, the researchers found that mentors needed both a bi-level knowledge base (knowledge of teaching and knowledge of mentoring) and a bi-focal perspective (focus on local and larger perspectives). Induction leaders believed that mentors needed knowledge of what diverse learners bring into classrooms; knowledge of pedagogy related to diversity and equity; and knowledge of oneself related to diversity and equity (Achinstein and Athanases, 2005).

As part of their study, Achinstein and Athanases (2005) also conducted a qualitative case study of one mentor-novice pair to examine what mentors' knowledge looked like in practice and its complexities. The induction program was designed to provide two years of support for beginning teachers with on-site weekly meetings with mentors and monthly seminars. In the case study, the mentor was able to focus on practices related to diversity and equity but faced multiple challenges. One challenge was finding a balance between sharing with their mentee how to improve their practice without reducing their self-efficacy and willingness to share vulnerabilities. Also, the mentee's principal and colleagues did not engage in practices to promote equity or support diverse learners; and this posed another challenge to the mentor. This finding is consistent with other literature that found that mentors not only need to have both the knowledge and skills to mentor novices for diversity, but they also should be able to challenge the structures, cultures, and conflicting messages present in schools that do not necessarily promote culturally responsive teaching or practices to improve equitable outcomes for diverse learners.

School culture, structures, and messages shared with novice teachers present a unique challenge to mentors especially with regard to culturally responsive teaching and promoting equitable outcomes for diverse learners. Dharan (2015) investigated how schools' cultures, practices, and policies seemed to affect beginning teachers' beliefs, attitudes, and practices in relation to student diversity. The author carried out a qualitative case study of five beginning primary teachers located in urban schools serving students of varying socioeconomic backgrounds in New Zealand. Dharan conducted interviews, classroom observations, and observations of meetings between novice teachers and their mentors. It is unclear how often these teachers worked with their mentors or the structure (i.e., frequency and types) of mentoring that took place. At the same time, Dharan explained that observations of mentor-novice conversations focused mainly on "established ways of working" (2015, p. 69). The mentors rarely focused on diverse students or the *why* and *who* and focused more on the *how* and *what* teaching practices already predominated the school culture.

Novice teachers in Dharan's (2015) study had broad perceptions of diversity that were influenced by their preparation program and they understood the importance of culture and individual diversity; the author notes that this finding is contrary to research that finds that many novice teachers are not exposed to nor prepared for diversity. However, after 12–18 months in the study, while these teachers maintained their definition and understanding of diversity, their instructional practices focused on homogenization through ability grouping based on assessments. In addition, their pedagogy focused on academic ability compared to social categories of teaching and learning. Dharan explained that the focus of these teachers on academic ability and homogenization of classroom practices was shaped by their schools' culture and structures that promoted such practices—and this included mentoring conversations. The author maintained that schools (particularly mentors) and teacher preparation programs need to work in conjunction in order to effectively induct novices to work effectively with diverse students (Dharan, 2015).

A study by Yendol-Hoppey et al. (2009) in the area of mentoring teachers for culturally responsive teaching and effectively working with diverse students echoed the studies mentioned above. In their qualitative study of an urban induction program, the authors focused on mentors' perspectives on how they conceptualized their work to support novice teachers, what challenges they faced working with novices, and how they addressed these challenges. The researchers focused on the Urban Mentoring Program (UMP), which was located in a large school district in the northeastern region of the US. The purpose of the UMP was to provide full-time mentors to schools that were in most need of mentoring services. The UMP selected schools based on the following criteria: 50% or more of the student population received free or reduced-price meals and the school had at least 10 first- and second-year teachers.

During the time of Yendol-Hoppey et al.'s study (2009), the UMP assigned 12 mentors to work in 15 schools and each was assigned to work with 12–15 teachers. The schools included two high schools, three middle schools, and seven elementary schools; most of the novice teachers entered teaching through alternative, non-college of education teacher preparation programs (e.g., Teach for America). Most of the mentors worked in a single school and they spent three-and-a-half days a week in their schools where they coached novice teachers. The researchers explained that UMP was based on the research literature on how to organize and structure induction programs. This included mentors having a daily presence in novices' classrooms, mentors not being part of the novice teacher evaluation process, and mentors utilizing self-selected mentoring practices based on school context and individual teacher needs. UMP mentors engaged with novice teachers in co-planning, preparing materials for instruction, collaborative study of novices' teaching practices, providing and accessing resources, preparing novices for performance evaluations, and coaching for instruction.

Yendol-Hoppey et al. (2009) primarily drew on one-on-one interview and focus group interview data from 12 mentors in the UMP program that they collected over the course of 16 months. The researchers focused on three mentors in their analyses. Echoing Achinstein and Barrett (2004) and Achinstein and Athanases (2005), Yendol-Hoppey et al. maintained that mentors needed to find balance between supporting novice teachers in survival mode (i.e., classroom management, the micropolitics of teaching, and an urban school bureaucracy) and focusing on deeper levels of pedagogical practice. The authors also argued that both mentors and novices need to be committed to social justice in order to effectively focus on both survival and success (i.e., addressing all students' learning and novice teachers' own professional learning). Lastly, the researchers also explained that if retaining teachers and developing their skills is the goal of induction programs, then mentors need to work closely with school leadership. This last assertion relates to Dharan's (2015) contention that school-based mentors need to work closely with teacher preparation programs in order to help foster reframing conversations.

In summary, these studies of mentoring novice teachers for culturally responsive teaching, equity, and social justice converge toward a number of themes. First, in order for mentors to effectively support novice teachers, they themselves need to have a deep understanding of diverse learners and pedagogical skills for teaching them. Second, mentors face multiple obstacles that make the work of preparing novices for diversity particularly challenging. These challenges typically include finding the right balance between supporting novice teachers with technical skills such as classroom management and student control and helping them develop a deeper knowledge of diverse students and pedagogical practices—this relates to what Achinstein and Athanases (2005) refer to as bi-level knowledge (knowledge of teaching and knowledge of learning).

The mentors in these studies also faced the pervasive challenge of dealing with conflicting messages from other school actors (i.e., novices' principals and colleagues). These school actors promoted practices that maintained the status quo—which focused on student control and increasing student test scores. Emphasizing these areas may inhibit beginning teachers from learning about their students' cultures, backgrounds, characteristics, and beliefs along with pedagogical practices that can promote deeper levels of student learning and not just skills to pass a test. In order to mentor novice teachers to effectively work with diverse students, mentors need to work in conjunction with school leadership and teacher preparation programs, and they need to move beyond classroom management and technical skills in order to focus more on curriculum and pedagogical practices (Achinstein and Athanases, 2005; Achinstein and Barrett, 2004; Dharan, 2015; Yendol-Hoppey et al., 2009).

Going forward, a few lines of research should be pursued. First, all of the studies presented here were qualitative and most featured interviews and observations. There is a need for larger-scale, quantitative studies that examine links between mentoring for diversity and outcomes related to diverse students' learning. Second, the studies featured here focused mainly on elementary teachers. There is a need for more research on secondary STEM teachers and how to support them with regard to equity and diversity. Third, while there are mentoring structures identified in these studies (e.g., weekly observations, reflection conversations, monthly seminars), there is a need for more studies on these and other mentoring structures and how they may affect beginning teacher outcomes. The field is starting to develop a strong understanding of what mentors need to know and the challenges they may face, but addressing these gaps will provide researchers and practitioners with a more robust understanding of how induction programs can effectively support novice teachers for diversity and enact skills and dispositions of culturally responsive and socially-just teaching.

Mentoring and induction programs and beginning STEM teacher effectiveness

While the second and third sections of this article described the structure, focus, and curriculum of beginning teacher induction programs in several countries, in this section we turn to a different question: whether or how teacher induction contributes to student achievement. Two recent large-scale, experimental studies in the US have investigated the impact of induction programs on the effectiveness of beginning elementary and middle school mathematics teachers (Glazerman et al., 2010; Schmidt et al., 2017). In one study, Glazerman et al. (2010) conducted a randomized control trial in 17 urban, medium-to-high-poverty school

districts that featured 1009 novice teachers in 418 elementary schools. The teachers taught kindergarten through grade 6; of the 1009 study participants, more than 900 (91%) taught mathematics and were responsible for students' mathematics outcomes. Each of the 418 schools was assigned at random to one of two treatment groups or a control group; one treatment group participated in the Educational Testing Service (ETS) Pathwise induction model and the other took part in the University of California Santa Cruz New Teacher Center (NTC) induction model. Beginning teachers in treatment schools in 10 of the 17 districts received comprehensive induction for one year only while novice teachers in treatment schools in the other 7 districts received comprehensive induction for two years (Glazerman et al., 2010).

As part of the Pathwise and NTC induction models, beginning teachers were assigned to full-time, trained mentors who observed their instruction and used formative teacher assessments to provide them with feedback. In addition, the novices received release time that enabled them to meet with their mentors. Mentors were usually assigned to a caseload of about 12–15 beginning teachers. Novice teachers in the control group received mentoring from a full-time teacher; in addition, beginning teachers and mentors in the control group were much less likely than those in the treatment group to receive release time and mentors were much less likely to participate in mentor training, to conduct classroom observations of their mentees, or to use a formative teacher assessment in their work with them. The teachers participated in the study through the end of their third year of full-time teaching and data was collected on the effects of induction on students' mathematics achievement as well as other outcomes. The authors presented results for the study participants' third year of teaching and reported that impacts on gains in mathematics scores for teachers in the seven two-year districts were positive and statistically significant. At the same time, they found no effect of comprehensive induction on gains in students' mathematics scores for teachers in the 10 one-year districts (Glazerman et al., 2010).

In a second study, Schmidt et al. (2017) carried out a randomized control trial in Broward County Public Schools (BCPS), Florida and Chicago Public Schools (CPS) that included 629 first- and second-year elementary and secondary mathematics teachers and 227 schools. The researchers randomly assigned 342 teachers in 108 schools at the school level to the treatment across the two districts and 287 in 119 schools to the control group. In terms of the treatment, the NTC induction model (described above) was provided to each beginning teacher in the treatment group during their first two years of teaching.

Similar to the Glazerman et al. study, novice teachers in the control group in the Schmidt et al. (2017) study received mentoring from a full-time teacher; novices and mentors in the control group were much less likely than those in the treatment group to receive release time; and mentors were much less likely to participate in mentor training, to conduct classroom observations of their mentees, or to use a formative teacher assessment in their work with them. The researchers examined the impact of two years of NTC induction support on the effectiveness of teachers in grades 4 through 8 in mathematics. They reported that students in grades 4 through 8 of teachers who received two years of induction support performed 0.15 standard deviation ($p < 0.01$) higher on average than students of control teachers; this is similar to moving from the 46th to the 52nd percentile (Schmidt et al., 2017).

Conclusion

During the past 20 years, beginning teacher induction programs for STEM teachers have received significant attention throughout the world. They are typically enacted in order to support novice teachers' development of high-quality instructional practices, help orient them to their schools, provide emotional support to them, and promote their retention. In this article, we drew on scholarship by Britton et al. (2003) to address a series of questions related to induction programs. In particular, in the second and third sections of the article, we examined the structure, focus and curriculum, purpose, and participants in induction programs in several countries and regions. In the fourth section, we turned to the question of the influence of induction programs on student achievement in mathematics.

In the second section, we presented brief descriptions of national approaches to induction in (a) three Asian countries/regions that have consistently performed at high levels on international comparisons of student achievement (i.e., PISA and TIMSS) and (b) two Western countries that regularly perform in the middle range on such assessments. In Singapore, Japan, and Shanghai, induction programs are generally based on national standards for teaching; are often coordinated with pre-service teacher education programs and/or teacher evaluation; link theory and practice; and reflect collective efforts to improve and professionalize teaching. In contrast, Australia and France lack national induction programs; instead, schools and districts in these countries have responsibility for supporting novice teachers and local induction programs are not tied to centralized standards for teaching or collective efforts to professionalize teaching. While many factors contribute to national performance on PISA and TIMSS, it seems that the approaches to induction (and more generally to teacher development) in Singapore, Japan, and Shanghai could be contributing to their students' consistently-high performance on these tests.

In the third section, we reviewed several empirical studies of induction programs that address issues related teaching for equity and diversity and came to two main conclusions. First, mentors are more likely to effectively help beginning teachers learn to enact culturally relevant teaching if the mentors themselves have developed a strong understanding of diverse learners and how to effectively teach them. Such learners can include students who are diverse with respect to race/ethnicity, language, sexual orientation, ability, and other characteristics. Second, in many school contexts, mentors potentially face a number of challenges in their efforts to help novice teachers address issues related to equity and diversity in their instruction; these can include, for example, norms among other teachers, expectations from school leaders, and pressure to follow a prescribed curriculum in mathematics and other subjects.

In the fourth section, we described two randomized control trial studies of the impact on students' mathematics achievement of comprehensive induction programs for beginning elementary and middle school mathematics teachers. In the first study, Glazerman et al. (2010) reported no effect of one year of comprehensive induction on students' mathematics achievement after the first year of teaching or the second year of teaching. On the other hand, the authors found that participating in two years of comprehensive induction was related to gains in students' mathematics achievement after the third year of teaching. In a somewhat similar finding, Schmidt et al. (2017) reported that participating in two years of comprehensive induction was related to gains in students' mathematics achievement after the second year of teaching. It may be that most beginning teachers face so many challenges as they transition to full-time teaching that they are generally less likely to benefit from comprehensive induction (in terms of improved effectiveness as measured by student performance on standardized tests) until their second year of teaching (Youngs et al., 2019). In addition, it may be that as mentors (who are assigned to work with beginning teachers in multiple schools) become more familiar with their mentees' school organizational contexts over time, they are in a better position to support them.

Appendix

Table 1 Overview of induction structures, targets, focus/curriculum, and purpose in six countries/regions.

	<i>Australia</i>	<i>Shanghai (China)</i>	<i>France</i>	<i>Japan</i>	<i>Singapore</i>	<i>California (U.S.)</i>
Structure	Five-year national probationary period; teachers must meet pedagogical content knowledge (PCK) and achievement standards; school or district determined	Regionally centralized; 240 h of training over first five years; additional 540 h of training available later	Recently shifted from nationally centralized to fully decentralized; no structure for induction	A total of 300 h in first year; 120 h based at school, four days per week; 180 h at MEXT training center, one day per week	Three-day national orientation; two years of school-based training with nationally trained mentors	Two years of on-site weekly meetings and monthly seminars; focused on California standards for the teaching profession
Level of experience targeted	First five years	First five years; senior teachers	Shifted from first-year "inductee" teachers to general professional development	First year	First two years	First two years
Focus and curriculum	Locally determined to support national PCK and achievement standards	Regionally-determined standards; data-driven improvement cycle	Shifted from focus on combining theory and practice to preparing for daily responsibilities such as classroom management	Nationally coordinated to support transition into everyday classroom responsibilities; lesson study	Entering national "professional fraternity of teachers;" combining theory with practice	Goal setting; sharing resources; lesson planning; classroom observations; modeling lessons; analysis of student work
Purpose	Teacher retention; professionalization	Building culture of collaboration	Shifted from contextualizing theory to general professionalization	Contextualizing theory in practice as a collective	Establishing and maintaining classroom excellence as a collaborative team	Improve instruction and teacher retention

References

- Achinstein, B., Athanases, S.Z., 2005. Focusing new teachers on diversity and equity: toward a knowledge base for mentors. *Teach. Teach. Educ.* 21 (7), 843–862.
- Achinstein, B., Barrett, A., 2004. (Re)framing classroom contexts: how new teachers and mentors view diverse learners and challenges of practice. *Teach. Coll. Rec.* 106 (4), 716–746.
- Britton, E., Paine, L., Pimm, D., Raizen, S., 2003. *Comprehensive Teacher Induction: Systems for Early Career Learning*. Kluwer Academic.
- Cornu, B., 2015. Teacher education in France: universitization and professionalization—from IUFMs to ESPEs. *Educ. Inq.* 6 (3), 289–307.
- Darling-Hammond, L., 2017. Teacher education around the world: what can we learn from international practice? *Eur. J. Teach. Educ.* 40 (3), 291–309.
- Dharan, V.M., 2015. Beginning teachers and diversity—why the need for extended critical professional support. *Asia Pac. J. Teach. Educ.* 43 (1), 61–74.
- Fujimura, Y., Mistilina, S., 2020. How Japanese education boards frame teacher education as a training, learning, and policy problem. *J. Educ. Teach.* 46 (1), 20–35.
- Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., Jacobus, M., Ali, M., 2010. Impacts of Comprehensive Teacher Induction: Final Results from a Randomized Controlled Study. U.S. Department of Education, Institute of Education Sciences.
- Howe, E.R., 2006. Exemplary teacher induction: an international review. *Educ. Philos. Theor.* 38 (3), 287–297.
- Mullis, I.V.S., Martin, M.O., Foy, P., Hooper, M., 2016. *TIMSS 2015 International Results in Mathematics*. TIMSS & PIRLS International Study Center, Boston College.
- OECD, 2018. *OECD Initial Teacher Preparation Study*. Author.

- OECD, 2019a. A Flying Start: Improving Initial Teacher Preparation Systems. Author.
- OECD, 2019b. PISA 2018 Results (Volume I): What Students Know and Can Do. PISA, OECD.
- Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., Tang, J.H., 2016. Highlights from TIMSS and TIMSS Advanced 2015: Mathematics and Science Achievement of U.S. Students in Grades 4 and 8 and in Advanced Courses at the End of High School in an International Context (NCES 2017-002). U.S. Department of Education, National Center for Education Statistics.
- Schleicher, A., 2016. Teaching excellence through professional learning and policy reform: lessons from around the world. Int. Summit Teach. Prof. 1–94.
- Schleicher, A., 2019. PISA 2018 Insights and Interpretations. OECD Publishing.
- Schmidt, R., Young, V., Cassidy, L., Wang, H., Laguarda, K., 2017. Impact of the New Teacher Center's new teacher induction model on teachers and students. SRI Int. 1–41. Appendices pp. A1-H-14.
- Wang, A.H., Coleman, A.B., Coley, R.J., Phelps, R.P., 2003. Preparing Teachers Around the World: Policy Information Report. Rep. No. ED 479903. Education Testing Service, pp. 1–47.
- Yendol-Hoppey, D., Jacobs, J., Dana, N.F., 2009. Critical concepts of mentoring in an urban context. N. Educat. 5 (1), 25–44.
- Youngs, P., Bieda, K., Kim, J., 2019. Teacher induction programs that lead to retention in the STEM teaching workforce. Am. Assoc. Adv. Sci. 1–41.
- Zhang, M., Kong, L., 2012. An exploration of reasons for Shanghai's success in the OECD Program for International Student Assessment (PISA) 2009. Front. Educ. China 7 (1), 124–162.

Design-based pedagogy: a heuristic for developing math and science teachers' professional practice

Adam Devitt^a and Catherine Milne^b, ^a Department of Teacher Education California State University, Stanislaus, Turlock, CA, United States; and ^b Department of Teaching and Learning, New York University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	292
Epistemological shifts for design-based pedagogy	292
Design-based pedagogy	293
Lesson planning as design partners	294
Implementing science lessons for ELLs and students with disabilities	295
Integrating robotics in ways to enhance professional learning	296
Conclusion	297
References	297

Introduction

Design-based pedagogy has been developed for fostering the professional practices of middle school math and science teachers. Our work originated out from a professional development summer workshop where we worked in collaboration with teachers to develop math and science lessons that integrated LEGO robotics. Design-based pedagogy was conceptualized out of Ann Brown's (2002) work on design-based research. She highlighted that the role of the teacher is to "entice students" and modify the curriculum in order to change classrooms, as she describes, "from academic work factories to learning environments that encourage reflective practice among students, teachers, and researchers" (Brown, 1992, p. 174). Our research enhances the design-based research literature by adding nuanced epistemological positions. We reposition teachers to be "intellectuals" and "critical leaders" as described by Henry Giroux (1985) and Darling-Hammond (2006), respectively. Drawing on these perspectives, we argue that teachers need to be integral at each stage of their professional work: including curriculum development, instructional design, and professional learning. In doing so, we find that teachers are better positioned to make adjustments in teaching when implementing lessons, modifying future lessons, providing rationales based on their experiences, and ultimately advancing their professional practice as math and science teachers (Devitt et al., 2019a,b).

Epistemological shifts for design-based pedagogy

In Science, Technology, Engineering and Mathematics (STEM) learning, teachers are required to implement curriculum and provide instruction in ways that fosters student learning. However, who makes the curriculum, who selects the instructional practices, and for what purposes (i.e., how learning is assessed) matters in conceptualizing the role of the teacher at each of these pedagogical stages and significantly impacts the quality of STEM learning that occurs. Giroux (1985) takes a strong stand saying, "One of the major threats facing prospective and existing teachers within the public schools is the increasing development of instrumental ideologies that emphasize a technocratic approach to both teacher preparation and classroom pedagogy argues that education is becoming increasingly commodified" (p. 47). In STEM education, this results in textbook companies becoming tasked with determining what math and science knowledge is important and how to impart knowledge on to students. Math and science curriculum become "teacher-proof", recipe-like STEM kits. Norton et al. (2007) describe their concerns with teachers implementing others' curriculum. They found that teachers have difficulty extracting important concepts from activities and fail to utilize the technological materials in ways that promote real conceptual development (Norton et al., 2007). This management style of pedagogy works to consolidate STEM knowledge, reduce complexity for consumption, and enable predefined disciplinary ideas and practices to be efficiently assessed on standardized tests—for the goal of controlling teacher behavior and making outcomes predictable (Giroux, 1985). Universal pre-packaged curriculums do not provide for pedagogical opportunities where localized teachers can flexibly design classroom learning experiences, excite young learners, and prepare them for college, careers, and civic responsibility in STEM areas (NGSS Lead States, 2013).

Instructional practices often endorsed and mandated by school districts contribute to the de-professionalization of teacher practice. In our own research, while engaging in collaborative lesson planning during a professional development on how to integrate LEGO Robotics into math and science lessons, teachers were quick to initiate the lesson planning process using terms like: "Two-fold instruction", "Do now", "recap", "Stop and jot", "Turn and talk", and "cold calling" (Devitt et al., 2018). These practices were popularized and endorsed by Doug Lemov (2015) which he describes as "specific, concrete, actionable techniques" for teaching. These techniques have been institutionalized and standardized across New York City and other large metropolitan areas and often serve as mandated behaviors to be observed on teacher evaluations. These practices are not inherently bad, and could in fact

promote learning in certain contexts. However, as abstracted teaching techniques geared toward supporting memorization, they do not prepare teachers to implement pedagogies that support high-quality learning as described by the Organization for Economic Co-operation and Development. In other words, when placed into a lesson plan with the expectation to be enacted with fidelity, teachers miss the contextual relevance of the learning as it unfolds in the classroom and fail to incorporate or develop the pedagogical practices needed to support STEM learning.

To meet the call set out by [NGSS Lead States \(2013\)](#) in the Next Generation Science Standards (NGSS)—the most current national science standards in the United States, we also frame what students are learning from a global perspective. In 2008, Organization for Economic Co-operation and Development (OECD), responsible for leading Program for International Student Assessment (PISA), described four key points applicable for quality STEM learning: (1) conceptual understanding; (2) connected and coherent knowledge, (3) authentic knowledge within a specific context of use, and (4) collaborative learning. To meet this call in STEM teaching, [Darling-Hammond \(2006\)](#) has argued that the United States educational system needs teachers to be the critical pedagogues in charge. This stance requires teachers to be the critical agents who led the curricular planning and instructional decision-making because they can be the more responsive within the complex and dynamic system that exists in STEM classrooms ([Darling-Hammond, 2006](#)).

Design-based pedagogy

One of the major goals that contributed to forming our research endeavor was to support teachers with design-based pedagogies in the context of integrating LEGO robotics and engineering when planning math and science lessons, developing instruction strategies, and becoming a stake-holder in their own professional learning. One of our key epistemological standpoints for design-based pedagogy is that teachers are stakeholders at every step of the pedagogical decision making process. Now, that does not mean that teachers are learning solely on their own. Rather, we found success in fostering teacher professional learning by positioning teachers, mathematics and science education researchers, and engineers as “design-partners” ([Devitt et al., 2018](#)). Each partner brought her or his personal areas of expertise, whether from experience, scholarship, or both, and contributed to the shared activities and purpose of developing math and science lessons that integrated LEGO robotics for middle school classrooms.

Over the course of a three-week professional development workshop, as teacher educators, we provided instruction on various learning theories and collaborated with middle school teachers on their on-going lesson planning to develop a sense in how to use new pedagogy to inform the lesson planning and instructional practice. We drew from the work on problem-based learning ([Kolodner et al., 2003](#)) and Technological Pedagogical Content Knowledge (TPACK) ([Mishra and Koehler, 2006](#)) in order to conceptualize how to design purposeful and authentic ways for using technology to enhance their math and science lessons. We explicitly wanted to avoid using technology as a flashy gimmick that provides entertainment but overshadows any learning of mathematics or science conceptual development—as found in the research of [Norton et al. \(2007\)](#). We drew from the work of Jean Lave on “situated learning” (1988). We provided learner-centered perspectives for designing engagement activities in mathematics and science lessons. A new phrase we integrated was how to reason *with context* and not just *in context* ([Lave, 1988](#)). In our professional development LEGO robots provided the context from which we developed mathematics and science lessons.

In order to analyze the intellectual contributions of teachers to their lesson planning and the artistic maneuvering when implementing their science and mathematics lessons, we adopted, and adapted, Donald Schön's (1983) model where he conceptualized teaching as a “professional practice”. One of the central tenants we adopted from Schön (1983) is that teachers need to be part of the “problem setting” stage. We interpret this to mean that, rather than having curriculum developers decide what is important to teach and subsequently providing pre-designed lesson plans and activities, the teachers need to be positioned from the start to take national and state standards, decide how best to translate the purpose in their given setting and to their specific audiences, and develop appropriate activities and instructional strategies to support learning. Schön (1983) sets forth a cogent heuristic for analyzing the professional work of teachers to illuminate, understand their overarching purpose (problem space), make curricular and instructional decisions based on their professional knowledge of their students and relevant social, political, or environment issues within which to couch the lessons (reframing the problem), and continuously monitor and make necessary, micro-level pedagogical changes if and as needed (experimenting and evaluating).

Within design-based pedagogy, we have developed guideposts on Schön's seminal book, *The Reflective Practitioner* (1983) to articulate the work of teachers as professional practice. We have found that it is both formative to help teachers construct science lessons as well as a useful analytical tool to recognize teachers' professional practice in times of action and reflection. Schön's (1983) three central guideposts include:

- (1) Identifying the problem space: Schön's first step of professional practice is to have teachers frame the initial problem. The “problem” is the ability to conceptualize or outline the dimensions of a theoretical space in which to enact pedagogical decisions. For teachers, this often means conceptualizing the depth and objectives stemming from national and state standards within their discipline. Within our professional development, teachers framed the problem space, for example about how to integrate LEGO robots into their science lessons.
- (2) Reframing the problem: Schön's second step of professional practice requires teachers to take on the task of reconceptualizing the problem space as a real-world, authentic issue that is contextualize for the purpose and relevance of her students' lived-

experiences and strengths—based on recommendations from problem-based learning (Kolodner et al., 2003). Also, teachers need to consider what lessons came before and what is coming in the future to maintain coherence in the curriculum.

- (3) **Experimenting and Evaluating:** The third guide post refers to how teachers put their pedagogical decision-making into practice. Teachers are tasked to continuously monitor and evaluate the quality of learning as the lesson unfolds. Teachers can make micro-level adjustments to the curriculum or in their instructional practice. These might look like teachers providing explanations of the context, prompting or probing students, assisting with the technology, etc. Based on what is learned from teaching the lessons, teachers refine what practices could be used in future iterations of science lessons (Barab and Squire, 2004).

In order to make the professional-level decisions, teachers utilize practices such as “reflection-in-action” and “reflection-on-action” and listen to how the teaching situations “talk back” (Schön, 1983). It is important to note that Schön (1983) comes from a phenomenological perspective and takes the stance that professional knowledge exercised in both technical critique and embodied practice—but placing extra emphasis on embodied decision making when it comes to in-the-moment pedagogical changes (Schön, 1983). Given the complex, dynamic, and contextual nature of teachers’ professional work, Schön’s (1983) model for professional practice has been instrumental for helping analyze, understand, and make the importance of the “intellectual” work of teachers. From these guideposts, we highlight the elements of how teachers and their “design partners” engage in the design-based pedagogy (see Fig. 1).

Lesson planning as design partners

As described earlier, teachers are best positioned to engage in the design process of developing and implementing science lessons when they are key stakeholders and designers from the initial conceptualization and planning stages. Design-based science lessons require goal-based, real-world scenarios for students to work toward (Schank et al., 1994). In our professional development project, LEGO robots played a key role. We advise that technology be utilized in a way that enhances the lesson, and not just become a “bell and whistle” engagement strategy. Technology integration ought to help students nuance the way in which they interact with the phenomena. New disciplinary ideas, tools, and terminology should be weaved into the lessons so that students capture the use and purpose within the context to foster scientific engagement and thinking (Vygotsky, 1978). Lessons should include multiple opportunities for students to interact with phenomena, plan and conduct investigations, and produce a meaningful artifact that can be shared or even kept (Papert and Harel, 1991). The artifact can be used to help communicate what they learned and how they developed their ideas throughout the process.

Teacher educators and researchers can help foster teachers’ abilities to develop these kinds of lessons as “design partners”. Rather than provide examples or pre-prepared curriculum, “design-partners” collaborate side-by-side to devise lessons at each step of the way. Together, “design-partners” examine state and national standards, choose relevant, real-world problems to reframe performance expectations or learning objectives. When teachers were provided an opportunity to collaborate and plan their lessons. They started by asking reflective questions: “what do we use as our resources” and “what do students need to know”? They used the Next Generation Science Standards (NGSS) as a starting point to construct their problem space. As part of the “reframing” process, one group of teachers devised an original scenario:

You are living in the year of 2150. Families are no longer required to make weekly visits to the grocery stores. The EV3 ShopNOW robot service provides families with home delivery. Three families have sent lists for groceries to be delivered. However, some robots are malfunctioning. The EV3 robots has encountered a defect that results in robots crashing into the cashier register. A force of 300 N or more will cause injury to the cashier. Your task is to investigate which cart mass will produce a force that the cashier can stop.

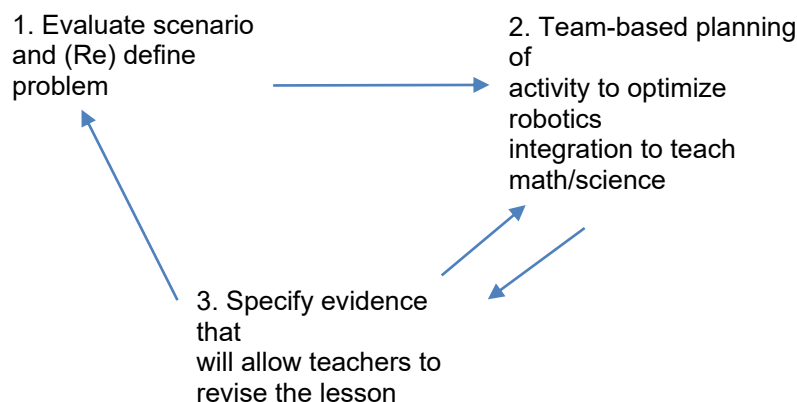


Fig. 1 Initial elements of design-based pedagogy.

Researchers on the project collaborated as “design-partners” to help revise the scenario. Together they co-constructed:

Right now companies are manufacturing and currently testing in local markets around the world—robots that can deliver up to three shopping bags' worth of goods within 30 min to homes within a 3-mile radius. The company, according to CEO Heinla, is “reinventing” the so-called last mile of transportation, and it's currently running tests with commercial partners such as London's takeout food delivery service Just Eats that involves the robots ferrying meals to customers. However, one of the challenges is making sure that the robot which collects the grocery items does not crash into the register. Your challenge is to explore the level of force involved and decide what is a safe level for the cashier.

The goal of revising the scenario was to make it as authentic to real-life as possible. First, we want science lessons to be realistic in the sense that the scenario actually occurs, or could happen, in the real life of students. In this scenario, we helped teachers conduct a Google search and select an example that exists in the real-world. Second, we wanted to maintain accountability to the agency of the phenomena. In this example, we helped teachers more accurately re-conceptualize the problem of how robots cause injury to cashiers, in order to make it more realistic. According to [discover.com](#), an article highlights that it takes about 4000 N to break a human leg. Therefore, 300 N as the original scenario was written, would not have accurately represented the phenomena of force and human injury. As design-partners, we helped teachers better understand that mathematical and science learning is more than just solving equations and problems, but also to help students learn about how to use equations and problem solving to better understand real life phenomena (Devitt et al., 2018).

Implementing science lessons for ELLs and students with disabilities

We learned that teacher professional practice is a complex and dynamic social phenomenon that is difficult to understand from the outside. As part of the professional development, after teachers taught science lessons that were designed during the summer workshops, we held follow-up reflective sessions. During these sessions, teachers shared stories and videos of how they implemented their lessons, highlighted strengths and challenges, and explained what they learned from engaging in the process. We found that using our framework of design-based pedagogy and the guideposts set out by Schön (1983), we were able to analyze teacher stories in ways that illuminated the multi-faceted layers of teachers' professional decision-making and evaluation. This was particularly fruitful to help us better understand how teachers implemented and refined their lessons. In particular, while hearing teacher stories, we were most interested in the parts where teacher moves and outcomes were discussed.

In Table 1, there is a list of quotes pulled out from a story told by Ms. Chang, a middle school science teacher who teaches in an area of New York City with high populations of Chinese-speaking students. She explained that she taught a lesson using the LEGO robots focusing on planets in the solar system. She showed a picture of a PowerPoint slide that had directions in both English and Chinese. Ms. Chang said, “*I translate the worksheet from English to Chinese so that my students can read. The only barrier for them to learn (science) is the language, otherwise they are super smart kids*”. Ms. Chang's story revealed that she modified her lesson to include Chinese characters to help students who struggle with the English language access the directions and lesson content. She implied that this was the only modification needed for her lesson as students were able to demonstrate learning once language barriers were removed.

In Table 2, there is a list of quotes pulled out from a story told by Mr. Wilson, a middle school science and special education teacher whose class consisted of 12 students with Individual Education Programs (IEPs). Within his story, he shares how he problematizes teaching about the solar system, while also considering the individual needs of students with disabilities who were in multiple grade-levels. He reframes the problem for discussing a specific lesson where he considered how to best use the LEGO robots. He planned for merging students' previous experiences with positive and negative integers and how to build on using computer programming (Scratch) to teach the concept of using “x and y” in a coordinate system. Mr. Wilson shared a multitude of strategies he tried in order to support student learning. For example, he practiced the coordinate system using 3D printing, he invited previous students back to his classroom to act as leaders and mentors to current students, and he allowed students to use the hallway to increase the space in which students can model the solar system. He evaluated the outcomes of his adjustments by saying that he was further along in the curriculum than previous years, he noticed that students have a “light go off”—indicating that students recognize their own “ah-ha” moments, and that students are engaged in quality scientific conversations.

Table 1 Example analysis of a middle school science teacher of English Language Learners.

The problem space	● How I get the lesson (cell cycle) is from the Google Drive ● I aligned the lesson to the scope and sequence. The cell cycle I learned from the last professional development [meeting].
Reframing the situation	● [Presented on PowerPoint slide] lab: Understanding the length of each phase of the cell cycle actual model. ● I translate the worksheet so my students can read.
Experimenting and evaluating	● The only barrier for them to learn is the language, otherwise they are super smart kids.

Table 2 Example analysis of a middle school science teacher of students with IEPs.

The problem space	● This year I was fortunate to have five, 12-to-1-to-1, configuration (i.e., 12 students with IEPs, one special education teacher and one paraprofessional). ● All the kids have IEPs with various disabilities. All my classes are mixed grades: sixth and seventh or seventh and eighth but also on second or third grade reading level. ● Focus on lessons on the solar system [PowerPoint slide title] ● There is such an interaction with whatever I use in the classroom, that there are hooks in and out of the program. I Look to hook them however I can so that they want to come to class so that they want to do whatever we are doing on that day.
Reframing the situation	● Although I'm teaching science, I have done work on positive and negative integers to get them comfortable. ● Because I'm using scratch [programming]—so the kids realize that there's an x and y thing happening, and if I can demonstrate that happening on the floor—that's a visualization of a concept that is somewhat difficult for them.
Experimenting and evaluating	● Students are doing two things: we started printing thing out on the 3D printer. The kids have to use TinkerCad. So they use 3D cadcam printing software, which is based on X,Y coordinates and they are printing off planets and putting them onto the robots. ● So we're using kids to come in a mentor other kids and we all know that kids learn best from each other. ● One of the other things I have done here, this is my second year here using the robots. I've developed leaders, I bring students from different classes in if I need a particular thing accomplished for that day. I've introduced programming for instance. So I have a couple of seventh and eighth graders—girls by the way—that come in and teach my sixth and seventh graders how to program. We will do, some time after the break: graphs. ● I bring the student programmer in and she helps them to understand how we can do this, how we can make it [the robot revolve] faster, how we can change the power, how we can change the wheel size—now we're talking about circumference, and radius, and diameter, there are formulas involved and I have students mentoring other students. ● We're further along this year than we were last year and I think that at least has something to do with my familiarity with things (i.e., robotics, programming—this was this teacher's second year in the program) and my willingness to trust the students—and I have some students who are working with me for a second year in a row. ● Now I have a team in the hallway and last year I could not let three kids in the hallway. ● At some point the light goes on after really accomplishing something. ● Now what they want to do is put a broomstick in the center and run all nine planets. Is it a true analysis of spherical rotations? Not at all, but if I get kids to have that level of conversations, it's a homerun.

The goal of highlighting two examples of how teachers implement science lessons is not to compare the quality of teaching. The purpose is to highlight that when teachers are positioned to develop their own lessons, they are capable and effective at making necessary adjustments to their curriculum and instructional practices as needed to the point where students more fully participate in the flow of the lesson and are able to demonstrate learning. Since the teachers selected the learning goals and subsequently developed the lessons, they were able to make the intellectual verifications of what quality science learning looked like, and subsequently connect their own successes to the adjustments they made to their curriculum and instruction. When teachers implement pre-packaged lessons with fidelity, the onus of learning often shifts to the full responsibility of the student. Failure is often attributed to deficiencies in student intellect, motivation, behavior, or other personal dispositions rather than assuming that the curriculum or instructional practices need adjustments. Design-based pedagogy enables teachers to better support and ensure the learning of our most marginalized students (Devitt et al., 2019a,b).

Integrating robotics in ways to enhance professional learning

Stories that teachers shared also revealed how teachers began to improve the quality of technological integration in ways that advanced students' development of conceptual understanding (Devitt, 2020). Alisha, a middle school math teacher described how she typically used the city-wide adopted mathematics curriculum for teaching her course. When teaching the concept: adding and subtracting positive and negative integers, she historically used number lines to demonstrate mathematical problems. Then, in light of learning how to use the LEGO robots and programming during the summer workshop, she wanted to plan a lesson that engaged students to think algebraically by using "in-and-out machines" or "in-and-out tables" in order to demonstrate the relationship between numbers and functions. She shared that by using the LEGO robot she could teach her sixth graders how to use higher grade-level mathematical tools (i.e., functions). Rather than having students use number lines, the robots could be programmed with functions and students could input numbers into the robot which would act as the "in-and-out machine"—and move according to the programmed function. Alisha expressed that students could reflexively engage in thinking about: *"What is it that the machine's doing to the number? What is happening mathematically?"* Alisha explained why the robots were so valuable to her nuanced way of teaching numbers and functions. She said, *"I think maybe just seeing visually how they [the robots] move along the number line I think it really clarified the error of, you know, you can count the numbers along the number line, as opposed to the jumps and the physical movement of the robot. And how that clarified it for them"*. In other words, students might make errors when using a numberline to practice the functions and not notice if they make mistakes. However, when the robot is programmed and students input numbers, students' expectations of where the robot should end up and where the robot ultimately ends up causes students to clarify their errors of thinking (Devitt, 2020).

Conclusion

In addition to conceptualizing teachers as “intellectuals” and “design partners” in the service of producing quality science and mathematics lessons while integrating robotics, with the design-based pedagogy model, we position teachers as agents in forwarding their professional practice. Two major goals are achieved when STEM teachers engage in professional practice that is reflective of the design-based pedagogy set forth in this book chapter. First, teachers produce high-quality STEM lessons that capitalize on integrating robotics and engineering. Because teachers are involved and invested in the construction of the entire lesson; including the planning, institutional design and implementation, and reflection on the overall process, they have the know-how to make micro-level adjustments as needed to support student learning and have full knowledge of the landscape from which to further their future lessons—key for promoting and developing professional practice. Second, as teachers have sound contextual understanding of the schools, community, and students’ needs and strengths, they are the most knowledgeable and positioned persons for designing the curricular and instructional practice for whole classes, and make the micro-level adjustments (i.e., prompts, probes, assists, etc.) for individual students. This pedagogy is especially promising for teachers to be most responsive to the quality and individualized education for English language learners and students with disabilities. It is my firm belief that we need to prepare teachers to develop pedagogical practice that best supports students on the margins from the ground up rather than start with a pre-packaged curriculum that needs to be “differentiated” for students who do not fit the intended mold.

References

- Barab, S., Squire, K., 2004. Design-based research: putting a stake in the ground. *J. Learn. Sci.* 13 (1), 1–14.
- Brown, A.L., 1992. Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.* 2 (2), 141–178.
- Darling-Hammond, L., 2006. Constructing 21st-century teacher education. *J. Teach. Educ.* 57 (3), 300–314.
- Devitt, A., Milne, C., Kapila, V., Ma, J., Iskander, M., 2018. A Design-Based Pedagogy: A Model for Collaborative Lesson Planning Integrating Robotics. National Association of Research for Science Teacher Education (NARST), Atlanta, Georgia. Conference.
- Devitt, A., Milne, C., Kapila, V., Ma, J., Iskander, M., 2019a. Teachers’ Use of Design-Based Pedagogy for Integrating Robotics as a Professional Practice in Middle School Math and Science Classrooms. American Educational Research Association (AERA), Toronto, Ontario. Conference.
- Devitt, A., Milne, C., Kapila, V., Ma, J., Iskander, M., 2019b. Design-Based Pedagogy: Integrating Robotics for Supporting Students with Disabilities in Middle School Science and Math. National Association of Research for Science Teacher Education (NARST), Atlanta, Georgia. Conference.
- Devitt, A., 2020. Using Teacher Narratives of Integrating LEGO Robotics as Assessment Tools and Evidence of Professional Learning. National Association of Research for Science Teacher Education (NARST), Portland, Oregon. Conference.
- Giroux, H.A., 1985. Teachers as transformative intellectuals. *Soc. Educ.* 49 (5), 376–379.
- Kolodner, J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Ryan, M., 2003. Problem-based learning meets case-based reasoning in the middle-school science classroom: putting learning by design (tm) into practice. *J. Learn. Sci.* 12 (4), 495–547.
- Lave, J., 1988. *Cognition in Practice*. Cambridge University Press, Cambridge.
- Lemov, D., 2015. *Teach Like a Champion 2.0: 62 Techniques that Put Students on the Path to College*. Jossey-Bass, San Francisco, CA.
- Mishra, P., Koehler, M.J., 2006. Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach. Coll. Rec.* 108 (6), 1017–1054.
- NGSS Lead States, 2013. Next Generation Science Standards: For States, by States. Retrieved from. www.nextgenscience.org.
- Norton, S.J., McRobbie, C.J., Ginns, I.S., 2007. Problem solving in a middle school robotics design classroom. *Res. Sci. Educ.* 37 (3), 261–277.
- OECD, 2008. Resolution of the Council on the Financing of Part 1 of the Budget of the Organization. OECD, Paris.
- Papert, S., Harel, I., 1991. Situating constructionism. *Constructionism* 36 (2), 1–11.
- Schank, R.C., Fano, A., Bell, B., Jona, M., 1994. The design of goal-based scenarios. *J. Learn. Sci.* 3, 305–346.
- Schön, D.A., 1983. *The Reflective Practitioner*. Temple Smith, London.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Mental Processes*. Harvard University Press, Cambridge, MA.

Promoting engineering research with entrepreneurship and industry experiences: a teacher professional development program

Purvee Chauhan and Vikram Kapila, Department of Mechanical and Aerospace Engineering, NYU Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: the STEM learning gap	298
Rationale for incorporating mechatronics and robotics with entrepreneurship experiences in classrooms	299
Overview, theoretical background, and design of the project	300
Theoretical background	300
Project-based learning	300
Robotics-based teacher professional development	301
5E instructional model	301
Next Generation Science Standards (NGSS)—alignment with teacher professional development program and 5E instructional model	301
Project design and activities	302
Guided training	302
Collaborative research	302
Lesson plan development workshop	303
Industry-based experiences	303
Classroom implementation support and academic year activities	303
Technology business idea competition	304
Description of the competition	304
Project results	305
Evaluation procedure	305
Evaluation outcomes and observations	305
Illustrative examples of student projects	307
Discussion	307
Conclusion	308
Appendix: A	309
Acknowledgments	310
References	310
List of relevant websites	311

Introduction: the STEM learning gap

The prevalence of and accelerating innovations in technology are transforming everyday human experiences. There has been a growing reliance on technology to facilitate and ease our daily activities, including healthcare, education, leisure, etc. Furthermore, COVID-19 pandemic has catalyzed the digital transformation and altered our perceptions about online channels for almost all the industries. The ever-changing nature of work, predictions about the future of work, and rapid proliferation of technology call for an increase in skilled Science, Technology, Engineering, and Mathematics (STEM) workforce. Unfortunately, the US education system is facing a substantial challenge in generating adequate numbers of qualified STEM professionals who can fulfill the demands of the future workforce. A closer look into the STEM pipeline reveals that only 20% of country's high school graduates are equipped to handle STEM education and a mere 10% of global science and engineering graduates are from the US (Herman, 2019). These dismal statistics indicate a major need to intervene and support schools, teachers, and students to acquire professional skills that contribute to STEM learning and a better-prepared workforce.

One barrier to a well-prepared STEM workforce stems from the lack of interest and perceived value of STEM learning among students. Even though students are exposed to a variety of technologies in their day-to-day activities, their STEM educational experiences fail in intentional use of modern technologies (Iskander et al., 2010). This lack of exposure to contemporary technologies in a learning context disengages learners from STEM disciplines. Moreover, it denies them meaningful experiences to hone professional skills relevant to a modern and technological workplace (Sobhan et al., 2007). Even as the interdisciplinary nature of STEM fields is widely recognized, students continue to receive siloed exposure to disciplinary content knowledge in STEM. Furthermore, a majority of students find it challenging to understand STEM concepts in an abstract manner and fail to see their real-world implications. This prevents them from deeply understanding the STEM topics and, thus, finding value in it. Unfortunately, 30% of students lose interest in science disciplines by grade 4 and nearly 50% of students lose interest in pursuing STEM careers by grade 8 (Hurst et al., 2019). Thus, most of the students disengage from STEM fields by their high school years.

Teacher knowledge and abilities are known to directly influence the achievement of students (Sanders and Rivers, 1996). Thus, it is paramount to have teachers who feel comfortable with and have the capacity for technology-enabled instruction so that students can be equipped to attain success in STEM disciplines. Yet, most teachers lack sufficient preparation in integrating STEM content in their lessons and find it challenging to adapt, contextualize, and demonstrate the use of contemporary technologies in their classrooms (Chen et al., 2009; Ottenbreit-Leftwich et al., 2018). Furthermore, it can be challenging for teachers to stay updated with the advancing hardware and software of these technologies (You and Kapila, 2017). This results in classrooms that are less engaging, accessible, and personally relevant to students (AeA, 2005). Teachers' lack of preparation can also result in absence of familiarity and knowledge of advanced technologies among them, which may, in turn, affect their self-efficacy.

It is pivotal to address the aforementioned barriers to student engagement and teacher preparation in STEM to meet academic demands of today's learners and for their career success. To generate enthusiasm and engagement among students in pursuing STEM fields, the curricula must incorporate modern and familiar technologies (Benitti, 2012; Riojas et al., 2011), be personally and culturally appropriate (Young et al., 2019), and embed real-life applications in students' own communities (Morales-Doyle, 2018). The STEM programs must invite students from early ages through developmentally appropriate and hands-on activities that build familiarity and promote knowledge development among learners. Moreover, professional development programs must empower teachers in developing comfort with modern technologies and promote technology-infused learning for their hands-on practice.

Rationale for incorporating mechatronics and robotics with entrepreneurship experiences in classrooms

The last few years have witnessed a noteworthy increase in the use of robotics for designing transformative interdisciplinary learning experiences (Alimisis, 2013). The emerging field where robots contribute toward development of learners' skills and competencies is called Educational Robotics (Eguchi, 2012). Robotics, for the purposes of teaching and learning, has been primarily introduced to develop learners' interest and involve them in topics that are challenging to teach via traditional methods (Whitman and Wither-spoon, 2003). Educational robotics fosters deeper understanding of concepts through hands-on, active engagement that gives learners the opportunities to collaborate and solve complex problems (Mosley and Kline, 2006). On the other hand, mechatronics, which is the synergistic integration of mechanical engineering, control theory, computer science, and electronics to manage complexity, uncertainty, and communication in engineered systems (Kapila and Lee, 2004), can also ignite the intellectual curiosity of participants and empower them to perform hands-on STEM learning, conduct engineering research, and foray into entrepreneurial explorations. Multiple studies have revealed that students who engaged in such a learning showed positive learning outcomes (Hartmann et al., 2007; Milto et al., 2002; Rusk et al., 2008). Among the myriad advantages of deploying mechatronics and robotics in classrooms, some are highlighted below.

- These disciplines are well suited for project-based explorations wherein participants learn about the course material and engage in critical thinking while solving problems, contextualized in a facilitated design project, that may entail multiple solutions (Barron et al., 1998).
- Mechatronics and robotics generate an authentic learning environment through real-world problem solving where students make connections between various STEM concepts, build their own knowledge, and showcase their learning (Eguchi, 2012). Such a learning is limited not just to STEM-based topics but is applicable to other disciplines including social science, literacy, etc. (Rahman et al., 2017).
- Robotics-enhanced experiences engender student-led learning that is relevant to them. Such a hands-on learning of abstract topics engages students, infuses enthusiasm in them, and motivates them to present their work (Brill et al., 2016).
- The interdisciplinary and active nature of robotics-based activities cultivates 21st century professional skills and computational thinking skills among learners (Rahman et al., 2018). Such a learning gives students multiple opportunities to collaborate and brainstorm on challenging problems and affords opportunities to iterate on their solutions (Eguchi, 2021).
- Educational robotics makes learning equitable and accessible to all learners. The hands-on exposure to STEM concepts is crucial in developing cognitive and social skills among learners (Alimisis, 2013) and, thus, is a critical contributor in bridging the STEM learning gap.

Even though robotics-based learning offers a compelling approach to engender deep understanding of concepts, its classroom adoption is often limited to robotic design and programming (Williams et al., 2011). Most programs fail to utilize robotics to its fullest potential by using it as a passive learning tool (Ghosh et al., 2020) or employing it for extra-curricular activities (Williams et al., 2012). This limited exposure to the foundational concepts of using technological artifacts can impact students' understanding and deter them from entering the STEM fields. Moreover, a well-informed and prepared STEM workforce requires students to not only have STEM content knowledge but to also be adequately equipped in the entrepreneurial landscape. Today's technology-driven entrepreneurial landscape calls for a large number of entrepreneur-engineers who are equipped to generate novel products for the innovation economy, especially due to a complete transformation of the technology landscape since the pandemic. The project team believes in the prospect of emergence of novel entrepreneurial ventures through the technological affordances of mechatronics and robotics.

The infusion of educational robotics and technology-based entrepreneurial innovations can be utilized to impact students' abilities, interest, and awareness in STEM disciplines and careers. When students are exposed to STEM fundamentals through mechatronics and robotics integrated with business and communication skills, they can envision concrete applications of their STEM

knowledge and skills. Research has shown positive impact of entrepreneurship-based education programs to increase learners' self-efficacy, persistence, creativity, and desire toward entrepreneurial careers (Beary, 2013; Cioffi et al., 2014). Moreover, mechatronics and robotics are still nascent disciplines that have the promise to solve varied problems and advance into successful entrepreneurial solutions. Mechatronics- and robotics-based designs can render experiences in systems integration and product development which can support learners in honing and implementing their creativity, innovation skills, and entrepreneurial mindset. Leveraging students' fascination toward robots, devising real-world simulations, and designing an entrepreneurial experience can be an effective approach for learners to envision themselves as contributors to the future STEM workforce.

Overview, theoretical background, and design of the project

The necessity to develop pedagogically-sound learning experiences in mechatronics and robotics, and integrate them with entrepreneurship, led to the design and implementation of research-informed teacher professional development workshops. The workshops sought to create for teachers a motivational and engaging program that includes mechatronics and robotics training; immersion in hands-on, collaborative research; opportunities to cultivate entrepreneurial skills; and interactions with industry engineers. To address the STEM learning gap, the following key activities were included in the summer workshops of six-week duration.

- Introduce the interdisciplinary fields of mechatronics and robotics to teachers for renewing their knowledge and skills in science, math, and research. To do so, the workshop created a structured and integrated learning environment consisting of training, mentoring, and real-world collaborative engineering research.
- Provide teachers with hands-on experiences, skills, and resources in engineering research, prototype product development, entrepreneurship, and professional engineering practices. Such an approach can prepare teachers to: integrate project-based learning in their science and math curriculum; enhance school labs with technology used by scientists and engineers; improve their communication, presentation, and business skills; and inspire their students, through examples of real-world applications, to pursue careers in STEM and create technology ventures.
- Conduct a student idea competition and create a high school summer internship program.
- Enable undergraduate and graduate engineering researchers to develop communication, leadership, and mentorship skills; a deeper appreciation of STEM teaching and learning; and contribute to the research on K-12 STEM teaching and learning.
- Enable project team to integrate their research, teaching, and curriculum and lab development efforts with K-12 education activities.

The six-weeklong summer professional development workshops and ongoing academic year support were founded on credible theoretical frameworks of project-based learning (PBL), robotics-based teacher professional development, 5E instructional model, and learning standards alignment. This novel intersection builds on the affordances of real-world applications of mechatronics, robotics, and entrepreneurship, and thus is valuable in positively shaping teacher preparation. The project incorporates a variety of thoughtfully conceptualized hands-on activities (see **Project design and activities** subsection) that culminate into a showcase of student projects through an entrepreneurial idea pitch competition. This work-based environment (Lave and Wenger, 1991) facilitates the accumulation of diverse experiences, familiarization with contemporary industry practices, and insight into disciplinary knowledge, all of which eventually translate into effective student learning. Thus, during the workshop, teachers were given opportunities to interact with professional engineers through industry site-visits and mentorship to enhance their practical knowledge of STEM disciplines. Under cognitive apprenticeship (Collins, 1991), such an experience allowed teachers to model the gained skills for students, improving their understanding of STEM topics as well as awareness of STEM pathways. Over the duration of 3 years, the workshops and academic year activities were attended by 30 teachers who were hosted by the faculty, researchers, and engineering graduate students at an engineering institution in North East US. A description of the theoretical background and project design are provided in the following subsections.

Theoretical background

To generate purposeful and engaging hands-on experiences, we utilized the interactions of mechatronics, entrepreneurship, and industry experiences grounded in the theoretical frameworks of PBL, robotics-based teacher professional development, 5E instructional model, and learning standards alignment. A brief review of the theoretical frameworks is provided below.

Project-based learning

PBL is an instructional approach where learners investigate or solve a challenge over a sustained duration of time and, in the process, acquire intended knowledge and skills. PBL promotes critical thinking, collaboration, enhanced understanding of concepts, and long-term retention of content (Melchior et al., 2004) and, thus, is well suited for integration with mechatronics and robotics. Such a learning renders opportunities for authentic problem solving (Eguchi, 2012) and inspires self-reflection. Instantaneous feedback in the form of mechatronics device performance or robot behavior can give students opportunities to iterate and revise their solution and, thus, engender an atmosphere for self-directed learning (Eguchi, 2012). Furthermore, the learning artifact designed via collaborative team efforts is shareable among a larger community (Carbonaro et al., 2004) and is an effective means to instill a sense of purpose in learners. While the improved STEM knowledge can aid students in developing their own products, their lack of

entrepreneurial knowledge may limit it from reaching the intended market. This necessitates the integration of entrepreneurship in the curriculum. Relevant project-based mechatronics and robotics experiences combined with business skills can empower students to envision real-world applications of the generated STEM knowledge, which, in turn, can help foster confidence and creativity among learners. In this effort, PBL was an integral part in the design and development of learning activities that sought to imbue intellectual curiosity through socially relevant problem-solving.

Robotics-based teacher professional development

Professional development is crucial in addressing the evolving needs of teachers' content knowledge, skill improvement, and building their confidence (Desimone, 2011). The ongoing advancement of technology and its integration in the process of teaching and learning has further necessitated concentrated efforts toward teacher upskilling (You and Kapila, 2017) as the complexities of such tools can make teachers reluctant in employing them in their classrooms (Moorhead et al., 2016). Professional development programs need to be designed to focus on pedagogical aspects of technology integration as opposed to using technology tools for administrative tasks (You and Kapila, 2017). Robotics-based professional development programs, specifically, must address teacher needs; generate familiarity and comfort in using robotics tools; support them in connecting robotic activities to curriculum (Norton et al., 2007); and build their skills, knowledge, and beliefs. Such programs must allow teachers to engage in hands-on explorations using the robotics tools to solve real-world problems that may be personally relevant to their students. Moreover, high-quality and effective professional development programs must last for sufficient duration of time, involve collective participation, endow active learning experiences, and be coherent in their design (Garet et al., 2001). For instance, effective professional development programs must move away from short, one-shot models that assume a passive role of teachers to more hands-on and constructivist learning models (Sparks, 2002; Harada, 2016). The plan and design of the workshops must adhere to curricular and learning standards (Allen and Penuel, 2015), be cognizant of students' cultures (Gay and Howard, 2000), and extend for a sufficient duration of time (Sparks, 2002). This effort was founded on a strong and structured professional development program that lasted for six weeks. Under the PBL framework, the participants were provided hands-on learning experiences, guided training, collaborative research work, and industry experiences. Moreover, teachers were periodically visited in their schools during the academic year to support their in-class and after-school activities that built on the summer workshop.

5E instructional model

The 5E instructional model, consisting of the Engage, Explore, Explain, Elaborate, and Evaluate components, is a pedagogical approach that aims to develop higher order thinking skills and render a conceptual change in the learners (Tanner, 2010). Through a structured and well-organized sequencing of its components, the 5E model allows the learners to surface their pre-conceptions, explore novel ways of approaching a problem, learn new concepts deeply, apply their learning to various situations, and evaluate their understanding (Tanner, 2010). The model is widely accepted for developing curriculum, planning lessons, designing teacher professional development programs, or designing informal learning experiences (Bybee et al., 2016). The adoption of 5E framework with fidelity can lead to improvement in student learning, motivation, and retention of concepts (Bybee et al., 2016). The 5E framework can work effectively in tandem with the Next Generation Science Standards by rendering lessons that incorporate engaging scientific practices, exploring research questions, conducting engineering design research, evaluating the results of new phenomena, etc. The 5E model addressing the learning dimensions of NGSS was a key component of this effort as it offered teachers a pragmatic methodology through a manageable and scientific sequence. Embedding such learning experiences in the classrooms can enable students to develop an appreciation for the new concepts through observation, inquiry, and analysis.

Next Generation Science Standards (NGSS)—alignment with teacher professional development program and 5E instructional model

The NGSS were devised to develop new ways of thinking and learning in STEM disciplines by alleviating the prevalent challenges of siloed focus on developing science content expertise and inadequate illustration of interconnections between the concepts. Organized in a unique three-dimensional format, the NGSS reimagined K-12 science in a model consisting of Disciplinary Core Ideas (DCI), Cross Cutting Concepts (CCC), and Science and Engineering Practices (SEP). DCI consist of science concepts and ideas students must learn, CCC are pervasive across subfields within science and allow students to apply their learning, and SEP are activities that help students understand the development of scientific knowledge and forms linkages between science and engineering (Bybee, 2014; Duncan and Cavera, 2015). The NGSS afford teachers with opportunities to flexibly envision, design, and enact novel lessons that prepare students for post-secondary education (Moore et al., 2015) and allow students to explore the real-world through the lens of scientists and engineers (Moore et al., 2015). Effective adoption of NGSS requires large-scale professional development programs to enable teachers to transition away from the old science standards (Wilson, 2013). Thus, programs that prepare teachers to understand and employ the NGSS are critical. This project's professional development program exposed participant teachers to effectively learn and apply the NGSS framework for addressing the challenge of designing standards-aligned curricula, instruction, and experiential learning. Moreover, the participating teachers received feedback and support from their peers and project team in improving their lesson and its enactment. The NGSS were integrated with the 5E framework for optimal exploration of scientific concepts that would expose learners to interconnected knowledge structures. Furthermore, this NGSS-plus-5E professional development workshop gave educators autonomy and flexibility to ideate, design, and implement novel lessons in classrooms.

Project design and activities

To maximize the learning potential using mechatronics, robotics, and entrepreneurship experiences, the summer workshop employed a scaffolded approach that included a two-weeklong “guided training” and a four-week “collaborative research” experience. Such an approach was deliberately followed to instill a level of confidence among teachers to engage with mechatronics and robotics technologies and research. Teachers also participated in a “lesson development workshop” on three days during the six-week professional development (one day every two weeks), that focused on aligning lessons to NGSS and 5E instructional model. Moreover, they gained familiarity with the “real-life experiences of industry professionals” through workshops or presentations by them as well as industry site visits. See Fig. 1 that illustrates teachers engaged in the diverse activities of the project. During the academic year, project team members periodically supported teachers as they implemented mechatronics and robotics-based lessons and activities. Students from teachers’ classrooms participated in a business idea competition that encouraged them to pitch economically viable technology solutions for real-world challenges.

The program started with an orientation and interaction between participating teachers and project team members. Teachers were also provided with the program agenda, demos of sample projects from earlier offerings, lab safety instructions, and a mechatronics and robotics learning kit composed of components, supplies, and study material. Following are details of the program elements.

Guided training

The first two-weeks of summer workshop involved guided training sessions that introduced teachers to the fundamentals of mechatronics and robotics (sensors, actuators, electronics, electro-mechanical components, microcontrollers, etc.) and entrepreneurship. Each day was divided into morning and afternoon sessions where teachers participated in interactive lectures, corresponding hands-on experiments, and discussions on classroom implementation of their learnings. Sessions on entrepreneurship included key concepts on assessment of business opportunities; strategies for managing financial, human, and technology resources; networking; business promotion; and sustainable growth. Teachers were divided in small groups during these sessions and provided opportunities to design their innovations, create their venture ideas, examine successful business models, and engage in networking. Experience of every small success, by themselves or by their peers, in myriad hands-on learning activities of guided training prepared teachers for collaborative research in mechatronics and robotics.

Collaborative research

Following the guided training, for the next four weeks, teachers experienced the process and challenge of conducting authentic engineering research by working individually or in teams, through project-based learning. They contributed to ongoing research projects

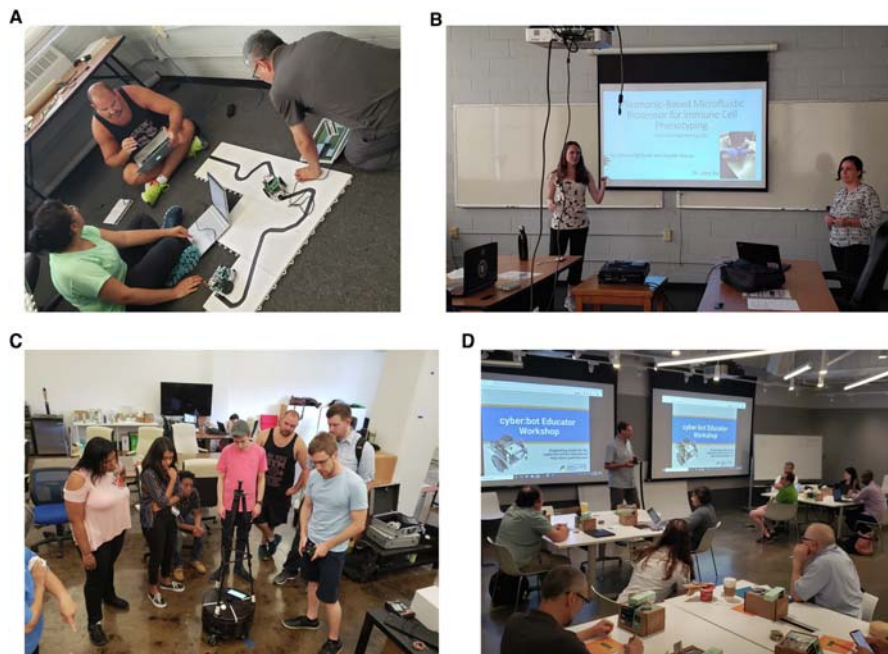


Fig. 1 Workshop participants: (A) conducting hands-on learning activities during guided training; (B) discussing their research; (C) visiting a startup robotics company; and (D) attending a robotics workshop by an industry professional.

involving graduate, undergraduate, and high school researchers, and faculty members in four different laboratories, i.e., dynamical systems, bioengineering, mechatronics, and robotics labs. These collaborations afforded the opportunity for teachers to observe how research is conducted and it allowed them to enact their newly acquired knowledge and skills from the guided training. They had ample opportunities to employ their learnings in authentic research settings while practicing the skills of communication, research, and teamwork.

Lesson plan development workshop

By participating in an NGSS-plus-5E lesson plan development workshop (for three days during the six-week period), teachers explored the 5E instructional model with the three-dimensional model of NGSS. This was done to allow the teachers to utilize their laboratory experiences from collaborative research to develop lesson plans grounded in the NGSS and 5E frameworks. The teachers first had an opportunity to support a group of teachers from another program by attending their lessons as students and providing them with feedback to improve lesson designs. In this process, they acquired skills to develop their own NGSS-plus-5E lesson designs. Attending these workshops also allowed the teachers to experience the real-world enactment of NGSS through the process of understanding how to approach research questions and examine the role of three dimensions (CCC, DCI, and SEP) in addressing these questions. They could transfer this learning from the workshop to their collaborative research, ensuring learning of research concepts. Following the workshop, teachers designed several novel lesson plans drawing knowledge from their research collaborations, i.e., the underlying concepts of research projects informed the NGSS-grounded lesson designs. For instance, a two teacher team conducted research on designing a user-friendly smartphone app to control a robotic car. The app consisted of various features that permit users to command the robot in a natural and intuitive manner with minimal effort. An Arduino microcontroller and a breadboard populated with required electronic and sensing components made up the control system hardware interfaced to the robotic car. Using the same principles of microcontrollers, electronics, sensing, programming, and a simple smartphone app, the teachers created a lesson plan that encouraged students to learn, develop, and apply their engineering skills to create a working windmill. The lesson embeds opportunities for students to improve their design to increase the windmill energy output (see [Table A.I](#), [Appendix A](#), for a detailed lesson plan). Another teacher collaborated on a research project to develop a rehabilitation system to monitor finger grip force for stroke patients. This system used a force sensor, a webcam, a rehabilitation device, all interfaced with an Arduino microcontroller, Unity game engine (for a user-friendly set up), and a machine learning algorithm to accurately measure the grip functionality of hand and arm movement. Equipped with her new understanding of force sensors, Arduino, and Unity gaming, the teacher created a lesson that encourages students to design different types of bridges and strategically mount force sensors to monitor the amount of load. The lesson is facilitated through examples of real bridge collapses, which can allow students to explore various causes that may lead to bridge failures (see [Table A.II](#), [Appendix A](#), for a detailed lesson plan).

Industry-based experiences

During the four weeks of collaborative research phase, teachers participated in workshops and talks by industry partners as well as informational visits to technology startup and industry sites. This afforded teachers with additional insights into industry operations and ability to associate practice with theory and *vice versa*. The visits provided teachers with exposure to startups, research facilities, and technology consulting companies that focused on generating a wide variety of engineering solutions. For instance, teachers visited: (i) a robotics-plus-virtual reality startup that supports real-estate marketing companies by creating 360 degree virtual walk-throughs; (ii) a physical rehabilitation research facility that is engineering and evaluating cutting-edge technologies for motor recovery; (iii) a research and development company that conducts innovative robotics engineering projects for several federal agencies; and (iv) a technology consulting startup of a graduate of the host engineering school. A two-day hands-on robotics workshop was conducted by an educational robotics company that allowed teachers to build and program a robot using a block-based programming interface. Moreover, a one-day hands-on microcontroller workshop was conducted by another educational electronics vendor that allowed teachers to interface a Micro:Bit device with a variety of components for their monitoring and control.

Classroom implementation support and academic year activities

After the six-weeklong summer professional development workshop, teachers enacted their NGSS and 5E aligned STEM lessons in classrooms or newly developed robotics clubs. Project faculty members and graduate researchers collaborated with and supported the teachers in planning and organizing their robotics activities. Graduate researchers' presence also allowed students participating in the mechatronics- or robotics-based business idea competition to consult and improve their engineering design and be more successful in such activities. Additionally, under the cognitive apprenticeship model ([Collins, 1991](#)), graduate researchers shared their research and professional experiences in STEM with the students. This rendered an opportunity for students to interact with researchers and learn more about STEM career pathways. Teachers sought to create and conduct varied activities to engage students during the academic year. For instance: (i) a teacher mentored a 12-member student team, a third of them girls, for a project-oriented educational robotics contest; (ii) two high-school teachers launched a robotics club with 17 boys and 23 girls who worked with various robotics kits; and (iii) another high school teacher worked with 42 boys and 39 girls on robotics. She

received 12 additional robotics kits from her school's support, allowing her students to learn and explore how to physically build robots, code, and control the robot motion. Students were also given opportunities to prototype and pitch their ideas in a technology business idea competition.

Technology business idea competition

At the end of each academic year, a technology-based business idea competition was organized whose purpose was to encourage students to explore real-world challenges in their communities, develop novel solutions for them, and present their innovations. All teachers who participated in summer workshops were challenged to mentor a team of students from their schools to partake in this activity and present their ideas in front of their peers and a panel of judges. Specifically, students were encouraged to seek out, from their home, neighborhood, school, or other familiar environments, a problem that could possibly be solved by employing mechatronics or robotics concepts. Developing and honing students' entrepreneurial skills was an important aspect of this effort, thus they were required to analyze and illustrate the market potential of their solution. Several illustrative mechatronics and robotics design submissions are reviewed in [Illustrative examples of student projects](#) subsection.

Description of the competition

Under the PBL framework, the competition was designed thoughtfully such that it integrates applications of various project elements including mechatronics, robotics, and entrepreneurship concepts in alignment with STEM learning and engineering design. Students used the robotics kits that were provided to their teachers for classroom use and provision of additional material for the prototype was limited to under \$20. Each team was represented by three to five members and the prize for the competition was in the form of a paid internship at a local startup business. Students prepared for the competition by participating in the following steps.

Identifying a problem: The competing teams identified problems in their home, school, or community environment that were deemed solvable by using an innovative technology product (specifically mechatronics or robots in this case). They conducted surveys to identify potential opportunities and their possible impact for the targeted environment. The survey constituted an important element for the competition, thus the students were encouraged to be diligent about it. Next, they formulated problem statements, brainstormed potential solution strategies, and selected one they deemed the best. Students were advised that the solutions can be in the form of an enhancement to an existing device, an incorporation of different technologies, or be entirely novel.

Designing and building: After identifying the problem and potential solution, teams designed and prototyped mechatronics or robotics devices. At this stage, they built, programmed, and tested the device ahead of the competition. Their mechatronics devices and robots were required to be mechanically stable and serve practical needs in real-world settings. The teams submitted videos demonstrating the performance of their prototypes. See [Fig. 2](#) for several robot designs entered in the competition.

Marketing: As a final step, teams were required to create marketing strategies for the designed prototype by employing concepts such as identifying consumer base, setting a product price, steps to market their product, advertising their product, budget, resources, etc. After conducting a thorough research on these aspects, the teams created a product pitch to persuade people to buy their product. They also gave their product a name, which was also their team's name and a part of their marketing pitch. Students shared all the relevant details in the form of a presentation and a reflective written assignment at the competition.

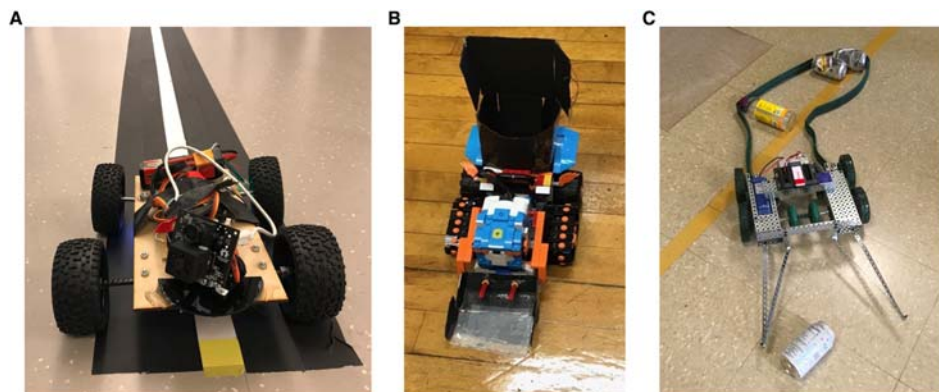


Fig. 2 Examples of student projects (A) TrackrBot, (B) Clara-BOT, and (C) RoBoiz.

Internship: Each year, one student from each of the two winning teams was awarded the opportunity to participate in a two-week, paid, hands-on, summer internship at a local company that supports young entrepreneurs. The winning teams and their teachers were free to nominate the internship recipients. This activity was incorporated to provide an intimate familiarity with technology startup industry to students so that they could experience and contribute to a company founded on the integration of entrepreneurship and technology. The project team expected this to be an effective contributor in shaping students' early experiences for gauging and developing engagement in technology startups.

Project results

Over the 3-year period of this effort, the summer workshop was attended by 30 teachers. The "self-identified" demographic information for teacher participants is summarized in [Table 1](#).

Evaluation procedure

The project evaluation was conducted by an external evaluator through the following activities.

- A survey was administered to the teachers to anonymously obtain responses and analyze data pertaining to their level of familiarity with the topics of mechatronics and entrepreneurship. The survey contained 35-items to assess teachers' familiarity with mechatronics skills, concepts, and devices introduced during the workshop. The survey also included seven items on familiarity with areas of business, intellectual property, and innovations. In addition to the familiarity surveys conducted by the external evaluator, the project team administered a technical quiz to assess teachers' math/science concepts and their understanding of engineering fundamentals *vis-à-vis* mechatronics. While the familiarity surveys were administered through an online mechanism, the technical quiz was conducted in a paper-pencil format. All pre-tests were conducted on the first day of the two-week hands-on training while all post-tests were conducted after the completion of the four-week collaborative research phase. This allowed the project team to assess any post-project improvements by comparing the pre-/post-tests. Note that teachers responded to all pre-/post-tests using self-selected unique three digit codes, which allowed anonymous data collection while permitting comparisons for pre-/post-tests for individual respondents.
- A survey was administered and analyzed concerning teachers' approaches in the educational process and their experiences during the program. To consider teachers' interests and concerns, they were prompted to indicate the level of priority for several areas in their teaching.
- Teachers were interviewed individually or in small groups and their research sites were visited to assess the authentic nature of their experiences. In the interviews, teachers detailed their accomplishments, experiences, and plans for future adoption of their learning for classroom. Visits to the research labs allowed the evaluator to observe teachers practicing what they had described during the interviews. Additional conversations with the project team members were also conducted.

Evaluation outcomes and observations

Substantial increases in teachers' familiarity with mechatronics skills and concepts and entrepreneurship topics were noted. Interview and lab visit data showcased teachers' enthusiasm in taking their learnings to classrooms for implementation. Further details are discussed below.

Mechatronics skills and concepts: While at the start of each summer workshop, teachers reported moderate levels of familiarity with the mechatronics topics, by the conclusion of each summer workshop, they reported higher levels of familiarity with the mechatronics topics. A 35-item survey assessed teachers' understanding of mechatronics concepts and over the 3 years of project implementation, teachers' self-perception improved from an average of 1.58/4.0 (pre-workshop) to 2.91/4.0 (post-workshop). The data broken down by each year of the project is reported in [Table 2](#) below.

Table 1 Demographic information of the participating teachers.

Year	Total	Gender		Race			
		Male	Female	White/Caucasian	African American	Asian	Others
Y1	9	6	3	4	4	0	1
Y2	11	4	7	5	3	3	0
Y3	10	6	4	6	0	3	1
Total	30	16	14	15	7	6	2

Table 2 Teacher's level for familiarity with the mechatronics topics.

	<i>Average of pre-test on teachers' self-perception of familiarity with project topics for mechatronics (4-point scale)</i>	<i>Average of post-test on teachers' self-perception of familiarity with project topics for mechatronics (4-point scale)</i>
Year 1	1.82	2.97
Year 2	1.25	2.72
Year 3	1.73	3.07

Table 3 Teacher's level for familiarity with the entrepreneurship topics.

	<i>Average of pre-test on teachers' self-perception of familiarity with project topics for entrepreneurship (4-point scale)</i>	<i>Average of post-test on teachers' self-perception of familiarity with project topics for entrepreneurship (4-point scale)</i>
Year 1	2.08	3.06
Year 2	1.25	3.05
Year 3	2.11	3.11

Entrepreneurship concepts: While at the start of each summer workshop, teachers reported moderate levels of familiarity with the entrepreneurship-based topics of the project, by the conclusion of each summer workshop, they reported higher levels of familiarity with the entrepreneurship topics. A seven-item survey assessed teachers' understanding of entrepreneurship concepts and over the 3 years of project implementation, teachers' self-perception improved from an average of 1.79/4.0 (pre-workshop) to 3.07/4.0 (post-workshop). The data broken down by each year of the project is reported in [Table 3](#) below.

Technical quizzes: Teacher's understanding of scientific, mathematical, and engineering foundations of mechatronics concepts improved based on the analysis of pre-/post-technical quizzes. The average of their scores over 3 years was: 45.90% (pre-workshop) and 59.65% (post-workshop); yielding 29.97% improvement.

- In Year 1, pre-to post-test scores changed from 49% to 64%, yielding an improvement of 30.6%. The analysis of a paired *t*-test on the quiz results produced a *t* value of 3.702 with a *p* value of 0.006. This outcome indicates that the null hypothesis can be rejected and the improvements in the quiz results are statistically significant at a level higher than 99%.
- In Year 2, pre-to post-test scores changed from 47.2% to 58.5%, yielding an improvement of 23.9%. The analysis of a paired *t*-test on the quiz results produced a *t* value of 3.38 with a *p* value of 0.007. This outcome indicates that the null hypothesis can be rejected and the improvements in the quiz results are statistically significant at a level higher than 99%.
- In Year 3, pre-to post-test scores changed from 41.67% to 57%, yielding an improvement of 36.8%. The analysis of a paired *t*-test on the quiz results produced a *t* value of 1.99 with a *p* value of 0.005. This outcome indicates that the null hypothesis can be rejected and the improvements in the quiz results are statistically significant at a level higher than 99%.

Approaches in educational processes: Results of the surveys indicated that teachers prioritized for their students: learning fundamentals; engaging in problem solving abilities; performing team work; honing communication skills; gaining business-related skills; practicing research skills; familiarizing with STEM-related careers; and staying abreast with technological innovation. While these are not areas in which significant pre/post changes were anticipated, Year 1 found that an increase from pre-to post-workshop survey results from an average of (3.81/5) to (4.46/5) yielding an increase of 0.65.

Interviews and lab visits: The evaluator noted that the teachers were enthusiastic in answering the questions and explaining what they had learned. During the collaborative research work, teachers explained their research topics and demonstrated what their mechatronics or robot designs were able to accomplish and described painstaking efforts in analyzing motion, fabricating devices, developing computer instructions, and testing electronic interfaces. A teacher emphasized on the importance of creating personalized assessment goals for students to support their learning progression and surfacing of misconceptions. Several teachers also shared that improvement in coding and programming capabilities is crucial for success in their classrooms. Based on their experiences with the professional development program, the teachers communicated the importance of real-world experiences in making learning relevant to classrooms, tailoring material to the context of their students, and employing experiences with entrepreneurship to make learning more meaningful. In discussing the take-away, teachers commented on their own learning experiences, particular technologies they plan to use (programming, soldering, 3D printing), and concrete projects. They also appreciated the opportunity of ongoing learning.

Illustrative examples of student projects

The competing student teams developed a variety of robotics and mechatronics-based solutions for problems in their environment. Several examples are highlighted below.

TrackrBot: Based on the surveys for their whole school and specifically the students in school's sport programs, this student team identified that athletes cannot keep pace with training and track their progress. To address this challenge, the students suggested a solution in the form of a robot that keeps a programmed pace profile for runners who are moving around the track and updates their progress on a smart-phone app. They created a mobile robot whose time, pace, and distance parameters can be specified based on the training profile for specific athletes such that when a runner finds the robot go past them, they would know to run faster. The robot was also integrated with a camera to record the runner's form. During the robot development process, students gained deeper understanding of STEM topics such as distance, speed, average, algorithm development through coding, and the functionality of various components used in the robot design, among others. Apart from the excellent prototype design, the team also provided a convincing presentation, an engaging video demonstration, a compelling business case, and well-considered marketing strategies to promote their product.

Clara-BOT: Based on the interviews of their peers, this team recognized a need to improve their school cafeteria that often gets littered around. A solution was to build a robot to pick up trash while also being educational and entertaining. The team created a robust engineering design of their Clara-BOT and endowed it with WiFi networking to communicate with other robots and facial recognition technology for school security application. In addition to a collaborative effort on their presentation and business pitch, the team showcased the robot as an environmentally-friendly solution for similar challenges in the future.

RoBoiz: This team sought to solve the problem of bottles and cans that litter the school floor by adopting the design thinking process to create their RoBoiz prototype. They considered the stages of "empathize" to identify the challenges in their school, "define" to formulate a problem statement, "ideate" to come up with potential solutions and robot designs, "prototype" to build their robot, and "test" to examine their robot design. Through this project, the students not only encouraged a cleaner environment but also promoted recycling. Their solution employed a wide range of STEM and entrepreneurship concepts such as waste collection and management, recycling, coding, environmental protection, marketing pitches, market research, user research, etc.

Similarly, by conducting surveys of their school peers or interviewing their classmates, many other teams identified challenges for entrepreneurial explorations under the umbrella of mechatronics and robotics technologies. For example, one student team opted to devise a "Silent Classroom" as a solution for distraction in learning due to disruptive noise or behavior inside or outside the classroom. The team proposed to use Bluetooth headphones integrated with a sensory system that can enable students to hear only their teacher's instruction by pressing a button on the sensory device. When students engage in group work, they have the option to hear their peers in close proximity by pressing another button on the sensory device. Apart from gaining a deep understanding of STEM concepts related to sound, engineering design, and coding, students also employed concepts of entrepreneurship by developing marketing strategies and a pitch video, all to emphasize the need of this solution and how it can help students, teachers, and schools. The "Anima" team identified the problem of lack of access to feminine hygiene products and the undue disruption and embarrassment that result when one goes about to obtain them. The team created a smart-phone app and a vending machine to facilitate the ease of availability of the product. The students can order the product using their student account through the app. Next, at the vending machine, students perform required touchpad interaction and present their ID for user identification and validation, following which the vending machine dispenses the items using the motors. In developing this prototype, the students utilized STEM concepts of coding to develop the app, and capacitors and touch sensors for the vending machine. Moreover, they also took a stance toward social justice by brainstorming a solution for a poignant problem such as lack of access to hygiene products for female students. The team also built a strong business case around the problem and suggested that such a solution can be adopted beyond schools in workplaces and public spaces that have women employees or visitors.

In written reflections about their project experiences, team members indicated that participation in this activity allowed them to learn how to address myriad social, organizational, and technical challenges. For example, students indicated that they learned teamwork strategies, including communication, brainstorming, and patience, to effectively plan, organize, schedule, and manage their time and activities. Moreover, they learned about robotics engineering and programming, 3D printing, hardware and software debugging, creating and editing videos, app development, etc.

Discussion

Based on the outcomes and observations of the research and evaluation of the program, an evident improvement in outcomes of teachers' self-perception of familiarity with mechatronics and entrepreneurship concepts was noticed. A growth in their understanding and use of technical and scientific terms was noted as teachers used new vocabulary terms related to programming, methods of measurement, etc., during the external evaluator's visit to their research lab sites and on the technical quiz. As they worked with industry professionals and researchers, teachers developed their work habits and mentioned the importance of planning step-by-step to understand about the underlying connections of novel concepts. Teachers were enthusiastic about the real-world aspect of project integration in their classrooms and received opportunities of understanding of scientific, mathematical,

and engineering foundations of mechatronics concepts. The emerging themes showcased a shift in leading teachers to “doing science” as opposed to reading about it. They highlighted that the integration of entrepreneurship can add a pragmatic dimension that would motivate their students. For instance, they explored 3D printers not just to fabricate simple items but to use them in conjunction with science-learning areas and contextualizing to introduce newer methods to students. They appreciated the ongoing learning opportunity provided to them via this professional development program and believed that they must themselves be engaged in learning. Some teacher testimonials are shared below.

- “Knowing more about business and industry will make abstract concepts more relevant to students by placing them in a context and showing how they may be impacted.”
- “The students are always buying the latest innovative devices. And they always ask teachers about how these devices function. It is wise for teachers to have a basic idea about the industry.”
- “By tying the math I teach to industry, students will see more of the real-world connections and find it more meaningful/purposeful to study.”
- “If I am able to determine the trend in any given industry, I will be able to guide my students more in what they should pursue for college and what skills to develop to help them become successful once they join the workforce.”

The teachers also appreciated the annual technology business idea competition as it provided an additional opportunity for students to apply what they had learned and present it to a larger group of audience. For instance, the teachers reflected as below.

- “Being able to watch them think deeply about an issue and work together to come up with a solution, is exactly what education should teach them. I tried to take a very hands-off approach to my students’ involvement in the competition, because I wanted to see what they could accomplish all on their own, and I have been fairly impressed considering that we don’t have a program or course of study or even a team dedicated to robotics and coding.”
- “As a result of our work together students demonstrate curiosity, stamina, the ability to rethink decisions and reiterate models. They engaged in the engineering design process and were readily able to work with a variety of classmates on varying assignments. They used their specific talents and gifts to help come up with solutions and solve problems.”

The graduate researchers’ school visits assisted teachers in organizing robotics activities and teams and enabled students interested in robotics to consult with researchers about expectations, commitments, and technical aspects to be successful in such activities. Collectively, these educators impacted over 500 students, with approximately 50% being women and a large proportion from groups underrepresented in STEM disciplines. These teachers acted as mentors to the robotics clubs and they developed and implemented lessons and activities based on the guided-training curriculum of the program. They used the mechatronics kits provided through the project and procured additional kits through their schools. The activities of student clubs were organized once or twice every week. The students actively worked on designing, programming, and testing robots. Through lessons, robotics clubs, and annual competitions, students explored multiple opportunities to design solutions for problems in their schools and communities, in turn developing knowledge and skills for STEM learning, engineering design, and entrepreneurship, all leading to greater engagement and interest in STEM pathways. Winners of the annual competitions were awarded internships following which one student reflected that the activities enabled her to improve research, reading, and writing skills and that she felt confident in finding relevant information and prioritizing important tasks a lot faster. Another student reflected about his experience of creating new opportunities for himself and following his dreams, even when they may not seem realistic. To summarize, over 3 years the project provided numerous, hands-on learning opportunities for teachers and students to become familiar with and experience mechatronics, robotics, and entrepreneurship concepts situated in real-life problems. The positive outcomes indicate that participants increased their knowledge, interest, and achievement in these concepts, with a potential of continued deployment in classrooms for the future.

Conclusion

This chapter discusses the design, implementation, and evaluation results of 3 years of a summer professional development program that focused on familiarizing teachers with mechatronics, robotics, and entrepreneurship concepts to facilitate meaningful and engaging STEM learning. During a six-week period, teachers immersed in hands-on learning, collaborative research, and industry-based experiences. Teachers also participated in a three-day lesson development workshop that focused on the design of NGSS-plus-5E lessons and their practice enactment with peers. The evaluation data showed an improvement in teachers’ familiarity with mechatronics, robotics, and entrepreneurship concepts, and an increased enthusiasm in implementing the real-world lessons in classrooms. Overall, 30 teachers received the guided training, research experiences, and industry exposure. Collectively, these educators impacted over 500 students, vast majority of who come from weaker socio-economic backgrounds, with approximately 50% being women and a large proportion from groups underrepresented in STEM disciplines. The teachers appreciated the project and continued to use the mechatronics and robotics kits provided to them beyond the program and appended the kits using resources from their own schools. The activities were organized 1–2 times a week after school hours and engaged the students in designing, programming, and testing robots and learning the various related STEM concepts.

Appendix: A

Table A.I NGSS-plus-5E lesson plan on alternative energy using windmills.

Lesson description: In this lesson students will utilize their engineering skills to create a working windmill. They will have opportunities to redesign their windmill to increase its energy output and share detailed information about how the windmill functions. After working with the windmill, students will be tasked with researching a specific form of alternative energy and present their findings to their peers.

Learning outcomes: Students will be able to:

- Explore the process of engineering design by creating a scaled model of a windmill and evaluate its energy output.
- Identify human impact on the environment by analyzing various data sources, including satellite imaging data.

Prior student knowledge:

- The concept of alternative forms and sources of energy
- Consequences of human impact on natural systems, e.g., deforestation, fossil fuel consumption, overpopulation, pollution, etc.
- Operation of a multimeter and other varied electrical components and equipment

Possible preconceptions/misconceptions:

- Electricity comes from the wall socket where you plug in and it is created (no conservation of energy)
- Alternative forms and sources of energy are not readily available for everyday use
- Wind turbines are not very large
- If it is not windy (wind power) or sunny (solar power) there will be no electricity available for those reliant on alternative energy sources

Performance expectation(s): HS-ESS3-4. Evaluate or refine a technological solution that reduces impacts of human activities on natural systems.

Science and engineering practices

Constructing explanations and

designing solutions

Disciplinary core ideas

ESS3.C: Human impacts

on earth systems

ETS1.B: Developing

possible solutions

Crosscutting concepts

Stability and change

Influence of science, engineering, and technology on society and the natural world

5E lesson plan:

Engage: Discuss in groups: What is the story behind turning on your light bulb? How many people does it take to turn on a light bulb?

Explore: Groups will prototype and design a scaled model of a windmill based on a scenario where the city mayor wants to use a new energy source using a turbine tower, motor, dowels, hub, multimeter, etc. This is to be done under a given budget to build a cost-effective turbine. Students will present their designs and run trials. Later, they will reflect on the following questions and upgrade their design.

- What was the dependent variable in this experiment? (e.g., voltage, energy produced, etc.)
- What were the variables you were manipulating? (e.g., type and amount of material, angle and surface area of blades, number of dowels, etc.)
- What design elements seemed to work best? What failed? Why do you think this occurred?

Explain: Understanding the windmill: parts (blades, low speed shaft, high speed shaft, anemometer, wind vane, controller, generator, yaw drive, brake) Questions: Why are the blades curved? What are some applications of windmills? What are some limitations of windmills? Which features are important for windmill design?

Elaborate: Research an alternative energy that is currently being used somewhere in the world and prepare a brief presentation. Explain how it compares to our traditional energy (burning of fossil fuels).

Evaluate: Present multiple iterations of the design and observations during each iteration. Explain the reason behind the changes in design. Submit reflections.

Table A.II NGSS-plus-5E lesson plan on bridges and forces.

Lesson description: In this lesson students will design scaled prototypes of different types of bridges and instrument them with force sensors. Students will analyze the effect of strategic placement of various force sensors, near the gusset plates of truss, on the structural integrity and stability of the bridge. As the students construct the prototype and instrument it, they will explore strategies to effectively monitor the amount of load forces in critical locations of the bridge.

Learning outcomes: Students will be able to:

- Explore and examine the causes of historical bridge collapses in the US.
- Calculate forces of compression and tension in their scaled prototype of a bridge.

Prior student knowledge:

- Concept of compressive and tensile forces on a bridge
- Knowledge of different types of bridges and historical bridge collapses in the US
- Familiarity with Arduino microcontroller interfacing and coding

Possible preconceptions/misconceptions:

- Only certain types of bridges can collapse
- It is difficult to determine forces acting on a bridge

Performance expectation(s): MS-ETS1-2. Evaluate competing design solutions using a systematic process to determine how well they meet the criteria and constraints of the problem.

(Continued)

Table A.11 NGSS-plus-5E lesson plan on bridges and forces.—cont'd

Science & engineering practices	Disciplinary core ideas	Crosscutting concepts
Engaging in argument from evidence	ETS1.B: Developing possible solutions	Influence of science, engineering, and technology on society and the natural world
5E lesson plan:		
Engage: Discuss in groups: (i) How do the forces of tension and compression affect a bridge? (ii) How do gravity and externally applied load, e.g., a truck carrying goods, affect a bridge?		
Explore: Using art supplies, students in small groups design scaled prototypes of different kinds of bridges, e.g., arch bridge, simple beam bridge, cantilever bridge, cable stayed bridge, suspension bridge, and truss bridge. Students calculate and quantify the force and resistance on the bridge and discuss solutions to design a better bridge. Based on their knowledge of compressive and tensile forces, they determine the location of force sensors (e.g., on street level above the gusset plate). Using the console of an Arduino microcontroller interfaced to the force sensors, students record the force sensor measurements corresponding to the placement of toy cars and trucks at different locations of their bridge prototype and use these measurements to calculate forces experienced by key elements of the bridge.		
Explain: Demonstrate an understanding of relevant vocabulary terms: tension, compression, buckling, load, live load, etc. Analyze graphs for different gusset plates and show an understanding of dependent and independent variables. Plot a graph of the demand to capacity ratio as a measure of bridge safety.		
Elaborate: Compare and contrast pros and cons of different kinds of bridges (simple beam bridges tend to sag in the middle, truss bridges can be strengthened by the assembly of interconnected triangles, etc.).		
Evaluate: Students will assume the role of an officer or a scientist, who evaluates the process of design and construction of new bridges, for the Federal Highway Administration or the National Transportation Safety Board. They will hypothesize about various sensors that can be used for condition monitoring of a bridge structure. Through an essay and a presentation, they will respond to the following prompts: Which kind of bridge has greater stiffness and why? Which type of bridge needs more support under the bridge deck and why? Can you comment on the tensile and compressive forces, including if they are same, for a: moveable, lift, bascule, or swing bridge? What is the bridge safety in terms of capacity ratio?		

Acknowledgments

This work is supported in part by the National Science Foundation grant RET Site EEC: 1542286. The authors thank the 30 teachers for their participation in this study. The authors also thank Dr. Sai Prasanth Krishnamoorthy, a doctoral student during this study, for his significant educational, instructional, mentoring, and research contributions to this effort. Moreover, the authors thank the following for their invaluable support: (i) the faculty mentors of teachers; (ii) graduate student collaborators of teachers; (iii) NYU Tandon School of Engineering's Center for K-12 STEM Education and Incubators Initiatives; (iv) NYU Entrepreneurial Institute; (v) industry partners; and (vi) external evaluator.

References

- AeA, 2005. Losing the Competitive Advantage? The Challenge for Science and Technology in the United States. American Electronics Association, Washington, DC.
- Alimisis, D., 2013. Educational robotics: open questions and new challenges. *Themes Sci. Technol. Educ.* 6, 63–71.
- Allen, C.D., Penuel, W.R., 2015. Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *J. Teach. Educ.* 66, 136–149.
- Barron, B.J., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., Bransford, J.D., 1998. Doing with understanding: lessons from research on problem- and project-based learning. *J. Learn. Sci.* 7, 271–311.
- Beary, V.E., 2013. The NFTE Difference: Examining the Impact of Entrepreneurship Education. NFTE. com.
- Benitti, F.B.V., 2012. Exploring the educational potential of robotics in schools: a systematic review. *Comput. Educ.* 58, 978–988.
- Brill, A.S., Elliott, C.H., Listman, J.B., Milne, C.E., Kapila, V., 2016. Middle School Teachers' Evolution of TPACK Understanding Through Professional Development. ASEE Annual Conference and Exposition.
- Bybee, R.W., Taylor, J.A., Gardner, A., Scotter, P.V., Powell, J.C., Westbrook, A., Landes, N., 2016. The BSCS 5E Instructional Model: Origins and Effectiveness. Colorado Springs, CO, Office of Science Education National Institutes of Health.
- Bybee, R.W., 2014. NGSS and the next generation of science teachers. *J. Sci. Teach. Educ.* 25, 211–221.
- Carbonaro, M., Rex, M., Chambers, J., 2004. Using LEGO robotics in a project-based learning environment. *Interact. Multimedia Electron. J. Comput. Enhanc. Learn.* 6, 55–70.
- Chen, F.H., Looi, C.K., Chen, W., 2009. Integrating technology in the classroom: a visual conceptualization of teachers' knowledge, goals and beliefs. *J. Comput. Assist. Learn.* 25, 470–488.
- Cioffi, N., Kulturel-Konak, S., Konak, A., 2014. "Anything is Possible"—Teaching Entrepreneurship in an Interactive K-12 Workshop, 2014 IEEE Integrated STEM Education Conference. IEEE, pp. 1–5.
- Collins, A., 1991. Cognitive apprenticeship and instructional technology. In: Idol, L., Jones, B.F. (Eds.), *Educational Values and Cognitive Instruction: Implications for Reform*. Desimone, L.M., 2011. A primer on effective professional development. *Phi Delta Kappan* 92, 68–71.
- Duncan, R.G., Cavera, V.L., 2015. DCIs, SEPs, and CCs, oh my!: understanding the three dimensions of the NGSS. *Sci. Teach.* 82, 67.
- Eguchi, A., 2012. Educational robotics theories and practice: tips for how to do it right. In: Barker, B.S. (Ed.), *Robots in K-12 Education: A New Technology for Learning*. IGI Global.
- Eguchi, A., 2021. Theories and practices behind educational robotics for all. In: Papadakis, S., Kalogiannakis, M. (Eds.), *Handbook of Research on Using Educational Robotics to Facilitate Student Learning*. IGI Global.
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., Yoon, K.S., 2001. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38, 915–945.
- Gay, G., Howard, T.C., 2000. Multicultural teacher education for the 21st century. *Teach. Educ.* 36, 1–16.
- Ghosh, S., Sabouri, P., Kapila, V., 2020. Examining the Role of LEGO Robots as Artifacts in STEM Classrooms (Fundamental). ASEE Annual Conference and Exposition.
- Harada, V.H., 2016. A practice-centered approach to professional development: teacher-librarian collaboration in capstone projects. *Sch. Libr. Res.* 19, 1–47.

- Hartmann, S., Wiesner, H., Wiesner-Steiner, A., 2007. Robotics and gender: the use of robotics for the empowerment of girls in the classroom. In: Zorn, I., Maass, S., Rommes, E., Schirmer, C. (Eds.), *Gender Designs IT: Construction and Deconstruction of Information Society Technology*. Springer.
- Herman, A., 2019. America's STEM crisis threatens our national security. *Am. Aff.* 3, 127–148.
- Hurst, M.A., Polinsky, N., Haden, C.A., Levine, S.C., Uttal, D.H., 2019. Leveraging research on informal learning to inform policy on promoting early STEM. *Soc. Pol. Rep.* 32, 1–33.
- Iskander, M., Kapila, V., Kriftcher, N., 2010. Using modern sensors in high school science labs to promote engineering careers. In: *GeoFlorida (Advances in Analysis, Modeling and Design)*.
- Kapila, V., Lee, S.-H., 2004. Science and mechatronics-aided research for teachers. *IEEE Control Syst. Mag.* 24, 24–30.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Melchior, A., Cutter, T., Cohen, F., 2004. Evaluation of FIRST LEGO League. Center for Youth and Communities, Heller Graduate School, Brandeis University, Waltham, MA.
- Milto, E., Rogers, C., Portsmouth, M., 2002. Gender Differences in Confidence Levels, Group Interactions, and Feelings About Competition in an Introductory Robotics Course. *IEEE Annual Frontiers in Education Conference*, p. F4C.
- Moore, T.J., Tank, K.M., Glancy, A.W., Kersten, J.A., 2015. NGSS and the landscape of engineering in K-12 state science standards. *J. Res. Sci. Teach.* 52, 296–318.
- Moorhead, M., Elliott, C.H., Listman, J.B., Milne, C.E., Kapila, V., 2016. Professional Development Through Situated Learning Techniques Adapted With Design-Based Research. *ASEE Annual Conference and Exposition*, p. 25967.
- Morales-Doyle, D., 2018. Students as curriculum critics: standpoints with respect to relevance, goals, and science. *J. Res. Sci. Teach.* 55, 749–773.
- Mosley, P., Kline, R., 2006. Engaging students: a framework using LEGO robotics to teach problem solving. *Inf. Technol. Learn. Perform J.* 24.
- Norton, S.J., McRobbie, C.J., Ginns, I.S., 2007. Problem solving in a middle school robotics design classroom. *Res. Sci. Educ.* 37, 261–277.
- Ottenbreit-Lefwich, A., Liao, J.Y.-C., Sadik, O., Ertmer, P., 2018. Evolution of teachers' technology integration knowledge, beliefs, and practices: how can we support beginning teachers use of technology? *J. Res. Technol. Educ.* 50, 282–304.
- Rahman, S.M., Krishnan, V.J., Kapila, V., 2017. Exploring the Dynamic Nature of TPACK Framework in Teaching STEM Using Robotics in Middle School Classrooms. *ASEE Annual Conference and Exposition*.
- Rahman, S.M., Chacko, S.M., Rajguru, S.B., Kapila, V., 2018. Fundamental: Determining Prerequisites for Middle School Students to Participate in Robotics-Based STEM Lessons: A Computational Thinking Approach. *ASEE Annual Conference and Exposition*.
- Riojas, M., Lysecky, S., Rozenblit, J., 2011. Educational technologies for precollege engineering education. *IEEE Trans. Learn. Technol.* 5, 20–37.
- Rusk, N., Resnick, M., Berg, R., Pezalla-Granlund, M., 2008. New pathways into robotics: strategies for broadening participation. *J. Sci. Educ. Technol.* 17, 59–69.
- Sanders, W.L., Rivers, J.C., 1996. *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.
- Sobhan, S., Yakubov, N., Kapila, V., Iskander, M., Kriftcher, N., 2007. Modern sensing and computerized data acquisition technology in high school physics labs. In: *Advances in Computer, Information, and Systems Sciences, and Engineering*. Springer.
- Sparks, D., 2002. *Designing Powerful Professional Development for Teachers and Principals*. National Staff Development Council, Oxford, OH.
- Tanner, K.D., 2010. Order matters: using the 5E model to align teaching with how people learn. *CBE-Life Sci. Educ.* 9, 159–164.
- Whitman, L.E., Witherspoon, T.L., 2003. Using LEGOs to Interest High School Students and Improve K12 STEM Education, 33rd ASEE/IEEE Frontiers in Education Conference. Boulder, CO, 76.
- Williams, K., Kapila, V., Iskander, M.G., 2011. Enriching K-12 Science Education Using LEGOs. *ASEE Annual Conference and Exposition*.
- Williams, K., Igel, I., Poveda, R., Kapila, V., Iskander, M., 2012. Enriching K-12 science and mathematics education using LEGOs. *Adv. Eng. Educ.* 3, n2.
- Wilson, S.M., 2013. Professional development for science teachers. *Science* 340, 310–313.
- You, H.S., Kapila, V., 2017. Effectiveness of Professional Development: Integration of Educational Robotics Into Science and Math Curricula. *ASEE Annual Conference*.
- Young, J.L., Young, J.R., Ford, D.Y., 2019. Culturally relevant STEM out-of-school time: a rationale to support gifted girls of color. *Roeper Rev.* 41, 8–19.

List of relevant websites

- ACT, 2018. The Condition of College and Career Readiness: National 2018. Available. <https://www.act.org/content/dam/act/unsecured/documents/cccr2018/National-CCCR-2018.pdf>.
- Barone, R., 2020. The State of STEM Education Told Through 16 Stats. Available. <https://www.idtech.com/blog/stem-education-statistics>.
- Chaney, B., 1995. Student Outcomes and the Professional Preparation of Eighth-Grade Teachers in Science and Mathematics. NSF/NELS: 88 Teacher Transcript Analysis. Available. <https://files.eric.ed.gov/fulltext/ED389530.pdf>.
- Lazio, R., Ford Jr., H., 2019. The U.S. Needs to Prepare Workers for STEM Jobs. Available. <https://tinyurl.com/Workers4STEM>.
- NGSS Lead States, 2013. Available. <https://www.nextgenscience.org/>.
- PBLworks Buck Institute for Education, What is PBL? PBLWorks. Available <https://www.pblworks.org/what-is-pbl>.
- STEM Teaching Tools, Teaching Tools for Science, Technology, Engineering and Math (STEM) Education. Available <https://stemteachingtools.org/sp/limfn>.
- Website of Mechatronics @ NYU Tandon, Mechanical and Aerospace Engineering Department, NYU Tandon School of Engineering. Available <http://engineering.nyu.edu/mechatronics/>.
- Weiner, B., 2018. Why the U.S. Has a STEM Shortage and How We Fix It (Part 1) [Online]. Available. <https://tinyurl.com/STEMShortage-Fix>.

CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program

Leticia Perez^a, Heather F. Clark^b, Roxana Hadad^a, Imelda Nava^a, and Maureen Giannotti^a, ^aUCLA Center X School of Education & Information Studies, Los Angeles, CA, United States; and ^bUCLA School of Education & Information Studies, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Dual challenge for computational thinking to address	312
The STEM + C3 urban teacher residency program	313
Preparing teachers with CT pedagogy	313
Preparing teachers with CT for equity pedagogy	314
Learning spaces for teacher candidates	316
How teachers connect CT, equity, and data	317
Conclusions	318
Acknowledgment	319
References	319

Dual challenge for computational thinking to address

Today's youth will live and work in a world that is heavily influenced by computing principles (Barr and Stephenson, 2011) and by the misleading data artifacts integrated in current computing that maintain and perpetuate social inequalities (Bradford, 2017; Robinson, 2019). This dual challenge is one that, in today's classrooms, educators are largely not being prepared to address in their teaching and therefore, students are largely not being prepared to encounter (Kessler, 2007; Fraillon et al., 2014). In this paper, we present computer science (CS) teacher preparation and orientation approaches, the role of social justice within computational thinking (CT) and CS spaces, an introduction to a CT for Equity Framework and an introduction to how one teacher education program at the University of California, Los Angeles (UCLA) is preparing educators to use CT to address this dual challenge.

CT is quickly becoming a prerequisite skill for most endeavors of the 21st century (Wing, 2008). In the last 20 years, nearly every field related to science and mathematics has seen the growth of CT, and sense-making around data is demanded of engaged community members (Buitrago Flórez et al., 2017; Utts, 2015). This shifting landscape requires that science and mathematics education shift toward CT, which in turn necessitates that teacher preparation programs respond by redesigning curriculum, clinical experiences, and mentor development. Computational thinking refers to the thought processes involved in expressing solutions in a way that can be carried out by a computer (Cuny et al., 2010; Aho, 2011; Lee, 2016), with the distinction that CT additionally describes many problem-solving skills involved in CS (CS for CA, 2021). CS is "the study of computers and algorithmic processes, including their principles, their hardware and software designs, their [implementation], and their impact on society" (Tucker et al., 2003, p. 6). The K-12 CS Framework places CT at the core of CS, with CT comprising four of its six practices and stating that "the most effective context and approach for developing computational thinking is learning computer science; they are intrinsically connected" (K-12 Computer Science Framework, 2016, p. 69).

Thanks to the "CS for All" movement, great effort has been put forth to promote CS education. While this effort occurs at the national-level, we situate our response in the context of the state of California. To increase the number of highly qualified CS teachers, California developed an alternative CS authorization route for teachers in 2019 that requires secondary teacher candidates to complete coursework in five domains of computing content. This coursework is common to (a) the introductory CS undergraduate programs and (b) the high school Advanced Placement course that introduces students to the foundational CS concepts of the field and the impact of computing on society. Many of the current secondary level CS supplementary authorization preparation programs across California are offered through a CS department, often in collaboration with the teacher education program faculty and staff, and offer coursework heavily situated within the domain of CS.

These programs are typically designed to fulfill both the credentialing requirements set forth by the California Teacher Credentialing Commission and expose teachers to introductory computing principles and practices typical for high school computing courses such as Exploring Computer Science or Advanced Placement Computer Principles. At present, mathematics teachers within California are permitted to teach CS courses regardless of their CS experience. California's CS Secondary Authorization increases content preparedness for math teachers and broadens the opportunity for other disciplines to provide CS instruction within their school sites. Historically, a) CS educator recruitment approaches provided in-service teachers professional development and curricular support that range from a few days to several weeks (Liu et al., 2011) or b) involved recruiting CS professionals to the classroom. States and local districts continue to innovate approaches to address the lack of CS teachers in secondary spaces. However, equity and social justice issues are rarely the central organizing theme for science, technology, engineering, and mathematics (STEM) teacher education programs and largely absent from CT/CS coursework (Mills and Ballantyne, 2016).

The foundational argument of this paper is that secondary CS courses and teacher preparation for CS courses are fundamentally inadequate to address the dual challenge of preparing students to use computing principles, and to do so in ways that apply a critical framework around societal inequities. We advance a framework of preparing educators to utilize CT to address these challenges. In the sections that follow, we first explain the structure of the program designed at UCLA by outlining our response to a set of barriers educators face in preparing students to critically use computing principles. This is followed by a brief set of examples of the pedagogical practices that teacher candidates have developed through their participation in UCLA's program. We contribute an approach to preparing and empowering social justice educators to support students understanding and using CT in their work and their lives.

The STEM + C3 urban teacher residency program

UCLA's Teacher Education Program (TEP) has responded to the need for more critical computing education through the development of an urban teacher residency program called STEM + C3. The three Cs referenced in the program name represent a Community of Practice, Computational Thinking, and Computer Science secondary authorization. This program aligns a humanizing teacher education program (Bartolome, 1994; del Carmen Salazar, 2013) with K-12 academic achievement standards and state and national content standards. STEM + C3 prepares approximately 16 new secondary mathematics and secondary science teachers each year and situates pre-service teachers' learning from university coursework in the context of mentors' classrooms in small autonomous schools within urban communities. Teacher candidates begin their student teaching experience on the first day of school and continue with the same mentor teacher and classes through the entire school year.

The residency program employs a gradual release of responsibility model (Fisher and Frey, 2013), where pre-service teachers observe during the first few months of their student teaching, co-plan, and co-teach during the next few months, and eventually take responsibility for two of their mentor's teaching periods for the remaining months of the school year. After one year of coursework and student teaching, pre-service teacher candidates obtain their preliminary teaching credential in a single subject as well as secondary authorization in CS, and begin their teaching career as full-time teachers of record. During the second year of the STEM + C3 program, teachers continue coursework toward the completion of a master's in education degree, as well as participate in a Community of Practice to support CT integration into their curriculum.

Preparing teachers with CT pedagogy

The design of UCLA's STEM + C3 program responds to several specific barriers educators face in developing pedagogical practices to teach CT and prepare students for lives in a world steeped in data and computing. Below, we explain three of those barriers followed by STEM + C3's response.

First, there are not enough well-trained teachers to meet the needs of schools. California created its own K-12 Computer Science Strategic Implementation plan, but currently faces a critical shortage of CS teachers (Sutcher et al., 2016) as schools struggle to respond to the increasing demand for CS courses in high school. In the year 2019–2020, 43 states reported teacher shortage in science and 47 in mathematics (including DC and Puerto Rico) (U.S. Department of Education Office of Postsecondary Education, 2000). The training for pre-service teachers is inadequate to fulfill this demand because the licensure pathways in many states do not require credentialed candidates to engage in CS coursework, learn CS pedagogies, or complete fieldwork in CS classrooms. As a result, there is a vacuum of quality CS teachers providing CS education for K-12 students, particularly students in high-need schools, which are far too often staffed by novice teachers (Borman and Dowling, 2008). In response, the STEM + C3 program is designed to train qualified STEM educators, 80% of whom are first generation educators of color who majored in STEM fields, and ensure retention of these teachers. All pre-service teacher candidates enroll in and complete four additional courses focused on computational thinking, programming in R, software design, and the design of the internet. This coursework preparation allows pre-service teacher candidates to apply for a supplementary secondary authorization in computer science.

Second, a focus on increasing the quantity of teachers qualified to teach a CS course is inadequate to ensure that all students learn about CT or computing principles. The Kapor Center reported that 42% of public high schools in CA offer at least one CS course, and provide enrollment access to approximately only 5% of high school students (Koshy et al., 2021). The report also stated that despite the \$65 million increase in CS funding, access to computing through stand-alone CS courses are not fulfilling a vision of "CS for All". In response, there has recently been a rally to integrate CT into core STEM courses and standards, in contrast to only having CT taught in a CS course. The goal of engaging in CT through core content areas is to increase student access and participation to CT and CS concepts and broaden both teacher and student perspectives around what CS can look and feel like (Weintrop et al., 2015). The meaningful use of CT opportunities in STEM curricula can deepen learning of mathematics and science concepts and reflect the many ways STEM professions are practiced (National Research Council, 2011; Sen-gupta et al., 2013).

The integration of CT into classroom standards is an important and necessary step to increasing student access and exposure to computing practices. The practice of "using mathematics and computational thinking" in Next Generation Science Standards (NGSS) reflects the growing importance of computation and digital technologies across the scientific disciplines (NGSS Lead States, 2013). The Common Core framework in mathematics also states that students should be able to use technological tools to deepen

their conceptual understanding ([National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010](#)). In addition, recent innovations in CS education such as Exploring Computer Science (ECS), Introduction to Data Science (IDS), and AP Computer Science Principles (APCSP) highlight the importance of inquiry and equity-based practices to broaden access and participation in computing. Typical high school CS curriculum exposes students to a range of computing concepts that use a programming language to construct coherence across topics. CT provides teachers an approach to connect across concepts in CS and other disciplines.

While expectations for teachers to use CT in math and science coursework are clear, the inclusion of these practices into standards, in and of itself, offers little guidance for teachers who will be required to realize them in their classrooms. As has been learned from teachers' efforts to enact other Science and Engineering Practices (SEPs) in the NGSS, ([Reiser, 2013](#)) teachers must be empowered with pedagogical practices and purposes to integrate CT in their classrooms. The STEM + C3 program trains teachers to integrate CT concepts into existing math and science teacher repertoires, especially in lower-level math and science courses, by helping to identify the overlap and the differences in CT practices with math or science practices. This is necessary because CT practices often overlap heavily with the practices of math and science, blurring the distinction between students' use of CT and students' disciplinary practices. Identifying the difference is necessary so teachers can meaningfully separate the practices and make those distinctions explicit to students. For example, in STEM + C3, teacher candidates explore Appendix F of the SEPs in the NGSS that states "Computers and digital tools can enhance the power of mathematics by automating calculations, approximating solutions to problems that cannot be calculated precisely, and analyzing large data sets available to identify meaningful patterns" ([National Research Council, 2013](#), p. 10).

Teachers new to CT often consider any use of technology as a CT practice or over-generalize CT as a simple problem-solving approach. For example, the steps to solve a mathematical problem algorithmically on paper versus writing an algorithm to solve the same problem are interrelated yet distinct tasks. In STEM + C3, teachers discover that examining and framing the NGSS's SEPs and the Common Core's Standards for Mathematical Practices with and without CT creates multiple opportunities to approach content from different perspectives and highlight distinctions between the disciplines. Identifying the overlap and unique distinctions are also necessary so teachers can engage in CT practices within typical math and science classroom contexts rather than within CS classrooms. STEM + C3 helps teachers see CT in computer-free tasks that are common in math or science classrooms, and further helps teachers identify and create "unplugged" CT opportunities ([Sim et al., 2021](#)).

Third, a focus on data and data tasks are absent in high school classrooms relative to the prevalence of data and data tasks in real-world contexts. The International Data Science Schools Project published a framework building a case for data science as a stand-alone course pathway that would ensure "sufficient understanding and appreciation of how data can be acquired and used to make decisions so that students can make informed judgments in their daily lives" ([IDSSP Curriculum Team, 2019](#), p. 7). Alternative mathematical pathways are an important part of increasing student CT and data literacy within STEM spaces. While there are notable examples fusing data, technology, and computing into lessons and courses (e.g., IDS, Concord Consortium Common Online Data Analysis Platform, Bootstrap, YouCubed), most students do not have access to such experiences. STEM + C3 responds by closely tying CT to data-related tasks, such as the data cycle or the statistical investigation cycle ([Bargagliotti et al., 2020](#)), with the goal of preparing students to "work with data and be critical consumers" ([Gould, 2021](#), p. S16).

Preparing teachers with CT for equity pedagogy

The design of UCLA's STEM + C3 program responds to additional barriers educators face in developing pedagogical practices that disrupt educational and social inequalities. Below, we explain two of those barriers followed by STEM + C3's response. We conclude this section by sharing the seven CT for Equity practices pre-service teacher candidates are taught in the STEM + C3 program.

First, CT/CS teacher education often fails to consider the diverse needs and contexts of schools and students which results in a lack of support in developing equity-oriented pedagogy. In a recent survey conducted by the Kapor Center and the Computer Science Teachers Association, 39% of CS teachers reported that did not see the importance of covering computing's role in perpetuating biases related to racism, sexism, and other social justice issues in the classroom ([Koshy et al., 2021](#)). A lack of curricular connection to students' lives has implications for student participation and outcomes, because students are more engaged with CS when it is presented alongside social justice issues ([Ryoo et al., 2019, 2020, 2021](#)). [Gordon and Heck \(2019\)](#) found that of CS teachers surveyed, "Fewer than one-quarter feel very well prepared to differentiate computer science instruction to meet the needs of diverse learners or to incorporate students cultural backgrounds into computer science instruction (p. 9)." The CS for Equity teaching practices and pedagogies support the integrations of student's lived experience and culture as an asset that also provides a differentiated pathway for students. The STEM + C3 program explicitly focuses on educational equity and justice, and provides pre-service teachers the opportunity to situate relevant social justice topics into curriculum, and use CT practices to open opportunities to explore issues of social justice within the classroom. This approach, called "CT for Equity", seeks to empower teachers to use CT in ways that explore, express, critique, and create artifacts about the world around them. It extends beyond the skills and meta-cognitive processes necessary for engaging in CT or computing tasks, and reaches toward classroom environments and curricula that affirm a student's rightful presence within the STEM classroom and the larger STEM community ([Barton and Tan, 2019](#)). CT for Equity entrusts teachers and students to use their data analysis and computing skills to develop and critique narratives about the

world around them as they become data storytellers who go beyond content, visualizations, and interpretation; they breathe life into the analysis, emphasizing the context, ethics, and implications of the data that are important to them and to their intended audiences. In STEM + C3, CT, data, and equity are inextricably linked, and this nexus serves as the coherent thread that weaves together different facets of our program.

A second barrier highlights that simply understanding CT terms (e.g., “abstraction” or “expressing algorithms”) does not clarify the purpose or process of learning CT or explicate the place of equity in that learning. Often, teachers new to CT and CS may perceive or position the incorporation of CT or computing concepts solely for career opportunity, as a means to recruit students into the STEM fields, or to increase enrollment. While CT plays a critical role in the STEM disciplines and fields, STEM + C3 is not focused on increasing workers in the STEM pipeline, but rather on CT as a tool for empowering students to use problem-solving and design skills that are relevant to the students’ content courses, and more importantly, to their lived experience.

STEM + C3 explicitly expands the purpose of CT in several ways. A means to foster CT for empowerment is through cognitive apprenticeship. Through cognitive apprenticeship (Collins et al., 1989), teachers scaffold the content and utilize modeling and coaching practices to assist students with developing the metacognition necessary to better connect computing and data with specific problem-solving approaches, all the while supporting the student’s sense of agency (Hadad, 2020). By encouraging the externalization of the “thinking” of CT, educators can reduce the mystery of how CT/CS practices are enacted and assist students in seeing themselves as computational thinkers.

CT embedded within data-rich experiences also provides opportunities to explore issues of social justice relevant to their student population and affirm all students’ identity and culture, and empower students to become agents of change. The interactional roles between data, CS, math and science situated within CT for Equity framework are reflected in Fig. 1.

Given the lack of familiarity with CT for many teachers, especially outside the context of a CS or programming curriculum, the STEM + C3 program developed basic definitions of CT practices as the foundation of teacher learning. CT for Equity is defined by seven practices, summarized in Fig. 2, which aim to expand CS access by creating different experiences for those learning about computing within their high school classes. This framework supports teachers in integrating CT concepts into existing required high school courses to increase access to computing and digital literacy. These definitions can be applied across many disciplines, have clear connections to data-centered tasks, and reflect a situated or critical lens (Kafai, 2016; Kafai et al., 2019). CT in these definitions are tools that assert a student’s rightful presence (Barton and Tan, 2019) in the classroom and an opportunity to liberate students from traditional pedagogical paradigms that privilege uniformity.

The definitions of these CT practices, as seen in Fig. 2, became the foundation for the STEM + C3 program and were developed in a multi-step iterative process that engaged the first cohort of teacher candidates. These functional definitions of CT support the NGSS SEPs as well as the California Common Core Standard (CCCS) Standards for Mathematical Practices, and build on existing knowledge largely shaped by International Society for Technology in Education CT Teacher and Leadership Toolkits and the K-12 Computational Thinking Frameworks. The CT practices hope to reflect how teachers make connections to data in STEM classrooms, leverage CT to support concept development, and use CT as a means to collaborate and communicate around issues of importance to students (K-12 Computational Thinking Framework, 2016). Specifically, they build on Jeanette Wing’s initial descriptors of CT (Wing, 2008) and recent findings from ECS and IDS evaluation reports (Gould et al., 2016) which indicate that ECS and IDS are powerful resources to improve STEM teacher preparation and development.

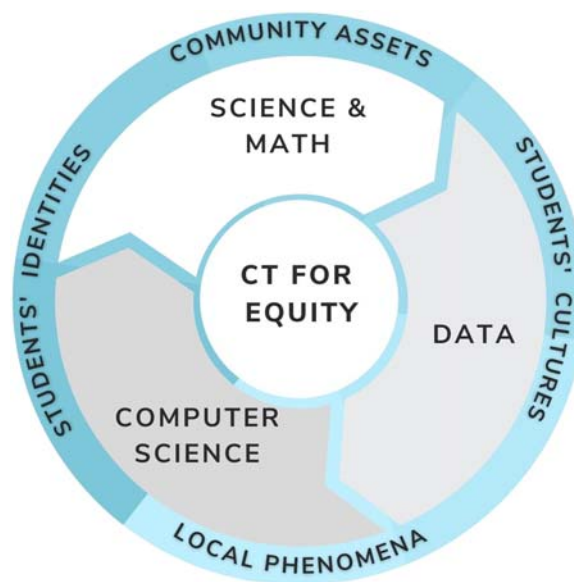


Fig. 1 Computational thinking for equity framework.

Computational Thinking for Equity CT for Equity seeks to empower students to use problem-solving and design skills that are informed by computing to explore, express, critique, and create artifacts about the world around them.	
Develop and Use Abstraction Consider and focus on key components and filter out unnecessary details to make problems easier to solve and includes acknowledging the benefits and consequence of generalizations and or lost details.	Algorithmic Thinking Recognize when a solution to a problem can be broken into step-by-step instructions. This set of instructions considers the audience, the purpose, and can be used by someone else or a computer.
Decomposition Break down a problem of interest into smaller, more manageable components, and the ability to describe this process in detail in order to make decisions, obtain help or find solutions.	Debugging and Evaluating The ability to identify and correct errors within an algorithm, and ensuring your solution is effective, feasible, and considers the impacts on various identities.
Exploring Equity Issues through Computational Thinking and Data Encourage alternative ways of solving problems. Promote inclusivity and provide universal access to all students, particularly those from underserved and intersectional populations.	Collaborate Around Computing and Data Work with others to understand, develop, and design computational artifacts or data visualizations to explore and explain phenomena, solve problems, or develop solutions that matter to students.
	Communicating with and About Computing and Data Describe, explain, and justify observations, patterns, and predictions around computational artifacts and data within the context of the problem as well as communicate applicable ethical considerations.

Fig. 2 Definitions of computational thinking practices used in the "CT for Equity" framework.

Learning spaces for teacher candidates

In this section, we outline the structure of the STEM + C3 program to explain the various learning spaces in which pre-service teacher candidates participate and how those spaces afford opportunities to learn the CT for Equity practices described above. The mission of both the UCLA TEP and STEM + C3 is to prepare teachers with the commitment, capacity, and resilience to work in and with the most underserved communities, and therefore, each of the five learning spaces share a central focus on achieving educational equity and social justice.

UCLA STEM + C3 builds on the knowledge and experience gained from the successful design and implementation of UCLA's IMPACT Urban Teacher Residency Program which was created in 2009 to develop highly prepared teacher leaders through a variety of coursework (language acquisition, learning theory, social justice) and field experiences redefining teacher quality (Kawasaki et al., 2020). These commitments and experiences structured the STEM + C3 program into two general categories: three academic learning spaces and two professional development spaces, as seen in Fig. 3. Academic learning spaces include UCLA TEP courses, a sequence of four CS courses, and the residency program in mentor teachers' classroom. Professional development spaces are the Partners in Practice, focused on reciprocal learning with mentors around questions of equity in education, and the Community of Practice, focused on experiencing and designing CT and data classroom lessons and pedagogies.

Teachers in the STEM + C3 program learn about CT for Equity primarily in the professional development space of the Community of Practice and the learning space of the CS coursework. The objectives of the Community of Practice are to a) operationalize the interdisciplinary nature of science, technology, engineering, and mathematics in STEM courses, b) integrate CT and CS concepts into NGSS or CCCS courses, and c) develop a critical mass of effective teacher-leaders who can bring high quality STEM teacher learning and professional development to scale. In a series of six workshops a year, teachers engage in lessons which integrate CT into math and science spaces and then reflect upon the pedagogical purposes, intent, and framing. In the second year of the program, teachers continue to explore problems of practice that center equity, data, and CT both within their Master's research project and the Community of Practice.

In the 12-months of the CS coursework, participants complete a sequence of courses that explores CT within the context of CS concepts such as data structures, foundation of programming using R, and the role of the internet, social media, and connectivity within digital devices. In the final course, participants work collaboratively to identify a large, multivariate dataset relevant to their

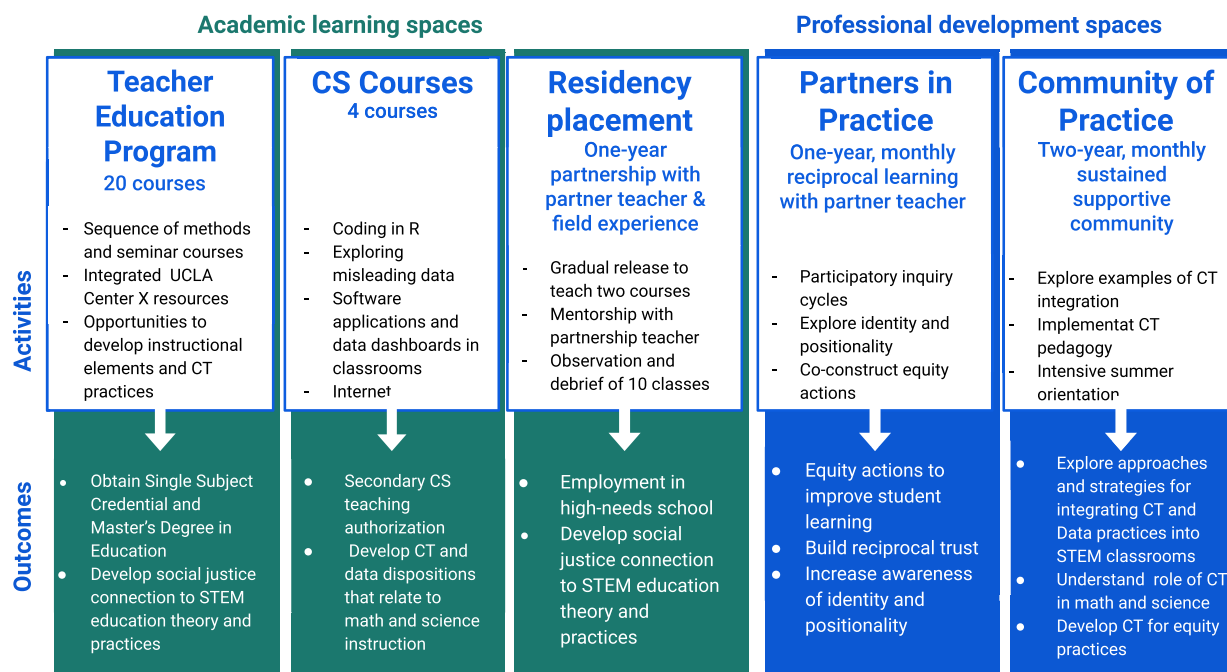


Fig. 3 Organizing structure of STEM + C3 program.

classroom content and context, and then design and build a Shiny web application (a data visualization dashboard). Teachers then share their data dashboards and accompanying lesson plan with each other in a showcase. Across this learning trajectory, CS and data opportunities within STEM classrooms are connected to CT.

The coursework sequence positions CT as a coherent thread that connects a student's STEM experiences together over the course of their K-12 education journey. Not only can CT connect together discrete CS experiences such as "hour of code" tasks to a robotics course, it is also used when programming sensors in physics or programming functions for different mathematical displays in trigonometry. At the center of all of these courses lies CT. Participants learn that CS is more than just developing students' capacity to use technology; through active engagement and deeper learning participants come to understand the role of CT, problem solving, algorithmic thinking, creativity, and collaboration in the context of CCCS for math and NGSS.

How teachers connect CT, equity, and data

To illustrate the impact of the STEM + C3 program on teachers' developing professional practices, we share selected examples of how preservice teacher candidates described the practices and purposes of integrating CT in their STEM courses. These examples come from pre-service teacher candidates who participated in the first cohort of the program and are drawn from their reflective writing in various learning spaces. The three examples (with pseudonyms for teacher candidates) were selected to highlight how teachers come to develop their own orientation to CT for Equity. Each excerpt was written in response to an open-ended prompt that invited teachers to reflect on how and why they integrate CT in their courses. In describing each of the excerpts from the pre-service teacher candidates, we link the candidates' ideas to the guiding framework of the STEM + C3 program to explicate the learning that this program facilitated.

By integrating CT students develop the tools they need to be a more critical member of society. The idea is that students don't just consume information but actually get to learn how information is gathered, analyzed and presented in order to critique the information that is constantly being communicated to them. One thing I really like about CT is that data practices can be very relevant to what is occurring in the students' community/world and therefore students are able to use real world experiences as they learn. *Jasmine, 12/2020 Programming Course Final*

CT makes STEM relevant and accessible to students by scaffolding problem-solving. Students can then connect this to problems of any size or relevance to their own lives. It also promotes students' independence, autonomy, self-driven inquiry etc because CT practices are skills that students can apply to many situations. *Ophelia, 3/2022 Software Course Final*

My students have interacted with computational thinking in many ways through different projects and activities in class. For example, my students did an endocrinology lab where they had to use their knowledge about feedback loops to figure out which specific hormones were injected into rats during a research study. My students actually worked to identify specifically how they engaged with the different computational practices: decomposition, algorithm thinking, debugging, pattern recognition, abstraction. Computational thinking allowed my students to think critically about solving problems. Computational thinking allowed my students to also find purpose in what they are doing or trying to accomplish during class. *Carissa, 2/2022 Community of Practice Reflection*

There are four aspects of Jasmine, Ophelia, and Carissa's reflective writing that we examine here to illustrate how they understand the purpose of CT and CT for Equity from their participation in STEM + C3. First, all three teacher candidates described a common purpose of supporting their students in becoming critical, independent problem solvers and citizens. Supporting students in developing problem solving was a universal goal for all ten candidates in this cohort (Clark et al., 2022). Problem solving is a foundational practice for both math and science disciplines and is a familiar entry point for teachers and reflects the ways in which the Community of Practice and CS Courses in STEM + C3 are designed around problems of practice that CT can address.

Second, Jasmine, Ophelia, and Carissa all wrote about how CT supports the relevance and importance of STEM learning and classroom tasks. These participants describe how students find relevance and a sense of purpose through connections to real-world and community issues. The STEM + C3 program designed a plethora of opportunities for teacher candidates to see the value of relevant teaching and to develop practices that supports relevant pedagogy. For example, STEM + C3 CS coursework kicked off with explorations designed to engage teacher candidates in dismantling STEM stereotypes and created counter-narratives through data and CT. Whether teachers explored intersectional representations of STEM community members, described search-engine bias within images tagged as "scientist", or used R language to explore Youth Risk Behaviors data from the Centers for Disease Control, every task in the CS courses was situated within the lives and experiences of both teachers and students. To communicate the commitment of the STEM + C3 program to equity and relevance, a major component of every CS course final was for teacher candidates to describe the implications of their new CS learning, how the CS learning was supported by CT for not only themselves, but also their urban student population.

Third, Jasmine wrote explicitly about the role of data in her teaching, and data practices are central to the learning experiences of candidates in the STEM + C3 program. In fact, all ten of the candidates in this cohort described relating course concepts to data, specifically providing data (primary or secondary) to students to achieve their instructional goals, and situating the analysis and interpretation of those data as necessary to learn concepts and/or engage in a science or math practice (Clark et al., 2022). Early on within the program, participants experienced how data storytelling through the data cycle could be used to structure classroom inquiry in a way that leveraged data as a tool to develop conceptual understanding in math or science. The enactments of CT through data varied among the teachers' reflections, whether it be with an exploratory data analysis task or examining techniques used to intentionally mislead audiences with data visualizations and representations. Interactions with data offered teachers an opportunity to simultaneously integrate CT, CS, and math/science concepts together. Data were treated as an object in which CT and Equity could be explored within the classroom. This approach is reflected in the emphasis of data across the four CS course sequence.

Fourth, Carissa described helping her students identify the differences and overlap between CT practices and science practices. This is exactly the skill she and her classmates learned for themselves, and she translated this to an activity for students. The practices detailed within NGSS and CCCS for math were written to be domain specific. While it may be easy for experts to identify how these practices are applied to other disciplines, it is not obvious for students. CT for Equity practices are generalizable and thus transfer to a variety of contexts beyond STEM to other content areas, as well as daily life experiences from finance to statistics around a favorite sports team or social media influencer.

The connections and distinctions between math/science practices and CT practices were modeled in a series of think-aloud activities by CS instructors, professional development providers in Community of Practice, and by the teachers themselves in their reflections. If students are to see themselves as capable users of technology and agents of change, they must be provided ample opportunities within a variety of contexts to incorporate CT into their repertoire of problem-solving approaches. The integration of CT into a student's academic experience may also positively impact the way they see themselves and their relationship to computing, perhaps inspiring future CS course enrollment or helping to foster a positive STEM identity.

Conclusions

This paper presented the UCLA STEM + C3 program's approach to teacher education which aims to prepare youth to live and work in a world that is heavily influenced by computing principles, and which highlights misleading data artifacts that maintain and perpetuate social inequalities. The guiding framework of the program is a set of seven CT for Equity practices that expand the notion of CT beyond problem solving assisted by a computer or other technology to include issues of educational and social equity and justice. The STEM + C3 program represents a major advancement in meaningfully preparing teachers to integrate CT and students' assets in core courses to support critical analysis and potential action in students' respective communities. Furthermore, experiences

such as STEM + C3 prepare teachers to advocate for the expansion of CS course offerings within their school sites, possibly leading to new types of computing units and courses that attract underrepresented student populations. The integration of CT within math and science classrooms provides students with exposure and confidence to pursue CS courses.

The STEM + C3 program will continue to be iterated upon and refined as it moves into its third year. It fills a void in a creative, novel way that is suited specifically for California. The field of CT education is currently in a similar place to the NGSS professional development in 2013, where a clear set of standards and framework had been laid out but education programs, professional learning providers, and educators lacked concrete tools and exemplars demonstrating authentic three-dimensional learning, where the disciplinary content knowledge, practices and cross-cutting concepts drive instruction. To continue this work, teachers will need curricula that model thoughtful integration of CT into math and science classrooms. There is a need for professional learning spaces in which educators critically examine the curriculum and its relation to their specific student population; teachers will need time and support to modify this curriculum for their particular context. CS education needs to better address the social inequities within computing and society at large. As we work toward the vision of “CS for All”, frameworks such as CT for Equity provide an opportunity to consider how we might go about doing this difficult work.

Acknowledgment

We would like to thank the invaluable intellectual labor and support offered by Suyen Machado, Monica Casillas, James Molyneux, and Beth Simon in the formation of the CS courses and the professional learning support provided by the UCLA Science project. This work could not be done without our amazing district liaisons, teacher candidate residents, and partner teachers who continue to inform this project. This work is supported by a grant from the US Department of Education (Award # U336S190038). The views and opinions expressed herein are those of the authors only, and do not represent the official views and opinions of the Dept. of Education.

References

- Aho, A.V., 2011. What is Computation? *Computation and Computational Thinking*, vol. 9. Ubiquity, pp. 1–8. <https://doi.org/10.1145/1895419.1922682>.
- Bargagliotti, A., Binder, W., Blakesley, L., Eusufzai, Z., Fitzpatrick, B., Ford, M., Huchting, K., Larson, S., Miric, N., Rovetti, R., Seal, K., 2020. Undergraduate learning outcomes for achieving data acumen. *J. Stat. Educ.* 28 (2), 197–211.
- Barr, V., Stephenson, C., 2011. Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *Acm Inroads* 2 (1), 48–54.
- Bartolome, L., 1994. Beyond the methods fetish: toward a humanizing pedagogy. *Harv. Educ. Rev.* 64 (2), 173–195.
- Barton, A., Tan, E., 2019. Designing for rightful presence in STEM: the role of making present practices. *J. Learn. Sci.* 28 (4–5), 616–658.
- Borman, G.D., Dowling, N.M., 2008. Teacher attrition and retention: a meta-analytic and narrative review of the research. *Rev. Educ. Res.* 78 (3), 367–409.
- Bradford, L., 2017. Fraudulent and misleading data. In: Allen, M. (Ed.), *The SAGE Encyclopedia of Communication Research Methods*, vol. 2. SAGE Publications, Inc., Thousand Oaks, CA, pp. 587–588.
- Buitrago Flórez, F., Casallas, R., Hernandez, M., Reyes, A., Restrepo, S., Danies, G., 2017. Changing a generation's way of thinking: teaching computational thinking through programming. *Rev. Educ. Res.* 87 (4), 834–860.
- del Carmen Salazar, M., 2013. A humanizing pedagogy: reinventing the principles and practice of education as a journey toward liberation. *Rev. Res. Educ.* 37 (1), 121–148.
- Clark, H.F., Perez, L., Nava, I., Howard, J., 2022. Computational Thinking Integration in STEM Pedagogy by Teacher Candidates. 95th National Association of Research on Science Teaching Conference: Unity and Inclusion for Global Scientific Literacy, Vancouver, Canada.
- Collins, A., Brown, J.S., Newman, S.E., 1989. Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics. In: *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser*, vol. 18, pp. 32–42.
- CSforCA, 2021. Computer Science Skills: Computational Thinking Explained. CSforCA News & Updates. Available at: <https://csforca.org/computer-science-skills-computational-thinking-explained/>. (Accessed 10 April 2022).
- Cuny, J., Snyder, L., Wing, J.M., 2010. Demystifying Computational Thinking for Non-computer Scientist. Unpublished manuscript in progress referenced in: <http://www.cs.cmu.edu/~CompThink/resources/TheLinkWing.pdf>.
- Fisher, D., Frey, N., 2013. Gradual Release of Responsibility Instructional Framework. IRA E-essentials, pp. 1–8.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt, E., 2014. Preparing for Life in a Digital Age: The IEA International Computer and Information Literacy Study International Report. Springer Nature, New York.
- Gordon, E.M., Heck, D.J., 2019. 2018 NSSME+: Status of High School Computer Science. Horizon Research, Inc., Chapel Hill, NC.
- Gould, R., 2021. Toward data scientific thinking. *Teach. Stat.* 43 (S1), S11–S22. <https://doi.org/10.1111/test.12267>.
- Gould, R., Machado, S., Ong, C., Johnson, T., Molyneux, J., Nolen, S., Tangmunarunkit, H., Trusela, A., Zanontian, L., 2016. Teaching data science to secondary students: the mobilize introduction to data science curriculum. In: Engle, J. (Ed.), *Promoting Understanding of Statistics About Society*. Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE), Berlin, Germany.
- Hadad, R., 2020. The Use of Makerspaces for the Development of Computational Thinking Skills and Dispositions. PhD Thesis. University of Illinois at Chicago, Chicago, IL. <https://doi.org/10.25417/uic.13475724.v1>.
- IDSSP Curriculum Team, 2019. Curriculum Framework for Introductory Data Science. Available at: https://www.dssp.org/files/IDSSP_Frameworks_1.0.pdf. (Accessed 10 April 2022).
- Kafai, Y.B., 2016. Moving from computational thinking to computational participation. *Commun. ACM* 59 (8), 26–27.
- Kafai, Y.B., Proctor, C., Lui, D., 2019. From theory bias to theory dialogue Embracing cognitive, situated and critical framings of computational thinking in K-12 CS education. In: *Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER '19)*. Association for Computing Machinery, New York, NY, pp. 101–109.
- Kawasaki, J., Quartz, K.H., Martinez, J.F., 2020. Using multiple measures of teaching quality to strengthen teacher preparation. *Educ. Pol. Anal. Arch.* 28 (128). <https://doi.org/10.14507/epaa.28.5001>.
- Kessler, G., 2007. Formal and informal preparation and teacher attitude toward technology. *Comput. Assist. Lang. Learn.* 20 (2), 173–188.
- Koshy, S., Hinton, L., Cruz, L., Scott, A., Flapan, J., 2021. The California Computer Science Access Report. Available at: https://www.kaporcenter.org/wp-content/uploads/2021/09/KC21007_CSCA_Access_Report.pdf. (Accessed 10 April 2022).

- K–12 Computer Science Framework, 2016. K-12 Computer Science Framework. Available at. <http://www.k12cs.org>. (Accessed 10 April 2022).
- Lee, I., 2016. Reclaiming the roots of CT. In: CSTA Voice: The voice of K-12 Computer Science Education and Its Educators, vol. 12, pp. 3–4.
- Liu, J., Lin, C.H., Hasson, E.P., Barnett, Z.D., 2011. Introducing Computer Science to K-12 Through a Summer Computing Workshop for Teachers. In Proceedings of the 42nd ACM Technical Symposium on Computer Science Education, March, Dallas, TX.
- Mills, C., Ballantyne, J., 2016. Social justice and teacher education: a systematic review of empirical work in the field. *J. Teach. Educ.* 67 (4), 263–276.
- National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010. Common Core State Standards for Mathematics. National Governors Association Center for Best Practices & Council of Chief State School Officers, Washington, DC.
- National Research Council, 2011. Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics. The National Academies Press, Washington, DC.
- National Research Council, 2013. APPENDIX F: Science and Engineering Practices in the Next Generation Science Standards Next Generation Science Standards: For States, by States. The National Academies Press, Washington, DC.
- NGSS Lead States, 2013. Next Generation Science Standards: For States, by States, vol. 1. The National Academies Press, Washington, DC.
- Reiser, B., 2013. What Professional Development Strategies Are Needed for Successful Implementation of the Next Generation Science Standards. Available at. <https://www.ets.org/Media/Research/pdf/reiser.pdf>. (Accessed 10 April 2022).
- Robinson, N.J., 2019. How Inequality Statistics Can Mislead You. *Current Affairs*. Available at. <https://www.currentaffairs.org/2019/04/how-inequality-statistics-can-mislead-you>. (Accessed 10 April 2022).
- Ryoo, J.J., Estrada, C., Tanksley, T., Margolis, J., 2019. Connecting Computer Science Education to Students' Passions: A Critical Step Toward Supporting Equity in CS Education. University of California, Los Angeles, Center X. Computer Science Equity Project.
- Ryoo, J.J., Tanksley, T., Estrada, C., Margolis, J., 2020. Take space, make space: how students use computer science to disrupt and resist marginalization in schools. *Comput. Sci. Educ.* 30 (3), 337–361.
- Ryoo, J.J., Morris, A., Margolis, J., 2021. What happens to the Raspado man in a cash-free society? Teaching and learning socially responsible computing. *J. Educ. Resour. Comput.* 21 (4), 1–28.
- Sengupta, P., Kinnebrew, J.S., Basu, S., Biswas, G., Clark, D., 2013. Integrating computational thinking with K-12 science education using agent-based computation: a theoretical framework. *Educ. Inf. Technol.* 18 (2), 351–380.
- Sim, T.Y., Teow, M., Lau, S.L., 2021. Unplugged Computational Thinking Activities Framework Development for Novice Programmer. 2021 IEEE International Conference on Computing (ICOCO), pp. 297–302. <https://doi.org/10.1109/ICOCO53166.2021.9673537>.
- Sutcher, L., Darling-Hammond, L., Carver-Thomas, D., 2016. A Coming Crisis in Teaching? Teacher Supply, Demand, and Shortages in the US. Learning Policy Institute, Palo Alto, CA.
- Tucker, A., Deek, F., Jones, J., McCowan, D., Stephenson, C., Verno, A., 2003. A Model Curriculum for K-12 Computer Science: Final Report of the ACM K-12 Task Force Curriculum Committee. Computer Science Teachers Association, New York.
- U.S. Department of Education Office of Postsecondary Education, 2000. Teacher Shortage Area. U.S. Department of Education. Available at. <https://tsa.ed.gov/#/reports>. (Accessed 10 April 2022).
- Utts, J., 2015. The many facets of statistics education: 175 years of common themes. *Am. Statistician* 69 (2), 100–107.
- Weintrop, D., Beheshti, E., Horn, M.S., Orton, K., Jona, K., Trouille, L., Wilensky, U., 2015. Defining computational thinking for mathematics and science classrooms. *J. Sci. Educ. Technol.* 25 (1), 127–147.
- Wing, J.M., 2008. Computational thinking and thinking about computing. *Phil. Trans. Math. Phys. Eng. Sci.* 366 (1881), 3717–3725.

The scholarship of teaching and learning: cases from one STEM department

Emily K. Faulconer and John Griffith, Department of STEM Education, Worldwide Campus, Embry-Riddle Aeronautical University, Daytona Beach, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	321
Analysis of recent SoTL and DBER research artifacts in STEM	322
Science	322
Statistics	323
Cross-disciplinary	323
Integration of research	324
Conclusion	324
References	325

Glossary

Asynchronous learning A course modality where instruction and learning do not occur in the same geographic location or the same time.

Disciplinary-Based Education Research (DBER) Systematic, empirical investigation to test theories relevant to teaching and learning within a single discipline, informed by disciplinary expertise

Evidence-based The practical application of rigorously evaluated current research to a concept, strategy, or design

Humanistic STEM (H-STEM) A blending the study of science, technology, engineering, and mathematics with interest in, and concern for, human affairs, welfare, values, or culture

Scholarship of Teaching and Learning (SoTL) Systematic, applied research to test theories relevant to teaching and learning, applicable across disciplines

Introduction

High profile STEM organizations have become vocal in advocating for improvements in STEM teaching and learning (Jamieson and Lohmann, 2012; Olson and Riordan, 2012). While STEM course enrollments have dropped slightly in the last two years, online course enrollments saw significant growth in this time frame (National Center for Education Statistics, 2019, 2020). There has been increased attention to best practices not just in online learning, but specifically in teaching STEM disciplines online. While the event was ultimately canceled due to the COVID-19 pandemic, The American Chemical Society's Biennial Conference on Chemical Education had a symposia titled "Advances in Online Chemistry Education". The 2020 Electronic Conference on Teaching Statistics (hosted by the Consortium for the Advancement of Undergraduate Statistics Education) held a panel for engaging diverse audiences through online learning.

The scholarship of teaching and learning (SoTL) offers a route to improve STEM teaching and learning, with notable benefits to students, instructors, and institutions. SoTL is often considered to have originated in 1990 with Ernest Boyer's book *Scholarship Reconsidered: Priorities of the Professoriate*. This book argued for the value of four aspects of scholarship, including scholarship of discovery, scholarship of teaching, scholarship of integration, and scholarship of application. While some disciplines had previously valued educational research, this public call to value the scholarship of teaching in higher education was heard; SoTL has flourished in the decades since. Two widely cited sources for best practices in SoTL are outlined in Table 1. SoTL activities generate evidence-based models of practice for colleagues and establish high-quality education for students (Bernstein, 2013).

Table 1 Best practices in SoTL.

	Glassick et al. (1997)	Felten (2013)
Focus	Clear goals	Inquiry focused on student learning
Literature review	Adequate preparation	Grounded in context
Methodology	Appropriate methods	Methodologically sound
		Conducted in partnership with students
Data analysis	Significant results	
Discussion	Reflective critique	
Dissemination	Effective presentation	Appropriately public

SoTL adheres to scholarly work standards as other forms of scholarship, requiring disciplinary expertise, innovation, replication, peer review, and impact (Diamond, 2002). The SoTL research process is similar to research in STEM Fields. SoTL research is a systematic inquiry, proceeding from identifying a problem, establishing a theoretical framework, designing the study using appropriate data collection methods (qualitative, quantitative, or mixed methods), analyzing data rigorously using statistical analysis where appropriate, and disseminating the work. Ethics must be considered in SoTL work as most studies will involve human subjects, necessitating review by an Institutional Review Board. Because the key goal of SoTL is to transform teaching in higher education (Huber and Hutchings, 2005), accessible dissemination is a key step in the SoTL process. Dissemination options are growing, including teaching conferences and journals. Some journals offer open access publication without large article processing charges. As with dissemination in all disciplines, venues must be thoroughly evaluated for quality and impact.

Felten argues that SoTL is a signature pedagogy of educational development that enacts practices, values, and assumptions in a way that aligns purpose and practice (Felten and Chick, 2018). Despite the documentation of effective pedagogies and interventions in traditional STEM teaching and learning, the rate of adoption and scale-up has been limited (Borrego and Henderson, 2014). It is reasonable to assume that online STEM faces the same limitations.

Despite the benefits of SoTL, some institutions have experienced the SoTL paradox, where the increased recognized value of SoTL work has not been matched with recognition for SoTL efforts in formal reward systems like promotion and tenure (Walker et al., 2008). Disciplinary-based education research (DBER) has also seen growth, and mismatched recognition within STEM (Dolan et al., 2018). This is in spite of many faculty engaging in SoTL and DBER research. According to one study that included STEM and non-STEM disciplines, nearly half of faculty promotions at a research university included at least one SoTL artifact (Marcketti and Freeman, 2016).

Solutions to the SoTL paradox are likely unique to the goals of an institution and, even within an institution, aspects of faculty promotion and tenure like load balance can vary (Cruz et al., 2009). Some institutions permit unequal value weighting of SoTL scholarship compared to scholarship of discovery for promotion and tenure decisions (Cruz et al., 2009). At one institution, the Engineering & Technology Department supported equal valuation both pre- and post-tenure while GeoSciences and Natural Resources Department supported unequal valuation pre-tenure with equal valuation post-tenure (Cruz et al., 2009). While some discussion of support for SoTL for online STEM faculty exists (Herman and Butler, 2019), at this time, little information is publicly available about the support for SoTL research (and the SoTL paradox) for tenured or tenure-track STEM faculty teaching fully online.

At Embry-Riddle's Worldwide Campus, the College of Arts and Sciences took proactive measures to address the SoTL paradox. Previously, the Math, Physical, and Life Sciences Department was comprised of faculty teaching STEM general education courses in mathematics, chemistry, life sciences, physics, computer science, and statistics, as well as applied sciences courses. While many of the faculty in the department were non-tenure track, those faculty on the tenure-track were engaged in SoTL and DBER research. To honor the research and teaching interests of faculty, the department was divided in 2019 to form the Department of STEM Education and the Department of Applied Sciences. Tenure line faculty in the Department of STEM Education could focus more on educational research in STEM and those in the Department of Applied Sciences would focus on disciplinary research. In 2019, Promotion and Tenure (P&T) policy documents were prepared and approved that support STEM-focused SoTL and DBER activities by permitting faculty to publish scholarship in professional spaces "as appropriate for the research focus defined by the candidate". Furthermore, this new P&T policy does not suggest unequal weighting of these efforts with STEM disciplinary works; faculty autonomy is supported, stating "no individual pursuit of research is the same as another". Further, according to this policy, publication venues must be refereed and demonstrate selectivity, prestige, and impact, as evidenced through bibliometrics. Many SoTL and DBER publication venues meet these criteria. The following section will present an analysis of various research artifacts generated by tenure-track faculty within the Department of STEM Education.

Analysis of recent SoTL and DBER research artifacts in STEM

Science

In Faulconer et al. (2019), two Department of STEM Education faculty collaborated to explore the influence of combining two practices—providing multiple attempts on summative assessments and providing high-quality immediate feedback—to explore the impact on student behaviors and performance. While this study focused on a Chemistry course, the concept is transferable to additional STEM disciplines. Therefore, the authors published in a SoTL journal rather than a DBER journal.

This paper meets Felten's good practices in SoTL (Felten, 2013). The work is an inquiry focused on student learning, with the specific goal of improving performance on summative assessments. The work is grounded in context by building on existing, relevant theory and prior research on multiple attempts and feedback. The work is methodologically sound through its use of a quasi-experimental design. The work is conducted in partnership with students in several ways. The study was reviewed by the institutional review board. An undergraduate student co-authored the paper. Additionally, the results have been communicated to students to encourage use of the multiple attempts. Finally, the conclusions of this work are appropriately public, published in a quality journal (2019 Metrics: Impact Factor 2.136, CiteScore 3.7, SNIP 1.985) managed by the reputable publisher Taylor & Francis.

The researchers used the learning management system to provide timely, actionable, specific feedback to students on their online quizzes. The option to re-take the quiz, pulling questions randomly from pools aligned with the learning objectives, turned the feedback into feedforward by allowing students to apply the knowledge gained. Each quiz attempt was unique, students earned their highest score from the multiple attempts, and there was no penalty for failure to use multiple attempts. They reported the

following conclusions: (1) students who did not initially earn an A were more likely to try again, (2) students tended to perform better on their second attempt in both chemistry lecture and laboratory courses (Faulconer et al., 2019). Interestingly, students who used two attempts did not outperform those who only used one attempt (Faulconer et al., 2019). It can be inferred, then, that the multiple attempts format allows students to close the performance gap. This is, however, a notable time investment for students, with those who used multiple attempts spending approximately twice as long on the assessment task as those who only used one attempt (Faulconer et al., 2019). In light of the outcome that time on task alone was not a good predictor of assessment performance (Faulconer et al., 2019), it further supports the notion that this pedagogical choice supports closure of a performance gap for students.

The impact of this research can be evaluated through several means, though traditional means such as the number of citations may not be appropriate (Dolan et al., 2018). The ultimate goal of SoTL and DBER is to improve student learning. Reach to practitioners is key, which cannot be evaluated by a citation count. For example, the most-downloaded article from the *Journal of College Science Teaching* has been accessed 11,682 times but, according to Google Scholar, this work has garnered 1330 citations. As of September 1, 2020, the Faulconer et al. (2019) article has an Altmetric of 2, with 210 views reported by the journal and 73 downloads reported by Scholarly Commons.

Statistics

In Bourdeau et al. (2017), one Department of STEM Education faculty member collaborated with colleagues within another department in the College of Arts and Sciences to explore the influence of targeted STEM educational videos on student understanding of key concepts (Bourdeau et al., 2017). While this pilot study took place in an introductory research course, many of the difficult topics targeted for video resource development were focused on experimental design and data analysis concepts related to statistical concepts. Specifically, videos were developed on the topics of categorical versus numerical variables, selecting an appropriate statistical test, p-values, and confidence intervals. While this work was largely research and statistics-focused, the idea is theoretically transferable to other disciplines. Therefore, this work was published in a SoTL journal rather than a DBER journal. Unfortunately, the generalizability of this study was limited due to a low response rate. The authors point to the potential stumbling blocks for implementing these types of interventions. For example, the non-mandatory nature of the videos may have reduced student access and perceived utility of the resource.

Felten's five good practices in SoTL are evident in this publication. This study directly focused on improving student learning by providing additional multimedia resources to support the most-difficult content topics in the course. This work is grounded in context, built upon previous research that demonstrates efficacy of multimedia use in the classroom. The authors ground their work in the theory of self-directed learning. Design-based research is a valid method for performing SoTL research. The student perspectives were measured using voluntary surveys. While this study secured institutional approval to protect human subjects, it goes beyond this minimum expectation by including students as partners in the inquiry through collection of their perspectives of the multimedia intervention. Finally, this work is appropriately public, published in an open-access journal indexed by ERIC and the Directory of Open Access Journals. While the journal is not listed in Scopus or Cabell's, the journal is published by the Eurasian Society of Educational Research.

While this paper has only garnered one citation, as of September 24, 2020, the article has been viewed by 307 people and downloaded by another 478 people, reported by the journal. These statistics demonstrate a reach to practitioners, which was the goal of the authors.

Cross-disciplinary

In Faulconer et al. (2021), two Department of STEM Education faculty collaborated with a colleague within another department as well as a colleague external to the institution to explore the use of course video trailers to contextualize the general education competencies, explaining how general the general education course was valuable to students in their academic and professional lives. This study included courses across multiple disciplines, including two mathematics courses, three science courses, an introductory research methods course, and an English composition course. The videos produced followed best practices, including use of a singular message focused on context, short duration (one to 4 min), and high production value (Goudsouzian, 2018; Truell, 2018). The cross-disciplinary nature of this work justified publication in a SoTL journal.

This paper meets Felten's good practices in SoTL. While this study did not focus specifically on disciplinary knowledge or skill development, it did explore student attitudes that connect to learning, which meets Felten's definition of inquiry focused on student learning (Felten, 2013). The work was grounded in relevant theory regarding the influence of course first impressions and contextualization on student attitudes, persistence, and performance (Wong, 2016; Wachen et al., 2011). The study was methodologically sound, using survey research to explore student opinions and attitudes regarding the course trailers. This research aligned with basic tenets of human subject research and was reviewed by the Institutional Review Board. Beyond that, the nature of the research methodology inherently engaged students in the inquiry process. This work is appropriately public, published in a quality journal published by Penn State University Press since 1999 (2019 Metrics: Impact Factor 0.23, CiteScore 0.4, SNIP 0.148).

Undergraduate students often misunderstand the nature and value of general education courses for their academic and professional growth. In this work, the researchers developed video course trailers that were designed specifically to contextualize the course in terms of a specific institutional general education competencies (i.e., transferable skills) developed throughout the course. Through the

videos, the researchers aimed to improve contextualization through four specific means: posing challenges and questions, reference to additional resources, redundant iterations of usefulness, and application of cited best practices in course trailer development. The influence of the videos was measured by comparing the change in student perceptions of the fit of the general education course into their academic and professional plans once videos were placed into the learning management system for each course. The study reported that, while a strong understanding of the fit of the general education course in both their academic and professional plans existed prior to implementation of course video trailers, there was a statistically significant improvement in this understanding through implementation of video course trailers (Faulconer et al., 2021). It is important to note that the predominance of physics students in the data may have skewed the results due to the aeronautical orientation of the university studied. However, taken in aggregate terms, the results seemed to support using videos to reinforce course value to students' personal and professional lives. Having just been published, the impact of this work is yet to be seen. Based upon the bibliometrics for the journal, notable citations of this work are not expected. While the readership for the journal is not publicly available, the journal is indexed in IBZ, MLA International Bibliography, and SCOPUS, making it available through certain institutional library collections.

Integration of research

The SoTL research led by faculty from ERAU Worldwide's Department of STEM Education has aided in course assessments and improvement of teaching practices. The concept of multiple attempts with feedforward has been implemented in six online STEM courses, including chemistry, environmental science, biology, and an interdisciplinary science course. Research that demonstrates applicability across multiple STEM disciplines is currently under review. The video course trailer study generated script templates and a best practices list that has been used to generate and evaluate videos in STEM and non-STEM courses in three other departments.

Demonstrating integration of research with service, a 2020 study led by the Department of STEM Education faculty reported on a survey of safety culture in academic laboratories (Faulconer et al., 2020). This lead author also serves as the Chair of the institution's Academic Safety Committee. This study helped support administrative actions to improve safety training, ultimately resulting in the development and launch of a fully online, asynchronous safety training for both students and faculty. This study also led to the subsequent publication of tandem editorials written in collaboration with faculty in another department on the topic of supporting safety culture in academia, including safety communication barriers (Faulconer and LeNoble, 2020) and giving a voice to faculty (LeNoble and Faulconer, 2020).

Another example of high-level integration started with two studies, with the first being a comparison of interests in and barriers to undergraduate research between distance and in-person students (Faulconer et al., 2020e) and the follow-on study comparing the same in online and residential faculty (Faulconer et al., 2020d). These studies laid the foundation for a successful funding proposal through the National Science Foundation's Improving Undergraduate Stem Education: Education and Human Resources program (NSF 19-601), which supports novel, creative, and transformative approaches to knowledge creation in undergraduate STEM teaching and learning. This project, launched in October 2020, will establish key supports for undergraduate research in a fully online engineering technology bachelor degree program with the goal of improving degree retention and persistence as well as improving STEM identity and attitudes. Supports include long-term research mentorship, synchronous and asynchronous online research skills workshops, online faculty professional development for research mentorship and supervision, and the development of a structured 9-week research-focused independent study course. This project supports teaching and service activities for multiple faculty members and will generate additional research scholarship.

Faculty within the department have also collaborated across departments to develop critical emphasis on humanistic STEM (H-STEM). A modern approach is humanistic STEM, defined as "a path blending the study of science, technology, engineering, and mathematics with interest in, and concern for, human affairs, welfare, values, or culture" (Bourdeau and Wood, 2019). Early work in this area explored the influence of infusing humanities into online STEM courses on student perceptions of connections between the course content and cross-disciplinary perspectives (Faulconer et al., 2020c). This effort led to efforts to infuse STEM into humanities courses, including efforts in a History of Aviation course. Faculty received funding from the Synergistic Undergraduate Mathematics via Multi-Institutional Interdisciplinary Teaching Partnerships (SUMMIT-P) grant to develop two math courses aligned with H-STEM practices, Learning to Reason: Art and Quotient and Learning to Reason: Commerce and Flux. The courses are fully online asynchronous courses taught over 9 weeks. Both courses were developed by a multidisciplinary team. Planned research will collect data from these courses to provide evidence to support H-STEM as a valid approach in teaching and learning. Departmental faculty are also seeking external funding to develop an H-STEM minor, including courses like Digital Humanities as well as Data Visualization and Storytelling with Data. Furthermore, the funding proposal also seeks to establish an H-STEM center to support a consortium of institutions engaged in H-STEM research.

Conclusion

While there has been a long period showing a positive growth trend in online education, this was significantly increased in 2020. We cannot assume faculty expertise in online pedagogy, learning management system use, and online curriculum design. There has been long-term resistance to offering STEM courses online. The provision of evidence-based models of excellence can support widespread adoption of the online modality for STEM courses among administrators and educators.

However, some STEM faculty may experience a SoTL paradox, where they do not have full academic freedom to pursue research lines in this area. Without administrative support through actions and policies, faculty may not see a good return on investment for their time on SoTL scholarly efforts. These research efforts offer benefits for students, faculty, and institutions. As evidenced from the research analyzed in this article, the SoTL research performed by fully online STEM faculty within the Department of STEM Education at Embry-Riddle's Worldwide Campus meets best practices and has implications beyond dissemination. The works supported improved teaching by providing evidence-based models of excellence in instructional design and teaching practices. Furthermore, SoTL research by STEM faculty supported high-level service activities for multiple STEM and non-STEM faculty members. As important as discipline-specific research is to the growth in the body of knowledge, SoTL research will help ensure that knowledge is effectively passed to the students we serve.

References

- Bernstein, D., 2013. How SoTL-active faculty members can be cosmopolitan assets to an institution. *Teach. Learn. Inq.* 1 (1), 35–40.
- Borrego, M., Henderson, C., 2014. Increasing the use of evidence-based teaching in STEM higher education: a comparison of eight change strategies. *J. Eng. Educ.* 103 (2), 220–252.
- Bourdeau, D.T., Wood, B.L., 2019. What is humanistic STEM and why do we need it? *J. Humanist. Math.* 9 (1), 205–216.
- Bourdeau, D.T., Roberts, D.L., Wood, B.L., 2017. A study of video-mediated opportunities for self-directed learning in required core curriculum. *Int. J. Educ. Method.* 3 (2), 85–91.
- Cruz, L., Ellem, J., Ford, G., Moss, H., White, B.J., 2009. Recognition and reward: SoTL and the tenure process at a regional comprehensive university. *MountainRise* 5 (3), 1–27.
- Diamond, R.M., 2002. Defining scholarship for the twenty-first century. *N. Dir. Teach. Learn.* 2002 (90).
- Dolan, E.L., Elliott, S.L., Henderson, C., Curran-Everett, D., John, K.St, Ortiz, P.A., 2018. Evaluating discipline-based education research for promotion and tenure. *Innovat. High. Educ.* 43, 31–39.
- Faulconer, E.K., Lenoble, C., 2020. Supporting safety culture in academia: safety communication barriers. *Department Chair* 30 (4), 15–16.
- Faulconer, E.K., Griffith, J.C., Frank, H., 2019. If at first you do not succeed: student behavior when provided feedforward with multiple trials for online summative assessments. *Teach. High. Educ.*
- Faulconer, E., Dixon, Z., Griffith, J.C., Frank, H., 2020. Surveying the safety culture of academic laboratories. *J. Coll. Sci. Teach.* 50 (2), 18–26.
- Faulconer, E., Griffith, J., Faulconer, L., Dixon, Z., 2021. A course in context: video course trailers. *J. Gen. Educ.* 69 (1–2), 60–85. <https://doi.org/10.5325/jgeneeduc.69.1-2.0060>.
- Faulconer, E., Wood, B., Griffith, J., 2020c. Infusing humanities in STEM education: student opinions of disciplinary connections in an introductory chemistry course. *J. Sci. Educ. Technol.* 29 (3), 340–345.
- Faulconer, E.K., Dixon, Z., Griffith, J., Faulconer, L., 2020d. Perspectives on Undergraduate Research Mentorship: a comparative analysis between online and traditional faculty. *Online J. Dist. Learn. Adm.* 23 (2).
- Faulconer, E.K., Griffith, J.C., Dixon, Z., Roberts, D., 2020e. Comparing online and traditional student engagement and perceptions of undergraduate research. *Scholarsh. Pract. Undergrad. Res.* 3 (3), 48–59.
- Felten, P., Chick, N., 2018. Is SoTL a signature pedagogy of educational development? *Improve Acad.* 37 (1), 4–16.
- Felten, P., 2013. Principles of good practice in SoTL. *Teach. Learn. Inq.* 1 (1), 121.
- Glassick, C.E., Huber, M.T., Maeroff, G.I., 1997. Scholarship assessed: evaluation of the professoriate. Special Report. Jossey Bass Inc., San Francisco, CA.
- Goudsouzian, L.K., 2018. Hollywood-style movie trailers increase student interest. *J. Microbiol. Biol. Educ.* 19 (1).
- Herman, C., Butler, D., 2019. Innovations in STEM distance education. *Open Learn.* 34 (1), 1–5.
- Huber, M.T., Hutchings, P., 2005. *The Advancement of Learning: Building the Teaching Commons*. Jossey-Bass, San Francisco, CA.
- Jamieson, L.H., Lohmann, J.R., 2012. *Innovation with Impact: Creating a Culture for Scholarly and Systematic Innovation in Engineering Education*. American Society for Engineering Education, Washington, DC.
- Lenoble, C., Faulconer, E.K., 2020. Supporting safety culture in academia; giving a voice to faculty. *Department Chair* 31 (1), 6–7.
- Marcketti, S.B., Freeman, S.A., 2016. SoTL evidence on promotion and tenure vitas at a research university. *J. Scholarsh. Teach. Learn.* 16 (5), 19–31.
- National Center For Education Statistics, 2019. *The Condition of Education 2019*. U.S. Department of Education, Washington, DC.
- National Center For Education Statistics, 2020. *Term Enrollment Estimates: Spring 2020*.
- Olson, S., Riordan, D.G., 2012. *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics*. President's Council of Advisors on Science and Technology, Washington, DC.
- Truell, M., 2018. (last update) Best Practices: Creating Video Course Trailers [Homepage of Duke University]. Available: <https://admin.trinity.duke.edu/communications/best-practices-creating-video-course-trailers>. (Accessed 8 May 2018).
- Wachen, J., Jenkins, D., Van Noy, M., 2011. Integrating basic skills and career-technical instruction: findings from a field study of Washington state's I-BEST model. *Community Coll. Rev.* 39, 136–159.
- Walker, J.D., Boepher, P., Cohen, B., 2008. The scholarship of teaching and learning paradox: results without rewards. *Coll. Teach.* 56 (3), 183–189.
- Wong, B.T.M., 2016. Factors leading to effective teaching of MOOCs. *Asian Assoc. Open Univ. J.* 11 (1), 105–118.

Universal design in STEM education

Sheryl E. Burgstahler, University of Washington, Seattle, WA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	326
The accommodation approach to access	327
The universal design approach to access	328
The UDE framework	329
UDE guidelines and practices in STEM education	331
Conclusion	332
Acknowledgments	332
References	332
Relevant websites	333

Introduction

The design of any product or environment involves myriad considerations—research basis, purpose, safety and other standards, flexibility, sustainability, and cost, to name a few. Many current design practices focus on the average user. In contrast, UD sets the goal to design products and environments suitable for individuals diverse with respect to gender identity, race, ethnicity, age, socioeconomic status, ability, learning style, and other characteristics. Originally applied to architecture and consumer product design, UD has more recently emerged as an approach for addressing diversity, equity, and inclusion in the design of IT as well as a wide range of on-site and online, formal and informal learning offerings at all educational levels.

Although student bodies at all levels have become increasingly diverse with respect to race, ethnicity, gender identity, ability, and other characteristics, significant gaps exist in the achievement of some postsecondary degrees that include those in STEM for marginalized groups such as students with disabilities compared with their peers without these identities ([National Science Foundation, 2021](#)). However, success stories demonstrate that opportunities do exist for students who successfully overcome barriers imposed by inaccessible facilities, curricula, websites, digital files, instructional practices, and other aspects of educational offerings ([DO-IT, 1993–2021](#); [Stern and Woods, 2001](#)). Factors that impact retention of a diverse STEM student body occur at both the individual (e.g., socioeconomic status, study skills) and institutional (e.g., institutional policies, instructional practices) levels ([Allen and Anderson, 2020](#); [Berge and Huang, 2004](#); [Braxton et al., 1997](#); [Tinto, 1993](#)). The author of this chapter addresses institutional issues by focusing on the design of formal and informal STEM educational offerings that are accessible to and inclusive of students with diverse characteristics, including disabilities.

The discussion below elucidates some of the challenges faced by STEM students with various levels of physical, sensory, learning, attention, and communication abilities and shares examples of ways an instructor or institution can proactively ameliorate potential barriers to these students.

- *Visual differences.* Web content with certain color combinations presents a barrier to access for people who are color-blind; webmasters can avoid this by ensuring that content and navigation do not require the distinction of colors alone. In a similar vein, since screen reader technology can only read aloud text, individuals who are blind cannot make sense of content within images; by including alternative text descriptions for these images, webmasters can bypass this barrier. Individuals who have low vision and use software to enlarge screen images can view only a small portion of content at one time; this problem is minimized when web pages are uncluttered and use large text and plain backgrounds.
- *Hearing differences.* Audio content is unavailable to students who are deaf; this barrier is avoided when videoconferencing, videos, and video clips are captioned or transcribed and sign language interpreters and real-time captions are provided.
- *Physical differences.* Some STEM labs and other facilities are not comfortable for students who are larger than average; providing a few larger chairs can solve this problem. Lab stations that are not wheelchair-accessible cannot be used by some students and instructors; designing workspaces with adjustable height surfaces is one way to address this problem. Some individuals use alternative input devices that emulate the functions of a standard keyboard but not a mouse; instructors can avoid erecting barriers by using websites and software applications that can be operated with the keyboard alone.
- *Learning differences.* Course content delivered in a single mode, such as by lecture, creates challenges to access for students with diverse learning styles, strengths, and preferences; offering learning opportunities in multiple ways (e.g., orally, in printed materials, in video presentations, in discussions) can address this access issues. Online courses that change formats quickly or crowd too much content on the screen can create confusion or overwhelm students who are neurodiverse (e.g., students on the autism spectrum, with some types of learning disabilities or attention deficits); this problem is resolved when instructors present content in consistent formats and patterns and on uncluttered pages. When time provided to complete an assignment is inadequate, students with learning disabilities that affect processing speed are placed at a disadvantage; presenting assignments

early allows all students to demonstrate their knowledge. Some students with learning disabilities that affect their ability to interpret written text are at a disadvantage when written materials are the only option for gaining content; offering these readings in electronic formats such as HTML or accessible PDF so that it can be read aloud with text-to-speech software benefits these students.

- *Attention differences.* Most classes include students with a wide range in attention spans; some may find it difficult to pay attention to a lecture or engage in a lab. Instructors can avoid erecting barriers to these students by providing printed instructions or an easy-to-follow lecture guide.
- *Communication differences.* Some communication differences are related to autism, language proficiency, and abilities to speak or hear. Communication methods that do not require the ability to hear or speak and allow for differences in time needed to compose a thought (e.g., Email, engagement in an online discussion forum) allow all participants to engage.

Addressing the role that the designs of physical spaces, courses, technology, and other products play in erecting barriers to some people supports social models of disability—e.g., with respect to gender identity, socioeconomic status, race, ethnicity, body size, and ability—as simply a normal state of the human experience. Some consider that “impairment is an individual limitation, while a disability is a socially imposed restriction. Not being able to walk is an impairment but lack of mobility is a disability (a socially created situation)” (Seale, 2013, p. 23). Not being able to see is an impairment, but not having access to website content using a screen reader is the result of poor web design. A wheelchair user unable to reach the equipment in a makerspace is a result of the inflexible design of the furniture and layout.

The accommodation approach to access

Approaches for providing access to education for individuals with disabilities have evolved over time. In the past some individuals with disabilities were excluded from educational opportunities or learned in environments separated from those for other students. In the US, increased demands for the inclusion of otherwise qualified students with disabilities in all learning opportunities came from injured veterans returning from World War II who were eligible for college funding through the GI Bill, parents who wanted more academic options for their children, and those inspired by the civil rights movements for women and minorities to extend similar rights to people with disabilities. These efforts led to mandates that protect the rights of children with disabilities. The Individuals with Disabilities Education Act (IDEA, 2004) ensures that all children with disabilities are entitled to a free appropriate public education to meet their unique needs and prepare them for further education, employment, and independent living. These efforts also led to more general civil rights protections that apply to all students with disabilities in formal and informal educational opportunities at all academic levels—Section 504 of the Rehabilitation Act of 1973 as well as the Americans with Disabilities Act of 1990 (ADA) and its 2008 Amendments (ADAA), which considers a person with a disability to be any individual who

- has a physical or mental impairment that substantially limits one or more major life activities of such individual;
- has a record of such an impairment; or
- is regarded as having such an impairment (ADAA, Section 12102).

According to the ADA, no otherwise qualified individual with a disability shall, solely by reason of that disability, be excluded from, denied the benefits of, or subjected to discrimination under educational programs and other covered entities. *Qualified* with respect to education available to the public means an individual with a disability who, with or without reasonable modifications to rules, policies, or practices, the removal of architectural, communication, or transportation barriers, or the provision of auxiliary aids and services, meets the essential eligibility requirements for the receipt of services or the participation in activities (ADA, Section. 12101). Similar laws that protect the rights of people with disabilities have been enacted in countries around the world.

Using an accommodation approach for addressing the inaccessible design of program offerings is deeply rooted in the culture of most educational institutions. An *accommodation* occurs when an adjustment or modification is made to a product or environment so that it is accessible to an individual with a disability. A focus on accommodations is grounded in a medical model (or deficit model) of disability, in which a professional identifies an individual’s functional limitations and prescribes a cure, rehabilitation, or adjustments that allow this person to fit into an established program (Jones, 1996; Loewen and Pollard, 2010). In K-12 education, parents, teachers, and counselors work together to ensure the student has an appropriate education; in higher education, typically the student must provide documentation of a disability to a specified office at the institution and request accommodations. The institutional representative approves accommodations considered reasonable and, with the permission of the student, shares this information with appropriate faculty and staff who implement the accommodations. Examples of accommodations include materials in alternate formats, extra time on assignments and exams, sign language interpreters and real-time captioners, and movement of classes to wheelchair-accessible locations.

It is estimated that most adult students potentially eligible for accommodations choose not to seek accommodations; similarly, often parents of students with disabilities in informal learning programs do not request accommodations. When asked why, some individuals with disabilities report embarrassment, concerns about discrimination, and cumbersome request processes. Negative aspects that occur when an accommodation-only approach is applied include the following:

- The process for securing accommodations marginalizes students with disabilities by requiring a segregated process for gaining access.
- The value associated with an accommodation does not extend to students with disabilities who choose not to self-disclose nor to other students in a class who might benefit from it.
- An accommodation does not always result in content and experiences equivalent to those of other students.
- An accommodation for one student does not in and of itself make an activity more accessible to other students.

The universal design approach to access

A variety of terminology—useable design, accessible design, inclusive design, design for all, barrier-free design, user-centered design, human-centered design, sustainable design, universal access, universal design—has been used to label proactive efforts that have been made to proactively design products and environments that avoid the creation of barriers to some people. They have in common a paradigm shift in design approaches from a focus on the average or deemed typical user to consideration of a broad range of human characteristics (Myers et al., 2013). Accessible design practices often limit their scope to access issues faced by people with disabilities. Useable design takes into account how easily users can learn how to operate a product, achieve their goals, and remember how to perform tasks when they return to the product at a later time. Not all accessible products are useable. For example, applications software could offer keyboard options in such a way that a person who is blind could in theory navigate it with a screen reader, but, make the product unusable because the options are too confusing for screen reader users. Inclusive design implies that a single product is appropriate for all potential users. However, designs that are technically accessible and useable may not be inclusive. For example, an entrance to a building that has stairs and a separate ramp for wheelchair users is accessible and useable but is not inclusive because all individuals approaching the building do not enter in the same way. A more inclusive design is a sloping ramp to the entrance that allows people of all abilities to approach and enter the building together.

Universal design (UD)—arguably the most comprehensive of all proactive approaches (Seale et al., 2020)—creates products that are accessible, useable, and inclusive, as illustrated in Fig. 1.

UD has a long and rich history. Marc Harrison (1928–1996), a professor of industrial engineering at the Rhode Island School of Design, was a pioneer in what was later called UD. Insight and inspiration for Harrison’s work came from engaging in years of rehabilitation after sustaining a traumatic brain injury as a child. He, like universal designers today, challenged design practices that focus on individuals of average ability and promoted the notion that products and environments should take into account the diversity of abilities in human beings. One of his many designs was a food processor with large and easily pressed buttons, large handles, and large, high-contrast labels that made it useable by more people (Hagley Museum and Library, n.d.). In the late 1970s, Ronald Mace—a wheelchair user, architect, product designer, and educator—coined the term universal design. At the Center for Universal Design (CUD) at North Carolina State University, he defined UD as “the design of products and environments to be useable by all people, to the greatest extent possible, without the need for adaptation or specialized design” (CUD, n.d.). He and his colleagues demonstrated how UD features could be integrated into a product or environment in such a way that they do not stand out and thus foster the social integration of people with disabilities and other marginalized groups. An example of UD is a sidewalk that has curb cuts to make it useable by people who are walking, using wheelchairs and walkers, pushing baby strollers, and rolling delivery carts. The common thread of all UD applications is that a diverse group of users can fully benefit from a product or environment in an inclusive setting (National Council on Disability, 2003).

Similarly, when instructors apply UD to a postsecondary course, they consider the great diversity of potential students. Considerations include level of ability to move, see, hear, read, learn, and process information; stature; age; race; ethnicity; culture;

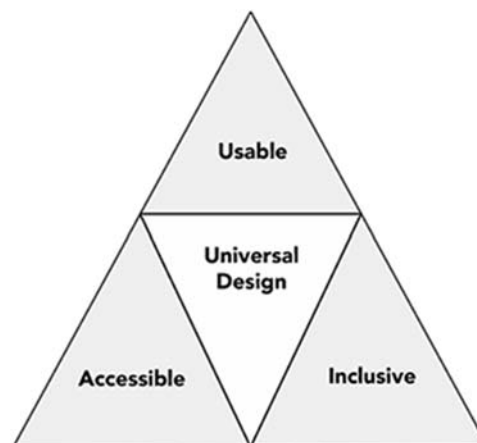


Fig. 1 Characteristics of a UD strategy: accessible, useable, and inclusive. Source: Burgstahler (2015, p. 15).

socioeconomic status; learning style and preference; dexterity; native language; intelligence; and gender. Considering that most characteristics (e.g., the ability to hear) can be measured on a continuum, the traditional view of a person “having a disability” or “not having a disability” is overly simplistic. Everyone has a unique combination of abilities that can change over the course of a lifetime.

In contrast to the limited accommodation approach to access, the UDE Framework

- is comprehensive enough to address issues of diversity, equality, accessibility, social integration, and community;
- goes beyond the design of learning activities to provide guidance for making physical spaces, services, and technology that make education more welcoming to, accessible to, and useable by all students;
- can be applied incrementally; and
- can be used to guide practices in specific academic areas such as those in STEM.

The UDE framework

A framework for the scope of any UD application area was developed by the author of this article (Burgstahler, 2020)—it includes scope, definition, principles, guidelines, practices, and process, as visualized in Fig. 2. The Framework can be fleshed out for any and all UDE applications. A scope could be narrowly focused (e.g., distance learning in STEM) or broadly focused (e.g., all design practices for teaching and learning, services, technologies, and physical spaces) design practices at a postsecondary institution) or more narrowly focused (e.g., courses in a specific academic field). The following paragraphs flesh out the UDE Framework for the scope of applications in this chapter, all UDE applications relevant to STEM, including on-site, online, and hybrid; formal and informal; and online and on-site offerings.

After the basic definition of UD was established, in the mid-90s, with funding from the U.S. Department of Education, a group of product developers, architects, environmental designers, and engineers at CUD established a set of seven principles of UD and related guidelines to consider in the design of any product or environment (CUD, n.d.).

- *Equitable use.* The design is useful and marketable to people with diverse abilities.
- *Flexibility in use.* The design accommodates a wide range of individual preferences and abilities.
- *Simple and intuitive use.* Use of the design is easy to understand, regardless of the user’s experience, knowledge, language skills, or current concentration level.
- *Perceptible information.* The design communicates necessary information effectively to the user, regardless of ambient conditions or the user’s sensory abilities.
- *Tolerance for error.* The design minimizes hazards and the adverse consequences of accidental or unintended actions. Discourage unconscious action in tasks that require vigilance.
- *Low physical effort.* The design can be used efficiently, comfortably, and with a minimum of fatigue.
- *Size and space for approach and use.* Appropriate size and space are provided for approach, reach, manipulation, and use regardless of the user’s body size, posture, or mobility (Story et al., 1998, pp. 34–35).

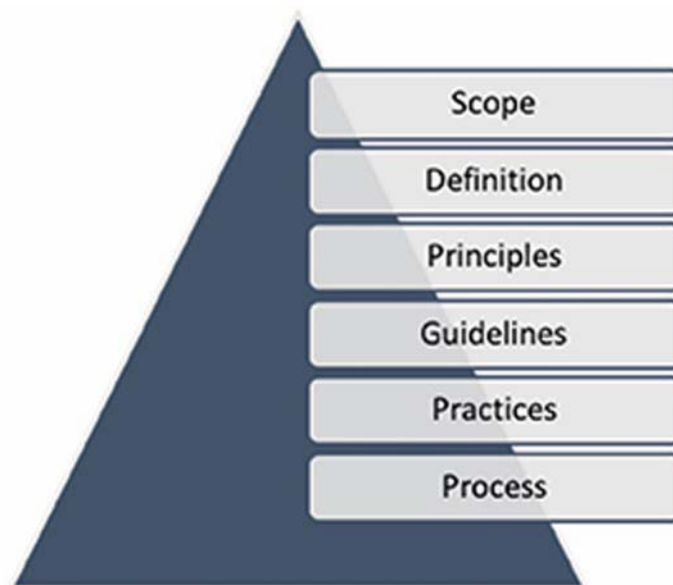


Fig. 2 Components of a UDE framework. Source: Burgstahler (2020, p. 36).

These general principles and guidelines can apply to any design discipline (Story et al., 1998, p. 32). From this foundation, the authors also expected that practitioners would develop design standards and compliance tests for specific fields. For example, principle 5, “Tolerance for error,” can be applied to a formal or informal online course—even though offering courses online was a rare practice when the principle was written—by providing helpful feedback that redirects a student who makes a specific selection error while engaging with the curriculum.

Some researchers and practitioners have used the seven principles of UD as a foundation on which to build additional principles, guidelines, and lists of promising practices to help practitioners implement UD in specific situations. With respect to IT, Vanderheiden and Vanderheiden (1992) developed specific guidelines for information technology. Later, IT accessibility leaders organized by the World Wide Web Consortium worked together to develop the Web Content Accessibility Guidelines (WCAG). These Guidelines dictate that all information and user interface components follow four guiding principles:

- *Perceivable*: Users must be able to perceive the content, regardless of the device or configuration they’re using.
- *Operable*: Users must be able to operate the controls, buttons, sliders, menus, etc., regardless of the device they’re using.
- *Understandable*: Users must be able to understand the content and interface.
- *Robust*: Content must be coded in compliance with relevant coding standards in order to ensure its accurately and meaningfully interpreted by devices, browsers, and assistive technologies (World Wide Web Consortium, 2018).

Although developed for web design, the WCAG principles, guidelines, and success criteria can also be applied to curriculum and materials, software, and other technologies (World Wide Web Consortium, 2013). Accessible practices that emerge from WCAG include providing alternative text to describe the content in images, captioning and audio describing videos, using headings at appropriate levels within documents and websites, using link text that describes its destination, and ensuring that the product can be operated with a keyboard alone.

Many UD-inspired approaches have emerged to specifically address instructional applications that include formal and informal learning (Burgstahler, 2015; Roberts et al., 2015). Each rests upon a common finding in educational research—learners are diverse with respect to their abilities, learning preferences, and responses to instruction. The most common UD-inspired principles applied to the design of curriculum and instruction are those developed by the Center for Applied Special Technology (CAST, n.d.). Applying Universal Design for Learning (UDL) principles requires that instructors offer students multiple means of engagement, representation, and action and expression throughout their learning experiences, as defined below.

- *Engagement*: For purposeful, motivated learners, stimulate interest and motivation for learning.
- *Representation*: For resourceful, knowledgeable learners, present information and content in different ways.
- *Action and expression*: For strategic, goal-directed learners, differentiate the ways that students can express what they know (Center for Applied Special Technology, 2018).

The combination of principles that underpin UD, WCAG, UDL can be applied to the design of all physical, technological, and pedagogical aspects of formal and informal, hybrid and fully online, and synchronous and asynchronous learning (University of Washington, n.d. b). Employing practices supported by this combination of principles for online learning technology and pedagogy ensure that

- students have multiple ways to gain knowledge, demonstrate what they have learned, and interact; and
- each practice and technology is accessible, useable, and inclusive.

Instructors can apply UD principles to

- the overall design of the course (e.g., choosing to employ accessible learning management system (LMS) features, carefully organizing the course so that it is easy to follow),
- curriculum (e.g., using accessibly designed documents, captioning videos, developing clear instructions for assignments),
- communication methods (e.g., allowing students to choose communication methods when meeting one-on-one with the instructor, clearly describing the questions to be addressed in a discussion thread), and
- the design of specific assignments (e.g., giving small group members guidelines for conducting accessible meetings, requiring students apply UD to the design of their projects).

It is good news to many instructors that UD, UDL, and WCAG principles do not replace principles of good design of STEM instruction, curriculum, and facilities. Rather, UDE complements them, in the same way that applying UD complements other accepted practices for sidewalk design. Similarly, applying UD does not require abandonment of adopted teaching and learning philosophies, theories, and practices such as cooperative learning, constructive learning, and learner-centered instruction (Burgstahler, 2020). Instead, STEM instructors can apply UDE to ensure that they employ multiple teaching practices and that each practice is accessible, useable, and inclusive. A reasonable approach for creating inclusive online learning pedagogy is for online learning designers and instructors to select evidence-based instructional practices and then apply relevant UD, UDL, and WCAG principles to each practice. This process is presented in Fig. 3.

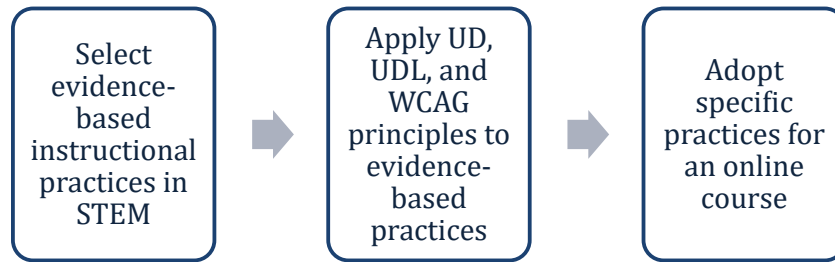


Fig. 3 An approach for applying UD to STEM curriculum and instruction.

UDE guidelines and practices in STEM education

Examples of formal and informal STEM education practices supported by the fourteen UDE principles are presented in **Table 1**.

There is no shortage of guidance to help course instructors develop formal and informal STEM offerings to be accessible to all students (Burgstahler, 2011, 2012, 2015, 2021; Higgins and Maxwell, 2021; Miller and Lang, 2016; Sweet, 2018). For example, the Access to Informal Science Learning (University of Washington, n.d. a) project, AccessISL, at the University of Washington developed a checklist of considerations for designing accessible and useable informal STEM learning (DO-IT, 2022). Addressing the questions presented in the checklist provides a good starting point for making informal learning activities, facilities, and resources accessible to and inclusive of everyone. A few examples follow.

Planning, Policies, and Evaluation. Consider disability and other diversity issues as you plan and evaluate offerings and activities.

- Are people with disabilities, racial and ethnic minorities, men and women, young and old, first generation, individuals with low income, and other groups represented in the project planning processes in numbers proportional to those of the whole campus or community?
- Are disability-related access issues and other diversity issues addressed in data collection, evaluation plans, and instruments?

Table 1 An example of a UDE practice supported by each UDE principle.

UDE principle	Example of UDE practice in STEM education
UD 1. Equitable use	<i>Overall design.</i> In designing and delivering all aspects of a science learning opportunity the designer or instructor maximizes its accessibility, usability, and inclusivity for students with diverse abilities, gender identities, races, ethnicities, interests, and other characteristics.
UD 2. Flexibility in use	<i>Science museum.</i> A bioengineering exhibit allows a visitor with diverse abilities to choose to read or listen to descriptions of the contents of display cases and to engage in hands-on activities.
UD 3. Simple and intuitive	<i>Assessment.</i> Testing in a computer science course is conducted in a predictable, straightforward manner.
UD 4. Perceptible information	<i>Lab.</i> An emergency alarm system in a biology lab has visual, aural, and kinesthetic characteristics.
UD 5. Tolerance for error	<i>Instructional software.</i> A physics tutorial provides guidance when a student makes an inappropriate selection.
UD 6. Low physical effort	<i>Curriculum.</i> Forensic science educational software includes on-screen control buttons that are large enough for students with limited fine motor skills to select.
UD 7. Size and space for approach and use	<i>Makerspace.</i> An adjustable table and flexible work area is useable by students who are right- or left-handed and have a wide range of physical abilities.
UDL 1. Multiple means of engagement	<i>Course.</i> Multiple examples in an environmental science class ensure relevance to a group with diverse interests.
UDL 2. Multiple means of representation	<i>Curriculum.</i> Students are shown a captioned video, assigned textbook reading, and given physical models to manipulate in order to learn a mathematics concept.
UDL3. Multiple means of action and expression	<i>Course project.</i> An assigned project in a general science class optimizes individual choice and autonomy.
WCAG 1. Perceivable	<i>Website.</i> A person who is blind and using a screen reader can access the content in images for an engineering course because text descriptions of content in images are provided.
WCAG 2. Operable	<i>Learning management system (LMS).</i> A person who cannot operate a mouse can navigate all content and operate all functions within a LMS used to deliver STEM courses by using a keyboard (or device that emulates a keyboard) alone.
WCAG 3. Understandable	<i>Instructional materials.</i> Definitions are provided for words, jargon, and idioms some students might not understand; all acronyms are spelled out when first used.
WCAG 4. Robust	<i>Forms.</i> Permission, lab reports, and other electronic forms can be completed using a wide range of devices, including screen readers and other ATs.

Information Resources and Technology. If your offerings use computers as information resources, ensure these systems employ accessible design, that staff members are aware of accessibility options, and systems are in place to make accommodations when requested.

- Do pictures in your publications and website include people with diverse characteristics with respect to race, gender, age, and disability?
- In key publications, do you include a statement about your commitment to access and procedures for requesting disability-related accommodations?
- Are all printed publications available (immediately or in a timely manner) in alternate formats?
- Do videos developed or used in the project have captions? Are they audio-described?
- Do you provide adequate work space for both left- and right-handed users?
- Are procedures in place for a timely response to requests for assistive technology?

Facilities and Offerings. Ensure that facilities, activities, materials, and equipment are physically accessible to and useable by all participants, and that all potential characteristics are addressed in safety considerations.

- Are all spaces welcoming, accessible, comfortable, and safe to a variety of abilities, racial and ethnic backgrounds, genders, and ages?
- Are aisles kept wide and clear of obstructions for the safety of users who have mobility or visual impairments?
- Are wheelchair-accessible and child-friendly restrooms with well-marked signs available in or near the facility?
- Is at least part of a service counter at a height accessible from a seated position?
- Are there ample high-contrast, large-print directional signs to and throughout the facility, including directions to accessible routes? When appropriate, are these signs marked in braille?

Staff. Make sure staff are prepared to work with all visitors and participants.

- Do staff members know how to respond to requests for disability-related accommodations, such as arranging for a sign language interpreter or providing a document in an alternative format?
- Are staff and contractors in specific assignment areas (e.g., web page development, video creation) knowledgeable about accessibility requirements and considerations?

Regardless of how many UDE practices are employed, although the need is minimized, reasonable accommodations may still be necessary to ensure full access and engagement to a particular student. For example, a student with a learning disability may require extra time on a test or assignment.

Conclusion

Practicing UDE instead of providing accommodations alone holds promise for making educational offerings more inclusive, both for students who disclose disabilities and request accommodations and for those with disabilities who do not, UDE practices benefits all students, including those with various learning styles and technological expertise, those whose native language is not English, those who are older than the average student, and those who are members of racial and ethnic minority groups. In summary UDE

- is applied proactively,
- makes a product or environment welcoming to all potential students,
- is useable by and accessible to people with a broad range of characteristics,
- can be implemented in incremental steps, and
- is offered in an inclusive setting.

Acknowledgments

This chapter is based on work supported by the National Science Foundation (Grant #1906147). Any opinions, findings, and conclusions or recommendations are those of the author and do not necessarily reflect the policy or views of the federal government, and you should not assume its endorsement.

References

- Allen, A.G., Anderson, S.C., 2020. Universal design for learning and instruction: overcoming barriers facing students with disabilities in colleges and universities. *J. Educ. Soc. Policy* 7 (4).
- Americans With Disabilities Act Amendments Act of 2008, 42 U.S.C.A. § 12102 Note.
- Americans With Disabilities Act of 1990, 42 U.S.C.A. § 12101 et seq.

- Berge, Z.L., Huang, Y., 2004. A model for sustainable student retention: a holistic perspective on the student dropout problem with special attention to e-learning. *DeosNews* 13 (5). Available at: http://www.ed.psu.edu/acsd/deos/deosnews/deosnews13_5.pdf.
- Braxton, J.M., Sullivan, A.V., Johnson, R.M., 1997. Appraising Tinto's theory of college student departure. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*, vol. XII. Agathon, New York, NY, pp. 107–164.
- Burgstahler, S., 2011. Universal design: implications for computing education. *J. Educ. Resour. Comput.* 11 (3).
- Burgstahler, S., 2012. Equal Access: Universal Design of Computer Labs. University of Washington, Seattle. Available at: <http://www.washington.edu/doiit/equal-access-universal-design-computer-labs>.
- Burgstahler, S. (Ed.), 2015. *Universal Design in Higher Education: From Principles to Practice*, second ed. Harvard Education Press, Cambridge.
- Burgstahler, S., 2020. *Creating Inclusive Learning Opportunities in Higher Education: A Universal Design Toolkit*. Harvard Education Press.
- Burgstahler, S., 2021. Twenty Tips for Teaching an Accessible Online Course. DO-IT, University of Washington, Seattle, WA. Available at: <https://www.washington.edu/doiit/20-tips-teaching-accessible-online-course>.
- Center for Applied Special Technology (CAST), 2018. Universal Design for Learning Guidelines Version 2.2. <http://udlguidelines.cast.org/more/research-evidence>.
- Center for Universal Design, (n.d.). History of Universal Design. https://projects.ncsu.edu/design/cud/about_ud/udhistory.htm.
- DO-IT, 1993–2021. DO-IT Snapshots. University of Washington, Seattle, WA. Available at: <http://www.washington.edu/doiit/programs/do-it-scholars/snapshots>.
- DO-IT, 2022. Equal Access: Universal Design of Informal Learning. University of Washington, Seattle. Available at: <https://www.washington.edu/doiit/equal-access-universal-design-informal-learning>.
- Hagley Museum and Library, (n.d.). The Marc Harrison Collection. Wilmington, DE: Hagley Museum and Library. Available at <http://www.hagley.org/>.
- Higgins, A.K., Maxwell, A.E., 2021. Universal design for learning in the geosciences for access and equity in our classrooms. *J. Appl. Instr.* 10 (1).
- Individuals With Disabilities Education Act (IDEA), 20 U.S.C. § 1400, 2004.
- Jones, S.R., 1996. Toward inclusive theory: disability as social construction. *NASPA J.* 33, 347–354.
- Loewen, G., Pollard, W., 2010. The social justice perspective. *J. Postsecondary Educ. Disabil.* 23 (1), 5–18.
- Miller, D.K., Lang, P.L., 2016. Using the universal design for learning approach in science laboratories to minimize student stress. *J. Chem. Educ.* 93 (11), 1823–1828.
- Myers, K.A., Lindberg, J.J., Nied, D.M., 2013. *Allies for Inclusion: Disability and Equity in Higher Education*. Wiley Periodicals, Inc., Hoboken, NJ.
- National Council on Disability, 2003. *People With Disabilities and Postsecondary Education—Position Paper*. Available at: <https://ncd.gov/publications/2003/people-disabilities-and-postsecondary-education-position-paper>.
- National Science Foundation, 2021. *Women, Minorities, and Persons With Disabilities in Science and Engineering*. National Science Foundation.
- Roberts, K.D., Satkyjgykova, M., Park, H.-J., 2015. Universal design for instruction in postsecondary education: a literature review of empirically based articles. In: Burgstahler, S. (Ed.), *Universal Design in Higher Education: From Principles to Practice*, pp. 65–80.
- Seale, J., Burgstahler, S., Havel, A., 2020. One model to rule them all, one model to bind them? A critique of the use of accessibility-related models in post-secondary education. *Open Learn.* 37 (1), 6–29.
- Seale, J.K., 2013. *E-learning and disability in higher education*. In: *Accessibility Research and Practice*. Routledge, Taylor & Francis, London.
- Section 504 of the Rehabilitation Act of 1973, as Amended 29 U.S.C. § 794.
- Stern, V., Woods, M., 2001. *Roadmaps and Rampways*. American Association for the Advancement of Science, Washington, DC.
- Story, M.F., Mueller, J.L., Mace, R.L., 1998. The principles of universal design and their application. In: Story, M.F., Mueller, M.L., Mace, R.L. (Eds.), *The Universal Design File: Designing for People of All Ages and Abilities*. Center for Universal Design, Raleigh, NC, pp. 32–36. Available at: http://www.ncsu.edu/ncsu/design/cud/pubs_p/docs/udfile/chap_3.pdf.
- Sweet, E., 2018. Accessibility in the laboratory. *ACS Symp. Ser.* 1272. <http://www.doi.org/10.1021/bk-2018-1272.ch001>.
- Tinto, V., 1993. *Leaving College: Rethinking the Causes and Cures of Student Attrition*, second ed. University of Chicago Press, Chicago, IL.
- University of Washington, (n.d. a). *Access to Informal STEM Learning (AccessISL)*. Seattle, WA: University of Washington. Available at <https://www.washington.edu/doiit/programs/accessisl>.
- University of Washington, (n.d. b). *The Center for Universal Design in Education*. Seattle, WA: University of Washington. Available at <https://www.washington.edu/doiit/programs/center-universal-design-education/overview>.
- Vanderheiden, G.C., Vanderheiden, K.R., 1992. *Guidelines for the Design of Consumer Products to Increase Their Accessibility to People With Disabilities or Who Are Aging (Working Draft 1.7)*. University of Wisconsin Trace Research and Development Center, Madison. Available at: http://trace.wisc.edu/docs/consumer_product_guidelines/toc.htm.
- World Wide Web Consortium, 2013. *Guidance on Applying WCAG 2.0 to Non-web Information and Communications Technologies*. <https://www.w3.org/TR/wcag2ict/>.
- World Wide Web Consortium, 2018. *Web Content Accessibility Guidelines, 2.1*. <http://www.w3.org/TR/wcag21/>.

Relevant websites

- Center for Universal Design, <https://projects.ncsu.edu/ncsu/design/cud/>.
- Web Content Accessibility Guidelines, <https://www.w3.org/WAI/standards-guidelines/wcag/>.
- Until Learning Has No Limits, <https://www.cast.org>.
- Center for Universal Design in Education, <https://www.washington.edu/doiit/programs/center-universal-design-education/overview>.

Designing equitable STEM education: Guidelines for parents, educators and policy-makers to reduce gender/racial achievement gaps

David Reilly^a and Aida Hurem^b, ^a School of Applied Psychology, Southport, QLD, Australia; and ^b School of Education and Professional Studies, Griffith University, Southport, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Gender stereotypes about mathematics and science	334
Parents and caregivers	336
Teachers and educators	337
Racial stereotypes about mathematics and science	337
Educational interventions to reduce gender/racial gaps	339
Spatial training	339
Values affirmation	340
Encouraging diversity and broadening the pool of students taking STEM coursework	341
References	342
Relevant websites	343

The importance of increasing the diversity of persons entering the STEM pipeline has long been recognized, but only recently has it reached a degree of prominence that educational institutions and policy-makers are now mandating concerned efforts. The fields of science, technology, engineering and mathematics have long battled a problem of diversity. However, the focus of these efforts has changed in recent decades: initially, the focus was primarily confined to the underrepresentation of women in STEM, but increasingly this has broadened to include racial diversity (and the intersection of gender and race), socioeconomic class, neurodiverse persons (Wei et al., 2013), sexual orientation and gender identity, as well as persons with a disability.

Part of the reason for this expansion of focus has certainly been social changes in attitudes toward diversity. However, another factor has been the acute shortage of STEM-trained workers for an ever-growing variety of professions and the economic imperative to address this problem. A dire shortage of workers at all ability levels with STEM-proficiencies in Western countries has been temporarily catered for by importing overseas talents. Still, industry, educators, and policy-makers increasingly acknowledge the need to broaden the pool of new entrants into the STEM-pipeline. Not every child will grow up to become a scientist or engineer, but science and technology literacy is regarded by educators and economists alike as just as important for occupational success as reading and mathematics literacy. And we all benefit from a scientifically literate populous that can use critical reasoning and evidence as part of the basis for decision-making.

This presents a singular challenge for educators: how can we design an educational environment that cultivates and nurtures an interest in STEM for *all* students, and sees them carry this interest forward into their high school and college years. Young children are innately curious and passionate about science, but many lose this interest and see science as “not for people like me” in their high school and college years. Educators are very adept at teaching scientific concepts and principles, especially when adequately resourced and trained. But a higher standard is called for if we are to design an educational environment that delivers on the goal of inclusiveness. Meeting the challenge of a broadened scientific pipeline requires more than just rote learning and instruction: as educators, we need to inspire interest in STEM. Our goal should be that every student with the inclination sees the value in learning these skills, and that a career using those skills is something attainable and viable for someone like them.

With this aim in mind, we can glean some valuable lessons and perspectives from the social sciences (and psychology in particular) for cultivating a more inclusive STEM learning environment. In this article, we offer an overview of how educational psychology's perspectives can contribute to the design of equitable and inclusive STEM education. We also offer some evidence-based interventions that have been shown to improve STEM outcomes for underrepresented demographic groups (such as for women, for minorities, and for those from socially disadvantaged backgrounds).

Gender stereotypes about mathematics and science

One of the strongest barriers to students developing and maintaining an interest in STEM are gender stereotypes about STEM. Gender stereotypes that mathematics and sciences are inherently masculine domains of ability are widely held by students and adults alike, and are often due to the influences of old societal and patriarchal beliefs. However the intensity of such beliefs does vary across individuals. Gender stereotypes are of particular interest to educational researchers because of all social identity groups (race, social class, etc.), gender applies to the greatest number of students; hence there has been greater research into this than other social categories.

Such stereotypes might be expressed as explicitly held beliefs (“men are just better at maths”), or they may be more implicit attitudes (such as stronger associations between science and masculinity than for femininity). They also operate at many levels. Firstly, we see a direct impact of endorsement of gender stereotypes on students’ attitudes toward mathematics and science, with boys reporting more positive attitudes to studying STEM and girls reporting lower self-confidence and self-efficacy. Younger children show no such gender differences, however. These only tend to emerge during early high school years which is also a time of intensifying gender conformity pressures from peers and society (Martin and Ruble, 2010). Lest, the reader assumes this might reflect “innate” differences in patterns of interest between males and females, we find a curious pattern when we investigate STEM attitudes and abilities internationally. Reilly et al. (2017b) reported an analysis of the Trends in Mathematics and Science Survey (TIMSS), which assessed over a quarter of a million students across 45 different nations. For OECD countries (which are composed of primarily Western and developed nations) boys did indeed report greater interest in STEM while girls reported lower self-efficacy, but in many Middle-Eastern and Asian countries this pattern was reversed: with girls reporting greater interest, self-efficacy, and achieving higher mathematics and science achievement scores.

Secondly, both attitudes toward STEM and self-efficacy beliefs are important drivers of student subject selection in high school. While mathematics and science are generally required subjects at younger ages, in the later years of high school students are offered greater flexibility in their coursework and can voluntarily select some or all of their coursework. Students face an even greater range of choice on entering college and university: research shows that those who hold more positive attitudes and feel more confident about their abilities are more likely to choose STEM coursework than those who do not (Eccles, 2013). If we want to boost the number of women pursuing STEM careers, then inclusive education must start first in the younger years.

Thirdly, endorsement of gender stereotypes about ability affects how students approach STEM, and reduces motivation when students encounter difficult content. When “stuck” with a concept or technique they are trying to master, female students who have high endorsement of stereotypes are more likely to believe they lack the capability to solve it. Endorsement of gender stereotypes also affects student performance in examinations and standardized tests of achievement. This is particularly so when the outcome of the test has higher stakes (such as a college entry exam, or a final exam).

So where do gender stereotypes come from, and how strong are they? At an early age children are naturally inquisitive and that curiosity extends to the natural and physical sciences. Younger boys and girls are equally as interested in science. But as children age, they are influenced by parents and caregivers who control the types of toys and activities provided to their children, then by educators and teachers who are often their first introduction to the world of science. They are also influenced by various agents of socialization such as popular culture and media: when these representations of scientists and engineers are not representative of the general population and show a lack of diversity, it can internalize a message to very malleable young minds that science is “not for people like me”. Evidence suggests that these messages are acquired quite early, and that children begin to implicitly associate males with science and girls with language. In older adolescents the strength of these associations increases quickly, and there are measurable differences in attitudes toward STEM and self-confidence which increase further still into young adulthood. Let us look at some of the research on the strength of these gender-STEM stereotypes, and their effect on educational outcomes.

For younger students for whom filling out surveys is not ideal, one way of measuring these implicit gender stereotypes is the “Draw-A-Scientist” task, where children are asked to draw a picture of what a scientist looks like. As you might imagine, such drawings often reflect depictions of scientists in popular culture: scientists wearing glasses and dressed in lab coats, often middle-aged or older (Ferguson and Lezotte, 2020). It’s a task that allows a child to project their worldview without being led by the researcher or other adults present. Invariably, the scientists drawn are typically depicted as male (often Caucasian), even when girls are asked to draw them. A review of research conducted over the span of 5 decades in US children across K-12 found that 73% of all scientist drawings were male (Miller et al., 2018). While this attests to gender-STEM stereotypes’ pervasiveness, optimistically, the review found some signs of progress with time. There was a small decline in the absolute number of male drawings in more recent years. This was attributed to an increased representation of women in STEM and female scientists in popular culture, which has been referred to as the “Scully Effect”. A recent study of women engaged in STEM careers found that 63% reported that they were influenced by the portrayal of Dana Scully in the *X-Files* (Geena Davis Institute on Gender in Media, 2018), highlighting the importance of availability of female role models in STEM. While media and popular culture do have a role to play in shaping more positive attitudes toward STEM, this in isolation will be insufficient to address the underrepresentation of women in STEM fields.

Educational researchers can use surveys with older students to measure students’ endorsements of gender stereotyped beliefs (“such as males are better suited to mathematics and science”), as well as student interest in learning about mathematics and science or even student confidence. This varies across individuals and so can predict individual differences in educational outcomes. For example, higher endorsement of gender stereotypes about female ability is strongly predictive of lower female performance (Steele, 1998; Spencer et al., 1999). This is especially so for high-stakes standardized tests (such as the SAT-Mathematics test used for college entry in the United States) or end-of-semester examinations. Gender differences are also frequently observed by middle-school in STEM interest and self-confidence for US students and increase with additional years of schooling (Miller et al., 2006). This finding has also been replicated cross-culturally in international studies of mathematics and science attitudes (Reilly et al., 2017b).

Even for students who do not endorse explicit gender-STEM stereotypes, nonconscious implicit beliefs may also contribute to their relationship with STEM (such as lowering confidence). These cannot be measured directly on a survey but can be inferred indirectly through a variety of mechanisms. In older students, we can use a projective measure like the Implicit Association Test (IAT), which has been used to measure the strength of the gender stereotype (math = male). Nosek et al. (2002) found that

stronger implicit beliefs in this stereotype were correlated with more negative attitudes toward maths in women but more positive attitudes for men. A later study that investigated this effect cross-culturally found a moderate correlation between national gender-math stereotypes (using IAT data) and the magnitude of gender differences in maths achievement found in individual countries on PISA assessments (Nosek et al., 2009). Stronger math = male implicit associations lead to larger gender gaps in mathematics on standardized tests. So, there is good evidence that these stereotypes are widely endorsed (at least in Western countries), and that they impact mathematics and science achievement scores. However, in cultures/educational environments where these implicit beliefs are lower, girls actually score higher than boys in STEM (Guiso et al., 2008; Reilly, 2012). The take-home message of this research is that gender differences in educational outcomes are not immutable; the educational environment a student is raised in matters, especially in relation to the likelihood of maintaining their initial interest into STEM which can further develop into adulthood. With that in mind, we offer some suggestions firstly to parents and caregivers, and secondly for educators and policy-makers.

Parents and caregivers

Parents and caregivers can play a pivotal role in contributing to the acquisition of gender stereotypes, especially with younger children, as parents largely determine the activities and resources provided to the children in their care. This manifests through the toys, leisure pursuits and educational enrichment experiences made available to children as well as their parent-child interactions. At very young ages gender differences in socialization practices (and in particular, differential reinforcement of gender-appropriate toys and activities) were once thought only to influence gender differences in social behavior (Leaper and Gleason, 1996). However, play is an important part of child development and can be an enriching educational experience (Lillard, 2015).

Encouraging a broad repertoire of toy play is important, as play stimulates cognitive development: it gives the opportunity to practice important skills such as creativity, problem-solving, spatial reasoning, etc. Of relevance to early childhood education, many stereotypically masculine childrens' toys (construction blocks, cars and trucks, computer games, etc.) provide additional opportunities to practice and train visual-spatial skills (Reilly et al., 2017a), as well as to a lesser extent mechanical reasoning. These lay down a foundation for quantitative reasoning (mathematics and science skills), and have been recognized as giving children a critical headstart in STEM. For example, Lowrie et al. (2017) report that providing additional visual-spatial training improves elementary childrens' mathematics performance, and many educators now advocate incorporating an element of visual-spatial training in mathematics programs (Newcombe and Frick, 2010). Toy selection alone will not dictate a child's life outcomes, but it will contribute to their initial spatial abilities and readiness for STEM.

Parents are often regarded as a child's first teacher, because the lessons they will learn in the classroom rests on a foundation of informal learning experiences that begin at home. As most parents tend to buy gender-consistent toys for their children, many girls miss out on masculine educational toys that promote spatial development. Parents and caregivers can make targeted selections though that will provide additional spatial learning experiences at home for both boys and girls (Reilly et al., 2017a). This can start with geometric shape toys and build up to construction blocks such as Legos. Parents can also engage directly with the child in shared play, for example, by using spatial language during block play (Ferrara et al., 2011; Jirout and Newcombe, 2015), encouraging children to play games that promote spatial awareness, using jigsaws, maps and mazes to facilitate spatial perception and navigation skills.

Children are naturally inquisitive explorers of the world around them, or what the French educational psychologist Jean Piaget called "little scientists". They rapidly soak up an understanding of the world around them at these younger ages, including language, social behavior, and even rudimentary scientific principles (if I push the cup of water off the table, it will fall). While children remain curious their mind is flexible; but at later ages the range of possibilities for what a child will become narrows. Once children reach adolescence, they experience a sharp increase in gender conformity pressures from their peers (Ruble et al., 2006). Children begin to self-select their interests and leisure activities, and this extends to academic domains and tentative occupational aspirations (such as exploring a career in science or technology fields).

Yet parents are not completely without influence at these ages either, and parental beliefs about the importance of mathematics and science still make a strong contribution to girls' and boys' attitudes toward STEM (Jacobs and Bleeker, 2004). Where there is even a small spark of interest and curiosity, parents can exert a powerful contribution by providing access to science books, games, and science kits. However, parents tend to provide boys with more science-based educational toys and activities (Jones et al., 2000; Alexander et al., 2012), such as magnets, magnifying glasses, electrical circuits and fuses, mechanical kits, etc. than to girls. Yet, all of these though are equally appropriate for daughters, are low in cost, and can be a useful way of engaging in parent-child discussions about how science can be applied in everyday life. Also important are providing enrichment experiences like trips to science museums, planetariums and zoos, especially for children from lower socioeconomic backgrounds who might not get to experience these in school trips (Gerde et al., 2021). Sadly, daughters generally engage in far less science-based extra-curricular activities than sons (Catsambis, 1995). Drawing attention to this, and educating parents and caregivers about ways in which they can engage their daughters in such science-based extracurricular activities is important as, in addition to providing a nurturing educational experience, it sends a tacit endorsement to the adolescent child that these things matter, regardless of whether they are boys or girls. These informal learning experiences will reap dividends in the years ahead and expand the range of possibilities they can explore.

Teachers and educators

Mathematics and science literacy are already important cornerstones in a child's education, and both are taught as compulsory courses in most jurisdictions for a large part of a student's academic career. However, as mentioned earlier, in adolescence we see a sharp decline in girls' STEM self-confidence and attitudes (Catsambis, 1995; Reilly et al., 2017b). Some educators might genuinely wonder whether this reflects "innate" differences in aptitude between girls and boys for STEM, especially if they have read studies reporting higher male performance on standardized tests of mathematics and science (Hedges and Nowell, 1995). While males as a group do tend to score somewhat higher on maths and science tests than females, the differences in averages tend to be relatively small in magnitude (Reilly et al., 2015). They are only really consequential when you look at the gender ratio of high-achieving students, where males outnumber females by a factor of 2 or more in both domains (i.e., there are twice as many males as females at this level). This can lead to exaggerated impressions among educators and policy-makers of vast differences; gender ratios in high-achievers are significant but not immutable. Additionally, for science achievement scores they vary across content area: gender gaps are largest for physics and earth sciences, but girls score equally as high as do boys in biology and life sciences (Reilly et al., 2015). Indeed, when we look to young adult samples, females often outnumber males in many fields of science such as biology, psychology, and medicine. Thus there is no evidence of a lack of aptitude, and as highlighted earlier in international studies, there are many countries where girls score substantially higher than boys in both mathematics and science (Reilly, 2012; Reilly et al., 2017b).

Rather, gender differences in maths and science achievement on standardized tests (e.g., Hedges and Nowell, 1995) appear to be due to a tendency for girls to underperform on such tests due to the phenomenon known as *stereotype threat* (Steele, 1997). This is where awareness of gender stereotypes about STEM undermines confidence and increases test anxiety, especially on "high-stakes" tests like end-of-year tests or college/tertiary entrance exams—it can become a self-fulfilling prophecy that leads to underperformance. When girls are assessed by other means (such as written reports, essays) they tend to score on par with boys or somewhat higher. This is reflected in the gender gap in school grades, where girls as a group score significantly *higher* than boys in mathematics and science coursework (Voyer and Voyer, 2014). This is true even at more advanced levels in high school and college, where women typically earn STEM higher grades than men. There are, however, areas of science where this is the exception, such as chemistry and physics, and where numbers of women are also comparatively fewer.

Gender stereotypes about STEM can be particularly harmful, because they reduce interest and confidence in pursuing further maths and science coursework (Shapiro and Williams, 2012), and also undermine test performance when girls and women's abilities are assessed in exam situations. Reminding students that these stereotypes hold little validity, paradoxically, can be counter-productive as it reinforces awareness of the existence of those stereotypes (Liben and Coyle, 2014). Many well-intentioned efforts at improving gender diversity in STEM, such as the practice of single-sex schooling (Halpern et al., 2011), actually draw attention to gender distinctions and creates greater disparities in outcomes. Fortunately though, there are a number of evidence-based interventions that have been demonstrated to improve the performance of girls and women in STEM subjects, and can be offered broadly to all students rather than selectively targeting a select few. They are also relatively brief in terms of time, effort, and resources required to deliver them. In addition, they have the added benefit of improving STEM motivation and achievement more broadly (rather than just performance scores). This has importance for STEM course-selection and decision making by students. We review these later in the article under interventions, after addressing racial stereotyping and STEM.

Racial stereotypes about mathematics and science

Stereotypes about intellectual ability and race also present a problem for educators seeking to increase STEM diversity. Their origins are complex and multifaceted, but it was not so very long ago that some psychologists argued that white students were born smarter than African-American students, such as in Herrnstein and Murray (1994) widely repudiated book *"The Bell Curve"*. While the American Psychological Association devoted a considerable volume of research refuting such claims (e.g., Kincheloe et al., 1997), most readers will be familiar with the term "the Bell Curve" even if they are vague on the specific details of the book. Similar arguments were made decades earlier against desegregated education in the United States, but if we look more broadly, there are negative racial stereotypes about many ethnic and minority groups cross-culturally: such is the nature of prejudice. It is not confined to a single ethnic group.

When any minority group is systemically devalued and told that they are lesser, it sets the ground for reduced confidence about their intellectual abilities in those students. Steele and Aronson (1995) termed this phenomenon *stereotype threat*, and first applied it with African-American students. When there are negatively held stereotypes about intellectual abilities, especially when student ability is examined in high-stakes testing situations, this can lead to diminished academic performance. The psychological processes involved are multifaceted: but stereotype threat lowers memory recall (Régner et al., 2010), reduces problem-solving and the number of questions student attempt (Rivardo et al., 2011), and decreases student motivation and effort. In a meta-analytic review of more than one hundred studies investigating stereotype threat, Nguyen and Ryan (2008) found similar decreases in test performance with ethnic minorities as those found in studies investigating gender and stereotype threat. It is particularly problematic when students are "primed" to consider their group identity: such as filling out a pre-test demographic questionnaire that inquires about ethnicity (as is often used in state and national testing), or when there are cues in the testing environment (for example, being the only member of that minority group present relative to a room of white peers).

However, stereotype threat also affects students outside of testing situations: self-beliefs about academic abilities are pivotal to student motivation and persistence in the face of difficulties. It is well established that attitudinal factors and emotions play an important part in academic performance. Many students internalize negative beliefs about the racial/ethnic group to which they belong to, and may see academic achievement incompatible with that identity. This results in lower motivation, lower attendance, and higher dropout rates. Consequentially, it leads to a further narrowing of the pool of potential STEM candidates selecting mathematics and science coursework in high school and then again in college/university when considering a STEM major. [Steele \(1997\)](#) writes that there is a “universal connection between stigmatization and poor school achievement” (p. 623), and reports many examples cross-culturally. There are even examples of stereotype threat in caste-based minorities of the same race, with similar-sized educational gaps to that found between African-American and Caucasian students, so the effect seems to apply to any stigmatized social group rather than reflecting any actual differences in aptitude.

At the heart of these stereotype threats, effects on student attitude and performance is the internalization of the message “I don’t belong here ... STEM isn’t for people like *me*”. Belonging is a fundamental human need ([Baumeister and Leary, 1995](#)), and feeling like you belong in an academic environment is critical for academic success. However, this is harder for underrepresented racial minority (URM) students to attain in a student body predominantly composed of white students and where they may lack established social networks such as in the transition to high school or college. [Walton and Cohen \(2007\)](#) examined students’ sense of belonging in undergraduate STEM students, finding that URM students rate their sense of belonging significantly lower than their peers. Challenging these student perceptions can be difficult, because targeted interventions that act on a specific ethnic/racial group can often highlight the salience of race (just as we see the same with targeting only female students mentioned earlier). One intervention though that shows promise directly targets a student’s sense of belonging, and acknowledges the difficulties URM students may be facing. [Walton and Cohen \(2007; Study 2\)](#) found that a 1 h information session completed as part of their orientation substantially reduced the racial achievement gap in STEM students. Students were presented with information on how common it is for students to feel out of place and like they don’t belong, regardless of race, including illustrative quotations from past students about how this improves over time and gets easier. It also presented the results of a study showing graphs of how these fears were common across racial and gender groups. This was delivered to *all* students in the study regardless of demographic category so as not to make group identity salient. The students were then tracked over the course of their first semester of study. While there was no substantive difference for white students, its effect on minority students was remarkable. Relative to controls, URM students receiving the intervention earned substantially higher GPAs ([Walton and Cohen, 2007](#)), and this effect persisted throughout the course of their degree ([Walton and Cohen, 2011](#)). Such interventions are brief and low-cost, so might be implemented in introductory STEM courses at colleges and universities. But they might equally be considered earlier in high school at the beginning of the school year.

Up until this point, we have focused primarily on attitudinal factors within students themselves. However, it would be remiss not also to consider the impact of contextual factors, which includes the attitudes of STEM teachers themselves. Many educators will be aware of the classic “*Pygmalion in the classroom*” study that showed just how teacher beliefs about their students’ potential translate into measurable differences in learning outcomes ([Rosenthal and Jacobson, 1968](#)). Even in the absence of overt racism, there is the potential for situational cues from STEM instructors to reinforce racial stereotypes. There is some evidence to suggest that these Pygmalion effects are stronger for URM students, and that URM students report their instructors as having lower expectations for them ([McKown and Weinstein, 2008](#)). It also illustrates how even educators who hold quite egalitarian racial attitudes might still reflect the prevailing implicit beliefs of their culture ([Greenwald et al., 2002](#)), through the burden of low expectations. Conversely, educators who can see the potential in their students (even when those students cannot yet see it themselves) can be transformational in inspiring STEM potential to blossom ([Rochmes, 2018](#)).

A comprehensive longitudinal study by [Canning et al. \(2019\)](#) investigated whether this might be the case for STEM courses, given the afore-mentioned implicit attitudes of scientists being white males. They examined racial gaps in STEM course grades across more than 150 different instructors, numerous STEM disciplines, and a sample of over 15,000 students. They also surveyed instructor beliefs about the malleability of intelligence: whether intelligence is fixed at birth, or whether intelligence can grow with effortful learning and practice. They found that racial achievement gaps were substantially larger when taught by instructors who held a “fixed” mindset compared to those instructors with a growth mindset. This was observed even after controlling for other influential demographic factors like gender, socioeconomic status, first-in-family to attend college, etc. There is also evidence that teaching students about a “growth mindset” may help change how they see themselves and their approach to learning STEM, especially for students entering high school ([Yeager et al., 2016](#)). Many students believe their ability to be fixed: if they can’t understand a particular task like algebra or chemistry now, then they won’t be able to understand it in the future. When this is reinforced by implicit cultural attitudes about intelligence and race, it can affirm a “fixed mindset” and discourage further practice. The argument has even been made that this should be taught to all students, but [Yeager and Dweck \(2020\)](#) note that it is particularly effective in cases where students (due to stereotyping) have low expectations of themselves.

If we want to broaden the diversity of STEM graduates to create a more inclusive STEM workforce, then we must work hard to capture their natural curiosity about STEM in early education, and then carefully nurture and safeguard it throughout their schooling career ([Herrera and Hurtado, 2011](#)). Studies generally find that URM students hold the same level of initial interest and enthusiasm for STEM as their white peers across middle and high school ([DeWitt et al., 2011](#)), as well as similar levels of parental encouragement and aspirations toward STEM. However, they also have less access to science enrichment activities outside of the school setting, and may lack the financial resources for technology that other children take for granted ([Archer et al., 2012](#)). It is here that educators and policy-makers can make a critical difference by providing funding and support for additional STEM enrichment experiences and technology access for the students in their care. [Archer et al. \(2012\)](#) identified ages 10–14 as a “critical

period” for URM students and STEM, during which additional science experiences and resources can help nurture their interests and talents. However, we would caution against interventions targeted at specific ethnic groups because of the unintended consequences of labeling; rather, students from any demographic group would benefit from increased technology access and STEM enrichment, such that there is a flow-on effect with increased ethnic representation.

Another factor identified as contributing to underrepresentation in STEM is that, despite equivalent interest in STEM, URM students tend to enroll in less challenging math and science coursework (Xie et al., 2015). They are also more likely to be placed in remedial and low-track academic streams. Undoubtedly, this has a profound impact on students’ academic self-image and motivation and leaves them unprepared to undertake further STEM coursework in college. Educational support and intervention prior to the age of voluntary course selection may be critical in building STEM aptitude and confidence, as well as psychoeducational interventions like values affirmation (see below).

Educational interventions to reduce gender/racial gaps

Frontline teachers and educators have the very best of intentions in promoting an inclusive learning environment but have only finite time, energy, and resources to do so. Educational fads come and go, with mixed results that do not always translate to improved learning outcomes (Dinham, 2017). Therefore, it is imperative that adopted educational interventions and strategies have a strong track record of proven results. This is referred to in psychology as evidence-based interventions. When conducting research in this area, educational psychologists largely write for a psychologically trained audience who can understand esoteric statistics and then publish in psychology journals hidden behind paywalls to which only universities usually have access. This disseminates information within the field of psychology well, but we often do a very poor job then of translating that body of literature into an accessible format for educators and policy-makers. The end-result means that some very promising educational interventions have been developed but never reach the coalface of teaching in schools. We seek to remedy this here today.

A number of educational interventions have been proposed for addressing racial, gender, and socioeconomic gaps in educational outcomes for STEM. Many though lack a firm and reliable evidence base, have relatively weak effects, or require considerable investment in resources or classroom time. However, two particular educational interventions that have withstood the crucible of scientific scrutiny, have been widely applied across a range of STEM disciplines (generalizing to all areas of mathematics and science), and are both low in cost and their demands on classroom time. Additionally, these interventions’ effects are so strong that they translate to enduring effects rather than temporary, transient effects. Both spatial training and values affirmation can be applied with any age group, and with limited resources. While we would highly recommend that policy-makers consider adopting them, their potency and low-cost mean that STEM educators could easily implement them in their own classroom settings.

Spatial training

Visual-spatial skills are the ability to perceive and understand spatial relationships, to visualize spatial stimuli such as objects, and to mentally manipulate or transform them in some way (such as rotating an object to imagine what it might look like from a different angle or perspective (Reilly et al., 2017a)). These skills are used quite heavily in certain areas of mathematics (like geometry and trigonometry), as well as more applied fields of science such as physics and chemistry. For several decades, evidence has been growing that visual-spatial reasoning lays down a foundation for quantitative reasoning (a collective term for mathematics and science thinking). Though the exact reasons for this are still unclear, more than 50 years of educational research demonstrates the relationship between spatial ability and science achievement scores is quite strong (Wai et al., 2009). Furthermore, it applies broadly across many different STEM disciplines including chemistry, physics, engineering, and computer science (Uttal and Cohen, 2012).

Typically, males have significantly higher visual-spatial ability, and the size of this gender gap is quite substantial. The reasons for this are complex and a matter of some debate in psychology (for a review see Halpern, 2011), but there is strong evidence that the amount of spatially-oriented play and leisure activities students participate in is positively correlated with visual-spatial ability (Doyle et al., 2012). As with other academic skills like reading, visual-spatial skills require considerable practice to develop fully. For most children and even many adults, visual-spatial skills are underdeveloped and require further environmental input in the form of spatial learning experiences. However, the gender gap in visual-spatial ability is substantial, and spatial training offers the potential to reduce this. Fortunately, there is a way to “shortcut” the many hundreds of hours such activities would normally require and boost visual-spatial reasoning considerably. Spatial reasoning skills are highly malleable and can be improved with training (for a review see Uttal et al., 2013). In a meta-analytic review of over two hundred studies, the average effect of training was moderately large ($d = 0.47$). While early education for spatial learning is desirable (Newcombe and Frick, 2010), students can benefit from spatial training at any age with the review reporting similarly sized effects in adolescent and adult samples alike. As spatial reasoning scaffolds later quantitative reasoning, many educators advocate providing additional spatial training. Again, we would caution though that this should be offered to all students, rather than selectively targeting females in isolation. Students from lower socioeconomic backgrounds, for example, also tend to have less developed spatial reasoning (Levine et al., 2005), and this may help level the playing field. Interventions can take the form of specific lessons teaching particular spatial skills, practicing spatial tasks like mental rotation and then receiving feedback on their performance, or even video games that have students perform spatial reasoning tasks. All offer fairly substantial improvements to visual-spatial reasoning skills (Uttal et al., 2013).

Educators might legitimately wonder whether it is worth allocating time and resources to spatial training for their students, and whether this actually translates to improved STEM outcomes. By way of example, we present two studies investigating the flow-on effect for mathematics achievement: one for younger children, and one for high school-aged students. A recent study by [Mix et al. \(2020\)](#) investigated whether providing several very brief, 30 min spatial training and practice sessions over the course of a week would boost elementary students' spatial and mathematics performance. The study recruited students from first and sixth grade, and the interventions were administered using iPad devices with which children were already familiar. What is notable about this particular study is the wide array of tasks they employed to measure mathematical ability (to test whether spatial training would transfer to all or just certain types of mathematics skills). These included number line estimation, arithmetic calculations without calculators, word problems, and basic algebra. Students showed a significant improvement in their mathematical performance for all tasks compared to a control group that did not receive spatial training. In another high-profile study, [Miller and Halpern \(2013\)](#) report the results of a longitudinal study into the benefits of offering spatial training in a group of STEM undergraduates. After experiencing six 2 h spatial training sessions, the gender gap in visual-spatial ability narrowed considerably relative to a control group that did not receive the training. Students were then followed over time, and their test scores in a college physics course were examined. Students who received spatial training scored significantly higher in their physics mid-terms, as well as the more demanding final exam. Students also reported greater self-confidence in their physics abilities relative to controls. Additionally, students who were enrolled in biology courses also went on to earn high grades than the control group. Similar results were reported by [Stieff et al. \(2014\)](#) with chemistry grades. These and other studies illustrate the benefits of spatial training for all students, and that such training translates to both improved mathematics and science achievement ([Sorby et al., 2018](#)).

Values affirmation

Gender and racial stereotypes about STEM prove particularly problematic for inclusive education. Even when taught by passionate educators who value the worth and dignity of every student, these stereotypes are still foremost in the minds of the students they teach: students are keenly aware that their identity (as a minority) differs from that of the typical STEM scholar. Thus, diversity training and affirmative action initiatives may be moot without also considering the social-psychological processes that affect student learning. As mentioned previously, attempts to "challenge" these stereotype beliefs can be problematic: even when explicit beliefs are refuted by counterexamples (such as the previously mentioned higher female grades in many STEM subjects), the implicit beliefs remain resistant to change. Indeed, they may even increase in intensity through the highlighting of race or gender identity. Rather, we need to inoculate students against these stereotypes and increase their resilience in the face of them.

Values affirmation is a brief yet effective intervention shown to deliver considerable increases in STEM achievement and motivation. More importantly, they can plausibly be administered to all students rather than singling out specific subgroups. After all, many students at one time or another feel like they don't belong, that the degree or courses they've chosen are not for "people like them", or what is commonly referred to as imposter syndrome. Though it may affect any type of student, women and ethnic minorities are greatly overrepresented in studies investigating the prevalence of these doubts ([Bravata et al., 2019](#)). While gender and racial subgroups are the most widely researched in this area, values affirmation is also equally applicable to other categories: students coming from a low socioeconomic status background, first-in-family to undertake college/tertiary studies, LGBTQIA, etc. As it targets motivation and strengthens resolve, it offers benefits for all students who receive the intervention.

The active ingredient in a values affirmation intervention is a reflective writing activity, typically set administered in class-time so that students give it serious consideration. [Miyake et al. \(2010\)](#) had students taking an introductory physics course at college complete a 15 min writing exercise in the first week of class. Students were told that communication is an important skill in a STEM degree and that rather than writing about physics (which they've yet to learn), they would instead practice writing about something they were already familiar with: themselves. In the values affirmation condition, they were asked to identify their most important personal values from a pre-defined list (such as inherent curiosity, their relationships with family and friends, being part of the community, their religious or spiritual values), and in a series of prompts were then asked to write why these were important to them. These served to remind the students of why they were here, and how they hoped to use the skills they were learning to have an impact. The mere act of writing an essay about their personal values appears to lead to greater tolerance for stress in the face of adversity and reaffirms people's individual worth ([Cohen and Sherman, 2014](#)). A control group completed a neutral writing task instead. Such exercises have been found to reduce racial achievement gaps in middle-school students ([Cohen et al., 2006](#)), but [Miyake et al. \(2010\)](#) were the first to examine its effect in reducing the gender gap in college students taking introductory STEM courses. Typically, the "washout" rate of these introductory courses is high, and they are sometimes referred to by educators as "weed-out classes" ([Chawla, 2020](#)). However, they can also be a bottleneck that greatly diminishes the diversity of students studying STEM: so even a very small impact of a values affirmation intervention could be greatly beneficial here. Student progress was then followed by the researchers over the course of the semester, and their final grades examined. The gender gap was eliminated entirely in those receiving the intervention, and the modal (most frequently occurring) grade for the values affirmation group raised from a C to a B. Improvements were highest though for women who endorsed had previously endorsed the belief that men are better suited to physics.

Given the simplicity of values affirmation exercises and the impressive effect on grades, there might be some natural skepticism among educators. After all, sometimes, psychological interventions can be due to sampling error (if the randomly selected control group was slightly lower in initial ability than the treatment group), and so may not be repeatable. Educational researchers set out to

replicate the values affirmation study with their own student samples. A replication study by [Kost-Smith et al. \(2012\)](#) found a similarly large effect on mid-term and physics finals at the University of Colorado, and as they had access to their student's SAT scores, the researchers were able to control for this statistically to ensure that the groups were equivalent at the outset. Other researchers have investigated whether values affirmation interventions might benefit other demographic groups, such as students coming from a lower socioeconomic background or belonging to an URM in STEM. [Harackiewicz et al. \(2016\)](#) applied a values affirmation intervention to four semesters of first-year biology students, finding that it reduced the achievement gap in students who were the first-in-their family to attend college and as well as URM students by about half a GPA point. Offering such interventions in introductory STEM courses might well make for a more inclusive, and more diverse student composition, as it buffers students against adversity and feeling like they don't belong. While it has been studied most heavily on boosting STEM achievement in female students and underrepresented ethnic minority students, it might well offer benefits for other demographic categories such as different sexual identities or orientations, students with a disability, etc.

Encouraging diversity and broadening the pool of students taking STEM coursework

Throughout this article, we have presented the perspectives of educational psychology research in the hope that this information will be of use to parents, educators, and policy-makers: all of whom have important roles to play in reaching a more diverse and inclusive STEM-trained society. We have focused most specifically on improving gender and racial diversity in STEM, and to a lesser degree socioeconomic status. Together these three areas have received the lion's share of attention in the research literature, and collectively represent the largest demographic groups. However, it would be remiss not to also briefly mention other subgroups, such as the underrepresentation of students with a disability ([Weinstein et al., 1987](#); [Xie et al., 2015](#)) and sexual minorities ([Sansone and Carpenter, 2020](#)) who are also less likely to participate in STEM. But imagine what a different world it would be today without the past contributions of scientists like Stephen Hawking (in the case of physical disabilities) or mathematician and father of artificial intelligence Alan Turing (who was jailed for homosexuality, but only after completing his work on breaking the Enigma encryption used by the Germans in WW2). Inclusive STEM education must make all willing students welcome: that requires a culture of support and respect that does not discount people because of group identity. But it must also consider the possibility that those who may differ from the prototypical image of a scientist may not yet recognize their own strengths and aptitudes, and may need encouragement and mentoring by parents and educators.

There are some scientific challenges that can eventually be solved by devoting sufficient computational or labor-force to it, albeit with some inefficiencies and diminishing returns. But there are other, more herculean, scientific problems that have remained intractable for many years. They require that spark of ingenuity, that leap of inspiration, that just might lie in the hands and heads of the next generation of students. So, with this in mind, there is one other category of student that we would urge parents and educators to consider encouraging and supporting in their STEM journey if indeed that's where their talents and interests lie. That would be the neurodiverse student who may take a particular interest in mathematics, science or technology.

For those unfamiliar with the term, neurodiversity is a broad catchall for individuals on the autism spectrum,¹ but also includes many other conditions such as attention-deficit disorder, dyslexia, dyscalculia, which educators frequently encounter. There is some evidence to suggest, for example, that autistic individuals are more likely than the general population to hold an interest in either mathematics or science, and as such, are highly prevalent in these fields. In a nationally representative sample of autistic students attending college, [Wei et al. \(2013\)](#) found that autistic individuals were significantly more likely to have chosen a STEM major than other disciplines, while [Baron-Cohen et al. \(2001\)](#) report that scientists score higher on the Autistic-Spectrum Quotient (a measure of autistic traits) than non-scientists. Similarly, [Shmulsky et al. \(2019\)](#) report in a survey of STEM educators that there is a high prevalence of autistic students among higher achievers, and that they often bring fresh and innovative perspectives. As they have found, the most common features in autistic individuals studying in STEM fields are creativity, abstract-thinking and critical reasoning.

These are exactly the attributes we need in the STEM workforce, and we would urge educators and career guidance counselors not to overlook the possibility that neurodiverse students have to offer. Additionally, there's quite a high prevalence of students with various levels of attention-deficit disorder, reading impairments, or even dyscalculia among college students ([DuPaul et al., 2009](#); [Nelson and Gregg, 2012](#)). Educational accommodations for college and university students are available, but our personal experiences in teaching said students is that they are often unaware these supports exist and are already high functioning independently. If they do not see it as a barrier to pursuing an education in STEM, then we'd urge educators at every stage (from early education to elementary and middle school, through to high school and beyond) to give such students the same attention, encouragement and mentoring they would any other and not discount them prematurely.

So, we would encourage parents, educators, and even students themselves to consider all demographic groups as potential STEM achievers, regardless of social class, or group identity. The many scientific, technological and engineering challenges facing the world today and in future decades will be best served by a diverse and inclusive STEM workforce bringing to bear their differing experiences and perspectives. While we have made great progress toward increasing STEM diversity, we have so very far yet to go.

¹In order to show respect for the diverse communities we write about in this article, we use the terms *autism spectrum* and *autistic*, rather than *autism spectrum disorder* as specified in the DSM-5 diagnostic manual. This is to ensure we respect preferences of the autism community and autistic advocates.

References

- Alexander, J.M., Johnson, K.E., Kelley, K., 2012. Longitudinal analysis of the relations between opportunities to learn about science and the development of interests related to science. *Sci. Educ.* 96, 763–786. <https://doi.org/10.1002/sce.21018>.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., Wong, B., 2012. Science aspirations, capital, and family habitus: how families shape children's engagement and identification with science. *Am. Educ. Res. J.* 49, 881–908.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., 2001. The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. <https://doi.org/10.1023/a:1005653411471>.
- Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>.
- Bravata, D.M., Watts, S.A., Keefer, A.L., Madhusudhan, D.K., Taylor, K.T., Clark, D.M., Nelson, R.S., Cokley, K.O., Hagg, H.K., 2019. Prevalence, predictors, and treatment of impostor syndrome: a systematic review. *J. Gen. Intern. Med.* 35, 1–24. <https://doi.org/10.1007/s11606-019-05364-1>.
- Canning, E.A., Muenks, K., Green, D.J., Murphy, M.C., 2019. STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Sci. Adv.* 5, eaau4734. <https://doi.org/10.1126/sciadv.aau4734>.
- Catsambis, S., 1995. Gender, race, ethnicity, and science education in the middle grades. *J. Res. Sci. Teach.* 32, 243–257.
- Chawla, D.S., 2020. Surviving weed-out classes in science may be a state of mind. *N. Y. Times*. Nov 16. Retrieved from: <https://www.nytimes.com/2020/11/16/science/weed-out-classes-stem.html>.
- Cohen, G.L., Sherman, D.K., 2014. The psychology of change: self-affirmation and social psychological intervention. *Annu. Rev. Psychol.* 65, 333–371. <https://doi.org/10.1146/annurev-psych-010213-115137>.
- Cohen, G.L., Garcia, J., Apfel, N., Master, A., 2006. Reducing the racial achievement gap: a social-psychological intervention. *Science* 313, 1307–1310.
- DeWitt, J., Archer, L., Osborne, J., Dillon, J., Willis, B., Wong, B., 2011. High aspirations but low progression: the science aspirations-careers paradox amongst minority ethnic students. *Int. J. Sci. Math. Educ.* 9, 243–271. <https://doi.org/10.1007/s10763-010-9245-0>.
- Dinham, S., 2017. The lack of an evidence bases for teaching and learning: fads, myths, legends, ideology and wishful thinking. *Prof. Voice* 11, 17–25.
- Doyle, R.A., Voyer, D., Cherney, I.D., 2012. The relation between childhood spatial activities and spatial abilities in adulthood. *J. Appl. Dev. Psychol.* 33, 112–120. <https://doi.org/10.1016/j.appdev.2012.01.002>.
- DuPaul, G.J., Weyandt, L.L., O'Dell, S.M., Varejao, M., 2009. College students with ADHD: current status and future directions. *J. Atten. Disord.* 13, 234–250.
- Eccles, J.S., 2013. Gender and achievement choices. In: Gershoff, E.T., Mistry, R.S., Crosby, D. (Eds.), *Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy*. Oxford University Press, New York.
- Ferguson, S.L., Lezotte, S.M., 2020. Exploring the state of science stereotypes: systematic review and meta-analysis of the Draw-A-Scientist checklist. *Sch. Sci. Math.* 120, 55–65. <https://doi.org/10.1111/ssm.12382>.
- Ferrara, K., Hirsh-Pasek, K., Newcombe, N.S., Golinkoff, R.M., Lam, W.S., 2011. Block talk: spatial language during block play. *Mind Brain Educ.* 5, 143–151. <https://doi.org/10.1111/j.1751-228X.2011.01122.x>.
- Geena Davis Institute on Gender in Media, 2018. The Scully Effect: I want to believe. In: STEM. Mount Saint Mary's University, Los Angeles, CA. Available from: <https://seejane.org/wp-content/uploads/x-files-scully-effect-report-geena-davis-institute.pdf>.
- Gerde, H.K., Pikus, A.E., Lee, K., Van Egeren, L.A., Huber, M.S.Q., 2021. Head Start children's science experiences in the home and community. *Early Child. Res. Q.* 54, 179–193.
- Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A., Mellott, D.S., 2002. A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychol. Rev.* 109, 3–25. <https://doi.org/10.1037/0033-295x.109.1.3>.
- Guiso, L., Monte, F., Sapienza, P., Zingales, L., 2008. Culture, gender, and math. *Science* 320, 1164–1165. <https://doi.org/10.1126/science.1154094>.
- Halpern, D.F., Eliot, L., Bigler, R.S., Fabes, R.A., Hanish, L.D., Hyde, J.S., Liben, L.S., Martin, C.L., 2011. The pseudoscience of single-sex schooling. *Science* 333, 1706–1707. <https://doi.org/10.1126/science.1205031>.
- Halpern, D.F., 2011. Sex Differences in Cognitive Abilities. Erlbaum, Mahwah, NJ.
- Harackiewicz, J.M., Canning, E.A., Tibbetts, Y., Priniski, S.J., Hyde, J.S., 2016. Closing achievement gaps with a utility-value intervention: disentangling race and social class. *J. Pers. Soc. Psychol.* 111, 745–765. <https://doi.org/10.1037/pspp0000075>.
- Hedges, L.V., Nowell, A., 1995. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. <https://doi.org/10.1126/science.7604277>.
- Hernstein, R.J., Murray, C., 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. The Free Press, New York.
- Herrera, F.A., Hurtado, S., 2011. Maintaining Initial Interests: Developing Science, Technology, Engineering, and Mathematics (STEM) Career Aspirations Among Underrepresented Racial Minority Students. Association for Educational Research Annual Meeting, New Orleans, LA. Retrieved from: <http://www.heri.ucla.edu/nih/downloads/AERA%202011%20-%20Herrera%20and%20Hurtado%20-%20Maintaining%20Initial%20Interests.pdf>.
- Jacobs, J.E., Bleeker, M.M., 2004. Girls' and boys' developing interests in math and science: do parents matter? *N. Dir. Child Adolesc. Dev.* 2004, 5–21. <https://doi.org/10.1002/cd.113>.
- Jirout, J.J., Newcombe, N.S., 2015. Building blocks for developing spatial skills evidence from a large, representative US sample. *Psychol. Sci.* 26, 302–310. <https://doi.org/10.1177/0956797614563338>.
- Jones, M.G., Howe, A., Rua, M.J., 2000. Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Sci. Educ.* 84, 180–192. [https://doi.org/10.1002/\(SICI\)1098-237X\(200003\)84:2<180::AID-SCE3>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<180::AID-SCE3>3.0.CO;2-X).
- Kincheloe, J.L., Steinberg, S.R., Gresson III, A.D., 1997. *Measured Lies: The Bell Curve Examined*. St. Martin's Press.
- Kost-Smith, L.E., Pollock, S.J., Finkelstein, N.D., Cohen, G.L., Ito, T.A., Miyake, A., Rebello, N.S., Engelhardt, P.V., Singh, C., 2012. Replicating a self-affirmation intervention to address gender differences: successes and challenges. *AIP Conf. Proc.* 231–234. Retrieved from: <https://aip.scitation.org/doi/pdf/10.1063/1.3680037>.
- Leaper, C., Gleason, J.B., 1996. The relationship of play activity and gender to parent and child sex-typed communication. *Int. J. Behav. Dev.* 19, 689–703.
- Levine, S.C., Vasilyeva, M., Lourenco, S.F., Newcombe, N.S., Huttenlocher, J., 2005. Socioeconomic status modifies the sex difference in spatial skill. *Psychol. Sci.* 16, 841–845. <https://doi.org/10.1111/j.1467-9280.2005.01623.x>.
- Liben, L.S., Coyle, E.F., 2014. Developmental interventions to address the STEM gender gap: exploring intended and unintended consequences. In: Liben, L.S., Bigler, R.S. (Eds.), *The Role of Gender in Educational Contexts and Outcomes*. Academic Press, San Diego, CA.
- Lillard, A.S., 2015. The development of play. In: Liben, L.S., Mueller, U. (Eds.), *Handbook of Child Psychology and Developmental Science*.
- Lowrie, T., Logan, T., Ramful, A., 2017. Visuospatial training improves elementary students' mathematics performance. *Br. J. Educ. Psychol.* 87, 170–186. <https://doi.org/10.1111/bjep.12142>.
- Martin, C.L., Ruble, D.N., 2010. Patterns of gender development. *Annu. Rev. Psychol.* 61, 353–381. <https://doi.org/10.1146/annurev.psych.093008.100511>.
- McKown, C., Weinstein, R.S., 2008. Teacher expectations, classroom context, and the achievement gap. *J. Sch. Psychol.* 46, 235–261.
- Miller, D.I., Halpern, D.F., 2013. Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learn. Individ. Differ.* 26, 141–152. <https://doi.org/10.1016/j.lindif.2012.03.012>.
- Miller, P.H., Slawinski Blessing, J., Schwartz, S., 2006. Gender differences in high-school students' views about science. *Int. J. Sci. Educ.* 28, 363–381.

- Miller, D.I., Nolla, K.M., Eagly, A.H., Uttal, D.H., 2018. The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US Draw-A-Scientist studies. *Child Dev.* 89, 1943–1955. <https://doi.org/10.1111/cdev.13039>.
- Mix, K.S., Levine, S.C., Cheng, Y.-L., Stockton, J.D., Bower, C., 2020. Effects of spatial training on mathematics in first and sixth grade children. *J. Educ. Psychol.* 113, 304–314. <https://doi.org/10.1037/edu0000494>.
- Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., Ito, T.A., 2010. Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science* 330, 1234–1237. <https://doi.org/10.1126/science.1195996>.
- Nelson, J.M., Gregg, N., 2012. Depression and anxiety among transitioning adolescents and college students with ADHD, dyslexia, or comorbid ADHD/dyslexia. *J. Atten. Disord.* 16, 244–254.
- Newcombe, N.S., Frick, A., 2010. Early education for spatial intelligence: why, what, and how. *Mind Brain Educ.* 4, 102–111. <https://doi.org/10.1111/j.1751-228X.2010.01089.x>.
- Nguyen, H.-H.D., Ryan, A.M., 2008. Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *J. Appl. Psychol.* 93, 1314–1334. <https://doi.org/10.1037/a0012702>.
- Nosek, B.A., Banaji, M.R., Greenwald, A.G., 2002. Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. <https://doi.org/10.1037/0022-3514.83.1.44>.
- Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N.M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., 2009. National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10593–10597. <https://doi.org/10.1073/pnas.0809921106>.
- Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Montell, J.M., Huguet, P., 2010. Individual differences in working memory moderate stereotype-threat effects. *Psychol. Sci.* 21, 1646–1648.
- Reilly, D., Neumann, D.L., Andrews, G., 2015. Sex differences in mathematics and science: a meta-analysis of National Assessment of Educational Progress Assessments. *J. Educ. Psychol.* 107, 645–662. <https://doi.org/10.1037/edu0000012>.
- Reilly, D., Neumann, D.L., Andrews, G., 2017a. Gender differences in spatial ability: implications for STEM education and approaches to reducing the gender gap for parents and educators. In: Khine, M.S. (Ed.), *Visual-Spatial Ability: Transforming Research Into Practice*. Springer International, Switzerland.
- Reilly, D., Neumann, D.L., Andrews, G., 2017b. Investigating gender differences in mathematics and science: results from the 2011 trends in mathematics and science survey. *Res. Sci. Educ.* 1–26. <https://doi.org/10.1007/s11165-017-9630-6>.
- Reilly, D., 2012. Gender, culture and sex-typed cognitive abilities. *PLoS One* 7, e39904. <https://doi.org/10.1371/journal.pone.0039904>.
- Rivardo, M.G., Rhodes, M.E., Camaione, T.C., Legg, J.M., 2011. Stereotype threat leads to reduction in number of math problems women attempt. *N. Am. J. Psychol.* 13, 5–16.
- Rochmes, J., 2018. The Power of Seeing Potential: Teachers' Beliefs About Overcoming Social Disadvantage and Student Achievement. Center for Education Policy Analysis: Stanford University, Stanford, CA. Available from: https://cepa-aws.gse.stanford.edu/sites/default/files/JRochmes_TeachersBeliefs-StudentOutcomes-CEPA_Dec2017.pdf.
- Rosenthal, R., Jacobson, L., 1968. *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development*. Holt, Rinehart & Winston, New York.
- Ruble, D.N., Martin, C.L., Berenbaum, S.A., 2006. Gender development. In: Eisenberg, N., Damon, W., Lerner, R.M. (Eds.), *Handbook of Child Psychology*, sixth ed. Wiley, Hoboken, NJ.
- Sansone, D., Carpenter, C.S., 2020. Turing's children: representation of sexual minorities in STEM. *PLoS One* 15, e0241596. <https://doi.org/10.1371/journal.pone.0241596>.
- Shapiro, J., Williams, A.M., 2012. The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex. Roles* 66, 175–183. <https://doi.org/10.1007/s11199-011-0051-0>.
- Shmulsky, S., Gobbo, K., Bower, M.W., 2019. STEM faculty experience teaching students with autism. *J. STEM Teach. Educ.* 53, 43–59.
- Sorby, S., Veirink, N., Streiner, S., 2018. Does spatial skills instruction improve STEM outcomes? The answer is “yes”. *Learn. Individ. Differ.* 67, 209–222. <https://doi.org/10.1016/j.lindif.2018.09.001>.
- Spencer, S.J., Steele, C.M., Quinn, D.M., 1999. Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. <https://doi.org/10.1006/jesp.1998.1373>.
- Steele, C.M., Aronson, J., 1995. Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>.
- Steele, C.M., 1997. A threat in the air: how stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. <https://doi.org/10.1037/0003-066X.52.6.613>.
- Steele, C.M., 1998. Stereotyping and its threat are real. *Am. Psychol.* 53, 680–681. <https://doi.org/10.1037/0003-066X.53.6.680>.
- Stieff, M., Dixon, B.L., Ryu, M., Kumi, B.C., Hegarty, M., 2014. Strategy training eliminates sex differences in spatial problem solving in a STEM domain. *J. Educ. Psychol.* 106, 390–402. <https://doi.org/10.1037/a0034823>.
- Uttal, D.H., Cohen, C.A., 2012. Spatial thinking and STEM education: when, why, and how?. In: *Psychology of Learning and Motivation*. Elsevier.
- Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., Newcombe, N.S., 2013. The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139, 352–402. <https://doi.org/10.1037/a0028446>.
- Voyer, D., Voyer, S., 2014. Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. <https://doi.org/10.1037/a0036620>.
- Wai, J., Lubinski, D., Benbow, C.P., 2009. Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. <https://doi.org/10.1037/a0016127>.
- Walton, G.M., Cohen, G.L., 2007. A question of belonging: race, social fit, and achievement. *J. Pers. Soc. Psychol.* 92, 82–96. <https://doi.org/10.1037/0022-3514.92.1.82>.
- Walton, G.M., Cohen, G.L., 2011. A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451.
- Wei, X., Jennifer, W.Y., Shattuck, P., McCracken, M., Blackorby, J., 2013. Science, technology, engineering, and mathematics (STEM) participation among college students with an autism spectrum disorder. *J. Autism Dev. Disord.* 43, 1539–1546. <https://doi.org/10.1007/s10803-012-1700-z>.
- Weinstein, R.S., Marshall, H.H., Sharp, L., Botkin, M., 1987. Pygmalion and the student: age and classroom differences in children's awareness of teacher expectations. *Child Dev.* 58, 1079–1093. <https://doi.org/10.2307/1130548>.
- Xie, Y., Fang, M., Shauman, K., 2015. STEM education. *Annu. Rev. Sociol.* 41, 331–357. <https://doi.org/10.1146/annurev-soc-071312-145659>.
- Yeager, D.S., Dweck, C.S., 2020. What can be learned from growth mindset controversies? *Am. Psychol.* 75, 1269–1284. <https://doi.org/10.1037/amp0000794>.
- Yeager, D.S., Romero, C., Paunesku, D., Hulleman, C.S., Schneider, B., Hinojosa, C., Lee, H.Y., O'Brien, J., Flint, K., Roberts, A., 2016. Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. *J. Educ. Psychol.* 108, 374–391. <https://doi.org/10.1037/edu0000098>.

Relevant websites

- A Biannual Report on the Representation of Women and Minorities Participating in the STEM Workforce. Each Report Shows Both the Progress Made, and How Far Remains Across Specific Scientific Disciplines.
- American Psychological Association's Guide to Stereotype Threat, www.apa.org/research/action/stereotype. Permanently archived at <https://web.archive.org/web/7/https://www.apa.org/research/action/stereotype>.
- National Science Foundation's Women, Minorities and Persons with Disabilities in Science and Engineering Biannual Report, HYPERLINK “

Gender and STEM career aspiration

Jingoo Kang, School of Applied Educational Science and Teacher Education, University of Eastern Finland, Joensuu, Finland

© 2023 Elsevier Ltd. All rights reserved.

Background	344
Relationships between abilities, interest, self-efficacy and STEM aspirations	344
Effect of perceived characteristics of STEM careers and lifestyle preferences on a career decision	345
Summary and implications	347
References	348

Background

In the contemporary world, since science and technology have become a foundation of overall industries, careers requiring STEM (science, technology, engineering, and mathematic) skills are becoming more abundant at the job markets (National Science Foundation, 2014). Given that the single most important factor that students consider in developing career aspirations is chances or size of future job markets, it was expected to have enough students who want to pursue STEM studies at college as the market grows (Sithole et al., 2017). However, ironically, in the last two decades, students, especially in secondary education have indicated a negative trend in their interest in science and STEM careers especially in many western developed countries (Kang et al., 2019a; OECD, 2016), and consequently, there is a lack of college students who pursue STEM majors such as physics, engineering, or computer science requiring advanced mathematic skills (Giffi et al., 2015). Interestingly, these are the majors that are dominated by males at higher education for the past twenty years (Diekman et al., 2010; National Science Foundation, 2014; Stoet and Geary, 2018). Thus, this lack of STEM workforce problem might be addressed with gender perspectives to improve the situation since this lack of women's participation slows national economic growth as well as gender equity (National Academy of Sciences et al., 2007).

So, why do female students try to avoid learning science and mathematical studies compared to males at school? In the past, some researchers suggested biological differences between female and male especially in brain structure; that is, males naturally have better mathematical abilities than females that make differences in career choices between gender in STEM fields (Ceci et al., 2009, 2014). However, recent national and international large-scale studies have indicated negligible gender differences in mathematics and science achievements; moreover, some countries, like Finland, indicate that girls are continuing to outperform boys in science (OECD, 2016). Indeed, girls are not only performing better in science but also other school subjects than boys at school (Stoet and Geary, 2018; Ullah and Ullah, 2019). Interestingly, however, girls have constantly indicated lower interest and self-efficacy in science and mathematics than boys (OECD, 2016; Stoet and Geary, 2018) that may impede them to pursue science careers. This seems to defy logic, given that interest and self-efficacy in science are known as strong indicators of achievement (Kang and Keinonen, 2017). Thus, there must be more complex interrelationships between achievement, interest, and self-efficacy in explaining gender differences in STEM career aspirations. Accordingly, in this review paper, we will discuss gender differences in abilities, interest, and self-efficacy, and how these constructs are interconnected with STEM career aspirations differently by gender based on recently published articles.

Another important factor related to the gender issue in STEM seems to lie in student perceptions of science careers (Archer et al., 2013) and their lifestyle preferences (Kang et al., 2019a). According to Archer et al. (2014) for instance, students still perceived the image of scientists as male, white, nerdy or geeky that are not favorable for females. Besides, student (mis)perceptions of lifestyle or working environment of STEM careers have indicated a negative effect on female students' pursuit of STEM careers. While females put much value on communal goals in their work and life than males, science was perceived as being uncommunal (Boucher et al., 2017); females in many cultures have more responsibilities in taking care of children so they prefer to work with high flexibility, while STEM careers are known to be spending much more time to work than other fields (Ferriman et al., 2009; Frome et al., 2006). Therefore, how to perceive such careers might affect students' study and career decision making processes stronger than the reality of the careers and the actual working environment, and these perceptions are highly related to cultural environments. Therefore, we will also discuss how these perceptions have affected students' STEM career choices and how they are interrelated with cultural aspects.

Relationships between abilities, interest, self-efficacy and STEM aspirations

In the last few decades, it has been reported that females did not select STEM-related studies and careers due to their lack of biological abilities in science and mathematics compared to males (Ceci et al., 2009). However, recent international large-scale studies such as Program for International Student Assessment (PISA) have reported that while female students indicated higher performance in science in many participating countries, males indicated higher self-efficacy and they pursue STEM careers more than females (OECD, 2016). Of course, within the same gender group, the one indicating a higher achievement in science presents higher

STEM aspiration, but when comparing different gender groups, this pattern is not revealed. Concerning this, [Stoet and Geary \(2018\)](#) found an interesting pattern from the PISA 2015 dataset that females indicated higher scores in Science and Reading than males, but the reading score gap between gender was higher than the science score gap. That is, although females outperform in science over boys, their strength in reading is relatively greater than in science. Accordingly, these relative personal abilities in other subjects may lead females to choose diverse options “in which they can gain the most benefit from their relative strength in reading” (p. 588). On the other hand, Science was usually the top strength of males comparing to other subjects and this led males to indicate higher self-efficacy than females. In the end, high abilities in other areas such as reading and low self-efficacy in science draw females away from STEM careers. Interestingly, this pattern is more and clearly presented in high gender-equal nations such as northern European countries like Finland (p. 591). Similarly, [Wang et al. \(2013\)](#) reported that, although high mathematical skills are a precursor of STEM career aspiration in general, individuals who had both high verbal and math skills are less likely to pursue STEM occupations than those who possessed only high math skills. Interestingly, there are more females in the group of high verbal and math skills but more males in the group of high math skills. However, males indicated higher self-concept in their math abilities than females and this self-concept in math influenced on STEM career choices for both females and males. Again, this study confirms that, females who have higher abilities in diverse subjects than males are likely to have more career options to choose and to possess lower self-efficacy in mathematics so that females may gravitate away more from STEM careers than males.

In addition, [Cardador et al. \(2020\)](#) argued that not only relative abilities but also relative or surplus interests in other areas may take women away from STEM careers. In their surplus model, they controlled abilities and examined the effect of two coexisting interests in STEM and other surplus areas such as Enterprising or Artistic interests. Then, they compared males and females to investigate whether these surplus interests moderated the relationships between gender and STEM interest. According to the results, females with high STEM interest chose disproportionately lower STEM careers as their surplus interest in other areas increased. Therefore, they concluded that “female adolescents with strong STEM-related interest should be less likely to pursue STEM careers when they also have strong interests in other areas, due to wider career options” (p. 1).

Surprisingly, a similar pattern was recently reported in Finland from very young children aged 8 to 10. [Oppermann et al. \(2021\)](#) conducted a longitudinal study for second and third graders in primary school concerning their interest and self-efficacy (as motivation) in science, mathematics, and Finnish language. According to the results, first, they found three motivational profiles as high motivation for all subjects, high motivation only in math, and low motivation for all subjects groups. Second, the high math motivation group indicated the highest STEM aspirations than the other two groups, and boys were overrepresented in this group. In addition, the stability of boys in the math motivation group was higher than girls between second to third grades, and “odds of students who remained in the math-motivated profile from second to third grade to name STEM aspirations were six times greater than for students who remained in the high motivation (for all subjects) profile” (p. 10). Thus, the results show that STEM careers are mostly aspired by students with single motivation in mathematics compared to students with multiple or surplus motivations probably due to the increased options for the latter group, and this pattern is already visible from elementary school students.

On the other hands, some studies show that science interest and self-efficacy affect career aspiration differently by gender. [Kang et al. \(2019b\)](#) conducted a longitudinal study between age 13 and 16 in Finland concerning causal effects of interest and self-efficacy on STEM career aspirations. According to the results, first, they found that self-concept stayed stably during the 3 years while interest increased significantly for both female and male students. Second, while both interest and self-efficacy influenced male student STEM aspirations, only interest was revealed as a strong predictor of the female student STEM aspirations while the effect of self-efficacy was negligible for females. Also, [Sax et al. \(2015\)](#) found that “the salience of MSC (math self-concept) in predicting STEM major selection has generally become weaker over time for women (but not for men). Ultimately, this suggests that women’s lower math confidence has become a less powerful explanation for their underrepresentation in STEM fields” (p. 813). Overall, these studies show that there are different effects of interest and self-efficacy on STEM aspirations by gender, and females may consider their internal satisfaction more than their perceived abilities when they aspire STEM careers.

Effect of perceived characteristics of STEM careers and lifestyle preferences on a career decision

Perceived characteristics of STEM careers including images of science and scientists and individual lifestyle preferences such as family-oriented are known to be strong predictors of STEM aspirations ([Archer et al., 2014](#); [Ferriman et al., 2009](#); [Kang et al., 2019a](#)). As adolescent’s career aspirations are influenced by contextual factors through socialization processes, the perceptions of specific careers and preferred lifestyle are also affected and formed by socio-cultural environments that individuals are surrounded ([McDaniel, 2016](#)). During the last decades, from career counseling and science education studies, it has been widely studied about the preferred working environment of men and women, and much research reported that while men prefer more things-oriented works such as making and developing a new product, women look for more people-oriented occupations aiming for communal goals such as helping, working and spending time with other people ([Diekman et al., 2011](#)). Also another study showed that men are more career-oriented and agentic, while women seem to be more holistic and attendant to family and the social well-being ([Ferriman et al., 2009](#)). [Diekman et al. \(2017\)](#) explain it under a goal congruity framework describing that a key factor affecting individual career pathways is the perceived alignment between social roles and career goals. However, since STEM careers have been (mis)perceived as being antithetical to communal goals and more things-oriented occupations, especially engineering and technology disciplines, women were likely to be deterred from STEM careers as well as communally-oriented men ([Diekman et al., 2010](#)) although STEM careers hold the key to supporting people and society in reality. Interestingly, college student choices of

not only STEM (Fuesting et al., 2017), but also Non-STEM majors are influenced by the perception of each major whether they are helpful to people (Ganley et al., 2018).

The development of these perceptions and preferences of STEM careers seems like to start from very early ages and continuously proceed through adolescents. According to Bian et al. (2017), children at ages 6 and 7 already indicated gender stereotypes such as *boys are smarter than girls*; these prejudices may influence children's interest in certain fields like STEM that are also perceived as the fields for smart people; consequently, it is "likely to narrow the range of careers they (girls) will one day contemplate" (p. 391). However, until at age 8, children draw roughly the same number of male and female scientists (Miller et al., 2018) although boys indicated higher self-concept and intrinsic value in mathematics than girls in this period (Oppermann et al., 2021). Thus, while these young children think that science and mathematics are more for boys, they do not have any bias on scientists' gender yet. Then, children begin to draw more male scientists during upper primary and lower secondary school, and by the time when they are in upper secondary school, the proportion of male scientists in their drawing becomes four times more than female scientists (Miller et al., 2018). Thus, it is plausible to assume that the earlier image of science for boys is transferred to the image of a scientist as males in the later phase. According to Archer and her colleagues, however, young students' views of scientist improved from 10 to 14 (Archer et al., 2013). That is, when they were in primary school, students perceived science as being hard and scientists as geeks, so girls did not want to have such careers that were unfavorable for them. However, when they got into the secondary school, students expressed much more positive views of scientists and their works for society. In addition, their follow-up study showed that student STEM career aspirations established at age 10 were fairly being stable until age 18 (Archer et al., 2020). Thus, they concluded that "low levels of science aspirations were not a consequence of ... the effect of negative views about scientists" (p. 5). That is, student STEM aspirations were being stable from age 10 to 18 while student views of scientists had improved; thus, there may be no significant relationships between aspirations and perceptions of STEM careers. However, what they have missed is that the views of scientists were already affected students' interest in STEM careers especially for young girls at age 10 as they perceived scientists as being "male", "smart", or "odd" which decreased their aspirations in STEM fields significantly at that moment. Since this low STEM aspiration at age 10 continued until they became age 18, it is hard to say that the negative views about scientists do not influence on STEM aspirations in future. Rather, it is better to conclude that the negative views (or prejudices) about scientists at an early age have a strong impact on their later STEM aspirations at age 18.

Perceived working environments is one of the critical factors affecting career decision and it is related to individual lifestyle preferences and goals such as having and caring children which is one of the most important concerns for women. Traditionally, in many cultures, girls have been expected to be a primary caregiver of their children and family in future, so it has affected girls to consider such careers perceived to be more flexible for the sake of family (Frome et al., 2006) and it leads them to be interested in a greater number of different careers than boys (Mendez and Crawford, 2002). In practice, unfortunately, since many STEM-related careers (usually male-dominated) do not (or are perceived not to) offer flexible working environments, girls and women who accept their role as the primary caretaker are not able to follow their career plans despite their abilities in STEM fields (Frome et al., 2006). According to Weisgram and Diekmann (2017), STEM occupations were perceived as incongruity with family goals, and this perception emerged more strongly with older age cohort. Also, Ferriman et al. (2009) reported that women in STEM desired more flexible, weekends free, and less working hours when they became parents while men did not show any significant changes before and after having children. Thus, as stated by Frome et al. (2006), desire for a family-flexible occupation and high time demands of occupation become strong predictors of women changing their aspirations out of STEM fields. Overall, this traditional gendered role particularly women as a primary caregiver of children may be related negatively to the perceptions (or realities) of STEM working environment so that there are few numbers of women in STEM occupations.

In a recent study, the similar pattern was also revealed from young students aged 13 in Finland as careers perceived as giving enough time for family indicated a significant relationship with their STEM interest for both girls and boys (Kang et al., 2019a). However, the results were surprising since girls' personal time orientation was positively connected to STEM aspirations. That is, Finnish girls perceived STEM careers as flexible in terms of time management so that they thought that they can combine STEM careers with family responsibilities. Given that Finland is well known as its high gender equal society and flexible working environments, this gives a good insight as to how the national and cultural environments can affect student STEM interest and future career perspectives. However, although expecting to work in STEM fields during adolescence plays a key role to lead them to the eventual STEM career (Legewie and DiPrete, 2014), since young adolescents' STEM aspirations are continuously changing during and after the secondary school as well as college, this positive perception of STEM careers as family friendly should be followed up carefully in Finland. For instance, Frome et al. (2006) argued that girls at 12th grade who were interested in STEM fields, decided to stop pursuing STEM careers at age 25 if they perceived the STEM environments as not family friendly while studying at college.

Also, at college, one of the raised gender issues in STEM seems to line in, so called, masculine and chilly climate "as fields with larger gender differences tend to be the ones that are perceived as more stereotypically masculine" (Ganley et al., 2018, p. 457) such as engineering or computer science. One major result or effect of this misperceived or practiced masculine culture at college is to attain and retain unqualified men over more qualified-women since "low-achieving men are both retained in and attracted to these fields during college" more than high-achieving women (Cimpian et al., 2020, p. 1319). Ganley et al. (2018) thus argued that the lower proportions of women in STEM fields are attributed to college student perceptions of such majors as being unwelcoming to women. They also showed that if women perceived such college majors do not discriminate against them, they may "choose majors that lead to higher paying careers" such as STEM (p. 478). Bloodhart et al. (2020) also showed that while women outperformed men in physical and life science at college, their abilities were undervalued by peers.

This “chilly climate” toward women in STEM fields has been well documented since the 1980s (Britton, 2017), but recent studies show that the situation has been improving gradually in both academy and society. In order to examine whether chilly climate affects women’s decision on switching majors from male-dominated fields, Riegle-Crumb et al. (2016) investigated whether women who entered male-dominated fields such as computer science or engineering were more likely to change their majors compared to women who were in other college majors. Surprisingly, according to the results, there were no differences between women of male-dominated fields and other general fields in terms of the switching decision. Specifically, 59% of women in the male-dominated fields stayed in their majors compared to 51% of women in other majors. Thus, they concluded that women’s decision to switch was not affected by the number of males in the specific majors, but more by the current “higher labor market returns” and changing social expectations toward “women to act as wage-earners in support of their family” (p. 445). Thus, this study shows the importance of increasing numbers of female students entering a college to study STEM, since once they begin, women are able to resist “chilly climate” faced and continue to complete it similar to women in other majors. Accordingly, they recommended that “focusing on understanding the psychological, social, and interactional processes that function early in the life course to deter females from ever considering or entering such fields, and subsequently designing effective programs and interventions at earlier stages, is likely to be most effective at promoting women’s presence in nontraditional fields” (p. 446). Concerning “chilly climate” at universities, William and Ceci also argued that (1) levels of job satisfaction of female STEM professors were similar to males’ as most of them were “somewhat to very” satisfied (Ceci and Williams, 2011), and (2) faculty of both math-intensive such as engineering and non-math-intensive fields such as biology preferred more women candidates than men (Williams and Ceci, 2015). Thus, they suggested that “it is an auspicious time to be a talented woman launching a STEM tenure-track academic career” (p. 5365).

Summary and implications

As mentioned, the gender issues in STEM are highly related to the lack of future STEM workforces that might affect each country’s economic and societal growth. Accordingly, in this short review paper, we have discussed the issues focusing on factors affecting women’s STEM career aspirations under two broad categories. However, there are still many other topics related to the gender disparities uncovered such as salary gaps between gender (Xu, 2015), academy cultures such as hiring opportunities (Eaton et al., 2020), or levels of gender equality of countries (Stoet and Geary, 2018). Also, Eaton et al. (2020) argued that gender issues in STEM must be addressed with racial perspectives since race also plays a critical role in STEM aspirations that are related to societal structures. However, considering the societal aspects of different countries, the issue becomes much more complex and inconclusive. As McDaniel (2010) argued, in countries where there is more gender equity, more women are expected to be incorporated into higher education. Thus, it is plausible to expect that countries with high levels with gender equality may indicate low gender gaps in terms of STEM aspirations. However, Sikora and Pokropek (2012) and Stoet and Geary (2018) found that gender gaps in students’ aspirations toward traditionally male-dominated fields were revealed more from higher gender-equality societies, like Northern European countries. Sikora and Pokropek (2012) explained this phenomenon based on self-expressive ideologies where advanced industrialized countries hold more than developing countries. On the other hand, Stoet and Geary (2018) described the results based on life-quality pressures as more gender-equal countries tend to have higher levels of social security while less gender-equal countries tend to possess lesser. Consequently, the girls in less gender-equal societies are attracted to focus more on science at school to pursue STEM careers that offer a relatively higher salary. Therefore, it might be important to take actual societal opportunity structures in different societies into account in examining the gender and STEM aspiration issues since the societal structures contribute to shaping career expectations. In addition, what makes it difficult to analyze the gender discrepancy in STEM is that “there are careers that are not STEM by definition, although they often require STEM skills” (Stoet and Geary, 2018, p. 592) such as health care industries including nursing and medicine where women make up the majority. It may describe partly concerning where women with strong STEM skills are located.

Thus, these multiple contributors and their complex interconnections make it a “wicked problem” and hard to give a one-shot solution on the gender issue in STEM fields. However, what many studies argue and agree upon the issue is the importance of early interventions to give proper impressions and information about STEM careers. For this, although special programs for girls may be able to be arranged out of school, it might be more effective to integrate those practices into everyday school science lessons. For instance, hands-on science activities at school are known to be a strong predictor of STEM aspirations both for girls and boys (Kang and Keinonen, 2017) but usually boys lead the inquiry processes actively while girls follow passively (Wieselmann et al., 2020), that might be one of the starting points that girls as well as boys perceive STEM as masculine. Therefore, teachers especially early childhood educators should remind the importance and long-term effects of science practices at school in terms of the gender disparity and should offer equal chances for both girls and boys to be an active participant and leader of the inquiry processes. By doing so, the teachers can give an impression to the young learners as science are gender-neutral and it might be the starting point to redress the gender equity issue in terms of STEM aspirations. Also, as argued by Stoet and Geary (2018) there are many girls who are talented in and enjoy science and mathematics but do not choose the STEM pathway. Thus, when we try to plan special campaigns and programs to increase women’s participation in STEM, it might be effective to target those girls who are obviously being lost but have high potential to return to the STEM pathway rather than targeting other girls who have strengths in another or surplus areas.

References

- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., Wong, B., 2013. "Not girly, not sexy, not glamorous": primary school girls' and parents' constructions of science aspirations. *Pedagog. Cult. Soc.* 21 (1), 171–194.
- Archer, L., DeWitt, J., Dillon, J., 2014. "It didn't really change my opinion": exploring what works, what doesn't and why in a school science, technology, engineering and mathematics careers intervention. *Res. Sci. Technol. Educ.* 32 (1), 35–55.
- Archer, L., Moote, J., Macleod, E., Francis, B., DeWitt, J., 2020. ASPIRES 2: Young People's Science and Career Aspirations, Age 10–19.
- Bian, L., Leslie, S.J., Cimpian, A., 2017. Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 355 (6323), 389–391.
- Bloodhart, B., Balgopal, M.M., Casper, A.M.A., Sample McMeeking, L.B., Fischer, E.V., 2020. Outperforming yet undervalued: undergraduate women in STEM. *PLoS One* 15 (6), e0234685.
- Boucher, K.L., Fuesting, M.A., Diekman, A.B., Murphy, M.C., 2017. Can I work with and help others in this field? How communal goals influence interest and participation in STEM fields. *Front. Psychol.* 8, 901.
- Britton, D.M., 2017. Beyond the chilly climate: the salience of gender in women's academic careers. *Gen. Soc.* 31 (1), 5–27.
- Cardador, M.T., Damian, R.I., Wiegand, J.P., 2020. Does more mean less? Interest surplus and the gender gap in STEM careers. *J. Career Assess.* <https://doi.org/10.1177/1069072720930658>.
- Ceci, S.J., Williams, W.M., 2011. Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U. S. A.* 108 (8), 3157–3162.
- Ceci, S.J., Williams, W.M., Barnett, S.M., 2009. Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135 (2), 218.
- Ceci, S.J., Ginther, D.K., Kahn, S., Williams, W.M., 2014. Women in academic science: a changing landscape. *Psychol. Sci. Publ. Interest* 15 (3), 75–141.
- Cimpian, J.R., Kim, T.H., McDermott, Z.T., 2020. Understanding persistent gender gaps in STEM. *Science* 368 (6497), 1317–1319.
- Diekman, A.B., Brown, E.R., Johnston, A.M., Clark, E.K., 2010. Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21 (8), 1051–1057.
- Diekman, A.B., Clark, E.K., Johnston, A.M., Brown, E.R., Steinberg, M., 2011. Malleability in communal goals and beliefs influences attraction to stem careers: evidence for a goal congruity perspective. *J. Pers. Soc. Psychol.* 101 (5), 902.
- Diekman, A.B., Steinberg, M., Brown, E.R., Belanger, A.L., Clark, E.K., 2017. A goal congruity model of role entry, engagement, and exit: understanding communal goal processes in STEM gender gaps. *Pers. Soc. Psychol. Rev.* 21 (2), 142–175.
- Eaton, A.A., Saunders, J.F., Jacobson, R.K., West, K., 2020. How gender and race stereotypes impact the advancement of scholars in STEM: professors' biased evaluations of physics and biology post-doctoral candidates. *Sex. Roles* 82 (3), 127–141.
- Ferriman, K., Lubinski, D., Benbow, C.P., 2009. Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: developmental changes and gender differences during emerging adulthood and parenthood. *J. Pers. Soc. Psychol.* 97 (3), 517.
- Frome, P.M., Alfeld, C.J., Eccles, J.S., Barber, B.L., 2006. Why don't they want a male-dominated job? An investigation of young women who changed their occupational aspirations. *Educ. Res. Eval.* 12 (4), 359–372.
- Fuesting, M.A., Diekman, A.B., Hudiburgh, L., 2017. From classroom to career: the unique role of communal processes in predicting interest in STEM careers. *Soc. Psychol. Educ.* 20 (4), 875–896.
- Ganley, C.M., George, C.E., Cimpian, J.R., Makowski, M.B., 2018. Gender equity in college majors: looking beyond the STEM/non-STEM dichotomy for answers regarding female participation. *Am. Educ. Res. J.* 55 (3), 453–487.
- Giffi, C., Dollar, B., Drew, M., McNelly, J., Carrick, G., Gangula, B., 2015. The Skills Gap in U.S. Manufacturing 2015 and Beyond. Retrieved from: <http://www2.deloitte.com/content/dam/Deloitte/us/Documents/manufacturing/us-pip-the-manufacturing-institute-and-deloitte-skills-gap-in-manufacturing-study.pdf>.
- Kang, J., Keinonen, T., 2017. The effect of inquiry-based learning experiences on adolescents' science-related career aspiration in the Finnish context. *Int. J. Sci. Educ.* 39 (12), 1669–1689.
- Kang, J., Hense, J., Scheersoi, A., Keinonen, T., 2019a. Gender study on the relationships between science interest and future career perspectives. *Int. J. Sci. Educ.* 41 (1), 80–101.
- Kang, J., Keinonen, T., Salonen, A., 2019b. Role of interest and self-concept in predicting science aspirations: gender study. *Res. Sci. Educ.* 1–23.
- Legewie, J., DiPrete, T.A., 2014. Pathways to science and engineering bachelor's degrees for men and women. *Sociol. Sci.* 1, 41–48.
- McDaniel, A., 2010. Cross-national gender gaps in educational expectations: the influence of national-level gender ideology and educational systems. *Comp. Educ. Rev.* 54 (1), 27–50.
- McDaniel, A., 2016. The role of cultural contexts in explaining cross-national gender gaps in STEM expectations. *Eur. Sociol. Rev.* 32 (1), 122–133.
- Mendez, L.M.R., Crawford, K.M., 2002. Gender-role stereotyping and career aspirations: a comparison of gifted early adolescent boys and girls. *J. Second. Gift. Educ.* 8 (3), 96–107.
- Miller, D.I., Nolla, K.M., Eagly, A.H., Uttal, D.H., 2018. The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US Draw-A-Scientist studies. *Child Dev.* 89 (6), 1943–1955.
- National Academy of Sciences, National Academy of Engineering, Institute of Medicine of the National Academies, 2007. *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. National Academies Press, Washington, DC.
- National Science Foundation, 2014. *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2013 (Special Report NSF 13-304)*. Retrieved from: <http://www.nsf.gov/statistics/wmpd/>.
- OECD, 2016. *PISA 2015 Results (Volume I). Excellence and Equity in Education*. OECD Publishing, Paris.
- Oppermann, E., Vinni-Laakso, J., Juuti, K., Loukomies, A., Salmela-Aro, K., 2021. Elementary school students' motivational profiles across Finnish language, mathematics and science: longitudinal trajectories, gender differences and STEM aspirations. *Contemp. Educ. Psychol.* 101927.
- Riegle-Crumb, C., King, B., Moore, C., 2016. Do they stay or do they go? The switching decisions of individuals who enter gender atypical college majors. *Sex. Roles* 74 (9–10), 436–449.
- Sax, L.J., Kanny, M.A., Riggers-Piehl, T.A., Whang, H., Paulson, L.N., 2015. "But I'm not good at math": the changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Res. High. Educ.* 56 (8), 813–842.
- Sikora, J., Pokropek, A., 2012. Gender segregation of adolescent science career plans in 50 countries. *Sci. Educ.* 96 (2), 234–264.
- Sithole, A., Chiyaka, E.T., McCarthy, P., Mupinga, D.M., Bucklein, B.K., Kibirige, J., 2017. Student attraction, persistence and retention in STEM programs: successes and continuing challenges. *High Educ. Stud.* 7 (1), 46–59.
- Stoet, G., Geary, D.C., 2018. The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychol. Sci.* 29 (4), 581–593.
- Ullah, R., Ullah, H., 2019. Boys versus girls' educational performance: empirical evidences from global north and global south. *Afr. Educ. Res. J.* 7 (4), 163–167.
- Wang, M.T., Eccles, J.S., Kenny, S., 2013. Not lack of ability but more choice: individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24 (5), 770–775.
- Weisgram, E.S., Diekman, A.B., 2017. Making STEM "family friendly": the impact of perceiving science careers as family-compatible. *Soc. Sci.* 6 (2), 61.
- Wieselmann, J.R., Dare, E.A., Ring Whalen, E.A., Roehrig, G.H., 2020. "I just do what the boys tell me": exploring small group student interactions in an integrated STEM unit. *J. Res. Sci. Teach.* 57 (1), 112–144.
- Williams, W.M., Ceci, S.J., 2015. National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proc. Natl. Acad. Sci. U. S. A.* 112 (17), 5360–5365.
- Xu, Y., 2015. Focusing on women in STEM: a longitudinal examination of gender-based earning gap of college graduates. *J. High Educ.* 86 (4), 489–523.

Learning to balance indigenous and exogenous knowledge systems for environmental decision-making in the Kumaon Himalayas

Sameer Honward, University at Buffalo, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	349
“Our land our life” and the educational context in the Kumaon Himalayas	349
Why is indigenous knowledge important to the educational system	350
What is the existing indigenous knowledge in the Kumaon Himalayas	351
Results of the Delphi	351
Question 1 on the Delphi	351
Question 2 on the Delphi	351
Question 3 on the Delphi	352
Shortcomings of the Delphi panel	352
Use of IK by youth for environmental decision-making	353
Coding the interviews and the focus group discussions	353
Results	354
Water management (IK)	354
Land management (IK)	354
Focus groups (water)	354
Interviews (water)	355
Focus groups (land)	355
Interviews (land)	355
Other related results	355
Discussion and conclusion	355
References	357

Introduction

Indigenous people around the world have struggled to keep their unique approaches to life and traditional knowledge systems intact with the advent of industrialization and, more recently, globalization (Agrawal, 2005). They have adapted and blended their values, beliefs and practices to subsist under a dominant mainstream society (Rogers, 2003). Most of this knowledge has its roots deep in the places they inhabit, and thus is extremely localized. Ethnographers and other social scientists have started to recognize that people in the “non-Western” world have their own science and practices grounded in this knowledge base (Sillitoe et al., 1998). Although indigenous knowledge systems are often very different from the widely accepted western knowledge systems, it has been observed that they play a crucial role in local-level decision-making within areas of agriculture, education, resource management, health care, and many other community-based activities (Kellner and Bosch, 2002).

This chapter discusses how youth in rural communities residing in the middle Himalayas use indigenous knowledge to support environmental decisions while addressing water and land use related concerns. The chapter serves to enrich our understanding of community decision-making, especially as connected to land use and water issues, but also helps us understand how youth in village communities in the Himalayan bioregion negotiate a balance between indigenous and non-indigenous knowledge. The chapter also sheds light on the decision making process situated within the context of the environmental education curriculum “our land our life” (the current environment education curriculum in public schools). The chapter discusses three different aspects of how local communities in Kumaon region negotiate indigenous and non-indigenous knowledge.

1. Identify indigenous knowledge in the Kumaon Himalayas;
2. Find out how much indigenous knowledge youth use for environmental decision making;
3. Discuss relevance of merging knowledge systems to the existing environmental education curriculum “our land our life”.

“Our land our life” and the educational context in the Kumaon Himalayas

With the advent of industrialization and colonization, formal education in most developing nations became a force to produce clerks, administrators and managers. This knowledge system based on western values did not hold the indigenous knowledge system in high regard. Western influxes thus created a myth that indigenous knowledge is unscientific and backward (Gupta,

2007), and the system of transferring indigenous knowledge from one generation to another began to fall apart. As a result, formal education became more classroom-focused, and soon the gap between community and schools began to increase (Niraula, 2007).

This phenomenon of the increasing gap between community and schools exists in the Kumaon region of the Himalayas. In order to help bridge this gap, the Uttarakhand Seva Nidhi (a local NGO) introduced an environment education curriculum into the school system. The curriculum is called *Our Land Our Life* (OLOL). OLOL focuses on local issues, and the embedded pedagogy within the curriculum tries to address the concerns of the rural people in the Kumaon Himalayas. The course currently runs in all the state government schools since the state government of Uttaranchal adopted this curriculum as the official environment education curriculum. The course is implemented from fifth grade to ninth grade (5 yrs) and tries to link people's livelihood issues as related to land, water, fodder, crops, trees and other ecological elements to formal education received in schools.

Even though this existing curriculum links local issues to formal schooling, the gap between communities and school still exists and although OLOL focuses on local environmental issues and discusses these issues within the context of local empowerment, it is the only part of the entire school curriculum that focuses on local issues. Therefore, in the Kumaon region, as well as in other parts of the rural Himalayas, in spite of OLOL being part of the official curriculum, the school is often perceived as an agent that introduces the youth to outside knowledge that is very different from the existing culture and tradition in the region (Jackson, 2003). School education is viewed as a vehicle that provides the youth with resources to help them make a life outside of their village community (Shiva, 2000). The youth who go through a formal education system lack the knowledge about traditional practices that feed into a system that provides sustainable livelihood to the village communities (Goonatilake, 2001). As Pande (2001, p. 48) points out, "In their haste to run away from the village, the young men and women do not seem to have the time to understand their own village and their own people, neither do they receive any orientation toward this in school." For example, in Maichun Village in the Kumaon region, *Palta* was a community activity that involved the entire village community coming together and making compost for their agricultural fields. The practice not only strengthened community bonds but also provided quality fertilizer for agriculture. Jackson (2004, p. 96) observes that "The young youth in the village do not see compost as a resource for sustainable agriculture. In fact, they are ashamed of working on the land: the girls for esthetic reasons (*nail paint would be spoiled* and the *compost stinks*-were some instant remarks from girls) and the boys for livelihood (*what will we do in the village? We go to the city, earn money and live comfortably*-the boys say). Several families in the village now complain of declining agricultural yields, so much so that "food is not even enough for 6 months in a year." The example clearly points out that since the indigenous knowledge is not passed on to the next generation through the intergenerational knowledge transfer system and since most of the formal educational system does not focus on traditional practices, the sustainable livelihood in the village is impacted. Pande (2001, p. 51) also adds, "These impacts were too small to be noticed in the village in the early stages and when they became apparent and obvious for everyone to notice them, it requires resources, the time, and knowledge to regenerate or improve them-a task that nobody in the village can do alone." Therefore, it is important to not only document the existing indigenous knowledge but also in looking at how the villagers might beneficially reincorporate this knowledge into their lives.

Why is indigenous knowledge important to the educational system

Indigenous knowledge (IK) can be defined as the process through which the natives of an area have built a relationship with their natural environment (Semali and Kincheloe, 1999). Thus, indigenous knowledge is developed over time and tied to a certain community. The community can be urban, rural, nomadic, or tribal (IIRR, 1996). Indigenous knowledge (IK) is dynamic, as each generation chooses to adapt the knowledge to suit their needs and values (Agrawal, 2005). But indigenous knowledge is unique to the place and thus is tied to the context. The IK system is often tied to the sustainable livelihood of the people. Generally, the system supports environmental sustainability because it represents generations of observations, analysis, and experimentation. IK provides a powerful problem-solving mechanism for local communities (Tella, 2007).

Understanding IK in traditional communities is fundamental to the design of any learning environment in the Kumaon region of the Himalayas (IIRR, 1996). Learning environments have to take into consideration several aspects of community. These include the community of the classroom, the school, and the connections between the school and the larger community, including the home. The importance of out-of-school learning becomes clear when one examines the relatively small amount of time spent in school compared to other settings. Activities in homes, community centers, and after-school clubs can have important effects on students' academic achievement (Bell et al., 2006). In the Kumaon region, learning also takes place when youth work on family farms, take cattle out for grazing and do landscape related work. Thus the effective instruction begins with taking into account what learners bring to the setting; this includes cultural practices and beliefs as well as knowledge of academic content. There are many studies that have shown that what people learn and how people learn is context-dependent (Mertl et al., 2007). Therefore, a learning environment that takes into consideration the context should help learners link ideas from ecology and formal science to their own lives (Burford et al., 2005). Finally, an educational initiative on sustainability that embraces IK would help learners negotiate different worldviews and value systems about development and livelihood (Palmer, 1998).

What is the existing indigenous knowledge in the Kumoan Himalayas

The issue of identifying knowledge types (indigenous, traditional ecological knowledge (TEK) or exogenous) is a complex process. There is very rarely a single instance that can be one knowledge type or another. In most instances there is an overlap. Semali and Kincheloe (1999) address this issue of complexity by pointing out that “Indigenous knowledge is an ambiguous topic that immediately places analysts on a dangerous terrain. Not only are scholars unsure what we’re talking about but many analysts are uncertain who should be talking about it.” Thus the first step is to ensure that there is consensus about the “definition” of indigenous knowledge/traditional ecological knowledge. Using the Delphi technique we identified what people in the region consider indigenous knowledge that can help with land and water resource management. The Delphi technique was used to facilitate information, opinions and judgments from a panel of experts to gain consensus on the issue of what Indigenous knowledge can be used to manage land and water resources (Dunham, 1996). The reason to focus on the above two issues was made because water—and land-related issues are important issues that are directly related to the livelihood and the quality of life of people in the region which relate to indigenous practices and maintain relevance today. Also water and land related issues have been identified by prior work in the region as key concerns, both by government and nonprofit workers, and by local inhabitants. Water touches every aspect of human life and, if not managed well, can create hardships. Similarly good land management practices are important so as to provide the people with food and shelter. By focusing on the types of knowledge people use to manage land and water, we were able to make suggestions toward designing better environmental education programs for helping communities in the region better their quality of life.

Thirteen participants/experts were identified and were requested to participate in this Delphi process. The participation was voluntary and no compensation was offered. The participants/experts were chosen by consulting with the officials of the Uttarakhand Seva Nidhi Paryavaran Shikshan Sanstha (Uttarakhand Environment Education Center—UEEC). The criteria for the selection of the participants were that the individuals would have experience working with the local communities on social, agricultural, economic and related environmental issues.

The Delphi procedure was carried out in three rounds. The first round the participants were asked three sets of questions. These questions were asked in interview format in Hindi and in English and were recorded on video. The questions were as follows:

1. What is your definition of indigenous knowledge?
2. What are some practices that use indigenous knowledge for water management? What is the specific indigenous knowledge used within these practices?
3. What are some practices that use indigenous knowledge for land management (as related mostly to forests and agriculture)? What is the specific indigenous knowledge used within these practices?

Results of the Delphi

The answers of all of the 13 participants were conceptually similar but keeping in mind that there was consensus for answer sheet/participant number 1, the answers given by this participant were taken as the result for research question number 1 and as the concluding result for the Delphi.

The results on the consensus building exercise (Delphi) was as follows:

Question 1 on the Delphi

What is your definition of indigenous knowledge?

Indigenous knowledge is knowledge that comes to you from your ancestors and that has evolved over time. It is knowledge that is attached to a certain place. Also there is a difference between traditional knowledge and traditional practices. Sometime when we follow a practice without understanding the logic behind the practice that is traditional practice, the actual logic behind the practice is traditional knowledge.

Question 2 on the Delphi

What are some practices that use indigenous knowledge for water management? What is the specific indigenous knowledge used within these practices?

1. Water is the basic necessity for life. The example of how the community traditionally manages water is “Naulas” (ground water springs). There is traditional knowledge associated with how to build a Naula. The structure is very important. The structure is such that it helps the ground water come to the surface and then stay above the surface.
2. The second thing that traditional knowledge does to manage water is keep the water sources clean. Often times the water sources are designated as religious and thus people have to go in them barefoot, which helps in keeping the source clean.
3. Collecting rainwater is another way we can manage water traditionally.

4. People used to plant native species (Oak and Deodhar) to make sure that there is maximum seepage of water into the soil, as planting these trees helps hold the soil together, thus leading to more water seepage into the ground and increasing the ground water table.
5. Digging trenches (khals) is another way to manage the water.
6. People also build small dams (choys) to stop flowing water
7. To keep the soil together and to help stop water runoff people around here build depressions (guls) around their farms.
8. There are traditional devices (Earthen pots) to store water and keep it clean.
9. There is encouragement to use running water rather than stored water because the logic is that running water has a natural filtration system

Question 3 on the Delphi

What are some practices that use indigenous knowledge for land management (as related mostly to forests and agriculture)? What is the specific indigenous knowledge used within these practices?

To keep the land healthy people use the following traditional practices:

1. Using organic compost.
2. Seed saving is traditionally considered good land management practice.
3. Mixed farming and crop rotation practices keep the land healthy.
4. Healthy forest keeps the farms healthy by providing more organic litter for decomposition, to make good compost.
5. Not cutting an entire tree for fodder. Cutting the lower branches of the tree keeps the tree from destruction and the more healthy the trees, the healthier the soil and better yield from the land.
6. Sharing labor for agriculture is a good land management practice as there is community participation and thus makes sure everyone keeps their piece of land healthy and bad practices do not spread.
7. Use organic pesticides and insecticides such as walnut leaves and neem leaves.

Thus the first part of the study helped identify and document the IK in the region.

Shortcomings of the Delphi panel

Although the Delphi panel identified indigenous knowledge and practices, we do not believe that they have identified all of the practices that are indigenous to the region. An example of this is the practice of community discussions to resolve issues faced by the community as a whole. The Delphi panel has not identified this practice as indigenous. But village communities are often times known to get together to resolve an issue faced by their community. The government of India in an attempt to restore traditional practices in village communities encourages *Panchayat Raj*, a form of local government that involves community discussions and resolving issues at community level. While this specific form of community discussion forum (Panchayat Raj) may not exist in the Kumaon region, other forms of community discussion forums may exist within the region. The issue is thus nuanced and complex. The practice of using formal community discussions to resolve issues may have existed in the past (before the British rule). It is documented that community discussion forums or *village sabhas* were part of society until 600 BCE (Mathew, 2000). After this, the subcontinent was broken up and different rulers/kings governed different parts (Mathew, 2000). During the British rule (which lasted about 150 yrs) all forms of local problem solving mechanisms were dismantled (Mathew, 2000). It was only after India regained independence that the local village governing systems were put back into practice. Thus the local community discussion forums exist in a different form than the ones that existed in pre British India, which brings up the issue of how indigenous practices evolve over time. The issue of the evolution of indigenous knowledge will be discussed later in the paper. For the purposes of this study, we have identified complex practices such as community discussions as a non-indigenous practice, not only because the Delphi panel did not identify it as indigenous knowledge or practice but also because the practice itself does not exist in its original form. There is documentation (Shiva, 2000) about how the caste system has hijacked the practice of community discussion forums, where upper caste members have an upper hand in decision-making processes. Thus since it does not exist in its true indigenous form, I have identified it as a non-indigenous practice.

Also some of the practices that the panel has identified as indigenous can also be identified as a part of the modern conservation movement that has been based on western ecosystem science. One prominent example is that of the practice of planting trees. The Delphi panel identified planting trees as an indigenous practice. This practice can be seen advocated in the modern conservation movement. The practice stemmed from the excessive deforestation that took place for developmental purposes. As the industrialization took priority, deforestation took place at a rate that was unsustainable and thus it gave rise to the need of tree planting (Govinda and Diwan, 2003). However, this is also an indigenous practice and is suggested to be a part of the sacred grove concept, where communities planted trees to appease the gods (Bhagwat et al., 2005). Thus for the purposes of this study, we have identified this particular practice of planting trees as indigenous knowledge.

We recognize that the issue of “what is indigenous and what is not” is far more complicated than the scope of this particular study, but the Delphi panel does provide a means to construct a reasonable list of indigenous practices within the region.

Use of IK by youth for environmental decision-making

The participants for the study were recruited from two villages east of the city of Almora. The first village was Chanoli and the second village was Maichun. The youth were recruited from two different high schools. The first school was Panvanaula high school and inter-college and the second school was the Garudabanj high school and inter-college (In India, high school runs through 10th grade, while inter college includes grades 11 and 12). Both these schools are run by the state government. Youth from Chanoli and Maichun attend the Panvanaula high school and inter college. The youth participation in the study was voluntary. Overall a total of 29 youth participated in this phase of the study. There were three main techniques used to identify youth's knowledge of the IK and the youth's use of IK for decision-making.

These three techniques are as follows:

- Focus groups: The student groups (5–10) students were identified with the help of the teachers.
- Role-play activity: a role-play activity was designed for the students in the focus groups. By having the students assume certain roles of stakeholders involved in the local decision-making processes, the students' ideas of how these processes function and what types of information is used was assessed.
- Semi-structured interviews: Semi-structured interviews provided useful information on how much IK they know and how they would use IK for potential decision making processes.

Coding the interviews and the focus group discussions

A total of 29 interviews were conducted with students from both the schools (13 from Garudabanj and 16 from Panvanaula). We also conducted three focus groups with the youth in the Panvanaula school and 4 focus groups with the youth in the Garudabanj school. Once all of the interviews and focus groups were transcribed every utterance was coded for the use of indigenous knowledge as identified by the Delphi.

The main categories of Indigenous knowledge for which every utterance was coded were as follows:

Traditional/indigenous knowledge with respect to water as found out in the results of the Delphi:

1. Use of *Naulas* (ground water springs)
2. Keeping water source clean.
3. Rain water harvesting.
4. Plant native species of trees (Oak and Deodhar)
5. Digging trenches (khals)
6. Build small dams –(choys)
7. Build depressions – (guls around their farms)
8. Use traditional devices to store water and keep it clean.
9. Use running water rather than stored water

Traditional/indigenous knowledge with respect to land management as found out in the results of the Delphi:

1. Using organic compost.
2. Seed saving is traditionally considered good land management practice.
3. Mixed farming and crop rotation practices.
4. Healthy forest keeps the farms healthy
5. Not cutting the entire tree. Cutting the lower branches.
6. Sharing labor for agriculture.
7. Use organic pesticides and insecticides such as walnut leaves and neem leaves.

The following is an example of the complete coding system with the main and the sub categories:

Participant: If the tap water runs out we go to the village leaders or government official and get them to fix the problem and in the mean time use the water from the naulas.

This utterance was coded for the main traditional knowledge category as **Use of naulas** and the utterance was further coded as **aware but do not think of as a viable solution to resolve the problem – 2**. The reason why this utterance was coded as 2 is because although the participant is using the naula, he/she feels like the solution to resolving their problem is that the government should fix the tap. A category 1 **Aware and think of as a viable solution** utterance is as follows *"We should use the naula water and take care of the naula"* and a category 3 utterance is *"We use the naula and the tap water as the tap water is convenient but we cannot use it for drinking, we prefer the naula water for drinking."*

After coding for traditional/indigenous knowledge use, The data was coded for other solutions. These solutions were called exogenous or non-traditional solutions. There were several of these other/exogenous/non traditional solutions that were offered by the youth. The main categories of exogenous/nontraditional solutions that the youth offered to resolve water problems in the village were as follows:

1. Educate the community members
2. Sit together for a meeting and collectively solve the problem.

3. Since youth study these issues in school adults need to listen to them.
4. Manage cattle grazing.
5. Petition the government/village chief to resolve the water issues.
6. Stop corruption
7. Build hand pumps and wells.
8. Maintain and repair existing pipelines.
9. Distribute/Ration water.

Since all the participants who mentioned these solutions thought of them as viable solutions, the utterances were sub-coded as 1- **Aware and think of as viable solution** and 4- **Did not mention at all**.

There were not many non-traditional solutions presented by participants to resolve some of the land management issues. Some of the non-traditional solutions presented by the youth to resolve the land related problems were:

1. Stop pollution
2. Human population control
3. Stop migration toward the city.

Results

Overall the youth indicated that the following solutions were important to address water and land related issues.

Water management (IK)

The number one solution that the students present in terms of managing water using indigenous knowledge is tree planting. In the focus groups (100%) and the interviews (83%), the students strongly stressed the need to plant native species of trees in order to hold the soil and increase water seepage into the ground water table.

Rain water harvesting, use of *naulas*, keeping water sources clean, and proper water storage were some other solutions that the students presented to solve the water issues in the community. Although aware of these traditional practices, not all of the students felt like these were viable solutions to resolve the problems related to water management.

Land management (IK)

The majority of the students during focus groups (86%) and interviews (83%) thought the practice of not cutting trees in their entirety is a solution to helping resolve land related environment problems. During focus group discussions, some students (14%) also thought that mixed farming, crop rotation, use of organic fertilizer, use of organic insecticide and seed saving are solutions to resolving land related issues. While during interviews, (7%) of students thought crop rotation and mixed farming were practices that lead to lesser land related problems. Also during interviews, (3%) of students thought using organic insecticide and fertilizer are good solutions to resolving land related problems faced by their community.

When it came to resolving land related issues, the students did not present many exogenous solutions to actively resolve the land related problems.

Focus groups (water)

During the focus groups the youth focused on following solutions for addressing water and land related issues:

The top two choices that youth mentioned as solutions to resolving water problems in their village are planting trees (100%) (native species, broad leaf trees –oak, deodhar, etc.) and keeping the water source clean (57.1%). Both of these are indigenous knowledge sources as indicated by the Delphi panel. Resolving problems through community discussions and meetings (57.1%) is also a top choice the students made to resolve water related problems faced by the community. This choice is based on outside knowledge as the Delphi panel did not identify it as indigenous knowledge. However, community meetings and problem solving is a known indigenous method in rural India (Shiva, 2000).

Also rain water harvesting (indigenous knowledge) and educating community members (outside knowledge) regarding water conservation and management issues were other choices made by 42.9% of students. 28.6% of students thought that using *naulas* (indigenous knowledge) and proper water storage devices (indigenous knowledge) was a good idea to solve some of the water related problems face by their community. This was followed by 14.3% of students suggesting that digging trenches (indigenous knowledge), adults listening to youth (exogenous knowledge), putting a halt to corruption (exogenous knowledge) and constructing new hand pumps and wells (exogenous knowledge) as ways to resolve water related problems faced by their communities.

Interviews (water)

The individual interviews also put forward a similar story as the focus groups. The top two choices that the students have indicated as practices to resolve water problems face by their community are planting trees (native species—broad leaf trees like oak and deodhar)—82.8% and keeping the water source clean (24.1%). Both of these are identified by the Delphi group as practices supported by indigenous/traditional knowledge. 20.7% of students think educating community members (outside knowledge/exogenous knowledge) is a good idea to resolve some of the water related problems in their community, and 13.8% of students think rain water harvesting (indigenous knowledge) is a good idea to resolve some of the water related problems faced by their communities. 10.3% of students think that proper water storage devices (indigenous knowledge), building new pipelines and repairing existing pipelines (exogenous knowledge) and fair distribution of water are good solutions to resolving water related problems faced by their communities. 6.9% of students think building small dams choys (indigenous knowledge) is a good idea while 3.4% of students think solving water related problems should include building hand pumps and wells (outside knowledge), asking the government to intervene (outside knowledge) and using *naulas* (indigenous knowledge). 6.9% of students think that their community should supplement the new methods such as pipes with traditional methods such as *naulas* to resolve some of the water related problems.

Focus groups (land)

As in the water related focus groups and interviews, the students used a combination of outside knowledge and indigenous knowledge to make their decisions with regards to issues presented to them in relation to the land in their community.

The top choice that the students have mentioned as solutions to resolving land related problems in their village are not cutting trees (85.7%) This is a practice based on indigenous knowledge sources as indicated by the Delphi panel. 14.3% of youth thought that keeping the forests healthy (indigenous knowledge), mixed farming (indigenous knowledge), seed saving (indigenous knowledge), organic insecticide use (indigenous knowledge) and crop rotation (indigenous knowledge) are good land management practices, and these should be used to resolve land related problems in their communities.

6.6% of students suggested that stopping pollution (exogenous knowledge) is a solution to resolving land related problems and 3.3% of students suggested population growth control (exogenous knowledge) as a solution to resolving land related problems face by their community. 14.3% of students thought mixing organic pesticides with new artificial pesticides was a solution to resolving some of the land related problems.

Interviews (land)

During individual interviews, 82.8% of students offered not cutting trees (indigenous knowledge) as a solution to resolving land related problems. The next best choice was to stop pollution (exogenous knowledge)—42.9%, followed by Crop rotation (indigenous knowledge)—14.3% and stopping out migration (exogenous knowledge)—14.3%.

6.9% of students suggested mixed farming (indigenous knowledge) and using organic insecticide (indigenous knowledge) as solutions to some of the land related problems. These were followed by 3.4% of students suggesting that keeping forests healthy (indigenous knowledge), using organic fertilizer (indigenous knowledge) and seed saving practices (indigenous knowledge) were good solutions to some of the land related problems in their community.

Other related results

In order to get the complete picture of how much knowledge is passed on to the youth and how much of this knowledge youth use to make current or potential decisions about the land and water situation in their village, it was also important to find out where is the source of all of this knowledge. Thus when asked the question “where did you learn about this specific practice?” 72.4% of students said that they learned all of their knowledge in the school, 13.8% identified the source of their knowledge as home, and 3.4% of students indicated that some of the knowledge they learned at home and some of it in school.

Discussion and conclusion

When asked about how they would go about solving water-related problems in their community, almost all of the students said that planting trees (native species) is the most important step. Their explanation for this was that the trees would hold the soil, which in turn will stop water runoff and increase groundwater tables so that the springs and rivers will have more water. This explanation corresponds to the explanation given for planting native species of trees in their environmental education curriculum *Our Land, Our Life*. The students also mentioned using *naulas*, harvesting rainwater, utilizing proper water storage and keeping water sources clean as other ways to resolve water-related issues. In terms of land related issues majority of the youth thought that not cutting trees in their entirety was the number one solution to resolve land-related issues. The explanation behind this solution was that cutting trees would lead to deforestation and thus will cause soil runoff, which would be harmful to the land. Other than the above, a few youth mentioned mixed farming, crop rotation, the use of organic insecticide, and seed-saving as solutions for resolving land-related problems in their community. Most students indicated that currently, they would use practices based on indigenous knowledge to resolve some of the water- and land-

related issues in their community. It also appears that the youth are aware and recommend certain practices based on indigenous knowledge to resolve issues related to water in their community and that they have learned about the majority of these practices in school. Based on their answers it also appears that the environment education curriculum "our land our life" is a major source of the knowledge that the youth are projecting when asked to solve water and land related environmental problems. This is reflected in the data where 72.4% of students mentioned that they learned most of these practices in school, and 13.8% of students indicated that they learned some of the practices at home and some at school.

However, the findings also indicate that even though the youth mention that most of the practices that they will use in resolving water- and land-related issues are traditional practices, it is not completely guaranteed that they would use them to resolve these issues in reality. This observation is supported by the data that indicates that when asked whether they prefer tap water or water from the naula, 95% of students said that they would prefer tap water. Jackson (2008) also addresses this issue of students not wanting to indulge in traditional practices, by pointing out that young people are more likely to choose urban/westernized ways of doing things when presented with a choice. Several families in the village now complain of declining agricultural yields, so much so that food is not even enough for 6 months in a year." This quote implies that the students do not want to indulge in the traditional practices.

It is not surprising that the students are indicating that they have learned about resolving water and land related issues from their curriculum *Our Land, Our Life*. *Our Land, Our Life* is a powerful curriculum developed by the Uttarakhand Seva Nidhi Paryavaran Shikshan Sanstha, a non-governmental organization working in the Kumaon ranges of the Himalayas. The curriculum tries to link classroom with actual environmental problems, identifying links between issues that reflect real-world situations and relating environmental education to the local community. The curriculum addresses not only environmental science but also environmental education for sustainability more generally. The curriculum was developed in a unique partnership with local villages and has a focus on indigenous knowledge. It may be possible that the curriculum *Our Land, Our Life* presents them with solutions that do not fit their reality. It also is possible that the youth gave the answers they have learned in school since the interviews were conducted in the school.

The question that arises is that how well the environmental education curriculum has fit the changing nature of the communities. The demographic of the village community has changed over time. The men in the villages often go to the urban centers to seek better economic opportunities (Pande, 2001). There is a need within these communities to urbanize, as that is recognized as being developed (Agrawal, 2005). Thus indigenous knowledge practices are often times adapted to the needs of development and urbanization. However, in the school, the youth are learning about indigenous knowledge systems and about how these knowledge systems are sustainable because they fit the context. So on one hand at home there is a push toward being more urbanized while on the other hand at school there is a push toward indigenous practices. This struggle shows in the data as the youth are not voicing the actual changes experienced within their communities.

There is also the issue of changing nature of indigenous knowledge systems. Along with changing communities the knowledge also changes. With immigration and emigration of knowledge the complexity surrounding what knowledge is indigenous for a particular generation remains a question that is yet to be answered. The Delphi panel has provided a starting point for the discussion of what is the indigenous knowledge in the region, and how indigenous is it? Although the study directly does not ask these questions, we feel they are questions that need to be addressed as communities that associate with this knowledge move in time and space. For example, the men in the village are oftentimes employed in the urban areas and interact with urban resources such as water taps. When they come back to the village, they seek these urban conveniences and try to adapt these conveniences to their context and culture. Thus in the process they add to the existing traditional knowledge systems. These adaptations are not always sustainable, but every once in while an adaptation such as a water tap next to a house becomes a sustainable practice. Then the question arises, when should we identify the water tap as indigenous. This is a complex question that is beyond the scope of this study, but like a restoration ecologist struggles with, how far back in time one should go in order to restore the ecosystem to its "native" state, similarly there is a struggle here as to how far back in time we should go in order to identify a knowledge practice as indigenous. The way indigenous knowledge evolves, and what constitutes indigenous knowledge, is a complex study that will require a different set of probes. But as we acknowledge the complexity of this issue, we would also like to state that we have tried to address the complexity of what is indigenous knowledge by gaining consensus from local experts. The Delphi panel does not address every single level of complexity of how these knowledge systems are constructed but provides a starting point and a good insight into the indigenous knowledge systems.

The school curriculum *Our Land, Our Life* needs to fit the changing face of knowledge. The curriculum needs to recognize that people move through time and space and knowledge itself will evolve over a period of time. If the curriculum hopes to revive the traditional knowledge systems within the youth without acknowledging the evolution of the knowledge base then we think there is a fair chance of it being rejected. Especially since the teachers (who come from an urban area) think of this curriculum as not "up to date" (field observations, 2005, 2006, 2007, 2008). By not recognizing the changing landscape of the indigenous knowledge systems, the curriculum is working against the intergenerational knowledge transfer. Especially since the knowledge the youth are getting at home is different from the knowledge they are learning in school.

The other relevant observation is that the youth were much more confident in presenting solutions to resolve water problems. When it came to answering questions related to resolving land issues, most youth seemed like they did not know much about how to resolve them, and thus used the information from their curriculum *Our Land, Our Life* to answer the question. This could be because most of the youth who go to school do not go into the fields with their parents (or their parents are not farmers). However, they use water daily and are much more aware of water-related issues than land-related issues.

Although *Our Land, Our Life* is a powerful curriculum that focuses on local indigenous practices, it needs to take into consideration the evolving nature of indigenous knowledge and the applicability of this knowledge. For example, newer ecological problems have arisen during the past 20 years. With problems such as climate change looming over the Himalayan communities, the curriculum needs to focus on adaptability rather than only indigenous knowledge. There is no indigenous knowledge that exists to deal with current environmental problems such as climate change. Thus the curriculum needs to keep up with the changing ecosystem to create solutions based on the evolution of indigenous knowledge.

In summary our data and observations suggest that the students are aware of the indigenous practices and some of the scientific reasons behind these practices but are less likely to use them in reality. There is evidence that youth are learning about different practices at the school and at home, but at this time they are unable to build bridges so as to carefully integrate the knowledge that they gain in all of these different settings. They seem to be thinking that the knowledge from home is not useful in school, and in school, they should only discuss things that they learn in the school curriculum.

It is important to bridge the gap between formal and informal learning as learners are constantly making sense of their environment in formal and as well as informal environments (Bransford et al., 2006). Thus, if the students do not connect in-school (formal) learning to what they learn at home (informal), then the students are going to find it difficult to construct scientifically sound and meaningful knowledge. The students' underdeveloped conceptual ecology with respect to environmental problem-solving solutions is certainly going to hamper the students' decision-making process. Thus, in order to close the gap between formal and informal learning and empower the students to make environmentally sound decisions about their local environment in the future, the designers of the curriculum *Our Land, Our Life* need to restructure the content of this curriculum. The curriculum itself needs to be realigned with the current environmental situation of the Kumaon Himalayas. This can be achieved by:

1. Evaluating the current environmental situation in the villages: This has to be done as problems such as global warming are starting to create severe water problems in the Himalayas, and these new environmental problems need new strategies of adaptation and survival.
2. Restructuring the curriculum *Our Land, Our Life* based on the environmental evaluation.
3. Conducting long-term ethnographic research with students to assess long-term learning and application of knowledge.
4. Preparing iterations of the curriculum based on these long-term studies.
5. Connecting the curriculum to other subjects and making it a true interdisciplinary curriculum. This would also involve organizing teacher trainings to have subject teachers collaborate with each other.

As the climate changes, the water and land resources in the Himalayas are changing at a rapid rate. There need to be strategies in place to help people in the Kumaon region cope with these changing times. The recommendations above are just a start, and there is plenty of more research and action needed to help the mountain environment and the communities that reside in the Kumaon Himalayas cope and adapt to the changing ecology of the Himalayas.

References

- Agrawal, A., 2005. Environmentalism: community, intimate government and the making of environmental subjects in the Kumaon, India. *Curr. Anthropol.* 46 (2), 1–23.
- Bell, P., Bricker, L.A., Lee, T.R., Reeve, S., Zimmerman, H.T., 2006. Understanding the cultural foundations of children's biological knowledge: insights from everyday cognition research. In: Barab, S.A., Hay, K.E., Hickey, D.T. (Eds.), *Proceedings of the Seventh International Conference of the Learning Sciences*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 1029–1035.
- Bhagwat, S.A., Kushalappa, C.G., Williams, P.H., Brown, N.D., 2005. A landscape approach to biodiversity conservation of sacred groves in the Western Ghats of India. *Conserv. Biol.* 19 (6), 1853–1862.
- Bransford, J., Vye, N., Stevens, R., Kuhl, P., Schwartz, D., Bell, P., Meltzoff, A., Barron, B., Pea, R., Reeves, B., Roschelle, J., Sabelli, N., 2006. Learning theories and education: toward a decade of synergy. In: Alexander, P., Winne, P. (Eds.), *Handbook of Educational Psychology*, second ed. Erlbaum, Mahwah, NJ, pp. 209–244.
- Burford, G., Ngila, L.O., Rafiki, Y., 2005. Education, indigenous knowledge and globalization. *S. Afr. J. Educ.* 5, 1–6.
- Dunham, R.B., 1996. The Delphi Technique. <http://www.scribd.com/doc/16939227/Delphi-Test>.
- Goonatilake, S., 2001. Anthropologizing Sri Lanka. In: *A Eurocentric Misadventure*. Indiana University Press, Indianapolis.
- Govinda, R., Diwan, R., 2003. *Community Participation and Empowerment in Primary Education*. Sage Publications, New Delhi; Thousand Oaks, CA.
- Gupta, A., 2007. *Going to School in South Asia*. Greenwood Press, Westport, CT.
- IIRR, 1996. *Recording and Using Indigenous Knowledge: A Manual*. International Institute for Rural Reconstruction (IIRR), Silang, Cavite, Philippines.
- Jackson, M., 2003. From practice to policy in environmental education. *S. Afr. J. Educ.* 20, 97–110.
- Jackson, M.G., 2004. Towards an authentic Indian environmentalism. *S. Afr. J. Educ.* 21, 94–107.
- Jackson, M.G., 2008. *Transformative Learning for a New Worldview: Learning to Think Differently*. Macmillan, New York.
- Kellner, K., Bosch, O.J.H., 2002. The value of local and indigenous knowledge for the development of information systems for conservation management. In: Lemons, J., Victor, R., Schaffer, D. (Eds.), *Conserving Biodiversity in Arid Regions – Best Practices in Developing Nations*. Kluwer Academic Publishers, Boston, pp. 493–513.
- Mathew, G., 2000. *Status of Panchayat Raj in States and Union Territories of India*. Concept Publishing, New Delhi, India.
- Mertl, V., McCarthy, L., Levas, S., Stevens, R., 2007. The Way She Does? No: Intergenerational Influences in Everyday Family Financial Matters. Paper Presented at The 12th Biennial European Association for Learning and Instruction (EARLI) Conference in Budapest, Hungary.
- Niraula, T., 2007. Schooling in Nepal. In: Gupta, A. (Ed.), *Going to School in South Asia*. Greenwood Publishing Group, Westport CT, pp. 126–142.
- Palmer, J.A., 1998. *Environmental Education in the 21st Century: Theory Practice Progress and Promise*. Routledge, New York, NY.
- Pande, A., 2001. Environmental education in rural central Himalayan schools. *J. Environ. Educ.* 32, 47–53.
- Rogers, E.M., 2003. *Diffusion of Innovation*, fifth ed. Free Press, New York.
- Semali, L., Kincheloe, J.L. (Eds.), 1999. *What is Indigenous Knowledge: Voices From the Academy*. Farmer Press, New York and London.
- Shiva, V., 2000. *Tomorrow's Biodiversity*. Thames & Hudson, London.
- Sillitoe, P., et al., 1998. The development of indigenous knowledge: a new applied anthropology [comments and reply]. *Curr. Anthropol.* 39 (2), 223–252.
- Tella, R., 2007. Towards promotion and dissemination of indigenous knowledge: a case of NIRD. *Int. Inf. Lib. Rev.* 39, 185–193.

STEM education and workforce development: the history, politics, and evidence

Hal Salzman^a and Daniel Douglas^b, ^a Rutgers University, Edward J. Bloustein School of Planning & Public Policy, John J. Heldrich Center for Workforce Development, New Brunswick, NJ, United States; and ^b Trinity College, Department of Sociology, Hartford, CT, United States

© 2023 Elsevier Ltd. All rights reserved.

The global nexus of STEM education and workforce development	358
Political history of STEM education and the workforce	359
What is STEM anyway? Conceptual and definitional issues	361
STEM demand and supply	362
Global STEM supply and domestic labor markets	363
Problems and problematics in the supply	365
Sub-baccalaureate STEM credentials	366
Mathematics coursework and the supply of STEM graduates	366
STEM attrition and STEM attraction	366
Conclusion	367
Acknowledgments	368
References	368
Relevant website	369

The global nexus of STEM education and workforce development

STEM workforce development is at once a nation-specific and a global enterprise. A nation's competitiveness and level of development are typically considered dependent on the strength of its science and technology capacity (e.g., [Romer, 1994](#)), and thus policy makers focus on various metrics about a nation's STEM workforces as indicative of its economic prowess. Similarly, the global migration of students and high-skilled workers has become an issue of national and international policy: "Receiving" countries encourage immigration into their colleges and STEM workforces, while "sending" nations view the loss of their potential workforce as a continuation of brain drain. However, some nations, such as India, have found global migration important for their outsourcing technology industries, or in the case of China, as a means of developing a high-skill workforce trained in cutting edge technology and science that can be recruited back to contribute to its own economic and technological development ([Lynn et al., 2012](#); [Lynn and Salzman, 2018](#); [Meil and Salzman, 2017](#); [Saxenian, 2006](#)).

Although STEM workforce development systems are integrated into a global network, they function differently around the world depending on the structure and quality of each nation's education system and the types of businesses in each country, along with national policies. Moreover, perspectives differ on whether each country's STEM system provides global or only national benefits. This debate has become a contentious public policy issue in the United States, where large global flows of STEM students and workers have become an integral component of its STEM workforce system. But too often these policy debates rest on simplistic notions about the role of STEM education and the STEM workforce, along with misconceptions about "national competitiveness" in these global systems. As a consequence, STEM and education policies often have the paradoxical effect of undermining a nation's, and the world's STEM workforce development and capacity.

Historically, nations benefitted from migration that contributed to their economic and social development and the STEM fields, in particular, benefited from selective global migration. In recent years, the creation of targeted and nonimmigrant education and work programs has again become a matter of policy and legislative debate. Scholars and lawmakers question whether some transnational flows of students and workers are counterproductive to the nation's indigenous STEM strength—either in loss of workers or in weakening labor markets in the receiving countries ([Borjas, 2016](#)) even if large supplies of guestworkers are profitable for individual firms. For example, the US Department of Labor calculates that the H-1B nonimmigrant temporary visa program, by supplying technology guestworkers at below market wages, creates a transfer payment from forgone worker wages to employers of \$3 billion to \$4 billion annually ([Employment and Training Administration, 2021](#)). As such, historical patterns of STEM education and workforce development may be changing, with different impacts from those of the past.

Further complicating simple formulations about the role of STEM workforce development in a nation's economy is the complex nature of global interdependencies in technology development. In the US, simplistic analyses have used metrics such as the number of engineering graduates to assess a country's presumed advantage or "competitiveness." The use of such indicators has led to dramatic proclamations: first, that the Soviet Union's large math and engineering cohorts would enable it to overtake US science and technology innovation; then a fear that Japan's prodigious engineer production would lead to its global dominance and, in the past decade, fears that China's large number of engineering graduates will imperil US competitiveness ([National Academy of Sciences et al., 2007](#); [Teitelbaum, 2014](#)). In these cases, the reports used questionable metrics and comparisons ([Berliner and Biddle, 1997](#); [Salzman and Lowell, 2008](#)) but, more importantly, they failed to consider the role of different STEM workforces in each

country's development. The number of engineers, for example, largely reflects the amount of construction in the country, however building roads, housing and factories is not considered an indicator of a nation's "global competitiveness" or innovation capacity (Lynn and Salzman, 2010; Kuehn and Salzman, 2018).

At the same time, a globally integrated STEM education system can provide strong global linkages in innovation networks, can facilitate the "sending" country's development through educating and training its workforce, and/or can contribute to global networks that enable outsourcing and offshoring. Such globalization can increase the fortunes of a firm while weakening the economy or welfare of its home country. Although the global STEM development system depends on industrialized countries' STEM education systems to supply globalized STEM enterprises, those "host" countries may not be the beneficiaries if firms globalize work or engage in global labor arbitrage to lower wages (Hira, 2008; Lynn and Salzman, 2018).

There is another challenge in aligning STEM education and workforce development: educational institutions vary greatly between and within countries, as do pedagogical practices. But STEM disciplines are assumed to largely, or at least conceptually, transcend national borders or politics. Thus, while education is a global enterprise, workforce development is largely a national initiative. As such, STEM education and workforce development have distinct, if interrelated purposes and are shaped by a complex set of national policies and global migration and economic activity. The focus of this chapter is on STEM education and workforce development in the United States in the context of global education migration and global enterprises that employ STEM graduates from US universities.

Political history of STEM education and the workforce

The need to educate and train students and workers to take up careers in fields now collectively referred to as STEM has been a long-standing concern of educators, development practitioners, analysts, and policymakers around the world. While STEM acronymically refers to Science, Technology, Engineering, and Mathematics, the meaning of STEM as a category of education, training, and occupations has evolved over time, as has the focus of related policymaking.

The STEM history as an acronym dates to circa 2001 when it was rearranged from its earlier formulation as "SMET" by the National Science Foundation (NSF), which was the successor to NSF's traditional classification, Science and Engineering (S&E). But STEM's pre-history begins shortly after the end of World War II, when science and engineering were recognized as vital to strengthening military capabilities. The destructive force of the Atomic Bomb was the most visible scientific achievement, but the second world war was fought with a vast number of scientific and technological advances, from radar and sonar to cryptography and operations management and logistics developed by the military's statistician "Whiz Kids". A new generation of military leaders and policy makers began to actively pursue an expanded role for government in scientific and technological development. This goal was immediately linked to a need to train a substantial supply of people who "understand the fundamental laws of nature and are skilled in the techniques of scientific research" (Bush, 1945; Zachary, 1999). And thus, in the mid-1950s, Congress asked the National Academies of Sciences to regularly conduct a census of the science and engineering workforce to have an ongoing assessment of the nation's science and technology capacity. Having sufficient numbers of scientists and engineers would be vital should there be another Manhattan project—which had developed the Atomic Bomb—and more generally necessary to advance the civilian and military development envisioned by the President's science adviser in the *Endless Frontier* (Bush, 1945).

From the beginning, the expansion of STEM supply was viewed as necessary not just for the military, but also for the private sector and advancing the nation's standard of living. Similarly, this period was the start of a perennial complaint that a wide range of the nation's ills—from military and economic weakness to an inability to compete with the technological prowess of, first the Soviet Union and later various Asian nations—came from the failure to produce enough graduates in mathematics, engineering and the sciences. In 1958, a five-part series in *Life* magazine compared the educational environments of two students, one in the US, the other in the Soviet Union, depicting the US school system as relatively lax and unfocused—particularly in the sciences. It appears this story was designed to raise Cold War fears, and it may not have been a faithful piece of journalism. Among a number of distortions, the Soviet student was likely a fiction of *Life* magazine (Bracey, 2009, p. 43), and the articles served to further its ultraconservative ideologue publisher Henry Luce's broader policy efforts to raise public fears about the Red Menace (Brinkley, 2010). Luce, along with other policymakers and corporate leaders, was battling President Eisenhower's efforts to reduce military spending and improve diplomatic relations with the Soviet Union. The attack on US education was part of the effort to shift US policy to confrontation with the Soviet Union by containing liberal arts education in the US that was, Luce argued, leading youth to question a hardline anti-Soviet attitude and toward liberal politics more generally (Bracey, 2009).

The most remembered, and misremembered event of the period was the successful launch of the *Sputnik* satellite in 1957 as part of the International Geophysical Year. Although the Soviet launch preceded by several months the American launch of a satellite that had a scientific mission, the political fallout lasted decades. Both nations had entered the International Geophysical Year (IGY) in 1957 with the jointly announced goal of each launching a satellite as global scientific and technological projects during IGY. The Soviets then sought the public relations coup of first-to-launch with an abbreviated scientific and technological mission. The US satellite that launched a few months later, the Jet Propulsion Lab's Explorer 1, had a cosmic ray detector that confirmed what became known as the Van Allen radiation belt. A month and a half after that, Vanguard I went into orbit to transmit upper atmosphere measurements, continuing to relay information for six years, and is still orbiting today. *Sputnik* was in orbit for three months and had only a beeping transmitter that sent a radio signal for a mere three weeks and provided no scientific information. Like much of STEM policy over the decades, the narrative of weakness and failure that

emerged from Sputnik had little correspondence to the evidence; Soviet satellite strength was transitory and overplayed, as was its technological and economic prowess more generally. Nonetheless, the STEM myths based on fear had a more important role than evidence in shaping US policies.

In 1958, Congressional education leaders, notably Alabama's Senator Lister Hill and Representative Carl Elliot were, after a decade of failed attempts, finally able to broker a political compromise to steer federal funds into education systems. Labeled the National Defense Education Act (NDEA), thrice-defeated federal education legislation was revived and, after a year of lobbying and public campaigns such as the *Life* magazine series, these Alabama congressmen navigated a difficult legislative path providing federal funds to schools that were controlled by states and local school boards without fear of federal regulation in matters such as desegregation following the 1955 *Brown v. Board of Education* Supreme Court ruling (Urban, 2010).

The NDEA provided funding for school buildings and programs in sciences and foreign languages. It provided federal support for higher education in the form of loans and fellowships with a priority to sciences and foreign languages, though a wide range of fields were supported as long as the college would administer loyalty oaths to its students; 166 colleges, however, publicly protested the provision, with 32 turning down all NDEA funding, including Harvard, Yale and Princeton (Clowse, 1982). Although the passage of NDEA is regarded as addressing a science and engineering shortage, the legislation came on the heels of the 1957 publication of a landmark study of *The Demand and Supply of Scientific Personnel* by National Bureau of Economic Research labor economists Blank and Stigler, who found no such shortage (Blank and Stigler, 1957).

Global politics have continued to be an integral part of the narrative about the importance of STEM and proposed legislative responses. As further history shows, global threats and claimed weaknesses are perennial, never solved, and conveniently raised to advance other political goals. It was thus, two decades later following a period of economic doldrums and declines in military spending and prestige that education once again took center stage in the nation's failures. The 1983 report entitled *A Nation at Risk* identified US schools as responsible for "the educational foundations of our society presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war" (National Commission on Excellence in Education, 1983). Perhaps not unintentionally, the report coincided with Reagan's highly controversial arms build-up through unprecedented peace-time deficit spending and funding for fantastical outer space technology development dubbed the "Star Wars" defense system, taking aim at the Soviet Union.

Once again, an educational failure report was followed by a serious study of the evidence that found the report's claims lacking. Most notably, the Department of Energy's Sandia Laboratories reexamination found the evidence weak and in error, and the prose hyperbolic (Carson et al., 1993). Nonetheless, the *Nation At Risk* report fueled proclamations of educational failure as the pivotal explanation for the economic declines of the previous decade (Freeman and Salzman, 2018; Weinstein, 1998) and added to the general fear of the Soviet Union.

In 2005, during the Bush administration's efforts to expand vouchers and charter schools to privatize education (Ravitch, 2020, 2014), and during a period of public doubt about the nation's military standing in the world, another commission rushed out a report that bore a title refashioned from Churchill's warning of *The Gathering Storm* about the Third Reich's ascendancy. The National Academies report, *Rising Above the Gathering Storm*, (RAGS), led by a defense contractor's retired CEO, advanced a similar argument to earlier reports—that US schools were failing in math and science relative to other countries, and that massive investment was needed in these subjects to maintain competitive advantage in an increasingly technology-driven global economy (National Academy of Sciences et al., 2007). The report's findings were echoed by numerous industry association reports and policy briefs, and provided support for Congress's passage of the America Competes Act in 2007. The Act had a broad agenda of funding science and technology research, and increasing NSF's budget, along with renewed emphasis on expanding the supply of STEM workers, especially increasing the supply of guestworkers, in response to a claimed shortage. At the same time, and once again, careful reviews of the evidence found the claimed shortages and loss of international competitiveness—to China, rather than the Soviet Union this time—was based on a flawed analysis of the data (Lynn and Salzman, 2010; Lowell and Salzman, 2007).

By 2007 when the report was issued, a booming economy made the claims of impending STEM failures less urgent than they had seemed following the dot-com crash earlier in the decade, when the crisis and failure reports began to proliferate (Teitelbaum, 2014). But a year later, the 2008 financial crisis put concerns about the US economy on center stage again.

In response to the financial crisis, an assemblage of science and industry advisers, along with many of the same authors of the 2007 RAGS report, offered in a 2010 follow-up the same argument of science and engineering declines in face of an ascendent China that portended an imminent threat to the nation like a "Category 5 hurricane." To the earlier report this one added further claims, again hyperbolic and fantastical, concluding the failure to respond to their first report could be seen as leading to outcomes that—quoting another 2001 report on which the same military contractor CEO served—were "second only to a weapon of mass destruction detonating in an American city ... " (National Academies, 2010, p. 66–67).

At the same time, President Obama assigned his President's Council of Advisors on Science and Technology (PCAST) the task of examining the future STEM needs in the economy. To bolster the mission of this group of science advisers, the President assembled a group of a dozen and a half industry executives and two labor representatives in an ad hoc Council on Jobs and Competitiveness, for which he tapped General Electric's CEO Jeff Immelt as chair. Their task was to formulate policies to address the causes of, and remedies to, the economic crisis and slow recovery. Remarkably, if consistent, the solution both councils identified was to increase the number of STEM graduates: the Jobs and Competitiveness council issued an interim report in June 2011 calling for 10,000 more engineers (Presidents Council on Jobs and Competitiveness, 2012). Shortly after that, the Science and Technology Advisors, in February 2012, called for one million more STEM graduates (Presidents Council of Advisors on Science and Technology, 2012).

Both of these proposals came a year after Apple's co-founder Steve Jobs told the president that Apple would have located its vast low-cost manufacturing operation, with 700,000 jobs, in the United States instead of China if only the company had been able to find enough US engineers to support its operations. And all these reports came on the heels of Obama's proclamation a year earlier that a new "Sputnik moment" had arrived, calling for more STEM workers to lead the economic recovery. Defying the evidence of vast numbers of STEM graduates without STEM jobs, the president, policy makers, the press and industry all united in their claim of STEM shortages. The shortage claim was once again used to advance political and policy goals of refocusing attention to failures of the education system rather than the financial system, and to renew calls for increases in lower paid guestworkers for the technology industries (Teitelbaum, 2014; Lynn and Salzman, 2011; Salzman, 2013).

STEM education and workforce development garners attention distinct from overall education and workforce development. In the post-war period, the strategic importance of science and engineering came from the obvious role it had in military weaponry as well as the general increase in the role of science and technology in the economy and everyday life. However, STEM policy was encumbered from the beginning with other political and policy objectives. As the focus of attention in response to broader anxieties—whether about international threats or domestic economic crises—STEM policy is seldom based on substantial empirical analysis. In particular, the persistent theme of STEM shortages is one of unfounded claims that often arise during a general period of economic or political turmoil; each and every time major reports on the STEM crisis have been issued, they have been found not to reflect reliable evidence.

Importantly, claims about STEM education and the workforce are often made with little conceptual clarity, and little empirical evidence, yet the STEM crisis remains a potent narrative that has wide appeal and broad applications in advancing policy in education and workforce development. It seems to matter not that, in more than half a century of STEM education and workforce shortage claims, the supporting evidence has never withstood careful examination; in fact, although "STEM" is a term widely used, it has no consistent or coherent definition.

What is STEM anyway? Conceptual and definitional issues

For all the political importance ascribed to STEM, it has come to describe a rather wide range of disciplines and occupations, inconsistently used from one study or policy to another, and all lacking a logically or empirically consistent or defensible definition. Does STEM denote fields of study? Occupations? Or a body of knowledge that is applied across any number of occupations? Indeed, all of these permutations are used in discussions and policymaking about STEM. The lack of a precise and agreed-upon definition of what counts as a STEM course, major, or occupation is a problem that, we find, would alter the evidence base that is used in many policy, funding, and education decisions. As it relates to education and workforce development, STEM is both a category of education and occupation. In this regard, there are two important sources of ambiguity in the definition of STEM.

The first of these is whether to include health majors and occupations under the heading of STEM. The first census of science and engineering (S&E) was developed by NSF for a series of indicators responding to Congress's request for an assessment of the nation's scientific and technical workforce. Because another US government agency—the National Institutes of Health (NIH)—is responsible for health and medicine, these fields are generally excluded from NSF's funding and data collection efforts. But as is obvious to even casual observers of education and workforce training, this distinction is more political than substantive. The curriculum for pre-medical studies overlaps significantly with the laboratory sciences of biology and chemistry, and health programs outside of premed have become increasingly reliant on mastery of scientific and technological content. The decision not to include graduates in health fields has an enormous impact on the reported size of a graduating STEM cohort or the STEM workforce (Oleson et al., 2014); including health graduates in STEM would increase the size of the STEM cohort by 20% or more. However, for the purpose of this analysis we follow nearly all other analyses of the STEM workforce and exclude health fields from our accounting of STEM education or the workforce.

A second problem relates to the definition of "technology"—the T in STEM. Programs of study in Computer Science and Information Technology may fit most neatly under the "technology" umbrella, but beyond this, what constitutes a technology program of study is unclear. In a rather remarkable expansion of the definition, the STEM label has been applied to HVAC installers and automotive technicians, and perhaps even includes janitors with the title "environmental engineer" (Rothwell, 2013). But apart from these ventures into empiricist absurdism, the technology workforce as part of the STEM classification refers to those with a college degree in a recognized technology field—computer science, manufacturing technology, etc.—or those in a professional occupation classified by the Bureau of Labor Statistics as technology fields such as computer programmers. When considering those with a technology degree, four-year and post-graduate degrees considered technology include computer science and engineering, while two-year Associates degrees also include machinists, repair technicians, etc. Health technicians such as radiology and EKG technicians, could be classified in either technology or health care. We developed a specific category of Health-tech for this group and do not include them in the STEM-technology grouping.

The estimates of demand for STEM workers and their potential supply, because of the diversity of the occupations, education, skills, and disciplines subsumed by the acronym, cannot be determined with any empirical or logical consistency or reliability. Workforce, education, and policy analysis thus needs to consider the specific disciplines and occupations that are relevant for a particular policy or program.

STEM demand and supply

To understand whether the supply of new workers is adequate depends on first identifying the number and trend of available jobs in STEM—demand—and the number of available STEM workers—supply. But since there isn't a standard or consistent definition of STEM, or of the specific qualifications needed for STEM jobs, discussions of STEM in the aggregate tend toward speculation and/or serving political purposes. For example, the more broadly one defines STEM, the larger the needed labor force, and thus the larger the potential to find a “shortage,” particularly—and paradoxically—when supply is not limited to a specific field of study. Estimates of the size of the STEM workforce vary dramatically based on whether these figures include healthcare occupations and “blue-collar” work involving machinery (e.g., advanced manufacturing), and on the decision whether to include jobs requiring less than a bachelor's degree (see below). One analysis found that estimates of the size of the US STEM workforce ranged from 5.4 million to over 26 million, depending on such operational choices (Oleson et al., 2014).

The largest category of US STEM workers, excluding health, is Information Technology (IT). In recent years, the IT industries have demand to fill about 50,000 new jobs per year (Zilberman and Ice, 2021) along with replacements for workers leaving these fields, with many of these jobs requiring less than a Bachelor's degree (see Fig. 1). In terms of supply, for the last eight academic years, the number of Bachelor's Degrees awarded in computer science awarded in the US has exceeded 50,000, and trended sharply upward, to over 95,000 in 2020. A similar trend of increased degree production is evident in engineering, where the number of new BAs increased from just over 75,000 in 2011 to over 129,000 in 2020. Demand for new engineers, however, is not expanding rapidly—that sector was only expected to have about 50,000 job openings per year, of which only 14,000 were new jobs and the balance replacements for retirements and those leaving the field (Torpey, 2018).

The technology industries pose a special case for estimating supply because the disciplinary background of that workforce is largely outside of computer science or IT fields, and even outside of STEM. Since nearly 30% of the IT workforce has less than a four-year college degree, and nearly half of those with a bachelor's degree did not major in computer science or engineering (see Figs. 1 and 2), the industry has a vast pool of potential workers to draw on for employment from the 470,000 STEM bachelors graduates, and 1.5 million non-STEM bachelors graduates, each year (U.S. Department of Education, 2020). The IT industry's very high turnover rates—reflecting its “burn and churn” employment practices that lead to low tenure—along with its preferences for hiring lower-cost guestworkers, may be factors that lead to the technology industry's stated demand as significantly higher than reflected in these projections (Lazonick et al., 2014; Salzman, 2016). Obviously IT occupations require a specific set of skills that may not be available in the population at large, but the rather modest needs of this industry for new hires, representing 8% of all industry annual workforce growth (Dubina et al., 2020), suggests the IT industry should be able to address its supply needs with no more difficulty than other STEM industries that require much higher levels of formal education and training (see Fig. 1).

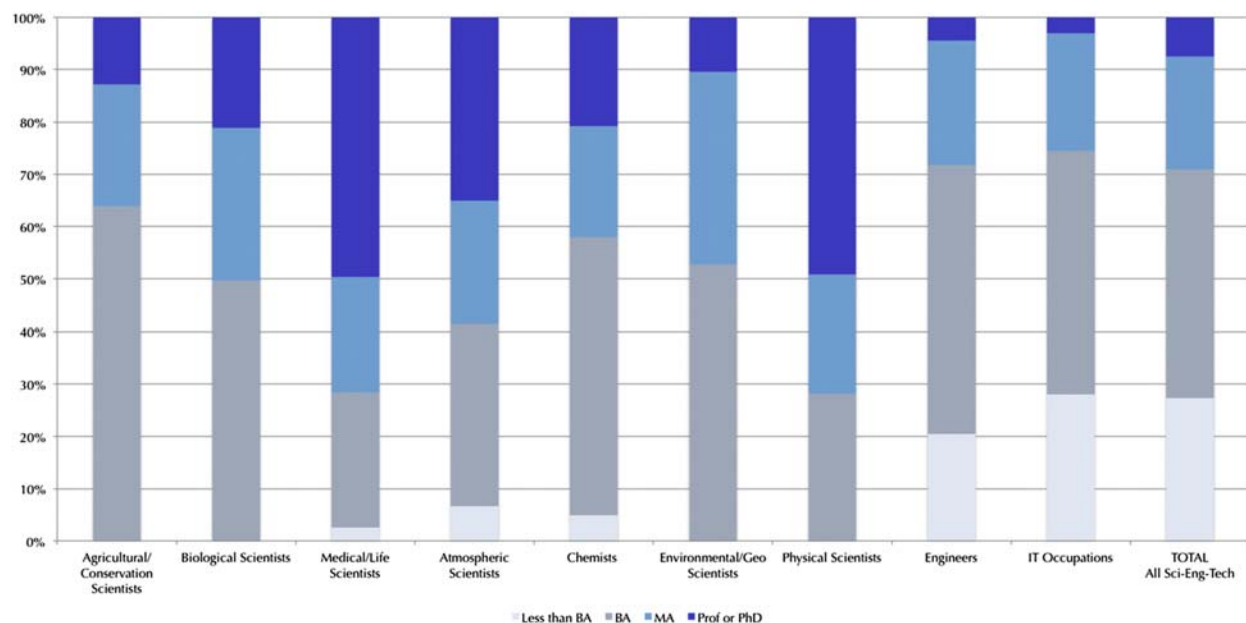


Fig. 1 STEM occupations by educational attainment. Authors' calculations based on: Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota, 2017. <https://doi.org/10.18128/D010.V7.0>.

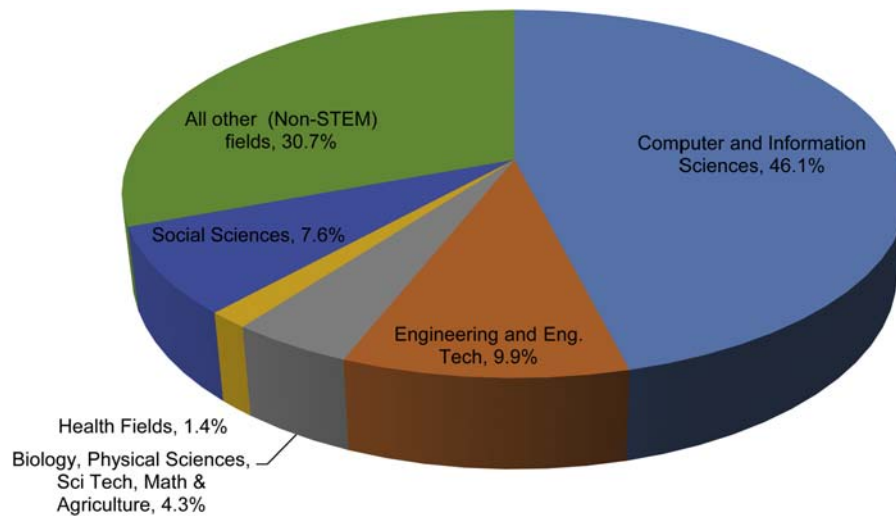


Fig. 2 IT workforce entrants' degree and field of education. Authors' tabulations based on: U.S. Department of Education, National Center for Education Statistics, Baccalaureate and Beyond Longitudinal Study (B&B), 2008/2012.

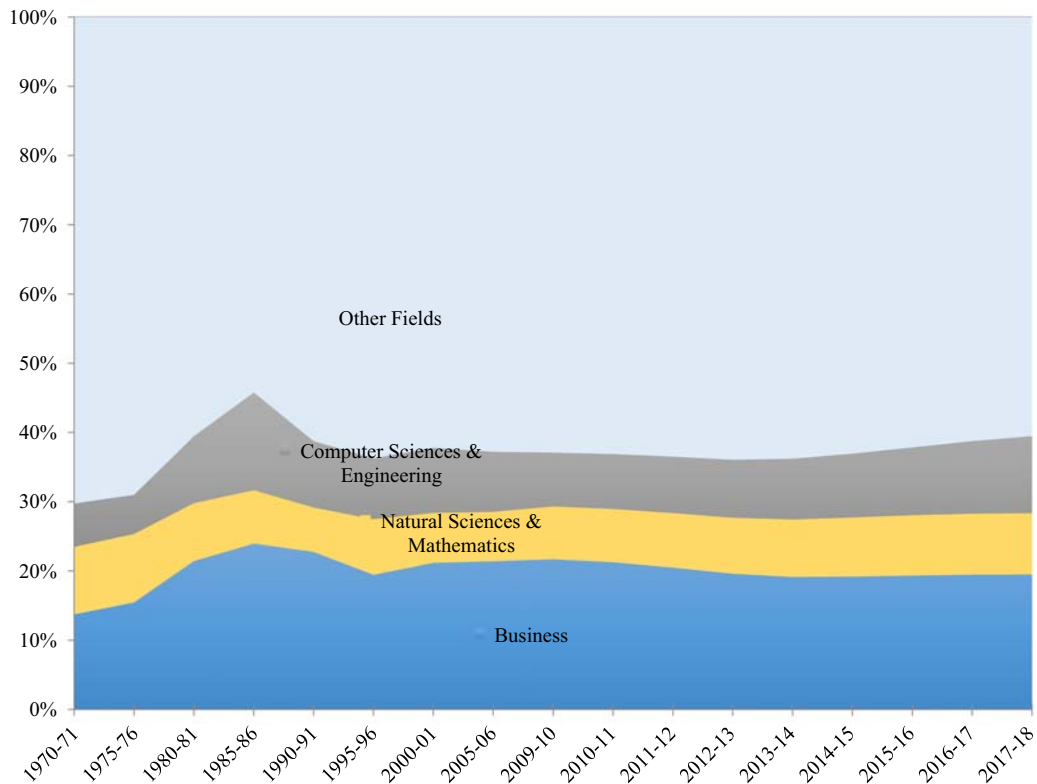


Fig. 3 Major fields of bachelor's graduates. Authors' tabulations based on: Digest of Education Statistics, 2015–2019 NCES Chapter 3.

Global STEM supply and domestic labor markets

The supply of domestic and global graduates varies quite dramatically by discipline and degree level, as do the labor markets for each occupation. Traditional calculations that examine just national education and labor markets without consideration of global flows do not accurately reflect STEM workforce supply and demand. Nor do analyses that examine these factors aggregated across disciplines or occupations provide much insight. In the US, the STEM supply and labor markets are structured to a significant degree by several guestworker policies enacted primarily for the technology industries and for computer science and engineering degrees, but also include the other STEM fields. These are guestworker work permits and visas known as the Optional Practical Training

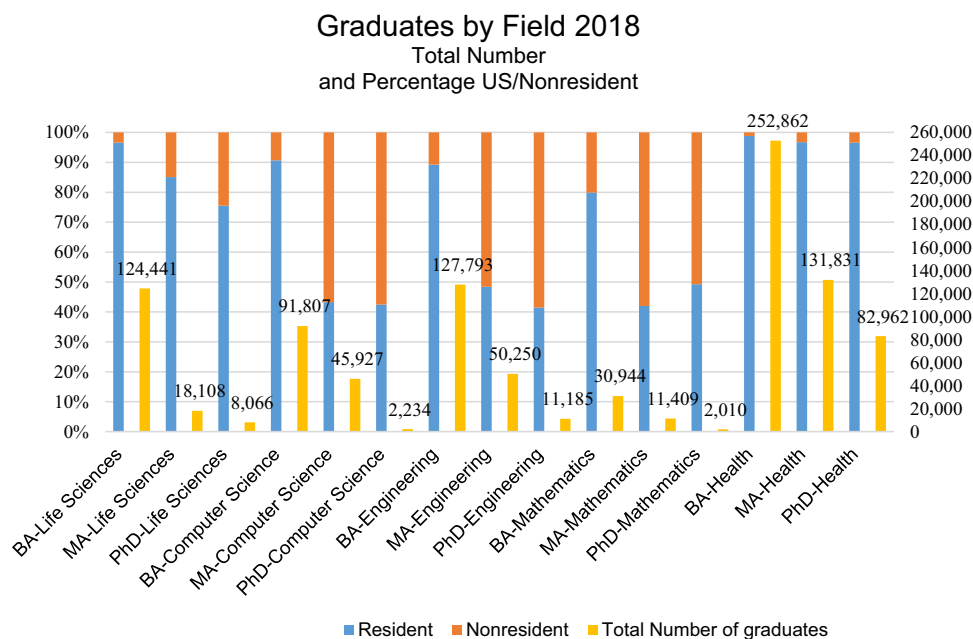


Fig. 4 Global student flows and graduate supply. Authors' tabulations based on: National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2018.

(OPT) student work permit, H-1B nonimmigrant high skill visa, and a few other categories such as the L-1 non-immigrant intra-company transfer visa. These visas are estimated to supply a guestworker labor pool equal to between one-half to two-thirds of all new entrants to the technology industries (Salzman, 2016; Rosenthal, 2021).

The supply of potential guestworkers, college graduates who qualify for a nonimmigrant visa or work permit, varies by degree level and field. As shown in Fig. 4, foreign students are a small share of bachelors graduates but comprise over half of Masters graduates in computer science, engineering and mathematics. Student flows into technology master's degree programs have increased dramatically in recent years. This reflects a change in US visa policy that vastly expanded the workforce entry options for foreign students. The two major routes into the workforce for foreign students during the 1990s and early 2000s had been the H-1B visa and, with much longer waits, permanent resident status known as the green card. Congress proposed legislation to expand these programs, in part fueled by the industry-driven shortage reports discussed earlier, but ultimately failed when the guestworker legislation was merged with broader immigration legislation (although part of the immigration bill, the H-1B visa and related visas and work authorizations, such as the OPT work authorization for the student F visa, are guest worker, nonimmigrant visas but often and erroneously described as immigration visas in the media and research reports). Consequently, to provide the technology industry greater access to guestworkers, the existing one-year work permit for all foreign graduates, known as the Optional Practical Training (OPT) work authorization, was extended to 17 months in 2008 for STEM graduates and then, in 2016, the Obama administration extended it to two years for STEM graduates, in addition to the one-year OPT for all graduates, thus creating a three-year work permit for STEM graduates (Rosenthal, 2021).

This liberalizing of the OPT work permits in 2016 led to a dramatic increase in the numbers of foreign students obtaining three-year work permits, nearly doubling from 2016 to 2019, and increasing 13-fold from 2009 (Fig. 5). As found in an analysis by Bloomberg journalist Rachel Rosenthal in a recent in-depth analysis:

With no annual cap, the number of work authorizations exploded, to 203,000 in 2019 from 83,000 in 2008, according to Immigration and Customs Enforcement, which oversees many aspects of the program. About 411,000 people had documentation to work under OPT in 2019, not including another 126,000 from a related pre-graduation version of the program. The number of initial approvals for OPT has exceeded those for H-1B visas every year since 2015, according to data from US Citizenship and Immigration Services.

Rosenthal (2021).

Since there is no cap on the number of OPT work authorizations for the F visa, and it has none of the minimal wage or work protections of the H-1B visas—thus allowing employers to hire them as low as the federal minimum wage of \$7.25 per hour—these graduates became the preferred source of new entrants. Universities recognized the revenue potential of the OPT program because graduates from their two-year Masters programs could then easily enter the US labor market for three years while trying to obtain the more limited number of six-year H-1B visas. There was rapid expansion of Masters programs in computer science and engineering targeted exclusively or near exclusively to foreign students who would pay full tuition as the price of entering the US labor market.

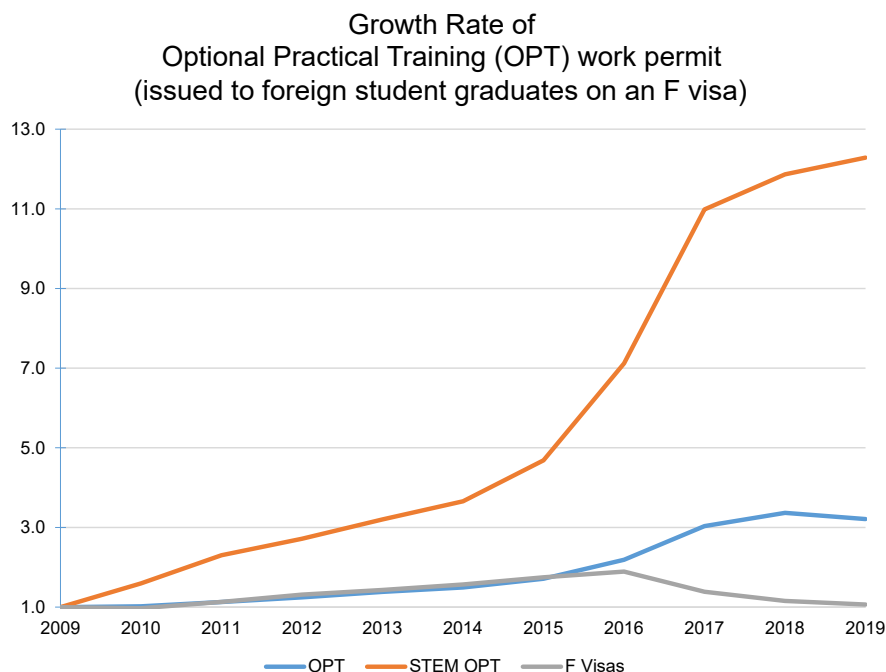


Fig. 5 Growth rate of OPT Calculations by authors; based on U.S. Immigration and Customs Enforcement SEVIS data 2008=1.0/F visa, OPT; 2009=1.0/OPT STEM; OPT STEM lagged 1 year—Issued at expiration of first year OPT.

Although some colleges with established and reputable programs created or expanded their Masters programs for foreign students, the vast majority of graduate supply in technology fields came from unranked and low-ranked programs: “More than 70% of nonresident computer science master’s degrees awarded in 2018 came from unranked programs or those ranked 50 and lower by US News and World Report. Just 17% came from schools ranked in the top 25” (Rosenthal, 2021). In the past decade the colleges with the largest cohorts of Computer Science graduates were the unranked universities of Wilmington University, Central Missouri State University, Southern Arkansas University, and University of Cumberlands, where over 90% of the graduates were nonresident, or foreign students. Overall, more than two-thirds of nonresident computer science Masters graduates were in programs that could be considered exclusionary, that were 75%–100% nonresident students. This level of exclusionary programs is specific to Masters and PhD programs in computer science, some fields of engineering, and mathematics (see Fig. 4).

It was in this way that legislation provided the technology industry the opportunity to hire guestworkers at below market wages (Hira and Costa, 2021) which, in turn, shifted the demand away from domestic to nonresident foreign computer science and engineering graduates. In response, some universities shifted their business model to create programs that exclusively, or near-exclusively recruited full-tuition-paying nonresident foreign students. Justifying the expansion of these programs, with a large share coming from universities without a history or evident credentials in these fields, the STEM shortage narrative provided justification for the policies that sustained these programs (Rosenthal, 2021; Salzman, 2016). It is, thus, the intersection of government policies, global labor supplies and the opportunity for global labor arbitrage, along with entrepreneurial initiative by some universities that combine to structure one segment of the technology labor market different from other STEM labor markets or STEM education.

In summary, the evidence does not support the longstanding narratives of STEM shortages or the predictions of dramatic increases in demand that cannot be met by existing education institutions to provide the necessary supply (Salzman and Benderly, 2019; Lynn et al., 2018). Considering the political and industry functions of the STEM shortage narrative, and the development of a technology student and worker supply system that is exclusionary and draws on neither the high-talent capacity nor the potential diverse domestic pool of students and workers, the evidence suggests that the global education and employment systems inhibit full development of the STEM workforce potential in the nation.

Problems and problematics in the supply

Up to this point, we have raised questions about the definition of STEM and assessed claims of shortages of college-educated STEM workers. But beyond the issue of relative levels of supply and demand, there is of course a need for a trained or re-trained STEM workforce at both the Baccalaureate and sub-Baccalaureate level. It is thus necessary to assess obstacles to producing that supply of workers at both levels. This is particularly important when considering factors that can facilitate or hinder the development of a diverse STEM workforce and of drawing on the full potential within the nation’s population.

Sub-baccalaureate STEM credentials

While most discussions of STEM workforce development emphasize the supply of bachelor's degree graduates, it is important to recognize that over a quarter of all STEM occupations require less than a bachelor's degree (Fig. 1). Furthermore, nearly half of students who ultimately earn a bachelor's degree in the US spend at least some time at two-year postsecondary institutions, community colleges (NSC Research Center, 2017). And as concerns about diversity in the STEM labor force grow, it is important to consider that community colleges serve a more diverse population of students than any other sector of postsecondary education, and at much lower cost, providing greater access to lower-income students. Community colleges are thus a critical part of the STEM worker supply in the US (National Academy of Engineering and National Research Council, 2012), and any consideration of the STEM workforce needs to consider the role played by these two-year institutions.

Although community colleges are more diverse and provide greater access to education than four-year colleges, the completion rates are less than a third of those in four-year colleges. One study of students in the midwestern state of Ohio found that among students who begin their college careers at two-year institutions with intention to pursue a STEM major, only 14% remained in these fields at the time of their last enrollment, as compared to 43% of four-year college students (Bettinger, 2010). More recent figures indicate that 69% of entering STEM students at two-year colleges had either switched majors or left college altogether (Chen, 2013). Importantly, attrition rates from STEM fields are not dissimilar from non-STEM fields (Chen, 2013). But the difference in attrition rates between two- and four-year programs suggests that there may be differing supply problems in these two sectors.

Mathematics coursework and the supply of STEM graduates

Attrition from STEM majors is exacerbated by a curriculum structured to filter out students through mathematics courses that are not always necessary for the major. These are known as gateway courses because they can be a barrier to entry or continuation in a STEM pathway. At both the baccalaureate and sub-baccalaureate levels, entry or advancement in STEM fields of study is often contingent upon completion of mathematics courses. Indeed, such "service courses" for other fields of study are the primary function of mathematics departments in US colleges and universities. Our analyses of college transcripts indicate that STEM bachelor's degree earners completed about 17 mathematics credits—the equivalent of about five courses—compared to 8 credits among non-STEM majors. Among non-mathematics STEM majors, students in engineering (22 credits), physics (17 credits) and computer science (16 credits) earned the most math credits, with life sciences majors earning the fewest (9 credits). Notably, courses in college-level math and calculus in US colleges have uniquely high rates of failure and withdrawal, at 10% each, even when compared to other STEM courses (ranging from 3–5%), suggesting that mathematics courses play a gatekeeper role in STEM fields (Douglas and Salzman, 2020).

Mathematics remedial courses serve an additional gatekeeper role at the sub-baccalaureate level. Early in the last decade, nearly 60% of entering US community college students were assessed as needing remedial instruction in mathematics; almost double the one-third of students needing remedial instructions at baccalaureate institutions (Chen and Simone, 2016). Remedial placement decisions are typically made on the basis of one-off standardized tests, even as rates of misplacement with such tests were found to be fairly high (Scott-Clayton et al., 2014). Rather than remediating identified mathematics deficits, these courses lead to a high rate of students leaving without completing any degree (Attewell et al., 2006; Chen and Simone, 2016). Indeed, the pre-college level mathematics courses have the lowest success rates of any course type in US postsecondary education; nationally representative student transcript data shows that only 57% of pre-college mathematics course attempts end in success—grades of C or better—compared to 79% of college physics courses and 77% of college English courses (Douglas and Salzman, 2020).

Colleges, especially community colleges where the problem is most severe, have recognized this issue, and have begun to respond. Reforms to remedial mathematics policy have ranged from reforming curriculum and pedagogy to altering placement mechanisms, to modifying remediation requirements or eliminating them altogether (Rutschow et al., 2019). One element of curricular reform is to align required mathematics content with students' intended fields of study, based on the assumption that different degrees will require different types of mathematical knowledge. Generally, this has meant that students intending to pursue non-STEM degrees are placed on a track to take introductory statistics or other non-algebra courses, while STEM-intending students are placed on a track to take college algebra and/or calculus. In general, these reforms have succeeded in reducing rates of remedial mathematics placement and increasing college completion (Center for the Study of Social Policy, 2016). The gatekeeper function of mathematics courses, to "weed out" and reduce the pool of STEM-eligible students, is a longstanding practice but is not grounded in evidence for its utility pedagogically or occupationally. Research on mathematics use among US workers indicates that a very small proportion of US workers—even those in occupations that require college degrees—make regular use of mathematics beyond simple computation (Douglas and Attewell, 2017; Handel, 2016). Alignment of mathematics requirements with students' college-level coursework and eventual occupations needs to be carefully considered, especially if mathematics coursework requirements are limiting supply by increasing attrition from STEM majors or college dropout altogether.

STEM attrition and STEM attraction

The traditional—and often unexamined—narrative of STEM education in US higher education is the so-called "leaky pipeline." This narrative and supporting analyses are based on the observation that many students who begin their postsecondary education career in STEM fields change to non-STEM fields of study and ultimately graduate with a non-STEM degree. This pattern of STEM attrition has been the subject of many studies, with particular focus given to the departure of female and underrepresented minority students

(Seymour and Hewitt, 1997; Crisp et al., 2009; Hill et al., 2010; Chen, 2013). Consequently, policymakers in the US have focused attention on retention, of plugging leaks in the pipeline by creating programs to retain students in STEM majors.

The focus on retention is a positive development to the extent that it spurs cultural changes in STEM fields, such as reducing discrimination against particular groups of students or improving pedagogy. However, it is also important to recognize that college is a process of discovery about interests and abilities, and some amount of change in field should be expected and is beneficial. Indeed, nationally representative survey data indicate that one-third of four-year college students change their major at least once within their first three years, and that students leaving STEM majors is about as common as students leaving non-STEM majors (Leu, 2017). The focus on STEM education as a “pipeline” that has only unidirectional “leaks” has led to a truncated analysis that fails to consider the “late entrants” who discover and move into a STEM field after their first years following an initial period of exploration.

In our analysis of this flow of students from college entry to graduation, we find that although a large share of four-year college freshmen leave STEM majors and switch to non-STEM majors, there is a slightly larger number of students who begin as non-STEM or as undeclared majors and are “late entrants” into STEM (Salzman et al., 2021). This results in the graduating STEM cohort becoming larger than the starting STEM cohort. The assumption that STEM completion is predicated on initial entry into STEM is thus not supported by this evidence, nor by an assessment of the “STEM-potential” of the non-STEM student population.

The large pool of non-STEM students who are STEM-potential is indicated by the significant share of students with high levels of mathematics coursetaking: there are twice as many non-STEM college majors as STEM majors with mathematics course credit levels at the STEM population median level of mathematics course credits (Douglas and Salzman, 2020). That is, the non-STEM population is large and a significant share of them take mathematics at the same level of intensity as half of the STEM population. Some fields such as engineering are more difficult to enter late, but the vast majority of STEM fields do not require, nor does the typical major have STEM course intensity that is beyond the capacity of late entrants.

Moreover, if fostering flows in and out of STEM results in better matching, it may lead to better performance, satisfaction, and tenure in STEM careers. Currently, only a quarter to a third of STEM-degree holders work in a STEM occupation, and 15% of those in a STEM occupation with a four-year degree do not have any STEM degree (Kannankutty, 2007; Landivar, 2013). Instead of trying to “plug the leaks,” or making the boundary less permeable, a more productive approach might be analysis of STEM attraction rather than of just attrition, focusing instead on ways to improve the exploration and matching processes of students, of understanding both factors that account for retention/persistence and the factors that attract students into STEM fields.

Finally, we may want to reconsider the assumption that STEM development and national competitiveness are best served by strengthening a pipeline that has high throughput with minimal flows across disciplinary boundaries. For the STEM-eligible population, of the pool of high-performing students, it may be the loose coupling between STEM disciplines and STEM careers that provides the US some of its dynamism, innovativeness, and creativity. Steve Jobs famously said that Apple, among the world’s most highly valued companies, represents the intersection between technology and the humanities. Before Jobs, Edwin Land, co-founder of Polaroid and a developer of the nation’s first advanced aerial imaging technology as well as key advisor in the founding of NASA, had discussed the importance of “standing at the intersection of humanities and science.” The porous boundaries of the US education and career system should be seen as a strength that supports and fosters the creativity of cross-disciplinary meandering.

Conclusion

Simple notions about STEM workforce policy fuel debates about national competitiveness the world over. In the United States, the STEM crisis theme is a perennial policy favorite, appearing every few years as an urgent concern in the nation’s competition with whatever other nation is ascendant, or as the cause of whatever problem is ailing the domestic economy. And the solution is always the same: increase the supply of STEM workers through expanding STEM education. Time and again, serious and empirically grounded studies find little evidence of any systemic failures or an inability of market responses to address whatever supply is required to meet workforce needs. The large share of the STEM workforce that requires advanced college degrees means that supply responses will be lagged by two or three years, but in decades of study there is no evidence of a failure to create whatever supply employers are seeking to hire at market rates (Freeman, 1976; Salzman, 2013; Lynn et al., 2018; Freeman and Salzman, 2018). Instead, the motivating interests of STEM workforce shortages or failures are found to be attempts to shape policies for other goals, most recently to lower the cost of labor through global labor arbitrage facilitated by policies responding to manufactured crises (Berliner and Biddle, 1997; Teitelbaum, 2014).

Although the goal of increasing support of STEM education and workforce development would seem to benefit from this increased attention and funding, even if the motivation is misguided, the seemingly paradoxical effect of these policies is to undermine the nation’s, and the world’s STEM workforce development and capacity for innovation. The expansion of guestworker policies and programs has created a shadow STEM workforce system of universities that create Masters programs to supply the lower cost, guestworker labor market (Hira and Costa, 2021; Rosenthal, 2021; Salzman, 2016), and that allows employers to pay below market wages—in the case of H-1B workers, transferring \$3 billion to \$4 billion per year of forgone wages from employees to employers (Employment and Training Administration, 2021). This global labor arbitrage system leads firms to disinvest in their STEM workforces and rely, instead, on a low-cost, high-churn model of employment (Lazonick et al., 2014). To the extent that

the quality of the STEM workforce is a factor in endogenous growth models explaining national economic performance, this STEM disinvestment might be contributing to the recent productivity and innovation declines (Gordon, 2016).

STEM education and workforce development have suffered in the wake of policies purported to be strengthening them. Education and workforce policy are part of the overall political process that reflects conflicting goals and are too often supported by analyses and policy reports that are narrowly focused without the necessary historical context or detailed empirical specificity required to develop effective policy. A broader perspective on the political context of STEM crisis reports, and more rigorous research and analyses will better inform our understanding of the STEM fields and how to best strengthen them.

Acknowledgments

This material is based upon work supported by the Alfred P. Sloan Foundation, Grant Nos. 2012-6-13 and G-2016-7310; and the National Science Foundation, Grant No. NSF-DGE-1561687. The authors are most grateful for the expert research assistance of Daniyal Rahim and Khudodod Khudododov.

References

- Attewell, P., Lavin, D., Domina, T., Levey, T., 2006. New evidence on college remediation. *J. High. Educ.* 77 (5), 886–924.
- Berliner, D.C., Biddle, B.J., 1997. *The Manufactured Crisis. Myths, Fraud, and the Attack of America's Public Schools*. Addison-Wesley, Reading, Mass.
- Bettinger, E.P., 2010. To be or not to be: major choices in budding scientists. In: Clotfelter, C. (Ed.), *American Universities in a Global Market*. University of Chicago Press, Chicago.
- Blank, D.M., Stigler, G.J., 1957. *The Demand and Supply of Scientific Personnel*. National Bureau of Economic Research, New York.
- Borjas, G.J., 2016. *We Wanted Workers: Unraveling the Immigration Narrative*. W.W. Norton & Company, New York.
- Bracey, G.W., 2009. *Education Hell: Rhetoric vs. Reality*. Educational Research Service, Alexandria, VA.
- Brinkley, A., 2010. *The Publisher: Henry Luce and His American Century*. Knopf, New York.
- Bush, V., 1945. *Science: The Endless Frontier*. United States Government Printing Office, Washington, DC.
- Carson, C., Huelskamp, R., Woodall, T., 1993. Perspectives on education in America: an annotated briefing. *J. Educ. Res.* 86 (5), 259–310.
- Center for the Study of Social Policy, 2016. *Carnegie Math Pathways: Case Study*. Center for the Study of Social Policy, Washington, DC.
- Chen, X., Simone, S., 2016. Remedial Course-taking at U.S. Public 2- and 4-Year Institutions: Scope, Experiences, and Outcomes. National Center for Education Statistics, Washington, DC.
- Chen, X., 2013. *STEM Attrition: College Students' Paths Into and Out of STEM Fields*. National Center for Education Statistics, Washington, DC.
- Clowse, B.B., 1982. *Brainpower for the Cold War: The Sputnik Crisis and National Defense Education Act of 1958*. Greenwood Publishing Group, United States.
- Crisp, G., Nora, A., Taggart, A., 2009. Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: an analysis of students attending a hispanic serving institution. *Am. Educ. Res. J.* 46 (4), 924–942.
- Douglas, D., Attewell, P., 2017. School mathematics as gatekeeper. *Socio. Q.* 58 (4), 648–669.
- Douglas, D., Salzman, H., 2020. Math counts: major and gender differences in college mathematics coursework. *J. High. Educ.* 91 (1), 84–112.
- Dubina, K.S., Kim, J., Rolen, E., Rieley, M.J., 2020. Projections Overview and Highlights, 2019–29. *Monthly Labor Review*, pp. 1–33.
- Employment and Training Administration, 2021. *Strengthening Wage Protections for the Temporary and Permanent Employment of Certain Immigrants and Non-immigrants in the United States; Delay of Effective Date*. Federal Information & News Dispatch, LLC, Washington.
- Freeman, R., Salzman, H. (Eds.), 2018. *U.S. Engineering in a Global Economy*. University of Chicago Press, Chicago.
- Freeman, R.B., 1976. A cobweb model of the supply and starting salary of new engineers. *Ind. Labor Relat. Rev.* 29 (2), 236–248.
- Gordon, R.J., 2016. *The Rise and Fall of American Growth*. Princeton University Press, Princeton.
- Handel, M.J., 2016. What do people do at work? *J. Labour Mark. Res.* 49 (2), 177–97. <https://doi.org/10.1007/s12651-016-0213-1>.
- Hill, C., Corbett, C., St Rose, A., 2010. *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. American Association of University Women, Washington, DC.
- Hira, R., Costa, D., 2021. The H-1B Visa Program Remains the "Outsourcing Visa": More than Half of the Top 30 H-1B Employers Were Outsourcing Firms. Available: <https://www.epi.org/blog/the-h-1b-visa-program-remains-the-outsourcing-visa-more-than-half-of-the-top-30-h-1b-employers-were-outsourcing-firms/>.
- Hira, R., 2008. *Outsourcing America: What's Behind Our National Crisis and How We Can Reclaim American Jobs*. AMACOM, New York.
- Kannankutty, N., 2007. Multiple Pathways to Science and Engineering Employment: Characteristics of the U.S. S&E Workforce. Paper Presented at Atlanta Conference on Science, Technology and Innovation Policy.
- Kuehn, D., Salzman, H., 2018. The engineering labor market: an overview of recent trends. In: Freeman, R., Salzman, H. (Eds.), *U.S. Engineering in a Global Economy*. University of Chicago Press, Chicago.
- Landivar, L.C., 2013. *Relationship Between Science and Engineering Education and Employment in STEM Occupations: American Community Survey Reports, ACS-23*. U.S. Census Bureau, Washington, DC.
- Lazonick, W., Moss, P., Salzman, H., Tulum, O., 2014. *Skill Development and Sustainable Prosperity: Cumulative and Collective Careers Versus Skill-Biased Technical Change*. Institute for New Economic Thinking.
- Leu, K., 2017. *Beginning College Students Who Change Their Majors Within 3 Years of Enrollment*. National Center for Education Statistics, Washington, DC.
- Lowell, B.L., Salzman, H., 2007. Into the Eye of the Storm: Assessing the Evidence on Science and Engineering Education, Quality, and Workforce Demand. The Urban Institute, Washington, D.C. <http://www.dx.doi.org.proxy.libraries.rutgers.edu/10.7282/T3X068P1>.
- Lynn, L., Salzman, H., 2010. The globalization of technology development: implications for U.S. skills policy. In: Finegold, D., Gatta, M., Salzman, H., Schurman, S.J. (Eds.), *Transforming the U.S. Workforce Development System: Lessons from Research and Practice*. Cornell University Press, Ithaca, NY, pp. 57–85.
- Lynn, L., Salzman, H., 2011. Is the president right when he says the United States needs 10,000 engineers a year? Why not let the market decide? *Manuf. Technol. News* 18 (17).
- Lynn, L., Salzman, H., 2018. Science and engineering "competitiveness": developing collaborative advantage in a global commons. In: Ettlie, J., Hira, R. (Eds.), *Engineering Globalization Reshoring and Nearshoring: Management and Policy Issues*. World Scientific, New Jersey, pp. 205–228.
- Lynn, L., Meil, P., Salzman, H., 2012. Reshaping global technology development: innovation and entrepreneurship in China and India. *J. Asia Business Stud.* 6 (2), 143–159.
- Lynn, L., Salzman, H., Kuehn, D., 2018. Dynamics of engineering labor markets: petroleum engineering demand and responsive supply. In: Freeman, R., Salzman, H. (Eds.), *U.S. Engineering in a Global Economy*. University of Chicago Press, Chicago.
- Meil, P., Salzman, H., 2017. Technological entrepreneurship in India. *J. Entrep. Emerg. Econ.* 9 (1), 65–84. <https://doi.org/10.7282/T3571FB9>.
- National Academy of Engineering and National Research Council, 2012. *Community Colleges in the Evolving STEM Education Landscape: Summary of a Summit*. National Academies Press, Washington, DC.

- National Academy of Sciences, National Academy of Engineering, Institute of Medicine, 2007. *Rising Above the Gathering Storm*. National Academies Press, Washington, DC.
- National Academy of Sciences, 2010. National Academy of Engineering, and Institute of Medicine, *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category, 5*. National Academies Press, Washington, DC.
- National Commission on Excellence in Education, 1983. *A Nation at Risk: The Imperative for Educational Reform: A Report to the Nation and the Secretary of Education*. United States Department of Education, National Commission on Excellence in Education, Washington, DC.
- NSC Research Center, 2017. *Two-Year Contributions to Four-Year Completions—2017*. National Student Clearinghouse, Herndon, VA.
- Oleson, A., Hora, M., Benbow, R.J., 2014. What is a STEM Job? How Different Interpretations of the Acronym Result in Disparate Labor Market Projections. Wisconsin Center for Education Research, Madison, WI.
- Presidents Council of Advisors on Science and Technology, 2012. *Report to the President Engage to Excel: Producing One Million Additional College Graduates With Degrees in Science, Technology, Engineering, and Mathematics*. Executive Office of the President, President's Council of Advisors on Science and Technology, Washington, DC.
- Presidents Council on Jobs and Competitiveness, 2012. *Road Map to Renewal*. The White House, President's Council on Jobs and Competitiveness, Washington.
- Ravitch, D., 2014. *Reign of error: the hoax of the privatization movement and the danger to America's Public Schools*. Vintage, New York.
- Ravitch, D., 2020. *Slaying Goliath: the passionate resistance to privatization and the fight to save America's Public Schools*. Vintage, New York.
- Romer, P.M., 1994. The origins of endogenous growth. *J. Econ. Perspect.* 8 (1), 3–22.
- Rosenthal, R., 2021. The STEM Graduate System is Broken. Here's How to Fix It. Bloomberg Opinion (March 10, 2021). <https://www.bloomberg.com/graphics/2021-opinion-optional-practical-training-problems-stem-graduates-deserve-better-jobs-opportunities/>.
- Rothwell, J., 2013. *The Hidden STEM Economy*. Brookings Institution, Washington, DC.
- Rutschow, E., Cormier, M.S., Dukes, D., Cruz-Zamora, D.E., 2019. *The Changing Landscape of Developmental Education Practices: Findings from a National Survey and Interviews With Postsecondary Institutions*. Center for the Analysis of Postsecondary Readiness, New York.
- Salzman, H., Benderly, B., 2019. STEM performance and supply: assessing the evidence for education policy. *J. Sci. Educ. Technol.* 28 (1), 9–25.
- Salzman, H., Lowell, L.L., 2008. Making the grade. *Nature* 453 (May 1).
- Salzman, H., Douglas, D., Khudododov, K., 2021. Reconceptualizing STEM college pathways and gender: who stays? who leaves? who graduates? Rutgers University, NJ.
- Salzman, H., 2013. What shortages? The real evidence about the STEM workforce. *Issues Sci. Technol.* 58–67.
- Salzman, H., 2016. The Impact of High-Skill Guestworker Programs and the STEM Workforce, Statement of Hal Salzman for the Hearing on "The Impact of High-Skilled Immigration on U.S. Workers". Submitted to the Senate Committee on the Judiciary, U.S. Senate, 25 February.
- Saxenian, A., 2006. *The New Argonauts*. Harvard University Press, Cambridge, Mass.
- Scott-Clayton, J., Crosta, P.M., Belfield, C.R., 2014. Improving the targeting of treatment: evidence from college remediation. *Educ. Eval. Pol. Anal.* 36 (3), 371–393.
- Seymour, E., Hewitt, N.M., 1997. *Talking About Leaving: Why Undergraduates Leave the Sciences*. Westview Press, Boulder, CO.
- Teitelbaum, M.S., 2014. *Falling Behind? Boom, Bust, and the Global Race for Scientific Talent*. Princeton University Press, Princeton, NJ.
- Torpey, E., 2018. Engineers: employment, pay, and outlook: Career Outlook. U.S. Bureau of Labor Statistics. Available: https://www.bls.gov/careeroutlook/2018/article/engineers.htm?view_full.
- Urban, W.J., 2010. *More than Science and Sputnik: The National Defense Act of 1958*. University of Alabama Press, Tuscaloosa.
- U.S. Department of Education, 2020. National Center for Education Statistics, Higher Education General Information Survey (HEGIS), "Degrees and Other Formal Awards Conferred" surveys, 1970–71 through 1985–86; Integrated IPEDS Fall 2000, Completions component. https://nces.ed.gov/programs/digest/d21/tables/dt21_322.10.asp.
- Weinstein, E., 1998. *How and Why Government, Universities, and Industry Create Domestic Labor Shortages of Scientists and High-Tech Workers*. National Bureau of Economic Research, Cambridge, MA.
- Zachary, G.P., 1999. *Endless Frontier*. The MIT Press, Cambridge, MA.
- Zilberman, A., Ice, L., 2021. Why Computer Occupations are Behind Strong STEM Employment Growth in the 2019–29 Decade: Beyond the Numbers. U.S. Bureau of Labor Statistics. Available: <https://www.bls.gov/opub/btn/volume-10/why-computer-occupations-are-behind-strong-stem-employment-growth.htm#:~:text=The%20U.S.%20Bureau%20of%20Labor,3.7%20percent%20for%20all%20occupations> [2021, May 6].

Relevant website

STEM Pathways Research, <https://go.rutgers.edu/STEMRU>.

Family influence and STEM career aspirations

Megan E. Ennes^a, M. Gail Jones^b, Heidi D. Cian^c, Remy Dou^d, Brian Abramowitz^e, Kathleen E. Bordewieck^f, and Kimberly L. Ideus^g, ^a Department of Natural History, University of Florida, Gainesville, FL, United States; ^b The Friday Institute for Educational Innovation, Department of STEM Education, North Carolina State University, Raleigh, NC, United States; ^c STEM Transformation Institute, Department of Teaching & Learning, Florida International University, Stanford, CA, United States; ^d STEM Transformation Institute, Department of Teaching & Learning, Florida International University, Miami, FL, United States; ^e Thompson Earth Systems Institute, Department of Natural History, University of Florida, Gainesville, FL, United States; and ^f College of Education, North Carolina State University, Raleigh, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	370
The role of family in STEM aspirations	371
Science capital	371
Family science habitus	372
Parental support	372
Early childhood experiences	373
Experiences with tools	373
Families and STEM hobbies	374
Gender differences	374
Families from marginalized communities and underrepresented groups in STEM	375
Modeling and predicting family influences on STEM career aspirations	376
Science identity	376
Science self-concept	376
Social cognitive career theory	376
Expectancy-value theory	377
Modeling STEM career aspirations	377
Assessing family influence and STEM career aspirations	377
ASPIRES	377
NextGen Scientist Survey	378
S-STEM	378
TOSRA	378
Conclusion	378
References	378

Introduction

Due in part to a worldwide concern over the number and diversity of students choosing to pursue careers in science, technology, engineering, and math (STEM), supporting and developing STEM interests in youth has become a high priority (e.g., [Moote et al., 2020](#); [Young et al., 2017](#)). Here we examine factors that are associated with family support for STEM career aspirations with a focus on “science,” particularly in light of recent work that has found associations between science capital, attitudes toward science, informal science learning, and “STEM” related outcomes, including interest and career aspirations ([Dou et al., 2019](#); [Moote et al., 2020](#); [Verdin, Godwin and Ross, 2018](#)). Many factors impact an individuals’ STEM interests and career aspirations, and families play a major role in the development of STEM interests (e.g., [Archer and DeWitt, 2016](#); [Jones et al., 2021](#); [Sha et al., 2016](#); [Silander et al., 2018](#)). In this chapter we examine how families influence the development of STEM interests and career aspirations of youth.

Before examining the role of the family in development of interests and aspirations, it is important to note that the concept of “family” is cultural and can vary between groups ([Tillman and Nam, 2008](#)). Formal definitions of family may not accurately reflect the various family types and may limit researchers’ “ability to fully examine patterns of exchange, support, and caregiving” ([Tillman and Nam, 2008](#), p. 368) that impact the members of the family. Further, different family members may play different roles in influencing aspirations. In interviews with college STEM students of color in the United States, [Rincón et al. \(2020\)](#) noted that mothers were especially influential in students’ decisions to pursue postsecondary education, fathers and siblings in their selection of STEM majors, and family members—including extended family such as aunts and uncles—and teachers in decisions to aspire to careers in STEM. When considering who is engaged in socializing children and who may impact their STEM interests and aspirations, it is important to think about who the family members may consider “family” to potentially include extended family as well as caregivers and unrelated individuals who are considered kin ([Yosso, 2005](#)).

The role of family in STEM aspirations

There is increasing evidence that an early introduction to science is important to support lifelong interests and exposure to potential careers (e.g., Cheryan et al., 2017; Jones et al., 2016). Even if a child does not seek to pursue a science career, enjoying science as a leisure activity, as well as being scientifically literate, equips individuals with a foundation for a science-rich life and the capacity to engage in decision-making in a world with an increasing science and technology focus (Sharon and Baram-Tsabari, 2020; Vahey et al., 2019). The development of science interests and aspirations tends to happen earlier for youth whose families are involved in the development of these interests (Dabney et al., 2013). However, family involvement in science has been shown to be related in many cases to the family's levels of science capital and family science habitus (Archer et al., 2015a; Claussen and Osborn, 2013; Sha et al., 2016).

Science capital

Science capital stems from Bourdieu's theory of social reproduction (Archer et al., 2015a). It refers to the economic, cultural, and social resources related to science to which a family has access (Archer et al., 2015a). This includes not only physical resources but also intangible resources such as relationships with people associated with science. Bourdieu (1986) described capital as representing the structure and function of the social world. He believed access to capital determines one's chance at being successful in one's endeavors.

Capital has three basic states: economic, social, and cultural (Fig. 1, Bourdieu, 1986). Economic capital is typically related to money and may also be seen as rights to property. Social capital is built of the social connections one has and is able to leverage. The cultural capital of an individual includes the knowledge, abilities, and cultural artifacts (such as books and art) an individual gains as part of their upbringing and education (Bourdieu, 1986). Science capital is a type of cultural and social capital that focuses on the value, knowledge, networks, and resources that individuals have related to science. Cultural capital is foundational for the development of habitus (Bourdieu, 1973; Chandler, 2013) which has also been proposed as an element of science capital (Archer et al., 2015a).

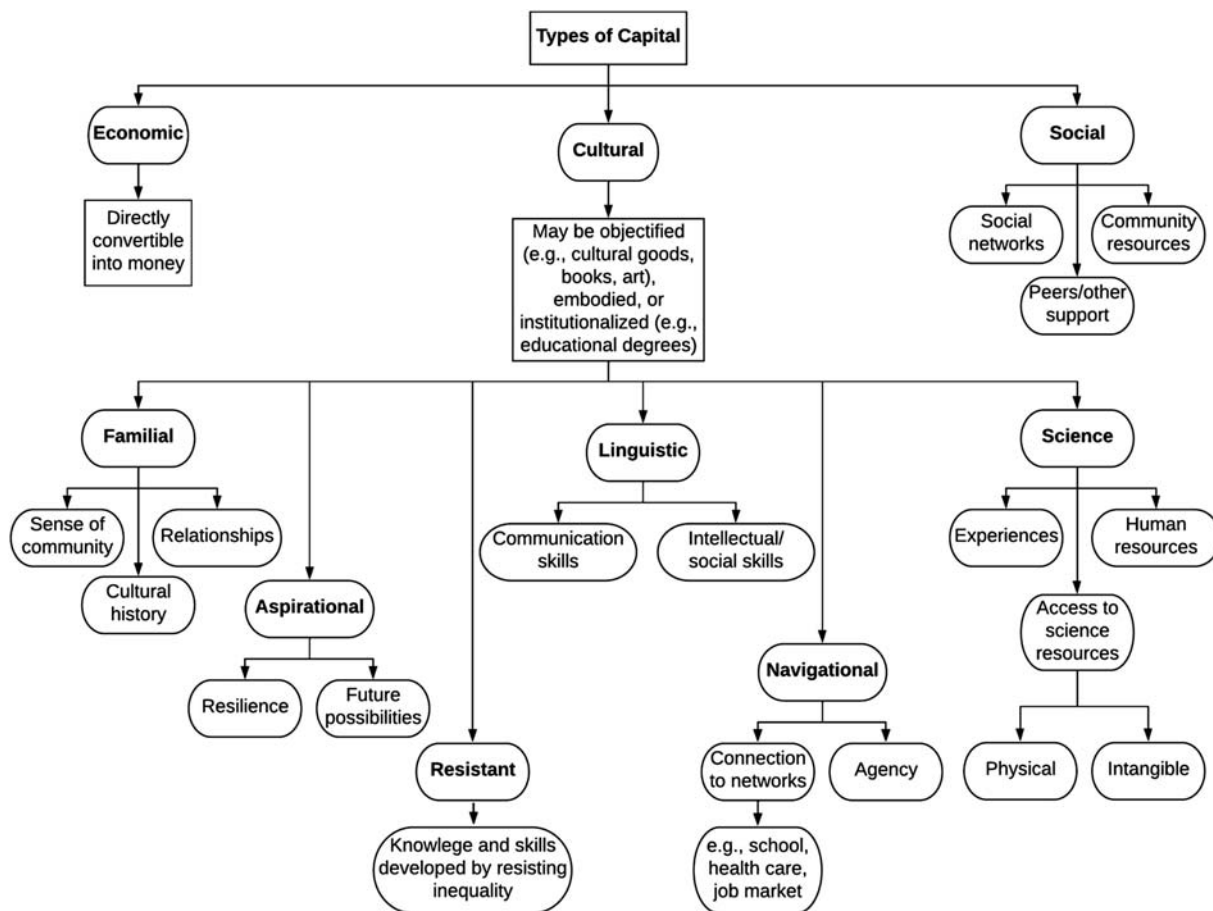


Fig. 1 Forms of capital. Adapted from Yosso (2005) and Archer and DeWitt (2016) as described in Ennes (2019).

Family science habitus

Habitus has been defined as a:

(s)ystem of lasting, transposable dispositions which, integrating past experiences, functions at every moment as a matrix of perceptions, appreciations, and actions and makes possible the achievement of infinitely diversified tasks, thanks to analogical transfers of schemes permitting the solution of similarly shaped problems (Bourdieu, 1977, p. 95).

From a social-psychology perspective, individuals have a series of internalized, cultural interests that are constantly competing with one another (Chandler, 2013). Cultural capital influences these field-specific interests (Chandler, 2013). A *field* is a socially structured environment that has its own set of rules and practices (Chandler, 2013). Science is one example of a field. The socialization of an individual into a particular field, such as science, at different points of their life will affect their habitus (Lizardo, 2004).

An individual's habitus is influenced by their experiences (Chandler, 2013; Dumais, 2002; McClelland, 1990), observations of others (McClelland, 1990), their place within society (or their family) and how they internalize that placement (Dumais, 2002; McClelland, 1990), race/ethnicity, gender, socioeconomic status (Archer et al., 2012), and capital (Reay, 1995). Habitus directs an individual's behaviors (Bourdieu, 1973), ways of speaking (Chandler, 2013; Silva, 2016), and belief systems (Archer and DeWitt, 2016; Silva, 2016). Additionally, habitus shapes what an individual believes is typical of who they are (Archer and DeWitt, 2016), as well as what is possible or acceptable for them to achieve in the future (Archer et al., 2012). Fig. 2 outlines several of the factors associated with habitus.

Archer et al. (2012) describe *family science habitus* as the way a family's capital (economic, social, and cultural) may influence the way the family thinks, identifies, values, talks about, and engages in science. These beliefs, values, and capital a family holds may influence the worldview that a child has regarding science (Archer et al., 2012). A family's habitus influences whether children are prepared to engage with the cultural capital of schools (Claussen and Osborne, 2013) or to engage in lifelong science habits and behaviors (Jones et al., 2016). A part of family science habitus, the types of experiences a family does or does not engage in, influences the science interests and career aspirations of the children (Jones et al., 2020). Fig. 3 summarizes the influences of parents and caregivers on the STEM interests of youth as outlined by Eccles (2015), with a focus on the role of capital and habitus as outlined in this chapter.

Parental support

Regardless of race or socioeconomic status, parents influence their children through their attitudes, encouragement, expectations, interests, aspirations, and role modeling (Dasgupta and Stout, 2014; Fisher and Padmawidjaja, 1999; Kiyama, 2010). They also affect their children's career aspirations through indirect pathways such as influencing their children's levels of self-efficacy, which influences children's career aspirations (Sawitri et al., 2015; Sha et al., 2016). Parental support allows children to explore careers and

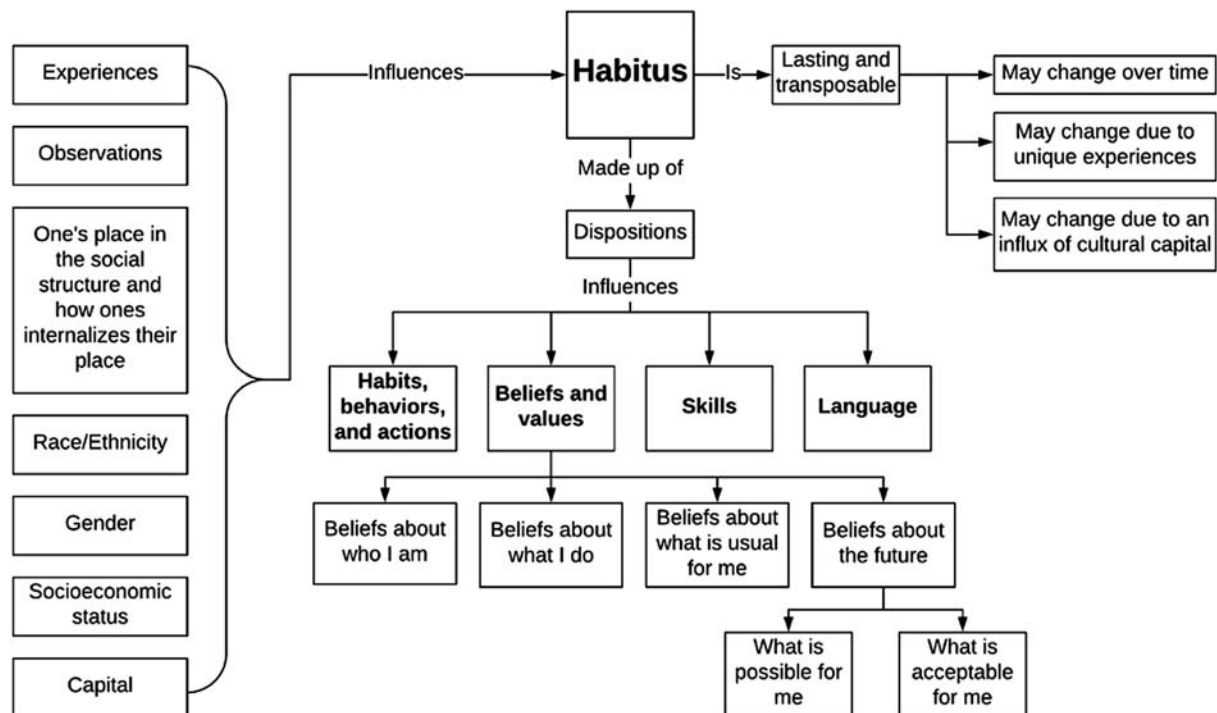


Fig. 2 Factors influencing and influenced by habitus as described in Ennes (2019).

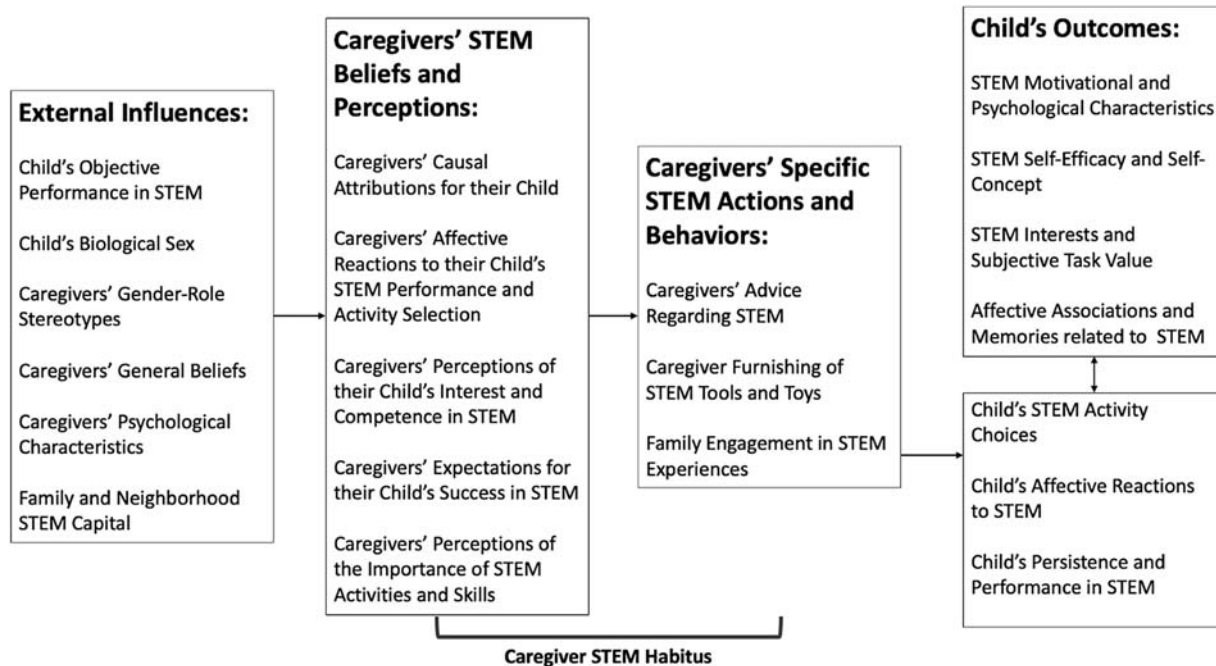


Fig. 3 Caregiver and parental influences on the STEM interests of youth as described by Eccles (2015).

may support their motivation and confidence in their career pursuits, leading to an increase in self-efficacy that helps them react positively to challenges to their career aspirations (Guan et al., 2016).

Parental support is important for all students, however, increasing this support for underrepresented youth may be particularly beneficial, especially as these students experience marginalization and discrimination in science contexts outside of the home (Cian et al., 2022; Kier and Blanchard, 2021). Researchers examining families from low socioeconomic contexts found that support from parents and caregivers may increase student science achievement for children (Perera, 2014). Additional research with African American students majoring in science at predominantly white universities said they persisted in their studies because their family communicated to them the value of education and held them to high standards (Nguyen et al., 2021; Russell and Atwater, 2005).

Early childhood experiences

There is increasing evidence that STEM interests and career aspirations begin early in childhood and not in the teen years, as was thought in earlier research (Auger et al., 2005; Moulton et al., 2018). As young children play, they develop an affinity for topics, such as science, that they find fun and engaging. Early engagement with science within the family comprises both family science habitus and contributes to the child's science capital.

Museums and informal STEM learning settings (e.g., after school centers, makerspaces, science media) have increasingly recognized the important role of the family in building STEM interests, and many of these centers offer play spaces related to STEM designed specifically for early childhood engagement (e.g., Museum of Life and Science in Durham, NC; Please Touch Museum in Philadelphia, PA). Although the recognition of early childhood experiences on STEM interests is growing among non-formal science centers, a survey of libraries and museums across Philadelphia showed that fewer than half identified working with young children as part of their mission and only a quarter considered it a "core service" (Sirinides et al., 2017). When parents do engage with their children in out-of-school settings, research suggests that parents support re-engagement in the science activities and help make connections of an activity with experiences and interests outside of the current activity (Pattison and Dierking, 2019).

Experiences with tools

Early childhood experiences also include opportunities beyond visiting informal STEM learning institutions. Recent research on the importance of family support of STEM interests found that parents and children have different perceptions about their access to science resources at home, with parents reporting having access to significantly more science tools in their home than their children (Ennes et al., 2022). This included tools such as rulers, measuring cups, and scales. Access to these types of foundational science tools are important for youth to develop their science process skills, and youth who do not have opportunities to develop these skills may be far behind their peers in their ability to engage in science inquiry lessons in school (Claussen and Osborne, 2013).

Experiences with tools were also seen as a vital part of a new model developed to explore the family influences on STEM aspirations (described below, [Jones et al., 2021](#)).

Families and STEM hobbies

The development of science interests in early childhood are often a result of family activities. In a study of nearly 3000 adult science hobbyists (defined as individuals engaging in science interests outside of work for pleasure), [Corin et al. \(2017\)](#) found that over half of the participants reported that their hobby began during their childhood. This study examined the development of ten different STEM hobby types. Corin et al. found that, across most of the hobby groups, the hobbyists reported their family had influenced their initial engagement with their hobby. Furthermore, the influence of the family was greater than that of friends, teachers, or other educators.

Similarly, [Jones et al. \(2019\)](#) examined the relationship between having a STEM hobby and selecting a STEM career. This study surveyed adult hobbyists and documented educational background, motivations to participate in a hobby, science identity, and influences of the hobby on career choice. Individuals who worked in STEM fields were more likely to report their career decisions were influenced by their family and their hobby interests and activities than those who did not work in STEM fields. When the STEM hobbyists who did not work in a STEM field were asked to rate factors that influenced that career decision, they reported their hobby interests began after their career decisions were made. Some of these non-STEM career hobbyists also indicated they felt they lacked the ability to pursue a career in the same area as their hobby.

Other research on the development of the interests of amateur astronomers and birders found that adult leisure interests began in childhood and were influenced by family members, having access to resources to explore the hobby (e.g., binoculars), and opportunities to experience the hobby ([Jones et al., 2017](#)). This study argued that the influences on life-long interests were driven by family members who provided support for interests, built self-efficacy for learning science, and supported motivation to continue to learn. The authors also suggest that these kinds of hobby groups provide individuals with a group identity and recognition that contributes to a sense of relatedness for STEM ([Jones et al., 2019](#)).

These studies suggest that leisure STEM activities during childhood, as well as participating in science alongside family members, teachers, and other educators influenced participants' STEM career decisions and leisure interests (see also [Dou and Cian, 2021](#)). In addition, this work showed that hobbyists who had STEM careers were more likely to report greater interests, ability, and identification with STEM subjects in school than those hobbyists who did not have a STEM career ([Jones et al., 2018](#)). These findings on hobbyists are congruent with career choice theory that maintains that career choice is influenced by interests (such as these hobby interests), outcome expectations, and self-efficacy ([Lent et al., 1994](#)). The role of leisure interests such as hobbies on the development of self-efficacy and identity for STEM has not been fully researched.

Gender differences

How parents engage their children in science can be influenced by the gender of the child (e.g., [Bhanot and Jovanovic, 2009](#); [Crowley et al., 2001](#); [Eccles, 2015](#); [Starr and Simpkins, 2021](#); [Tenenbaum and Leaper, 2003](#)). As suggested in [Fig. 3](#), [Eccles \(2015\)](#) theorized that a child's biological sex impacts the behaviors and beliefs of their parent or caregiver, which in turn affects the child's beliefs about what is possible and appropriate for them. Research on middle school students in the United States found that parents of boys were more likely to overestimate their child's ability in science than parents of girls ([Bhanot and Jovanovic, 2009](#)). This study found that when parents expressed that science was important for their daughter, the daughter valued science more; however, parent beliefs did not impact boys' value of science ([Bhanot and Jovanovic, 2009](#)). Additionally, the researchers found that when mothers had conversations with their daughters about the value of science careers and science interests, daughters valued science more highly. Other studies have found that parents explain more information to their sons when discussing science than they do with their daughters ([Crowley et al., 2001](#)). This body of work highlights the importance of family habitus that includes conversations around science particularly for families with girls.

A recent study examining the gender stereotypes that United States high school students held for math and science found that the beliefs students' held were positively correlated with their parent's beliefs ([Starr and Simpkins, 2021](#)). The parents in this study held fairly traditional beliefs that boys were better than girls in science, and by the time the students reached grade 11, the students' beliefs more closely reflected that of their parents. However, the gender stereotypes were fairly fluid between grades 9 and 11, suggesting that parents play a significant role in gender stereotypes and identities in high school ([Starr and Simpkins, 2021](#)). The authors concluded: "Parents' own gender stereotypes relate to their children's gender stereotypes and math/science identity, even as late as 11th grade" ([Starr and Simpkins, 2021](#), p. 293).

Another study in Australia found that parents of boys who were interested in science fields were more likely to have university aspirations for their children than parents of girls interested in science careers ([Lloyd et al., 2018](#)). The authors of this study suggested parents should pay close attention to the type of science role models they offer their children and that they themselves need to be aware of potential role models. The authors recommended that interventions not only target children but also parents so they can support their children's science interests regardless of gender ([Lloyd et al., 2018](#)). This supposition is supported by the work of [Archer et al. \(2013\)](#). In interviews with 10–14 year-old English girls and their parents, the researchers observed that parents expressed incongruity between "girliness" and interest in STEM careers, which were positioned in masculine terms, and that these perceptions were reflected in how their daughters thought of themselves in relation to science careers.

Among Latine youth, research has also suggested that the gender of the parent can also play a role in the ways youth, regardless of gender, engage in science related conversations and activities in the home. [Cian et al. \(2022\)](#) point to evidence suggesting that paternal caregivers tend to adopt stereotypical STEM “expert” like behaviors, such as providing factual knowledge to in-the-moment contexts, while maternal caregivers tend to identify and guide their children toward resources from which they could draw STEM related knowledge, such as by making relational connections to STEM professionals within their social networks.

Families from marginalized communities and underrepresented groups in STEM

Research across various contexts and geographies often presents family science capital as a characteristic of white or European, middle-to-upper class homes more so than the homes of children from marginalized groups in STEM fields. As alluded to earlier, Archer and colleagues’ ([Archer et al., 2015a](#)) conceptualization of science capital encompasses youths’ and families’ science related literacies, dispositions, perceptions of labor market value, practices, out-of-school participation, and social capital, which have exhibited repeated associations with demographic variables like race, ethnicity, and gender. For instance, in a study of black students and parents’ perceptions of science and science careers among children aged 10–14 in the United Kingdom, [Archer et al. \(2015b\)](#) noted that participants’ parents had less access to science material and science educational capital than their white counterparts whose children were more likely to access academic support programs, have family members who were highly educated, and understand how to navigate educational systems. While the authors noted that all families were optimistic about their children’s future, the support from black families was “often restricted to providing moral support, encouragement, guidance, and motivation, rather than subject expertise, social contacts, or expensive extracurricular activities” ([Archer et al., 2015b](#), p. 212) in relation to their participation in science activities.

Associations between family science capital and factors like race and ethnicity in particular do not always exhibit monolithic outcomes. [Wong’s \(2016\)](#) interviews with 46 students ages 11–14 revealed nuances in these relationships and the complexity of the influence of family capital in developing and maintaining career aspirations. For instance, black Caribbean boys in the study were among the few to express interest in careers as research scientists and that interest appeared to be sustained by their family capital such as science toys, books, and trips to museums. But, their interest was classified as “minimal,” meaning they actually mentioned more non-science aspirations than science aspirations and Wong noted that it was unclear where the science aspirations ranked among the students’ other career interests. Still, these students reported below-average grades in science but still held science-related career aspirations, suggesting a sustaining influence of family capital. [Jones et al.’s \(2020\)](#) validation study of a science capital survey with middle school students ($N = 889$) exhibited high correlations (0.67) with youths’ expectancy values, which included their science-related self-concepts and perceptions of how others recognize them in science, as well with youths’ perceptions of the value their families place on science and science careers. Yet, resonating with [Wong’s \(2016\)](#) study, their instrument exhibited only partial scalar invariance with regard to participants’ gender and ethnic scalar invariance, suggesting that across these demographic factors, subgroups of individuals differ in how they interpret and respond to one or more of the science capital survey items.

[Nguyen et al.’s \(2021\)](#) interviews with black women STEM students attending Historically Black Colleges or Universities (HBCUs) illustrated how parents capitalized on everyday opportunities to engage children with STEM, which “gave their children access to understanding of the importance of STEM in daily life” (p. 890). In a survey study of Hispanic youths attending a Hispanic Serving Institution (HSI), [Dou and Cian \(2022\)](#) similarly reported on the contributive effect of everyday science talk with close family members on STEM aspirations. While these two studies describe parent-child interactions that were conducive to STEM pursuits, others have noted that families communicate support for STEM aspirations non-verbally. [Avraamidou \(2022\)](#) relates the story of Zehra, a physics instructor who grew up in Turkey before being educated in the United States and moving to Europe to work. Zehra’s family communicated support for her aspirations by deciding not to oppose them, which was unusual for a girl in their cultural context.

Many of the findings presented in these and similar studies complement [Gonsalves et al. \(2021\)](#) claim that “students’ accumulation of science capital is not sufficient for engagement in postsecondary science” (p. 30). Researchers have further identified resources that families use to support child STEM aspirations that are unique to marginalized and underrepresented families. [Rodriguez et al. \(2021\)](#) observed that Latina college students attending a Predominantly White Institution (PWI) in the United States invoked *familismo*, or “a strong attachment to nuclear and extended family” (p. 1778) as a motivator for their aspirations. Students discussed feeling that their accomplishments were shared by their families, a sense of support their families communicated when they felt challenged, and a willingness to acquire a career that enabled them to provide for their families. [Yosso \(2005\)](#) introduced the concept of Community Cultural Wealth (CCW) by looking at cultural capital with a critical race theory (CRT) lens. CCW highlights six forms of capital: aspirational, or the capacity to aspire in spite of obstacles; linguistic, or fluency in multiple languages or dialects; familial, or shared cultural knowledge and history; navigational, or capacity to find one’s way through complex social systems; and resistant, or the acquisition of skills through facing inequity ([Fig. 1](#)). Empirical research has further developed these ideas. For instance, in a phenomenological study of 38 black and Latinx STEM college students in the United States, [McGee and Bentley \(2017\)](#) noted that their participants were driven to STEM careers by an “equity ethic” such that they viewed STEM careers as opportunities to serve marginalized communities and disrupt inequities—an ethic that the authors noted could be contributed to their belonging to collectivist cultures that are “likely to place a higher priority on group goals, needs, beliefs, and expectations” (p. 6) and have experienced more “violence or adversity” (p. 7, citing [Vollhardt, 2009](#), p. 53).

Another area that has received considerable attention is the family capital that is unique to families with immigration histories. Erel (2010) used a case study of three migrant women to illustrate how, rather than translocating capital from one country to another—which may or may not have value in the migrant location—new forms of capital are forged through the migration process and existing forms of capital are renegotiated for fit in the new context. Further research has supported this suggestion. A student in a study of interactions between immigrant fathers and their children in a family workshop related his father's ability to help him in science to his father's past work in construction and "on the land" (Allestaht-Snider et al., 2020), evidencing both this student's respect for his father's former career and its relationship to science and a broader concept of what could constitute as a career that supplied family science capital. Latine college students interviewed in a study by Herrera and Kovats Sánchez (2022) associated their desire to pursue STEM careers with their parents' immigration stories. However, in interviews with second-generation migrant college STEM students in the United States, Cian et al. (2022) noted that students' dismissal of their mothers' STEM expertise was related to their mothers' abandonment of a STEM career that they had held in their migrant country. Several researchers have noted that the value migrant families place on STEM careers relates to their children's science-related outcomes. Gokpinar and Reiss' (2016) study of Turkish immigrant students highlights how parental attitudes toward science contribute to children's science interests and attainment even across families with "very similar [socioeconomic] and parental, ethnic and sociological backgrounds" (p. 1294).

Modeling and predicting family influences on STEM career aspirations

Researchers have recently begun to model the role that the family plays in the development of career aspirations and interests in science. These models are based on various theoretical frameworks that have been used to assess the STEM career aspirations of youth.

Science identity

Several of these frameworks examine career aspirations from identity or self-concept perspectives. As described by Hazari et al. (2010), science identity is made up of an individual's interest (their desire to learn about science), the recognition they receive from others, their performance (an individual's belief about their ability to perform specific tasks), and their competence (an individual's belief about their ability to understand science). Students may have confidence in their academic abilities but still feel unwelcome in STEM spaces because of individual characteristics such as gender or race/ethnicity. If an environment is not purposefully inclusive, individuals can feel like they do not see themselves as working in that environment or that they lack the requisite identity to participate. STEM engagement and academic achievement can be affected when students do not identify with the STEM disciplines such as science (Kang, 2019; Singh et al., 2002). Studies have shown participation in informal STEM experiences, often with family, develops a student's scientific identity and promotes their interest as a future STEM professional (Adams and Gupta, 2013; Dasgupta and Stout, 2014; Dou et al., 2019; Goff et al., 2020). In addition to educational settings, informal experiences predicted participants' STEM identity. Simple activities such as talking with friends and family about science are predictive of STEM identity in college (Dou et al., 2019). In a study with nearly 16,000 participants examining how STEM identity, experience, and career aspirations relate, Dou et al. (2019) found that participants' probability of choosing a STEM profession increased by 85% for every one-point higher on their STEM identity scale.

Science self-concept

Closely related to identity and overlapping with self-efficacy, academic self-concept is defined as a person's perceptions about themselves and their knowledge competence in academic situations (Bong and Skaalvik, 2003; Byrne, 1984; Shavelson and Bolus, 1982; Wigfield and Karpachian, 1991). Awareness of one's academic ability is directly related to how confident the student feels in pursuing a STEM career. Studies have shown that students who also experience science outside of school with their families have a stronger positive self-perception within STEM (Kang et al., 2019). Students with positive academic self-concept reported higher persistence, goal-setting, intrinsic motivation, and STEM career aspirations (Bong and Skaalvik, 2003).

Academic self-concept is often associated with specific school subjects. In a study looking at low-income communities across the United States, Kang et al. (2019) found self-concept in science is a strong predictor of girls' STEM-related career aspirations. Self-concept in mathematics and science self-concept are regarded as significant indicators of a student's professional pursuit across all STEM fields (Bong and Skaalvik, 2003; Kang et al., 2019; Sax et al., 2015). Overall, a student's academic self-concept, coupled with a personal interest in STEM, increases their probability of choosing a career within STEM (Adams and Gupta, 2013).

Social cognitive career theory

One of the earliest theories applied to career aspirations is that of social cognitive career theory (SCCT), proposed by Lent et al. (1994). This theory suggests that beliefs related to self-efficacy, outcome expectations, and goal-related beliefs influence career decisions. This theory has been applied in a variety of studies in different contexts (e.g., Byars-Winston and Rogers, 2019; Dutta et al.,

2015). SCCT has been used to predict the role of interests, social supports (e.g., family), social barriers, self-efficacy, and outcome expectations (Lent et al., 2008) on career aspirations.

Expectancy-value theory

Building on similar expectation-related constructs that are included in SCCT, expectancy value theory (EVT) has emerged as a useful and valid tool to predict achievement-related choices such as career selection (Eccles and Wigfield, 2020). This comprehensive theory allows researchers to account for individual factors such as aptitude, gender, race/ethnicity, prior experiences, and cultural variables such as family demographics and social role systems as well as between-person variables such as perceptions of socializer's beliefs, social roles, affective reactions and memories, as well as self-schemata (identity and self-concept). According to this model, these in-person and between-person variables contribute to an individual's subjective task value, or the perception of the attainment value, utility value, and costs for making an achievement-related choice such as a career choice. This model has become widely applied and validated in studies across different disciplines and contexts (e.g., Ball et al., 2016; Fielding-Wells, 2017; Watkinson et al., 2005).

Modeling STEM career aspirations

Recent research on the influence of the family on students' career aspirations (Jones et al., 2021, 2020) applied constructs of science capital and habitus with factors of expectancy value theory. Using a national sample and structural equation modeling, Jones et al. (2021) found science capital factors (Tool Access, STEM Experiences, and access to STEM practitioners) were influential in predicting students' Science Achievement Value and students' perceptions of Family Science Achievement Value (the degree students thought their families valued science) (Fig. 4).

This model further showed that Family Science Achievement Value along with Science Achievement Value are associated with Future Science Task Value (i.e., career aspirations). The study examined how those out-of-school variables such as access to tools and knowing people who worked in STEM fields influenced students' perceptions of the value of science for their future. Models such as these described above allow us to predict and identify career choice variables and to design new forms of intervention that can support students' budding interests in science and other related fields.

Assessing family influence and STEM career aspirations

In this final section we describe a selection of instruments that specifically assess familial influences on STEM career aspirations. The following surveys were strategically chosen as they are all validated measures.

ASPIRES

The ASPIRES survey (Archer and DeWitt, 2013) was validated and was developed to improve researchers' understanding of youth science career aspirations. Youth in the United Kingdom completed Likert scale survey items on a variety of evaluative topics including, but not limited to, aspirations in science, attitudes toward school science, images of scientists, and parental expectations. The target population for this survey is young people ages 10–14.

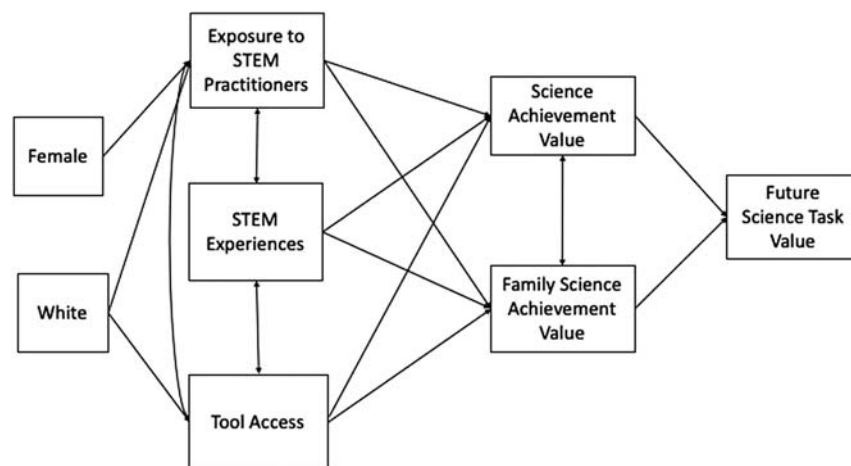


Fig. 4 Structural equation model depicting factors that predict Future Science Task Value. Adapted from Jones et al. (2021).

NextGen Scientist Survey

The NextGen Scientist Survey is a tool that measures the accessibility of factors related to science capital and habitus (Jones et al., 2021, 2020). More specifically, the aspects of the survey include science experiences, science achievement value, perceptions of family science achievement value, and future science task value (Jones et al., 2021). The NextGen Scientist Survey has been validated for upper elementary students (age 8–10; Jones et al., 2022) and middle grades (age 11–13; Jones et al., 2021, 2021) in the United States. The validation of the parent version of the survey is forthcoming.

S-STEM

The S-STEM survey (Unfried et al., 2015) measures attitudes toward and career interest in STEM. The survey has two versions that have been validated for upper elementary students (age 8–10) as well as for middle and high school (age 11–18). The survey is divided into four sections that assesses STEM attitudes, confidence, performance expectations, career aspirations, and access to STEM professionals.

TOSRA

The Test of Science-Related Attitudes (TOSRA; Fraser, 1978) assesses students' (secondary school students, age 12–15) attitudes toward science. TOSRA is composed of several domains which include enjoyment of science lessons, leisure interest in science, and career interest in science, just to name a few. It can be used among both students and caregivers (Fraser, 1981).

Conclusion

As noted in this chapter, research has shown that youth do not develop science interests or STEM career aspirations on their own. Rather, they form it in a family context. Families play an important role in helping their children develop their ideas of what they could, should, or might do in the future as it relates to science and STEM careers (Archer et al., 2012). This is done through early childhood experiences, out-of school family activities, engagement with tools, and exposure to STEM hobbies. These experiences have been shown to influence the self-efficacy (e.g., Bandura, 1978), identity (e.g., Hazari et al., 2010), and academic self-concept (e.g., Bong and Skaalvik, 2003) of youth related to science. Research is emerging that shows that by engaging children in interventions that also involve the whole family, the child's developing science interests and career aspirations are supported (Ennes et al., under review; Ennes and Jones, 2021). These interventions can be examined using several theoretical frameworks including social cognitive career theory (e.g., Lent et al., 1994) and expectancy-value theory (Eccles and Wigfield, 2020). Recently, new models have also been developed to better understand the family influences on STEM career aspirations (e.g., Jones et al., 2021). Families are responsible for the early introduction of children to science which can support lifelong interests and the development of career aspirations (e.g., Cheryan et al., 2017; Jones et al., 2016). Given the increasing dependency of society on science and technology, children, families, and the broader society can benefit from a focus on family engagement in science as these efforts enhance STEM literacy and support formal and informal interests (Sharon and Baram-Tsabari, 2020; Vahey et al., 2019).

References

- Adams, J.D., Gupta, P., 2013. "I learn more here than I do in school. Honestly, I wouldn't lie about that": creating a space for agency and identity around science. *Int. J. Crit. Pedagogy* 4 (2), 87–104. <http://partnershipsjournal.org/index.php/ijcp/article/viewFile/401/369>.
- Allexsaht-Snider, M., Vazquez Dominguez, M., Buxton, C., Karsli, E., 2020. Figured worlds of immigrant fathers, sons, and daughters in steps to college through science bilingual family workshops. *Gend. Educ.* 32 (3), 311–327. <https://doi.org/10.1080/09540253.2017.1343936>.
- Archer, L., DeWitt, J., 2013. ASPIRES: Young People's Science and Career Aspirations, Age 10–14. Final Project Report. Department of Education and Professional Studies, King's College London, London.
- Archer, L., DeWitt, J., 2016. Understanding Young People's Science Aspirations: How Students Form Ideas About "Becoming a Scientist.". Routledge, London. <https://doi.org/10.4324/9781315761077>.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., Wong, B., 2012. Science aspirations, capital, and family habitus: how families shape children's engagement and identification with science. *Am. Educ. Res. J.* 49 (5), 881–908. <https://doi.org/10.3102/0002831211433290>.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., Wong, B., 2013. 'Not girly, not sexy, not glamorous': primary school girls' and parents' constructions of science aspirations. *Pedagogy Cult. Soc.* 21 (1), 171–194. <https://doi.org/10.1080/14681366.2012.748676>.
- Archer, L., Dawson, E., DeWitt, J., Seakins, A., Wong, B., 2015a. Science capital: a conceptual, methodological, and empirical argument for extending Bourdieusian notions of capital beyond the arts. *J. Res. Sci. Teach.* 52 (7), 922–948. <https://doi.org/10.1002/tea.21227>.
- Archer, L., DeWitt, J., Osborne, J., 2015b. Is science for us? Black students' and parents' views of science and science careers. *Sci. Educ.* 99 (2), 199–237. <https://doi.org/10.1002/sce.21146>.
- Auger, R.W., Blackhurst, A.E., Wahl, K.H., 2005. The development of elementary-aged children's career aspirations and expectations. *Prof. Sch. Counsel.* 8 (4), 322–329. <https://www.jstor.org/stable/42732626?seq=1>.
- Avramidou, L., 2022. Identities in/out of physics and the politics of recognition. *J. Res. Sci. Teach.* 59 (1), 58–94. <https://doi.org/10.1002/tea.21721>.
- Ball, C., Huang, K.T., Cotten, S.R., Rikard, R.V., Coleman, L.O., 2016. Invaluable values: an expectancy-value theory analysis of youths' academic motivations and intentions. *Inf. Commun. Soc.* 19 (5), 618–638. <https://doi.org/10.1080/1369118X.2016.1139616>.
- Bandura, A., 1978. Self-efficacy: toward a unifying theory of behavioral change. *Adv. Behav. Res. Ther.* 1 (4), 139–161. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4).

- Bhanot, R.T., Jovanovic, J., 2009. The links between parent behaviors and boys' and girls' science achievement beliefs. *Appl. Dev. Sci.* 13 (1), 42–59. <https://doi.org/10.1080/10888690802606784>.
- Bong, M., Skaalvik, E.M., 2003. Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15. <https://doi.org/10.1023/A:1021302408382>.
- Bourdieu, P., 1973. *Cultural Reproduction and Social Reproduction*. Tavistock, London, p. 178.
- Bourdieu, P., 1977. *Outline of a Theory of Practice* (No. 16). Cambridge University Press, Cambridge.
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J. (Ed.), *Handbook of Theory and Research for the Sociology of Education*. Greenwood Press, New York, pp. 241–258.
- Byars-Winston, A., Rogers, J.G., 2019. Testing intersectionality of race/ethnicity $\times$ gender in a social-cognitive career theory model with science identity. *J. Counsel. Psychol.* 66 (1), 30. <https://psycnet.apa.org/doi/10.1037/cou0000309>.
- Byrne, B.M., 1984. The general/academic self-concept nomological network: a review of construct validation research. *Rev. Educ. Res.* 54 (3), 427–456. <https://doi.org/10.3102/00346543054003427>.
- Chandler, B., 2013. The subjectivity of habitus. *J. Theor. Soc. Behav.* 43 (4), 469–491. <https://doi.org/10.1111/jtsb.12025>.
- Cheryan, S., Ziegler, S.A., Montoya, A.K., Jiang, L., 2017. Why are some STEM fields more gender balanced than others? *Psychol. Bull.* 143 (1), 1–35. <https://doi.org/10.1037/bul0000052>.
- Cian, H., Dou, R., Castro, S., Palma-D'souza, E., Martinez, A., 2022. Facilitating marginalized youths' identification with STEM through everyday science talk: the critical role of parental caregivers. *Sci. Educ.* 106 (1), 57–87. <https://doi.org/10.1002/sce.21688>.
- Claussen, S., Osborne, J., 2013. Bourdieu's notion of cultural capital and its implications for the science curriculum. *Sci. Educ.* 97 (1), 58–79. <https://doi.org/10.1002/sce.21040>.
- Corin, E.N., Jones, M.G., Andre, T., Childers, G.M., Stevens, V., 2017. Science hobbyists: active users of the science-learning ecosystem. *Int. J. Sci. Educ.* 7 (2), 161–180. <https://doi.org/10.1080/21548455.2015.1118664>.
- Crowley, K., Callanan, M.A., Tenenbaum, H.R., Allen, E., 2001. Parents explain more often to boys than to girls during shared scientific thinking. *Psychol. Sci.* 12 (3), 258–261. <https://doi.org/10.1111/1467-9280.00347>.
- Dabney, K.P., Chakraverty, D., Tai, R.H., 2013. The association of family influence and initial interest in science. *Sci. Educ.* 97 (3), 395–409. <https://doi.org/10.1002/sce.21060>.
- Dasgupta, N., Stout, J.G., 2014. Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Pol. Insights Behav. Brain Sci.* 1 (1), 21–29. <https://doi.org/10.1177/2372732214549471>.
- Dou, R., Cian, H., 2021. The relevance of childhood science talk as a proxy for college students' STEM identity at a Hispanic Serving Institution. *Res. Sci. Educ.* 50, 1–13. <https://doi.org/10.1007/s11165-020-09928-8>.
- Dou, R., Cian, H., 2022. Constructing STEM identity: an expanded structural model for STEM identity research. *J. Res. Sci. Teach.* 59, 458–490. <https://doi.org/10.1002/tea.21734>.
- Dou, R., Hazari, Z., Dabney, K., Sonnert, G., Sadler, P., 2019. Early informal STEM experiences and STEM identity: the importance of talking science. *Sci. Educ.* 103 (3), 623–637. <https://doi.org/10.1002/sce.21499>.
- Dumais, S.A., 2002. Cultural capital, gender, and school success: the role of habitus. *Sociol. Educ.* 75 (1), 44–68. <https://doi.org/10.2307/3090253>.
- Dutta, A., Kang, H.J., Kaya, C., Benton, S.F., Sharp, S.E., Chan, F., da Silva Cardoso, E., Kundu, M., 2015. Social-cognitive career theory predictors of STEM career interests and goal persistence in minority college students with disabilities: a path analysis. *J. Vocat. Rehabil.* 43 (2), 159–167. <https://psycnet.apa.org/doi/10.3233/JVR-150765>.
- Eccles, J.S., Wigfield, A., 2020. From expectancy-value theory to situated expectancy-value theory: a developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* 61, 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>.
- Eccles, J.S., 2015. Gendered socialization of STEM interests in the family. *Int. J. Gen. Sci. Technol.* 7 (2), 116–132. <http://genderandset.open.ac.uk/index.php/genderandset/article/view/419>.
- Ennes, M., Jones, M.G., Childers, G., Cayton, E., Chesnutt, K., (2022, accepted). Children and parents' perceptions of access to science tools at home and their role in science self-efficacy.
- Ennes, M., Jones, M.G., Childers, G.M., Cayton, E.M., Chesnutt, K.M., (under review). Tools of the Trade: How Science Tool Experiences Contribute to Youths' Future Science Interests.
- Ennes, M., 2019. *Building Science Capital and Family Habitus Using a Systems Approach*. Doctoral dissertation, North Carolina State University. <http://www.lib.ncsu.edu/resolver/1840.20/36356>.
- Ennes, M., Jones, M.G., 2021. Building a science community: family science programs. *Science Scope* 4 (4), 29–35. <https://www.nsta.org/science-scope/science-scope-marchapril-2021/building-science-community>.
- Erel, U., 2010. Migrating cultural capital: Bourdieu in migration studies. *Sociology* 44 (4), 642–660. <https://doi.org/10.1177/0038038510369363>.
- Fielding-Wells, J., O'Brien, M., Makar, K., 2017. Using expectancy-value theory to explore aspects of motivation and engagement in inquiry-based learning in primary mathematics. *Math. Educ. Res. J.* 29 (2), 237–254. <https://doi.org/10.1007/s13394-017-0201-y>.
- Fisher, T.A., Padmawidjaja, I., 1999. Parental influences on career development perceived by African American and Mexican American college students. *J. Multicult. Counsel. Dev.* 27 (3), 136–152. <https://doi.org/10.1002/j.2161-1912.1999.tb00220.x>.
- Fraser, B.J., 1978. Development of a test of science-related attitudes. *Sci. Educ.* 62, 509–515. <https://doi.org/10.1002/sce.3730620411>.
- Fraser, B.J., 1981. *Test of Science-Related Attitudes (TOSRA) Handbook*. Australian Council for Educational Research.
- Goff, E.E., Mulvey, K.L., Irvin, M.J., Hartstone-Rose, A., 2020. The effects of prior informal science and math experiences on undergraduate STEM identity. *Res. Sci. Technol. Educ.* 38 (3), 272–288. <https://doi.org/prox.lib.ncsu.edu/10.1007/s11162-015-9375-x>.
- Gokpinar, T., Reiss, M., 2016. The role of outside-school factors in science education: a two-stage theoretical model linking Bourdieu and Sen, with a case study. *Int. J. Sci. Educ.* 38 (8), 1278–1303. <https://doi.org/10.1080/09500693.2016.1188332>.
- Gonsalves, A.J., Cavalcante, A.S., Sprowls, E.D., Iacono, H., 2021. "Anybody can do science if they're brave enough": understanding the role of science capital in science majors' identity trajectories into and through postsecondary science. *J. Res. Sci. Teach.* 58 (8), 1117–1151. <https://doi.org/10.1002/tea.21695>.
- Guan, P., Capezio, A., Restubog, S.L.D., Read, S., Lajom, J.A.L., Li, M., 2016. The role of traditionality in the relationships among parental support, career decision-making self-efficacy and career adaptability. *J. Vocat. Behav.* 94, 114–123. <https://doi.org/10.1016/j.jvb.2016.02.018>.
- Hazari, Z., Sonnert, G., Sadler, P.M., Shanahan, M.C., 2010. Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study. *J. Res. Sci. Teach.* 47 (8), 978–1003. <https://www.jstor.org/stable/43631586>.
- Herrera, F., Kovats Sánchez, G., 2022. Curando la comunidad [Healing the community]: community-centered STEM identity. *J. Hisp. High. Educ.* <https://doi.org/10.1177/15381927211069543>, 15381927211069543.
- Jones, M.G., Corin, E.N., Andre, T., Childers, G.M., Stevens, V., 2016. Factors contributing to lifelong science learning: amateur astronomers and birders. *J. Res. Sci. Teach.* 54 (3), 412–433. <https://doi.org/10.1002/tea.21371>.
- Jones, M.G., Corin, E., Andre, T., Childers, G., Stevens, V., 2017. Factors contributing to lifelong science learning: amateur astronomers and birders. *J. Res. Sci. Teach.* 54 (3), 412–433. <https://doi.org/10.1002/tea.21371>.
- Jones, M.G., Childers, G., Andre, T., Corin, E.N., Hite, R., 2018. Citizen scientists and non-citizen scientist hobbyists: motivation, benefits, and influences. *Int. J. Sci. Educ.* 8 (4), 287–306. <https://doi.org/10.1080/21548455.2018.1475780>.
- Jones, M.G., Childers, G., Corin, E., Chesnutt, K., Andre, T., 2019. Free choice learning and STEM career choice. *Int. J. Sci. Educ.* 9 (1), 29–39. <https://doi.org/10.1080/21548455.2018.1534024>.
- Jones, M.G., Ennes, M., Weedfall, D., Chesnutt, K., Cayton, E., 2020. The development and validation of a measure of science capital, habitus, and future science interests. *Res. Sci. Educ.* 51 (6), 1549–1565. <https://doi.org/10.1007/s11165-020-09916-y>.

- Jones, M.G., Chesnutt, K., Ennes, M., Mulvey, K.L., Cayton, E., 2021. Understanding science career aspirations: factors predicting future science task value. *J. Res. Sci. Teach.* 58 (7), 937–955. <https://doi.org/10.1002/tea.21687>.
- Jones, M.G., Chesnutt, K., Ennes, M., Macher, D., Paechter, M., 2022. Measuring Science Capital, Science Attitudes, and Science Experiences in Elementary and Middle School Students.
- Kang, H., Calabrese Barton, A., Tan, E., Simpkins, S.D., Rhee, H.-Y., Turner, C., 2019. How do middle school girls of color develop STEM identities? Middle school girls' participation in science activities and identification with STEM careers. *Sci. Educ.* 103, 418–439. <https://doi.org/10.1002/sce.21492>.
- Kier, M.W., Blanchard, M.R., 2021. Eliciting students' voices through STEM career explorations. *Int. J. Sci. Math. Educ.* 19 (1), 151–169. <https://doi.org/10.1007/s10763-019-10042-z>.
- Kiyama, J.M., 2010. College aspirations and limitations: the role of educational ideologies and funds of knowledge in Mexican American families. *Am. Educ. Res. J.* 47 (2), 330–356. <https://doi.org/10.3102/0002831209357468>.
- Lent, R.W., Brown, S.D., Hackett, G., 1994. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45 (1), 79–122. <https://doi.org/10.1006/jvbe.1994.1027>.
- Lent, R.W., Lopez Jr., A.M., Lopez, F.G., Sheu, H.B., 2008. Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *J. Vocat. Behav.* 73 (1), 52–62. <https://doi.org/10.1016/j.jvb.2008.01.002>.
- Lizardo, O., 2004. The cognitive origins of Bourdieu's habitus. *J. Theor. Soc. Behav.* 34 (4), 375–401. <https://doi.org/10.1111/j.1468-5914.2004.00255.x>.
- Lloyd, A., Gore, J., Holmes, K., Smith, M., Fray, L., 2018. Parental influences on those seeking a career in STEM: the primacy of gender. *Int. J. Gend. Sci. Technol.* 10 (2), 308–328. <http://genderandset.open.ac.uk/index.php/genderandset/article/view/510>.
- McClelland, K., 1990. Cumulative disadvantage among the highly ambitious. *Sociol. Educ.* 63 (2), 102–121. <https://doi.org/10.2307/2112857>.
- McGee, E., Bentley, L., 2017. The equity ethic: black and Latinx college students reengineering their STEM careers toward justice. *Am. J. Educ.* 124 (1), 1–36. <https://psycnet.apa.org/doi/10.1086/693954>.
- Moote, J., Archer, L., DeWitt, J., MacLeod, E., 2020. Science capital or STEM capital? Exploring relationships between science capital and technology, engineering, and maths aspirations and attitudes among young people aged 17/18. *J. Res. Sci. Teach.* 57 (8), 1228–1249. <https://doi.org/10.1002/tea.21628>.
- Moulton, V., Flouri, E., Joshi, H., Sullivan, A., 2018. Individual-level predictors of young children's aspirations. *Res. Pap. Educ.* 33 (1), 24–41. <https://doi.org/10.1080/02671522.2016.1225797>.
- Nguyen, T.H., Gasman, M., Washington Lockett, A., Peña, V., 2021. Supporting Black women's pursuits in STEM. *J. Res. Sci. Teach.* 58 (6), 879–905. <https://doi.org/10.1002/tea.21682>.
- Pattison, S.A., Dierking, L.D., 2019. Early childhood science interest development: variation in interest patterns and parent–child interactions among low-income families. *Sci. Educ.* 103 (2), 362–388. <https://doi.org/10.1002/sce.21486>.
- Perera, L.D.H., 2014. Parents' attitudes towards science and their children's science achievement. *Int. J. Sci. Educ.* 36 (18), 3021–3041. <https://doi.org/10.1080/09500693.2014.949900>.
- Reay, D., 1995. "They employ cleaners to do that": habitus in the primary classroom. *Br. J. Sociol. Educ.* 16 (3), 353–371. <https://doi.org/10.1080/0142569950160305>.
- Rincón, B., Fernández, É., Hinojosa, J.K., 2020. "I wanted to follow in her footsteps": activating, nurturing, and extending community cultural wealth for students of color entering STEM pathways. *Teach. Coll. Rec.* 122 (9), 1–36. <https://doi.org/10.1177/016146812012200903>.
- Rodríguez, S., Pilcher, A., García-Tellez, N., 2021. The influence of familismo on Latina student STEM identity development. *J. Latinos Educ.* 20 (2), 177–189. <https://doi.org/10.1080/15348431.2019.1588734>.
- Russell, M.L., Atwater, M.M., 2005. Traveling the road to success: a discourse on persistence throughout the science pipeline with African American students at a predominantly white institution. *J. Res. Sci. Teach.* 42 (6), 691–715. <https://doi.org/10.1002/tea.20068>.
- Sawitri, D.R., Creed, P.A., Zimmer-Gembeck, M.J., 2015. Longitudinal relations of parental influences and adolescent career aspirations and actions in a collectivist society. *J. Res. Adolesc.* 25 (3), 551–563. <https://doi.org/10.1111/jora.12145>.
- Sax, L.J., Kanmy, M.A., Riggers-Piehl, T.A., Whang, H., Paulson, L.N., 2015. "But I'm not good at math": the changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Res. High. Educ.* 56 (8), 813–842. <https://doi.org/10.1007/s11162-015-9375-x>.
- Schavelson, R.J., Bolus, R., 1982. Self-concept: the interplay of theory and models. *J. Educ. Psychol.* 74 (1), 3–17. <https://psycnet.apa.org/doi/10.1037/0022-0663.74.1.3>.
- Sha, L., Schunn, C., Bathgate, M., Ben-Eliyahu, A., 2016. Families support their children's success in science learning by influencing interest and self-efficacy. *J. Res. Sci. Teach.* 53 (3), 450–472. <https://doi.org/10.1002/tea.21251>.
- Sharon, A.J., Baram-Tsabari, A., 2020. Can science literacy help individuals identify misinformation in everyday life? *Sci. Educ.* 104 (5), 873–894. <https://doi.org/10.1002/sce.21581>.
- Silander, M., Grindal, T., Hupert, N., Garcia, E., Anderson, K., Vahey, P., Pasnik, S., 2018. What Parents Talk About When They Talk About Learning: A National Survey About Young Children and Science. Education Development Center, Inc. <https://eric.ed.gov/?id=ED603163>.
- Silva, E.B., 2016. Habitus: beyond sociology. *Socio. Rev.* 64 (1), 73–92. <https://doi.org/10.1111/1467-954X.12345>.
- Singh, K., Granville, M., Dika, S., 2002. Mathematics and science achievement: effects of motivation, interest, and academic engagement. *J. Educ. Res.* 95 (6), 323–332. <https://doi.org/10.1080/00220670209596607>.
- Sirinides, P., Fink, R., DuBois, T., 2017. A study of early learning services in museums and libraries. *Early Child. Educ. J.* 45 (4), 563–573. <https://doi.org/10.1007/s10643-016-0820-z>.
- Starr, C.R., Simpkins, S.D., 2021. High school students' math and science gender stereotypes: relations with their STEM outcomes and socializers' stereotypes. *Soc. Psychol. Educ.* 24 (1), 273–298. <https://doi.org/10.1007/s11218-021-09611-4>.
- Tenenbaum, H.R., Leaper, C., 2003. Parent-child conversations about science: the socialization of gender inequities? *Dev. Psychol.* 39 (1), 34–47. <https://doi.org/10.1037/0012-1649.39.1.34>.
- Tillman, K.H., Nam, C.B., 2008. Family structure outcomes of alternative family definitions. *Popul. Res. Pol. Rev.* 27 (3), 367–384. <https://doi.org/10.1007/s11113-007-9067-0>.
- Unfried, A., Faber, M., Stanhope, D.S., Wiebe, E., 2015. The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *J. Psychoeduc. Assess.* 33 (7), 622–639. <https://doi.org/10.1177/0734282915571160>.
- Vahey, P., Vidiksis, R., Adair, A., 2019. Increasing science literacy in early childhood: the connection between home and school. *Am. Educ.* 42 (4), 17. <https://eric.ed.gov/?id=EJ1200687>.
- Verdín, D., Godwin, A., Ross, M., 2018. Stem roles: how students' ontological perspectives facilitate STEM identities. *J. Pre-Coll. Eng. Educ. Res.* 8 (2), 31–48. <https://doi.org/10.7771/2157-9288.1167>.
- Vollhardt, J.R., 2009. Altruism born of suffering and prosocial behavior following adverse life events: a review and conceptualization. *Soc. Justice Res.* 22 (1), 53–97. <https://doi.org/10.1007/s11211-009-0088-1>.
- Watkinson, E.J., Dwyer, S., Nielsen, A.B., 2005. Children theorize about reasons for recess engagement: does expectancy-value theory apply? *Adapt. Phys. Act. Q.* 22 (2). <https://doi.org/10.1123/apaq.22.2.179>.
- Wigfield, A., Karpachian, M., 1991. Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educ. Psychol.* 26 (3–4), 233–261. <https://doi.org/10.1080/00461520.1991.9653134>.

- Wong, B., 2016. Minority ethnic students and science participation: a qualitative mapping of achievement, aspiration, interest and capital. *Res. Sci. Educ.* 46 (1), 113–127. <https://doi.org/10.1007/s11165-015-9466-x>.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91. <https://doi.org/10.1080/1361332052000341006>.
- Young, J.R., Ortiz, N., Young, J.L., 2017. STEMulating interest: a meta-analysis of the effects of out-of-school time on student STEM interest. *Int. J. Educ. Math. Sci. Technol.* 5 (1), 62–74. <https://www.ijemst.org/index.php/ijemst/article/view/109>.

Secondary school students' STEM career aspirations

Kathryn Holmes, Center for Educational Research, School of Education, Western Sydney University, Kingswood, NSW, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	382
Formation of STEM aspirations	382
Factors related to STEM aspirations	383
Individual factors	383
Other influences	383
Theoretical perspectives	384
Recommendations for further research	385
References	386

Introduction

In the US alone STEM jobs are predicted to grow by 8.8% between 2017 and 2029 (Zilberman and Ice, 2021). Similarly, in Australia, between 2014 and 2019, employment in STEM occupations grew by 19.7% which is 1.9 times higher than other occupational groups (Department of Education, Skills and Employment, 2020). This expected growth is offset by skills shortages in STEM areas, leading to employer difficulties in filling positions (Department of Education, Skills and Employment, 2020). Difficulties in attracting interest in STEM is further exacerbated by lower participation rates in some STEM areas (for example, computing and engineering) for females and particular cultural and ethnic groups (Fry et al., 2021).

The participation rates of females in STEM higher education and in some STEM careers consistently lags behind that of males. In Australia women made up less than a quarter of students studying in STEM fields in 2019 and 5 years after graduation males with a STEM qualification were 1.8 times more likely to be working in a STEM-qualified occupation compared to their female peers (Department of Industry, 2021). Globally there are more women than men enrolled in higher education, but these women are less likely to enroll in STEM fields, particularly in engineering, computer science and physics (The World Bank, 2021).

Aspirations for future study and careers begin forming in childhood and trajectories are often solidified in early adolescence (Berger et al., 2020). It is important, therefore, to understand the factors that can support or undermine aspirations in the STEM fields during the secondary school years. Some factors are related to individuals and their concept of themselves as STEM learners, while others sit at the school or community levels, interacting with individual factors to enable or disable the development of strong STEM aspirations throughout the school years.

Formation of STEM aspirations

The STEM acronym represents a broad range of career options, with some research studies focusing solely on careers in science, technology, engineering and mathematics (Holmes et al., 2018; Luo et al., 2021), while others include careers in related fields such as health care and medicine (Sadler et al., 2012; Sáinz and Müller, 2018). Similarly, some studies focus on attitudes toward STEM subjects and skills generally during school (Atabey and Topcu, 2021) while others examine particular STEM subjects and their influence on STEM aspirations (Ardies et al., 2015; Blotnick et al., 2018). This wide variation in the interpretation of what STEM represents presents a level of complexity when trying to disentangle the relevant factors at play in shaping aspirations for STEM careers.

Aspirations for STEM careers are seeded during the school years. In a longitudinal study (2012–2015) of 6492 students in Years 3–12 in Australia, STEM aspirations were examined in relation to a broad range of student and school level factors (Holmes et al., 2018). The study was shaped by a comprehensive theoretical framework drawing on notions of human, social, cultural and economic capital and operationalized at the student and the school level. In this study STEM careers were defined as those requiring a tertiary qualification in science, technology, engineering or mathematics and of the total sample 894 (13.8%) students expressed a desire for this type of career. However, these aspirations were unevenly distributed in number, and by gender, across the four STEM domains with 553 (47% female) students expressing interest in a science career, 207 (6% female) in a technology career, 396 (91% female) in engineering and only 8 (33% female) in a mathematics career. Male students were three times more likely to aspire to a STEM career than females, but aspirations for STEM careers did increase in the older cohorts of students compared to the younger cohorts. Students in the highest quartile for prior achievement in reading and numeracy were also more likely to aspire to STEM careers. Remarkably, student socio-economic status or Indigenous status were not significant variables in the presence of other predictors. However, the study did find that students with a parent in a STEM occupation were more likely to aspire to a STEM career, as were those in the higher quartiles for cultural capital (a measure based on how frequently students attended cultural events, galleries, museums and libraries) (Holmes et al., 2018).

In the US, ethnicity has emerged as a factor related to STEM aspirations with Asian students aspiring to STEM in greater proportions than other racial groups (Edwin et al., 2019). This finding was reinforced by a large study of US ninth grade students, which indicated that STEM aspirations were influenced by race, gender, socioeconomic status, math interest and science self-efficacy (Mau and Li, 2018). The authors found that being male, white, having a higher SES, having greater parental educational expectations, higher maths interest and science self-efficacy was related to aspirations for STEM careers. Female student aspirations were found to be hindered by a lack of confidence, unfair treatment and discrimination, financial pressures and biased stereotypes (Mau and Li, 2018).

STEM aspirations are influenced by factors that can be broadly categorized as being related to the student at the individual, family, school or broader community levels. At the individual level, student demographic factors and their attitudes to STEM subjects have been examined in relation to their aspirations (Ahmed and Mudrey, 2019). At the family level, the influence of parents is recognized as being an important influential factor in the formation of young peoples' aspirations for STEM careers (Lloyd et al., 2018). At the school level various practices can have an influence on student attitudes to STEM (Reinhold et al., 2018) and at the broader community level there has been a strong focus on stereotypes that are prevalent in relation to the study of STEM subjects and STEM careers (Makarova et al., 2019). The influence of these proximal and distal factors will be discussed in the following sections.

Factors related to STEM aspirations

Individual factors

Numerous studies have examined student attitudes to STEM careers and STEM subjects at the secondary school level. Many of these studies have drawn on existing datasets generated by international testing schemes such as PISA (Program for International Student Assessment). Using PISA data a US study of 5611 15 year olds found that students who enjoy science and perceive it as useful were more likely to aspire to STEM careers (Ahmed and Mudrey, 2019). Similarly, using PISA data from the United Arab Emirates, Cairns and Dickson (2021) found that STEM career aspirations for both genders were related to enjoyment of science, science self-efficacy, interest in science and instrumental motivation. It matters, therefore, how students feel about science while at school, with more positive attitudes being related to the development of aspirations for future STEM engagement.

Of the STEM subjects, attitudes toward, and participation in mathematics has been consistently found to be related to STEM aspirations (Sax et al., 2015; Wang et al., 2017). In a large Canadian study of 1448 students in grades 7–9, those with low mathematics self-efficacy were found to have declining interest in STEM careers over time (Blotnick et al., 2018). In a Finnish study of elementary students' motivation across three subjects (Finnish languages, science and mathematics), those who were "math-motivated" were more strongly associated with aspirations for STEM occupations (Oppermann et al., 2021). In addition to attitudes toward mathematics, the level of mathematics studied in school has also been found to have a major influence on STEM aspirations. In a very large ($n = 650,000$) Israeli study over 15 years, the choice of more advanced levels of mathematics in school, rather than the level of success in those levels, was found to be the best predictor of choosing STEM as a major in secondary school, completing a STEM bachelor's degree and succeeding in that degree, and of making a STEM career choice (Kohen and Nitzan, 2022). Helping students to aspire to higher levels in mathematics during school and keeping them motivated and confident in mathematics, therefore, can better prepare them to aspire to STEM careers. However, by contrast, there is evidence that the relationship between mathematics interest and careers in the health area is not as strong as for other STEM careers (Wiebe et al., 2018).

STEM aspirations are also influenced by attitudes toward science as a school subject (Mau and Li, 2018; Sahin and Waxman, 2021). In a large study of almost 3000 middle school students in California, students who were "activated" toward science were more likely to have a greater affinity toward STEM careers (Dorph et al., 2018). In this study science learning activation was operationalized a multi-faceted construct including affinity, fascination, valuing, competency beliefs and scientific sensibility. Interestingly, the valuing of science was correlated more with career aspirations in the fields of technology and engineering, than careers in mathematics and science (Dorph et al., 2018). While attitudes to science generally are important for STEM aspirations, the nature of that science interest is also important, as it has also been found that those interested in physics at the start of secondary school had the greatest retention in STEM study throughout school (Sadler et al., 2012). Also, differences have been found in the factors related to intention to study the various science disciplines such as biology, chemistry and physics. In a study of 125 adolescent girls, the intention to study biology was predicted by self-efficacy in biology and social support from teachers. In this study, intentions to study chemistry and physics were predicted by self-efficacy in each subject respectively, but also by the degree to which the students endorsed a growth mindset (Mackenzie et al., 2021).

Other influences

Student attitudes are strongly related to the formation of STEM aspirations and are often the key determinant in whether a student pursues STEM study and ultimately a STEM career. These attitudes, however, are influenced by a wide range of factors including family, schooling, and community attitudes.

At the family level there is evidence that student STEM aspirations are positively related to the STEM occupation of parents (Holmes et al., 2018) and that fathers are more likely to have a STEM qualification than mothers (Department of Industry, 2021). The influence of family members could also be dependent on context, with mothers found to be more influential on student

aspirations for STEM in Turkey when compared to the United States (Bahar and Adiguzel, 2016). The level of education of mothers has also been found to be related to STEM aspirations in a Spanish study of almost 800 students (Sáinz and Müller, 2018). In this study, boys with mothers with intermediate education levels were more likely to aspire to STEM careers in architecture and technology, and girls with mothers with high levels of education were more likely to aspire to STEM-related health careers. In a study of Australian students, students aspiring to STEM careers were typically higher achieving than their peers and more than 90% of their parents had aspirations for them to attend university (Lloyd et al., 2018). Parents have also been found to be influential in the subject choices that students make at school (Sahin and Waxman, 2021) which is important given that subject choice in school has been found to be related to retention in STEM in the long term (Kohen and Nitzan, 2022).

Schools have a significant impact on the formation of aspirations for students' future study and career choices. Students in particular types of schools have been found to hold STEM aspirations, particularly if they were male, in a private school with high socioeconomic status and with high mathematics self-efficacy (Ketenci et al., 2020). A further study using PISA data from 55 countries confirmed the relative advantage of students in private schools in developing STEM aspirations, although this advantage was stronger for males than females (Mann et al., 2015).

It is not only school type that is related to student STEM aspirations but also the initiatives that schools offer to encourage STEM participation and build self-efficacy in STEM subjects, with many of these initiatives focusing on supporting female aspirations in STEM. Some success has been found through the use of STEM career scenarios and active engagement with science experts and scientific practices (Drymiotou et al., 2021). A 9 week intervention designed to improve career decision making self-efficacy and STEM self-efficacy was found to increase these outcome measures with a sample of 88 high school girls (Falco and Summers, 2019). However, there is also evidence that improving motivation and attitudes to STEM may not be sufficient to ensure STEM participation for girls. An intervention designed to highlight the opportunity to fulfill communal goals in mathematics and science did increase interest in STEM careers for high school girls (Fuesting et al., 2017). Girls have been shown to be more motivated by altruistic goals than boys in expressing interest in future careers (Gore et al., 2015).

STEM career aspirations can also be impacted when students engage with informal STEM learning or higher education institutions. A large scale ($n > 15,000$) retrospective study of college students found that campus visits during their schooling, when they involved meetings with college professors, increased their odds of expressing STEM career aspirations (Kitchen et al., 2018, 2020). A summer learning experience focused on connecting STEM to students' everyday lives with authentic hands-on experiences, exposing them to a broad range of STEM careers and highlighting how STEM careers can help improve lives, was found to increase interest in STEM in middle school students (Maiorca et al., 2021) as did a summer STEM camp (Vela et al., 2020) and other out-of-school activities (Donmez, 2021).

Beyond educational institutions it is important to recognize the impact of cultural norms and stereotypes on the formation of STEM aspirations, particularly for females. While self-efficacy in STEM subjects has been consistently demonstrated to be predictive of STEM career interest, there is evidence that negative STEM stereotypes can impact on mathematics and science self-efficacy (Gariott et al., 2017; Luo et al., 2021). However, it is not only stereotypical beliefs about STEM careers that can impact on STEM aspirations. School subjects also have varying degree of masculinity associated with them. In a study involving 1364 Swiss secondary students, mathematics was considered to be the most masculine subject, followed by physics and chemistry respectively (Makarova et al., 2019). These perceptions of STEM subjects as gendered to varying degrees has the potential to influence the subject choices of students at key points during their education.

The research on the development of STEM aspirations paints a complex picture of interacting variables that act to support or discourage students toward STEM study and careers. It is clear that STEM self-efficacy and achievement in STEM subjects is important, but attitudinal variables such as enjoyment, interest in, and valuing of STEM are also significant for shaping students' futures in STEM fields. Also important is the degree to which students have accurate information about the broad range of STEM occupations and the veracity of common STEM stereotypes.

Many of the influences discussed in this chapter were outlined in a comprehensive review article examining the factors that are important in study choice and career development in STEM for 8–16 year olds (van Tuijl and van der Molen, 2016). The authors recommend three, potentially malleable, factors that could form the basis for action to improve STEM participation beyond schooling. Firstly, parents', teachers' and children's knowledge of the STEM field needs to be improved, including by allowing children to experience STEM activities that are relevant to their lives, interests and abilities. Also, the authors point to the need to make parents aware of the extent to which children circumscribe STEM aspirations early in their schooling, and the need to ensure that students have knowledge of the vast range of STEM careers. Secondly, the study points to the need to be aware of and to counter the affective aspects of STEM, largely influenced by STEM stereotypes that reinforce STEM as a "male- and things-oriented" field. Thirdly, the review identifies the importance of individual factors such as self-efficacy and ability beliefs, highlighting the need to turn entity beliefs (that ability is fixed) into incremental beliefs (that competencies can be developed over time) in STEM subjects (van Tuijl and van der Molen, 2016). These three foci have been confirmed in the research included in this chapter and provides a useful framework for those concerned with improving STEM outcomes for a broader range of students.

Theoretical perspectives

There are a range of theoretical perspectives employed by researchers to understand the complexities that surround the formation of aspirations for particular types of careers. Those most commonly employed include Social Cognitive Career Theory

(SCCT), developed by [Lent et al. \(1994\)](#), Expectancy-Value Theory, developed by [Eccles et al. \(1983\)](#) and the Theory of Circumscription and Compromise ([Gottfredson, 1981](#)).

Social Cognitive Career Theory ([Lent et al., 1994](#)) aims to capture how career development occurs by examining the connections between cognitive and interpersonal factors, alongside self-directed and externally imposed influences. The theory has its origins in Albert Bandura's general social cognitive theory which emphasized the relationship between "self-referent thought and social processes in guiding human behavior" ([Lent et al., 1994](#), p. 258). The theory posits that the personal determinants of career development can be represented by three variables: self-efficacy, outcome expectations and personal goals. These individual variables are then situated with contextual and experiential factors such as gender, culture, support systems and barriers ([Lent et al., 1994](#)). Furthermore students' learning experiences play a "pivotal role as a mediator between personal backgrounds and socio-cognitive mechanisms in SCCT" ([Kang and Keinonen, 2017](#), p. 1672).

The SCCT model has informed the development of a career-oriented learning environments as a means of assisting with career decision making ([Drymiotou et al., 2021](#)). The contextual features that were effective in the development of STEM aspirations were found to be learning experiences which: actively involved students in problem solving; established connections with authentic workplace environments; and included interactions with STEM experts ([Drymiotou et al., 2021](#)). SCCT links self-efficacy and career interest and therefore, learning experiences designed to improve student self-efficacy have been found to impact on students' interests and future aspirations, alongside the influence of parents and families ([Murcia et al., 2020](#); [Maiorca et al., 2021](#); [Sahin and Waxman, 2021](#)).

Expectancy-Value Theory (EVT) is another theoretical framework commonly employed in studies of student aspirations and career interest. The theory suggests that individuals show motivation to pursue particular choices depending on their expectation of achieving that outcome and the value that they place on that outcome ([Wille et al., 2020](#)). An individual's expectancy for success is closely linked to their self-concept ([Eccles and Wigfield, 2002](#)) and their value beliefs can be understood as a combination of intrinsic value, attainment value, utility value and associated costs ([Wille et al., 2020](#)). EVT constructs have been demonstrated to predict student achievement and career interests in STEM ([Parker et al., 2014](#); [Wang et al., 2017](#)).

Gottfredson's theory of circumscription and compromise focuses on the development of aspirations through a consideration of the compatibility of one's self-concept with a given occupation ([Bittinger et al., 2021](#)). Gottfredson perceived self-concept to include the beliefs that people hold about themselves, their interests and their place in society ([Cochran et al., 2011](#)). The theory proposes that children, as they move into adolescence, circumscribe their potential future aspirations based on social factors such as the gendered perceptions of certain careers and the social prestige of those careers ([Gottfredson, 1981](#); [Cochran et al., 2011](#)). Gendered perceptions of STEM careers are well established and are reflected in the statistics presented at the beginning of this chapter.

The three theories discussed in this section have similarities in the ways that they explain how individuals develop aspirations and make career choices. They each recognize the interplay between individuals and the social worlds that they inhabit, whether it be the influence of parents and families or the impact of school learning experiences on the choices made by individuals. In this way, each provides an appropriate lens through which to view the formation of aspirations and to ensure that efforts to increase STEM aspirations are suitably targeted toward factors that can make a difference, particularly for under-represented groups.

Recommendations for further research

The research included in this chapter reveals a largely consistent picture of the factors that are important for the development of aspirations for further STEM study and STEM careers. It is clear that self-efficacy in STEM subjects, particularly for mathematics, is vital if students are to continue to study STEM and that interest, enjoyment and valuing of STEM subjects is also significant. Beyond the individual, however, the influence of family and dominant societal norms and stereotypes, also play a role in student choices as they plan for their futures. The ongoing perceived "maleness" of some STEM subjects, such as physics for example, and of STEM careers in technology and engineering, continues to work against the many efforts that are being made to encourage more girls and young women to persist in studying STEM subjects.

The STEM acronym can be both a blessing and a curse for addressing the issues that remain. While it has been useful to consider the science, technology, engineering, and mathematics disciplines as a cohesive group, allowing for significant attention to be drawn to the STEM sector as a whole, it is clear that there are substantial differences between the S, T, E and M disciplines. Similarly, there are differences within disciplines, for example in science, between biology, chemistry, and physics. Future research into the development and support for STEM aspirations should take a more fine-grained approach, carefully delineating between STEM disciplines and sub-disciplines so that programs designed to address under-representation can be better targeted. Not all STEM disciplines are affected by gendered stereotypes, for example, and mathematics, while vitally important in some STEM disciplines is less important in others. If we are to increase interest and participation in STEM generally, while addressing the need to improve the under-representation of particular groups of students, then our research base also needs to diversify to better understand the inequities across the STEM landscape and the most promising ways to address them.

References

- Ahmed, W., Mudrey, R.R., 2019. The role of motivational factors in predicting STEM career aspirations. *Int. J. Sch. Educ. Psychol.* 7 (3), 201–214. <https://doi.org/10.1080/21683603.2017.1401499>.
- Ardies, J., De Maeyer, S., Gijbels, D., 2015. A longitudinal study on boys' and girls' career aspirations and interest in technology. *Res. Sci. Technol. Educ.* 33 (3), 366–386. <https://doi.org/10.1080/02635143.2015.1060412>.
- Atabey, N., Topcu, M.S., 2021. The relationship between Turkish middle school students' 21st century skills and STEM career interest: gender effect. *J. Educ. Sci. Environ. Health* 7 (2), 86–103.
- Bahar, A., Adiguzel, T., 2016. Analysis of factors influencing interest in STEM career: comparison between American and Turkish high school students with high ability. *J. STEM Educ.* 10, 64–69.
- Berger, N., et al., 2020. Charting career aspirations: a latent class mixture model of aspiration trajectories in childhood and adolescence. *Aust. Educ. Res.* 47 (4), 651–678. <https://doi.org/10.1007/s13384-019-00363-x>.
- Bittinger, J.D., Wells, R.S., Kimball, E.W., 2021. STEM career aspirations for high school students with individualized education programs. *Educ. Forum* 85 (1), 49–62. <https://doi.org/10.1080/00131725.2020.1772928>.
- Blotnick, K.A., et al., 2018. A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *Int. J. STEM Educ.* 5 (1), 22. <https://doi.org/10.1186/s40594-018-0118-3>.
- Cairns, D., Dickson, M., 2021. Exploring the relations of gender, science dispositions and science achievement on STEM career aspirations for adolescents in public schools in the UAE. *Asia-Pac. Educ. Res.* 30 (2), 153–165. <https://doi.org/10.1007/s40299-020-00522-0>.
- Cochran, D.B., et al., 2011. Adolescent occupational aspirations: test of Gottfredson's theory of circumscription and compromise. *Career Dev. Q.* 59 (5), 412–427. <https://doi.org/10.1002/j.2161-0045.2011.tb00968.x>.
- Department of Education, Skills and Employment, 2020. STEM Jobs Growing Almost Twice as Fast as Other Jobs. Department of Education, Skills and Employment scheme—AGLSTERMS.AglAgent; corporateName=Department of Education, Skills and Employment; address=50 Marcus Clarke St, Canberra City, ACT 2601; contact=+61 1300 566 046. Available at: <https://www.dese.gov.au/newsroom/articles/stem-jobs-growing-almost-twice-fast-other-jobs>. (Accessed 4 May 2022).
- Department of Industry Science, 2021. STEM Equity Monitor. Department of Industry, Science, Energy and Resources. Available at: <https://www.industry.gov.au/data-and-publications/stem-equity-monitor>. (Accessed 24 March 2022).
- Donmez, I., 2021. Impact of out-of-school STEM activities on STEM career choices of female students. *Eur. J. Educ. Res.* [Preprint]. Available at <https://eric.ed.gov/?id=EJ1284238>. (Accessed 7 January 2022).
- Dorph, R., et al., 2018. When I grow up: the relationship of science learning activation to STEM career preferences. *Int. J. Sci. Educ.* 40 (9), 1034–1057. <https://doi.org/10.1080/09500693.2017.1360532>.
- Drymiotou, I., Constantinou, C.P., Avraamidou, L., 2021. "Enhancing students' interest in science and understandings of STEM careers: the role of career-based scenarios". *Int. J. Sci. Educ.* 43 (5), 717–736. <https://doi.org/10.1080/09500693.2021.1880664>.
- Eccles, J., et al., 1983. Expectancies, values and academic behaviors. In: *Achievement and Achievement Motives*. W.H. Freeman, San Francisco.
- Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53 (1), 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>.
- Edwin, M., Prescod, D.J., Bryan, J., 2019. "Profiles of high school students' STEM career aspirations". *Career Dev. Q.* 67 (3), 255–263. <https://doi.org/10.1002/cdq.12194>.
- Falco, L.D., Summers, J.J., 2019. Improving career decision self-efficacy and STEM self-efficacy in high school girls: evaluation of an intervention. *J. Career Dev.* 46 (1), 62–76. <https://doi.org/10.1177/0894845317721651>.
- Fry, R., Kennedy, B., Funk, C., 2021. STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity. Pew Research Center Science & Society, 1 April. Available at: <https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/>. (Accessed 4 May 2022).
- Fuesting, M.A., Diekmann, A.B., Hudiburgh, L., 2017. From classroom to career: the unique role of communal processes in predicting interest in STEM careers. *Soc. Psychol. Educ.* 20 (4), 875–896. <https://doi.org/10.1007/s11218-017-9398-6>.
- Garriott, P.O., Hultgren, K.M., Frazier, J., 2017. STEM stereotypes and high school students' math/science career goals. *J. Career Assess.* 25 (4), 585–600. <https://doi.org/10.1177/1069072716665825>.
- Gore, J., et al., 2015. Socioeconomic status and the career aspirations of Australian school students: testing enduring assumptions. *Aust. Educ. Res.* 42 (2), 155–177. <https://doi.org/10.1007/s13384-015-0172-5>.
- Gottfredson, L., 1981. Circumscription and compromise: a developmental theory of occupational aspirations. *J. Counsel. Psychol. Monogr.* 28 (6), 545–579.
- Holmes, K., et al., 2018. An integrated analysis of school students' aspirations for STEM careers: which student and school factors are most predictive? *Int. J. Sci. Math. Educ.* 16 (4), 655–675. <https://doi.org/10.1007/s10763-016-9793-z>.
- Kang, J., Keinonen, T., 2017. The effect of inquiry-based learning experiences on adolescents' science-related career aspiration in the Finnish context. *Int. J. Sci. Educ.* 39 (12), 1669–1689. <https://doi.org/10.1080/09500693.2017.1350790>.
- Ketenci, T., Leroux, A., Renken, M., 2020. Beyond student factors: a study of the impact on STEM career attainment. *J. STEM Educ. Res.* 3 (3), 368–386. <https://doi.org/10.1007/s41979-020-00037-9>.
- Kitchen, J.A., Sonnet, G., Sadler, P., 2020. Campus visits: impact of a college outreach strategy on student STEM aspirations. *J. Student Aff. Res. Pract.* 57 (3), 266–281. <https://doi.org/10.1080/19496591.2019.1653312>.
- Kitchen, J.A., Sonnet, G., Sadler, P.M., 2018. The impact of college- and university-run high school summer programs on students' end of high school STEM career aspirations. *Sci. Educ.* 102 (3), 529–547. <https://doi.org/10.1002/sce.21332>.
- Kohen, Z., Nitzan, O., 2022. Excellence in mathematics in secondary school and choosing and excelling in STEM professions over significant periods in life. *Int. J. Sci. Math. Educ.* 20 (1), 169–191. <https://doi.org/10.1007/s10763-020-10138-x>.
- Lent, R.W., Brown, S.D., Hackett, G., 1994. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45 (1), 79–122. <https://doi.org/10.1006/jvbe.1994.1027>.
- Lloyd, A., et al., 2018. Parental influences on those seeking a career in STEM: the primacy of gender. *Int. J. Gender Sci. Technol.* 10 (2), 308–328.
- Luo, T., et al., 2021. STEM stereotypes predict students' STEM career interest via self-efficacy and outcome expectations. *Int. J. STEM Educ.* 8 (1), 36. <https://doi.org/10.1186/s40594-021-00295-y>.
- Mackenzie, E., Berger, N., Holmes, K., 2021. Predicting adolescent girls' intentions to study science in senior high school. *Issues Educ. Res.* 31 (2), 574–585.
- Maiorca, C., et al., 2021. Informal learning environments and impact on interest in STEM careers. *Int. J. Sci. Math. Educ.* 19 (1), 45–64. <https://doi.org/10.1007/s10763-019-10038-9>.
- Makarova, E., Aeschlimann, B., Herzog, W., 2019. The gender gap in STEM fields: the impact of the gender stereotype of math and science on secondary students' career aspirations. *Front. Educ.* 4, 60. <https://doi.org/10.3389/educ.2019.00060>.
- Mann, A., Legewie, J., DiPrete, T.A., 2015. The role of school performance in narrowing gender gaps in the formation of STEM aspirations: a cross-national study. *Front. Psychol.* 6, 171. <https://doi.org/10.3389/fpsyg.2015.00171>.
- Mau, W.-C.J., Li, J., 2018. Factors influencing STEM career aspirations of underrepresented high school students. *Career Dev. Q.* 66 (3), 246–258. <https://doi.org/10.1002/cdq.12146>.

- Murcia, K., Pepper, C., Williams, J., 2020. Youth STEM career choices: what's influencing secondary students' decision making. *Issues Educ. Res.* 30 (2), 593–611. <https://doi.org/10.3316/informit.266265177906130>.
- Oppermann, E., et al., 2021. Elementary school students' motivational profiles across Finnish language, mathematics and science: longitudinal trajectories, gender differences and STEM aspirations. *Contemp. Educ. Psychol.* 64, 101927. <https://doi.org/10.1016/j.cedpsych.2020.101927>.
- Parker, P., et al., 2014. Predicting career aspirations and university majors from academic ability and self-concept: a longitudinal application of the internal–external frame of reference model. In: Schoon, I., Eccles, J.S. (Eds.), *Gender Differences in Aspirations and Attainment: A Life Course Perspective*. Cambridge University Press, Cambridge, pp. 224–246. <https://doi.org/10.1017/CBO9781139128933.015>.
- Reinhold, S., Holzberger, D., Seidel, T., 2018. Encouraging a career in science: a research review of secondary schools' effects on students' STEM orientation. *Stud. Sci. Educ.* 54 (1), 69–103. <https://doi.org/10.1080/03057267.2018.1442900>.
- Sadler, P.M., et al., 2012. Stability and volatility of STEM career interest in high school: a gender study. *Sci. Educ.* 96 (3), 411–427. <https://doi.org/10.1002/sce.21007>.
- Sahin, A., Waxman, H.C., 2021. Factors affecting high school students' stem career interest: findings from a 4-year study. *J. STEM Educ. Innov. Res.* 22 (3). Available at: <https://www.jstem.org/jstem/index.php/JSTEM/article/view/2472>. (Accessed 7 January 2022).
- Sáinz, M., Müller, J., 2018. Gender and family influences on Spanish students' aspirations and values in stem fields. *Int. J. Sci. Educ.* 40 (2), 188–203. <https://doi.org/10.1080/09500693.2017.1405464>.
- Sax, L.J., et al., 2015. "But I'm not good at math": the changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Res. High. Educ.* 56 (8), 813–842. <https://doi.org/10.1007/s11162-015-9375-x>.
- The World Bank, 2021. The Equality Equation: Advancing the Participation of Women and Girls in STEM—The World Bank Group, World Bank. Available at: <https://www.worldbank.org/en/news/infographic/2021/04/19/the-equality-equation-advancing-the-participation-of-women-and-girls-in-science-technology-engineering-and-mathematics-s>. (Accessed 27 April 2022).
- van Tuijl, C., van der Molen, J.H.W., 2016. Study choice and career development in STEM fields: an overview and integration of the research. *Int. J. Technol. Des. Educ.* 26 (2), 159–183. <https://doi.org/10.1007/s10798-015-9308-1>.
- Vela, K.N., Pedersen, R.M., Baucum, M.N., 2020. Improving perceptions of STEM careers through informal learning environments. *J. Res. Innov. Teach. Learn.* 13 (1), 103–113. <https://doi.org/10.1108/JRIT-12-2019-0078>.
- Wang, M.-T., Ye, F., Degol, J.L., 2017. Who chooses STEM careers? Using a relative cognitive strength and interest model to predict careers in science, technology, engineering, and mathematics. *J. Youth Adolesc.* 46 (8), 1805–1820. <https://doi.org/10.1007/s10964-016-0618-8>.
- Wiebe, E., Unfried, A., Faber, M., 2018. The relationship of STEM attitudes and career interest. *Eurasia J. Math. Sci. Technol. Educ.* 14 (10). <https://doi.org/10.29333/ejmste/92286>.
- Wille, E., et al., 2020. It takes two: expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemp. Educ. Psychol.* 61, 101858. <https://doi.org/10.1016/j.cedpsych.2020.101858>.
- Zilberman, A., Ice, L., 2021. Why Computer Occupations Are Behind Strong STEM Employment Growth in the 2019–29 Decade : Beyond the Numbers. U.S. Bureau of Labor Statistics. Available at: <https://www.bls.gov/opub/btn/volume-10/why-computer-occupations-are-behind-strong-stem-employment-growth.htm>. (Accessed 4 May 2022).

Self-efficacy in STEM

Martha Escobar^a, Haylee M. Majewski^a, Mohammed Qazi^b, and Yasmeen Rawajfih^c, ^aDepartment of Psychology, Oakland University, Rochester, MI, United States; ^bDepartment of Mathematics, Tuskegee University, Tuskegee, AL, United States; and ^cDepartment of Computer Science, Tuskegee University, Tuskegee, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	388
Self-efficacy is a construct about the self	388
Development of self-efficacy	389
Sources of self-efficacy	389
Reciprocal determinism and self-efficacy	389
Measurement issues	390
Academic self-efficacy	390
Self-efficacy as a predictor of academic performance	390
Academic self-efficacy and academic self-regulation	390
Self-efficacy in STEM	391
Gender differences in self-efficacy and STEM achievement	391
Educational strategies to develop and strengthen STEM self-efficacy	392
Use of mastery experiences	392
Use of vicarious learning	392
Use of social persuasion	392
Other strategies	392
Conclusion	393
Acknowledgment	393
References	393
Relevant website	394

Introduction

Self-efficacy can be defined as people's "beliefs in their capabilities to produce desired effects by their own actions" (Bandura, 1997, p. vii). That is, self-efficacy is not one's awareness about existing skills or capabilities, it is the *belief* of what one can do with those skills and capabilities. Self-efficacy beliefs are specific to a task within a specific domain (e.g., Pajares, 2005). This means that self-efficacy is *not* a personality trait. Personality involves general and stable patterns of thinking, feeling and behaving; in contrast, self-efficacy can be specific to a particular skill or domain. Indeed, measures of self-efficacy for specific domains can be more predictive than general measures of self-efficacy (e.g., Bandura, 1997; Maddux, 1995). Because self-efficacy is a belief, it can also be relatively independent of objective performance, being more consistent with the subject's own *interpretation* of that performance.

Self-efficacy is a construct about the self

Self-efficacy is related to, but different from, other psychological constructs about the self, such as self-esteem and self-concept. *Self-esteem* is a global evaluation of the self that involves strong emotional components (e.g., self-like; Gist and Mitchell, 1992), whereas *self-concept* is related to domain-related perceptions about the self that can include evaluative and/or emotional components (e.g., "I am good in science"). In contrast, *self-efficacy* is specific to a task or goal (e.g., "I will do good in the taxonomy test"), does not involve a self-evaluation or self-like component, and is independent of others' capabilities (Bong and Skaalvik, 2003). Self-efficacy is strongly associated with self-concept, but they do not necessarily vary together. For example, high self-efficacy for different science tasks may contribute to a high science self-concept, but an individual can have a high science self-concept and low self-efficacy related to a particular task or assignment. Self-efficacy is more predictive of performance than self-concept (Britner and Pajares, 2006), possibly because it modulates both self-regulation (self-monitoring and assessment; Zimmerman, 2000) and cognitive engagement (Pintrich, 2003; Zimmerman, 2000) during task completion. In a similar way, self-efficacy (associated with performance) may vary with self-esteem (associated with self-worth), but they can also be dissociable. Self-esteem can be categorized as a trait-level construct, so one can have high overall self-esteem while simultaneously having low self-efficacy for a particular task (Richardson et al., 2012).

Development of self-efficacy

Sources of self-efficacy

According to Bandura (1977), self-efficacy beliefs are derived from four primary sources:

- Mastery experience* refers to the outcomes of previous task performance. These outcomes can act as reinforcers or punishers of self-efficacy, with successful outcome attainment increasing self-efficacy, and failures to attain outcome lowering it.
- Vicarious experience* refers to observations of others performing a task. Characteristics of the actor, and whether the individual can identify with the actor, influence the impact of these vicarious experiences.
- Social persuasion* refers to the feedback of other individuals (particularly those who are influential to the individual) regarding one's performance.
- Physiological and emotional reactions* refer to the individual's interpretation of the physiological reactions that are triggered by the task (e.g., sweaty palms, fast heart rate, "butterflies" in the stomach).

Mastery experience and physiological/emotional reactions constitute intrapersonal sources of self-efficacy, whereas vicarious experience and social persuasion constitute interpersonal sources of self-efficacy. Intrapersonal and interpersonal sources can interact to determine self-efficacy. However, despite strong correlations among the four sources of self-efficacy listed above (mastery experience, vicarious experience, social persuasion, and physiological reactions; e.g., Klassen, 2004), their relative contribution to self-efficacy is not equal. For example, mastery experience is a more robust determinant of self-efficacy than vicarious experiences, which are in general stronger determinants of self-efficacy than social persuasion. However, social persuasion can be a very robust determinant if the person providing the feedback is a person of authority, highly trusted, or considered an expert. Furthermore, mastery experience influences self-efficacy only if successes and failures are attributed to one's own behavior (Maddux, 2012). Mastery experience tends to have a linear relationship with self-efficacy (the more experiences with successful completion of a task, the higher the self-efficacy), whereas physiological reactions may not have such a linear relationship with self-efficacy (e.g., moderate stress can improve performance but high levels of stress usually impair it).

Even though the sources of self-efficacy are presented as individual factors, each of these sources is not experienced in isolation. For example, if a student is taking a test, the mastery experience is likely to be accompanied by physiological and emotional reactions (e.g., anxiety). The student is also likely to see others attempting or completing the same test, and will receive feedback from the teacher regarding performance in the test. Thus, self-efficacy will not be determined by a single source; rather, it is the integration and interpretation of these four sources of information that will eventually determine self-beliefs.

There is also a reciprocal relationship between performance and self-efficacy. Students with high levels of self-efficacy set more challenging goals for themselves, and achievement of those goals can boost self-efficacy (conversely, failure to attain goals can decrease self-efficacy). Furthermore, successes that result from overcoming challenges may have a stronger impact on self-efficacy than successes that are easily won. Because self-efficacy is a belief, mastery itself does not lead to strong self-efficacy; rather, it is the interpretation of the task in terms of context, previously held beliefs, perceived difficulty, and perceived effort that determines its impact on self-efficacy. This means that if the individual does not have prior experience with the task or their own abilities, vicarious learning or social persuasion may become more relevant than mastery experience (Britner and Pajares, 2006).

Reciprocal determinism and self-efficacy

Bandura (e.g., 1978) suggested that behavior influences and is influenced by personal factors and the social environment. In this *reciprocal determinism* framework, behavior (e.g., performance, skills, practices) is determined by the individual (e.g., self-efficacy, cognitive abilities, attitudes) and the environment (e.g., imposed environments, selected environments, constructed environments). For example, a student who does not feel confident about his or her capacity to pass a math course (individual) may not turn in an assignment (behavior) leading to a negative comment from the teacher (environment), which may lead to the student feeling even less confident about his or her capacity to pass the math course (see Fig. 1).

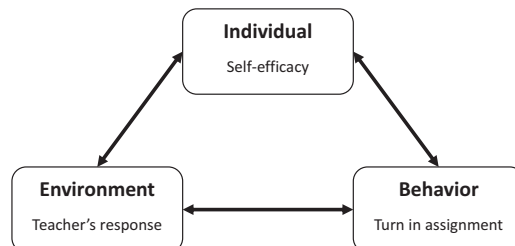


Fig. 1 Reciprocal determinism. Behavior (e.g., a student turning in an assignment for a class) is determined by individual (e.g., student's self-efficacy) and environmental (e.g., teacher's response to the student) factors. In turn, the consequences of behavior can impact individual and environmental factors. See text for details.

Measurement issues

Several scales have been developed to measure the four sources of self-efficacy (see **Sources of self-efficacy**; e.g., Britner and Pajares, 2006; Hampton, 1998; Lent et al., 1991), but they have proven unreliable to assess the impact of interpersonal sources of self-efficacy (vicarious experiences and social persuasion; Ahn et al., 2017; Usher, 2009). Vicarious experience measures typically rely on questions such as, “Seeing adults do well in math pushes me to do better,” which effectively assess both the vicarious experience and the resulting self-belief (Ahn et al., 2017). Similarly, social persuasion measures typically include questions such as, “My classmates said that I understood everything taught in class,” which not only assesses the impact of feedback but also corroborate self-efficacy information (Ahn et al., 2017). Thus, self-reports of interpersonal sources of self-efficacy can be biased by personal self-efficacy beliefs (Ahn et al., 2017; Zander et al., 2020). Recently, Zander et al. (2020) suggested that information from the subjects’ social environment can be used to gain a more reliable estimate of the impact of interpersonal sources of self-efficacy. For example, rather than assessing vicarious experiences by asking a student whether seeing adults do well makes the student want to do better, they proposed to ask students who would they select as a role model, and then assess the model’s competence-related status (i.e., the model’s grade in the course). Similarly, rather than assessing social persuasion by asking a student whether her classmates think she understood a particular topic, Zander et al. proposed asking the classmates whether they would approach the student to ask her questions about the class topic. Unfortunately, these measures have their own limitations. For example, intentions to ask for help do not necessarily reflect actual requests for help, thus decreasing the predictive value of the measure.

Academic self-efficacy

Self-efficacy as a predictor of academic performance

When considering academic performance, self-efficacy can be divided into *performance self-efficacy* (beliefs about one’s capacity derived from past experiences) and *academic self-efficacy* (beliefs about one’s capacity derived from perceived competencies). Performance self-efficacy is useful when encountering familiar tasks, whereas academic self-efficacy is useful when encountering unfamiliar tasks, for which past experiences provide little information (Richardson et al., 2012).

Self-efficacy develops partially based on feedback regarding performance (Fig. 1). For students, performance feedback usually takes the form of grades, and as feedback accumulates (more grades are accrued in a particular subject), self-efficacy stabilizes. Beliefs about competency are important to set both short-term (e.g., passing an exam) and long-term (e.g., passing a course) academic goals, and influence the interpretation of mismatches between goals and performance. These mismatches can be interpreted in different manners based upon levels of self-efficacy. For example, a student with high self-efficacy obtaining a lower grade (performance) than desired (goal) could interpret the mismatch as a result of behavioral factors (e.g., not studying hard enough), whereas a student with low self-efficacy could interpret it as a result of individual factors (e.g., lack of capacity). This different interpretation would then strengthen or weaken self-efficacy (respectively), leading to differential goal setting and effort to attain those goals. Achieving a long-term goal requires attainment of many short-term goals; thus, self-efficacy can be continuously strengthened or weakened on the path to obtaining those goals (Rittmayer and Beier, 2009).

Students’ beliefs about their competency are highly predictive of their grade point average (GPA; Pajares and Miller, 1995). In turn, grades are powerful determinants of self-efficacy, depending on whether the student interprets the grade obtained as a success (“I studied hard and got a B!”) or a failure (“I got a B but I should have gotten an A”). Academic successes and failures may have different impacts on self-efficacy. Good grades seem to be less influential for the development of self-efficacy than bad grades; that is, good grades will not strengthen self-efficacy as much as bad grades will weaken self-efficacy (Crocker et al., 2003). Similarly, negative or discouraging appraisals can weaken self-efficacy more than positive or encouraging appraisals can strengthen self-efficacy (Britner and Pajares, 2006). The type of feedback that is delivered in the academic environment also appears to have a critical role in academic achievement: Constructive feedback (e.g., references to the accuracy of the work completed) has a more positive impact on self-efficacy than feedback about deficiencies (e.g., references to the things missed in the work completed; Britner and Pajares, 2006).

Academic self-efficacy and academic self-regulation

Self-efficacy is an essential component of self-regulation (e.g., Barone et al., 1997), the processes that allow an individual to decide how to best implement his or her efforts. Self-regulation depends on individuals setting goals for performance, which can motivate and guide action, and on self-evaluations of such performance, which result in emotional responses that can enhance or disrupt self-regulation. Self-efficacy beliefs determine the extent to which the goals set are challenging, the choice of activities to be performed to attain those goals, and how discrepancies between goals and performance are interpreted (Maddux, 2012). Pintrich (2004) suggested that self-regulation is related to the constructs of goal orientation, approach goals, academic intrinsic motivation, and academic self-efficacy. Pintrich’s model can be assessed with the Motivated Strategies for Learning Questionnaire (MSLQ), which assesses the use of cognitive strategies such as rehearsal, organization, critical thinking, and some aspects of metacognition. Academic self-efficacy would be directly related to effort regulation, which determines persistence in the face of challenges (effort regulation can be viewed as similar to achievement motivation; Richardson et al., 2012). Thus, self-efficacy is intrinsically related to

goal-setting and attainment, the choice of behaviors that will be used to pursue those goals, the self-evaluation of the outcomes of one's efforts, and persistence toward those goals.

Self-efficacy in STEM

Within any given domain, self-efficacy is a powerful predictor of motivation, persistence, and task performance (Bandura and Locke, 2003). When it comes to the domain of science, engineering, technology, and mathematics (STEM), self-efficacy is a better predictor of task performance than past performance in objective assessments, such as tests and assignments (Bandura, 1986; Bandura and Locke, 2003; Larson et al., 2015). Persistence in STEM fields also shows a robust positive association with self-efficacy (Shaw and Barbuti, 2010). Starting in middle school, science self-efficacy is strongly associated with science achievement (Britner and Pajares, 2006), and self-efficacy scores become one of the most reliable predictors of achievement and engagement with science activities in high school (Lau and Roeser, 2002) and college students (Gwilliam and Betz, 2001).

Self-efficacy scores are not only positively correlated to science and math grades (Britner and Pajares, 2006), they are also positively correlated to effort, persistence, engagement and interest in science (Schunk and Pajares, 2002; Zimmerman, 2000). The reciprocal relationship between performance and self-efficacy is mirrored by a reciprocal relationship between interest and self-efficacy. Successful performance in a task leads to increased interest and future pursuit of that task, often at a more challenging level. Indeed, high self-efficacy leads to engagement and persistence in challenging tasks (Komarraju and Nadler, 2013; Pajares, 1996) and self-efficacy is more strongly associated to interest than actual skill or ability (Bandura, e.g., 1991). For example, successful completion of a basic math class may boost interest in math and pursuit of an intermediate math class; indeed, individuals with high self-efficacy tend to choose challenging classes and not view the challenge as a threat (Watt, 2006).

Importantly, as is the case for other forms of self-efficacy, academic self-efficacy is not general to a domain, but specific to tasks within that domain. For example, performance in solving math problems is more strongly associated with students' confidence in their capacity to solve math problems than with their confidence in their capacity to perform math-related tasks or succeed in math-related courses, whereas confidence to succeed in math-related courses was more strongly associated with students' choice of math majors than with their confidence in solving math problems or performing math-related tasks (Pajares and Miller, 1995). However, overall math self-efficacy is not only a strong predictor of math performance, but also a predictor of choice of science-related courses and careers (Lent et al., 1991).

Gender differences in self-efficacy and STEM achievement

It is well known that women and certain racial/ethnic groups are severely underrepresented in STEM careers (National Science Board, 2020). Because self-efficacy beliefs are robust predictors of STEM success, it is not surprising that self-efficacy is a focus for educators who wish to increase engagement and performance in STEM. Among middle schoolers, female students have higher science self-efficacy scores (and hence, higher science grades) than male students (Britner and Pajares, 2001). Although these gender differences are not evident among high schoolers (Lau and Roeser, 2002), the opposite pattern begins to emerge and solidify as students transition to tertiary education. Gender differences in self-efficacy seem to be discipline specific, with self-efficacy for males being higher than for females in "hard" STEM disciplines such as physics, chemistry, mathematics, and engineering, equivalent across genders in life sciences such as biology, and higher for females than males in non-STEM disciplines such as English (Whitcomb et al., 2020). Thus, underrepresentation of women in some STEM disciplines is associated with low self-efficacy scores for those disciplines.

In 2014–2015, women obtained 58.01% of bachelor's (BS) degrees in the biosciences and 54.58% of BS degrees in the social sciences. In contrast, women obtained only 19.8%, 18.1%, and 41.7% of BS degrees in engineering, computer science, and mathematics, respectively (National Center for Science and Engineering Statistics, 2017). Notably, while the gender gap in math achievement has grown smaller, females are still overall less likely to choose STEM careers than males (e.g., Hyde, 2014; Zander et al., 2020). This is a paradox, because math is a foundational discipline for all other STEM fields, and achievement and interest should be expected to vary together.

Traditionally, women have been seen as innately less capable of excelling in math, an assumption that has been challenged by the progressively closing gender gap in math achievement (e.g., Hyde, 2014). Social gender stereotypes are a clear determinant of math and STEM choice, with gender gaps being almost nonexistent in countries in which these stereotypes are less prominent (Glazer, 2019). Gender biases create *stereotype threats*, in which individuals fear confirming the stereotypes that the social group has about them. Biases and stereotypes are environmental factors that can powerfully impact math and STEM self-efficacy in women. As mentioned above, self-efficacy is usually a better predictor of performance than past objective evaluations (Larson et al., 2015; Pajares, 1996). Thus, even if female students receive math grades that are equivalent to those of male students, they may be more likely to hold beliefs about insufficient math abilities than male students (e.g., Else-Quest et al., 2013). Thus, female students do not lack the capacity to do STEM; rather, they lack the belief that they are able to do STEM (Rittmayer and Beier, 2009). (Note that the lack of belief in one's capacity could be viewed as associated to the so-called *imposter syndrome*, in which individuals who are high achieving doubt their capacities and feel like a "fraud" for their accomplishments).

Mastery experiences increase academic self-efficacy for both male and female students, but it seems that mastery experiences (intrapersonal sources) are most important for male students, whereas social persuasion and vicarious learning (interpersonal

sources) appear to be most important for female students (Butz and Usher, 2015; Sawtelle et al., 2012; Usher and Pajares, 2008). Mentors and role models can be powerful sources of social persuasion, and experiences with role models who belong to one's same gender and/or racial group increase their persuasive value. Observation of *female* role models in STEM is an especially robust determinant of STEM self-efficacy for female students (Seymour, 1995; Zeldin and Pajares, 2000). However, any mentorship by a STEM expert can have a powerful impact on students' STEM self-efficacy because such mentors provide encouragement and recognition (Carlone and Johnson, 2007; Stout et al., 2011) and provide intellectual discussions that can boost self-efficacy through mastery performance (Rittmayer and Beier, 2009).

Male and female students may also engage in different cognitive strategies to interpret successes and failures that may determine how their academic beliefs are impacted by mastery experience. Male students are more likely than female students to engage in *self-enhancement* (positive views of the self that are not consistent with performance; Kurman, 2004), and these enhanced views of the self can in turn enhance self-efficacy beliefs for male students, but not their female counterparts.

Educational strategies to develop and strengthen STEM self-efficacy

Use of mastery experiences

Educational programs can be used to help students develop and strengthen their beliefs in their own capacities. Self-efficacy can be strengthened by providing students with *mastery experiences* through interventions such as active and hands-on learning (Kuchynka et al., 2021), structured goal-setting and encouragement to develop self-regulation strategies (e.g., Dunlap, 2005), and proximal goal setting with increasing-difficulty tasks that ensure continuous success (e.g., Luzzo et al., 1999). At a practical level, this means that activities that encourage the development of self-regulation, are age- and ability-appropriate to ensure and encourage success, and which can be broken down into small, short-term goals should be beneficial for the development of self-efficacy (Rittmayer and Beier, 2009). Other strategies may include providing students with immediate feedback related to their performance in order to adjust their self-efficacy beliefs to their actual mastery performance. Considering that girls tend to have lower STEM self-efficacy and self-enhance less than boys, immediate feedback regarding good performance may allow them to incorporate a successful experience more effectively than receiving delayed feedback (Zander et al., 2020).

Use of vicarious learning

Vicarious learning experiences are especially relevant for young women, and their impact is usually mediated by whether the students can identify with the model (i.e., if they view themselves as similar to the model; Zeldin and Pajares, 2000), and whether the vicarious experiences are provided along with other interventions (e.g., mastery experiences; Luzzo et al., 1999). Formal and informal environments can provide a variety of vicarious STEM experiences, including interactions with STEM professionals in structured and unstructured situations, mentoring by STEM professionals, shadowing STEM professionals at work, audiovisual materials that expose students to STEM professionals' life experiences, and workshops and summer or after-school camps (American Association of University Women, 2004). Female students appear to benefit most from mentoring if inequalities in STEM achievement are described explicitly as part of interventions that expose them to female STEM role models (Weisgram and Bigler, 2007). Possibly, explicit discussion of gender inequalities in STEM allows students to re-evaluate and adjust their beliefs about their own STEM capacities. At a practical level, group work in which a peer with high ability to complete the task, and interactions with role models (e.g., older students and STEM professionals) with whom the students can identify (both in terms of race and/or gender) should encourage the development of self-efficacy (Rittmayer and Beier, 2009). This is especially important for female students.

Use of social persuasion

Social persuasion can be used to support other interventions aimed at increasing STEM self-efficacy. Positive feedback from individuals who are relevant to the student can dramatically impact their beliefs about their own abilities. Parents' expectations about their children's performance in academic tasks can be reliable predictors of the child's objective performance (Vekiri and Cronaki, 2008) and choice of STEM careers (Bleeker and Jacobs, 2004). Positive feedback can increase the impact of mastery experiences, especially if students receive support to interpret performance failures as opportunities for learning (rather than threats to self-esteem; Yeager and Dweck, 2012; Zander et al., 2020). Constructive feedback (e.g., comments about strategies for improvement) has a more positive impact on self-efficacy than feedback about what has not yet been accomplished (e.g., comments about deficiencies; Britner and Pajares, 2006). At a practical level, feedback should be framed in positive terms, encourage students to view failures as part of the learning process, provide information that allows students and relevant others (e.g., parents) to become involved in the learning process, as well as encourage participation in extra-curricular activities in which positive feedback is likely (Rittmayer and Beier, 2009). Importantly, feedback should also be *realistic*; that is, it should be based on the student's actual skills and capacities. Unrealistic feedback can create unrealistic expectations, lead students to set goals that are not attainable, and which can possibly result in performance failures. Even if such unrealistic feedback is positive, it can have negative consequences for self-efficacy.

Other strategies

Some researchers have proposed the use of imaginal experiences in situations in which performing the task or observing others performing the task is not possible. This could be achieved either by imagining situations based on prior experiences, or by having

another individual guide the imaginal experience (Williams, 1995). However, the impact of such imagined experiences on the development of self-efficacy beliefs is likely to be minor.

Conclusion

Academic performance appears to be closely related to self-efficacy beliefs. High levels of academic self-efficacy are strongly associated with the use of self-regulation strategies, high levels of cognitive effort, and persistence to achieve increasingly more challenging academic goals. Academic self-efficacy is a strong predictor of grades, especially in science subjects. Feedback about performance (grades) can strengthen or weaken self-efficacy, as mastery experiences are one of its most powerful sources. Negative performance evaluations seem to outweigh positive performance evaluations, suggesting that educational interventions that present material in an incremental fashion or in a way that guarantees mastery should strengthen self-efficacy. This should create a positive feedback cycle: Self-efficacy is related to the choice of behaviors used to pursue academic goals, self-evaluation of performance, and persistence toward those goals. That is, self-efficacy is relevant for both self-regulation of academic behaviors and metacognition about one's performance.

STEM Self-efficacy is a better predictor of STEM performance than the results of previous objective evaluations (e.g., scholastic assessments such as tests and assignments). STEM self-efficacy is also associated with motivation, interest, and persistence in STEM disciplines, suggesting that interventions that increase self-efficacy should also increase interest and persistence in STEM. Indeed, differences in self-efficacy appear to be one of the reasons for differences in the desire to pursue STEM careers and STEM achievement among male and female students, despite equivalent objective capacity (i.e., grades). Interventions that increase mastery experiences do not appear to be enough to encourage female students to persist in STEM; rather, female students appear to be more sensitive to the impact of vicarious experiences and social persuasion, especially if those vicarious experiences are by a model who resembles them (i.e., a female), and if the persuasive messages are provided by someone highly relevant to them (e.g., a respected STEM professional, parents).

Importantly, although self-efficacy has been reviewed here as an isolated factor associated to academic motivation and achievement, self-efficacy interacts with many other cognitive and emotional factors that determine academic motivation. Strategies to incorporate self-efficacy into educational strategies should take advantage of such interactions to foster their effectiveness.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grants No. HRD-1719423 to Oakland University and HRD-1719433 to Tuskegee University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Francisco Arcediano for comments on an early version of this manuscript.

References

- Ahn, H.S., Bong, M., Kim, S.-I., 2017. Social models in the cognitive appraisal of self-efficacy information. *Contemp. Educ. Psychol.* 48, 149–166.
- American Association of University Women (AAUW), 2004. *Under the Microscope*. AAUW, Washington, DC.
- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215.
- Bandura, A., 1978. The self system in reciprocal determinism. *Am. Psychol.* 33, 344–358.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, NJ.
- Bandura, A., 1991. Social cognitive theory of self-regulation. *Organ. Behav. Hum. Decis. Process.* 50, 248–287.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman, New York.
- Bandura, A., Locke, E.A., 2003. Negative self-efficacy and goal effects revisited. *J. Appl. Psychol.* 88, 87–99.
- Barone, D., Maddux, J.E., Snyder, C.R., 1997. *Social Cognitive Psychology: History and Current Domains*. Plenum, New York.
- Bleeker, M.M., Jacobs, J.E., 2004. Achievement in math and science: do mothers' beliefs matter 12 years later? *J. Educ. Psychol.* 96, 97–109.
- Bong, M., Skaalvik, E.M., 2003. Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15, 1–40.
- Britner, S.L., Pajares, F., 2006. Sources of science self-efficacy beliefs of middle school students. *J. Res. Sci. Teach.* 43, 485–499.
- Britner, S.L., Pajares, F., 2001. Self-efficacy beliefs, motivation, race, and gender in middle school science. *J. Women Minorities Sci. Eng.* 7, 271–285.
- Butz, A.R., Usher, E.L., 2015. Salient sources of early adolescents' self-efficacy in two domains. *Contemp. Educ. Psychol.* 42, 49–61.
- Carlone, H.B., Johnson, A., 2007. Understanding the science experiences of successful women of color: science identity as an analytic lens. *J. Res. Sci. Teach.* 44 (8), 1187–1218.
- Crocker, J., Karpinski, A., Quinn, D.M., Chase, S.K., 2003. When grades determine self-worth: consequences of contingent self-worth for male and female engineering and psychology majors. *J. Pers. Soc. Psychol.* 85, 507–516.
- Dunlap, J.C., 2005. Problem-based learning and self-efficacy: how a capstone course prepares students for a profession. *Educ. Technol. Res. Dev.* 53, 65–85.
- Else-Quest, N.M., Mineo, C.C., Higgins, A., 2013. Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychol. Women Q.* 37, 293–309.
- Gist, M.E., Mitchell, T.R., 1992. Self-efficacy: a theoretical analysis of its determinants and malleability. *Acad. Manag. Rev.* 17, 183–211.
- Glazer, A., March 10, 2019. National Mathematics Survey. MIT Women in Mathematics. Available at: <https://math.mit.edu/wim/2019/03/10/national-mathematics-survey/>. (Accessed 10 June 2022).
- Gwilliam, L.R., Betz, N.E., 2001. Validity of measures of math- and science-related self-efficacy for African Americans and European Americans. *J. Career Assess.* 9, 261–281.
- Hampton, N.Z., 1998. Sources of academic self-efficacy scale: an assessment tool for rehabilitation counselors. *Rehabil. Counsel. Bull.* 41, 260–277.
- Hyde, J.S., 2014. Gender similarities and differences. *Annu. Rev. Psychol.* 65, 373–398.

- Klassen, R., 2004. A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo non-immigrant early adolescents. *J. Educ. Psychol.* 96, 731–742.
- Kumaraju, M., Nadler, D., 2013. Self-efficacy and academic achievement: why do implicit beliefs, goals, and effort regulation matter? *Learn. Indiv. Differ.* 25, 67–72.
- Kuchynka, S., Reifsteck, T.V., Gates, A.E., Rivera, L.M., 2021. Developing self-efficacy and behavioral intentions among underrepresented students in STEM: the role of active learning. *Front. Educ.* 6, 668239.
- Kurman, J., 2004. Gender, self-enhancement, and self-regulation of learning behaviors in junior high school. *Sex. Roles* 50, 725–735.
- Larson, L.M., Pesch, K.M., Surapaneni, S., Bonitz, V.S., Wu, T.-F., Werbel, J.D., 2015. Predicting graduation: the role of mathematics/science self-efficacy. *J. Career Assess.* 23, 399–409.
- Lau, S., Roeser, R.W., 2002. Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educ. Assess.* 8, 139–162.
- Lent, R.W., Lopez, F.G., Bieschke, K.J., 1991. Mathematics self-efficacy: sources and relation to science-based career choice. *J. Counsel. Psychol.* 38, 424–430.
- Luzzo, D.A., Hasper, P., Albert, K.A., Bibby, M.A., Martinelli, E.A., 1999. Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students. *J. Counsel. Psychol.* 46, 233–243.
- Maddux, J.E., 1995. Self-efficacy theory: an introduction. In: Maddux, J.E. (Ed.), *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application*. Plenum, New York, pp. 3–36.
- Maddux, J.E., 2012. Self-efficacy: the power of believing you can. In: J. Lopez, S., Snyder, C.R. (Eds.), *The Oxford Handbook of Positive Psychology*, second ed. Oxford University Press.
- National Science Board (NSB), 2020. The State of U.S. Science and Engineering 2020. Available at: <https://ncses.nsf.gov/pubs/nsb20201>. (Accessed 10 June 2022).
- NCES (National Center for Science and Engineering Statistics), January 2017. Women, Minorities, and Persons With Disabilities in Science and Engineering. National Science Foundation (NSF 17-310). Available at: <https://www.nsf.gov/statistics/2017/nsf17310/digest/fod-women/mathematics-and-statistics.cfm>. (Accessed 10 June 2022).
- Pajares, F., 1996. Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578.
- Pajares, F., 2005. Gender differences in mathematics self-efficacy beliefs. In: Gallagher, A.M., Kaufman, J.C. (Eds.), *Gender Differences in Mathematics: An Integrative Psychological Approach*. Cambridge University Press, New York, pp. 294–315.
- Pajares, F., Miller, M.D., 1995. Mathematics self-efficacy and mathematics performances: the need for specificity of assessment. *J. Counsel. Psychol.* 42, 190–198.
- Pintrich, P.R., 2003. A motivational science perspective on the role of student motivation in learning and teaching contexts. *J. Educ. Psychol.* 95, 667–686.
- Pintrich, P.R., 2004. A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ. Psychol. Rev.* 16, 385–407.
- Richardson, M., Abraham, C., Bond, R., 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–387.
- Rittmayer, A.D., Beier, M.E., 2009. Self-Efficacy in STEM. SWE-AWE CASEE Overviews. Available at: http://aweonline.org/arp_selfefficacy_overview_122208.pdf. (Accessed 10 June 2022).
- Sawtelle, V., Brewster, E., Kramer, L.H., 2012. Exploring the relationship between self-efficacy and retention in introductory physics. *J. Res. Sci. Teach.* 49, 1096–1121.
- Schunk, D.H., Pajares, F., 2002. The development of academic self-efficacy. In: Wigfield, A., Eccles, J.S. (Eds.), *Development of Achievement Motivation: A Volume in the Educational Psychology Series*. Academic Press, San Diego, CA, pp. 15–31.
- Seymour, E., 1995. The loss of women from science, mathematics, and engineering undergraduate majors: an explanatory account. *Sci. Educ.* 79, 437–473.
- Shaw, E.J., Barbuti, S., 2010. Patterns of persistence in intended college major with a focus on STEM majors. *NACADA J.* 30, 19–34.
- Stout, J.G., Dasgupta, N., Hunsinger, M., McManus, M.A., 2011. STEMing the tide: using ingroup experts to inoculate women's self-concept in Science, Technology, Engineering, and Mathematics (STEM). *J. Pers. Soc. Psychol.* 100 (2), 255–270.
- Usher, E.L., Pajares, F., 2008. Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78, 751–796.
- Usher, E.L., 2009. Sources of middle school students' self-efficacy in mathematics: a qualitative investigation. *Am. Educ. Res. J.* 46, 275–314.
- Vekiri, I., Chronaki, A., 2008. Gender issues in technology use: perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Comput. Educ.* 51, 1392–1404.
- Watt, H.M.G., 2006. The role of motivation in gendered educational and occupational trajectories related to maths. *Educ. Res. Eval.* 12, 305–322.
- Weisgram, E.S., Bigler, R.S., 2007. Effects of learning about gender discrimination on adolescent girls' attitudes toward and interest in science. *Psychol. Women Q.* 31, 262–269.
- Whitcomb, K.M., Kalender, Y., Nokes-Malach, T.J., Schunn, C., Singh, C., 2020. *Int. J. Eng. Educ.* 36, 1996–2014.
- Williams, S.L., 1995. Self-efficacy, anxiety, and phobic disorders. In: Maddux, J.E. (Ed.), *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application*. Plenum, New York, pp. 69–107.
- Yeager, D.S., Dweck, C.S., 2012. Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47, 302–314.
- Zander, L., Höhne, E., Harms, S., Pfost, M., Hornsey, M.J., 2020. When grades are high but self-efficacy is low: unpacking the confidence gap between girls and boys in mathematics. *Front. Psychol.* 11, 552355.
- Zeldin, A.L., Pajares, F., 2000. Against the odds: self-efficacy beliefs of women in mathematical, scientific, and technological careers. *Am. Educ. Res. J.* 37, 215–246.
- Zimmerman, B.J., 2000. Self-efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25, 82–91.

Relevant website

National Science Board (NSB), 2018. Science and Engineering Indicators 2018 (NSB-2018-1). Available at: <https://www.nsf.gov/statistics/indicators>. (Accessed 10 June 2022).

Positioning diverse students for science, technology, engineering, and mathematics (STEM) careers: reinforced learning ecosystems for STEM success

Breonte S. Guy, Winston-Salem State University, Winston-Salem, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	395
Positioning a diverse STEM workforce for STEM careers	396
Inclusive STEM	396
The future: strengthening the STEM learning ecosystem	397
Conclusions and recommendations	399
Acknowledgments	399
References	400

Introduction

What did you want to be when you grew up? Did your career dreams come true? Did your parents help you on your journey toward becoming what you wanted to be? Were your extended family, schools, or other external community, state, and government resources helpful in supporting your interests?

Like many children growing up, my dream was to become a veterinarian. This dream was somewhat logical for a child growing up in an agricultural community in the deep south. Family and neighbors nurtured my love of animals as best they could. They drilled me on different breeds of companion animals and livestock, posed questions about other species, requested help identifying unusual characteristics of miscellaneous critters they would come across in our rural environment, invited me over to watch their companion animals and livestock births/hatchings, and even organized an opportunity for me to shadow a local veterinarian, which made me feel more connected to veterinary medicine than ever before. These early learning experiences shaped my views about science. However, being raised by my grandparents, who had only sixth and eighth-grade educations, foretold limited options for parlaying my love of animals into a career in veterinary medicine.

Although they did not have formal knowledge of veterinary medicine, my grandparents relied on our community for resources to help support my interests. In addition, they ensured that my teachers were aware of my interests and abilities and taught me the importance of believing in myself. But negative perceptions about what a Black boy from rural Alabama was capable of were powerful, so my primary interest in STEM became more secondary. Signals received in more formal learning environments told me I was not the type of child who should reasonably expect to become a veterinarian because I was neither White nor naturally talented at math. I am Black and struggle with math. My passion for veterinary medicine became less intense with each year due to the enormous burden of not fitting the mold and lacking knowledge and resources. My personal experience illustrates the tension between position, identity, and resources. It lends credence to the notion that STEM participation requires an interconnected ecosystem ripe with opportunities for formal and informal inclusive learning and support.

Today, I am not a veterinarian, but I am a scientist. Over the past seven years, my collaborative interdisciplinary teams have been awarded several federal grants from the National Science Foundation to support STEM learning, performance, and persistence for diverse students. So, as a researcher, I know the tools necessary for solidifying learners' STEM intentions and commitment. As an underrepresented minority (URM), I know barriers to educational persistence erode STEM interest and engagement. My experience as a researcher and a URM impacts my understanding of factors that promote STEM career awareness and participation among underserved populations and guides how the teams I work with approach designing and implementing early interventions that integrate and assess technologies for increasing STEM participation.

To close gaps in American competitiveness in the global STEM market, to increase representation of racial and ethnic minorities, and to encourage more women to participate, we must nurture STEM learning ecosystems in ways that empower young learners with STEM interests. To help you better understand how we can reframe STEM as a more viable career option for URM students, I provide recommendations based on lessons learned from my research. Hopefully, this provides a lens through which others can better understand factors that position underserved students for success in STEM and lead toward STEM careers.

Explanations for the lack of URM representation in STEM education and workforce linked with inadequate preparatory STEM education offer limited insights into relevant psychosocial factors that undergird URM interest in STEM participation. Here, I surmise that tending to students' learning ecologies provides inroads toward renegotiating longstanding science stereotypes and that both formal and informal experiences inside and outside the classroom can shift URM students' identities and position them for transformative career opportunities.

Positioning a diverse STEM workforce for STEM careers

Current trends indicate URM STEM participation drops perniciously from early k-12 education throughout postsecondary education due mainly to pervasive science stereotypes, which results in URM students having difficulty developing identities as scientists. Changing dominant cultural narratives about underrepresented minorities in STEM will require early intervention and intentional tending to their learning ecologies, as STEM researchers link pervasive cultural stereotypes regarding science to early declines in STEM interest and participation (Totonchi et al., 2021; Dare and Roehrig, 2020; Battel et al., 2021). Premature disengagement from STEM also leads to disproportionate rates of URM students attriting from the STEM workforce pipeline (Almeda and Baker, 2020). STEM careers are lucrative and in high demand. However, current trends in government and private sector STEM workforce needs point to field and degree-specific disparities between available jobs and qualified employees. Positioning URM students for STEM careers requires systematic, intentional diversification of the STEM workforce to increase the available pool of eligible STEM employees ready to serve a global imperative driving STEM expansion and American competitiveness.

Much has been written about the dire need to increase STEM participation, and America clearly cannot afford to lose its competitive edge in the global STEM market. Expected growth in STEM employment necessitates expanding the pool of qualified STEM workers (Langdon et al., 2011; Carnevale et al., 2011). Between May 2009 and May 2015, there was a 10.5% increase in STEM occupations (Fayer et al., 2017), and STEM occupations are expected to grow by 8.9% from 2014 to 2024 as compared to 6.4% for non-STEM jobs (Noonan, 2017). This growth equates to roughly 1.5 million job openings for selected STEM occupations between 2012 and 2022 (Vilorio, 2014). Considering increases in STEM job creation, statistics from the National Science Board (NSB, 2016) indicate that the STEM workforce is aging, with workers aged between 51 and 75 years increasing from 20% of the STEM workforce in 1993 to 34% in 2013. These data show that with increasing job growth and an aging STEM workforce, large increases in available jobs are unavoidable. For example, computer-based occupations can expect over 1 million job openings between 2014 and 2024 (Fayer et al., 2017). Intentional diversification of the STEM education-to-workforce pipeline is a promising method for increasing STEM participation and racial/ethnic parity in STEM careers.

Together, Blacks, Hispanics, and American Indians/Alaska Natives earned 28% of science and engineering degrees awarded to US citizens and permanent residents in 2012 at the associate's level, 20% at the bachelor's level, 18% at the master's level, and 12% at the doctoral level (NSF, 2016). Black students are underrepresented across STEM fields of study, with 6% degree attainment in the physical sciences, 4% degree attainment in engineering, and 5% degree attainment in mathematics (Gasman et al., 2017). These data underscore the absence of diversity in the STEM education-to-employment pipeline. Hispanics comprised more than 5% of workers in S&E occupations in 2010, whereas Blacks comprised 5%, American Indians/Alaska Natives comprised 0.2%, and Native Hawaiians/other Pacific Islanders comprised 0.2% (NSF, 2016). Although Blacks comprise 5% of S&E workers, they represent 12% of the working-age population. Asians represent 19% of S&E workers, but they only represent 5% of the working-age population (NSF, 2016). This racial disparity in STEM employment versus STEM career opportunity indicates expansive space for Black and other URM workers to fill available gaps, which could improve STEM workforce parity.

The United States Innovation and Competition Act of 2021 established a Directorate for Technology, Innovation, and Partnership at the National Science Foundation to "accelerate the development of new technologies and products that improve Americans' way of life, grow the economy and create new jobs, and strengthen and sustain US competitiveness for decades to come (National Science Foundation, 2022)." Positioning students from underserved backgrounds to pursue educational opportunities that will prepare them for future career success is critical for breaking down educational and economic empowerment barriers and helping ensure American STEM competitiveness. Exposing URM students to STEM careers, which are lucrative and abundant, is one resourceful way to introduce them to career options that they may not have considered previously.

The extant literature has identified numerous explanations for the continuing underrepresentation of URM people in STEM education and the workforce, ranging from poor academic preparation and low degree completion rates (National Science Foundation, 2016) to racial isolation and lack of institutional support (Rockinson-Szapkiw et al., 2021; Brown et al., 2016). These explanations emphasize a need to increase STEM participation to reposition America in the global STEM market. Some argue that traditional assessments of STEM performance measure few cognitive abilities and do not examine students' other artistic, technological, critical thinking, or social skills, which are essential 21st-Century Skills. As a result, educators often label URM students as lower achievers, which may limit their future potential (Al-Balushi and Al-Aamri, 2014; Sumarni and Kadarwati, 2020). At the undergraduate level, introductory STEM courses may discourage STEM retention due to high competition and unengaging pedagogy (Gasiewski et al., 2012; Seymour and Hewitt, 1997). However, hands-on, laboratory-based curricula have positively impacted STEM persistence, such that "lab-based education provides students with structured opportunities to develop a STEM identity and gain crucial research skills along the way" (Gasman et al., 2017, p. 15). In addition, academically engaged students are more likely to persist (Nelson Laird et al., 2008). For Black males with STEM interests, who are among the least represented in STEM careers (NSF, 2016), there should be a particular emphasis on nurturing formal and informal interactions to help integrate them into the culture of STEM to help them see STEM as a viable career option (Strayhorn, 2015).

Inclusive STEM

Designing an inclusive STEM curriculum to excite URM students about STEM and expose them to impactful careers could help reshape their identities and increase the number of STEM graduates who are interested in and prepared for STEM careers or graduate

study. Working with researchers from exercise physiology, psychology, biology, and chemistry, one of my collaborative teams set out to infuse laboratory-bench skills and summer research internships into the exercise physiology curriculum at a minority-serving institution (Brown et al., 2018). We used a novel series of chemistry, biology, and biophysics laboratory-bench modules to introduce students to cellular and molecular innovations in STEM that paralleled the evolution of bioengineering in exercise physiology curriculum. Regenerative medicine attracts those with STEM interests, and the field has received considerable attention due to the far-reaching implications of growing tissues and organs in a laboratory setting. We leveraged the excitement surrounding the field of regenerative medicine to engage students. We provided a subset of them with intensive summer research experiences with diverse partners around America like the Wake Forest Institute for Regenerative Medicine and research centers at Miami University (Ohio) and North Carolina State University.

Underrepresented students with STEM career interests face a fundamental disconnect between their academic identity and inclusion in science, such that group membership in STEM has traditionally been framed as White and male, which studies (Hurtado et al., 2008; Chang et al., 2008; Johnson and Elliott, 2020) indicate erodes the potential for URM students to integrate into the culture of STEM fully. Improving URM students' science attitudes could help to bridge that gap. Increasing engagement in a science classroom supports positive science attitude development, which, according to previous studies (Else-Quest et al., 2013) may have positive effects on STEM achievement and persistence. Our proposed psychosocial model (see Fig. 1) linked laboratory-bench techniques to cognitive and behavioral domains. These cognitive and behavioral domains represent pivots for change, such that improving students' attitudes toward science, increasing their confidence to complete science-related tasks, and engaging them in laboratory-based research serve to increase intentions to engage in undergraduate research and to persist in STEM. We positioned URM learners for STEM success by considering the multitude of cognitive and behavioral domains and how our curricular adjustments could support growth in those domains.

Attitudes are highly malleable and contextualized and can be changed due to internal and external factors (Gonzalez et al., 2021). The way URM students see themselves in relation to others, particularly those in STEM fields, changes depending on their narratives about minorities in STEM. Often, STEM learning interventions rely on increased social support to help students overcome perceived deficits. Students who perceive they are receiving adequate social support hold better science attitudes and more positive perceptions of their abilities in math and science (Rice et al., 2013; Rosenthal et al., 2011), and women's STEM identity development is significantly impacted by support factors and early exposure to STEM (Stitt Richardson et al., 2019). Peers or professionals can provide social support within the academic context as role models and mentors that help enhance learners' sense of security and belongingness in STEM (Taylor and Lobel, 1989), and can be particularly impactful when those mentors share lived experiences with URM students. Reducing the barriers to engaging with faculty members for women and URM students can help increase social support that transforms outcomes for URM students (Hurtado et al., 2011). However, attitudes and identity are only two factors in the STEM cultural backdrop that contribute to persistence for URM students. Innovative approaches to exciting students about STEM learning must also be rich in technology.

The future: strengthening the STEM learning ecosystem

My interest in examining issues at the intersection of STEM education, social, behavioral, and economic sciences emerged organically through lived experience and academically through promoting STEM careers to broaden participation among underserved populations. Theoretically, increasing America's STEM workforce requires intentional tending to students' STEM learning ecologies, early support for STEM interest, and technology-rich experiences that excite and engage learners. This may be particularly important for young learners at schools federally designated as "at risk." To meet the needs of diverse 21st century learners, my team and I developed stakeholder-vetted, culturally responsive STEM learning modules that align with Geneva Gay (Gay, 2002) and Gloria Ladson-Billings' (Ladson-Billings, 1995) Culturally Responsive Teaching and Culturally Relevant Pedagogy frameworks.

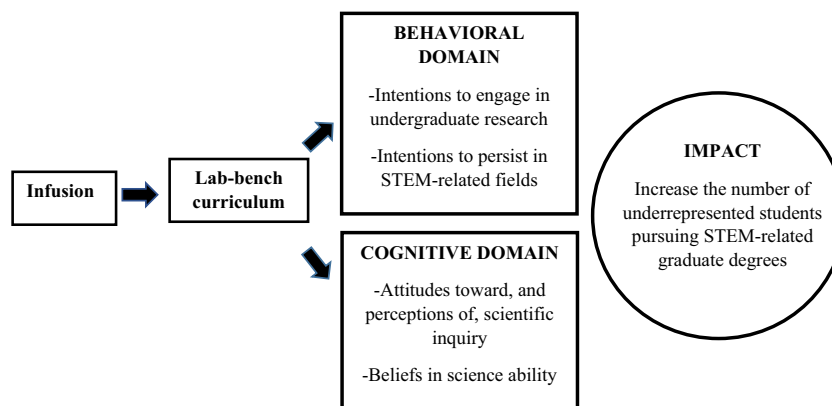


Fig. 1 Psychosocial framework for promoting stem persistence.

Embedding high-impact, technology-rich experiences into students learning ecology in culturally responsive ways might hasten learners' identity shift, supporting their excitement for STEM and their persistence toward STEM careers.

Our overarching goal for developing these STEM learning modules was to provide opportunities to embed third–fifth-grade students in intensive, technology-rich experiences that fostered STEM interest and engagement and to expose students to a range of STEM careers and promote positive STEM identities with our culturally relevant STEM career-based modules. In addition, we strove to increase students' awareness of STEM occupations, support student engagement via culturally relevant learning activities that enhance learners' identity, motivation, and mindset while also providing students opportunities to learn and practice those interpersonal and intra-personal skills necessary for the STEM workplace. Ultimately, we sought to advance our understanding of fostering student interest and capacities to participate in the STEM workforce.

Toward achieving our stated goals, we developed focus group protocols to extract valuable feedback from Winston-Salem/Forsyth County School teachers, third-fifth grade students, their parents, our STEM career partners, and Winston-Salem State University's STEM Engagement Faculty. Teachers noted that their efforts to expand STEM learning felt stifled by the lack of parental involvement, lack of funding, administrative red tape, and the coronavirus pandemic. Administrators reflected on the difficulty that the global COVID-19 pandemic presented, applauding us for delivering flexible, timely, and exciting modules but lamenting how they could not force teachers, students, or parents to engage during the pandemic. Student feedback showed a clear preference for hands-on STEM learning.

As one component of integrating new technologies to better position students for STEM success, we digitized the Draw-a-Scientist test (Chambers, 1983), a projective test used to assess students' stereotypic perceptions of scientists. Students' Draw-a-Scientist Test (DAST) drawings indicated an increase in female representation, as more of their "scientists" were women, which aligns well with recent DAST literature (Miller et al., 2018). However, that result was complicated because all current science teachers at partner schools were women; thus, students' impressions of what a scientist looks like may have been influenced by their experiences with available science teachers. However, students also drew more diverse characters in diverse environments (e.g., not limiting their scientists' work to indoor environments but also drawing scientists working outdoors). In addition, parents indicated that they needed more education about STEM career options. Still, they wanted their students to have a broad range of career options and be happy with whatever career they chose. These preliminary qualitative findings speak to a dire need for STEM scholars to better understand students' STEM learning ecosystems, which beckons opportunities to collaborate, pool resources, and examine issues at the intersection of STEM education and economics more thoroughly.

My formal training is in community psychology. Because my interests center on promoting best practices for encouraging STEM persistence toward STEM careers, I find the interconnectedness of students' leaning ecologies to be of particular importance. Traditionally, psychology has focused on the individual level of analysis to understand student learning better; however, STEM researchers and interventionists must consider the entirety of students' STEM learning ecosystem, including the learner, their parents, teachers, schools, communities and policies. Reinforcing students learning ecologies gives us inroads to preserve their STEM interests and identities, expose them to STEM career options, and support learning. Uri Bronfenbrenner's Ecological Systems Theory (Bronfenbrenner, 1979) provided the framework for our endeavor. Ultimately, our goal was to create an interconnected, synergistic STEM Learning Ecosystem to facilitate STEM career engagement and strengthen students' STEM identities for third–fifth-grade students, their parents, and underserved schools. In alignment with Bronfenbrenner's bioecological model, we aimed to impact students' microsystems by linking them, their families, and their teachers/schools with culturally relevant and responsive modules for exploring STEM career options. At the mesosystem level, we aimed to connect parents and schools better, increase parental involvement, and build a bridge between STEM learning at school and home. At the macrosystem level, our project targeted schools situated in economically depressed regions catering to underserved students. We deeply considered students' and parents' economic and social pressures to design a STEM curriculum that positively altered their future educational and career trajectories. By building authentic relationships with underserved schools, parents of students studying at underserved schools, teachers and administrators, and community partners, we sought to transform long-held dominant cultural narratives about science and build a pipeline to funnel underserved PreK-12 students toward STEM degrees and careers.

This interdisciplinary project has far-reaching impacts on multiple disciplines and impacts PreK-12 education, postsecondary education, and expands technology use/innovation across the academic, community, and professional settings. This exploratory project aimed to develop culturally relevant, responsive, technology-rich STEM Learning Modules that could be deployed in various settings to positively impact students' learning ecosystems. Our focus on improving learning strategies for students, parents, and teachers impacts STEM education and changes the narrative on STEM career access for underserved students and positions them for lucrative STEM careers. We sought to challenge long-held STEM narratives that exclude persons of color and make it difficult for them to see themselves as scientists, impacting their identity as scientists. Revamping and centering STEM learning in the context of culturally responsive pedagogy should improve key outcomes for underserved students, expand their awareness of STEM careers, and advance social justice in education. To be clear, both psychology and education are impacted by this project positively. Psychology is enriched because refocusing on student's culture and relatable interests and showing them role models from their ethnicity and culture will increase their motivation and perseverance. Education is enriched because using culturally relevant modules to teach will improve teacher-student relationships, regardless of either party's ethnicity/culture, and improve student performance because the material will be more engaging to them.

Conclusions and recommendations

Contemporary trends in STEM education offer new insights into best practices for exciting students about STEM. Exposing them to STEM careers early in the pipeline is a practical and effective way to ensure that they have the time, tools, and resources to attain STEM careers in the future. Of the more important lessons learned from my research is that retaining students in undergraduate STEM courses is much different than retaining students in the STEM education-to-career pipeline. Undergraduate students who choose STEM education have already been supported through the more dangerous early years where their STEM participation could be compromised, so my recommendations for positioning students for STEM careers center on tending to students' early learning ecologies, paving inroads toward reconciling longstanding science stereotypes, and offering experiences that can shift URM students' identities and prepare them for transformative career opportunities. Early support for STEM learning patches holes in the leaky STEM education-to-career pipeline, thus positioning more students for exciting, impactful, transformative STEM careers of the future. To this end, the following recommendations are provided as pillars for research and implementation to help better position students for STEM careers:

1. Work collaboratively across disciplines to elucidate and remediate factors that influence STEM learning and efficacy. Collaborative, interdisciplinary teams are powerful. No individual has all the answers to the pressing questions we have covered in this article but working collaboratively toward a goal of increasing STEM persistence and performance to position students for STEM success is a viable tool for delivering transformative research and designing impactful interventions. Working in silos is counterproductive for positioning diverse students for success.
2. Ensure your work is stakeholder informed and vetted. Despite my strong focus on delivering transformative STEM opportunities for students, I am not a parent and have limited knowledge about what parents want for their children in STEM education. Therefore, I advocate for working closely with not only students as key stakeholders, but also parents, teachers, and administrators. Everyone who interacts with the student has vital information to share about the students' needs, how those needs can best be met, and what resources have already been made available. This is important, because doubling up on resources may or may not be beneficial to students. For example, in our work in third–fifth-grade classrooms, we found that although our target schools were labeled as “at-risk”, previous federal dollars had equipped classrooms with good technology. Each student had digital tablets, smart boards and televisions were in every classroom, and microscopes, beakers, and other tools for science were abundant. It would have been futile to provide more of these tools for science, so working with key stakeholders informed my team about what needs had already been met so that we could meet unmet needs.
3. Although equal numbers of learners have STEM interest during early education (Wang and Degol, 2017), that interest declines precipitously for students with lower levels of STEM self-efficacy (Blotnicky et al., 2018). To support students learning efficacy and position them for STEM careers, we must intervene early and stay the course. We know that science stereotypes strongly impact students' perceptions of scientists, and that concrete perceptions of scientists are apparent as early as the second grade (Chambers, 1983). Early intervention is key, but we must continue supporting STEM engagement across the education-to-career pipeline. Positioning diverse students for STEM careers requires constant nurturing along the developmental pathway to STEM careers.
4. Provide teachers with comprehensive training. Ensuring teachers are trained on contemporary practices for engaging and impactful science for early learners is key to positioning those learners for the STEM jobs of the future. Further, teachers need training in inclusive STEM to support underserved learners' STEM identities. Finally, because technologies evolve rapidly, teachers must also stay abreast of new technologies for STEM learning, be proficient at using those technologies, and comfortable integrating those technologies into sometimes congested daily lesson plans.
5. Communities harbor human capital and empowering those communities to advocate for necessary resources to support human capital is beneficial to individuals, communities, and society. In terms of advocating for STEM, communities may benefit from clear messaging that conveys the transformative power of STEM, how science impacts communities, and how federal and local governments are doling out resources. In speaking with parents, we have found that they were rarely informed about the bevy of resources available to them for their young learners.

Pervasive stereotypes about STEM and a lack of access to necessary resources erodes underserved students' love of STEM learning and persistence toward STEM careers. The imperative driving global STEM competitiveness is ripe with opportunities for those from underserved backgrounds to produce innovations that benefit our global society, so ensuring that the ecological systems in which they learn are geared toward successfully positioning them for the STEM careers of the future is key. Forging a future that is more technologically advanced, inclusive, and just will require a diverse input. Working collaboratively with stakeholders, researchers and implementation scientists can positively impact global STEM outcomes by providing and assessing best-in-class practices for increasing representation in STEM and preparing diverse learners for 21-century careers.

Acknowledgments

This work was made possible by generous support from the National Science Foundation (NSF HBCU-UP TIP 1533476 and NSF ITEST 1850283).

References

- Al-Balushi, S.M., Al-Aamri, S.S., 2014. The effect of environmental science projects on students' environmental knowledge and science attitudes. *Int. Res. Geogr. Environ. Educ.* 23 (3), 213–227.
- Almeda, M., Baker, R.S., 2020. Predicting student participation in STEM careers: the role of affect and engagement during middle school. *J. Educ. Data Mining* 12 (2), 33–47.
- Battel, K., Foster, N., Barroso, L.V., Bhaduri, S., Mandala, K., Erickson, L., 2021. October. "We make the village"-Inspiring STEM among young girls and the power of creative engineering education in action. In: 2021 IEEE Frontiers in Education Conference (FIE). IEEE, pp. 1–7.
- Blotnick, K.A., Franz-Odenaal, T., French, F., Joy, P., 2018. A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *Int. J. STEM Educ.* 5 (1), 1–15.
- Bronfenbrenner, U., 1979. Contexts of child rearing: problems and prospects. *Am. Psychol.* 34 (10), 844.
- Brown, B.A., Henderson, J.B., Gray, S., Donovan, B., Sullivan, S., Patterson, A., Wagstaff, W., 2016. From description to explanation: an empirical exploration of the African-American pipeline problem in STEM. *J. Res. Sci. Teach.* 53 (1), 146–177.
- Brown, J.M., Guy, B.S., Henderson, D.X., Ebert, C.E., Harp, J., Markert, C.D., 2018. Regenerative medicine: a vehicle to infuse laboratory-bench modules into an exercise physiology curriculum. *Adv. Physiol. Educ.* 42 (1), 32–42.
- Carnevale, A.P., Smith, N., Melton, M., 2011. STEM: Science Technology Engineering Mathematics. Georgetown University Center on Education and the Workforce.
- Chambers, D.W., 1983. Stereotypic images of the scientist: the draw-a-scientist test. *Sci. Educ.* 67 (2), 255–265.
- Chang, M.J., Cerna, O., Han, J., Saenz, V., 2008. The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *Rev. High. Educ.* 31 (4), 433–464.
- Dare, E.A., Roehrig, G.H., 2020. Beyond Einstein and Explosions: understanding 6th grade girls' and boys' perceptions of physics, school, school science, and STEM careers. *J. Women Minorities Sci. Eng.* 26 (6).
- Else-Quest, N.M., Mineo, C.C., Higgins, A., 2013. Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychol. Women Q.* 37 (3), 293–309.
- Fayer, S., Lacey, A., Watson, A., 2017. BLS Spotlight on Statistics: STEM Occupations-Past, Present, and Future.
- Gasiewski, J.A., Eagan, M.K., Garcia, G.A., Hurtado, S., Chang, M.J., 2012. From gatekeeping to engagement: a multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Res. High. Educ.* 53 (2), 229–261.
- Gasman, M., Nguyen, T.H., Conrad, C.F., Lundberg, T., Commodore, F., 2017. Black male success in STEM: a case study of Morehouse College. *J. Divers. High. Educ.* 10 (2), 181.
- Gay, G., 2002. Preparing for culturally responsive teaching. *J. Teach. Educ.* 53 (2), 106–116.
- Gonzalez, A.M., Steele, J.R., Chan, E.F., Lim, S.A., Baron, A.S., 2021. Developmental differences in the malleability of implicit racial bias following exposure to counterstereotypical exemplars. *Dev. Psychol.* 57 (1), 102.
- Hurtado, S., Eagan, M.K., Cabrera, N.L., Lin, M.H., Park, J., Lopez, M., 2008. Training future scientists: predicting first-year minority student participation in health science research. *Res. High. Educ.* 49 (2), 126–152.
- Hurtado, S., Eagan, M.K., Tran, M.C., Newman, C.B., Chang, M.J., Velasco, P., 2011. "We do science here": underrepresented students' interactions with faculty in different college contexts. *J. Soc. Issues* 67 (3), 553–579.
- Johnson, A., Elliott, S., 2020. Culturally relevant pedagogy: a model to guide cultural transformation in STEM departments. *J. Microbiol. Biol. Educ.* 21 (1), 05.
- Langdon, D., McKittrick, G., Beede, D., Khan, B., Doms, M., 2011. STEM: Good Jobs Now and for the Future. ESA Issue Brief# 03-11. US Department of Commerce.
- Ladson-Billings, G., 1995. But that's just good teaching! the case for culturally relevant pedagogy. *Theor. Pract.* 34 (3), 159–165.
- Miller, D.I., Nolla, K.M., Eagly, A.H., Uttal, D.H., 2018. The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US Draw-a-Scientist studies. *Child Dev.* 89 (6), 1943–1955.
- National Science Foundation, 2016. Science and Engineering Indicators 2016. Digest (NSB-2016-2) | January 2016.
- National Science Foundation, 2022. NSF Establishes New Directorate for Technology, Innovation and Partnership. Available at: <https://beta.nsf.gov/news/nsf-establishes-new-directorate-technology-innovation-and-partnerships#hero-credit-caption>. (Accessed 30 March 2022).
- Nelson Laird, T.F., Chen, D., Kuh, G.D., 2008. Classroom practices at institutions with higher-than-expected persistence rates: what student engagement data tell us. *N. Dir. Teach. Learn.* 115, 85–99.
- Noonan, R., 2017. STEM Jobs: 2017. Update. ESA Issue Brief# 02-17. US Department of Commerce.
- Rice, L., Barth, J.M., Guadagno, R.E., Smith, G., McCallum, D.M., 2013. The role of social support in students' perceived abilities and attitudes toward math and science. *J. Youth Adolesc.* 42 (7), 1028–1040.
- Richardson, R.L.S., Guy, B.S., Perkins, K.S., 2019. "I am committed to engineering": the role of ego identity in Black women's engineering career persistence. *J. Negro Educ.* 88 (3), 281–296.
- Rockinson-Szapkiw, A., Watson, J.H., Gishbaugh, J., Wendt, J.L., 2021. A case for a virtual STEM peer-mentoring experience for racial and ethnic minority women mentees. *Int. J. Mentor. Coach. Educ.* 10 (3), 267–283.
- Rosenthal, L., London, B., Levy, S.R., Lobel, M., 2011. The roles of perceived identity compatibility and social support for women in a single-sex STEM program at a co-educational university. *Sex. Roles* 65 (9), 725–736.
- Seymour, E., Hewitt, N.M., 1997. Talking About Leaving, vol. 34. Westview Press, Boulder, CO.
- Strayhorn, T.L., 2015. Factors influencing Black males' preparation for college and success in STEM majors: a mixed methods study. *West. J. Black Stud.* 39 (1), 45.
- Sumarni, W., Kadarwati, S., 2020. Ethno-stem project-based learning: its impact to critical and creative thinking skills. *J. Pendidik. IPA Indones.* 9 (1), 11–21.
- Taylor, S.E., Lobel, M., 1989. Social comparison activity under threat: downward evaluation and upward contacts. *Psychol. Rev.* 96 (4), 569.
- Totonchi, D.A., Perez, T., Lee, Y.K., Robinson, K.A., Linnenbrink-Garcia, L., 2021. The role of stereotype threat in ethnically minoritized students' science motivation: a four-year longitudinal study of achievement and persistence in stem. *Contemp. Educ. Psychol.* 67, 102015.
- Vilorio, D., 2014. STEM 101: Intro to tomorrow's jobs. *Occup Outlook Q* 58 (1), 2–12.
- Wang, M.T., Degol, J.L., 2017. Gender gap in science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Educ. Psychol. Rev.* 29 (1), 119–140.

Student motivation in STEM: factors related to and measurement of STEM motivation

Tian Luo^a, Lilith Rüschenpöhler^b, and Jian Wang^{c,*}, ^a College of Teacher Education, Capital Normal University, Beijing, P.R. China; ^b Ludwigsburg University of Education, Ludwigsburg, Germany; and ^c College of Life Sciences, Beijing Normal University, Beijing, P.R. China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	401
The role of motivation in STEM education	401
Theoretical foundation of motivation in STEM	401
Toward a conceptualization of motivation in STEM	402
Factors related to student motivation in STEM	402
Gender, age, family, culture, and student motivation in STEM	403
Formal and informal STEM learning that promotes student motivation in STEM	404
Measurement of motivational constructs in STEM	405
Concluding remarks	405
References	406

Introduction

Science, technology, engineering, and mathematics (STEM) have been recognized as essential for the technological development and economic growth of a nation (Marginson et al., 2013). Compared to other occupations, careers in STEM domains are rapidly growing in numbers and well-paid. The need of talents in STEM disciplines cause the promotion STEM education in many countries, including the US (National Science Board, 2007), South Korea (Korean Ministry of Education, Science and Technology, 2011), Australia (Education Council, 2015), Singapore (STEM Inc, 2019), and China (National Institute of Educational Science, P. R. China, 2017). STEM education was firstly promoted in higher education in the US, emphasizing the importance of workforce readiness in STEM disciplines, and later at the K-12 stage as well. Presently, STEM education in K-12 not only includes education in each discipline but also integrated STEM education. In other words, STEM is not only taught as discrete subjects but also applied as problem-based learning or project-based learning in real-world context, where usually science, technology, engineering, and mathematics were flexibly applied to resolve problems or questions.

The role of motivation in STEM education

Motivation is a driving force for students' enrollment and participation in STEM learning and career development. Zimmerman and Schunk (2008) summarized that motivation could affect the what, when, and how in students' learning. In STEM disciplines, motivation could positively predict students' learning persistence and future STEM career choices (Bahar and Adiguzel, 2016; Dabney et al., 2012; Maltese et al., 2014; Maltese and Tai, 2011; Robnett and Leaper, 2013; Wang, 2013).

Furthermore, motivation is also a crucial educational outcome in STEM education. Many government initiatives regarding STEM education have put an emphasis on promoting students' STEM motivation (Korea Foundation for the Advancement of Science and Creativity, 2016; National Institute of Educational Sciences, 2017; National Research Council, 2014). Specifically, nurturing students' interest in STEM learning, willingness of entering STEM careers, and STEM identity are highlighted (National Research Council, 2014). For example, in *STEM integration in K-12 education: Status, prospects, and an agenda for research*, these outcomes include STEM course-taking, educational persistence, and graduation rates, STEM-related employment, STEM interest, and development of STEM identity.

Theoretical foundation of motivation in STEM

The definitions of motivation vary and there has been much disagreement about it in literature. According to the definition by Schunk et al. (2014), motivation should be viewed as "the process whereby goal-directed activities are instigated and sustained" and modern motivation theories recognize that such a process involves cognition, affect, and behaviors. This definition emphasizes that the nature of motivation as a process that can only be inferred from individuals' actions (e.g., effort, persistence, choice of tasks) or self-reports.

*Corresponding author

Motivation for STEM learning and motivation for STEM careers are both important topics in the literature for STEM education, both of which are important educational outcomes. To examine motivation for STEM learning, several theories are very informative, including self-determination theory (Ryan and Deci, 2000), self-regulated learning framework (Pintrich, 1999), attribution theory (Kelley, 1967; Weiner, 1972), intrinsic and extrinsic motivation (Ryan and Deci, 2000), etc. In addition, as STEM education has an emphasis on developing students' career interest, motivation theories related to career development, including theories of reasoned action and planned behavior (Ajzen, 1985), social-cognitive career theory (Lent et al., 1994) and expectancy-value theory (Wigfield and Eccles, 2000) are also frequently applied in research in STEM education. These theories provide different lenses to understand the external and internal factors that pull or push individuals to do an action (Hernandez et al., 2013).

Due to the limited length of this article, only two frequently used theories, namely social-cognitive career theory (Lent et al., 1994) and expectancy-value theory (Wigfield and Eccles, 2000) are introduced. These two theories, both derived from social cognitive theory developed by Bandura (1986) are not only applied to explain students' development of STEM learning goals but also STEM career goals and choices.

In social cognitive career theory (SCCT), the career-related outcomes (e.g., interest in science learning, career goal as an engineer, etc.) are determined by the "transactions occurring between persons and their environments" (Lent et al., 2002). SCCT proposed several related models explaining what factors promote individuals' interest and goals (Lent et al., 2002). In these models, the core factors incorporate self-efficacy (i.e., beliefs about an individual's own "capabilities to produce effects") and outcome expectations (i.e., expected results of a particular action), both of which are important contributors to interest and goals in one's learning and in careers, which may influence one's activity selection and practice. And one's activities could promote his/her goal attainment, which may result in providing more sources for self-efficacy and outcome expectations. In addition, background factors, such as gender and family, shed influence on students' learning experiences and shape their choices and outcomes. According to the review by van Tuijl and van der Molen (2016) and the meta-analysis by Lent et al. (2018), the proposed interest and choice model in SCCT has received much empirical support in STEM domains, most of which are quantitative studies focusing on college students.

Similar to SCCT, expectancy-value theory (EVT), which is built upon Atkinson's (1957) and Bandura's (1986) work, emphasized the contributing role of expectancies for success (i.e., the perceived level of success in an activity) and subjective task value (i.e., the perceived value of a task or activity) to students' learning and career goals. In this regard, EVT has a similar foundation with SCCT and some researchers (Parr and Bonitz, 2015) used both approaches in a complementary manner to frame their research. Compared to SCCT, EVT explains the students' forming of motivation-related perceptions and interpretations of experiences in a more detailed and comprehensive manner. For example, EVT provides explanations of how gender role stereotypes and cultural stereotypes of subjects and occupations may have impact on students' choices indirectly through several constructs. In addition, as Eccles and Harold (1991) initially developed the EVT model to explain gender differences in elementary students' and adolescents' mathematics expectancies and the values that affect their choices regarding courses and majors, EVT is very suitable to apply to educational research, including those at the K-12 stage.

From the different theoretical lenses, it is evident to summarize that experiences are the main sources for STEM motivation and that students' interest in STEM subjects/domains are predictive of their STEM career choices. These associations have been supported by many empirical studies in STEM education (Dabney et al., 2012; Lent et al., 2018; Maltese and Tai, 2011). The theoretical lenses are not only informative for researchers to examine, analyze, and explore the formation and maintenance of STEM motivation but also can serve as scaffolds for designing interventions (e.g., Harackiewicz et al., 2012).

Toward a conceptualization of motivation in STEM

Motivation in STEM is thus conceptualized in this article as follows: motivation in STEM is the process in which the STEM learning and pursuit of STEM careers are instigated and sustained. Here, STEM learning involves learning in science, technology, engineering, and technology domains. And STEM careers include careers that use a lot of science, technology, engineering, and mathematics in their work, including scientists, engineers, technologists, etc.

Motivation is a very broad concept with a long history of research and many paradigms (Hernandez et al., 2013). And thus, motivation here is not viewed as a single construct but a spectrum of constructs, which could be named as motivational constructs or motivational variables/factors. These constructs include (but are not limited to) self-efficacy, expectancies, values, goal orientations, and interest.

Factors related to student motivation in STEM

There have been numerous studies on how various external factors, including demographic factors such as gender, age, family background, such as socioeconomic status and parental occupations, school contexts, culture, and learning experiences may impact students' STEM motivation.

A well-known notion named "science capital" promoted by Archer et al. (2014) is a theoretical framework to examine how personal and background factors are related to science aspirations. It is a conceptual framework with a series of variables (i.e., everyday science engagement, future science aspirations, informal science activities, parental attitudes and practices, science teachers and lessons, self-efficacy in science, utility of science qualifications, valuing museums/museum experiences,

valuing science and scientists) to explain the “patterned natures of aspirations and educational participation” of students from age 11 to 16. A further study by [Moote et al. \(2020\)](#) suggested that science capital is strongly related to engineering aspirations but not strongly related to mathematics and technology aspirations. This suggests that science capital may need adaptations to better understand STEM motivation, but it does provide a good example of structuring a family of personal and background factors to understand STEM motivation.

Gender, age, family, culture, and student motivation in STEM

STEM is a field in which important gender differences appear. In general, boys seem to be more motivated for STEM ([Desy et al., 2011](#)). A quantitative study involving 6492 students from grade 3 to 12 indicated that being male is a significant predictor of STEM career aspirations in general ([Holmes et al., 2018](#)). However, differences exist between the STEM disciplines. Health and biology related fields are preferred by girls while physics and engineering related fields are preferred by boys ([Buccheri et al., 2011](#); [Koul et al., 2012](#)). Chemistry seems to be situated in-between these fields but nevertheless preferred by boys ([Cheung, 2009](#)) although the gender effect is not always significant ([Cheung, 2018](#)). The gender gaps and their difference across subject/career domains may be related to the gender-related occupational stereotypes of different STEM subject/careers held by students. And students’ stereotypical views may be influenced by their parents and teachers. As argued in the review by [Van Tuijl and van der Molen \(2016\)](#), teachers’ and parents’ stereotypical values regarding STEM subjects and careers need to be tackled.

Age seems to have an effect on STEM motivation as well. First, gender differences seem to be larger among elder students than among younger school children ([Baram-Tsabari and Yarden, 2011](#)). Further, motivation for STEM subjects declines as students grow older ([Beghetto et al., 2011](#); [Güngören, 2009](#); [Sorge, 2007](#); [Jacobs et al., 2002](#); [Lofgran et al., 2015](#)). This could be an effect of the fact that students dislike school science as [Potvin and Hasni \(2014\)](#) argue. However, it could also be an effect of identity formation among young people. It has been shown that on average, academic self-concepts decline over time because the young people’s individual personalities become more differentiated and more pronounced ([Marsh, 1989](#)). This is a phenomenon observed in general educational psychology which applies also to the field of STEM. Students develop identities (very pronounced motivations) in certain fields while in the remaining areas, motivation falls. This could explain the finding that motivation declines when looking at a group of students. Despite the general tendency of declining motivation in STEM, there are still some studies indicating contrary or mixed results ([Karaarslan and Sungur, 2011](#); [Luo et al., 2019](#); [Fortus and Vedder-Weiss, 2014](#)), which involve different sample choices, age span, and motivational constructs.

One popular metaphor related to the change of STEM aspirations/career choices with age is that of the “leaky STEM pipeline” (e.g., [Ginorio et al., 2002](#); [van den Hurk et al., 2019](#)): many students are interested in STEM but as students grow older, more and more decide not to continue pursuing a STEM career ([Sheldrake, 2020](#)). The metaphor is often used comparing the number of STEM freshmen with university graduates in STEM. A critique of this metaphor of the “leaky pipeline” (see e.g., [Mendick et al., 2017](#)) is that programs designed to fill the cracks in the STEM pipeline focus on supplying the industry with workforce and not on serving the people. However, many scholars adopt this metaphor with the goal to empower groups that are currently underrepresented, in order to support their inclusion in a job market which presents opportunities that could be interesting to them. Still, it seems important to look at the ultimate goals of studies in this area, even if these are not explicitly named.

The influential factor of the family context has long been recognized in motivation theories and career development theories ([Brown, 2002](#); [Ing, 2014](#)). Parents with STEM occupations ([Craig et al., 2018](#)) are predictive of students’ STEM career choices, such as choosing STEM majors in college. Researchers found that parental motivational practices can promote students’ persistence in STEM careers in adulthood through a longitudinal study using a nationally representative sample ([Ing, 2014](#)). An experimental study done by [Harackiewicz et al. \(2012\)](#) indicates that interventions informing parents about the importance of science and mathematics courses (via brochures and websites) may enable the high school students to choose more science and mathematics courses. To summarize, researchers have found that family factors have influence on students’ STEM career interest and such effect can be used to increase students’ STEM motivation.

Culture could play an important role in STEM motivation although knowledge in this area is still limited. The studies on motivation in STEM are not equally distributed around the globe. Research in this field is mainly from the US, Europe, Australia, and Israel and often based on data from these countries ([Potvin and Hasni, 2014](#)). This implies that knowledge on students’ motivation in STEM from other areas is limited. This opens the field for interesting investigations which could help to better understand, for instance, the prevailing gender differences in STEM. There is evidence that the gender relations differ depending on the cultural backgrounds of the students ([Rüschpöhler and Markic, 2019, 2020](#)). This suggests that the associations between the STEM fields and gender are products of the local culture. Learning from other cultures in which these associations are different could be insightful. For example, it seems like in emerging countries, STEM motivation is more pronounced than in developed countries and gender differences are smaller ([Shin et al., 2018](#)). Understanding these phenomena could provide interesting insights.

Another area in which new developments are made is that of identity research in STEM education. One of the reasons for the interest in science identities is the doing-being divide which has been observed: students enjoy doing science but they do not aspire to a career in the field ([Archer et al., 2010](#); [Zimmerman, 2012](#)). Identity formations which are important for choosing and persisting in a STEM career path thus seem to demand a far deeper personal connection with STEM than only interest or positive attitudes. For instance, it has been shown that science identity formation is easiest for middle-class boys belonging to the dominant ethnic group ([Archer et al., 2014](#); [Carlone et al., 2015](#)). Thus, this depends on gender, social class, and race. Identity research in STEM could provide explanations for the underrepresentation of certain groups, namely females ([OECD, 2009](#)) and non-dominant ethnic groups ([OECD, 2008](#)) in STEM fields. Identity research constitutes an interesting approach to STEM motivation research.

Further, it has been argued that it could be more insightful to concentrate on smaller entities when investigating student motivation. In a literature review, [Potvin and Hasni \(2014\)](#) found that gender differences appear more acutely when looking at particular themes in the science subjects and not at a general science or STEM motivation. They refer to a study showing that some students who have very little interest in the general field of STEM do have very pronounced interest in specific, narrow areas ([Yang, 2010](#)).

Formal and informal STEM learning that promotes student motivation in STEM

Literature on increasing STEM motivation through education for STEM is abundant. The reasons for this interest are no doubt manifold but certain rationales seem to underlie many parts of the research in this field. Both formal and informal STEM activities are crucial for initiating interest in STEM domains ([Young et al., 2017](#)) and the trigger for students' STEM motivation can start early in childhood ([Van Tuijl and van der Molen, 2016](#)) and happen across a wide age spectrum ([Maltese et al., 2014](#)). This argument has been supported by several studies exploring how students develop their motivation in STEM using a retrospective approach ([Dabney et al., 2012](#); [Maltese et al., 2014](#); [Maltese and Tai, 2010](#)). For example, [Maltese et al. \(2014\)](#) compared the reported experiences of STEM majors and non-STEM majors and summarized that the pathways to STEM are very diverse without a clear preferential way. [Maltese and Tai \(2010\)](#) analyzed experiences reported by scientists and graduate students and found that both school-related activities and self-initiated activities accounted for their interest in science. And it is worth noting that 65% of the participants reported an initial interest in science before middle school.

Among students' educational experiences, some free-choice STEM activities, such as assembling machines ([Maltese et al., 2014](#)), participation in STEM clubs and competitions ([Sahin, 2013](#)), talking to others about science and reading/watching science materials ([Dou et al., 2019](#)) are predictive of students' future STEM career choices. On the other hand, researchers found that participation in advanced placement (AP) courses and students' achievement are not predictive of students' choices of STEM major ([Maltese and Tai, 2011](#)). It seems that the experience of STEM learning which is motivated by students' inner interest rather than extrinsic motivation is more closely related to students' future STEM career choices.

One important consideration in the literature regarding interventions and STEM motivation are methodological difficulties. Articles about interventions which did not increase student motivation are scarce ([Potvin and Hasni, 2014](#)) although negative results would be very insightful for differentiating between interventions that work and those that do not. Many extracurricular programs like STEM camps (e.g., [Halonen and Aksela, 2018](#)), trips to the industry or to university (e.g., [Glowinski and Bayrhuber, 2011](#)), museum visits (e.g., [Martin et al., 2016](#)), competitions ([Höffler et al., 2017](#)) etc. seem to be beneficial for motivation in STEM. They allow for science learning in non-formal settings which could promote a transition of science—framed as being a school subject—toward science, framed as being a field of leisure interest. However, one problem with these studies is that they usually concern selected students who have opted to participate in the activity in their spare time. It is not possible to derive conclusions about how to increase STEM motivation among those who do not opt for this kind of programs. Furthermore, showing that an intervention did increase motivation does not yet allow to infer how this influence might be wielded. Making sure to avoid prior selections ([Potvin and Hasni, 2014](#)) could be one factor for improving research in this field. Mixed methods design studies with quantitative assessments of the strength of motivation in STEM and qualitative analyses of how motivation is increased could be interesting as well.

Among the multitude of factors that might make interventions successful in increasing motivation in STEM, several factors seem to be commonly recognized in current research literature. First, making STEM education relevant seems to be an important factor. Relevance is a term that often remains undefined and unconceptualized. It is sometimes translated with “contextualization” in STEM education although the terms differ from one another. Relevance, according to the model by [Stuckey et al. \(2013\)](#), is a three-dimensional construct comprising the dimensions of relevance for the student as an individual, for the student as a member of society, and for the student as a person who needs to qualify for a future job. Contextualization is a technique that can increase relevance on different dimensions. Potvin and Hasni conclude in their review that interventions containing contextualization “reported positive effects but from very different approaches” (2014, p. 105). This could be a reflection of a lack of differentiation between the aspects of “contextualization”. However, contextualized STEM learning seems to be beneficial for motivation (for an overview see [King, 2012](#)). The perceived connections to real life could predict students' career choice about whether to enter a science major at the college level ([Hutchinson-Anderson et al., 2015](#)).

Some STEM interventions increase relevance by incorporating information and participation of STEM professionals in the curriculum/activities. Some researchers provide students with videos about STEM professionals' work ([Colston et al., 2017](#); [Wyss et al., 2012](#)) and some researchers invite STEM professionals to give seminars ([Cantrell and Ewing-Taylor, 2009](#)). In addition, some researchers invite STEM professionals to work with students ([Barab and Hay, 2001](#)). These approaches may enable students to understand how STEM knowledge and skills serve the society and encourage them to pursue STEM careers in the future. The results showed that, students' exposure to interviews about STEM professionals contributes to an increase in career interest ([Colston et al., 2017](#)).

Second, besides increasing the relevance for the students, researchers have argued that the level of integration of STEM domains seems to be beneficial for motivation in STEM fields ([Thibaut et al., 2018](#)). Solving real-world STEM problems in STEM class usually requires students to apply knowledge and skills from several disciplines in a flexible way, making STEM education naturally related to the integration of knowledge and skills in science, technology, engineering and mathematics. [Yoon et al. \(2014\)](#) found that interventions working with science, technology, and engineering as an integrated subject could promote elementary students' engineering career identities compared to interventions without integration. A meta-summary of qualitative findings regarding

STEM education showed that it could promote students' willingness of choosing STEM professions (Kanadli, 2019). In addition, there are also some meta-analysis studies supporting the argument that STEM integration is beneficial for students' mathematics achievement (Mustafa et al., 2016), or more effective in promoting achievement in general, compared to other forms of integration (e.g., E-M and M-S-T) in integrated STEM education (Becker and Park, 2011). And these effects of STEM integration on students' achievement may further help build STEM motivation.

Third, inquiry-based learning may promote STEM motivation. Inquiry-based learning is one of the major approaches in STEM education (Thibaut et al., 2018) and has a long tradition in science and mathematics education, which is based on students engaging in a research process. There are different levels of inquiry-based learning which allow for more or less degrees of freedom and structuration by the teacher (see Rönnebeck et al., 2016 for an overview). Studies have indicated positive effects of inquiry-based learning on student motivation (Cairns and Areepattamannil, 2019; Chin and Osborne, 2008). Besides these three important factors (contextualization, inquiry-based learning, and the level of integration of the STEM subjects into one single domain), many other aspects in learning might be important for promoting STEM motivation. For further literature, see the review by Potvin and Hasni (2014).

Measurement of motivational constructs in STEM

Most existing measures for motivational constructs in STEM focus on the cognitive or affective domain of motivation. There are already many existing instruments for measuring science motivation (e.g., Glynn et al., 2009; Tuan et al., 2005), technology motivation (Liou and Kuo, 2014), engineering motivation (e.g., Mamaril et al., 2016), and mathematics motivation (e.g., Lim and Chapman, 2015), and some motivation-related measures for STEM as four separate-but-related domains (Guzey et al., 2014; Kier, 2014; Luo et al., 2019; Oh et al., 2013; Toker and Ackerman, 2012; Tyler-Wood et al., 2010; Unfried et al., 2015). The following review will put more emphasis on a few measures for STEM as four related domains.

As most of the existing measuring tools target different sub-categories of motivation, or some motivational constructs, the existing measurement tools related to STEM motivation also vary regarding their names, structures, and functions. For example, Oh et al. (2013) developed the Educational and Career Interest in STEM measure for high school students, which contains 12 items measuring students' interest in the four STEM domains respectively and 8 items measuring their interests in STEM as a unified concept. The tool has been validated on high school students with a Cronbach's alpha of 0.88 and three factors, namely educational and career interest in science, technology, and mathematics. The shortcoming of the measure is that the items regarding engineering were deleted finally, which may be caused by students' unfamiliarity with engineering.

In addition, Kier et al. (2014) developed a survey called STEM career interest survey (STEM-CIS) based upon SCCT. The STEM-CIS has items on personal goals, outcome expectations, interests, personal inputs, and contextual supports and barriers. The survey has been validated on 4th to 8th grade students and factor analysis confirmed that it has four subscales for science, technology, engineering, and mathematics respectively with a total Cronbach's alpha of the subscales ranging from 0.77 to 0.89.

Only a few measures can be used to examine the behavior aspect of motivation. Luo et al. (2019) developed an instrument measuring students' continuing motivation in STEM (STEM CM) with reference to the science continuing motivation scale developed by Fortus and Vedder-Weiss (2014). STEM continuing motivation is a specific type of motivation that exhibits students' tendencies of engaging in free-choice STEM-related activities. The STEM CM measure was validated using Rasch analysis on 7th and 8th grade students and has four subscales on science, technology, engineering, and mathematics respectively, with a total Cronbach's alpha of 0.91.

The above-mentioned measures all treat STEM as four separate but related subjects. There are also a few measures for motivational constructs that treat STEM as an integrated field (Luo et al., 2021; Milner et al., 2014; Nugent et al., 2015). These measures usually treat STEM learning as learning through practices that incorporate multiple STEM subjects. For example, the self-efficacy for STEM activities (STEMaSE) survey was developed and validated on upper-elementary students to evaluate their self-efficacy in STEM activities through practices such as "to test and compare different solutions" and "to represent the data with graphs" (Luo et al., 2021). Moreover, the STEM Career Interest Test and STEM Career Self-Efficacy Test developed for college students measure their career interest and self-efficacy in tasks such as "measure the speed of electrons" and "design a diagnostic routine for a power plant" (Milner et al., 2014). All these three surveys have one factor, according to exploratory factor analysis and confirmatory factor analysis, suggesting students' self-efficacy and career interest in these tasks exhibited an integrated nature when measured.

Concluding remarks

To date, we have learned from literature that STEM motivation is a driving force and an important outcome of STEM learning. It may start early in an individual's life and there is not a singular pathway but various ways for its development. There are gender differences in STEM motivation in different STEM disciplines. Generally, students' STEM motivation seems to decline with age. Family background as well as parental influence seem to play a role in students' STEM motivation whereby parent education may help promote students' STEM motivation.

Abundant literature has shown the positive influence of both informal and formal education on STEM motivation. However, not much is yet known about which specific factors in the integrated STEM interventions work on students' motivation. Better contextualizing (including linking STEM learning with information about STEM careers), a higher level of STEM integration, and the use of the inquiry-based approach seem to have positive effects on students' motivation. Future studies are needed to elucidate what interventions do and which do not promote STEM motivation.

Some suggestions for future studies are proposed as follows. First, there are many widely applied theories (e.g., EVT, etc), or perspectives (e.g., pipeline metaphor, science capital) that can be applied in STEM education research as frameworks to analyze the formation, maintenance, and shifts in students' motivation. The usage of such theories and perspectives may situate the research to a well-grounded foundation of previous studies and may serve as reference in designing interventions promoting students' STEM motivation.

Second, many previous studies regarding STEM motivation do not explicitly specify the level of integration or the levels of integration are inconsistent between the intervention and the measurement. The level of integration in STEM education suggested by English (2016) can be used as a reference. According to English (2016), the integration level of STEM ranges from disciplinary to multidisciplinary, interdisciplinary, and transdisciplinary. Disciplinary STEM indicates that knowledge and skills are acquired in each of the four domains in STEM separately. The second level of integration, multidisciplinary, means that knowledge and skills are learned from each domain but under a common theme. The third level is named interdisciplinary, in which knowledge and skills that are related are learned from more than one discipline to deepen the understanding. The highest level of integration is named transdisciplinary, which means that knowledge and skills are learned from more than one discipline and are applied to real-world problems or projects. For future studies, it will be a good practice to identify the level of STEM integration as a theoretical stance and to keep the level of integration consistent between the interventions (if applicable) and the measurement.

So far, many studies regarding STEM motivation treat STEM as disciplinary or multidisciplinary concepts. With the development of STEM education, more studies regarding higher levels of STEM integration (i.e., interdisciplinary or transdisciplinary) will be needed as they are promoted in the education community as a future trend. Nevertheless, pursuing higher levels of STEM integration, on the other hand, increases the difficulty in designing interventions and measuring instruments.

References

- Ajzen, I., 1985. From intentions to actions: a theory of planned behavior. In: Kuhl, J., Beckmann, J. (Eds.), *Action Control*. Springer, pp. 11–39. https://doi.org/10.1007/978-3-642-69746-3_2.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., Wong, B., 2010. "Doing" science versus "being" a scientist: examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Sci. Educ.* 94 (4), 617–639. <https://doi.org/10.1002/sce.20399>.
- Archer, L., DeWitt, J., Willis, B., 2014. Adolescent boys' science aspirations: masculinity, capital, and power. *J. Res. Sci. Teach.* 51 (1), 1–30. <https://doi.org/10.1002/tea.21122>.
- Atkinson, J.W., 1957. Motivational determinants of risk-taking behavior. *Psychol. Rev.* 64 (6), 359–372. <https://doi.org/10.1037/h0043445>.
- Bandura, A., 1986. *Social foundations of thought and action*. NJ: Prentice Hall, Englewood Cliffs.
- Bahar, A., Adiguzel, T., 2016. Analysis of factors influencing interest in STEM career: comparison between American and Turkish high school students with high ability. *J. STEM Educ.* 17 (3), 64–69.
- Barab, S.A., Hay, K.E., 2001. Doing science at the elbows of experts: issues related to the science apprenticeship camp. *J. Res. Sci. Teach.* 38 (1), 70–102. [https://doi.org/10.1002/1098-2736\(200101\)38:1<70::AID-TEA5>3.0.CO;2-L](https://doi.org/10.1002/1098-2736(200101)38:1<70::AID-TEA5>3.0.CO;2-L).
- Baram-Tsabari, A., Yarden, A., 2011. Quantifying the gender gap in science interests. *Int. J. Sci. Math. Educ.* 9 (3), 523–550.
- Becker, K., Park, K., 2011. Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: a preliminary meta-analysis. *J. STEM Educ.* 12 (5 & 6), 23–37.
- Beghetto, R.A., Kaufman, J.C., Baxter, J., 2011. Answering the unexpected questions: exploring the relationship between students' creative self-efficacy and teacher ratings of creativity. *Psychol. Aesthet. Creat. Arts* 5 (4), 342–349. <https://doi.org/10.1037/a0022834>.
- Brown, D., 2002. *Career Choice and Development*. John Wiley & Sons.
- Buccheri, G., Gürber, N.A., Brühwiler, C., 2011. The impact of gender on interest in science topics and the choice of scientific and technical vocations. *Int. J. Sci. Educ.* 33 (1), 159–178. <https://doi.org/10.1080/09500693.2010.518643>.
- Cairns, D., Areepattamannil, S., 2019. Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Res. Sci. Educ.* 49 (1), 1–23. <https://doi.org/10.1007/s11165-017-9639-x>.
- Cantrell, P., Ewing-Taylor, J., 2009. Exploring STEM career options through collaborative high school seminars. *J. Eng. Educ.* 98 (3), 295–303. <https://doi.org/10.1002/j.2168-9830.2009.tb01026.x>.
- Carlone, H.B., Webb, A.W., Archer, L., Taylor, M., 2015. What kind of boy does science? A critical perspective on the science trajectories of four scientifically talented boys. *Sci. Educ.* 99 (3), 438–464. <https://doi.org/10.1002/sce.21155>.
- Cheung, D., 2009. Students' attitudes toward chemistry lessons: the interaction effect between grade level and gender. *Res. Sci. Educ.* 39 (1), 75–91. <https://doi.org/10.1007/s11165-007-9075-4>.
- Cheung, D., 2018. The key factors affecting students' individual interest in school science lessons. *Int. J. Sci. Educ.* 40 (1), 1–23. <https://doi.org/10.1080/09500693.2017.1362711>.
- Chin, C., Osborne, J., 2008. Students' questions: a potential resource for teaching and learning science. *Stud. Sci. Educ.* 44 (1), 1–39. <https://doi.org/10.1080/03057260701828101>.
- Colston, N., Thomas, J., Ley, M.T., Ivey, T., Utley, J., 2017. Collaborating for early-age career awareness: a comparison of three instructional formats. *J. Eng. Educ.* 106 (2), 326–344. <https://doi.org/10.1002/jee.20166>.
- Craig, C.J., Verma, R., Stokes, D., Evans, P., Abrol, B., 2018. The influence of parents on undergraduate and graduate students' entering the STEM disciplines and STEM careers. *Int. J. Sci. Educ.* 40 (6), 621–643. <https://doi.org/10.1080/09500693.2018.1431853>.
- Dabney, K.P., Tai, R.H., Almarode, J.T., Miller-Friedmann, J.L., Sonnet, G., Sadler, P.M., Hazari, Z., 2012. Out-of-school time science activities and their association with career interest in STEM. *Int. J. Sci. Educ.* 2 (1), 63–79. <https://doi.org/10.1080/21548455.2011.629455>.

- Desy, E.A., Peterson, S.A., Brockman, V., 2011. Gender differences in science-related attitudes and interests among middle school and high school students. *Sci. Educ.* 20 (2), 23–30.
- Dou, R., Hazari, Z., Dabney, K., Sonnert, G., Sadler, P., 2019. Early informal STEM experiences and STEM identity: the importance of talking science. *Sci. Educ.* 103 (3), 623–637.
- Eccles, J.S., Harold, R.D., 1991. Gender differences in sport involvement: applying the Eccles' expectancy-value model. *J. Appl. Sport Psychol.* 3 (1), 7–35.
- Education Council, 2015. National STEM School Education Strategy: A Comprehensive Plan for Science, Technology, Engineering and Mathematics Education in Australia. Education Council (of Australian Education Ministers).
- English, L.D., 2016. STEM education K-12: perspectives on integration. *Int. J. STEM Educ.* 3 (1), 3. <https://doi.org/10.1186/s40594-016-0036-1>.
- Fortus, D., Vedder-Weiss, D., 2014. Measuring students' continuing motivation for science learning. *J. Res. Sci. Teach.* 51 (4), 497–522. <https://doi.org/10.1002/tea.21136>.
- Ginorio, A.B., Huston, M., Frevert, K., Seibel, J.B., 2002. The rural girls in science project: from pipelines to affirming science education. *J. Women Minor. Sci. Eng.* 8 (3–4), 22.
- Glowinski, I., Bayrhuber, H., 2011. Student labs on a university campus as a type of out-of-school learning environment: assessing the potential to promote students' interest in science. *Int. J. Environ. Sci. Educ.* 6 (4), 371–392.
- Glynn, S.M., Taasobshirazi, G., Brickman, P., 2009. Science motivation questionnaire: construct validation with nonscience majors. *J. Res. Sci. Teach.* 46 (2), 127–146. <https://doi.org/10.1002/tea.20267>.
- Güngören, S., 2009. The Effect of Grade Level on Elementary School Students' Motivational Beliefs in Science. Master's Dissertation. Middle East Technical University, Ankara, Turquia.
- Guzey, S.S., Harwell, M., Moore, T., 2014. Development of an instrument to assess attitudes toward science, technology, engineering, and mathematics (STEM). *Sch. Sci. Math.* 114 (6), 271–279. <https://doi.org/10.1111/ssm.12077>.
- Hälonen, J., Aksela, M., 2018. Non-formal science education: the relevance of science camps. *Lumat* 6 (2), 64–85.
- Harackiewicz, J.M., Rozek, C.S., Hulleman, C.S., Hyde, J.S., 2012. Helping parents to motivate adolescents in mathematics and science: an experimental test of a utility-value intervention. *Psychol. Sci.* 23 (8), 899–906. <https://doi.org/10.1177/0956797611435530>.
- Hernandez, P.R., Schultz, P., Estrada, M., Woodcock, A., Chance, R.C., 2013. Sustaining optimal motivation: a longitudinal analysis of interventions to broaden participation of underrepresented students in STEM. *J. Educ. Psychol.* 105 (1), 89–107. <https://doi.org/10.1037/a0029691>.
- Höfler, T.N., Bonin, V., Parchmann, I., 2017. Science vs. sports: motivation and self-concepts of participants in different school competitions. *Int. J. Sci. Math. Educ.* 15 (5), 817–836. <https://doi.org/10.1007/s10763-016-9717-y>.
- Holmes, K., Gore, J., Smith, M., Lloyd, A., 2018. An integrated analysis of school students' aspirations for STEM careers: which student and school factors are most predictive? *Int. J. Sci. Math. Educ.* 16 (4), 655–675. <https://doi.org/10.1007/s10763-016-9793-z>.
- Hutchinson-Anderson, K., Johnson, K., Craig, P.A., 2015. Students' perceptions of factors influencing their desire to major or not major in science. *J. Coll. Sci. Teach.* 45 (2), 78–85.
- Ing, M., 2014. Can parents influence children's mathematics achievement and persistence in STEM careers? *J. Career Dev.* 41 (2), 87–103. <https://doi.org/10.1177/0894845313481672>.
- Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S., Wigfield, A., 2002. Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73 (2), 509–527. <https://doi.org/10.1111/1467-8624.00421>.
- Kanadli, S., 2019. A meta-summary of qualitative findings about STEM education. *Int. J. InStruct.* 12 (1), 959–976.
- Karaarslan, G., Sungur, S., 2011. Elementary students' self-efficacy beliefs in science: role of grade level, gender, and socio-economic status. *Sci. Educ. Int.* 22 (1), 72–79.
- Kelley, H.H., 1967. Attribution theory in social psychology. *Nebr. Symp. Motiv.* Paper 15, 192–238.
- Kier, M.W., Blanchard, M.R., Osborne, J.W., Albert, J.L., 2014. The development of the STEM career interest survey (STEM-CIS). *Res. Sci. Educ.* 44 (3), 461–481. <https://doi.org/10.1007/s11165-013-9389-3>.
- King, D., 2012. New perspectives on context-based chemistry education: using a dialectical sociocultural approach to view teaching and learning. *Stud. Sci. Educ.* 48 (1), 51–87. <https://doi.org/10.1080/03057267.2012.655037>.
- Korea Foundation for the Advancement of Science and Creativity, 2016. Introduction to STEAM Education. KOFAC, Seoul.
- Korean Ministry of Education, Science and Technology, 2011. The Second Basic Plan to Foster and Support the Human Resources in Science and Technology (2011–2015). MEST, Seoul.
- Koul, R., Roy, L., Lerdpornkulrat, T., 2012. Motivational goal orientation, perceptions of biology and physics classroom learning environments, and gender. *Learn. Environ. Res.* 15 (2), 217–229. <https://doi.org/10.1007/s10984-012-9111-9>.
- Lent, R.W., Brown, S.D., Hackett, G., 1994. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122. <https://doi.org/10.1006/jvbe.1994.1027>.
- Lent, R.W., Brown, S.D., Hackett, G., 2002. Social cognitive career theory. In: *Career Choice and Development*. John Wiley & Sons, Inc, pp. 255–311.
- Lent, R.W., Sheu, H., Miller, M., Cusick, M., Penn, L., et al., 2018. Predictors of science, technology, engineering, and mathematics choice options: a meta-analytic path analysis of the social-cognitive choice model by gender and race/ethnicity. *J. Counsel. Psychol.* 65 (1), 17–35. <https://doi.org/10.1037/cou0000243>.
- Lim, S.Y., Chapman, E., 2015. Adapting the academic motivation scale for use in pre-tertiary mathematics classrooms. *Math. Educ. Res. J.* 27 (3), 331–357. <https://doi.org/10.1007/s13394-014-0140-9>.
- Liou, P., Kuo, P., 2014. Validation of an instrument to measure students' motivation and self-regulation towards technology learning. *Res. Sci. Technol. Educ.* 32 (2), 79–96. <https://doi.org/10.1080/02635143.2014.893235>.
- Lofgran, B.B., Smith, L.K., Whiting, E.F., 2015. Science self-efficacy and school transitions: elementary school to middle school, middle school to high school. *Sch. Sci. Math.* 115 (7), 366–376. <https://doi.org/10.1111/ssm.12139>.
- Luo, T., Wang, J., Liu, X., Zhou, J., 2019. Development and application of a scale to measure students' STEM continuing motivation. *Int. J. Sci. Educ.* 41 (14), 1885–1904. <https://doi.org/10.1080/09500693.2019.1647472>.
- Luo, T., So, W.M.W., Li, W.C., Yao, J., 2021. The development and validation of a survey for primary students' self-efficacy in STEM activities. *J. Sci. Educ. Technol.* 30 (4), 1–12.
- Maltese, A.V., Tai, R.H., 2010. Eyeballs in the fridge: sources of early interest in science. *Int. J. Sci. Educ.* 32 (5), 669–685. <https://doi.org/10.1080/09500690902792385>.
- Maltese, A.V., Tai, R.H., 2011. Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among US students. *Sci. Educ.* 95 (5), 877–907. <https://doi.org/10.1002/sce.20441>.
- Maltese, A.V., Melki, C.S., Wiebke, H.L., 2014. The nature of experiences responsible for the generation and maintenance of interest in STEM. *Sci. Educ.* 98 (6), 937–962. <https://doi.org/10.1002/sce.21132>.
- Mamaril, N.A., Usher, E.L., Li, C.R., Economy, D.R., Kennedy, M.S., 2016. Measuring undergraduate students' engineering self-efficacy: a validation study. *J. Eng. Educ.* 105 (2), 366–395. <https://doi.org/10.1002/jee.20121>.
- Marginson, S., Tytler, R., Freeman, B., Roberts, K., 2013. STEM: Country Comparisons: International Comparisons of Science, Technology, Engineering and Mathematics (STEM) Education. Final report. Australian Council of Learned Academies, Melbourne, VIC.
- Marsh, H.W., 1989. Age and sex effects in multiple dimensions of self-concept: preadolescence to early adulthood. *J. Educ. Psychol.* 81 (3), 417–430. <https://doi.org/10.1037/0022-0663.81.3.417>.
- Martin, A.J., Durksen, T.L., Williamson, D., Kiss, J., Ginns, P., 2016. The role of a museum-based science education program in promoting content knowledge and science motivation. *J. Res. Sci. Teach.* 53 (9), 1364–1384. <https://doi.org/10.1002/tea.21332>.
- Mendick, H., Berge, M., Danielsson, A., 2017. A critique of the stem pipeline: young people's identities in Sweden and science education policy. *Br. J. Educ. Stud.* 65 (4), 481–497. <https://doi.org/10.1080/00071005.2017.1300232>.

- Milner, D.I., Horan, J.J., Tracey, T.J.G., 2014. Development and evaluation of STEM interest and self-efficacy tests. *J. Career Assess.* 22 (4), 642–653. <https://doi.org/10.1177/1069072713515427>.
- Moote, J., Archer, L., DeWitt, J., MacLeod, E., 2020. Comparing students' engineering and science aspirations from age 10 to 16: investigating the role of gender, ethnicity, cultural capital, and attitudinal factors. *J. Eng. Educ.* 109 (1), 34–51.
- Mustafa, N., Ismail, Z., Tasir, Z., Mohamad Said, M.N.H., 2016. A meta-analysis on effective strategies for integrated STEM education. *Adv. Sci. Lett.* 22 (12), 4225–4228. <https://doi.org/10.1166/asl.2016.8111>.
- National Institute of Educational Sciences, P. R. of China, 2017. White Paper on STEM Education in China. National Institute of Educational Sciences, Beijing. Available at: <http://beed.Asia/wpcontent/uploads/2017/06/%E4%B8%AD%E5%9B%BDSTEM%E6%95%99%E8%82%B2%E7%99%BD%E7%9A%AE%E4%B9%A6%EF%BC%88%E7%B2%BE%E5%8D%8E%E7%89%88%EF%BC%89.pdf>.
- National Research Council, 2014. STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research. National Academies Press, Washington, DC.
- National Science Board (US), 2007. A National Action Plan for Addressing the Critical Needs of the US Science, Technology, Engineering, and Mathematics Education System. National Science Foundation.
- Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C.R., et al., 2015. A model of factors contributing to STEM learning and career orientation. *Int. J. Sci. Educ.* 37 (7), 1067–1088. <https://doi.org/10.1080/09500693.2015.1017863>.
- OECD, 2008. Encouraging Student Interest in Science and Technology Studies. OECD Publishing, Paris.
- OECD, 2009. Chart A4.6 Tertiary Graduates in Science-Related Fields Among 25-34 Year-Olds in Employment, by Gender (2009). In: Education at a Glance 2011. Retrieved September 4, 2018, from: <http://statlinks.oecdcode.org/962011041P1G020.XLS>.
- Oh, Y.J., Jia, Y., Lorentson, M., LaBanca, F., 2013. Development of the educational and career interest scale in science, technology, and mathematics for high school students. *J. Sci. Educ. Technol.* 22 (5), 780–790. <https://doi.org/10.1007/s10956-012-9430-8>.
- Parr, A.K., Bonitz, V.S., 2015. Role of family background, student behaviors, and school-related beliefs in predicting high school dropout. *J. Educ. Res.* 1–11. <https://doi.org/10.1080/00220671.2014.917256>.
- Pintrich, P.R., 1999. The role of motivation in promoting and sustaining self-regulated learning. *Int. J. Educ. Res.* 31 (6), 459–470. [https://doi.org/10.1016/S0883-0355\(99\)00015-4](https://doi.org/10.1016/S0883-0355(99)00015-4).
- Potvin, P., Hasni, A., 2014. Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. *Stud. Sci. Educ.* 50 (1), 85–129. <https://doi.org/10.1080/03057267.2014.881626>.
- Robnett, R.D., Leaper, C., 2013. Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. *J. Res. Adolesc.* 23 (4), 652–664. <https://doi.org/10.1111/jora.12013>.
- Rönnebeck, S., Bernholt, S., Ropohl, M., 2016. Searching for a common ground: a literature review of empirical research on scientific inquiry activities. *Stud. Sci. Educ.* 52 (2), 161–197. <https://doi.org/10.1080/03057267.2016.1206351>.
- Rüschenhöfler, L., Markic, S., 2019. A mixed methods approach to culture-sensitive academic self-concept research. *Educ. Sci.* 9 (3), 240–256. <https://doi.org/10.3390/educsci9030240>.
- Rüschenhöfler, L., Markic, S., 2020. Secondary school students' chemistry self-concepts: gender and culture, and the impact of chemistry self-concept on learning behavior. *Chem. Educ. Res. Pract.* 21 (1), 209–219. <https://doi.org/10.1039/C9RP00120D>.
- Ryan, R.M., Deci, E.L., 2000. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Educ. Psychol.* 25 (1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>.
- Sahin, A., 2013. STEM clubs and science fair competitions: effects on post-secondary matriculation. *J. STEM Educ.:Innov. Res.* 14 (1), 5–11.
- Schunk, D.H., Meece, J.R., Pintrich, P.R., 2014. Motivation in Education: Theory, Research, and Applications. Pearson Higher Ed.
- Sheldrake, R., 2020. Changes in children's science-related career aspirations from age 11 to age 14. *Res. Sci. Educ.* 50 (4), 1435–1464.
- Shin, S., Rachmatullah, A., Roshayanti, F., Ha, M., Lee, J.-K., 2018. Career motivation of secondary students in STEM: a cross-cultural study between Korea and Indonesia. *Int. J. Educ. Vocat. Guid.* 18 (2), 203–231. <https://doi.org/10.1007/s10775-017-9355-0>.
- Sorge, C., 2007. What happens? Relationship of age and gender with science attitudes from elementary to middle school. *Sci. Educ.* 16 (2), 33–37.
- STEM Inc, 2019. About STEM Inc. Retrieved January 2, from: <https://www.science.edu.sg/stem-inc/>.
- Stuckey, M., Hofstein, A., Mamlok-Naaman, R., Eilks, I., 2013. The meaning of "relevance" in science education and its implications for the science curriculum. *Stud. Sci. Educ.* 49 (1), 1–34. <https://doi.org/10.1080/03057267.2013.802463>.
- Thibaut, L., Knipprath, H., Dehaene, W., Depaepe, F., 2018. The influence of teachers' attitudes and school context on instructional practices in integrated STEM education. *Teach. Educ.* 71, 190–205.
- Toker, Y., Ackerman, P.L., 2012. Utilizing occupational complexity levels in vocational interest assessments: assessing interests for STEM areas. *J. Vocat. Behav.* 80 (2), 524–544. <https://doi.org/10.1016/j.jvb.2011.09.001>.
- Tuan, H.L., Chin, C.C., Shieh, S.H., 2005. The development of a questionnaire to measure students' motivation towards science learning. *Int. J. Sci. Educ.* 27 (6), 639–654. <https://doi.org/10.1080/0950069042000323737>.
- Tyler-Wood, T., Knezek, G., Christensen, R., 2010. Instruments for assessing interest in STEM content and careers. *J. Technol. Teach. Educ.* 18 (2), 345–368.
- Unfried, A., Faber, M., Stanhope, D.S., Wiebe, E., 2015. The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *J. Psychoeduc. Assess.* 33 (7), 622–639.
- Van den Hurk, A., Meelissen, M., van Langen, A., 2019. Interventions in education to prevent STEM pipeline leakage. *Int. J. Sci. Educ.* 41 (2), 150–164. <https://doi.org/10.1080/09500693.2018.1540897>.
- Van Tuijl, C., van der Molen, J.H.W., 2016. Study choice and career development in STEM fields: an overview and integration of the research. *Int. J. Technol. Des. Educ.* 26 (2), 159–183. <https://doi.org/10.1007/s10798-015-9308-1>.
- Wang, X., 2013. Why students choose STEM majors: motivation, high school learning, and postsecondary context of support. *Am. Educ. Res. J.* 50 (5), 1081–1121. <https://doi.org/10.3102/0002831213488622>.
- Weiner, B., 1972. Attribution theory, achievement motivation, and the educational process. *Rev. Educ. Res.* 42 (2), 203–215. <https://doi.org/10.2307/1170017>.
- Wigfield, A., Eccles, J.S., 2000. Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25 (1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>.
- Wyss, V.L., Heulskamp, D., Siebert, C.J., 2012. Increasing middle school student interest in STEM careers with videos of scientists. *Int. J. Environ. Sci. Educ.* 7 (4), 501–522.
- Yang, L.-H., 2010. Toward a deeper understanding of student interest or lack of interest in science. *J. Coll. Sci. Teach.* 39 (4), 68–77.
- Yoon, S.Y., Dyehouse, M., Lucietto, A.M., Diefes-Dux, H.A., Capobianco, B.M., 2014. The effects of integrated science, technology, and engineering education on elementary students' knowledge and identity development. *Sch. Sci. Math.* 114 (8), 380–391. <https://doi.org/10.1111/ssm.12090>.
- Young, J., Ortiz, N., Young, J., 2017. STEMulating interest: a meta-analysis of the effects of out-of-school time on student STEM interest. *Int. J. Educ. Math. Sci. Technol.* 5 (1), 62–74.
- Zimmerman, B.J., Schunk, D.H., 2008. Motivation: an essential dimension of self-regulated learning. In: *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Lawrence Erlbaum Associates Publishers, pp. 1–30.
- Zimmerman, H.T., 2012. Participating in science at home: recognition work and learning in biology. *J. Res. Sci. Teach.* 49 (5), 597–630. <https://doi.org/10.1002/tea.21014>.

Integrated STEM curriculum and its assessment

Young Ae Kim^a, Eunjeong Kim^b, and J. Steve Oliver^c, ^aUniversity of Arizona, Tucson, AZ, United States; ^bKorea Foundation for the Advancement of Science & Creativity, Seoul, Korea; and ^cThe University of Georgia, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	409
Background and defining STEM education	410
Conceptualizing “integrated” STEM learning	410
Engineering design as a key integrator in STEM curricula	411
Innovative approaches to STEM curriculum	411
Assessment of student learning in integrated STEM	412
STEM competence	412
Degree of integration and STEM assessment	413
Approaches to assessment of integrated STEM	414
STEM assessment and teacher education	415
A conceptual model for integrated STEM	415
Conclusion	417
References	418

Introduction

Interest and popularity of STEM education that integrates or combines science, technology, engineering, and mathematics has been raised among scholars and educators for teaching and learning in the past few decades. The demand for a STEM-educated workforce has gained increased emphasis from urgent global challenges, including climate change, population, resource management, health and energy issues (Bybee, 2010). All of these complex global issues which can only be addressed by the integration of knowledge and skills of multiple disciplines (Thomas and Watters, 2015; Kelley and Knowles, 2016; Roehrig et al., 2021), require building capacity and international scientific engagement in STEM education (Clark, 2014). In the United States, national policy documents National Academy of Science and National Academy of Engineering, and Institute of Medicine (2007), Next Generation Science Standards (NGSS) (NGSS Lead States, 2013) call for reform to K-12 STEM Education which aim to prepare next generation students for STEM careers and professions (Dori et al., 2018; Roehrig et al., 2021; Shernoff et al., 2017). In particular, since science education included engineering practice and content in the NGSS, how to incorporate integrated multi-disciplinary content into science education has become a key issue (Kim and Oliver, 2018). Meanwhile, some scholars and practitioners have been concerned about over-emphasis on the STEM career pathways and STEM-capable workforce readiness, and application of STEM knowledge with economic perspective of STEM education, and have led the discussions of STEM literacy (Mohr-Schroeder et al., 2020).

To support all diverse students' equity and development of real-world problem-solving competencies and informed decision-making in everyday lives (NGSS Lead States, 2013; ITEEA, 2007; ABET, 2013; NCTM, 2000; NAE and NRC, 2014; NRC, 2012), STEM literacy has also received attention among scholars and educators in STEM education. Mohr-Schroeder et al. (2020) suggested a definition of STEM literacy as “dynamic process and ability to apply, question, collaborate, appreciate, engage, persist, and understand the utility of STEM concepts and skills to provide solutions for STEM-related personal, societal, and global challenges that cannot be solved using a single discipline (p.33)”, highlighting the needs of integration among disciplines to solve problems. There is no one-agreed definition of STEM literacy due to its dynamic and responsive nature in our continuously changing society (Cavalcanti, 2017; Cavalcanti and Mohr-Schroeder, 2019; Zollman, 2012; Mohr-Schroeder et al., 2020). However, one obvious commonality of defining STEM literacy is that STEM literacy should be considered as an integrated literacy (Zollman, 2012). In the broader sense of integrated literacy, we also found many schools are incorporating non-STEM disciplines, exploring the integration of STEM and non-STEM disciplines such as art and humanities and social sciences (e.g., STEAM movement is referring to the case of integration STEM and arts). Some scholars express the view that the STS (Science, Technology and Society) or SSI (Socio-Scientific Issues) movement can be considered as integrated STEM/STEAM education (Tarnoff, 2010). Further, Owens and Sadler (2020) highlighted that SSI can serve as meaningful contexts in STEM curriculum as SSI situates learners to solve real-world problems in an integrated manner (Zeidler, 2014).

Integrated approaches to STEM education have been recognized for their significant advantages in that students can apply multi-disciplinary concepts, critical thinking, collaboration and creativity to real-world problems (Burrows and Slater, 2015; Shernoff et al., 2017). However, research have found that making interrelationships/crosscutting connections among the STEM disciplines are complex and requires deliberate, purposeful and strategic approaches to teach STEM content in implementation, so that students understand how STEM knowledge is conceptually and contextually linked and applied to real-world problems (Kelley and Knowles, 2016; Roehrig et al., 2021). While integrated approaches to STEM education have been emphasized with reform efforts, there are still not many well-developed curriculums, and an integrated way into classroom instructions and assessment

is uncertain and needs a clearer picture of the integrated STEM. Despite the recognized benefits of integrated STEM education, a number of barriers were also identified and discussed in teaching integrated STEM curriculum of many STEM teachers and professionals (e.g., lack of confidence in teaching integrated STEM curricula) (Cunningham and Carlsen, 2014; Cunningham et al., 2005; Honey et al., 2014; Kim and Oliver, 2018; Brophy et al., 2008; Roehrig et al., 2021).

In this article, we begin with a brief review of the ways in which integrated STEM has been conceptualized in the research literature, followed by an overview of research that has focused on the way curriculum and assessment of integrated STEM has developed. Since our conceptions of assessment are not separable from the curriculum and instruction, we will describe assessment in alignment with implementation of curriculum and instruction and present a model that aims to suggest how the four disciplines of STEM can get equal emphasis in classrooms. Concluding this article, we will discuss implications on teacher education and professional development in relation to the assessment of integrated STEM.

Background and defining STEM education

Influences from changes in the broader society have led to changes in educational goals and curriculum reforms (DeBoer, 2014). In the 1983 report, *A Nation at Risk*, the National Commission on Excellence in Education (NCEE) in the US suggested a need for improvement of mathematics and science education with emphasis on higher expectations for students and also spurred educational reforms to integrate or make connections among curriculum (Jackson et al., 2020; DeBoer, 2014). In 1990s, while the discussions on the nature of integration were centered around the integration within science and mathematics, there were some traces of the early ideas of STEM integration of technology, engineering, and mathematics in science (*Science for All Americans*) (Herschman, 1988) and even included social studies into that integration (Anon, 1992; Jackson et al., 2020). In this time of the curriculum development, various curriculum design formats for integration of science and mathematics have been suggested: independent science or mathematics, science focus (math as a subset of science), mathematics focus (science as a subset of mathematics), some overlap between science and math, and a major overlapped program (Czerniak et al., 1999; Jacobs, 1989; Lonning and Defranco, 1997). Davison et al. (1995) suggested five types of integration for developing integrated curriculum: discipline specific (e.g., integration within a discipline), content specific (e.g., integration of mathematical and scientific concepts), process (e.g., integration of scientific processes such as collecting data and mathematics standards such as statistics), methodological (e.g., integration of the learning cycle or models for teaching and learning science and mathematics), and thematic integration (e.g., thematic topics for integration between disciplines).

Later on, until the 2000s, STS education was focused on the inclusion of technology or engineering into science and mathematics, and little attempt was made to create an integrated way of STEM learning in terms of equal attention to two or more disciplines (Bybee, 2010; Sanders, 2009; Wang et al., 2011; Kelley and Knowles, 2016). Bybee (2010) made the point that the term STEM has been used “as a generic label for any event, policy, program, or practice that involves one or several of the STEM disciplines” (p. 30) and the education community needed to clarify the meaning of STEM beyond the general label. Bybee (2010) asserted that STEM should mean increased emphasis on technology and engineering in K-12 education, and an integrated curricular approach to deal with global issues such as energy efficiency or environmental quality.

Sanders (2009) presented the meaning of STEM education focusing on the integration of subjects as approaches of “teaching and learning between/among any two or more of the STEM subject areas, and/or between a STEM subject and one or more other school subjects” (p. 21). In addition to the idea of combining STEM disciplines, Stohlmann et al. (2012) made the point that integrated STEM education is not necessarily involved in all four STEM disciplines and can be utilized as extended across multiple class sessions rather than integration into one class. Moore et al. (2014a,b) also addressed STEM integration as combinations of some or all of the STEM disciplines and emphasized the connections between the subjects in real-world problems. Kelley and Knowles (2016) more organically defined integrated STEM education as “the approach to teaching the STEM content of two or more STEM domains, bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning” (p. 3).

Conceptualizing “integrated” STEM learning

Beyond the definitions, Kelley and Knowles (2016) also proposed a conceptual framework for integrated STEM education to provide an ideal picture with full potential of the integrated STEM focused on secondary education. This conceptual framework includes four main components as an integrated system: *engineering design, scientific inquiry, technology literacy, and mathematical thinking* by engaging students in a *community of practice in the situated STEM learning*. Moore et al. (2014b) suggested six tenets for designing integrated STEM curriculum which emphasized the engineering design process in the STEM curriculum: A motivating and engaging context, an engineering design challenge, the opportunity to learn from failure through redesign, math and/or science content, student-centered pedagogies, and an emphasis on teamwork and communication.

In the book, *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research* (NAE and NRC, 2014) introduced a comprehensive framework for integrated STEM education in K-12 education. The committee identified the learning outcomes of *integrated STEM education including* 21st century competencies, STEM interest, STEM identity, ability to make connections among STEM disciplines for students, as well as STEM-related employment or STEM graduate rates.

Table 1 Types and degrees of STEM integration.

<i>Types and degrees of STEM integration</i>	<i>General description</i>
Monodisciplinary	Content learned in separate/individual disciplinary classrooms
Multidisciplinary	Content learned separately but connected through a common theme or real-world issue Each discipline would be identifiable within the curriculum (Lederman and Niess, 1997) Can take place in a single classroom or through curricular alignment across multiple subject area classes (Drake, 1998; Fogarty, 1991)
Interdisciplinary	Focus on connections between disciplines (Gao et al., 2020), and interdisciplinary content, skills and practices from two or more disciplines connected through a common theme or problem or issue (Roehrig et al., 2021).
Transdisciplinary	Each discipline would be difficult to distinguish from one another (Lederman and Niess, 1997). Beyond the disciplinary constraints (Gao et al., 2020) and not much attention to traces of the individual disciplines (Klein, 2008) Content from two or more disciplines are applied to real-world problems, with the learning focus on the real-world problem or issues (not different subject areas) “Combining ideas and subject matters of different disciplines into a seamless whole” (Lederman and Niess, 1997).

This summary was constructed adapted from Roehrig et al. (2021) and Gao et al. (2020)'s categorizations of the nature of and levels of STEM integration (which initially adapted from Vasquez et al., 2013) as a continuum from disciplinary to transdisciplinary as the integration increases.

Various terms on the integrated STEM education have been used to describe the connections (e.g., multidisciplinary or cross-disciplinary, etc.) within and across STEM disciplines. Although the complexity and ambiguity with diverse views on the integrated STEM, there are four main categorizations have been identified based on the nature of and levels of STEM integration from the literature: (mono-)disciplinary, multidisciplinary, interdisciplinary, and transdisciplinary approaches (Drake, 1998; Lederman and Niess, 1997; Honey et al., 2014; Moore and Smith, 2014; Vasquez et al., 2013; Gao et al., 2020; Roehrig et al., 2021). Table 1 shows a summary of the categorization with the types and degrees of STEM integration.

Due to the diverse views on the nature and development of proficiencies in STEM education, a focus on the nature of integration has been increased and generated needs for further research in STEM education (English, 2016; Marginson et al., 2013).

Engineering design as a key integrator in STEM curricula

Many researchers have shown that, in the development of the integrated STEM curriculum, engineering design serves as an essential vehicle, entry point, catalyst, or integrator for all STEM disciplines in that engineering design provides a platform for equal multidisciplinary connections and applications of core concepts and skills in real-world situations (NAE and NRC, 2014; NRC, 2012; Frykholm and Glasson, 2005; Barnett and Hodson, 2001; Moore et al., 2014a,b; Kelley and Knowles, 2016; Kim et al., 2019b). In this line of research, Moore and colleagues have provided empirical findings that engineering is an essential driver or connector that can integrate STEM disciplines in K-12 education (Glancy et al., 2017; Guzey et al., 2014; Moore and Smith, 2014; Moore et al., 2014a,b; Shernoff et al., 2017). Further, engineering design tasks offer a consistent storyline (Roehrig et al., 2021) bound by STEM practices (e.g., engineering practices) which function as a contextual integrator within curriculum (Kelley and Knowles, 2016; Reynante et al., 2020; Roehrig et al., 2021).

Roehrig et al. (2021) emphasized the *curricular coherence* of STEM units. The authors argued that the coherence of ideas within curriculum is a critical factor for students to develop explanations and conceptual understanding when those are meaningfully linked and richly integrated. In the article, Roehrig et al. (2021) proposed a visualized tool (i.e., CFG, *Conceptual Flow Graphic*) which organized chronologically with main concepts, connections among concepts and lessons, and strength of connections to assess the quality integration of STEM curricula. In an integrated STEM curriculum, it is important to give students the opportunity to understand the connections between STEM disciplines (English, 2016; Honey et al., 2014; Moore et al., 2014a), and the opportunity to explicitly recognize how STEM disciplines are integrated in the contexts (Reynante et al., 2020; Roehrig et al., 2021).

Innovative approaches to STEM curriculum

The development of an integrated STEM curriculum requires complex and cohesive organization of STEM disciplines through tangible forms for implementation such as project-, problem-, or theme-based learning approaches. In turn, the tasks in these innovative curricula lead to assessment of student learning through the tangible products produced (e.g., design product in an engineering design project). Regarding the instructional design for the integrated STEM, problem-based learning and project-based learning were adopted as a form of “engineering design challenges” in most cases (Harwell et al., 2015; Hmelo et al., 2000; Mehalik et al., 2008; Moore et al., 2014a; Sadler et al., 2000).

Problem-based learning and project-based learning have many similarities, and both approaches involve student-centered learning and helping students to develop problem solving and higher order thinking abilities. Problem-based learning promotes; (1) learning organized content, (2) problem-solving abilities, and (3) independent learning skills and self-reflection, as it is a student-centered learning approach in which teachers act as facilitators (Hung et al., 2008). However, there are mixed research results concerning the impact of problem-based learning on students' academic achievement. Several studies have shown that when students engaged in problem-based learning curricula, no significant difference or even underperformance on assessments of content knowledge acquisition were found compared to students who were taught using a traditional teaching method (Gallagher and Stepien, 1996; Mergendoller et al., 2000; Perrenet et al., 2000).

Compared with problem-based learning, project-based learning emphasizes developing teamwork and communication skills through group work. Project-based learning typically takes a longer time, focuses on the application of knowledge (Mills and Treagust, 2003) and concerns the development of cognitive and practical skills than simple acquisition of content (Heywood, 2005). Project-based STEM learning approaches have been considered as student-centered, meaningful, engaging, authentic and collaborative learning experiences, and recognized as a powerful method for; (1) integrating knowledge from multiple disciplines (Bell, 2010; Heywood, 2005; Mills and Treagust, 2003), (2) acquiring a deep understanding of science knowledge by actively constructing knowledge (Krajcik and Blumenfeld, 2006), (3) engaging in self-evaluation of both learning processes and social interactions (Bell, 2010), (4) promoting students' creativity, critical thinking and problem-solving skills through situated inquiry (Stohlmann et al., 2012; Shernoff et al., 2017; Moore et al., 2014a,b; Krajcik and Blumenfeld, 2006) and (5) use of learning technology in the integrated STEM instructions (Krajcik and Blumenfeld, 2006). It is clear that all these learning outcomes that students can develop through project work cannot be assessed through traditional assessment techniques (e.g., written examinations) such as planning or synthesizing abilities.

Assessment of student learning in integrated STEM

Assessment is an essential part of education for both teachers and students in that assessment clarifies learning goals, performance expectations and information on what and how to assess (Biggs, 2003; Douglas et al., 2020). Assessment can also be used to evaluate curriculum or effectiveness of teachers beyond interpretation of individual students' performances. Assessment can be defined as "a formal way to determine students' status with respect to variables of interest" (Popham, 2010, p. 7) as well as a way to determine where students stand in relation to the learning objectives to make instructional decisions to advance student learning (Black and William, 2009; National Research Council, 2001). The latter assessment is called as formative assessment focusing on the students' dynamic learning progress, which can be implemented in a formal and informal manner when teachers gather and analyze assessment data in the classrooms (Bell and Cowie, 2001; Cowie and Bell, 1999; Furtak, 2012; Araceli Ruiz-Primo and Furtak, 2006; Kim et al., 2019a,b). We submit that the assessment of integrated STEM learning can (and should) be also assessed both formally and informally, as well as in summative and formative ways. Pellegrino et al. (2001) highlighted assessment as "a process of reasoning from evidence" of what students know and what students can do, proposing closely linked three main components of assessment: *cognition*, *observation* and *interpretation*. *Cognition* refers to the knowledge, skills and abilities students develop in the target domains. The *observation* refers to the types of assessment tasks to be able to collect evidence of the targeted competence, and the *interpretation* refers to the process of reasoning about students' competence based on the observed data of student performance (Pellegrino et al., 2001; Pellegrino, 2014; Gao et al., 2020; Douglas et al., 2020). These components should be triangulated for proper interpretation of the observed student learning data in relation to the targeted domains in the assessment.

Valid assessment of student learning in STEM is key to determining the success of an innovative curriculum and student performance that are aligned to the spirit of integrated STEM. The *Committee on Integrated STEM Education* made recommendations to guide multiple stakeholders in designing initiatives and developing curriculum and assessments for integrated STEM, as well as further research agenda to support the integrated STEM education (NAE and NRC, 2014). Empirical research support in this area of the integrated STEM learning and assessment needs to be tested and investigated. In particular, it has been a challenge to develop valid and reliable assessment tools in relation to the integrated STEM learning outcomes.

How researchers have conceptualized the integrated STEM, and developed and implemented curriculum are important to better understand the assessment of the integrated STEM learning outcomes in alignment with the learning goals. In the next section, we will review the research that has focused on the assessment of integrated STEM learning and conclude with the discussion of the implication of integrated STEM assessment.

STEM competence

To assess integrated STEM learning, we need to identify what to assess and what we count as evidence of accomplishment of integrated STEM learning outcomes. Bybee (2010) addressed STEM competencies for students to achieve in the 21st century which included identifying STEM issues, explaining issues from STEM perspectives, and using STEM information.

Douglas et al. (2020) presented a conceptual framework of "integrated STEM competence" as "the ability to combine knowledge, skills and abilities (KSAs) from one or more disciplines to solve problems across a wide range of authentic contexts" (p. 234). These authors identified that STEM competence includes *multidimensions* such as *knowledge*, *skills*, *processes* dimensions

to engage in ideas, concepts and practices of STEM to solve authentic problems, reflected from multiple reform documents (NRC, 2000; Krzywacki et al., 2012), however, the multidimensions have not been agreed upon among the reform documents. For example, the *Framework for K-12 Science Education* (NRC, 2012) presented three dimensions of science teaching: *disciplinary core ideas*, *crosscutting concepts* and *science and engineering practices* which serve as learning objectives within science. Although no single approach was acknowledged on how to integrate the three dimensions into standards, curriculum, instruction and assessment, it was clearly stated that “performance expectations must be designed to gather evidence of students’ ability to apply the practices and their understanding of the crosscutting concepts in the contexts of specific applications in multiple disciplinary areas” (NRC, 2012, p. 218). This document introduced the “science and engineering practices” as an integration of multiple disciplinary areas, and also emphasized the importance of the integration of understanding of concepts and practices. Gao et al. (2020) classified the learning outcomes of the integrated STEM programs as four dimensions: *knowledge, skills, practice and affective domain* based on the learning objectives of Bloom’s taxonomy, examining research in assessment of integrated STEM combined with the levels of integration among disciplines.

Gane et al. (2018) also confirmed that these integrated competences are multidimensional and claimed they should be assessed concurrently. This suggests that students can (and will) demonstrate targeted multidimensions in their performance, at the same time, the assessment tasks should be designed to gather evidence of the multidimensions of student competencies related to the integrated STEM learning (Gane et al., 2018; Harris et al., 2019; Douglas et al., 2020).

Although models and frameworks have been proposed for the assessment of integrated STEM and its multidimensions, we acknowledge that concrete examples with practical standards and rubrics with meaningful connections among multiple disciplines are still rare, and need further research in this area. In addition, the STEM competences can be promoted “holistically” from the engagement of STEM practices, exploring the relationship between ideas within and across multiple STEM disciplines (NRC, 2014; Douglas et al., 2020). Although STEM curricula have been increased and implemented in schools, research has reported that very few STEM assessments are conducted in an integrated way, as well as merely aligned with curriculum and instruction, assessing memorized content and practices (Douglas et al., 2017).

Degree of integration and STEM assessment

The benefits of the integrated STEM curriculum in student learning have been emphasized, and the ability to make connections among multiple disciplines is considered as one of the most important learning outcomes in STEM education (NAE and NRC, 2014). However, research in the area of integrated STEM assessment indicated there is little scholarship on the explicit assessment of students’ understanding of connections among multiple disciplines in assessment tasks.

Gao et al. (2020) conducted a comprehensive systematic review of 49 empirical research on assessment of interdisciplinary STEM education focused on secondary and tertiary level for the last two decades. These authors analyzed the empirical studies, creating a matrix with two main dimensions for the analysis: (1) the nature of disciplines (monodisciplinary, interdisciplinary and transdisciplinary) and (2) learning objectives (knowledge, skills, practice and affective domain). The authors found, although explicit and purposeful connections among the STEM disciplines in STEM curricula have been emphasized in the initiatives and standards (Moore et al., 2014a,b; Honey et al., 2014; English, 2016; Reynante et al., 2020; Roehrig et al., 2021), most studies on the assessment of STEM focused on objectives within individual monodisciplines, such that monodisciplinary knowledge was assessed within 38.8% of the total articles, and monodisciplinary affective domains within 32.7% of the total articles. This was true even in cases of the STEM curricula implemented in integrated ways. In terms of dimensions of assessment, it is also reported that knowledge and affective domains were paid attention to dominantly in the assessment of STEM. In particular, the affective domain was the most popular targeted dimension for STEM assessment, including awareness, beliefs, motivation, self-efficacy, interest toward STEM disciplines, and willingness to major in STEM or attitudes toward STEM related careers.

In cases of connections among STEM disciplines, the analysis showed that science and engineering were the dominant disciplines in studies on assessment of integrated STEM (82% of the total articles included both science and engineering). It was reported that interdisciplinary categories (39% of the total articles) were reported least in assessment in relation to the degree of integration. The transdisciplinary category was the second most popular in the degree of integration: transdisciplinary affective domains (36.7% of the total articles) and transdisciplinary skills (22.4% of the total articles). These findings highlight that not only interdisciplinary and transdisciplinary assessment of student understanding, but also skills or practices of integrated STEM disciplines are less targeted which indicates that the assessments are not well aligned with their curriculum designs and key expected outcomes in the integrated STEM education.

Many scholars have discussed the assessment issues that arise from the study of innovative teaching strategies in educational communities (Brophy et al., 2008; Hung et al., 2008; Mills and Treagust, 2003). These researchers argue that anticipated student outcomes from problem-based or project-based learning, such as high-level thinking skills, motivation to learn, teamwork, or communication skills, cannot be assessed properly by using traditional assessment methods. On the other hand, school administrators need clear and unequivocal evidence of student achievement that show the effects of innovative strategies. However, scholars have acknowledged that teachers have difficulty in assessing student outcomes when they are implementing a new instructional method. Ideally assessments need to be aligned with the curriculum which reflect learning objectives. However, scholars and educators have been concerned about the misalignment between the curriculum and assessment in education and claimed the primary

source of performance expectations and learning outcomes should be learning goals and tasks in the curriculum (Yager and Tamir, 1993; McTighe and Wiggins, 2004). In a similar manner, it is essential to clarify assessment domains and evidences of desired competence that show effectiveness and improvement in students' abilities with STEM curricula. However, the concerns on the discrepancy between the goals of developing student competence in integrated STEM and STEM assessment, which have focused on isolated factual knowledge and procedural knowledge of STEM have been raised (Pellegrino, 2014; Douglas et al., 2020).

Approaches to assessment of integrated STEM

The development of appropriate assessment methods and knowing how to assess is an important issue in STEM education. From Gao et al. (2020) analysis, the studies in the assessment of integrated STEM were conducted more in informal settings (57% of the total articles) than formal settings (43% of the total articles). Within secondary and tertiary grade levels, most research was studied in secondary levels (86% of the total articles).

Diverse formats were utilized to assess STEM learning outcomes from researchers, including *questionnaire/test* (Riskowski et al., 2009; Musavi et al., 2018; Nugent et al., 2015), *interviews (individual or focus group)* (Lou et al., 2017), *observation, artifacts (e.g., worksheets, student written work) and reflections* (Dierdorp et al., 2014; Chen and Lin, 2019), *design product analyzing* (Fan and Yu, 2017), *video analysis* (English et al., 2017), etc. Also, the STEM related major application/graduate rates were considered as outcomes of STEM curriculum and programs. The most used format was questionnaire/test (76%). The second most used method was interviews (37% of the total articles) for either individual or focus group interviews. The classroom observation is also a good method to measure, particularly skills or processes of the STEM task. Some researchers tried to use video analysis of student performance for assessing STEM practices, however, this method seems impractical in the classroom assessment use due to time constraints. In addition, students' written work and design products as a result of tasks in a project have been used as evidence of student learning. Students' class artifacts such as worksheets and reflections that were completed during the process of design tasks to assess student learning formatively.

At the elementary level, considerable integrated STEM curricula and assessment have been implemented as project-based learning. For example, a project, Engineering is Elementary (EiE), developed assessment methods to measure students' understanding of engineering and technology (Cunningham et al., 2005). The assessment consisted of open-ended questions, including images of people at work, that asked about what engineers do and what technology is. This study revealed elementary students' lack of understanding about engineering, such as perception of it being male-oriented work. Other researchers also have been focused on assessing students' perceptions about the STEM disciplines or its impact on science knowledge (Apedoe et al., 2008; Brophy et al., 2008; Cunningham et al., 2005; Mehalik et al., 2008; Sadler et al., 2000).

With examples, we would like to look at assessment methods a little more closely to see how integrated STEM can be assessed at the practical classroom level. Riskowski et al. (2009) developed a design project of a practical water purification system in an eighth-grade science course, creating a table top apparatus/portable drinking water device to purify a contaminated water sample. In this study, students (1) worked in small groups and student interaction and participation were considered as a significant component of the project, (2) developed a design artifact which was considered as being reflected their understanding of water quality concepts and (3) shared their findings through a presentation with the entire class and peers were encouraged to discuss advantages and disadvantages of others' designs. These students were assessed through questionnaires multiple times in the process of the project which can be used as good examples of assessment tools using questionnaires in practical classroom use. The researchers conducted a pre/post-test including (1) true/false questions on water quality, (2) open-ended questions on how students view water quality issues in the US and (3) design questions requiring students to describe and explain their designed devices. The design question is as follows: "IDENTIFY what makes the water unsafe to drink and EXPLAIN how your design ensures safe drinking water. Use any means necessary to explain your answer, such as drawings, maps, or sentences" (p. 187). The authors presented the grading rubric and sample response for their analysis which can be used for practical assessment rubric in the classroom as well. The highest score was 4 (0–4 scale) for "Student identifies all 3 contaminants. Uses evidence to justify and explain design." (p. 187). The questionnaires can be effective when they require students' explanation and reasoning. In this study, the authors didn't include the evaluation of the design product as a data source, but the assessment of the design product can be used as an assessment tool as well. In Fan and Yu (2017)'s study, student generated design products (designing a mechanical toy) were assessed. The design products were assessed focused on functionality, mechanical design, and materials used in the design through a five-point scale. Kim and Oliver (2018) developed a water purifier design activity in a preservice teacher education program to guide how to implement and assess student generated water purifiers focused on modeling practice. The study showed the products can be assessed the aspect of functionality, designing with limited materials efficiently as well as the process of the design with modeling on the relationship among degree of purification, water flow rate and the grain size of activated charcoal through a form of drawing and writing of reasoning of the relationship of the factors.

Various approaches can be adopted for assessing student learning outcomes in integrated STEM education. Although we believe that research on assessment of integrated STEM learning can guide the development of effective and practical assessment methods to measure students' learning gains, developing the assessment tools highlighting the benefits of integrated STEM education still remains as challenges. For example, how can we answer our questions such as "what would be the best way to measure students' inter- or multi-disciplinary knowledge/skills/processes in a practical classroom STEM project? What methods can be adopted as formative and summative assessment to give valuable feedback to students in the integrated STEM?" Yet, we need much further investigations in this area.

STEM assessment and teacher education

STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research (NAE and NRC, 2014) suggested “increased STEM content and pedagogical content knowledge (PCK)” and “changes in practice” as goals and outcomes for educators. How teachers understand the integrated STEM curricula and implement STEM instructions effectively is significant for students’ successful STEM learning. In our educational system, as teachers are the most basic and pivotal resource in STEM learning, enhancing teacher’s PCK to teach integrated STEM is critical (Bybee, 2014). In turn, the integrated STEM approach requires new professional development opportunities for both in-service and pre-service teachers, and changes in teacher preparation programs. As a component of PCK, teachers’ knowledge of assessment includes knowledge of student learning as well as knowledge of assessment methods (Tamir, 1988; Park and Oliver, 2008a,b; Sun and Strobel, 2014). A few studies have shown that when teachers encountered assessment difficulties due to unfamiliar subject matter, they employed different assessment methods for two purposes: to know their students’ understanding and learning needs and to encourage and improve their students’ learning (Park and Oliver, 2008a; Sun and Strobel, 2014). The elementary teachers who participated in professional development (PD) programs for practicing engineering teaching had difficulties assessing students’ engineering learning using traditional tools such as rubrics or questions (Sun and Strobel, 2014). After trying traditional/formal assessment methods, the teachers employed different informal assessment methods in their engineering teaching such as questioning, journals, observations, or flow maps. The teachers’ usage of the different assessment methods depended on the assessment purposes. For example, some teachers asked their students to create their own criteria for assessing their engineering products in order to encourage their students to focus on the learning goals. The teachers developed knowledge of assessment through teaching practices within the specific teaching context (Kim, 2019).

How researchers have conceptualized the integrated STEM, and developed and implemented curriculum are important to better understand the assessment of the integrated STEM learning outcomes in alignment with the learning goals. In the next section, we will discuss the tensions that exist in the school level implementation of the integrated STEM and propose a conceptual model of the integrated STEM curriculum which can guide assessment of STEM.

A conceptual model for integrated STEM

If we agree that at its most fundamental core, STEM is about the integration of four domains, then successful STEM pedagogy must be conceptualized as consonant with deep approaches to learning. Although there is a great challenge that arises from the effort to prepare instructional sessions that support learning of all students in a deep manner, the greater challenge comes from preparing the assessments that will validly measure the integrated outcomes.

We want to highlight tensions that exist in school-based implementations of STEM, and propose that a new model of STEM education should be adopted such that the *Integration* of science, technology, engineering and mathematics is the primary descriptor of STEM and that each of these four areas should play an equal role in its characterization. The implication of the goals for STEM with its fully integrated and equilibrated design suggest a series of tensions that are suggested in the day to day life of science teachers and the schools in which they live. In Fig. 1, we have identified these tensions such that a tension is defined in terms of a metaphor of pulling forces that are exerted between two end-points of an argument. We are attempting to create a dialectic so as to contrast these end-points. The new STEM model considers each of the tensions and their impact on the courses that students take.

The traditional conceptualization of STEM has science at the center of all classroom activities and the other areas (technology, mathematics and engineering) feeding into it from outside or being fed by science. However, the new conceptualization of STEM puts science on an equitable footing with the other areas and integrated STEM as the instructional outcomes that arise from the integration of the four S-T-E-M units. In this new conceptualization, each of the four domains (S-T-E-M) can serve the instructional goals as one of three spaces for the construction of objectives or the enactment of teaching. These new spaces are: (1) as subject matter content; (2) as enabler of learning activities; and (3) as conceptual workspace in which the new learning can happen (See Fig. 2). This model allows each of STEM disciplines to serve as content and supports students to construct knowledge rather than just use contextualized examples. It is most easily understood how each of the four domains may be a subject matter content area; the other two functions are more abstract. Mathematics or technology, in very different ways, illustrate the “enabler” function. Frequently, students use technology as a component of the learning process, but in most cases, the technology is simply serving an instrumental function that allows the work to happen. This function is different in this model from the “workspace” in which the domain acts as a space within which the students construct the STEM knowledge. For instance, the difference might be clarified by comparing how mathematics serves to *enable* the learner to calculate values related to understanding Newton’s Laws of motion. No goal for new mathematics learning is necessarily present. In a laboratory inquiry setting, science may serve as a workspace by allowing students to learn about mathematics by creating a subject matter context in which learning takes place. In this situation, learning about both mathematics and science would be anticipated. Computer programming within the domain of technology might also have this function and serve as a workspace rather than as just an enabler. When serving as an enabler, the domain’s activity is that of a metaphorical catalyst left unchanged by the learning activity. When serving as a workspace, the domain is the site of integration of the four S-T-E-M domains into this new integrated knowledge. Throughout this conception of STEM pedagogy, each of the four domains may serve as subject matter content (Kim et al., 2016).

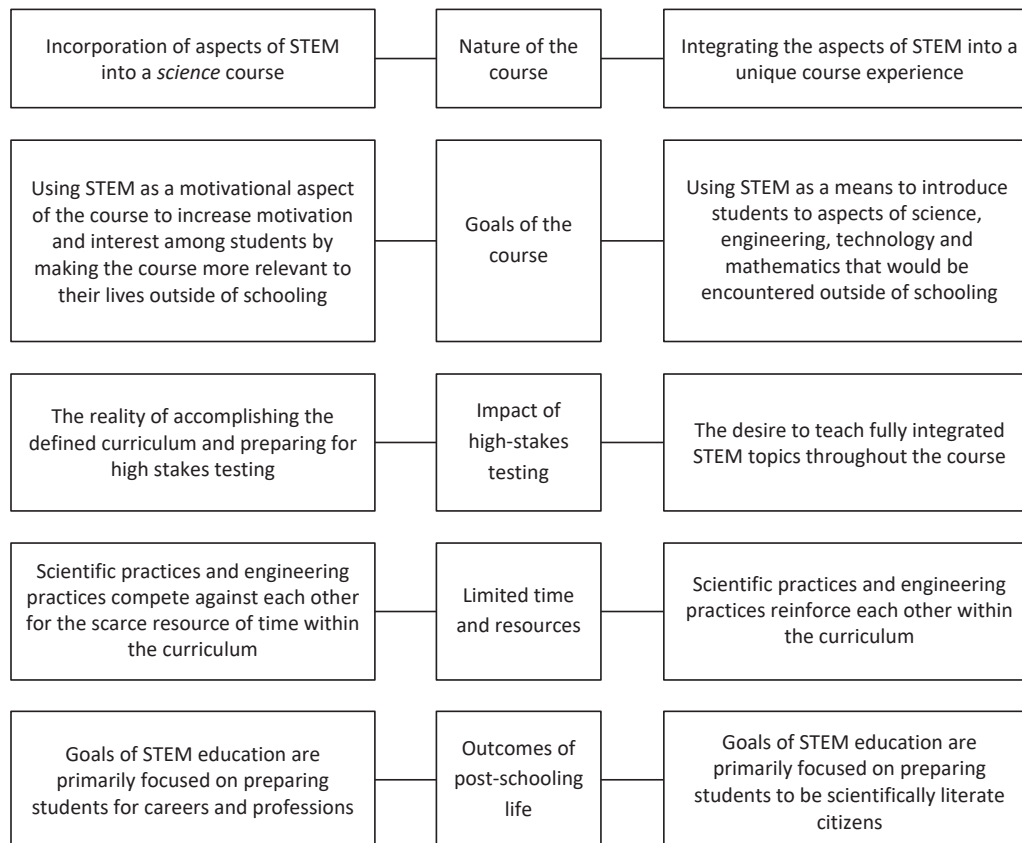


Fig. 1 Tensions that exist along 5 dimensions of school-based implementations of STEM. Adapted from Kim et al. (2016).

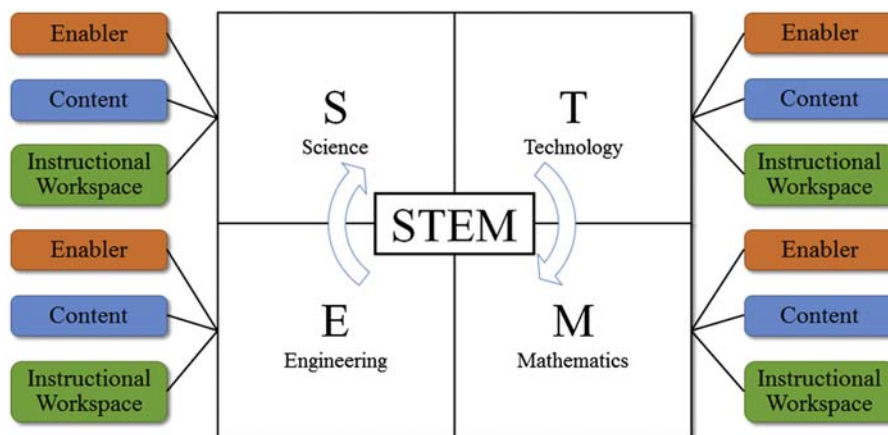


Fig. 2 Schematic representation of a new model for STEM emphasizing a multidimensional, reciprocal, and symmetric relationship between the four domains. Adapted from Kim et al. (2016).

The new STEM model is aligned with conducting assessments based on the SOLO Taxonomy that accounts for the higher-level learning outcomes and student interest and motivation. SOLO is an acronym that stands for *Structure Observation of Learning Outcomes* (Biggs and Collis, 1982). Over several decades, researchers in science education and other areas have been developing and evaluating assessments using the SOLO Taxonomy framework (Hodges and Harvey, 2003; Minogue and Jones, 2009; Padiotis and Mikropoulos, 2010). The framework identifies student outcomes using an ordinal series that begins with the lowest level of accomplishment, i.e., unistructural, to the highest, i.e., extended abstract. The advantage of using this type of framework within which to create assessments and to evaluate open-ended responses can be seen in the article by Minogue and Jones (2009). They wrote of how the challenge to students of cellular biology is magnified if they are unable to integrate

the relationship of the cell structures and the functions. This type of integration is the hallmark of STEM and of the NGSS, and thus must be assessed by tools whose potential to attribute complex and deep responses to student knowledge has demonstrable validity (Kim et al., 2016).

The existence of STEM as a revised or modified type of science course will likely not be adequate to accomplish the goals of STEM as described in either STEM-related scholarship or national reform documents. The existence of tensions such as those identified in Fig. 1 pose very real constraints on the courses currently taught in schools. But it is possible to conceive of a STEM course that would go beyond those constraints and allow the four disciplines to have an equal footing. The model presented here is a first effort to consider how that equal footing and STEM as it is often conceptualized in the scholarship of our field might be pushed into existence.

Conclusion

In this article, we reviewed ways in which integrated STEM has been conceptualized and how integrated STEM curriculum and assessment have developed to this point. Also, we discussed the current status of research on integrated STEM assessment and teachers' professional development in relation to teaching and assessment of integrated STEM. The literature on the assessment of STEM learning outcomes reveals several barriers and challenges. Based on the review, we discuss some implications on teacher education and make suggestions to help our educational community advance integrated STEM education and assessment as follows:

- In relation to the integrated STEM learning goals and outcomes as described in the *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research* (NAE and NRC, 2014), the ability to make connections among STEM disciplines and understand the type of STEM connections are clearly stated as learning outcomes. However, there have been challenges in assessing multidisciplinary learning and the key connections among STEM disciplines to be assessed.
- Most STEM educational programs stated the orientations of their curriculum as integrated approaches to improve students' understanding of connections and multidisciplinary content knowledge and skills, however, rare are the attempts, as recorded in the research literature, to fulfill the assessment goals. As Gao et al. (2020) showed, though the stated focus of assessment is on integrated STEM, assessment of students' content or affective gains are conducted within individual disciplines even though many cases of STEM instructions were implemented in an integrated way. Also, there are few attempts to give students the opportunity to understand the connections between STEM disciplines explicitly and measure those metacognitive learning on the nature of connections among STEM disciplines (Wengrowicz et al., 2018).
- Moore et al. (2014b) clearly emphasized the practices in the process of integrated STEM curriculum (e.g., redesign or modification opportunity in engineering design) as well as social aspects such as teamwork and communication as important curriculum design components. There is little research or conceptual framework on how to assess practices/processes in STEM learning. In the same vein, there have been challenges in assessing teamwork/communication and collaborative problem solving in STEM education with concrete models and practical standards and/or rubrics.
- On a practical level, developing assessment tools and rubrics to guide school teachers for integrated STEM learning has been a critical issue. In particular, the assessment instruments and rubrics on integration are rarely developed due to the attention to the assessment of specific concepts and procedures have been paid in research and practice (NRC, 2014). Many project-based STEM learning approaches may end up with school-based projects with emphasis on student engagement but with limited assessment opportunities to measure important learning goals in STEM education.
- The integrated STEM learning can be measured as formative and summative assessment (NRC, 2014) in both formal and informal ways in relation to the curriculum development, course sequence or designing tasks with tangible learning outcomes. Teachers should be able to utilize the formative assessment in their projects and classroom lessons which enables teachers to characterize current student understanding and skills in relation to the integrated STEM learning objectives. Thus, the formative assessment information can be used to tailor and improve integrated STEM instruction (Heritage, 2013; Wiliam, 2011). However, many pre-service and in-service teachers need support on how to use diverse purposes and forms of assessment such as observations, video analysis or reflections of the integrated STEM learning practices as needed.
- Current teacher education has been considered as insufficient for ensuring teachers' readiness for teaching and assessing the integrated STEM in classrooms (Shernoff et al., 2017). This highlights that teachers need support the opportunities to be engaged in learning experiences to develop teacher knowledge and practices in STEM teaching, in particular, integrated STEM PCK which comprehensively includes orientations to teaching STEM integrated ways, knowledge of STEM curriculum, knowledge of instructional strategies, knowledge of assessment of integrated STEM and knowledge of student understanding when they are engaged in STEM curriculum (e.g., strengths and difficulties in STEM project) (Adapted from Park and Oliver, 2008a,b; Van Driel et al., 2001).
- The development of an integrated STEM curriculum should lead to the application of core concepts from multiple disciplines and STEM practices. Teachers should give students the opportunity to demonstrate their ability to design and perform STEM tasks in problem-solving or projects. In this way, students can develop a deeper understanding of the whole process of the STEM practices, and teachers should base their assessment on the tangible products students create through implementing such innovative tasks.

- Finally, the researchers found that integrated STEM learning and assessment are still fairly new initiatives and the field must continue to strive for a clearer picture of it, thus, the readiness of teachers, administrators and even students still needs support in many ways. The current status of teacher education highlights the challenges with misalignment among curricular development, teachers and administrators' understanding of STEM in implementation and assessment at schools and classroom levels. Still the attention to large-scale tests with multiple-choice items have been paid by administrators dominantly, which cannot be adequate for assessing critical thinking and creativity.

To achieve students' anticipated learning outcomes in the integrated STEM learning, appropriate assessment models and techniques that validly demonstrate the powerful learning gains resulting from this integrated approach are of great importance. It is clear that if there are no evaluation results that show the potential for validly measuring student achievement, the reform cannot move forward. This lack is compounded by the difficulty of assuring teachers of the value of innovative STEM curriculum and instructional strategies. It is essential to provide specific ways to assess student outcomes that are direct results of the implementation of STEM teaching and learning approaches to give valuable feedback to students, as well as support the professional development of teachers to enhance PCK in integrated STEM.

References

- Accreditation Board for Engineering and Technology (ABET), 2013. Criteria for Accrediting Computing Programs. Retrieved from. nse.indiana.edu/institute/documents/accred/spec_2013/ABET_COMPSCI2013.pdf.
- Anon, 1983. A Nation at Risk: The Imperative for Educational Reform: A Report to the Nation and the Secretary of Education, United States Department of Education. The Commission, Washington, D.C. [Supt. of Docs., U.S. G.P.O. distributor].
- Anon, 1992. AAAS 1993: the synergy of the sciences. Boston 11–16 Februar Science 257 (5078), 1965.
- Apedoe, X.S., Reynolds, B., Ellefson, M.R., Schunn, C.D., 2008. Bringing engineering design into high school science classrooms: the heating/cooling unit. *J. Sci. Educ. Technol.* 17 (5), 454–465.
- Araceli Ruiz-Primo, M., Furtak, E.M., 2006. Informal formative assessment and scientific inquiry: exploring teachers' practices and student learning. *Educ. Assess.* 11 (3–4), 237–263.
- Barnett, J., Hodson, D., 2001. Pedagogical context knowledge: toward a fuller understanding of what good science teachers know. *Sci. Educ.* 85 (4), 426–453.
- Bell, B., Cowie, B., 2001. The characteristics of formative assessment in science education. *Sci. Educ.* 85 (5), 536–553.
- Bell, S., 2010. Project-based learning for the 21st century: skills for the future. *Clear. House* 83 (2), 39–43.
- Biggs, J.B., Collis, K.F., 1982. Evaluating the Quality of Learning: The SOLO Taxonomy. Academic Press, New York.
- Biggs, J., 2003. Aligning teaching for constructing learning. *Higher Educ. Acad.* 1 (4).
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21, 5. <https://doi.org/10.1007/s11092-008-9068-5>.
- Brophy, S., Klein, S., Portsmore, M., Rogers, C., 2008. Advancing engineering education in P-12 classrooms. *J. Eng. Educ.* 97 (3), 369–387.
- Burrows, A., Slater, T., 2015. A proposed integrated STEM framework for contemporary teacher preparation. *Teach. Educ. Pract.* 28 (2/3), 318–330.
- Bybee, R.W., 2010. Advancing STEM education: a 2020 vision. *Technol. Eng. Teach.* 70 (1), 30.
- Bybee, R.W., 2014. NGSS and the next generation of science teachers. *J. Sci. Teach. Educ.* 25 (2), 211–221. <https://doi.org/10.1007/s10972-014-9381-4>.
- Cavalcanti, M., Mohr-Schroeder, M.J., 2019. Toward a common vision of STEM literacy. In: Sahin, A., Mohr-Schroeder, M.J. (Eds.), *STEM Education 2.0. Myths and Truths: What Did 10 Years of STEM Education Research in K12 Teach Us?* Koninklijke Brill NV, Leiden, The Netherlands, pp. 3–21. <https://doi.org/10.1163/9789004405400>.
- Cavalcanti, M.A.L., 2017. Assessing STEM Literacy in an Informal Learning Environment (Unpublished doctoral dissertation). University of Kentucky, Lexington, KY. Retrieved from. https://uknowledge.uky.edu/edsc_etds/22/.
- Chen, C., Lin, J., 2019. A practical action research study of the impact of maker-centered STEM-PjBL on a rural middle school in Taiwan. *Int. J. Sci. Math. Educ.* 17, 85–108.
- Clark, J.V. (Ed.), 2014. Closing the Achievement Gap from an International Perspective: Transforming STEM for Effective Education. Springer Science & Business Media.
- Cowie, B., Bell, B., 1999. A model of formative assessment in science education. *Assess Educ. Princ. Pol. Pract.* 6 (1), 101–116.
- Cunningham, C.M., Carlsen, W.S., 2014. Teaching engineering practices. *J. Sci. Teach. Educ.* 25 (2), 197–210.
- Cunningham, C., Lachapelle, C., Lindgren-Streicher, A., 2005. Assessing Elementary School Students' Conceptions of Engineering and Technology. Association for Engineering Education. Engineering Library Division Papers, p.10.227.1.
- Czerniak, C.M., Weber, W.B., Sandmann, A., Ahern, J., 1999. A literature review of science and mathematics integration. *Sch. Sci. Math.* 99 (8), 421–430.
- Davison, D.M., Miller, K.W., Metheny, D.L., 1995. What does integration of science and mathematics really mean? *Sch. Sci. Math.* 95 (5), 226–230.
- DeBoer, G.E., 2014. The history of science curriculum reform in the United States. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, Volume II. Routledge, London, pp. 573–592.
- Dierdorff, A., Bakker, A., Van Maanen, J.A., Eijkelhof, H., 2014. Meaningful statistics in professional practices as a bridge between mathematics and science: an evaluation of a design research project. *Int. J. STEM Educ.* 1 (9), 1–15.
- Dori, Y.J., Mevarech, Z.R., Baker, D.R., 2018. Cognition, metacognition, and culture in STEM education. *Innov. Sci. Educ. Technol.* 24.
- Douglas, K.A., Moore, T.J., Merzdorf, H., Li, T., Johnston, A., 2017. A Content Analysis of How Engineering is Assessed in Published Curricula. Proceedings of the American Society for Engineering Education. ASEE-American Society for Engineering Education, New Orleans, LA.
- Douglas, K., Gane, B.D., Neumann, K., Pellegrino, J.W., 2020. Contemporary Methods of Assessing Integrated Stem Competencies. *Handb. Res. STEM Educ.* 234–254. New York: Routledge.
- Drake, S.M., 1998. Creating Integrated Curriculum: Proven Ways to Increase Student Learning. Corwin Press, Thousand Oaks, CA.
- English, L.D., King, D., Smeed, J., 2017. Advancing integrated STEM learning through engineering design: sixth-grade students' design and construction of earthquake resistant buildings. *J. Educ. Res.* 110 (3), 255–271.
- English, L., 2016. STEM education K-12: perspectives on integration. *Int. J. STEM Educ.* 3 (3), 1–8.
- Fan, S.C., Yu, K.C., 2017. How an integrative STEM curriculum can benefit students in engineering design practices. *Int. J. Technol. Des. Educ.* 27 (1), 107–129.
- Fogarty, R., 1991. Ten ways to integrate curriculum (integrating the curriculum). *Educ. Leader* 49 (2), 61.
- Frykholm, J., Glasson, G., 2005. Connecting science and mathematics instruction: pedagogical context knowledge for teachers. *Sch. Sci. Math.* 105 (3), 127–141.
- Furtak, E.M., 2012. Linking a learning progression for natural selection to teachers' enactment of formative assessment. *J. Res. Sci. Teach.* 49 (9), 1181–1210.
- Gallagher, S.A., Stepien, W.J., 1996. Content acquisition in problem-based learning: depth versus breadth in American studies. *J. Educ. Gift.* 19 (3), 257–275.
- Gane, B.D., Zaidi, S.Z., Pellegrino, J.W., 2018. Measuring what matters: using technology to assess multidimensional learning. *Eur. J. Educ.* 53, 176–187. <https://doi.org/10.1111/ejed.12269>.

- Gao, X., Li, P., Shen, J., Sun, H., 2020. Reviewing assessment of student learning in interdisciplinary STEM education. *Int. J. STEM Educ.* 7 (1), 1–14.
- Glancy, A.W., Moore, T.J., Guzey, S., Smith, K.A., 2017. Students' successes and challenges applying data analysis and measurement skills in a fifth-grade integrated STEM unit. *J. Pre-Coll. Eng. Educ. Res.* 7 (1), 68–75. <https://doi.org/10.7771/2157-9288.1159>.
- Guzey, S.S., Tank, K., Wang, H.H., Roehrig, G., Moore, T., 2014. A high-quality professional development for teachers of grades 3–6 for implementing engineering into classrooms. *Sch. Sci. Math.* 114 (3), 139–149.
- Harris, C.J., Krajcik, J.S., Pellegrino, J.W., DeBarger, A.H., 2019. Designing knowledge-in-use assessments to promote deeper learning. *Educ. Meas.: Issues Pract.* <https://doi.org/10.1111/emip.12253>.
- Harwell, M., Moreno, M., Phillips, A., Guzey, S.S., Moore, T.J., Roehrig, G.H., 2015. A study of STEM assessments in engineering, science, and mathematics for elementary and middle school students. *Sch. Sci. Math.* 115 (2), 66–74.
- Heritage, M., 2013. *Formative Assessment in Practice: A Process of Inquiry and Action*. Harvard Education Press, Cambridge, MA, USA.
- Herschman, A., 1988. 1989 AAAS annual meeting. *Science* 242 (4878), 587–601.
- Heywood, J., 2005. *Engineering Education: Research and Development in Curriculum and Instruction*. John Wiley and Sons.
- Hmelo, C.E., Holton, D.L., Kolodner, J.L., 2000. Designing to learn about complex systems. *J. Learn. Sci.* 9 (3), 247–298.
- Hodges, L.C., Harvey, L.C., 2003. Evaluation of student learning in organic chemistry using the SOLO taxonomy. *J. Chem. Educ.* 80 (7), 785.
- Honey, M., Pearson, G., Schweingruber, H. (Eds.), 2014. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. National Academies Press, Washington, DC.
- Hung, W., Jonassen, D.H., Liu, R., 2008. Problem-based learning. *Handb. Res. Educ. Commun. Technol.* 3 (1), 485–506.
- International Technology and Engineering Educators Association (ITEEA), 2007. Standards for Technological Literacy: Content for the Study of Technology. Retrieved from. <https://www.iteea.org/File.aspx?id=67767&v=b26b7852>.
- Jackson, C., Mohr-Schroeder, M., Moore, T., English, L., 2020. History of integrated STEM curriculum. In: Jackson, C., Tank, K.M., Appelgate, M.H., Jurgenson, K., Delaney, A., Erden, C. (Eds.), *Handbook of Research on STEM Education*. Routledge, New York, pp. 169–183.
- Jacobs, H.H., 1989. *Interdisciplinary Curriculum: Design and Implementation*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Kelley, T.R., Knowles, J.G., 2016. A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3 (1), 1–11.
- Kim, Y., Oliver, J.S., 2018. Supporting pre-service teachers' use of modeling: building a water purifier. *Innov. Sci. Teach. Educ.* 3 (1). Retrieved from. <https://innovations.theaste.org/supporting-preservice-teachers-use-of-modeling-building-a-water-purifier/>.
- Kim, E., Oliver, J.S., Jackson, D.J., 2016. Connecting the Imperatives of STEM, NGSS, Deep Learning and Assessment: A Conceptual Paper. National Association for Research in Science Teaching, Baltimore, Maryland.
- Kim, Y., Cox, J., Southard, K.M., Elfring, L., Blowers, P., Talanquer, V., 2019a. Learning researchers: promoting formative assessment in STEM courses. *J. Coll. Sci. Teach.* 48 (5).
- Kim, E., Oliver, J.S., Kim, Y.A., 2019b. Engineering design and the development of knowledge for teaching among preservice science teachers. *Sch. Sci. Math.* 119 (1), 24–34.
- Kim, E., 2019. A Descriptive Study of a Middle School Faculty's Implementation of an Integrated Stem Curriculum. Doctoral dissertation. University of Georgia.
- Klein, J.T., 2008. Evaluation of interdisciplinary and transdisciplinary research: a literature review. *Am. J. Prev. Med.* 35 (2S), 116–123.
- Krajcik, J.S., Blumenfeld, P., 2006. Project based learning. In: Sawyer, R.K. (Ed.), *Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York.
- Krzywacki, H., Koistinen, L., Lavonen, J., 2012. Assessment in Finnish mathematics education: various ways, various needs. In: *Proceedings of the 12th International Congress on Mathematical Education: Topic Study Group*, vol. 33. Korea, Seoul, pp. 6661–6670. Retrieved from. <http://www.icme12.org/upload/UpFile2/TSG/0732.pdf>.
- Lederman, N.G., Niess, M.L., 1997. Integrated, interdisciplinary, or thematic instruction? Is this a question or is it questionable semantics? *Sch. Sci. Math.* 97 (2), 57–58.
- Lonning, R.A., DeFranco, T.C., 1997. Integration of science and mathematics: a theoretical model. *Sch. Sci. Math.* 97 (4), 212–215.
- Lou, S., Chou, Y., Shih, R., Chung, C., 2017. A study of creativity in CaC₂ steamship-derived stem project-based learning. *Eurasia J. Math. Sci. Technol. Educ.* 13 (6), 2387–2404.
- Marginson, S., Tytler, R., Freeman, B., Roberts, K., 2013. *STEM: Country Comparisons*. Australian Council of Learned Academies, Melbourne, Australia.
- McTighe, J., Wiggins, G., 2004. *Understanding by Design Professional Development Workbook*. Association for Supervision and Curriculum Development (ASCD), Alexandria, VA.
- Mehalik, M.M., Doppelt, Y., Chuun, C.D., 2008. Middle-school science through design-based learning versus scripted inquiry: better overall science concept learning and equity gap reduction. *J. Eng. Educ.* 97 (1), 71–85.
- Mergendoller, J.R., Maxwell, N.L., Bellisimo, Y., 2000. Comparing problem-based learning and traditional instruction in high school economics. *J. Educ. Res.* 93 (6), 374–382.
- Mills, J.E., Treagust, D.F., 2003. Engineering education—is problem-based or project-based learning the answer. *Australas. J. Eng. Educ.* 3 (2), 2–16.
- Minogue, J., Jones, G., 2009. Measuring the impact of haptic feedback using the SOLO taxonomy. *Int. J. Sci. Educ.* 31 (10), 1359–1378.
- Mohr-Schroeder, M.J., Bush, S.B., Maiorca, C., Nickels, M., 2020. Moving toward an equity-based approach for STEM literacy. In: *Handbook of Research on STEM Education*. Routledge, pp. 29–38.
- Moore, T.J., Smith, K.A., 2014. Advancing the state of the art of STEM integration. *J. STEM Educ.* 15 (1), 5.
- Moore, T.J., Glancy, A.W., Tank, K.M., Kersten, J.A., Smith, K.A., 2014a. A framework for quality K-12 engineering education: research and development. *J. Pre-Coll. Eng. Educ. Res.* 4 (1), 1–13.
- Moore, T.J., Stohlmann, M.S., Wang, H.H., Tank, K.M., Glancy, A.W., Roehrig, G.H., 2014b. Implementation and integration of engineering in K-12 STEM education. In: *Engineering in Pre-college Settings: Synthesizing Research, Policy, and Practices*. Purdue University Press, pp. 35–60.
- Musavi, M., Friess, W.A., James, C., Isherwood, J.C., 2018. Changing the face of stem with stormwater research. *Int. J. STEM Educ.* 5 (1), 2.
- National Academy of Engineering (NAE), National Research Council (NRC), 2014. *STEM integration in K-12 education: status, prospects, and an agenda for research*. In: Committee on Integrated STEM Education. National Academy of Engineering and National Research Council of the National Academies. The National Academies Press, Washington, D.C.
- National Academy of Science, National Academy of Engineering, and Institute of Medicine, 2007. *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. The National Academies Press, Washington D.C.
- National Council of Teachers of Mathematics (NCTM), 2000. *Principles and Standards for School Mathematics*. Author, Reston, VA.
- National Research Council, 2000. *Inquiry and the National Science Education Standards: A guide for Teaching and Learning*. National Academies Press.
- National Research Council, 2001. *Classroom Assessment and the National Science Education Standards*. National Academy Press, Washington, DC.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- National Research Council, 2014. *Developing Assessments for the Next Generation Science Standards*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/18409>.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, By States*. The National Academies Press, Washington, DC.
- Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C.R., Nelson, C., 2015. A model of factors contributing to stem learning and career orientation. *Int. J. Sci. Educ.* 37 (7), 1067–1088.
- Owens, C., Sadler, T.D., 2020. Socio-scientific issues as contexts for the development of stem literacy. In: Johnson, C.C., et al. (Eds.), *Handbook of Research on STEM Education*. Routledge, New York, pp. 210–222.
- Padiotis, I., Mikropoulos, T.A., 2010. Using SOLO to evaluate an educational virtual environment in a technology education setting. *J. Educ. Technol. Soc.* 13 (3), 233–245.
- Park, S., Oliver, J.S., 2008a. National Board Certification (NBC) as a catalyst for teachers' learning about teaching: the effects of the NBC process on candidate teachers' PCK development. *J. Res. Sci. Teach.* 45 (7), 812–834.
- Park, S., Oliver, J.S., 2008b. Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Res. Sci. Educ.* 38 (3), 261–284.
- Pellegrino, J.W., Chudowsky, N., Glaser, R., 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academy Press, Washington, DC.

- Pellegrino, J.W., 2014. Assessment as a positive influence on 21st century teaching and learning: a systems approach to progress. *Educ. Psychol.* 20 (2), 65–77. <https://doi.org/10.1016/j.pse.2014.11.002>.
- Perrenet, J.C., Bouhuijs, P.A.J., Smits, J.G.M.M., 2000. The suitability of problem-based learning for engineering education: theory and practice. *Teach. High. Educ.* 5 (3), 345–358.
- Popham, W.J., 2010. *Classroom Assessment: What Teachers Need to Know*. Pearson/Allyn and Bacon, Boston, MA.
- Reynante, B.M., Selbach-Allen, M.E., Pimentel, D.R., 2020. Exploring the promises and perils of integrated STEM, through disciplinary practices and epistemologies. *Sci. Educ.* <https://doi.org/10.1007/s11191-020-00121-x>.
- Riskowski, J.L., Todd, C.D., Wee, B., Dark, M., Harbor, J., 2009. Exploring the effectiveness of an interdisciplinary water resources engineering module in an eighth grade science course. *Int. J. Eng. Educ.* 25 (1), 181–195 (15).
- Roehrig, G.H., Dare, E.A., Ring-Whalen, E., Wieselmann, J.R., 2021. Understanding coherence and integration in integrated STEM curriculum. *Int. J. STEM Educ.* 8 (1), 1–21.
- Sadler, P.M., Coyle, H.P., Schwartz, M., 2000. Engineering competitions in the middle school classroom: key elements in developing effective design challenges. *J. Learn. Sci.* 9 (3), 299–327.
- Sanders, M.E., 2009. STEM, STEM education, STEMmania. *Technol. Teach.* 68 (4), 20–26.
- Sherhoff, D.J., Sinha, S., Bressler, D.M., Ginsburg, L., 2017. Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *Int. J. STEM Educ.* 4 (1), 1–16.
- Stohlmann, M., Moore, T.J., Roehrig, G.H., 2012. Considerations for teaching integrated STEM education. *J. Pre-Coll. Eng. Educ. Res.* 2 (1), 4.
- Sun, Y., Strobel, J., 2014. From knowing-about to knowing-to: development of engineering pedagogical content knowledge by elementary teachers through perceived learning and implementing difficulties. *Am. J. Eng. Educ.* 5 (1), 41.
- Tamir, P., 1988. Subject matter and related pedagogical knowledge in teacher education. *Teach. Teach. Educ.* 4, 99–110.
- Tarnoff, J., 2010. STEM to STEAM—Recognizing the Value of Creative Skill in the Competitiveness Debate. The Huffington Post. Retrieved from. http://www.huffingtonpost.com/john-tarnoff/steam-to-steamrecognizing_b_756519.html.
- Thomas, B., Watters, J.J., 2015. Perspectives on Australian, Indian and Malaysian approaches to STEM education. *Int. J. Educ. Dev.* 45, 42–53.
- Van Driel, J.H., Beijard, D., Verloop, N., 2001. Professional development and reform in science education: the role of teachers' practical knowledge. *J. Res. Sci. Teach.* 38, 137–158.
- Vasquez, J., Sneider, C., Comer, M., 2013. *STEM Lesson Essentials, Grades 3–8: Integrating Science, Technology, Engineering, and Mathematics*. Heinemann, Portsmouth, NH.
- Wang, H.H., Moore, T.J., Roehrig, G.H., Park, M.S., 2011. STEM integration: the impact of professional development on teacher perception and practice. *J. Pre-Coll. Eng. Educ. Res.* 1, 1–13. <https://doi.org/10.5703/1288284314636>.
- Wengrowicz, N., Dori, Y.J., Dori, D., 2018. Metacognition and Meta-assessment in engineering education. In: *Cognition, Metacognition, and Culture in STEM Education. Innovations in Science Education and Technology*. Springer International Publishing, Cham, pp. 191–216.
- William, D., 2011. *Embedded Formative Assessment*. Solution Tree Press, Bloomington, IN, USA.
- Yager, R.E., Tamir, P., 1993. STS approach: reasons, intentions, accomplishments, and outcomes. *Sci. Educ.* 77 (6), 637–658.
- Zeidler, D.L., 2014. Socioscientific issues as a curriculum emphasis: theory, research and practice. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, Volume II. Routledge, New York, NY, pp. 697–726.
- Zollman, A., 2012. Learning for STEM literacy: STEM literacy for learning. *Sch. Sci. Math.* 112 (1), 12–19.

Measurement instruments of STEM affective learning: a systematic review

Jeanna R. Wieselmann^a, Emily A. Dare^b, Gillian H. Roehrig^c, and Elizabeth A. Ring-Whalen^d, ^a Department of Teaching and Learning, Southern Methodist University, Dallas, TX, United States; ^b Department of Teaching and Learning, Florida International University Miami, Miami, FL, United States; ^c Department of Curriculum and Instruction, University of Minnesota, Learning and Environmental Sciences, St. Paul, MN, United States; and ^d Department of Education, St. Catherine University, St. Paul, MN, United States

© 2023 Elsevier Ltd. All rights reserved.

Affective learning	421
Methods	422
Search results and discussion of constructs	424
Attitudes	424
Self-efficacy	430
Interest	430
Motivation	435
Identity	435
Discussion, additional considerations, and limitations	435
References	438

With ongoing demand for expertise in science, technology, engineering, and mathematics (STEM), recent educational reforms in the United States and around the world have called for improved STEM education (e.g., [Australian Curriculum, Assessment, and Reporting Authority, 2016](#); [European Commission, 2015](#); [Hong, 2017](#); [National Research Council \[NRC\], 2011](#); [NRC, 2012](#)). In addition to the need for workers within STEM fields, scientific and technological literacy are increasingly recognized as paramount to informed decision-making for all individuals living in the 21st century ([NRC, 2011](#)). Thus, providing quality STEM learning opportunities to all students, regardless of whether they intend to pursue STEM careers, is a priority in education.

Affect is significant with respect to STEM learning because of the links to learning and behavior (e.g., [Bandura, 1997](#); [Dweck and Leggett, 1988](#); [Koballa and Glynn, 2007](#); [Pintrich et al., 1993](#)). Affective learning includes feelings, emotions, attitudes, interest, self-efficacy, motivation, and identity (e.g., [Ellis and Fouts, 1996](#); [English and English, 1958](#); [Krathwohl et al., 1964](#); [Tytler, 2014](#)). A wide range of instruments have been developed to measure affective learning in STEM; however, with a range of different affective constructs, disciplinary foci, target age groups, and quality of instrumentation, selecting an appropriate instrument for a particular study can be challenging. In addition, because the validity and reliability of an instrument should be reestablished when the instrument is modified or used with a new population ([Koballa and Glynn, 2007](#)), selecting an appropriate instrument is further complicated. This article presents a systematic review of existing instruments that aim to measure affective learning across STEM domains among students in elementary through postsecondary contexts. We aim to provide researchers with a broad range of instruments, including modifications and translations, that will be useful to an international audience of researchers.

Affective learning

Affect has long been considered in relation to learning but has been challenging to define. A number of communities of researchers and practitioners, from psychology to engineering, are interested in the topic of affective learning, resulting in a lack of consistent terminology or theory related to affect in learning ([Meredith et al., 1997](#); [Tytler, 2014](#)). [Krathwohl et al. \(1964\)](#) identified three domains of learning objectives in education: cognitive, affective, and psychomotor. They defined the affective domain as emphasizing “a feeling or tone, an emotion, or a degree of acceptance or rejection” (p. 7). [English and English \(1958\)](#) defined affect as relating to a “feeling, emotion, mood, [or] temperament” (p. 15). [Ellis and Fouts \(1996\)](#) expanded these definitions to include the consideration of attitudes. Because attitudes are feelings that are shaped and endure over time rather than fleeting emotions ([McLeod, 1992](#); [Petty and Cacioppo, 1981](#)), they may hold greater potential to influence motivation, behavior, and learning. Within STEM disciplines, affective learning can be further narrowed and defined. For example, [Tytler \(2014\)](#) identified the following constructs that are central to science education: (a) attitudes toward science, scientists, and school science (b) interest in science and science-related activities or careers (c) self-efficacy in science, and (d) motivation to do and learn science.

Despite a lack of consensus in defining affective learning, it is largely accepted as an important aspect of learning. Attitudes, interests, and beliefs are known to influence students’ thinking, motivation, and behavior (e.g., [Bandura, 1997](#); [Dweck and Leggett, 1988](#); [Koballa and Glynn, 2007](#)). For example, [Pintrich et al. \(1993\)](#) found that attitudinal and motivational constructs moderated learners’ conceptual change in science, thereby influencing science learning outcomes. Personal interest has a greater influence on high school and college students’ career choices than both earning potential and parental influence ([Hall et al., 2011](#)). Interest is a more significant predictor of enrollment in STEM courses than either prior enrollment or achievement in science or mathematics

(Maltese and Tai, 2011) and heavily influences persistence in STEM (Maltese and Cooper, 2017). Guided by Tytler's framing of these constructs and our own prior research (e.g., Dare and Roehrig, 2016, 2020; Smith-Colin and Wieselmann, 2020; Wieselmann et al., 2020a,b), this systematic review of instruments that measure STEM affective learning focuses on attitudes, self-efficacy, interest, and motivation. In addition, we consider identity, or a person's conception of self that is formed through practice and impacted by social setting, power, privilege, and oppression (Calabrese Barton et al., 2013), as central to affective learning.

Methods

We structured our systematic review based on Newman and Gough (2020) recommendations. First, we developed our research question: What instruments developed between 1995 and 2020 measure STEM affective learning among students in K-16 contexts and provide psychometric evidence of reliability and validity? Second, we defined our inclusion criteria:

- Published between 1995 and 2020
- Focus on preK-16 students
- Focus on STEM (one or more disciplines)
- Focus on one or more of the following constructs: attitudes, self-efficacy, interest, motivation, and identity
- Reports on the development of an original survey (not using existing instrument solely for research purposes)
- Provides some evidence of statistical validity and reliability
- Focus is not specific to a particular intervention or course
- Published in a peer-reviewed academic journal
- Article published in English (instrument in another language okay)

With these inclusion criteria in mind, we developed our search strategy and carried out the search in September 2020. We searched three electronic databases: EBSCO Education Source, APA PsychInfo, and Education resources Information Center (ERIC). Several combinations of search terms were tested in preparation for the final search. The search terms that yielded the most relevant results are shown in Table 1. In addition to these search terms, we used the search functions to limit publication dates to 1995-present and restrict search results to publications in peer-reviewed academic journals.

With the resulting publications, we conducted several rounds of review, coding, and analysis of the studies (see Fig. 1). We initially screened the titles and abstracts to determine if they still fit within our inclusion criteria. The full texts were located and downloaded for all publications that appeared to fit our inclusion criteria. Each publication was examined in more depth for adherence to our inclusion criteria and a critical appraisal of their quality (Petticrew and Roberts, 2006). Our critical appraisal focused on two key evidence standards, identified by Gough et al. (2019). First, we considered the suitability of method. In response to critiques of attitudes instruments (e.g., Munby, 1982, 1983), Kind et al. (2007) developed guidelines for the construction of attitudes measures. These guidelines are also useful when considering the development of other affective instruments and include having a strong theoretical grounding with clear descriptions of constructs, demonstrating reliability by confirming the internal consistency of the construct (e.g., using Cronbach's alpha), providing validity evidence using multiple methods, including psychometric analyses, and taking care when combining separate constructs to form a single scale (Kind et al., 2007). With these guidelines framing our review, we paid careful attention to whether the authors of each study reported psychometric properties to provide some evidence of reliability and validity for their instrument.

In order to determine the validity of an instrument, it is necessary to consider the full range of evidence presented by the authors and whether this evidence supports the intended interpretation of a score for a certain purpose (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014). Researchers (e.g., Messick, 1989; Osterlind, 2006) have highlighted the importance of content, construct, and criterion validity. Content validity refers to whether the instrument represents all relevant aspects of the construct it intends to measure. Construct validity refers to whether an instrument represents the construct intended to be measured, including alignment to relevant theory and theoretical expectations. All of the studies that met our inclusion criteria relied on theory or prior instrumentation to some degree in creating the current instrument, demonstrating some level of construct validity. In addition, two particular types of construct validity were considered. First, convergent validity refers to the correlation between scores on the current instrument and scores on similar

Table 1 Search terms used for database search.

Search term	Location
(Stem OR science OR technology OR engineering OR math*)	Abstract
(Instrument OR survey)	Abstract
(Attitude OR interest OR self-efficacy OR motivation OR identity OR affect*)	Abstract
Student	Abstract
(Valid* OR reliab*)	Anywhere

Is used to broaden a search by finding words that start with the same letters. For example, reliab would find reliable, reliability, etc.

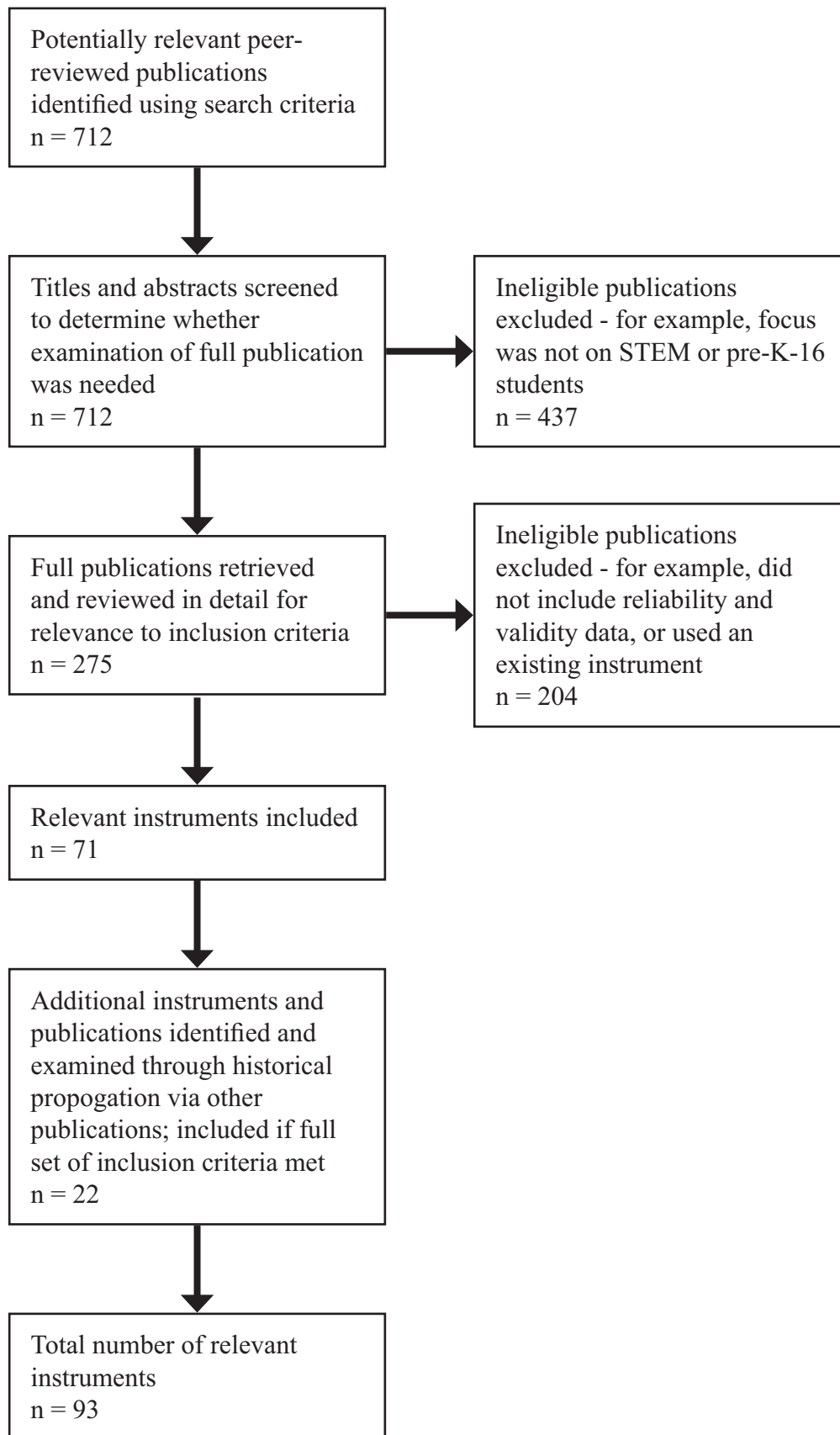


Fig. 1 Search process.

instruments. Second, divergent or discriminant validity involves looking for the absence of a correlation between the current instrument and another, conceptually unrelated, instrument. Criterion validity refers to how closely the current instrument relates to another standard of comparison and includes two types. First, concurrent validity involves comparing scores on the current instrument to scores on another established instrument administered around the same time. Second, predictive validity involves comparing scores on the current instrument to some later outcome or measure to determine whether the instrument has predictive value. Face validity is sometimes used to establish the validity of an instrument, but because it is related to the *appearance* of the instrument as appropriate for measuring the construct rather than psychometric evidence, we did not consider face validity alone as sufficient validity evidence (Munby, 1982). We looked for psychometric evidence, such as exploratory factor analysis (EFA), confirmatory factor analysis (CFA), Rasch analysis, principal component analysis (PCA), and other statistics in determining whether an instrument met our inclusion criteria.

Our second key quality evidence standard, identified by Gough et al. (2019), involved examining the instrument's relevance to our research question. In particular, we ensured that the instrument was STEM-focused, excluding broader instruments that focused on things like students' general motivation to learn. We examined the instrument's relevance across contexts and excluded any instruments that were tightly linked to a particular intervention or course. We also ensured that the instrument was aligned to one or more of the five affective constructs within the scope of this study: attitudes, self-efficacy, interest, motivation, and identity.

We extracted information about each instrument that met all inclusion criteria following this deeper review of the full publication. This information included the instrument name, authors, relevant STEM discipline(s), affective construct(s) measured, grade band(s) of intended use, country(ies) of students included in the development of the instrument, including its validation studies, and the reliability and validity information that was reported.

The information extraction process revealed that some of the publications our search produced used or referenced previously published instruments, so we followed the citations back to the original sources of the instruments, using a process of historical propagation to help minimize the likelihood that our search criteria and process overlooked key instruments. In examining the sources of the original instruments, we maintained all of our inclusion criteria, with two exceptions. First, if an instrument was originally published prior to 1995 but was revalidated or underwent validity tests due to modifications (e.g., translations to another language, short form, etc.) between 1995 and 2020, it was included in our findings. Second, the Pupils' Attitudes Toward Technology (PATT) instrument (Raat and de Vries, 1986) was not originally published in a peer-reviewed journal. However, because of its prevalent use in countries around the world, future validation studies that were published in journals, and the relatively low number of instruments focused on technology, this particular instrument was included in our findings.

With our search process complete, we verified the information that had been entered about each relevant instrument. Each instrument was verified by a member of the research team who had not originally entered the instrument's information. Finally, we synthesized the results of our search to answer our research question.

Search results and discussion of constructs

Our systematic review process resulted in a total of 93 distinct instruments that measure STEM affective learning and met our inclusion criteria. A total of 44 instruments included attitudes as one of the primary affective constructs measured. Twenty of the instruments focused on self-efficacy, 12 on interest, 11 on motivation, and 10 on identity. These instruments are presented in the following sections, organized by the key constructs our review focused on: attitudes, self-efficacy, interest, motivation, and identity. It is important to note that three of the instruments focused on more than one of our focus constructs and are therefore included in multiple tables. Many of the instruments also included additional affective constructs as factors or sub-scales or considered other aspects of affect as constituent of the broader construct being measured; instruments are reported in the table or tables that most closely align with the overall focus of the instrument. Each table is organized first by STEM discipline, then alphabetically by author(s). Instruments that focus on multiple disciplines or STEM are located at the bottom of the tables. Additional publications that utilized or modified the existing instruments are included as separate columns within the row for the original instrument.

Attitudes

Research on student attitudes related to science and mathematics dates back over 50 years. The vast majority of instruments assess student attitudes toward science or mathematics; however, also present in the literature are instruments that measure scientific attitudes (e.g., Moore and Sutman, 1970) and attitudes about learning science or mathematics (e.g., Adams et al., 2006). Our search resulted in 44 instruments focused on attitudes, shown in Table 2.

Broadly, attitudes refer to a favorable or unfavorable feeling toward something and are composed of three elements: cognitive or evaluative beliefs, affective feelings and emotions, and behavioral action tendencies (Cheung, 2009; McGuire, 1985). While some attitudes instruments reside completely within the affective domain, most attitudes surveys address both affective and behavioral components because the primary goal of attitudes research is often to predict student behavior (e.g., Koballa, 1988; Osborne et al., 2003). Researchers with a strong theoretical grounding for their instrument development (e.g., Abd-El-Khalick et al., 2015; Summers and Abd-El-Khalick, 2018; Teo and Noyes, 2008) used theories of reasoned action (Ajzen and Fishbein, 1980) and, more recently, theories of reasoned action and planned behavior (Ajzen and Fishbein, 2005) to argue for the inclusion of career-oriented items.

Table 2 Attitudes instruments.

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Arabic-speaking students' attitudes toward science survey (ASSASS) ^b	Abd-El-Khalick et al. (2015)	Said et al. (2016) , Summers et al. (2019)	Science	Elementary, middle school, high school	Qatar	Expert judgment EFA CFA	Scale reliability = 0.61–0.87	
Colorado learning attitudes about science survey (CLASS) ^b	Adams et al. (2006)	Heredia and Lewis (2012) , Mason and Bertram (2018) , Miranda and Rada (2018)	Science	Postsecondary	US	Expert judgment Factor Analysis	Test-retest reliability = 0.88–0.99	Douglas et al. (2014) —short form Barbera et al. (2008) —CLASS-Chem (chemistry version) Semsar et al. (2011) —CLASS-Bio (biology version) Zwickl et al. (2014) —E-CLASS (experimental physics version) Dorn and Elliott Tew (2015) —computing attitudes survey (CAS) adaptation Code et al. (2016) —mathematics attitudes and perceptions survey (MAPS) adaptation
Astronomy and science student attitudes (ASSA) ^b	Bartlett et al. (2018)		Science (astronomy)	High school	Australia	Factor analysis Construct validity	Cronbach's alpha = 0.810–0.943	
Attitude toward the subject of chemistry inventory (ASCI) ^b	Bauer (2008)		Science (chemistry)	Postsecondary	US	Exploratory factor analysis Convergent validity	Cronbach's alpha = 0.77–0.83 Test-retest correlation = 0.47–0.74	Xu and Lewis (2011) —ASCIv2 short version Rocabado et al. (2019) —ASCIv3 short version with different item order Montes et al. (2018) —adapted ASCIv2 for Latin American context Kahveci (2015) —Turkish adaptation of ASCIv2
Attitudes to school science and science ^b	Bennett and Hogarth (2009)		Science	Middle school, high school	England, Wales	Expert judgment Construct validity	85% Agreement between free and fixed responses	
Attitude toward learning of scientific subjects (ATLoSS) ^b	Caputo (2017)		Science	Middle school	Italy	Expert judgment Content validity Criterion validity	Cronbach's alpha = 0.864	

(Continued)

Table 2 Attitudes instruments.—cont'd

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Attitude toward science test (ATST) ^b	Chin and Lim (2016)		Science	High school	Malaysia	Rasch analysis	Cronbach's alpha = 0.94 Person reliability = 0.93	
Scientific-mind (SCIMIND) measurement model ^b	Chokchai and Pupat (2018)		Science	Middle school	Thailand	Index of item objective congruence Expert judgment CFA	Cronbach's alpha = 0.73, 0.82	
Chemistry attitudes and experiences questionnaire (CAEQ) ^{b,c}	Dalgety et al. (2003)		Science (chemistry)	Postsecondary	New Zealand	Expert judgment Convergent validity Discriminant validity Factor analysis	Cronbach's alpha (for attitudes scale at start and end of semester) = 0.76–0.80	
Science aspirations ^{a,b}	Dewitt et al. (2011)		Science	Middle school	England	PCA	Cronbach's alpha = 0.512–0.869	
Attitudes toward climate change and science instrument (ASCI)	Dijkstra and Goedhart (2012)		Science (climate change)	Middle school, high school	France, Norway, Italy, Netherlands and Spain	Expert judgment Principal component analysis Convergent validity	Cronbach's alpha = 0.71–0.87	
Test of science related attitudes (TOSRA) ^b	Fraser (1978)		Science	Middle school, high school	Australia	Expert judgment Discriminant validity	Cronbach's alpha = 0.64–0.93 Test-retest reliability = 0.69–0.84	Villafañe and Lewis (2016) —short version Spinner and Fraser (2005) —test of mathematics related attitudes (TOMRA) Cheung (2009) —attitude toward chemistry lessons scale (ATCLS) Achor et al. (2019) —students' attitudes questionnaire (SAQ)—Nigerian version
Attitudes toward science inventory (ATSI) ^b	Gogolin and Swartz (1992)		Science	Postsecondary	US (assumed)	Construct validity	Coefficient alpha = 0.73–0.90	Weinburgh and Steele (2000) —modified attitudes toward science inventory (mATSI)
Student interests and motivation in science questionnaire (SIMSQ) ^b	Hassan (2008)		Science	High school, postsecondary	Australia	Concurrent validity EFA	Cronbach's alpha = 0.88	
Attitude toward learning science ^{a,b}	Hassan (2018)		Science	Middle school	Malaysia	CFA	Coefficient alpha = 0.908	

School science attitude survey (SSAS) ^b	Kennedy et al. (2016)		Science	Middle school, high school	Australia	Expert judgment PCA	Coefficient alpha = 0.90–0.98	Toma and Meneses Villagra (2019)—Spanish version (S-SSAS)
Attitude to science measures ^b	Kind et al. (2007)	Barmby et al. (2008)	Science	Middle school	England	Factor analysis	Cronbach's alpha = 0.72–0.94	Kaya (2012)—Turkish version
Students' attitudes toward science ^{a,b}	Nadirova and Burger (2008)		Science	Elementary (upper)	Canada	Expert judgment EFA Discriminant validity	Cronbach's alpha = 0.51–0.90	
Attitude toward science scale ^b	Ozcan and Koca (2020)		Science	Middle school	Turkey	EFA CFA Expert judgment	Cronbach's alpha = 0.72–0.91	
Attitude to science scales ^{a,b}	Pell and Jarvis (2001)		Science	Middle school	England	Factor analysis	Reliability = 0.81	
Changes in attitude about the relevance of science (CARS) questionnaire	Siegel and Ranney (2003)		Science	High school	US (assumed)	Rasch analysis	Cronbach's alpha = 0.91	
Simpson–Troost attitude questionnaire (STAQ) ^b	Simpson and Troost (1982)		Science	Middle school, high school	US	Expert judgment Factor analysis	Internal consistency reliability = 0.90+ (for the scales reported)	Owen et al. (2007)—STAQ short form Liaghatdar et al. (2011)—Persian version
Behaviors, related attitudes, and intentions toward science (BRAINS) survey ^b	Summers and Abd-El-Khalick (2018)		Science	Middle school, high school (grades 5–10)	US	Expert judgment CFA	Scale reliability = 0.70–0.91	
Asian student attitudes toward science class survey (ASATSCS) ^b	Wang and Berlin (2010)		Science	Elementary	Taiwan	Expert judgment PCA	Cronbach's alpha = 0.93	
Three-dimensions of student attitude toward science (TDSAS) ^b	Zhang and Campbell (2011)		Science	Elementary	China	Concurrent validity CFA	Cronbach's alpha = 0.65–0.83	
Math and Me survey ^b	Adelson and McCoach (2010),(2011)	Aktas and Tabak (2018)	Mathematics	Elementary (grades 3–6)	US	Expert judgment EFA CFA Item response theory Discriminant validity PCA	Cronbach's alpha = 0.917–0.939	
Chamberlin affective instrument for mathematical problem solving (CAIMPS) ^{b,c}	Chamberlin and Powers (2013)		Mathematics	Elementary (upper), middle school	US		Coefficient alpha = 0.805–0.918 (0.918 for attitudes factor)	Chamberlin et al. (2017)—CFA validation study

(Continued)

Table 2 Attitudes instruments.—cont'd

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Fennema-Sherman mathematics attitude scales (FSMAS) ^b	Fennema and Sherman (1976), (1977)	Casem (2016), Ren et al. (2016)	Mathematics	High school	US	Expert judgment PCA	Split-half reliability ≤ 0.89	Kahveci (2010) —Turkish version for technology
Student survey on attitudes toward mathematics teaching-learning process ^{a,b}	Mutohir et al. (2018)		Mathematics	Middle school	Indonesia	EFA	Cronbach's alpha = 0.599–0.648	
Survey of attitudes toward statistics (SATS-28) ^b	Schau et al. (1995)	Ayebo et al. (2020)	Mathematics (statistics)	Postsecondary	US	Expert judgment CFA Concurrent validity	Cronbach's alpha = 0.64–0.85	da Silva et al. (2015) —Portuguese version Zimprich (2012) —Swiss version Wang et al. (2018) —validation study with students from US and China Saidi and Siew (2019) —validation study in Malaysia Stanisavljevic et al. (2014) —Serbian version Krishnan and Idris (2015) —matriculation sample
Attitudes toward mathematics inventory (ATMI) ^b	Tapia and Marsh (2004)	Davadas and Lay (2018), Olango (2015)	Mathematics	High school, postsecondary	US	EFA	Cronbach's alpha = 0.97 Test-retest reliability = 0.89	Lim and Chapman (2013) —short ATMI Tabuk (2018) - Turkish version Huang and Lin (2015) —attitude toward calculus inventory (ATCI)
Utley geometry attitude scale (UGAS) ^b	Utley (2007)		Mathematics (geometry)	Postsecondary	US	Expert judgment Concurrent validity EFA PCA	Cronbach's alpha = 0.96	Avcu and Avcu (2015) —Turkish version
Scale for assessing attitudes toward mathematics in secondary education (SATMAS) ^b	Yáñez-Marquina and Villardón-Gallego (2016)		Mathematics	High school	Spain	Expert judgment Discriminant validity Criterion validity	Cronbach's alpha = 0.78–0.93 Composite reliability = 0.78–0.93	
Attitudes toward mathematics scale (ATMS) ^b	Zakariya (2017)		Mathematics (calculus)	High school	Nigeria	EFA PCA	Cronbach's alpha = 0.91	

Engineering interest and attitudes (EIA) ^b	Lachapelle and Brennan (2018)	Engineering	Elementary	US	Expert judgment EFA CFA PCA	Cronbach's alpha discussed but value not reported	
Elementary student coding attitudes survey (ESCAS) ^b	Mason and Rich (2020)	Technology	Elementary (upper)	US	Expert judgment CFA Structural equation modeling	Cronbach's alpha = 0.653–0.933	
Pupils' attitudes toward technology (PATT) ^b	Raat and de Vries (1986)	Technology	Middle school	Netherlands, Global	Expert judgment Factor analysis	Cronbach's alpha discussed but value not reported	Bame and Duggar (1989) —PATT-USA Ardies et al. (2013) —PATT short questionnaire (PATT-SQ) Flemish version Svenningsson et al. (2018) —PATT-SQ Swedish version Becker and Maunsaiyat (2002) —technology attitude and concept scale, Thai version (TACS-Thai)
Computer attitude measure for young students (CAMYS) ^b	Teo and Noyes (2008)	Technology	Middle school	Not specified	CFA Structural equation modeling Criterion validity	Coefficient alpha = 0.85	Asil et al. (2014) —revalidation study in Singapore
STEM learning environment and attitude scales ^{a,b}	Koul et al. (2018)	Engineering, technology	Elementary, middle school	Australia	Factor analysis	Cronbach's alpha = 0.88–0.94	
Biotechnology attitude questionnaire (BAQ)	Erdogan et al. (2009)	Science, technology (biotechnology)	Postsecondary	Turkey	Expert judgment Factor analysis Discriminant validity	Cronbach's alpha = 0.82	
Nanobiotechnology attitude scale ^b	Gul (2017)	Science, technology (nanobiotechnology)	Postsecondary	Turkey	Expert judgment EFA	Cronbach's alpha = 0.93	
Student attitudes toward STEM ^{a,b}	Guzey et al. (2014)	STEM	Elementary (upper), middle school	US	Expert judgment EFA	Cronbach's alpha = 0.91	
Student attitude toward STEM ^b	Mahoney (2010)	STEM	High school	US	Expert judgment PCA Concurrent validity	Cronbach's alpha = 0.88–0.97	
Student attitudes toward science, technology, engineering, and math (S-STEM) ^b	Unfried et al. (2015)	Luo et al. (2019) STEM	Elementary (upper), middle school, high school	US	Expert judgment EFA CFA	Cronbach's alpha = 0.89–0.92 (middle/high school), 0.83–0.87 (elementary school)	

^aInstrument was not clearly named in the publication.

^bAdditional affective measures are included as sub-scales or as part of broader constructs being measured.

^cInstrument included in another of the tables in this article because it addresses additional constructs.

Researchers agree that attitude is a complex and multi-dimensional construct, which is reflected in the development of instruments with multiple scales. Common dimensions of attitudes included enjoyment, confidence, importance or relevance, and usefulness (e.g., Bauer, 2008; Chin and Lim, 2016; Fraser, 1978; Gogolin and Schwartz, 1992; Ozcan and Koca, 2020; Tapia and Marsh, 2004; Teo and Noyes, 2008; Zhang and Campbell, 2011). However, additional dimensions are present in different attitude instruments, such as perception of the teacher, anxiety, and attitudes toward scientific inquiry and other pedagogical approaches (e.g., Adams et al., 2006; Bauer, 2008; Chin and Lim, 2016; Gogolin and Schwartz, 1992; Nadirova and Burger, 2008; Tapia and Marsh, 2004; Zhang and Campbell, 2011).

A common critique of attitudes instruments is that most researchers do not provide well-defined theoretical frameworks and definitions of constructs (e.g., Munby, 1982; Summers and Abd-El-Khalick, 2018). Thus, the construct of attitudes remains “somewhat nebulous and may be defined or framed in different ways depending on the purpose and perspective of the researcher(s) involved” (Summers and Abd-El-Khalick, 2018, p. 179). As such, a seemingly common factor like “usefulness” could focus on the usefulness or value of science to society, or alternatively, on the usefulness of science to individuals for their future careers.

In addition, concerns about theoretical grounding and definition of constructs are more complicated when considering the development of an instrument to measure attitudes toward STEM. None of the included papers that proposed a STEM attitudes instrument defined STEM. Further, proposed instruments often simply measured the same set of items separately for each discipline, and the treatment of technology differed across instruments. For example, it was sometimes assessed separately and other times included as part of a single scale with engineering.

Self-efficacy

Self-efficacy refers to individuals’ beliefs about their personal capability in a domain, and such beliefs are closely linked to what individuals believe to be the likely outcomes of their actions (Bandura, 1977). Our systematic review resulted in 20 instruments that focused on self-efficacy, shown in Table 3. Looking across STEM self-efficacy instruments, researchers relied upon common theoretical frameworks as they operationalized self-efficacy, consistently drawing upon Bandura’s self-efficacy theory. In particular, Bandura (1997) proposed four sources that influence self-efficacy beliefs: enactive mastery experiences, vicarious experiences, verbal persuasion, and affective states. Some researchers (e.g., Byars-Winston et al., 2016; Lent et al., 1991; Usher and Pajares, 2009) used these sources of self-efficacy beliefs to develop survey sub-scales or items.

In addition to Bandura’s self-efficacy theory, instruments often were framed based on Bandura’s (1986) social cognitive theory (e.g., Ibrahim et al., 2016; Lent et al., 1991; Usher and Pajares, 2009) or social cognitive career theory (SCCT; e.g., Byars-Winston et al., 2016; Hiller and Kitsantas, 2016; Nugent et al., 2015). SCCT considers how self-efficacy, outcome expectations, contextual supports and barriers, and personal inputs interact to influence academic and career interests and choices (Lent et al., 1994).

Bandura (1997, 2006) posited that self-efficacy is multi-dimensional and domain-specific. Thus, a general self-efficacy instrument is not sufficient, and researchers have developed instruments for a variety of contexts ranging in degree of specificity. For example, within science, there are instruments that aim to measure self-efficacy within the general discipline of science (e.g., Byars-Winston et al., 2016; Lin et al., 2013; Lin and Tsai, 2013), the field of biology (Baldwin et al., 1999), and specifically in relation to citizen science experiences (Hiller and Kitsantas, 2016). Within technology, instruments tended to be highly task-specific, for example, exploring self-efficacy in human-robot-interactions (Rosenthal-Von Der Pütten and Bock, 2018) and in computer programming (Tsai et al., 2019). The relative lack of instruments as well as the high degree of specificity of those measuring self-efficacy in technology aligns with ongoing debate about defining technology within the context of STEM education (e.g., Ellis et al., 2020). Therefore, researchers who hope to use an existing self-efficacy instrument will need to consider the domain of interest and whether existing options are an adequate fit given their research focus. Notably, only one of the instruments that resulted from this systematic review (Chamberlin and Powers, 2013) included elementary students as part of the intended audience, demonstrating a gap in instrumentation.

Interest

Our search resulted in 12 instruments that measure STEM interest, shown in Table 4. Looking across these instruments, two clear patterns were present in how interest was operationalized. In some cases (e.g., Lamb et al., 2012; Romine et al., 2016), the instrument was used to measure interest in the STEM discipline, nearly always including considerations of external factors that may impact interest (e.g., in-school learning, out-of-school experiences, parents, peers, etc.). These instruments were grounded in psychological underpinnings of interest development, for example, the four stages of interest development originally proposed by Hidi and Renninger (2006). Interest was also operationalized with respect to interest in STEM careers (e.g., Kier et al., 2014; Roller et al., 2020). With this in mind, these instruments tended to draw upon SCCT (Lent et al., 1994) as a theoretical framework and guide for survey factors or sub-scales. Because of this, interest in STEM was nearly always a separate factor or construct in the instrument. Although most of the latter group of instruments focused on interest in STEM careers were targeted toward older students (high school and post-secondary), this was not always the case; for example, Roller et al. (2020) leveraged SCCT to develop an instrument to examine student interest and choices in STEM careers for students in grades 4–12.

Table 3 Self-efficacy instruments.

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Biology self-efficacy scale	Baldwin et al. (1999)		Science (biology)	Postsecondary	US	PCA CFA Concurrent validity	Cronbach's alpha = 0.088–0.089	
Predictors of academic persistence for historically underrepresented racial/ethnic undergraduates in science ^{a,b}	Byars-Winston et al. (2016)		Science	Postsecondary	US	Expert judgment CFA	Coefficient alpha = 0.673–0.910	
Chemistry attitudes and experiences questionnaire (CAEQ) ^{b,c}	Dalgety et al. (2003)		Science (chemistry)	Postsecondary	New Zealand	Expert judgment Convergent validity Discriminant validity Factor analysis	Cronbach's alpha (for self-efficacy scale at start and end of semester) = 0.93–0.96	
Citizen science self-efficacy scale (CSSES)	Hiller and Kitsantas (2016)		Science (citizen science)	Middle school	US	Expert judgment EFA CFA Concurrent validity	Cronbach's alpha = 0.89	
McGill self-efficacy for inquiry engagement (McSELFIE) ^b	Ibrahim et al. (2016)		Science	Postsecondary	US (assumed)	Expert judgment CFA	Cronbach's alpha = 0.96	
Science learning self-efficacy (SLSE) instrument	Lin et al. (2013) , Lin and Tsai (2013)	Lin et al. (2015) , Wang and Tsai (2020)	Science	Middle school	Taiwan, Singapore	Expert judgment Concurrent validity PCA EFA	Cronbach's alpha = 0.85–0.094	Zheng et al. (2018) —self-efficacy of learning science (SELS)—Mandarin
Mathematics self-efficacy scale ^b	Betz and Hackett (1983)	Olango (2016)	Mathematics	Postsecondary	US (assumed)	Item analyses	Coefficient alpha = 0.96	Kranzler and Pajares (1997) —mathematics self-efficacy scale—revised (MSES-R) Britner and Pajares (2006) —sources of science self-efficacy scale Chamberlin et al. (2017) —CFA validation study
Chamberlin affective instrument for mathematical problem solving (CAIMPS) ^{b,c}	Chamberlin and Powers (2013)		Mathematics	Elementary (upper), middle school	US	PCA	Coefficient alpha = 0.805–0.918 (0.908 for self-efficacy factor)	
Sources of mathematics self-efficacy scales	Lent et al. (1991)	Lent et al. (1996)	Mathematics	Postsecondary	US	Convergent validity	Coefficient alpha = 0.56–0.92 Test-retest reliability = 0.72–0.96	Lopez and Lent (1992) —high school version
Mathematics self-efficacy ^{a,b}	Marat (2005)		Mathematics	Middle school, high school	New Zealand	Factor analysis	Cronbach's alpha = 0.77–0.96	

(Continued)

Table 3 Self-efficacy instruments.—cont'd

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Sources of middle school mathematics self-efficacy scale ^b	Usher and Pajares (2009)	Khing and Guan (2019)	Mathematics	Middle school	US	Expert judgment EFA CFA Convergent validity Concurrent validity	Cronbach's alpha = 0.84–0.88	Kontas and Ozcan (2017)—Turkish version Mozahem et al. (2020)—Lebanese version
Calculus self-efficacy inventory (CSEI)	Zakariya et al. (2019)		Mathematics (calculus)	Postsecondary	Norway (assumed)	EFA Concurrent validity	Ordinal coefficient alpha = 0.91	
Engineering design self-efficacy instrument ^{a,b}	Carberry et al. (2010)		Engineering	Postsecondary	US (assumed)	Concurrent validity EFA CFA	Cronbach's alpha = 0.940–0.967	
Longitudinal assessment of engineering self-efficacy (LAESE) ^b	Marra et al. (2005), (2009)		Engineering	Postsecondary	US	Expert judgment Factor analysis	Cronbach's alpha = 0.72–0.87	Yoon and Sorby (2020)—LAESE V3.0
Self-efficacy in human-robot-interaction (SE-HRI) scale—German and english versions	Rosenthal-Von Der Pütten and Bock (2018)		Technology	Postsecondary	Germany	EFA CFA PCA Convergent validity Discriminant validity	Cronbach's alpha = 0.94	
Computer programming self-efficacy scale (CPSES) ^b	Tsai et al. (2019)		Technology	Middle school, high school	Taiwan (assumed)	EFA	Cronbach's alpha = 0.96	
Middle school self-efficacy scale (MSES) ^b	Fouad et al. (1997)		Science, mathematics	Middle school	US	Expert judgment CFA	Coefficient alpha = 0.56–0.84	
Technology and design course self-efficacy scale	Ozden et al. (2019)		Technology, engineering	Middle school	Cyprus	Expert judgment EFA CFA	Cronbach's alpha = 0.93	
STEM career self-efficacy test (SCS-ET) STEM; STEM occupational self-efficacy test (SOS-ET) ^b	Milner et al. (2014)		STEM	Postsecondary	US	EFA CFA	Cronbach's alpha: SCS-ET long = 0.97 SCS-ET short = 0.90 SOS-ET long = 0.97 SOS-ET short = 0.91	
Factors contributing to STEM learning and career orientation ^{a,b,c}	Nugent et al. (2015)		STEM	Middle school	US	CFA Structural equation modeling	Coefficient alpha = 0.463–0.930	

^aInstrument was not clearly named in the publication.^bAdditional affective measures are included as sub-scales or as part of broader constructs being measured.^cInstrument included in another of the tables in this article because it addresses additional constructs.

Table 4 Interest instruments.

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Science interest survey	Lamb et al. (2012)		Science	Middle school, high school	US	Expert judgment PCA CFA Item response theory	Cronbach's alpha = 0.72 Person separation index = 8.75	
Assessment of interest in medicine and science (AIMS)	Romine et al. (2016)		Science	Middle school	US (assumed)	PCA CFA Differential item functioning Rasch analysis PCA	Rasch person-measurement reliability = 0.86–0.90	
Chamberlin affective instrument for mathematical problem solving (CAIMPS) ^{b,c}	Chamberlin and Powers (2013)		Mathematics	Elementary (upper), middle school	US		Coefficient alpha = 0.805–0.918 (0.918 for interest factor)	Chamberlin et al. (2017) —CFA validation study
Value, interest, and expectations scale (VIES) ^b	Appianing and Van Eck (2015)		Technology (computer technology)	Postsecondary	US	Expert judgment EFA	Cronbach's alpha = 0.87–0.90	Appianing and Van Eck (2018) —value-expectancy STEM assessment scale (VESAS)—STEM focus
Student interest in technology and science (SITS) ^b	Romine et al. (2014)		Technology, science	High school	US	Concurrent validity Expert judgment EFA CFA Item response theory	Cronbach's alpha = 0.8	Bodzin et al. (2020) —student interest in science, technology and geospatial technology (STEM-GEO)
Measure of STEM interest for adolescents ^a	Falk et al. (2016) , Staus et al. (2020)		STEM	Middle school	US	Concurrent validity Expert judgment PCA	Cronbach's alpha = 0.69–0.89	
Interest in STEM career questionnaire ^b	Halim et al. (2018)		STEM	High school	Malaysia	EFA CFA	Cronbach's alpha = 0.817–0.933	
STEM career interest survey (STEM-CIS) ^b	Kier et al. (2014)		STEM	Middle school	US	Structural analysis CFA	Cronbach's alpha = 0.77–0.89	Mau et al. (2019) —Chinese version Koyunlu Unlu, Dokme and Unlu (2016) —Turkish version Wieselmann et al. (2018) —Japanese version

(Continued)

Table 4 Interest instruments.—cont'd

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
STEM career interest test (SCIT); STEM occupational interest test (SOIT) ^b	Milner et al. (2014)		STEM	Postsecondary	US	Concurrent validity EFA CFA	Cronbach's alpha: SCIT long = 0.96 SCIT short = 0.91 SOIT long = 0.95 SOIT short = 0.87	
Factors contributing to STEM learning and career orientation ^{a,b,c}	Nugent et al. (2015)		STEM	Middle school	US	CFA Structural equation modeling	Coefficient alpha = 0.463–0.930	
Educational and career interest in STEM measure	Oh et al. (2013)		STEM	High school	US	Concurrent validity Expert judgment EFA CFA	Cronbach's alpha = 0.88	
Student interest and choice in STEM (SIC-STEM) survey 2.0 ^b	Roller et al. (2020)		STEM	Elementary (upper), middle school, high school	US	EFA Structural equation modeling	Cronbach's alpha = 0.84–0.89	

^aInstrument was not clearly named in the publication.

^bAdditional affective measures are included as sub-scales or as part of broader constructs being measured.

^cInstrument included in another of the tables in this article because it addresses additional constructs.

Motivation

Our systematic review resulted in 11 instruments that focused on STEM motivation, shown in [Table 5](#). Across these instruments, several theoretical frameworks were commonly used to operationalize motivation. Self-determination theory ([Deci and Ryan, 1985](#)), which distinguishes between intrinsic and extrinsic motivation and focuses on individuals' desire to fulfill fundamental needs for competence, autonomy, and relatedness, frequently served as a theoretical framework for STEM motivation instruments (e.g., [Glynn et al., 2009](#); [Hammoudi, 2020](#); [Lim and Chapman, 2015](#); [Skinner et al., 2017](#)). Additional instruments (e.g., [Glynn et al., 2009](#); [Liou and Kuo, 2014](#)) utilized social cognitive theory ([Bandura, 1986, 2001](#)), while others (e.g., [Glover and Bodzin, 2019](#); [McCormick et al., 2015](#)) used expectancy-value theory ([Eccles and Wigfield, 1995, 2002](#)). Given the connections between social cognitive theory, expectancy-value theory, and self-efficacy, it is perhaps unsurprising that self-efficacy was frequently included in the theoretical framing or as a subscale or factor in various instruments (e.g., [Glynn et al., 2009](#); [Hammoudi, 2020](#); [Liou and Kuo, 2014](#); [Tuan et al., 2005](#)). Thus, researchers who are interested in measuring self-efficacy may also find useful items or instruments in [Table 5](#). The majority of the STEM motivation instruments were focused on science and mathematics, with only one addressing technology and one addressing engineering. In addition, only two of the instruments ([Orosco, 2016](#); [Sha et al., 2015](#)), were designed for use by elementary students and two for use by middle school students ([Tuan et al., 2005](#); [Sha et al., 2015](#)).

Identity

The construct of STEM identity has notably only been studied quantitatively for the last decade or so. Our systematic review produced 10 instruments focused on measuring STEM identity, shown in [Table 6](#). These instruments primarily drew upon [Carlone and Johnson \(2007\)](#) qualitative examination of science identity, which provided a strong theoretical framework for researchers. Expectancy-value theory (e.g., [Eccles et al., 1998](#)), which emphasizes the importance of both expectations and values in shaping learning behaviors, and SCCT (e.g., [Lent et al., 1994](#)) were also used. However, the science identity framework was only adopted for science and engineering spaces; notably, mathematics and technology identity are missing from this set of instruments. Although [Dou et al. \(2019\)](#) and [Paul et al. \(2020\)](#) focus on "STEM identity," their instruments do not break down the acronym into sub-scales by discipline, but rather collapse science, technology, engineering, and mathematics into a single domain. Because the nature of identity includes interest as a main theoretical construct, interest in STEM fields (rather than in STEM careers), was included as a survey sub-scale or construct. It is interesting to note that with the exception of [Paul et al. \(2020\)](#) and [Taconis et al. \(2010\)](#), classical test theories (e.g., EFA, CFA) were used to develop the instruments, potentially due to the still exploratory nature surrounding understanding the complexity of STEM identities.

Discussion, additional considerations, and limitations

Our search results indicate the availability of a wide array of instruments that measure STEM affective learning. Of the 93 distinct instruments that measure STEM affective learning, 41 focus on science, 18 on mathematics, eight on engineering, and seven on technology. Additionally, another 19 instruments included at least two STEM disciplines, with 13 of them reporting a focus on STEM. The 13 instruments that explicitly focused on "STEM" affective learning rather than one or more of the STEM disciplines were found in relatively recent publications, the oldest of which was published in 2010, and they were also more prevalent in measuring interest than other affective constructs. The development of new, STEM-focused instruments coincides with the increasing focus on STEM education over the past decade.

Our results also explored the availability of STEM affective survey instruments across grade bands. Seventeen of the instruments were intended for use with elementary students, 37 with middle school students, 30 with high school students, and 31 with post-secondary participants. Some instruments were intended for use across multiple grade bands. Perhaps unsurprisingly given potential limitations in relation to students' reading ability, the fewest instruments existed at the lower elementary level. While some of the existing instruments were designed for use across multiple grade bands, it is important to consider the appropriateness of the instrument for the particular ages of interest and context. Those who use any existing instrument should carefully examine the items in the instrument itself rather than relying solely on the authors' recommendations for appropriate grade levels.

The instruments also measured a number of different affective constructs, with student attitudes being the most common construct, likely due to its longer history in the field of educational research. However, many of the instruments measured multiple constructs, reflecting both practical and theoretical connections among affective constructs. Three of these instruments ([Chamberlin and Powers, 2013](#); [Dalgety et al., 2003](#); [Nugent et al., 2015](#)) included two or more of the key constructs we explored as primary constructs of focus and were therefore included in multiple tables. However, there is evidence of theoretical overlap between constructs even among instruments that were placed in only one table. For example, an instrument developed based on SCCT, such as the STEM Career Interest Survey ([Kier et al., 2014](#)), may present student interest as the primary affective construct while also including items designed specifically to relate to self-efficacy, outcome expectations, and other affective constructs. The Student Interests and Motivation in Science Questionnaire (SIMSQ; [Hassan, 2008](#)) considered interest and motivation within its scales, but its overall theoretical grounding and use of existing instruments for item generation aligned more closely with the attitudes construct, so it was included in [Table 2](#).

It is also important to note that the approaches to statistical validity and reliability varied across instruments. Our inclusion criteria required some psychometric evidence of validity and reliability, but the reported studies provided different forms of

Table 5 Motivation instruments.

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Value-achievement-cost (VAC) survey ^b	Glover and Bodzin (2019)		Science (health sciences)	High school	US	Convergent validity	Cronbach's alpha = 0.90	
Science motivation questionnaire (SMQ) ^b	Glynn et al. (2009)	You et al. (2018)	Science	Postsecondary	US	EFA Criterion validity	Cronbach's alpha = 0.91	Glynn et al. (2011)—SMQ II Salta and Koulougliotis (2015)—Greek version of SMQ II Rachmatullah et al. (2018)—Korean and Indonesian versions of SMQII Kwon (2016)—Korean version of SMQ II focused on technology Dong et al. (2020)—Chinese version (SMQ II-C)
Science choice preference scale ^b	Sha et al. (2015)		Science	Elementary (upper), middle school	US	Factor analysis Item response theory Concurrent validity Differential item functioning	Cronbach's alpha = 0.78–0.91	
Self-determination, purpose, identity, and engagement in science (SPIRES) ^b	Skinner et al. (2017)		Science	Postsecondary	US	CFA Predictive validity	Cronbach's alpha = 0.67–0.94 McDonald's omega = 0.74–0.95 Cross-time stability = 0.49–0.79	
Students' motivation toward science learning (SMTSL) scale ^b	Tuan et al. (2005)	Köksal (2013)	Science	Middle school, high school	Taiwan	Factor analysis Expert judgment Criterion validity Discriminative validity	Cronbach's alpha = 0.89	Yilmaz and Çavaş Huyugüzel (2007)—Turkish version Dermitzaki et al. (2013)—Greek version
Mathematics motivation and self-concept survey ^b	Hammoudi (2020)		Mathematics	Postsecondary	UAE	Factor analysis	Cronbach's alpha = 0.79–0.94	
Academic motivation toward mathematics scale (AMTMS) ^b	Lim and Chapman (2015)		Mathematics	High school	Singapore	EFA CFA	Coefficient alpha = 0.83–0.91 Test-retest reliability = 0.71–0.75	
Mathematics value inventory (MVI) ^b	Luttrell et al. (2010)		Mathematics	Postsecondary	US	Expert judgment PCA Discriminant validity	Cronbach's alpha = 0.87–0.95 Test-retest reliability (correlation) $r = 0.96$	Andrews et al. (2017)—math-biology values instrument (MBVI) Rodriguez-Ayan and Rico (2015)—Spanish version
Beliefs, engagement, and attitude math motivation scale (BEAMMS) ^b	Orosco (2016)		Mathematics	Elementary (grades 2–3)	US	Factor analysis Differential item functioning	Coefficient alpha = 0.85	
Student motivation to engage in sustainable engineering ^{a,b}	McCormick et al. (2015)		Engineering	Postsecondary	US (assumed)	Expert judgment PCA	Cronbach's alpha = 0.785–0.957	
Motivation and self-regulation toward technology learning (MSRTL) ^b	Liou and Kuo (2014)		Technology	High school	Taiwan	Expert judgment EFA CFA	Cronbach's alpha = 0.83–0.91	

^aInstrument was not clearly named in the publication.^bAdditional affective measures are included as sub-scales or as part of broader constructs being measured.

Table 6 Identity instruments.

<i>Instrument name</i>	<i>Citation of instrument development</i>	<i>Citation(s) that used the instrument</i>	<i>STEM discipline(s)</i>	<i>Primary grade band(s) of intended use</i>	<i>Country(ies) of students included in development</i>	<i>Validity evidence</i>	<i>Reliability evidence</i>	<i>Modifications and translations of original instrument</i>
Science identities, expectations of success in science, values of science, and environmental attitudes (SIEVEA) ^b	Aghekyan (2019)		Science	High school	US	Concurrent validity EFA CFA	Composite reliability = 0.727–0.882	
Persistence research in science & engineering (PRISE) survey ^b	Hazari et al. (2010)	Godwin et al. (2016)	Science (physics)	Postsecondary	US	Concurrent validity Expert judgment Factor analysis	Cronbach's alpha (for physics identity construct) = 0.83 Test-retest reliability coefficient = 0.7	
Science orientedness of students' identities ^{a,b}	Taconis et al. (2010)		Science	High school	Netherlands	Expert judgment Predictive validity	Cronbach's alpha = 0.88	
Science identity scale ^{a,b}	Vincent-Ruz and Schunn (2018)		Science	Middle school	US	EFA CFA Discriminant validity Predictive validity	Cronbach's alpha = 0.84	
Engineering identity development scale (EIDS) ^b	Capobianco et al. (2012)		Engineering	Elementary	US	Expert judgment EFA CFA	Cronbach's alpha = 0.58–0.760	
Sustainability and gender in engineering (SaGE) survey ^b	Godwin et al. (2013), (2016)		Engineering	Postsecondary	US	Expert judgment EFA CFA Structural equation modeling Discriminant validity	Test-retest –Pearson's correlation = 0.732	
Identity and persistence in engineering survey ^b	Patrick et al. (2018)		Engineering	Postsecondary	US	Concurrent validity EFA	Cronbach's alpha = 0.726–0.864	
Role-identity survey in engineering (RIS-E) ^b	Paul et al. (2020)		Engineering	Elementary	US	Concurrent validity Expert judgment EFA PCA CFA Item response theory	Cronbach's alpha = 0.818–0.921	
STEM identity ^{a,b}	Dou et al. (2019)		STEM	Postsecondary	US	EFA CFA	Cronbach's alpha = 0.97	
Role-identity survey in STEM (RIS-STEM) ^b	Paul et al. (2020)		STEM	Elementary	US	Concurrent validity Expert judgment CFA Item response theory	Cronbach's alpha = 0.787–0.884	

^aInstrument was not clearly named in the publication.^bAdditional affective measures are included as sub-scales or as part of broader constructs being measured.

evidence to various degrees of detail. Reliability was most frequently reported using Cronbach's alpha as a measure of internal consistency. Some publications reported test-retest reliability, with the same participants taking the survey on multiple occasions to demonstrate consistency in their responses. Validity was most often reported based on results from factor analysis (both exploratory and confirmatory). Some publications reported a Rasch analysis or convergent validity results in which participants responded to the new survey instrument as well as another, previously validated, instrument that measured the same construct to determine whether their response patterns were consistent across the different measures. Panels of experts were often used to establish validity, but particular researchers' approaches to using expert judgments varied and were reported in relation to face, content, and construct validity. In addition, the number of participants used for validation statistics varied widely across studies.

Rather than establishing clear boundaries for the amount of psychometric evidence required for the inclusion of an instrument, we included an instrument as long as it provided some psychometric evidence of both validity and reliability and met our other inclusion criteria. Thus, we highly encourage those who are considering using any of these instruments to examine the full publication and determine whether it meets their threshold of reliability and validity evidence for the intended purposes. For example, the Engineering Identity Development Scale (EIDS; Capobianco et al., 2012) has been criticized for not actually measuring identity, but rather focused on students' knowledge of engineering careers (e.g., Paul et al., 2020), so researchers aiming to measure engineering identity would be well-served to closely examine the original manuscript to determine if the construct of identity is operationalized in a manner that fits their needs. For further guidance on what should be considered when determining the validity of an instrument, we recommend consulting the *Standards for Education and Psychological Testing* chapter focused on validity (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014).

Although we used rigorous search procedures and carefully selected our search terms, it is important to acknowledge the limitations of this work. Despite our search of multiple databases and our consideration of references cited in the full publications we reviewed, it is possible that relevant peer-reviewed journal publications are missing from this review. However, the number of instruments presented provides a range of options for researchers in STEM education and can serve as a useful starting point in selecting a survey instrument.

We also recognize that some high-quality instruments were likely excluded from our results due to our requirement that they be published in peer-reviewed academic journals. Instruments that were published in dissertations, books, conference presentations, or other reports were not included, and these instruments may serve as additional resources for researchers or others considering STEM affective learning. For example, the Relevance of Science Education (ROSE) study (Schreiner and Sjøberg, 2004) and its associated questionnaire address student experiences and interests in science but was not included in this review. This survey is available in multiple languages and has been used in many countries around the world in international comparisons, so it may be of use to researchers. In addition, the STEM Learning and Research (STELAR) Center contains a repository of instruments that have been submitted by the National Science Foundation Innovative Technology Experiences for Students and Teachers (ITEST) projects for their research or evaluation (see <http://stelar.edc.org/resources/instruments>). This may be a useful tool for finding additional instruments aligned with particular project or research goals, but many of these instruments were not published in peer-reviewed journals.

In addition, we acknowledge that 54 of the instruments included in our results focused on students from the United States in their development. While we made every effort to include instruments developed in a diverse range of international contexts and languages, we were limited by our own language proficiencies to those that published development information, including reliability and validity details, in English. Our selection of instruments does include a number of surveys available in languages other than English, but there are certainly additional instruments that were overlooked due to this inclusion criterion. We have included citations for instrument translations that we discovered in our systematic review, but we have only reported validity and reliability evidence for the original instrument. Those who choose to use a translation should carefully examine the cited article to determine whether the instrument would be valid and reliable given the intended use.

Finally, while this review focused on quantitative measures of STEM affective learning, the limitations of a survey approach must be considered. Surveys can be an efficient means of collecting consistent data from a large number of participants, but they are inherently limited by the items included. We believe mixed methods approaches that consider complementary qualitative data (e.g., interviews, observations, etc.) in addition to quantitative survey data result in stronger research designs with a more thorough understanding of the phenomenon of STEM affective learning.

With strong connections to thinking, motivation, and behavior (e.g., Bandura, 1997; Dweck and Leggett, 1988; Koballa and Glynn, 2007; Pintrich et al., 1993), affective learning in STEM contexts is undoubtedly of great importance. Students' attitudes, self-efficacy, interest, motivation, and identity must be fostered through positive STEM learning experiences in order to work toward the goals of STEM literacy for all students as well as knowledgeable STEM workers. A vast number of survey instruments have been developed to measure STEM affective learning, and in combination with other approaches to research and evaluation, they can be used to guide the development of activities, courses, and programs that are supportive of all students' STEM affect.

References

- Abd-El-Khalick, F., Summers, R., Said, Z., Wang, S., Culbertson, M., 2015. Development and large-scale validation of an instrument to assess Arabic-speaking students' attitudes toward science. *Int. J. Sci. Educ.* 34, 2637–2663.
- Achor, E.E., Danjuma, I.M., Orji, A.B.C., 2019. Classroom interaction practices and students' learning outcomes in physics: implication for teaching-skill development of physics teachers. *J. Educ. E-Learn. Res.* 6, 96–106.

- Adams, W.K., Perkins, K.K., Podolefsky, N.S., et al., 2006. New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes About Science Survey. *Phys. Rev. ST Phys. Educ. Res.* 2, 010101.
- Adelson, J.L., McCoach, D.B., 2010. Measuring the mathematical attitudes of elementary students: the effects of a 4-point or 5-point Likert-type scale. *Educ. Psychol. Meas.* 70, 796–807.
- Adelson, J.L., McCoach, D.B., 2011. Development and psychometric properties of the Math and Me Survey: measuring third through sixth graders' attitudes toward mathematics. *Meas. Eval. Counsel. Dev.* 44, 225–247.
- Aghekyan, R., 2019. Measuring high school students' science identities, expectations of success in science, values of science and environmental attitudes: development and validation of the SIEVEA survey. *Sci. Educ. Int.* 30, 342–353.
- Ajzen, I., Fishbein, M., 1980. *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall, Englewood-Cliffs, NJ.
- Ajzen, I., Fishbein, M., 2005. The influence of attitudes on behavior. In: Albarracín, D., Johnson, B.T., Zanna, M.P. (Eds.), *The Handbook of Attitudes*. Erlbaum, Mahwah, NJ, pp. 173–221.
- Aktas, M.C., Tabak, S., 2018. Turkish adaptation of Math and Me Survey: a validity and reliability study. *Eur. J. Educ. Res.* 7, 707–714.
- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (Eds.), 2014. *Standards for Education and Psychological Testing*. American Educational Research Association, Washington, DC.
- Andrews, S.E., Runyon, C., Aikens, M.L., 2017. The math-biology values instrument: development of a tool to measure life science majors' task values of using math in the context of biology. *CBE Life Sci. Educ.* 16, ar45.
- Appianing, J., Van Eck, R.N., 2015. Gender differences in college students' perceptions of technology-related jobs in computer science. *Int. J. Gen. Sci. Technol.* 7, 29–56.
- Appianing, J., Van Eck, R.N., 2018. Development and validation of the Value-Expectancy STEM Assessment Scale for students in higher education. *Int. J. STEM Educ.* 5.
- Ardies, J., De Maeyer, S., Gijbels, D., 2013. Reconstructing the pupils attitude towards technology-survey. *Design Technol. Educ.* 18, 8–19.
- Asil, M., Teo, T., Noyes, J., 2014. Validation and measurements invariance of the computer attitude measure for young students (CAMYS). *J. Educ. Comput. Res.* 51, 49–69.
- Australian Curriculum, Assessment, and Reporting Authority, 2016. ACARA STEM Connections Project Report. Retrieved from: <https://www.australiancurriculum.edu.au/media/3220/stem-connections-report.pdf>.
- Avcu, R., Avcu, S., 2015. Turkish adaptation of Utley geometry attitude scale: a validity and reliability study. *Eur. J. Educ. Res.* 58, 1–24.
- Aybeo, A., Bright, J., Ballam, C., 2020. Examining the factor structure of the survey of attitudes towards statistics among undergraduate health science students. *Int. Electron. J. Math. Educ.* 15.
- Baldwin, J.A., Erbert-May, D., Burns, D.J., 1999. The development of a college biology self-efficacy instrument for nonmajors. *Sci. Educ.* 83, 397–408.
- Bame, E.A., Dugger Jr., W.E., 1989. Pupils' attitude toward technology: PATT-USA. In: PATT 4 Conference Proceedings. PATT-Foundation, Eindhoven, Netherlands, pp. 309–319.
- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215.
- Bandura, A., 1986. *Social Foundations of Thought and Action*. Prentice-Hall, Englewood Cliffs, NJ.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman, New York, NY.
- Bandura, A., 2001. Social cognitive theory: an agentic perspective. *Annu. Rev. Psychol.* 52, 1–26.
- Bandura, A., 2006. Guide for constructing self-efficacy scales. In: Pajares, F., Urdan, T. (Eds.), *Self-efficacy Beliefs of Adolescents*, vol. 5. Information Age, Greenwich, CT, pp. 307–337.
- Barbera, J., Adams, W.K., Wieman, C.E., Perkins, K.K., 2008. Modifying and validating the Colorado learning attitudes about science survey for use in chemistry. *J. Chem. Educ.* 85, 1435–1439.
- Barnby, P., Kind, P., Jones, K., 2008. Examining changing attitudes in secondary school science. *Int. J. Sci. Educ.* 30, 1075–1093.
- Bartlett, S., Fitzgerald, M.T., McKinnon, D.H., Danaia, L., Lazendic-Galloway, J., 2018. Astronomy and Science Student Attitudes (ASSA): a short review and validation of a new instrument. *J. Astron. Earth Sci. Educ.* 5.
- Bauer, C.F., 2008. Attitude towards chemistry: a semantic differential instrument for assessing curriculum impacts. *J. Chem. Educ.* 85, 1440–1445.
- Becker, K.H., Maunsaiyat, S., 2002. Thai students' attitudes and concepts of technology. *J. Technol. Educ.* 13, 6–20.
- Bennett, J., Hogarth, S., 2009. Would YOU want to talk to a scientist at a party? High school students' attitudes to school science and to science. *Int. J. Sci. Educ.* 31, 1975–1998.
- Betz, N.E., Hackett, G., 1983. The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *J. Vocat. Behav.* 23, 329–345.
- Bodzin, A., Hammond, T., Fu, Q., Farina, W., 2020. Development of instruments to assess students' spatial learning attitudes (SLA) and interest in science, technology and geospatial technology (STEM-GEO). *Int. J. Educ. Methodol.* 6, 67–81.
- Britner, S.L., Pajares, F., 2006. Sources of science self-efficacy beliefs of middle school students. *J. Res. Sci. Teach.* 43, 485–499.
- Byars-Winston, A., Rogers, J., Branchaw, J., et al., 2016. New measures assessing predictors of academic persistence for historically underrepresented racial-ethnic undergraduates in science. *CBE Life Sci. Educ.* 15, ar32.
- Calabrese Barton, A., Kang, H., Tan, E., et al., 2013. Crafting a future in science: tracing middle school girls' identity work over time and space. *Am. Educ. Res. J.* 50, 37–75.
- Capobianco, B.M., French, B.F., Diefes-Dux, H.A., 2012. Engineering identity development among pre-adolescent learners. *J. Eng. Educ.* 101, 698–716.
- Caputo, A., 2017. A brief scale on Attitude Toward Learning of Scientific Subjects (ATLoSS) for middle school students. *J. Educ. Cult. Psychol. Stud.* 16, 57–76.
- Carberry, A.R., Lee, H.-S., Ohland, M.W., 2010. Measuring engineering design self-efficacy. *J. Eng. Educ.* 99, 71–79.
- Carlone, H.B., Johnson, A., 2007. Understanding the science experiences of successful women of color: science identity as an analytic lens. *J. Res. Sci. Teach.* 44, 1187–1218.
- Casem, R.Q., 2016. Effects of flipped instruction on the performance and attitude of high school students in mathematics. *Eur. J. STEM Educ.* 1, 37–44.
- Chamberlin, S.A., Powers, R.A., 2013. Assessing affect after mathematical problem solving tasks: validating the Chamberlin affective instrument for mathematical problem solving. *Gift. Educ. Int.* 29, 69–85.
- Chamberlin, S.A., Moore, A.D., Parks, K., 2017. Using confirmatory factor analysis to validate the Chamberlin affective instrument for mathematical problem solving with academically advanced students. *Br. J. Educ. Psychol.* 87, 422–437.
- Cheung, D., 2009. Developing a scale to measure students' attitudes toward chemistry lessons. *Int. J. Sci. Educ.* 31, 2185–2203.
- Chin, S.F., Lim, H.L., 2016. Validation of an adapted instrument to measure students' attitude towards science. *Asia Pac. Forum Sci. Learn. Teach.* 17.
- Chokchai, O., Papat, P., 2018. A student's scientific mind: a confirmatory factor analysis. *Malaysian J. Learn. Instr.* 15, 1–31.
- Code, W., Merchant, S., Maciejewski, W., Thomas, M., Lo, J., 2016. The Mathematics Attitudes and Perceptions Survey: an instrument to assess expert-like views and dispositions among undergraduate mathematics students. *Int. J. Math. Educ. Sci. Technol.* 47, 917–937.
- da Silva, J.M.T., Oliveira, A.L., Miguel, J.P., 2015. Cross-cultural adaptation and validation of a measure of attitudes toward statistics. *Rev. Iberoam. Diagn. Ev.* 39, 102–112.
- Dalgety, J., Coll, R.K., Jones, A., 2003. Development of chemistry attitudes and experiences questionnaire (CAEQ). *J. Res. Sci. Teach.* 40, 649–668.
- Dare, E.A., Roehrig, G.H., 2016. "If I had to do it, then I would": understanding early middle school students' perceptions of physics and physics-related careers by gender. *Phys. Rev. Phys. Educ. Res.* 12, 020117.
- Dare, E.A., Roehrig, G.H., 2020. Beyond Einstein and explosions: understanding 6th grade girls' and boys' perceptions of physics, school science, and STEM careers. *J. Women Minor. Sci. Eng.* 26, 541–577.
- Davadas, S.D., Lay, Y.F., 2018. Factors affecting students' attitudes toward mathematics: a structural equation modeling approach. *Eurasia J. Math. Sci. Technol. Educ.* 14, 517–529.
- Deci, E.L., Ryan, R.M., 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum, New York, NY.
- Dermizaki, I., Stavroussi, P., Vavougios, D., Kotsis, K.T., 2013. Adaptation of the students' motivation towards science learning (SMTSL) questionnaire in the Greek language. *Eur. J. Psychol. Educ.* 28, 747–766.

- Dewitt, J., Archer, L., Osborne, J., et al., 2011. High aspirations but low progression: the science aspirations-career paradox amongst minority ethnic students. *Int. J. Sci. Math. Educ.* 9, 243–271.
- Dijkstra, E.M., Goedhart, M.J., 2012. Development and validation of the ACSI: measuring students' science attitudes, pro-environmental behavior, climate change attitudes and knowledge. *Environ. Educ. Res.* 18, 733–749.
- Dong, Z., Li, M., Minstrell, J., Cui, Y., 2020. Psychometric properties of Science Motivation Questionnaire II-Chinese version in two waves of longitudinal data. *Psychol. Sch.* 57, 1240–1256.
- Dorn, B., Elliott Tew, A., 2015. Empirical validation and application of the computing attitudes survey. *Comput. Sci. Educ.* 25, 1–36.
- Dou, R., Hazari, Z., Dabney, K., Sonnet, G., Sadler, P., 2019. Early informal STEM experiences and STEM identity: the importance of talking science. *Sci. Educ.* 103, 623–637.
- Douglas, K.A., Yale, M.S., Bennett, D.E., Haugan, M.P., Bryan, L.A., 2014. Evaluation of Colorado learning attitudes about science survey. *Phys. Rev. ST Phys. Educ. Res.* 10, 020128.
- Dweck, C.S., Leggett, E.L., 1988. A social-cognitive approach to motivation and personality. *Psychol. Rev.* 95, 256–273.
- Eccles, J.S., Wigfield, A., 1995. In the mind of the actor: the structure of adolescents' achievement task values and expectancy-related beliefs. *Pers. Soc. Psychol. Bull.* 21, 215–225.
- Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132.
- Eccles, J.S., Wigfield, A., Schiefele, U., 1998. Motivation to succeed. In: Damon, W., Eisenberg, N. (Eds.), *Handbook of Child Psychology: Social, Emotional, and Personality Development*, fifth ed. John Wiley & Sons, Hoboken, NJ, pp. 1017–1095.
- Ellis, A.K., Fouts, J.T., 1996. *Handbook of Educational Terms and Applications*. Eye on Education, Princeton, NJ.
- Ellis, J., Wieselmann, J., Sivaraj, R., et al., 2020. Toward a productive definition of technology in science and STEM education. *Contemp. Issues Technol. Teach. Educ.* 20, 472–496.
- English, H.B., English, A.C., 1958. *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms*. Longmans, Green, New York, NY.
- Erdogan, M., Özel, M., Uşak, M., Prokop, P., 2009. Development and validation of an instrument to measure university students' biotechnology attitude. *J. Sci. Educ. Technol.* 18, 255–264.
- European Commission, 2015. *Science Education for Responsible Citizenship*. European Union, Brussels, Belgium.
- Falk, J.H., Staus, N., Dierking, L.D., et al., 2016. Understanding youth STEM interest pathways within a single community: the Synergies project. *Int. J. Sci. Educ. Part B* 6, 369–384.
- Fennema, E., Sherman, J.A., 1976. Fennema-Sherman Mathematics Attitudes Scales: instruments designed to measure attitudes toward the learning of mathematics by females and males. *J. Res. Math. Educ.* 7, 324–326.
- Fennema, E., Sherman, J., 1977. Sex-related differences in mathematics achievement, spatial visualization and affective factors. *Am. Educ. Res. J.* 14, 51–71.
- Fouad, N.A., Smith, P.L., Enochs, L., 1997. Reliability and validity evidence for the middle school self-efficacy scale. *Meas. Eval. Counsel. Dev.* 30, 17–31.
- Fraser, B.J., 1978. Development of a test of science-related attitudes. *Sci. Educ.* 62, 509–515.
- Glover, K.R., Bodzin, A., 2019. Psychometric testing of a value-achievement-cost motivation survey for 12th grade health sciences students for use in simulation-based-games. *Simulat. Gaming* 50, 789–811.
- Glynn, S.M., Taasoobshirazi, G., Brickham, P., 2009. Science motivation questionnaire: construct validation with nonscience majors. *J. Res. Sci. Teach.* 46, 127–146.
- Glynn, S.M., Brickman, P., Armstrong, N., Taasoobshirazi, G., 2011. Science Motivation Questionnaire II: validation with science majors and nonscience majors. *J. Res. Sci. Teach.* 48, 1159–1176.
- Godwin, A., Potvin, G., Hazari, Z., 2013. The development of critical engineering agency, identity, and the impact on engineering career choices. In: *Proceedings of the 2013 American Society for Engineering Annual Conference and Exposition*. ASEE, Atlanta, GA.
- Godwin, A., Potvin, G., Hazari, Z., Lock, R., 2016. Identity, critical agency, and engineering: an affective model for predicting engineering as a career choice. *J. Eng. Educ.* 105, 312–340.
- Gogolin, L., Swartz, F., 1992. A quantitative and qualitative inquiry into the attitudes toward science of nonscience college students. *J. Res. Sci. Teach.* 29, 487–504.
- Gough, D., Thomas, J., Oliver, S., 2019. Clarifying differences between reviews within evidence ecosystems. *Syst. Rev.* 8, 170–184.
- Gul, S., 2017. Development of an attitude scale to measure the undergraduate students' attitudes towards nanobiotechnology. *J. Sci. Educ. Technol.* 26, 519–533.
- Guzey, S.S., Harwell, M., Moore, T., 2014. Development of an instrument to assess attitudes toward Science, Technology, Engineering, and Mathematics (STEM). *Sch. Sci. Math.* 114, 271–279.
- Halim, L., Rahman, N.A., Wahab, N., Mohtar, L.E., 2018. Factors influencing interest in STEM careers: an exploratory factor analysis. *Asia Pac. Forum Sci. Learn. Teach.* 19.
- Hall, C., Dickerson, J., Batts, D., Kauffmann, P., Bosse, M., 2011. Are we missing opportunities to encourage interest in STEM fields? *J. Technol. Educ.* 23, 32–46.
- Hammoudi, M.M., 2020. Measurement of students' mathematics motivation and self-concept at institutions of higher education: evidence of reliability and validity. *Int. J. Math. Educ. Sci. Technol.* 51, 63–86.
- Hassan, G., 2008. Attitudes toward science among Australian tertiary and secondary school students. *Res. Sci. Technol. Educ.* 26, 129–147.
- Hassan, S.S.S., 2018. Measuring attitude towards learning science in Malaysian secondary school context: implications for teaching. *Int. J. Sci. Educ.* 40, 2044–2059.
- Hazari, Z., Sonnet, G., Sadler, P.M., Shanahan, M.-C., 2010. Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study. *J. Res. Sci. Teach.* 47, 978–1003.
- Heredia, K., Lewis, J.E., 2012. A psychometric evaluation of the Colorado Learning Attitudes about Science Survey for use in chemistry. *J. Chem. Educ.* 89, 436–441.
- Hidi, S., Renninger, K.A., 2006. The four-phase model of interest development. *Educ. Psychol.* 41, 111–127.
- Hiller, S.E., Kitsantas, A., 2016. The validation of the citizen science self-efficacy scale (CSSES). *Int. J. Environ. Sci. Educ.* 11, 543–558.
- Hong, O., 2017. STEAM education in Korea: current policies and future directions. *Pol. Traject. Initiat. STEM Educ.* 8, 92–102.
- Huang, Y.-C., Lin, S.-H., 2015. Development and validation of an inventory for measuring student attitudes toward calculus. *Meas. Eval. Counsel. Dev.* 48, 109–123.
- Ibrahim, A., Aulls, M.W., Shore, B.M., 2016. Development, validation, and factorial comparison of the McGill Self-Efficacy of Learners for Inquiry Engagement (McSELFIE) survey in natural science disciplines. *Int. J. Sci. Educ.* 38, 2450–2476.
- Kahveci, M., 2010. Students' perceptions to use technology for learning: measurement integrity of the modified Fennema-Sherman Attitudes Scales. *Turkish Online J. Educ. Technol.* 9, 185–201.
- Kahveci, A., 2015. Assessing high school students' attitudes toward chemistry with a shortened semantic differential. *Chem. Educ. Res. Pract.* 16, 283–292.
- Kaya, H., 2012. An investigation into upper elementary students' attitudes towards science. *Eur. J. Phys. Educ.* 3, 22–33.
- Kennedy, J.P., Quinn, F., Taylor, N., 2016. The school science attitude survey: a new instrument for measuring attitudes towards school science. *Int. J. Res. Method Educ.* 39, 422–445.
- Khing, S.B., Guan, T.E., 2019. Development of a mathematics self-efficacy scale: a Rasch validation study. *J. Appl. Meas.* 20, 184–205.
- Kier, M.W., Blanchard, M.R., Osborne, J.W., Albert, J.L., 2014. The development of the STEM career interest survey (STEM-CIS). *Res. Sci. Educ.* 44, 461–481.
- Kind, P., Jones, K., Barnby, P., 2007. Developing attitudes towards science measures. *Int. J. Sci. Educ.* 29, 871–893.
- Koballa Jr., T.R., Glynn, S.M., 2007. Attitudinal and motivational constructs in science learning. In: Abell, S.K., Lederman, N.G. (Eds.), *Handbook of Research on Science Education*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ, pp. 75–102.
- Koballa, T.R., 1988. Attitude and related concepts in science education. *Sci. Educ.* 72, 115–126.
- Köksal, M.S., 2013. Comparison of gifted and advanced students on motivation toward science learning and attitude toward science. *J. Am. Acad. Spec. Educ. Prof.* 100–112.
- Kontas, H., Özcan, B., 2017. Adapting sources of middle school mathematics self-efficacy scale to Turkish culture. *Int. J. Eval. Res. Educ.* 6, 288–294.

- Koul, R.B., Fraser, B.J., Maynard, N., Tade, M., 2018. Evaluation of engineering and technology activities in primary schools in terms of learning environment, attitudes and understanding. *Learn. Environ. Res.* 21, 285–300.
- Koyunlu Unlu, Z., Dokme, I., Unlu, V., 2016. Adaptation of the science, technology, engineering, and mathematics career interest survey (STEM-CIS) into Turkish. *Eur. J. Educ. Res.* 63, 21–36.
- Kranzler, J.H., Pajares, F., 1997. An exploratory factor analysis of the mathematics self-efficacy scale revised (MSES-R). *Meas. Eval. Counsel. Dev.* 29, 215–228.
- Krathwohl, D.R., Bloom, B.S., Masia, B.B., 1964. Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook II: Affective Domain. David McKay, New York, NY.
- Krishnan, S., Idris, N., 2015. Investigating SATS-36 for a matriculation sample. *Int. J. Eval. Res. Educ.* 4, 84–88.
- Kwon, H., 2016. Middle school students' motivation for learning technology in South Korea. *Eurasia J. Math. Sci. Technol. Educ.* 12, 1033–1046.
- Lachapelle, C.P., Brennan, R.T., 2018. An instrument for examining elementary engineering student interests and attitudes. *Int. J. Educ. Math. Sci. Technol.* 6, 221–240.
- Lamb, R.L., Annetta, L., Meldrum, J., Vallett, D., 2012. Measuring science interest: Rasch validation of the science interest survey. *Int. J. Sci. Math. Educ.* 10, 643–668.
- Lent, R.W., Lopez, F.G., Bieschke, K.J., 1991. Mathematics self-efficacy: sources and relation to science-based career choice. *J. Counsel. Psychol.* 38, 424–430.
- Lent, R.W., Brown, S.D., Hackett, G., 1994. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122.
- Lent, R.W., Lopez, F.G., Brown, S.D., Gore, P.A., 1996. Latent structure of the sources of mathematics self-efficacy. *J. Vocat. Behav.* 49, 292–308.
- Liaghatdar, M.J., Soltani, A., Abedi, A., 2011. A validity study of attitudes toward science scale among Iranian secondary school students. *Int. Educ. Stud.* 4, 36–46.
- Lim, S.Y., Chapman, E., 2013. Development of a short form of the attitudes toward mathematics inventory. *Educ. Stud. Math.: Int. J.* 82, 145–164.
- Lim, S.Y., Chapman, E., 2015. Adapting the academic motivation scale for use in pre-tertiary mathematics classrooms. *Math. Educ. Res. J.* 27, 331–357.
- Lin, T.-J., Tsai, C.-C., 2013. A multi-dimensional instrument for evaluating Taiwanese high school students' science learning self-efficacy in relation to their approaches to learning science. *Int. J. Sci. Math. Educ.* 11, 1275–1301.
- Lin, T.-J., Tan, A.L., Tsai, C.-C., 2013. A cross-cultural comparison of Singaporean and Taiwanese eighth graders' science learning self-efficacy from a multi-dimensional perspective. *Int. J. Sci. Educ.* 35, 1083–1109.
- Lin, T.-C., Liang, J.-C., Tsai, C.-C., 2015. Conceptions of memorizing and understanding in learning, and self-efficacy held by university biology majors. *Int. J. Sci. Educ.* 37, 446–468.
- Liou, P.-Y., Kuo, P.-J., 2014. Validation of an instrument to measure students' motivation and self-regulation towards technology learning. *Res. Sci. Technol. Educ.* 32, 79–96.
- Lopez, F.G., Lent, R.W., 1992. Sources of mathematics self-efficacy in high school students. *Career Dev. Q.* 41, 3–12.
- Luo, W., Wei, H.-R., Ritzhaupt, A.D., Huggins-Manley, A.C., Gardner-McCune, C., 2019. Using the S-STEM survey to evaluate a middle school robotics learning environment: validity evidence in a different context. *J. Sci. Educ. Technol.* 28, 429–443.
- Luttrell, V.R., Callen, B.W., Allen, C.S., et al., 2010. The mathematics value inventory for general education students: development and initial validation. *Educ. Psychol. Meas.* 70, 142–160.
- Mahoney, M.P., 2010. Students' attitudes toward STEM: development of an instrument for high school STEM-based programs. *J. Technol. Stud.* 36, 24–34.
- Maltese, A.V., Cooper, C.S., 2017. STEM pathways: do men and women differ in why they enter and exit? *AERA Open* 3, 1–16.
- Maltese, A.V., Tai, R.H., 2011. Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among U.S. students. *Sci. Educ.* 95, 877–907.
- Marat, D., 2005. Assessing mathematics self-efficacy of diverse students from secondary schools in Auckland: implications for academic achievement. *Issues Educ. Res.* 15, 37–68.
- Marra, R.M., Schuurman, M., Moore, C., Bogue, B., 2005. Women engineering students' self-efficacy beliefs—the longitudinal picture. In: *Proceedings of the 2005 American Society for Engineering Education Annual Conference and Exposition*. ASEE, Portland, OR.
- Marra, R.M., Rodgers, K.A., Shen, D., Bogue, B., 2009. Women engineering students and self-efficacy: a multi-year, multi-institution study of women engineering student self-efficacy. *J. Eng. Educ.* 98, 27–38.
- Mason, A.J., Bertram, C.A., 2018. Consideration of learning orientations as an application of achievement goals in evaluating life science majors in introductory physics. *Phys. Rev. Phys. Educ. Res.* 14, 010125.
- Mason, S.L., Rich, P.J., 2020. Development and analysis of the elementary student coding attitudes survey. *Comput. Educ.* 153, 103898.
- Mau, W.-C., Chen, S.J., Lin, C.-C., 2019. Assessing high school student's STEM career interests using a social cognitive framework. *Educ. Sci.* 9.
- McCormick, M., Bielefeldt, A.R., Swan, C.W., Paterson, K.G., 2015. Assessing students' motivation to engage in sustainable engineering. *Int. J. Sustain. High Educ.* 16, 136–154.
- McGuire, W.J., 1985. Attitudes and attitude change. In: *Kindzey, G., Aronson, E. (Eds.), Handbook of Social Psychology*, third ed. Random House, New York, NY, pp. 233–346.
- McLeod, D.B., 1992. Research on affect in mathematics: a reconceptualization. In: *Grouws, D.A. (Ed.), Handbook of Research on Mathematics Teaching and Learning*. Macmillan, New York, NY, pp. 575–596.
- Meredith, J.E., Fortner, R.W., Mullins, G.W., 1997. Model of affective learning for nonformal science education facilities. *J. Res. Sci. Teach.* 34, 805–818.
- Messick, S., 1989. Validity. In: *Linn, R.J. (Ed.), Educational Measurement*. Macmillan, New York, NY, pp. 13–103.
- Milner, D.I., Horan, J.J., Tracey, T.J.G., 2014. Development and evaluation of STEM interest and self-efficacy tests. *J. Career Awareness* 22, 642–653.
- Miranda, J.C., Rada, T., 2018. An examination of the beliefs about physics and learning physics among engineering students. *Int. J. Eng. Educ.* 34, 1427–1435.
- Montes, L.H., Ferreira, R.A., Rodríguez, C., 2018. Explaining secondary school students' attitudes towards chemistry in Chile. *Chem. Educ. Res. Pract.* 19, 533–542.
- Moore, R.W., Sutman, F.X., 1970. The development, field test and validation of an inventory of scientific attitudes. *J. Res. Sci. Teach.* 7, 85–94.
- Mozahem, N.A., Boulad, F.M., Ghanem, C.M., 2020. Secondary school students and self-efficacy in mathematics: gender and age differences. *Int. J. Sch. Educ. Psychol.* <https://doi.org/10.1080/21683603.2020.1763877>.
- Munby, H., 1982. The impropriety of “panel of judges” validation in science attitude scales: a research comment. *J. Res. Sci. Teach.* 19, 617–619.
- Munby, H., 1983. Thirty studies involving the “scientific attitude inventory”: what confidence can we have in this instrument? *J. Res. Sci. Teach.* 20, 141–162.
- Mutohir, T.C., Lowrie, T., Patahuddin, S.M., 2018. The development of a student survey on attitudes towards mathematics teaching-learning processes. *J. Math. Educ.* 9, 1–14.
- Nadirova, A., Burger, J., 2008. Evaluation of elementary students' attitudes toward science as a result of the introduction of an enriched science curriculum. *Alberta J. Educ. Res.* 54, 30–49.
- National Research Council, 2011. *Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics*. National Academies Press, Washington, DC.
- National Research Council, 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press, Washington, DC.
- Newman, M., Gough, D., 2020. Systematic reviews in educational research: methodology, perspectives and application. In: *Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., Buntins, K. (Eds.), Systematic Reviews in Educational Research*. Springer VS, Wiesbaden, Germany, pp. 3–22.
- Nugent, G., Barker, B., Welch, G., et al., 2015. A model of factors contributing to STEM learning and career orientation. *Int. J. Sci. Educ.* 37, 1067–1088.
- Oh, Y.J., Jia, Y., Lorentson, M., LaBanca, F., 2013. Development of the educational and career interest scale in science, technology, and mathematics for high school students. *J. Sci. Educ. Technol.* 22, 780–790.
- Olango, M., 2015. Effects of selected affective variables on mathematics achievement of freshmen science and engineering students: the case of Hawassa University. *Afr. Educ. Res. J.* 3, 255–268.
- Olango, M., 2016. Mathematics anxiety factors as predictors of mathematics self-efficacy and achievement among freshmen science and engineering students. *Afr. Educ. Res. J.* 4, 109–123.
- Orosco, M.J., 2016. Measuring elementary student's mathematics motivation: a validity study. *Int. J. Sci. Math. Educ.* 14, 945–958.

- Osborne, J., Simon, S., Collins, S., 2003. Attitudes towards science: a review of the literature and its implications. *Int. J. Sci. Educ.* 25, 1049–1079.
- Osterlind, S.J., 2006. *Modern Measurement: Theory, Principles, and Applications of Mental Appraisal*. Pearson, Upper Saddle, NJ.
- Owen, S.V., Toepferwein, A., Marshall, C.E., et al., 2007. Finding pearls: psychometric reevaluation of the Simpson-Troost Attitude Questionnaire (STAQ). *Sci. Educ.* 92, 1076–1095.
- Ozcan, H., Koca, E., 2020. Development of the attitude towards science scale: a validity and reliability study. *Eur. J. Educ. Res.* 85, 109–134.
- Ozden, C., Tezer, M., Atasoy, R., 2019. Developing technology and design course self-efficacy scale: a validity and reliability study. *World J. Educ. Technol.* 11, 186–197.
- Patrick, A.D., Borrego, M., Prybutok, A.N., 2018. Predicting persistence in engineering through an engineering identity scale. *Int. J. Eng. Educ.* 34, 351–363.
- Paul, K.M., Maltese, A.V., Valdivia, D.S., 2020. Development and validation of the role identity surveys in engineering (RIS-E) and STEM (RIS-STEM) for elementary students. *Int. J. STEM Educ.* 7.
- Pell, T., Jarvis, T., 2001. Developing attitude to science scales for use with children of ages from five to eleven years. *Int. J. Sci. Educ.* 23, 847–862.
- Petticrew, M., Roberts, H., 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell Publishing, Malden, MA.
- Petty, R.E., Cacioppo, J.T., 1981. *Attitude and Persuasion: Classic and Contemporary Approaches*. Routledge, New York, NY.
- Pintrich, P.R., Smith, D.A.F., Garcia, T., McKeachie, W.J., 1993. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educ. Psychol. Meas.* 53, 801–813.
- What do girls and boys think of technology? In: Raat, J.H., de Vries, M.J. (Eds.), 1986. Report PATT-Workshop, March 6–11, 1986. Eindhoven University of Technology, Eindhoven, Netherlands.
- Rachmatullah, A., Roshayanti, F., Shin, S., Lee, J.-K., Ha, M., 2018. The secondary-student science learning motivation in Korea and Indonesia. *Eurasia J. Math. Sci. Technol. Educ.* 14, 3123–3141.
- Ren, L., Green, J.L., Smith, W.M., 2016. Using the Fennema-Sherman mathematics attitude scales with lower-primary teachers. *Math. Educ. Res. J.* 28, 303–326.
- Rocabado, G.A., Kilpatrick, N.A., Mooring, S.R., Lewis, J.E., 2019. Can we compare attitude scores among diverse populations? An exploration of measurement invariance testing to support valid comparisons between Black female students and their peers in an organic chemistry course. *J. Chem. Educ.* 96, 2371–2382.
- Rodríguez-Ayan, M.N., Rico, M.E.S., 2015. Validation of the Spanish version of the Mathematics Value Inventory (MVI) among Uruguayan University Students. *Psicol. Reflex. Crit.* 28, 678–689.
- Roller, S.A., Lampley, S.A., Dillihunt, M.L., et al., 2020. Development and initial validation of the Student Interest and Choice in STEM (SIC-STEM) Survey 2.0 instrument for assessment of the social cognitive career theory constructs. *J. Sci. Educ. Technol.* 29, 646–657.
- Romine, W., Sadler, T.D., Presley, M., Klosterman, M.L., 2014. Student Interest in Technology and Science (SITS) survey: development, validation, and use of a new instrument. *Int. J. Sci. Math. Educ.* 12, 261–283.
- Romine, W.L., Miller, M.E., Knese, S.A., Folk, W.R., 2016. Multilevel assessment of middle school students' interest in the health sciences: development and validation of a new measurement tool. *CBE Life Sci. Educ.* 15, 1–13.
- Rosenthal-Von Der Pütten, A., Bock, N., 2018. Development and validation of the self-efficacy in human-robot-interaction scale (SE-HRI). *ACM Trans. Hum. Robot Interact.* 7, 21.
- Said, Z., Summers, R., Abd-El-Khalick, F., Wang, S., 2016. Attitudes toward science among grades 3 through 12 Arab students in Qatar: findings from a cross-sectional national study. *Int. J. Sci. Educ.* 38, 621–643.
- Saidi, S.S., Siew, N.M., 2019. Investigating the validity and reliability of survey attitude towards statistics instrument among rural secondary school students. *Int. J. Educ. Methodol.* 5, 651–661.
- Salta, K., Koulougliotis, D., 2015. Assessing motivation to learn chemistry: adaptation and validation of Science Motivation Questionnaire II with Greek secondary school students. *Chem. Educ. Res. Pract.* 16, 237–250.
- Schau, C., Stevens, J., Dauphinee, T.L., Del Vecchio, A., 1995. The development and validation of the survey of attitudes toward statistics. *Educ. Psychol. Meas.* 55, 868–875.
- Schreiner, C., Sjøberg, S., 2004. Sowing the Seeds of ROSE: Background, Rationale, Questionnaire Development and Data Collection for ROSE (The Relevance of Science Education)—A Comparative Study of Students' Views of Science and Science Education. Unipub AS, Oslo, Norway.
- Semsar, K., Knight, J.K., Birol, G., Smith, M.K., 2011. The Colorado Learning Attitudes about Science Survey (CLASS) for use in biology. *CBE Life Sci. Educ.* 10, 268–278.
- Sha, L., Schunn, C., Bathgate, M., 2015. Measuring choice to participate in optional science learning experiences during early adolescence. *J. Res. Sci. Teach.* 52, 686–709.
- Siegel, M.A., Ranney, M.A., 2003. Developing the Changes in Attitude about the Relevance of Science (CARS) questionnaire and assessing two high school science classes. *J. Res. Sci. Teach.* 40, 757–775.
- Simpson, R.D., Troost, K.M., 1982. Influences on commitment to and learning of science among adolescent students. *Sci. Educ.* 66, 763–781.
- Skinner, E., Saxton, E., Currie, C., Shusterman, G., 2017. A motivational account of the undergraduate experience in science: brief measures of students' self-system appraisals, engagement in coursework, and identity as a scientist. *Int. J. Sci. Educ.* 39, 2433–2459.
- Smith-Colin, J.A., Wieselmann, J.R., 2020. Effective identity-safety cues for assuaging social identity threat of young Black girls in STEM. In: *Proceedings of the 2020 American Society for Engineering Education Virtual Conference*. ASEE.
- Spinner, H., Fraser, B.J., 2005. Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development. *Int. J. Sci. Math. Educ.* 3, 267–293.
- Stanisavljevic, D., Trajkovic, G., Marinkovic, J., et al., 2014. Assessing attitudes towards statistics among medical students: psychometric properties of the Serbian version of the Survey of Attitudes Towards Statistics (SATS). *PLoS One* 9, e112567.
- Staus, N.L., Lesseig, K., Lamb, R., Falk, J., Dierking, L., 2020. Validation of a measure of STEM interest for adolescents. *Int. J. Sci. Math. Educ.* 18, 279–293.
- Summers, R., Abd-El-Khalick, F., 2018. Development and validation of an instrument to assess student attitudes toward science across grades 5 through 10. *J. Res. Sci. Teach.* 55, 172–205.
- Summers, R., Wang, S., Abd-El-Khalick, F., Said, Z., 2019. Comparing Likert scale functionality across culturally and linguistically diverse groups in science education research: an illustration using Qatari students' responses to an attitude toward science survey. *Int. J. Sci. Math. Educ.* 17, 885–903.
- Svenningsson, J., Hultén, M., Hallström, J., 2018. Understanding attitude measurement: exploring meaning and use of the PATT short questionnaire. *Int. J. Technol. Des. Educ.* 28, 67–83.
- Tabuk, M., 2018. Turkish adaptation of the short form of attitude towards mathematics scale. *Mersin Univ. J. Fac. Educ.* 14, 84–95.
- Taconis, R., de Putter-Smits, L.G.M., Henry, S., den Brok, P.J., Beijard, D., 2010. The construction of a questionnaire to evaluate the science orientedness of students' identities as learners from a cognitive perspective. *Educ. Res. Eval.* 16, 383–400.
- Tapia, M., Marsh II, G.E., 2004. An instrument to measure mathematics attitudes. *Acad. Exchange Q.* 8, 16.
- Teo, T., Noyes, J., 2008. Development and validation of a computer attitude measure for young students (CAMYS). *Comput. Hum. Behav.* 24, 2659–2667.
- Toma, R.B., Meneses Villagra, J.A., 2019. Validation of the single-items Spanish-School Science Attitude Survey (S-SSAS) for elementary education. *PLoS One* 14, e0209027.
- Tsai, M.-J., Wang, C.-Y., Hsu, P.-F., 2019. Developing the computer programming self-efficacy scale for computer literacy education. *J. Educ. Comput. Res.* 56, 1345–1360.
- Tuan, H.-L., Chin, C.-C., Shieh, S.-H., 2005. The development of a questionnaire to measure students' motivation towards science learning. *Int. J. Sci. Educ.* 27, 639–654.
- Tytler, R., 2014. Attitudes, identity, and aspirations toward science. In: Lederman, N.G., Abell, S.K. (Eds.), *Handbook of Research on Science Education*, vol. 2. Routledge, New York, NY, pp. 82–103.
- Unfried, A., Faber, M., Stanhope, D.S., Wiebe, E., 2015. The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *J. Psychoeduc. Assess.* 33, 622–639.
- Usher, E.L., Pajares, F., 2009. Sources of self-efficacy in mathematics: a validation study. *Contemp. Educ. Psychol.* 34, 89–101.
- Utley, J., 2007. Construction and validity of geometry attitude scales. *Sch. Sci. Math.* 107, 89–93.

- Villafañe, S.M., Lewis, J.E., 2016. Exploring a measure of science attitude for different groups of students enrolled in introductory college chemistry. *Chem. Educ. Res. Pract.* 17, 731–742.
- Vincent-Ruz, P., Schunn, C.D., 2018. The nature of science identity and its role as the driver of student choices. *Int. J. STEM Educ.* 5.
- Wang, T.-L., Berlin, D., 2010. Construction and validation of an instrument to measure Taiwanese elementary students' attitudes toward their science class. *Int. J. Sci. Educ.* 32, 2413–2428.
- Wang, Y.-L., Tsai, C.-C., 2020. An investigation of Taiwanese high school students' basic psychological need satisfaction and frustration in science learning contexts in relation to their science learning self-efficacy. *Int. J. Sci. Math. Educ.* 18, 43–59.
- Wang, P., Palocsay, S.W., Shi, J., White, M.M., 2018. Examining undergraduate students' attitudes toward business statistics in the United States and China. *Decis. Sci. J. Innovat. Educ.* 16, 197–216.
- Weinburgh, M.H., Steele, D., 2000. The modified attitudes toward science inventory: developing an instrument to be used with fifth grade urban students. *J. Women Minor. Sci. Eng.* 6, 87–94.
- Wieselmann, J.R., Saito, T., Okumura, J.-I., 2018. Japanese adaptation of the STEM career interest survey (STEM-CIS). In: Presentation at the East-Asian Association for Science Education International Meeting. Hualien, Taiwan.
- Wieselmann, J.R., Duschl, R., Reiley, K., Berry, K., 2020a. Student dispositions toward STEM: exploring an engineering summer camp for underrepresented students. In: Proceedings of the 2020 American Society for Engineering Education Virtual Conference. ASEE.
- Wieselmann, J.R., Roehrig, G.H., Kim, J.N., 2020b. Who succeeds in STEM? Elementary girls' attitudes and beliefs about self and STEM. *Sch. Sci. Math.* 120, 233–244.
- Xu, X., Lewis, J.E., 2011. Refinement of a chemistry attitude measure for college students. *J. Chem. Educ.* 88, 561–568.
- Yáñez-Marquina, L., Villardón-Gallego, L., 2016. Attitudes towards mathematics at secondary level: development and structural validation of the Scale for Assessing Attitudes Towards Mathematics in Secondary Education (SAT-MAS). *Electron. J. Res. Educ. Psychol.* 14, 557–581.
- Yilmaz, H., Çavaş Huyugüzel, P., 2007. Reliability and validity study of the Students' Motivation Toward Science Learning (SMTSL) questionnaire. *Elem. Educ. Online* 6, 430–440.
- Yoon, S.Y., Sorby, S.A., 2020. Rescaling the longitudinal assessment of engineering self-efficacy V3.0 for undergraduate engineering students. *J. Psychoeduc. Assess.* 38, 209–221.
- You, H.S., Kim, K., Black, K., Min, K.W., 2018. Assessing science motivation for college students: validation of the Science Motivation Questionnaire II using the Rasch-Andrich rating scale model. *Eurasia J. Math. Sci. Technol. Educ.* 14, 1161–1173.
- Zakariya, Y.F., Goodchild, S., Bjørkestøl, K., Nilsen, H.K., 2019. Calculus self-efficacy inventory: its development and relationship with approaches to learning. *Educ. Sci.* 9.
- Zakariya, Y.F., 2017. Development of Attitudes Towards Mathematics Scale (ATMS) using Nigerian data—factor analysis as a determinant of attitude subcategories. *Int. J. Prog. Educ.* 13, 74–84.
- Zhang, D., Campbell, T., 2011. The psychometric evaluation of a three-dimension elementary science attitude survey. *J. Sci. Teach. Educ.* 22, 595–612.
- Zheng, L., Dong, Y., Huang, R., Chang, C.-Y., Bhagat, K., 2018. Investigating the interrelationships among conceptions of, approaches to, and self-efficacy in learning science. *Int. J. Sci. Educ.* 40, 139–158.
- Zimprich, D., 2012. Attitudes toward statistics among Swiss psychology students. *Swiss J. Psychol.* 71, 149–155.
- Zwicky, B.M., Hirokawa, T., Finkelstein, N., Lewandowski, H.J., 2014. Epistemology and expectations survey about experimental physics: development and initial results. *Phys. Rev. ST Phys. Educ. Res.* 10, 010120.

Using flexible thinking to assess student sensemaking of phenomena in project-based learning

E.C. Adah Miller^a, T. Li^b, I.C. Chen^b, and S.K. Codere^b, ^a University of Wisconsin-Madison, Madison, WI, United States; and ^b Create for STEM Institute, University of Michigan State, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Problem and objective	444
Significance of work	445
Theoretical framework	445
Design process	446
Objective 1. Elaboration and synthesis of unit learning goals (steps 1 and 2 in)	446
Objective 2. Creation of flexible task performance expectations, flexible performance tasks, evidence statements, and rubrics (steps 3 through 6 in)	448
Objective 3. Testing and redesign of performance tasks (step 7)	449
Develop rubrics for depth, flexibility and robust community perspective	449
Depth of knowledge	449
Robust community perspective	450
Designed for flexibility	450
Methodology and analysis	450
Context and participation	450
Participation	451
Sample	451
Data sources	451
Cognitive interviews. Pre and post-unit assessment	451
Data analysis	452
Analysis of student data	452
Example of student response on the squirrel post-unit assessment focus group	452
Response analysis	452
Context	452
Depth of knowledge as evidence of knowledge in use	452
Flexible thinking as evidence of knowledge in use	452
Robust community perspective	454
Major findings	454
Discussion and significance	456
Acknowledgment	456
References	456

Problem and objective

The goals of science education move toward developing all student knowledge-in-use (NRC, 2012; NGSS Lead States, 2013). The knowledge-in-use perspective reflects an increased awareness by learning scientists, science educators and policymakers of the proficiencies that learners will need as global citizens in the 21st Century (Choi et al., 2011; OECD, 2016) in order to address the current and emerging challenges facing our world. All learners will need to use and apply knowledge in novel ways to solve important problems and make decisions that affect their personal lives and our global society. *The Framework for K-12 Science Education* (hereafter: *The Framework*, NRC, 2012) advances this knowledge-in-use perspective of learning. Knowledge-in-use focuses on building and refining knowledge, solving problems, investigating the natural world and learning more when needed and results in deeper, more sophisticated knowledge that learners can apply. This knowledge-in-use perspective is the basis for the *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013) in the US and other national standards throughout the globe, such as the Finnish standards (Finnish National Board of Education, 2016) and German standards (Kulgemeyer and Schecker, 2014). Knowledge-in-use is also the perspective taken by PISA (OECD, 2016). With knowledge-in-use, students can develop deep, meaningful learning in which they use ideas to solve problems and make decisions by building and applying knowledge (Pellegrino and Hilton, 2012). However, how to support student equitable knowledge-in-use development with sensemaking has been a challenge in the science education field for a while (NGSS Lead States, 2013). To do that, *The Framework* and the NGSS suggest that three-dimensional (3D) learning could be one way to achieve this goal (NRC, 2012; NGSS Lead States, 2013). That is, student knowledge-in-use can be developed through using and applying knowledge in the context of three dimensions of science knowledge—science and engineering practices

(SEPs), crosscutting concepts (CCCs), and disciplinary core ideas (DCIs)—to make sense of phenomena or solve problems (NRC, 2012). With so doing, students can immerse in sensemaking process, which means they can not only connect their existing knowledge with new knowledge but can also apply what their scientific knowledge into different contexts and some new domains flexibly.

One main goal of science education is to ensure all students have the learning opportunities to learn, regardless of gender, race, ethnicity, socioeconomic status, disabilities, and English Language proficiency. To reach this inclusive goal, the performance expectations (PEs), an integration of a scientific practice, disciplinary core idea (science idea), and crosscutting concept, articulated in the NGSS are fundamental parts to support student sensemaking process and then further lead to knowledge-in-use development. The NGSS PEs promote the development of an ability to approach unfamiliar phenomena and solve complex problems with explanations and solutions. These processes involve consideration of possible viable approaches and use of a variety of science ideas and practices. A student (or a scientist) would have the capacity to sift through various ideas related to the phenomenon, for example, that could be explanatory, and ultimately, arrive at more than one acceptable explanation. In this way, the PEs, taken together, require learners to apply flexible thought. Acknowledging the range of probable and useful explanations and solutions to problems, and that there is more than one appropriate approach for each specific phenomenon or problem, is an intellectual and creative endeavor required of the next generation of scientifically literate global citizens.

Designing assessment tasks that can measure the three dimensions of scientific proficiency is a formidable task, and the field continues to struggle with how to construct assessment tasks that align with the knowledge-in-use perspective and that challenge students to solve problems and make sense of phenomena (Pellegrino and Hilton, 2012). Our vision is that assessment tasks designed with this knowledge-in-use perspective must engage learners in a complex or multi-layered phenomenon that gives rise to multiple avenues for sensemaking and affordances for using the three dimensions, similar to those in the real world. The tasks assess students' capacity to apply multiple practices, various elements of disciplinary core ideas, and several elements of crosscutting concepts as affordances for sensemaking. In this research, we explore how to construct and use knowledge-in-use assessment tasks and analyze how learners respond to such tasks.

Our work takes place in an elementary project-based learning (PBL) environment where students focus on responding to driving questions. We are developing and using knowledge-in-use assessment tasks that allow students to use multiple practices, CCCs and DCIs as tools to make sense of unfamiliar phenomena in order to promote flexible thinking, depth of knowledge, and community perspective (Pellegrino and Hilton, 2012; Baines et al., 2017; Spiro et al., 1991). We define flexible thinking as using multiple elements of DCIs, CCCs and SEPs to explain complex phenomena and examining alternative explanations for the same phenomenon (Ford and Wargo, 2012). It is important to note that to be a flexible thinker, students need to be able to access and evaluate all of the ideas that they may use to make sense of the phenomenon, and then make use of them as tools for sensemaking. Students would need to consider and discard some of the ideas as not applicable and apply others. We call this using of ideas, practices and CCCs affordances for sensemaking. We ask if this capacity can be measured, and if it can be developed over time through PBL and phenomenon-based units. Learners can draw from multiple science ideas to make sense of and explain a phenomenon (Miller et al., 2021).

Our research questions include: (1) How can we design assessment tasks that measure students' flexible use of knowledge to solicit and select from an array of the dimensions of scientific knowledge to make sense of an unfamiliar phenomenon or solve problems? and (2) How can we gather information to measure students' capacity to engage in socially situated practice of considering and weighing possible alternatives for making sense of the phenomenon under study?

Significance of work

Articulating an assessment process that will result in assessment tasks that provide robust and critical dialogic sensemaking for students and enable them to use their knowledge to think more flexibly and innovatively, closely resembling scientific and engineering endeavors. The assessment argument we present articulates a dynamic process for measuring scientific sensemaking as called for by policy documents throughout the globe (NRC, 2012; Finnish National Board of Education, 2016; Kulgemeyer and Schecker, 2014; OECD, 2016) that call for students to use their knowledge to solve problems, make decisions and think innovatively.

Knowledge-in-use goals provide new opportunities to support students in building capacity to approach unfamiliar problems and phenomena with "a developing set of ideas with explanatory possibilities" (Weick, 1995). However, assessing knowledge-in-use is challenging. This research presents one systematic method for designing assessment tasks that will engage learners in using multiple SEPs, DCIs and CCCs as affordances to make sense of complex phenomena. The tasks engage students in the figuring out process as well as gauge their depth and flexibility of knowledge, and connections to science as a social endeavor developed by and for communities.

Theoretical framework

Although policy documents from many countries describe a vision and process for learning that is situated in community and call for evaluation of a multiplicity of alternatives in science (Ford and Wargo, 2012), the interpretation of the NGSS performance expectations (PEs) in the United States for assessment purposes has focused primarily on using one SEP, an element of a DCI and a crosscutting concept. Such an approach focuses on learners applying their knowledge in a narrowly conceived area. Harris

et al. (2016) presented an approach based on the principles of evidence-centered design (ECD) (Mislevy and Haertel, 2006) to develop NGSS aligned assessment tasks for classroom use. The ECD approach emphasizes providing coherent and logical relationships among essential elements for designing assessments: (a) claims (learning goals) expressed as task performances; (b) evidence that reveals that the target learning goals are met; and (c) features of tasks that should elicit those behaviors or performances. Once these three conditions are met, this process can result in developing tasks that are highly-specific tasks, aligned to the standard or a discrete bundle of standards. The process results in evidence to support claims regarding student knowledge, and tasks to garner that information. This prior work focused on creating assessment tasks that can be used in classrooms for formative purposes and has made contributions in designing knowledge-in-use tasks. Focused on rich phenomena, these tasks assess the three dimensions of scientific proficiency—a science or engineering practice and an element of a DCI and CCC—to provide teachers with information on growth toward specific performance expectations. And while it is important to develop formative assessment tasks that can provide teachers and students with information on progress toward the standards, this work does not require learners to understand and apply multiple SEPs, elements of the DCIs and CCCs affordances to make sense of complex and novel phenomena.

We build on this prior work to design assessment tasks that better reflect knowledge-in-use including how scientists approach problems and how science is situated in the community. These assessments require learners to consider multiple scientific practices, multiple elements of DCIs and CCCs to make sense of phenomena or solve problems. Our approach builds from situated cognition and social constructivist theory of learning that propose that communities are complex systems in which the practices developed and refined within the community are responsive to the shifting goals of that community, and thus are created by and for the participants in the community (Lave and Wenger, 1991). The theory purports that sensemaking is an ongoing dialog and as such, making sense of phenomena is dynamic and messy (Miller and Januszyk, 2014). Social groups that consider and evaluate more ideas as potentially valuable have higher success rates in solving novel problems (Barron, 2003). While our curriculum resources provide goal statements that suggest specific paths for making sense of phenomena, incorporating assessment elements that reflect the *context-rich* processes of the learning allows for a more accurate portrayal of the figuring out process and thus of developing science competencies (Evans, 1999). Thus, building assessment tasks that require a multiplicity of approaches to solving the problem or making sense of a phenomenon more authentically matches the language and activities of sensemaking of STEM professionals. These assessments also provide more access to underrepresented groups such as English Learners (Moschkovich, 2007) and other underrepresented demographic groups in STEM who draw from diverse intellectual resources (Ladson-Billings, 2008; Pacheco, 2010). As students appropriate the scientific practices and science ideas, they negotiate and revise them for immediate use, a process which both improves performance and promotes a deepening understanding of what it means to “do science” (Ryu and Sandoval, 2012). Furthermore, to identify students’ strengths, we take into account how the assessment task is placed within the lived social practices of the participant and his or her interpretations of those practices (Gee, 2010; Shepard et al., 2017). As espoused in the *Framework* (NRC, 2012), to form useable understanding, knowing and doing cannot be separated, but rather must be experienced in an integrated fashion that allows for problem solving, decision making, explaining real-world phenomena, and connecting new ideas within a community of practice.

Design process

Our method for assessment development uses an iterative design-based research (Design-Based Research Collective, 2003) process in which we iteratively revise our assessment tasks based upon classroom use. Our methods for designing tasks include a seven-step process and meet three overarching objectives. We begin with unpacking in order to (1) elaborate and synthesize unit learning goals to (2) develop enduring understandings and affordances, (3) create flexible task performances, (4) design performance tasks based on related phenomena and the enduring understanding, (5) select appropriate flexible task performances and develop evidence statements, and (6) rubrics for scoring student performance tasks and evaluating test design, redesign the tasks to ensure that tasks elicit evidence of student learning, and (7) engage in iterative testing and redesign. Fig. 1 presents a schematic of the design process. The figure attempts to show the iterative and dynamic nature of the design process.

Objective 1. Elaboration and synthesis of unit learning goals (steps 1 and 2 in Fig. 1)

We first unpack and describe the array of performance expectations that students can draw from to make sense of phenomena. In unpacking we include determining the ideas that are critical for future application (knowledge-in-use) to solve problems and explain phenomena. The process involves breaking apart and expanding on the various ideas to elaborate important learning for the students at that grade level (Harris et al., 2016).

Each 6-week ML-PBL unit includes Learning Set Performances (LSP), which frame the broader three-dimensional learning goals of the unit, and provide unit coherence. Like the PEs from which they are derived, the LSPs are performance-based and engage learners in tasks to build knowledge-in-use toward PEs. The three dimensions of scientific knowledge from the LSPs stem from the initial unpacking of the NGSS performance expectations. The LSP captures more than one practice (or element of the practice), at least one DCI element, and at least one CCC element. It summarizes the LPs of the lessons in the Learning Set.

We consider the DCIs, SEPs, CCCs that make up each of the LSPs as an array of *possible affordances* for making sense of an unfamiliar phenomena. These affordances represent *the combination of experiences with the phenomenon and driving questions, and flow of the*

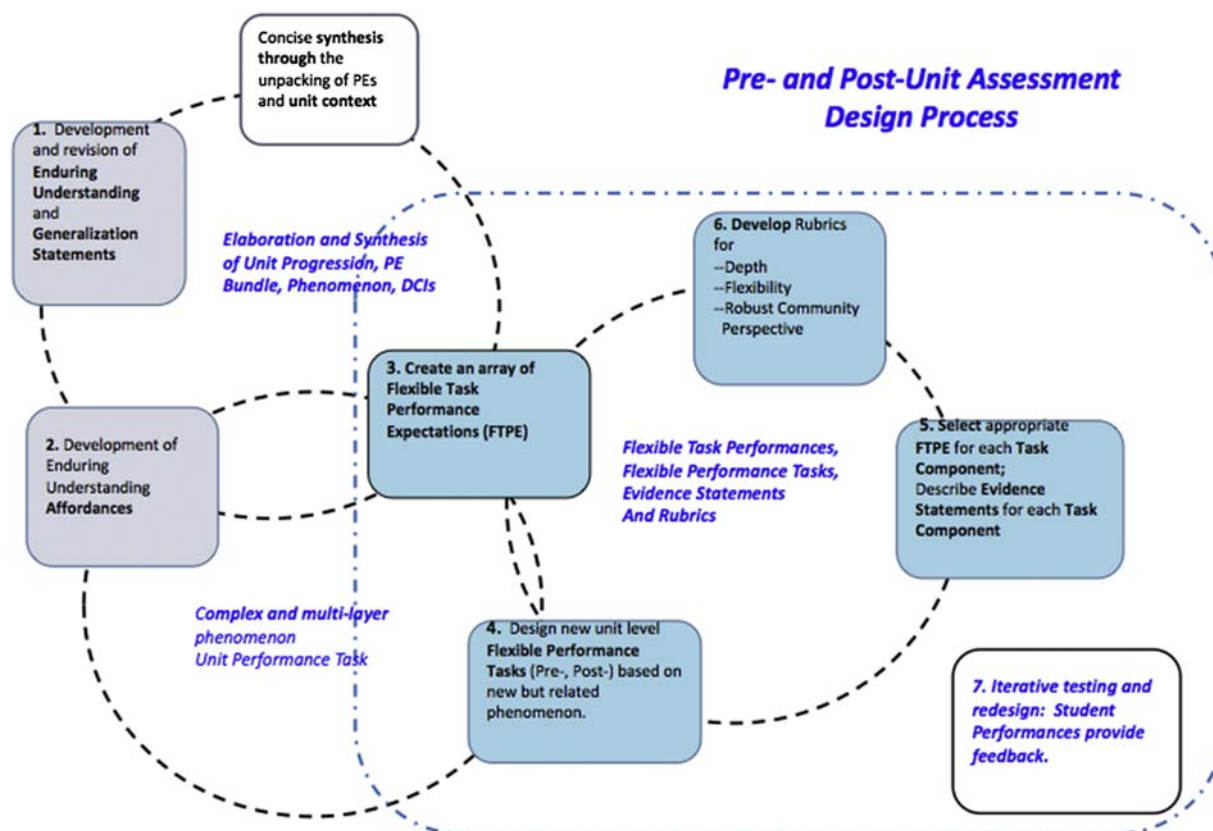


Fig. 1 Design process with seven steps to address three objectives.

unit, which can be represented in the Table one, below. We synthesize the elements of the LSPs and take into consideration the context of the unit to articulate an enduring understanding (Wiggins et al., 2005; Erickson, 2006) that reflects the big ideas students should take away from engaging in the unit and use to make sense of phenomena or to solve problems they encounter throughout their lives. Table 1 provides an example of the Enduring Understanding Affordances from the Squirrel Unit. The DCI elements from the LSPs are captured in the enduring understanding statement.

Evidence that students have developed knowledge-in-use of the enduring understanding is demonstrated by: *depth* of useable knowledge of affordances; *flexibility* in applying an array of DCIs, SEPs, and CCCs to make sense of phenomena; and a *robust community perspective* in which students make explicit connections to their local environment or social community.

Table 1 Enduring understanding affordances.

Practices from LSPs		Enduring understanding		CCCs from LSPs
- Observe squirrels (investigation) - Develop and revise a model - Construct claims using a model - Develop a model to explain - Analyze data from fossils - Argue from evidence - Construct claims with evidence	Using these ideas	- Species can survive hundreds of millions of years because organisms in that species are either very adaptable to changes in the environment (including changes in climate and changes in the other organisms) or because of wide diversity among them. - Generalizations: survival depends on change (adaptation). Change in the environment (over time) causes change in populations of organisms.	Using these ideas	- Systems and system models - Cause and effect - Structure and function - Stability and change - Scale, proportion and quantity - Patterns

Students use a flexible and robust array of affordances for making sense of a new but related phenomenon.

Objective 2. Creation of flexible task performance expectations, flexible performance tasks, evidence statements, and rubrics (steps 3 through 6 in Fig. 1)

While students have the opportunity to demonstrate learning connected to the phenomena throughout each unit, in the assessment, they apply an array of DCIs, SEPs, and CCCs to related but different phenomena that align to the unit driving question. The assessment phenomena must be complex and *multi-layered* (i.e., give rise to multiple avenues for sensemaking and affordances for using the three dimensions), *engaging and promoting wonderment* (i.e., student interviews show that students are motivated by the phenomena) and *related* to the unit level phenomenon (i.e., The squirrel unit is focused on looking at structures of fossils and comparing them to extant organisms, thinking about how they have changed over time to respond to needs; the related phenomenon is a change in habitat and predicting the physical response of living organisms). **Table 2** illustrates the development of related phenomenon for engagement with a new but related phenomenon in the assessment tasks.

Next, we use the phenomena to create *Flexible Task Performance Expectations (FTPE)* and *Performance Tasks*. **Fig. 2** shows an excerpt from one section of the third grade Squirrel Unit Assessment. **Figs. 3** and **4** show two more excerpts from the same assessment task.

It is important to note that the assessment offers students the opportunity to show flexible thinking and useable knowledge in more than one way. First, they may show two different DCI affordances to respond to Flexible Task Performance 1. For example a student might be able to explain how the squirrel would survive even if all the trees were cut down because they eat many different things—the idea that they are generalists (ESS3.A). Next they might be able to say that squirrels leave the area and move on (LS2.C).

Next, we point out that the students have the opportunity to demonstrate useable knowledge in the last prompt with the second new but related phenomenon. A student might use the idea that animals can adapt to changes in their environment over thousands of years (LS4.A or LS4.C).

Fig. 1 Assessment excerpt from the Squirrel Unit Assessment (chart and other text is not shown) Image attribution: Deforestation at Loch Thom <http://www.geograph.org.uk/photo/2298274>.

Each performance task is informed by the evidence of student learning: they must feature one or more practices, allow for various levels of depth and three dimensional applications for sensemaking. This means that a student could explain a phenomenon using a model, data analysis or a combination of scientific practices, and they could also draw upon the idea of adaptation, ecological systems, or animal needs for survival or a combination of these DCIs, and cause and effect, systems and systems thinking, and structure and function or combination of these CCCs to perform the figuring out.

We prompted for Robust Community Perspective with the following prompt: “How do people learn about squirrels and their changing habitats? How do you think learning about squirrels and their habitats can **help people**?”

Our aim with this question is to elicit understandings that students have developed about how science involves community-based practices, how science as an endeavor informs the community. We hope to help students build links between home,

Table 2 Develop new but related phenomenon.

Anchoring phenomenon		New but related phenomenon		
	We watch a movie that shows that a squirrel-like organisms lived at the same time as the stegosaurus and then look for squirrels and stegosaurus outside. They ask— <i>Why do I see so many squirrels but I can't find any stegosauruses?</i>	Two new but related phenomena: the squirrel's environment changes drastically. 	- Performance phenomenon must be related, rich, accurate, multi-layered and described - Performance phenomenon must be familiar, place based and accessible - Invoke wonder what will happen to the squirrel? Will it survive? What about future squirrels?	Related to anchoring phenomenon Invoke wonder —how would the squirrel population survive? Links to the DQ —the environment changes causing the squirrel to adapt to promote future species' success. The phenomenon is presented in a different context.
	Squirrel-like organisms live now and also during Jurassic.	- All the trees are cut down; - There is flooding for thousands of years and squirrels still live in the area.		

Think about Eastern Grey Squirrels and what they need from the forest.
Look at the picture on the left. If someone cuts down all the trees, what could happen to the Eastern Grey Squirrels who live in this area?
Draw a model to show what may happen to the squirrels. Make sure you show why you think that will happen. Use evidence from your experiences, the chart, and the text.

Fig. 2 Flexible performance task (FPT) 1.

You shared one idea of what could happen to the squirrels if the trees were cut down. Now, think of a very **different** idea about what might happen to the squirrels and explain your thinking.

Fig. 3 Flexible performance task (FPT) 2.

Look at the picture on Page 2 and imagine that the land flooded. Now there is so much water covering the land.

Imagine that Eastern Grey Squirrels are still living in the area after millions of years. What are possible changes you might see in Eastern Grey Squirrels that survived? How would those changes help them survive?

Write your answer and draw a picture.

Fig. 4 Cross-unit flexible performance task (FPT) 3.

community and science. For useable knowledge students need to become aware of how science practices develop over time based on local community needs (i.e., such as curiosity (Manz, 2012) or motivated by student sense-making (Berland and Hammer, 2012)). We hope to begin to elicit information regarding students' developing understanding about the relationship of scientific discovery and the larger community—scientific questions arise from community-based interests. The problems encountered by the larger community in the natural and design world inform and drive this mutual relationship. As we center all of our units in place-based and relevant driving questions, which are meant to blur the lines between classroom, home and community and science, we expect to see growth in sophistication for this question among students over time.

Objective 3. Testing and redesign of performance tasks (step 7)

We used initial evidence of student learning—for depth, flexibility and robust community perspective—to iteratively revise and refine assessment tasks based on open-ended cognitive, focus group interviews (Ginsburg, 1997) and through pre- and post-test enactment of the performance tasks during field study.

We made changes to the tasks based on the focus group interviews. While we made many revisions based on the focus groups, they were insubstantial. The majority of the changes were responding to modification in language (word use), number of lines for prompts, and readability, etc. Students reported that the tasks were “*really fun*,” and “*interesting*,” and that they would be able to “*learn from the test*” (Focus group, May 24, 2017).

Evidence of student learning incorporated both cognitive and sociocultural aspects in the following ways: (1) students make sense of phenomena using the three dimensions to guide the process, (2) students flexibly use the sensemaking tools (array of DCIs, SEPs and CCCs) for figuring out the phenomenon, and (3) students describe scientific decisions as being situated in the community, which serves as the both the motivator for science and gauge of outcomes (Zeidler et al., 2005).

Develop rubrics for depth, flexibility and robust community perspective

The rubrics were developed from our theoretical framework, appendices of the NGSS (NGSS Lead States, 2013), and an iterative analysis of the data (Miles and Huberman, 1994). Each rubric describes the (1) *depth* of three dimensional application to the task across responses to the complex task, (2) *flexibility* across at least one DCI or CCC, or two dimensions as an affordance for figuring out, and (3) reference to science as a *community*-embedded, or community-driven activity. Table 5 provides an example of the Depth Rubric. In the excerpt above (Fig. 1), students can make use of the following practices: construct and use a model to explain, analyze data, and use of evidence; elements of the DCIs of adaptation, survival in changing environments and ecological systems; and the crosscutting concepts of cause and effect, systems and system models, and structure and function.

We used the rubrics to score student responses.

Depth of knowledge

On the pre-assessment, students expressed simple ideas and were given a Depth of Knowledge designation of *Partial Affordance*. On the post-assessment, many students' responses showed depth of knowledge and flexible thinking and were given a Depth of Knowledge designation of *Quality* or *High Quality Affordance* (see Table 3). Below we provide an example of a student response on the post-unit assessment.

Table 3 Depth rubric (example).

<i>Partial affordance</i>	<i>Partial/complete affordance</i>	<i>Quality affordance</i>	<i>High quality affordance</i>
1. The squirrel will die/the squirrel will die because it needs nuts. Partial CCC cause and effect full element of EU Partial EU Partial claim	2. The squirrel will die because the squirrel meets its needs using things from the environment and the environment changes and the squirrel won't meet those needs and it will not survive. Partial CCC cause and effect full element of EU Partial claim	3. The squirrel can survive extreme circumstances (and is a generalist) and can adapt by doing something else. It will probably out a way to meet its needs. We saw that squirrels can survive on the roof of houses and eat berries. Full CCC cause and effect full element of EU Full claim	4. (As in 3, but at the species scale.) The squirrel will not only survive but it can have babies who will also survive and they will have behavior that is similar to the squirrel who survives by eating berries and living on a roof. Adds one or more dimensions to the 3D to provide depth. In this case added another element of the EU, and CCC (Scale, proportion and quantity)
1. The environment changed a lot and this impacts this squirrel. Partial CCC stability and change partial element of the EU partial claim	2. The squirrel lives in different kinds of environments over lots of years. This is because squirrels have a lot of different physical and/or behavioral characteristics. Partial CCC stability and change full element of EU partial claim	3. As the environment has changed, some animals like the squirrel have changed. As the squirrel changes over time it has become really different (like the marmot and the gray squirrel) as a response to meet its needs as a population. Full CCC stability and change full element of EU full claim	4. As in 3, but includes the idea that the squirrel also has changed the environment. Adds one or more dimensions to the 3D to provide depth. In this case added another element of the EU, CCC (systems models)

Robust community perspective

Because this performance task aims to elicit students' sensemaking of the context for science knowledge (social, environmental, historical, or any other context applied). There are three criteria for Robust Community Perspective (see [Table 4](#)).

1. Students provide evidence of community perspective for science, i.e., community from which they derived information (by referencing class or home/community, referencing text), and/or community with whom they are sharing information (by considering audience in their model and reasoning).
2. Students refer to social, critical, and historical perspectives or place and context.
3. Students include thoughts about science community and agreed-upon norms and practices.

Designed for flexibility

Students show multiple avenues of engagement using more than one element of Enduring Understanding, and CCC with practice in an answer to present possibilities or different outcomes or perspectives.

Methodology and analysis**Context and participation**

Our work is situated in upper elementary project-based learning environments ([Blumenfeld et al., 1991](#)) located within 12 school districts across two Midwestern states, in which learners make sense of phenomena or solve problems collaboratively by making use of the three dimensions of the NGSS. In the second year, the study involved 12 teachers and their classes () and the third year, 21

Table 4 Robust community perspective rubric.

2. Quality (more than one) - Mentions context of science work, as social practice, in class or society. - Indicates awareness of audience in communicating response. - Reflects that science endeavor impacts the world.	1. Moving toward (includes one) - Mentions context of science work, as social practice, in class or society. - Indicates awareness of audience in communicating response. - Reflects that science endeavor impacts the world.	0. Does not demonstrate community perspective - Missing mention of context of science work, as social context in class or society. - Missing indication of awareness of audience in response. - Missing reflection of science endeavor having impact in the world.
---	--	---

teachers and their classes. Our study presented here took part in the second and third year of the Multiple Literacies in Project-based Learning Project (ML-PBL; Krajcik et al., 2015; Miller and Krajcik, 2019) in which we designed, enacted and revised third, fourth and fifth grade project-based resources for meeting the NGSS performance expectations as well as Common Core State Standards in English Language Arts and Mathematics. Students engage in finding answers to driving questions by *collaboratively* building artifacts, which are meaningful to the classroom community, and often centered on authentic problems. Driving questions are designed to promote student engagement and student learning throughout the project. Moreover, driving questions are broad enough to allow learners to ask sub-questions and pursue solutions to those questions throughout the project. For the research reported in this paper, our work takes place in 4 of the third grade classrooms in which teachers are enacting ML-PBL units.

Participation

In this paper, we report on data from one-third grade unit: *Why do I see so many squirrels, but I can't find any stegosauruses?* The Squirrel unit focuses on adaptation, diversity, survival, fossils and extinction. In the squirrel unit, the students first observe squirrels and make a model of how squirrels survive in their environment, slowly adding ideas about specialized structures and ecological systems. Then they examine fossils and consider how fossils, specifically fossils of prehistoric rodent-like organisms provide clues about the structures and ecological systems of the Jurassic period. Finally, the students collaborate on a presentation (drawing, model, story or play) that explains how the early rodent might have survived when the stegosaurus died out.

Sample

The samples of this study come from three classrooms in years 2 and 3 of the 5-year curriculum development project. In two of the classes, the teachers were second year enactors of ML-PBL and in the third classroom, the teacher was in first year enactment. In the first classroom (second year enactment, urban), 100% of students are economically disadvantaged, and the majority of the students speak a language other than English in the home. The other two classes are situated in rural schools with a majority of white students and high poverty (<75% economically disadvantaged). The assessments in both years were given with a scripted slide presentation to reduce any confusion the students might encounter with directions, with English language and/or wording.

Data sources

Cognitive interviews. Pre and post-unit assessment

The data for this study are derived from a pre- and post-unit assessment of the Grade 3 ML-PBL Units: *Why do I see so many squirrels, but I can't find any stegosauruses?* To assure usability of the tasks, we performed two sets of semi-structured cognitive task interviews with focus groups (Ginsburg, 1997). We videotaped and transcribed eight groups of two students each, working through the tasks collaboratively. Lave describes learning as “change in participatory practice” (1991). As “individual” learning cannot be understood apart from larger social and historical influences (Gutiérrez et al., 1999), and since we hope to gain access into changes to the dialogic processes of sensemaking, we sought to understand how the students use their meaning-making resources in context, and over time change their understanding of participatory practice. Thus, semi-structured interviews provide access to the students' various resources and socially developed constructs of practices for figuring out (Barron, 2003; Carlone and Johnson, 2007) before participating in the units, and then as they become (re) negotiated over time.

The students were selected by the teacher to represent a range of science, math and language interests and proficiencies. Four of the sixteen students were English language learners. The students made sense of the tasks and orally explained their thinking, providing in-the-moment access to their thinking processes (Webb, 1982). The task was designed to indicate the process that underlies the dynamic shared learning (a shared-knowledge product) of two subjects working together to solve a problem. We transcribed the interviews to understand the learning of the two individuals working together and how their shared knowledge influences how they make sense of the task (Chi, 1997). We used these findings to revise the tasks. Table 5 presents sample questions from the semi-structured cognitive task interviews.

Finally, we tested the tasks pre- and post-unit enactment with three classrooms with approximately 18 students in each classroom. Immediately after the task, we administered four to six semi-structured interviews about the assessment task and how it felt to take the assessment. Table 6 presents several questions that were asked of students after completing the post task.

Table 5 Example of the cognitive task interview.

Q1	Please rephrase the prompt or task in your own words? Do you have any questions?
Q2	Are there any parts of this section that you find confusing? Do you have any questions?
Q3	This task asks you to _____. What do you think this means you need to do?

Table 6 Example of the semi-structured task interview.

Q1	How did that (assessment) go for you?
Q2	What did you find challenging? What did you find interesting?
Q3	Were you able to show what you know?

Data analysis

Analysis of student data

We rely on mixed methods (Creswell and Clark, 2017) emphasizing both qualitative understandings of how students use talk to organize their thinking (Silverman, 2006), and a quantitative analysis of artifacts for depth, flexibility and use of community resources. We analyzed the assessments of 60 students for depth, flexibility, and use of community resources. We examined internal and shared processes (collaborative knowledge building) through discourse, analysis of semi-structured interviews of focus groups, artifact collection, and the analysis of the assessment.

Example of student response on the squirrel post-unit assessment focus group

After introducing students to and providing information about the phenomenon (2 species of squirrels that live in different areas), and prompting students to think about similarities and differences as well as how the squirrels meet their needs, we asked the question (as in the example above): “If someone cuts down all the trees, what could happen to the Eastern Gray Squirrels who live in this area? Use evidence from your experiences, the chart, and the text.”

Focus Group students’ written responses: *The squirrel will not be able to hide and the hawks will eat them. No more squirrel will survive here. Squirrel have to have trees to hide and survive. Some squirrels will move to a new forest and live there.*

The task continued with the following prompt: “You shared one idea of what might happen. Now share of a very different idea of what could happen to the population of squirrels? Use evidence in your answer.”

Focus Group students’ written responses: *Some of the squirrels will hide in the chopped down trees and they bury the nuts and those seeds will grow into trees. The squirrels might survive because they can adapt. Some squirrels live in (student’s) attic and some can live in the rain gutter.*

Response analysis

Context

We analyzed pre- and post-unit assessments from classrooms in three school districts. We did not however help the students read the chart provided in the first part of the assessment (not pictured here). Students were given the chance to ask clarifying questions about wording, directions, and the teachers were provided instruction as to how to solicit and answer all questions equally across settings.

Evidence that students have developed knowledge-in-use of the enduring understanding is demonstrated by: *depth* of useable knowledge; *flexibility* in applying an array of DCIs, SEPs, and CCCs to make sense of phenomena; and a robust *community perspective*, in which students make explicit connections to their local environment or social community.

Depth of knowledge as evidence of knowledge in use

Based on our rubric, across the pre- and post-assessment artifacts examined for this study, the overwhelming majority of students developed in depth of knowledge as shown in the application of one Practice, CCC and element of the DCI to address the question. As indicated in Table 7, students across the three classrooms demonstrated growth in depth of knowledge.

Flexible thinking as evidence of knowledge in use

We found that students “tried on” various approaches to solving the problem or making sense of the phenomenon and that these “trying on” approaches were often, but not always, correlated with depth of knowledge. Students were able to write *different possible ideas* as a demonstration of their understanding of the affordances that different SEPs, DCIs or CCCs offer the sensemaking process, or in other words that students showed evidence of flexible thinking. Based on the focus group, cognitive interviews and field tests, we saw an increase in flexible thinking after instruction in all classrooms, but more in some than others. In the example above, the student who first suggested that the squirrels would not survive describes another possible scenario in which the squirrels plant nuts that grow into trees or that the squirrels adapt to the new environment.

In Fig. 5, we see an example of flexible thinking across questions 1 and 2 for the **same phenomenon**. The student used ideas of the environment providing needs for the animal to survive, and then approached the question again using ideas of adaptation. We see two DCI components used as affordances for sensemaking (*LS2.C Ecosystem Dynamics* and *LS4.C Adaptation*); and two different CCC’s (*structure and function* and *cause and effect* respectively) and two practices. The ideas show depth of thinking for one DCI (*LS4.D Biodiversity*), but do not use correct science ideas for a second DCI (*LS4.C Adaptation*). However, the understanding of affordance of the DCI *LS4.C Adaptation* does not correspond with evidence of depth of that idea. We noticed across the artifacts that flexibility did not always correspond with depth of DCI component, CCC or practice use.

Table 7 Growth in depth of student knowledge.

Site 1	Growth in depth 89%
Site 2	Growth in depth 88.8%
Site 3	Growth in depth 94.7%

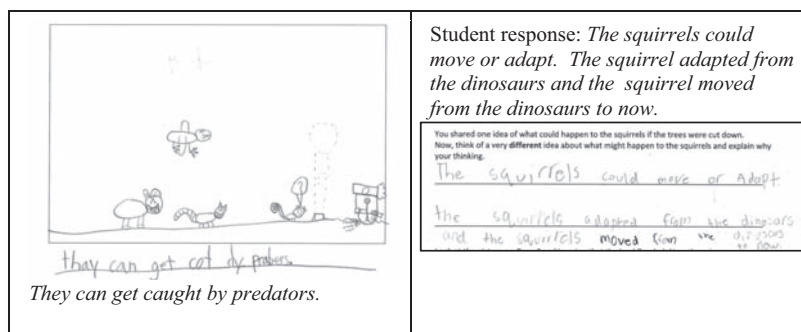


Fig. 5 Example of flexible thinking across Q1 and 2 for the **same phenomenon**. The student used ideas of the environment providing needs for the animal to survive, and then approached the question using ideas of adaptation.

In **Fig. 6**, note an artifact from a different student from class 3, and a non-example of flexible thinking across Q1 and 2 and across the assessment tasks with two related phenomena. This student is considered to have not yet reached “flexible thinking” because s/he responded with the same DCI component (*ecosystem dynamics LS2.C*) for each of the tasks for the first phenomenon, including the prompt for flexible thinking. This is the same student responding to the new prompt related to the new (second) phenomenon (below). They use the same idea and CCC as they did for the first phenomenon. (In both cases, we prompt for the same practice.)

In the following example, **Fig. 7**, the student showed evidence of flexible thinking to make sense of the different but related phenomenon, the first where there is an immediate change to the environment, and the other where the squirrel survived in an environment with different features, over millions of years.

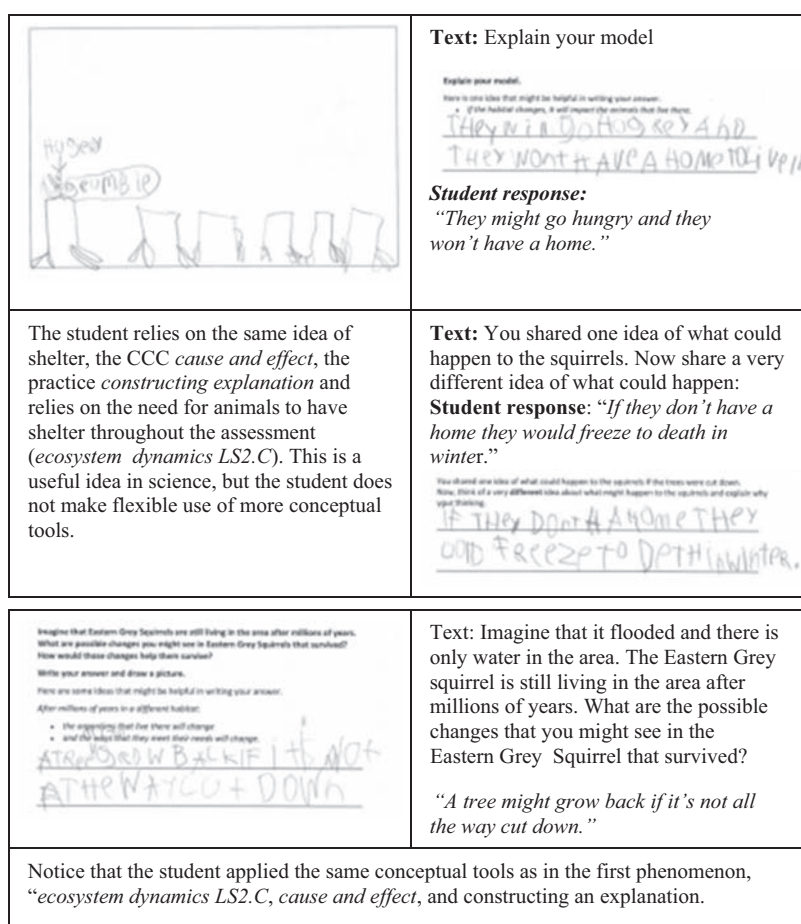


Fig. 6 Non-example of flexible thinking The same DCI component, CCC and practice (modeling is shown as prompted, but the student relied heavily on the explanation) is applied across the prompts.

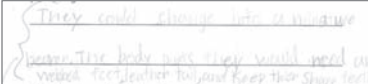
Flexible Performance Task 1. First Phenomenon.  Text Prompt: All the tree are cut down in the area. What might happen to the squirrel who live there? Student Response: Model: <i>The trees are cut down and there is no food. Explanation, "They need the nuts from the trees, the squirrels might die."</i>	Flexible Performance Task 2. Second Phenomenon. Text Prompt: Imagine that it flooded and there is only water in the area. The Eastern Grey squirrel is still living in the area after millions of years. What are the possible changes that you might see in the Eastern Grey Squirrel that survived?  Student response: <i>"They could change into a miniature beaver. The body parts they would need are webbed feet, leather tail, and keep their sharp teeth."</i>
---	---

Fig. 7 Example of flexible thinking across the flexible task performances (across the different phenomena). Use of different CCC (structure and function and cause and effect respectively) and DCI across the assessment items (ecosystem dynamics LS2.C and LS4.C Adaptation respectively).

Robust community perspective

While in the focus interviews, we did not prompt specifically for community perspective, some students brought forth evidence of this understanding (i.e., "We talked about this in class, when we looked at skulls. Animals fit the place they live in" and "It's good for people to have trees"), which initiated new prompts in our redesign. Community connections were less evident in a majority of focus group interviews. As we refine these assessments, we will elicit these connections more explicitly.

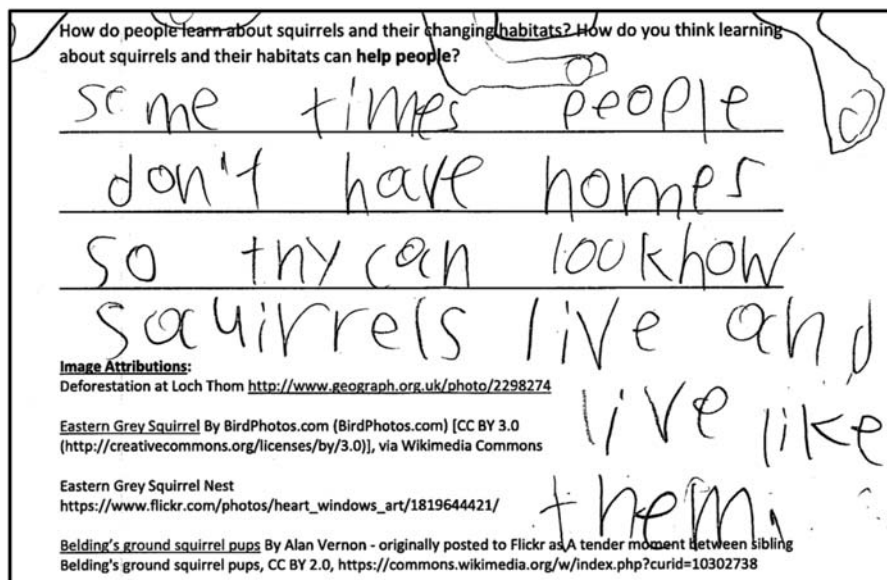
Fig. 8 shows one example of a student's response, which demonstrates full affordance of community perspective.

This example shows two items in our rubric: Mentions context of science work, as social practice, in class or society; Reflects that science endeavor impacts the world. Overall we saw that students gained in Robust Community Perspective across the classes, but they developed this criteria for Knowledge -in-Use differently.

Table 8 shows pre- and post-assessment scores on community perspective at the three sites.

Major findings

Our analysis shows that our flexible task performances can elicit depth of knowledge as evidenced by full application and flexible use of elements of the SEPs, DCIs and CCCs, to make sense of complex phenomena. In addition, they can elicit change in the three domains of depth of knowledge, flexibility in use of ideas and robust community perspective.



How do people learn about squirrels and their changing habitats? How do you think learning about squirrels and their habitats can help people?

some times people
don't have homes
so they can look how
squirrels live and
live like
them.

Image Attributions:
Deforestation at Loch Thom <http://www.geograph.org.uk/photo/2298274>
Eastern Grey Squirrel By BirdPhotos.com (BirdPhotos.com) [CC BY 3.0 (<http://creativecommons.org/licenses/by/3.0/>)], via Wikimedia Commons
Eastern Grey Squirrel Nest https://www.flickr.com/photos/heart_windows_art/1819644421/
Belding's ground squirrel pups By Alan Vernon - originally posted to Flickr as A tender moment between sibling Belding's ground squirrel pups, CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=10302738>

Fig. 8 An example of full affordance of community perspective.

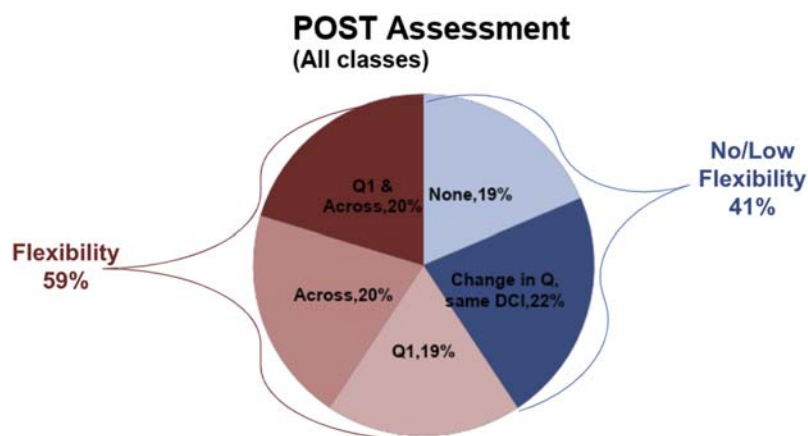
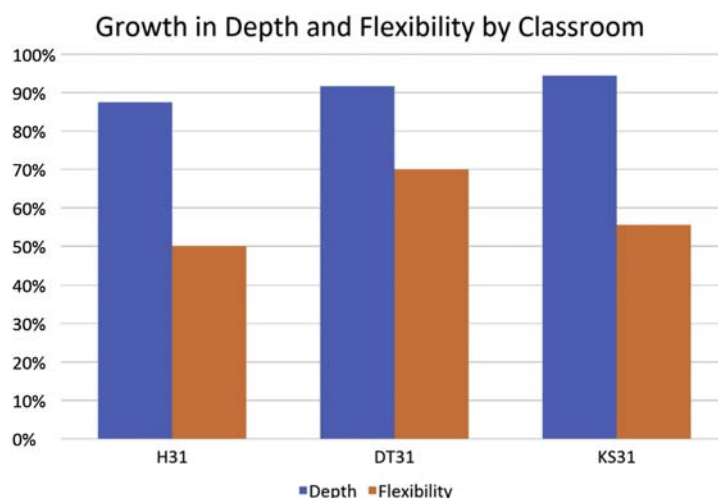
Table 8 Pre and post scores on community perspective in the three sites.

Site 1	Evidence of community perspective pre 20%/post 100%
Site 2	Evidence of community perspective pre 31.8%/post 88%
Site 3	Evidence of community perspective pre 37%/post 87.5%

Flexible thinking growth data across three classes show there were students in each class who showed evidence of flexible thinking. Across the three classrooms, 59% showed some evidence of flexible thinking (Figs. 9 and 10). We also see that in one classroom, a small majority of students responded to the prompts with flexibility and one in which there was a minority of students who responded to the prompts with flexible thinking.

As seen in Fig. 9, in each classroom the majority of students developed depth of sophistication of ideas. In terms of flexibility, the amount of growth seems dependent on in class contextual factors, (i.e., such as teaching style or classroom makeup.) Most of the growth is seen as occurring in **sophistication of ideas** across the three classes.

In School 1, the majority of the students did not grow in Flexible thinking, but grew in sophistication of ideas and practices. In school 2, in contrast, the majority of the students developed in depth as well as in flexibility, and in school 3, the same number of students developed knowledge-in-use as those who did not. These data suggest that there are contextual factors that influence this growth in knowledge-in-use, and that it is not necessarily correlated to growth in sophistication of ideas.

**Fig. 9** Pie chart showing percentage of flexible thinking $n = 59$.**Fig. 10** Learners' depth of knowledge and flexible use of affordances increased over time.

Discussion and significance

Assessment of “what we can measure” informs instruction and teacher noticing (Shepard et al., 2017). We keep this in mind as we develop flexible performance tasks designed to socially meaningful context-rich project-based learning environments and to the NGSS. We integrate ideas from the practice world of scientists, with the context of PBL, and the knowledge-in-use perspective to inform teachers about what evidence of students’ learning we care about (Biesta, 2015)—depth, flexibility and community perspective. The knowledge-in-use perspective provides new opportunities to support students in building capacity to approach unfamiliar problems and complex and multi-layered phenomena using a developing set of ideas that allow them to provide explanatory possibilities (Weick, 1995). The expansive approach to assessment highlights students’ potential, particularly those with diverse intellectual and linguistic resources (see, Miller and Reigh, 2020).

However, assessing knowledge-in-use is challenging. Building on previous work (Miller et al., 2018), we develop a systematic process to the design of performance tasks that can provide evidence of depth of knowledge, flexible thinking and community perspective. These flexible performance tasks were tested in elementary classrooms using project-based materials designed to meet the performance expectation of the NGSS and the intent of the Framework for K-12 Science Teaching. Our findings demonstrate that the flexible performance tasks can elicit depth of student knowledge, flexibility in thinking, and community perspective when administered in PBL classrooms. However, based on our findings, questions emerged such as why some classrooms fostered more growth than others in knowledge-in-use, and which instructional practices and contexts promote greater development of knowledge in use. We wonder about the relationship between the richness of sensemaking discourse that happens in the classroom and the development of knowledge-in-use. Our hope is that the flexible performance tasks gather understanding of what students know, but also show an understanding that learning is inseparable from what students can perform in the PBL socially situated context with a multiplicity of perspectives and community-grounded motivators for the science work underway (Gee, 2010).

Acknowledgment

This work was funded by the Lucas Education Research, a division of the George Lucas Educational Foundation, APP# 13987. The ideas and opinions expressed in the paper are those of the authors and not the George Lucas Educational Foundation.

References

- Baines, A., DeBarger, A.H., De Vivo, K., Warner, N., 2017. Why is Social and Emotional Learning Essential to Project-Based Learning. LER Position Paper, 2.
- Barron, B., 2003. When smart groups fail. *J. Learn. Sci.* 12 (3), 307–359.
- Berland, L.K., Hammer, D., 2012. Framing for scientific argumentation. *J. Res. Sci. Teach.* 49 (1), 68–94.
- Biesta, G., 2015. What is education for? On good education, teacher judgment, and educational professionalism. *Eur. J. Educ.* 50 (1), 75–87.
- Blumenfeld, P.C., Soloway, E., Marx, R.W., Krajcik, J.S., Guzdial, M., Palincsar, A., 1991. Motivating project-based learning: sustaining the doing, supporting the learning. *Educ. Psychol.* 26 (3–4), 369–398.
- Carlone, H.B., Johnson, A., 2007. Understanding the science experiences of successful women of color: science identity as an analytic lens. *J. Res. Sci. Teach.* 44 (8), 1187–1218.
- Choi, K., Lee, H., Shin, N., Kim, S.W., Krajcik, J., 2011. Re-conceptualization of scientific literacy in South Korea for the 21st century. *J. Res. Sci. Teach.* 48 (6), 670–697.
- Creswell, J.W., Clark, V.L.P., 2017. *Designing and Conducting Mixed Methods Research*. Sage Publications, Los Angeles, CA.
- Design-Based Research Collective, 2003. Design-based research: an emerging paradigm for educational inquiry. *Educ. Res.* 32 (1), 5–8.
- Erickson, H.L. (Ed.), 2006. *Concept-Based Curriculum and Instruction for the Thinking Classroom*. Corwin Press, Thousand Oaks, California, USA.
- Evans, J., 1999. Building bridges: reflections on the problem of transfer of learning in mathematics. *Educ. Stud. Math.* 39 (1), 23–44.
- Finnish National Board of Education, 2016. *National Core Curriculum for General Upper Secondary Schools 2015*. Finnish National Board of Education (FNBE), Porvoo Kirjakeskus, Helsinki. http://www.oph.fi/saadokset_ja_ohjeet/opetusuunnitelmien_ja_tutkintojen_perusteet/lukiokoulutus/ops2016/103/0/lukion_opetusuunnitelman_perusteet_2015.
- Ford, M.J., Wargo, B.M., 2012. Dialogic framing of scientific content for conceptual and epistemic understanding. *Sci. Educ.* 96 (3), 369–391.
- Gee, J.P., 2010. A situated-sociocultural approach to literacy and technology. In: *The New Literacies: Multiple Perspectives on Research and Practice*. Guilford Publications, New York, NY, pp. 165–193.
- Ginsburg, H., 1997. *Entering the Child’s Mind: The Clinical Interview in Psychological Research and Practice*. Cambridge University Press, Cambridge, MA.
- Gutiérrez, K.D., Baquedano-López, P., Tejeda, C., 1999. Rethinking diversity: hybridity and hybrid language practices in the third space. *Mind Cult. Activ.* 6 (4), 286–303.
- Harris, C.J., Krajcik, J.S., Pellegrino, J.W., McElhane, K.W., 2016. *Constructing Assessment Tasks that Blend Disciplinary Core Ideas, Crosscutting Concepts, and Science Practices for Classroom Formative Applications*. SRI International, Menlo Park, CA.
- Krajcik, J.S., Palincsar, A., Miller, E., 2015. *Multiple Literacies in Project-Based Learning*. Lucas Education Research, a division of the George Lucas Educational Foundation, Michigan State University, East Lansing, MI.
- Kulgemeyer, C., Schecker, H., 2014. Research on educational standards in German Science Education—towards a model of student competences. *Eurasia J. Math. Sci. Technol. Educ.* 10 (4), 257–269.
- Ladson-Billings, G., 2008. Yes, but how do we do it? Practicing culturally relevant pedagogy. In: Landsman, J., Lewis, C.W. (Eds.), *City Kids, City Schools: More Reports From the Front Row, White Teachers/Diverse Classrooms: A Guide to Building Inclusive Schools, Promoting High Expectations and Eliminating Racism*. Stylus Publishers, Sterling, VA, pp. 29–42, pp. 162–177.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, MA.
- Manz, E., 2012. Understanding the codevelopment of modeling practice and ecological knowledge. *Sci. Educ.* 96 (6), 1071–1105.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, second ed. Sage, Thousand Oaks, CA.
- Miller, E., Janusz, R., 2014. The NGSS case studies: all standards, all students. *Sci. Child.* 51 (5), 10.
- Miller, E.C., Krajcik, J.S., 2019. Promoting deep learning through project-based learning: a design problem. *Discip. Interdiscip. Sci. Educ. Res.* 1 (1), 1–10.
- Miller, E., Reigh, E., 2020. Interactional Approach to Language: Supporting All Students to Engage in Equitable Sensemaking. *International Society of Learning Sciences (ICLS)*.

- Miller, E., Codere, S., Krajcik, J., 2018. Developing Assessment Tasks to Promote Student Sensemaking of Phenomena and Flexible Thinking. International Society of the Learning Sciences, Inc. [ISLS].
- Miller, E.C., Severance, S., Krajcik, J., 2021. Motivating teaching, sustaining change in practice: design principles for teacher learning in project-based learning contexts. *J. Sci. Teach. Educ.* 32 (7), 757–779.
- Mislevy, R.J., Haertel, G.D., 2006. Implications of evidence-centered design for educational testing. *Educ. Meas.* 25 (4), 6–20.
- Moschkovich, J.N., 2007. Beyond words to mathematical content: assessing English learners in the mathematics classroom. *Assess. Math. Prof.* 53, 345–352.
- National Research Council, 2012. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. National Academies Press, Washington, DC.
- NGSS Lead States, 2013. Next Generation Science Standards: For States, by States. National Academies Press, Washington, DC.
- OECD, 2016. PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy. OECD Publishing, Paris.
- Pacheco, M., 2010. Performativity in the bilingual classroom: the plight of English learners in the current reform context. *Anthropol. Educ. Q.* 41 (1), 75–93.
- Pellegrino, J.W., Hilton, M.L., 2012. In: National Research Council (Ed.), Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century. National Academies Press.
- Ryu, S., Sandoval, W.A., 2012. Improvements to elementary children's epistemic understanding from sustained argumentation. *Sci. Educ.* 96 (3), 488–526.
- Shepard, L.A., Penuel, W.R., Davidson, K.L., 2017. Design principles for new systems of assessment. *Phi Delta Kappan* 98 (6), 47–52.
- Silverman, D., 2006. Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction. Sage, New York, NY.
- Spiro, R.J., Feltovich, P.J., Jacobson, M.J., Coulson, R.L., 1991. Knowledge representation, content specification, and the development of skill in situation-specific knowledge assembly: some constructivist issues as they relate to cognitive flexibility theory and hypertext. *Educ. Technol.* 31 (9), 22–25.
- Webb, N.M., 1982. Student interaction and learning in small groups. *Rev. Educ. Res.* 52 (3), 421–445.
- Weick, K.E., 1995. Sensemaking in Organizations, vol. 3. Sage, New York, NY.
- Wiggins, G.P., Wiggins, G., McTighe, J., 2005. Understanding by Design. ASCD, Alexandria, VA.
- Zeidler, D.L., Sadler, T.D., Simmons, M.L., Howes, E.V., 2005. Beyond STS: a research-based framework for socioscientific issues education. *Sci. Educ.* 89 (3), 357–377.

Neurocognitive and physiological measurement of STEM learning processes

Thomas Delahunty, Department of Education, Maynooth University, Maynooth, Co. Kildare, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction: a primer on the necessity for neuro and physiological tools	458
Applications of neurocognitive methods to educational research in STEM	458
Research approaches using electroencephalography	459
Investigating cognitive strategies: the case of arithmetic	459
Exploring individual/group differences in STEM: the case of spatial reasoning	460
Uncovering causal explanations in STEM: the case of problem conceptualization	462
Functional imaging approaches in STEM educational research	462
Investigating cognitive strategies: mathematical problem solving	463
Exploring individual differences and causality: the case of scientific innovative problem solving	464
Concluding remarks on fMRI for STEM education research	465
Other physiological measures of learning processes	465
Physiology of emotion: heart rate and electrodermal approaches	465
Observing physiological differences in expertise: eye-tracking approaches	467
Concluding remarks on physiological approaches in STEM educational research	468
Broader considerations for application	468
Conclusion	469
References	469
Further reading	471

Introduction: a primer on the necessity for neuro and physiological tools

Characterizing the process of learning in STEM education is an extremely difficult undertaking and this is partly due to the myriad of knowledge types and conceptions involved in defining the field itself. Acknowledging the complexity, which characterizing learning in STEM education entails, is important to considerations of applying neurological and physiological methods in this field of research. A defining characteristic of learning in STEM education is the focus on developing capability, which is more concerned with knowledge utility, application and adaptability than acquisition (Kimbell and Stables, 2008). This focus has given rise to the notion of STEM learning as a complex adaptive system (c.f. Stephens and Richey, 2011). These systems result in emergent behavior occurring in a non-linear fashion, involving interaction between agents resulting in more complex learning than could be achieved in linear individualistic fashions (Jacobson et al., 2019).

Pertinent to this review is the concept of nested systems among CAS (Jacobson et al., 2016). A CAS can be treated as various complex systems within one another, such as the classroom environment, a group of students in a project-based learning scenario, the individual student within this group, and the individual brain of this student. If we accept this analogy as appropriate to characterizing learning in STEM, then how exactly do we achieve a complete sketch of the process of learning? For example, problem-based learning is an oft cited pedagogical approach used in many STEM contexts and across all levels of educational provision (Williams et al., 2008; Newhouse, 2016; Delahunty, 2019). Problem solving is a complex higher-level set of cognitive processes (Stillings et al., 1995) and has a significant history in psychology and educational research. Of important note stemming from this history is that the predominant research focus, in the area of problem solving in education, relates to the overt behaviors and heuristics in the process of solving problems (Ohlsson, 2012; Jakel and Schreiber, 2013). As delineated by Ohlsson (2012), this is likely due to the dominance that behaviorism exerted on educational and psychological research. An important dearth in our understanding that arises from this focus on heuristics is the tacit cognitive phenomena that underpin problem solving that are not directly observable in overt behavior or performance. To do this we need an enhanced repertoire of research methods that allow us to observe the “unobservable”.

Applications of neurocognitive methods to educational research in STEM

There are several considerations that must be addressed in the application of any neurocognitive method to research in education. Of course, it goes without saying that ethical considerations must be to the forefront of the researcher's mind. It is important to briefly delineate some core principles that must be considered by the researcher in contemplating the application of neurocognitive or physiological measures to their research. Some of these have been primary rallied in criticisms of the field of educational neuroscience in general. For example, neuromyths are often utilized to exemplify the miscommunication or misapplication between

neuroscience and education. These are defined as misunderstandings about the brain and education held among teachers or the wider public which do not stand up to scrutiny (Gardner, 2019). Another criticism that has been aimed at the role of neurocognitive data in educational contexts has to do with the “seductive allure” of neuroscience explanations (Weisberg et al., 2008). This broadly relates to the phenomenon that evidence is more likely to be believed if the presentation of it is accompanied by complex brain images. Both of these examples highlight the issue of framing the research itself and appropriately weighting the methods within the study.

Much of the framing issue with regards to research using neurocognitive methods again comes back to how learning is conceptualized as a set of processes. Howard-Jones (2011) reminds us that neuroscience and education both have distinct heritages and philosophies such as the positivist images of the controlled experiments in the neuroscience lab versus the social complexities of constructivist learning in the classroom. What is critical to remember here is that neither perspective alone can explain the totality of any learning mechanism or process and what is required is an interdisciplinary mindset to overcome the growing pains of the discipline of educational neuroscience (Thomas et al., 2019). Pragmatically, this means that the STEM education researcher does not start with the intention of utilizing a neurocognitive method but instead critically frames their research, using appropriate theory and evidence, to focus on the aspect of the learning process that they wish to investigate. Once the research questions are posed it can then be decided if the appropriate neurocognitive method is required.

In what follows I review a series of studies using appropriate alignment of neurocognitive methods for STEM educational research that sought to investigate STEM-related learning processes. It would be impossible to review a large series of studies in the space afforded here so I have tried to focus on studies that have carefully selected a neurocognitive tool with the aim of elucidating a STEM learning process typically unattainable (or attainable with great difficulty and potential bias) by direct observation. Aligning with calls for interdisciplinarity between neuroscience and education (e.g. Thomas et al., 2019; Howard-Jones, 2011), my intention is to demonstrate how a neurocognitive or physiological method can augment the typical repertoire of educational research methods. Again, given the limited space afforded in this article, I will focus on the two most popular and utilised approaches; EEG and fMRI.

Research approaches using electroencephalography

Electroencephalography (EEG) is a neuroscientific tool that measures electrical activity, generated from the cerebral cortex, through a series of sensors on the human scalp (Rowan and Tolunsky, 2003). It is a method that offers a high degree of temporal resolution allowing the accurate detection of the onset of significant behavioral events in the neurological signal recorded (Delahunty et al., 2018) often within milliseconds of the occurrence (Banich and Compton, 2011). There are a variety of EEG devices available ranging from the high-density versions associated with laboratory settings to the more recent evolution of portable wireless EEG devices (See Fig. 1).

Clearly the high-density devices have higher spatial accuracy given the increased number of sensors available to the researcher however, they are restricted to the laboratory setting.

Investigating cognitive strategies: the case of arithmetic

Core to the study of problem-solving in STEM education research is the desire to uncover evidence related to the cognitive strategies that underpin students’ performances. As discussed in other research (e.g. Delahunty, 2019; Delahunty et al., 2018), typical educational and psychological methods rely almost entirely on indirect evidence to determine evidence of these various strategies. The issue with using indirect evidence such as visual-verbal protocol or analysis of performance data is that the underlying cognitive mechanisms that led to the visible behavior or external representation remain unknown, or at the very least inferred with large potential for error. Seeking direct evidence of the cognitive bases of these problem-solving strategies is precisely where a tool such as EEG can add much value to a research project.

Hinault and Lemaire (2016) cite the case of arithmetic strategies in mathematical problem solving as an area where EEG can unpack some of the hidden underlying cognition that underpins students’ arithmetic. In particular, arithmetic reasoning episodes have short latencies and are often tacit, meaning that students will not be able to report on their approaches accurately, for example



Fig. 1 High density lab-based EEG (Left) and portable wireless EEG headset (right).

in a think aloud method (Thevenot et al., 2010). There are several arithmetic strategies commonly reported in the literature. Two common strategies are procedural (e.g. decomposing problems into smaller facts, $8 + 13$ decomposed into $8 + 3 + 10$) or retrieval (e.g. 40 for 10×4) (Hinault and Lemaire, 2016). These examples should emphasize how difficult it is to capture direct evidence of the cognitive processing occurring during such a short problem epoch.

In a study with 25 university students, Grabner and De Smedt (2011) used EEG in conjunction with verbal reports to explore the cognitive bases of arithmetic reasoning. They found that reports of the retrieval strategy were accompanied by theta band synchronization over the left parietooccipital areas of the cortex (a cortical area implicated in mathematical reasoning in adults (Cf. Wendelken, 2015)) compared to procedural strategies, which demonstrated higher levels of alpha band desynchronization in the right parietooccipital areas (ibid). Theta band frequencies are often associated with retrieval from long-term memory (Klimesch, 1999) whereas alpha dyssynchronization or decreases can reflect cognitive activation and increased cognitive load (Neubauer et al., 2006). This study by Grabner and De Smedt (2011) demonstrates that the arithmetic strategies involve distinct cognitive networks in the cortex and support the hypothesis that retrieval strategies are less cognitively demanding than procedural strategies.

Building on this evidence, research can then move on to look at the cognitive demands and strategy elicitation of task types. Take for example the idea that a teacher may want to engage students in the procedural strategy of arithmetic problem solving rather than relying on retrieval, which although cognitively less demanding, may prevent the progression of students to solving more complex tasks. How then do they know that the textbook problems will do this? This is where a research study with the neurocognitive lens of EEG may be useful. De Smedt et al. (2009) conducted such an experimental study, manipulating arithmetic problem size with a sample of university students. In this work they found that the larger problems, being logically more demanding, elicited procedural strategy use shown by desynchronization of the alpha band in bi-lateral parietooccipital areas. Zarjam et al. (2013) conducted a similar study and found increased levels of desynchronization of alpha and delta frequency bands over the frontal lobes (known seat of working memory) for the larger and more complex arithmetic problems. This desynchronization, being an indice of cognitive activation and load in the alpha band, highlights a greater demand being placed on the working memory centers. Moreover, the discernible patterns of activity and engagement dependent on the type of task, allows the researcher to evaluate these designs with high accuracy and ultimately draw out the implications for educational practices.

Lastly, moving to a longitudinal analysis of how learning and strategy use in arithmetic interact over time, Soltanlou et al. (2019) recorded EEG while children solved a series of multiplication tasks. They found that over time the students transitioned from a procedural strategy to a retrieval strategy, indicated by progressively increasing levels of theta power in successive trials. Theta band power and desynchronisation correspond to the retrieval of content from semantic memory networks (Chiang et al., 2016). Therefore we see that as students become more proficient in a particular element of mathematics, they rely less on effortful procedural cognition and more on retrieval strategies. Taken together, these studies highlight that it is possible to excise neural data on both cognitive load and strategy use with the EEG method in response to problem task design parameters. This has both utility for the STEM education researcher and practical application in the learning sciences for the design of pedagogical approaches and task materials.

Exploring individual/group differences in STEM: the case of spatial reasoning

The study of individual differences in reasoning and problem solving has led to much of our understanding as to how individuals approach their learning and how, as educators, we should direct our pedagogy to alleviate potential barriers to student progression. Perhaps the best known case of group differences in STEM education is the consistently reported sex differences in spatial cognition favoring males, particularly in mental rotation (e.g. Sorby, 2009; Casey, 2013; Halpern and Collaer, 2005). I will not belabor the relationship of spatial talent to STEM success, but these individual sex differences do help in understanding some of why female representation in STEM is generally poor. Now to the complicated part, why do females perform worse in mental rotation (MR)? There are many possible reasons for this ranging from developmental differences in the sociocultural upbringings of males and females (Sorby, 1999) to biological differences in working memory capacity (Kaufman, 2007).

Another well founded difference in males and females is in strategy use. Numerous studies report a general phenomenon of males using holistic strategies whereas females use more analytical approaches during mental rotation problems (e.g. Peña et al., 2008; Heil and Jansen-Osmann, 2008). While it is tempting to assume a neat divide in strategy preference according to gender grouping, the research on strategy types is more varied. In a study with 1695 second level students completing a MR test, Geiser et al. (2006) implemented a latent class analysis to determine 5 potential classes of strategies that were evident ranging from holistic to analytical categories. In a more recent study, Hegarty (2018) utilized the same MR tasks as Geiser and colleagues with a group of undergraduate students and collected concurrent verbal protocol data. During analysis, Hegarty (2018) found that most participants, regardless of gender, utilized a blend of holistic and analytic strategies. This suggests that the effect of strategy use in explaining sex differences in MR is complex and again there are issues evident in the use of behavioral strategies in inferring the existence of strategy differences.

How does spatial reasoning map to EEG data?

The application of EEG as an alternative lens in this particular instance of STEM education has uncovered more nuances that further our understanding of this critical issue. In mapping spatial reasoning to its neurological bases, a number of key networks and frequencies have been shown to indicate elements of spatial cognitive processing in EEG data. In particular, the desynchronization of the alpha frequency in the parietooccipital regions of the cortex is thought to be indicative of visuospatial cognition (Cabeza and Nyberg, 2000; Rescher and Rapplesberger, 1999; Klimesch, 1997).

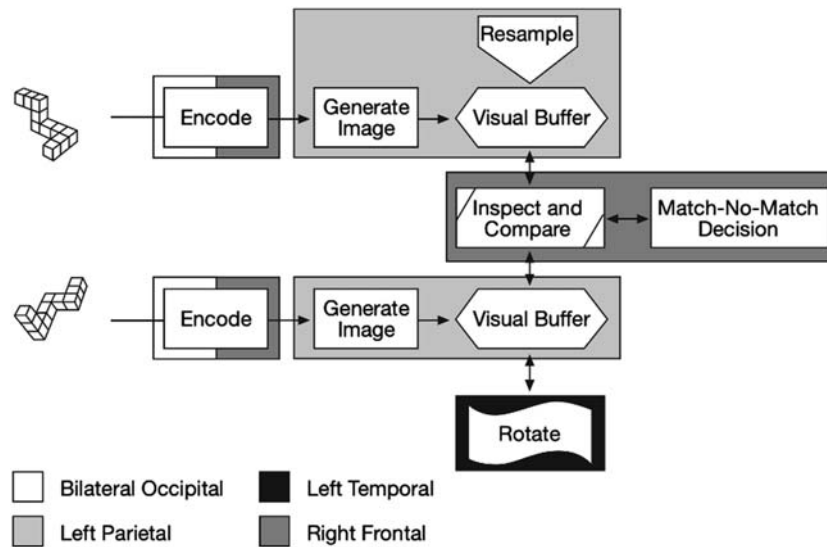


Fig. 2 Hypothesized model of mental rotation with cortical mappings. From Gill, H.S., O'Boyle, M.W., Hathaway, J., 1998. Cortical distribution of EEG activity for component processes during mental rotation. *Cortex* 34, 707–718.

Specific to MR, which is often the spatial reasoning process associated with the majority of sex differences (Casey, 2013), Gill et al. (1998) found that MR resulted in increased alpha desynchronization in the right frontal and left parietal areas. This points to both an active encoding, maintenance, manipulation and comparison of base stimuli with target stimuli (see Fig. 2) that is achieved through functional connection of working memory and visuospatial areas of the cortex (Gill et al., 1998; Delahunty et al., 2020; Klimesch, 1999). Increased levels and duration of alpha suppression are found with increased angular disparities in MR tasks (Michel et al., 1994).

Examining sex differences with EEG methods

Coupled with strategy use, as a potential mediator of sex differences in MR, some research has found that the modality in which tasks are presented moderates this effect further. Research by Parsons (2004) found that the sex difference completely disappeared when stimuli were presented in a true 3-D format (virtual environment) compared to a paper and pencil version (2-D representations of 3-D stimuli). So, this begs the question of whether the sex difference is truly confined to MR or abstractions of rotated stimuli, especially given the malleability of MR skills (c.f. Uttal et al., 2013)? Again, a neurocognitive lens may add further clarity. Rescher and Rapplesberger (1999) were among the first to begin looking at potential differences in EEG between males and females during a MR task. They found distinct patterns of cognitive activation in males and females with males displaying more coherence in alpha band in frontal and parietal areas and females displaying more overall brain activity, especially in the posterior areas. This suggests that females have to exert significantly higher levels of activation during MR compared to males but also that a different cortical network is implicated and differentiated by sex.

In a more recent study, Gootjes et al. (2008) used EEG with a group of young adults and found a significant effect of performance and connectivity of cortical regions. The authors found that males generally demonstrated higher functional connectivity in the temporal-parietal areas whereas females demonstrated more coherence in frontal areas. Sharma et al. (2018) further examined these differences as a function of task difficulty and utilizing the superior advantage of temporal resolution that EEG offers was able to distinguish sex differences in the sub-processes of MR. Based on the EEG data, Sharma et al. (2018) outlined the sub-cognitive processes of the visuospatial attention allocation network (VSAN), involved in perceptual encoding and stimuli representation, and the mental rotation network (MRN), involved in the spatial transformation of encoded stimuli. Results indicated that males demonstrate advantages over females in the VSAN early in the process of MR with superior bottom-up encoding, evident in parietal-temporal activity whereas females displayed a top-down strategy with parietal activation indicating they may have put more effort into integrating the complex information of the stimuli. In addition, males were found to recruit more right lateralized cortical sites whereas females indicated increased left hemispheric activity, indicating the potential adoption of holistic (right) versus analytical (left) approaches. Interestingly, this pattern has also been found when comparing those in STEM disciplines with humanities or social sciences (e.g. Li and Boyle, 2013). Consequently, it is apparent that different cognitive networks are implicated and this goes some way to explaining the causality of these sex differences in spatial cognition.

EEG allows the researcher to uncover further evidence of strategy use on the cognitive level, as well as potential cortical differences, which help in further understanding the etiology of these well-established group differences. Causality is often lamented as the ultimate aim in a researcher's pursuit of understanding individual differences, but is notoriously complex and difficult to ascertain. Hopefully it is becoming apparent how this neurocognitive tool can expose a layer of understanding in this domain of

education that has traditionally been somewhat unattainable. To further illustrate this potential, let's briefly turn to another useful application of this neurocognitive lens in exploring the issue of the etiological underpinnings of the relationship between spatial ability and STEM success (c.f. [Delahunty et al., 2016](#)).

Uncovering causal explanations in STEM: the case of problem conceptualization

While it is undoubtedly the case that spatial talent plays a key role in success in STEM education, there are no clear explanations as to why this is the case. While the causal relationship between spatial talent and success in STEM is likely multifaceted, a promising area that demonstrates some explanation is in the area of problem solving and particularly the early stages of the problem-solving process. A recent study by [Duffy et al. \(2020\)](#) presents correlational evidence for the role of spatial ability (measured by the Purdue Spatial Visualization Test: Rotations) and problem representation in word problem solving. This is an interesting finding with a sample of engineering students, and taken at a cursory level certainly seems promising in explaining some of this causal relationship we desire. Unfortunately the study suffers from theoretical and empirical limitations. This study would have benefited from a more critical consideration of the area of problem representation which the authors seem to conflate with the external representation only (via student solutions) despite the known phenomenon of internal-external interaction in the problem representation phase ([Delahunty et al., 2020](#); [Schnotz et al., 2010](#)). There is also a failure to control for confounding variables in this study such as reading ability which has actually been shown to be a more significant predictor of word problem solving than spatial cognition ([Kyttälä and Björn, 2014](#)), which is also supported by fMRI data ([Newman et al., 2011](#)). But in fairness to the authors, the internal evidence underpinning problem representation is a very difficult thing to directly observe.

In addressing this issue, and the general issue of problem-solving research in failing to address the early components of the process ([Ohlsson, 2012](#); [Barsalou, 2009](#)), the author's own work has introduced the notion of "problem conceptualization". Problem conceptualization can be operationally defined "as the period of cognitive processing occurring before the externalization of any representation (e.g. a sketch, mathematical expression, verbal communication etc.)" ([Delahunty et al., 2020](#), p. 399). Based on this definition, it is clearly a difficult undertaking to observe cognitive activity that occurs in this process to determine strategy use, that eventually leads to the externalized (or purely internalized) problem representation. In this work ([Delahunty, 2019](#); [Delahunty et al., 2015, 2018, 2020](#)), we have successfully utilized EEG for just this process.

For example, in [Delahunty et al. \(2020\)](#) EEG is applied as part of a mixed-methods study and also synthesized with video protocols of solutions and interviews with participants to explore exactly how word problems are conceptualized in the first seconds of a task. In this work, the EEG data yielded evidence of the activation of significant visuospatial cortical networks (i.e. frontal and parietal regions in theta and alpha frequencies) in the first few seconds of the problem-solving episode, prior to any externalized activity. This suggests that visuospatial cognition is involved within the internal aspects of problem conceptualization of word problems, prior to the representation phase of problem solving. This then, not only supports the importance of spatial cognition to problem representation but suggests that its role is operational in the immediate seconds of perception, prior to representation occurring. This again exemplifies the causal explanatory potential of the method but what is most noteworthy here is in the balancing of methods in our research. While in many cases the interviews that we have conducted about participants' solutions have aligned with the EEG data, there are cases when the EEG data revealed cognitive processes that were not recognized by participants. Synthesizing the quantitative cognitive EEG tool with other phenomenological approaches, such as interviews and stimulated recall episodes, connects the biological-cognitive with the sociocultural-sociohistorical, allowing a more comprehensive insight into causal relationships. For the reader interested in more on this "Cog-Phenom" approach see [Delahunty et al. \(2020, 2018\)](#).

Concluding remarks on EEG for STEM education research

As the previous sections have shown, there is a wide array of application for EEG in the investigation of learning related processes in STEM. EEG has advantages in temporal resolution being able to demonstrate specific reactions to stimuli or the cognitive activity underlying behavioral events. The advent of portable commercial headsets means that this technology may in time find further utility in more ecologically valid manners of research in STEM. There are however some limitations that are noteworthy. For example, the data processing is time consuming and complex and the collection of data is a very sensitive process susceptible to many artifacts. EEG is also quite limited in its spatial resolution and in order to gain more detailed and accurate insights into brain regions associated with learning, the researcher must consult functional imaging methods.

Functional imaging approaches in STEM educational research

This category of imaging methods utilizes the changes in blood flow and metabolic rate to determine areas of activation within the brain. The most commonly used strategy to achieve this is functional magnetic resonance imaging (fMRI) which assesses changes in the oxygenation of blood in the brain ([Banich and Compton, 2011](#)). The use of fMRI has a number of advantages including the ability to measure deep brain structures (see [Fig. 3](#)) such as activity in the hippocampus ([Ward, 2010](#)). Below I review and showcase a number of relevant studies aligning with the core areas of investigating strategies, exploring individual differences and aiming at causal explanations as I have with EEG in the previous section.

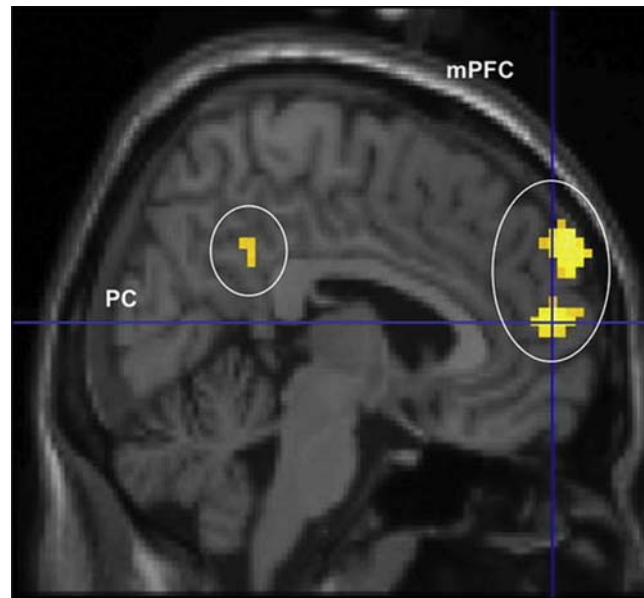


Fig. 3 fMRI scan showing activation in structure of the brain (“ENG111: Figure 9.7a” by Rosenfeld Media is licensed under CC BY 2.0).

Investigating cognitive strategies: mathematical problem solving

Picking up from where we left off with the topic of word problem solving in the last section, fMRI studies in the last few years have allowed us some further insight into the complexity of cognition involved. For example, [Newman et al. \(2010\)](#) implemented a study which sought to examine the neurological bases of differences of problem-solving strategy in word problems. They explicitly instructed two groups of university students to use either a spatial (e.g. imagine yourself moving backward and forward on the number line) or algebraic (e.g. representing the verbal information in equation format) approaches and varied task difficulty. They found greater overall activation in key areas such as the parietal and left pre-frontal cortices for the algebraic strategy, which tells us that this approach is more cognitively demanding than the spatial approach. Interestingly, and a key advantage of fMRI being the superior spatial resolution, they were able to compare individuals based on strategy preference and found that those with a preference for algebra had higher activations of the anterior cingulate cortex which is associated with cognitive control ([Mohanty et al., 2007](#)). This finding coupled with the superior working memory capacity of this group (discerned from a reading span test) supports the idea that the algebraic group had a higher working memory capacity and hence more cognitive control in the more demanding algebraic approach. This gives us some further insight into the relationship between strategy preferences and underlying cognitive architecture of individuals.

Whereas EEG data relies on the inference of cognitive functions based on frequency analysis and location, fMRI allows direct observation of spatially accurate cortical activation, including areas within the brain such as the limbic system. This has led to enhanced understanding of different elements of memory function in areas such as arithmetic problem solving. For example, [Grabner et al. \(2009\)](#) found increased activation of the left angular gyrus in adults who self-reported the use of retrieval strategies during arithmetic problem solving. While many studies up to that point had linked this area to the process of arithmetic problem solving (e.g. [Ischebeck et al., 2007](#)), the exact function remained elusive. In the study, [Grabner et al. \(2009\)](#) were able to isolate this region as unique to those who adopted retrieval (compared to procedural) strategies (see [Fig. 4](#)) providing robust evidence for this location in arithmetic fact retrieval and distinguishing retrieval from procedural strategies as involving less cognitive networks. This evidence supports the idea of increasing neural efficacy as a result of learning and again allows STEM educational researchers deeper insights to inform task or pedagogical design.

Lastly, as discussed in previous sections, access to the internal reasoning that occurs during the representation of word problems is notoriously difficult and has led to much debate. In exploring the differences in solving word versus number/equation tasks [Newman et al. \(2011\)](#) found that both strategies involve similar processes on the cognitive neurological level. These included activations of the posterior parietal cortex and interparietal sulcus, which have been associated with problem representation ([Loetscher et al., 2007](#); [Alfred et al., 2018](#)). However, [Newman et al. \(2011\)](#) demonstrated that the word problems elicited significant activation of language centers (Broca’s and Wernicke’s areas) whereas the number problems recruited more visuospatial locations (see [Fig. 5](#)). This nuanced detail provides further evidence of the underlying cognitive processes required for the successful solving of various problem types. This in turn informs the educational researcher and teacher in terms of pedagogical strategies to alleviate barriers to the successful engagement in problem solving by students (e.g. strategies targeted at semantic analysis of word problems vs. visual representations of numeric problems).

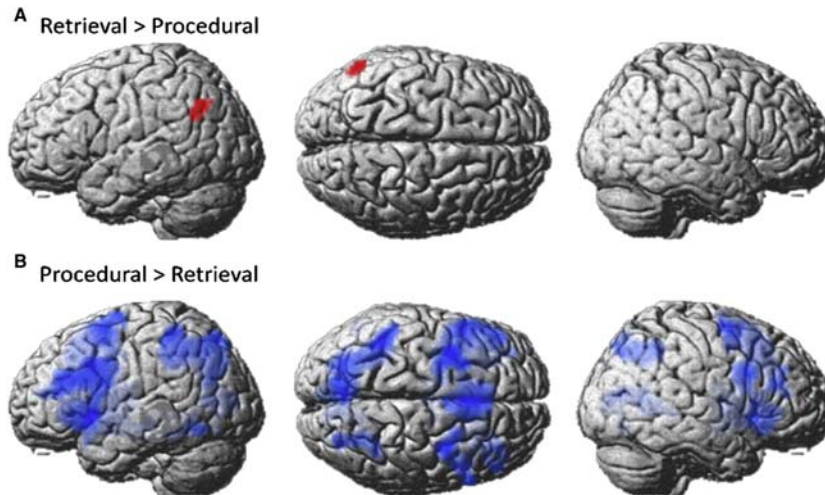


Fig. 4 Illustrating differential patterns of activation with retrieval strategies activating left angular gyrus. Data from Grabner, R.H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., Neuper, C., 2009. To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* 47(2), 604–608.

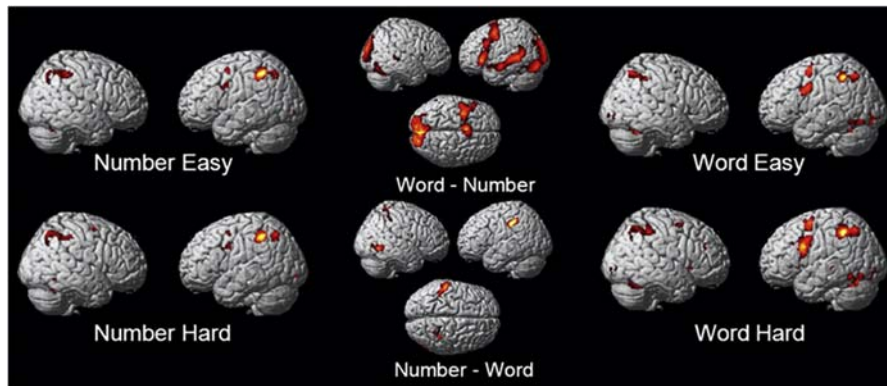


Fig. 5 Activation maps from Newman et al. (2011) depicted on the left are number problems and right word problems activations. Center are main effects of problem type.

Exploring individual differences and causality: the case of scientific innovative problem solving

Within the realm of science education, a core goal is to develop students inventive and innovative problem-solving capacities (Banks and Barlex, 2014). A core reasoning ability in this context is that of analogy or analogical reasoning (AR) (Krumnack et al., 2013; Brookman-Byrne et al., 2019). While AR has a role to play in many higher-level cognitive processes the actual process itself is complex and research methodologies typically employ regression type designs (e.g. Stamenković et al., 2019) where intelligence tests are used to predict performance or self-report measures such as interviews (e.g. Mozzer et al., 2013) where participants are asked to describe their approaches. Again, these represent a somewhat detracted set of evidence to observe individual differences. While there is considerable debate as to the true validity of intelligence tests with regards their relevance to applied activities (e.g. Richardson and Norgate, 2015) they have been extremely useful in illustrating the types of abilities that underpin success in AR.

In order to fully understand individual differences though, we need to have a look at the manifestation of these intelligence differences in the applied sense and in real time. This is again where a neurocognitive lens can be helpful. It is worth noting at this point that much fruitful work has been done on mapping the neurological bases, including individual differences, of intelligence and for further insight the reader should consult the work of Richard Haier (e.g. Haier, 2011; Haier et al., 2009; Haier et al., 2004). Specific to scientific problem solving this has also been done by Tong et al. (2019) who found increased gray matter volume in the right superior temporal gyrus of young adults more adept at innovative problem solving. This was hypothesized to lead to their superior ability in the semantic integration of heuristic prototypes, key to insightful problem solving (Yang et al., 2016).

In examining individual differences in AR from the neurocognitive perspective Hammer et al. (2019), using fMRI with a large sample, found that individual differences were represented in the executive function and visuospatial networks (prefrontal cortex and lingual gyrus). Of important note was the finding that participants success in AR was contingent on their cognitive skill

(ability to activate) these cortical networks. This not only demonstrates how fMRI differentiated real-time cognitive networks of individual differences but highlights the area of visuospatial cognition once again as an important underpinning mechanism in STEM problem solving. Again, we can see the utility of neurocognitive perspectives in highlighting causal mechanisms of STEM learning. For more on the role of analogical reasoning in supporting science and math problem solving see [Brookman-Byrne et al. \(2019\)](#).

Concluding remarks on fMRI for STEM education research

The use of fMRI for educational research in STEM has many key benefits as can be seen from this brief review. There are however limitations in the capturing of functional data as what is measured is a hemodynamic response which is much slower (normally occurs approximately 2 s after the event) ([Freeman and Quiroga, 2013](#)). This presents issues for studies looking at particular stages of performance in applied tasks, as the data is an average over an entire episode ([Banich and Compton, 2011](#)). Again, this is the trade-off of advantages/disadvantages with EEG. In addition, fMRI requires advanced equipment, normally in the form of medical grade scanners (see [Fig. 6](#)), to undertake and is typically limited to the laboratory. This presents limitations with its application in applied settings such as the science laboratory once again bringing into question issues of validity. However, an emerging neurocognitive method that overcomes this limitation is functional near-infrared spectroscopy which provides spatially accurate measurement of hemodynamic brain responses and can be applied in “unconstrained” contexts. For more on this see [Balardin et al. \(2017\)](#).

Other physiological measures of learning processes

While I have devoted much of this section to the unpacking of some of the potential neurocognitive applications in STEM education research given their complexity, there are other methods available to the researcher in the wider physiological sphere. In the previous sections I have concentrated only on the cognitive processes underpinning typical learning processes in STEM however this is not to say that emotional responses cannot be captured. As we know, emotions, and their impact/association with cognition, have a large role in learning success ([Scrimin et al., 2016](#)) so in this section I will also look at those studies examining emotional aspects of learning with physiological lenses. Here I will try to demonstrate a variety of methods rather than the same depth as I have previously with the neurocognitive tools.

Physiology of emotion: heart rate and electrodermal approaches

Of all emotional variables that can impact the learning process in STEM, mathematics anxiety is possibly the best known and most pervasive area (see [Namkung et al., 2019](#)). While it is clear that math anxiety has a negative effect on the learning process, other research has shown its potential physical impact on health. For example, [Reims et al. \(2004\)](#) found that during a mental arithmetic test, adults’ blood pressure increased and demonstrated poorer recovery in those that had pre-test high blood pressure readings. This has obvious health implications but what about earlier in educational development, particularly in the classroom, and how could it be investigated? [Hunt et al. \(2017\)](#) conducted a study using a sample of primary school children and asking them to wear a portable blood pressure cuff (see [Fig. 7](#)) while solving arithmetic problems. They found that heart rate and blood pressure increased in students with increasing task difficulty (number of digits) but most significantly they found a positive correlation between mathematics anxiety and higher physiological readings (see [Table 1](#)). This highlights potential concerns for the well-being of pupils in



Fig. 6 fMRI scanner (“File:190603 functional magnetic resonance imaging at the Imperial Center for Psychedelic Research. jpg” by Thomas Angus, Imperial College London is licensed under CC BY-SA 4.0).



Fig. 7 Portable blood pressure monitor.

Table 1 Means (and SDs) of heart rate, sysBP and diaBP as a function of problem type.

		Baseline	2-Digit problems	3-Digit problems
Measure	Heart rate	92.00 (15.63)	89.18 (13.17)	89.75 (12.97)
	SysBP	105.51 (16.82)	105.56 (17.29)	109.36 (17.01)
	DiaBP	72.26 (16.98)	74.19 (14.21)	80.87 (17.00)

From Hunt, T.E., Bhardwa, J., Sheffield, D., 2017. "Mental arithmetic performance, physiological reactivity and mathematics anxiety among U.K. primary school children". *Learn. Indiv. Diff.* 57, 129–132.

this context and one which could not be inferred directly for the typically utilized method of self-report measures in mathematics anxiety research. This is potentially useful to the researcher as it is now possible to record these data within the social context of the school and classroom. This offers the potential to record quantitative physiological indices of emotional responses to phenomena of interest, and in a manner conducive to high levels of ecological validity or as part of a multidimensional approach to the measurement of emotional phenomena. As well as furthering our understanding of these emotional components, cognitive processes are also manifested with correlated physiological indicators (e.g. Obradović and Finch, 2017). This again gives the researcher another suite of tools in which to capture data on cognitive processes related to learning in STEM contexts.

Assessing engagement is an important aspect of educational research given the implications that this data would have for teaching and learning designs among other areas. Engagement can be captured in real-time using a physiological approach. Pijera-Díaz et al. (2018) wished to observe arousal states during physics lessons with high-school students and utilized an electrodermal measure. To do this, they employed the Empatica E4 wristband, which measures electrodermal activity from the surface of the participant's skin with accuracy (McCarthy et al., 2016). This device takes the form of a wrist wearable sensor similar to a watch. Pijera-Díaz et al. (2018) found that a predominant occurrence of low arousal was indicated in most lessons and during an exam setting that arousal increased and was correlated with success. These findings not only provide an indicator of task engagement, allowing, for example, a learning sciences researcher to implement a design-based approach to a new task format, but also gives us an insight into a potential threshold of cognitive engagement for academic performance. Although work using in this technology is relatively new in the context of STEM education research, there are some emerging projects appearing in the literature using this technique. For example in engineering education research (e.g. Khan et al., 2019; Villanueva et al., 2018).

As well as investigating cognitive components of task engagement, these wearable sensor devices can also be implemented in the context of quasi-experimental or field experiment research designs. Geršak et al. (2020) demonstrate an example of this where they studied two groups of second grade children who were taught under classical and movement-based pedagogical methods (based on embodied cognition), respectively. The researchers' goal was to explore the impact of physical activity on geometric problem solving and collected multidimensional physiological data from pupils including electrodermal activity, energy expenditure and skin temperature. These were captured using a wearable monitor and paired with an accelerometer to capture activity levels. This allowed Geršak et al. (2020) to determine that the movement group did demonstrate higher physical activity compared to the control, showed higher cognitive activation during the geometry lectures (measured by electrodermal response) and retained knowledge at a better rate 6 months later. Now caution must be exercised here as no effect sizes have been presented by Geršak et al. (2020) so we cannot determine the true effects of the experiment. However, the study does demonstrate the potential value such physiological methods can add to pedagogical research, particularly as it applies to experimentation with specific teaching and learning practices.

Observing physiological differences in expertise: eye-tracking approaches

Eye-tracking or pupillometry has emerged in recent years with various applications in educational research. Eye-tracking works by monitoring the fixations of individuals' gazes and also their pupillary responses (such as dilation as a result of physiological arousal) which can be used to infer cognitive processes such as cognitive load (Bianchi et al., 2020). As an applied example, within the field of teacher education there is always an interest in examining elements of teacher expertise in order to inform practice which traditionally relies on approaches such as observation of novice and expert teachers in practice. Cortina et al. (2015) used eye-tracking, worn by expert and novice teachers during class (see Fig. 8), to assess this topic. They found that in all, expert teachers fixated less on individual students indicating increased monitoring of overall classroom activity compared to those fixations of the novice teachers. However, as the authors integrated the tool with observational data of classroom behavior, they found that the higher fixations were not correlated with behavioral issues of individual students but instead related to the giving of quality feedback to students and engaging in discussion on subject matter. This allowed the researcher a more nuanced look into the manifestation of group differences in teaching approaches, beyond the more obvious hypotheses on classroom management issues.

Eye-tracking works by monitoring the fixations of individuals' gazes and also their pupillary responses (such as dilation as a result of physiological arousal) which can be used to infer cognitive processes such as cognitive load (Bianchi et al., 2020). As an applied example, within the field of teacher education there is always an interest in examining elements of teacher expertise in order to inform practice which traditionally relies on approaches such as observation of novice and expert teachers in practice. Cortina et al. (2015) used eye-tracking, worn by expert and novice teachers during class, to assess this topic. They found that in all, expert teachers fixated less on individual students indicating increased monitoring of overall classroom activity compared to those fixations of the novice teachers. However, as the authors integrated the tool with observational data of classroom behavior, they found that the higher fixations were not correlated with behavioral issues of individual students but instead related to the giving of quality feedback to students and engaging in discussion on subject matter. This allowed the researcher a more nuanced look into the manifestation of group differences in teaching approaches, beyond the more obvious hypotheses on classroom management issues.

One very useful benefit of eye tracking methods today, is their potential application to a wide variety of research participants. For example, application of some of the previously explored tools to young children in pre-school settings may prove problematic. For instance, even with mobile EEG, there is still quite a significant headset to be worn and the method itself is quite sensitive to extraneous factors such as excessive movement or physical interference. This in all likelihood rules this technique of capturing cognitive data an unsuitable approach. Eye-tracking however can be applied due to its additional robustness and variety of means data gathering. Today, eye tracking devices come in a variety of forms ranging from the glasses in Fig. 8 to remote eye tracking devices, such as the one shown in Fig. 9. Here the eye tracking device is integrated with the screen to be used for displaying research stimuli and the position of the child's seat is also depicted (c.f. Hessels and Hooge, 2019).

These types of technology clearly open up a myriad of possibilities for the STEM education researcher interested in cognitive process that underpin learning phenomena of interest. Given the interest that STEM education researchers have in spatial cognition, it would be useful to have a technique such as eye-tracking available in order to gather data on the early stages of skill development, say in preschoolers. Verdine et al. (2017) conducted such a study with 79 three-year olds and aimed to observe the interaction of socioeconomic status (SES) with differences in the cognitive processing of geometric shape names. The study employed an eye-tracking approach using an observational setup similar to that in Fig. 9. The results indicated that students in the lower SES took significantly longer to fixate on target geometric shapes and spent less time looking at them, revealing a distinct processing difference. Of most interest, performance on this task was predictive of spatial ability at age 5, suggesting that the ability to process geometric shape names may be foundational for later spatial development. This suggests that, like many wider areas of



Fig. 8 Eye-tracking wearable headset ("Eye tracking study" by Nestlé is licensed under CC BY-NC-ND 2.0).

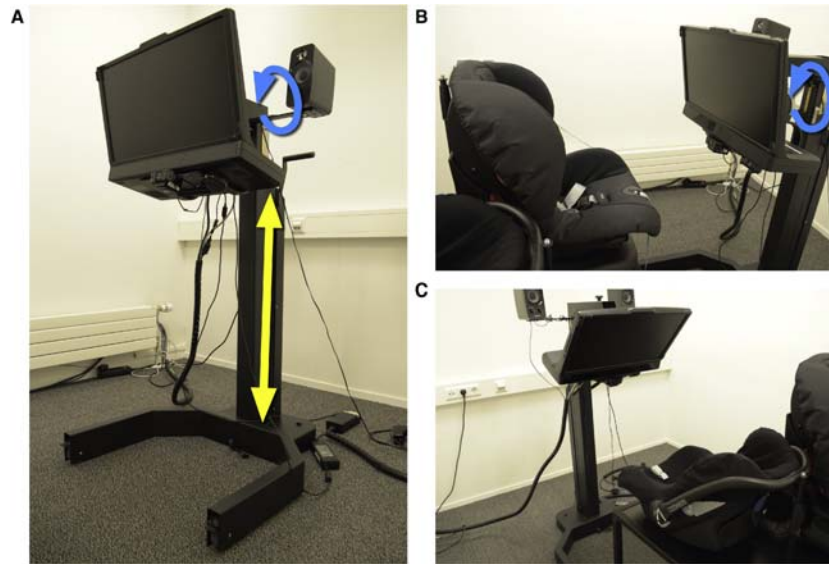


Fig. 9 Remote eye-tracking setup where the eye-tracker is mounted to the screen frame. From Hessels, R.S., Hooge, I.T.C., 2019. Eye tracking in developmental cognitive neuroscience – the good, the bad and the ugly. *Dev. Cogn. Neurosci.* 40, 100710. <https://doi.org/10.1016/j.dcn.2019.100710>, p. 40.

STEM education (e.g. Betancur et al., 2018), that SES is another variable that impacts the development of spatial cognitive abilities and from a young age. It also gives some insight into the potential etiological precursors to the successful development of spatial talents. These example studies demonstrate the broad applicability and rich data that the eye-tracking approach can generate.

Concluding remarks on physiological approaches in STEM educational research

As this brief overview shows, there are many advantages to the implementation of physiological methods to STEM learning investigations. They are in general less invasive than their neurocognitive counterparts and thus allow more widespread application. They are also applicable to a wider variety of potential participants, particularly young children. They do however suffer from the limitation that the researcher is not observing direct evidence of cognition per se, but must make inference of the underlying cognitive activity exemplified in physiological data. However, physiological data collection methods can be integrated with neurocognitive methods through software platforms such as iMotions (<https://imotions.com>) thereby overcoming some of these limitations. These approaches offer the STEM education researcher the potential to capture quantitative evidence of the manifestation of emotional responses, as well as data on cognition processes, of applied activities such as those undertaken in the school, classroom, lecture hall or even the home.

Broader considerations for application

This article has presented a broad overview of some of the neurocognitive and physiological approaches that have been applied to the field of STEM education research in recent years. While it gives a broad sample of the work carried out under a variety of themes and encompassing several phenomena of interest to the STEM education researcher, it is not intended to be an all-encompassing critical review and the reader should dig further into the studies and topics outlined for application in their own work. These neurocognitive and physiological methods have been applied to a relatively small number of topics in the field, but as the technology develops the application of these approaches may become even more widespread. Nonetheless there are some considerations worthy of noting.

There is much debate and controversy about the applicability of neuroscience concepts to education. One such area of contention is in the potential for these methods to assume the dominant explanatory position in education, at the expense of sociocultural perspectives (cf. Hall et al., 2014). It is important to reiterate at this point, that this is not the case and at all times in this article I have attempted to demonstrate how a particular neuroscientific or physiological approach can augment our understandings from wider theoretical perspectives. The key in all of this is in achieving an appropriate balance of methods and drawing conclusions and implications with cognisance that neuroscientific or physiological explanations can never explain the totality of learning. This issue in part relates to how one analogizes learning, as briefly mentioned earlier in the article where I explored the notion of a complex adaptive systems (for more depth on this discussion of analogies please see Delahunty and Kimbell (2021)), but also relating to how we conceptualize educational research as a discipline.

Without delving into a nuanced discussion of this issue here it is useful to highlight that, in the Anglo-American tradition, educational research has been conceived of as based on investigative lenses from the foundational disciplines of sociology, philosophy, psychology and history. Lawn and Furlong (2009) would argue that these disciplinary contributions have waned in recent times, but Biesta (2020) reminds us that in this Anglo-American tradition it is arguable as to whether educational research can be classified as a basic discipline in its own right and therefore must be multi/inter-disciplinary (pp. 86–87). The danger that emerges then in the context of STEM education research, and aligning with the critique levied by Hall et al. (2014), is the potential for premature application of insight from neurocognitive or physiological research to explain or inform learning practices. This is a salient concern for any context of education and educational research given the desire of policy makers to utilize, supposedly objective (Gillborn et al., 2018), quantitative data to derive educational policies. This worry is heightened when the quantitative data in question, is derived from, say, a neuroscientific perspective given the seductive allure of these data in supporting particular ideas (Weisberg et al., 2008). However, returning to the idea of foundational disciplines of educational research, it is doubtful that educationalists, who oppose the idea of neuroscience and physiological methods, would levy the same complaints to the field of psychology and its impact on educational practices.

Of the disciplines, psychology has had one of the strongest contributions and impacts on educational research but even this field of research is somewhat limited itself, when applied to questions of educational research. To take an example from Thomas et al. (2019), consider the widely circulated analogy of the computer metaphor to describe fundamental components of learning such as working memory. As this process assumes a general, context independent architecture, one would hypothesize that gains in one domain would translate to wider areas. However, it is well supported that such cases of “far-transfer” are extremely rare (Bransford and Schwartz, 1999). The implications here are that the brain must therefore utilize mechanisms of cognition that are less general than currently considered in cognitive theory and so there are “aspects of learning that simply do not make sense except with respect to the brain” (Thomas et al., 2019, p. 478). Taking this idea further, the situated or contextually specific nature of learning processes actually strengthens the role of sociocultural perspectives in explaining educational realities, but in order to gain a complete picture, the place of neurocognitive and physiological methods is underscored.

Conclusion

In order to achieve more complete understandings of STEM learning processes, educational researchers are required to adopt an open disposition to interdisciplinary collaborations. This must be conducted while maintaining a critical stance acknowledging the specific aspects of the learning process that the methods, reviewed in this article, are able to capture data on, and to not put so much weight on these data at the expense of missing vital perspectives offered by other disciplinary lenses. Neuroscientific data is particularly sensitive given the seductive appeal that these types of data have in the political landscape of academic publishing and funding applications. Therefore, I wish to conclude with a consideration for the STEM education researcher, inspired by Biesta’s (2020) discussion on the culture of measurement in education. That is, when considering the application of these methods to a research study ask yourself, are you measuring what you value or valuing what you measure? Put more plainly, start with the phenomenon of interest and after careful theorizing and initial research, then ask yourself if one of these methods will be illuminating in your particular research study.

References

- Alfred, K.L., Connolly, A.C., Kraemer, D.J.M., 2018. Putting the pieces together: generating a novel representational space through deductive reasoning. *NeuroImage* 183, 99–111.
- Balardin, J.B., Zimeo Morais, G.A., Furucho, R.A., Trambaiolli, L., Vanzella, P., Biazoli, C., Sato, J.R., 2017. Imaging brain function with functional near-infrared spectroscopy in unconstrained environments. *Front. Hum. Neurosci.* 11 (258).
- Banich, M.T., Compton, R.J., 2011. *Cognitive Neuroscience*. Cengage, Wadsworth, USA.
- Banks, F., Barlex, D., 2014. *Teaching STEM in the Secondary School*. Routledge, London.
- Barsalou, L.W., 2009. Situating concepts. In: Robbins, P., Aydede, M. (Eds.), *The Cambridge Handbook of Situated Cognition*. Cambridge University Press, New York.
- Betancur, L., Votruba-Drzal, E., Schunn, C., 2018. Socioeconomic gaps in science achievement. *Int. J. STEM Educ.* 5 (1), 1–25. <https://doi.org/10.1186/s40594-018-0132-5>.
- Bianchi, L.J., Kingstone, A., Risko, E.F., 2020. The role of cognitive load in modulating social looking: a mobile eye tracking study. *Cogn. Res.* 5 (1), 1–21.
- Biesta, G., 2020. *Educational research: An unorthodox introduction*. Bloomsbury Publishing.
- Bransford, J.D., Schwartz, D.L., 1999. Rethinking transfer: a simple protocol with multiple implications. *Rev. Res. Educ.* 24, 61–100.
- Brookman-Byrne, A., Mareschal, D., Tolmie, A.K., Dumontheil, I., 2019. The unique contributions of verbal analogical reasoning and nonverbal matrix reasoning to science and maths problem-solving in adolescence. *Mind Brain Educ.* 13 (3), 211–223.
- Cabeza, R., Nyberg, L., 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12 (1), 1–47.
- Casey, B.M., 2013. Individual and group differences in spatial ability. In: *Handbook of Spatial Cognition*. American Psychological Association, Washington, DC, US, pp. 117–134.
- Chiang, H.-S., Eroh, J., Spence, J.S., Motes, M.A., Maguire, M.J., Krawczyk, D.C., Brier, M.R., Hart, J., Kraut, M.A., 2016. Common and differential electrophysiological mechanisms underlying semantic object memory retrieval probed by features presented in different stimulus types. *Int. J. Psychophysiol.* 106, 77–86. <https://doi.org/10.1016/j.jpsycho.2016.06.011>.
- Cortina, K.S., Miller, K.F., McKenzie, R., Epstein, A., 2015. Where low and high inference data converge: validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *Int. J. Sci. Math. Educ.* 13 (2), 389–403.
- De Smedt, B., Grabner, R.H., Studer, B., 2009. Oscillatory EEG correlates of arithmetic strategy use in addition and subtraction. *Exp. Brain Res.* 195 (4), 635–642.
- Delahunty, T., Kimbell, R., 2021. (Re)framing a philosophical and epistemological framework for teaching and learning in STEM: emerging pedagogies for complexity. *Br. Educ. Res. J.* 47 (3), 742–769. <https://doi.org/10.1002/berj.3706>.

- Delahunty, T., Seery, N., Lynch, R., 2015. Spatial skills and success in problem solving within engineering education. In: 6th Research in Engineering Education Symposium, DIT, July 13–15.
- Delahunty, T., Sorby, S., Seery, N., Pérez, L., 2016. Spatial skills and success in engineering education: a case for investigating etiological underpinnings. In: 70th Engineering Design Graphics Division Annual Mid-Year Conference, Daytona Beach, Florida, January 24–26.
- Delahunty, T., Seery, N., Lynch, R., 2018. Exploring the use of electroencephalography to gather objective evidence of cognitive processing during problem solving. *J. Sci. Educ. Technol.* 27, 114–130. <https://doi.org/10.1007/s10956-017-9712-2>.
- Delahunty, T., Seery, N., Lynch, R., 2020. Exploring problem conceptualization and performance in STEM problem solving contexts. *Instr. Sci.* 48 (4), 395–425. <https://doi.org/10.1007/s11251-020-09515-4>.
- Delahunty, T., 2019. Enhancing the teaching of problem-solving in technology education. In: Williams, P., Barlex, D. (Eds.), *Explorations in Technology Education Research: Contemporary Issues in Technology Education*. Springer, Singapore. https://doi.org/10.1007/978-981-13-3010-0_10.
- Duffy, G., Sorby, S., Bowe, B., 2020. An investigation of the role of spatial ability in representing and solving word problems among engineering students. *J. Eng. Educ.* 109 (3), 424–442.
- Freeman, W.J., Quiroga, R.Q., 2013. Frequency analysis. In: *Imaging Brain Function With EEG*. Springer, pp. 21–36.
- Gardner, H., 2019. "Neuromyths": a critical consideration. *Mind Brain Educ.* 14 (1), 2–4.
- Geisser, C., Lehmann, W., Eid, M., 2006. Separating "rotators" from "nonrotators" in the mental rotations test: a multigroup latent class Analysis. *Multivariate Behav. Res.* 41 (3), 261–293.
- Gersák, V., et al., 2020. Use of wearable devices to study activity of children in classroom: case study — Learning geometry using movement. *Comput. Commun.* 150, 581–589. <https://doi.org/10.1016/j.comcom.2019.12.019>.
- Gill, H.S., O'Boyle, M.W., Hathaway, J., 1998. Cortical distribution of EEG activity for component processes during mental rotation. *Cortex* 34, 707–718.
- Gillborn, D., Warrington, P., Demack, S., 2018. QuantCrit: education, policy, 'Big Data' and principles for a critical race theory of statistics. *Race Ethn. Educ.* 21 (2), 158–179. <https://doi.org/10.1080/13613324.2017.1377417>.
- Gootjes, L., Bruggeling, E.C., Magnee, T., Van Strien, J.W., 2008. Sex differences in functional connectivity during mental rotation: an EEG study. *Int. J. Psychophysiol.* 69 (3), 228.
- Grabner, R.H., De Smedt, B., 2011. Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biol. Psychol.* 87 (1), 128–136.
- Grabner, R.H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., Neuper, C., 2009. To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* 47 (2), 604–608.
- Haier, R.J., Jung, R.E., Yeo, R.A., Head, K., Alkire, M.T., 2004. Structural brain variation and general intelligence. *NeuroImage* 23 (1), 425–433.
- Haier, R.J., Colom, R., Schroeder, D.H., Condon, C.A., Tang, C., Eaves, E., Head, K., 2009. Gray matter and intelligence factors: is there a neuro-g? *Intelligence* 37 (2), 136–144.
- Haier, R.J., 2011. Biological Basis of Intelligence. *The Cambridge handbook of intelligence*, pp. 351–370.
- Hall, K., Curtin, A., Rutherford, V., 2014. *Networks of Mind: Learning, Culture, Neuroscience*. Routledge.
- Halpern, D.F., Collaer, M.L., 2005. Sex differences in visuospatial abilities. In: Shah, P., Miyake, A. (Eds.), *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press, New York, pp. 170–212.
- Hammer, R., Paul, E.J., Hillman, C.H., Kramer, A.F., Cohen, N.J., Barbey, A.K., 2019. Individual differences in analogical reasoning revealed by multivariate task-based functional brain imaging. *NeuroImage* 184, 993–1004.
- Hegarty, M., 2018. Ability and sex differences in spatial thinking: what does the mental rotation test really measure. *Psychon. Bull. Rev.* 25, 1212–1219.
- Heil, M., Jansen-Osmann, P., 2008. Sex differences in mental rotation with polygons of different complexity: do men utilize holistic processes whereas women prefer piecemeal ones? *Q. J. Exp. Psychol.* 61 (5), 683–689.
- Hessels, R.S., Hooge, I.T.C., 2019. Eye tracking in developmental cognitive neuroscience – the good, the bad and the ugly. *Dev. Cogn. Neurosci.* 40, 100710. <https://doi.org/10.1016/j.dcn.2019.100710>.
- Hinault, T., Lemaire, P., 2016. What does EEG tell us about arithmetic strategies? A review. *Int. J. Psychophysiol.* 106, 115–126.
- Howard-Jones, P.A., 2011. A multiperspective approach to neuroeducational research. *Educ. Philos. Theor.* 43 (1), 24–30. <https://doi.org/10.1111/j.1469-5812.2010.00703.x>.
- Hunt, T.E., Bhargava, J., Sheffield, D., 2017. Mental arithmetic performance, physiological reactivity and mathematics anxiety amongst U.K. primary school children. *Learn. Individ. Differ.* 57, 129–132.
- Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., Delazer, M., 2007. Imaging early practice effects in arithmetic. *NeuroImage* 36 (3), 993–1003.
- Jacobson, M.J., Kapur, M., Reimann, P., 2016. Conceptualizing debates in learning and educational research: toward a complex systems conceptual framework of learning. *Educ. Psychol.* 51 (2), 210–218.
- Jacobson, M.J., Levin, J.A., Kapur, M., 2019. Education as a complex system: conceptual and methodological implication. *Educ. Res.* 48 (2), 112–119.
- Jakel, F., Schreiber, C., 2013. Introspection in problem solving. *J. Prob. Sol.* 6 (1), 20–33.
- Kaufman, S.B., 2007. Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity? *Intelligence* 35 (3), 211–223.
- Khan, T.H., Villanueva, I., Vicioso, P., Husman, J., 2019. Exploring relationships between electrodermal activity, skin temperature, and performance during. In: 2019 IEEE Frontiers in Education Conference (FIE), pp. 1–5. <https://doi.org/10.1109/FIE43999.2019.9028625>.
- Kimbell, R., Stables, K., 2008. *Researching Design Learning: Issues and Findings From Two Decades of Research and Development*. Springer Science and Business Media B.V.
- Klimesch, W., 1997. EEG-alpha rhythms and memory processes. *Int. J. Psychophysiol.* 26, 319–340.
- Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29 (2), 169–195.
- Krumnack, U., Kühnberger, K.-U., Schwering, A., Münster, O.O., Besold, T.R., 2013. Analogies and analogical reasoning in invention. In: *Encyclopaedia of Creativity, Invention, Innovation and Entrepreneurship*, pp. 56–62.
- Kyttälä, M., Björn, P.M., 2014. The role of literacy skills in adolescents' mathematics word problem performance: controlling for visuo-spatial ability and mathematics anxiety. *Learn. Individ. Differ.* 29, 59–66.
- Lawn, M., Furlong, J., 2009. The disciplines of education in the UK: between the ghost and the shadow. *Oxf. Rev. Educ.* 35 (5), 541–552. <https://doi.org/10.1080/03054980903216283>.
- Li, Y., Boyle, M.O., 2013. How sex and College major relate to mental rotation accuracy and preferred strategy: an electroencephalographic (EEG) study. *Psychol. Rec.* 63 (27–42).
- Loetscher, T., Bockisch, C.J., Brugger, P., 2007. Looking for the answer: the mind's eye in number space. *Neuroscience* 151 (3), 725–729.
- McCarthy, C., Pradhan, N., Redpath, C., Adler, A., 2016. Validation of the Empatica E4 wristband. In: 2016 IEEE EMBS International Student Conference (ISC), pp. 1–4. <https://doi.org/10.1109/EMBSISC.2016.7508621>.
- Michel, C.M., Kaufman, L., Williamson, S.J., 1994. Duration of EEG and MEG α suppression increases with angle in a mental rotation task. *J. Cogn. Neurosci.* 6 (2), 139–150.
- Mohanty, A., Engels, A.S., Herrington, J.D., Heller, W., Ringo Ho, M.-H., Banich, M.T., Webb, A.G., Warren, S.L., Miller, G.A., 2007. Differential engagement of anterior cingulate cortex subdivisions for cognitive and emotional function. *Psychophysiology* 44 (3), 343–351.
- Mozzer, N.B., Mozzer, N.B., Justi, R., Justi, R., 2013. Science teachers' analogical reasoning. *Res. Sci. Educ.* 43 (4), 1689–1713.
- Namkung, J.M., Peng, P., Lin, X., 2019. The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. *Rev. Educ. Res.* 89 (3), 459–496.
- Neubauer, A.C., Fink, A., Grabner, R.H., 2006. Sensitivity of alpha band ERD to individual differences in cognition. *Prog. Brain Res.* 159, 167–178.
- Newhouse, C.P., 2016. STEM the boredom: engage students in the Australian curriculum using ICT with problem-based learning and assessment. *J. Sci. Educ. Technol.* 26 (1), 44–57.

- Newman, S.D., Pruce, B., Rusia, A., Burns, T., 2010. The effect of strategy on problem solving: an fMRI study. *J. Prob. Sol.* 3 (1), 1–26.
- Newman, S.D., Willoughby, G., Pruce, B., 2011. The effect of problem structure on problem-solving: an fMRI study of word versus number problems. *Brain Res.* 1410, 77–88.
- Obradović, J., Finch, J.E., 2017. Linking executive function skills and physiological challenge response: piecewise growth curve modeling. *Dev. Sci.* 20 (6), e12476. <https://doi.org/10.1111/desc.12476> n/a.
- Ohlsson, S., 2012. The problems with problem solving: reflections on the rise, current status, and possible future of a cognitive research paradigm. *J. Prob. Sol.* 5 (1), 101–128.
- Parsons, T., 2004. Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia* 42 (4), 555–562.
- Peña, D., Contreras, M.J., Shih, P.C., Santacreu, J., 2008. Solution strategies as possible explanations of individual and sex differences in a dynamic spatial task. *Acta Psychol.* 128 (1), 1–14.
- Pijera-Díaz, H.J., Drachsler, H., Kirschner, P.A., Järvelä, S., 2018. Profiling sympathetic arousal in a physics course: how active are students? *J. Comput. Assist. Learn.* 34 (4), 397–408.
- Reims, H., Sevre, K., Fossum, E., Høiegg, A., Eide, I., Kjeldsen, S., 2004. Plasma catecholamines, blood pressure responses and perceived stress during mental arithmetic stress in young men. *Blood Pressure* 13 (5), 287–294.
- Rescher, B., Rapplesberger, P., 1999. Gender dependent EEG-changes during a mental rotation task. *Int. J. Psychophysiol.* 33, 209–222.
- Richardson, K., Norgate, S.H., 2015. Does IQ really predict job performance? *Appl. Dev. Sci.* 19 (3), 153–169.
- Rowan, A.J., Tolunsky, E., 2003. *Primer of EEG*. Butterworth Heinemann, Pennsylvania.
- Schnitz, W., Baadte, C., Uller, A.M., Rasch, R., 2010. Creative thinking and problem solving with depictive and descriptive representations. In: Verschaffel, L., DeCorte, E., deJong, T., Ellen, J. (Eds.), *Use of Representations in Reasoning and Problem Solving*. Routledge, New York, pp. 11–35.
- Scrimin, S., Altoè, G., Moscardino, U., Pastore, M., Mason, L., 2016. Individual differences in emotional reactivity and academic achievement: a psychophysiological study. *Mind Brain Educ.* 10 (1), 34–46.
- Sharma, G., Anto, A., Singh, V., 2018. Sex differences in mental rotation: cortical functional connectivity using direct transfer function. *Biomed. Signal Process. Contr.* 40, 425–432.
- Soltanlou, M., Artemenko, C., Dresler, T., Fallgatter, A.J., Nuerk, H.-C., Ehlis, A.-C., 2019. Oscillatory EEG changes during arithmetic learning in children. *Dev. Neuropsychol.* 44 (3), 325–338.
- Sorby, S.A., 1999. Developing 3-D spatial visualization skills. *Eng. Des. Graph. J.* 63 (2), 21–32.
- Sorby, S., 2009. Educational research in developing 3-D spatial skills for engineering students. *Int. J. Sci. Educ.* 31 (3), 459–480.
- Stamenković, D., Ichien, N., Holyoak, K.J., 2019. Metaphor comprehension: an individual-differences approach. *J. Mem. Lang.* 105, 108–118.
- Stephens, R., Richey, M., 2011. Accelerating STEM capacity: a complex adaptive system perspective. *J. Eng. Educ.* 100 (3), 417–423.
- Stillings, N.A., Weisler, S.E., Chase, C.H., Feinstein, M.H., Garfield, J.L., Rissland, E.L., 1995. *Cognitive Science: An Introduction*. MIT Press, London.
- Thevenot, C., Castel, C., Fanget, M., Fayol, M., 2010. Mental subtraction in high- and lower skilled arithmetic problem solvers: verbal report versus operand-recognition paradigms. *J. Exp. Psychol. Learn. Mem. Cogn.* 36 (5), 1242–1255.
- Thomas, M.S.C., Ansari, D., Knowland, V.C.P., 2019. Annual research review: educational neuroscience: progress and prospects. *J. Child Psychol. Psychiatr.* 60 (4), 477–492. <https://doi.org/10.1111/jcpp.12973>.
- Tong, D., Lu, P., Li, W., Yang, W., Yang, Y., Yang, D., Qiu, J., Zhang, Q., 2019. Critical thinking and regional gray matter volume interact to predict representation connection in scientific problem solving. *Exp. Brain Res.* 237 (8), 2035–2044.
- Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., Newcombe, N.S., 2013. The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139 (2), 352–402.
- Verdine, B.N., Bunker, A., Athanasopoulou, A., Golinkoff, R.M., Hirsh-Pasek, K., 2017. Shape up: an eye-tracking study of preschoolers' shape name processing and spatial development. *Dev. Psychol.* 53 (10), 1869–1880. <https://doi.org/10.1037/dev0000384>.
- Villanueva, I., Campbell, B.D., Raikes, A.C., Jones, S.H., Putney, L.G., 2018. A multimodal exploration of engineering students' emotions and electrodermal activity in design activities. *J. Eng. Educ.* 107 (3), 414–441. <https://doi.org/10.1002/jee.20225>.
- Ward, J., 2010. *The Student's Guide to Cognitive Neuroscience*, second ed. Psychology Press, East Sussex.
- Weisberg, D.S., Keil, F.C., Goodstein, J., Rawson, E., Gray, J.R., 2008. The seductive allure of neuroscience explanations. *J. Cogn. Neurosci.* 20 (3), 470–477.
- Wendelken, C., 2015. Meta-analysis: how does posterior parietal cortex contribute to reasoning? *Front. Hum. Neurosci.* 8. <https://doi.org/10.3389/fnhum.2014.01042>.
- Williams, J., Iglesias, J., Barak, M., 2008. Problem based learning: application to technology education in three countries. *Int. J. Technol. Des. Educ.* 18, 319–335.
- Yang, W., Dietrich, A., Liu, P., Ming, D., Jin, Y., Nusbaum, H.C., Qiu, J., Zhang, Q., 2016. Prototypes are key heuristic information in insight problem solving. *Creativ. Res. J.* 28 (1), 67–77.
- Zarjam, P., Epps, J., Chen, F., Lovell, N.H., 2013. Estimating cognitive workload using wavelet entropy-based features during an arithmetic task. *Comput. Biol. Med.* 43 (12), 2186–2195.

Further reading

- Bermúdez, J.L., 2014. *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge University Press, United Kingdom.
- Blakemore, S.-J., Frith, U., 2005. The learning brain: lessons for education: a précis. *Dev. Sci.* 8 (6), 459–471.
- Kosslyn, S.M., Ganis, G., Thompson, W.L., 2001. Neural foundations of imagery. *Nat. Rev.* 2, 635–642.
- OECD, 2007. *Understanding the Brain: The Birth of a Learning Science*. OECD Publishing, Paris.

FOURTH EDITION

International Encyclopedia of Education

Qualitative, Multimethod,
and Mixed Methods Research

VOLUME
12

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editor

Nataliya Ivankova



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 12

Qualitative, Multimethod, and Mixed Methods Research

VOLUME 12 EDITOR

Nataliya V Ivankova



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

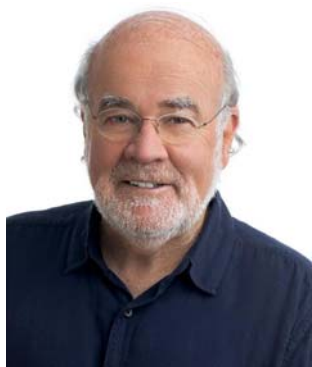
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



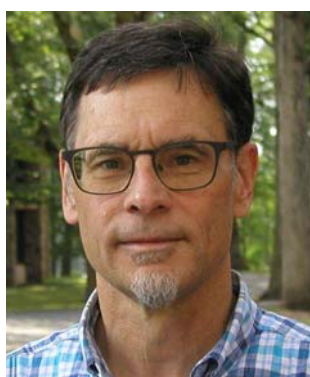
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 12: Qualitative, Multimethod, and Mixed Methods Research

Nataliya V Ivankova

© 2023 Elsevier Ltd. All rights reserved.

Introduction

Research in education has a long history of applying different methods and forms of inquiry to address complex education issues. While a quantitative research approach with its focus on standardized measurements and objective scores has prevailed in education research, other forms of inquiry, such as qualitative and mixed methods research, have gained popularity in education studies. Qualitative research that aims at exploring individuals' experiences with and perceptions of social phenomena has achieved recognition as a means of exploring and understanding the complexity of the teaching and learning process. Its rightful position in international education research was captured through a representative number of articles addressing philosophical and methodological aspects of qualitative research in the previous, third edition of the *International Encyclopedia of Education*. Many of the chapters in the current volume provide updated information on those topics, but also reflect the new developments in conceptualization and conduct of qualitative research in education.

Recently, mixed methods research has been added to the toolkit of education researchers. Mixed methods research quickly became the methodology of choice due to its ability to meaningfully integrate qualitative stakeholder engagement methods with quantitative outcome-focused approaches to promote an integrative thinking for addressing complex education issues. The growing popularity of mixed methods in education research is reflected in the number of chapters in this volume that are devoted to conceptual and procedural issues of designing, conducting, and reporting mixed methods studies in education. In contrast, mixed methods research was represented by one article written by A. Tashakkori and I. Newman in the previous, third edition of the *International Encyclopedia of Education*.

While there is no clear distinction between mixed methods and multimethod research, and mixed methods scholars' perspectives vary in this regard, we acknowledge the existence of fundamental similarities between these two approaches. They both aim at a deeper, comprehensive, and fuller understanding of the problem and include multiple forms of evidence (Hunter and Brewer, 2015; Flick, 2017), which are integrated to produce knowledge that is more than the sum of the individual quantitative and qualitative components (Fetters and Freshwater, 2015) (see J. Schoonenboom's chapter on mixed methods and multimethod research in education in this volume for a more detailed discussion). Moreover, as claimed by A. Skukauskaite in this volume, some traditional forms of qualitative inquiry are multimethod in nature because they rely on combination of alternative methods to produce more in-depth conclusions about the studied problem. So, we encourage the readers to think about these forms of inquiry from a pragmatic perspective with the focus on the value added that the combination of multiple methods offers to understanding education issues.

This volume consists of 75 chapters that address different aspects of qualitative, multimethod, and mixed methods research from philosophical and conceptual issues, to design and application of methods, and to presentation and reporting of data and findings. Consistent with the international focus of the Encyclopedia, the volume includes chapters written by the authors from 21 countries representing 5 continents. Of 128 authors who contributed chapters to this volume, about 48% are international. This speaks to the popularity of qualitative, multimethod, and mixed methods research in education across different countries and continents. Such broad international representation across the chapters also helps capture variability of methodological perspectives related to conceptualizing and using these research approaches shaped by unique historical developments, epistemological preferences, and contextual influences observed in these and similar countries. What is common across the authors is the recognition of the many advantages that such methods provide to address a wide range of education issues that require in-depth exploration of the users' experiences and viewpoints as well as a comprehensive treatment of the problem from multiple angles and multiple perspectives.

The chapters are organized into eight sections, three of which are devoted to qualitative research and five sections focus on multimethod and mixed methods research. The following is a detailed account of each section to help orient the readers to their structure and content.

The first section includes eight chapters that address philosophical and theoretical underpinnings of qualitative research. The first chapter under the same title by E. Boswell and W. Babchuk sets the tone for the whole volume by addressing the role of philosophy and theory in research in general and qualitative research, in particular, with implications for multimethod and mixed methods research. The next chapter by M. Hammersley provides a comprehensive historical overview of the use of qualitative research methods in education and sets the stage for the discussion of different qualitative research approaches and their applications in education research in the chapter by P. Delmas and R. Giles. A. Ryen engages the readers in a discussion of ethics and politics in qualitative research and their influence on how qualitative results are viewed and interpreted by different audiences. The other four chapters focus on applications of different theoretical perspectives to guide qualitative studies in education, such as the role of critical theory in qualitative research in education (J. Ryoo and J. Crawford), and examples of applications of feminist and gender theory (C. McLewis, J. Gutzwa, and S. Hurtado), critical race theory (T. Loder-Jackson, M. Bodine Al-Sharif, C. Jones, M. Files, and C. Wiggins), and queer theory (B. Wozolek). The importance of these theoretical perspectives for informing qualitative inquiry has been well recognized (Creswell and Poth, 2018; Kincheloe and McLaren, 2011; Winkle-Wagner et al., 2019).

The next section consisting of 12 chapters focus on different qualitative research designs and approaches, the design and procedural features, and illustration of their use in research in education. The first six chapters address the established qualitative research designs and approaches, often referred to as “traditions of qualitative inquiry” (Creswell and Poth, 2018; Hays and Singh, 2011), and include case study (R. Stake and M. Visse), grounded theory (W. Babchuk and E. Boswell), phenomenology (N. Friesen), narrative (S. Chung and J. Clandinin), ethnography (A. Skukauskaite), and autoethnography (S. Pensoneau-Conway). Each chapter offers a theoretical foundation and a practical application of the discussed approach capitalizing on the new developments in conceptualizing and applying these designs while illustrating innovative thinking. Other chapters in this section provide a more extended perspective on applications of qualitative research approaches in pursuit of specific purposes, such as responding to cultural diversity through action and community-based participatory research (R. Bead, A. Ortiz Aragon, and E. Stringer), addressing community needs by engaging in an oral tradition of storytelling (C. Nakhid and A. Fernandez Santana), and using discourse analysis as a space of signification through communication and meaning making (L. Suci). An innovative chapter by J. Greenwood explores interconnections between art and qualitative research, while a chapter by M. Danesi provides insights into semiotics as a form of qualitative inquiry. The chapter by M. Patton underscores the utility of qualitative research for evaluating educational interventions by connecting evaluation questions to intended users and selecting appropriate methods.

The third section focuses on qualitative research methods and their application in education research from a new or updated angle. Each of the 13 chapters discusses a specific qualitative method or strategy, such as approaches to sampling and collecting qualitative data (M. Koro, N. Fairchild, A. Benozzo, and T. Löytönen), using qualitative interviews (S. Brinkmann), conceptualizing focus groups (A. Welker and G. Kamberelis), doing naturalistic observations (J. Schostak), including document analysis (T. Rapley, K. Ardron, D. Littlefair, K. Mulholland, and N. Jenkins), engaging with photo voice (H. Meyer, L. Dell, and R. Gibson), and conducting participatory photography (G. Holm). Several chapters address different approaches to qualitative data analysis including inductive thematic analysis (P. Mihos) and content analysis (P. Mayring). The chapters discussing rapid qualitative research (C. Vindrola-Padros, G. Chisnall, and S. Kumpunen) and conducting qualitative research in online spaces (H. Gerber) emphasize the advances in the application of the established qualitative methods in new research environments. The last two chapters in this section address important issues of establishing quality in qualitative research (L. Cain, R. Williams, and V. Bradshaw) and the advantages and caveats of triangulating multiple methods and data in qualitative studies (A. Pratesi), thus providing a transition to multimethod and mixed methods research which is the focus of the next four sections in this volume.

Section four opens with the chapter by J. Schoonenboom, in which she discusses the foundational aspects of multimethod and mixed methods research and outlines the design decisions important for collecting, combining, and developing multimethod and mixed methods datasets. The next chapter by A. Tashakkori and late I. Newman is an updated chapter from the third edition of the Encyclopedia. In this chapter, the authors provide a foundational overview of mixed methods research using the new developments in the mixed methods field and discuss the advantages of integrating multiple methods in research in education. Five chapters that follow explore philosophical and conceptual premises of mixed methods research offering current perspectives on philosophical underpinnings of the mixed methods approach (P. Shannon-Baker), conceptual frameworks to guide applications of mixed methods research (V. Plano Clark, N. Ivankova, and N. Yang), transformative mixed methods research (D. Mertens), and the role of culture in conducting mixed methods studies in education written by a team of researchers from Jamaica and led by L. Cook. The chapter by E. Creamer delves into the concept of integration in mixed methods research and the philosophical assumptions that facilitate meaningful integration of multiple data, methods, and approaches in mixed methods and multimethod studies.

Section five is devoted to the methodological aspects of designing integrated multimethod and mixed methods studies. This topic has received a lot of attention in the mixed methods literature with discussions ranging from advancing mixed methods design typologies (Creswell and Plano Clark, 2018; Greene, 2007; Morse and Niehaus, 2009; Tashakkori et al., 2021) to typology-free alternative viewpoints (Guest, 2013; Maxwell, 2013; Nastasi et al., 2010). Ten chapters in this volume provide current thinking on different topics within the spectrum of the mixed methods research process and the “how-to” of a mixed methods study. C. Tajima and M. Fetter’s chapter offers an overview of mixed methods research designs using a typology-based approach, whereas J. Maxwell’s chapter on designing integrative mixed methods research presents an interactive concept of design that emphasizes the unique role of the study components and their interaction in the planning and conduct of a mixed methods project. B. Smit, V. Scherman, and L. Liebenberg discuss the role of the qualitative research approach in designing mixed methods studies, and P. Pluye and Q. Hong define and illustrate the concepts of convergence and divergence in mixed methods research. Other chapters in this section cover the logistics of designing and implementing mixed methods studies, including mixed methods sampling and data collection (N. Leech and C. Donovan) and integration strategies (J. Schoonenboom), approaches to mixed methods data analysis (S. Vogl) and mixed methods analysis software (U. Kuckartz and S. Rädiker), quality assurance in mixed methods research (K. Collins), and ethical dilemmas arising from integration of multiple methods (C. Poth, E. Creamer, and L. Cain).

The sixth section combines 10 chapters that illustrate the intersection of mixed methods research with other approaches and designs to address complex education research problems. A concept of methodological intersection has recently emerged in the mixed methods literature to describe research situations when two or more research approaches are intentionally joined within a study design (Plano Clark and Ivankova, 2016). Three chapters discuss the logistics, advantages, and challenges of intersecting mixed methods research with traditional qualitative research designs, such as case study (V. Plano Clark, L. Foote, and J. Walton), grounded theory (E. Creamer), and phenomenology (I. Mayoh and J. Colditz). A chapter by H. White and Y. Ting illustrates the advantages of intersection of mixed methods with quantitative randomized controlled trials in education. At the methodological level, D. Mertens addresses the long-established practices of intersecting mixed methods with evaluation, and N. Ivankova and N. Wingo provide a framework for intersecting mixed methods with action and community-based participatory research approaches.

More recent developments in the intersection of mixed methods with other methodologies include combining mixed methods research with network analysis (D. Froehlich), Delphi method (M. Taghipoorreynah), qualitative comparative analysis (H. Kane and L. Kahwati), and translational research (N. Ivankova, J. Anderson, I. Herbey, L. Roussel, and D. Kim). These chapters illustrate how applying mixed methods research can enhance established designs and methodologies to better address education problems.

Section seven includes nine chapters showcasing different applications of multimethod and mixed methods research in education studies. A chapter by C. Poth and E. Bullock specifically focuses on mixed methods research design practices to address complexity in education. K. Niglas and coauthors describe the logistics of conducting large-scale mixed methods studies in education, while D. Watkins and N. Johnson discuss advancing education research by using mixed methods with existing data. Methodological flexibility of mixed methods research provides a platform for its application in developing culturally tailored interventions (B. Nastasi) and for enhancing systematic literature reviews (P. Onghena). Other chapters provide insights into innovative applications of mixed methods research in combination with other methods, such as geographic information systems (N. Bagheri), art and artistic practices (P. Shannon-Baker and L. Smith), survey methods (J. Hitchcock and A. Crean Davis), and web-based technologies (J. Salmons).

The last section consisting of six chapters provides applied information on how to propose, report, and visualize qualitative, multimethod, and mixed methods research in education. Chapters on developing a qualitative research proposal by M. Weaver-Hightower and on writing a qualitative research report by J. Drisko offer practical advice to those who plan to propose a qualitative study or need to publish findings based on qualitative data. The latter chapter also describes the process of submitting a qualitative manuscript to a journal, explains a peer review process, and suggests strategies for optimizing article acceptance rates. Two chapters provide advice on writing a successful mixed methods research proposal (J. DeCuir-Gunby and W. McCoy) and on preparing journal manuscripts based on mixed methods studies (J. Gothberg and A. Sterenberg Mahon) that are, reportedly, more difficult to publish due to their length and scope (Bazeley, 2015; Creswell, 2022). This section concludes with two chapters that address different types of qualitative and mixed methods visualizations that enhance reporting by using joint displays (T. Guetterman) and improve understanding of information by developing visuals to illustrate the methodological features of a mixed methods study (M. McCrudden).

We hope that the collection of the chapters in this volume responds to the many interests and needs in qualitative, multimethod, and mixed methods research of the Encyclopedia readers. A balanced approach to presenting theoretical and methodological information and practical “how-to” applications in this volume chapters aims at ensuring the readers’ diverse needs are met and serves as a safeguard for their understanding and applying the information. To optimize the learning experience and to expand the readers’ knowledge on the topics of interest, we cross-listed the chapters where possible across the volume, thereby providing the readers with points of reference to other chapters that discuss the same or similar topics in more depth. In conclusion, we wish the readers an enjoyable journey into qualitative, multimethod, and mixed methods research that the chapters in this Encyclopedia volume offer.

References

- Bazeley, P., 2015. Writing up multimethod and mixed methods research for diverse audiences. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multi-method and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, UK, pp. 296–313.
- Creswell, J.W., 2022. *A Concise Introduction to Mixed Methods Research*, second ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 4th ed. Sage Publications.
- Fetters, M.D., Freshwater, D., 2015. The 1 + 1 = 3 integration challenge. *J. Mixed Methods Res.* 9 (2), 115–117.
- Flick, U., 2017. Mantras and myths: the disenchantment of mixed-methods research and revisiting triangulation as a perspective. *Qual. Inq.* 23, 46–57.
- Greene, J.C., 2007. *Mixed methods in social inquiry*. Jossey-Bass, San Francisco.
- Maxwell, J.A., 2013. In: *Qualitative Research Design: An Interactive Approach*, third ed. Sage, Thousand Oaks, CA.
- Guest, G., 2013. Describing mixed methods research: an alternative to typologies. *J. Mixed Methods Res.* 7 (2), 141–151.
- Hays, D.G., Singh, A.A., 2011. *Qualitative Inquiry in Clinical and Educational Settings*. Guilford Press.
- Hunter, A., Brewer, J., 2015. Designing multimethod research. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, U.K., pp. 185–205.
- Kincheloe, J.L., McLaren, P., 2011. Rethinking critical theory and qualitative research. In: *Key Works in Critical Pedagogy*. Brill, pp. 285–326.
- Morse, J.M., Niehaus, L., 2009. *Mixed Methods Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Nastasi, B.K., Hitchcock, J.H., Brown, L.M., 2010. An inclusive framework for conceptualizing mixed methods design typologies: Moving toward fully integrated synergetic research models. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 305–338.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Winkle-Wagner, R., Lee-Johnson, J., Gaskew, A. (Eds.), 2019. *Critical Theory and Qualitative Data Analysis in Education*. Routledge, New York.

This page intentionally left blank

CONTRIBUTORS TO VOLUME 12

Jami L Anderson

School of Health Professions, The University of Alabama at Birmingham, Birmingham AL, USA

Kevin Ardron

Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom

Wayne A Babchuk

Department of Educational Psychology, University of Nebraska-Lincoln, Lincoln, NE, United States

Nazgol Bagheri

University of Texas at San Antonio, San Antonio, TX, United States

Rosalind Beadle

Flinders University, Bedford Park, SA, Australia

Angelo Benozzo

University of Turin, Turin, Italy

Mary Ann Bodine Al-Sharif

The University of Alabama at Birmingham, School of Education, Birmingham, AL, United States

Eileen Boswell

Department of Teaching, Learning, and Teacher Education, University of Nebraska-Lincoln, Lincoln, NE, United States

Victoria Bradshaw

Sport Psychology and Motor Behavior, University of Tennessee, Knoxville, TN, United States

Svend Brinkmann

Aalborg University, Denmark

Emma P Bullock

Department of Mathematics and Statistics, College of Science and Engineering Technology, Sam Houston State University, Huntsville, TX, United States

Leia K Cain

Department of Evaluation, Statistics, and Methodology, University of Tennessee, Knoxville, TN, United States; and Department of Educational Psychology & Counseling, University of Tennessee, Evaluation, Statistics, and Methodology, Knoxville, TN, United States

Georgia Chisnall

Department of Targeted Intervention, University College London (UCL), London, United Kingdom

Simmee Chung

Concordia University of Edmonton, Edmonton, AB, Canada

D Jean Clandinin

University of Alberta, Edmonton, AB, Canada

Jason Colditz

University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

Kathleen MT Collins

Department of Curriculum & Instruction, University of Arkansas, Fayetteville, AR, United States

Lorraine D Cook

School of Education, University of the West Indies, Mona, Jamaica

Jenifer Crawford

University of Southern California, Los Angeles, CA, United States

Elizabeth G Creamer

School of Education, Virginia Tech, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States

Marcel Danesi

Department of Anthropology, University of Toronto, Toronto, ON, Canada

Allison Crean Davis
Westat, Rockville, MD, United States

Jessica T DeCuir-Gunby
University of Southern California, Los Angeles, CA, United States

Laura Dell
University of Cincinnati, College of Education, Criminal Justice & Human Services, Cincinnati, OH, United States

Peggy M Delmas
University of South Alabama, Department of Leadership & Teacher Education, Mobile, AL, United States

Courtney Donovan
University of Colorado Denver, Denver, CO, United States

James Drisko
Smith College, Northampton, MA, United States

Nikki Fairchild
Research and Innovation, School of Education and Sociology, University of Portsmouth, Portsmouth, United Kingdom

Michael D Feters
Mixed Methods Program, Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

Martez D Files
University of Pittsburgh, School of Education, Pittsburgh, PA, United States

Lori A Foote
School of Education, University of Cincinnati, Cincinnati, OH, United States

Norm Friesen
Boise State University, College of Education, Boise, ID, United States

Dominik E Froehlich
University of Vienna, Vienna, Austria

Nollaig Frost
University College Cork, School of Applied Psychology, Ireland

Hannah R Gerber
Sam Houston State University, Huntsville, TX, United States; and University of South Africa, Pretoria, South Africa

Randall Gibson
University of Cincinnati, School of Education, Cincinnati, OH, United States

Rachel L Giles
University of South Alabama, Department of Leadership & Teacher Education, Mobile, AL, United States

June E Gothberg
Cornell University, Battle Creek, MI, United States

Janinka Greenwood
University of Canterbury, Whangarei, New Zealand

Timothy C Guetterman
University of Michigan, Ann Arbor, MI, United States

Justin A Gutzwa
Department of Physics and Astronomy, University of Utah, Salt Lake City, UT, United States

Martyn Hammersley
WELS, The Open University, Milton Keynes, United Kingdom

Tashane Haynes-Brown
School of Education, University of the West Indies, Mona, Jamaica

Ivan I Herbey
School of Medicine, The University of Alabama at Birmingham

John H Hitchcock
Westat, Rockville, MD, United States

Gunilla Holm
Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

Quan Nha Hong
School of Rehabilitation, Université de Montréal, Montreal, QC, Canada

Ingrid E Hunt-Anderson
School of Education, University of the West Indies, Mona, Jamaica

S Hurtado
Department of Education, University of California. Los Angeles, CA, United States

Nataliya V Ivankova
School of Health Professions, The University of Alabama at Birmingham, Birmingham, AL, United States

K Neil Jenkins
School of Geography, Politics & Sociology, Newcastle University, Newcastle upon Tyne, United Kingdom

Natasha C Johnson
Columbia University School of Social Work, New York, NY, United States

Christopher E Jones

The University of Alabama at Birmingham, Office of Diversity, Equity, and Inclusion, Birmingham, AL, United States

Leila Kahwati

School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States; and RTI international, North Carolina, United States

Vimala Judy Kamalodeen

School of Education, University of the West Indies, St. Augustine, Trinidad and Tobago

George Kamberelis

Education Department, Western Colorado University, Gunnison, CO, United States

Heather Kane

School of Professional Advancement, Tulane University, New Orleans, LA, United States; and RTI international, North Carolina, United States

Daniel Kim

College of Business, Idaho State University, Pocatello, ID

Mirka Koro

University of Helsinki, Helsinki, Finland

Udo Kuckartz

Philipps University Marburg, Berlin, Germany

Stephanie Kumpunen

Department of Applied Health Research, University College London (UCL), London, United Kingdom

Lamoine Samuel Lee

School of Education, University of the West Indies, Mona, Jamaica

Nancy L Leech

University of Colorado Denver, Denver, CO, United States

Linda Liebenberg

Faculty of Graduate Studies, Dalhousie University, Halifax, NS, Canada

David Littlefair

Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom

Tondra L Loder-Jackson

The University of Alabama at Birmingham, School of Education, Birmingham, AL, United States

Krista Loogma

Tallinn University, Tallinn, Estonia

Teija Löytönen

University of Arts Helsinki, Theatre Academy, Helsinki, Finland

Joseph Alex Maxwell

George Mason University, Fairfax, VA, United States

Joanne Mayoh

Bournemouth University Business School, Bournemouth University, Poole, United Kingdom

Philipp AE Mayring

Alps-Adria University, Association for Supporting Qualitative Research ASQ, Klagenfurt, Austria

Whitney N McCoy

Duke University, Durham, NC, United States

Matthew T McCrudden

Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA, United States

Channel C McLewis

Department of Education, University of California. Los Angeles, CA, United States

Donna M Mertens

Gallaudet University, Rockville, MD, United States

Helen Meyer

University of Cincinnati, School of Education, Cincinnati, OH, United States

Paul Mihás

Howard W. Odum Institute for Research in Social Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Rain Mikser

Tallinn University, Tallinn, Estonia

Kirstin Mulholland

Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom

Camille Nakhid

Auckland University of Technology, Auckland, New Zealand

Bonnie Kaul Nastasi

Tulane University, New Orleans, LA, United States

Isadore Newman

University of Akron, Akron, OH, United States

Katrin Niglas

Tallinn University, Tallinn, Estonia

Patrick Onghena

KU Leuven, Faculty of Psychology and Educational Sciences, Leuven, Belgium

Alfredo Ortiz Aragón
University of the Incarnate Word, San Antonio, TX, United States

Tiiu Õun
Tallinn University, Tallinn, Estonia

Michael Quinn Patton
Claremont Graduate University, Claremont, CA, United States

Sandra L Pensoneau-Conway
Southern Illinois University, Carbondale, IL, United States

Vicki L Plano Clark
School of Education-Research Methods, University of Cincinnati, Cincinnati, OH, United States

Pierre Pluye
Department of Family Medicine, McGill University, Montreal, QC, Canada

Cheryl N Poth
Department of Educational Psychology, Faculty of Education, University of Alberta, Center for Research in Applied Measurement and Evaluation, Edmonton, AB, Canada

Alessandro Pratesi
Department of Social and Political Sciences, University of Florence, Florence, Italy

Tim Rapley
Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom

Linda Roussel
Cizik School of Nursing at UTHHealth, The University of Texas Health Science Center

Stefan Rädiker
Methoden-Expertise.de, Verden, Germany

Anne Ryen
University of Agder, Norway

Jean J Ryoo
University of California, Los Angeles, CA, United States

Janet Salmons
Research Community Manager SAGE Methodspace, SAGE Publishing, London, United Kingdom

Anabel Fernández Santana
COMET - Te Hononga Akoranga, Auckland, New Zealand

Vanessa Scherman
Department of Psychology of Education, UNISA, Pretoria, South Africa

Judith Schoonenboom
University of Vienna, Department of Education, Vienna, Austria

J Schostak
Manchester Metropolitan University, Manchester, United Kingdom

Peggy Shannon-Baker
Georgia Southern University, Statesboro, GA, United States

Audra Skukauskaitė
University of Central Florida, Orlando, FL, United States

Brigitte Smit
The University of Johannesburg, Faculty of Education, South Africa

Lyric Smith
Georgia Southern University, Statesboro, GA, United States

Robert Stake
University of Illinois at Urbana-Champaign, Urbana, IL, United States

Alycia J Sterenberg Mahon
Western Michigan University, Kalamazoo, MI, United States

Ernest Stringer
Independent Scholar, Fremantle, West Australia, Australia

Lavinia Suci
Politehnica University Timisoara, Timisoara, Romania

Mehdi Taghipoorreynh
University of Applied Science and Technology (UAST), Tehran, Iran

Chihiro Tajima
Mixed Methods Program, Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States; and Gakushuin Women's College, Shinjuku, Tokyo, Japan

Abbas Tashakkori
Professor Emeritus, University of North Texas, Denton, TX, United States

Inge Timoštšuk
Tallinn University, Tallinn, Estonia

Meril Ümarik
Tallinn University, Tallinn, Estonia

Cecilia Vindrola-Padros
Department of Targeted Intervention, University College London (UCL), London, United Kingdom

Merel Visse
Drew University, Madison, NJ, United States

Susanne Vogl
University of Stuttgart, Department of Social Sciences, Stuttgart, Germany

Janet B Walton
College of Education, NC State University, Raleigh, NC, United States

Daphne C Watkins
University of Michigan, School of Social Work, Ann Arbor, MI, United States

Marcus B Weaver-Hightower
Virginia Tech, Blacksburg, VA, United States

Leemoy Weaver
Sociology, Psychology and Social Work, University of the West Indies, Mona, Jamaica

Steve R Weaver
The UWI School of Nursing, Mona, Jamaica

Alyson Welker
Department of English, Colorado State University, Fort Collins, CO, United States

Howard White
Evaluation and Evidence Synthesis, Global Development Network, Delhi, New Delhi, India

Charletta N Wiggins
The University of Alabama at Birmingham, The Classic Upward Bound Academy, Birmingham, AL, United States

Rachel E Williams
Sport Psychology and Motor Behavior, University of Tennessee, Knoxville, TN, United States

Nancy P Wingo
School of Nursing, The University of Alabama at Birmingham, Birmingham, AL, United States

Boni Wozolek
Penn State University, Abington College, Abington, PA, United States

Na Yang
University of Cincinnati, School of Education-Second Language Acquisition, University of Cincinnati, Cincinnati, OH, United States

Ting Yang
Independent Evaluation Consultant, Beijing, China

This page intentionally left blank

CONTENTS OF VOLUME 12

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 12: Qualitative, Multimethod, and Mixed Methods Research</i>	<i>xix</i>
<i>Contributors to Volume 12</i>	<i>xxiii</i>

PHILOSOPHICAL AND THEORETICAL UNDERPINNINGS OF QUALITATIVE RESEARCH

Philosophical and theoretical underpinnings of qualitative research <i>Eileen Boswell and Wayne A Babchuk</i>	1
The history of qualitative research in education <i>Martyn Hammersley</i>	14
Qualitative research approaches and their application in education <i>Peggy M Delmas and Rachel L Giles</i>	24
Ethics and politics in qualitative research <i>Anne Ryen</i>	33
Critical theory in qualitative research in education <i>Jean J Ryoo and Jenifer Crawford</i>	45
Feminist and critical approaches to gender in qualitative research <i>Channel C McLewis, Justin A Gutzwa, and S Hurtado</i>	55
Critical race theory and educational research utilizing qualitative methods <i>Tondra L Loder-Jackson, Mary Ann Bodine Al-Sharif, Christopher E Jones, Martez D Files, and Charletta N Wiggins</i>	67
Queer theory and qualitative research <i>Boni Wozolek</i>	78

QUALITATIVE RESEARCH APPROACHES AND DESIGNS

Case study research <i>Robert Stake and Merel Visse</i>	85
--	----

Ethnography: foundations, challenges, and spaces of possibilities <i>Audra Skukauskaitė</i>	92
Autoethnography: storied scholarship <i>Sandra L Pensoneau-Conway</i>	102
Grounded theory <i>Wayne A Babchuk and Eileen Boswell</i>	107
Narrative inquiry <i>Simmee Chung and D Jean Clandinin</i>	123
Phenomenology and education: researching pedagogical experience <i>Norm Friesen</i>	131
Qualitative research approaches and designs: discourse analysis <i>Lavinia Suci</i>	141
Qualitative evaluation <i>Michael Quinn Patton</i>	150
Arts-based research: inquiry through, in, and as art <i>Janinka Greenwood</i>	159
Responding to cultural diversity through action and community-based participatory approaches in qualitative research <i>Rosalind Beadle, Alfredo Ortiz Aragón, and Ernest Stringer</i>	169
Affirming methodologies—the Caribbean oral tradition of Ole Talk <i>Camille Nakhid and Anabel Fernández Santana</i>	179
Semiotics in the context of qualitative research <i>Marcel Danesi</i>	189

QUALITATIVE RESEARCH METHODS

Qualitative sampling and qualitative data <i>Mirka Koro, Nikki Fairchild, Angelo Benozzo, and Teija Löytönen</i>	198
Qualitative interviews <i>Svend Brinkmann</i>	210
Imagining and reimagining focus group research in education <i>Alyson Welker and George Kamberelis</i>	219
Naturalistic observation <i>J Schostak</i>	250
Photovoice with children and youth <i>Helen Meyer, Laura Dell, and Randall Gibson</i>	260
Participatory photography: high school students' awareness of privileges in Nordic welfare states <i>Gunilla Holm</i>	270
Document analysis <i>Tim Rapley, Kevin Ardron, David Littlefair, Kirstin Mulholland, and K Neil Jenkins</i>	285
The ontological imperative in conducting qualitative research in online spaces <i>Hannah R Gerber</i>	292

Qualitative research methods: approaches to qualitative data analysis	302
<i>Paul Mihas</i>	
Qualitative content analysis	314
<i>Philipp AE Mayring</i>	
Rapid qualitative research	323
<i>Georgia Chisnall, Stephanie Kumpunen, and Cecilia Vindrola-Padros</i>	
Establishing quality in qualitative research: trustworthiness, validity, and a lack of consensus	336
<i>Leia K Cain, Rachel E Williams, and Victoria Bradshaw</i>	
Triangulation: uses, abuses and recent developments	351
<i>Alessandro Pratesi</i>	

MIXED METHODS AND MULTIMETHOD APPROACHES TO RESEARCH

Mixed methods and multimethod research in education: six design decisions for collecting, combining, and developing datasets	361
<i>Judith Schoonenboom</i>	
Foundations of mixed methods research	372
<i>Abbas Tashakkori and Isadore Newman</i>	
Philosophical underpinnings of mixed methods research in education	380
<i>Peggy Shannon-Baker</i>	
Frameworks for conceptualizing mixed methods research	390
<i>Vicki L Plano Clark, Nataliya V Ivankova, and Na Yang</i>	
Transformative mixed methods research	402
<i>Donna M Mertens</i>	
Integration in mixed method research	411
<i>Elizabeth G Creamer</i>	
Cultural studies in education: insights for the mixed methods researcher	417
<i>Loraine D Cook, Vimala Judy Kamalodeen, Ingrid E Hunt-Anderson, Tashane Haynes-Brown, Leemoy Weaver, Steve R Weaver, and Lamoine Samuel Lee</i>	

DESIGNING INTEGRATIVE MIXED METHODS RESEARCH

Mixed methods research designs	426
<i>Michael D Fetters and Chihiro Tajima</i>	
Designing integrative mixed methods research	441
<i>Joseph Alex Maxwell</i>	
Ten mixed methods integration strategies for obtaining a detailed understanding	450
<i>Judith Schoonenboom</i>	
Convergence and divergence in mixed methods research	462
<i>Pierre Phuye and Quan Nha Hong</i>	
Thinking qualitatively in the context of mixed methods research: implications for study design	478
<i>Brigitte Smit, Vanessa Scherman, and Linda Liebenberg</i>	
Mixed methods sampling and data collection	485
<i>Nancy L Leech and Courtney Donovan</i>	

Strategies to integrative mixed methods analysis <i>Susanne Vogl</i>	491
Using software for mixed methods analysis <i>Udo Kuckartz and Stefan Rädiker</i>	500
Role of quality criteria in integrative mixed methods research <i>Kathleen MT Collins</i>	513
Promising practices for addressing unique ethical dilemmas arising from integration in mixed methods research <i>Cheryl N Poth, Elizabeth Creamer, and Leia K Cain</i>	522

INTERSECTION OF MIXED METHODS RESEARCH WITH OTHER APPROACHES

Mixed methods and evaluation <i>Donna M Mertens</i>	531
Combining mixed methods and case study research <i>Vicki L Plano Clark, Lori A Foote, and Janet B Walton</i>	538
Intersecting mixed methods with phenomenology <i>Joanne Mayoh, Jason Colditz, and Nollaig Frost</i>	550
Mixed methods and grounded theory <i>Elizabeth G Creamer</i>	558
Intersecting mixed methods with action and community-based participatory research approaches <i>Nataliya V Ivankova and Nancy P Wingo</i>	567
Mixed methods and qualitative comparative analysis <i>Heather Kane and Leila Kahwati</i>	581
Intersecting mixed methods with translational research: implications for educational research and practice <i>Nataliya V Ivankova, Jami L Anderson, Ivan I Herbey, Linda Roussel, and Daniel Kim</i>	588
Mixed methods in education RCTs <i>Howard White and Ting Yang</i>	599
Mixed methods and the Delphi method <i>Mehdi Taghipoorreyneh</i>	608

APPLICATIONS OF MIXED METHODS RESEARCH IN EDUCATION

Mixed methods research design practices to address complexity in education <i>Cheryl N Poth and Emma P Bullock</i>	615
Conducting large-scale mixed methods studies in education <i>Katrin Niglas, Rain Mikser, Krista Loogma, Meril Ümarik, Inge Timoštšuk, and Tiia Õun</i>	626
Advancing education research through mixed methods with existing data <i>Daphne C Watkins and Natasha C Johnson</i>	636
Using mixed methods research with Geographic Information Systems (GIS) <i>Nazgol Bagheri</i>	645

Mixed methods research synthesis for literature reviews <i>Patrick Onghena</i>	655
Mixed methods in survey research <i>John H Hitchcock and Allison Crean Davis</i>	666
Mixed methods in culturally tailored interventions <i>Bonnie Kaul Nastasi</i>	678
INTERSECTION OF MIXED METHODS RESEARCH WITH OTHER APPROACHES	
Mixed methods and social network analysis <i>Dominik E Froehlich</i>	685
APPLICATIONS OF MIXED METHODS RESEARCH IN EDUCATION	
Mixed methods and web-based technologies for data collection <i>Janet Salmons</i>	693
Integrating art and artistic practices in mixed methods research in education <i>Lyric Smith and Peggy Shannon-Baker</i>	702
VISUALIZING AND WRITING-UP QUALITATIVE, MULTIMETHOD AND MIXED METHODS RESEARCH	
Developing qualitative research proposals <i>Marcus B Weaver-Hightower</i>	711
Writing up qualitative research <i>James Drisko</i>	721
Mixed methods research proposal <i>Jessica T DeCuir-Gunby and Whitney N McCoy</i>	729
Reporting mixed methods research studies <i>June E Gothberg and Alycia J Sterenberg Mahon</i>	739
Joint displays in mixed methods research <i>Timothy C Guetterman</i>	754
Visual displays in mixed methods research: integration of methods <i>Matthew T McCrudden</i>	763

This page intentionally left blank

Philosophical and theoretical underpinnings of qualitative research

Eileen Boswell^a and Wayne A. Babchuk^b, ^a Department of Teaching, Learning, and Teacher Education, University of Nebraska-Lincoln, Lincoln, NE, United States; and ^b Department of Educational Psychology, University of Nebraska-Lincoln, Lincoln, NE, United States

© 2023 Elsevier Ltd. All rights reserved.

Chapter introduction	2
Chapter overview	3
Part I. Philosophical assumptions of qualitative research	3
Introduction to Part I	3
Distinguishing ontology from epistemology	3
Logics employed in qualitative research	4
Generalizability and transferability	4
Power, positionality, and reflexivity in qualitative research	4
Summary of Part I	5
Part II. Theory and theories in qualitative research	5
Introduction to Part II	5
Historical roots of grand theories	5
Major taxonomies of qualitative research	6
Qualitative research after positivism	6
Summary of Part II	7
Part III. Applying theory and philosophy in qualitative research	7
Values as manifest in theoretical frameworks	8
Axiological foundations of specific qualitative methodologies	8
Ethnography	8
Phenomenology	8
Grounded theory	9
Additional qualitative sub-approaches as values-driven	9
Operational values in specific methods and procedures	10
Sampling	10
Interviews	10
Participant observation	10
Member checking	10
Memo-writing and bracketing	10
Importing qualitative sensibilities into mixed methods research	11
Summary of Part III	11
Coda: future directions in qualitative research	11
References	12
Further reading	13

Glossary

Abduction inferring an apt conclusion based on the iterative evaluation of data against possible, plausible explanations

Audit trail a set of documents and resources showing the timing and rationale for key decisions in the research process

Axiology the nature of, and study of, values or of the origin of the values conferred on something

Bracketing the practice of documenting—and setting aside—personal views and experiences related to the subject of inquiry, particularly popular in some forms of phenomenological research

Deduction reasoning from the general to the specific, often in the form of strict hypothesis testing that aims to predict something about a population based on a sample

Emic denotes the ‘insider’ perspective of participants at the core of the study

Epistemology the study of knowledge and ways of knowing, and of the nature of evidence or proof

Ethnography the study of a culture-sharing group or some aspect thereof; a term representing both a process and a product of research (*an ethnography*)

Etic denotes the ‘outsider’ perspective of the researcher

Induction reasoning from the specific to the general, often in the form of exploratory research questions that seek to build more knowledge about an understudied phenomenon or community

Member checking any of multiple respondent validation procedures that the researcher uses to verify their findings and interpretations

Memo reflexive documentation of a researcher's insights, methods, decisions, and possible biases throughout a project

Methods specific procedures chosen in alignment with the guiding philosophical assumptions and central questions of the researcher

Methodology a comprehensive approach to research that encompasses the abstract and concrete components of research design and implementation including ontology, epistemology, axiology, and methods

Ontology the study of perceptions of reality and the nature of being

Paradigm an overarching worldview that guides research processes and encompasses a researcher's ontology, epistemology, axiology, and methodology

Positionality a researcher's practice of considering the implications of power dynamics in the research process and how these dynamics produce advantages, disadvantages, biases, or limitations

Positivism the paradigmatic view that a singular, objective reality exists and that it can be accurately observed and measured via scientific inquiry

Post-positivist describes the assertion that a singular objective reality exists but can only be approximated through scientific inquiry

Qualitative research an umbrella term used to designate related approaches that emphasize inductive reasoning, collecting data in natural settings, and understanding participants' points of view

Reflexivity the process whereby a researcher reflects on their own social location, possible biases, and unique insights

Symbolic Interactionism the paradigmatic view that assumes one's social experience are mediated through social interaction mediated by personal and social interpretations of language and other symbols

Teleology the study of time, completion, temporality; perspectives on the contemplation of the end-points of an event or state of being

Theory a way of explaining observed reality and proposing relationships among concepts, people, groups, or sets of conditions

Chapter introduction

Qualitative inquiry is a naturalistic approach to empirical research, one that privileges a participant's perspective over that of the researcher. Qualitative practitioners affirm that the expertise that inheres in individuals and communities holds a wealth of potential for social research. By plumbing the depths of people's lived realities, qualitative research constitutes an equal and worthy counterpart to quantitative sciences and other research methodologies. **Qualitative research** is "an umbrella term used to designate a family of approaches that emphasize inductive reasoning, collecting data in natural settings, and understanding participants' points of view" (Babchuk, 2019, p. 1). Although there are arguably minor distinctions between the terms *qualitative inquiry* and *qualitative research*, these terms are used synonymously throughout this chapter.

Sometimes the task of understanding an abstract set of concepts is made easier by contrasting these concepts with related, but divergent ideas. Although most qualitative scholars view qualitative inquiry as equally valuable to quantitative research, it can be useful to describe the two in terms of how they differ, in order to more fully emphasize the unique features of a qualitative orientation that set it apart and add to its value. Therefore, passages that set qualitative approaches against a backdrop of quantitative science do so only in the hope that readers more familiar with quantitative designs can draw on this knowledge to learn the dialectical ways to understand what qualitative research offers. Given that the history of qualitative research developed partly as a response to the dominance of quantitative approaches, the two are juxtaposed at times throughout this chapter in order to give readers opportunities to compare and contrast them before the introduction of mixed methods at the conclusion of the chapter.

What distinguishes a qualitative approach from others is an emphasis on the perspective and agency of participants in a particular population or culture-sharing group, and sometimes a commitment to participants' direct participation in planning, designing, and implementing research. Additionally, qualitative science is distinct from quantitative science in that it does not rely on hypotheses to guide the research process, nor does it typically attempt to establish causal connections among variables. Erickson (2018) has written that, "The qualitative researcher first asks, 'What are the kinds of things (material and symbolic) to which people in this setting orient as they conduct everyday life?'" (p. 36). This style of inquiry prioritizes language and discourse, both in data collection methods such as extensive interviews as well as in creating research reports and other deliverables that are produced as outcomes of a study.

Participant observation methods (see *Naturalistic Observation*, this volume) and unstructured or semi-structured interviews are popular methods of qualitative data collection that produce the rich, detailed verbiage and "thick description" that are trademarks of qualitative research (Geertz, 1973). Prolonged engagement in the field and immersive, comprehensive data analysis are part and

parcel of a qualitative approach, as are researchers' candid reflections on their role in shaping research findings and interpretations thereof. In the qualitative realm, researchers acknowledge that people are experts on their own lives (Delpit, 1988, p. 297). The participant functions as the expert in the study. This contrasts with quantitative paradigms wherein the researcher commonly adopts a top-down, paternalistic stance. This stance has been imagined as the distant or impersonal researcher donning the mantle of power to administer a 'treatment' to a 'subject' or treating people as anonymous, extracted data sets.

A qualitative researcher strives to adopt a deeply humanist ethos and a profound respect for the participants as complex and holistic beings worthy of dignity throughout the research process. This is sometimes referred to as practicing a duty of care, or an "ethic of care" (Matteson and Lincoln, 2009). Honor, respect, and a collaborative spirit join the qualitative researcher with participants in a dialogic exchange that does not shy away from the fact that the research experience itself is intersubjective. An interview, for example, is a co-construction of knowledge between individuals socialized through intersecting lenses of class, race, gender, culture, and many other aspects of biography that fashion both interviewer and respondent perspectives.

Chapter overview

Part I of this chapter explains some philosophy concepts that are foundational to the assumptions guiding qualitative designs. In Part II, these assumptions are shown to have guided the historical development of overarching grand theories and paradigmatic taxonomies. The application of theory and philosophy to contemporary qualitative methodology is taken up in Part III, including how such values and practices are incorporated into mixed methods research. The chapter concludes with some ideas about future directions within qualitative inquiry.

Part I. Philosophical assumptions of qualitative research

Introduction to Part I

Imbricated with some of the broad theories in qualitative inquiry are the guiding philosophical assumptions that underly most qualitative projects, even if they remain unacknowledged or minimally articulated. These include a preference for inductive over deductive reasoning, a commitment to reflecting on the researcher's unique skills and possible limitations, candor and transparency about the research process, and humanist ethics that course throughout all qualitative inquiry. The term *qualitative research* entered the social sciences in the 1960s (Bogdan and Biklen, 2007, p. 2). Early qualitative innovators emphasized "human flexibility, the importance of the genesis of the self, the definition of the situation, and the role of the community in the social process" (Deegan, 2007, p. 14). Rather than isolating variables in a laboratory for precise quantitative measurement, qualitative approaches embrace the complexity of the human condition and endeavor to understand the fundamental processes and conditions that knit people together into communities and culture-sharing groups. These features position qualitative researchers to offer detailed and context-specific information about known and novel situations across diverse communities and in variegated circumstances of life and work. A qualitative orientation is moreover a natural outgrowth of many academic disciplines, provided that the specific philosophical assumptions of a discipline are interrogated before designing research.

Distinguishing ontology from epistemology

Questions about the nature of knowledge behoove us to make ourselves aware of the limits of our knowledge (Schneeg, 2014, p. 46). **Epistemology** is the study of ways of knowing, and of what counts as evidence in various research paradigms. Epistemologies are not the only philosophical basis for qualitative research approaches, however. Axiology, teleology, and especially **ontology**—the nature of being and the study of that which is 'real'—all matter as philosophical pillars supporting the epistemological foundations of a research study. Epistemology presupposes ontology, but many students of qualitative research learn these in reverse order. First, one must conceive of an entity and its indicators in whatever reality one adopts. Only then can one consider the epistemological axes of that reality—how we come to know of, and prove, its existence. These axes include the nature of facts, knowledge, evidence, and observation: what counts as knowing and how we go about gathering and verifying knowledge.

Humans are socialized to accept some forms of knowledge uncritically, and thus epistemology is often obscured by strong social norms reinforcing what is taken for granted by one's surrounding community. For example, the concept of a 'day' has ontological roots in the regular cycle of the sun; however, dividing this concept into 'hours' reveals how society has organized and expressed our experience of the solar cycle—one which we could easily conceptualize in other ways. Whereas solar and lunar cycles are ontological, the concept of a seven-day week is purely epistemological; there is nothing inherent in nature that distinguishes a weekday from a weekend. Only habit, indoctrination, and social organization accomplish this cultural feat.

A 'weekday' is an episteme: a tool of thought that humans use to think about and talk about reality. Some tools of thought are innocuous but no less revealing of the biases limiting those who coin them. An innocent example that illustrates the creation and application of an episteme is the term 'ornamental shrub'—it defines whole categories of plant species based on human uses (or lack thereof) for those species, rather than in line with botanical characteristics. By circulating these terms, the creators of such nomenclature trade ontology for epistemology: they conflate the nature of the thing itself with how they label that thing within their chosen, culturally conditioned, and narrow-sighted knowledge systems.

The difference between ontology and epistemology might at first strike readers as a heady philosophical conundrum or perhaps a trifle, but qualitative researchers would be well-advised to devote some energy to clarifying which concepts and definitions are embedded in a project's design premise and whether the participants, funders, reviewers, and other stakeholders share those ideas. It is also crucial to anticipate the offense taken, or even backlash, when controversial assumptions belie an ethnocentric or otherwise limited bias that undermines the inclusion and emancipation at the heart of qualitative research.

Logics employed in qualitative research

Accompanying the static assumptions that qualitative researchers bring into their work with them are philosophical processes, or logics, that structure their reasoning and decision-making throughout a project. A **logic** is merely a system of inputs and outputs, and more than one type of logic might be needed to move the project—especially the data analysis—from beginning to end. Two basic logical processes are **induction** and **deduction**, with the latter being the default approach to quantitative science. The former, inductive reasoning, is a signature element of qualitative research, and there is room for other forms of reasoning to support induction, and even to validate it. These include abduction, retroduction, and transduction (see [Lipscomb, 2012](#); [Belfrage and Hauf, 2017](#); and [Visokolskis, 2021](#), respectively).

Although the terms are applied slightly differently in syllogistic and rhetorical applications than in this chapter, the basic meanings of induction and deduction hold in qualitative research: induction moves from the specific to the general while deduction moves from the general to the specific. Whereas deductive reasoning is well suited to a quantitative project that aims to predict the behavior of a population based on the results of a sample, inductive reasoning provides a path for gathering specific data, analyzing it, and then making generalizations that may or may not be predictive of people outside the study sample. Induction may be used to generate hypotheses, and deduction is apt for testing those hypotheses. Induction restrains itself from presupposing too much about the sample or the data and is better applied to research questions and problems requiring intellectual exploration. Deduction, conversely, would presuppose a great deal in order to invent a treatment, experiment, or intervention that hopes to confirm (or deny) a pre-established prediction through specific hypotheses identified prior to data collection.

In qualitative research generally—and grounded theory studies in particular—an additional form of logic, **abduction**, is also practiced ([Bryant, 2017](#); [Hadley, 2017](#)). Abductive reasoning channels one's ability to infer a well-fit, plausible conclusion based on available data and to eliminate ill-fit conclusions along the way. While it cannot produce a neat or tidy logical proof, abduction can offer solid reasoning for how one arrived at one's conclusions. Abductive reasoning, such as depicted in Babchuk and Boswell's theory-tracker tool (see *Grounded Theory*, this volume), demands that a researcher iteratively induce and deduce as data are collected and analyzed. Data that ground a theory are examined through a filter of possible/plausible theoretical explanations; from there, the explanations that lack adequate support are abandoned until a well supported or 'saturated' theory can be achieved—one that never loses sight of the need to be 'fit' to the data collected.

Generalizability and transferability

Unlike a quantitative approach, which aims to generalize results from a sample to a population, a qualitative approach seeks nuanced conclusions about a relatively small sample and is not explicitly designed to extend findings to the general population. Nevertheless, a qualitative researcher may assert some generalizations about the sample, or about the application of a particular theoretical framework to a research problem. Thus, it would be an oversimplification to say that qualitative projects do not generalize—they do, but not in the same way that quantitative projects do, and not nearly as much. In [Lincoln and Guba's \(1985\)](#) iconic work on trustworthiness, they used the term *transferability*, a term which in qualitative research loosely parallels that of *generalizability* in quantitative research, and they suggest strategies to achieve it (see for example, [Merriam and Tisdell, 2016](#)). It is therefore incumbent on qualitative researchers to explain how their findings do and do not apply to other samples or sub-populations, as well as to reflexively critique their own methods and assumptions. Moreover, qualitative researchers must document these decisions and assumptions through an audit trail, as described next.

Philosophically, qualitative research departs from quantitative and other positivist-influenced approaches in that the researcher is considered the "instrument" of research ([Creswell, 2014](#), p. 185) and can be critiqued in this capacity. One way these assessments are undertaken is to evaluate the **audit trail** chronicling the research process. In the event of an audit for any reason, the researcher needs to be able to produce evidence and documentation of each step of the data collection and analysis. This could be in the form of voice memos, correspondence with an advisor as major decisions are taken, read-only data sets that have not been emended, audio/video recordings of interviews, correspondence with key informants, or other items that show the steps and rationale for the project. Establishing which methods will be used should be done early in the process and with transparency as the guiding principle, as opposed to reproducibility or replication, which may the goal of an audit trail or research protocol in the quantitative sciences.

Power, positionality, and reflexivity in qualitative research

Among the philosophical assumptions underlying any qualitative enterprise is that the participant is the expert and the researcher must subordinate themselves. Astute qualitative researchers take note of power dynamics through a reflexive process of positioning themselves in the research. The concepts of positionality and reflexivity thus form a "qualitative imperative" ([Boswell, 2021b](#), p. 13)

in social research. In addition to the ‘usual suspects’ of race, class, and gender, one’s research is also subject to both additive and subtractive properties of one’s religious upbringing and activities; military or veteran status; housing, mobility, or incarceration circumstances; caste or tribal affiliation; and numerous manifestations of privilege and birthright. Charmaz (2014) moreover argues that, “Data are never entirely raw. Recording data alone confers interpretations on them because we frame these data through our use of language and understandings about the world” (p. 54). Being reflective about uses of language is an excellent starting point for the practice of reflexive qualitative research.

Reflexivity is a necessary component of the qualitative research process, one that provides “heightened awareness of the self in the process of knowledge creation, and a clarification of how one’s beliefs have been socially constructed” (Grbich, 2013, p. 113). Defined as “researching the self, researching the self in relation to others, engaged reflection and representation, and shifting from the self to system” (Milner, 2007, p. 388), reflexivity in the design and conduct of qualitative research additionally entails scrutinizing one’s positionality as a researcher. **Positionality** is “thinking about the ways in which our social positions shape not only our interactions in the field but also what we see and hear” (Reyes, 2018, p. 212). Disclosures of a researcher’s position relative to participants, data, and the research process are a needed ingredient in strong qualitative studies. These are necessary inclusions because qualitative research “locates the observer in the world” (Strauss and Corbin, 1994, p. 3). Without positionality and reflexivity, research is structured around “inquirer-oriented power” (Cannella and Lincoln, 2011, p. 82). Reflexivity and positionality are therefore paramount in qualitative research, and omitting a disclosure of investigators’ personal contexts is contrary to many best practices within social research.

Summary of Part I

Part I of this chapter has outlined some general philosophical concepts that undergird qualitative approaches to research, irrespective of which theoretical framework(s) one might choose for a research project. In the following sections, the basics of ontology and epistemology will be shown to configure various grand theories that have been instrumental in proliferating the qualitative applications that will be described in the third part of the chapter.

Part II. Theory and theories in qualitative research

Introduction to Part II

There exist broad theoretical and philosophical foundations to qualitative research generally, as well as narrower theoretical frameworks that led to the formation of specific research traditions. This chapter focuses on the broad, macro-level theories, with selected examples of lower-level theories integrated to exemplify contemporary qualitative applications. For instance, many qualitative projects could employ feminism as an overarching theoretical foundation and the project might also use theories within a discipline, such as assessment theories within the field of education or patient-compliance theories within the health sciences.

A **theory** is any articulated way of explaining observed reality—one that usually proposes a relationship between two or more concepts, entities, or sets of conditions. Theories that gain a lot of publicity, whether because they are universally accepted or highly controversial, might be known to a great segment of the general public (e.g., the theory of natural selection) whereas other theories may only be known to specialists in particular fields within a discipline (e.g., theories of how best to teach world languages to adult learners). Throughout this chapter, the emphasis is on ‘grand theories’ or formal, macro-level theories that propose large-scale and wide-sweeping social implications of the status quo.

Historical roots of grand theories

The history of qualitative research is one lens through which to peruse the trajectory of disciplinary communities and divisions, and of conceptions of research from the Enlightenment through the new millennium. Some of the major changes that have contributed to the current state of qualitative research include macro-theoretical assumptions guiding a study; the race, gender, and ethnic background of the researchers as well as of the participants; an increased focus on the potential reciprocal benefits for researchers and participants; and the socio-historic and geopolitical contexts in which research is conceived, funded, and undertaken. These forces continue to sculpt research methodologists and the students to whom they impart their methodologies in a brave new world in which data objects can be tweets, interviews can be conducted remotely, and all manner of artifacts can be crowdsourced in an instant. The *zeitgeist* or ‘spirit of the times’ hews all academic trends to some extent, insofar as it renders certain topics fundable or unfundable, alluringly polemic, or exceedingly popular when a given topic or methodology garners ample airtime at conferences and in themed publications.

Among the major players whose scholarship influenced the early formulations of qualitative social research are Franz Boas, Bronislaw Malinowski, Ernest Burgess, Robert Park, Clifford Geertz, and Margaret Mead. The academic climate in the 1850–1900s in particular was a lightning rod for the burgeoning research approaches that characterize qualitative inquiry today, as influenced by the writings of Weber, Dilthey, Simmel, Husserl, Marx and other timeless icons of the late 19th and early 20th centuries. Just as the *zeitgeist* had an influence on qualitative inquiry over the last century and a half, the *ortgeist* or ‘spirit of the place’ was also of import. In the United States, new directions in **methodology** took researchers from ethnographic field studies in distant locales to research at home in urban and rural settings, where they discovered more and more about the American way of life and the people living it.

Crime (and criminal justice) as a focus of study saw its heyday as qualitative research blossomed in Chicago and other urban centers, while studies of race and poverty knew no limits—rural or urban. Migration trends in the early twentieth century aided social scientists in crafting innovative research questions about integration, segregation, and assimilation in U.S. neighborhoods, in addition to casting new light on the intricacies of public spaces as more and different cultures intermixed therein. Education research moved these questions into the schoolhouse, with more studies focusing on newcomer students and their immigrant families.

The Chicago School of Sociology, together with John Dewey's pragmatic and experiential focus in education, set the stage for deeper examination of individuals' lived experiences as the center of inquiry. The school of thought known as **Symbolic Interactionism** (Mead and Schubert, 1934; Blumer, 1969, 1980) adapted ethnographic methods to "explain why both actors and processes are as they are" (Vidich and Lyman, 2000, p. 38). This focus on the 'actor' highlighted the individual psyche as a worthwhile avenue for qualitative exploration. Symbolic Interactionists contended that one's social experience is mediated by personal and social interpretations of language and other symbols, leading to a watershed for methodologies that prioritized "the meaning people give to their experience" (Bogdan and Biklen, 2007, p. 27). Allowing for subjective and interpretive definitions of reality proved to be the ontological bell that could not be unrung: interpretivist paradigms led to the growth of an experience-centered methodology known as phenomenology.

Throughout this era, thick description and detailed documentation of human behavior were the order of the day in qualitative research. Ethnographers in both sociology and anthropology took pains to analyze multiple sources of data including interviews, maps, audio samples, photography, and meticulous biographical and kinship data—thereby establishing lineage, and linking habits with heritage. Completeness of data and compelling detail became standard as more and more populations under examination were seen to be novel, neglected, or otherwise warranting intensive study. During this period, women were beneficiaries of the propagation of new and exciting pathways in qualitative research—both as practitioners of the research process and as groups/individuals newly represented in qualitative studies. Feminist research, among other major paradigms, is taken up in the following sections.

Major taxonomies of qualitative research

Some theoretical perspectives have been sufficiently developed to have birthed their own methodologies, research communities, journals, conferences, and social movements. These high-level frameworks can be referred to as **paradigms**. A paradigm is a "unitary package of beliefs about science and scientific knowledge" (Crotty, 1998, pp. 34–35). Students new to qualitative research could also think of a paradigm as an overarching worldview that guides research processes and incorporates a researcher's ontology, epistemology, axiology, and methodology. The postmodern paradigm, for example, entails particular beliefs about the nature of reality, what counts as evidence or knowledge, what principles are valued in research and in society, as well as ways to go about studying social phenomena.

Authors of popular works on qualitative inquiry delineate their taxonomies of research paradigms in many different ways. Merriam and Tisdell (2016) note that, "there is almost no consistency across writers in how this aspect of qualitative research is discussed" (p. 8). For example, Neuman (2011) lists three approaches to social science (positivist, interpretive, and critical social science), each with its own set of philosophical assumptions and principles as well as its own stance on how to do research. He includes feminist and postmodern research as separate categories. Merriam and Tisdell (2016) organize this material into four philosophical traditions (positivist, interpretive, critical, and poststructural/postmodern), while Hatch (2002) compares five research paradigms (positivist, post-positivist, constructivist, critical/feminist, and poststructuralist). Lincoln et al. (2018)—in what can be viewed as the seminal work on this topic—review five (positivism, post-positivism, critical theory, constructivism, and participatory).

Bogdan and Biklen (2007) discuss five sets of theoretical underpinnings of qualitative research (positivism, phenomenology, symbolic interaction, ethnomethodology, contemporary trends), while Creswell and Poth (2018) divide their text into philosophical assumptions (ontology, epistemology, axiology, methodology) and interpretive frameworks (post-positivism, social constructivism, transformative frameworks, postmodern perspectives, pragmatism, feminist theories, critical theory and critical race theory, queer theory, and disability theory) in order to facilitate the reader's understanding of qualitative inquiry. Crotty (1998) presents positivism, post-positivism, constructionism, interpretivism, critical inquiry, feminism, and postmodernism as the major paradigms. Despite considerable variation, all of these scholars argue for the importance of understanding paradigmatic tenets of the research enterprise.

Positivism—the paradigmatic view that a singular, objective reality exists and can be accurately measured and observed—is the oldest of these philosophical traditions with its origins in the work of Hume, Mill, Comte, and Durkheim in the 18th and 19th centuries, and that of sociologists such as Robert Merton and Talcott Parsons in the twentieth (see especially Erickson, 2018; Neuman, 2011). Comte (1858) is credited for coining the term "sociology" to advocate for the scientific study of social behavior (borrowed from lab sciences) that came home to roost in the iconic *The Rules of Sociological Method* (Durkheim, 1895), which stands as the first truly empirical and objective study of social phenomena.

Qualitative research after positivism

In contrast to the well-worn objectivist grooves in turn-of-the-century social sciences, many qualitative researchers beginning in the early twentieth century began to question positivism, to the point of using qualitative research as a response to what was absent

from laboratory settings: the **emic**, or insider's expert knowledge and perspective (Merriam and Tisdell, 2016, p. 16). An excellent example of this is Edmund Husserl, a mathematician, who desired a deeper understanding of shared meaning through qualitative methods than that afforded by quantitative designs. This swing away from objectivist thinking toward abstraction created openings within qualitative research for original, post-positivist epistemological frameworks from subjectivism and interpretivism to constructivism, postmodernism, and critical paradigms that scrutinize the power dynamics underlying all social inquiry (Crotty, 1998; Neuman, 2011).

Many contemporary scholars lean more toward a **post-positivist** worldview that, unlike positivism, assumes there exists a knowable, objective, patterned, and stable reality, and which avers that social scientists can only approximate an objective reality through structured and systematic approaches that mirror those historically considered successful in the natural sciences. Importantly, the so-called 'qualitative revolution' in the social sciences, beginning in the 1960s, ran parallel to a 'quantitative revolution' in anthropology wherein many cultural anthropologists adopted quantitative empiricism in fields such as Human Behavioral Ecology to lend what they perceived as more scientific credibility to the study of human behavior (Chagnon and Irons, 1979). Post-positivist worldviews have also been adopted by many qualitative scholars who advocate rigorous and systematic methods in their research and often favor the use of computer assisted qualitative data analysis software and the counting of codes in their analyses.

In post-positivist qualitative research, many projects are subtly or not so subtly influenced by macro-level theories or 'grand theories' that arose in response to positivism and were enhanced by the variation in both where and when they developed, and by whom. Critical theory, postmodern theory, and feminist theories all bear the imprint of the particular political, class, race, and gender dynamics of the eras in which they came into full use in social research. Critical theory or critical social science is a term for approaches that disrupt the status quo by empowering historically oppressed and marginalized groups through research, and achieve a more democratic and equitable society where all individuals can realize their full human potential. Origins of critical theory are credited to Marx, Freud, Fromm, Adorno, and Marcuse, and more contemporary proponents such as Habermas, Freire, Bourdieu, and Tuhiwai Smith (Neuman, 2011).

Feminism can be located within the main framework of critical theory and, similar to some of the other philosophical orientations discussed above, it critically questions masculinist research that directed early scholarship in the study of social behavior. Scholars in the feminist tradition see positivism and its notions of objectivity, logic, rationality, control, claims of value-neutrality, and individual competition as representative of a male point of view that obstructs and underplays the importance of accommodation, interconnectedness, trust, and mutual obligation in the social activities that bevel human interaction (Creswell and Poth, 2018; Neuman, 2011). Feminist scholars pose research questions that expose power relationships, social positioning, and institutional sexism to help transform society and achieve equitable conditions for all. Feminist researchers often employ action-oriented research methods that build solidarity with participants and that are designed to facilitate personal and societal change. Qualitative scholars in the feminist tradition include Margaret Mead, Kimberlé Crenshaw, Patricia Hill Collins, Mary Field Belenky, Carol Gilligan, and Sarah Ahmed, among many others.

Postmodernism in qualitative research is closely tied to the larger postmodern movement across philosophy, the arts, music, and literature, and it shares some overlapping principles with critical theory and feminism. These include, for example, the use of language as a tool of repression of historically marginalized groups, and the rejection of more traditional approaches to the study of social behavior that is viewed as male-dominated, Eurocentric and biased in helping institutionalize and sustain extant power imbalances. In this vein, postmodernists have viewed traditional explanations as forms of critical fiction or myth in that they were so heavily biased by the authors of ethnographic reports that one needs to deconstruct these texts for layered meanings and reject the notion of a single, objective truth (Creswell and Poth, 2018; Neuman, 2011). To them, multiple truths exist. Postmodernists question the grand narratives put forth in those reports and the objective portrayal of the 'other' as historically seen through a White, masculinist, Western lens.

Interrogating the limited perspectives of European research methodologies in the twentieth century qualitative research boom has led to further critique of the preponderance of social science findings that are based solely on samples assumed to be Western, educated, industrialized, rich, and democratic—or "W.E.I.R.D." as a shorthand (Henrich, 2020; Henrich et al., 2010a,b). Breaking the positivist mindset open enables other sets of values to manifest themselves in qualitative inquiry. The third part of this chapter will take up how values guide the research process.

Summary of Part II

The above sections have outlined how paradigms and specific theories originated and how they continue to usher qualitative inquiry in emancipatory directions, to a greater or lesser extent, depending on one's theoretical inclinations and choices. Major qualitative research paradigms such as feminism weave together particular ontologies and epistemologies, often in reaction to previous paradigms' shortcomings. In the following section on research applications, the role of values—axiology—will be in the spotlight, illuminating how values guide each step of the research process.

Part III. Applying theory and philosophy in qualitative research

Methodological sub-approaches are not born into a vacuum; rather, they are forged in the historical and political contexts of a particular time and place, in response to what has come before and challenges thereto. Historical forces drive which research gets funded,

approved, and ultimately completed and published. Imbued throughout all of these decisions are values—personal, institutional, communal, cultural, and political. The study of values in a philosophy or research context is called **axiology** and—together with ontology and epistemology, as outlined above—axiological drivers of research are integral to one’s understanding of contemporary qualitative applications. Consider, for instance, the value of having one’s research make a positive impact in the world; this is not a foregone conclusion for all qualitative researchers. In fact, some would caution against research as advocacy while others wholeheartedly embrace justice-oriented outcomes as a primary aim of their projects.

While opinions differ on whether qualitative researchers should act as advocates for the communities they conduct research alongside, many believe that qualitative research should strive for dignity, respect, and a generous empathy for all individuals participating in the research process. These values underly qualitative inquiry and are evident in the specific theories, methodologies, and methods chosen by researchers. Whereas a quantitative experiment may be “reductive” of human experience (Richards and Morse, 2012), qualitative approaches view human lives as each being situated in diverse particularities. For this reason, qualitative applications require a thorough and comprehensive commitment on the part of the researcher to truly understand and honor participants, to authentically contemplate the implications of the research, and to mitigate the possible harms—literal or representational—of any study being conducted. Axiology, therefore, can be detected at the levels of theory, methodology, and methods, as explained in the following sub-sections.

Values as manifest in theoretical frameworks

As far back as the sociological writings of Du Bois (1898, 1935, 1897), qualitative inquiry has played a crucial role in conveying the lived realities of African Americans and other systemically oppressed or subjugated communities. Bodies of work leveraging such critical epistemologies include general feminist approaches and their subdomains such as Black feminist thought (Berry, 2010; Dillard, 2000), Chicana feminist epistemologies (Anzaldúa, 1987, 2009; Delgado Bernal, 2002), and other criticalities that question and dismantle colonialist and normative perspectives in education and in the social sciences writ large (Delgado and Stefancic, 2017; Patel, 2015; Tuck and Yang, 2012). While space does not permit a thorough analysis of the values of every theory used in qualitative applications, students and novice researchers should keep in mind that being able to articulate one’s values as a researcher and how these inform the choice of a theoretical framework is prerequisite to selecting a methodological sub-approach and research design.

Axiological foundations of specific qualitative methodologies

Increasingly, qualitative research is seen as a venue for amplifying and uplifting individuals, communities, and perspectives that have historically been devalued, dismissed, distorted, or silenced in a given population or culture. Owing to the emic imperative in qualitative research, and embracing the “centrality of experiential knowledge” that derives from many critical race theoretical applications (Pérez Huber, 2009; Solórzano and Yosso, 2016), qualitative inquiry is uniquely poised to elevate scholarship that counters epistemological racism and shortcomings of literature that focus only on the Global North and its limited viewpoints (Kubota, 2020; Scheurich and Young, 1997). From literature reviews and citation practices, to data collection methods, to suggestions for future studies, qualitative research increasingly embodies values of pluralism and antiracism. Arguably, credit for the development of these values in the history of social research goes to the methodology known as Ethnography, explored next.

Ethnography

Ethnography is the study of a culture-sharing group or some aspect of a culture-sharing group and, like some of the other qualitative approaches outlined below, it represents both a process and a product of the research (*an* ethnography). There are currently many forms of Ethnography practiced by qualitative researchers including critical, feminist, autoethnography, life history, ethnographic case study, and ethnographic grounded theory (see *Grounded Theory*, this volume; see also Creswell and Guetterman, 2019). Educational anthropology focusing on the use of ethnographic methods to study education topics emerged in the 1950s, led by George and Louise Spindler, Jules Henry and others, and has grown in popularity over time, having been applied in many and varied educational settings (Creswell and Guetterman, 2019; Erickson, 2018) and foregrounding participant viewpoints.

Field researchers often distinguish between emic perspectives, representing those of the participants, and etic views espoused by the researchers themselves (see Lévi-Strauss (1973) for a discursive history of these two terms). Most ethnographic research allows for a blend of these vantage points. Looking at the history of ethnographic fieldwork, one witnesses a sea change from more etically-driven approaches to emically-based accounts as fieldworkers have underscored the importance of more equitable partnerships and realization of potential benefits for participants. Accordingly, many contemporary ethnographers value transferring the authority from the researcher to the participants—and to the communities they represent—in order to distance themselves from the colonial stigma that has haunted previous ethnographers.

Phenomenology

German philosopher Wilhelm Dilthey laid the foundation for interpretivist paradigms by introducing the notion of *Verstehen* (‘understanding’) to recast reality as being composed, at minimum, of “natural reality and social reality” (Crotty, 1998, p. 67; see also Neuman, 2011). The German philosophical tradition incubated this idea throughout the early 1900s, bringing forth what would later become phenomenological methodology. In the 1960s and 70s, Phenomenology gained popularity, shepherded

by Edmund Husserl—who favored a descriptive or *transcendental* approach—and Martin Heidegger who advanced an interpretive or *hermeneutic* approach. For both of them, the essence of a person's experience was valued as the object of study, shifting research foci from the external and seen toward the unseen and ineffable elements of human existence. Similarly, Mead and Blumer (see Blumer, 1980) promoted Symbolic Interactionism that rested on the idea that people construct reality through the use of symbols, meaning-making, action, and interaction. While interpretivist thinking was inspiring new phenomenological methods and methodologists, a divergent sub-approach that privileges data over interpretation was arriving on the qualitative scene.

Grounded theory

As new targets of focus flourished in qualitative research, so did ways of using the research process to grow or refine current theories across the disciplines. The development of Grounded Theory methodology is one such advance. Grounded Theory values systematic and iterative methods that guide researchers to better understand a social or social-psychological process via a theoretical sample that then undergoes *constant comparative* analysis (Glaser, 1965) until thematic *saturation* is achieved (Hood, 2007, p. 154). Evident in this methodological sub-approach, and at times contested, is the value of collecting a substantial amount of data—whether that be in the number of interviews conducted (Guest et al., 2006), the sampling strategy (Draucker et al., 2007), or the richness or “information power” of the data set (Malterud et al., 2016). Readers can imagine without difficulty that some qualitatively-inclined researchers may take issue with Grounded Theory's emphasis on *quantity* as a valuable dimension of data composition.

Data selection procedures within a Grounded Theory approach signified a departure from the way sites or data sets were chosen for the core ethnographies. Data for the former are chosen based on their potential for grounding heretofore unknown or underdeveloped theory, whereas the latter (ethnographic) data were prized for giving holistic insight into particular groups of people and communities. Hood (2007) notes that, “This focus on emerging theory and theory driven data collection is very different from data collection in conventional ethnography which strives toward thick description of all that can be observed in a given setting regardless of theoretical relevance” (p. 155). Whereas ethnographic fieldwork is characterized by prolonged engagement with a community, some grounded theorists hold as their primary methodological value a prolonged engagement with the data.

We see here that the thick description once honed so carefully by the Chicago School of Sociology, in an effort to position qualitative research as equally ‘rigorous’ to the quantitative studies dominating other scientific fields at the time, came to be of lesser importance during the late 1960s and early 1970s when Grounded Theory was enjoying its halcyon days. That the hallmarks of early qualitative research could move aside in favor of Grounded Theory's systematic approach arguably demonstrates that qualitative research was undergoing a renaissance of sorts—the thick description of the 1920s and 30s' Chicago ethnographers was no longer necessary to evince the rigor of the qualitative approach.

Additional qualitative sub-approaches as values-driven

All sub-approaches within qualitative research showcase values to some extent in their applications. Narrative Inquiry focuses on the participant's personal knowledge as a source of rich data; however, the outcome of the study is seen as a co-creation of both researcher and participant. Some of the philosophical assumptions that define the contours of Narrative Inquiry as a qualitative methodology include temporal experience as fluid, and personal reflection as pivotal, in people's own evolving comprehension of their circumstances. Narrative Inquiry therefore values an alternative, non-realist *teleology*, or the nature of the concept of time.

Along these teleologic lines, a lesser-known qualitative methodology, *Currere*, “foregrounds the relationship between narrative (life history) and practice, and provides opportunities to theorize particular moments in one's educational history, to dialogue with these moments, and examine possibilities for change” (Kanun, 2006, p. 104). From the Latin for ‘to run the course’, this method underscores biography, autobiography, and temporality as lenses through which to view personal experience, and it embodies “the self-conscious conceptualization of the temporal” in a four-step method of reflexive methodology that progresses from *regressive*, to *progressive*, to *analytical*, and finally to *synthetical* inquiry (Pinar, 1994, p. 19). Telicity in *Currere* and Narrative Inquiry is more relativist in nature, whereas Action Research—described next—takes a more realist approach to the concept of time.

Action Research (see *Action and Community-Based Participatory Approaches in Qualitative Research*, this volume) encompasses a wide range of approaches spanning diverse fields such as education, social work, organizational psychology, and other social and health sciences. Broadly defined, it is a systematic, self-reflective, and cyclical process that strives to collaboratively engage researchers and participants toward the goal of generating practical solutions to problems of mutual concern, and spurring positive, long-term change (Babchuk et al., 2021; Brydon-Miller et al., 2011; Creswell and Guetterman, 2019; Grimwood, 2015; Ivankova, 2014). Sub-approaches of Action Research include Participatory Action Research, which values full democratic participation of stakeholders, often involving an explicit focus on political engagement, collaboration, and social justice; and Community-Based Participatory Research, often used in health sciences and aspiring for full community participation in all parts of the research. These forms of Action Research promote equitable partnerships among stakeholders (*i.e.*, researchers, organizational representatives, and members of the community) in all phases of the research process including selecting a research question and the design, implementation, evaluation, and dissemination of the research itself (Israel et al., 2008; Wallerstein and Duran, 2017).

Just as specific methodologies are the product of the special combinations of ontology, epistemology, and axiology that a research community supports, so too are the quotidian practices and procedures that make up a research project start to finish. In the following sections, the values manifest in choosing methods such as member checking, memoing, and more are discussed as they relate to philosophical assumptions about research, researchers, and participants.

Operational values in specific methods and procedures

The specific **methods** employed in qualitative research ought to be deliberately chosen for their appropriateness to the overall design, and for their ability to answer the research question(s) within the timeline and other constraints established. Whatever combination of sampling, data collection, analysis, and write-up methods qualitative researchers choose, project leaders should be able to rationalize why and how each design decision was made. Values are operating beneath the surface of each qualitative application, whether acknowledged or not. The following sub-sections attend to specific procedures in the research process, noting what each one says about values in qualitative inquiry.

Sampling

Although qualitative findings are not intended to be generalized from a sample to a population, thoughtful and rationalized sampling in a qualitative study (see *Qualitative Sampling and Qualitative Data*, this volume) is an opportunity to add rigor and quality, and to enhance the possible application of a study's conclusions. In some sub-approaches, such as Grounded Theory, sampling is an integral part of the data collection and analysis strategy throughout the process. In focus group methodology (see *Focus Groups*, this volume), the make-up of the sample can make all the difference in whether rich data are collected or not. Many qualitative studies rely on purposeful sampling of individuals or communities who possess intimate knowledge of the subject matter. As such, sampling in qualitative research is philosophically aligned with the belief that people's expertise with ordinary facets of their lives is the reason for their selection as participants. The ways of defining the characteristics of a purposeful sample—as well as the number of people in that sample—are contested matters that vary across and within particular research designs.

Interviews

Whereas many qualitative researchers will instinctively lean toward interviews as the primary means of data collection, this choice should not be made unthinkingly, and it is not without alternatives. Interviewing is a common, popular, and powerful means of gathering data from human participants. Infused with the emic lens resident in all qualitative approaches, interviews showcase individuals' perspectives in their own words. However, Kvale (1994) and others have questioned interviewing as a go-to method, arguing that a theoretical or scientific basis for conducting interviews is generally lacking in most research paradigms. Interviews as the principal form of data collection have also been rightly accused of an ableist bias (Kerschbaum and Price, 2017), depending as they do on eye contact, physical or remote proximity and, in most cases, synchronous, spoken language intelligible to both interviewer and interviewee.

Participant observation

Many qualitative studies, especially those using ethnographic methodology, rely in part on participant observation as one data collection method. This means that researchers immerse themselves in the activity, locale, or experience being studied and that they take part as authentically as possible in order to better understand the community they are profiling. Proximity, immersion, detail, reflection, and personal humility while embedding oneself in the study area are highly valued in this method. Although researchers must be judicious in drawing conclusions based on their direct experience of the phenomenon or community under study, the real-time simulation of a cultural phenomenon allows the researcher to gain insights that may not be otherwise available. To balance their personally drawn observations with the views of participants, and to show transparency in the research process, many qualitative researchers include a member checking procedure, as described below.

Member checking

One of the ways qualitative researchers enact values of rigor and transparency is to commit to **member checking** or similar forms of respondent validation, especially when findings from participant interviews are the sole means of data collection. There are numerous and divergent viewpoints on member checking, with some methodologists describing the issue as a “chaos of considerations” (Boswell and Perez, 2021). These include methodological, epistemological, and population- or topic-specific concerns. Member checking also reveals ethical issues such as determining whose interests are served by conducting member checking, and how one might proceed when the interests of the participant, the researcher, and the academic institution sponsoring the research appear to conflict. Some have also questioned whether member checking can cause harm to participants (see especially Hallett, 2013). Thus, one's philosophical assumptions about the nature of truth and the implications of the research situation itself greatly shape research dynamics throughout the process.

Memo-writing and bracketing

There are additional ways that a researcher can apply the values of transparency and humility by acknowledging the limits of their own perspective, although the researcher's private reflections are also highly valued in many methods and methodologies. **Bracketing** or *epoché*, from the Greek for ‘suspension of judgment’, is an ancient Stoic practice requiring the researcher to become aware of—and hold in abeyance—their personal views and experiences related to the subject of inquiry. Many qualitative researchers similarly rely on margin notes and **memos** to document their thoughts as they go, and to brainstorm possible limitations in their approach to the research. Although writing research memos is at times a perfunctory task, this method is grounded in a number of axiological dimensions of naturalistic inquiry. Among them are the need for the researcher to document insights, musings, and possible biases. This reflective and reflexive process manifests the qualitative belief in the researcher as the main instrument

of inquiry, as well as the qualitative imperative for one to consider how one's own social location influences the project. Memos add methodological rigor to the research process and they form part of the audit trail used to assess the researcher and the findings.

Importing qualitative sensibilities into mixed methods research

The philosophical underpinnings of qualitative research have bearing on how qualitative strands are integrated within mixed methods and multimethod designs. Mixed methods practitioners must value numbers, stories, and teamwork in order to transcend the limitations of a single methodology. As argued by Johnson and Walsh (2019), Mixed Methods Research (MMR) rests on the "collection and integration of qualitative data, logics, and philosophies" (p. 521). Because MMR is a "team sport" (Curry et al., 2012) philosophical tension, if not outright conflict, is to be expected. However, the tensions that inhere in mixing qualitative and quantitative traditions can be productive venues for researchers to create a whole that is greater than the sum of its parts. This aligns with the basic methodological rationale of MMR, wherein both the design elements and their execution are undertaken in such a way as to capitalize on the assets of both methodologies while "neutralizing the weaknesses" of each approach (Creswell and Plano Clark, 2018, p. 546).

As a developing field, MMR serves as a role model for articulating whether and why signature qualitative elements should be addressed, including memo-writing, member checking, positionality, and reflexivity, but these qualitative components are not included in MMR manuscripts as regularly as they are in qualitative studies (Boswell, 2021a). Because MMR is often practiced in multidisciplinary teams, "methodological camps" (Curry et al., 2012, p. 10) can inhibit and ultimately disrupt inquiry. Conflicting philosophical assumptions could surface even under the best of circumstances. Due to the breadth of disciplines and philosophies practicing mixed methods research, these designs invite questions about researchers' biases and other backgrounding assumptions, in much the same way that qualitative designs invite reflexivity and disclosures about researchers' own backgrounds and limitations. Historically, some have argued that mixed designs owe their legacy to experimental approaches and "take qualitative methods out of their natural home" (Denzin and Lincoln, 2018, p. 277). For the sake of robust collaboration however, and to ensure the care and feeding of the *dialectical pluralism* at the heart of mixed paradigms (Greene, 2015; Johnson, 2017), mixed methodologists should encourage their diverse research traditions to embrace, and augment, their combined strengths. A qualitative ethos built into a mixed design not only supports a particular project, but can help non-qualitative researchers on an MMR team understand the humanist values at the heart of naturalistic inquiry.

Summary of Part III

A researcher's philosophical assumptions, theories, and values are on display in any qualitative application. The choice of a particular theoretical framework, qualitative design, and the step-by-step methods and procedures selected enact the ontological, epistemological, and axiological stances of the person or people responsible for conducting a qualitative or mixed methods project.

Coda: future directions in qualitative research

Where is qualitative research headed in the pandemic-era, politically polarized 21st century? Current trends presage innovative qualitative and multi-method, collaborative studies suffused with decolonial, postmodern, womanist and post-qualitative works that upend outdated methodologies and supplant them with more participatory methods, social justice themes, and emancipatory outcomes. Approaches such as Black feminist thought and Chicana/o epistemologies overturn longstanding and unexamined Eurocentric theories. In the next generation of qualitative inquiry, methodologists must navigate the antiracist turn in social research, as well as learning to wed multiple ontological and epistemological stances harnessed in service of currently unimagined prosocial aims.

At the same time, there is pushback from segments of academia and public policy that consider whole fields—such as queer studies (see *Queer Theory and Qualitative Research*, this volume), discourse analysis, or critical race theory—to be leftist 'fringe' pursuits plagued by shoddy methods and intellectual navel-gazing. The Sokal Squared Hoax of 2018 (Emba, 2018; Kafka, 2018; Melchior, 2018) shone a light on insular academic communities, which some feel have a penchant for so-called *grievance studies* that shallowly (in the view of some 'anti-woke' critics such as the founders of the new University of Austin in the state of Texas, USA) make all scholarship about dismantling patriarchies. Geertz (2012) referred to these as the 'johnny-come-lately' disciplines—areas such as gender studies, and ethnic studies that comprise the most important subjects on college campuses, as well as the most divisive. Unfortunately, this has inserted an imaginary wedge between the pursuits of rigorous research and justice-oriented social movements.

Currently, in the United States, entrenched and vehement political divisions are accelerating the closure, discrediting, or censoring (Svrluga and Rozsa, 2022) of many collegiate programs and departments dedicated to ethnic studies, gender studies and other newly controversial topics. (Dreger, 2015) telegraphed that this would result in some academicians seeking "regression to the safe" in terms of declaring research priorities (p. 258), but she urges the academy to remember that, "Science and social justice require each other to be healthy, and both are critically important to human freedom" (p. 11). Rather than putting justice and science at odds with one another, qualitative research from this point forward is well positioned to embody their confluence.

References

- Anzaldúa, G., 2009. *The Gloria Anzaldúa Reader*. Duke University Press.
- Anzaldúa, G., 1987. *Borderlands: la frontera*. Aunt Lute, San Francisco.
- Babchuk, W.A., 2019. Fundamentals of qualitative analysis in family medicine. *J. Fam. Med. Community Health* 7 (2).
- Babchuk, W.A., Brand, L., Guetterman, T.C., 2021. Improving research-based practice through qualitative inquiry: a community-based study of minority health care. *J. Ethnogr. Qual. Res.* 15 (2), 99–116.
- Belfrage, C., Hauf, F., 2017. The gentle art of retrodution: critical realism, cultural political economy and critical grounded theory. *Organ. Stud.* 38, 251–271.
- Berry, T., 2010. Critical race feminism. *Encycl. Curric. Stud.* 1, 151–152.
- Blumer, H., 1969. *Symbolic interactionism*, 9. Prentice-Hall, Englewood Cliffs, NJ.
- Blumer, H., 1980. Mead and Blumer: the convergent methodological perspectives of social behaviorism and symbolic interactionism. *Am. Sociol. Rev.* 409–419.
- Bogdan & Biklen, 2007. *Qualitative Research for Education: An Introduction to Theories and Methods*, fifth ed. Pearson A & B, Boston.
- Boswell, E., 2021a. The Qualitative Imperative: Searching for Positionality and Reflexivity in Mixed Methods-Grounded Theory Studies [Virtual Conference Presentation]. American Educational Research Association.
- Boswell, E., 2021b. Again Awake: A White Researcher's Iterative Positioning for Entering Black Spaces. *The Nebraska Educator* 6. <https://doi.org/10.32873/unl.dc.ne020>.
- Boswell, E., Perez, A., 2021. Member Check Interviews in Qualitative Research: Ethics, Procedures, and Practices [Virtual Conference Presentation]. International Congress of Qualitative Inquiry.
- Bryant, A., 2017. *Grounded Theory and Grounded Theorizing. Pragmatism in Research Practice*. Oxford University Press, Oxford.
- Brydon-Miller, M., Kral, M., Maguire, P., Noffke, S., Sabhlok, A., Denzin, N., Lincoln, Y., 2011. Jazz and the Banyan Tree. *Handbook of Qualitative Research*. SAGE, Thousand Oaks, CA, pp. 387–400.
- Cannella, G.S., Lincoln, Y.S., 2011. Ethics, research regulations, and critical social science. *The Sage Handbook of Qualitative Research*, fourth ed. SAGE, pp. 81–90.
- Chagnon, N.A., Irons, W., 1979. *Evolutionary Biology and Human Social Behavior: An Anthropological Perspective*. Duxbury Press, North Scituate, MA.
- Charmaz, K., 2014. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*, second ed. SAGE, London.
- Comte, A., 1858. *The Positive Philosophy of Auguste Comte*. Blanchard.
- Creswell, J.W., 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, fourth ed. SAGE Publications, Thousand Oaks, CA.
- Creswell, J., Guetterman, T., 2019. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, sixth ed. Pearson Education, New York.
- Creswell, J.W., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE, Thousand Oaks, CA.
- Creswell, J.W., Poth, C., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. SAGE.
- Crotty, M., 1998. *The Foundations of Social Research: Meaning and Perspective in the Research Process*. SAGE.
- Curry, L.A., O'Cathain, A., Plano Clark, V.L., Aroni, R., Fetter, M., Berg, D., 2012. The role of group dynamics in mixed methods health sciences research teams. *J. Mix. Methods Res.* 6, 5–20.
- Deegan, M.J., 2007. The Chicago school of ethnography. In: Atkinson, P.A., Coffee, A.J., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*. SAGE, Thousand Oaks, CA.
- Delgado Bernal, D., 2002. Critical race theory, Latino critical theory, and critical raced-gendered epistemologies: recognizing students of color as holders and creators of knowledge. *Qual. Inq.* 8, 105–126.
- Delgado, R., Stefancic, J., 2017. *Critical Race Theory: An Introduction*. NYU Press, New York, NY.
- Delpit, L., 1988. The silenced dialogue: power and pedagogy in educating other people's children. *Harv. Educ. Rev.* 58, 280–299.
- Denzin, N.K., Lincoln, Y.S., 2018. Chapter 1: the discipline and practice of qualitative research. In: *The SAGE Handbook of Qualitative Research*. SAGE, pp. 1–26.
- Dillard, C.B., 2000. The substance of things hoped for, the evidence of things not seen: examining an endarkened feminist epistemology in educational research and leadership. *Int. J. Qual. Stud. Educ.* 13 (6), 661–681.
- Draucker, C.B., Martsof, D.S., Ross, R., Rusk, T.B., 2007. Theoretical sampling and category development in grounded theory. *Qual. Health Res.* 17, 1137–1148.
- Du Bois, W.B., 1898. The study of the Negro problems. *Ann. Am. Acad. Polit. Soc. Sci.* 11, 1–23.
- Du Bois, W.E.B., 1935. Does the Negro need separate schools? *J. Negro Educ.* 4 (3). <https://doi.org/10.2307/2291871>.
- Du Bois, W.E.B., 1897. *Strivings of the Negro people*. Atlantic 80.
- Durkheim, É., 1895. *Rules of Sociological Method* [first published as *Les Règles de la méthode sociologique*]. Simon and Schuster.
- Dreger, A., 2015. *Galileo's Middle Finger: Heretics, Activists, and One Scholar's Search for Justice*. Penguin Books.
- Emba, C., 2018. The 'Sokal Squared' Hoax Sums Up American Politics. *Washington Post*.
- Erickson, F., 2018. A history of qualitative inquiry in social and educational research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 36–65.
- Geertz, C., 1973. *The Interpretation of Cultures*, vol. 5019. Basic Books.
- Geertz, C., 2012. *Available Light*. Princeton University Press.
- Glaser, B.G., 1965. The constant comparative method of qualitative analysis. *Soc. Probl.* 12, 436.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods research merger. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, pp. 606–615.
- Grimwood, B.S., 2015. Participatory action research: democratizing knowledge for social justice. In: *Fostering Social Justice through Qualitative Inquiry*. Routledge, pp. 196–217.
- Grbich, C., 2013. *Qualitative Data Analysis: An Introduction*, second. SAGE, London.
- Guest, G., Bunce, A., Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18, 59–82.
- Hadley, G., 2017. *Grounded Theory in Applied Linguistics Research: A Practical Guide*. Routledge, New York, NY.
- Hallett, R.E., 2013. Dangers of member checking. In: *The Role of Participants in Education Research*. Routledge, pp. 40–50.
- Hatch, J.A., 2002. *Doing Qualitative Research in Education Settings*. SUNY Press.
- Henrich, J., 2020. *The WEIRD People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Penguin, UK.
- Henrich, J., Heine, S.J., Norenzayan, A., 2010a. The weirdest people in the world? *Behav. Brain Sci.* 33 (2–3), 61–83.
- Henrich, J., Heine, S.J., Norenzayan, A., 2010b. Most people are not WEIRD. *Nature* 466 (7302), 29.
- Hood, J.C., 2007. Orthodox vs. power: the defining traits of grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. SAGE, Thousand Oaks, CA, pp. 151–164.
- Israel, B.A., Schulz, A.J., Parker, E.A., Becker, A.B., Allen III, A.J., Guzmán, J.R., 2008. Critical issues in developing and following CBPR principles. *Community Based Participatory Research for Health: From Outcomes to Process*. Jossey-Bass, San Francisco, CA.
- Ivankova, N.V., 2014. *Mixed Methods Applications in Action Research*. SAGE.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11, 156–173.
- Johnson, R.B., Walsh, I., 2019. Mixed grounded theory: merging grounded theory with mixed methods and multimethod research. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Current Developments in Grounded Theory*. SAGE, Thousand Oaks, CA, pp. 517–531.
- Kafka, A.C., 2018. 'Sokal Squared': Is Huge Publishing Hoax 'Hilarious and Delightful' or an Ugly Example of Dishonesty and Bad Faith? *The Chronicle of Higher Education*.

- Kanu, Y., 2006. 'Currere' to the rescue? Teachers as 'amateur intellectuals' in a knowledge society. *J. Can. Assoc. Curric. Stud.* 4 (2).
- Kerschbaum, S.L., Price, M., 2017. Centering disability in qualitative interviewing. *Res. Teach. Engl.* 52 (1), 98–107.
- Kubota, R., 2020. Confronting epistemological racism, decolonizing scholarly knowledge: race and gender in applied linguistics. *Appl. Linguist.* 41, 712–732.
- Kvale, S., 1994. Ten standard objections to qualitative research interviews. *J. Phenomenol. Psychol.* 25, 147–173.
- Lévi-Strauss, C., 1973. Structuralism and ecology. *Soc. Sci. Inf.* 12 (1), 7–23.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. SAGE.
- Lincoln, Y.S., Lynham, S.A., Guba, E.G., 2018. Paradigmatic controversies, contradictions, and emerging confluences revisited. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 108–150.
- Lipscomb, M., 2012. Abductive reasoning and qualitative research. *Nurs. Philos.* 13, 244–256.
- Malterud, K., Siersma, V.D., Guassora, A.D., 2016. Sample size in qualitative interview studies: guided by information power. *Qual. Health Res.* 26, 1753–1760.
- Matteson, S.M., Lincoln, Y.S., 2009. Using multiple interviewers in qualitative research studies: the influence of ethic of care behaviors in research interview settings. *Qual. Inq.* 15, 659–674.
- Mead, G.H., Schubert, C., 1934. *Mind, Self and Society*, 111. University of Chicago Press, Chicago.
- Melchior, J.K., 2018. Fake news comes to academia. *Wall Street Journal* A15.
- Merriam, S., Tisdell, E.J., 2016. *Qualitative Research: A Guide to Design and Implementation*, fourth ed. John Wiley and Sons, San Francisco.
- Milner, H.R., 2007. Race, culture, and researcher positionality: working through dangers seen, unseen, and unforeseen. *Educ. Res.* 36, 388–400. <https://doi.org/10.3102/0013189X07309471>.
- Neuman, W.L., 2011. Chapter 4: the meanings of methodology in social research methods. In: *Qualitative and Quantitative Approaches*. Allyn & Bacon, Boston, pp. 90–122.
- Patel, L., 2015. *Decolonizing Educational Research: From Ownership to Answerability*. Routledge.
- Pérez Huber, L., 2009. Challenging racist nativist framing: acknowledging the community cultural wealth of undocumented Chicana college students to reframe the immigration debate. *Harv. Educ. Rev.* 79, 704–730.
- Pinar, W.F., 1994. The method of "Currere". *Counterpoints* 2, 19–27.
- Reyes, V., 2018. Three models of transparency in ethnographic research: naming places, naming people, and sharing data. *Ethnography* 19, 204–226. <https://doi.org/10.1177/1466138117733754>.
- Richards, L., Morse, J.M., 2012. *Readme First for a User's Guide to Qualitative Methods*. SAGE.
- Scheurich, J.J., Young, M.D., 1997. Coloring epistemologies: are our research epistemologies racially biased? *Educ. Res.* 26, 4–16.
- Schnegg, M., 2014. The nature and validation of knowledge. In: *Handbook of Methods in Cultural Anthropology*, second ed. Rowman and Littlefield, Lanham, MD, pp. 21–53.
- Solórzano, D.G., Yosso, T.J., 2016. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1). <https://doi.org/10.1177/107780040200800103>.
- Strauss, A.L., Corbin, J., 1994. Grounded theory methodology: an overview. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. SAGE, Thousand Oaks, CA.
- Svrluga, S., Rozsa, L., 2022. In Florida, DeSantis's plans for colleges rattle some academics. *Higher Education | The Washington Post*. Retrieved July 16, 2022 from: <https://www.washingtonpost.com/education/2022/07/01/desantis-florida-universities-laws/>.
- Tuck, E., Yang, K.W., 2012. Decolonization is not a metaphor. *Decolo Indigen. Edu. Soc.* 1, 1–40.
- Vidich, A.J., Lyman, S.J., 2000. Qualitative methods: their history in sociology and anthropology. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 37–82.
- Visokolskis, S., 2021. Insightful yet inferential creativity: transduction as a derivation of abduction. In: *Proceedings of the Thirteenth Latin American Workshop on Logic/Languages, Algorithms and New Methods of Reasoning*. CEUR Workshop Proceedings, pp. 23–38.
- Wallerstein, N., Duran, B., 2017. *The Theoretical, Historical and Practice Roots of CBPR*. Community-Based Participatory Research for Health: Advancing Social and Health Equity. Jossey-Bass, San Francisco, pp. 17–29.

Further reading

- Bogdan, Biken, 2007. *Qualitative Research for Education: An Introduction to Theories and Methods*, fifth ed. Pearson, Boston.
- Charmaz, K., 2014. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*, second ed. Sage, London.
- Corbin, J., Strauss, A.L., 2015. *Basics of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. Sage, Thousand Oaks, CA.
- Crotty, M.J., 1998. *The Foundations of Social Research: Meaning and Perspective in the Research Process*. Sage, Thousand Oaks, CA.
- Glaser, B.G., Strauss, A.L., 1967. *The Discovery of Grounded Theory*. Aldine, Chicago.
- Guetterman, T.C., Babchuk, W.A., Howell Smith, M.C., (in preparation). *Intersecting Mixed Methods Research with Qualitative Designs: Principles and Practices*, Mixed Methods Research Series. Sage, Thousand Oaks, CA.
- Merriam, S.B., Tisdell, E.J., 2016. *Qualitative Research: A Guide to Design and Implementation*, fourth ed. John Wiley & Sons, San Francisco.
- Schnegg, M., 2014. The nature and validation of knowledge. In: *Handbook of Methods in Cultural Anthropology*, second ed. Rowman and Littlefield, Lanham, MD, pp. 21–53.

The history of qualitative research in education

Martyn Hammersley, WELS, The Open University, Milton Keynes, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

The early history of qualitative research in education	14
Differences in orientation among qualitative researchers	17
Conclusion	19
References	19

Over the past century, there has been a huge amount of qualitative research concerned with educational institutions and processes, carried out in many countries. Indeed, education has been a leading field in the development and application of qualitative approaches. In this chapter I will outline the history of this body of work, sketching the range of topics investigated, the diverse forms that qualitative inquiry has taken, and some of the disputes to which these have given rise.¹

Before starting on this, however, it is important to note that rather different histories could be written, depending upon how the words “qualitative” and “education” are interpreted; and, also, according to which countries are covered. I will adopt a relatively broad conception of what counts as qualitative research: as involving the collection of data that are not pre-structured in terms of researcher-defined categories, so that analytic interpretations must be developed during the course of inquiry on the basis of the data. These data may take the form of open-ended fieldnotes, and/or audio- or video-recordings, produced by the researcher via observation or interviews, as well as already available or elicited documentary material (text, audio, and/or visual, still and moving) (Hammersley, 2013a). I will also employ a relatively wide conception of what would count as *educational* research. This can vary in terms of its substantive focus, since education takes place not just in schools and in further and higher education institutions but in other settings too—such as homes, workplaces, and recreational facilities—as well as in networks of various kinds, online as well as offline. The focus may also vary, of course, according to whether it is on teachers, on learning processes, on the motivations and orientations of students, on the role of technologies in facilitating learning or working against it, and so on. As we shall see, what counts as *educational* research has also been a matter of dispute in methodological and theoretical terms, for instance according to whether it simply provides education-relevant knowledge or must be geared *directly* to bringing about educational improvement or change.

While adopting a relatively liberal approach in these terms, I will limit myself to English-language qualitative research, and primarily to that produced in the UK or the US. There is little good justification for these restrictions, my only excuses are the limits to my expertise and the space available.² Even within these limits, given the sheer amount of qualitative research that has been produced, coverage will necessarily be highly selective, designed primarily to give a sense of the range of types of work that emerged at different times; so, what is offered here is certainly not a fully comprehensive account. At the same time, it should provide a reasonably representative picture.

Today, qualitative research in education covers an extremely broad range of topics, takes a very wide variety of forms, and operates on the basis of a diverse range of fundamental assumptions about its purposes, about the nature of the phenomena being investigated, and about how they can best be understood. As we shall see, while the starting point was not some single coherent conception of what constituted qualitative educational inquiry, the proliferation of different approaches and the depth of disagreements has increased sharply over time. In large part, then, this account is a history of that process.

The early history of qualitative research in education

There are, of course, no absolute starting points in history. Qualitative research in education could be traced back at least to the child study movement in the 19th century (see, for example, Preyer and Brown, 1888), including Charles Darwin’s “A biography of an infant”, based on observations of one of his own children (see Hellal and Lorch, 2010). A little later, we could refer to Waller’s (1932) *The Sociology of Teaching*, which drew on his father’s experience as a school superintendent, his own experience teaching in a military academy, as well as case materials provided by teachers (see Everhart, 1989; Willower and Boyd, 1989). However, it was not until the 1950s that we find the beginnings of any sustained tradition of qualitative research in education, and so that is where I will begin.

¹For general accounts of the history of educational research, see de Landsheere (1988) and Lagemann (2000). For accounts of the development of ethnography in the field of education, see Woods (1984) and Hammersley (2018).

²Some sense of developments in other parts of the world can be gained from Milstein and Clemente (2018), Egger and Unterwieser (2018), and Modiba and Stewart (2018). These articles focus specifically on educational ethnography in, respectively, Latin America, Germany and Switzerland, and Africa.

This arose at a time when the dominant kinds of quantitative work in the field of education, and elsewhere, were coming under increasing criticism for failing to meet the rigorous standards they had set themselves and for frequently producing inconsistent or trivial results. Some argued that this stemmed from a fundamental mismatch between the, broadly positivist, philosophy of science that underpinned this work and the very nature of human social life (see, for instance, Bruyn, 1966; Lincoln and Guba, 1985). But also, around this time, there were growing doubts within the philosophy of science about whether positivism offered an accurate picture even of natural science, the most significant example being the revisionist account of the development of scientific knowledge provided by Thomas Kuhn (1962).

In the United States, an important initial development was ethnographic investigations by anthropologists of the relationship between schools and local North American indigenous communities (Spindler, 1955; Dumont and Wax, 1971; Wax et al., 1971; Erickson, 1977; McDermott, 2008).³ This approach was later extended to investigate schooling in other types of local community (Wylie, 1964; Warren, 1967), distinctive forms of religion-based schooling (Peshkin, 1978), the role of key actors within a school (Wolcott, 1984), and educational inequalities within urban communities (Ogbu, 1974). Around the same time, in the field of sociology, drawing on the tradition of the Chicago School, Howard S. Becker (1951, 1952a,b, 1953) carried out a study of school-teachers which significantly informed later work, such as that of Lortie (1975, 2009). Adopting a very different focus, Cicourel and Kitsuse (1963) applied labeling theory and ethnomethodology in studying the role that counsellors played in shaping students' careers in a US school (see also Erickson and Shultz, 1981a). Cicourel followed this up later with a study of the role of language in educational decision-making and testing processes (Cicourel et al., 1974). Outside of anthropology and sociology there were some other pioneering qualitative studies of schools in the US, such as those of Jackson (1968) and Smith and Geoffrey (1968). The first was one of a number of studies that set out to document the reality of life in (American) schools, as against prescriptive accounts. The second was a much more methodologically deliberate study carried out by an observer initially trained in psychology but alienated from some of the main trends within that discipline, collaborating with the practicing teacher whose classroom he was studying. A little later Rist (1973), Cusick (1973, 1983) and Everhart (1983) provided further participant observation studies in schools.

In the UK, the origins of qualitative research are usually traced back to the work of Lacey (1966, 1970), Hargreaves (1967), and Lambart (1976, 1982), which was influenced by both anthropology and sociology, and focused on the differentiation of pupils—in the form of streaming and setting (tracking)—and the consequences of this. And others followed up with similar studies: Ball (1981), Abraham (1995). Later, Hargreaves adopted a rather different approach, applying labeling theory and ethnomethodological ideas in investigating *Deviance in Classrooms* (Hargreaves et al., 1975; Hargreaves, 1976). This reflected a significant shift in orientation within the field in the UK in the early 1970s and there was a considerable expansion in the amount of qualitative research carried out at that time (see Delamont et al., 1993). This included what was labeled “the new sociology of education” (Young, 1971; Sharp and Green, 1976); the development of ethnographic work drawing on symbolic interactionism and social constructionism (Delamont, 1976; Hammersley, 1974, 1976; Stubbs and Delamont, 1976; Woods, 1979; Woods and Hammersley, 1977; Hargreaves and Woods, 1984; Hammersley and Woods, 1984); the work of the Birmingham Centre for Contemporary Cultural Studies, notably Willis's *Learning to Labour* (1977); and the emergence of qualitative approaches in the field of educational evaluation, such as Parlett and Hamilton's (1972, 1977) “illuminative evaluation”, and various forms of “case study” (Stake, 1995; Walker, 1978; MacDonald, 1974).

From the 1970s onwards, in both the US and the UK, there was also work by linguists, sociolinguists, and conversation analysts focused on describing the structure and process of classroom discourse (for instance, Cazden et al., 1972; Sinclair and Coulthard, 1975; Stubbs, 1975, 1976; Payne, 1976; Mehan, 1979; McHoul, 1978; Green and Wallat, 1981, Green 1983), as well as investigating the relationship between language use at home and at school (for instance Brice Heath, 1982, 1983; Philips, 1983). Indeed, around this time there was considerable interplay between linguistics, anthropology, sociolinguistics, and sociology within the field of education. And this is an area where work has continued to develop, for example resulting more recently in the emergence of linguistic ethnography (Rampton, 2006; Maybin, 2006; Copland and Creese, 2015).

While much qualitative research in the 1970s and 80s focused on documenting what was happening in schools and classrooms, there were other sorts of work too. For example, there was some historical investigation of how various school subjects had become established in the curriculum, and how they had changed over time (Goodson and Ball, 1984; Cooper, 1985). Another new development was life history investigations of teachers' careers, showing the effects of external events and personal change over time, and documenting variation in career patterns, notably those related to gender (see Sikes et al., 1985).

From the 1960s onward, there were also some pioneering studies in the field of higher education, notably Becker et al.'s *Boys in White* (1961), a study of a medical school, and *Making the Grade* (1968), an investigation of the experiences and perspectives of undergraduate students in a US university. Meanwhile, Olesen and Whittaker (1968) carried out a similar study of the education of nurses. These investigations stimulated later work, such as Atkinson's (1997) investigation of medical students' experience of clinical training. There was also the pioneering work of Miller and Parlett (1974) on students' orientations toward assessment and examinations. And, in the 1970s, phenomenological approaches to student learning were developed in the field of higher education, under the heading of “phenomenography” (Marton, 1981; Gibbs et al., 1982). There were also a few studies of informal educational processes (see Becker, 1972; Geer, 1972, and other articles in the same issue of the *American Behavioral Scientist*; Lave and Wenger, 1991).

³Long before this some anthropologists had given attention to educational processes, a famous example being Mead's (1928) *Coming of Age in Samoa*. For an account of the history of anthropological work on education, see Mills (2012).

During the remainder of the 20th century and into the 21st, qualitative research expanded hugely, covering many aspects of education. Initially, at least, much of it focused on publicly-funded schooling, and a great deal of it on the secondary level, but even early on there were exceptions. For example, there was the work of Pollard and others on primary schools, some of this taking the form of longitudinal ethnography (see [Berlak and Berlak, 1981](#); [Pollard, 1985, 1996, 1999](#)); see also Andy Hargreaves' (1986) research on middle schools. There were studies of the teaching of particular subjects in primary school, notably science ([Driver, 1983](#); [Driver et al., 1985](#); [Traianou, 2006](#)). There were also some studies of elite private schools ([Delamont, 1973](#); [Howard, 2008](#); [Gaztambide-Fernandez, 2009](#); [Khan, 2011](#)) as well as of vocational training courses ([Bates, 1990](#); [Bates and Riseborough, 1993](#)). In addition, there were a few investigations of teaching and learning outside of school or college settings, for instance Alim's (2011) investigation of Hip Hop practice, [Atkinson et al.'s \(2013\)](#) research on operatic masterclasses, and [Delamont et al.'s \(2017\)](#) account of the teaching and learning of the Brazilian martial art Capoeira.

Perhaps not surprisingly, over time, there was increasing diversity in approach, even within particular disciplinary fields (see, for instance, [Eisenhart's, 2018](#) discussion of the anthropology of education). Key aspects of this were the emergence of critical ethnography ([Anderson, 1989](#); [Carspecken, 1996](#); [Beach and Vigo-Arrazola, 2020](#)), and the development of participatory forms of inquiry, notably in the field of Childhood Studies (see, for instance, [Cuevas-Parra and Tisdall, 2019](#)). Both of these were influenced by the development of feminist ideas about social research methodology (see [Fonow and Cook, 1991](#)), and other theoretical and methodological developments.

A central theme in much work in the sociology of education, and in "critical ethnography", was the role of schooling in reproducing social inequalities. Initially, the focus was on working-class underachievement, and this continued to be a major topic (see, for instance, [Dubberley, 1988](#); [Lareau, 1989, 2003](#); [Weis, 1990, 2004](#); [Weis and Fine, 2000](#)). However, from the 1980s, feminist researchers began to use qualitative methods to investigate the differential treatment of girls and boys in schools, and its effects (see, for example, [Deem, 1980](#); [Delamont, 1980](#)).⁴ This was often conceptualized as curricular and pedagogical discrimination against girls. For example, [Clarricoates \(1980\)](#) explored how primary teachers' gendered expectations shaped their treatment of children in the classroom. Coming from a rather different angle, and drawing on the techniques of sociolinguistics, [French and French \(1984\)](#) investigated the processes by which boys—or, rather, *some* boys—in primary schools gained greater attention from teachers than did girls. There was also exploration of the perspectives and strategies employed by female students to manage the conflicting expectations they faced and achieve their goals (see, for instance, [Measor, 1983](#)), and this led to a considerable amount of work looking at how different versions of femininity were constructed in school contexts and beyond ([Renold, 2001](#)) as well as research on girls' friendships ([Hey, 1997](#); [Gordon et al., 2001](#); [George, 2007](#)) and how they construct gender identities ([Francis, 1998](#)). Later there was also investigation of hegemonic and subordinated forms of masculinity ([Mac an Ghaill, 1994](#); [Renold, 1995](#); [Skelton, 1996, 2001](#); [Kenway et al., 2006](#); [Swain, 2006](#)).⁵ In addition, some attention started to be given to the experiences of Lesbian, Gay, Bisexual, and Transgender students in schools and colleges (see, for instance, [Youdell, 2009](#)).

In the UK, the move away from an almost exclusive focus on social class also led to investigations of the experiences of ethnic minority children in schools. As in the case of feminist research, this was often carried out within the framework of an explicit political orientation, this time an anti-racism which viewed official multiculturalism as an ideological cover for continued prejudice and discrimination, especially of an institutional kind. Much research set out to document discrimination against African Caribbean and South Asian students ([Wright, 1967](#); [Gillborn, 1990](#)). In the US similar studies developed under the heading of Critical Race Studies (see [Lynn and Parker, 2006](#)). There was also investigation of the perspectives and strategies employed by minority ethnic students. For example, [Fuller \(1980\)](#) studied the ways in which a group of African-Caribbean girls adapted to their situation in school, seeking academic success while minimizing their conformity to the demands of teachers. [Mac an Ghaill \(1988\)](#) examined the relations between students from different ethnic groups and teachers in a boys' secondary school and in a sixth form college. In the process, he documented the stereotypes on which teachers operated and the range of adaptations adopted by students, these influenced not just by ethnicity but also by gender. This early work stimulated a whole range of later studies exploring the experiences and perspectives of ethnic minority children in UK schools (see, for example, [Haw, 1998](#); [Bhatti, 1999](#)), as well as at how these influenced subsequent careers ([Saeed, 2021](#)).

During the final decades of the 20th century, and into this century, another major focus for qualitative research in education was the nature and effects of education policies. In a unique investigation [Ball \(1990\)](#) studied the policymaking process by interviewing politicians and officials who had been involved in formulating and implementing key legislation in the UK, and he and others carried out a series of studies of UK policies and their impact (for instance [Gewirtz et al., 1995](#); [Ball et al., 2000](#); [Maguire et al., 2001](#); [Perryman et al., 2011, 2017](#)). A considerable body of work investigated the adaptations of schools and teachers to new policy initiatives and their implications for the nature of teaching and learning ([Woods et al., 1997](#); [Jeffrey and Woods, 1998](#)). Much of this research was concerned with the constraints that various national-level reforms imposed on more progressive forms of education, and the extent to which they worsened existing inequalities. There was also investigation of the impact of school inspections ([Jeffrey and Woods, 1998](#)) and of the effects of stress on teachers generated by rolling waves of new policy initiatives ([Troman and Woods, 2001](#)).

⁴It is also worth mentioning that most previous ethnographic work had been located in secondary schools, whereas feminists gave as much if not more attention to primary schools.

⁵For useful collections illustrating the nature of the work that was done subsequently, see [Walker and Barton \(1983\)](#), [Woods and Hammersley \(1993\)](#), [Skelton and Francis \(2003\)](#).

While I have tried to cover the areas in which qualitative research was most concentrated, there were many other topics investigated, from research on the changing role of technology in schools (see, for examples, [Littleton, 2010](#)) to the character of educational testing (see, for instance, [Cooper and Dunne, 2000](#)). But equally important has been the proliferation of new theoretical and methodological approaches, and this will be my focus in the next section.

Differences in orientation among qualitative researchers

There have been various attempts to conceptualize the increasing diversity in approach among qualitative researchers in education. These operate at different levels, or address it from different angles, and there is also variation according to which country is focused on. At a relatively specific level, [Jacob \(1987\)](#) and [Delamont et al. \(1993\)](#) sketched the various approaches to be found in the United States and the UK, respectively, in the period up to the late 1980s. Subsequently, many other kinds of work differentiated at this relatively specific level have been added, such as Activity Theory ([Roth et al., 2012](#)), Autoethnography ([Goode, 2019](#)), the use of Focus Group Methodology ([Williams and Katz, 2001](#)), visual methods ([Prosser, 2007](#)), and Qualitative Comparative Analysis ([Cooper and Glaesser, 2012](#)).

At a deeper, more philosophical, level, distinctions were often drawn between broad “paradigms”, such as “positivist/post-positivist”, “critical”, and “interpretive” (see [Eisner and Peshkin, 1990](#); [Guba, 1990](#); see also [Hammersley, 1992](#); [Piper and Stronach, 2004](#)). The term “paradigm” was taken from [Kuhn’s \(1962\)](#) work on the history of natural science, but interpreted in a broad way as referring to a set of fundamental assumptions—about the world and how to understand it, as well as about the proper purpose or intended product of research—that could underpin social scientific work. The post-positivist paradigm was portrayed as retaining a commitment to the model of natural science, though interpreting this in a different and more flexible way than positivism. The “critical” approach treats research as properly engaged in challenging the socio-political status quo, focusing on social inequalities, whether in terms of social class, gender, race/ethnicity, sexual orientation, or ability/disability. The interpretivist paradigm places emphasis on the task of understanding the perspectives of the various actors involved in the education process and how these shape their actions. While these broad differences in orientation continue to be relevant today, further differentiation within and beyond them has rendered the field even more complicated. This complexity is nicely illustrated by the fact that influential categories of qualitative inquiry—such as “ethnography”, “action research”, and “discourse analysis”—each contain multiple variants, often differing in fundamental respects.

Not surprisingly given this process of diversification, there have been recurrent disputes among qualitative researchers relating to methodological and theoretical matters. These began in the 1950s within the anthropology of education over the “culture and personality” perspective that was adopted initially in this field. While later work maintained the focus on culture, some of it adopted a more micro-focused approach, and rejected a psychological in favor of a more sociological orientation (see [McDermott, 1977](#); [Erickson and Shultz, 1981b](#); [Garcez, 2008](#)). These moves generated a significant debate with anthropological advocates of holism. They insisted that what happens in particular situations on particular occasions cannot be understood without being located in an understanding of the wider community, the whole society, and perhaps even global processes and relationships (see, for instance, [Lutz, 1981](#)). And, along with this, there were also some who emphasized the need for cross-cultural, comparative analysis. But many insisted that detailed analysis of processes of social interaction is required if these are to be understood, raising questions (implicitly or explicitly) about the validity of claims about macro-level contexts. Over the second half of the 20th century and into the 21st, there has been continuing tension and oscillation between these two emphases.

This issue, frequently conflated with differential weight given to “structure” and “agency”, also arose within the sociology of education in the 1970s and 80s. This reflected the growing influence of Marxism, in the wake of the new sociology of education. Some Marxists treated schools as locked into the reproduction of capitalist societies (see, for instance, [Anyon, 1980, 1981](#)), but many qualitative researchers in the UK, while not denying the pressures on schools to serve this function, emphasized the agency of teachers and students and the significance of processes of social interaction (see [Hammersley, 1980](#)). In addition, they complained that Marxist analyses failed to engage sufficiently with the problems involved in producing and analyzing data about school processes, tending to present just-so stories derived from Marxist theory rather than well-researched accounts ([Hargreaves, 1982](#)). A famous example that stimulated considerable controversy was [Willis’s \(1977\) *Learning to Labor*](#), a study of a small number of working-class boys with an oppositional attitude to school. While Willis stressed the agency of the boys, he portrayed them as exercising that agency in such a way as to condemn themselves to working class jobs, thereby reproducing capitalism (for critical assessments, see [Turner, 1979](#); [Walker, 1985](#); [McGrew, 2011](#)).

Somewhat similar debates took place in the United States, one example being a dispute between [Cusick \(1985\)](#) and [Everhart \(1985\)](#), ([Cusick and Everhart, 1985](#)). Both these authors had produced participant-observation studies of schools, and a journal assigned them to review one another’s books. Cusick complained that in Everhart’s account: “Although students are described and presented in a human and believable manner, administrators and teachers are described only through the eyes of the students and the Marxist perspective. In effect they are presented as mere unwitting tools of capitalist oppression” ([Cusick, 1985:71](#)). Meanwhile, Everhart criticized Cusick’s “overly-rigid adherence to a functionalist approach”, and a failure to “address at all the manner in which [the school] system works within the context of the larger sociopolitical order” ([Everhart, 1985:75](#)).

Another area of dispute was the relationship between research and educational practice. An influential strand of research coming out of the educational evaluation movement promoted various forms of action research ([Hendricks, 2019](#)). As part of this it was often argued that education research must be *educative* in character, in the sense of directly serving the improvement of educational

practice, rather than simply contributing to knowledge about educational processes, institutions, and outcomes. This sparked responses from some qualitative researchers committed to more “academic” forms of research, who questioned the quality of much that came under the heading of educational evaluation and action research (Atkinson and Delamont, 1985) and even whether action research could be a coherent enterprise (Hammersley, 2004). As already noted, though, action research fragmented into a range of competing approaches that were incompatible with one another in significant respects, from those adopting a “critical” orientation, according to which action research could challenge capitalist society, to those who viewed it entirely in terms of the educational development of individual teachers or the solution of practical classroom problems (see Kemmis, 1988).

In the 1980s there were debates sparked by feminism and anti-racism, with criticism directed at mainstream research for neglecting gender and race/ethnicity, and thereby overlooking significant forms of inequity. However, there was also some counter-criticism, with critics arguing that much research guided by feminism or anti-racism was speculative in character, substituting political assumptions for careful and systematic empirical investigation of what happens in schools, and of the complexities involved in the production of inequalities in educational outcomes (see Foster et al., 1996). For example, Foster (1990) challenged the adequacy of some of the work that claimed to document discrimination against ethnic minority students by teachers in UK schools, pointing to serious weaknesses in the evidence on which conclusions were based, and arguing that a bias was operating in the interpretation of this evidence.⁶ Foster’s criticisms of anti-racist research, supported by others (Gomm, 1993; Hammersley and Gomm, 1993; Foster et al., 1996), were met with a counter-charge of “methodological purism”: of using methodological arguments to discredit evidence of discrimination, as a result of unconscious racism (Gillborn and Drew, 1993; Troyna, 1993; Gillborn, 1995). The issues involved here were important ones, but little progress was made toward resolving the disagreement. Indeed, the debate raised questions about whether there was any scope for dialogue between “critical” and analytic approaches; some commentators appeared to remove any possibility of this (see, for instance, Scheurich and Young, 1997), reflecting the growing influence of Critical Race Theory (Taylor et al., 2016; Dixon et al., 2017).

From the 1980s onwards, research on education, as in other fields, was influenced by the range of ideas often given the label “postmodernism” (see Stronach and MacLure, 1997). These amounted to a selective appropriation and interpretation of influential French philosophical thinking from the 1960s and early 70s, particularly the work of Foucault and Derrida. One central theme was the role of various types of institution, including schools, in normalizing behavior by pedagogic or therapeutic means—this often done in the name of scientific expertise. Postmodernism highlighted the rhetoric and functioning of social science in this process, and reinforced a commitment on the part of many researchers to the championing of marginalized groups and their “subjugated forms of knowledge”. As part of this, claims to scientific expertise were attacked, not least through questioning the very possibility of the kind of knowledge science purports to offer, in which facts are presented on the basis of scientifically-controlled observation and analysis of one sort or another. In the firing line here was not just quantitative research but also many forms of qualitative inquiry. Any capacity of language to represent phenomena existing independently of it, or indeed to anchor and preserve meanings in the manner required for this, came to be denied by some educational researchers. The artifactual, if not fictional, character of all accounts of the world was emphasized, with literature and art increasingly treated as the appropriate model rather than science; though the rhetorical value of the latter continued to be recognized. The tension that had always been present within qualitative research between science and art was given a new twist, in fact several twists (see, for example, some of the articles in Reid et al., 2014).

An important effect of postmodernism was to increase the amount of attention given to the role of discourse (MacLure, 2003): official accounts, especially, but also those of informants in interviews, were no longer to be treated as accurate or inaccurate representations of the world but rather investigated for the work that they do in constituting and sustaining particular “truth regimes” within educational institutions and elsewhere—especially ones taken to be unjust. Thus, discourse analysis was often used to serve the same “critical” function as earlier forms of Marxist ideology critique, even though the previously assumed contrast between ideological and objective scientific accounts had been abandoned. Another effect of postmodernism was an exploration of alternative ways of pursuing and presenting the product of educational inquiry, for example the use of poetry, dance, and drama, much of this coming to be included under the heading of “arts-based research” (Bagley and Cancienne, 2002; Atkinson, 2012; Sikes, 2012; Leavy, 2018).

In the 1990s and into the 21st century a major conflict was prompted by the evidence-based practice movement and its promotion of Randomized Controlled Trials (RCTs) and systematic reviews as the “gold standard” source of evidence for educational policymaking and practice. Here, the older debate between quantitative and qualitative approaches was revived, albeit in a somewhat different form. There was growing public criticism of qualitative research in some influential quarters, on the grounds that it lacked scientific validity and that it did not serve educational policymaking and practice effectively (Hargreaves, 1996; see also Hammersley, 1997; Hillage et al., 1998; Tooley and Darby, 1998; National Research Council, 2002, Hammersley 2007). In the United States, a key development was the reauthorization of the Elementary and Secondary Education Act (“No Child Left Behind”) in 2001, which prioritized RCTs as a source of evidence in relation to educational practice. Increasingly, the use of RCTs was institutionalized by government in both the US and the UK (Hammersley, 2013b). Qualitative researchers reacted sharply against what they regarded as a devaluing of their work (see Lather, 2004a,b; Denzin and Giardina, 2006; Torrance, 2011). In the UK, the Economic and Social Research Council’s *Teaching and Learning Research Program* (TLRP), initiated in 2000 and ending in 2011, boosted investment in education research designed to make a practical contribution to educational practice; though quite a lot of this research was

⁶For full references to Foster’s work, and a discussion of the debate, see Hammersley (1995: ch. 4).

qualitative in character. Subsequently, sponsored by the UK Government, the Education Endowment Foundation has funded a considerable amount of research aimed at providing evidence about the effectiveness of various educational strategies and practices, much of this relying on RCTs.

In the wake of attempts to defend qualitative research against the resurgence of positivist ideas, along with a funding emphasis on quantitative method and the rise of the mixed methods movement, some qualitative researchers have raised the prospect of a move to the “post-qualitative” (Lather and St. Pierre, 2013; St. Pierre, 2019). They have pointed out how, to a large extent, the distinction between quantitative and qualitative approaches, and the opposition between positivism and other research philosophies, is founded on the assumption of a fundamental distinction between the physical phenomena that the natural sciences study and human social life. Some recent ideas, often going under the headings of “new materialisms” and “post-humanism”, question the previously assumed ontological basis for that distinction, insisting that positivist assumptions about the nature of physical phenomena do not take account of more recent developments in the natural sciences—for example quantum physics and epigenetics in biology. Involved here is an appeal to vitalist conceptions of matter, with the latter viewed as having agential capabilities that previous approaches to qualitative research had tended to restrict to human beings. There is also a link here to a valorization of indigenous cultural beliefs (Rosiek et al., 2020). In part this amounts to a reaction against the previous emphasis on discourse by qualitative researchers, characteristic of work influenced by post-structuralism and postmodernism, in favor of a focus on material processes. Thus, post-humanism rejects the idea that reality is constituted in and through the discursive or cognitive practices employed by human beings, and especially those of Western intellectuals. There is a focus on the role of technologies and material artifacts, and a recognition of both the ways in which nature (for instance the world’s climate) is being changed by human beings and how human social life is interpenetrated by materialities, with technologies no longer being mere tools but actually doing work that humans previously did, such as EdTech teaching programs, and increasingly reshaping human bodies in various ways from hearing implants to wearable computing and embedded microchips to control behavior. The post-qualitative movement is by no means single-stranded, and at present it is difficult to know what consequences it may have for qualitative research in education. These developments bring us up to the present-day; and, if we were to take the phrase “post-qualitative” literally, to the end of the history of qualitative research.

Conclusion

In this chapter I have outlined some of the history of qualitative inquiry in the field of education, and sought to provide a sense of the range of different areas of work and the proliferation of methodological and theoretical approaches that have developed there. As part of this, I looked at several of the disputes among qualitative researchers that have taken place over the years, these reflecting quite fundamental differences in assumptions about the nature of social and educational phenomena, how these can best be understood, and what the function of research ought to be. It is important to be aware of how current ideas and practices developed, since this will enable us to understand them better. It may also tell us to what extent we are merely treading water, or going round in circles, rather than making progress. Even aside from this, when doing research it is necessary to have a clear sense of the rationale for the kind of work one is doing in the context of the whole range of alternative approaches, their relative strengths and weaknesses, affordances and problems.

A key issue that arises from this history is whether the deep pluralism that now prevails was an inevitable product of the expansion of educational research, and of social inquiry more generally, or whether rather more agreement about fundamental issues could have been, and perhaps still can be, achieved. But there have also been disagreements about whether such pluralism is desirable or undesirable (Hodkinson, 2004; Hammersley, 2005; Lather, 2006). While the tension with quantitative method remains for many qualitative researchers, there is now little agreement among them about theoretical and methodological issues; and, as we saw, some erstwhile qualitative researchers wish to abandon or transcend it to enter a new post-qualitative age. The future is, as always, uncertain, but in facing it we would do well to learn what we can from the past.

References

- Abraham, J., 1995. *Divide and School: Gender and Class Dynamics in Comprehensive Education*. Falmer, London.
- Alim, H.S., 2011. Hip Hop and the politics of ill-literacy. In: Levinson, B., Pollock, M. (Eds.), *A Companion to the Anthropology of Education*. Wiley-Blackwell, Malden, MA.
- Anderson, G.L., 1989. Critical ethnography in education: origins, current status, and new directions. *Rev. Educ. Res.* 59 (3), 249–270.
- Anyon, J., 1980. Social class and the hidden curriculum of work. *J. Educ.* 162 (1), 67–92.
- Anyon, J., 1981. Social class and school knowledge. *Curric. Inq.* 11 (1), 3–42.
- Atkinson, P., Delamont, S., 1985. Bread and dreams or bread and circuses? A critique of “case study” research in education. In: Shipman, M. (Ed.), *Educational Research: Principles, Policies and Practices*. Falmer Press, Lewes.
- Atkinson, P., Watermeyer, R., Delamont, S., 2013. Expertise, authority and embodied pedagogy: operative masterclasses. *Br. J. Sociol. Educ.* 34 (4), 487–503.
- Atkinson, P., 1997. *The Clinical Experience: The Construction and Reconstruction of Medical Reality*, second ed. Ashgate, Farnham, UK.
- Atkinson, P., 2012. The literary and rhetorical turn. In: Delamont, S. (Ed.), *Handbook of Qualitative Research in Education*. Edward Elgar, Cheltenham.
- Bagley, C., Cancienne, M.B. (Eds.), 2002. *Dancing the Data*. New York, Peter Lang.
- Ball, S.J., Maguire, M.M., Macrae, S., 2000. *Choices, Pathways and Transitions Post-16: New Youth, New Economies in the Global City*. RoutledgeFalmer, London.
- Ball, S.J., 1981. *Beachside Comprehensive: A Case-Study of Secondary Schooling*. Cambridge University Press, Cambridge.
- Ball, S.J., 1990. *Politics and Policy Making in Education*. Routledge, London.

- Bates, I., Riseborough, G. (Eds.), 1993. *Youth and Inequality*. Open University Press, Buckingham.
- Bates, I., 1990. "No roughs and no really brainy ones": the interaction between family background, gender, and vocational training on a BTEC fashion design course. *J. Educ. Work* 4 (1), 79–90.
- Beach, D., Vigo-Arrazola, M., 2021. Critical ethnographies of education and for social and educational transformation: a meta-ethnography. *Qual. Inq.* 27 (6), 677–688.
- Becker, H., Geer, B., Hughes, E.C., Strauss, A.L., 1961. *Boys in White. Student Culture in Medical School*. University of Chicago Press, Chicago.
- Becker, H., Geer, B., Hughes, E.C., 1968. *Making the Grade: The Academic Side of College Life*. John Wiley and Sons, New York.
- Becker, H.S., 1951. Role and Career Problems of the Chicago Public School Teacher. Unpublished PhD Thesis. University of Chicago.
- Becker, H.S., 1952a. The career of the Chicago public school teacher. *Am. J. Sociol.* 57, 470–p477.
- Becker, H.S., 1952b. Social class variations in the teacher-pupil relationship. *J. Educ. Sociol.* 25, 451–p465.
- Becker, H.S., 1953. The teacher in the authority system of the public school. *J. Educ. Sociol.* 27, 128–p141.
- Becker, H.S., 1972. A school is a lousy place to learn anything in. *Am. Behav. Sci.* 16 (1), 85–105.
- Berlak, A., Berlak, H., 1981. *Dilemmas of Schooling*. Methuen, New York.
- Bhatti, G., 1999. *Asian Children at Home and at School*. Routledge, London.
- Brice Heath, S., 1982. Questioning at home and at school: a comparative study. In: Spindler, G. (Ed.), *Doing the Ethnography of Schooling: Educational Anthropology in Action*. Holt, Rinehart and, New York, pp. 102–131.
- Brice Heath, S., 1983. *Ways With Words: Language, Life and Work in Communities and Classrooms*. Cambridge University Press, Cambridge.
- Bruyn, S., 1966. *The Human Perspective in Sociology*. Prentice-Hall, Englewood Cliffs, NJ.
- Carspecken, P., 1996. *Critical Ethnography in Educational Research: Atheoretical and Practical Guide*. Routledge, New York.
- Cazden, C.B., John, V.P., Hymes, D. (Eds.), 1972. *Functions of Language in the Classroom*. Teachers College Press, New York.
- Cicourel, A.V., Kitsuse, J.I., 1963. *The Educational Decision-Makers*. Bobbs-Merrill, New York.
- Cicourel, A.V., Jennings, K.H., Jennings, S.H.M., Leiter, K.C.W., MacKay, R., Mehan, H., Roth, D., 1974. *Language Use and School Performance*. Academic Press, New York.
- Claricoates, K., 1980. In: Deem, R. (Ed.), *The Importance of Being Ernest ... Emma ... Tom ... Jane. The Perception and Categorization of Gender Conformity and Gender Deviation in Primary Schools*.
- Cooper, B., Dunne, M., 2000. *Assessing Children's Mathematical Knowledge: Social Class, Sex and Problem-Solving*. Open University Press, Buckingham.
- Cooper, B., Glaesser, J., 2012. Qualitative work and the testing and development of theory: lessons from a study combining cross-case and within-case analysis via Ragin's QCA. *Forum Qual. Soc. Res.* 13 (2), 4. <http://nbn-resolving.de/urn:nbn:de:0114-fqs120247>.
- Cooper, B., 1985. *Renegotiating Secondary School Mathematics: A Study of Curriculum Change and Stability*. Falmer Press, Brighton.
- Copland, P., Creese, A., 2015. *Linguistic Ethnography: Collecting, Analyzing and Presenting Data*. Sage, London.
- Cuevas-Parra, P., Tisdall, E.K.M., 2019. Child-led research: questioning knowledge. *Soc. Sci.* 8, 44. Available at: <https://www.mdpi.com/2076-0760/8/2/44>.
- Cusick, P., Everhart, R., 1985. The everhart/cusick reviews: a brief exchange. *Anthropol. Educ. Q.* 16 (3), 246–248.
- Cusick, P., 1973. *Inside the High School*. Holt, Rinehart and Winston, New York.
- Cusick, P., 1983. *The Egalitarian Ideal and the American High School: Studies of Three Schools*. Longman, New York.
- Cusick, P., 1985. Review of reading, writing and resistance: adolescence and labor in a junior high school by Robert B. Everhart. *Anthropol. Educ. Q.* 16 (1), 69–72.
- Deem, R. (Ed.), 1980. *Schooling for Women's Work*. Routledge, London.
- de Landsheere, G., 1988. History of educational research. In: Keeves, J. (Ed.), *Educational Research, Methodology, and Measurement*. Pergamon, Oxford.
- Delamont, S., Atkinson, P., Hammersley, M., 1993. Qualitative research traditions. In: Hammersley, M. (Ed.), *Educational Research: Current Issues*. Paul Chapman, London.
- Delamont, S., Stephens, N., Campos, C., 2017. *Embodying Brazil: An Ethnography of Diasporic Capoeira*. Routledge, London.
- Delamont, S., 1973. *Academic Conformity Observed: Studies in the Classroom*. Unpub. PhD thesis. University of Edinburgh. Available at: <https://era.ed.ac.uk/handle/1842/17148>.
- Delamont, S., 1976. *Interaction in the Classroom*. Methuen, London.
- Delamont, S., 1980. *Sex Roles and the School*. Allen and Unwin, London.
- Denzin, N., Giardina, M., 2006. *Qualitative Inquiry and the Conservative Challenge*. Left Coast Press, Walnut Creek, CA.
- Dixon, A., Rousseau Anderson, C., Donnor, J. (Eds.), 2017. *Critical Race Theory and Education*. Routledge, New York.
- Driver, R., Guesne, E., Tiberghien, A. (Eds.), 1985. *Children's Ideas in Science*. Open University Press, Buckingham.
- Driver, R., 1983. *The Pupil as Scientist*. Open University Press, Buckingham.
- Dubberley, W., 1988. Social class and the process of schooling: a case study of a comprehensive school in a mining community. In: Green, A., Ball, S. (Eds.), *Progress and Inequality in Comprehensive Education*. Routledge, London.
- Dumont, R.V., Wax, M.L., 1971. Cherokee school society and the intercultural classroom. In: Cosin, B.R., Dale, I.R., Esland, G.M., Swift, D.F. (Eds.), *School and Society*. Routledge and Kegan Paul, London.
- Egger, A.S., Unterwiesing, G., 2018. Ethnographic research in schools: historical roots and developments with a focus on Germany and Switzerland. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley, Hoboken, NJ.
- Eisner, E., Peshkin, A. (Eds.), 1990. *Qualitative Inquiry in Education*. Teachers College Press, New York.
- Eisenhart, M., 2018. Changing conceptions of culture and ethnography in anthropology of education in the United States. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley, Hoboken, NJ.
- Erickson, F., Shultz, J., 1981a. The Counselor as Gatekeeper: Social Interaction in Interviews. Academic Press, Orlando, FL.
- Erickson, F., Shultz, J., 1981b. When is a context? Some issues and methods in the analysis of social competence. In: Green, J., Wallat, C. (Eds.), *Ethnography and Language in Educational Settings*. Ablex, Norwood, NJ, pp. 147–160.
- Erickson, F., 1977. Some approaches to inquiry in school-community ethnography. *Anthropol. Educ. Q.* 8, 58–69.
- Everhart, R., 1983. *Reading, Writing and Resistance: Adolescence and Labor in a Junior High School*. Routledge and Kegan Paul, Boston, MA.
- Everhart, R., 1985. Review of the egalitarian ideal and the American high school: studies of three schools by Philip A. Cusick. *Anthropol. Educ. Q.* 16 (1), 72–75.
- Everhart, R., 1989. Science and the person: the research of Willard Waller. *Int. J. Qual. Stud. Educ.* 2 (2), 145–156.
- Fonow, M., Cook, J. (Eds.), 1991. *Beyond Methodology: Feminist Scholarship as Lived Research*. Indiana University Press, Bloomington, IND.
- Foster, P., Gomm, R., Hammersley, M., 1996. *Constructing Educational Inequality*. Falmer, London.
- Foster, P., 1990. Cases not proven: an evaluation of two studies of teacher racism. *Br. Educ. Res. J.* 16 (4), 335–348.
- Francis, B., 1998. *Power Plays: Primary School Children's Construction of Gender, Power, and Adult Work*. Trentham, London.
- French, J., French, P., 1984. Gender imbalances in the primary classroom. *Educ. Res.* 26 (2), 127–136.
- Fuller, M., 1980. In: Deem, R. (Ed.), *Black Girls in a London Comprehensive School*.
- Garcez, P.M., 2008. Microethnography in the classroom. In: Hornberger, N.H. (Ed.), *Encyclopedia of Language and Education*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30424-3_259.
- Gaztambide-Fernandez, R., 2009. *The Best of the Best: Becoming Elite at an American Boarding School*. Harvard University Press, Cambridge, MA.
- Geer, B., 1972. Editorial foreword: the story of the project. *Am. Behav. Sci.* 16 (1), 3–14.
- George, R., 2007. *Girls in a Goldfish Bowl*. Sense Publishers, Rotterdam.
- Gewirtz, S., Ball, S., Bowe, R., 1995. *Markets, Choice and Equity in Education*. Open University Press, Buckingham.

- Gibbs, G., Morgan, A., Taylor, E., 1982. A review of the research of Ference Marton and the Goteborg Group: a phenomenological research perspective on learning. *High Educ.* 11, 123–145.
- Gillborn, D., Drew, D., 1993. The politics of research: some observations on "methodological purity". *New Commun.* 19 (2), 354–360.
- Gillborn, D., 1990. 'Race', Ethnicity and Education. Routledge, London.
- Gillborn, D., 1995. Racism and Anti-racism in Real Schools. Open University Press, Buckingham.
- Gomm, R., 1993. Figuring out ethnic equity. *Br. Educ. Res. J.* 19 (2), 149–165.
- Goode, J., 2019. *Clever Girls: Autoethnographies of Class, Gender and Ethnicity*. Palgrave Macmillan/Springer Nature, Basingstoke.
- Goodson, I., Ball, I. (Eds.), 1984. *Defining the Curriculum: Histories and Ethnographies*. Falmer Press, Lewes.
- Gordon, T., Holland, J., Lahelma, E., 2001. Friends or foes? Interpreting relations between girls in school. In: Walford, G., Hudson, C. (Eds.), *Gender and Sexualities in Educational Ethnography, Studies in Educational Ethnography*, vol. 3. Emerald Group Publishing Limited, Bingley, pp. 7–25.
- Green, J.L., Wallat, C. (Eds.), 1981. *Ethnography and Language in Educational Settings*. Ablex, Norwood, NJ.
- Green, J.L., 1983. Research on teaching as a linguistic process: a state of the art. In: Gordon, E.W. (Ed.), *Review of Research in Education*, vol. 10. American Educational Research Association, Washington, DC, pp. 151–252.
- Guba, E. (Ed.), 1990. *The Paradigm Dialog*. Sage, Newbury Park, CA.
- Hammersley, M., Gomm, R., 1993. A response to Gillborn and Drew on "race," class, and school effects. *New Commun.* 19 (2), 348.
- Hammersley, M., Woods, P. (Eds.), 1984. *Life in Classrooms*. Open University Press, Milton Keynes.
- Hammersley, M., 1974. The organization of pupil participation. *Sociol. Rev.* 22 (3), 355–368.
- Hammersley, M., 1976. The mobilization of pupil attention. In: Hammersley, M., Woods, P. (Eds.), *The Process of Schooling*. Routledge and Kegan Paul, London.
- Hammersley, M., 1980. On interactionist empiricism. In: Woods, P. (Ed.), *Pupil Strategies: Explorations in the Sociology of the School*. Croom Helm, London. Reissued by Routledge, 2012.
- Hammersley, M., 1992. The paradigm wars: reports from the front. *Br. J. Sociol. Educ.* 13 (1), 131–143.
- Hammersley, M., 1995. *The Politics of Social Research*. Sage, London.
- Hammersley, M., 1997. Educational research and teaching: a response to David Hargreaves' TTA lecture. *Br. Educ. Res. J.* 23 (2), 141–161.
- Hammersley, M., 2004. Action research: a contradiction in terms? *Oxf. Rev. Educ.* 30 (2), 165–181.
- Hammersley, M., 2005. Countering the "new orthodoxy" in educational research: a response to Phil Hodgkinson. *Br. Educ. Res. J.* 31 (2), 139–155.
- Hammersley, M. (Ed.), 2007. *Educational Research and Evidence-Based Practice*. Sage, London.
- Hammersley, M., 2013a. *What is Qualitative Research?* Bloomsbury, London.
- Hammersley, M., 2013b. *The Myth of Research-Based Policy and Practice*. Sage, London.
- Hammersley, M., 2018. Ethnography of schooling in England: a history and assessment of its early development. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley, Hoboken, NJ.
- Hargreaves, A., Woods, P. (Eds.), 1984. *Classrooms and Staffrooms*. Open University Press, Milton Keynes.
- Hargreaves, D.H., Hester, S., Mellor, F., 1975. *Deviance in Classrooms*. Routledge and Kegan Paul, London.
- Hargreaves, D.H., 1967. *Social Relations in a Secondary School*. Routledge and Kegan Paul, London.
- Hargreaves, D.H., 1976. Reactions to labelling. In: Hammersley, M., Woods, P. (Eds.), *The Process of Schooling*. Routledge and Kegan Paul, London.
- Hargreaves, A., 1982. Resistance and relative autonomy theories: problems of distortion and incoherence in recent marxist analyses of education. *Br. J. Sociol. Educ.* 3 (2), 108–126.
- Hargreaves, A., 1986. *Two Cultures of Schooling: The Case of Middle Schools*. Brighton, Falmer.
- Hargreaves, D.H., 1996. Teaching as a Research-Based Profession: Possibilities and Prospects (Annual Lecture). Teacher Training Agency, London [Reprinted, along with some responses, in Hammersley ed. 2007].
- Haw, K., 1998. *Educating Muslim Girls*. The Open University Press, Buckingham.
- Hellal, P., Lorch, M., 2010. Darwin's contribution to the study of child development and language acquisition. *Lang. Hist.* 53 (1), 1–14. <https://doi.org/10.1179/175975310X12640878626147>.
- Hendricks, C., 2019. History of action research in education. In: Mertler, C. (Ed.), *The Wiley Handbook of Action Research in Education*. Wiley-Blackwell, Hoboken, NJ.
- Hey, V., 1997. *The Company She Keeps*. Open University Press, Buckingham.
- Hillage, J., Pearson, R., Anderson, A., Tamkin, P., 1998. *Excellence in Research on Schools*. Department for Education and Employment, London.
- Hodkinson, P., 2004. Research as a form of work: expertise, community and methodological Objectivity. *Br. Educ. Res. J.* 30 (1), 9–26.
- Howard, A., 2008. *Learning Privilege: Lesson of Power and Identity in Affluent Schooling*. Routledge, New York.
- Jackson, P.W., 1968. *Life in Classrooms*. Holt, Rinehart and Winston, New York.
- Jacob, E., 1987. Qualitative research traditions. *Rev. Educ. Res.* 57 (1), 1–50.
- Jeffrey, B., Woods, P., 1998. *Testing Teachers: The Effects of School Inspections on Primary Teachers*. Routledge, London.
- Kemmis, S., 1988. Action research. In: Keeves, J. (Ed.), *Educational Research, Methodology, and Measurement*. Pergamon, Oxford.
- Kenway, J., Kraack, A., Hickey-Moody, A., 2006. *Masculinity Beyond the Metropolis*. Palgrave Macmillan, Basingstoke.
- Khan, S., 2011. *Privilege: The Making of an Adolescent Elite*. Princeton University Press, Princeton, NJ.
- Kuhn, T., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Lacey, C., 1966. Some sociological concomitants of academic streaming in a grammar school. *Br. J. Sociol.* XVII (3), 245–262.
- Lacey, C., 1970. *Hightown Grammar*. Manchester University Press, Manchester.
- Lagemann, E.C., 2000. *An Elusive Science: The Troubling History of Education Research*. University of Chicago Press, Chicago, IL.
- Lambart, A., 1976. The sisterhood. In: Hammersley, M., Woods, P. (Eds.), *The Process of Schooling*. Routledge and Kegan Paul, London.
- Lambart, A., 1982. Expulsion in context: a school as a system in action. In: Frankenberg, R. (Ed.), *Custom and Conflict in British Society*. Manchester University Press, Manchester.
- Lareau, A., 1989. *Home Advantage*. Falmer, London.
- Lareau, A., 2003. *Unequal Childhoods*. University of California Press, Berkeley.
- Lather, P., St Pierre, E., 2013. Post-qualitative research. *Int. J. Qual. Stud. Educ.* 26 (6), 629–673.
- Lather, P., 2004a. This is your father's paradigm: governmental intrusion and the case of qualitative research in education. *Qual. Inq.* 10 (1), 15–34.
- Lather, P., 2004b. Scientific research in education: a critical perspective. *Br. Educ. Res. J.* 30 (6), 759–771.
- Lather, P., 2006. Paradigm proliferation as a good thing to think with: teaching research in education as a wild profusion. *Int. J. Qual. Stud. Educ.* 19 (1), 35–57.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Leavy, P. (Ed.), 2018. *Handbook of Arts-Based Research*. Guildford Press, New York.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage, Newbury Park, CA.
- Littleton, K., 2010. Research into teaching with whole-class interactive technologies: emergent themes. *Technol. Pedagog. Educ.* 19 (2), 285–292.
- Lortie, D.C., 1975. *Schoolteacher: A Sociological Study*. University of Chicago Press, Chicago, IL.
- Lortie, D.C., 2009. *School Principal: Managing in Public*. University of Chicago Press, Chicago, IL.
- Lutz, F., 1981. In: Green, J., Wallat, C. (Eds.), *Ethnography: The Holistic Approach to Understanding Schooling*.
- Lynn, M., Parker, L., 2006. Critical race studies in education: examining a decade of research on US schools. *Urban Rev.* 38 (4), 257–290.

- Mac an Ghaill, M., 1988. *Young, Gifted and Black*. Open University Press, Buckingham.
- Mac an Ghaill, M., 1994. *The Making of Men: Masculinities, Sexualities and Schooling*. Open University Press, Buckingham Philadelphia.
- MacDonald, B., 1974. Education and the control of education. In: MacDonald, B., Walker, R. (Eds.), *Innovation, Evaluation, Research and the Problem of Control*. Centre for Applied Research in Education, University of East Anglia.
- MacLure, M., 2003. *Discourse in Educational and Social Research*. Open University Press, Buckingham.
- Maguire, M., Ball, S., Macrae, S., 2001. "In all our interests": internal marketing at Northwark Park School. *Br. J. Sociol. Educ.* 22 (1), 35–50.
- Marton, F., 1981. Phenomenography—describing conceptions of the world around us. *Instr. Sci.* 10 (1981), 177–200.
- Maybin, J., 2006. *Children's Voices: Talk, Knowledge and Identity*. Palgrave Macmillan, Basingstoke.
- McDermott, R.P., 1977. Social relations as contexts for learning in school. *Harv. Educ. Rev.* 47, 198–213.
- McDermott, R.P., 2008. Reading George Spindler. *Anthropol. Educ. Q.* 39 (2), 117–126.
- McGrew, K., 2011. A review of class-based theories of student resistance in education: mapping the origins and influence of learning to labor by Paul Willis. *Rev. Educ. Res.* 81 (2), 234–266.
- McHoul, A., 1978. The organization of turns at formal talk in the classroom. *Lang. Soc.* 7 (2), 183–213.
- Mead, M., 1928. *Coming of Age in Samoa*. William Morrow, New York.
- Measor, L., 1983. Gender and the sciences: pupils' gender-based conceptions of school subjects. In: Hammersley, M., Hargreaves, A. (Eds.), *Curriculum Practice*. Falmer, London.
- Mehan, H., 1979. *Learning Lessons: Social Organization in the Classroom*. Harvard University Press, Cambridge, MA.
- Miller, C.M.L., Parlett, M., 1974. *Up to the Mark: A Study of the Examination Game*. Society for Research into Higher Education. Monograph 21.
- Mills, D., 2012. Anthropology and education. In: Delamont, S. (Ed.), *Handbook of Qualitative Research in Education*. Edward Elgar, Cheltenham.
- Milstein, D., Clemente, A., 2018. Latin American educational ethnography. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley, Hoboken NJ.
- Modiba, M., Stewart, S., 2018. Ethnography and education in an African context. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley, Hoboken NJ.
- National Research Council, 2002. In: Shavelson, R.J., Towne, L. (Eds.), *Scientific Research in Education*. National Academies Press, Washington, DC.
- Ogbu, J., 1974. *The Next Generation: An Ethnography of Education in an Urban Neighborhood*. Academic Press, New York.
- Olesen, V., Whittaker, E., 1968. *The Silent Dialogue*. Jossey-Bass, San Francisco, CA.
- Parlett, M., Hamilton, D., 1972. *Evaluation as Illumination: A New Approach to the Study of Innovation Programmes*. Scottish Centre for Research in Educational Sciences, Edinburgh.
- Parlett, M.R., Hamilton, D., 1977. Evaluation in illumination: a new approach to the study of innovative programs. In: Hamilton, D. (Ed.), *Beyond the Numbers Game*. Macmillan Education, London.
- Payne, G.C.F., 1976. Making a lesson happen: an ethnomethodological analysis. In: Hammersley, M., Woods, P. (Eds.), *The Process of Schooling*. Routledge and Kegan Paul, London.
- Perryman, J., Ball, S., Maguire, M., Braun, A., 2011. Life in the pressure cooker: school league tables and English and mathematics teachers' responses to accountability in a results-driven era. *Br. J. Educ. Stud.* 59 (2), 179–195.
- Perryman, P., Ball, S., Braun, A., Maguire, M., 2017. Translating policy: governmentality and the reflective teacher. *J. Educ. Pol.* 32 (6), 745–756.
- Peshkin, A., 1978. *Growing Up American: Schooling and the Survival of Community*. University of Chicago Press, Chicago.
- Phillips, S., 1983. *The Invisible Culture: Communication in School and Community on the Warm Springs Indian Reservation*. Longman, New York.
- Piper, H., Stronach, I. (Eds.), 2004. *Educational Research: Difference and Diversity*. Ashgate, Aldershot.
- Pollard, A., 1985. *The Social World of the Primary School*. Holt, Rinehart and Winston, London.
- Pollard, A., 1996. *The Social World of Children's Learning: Case Studies of Pupils From Four to Seven*. Continuum, London.
- Pollard, A., 1999. *The Social World of Pupil Careers: Strategic Biographies Through Primary School*. Cassell, London.
- Preyer, W.T., Brown, H.W., 1888. *The Mind of the Child. Observations Concerning the Mental Development of the Human Being in the First Years of Life*, 2 vols. D. Appleton & Co., New York.
- Prosser, J., 2007. Visual methods and the visual culture of schools. *Vis. Stud.* 22 (1), 13–30.
- Rampton, B., 2006. *Language in Late Modernity: Interaction in an Urban Classroom*. Cambridge University Press, Cambridge.
- Reid, A., Hart, E.P., Peters, M., 2014. *A Companion to Research in Education*. Springer, Dordrecht.
- Renold, E., 1995. *Girls, Boys, and Junior Sexualities*. Routledge, London.
- Renold, E., 2001. Square-girls', femininity and the negotiation of academic success in the primary school. *Br. Educ. Res. J.* 27 (5), 577–588.
- Rist, R., 1973. *The Urban School: A Factory for Failure*. Routledge, London.
- Rosiek, J.L., Snyder, J., Pratt, S.L., 2020. The new materialisms and indigenous theories of non-human agency: making the case for respectful anti-colonial engagement. *Qual. Inq.* 26 (3–4), 331–346.
- Roth, W.-M., Radford, L., LaCroix, L., 2012. Working with cultural-historical activity theory. *Forum Qual. Soc. Res.* 13 (2), 23. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1202232>.
- Saeed, A., 2021. *Education, Aspiration and Upward Social Mobility: Working Class British Pakistani Women*. Palgrave Macmillan, Basingstoke.
- Scheurich, J., Young, M., 1997. Coloring epistemologies: are our research epistemologies racially biased? *Educ. Res.* 26 (4), 4–16.
- Sharp, R., Green, A., 1976. *Education and Social Control*. Routledge and Kegan Paul, London.
- Sikes, P., Measor, L., Woods, P., 1985. *Teacher Careers*. Falmer, London.
- Sikes, P., 2012. The literary turn: fictions and poetry. In: Delamont, S. (Ed.), *Handbook of Qualitative Research in Education*. Elgar, Cheltenham.
- Sinclair, J., Coulthard, M., 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford University Press, Oxford.
- Skelton, C., Francis, B. (Eds.), 2003. *Boys and Girls in the Primary Classroom*. Open University Press, Maidenhead.
- Skelton, C., 1996. Primary boys and hegemonic masculinities. *Br. J. Sociol. Educ.* 18 (3), 349–p370.
- Skelton, C., 2001. *Schooling the Boys: Masculinities and Primary Education*. Open University Press, Buckingham.
- Smith, L.M., Geoffrey, W., 1968. *The Complexities of an Urban Classroom. An Analysis Toward a General of Teaching*. Holt, Rinehart and Winston, New York.
- Spindler, G. (Ed.), 1955. *Anthropology and Education*. Stanford University Press, Stanford, CA.
- St Pierre, E., 2019. Post qualitative inquiry, the refusal of method, and the risk of the new. *Qual. Inq.* <https://doi.org/10.1177/1077800419863005>.
- Stake, R., 1995. *The Art of Case Study*. Sage, Thousand Oaks, CA.
- Stronach, I., MacLure, M., 1997. *Educational Research Undone: The Postmodern Embrace*. Open University Press, Buckingham.
- Stubbs, M., Delamont, S. (Eds.), 1976. *Explorations in Classroom Observation*. Wiley, Chichester.
- Stubbs, M., 1975. Teaching and talking. In: Chanan, G., Delamont, S. (Eds.), *Frontiers of Classroom Research*. National Foundation for Educational Research, Slough.
- Stubbs, M., 1976. Keeping in touch: some functions of teacher talk. In: Stubbs, M., Delamont, S. (Eds.), *Explorations in Classroom Observation*. Wiley, London.
- Swain, J., 2006. Reflections on patterns of masculinity in school settings. *Men Masc.* 8 3, 331–349.
- Taylor, E., Gillborn, D., Ladson-Billings, G. (Eds.), 2016. *Foundations of Critical Race Theory in Education*. Routledge, London.
- Troyna, B., 1993. Underachiever or misunderstood? A reply to Roger Gomm. *Br. Educ. Res. J.* 19 (2), 167–174.
- Tooley, J., Darby, D., 1998. *Educational Research: A Critique: A Survey of Published Educational Research*. Ofsted, London.

- Torrance, H., 2011. Qualitative research, science and government: evidence, criteria, policy and politics. In: Denzin, N., Lincoln, Y. (Eds.), *Sage Handbook of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA.
- Traianou, A., 2006. *Understanding Teacher Expertise in Primary Science*. Sense Publishers, Rotterdam.
- Troman, G., Woods, P., 2001. *Primary Teachers' Stress*. RoutledgeFalmer, London.
- Turner, G., 1979. Review of Paul Willis *Learning to Labor*. *Sociology* 13 (2), 336–338.
- Walker, S., Barton, L. (Eds.), 1983. *Gender, Class, and Education*. Falmer, London.
- Walker, R., 1978. The conduct of educational case studies: ethics, theory and procedure. In: Dockrell, B., Hamilton, D. (Eds.), *Rethinking Educational Research*. Hodder and Stoughton, London.
- Walker, J.C., 1985. Rebels without applause? A critique of resistance theory in Paul Willis's ethnography of schooling. *J. Educ.* 167 (2), 63–83.
- Waller, W., 1932. *The Sociology of Teaching*. Wiley, New York.
- Warren, R.L., 1967. *Education in Rebhausen: A German Village*. Holt, Rinehart and Winston, New York.
- Wax, M., Diamond, S., Gearing, F., 1971. *Anthropological Perspectives on Education*. Basic Books, New York.
- Weis, L., Fine, M. (Eds.), 2000. *Construction Sites: Excavating Race, Class, and Gender Among Urban Youth*. Teachers College Press, New York.
- Weis, L., 1990. *Working Class Without Work: High School Students in a Deindustrializing Economy*. Routledge, New York.
- Weis, L., 2004. *Class Reunion: The Remaking of the American White Working Class*. Routledge, New York.
- Williams, A., Katz, L., 2001. The use of focus group methodology in education: some theoretical and practical considerations. *Int. Electron. J. Leader. Learn* 5 (3). Available at: https://static1.squarespace.com/static/5e17a6b1faf6d13c584588e4/t/5e2f4b31d6c2003bc0798cb4/1580157746598/The_Use_of_Focus_Group_Methodology_in_Education_So.pdf.
- Willis, P., 1977. *Learning to Labor*. Saxon House, Farnborough.
- Willower, D., Boyd, W. (Eds.), 1989. *Willard Waller on Education and Schools: A Critical Appraisal*. McCutchan, Berkeley.
- Wolcott, H., 1984. *The Man in the Principal's Office: An Ethnography*. Waveland, Prospect Heights, IL.
- Woods, P., Hammersley, M. (Eds.), 1977. *School Experience*. Croom Helm, London.
- Woods, P., Hammersley (Eds.), 1993. *Gender and Ethnicity in Schools*. Routledge, London.
- Woods, P., Jeffrey, B., Troman, G., Boyle, M., 1997. *Restructuring Schools, Reconstructing Teachers: Responding to Change in the Primary School*. Open University Press, Buckingham.
- Woods, P., 1979. *The Divided School*. Routledge and Kegan Paul, London.
- Woods, P., 1984. Educational ethnography in Britain. *J. Thought* 19 (2), 75–94.
- Wright, H.F., 1967. *Recording and Analyzing Child Behavior*. Harper and Row, New York.
- Wylie, L., 1964. *Village in the Vaucluse*, second ed. Harvard University Press, Cambridge, MA.
- Yodell, D., 2009. Queer outings: uncomfortable stories about the subjects of post-structural school ethnography. *Int. J. Qual. Stud. Educ.* 23 (1), 87–100.
- Young, M.F.D. (Ed.), 1971. *Knowledge and Control*. Collier-Macmillan, London.

Qualitative research approaches and their application in education

Peggy M. Delmas and Rachel L. Giles, University of South Alabama, Department of Leadership & Teacher Education, Mobile, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

Case study	24
Phenomenology	26
Grounded theory	27
Narrative inquiry	27
Ethnography	28
Discourse analysis	29
Participatory action research	29
Conclusion	30
References	30

For scholars who are not familiar with or have not previously conducted qualitative research, the options available for framing and carrying out a study can seem bewildering. It is important at the outset to define qualitative research. [Creswell and Poth \(2018\)](#) stated that qualitative research:

... begins with assumptions and the use of interpretive/theoretical frameworks that inform the study of research problems addressing the meaning individuals or groups ascribe to a social or human problem. To study this problem, qualitative researchers use an emerging qualitative approach to inquiry, the collection of data in a natural setting sensitive to the people and places under study, and data analysis that is both inductive and deductive and establishes patterns or themes. The final written report or presentation includes the voices of participants, the reflexivity of the researcher, a complex description and interpretation of the problem, and its contribution to the literature or a call for change (p. 43).

Our colleagues, through accompanying chapters in this volume will describe the philosophical underpinnings of qualitative research as well as the history of qualitative research in education. In this chapter however, our purpose is to simplify the choice of methods for scholars by presenting a more easily understandable overview of qualitative research approaches. We choose the term qualitative approaches ([Creswell and Poth, 2018](#)) here to refer to the ways in which the research can be organized or framed, although in the literature other terms appear such as varieties ([Tesch, 1990](#)), categories ([Slife and Williams, 1995](#)), traditions ([Creswell, 1998](#)), methods ([Morse and Richards, 2002](#)), and strategies of inquiry ([Denzin and Lincoln, 2005](#)).

For new qualitative researchers the problem is not that there is not enough information on the approaches available, but that there exists an almost baffling array and volume of literature. Further, within the literature of many approaches is a tangled history of the development, key scholars, and resulting variations of the approaches, which can lead to the confusion and frustration of many beginning researchers. Thus, we have approached writing this chapter from the viewpoint of an academic (first author) who assists students with doctoral research in the field of education and a graduate student (second author) who is currently conducting doctoral research in education. In our discussion of each qualitative approach, we have taken pains to be clear on its definition and characteristic features, to provide a brief background, including key scholars, to describe its evolution and resulting variants of the approach, to discuss its strengths and weaknesses, and to provide examples of studies in education using the approach. Additionally, in the spirit of this international publication we have tried to include studies from around the world representative of the various qualitative approaches as applied in the field of education. For the sake of simplification and due to space constraints, we will discuss the five established qualitative approaches (i.e., case study, phenomenology, grounded theory, narrative inquiry, and ethnography) ([Creswell and Poth, 2018](#)), as well as two additional qualitative approaches (i.e., discourse analysis and participatory action research) more commonly seen in educational research studies. [Table 1](#) provides an overview of the qualitative approaches, key characteristics, and representative studies in this chapter.

Case study

Case study is a popular method of conducting research as evidenced by our recent database query which yielded 465 English language dissertations with “case study” in the title published in the last 5 years. Case studies have been used in a number of disciplines and fields of study including psychology, anthropology, sociology, law, and medicine. A case study is defined as the in-depth study of a case or cases in a real-life, contemporary context or setting ([Yin, 2014](#)). Case studies have historically been used to answer how or why questions ([Yin, 2014](#)). A case, or the object of the study, can be an organization, program, or

Table 1 Overview of qualitative research approaches, key characteristics, and representative studies.

<i>Qualitative research approach</i>	<i>Key characteristics</i>	<i>Representative studies</i>
Case study	A case study is defined as the in-depth study of a case or cases in a real-life, contemporary context or setting (Yin, 2014). Case studies have been used in a number of disciplines and fields of study including psychology, anthropology, sociology, law, and medicine. Case studies have historically been used to answer how or why questions (Yin, 2014).	Single case study: Hoppey and McClesky (2013) Multiple case study: Delmas and Ivankova (2018)
Phenomenology	Phenomenology is a method in which researchers make it their focal point to study human experiences, including the ways people conceptualize their lived, shared experiences (Braun and Clarke, 2013; Sokolowski, 2000). Phenomenology concentrates on the perceptions and communications of people regarding certain events or objects. This approach can be used to analyze individual cases or to develop cases from a small group of research participants (Braun and Clarke, 2013).	Transcendental phenomenology: Moerer-Urdahl and Creswell (2004) Hermeneutic phenomenology: DeGagne and Walters (2010)
Grounded theory	Grounded theory is defined as the discovery of theory from data (Glaser and Strauss, 1967). This approach focuses on a certain development or activity that is unfolding, ideally leading to the researcher's creation of a theory regarding this series of events or process (Creswell and Poth, 2018). Essentially, the theory centers on the researcher's end goal of forming a theoretical frame which sheds light upon a particular phenomenon, or series of events, including how the phenomenon began, grew, and changes (Charmaz, 2014).	Classic grounded theory: Chametzky (2015) Straussian grounded theory: Vassilakaki and Johnson (2015) Constructivist grounded theory: Karpouza and Emvalotis (2019)
Narrative inquiry	Narrative is both a phenomenon and a method (Connelly and Clandinin, 1990). Narrative inquiry, also called narrative research is the study of people's experiences. Narratives or stories are the object of analysis (Caine et al., 2013). The experience can be studied by "listening, observing, living alongside another, writing, and interpreting texts" (Clandinin, 2006, p. 46).	Autobiography: Grace (2006) Biographical study: Stroobants (2005) Life history: Tsui (2007) Oral history: Delaney and Raftery (2020) Testimonio: Cordova and Knecht (2019) Visual narrative inquiry: Sairanen and Kumpulainen (2014) Collaborative autoethnography: Chang et al. (2014) Critical ethnography: Smyth (2016) Institutional ethnography: Lund (2012)
Ethnography	Ethnography is employed when the researcher's goal is to document what actually happens in a natural setting. The data collection typically occurs over a long period of time, and includes a range of data types, with a heavy emphasis on participant observation. Researchers agree that ethnography does not focus on the culture of a group, but rather on the patterns of a group (Wolcott, 2008), or the beliefs, behaviors, language, and interactions of a group which shares the same culture (Creswell and Poth, 2018).	Descriptive discourse analysis: Iverson (2016) Critical discourse analysis Thompson (2002)
Discourse analysis	As a research approach discourse analysis has application across a broad array of fields of study, including communications, anthropology, sociology, and nursing. The two primary types of discourse analysis are descriptive and critical. Descriptive discourse analysis seeks to understand how language influences people and situations. Critical discourse analysis seeks to expose through description and interpretation the ways that social inequalities are constructed, maintained, and legitimized (Mullet, 2018).	Participatory action research: Ditrano and Silverstein (2006), Halliday et al. (2019)
Participatory action research	The term participatory action research (PAR) is considered an umbrella term because it encompasses a range of approaches to participatory action-oriented studies (Kindon et al., 2007). Participatory action research combines elements of community-based research, popular education, and action for social change (Lewin, 1946; Williams and Brydon-Miller, 2004). Brydon-Miller and Maguire state that PAR is "an openly and unapologetically political approach to knowledge creation through and for action" (2009, p. 79). PAR is a collaboration between researchers and participants whereby a problem or issue is investigated with the purpose of finding a solution or seeking improved conditions. Key to PAR is its central feature of democratizing research through a partnership between researchers and participants.	

process (Yin, 2003), instant, event, occasion, or person (Schwandt, 2015). The defining characteristic of a case study is a delimiting of the case (Merriam, 2001). A case is a bounded system. In other words, it is bounded by time and place. Another hallmark of case study is its use of multiple sources of information (Creswell and Poth, 2018) such as interviews, observations, documents, artifacts, and audiovisual material. These multiple sources of information should lead the researcher to the same findings, thus aiding in triangulation and helping to establish the credibility of the research (Yazan, 2015).

There are two types of case designs, case studies and multiple case studies, depending on the number of cases being studied. Case studies may be further classified as exploratory, descriptive, or explanatory (Yin, 2014), or intrinsic or instrumental (Stake, 1995). Exploratory case studies are used as a means to identify the research questions or procedures that will be employed in a future study. Descriptive case studies aim to describe the phenomenon, or case, in its natural context. Explanatory case studies aim to reveal how a condition or situation arose. An intrinsic case study emerges from the need to learn about a particular case, whereas an instrumental case study is employed when one seeks to learn about a larger issue or phenomenon by exploring one case.

Some of the strengths of the case study approach are that it can aid in the discovery of the unexpected and unusual, facilitate an understanding of complex relationships (Hodkinson and Hodkinson, 2001), and provide an in-depth grasp of a case. Weaknesses or challenges inherent in the case study approach include difficulty in identifying the case and determining its delimitations, and the potential for accumulating an unmanageable amount of data.

Yazan (2015) offers a comparative analysis of three case study approaches by leading methodologists which may be useful for scholars considering designing a case study. Examples of two different types of educational case studies are described here as illustrations of single and multiple case studies. Hoppey and McClesky (2013) used a single case study to examine the “role of the principal in supporting school improvement during an era of high-stakes accountability” (p. 247). In this study the lived experience of the principal was the case, bounded by time (one school year) and place (the elementary school in rural Florida). Delmas and Ivan-kova (2018) employed a descriptive, multiple case study in their examination of the academic experiences of three Catholic nuns serving as faculty members at universities in the American South. In their research, the academic experiences of each woman served as a case, and was bounded by time (February 2005–April 2006) and place (the geographic boundaries of the research sites in the South).

Case studies are a popular qualitative research approach in education. Their pervasiveness means that most researchers are likely to be familiar with case studies. The case study approach allows the researcher to develop a deep understanding of the object of study, whether the reason for the research is to learn more about a unique phenomenon or to investigate a specific problem or issue. To learn more about case studies, see the chapter on Case Study Research by R. Stake in this volume.

Phenomenology

Phenomenology is a method in which researchers focus on studying human experiences, including the way people conceptualize their lived, shared experiences (Braun and Clarke, 2013; Sokolowski, 2000). Phenomenology concentrates on the perceptions and communications of people regarding certain events or objects. This approach can be used to analyze individual cases or to develop cases from a small group of research participants (Braun and Clarke, 2013).

Phenomenology's popularity in the US has waxed and waned with initial enthusiasm for the method from scholars in the 1960s and a resurgence of interest from cognitive scientists in the 1990s (Gallagher, 2012). Dall'Alba (2009, p. 7) noted that “as the impact of phenomenology extended beyond philosophy to fields such as sociology, education, anthropology, politics, linguistics, psychology and feminist studies, it has challenged taken-for-granted assumptions and prompted new insights into what it means to live, work, play and learn in our world.”

Braun and Clarke note that while there are a number of types of phenomenology, the method can generally be described as concerned with understanding people's subjective experiences (2013). The two primary types of phenomenology are transcendental (also referred to as empirical or psychological) and hermeneutic. Transcendental phenomenology is built on the philosophical work of Husserl (1931) and applied specifically as a qualitative approach by Moustakas (1994). It is descriptive and focuses on how experiences are structured and organized. This approach “analyzes the essences perceived by consciousness with regard to individual experiences” (Padilla-Diaz, 2015, p. 103). Hermeneutic phenomenology, based on the work of Heidegger (1972) is defined as research centered on personal experiences and requiring interpretations or descriptions of the meanings and significance of phenomena which participants experience (Padilla-Diaz, 2015). This approach requires a historical analysis of the phenomenon and its effects on participants.

An example of a transcendental phenomenology in education is Moerer-Urdahl and Creswell's (2004) study of nine participants in a US college mentoring leadership program which sought to understand the meaning of the mentors' experiences. This study may be especially helpful to researchers planning to undertake a transcendental phenomenology as it describes in detail the process of organizing and analyzing data according to Moustakas's (1994) tradition. While hermeneutic phenomenology is used much more commonly in the health professions, DeGagne and Walters (2010) have applied the approach to a study of the lived experiences of 11 online educators from varying locations in the US. From the study results, the authors proposed recommendations for the improvement of online learning communities. Padilla-Diaz (2015) offers a clear discussion of basic concepts and essential elements of phenomenology for those seeking further guidance to this qualitative research approach.

While phenomenology allows for the in-depth exploration of the meaning of a lived experience or a phenomenon using a small group of participants, there are challenges to this approach. Finding participants who have experienced exactly the same phenomenon may prove difficult. Additionally, the bracketing (also called *epoché*) of the researcher's experiences and assumptions, required in transcendental phenomenology, may be onerous to inexperienced scholars.

Phenomenology is a versatile qualitative approach. It can be employed to explore a single concept or to analyze that concept as it is experienced by a group of people. Phenomenology can assist in discovery of new facets of experiences and phenomena that are so common as to be considered mundane. As Lavery (2003) explained, phenomenology can aid in a re-examination of an experience or phenomenon that may reveal different perspectives, helping to cast them in a new light. For more information about phenomenology, see the chapter on Phenomenology and Education: Researching Pedagogical Experience by N. Friesen in this volume.

Grounded theory

Grounded theory is defined as the discovery of theory from data (Glaser and Strauss, 1967). This approach focuses on a certain development or activity that is unfolding, ideally leading to the researcher's creation of a theory regarding this series of events or process (Creswell and Poth, 2018, p. 83). Essentially, the theory centers on the researcher's end goal of forming a theoretical frame which sheds light upon a particular phenomenon, or series of events, including how the phenomenon began, grew, and changes (Charmaz, 2014). Because of this, grounded theory can lead to the development of a useful theoretical frame which helps to explain an incomplete or disorganized phenomenon (Corbin and Strauss, 2015). Grounded theory's application has grown widely and now appears in numerous areas of study including education, medicine, nursing, psychology, social work, religion, gender studies, and anthropology. Morse (2009) speculates that grounded theory may be the most commonly used qualitative research method.

Two American sociologists, Barney G. Glaser and Anselm L. Strauss, are considered the fathers of grounded theory (Kenny and Fourie, 2014). They purposefully named the method *grounded theory* in order to describe its intent to "ground" theory in empirical research. Conducting a classical grounded theory study involves an iterative process of coding, memoing, sorting, conceptualization, and constant comparison. Over time the approaches that Glaser and Strauss took to grounded theory diverged, resulting in classical grounded theory and Straussian grounded theory. Straussian grounded theory grew out of a collaboration between Strauss and Corbin (1990), and relied on an analytical coding framework designed to systematically infer theory from data (Kenny and Fourie, 2014). A student of Glaser and Strauss, Kathy Charmaz, is associated with a third variant of grounded theory, called constructivist grounded theory (Charmaz 2006, 2014), which highlights the phenomenon under investigation rather than the methods utilized to study it (Charmaz, 2005).

Hussein et al. (2014) state that the strengths of grounded theory include its richness of data, cultivation of creativity, intuitive appeal, capacity for conceptualization, and systematic approach to analyzing data. Alternatively, the authors list weaknesses of the theory as "exhaustive process, potential for methodological errors, reviewing the literature without developing assumptions, multiple approaches to GT, limited generalizability" (p. 3). As a remedy to aid in understanding the multiple approaches, Kenny and Fourie's (2014) article does an admirable job of laying out the history of grounded theory and tracing its development through the three variants of grounded theory. Additionally, Tan's (2010) article may be useful to scholars new to the grounded theory as it attempts to clarify common confusions with the approach through an examination of the author's doctoral study investigating post-secondary students' learning and knowledge sharing as evidenced through blogging.

Following are examples of the three common variants of grounded theory as used to frame research studies. Chametzky (2015) used a classic grounded theory approach to understand the issues of contingent, part-time faculty in US higher education, with a sample of 11 participants. The resulting theory, surviving situational suffering, described how contingent faculty dealt with their primary concern—remaining employed—while simultaneously operating under conditions of lesser appreciation, gratitude, and utilization (Chametzky, 2015). An example of a Straussian grounded theory study can be found in the research of Vassilakaki and Johnson (2015) which used 24 undergraduate and graduate participants from Manchester Metropolitan University in the UK to identify elements of the user experience in information searching on an experimental multilingual information retrieval system. Karpouza and Emvalotis (2019) employed a constructivist grounded theory using 20 teacher educators and 25 graduate students from master's programs in Educational Sciences in Greek universities to explore the teacher-student relationship in graduate education. To learn more about grounded theory, see the chapter on Grounded Theory by W. Babchuk in this volume.

Narrative inquiry

Narrative is both a phenomenon and a method (Connelly and Clandinin, 1990). Narrative inquiry, also called narrative research is the study of people's experiences. Narratives or stories are the object of analysis (Caine et al., 2013). The experience can be studied by "listening, observing, living alongside another, writing, and interpreting texts" (Clandinin, 2006, p. 46). Researchers in the social sciences began using narrative inquiry in the late 1980s. Connelly and Clandinin (1990) first used the approach in education to recount the personal stories of teachers. The types of data used in narrative inquiry, which Clandinin and Connelly (2000) call *field texts* include photographs, field notes, conversation transcripts, and interview transcripts. To the list of potential data sources Savin-Baden and Niekirk (2007) added letter writing, and autobiographical and biographical writing.

Types of narrative inquiry include: autobiography, where the narrative is created by the researcher from their experiences; biographical study, where the researcher reports on the life of a participant; life history, which recounts a participant's entire life; oral history, which is composed of the reflections of one or more participants on events and their causes and effects; testimonio, which is reported in first-person by the participant and urges the reader into acknowledgment of the urgency of addressing social wrongs; and visual narrative inquiry, where images are analyzed alongside text.

Savin-Baden and Nierkerk (2007) have pointed to the ease in getting people to talk and the ease in accumulating data as strengths of the narrative inquiry approach. A counterpoint is that the data can be cumbersome and participants may not report their experiences in a linear or chronological fashion. Additionally, Clandinin and Connelly (2000) caution that power relations between the researcher and participant may pose a challenge to the narrative inquiry approach. Because of its characteristic of research over time, narrative inquiry has been recognized as a particularly effective methodology in the field of medicine where it can be used to tell the stories of individuals with chronic illness (Bleakley, 2005). Narrative inquiry is also noted as being effective in educational leadership to aid individuals in career path exploration and identity formation (Floyd, 2012).

Examples of the six types of narrative inquiry previously described are given here as starting points for scholars who wish to learn more about this research approach. Grace (2006) employed queer autobiography to link the personal to positional practice as an educator and gay man living in Canada. Stroobants (2005) reflected on the stories told by both participants and the researcher as well as the learning processes that they went through during a narrative biographical study of eight women who had participated in adult education initiatives. To explore the processes of teacher identity formation, Tsui (2007) used life history to examine the professional identity of an English as a Foreign Language teacher in the People's Republic of China. Through oral history accounts Delaney and Raftery (2020) provided insight into the ways in which Ireland's Free Education Scheme of 1967 impacted the lives of 16 Catholic sister educators with leadership/management responsibilities. Employing testimonios, Cordova and Knecht (2019) explored the perceptions of three Chicana/o doctoral students in an US educational leadership program regarding the intersection of their gender, race, and ethnicity as it related to their academic advancement and leadership development. Finally, Sairanen and Kumpulainen (2014) used visual narrative inquiry to explore the sense of agency in educational settings for five Finnish children in pre-school and first grade.

The narrative inquiry approach is best used when the researcher wants to explore the life experiences of a person or persons. It is achieved through the participant telling stories about some aspect of their life. The researcher's responsibility is to organize the stories into a meaningful, typically chronological format that represents the participant's experiences. The end result often helps provide insight not only into the participant's life, but also the society, culture, or era in which their experiences occurred. Further information on narrative inquiry can be found in D.J. Clandinin's Narrative Inquiry chapter in this volume.

Ethnography

Like many other qualitative methods, ethnography is both a research approach and a research product. Researchers agree that ethnography does not focus on the culture of a group, but rather on the patterns of a group (Wolcott, 2008), or the beliefs, behaviors, language, and interactions of a group which shares the same culture (Creswell and Poth, 2018). Ethnography is employed when the researcher's goal is to document what actually happens in a natural setting. The data collection typically occurs over a long period of time, and includes a range of data types, with a heavy emphasis on participant observation. Ethnography evolved from the field of anthropology in the early 20th century and its use in education developed in the 1960s and 1970s.

Hammersley (2018) notes over 40 adjectives used to describe various types of ethnographic studies (e.g., collaborative, feminist, insider, longitudinal, rural). However, some of the more commonly encountered types of ethnography in education include autoethnography, critical ethnography, and institutional ethnography. Autoethnography is usually written in first-person and allows the author to examine their personal reflections to illuminate a particular event, culture, or institution. It results in the author's deeper understanding of self (Skott-Myhre et al., 2012). Critical ethnography uncovers the relationships between structures, such as cultural and political institutions and norms, and the agency of individuals and groups (Castagno, 2012). According to Anderson "... the overriding goal of critical ethnography is to free individuals from sources of domination and repression" (1989, p. 249). Institutional ethnography was developed and named by Canadian sociologist Dorothy E. Smith in the 1980s. Institutional ethnographies are created through "the examination of work processes and how they are coordinated, typically through texts and discourses ..." (Devault, 2006, p. 294).

Strengths of ethnography as a research approach include that it can be carried out in an array of settings by an individual researcher, and that it "relies essentially on a human observer to observe humans" (Wolcott, 2008, p. 66). The duration of an ethnographic study is highly variable, ranging from years, sometimes spanning a researcher's entire career, to a much shorter period of a few days to a month, in what Jeffrey and Troman have called a "compressed time mode," which requires the near constant presence of the researcher at the research site (2004, p. 538). A criticism of ethnographies created from shorter time in the field is that they are superficial because the researcher has not had extensive exposure to participants and a fuller range of their behaviors. Access can also be an issue for ethnographic researchers. For researchers in educational settings, the length of time when one can be at the site is controlled by the academic calendar. In other words, there is nothing to observe when school is not in session. Further, research sites which involve the observation of minor children can require negotiation of permissions from multiple sources and levels of administration.

Examples of three types of ethnography commonly encountered in education are given here. In a collaborative autoethnography, [Chang et al. \(2014\)](#) explored how mentoring was experienced and influenced the leadership development of 14 U.S. higher education leaders of color working in faith-based institutions. [Smyth \(2016\)](#) produced a critical ethnography of student disengagement by Australian youth who leave formal education as an act of political resistance. Finally, [Lund \(2012\)](#) created an institutional ethnography of a Finnish university from the standpoint of female academics who are disadvantaged in trying to achieve the “ideal academic” status by publishing in top journals (p. 218). A. Skukauskaitė’s chapter on Ethnography in this volume offers further information pertaining to this method.

Discourse analysis

Discourse analysis, which in its most basic form is the study of language, is derived from the field of linguistics. Writing for the *Linguistic Society of America*, Tannen notes that discourse analysis may be defined as “the analysis of language ‘beyond the sentence’” (n.d., Discourse Analysis section, para. 1). American scholar Zellig Harris is often credited with coining the term *discourse analysis* ([Salkie, 2009](#); [Stanlaw, 2021](#)). He wrote that discourse analysis was valuable in order to study the exemplification of grammar and to focus on how a text sends its messages rather than what the messages within the text may be ([Harris, 1952](#)). Contemporary researchers acknowledge that there are countless variations of definitions for the term discourse analysis, differing in relation to the field of study, specific usage, and scholar’s stances on the method ([Tannen et al., 2015](#)). As a research approach discourse analysis has application across a broad array of fields of study, including communications, anthropology, sociology, and nursing. The 1970s and 1980s saw educators begin to employ discourse analysis in classrooms focusing, for example, on second language learning ([Cohen et al., 1979](#); [Larsen-Freeman, 1980](#); [Lezberg and Hilferty, 1978](#)). Starting in the 1990s, educators used discourse analysis in a critical way to explore the relationship between language and society ([Rogers et al., 2005](#)).

The two primary types of discourse analysis are descriptive and critical. Descriptive discourse analysis seeks to understand how language influences people and situations. Critical discourse analysis seeks to expose through description and interpretation the ways that social inequalities are constructed, maintained, and legitimized ([Mullet, 2018](#)). [Fairclough \(1985\)](#) holds that critical discourse requires an investigation of the ways that verbal interactions are determined by social structures as well as the effects of those verbal interactions on social structures. [Iverson \(2016\)](#) provides a descriptive discourse analysis of sexual assault policies in US higher education using a sample of 22 sexual violence policies from recipients of a 2012 U.S. Department of Justice’s Office of Violence Against Women campus grants. Much of the discourse analysis scholarship in education employs a critical discourse analysis approach. An example of critical discourse analysis can be found in [Thompson’s \(2002\)](#) study conducted with a single music teacher at a government school in Queensland, Australia which “revealed a version of world music based on an ideology of difference” (p. 19). Additionally, [Saarinen’s \(2008\)](#) article on the position of text and discourse analysis in higher education research, which focuses on the construct of quality in higher education policy, may be instructive to researchers seeking further direction on how to use critical discourse analysis. For a more in-depth exploration of discourse analysis, see L. Suciū’s Discourse Analysis chapter in this volume.

Though discourse analysis has become increasingly popular as a qualitative research approach, due in part to its potential for application across a wide array of fields, there are critics of the method. Researchers have voiced concern over its conceptual confusion and frequent conflation of text and discourse ([Schiffrin et al., 2001](#); [Widdowson, 1995](#)). While discourse analysis offers a powerful approach for investigating the use and effects of language, a lack of structure can pose a potential problem for researchers. The creation of new types of discourse (e.g., social media) and technology will no doubt continue to generate an interest from scholars in this “well-established but still burgeoning field” ([Tannen and Hamilton, 2015](#), p. xx) of discourse analysis.

Participatory action research

The term participatory action research (PAR) is considered an umbrella term because it encompasses a range of approaches to participatory action-oriented studies ([Kindon et al., 2007](#)). While participatory action research can involve quantitative data, our focus here is on PAR as a qualitative approach. Deriving from [Lewin’s \(1946\)](#) work on action research, participatory action research combines elements of community-based research, popular education, and action for social change ([Adelman, 1993](#); [Williams and Brydon-Miller, 2004](#)). Brydon-Miller and Maguire state that PAR is “an openly and unapologetically political approach to knowledge creation through and for action” (2009, p. 79). Essentially, PAR is a collaboration between researchers and participants whereby a problem or issue is investigated with the purpose of finding a solution or seeking improved conditions. Key to PAR is its central feature of democratizing research through a partnership between researchers and participants.

The term “participatory research,” the forerunner to PAR, was first used in the early 1970s in Tanzania to describe various community-based approaches to knowledge creation ([Hall, 2005](#)). Since that time PAR has become increasingly popular, particularly within the social sciences. Social activist, author, and educator [Freire \(2005\)](#) whose writings helped form much of PAR’s theoretical framework ([Cahill, 2007](#); [Jacobs, 2016](#)), including “dehumanization; the difference between subjects versus objects in research; co-intentional education, and the ‘banking’ concept of education” was also instrumental in bringing the research approach to educational settings ([Jacobs, 2018](#), p. 45). [Brydon-Miller and Maguire \(2009\)](#) state that PAR works under the concept that ordinary people can learn and transform their own situations and lives via education, action, and research. The authors maintain that

PAR works in education because it “openly challenges existing structures of power and creates opportunities for the development of innovative and effective solutions to the problems facing our schools and communities” (Brydon-Miller and Maguire, 2009, p. 81). PAR is a research strategy which rests on a philosophical foundation of mutual respect and cooperation between the researcher and participants. As such it has been identified as one methodological research approach that is appropriate for historically marginalized or disenfranchised groups such as Aboriginal peoples and communities (Bennett, 2004).

An attractive feature of PAR for researchers, particularly new researchers, is that there are no “hard and fast rules” for how the approach should be implemented, making it easily adaptable to a variety of circumstances and settings (Bennett, 2004, p. 23). When applied in school settings, PAR helps address power imbalances between teachers and students by acknowledging the contributions of students rather than maintaining the traditional view that all knowledge is held by teachers (Brydon-Miller and Macguire, 2009; Jacobs, 2016). Related to this concept of power imbalance is the need for educators and researchers involved in PAR to critically examine their role in the educational process as well as their role in the research study (Brydon-Miller and Macguire, 2009). Failure to thoroughly consider researcher role and positionality can be a potential weakness in the use of PAR. When working in non-school based community settings, researchers must be knowledgeable of the community and sensitive to the agendas of participants. The amount of time that can be required to conduct a PAR study must also be carefully considered and weighed against the realities of the researcher’s responsibilities and time constraints (MacDonald, 2012).

Ditrano and Silverstein (2006) offer an example of PAR in their study involving parents of children with emotional disabilities and a school psychologist in upstate New York. At the outset of the study, the nine parent participants had little understanding of how the school district’s special education system worked and were not connected to support networks. Through meetings over the course of 5 months in which participants got to dialog and think critically about the problems they wanted to address, they moved from passivity to advocacy and action in the school community. Halliday et al. (2019) found that involving students in PAR was promising in helping to understand student well-being. Their research included ten middle-school students, ages 9–11, from a publicly funded Australian school involved in a participatory action research project regarding positive education. Also, Bengele and Schuch (2018) offer insight into conducting dissertations using a PAR approach, including benefits that graduate students might gain from applying the method. E. Stringer’s Action and Community-based Participatory Approaches in Qualitative Research chapter in this volume provides an expanded discussion of PAR.

Conclusion

This chapter was written as an overview or as a starting point to introduce new researchers to seven of the qualitative approaches to research available and applicable to studies in education. Our goal was to synthesize the literature on the five established qualitative approaches of case study, phenomenology, grounded theory, narrative inquiry, and ethnography, as well as two additional qualitative approaches more commonly seen in educational research—discourse analysis and participatory action research—and to present a concise, organized summary of each. With these summaries we hope that the potential for research using one or more of the qualitative approaches begins to take shape in the reader’s imagination. We encourage readers who are inspired by one or more of these approaches and seeking further information to consult the other chapters in this volume, such as M. Hammersley’s The History of Qualitative Research in Education.

References

- Adelman, C., 1993. Kurt Lewin and the origins of action research. *Educ. Action Res.* 1 (1), 7–24. <https://www.tandfonline.com/doi/abs/10.1080/0965079930010102>.
- Bengele, T., Schuch, C., 2018. Integrating participatory action research into graduate geography studies: a tale of two dissertations. *J. Geogr. High. Educ.* 42 (4), 617–629. <https://doi.org/10.1080/03098265.2018.1514589>.
- Bennett, M., 2004. A review of the literature on the benefits and drawbacks of participatory action research. *First Peoples Child Fam. Rev.* 14 (1), 109–122. <https://doi.org/10.7202/1069582ar>.
- Bleakley, A., 2005. Stories as data, data as stories: making sense of narrative inquiry in clinical education. *Med. Educ.* 39 (5), 534–540. <https://doi.org/10.1111/j.1365-2929.2005.02126.x>.
- Braun, V., Clarke, V., 2013. *Successful Qualitative Research: A Practical Guide for Beginners*. Sage, Los Angeles, CA.
- Brydon-Miller, M., Maguire, P., 2009. Participatory action research: contributions to the development of practitioner inquiry in education. *Educ. Action Res.* 17 (1), 79–93. <https://doi.org/10.1080/09650790802667469>.
- Cahill, C., 2007. Participatory data analysis. In: Kindon, S., Pain, R., Kesby, M. (Eds.), *Participatory Action Research Approaches and Methods: Connecting People, Participation and Place*. Routledge, London, pp. 181–187.
- Caine, V., Estefan, A., Clandinin, D.J., 2013. A return to methodological commitment: reflections on narrative inquiry. *Scand. J. Educ. Res.* 57 (6), 574–586. <https://doi.org/10.1080/00313831.2013.798833>.
- Castagno, A.E., 2012. What makes critical ethnography “critical.” In: Lapan, S.D., Quartaroli, M.T., Riemer, F.J. (Eds.), *Qualitative Research: An Introduction to Methods and Designs*. Jossey-Bass, pp. 373–390. http://www.ru.ac.bd/wp-content/uploads/sites/25/2019/03/407_02_00_Lapan-Qualitative-Research_-An-Introduction-to-Methods-and-Designs-2011.pdf#page=405.
- Chametzky, B., 2015. Surviving situational suffering: a classic grounded theory study of post-secondary part-time educators in the United States. *Grounded Theor. Rev.* 14 (1), 26–40. <http://groundedtheoryreview.com/wp-content/uploads/2015/06/Surviving-Situational-suffering-06192015.pdf>.
- Chang, H., Longman, K.A., Franco, M.A., 2014. Leadership development through mentoring in higher education: a collaborative autoethnography of leaders of color. *Mentor. Tutoring* 22 (4), 373–389. <https://doi.org/10.1080/13611267.2014.945734>.
- Charmaz, K., 2005. Applications for advancing social justice studies. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Sage Handbook of Qualitative Research*. Sage Publications, Thousand Oaks, CA, pp. 507–535.

- Charmaz, K., 2006. *Constructing Grounded Theory*. Sage, London.
- Charmaz, K., 2014. *Constructing Grounded Theory*, second ed. Sage, Los Angeles, CA.
- Clandinin, D.J., Connelly, F.M., 2000. Narrative Inquiry: Experience and Story in Qualitative Research. Jossey-Bass.
- Clandinin, D.J., 2006. Narrative inquiry: a methodology for studying lived experience. *Res. Stud. Music Educ.* 27 (1), 44–54. <https://doi.org/10.1177/1321103X060270010301>.
- Cohen, A., Glasman, H., Rosenbaum-Cohen, P.R., Ferrara, J., Fine, J., 1979. Reading English for specialized purposes: discourse analysis and the use of student informants. *TESOL Q.* 13 (4), 551–564. <https://www.jstor.org/stable/3586449>.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14. <https://doi.org/10.3102/0013189X019005002>.
- Corbin, J.M., Strauss, A.L., 2015. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, fourth ed. Sage, Los Angeles, CA.
- Cordova, A.J., Knecht, L.M., 2019. Liminal knowledge: positioning intersectionality in academia. *Cult. Stud. Crit. Methodol.* 19 (3), 203–213. <https://doi.org/10.1177/1532708618819635>.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among the Five Approaches*, fourth ed. Sage, Los Angeles, CA.
- Creswell, J.W., 1998. *Qualitative Inquiry and Research Design: Choosing Among the Five Traditions*. Sage Publications, Thousand Oaks, CA.
- Dall'Alba, A.G., 2009. Phenomenology and education: an introduction. *Educ. Philos. Theor.* 41 (1), 7–9. <https://doi.org/10.1111/j.1469-5812.2008.00479.x>.
- De Gagne, J.C., Walters, K.J., 2010. The lived experience of online educators: hermeneutic phenomenology. *J. Online Learn. Teach.* 6 (2), 357–366. https://jolt.merlot.org/vol6no2/degagne_0610.pdf.
- Delaney, C., Raftery, D., 2020. Becoming “everything rolled into one”: how sister-teachers experienced the immediate impact of the free education scheme, c. 1958–1968, an oral history approach. *J. Educ. Adm. Hist.* <https://doi.org/10.1080/00220620.2020.1854204>. Advance online publication.
- Delmas, P.M., Ivankova, N.V., 2018. Sisters in the sacred grove: Catholic women religious as faculty members at public universities. *Qual. Rep.* 23 (2), 350–368. <https://nsuworks.nova.edu/tqr/vol23/iss2/5/>.
- Denzin, N.K., Lincoln, Y.S., 2005. *The Sage Handbook of Qualitative Research*, fourth ed. Sage, Los Angeles.
- DeVault, M.L., 2006. Introduction: what is institutional ethnography? *Soc. Probl.* 53 (3), 294–298. <https://www.jstor.org/stable/10.1525/sp.2006.53.3.294>.
- Ditrano, C.J., Silverstein, L.B., 2006. Listening to parents’ voices: participatory action research in the schools. *Prof. Psychol. Res. Pract.* 37 (4), 359. <https://doi.org/10.1037/0735-7028.37.4.359>.
- Fairclough, N.L., 1985. Critical and descriptive goals in discourse analysis. *J. Pragmat.* 9 (6), 739–763. [https://doi.org/10.1016/0378-2166\(85\)90002-5](https://doi.org/10.1016/0378-2166(85)90002-5).
- Floyd, A., 2012. Narrative and life history. In: Briggs, A.R.J., Coleman, M., Morrison, M. (Eds.), *Research Methods in Educational Leadership & Management*. Sage, Los Angeles, CA, pp. 223–235.
- Freire, P., 2005. *Pedagogy of the Oppressed*. Continuum, New York.
- Gallagher, S., 2012. What is phenomenology? *Phenomenology*. Palgrave Macmillan, pp. 7–18.
- Glaser, B.G., Strauss, A.L., 1967. *The Discovery of Grounded Theory*. Aldine Publishing Company, Hawthorne, NY.
- Grace, A., 2006. Writing the queer self: using autobiography to mediate inclusive teacher education in Canada. *Teach. Teach. Educ.* 22. <https://doi.org/10.1016/j.tate.2006.04.026>, 826–824.
- Hall, B.L., 2005. In from the cold? Reflections on participatory research from 1970–2005. *Convergence* 38 (1), 5–24.
- Halliday, A.J., Kern, M.L., Garrett, D.K., Turnbull, D.A., 2019. The student voice in well-being: a case study of participatory action research in positive education. *Educ. Action Res.* 27 (2), 173–196. <https://doi.org/10.1080/09650792.2018.1436079>.
- Hammersley, M., 2018. What is ethnography? Can it survive? Should it? *Ethnogr. Educ.* 13 (1), 1–17. <https://doi.org/10.1080/17457823.2017.1298458>.
- Harris, Z.S., 1952. Discourse analysis. *Language* 28 (1), 1. <https://doi.org/10.2307/409987>.
- Heidegger, M., 1972. *On Time and Being* (J. Stambaugh, Trans.). Harper & Row, New York.
- Hodkinson, P., Hodkinson, H., 2001. The Strengths and Limitations of Case Study Research. http://education.exeter.ac.uk/tlc/docs/publications/LE_PH_PUB_05.12.01.rtf.
- Hoppey, D., McLeskey, J., 2013. A case study of principal leadership in an effective inclusive school. *J. Spec. Educ.* 46 (4), 245–256. <https://doi.org/10.1177/0022466910390507>.
- Hussein, M.E., Hirst, S., Salyers, V., Osuji, J., 2014. Using grounded theory as a method of inquiry: advantages and disadvantages. *Qual. Rep.* 19 (27), 1–15. <http://www.nova.edu/ssss/QR/QR19/el-hussein13.pdf>.
- Husserl, E., 1931. *Ideas* (W.R. Boyce Gibson, Trans.). George Allen & Unwin, London.
- Iverson, S.V., 2016. A policy discourse analysis of sexual assault policies in higher education. In: Wooten, S.C., Mitchell, R.W. (Eds.), *The Crisis of Campus Sexual Violence: Critical Perspectives on Prevention and Response*. Routledge, New York, pp. 15–32.
- Jacobs, S.D., 2016. The use of participatory action research within education—Benefits to stakeholders. *World J. Educ.* 6 (3), 48–55. <https://doi.org/10.5430/wje.v6n3p48>.
- Jacobs, S.D., 2018. A history and analysis of the evolution of action and participatory action research. *Can. J. Action Res.* 19 (3), 34–52. <https://doi.org/10.33524/cjar.v19i3.412>.
- Jeffrey, B., Troman, G., 2004. Time for ethnography. *Br. Educ. Res. J.* 30 (4), 535–548. <https://doi.org/10.1080/0141192042000237220>.
- Karpouza, E., Emvalotis, A., 2019. Exploring the teacher-student relationship in graduate education: a constructivist grounded theory. *Teach. High. Educ.* 24 (2), 121–140. <https://doi.org/10.1080/13562517.2018.1468319>.
- Kenney, M., Fourie, R., 2014. Tracing the history of grounded theory methodology: from formation to fragmentation. *Qual. Rep.* 19 (52), 1–9. <https://nsuworks.nova.edu/tqr/vol19/iss52/1>.
- Kindon, S., Pain, R., Kesby, M. (Eds.), 2007. *Participatory Action Research Approaches and Methods: Connecting People, Participation and Place*. Routledge, London.
- Larsen-Freeman, D., 1980. *Discourse Analysis in Second Language Research*. Newbury House Publishers, Inc, Rowley, MA.
- Laverty, S.M., 2003. Hermeneutic phenomenology and phenomenology: a comparison of historical and methodological considerations. *Int. J. Qual. Methods* 2 (3), 21–35. <https://journals.sagepub.com/doi/pdf/10.1177/160940690300200303>.
- Lewin, K., 1946. Action research and minority problems. In: Lewin, G.W. (Ed.), *Resolving Social Conflicts*. New York.
- Lezberg, A., Hilferty, A., 1978. Discourse analysis in the reading class. *TESOL Q.* 12 (1), 47–55. <https://www.jstor.org/stable/3585790>.
- Lund, R., 2012. Publishing to become an “ideal academic”: an institutional ethnography and a feminist critique. *Scand. J. Manag.* 28 (3), 218–228. <https://doi.org/10.1016/j.scaman.2012.05.003>.
- MacDonald, C., 2012. Understanding participatory action research: a qualitative research methodology option. *Can. J. Action Res.* 13 (2), 34–50. <https://doi.org/10.33524/cjar.v13i2.37>.
- Merriam, S.B., 2001. *Qualitative Research and Case Study Applications in Education*. Jossey-Bass, San Francisco, CA.
- Moerer-Urdahl, T., Creswell, J., 2004. Using transcendental phenomenology to explore the “ripple effect” in a leadership mentoring program. *Int. J. Qual. Methods* 3 (2), 1–28. <https://doi.org/10.1177/160940690400300202> (27 pages).
- Morse, J.M., Richards, L., 2002. *Readme First for a User’s Guide to Qualitative Methods*. Sage Publications, Thousand Oaks, CA.
- Morse, J.M., 2009. Tussles, tensions, and resolutions. In: Morse, J.M., Stern, P.N., Corbin, J., Bowers, B., Charmaz, K., Clarke, A.E. (Eds.), *Developing Grounded Theory: The Second Generation*. Left Coast Press, Inc, Walnut Creek, CA, pp. 13–19.
- Moustakas, C., 1994. *Phenomenological Research Methods*. Sage Publications, Thousand Oaks, CA.
- Mullet, D.R., 2018. A general critical discourse analysis framework for educational research. *J. Adv. Acad.* 29 (2), 116–142. <https://doi.org/10.1177/1932202X18758260>.
- Padilla-Díaz, M., 2015. Phenomenology in educational qualitative research: philosophy as science or philosophical science. *Int. J. Educ. Excell.* 1 (2), 101–110. <https://doi.org/10.18562/IJEE.2015.0009>.

- Rogers, R., Malancharuvil-Berkes, E., Mosley, M., Hui, D., O'Garro Joseph, G., 2005. Critical discourse analysis in education: a review of the literature. *Rev. Educ. Res.* 75 (3), 365–416. <https://www.jstor.org/stable/3515986>.
- Saarinén, T., 2008. Position of text and discourse analysis in higher education policy research. *Stud. High. Educ.* 33 (6), 719–728. <https://doi.org/10.1080/03075070802457090>.
- Sairanen, H., Kumpulainen, K., 2014. A visual narrative inquiry into children's sense of agency in preschool and first grade. *Int. J. Educ. Psychol.* 3 (2), 141–174. <https://doi.org/10.4471/ijep.2014.09>.
- Salkie, R., 2009. Harris. In: Cobley, P. (Ed.), *The Routledge Companion to Semiotics* [Electronic Resource], 207. Routledge, New York.
- Savin-Baden, M., Niekerk, L.V., 2007. Narrative inquiry: theory and practice. *J. Geogr. High. Educ.* 31 (3), 459–472. <https://doi.org/10.1080/03098260601071324>.
- Schiffrin, D., Tannen, D., Hamilton, H.E. (Eds.), 2001. *The Handbook of Discourse Analysis*. Blackwell Publishers, Malden, MA.
- Schwandt, T.A., 2015. *The Sage Dictionary of Qualitative Inquiry*. Sage, Los Angeles, CA.
- Skott-Myhre, K., Weima, K., Gibbs, H., 2012. *Writing the Family: Women, Auto-ethnography, and Family Work*, vol. 80. Sense Publishers, Rotterdam, The Netherlands.
- Slife, B.D., Williams, R.N., 1995. *What's Behind the Research? Discovering Hidden Assumptions in the Behavioral Sciences*. Sage Publications, Thousand Oaks, CA.
- Smyth, J., 2016. Puncturing notions of precarity through critical educational research on young lives in Australia: towards a critical ethnography of youth. *Ethnogr. Educ.* 11 (2), 129–141. <https://doi.org/10.1080/17457823.2015.1040429>.
- Sokolowski, R., 2000. *Introduction to Phenomenology*. Cambridge University Press, New York, NY.
- Stake, R.E., 1995. *The Art of Case Study Research*. Sage, Thousand Oaks, CA.
- Stanlaw, J., 2021. In: Harris, Z. (Ed.), *The International Encyclopedia of Linguistic Anthropology*. John Wiley & Sons, Hoboken, NJ, pp. 1–3.
- Strauss, A., Corbin, J., 1990. *Basics of Qualitative Research*. Sage, Newbury Park, CA.
- Stroobants, V., 2005. Stories about learning in narrative biographical research. *Int. J. Qual. Stud. Educ.* 18 (1), 47–61. <https://doi.org/10.1080/09518390412331318441>.
- Tan, J., 2010. Grounded theory in practice: issues and discussion for new qualitative researchers. *J. Doc.* 66 (1), 93–112. <https://doi.org/10.1108/00220411011016380>.
- Tannen, D., Hamilton, H.E., 2015. Preface to the second edition. In: Tannen, D., Hamilton, H.E., Schiffrin, D. (Eds.), *The Handbook of Discourse Analysis* [Electronic Resource], second ed. John Wiley and Sons, Chichester, West Sussex, UK. Incorporated.
- Tannen, D., Hamilton, H.E., Schiffrin, D. (Eds.), 2015. *The Handbook of Discourse Analysis* [Electronic Resource], second ed. John Wiley and Sons, Chichester, West Sussex, UK. Incorporated.
- Tannen, D., n.d. Discourse analysis—what speakers do in conversation. *Ling. Soc. Am.* <https://www.linguisticsociety.org/resource/discourse-analysis-what-speakers-do-conversation>.
- Tesch, R., 1990. *Qualitative Research: Analysis Types and Tools*. The Falmer Press, New York.
- Thompson, K., 2002. A critical discourse analysis of world music as the “other” in education. *Res. Stud. Music Educ.* 19 (1), 14–21. <https://doi.org/10.1177/1321103X020190010301>.
- Tsui, A.B., 2007. Complexities of identity formation: a narrative inquiry of an EFL teacher. *TESOL Q.* 41 (4), 657–680. <https://doi.org/10.1002/j.1545-7249.2007.tb00098.x>.
- Vassilakaki, E., Johnson, F., 2015. The use of grounded theory in identifying the user experience during search. *Libr. Inf. Sci. Res.* 37 (1), 77–87. <https://doi.org/10.1016/j.lisr.2014.06.006>.
- Widdowson, H.G., 1995. Discourse analysis: a critical view. *Lang. Lit.* 4 (3), 157–172. <https://doi.org/10.1177/096394709500400301>.
- Williams, B.T., Brydon-Miller, M., 2004. Changing directions: participatory action research, agency, and representation. In: Brown, S.G., Dobrin, S.I. (Eds.), *Ethnography Unbound: From Theory Shock to Critical Praxis*. State University of New York Press, pp. 241–258. muse.jhu.edu/book/4799.
- Wolcott, H.F., 2008. *Ethnography: A Way of Seeing*. AltaMira Press, Lanham, MD.
- Yazan, B., 2015. Three approaches to case study methods in education: Yin, Merriam, and Stake. *Qual. Rep.* 20 (2), 134–152. <http://www.nova.edu/ssss/QR/QR20/2/yazan1.pdf>.
- Yin, R.K., 2003. *Case Study Research: Design and Methods*, third ed., vol. 5. Sage, Los Angeles, CA.
- Yin, R.K., 2014. *Case Study Research: Design and Methods*, fifth ed. Sage, Los Angeles, CA.

Ethics and politics in qualitative research

Anne Ryen, University of Agder, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	34
A brief history of research ethics	35
From illegitimate medical research to formal research ethics	35
From medicine to the social sciences	35
Anchors: from philosophy to paradigms, worldviews, and pathways	36
From abstract principles to research practice	36
Philosophical wealth: ontologies and epistemologies from across the world	36
Brief reminders on traditions and practices	37
First reminder: histories and traditions in qualitative research	37
Second reminder: what to do and what you do	38
Politics and ethics entangled	38
Positivism: external reality and pre-packed ethical issues	38
Ethnomethodology: doing being ethical as collaborative accomplishment	39
Constructivism: rejecting old scripts	40
Feminist and indigenous research: community grounded research ethics	41
Conclusions	42
References	42

Glossary

Qualitative research There is no general agreement about the nature of qualitative research (Silverman, 2021, p. 3). According to the definition which privileges “experience”, qualitative research is concerned with “subjective meaning” as in interpretive research and as opposed to “facts” in quantitative research. In an alternative definition, qualitative research focuses on behavior in naturalistic settings as in ethnomethodology and conversational analysis. In line with this, Silverman argues qualitative research is “theoretically driven, complements quantitative research by [exploring] how social phenomena are constituted in real time, is as much about social practices as about experience, and should be a credible, rigorous enterprise (2021, pp. 4–5). Contrary to this, Denzin and Giardina argue that qualitative inquiry is about changing the world, about advocacy ... that contributes to a public conversation ... a dialogue that puts the very notions of democracy and freedom, citizen and patriot, into play ... the performative and the political intersect on the terrain of a praxis-based ethic” (2012, p. 19). These definitions have important consequences for the practical conduct of qualitative research practice and ethics

Ethics “Ethics” can mean a field of study as well as “the set of ethical principles that should be taken into account when doing social research” (Hammersley and Traianou, 2012, p. 17). “Ethical” can refer to “The realm of considerations relevant to determining what is good or bad, right or wrong ... Here the contrast is with *non-ethical* matters, ...” and “What is good or right, as contrasted with the *unethical*—what is bad or wrong” (Hammersley and Traianou, 2012, p. 17). Ethical concerns in social research commonly refer to certain recognized distinct ethical principles “along with ideas about how these relate to the practice of research” (Hammersley and Traianou, 2012, p. 17). These principles include minimizing harm, respecting participants’ autonomy, and their privacy, offering reciprocity and treating people equitably (Hammersley and Traianou, 2012, points 1–5). The classic informed consent, confidentiality, and anonymity are part of these (Hammersley and Traianou, 2012, chapters 4 and 5)

Paradigms Guba and Lincoln define paradigms as “Basic belief systems based on ontological, epistemological, and methodological assumptions ... The beliefs are basic in the sense that they must be accepted simply on faith ... there is no way to establish their truthfulness” (1996, p. 105). They argue paradigms are not open to proofs in any conventional sense (1996, p. 108), and point to the consequences for practical research, interpretation of findings, policy choices (1996, p. 112) and ethical issues. To illustrate, in positivism ethics is externally audited by codes and committees, in critical theory it is (also) intrinsic by recognizing values and historical situatedness. It is also intrinsic in constructivism by additionally including participant values, but the personal interaction may bring other problems

Worldviews Guba and Lincoln claims that “A paradigm Represents a worldview that defines, for its holder, the nature of the “world”, the individual’s place in it, and the range of possible relationships to that world and its parts ...” (1996, p. 107). As an example, indigenous research illustrates a relationship with the world different from in classic western research, and “deploy interpretive strategies and skills fitted to the needs, language, and traditions of their respective indigenous community ...” (Denzin et al., 2008, p. 11) based on “a commitment to an indigenist outlook” which makes indigenous pedagogies “grounded in an oppositional consciousness” (both 2008, p. 10)

Method talk In the opening of their book, Gubrium and Holstein argue that “the languages of the social sciences shapes how researchers view the world ...” and describe method talk as “the working vocabulary of research procedure” (1997, p. 3). They illustrate by “One familiar idiom directs researchers to a real world of facts and figures. A second urges them to focus on how the real world is conceptualized or formed. Other research vocabularies stake out different empirical terrain. If we accept the tenet that knowledge is humanly produced, it’s not hard to conceive of sites of method talk as knowledge factories—places where the work of knowledge construction takes place. Extending the metaphor, colleges and universities stand out as production sites ...” (1997, p. 3). This illustrates that also paradigms and worldviews shape the way researchers talk about research which make their proponents use alternative vocabularies as in alternative approaches to educational or pedagogic research

According to Gubrium and Holstein’s “idioms of reality”, “It considers the way the language of qualitative method relates to how researchers view and describe social life.” They refer to Geertz’ “The strange idea that reality has an idiom ... leads to an even stranger idea that, if literalism is lost, so is fact” (1997, p. 3). They also refer to Wittgenstein’s “language games” ... in the sense that the manner in which one proceeds to do social research organizes the empirical contours of what is under investigation” (1997, p. 5). Due to the diversity in qualitative research ... each variant is conducted in, a language or idiom of its own. Accordingly, each idiom, presents a distinct reality, virtually constituting its empirical horizons” (1997, p. 5).

Introduction

Research and the research process are shaped struggles over legitimacy and violations. This comes with consequences to research ethics, and with the many positions within qualitative research, as there is no fixed research ethics template (Ryen, 2021). It is by far too complex. Hammersley and Traianou remind us that “there are no single agreed sense given to the term” (2021a, p. 16). Qualitative researchers typically meet up with and work closely with people and write about their lives from settings and contexts that may rhyme, though never fully repeat themselves. This has consequences on how to think about research ethics, particularly when going from theory to the less predictable research practice. When qualitative researchers argue that reality is social, this makes truth a much more complex phenomenon compared to the classic sciences (Denzin and Lincoln, 2005, Guba and Lincoln, 1994 on paradigms, Gubrium and Holstein, 1997 on worldviews and how to talk about qualitative research). This is amplified by the many parallel and conflicting positions in philosophy and the theory of science, which invite alternative positions on issues, such as what research is all about, images of the involved and researchers’ ethical obligations in their relationships with their study participants, including that of writing up their research. This has made the development of qualitative research and formal research ethics far from a linear, unproblematic process within static contexts, but one of critical dialogue and controversies that now range from theoretical concerns to activism and politization, as in class, gender and race studies (Denzin and Gardina, 2009 on qualitative inquiry as hope, Denzin and Gardina, 2012 as advocacy, Haggerty, 2004 on the problems with institutional research ethics government, Smith, 2005 on indigenous research and research ethics, Skeggs, 2010 on feminism and social transformation, Silverman, 2017a,b on research ethics in African contexts). This has made qualitative research a place for reflections on the meeting between ideals and research practice all throughout the research process, as well as social science research permeated by politics at all levels (Christians, 2005 on research and politics, House, 2005 on qualitative evaluation). It has also influenced all research activities, including on how to think about them in ethical research terms, which has made research ethics itself into an arena under constant scrutiny (Fakhry, 1998a on ethics in Islamic philosophy, the many chapters on research ethics issues in Mertens and Ginsberg, 2009). The emergence of new -isms, post-, method talks and pathways in qualitative research raises new takes on old issues by posing new questions to former ways of thinking about research (Carlson et al., 2021, Kuntz and St. Pierre, 2021a,b) and ethics (Tesar, 2021). In a nutshell, this illustrates how ethics is firmly entangled in the wider politics of qualitative research. In the next five sections, this chapter will raise and illustrate such issues, and invite the reader to further reflections relevant to her/his own research.

We begin with providing a brief history of research ethics in the West to help illustrate how the standard code of research ethics came into being and into the social sciences, its justification and the consequences for research. This takes us to the next section on the theory of science and worldviews, in which research itself and ethical ways of thinking are anchored. Research ethics is an exemplary illustration of the impact of politics with the constant friction between institutional requirements and alternative approaches in qualitative research, thus rejecting quick-fix recipes that external committees can follow and tick off, as in much positivist anchored thinking. Then, in the next section, follows two reminders of variations across global regions and disciplines. The third section will in more concrete ways illustrate the implications of philosophy all the way through to research ethics in some paradigms and method talks. This demonstrates in a more systematic way that to think differently about reality comes with alternative ethical research codes anchored in alternative, contrasting worldviews. Denzin and Lincoln (1994, p. 5) reminded us many years ago that the differences we see between them in addressing the same issues return to the politics of research, and to the question of who has the power to claim the right answers to these differences, including ethical research issues. Yet decades later, and as always,

ontologies have taken qualitative research in new directions. We argue that the question of knowledge production makes research ethics a complex and controversial issue. It also means that to learn about research ethics in research practice invites students to become active participants in inquiring into knowledge and knowing, in addition to the impact to their own studies. It also invites the major question: What is research and research ethics all about? In the brief conclusion we reflect on the many turns and world-views in qualitative research. Each new turn brings something new and each new disrupts politics and ethics.

A brief history of research ethics

The last decades came with more external and extensive formalities in research ethics as a response to the abuse and misdemeanors, especially in the name of medical research, though often in collaboration with more powerful co-actors. This provoked reactions that demanded preventive measures with regard to certain questions. This section raises some of these, including: Why all this focus on research ethics? How did it enter the social sciences? Why so much focus on research participants?

From illegitimate medical research to formal research ethics

The introduction of more formal ethical research codes dates to the post-war (WW II) efforts to put an end to illegitimate experiments on human beings in concentration camps during the Nazi regime, under the heading of medical research (Lefor, 2005). This is partially correct, although the roots go further back to when German physicians, who a couple of decades earlier were accused of unethical medical practices in their work with the Nazi ideology of racial hygiene, to keep the Aryan race “pure” through the use of biology (Ruyter, 2003). All criticism of their experiments was ignored. In general, prisoners of war were considered to be an easy target for experiments, but their many experimentation studies also included children, Jews, homosexuals, the disabled, Romani and other ethnic folk groups, etc. The issue of consent was regarded as irrelevant, as prisoners were forced to participate and were often not aware of their participation (such as with the radiation experiments), and with no option to withdraw (Caplan, 1992).

After the war, this led to The Nuremberg Code as part of the court settlement in Nürnberg 1947. The Nuremberg Code or Codex (Nuremberg Code, 1996) consists of 10 rules to prevent such abuses from ever happening again. The first code makes voluntary consent essential, the ninth that the involved should be free to end participation at any time, and the 10th obliging the researcher to end the project in case of harm of any kind to the participant. The focus on protecting participants was a direct response to the historical context, but according to Katz (1996, p. 1662) not strictly confined to this context, as the US judges had more universal relevance in mind.

Despite this, more studies also unintentionally came to put research ethics on the agenda. This included the Manhattan project, which manufactured nuclear weapons. As such, the project was seen as a success, although with devastating effects to Hiroshima and Nagasaki and much more. This makes us ask: When is a research project a success? Does that include how the knowledge is being applied? The above-mentioned project involved researchers, universities, private companies, state leaders, the military, several countries, secrecy and voluntary censorship. This also invites a variety of critical reflections, such as whose past is it (Chakrabarty, 1992), on the rhetoric of historical representation (Rigney, 1990), and most importantly, how does this make us think about research ethics?

Another classic is the “Tuskegee Syphilis Study” 1932–72, with 600 impoverished African American sharecroppers (Reverby, 2009) from Alabama. This was a United States Public Health Service study in collaboration with Tuskegee University. The study was a grave violation of more ethical research issues, and with severe consequences to the involved, their wives and children. This later paved the way for informed consent and US institutional review boards (IRBs) as their ethical research committees. A similar study took place in Guatemala, with several other medical or health research projects taking place across global regions before and after the same war, which involved vulnerable or easy targeted populations, such as prisoners, patients and soldiers. Formal reactions followed.

From medicine to the social sciences

These experiences “paved the way for internal regulations in biomedical research with the Genève Declaration (WMA, 2021a) and the Helsinki Declaration (WMA, 2021b), both passed by the World Medical Association (WMA) in 1947 and 1964, respectively, and both with subsequent versions. Still, further medical misdemeanors followed, which caused the WMA to also recommend formally organized independent ethical research committees (1975) (Ryen, 2011, pp. 417–418). Punch (1994, pp. 88–94) offers a relevant overview over some of the much criticized studies in US’ 1950s–70s social science, including covert research in religious congregations, deception by membership in a sect, a CIA co-funded project where they secretly gave LSD to visitors to a brothel before filming the results by hidden camera, the use of secret microphones to record jury deliberations, Milgram’s classic psychological experiments on authority, in which participants caused “pain” to others (who knew it was fake) in a laboratory setting. Participants were not informed of this. Other countries have their own examples.

The criticism by contemporary researchers also shows that despite international declarations misdemeanors still occurred, that external organizations were also involved, and that studies did cause reactions from both within the research communities and

beyond. There are parallel histories in other countries that paved the way for their own formalized versions of the research ethics regime. Despite common foci, they display variations. There is no overview; therefore, each of us will have to explore the history of our regime. This may better prepare us for potential encounters related to issues of written consent, trust and the assumption of anonymity as the preferred option, and more. Taken together, the above history has nurtured responses (national laws and international conventions), invited research ethics literature (books and special journal issues, separate textbook chapters) and made it mandatory in applications for much research funding. It also has also entered methods courses, insofar as thinking one can strengthen research ethics through education, more awareness, and an improved understanding of such issues. Research ethics is not about laws, although the authority has come to rest with professionalism, and by an externally imposed pressure, until new paradigms in qualitative research emerged that saw ethical care as an internally imposed obligation based on more abstract principles, which commits the morally responsible qualitative researcher. However, there has been an increase in the numbers and issues covered in guidelines, as well as a more intense surveillance. As often described, the net was cast wide, i.e., ethics creep (Haggerty, 2004; Katz, 2007), with new actors entering the stage.

Anchors: from philosophy to paradigms, worldviews, and pathways

All research is anchored in abstract principles that shape how we see the world and explore it; just as qualitative and quantitative research are anchored in different principles, the same goes for the many different approaches within qualitative research. In this section, we will work our way through from the abstract to the concrete and offer a few illustrations to show there is some orderliness in the apparently bewildering chaos in a field where arguments are constantly sharpened, and new directions emerge. The important thing is to understand the longer lines, and to ask questions about our own position and how this informs the politics of qualitative research as we go from philosophy to our own research—research ethics included.

From abstract principles to research practice

These principles or beliefs refer to more abstract ideas about ontology and epistemology, with both informing methodology. Ontology is the philosophical study of the essential nature of reality (Hoffman and Kumar, 2020), or of being (Simons, 2015). This is the most abstract principle and deals with the nature of reality, and belongs to metaphysics, that part of philosophy which examines the fundamental nature of reality. This informs epistemology or the theory of knowledge, which is the philosophical study of the nature, origin and limits of human knowledge (Jupp, 2006), and comes from the Greek terms *epistēmē* (“knowledge”) and *logos* (“reason”) (Stroll and Martinich, 2021), which is concerned with the questions, what is knowledge, how do we acquire knowledge, or how to know the world. In combination, ontology and epistemology constitute the framework for methodology, which is the study of how to do research or the principles that guide our research practices or simply put: How is the researcher going to find out whatever s/he wants to find out, or know? In brief, ontology and epistemology confine the answer to methodology. A belief in a “real” reality demands we leave values behind, as in quantitative research with its various techniques, whereas answers to an ontology that assumes a reality as interactionally accomplished leads to a qualitative methodology with quite different methods. Research methods represent the most concrete level and refer to the specific tools and procedures we use to collect and analyze data. Our position in this wider landscape informs the research principles and procedures we use in practice just as in feminist research on women’s experiences which politicized research. This calls us to reflect on the question: What is my position in this landscape?

Firstly, taken together, the answers to the researcher’s ontological, epistemological and methodological basis make up nets described as paradigms or interpretive frameworks that guide research (Denzin and Lincoln, 1994, p. 13) or worldviews, which specify their own directions or positions in qualitative research that guide research practice. Others refer to method talks (Gubrium and Holstein, 1997, p. 100). They all bring their own language and concepts, which work as rather distinct approaches to inquire into social phenomena and make them recognizable to us as realities of some kind out there.

This is the topic of the next section. Importantly, these nets or various positions also inform us as how to think about research ethics and teach us how to be morally committed researchers. As we shall see, this makes research ethics varied, complex and intricate, a point often ignored in institutional ethical research arrangements.

Globalization made the world smaller, but also more diverse.

Philosophical wealth: ontologies and epistemologies from across the world

All societies come with a variety in formalized philosophical traditions and systematic thinking, whether referred to as “philosophy” or not (Ames, 1998). Global regions show long oral and written traditions, which in their own ways inform the politics of qualitative research, and how to work ethically in their impact fields. Philosophy has been sorted by its position in historical time as in Ancient- and Modern philosophy, but there has also been a tendency to appeal to regional categories such as Western, non-Western, Eastern or Oriental philosophy, or even the more specific Greek, Chinese and other philosophies. Such categorization may serve its functions, yet also comes with a tilt toward undue simplification. History shows that regimes, philosophers, and ideas have traveled along with the history of conquests and traveling routes, where philosophies were inspired by theories from across regions. Philosophical works were also translated, such as Aristotle’s works into Arabic, which caused Greek philosophy to have an impact on Islamic philosophy (Fakhry, 1998a,b). Further, Copernicus traveled to Italy and encountered Islamic science before

he developed his revolutionary idea of no longer seeing the Earth, but rather the Sun as the center of the cosmos. Moreover, the Prague astronomical clock, or the Prague Orloj, was inspired by Islamic science. In social science the history is more complex, and as always also due to internal controversies also outside European philosophy (as reflected in [Sardar, 2006](#)).

Despite this richness, and despite the many diasporas, courses and studies in Western higher education have been criticized for a tendency to focus on philosophers from their own region without being explicit as in the difference between “Philosophy” and “European Philosophy”, until philosophies from outside Europe eventually entered Western universities ([Van Norden, 2017](#); [Kim, 2013](#)) and inspired dialogues across regions ([Angle and Slote, 2013](#); [Taebnia, 2021](#)). This reminds us of the very long history of science and of philosophy beyond Europe ([Park, 2013](#)), as in the case of prominent Arab historian Ibn Khaldun (1332–1406) and his work on the philosophy of history ([Khaldun, 2012](#)), described as a landmark of social thought ([Issawi and Leaman, 1998](#)).

Even though this chapter draws on Western thinking, the essential message is that research ethics is philosophically rooted, that the multitude of theories reflect alternative and conflicting positions with their own assumptions pertaining to ethical research issues in qualitative research. The risk is the further universalization of research ethics, one of the classic objections to how it has institutionally been managed regardless of research participants, contextual matters and translational issues ([Wodak, 2011](#)). This includes the criticism of hierarchies and hegemonic knowledge, in this case with committees over qualitative researchers ([Hammersley, 2009](#)).

This reminds us that philosophical and cultural roots also inform our study participants in everyday matters, such as relations with strangers, respect and communication, as it plays in social meetings between participants. If unfamiliar, we as researchers need more time and experience to learn doing being moral inquirers ([Ryen, 2001](#) and [2007](#)). If politeness to guests overrides straightforward speaking, then answers to our simple questions will need interpretation to achieve a mutual understanding. It may be helpful to read about other researchers’ experiences in the relevant culture in order to help prepare us prior to going into the field, such as the many chapters in [Barnard and Wang \(2020\)](#), the citations from others’ studies (as in [Silverman, 2013](#) or [Johnson, 2020](#)), and when writing up your document ([Baker et al., 2016](#) in nursing).

Brief reminders on traditions and practices

Qualitative research means different things to different people, and [Denzin and Lincoln \(2005, p. xv\)](#) remind us that the discourse in qualitative interpretive research moves in different directions at the same time, which creates “new spaces, new possibilities, and new formations for qualitative research methods while closing down others”. This tells us that the social in social science makes it incongruent with dominant principles in classic science, and that the internal diversification within qualitative research is a continuous process. Worldviews address the world and the same issues in it, but in systematically diverse ways, always returning to the politics of research which accordingly frame research ethics.

As part of this, globalization has reminded us of the universal/local disputes, and that qualitative research is always situated. It takes place with people, somewhere and within certain contexts. This calls us to be aware of what these are, as well as the implications to our own studies.

First reminder: histories and traditions in qualitative research

In qualitative research, conflicts between distinct parties with incompatible arguments are the norm and not the exception, as we saw with regional, national and disciplinary traditions and dividers. New thinking occurs at the turning point between theories, not in the comfort zone. What feminism saw as constraints in positivism defined the fields and took them in different directions and still does. This is how the basic metaphysic belief systems lead to paradigms or worldviews ([Denzin and Lincoln, 1994, p. 107](#), [Guba and Lincoln, 1994, p. 107](#)). People within each talk about the world and about research in their own ways which has caused [Gubrium and Holstein \(1997\)](#) describe these as “method talks”. Each brings its own rhetoric particularities and idioms that reflect its philosophical position and constitutes social reality its own way.

The history of qualitative research has been categorized into successive phases, although authors tend to talk from their own site location as if these phases and time of origin are universal (as in [Denzin and Lincoln’s “moments” in 1994, pp. 7–11](#), and in [Vidich and Lyman’s “Qualitative Methods: Their History in Sociology and Anthropology” 1994, pp. 23–59](#)). Both offer great insight, but into aspects of the American history of qualitative research. This is not about persons, but instead what [Atkinson et al. \(2011, p. 467\)](#) refer to as “a pervasive problem in the international circuits of awareness” in their article on European traditions in biographical research. They point to the problem that researchers tend to only know their own national context. [Atkinson](#) previously ([2005, p. 2](#)) reminded us that:

While the global character of academic publishing has meant that English-language research communities have dominated much of the discourse, and American work has been especially prominent, in fact there are strong and distinctive national as well as disciplinary traditions.

Due to a lack of more publications, the implication of these traditions to research ethics is not yet clear, so therefore it is all the more important that we investigate into and reflect on the relevance to our own research. Textbooks may easily make us assume they are universally applicable as they are.

Second reminder: what to do and what you do

Official ethical research guidelines are general in nature, and for practical reasons. However, encounters are emergent and cannot be foreseen differently from what ethical research forms make us assume. This is about principles and practices. How we see the world, how we respond and how we represent it is also informed by cultural and other scripts. This also applies to research ethics, though it is often ignored in contemporary ethical research regimes.

A philosophy that differentiates between descriptive and normative ethics is most relevant here. The concepts are telling. [Kitchener and Kitchener \(2009, p. 5\)](#) put it simply:

Descriptive ethics studies how people actually behave and what ethical values they actually hold. This area belongs to empirical social science. Normative ethics is concerned with the questions: How should an individual behave; what properties are valuable and good? In short, whereas descriptive ethics studies the “is” side of the is-ought distinction, normative ethics studies the “ought” side.

This makes applied ethics deal with “principles ... from normative ethics to resolve specific moral issues in concrete and particular settings” and “social science research ethics an example of applied ethics” (2009, pp. 5–6). Applied ethics sorts itself under normative ethics because it deals with practical questions or principles used to solve concrete ethical issues ([Blackburn, 2016](#); [Crisp, 2011](#)). Worldviews and cultures diverge in what we see, and how we see it. What some see as good and acceptable, others may find immoral. It follows from this that we also diverge in how to resolve specific moral issues in concrete and particular settings. This is reflected in the contrasts between the paradigms and worldviews. This causes ethics and politics in qualitative research to become entangled, which accounts for the variation in what appears as ethical issues or not, and morally acceptable responses or not. [Elizabeth Murphy and Robert Dingwall \(2010, p. 339\)](#) remind us:

The ethics and the politics of ethnography are not clearly separable. Questions about the right way to treat each other as human beings, within a research relationship, are not wholly distinct from questions about the values which should prevail in a society, and the responsibility of the social scientist to make, or refrain, judgements about these ... Our assumptions about the nature of reality, the possible knowledge of that reality, the truth claims and so on, all have significant implications for our judgements about the ethnographer’s responsibilities The lack of consensus about methodology in and about ethnography, is reflected in discussions about its ethics.

This is not confined to ethnography, but ethnography appears as more complex compared to other methods, due to the character of it and the many questions that follow. It also illustrates the division between positivism and the many constructivist approaches. Positivist principles constitute what qualitative researchers initially reacted against, and nurtured alternatives. We will now illustrate the implications.

Politics and ethics entangled

Let us briefly recall the major idea in- and concepts associated with positivism, which has gained renewed relevance albeit now as politicians’ favorite. The reactions to the implications of positivist philosophical anchors, and the associated politics of research, initiated the responses and alternatives in constructivist research. It is this that makes “qualitative research mean[s] different things to different people,” which we will now illustrate.

Positivism: external reality and pre-packed ethical issues

Classic positivist ontology is inspired by the philosopher René Descartes on reason that developed the natural sciences. This rationality comes with an ontology that reality is “real,” externally out there. Epistemologically, this implies a dual thinking with methodological approaches that look for truth/not truth, or for the ways things are (and not). This is the background to the scientific revolution that followed, in which William [Bristow \(2017\)](#) reminds us that to Descartes, “all knowledge, in a relatively simple and elegant rationalist metaphysics provides a model of a rigorous and complete secular system of knowledge,” which led to “The Age of Reason” or the idea of Enlightenment in the 18th century. Reason came with attacks on dogmas in science, metaphysics and religion, and with a strong belief in the intellect to acquire a systematic knowledge of nature. Empiricism is another keyword for that era, along with a strong optimism. This simply means that natural science slowly departed from metaphysics and became the dominant worldview in all research. Even so, there was also a widespread skepticism within philosophy, which makes these “new” ideas by far more complex than what is often presented in the social sciences. The essence is that it marks a profound change in thinking about nature, and in how to obtain knowledge about the world. This is a simple description of positivist quantitative research (see [Christians, 2005](#), pp. 139–144 on the Enlightenment, autonomy, Rousseau and Mill), but it captures the key elements in our context.

In quantitative research, this has been translated into guidelines to help provide analytic results we can generalize beyond the sample in search of true, lasting knowledge. This is based on the assumptions that there is a stable and external reality. This ontology

demands control of all matters that may cause bias and skewed results, or in other ways interfere with research as neutral and objective. The stimuli-response model demands all respondents to be exposed to the same question. This explains the weight given to find the correct question formulation in survey research. It also explains the weight on instructions and training in correct interview behavior to maintain a distance to the respondent to avoid influence. Some textbooks are even more specific through instructions on a dress code to appear as neutral and minimizing outside informal talk beyond the interview questions, and to control your emotions. All of this is considered as a possible source of contaminating data. Such instructions reflect a particular image of the world, of the participant and of normative principles of our relationship with those involved in our studies. To assume that knowledge is objective makes the researcher into the authoritative knower also when writing up research. In this paradigm, the importance with participants is to gain access to the relevant data they already carry with them, hence data “collection.” Our interest is with them as containers or “vessels of answers” (Gubrium and Holstein, 1997, p. 65), and it follows from the according epistemology to keep a distance, to not get involved, to control one’s feelings and to never become emotionally involved. This makes research ethics external and more straightforward as in much contemporary formal research ethics. However, in contrast to historical violations, participants are now entitled to become informed about the study and to be asked for their consent to participate, trusting that the researchers will treat the data as agreed upon and otherwise respect them. They are also free to drop out whenever they desire. In this worldview, ethical research protocols share principles based on the same ontological anchor as in biomedicine, with laboratory studies with individual participants in a rather manageable process. To illustrate, the US history of evaluation has been quantitatively dominated with a view on science and morality as separate and on evaluation as apolitical. This caused House (2005) to argue that social justice concerns have been “routinely omitted from evaluation discussions” (p. 1074) due to the evaluators’ unfamiliarity with philosophy and the affiliations with “politics” (p. 1075). And, by his “interest in the normative political characteristics of ethics in evaluation” (p. 307), Schwandt (2018) challenges this apolitical character of professional ethics in traditional evaluation. Another classic, the PISA (Program for International Student Assessment), is “the largest international study of education, examining the reading, mathematics and science achievements of students as they approach the end of compulsory schooling” (<https://www.erc.ie/studies/pisa/>). Here, Sjöberg and Jenkins (2022) point to the political character by the impacts on national curriculum policies, through politicians’ suggestions to modify school curricula to maximize PISA performance (p. 10).

Qualitative researchers see positivism differently. Post-positivist, constructivist researchers argue that it is reductionist, deterministic and simply naïve.

Ethnomethodology: doing being ethical as collaborative accomplishment

Professions such as nursing and social work tend to refer to phenomenology that draws on Husserl and Heidegger. They also refer to Alfred Schütz on interpretation when they interview about meaning to understand experience—a commonsense notion, but David Silverman argues they have misunderstood Schütz who would study commonsense reasoning as a topic “rather than to rely on it as a tacit resource” (D. Silverman, 2017, p. 148) so crucial to Harvey Sacks who developed ethnomethodology further. This reminds us of the complexity of philosophical theories. Sacks (Silverman, 1998) in his critical dialogue with Schütz’s theory, moved away from phenomenology and developed ethnomethodology where researchers analyze taped talk without meeting their study participants. The interest is with ordering or the practices through which lived order is organized (hence the “doing being”), according to Pollner and Emerson, “a measure of analytic distance” (2010, p. 124). This makes certain ethical strengths often pass without seeing.

Firstly, the distance to the members involved makes the classic ethnomethodologist stay indifferent to whatever participants say, but also become worried about getting involved with members’ concerns. This is easier when analyzing taped talk compared to when working ethnographically. Secondly, in ethno-related research on taped data there is much member interaction, but no researcher-participant interaction because they simply do not meet. This reduces ethical risks. On the other hand, critics would argue that they do meet, but their relationship is a textual one. They would also argue this still leaves the researcher in control of analysis and the final text. The researcher also negotiates access to the organization and gatekeepers, who then negotiate with potential participants. This causes ethnomethodology’s indifference to be conditional, as they do not escape research ethics issues.

Still, this approach brings an ethical research strength that often passes unnoticed. We may draw on Paul Atkinson and David Silverman (1997), who point to the prevalent failure not to recognize the “work” participants do in most interview studies. To help illustrate this, patients in the emergency room often have several problems, while the doctors’ “What can I help you with?” usually diverts only one of the patient’s problems. In John Heritage’s sequential analysis of talk in the emergency room (in Silverman, 2007, p. 116), he finds that doctors tend to ask patients one of two questions:

Do you have any other medical problems?

Do you have some other medical problems?

The analysis shows that “any questions” tend to elicit much fewer questions compared to “some questions” (not shown here). Such details only available and noticeable on tape tell us that patients tend to hear “any” and “some” in quite different ways. “Any” is heard as if the doctor expects no further questions, as opposed to “some.” This makes “some questions” more patient-friendly. The

interaction is shared with the reader, as the patient's talk is just as important as the doctor's talk, and their joint *in situ* data (usually not shown in other method talks) make the analyst avoid jumping to conclusions based on interpretations of patients' stories about their experiences. All talk is available to the reader and the analytic strategy does not interpret talk, but rather members' methods in creating social order. Extracts of sequential talk also make it easy for doctors to recognize themselves in such natural contexts (though the simple example above illustrates our point) (see Silverman, 2007, p. 116). For more illustrations, see Drew and Heritage, 1992, Ryen, 2009 for Sacks' MCD analysis, Heritage, 2018, Drew, 2018. This shows that data on "knowledge-in-action" also makes it easier to invite the professionals to discuss how to improve their own practice to the benefit of the patients.

Thirdly, this illustrates the entanglement of epistemology, analytic strategy, and research ethics, which invites the question: Can an analytic strategy itself be unethical? (And, what about positivism?). In this view, much feminist research concerned with voices is regarded as analytically inadequate due to treating these as unmediated descriptions used as a resource for the researcher. When ethnomethodologists argue this as being naïve, it reflects an epistemological disagreement. When feminist research assumes a special ability in "unmasking delusions" as Silverman describes it (1993, p. 179), their failure, he continues (1993, p. 181) is "too often ignoring what people actually are saying and doing." He reminds us (1993, p. 181) "not to confuse actor's accounts with sociological [or other disciplines'] analysis" (for further discussions of interpretivist, constructivist thinking and philosophical roots, see Schwandt, 1994, pp. 118–137). They would argue that all qualitative research that assumes responses to be authentic and simple truth claims fails. A major argument is that data must be analyzed, and not simply reported. We will later see that critical feminism counters such accusations. Fourthly, through the apparent micro-orientation of this philosophical position, it offers us the ability to analyze how the wider contexts, such as school policies, advise dealing with Covid-19, or that everyday circulating images and concepts enter ordinary talk by studying how social reality gets accomplished. This makes the argument that it ignores power irrelevant, just as we saw above in the example on making the hidden in doctor-patient talk visible.

As referred to, such detailed interactional data are frequently used for educational purposes with professionals and others. Conversational analyses (CA) of access to public services in emergency service calls in general (Svennevig, 2012) or calls with native—and non-native speakers in which language differences may complicate the interaction between call takers and callers (Garcia, 2021), illustrate this well. To learn how to improve such interactions will better prepare call takers to avoid occurring challenges in such calls. Others have investigated how couples manage conflicting knowledge claims when one of the persons has dementia (PWD) (Landmark, 2021). Such analyses will help educate carers and family members in how to deal with such issues and still treat PWDs as competent conversational participants (p. 208).

When researchers within this method talk work ethnographically, they will face much the same ethical challenges that others do (for references see Pollner and Emerson, 2010, for analyzing mobile talk see Ryen, 2008, research ethics in ethnography work 2012, and see Roulston, 2018 on her students' experiences in online coursework). This opens to more nuanced analyses of power plays, and how hierarchical thinking plays out in practice (Ryen, 2019a). This framework is also employed in feminist research on focus groups (Wilkinson, 2021) and as we saw in Conversational Analysis (Kitzinger, 2000 and 2007).

Constructivism: rejecting old scripts

Reality as interactional products makes ethical issues emergent and most evident in ethnographic work. This makes qualitative research a most heterogeneous category. Both ethnomethodology and the social construction of reality are seen as disciplinary extensions beyond the social phenomenology of Alfred Schütz and developed further by Thomas Luckman and Peter Berger as a study of things as they appear through the acting subject's perspective. This rejects the positivist claim of objectivity and makes knowledge a human construct. We recognize this in philosophy of science by Thomas Kuhn on changes in paradigms, in Wittgenstein on language games, and in Joe L. Kincheloe's influential critical constructivist epistemology in education (Kincheloe and McLaren, 2005, pp. 303–342). It also implies a criticism of standardized ethical guidelines as universal and with standardized solutions readily available to external persons. Rather, as interactional outcomes, ethical issues are emergent throughout the whole research process. This calls for self-reflectivity both within and across cultural contexts and calls for reflection on issues like (Ryen, 2019a, p. 192): "Who am I to others? What am I a part of? Do I want that?", and Fazal Rizvi who claims that "Knowledge claims are, if you like, preceded by questions" (2009, p. 112) such as "who is this 'our' that is being talked about?" and "does it include me?"—and if it doesn't then, "how is language being used to marginalized people?" (2009, p. 109). Ethnographers argue their research ethical strength is with bringing in contextual knowledge which may uncover issues like hidden power. This also deals with protecting participants from harm and their concern with participant rights. However, the individual—collective of rights reflects a historic moment in the Western society with autonomy and self-determination, and reminds us of the local contextual of contemporary codes and regulations. To illustrate, by ignoring ethnographers' sensitivities to methods specific issues, they may rather increase the risk of harm.

The historical background to formal research ethics paved the way for a focus on the classic concerns with codes and consent, confidentiality, and trust (Faden and Beauchamp, 1986). Ethnographic work has showed these as much more complex than assumed (Stake and Jegatheesan, 2008; Ryen, 2021) captured in Guillemin and Gillam's (2004, Baker et al., 2016) "ethically important moments" (p. 262) in "ethics in practice" (p. 264, Renold et al. 2008) different from committee talk. Consent presupposes knowledge about what it means to participate, what it is that you consent to, that you can evaluate benefits and risks, understand what it implies to consent when reusing archived material if relevant, and the symbol of signing a document. If local procedures allocate actual consent with a gate keeper the potential participant finds it hard to oppose, things get complex. We simply get entangled in local authority—and loyalty webs despite participant consent as captured in Sabari and Ben-Yehoshua's title

(2017) "I'll sue you if you publish my wife's interview" on life story research in Israel. The emergent of ethnographic encounters have made ethnographers rather prefer a process consent procedure over an initial consent form at the start. Other issues such as data ownership and control over analysis and interpretation linger on (Ryen, 2021, p. 37). Further, not all regimes demand written consent, and historic memories of betrayal and "signing away" may invoke distrust with signatures. This reminds us that research ethics with confidentiality/anonymity and trust are more complex than assumed in formal codes based on a quite different ontology (Ryen, 2021, pp. 39–41, see Olesen 2005, pp. 235–278 for challenges in feminist research, and Lincoln, 2009, pp. 150–169 on the many shifts in ethical frameworks in qualitative research, Mertens and Ginsberg, 2009). Codes assume that participants want their participation to remain confidential, but some 'do want their names in the report in accordance with regional traditions, to prohibit their stories from being taken away from them or for assuming being published may open opportunities. To ignore such requests may appear as patriarchal and disregard of local preferences, or the researcher may wrongly assume lack of awareness of potential consequences of disclosure, or of other risks. Despite consent, confidentiality may be tricky in small places and participants easily detected by journalists, or people may react to the report for other reasons (Ellis, 1995). Murphy and Dingwall (2010, p. 340) also remind us to be cautious of any cathartic effects of research interaction often claimed by feminists. And, if ethnographers document how participants adapt to structural and institutional pressures, people with power may use the report "to control and manipulate those without" (2010, p. 341). This shows that ethnographers have discussed interpretive and representational authority even prior to the postmodernist turn. This reminds us of the regional of Western philosophical anchors and ethical regulations (see Liamputtong, 2012, Nortjé et al. 2019; Trimble and Fisher 2006; H. Silverman, 2017). Such experiences and arguments reflect the close and complex relationship between philosophical anchors, politics, and ethics in qualitative research (Bauman, 2005, pp. 1083–1098, Christians, 2005, pp. 139–164). The politics of constructivist ethnographic research make visible the intricate of field relations and the subtleties and messiness of fieldwork (Delamont, 2007, pp. 2–4).

Stanton Wortham and Kara Jackson remind us that there is a vast variety in the certainties that constructionists presuppose, including in educational constructionism (2008). They remind us that, "Various objects are constructed in educational processes, such as the identities of teachers and learners, the subject matter learned, and the social structures produced and reproduced"; this takes place through mechanisms at different organizational levels and include "psychological, interactional, cultural, and social elements" (p. 107), and that such studies may contribute to transform education (p. 107). Gomez (2016) on working with the Roma Community, and Hurd (2019) on mixed identity illustrate Wortham and Jackson's points above well.

The politics and ethics of much interpretive research overlaps much with that of more action-oriented approaches who more explicitly flag their transformative aim. However, as to the very pursuit of research, they follow separate pathways. This calls for a sub-section of their own.

Feminist and indigenous research: community grounded research ethics

Joe L. Kincheloe's argued his critical constructivist epistemology is important to becoming an educated person and to social change and—justice. These political and pedagogical arguments became crucial for the movement of critical constructivism in the 1990s onwards concerned with political struggles and power in knowledge construction, and the perspectives of oppressed peoples. It makes emancipation or transformation the very task of research as in feminist and indigenous research based on his critical ontology which brings a new constellation of politics and research ethics with research endowed with a moral obligation toward social justice based on their own worldview free from the burden of classic academic scripts (Hesse-Biber and Nagy, 2011; Hall, 2021; Kegel, 2011; Loomba and Lukose, 2012). With the researcher as bricoleur (Kincheloe and McLaren, 2005, pp. 316–317) s/he will use all tools available to unpack the relationship between his/her way of seeing and the social location of his/her personal history. This makes the relationship between individuals and their contexts crucial to uncover the invisible inherit in power and culture (2005, p. 317). This includes to become aware of epistemological presuppositions that inform both our research and our own thinking (2005, p. 305) as with the authoritative research ethics codes and calls for Freire's pedagogy of the oppressed to eradicate perpetuating oppression. It argues truth as discursively situated and implicated in power relations (2005, p. 327, Jakobson, 2016). This applies to both feminist and indigenous research as reflected in their ethics of care.

Feminist prescriptive ethics implies to treat participants with respect in all phases of the research process and includes reciprocity, honesty, accountability, responsibility, equality etc. (Skeggs, 2010, p. 433). The American grown ethics of care is a feminist relational philosophical perspective and a distinct normative moral theory based on the experience of care and care work adopted by many American feminist and other "post-" researchers. It has faced much criticism by feminists and others and seen as regional by reflecting a particular culture that appeals more to Americans than to Europeans as in Hammersley and Traianou's criticism (Hammersley and Traianou, 2012). Feminist research, always concerned with power—how it works and how to challenge it (Skeggs, 2010, p. 426, Atanga, 2013), defends values of reciprocity, equalizing power control, and conscientization (see Skeggs, 2010, p. 435 for criticism). These values also inspired indigenous research though feminist research has pinpointed some more problematic links between politics and ethics related to the need to translate women's voices from talk to text. This refers to the issue of texts as representations of true experiences and relates to the meeting between epistemological-analytical and activism. In practice, some handled this by looking at discourses and ways women represented their stories, others kept them "raw" due to their political potentials. These debates later initiated new experimental representational practices including poetry, theater, assemblages etc. which enlarged the cleft between other epistemologies and feminist activism, but also inspired indigenous researchers to reject old epistemological constraints based on the much-criticized Enlightenment period (Israel, 2010). Feminist and indigenous researchers both use past experiences in their fight for a different future. They draw on their own history and

experiences from which new, different ethical principles grew in respect of feminist and indigenous research as a counterhegemonic activity for social justice. The latter speaks from within an interpretive community based on its own historical research traditions and claims a need not only to resist, but also for reordering and redefining themselves as “we” see “ourselves”, not as another “other” (see Ryen, 2019b). Indigenous scholars and activists including Linda Tuhiwai Smith (2005/2012), Bishop (2008), Kovach (2009), Marie Battiste (as in Battiste and Henderson, 2000, Denzin et al., 2008 Part IV Power, Truth, Ethics, and Social Justice pp. 439–458) insist on indigenous peoples’ control over their knowledge and heritage based on guidelines developed by their tribal communities and indigenous worldviews. Denzin and Giardina (2012, p. 20) concerned with pedagogics refer to a praxis-based ethics, an ethics that resists oppression and which enacts a hope of possibility “grounded in performative practices that embody love, hope, care, and compassion”. This demands that indigenous people—not only individuals—shall be informed that research is being carried out on, among and with them, researchers will need training in relevant research protocols, and it is required that the community shall benefit from research. Despite internal variations, indigenous researchers have for political reasons decided to speak with one voice just as in early feminism. It is no longer about interpretations, but about change (Denzin and Gardina, 2009, p. 13). This turn, as with all battles, undermines internal relations more beneficial to some than to others, and this social justice refers to protecting the many over the few. So far, researchers and participants will have to comply with that.

Storytelling has been important in both feminist and indigenous research. To help illustrate, in her article, “de-colonizing practices: Negotiating narratives from racialized and gendered experiences of education”, Phoenix (2009) refers to postcolonial feminist work when she draws on Caribbean serial migrants’ narrative accounts to “examines how (post)colonial legacies intersect with racialization and gender in experiences of education to produce subjectification as racialized and gendered subjects” (p. 102). Also, indigenous research draws on storywork as in Pedri-Spade’s (2016) performance ethnography, where Anishinabe participants share stories, sing and drum. This is part of a critical pedagogy with an emancipatory potential, in which both analytic quality and research ethics rest with indigenous ways of knowing and doing.

Conclusions

We are once more reminded of Murphy and Dingwall on “The ethics and the politics of ethnography are not clearly separable ...”. Changes in images of the participant, of research relations and of the purpose of research, make research ethics more complex and the dominant guidelines less adequate. In “Proclaiming the New” (Kuntz and St. Pierre, 2021b in their 2021a) Marek Tesar in his « -Philosophy as a Method» (2021) argues that “the core of philosophy as a method is the ethical relationship with a thought” (p. 545). He urges the reader to work with original texts. He turns and returns to Decartes, Dewey, Freire, Deleuze and many more, and by “What is Philosophy?” (p. 252) he emphasizes the question mark and asks: “After the linguistic turn, the cultural turn, the historical turn, the post-structural turn, and the post-human turn, where do we turn?” (p. 252). The question mark is important because these turns is what drives qualitative research. Each turn destabilizes politics and ethics and brings something new.

References

- Ames, R.T., 1998. East Asian Philosophy, Routledge Encyclopedia of Philosophy. Taylor and Francis. <https://doi.org/10.4324/9780415249126-G218-1>. <https://www.rep.routledge.com/articles/overview/east-asian-philosophy/v-1>.
- Angle, S.C., Slote, M. (Eds.), 2013. *Virtue Ethics and Confucianism*. Routledge, Abington and NY.
- Atanga, L.L., 2013. African feminism? In: Atanga, L.A., Ellece, E.S., Litosselti, L., Sunderland, J. (Eds.), *Gender and Language in Sub-saharan Africa. Tradition, Struggle and Change*. John Benjamins Publishing Company.
- Atkinson, P., 2005. Qualitative research—unity and diversity [25 paragraphs]. *Forum Qual. Soc. Res.* 6 (3). Art. 26.
- Atkinson, P., Silverman, D., 1997. Kundera’s immortality: the interview society and the invention of the self. *Qual. Inq.* 3 (3), 304–325.
- Atkinson, P., Seweryn, A., Tirini, S., 2011. Knowing selves: biographical research and European traditions. *Int. Rev. Qual. Res.* 4 (4), 461–487.
- Baker, L., Phelan, S., Snelgrove, R., Varpio, L., Maggi, J., Ng, S., 2016. Recognizing and responding to ethically important moments in qualitative research. *J. Grad. Med. Educ.* 8 (4), 607–608.
- Barnard, R., Wang, Y. (Eds.), 2020. *Research Ethics in Second Language Education Universal Principles, Local Practices*. Routledge.
- Battiste, M., Henderson, J.Y., 2000. *Protecting Indigenous Knowledge and Heritage: A Global Challenge*. Purich Press, SK.
- Bauman, Z., 2005. Afterthought: on writing; on writing sociology. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. 1089–1098.
- Bishop, R., 2008. Te kotahitanga: kaupapa māori in mainstream classrooms. In: Denzin, N.K., Lincoln, Y.S., Smith, L.T. (Eds.), *Handbook of Critical and Indigenous Methodologies*, first ed. SAGE, Los Angeles, pp. 439–458.
- Blackburn, S., 2016. Ethics applied. In: *The Oxford Dictionary of Philosophy*. Oxford University Press. <https://doi.org/10.1093/acref/9780199541430.001.0001>.
- Bristow, W., 2017. Enlightenment. In: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. URL. <https://plato.stanford.edu/archives/fall2017/entries/enlightenment/>.
- Caplan, A.L., 1992. When Medicine Went Mad. Bioethics and the Holocaust. Humana Press, Totowa, NJ.
- Carlson, D.L., Wells, T.C., Mark, L., Sandoval, J., 2021. Introduction: working the tensions of the post-qualitative movement in qualitative inquiry. *Qual. Inq.* 27 (2), 151–157.
- Chakrabarty, D., 1992. Postcoloniality and the artifice of history: who speaks for “Indian” pasts? *Representations* 37, 1–26.
- Christians, C.G., 2005. Ethics and politics in qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. 139–164.
- Crisp, R., 2011. Applied ethics. In: *Ethics*. Routledge Encyclopedia of Philosophy, Taylor and Francis. <https://doi.org/10.4324/9780415249126-L132-2>. <https://www.rep.routledge.com/articles/overview/ethics/v-2/sections/ethics-and-metaethics>.
- Delamont, S., 2007. Arguments against auto-ethnography. *Qual. Res.* 4, 2–4.
- Denzin, N.K., Gardina, M.D., 2009. Introduction. Qualitative inquiry and social justice. Towards a politics of hope. In: Denzin, N.K., Gardina, M.D. (Eds.), *Qualitative Inquiry and Social Justice. Towards a Politics of Hope*. Left Coast Press, Walnut Creek, California, pp. 11–50.

- Denzin, N.K., Giardina, M.D., 2012. Introduction. Qualitative inquiry and the politics of advocacy. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry and the Politics of Advocacy*. Left Coast Press, Walnut Creek, pp. 9–38.
- Denzin, N.K., Lincoln, Y.S., 1994. Introduction. Entering the field of qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, first ed. SAGE, Thousand Oaks, pp. 1–17.
- Denzin, N.K., Lincoln, Y.S., 2005. Preface. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. ix–xix.
- Denzin, N.K., Lincoln, Y.S., Smith, L.T. (Eds.), 2008. *Handbook of Critical and Indigenous Methodologies*, first ed. SAGE, Los Angeles, pp. ix–xv.
- Drew, P., 2018. Epistemics in social interaction. *Discourse Stud.* 20 (1), 163–187.
- Drew, P., Heritage, J. (Eds.), 1992. *Talk at Work: Interaction in institutional settings*. (Studies in interactional sociolinguistics, 8.) Cambridge University Press, Cambridge & New York.
- Ellis, C., 1995. Emotional and ethical quagmires in returning to the field. *J. Contemp. Ethnogr.* 24, 68–98.
- Faden, T., Beauchamp, L., 1986. *History and Theory of Informed Consent*. Oxford University Press, NY.
- Fakhry, M., 1998a. Ethics in Islamic philosophy. In: *Routledge Encyclopedia of Philosophy*. Taylor and Francis. <https://doi.org/10.4324/9780415249126-H018-1>.
- Fakhry, M., 1998b. Greek philosophy: impact on Islamic philosophy. In: *Routledge Encyclopedia of Philosophy*. Taylor and Francis.
- Garcia, A.C., 2021. Interactional challenges for non-native speakers of English telephone calls. *J. Pragmat.* 193, 222–234.
- Gomez, A., 2016. Communicative methodology: researching “with” rather than “on” the romà community. *Int. Rev. Qual. Res.* 9 (2), 178–184.
- Guba, E.G., Lincoln, Y.S., 1994. *Competing Paradigms in Qualitative Research*.
- Gubrium, J.F., Holstein, J.A., 1997. *The New Language of Qualitative Method*. Oxford University Press, Oxford.
- Guillemin, M., Gillam, L., 2004. Ethics, reflexivity, and “ethically important moments” in research. *Qual. Inq.* 10 (2), 261–280.
- Haggerty, K.D., 2004. Ethics creep: governing social science research in the name of ethics. *Qual. Sociol.* 27 (4), 391–414.
- Hall, K.Q. (Ed.), 2021. *The Oxford Handbook of Feminist Philosophy*. Oxford University Press Inc, NY.
- Hammersley, M., 2009. Against the ethicists: on the evils of ethical regulation. *Int. J. Soc. Res. Methodol.* 12 (3), 211–225.
- Hammersley, M., Traianou, A., 2012. An Alternative Ethics? Justice and care as guiding principles for qualitative research. *Socio. Res. Online* 19 (3), 104–117.
- Heritage, J., 2018. The ubiquity of epistemics: a rebuttal to the “epistemics of epistemics” group. *Discourse Stud.* 20 (1), 14–56.
- Hesse-Biber, Nagy, S. (Eds.), 2011. *Handbook of Feminist Research. Theory and Praxis*, second ed. SAGE, Los Angeles.
- Hoffman, S.G., Kumar, V., 2020. Ontology. In: Atkinson, P., Delamont, S., Cernat, A., Sakshaug, J.W., Williams, R.A. (Eds.), *Handbook of Ethnography*. SAGE Research Methods.
- House, E.R., 2005. Qualitative evaluation and changing policy. In: Denzin, N., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. 1069–1081.
- Hurd, E., 2019. *The reflexivity of pain and privilege. Auto-Ethnographic Collections of Mixed Identity*. Brill Sense, Leiden.
- Israel, J., 2010. *A Revolution of the Mind: Radical Enlightenment and the Intellectual Origins of Modern Democracy*. Princeton.
- Issawi, C., Leaman, O., 1998. Ibn Khaldun, “abd al-rahman” (1332–1406). In: *Routledge Encyclopedia of Philosophy*. Taylor and Francis. <https://doi.org/10.4324/9780415249126-H024-1>.
- Jakobsen, H., 2016. How violence constitutes order: consent, coercion, and censure in Tanzania. In: *Violence against Women*, vol. 24, p. 2018. <https://doi.org/10.1177/1077801216678091>. Issue 1.
- Johnson, C., 2020. Taoist leadership ethics. *J. Leadersh. Organ. Stud.* 7 (1), 82–91.
- Jupp, V., 2006. Epistemology. In: *The SAGE Dictionary of Social Research Methods 1-0*. SAGE Publications, Ltd, London.
- Katz, J., 1996. The Nuremberg code and the Nuremberg trial: a reappraisal. *JAMA* 276 (20), 1662–1666.
- Katz, J., 2007. Towards a natural history of ethical censorship. *Law Soc. Rev.* 41 (4), 797–810.
- Kegel, C.M., 2011. Global feminism. In: Edelglass, W., Garfield, J.L. (Eds.), *The Oxford Handbook of World Philosophy* (Chapter 39).
- Khaldun, I., 2012. *Al-Muqaddimah. Boken om lærdommene, med en introduksjon til verdens historie*, pp. 1–2 (Translated by Magid Al-Araki. Oslo: Pax Forlag).
- Kim, C.-T., 2013. Justificatory structures: taoist and confucian ethics. *J. East-West Thought* 3, 63–78.
- Kincheloe, J.L., McLaren, P., 2005. Rethinking critical theory and qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. 303–342.
- Kitchener, K.S., Kitchener, R.F., 2009. Social science ethics. Historical and philosophical issues. In: Mertens, D.M., Ginsberg, P. (Eds.), *The Handbook of Social Science Research Ethics*, first ed. SAGE, Los Angeles, pp. 5–22.
- Kitzinger, C., 2000. Doing feminist conversation analysis. *Fem. Psychol.* 10 (2), 163–193.
- Kitzinger, C., 2007. Feminist approaches. In: Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*, first ed. SAGE, Los Angeles, pp. 113–128.
- Kovach, M., 2009. *Indigenous Methodologies: Characteristics, Conversations, and Contexts*. University of Toronto Press, Toronto, Canada.
- Kuntz, A.M., St. Pierre, E.A., 2021a. Special issue: new approaches to inquiry. *Qual. Inq.* 27 (5), 475–479.
- Kuntz, A.M., St. Pierre, E.A., 2021b. Proclaiming the new: an introduction. *Qual. Inq.* 27 (5), 475–479.
- Landmark, A.M.D., 2021. Couples living with dementia managing conflicting knowledge claims. *Discourse Stud.* 23 (2), 191–202.
- Lefor, A.T., 2005. Scientific misconduct and unethical human experimentation: historic parallels and moral implications. *Nutrition* 21 (7–8), 878–882.
- Liamputtong, P., 2012. *Performing Qualitative Cross-Cultural Research*. Cambridge University Press, Cambridge.
- Lincoln, Y.S., 2009. Ethical practices in qualitative research. In: Mertens, D.M., Ginsberg, P. (Eds.), *The Handbook of Social Science Research Ethics*, first ed. SAGE, Los Angeles, pp. 150–169.
- Loomba, A., Lukose, R.A. (Eds.), 2012. *South Asian Feminisms*. Duke University Press, Durham.
- Mertens, D.M., Ginsberg, P., 2009. Section II. Perspectives on ethical regulation. In: *The Handbook of Social Science Research Ethics*, first ed. SAGE, Los Angeles, pp. 103–183.
- Murphy, E., Dingwall, R., 2010. The ethics of ethnography. In: Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*, third ed. SAGE, Los Angeles, pp. 339–351.
- Nortjé, N., Visagie, R., Wessels, J.S., 2019. Social science research ethics in Africa. Vol. 7. In: Hunter, D., McMillan, J., Weijer, C. (Eds.), *Research Ethics Forum*. Springer Nature Switzerland AG. Springer International Publishing. Nuremberg Code. https://en.wikipedia.org/wiki/Nuremberg_Code.
- Nuremberg Code, 1996. *Br. Med. J.* 313, 1448. No. 7070. https://media.tghn.org/medialibrary/2011/04/BMJ_No_7070_Volume_313_The_Nuremberg_Code.pdf.
- Olesen, V., 2005. Feminist qualitative research in the millennium’s first decade: developments, challenges, prospects. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. SAGE, Thousand Oaks, pp. 235–278.
- Park, P.K.J., 2013. *Africa, Asia, and the History of Philosophy: Racism in the Formation of the Philosophical Canon, 1780–1830*. State University of New York Press, NY.
- Pedri-Spade, C., 2016. “The drum is your Document”: decolonizing research through Anishinabe song and story. *Int. Rev. Qual. Res.* 9 (4), 385–406.
- Phoenix, A., 2009. De-colonising practices: negotiating narratives from racialised and gendered experiences of education. *Race Ethn. Educ.* 12 (1), 101–114.
- PISA, Programme for International Student Assessment <https://www.oecd.org/pisa/>.
- Pollner, M., Emerson, R.M., 2010. Ethnomethodology and ethnography. In: Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*, third ed. SAGE, Los Angeles, pp. 118–135.
- Punch, M., 1994. Politics and ethics in qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, first ed. SAGE, Thousand Oaks, pp. 83–97.
- Renold, E., Holland, S., Ross, N.J., Hillman, A., 2008. “Becoming participant”. Problematising “Informed consent” in participatory research with young people in care. *Qual. Soc. Work* 7 (4), 427–447.
- Reverby, S., 2009. *Examining Tuskegee: The Infamous Syphilis Study and its Legacy*. The University of North Carolina, Chapel Hill.

- Rigney, A., 1990. *The Rhetoric of Historical Representation: Three Narrative Histories of the French Revolution*. Cambridge University Press, Cambridge.
- Rizvi, F., 2009. Mobile minds. In: Kenway, J., Fahey, J. (Eds.), *Globalizing the Research Imagination*. Routledge, London, pp. 101–114.
- Roulston, K., 2018. Qualitative interviewing and epistemics. *Qual. Res.* 18 (3), 322–341.
- Ruyter, K.W., 2003. Forskningsetikkens historie. In: Ruyter, K.W. (Ed.), *Forskningsetikk: beskyttelse av enkeltpersoner og samfunn*, first ed. Gyldendal Norsk Forlag, Oslo, pp. 17–25. <https://www.forskningsetikk.no/ressurser/fbib/systematiske-historiske/historie/>.
- Ryen, A., 2001. Cross-cultural interviewing. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research*, first ed. SAGE, Thousand Oaks, US, pp. 335–353.
- Ryen, A., 2007. Do Western research ethics work in Africa? A discussion about not taking “the taken-the-granted” for granted. *Mosenodi* 5 (1), 31–45.
- Ryen, A., 2008. Trust in cross-cultural research: the puzzle of epistemology, research ethics and context. *Qual. Soc. Work* 7 (4), 448–465.
- Ryen, A., 2009. Ethnography. Constitutive practice and research ethics. In: Mertens, D.E., Ginsberg, P.E. (Eds.), *Handbook of Social Research Ethics*, first ed. SAGE, Los Angeles, pp. 229–242.
- Ryen, A., 2011. Ethics and qualitative research. In: Silverman, D. (Ed.), *Qualitative Research*, third ed., pp. 416–438.
- Ryen, A., 2012. Assessing the risk of being interviewed. In: *The SAGE Handbook of Interview Research: The Complexity of the Craft Second*. SAGE Publications, ISBN 978-1-4129-8164-4, pp. 477–493. Chap. 33.
- Ryen, A., 2019a. The ubiquitous rank. Some reflections on walking on thin ice. In: Hurd, E. (Ed.), *The Reflexivity of Pain and Privilege. Auto-Ethnographic Collections of Mixed Identity*, first ed. Brill Sense, pp. 191–209.
- Ryen, A., 2019b. Indigenous methods. In: Atkinson, P., Delamont, S., Cernat, A., Sakshaug, J.W., Williams, R.A. (Eds.), *SAGE Research Methods*.
- Ryen, A., 2021. Research ethics: between care and control. In: Silverman, D. (Ed.), *Qualitative Research*, fifth ed. SAGE, Los Angeles, pp. 35–49.
- Sabar, G., Ben-Yehoshua, N.S., 2017. “I’ll sue you if you publish my wife’s interview”: ethical dilemmas in qualitative research based on life stories. *Qual. Res.* 17 (4), 408–423.
- Sardar, Z., 2006. *How Do You Know: Reading Ziauddin Sardar on Islam, Science and Cultural Relations?* Pluto Press, London.
- Schwandt, T.A., 1994. Constructivist, interpretivist approaches to human inquiry. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, first ed. SAGE, Thousand Oaks, pp. 118–137.
- Schwandt, T., 1993. Acting together in determining value: a professional ethical responsibility of evaluators. *Eval.* 20 (3), 306–317.
- Silverman, D., 1993. *Interpreting Qualitative Data*. SAGE, London.
- Silverman, D., 1998. *Harvey Sacks. Social Science and Conversation Analysis*. Polity Press, Cambridge.
- Silverman, D., 2007. *A Very Short, Fairly Interesting and Reasonably Cheap Book about Qualitative Research*. SAGE, Los Angeles.
- Silverman, D., 2013. *Doing Qualitative Research: A Practical Handbook*, fourth ed. SAGE, London.
- Silverman, D., 2017. How was it for you? The interview society and the irresistible rise of the (poorly analyzed) interview. *Qual. Res.* 17 (2), 139–143.
- Silverman, H. (Ed.), 2017. *Research Ethics in the Arab Region*. Vol. 5. In: Hunter, D., McMillan, J., Weijer, C. (Series editors) *Research Ethics Forum*. Springer International Publishing, Springer Nature Switzerland AG.
- Silverman, D. (Ed.), 2021. *Qualitative Research*, fifth ed. SAGE, Los Angeles.
- Simons, P.M., 2015. Ontology. In: *Encyclopedia Britannica*, 19 Jan. <https://www.britannica.com/topic/ontology-metaphysics>.
- Sjøberg, S., Jenkins, E., 2022. PISA: a political project and a research agenda. *Stud. Sci. Educ.* 58 (1), 1–14. <https://doi.org/10.1080/03057267.2020.1824473>.
- Skeggs, B., 2010. Feminist ethnography. In: Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*, third ed. SAGE, Los Angeles, pp. 426–442.
- Smith, L.T., 2005. *Decolonizing Methodologies, Research and Indigenous Peoples*. Zed Books Ltd, London.
- Stake, R., Jegatheesan, B., 2008. Access, a zone of comprehension, and intrusion. In: Jegatheesan, J. (Ed.), *Access: A Zone of Comprehension, and Intrusion*. JAI Press, Emerald, UK.
- Stroll, A., Martinich, A.P., 2021. “epistemology”. *Encyclopedia Britannica*. <https://www.britannica.com/topic/epistemology>. (Accessed 13 May 2022).
- Svennevig, J., 2012. On being heard in emergency calls. The development of hostility in a fatal emergency call. *J. Pragmat.* 44 (11), 1393–1412.
- Taebnia, V., 2021. Differences and similarities between the later Wittgenstein’s philosophy of religion and the Islamic mystical tradition. *Asian Philos.* 1–17. <https://doi.org/10.1080/09552367.2021.1917157>.
- Tesar, M., 2021. Philosophy as a method. Tracing the histories of Intersections of “philosophy,” “methodology,” and “education”. *Qual. Inq.* 27 (5), 544–553.
- Trimble, J.E., Fisher, C.B., 2006. *The Handbook of Ethical Research with Ethnocultural Populations & Communities*. SAGE Research Methods.
- Van Norden, B.W., 2017. *Taking Back Philosophy: A Multicultural Manifesto*. Columbia University Press, NY.
- Vidich, A.J., Lyman, S.M., 1994. Qualitative Methods: their history in sociology and anthropology. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, first ed. SAGE, Thousand Oaks, pp. 23–59.
- Wilkinson, S., 2021. Analysing focus group data. In: Silverman, D. (Ed.), *Qualitative Research*, fifth ed. SAGE, Los Angeles, pp. 87–104.
- WMA. The World Medical Association, 2021a. Declaration of Geneva. In: The “Modern Hippocratic Oath”. <https://www.wma.net/what-we-do/medical-ethics/declaration-of-geneva/>.
- WMA, 2021b. WMA Declaration of Helsinki 1975. <https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/doh-oct1975/>.
- Wodak, R., 2011. Complex texts: analysing understanding, explaining and interpreting meanings. *Discourse Stud.* 13 (5), 623–633 (Special Issue on Hermeneutics and Discourse Analysis).
- Wortham, S., Jackson, K., 2008. Educational constructionisms. In: Holstein, J.A., Gubrium, J.F. (Eds.), *Handbook of Constructionist Research*, first ed., pp. 107–127.

Critical theory in qualitative research in education

Jean J. Ryoo^a and Jenifer Crawford^b, ^a University of California, Los Angeles, CA, United States; and ^b University of Southern California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

J.J. Ryoo, J. Crawford, Critical Theory, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 348–353, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01388-9>.

Introduction	46
Critical theory in qualitative research in education	46
Critical theory: an overview	46
A brief history of critical theory and the Frankfurt School	46
Critical theory in qualitative educational research	50
What makes research “critical”?	50
Contemporary critical accounts of schools in the United States	50
Applications of critical theory in qualitative educational research	51
Critical ethnography	51
Duoethnography	51
Counter story-telling	51
Participatory action research	52
Conclusion	52
References	52
Relevant websites	54

Glossary

Counter Storytelling A narrative approach to qualitative research that draws on Critical Race Theory to unmask racialized experiences and reconstructs narratives from the views of the oppressed to arrive at a more complete understanding of oppression and social change

Critical Ethnography A qualitative research tradition with roots in the 1930s community sociologists and in the 1960s action anthropologists that contribute to critiques of racial and class inequalities with research narratives of oppressed communities that were more asset-based rather than deficit oriented

Critical pedagogy A teaching approach that brings together critical theory and educational practice in order to help students question and challenge dominant beliefs and practices for higher social, critical consciousness

Critical theory A processes of critique of knowledge and power aimed to transform practices in society. Critical theory is not a research method, but has lead philosophers and qualitative researchers around the globe to challenge oppression, question neutrality, and engage in collective social action toward justice.

Dialectic Initially conceptualized as a Socratic technique of employing discussion, reasoning, and, therefore, dialog to expose false beliefs, Hegel applied the dialectic to mean a critical investigation of the process of change resulting in the theoretical application of such a process to the social sciences. Domination—The exercise of ruling power with supremacy or preeminence over another

Duoethnography A qualitative ethnographic approach that engages researchers in a dialog with each other and educational texts or artifacts that examines power, privilege and improving conditions of teaching and learning

Exploitation The act of making use of another in an unfair or mean manner for one’s own advantage. Inquiry—A systematic investigation and examination of facts or principles

Participatory Action Research An approach to qualitative research that involves researchers and community or school stakeholders collaboratively questioning and acting on structural issues of power and the inequitable distribution of social and material resources through solidarity work

Praxis A union of action and reflection and of theory and practice; critical praxis in education refers to a threefold process of self-reflection, reflective action, and collective reflective action for social change

Introduction

Critical theory in qualitative research in education

Critical theory, involves critique of knowledge and power aimed to transform practices in society. It has a long history of informing and shaping diverse qualitative research practices. Critical theory itself is not a research method, but a way to think about, know, and relate to the world with an eye toward inequitable social relations (Kinchloe and McLaren, 2011). Such a theoretical lens—both for educational researchers and otherwise—has proven valuable in: (1) exploring power relations that lead to unjust distribution of social and material resources by institutions, groups, and individuals based on economic, racial, and gender hierarchies (among others); (2) challenging qualitative researchers to question more positivist approaches that can inaccurately assume scientific neutrality among objective observers by engaging researchers in more reflexive research practices that question the biases from which they view the world; and (3) situating collective social action that challenges oppression. Through using critical theory, qualitative researchers can push educational reform for the better of all humanity.

Critical theory: an overview

Although it is difficult to point to one, universally accepted critical theory that all critical philosophers have embraced, critical theory has its roots in both the legacy of Western theoretical traditions including the Frankfurt School, and has been enriched and extended by philosophers in global contexts through growing schools of thought and processes of critique. Globally, critical theories share epistemic and ontological orientations toward transformative knowledge and practices.

As a philosophy within the social sciences, critical theory often refers to the prolific work of several generations of German thinkers belonging to Western European and Marxist traditions, many of whom are members of the Frankfurt School. Philosophers of the Frankfurt School moved away from traditional theory by seeking to reconstruct the meaning of human emancipation radically through what Max Horkheimer has described as “liberat[ing] human beings from the circumstances that enslave them” (Horkheimer, 1972: 244). Critical theory is liberating through the processes of historicizing, critiquing, and exposing the relationships of domination and subordination and the contradictions of human experiences. Such critique focuses on deconstructing how knowledge is produced, whose knowledge is valued, and how control of such knowledge equates to power in society. This critique prevents people from becoming slaves to the ideas and activities that societal institutions push upon them. Because of this search for a theory that—through critique of social constructs—could articulate freedom for all, while arguing against the suppression of human consciousness, multiple and diverse critical theories have developed since the birth of the Frankfurt School, including (but not limited to): critical race theory, critical pedagogy, critical literacy, critical feminist theory, etc. (see the section titled **Global critical philosophies**).

Viewed as a process of critique or method, we may also define critical theory as a type of self-conscious means of addressing the need for dialog and a dialectic approach around transformative, emancipatory practice. Such a process of critique pushes theorists to question continuously how they look at—and depict—the world compared to how the world actually exists, and, as a result, challenges theorists to question the very theories they form through such a process. According to critical theory, when a philosophy fails to question wider society and problematize oppression, it has become an ideology in need of critique. The Frankfurt School valued the use of such inquiry or critique as a tool for praxis—therefore, interrogating and engaging the relationship between theory and society.

A brief history of critical theory and the Frankfurt School

The best way to understand critical theory is to first locate it within Western philosophy’s historical trajectory. In doing so, one finds that we can trace critical explorations of knowledge production that form the framework of critical theory today back as far as 2500 years before the first critical scholars.

European classical philosophers

We can find a critical approach to understanding the relationship between knowledge and power in society in the work of the Greek philosophers Socrates (469–399 BC), Plato (428/427–348/347 BC), and Aristotle (384–322 BC). Through Plato’s dialogues—most notably *Gorgias* and *The Apology*—one can see how Socrates (Plato’s mentor) questioned the way the Sophists (classical Greek philosophers who exercised great power in Athens) used rhetoric to persuade people to manufacture and believe in false truths, rather than use rhetoric to search for ultimate truth for the better of humankind. Similarly, Plato’s determination to record Socrates’ stories through the written word—despite Socrates’ skepticism about the technology of writing—reflects how Plato recognized the importance of language as a meaning-making tool for the development of critical thought. Following Plato, Aristotle theorized and deconstructed the power of knowledge creation in language manipulation through *The Art of Rhetoric* and *Poetics*. Aristotle outlines the process of writing persuasive speeches that appeal to *logos* (reason), *pathos* (emotion), and *ethos* (confidence in the orator’s personal character). Although thousands of years would pass between the lives of these Greek philosophers and the foundations of critical theory which grew out of the Enlightenment, the examination of linguistic tools in knowledge production that were explored in Socrates, Plato, and Aristotle’s work proved invaluable to Western philosophy traditions.

Kant, Hegel, Marx, and modern foundations of critical philosophy

Following the Middle Ages in Europe—that marked the hegemony of the Catholic Church—and the European Renaissance, that marked the break from this Church’s control, came the birth of the Enlightenment in Europe and America. During this era, Immanuel Kant (1721–1804) highlighted the value of critique as the only way to find true knowledge and combat illusory ideologies by separating the self from the dogmas surrounding us and by questioning those concepts that are presented to us as facts. Georg Wilhelm Friedrich Hegel (1770–1831) also theorized about knowledge production but wished to push beyond Kant’s separation of mind and self when deconstructing ideologies. Instead, Hegel sought to historicize the process of arriving at truth. Hegel viewed history as progress toward a greater humanity—and Utopia—through a higher consciousness of self in and with the world. Dialectic reasoning is the path to this truth and involves engaging in dialog to overcome the contradictions between thesis and antithesis toward a synthesis.

Following Hegel, Karl Marx’s (1818–83) work with Friedrich Engels regarding the historical and material conditions of capitalism is the most influential in shaping critical theory. Marx explained how the bourgeoisie (ruling class) accumulate wealth through the alienation and exploitation of the proletariat (working class). Capitalist ideologies reinforce false consciousness and maintain power in the hands of the ruling class. Thus, Marx provided the basic idea on which critical theory is constructed: those who have economic power in society also control the cultural superstructure of that society. Marx recognized how production and consumption of media (through language) controlled class consciousness and how, simultaneously, such a tool could also be used against oppressive capitalist ideologies for a better humanity.

The Frankfurt School

The original home of the Frankfurt School was in the Institute for Social Research (Das Institut für Sozial-forschung) that was founded in Frankfurt, Germany, in February 1923, but was forced to move under Nazi threat in 1933. In 1930, the Institute came under the directorship of Max Horkheimer and grew to include critical theorists such as Erich Fromm, Herbert Marcuse, Theodor Adorno, and Walter Benjamin who shifted its theoretical focus from an analysis of bourgeois society’s socioeconomic substructures to bourgeois cultural superstructures. Although primarily influenced by Marxism, the Frankfurt School rejected an orthodox reading of Marx and represented a different form of Marxism than their Bolshevik contemporaries, as they did not envision an inevitable proletariat revolution nor focus solely on economic production. Instead, the Frankfurt School sought to understand the relationship between theory and society through a dialectical framework that also examined the realm of culture “including not only the so-called spiritual contents of science, art, and religion, but also law, ethics, fashion, public opinion, sport, amusement, lifestyle, etc.” (Horkheimer, 1972). Thus, the Frankfurt School became both a representation and critique of Western Marxism.

Central to the Frankfurt School’s articulation of critical theory was its critique of the Enlightenment’s conceptualization of instrumental reason, rationality, and positivism (as explored by diverse thinkers including Saint-Simon, Comte, the Vienna Circle, and others). The Frankfurt School believed there was a breakdown in reason due to a “crisis of science” that reduced the concept of knowledge to a limited scientific method of “description, classification, and generalization of phenomena, with no care to distinguish the unimportant from the essential” that could only be resolved through the creation of “a correct theory of the present social situation” (Horkheimer, 1972). Echoing Nietzsche’s belief that it was not science but the scientific method that proved triumphant during the nineteenth century, the Frankfurt School sought to rescue the world from the basic assumptions of logical positivism—that valued objective observation as the only way to produce truth—by envisioning critical theory as an improved approach to human reasoning, supporting a more self-conscious critique of human action for transformation. The Frankfurt School believed that positivist rationality and its narrow view of science endangered human ability to think critically and subjectively regarding the world—freezing human history behind selected, organized, and fetishized facts in false neutrality, rather than recognizing how natural reality and science relate both to human consciousness and lack of consciousness as well as subjectivity and objectivity. The Frankfurt School articulated how those engaged in using positivist rationality were not necessarily conscious of how such an approach could maintain an oppressive power imbalance or the status quo in work labeled as scientific and thus urged philosophers to consider their positionality as producers of knowledge.

In order to address these problems with logical positivity and Enlightenment rationality, the Frankfurt School formulated a critical theory that first recognized the relationship between individuals and society. In this way, critical theory develops a capacity for meta-theory by acknowledging the cultural–historical positionality of interested, non-neutral parties whose political or social interests affect how they navigate and present knowledge or information. With such a self-critical and self-conscious lens, critical theory engages in dialectical thought that lives within the tension between critique and theoretical reconstruction. As noted by Marcuse (1960), “Dialectical thought starts with the experience that the world is unfree; that is to say, man and nature exist in conditions of alienation, exist as ‘other than they are.’ ... Thought ‘corresponds’ to reality only as it transforms reality by comprehending its contradictory structure.” Dialectical methodologies synthesize philosophy, the social sciences, and radical politics to help illustrate how society, state, economy, culture, and the individual are all related to one another. Thus, critical theory should work toward an improved society devoid of injustice and oppression and “emancipation from slavery” (Horkheimer, 1972) through equal engagement of theory and practice in critical praxis. While the concept of “praxis” dates back to ancient Greek philosophy, in the context of critical theory praxis is the changing of circumstance and human activity (Marx and Engels, 1998) that seeks “to liberate human beings from the circumstances that enslave them” (Horkheimer, 1982: 244).

The Frankfurt School also used critical theory to reject essentializing and hierarchical perspectives of culture that were supported by sociologists and orthodox Marxists of the time. The Frankfurt School rejected definitions of culture as a stagnant, objectified ideology, and instead embraced culture as a dynamic political and economic process in the development of human relationships.

Theodor Adorno (1903–69) explored how major culture industries used popular culture—such as film and radio—to advance the oppressive ideologies of the dominant, capitalist society. Through Adorno’s work, critical theory grew to question the dehumanizing relationship between media culture and its consumers, while simultaneously recognizing the power of using new media literacies against such control. In *One Dimensional Man* (Marcuse, 1964), Herbert Marcuse (1898–1979) critiques industrialism’s oppression of individuals for the maintenance of systems of production. Through this work, Marcuse pushed critical theory to demonstrate how a system of production controls individual consumption. In other work, Marcuse also explain elucidate the purpose of dialectical thought as a connection between thought and action in *Reason and Revolution* (Marcuse, 1960), where he writes how “Thought ‘corresponds’ to reality only as it transforms reality by comprehending its contradictory ... nature ... to break down the self-assurance and self-contentment of common sense, to undermine the sinister confidence in the power and language of facts.” Max Horkheimer’s (1895–1973) work focused primarily on the function of reason within philosophy and the creation of oppressive ideologies in larger society particularly significant during a period in which the Nazis were using evil and dehumanizing methods to manipulate society to engage in the mass murder and torture of Jewish, LGBTQ, and other minoritized people. Horkheimer used critical theory as a tool to fight destructive mass ideologies in a rapidly changing world. Walter Benjamin (1892–1940) engaged in critiques of literary and artistic work—building on aesthetic theory while pushing critical theory to discuss the realm of culture and its historical context. Jürgen Habermas (born in 1929) contributed to critical theory in his development of the theory of communicative reason or communicative rationality through discourse ethics and idea exchange. Horkheimer defines a truly critical theory as one that “has as its object human beings as producers of their own historical form of life” (Horkheimer, 1993: 21). Thus, critical theory gives people the agency to challenge taken-for-granted assumptions through tools such as dialectical reasoning and praxis that promote free-thinking and self-determination for all individuals in society.

Critical theory involves engagement in dialectic and critical praxis toward critical consciousness. Critical consciousness involves thinking about power through the analysis of oppression in the world/society/community, identifying one’s identity and naming their role in oppression and engaging in individual and collective social action toward more just practices and policies (Nieto and McDonough, 2011). From this perspective, oppression often relates to the current function of school and other institutions (e.g., family, criminal justice, health care, places of worship) to reproduce or challenge that oppression and hierarchies based on race, social class, and other demographic categories. Critical consciousness requires socio-cultural-political competencies and practices for educational researchers that include political clarity (Bartolomé, 2004) and critical social consciousness (Grant and Agosto, 2008).

Global critical philosophies

Although critical theory itself was born of a Western and European philosophical history, other influential philosophies have also developed (and often preceded) worldwide that we should recognize as part of the critical theory canon. Some such traditions have borrowed from Western philosophies while others have developed more independently and, while not all such work is recognized by North American and European scholars due to language barriers, racism, and imbalanced power relationships between Western imperialists and scholars of other nations, in this section we aim to bring some of these global critical philosophies to the mainstream of Eurocentric critical thought.

African philosophical tradition is rich and varied, with central and longstanding tenets of communalism, justice, and care informing tenets of critical western philosophers like Marx and Horkheimer by offering a vision for interrogating knowledge, power, and the role of individuals working toward liberation through collective social action (Crawford et al., 2009). A critical approach to understanding a communalistic ethos inextricably connects the individual and the larger community and is extrapolated in African socialism by leaders like Julius Nyerere in Tanzania, Kwame Nkrumah in Ghana, among others (Senghor, 1964). Fu-Kiau (2001) points to myriad examples of traditional African communalistic societies where “[e]verybody felt responsible to everybody else in the community and its neighborhood. When a community member suffered, it was the community as a whole that suffered” (p. 50). This communalistic ethos aligns with Horkheimer’s criteria for critical praxis, including a focus on being: (1) explanatory (how people should organize communities), (2) practical (what individuals do in that community), and (3) normative (traditional mores for equitable distribution of power). The ethos of communalism is well-documented in the customs and proverbs of Africa and her diaspora (Ikuenobe, 2006). For example, the proverbs of the Akan people of Ghana, West Africa, describe the interdependence of the individual and collective members of the community where individualism harms not only the community but the individual who receives the benefit: “Obiako di’wo a, etoa ne yam” (“If one alone eats the honey, it plagues his stomach”) (Danquah, 1968, 195).

However, across the continent agents of European colonialism attacked African communalism and justice. Frantz Fanon (1925–61) was a critical and transformative intellectual in the Algerian resistance to such colonialism experienced under French rule. Believed to be influenced by the 1930s Négritude movement—that rejected French colonial racism and was led by Martinican poet Aimé Césaire, Guianan Léon Damas, and future Senegalese President Léopold Sédar Senghor—Fanon explored how the psychology of oppression and the impact of colonialism affected the mental health of colonized people in works like *A Dying Colonialism* (Fanon, 1959) or *Black Skin, White Masks* (Fanon, 1967). By critiquing the way colonial powers control colonial ideology production and dissemination and how media is used to manipulate colonized people into believing in Western racial hierarchies of language, custom, or skin color, Fanon illustrates how indigenous peoples can deconstruct dominant language and texts toward

ideology change for liberation. Fanon's work highly influenced the revolutionaries of other nations—most notably Fidel Castro in Cuba, Paulo Freire in Brazil, and the Black Panthers in the United States.

Former Teacher, Tanzanian President, and African Nationalist and Socialist, Julius Nyerere argued that in the communalistic society, the “educated know themselves to be an integral part of the nation and recognize the responsibility to give greater service [commensurate with] the greater opportunities they have had” (Nyerere, 1967, p. 28). The ability of the educated to serve “the good” of the community is directly related to the type of education they have received. Education has to liberate both the mind and the body of man. It has to make him more of a human being because he is aware of his potential as a human being, and is in a positive, life-enhancing relationship with himself, his neighbor, and his environment. Education must therefore enable people to throw off the impediments to freedom which restrict their full physical and mental development (Nyerere, 1974, p. 3). Thus, the purpose of critical education is relational and contextual. Education has a duty to liberate where schooling is neither fair nor just and the society does not present youth with equal opportunities to learn and to form positive relationships with self, community, and society.

The communalistic justice and care model of achieving “the good” centers on the critical practice of simultaneously considering the needs of the individual and the community (Ekei, 2001; Ikuenobe, 2006). In fact, Walker et al. (2004), point to a dialectic ethical space where “[T]he delineation of self-care and community care into isolated categories fashions a false dichotomy” (p. 11). Within the African Diaspora, critical theorists use communalism to re-imagine liberation in communities and education practices. W.E.B. DuBois (1868–1963) in his famous work—*The Souls of Black Folk* (DuBois, 1903) explores Black individuals and communities' experience in America. DuBois theorizes about “double consciousness” experienced in the African American who sees herself/himself through the gaze of those who hate people of color, and describes the torment this places on the spirits of all people in a racist society. DuBois recognizes how ideologies of race are formed through dominant languages, but how a critical approach to education could help transcend such oppressive ideologies and lead to social change. Following the Harlem Renaissance and the growth of transformative art and literary works that helped engage humanity in a critique of America's anti-Black social structures and institutions, Carter G. Woodson (1875–1950) published *The Miseducation of the Negro* (Woodson, 1933) which challenged the ways classical education incorporated forms of scientific racism by revising world histories to enforce ideas about the racial inferiority of people of color. By removing African Americans and African peoples from school knowledge regarding history and science, classical education dehumanized African Americans and continued to oppress and mold them to the racist, hierarchical American fabric. Woodson calls for a critical education in which people of color could question the schooling they receive for individual and community liberation. Such works not only influenced innumerable educators, scholars, artists, and activists toward a critique of society for social change as is visible in the work of Frederick Douglass, Langston Hughes, Marcus Garvey, Malcolm X, and the Black Panther Movement—all of whom engaged in challenging how knowledge is produced and used to manipulate truth in a racially hierarchical society.

During the 1980s, other scholars—such as Edward Said and Homi Bhabha—pushed critical theorists to think in new ways by developing studies in colonialism and postcolonialism. Recognizing how Western colonialism maintained its hegemony through the control of knowledge and knowledge production, postcolonial theorists engage critical theory in ways most applicable to the current reality of globalization. Although often associated with postmodern theory in its method under the influence of Jacques Derrida's concepts of deconstruction, Jacques Lacan's psychoanalysis, and Michel Foucault's discourse analysis, postcolonial theory still relates to critical theory in how colonized, indigenous peoples have found an agency in critiquing the oppressive societies in which they live. In *Orientalism* (Said, 1978), Said critiques how the West continues to romanticize Asian and Middle Eastern nations as the archaic and unchanging “other”, while using such false ideologies to justify continued imperialism throughout the Orient. Similarly, in *Nation and Narration* (Bhabha, 1990), Bhabha problematizes how postcolonial countries continue to be treated as one, homogeneous group by their ex-colonizers, with concepts of nationhood narrativized through power imbalances between ex-colonizing nations and their ex-colonial states.

Of course, Latin America and Central American development of new critical theories can be traced through the work of Liberation Theology that united indigenous people with Catholic clergy against the oppressive, racist structures of the European Catholic church, or in the work of Ernesto Ché Guevara—Minister of Education in post-revolution Cuba and organizer of the guerrilla mountain schools—but the most notable development of critical theory in education can be found in the work of Paulo Freire (1921–97). Freire, a former teacher like his Black contemporaries Milton Santos and Abdias Nascimento, form part of a tradition of critical activist scholars that extend Western Critical Theories including Marxism and the Frankfurt School to interrogate oppression and imagine more just urban spaces, popular education, and cultural production (Almada, 2013; Santos, 2000; Silva, 2010). For example, Freire's most famous work—*The Pedagogy of the Oppressed*—explores the ways the educational system proves central to the dehumanizing processes of capitalist and neocolonial societies through banking education that treats students as empty vessels that passively collect the oppressive norms and knowledge of teachers. Freire focused on literacy and literacy education while teaching illiterate adults how to interrogate and deconstruct dominant texts within a racist, classist Brazil. Freire believed in the power of praxis and dialog to push critical thought against oppressive forces that control schooling structures. Critical Black theorists in Brazil like Oliveira and Ferreira (2020) have also drawn on African philosophy in dialog with Latin American philosophies of liberation to create a critical ancestral philosophy that is a communal construct respecting wisdom of elders in their experiences toward liberation and the social good in the community that can guide Afro-Brazilians enacting racial and social justice in education and social movements (Oliveira, 2018).

While critical theory has surely influenced and been influenced by revolutionaries, scholars, and activists throughout East Asia, the continued imbalance of power relationships that maintain hierarchies regarding race, class, color, and language in the Western

world continue to negate and oppress East Asian critical scholarship. From our limited ability to explore what is prioritized in English-language academic sources, important dialog has taken place around the concept of second modernity (Beck, 2008; Beck et al., 1995; Beck and Grande, 2010) and capitalist expansion within the context of Bureaucratic Authoritarian states in locations such as Korea or Japan (See, for example, Han, 1997, 2007, 2009). But without translation and prioritization of Eastern knowledge in Western academia, much is being lost in current explorations of critical theory. Critical theorists should immediately address this issue of exclusion of non-western philosophers.

Returning to the West and to the United States, critical theory has been expanded upon, questioned, and integrated in important ways through critical Feminist Theories, Black Feminist Theory and Critical Race Theory that can powerfully inform qualitative research methodologies. Within Feminist Theories, Feminist Standpoint Theorists (Harding, 1991; Hartsock, 2003) have emphasized the need to produce knowledge grounded within the experiences of those who have undergone collective political struggle and been most marginalized from institutional knowledge creation processes. Thus, starting from the lived experiences of women, for example, can prevent research that separates women from their realities toward developing more relevant understandings of how power is constructed and maintained by patriarchal structures. Black Feminist Theory—rooted in Black Feminism that has existed since the time of slavery in the US—centers Black women’s experiences and perspectives, identifying the primacy of their contributions to our world from unique experiences at the intersections of race, gender, class, and sexual identity. Valuing the intersectionality (Crenshaw, 2017) of Black female identity, critical Black Feminist Theory offers much needed analyses of how patriarchy and white supremacy create the oppressive institutions and contexts within which current society operates (see Collins, 2019). Critical Race Theory—building off of early 1980s critical legal studies that centered race in its analyses of law and institutional structures—and LatCrit Theory allow us to examine and challenge the epistemological and methodological foundations of educational research as well as analyze educational institutions and practices through the lens of race and racism (see Solorzano and Yosso, 2000; Solorzano and Bernal, 2001). CRT advances storytelling and counter-storytelling as important qualitative methodologies with which to explore issues of white supremacy, power, and oppression (Solorzano and Yosso, 2001).

Critical theory in qualitative educational research

In qualitative educational research, critical praxis is the reassertion of research design, implementation, and distribution for a more human world at the individual and social levels, where the simultaneous changing of circumstances and self-change occur through self-reflection, reflective action, and collective reflective action (Freire, 1970). Critical qualitative education researchers enact critical praxis to support students through the co-construction of educational practices that allow people to acquire, analyze, and produce both social and self-knowledge (Ryoo et al., 2009).

What makes research “critical”?

Critical theorists who conduct qualitative research in education share similar ways of knowing and challenging inequitable distribution of power and privileged or marginalized ways of being within social relations. Marx (1969), in his 11th Thesis on Feuerbach reminds us of the epistemological and ontological telos of a critical knowledge production is political action, “Philosophers have hitherto only interpreted the world in various ways; the point is to change it.” This statement reflects the shared value that critical theorists conducting qualitative research share of critique of oppression indexed in current social relations and transformative social actions. While the type and amount of change depends on a critical theorist’s orientation, all educational researchers employing critical theory concur that research, knowledge production, and educational practices are political (Chun and Morgan, 2019). Zinn (2018) popularized this position when he said that, “you can’t be neutral on a moving train.” If a qualitative researcher is not actively contesting oppression through anti-racist and anti-capitalist research practices, then they are perpetuating white supremacy and economic exploitation. All phases of qualitative research design, data collection and analysis, and knowledge dissemination benefit from critical consciousness and critical praxis (Kincheloe and McLaren, 2011). To transform systems, critical theorists in qualitative education research engage in iterative processes of critical consciousness and critical praxis (as described above).

Contemporary critical accounts of schools in the United States

Given the current structure of public primary and secondary schools in the United States, a commodification of learning and the learner often occurs when educational researchers function as dutiful knowledge producers carrying out mandates generated from the dominant elite and their funding sources with little input from stakeholders in schools (Kincheloe and McLaren, 2011; McLaren and Giarelli, 1995). These mandates include implementing scripted-curricula in overcrowded classrooms and administering high-stakes assessments and zero tolerance discipline policies that bolster opportunity gaps and cause students to be compliant “objects” during the learning process (Apple, 2004). Not coincidentally, these educational practices are widely supported by money interests and the state (Marx, 2019; see also Marx, 1972).

Lave and McDermott (2002) illustrate further that, “Just as the result of alienated labor is embodied in the things produced, so the object of alienated learning becomes material in the things learned The learner becomes all the poorer the more he becomes subject to the whim of the educational system” (p. 34). In this way the object of alienated research becomes material in the

knowledge produced, whereas the researcher becomes poorer the more they are subject to conducting research that does not support critical consciousness or praxis that challenges scripted curricula and standardized assessments and a host of other institutional practices that support the status quo of racial and class oppression in schools.

However, there is a long history of activists and critical researchers who have been able to negotiate structural and institutional constraints in schools through their writing and research while bringing about radical changes in society (Woodson, 1933; DuBois, 1903; Douglass, 2018). Critical educational researchers recognize how traditional schooling works against the interests of those students who are most vulnerable in society by reproducing class differences and racialized inequality (McLaren, 2006; Giroux, 1983; Shor, 1992; Darder, 2011). Critical qualitative educational researchers are committed to creating an emancipatory culture of schooling that counters alienation. Through critical consciousness and praxis, critical qualitative researchers aim to resolve the tensions inherent in researching schooling against the status quo amid educational and research institutions that support it.

Applications of critical theory in qualitative educational research

Kincheloe and McLaren (2011) argue that the application of critical theory in qualitative educational research centers on a transformative praxis that works to understand and alleviate oppression. Critical qualitative educational researchers move beyond empirical descriptions of “what is?” to ask, “how what is has come to be, whose interests are served by particular institutional arrangements, and where our own frames of reference come from” in order to enact “critical world making” with others (p. 315). While an exhaustive review of all educational qualitative research that interrogates injustice and aims toward transformative praxis is beyond this brief chapter, in this section we will review key methodological approaches and examples of qualitative education studies informed by critical theory in the United States. These methods grow from critical ethnography and include duoethnography, counter-storytelling, and participatory action research.

Critical ethnography

Beginning in the 1930s, community sociologists and in the 1960s action anthropologists contributed to critiques of racial and class inequalities with research narratives of oppressed communities that were more asset-based rather than deficit oriented (Vidich and Lyman, 2000). One contemporary design sociologist, Hugh Mehan, extended critiques of inequity in public education to action by co-creating a public school based on equitable policies (like detracking and community engagement) and rigorous and inclusive teaching (Mehan, 2008). More recent activist anthropologists expand collaborative research design and execution with multimodal mediums of dissemination of their studies (Foley and Valenzuela, 2004; see special issue of the *Anthropology and Education Quarterly* 2008). This approach was used by Bowles and Gintis (1976) in their seminal study of the social reproductive function of school for working class students and by Valenzuela (1999) to look at the way high school assimilation policies and practices takes away US–Mexican youths social capital in a public high school. Three examples of specific approaches aligned with a critical theory that follow the basic tenets of critical ethnography are duoethnography, counter story-telling, and participatory action research.

Duoethnography

Duoethnographers enact social justice by examining power, privilege and improving conditions of teaching and learning. The process of duoethnography involves the following six iterative steps: (1) select a phenomenon, (2) analyze the phenomenon, (3) engage in abduction (explanatory reasoning), (4) situate the data, (5) critique that data, (6) develop an insight (Sawyer and Norris, 2013). Researchers analyze themselves as text in context in a recursive process grounded in trust and a commitment to question emergent common understandings through a dialog. This approach was used by Diversi and Moreira (2009) who worked with street children in Brazil, Sawyer and Liggett (2016) who explored postcolonial education, Koonce and Lewis (2020) who explored critical care in their university classrooms, and McDermott et al. (2012) who interrogated a curriculum of beauty. These approaches align with the telos of critical theory with explicit emphasis on praxis through a collective reflective action, or what Freire (1970) calls conscientization and Pinar (1975) describes as *currere*. This process calls for critical self-analysis and unlearning of internalized colonial, neocolonial, neoliberal and other oppressive discourses and narratives toward what Said (1993) calls the counterpunctual narrative in order to find a “new way to imagine cultural worlds with the tension of standardized and often neoliberal discourses” (Sawyer and Norris, 2013, p. 8).

Counter story-telling

Many critical race theorists conduct qualitative education research using counter story-telling. This approach to research draws on Critical Race Theory to unmask racialized experiences by showing how race organizes non-racial phenomena, policies, practices, and interactions. Researchers using counter story-telling methods begin with the assumption that race is a narrative with social and material consequences. Narrative has been important in constructing and deconstructing white supremacy throughout the history of the United States (Baszile, 2015) and is also particularly important to indigenous epistemologies (Brayboy, 2005). Race stories construct our actions and beliefs. The practice of counter storytelling reconstructs narratives from the views of the oppressed to arrive at a more complete understanding of oppression and social change (Solórzano and Yosso, 2002).

Counter story-telling studies have been used to study race and racism in education from the point of view of Black students (DeCuir and Dixon, 2004; Howard, 2008), students engaged in action research (Goessling, 2018; Hauber-Özer et al., 2021), parents and families of Black children in public schools (Lewis and Adkins-Sharif, 2020; Reynolds, 2010), Black K-12 teachers (Cook and Dixon, 2013) and school leaders (Parker and Villalpando, 2007), Latinx graduate students (Solórzano and Yosso, 2001), and Black

university faculty (Griffin et al., 2014). For example, Howard (2008) analyzed Black male's stories of their schooling experiences of teachers and school administrator's racialized and deficit perceptions of them and how they worked to "eradicate negative racial stereotypes" of Black males in schools (p. 954). Croom and Marsh (2016) share studies from university faculty and administrators who use counter storytelling to enact greater justice in school policies, labor practices, academic curriculum, and classroom pedagogy. These myriad counter story-telling qualitative studies across educational contexts share a commitment to Black educational officials, learners, and community to name their own reality, challenge individual and institutional racism, and re-imagine more racially just schools.

Participatory action research

Conducting qualitative educational research through critical inquiry to liberate learners and transform schools is central to participatory action research (PAR) designs (Kemmis et al., 2013). Participatory action research with a critical lens involves educators examining their own practice in collaboration with their students to question and interrogate structural issues of power and the inequitable distribution of social and material resources through solidarity work (Martínez-Roldán and Quiñones, 2016; Martinez, 2017).

One example of an emerging trend aligned with critical theory under the umbrella of PAR is Youth Participatory Action Research (YPAR) (Mirra et al., 2015). Critical qualitative researchers have explored YPAR as approaches to a pedagogy of resistance (Camarota and Fine, 2010), multicultural education (Irizarry, 2009), service learning (Schensul and Berg, 2004), civic engagement and advocacy (Kirshner, 2010), and policy-making (Warren and Marciano, 2018). A systematic review of 63 YPAR studies found that most youth who participated in these studies gained agency and leadership, as well as academic or career skills (Anyon et al., 2018). For example, Akom et al. (2016) describe how young researchers in Oakland, California used digital technology with local community members that "sparked a food revolution" by identifying food insecurity and working with schools and community leaders and organizations to improve access to healthy food in their neighborhood. These studies share the fact that they are community-based and involve youth in identifying and acting on social, economic, cultural and political issues with a goal of more just access, representation, experience, and policies.

Conclusion

Here we have explored the history and key ideas of critical theory, as well as its application to education and qualitative educational research. Critical theory in qualitative research supports the interrogation of oppression, the analysis of power structures, and ultimately an engagement in social action toward creating more just learning contexts and communities. Qualitative educational researchers apply critical theory through interrogating (a dialogical reasoning of) the relationship between ways of knowing and being, oppressive schooling and educational practices, and unjust social relations in schools and other social institutions. Critical qualitative educational researchers critique exploitation in schools and work collectively to assert teachers', students' and communities' humanity through research that often includes social action. As the field continues to grow, critical qualitative educational researchers in the United States should prioritize Africa, Asian, and Latin America (among other) critical philosophies, research methods and general approaches to knowledge production. Qualitative educational researchers applying critical theories will continue to strive for solidarity with local and global communities of researchers, educators and learners to critique exploitation and humanize classrooms, schools, and societies.

References

- Akom, A., Shah, A., Nakai, A., Cruz, T., 2016. Youth participatory action research (YPAR) 2.0: how technological innovation and digital organizing sparked a food revolution in East Oakland. *Int. J. Qual. Stud. Educ.* 29 (10), 1287–1307.
- Almada, S., 2013. *Abdias Nascimento*. Selo Negro, São Paulo.
- Anyon, Y., Bender, K., Kennedy, H., Dechants, J., 2018. A systematic review of Youth Participatory Action Research (YPAR) in the United States: methodologies, youth outcomes, and future directions. *Health Educ. Behav.* 45 (6), 865–878.
- Apple, M.W., 2004. Creating difference: neo-liberalism, neo-conservatism and the politics of educational reform. *Educ. Pol.* 18 (1), 12–44.
- Bartolomé, L., 2004. Critical pedagogy and teacher education: radicalizing prospective teachers. *Teach. Educ. Q.* 31 (1), 97–122.
- Baszile, D.T., 2015. Rhetorical revolution: critical race counterstorytelling and the abolition of white democracy. *Qual. Inq.* 21 (3), 239–249.
- Beck, U., 2008. *Critical Theory of World Risk Society*. SNU Public Lecture.
- Beck, U., Grande, E., 2010. Varieties of second modernity: the cosmopolitan turn in social and political theory and research. *Br. J. Sociol.* 61 (3), 409–443.
- Beck, U., Giddens, A., Lash, S., 1995. *Reflexive Modernization*. Polity, Cambridge.
- Bhabha, H. (Ed.), 1990. *Nation and Narration*. Routledge, New York.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America*. Basic Books, New York.
- Brayboy, B.M.K.J., 2005. Toward a tribal critical race theory in education. *Urban Rev.* 37 (5), 425–446.
- Camarota, J., Fine, M., 2010. Youth participatory action research: a pedagogy for transformational resistance. In: *Revolutionizing Education*. Routledge, New York, pp. 9–20.
- Chun, C.W., Morgan, B., 2019. Critical research in English language teaching. In: *Second Handbook of English Language Teaching*, pp. 1091–1110.
- Collins, P.H., 2019. *Intersectionality as Critical Social Theory*. Duke University Press.
- Cook, D.A., Dixon, A.D., 2013. Writing critical race theory and method: a composite counterstory on the experiences of Black teachers in New Orleans post-Katrina. *Int. J. Qual. Stud. Educ.* 26 (10), 1238–1258.

- Crawford, J., Gyamfi, N., McLaren, P., 2009. Unplugging the stomach: curing the University of California admissions policy with an ethic of communal justice and care. In: *Education and Hope in Troubled Times*. Routledge, New York, pp. 257–274.
- Crenshaw, K.W., 2017. *On Intersectionality: Essential Writings*. The New Press, New York.
- Croom, N.N., Marsh, T.E. (Eds.), 2016. *Envisioning Critical Race Praxis in Higher Education Through Counter Storytelling*. IAP, Charlotte.
- Danquah, J.B., 1968. *The Akan Doctrine of God: A Fragment of Gold Coast Ethics and Religion*, second ed. Cass, London.
- Darder, A., 2011. *A Dissident Voice: Essays on Culture, Pedagogy, and Power*. Peter Lang, New York.
- DeCuir, J.T., Dixon, A.D., 2004. "So when it comes out, they aren't that surprised that it is there": using critical race theory as a tool of analysis of race and racism in education. *Educ. Res.* 33 (5), 26–31.
- Diversi, M., Moreira, C., 2009. *Between Talk: Decolonizing Knowledge Production, Pedagogy, and Praxis*. Left Coast Press, Walnut Creek, CA.
- Douglass, F., 2018. *Frederick Douglass Papers*. Yale University Press, New Haven.
- DuBois, W.E.B., 1903. *The Souls of Black Folks*. Fawcett, Greenwich, CT.
- Ekei, J.C., 2001. *Justice in Communalism: A Foundation of Ethics in African Philosophy*. Realm Communications, Lagos, Nigeria.
- Fanon, F., 1959. *A Dying Colonialism*. Grove Press, New York.
- Fanon, F., 1967. *Black Skin, White Masks*. Grove Press, New York.
- Foley, D., Valenzuela, A., 2004. Critical ethnography: the politics of collaboration. In: Denzin, N., Lincoln, Y. (Eds.), *The Handbook for Ethnography*. Sage Publications, Thousand Oaks, CA, pp. 217–234.
- Freire, P., 1970. *Pedagogy of the Oppressed*. Continuum, New York.
- Fu-Kiau, K.K.B., 2001. *African Cosmology of the Bantu-Kongo: Tying the Spiritual Knot: Principles of Life & Living*, second ed. Athelia Henrietta Press, Brooklyn, NY.
- Giroux, H., 1983. *Theory and Resistance in Education*. Bergin and Garvey, South Hadley, MA.
- Goessling, K.P., 2018. Increasing the depth of field: critical race theory and photovoice as counter storytelling praxis. *Urban Rev.* 50 (4), 648–674.
- Grant, C.A., Agosto, V., 2008. Teacher capacity and social justice in teacher education. In: Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D.J., Demers, K.E. (Eds.), *Handbook of Research on Teacher Education: Enduring Questions and Changing Contexts*, third ed. Routledge, pp. 175–200.
- Griffin, R.A., Ward, L., Phillips, A.R., 2014. Still flies in buttermilk: Black male faculty, critical race theory, and composite counter storytelling. *Int. J. Qual. Stud. Educ.* 27 (10), 1354–1375.
- Han, S., 1997. The political economy and moral institutions: the formation of the middling grassroots in Korea. *Humboldt J. Soc. Relat.* 23 (1 and 2), 71–89.
- Han, S., 2007. Paradoxical modernity and the quest for a neo-communitarian alternative. *Dev. Soc.* 36 (1), 103–128.
- Han, S., 2009. Social transformation in contemporary Korea: three prime movers in a contested civil society. In: Giri, A. (Ed.), *The Modern Prince and the Modern Sage: Transforming Power*. Sage, London.
- Harding, S., 1991. *Whose Science/Whose Knowledge?* Open University Press, Milton Keynes.
- Hartsock, N., 2003. The feminist standpoint: developing the ground for a specifically historical materialism. In: Harding, S., Hintikka, M.B. (Eds.), *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, second ed. Kluwer Academic Publishers, Boston.
- Hauber-Özer, M., Call-Cummings, M., Hassell-Goodman, S., Chan, E., 2021. Counter-storytelling: toward a critical race praxis for participatory action research. *Int. J. Qual. Stud. Educ.* 1–16.
- Horkheimer, M., 1972. *Critical Theory*. Seabury Press, New York.
- Horkheimer, M., 1982. *Critical Theory: Selected Essays*. Continuum, New York.
- Horkheimer, M., 1993. *Between Philosophy and Social Science*. MIT Press, Cambridge.
- Howard, T., 2008. Who really cares? The disenfranchisement of African American males in preK-12 schools: a critical race theory perspective. *Teach. Coll. Rec.* 110 (5), 954–985.
- Ikuenobe, P., 2006. *Philosophical Perspectives on Communalism and Morality in African Traditions*. Lexington Books, Lanham, MD.
- Irizarry, J.G., 2009. Reinventing multicultural education through youth participatory action research. *Multicult. Perspect.* 11 (4), 194–199.
- Kemmis, S., McTaggart, R., Nixon, R., 2013. *The Action Research Planner: Doing Critical Participatory Action Research*. Springer Science & Business Media, Berlin.
- Kincheloe, J.L., McLaren, P., 2011. Rethinking critical theory and qualitative research. In: *Key Works in Critical Pedagogy*. Brill, Sense, pp. 285–326.
- Kirshner, B., 2010. Productive tensions in youth participatory action research. *Yearbk. Natl. Soc. Stud. Educ.* 109 (1), 238–251.
- Koonce, J.B., Lewis, K.A., 2020. Culturally relevant care through the lens of duoethnography. *Qual. Rep.* 25 (6), 1721A–1735A.
- Lave, J., McDermott, R., 2002. Estranged labor learning. *Outlines Crit. Pract. Stud.* 4 (1), 19–48.
- Lewis, N.D., Adkins-Sharif, J., 2020. Allying and leading through racial realities. *J. Educ. Human Resour.* 38 (2), 194–212.
- Marcuse, H., 1960. *Reason and Revolution*. Beacon Press, Boston, MA.
- Marcuse, H., 1964. *One-Dimensional Man*. Beacon Press, Boston.
- Martinez, D.C., 2017. Imagining a language of solidarity for Black and Latinx youth in English language arts classrooms. *Engl. Educ.* 49 (2), 179–196.
- Martinez-Roldán, C.M., Quiñones, S., 2016. Resisting erasure and developing networks of solidarity: testimonios of two Puerto Rican scholars in the academy. *J. Lang. Ident. Educ.* 15 (3), 151–164.
- Marx, K., 1969. *Theses on Feuerbach, 1888*. Retrieved from: <https://www.marxists.org/archive/marx/works/1845/theses/theses.htm>.
- Marx, K., 1972. *Critique of the Gotha Programme (1875)*. People's Publishing House, Peking.
- Marx, K., 2019. *Capital: Volume One*. Courier Dover Publications, Mineola.
- Marx, K., Engels, F., 1998. *The Communist Manifesto: A Modern Edition*. Verso, Brooklyn.
- McDermott, M., Shelton, N.R., Mogge, S.G., 2012. Preservice teachers' perceptions of immigrants and possibilities of transformative pedagogy: recommendations for a praxis of "critical aesthetics". *Int. J. Multicult. Educ.* 14 (2).
- McLaren, P. (Ed.), 2006. *Rage + Hope: Interviews With Peter McLaren on War, Imperialism, + Critical Pedagogy*, vol. 295. Bern, Peter Lang.
- McLaren, P., Giarelli, J.M. (Eds.), 1995. *Critical Theory and Educational Research*. SUNY Press, New York.
- Mehan, H., 2008. Engaging the sociological imagination: my journey into design research and public sociology. *Anthropol. Educ. Q.* 39 (1), 77–91.
- Mirra, N., Garcia, A., Morrell, E., 2015. *Doing Youth Participatory Action Research: Transforming Inquiry With Researchers, Educators, and Students*. Routledge, New York.
- Nieto, S., McDonough, K., 2011. Placing equity front and center revisited. In: *Studying Diversity in Teacher Education*, pp. 363–384.
- Nyerere, J.K., 1967. *Education for Self-Reliance*. Government Information Services, Dar es Salaam.
- Nyerere, J.K., 1974. *Education Must Liberate Man: President Nyerere's Speech to the Dag Hammarskjöld Seminar on Education and Training, 1974*. Government Information Services, Dar es Salaam.
- Oliveira, E.D.D., 2018. *Filosofia da ancestralidade: corpo e mito na filosofia da educação brasileira*. Tese (Doutorado) - Universidade Federal do Ceará, Programa de Pós-graduação em Educação Brasileira. Fortaleza (CE), p. 2005.
- Oliveira, E., Ferreira, L.C., 2020. Filosofia da Ancestralidade: diálogo Entre a Filosofia da Libertação e a Filosofia Africana. *Revista Ideação* 1 (41), 219–253.
- Parker, L., Villalpando, O., 2007. A racialized perspective on education leadership: critical race theory in educational administration. *Educ. Adm. Q.* 43 (5), 519–524.
- Pinar, W., 1975. Currere: toward reconceptualization. In: *Curriculum Theorizing: The Reconceptualists*, pp. 396–414.
- Reynolds, R., 2010. "They think you're lazy," and other messages Black parents send their Black sons: an exploration of critical race theory in the examination of educational outcomes for Black males. *J. Afr. Am. Males Educ.* 1 (2), 144–163.
- Ryoo, J.J., Crawford, J., Moreno, D., McLaren, P., 2009. Critical spiritual pedagogy: reclaiming humanity through a pedagogy of integrity, community, and love. *Power Educ.* 1 (1), 132–146.

- Said, E., 1978. *Orientalism*. Vintage, New York.
- Said, E., 1993. *Culture and Imperialism*. Knopf, New York.
- Santos, M., 2000. Por uma outra globalização: Do pensamento único à consciência universal. Record, Rio de Janeiro.
- Sawyer, R.D., Liggett, T., 2016. Postcolonial education: using a duoethnographic lens to explore a personal curriculum of post/decolonization. In: *Duoethnography*. Routledge, New York, pp. 71–88.
- Sawyer, R.D., Norris, J., 2013. *Duoethnography*. Oxford University Press, Oxford.
- Schensul, J.J., Berg, M., 2004. Youth participatory action research: a transformative approach to service-learning. *Mich. J. Community Serv. Learn.* 10 (3), 76–88.
- Senghor, L.S., 1964. On African Socialism. Trans. Mercer Cook. Pall Mall, London.
- Shor, I., 1992. *Culture Wars: School and Society in the Conservative Restoration*. University of Chicago Press, Chicago.
- Silva, R.F.D., 2010. Trajetórias de dois intelectuais negros brasileiros: Abdias Nascimento e Milton Santos. Retrieved from. <http://www.repositorio.ufba.br/ri/handle/ri/8585>.
- Solórzano, D.G., Bernal, D.D., 2001. Examining transformational resistance through a critical race and LatCrit theory framework: Chicana and Chicano students in an urban context. *Urban Educ.* 36 (3), 308–342.
- Solórzano, D.G., Yosso, T.J., 2000. Maintaining social justice hopes within academic realities: a Freirean approach to critical race/LatCrit pedagogy. *Denv. UL Rev.* 78, 595.
- Solórzano, D.G., Yosso, T.J., 2001. Critical race and LatCrit theory and method: counter-storytelling. *Int. J. Qual. Stud. Educ.* 14 (4), 471–495.
- Solórzano, D.G., Yosso, T.J., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1), 23–44.
- Valenzuela, A., 1999. *Subtractive Schooling: US–Mexican Youth and the Politics of Caring*. State University of New York Press, Albany.
- Vidich, A.J., Lyman, S.M., 2000. Qualitative methods: their history in sociology and anthropology. *Handb. Qual. Res.* 2, 37–84.
- Walker, S., Vanessa, S., Snarey, J.R., 2004. *Race-ing Moral Formation: African American Perspectives on Care and Justice*. Teachers College Press, New York.
- Warren, C.A., Marciano, J.E., 2018. Activating student voice through Youth Participatory Action Research (YPAR): policy-making that strengthens urban education reform. *Int. J. Qual. Stud. Educ.* 31 (8), 684–707.
- Woodson, C.G., 1933/1990. *The Miseducation of the Negro*. Africa World Press, Washington, DC.
- Zinn, H., 2018. *You Can't Be Neutral on a Moving Train: A Personal History*. Beacon Press, Boston.

Relevant websites

- Critical Theory Institute, University of California, Irvine. <http://www.humanities.uci.edu>.
- Illuminations, The University of Texas at Arlington. <http://www.uta.edu>.
- Marxists Internet Archive. <http://www.marxists.org>.
- Young People Empowered to Change the World, University of California, Berkeley. <http://yparhub.berkeley.edu/>.

Feminist and critical approaches to gender in qualitative research

Channel C. McLewis^a, Justin A. Gutzwa^b, and S. Hurtado^a, ^a Department of Education, University of California, Los Angeles, CA, United States; and ^b Department of Physics and Astronomy, University of Utah, Salt Lake City, UT, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	55
Adopting a transformative paradigm	55
Disrupting dominant ideologies in qualitative research	56
Feminist epistemologies and theories	57
Centering women of color feminisms	57
The role of feminist and critical social theories	58
Methodology, methods, and analysis	59
Methodologies	59
Methods	59
Data analysis and interpretation	60
Positionality and reflexivity	60
The power of positionality in praxis	61
McLewis' positionality applied: othersisters in feminist research	62
Gutzwa's positionality applied: emic and etic flux in the field	62
Conclusion	63
References	64

Introduction

While qualitative research that centers gender in education can be conducted using commonly-used paradigms, including postpositivist, social constructivist, and pragmatic paradigms, much of those approaches are conducted without regard to existing power dynamics or concern for liberatory or emancipatory aims for participants (Creswell and Plano Clark, 2017; Hurtado, 2015). This chapter focuses on those methodologies guided by transformative paradigmatic approaches that hold potential for the researcher to become an “instrument of social change” (Merten, 2009, p. 11), empowering marginalized communities and seeking social justice as a key aim for participants. Influenced by social and intellectual movements, researchers became frustrated with existing paradigms and research methods that ignored or obscured the power dimensions that influence the lives of participants and their contexts, which privileged researchers’ epistemologies while undervaluing participants’ voice, active resistance, and agency.

With a focus on the perspectives of scholars conducting research with underserved and marginalized communities, this chapter illuminates how critical and feminist theories and methodologies enrich qualitative education research. Because feminist research approaches are comprehensively captured in other volumes (Hesse-Biber, 2007, 2014), we offer a brief overview of feminist research within the context of critical social theories and provide select examples that demonstrate how critical and feminist theories and methodologies are used in education research. Next, we share personal researcher accounts on conducting qualitative research in education with gender identity and intersections with racial/ethnic groups. We particularly focus on the roles of positionality and reflexivity as insights into the complexities and possibilities to advance a social justice agenda and perform transformative educational research. In doing so, we provide insights into how conducting gender identity research in education can also be transformative for researchers as they become more conscious of their role in relation to marginalized communities.

Adopting a transformative paradigm

Recognizing the persistent educational and societal inequities, Rios-Aguilar and Neri (2021) call for “educational research that contributes to revitalized critique and transformative pro-action against educational reproduction of existing inequalities” (p. 2). In working toward proactive solutions, the key elements of the transformative approach indicate a concern for social justice and the welfare of marginalized communities that permeate the entire research process. The transformative paradigm is a “metaphysical” umbrella and framework for basic beliefs, epistemologies, and research practices that recognize marginalized identities and are guided by critical theory, feminist theory, critical race theory, queer theory, decolonial, and culturally responsive approaches (Merton, 2009, p. 13). It is guided by ethical concerns that show respect for cultural norms and conditions, rejects deficit assumptions, and identifies participants’ strengths and critical resilience. Researchers working in this paradigm recognize the co-existence of multiple realities shaped by social, political, gender, racial, and economic contexts and conditions that reveal unequal power relationships. They recognize epistemological differences and similarities that shape their role in relation to participants, requiring the development of a trusting relationship. Merten (2009) indicates that the transformative paradigm requires qualitative methods (although mixed methods can also be used) to address cultural complexity. Using qualitative methods with a critical perspective

also accounts for contexts and histories that involve multiple forms of oppression, remaining cognizant of dominant perspectives that permeate methods and decisions that can adversely affect participants and social justice aims.

Advancing a transformative paradigm requires the use of approaches that challenge dominant ideologies and traditional ways and methods of knowing. Critical social theories employ culturally situated frameworks and intentionally address systems of oppression and power. The interrogation of power and subordination is not only reflected in critiques of the social organization of society but also in knowledge structures. Cynthia Dillard (2000) explains that the construction of knowledge is not neutral, “all research is social construction and a cultural endeavor,” (p. 662) in which she underscores how research is shaped by and shapes ideologies. For example, ideologies such as whiteness, racism, and patriarchy are rendered normative and permeate traditional approaches to research (Allen et al., 2019; Bhattacharya, 2019; Evans-Winters and Esposito, 2019; García and Mayorga, 2018; Gillborn, 2010; Harding, 1987; Zuberi and Bonilla-Silva, 2008). Bhattacharya (2019) contends that the overwhelming representation of white faculty who teach qualitative research and the subsequent lack of qualitative methods training that interrogates whiteness demonstrates how “qualitative research is mostly celebrated through a form of whiteness” that itself “cannot generate any knowledge that is completely devoid of whiteness” (p. 113). Thus, while qualitative research is typically intended to be sensitive to participants’ experiences, traditionally accepted modes of knowledge production are inextricable from whiteness, patriarchy, and other forms of domination.

Disrupting dominant ideologies in qualitative research

The perpetuation of dominant ideologies in traditional qualitative research approaches and training (re)produces systems of subjugation and inequity. Fundamentally, research is a mechanism to create and disseminate knowledge that can shape social life, including influencing or hindering life opportunities. Such a process necessitates the transformation of the procedures entailed in research inquiry to disrupt the colonial norms of white supremacy, heteropatriarchy, and other intersecting systems of domination imbued within research practices (Dillard, 2000; Evans-Winters and Esposito, 2019; Smith, 2012). The role of ideology in research calls for attention to research practices, which include what is represented as knowledge, the process of knowledge production, and who is authorized to be bearers and producers of knowledge.

Critical social theories and frameworks emphasize how ontological and epistemological orientations inform theory, as well as methodology, methods, and analysis. Examples of critical social theories, frameworks, and methodologies that are attentive to oppressive social structures and/or empowerment for minoritized groups include community cultural wealth (Yosso, 2005), critical race theory in education (Gillborn and Ladson-Billings, 2010; Ladson-Billings and Tate, 1995; Lynn and Dixon, 2021; Patton, 2016; Solórzano, 1997), decolonizing and anticolonial methodologies (Bhattacharya, 2019; Patel, 2014; Smith, 2012), funds of knowledge (Kiyama and Rios-Aguilar, 2017), funds of identity (Esteban-Guitart, 2016), racial capitalism (Robinson, 2020), and racial formation theory (Omi and Winant, 2014). Though varied in ontological histories, epistemic standpoints, and approaches, a feminist inquiry is another example of research that critiques oppressive social structures and seeks the empowerment of participants (Collins, 2000; Harding, 1987; Naples and Gurr, 2014).

Feminist theories and methodologies have the potential to transform education research and praxis. “Feminist research positions gender as the categorical center of inquiry and the research process,” in which “feminist researchers use gender as a lens through which to focus on social issues” (Hesse-Biber, 2014, p. 3). While some employments of feminist thought solely center on the experiences of cisgender women through the inappropriate equation of biological sex and gender identity, such as trans exclusionary radical feminism (Gutzwa, 2021a), feminist research does not inherently reinforce narrow categorizations and conceptions of biological differences. Instead, Feminists of Color, such as Black Feminist Cynthia Dillard (2000), extend researchers’ and participants’ ontology and epistemology as grounded in community:

When we begin to move beyond race/ethnicity and gender as biological constructions to more culturally engaged explanations of being human, and when we seek to examine the origins of such knowledge constructions as to the very nature of how reality is known (its patterns of epistemology), we will find that what constitutes knowledge depends profoundly on the consensus and ethos of the community in which it is grounded (p. 661–662).

Similarly, Evans-Winters and Esposito (2019) explain how critical race feminist frameworks such as intersectionality are useful in critiquing social order and hierarchy:

It is through a shared understanding of the social world that women of color, the indigenous, racial and ethnic minoritized people, queer and non-gender conforming people, and the poor, especially, can draw upon intersectionality as an analytical tool in the examination of (and resistance to) systems of power, privilege, and domination. (p. 56)

Feminist research is a mode of analysis to identify and challenge the ways that hegemonic, patriarchal systems of domination impact communities broadly. Feminists of Color point out that this challenge of pervasive patriarchal violence also critiques how reliance on narratives of whiteness excludes the realities of Black, Indigenous, and other People of Color (Lorde, 1984a). For example, Dillard (2000) contends that endarkened feminists engage in a “social critique” that “is focused on the violence

perpetrated in the universal generalization from the particular White male knowledge of the nature of reality to describe everyone's realities" (p. 671).

As such, the use of gender as a category necessitates nuance given people's diverse worldviews and varied social positions shaped by social contexts and power structures (Collins, 2019a). Across methodological approaches, such nuance is falsely viewed as difficult to actualize, as white, cisgender, heterosexual scholars often forego the intentionality required in discussing minoritized identities when whiteness is no longer a normative focal point of discourse. In quantitative research, using gender as a categorical variable can potentially essentialize the status of a gendered group because measures and statistical analyses may not have the explanatory power to consider exogenous factors that qualitative approaches can explore (Catalano, 2017; McLewis, 2021; Patton et al., 2016). For example, research on gender differences that compare groups at the aggregate level presumes a universal lived experience driven by one's social identity. Even within qualitative research, essentialist notions of gender can be reinforced when scholars do not consider the variation in lived experiences across and within gendered groups (Patton and Njoku, 2019; Stewart, 2022).

Moreover, viewing all feminist research through the reductionist perspective of feminism as an inquiry into cisgender women's issues succumbs to hegemonic descriptors of gender binaries, namely cisgender women and men. In articulating the uses of feminist frameworks in qualitative research, Creswell and Creswell (2018) simplify the focus of this work as women in context: "Feminist perspectives view as problematic women's diverse situations and the institutions that frame those situations. Research topics may include policy issues related to realizing social justice for women in specific contexts or knowledge about oppressive situations for women" (p. 62). This philosophical assumption implies that the utility of feminist frameworks is limited to exploring the experiences of cisgender women. Reducing feminist research on social issues related to women discounts the complexity and evolving conceptions of gender and erases the experiences of sexually minoritized groups (Nicolazzo, 2019; Stewart, 2022), which is ultimately antithetical to the advancement of a transformative paradigm. The oversimplification of feminist research as the study of women's issues also dismisses how a gendered lens can enrich inquiry on social dilemmas and inequities that fortify marginalization and domination. For instance, a gendered lens and/or feminist frameworks may be used to examine men, masculinity, and power relations (Okello, 2019); similarly, when theorized in ways that affirm the socially constructed nature of gender as an identity category, feminist frameworks are well-suited to explore the lived realities of those who hold gender identities beyond the cisgender binary (Carrera-Fernández and DePalma, 2020). In what follows, we discuss how utilizing expansive approaches to feminist epistemologies and theories can effectively center minoritized groups in qualitative inquiry.

Feminist epistemologies and theories

In centering the perspectives of minoritized groups in critical social and feminist research, entry into interrogating ways of knowing in education research reckons with the fraught history of white supremacy and settler colonialism that has prioritized white-dominant social constructions of gender that have historically excluded and exploited Black, Indigenous, and other People of Color (Arvin et al., 2013; Smith, 2012). For example, though intersectionality is now popularized in academic research, Black women have long undertaken an intersectional analysis in intellectual and social justice projects:

Before critical legal scholar and Black feminist Kimberlé Crenshaw coined the term that emerged in academic and legal discourse as intersectionality in the late 1980s (Collins 2019a), Black women were describing and documenting how the intersection of race, gender, class, and sexuality shaped their treatment in US society. In 1861, Harriet Ann Jacobs introduced the enslaved Black woman as the subject while exposing the ills of American slavery in *Incidents in the Life of a Slave Girl*. Early narratives and writings of scholars and activists like Sojourner Truth, Maria Stewart, Anna Julia Cooper, Ida B. Wells, Mary Church Terrell, Frances Beal, and Angela Davis also addressed the structural oppression of Black women.

McLewis (2021), p.136.

Thus, Black Feminists have drawn upon the lessons and knowledge produced by Black women that are rendered unorthodox and not historically welcomed in academia or mainstream white feminist traditions.

Centering women of color feminisms

The feminist identification has been interrogated, contested, and even refuted because of the history of marginalizing racial, working-class, queer, transgender, and non-conforming gender identities and their prioritization of concerns regarding cis-white feminist social movements organizations and research (Evans-Winters, 2019; Moraga and Anzaldúa, 1983). For example, the Combahee River Collective, a collective of Black feminist activists, was formed in 1974, because of their status and experiences of being Black lesbians and socialists marginalized in feminist and civil rights movements. The Combahee River Collective called for the "development of integrated analysis and practice based upon the fact that the major systems of oppression are interlocking" (Hull et al., 1982, p. 13). Similarly, within the research and scholarly communities, minoritized groups have challenged the marginalizing structure of traditional approaches to feminist theory.

Faced with simultaneously being suppressed within feminist research traditions and their respective academic disciplines and fields, Women of Color sought their own cultural epistemologies and standpoints to develop theories, frameworks, and praxis that advanced liberatory practices (Arvin et al., 2013; Crenshaw, 1989; Dill and Kohlman, 2012; Guy-Sheftall, 1995). For example,

intersectionality as a social justice project (Collins, 2019b), analytical framework (Harris and Patton, 2018; Núñez, 2014), and method (Evans-Winters and Esposito, 2019; Museus and Griffin, 2011) advance an analytic that accounts for the nuanced ways people holding multiple, intersecting minoritized identities experience systemic domination structurally, politically, and representationally (Crenshaw, 1989, 1991). Other examples include Critical Race Feminism in legal studies, though now also used in education research (Wing, 2003), and Black Feminist Sociology (Brewer, 1989; Luna and Pirtle, 2021). Within the field of education, Scholars of Color have employed interdisciplinary and critical approaches to advance feminist epistemologies and theories that are attentive to how gender interlocks with other social identities in education contexts. For instance, Delgado Bernal (1998) contends Chicana Feminist Epistemology in education is a “means to resist epistemological racism and to recover untold histories” (p. 556). Among Black feminist theorists, Cynthia Dillard (Dillard, 2000) offers endarkened feminist epistemology which posits

Reality is known when based in the historical roots of Black feminist thought, embodying a distinguishable difference in cultural standpoint, located in the intersection/overlap of the culturally constructed socializations of race, gender, and other identities and the historical and contemporary contexts of oppressions and resistance for African-American women. (p. 662)

Native feminist theory calls for the need to engage with settler colonialism which is too routinely omitted in race and gender discourses (Arvin et al., 2013). Native feminist theorists

actively seek alliances in which differences are respected and issues of land and tribal belonging are not erased in order to create solidarity, but rather, relationships to settler colonialism are acknowledged as issues that are critical to social justice and political work that must be addressed.

Arvin et al. (2013), p.19.

These examples are a select few approaches to feminist theory in education research and are by no means an exhaustive list or fully capture the ontological history and breadth of diverse thoughts and practices within each.

The role of feminist and critical social theories

Critical social and feminist theory account for how epistemology and ontology provide a worldview that informs the research process, including the framing of research questions, the data collection process, the intentionality in data analysis and dissemination, and the justification and contributions for the inquiry. Venus Evans-Winters (2019) describes Black feminism as a mosaic that enriches qualitative research. “Black feminist scholars bring a wealth of knowledge, skills, talents, and experiences into the research process. These bits of experiences mold together to construct our multiple identities. And, from these multiple identities, yields a creative, distinctly mosaic worldview” (p. 15). Education researchers have employed Women of Color feminism and epistemologies to disrupt the race and gender neutrality that permeates scholarship. For example, McLewis (2021) applied Collins (2000) matrix of domination to challenge traditional approaches to studying college “choice.” Others have used Women of Color feminism and epistemologies to develop new theories and frameworks such as Black Girl Cartography (Butler, 2018) and Chicana Motherwork (Caballero et al., 2019) to defy mainstream knowledge structures.

Subsets of feminist theories have also long been inclusive of queer and trans identities. Foundational Black and Chicana feminists, such as Audre Lorde and Gloria Anzaldúa, centered their queer identities in their theorization, carving feminist inquiry as a space to explore the ways queerness intersects with race and gender. Queer scholars of color similarly draw on the work of Women of Color feminisms in constructing critical frameworks that speak to intersections of identities. Brockenbrough (2015), for example, draws from Black feminist scholarship (namely intersectionality) in adapting a Queer of Color critique to educational research. Similarly, at the intersection of race, gender, and sexuality, the critical framework of Jotería Studies is a theoretical framework to examine the experiences of queer Latinx students (Duran et al., 2020). The work of Transfeminists of Color, such as Koyama (2003, 2020), has been employed in education scholarship to demonstrate the ways “schools are rife with gender category oppression and gender transgression oppression” and argue that “teacher educators prepare teachers to challenge the gender oppression matrix in their classrooms” (Rands, 2009, p. 423).

Beyond feminist paradigms, other critical frameworks have been employed to explore the lived realities of both queer and transgender, nonbinary, and gender-nonconforming communities in education research. For one, queer theory has become increasingly commonplace in higher education research centering queer communities (Renn, 2010) and in student affairs research and praxis (Kasch, 2013). Queer theory allows “a critique of societal constructs surrounding sexuality and gender, namely how these constructs determine what is deemed normal” (Goodman and Garcia, 2021, p. 4), discourses that have historically been utilized to police stakeholders at a variety of levels in education and across various educational sectors (Renn, 2010). Scholars have employed more traditional understandings of queer theory in the exploration of queer and trans student experiences across educational sectors (e.g., Abes and Kasch, 2007; Schmidt, 2017), as well as frameworks that advance queer theory to accurately reflect the realities of queer Communities of Color, particularly Black, Indigenous, and other trans People of Color. For example, Queer of Color critique has emerged in education research to critically center queer and trans people of color (QTPOC) narratives in K-12 (Brockenbrough, 2015) and postsecondary education scholarship, such as Duran and Garcia’s (2021) exploration of how Queer Women of Color in

culturally-based sororities experience communal inclusion and exclusion in social media. Critical trans politics, which emerged from the legal scholarship of Dean Spade (2015) and draws from both Women of Color feminisms and queer theory, has similarly informed scholarship centering on the experiences of trans students (Nicolazzo, 2017) and faculty (Jourian et al., 2015) in higher education. Community cultural wealth has also been used as a critical framework in scholarship centering queer and trans-Latinx communities in postsecondary education, such as Duran and Pérez II's (2018) work on how chosen *familia* develops and strengthens queer Latino men's various forms of capital.

Aforementioned are select examples of critical social and feminist theories that can be utilized in qualitative research to illuminate and respond to the concerns of minoritized groups that are often erased from "traditional" ways of knowing. We offer these examples to demonstrate the boundless potential of critical social and feminist theory in transforming qualitative inquiry. In the next section, we argue the importance of infusing critical approaches such as these into all stages of qualitative research design and data collection.

Methodology, methods, and analysis

Fundamental in critical social and feminist theories is that research is culturally situated. Considerations on how methodology and methods are carried out that affirm both the researcher and participant as knowledge producers are central. Scholars oftentimes interchange the terms "methodology" and "methods" (Mackenzie and Knipe, 2006); as such, we begin this section by offering concrete definitions of both research methodologies and methods offered by Harding (1987) to highlight the nuances between these concepts.

Methodology is "a theory and analysis of how research does or should proceed" (Harding, 1987, p. 3). Methodologies are conceptualized as research strategies of inquiry that are connected to the epistemologies, theories, and methods a researcher employs by their research design (Denzin and Lincoln, 2005). These strategies of inquiry "comprise a bundle of skills, assumptions, and practices" that shape how the researcher navigates "from paradigm to the empirical world" (Denzin and Lincoln, 2005, p. 25). Though often conflated with various aspects of the research process, methods as a term is used to describe "a technique for (or way of proceeding in) gathering evidence" (Harding, 1987, p. 2). Further integral to the research design process is data analysis and interpretation. In this section, we explore modalities employed by feminist and critical social theorists to integrate theory and method in qualitative research.

Methodologies

Within the field of education, critical and feminist scholars use various methodologies and methods to examine gender in education research. While some scholars have used more traditional qualitative inquiry approaches (Creswell and Creswell, 2018), these research strategies have their own histories, approaches, and canonical texts rooted in the white cisheteropatriarchy, ideologies that are not in alignment with existing and burgeoning social issues and concerns (Denzin and Lincoln, 2005). Critics have revamped research designs that encompass critical perspectives (i.e., critical race grounded theory, endarkened narrative inquiry, feminist ethnography, queer phenomenology). Some scholars have even innovated alternative methodological approaches that characteristically reimagine research inquiry, such as Okello's (2019) Black feminist performance auto/ethnographic approach or Johnson's (2015) *sista* circle methodology.

Such transformative work proposes alternatives to white male and cisgender orientations of knowledge production that are accustomed practices that are culturally familiar to researchers and/or participants and communities. For People of Color, and/or queer, trans, and other minoritized scholars, this often requires decentering disciplinary standards of the research process. As Bhattacharya (2019) contends, "if we accept that qualitative research is a colonizing enterprise, then we must actively construct new approaches to inquiry that are consistent with our ontoepistemic orientation and that honor the communities with whom we work" (p. 114). This involves rigorously engaging with robust modes of knowledge production, analytical practices, and methodological sensibilities that are attuned to other ways of knowing and being that may be cultural, indigenous, ancestral, spiritual, or activist.

Methods

Further, pushing against research boundaries and refuting research standards reinforced in positivist paradigms, critical social theorists invite methods and tools that disrupt how we conceive what is research data, what is considered means for data collection, and approaches to engaging research participants and communities. As defined above, research methods describe the specific tools used by the researcher to proceed in data collection (Harding, 1987). Methods can include augmented fotonovelas (Hidalgo, 2015), counternarratives and counterstories (Jones, 2021), native texts (Sabzalian (2016), photovoice (Phelps-Ward et al., 2021), *platicas* (Flores and Morales, 2021), poetic elicitation of identities (Gutzwa, 2021b, 2022), portraiture (Dixon et al., 2005; Lawrence-Lightfoot and Davis, 1997), *sista* circles (Brown et al., 2021; Johnson, 2015; Palmer, 2021), storytelling (Toliver, 2021), testimonios (Delgado Bernal et al., 2012; Huber, 2009), walking interviews (Harris, 2019), writings (Muhammad, 2012), and the use of zines (Honma, 2016) to amplify unheard voices.

Researchers employ these methodological techniques that are responsive to the historical, geographical, social, cultural, and political contexts that shape participants' lived experiences and encourage their expression of resistance. For example, critical race feminists use counterstories or counternarratives to challenge the power relations that uphold dominant forms of knowledge. This acknowledges and legitimizes minoritized communities' epistemologies, as not all ways of knowing are equally valued in research or societal contexts. For example, [Jones \(2021\)](#) uses counterstorytelling to examine how Black women negotiate gendered racial microaggressions at historically white postsecondary institutions. Another example is the use of testimonios, both a methodology and method grounded in feminist tradition as a "pedagogical, methodological, and activist approach to social justice that transgresses traditional paradigms in academia" and "challenges objectivity by situating the individual in communion with a collective experience marked by marginalization, oppression, or resistance" ([Delgado Bernal et al., 2012](#), p. 363).

Data analysis and interpretation

At their core, interpretations of qualitative data are "constructive," "endlessly creative and interpretive," and stem from analytic processes that are "both artistic and political" ([Denzin and Lincoln, 2005](#), p. 26). Although feminist and critical social approaches to data collection are plentiful, guides for feminist and critical analysis of qualitative data are notably more elusive ([Bhattacharya, 2019](#); [Maynard, 2004](#)). This may be because there is no one prescribed form of data analysis that is specific to feminist theory; instead, feminist perspectives are employed in the data analysis process to encompass how researchers make meaning of social life in relation to gender, power hierarchies, and systems of domination. Researchers are often implored to marry their theories and epistemologies with every element of research design ([Grant and Osanloo, 2014](#)), however most methodologists discuss the importance of intertwining theory and analysis without instructing on how to effectively do so (e.g., [Saldaña, 2020](#)) or discuss the importance of data analysis in generating future theory. In this section, we work to offer guidance on the integration of critical social and feminist theories with data analysis.

Data analysis encompasses what [Miles et al. \(2020\)](#) refer to as "three concurrent flows of activity: data condensation, data display, and conclusion drawing/verification" (p. 8). Qualitative data analysis comprises of several traditionally accepted strategies, ranging from jottings and reflective memoing ([Saldaña, 2020](#)) to coding ([Saldaña, 2015](#)) and member checking ([Birt et al., 2016](#)), that are ideally intentionally selected by the researcher: "qualitative research methods have theoretical underpinnings that guide the study with emphasis on the development of a theory a posteriori during the data analysis phase" ([Grant and Osanloo, 2014](#), p. 21). For example, qualitative scholars guided by Black feminist epistemologies strive to construct their data analysis in ways that honor the ideals of reflexivity, identity salience, and self-authorship which Black feminism prioritizes, such as by adopting in vivo or open coding strategies that honor the self-authored narratives of participants ([Clemons, 2019](#)). Integrating participants as co-researchers in the analysis process is another way Black feminist scholars have worked to critically transform data analysis, with [Evans-Winters \(2019, p. 5\)](#) offering an understanding of qualitative data analysis as "a co-constructed piece of artwork" between her and her participants (p. 5). Similarly, researchers guided by Black queer studies, such as epistemologies of BlackLove, strive to prioritize relationality, communal solidarity, and ancestral legacies ([McGuire et al., 2022](#)).

These examples highlight feminist and critical social epistemologies as powerful tools that can reject white, cisheteronormative, patriarchal understandings of knowledge production traditionally prioritized by qualitative research ([Bhattacharya, 2019](#); [Robinson, 2022](#)). As such, while coding and member checking are some of the analytic tools available to qualitative researchers, those guided by feminist and critical social theories must also consider expanding beyond those strategies of data analysis that are reified as "rigorous." [Robinson \(2022\)](#), for example, offers their understanding of non-binary knowledge production as an epistemology that "captures my state of thinking about my own experiences of marginalization and discrimination and a way that exposes and disrupts cisheteronormative and heteronormative ways of knowing and writing" (p. 3); in doing so, they suggest that writing and theorizing about silencing in one's lived experiences is one potential analytic tool that can challenge the hegemonic prioritization of disclosure in qualitative inquiry.

Such strategies that connect the personhood of a researcher, the participants they are in community with, and the epistemologies they employ can be adopted by all: as [Robinson \(2022\)](#) reminds, "if they experience gender recognition, they too can theorize what this recognition means in relation to their field sites, methods, and the people in their study" (p. 15). Thoughtful interrogation of both positionality and reflexivity is therefore a vital component of feminist and critical social approaches to data analysis, as the "complex, messy, and multilayered" approaches researchers take in interpreting data reflect the complexities and nuances across their in/visible social identities ([Evans-Winters and Esposito, 2019](#), p. 55). In the next section, we discuss the importance of positionality and reflexivity in feminist and other critical social qualitative research, providing examples of positionality in praxis as reflections of the ways we as authors have individually engaged our shifting and evolving positionalities in our work.

Positionality and reflexivity

Critical and feminist approaches to qualitative research contend knowledge is socially constructed and dependent on the perspectives that are being privileged ([Collins, 2019b](#); [Dillard, 2000](#); [Harding 1992](#)). In either form, quantitative or qualitative, research is a political project when considering power relations and the impact of research in shaping social worlds. As [Denzin and Giardina \(2019\)](#) argue, "the researcher is historically and locally situated within the very processes being studied" (p. 6), meaning that one's social identities, worldviews, and perspectives will always permeate their approach to research. Challenging the positivist premise of

objectivity in research, critical and feminist theories contest the notion that there is a universal truth or neutral social scientist. Harding (1992) states strong objectivity can be achieved by challenging power dynamics and notions of value-free science. Specifically, she contends the use of standpoint theories can facilitate getting closer to the truth through the inclusion and centering of minoritized perspectives. However, some feminist and postmodernist critics suggest the search for the truth and situated knowledge in standpoint theories even with strong objectivity or reflexivity still subscribe to traditional empiricist approaches and undermine the fluidity in social positions (Naples and Gurr, 2014; Pillow, 2003; Sabzalian, 2018).

Though ideological and theoretical approaches to understanding knowing vary (e.g., differences between standpoint and post-modern feminist approaches), situated perspectives and knowledge involve reckoning with individuals' multiple lived realities that are derived from varied social positions and contexts. As such, when engaging in meaning-making and knowledge production, it is imperative to consider the role of the researcher, their identities, and their world views. Positionalities are "the ways that researchers' cultures work as lenses through which they experience the world, and how they envision and conduct research" (Rogeman et al., 2016, p.44). Among some qualitative scholars, one's positionality can strengthen the research process, particularly when there is a shared identity between the researchers and the participants. This presumption is informed by the dominance of white thought in academic disciplines, the suppression of knowledge produced by marginalized scholars, and epistemic oppression (Collins, 2000; Dotson, 2014). For example, Collins (1986) contends that an outsider *outsider within* can benefit social science research because it "might generate a distinctive standpoint vis-a-vis existing sociological paradigms" (p. S16). Similarly, Delgado Bernal (1998) illustrates cultural intuition is "nurtured through our personal experiences (which are influenced by ancestral wisdom, community memory, and intuition), the literature on and about Chicanas, our professional experiences, and the analytic process we engage in when we are in a central position in our research and our analysis" (p. 567–568).

However, others contest the binaries of outsider versus insider status with origins in anthropology. Applying Linda Tillman' (2002) framework for culturally sensitive research, Rogeman and colleagues (2016) argue "cultures [are] always present throughout research processes and not an additive feature distinct from research processes" (p. 45). Rather than subscribing to fixed notions of outsiders and insiders, one can employ considerations on how the cultural knowledge of the researcher and participant interact. Culturally sensitive research prioritizes centering participants' cultural, co-constructed knowledge through analysis that is reciprocal, ethical, and responsive to power dimensions. While some scholars challenge the dichotomy of outsiders versus insiders in research, some critical social theorists account for the complexities of conducting within-group research in response to overly essentializing and romanticizing in-group status. For example, Collins' (1986, 2000; 2019a) understanding of *outsiders within* emphasizes the importance of antiessentialism by accounting for how social identities are situated in a social context and vary across matrices of domination.

While positionality involves reflection on one's identity and social position shapes their understanding of the social world and impacts the research, reflexivity is practiced (Hesse-Biber, 2014). Sharlene Hesse-Biber (2014) defines reflexivity as "a process by which researchers recognize, examine, and understand how their social background, location, and assumptions can influence the research" (p. 3). Other feminist researchers describe reflexivity "as critical reflection on the ways that researchers, research participants, their communities, and academic audiences create, use, sustain, and reconfigure power within complex, dynamic relationships" (Knight et al., 2004, p. 391). This involves moving beyond simply acknowledging one's positionality but integrating throughout the research process active interrogation on how it influences the construction and representation of knowledge (Knight et al., 2004; Pillow, 2003). "Reflexivity under feminism is not only about investigating the power embedded in one's research but is also about doing research differently." (Pillow, 2003, p. 178). Examining the use and tension of reflexivity as a methodological tool for representation and legitimacy, Pillow (2003) problematizes four ways in which reflexivity is commonly practiced: reflexivity as recognition of self, reflexivity as recognition of other, reflexivity as truth, and reflexivity as transcendence. Using postmodernism, Pillow (2003) posits a need for "uncomfortable reflexivity" as "a reflexivity that seeks to know while at the same time situates this knowing as tenuous" and "rendering the knowing of their selves or their subjects as uncomfortable and uncontainable (p. 188).

The power of positionality in praxis

In sum, researcher positionality and reflexivity are viewed as instrumental in inquiry. The history of exclusion and exploitation in research and science has necessitated reflexivity as a mechanism, though not solely sufficient, to interrogate Western research traditions and dominant knowledge systems. Reflection of researcher positionality and the practice reflexivity are not limited to qualitative inquiry nor addressing status differences and cultural incongruence between researchers and participants. The politics of research compels the need to undertake positionality and reflexivity as integral components of research processes, whether it is quantitative or qualitative inquiry, "white-on-white research" (Phillippo and Nolan, 2022), or research among communities with shared identities. Further, effective engagement of positionality and reflexivity in both the research and writing processes can help challenge white, cisheteronormative logics of identity: as Robinson (2022) reminds, it is possible "to improve reductionist writing about race, gender, and sexuality is in how we do positionality and reflexivity [...] We can turn to embodiment to write positionality in a more nuanced manner, away from static, essentialist notions of identity" (p. 10).

As shown, the methodological process in how positionality and reflexivity are practiced in qualitative research varies. However, essential to transformative research is deep consideration and action that challenges the false premise of the neutral and objective researcher and attention to how research can advance social change (Milner, 2007). As researchers, our positionalities are always with us. Our identities and backgrounds do not leave our psyches as we engage with participants and analyze data. As Rogeman and colleagues (2016) argue, "we take up grappling with positionality not as fixed and located in a physical space where research

occurs, but rather, as it travels with us, within and across multiple social and geographic locations and communities" (p. 47). In the following sections, two authors of this chapter, McLewis and Gutzwa, demonstrate how they have grappled with these realities when conducting research with minoritized communities in education contexts by describing how their positionalities traveled with them throughout the research process.

McLewis' positionality applied: othersisters in feminist research

Employing Critical Race Theory and Black Feminist Thought, my research examined how gendered racism contributes to Black women's pathways to college and their perceptions of their college decisions. I investigated how interlocking forms of oppression operate at both the macro (structural) and micro (interpersonal) level shaped how Black women made decisions about which college to attend and their perceptions of their "choices." My *outsider within* status and identity as a Black woman conducting research with and for Black women granted entry to engage with participants as they storied their educational trajectories. As an outsider, the lack of rapport with prospective participants caused concern about how open students would be to volunteering to participate in my study. After all, rapport is often deemed keen to successfully recruit and engage research participants. Moreover, I was a graduate student and did not attend their campus for my undergraduate education. Surprisingly, however, I found ease in interviewing other Black women. One by one, I would invite women to share about their trajectories to college, beginning with the question: "Tell me about yourself. What's your age, major, what city are you from, and your future aspirations?" Within minutes, students were sharing with me, a perfect stranger, intimate details about their lived experiences that entailed accounts of family drama, secrets, the trauma of sexual violence, challenges and triumphs of navigating educational spaces, and their hopes and dreams. Findings illuminate the extent to which they trusted me and the lack of campus-based spaces and support available specific to Black women's unique needs.

Black feminists describe how a Black feminist standpoint provides a way of knowing and examining research data that is obscured in more traditional research approaches. Our shared identities and similar trajectories made the strange familiar. Their telling of issues related to racial microaggressions from white faculty, predatory male peers disguised as mentors, and colorism and dating retold experiences my Sista-friends and I encountered as undergraduates. In interviewing Black undergraduate women about their college-going experiences, I found that my commitment to social justice, my sense of moral and ethical obligations, and my positionality as the eldest daughter blurred the boundaries of the objective researcher. Dillard (2000) emphasizes that "all views expressed and actions taken related to educational inquiry arise from a personally and culturally defined set of beliefs that render the researcher *responsible* to the members and the well-being of the community from which their very definition arises" (p. 672). As they shared their college experiences, there were instances positions as a researcher shifted into an "othersister" role to assist Black undergraduate women in navigating the university.

Within Black Feminist Thought, othermothers (James, 1993) are "women who assist bloodmothers by sharing mothering responsibilities" (Collins, 2000, p. 178). Collins (2000) describes how historically within the Black community, the realities of intersecting systems of oppression required othermothers, such as extended family members, neighbors, and networks, to play a significant role in providing familial and communal support. Pertinent to this discussion is how community, support, fictive kinship, and social responsibility shape the relationships Black women have with members of broader Black communities. The sentiment of othermothers is thus helpful in conceptualizing the role of Black women and other women of color scholars working with research communities, especially those that reflect their own social identities. I offer "othersisters" to designate the relationship between women of color and their respective research communities.

The role of othersisters in research settings entails the social responsibility associated with conducting research with communities one cares about. My decision to intrude on the research process was guided by the sense of responsibility to ensure that these women were equipped to be academically and socially successful at the university. For example, some of the women I interviewed described challenges in meeting other Black students given the university was a predominately white institution and lacked a Black cultural student center. With my familiarity with cultural affinity and underground social network groups, I shared information about resources on campus. Other actions included sharing information about Title IX offices, providing advice on courses and professors to seek out, assisting in completing financial aid documents, and even scholarship award information to recipients who were unaware that they were fully funded and were working part-time to support their studies—a revelation and indicator of the limited college knowledge and institutional support Black women receive before and after entering a selective university that is based on white middle-class assumptions regarding their needs.

Gutzwa's positionality applied: emic and etic flux in the field

In the early stages of my doctoral career, I embarked on a critical, qualitative study which explored trans collegians' experiences navigating collegiate classroom environments (Gutzwa, 2021b, 2022). At the time, I identified as a cisgender, white, queer settler, placing me firmly outside of the communities the study sought to engage. As such, the power and privilege which my social identities afforded me were constant point of reflection for me in the field. I thought long and hard—to the point of almost stopping—about why I was inspired to do this work, what place I had in working with communities I myself was not a part of, and how I could be furthering damage against an already heavily systemically minoritized community. One way I worked to disrupt the power balance of the research setting was through the use of creative methods, including the writing of poems and the drawing of identity maps, that aimed to decenter traditional white, cisgender modes of knowledge production (Bhattacharya and Payne, 2016; Gutzwa, 2022). Through collecting what Esteban-Guitart (2016) describes as "testimonies of identity" that can illuminate the ways of

knowing students develop throughout their lives and bring with them to educational spaces, this asset-based approach prioritized self-authored narratives of trans student experiences in research design.

Before beginning data collection for the study, it was suggested by a peer that I offer to participate in the creation of “I am from ...” poems and drawings alongside my participants as a way to build a relationship and present my own ways of knowing to them. I still remember the first interview I conducted for the project. When describing the poem elicitation activity to my participant, a nonbinary Student of Color, I offered to complete the exercise alongside them. Immediately, they let out a sigh of relief, saying that my doing so would make them feel more comfortable. Through that poem and the questions my participant asked me that followed, I explained my history with mental health in college, my neurodivergence, my queerness—things that I am open about conversationally, but rarely in such a raw, visceral way. While I believe it impacted my participant’s willingness to share their experiences with me, this was the first moment where I felt conducting qualitative research could also become an act of healing for me, the researcher.

Over the course of the interviews and conversations with my participants in this study, I began to slowly write more through my poems and draw more through my circles than I had ever verbalized before. In essence, I came out to myself as nonbinary through my writing, through these poems, while collecting data. My positionality evolved and shifted every day in the field. Hearing students discuss their experiences with gender and simultaneously exploring sides of my memories and heart that had been sealed away through the trauma of my lived experiences, I came to understand that I no longer identified as a cis man. Through my work, I found myself in the community I was engaging with: My methods were as liberatory for me as they were for my participants.

In her essay and call to action “The transformation of silence into language and action,” Black feminist Audre Lorde (1984b) argued that all feminists ought to be united by their strifes in “reclaiming ... that language that has been made to work against us” (p. 43). Employing decolonial methods in my research methods through a critical, queer, feminist lens enabled me to actualize Lorde’s words—not just to reclaim words that subjugate my communities and the communities I work with, but also to reclaim the language that has been weaponized against me throughout the course of my life. Through taking a critical approach in understanding gender in my research, I not only build rapport with participants but also intimately interrogated my positionality in the field. This reflexive process further embodies Robinson’s (2022) non-binary knowledge production: in writing this thick description of my own positionality, “I follow the queer routes that my mind and life take me”, highlighting the ways the evolution of my identities and my relationality to those who I engaged in research with both shapes my meaning making of trans experiences and serves as a rejection of hegemonic white, cisheteronormative tensions in qualitative research (p. 3).

Conclusion

In this chapter, we explored how critical social and feminist theories can be employed to advance research and the complexity of gender in the field of education. We complicate the examination of gender in research that fails to interrogate how gender is socially constructed and its intersection with other forms of domination and power. As stated by Stewart (2022), “Containing (limiting) diversity, in ways that reify normative understandings of identity and diversity, undermines the ability of researchers to uncover findings that would lead to implications capacious enough to effect transformation and disrupt structures of oppression.” (p. 62). Advancing a transformative research agenda insists on disrupting oppressive ideologies and assumptions embedded in research practices that are implemented by researchers. While we present critical social and feminist theories as frameworks that can be utilized when forging transformative research, we caution the mere application of these theories (or methods) without intentional integration of disruptive practices throughout research processes.

Addressing the liberatory and empowering aims of a transformative research agenda opens opportunities for gender-based research in education to become a mutually dialogic process, potentially becoming as transformative for researchers as it is for participants in building resistance to dominant ideologies and frames. Fundamental to this work, however, is that researchers interrogate their own investment in the inequities they seek to study and engage in greater self-awareness regarding their own identities, epistemologies, and relationship with participants. This process requires much more than simply laying bare one’s biases in a positionality statement to increase the trustworthiness of a study. Intentionally interrogating one’s positionality necessitates also considering how the researcher sees their role as an agent of change, seeks to disrupt reinforcing systems of oppression to extend educational opportunities, improves outcomes for minoritized groups and communities, and dares to reimagine new possibilities of education environments and research. More examples and descriptions of methods for increasing reflexivity and ways to debrief that liberate the researcher from dominant assumptions are needed in the literature.

In contrast, an increasing number of creative methods have arisen to amplify the voices of gender-identity groups in the literature, conveying their perspectives and agency. These methods and methodologies are oftentimes not taught in traditional methods classes, but instead are more often derived from the cultural communities that researchers and participants belong to which are then subsequently shared among and adopted by communities of researchers. These communities, comprised of both knowledge constructors and seekers, provide the energy and insight needed to dismantle interlocking forms of oppression, ultimately harnessing the liberatory and transformative potential of educational research and praxis.

References

- Abes, E.S., Kasch, D., 2007. Using queer theory to explore lesbian college students' multiple dimensions of identity. *J. Coll. Stud. Dev.* 48 (6), 619–636.
- Allen, W., Jones, C., McLewis, C., 2019. The problematic nature of racial and ethnic categories in higher education. *Am. Council Educ.* 15–19.
- Arvin, M., Tuck, E., Morrill, A., 2013. Decolonizing feminism: challenging connections between settler colonialism and heteropatriarchy. *Fem. Form.* 8–34.
- Bhattacharya, K., 2019. Theorizing from the streets: de/colonizing, contemplative, and creative approaches and consideration of quality in arts-based qualitative research. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative, and Methodological Reflection*. Routledge, pp. 109–125.
- Bhattacharya, K., Payne, R., 2016. Mixing mediums, mixing selves: arts-based contemplative approaches to border crossings. *Int. J. Qual. Stud. Educ.* 29 (9), 1100–1117.
- Birt, L., Scott, S., Cavers, D., Campbell, C., Walter, F., 2016. Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qual. Health Res.* 26 (13), 1802–1811.
- Brockenbrough, E., 2015. Queer of Color agency in educational contexts: analytical frameworks from a Queer of Color critique. *Educ. Stud.* 51 (1), 28–44.
- Brewer, R.M., 1989. Black women and feminist sociology: the emerging perspective. *Am. Sociol.* 20 (1), 57–70.
- Brown, L.C., Williams, B.M., Williams, Q.A.S., 2021. Melanin messages: Black college women's experiences and reflections on navigating colorism. *J. Div. High. Educ.* <https://doi.org/https://doi.org/10.1037/dhe0000347>. (Advanced online publication).
- Butler, T.T., 2018. Black girl cartography: Black girlhood and place-making in education research. *Rev. Res. Educ.* 42 (1), 28–45.
- Caballero, C., Martínez-Vu, Y., Pérez-Torres, J., Téllez, M., Vega, C., Castillo, A., 2019. *The Chicana Motherwork Anthology*. University of Arizona Press, Tucson.
- Carrera-Fernández, M.V., DePalma, R., 2020. Feminism will be trans-inclusive or it will not be: why do two cis-hetero woman educators support trans feminism? *Socio. Rev. Monogr.* 68 (4), 745–762.
- Catalano, D.C.J., 2017. Resisting coherence: trans mens' experiences and the use of grounded theory methods. *Int. J. Qual. Stud. Educ.* 30 (3), 234–244.
- Clemons, K.M., 2019. Black Feminist Thought and Qualitative Research in Education. *Oxford Research Encyclopedia of Education*.
- Collins, P.H., 1986. Learning from the outsider within: the sociological significance of Black feminist thought. *Soc. Probl.* 33 (6), s14–s32.
- Collins, P.H., 2019a. Intersectionality as Critical Social Theory. *Duke University Press*.
- Collins, P.H., 2019b. The difference that power makes: intersectionality and participatory democracy. In: Hankivsky, O., Jordan-Zachery, J.S. (Eds.), *The Palgrave Handbook of Intersectionality in Public Policy*. Palgrave Macmillan, pp. 167–192.
- Collins, P.H., 2000. Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment. *Routledge*.
- Crenshaw, K.W., 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, pp. 139–167.
- Crenshaw, K.W., 1991. Mapping the margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1299.
- Creswell, J.W., Creswell, J.D., 2018. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage Publications.
- Delgado Bernal, D., 1998. Using a Chicana feminist epistemology in educational research. *Harv. Educ. Rev.* 68 (4), 555–582.
- Delgado Bernal, D., Burciaga, R., Flores Carmona, J., 2012. Chicana/Latina testimonios: mapping the methodological, pedagogical, and political. *Equity Excell. Educ.* 45 (3), 363–372.
- Denzin, N.K., Giardina, M.D., 2019. Introduction: qualitative inquiry at a crossroads. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative, and Methodological Reflection*. Routledge, pp. 1–16.
- Denzin, N.K., Lincoln, Y.S., 2005. Introduction: the discipline and practice of qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*. SAGE Publications, pp. 1–32.
- Dill, B.T., Kohlman, M.H., 2012. Intersectionality: a transformative paradigm in feminist theory and social justice. *Handbook of Fem. Res. Theory Praxis* 2, 154–174.
- Dillard, C.B., 2000. The substance of things hoped for, the evidence of things not seen: examining an endarkened feminist epistemology in educational research and leadership. *Int. J. Qual. Stud. Educ.* 13 (6), 661–681.
- Dixon, A.D., Chapman, T.K., Hill, D.A., 2005. Research as an aesthetic process: extending the portraiture methodology. *Qual. Inq.* 11 (1), 16–26.
- Dotson, K., 2014. Conceptualizing epistemic oppression. *Soc. Epistemol.* 28 (2), 115–138.
- Duran, A., Garcia, C., 2021. Social media as a source of (dis)connection for queer women of color in culturally-based sororities. *Oracle* 16 (1), 19–35.
- Duran, A., Orozco, R.C., Gonzalez, S.A., 2020. Imagining the future of *jotería* studies as a framework in the field of higher education. *Assoc. Mexican Am. Educ. J.* 14 (2), 67–86.
- Duran, A., Pérez II, D., 2018. The multiple roles of chosen familia: exploring the interconnections of queer Latino men's community cultural wealth. *Int. J. Qual. Stud. Educ.* 32 (1), 1–18.
- Esteban-Guitart, M., 2016. *Funds of Identity: Connecting Meaningful Learning Experiences in and outside of School*. Cambridge University Press.
- Evans-Winters, V.E., 2019. Black Feminism in Qualitative Inquiry: A Mosaic for Writing Our Daughter's Body. *Routledge*.
- Evans-Winters, V.E., Esposito, J., 2019. Intersectionality in education research: methodology as critical inquiry and praxis. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative, and Methodological Reflection*. Routledge, pp. 52–64.
- Flores, A.I., Morales, S., 2021. A Chicana/Latina feminist methodology: examining *pláticas* in educational research. *Handbook Latinos Educ.* 35–45.
- Garcia, N.M., Mayorga, O.J., 2018. The threat of unexamined secondary data: a critical race transformative convergent mixed methods. *Race Ethn. Educ.* 21 (2), 231–252.
- Gillborn, D., 2010. The colour of numbers: surveys, statistics and deficit thinking about race and class. *J. Educ. Pol.* 25 (2), 253–276.
- Gillborn, D., Ladson-Billings, G., 2010. Critical race theory. *Int. Encyclopedia Educ.* 6, 341–347.
- Goodman, M.A., Garcia, C., 2021. From summer pride to transgender day of visibility: a review of sorority/fraternity social media involving LGBTQ+ identities and issues. *Oracle* 16 (1), 66–82.
- Grant, C., Osanloo, A., 2014. Understanding, selecting, and integrating a theoretical framework in dissertation research: creating the blueprint for your "house". *Admin. Issues J. Connect. Educ. Pract. Res.* 4 (2), 12–26.
- Gutzwa, J.A., 2021a. Trans exclusionary radical feminists (TERFs). In: Strunk, K.K., Shelton, S.A. (Eds.), *Encyclopedia of Queer Studies in Education*. Brill, pp. 695–698.
- Gutzwa, J.A., 2021b. "It's not worth me being who I am": exploring how trans* collegians navigate classroom experiences through a funds of identity lens. *J. Women Gend. High. Educ.* 14 (3), 302–323.
- Gutzwa, J.A., 2022. "It's survival mode": exploring how an Indigenous Trans* Student of Color (per)forms identity while transgressing space. *J. Diver. High. Educ.* <https://doi.org/https://doi.org/10.1037/dhe0000410>. (Advance online publication).
- Guy-Sheftall, B., 1995. *Words of Fire: An Anthology of African-American Feminist Thought*. The New Press.
- Harding, S.G. (Ed.), 1987. *Feminism and Methodology: Social Science Issues*. Indiana University Press.
- Harding, S., 1992. Rethinking standpoint epistemology: what is "strong objectivity?". *Centen. Rev.* 36 (3), 437–470.
- Harris, J.C., 2019. Whiteness as structuring property: multiracial women students' social interactions at a Historically White Institution. *Rev. High. Educ.* 42 (3), 1023–1050.
- Harris, J.C., Patton, L.D., 2018. Un/doing intersectionality through higher education research. *J. High Educ.* 1–26.
- Hesse-Biber, S.N. (Ed.), 2007. *Handbook of Feminist Research*. Sage Publications, Thousand Oaks, CA.
- Hesse-Biber, S.N., 2014. A re-invitation to feminist research. In: Hesse-Biber, S.N. (Ed.), *Feminist Research Practice: A Primer*. Sage Publications, pp. 1–13.
- Hidalgo, L., 2015. Augmented fotonovelas: creating new media as pedagogical and social justice tools. *Qual. Inq.* 21 (3), 300–314.
- Honma, T., 2016. From archives to action: zines, participatory culture, and community engagement in asian America. *Radic. Teach.* 105, 33–43.

- Huber, L.P., 2009. Disrupting apartheid of knowledge: testimonio as methodology in Latina/o critical race research in education. *Int. J. Qual. Stud. Educ.* 22 (6), 639–654.
- Hull, G.T., Scott, P.B., Smith, B. (Eds.), 1982. *All the Women Are White, All the Blacks Are Men, but Some of Us Are Brave: Black Women's Studies*. The Feminist Press.
- Hurtado, S., 2015. The transformative paradigm: principles and challenges. In: Aleman, A., Posner, B.P., Bensimon, E.M. (Eds.), *Critical Approaches to the Study of Higher Education*. Johns Hopkins University, Baltimore, MD, pp. 285–307.
- James, S.M., 1993. A possible Black feminist link to social transformation. In: *Theorizing Black Feminisms: The Visionary Pragmatism of Black Women*, pp. 44–54.
- Jones, A.M., 2021. Conflicted: how Black women negotiate their responses to racial microaggressions at a historically white institution. *Race Ethn. Educ.* 1–16.
- Johnson, L., 2015. *Using Sista Circles to Examine the Professional Experience of Contemporary Black Women Teachers in Schools: A Collective Story about School Culture and Support*. Unpublished doctoral dissertation. University of Georgia.
- Jourian, T.J., Simmons, S.L., Devaney, K.C., 2015. "We are not expected": trans* educators (re)claiming space and voice in higher education and student affairs. *TSQ: Transgender Stud. Quart.* 2 (3), 431–446.
- Kasch, D., 2013. Queer theory. In: Jones, S.R., Abes, E.S. (Eds.), *Identity Development of College Students: Advancing Frameworks for Multiple Dimensions of Identity*. Jossey-Bass, pp. 191–212.
- Kiyama, J.M., Rios-Aguilar, C., 2017. Conclusion: the future of funds of knowledge in higher education. In: Kiyama, J.M., Rios-Aguilar, C. (Eds.), *Funds of Knowledge in Higher Education: Honoring Students' Cultural Experiences and Resources as Strengths*. Routledge, pp. 189–196.
- Knight, M.G., Bentley, C.C., Norton, N.E., Dixon, I.R., 2004. (De) constructing (in) visible parent/guardian consent forms: negotiating power, reflexivity, and the collective within qualitative research. *Qual. Inq.* 10 (3), 390–411.
- Koyama, E., 2003. The transfeminist manifesto. In: Dicker, R., Piepmeyer, A. (Eds.), *Catching a Wave: Reclaiming Feminism for the Twenty-First Century*. Northeastern University Press.
- Koyama, E., 2020. Whose feminism is it anyway? The unspoken racism of the trans inclusion debate. *Socio. Rev. Monogr.* 68 (4), 735–744.
- Ladson-Billings, G., Tate, W.F., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97 (1), 47–68.
- Lawrence-Lightfoot, S., Davis, J.H., 1997. *The Art and Science of Portraiture*. Jossey-Bass, San Francisco.
- Lorde, A., 1984a. An open letter to Mary Daly. In: Lorde, A. (Ed.), *Sister Outsider: Essays and Speeches*. Ten Speed Press, pp. 66–71.
- Lorde, A., 1984b. The transformation of silence into language and action. In: Lorde, A. (Ed.), *Sister Outsider: Essays and Speeches*. Ten Speed Press, pp. 40–44.
- Luna, Z., Pirtle, W., 2021. *Black Feminist Sociology: Perspectives and Praxis*. Routledge.
- Lynn, M., Dixon, A.D., 2021. *Handbook of Critical Race Theory in Education*. Taylor & Francis.
- Mackenzie, N., Kriple, S., 2006. Research dilemmas: paradigms, methods and methodology. *Issues Educ. Res.* 16, 1–11.
- Maynard, M., 2004. Feminist issues in data analysis. In: Hardy, M., Bryman, A. (Eds.), *Handbook of Data Analysis*. SAGE Publications, pp. 131–145.
- McGuire, K.M., Edwards, K.T., Dancy II, T.E., 2022. BlackLove stories. In: Tachine, A.R., Nicolazzo, Z. (Eds.), *Weaving an Otherwise: In-Relations Methodological Practice*. Stylus Publishing, pp. 127–143.
- McLewiss, C.C., 2021. The Limits of choice: a Black feminist critique of college "choice" theories and research. In: Perna, L.W. (Ed.), *High Educ. Handb. Theor. Res.* 36, 105–160.
- Mertens, D.M., 2009. *Transformative Research and Evaluation*. Guilford Press, New York.
- Miles, M.B., Huberman, M., Saldaña, J., 2020. *Qualitative Data Analysis: A Methods Sourcebook*, fourth ed. SAGE Publications.
- Milner IV, H.R., 2007. Race, culture, and researcher positionality: working through dangers seen, unseen, and unforeseen. *Educ. Res.* 36 (7), 388–400.
- Moraga, C., Anzaldúa, G., 1983. *This Bridge Called My Back: Writings by Radical Women of Color*. Kitchen Table: Women of Color Press, Brooklyn, NY.
- Muhammad, G.E., 2012. Creating spaces for Black adolescent girls to "write it out!". *J. Adolesc. Adult Literacy* 56 (3), 203–211.
- Museum, S.D., Griffin, K.A., 2011. Mapping the margins in higher education: on the promise of intersectionality frameworks in research and discourse. *N. Dir. Inst. Res.* (151), 5–13.
- Naples, N.A., Gurr, B., 2014. Feminist empiricism and standpoint theory. In: Hesse-Biber, S.N. (Ed.), *Feminist Research Practice: A Primer*. Sage Publications, pp. 14–41.
- Nicolazzo, Z., 2017. Trans* in College: Transgender Students' Strategies for Navigating Campus Life and the Institutional Politics of Inclusion. Stylus, Sterling, VA.
- Nicolazzo, Z., 2019. Starting with trans*: the effects of methodological violence through gender categorization in higher education research. In: Henderson, E.F., Nicolazzo, Z. (Eds.), *Starting with Gender in International Higher Education Research: Conceptual Debates and Methodological Considerations*. Routledge, New York, NY, pp. 144–160.
- Núñez, A.M., 2014. Advancing an intersectionality framework in higher education: power and Latino postsecondary opportunity. In: Paulsen, M.B. (Ed.), *Higher Education: Handbook of Theory and Research*, pp. 33–92.
- Okello, W., 2019. "I wanted the world to see": Black feminist performance auto/ethnography. In: Denzin, N.K., Giardina, M.D. (Eds.), *Qualitative Inquiry at a Crossroads: Political, Performative, and Methodological Reflection*. Routledge, pp. 32–47.
- Omi, M., Winant, H., 2014. *Racial Formation in the United States*. Routledge.
- Palmer, D., 2021. I Got a Lot to Be Mad about: Sista Circles as a Counterspace for Black Women Doctoral Students (Order No. 28718315). [Doctoral dissertation, Indiana University]. ProQuest Dissertations and Theses Global.
- Patel, L., 2014. Countering coloniality in educational research: from ownership to answerability. *Educ. Stud.* 50 (4), 357–377.
- Patton, L.D., 2016. Disrupting postsecondary prose: toward a Critical Race Theory of higher education. *Urban Educ.* 51 (3), 315–342.
- Patton, L.D., Crenshaw, K., Haynes, C., Watson, T.N., 2016. Why we can't wait: (Re) examining the opportunities and challenges for Black women and girls in education (Guest Editorial). *J. Negro Educ.* 85 (3), 194–198.
- Patton, L.D., Njoku, N.R., 2019. Theorizing Black women's experiences with Institution sanctioned violence: a #BlackLivesMatter imperative toward Black liberation on campus. *Int. J. Qual. Stud. Educ.* 32 (9), 1162–1182.
- Phelps-Ward, R., Latz, A.O., Turner Kelly, B., Kortegast, C., 2021. Re-examining and reimagining power in participatory visual methodologies: a collaborative autoethnography. *Int. J. Qual. Stud. Educ.* 1–18.
- Phillippo, K.L., Nolan, J.L., 2022. White-on-white research: a study of white qualitative researcher positionality among white participants. *Int. J. Qual. Stud. Educ.* 1–19.
- Pillow, W., 2003. Confession, catharsis, or cure? Rethinking the uses of reflexivity as methodological power in qualitative research. *Int. J. Qual. Stud. Educ.* 16 (2), 175–196.
- Rands, K.E., 2009. Considering transgender people in education: a gender-complex approach. *J. Teach. Educ.* 60 (4), 419–431.
- Renn, K.A., 2010. LGBT and other queer research in higher education: the state and status of the field. *Educ. Res.* 39 (2), 132–141.
- Rios-Aguilar, C., Neri, R.C., 2021. Funds of Knowledge, Community Cultural Wealth, and the Forms of Capital: Strengths, Tensions, and Practical Considerations. *Urban Education*.
- Robinson, B.A., 2022. Non-binary embodiment, queer knowledge production, and disrupting the cisnormative field: notes from a trans ethnographer. *Journal of Men's Studies*. Advance online publication.
- Robinson, C.J., 2020. *Black Marxism, Revised and Updated Third Edition: The Making of the Black Radical Tradition*. UNC Press Books.
- Roegman, R., Knight, M.G., Taylor, A.M., Watson, V.W., 2016. From microscope to mirror: doctoral students' evolving positionalities through engagement with culturally sensitive research. *Int. J. Qual. Stud. Educ.* 29 (1), 44–65.
- Sabzalian, L., 2016. Native feminisms in motion. *Engl. J.* 23–30.
- Sabzalian, L., 2018. Curricular standpoints and native feminist theories: why native feminist theories should matter to curriculum studies. *Curric. Inq.* 48 (3), 359–382.
- Saldaña, J., 2015. *The Coding Manual for Qualitative Researchers*, third ed. SAGE Publications.
- Saldaña, J., 2020. Qualitative data analysis strategies. In: Leavy, P. (Ed.), *The Oxford Handbook of Qualitative Research*, second ed. Oxford University Press.
- Schmidt, S.J., 2017. Genderplay and queer mapping: heterotopia as sites of possible gender reform and spatial reconstruction. In: Ares, N., Buendía, E., Helfenbein, R. (Eds.), *Deterritorializing/reterritorializing: Critical Geography of Educational Reform*. Sense Publishers, pp. 209–232.
- Smith, L.T., 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*, second ed. Zed Books.
- Solorzano, D.G., 1997. Images and words that wound: critical race theory, racial stereotyping, and teacher education. *Teach. Educ. Q.* 5–19.

- Spade, D., 2015. *Normal Life: Administrative Violence, Critical Trans Politics, and the Law*, second ed. Duke University Press.
- Stewart, D.-L., 2022. Performing goodness in qualitative research methods. *Int. J. Qual. Stud. Educ.* 35 (1), 58–70.
- Tillman, L., 2002. Culturally sensitive research approaches: an African-American perspective. *Educ. Res.* 31 (9), 3–12.
- Wing, A.K. (Ed.), 2003. *Critical Race Feminism: A Reader*, second ed. New York University Press, New York.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8 (1), 69–91.
- Zuberi, T., Bonilla-Silva, E. (Eds.), 2008. *White Logic, White Methods: Racism and Methodology*. Rowman & Littlefield Publishers.

Critical race theory and educational research utilizing qualitative methods

Tondra L. Loder-Jackson^a, Mary Ann Bodine Al-Sharif^a, Christopher E. Jones^b, Martez D. Files^c, and Charletta N. Wiggins^d

^aThe University of Alabama at Birmingham, School of Education, Birmingham, AL, United States; ^bThe University of Alabama at Birmingham, Office of Diversity, Equity, and Inclusion, Birmingham, AL, United States; ^cUniversity of Pittsburgh, School of Education, Pittsburgh, PA, United States; and ^dThe University of Alabama at Birmingham, The Classic Upward Bound Academy, Birmingham, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update and a synthesis of the following two chapters: R. Vue and C. B. Newman, Critical Race Theory in Education Research, In International Encyclopedia of Education (Third Edition), pp. 779–786, Edited by: Penelope Peterson, Eva Baker, Barry McGaw, Elsevier, 2010, <https://doi.org/10.1016/B978-0-08-044894-7.01738-3>; D. Gillborn and G. Ladson-Billings, Critical Race Theory, In International Encyclopedia of Education (Third Edition), pp. 341–347, Edited by: Penelope Peterson, Eva Baker, Barry McGaw, Elsevier, 2010, <https://doi.org/10.1016/B978-0-08-044894-7.01530-X>.

Introduction	68
Critical race theory under attack	68
Critical race theory's definition and legal foundations	68
Critical legal theory	69
Centrality of race	70
The centrality of experiential knowledge	70
Researcher positionality	71
Narratives, stories, and counterstories	72
Branches of critical race theory and its applications in disciplines beyond law	72
Critical race theory in educational research	72
Reframing dominant educational theory through a critical race perspective	73
CRT and qualitative methods in educational research	73
Implications of CRT for future educational scholarship	74
Conclusion	75
References	75
Further reading	77

Glossary

Colorblindness or colorblind ideology Refers to a worldview that disregards race or racial characteristics. In practice it is a method of disregarding race as a criterion used for selection of participation or receipt of services

Critical race theory A multi-disciplinary approach rooted in critical legal theory that combines social activism with a critique of the fundamental role played by racism, white supremacy, and white privilege in shaping contemporary societies

Interest convergence principle An analysis, first proposed by critical legal scholar Derrick Bell, noting that advances in race equality have historically come about only when white elites see the changes as in their own interest; especially as a means of limiting civil unrest and protecting against fundamental societal change

Intersectionality An approach that is especially sensitive to the ways in which multiple axes of oppression interact, e.g., analyzing how racism operates within and across factors such as social class and gender

Jurisprudence Refers to the theory or philosophy of law. In this article it primarily refers to the dominant legal thought which excludes and marginalizes racial minorities in the United States

Meritocracy Refers to a system of rewards based on the ability and talent of individuals, leaving unaccounted for the structural advantages and disadvantages conferred to different groups in society

Positivism As it concerns social reality and knowledge is the notion that both exist objectively and independently of social actors

Race Refers to a socially constructed classification used to categorize people into groups based on phenotype and characteristics assumed to be inherently shared

White privilege Any phenomena, whether individual, structural, political, economic, or social, that serve to benefit white people while oppressing people of color. The tangible benefits of access to resources and social rewards which whites receive, coupled with the power to shape the norms and values of society, unconsciously or consciously, by virtue of their skin color

White supremacy Commonly refers to individuals and groups who engage in the crudest, most obvious acts of race hatred (such as the Ku Klux Klan and other extremist nationalists and neo-Nazis). In critical race theory the term refers to a regime of assumptions and practices that constantly privilege the interests of white people but are so deeply rooted that they appear normal to most persons in the culture

Introduction

There has probably never been a more apropos time for education scholars to define what critical race theory (CRT) is (and is not), to revisit its historical foundations and original intent, and to trace its application in educational research. Specifically, this chapter updates and synthesizes two extant chapter reviews of CRT in this encyclopedia series (see “Tagline”) and pushes further the exploration of CRT as a theoretical framework for educational research that utilizes qualitative methods. Although an exhaustive list of the theoretical elements and arguments of CRT scholars is beyond the scope of this article (see [Ledesma and Calderon, 2015](#)), the key principles presented provide a basic understanding of its theoretical underpinnings. However, since CRT became such a political lightning rod under the United States presidential administration of Donald J. Trump (2017–2021), and is currently being contested vociferously in state gubernatorial, legislative, and educational bodies across the country ([Sawchuk, 2021](#), May 18; [Wallace-Wells, 2021](#), June 18), we begin with a brief overview of the present context. Second, we define CRT and review its historical and legal foundations. Third, we offer a description of CRT’s resonance in fields beyond law, namely education, and present examples of qualitative research where it is employed to elucidate issues in this field. Finally, we conclude with implications for the future of educational research framed by CRT and qualitative methods.

Critical race theory under attack

The term “critical race theory” has been bulldozed out of the confines of academia and into the mainstream of American political discourse since the Trump era ([Sawchuk, 2021](#), May 18; [Wallace-Wells, 2021](#), June 18). Under the Trump administration, with calculated counsel from Christopher Rufo, a Seattle-based conservative think tank consultant once profiled on Fox News, this discourse took the decidedly polemical turn of denouncing CRT tenets and any acknowledgment of white privilege in federal racial sensitivity training programs wholesale ([Wallace-Wells, 2021](#), June 18). Namely, *Executive Order on Advancing Racial Equity and Support for Underserved Communities through the Federal Government* (<https://www.archives.gov/presidential-libraries/archived-websites>) issued by the Director of the Office of Budget and Management, ordered all federal agencies to cease all racial sensitivity trainings on white privilege and critical race theory citing them as “divisive anti-American propaganda.” The memo took specific aim at any CRT training espousing that “virtually all White people contribute to racism” or “benefit from racism.”

Notably, since January 2021, thirty-five states have introduced bills or taken steps to restrict teachers from teaching CRT ([Schwartz, 2022](#), February 4; [Associated Press, 2022](#), January 19). Conservative gubernatorial candidates have been hedging their bets on winning campaigns partly on *anti-CRT* platforms with the most recent election upset leveled by Virginia’s Republican Governor Glenn Youngkin, who has vowed to abolish CRT in public education and government ([Beauchamp, 2021](#), November 4). In response, many K-12 and higher educators and administrators across the country are confused, worried, and tentative about whether and how they should be teaching about race and racism in United States history and current events. There are mounting concerns that *anti-CRT* backlash is having a “chilling effect” not only on using CRT frameworks but also on the teaching and training of diversity, equity, and inclusion topics writ large, especially those focused on race, gender expression, and sexuality ([Harvard Law Review, 2020](#), February 10; [Collman, 2021](#), September 2). Yet many educators are joining ranks with legal, civic, and academic organizations to fight back.

This polarizing context underscores the importance of clarifying the definition, origins, and significance of CRT in educational research, particularly to counter myths and misunderstandings of CRT. Three myths or misunderstandings seem particularly common. The first is that CRT views race as the source of all major inequalities. This myth will be addressed in a subsequent section on intersectionality. The second is that CRT views white people as uniformly racist and privileged. While CRT does implicate racism as an endemic system that benefits whites uniformly, this does not mean that it does not recognize how poverty, gender, and other forms of oppression differentiate the experiences of whites. It also does not support the view that whites cannot be genuinely invested in *anti-racist* causes. Finally, there is the myth that CRT promotes a sense of hopelessness and pessimism. Although scholars and individuals who view racism as endemic to society within the United States tend to offer many critiques, they also hold out hope for a more democratic, just, and inclusive society if indeed the root causes of inequality and injustice are called out and sufficiently addressed. Delgado and Stefancic responded to this myth by inquiring, “Is medicine pessimistic because it focuses on diseases and traumas?” (2001, p. 13).

Critical race theory’s definition and legal foundations

CRT recognizes racism as systemic and endemic to American society across all spheres—education, housing, employment, health-care, etc. ([NAACP Legal Defense Fund, n. d.](#)). It is an academic, multi-disciplinary approach rooted in critical legal theory that combines social activism with a critique of the fundamental role played by racism, white supremacy, and white privilege in shaping contemporary societies. More broadly, CRT is an evolving movement formalized in the legal field during the 1970s in order to grapple with the nexus between race, racism, and American jurisprudence, in a more constructive manner ([Bell, 1987, 1992](#); [Crenshaw, 1988](#)). The term critical race theory was coined in the 1980s by legal scholars, mostly people of color, who were dissatisfied with existing approaches to the field.

CRT's emergence represents a response to the retrenchment of civil rights reform. This accounts for critical race scholars' critique of civil rights doctrine (which espouses equal treatment) and dissatisfaction with traditional legal reform. Both legal ideals were viewed as inadequate for the emancipation of people of color and susceptible to discourses supporting neoconservative agendas to reinstate racial subordination. Accordingly, CRT was formulated by scholars whose primary concern was advancing a theory and method for examining race and overcoming racial inequality through critical analysis of legal doctrines.

Though critical race theorists vary in the scope of their study, at least five central tenets guide its use in education (i.e., theory, research, pedagogy, curriculum, and policy): "Centrality of race and racism and their intersectionality with other forms of subordination, the challenge to dominant ideology, the commitment to social justice, the centrality of experiential knowledge, and the transdisciplinary perspective" (Solórzano et al., 2000, p. 63). The use of CRT in education is appropriate for several reasons; however, the main justification extends its emergence out of critical legal studies (CLS) both in its points of alignment and departure. Thus, before addressing CRT in education some discussion of its emergence in critical legal theory merits attention.

Critical legal theory

CRT is considered to have its origins in a departure from the theorizing of CLS. While the critical race movement's inception is marked by critiques of legal scholars, most of whom were scholars of color (such as Derrick Bell, Richard Delgado, Allan Freeman, and later scholars, such as Kimberle' Williams Crenshaw, Angela Harris, and Mari Matsuda) within the legal studies movement, the formulation of CRT is inherently informed by a myriad of previous movements—including (among others) feminism and the spirit of the civil rights movement encompassing a long legacy of resistance to racial subordination (Delgado and Stefancic, 2001)—and literature across multiple disciplines (Yosso, 2005). However, legal scholars were instrumental in informing the constructs that ultimately have come to define the critical race movement and theory.

The emergence of CRT is more than simply an extension of CLS: it serves as a critique. Although critical race theories and critical legal theories are separate and in contention with one another, they share some notions about how law operates in society. Broadly, the critical legal movement emerged to counter presumptions that the law operates in a neutral and objective manner (Hunt, 1986). Specifically, it advocates notions of realism, a movement "developed out of the dissatisfaction with tenets of classical legal thought that cast judicial decision as the product of reasoning from a finite set of determined rules" (White, 1972 as cited in Tate, 1997, p. 207). Thus, CLS provides an analytical lens to critically examine legal ideology and discourse as a mechanism to socially reproduce and legitimate inequality within society (Hunt, 1986; Tate, 1997). In alignment with this tradition, Tate's (1997) review of CRT origins notes that critical legal theories provide two important critiques that would ultimately inform CRT: these include a critique of formalism and objectivism. The first is characterized by "impersonal purposes, policies, and principles that serve as the foundation for legal reasoning" while the latter assumes the legal system provides an intelligible moral order (Tate, 1997, p. 208). Therefore, critical legal theorists offer a critique of law which acknowledges legal ideology and legal structures of authority as producing and maintaining social inequality, as opposed to functionalist views viewing social inequality as natural. This critique has served critical race theorists in their interrogation of inconsistencies within legal doctrines, specifically explanations of the contradictions and therefore negative consequential effects of civil rights law on progressive racial reform (i.e., Bell, 1976, 1980; Freeman, 1995). For these reasons, Matsuda (1995) described, "[t]he willingness to at least consider the utopian prospect and the enthusiastic criticism of existing conditions of racism and poverty attract nonwhite readers to CLS" (p. 64). Given the legacy of racial inequality within the United States, largely maintained through the construction of laws, critical legal theories share an important perspective with the critical race theorist, which is necessary for understanding the nexus between law and oppression.

Despite its utility, critical race theorists point to limitations of critical legal theories, which render them inadequate for understanding experiences of populations of color. Dalton (1995) wrote, as scholars of color within the CLS movement they "are always invited in for tea, but rarely invited to stay for supper, lest we use the wrong intellectual fork" (p. 80), specifically speaking to the limits and determinism that CLS's hegemonic view of rights doctrine poses for people of color. The collective response to sentiments of disempowerment and intellectual differentiation resulted in CRT; that is, a movement which foregrounds the lived experiences of people of color in order to both acknowledge their marginalization in American jurisprudence but also enable a vision of law that holds the potential for social justice in terms of equal rights, status, and material outcomes. While there are multiple theoretical discontents, including inability to offer structures for social change, a central critique that critical race theorists have of critical legal theories is the assumption that racism is the same as other forms of oppression (Delgado and Stefancic, 1993). Utilizing legal theorists' critiques of legal doctrine as a foundation for their own work, critical race theorists brought the racial dimension to the forefront of the analysis: race, they assert, is crucial for understanding and ultimately overcoming inequality given racial saliency in the United States.

Grounded in traditions of the critical legal theorist, Bell's early writings (including Bell, 1976, 1980) have been identified as "the intellectual starting point of Critical Race Theory" (Crenshaw et al., 1995, p. 3). Crenshaw and colleagues believe that while Bell's critiques of civil rights discourse represent critical legal critiques of the time, his use of a race-conscious perspective that (1) focused specifically on the effect on Blacks and (2) presented legal decisions as reflective of white racial interests, was a departure in perspective, marking the development of CRT. Specifically, Bell's (1980) concept of interest convergence—the argument that basic rights are provided only when they meet the shared interests of and do not harm the benefits of whites—was significant in propelling forward the critical race movement, especially as scholars of color worked to construct a theory in which race centered as an important, necessary basis of analysis.

Centrality of race

The centrality of race is one of the distinguishing factors of CRT (Delgado and Stefancic, 1993; Solórzano et al., 2000). Crenshaw (1988) pointed to the importance of race as a basis of analysis, given the legacy of racial oppression in the United States. She argued to ignore claims that race is illogical, especially when race has become so deeply embedded in how institutions operate to produce social inequality. For that reason, critical race theorists argue that race is significant and excluding race from the analysis merely veils rather than halts its influence in research, policy, and practice. By bringing race to the forefront of the analysis, critical race theorists explicate the presence use of race. Therefore, revealing race consciousness will illuminate what has traditionally been ignored or hidden in past analysis.

By evoking principles of equality while ignoring the cumulated effects of race reification in material conditions, the concept of colorblindness is deployed to eliminate progressive gains of racial minorities. A well-known and debated contemporary example of this is present in debates over the issue of affirmative action in United States higher education. Opponents of affirmative action argue that race-based sanctions are unjust because they betray a fair approach by privileging people of color over whites. To ground their arguments, opponents call forth the contradictions that affirmative action poses to treating all people as equal. Embedded within this discourse is the notion of a meritocracy, which is a system of rewards believed to be based on demonstrated talent and ability of an individual. Both simultaneously work to undermine social reform addressing race inequities. The use of colorblind ideology to dismantle affirmative action policies exemplifies critical race legal scholars' critiques of civil rights doctrine, and more broadly liberalism, which is ambiguous and potentially serves to disarm progressive efforts toward social justice. In principle alone, arguments grounded in colorblind discourses may appear to adhere to moral reasoning; however, this rationale leaves accounted for the historical and sociocultural contexts that justify the need for affirmative action practices.

Crenshaw (1988) addressed the notion of equality by framing it in terms of equal process and equal outcomes. Using the terms restrictive and expansive equality Crenshaw argued that equal processes are not always indicative of equal outcomes. In accord with the earlier work of Freeman (1995), who used the competing concepts of perpetrator and victim perspectives of racism to explicate race (and consequently non-neutral nature) in law, Crenshaw pointed to the absence of historical continuity in rationalizing colorblind approaches to equality. The notion of equality as used to support the elimination of race-based initiatives fails to acknowledge the presence of different opportunities, erroneously positioning whites as the true victims. As such, race as used in the case of explaining affirmative action, and more generally law in the United States is important because it explicates the presence of race and the role of racism in ideology, structures, and practices.

Similarly, Tate (1997) raised the issue of race as critical for understanding law, education, and society. Arguments for including race are based on the premise that a colorblind approach requires a consciousness of race to be able to reject its consideration (Aleinikoff, 1991 as cited in Tate, 1997, p. 203). In this regard, ignoring race merely disguises rather than eliminates its significance to debates. In citing Justice Harry Blackmun's dissenting opinion in the *Bakke* decision, this argument is exemplary of critical race thought: "In order to get beyond racism [to color blindness], we must first consider race. There is no other way" (Regents of the University of California vs. *Bakke*, 1978, p. 407 as cited in Tate, 1997, p. 203). More importantly, recognizing that race exists unveils the construction of whiteness as normative. Critical race theorists put forth race as a strong analytical tool for understanding and overcoming racism. Therefore, CRT's advancement in the legal field serves to further deconstruct the law, which was formed (and evolved) to maintain racial privileges of one group and consequently legalize the oppression of another; all factors which betray the objectiveness it presupposes.

Although race is a central feature of CRT critical race scholars recognize the role of multiple forms of subordination including class, gender, and sexuality operating in society and law. Therefore, embedded within the critical race framework is a rejection and critique of essentialism, that is, the presumption of a monolithic experience for a particular group with a number of shared cultural characteristics (Delgado and Stefancic, 2001). *Anti-essentialism* is the notion there is no singular experience for people of color, Black people, or any other socially-constructed group. To further advance the *anti-essentialism* thesis, critical race feminist scholars proposed intersectionality to more meaningfully account for the unique ways in which multiple forms of oppression interact to shape the lived experiences of those with multiple identity (Collins, 1996). Intersectionality does not merely suggest that multiple identities are additive, but that their intersectional nature create a unique social location of marginality as well as the potential for ground breaking social and political empowerment. For example, Crenshaw (1995) described how both feminists and antiracist discourses have failed to adequately consider women of color's intersectional identities. Notably, other Black feminist and womanist scholars and authors have addressed this failure in seminal texts such as *Black Feminist Thought*, *Ain't I a Woman*, *Sister Outsider*, and *In Search of Our Mother's Gardens* (Collins, 2000; Hooks, 1981; Lorde, 1984; and Walker, 1983). Therefore, the centrality of race is not to minimize the role of these other forms of oppression, but instead to recognize its social construction largely through legal processes (among other means), such as the formation of the Constitution, which continue to influence society. Critical race scholarship works under the notion that racial subordination is intricately linked to other forms of oppression; thus, dismantling racial inequity is a step toward achieving social injustice.

The centrality of experiential knowledge

Within the CRT movement is the commitment to voice scholarship, which centers experiential knowledge as important. Critical race scholars emphasize the importance of localized narratives, which helps to contextualize and provide understanding. This incorporates a methodological and theoretical approach of perspective taking that requires "looking to the bottom,"—the oppressed of

society—to effectively elucidate and appreciate the complexities of their struggle, both material and ideological (Matsuda, 1995). Further, Delgado (1989) argued that a method of including the experiential knowledge of people of color counters the way traditional legal scholarship conducts analysis. Given that many legal scholars embrace universalism over particularity, the perspective tends to minimize anything that is historical, contextual, or specific with the unscholarly descriptors of literacy or personal (Tate, 1997, p. 220). Critical race theorists argue for the inclusion of these factors in helping to explain the realities of people of color, which are often marginalized in the presentation of law and society generally.

Bringing attention to whose voice is included and reflected in the dominant narrative points to the larger issue of what is considered legitimate knowledge in scholarship. Bringing radicalized lenses to the discussion not only gives voice to marginalized communities, but also offers a critique of the dominant ideology of what is understood as valid knowledge. The inclusion of narratives of people of color in understanding and reforming law counters the dominant discourse and challenges the racial construction of whiteness as the norm by which standards and merit are judged. Universal truths and methods of reasoning as they have been constructed in law are inherently biased toward privileging the dominant group. This is exemplified in Freeman's (1995) analysis of court decisions on civil rights cases of the past century, asserting that the court often adopted a perpetrator perspective of racial discrimination—viewing “racial discrimination not as a social phenomenon but merely as the misguided conduct of particular actors” (p. 30)—a process that consequently negates the victim perspective of discrimination which names the inequities in outcomes and conditions. Accordingly, CRT scholars emphasize subjectivity in experiences, which account for the differences in reality (Delgado, 1989). Thus, where they previously have been silenced, centering on the lived experiences of people of color, voice is a crucial concern of analysis and methodology, both for legal policy and for educational research.

Researcher positionality

Related to this concept of voice, as it involves inclusion of people of color, is the role of the scholar and researcher. This aspect has been addressed by critical race scholars, who find themselves as outside scholars, marginal within their fields of study (Delgado, 1984, 1992). Specifically, scholars like Delgado (1984) have called attention to how the ethos of academia perpetuates dominant narratives and inadvertently works to de-legitimize and silence scholars of color, particularly in discussions concerning minority rights (i.e., civil rights). Additionally, this critique has led to employing methodology which includes reflexive processes of self-interrogation (particularly for white scholars, discussed in more detail below), interrogating how scholars' worldviews inform, understand, and contextualize the textual product. These aspects of theory and methodology are at the heart of CRT.

Notably, researchers must be aware of their own positionality—that is, awareness and understanding of how their racial and/or ethnic identities may position them in a larger system of privilege, power, and/or oppression that affects their own personal experiences, the perceptions of their research participants' experiences, and the possibility of creating deficit discourse within their work (Coghlan and Brydon-Miller, 2014; Milner, 2007). Milner (2007) noted that it was not so much that researchers must represent the identities both racially and culturally of their study participants, but more so that they must “be actively engaged, thoughtful, and forthright regarding tensions that can surface when conducting research where issues of race and culture are concerned” (p. 388). This is important because the experiences of people of color are shaped by the various attributes of their identities and most significantly by their race, ethnicity, and cultural heritage. Therefore, it is important for researchers to be conscientious of the potential for implicit bias, that is, the unconscious act of prejudice and/or stereotyping, to appear within their work (Brownstein, 2019). Hence, researchers must be held accountable for the discourse they use to present their work and how it represents their participants and their participants' communities.

For this reason, much discussion is needed concerning the presence and approach of white scholars utilizing CRT as advocates against racial injustice. Bergerson (2003) learned through discussions with her colleagues of color that “for whites to move into the area of CRT would be a form of colonization in which we [white scholars] would take over CRT to promote our own interests or recenter our positions while attempting to ‘represent’ people of color” (p. 52). Because whiteness allows one to remove an awareness of race, it also creates a power dynamic where whites are able to define the attributes of other races. Therefore, whiteness becomes the norm and can give a false sense of colorblindness—the idea that race no longer matters (Bonilla-Silva, 2013). As noted previously, the notion of colorblindness provides an avenue for racism to persist within a white-normed environment that gives an inequitable preference to people of color who appear whiter (Bergerson, 2003). Unfortunately, this also attributes further to issues of colorism—giving preferential treatment to those of lighter skin tone by people of color within communities of color (Glenn, 2008; Wiggins, 2021).

Therefore, white scholars must rethink their positioning when conducting research within communities of color and their utilization of CRT. In particular, white scholars must acknowledge their privileged perspectives and the benefits that they receive due to their whiteness (Bergerson, 2003). In addition, white scholars must center their concepts of race by seeing whiteness as a non-neutral race; hence, dispelling the concept of colorblindness (Bonilla-Silva, 2013). People of color developed CRT to provide counter-narratives and give understanding of their lived experiences (DeCuir-Gunby et al., 2019; Delgado and Stefancic, 2017). Thus, white scholars must utilize the tenets of CRT as a way to inform their concepts of race and as advocates by insisting “that CRT is a legitimate method and lens for studying racism” and therefore, normalizing its application by citing those who have used it (Bergerson, 2003, p. 60). Because CRT can be utilized to create an understanding of the lived experiences of people of color and how race-neutral policies and practices continue to perpetuate racism within their lives and communities (Parker, 2015; Parker and Lynn, 2002), scholars who utilize CRT must recognize their privilege, position to power, and/or ability to oppress in order to ensure that their work does not further acts of racism but creates change.

Narratives, stories, and counterstories

The critical race movement was founded on the works of scholars such as Derrick Bell and Richard Delgado, who employed and established a narrative tradition, which departed from mainstream legal scholarship and was more commonly found in literature. These include the use of “parables, chronicles, stories, counterstories, poetry, fiction, and revisionist histories” (Ladson-Billings and Tate, 1995). These narrative forms enable scholars to challenge dominant discourses that propagate stereotypes and distorted images of people of color. Appropriately termed counterstories, they exist as a method for conveying the untold experiences of marginalized groups, and also function as a tool for analyzing dominant narratives presumed to be reality. Strauss and Corbin (1990) were early pioneers of this process by coining early forms of this approach as theoretical sensitivity. This sensitivity is achieved by collecting data from sources that have been collected by the true lived experience surrounding the phenomenon being explored, (Solórzano and Yosso, 2002). Data sources within counter-storytelling methods include the researcher’s own experiences and data collected from the research process itself. Even though counter-storytelling is the method often utilized by qualitative CRT scholars, other methods have included first person narratives, critical race grounded theory, and problem-posing methodology (Smith-Maddox and Solórzano, 2002).

Among the reasons for utilizing counterstories include (1) community-building; (2) challenging the perceived wisdom of the those at the center of society; (3) self-preservation for marginalized groups; (4) constructing a richer description of issues and stories that reality alone may be unable to provide; and (5) providing a context for understanding and challenging established belief systems (Delgado, 1989; Lawson, 1995 as cited in Solórzano and Yosso, 2002). While stories are a critical part of the critical race tradition, their use is highly contentious and is not universally accepted as appropriate for the academy. Still, CRT scholars argue that this method is valid, and critiques originate from a positivist paradigm which privileges the normative nature of whiteness as the status quo. Further, they argue that counterstories are necessary particularly since reality is socially constructed and therefore stories from people of color’s lived experiences can illuminate their circumstances, since the experiences of whites are privileged in the dominant or majoritarian narrative (Delgado, 1989).

Branches of critical race theory and its applications in disciplines beyond law

While CRT is largely used in the legal field, the theory has been adopted by scholars in fields such as sociology (Christian et al., 2019). The approach grew in popularity and transcended its disciplinary boundaries, moving into educational studies in the mid-1990s (Ladson Billings and Tate, 1995). Scholars in these fields continue to utilize and advance CRT’s core tenets—most notably the centrality of race and experiential knowledge. Therefore, it is a reductionist misrepresentation to declare that CRT is confined solely to the legal academic realm. Although CRT originated in the legal field, the theory has made its way into multiple fields. In addition to sociology and education, due to its interdisciplinary focus, CRT literature also overlaps with literature from fields such as women studies, ethnic studies, and history.

Critical race theory has also been advanced to offer new insights and innovative critiques of previous theories as well as issues within these fields. For example, CRT as a theoretical construct has expanded to incorporate, or further illuminate among extant theories, the experiences of women (FemCrit, Black feminist thought, and womanism), Latinas/os (LatCrit), Native Americans (TribalCrit), Asian Americans (AsianCrit), sexualities, genders, and other social identities that fall outside normative identities (QueerCrit). Even white scholars have drawn inspiration from CRT (WhiteCrit) to examine and challenge white privilege and supremacy often under the umbrella of “whiteness” studies. These branches of CRT move beyond the Black/white binary that CRT critiques tended to follow, understandably, as their critiques were originally focused on civil rights legislation (Yosso, 2005). Therefore, CRT has become more nuanced as it is applied to multiple experiences of people of color.

Yet, just as each of these expansions represents the varied issues and priorities within and between communities of color and socially marginalized groups, CRT scholars acknowledge, “the institution of slavery is foundational in American racism” and instructive for the struggle against multiple oppressive “isms” beyond the Black–White paradigm (Matsuda, 2002, p. 397). Notably, the recent emergence of a Black Critical Race Theory or BlackCrit perspective recenters concerns about Black people’s experiences in an *anti*-Black society (Dumas and ross, 2016). BlackCrit asserts that *anti*-Blackness is endemic in global society and is in tension with notions of multiculturalism. Moving beyond what some observers view as a pessimistic tenor of CRT, BlackCrit encourages Blacks to envision—even fantasize—a liberatory future unfettered by the legacy of slavery and oppression.

Critical race theory in educational research

CRT is used in education to frame how institutions maintain racism, sexism, and classism through structures, processes, and discourses. Ladson-Billings and Tate (1995) initially introduced CRT to the field of education, arguing that race has gone virtually untheorized in understanding issues of education. Together, they put forth CRT as a theory for understanding and overcoming school and educational inequity. Since then, a growing number of scholars began employing CRT in education to elucidate how educational practices and widely accepted theories minimize the importance of race and/or perpetuate a system which oppresses, subordinates, and marginalizes African American, Latino/a, Asian American, and Native American students. Solórzano and Yosso (2001) contended that stereotypes and deficit frames used in elementary and secondary educational practices reproduce social inequities. Educational practitioners, who subscribe to these stereotypes, suggest that students of color are intellectually inferior, lack

motivation, and are destined to become engaged in criminal activity. Additionally, educational practitioners have used deficit-framed perspectives to explain the academic underachievement of students of color. These frameworks often put the onus of blame on student and/or his or her culture for poor academic performance. For example, deficit-frame employers may state, "African Americans (or Latinos) do not perform academically because African American (or Latino) culture does not emphasize education." Clearly, the unequal resources, the quality of education, and societal inequities are not considered in these deficit frames. As a result, these educational inequities may be manifested in low expectations for African American, Latino/a, and Native American students, tracking, remediating the curriculum, and limiting access to high-level or advanced placement courses (Solórzano and Yosso, 2001; Darling-Hammond, 2001).

Reframing dominant educational theory through a critical race perspective

CRT is used to reconceptualize dominant educational theory, which has been framed around European American cultural values. A noteworthy example is Yosso (2005), who uses CRT to reframe the Bordieuan theory of cultural capital, which is a narrowly defined theory based on "White, middle class values, and is more limited than wealth—one's accumulated assets and resources" (p. 77). Yosso expanded this view to include cultural capital more representative of communities of color, which include aspirational capital (maintain hopes and dreams), familial capital (cultural knowledge and community history), linguistic capital (skill gained from communicating in more than one language), social capital (networks and community resources), navigational capital (negotiate social institutions), and resistant capital (knowledge and skill gained from opposing inequality). Collectively, Yosso referred to this new cultural capital frame as a community of cultural wealth.

Similarly, Patton et al. (2007) used a critical race frame to problematize racelessness (p. 43) in prevailing college student development theories. An example of a raceless college student development theory is in Chickering and Reisser's (1993) *Education and Identity*, which gives limited treatment to the role of race in identity development. Additionally, Patton et al. pointed out that some scholars have used homogeneously white samples to inform the theories they developed. Subsequently, these theories are applied to college students in general and do not account for the experiences of college students of color. Patton et al. suggested "Color-blind racism and racial indifference must consistently be challenged through exposing the manner in which racial advances often come at the cost of promoting or feeding into White self-interests" (p. 43).

Files' (2021) dissertation, "Mothering Ourselves to Wholeness," worked to uncover the myriad ways Black mothers in life and literature have forged communal wholeness, protected the most vulnerable, and healed harm outside and within communities. In this work, the conceptual framework of Black Mothering Praxis (BMP) constellates the communal practices that counter exploitation, center care, bolster accountability, and foster wholeness in a world that often breaks and fragments Black bodies and beings. Integrating Critical Race Parenting (CRP), Black Feminist Thought (BFT), and Black Mothering (BM) as theoretical frameworks, this research began with identifying the various ways Black mothers have been read as a "problem" in research. It contends that even though BFT has, in some ways, become more mainstream, racist misreadings of Black mothers as deficient and unfit persist. Files (2021) noted how numerous researchers have worked to address Black mothers' misreadings, stereotyping, framing, criminalization, and dehumanization. However, according to Files (2021), few studies move beyond the framing and re-imagining of Black motherhood and Black mothering and build a path for exploring how Black mothers' portrayal in literary texts can be connected to the narratives that real-world Black mothers tell. This research engages those narratives to instruct us in creating more livable and loving communities, based on the Black Mothering Praxis of rural, urban, and suburban Black mothers, and the Black mother characters that show up in Black women's literature. The research gathered the narratives of fourteen socially and politically engaged Black mothers and thematically analyzed those narratives to find particularities and commonalities. The themes emerged and highlighted connection to elders, self-care, and freedom dreams as central to the Black mothers' stories. Moreover, Files (2021) engaged in rigorous Black literary critic to understand the ways Black mothers are represented in fictive narratives. Drawing from Black feminism(s) and critical race theory, the research concludes that Black mothering practices activate freedom, resist oppression, and restore community wholeness. Finally, the work demonstrated how Black mothers' life stories in literature and across regional contexts clear pathways for fostering communal care and accountability.

Exemplified through the works reviewed, CRT can aid scholars in reshaping the discourse around students, families, and communities of color in educational research. Previously, the norms and values of dominant cultural groups were juxtaposed with those of communities of color, resulting in an assimilationist approach to dealing with students of color. This approach suggests that students must discard cultural values not aligned with the dominant frame in order to succeed. As a result, students' (of color) culture is to blame for their academic underperformance and not the system maladaptive to multiple points of view. CRT has given scholars and practitioners the tools to reevaluate theoretical perspectives and educational practices from multiracial and multicultural perspectives.

CRT and qualitative methods in educational research

Educational researchers have effectively utilized CRT to frame empirical research applying qualitative methods. For example, CRT-based research has been used to explain the importance of family and culture in education, the experiences of Black males in higher

education, the role of racial micro-aggressions in the perceived college-campus climate of students of color, and intra-racial concerns among Black collegiate women. In these contexts, CRT goes beyond simply describing a problem or challenges associated with educational outcomes and experiences; it explicates the sources of persisting problems, which is very often grounded in systemic racism and oppression.

Focusing on ethnographic studies of Latino/a families and education, Villenas and Deyhle (1999) used CRT to position schools' and larger society's negative perceptions of cultural differences in family socialization and education within a framework of power relations and the castification of Latinos/as in the United States. They found that Latino families play key roles resisting assaults on Latino/a culture. The history of Latinos as taught from the dominant White ideology, propagates views of Latinos/as illegal immigrants and foreigners. The ethnographic studies demonstrate how Latino/a families can take collective action to transform the schooling of their children. Clearly, CRT can be used to identify counternarratives that "challenge dominant culture's 'stock stories' which legitimize the world views of White, upper-class United States" (p. 441).

CRT perspectives and qualitative methods have been used to examine the experiences of Black males within educational institutions (Ladson-Billings and Tate, 2006; Hiraldo, 2010). In particular, CRT has helped to provide counternarratives to demonstrate Black male student success as well as challenge the dominant culture's deficit-oriented narratives surrounding Black male achievement (Jones, 2021). Harper (2009) used CRT's counternarratives to understand the experiences of high achieving Black male college students and to demonstrate their resilience in spite of the various barriers placed on them by administrators and educators alike. By capturing firsthand narratives of Black males, CRT has contributed to a research paradigm that is no longer focused on the deficit narratives which have been the focus of previous Black male research (Harper, 2009). Centering the voices of Black males acknowledges their resiliency and supports solving the challenges they face within education (Jones, 2021).

A CRT perspective can also elucidate the atmospheric nature of racism in educational institutions as hostile climates for students of color. Solórzano et al. (2000) used CRT to describe the experiences of African American students attending predominantly white research universities. Participants in this study shared their experiences with everyday subtle forms of racism and stereotypes (micro-aggressions). Examples of these micro-aggressions include feeling invisible as one of only a few African Americans in a particular college classroom, faculty having low expectations, challenges in joining study groups with non-African American students and feeling a sense of hyper-surveillance because of perceived criminality. Micro-aggressions created a negative campus racial climate, which led to feelings of frustration, isolation, self-doubt, and exhaustion. Many students in this study coped with the negative campus racial climate by creating counterspaces where they could be surrounded by supportive peers and collectively reject the notions of intellectual inferiority and stereotypes of deviance or criminality.

Finally, CRT coupled with qualitative methods can address intra-racial concerns, of course, understood within the context of racism, white supremacy, and white privilege. According to Wiggins (2021), colorism perpetuates the same variables that drive the principles of racism and has led to disparate treatment, lower self-esteem, and negative perceptions within communities of color, particularly among Black women, for decades. Since Black women are considered second class citizens, behind both men and white women, it is vital to acknowledge that their experiences with racism, compounded by colorism, are most salient. In her qualitative research study entitled *Shades of Black: Black Collegiate Women and the Impact of Colorism in Higher Education*, Wiggins set out to: (1) explore colorism as a phenomenon while also investigating its impact on the higher education attainment of Black collegiate women, (2) amplify these women's voices regarding their personal experiences with colorism; and (3) generate culturally pertinent information that can be used to raise awareness, provide empowerment, and create social change regarding how colorism is viewed. In order to answer the question of how Black women narrate their experiences with colorism in higher education at a predominantly white urban public state institution in the South, Wiggins clearly conveyed an understanding of these Black collegiate women's perceptions by drawing on literature addressing the historical foundations of colorism, impacts of colorism on education, and the psychological impacts of colorism on perceptions of intelligence, self-worth, and self-efficacy. Wiggins used critical race theory and Black feminist thought—which combined, inform critical race feminisms—to acknowledge the participants' personal experiences, while focusing on the significant impacts of childhood trauma, overt racism, and the reality of being Black. Her study ultimately created an enhanced appreciation of Black collegiate women's personal experiences, challenges, and perspectives as they navigate a predominantly white urban public state institution in the South. Wiggins (2021) noted that as foundational understandings and significances of the impact of colorism are established, this should inspire critical discussions within academic settings to encourage the exploration and collaboration to end colorism.

Implications of CRT for future educational scholarship

While CRT has emerged as a powerful framework for understanding educational issues, it is yet to be fully realized. Scholars who employ a CRT framework most notably utilize the concept of permanence of racism and voice scholarship through counter-storytelling. For example, voice scholarship has included examining the experiences of students of color as well as the inclusion of scholars of color in theorizing space. However, the presence of an internal struggle over the materialist concerns of CRT has pressing implications for CRT scholars. Namely, it points to CRT inquiry, which has transcended from its material origins to its ideological concerns. Yet, these two forms of inquiry are inextricably connected, merely posing as two sides of the same coin. For example, while Crenshaw (1995) acknowledged identities as ideological in nature, she argued their consequences on lived experiences necessitate their consideration, in order to explicate and positively change one's marginalized condition by embracing and enacting on one's identity. In other words, although race is socially constructed and not a social fact in and of itself, the material consequence of race conveys its social significance. Therefore, challenging both the social construction of race and the enactment of

power based on race is significant. Still, the concern that scholarship in which CRT engages may never reach those who are most affected, the very populations for which CRT is meant to account for (Ladson-Billings, 1999), exists as a valid concern in which scholars must consciously engage themselves. One challenge then is in examining and asserting how ideological struggles are linked to material consequences. Harris' (1995) concept of whiteness as property, further articulated by Ladson-Billings and Tate (1995), presented the opportunity to identify (ideological and structural) processes that consequently result in concrete differences in schooling opportunities. Yet, it has been underutilized, presenting a gap for future work to fill.

Additionally, Bell's (1980) concept of interest convergence also has important implications for understanding numerous issues, including the discourse centered on Asian Americans, especially within the context of current affirmative action debates. A CRT approach which recognizes and examines how discursive practices construct, impose, and empower identity as well as shape definitions of minority, diversity, or social justice are important for understanding contemporary debates over the allocation of college-admission slots. First, it can challenge the social construction of race and further examine how particular definitions of race are privileged in certain instances and used to inform a limited definition and application of diversity that ultimately support conservative agendas (Omi and Winant, 1994). A CRT perspective can also elucidate the larger social context in which education issues are understood, explicating how international relations often shape experiences of Asian Americans and people of color in the US that ultimately support conservative agendas (Omi, 2008). For example, the perception of a peril of Asian Americans invading and overrunning higher education institutions, especially in light of recent federal court decisions involving race-conscious admissions at Harvard University and University of North Carolina, requires a critical race approach to elucidate that this perception mirrors the larger societal sentiments seen in international relations (Jaschik, 2020; Liptak and Hartocollis, 2022). Finally, CRT and interest convergence can further aid, particularly in understanding how, despite these sentiments, whiteness can be thrust upon Asian Americans (Omi, 2008; Liptak and Hartocollis, 2022) when it aids in the larger social and political interests of whites. The case of Asian Americans serves as just one example of how CRT theoretical constructs can elucidate and explicate how ideologies of race and discursive practices have real consequences and are directly connected to material interests.

Conclusion

To conclude, education scholars employ CRT to understand school inequity and educational issues while advancing the evolution of CRT in significant ways. Given its origins in and departure from critical legal theories, aimed at understanding the interplay of race, racism, and the law, a critical race analysis is appropriate for the field of education on both philosophical and historical grounds. When considering that law was used to, first, ban African Americans and other marginalized groups from obtaining education; then, institute segregation and later, repeal initiatives such as affirmative action aimed at addressing the legacy of racism, the use of CRT in education is logical if the ultimate goal is to overcome societal inequity. While these reasons alone represent a compelling case for CRT, the need for a theory which advances race as an analytical tool for understanding issues in education treads much deeper. It calls into question the ideologies and jurisprudence which informs and sustains these laws. Much like the legal scholars of color who were instrumental in creating a CRT framework, the allure of CRT for scholars in education is the possibility for liberation. Specifically, we are drawn to the idea that although schools and schooling practices (curriculum, instruction, pedagogy, etc.) are sites and tools of social reproduction, we remain hopelessly optimistic in achieving a transformative education: one that is grounded in our experiences, inclusive of our cultural knowledge, and intentionally fosters social praxis for the emancipation of marginalized communities. For these reasons, scholars of color in education find solace in the discontent, struggle, and discourse of legal scholars of color within their field, (not merely in their collective identity as people of color, but more principally) in their criticisms, yet unyielding commitment in the ideals of and belief in the very institution in which they hope to transform (theoretically, methodologically, and practically). Finally, and perhaps most important, CRT challenges scholars to interrogate their own socialization into the academy and confront their possible roles in reproducing dominant discourses. Therefore, CRT has far-reaching implications for theory, research, and practice in education. Yet, recent attacks on CRT bear close watching coupled with strategic, constructive, and bold responses to preserve its integrity and frankly, its survival, as a relevant theoretical framework in educational research utilizing qualitative methods.

References

- Aleinikoff, A.T., 1991. A case for race-consciousness. *Columbia Law Rev.* 91, 1060–1123.
- Associated Press, 2022. Florida Bill Bans Making White People Feel "Discomfort" in Racial Discrimination Training. *Al.com*. <https://www.al.com/news/2022/01/florida-bill-bans-making-white-people-feel-discomfort-in-racial-discrimination-training.html>. (Accessed January 2022).
- Bakke, V., 1978. Regents of the University of California v. Bakke, 438 U.S. 265.
- Beauchamp, Z., 2021. Did Critical Race Theory Really Swing the Virginia Election? *Vox*. <https://www.vox.com/policy-and-politics/2021/11/4/22761168/virginia-governor-glenn-youngkin-critical-race-theory>. (Accessed January 2022).
- Bell, D.A., 1976. Serving two masters: integration ideals and client interests in school desegregation litigation. *Yale Law J.* 84, 470–516.
- Bell, D.A., 1980. Brown v. Board of Education and the interest convergence dilemma. *Harv. Law Rev.* 93, 518–533.
- Bell, D.A., 1987. *And We are Not Saved: The Elusive Quest for Racial Justice*. Basic Books, New York.
- Bell, D.A., 1992. *Faces at the Bottom of the Well: The Permanence of Racism*. Basic Books, New York.
- Bergerson, A.A., 2003. Critical race theory and white racism: is there room for white scholars in fighting racism in education? *Qual. Stud. Educ.* 16, 51–63.
- Bonilla-Silva, E., 2013. *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*. Rowman and Littlefield, Lanham, MD.

- Brownstein, M., 2019. Implicit Bias. In: Zalta, E.N., (ed.), Stanford Encyclopedia of Philosophy. Retrieved on December 10, 2020 from <https://plato.stanford.edu/archives/fall2019/entries/implicit-bias>.
- Chickering, A.W., Reisser, L., 1993. Education and Identity, second ed. Jossey-Bass, San Francisco, CA.
- Christian, M., Seamster, L., Ray, V., 2019. New Directions in Critical Race Theory and Sociology: Racism, White Supremacy, and Resistance. *American Behavioral Scientist*, pp. 1–10.
- Coghlan, D., Brydon-Miller, M. (Eds.), 2014. The Sage Encyclopedia of Action Research. Sage, Thousand Oaks, CA.
- Collins, P.H., 1996. What's in a name? Womanism, Black feminism, and beyond. *Black Scholar* 26, 9–17.
- Collins, P.H., 2000. Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment, second ed. Routledge, New York.
- Collman, A., 2021. Teachers Say the Conservative Backlash to Critical Race Theory is Overblown, But Fear It Will Have a "Chilling Effect" on Education. *Insider*. <https://www.insider.com/critical-race-theory-debate-teachers-fear-chilling-effect-on-education-2021-8>.
- Crenshaw, K., 1988. Race, reform, and retrenchment: transformation and legitimation in antidiscrimination law. *Harv. Law Rev.* 101, 1331–1387.
- Crenshaw, K., 1989. Mapping the margins: intersectionality, identity politics and violence against women of color. In: Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York, pp. 357–383.
- Crenshaw, K., Gotanda, N., Peller, G., Thomas, K., 1995. *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York.
- Dalton, H.L., 1995. The clouded prism: minority critique of the critical legal studies movement. In: Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York, pp. 80–84.
- Darling-Hammond, L., 2001. New standards and old inequalities: school reform and the education of African-American students. *J. Negro Educ.* 69, 263–285.
- DeCuir-Gunby, J.T., Chapman, T.K., Schutz, P.A., 2019. Understanding Critical Race Research Methods and Methodologies: Lessons From the Field. Routledge.
- Delgado, R., 1984. The imperial scholar: reflections on a review of civil rights literature. *Univ. Penn. Law Rev.* 132, 561–578.
- Delgado, R., 1989. Storytelling for oppositionists and others: a plea for narrative. *Mich. Law Rev.* 46, 665–676.
- Delgado, R., 1992. The imperial scholar revisited: how to marginalize outside writing, ten years later. *Univ. Penn. Law Rev.* 140, 1349–1372.
- Delgado, R., Stefancic, J., 1993. Critical race theory: an annotated bibliography. *Va. Law Rev.* 79, 461–516.
- Delgado, R., Stefancic, J., 2001. *Critical Race Theory: An Introduction*. New York University Press, New York.
- Delgado, R., Stefancic, J., 2017. *Critical Race Theory: An Introduction*, third ed. New York University Press, New York.
- Dumas, M.J., Ross, K. M., 2016. "Be real Black for me": imagining BlackCrit in education. *Urban Educ.* 51, 415–442.
- Files, M.D., 2021. Mothering Ourselves to Wholeness: Employing Black Feminist Criticism and Critical Narrative Methods to Trace Black Maternal Archetypes and Restore the Praxes of Socio-politically Engaged Black Mothers in Rural, Suburban, and Urban Contexts (Order No. 28867079). Available From Dissertations and Theses @ University of Alabama at Birmingham. (2622294739).
- Freeman, A.D., 1995. Legitimizing racial discrimination through antidiscrimination law: a critical review of Supreme Court doctrine. In: Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York, pp. 29–45.
- Glenn, E.N., 2008. Yearning for lightness: transnational circuits in the marketing and consumption of skin lighteners. *Gen. Soc.* 22, 281–302.
- Harper, S.R., 2009. Niggers no more: a critical race counternarrative on Black male student achievement at predominantly white colleges and universities. *Int. J. Qual. Stud. Educ.* 22, 697–712.
- Harris, C.I., 1995. Whiteness and property. In: Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), *Critical Race Theory: The Key Writings That Formed the Movement*. The New Press, New York, pp. 357–383.
- Harvard Law Review, 2020. First Amendment: Religion the Establishment Clause and the Chilling Effect. <https://harvardlawreview.org/2020/02/the-establishment-clause-and-the-chilling-effect/>.
- Hirald, P., 2010. The role of critical race theory in higher education. In: *The Vermont Connection*, vol. 31, p. 7.
- Hooks, B., 1981. *Ain't I a Woman? Black Women and Feminism*. South End Press, Boston, MA.
- Hunt, A., 1986. The theory of critical legal studies. *Oxf. J. Leg. Stud.* 6, 1–45.
- Jaschik, S., 2020. Appeals Court Backs Harvard on Affirmative Action. *Inside Higher Ed*. <https://www.insidehighered.com/admissions/article/2020/11/16/appeals-court-backs-harvard-affirmative-action>.
- Jones, C.E., 2021. Man Enough: Exploring the Gendered Experiences of High-Achieving Black Men in Higher Education. Order No. 28644773. The University of Alabama at Birmingham. Ann Arbor: ProQuest. Web. 21 Jan. 2022.
- Ladson-Billings, G., 1999. Just what is critical race theory, and what's it doing in a nice field like education? In: Parker, L., Doyle, D., Villenas, S. (Eds.), *Race is ... Race Isn't: Critical Race Theory and Qualitative Studies in Education*. Westview Press, Boulder, CO, pp. 7–30.
- Ladson-Billings, G., Tate IV, W.F., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97, 47–68.
- Ladson-Billings, G., Tate, W.F., 2006. Toward a critical race theory of education. In: Dixon, A.D., Rousseau, C.K. (Eds.), *Critical Race Theory in Education: All God's Children Got a Song*, first ed. Routledge, New York, pp. 11–30.
- Lawson, R.J., 1995. Critical race theory as praxis: a view from outside to outside. *Harv. Law J.* 38, 353–370.
- Ledesma, M.C., Calderon, D., 2015. Critical race theory in education: a review of past literature and a look into the future. *Qual. Inq.* 21, 206–222.
- Liptak, A., Hartocollis, A., 2022. Supreme Court will hear challenge to affirmative action at Harvard and U.N.C. *The New York Times*. <https://www.nytimes.com/2022/01/24/us/politics/supreme-court-affirmative-action-harvard-unc.html>. (Accessed June 2022).
- Lorde, A., 1984. *Sister Outsider*. Ten Speed Press, New York.
- Matsuda, M., 1995. Looking to the bottom: critical legal studies and reparations. In: Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York, pp. 63–79.
- Matsuda, M., 2002. Beyond, and not beyond, black and white: deconstruction has a politics. In: Valdes, F., Culp, J.M., Harris, A.P. (Eds.), *Crossroads, Directions, and a New Critical Race Theory*. Temple University Press, Philadelphia, PA, pp. 393–398.
- Milner, H.R., 2007. Race, culture, and researcher positionality: working through dangers seen, unseen, and unforeseen. *Educ. Res.* 36, 388–400. <https://doi.org/10.3102/0013189X07309471>.
- NAACP Legal Defense Fund. Critical Race Theory: Frequently Asked Questions, <https://www.naacpldf.org/critical-race-theory-faq/>, Accessed January 2022.
- Omi, M., 2008. Asian-Americans: the unbearable whiteness of being? *Chron. High. Educ.* 55. <http://chronicle.com/article/Asian-Americans-the/14271>. (Accessed October 2009).
- Omi, M., Winant, H., 1994. *Racial Formation in the United States: From the 1960s to the 1990s*, second ed. Routledge, New York.
- Parker, L., 2015. Critical race theory in education and qualitative inquiry: what each has to offer each other now? *Qual. Inq.* 21, 199–205. <https://doi.org/10.1177/1077800414557828>.
- Parker, L., Lynn, M., 2002. What's race got to do with it? Critical race theory's conflicts with and connections to qualitative research methodology and epistemology. *Qual. Inq.* 8, 7–22. <https://doi.org/10.1177/107780040200800102>.
- Patton, L.D., McEwen, M., Rendon, L., Howard-Hamilton, M., 2007. Critical race perspectives on theory in student affairs. In: Harper, S.R., Patton, L.D. (Eds.), *Responding to the Realities of Race on Campus, New Directions for Student Services*. Jossey-Bass, San Francisco, CA, pp. 39–53. No. 120.
- Sawchuk, S., 2021. What is critical race theory, and why is it under attack? *Education Week*. <https://www.edweek.org/leadership/what-is-critical-race-theory-and-why-is-it-under-attack/2021/05>.
- Schwartz, S., 2022. Map: where critical race theory is under attack. *Education Week*. <https://www.edweek.org/policy-politics/map-where-critical-race-theory-is-under-attack/2021/06>.

- Smith-Maddox, R., Solórzano, D.G., 2002. Using critical race theory, Paulo Freire's problem-posing method, and case study research to confront race and racism in education. *Qual. Inq.* 8, 66–84.
- Solórzano, D.G., Yosso, T.J., 2001. From racial stereotyping and deficit discourse toward a critical race theory in teacher education. *Multicult. Educ.* 9, 2–8.
- Solórzano, D.G., Yosso, T.J., 2002. A critical race counter story of race, racism and affirmative Action. *Equity Excell. Educ.* 35, 155–168.
- Solórzano, D.G., Ceja, M., Yosso, T.J., 2000. Critical race theory, racial microaggressions, and campus racial climate: the experiences of African American college students. *J. Negro Educ.* 69, 60–73.
- Strauss, A.L., Corbin, J.M., 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage Publications, Thousand Oaks.
- Tate, W.F., 1997. Critical race theory and education: history, theory, and implications. *Rev. Res. Educ.* 22, 195–247.
- Villenas, S., Deyhle, D., 1999. Critical race theory and ethnographies challenging the stereotypes: Latino families, schooling, resilience, and resistance. *Curric. Inq.* 29, 413–445.
- Walker, A., 1983. *In Search of Our Mothers' Gardens: Womanist Prose*. Harcourt Brace Jovanovich, San Diego, CA.
- Wallace-Wells, B., 2021. How a conservative activist invented the conflict over critical race theory. *The New Yorker*. <https://www.newyorker.com/news/annals-of-inquiry/how-a-conservative-activist-invented-the-conflict-over-critical-race-theory>. (Accessed January 2022).
- White, G., 1972. From sociological jurisprudence to realism: jurisprudence and social change in early twentieth-century America. *Va. Law Rev.* 58, 999–1028.
- Wiggins, C.N., 2021. *Shades of Black: Black Collegiate Women and the Impact of Colorism in Higher Education* (Publication No. 28860368) [Doctoral Dissertation, University of Alabama at Birmingham]. ProQuest Dissertations and Theses Global.
- Yosso, T.J., 2005. Whose culture has capital? A critical theory discussion of community cultural wealth. *Race Ethn. Educ.* 8, 69–91.

Further reading

- Bernal, D.D., 2002. Critical race theory, Latino critical theory, and critical raced-gendered epistemologies: recognizing students of color as holders and creators of knowledge. *Qual. Inq.* 8, 105–126.
- Bonilla-Silva, E., 2017. *Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in the United States*, fifth ed. Rowman and Littlefield, Lanham, MD.
- Cannella, G.S., Steinberg, S.R. (Eds.), 2012. *Critical Qualitative Research Reader*. Peter Lang, New York.
- Crenshaw, K., Gotanda, N., Peller, G., Thomas, K. (Eds.), 1996. *Critical Race Theory: The Key Writings That Formed the Movement*. New Press, New York.
- Delgado, R. (Ed.), 1995. *Critical Race Theory: The Cutting Edge*. Temple University Press, Philadelphia, PA.
- Dixon, A.D., Ladson-Billings, G.J., Suarez, C.E., Trent, W.T., Anderson, J.D., 2020. Condition or Process? Researching Race in Education. American Educational Research Association, Washington, D. C.
- Feagin, J.R., 2006. *Systemic Racism: A Theory of Oppression*. Routledge, New York.
- Hannah-Jones, N., 2021. *The 1619 Project: A New Origin Story*. The New York Times Company, New York.
- Harper, S.R., Patton, L.D., 2007. Responding to the Realities of Race on Campus *New Directions for Student Services*, No. 120. Jossey-Bass, San Francisco, CA.
- Kendi, I.X., 2016. *Stamped From the Beginning: The Definitive History of Racist Ideas in America*. Nation Books, New York.
- Kendi, I.X., 2019. *How to Be an Antiracist*. One World, New York.
- Ladson-Billings, G., 2021. *Critical Race Theory in Education: A Scholar's Journey*. Teachers College Press, Columbia University, New York.
- Lynn, M., Dixon, A.D. (Eds.), 2013. *Handbook of Critical Race Theory in Education*. Routledge, New York.
- Lynn, M., Yosso, T.J., Solórzano, D.G., Parker, L., 2002. Critical race theory and education: qualitative research in the new millennium. *Qual. Inq.* 8, 3–6.
- National Commission on Asian American and Pacific Islander Research in Education, 2008. *Asian Americans and Pacific Islanders. Facts, Not Fiction: Setting the Record Straight*. The College Board, New York.
- Parker, L., Deyhle, D., Villenas, S. (Eds.), 1999. *Race is ... Race Isn't: Critical Race Theory and Qualitative Studies in Education*. Westview Press, Boulder, CO.
- Solórzano, D.G., 1998. Critical race theory, race and gender microaggressions, and the experience of Chicana and Chicano scholars. *Qual. Stud. Educ.* 11, 121–136.
- Solórzano, D.G., Yosso, T.J., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8, 23–44.
- Taylor, E., Gilborn, D., Ladson-Billings, G., 2009. *Foundations of Critical Race Theory in Education*. Routledge, New York.
- Wilkerson, I., 2020. *Caste: The Origins of Our Discontents*. Random House, New York.

Queer theory and qualitative research

Boni Wozolek, Penn State University, Abington College, Abington, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

The complexities of “queer”?	79
Queering qual: the re-mapping of research	80
Conclusion: what’s next?	82
Acknowledgment	83
References	83

Looking inward and outward

at once before and after

seeking a now that can breed futures

like bread in our children’s mouths

so their dreams will not reflect

the death of ours

Lorde (1978/1997, p. 235).

To begin an ethnographic project with a goal, with an object of research and a set of presumptions, is already to stymie the process of discovery; it blocks one’s ability to learn something that exceeds the framework with which one enters.

Halberstam (2011, p. 12).

There is no one way to “queer” research for, as queer theorists have long argued, the process of queering is fluid and liminal (Butler, 1993; Cohen, 2000; Halperin, 1995; Johnson, 2016). To “queer” research is a commitment to contradictions that recognizes the multiplicity of experiences, and troubles sociocultural, political, and historical norms and values (Corber and Valocci, 2003; Epstein, 2005; Jagose, 1996). Queering research is therefore a critical entanglement with making the strange familiar and the familiar strange (Johnson, 2001; Spindler and Spindler, 1982) in that it is omnidirectional—at once looking reflexively, as Lorde (1978/1997) argues above, while seeking to exceed (Halberstam, 2011) and abolish (Love, 2019) oppressive frameworks across scholarly dialogs.

Both queer theory and qualitative research have longstanding and complex histories. It is therefore impossible in one encyclopedia chapter to capture queer theory and qualitative research within their respective entireties. Queer theory and its relationships to communities and people who identify as either lesbian, gay, bisexual, transgender, queer/questioning, intersex, asexual, two spirit, or some combination of these identities (LGBTQIA2S++) are, for example, intricate due to multifaceted identities that exist within and across queer people, communities, and places. It is important to pause at this juncture and note three points. First, the use of “relationship” is intentional as too often when one thinks through, about, and with queerness, the idea is frequently ascribed to/with queer identities. The inclusion of the term “relationships” is to intentionally consider that one does not need to be queer to have a relationship with queerness or to be in a queer relationship. To assume one identifies within and/or across the LGBTQIA2S++ community may reinscribe the very cishetero norms and values that queer theories are meant to interrupt. Second, throughout this chapter, the author uses “queer” to broadly describe people across LGBTQIA2S++ communities as a generally inclusive term. However, while used in the spirit of inclusion, it should be noted that any umbrella term can essentialize people and falsely sew together identities that are not necessarily related to how one is identified or aligned with how they might identify themselves.

Finally, rather than using terms like “homophobia,” “transphobia,” and the like, this chapter will use terms like “prejudice” and “bias” to describe the negative, often violent, attitudes and feelings espoused by people who dislike and dehumanize queer

people. For example, the Diagnostic and Statistical Manual of Mental Disorders (DSM) in places like the United States was historically used to maintain marginalization of queer people by pathologizing sexual orientations and gender identities as disorders that, in turn, allowed family, friends, and agents of the state to forcibly institutionalize both queer identifying and queer identified people. Although the term “homosexuality” was removed from the DSM-2 in 1973, the association between queer people and a diagnosis of being “disturbed” had already caused considerable cultural damage. For example, as of 2021, there are still several states in the United States that have sodomy laws that, although not enforced, are still legal doctrines dehumanize and despoil queer energies, rendering them “illegal” and equating them with rape, bestiality, and pedophilia (Dixit, 2020). The DSM does not, however, include prejudice against minoritized people and communities as a diagnosable disorder. The characterization of “phobias” in the DSM (e.g., animal, situational, environmental, or bodily phobias are the main phobia categories in the DSM-5), overlooks how the fear of people can be psychologically detrimental. Further, phobias in the DSM are pathologized by considering an individual’s experience with diagnoses like arachnophobia, rather than considering broader sociocultural norms around spiders. The use of “bias” or “prejudice” is intended to simultaneously attend to these histories by considering how biases are at once rooted in individual experiences while acknowledging how they are enmeshed with sociopolitical and cultural norms.

Similar to the complex nature of queer theory, qualitative research is knotted (Nespor, 1997), with subsets of methodological approaches and theoretical frameworks. Further, there are varying positions and positionalities on both queer theories and qualitative research that are contextually dependent. For example, thinking about how these fields feed each other in a nation where sexual and gender prejudice often galvanize violence and encourage social and political exclusion might be a very different dialog than thinking about a nation that seeks inclusion across identities. Rather than working impossibly to not get it all wrong, as Wolcott (1994) might argue, this chapter will first trace the contours of queer theories before turning to some significant, broadly painted, dialogs within qualitative research without centering on one particular methodology. Finally, this chapter will turn to the imbrications of queer theory and qualitative research before concluding with some important implications that arise from the intersections of scholarly dialogs.

The complexities of “queer”?

In a contemporary sense, the term “queer” is rather fluid. In the English language, it can be used as a noun, and adjective, or a verb (It is important to acknowledge that queer understandings are not the same across generations, languages, and cultures. While English has been emphasized here, the author would also like to note the diversity even among English language speakers when it comes to cultural and often generational understandings of “queer.” The point is to, on one hand, acknowledge the author’s positionality as someone who, while bilingual, is writing in English and speaking from a Brown, Generation Y, and queer perspective within the United States. On the other, this is an attempt to acknowledge the often-monolithic dialogs on queerness that can come from English-normative ways of knowing, being, and doing). As a verb it can mean to look at something through a different lens; to mobilize bodies—both human and nonhuman—in ways that trouble or unsettle assumptions (Spargo, 1999). “Queer” as an adjective not only describes but can also be understood within its knotted nature of re-claiming queer-ness as both adjective and noun—or how one identifies versus how one can be identified (Johnson, 2001; Rand, 2014). Yet, the term “queer” remains generationally and contextually based in ways that are significant. In addition, as mentioned in the introduction, while queer is often a term used to broadly describe the LGBTQIA2S++ community, the use of an umbrella term can be incidentally marginalizing, reinforcing the very norms it is meant to disrupt. For example, it is not unusual to hear an individual referred to as “LGBTQ.” This assumes one individual identifies within the entire community, unintentionally flattening the diversity of experiences and perspectives across LGBTQIA2S++ communities. Queer theory is similarly fluid, wrought with potential pitfalls of essentialization, and can be understood as an engagement and entanglement with and across critical practices and dialogs. Queer theorists have—to name a few complicated and interrelated conversations—argued the necessity of queer bodies in destabilizing systems of power, queering as a way to re-consider performance and performativity (e.g., Muñoz, 2006; Sedgwick, 2003), and queer frameworks as significant in re-reading texts (Butler, 1993; Norton, 2016), across their broad conceptualizations.

Part of queer theory’s fluidity lies in the fact that some scholars who are considered queer theorists were retroactively brought into the field after Teresa de Lauretis (1991) coined the term. De Lauretis initially used “queer theory” as a provocation that was meant to unsettle gay and lesbian studies as they existed at the time. Halperin (2008) argues that queer theory was meant provoke dialogs that would encourage scholars to make “theory queer (to challenge the heterosexist underpinnings and assumptions of what conventionally passed for ‘theory’ in academic circles) and to queer theory (to call attention to everything that is perverse about the project of theorizing sexual desire and sexual pleasure)” (p. 340). De Lauretis’ call to the fields of gay and lesbian studies was, from her perspective, insufficiently answered and, barely 3 years after the initial conference where she introduced the term, she stepped back from the field, explaining that it had been picked up in ways that only reified the very institutions it was meant to resist and interrupt (de Lauretis, 1994).

Despite de Lauretis’ movement away from the field, the project of continuing to define queer theory continued. Although scholars worked to retroactively trace the contours of queer theory through its sociohistorical, cultural, and political implications (e.g., Berlant and Warner, 1995), others used seminal work to think about how the field might be shaped (see, for example, Butler’s (1990) *Gender Trouble*, Sedgwick’s (1990) *Epistemology of the Closet*, or Rich’s (1980) *Compulsory Heterosexuality and Lesbian Existence*).

Consequently, this process of tracing and shaping the field left an “undesirable and misleading effect of portraying all previous work in lesbian and gay studies as under-theorized ... narrowing the scope of queer studies by privileging its theoretical register, restricting its range, and scaling down its interdisciplinary ambition” (Halperin, 2008, p. 341).

Through the process of de- and re-territorializing the field of queer theory, a space for critique of the field itself among queer theorists was constructed. These critiques included scholars considering the voices present and those (un)intentionally silenced. This included dialogs on genders in their multiple forms and expressions (e.g., Halberstam, 1998), sexual orientations (e.g., Dilley, 1999), races (e.g., Johnson, 2001), and, later on, intersex bodies and beings (Morland, 2009), and dis-abilities (McRuer, 2006), to name a few. The field also wrestled with its dis-connection with queer activism. Activist groups, from the Mattachine Society in the United States (established in 1950), to the United Kingdom’s Stonewall (established in 1989), to Italy’s Circolo di cultura omosessuale Mario Mieli (established in 1983), to Iceland’s Samtökin (established in 1978), among others, existed long before queer theory was codified in academic circles. This scholar-activist split was, as scholars have argued, potentially detrimental to early policy work across global and local contexts (Rand, 2014). Further, queer politics tend to be organized around the narrow notion that queer activism was a direct response to Reagan-era responses to the HIV/AIDS epidemic.

In the United States there is truth to the galvanizing response that Reagan-era politics created in queer movements. However, it is equally important to recognize radical global movements against racism, sexism, and ableist ideologies that were central to queer activist work that has also become central to scholarly dialogs. As Cohen (2019) argues, forgetting this can cause an erasure that “inhibits our ability to know what happened before us” (p. 141). Cohen continues by explaining there were, for example, several Black radical individuals like Audre Lorde, Marlon Riggs, Cheryl Clarke, and Pat Parker who were active in documenting, fighting, and providing organizational structure to perusing liberation and drawing attention to anti-queer inequities. The work of these activists has also become central to the theorization of queer theory as it exists in contemporary scholarly conversations (e.g., Mayo and Rodriguez, 2019; Pinar, 2012; Rodriguez et al., 2016).

The troubling of essentialist and limiting notions within the field has been significant in how scholars have attended to both political implications of the field and the broader cultural impacts of the field. Specifically, there has been an attention to how queer theory has been defined, mapped, and de/territorialized within the academy and its broader communities (Talbot and Steinberg, 2000). This had led to more complex theorizations of intersectionality (Crenshaw, 1991; Jones and Abes, 2013) within the field. For example, in 2001, E. Patrick Johnson critiqued queer theory for suppressing dialogs of race and class. Johnson proposed “quare theory” as one way to actively attend to the political power inherent in racist and anti-queer spaces and places. Similarly, decolonial and indigenous scholars have proposed a queer-decolonial lens (e.g., Driskill, 2011; Pereira, 2015; Singh, 2017) as a way to create and deepen a presence of GLBTQ2 people and communities in queer theory and broader policy dialogs. Finally, scholars attending to thin-normativity have used Sedgwick’s (1993) call for queer theory to provide a “spin outward along dimensions that cannot be subsumed under gender or sexuality” (p. 9) to include voices on fat-studies, a metaphorical queering of the fat body (e.g., Pausé et al., 2016).

While these movements represent a sliver of the critical conversations within the field as it continues to touch and be touched across scholarly work, it is important to note that inherent nature of the field continues to open up spaces of possibility within critique and re-thinking what it means to be a “queer theorist” and engage in queer theory. This chapter will now turn to qualitative research as it is entangled with queer theory before concluding with implications for the field(s).

Queering qual: the re-mapping of research

Broadly, the field of qualitative research has worked with/in tensions, hesitations, and contradictions (Denzin and Lincoln, 2008). Within the field, researchers have attempted to decolonize the process and presentation of collecting and disseminating empirical data (Smith, 2013) while attending to issues like the inherent racialized (e.g., Solórzano, 1998; Steinberg and Cannella, 2012), gendered (e.g., Ribbens and Edwards, 1997; Rogers, 2020), ableist (e.g., Annamma, 2016; Lester and Nusbaum, 2018), linguistic (e.g., Heigham and Croker, 2009), and heteronormative (e.g., Spencer et al., 2014) ways of being, doing, and knowing that were normalized early on within qualitative research. In addition to these internal tensions, qualitative research also has longstanding histories of interrupting norms centered on quantitative research that rank numerically driven research as superior, often arguing that it is a more “valid” form of data collection and dissemination of ideas (Behar and Gordon, 1995; Denzin and Lincoln, 2008).

Much like queer theory, qualitative research had a relationship with queer scholars, queer communities, and queer/d ideas prior to the field being named and recognized across academic disciplines. For example, the 1990s includes a rather extensive list of scholars who were working with/in queer scholarship (e.g., Britzman, 1995; Kumashiro, 1999; Jagose, 1996; Lather and Smithies, 1997). While these scholars might not have necessarily identified themselves as “queer theorists” at the time, their work has been identified by the field as significant in forming and informing the work of queer theory and qualitative research. Similarly, as with all qualitative research, queer(d) narratives are often analyzed though work that, while not qualitative inquiry, speaks to and across fields of empirical research. This is significant in that queer theory informs qualitative researchers in the same way that qualitative research can inform queer theory. This reciprocal relationship between queer theory and qualitative research is central to, as Kincheloe and McLaren (2011) argue, unsettling marginalizing conceptual frameworks through the nexus of critical theories and qualitative research. Queer theory can, for example deconstruct ideas and ideals within qualitative research while qualitative research can ground queer theory in material concerns of cishetero oppressions.

One important factor to the growth of queer voices and narratives in qualitative research lies within the increased visibility of LGBTQIA2S++ people and communities across contexts (Sears, 2005). Although telling vulnerable stories is always a brave act, there is something significant to having access to queer experiences and, especially for researchers to have those narratives valued across academic spaces and places (Levy and Johnson, 2012). This is not to say that those who do queer qualitative work do not still face oppressions in the form of gaslighting (Wozolek, 2018), being dismissed (Allen and Rasmussen, 2015) and facing normalized dehumanization (Kessi, 2018) across higher education. Yet, the ground that forefolx paved in terms of liberatory policies and practices has made studying and reporting queer narratives through qualitative research far more accepted and, in certain contexts, highly valued, than it was when queer theory was in its infancy.

Queer theory as it is related and, for some scholars, integral to qualitative research is not as much a methodological turn but an attention to how traditional methodologies can be queered. For example, Lambevski (1999) argued that researchers are not often transparent about the tensions of personal experiences and identities as they intersect with one's research. While this might seem as incredible in a contemporary era where situating oneself is expected in qualitative research, and vulnerability has become a well-trodden dialog (e.g., Behar, 1996; Geertz, 1973; Rosaldo, 1993), it is significant to recall that despite white male privilege as an academic, scholars like Humphreys (1970) were unable to write about their own sexuality in their research. Although, as Plummer (2008) points out, despite the absence of identity in Humphreys' work, there is a footnote about the taste of semen so, as many queer authors do, Humphreys leaves a trail of breadcrumbs toward queer understandings and identities. The work of scholars like Humphreys paved the way for others, like Miller (1998) whose part qualitative research, part autobiographical work is intensely personal as he discusses gay culture in Broadway. Others, like Weinberg's (1995) edited collection that focuses on bondage, discipline, dominance and submission, and sadomasochism (BDSM) culture and, while now slightly dated, set the stage for sociological rather than psychological conversations about kink-positive culture and identities (e.g., Carlström, 2019; Fennell, 2018) through qualitative research and other significant theoretical frameworks.

Along similar lines, queer theory offers a critical lens through which foundational methodological processes are carried out. Regardless of one's identity, or how they are identified, using queer theory as a lens through which one conducts qualitative research can, and to some degree should, deeply impact a method. Wells et al. (2012), use fieldnotes as but one example of the effect queer theory might have on ethnographic research. "Queer theory," they argue, "asks the researcher to put power relations ... in the forefront of her mind not only during analysis, but also during data collection. The fieldnotes, then, should include information about actions, dialogs, events, and so on in relation to this axis of social relations" (p. 121). Although the authors foreground sexuality, it is important to understand the move in contemporary queer theory to, as Sedgwick (1993) argues, allow critical queer analysis to "spin outward along dimensions" (p. 9) to include queering qualitative inquiry across its methodological frameworks and processes.

It should be noted that while it's significant to have representation in terms of ideas and voices, there are times when identity politics have been a detriment to who can collect queer narratives and who is seen as a person who can theorize them ways that are recognized as "valid." For example, Youdell (2009) illustrates this tension between perceived researcher identity and local actors in her article "Queer Outings" in which she discusses a context where students read her involvement with a boyfriend as a claim to heterosexuality. Although bisexual, and a person who rejects the idea of the closet, she recognized how her "passing" as straight impacted her research and relationship with participants. Other scholars have argued that prevalent identity politics that dictate straight researchers are unable to participate in anti-normative scholarship detrimental (Allen, 2010). Although, to be clear, it is important in contexts like Youdell's research for students to understand queerness as normal, and researcher identity matters, this brings into focus how scholars might take up Britzman's (1995) dialog to "stop reading straight" in ways that both acknowledge the significance of identities while resisting a kind of identity politics that can reinforce the very kinds of exclusion that queer theory is meant to interrupt, particularly as it intersects with qualitative research.

It should be noted that while Britzman called for scholars to attend to what it might mean to "read straight," this does not mean that cishetero identifying people cannot engage in queer theory, qualitative research done through a queer theory lens, or with queer identifying participants. One can argue that contemporary researchers often find themselves in a conundrum driven by identity politics. Qualitative research has histories of marginalizing people and communities as, historically, straight-presenting, white people would exoticize and Other people of color, economically disenfranchized people, queer communities, and the like. One only needs to reflect on early ethnographic work to understand how othering was normalized and, just as importantly, how ethnographers like Geertz (1973), Behar (1996), and Agar (1996), among others, attended to these normalized forms of dehumanization in their research as an attempt to ethically (re)orient qualitative research. Inherent in these dialogs are questions of voice and transparency that are critical to qualitative research and, when implemented, can be considered a way of queering qualitative research and scholarly orientations that caused harm across people and contexts (Jones and Harris, 2018). This does not mean that one's identity need to align with the participants in a study but, rather, like all strong forms of qualitative research, that one's positionality should be clear in their writing, regardless with whom and where data was collected. To be clear, this is not a call for people to out themselves because it is important for any person to be able to reveal facets of their identity on their own terms. It is, however, a call for transparency about one's ethics and positionality.

While the idea of queering qualitative inquiry has underscored the significance of how people and communities are represented in research, it is important to remember that queer theory exists as a theoretical framework that acts as a lens through which qualitative research is carried out, analyzed, and presented. Scholars have used this lens as one way to foreground the political (Plummer, 2008) and to refuse traditional methods to propagate a "certain disloyalty to conventional disciplinary methods"

(Halberstam, 1998, p. 10). Though used frequently as a textual turn in analyzing films, literature, television, and the like (e.g., Doty, 1993; Sedgwick, 1985), others have used queering methodology as a way to subvert traditional qualitative scholarship that often silenced LGBTQIA2S++ people and communities—a way to galvanize queer folx and refuse ideals of fixed identities (Gamson, 1995; Quinn and Meiners, 2009; Sykes, 2011; Whitlock, 2013). What queer theory offers in terms of political ideas and ideals is yet another way to reject modernist tendencies to position the researcher as detached and objective (Wells et al., 2012). This is not unlike other forms of qualitative research in general and, specifically, ethnographic research and other forms of narrative inquiry that acknowledge subjectivities as central to qualitative inquiry as a whole. The queer ethos in qualitative research is therefore defined by (re)imaginings, (re)orientations, and challenges to dominant discourses.

Finally, queer theorists who work through qualitative lenses in general, and educational research specifically, have argued that queer theory as it intersects with qualitative research should emphasize the importance of destabilizing norms and values (e.g., Jagose, 1996; Rodriguez and Pinar, 2007). However, there is a tension between educational research as a way to understand, recognize, and affirm queer youth experience, and the essentialization that often happens when data is deemed generalizable (Chang, 2005) across contexts and experiences. To this point, Wells et al. (2012) argue that researchers often “actively, although unknowingly, engage in normalizing discourses that structure their social contexts ... the social therefore arises in people’s activities and through the ongoing and purposeful concerting and coordinating of those activities” (p. 118). Although it is significant to trouble the sociocultural practices that have normalized the closet (Kissen, 2002; Sedgwick, 1990), it is equally as significant that qualitative research does not use the space to reify normalized views of queerness or reinforce binaried ways of being, knowing, and doing.

Conclusion: what’s next?

Queerness can be understood as a space of possibility (Edelman, 1998; Weiss, 2016). To queer ideas means to engage critically; interpreting moments, thoughts, data, bodies, and the like from multiple directions and perspectives while rejecting traditional norms and values. It is troubling of sorts (Butler, 1990) that deconstructs binaries in and across cishetero patriarchies. Yet, queer theory is not siloed. It emerges with and from seeds sown in fields like Black Cultural Studies (e.g., Mercer, 1994), decolonial studies (e.g., Pérez, 1999), critical race theory (e.g., Harper et al., 1997), and indigenous studies (e.g., Driskill, 2011), among others. To engage with queer theory, then, does not mean one must identify with or have personal histories with queerness. Nor does it mean that a researcher is positioned within an “either/or” of theoretical frameworks. Instead, the richness of queer theory, like the richness of strong qualitative research, can exist in nested and layered identities and curiosities that are ethically attuned to queer lives.

The relationship between queer theory and qualitative research is significant for at least the following reasons. First, the tension between queering qualitative inquiry and not allowing such work to (re)create binaries can be understood as the very foundation queered qualitative research. Second, and related to the potential pitfalls of the first point, queer researchers have used their work to call for and create contexts that foreground equity and access (e.g., Love, 2019; Mayo, 2006; Tierney, 1997; Wozolek, 2018). While this orientation to action, what Lather (1986) discussed as catalytic validity, or what Wells et al. (2012) discuss as empowerment-as-action, has its limitations, it is significant in terms of how theory and practice can move not only the field but broader sociocultural and political norms. Third, queer theory as it relates to qualitative research is yet another tool for researchers to use in terms of reflexivity across researcher and participant identities and they are formed and inform their imbricated contexts.

Queer theory and qualitative research, as they exist separately and within an entanglement, continue to be challenged by forms of oppression that exist both within and outside of their respective and interrelated fields. There is however, as scholars have argued, always greater in-group than between group difference (Gershon, 2017; Tatum, 2003; Varenne and McDermott, 1998). Such differences tend to be sorted through oppression that helps gain cultural capital (Bourdieu, 1993; Freire, 1970; Wozolek, 2019) and maintains lines of power (Foucault, 1978). Within queer communities, this tends to be a focus on, for example, sexualities over genders or the silencing of queer Black, Indigenous, people of color in ways that narrow the field rather than allow it to exist as a pluralistic bloom space (Stewart, 2010) of possibility. As researchers continue to use queer theory and its interrelated narratives to disrupt anti-queer norms and values through qualitative research, the field itself must continue to disrupt the very lines of oppression that reinscribe marginalization. Such forms of oppression are often born within and across in-group narratives as they are imposed through out-group messages, politics, and cultural ideals. Similar to dialogs within critical race theory (e.g., Bell, 1995; Delgado and Stefancic, 2017) and critical race pedagogies (e.g., Freire, 1970), both in- and out-group marginalization only serve to reinforce power structures in ways that will allow oppression to continue largely unabated.

In addition, queer theory as a theoretical framework for qualitative research must also think critically about making sure that queer narratives themselves are diverse, not just in terms of sociocultural precepts but in terms of representations of the many conditions under which queerness exists. There are, for example, many necessary queer narratives about exclusion, silencing, and violence that happen when White, cishetero, and patriarchal ideas and ideals collide with queer ways of being, knowing, and doing (e.g., Brockenbrough, 2018; McCready, 2010), there remains a need for more narratives about queer excellence and accomplishment (Rosenberg, 2018).

Queer theory is certainly far from the only way that qualitative researchers can resist and refuse objectivity. However, it continues to be a powerful tool for researchers to explore spaces of possibility that can abolish oppressive norms and values that are present not only in everyday experiences but also throughout the academy (Allen, 2010; Dilley, 1999; Gilbert, 2014). Critical queer theory's attention to power and politics in terms of silencing across contexts continues to be one way that these fields are significant in creating and maintaining equity and access (Chang, 2005) to address normalized oppressions. Returning to the quotes by Halberstam (2011) and Lorde (1978/1997) that opened this chapter, queer(d) qualitative research is therefore significant in "looking inward and outward ... to seeking a now that can breed futures" (Lorde, 1978/1997, p. 235) while letting go of "presumptions [that] block one's ability to learn something that exceeds" (Halberstam, 2011, p. 12) the past, present, and projected futures that we have, do, and have not even begun to imagine.

Acknowledgment

My gratitude to my forefolx—those who came before me and not only whose work has been central in my intellectual growth but to those whose bodies and beings in the academy have made my ways of being, knowing, and doing as a Brown, queer woman possible in and across academic spaces and places. In addition, to the contemporary scholars who continue to be my chosen family, my continual appreciation and love.

References

- Agar, M.H., 1996. *The Professional Stranger: An Informal Introduction to Ethnography*, second ed. Academic Press.
- Allen, L., Rasmussen, M.L., 2015. Queer conversation in straight spaces: an interview with Mary Lou Rasmussen about queer theory in higher education. *High. Educ. Res. Dev.* 34 (4), 685–694.
- Allen, L., 2010. Queer (y) ing the straight researcher: the relationship (?) between researcher identity and anti-normative knowledge. *Fem. Psychol.* 20 (2), 147–165.
- Annamma, S.A., 2016. *DisCrit: Disability Studies and Critical Race Theory in Education*. Teachers College Press, New York, NY.
- Behar, R., Gordon, D.A. (Eds.), 1995. *Women Writing Culture*. University of California Press, Berkeley, CA.
- Behar, R., 1996. *The Vulnerable Observer: Anthropology That Breaks Your Heart*. Beacon Press, Boston, MA.
- Bell, D.A., 1995. Who's Afraid of Critical Race Theory. *University of Illinois Law Review*, p. 893.
- Berlant, L., Warner, M., 1995. Guest column: what does queer theory teach us about X? *Publ. Modern Lang. Assoc. Am.* 110 (3), 343–349.
- Bourdieu, P., 1993. *The Field of Cultural Production*. Polity Press, Cambridge, UK.
- Britzman, D.P., 1995. Is there a queer pedagogy? Or, stop reading straight. *Educ. Theor.* 45 (2), 151–165.
- Brockenbrough, E., 2018. *Black Men Teaching in Urban Schools: Reassessing Masculinity*. Routledge, New York, NY.
- Butler, J., 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, New York, NY.
- Butler, J., 1993. *Bodies That Matter: On the Discursive Limits of "Sex"*. Routledge, New York, NY.
- Carlström, C., 2019. BDSM, becoming and the flows of desire. *Cult. Health Sex.* 21 (4), 404–415.
- Chang, Y.K., 2005. "Through queer" eyes: critical educational ethnography in queer studies. *Rev. Educ. Pedagog. Cult. Stud.* 27 (2), 171–208.
- Cohen, C.J., 2000. Punks, bulldaggers, and welfare queens: the radical potential of queer politics. In: Richardson, L., Taylor, V., Whittier, N. (Eds.), *Feminist Frontiers*. McGraw Hill, New York, NY, pp. 495–511.
- Cohen, C., 2019. The radical potential of queer? Twenty years later. *GLQ A J. Lesbian Gay Stud.* 25 (1), 140–144.
- Corber, J., Valocchi, S., 2003. *Queer Studies: An Interdisciplinary Reader*. Blackwell Publishing, Malden, MA.
- Crenshaw, C., 1991. Mapping the margins: intersectionality, identity, and violence against women of color. *Stanford Law Rev.* 43 (6), 1241–1300.
- de Lauretis, T., 1991. Queer theory: lesbian and gay sexualities. *Differences* 3 (2), iii–xviii.
- de Lauretis, T., 1994. Habit changes. *Differences* 6 (2–3), 296–313.
- Delgado, R., Stefancic, J., 2017. *Critical race theory*. In: *Critical Race Theory*, third ed. New York University Press.
- Denzin, N., Lincoln, Y., 2008. *The Landscape of Qualitative Research*. Sage Publications, Los Angeles, CA.
- Dilley, P., 1999. Queer theory: under construction. *Int. J. Qual. Stud. Educ.* 12 (5), 457–472.
- Dixit, P., 2020. Navtej Singh Johar v Union of India: decriminalising India's sodomy law. *Int. J. Hum. Right.* 24 (8), 1011–1030.
- Doty, A., 1993. *Making Things Perfectly Queer: Interpreting Mass Culture*. University of Minnesota Press, Minneapolis, MN.
- Driskill, Q.L. (Ed.), 2011. *Queer Indigenous Studies: Critical Interventions in Theory, Politics, and Literature*. University of Arizona Press, Tucson, AZ.
- Edelman, L., 1998. The future is kid stuff: queer theory, disidentification, and the death drive. *Narrative* 6 (1), 18–30.
- Epstein, S., 2005. A queer encounter. In: Steele, T.L. (Ed.), *Sex, Self, and Society: The Social Context of Sexuality*. Thompson Wadsworth, Belmont, CA, pp. 68–73.
- Fennell, J., 2018. *Coed Naked: Insider Ethnography in the BDSM Subculture*. SAGE Publications Ltd, Thousand Oaks, CA.
- Foucault, M., 1978. *The History of Sexuality, Part I*. Translated by R. Hurley. Vintage Books, New York, NY.
- Freire, P., 1970. *Pedagogy of the Oppressed*. M. B. Ramos, Trans. Continuum, New York, NY.
- Gamson, J., 1995. Must identity movements self-destruct? A queer dilemma. *Soc. Probl.* 42 (3), 390–407.
- Geertz, C., 1973. *The Interpretations of Cultures*. Basic Books, New York, NY.
- Gershon, W.S., 2017. *Curriculum and Students in Classrooms: Everyday Urban Education in an Era of Standardization*. Lexington Books, Landham, MD.
- Gilbert, J., 2014. *Sexuality in School: The Limits of Education*. University of Minnesota Press, Minneapolis, MN.
- Halberstam, J., 1998. *Female Masculinity*. Duke University Press, Durham, NC.
- Halberstam, J., 2011. *The Queer Art of Failure*. Duke University Press, Durham, NC.
- Halperin, D.M., 1995. *Saint Foucault: Towards a Gay Hagiography*. Oxford University Press, New York, NY.
- Halperin, D.M., 2008. The normalization of queer theory. *J. Homosex.* 45 (2–4), 339–343.
- Harper, P.B., McClintock, A., Muñoz, J.E., Rosen, T., 1997. Queer transxions of race, nation, and gender: an introduction. *Soc. Text* 1 (2), 1–4.
- Heigham, J., Croker, R. (Eds.), 2009. *Qualitative Research in Applied Linguistics: A Practical Introduction*. Palgrave, New York, NY.
- Humphreys, L., 1970. *Tearoom Trade: Impersonal Sex in Public Places*. Transaction Publishers, Piscataway, NJ.
- Jagose, A., 1996. *Queer Theory: An Introduction*. New York University Press, New York, NY.
- Johnson, E.P., 2001. "Quare" studies, or (almost) everything I know about queer studies I learned from my grandmother. *Text Perform. Q.* 21 (1), 1–25.
- Johnson, E.P. (Ed.), 2016. *No Tea, No Shade: New Writings in Black Queer Studies*. Duke University Press, Durham, NC.
- Jones, S.R., Abes, E.S., 2013. *Identity Development of College Students: Advancing Frameworks for Multiple Dimensions of Identity*. Jossey-Bass, San Francisco, CA.

- Jones, S.H., Harris, A.M., 2018. *Queering Autoethnography*. Routledge.
- Kessi, S., 2018. Photovoice as a narrative tool for decolonization: black women and LGBT student experiences at UCT. *S. Afr. J. High. Educ.* 32 (3), 101–117.
- Kincheloe, J.L., McLaren, P., 2011. Rethinking critical theory and qualitative research. In: Hayes, K., Steinberg, S.R., Tobin, E. (Eds.), *Key Works in Critical Pedagogy*. Brill Press, Leiden, Netherlands, pp. 285–326.
- Kissen, R.M. (Ed.), 2002. *Getting Ready for Benjamin: Preparing Teachers for Sexual Diversity in the Classroom*. Rowman & Littlefield Publishers, Landham, MD.
- Kumashiro, K.K., 1999. Supplementing normalcy and otherness: queer Asian American men reflect on stereotypes, identity, and oppression. *Int. J. Qual. Stud. Educ.* 12 (5), 491–508.
- Lambevski, S.A., 1999. Suck my nation-masculinity, ethnicity and the politics of (homo) sex. *Sexualities* 2 (4), 397–419.
- Lather, P.A., Smithies, C.S., 1997. *Troubling the Angels: Women Living With HIV/AIDS*. Westview Press, Boulder, CO.
- Lather, P., 1986. Issues of validity in openly ideological research: between a rock and a soft place. *Interchange* 17 (4), 63–84.
- Lester, J.N., Nusbaum, E.A., 2018. "Reclaiming" disability in critical qualitative research: introduction to the special issue. *Qual. Inq.* 24 (1), 3–7.
- Levy, D.L., Johnson, C.W., 2012. What does the Q mean? Including queer voices in qualitative research. *Qual. Soc. Work* 11 (2), 130–140.
- Lorde, A., 1978/1997. *The Collected Poems of Audre Lorde*. W.W. Norton Publishers, New York, NY.
- Love, B., 2019. *We Want to Do More Than Survive: Abolitionist Teaching and the Pursuit of Educational Freedom*. Beacon Press, Boston, MA.
- Mayo, C., Rodriguez, N.M., 2019. *Queer Pedagogies: Theory, Praxis, Politics*. Springer Publishing, Switzerland.
- Mayo, C., 2006. Pushing the limits of liberalism: queerness, children, and the future. *Educ. Theor.* 56 (4), 469–486.
- McCready, L.T., 2010. *Making Space for Diverse Masculinities: Difference, Intersectionality, and Engagement in an Urban High School*. Peter Lang, New York, NY.
- McRuer, R., 2006. *Crip Theory: Cultural Signs of Queerness and Disability*. New York University Press, New York, NY.
- Mercer, K., 1994. *Welcome to the Jungle: New Positions in Black Cultural Studies*. Routledge.
- Miller, D., 1998. *Place for Us: Essay on the Broadway Musical*. Harvard University Press, Cambridge, MA.
- Morland, I., 2009. What can queer theory do for intersex? *GLQ A J. Lesbian Gay Stud.* 15 (2), 285–312.
- Muñoz, J.E., 2006. Stages: queers, punks, and the utopian performance. In: Madison, D.S., Hamera, J. (Eds.), *The Handbook of Performance Studies*. Sage, Thousand Oaks, CA, pp. 9–20.
- Nespor, J., 1997. *Tangled up in School: Politics, Space, Bodies, and Signs in the Educational Process*. Routledge, New York, NY.
- Norton, R., 2016. *Myth of the Modern Homosexual: Queer History and the Search for Cultural Unity*. Bloomsbury Publishing, London, UK.
- Pausé, C., Wykes, J., Murray, S. (Eds.), 2016. *Queering Fat Embodiment*. Routledge, New York, NY.
- Pereira, P.P.G., 2015. Queer decolonial: quando as teorias viajam. *Contemporânea-Revista de Sociologia da UFSCar* 5 (2), 411.
- Pérez, E., 1999. *The Decolonial Imaginary: Writing Chicanas Into History*. Indiana University Press.
- Pinar, W.F. (Ed.), 2012. *Queer Theory in Education*. Routledge, New York, NY.
- Plummer, K., 2008. Critical humanism and queer theory: living with the tensions. In: Denzin, N., Lincoln, Y. (Eds.), *The Landscape of Qualitative Research*. Sage Publications, Los Angeles, CA, pp. 477–500.
- Quinn, T., Meiners, E.R., 2009. *Flaunt It!: Queers Organizing for Public Education and Justice*. Peter Lang, New York, NY.
- Rand, E.J., 2014. *Reclaiming Queer: Activist and Academic Rhetorics of Resistance*. University of Alabama Press, Tuscaloosa.
- Ribbens, J., Edwards, R. (Eds.), 1997. *Feminist Dilemmas in Qualitative Research: Public Knowledge and Private Lives*. Sage, London, UK.
- Rich, A., 1980. Compulsory heterosexuality and lesbian existence. *Signs* 5 (4), 631–660.
- Rodriguez, N.M., Pinar, W.F. (Eds.), 2007. *Queering Straight Teachers: Discourse and Identity in Education*. Peter Lang, New York, NY.
- Rodriguez, N.M., Martino, W.J., Ingre, J.C., Brockenbrough, E., 2016. *Critical Concepts in Queer Studies and Education*. Palgrave Macmillan, New York, NY.
- Rogers, B.A., 2020. Researching while queer: a research note about a genderqueer lesbian conducting qualitative research in the Southeastern United States. *Int. J. Soc. Res. Methodol.* 1–7.
- Rosaldo, R., 1993. *Culture & Truth: The Remaking of Social Analysis*. Beacon Press, Boston, MA.
- Rosenberg, S., 2018. Coming in: queer narratives of sexual self-discovery. *J. Homosex.* 65 (13), 1788–1816.
- Sears, J.T. (Ed.), 2005. *Gay, Lesbian, and Transgender Issues in Education: Programs, Policies, and Practices*. Routledge, New York, NY.
- Sedgwick, E.K., 1985. *Between Men: English Literature and Male Homosocial Desire*. Columbia University Press, New York, NY.
- Sedgwick, E.K., 1990. *Epistemology of the Closet*. University of California Press, Berkeley, CA.
- Sedgwick, E.K., 2003. *Touching Feeling: Affect, Pedagogy, Performativity*. Duke University Press, Durham, NC.
- Singh, J., 2017. *Unthinking Mastery: Dehumanism and Decolonial Entanglements*. Duke University Press, Durham, NC.
- Smith, L.T., 2013. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books Ltd, London, UK.
- Solórzano, D.G., 1998. Critical race theory, race and gender microaggressions, and the experience of Chicana and Chicano scholars. *Int. J. Qual. Stud. Educ.* 11 (1), 121–136.
- Spargo, T., 1999. *Foucault and Queer Theory*. Icon books, Cambridge.
- Spencer, R., Pryce, J.M., Walsh, J., 2014. Philosophical approaches to qualitative research. In: Leavy, P. (Ed.), *The Oxford Handbook of Qualitative Research*. Oxford University Press, New York, NY, pp. 81–98.
- Spindler, G., Spindler, L., 1982. Roger Harker and Schonhausen: from the familiar to the strange and back again. In: Spindler, G. (Ed.), *Doing the Ethnography of Schooling*. Holt, Rinehart and Winston, New York, NY.
- Steinberg, S.R., Cannella, G.S., 2012. *Critical Qualitative Research Reader*. Critical Qualitative Research. Peter Lang, New York, NY.
- Stewart, K., 2010. Worlding refrains. In: Gregg, M., Seigworth, G.J. (Eds.), *The Affect Theory Reader*. Duke University Press, Durham, NC, pp. 339–353.
- Sykes, H.J., 2011. *Queer Bodies: Sexualities, Genders, & Fatness in Physical Education*. Peter Lang, New York, NY.
- Talbot, S., Steinberg, S.R., 2000. *Thinking Queer: Sexuality, Culture, and Education*. Peter Lang, New York, NY.
- Tatum, B.D., 2003. "Why are All the Black Kids Sitting Together in the Cafeteria?": And Other Conversations About Race. Basic Books, New York, NY.
- Tierney, W.G., 1997. *Academic Outlaws: Queer Theory and Cultural Studies in the Academy*. Sage Publishing, Thousand Oaks, CA.
- Varenne, H., McDermott, R., 1998. *Successful Failure: The School America Builds*. Westview, Boulder, CO.
- Weinberg, T.S., 1995. *S & M: Studies in Dominance & Submission*. Prometheus Books, Amherst, NY.
- Weiss, M., 2016. Always after: desiring queerness, desiring anthropology. *Cult. Anthropol.* 31 (4), 627–638.
- Wells, C., MacLeod, A., Frank, B., 2012. Queer theory, ethnography and education. In: Delamont, S. (Ed.), *Handbook of Qualitative Research in Education*. Edward Elgar, Cheltenham, UK, pp. 108–125.
- Whitlock, R.U. (Ed.), 2013. *Queer South Rising: Voices of a Contested Place*. Information Age Press, Charlotte, NC.
- Wolcott, H.F., 1994. *Transforming Qualitative Data*. Sage Publications, Thousand Oaks, CA.
- Wozolek, B., 2018. Gaslighting queerness: schooling as a place of violent assemblages. *J. LGBT Youth* 15 (4), 319–338.
- Wozolek, B., 2019. M/othering redux: a reading of curriculum theory to get through sleepless nights. *J. Gend. Stud.* 28 (8), 873–882.
- Youde, D., 2009. Queer outings: uncomfortable stories about subjects of post-structural school ethnography. *Int. J. Qual. Stud. Educ.* 23 (1), 87–100.

Case study research

Robert Stake^a and Merel Visse^b, ^a University of Illinois at Urbana-Champaign, Urbana, IL, United States; and ^b Drew University, Madison, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

The disciplined study of cases in education	85
Case study design	86
Size	86
Choosing cases	86
Types	86
Narrative study	87
Issues	87
Activities	87
Triangulation	87
Humanistic study	87
An example of case study: reader focused writing program	88
Indianapolis Regional Office, Robert Stake (Stake et al., 1997)	88
The station	88
RFW training	89
Obstacles to implementation	89
The image of service	90
Conclusion	90
References	91

What is that, silent as a fox, sneaking through the research woods? Can't really tell. Could be fox. So hard to define, yet fundamental to the methods of inquiry. Seen differently in different contexts, for different purposes, and yet the favorite of doctoral students, at least in certain fields. Not really foxy, but an approach to inquiry with promise of being vivid, protective, and valid, capturing the movement and complexity of the case and its habitat.

The disciplined study of cases in education

Education is a phenomenon resounding wonder. The wonderful doesn't hide the dread and boredom and discrimination of classroom or lifelong learning, but the puzzlement of the unknown (and unknowable) is ever a challenge. It is the mystery of metaphors and minus numbers and microbiologics that ever keeps our minds off balance. Education is an everybody's venue of inquiry, charming and desolate. With method or not, we grasp at meanings, seeking to find a fit for the oncoming strangeness. The strangeness is often a something, and we need to know that something better. It can be a case, a single case—stirring our wonder—and before test time or deadline, we need to know it better, a case to be studied.

Belonging to a discipline or not, the case study is usually the study of one thing. For the uniqueness or generality, we seek to understand more about that thing. Sometimes that thing is compared. Other times not, but we do not primarily study a case to understand other cases. The case is a bounded system, a special something to be studied, an entity such as a child, a teacher, a classroom or a program. A phenomenon or a relationship may also be studied indirectly, but the case is the entity to be studied directly. Take for example, a manager, as the case. As with other cases, the manager has activities, problems, decisions, relationships, all of which may need to be looked into for a study of the case. The particular case has, in some ways, a unique life, although many parts of it may be ordinary. We study the ordinary but especially the unusual. The case is something we do not sufficiently understand, and studying it means we give it special care.

Case study is a method of inquiry, or a collection of methods, used to dispel and to expand the wonder of understanding the individual case. Sometimes because we have questions, making the case study instrumental. Other times because the case presents itself to us with an intrinsic value, we call that an intrinsic case study. Either instrumental or intrinsic, a case study in different places goes by different names (Feagin et al., 1991). Sometimes it's a diary, sometimes a transcript, a course file, sometimes an obituary. Biographers call a case study a biography, cartographers a map; all of us call a history a history. It's a history of something, a history of a case, brief or unending. There should be no pressure on us all to call it the same thing. Or to define it the same. In a distinguished career, sociologist Howard Becker (1992) regularly studied cases, but preferred to call his case study, "fieldwork." Lots of tabs, still, the discipline of case study, as a scholarly research method, calls for us to recognize some important descriptors to be selected and certain happenings to be recorded, so as to understand the case better.

Much scholarly study is aimed at theorizing, generalizing, to understand causality and interdependence. That is not the strength of case study, although it happens (Yin, 2013; Ragin and Becker, 1992). Surveys and experiments are mainly for the purpose of comparison and explanation, sometimes found useful in case studies (as in the example of the Veterans' Regional Center, later in these pages), but understanding an individual case, even a case inclusive of many places and times and people, concentrates

more on the *activity* and *context* of that one case. A case cannot be well understood without some comparison and sense of agency, but the usual purpose is to come to better comprehend that designated case, that bounded case.

Case study is sometimes thought to be more on the side of personalizing an inquiry, and a tally might find it so. But personalization is not a definitive characteristic. Personalizing does not necessarily make a study a case study. It could, for example, be an autoethnography or narrative without being a case study. For a case study to be a case study, there needs to be a case with clear boundaries and issues, as well as an interaction with a context.

An issue is a matter of contention which is of special concern or importance. In case study, an issue has to do with the functioning of the case, particularly reflecting one or more of its purposes, an exercise in management by those in control. Certain choices or actions are taken, with uncertainty as to the implication or outcome, with different effects on different constituents likely. Here is a visualization that may help.

Personalization by extending intimacy and compassion into the study, may open important facets of understanding, but may threaten the intersubjectivity of the study. In case studies, intersubjectivity is a matter for researchers to negotiate with their audiences, being informative as to proximity of observation and depth of questioning. Case study can be a ground for compassion but the method tolerates indignation and intrusion if they better get the story told (Merriam, 1998).

Observation is a high priority method for the case researcher, often more important than interviewing or document examination. The quality of observation depends on discipline and repetition. One needs to think through, in advance, some of which may come to pass. One needs to be part of the context of the case more than once. Classification of issues and unfolding insights about those issues is needed to tease out subtle tendencies. Without deep involvement, not meaning participation but meaning pondering, one cannot notice nuances and tease out meanings. Much of the budget may be consumed bring in watchful eyes and note-taking laptops, resulting in little remaining for persons and sites and occasions to be recorded. The generality of aspects within the case (everywhere true? always true?) require multiple kinds of evidence, often relying on the observations and recollections of others than researchers. Existing documents, photos, recordings can reward a diligent search. This is not a case of law: evidence is defined not only by the observer, but by the reader, given thoughtful help and cautionary advice by the author. So case study uses a kaleidoscope of approaches, from narrative to ethnographic, and a kaleidoscope of methods, from observation to interviewing. Any style of inquiry may be used—it should fit well the audience and purpose of the study (Stake, 2005).

That may be giving too much leeway to the researcher. Just as there are questions of validity in studies relying heavily on measurement and statistical analysis, validity is an issue with case studies as well. Triangulation is the term used by many researchers to draw evidence from different vantage points to show accuracy of data and strength and weakness of the finding. The data in all studies can be interpreted in multiple, even contradictory, ways (Creswell and Poth, 2018). And the researcher has an obligation to search how the representation of the case can be differently interpreted and misinterpreted, and to say so in the report. In the case study example (several pages ahead), the reader might conclude that central office (federal) administrators were insensitive to regional office communications with veterans, and perhaps the researchers insufficiently looked into evidence of that perception.

The Holy Bible and Koran are case studies, possibly long ago at risk of being turned down by a doctoral committee and the journal editor. Truth and faith need to be part of the review, and so with other accounts of experience. The readers have their own affairs to keep in order, advocacies, reputations and page limits, but stories need to be told, and meanings questioned. Mistakes will be made and beliefs pondered. Case studies, spirited and perfunctory, are in little danger of being superseded. But in an age of free speech and political misinformation, a discipline of inquiry works against amusement and persuasion (Kemmis, 1980).

Case study design

Size

A case study project or its report can be large or small. The study can be personable or austere, chatty or scholarly, activist or dispassionate. In research circles, it is likely to be more respected when dealing with scholarly relationships and circumstances (Geertz, 1973). Political content such as street protests can increase the report's circulation. (MacDonald et al., 1982). But an hour of pondering the case will sometimes be the right measure.

Choosing cases

For the teacher, the program evaluator and the auditor, their cases are usually known or assigned in advance. Their cases often have been of interest before any formal study begins. For educational researchers, case studies usually require cases to be chosen. At the outset of a case study, sometimes a phenomenon such as discrimination or on-line testing is identified—the cases are opportunities to study it. Understanding critical phenomena and emerging problems may depend on choosing the cases well. Even when the cases are decided in advance, the researchers may need to select persons, places, events, and other “particulates” to observe.

Types

Many have categorized case studies (Hamilton, 1980; Schwandt, 2000; Sjoberg et al., 1991; Stenhouse, 1979). Stake (2005) identified three types. An *intrinsic case study* is undertaken because the researcher wants better understanding of a particular case. Here the case does not represent other cases, nor does it need to lead toward theory. It does not illustrate a particular feature or problem. An

intrinsic case study is often chosen because of its particularity *and* ordinariness. The case itself is of interest. It is cared about. It sticks out as the one to study. The second type is an *instrumental case study*. This type provides insight into an issue with an aim to generalize, to transfer insights to other contexts. The case is of secondary interest, it plays a supportive role. It facilitates our understanding of something else, again maybe discrimination or on-line testing. And third, collective case study, also called multiple case study, encircles a number of cases, separately and together, in order to *investigate* a hypothesis, gathering, phenomenon, or condition. Here, researchers strive for a common understanding, or they want the case study to contribute to the development of theory (Gerring, 2017).

Narrative study

Case studies and narrative go hand in hand. Case studies are represented through thick descriptions using narrative vignettes to convey experiential data, sometimes visuals, but also statistical data, if relevant. Narratives situate the study and its experiential data. Significant events, times and places and any memorable contexts serve to make many case reports easier to understand. A strong story is the vehicle in many biographies, histories, and developmental records. Interviews often serve well in data gathering. Dialogic presentation, conversational format, is popular with many case researchers (Smith, 1994; Clandinin and Connelly, 1999; Silverman, 2000).

Issues

A research case study does not seek to understand a phenomenon in order to control it. Seldom to predict or find the basic truth in it, such as an essence. When we study a case, we are sensitive to learning about its discrete units and their relationships. They show us the pathway to understanding issues that are at stake. A case study is not so relevant when we want to understand broad information questions such as “How were these programs implemented?” and “Who cares for most of the elderly?” Case study leans toward other questions, such as “How have they tailored this particular care to women who home-school their children?” or “What happened to the elderly when their residence home (the case) participated in a community-outreach program?” Case studies are organized around *issues*. Issues are complex, situated, problematic relationships. They are themes, arguments, trade-offs. We cannot know all the issues before we enter the context of the case. We gradually learn to formulate better questions, driven by issues. Issues invite attention to ordinary experience but also to the language and understanding of the common disciplines of knowledge, such as anthropology, economics, political theory, and literary criticism.

Many qualitative case researchers orient to complexities connecting ordinary practice to the abstractions and concerns of academic theories. This dialectic connection is present in the single case study, but does not replace it as focus. There is a fine line and tension between practice and theory, needing the researchers’ attention. The themes and issues used to organize the study may or may not be the ones used to report the case to others. Observing and analyzing work are different work from presenting the case report. At the end, it may be the anticipated issues of readers that will structure the report (Becker, 1992).

Activities

Usually, researchers need to find out firsthand what the case does (did)—its activity, its functioning. Sometimes the most direct answers come from observing the actions of people, the operation, if in the context of a program, or asking those who saw it (Erickson, 1986). This includes people as cases. Almost always, the case’s activities are expected to be influenced by contexts; understandings need contexts. Therefore, contexts need to be studied and described, whether or not evidence of influence is to be found. People simultaneously experience many things. The researcher lets mind and eye scan a large number of happenings, variables, and contexts, and makes tough decisions as to what to dig into.

Triangulation

Case researchers are a heterogeneous lot, but most are disposed to tell the truth, even though that is not the purpose of case study. They selectively choose which (narrative) truth to tell, but they choose not to deceive and seek redundant confirmation of their main interpretations. By observations that approach from three or more directions or data sources to support their claims, they offer confirmations. They call it, “triangulation.” They mean pretty much what measurements specialists mean by “validity” (Campbell, 1975; Maxwell, 1992).

Humanistic study

Many intrinsic case studies have a humanistic leaning, but it need not be. Particularity or ordinariness does not immediately mean we pay attention to human experience, relationships or issues of empathy and compassion. A humanistic lens would highlight human welfare and progress, dignity, dialog, openness. Even when a case study is not seen to be “humanistic”, we can still look at what happens when people are in play, what human activity is going on, and what meanings unfold in the process.

An example of case study: reader focused writing program

For these pages on case study, we have chosen to include a half-shortened case report of a Reader Focused Writing program as an example (Stake et al., 1997). As an example of an intrinsic case study used in an administrative situation, it is the report of a formal, contracted evaluation of a training program run by the US Veterans Benefits Administration, a federal agency. Not much can be learned about that particular training program from just one short report. But something can be learned about the method from a single case study. The purpose here is to illustrate choices, hazards, and opportunities in doing a case study for institutional purposes. Much case study research will be with cases of persons, but any program or initiative can be a case to be studied, and the case study below should help us understand organizational complexities.

The case (here an evaluand) was a special all-states program to improve letter writing to veterans regarding housing, education, and other earned benefits. The evaluation was conducted for the Veterans Benefits Administration by a University of Illinois CIRCE team during the fall of 1997. A four-member team of evaluation specialists was led by Stake et al. (1997). In addition to Indianapolis, evaluative case studies were conducted at St. Petersburg, Boise, Baltimore and Denver. Five of the fifty regional offices were chosen for data gathering, and each was treated as a case, so the collection could be called a multiple case study (Stake, 2006). The report was a report on the quality of the training nationally. In addition to these five case studies—making the report a multiple case study report—national data gathering and analysis included veteran interviews, a staff survey, pre- and post-training letter comparisons, a regional director survey and administrator interviews. The several tasks are shown on the next page. The responsibilities of team members are indicated by the sectors and the case study sites.

The Reader Focused Writing program was ordered by Congress because VA letters to vets are periodically complained about for being inconsiderate, incomprehensible and legalistic. Congress ordered the VA to fix it. In this case, 656 participants took a seven-day satellite training to improve their letter writing to veterans. From the case we learn about the tension between adhering to protocols, “indicators of production”, leading to not so good letters, and what needed to change to produce a better letter. Trainees were torn between doing what was required, or doing what was needed, at the expense of productivity—a struggle that we all know so well and want to escape from. Here, we already notice several issues, such as the tension between needs for communication and staff time available for communicating.

Indianapolis Regional Office, Robert Stake (Stake et al., 1997)

Upstairs on the third floor of the Federal Building, I inquired about services to Indiana veterans. The workers there were aware that from time to time the federal Veterans Benefit Administration was accused of writing incomprehensible letters. They knew from experience that over the years, many letters had been bad, filled with legal jargon and bureaucratic syntax. Furthermore, they knew that they were part of a political context. But when they discussed their work, they spoke consistently without reference to the media, Congressional oversight, or what the public thought of them. There was a world of veterans, many of whom they had talked to on the phone, a world that included loan agencies, hospitals, and colleges. They referred often to the Central Office in Washington, but in organizational terms, not political.

The station

The VBA Regional Office in Indianapolis provided services. For disability compensation, pensions, vocational rehabilitation, and backing for home loans. Educational and insurance benefits were handled at St. Louis and Cleveland, respectively. As the division names imply, these services are quite different, and specialized within, so that the responsibilities of different workers ranged greatly from room to room across the station. Staff size was about 145, down from a high of 200 three years earlier. Loan Guaranty currently had 30 positions; Vocational Rehabilitation and Counseling, 11. What had been the largest group had become larger, with 81 positions, as Compensation and Pension merged with Veterans Services. C & P was where adjudicators and rating specialists acted upon claims for support due to physical disability resulting from military service. The reorganization brought with it the expectation that more personal communication practices, via telephone and face-to-face counseling, would reduce the heavy load of adjudication letter-writing (although paper trails would not disappear). The local work in Loan Guaranty was changing as Indianapolis became the national center for a portfolio loan program. Down-sizing (called “right-sizing” by some) meant not only a diminution of time available to handle a claim but an increase in claims handled at a national rather than regional center.

The director of this Regional Office was Dennis Wyant, heading the office since mid-1994 after many years in the Central Office where he had been national director of Vocational Rehabilitation and Counseling. Wyant held a doctorate in Education and was a certified rehabilitation counselor. Despite being legally blind, Wyant kept an eye personally on what was happening around the station. He was a “people person,” a “problem-solver,” neither confrontive nor authoritarian. He heard well what his chiefs and assistants had to say about running the programs. He expressed strong support of Reader Focused Writing as a step in the improvement of services to veterans.

Claims for benefits are made by veterans filling out forms. Most veterans have made contact with one of the service organizations, the American Legion and Veterans of Foreign Wars, for example, and many first encounter VBA with a form already filled in. Some requests are taken by telephone or over the counter, but quickly forms become the currency. Some forms are supplemented with records from doctors or campus admissions offices or county offices and such. The evidence is reviewed and a reply prepared.

For a disability compensation claim, an adjudicator will recognize which on the menu of computer letters should be called up, and by using Microsoft Word and PCGL (Personal Computer Generated Letters) and by filling in blanks, modifying language, and attaching forms and pamphlets, compose a reply to the claimant.

A quality assurance review annually examines a sample of cases from the compensation and pension files, 106 were selected for 1996, of which 90 were reviewed by a federal team. Attention is given to the appropriateness of findings and not to the appropriateness of the communication. At Indianapolis, for a while, different divisions were conducting their own quality assurance operations but reports are no longer required and the practice waned.

RFW training

On September 24, 1996, Wyant wrote to his division chiefs informing them of the Reader Focused Writing schedule as follows:

We are nearing our scheduled time for Indianapolis to receive RFW training. Each RFW Tools Course consists of seven sessions; approximately 3 hours each of core instruction and a 2–3 hours follow-up approximately 6–8 weeks following the class. The 3 hours per day includes 2 hours of broadcast instruction plus 1 h of class. Exercises and practice. Since we have eight keypads, we can train eight during the morning and eight during the afternoon.

Such lengthy all-staff training had not happened previously at the station. Opposition to the plan among the chiefs and among the workers was considerable, voiced as putting too much strain on the workload, but also with skepticism that this professional development workshop, “as always before,” one trainee said, would be anything more than consciousness-raising. It also was recognized that, for most of the participants, letter-writing to veterans was not a large part of their job description. But opposition melted with realization of the generality of the communication principles and the appeal of the composition techniques, warmed also by opportunity to work on interesting tasks with staffers from other divisions and by the high quality of the broadcasts. In retrospect, essentially all participants considered the training relevant, practice-changing, and personally upgrading.

For 2 hours a day in a corner of the station’s large conference room, the eight participants for each session watched—via satellite screen—Melodee Mercer, a loan guarantee specialist in Philadelphia and member of the VBA communications task force coaching the RFW and interacting on touch pads. “On page 48 of your manual, in the letter from Timothy Morrell, what implied questions can you identify? Your light is on, Frank at Detroit, what questions are implied in his letter?” The individual touch pad connections failed often, but ordinary phones then were substituted. The third hour was used for group activities. The exercises were well developed, thought-provoking, especially in that they gave the specialists opportunity to see how co-workers in other divisions resolved content problems. No further broadcasts were scheduled and even the six-weeks-later sessions disappeared.

Obstacles to implementation

One of the primary findings from our fieldwork is that there was not a close match between the training and the actual work at Regional Offices. It has already been mentioned that the work is diverse and specialized, yet the training was the same for all participants.

Regional Office staff members who participated in VBA training in October 1996, were unanimous in considering quality of letter-writing to veterans an important issue and that writers should strive to orient more to the veteran’s needs. All of the 29 but two said that this VBA approach made sense to them and all but three said that VBA was *not* making too much of a fuss about improving letter writing.

In the interviews as well as in the surveys, it was apparent that the staffers supported a continuing effort to improve letter-writing. An interviewee said.

It has been a long time since we have taken a hard look at letter writing and communication skills in general. Some letters we send out are confusing, legalistic. The Agency as a whole needs to provide the legal angle, and that still is the dominant attitude across the Agency. We are not allowed to advocate for the veteran or to gather information to advance his claim. RFW training is the first time the Agency clearly provided a pro-veteran view. It was long overdue.

In interviews with six of these respondents, it was also apparent that they felt that the quality of letter writing had improved considerably in the last several years. When asked why, two or three said, “The main reason is that the standard computer letters that we draw from are much better than they were”. Twenty-four indicated they had applied RFW ideas to their own work, four of those said “often.”

We asked these 29 trainees what portion of their work time they presently spent writing letters to veterans. Nine of them indicated they did not do that at all and the median for the group was doing it only 5% of the time. Only three of them indicated that they spent half their time or more writing letters to veterans.

In two separate ways, we asked them the average amount of time in the last year spent on developing RFW skills. The median was 1–2 hours per month. Almost half of them had participated in earlier VBA efforts to improve letter writing. About a third of them had once had a formal writing-improvement course. When asked about the need for inservice training of all kinds, the median of the group responses was that 4 h a month should be set aside.

As to the training activity, 26 of the 29 checked that it was well organized. The same number said they learned much about better letter-writing. Just over half said the videos and materials were of high quality, five said low. About half of them said that the quality of the contribution of the on-site instructor was low.

As to implementation, 20 indicated that it was impractical to implement RFW in the ways illustrated in the course, apparently meaning that there were just very few occasions for start-from-scratch letter writing, and too little time to do it. When asked if the organization of the Regional Office had changed in any way to support RFW, 17 said “no” and another 11 said they didn’t know. In follow-up, they assured us that the officers had clearly indicated their support for RFW letter writing but, as they saw it, the organization had not changed to make it happen, and most did not see ways that it should change.

When asked if simpler letters trying to deal with a complex situation sometimes steer the veteran wrong, chiefs and workers alike said they could not identify cases of that happening. One added, “The complex information is in the attachments and we always tell them how to ask for more information.”

One thing apparent in discussions at Indianapolis was a perceived mismatch between the training as designed and work methods and workload of the staff. The training apparently did not sufficiently acknowledge the diversity of communication responsibilities within VBA or sufficiently point to the generality of the RFW approach for communication in general. Two suggestions arising again and again were that in their own communications by letter, the Central Office should exemplify Reader Focused Writing and that the computer letters and the PCGL programs should be further upgraded along RFW lines.

The image of service

Even under careful probing, the staff of the Indianapolis Regional Office revealed a strong work ethic and a sincerity in being of service to Veterans. But they also were careful to maintain a good public image. They were highly aware of the low regard many citizens have for federal employees, and they themselves decried bureaucracy and political motive, for, in their own eyes, they were “Hoosiers,” not bureaucrats (Hoosiers are loyal supporters of Indiana).

Still they could not overlook the fact that in Oklahoma City, a federal services building of vintage and set-aside like theirs, had been blown sky-high. When I asked why the empty slot on the directory where earlier the Director’s Office had been identified (Room 397) and when I asked why a punch-code lock on the Door to the Director’s suite, I was told they were changed when an angry veteran had threatened Rowland Christian, the Assistant Director. In an age when citizens daily watch television actors acting out their rage, these public servants know that violence can quickly come to them.

But they also know all but a few of the thousands they serve are reasonable men and women seeking only their entitlements. They are aware that the laws awarding those benefits are complex but have a confidence that their legal and technical people have appropriately categorized the conditions that establish and differentiate eligibilities. They take pride in their efforts to communicate and feel that the changes in rules and word processing in the last few years have made almost all of their letters comprehensible. They would like better form letters and even more training and work time to purge their letters of typographical errors, inconsistencies and obfuscations (For the Encyclopedia, the report was cut short at this point).

Here in this report, the case is the Reader Focused Writing Program, a bounded training system. Our study highlighted issues such as (1) the mis-match between the training and the actual work at the Regional Office and (2) the difficulties of following limits to benefits earned by veterans still living the effects of war and (3) the realization that some claims will be fraudulent claims. Pressing hard were the cold facts of eligibility, as determined by Congress, visibly displayed against the reality of difficulty, for many, of getting a loan, a hospital bed, or further education. The staff coped with issues of race and politics mixed with issues of poverty and imperfect records. The needed evidence was not easily found in physical examinations, documents and interviews, much needing repeated observation and scrutiny and further inquiry. The ability of case study to capture these human conditions came to mind again as VBA staff members and others read this report of the Indianapolis Service Center and VA services at large.

What more is illustrated there in the example? Looking at just the half report shown here, we see description of the diverse work of the staff and the activity of the training program. We recognize that several research methods were used: observation of staff at work and during training; interview of director, training officer, and workers; document review, opinion surveys, comparison of real letters written before and after training; and evaluation of training exercises. We triangulated the views of students, supervisors, and administrators as to barriers and facilitations to better letter writing. We sought a historical perspective through past reports. The four field workers met to discuss their experiences at the service centers.

The five cases were selected to provide differentiation in seniority, mix of veterans served, and ease of access. They were first treated as single case studies, but merged for treatment as a collective case study. The training activity provided a story scenario with problems to be solved and a narrative to report. In the weeks following the training it became apparent that the rate of portfolios to be processed in a day was not going to be reduced and that reader-focused writing on average would lower the productivity. The central administration may have known that all along, but they supported the training to see what would happen. The hope remained that staffing requests would be successful. They were awarded recognition in the Executive Department for a vigorous effort to do better.

Conclusion

In this chapter, we highlighted case study as an approach to inquiry, sometimes consisting of a single method, such as observation, other times a collection of methods, used to deepen our understanding of the individual case. Case study is useful in spaces of education and other social services such as health care. Other contexts, like for-profit organizations, benefit from case studies too. Everywhere where there is wonder about one or multiple cases, case study helps us understand the scene and the stories of those

involved. It acknowledges their experiences and viewpoints, it cares and empathizes, making it a humanistic endeavor. In all situations, the researcher tries to preserve the multiple realities, the differences and even contradictory views of what is happening. Of course, there are many other good approaches and ours is just one of many, stressing the particularistic, naturalistic fieldwork, keeping a sharp eye, and staying open to new impressions.

What is foxy about research is not its stealth or its sheen. It is its contribution to its field. Educational research has helped little to make education better, but when vivid, protective and valid, it makes education better understood.

References

- Becker, H.S., 1992. Cases, causes, conjunctures, stories, and imagery. In: Ragin, C.C., Becker, H.S. (Eds.), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge University Press, Cambridge, pp. 205–216.
- Campbell, D.T., 1975. Degrees of freedom and case study. *Comp. Polit. Stud.* 8, 178–193.
- Clandinin, J., Connelly, M., 1999. *Narrative Inquiry*. Jossey Bass, San Francisco.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. Sage, Thousand Oaks, CA.
- Erickson, F., 1986. Qualitative methods in research on teaching. In: Wittrock, M.C. (Ed.), *Handbook of Research on Teaching*, third ed. Macmillan, New York, pp. 119–161.
- Feagin, J.R., Orum, A.M., Sjöberg, G., 1991. *A Case for the Case Study*. University of North Carolina Press, Chapel Hill.
- Geertz, C., 1973. Thick description: toward an interpretive theory of culture. In: Geertz, C. (Ed.), *The Interpretation of Cultures*. Basic Books, New York, pp. 3–30.
- Gerring, J., 2017. *Case Study Research: Principles and Practices*. Cambridge University Press.
- Hamilton, D., 1980. Some contrasting assumptions about case study research and survey analysis. In: Simons, H. (Ed.), *Towards a Science of the Singular*. University of East Anglia, Centre for Applied Research in Education, Norwich, pp. 76–92.
- Kemmis, S., 1980. The imagination of the case and the invention of the study. In: Simons, H. (Ed.), *Towards a Science of the Singular*. University of East Anglia, Centre for Applied Research in Education, Norwich, pp. 93–142.
- MacDonald, B., Adelman, C., Kushner, S., Walker, R., 1982. *Bread and Dreams: A Case Study in Bilingual Schooling in the USA*. University of East Anglia, Centre for Applied Research in Education, Norwich.
- Maxwell, J.A., 1992. Understanding and validity in qualitative research. *Harv. Educ. Rev.* 63, 279–300.
- Merriam, S.B., 1998. *Qualitative Research and Case Study Applications in Education*. Wiley.
- Ragin, C.C., Becker, H.S., 1992. *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge University Press, Cambridge.
- Schwandt, T.A., 2000. Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructionism. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, Newbury Park, CA.
- Silverman, D., 2000. Analyzing talk and text. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, Newbury Park, CA.
- Sjöberg, G., Williams, N., Vaughan, T.R., Sjöberg, A., 1991. The case approach in social research: basic methodological issues. In: Feagin, J.R., Orum, A.M., Sjöberg, G. (Eds.), *A Case for the Case Study*. University of North Carolina Press, Chapel Hill, pp. 27–79.
- Smith, L.M., 1994. Biographical method. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 286–305.
- Stake, R.E., 2005. Qualitative case studies. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Sage Handbook of Qualitative Research*, third ed. Sage, Thousand Oaks, CA, pp. 443–466.
- Stake, R.E., 2006. *Multiple Case Study Analysis*. Guilford, NY.
- Stake, R.E., Davis, R., Guynn, S., 1997. Evaluation of Reader Focused Writing. <http://www.ed.uiuc.edu/circe/RFW/index.html>.
- Stenhouse, L., 1979. *The Study of Samples and the Study of Cases*. Presidential Address to the Annual Conference of the British Educational Research Association.
- Yin, R.K., 2013. *Case Study Research and Applications: Design and Methods*, sixth ed. Sage, Newbury Park, CA.

Ethnography: foundations, challenges, and spaces of possibilities

Audra Skukauskaitė, University of Central Florida, Orlando, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	92
Historical foundations and challenges in understanding of ethnography and its variations	92
Ethnography's roots and contemporary quandaries	92
Variety of ethnographies	93
What counts as ethnography?	93
Attempts at defining ethnography	94
Ethnographic logic of inquiry	95
In or out of an ethnographic space: ethnography, ethnographic studies, and use of methods derived from ethnography	96
Primary goals for ethnography: ethnography of, in, and for education (or other disciplines)	97
Power and potentials of ethnography	98
References	99

Introduction

Ethnography is a systematic approach to studying everyday life a social group co-constructs through their actions and interactions over time. Ethnographers seek to answer the question “What is happening here?” from insider-driven, culturally relevant points of view. Ethnography is not a mere method for data collection or fieldwork. It is a way of seeing (Frank, 1999; Wolcott, 2008), being (Heath and Street, 2008; McCarty, 2014), languaging (Beach and Bloome, 2019), theorizing (Bloome et al., 2005; Kelly and Green, 2019a), living (Skukauskaitė, 2021) and knowing (Agar, 2006b). Ethnography is an epistemology (Green et al., 2012) and philosophy of research (Anderson-Levitt, 2006) through which ethnographers come to know and represent the complex social worlds they study. Driven by a particular kind of iterative, recursive, and abductive logic of inquiry with an emphasis on understanding different points of view (Agar, 2006b), ethnography illuminates the complex dynamic systems (Agar, 2004; Anderson-Levitt, 2006) constituting everyday life in the multilayered contexts of people’s actions and interactions with each other and with the historical, material, ecological, and non-human actors and resources.

Starting from a position of ethnography as epistemology, a way of knowing, in this chapter I briefly overview key historical foundations and challenges created as ethnography expanded and scholars developed a variety of labels and options for ethnographic work. I then explore the ethnographic space of possibilities driven by an underlying logic, followed by an explication of two kinds of ethnographic research that fit and one kind of research that may refer to ethnography but does not fit in the ethnographic space. Subsequently, I outline ethnography’s primary goals and end the chapter with considerations about ethnography’s potentials within and beyond the academic spaces.

Historical foundations and challenges in understanding of ethnography and its variations

Ethnography’s roots and contemporary quandaries

Ethnography has a long history in the social sciences and humanities and has faced numerous challenges in over a century of its use; nevertheless, it persists and is widely used in education and across disciplines as a systematic and empirical inquiry for exploring the processes and practices people culture-into-being (Skukauskaitė and Green, 2023, based on the view of culture as a verb by Street, 1993) as they act and interact with each other moment-by-moment and over time. The roots of ethnography are usually traced to anthropology and the methods of fieldwork cultural anthropologists in the late 19th-early 20th century utilized to study groups of people (“cultures”) unfamiliar to the “western” audiences. As sociologists in the early 20th century employed ethnography to study groups closer to home in modern societies, they adopted and adapted many of the anthropological fieldwork techniques and brought in social and critical theories to the study of everyday life. (For excellent exposés of ethnography’s histories, as presented from different disciplinary and theoretical frames, see, for example Atkinson et al., 2007; Beach et al., 2018b; Hackett and Hayre, 2020; McCarty and Liu, 2017, among others).

The tendency toward colonialist, Euro- and US-centered agendas of early anthropologists has been widely criticized (Beach, 2017; Bejarano et al., 2019; Hammersley, 2006; Marcus, 2008) and has led to “crises of representation” and the reflexive turn in anthropology, sociology, education and other disciplines where ethnography is widely used (Green and Bridges, 2018; Marcus, 2021). The critiques of the euro-, western-, anglo-, white-, male- and elite-centric research agendas and narrow, often faulty constructions of “culture” as a stable and geographically bounded unit (Agar, 2006a, 2019; Eisenhart, 2018), opened doors for varied perspectives and theories which have expanded ethnographic epistemology and practice (see, for example, volumes by Anderson-Levitt, 2011; Beach et al., 2018b). Contemporary ethnographies acknowledge the early roots in anthropology and

sociology (and precursors in the work of ancient historians such as Tacitus) but also question the concepts such as “culture” and “participant observation” or “insider–outsider”. Even “culture”, a concept key to ethnographic scholarship, “is a mess” (Agar, 2006a) and has no agreed-upon definition. This situation opens doors to creative and locally responsive re-conceptualizations of culture, or culture-in-the-making, as a process, a verb, a system of beliefs, patterned practices, language, and/or a combination of a variety of aspects of life of a group of people engaged in some common activity (Agar, 2019). The lack of clear definitions enhances researcher responsibility for transparency of their logics-of-inquiry and reporting, while also creating spaces for learning *from and with* the people of how the people conceptualize their own cultures, both locally situated and globally dispersed.

Variety of ethnographies

Currently, ethnography is not tied to a particular disciplinary home or national or organizational origin. It is an epistemology for inquiry into human social worlds and has been used in many countries, representing diverse languages and peoples around the globe (e.g., Anderson-Levitt and Rockwell, 2017; Milstein and Clemente, 2018; Modiba and Stewart, 2018). Ethnography is widely utilized across varied disciplines, including education, law, medicine, hospitality, marketing and many others. A variety of ethnographies are now available, and a general search can result in a plethora of adjectives for ethnography. In Table 1, I listed some of the more common labels for ethnography and classified them based on their focus.

Table 1 illuminates one potential way of classifying varied ethnographies based on their overarching focus, such as the focus on site, discipline, or phenomenon, among other possibilities. Many other kinds of ethnographies and classifications exist, and I suggest readers explore other possibilities or create their own organizational tools. In later sections of this chapter, I introduce additional classifications based on fit within an ethnographic space and disciplinary purposes in education.

What counts as ethnography?

This variety of available ethnographies dispersed across disciplines and countries makes visible the potentials of ethnography while also bringing to the fore questions of “what counts as ethnography?” or what are the parameters or characteristics of ethnography? These questions became particularly astute since the 1980s, when the adaptations of methods deriving from ethnography resulted in work that had little resemblance to the principles, practices, and logic of ethnography (Agar, 2001; Hammersley, 2006; Zaharlick and Green, 1991). One of the major influences in this dispersal of ethnography was the spreading popularity of qualitative research. While the growth of qualitative methodologies over the past five+ decades has brought positive developments and new perspectives to the study of complex social worlds, the association of ethnography with qualitative research created new challenges for ethnography (Hammersley, 2018). Ethnography has been an epistemology that enabled ethnographers to use a variety of methods (now called qualitative and quantitative) and dynamic design to explore many possible angles of overarching research questions. Ethnography was (and mainly still is) a multi-method approach (Agar, 2001; Walford, 2020) and even those from “traditional behavioral social science” (Agar, 2013) and funding organizations acknowledge it as scientific, systematic and rigorous (Shavelson and Towne, 2002).

Unfortunately, over the last few decades, the conflation of ethnography with “qualitative methods” has often led to misunderstandings and misuse of ethnography, overtime impacting its reputation and use. Without understanding the histories, theories, and epistemology of ethnography (Birdwhistell, 1977), too many researchers started calling their “qualitative” studies “ethnographic” simply because they used some fieldwork techniques or interviews, or because ethnography was viewed as a legitimate, systematic

Table 1 Kinds of ethnographies classified by focus.

<i>Classification focus</i>	<i>Kinds of ethnographies</i>
Site	organizational; urban; school; classroom; multi-sited; cyber; virtual
Discipline or field	educational; medical; ethnomethodological; linguistic; sociolinguistic
Theories or epistemology	critical; realist; cognitive; feminist; postmodern
Primary mode of inquiry	visual; sensory; video; hybrid; multimodal
Scope	focused; macro-; micro-
Phenomenon of focus	interactional; educational; performance; design; - of communication; - of ...
Relationship (or person) focus	collaborative; participatory; insider; duo ^a -, auto ^a -
Purpose	practical; applied; rapid; in/of/for education

^aDuo- and auto-ethnographies tend to be reflective accounts of an individual and may or may not study cultural groups or contribute to understanding of the patterns of activity, language, and participation in a cultural group. Auto-ethnographies which focus primarily on the author and their self-examination, personal experiences, and reflection may be closer to autobiographical or narrative accounts rather than ethnography.

research approach with which they sought to be associated. This led to “blitzkrieg ethnographies” (Rist, 1980) and a variety of studies that claim to be ethnographies but have very little to nothing ethnographic about them (Agar, 2001; Atkinson, 2015).

This spread and misuse precipitated decades of discussions and debates of what is and is not “real” ethnography. Anthropologist Michael Agar argued, the question of what is “real ethnography” is flawed since it assumes a particular definition and a predefined set of methods rather than a dynamic, adaptable logic of inquiry for studying complex social life (Agar, 2006b). Nevertheless, many have tried to define and establish parameters for ethnography, as briefly introduced in the next section. To this day, there are no clear-cut answers and no black-and-white decision points of what is “real” ethnography; nevertheless, there are a number of explicit descriptions of expectations for work that fits within the ethnographic space of acceptable ethnographies.

Attempts at defining ethnography

Agar (2006b) argued, “more than one ethnography is possible, *not all ethnographies are acceptable*” (par. 26, italics in original). What is acceptable has been presented differently by different ethnographers, groups, and publication outlets. At its most basic definition, ethnography is writing about a group of people, based on the etymology of the term derived from the Greek *ethnos*, referring to a people or culture, and *graphy*, referring to writing. Varied definitions, presentations of history, and multiple perspectives on ethnography abound in ethnography-focused journals and multidisciplinary volumes such as the *Handbook of Ethnography* (Atkinson et al., 2007), *International Encyclopedia of Anthropology* (Callan and Coleman, 2018), or *Wiley Handbook of Ethnography of Education* (Beach et al., 2018b), among others. Here, instead of reiterating the histories the readers can find in many other volumes, I present two prominent explications of ethnography and include long quotes from the original texts so the reader can follow the argument as it was constructed. The first, from the *Ethnography and Education* journal, outlines parameters for ethnography as conceptualized by journal editors who work at the intersections of sociology and education in the UK and Europe. The second is a description of ethnography offered by the American Anthropological Association in a statement for Institutional Review Boards. The commonalities between these two presentations of ethnography in two different organizations, countries, and kinds of publications lay the ground for understanding ethnography as an epistemology, with a particular kind of logic, not a mere method or set of criteria, – an argument to be developed in the subsequent section.

In the inaugural issue of the *Ethnography and Education* journal, the editors (Troman et al., 2006) outlined “the key elements of ethnographic research applied to the study of education contexts” (p. 1, bullets in the original):

- the focus on the study of cultural formation and maintenance;
- the use of multiple methods and thus the generation of rich and diverse forms of data;
- the direct involvement and long-term engagement of the researcher(s);
- the recognition that the researcher is the main research instrument;
- the high status given to the accounts of participants’ perspectives and understandings;
- the engagement in a spiral of data collection, hypothesis building and theory testing – leading to further data collection; and
- the focus on a particular case in depth but providing the basis for theoretical generalization.

These key elements emphasize the focus on culture, multiple methods, long-term engagement in the field, participant accounts, and the iterative processes of working with the data while recognizing the influence of the researcher throughout the process. These elements, drawn from a history of ethnographic research in sociology in the UK are commonly cited as parameters for work that would be considered ethnographic and publishable in the *Ethnography and Education* journal.

Anthropological perspectives on ethnography have much in common with those of sociologists, even if the emphasis on the role of theory and engagement with the people may differ in some ways (Delamont and Atkinson, 1980; Green and Bloome, 1997). The American Anthropological Association (AAA), which includes the Council for Anthropology and Education whose members have been instrumental in developing ethnographic work in education over the past century, has provided an overview of ethnography in a *Statement on Ethnography and Institutional Review Boards* (American Anthropological Association, 2004). Posing a question “What is ethnography?” the Statement explains:

Ethnography involves the researcher’s study of human behavior in the natural settings in which people live. Specifically, ethnography refers to the description of cultural systems or an aspect of culture based on fieldwork in which the investigator is immersed in the ongoing everyday activities of the designated community for the purpose of describing the social context, relationships and processes relevant to the topic under consideration. Ethnographic inquiry focuses attention on beliefs, values, rituals, customs, and behaviors of individuals interacting within socioeconomic, religious, political and geographic environments. Ethnographic analysis is inductive and builds upon the perspectives of the people studied. Ethnography emphasizes the study of persons and communities, in both international and domestic arenas, and involves short or long-term relationships between the researcher and research participants.

Multiple methods are used in ethnographic research. These include but are not limited to the following: unobtrusive direct observation, participant observation, structured and unstructured interviewing, focused discussions with individuals and community members, analysis of texts, and audio-visual records. Ethnographic methods can be employed in non-traditional ways in interdisciplinary projects that bridge the sciences and humanities.

The complexity and length of ethnographic research engenders an approach to ethics that is both dynamic and flexible. The process of obtaining informed consent may be continuous and incremental throughout the course of the research, and review of consent obtained may be periodic.

In this statement, AAA emphasizes the focus on human behavior, rituals and values in natural settings, fieldwork through immersion, focus on actions, relationships, and belief systems contextualized within multilayered environments. The Statement also emphasizes the relationality between the researcher and participants, the building of explanations from participant perspectives as well as the dynamic and flexible inquiry process, which includes the use of multiple methods to develop empirically grounded representations “of cultural systems or an aspect of culture.” The last part of the AAA definition of ethnography emphasizes the complexity of ethnographic research and the need for a continuous and flexible approach to ethics. This statement has much in common with the key elements of ethnography outlined in the *Ethnography & Education* journal. Both emphasize:

- over-time immersion in the field,
- study of cultural groups and processes,
- utilization of multiple methods,
- multiple perspectives,
- a dynamic nonlinear process of data collection and analysis,
- primacy of participant accounts, and
- ongoing relationship-building and ethics.

While many ethnographers draw on these elements to distinguish what is and is not ethnography, others have critiqued these elements as too prescriptive and/or too rigid to respond not only to the new developments in theories, technologies and ways of thinking, but also to the contemporary life with its constant movement and blurry spatial and temporal boundaries locally and globally (Agar, 2019; Beach et al., 2018a). Debates about such parameters point to ambiguities in concepts such as: culture, direct involvement, long-term, hypothesis, role of theory, and positionality of the researcher (Agar, 2006a, 2006b; Eisenhart, 2018). As one way of addressing these debates, Agar (2006b) has offered a way of looking at ethnography through the lens of its logic of inquiry rather than through the particular parameters that may or may not apply in contemporary ethnographic studies.

Ethnographic logic of inquiry

Drawing on examples of different of ethnographic studies conducted at the same site with different representations of the culture studied, Agar (2006b) explained how more than one ethnography is possible in any particular site. The ethnographies “will be different because of different combinations of ethnographer and community, different ways that a study moves, different choices and different contingencies along the way, different events in the world around the study – any or all of these can change the *trajectory* of a study over time. More than one path is possible” [par.23]. However, not every path and resulting study is acceptable as ethnography. Instead of parameters such as the ones introduced in the previous section, Agar proposed the logic of inquiry and points of view examined as ways of exploring which kinds of studies belong in an ethnographic space. He used the idea of an ethnographic space of possibilities to highlight that there are no clear binary points determining what is and is not ethnography. Instead, the logic and perspectives taken can help determine if a study is part of the space of acceptable ethnographic possibilities.

Agar (2006b) argued that the *iterative, recursive and abductive logic* and the *emphasis on the points of view different from those of an ethnographer* are key determinants of a study’s fit within an ethnographic space. The emphasis on the iterative, recursive, and abductive logic echoes the parameters for the “spiral of data collection, hypothesis building and theory testing – leading to further data collection” as presented by the *Ethnography and Education* editors (Troman et al., 2006) and *American Anthropological Association’s* (2004) call for a “dynamic and flexible” approach to ethics due to the “complexity and length” as well as the “inductive” process of an ethnographic study. The ethnographic logic refers to *iterative* processes needed for ethnographers to explore different perspectives at multiple levels of scale, *recursive* actions in building new understandings while retaining the contextual foundations, and *abductive* logic which allows the researcher to chase the rich points, develop new understandings, and fold them into the developing representations of the ways people engage in culturing processes moment-by-moment and over time (Skukauskaitė and Green, 2023). Agar (2006b, 2013) emphasizes the abductive, rather than inductive, logic of inquiry since, unlike the closed systems of inductive and deductive research processes, abduction enables multi-directional, nonlinear and dialogic explorations of:

- existing, developing, and new theories;
- differences in perspectives between the ethnographer and people studied as well as among the people;
- varied sources of data iteratively generated throughout nonlinear ethnographic processes;
- adaptable and developing methods of analyses driven by the iterative, recursive and abductive logic;
- emphasis on the point of views of the studied group’s members;
- change moment-by-moment and overtime; and
- polyvocal ethnographic representations that are culturally and contextually responsive to the group studied and accessible to the ethnographer’s audiences.

The complex dynamic processes of ethnographic inquiry also maintain the emphasis on the point of view different from that of an ethnographer (Agar, 2006b). Ethnographers do not test preexisting theories. They go into the field with the overarching question “What is going on here?” and seek to understand the often invisible processes and practices of everyday life. Ethnographers work at uncovering the taken-for-granted, mundane, invisible aspects constituting the complex social life of a group, thus making the ordinary extraordinary and the extraordinary (or unusual) ordinary and understandable.

Ethnographers take a learning stance and seek to learn something new with and from the people (Agar, 1996; Baker et al., 2023; Skukauskaitė and Green, 2023); they ask contextually grounded questions of who does and says (or cannot do/say) what, in what ways, when, where, with whom and/or what resources, for what purposes, under what circumstances, and with what outcomes and intended or unintended consequences (Green et al., 2012; Skukauskaitė and Rupšienė, 2023). In exploring these questions, the ethnographer looks for rich points (Agar, 1994), or moments that indicate differences in points of view between the insiders and the understandings of the ethnographer. Noticing these differences, surprises, and shifts in frame of reference, ethnographers chase the varied rich points utilizing the iterative, recursive and abductive logic until an explanation grounded in insider actions, language and perspectives can be constructed to demonstrate new knowledge (Green et al., 2020; Skinner, 2023).

The ethnographic logic of inquiry is iterative, recursive and abductive, and seeks new understandings arising from encounters with different perspectives. Based on Agar’s writing, the logic and point of view create the boundaries for the ethnographic space in which varied parameters and methods could be used in myriad ways to construct trajectories of an ethnographic study. One additional way of distinguishing what work belongs in the ethnographic space is to consider the study’s design and epistemological foundations.

In or out of an ethnographic space: ethnography, ethnographic studies, and use of methods derived from ethnography

Ethnographers across disciplines have long argued that ethnography is not a method but an epistemology, logic of inquiry, philosophy and a way of thinking, knowing, and living (e.g., Agar, 2001; Anderson-Levitt, 2006; Atkinson, 2017; Green et al., 2012). Using only methods, such as participant observation or artifact analysis, derived from anthropological or sociological ethnographic traditions does not make a study ethnographic, however much a scholar may want to associate themselves with ethnographic scholarship. Green and Bloome (1997) proposed a helpful distinction between two ways of engaging in ethnography that do fit within the ethnographic space and one choice for research that does not. *Doing ethnography* and *adopting an ethnographic perspective* are the kinds of studies within an ethnographic space of possibilities. However, studies that simply *use methods* derived from ethnography but do not encompass ethnographic goals and logic are not part of the space.

Doing ethnography (also called *comprehensive* or *full ethnography*) entails engaging in a study of a cultural group over time, conceptualizing culture(s) through particular theoretical and disciplinary perspectives, and conducting research through long-term participation and observation in the field. The scope of “long-term” may differ based on disciplinary parameters of the scholar and/or the studied group’s natural cycles of activity (Bloome et al., 2005; Kelly and Green, 2019b), such as, for example, a semester-long course in higher education or a year in a K-12 school. The “field” may be dispersed across physical and virtual sites (Eisenhart, 2017; Erstad and Sefton-Green, 2013), and observation, interviewing, and artifact collection may take many forms (including video, photo, and digital tools) in response to the studied group’s patterns of life (Baker et al., 2008; Bridges, 2023; Skukauskaitė and Sullivan, 2023; Tobin, 2019). In Agar’s terms, many “trajectories” are possible, but overall, a full-scale ethnography requires a study of “culture” (however conceptualized, Agar, 2006a, 2019) and broad, holistic analysis of patterns of everyday life in all its complexity and multi-layered contexts. The “whole” studied could be a whole society (as in classical ethnographies), a particular organization such as a school or business, or a particular classroom studied over time to uncover patterns of practice and insider understandings about life in their group.

Adopting an ethnographic perspective or *engaging in an ethnographic study* entails a more focused exploration of particular aspects of the group’s everyday life while still following the iterative, recursive and abductive logic, and drawing on theories of culture to conceptualize and represent the patterns of activity of the group studied. An ethnographic study, like a comprehensive ethnography, entails participating in the field over time, relationship building with the insiders (American Anthropological Association, 2004; Skukauskaitė and Sullivan, 2023), utilizing multiple methods (Agar, 2001; Walford, 2020), and chasing the rich points (Agar, 1994; Skinner, 2023) to construct understandings from insider points of view. Ethnographic studies may focus on particular activities or practices (e.g., specific literacy or mathematics events in a classroom, such as morning circle or number talks), particular places or sites (e.g., a bird identification group on a digital platform), or specific phenomena (such as gendering in science classrooms, or reflexivity in a doctoral research course). No matter the focus and scope of time, a study which adopts an ethnographic perspective and fits in the ethnographic space will draw on concepts of culture, engage in the iterative, recursive and abductive logic, and construct new knowledge responsive to insider points of view and contexts impacting the group studied.

A kind of study that does not fit into an ethnographic space is one in which the scholars merely *use methods* derived from ethnographic traditions for purposes not aligned with ethnographic logic and goals. Use of methods does not make a study ethnographic. Methods are tools for data collection and analysis and can be used in varied ways across different kinds of research designs. For example, fieldwork and observation methods, often associated with ethnography, can and have been used in studies ranging from experimental designs and process-product studies (Evertson and Green, 1986; Rex et al., 2006) to video-based qualitative studies in teacher education (Kumpulainen et al., 2012; Richards et al., 2021). Use of participant observation, conversational

interviewing, field notes, artifact analysis (among other methods used in ethnography), or a combination of methods without the underlying ethnographic logic and focus on cultural patterns of life, is outside the boundaries of the ethnographic space. Qualitative interview studies of individuals and their perspectives about something, observational studies of predetermined phenomena and variables, and multi-method or mixed method studies of populations or interventions are not ethnographic.

Doing ethnography or engaging in an ethnographic study requires understanding of the histories, theories, and epistemology of ethnography; employing the iterative, recursive and abductive logic; and seeking to construct new knowledge by examining differences in ethnographer's and participants' points of view. Ethnographers study complex patterns of practice a cultural group (not individuals by themselves) co-constructs through their discourse and actions in the flow of their everyday life, situated in multilayered contexts visible and invisible to insiders (Skukauskaitė and Girdzijauskienė, 2021). Ethnography is an epistemology, not a method. As an epistemology, it calls for theoretically grounded, consistent, and transparent logic of inquiry guided by particular ethnographic goals and ways of representing the complexities of everyday life. In the next section I introduce three kinds of possible goals of educational ethnography. While I draw on scholars in education, the distinctions they make are applicable for ways of thinking about the purposes of ethnography in other disciplines as well.

Primary goals for ethnography: ethnography *of*, *in*, and *for* education (or other disciplines)

In 1997, Green and Bloome (1997) proposed two ways of considering ethnographic spaces for those working in Education and in 2018, Bloome and colleagues (Bloome et al., 2018) expanded the arguments to Anthropology. Green and Bloome directed scholars' attention to a small word – a preposition (*of*, *in*, *for*) – whose use can help determine intended differences among goals of educational ethnographies. They argued that ethnography *of* education is typically conducted by scholars outside the discipline of education. Those scholars may use educational settings such as schools as a research site and study phenomena of interest to their particular disciplines (e.g., a psychologist studying motivation in a school setting, a kinesiologist studying physical activity of children at recess, or a sociologist studying formal and informal organizational/group structures within an educational site). Ethnography *of* education is driven by the histories, theories and goals of the scholars' home disciplines and theories outside education.

In contrast, ethnography *in* education is conducted by scholars within education. Education researchers use the knowledge, history and theories of education to explore questions of relevance to educational practice, policy, and/or theory. Ethnographers *in* education work *with* teachers, students, parents, administrators and other stakeholders to answer questions arising from engagement in the work of teaching, learning, and leading in educational settings. After all, as Ball and Forzani (2007) argued, what makes education research educational is primarily the education scholars' ability to "probe the instructional dynamic *inside* educational situations and problems" (p. 530, italics mine) and examine this complex "instructional dynamic" from multiple perspectives. Probing inside and examining educational processes and practices in context and from multiple points of view responsive to the insiders may in turn aid in informing and improving educational practice. Ethnography *in* education involves exploring these instructional dynamics and their contexts from insider questions and points of view, including perspectives and actions of teachers, students, administrators, parents, and others. While both ethnography *of* and *in* education (or any other discipline) draw on the history; cultural theories; iterative, recursive, and abductive logic; and methods stemming from ethnography's roots in anthropology and sociology, the goals, phenomena investigated, theories utilized, and involvement of and with the people will differ.

Exploring how anthropologists have studied education, Bloome and colleagues (2018) proposed a similar heuristic, adding *for* to the *of* and *in* prepositions: anthropology *of* education, *in* education, and *for* education. Writing for anthropology, Bloome et al. (2018) explained, "Anthropology of education refers to theory building in anthropology; anthropology in education refers to theory building in education; and anthropology for education refers to the enactment of educational change" (p. 2). In their encyclopedia chapter, Bloome et al. demonstrated how intersections of anthropology and education in different ways have enabled different kinds of theorizing and understanding of both anthropology and education. Here anthropology as a logic and process of inquiry is synonymous with ethnography; such use is not uncommon in the discipline of anthropology (e.g., Anderson-Levitt, 2011; Anderson-Levitt and Rockwell, 2017). Explaining anthropology (ethnography) *in*, *of*, and *for* education, Bloome et al. (2018) summarized the contributions of anthropology (ethnography) as follows (p. 7, bullets and italics mine):

- Scholarship in anthropology *of* education has deepened the understanding of how schooling and other educational processes contribute to and are influenced by socialization and cultural transmission process more generally.
- Scholarship in anthropology *in* education has built on and incorporated theoretical constructs, epistemologies, and methodologies from anthropology, reconceptualizing core education concepts and crafting a laminated theoretical framework that includes anthropological theories, the intellectual histories of educational research, other education-related disciplines, and the lived experiences of educational researchers and educators.
- Studies associated with anthropology *for* education have crafted instructional theories, models and cases that honor and build on the cultural diversity of students' communities and challenge deficit models of students from nondominant communities; they have also connected academic teaching and learning with critical social justice.

The distinctions in the goals, histories, theories, and disciplinary roots of varied ethnographies uncovered through the change of a pronoun are useful heuristics – processes of enabling discovery and learning of something nuanced that otherwise remains invisible or is taken for granted. Table 2 presents a summary of the distinctions among the underlying goals of different kinds of ethnographies within and across disciplines.

Table 2 Differentiating primary goals of ethnography within a discipline, such as education

<i>Ethnography</i>	<i>Who conducts</i>	<i>Question origins</i>	<i>Example</i>
<i>of</i>	Scholars outside the discipline	Theories from home discipline	Sociologist studying group composition and structures on a school campus
<i>in</i>	Scholars within the discipline	Problems of practice identified in collaboration with insiders	Education researcher studying how children co-construct their understanding of mathematics in an elementary classroom
<i>for</i>	Scholars within or outside the discipline, often in collaboration with community	Need for change in practice and/or theories within the discipline	A team of researchers, teachers, and community members collaborating on a community based social justice project

The heuristics of changing the pronoun when conceptualizing *ethnography of/in-/for*-a particular discipline, enable exploring the nuanced differences in the ways ethnographies are conceptualized, carried out and presented for particular audiences across fields of study. However, as [Bloome et al. \(2018\)](#) and [Green and Bloome \(1997\)](#) acknowledge, in practice, the distinctions blur as researchers collaborate and draw on knowledge and theories from across disciplines:

Social scientists *of education* often collaborate with researchers and practitioners in education and employ many kinds of knowledge generated within the disciplines *of education*; and those *in education* often employ frames, modes of inquiry and findings from disciplines within the fields of anthropology and sociology. Thus, rarely is any ethnography or ethnographic study *of education* or only *in education*. Nonetheless, the distinction is useful in examining the intellectual sites within which ethnography and ethnographic studies on education have been located (p. 186, emphasis in original).

[Green and Bloome \(1997\)](#) writing for education, [Bloome and colleagues \(2018\)](#) writing for anthropology, [Heath and Street \(2008\)](#) writing for language and literacy studies, and [McCarty \(2014\)](#) writing for educational linguistics (among many other authors writing for diverse disciplines) all highlight the need to explore disciplinary, theoretical, historical and epistemological foundations of ethnography as it is enacted in particular settings, with particular groups of people, and in particular ways. These scholars also emphasize the potentials of working across disciplines and collaborating with others to understand the educational phenomena and social life in multifaceted, layered ways.

Ultimately, despite differences, ethnographies which build on the anthropological and/or sociological roots, engage in the ethnographic logic of inquiry and seek to understand the patterned processes and practices of social life in its situated complexity over time. Ethnography focuses on “what is going on?” in a particular group moment by moment and over time, in the complex dynamic contexts shaping and being shaped by the actions and interactions of varied configurations of people within and beyond the group. I end the chapter with a summary about the kinds of questions ethnography can answer and how it has and can be used beyond publishing-focused academic spaces.

Power and potentials of ethnography

Described throughout this chapter are perspectives that demonstrate the complexity, variation, challenges and possibilities of the ethnographic epistemology within and across fields. At the core of ethnography is its iterative, recursive and abductive logic of inquiry and emphasis on understanding cultural patterns from the points of view of people creating those cultural patterns within particular groups, in specific sites and situations, through their actions and interactions moment-by-moment and over time. Anthropologist of education Kathryn Anderson-Levitt, in a chapter on Ethnography (2006) for the *Handbook of Complementary Methods in Education Research* ([Green et al., 2006](#)) provided an overview what ethnography can offer. In posing the question “what is ethnography good for?”, she explained:

It is important to understand what kinds of questions ethnography can address and what it cannot. Ethnography is useful, first, for discovering what meanings different actors are making of a situation. We cannot solve a problem or evaluate an attempted solution if we do not understand what the real problem is or how actors define it. As [Erickson and Gutiérrez \(2002\)](#) put it, “a logically and empirically prior question to ‘Did it work?’ is ‘What was the “it”?” (p. 21). Second, ethnography is useful for developing a valid understanding of local situations in all their complexity. Ethnography provides the opportunity to explore actions in their wider context and thus to describe the real-world complexity of human behavior. Third, because ethnography takes time, it gives us the opportunity to observe and understand processes as they happen. In short, ethnography is good for asking, “What is going on there?” It is a good method when we are not sure exactly what the problem is. It is good for asking, “How does this happen?” and “what does this mean?” (p. 282).

In providing the long quote here I demonstrate how an anthropologist of education who has engaged in ethnographic studies around the world, presents ethnographic potentials for uncovering complex, situated, real-world understandings of processes, practices, and discourse constituting “what is going on” and what it means for the people within the group. Ethnography’s potentials to uncover the complexity of “human behavior” and overtime processes in the wider contexts influencing and influenced by the actions of the people also invoke questions of power and who gets to conduct and participate in ethnographies.

Anderson-Levitt (2006) argued that “Because ethnography requires the researcher to understand how insiders make sense of a situation, ethnographers can draw attention to and express publicly points of view of less powerful players” (p. 283). While illuminating the lives and perspectives of the “less powerful” is an important contribution to understanding the complexities of social life in diverse social structures and cultures, she also cautions that “studying *only* the powerless can have the unintended consequence of making their behavior more visible to those in power and therefore easier to control it; it can also unwittingly collaborate in keeping the operation of power out of sight” (p. 283). Therefore, ethnographers, individually and collectively, need to study the social and cultural construction of life across varied social levels and power structures.

Considering what, whom, where, and in what ways to study, ethnographers also need to be aware of their own powers and potential influences, and thus work toward power diffusion or sharing when possible. Ethnographers start with a learner stance and position people studied as experts (Agar, 1980; Baker et al., 2023) of their lives and cultures-in-the-making (Collins and Green, 1992). Members of the group are cultural guides who choose if, when, and in what ways to help the ethnographer learn about their worlds (Skukauskaitė and Rupšienė, 2023). Listening to the people, inviting their collaboration and guidance, and partnering in the research process can diffuse some of the power differentials between an academic ethnographer and the members of the group studied (Guerrero et al., 2023; Skukauskaitė and Sullivan, 2023). Over the past few decades more work has shown the agency and power of research partners, including children, students, and teachers (e.g., Egan-Robertson and Bloome, 1998; Guerrero et al., 2023) who sometimes choose to become co-authors on research publications and presentations. But academic, publishing-focused spaces, however significant they may be to the academics, often are not the spaces of meaning or relevance to the cultural groups studied. If ethnographers involve research partners throughout the ethnographic process in transparent ways, the partners/participants construct new understandings of how ethnography can be helpful for their own purposes, be it teaching, understanding systems, learning from and with other points of view, collaborating, and/or exploring taken-for-granted aspects of daily life. The people become ethnographers of their own lives and social groups.

The power of ethnography is in its potential for the people to take it up in their own work and spaces of significance and to (re) shape their own cultural worlds by understanding how and why the systems of which they are part function. Insiders have the power to change their own practices and lives, and ethnography can be a resource (Green and Bloome, 1997) through which they develop understandings of what is happening, in what ways, when and where, in what situations, with whom, and with what intended or unintended outcomes and consequences. As Erickson and Gutteriez (2002) argued, understanding what “it” is and how “it” works in specific, real-world, contextualized lives of people acting and interacting with each other and their social and structural systems, opens doors for changing the “it”. In education, teachers, teacher educators, students, and principals have been taking up ethnography as a process of inquiry, learning, and change (Egan-Robertson and Bloome, 1998; Yeager, 2023). They demonstrate that ethnography is more than a research method – it is an epistemology, a way of knowing, and a way of living and understanding the complex changing worlds in which we live and create. Ethnography listens and honors the work and lives of people in culturally responsive, relational, situated, grounded, and transparent ways.

References

- Agar, M., 1980. *The Professional Stranger: An Informal Introduction to Ethnography*. Academic Press.
- Agar, M., 1994. *Language Shock: Understanding the Culture of Conversation*. Quill.
- Agar, M., 1996. *The Professional Stranger: An Informal Introduction to Ethnography*, second ed. Academic Press.
- Agar, M., 2001. Ethnography. In: *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, pp. 4857–4862.
- Agar, M., 2004. We have met the other and we’re all nonlinear: ethnography as a nonlinear dynamic system. *Complexity* 10 (2), 16–24.
- Agar, M., 2006a. Culture: can you take it anywhere? *Int. J. Qual. Methods* 5 (2), 1–12. http://www.ualberta.ca/~iiqm/backissues/5_2/PDF/agar.pdf.
- Agar, M., 2006b. An ethnography by any other name *Forum Qual. Soc. Res.* 7 (4). <http://www.qualitative-research.net/fqs>.
- Agar, M., 2013. *The Lively Science: Remodeling Human Social Research*. Mill City Press.
- Agar, M., 2019. *Culture: How to Make it Work in a World of Hybrids*, Kindle ed. Rowman & Littlefield.
- American Anthropological Association, 2004. American Anthropological Association Statement on Ethnography and Institutional Review Boards. American Anthropological Association.
- Anderson-Levitt, K.M., 2006. Ethnography. In: Green, J.L., Camilli, G., Elmore, P.B. (Eds.), *Handbook of Complementary Methods in Education Research*. Lawrence Erlbaum & Associates for AERA, pp. 279–296.
- Anderson-Levitt, K.M. (Ed.), 2011. *Anthropologies of Education: A Global Guide to Ethnographic Studies of Learning and Schooling*. Berghahn Books.
- Anderson-Levitt, K.M., Rockwell, E., 2017. Comparing Ethnographies: Local Studies of Education across the Americas. American Educational Research Association.
- Atkinson, P., 2015. *For Ethnography*. Sage.
- Atkinson, P., 2017. Thinking Ethnographically. Sage. <https://doi.org/10.4135/9781473982741>.
- Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), 2007. *Handbook of Ethnography*. Sage.
- Baker, W.D., Green, J.L., Skukauskaitė, A., 2008. Video-enabled ethnographic research: a microethnographic perspective. In: Walford, G. (Ed.), *How to Do Educational Ethnography*. Tufnell Press, pp. 77–114.
- Baker, W.D., Machtmes, K., Green, J.L., 2023. On ethnographer-as-learner and theory builder. In: Skukauskaitė, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 43–69.
- Ball, D.L., Forzani, F.M., 2007. What makes education research “educational”? *Educ. Res.* 36 (9), 529–540. <https://doi.org/10.3102/0013189x07312896>.

- Beach, D., 2017. International trends and developments in the ethnography of education. *Acta Paedagog. Vilnensia* 39 (2), 15–30. <https://doi.org/10.15388/ActPaed.2017.39.11455>.
- Beach, D., Bagley, C., Marques da Silva, S., 2018a. Ethnography of education: thinking forward, looking back. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell, pp. 515–532.
- Beach, D., Bagley, C., Marques da Silva, S. (Eds.), 2018b. *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell.
- Beach, R., Bloome, D. (Eds.), 2019. *Language Relations for Transforming the Literacy and Language Arts Classroom*. Routledge.
- Bejarano, C.A., García, M.A.M., Juárez, L.L., Goldstein, D.M., 2019. Decolonizing Ethnography: Undocumented Immigrants and New Directions in Social Science. Duke University Press. <https://doi.org/10.2307/j.ctv11smmv5>.
- Birdwhistell, R.L., 1977. Some discussion of ethnography, theory and method. In: Brockman, J. (Ed.), *About Bateson: Essays on Gregory Bateson*. E.P. Dutton, pp. 103–144.
- Bloome, D., Beauchemin, F., Brady, J., Buescher, E., Kim, M.-Y., Schey, R., 2018. Anthropology of education, anthropology in education, and anthropology for education. In: Callan, H. (Ed.), *The International Encyclopedia of Anthropology*. John Wiley & Sons, Ltd, pp. 1–10. <https://doi.org/10.1002/9781118924396.wbiea2140>.
- Bloome, D., Carter, S.P., Christian, B.M., Otto, S., Shuart-Faris, N., 2005. Discourse Analysis and the Study of Classroom Language and Literacy Events: A Microethnographic Perspective. Lawrence Erlbaum.
- Bridges, S., 2023. Video-enabled educational ethnographies: the centrality of recordings in an interactional ethnography. In: Skukauskaite, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 91–113.
- Callan, H., Coleman, S. (Eds.), 2018. *The International Encyclopedia of Anthropology*. <https://doi.org/10.1002/9781118924396>.
- Collins, E.C., Green, J.L., 1992. Learning in classroom settings: making or breaking a culture. In: Marshall, H. (Ed.), *Redefining Student Learning: Roots of Educational Restructuring*. Ablex, pp. 59–85.
- Delamont, S., Atkinson, P., 1980. The two traditions in educational ethnography: sociology and anthropology compared. *Br. J. Sociol. Educ.* 1 (2), 139–152. <https://doi.org/10.1080/0142569800010201>.
- Egan-Robertson, A., Bloome, D. (Eds.), 1998. *Students as Researchers of Culture and Language in Their Own Communities*. Hampton Press.
- Eisenhart, M., 2017. A matter of scale: multi-scale ethnographic research on education in the United States. *Ethnogr. Educ.* 12 (2), 134–147. <https://doi.org/10.1080/17457823.2016.1257947>.
- Eisenhart, M., 2018. Changing conceptions of culture and ethnography in anthropology of education in the United States. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell, pp. 153–172.
- Erickson, F., Gutierrez, K., 2002. Culture, rigor, and science in educational research. *Educ. Res.* 31 (8), 21–24. <https://doi.org/10.3102/0013189x031008021>.
- Erstad, O., Sefton-Green, J. (Eds.), 2013. *Learning Lives: Transactions, Digital Culture, and Learner Identity*. Cambridge University Press.
- Evertson, C.M., Green, J.L., 1986. Observation as inquiry and method. In: Wittrock, M. (Ed.), *Handbook of Research on Teaching*, third ed. American Educational Research Association, pp. 162–213.
- Frank, C., 1999. *Ethnographic Eyes: A Teacher's Guide to Classroom Observation*. Heinemann.
- Green, J.L., Baker, W.D., Chian, M., Vanderhoof, C.M., Hooper, L., Kelly, G.J., Skukauskaite, A., Kalainoff, M., 2020. Studying the over-time construction of knowledge in educational settings: a microethnographic-discourse analysis approach. *Rev. Res. Educ.* 44, 161–194. <https://doi.org/10.3102/0091732X20903121>.
- Green, J.L., Bloome, D., 1997. Ethnography and ethnographers of and in education: a situated perspective. In: Flood, J., Heath, S.B., Lapp, D. (Eds.), *Handbook of Research on Teaching Literacy through the Communicative and Visual Arts*. International Reading Association & MacMillan, pp. 181–202.
- Green, J.L., Bridges, S.M., 2018. Interactional ethnography. In: Fischer, F., Hmelo-Silver, C.E., Goldman, S.R., Reimann, P. (Eds.), *International Handbook of the Learning Sciences*. Routledge, pp. 475–488.
- Green, J.L., Camilli, G., Elmore, P.B., with, Skukauskaite, A., Grace, E. (Eds.), 2006. *Handbook of Complementary Methods in Education Research*. Lawrence Erlbaum & Associates for AERA.
- Green, J.L., Skukauskaite, A., Baker, W.D., 2012. Ethnography as epistemology: an introduction to educational ethnography. In: Arthur, J., Waring, M.J., Coe, R., Hedges, L.V. (Eds.), *Research Methodologies and Methods in Education*. Sage, pp. 309–321.
- Guerrero, A.L., Peña, I.N., Dantas-Whitney, M., 2023. Collaborative ethnography with children: building intersubjectivity and co-constructing knowledge of place. In: Skukauskaite, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 163–182.
- Hackett, P.M.W., Hayre, C.M. (Eds.), 2020. *Handbook of Ethnography in Healthcare Research*. Taylor and Francis. <https://doi.org/10.4324/9780429320927>.
- Hammersley, M., 2006. Ethnography: problems and prospects. *Ethnogr. Educ.* 1 (1), 3–14. <https://doi.org/10.1080/17457820500512697>.
- Hammersley, M., 2018. What is ethnography? Can it survive? Should it? *Ethnogr. Educ.* 13 (1), 1–17. <https://doi.org/10.1080/17457823.2017.1298458>.
- Heath, S.B., Street, B.V., 2008. *On Ethnography: Approaches to Language and Literacy Research*. Teachers College Press.
- Kelly, G.J., Green, J.L., 2019a. Framing issues of theory and methods for the study of science and engineering education. In: Kelly, G.J., Green, J.L. (Eds.), *Theory and Methods for Sociocultural Research in Science and Engineering Education*. Routledge, pp. 1–28.
- Kelly, G.J., Green, J.L. (Eds.), 2019b. *Theory and Methods for Sociocultural Research in Science and Engineering Education*. Routledge.
- Kumpulainen, K., Toom, A., Saalasti, M., 2012. Video as a potential resource for student teachers' agency work. In: Hjørne, E., Van Der Aalsvoort, G., De Abreu, G., Hjørne, E. (Eds.), *Learning, Social Interaction and Diversity - Exploring Identities in School Practices*. Sense, pp. 169–188.
- Marcus, G.E., 2008. The end(s) of ethnography: social/cultural anthropology's signature form of producing knowledge in transition. *Cult. Anthropol.* 23 (1), 1–14.
- Marcus, G.E., 2021. *Ethnography through Thick and Thin*. Princeton University Press.
- McCarty, T.L., 2014. Ethnography in educational linguistics. In: Bigelow, M., Enns-Kananen, J. (Eds.), *The Routledge Handbook of Educational Linguistics*. Routledge, pp. 23–37. <https://doi.org/10.4324/9781315797748.ch2>.
- McCarty, T.L., Liu, L., 2017. Ethnography of language policy. In: King, K.A., Lai, Y.-J., May, S. (Eds.), *Research Methods in Language and Education*. Springer International Publishing, pp. 53–66. https://doi.org/10.1007/978-3-319-02249-9_4.
- Milstein, D., Clemente, A., 2018. Latin American educational ethnography. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell, pp. 195–214.
- Modiba, M., Stewart, S., 2018. Ethnography and education in an African context. In: Beach, D., Bagley, C., Marques da Silva, S. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell, pp. 257–322.
- Rex, L.A., Steadman, S.C., Graciano, M.K., 2006. Researching the complexity of classroom interaction. In: Green, J.L., Camilli, G., Elmore, P.B. (Eds.), *Handbook of Complementary Methods in Education Research*. Lawrence Erlbaum & Associates for AERA, pp. 727–773.
- Richards, J., Altshuler, M., Sherin, B.L., Sherin, M.G., Leatherwood, C.J., 2021. Complexities and opportunities in teachers' generation of videos from their own classrooms. *Learn. Cult. Soc. Interact.* 28, 100490. <https://doi.org/10.1016/j.lcsi.2021.100490>.
- Rist, R.C., 1980. Blitzkrieg ethnography: on the transformation of a method into a movement. *Educ. Res.* 9 (2), 8–10. <https://doi.org/10.2307/1174643>.
- Shavelson, R.J., Towne, L. (Eds.), 2002. *Scientific Research in Education*. National Academy Press.
- Skinner, K., 2023. Anchoring analysis in rich points. In: Skukauskaite, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 242–265.
- Skukauskaitė, A., 2021. Becoming an ethnographer: living, teaching, and learning ethnographically. In: Sancho-Gil, J.M., Hernández- Hernández, F. (Eds.), *Becoming an Educational Ethnographer: The Challenges and Opportunities of Undertaking Research*. Routledge, pp. 52–63.
- Skukauskaitė, A., Girdzijskienė, R., 2021. Video analysis of contextual layers in teaching-learning interactions. *Learn. Cult. Soc. Interact.* 29. <https://doi.org/10.1016/j.lcsi.2021.100499>.

- Skukauskaitė, A., Green, J.L., 2023. Ethnographic spaces of possibilities: interactional ethnography in focus. In: Skukauskaitė, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 1–18. <https://doi.org/10.4324/9781003215479>.
- Skukauskaitė, A., Rupšienė, L., 2023. Understanding interactional ethnography as a languaculture with a bilanguacultural guide. In: Skukauskaitė, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 21–42.
- Skukauskaitė, A., Sullivan, M., 2023. Conversational interviewing grounded in interactional ethnographic principles. In: Skukauskaitė, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 114–141.
- Street, B., 1993. Culture is a verb. In: Graddol, D. (Ed.), *Language and Culture. Multilingual matters/BAAL*, pp. 23–43.
- Tobin, J., 2019. The origins of the video-cued multivocal ethnographic method. *Anthropol. Educ. Q.* 50 (3), 255–269. <https://doi.org/10.1111/aeq.12302>.
- Troman, G., Gordon, T., Jeffrey, B., Walford, G., 2006. Editorial. *Ethnogr. Educ.* 1 (1), 1–2. <https://doi.org/10.1080/17457820500512671>.
- Walford, G., 2020. Ethnography is not qualitative. *Ethnogr. Educ.* 15 (1), 122–135. <https://doi.org/10.1080/17457823.2018.1540308>.
- Wolcott, H.F., 2008. *Ethnography: A Way of Seeing*, second ed. AltaMira Press.
- Yeager, E.V., 2023. Interactional ethnography as a resource for learning in K12: building communities of inquiry. In: Skukauskaitė, A., Green, J.L. (Eds.), *Interactional Ethnography: Designing and Conducting Discourse-Based Ethnographic Research*. Routledge, pp. 285–298.
- Zaharlick, A., Green, J.L., 1991. Ethnographic research. In: Flood, J., Jensen, J., Lapp, D., Squire, J.R. (Eds.), *Handbook on Teaching the English Language Arts*. Macmillan, pp. 205–225.

Autoethnography: storied scholarship

Sandra L. Pensoneau-Conway, Southern Illinois University, Carbondale, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	102
Autoethnography defined	102
Autoethnography: an historical view	103
Autoethnography and voice	103
Evocative autoethnography	104
The what of autoethnography	105
Outlets for autoethnographic scholarship	105
Autoethnographic futures	105
References	105

Introduction

The storied and situated self.

The personal and political process.

The micro and macro enmeshed.

The theoretical and theorized thickness.

Autoethnography articulates the intersection of cognitive and somatic experience. Through narrative craft, it stories a life lived, and in so doing, engages theory to make sense of how that life is socio-historically situated. It demonstrates the systemic consequences on everyday interactions, while simultaneously showing how a constellation of interactions both sustain and change those very systems. Importantly, autoethnography provides a snapshot of understanding and critique, embodying Fisher's (1984) description of *homo narrans* (p. 6), humans as narrating (and narrated) beings.

In this chapter, I first provide a broad definition of autoethnography in order to orient the reader to this method. Next, I briefly trace the history of autoethnography. I then explicate the notion of autoethnographic voice, while overviewing common instantiations of voice in autoethnographic work. I outline the evocative element of autoethnography before identifying the sort of content that autoethnographers write about. I end by providing a snapshot of where one might look to further explore autoethnography as research practice.

Autoethnography defined

Broadly speaking, autoethnography relies upon conventions of creative and performative writing (Pollock, 1998) to story personal experience while simultaneously (though perhaps implicitly) theorizing such experience. Autoethnographers believe that experience is a valid source of data. Situating such data within a variety of contexts such as historical (what historical conditions have contributed to how one understands and interprets this experience?), cultural (how do cultural assumptions and affiliations impact one's understanding and interpretation of this experience?), and political (how do power imbalances inform how one understands and interprets this experience?), highlights the relationship between the personal and the cultural. At the same time, autoethnographers understand that these elements (e.g., personal, historical, cultural, political, economic, sociological, and so on) can never be so easily compartmentalized. In short, autoethnography is a way of theorizing the world at large through the lens of the personal. Theory helps us explain; story helps us personalize theory.

Autoethnographers often make use of conventions of creative writing in their scholarship. First-person voice is a common element, though purposeful use of second and third voice can have a powerful impact [see, for example, Pelias (2000), for a compelling use of second person]. Dialogue may make its way into an autoethnographic essay [see, for example, Pensoneau-Conway and Valenta (2021) for how they use dialogue to show their relationship development]. Variation in the use of font, spacing, boldface and italics, and other typesetting options brings nuance to an autoethnography, and invites the reader to reflect upon not just the overall message of a piece, but also what particular differences may symbolize about the intention of the author, and for the interpretation the reader makes.

For example, Gingrich-Philbrook's (2015) chapter opens in third person: "One hears a lot about home" (p. 199), which serves (to this reader) as a broad opening, allowing any reader to place themselves in the subject position of "one." The second sentence shifts into first-person plural and third-person plural: "If we have to go there, for example, they have to take us in" (p. 199). The use of first-person plural creates a sense of belonging for those who may identify with the "we," along with a sense of separation between "we" and "they." While reference to "me" is made in the first paragraph, Gingrich-Philbrook gives a greater sense that this is a personal essay in the first sentence of the second paragraph: "Working on this essay, I learned I am the very last person

a reader should trust to articulate what home does or does not comprise or afford—at least that ‘home’ normative forms of family and patriarchy want us to think of when we hear that word” (p. 199–200). Here, Gingrich-Philbrook further strengthens the connection between himself and the reader (“I” and “we”) while also invoking cultural manifestations of patriarchy and the construct of *family*. This example demonstrates the care and consideration autoethnographers take with each word, as it stands alone and as it exists in relation to other words, all in the service of forwarding a particular theme or message, while still remaining open for the reader to interpret the text in their own particular way, as well.

Autoethnography: an historical view

Autoethnography’s origins are somewhat contested. The term itself—“auto” + “ethno” + “graphy,” or, the self writing about culture—has roots in ethnographic research practice. David Hayano (1979) is credited with bringing autoethnography into effectual visibility in the field of anthropology, a fitting location given anthropology’s focus on culture. Hayano identifies his first encounter with the term as being in a 1966 structuralism seminar. In that setting, the lecturer, Sir Raymond Firth, used *autoethnography* to refer to an element of Jomo Kenyatta’s 1938 study of his own Kikuyu people. The 1938 occurrence referenced the “question of judging the validity of anthropological data by assessing the characteristics, interests, and origin of the person who did the fieldwork” (p. 100). In this sense, autoethnography referred to the anthropologist’s concern with how an ethnographer writes ethnographically about one’s “own people” (p. 99). Yet, in Hayano’s explication of autoethnography, there is still the prominent element of a group, of which the ethnographer is but a part. The ethnographer recognizes themselves as a member of the group, yet the group still remains the focus. The self comes as a result of the group, rather than starting with the self.

Others credit literary and feminist scholar Françoise Lionnet (1989) with using *autoethnography* in a way that’s more in line with autoethnography of the past 30 years. Lionnet used the term to describe the 1942 autobiography of writer and anthropologist Zora Neale Hurston, *Dust Tracks on a Road*. Lionnet writes, “what I would like to suggest here is that *Dust Tracks* amounts to autoethnography, that is, the defining of one’s subjective ethnicity as mediated through language, history, and ethnographical analysis; in short, that the book amounts to a kind of ‘figural anthropology’ of the self” (p. 99). In this way, the self provides an account of their subjective location within a group, rather than foregrounds the group that happens to be comprised of selves. This situated self—different from a group/culture of selves—is the focus of autoethnography as it is most commonly manifested today.

And yet, the elements of self and other, self and group, self and culture cannot be so easily disconnected. Indeed, autoethnography aims to illuminate culture through the narration of experience, and is “an evocatively rendered, aesthetically compelling, and revelatory encounter” (Adams and Holman Jones, 2008, p. 375). It is famously described by autoethnographic pioneer Carolyn Ellis (2004) as a “back and forth”:

Back and forth autoethnographers gaze: First they look through an ethnographic wide angle lens, focusing outward on social and cultural aspects of their personal experience; then, they look inward, exposing a vulnerable self that is moved by and may move through, refract, and resist cultural interpretations. As they zoom backward and forward, inward and outward, distinctions between the personal and cultural become blurred, sometimes beyond distinct recognition. (pp. 37–38)

And it is this back and forth that distinguishes autoethnography from autobiography.

Autoethnography and voice

The autoethnographer’s voice functions as the mechanism by which the back and forth happens, by which culture is communicated. The *what* of the voice becomes important, and sets autoethnography apart from other life writings. In other words, the “ethno” serves as a key marker of autoethnography. The crisis of representation is but one condition that has created the possibility for autoethnography. Pensoneau-Conway and Toyosaki (2011) explain that autoethnography (along with more critical ethnographic work) responds to the crisis of representation, in which scholars are no longer able to ignore the politics of representation (of the self and the other), the interdisciplinarity of scholarship, and the colonizing history of scholarly practices (relevant to this discussion, of ethnography in particular). They trace the development of autoethnography in this light, acknowledging the growing skepticism of the ethics of more traditional ethnographic practices and a healthy questioning of the influences of the researcher on all aspects of the research process. Autoethnography provides one response to the concern regarding the relationship between self and other, by positioning the self *as* other. Thus, the autoethnographic voice provides rich insight into the makings of experience, and autoethnographers understand *voice* to be noun and verb.

Both singular and plural, autoethnography voices experience from multiple perspectives. In solo-authored work, readers may encounter the author writing from a seemingly singular voice, as in Duff’s (2021) narratives surrounding the COVID-19 pandemic, Adams’s (2011) exploration of the closet, or LeMaster’s (2013) evocative work living on/in the border of gender, race, and family. Or, readers will meet the multiple selves of that author, who writes as a multi-voiced individual, as in Carol Rambo’s (1995, 1996) work, known for the usage of the “layered account.”

The “auto” in autoethnography often points to a singular self and/or singular author. The question of authorship raises interesting considerations for autoethnography, and for autoethnographers. Postmodernism inspires some of these questions, including:

- Who owns a story?
- Does a story necessarily have more than one author, since no experience can be entirely isolated from other actors/authors?
- Does the *auto* in autoethnography preclude the question of multi-authorship?
- Does a fractured sense of self—that the self can never be contained into one coherent subject—mean that all stories are necessarily multi-authored?

There are many ways to answer these questions. Some scholars have developed methods resonant with autoethnography, but that address the question of solo-authorship. For example, [Alexander et al. \(2012\)](#) use “tri-autoethnography” to bring their three voices together in a beautifully told exploration of family relationship patterns. [Gale et al. \(2012\)](#) poetically and dialogically paint the collaborative writing process. [Toyosaki et al. \(2009\)](#) developed community autoethnography as an engagement in “dialogical theorization” of whiteness, in order to honor the relational aspect of experience. Similarly, [Chang et al. \(2013\)](#) offer an in-depth methodological primer on collaborative autoethnography, though not to diffuse the specificity of autoethnography. Rather than understanding autoethnography as only a solo practice (a focus on the “auto,”), these collaborators acknowledge the contribution of the “ethno” (culture/group) to the “auto.” It’s quite common to find autoethnographers blending their stories with those of others, as [Boylorn \(2013\)](#) did in *Sweetwater: Black women and narratives of resilience*. These works answer the call put forth by [Ede and Lunsford \(2001\)](#) to “enact contemporary critiques of the author and of the autonomous individual through a greater interest in and adoption of collaborative writing practices” (pp. 355–356). In some of these examples, authors use autoethnography as part of a multi-method approach to their work. In other cases (such as with Alexander, Moreira, and Kumar; and Toyosaki, Pensoneau-Conway, Wendt, and Leathers), the authors claim that their work still constitutes autoethnography, but in such a way that highlights the intersubjective nature of selfhood.¹

Additionally, autoethnographic voice manifests in a variety of forms. Though much scholarship is found in a more traditional written format, even on the page, autoethnography can look very different from one to the next. [Spry \(2011\)](#) plays with space and place in her poetic autoethnography, as does [Cosenza’s \(2013\)](#) autoethnographic performance of queer futurity. Cosenza’s piece also demonstrates the literal embodiment of autoethnography, in its mediated form. On the surface, this seems contradictory to my earlier gesture toward “graphy” as “writing.” However, embodied/performed autoethnography works against the scriptocentrism that critical scholars, especially in the humanities [e.g. [Conquergood \(2002\)](#)], have argued against. “More often than not, subordinate people experience texts and the bureaucracy of literacy as instruments of control and displacement” ([Conquergood, 2002](#), p. 147). Indeed, Cosenza specifically puts her autoethnography on its feet as a form of resistance to the ways text had been used against her as a dyslexic scholar. These scholars, and others, co-opt the written form, in a sense, to allow access to different forms of autoethnography.

Regardless of the shape of autoethnography, such work asks readers/audience members to sit with questions of voice and authorship. It comes to a question of who has the authority to narrate experience, and thus, narrate culture. Whose story is it? Is it the story of the self? And in a poststructuralist turn, is there a “self” who can speak in the first place ([Nelson, 1985](#))? Rather than pose an answer to these questions, the autoethnographic enterprise plays with and explores these questions as part of its practice.

Evocative autoethnography

Autoethnographic writing is highly evocative and personal. While some may advocate for a more analytical and distanced narrative ([Anderson, 2006](#)), others would argue that these two (evocative writing and analysis) are not mutually exclusive. Indeed, autoethnography “hinges on the push and pull between and among analysis and evocation, personal experience and larger social, cultural, and political concerns” ([Adams and Holman Jones, 2008](#), p. 374). As [Berry \(2006\)](#) explains, through telling an evocative story autoethnographically, a scholar creates a communicative interaction with a reader, an interaction that can speak loudly about (and thus, analyze) a social phenomenon. Further, not only does embodied, lived experience constitute valid data for analyzing cultural phenomena, but for many autoethnographers, the writing process itself is the act of inquiry. That is, rather than being a *representation* of data collection and analysis (as it might be for more conventionally ethnographic approaches), autoethnographers understand the writing/crafting process as illuminating cultural patterns and systems that impact how we see ourselves and others ([Pollock, 1998](#); [Richardson and St. Pierre, 2005](#)). These insights often come within the act of writing, rather than the writing only being the representation of these insights.

¹While it is beyond the scope of this chapter to reference all instantiations of autoethnography and approaches that may not be autoethnography proper, but rather, autoethnography adjacent or may use elements of autoethnography, I appreciate a reviewer’s point that such a broad approach to what may be considered autoethnography risks giving up the specificity of the method.

The what of autoethnography

Autoethnographic narrative seeks to articulate how the body moves through the world, in all of its shifting identifications and manifestations. Through story, the autoethnographer narrates the process of how identity is constituted in and through the body. As [Gannon \(2006\)](#) explains, “The authority for the story begins with the body and memories of the autoethnographic writer at the scene of lived experience” (p. 475). And in autoethnographic scholarship, that experience runs the gamut from the constitution of a gay male body’s becoming ([Berry, 2007](#)); the transformative experience of recognizing and owning White privilege ([Boyd, 2008](#)); the tricky, painful, and necessary experience of gender transition and family dynamics ([Dent, 2002](#)); the grief process ([Ellis, 2009](#)); narrative reframing of abuse ([Kiesinger, 2002](#)); disability and qualitative research ([Polczyk, 2012](#)); and mental health ([Collins, 2017](#); [Iago, 2002](#)).

Specific to the field of education, autoethnographers have taken up a number of topics, ranging from the academic as tourist ([Pelias, 2003](#)); the relational intricacies of a teaching and advising relationship ([Pensoneau-Conway and Valenta, 2021](#)); identity politics in higher education ([Toyosaki, 2007](#)); self-disclosure, sexual orientation, and the classroom ([Gust, 2007](#)); professional life of teacher educators ([Hayler, 2011](#)); disability and educational contexts ([Smith, 2013](#)); and professional development of and barriers faced by leaders of color in academia ([Chang et al., 2014](#)). Still, autoethnography remains underexplored and underutilized in K-12 education in the US.

Outlets for autoethnographic scholarship

Interested scholars have quite a number of outlets for finding and publishing autoethnographic work. These include *Journal of Contemporary Ethnography* (Sage), *Journal of Autoethnography* (University of California Press), *Qualitative Inquiry* (Sage), and *Cultural Studies* ← → *Critical Methodologies* (Sage). *Liminalities: A Journal of Performance Studies* (Creative Commons) regularly publishes mediated autoethnographies. In addition, the International Association for Autoethnography and Narrative Inquiry (IAANI, n.d.) hosts an annual Symposium on Autoethnography and Narrative (formerly the Doing Autoethnography conference). Likewise, the *Critical Autoethnography* (n.d.) group also hosts a regular conference. The International Congress of Qualitative Inquiry is very welcoming toward autoethnographic work, and includes an autoethnography “Special Interest Group.” Finally, books such as *Doing Autoethnography* ([Pensoneau-Conway et al., 2017](#)), *Sage Handbook of Qualitative Methods* ([Denzin and Lincoln, 2017](#); this *Handbook* has been publishing chapters related to autoethnography at least since the second edition in 2000) and *Handbook of Autoethnography* ([Adams et al., 2022](#); [Holman Jones et al., 2013](#)) have all given visibility and legitimacy to autoethnography as research method.

Autoethnographic futures

In more recent years, scholars have begun to ask hard questions of autoethnography, questions that aim to bring into relief the multitude of marginalized voices and experiences. In calling for autoethnography as a praxis of social justice, [Toyosaki and Pensoneau-Conway \(2013\)](#) argue that “while social justice may be fundamentally *social* [and autoethnography may seem to be fundamentally *individual*], our argument for autoethnography as the praxis of social justice entails an examination of the self who engages in social justice/responds to social injustice” (p. 560). Autoethnography is beginning to heed the call for a more critical perspective, one that can respond to the current cultural, global, digital, and political moment. Such works include *Critical Autoethnography: Intersecting Cultural Identities in Everyday Life* ([Boylorn and Orbe, 2021](#)), along with [Atay’s \(2020\)](#) attention to mediated autoethnography and transnational and postcolonial autoethnographies ([Atay, 2018](#)), and [Parks’s \(2021\)](#) intersectional autoethnography of race and disability.

While autoethnography is enjoying relative acceptance in academia, questions remain about the existence of a self, the role of the self in the construction of the other (and vice versa), and the limits of scholarly practice. While no method alone can (or should) claim to be able to respond adequately to any of these questions, autoethnography finds hope in the stories that work toward some semblance of resolution. Embracing the ontology of *homo narrans* may provide a humanizing perspective to a scholarly orientation toward understanding social phenomenon. The generative component to storytelling may just be a key to seeing one another more fully, more humanely, and more justly.

References

- Adams, T.E., 2011. *Narrating the Closet: An Autoethnography of Same-Sex Attraction*. Routledge.
- Adams, T.E., Holman Jones, S., 2008. Autoethnography is queer. In: Denzin, N.K., Lincoln, Y.S., Smith, L.T. (Eds.), *Handbook of Critical and Indigenous Methodologies*. Sage, pp. 373–390.
- Adams, T.E., Ellis, C., Holman Jones, S. (Eds.), 2022. *Handbook of Autoethnography*, second edition. Routledge.
- Alexander, B.K., Moreira, C., Kumar, h.s., 2012. Resisting (resistance) stories: a tri- autoethnographic exploration of father narratives across shades of difference. *Qual. Inq.* 18 (2), 121–133. <https://doi.org/10.1177/1077800411429087>.
- Anderson, L., 2006. Analytic autoethnography. *J. Contemp. Ethnogr.* 35 (4), 373–395. <https://doi.org/10.1177/0891241605280449>.
- Atay, A., 2018. Journey of errors: finding home in academia. *Cultural Studies* ← → *Critical Methodologies* 18 (1), 16–22. <https://doi.org/10.1177/1532708617734009>.

- Atay, A., 2020. What is cyber or digital autoethnography? *Int. Rev. Qual. Res.* 13 (3), 267–279. <https://doi.org/10.1177/1940844720934373>.
- Berry, K., 2006. Implicated audience member seeks understanding: reexamining the “gift” of autoethnography. *Int. J. Qual. Methods* 5 (3), 1–2. Available at: <https://journals.sagepub.com/doi/10.1177/160940690600500309>.
- Berry, K., 2007. Embracing the catastrophe: gay body seeks acceptance. *Qual. Inq.* 13 (2), 259–281. <https://doi.org/10.1177/1077800406294934>.
- Boyd, D., 2008. Autoethnography as a tool for transformative learning about White privilege. *J. Transformative Educ.* 6 (3), 212–225. <https://doi.org/10.1177/1541344608326899>.
- Boylorn, R.M., 2013. *Sweetwater: Black Women and Narratives of Resilience*. Peter Lang.
- Boylorn, R.M., Orbe, M.P. (Eds.), 2021. *Critical Autoethnography: Intersecting Cultural Identities in Everyday Life*, second ed. Routledge.
- Chang, H., Ngunjiri, F.W., Hernandez, K.-A.C., 2013. *Collaborative Autoethnography*. Routledge.
- Chang, H., Longman, K.A., Franco, M.A., 2014. Leadership development through mentoring in higher education: a collaborative autoethnography of leaders of color. *Mentor. Tutoring: Partnersh. Learn.* 22 (4), 373–389. <https://doi.org/10.1080/13611267.2014.945734>.
- Collins, D., 2017. The art of faking a smile. In: Pensoneau-Conway, S.L., Adams, T.E., Bolen, D.M. (Eds.), *Doing Autoethnography*. Sense, pp. 201–210.
- Conquergood, D., 2002. *Performance Studies: Interventions and Radical Research*. TDR, 2nd ed. 46. pp. 145–156, The MIT Press.
- Cosenza, J., 2013. Once upon a time: looking to the ecstatic past for queer futurity. *Liminalities* 9 (2), 28–38. Available at: <http://liminalities.net/9-2/cosenza.html>.
- Critical Autoethnography, (n.d.). Home [Facebook Page]. Facebook. Available at: <https://www.facebook.com/groups/708668335919178>.
- Dent, B., 2002. Border crossings: a story of sexual identity transformation. In: Bochner, A.P., Ellis, C. (Eds.), *Ethnographically Speaking: Autoethnography, Literature, and Aesthetics*. Sage, pp. 191–200.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2017. *The SAGE Handbook of Qualitative Research*, second ed. Sage.
- Duff, M., 2021. A town neglected: living in Harlan, Kentucky during COVID-19. *J. Autoethnog.* 2 (2), 221–225. <https://doi.org/10.1525/joae.2021.2.2.221>.
- Ede, L., Lunsford, A.A., 2001. Collaboration and concepts of authorship. *PMLA* 116 (2), 354–369. Available at: <https://www.jstor.org/stable/463522>.
- Ellis, C., 2004. *The Ethnographic I: A Methodological Novel about Autoethnography*. AltaMira.
- Ellis, C., 2009. Learning to be “with” in personal and collective grief. In: *Revisions: Autoethnographic Reflections on Life and Work*. Left Coast Press, pp. 259–286.
- Fisher, W.R., 1984. Narration as a human communication paradigm: the case of public moral argument. *Commun. Monogr.* 51 (1), 1–22. <https://doi.org/10.1080/03637758409390180>.
- Gale, K., Martin, V., Sakellariadis, A., Speedy, J., Spry, T., 2012. Collaborative writing in real time. *Cult. Stud.* ← → *Crit. Methodol.* 12 (5), 401–407. <https://doi.org/10.1177/1532708612453006>.
- Gannon, S., 2006. The (im)possibilities of writing the self-writing: French poststructuralism theory and autoethnography. *Cult. Stud.* ← → *Crit. Methodol.* 6 (4), 474–495. <https://doi.org/10.1177/1532708605285734>.
- Gingrich-Philbrook, C., 2015. On Dorian Street. In: Chawla, D., Holman Jones, S. (Eds.), *Stories of Home: Place, Identity, Exile*. Lexington Books, pp. 199–214.
- Gust, S.W., 2007. “Look out for the football players and the frat boys”: autoethnographic reflections of a gay teacher in a gay curricular experience. *Educ. Stud.* 41 (1), 43–60. <https://doi.org/10.1080/00131940701308999>.
- Hayano, D.M., 1979. Auto-ethnography: paradigms, problems, and prospects. *Hum. Organ.* 38 (1), 99–104. Available at: <https://www.jstor.org/stable/44125560>.
- Hayler, M., 2011. *Autoethnography, Self-Narrative and Teacher Education*. Springer.
- Holman Jones, S., Adams, T.E., Ellis, C. (Eds.), 2013. *Handbook of Autoethnography*. Left Coast Press.
- IAANI, (n.d.). Home [Facebook Page]. Facebook. Available at: <https://www.facebook.com/groups/doingautoethnography>.
- Jago, B.J., 2002. Chronicling an academic depression. *J. Contemp. Ethnogr.* 31 (6), 729–757. <https://doi.org/10.1177/08912402237823>.
- Kiesinger, C.E., 2002. My father’s shoes: the therapeutic value of narrative reframing. In: Bochner, A.P., Ellis, C. (Eds.), *Ethnographically Speaking: Autoethnography, Literature, and Aesthetics*. Left Coast Press, pp. 95–114.
- LeMaster, B., 2013. Telling multiracial tales: an autoethnography of coming out home. *Qual. Inq.* 20 (1), 51–60. <https://doi.org/10.1177/107780041358532>.
- Lionnet, F., 1989. Autoethnography: an an-archic style of Dust Tracks on a road. In: *Autobiographical Voices: Race, Gender, and Self-Portraiture*. Cornell University Press, pp. 97–129.
- Nelson, C., 1985. Poststructuralism and communication. *J. Commun. Inq.* 9 (2), 2–15. <https://doi.org/10.1177/019685998500900200>.
- Parks, E.S., 2021. The hybrid I/eye: a critical autoethnography of intersecting race and disability identities. *J. Autoethnog.* 2 (1), 26–38. <https://doi.org/10.1525/joae.2021.2.1.26>.
- Pelias, R., 2000. The critical life. *Commun. Educ.* 49 (3), 220–228. <https://doi.org/10.1080/03364520009379210>.
- Pelias, R., 2003. The academic tourist: an autoethnography. *Qual. Inq.* 9 (3), 369–373. <https://doi.org/10.1177/1077800403251759>.
- Pensoneau-Conway, S.L., Toyosaki, S., 2011. Automethodology: tracing a home for praxis-oriented ethnography. *Int. J. Qual. Methods* 10 (4), 378–399. Available at: <http://ejournals.library.ualberta.ca/index.php/IJQM/article/view/10581/9361>.
- Pensoneau-Conway, S.L., Valenta, D.J., 2021. Embracing autoethnographic anxiety: the joyous potential of teaching and advising relationships. In: Adams, T.E., Holman Jones, S., Ellis, C. (Eds.), *Handbook of Autoethnography*, second ed. Routledge, pp. 229–240.
- Pensoneau-Conway, S.L., Adams, T.E., Bolen, D.M. (Eds.), 2017. *Doing Autoethnography*. Sense.
- Polczyk, P., 2012. Autoethnography as an accessible method of research. *Forum Oświatowe/Educ. Forum* 24 (2), 175–182. Available at: <https://forumswiatowe.pl/index.php/czasopismo/article/view/20>.
- Pollock, D., 1998. Performing writing. In: Phelan, P., Lane, J. (Eds.), *The Ends of Performance*. New York University Press, pp. 73–103.
- Rambo Ronai, C., 1995. Multiple reflections of child sex abuse: an argument for a layered account. *J. Contemp. Ethnogr.* 23 (4), 395–426. <https://doi.org/10.1077/089124195023004001>.
- Rambo Ronai, C., 1996. My mother is mentally retarded. In: Ellis, C., Bochner, A.P. (Eds.), *Composing Ethnography: Alternative Forms of Qualitative Writing*. AltaMira, pp. 109–131.
- Richardson, L., St Pierre, E.A., 2005. Writing: a method of inquiry. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *SAGE Handbook of Qualitative Methods*, third ed. Sage, p. 595.
- Smith, P., 2013. *Both Sides of the Table: Autoethnographies of Educators Learning and Teaching With/in [dis]ability*. Peter Lang.
- Spry, T., 2011. *Body, Paper, Stage: Writing and Performing Autoethnography*. Routledge.
- Toyosaki, S., 2007. Communication sensei’s storytelling: projecting identity into critical pedagogy. *Cult. Stud.* ← → *Crit. Methodol.* 7 (1), 48–73. <https://doi.org/10.1177/1532708606288643>.
- Toyosaki, S., Pensoneau-Conway, S.L., 2013. Autoethnography as a praxis of social justice: three ontological contexts. In: Holman Jones, S., Adams, T.E., Ellis, C. (Eds.), *Handbook of Autoethnography*. Left Coast Press, pp. 557–575.
- Toyosaki, S., Pensoneau-Conway, S.L., Wendt, N.A., Leathers, K., 2009. Community autoethnography: compiling the personal and resituating whiteness. *Cult. Stud.* ← → *Crit. Methodol.* 9 (1), 56–83. <https://doi.org/10.1177/1532708608321498>.

Grounded theory

Wayne A. Babchuk^a and Eileen Boswell^b, ^a Department of Educational Psychology, University of Nebraska-Lincoln, Lincoln, NE, United States; and ^b Department of Teaching, Learning, and Teacher Education, University of Nebraska-Lincoln, Lincoln, NE, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	108
Overview of article	108
Origins and history of grounded theory	109
Diverging interpretations	109
Grounded theory's contested family of approaches	110
Conducting grounded theory research in education	111
Coding	113
Theoretical sampling and theory development	113
Number of interviews	113
Tracking the developing theory	113
Grounded theory purpose statements and research questions	114
From purpose to process: grounded theory methods in practice	114
Evaluating a grounded theory	115
Examples of grounded theory studies in education	116
Future directions	116
Ethnographic grounded theory	116
Grounded theory and mixed methods designs	117
Mixed methods-grounded theory exemplars	118
Ontological and epistemological evolution	118
Critical theoretical applications	118
Critical race theory and grounded theory methodology	119
Conclusion	119
Acknowledgments	119
References	119
Recommended grounded theory texts	122

Glossary

Abduction A form of reasoning where the researcher considers various explanations for observed data and then tests these explanations against the data to ultimately yield the most plausible explanation of the observed phenomenon (See Babchuk, 1997; Bryant and Charmaz, 2007a,b; Guetterman et al., 2019; Guetterman et al., in preparation)

Axial coding Strauss' (1987) and Strauss and Corbin's (1990) term for the delineation of hypothetical relationships between a category and its subcategories, and between the category and other categories through use of the coding paradigm (i.e., conditions, context, action/interactional strategies, and consequences)

Category A set of codes that share common properties. Categorizing is the process of abstracting these codes into an analytic concept or pattern

Coding The data analysis process of applying descriptive labels to data or data sets for analytic purposes to organize data at increasingly higher levels moving from codes to categories, concepts, or themes to theory. Grounded theorists advocate different forms of coding from open, initial, axial, and substantive coding to increasingly selective, focused, or theoretical techniques used to develop the emerging theory

Conditional/consequential matrix A heuristic tool diagramming the relationships between macro and micro conditions and consequences

Constant comparative/-ison An analytical ongoing and iterative process central to the grounded theory methodology widely incorporated by other qualitative approaches in which the researcher continually compares emerging data in increasingly abstract levels to develop concepts, categories, and theory

Ethnographic grounded theory A hybrid qualitative research approach that systematically combines elements of ethnography and grounded theory

Grounded theory An overarching term denoting a specific research design geared at developing a midrange theory, explanation, typology or conceptual framework grounded in the data as well as the end product of this form of research

Maximum variation First identified as a concept by Glaser and Strauss (1967), this is the principle of purposefully sampling data from people or cases exhibiting widely varying circumstances in order to cast as wide a net as possible to study the general characteristics of a phenomenon before narrowing, via **theoretical sampling**, to characteristics befitting the model in development

Memoing or memo-writing A key part of the grounded theory process leading toward theory-generation; a memo can be reflective, theoretical, procedural, interpretive, or linked to a code or category; memo-writing happens in iterative cycles together with data collection and coding

Middle-range theory (c.f., grand or “formal” theories) A primary aim in most grounded theory studies capturing “abstract renderings of specific social phenomena” (Charmaz, 2014, p. 9); contrasts with contrast to so-called “grand” theories such as Marxism or feminist theory; Charmaz (2014) adds other grounded theory aims such as substantive theorizing, extending extant theory with new concepts, and challenging extant theory (pp. 248–259)

Open coding The initial coding process in grounded theory which involves the breaking down, analysis, comparison, and categorization of data (Strauss and Corbin, 1990, p. 61). In constructivist applications this stage is referred to as *initial* coding (Charmaz, 2014) but is much the same process

Selective coding The process by which categories are related to the core category, and these relationships are validated against the data. In systematic or Straussian grounded theory, this becomes the basis for the grounded theory

Theory A substantive or formal explanation or model of the social psychological phenomenon being studied that emerges from data analysis throughout the research process

Theoretical sampling A unique form of purposive sampling that involves the ongoing selection of sites, participants, and events that are strategically selected to assist the development of theoretical categories and concepts to enrich the emerging theory

Theoretical saturation (sufficiency) A condition of a grounded theory study when newly collected data do not yield additional insights to the further development of the relationships between categories and development of the emerging theory signaling the end of the data collection process

Theoretical sensitivity The ability of the researcher, through the specific grounded theory analytic procedures chosen, to recognize the theoretical direction(s) and relevant data and reflect upon it with the help of theoretical terms or concepts

Introduction

Grounded theory is an inductive—or perhaps more accurately—an abductive method aimed at generating theory from empirical data collected in the field. Often viewed as both a process and product of social research, grounded theory has been argued to be the most utilized contemporary qualitative research methodology (Babchuk, 1997; Bryant and Charmaz, 2007a; Guetterman et al., 2019) but is not limited solely to the collection and analysis of qualitative data having been employed in both quantitative and mixed methods research (Glaser, 2008; Glaser and Strauss, 1967; Holton and Walsh, 2017; Johnson et al., 2010; Johnson and Walsh, 2019; Walsh, 2014, 2015). Due in part to its positivist/post-positivist epistemological origins and the systematic and rigorous foundational aspects advanced by its co-founders, it can be considered a particularly appealing “transition methodology” for quantitatively trained scholars interested in adopting qualitative methodologies for their research or incorporating it in mixed methods designs (Guetterman et al., 2019; Howell Smith et al., 2020; Johnson and Walsh, 2019; Johnson et al., 2010). Although its applications in real world settings can prove challenging to both novice and more experienced researchers faced with a wide array of conflicting interpretations as to how best to utilize this methodology, taken together the various iterations of this methodology have proven to be particularly successful over time and across disciplines for advancing social research (Charmaz, 2014; Charmaz and Bryant, 2010).

Overview of article

This article begins with an overview of the origins and history of grounded theory followed by a discussion of grounded theory’s extended family of approaches that have been widely adopted by contemporary researchers. We then focus on conducting grounded theory research in education by considering shared core attributes of this methodology, procedures for conducting grounded theory research, criteria for evaluating grounded theory studies, and suggestions for getting started in grounded theory research. We conclude this article through a discussion of future directions where we consider other potentially fruitful applications of grounded theory including **ethnographic grounded theory**,¹ mixed methods-grounded theory, and designs that leverage grounded theory in

¹Bold-face terms are defined in the Glossary section.

service of critical race theory and social justice aims. It is our intent to provide a comprehensive yet succinct overview of grounded theory in order to introduce readers to this popular approach and to point those new to this methodology in the right direction as they consider taking on this form of inquiry.

Origins and history of grounded theory

In 1960, the University of California-San Francisco (UCSF) School of Nursing hired Anselm Strauss to teach medical sociology and serve as inaugural Director of the School's newly minted Department of Social and Behavioral Sciences (Birks and Mills, 2015). He helped establish the first graduate program in nursing in the US to guide nurses in their Ph.D. research, a position he held until his retirement in 1987 (Babchuk, 2011). Strauss was a byproduct of the iconic Chicago School of sociology where his training was strongly permeated by an emphasis on naturalistic (ethnographic) fieldwork, symbolic interactionism, and pragmatism via the influences of Pierce, Dewey, Mead, Park, and Blumer (Bryant, 2017; Hadley, 2017; Olson, 2018). To help in these tasks and in conducting research as part of a 4-year grant on death and dying, Strauss enlisted the aid of Barney Glaser, who was fresh off his doctoral work at the prestigious Department of Sociology at Columbia University, where scholars such as Paul Lazarsfeld and Robert Merton were advancing quantitative research methodologies and pioneering "theories of the middle range" (Babchuk, 2011, emphasis added).

Accompanied by Jeanne Quint Benoliel, then a doctoral student in nursing—who seems to have played an important yet largely uncredited role in the methodology's development, analogous to that of Rosalind Franklin in the discovery of DNA—Glaser and Strauss conducted collaborative research on death and dying in hospitals in order to forge a new approach to social research they labeled grounded theory. This methodology was explicitly designed to systematically generate theory from empirical data collected in the field (Babchuk, 2011; Bryant, 2017, 2019). Formally introduced in *The Discovery of Grounded Theory: Strategies for Qualitative Research* (1967), the methodology [that] Glaser and Strauss sought to develop had to be "useful for both verification and generation of theory" through employing quantitative and qualitative data with a particular emphasis on qualitative research "because the crucial elements of sociological theory are often found best within a qualitative method" (p. 18). In *The Discovery of Grounded Theory*, they sought to advance a systematic qualitative approach equivalent to the predominant quantitative methodologies of the era outlining fundamental principles of grounded theory including the constant comparative method, **theoretical sampling** and **saturation**, simultaneous data collection and analysis, and **memo writing** (Glaser and Strauss, 1967; Guetterman et al., 2019; and see Glaser, 1965).

Unfortunately, *The Discovery of Grounded Theory* (1967), together with their two other prominent early texts, *Awareness of Dying* (1965) and *Time for Dying* (1968), provided little guidance on how to employ these fundamental principles to successfully conduct grounded theory research (Bryant, 2017). Unbeknownst to Glaser and Strauss, or anyone at this time, underlying this collaboration and the publication of the "death and dying trilogy" were fundamental differences in how these two scholars envisioned the use of this methodology to advance social research. They taught and worked together for almost 10 years before Glaser—after being denied tenure at UCSF—left to conduct his own research and establish his own publishing company he titled Sociology Press in Mill Valley, California in 1970.

Diverging interpretations

Glaser and Strauss parted ways. Their grounded theory "divorce" was of a magnitude rarely seen in academe, and one which has been well documented in the grounded theory literature as well as being the subject of the first author's doctoral dissertation research (Babchuk, 1996, 1997, 2008, 2011; Birks and Mills, 2015; Guetterman et al., 2019; Heath and Cowley, 2004; Kelle, 2007). As will be discussed in more detail in the next section, both Glaser and Strauss went on to promote their own versions of grounded theory that shared elements of the original design but differed substantially on key attributes particularly relating to the challenging process of data analysis. Others would ultimately offer their own epistemological and operationally repositioned versions to constitute grounded theory's "family of methods" a term first coined by Anthony Bryant (Babchuk, 2011; Bryant, 2017, 2019; Bryant and Charmaz, 2007a), "family of variants" (Bryant, 2019), "compendium of methods" (Corbin, 2009), "constellation of methods" (Charmaz, 2014), or "package of methods" (Birks and Mills, 2015). Although several scholars advocate for three major grounded theory approaches including (1) Classic, Emergent or Traditional (Glaserian) Grounded Theory, (2) Systematic (Straussian) Grounded Theory, and (3) Constructivist (Charmazian) Grounded Theory (Bryant, 2017, 2019; Creswell and Poth, 2018) we include Adele Clarke's (4) Postmodern/Situational as a fourth variant (Babchuk, 2011; and see Reichertz, 2007, who discussed five variants). Other sub-approaches of this methodology that deserve consideration are Dimensional Analysis (Bowers and Schatzman, 2009) and Hadley's (2017, 2019) Critical Grounded Theory. Although each approach advances its own unique recommendations for conducting grounded theory research, most incorporate several core features of grounded theory as originally envisioned by the co-founders (Charmaz, 2014; Guetterman et al., 2019).

As stated in the introduction to this article, grounded theory is often considered the mostly widely utilized contemporary qualitative research methodology, although systematic reviews of the grounded theory literature (Babchuk, 1997; Charmaz, 2014; Guetterman et al., 2019) have indicated that many studies purporting to be grounded theory analyses should perhaps not be accorded the grounded theory label as will be discussed in more detail later in this article. Regardless, the historical significance of grounded theory methodology cannot be understated, as Merriam and Tisdell (2016) have argued that *The Discovery of Grounded Theory* (1967)

represents one of the two most influential works in the history of the qualitative tradition (Babchuk, 2011), and Charmaz (2000) proclaimed that this work “served at the front of the qualitative revolution” (p. 509) that began in the 1960s. Similarly, Denzin and Lincoln’s (2005) “Second Movement of Qualitative Research” or the “Modernist Phase”—also referred to by them as the “golden age of rigorous qualitative research”—was “bracketed in sociology by *Boys on White* (Becker et al., 1961) at one end and the *Discovery of Grounded Theory* (Glaser and Strauss, 1967) at the other” (p. 17, and see Babchuk, 2011; Birks and Mills, 2015).

Grounded theory’s contested family of approaches

The early collaborative works of Glaser and Strauss (1965, 1967, 1968, 1971) provided little guidance to the aspiring grounded theorist on how to design and conduct this form of research. To make explicit grounded theory’s fundamental procedures outlined in the *Discovery of Grounded Theory* (1967) and to better train doctoral students in the use of this methodology, Anselm Strauss alone (1987) and with former UCSF doctoral student Juliet Corbin (1990, 1994, 1998) advanced their own highly structured and systematic guidelines. Influenced by Glaser’s classic text *Theoretical Sensitivity* (1978) and his notions of theoretical coding and coding families, Strauss—and later Strauss and Corbin—put more emphasis on theory verification rather than theory generation and they employed open, axial, and selective coding procedures (Charmaz, 2014; Kelle, 2019; Strauss and Corbin, 1990). Central to this approach is the axial coding process that is a procedure that further extends the initial open coding phase by systematically exploring the relationships between categories and sub-categories through the use of a coding paradigm specifying causal conditions, contextual and intervening conditions, and consequences (Babchuk, 2008; Creswell and Guetterman, 2019; Strauss and Corbin, 1990). This process sets the stage for selective coding and the identification of a core category of the studied phenomenon integrating the categories and leading to the generation of theory. Introduced in 1990 and later elaborated in the second edition (1998) of their text, Strauss and Corbin also introduced their conditional (later labeled conditional/consequential) matrix, a heuristic tool diagramming the relationships between macro and micro conditions and consequences (Strauss and Corbin, 1990, 1998). An example of Systematic or Straussian Grounded Theory is Lovell’s (2016) study geared at explaining parental involvement in childhood health and nutrition of low-income Early Head Start families.

Strauss and Corbin’s texts were generally well received as they represented the first attempts to provide systematic guidelines as to how to conduct grounded theory analysis, and they were further endorsed by Creswell in multiple editions of his widely read *Qualitative Research Inquiry and Research Design: Choosing among Five Approaches* texts (see Creswell and Poth, 2018). Firmly believing that Strauss and Corbin’s (1990) interpretation of the grounded theory methodology deviated too greatly from the tenets laid out in *The Discovery of Grounded Theory* (1967), Glaser labeled their approach “full conceptual description” and in response, after Strauss and Corbin ignored his requests for them to pull the text, he published his own corrective text through his publishing house Sociology Press—*Basics of Grounded Theory Analysis* (1992)—he claimed could be used to set “the average researcher back on the correct track to generating a grounded theory” (p. 6). Glaser forcefully argued that Strauss and Corbin’s coding procedures and paradigm were too prescriptive and structured, and that these represented a type of approach he and Strauss had originally attempted to distance themselves from in their 1960s work recommending procedures that were forced rather than emergent. Systematic Grounded Theory has historically been adopted for use by graduate students and less experienced researchers who can use the relatively straightforward set of procedures as outlined by Strauss and Corbin, but this formulaic approach may be potentially too constrictive for some and seems to have declined in popularity in recent years.

In contrast, Glaser’s advanced a version of grounded theory he believed to represent a seamless development from *The Discovery of Grounded Theory*, extending to *Theoretical Sensitivity* (1978), to a series of his more contemporary works published through Sociology Press. In *Theoretical Sensitivity*, Glaser outlines his theoretical coding model which helps the analyst to further structure emerging categories by drawing upon 18 theoretical coding families (further expanded and modified in 1998). Theoretical codes act to conceptualize how the initial substantive codes developed during open coding relate to each other as hypotheses to be integrated in the emerging theory (Charmaz, 2014; Glaser, 1978; Kelle, 2007). Theoretical sensitivity—the ability for the analyst to recognize and reflect upon relevant empirical data during analysis—is key to successfully utilizing the methodology. In his version, Glaser places more emphasis on theory generation than theory verification through constant comparison, open, selective, and theoretical coding, and much inductive thinking (Hadley, 2017). Perceived by some as too dense and formidable for many researchers to tackle, Glaserian Grounded Theory has gained considerable prevalence in recent years thanks in part to the influence of contemporary scholars such as Martin and Gynnild (2011) and Holton and Walsh (2017) who demonstrated the potential of this approach. An example of Classic or Glaserian Grounded Theory is Giske and Artinian’s (2007, 2008) study of coping by patients in a gastroenterology ward (and see Bryant, 2019, for further discussion of this study).

Glaser and Strauss are considered the first generation of grounded theorists and the founders of this methodology. However, a second generation represented primarily by some of their former students at UCSF including Juliet Corbin, Kathy Charmaz, Adele Clarke, Phyllis Noerager Stern, and others (Stern and Porr, 2011) have reshaped the grounded theory landscape. The work of Charmaz (2000, 2006, 2014) and Clarke (2005, 2009, 2019) in particular, has been seen by many as more aligned with epistemological, theoretical, and methodological advancements taking place in the social and behavioral sciences over the past few decades (Babchuk, 2008; Bryant, 2002, 2017, 2019; Bryant and Charmaz, 2007b; Charmaz, 2014).

Rejecting objectivist, positivist/post-positivist assumptions she believed were at the epistemological core of both the Glaserian and Straussian approaches, Charmaz developed her own constructivist brand of grounded theory (2000, 2006, 2007, 2014) and was later joined by Antony Bryant who helped champion and extend her approach (Bryant, 2017, 2019; Bryant and Charmaz,

2007a,b, 2019). This approach sidesteps the pure inductivism of Glaserian Grounded Theory and the rigid formulaic procedural guidelines of Straussian Grounded Theory. Instead, it advances a more democratic approach aimed at co-constructing rather than discovering or verifying grounded theories through shared interactions between researchers and multiple social actors (Charmaz, 2014). Charmaz also rejects Glaser's viewpoint that grounded theory should evoke one core category or process. She embraces the notions of multiple ways of knowing (following Strauss), a focus on basic social processes, and rejects a heavy reliance on dogmatic methodological adherence to produce theory (Charmaz, 2014; Hadley, 2017). She also incorporates open coding techniques she refers to as *initial* coding, followed by a secondary focused coding stage that collapses codes into similar categories, and then theoretical coding to further develop the emerging theory. Charmaz is a strong advocate of using gerunds (typically words ending in “-ing” in English that are verbs acting as nouns) in coding (as recommended by Glaser (1978)) and relies upon the core processes of constant comparison, theoretical sampling, and memoing in conducting grounded theory research. In her version of grounded theory, the researcher acknowledges her or his key role in the research itself and at the same time elevates the participant to more equal standing in the co-construction of mutual derived explanations of the studied social phenomenon. Unlike previous approaches, Charmaz maintained that the data do not speak for themselves and instead rest on the interactions of researchers and participants (Olson, 2018). An example of Constructivist or Charmazian Grounded Theory is Lynch-O'Brien et al. (2021) who developed a substantive theory of “transference” to explain “the social process by which educational and attitudinal impacts intended by program leaders for program participants are filtered by citizen scientists and transferred to others” (p. 1) in a citizen science program.

Also rejecting the objectivist approaches of grounded theory's co-founders, Adele Clarke promotes her own well-conceptualized version of postmodern/situational grounded theory she labeled Situational Analysis. Drawing from Chicago School symbolic interactionism, the Straussian notion of social world/arenas theory, the post-structural theorizing of Michel Foucault and other post-modern scholars, and assumptions driving constructivist grounded theory, Clarke's Situational Analysis (Clarke, 2005, 2019; Clarke and Friese, 2007; Clarke et al., 2016) employs the use of cartographic mapping techniques and the making of situational, world/arenas, and positional maps to best illuminate social situations of power and oppression viewed holistically from multiple overlapping lenses. Clarke also rejects oversimplified Glaserian and Straussian notions of a single social process, “focusing instead on analyzing relationalities both within and across three kinds of maps” (Clarke, 2019, p. 16). In other words, Clarke is more focused on the social situation than on the analysis of basic social processes. An instance of this approach comes from the work of Pérez and Cannella (2013) who “provide examples of situational mapping from a study focusing on disaster capitalism and the privatization of the public education system in post-Katrina New Orleans” (p. 505).

In addition to varying epistemological, theoretical, and methodological interpretations relating to the design and conduct of grounded theory research outlined above, there have been numerous other controversies underpinning this approach over the past fifty years. These include the nature of the research process—inductive, deductive, and abductive—the timing of how relevant literature should be incorporated in grounded theory research, conflicting interpretations of meaning and applications of the terms “data” and “theory,” use of theoretical sampling and sample size, theoretical saturation, positionality of the researcher, what constitutes a grounded theory study and what the findings should look like (e.g., theory, typology, model, etc.), and evaluative criteria on how to assess the integrity and contribution of a grounded theory study (Babchuk, 1997, 2008; Bryant, 2002, 2017; Bryant and Charmaz, 2007b; Charmaz, 2014; Clarke, 2005; Goulding, 2002; Hood, 2007; Reichertz, 2007). All of these dimensions are best taken into account when undertaking a grounded theory study or one's dissertation research.

Conducting grounded theory research in education

A number of grounded theory scholars have identified “core” (Bryant, 2019, pp. 49–50; Hadley, 2017, p. 59), “essential” (Birks and Mills, 2015, p. 9), “hallmark” (Guetterman et al., 2019, p. 182), “key” (Creswell and Guetterman, 2019, p. 442), and “defining” (Charmaz, 2014, p. 7; Creswell and Poth, 2018, pp. 83–84) features shared by most if not all of the grounded theory variants discussed above that together distinguish grounded theory from other qualitative methodologies. Table 1 below lists the key characteristics of grounded theory methodology alongside tips for practical application of each. (See also Glossary section above.)

The key characteristics include simultaneous and iterative data collection and analysis; the **constant comparison** method involving the making of comparisons throughout analysis at increasingly more abstract levels to develop the relationships between categories, concepts, and theory; theoretical sampling as a form of purposive sampling involving the selection of sites, participants, activities, etc., to assist in the development of theoretical categories and concepts aimed at **theory** construction; and memo-writing whereby the researcher engages in a reflective process on the data to help guide the ongoing analysis. Other important grounded theory features include delaying the literature review until analysis is already underway; theoretical saturation that signals the end of data collection when no new properties or theoretical insights are identified that contribute to the development of theoretical constructs; **theoretical sensitivity** referring to the researcher's ability to recognize and reflect upon relevant empirical data during analysis; and the end or byproduct of the research represented by the development of a substantive or formal theory, explanation, or model to explain the studied phenomenon. As is the case with all qualitative approaches, the grounded theory coding process is at the heart of methodological operations: coding is the data analysis process of applying descriptive labels to data or data sets for analytic purposes in order to organize data at increasingly higher levels of abstraction moving from codes, to categories, to concepts, to theory.

Table 1 Key characteristics of grounded theory research design.

<i>Characteristic</i>	<i>Description and importance</i>	<i>Tips for application</i>
Constant comparison	A time-tested analytical strategy adapted by grounded theorists that serves as a foundational pillar of the method in which emerging data are continually compared to each other at increasingly more abstract levels to further develop categories, concepts, and theory	Denzin and Lincoln (2018) advocate an idealized process whereby the researcher compares “data with data, data with code, code with code, code with category, and category with category” along the way to saturation. Gibbs (2018) offers specific memo prompts, tests, and exercises such as “waving the red flag” and “far out comparisons” to identify typical and atypical cases (pp. 67–70), and Boswell (in preparation) offers guidance on leveraging coding software for constant comparison in intersectional analyses.
Category	A type or grouping of similar or closely related codes or concepts generated through inductive data analysis. Categorizing is the process of organizing or abstracting these codes into an analytic concept or pattern	Birks and Mills (2015) present the historical usage of this term by grounded theorists (pp. 86–87) and point out that concepts, codes, and categories are often interchangeable terms in the grounded theory literature.
Coding (open, initial, theoretical, axial, selective, and focused)	The data analysis process of assigning labels and organizing data at successively higher analytical levels in grounded theory.	Charmaz (2006, 2014) and Corbin and Strauss (2014) provide examples specifically for grounded theory research that Saldaña (2021) also elaborates in his extensive overview of coding in qualitative research.
Iterative data collection and analysis	The imbricated process of collecting and analyzing data concurrently in order to leverage the analysis for more and more targeted and theoretical sampling of relevant, information-rich cases to support or refine developing theoretical categories	During ongoing transcription and coding of, for example, interview data, the researcher narrows the interview focus and continues to collect and analyze data as the theoretical model is developing.
Memoing	A well-established research strategy in the social sciences adopted by grounded theorists to refer to the researcher’s ongoing process of reflection throughout the study to help guide the emerging analysis	Types of memos include process memos, code memos, coding memos, pre-theoretical memos, lit memos, decision memos, and reflexive memos. Gibbs (2018) includes several provocative memo prompts (see especially Chapter 4) and Charmaz (2014) lists questions to ask of the data and other memo prompts (pp. 171–190).
Theoretical sampling	An often ignored but foundational form of purposeful sampling used in grounded theory referring to the ongoing selection or revisiting of participants, sites, and events directly aimed at the analytical development of theoretical categories and concepts	Researchers are encouraged to track which questions are added to an interview protocol as a direct result of previous data collection and the developing theoretical direction of the study. After each interview, researchers are also encouraged to note which of the additional questions were asked/answered in each interview in order to chart a visual representation of how data were sampled throughout the study. Draucker et al. (2007) offer an alternative visualization whereby the data objects used at each stage of sampling are tracked as input funnels (p. 1141).
Theoretical saturation	The stage of analysis that signals the end of data gathering in grounded theory analysis when no new properties are emerging that contribute to the further elaboration of theoretical constructs.	The visual representations indicated above will help the researcher “zoom in” and “zoom out” to visually see which categories are saturating.
Theoretical sensitivity	The ability and aptitude of the researcher to detect or identify multiple and variegated theoretical directions during iterative stages of data collection and analysis.	Striking a delicate balance between contemplating extant theory while also not being bound by it will enhance a researcher’s sensitivity to possible theoretical models; vast general knowledge across disciplines, paradigms, and philosophical/historical viewpoints (without strict commitments to any one ontology/epistemology) can help a researcher remain open to theoretical positions and dispositions that help to ground a theory.
Theory	Adapted by grounded theorists to represent a substantive or formal explanation or model of the central phenomenon being studied that emerges from data analysis throughout the research process	Saldaña (2021) provides helpful guidance on how second cycle codes can be joined by verbs that then help to flesh out the theoretical model, based on the stated relationships between categories (pp. 347–359) such as <i>X contributes to Y</i> , or <i>A varies when B</i> .

Adapted from [Guetterman et al. \(2019\)](#).

Coding

Grounded theorists advocate different forms of coding from open, initial, axial, and substantive coding to increasingly selective, focused, or theoretical techniques used to develop the emerging theory. Open coding refers to the first phase of coding procedures where the analyst assigns codes/descriptors to text passages. Charmaz (2014) uses the term initial coding for this first phase of analysis and advocates using gerunds and line-by-line coding of text passages. Axial coding denotes Strauss and Corbin's (1990, 1998) second phase or type of coding where the researcher views a category as an axis and delineates relationships around it using a coding paradigm. Selective coding is the process of choosing a central or core category and then systematically relating it to other categories (Strauss and Corbin, 1990, 1998). Focused coding is the process of identifying the most frequent or important codes to further test against the data to arrive at analytical or focused codes that serve as the basis for theory development (Charmaz, 2014). Theoretical coding is the process by which substantive theoretical codes are related to the core category that emerges through data analysis (Glaser, 1978, 1992). As we have discussed, these coding procedures differ in the various grounded theory sub-approaches. Additionally, coding schema are almost as numerous and diverse as the data sets they are meant to analyze: Saldaña (2021) outlines the rationale for, and applications of, more than 25 coding methods. This is generally an intuitive process whereby potential grounded theorists experiment with different coding techniques as they learn to master strategies that are most effective to accomplish their research goals (see Bryant, 2017, for a detailed discussion of coding and how doctoral advisees approached this process).

A number of scholars have outlined the basic procedures or strategies to conduct grounded theory analysis, each of them incorporating the core features outlined above (Birks and Mills, 2015; Bryant, 2017; Charmaz, 2014; Creswell and Guetterman, 2019; Creswell and Poth, 2018; Hadley, 2017; Strauss and Corbin, 1990, 1998) and in many ways adhere to the general structure of conducting qualitative research (see Babchuk, 2019). The researcher must first address if grounded theory is an appropriate (or the most appropriate) design to address the research problem. Grounded theory has historically been the methodology of choice when the researcher is interested in developing a theoretical framework or explanation of a basic social or social-psychological process, often when there is little known about the central phenomenon of study. In consideration of the researcher's own positionality, their epistemology, and their theoretical framework and its potential fit with the research question, the researcher may benefit by selecting one of the sub-approaches to grounded theory outlined in this article and follow the suggestions associated with it in the design and conduct of the research.

Theoretical sampling and theory development

One core feature of grounded theory that helps distinguish it from other qualitative approaches is the process of theoretical sampling (discussed above) guiding the strategic selection of sites, individuals, and events to further the development of theoretical categories and concepts to enrich theory building. Draucker et al. (2007) provide comprehensive details on both the principles of theoretical sampling as well as examples from their own studies. An additional feature attributed to grounded theory research that has been widely incorporated across qualitative designs is **maximum variation** sampling. This procedure involves the selection of a range of participants, groups, sites, or activities that differ on selected dimensions that are germane to the research question (e.g., age, location, past experience), and it enables the researcher to develop a more complete picture and elucidate the boundaries of the studied phenomenon.

Number of interviews

A burning question in grounded theory methodology, rooted in quantitative traditions and the need for a study to be adequately *powered* is the issue of how much data, or how many instances or information-rich cases are sufficient in order for theoretical saturation to obtain. Guest et al. (2006) indicate that hindsight tends to reveal that saturation was reached after the 12th interview, with no new categories added after the sixth. However, Charmaz (2014) took issue with Guest et al.'s approach, arguing that they were aiming to saturate data, not categories (p. 106). The saturation of categories rather than data constitutes an important distinction between grounded theory and other qualitative approaches as the researcher strives to saturate categories in the development of theory versus saturation of data per se. Creswell and Poth (2018) recommend 20–30 participants for a grounded theory study whereas some suggest as many as 50 (Birks and Mills, 2015). In Guetterman (2015) oft-cited analysis of sampling strategies employed in contemporary qualitative research studies, he found that the mean sample size in his article pool was 59 for those researchers who employed grounded theory (p. 12).

Beyond mere numbers, Malterud et al. (2016) advance a concept of “information power” (p. 1753) which determines saturation thresholds for saturations based on a combination of the study's particular aims, sample heterogeneity, the richness of interview data and the specific analytic strategies employed (p. 1756). There is thus no “one size fits all” prescription for achieving saturation through sample size. Sample size is a matter of some debate in grounded theory, but the researcher should aspire to achieve theoretical saturation to signal the end of the data collection process.

Tracking the developing theory

In order to keep a running record of the theoretical sampling process, and the progress of the developing theoretical directions, researchers may be best served to devise a convenient tool for themselves, such as the sample theory tracker in Fig. 1 below.

Fig. 1 shows a visual depiction of how developing theories can be tracked. Grounded theorists are encouraged to use some version of this tool as part of their research audit trail, as well as to have a record of how theoretical sampling decisions were

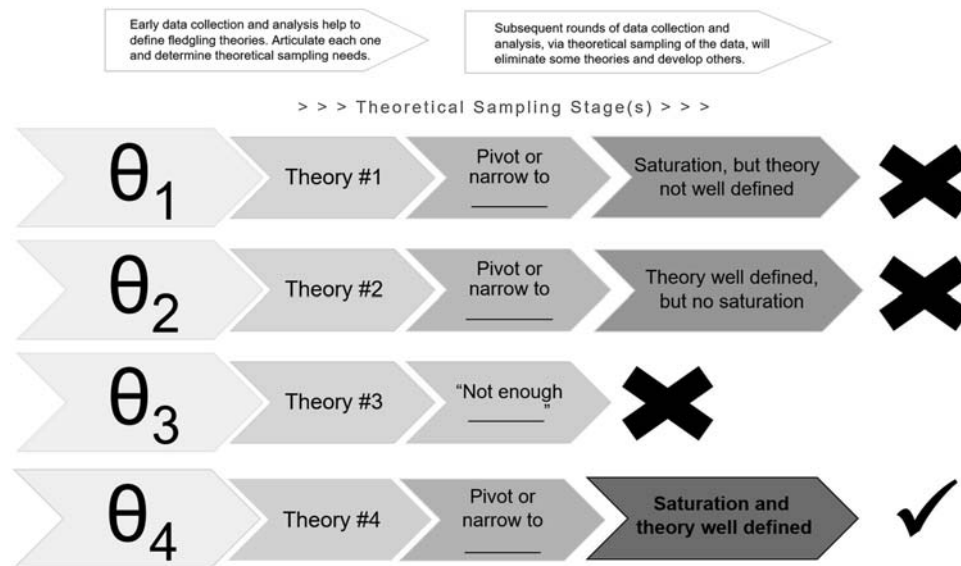


Fig. 1 Sample theory tracker.

made and whether they led to further theory development. In this image, Theory #3 (labeled as Θ_3 in the diagram) is abandoned early in the process because the researcher does not have access to enough information-rich cases from which to theoretically sample. Theories #1 and #2 are further fleshed out through the theoretical sampling process, but the former does not lead to a well-defined theory (despite saturation) and the latter does not saturate. Theory #4 is adopted in this example because it is both well-defined and saturation obtains.

Grounded theory purpose statements and research questions

At the heart of the research design is the purpose statement and central research question and sub-questions. In writing a grounded theory purpose statement it is advised that the researchers state that they are using a grounded theory approach to study a particular phenomenon, process, or action and, if possible, also indicate the participants and site(s) of the research. A grounded theory central question is a broad, overarching question that restates the purpose statement and should indicate that the research is designed to yield a theoretical explanation of the social psychological process, event, or action (see Creswell and Poth, 2018, p. 132 for a grounded theory purpose statement sample script, and p. 135 for a grounded theory central question sample script). The sub-questions break down the central question into several component parts. Table 2 provides some exemplars of grounded theory (GT) purpose statements to help guide researchers in formulating these in their research designs.

From purpose to process: grounded theory methods in practice

Data collection in grounded theory relies on methods typically employed in qualitative research including interviewing, participant observation, document analysis, audiovisual and other materials, etc. (see Birks and Mills, 2015, p. 65) but grounded theory has historically relied most heavily on interviews as the most common data collection strategy (Creswell and Guetterman, 2019; Guetterman, 2015) and Charmaz (2014) provides as the rationale the fact that interviewing allows for optimum control as to which data are generated in the course of the study (pp. 87–88). Semi-structured interview protocols typically consist of initial, intermediate, and closing or ending questions. Charmaz (2014) provides examples of these three types of questions (pp. 67–67) in a study about a life change. Typically, initial questions are open-ended and are designed to initiate the conversation about the studied phenomenon in broad, general terms. Intermediate questions delve more deeply into answering the study's research questions in more detail, whereas closing questions can serve to elaborate and revisit the information provided in the intermediate phase and provide an opportunity for the participants to raise issues that they believe were omitted or not fully developed in the previous sections and underscore important points from their perspective.

Utilizing the constant comparative method, memo-writing, and one's ability to employ theoretical sensitivity, the grounded theorist undertakes data analysis through the use of open or initial coding procedures and then more analytical intermediate and advanced coding techniques (see Birks and Mills, 2015, Ch. 6 for an excellent discussion of these processes)—as specified by the grounded theory sub-approach employed—leading to the development of a theoretical explanation or model. One also needs to take into account standards of validation (Creswell and Poth, 2018) or trustworthiness (Lincoln and Guba, 1985) well-documented in the qualitative literature (see Babchuk et al., 2017; Creswell and Guetterman, 2019). Finally, when the researcher writes the grounded theory research report it is important to rationalize the methodological decisions that have been made and

Table 2 Examples of grounded theory purpose statements.

Author	Purpose statement	GT sub-approach
Obmerga (2021, p 1)	To provide a model which would uncover the process of sensemaking of communitarian values to enrich the transformational leadership attributes of the millennial academic supervisors	Glaserian GT
Boyer-Thurgood (2017, p. 2)	To expand the explanations of and build theory about educational applications containing virtual manipulatives	Systematic GT
Chen and Phillips (2018, p. 80)	To explore the roles of influential factors among teacher–child relationships from teachers’ perspectives and to pave the way for further investigations	Constructivist GT
Ball (2013, p. 293)	To generate a theoretical model that explains how EI is used in nursing by accelerated baccalaureate nursing students	Mixed-methods GT for theory-generation
Thornberg (2018, p. 145)	In the present qualitative study, I have employed ethnographic fieldwork (Hammersley and Atkinson, 2007; Schatzman and Strauss, 1973) guided by grounded theory methods (Charmaz, 2014; Glaser and Strauss, 1967) in order to investigate the social dynamics of school bullying.	Ethnographic grounded theory

implemented; the findings (theory, model, hypothesis);-visual representations of data (tables, graphs, figures, diagrams, etc.);-limitations and scope delimitations;-the implications of the findings;-and how the study fits into the larger literature base (see Hadley, 2017, for a comprehensive discussion on writing and presenting a grounded theory, and also Strauss and Corbin, 1990, pp. 229–230). This includes appropriate disclosure of “negative cases” or discordant data (see Gibbs, 2018, p. 131 and Brewer, 2000, pp. 132–33, as well as Creamer, 2018b for handling discordant data in mixed methods-grounded theory designs). Grounded theory researchers have provided guidelines for converting theses or dissertations into publishable articles (Strauss and Corbin, 1990), and identifying journals wherein to publish grounded theory work (Charmaz, 2014, p. 301; Creswell and Baez, 2021, pp. 253–256).

Evaluating a grounded theory

Several scholars have offered criteria for assessing or evaluating the quality of grounded theory research. Glaser (1978) advanced criteria of fit, work, relevance, and modifiability (pp. 4–5), while Strauss and Corbin (1990) also provided guidelines on how to evaluate a systematic grounded theory study (p. 253). Charmaz (2014) offered her own criteria for evaluating grounded theory studies including credibility, originality, resonance, and usefulness (pp. 337–338). In another thorough treatment of criteria for a grounded theory, Creswell and Guetterman (2019) provide indicators of lower and higher quality criteria relating to the utilization of the key elements of a grounded theory study (p. 455). Bryant (2017) also discusses quality criteria in detail (p. 363), as do Charmaz and Thornberg (2021) who developed a checklist of “flexible guidelines” (p. 17–18) to assess the quality of grounded theory research in general, and constructivist grounded theory in particular.

Reviews of the grounded theory literature over time indicate that those purporting to use grounded theory often do not attend to the suggested procedures recommended by grounded theorists, or selectively or partially use grounded theory methods and/or don’t report specific details of their methodology in their published work (Babchuk 1996, 1997, 2008, 2011; Guetterman et al., 2019). On this basis, several researchers have offered recommendations for educators interested in conducting grounded theory research (Babchuk, 1997, 2008, 2011; Bryant, 2017; Guetterman et al., 2019). These include acquiring a basic understanding of qualitative methods/theory and research including differences between quantitative and qualitative approaches as well as similarities and differences of several of the popular contemporary approaches (see Creswell and Poth, 2018 or Merriam and Tisdell, 2016). The aspiring grounded theorist should also read key works in the grounded theory tradition (see *Recommended Readings Texts* section below) to acquire an understanding of this methodology and its variants to help inform choices when selecting a sub-approach for the research. We recommend selecting one of the grounded theory variants discussed in this article and following the guidelines offered by the proponent(s) of the sub-approach. Bryant (2017) argues for a less methodologically dogmatic view that encourages students to figure out their own interpretation and application of the grounded theory methodology when conducting their research whereas Holton and Walsh (2017) advise the novice grounded theorists to develop a research stance.

Whereas some researchers choose an à la carte approach to grounded theory—adhering to only those methods that are attractive or convenient for them—attending to the full complement of grounded systematic and rigorous procedures, as well as the key characteristics outlined above, better enables the researcher to more fully develop categories, concepts, and theory. It may also be beneficial to delay an extensive use of research literature until one is analyzing and writing up the findings and developing the theoretical framework. As in all qualitative research, the researcher needs to position oneself in the study and clearly articulate this to the reader. Moreover, the researcher is not just another consideration to be taken into account but an active participant in the totality of the

research process from initial design to the dissemination of the research to a broader audience. It is important that the analyst be as transparent as possible and provide as much detail as logistically feasible so that the reader can best understand the design decisions that were made.

Examples of grounded theory studies in education

There is a robust literature base employing grounded theory methodology to study educational issues and topics (see Babchuk, 1997; Birks and Mills, 2015; Bryant, 2017; Charmaz, 2014; Creswell and Guetterman, 2019). Following Glaser and Strauss (1967) and Glaser's (1998) classic grounded theory approach, Bahng and Lee (2017) utilized the constant comparative method and open, selective, and theoretical coding practices to develop themes and theory explaining the professional journey of "elementary teacher candidates (ETC) both as learners and as teachers by exploring their learning experiences and practices" (p. 225) regarding their use of a virtual reality platform. Another example comes from Creswell and Brown (1992) who, drawing upon Strauss and Corbin (1990) grounded theory procedures of open, axial, and selective coding, interviewed 33 department chairpersons to develop a model specifying their roles in enhancing faculty research productivity.

Also relying upon Strauss and Corbin's (1998) systematic grounded theory, Komives et al. (2005) developed a conceptual model to explain the development of leadership identity among college students that illustrated processual change from a leader-centric view to a more collaborative and relational process over time. Stating they were using a constructivist grounded theory approach, Williamson et al. (2012) employed Strauss and Corbin's (1998) open, axial, and selective coding procedures to develop a substantive theoretical model of meaning making among high functioning 7–13-year-old children with autism spectrum disorder. Their grounded theory yielded three reading comprehension profiles explaining this process.

Thornberg (2018) employed Charmaz's constructivist grounded theory—utilizing theoretical sampling, constant comparison, memo-writing, and initial, open, and theoretical coding—in an ethnographic study of three urban public primary schools in Sweden. In what could be classified as a grounded theory ethnography (see below), this researcher provided an interpretative portrayal of school bullying at the macro, meso, and micro levels. In our final example in this section, Vaccaro et al. (2018) followed Charmaz (2014) initial, focused, and theoretical coding procedures, theoretical sampling, and theoretical saturation to develop a multifaceted grounded theory model of the purpose development process among college students with a wide array of disabilities. These researchers documented the developing narrative sense of self among participants explaining how these individuals negotiated this process.

Future directions

The use of grounded theory continues to grow across disciplines and subject areas, having been incorporated by scholars and practitioners in a wide array of fields to explore myriad subject areas. Earlier in this article we presented four contemporary variants or sub-approaches of grounded theory research. Two additional variants of grounded theory are ethnographic grounded theory, which we believe holds great potential for future research, and mixed methods-grounded theory, which is already widely adopted and rapidly gaining traction as the mixed methods movement has become increasingly embraced by scholars in education and the social and health sciences.

Various forms of comparative analysis and emergent theorizing have a rich and long history in ethnographic research and are well documented in the literature (see Jaccard and Jacoby, 2009; Salzman, 2011). In *The Discovery of Grounded Theory* (1967), Glaser and Strauss made explicit ties between constant comparison, grounded theory, and ethnographic research (Babchuk and Hitchcock, 2013; Glaser and Strauss, 1967). Here, Glaser and Strauss incorporated a comparative analysis accounting checklist to systematically review historical applications of the comparative method in published research by sociologists and anthropologists. Using criteria outlined in this checklist, they assigned comparative research studies from these disciplines into one of three categories: (1) Verification of Theory, (2) Assumed Verification plus Limited Generation, and (3) Comparisons for Generation. The third category represented examples of comparative research most akin to their grounded theory approach. These included such classics as Goffman's (1959) *Presentation of Self in Everyday Life*, Etzioni's (1961) *A Comparative Analysis of Complex Organizations*, Geertz (1963) *Peddlers and Princes*, and, most notably, Evans-Pritchard's (1937) *Witchcraft, Oracles, and Magic among the Azande* (Babchuk and Hitchcock, 2013; Glaser and Strauss, 1967).

Ethnographic grounded theory

Since the publication of *The Discovery of Grounded Theory* (1967) some researchers have underscored potential advantages of blending traditional features of ethnography and grounded theory's theory-methods package to be utilized for ethnographic and cross-cultural research (Aldiabat and Navaenec, 2011; Babchuk and Hitchcock, 2013; Barnes, 1996; Charmaz, 2006, 2014; Charmaz and Mitchell, 2001; Gales, 2003; Goulding, 2002; Pettigrew, 2000; Phelps and Horman, 2010; Sheridan and Storch, 2009; Timmermans and Tavory, 2007). Systematically combining elements of grounded theory and ethnography can be envisioned as representing a new hybrid qualitative approach. We have chosen to use the label **ethnographic grounded theory** herein to represent this hybrid approach.

Ethnographic grounded theory has historically been referred to in the grounded theory literature as "grounded ethnography" (Battersby, 1981, p. 93), "grounded theory ethnography" (Charmaz, 2006, p. 22), and "grounded theory in ethnography" (Charmaz, 2014, p. 35) and has been effectively utilized across disciplines and subject areas. Lauding the potential benefits of such

a merger, Battersby (1981) argued that grounded theory can “extend the limited theoretical component of ethnography” (p. 259), while Pettigrew posited that merging grounded theory and ethnography yields a level of detail “unavailable from other methodologies” (p. 260). Similarly, Sheridan and Storch (2009) upheld the potential of grounded theory to make an important contribution to intercultural research.

In the most comprehensive treatment of ethnographic grounded theory, Charmaz and Mitchell (2001) and Charmaz (2006, 2014) posited that this hybrid approach, by way of the and the constant comparative method, enables ethnographers to systematically compare data throughout the research rather than after data collection, use data to compare with emerging categories, demonstrate relationships among concepts and categories, achieve greater active involvement by ethnographers in the research process, and help elevate descriptive data analysis to veritable explanation and theory building. Moreover, Charmaz (2014) posits that the approach prevents ethnographers from uncritical acceptance of the viewpoints of participants, fieldwork lacking in direction or focus, and an over-reliance pre-existing etically derived concepts or categories. In effect, grounded theory ethnography has the potential to extend the theoretical component of ethnography that has often relied heavily upon description and the use of a priori theory. It can help shift analysis from description to explanation and theory building while reducing ethnocentrism by allowing for the emergence of culture-specific concepts rather than the dependence or over-reliance on pre-determined hypotheses or frameworks of the discipline (Babchuk and Hitchcock, 2013; Spradley, 1980). Thornberg (2018) leveraged ethnographic fieldwork including interviews and observations to ground a theory of bullying in Swedish public primary schools. Bartholomew, in a series of publications (Bartholomew, 2017, 2018; Bartholomew and Genz, 2019) has employed ethnographic grounded theory to study to mental health issues among the Awambo of Northern Namibia.

Grounded theory and mixed methods designs

In *The Discovery of Grounded Theory* (1967) Glaser and Strauss also emphasized the potential for grounded theory to be used with both quantitative and qualitative data. Since that time, scholars have argued for the effectiveness of grounded theory to be used in pluralistic designs including multimethod qualitative research and mixed methods research applications (Babchuk, 2015; Birks and Mills, 2015; Creamer, 2018a,b, 2021; Holton and Walsh, 2017; Johnson et al., 2010; Johnson and Walsh, 2019; Charmaz, 2014; Guetterman et al., 2019; Walsh, 2015). Mixed methods-grounded theory (MM-GT), a term coined by Johnson et al. (2010), pragmatist grounded theory (Babchuk, 2015), and mixed grounded theory (Johnson and Walsh, 2019) are roughly synonymous labels referring to the use of mixed methods in the service of grounded theory research. Based on Johnson et al. (2010) pioneering work, and that of Guetterman et al. (2019), we will use the term *mixed methods-grounded theory* in this article.

Mixed methods, or what has been called the third major research paradigm (Johnson and Onwuegbuzie, 2004), involves the systematic integration of quantitative and qualitative data to achieve a deeper understanding of the studied phenomenon than could be realized through a single method or research strand alone (Creswell and Plano Clark, 2018; Guetterman et al., 2019; Plano Clark and Ivankova, 2016). Traditionally, in a mixed design, qualitative and quantitative strands are strategically sequenced and combined in a mixed methods study through *integration* at the design, procedural, and/or interpretive levels (Fetters et al., 2013). Mixed methods designs have been categorized—in one widely used typology—into the three core designs of *explanatory sequential* (initial quantitative strand followed by a qualitative strand), *exploratory sequential* (initial qualitative strand followed by a quantitative strand), and *convergent* (quantitative and qualitative strands conducted independently, after which findings and results are merged), a typology explicated in detail by Plano Clark and Ivankova (2016) as well as Creswell and Plano Clark (2018). Here also exist advanced designs that showcase more complex configurations such as embedding, or more nuanced methodologies such as mixed social justice or action research designs (Creswell and Poth, 2018; Guetterman et al., 2019; Lingard et al., 2008; Plano Clark and Ivankova, 2016). Consistent with grounded theory’s historical goals to be used for either theory generation or theory verification, MM-GT designs can be configured to accomplish both these aims within one study: to use a grounded theory strand to generate or develop theory that is subsequently tested in a quantitative strand, or to follow a quantitative strand with a grounded theory one that plumbs for further detail on the preceding quantitative results.²

MM-GT offers many advantages for researchers interested in applying mixed methods approaches to the study of social phenomena. This approach is often characterized by interdisciplinarity and team collaborations whereby multiple researchers blend their perspectives and skillsets to achieve a “both-and” advantage in their studies (Johnson and Walsh, 2019, p. 524). According to Charmaz (2014), “grounded theory can contribute to each form of mixed methods research” and (p. 324) “disparate findings, particularly between interpretive understandings and measurable responses, can be significant and certainly merit further inquiry” (p. 325). MM-GT’s unique contribution to social research is the many options it features for combining multiple logics of inquiry including induction, abduction, and deduction—in an elegant alignment with grounded theory’s embrace of these multiple logics. Team grounded theory research can especially accommodate these logical approaches provided that they are deliberately harmonized and do not pose ontological or epistemological conflicts for diverse researchers combining their efforts on a jointly conceived project. Other advantages of employing MM-GT research over qualitative grounded theory used in isolation include the possibility of increasing funding opportunities, as funding agencies may prioritize quantitative research over stand-alone qualitative designs. In

²There are also instances where a quantitative grounded theory strand can be combined with a qualitative one as part of a mixed study. Johnson and Walsh (2019) offer several permutations including concurrent QUAN and QUAL GT strands, an exploratory sequence in which a QUAL GT strand precedes a QUAN GT strand, or an explanatory sequence in which QUAN GT leads to a QUAL GT phase, among other designs. Any MM-GT design counts this combination among its affordances, but the sequencing should be determined by the specific aims of the researchers.

a similar vein, mixed methods can enhance credibility of qualitative research in fields that are historically quantitative in nature, can be particularly appealing to quantitatively-trained researchers interested in expanding their toolkit as grounded theory has historically been considered more rigorous and systematic than most other qualitative approaches, and can bolster substantive or formal theory development through generation and testing of theoretical propositions in the same study (Babchuk, 2015; Guetterman et al., 2019). Along these lines, Creamer's (2018b) identified advantages of MM-GT include constructing a multilayer theory, addressing discordance between qualitative findings and quantitative results in the development of a more complex theory, and both generating and testing a theory within a study.

There are some special considerations when undertaking MM-GT studies. Researchers must each commit to a thorough and reflexive exploration of their positionality in relation to participants, data, and other aspects of the research process. These efforts toward positioning, when considering ethics in grounded theory research, are specialized in grounded theory, owing to the range of philosophies that undergird different grounded theory approaches. On an interdisciplinary MM-GT research team, it behooves each researcher to authentically and continually recalibrate one's positions and reflections to identify areas of possible bias as well as "dangers unseen and unforeseen" (Milner, 2007). The Glaserian notion that "themes emerge" inductively from the data on their own, for example, poses philosophical challenges to interpretivists who approach the research with an expanded view of the role of the researcher as an instrument of inquiry. Given that "mixed methods may divide, collide, or cohere" (Charmaz, 2014, p. 324) under the best of circumstances, special care must be taken by mixed research teams intersecting grounded theory with other designs in order to clearly and appropriately imbricate multiple ontologies, epistemologies, and methodologies in a single study.

Unfortunately, however, a review of the MM-GT literature revealed that methodological details are most often incomplete or vague in this form of research offering little guidance for researchers wishing to employ this approach for their own research (Guetterman et al., 2019; Howell Smith et al., 2020). To help address this limitation, a Mixed-Methods Grounded Theory Checklist for conducting MM-GT is provided in Guetterman et al. (2019, p. 192), and Howell Smith et al. (2020, p. 201), also offer a comprehensive list of best practices for conducting MM-GT research (see also Creamer, this volume).

Mixed methods-grounded theory exemplars

Examples of mixed methods-grounded theory studies in the literature include Ball's (2013) *explanatory sequential* design in which a quantitative measure of emotional intelligence was followed by a qualitative, constructivist grounded theory strand in which the researcher conducted interviews to generate a theoretical model of how nurses leverage emotional intelligence in their work. Berman (2017) employed an exploratory sequential mixed design to first ground a theory about researchers' data management practices and then utilize the resulting themes to develop a quantitative survey instrument that measured six dimensions of researchers' behaviors. The results of both strands were integrated into a presentation known as a *visual joint display*, which is recommended for communicating the results of mixed methods research (Guetterman et al., 2015). Howell-Smith et al. (2020) also utilized an exploratory sequential instrumental development design to develop and test a theoretical model to explain how undergraduate engineering students develop an interest in pursuing a Ph.D. in engineering.

Ontological and epistemological evolution

The application of grounded theory in mixed designs showcases the inherent nature of this methodology to act in nimble ways, aligning itself with many and variegated epistemologies and ontologies. Grounded theory is not, however, infinitely flexible and if it were, it could stand to lose its rigor and value. One can imagine that a classic, Glaserian grounded theory where "themes emerge" of their own accord—with the grounded theorist playing the role of neutral observer—would be ill fit to a project resting on a more relativist or postmodern view of reality and the role of the researcher. Likewise, constructivist notions—take to their logical extreme—may refute the existence of a shared set of observable facts. This poses an acute ontological dilemma (Boswell and Babchuk, 2022) as grounded theory enters its sixth decade and transforms—within reason—to adapt to new and renewed ways of knowing, and other seeing the world. In particular, combining constructivist GT with particular critical theoretical applications presents both challenges and opportunities for the future of this methodology.

Critical theoretical applications

Oliver (2012) makes the case for a critical realist brand of grounded theory, seeking to blend the best of both approaches toward the emancipatory aims of critical social work while embracing realist ontologies of social ills. Hadley (unpublished manuscript) outlines what might be at stake for grounded theorists—particularly constructivists—to explore ontologies and epistemologies above and beyond those that have undergirded the main grounded theory sub-approaches. For instance, a critical realist approach would have researchers differentiate ontology from epistemology, espousing a mind-independent reality that can be fallibly known. This has implications for grounded theory methodologists because it sets up a question as to whether or not one must be constructivist in belief, for example, in order to rigorously apply constructivist grounded theory methods. While embracing multiple ways of knowing and deep reflexivity, constructivist grounded theory could be rightly judged for leaning too heavily toward multiple social realities, and away from empirical facts about systemic forms of oppression and abuses of power, as centered in critical theoretical applications.

In her argument that constructivist grounded theory is well poised toward social justice aims and aligned with critical theories, Charmaz (2020) praises her sub-approach as better attending to researchers' reflexive positioning in relation to participants, particularly those from marginalized or minoritized backgrounds. She argues that research in the public sphere benefits from the

increased scrutiny of the role, possible biases, and background of the researcher. Indeed, constructivist grounded theory augments the qualitative imperative to acknowledge the researcher as “instrument” of inquiry (Creswell, 2014, p. 185). But reflexivity can also be taken to an extreme that reduces the usefulness of the grounded theory generated (Gibson, 2007, p. 440). When challenged as to whether critical inquiry can be adequately accomplished through the interpretivist tradition, Charmaz signals her disagreement with scholars advocating a critical realist ontology (2020, p. 174). In order for grounded theory to continue to grow and realize its full potential, however, it must accommodate the empirical realism that inheres in a specific critical framework, critical race theory.

Critical race theory and grounded theory methodology

Critical race theory (CRT) reshaped legal studies in the United States in the 1970s and 80s, endorsing an empirical realism that forces individuals and institutions to attend to the systematic results of raced, classed, gendered, and ableist assumptions operating beneath social agreements large and small. In the ensuing years, CRT was introduced to the field of education by Ladson-Billings and Tate (1995), who argued that educational inequalities in the US are not surprising or unexpected, and in fact that these gross inequities are “a logical and predictable result of a racialized society in which discussions of race and racism continue to be muted and marginalized” (p. 47). While Hadley (2017, 2019) advocates for a critical grounded theory and Charmaz (2011) espoused social justice aims, neither directly applies CRT into a grounded theory methodology. Grounded theory must be prepared to embrace the crucial antiracist work that requires CRT-informed methodologies both quantitative and qualitative. The next “turn” in the grounded theory universe may prove to be the *antiracist turn*.

Malagon, Pérez Huber, and Velez telegraphed this conversation in 2009, concluding that, “CRT scholarship must not only direct energy to substantive knowledge generation but must also reveal how an anti-oppressive research process can aid in achieving our central goals for racial and social justice” (p. 269). A decade later, Vanidestine and Aparicio (2019) used what they termed a “social justice approach to grounded theory” to study race and racism. What remains to be seen is how education researchers and other social scientists will resolve the ontological conflicts between the interpretivist disposition toward relativism and CRT’s need for an empirical realism that acknowledges shared facts of systemic racism and what Coates (2017) calls “the bitter calculus of oppression” (p. 100). Some promising integrations include grounded visualizations that display *situated knowledges* as well as a process known as “ground truthing” (Vélez and Solórzano, 2019)—both tools of critical race spatial analysis and both poised to bridge the history of grounded theory with the future of innovative data collection and sharing.

Conclusion

Grounded theory has been considered the most popular contemporary qualitative research methodology. Although it can also be used for quantitative applications, it is often selected as the qualitative method of choice in the blossoming third research movement of mixed methods research. Harkening back to the epistemological “paradigm wars” relating to an earlier era of qualitative research debate (Denzin and Lincoln, 2018), conflicting interpretations of the grounded theory methodology and its application are less acute than they were in earlier decades with researchers instead focusing on their own potentially successful interpretations of the grounded theory methodology rather than further politicizing these arguments (Bryant, 2017). Grounded theory has grown exponentially in its use across disciplines, settings, and problem areas over the past fifty plus years and it continues to hold great potential for the future of educational and social research. The research methodology literature owes a great intellectual debt to the genius of Barney Glaser and Anselm Strauss as well as those that continue to carry on the grounded theory tradition.

Acknowledgments

We would like to thank Dr. Gregory Hadley, Dr. Timothy C. Guetterman, and Analay Perez for their contributions to this manuscript.

Dedication: We dedicate this article to the first author of the last edition’s article on grounded theory of this Encyclopedia, Kathy Charmaz, whose guidance and influence has been immeasurable on a generation of grounded theorists including the co-authors of this manuscript.

References

- Aladiabat, K., Le Navenec, C., 2011. Clarification of the blurred boundaries between grounded theory and ethnography. Differences and similarities. Turkish Online J. Qual. Inq. 2 (3), 1–13.
- Babchuk, W., Hitchcock, R., 2013. Grounded theory ethnography: merging methodologies for advancing naturalistic inquiry. In: Proceedings of the 54th Annual Adult Education Research Conference. University of Missouri-St. Louis, St. Louis, Missouri, pp. 10–16.
- Babchuk, W.A., Guetterman, T.C., Garrett, A.L., 2017. A horse of a different color: establishing validity and reliability in qualitative research. In: Proceedings of the 36th Annual Research-To-Practice Conference in Adult and Higher Education. University of Central Oklahoma, Memphis, TN.
- Babchuk, W.A., 1996. Glaser or Strauss? Grounded theory and adult education. In: Dirkx, J.M. (Ed.), Proceedings of the 15th Annual Midwest Research-To-Practice Conference in Adult, Continuing, and Community Education. University of Nebraska-Lincoln, Lincoln, Nebraska, pp. 1–6.
- Babchuk, W.A., 1997. The Rediscovery of Grounded Theory: Strategies for Qualitative Research in Adult Education. Doctoral Dissertation. The University of Nebraska-Lincoln, Lincoln, Nebraska.

- Babchuk, W.A., 2008. Variations on a theme revisited: operationalizing grounded theory for research and practice. In: Rowland, M.L. (Ed.), *Proceedings of the 27th Annual Midwest Research-To-Practice Conference in Adult, Continuing, and Community Education*. Western Kentucky University, Bowling Green, Kentucky, pp. 7–13.
- Babchuk, W.A., 2011. Grounded theory as a “family of methods”: a genealogical analysis to guide research. *US China Educ. Rev.* 1, 383–388.
- Babchuk, W.A., 2015. Pragmatist grounded theory: advancing mixed methods for educational inquiry. In: *Proceedings of the 34th Annual Research-to-Practice in Adult and Higher Education*. University of Central Oklahoma, Oklahoma City, Oklahoma. November 15–17.
- Babchuk, W.A., 2019. Fundamentals of qualitative analysis in family medicine. *J. Fam. Med. Community Health* 7, 000040.
- Bahng, E., Lee, M., 2017. Learning experiences and practices of elementary teacher candidates on the use of emerging technology: a grounded theory approach. *Int. Electron. J. Environ. Educ.* 10, 225–241.
- Ball, L.S., 2013. Accelerated baccalaureate nursing students use of emotional intelligence in nursing as “caring for a human being”: a mixed methods grounded theory study. *Int. J. Nurs. Educ. Scholarsh.* 10, 293–300. <https://doi.org/10.1515/ijnes-2013-0015>.
- Barnes, D.M., 1996. An analysis of the grounded theory method and the concept of culture. *Qual. Health Res.* 6, 429–441. <https://doi.org/10.1177/104973239600600309>.
- Bartholomew, T.T., Gentz, S.G., 2019. “How can we help you”: mental health practitioners’ experiences of service provision in northern Namibia. *Cult. Med. Psychiatr.* 43 (3), 496–518.
- Bartholomew, T.T., 2017. Madness (Eemwengu) and its sources: conceptualizing mental illness in Namibian Ovambo culture. *J. Cross Cult. Psychol.* 48 (3), 421–437.
- Bartholomew, T.T., 2018. Beliefs about the treatment of mental illness among the Namibian Aawambo: an exploratory study. *Int. Perspect. Psychol.* 7 (4), 258–275.
- Battersby, D., 1981. The use of ethnography and grounded theory in educational research. *McGill J. Educ.* 16, 91–98.
- Becker, H.S., Geer, B., Hughes, E.C., Strauss, A.L., 1961. *Boys in White: Student Culture in Medical School*. University of Chicago Press, Chicago.
- Berman, E.A., 2017. An exploratory sequential mixed methods approach to understanding researchers’ data management practices at UVM: integrated findings to develop research data services. *J. eSci. Librariansh.* 6, e1104. <https://doi.org/10.7191/jeslib.2017.1104>.
- Birks, M., Mills, J., 2015. *Grounded Theory: A Practical Guide*, fifth ed. Sage, Thousand Oaks, CA.
- Boswell, E., Babchuk, W.A., 2022. Constructing Compatibility Among Race, Reflexivity, and Grounded Theory: A Critical Methodological Analysis of Questionable Theory-Method Combinations. Presentation to the American Educational Research Association Annual Meeting, San Diego, CA.
- Boswell, E., (in preparation). Constant Comparison in Grounded Theory Intersectionality Analyses: Tales From the Field.
- Bowers, B., Schatzman, L., 2009. Dimensional analysis. In: Morse, J.M., Stern, P.N., Corbin, J., Bowers, B., Charmaz, K., Clarke, A.E. (Eds.), *Developing Grounded Theory: The Second Generation*. Left Coast Press, Walnut Creek, CA, pp. 86–126.
- Boyer-Thurgood, J.M., 2017. The Anatomy of Virtual Manipulative Apps: Using Grounded Theory to Conceptualize and Evaluate Educational Apps that Contain Virtual Manipulatives.
- Brewer, J., 2000. *Ethnography*. McGraw-Hill Education, UK.
- Bryant, A., Charmaz, K., 2007a. Grounded theory in historical perspective: an epistemological account. In: *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 31–57. <https://doi.org/10.4135/9781848607941.n1>.
- Bryant, A., Charmaz, K., 2007b. Introduction: grounded theory research: methods and practices. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 1–28.
- Bryant, A., Charmaz, K., 2019. *The Sage Handbook of Current Developments in Grounded Theory*. SAGE Publications. <https://doi.org/10.4135/9781526485656>.
- Bryant, A., 2002. Re-grounding grounded theory. *J. Inf. Technol. Theor. Appl.* 41, 25–42.
- Bryant, A., 2017. *Grounded Theory and Grounded Theorizing: Pragmatism in Research Practice*. Oxford University Press, Oxford.
- Bryant, A., 2019. *Varieties of Grounded Theory*. Sage, Thousand Oaks, CA.
- Charmaz, K., Bryant, A., 2010. Grounded theory. In: Peterson, P., Baker, E., McGraw, B. (Eds.), *International Encyclopedia of Education*. Elsevier, Oxford, UK, pp. 406–412.
- Charmaz, K., Mitchell, R.G., 2001. Grounded theory in ethnography. In: Atkinson, P., Coffey, A., Delamont, S., Lofland, J., Lofland, L. (Eds.), *Handbook of Ethnography*. Sage, Thousand Oaks, CA, pp. 160–174.
- Charmaz, K., Thornberg, R., 2021. The pursuit of quality in grounded theory. *Qual. Res. Psychol.* 18, 305–327. <https://doi.org/10.1080/14780887.2020.1780357>.
- Charmaz, K., 2000. Grounded theory: objectivist and constructivist methods. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, Calif, pp. 509–535.
- Charmaz, K., 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage.
- Charmaz, K., 2011. Grounded theory methods in social justice research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 359–380.
- Charmaz, K., 2014. *Constructing Grounded Theory*. SAGE.
- Charmaz, K., 2020. “With constructivist grounded theory you can’t hide”: social justice research and critical inquiry in the public sphere. *Qual. Inq.* 26, 165–176. <https://doi.org/10.1177/1077800419879081>.
- Chen, S., Phillips, B., 2018. Exploring teacher factors that influence teacher-child relationships in Head Start: a grounded theory. *Qual. Rep.* 23, 80–97.
- Clarke, A.E., Friese, C., 2007. Grounded theory using situational analysis. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 363–397.
- Clarke, A.E., Friese, C., Washburn, R. (Eds.), 2016. *Situational Analysis in Practice: Mapping Research With Grounded Theory*. Routledge, New York.
- Clarke, A.E., 2005. *Situational Analysis Grounded Theory After the Postmodern Turn*. SAGE, Thousand Oaks, CA.
- Clarke, A.E., 2009. From grounded theory to situational analysis. In: Morse, J.M., Stern, P.N., Corbin, J., Bowers, B., Charmaz, K., Clarke, A.E. (Eds.), *Developing Grounded Theory: The Second Generation*. Walnut Tree Press, pp. 194–235.
- Clarke, A.E., 2019. Situating grounded theory and situational analysis in interpretive qualitative inquiry. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Current Developments in Grounded Theory*. Sage, Thousand Oaks, CA, pp. 3–48.
- Coates, T.-N., 2017. The legacy of Malcolm X. In: *We Were Eight Years in Power: An American Tragedy*. One World/Ballantine.
- Corbin, J., Strauss, A., 2014. *Basics of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA.
- Corbin, J., 2009. Taking an analytical journey. In: Morse, J.M., Stern, P.N., Corbin, J., Bowers, B., Charmaz, K., Clarke, A.E. (Eds.), *Developing Grounded Theory: The Second Generation*. Left Coast Press, Walnut Creek, CA, pp. 35–54.
- Creamer, E.G., 2018a. An Introduction to Fully Integrated Mixed Methods Research. Sage, Thousand Oaks, CA.
- Creamer, E.G., 2018b. Enlarging the conceptualization of mixed method approaches to grounded theory with intervention research. *Am. Behav. Sci.* 62, 919–934.
- Creamer, E.G., 2021. *Advancing Grounded Theory With Mixed Methods*. Routledge.
- Creswell, J.W., Báez, J.C., 2021. *30 Essential Skills for the Qualitative Researcher*, second ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Brown, M.L., 1992. How chairpersons enhance faculty research: a grounded theory study. *Rev. High. Educ.* 16, 41–62.
- Creswell, J., Guetterman, T., 2019. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, sixth ed. Pearson Education, New York, US.
- Creswell, J.W., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, fourth ed. Sage Publications.
- Denzin, N.K., Lincoln, Y.S., 2005. Introduction: the discipline and practice of qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 1–32.
- Denzin, N.K., Lincoln, Y.S., 2018. *The Sage Handbook of Qualitative Research*, fifth ed. Sage.
- Draucker, C.B., Martsof, D.S., Ross, R., Rusk, T.B., 2007. Theoretical sampling and category development in grounded theory. *Qual. Health Res.* 17, 1137–1148.

- Etzioni, A., 1961. *A Comparative Analysis of Complex Organizations*. The Free Press, New York.
- Evans-Pritchard, E.E., 1937. *Witchcraft, Oracles and Magic Among the Azande*. Clarendon Press, Oxford.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48, 2134–2156.
- Gales, L.M., 2003. Linguistic sensitivity in cross-cultural organizational research: positivist/post-positivist and grounded theory approaches. *Lang. Intercult. Commun.* 3, 131–140.
- Geertz, C., 1963. *Peddlers and Princes: Social Development and Economic Change in Two Indonesian Towns*. University of Chicago Press, Chicago.
- Gibbs, G.R., 2018. *Analyzing qualitative data*. Sage.
- Gibson, B., 2007. Accommodating critical theory. *The Sage handbook of grounded theory*, pp. 436–453.
- Giske, T., Artinian, B., 2007. A personal experience of working with classical grounded theory: from beginner to experienced grounded theorist. *Int. J. Qual. Methods* 6, 67–80.
- Giske, T., Artinian, B., 2008. Patterns of “balancing between hope and despair” in the diagnostic phase: a grounded theory study of patients on a gastroenterology ward. *J. Adv. Nurs.* 62, 22–31.
- Glaser, B.G., Strauss, A.L., 1965. *Awareness of Dying*. Aldine, Chicago.
- Glaser, B.G., Strauss, A.L., 1967. *The Discovery of Grounded Theory*. Aldine, Chicago.
- Glaser, B.G., Strauss, A.L., 1968. *Time for Dying*. Aldine, Chicago.
- Glaser, B.G., Strauss, A.L., 1971. *Status Passage*. Aldine Atherton, Chicago.
- Glaser, B.G., 1965. The constant comparative method of qualitative analysis. *Soc. Probl.* 12, 436–445.
- Glaser, B.G., 1978. *Theoretical Sensitivity*. Sociology Press, Mill Valley, CA.
- Glaser, B.G., 1992. *Emergence vs. Forcing: Basics of Grounded Theory Analysis*. Sociology Press, Mill Valley, CA.
- Glaser, B.G., 1998. *Doing Grounded Theory: Issues and Discussions*. The Sociology Press, Mill Valley, CA.
- Glaser, B.G., 2008. *Doing Quantitative Grounded Theory*. Sociology Press, Mill Valley, CA.
- Goffman, E., 1959. *The Presentation of Self in Everyday Life*. Doubleday, New York.
- Goulding, C., 2002. *Grounded Theory: A Practical Guide for Management, Business, and Market Researchers*. Sage, London.
- Guest, G., Bunce, A., Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18, 59–82.
- Guetterman, T., Creswell, J.W., Kuckartz, U., 2015. Using joint displays and MAXQDA software to represent the results of mixed methods research. In: *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*, pp. 145–175.
- Guetterman, T.C., Babchuk, W.A., Howell Smith, M.C., Stevens, J., 2019. Contemporary approaches to mixed methods—grounded theory research: a field-based analysis. *J. Mix. Methods Res.* 13, 179–195. <https://doi.org/10.1177/1558689817710877>.
- Guetterman, T.C., Babchuk, W.A., Howell Smith, M.C., (in preparation). *Intersecting Mixed Methods Research With Qualitative Designs: Principles and Practices*, Mixed Methods Research Series. Sage, Thousand Oaks, CA.
- Guetterman, T.C., 2015. Descriptions of sampling practices within five approaches to qualitative research in education and the health sciences. *Forum Qual. Soc. Res.* 16, Art. 25.
- Hadley, G., 2017. *Grounded Theory in Applied Linguistics Research: A Practical Guide*. Routledge, New York, NY.
- Hadley, G., 2019. Critical Grounded Theory. *The SAGE Handbook of Current Developments in Grounded Theory*, pp. 564–592.
- Hammersley, M., Atkinson, P., 2007. *Ethnography: Principles in Practice*, 3rd ed. Routledge, New York.
- Heath, H., Cowley, S., 2004. Developing a grounded theory approach: a comparison of Glaser and Strauss. *Int. J. Nurs. Stud.* 41, 141–150.
- Holton, J.A., Walsh, I., 2017. *Classic Grounded Theory: Applications With Qualitative and Quantitative Data*. Sage, Thousand Oaks, CA.
- Hood, J.C., 2007. Orthodoxy vs. power: the defining traits of grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 151–164.
- Howell Smith, M.C., Babchuk, W.A., Stevens, J., Garrett, A.L., Wang, S.C., Guetterman, T.C., 2020. Developing scales for a new measurement model: modeling the use of mixed methods—grounded theory. *J. Mix. Methods Res.* 14 (2), 184–206. <https://doi.org/10.1177/1558689819872599>.
- Jaccard, J., Jacoby, J., 2009. *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists*. Guilford Press, New York, NY.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Johnson, R.B., Walsh, I., 2019. Mixed grounded theory: merging grounded theory with mixed methods and multimethod research. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Current Developments in Grounded Theory*. Sage, Thousand Oaks, CA, pp. 517–531.
- Johnson, R.B., McGowan, M.W., Turner, L.A., 2010. Grounded theory in practice: is it inherently a mixed method? *Res. Sch.* 17, 65–78.
- Kelle, U., 2007. The development of categories: different approaches in grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 191–213.
- Kelle, U., 2019. The status of theories and models in grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Current Developments in Grounded Theory*. Sage, Thousand Oaks, CA, pp. 68–88.
- Komives, S.R., Owen, J.E., Longenecker, S.D., Mainella, F.C., Osteen, L., 2005. Developing a leadership identity: a grounded theory. *J. Coll. Student Dev.* 46, 593–611.
- Ladson-Billings, G., Tate, W., 1995. Toward a critical race theory of education. *Teach. Coll. Rec.* 97, 47–68.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage, Newbury Park, CA.
- Lingard, L., Albert, M., Levinson, W., 2008. Grounded theory, mixed methods, and action research. *BMJ* 337.
- Lovell, J.L., 2016. How parents process child health and nutrition information: a grounded theory model. *Appetite* 97, 138–145.
- Lynch-O'Brien, L.I., Babchuk, W.A., Dauer, J.M., Heng-Moss, T., Golick, D., 2021. Transference of citizen science program impacts: a theory grounded in public participation in scientific research. *Diversity* 13, 339.
- Malagon, M.C., Pérez Huber, L., Velez, V.N., 2009. Our experiences, our methods: using grounded theory to inform a critical race theory methodology. *Seattle J. Soc. Just.* 8, 253.
- Malterud, K., Siersma, V.D., Guassora, A.D., 2016. Sample size in qualitative interview studies: guided by information power. *Qual. Health Res.* 26, 1753–1760.
- Martin, V.B., Gynnild, A. (Eds.), 2011. *Grounded Theory: The Philosophy, Method, and Work of Barney Glaser*. BrownWalker Press, Boca Raton, FL.
- Merriam, S.B., Tisdell, E.J., 2016. *Qualitative Research: A Guide to Design and Implementation*, fourth ed. John Wiley & Sons, San Francisco.
- Milner IV, H.R., 2007. Race, culture, and researcher positionality: working through dangers seen, unseen, and unforeseen. *Educ. Res.* 36, 388–400.
- Obmerga, M.E., 2021. For whom the bell (really) tolls: a grounded theory of millennial academic supervisors' sensemaking of communitarian values as a springboard to enrich their transformational leadership attributes. *Int. J. Leader. Educ.* 1–33. <https://doi.org/10.1080/13603124.2020.1862919>.
- Oliver, C., 2012. Critical realist grounded theory: a new approach for social work research. *Br. J. Soc. Work* 42, 371–387.
- Olson, K., 2018. Ethical issues in grounded theory. In: Iphofen, R., Tolich, M. (Eds.), *The SAGE Handbook of Qualitative Research Ethics*. Sage, Thousand Oaks, CA, pp. 486–494. <https://doi.org/10.4135/9781526435446.n33>.
- Pérez, M.S., Cannella, G.S., 2013. Situational analysis as an avenue for critical qualitative research: mapping post-Katrina New Orleans. *Qual. Inq.* 19, 505–517.
- Pettigrew, S.F., 2000. Ethnography and grounded theory: a happy marriage? *Adv. Consum. Res.* 27, 256–260.
- Phelps, A.F., Horman, M.J., 2010. Ethnographic theory-building research in construction. *J. Construct. Eng. Manag.* 136, 58–65.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Plano Clark, V.L., Creswell, J.W., Green, D.O., Shope, R.J., 2008. Mixing quantitative and qualitative approaches. In: Hesse-Biber, S.N., Leavy, P. (Eds.), *Handbook of Emergent Methods*. The Guilford Press, New York, NY, pp. 363–388.
- Reichert, J., 2007. Abduction: the logic of discovery of grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. Sage, Thousand Oaks, CA, pp. 214–228.
- Saldaña, J., 2021. *The Coding Manual for Qualitative Researchers*, fourth ed. Sage.

- Salzman, P.C., 2011. *Classic Comparative Anthropology: Studies From the Tradition*. Waveland Press, Long Grove, Illinois.
- Sheridan, V., Storch, K., 2009. Linking the intercultural and grounded theory: methodological issues in migration research. *Forum Qual. Soc. Res.* 10, 1–20.
- Spradley, J.P., 1980. *Participant Observation*. Holt, Rinehart, and Winston, New York.
- Stern, P.N., Porr, C.J., 2011. *Essentials of Accessible Grounded Theory*. Left Coast Press, Walnut Creek, CA.
- Strauss, A., Corbin, J., 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage, Newbury Park, CA.
- Strauss, A.L., Corbin, J., 1994. Grounded theory methodology: an overview. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA.
- Strauss, A., Corbin, J., 1998. *Basics of Qualitative Research Techniques*. Sage, Thousand Oaks, CA.
- Strauss, A.L., 1987. *Qualitative Analysis for Social Scientists*. Cambridge University Press, New York.
- Thornberg, R., 2018. School bullying and fitting into the peer landscape: a grounded theory field study. *Br. J. Sociol. Educ.* 39, 144–158.
- Timmermans, S., Tavory, I., 2007. Advancing ethnographic research through grounded theory practice. In: Bryant, A., Charmaz, K. (Eds.), *The Sage Handbook of Grounded Theory*. SAGE Publications, Thousand Oaks, CA, pp. 493–512.
- Vaccaro, A., Kimball, E.W., Moore, A., Newman, B.M., Troiano, P.F., 2018. Narrating the self: a grounded theory model of emerging purpose for college students with disabilities. *J. Coll. Student Dev.* 59, 37–54.
- Vanidestine, T., Aparicio, E.M., 2019. How social welfare and health professionals understand “race,” racism, and whiteness: a social justice approach to grounded theory. *Soc. Work Public Health* 34, 430–443.
- Vélez, V.N., Solórzano, D.G., 2019. Critical race cartographies: exploring map-making as anti-racist praxis. *Understanding critical race research methods and methodologies: lessons from the field*, pp. 150–165.
- Walsh, I., 2014. A strategic path to study IT use through users’ IT culture and IT needs: a mixed-method grounded theory. *J. Strat. Inf. Syst.* 23, 146–173.
- Walsh, I., 2015. Using quantitative data in mixed-design grounded theory studies: an enhanced path to formal grounded theory in information systems. *Eur. J. Inf. Syst.* 24, 531–557.
- Williamson, P., Carnahan, C.R., Jacobs, J.A., 2012. Reading comprehension profiles of high-functioning students on the autism spectrum: a grounded theory. *Except. Child.* 78, 449–469.

Recommended grounded theory texts

- Birks, M., Mills, J., 2015. *Grounded Theory: A Practical Guide*, fifth ed. Sage, Thousand Oaks, CA.
- Bryant, A., 2017. *Grounded Theory and Grounded Theorizing: Pragmatism in Research Practice*. Oxford University Press, Oxford.
- Charmaz, K., 2014. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*, second ed. Sage, London.
- Corbin, J., Strauss, A., 2015. *Basics of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA.
- Creamer, E.G., 2021. *Advancing Grounded Theory with Mixed Methods*. Routledge.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. Sage, Thousand Oaks, CA.
- Glaser, B.G., Strauss, A.L., 1967. *The Discovery of Grounded Theory*. Aldine, Chicago.
- Hadley, G., 2017. *Grounded Theory in Applied Linguistics Research: A Practical Guide*. Routledge, New York, NY.
- Holton, J.A., Walsh, I., 2017. *Classic Grounded Theory: Applications With Qualitative and Quantitative Data*. Sage, Thousand Oaks, CA.

Narrative inquiry

Simmee Chung^a and D. Jean Clandinin^b, ^aConcordia University of Edmonton, Edmonton, AB, Canada; and ^bUniversity of Alberta, Edmonton, AB, Canada

© 2023 Elsevier Ltd. All rights reserved.

D.J. Clandinin, J. Huber, Narrative Inquiry, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 436–441, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01387-7>.

Narrative inquiry on the research landscape	123
Ontological and epistemological considerations in narrative inquiry	124
Thinking narratively: three-dimensional narrative inquiry space	124
Engaging participants in narrative inquiry	125
Negotiating ongoing reflexivity: narrative beginnings	125
Negotiating living, telling, retelling and reliving stories of experience	125
Negotiating being in the field	126
Negotiating from field to field texts	126
Negotiating from field texts to interim research texts	126
Negotiating from interim research texts to final research texts	126
Negotiating working with response communities	127
Design considerations	127
Justifications for narrative inquiry	127
Personal justification	128
Practical justifications	128
Social justifications	128
Touchstones of narrative inquiry	128
Relational ethical considerations	129
Conclusion	129
References	129
Further reading	130

Narrative inquiry on the research landscape

Narrative inquiry, first named in [Connelly and Clandinin \(1990\)](#), *Stories of Experience and Narrative Inquiry*, is a qualitative methodology in the field of social science research ([Clandinin, 2007](#); [Pinnegar and Daynes, 2007](#)). Narrative inquirers understand that “people shape their daily lives by stories of who they and others are and as they interpret their past in terms of these stories. Story ... is a portal through which a person enters the world and by which their experience of the world is interpreted and made personally meaningful” ([Connelly and Clandinin, 2006](#): 375). Narrative inquiry is both a way of thinking about experience and a methodology for the study of experience.

Foundational to narrative inquiry is [Dewey’s \(1938\)](#) conception of experience, that is, that experience is “a changing stream that is characterized by continuous interaction of human thought with our personal, social, and material environment” ([Clandinin and Rosiek, 2007](#):39). Further to this, it is important to recognize that a narrative view of experience is “not only on individuals’ experience but also on the social, cultural, and institutional narratives within which individuals’ experiences are constituted, shaped, expressed, and enacted” ([Clandinin and Rosiek, 2007](#): 42). Narrative inquirers study a person’s experience “that is storied both in the living and telling and that can be studied by listening, observing, living alongside another, and writing, and interpreting texts” ([Clandinin and Rosiek, 2007](#): 43). Thus, experience is understood as a storied phenomenon in that human beings compose lives that are shaped by their experiences within personal, familial, social, institutional, professional, linguistic, cultural, and historical narratives. A widely accepted definition of narrative inquiry is that it is a way of understanding experience, and is

a collaboration between researcher and participants over time, in a place or series of places, and in social interaction with milieus. An inquirer enters this matrix in the midst and progresses in the same spirit, concluding the inquiry still in the midst of living and telling, reliving and retelling, the stories of the experiences that made up people’s lives, both individual and social.

[Clandinin and Connelly \(2000: 20\)](#).

In this chapter, we describe the relational ontological and epistemological commitments of narrative inquiry highlighting the importance of working with relational ethics. Narrative inquiry works from a relational ontology that attends to people’s ordinary experiences in their everyday lives. We make explicit the narrative view of experience grounded in [Dewey’s \(1938\)](#) theory of

experience with its two criteria of experience, continuity and interaction, along with his understanding of situation. Dewey's theory of experience provides the theoretical foundation of the three-dimensional narrative inquiry space comprising temporality, sociality, and place. We describe the ways in which narrative inquirers engage with participants including negotiations of ongoing researcher reflexivity, the processes of living, telling, retelling and reliving stories of experience as narrative inquirers negotiate being in the field, moving from field to field texts, interim and final research texts. We detail the design elements of narrative inquiry that follow from a relational epistemology, ontology and ethics including considerations in methods that can be used as a researcher moves from field to field texts, and research texts. Woven throughout the chapter is the foundation for the touchstones or criteria for assessing a narrative inquiry. Relational ethical considerations and working with response communities are also discussed.

Ontological and epistemological considerations in narrative inquiry

Narrative inquiry has its ontological and epistemological roots in the pragmatic philosophy of educational philosopher John Dewey (1938, 1958). Drawing upon Dewey's work, narrative inquiries start with the view that experience is "the fundamental ontological category from which all inquiry—narrative or otherwise—proceeds" (Clandinin and Rosiek, 2007: 39). Experience has important ontological distinctions for narrative inquirers in that, within a Deweyan view, experience is a transaction between what each person thinks, what he or she shares with others, how she or he encounters and relates to others and the contexts in which each person lives. Experience is an active and recursive immersion in the field of experience—in life (Dewey, 1896/1975). The epistemological and ontological commitments in narrative inquiries require attention to the relational, ethical, contextual, and temporal unfolding of storied lives. A relational ontology shapes the relationship between knower and what is known, between knowing and action, between how one knows and what one knows, as well as the relationship between knowing and silence.

Narrative inquirers recognize that human beings are always in the midst of experience (Caine et al., 2013a). Part of the ontological commitment holds an understanding that to come alongside "as narrative inquirers we do not remove ourselves from our research puzzles and observe a phenomena or analyze people, but rather, we become a part of the ongoing negotiation of making meaning of experience" (Clandinin and Caine, 2013:177). Therefore, narrative inquirers understand their involvement requires that their experiences are also under study. It is not a study of the other but reflects the co-composed understandings of researchers and participants' experiences (Chung, 2018; Clandinin and Connelly, 2000). As Clandinin and Connelly (2000) wrote, narrative inquirers are "complicit in the world we study, we need to remake ourselves as well as offer up research understanding that could lead to a better life" (p. 61).

Narrative inquirers work from an epistemological perspective that reduces the distance or gap between what happens in the world and how people think about and thus know it. Knowledge arises through, and is modified by, transactions (Dewey, 1938) among thought, social context, time, and place. Knowledge is formed in response to a person actively adapting to, or engaging with, his or her environment and is thus linked to experience. Narrative inquirers build on Dewey's view that people are always positioned on a continuum of experiences, and emphasize that "each point has a past experiential base and leads to an experiential future" (Clandinin and Connelly, 2000: 2).

Thinking narratively: three-dimensional narrative inquiry space

Dewey's (1938) criteria of experience-continuity and interaction enacted in situations-provides the foundation for attending to experience. The criterion of continuity (Dewey, 1938) highlights the unfolding temporality of experience with an understanding that past experiences live in, and shape, future experiences. Continuity points toward how "independent of desire or intent, every experience lives in future experiences" (27) and that "every experience both takes up something from those which have gone before and modifies in some way the quality of those which come after" (35). In this transactional view of experience, relations are built "between a human being and her environment—her life, community, world" (Clandinin and Rosiek, 2007: 39). The conception of reality and knowledge is relational, temporal, and continuous between researchers and participants. Dewey's (1938) other criterion of experience, interaction, which is inseparable from the concept of situation, highlights that "an experience is always what it is because of a transaction taking place between an individual and what, at the time, constitutes his environment" (43).

Narrative Inquiry is conceptually grounded in Dewey's (1938) notion of the dynamic interplay and interaction of the two criteria. This provides groundings for what narrative inquirers name as the three-dimensional narrative inquiry space with its dimensions of temporality, sociality, and place. Attending to experience within the three-dimensional space allows narrative inquirers to study the complexity of the relational composition of people's lived experiences, in the past and in the present, as well as to imagine the future possibilities of these experiences. Within the three-dimensional space, researchers look simultaneously backward and forward (temporality), inward and outward (sociality), with attention to place(s).

Within the temporal dimension, attention is directed toward the past, present, and future of people, places, things, and events, under study. The importance of temporality in narrative inquiry is informed by philosophical views of experience where the "formal quality of experience through time is [seen as] inherently narrative" (Crites, 1971: 291). Working from Carr's (1986) view that "we are composing and constantly revising our autobiographies as we go along" (76), narrative inquirers need to attend to the temporality of their own and participants' lives, as well as to the temporality of places, things, and events.

Narrative inquirers also attend to both personal conditions and, simultaneously, to social conditions, that is with attentiveness inward to “feelings, hopes, desires, esthetic reactions and moral dispositions” (Connelly and Clandinin, 2006: 480) and outward to the milieu, the conditions under which people’s experiences and events are unfolding. These social conditions are understood, in part, in terms of cultural, social, institutional, and linguistic narratives. Connelly and Clandinin (2006) define place as “the specific concrete, physical and topological boundaries of place or sequences of places where the inquiry and events take place” (480). Drawing on Marmon Silko (1996), experiences are always situated in a particular place or in places, and within the stories we tell of these experiences.

Working within the three-dimensional narrative inquiry space also requires that we think narratively about a phenomenon, that is, that we think within dimensions of temporality, sociality, and place. Narrative inquiries highlight the shifting, changing, personal, and social nature of the phenomenon under study. Thinking narratively about a phenomenon challenges views of a phenomenon as being fixed and unchanging throughout an inquiry. Framing a research puzzle is part of thinking narratively. Each narrative inquiry is composed around a particular wonder and, rather than framing a research question with an expectation of an answer, narrative inquirers frame a research puzzle that carries with it “a sense of a search, a ‘re-search’, a searching again”, “a sense of continual reformulation” (Clandinin and Connelly, 2000: 124).

Engaging participants in narrative inquiry

As narrative inquirers work to live out the relational aspects of each inquiry, it is important to attend to the negotiating that is required throughout an inquiry: negotiating ongoing reflexivity, negotiating living, telling, retelling and reliving stories, and negotiation within each phase of the inquiry through being in the field and composing field texts, interim research texts and final research texts, negotiating the ethical relations of being alongside as well as the work with response communities.

Negotiating ongoing reflexivity: narrative beginnings

Engaging in narrative inquiry involves ongoing reflexivity, a process that begins with composing narrative beginnings. As Dubnewick et al. (2017) show, reflexivity in the form of autobiographical narrative inquiries comes out of the relational ontological commitments of narrative inquiry. Reflexivity, in the form of narrative beginnings, is a process in which narrative inquirers write autobiographically, telling stories of earlier landscapes which may provide a frame for the personal, practical, and social justifications of why the researcher is undertaking the study. However narrative beginnings include more than telling or writing stories. Part of composing narrative beginnings involves inquiry into the stories within the three-dimensional narrative inquiry space.

The research puzzle often takes shape and becomes more visible during the construction of narrative beginnings. Greene (1995) emphasizes the importance of reflecting on earlier landscapes or what she describes as our “rememory” (82) of them. It is understood that we can never truly return to earlier landscapes, however, we are able to narratively reflect on them from our present, perceived vantage point. Greene illuminates the importance of returning to earlier landscapes as she notes, “we can only become present to them by reflecting on them” (1995: 73).

Understanding life as narrative led Bruner (2004) to posit that “the stories we tell about our lives ... [are] our ‘autobiographies’” (691). Yet, narrative inquirers understand that telling stories is not an untethered process. How people tell their stories and what their stories tell is shaped by “cultural conventions and language usage ... [and] reflect the prevailing theories about ‘possible lives’ that are part of one’s culture” (694).

Narrative beginnings are a kind of autobiographical narrative inquiry similar to autoethnographic inquiry. Autobiographical narrative inquiry, “the interpretation and writing of the personal past ... is ... a product of the present and the interests, needs, and wishes that attend it. This present, however—along with the self whose present it is—is itself transformed in and through the process” (Freeman, 2007:137–138).

Negotiating living, telling, retelling and reliving stories of experience

Within narrative inquiry we work with four key terms: living, telling, retelling, and reliving. We live out stories and then tell stories of our living. In this process, inquirers continue to live their stories, even as, over time, they tell stories of their experiences. While most narrative inquiries begin with telling stories, that is, with a researcher interviewing or having conversations with participants who tell stories of their experiences, “a more difficult, time-consuming, intensive, and yet, more profound method is to begin with participants’ living because in the end, narrative inquiry is about life and living” (Connelly and Clandinin, 2006: 478).

We also retell stories through processes of inquiring into the lived and told stories. Narrative inquirers engage in inquiry with themselves and participants as they retell their stories. As we retell our lived and told stories, there are possibilities for change, that is, for reliving the retold stories. Through retelling and reliving stories of experience, we begin to restory ourselves and imagine how we might live differently in the world, and perhaps “shift the institutional, social, and cultural narratives in which we are embedded” (Clandinin, 2013: 34).

Negotiating being in the field

Narrative inquiry is a process of entering into lives in the midst of each participant's and each inquirer's life. Narrative inquirers intentionally come into relation with participants and come alongside another's life while negotiating an ongoing relational space with participants; "a relational space we call the field.... In this way, there are possible starting points for narrative inquiry: listening to individuals tell their stories and living alongside participants as they live and tell their stories" (Clandinin, 2013: 45).

What this draws attention to is the importance of acknowledging the ongoing temporality of experience when it is understood narratively. Engaging in narrative inquiries is a recursive process of being in the field, composing field texts, drafting and sharing interim research texts, and writing research texts. Lessard et al. (2020) point to the importance of being in relation throughout the research process, that is, to ways in which we continue to puzzle alongside each other and alongside participants. Without continuing to puzzle alongside each other and participants, we are in danger of becoming certain and of losing our inquiry edges. In becoming certain, we lose "the value of trying to see otherwise, of seeing on the peripheries of what is possible." Part of narrative inquiry is an openness to going into the field without tightly rehearsed scripts. This openness to otherwise "often bumps up with other types of research that may prefer specific structured questioning techniques, narrow timelines, and ideas of how to engage" (4).

Negotiating from field to field texts

Being in the field involves settling into and attending to the temporal unfolding of lives. When we engage with participants in telling their stories, Sarris reminds us that stories are often not shared in "chronological sequence" (1993: 1). In living alongside participants, hooks (1997) cautions us that people's lived and told stories are not linear—they do not necessarily "move from point A to point B" (24). The narrative qualities of lived and told stories arise from the temporal nature of experience in which people are simultaneously participants in, and tellers of, their life stories (Carr, 1986).

There are multiple ways to create field texts (data) from studying experiences of participants and inquirers in a narrative inquiry. Field texts can include transcripts of conversations, field notes, stories, memory-box artifacts, photographs, documents, art work, journals, and other texts that are composed by narrative inquirers and participants to represent aspects of lived experience. Whether narrative inquirers are listening to participants' told stories or living alongside participants as their lives unfold in particular contexts, interpretation is an essential, ongoing aspect. Being attentive to the relational aspects of working with participants within the conceptual frame of the three-dimensional space requires that narrative inquirers and participants acknowledge that they are always interpreting their pasts from their present vantage points.

Negotiating from field texts to interim research texts

Bateson (1989) writes, "Dissection is an essential part of scientific method, and it is particularly tempting to disassemble a life composed of odds and ends, to describe the pieces separately" (10). Within narrative inquiry we understand experience within the wholeness of a life lived and told rather than as fragments of a person's life. We work to resist the temptation to "disassemble a life." Gergen (2003) also cautions that an "analytical method of deconstructing stories into coded piles" could undermine "the aims of the research" (272) by directing attention away from thinking narratively about experience.

Through attending to dimensions of temporality, sociality and place, narrative inquirers and participants shape field texts into interim research texts. To move from field texts to interim research texts, a narrative inquirer begins by looking at all field texts as an "archival task," that is, they ensure they have "read and reread all the field texts and in some way sorted them" (Clandinin and Connelly, 2000: 131). This movement from field texts to interim research texts is a time marked by tension and uncertainty.

While some interpretations are underway as the inquiry is lived with participants in the field, at some point there is a move away from the intensive contact with participants. In composing interim research texts, narrative inquirers continue to think narratively, that is, positioning field texts within the three-dimensional inquiry space. As inquirers continue to compose interim research texts, usually called narrative accounts, they may discern resonant threads, or patterns, that reverberate across the lived and told stories. With this intentional focus on threads or patterns, inquirers can follow "particular plotlines that threaded or wove over time and place through an individual's narrative account" (Clandinin, 2013: 132).

Interim research texts are often partial texts. Bringing back interim research texts to further negotiate with participants around unfolding experiences is part of the process of composing interim research texts. This ongoing dialog with participants can lead to more intensive work with participants if further field texts are needed in order to compose a more complex account of the participants' experiences alongside the researcher.

Negotiating from interim research texts to final research texts

As narrative inquirers move from interim research texts to composing final research texts, there is another level of negotiation as the inquirer works to write the most telling aspects of the experiences. Often what is shared in research texts is only a portion of the overall data. These negotiations are within the inquirer as s/he looks at the experiences represented in the interim research texts as well as to the literature around the phenomenon.

Working with metaphors, creating visual and textual collages, found poetry, word images, art work, conversations, archival and purpose-taken photographs and other artifacts, narrative inquirers create research texts in order to show the complex, multilayered, storied experiences of participants and inquirers. Final research texts are the texts we make “visible to public audiences, unknown audiences who may be far removed from the lived and told experience of participants” (Clandinin, 2013: 50). As narrative inquirers compose final research texts, they return to the personal, practical, and social justifications of the work.

In narrative inquiry, “ethical matters need to be narrated over the entire narrative process” (Clandinin and Connelly, 2000: 170). In honoring the relational aspect of each inquiry, final research texts are shared with participants whenever possible but are not composed with them. Clandinin and Connelly (2000) speak to narrative inquirers’ ethical responsibilities to, and with, participants throughout the research process. Ongoing negotiation with participants around interim research texts allows narrative inquirers to create final research texts that critically represent narrative inquirers’ and participants’ experiences while also maintaining each person’s integrity and their relationships into the future. When, for example, a narrative inquiry makes visible the bumping up of participants’ lives with dominant cultural, social, and institutional narratives, working with fictionalization in research texts can enable stories to be told without harming participants’ lives or the relationships composed between narrative inquirers and participants.

Signature and voice both shape research texts. We attend carefully in order not to allow the inquirer’s voice to write over participants’ voices in the final research texts. Issues of audience are also important as narrative inquirers reflect on the purpose of their research. Final research texts are “shared in ways that help readers question their own stories, raise their own questions about practice and ... to foster reflection, storytelling and restorying for readers (Connelly and Clandinin, 1990: 20).

Negotiating working with response communities

It is often helpful for narrative inquirers to participate in response communities where, alongside other narrative inquirers as well as people from diverse backgrounds, they share and respond to one another’s thinking or writing in progress, in ways that are attentive to the lives being represented. Response communities may profoundly shape each aspect of a narrative inquiry, noting the importance of members being attentive to “who we each are and are becoming in the research and how relationships shape how we come to conceptualize research puzzles as well as how we live alongside research participants. How we begin shapes the unfolding of our inquiries, it shapes what is possible” (Lessard et al., 2020: 2).

There are seven interwoven features of response communities (Caine et al., 2018) that are necessary for a response community to be a space that sustains, allows for the giving back of stories, and allows researchers to become visible in the ways we live out our stories as people in relation. Within narrative inquiry, response communities are embedded within the social narratives of researchers’ lives and help researchers understand themselves “within what Arendt (1958) would call public spaces. Response communities help us face the world in a different way, one that makes different social, political relationships possible” (Caine et al., 2020: 5). Response communities also help make visible other considerations that narrative inquirers need to attend to when coming alongside participants. As well, response communities work to caution narrative inquirers about the importance of slowing down, of continuing to find ways to be in relation, and of working to develop relationships with participants as ones marked by a sense of mutuality, respect, and reciprocity. In order to slow down and move in more relational ways, “it is important to recognize that the questioning of who we are in each situation, in each relationship, and in each place is part of the ways narrative inquiry is lived” (Lessard et al., 2020: 4).

Before, during, and after research projects end, narrative inquirers continue to engage in conversation with response communities and to stay in relation with former participants. While it matters how we begin a narrative inquiry, it also matters how an inquiry shapes forward looking stories for participants and for inquirers. The relational process of engaging in a narrative inquiry alongside participants carries a responsibility and understanding that “despite a focus on particular phenomena we are interested in understanding, we do not lose sight of each other and of how our lives are shaping the research and how the research is shaping us (Chung 2018; Clandinin et al., 2018)” (Lessard et al., 2020: 2).

Design considerations

There are important considerations in designing and engaging in narrative inquiry. It is crucial that narrative inquirers are able to articulate the purposes and justifications for their research. The criteria for judging narrative inquiries, what are termed the touchstones, will also be explored as part of the research design considerations.

Justifications for narrative inquiry

In narrative inquiry it is important to think about justifying the research in three ways—personal, practical, and social. The researcher’s justifications for engaging in the research are embedded in the narrative beginnings which provide the rationale of how, and why the narrative inquirer came to their research puzzle. At the end of the study, the researcher returns to the justifications to retell, relive, and re-imagine new possibilities as they think with the stories lived and told by participants. Below is a more detailed description of each justification.

Personal justification

Narrative inquirers begin with personal justifications, that is, by justifying the inquiry in the context of their own life experiences and what personally brings them to their research puzzle. It is important to inquire into who inquirers see themselves as being, and becoming, within the inquiry in order to be wide awake (Greene, 1995) to the experiences of those we come alongside. The narrative inquirer positions oneself as part of the inquiry, using “I” to share their personal justification. Narrative inquiries including theses and dissertations begin with autobiographical stories lived and told by the researcher which are guided by the three-dimensional Inquiry space. The personal narratives and reasons which have inspired or provoked the narrative inquirer for undertaking the research are shared at the onset and returned to at the end of the study. For instance, in Chung’s (2016) narrative inquiry alongside Indigenous youth and their families, she begins with an inquiry into her lived and told stories of experience of growing up as a child newly immigrated to Canada. Thinking with stories of the youth, she returns at the conclusion of the inquiry to look at the ways her early experiences shaped her understandings of belonging and identity-making (Chung, 2016).

Practical justifications

Narrative inquirers also justify their research practically, that is, by explicitly explaining the possibility of how the inquiry might shift or change one’s practice. As part of the practical justification, the researcher may provide pragmatic reasons for undertaking the narrative inquiry as it relates to their practice. For example, practical justifications are sometimes made in narrative inquiries around questions of ameliorating early career teacher attrition (Schaefer et al., 2014) or in medical education around puzzles concerning the conditions under which medical residents engage in reflecting on their clinical practice (Clandinin and Cave, 2008).

Social justifications

Narrative inquiries must also attend to social justifications in terms of addressing the so-what and who-cares questions important in all research undertakings (Clandinin and Connelly, 2000). Social justifications can be understood in terms of theoretical justification as well as social action and policy justifications. Theoretical justification is understood in terms of new methodological knowledge and/or advancing disciplinary knowledge. The social justification for the research may provoke social actions and/or policy shifts or changes that impact more than the researcher and participant, but also specific demographics, communities, or potentially society as a whole. For example, see studies on early school leavers (Caine et al., 2013b) which offers understandings about the diversity of youth’s experiences of leaving school that lead to changes in school policies and practices, and studies on narrative care (Blix et al., 2020) which offers new concepts for understanding care for elderly and others who need health support.

Touchstones of narrative inquiry

Criteria for judging narrative inquiries follow from the definition of narrative inquiry and the conceptual frame for thinking narratively. To consider the criteria for assessing the quality of a study, narrative inquiries need to be considered with attention to the ontological, methodological, and ethical commitments lived out in a narrative inquiry. Clandinin and Caine (2013) developed a working set of “touchstones” or quality markers, to be used as criteria for assessing narrative inquiries. They draw attention to a touchstone as a quality or example that is used to test excellence or genuineness. Using the metaphor of a hard black rock which may be used to test the quality of gold or silver, the streaks left on the stone by one of these metals is compared with that of a standard alloy. Clandinin and Caine (2013) discussed the critical components of a narrative inquiry and identified twelve kinds of streaks to mark the touchstones.

As Clandinin and Caine developed these touchstones, they turned to the work of other researchers whose work provided exemplars of narrative inquiry as well as to the theoretical roots of narrative inquiry (Caine et al., 2022). The touchstones outlined include:

- A. attentiveness to relational responsibilities, which turns our attention to the need to answer to those with whom we are in relation, to be responsive to others and to the communities of which we are part;
- B. understanding of what it means to engage in the temporal midst, that is, to temporality within, and across, lives, within storied memories, and the making of new memories and with attention to temporality that stretches backwards and forwards;
- C. the centrality of negotiation of relationships at all phases of an inquiry;
- D. the importance of engaging in autobiographical narrative inquiries into who we are, and are becoming, as narrative inquirers. We need to pay close attention to who we are in the inquiry and to understand that we are part of the storied landscapes we are studying;
- E. processes of negotiating entry to the field understood as a co-composed inquiry space between researcher and participant;
- F. attention to the move from field to field texts, and the move from field texts to interim and final research texts;
- G. the need to represent narratives of experience in ways that attend to temporality, sociality, and place;
- H. the place of relational response communities;
 - I. the delineation of personal, practical, social justifications at the beginning and conclusion;
 - J. attentiveness to voices and multiple audiences including participants; and
 - K. commitment to understanding unfolding lives in motion or in the midst.

Relational ethical considerations

Narrative inquirers follow the legal and procedural aspects of ethics outlined by institutional research boards. Ethical considerations are of vital importance because of the relational ontological commitments and thus requires that narrative inquirers undertake research with an understanding of relational ethics. Relationships marked by mutuality, reciprocity, and respect between participants and researchers are central. Clandinin et al. (2018: 198) note that working with relational ethics asks inquirers “to recognize that we live in, and co-create with participants, spaces of uncertainty or liminality, that both defy simply answers to questions of living and simultaneously acknowledge not knowing as part of living ethically in which we, as researchers, take care to position ourselves in places of dis/ease.”

An ethic of care (Noddings, 1984) is a necessary starting point for narrative inquiry, however, narrative inquirers move beyond an ethic of care to relational ethics. While an ethic of care is lived with participants throughout an inquiry, “an ethic of care is not sufficient alone to sustain a narrative inquiry relationship” (Caine et al., 2019:4). Within narrative inquiry, we cannot think with only a personal ethic but must work with a social ethic that allows us to see ourselves as part of larger communities with responsibilities and commitments to more than ourselves and our personal and familial relationships. This move to include both an ethic of care and relational ethics involves a shift to becoming together with responsibility to, and for, each other.

While narrative inquiry is a relational research methodology, it is also “a transaction between people which makes ethical issues and concerns about living well with others central to the inquiry” (Caine et al., 2013a: 580). Therefore, as narrative inquirers live out an ethics of care which directs us “to attend, to listen, and to respond as positively as possible” (Noddings, 2012: 109), narrative inquirers also live out a relational ethics which directs attention to the complex dimensions within research relationships. Caine et al. (2013a) note the ways in which a relational ontology shapes our work with participants’ experiences as we move to issues of representation in research texts. Representations that arise from our relational ontology necessitate “our living with the unfitting story rather than with attempts to tame, sanitize, or analyze. As narrative inquirers we attend to difficult stories and experiences, we stay with them; we dwell alongside potential participants in possible ways to retell them” (581).

The relational aspects compel narrative inquirers to pay attention to particular ethical matters as research texts are written and shared with public audiences. Narrative inquirers understand that a person’s lived and told stories are who they are and who they are becoming and that these stories sustain them. This understanding shapes the necessity of writing research texts that respectfully represent participants’ lived and told stories.

Conclusion

Engaging in narrative inquiries requires careful planning with time to develop and sustain relationships. It requires intentionality, reciprocity, mutual respect, and openness to co-composing inquiry spaces alongside participants. This chapter describes the theoretical foundations of Narrative Inquiry, the processes of thinking narratively, and processes for designing and engaging in an inquiry, and criteria or touchstones for assessing an inquiry. However, the heart of Narrative Inquiry is relational ethics (Clandinin et al., 2018). This relational ethical stance comes with the responsibility and commitment from researchers and participants to co-compose and negotiate spaces within the re/search to retell, relive, and re-imagine new possibilities and ways of living responsibly, individually and collectively, in the world.

References

- Arendt, H., 1958. *The human condition*. The University of Chicago Press, Chicago.
- Bateson, M.C., 1989. *Composing a Life*. Atlantic Monthly Press, New York.
- Blix, B., Berendonk, C., Clandinin, D.J., Caine, V., 2020. The necessity and possibilities of playfulness in narrative care with older adults. *Nurs. Ethics*. <https://doi.org/10.1111/nin.12373>.
- Bruner, J., 2004. Life as narrative. *Soc. Res.* 71 (3), 691–710.
- Caine, V., Estefan, A., Clandinin, D.J., 2013a. A return to methodological commitments: reflections on narrative inquiry. *Scand. J. Educ. Res.* 57 (6), 574–586.
- Caine, V., Lessard, S., Steeves, P., Clandinin, D.J., 2013b. Narrative understandings of lives lived in and out of schools. In: Tillizek, K., Ferguson, B. (Eds.), *Youth, Education, and Marginality: Local and Global Expressions*. Wilfred Laurier University Press, Waterloo, ON, pp. 197–216.
- Caine, V., Lessard, S., Clandinin, D.J., 2018. Lingering departures: reverberations through, and in, narrative inquiry response communities. In: Hanne, M., Kaal, A.A. (Eds.), *Narrative and Metaphor in Education: Look Both Ways*. Routledge, Abingdon and New York, NY, pp. 278–290 (Chapter 20).
- Caine, V., Chung, S., Steeves, P., Clandinin, D.J., 2019. The necessity of a relational ethics alongside Noddings’ ethics of care in narrative inquiry. *Qual. Res.* <https://doi.org/10.1177/1468794119851336>. Retrieved from.
- Caine, V., Clandinin, D.J., Lessard, S., 2020. Considering response communities: spaces of appearance in narrative inquiry. *Qual. Inq.* 1–6. <https://doi.org/10.1177/1077800420948105>.
- Caine, V., Clandinin, D.J., Lessard, S., 2022. *Narrative Inquiry: Theoretical Roots*. Bloomsbury Publishing, London.
- Carr, D., 1986. *Time, Narrative, and History*. Indiana University Press, Bloomington, IN.
- Chung, S., 2016. *A Narrative Inquiry Into Aboriginal Youth and Families’ Experiences of Belonging as Interwoven With Identity Making*. Doctoral Dissertation. University of Alberta. Retrieved from. <https://era.library.ualberta.ca/items/b67d1949-3b75-4a4a-950e-2baebc104bae>.
- Chung, S., 2018. Education is ceremony: thinking with stories of Indigenous youth and families. *LEARNing Landscapes* 11 (2), 93–108.
- Clandinin, D.J., Caine, V., 2013. Narrative inquiry. In: Trainor, A.A., Graue, E. (Eds.), *Reviewing Qualitative Research in the Social Sciences*. Routledge, New York, pp. 166–179.
- Clandinin, D.J., Cave, M., 2008. Creating pedagogical spaces for developing doctor professional identity. *Med. Educ.* 42 (8), 765–770.

- Clandinin, D.J., Connelly, F.M., 2000. *Narrative Inquiry: Experience and Story in Qualitative Research*. Jossey-Bass, San Francisco, CA.
- Clandinin, D.J., Rosiek, J., 2007. Mapping a landscape of narrative inquiry: borderland spaces and tensions. In: Clandinin, D.J. (Ed.), *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage, Thousand Oaks, CA, pp. 35–75.
- Clandinin, D.J., Caine, V., Lessard, S., 2018. *Relational Ethics in Narrative Inquiry*. Routledge, London.
- Clandinin, D.J. (Ed.), 2007. *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage, Thousand Oaks, CA.
- Clandinin, D.J., 2013. *Engaging in Narrative Inquiry*. Left Coast Press, Walnut Creek, CA.
- Connelly, F.M., Clandinin, D.J., 1990. Stories of experience and narrative inquiry. *Educ. Res.* 19 (5), 2–14.
- Connelly, F.M., Clandinin, D.J., 2006. Narrative inquiry. In: Green, J., Camilli, G., Elmore, P. (Eds.), *Handbook of Complementary Methods in Education Research*. Erlbaum, Mahwah, NJ, pp. 375–385.
- Crites, S., 1971. The narrative quality of experience. *J. Am. Acad. Relig.* 39 (3), 291–311.
- Dewey, J., 1896/1975. The reflex arc concept in psychology. In: Boydston, J.A. (Ed.), *The Early Works of John Dewey, Volume 5, 1882–1898: Early Essays, 1895–1898*. University of Southern Illinois Press, Carbondale, IL.
- Dewey, J., 1938. *Experience and Education*. Collier Books, New York, NY.
- Dewey, J., 1958. *Experience and Nature*. Dover, New York, NY.
- Dubnewick, M., Clandinin, D.J., Lessard, S., McHugh, T., 2017. The centrality of reflexivity towards narrative beginnings: towards living reconciliations. *Qual. Inq.* 1–8. <https://doi.org/10.1177/1077800417727762>.
- Freeman, M., 2007. Autobiographical understanding and narrative inquiry. In: Clandinin, D.J. (Ed.), *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage, Thousand Oaks, CA, pp. 120–145.
- Gergen, M., 2003. Once upon a time. A narratologist's tale. In: Daiute, C., Lightfoot, A. (Eds.), *Narrative Analysis: Studying the Development of Individuals in Society*. Sage, Thousand Oaks, CA, pp. 267–285.
- Greene, M., 1995. *Releasing the Imagination: Essays on Education, the Arts, and Social Change*, first ed. Jossey-Bass, San Francisco, CA.
- hooks, B., 1997. *Wounds of Passion: A Writing Life*. Henry Holt, New York.
- Lessard, S., Kootenay, I., Whiskeyjack, F., Chung, S., Caine, V., Clandinin, D.J., 2020. Working with indigenous elders in narrative inquiry: reflections and key considerations. *Qual. Inq.* 11 (2), 93–108.
- Marmon Silko, L., 1996. *Yellow Woman and a Beauty of the Spirit: Essays on Native American Life Today*. Touchstone, New York.
- Noddings, N., 1984. *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press, Berkeley.
- Noddings, N., 2012. *Peace Education: How We Come to Love and Hate War*. Cambridge University Press, Cambridge.
- Pinnegar, S., Daynes, G., 2007. Locating narrative inquiry historically: thematics in the turn to narrative. In: Clandinin, D.J. (Ed.), *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage, Thousand Oaks, CA, pp. 3–34.
- Schaefer, L., Downey, C.A., Clandinin, D.J., 2014. Shifting from stories to live by to stories to leave by: early career teacher attrition. *Teach. Educ. Q.* 41 (1), 9–27.

Further reading

- Anzaldúa, G., 1987. *Borderlands/La Frontera: The New Mestiza*. Aunt Lute Books, San Francisco, CA.
- Caine, V., Clandinin, D.J., Lessard, S., 2022. *Narrative Inquiry: Theoretical Roots*. Bloomsbury Publishing, London.
- Charon, R., Montello, M. (Eds.), 2002. *Stories Matter: The Role of Narrative in Medical Ethics*. Routledge, New York.
- Clandinin, D.J. (Ed.), 2007. *Handbook of Narrative Inquiry: Mapping a Methodology*. Sage, Thousand Oaks, CA.
- Coles, R., 1989. *The Call of Stories: Teaching and the Moral Imagination*. Houghton Mifflin, Boston.
- Downey, C.A., Clandinin, D.J., 2010. Narrative inquiry as reflective practice: tensions and possibilities. In: Lyons, N. (Ed.), *The Handbook of Reflection and Reflective Inquiry*. Springer, New York, NY, pp. 385–400.
- Kerby, A.P., 1991. *Narrative and the Self: Studies in Continental Thought*. Indiana University Press, Bloomington, IN.
- King, T., 2003. *The Truth About Stories: A Native Narrative*. House of Anansi Press, Toronto, ON.
- Lugones, M., 1987. Playfulness, “world”-travelling, and loving perception. *Hypatia* 2 (2), 3–19.
- Paley, V.G., 1997. *The Girl With the Brown Crayon*. Harvard University Press, Cambridge, MA.
- Sarbin, T.R., 2004. The role of imagination in narrative construction. In: Daiute, C., Lightfoot, C. (Eds.), *Narrative Analysis: Studying the Development of Individuals in Society*. SAGE, Thousand Oaks, CA, pp. 5–20.
- Steeves, P., 2006. Sliding doors—opening our world. *Equity Excell. Educ.* 39 (2), 105–114.

Phenomenology and education: researching pedagogical experience

Norm Friesen, Boise State University, College of Education, Boise, ID, United States

© 2023 Elsevier Ltd. All rights reserved.

Phenomenology, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 449–455, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01539-6>.

History: Husserl, Heidegger & Merleau-Ponty	131
From philosophy to educational research	133
Phenomenology and methodology: Amedeo Giorgi and Max van Manen	134
Amedeo Giorgi: “scientific” psychological phenomenology	135
Max van Manen: hermeneutic pedagogical phenomenology	136
Giorgi, van Manen, and contemporary qualitative research: an evaluation	138
Pedagogical phenomenology: possible futures	139
References	139

Phenomenology is the study of experience as it is concretely lived, a study that works to develop general structures or interpretations from this concrete particularity. Pedagogy, on the other hand, can be delineated through the “pedagogical difference” that separates the experience of the educator from the experience of those being educated. Phenomenology and pedagogy converge not just in their interest in particular experiences on either side of this divide; both imply a position or an “attitude” from which everyday pedagogical experience can be understood.

Phenomenology is a philosophical tradition and an approach to qualitative research that has been adapted to the study of lived experience in fields as diverse as psychology, psychiatry and sociology. In both philosophy and qualitative research, phenomenology is seen as embodying a method with multiple stages or facets, known collectively as the reduction, bracketing or *epoché*. Today phenomenology designates an active and varied international field of research and theory, one that continues to address urgent issues and concerns in education and in a wide variety of other domains.

Edmund Husserl (1859–1938), the founder of phenomenology, conceived of it in explicit opposition to the specialized natural sciences as epitomized in his famous call to “go back to the things themselves” (1900/2001, p. 168). Phenomenology focuses on *subjectivity*, on experience as it unfolds, on the contents of consciousness, rather than on the *objectivity* claimed by scientific experiments or psychological tests and measures. Phenomenology, Husserl writes, describes “objects”—today: phenomena or experiences—as “first constituted as being, what they are for us, and as what they count as for us, in varying forms of ... intention” (1913/2001, p. 275). Our immediate experience of the world is not of an accumulation of indifferent facts and data; it is already “a world of objects with values, a world of goods, a practical world,” a lifeworld (1913/2001, p. 53, emphasis in original). Subject and world are not separated, but are inextricably connected through *intentionality*, the fact that our perceptions, awareness and actions are always directed, are always *about* something. In everyday life, a tree is shade on a hot summer day just as a building doorway can be a refuge on a windy or rainy one. Phenomenology first carefully separates experience as lived from this everyday “aboutness” and asks (for example) about lived objects or experiences themselves (investigating, for example, refuge, exposure, or sheltering trees or doorways).

To understand phenomenology in its pedagogical dimension, it is important to recognize that although the educator and young person are separated by the pedagogical difference, they are united by a relation with particularly pedagogical characteristics and by their immersion in a common lifeworld. Phenomenology, then, can study a child’s book or toy not merely a way to keep them preoccupied or out of trouble—as a busy parent might initially intend. The experience of a book or a toy can instead be revealed as a way of focusing and directing both attention and behavior, structuring the child’s relationship to time, space, and to others around them. A specifically *pedagogical* phenomenology is thus for example able to study childhood risks (e.g. on playgrounds) as something important for their “becoming at home in the world” (Smith, 1998, p. 182; see also Dall’Allba, 2009; Friesen et al., 2012, for further examples).

This chapter begins with brief histories of the development and expansion of phenomenology, both as a philosophy and a research approach in education. It then offers an overview of recent and contemporary variations of phenomenology for research in education and psychology, including a critical evaluation of their deviation from contemporary qualitative validation techniques. It concludes by describing examples of recent developments related to phenomenology and pedagogy.

History: Husserl, Heidegger & Merleau-Ponty

Edmund Husserl conceived of phenomenology very ambitiously. He saw it as nothing less than a new science to rival all others: as a presuppositionless undertaking based on a “pure transcendental ego” that can—through a method that is itself “transcendental”—intuit and describe “pure essences” and develop from them “the most general structure of this pure consciousness” (1913/

2001, pp. xxi, 11). Although such terms and phrases may seem abstruse if not also philosophically problematic, this chapter illustrates that versions of each still play important role in phenomenology today. In his book *Ideas: General Introduction to Pure Phenomenology*, Husserl, like Descartes, takes as his “first consideration” a perceiving self, sitting in a room (Descartes 1637/2006, p. 13; cf. Husserl, 1913/2001, p. 53), reflecting and observing, and extrapolating from this an ideal self (i.e. ego) able to apprehend objects without prejudice or bias. This is a self that is able to fully carry out Husserl’s phenomenological method, the reduction, in its three parts: (1) the “eidetic reduction” which varies the phenomenon in question (in one’s imagination or via concrete examples), comparing and contrasting it in order to establish both what it is and what it is *not*¹; (2) the skeptical *epoché* (also known as *Deskription*), which is the provisional *exclusion* of existing and possible interpretations, theories and explanations to allow the phenomenon appear for what it is; (3) the rejection of the “natural attitude” of everyday purposeful engagement with a relatively impartial, contemplative “phenomenological attitude” (pp. 51–53, 112–114). This three-stage method, Husserl reasons, could theoretically be applied to any given domain of empirical and ideal objects (e.g. concepts) to produce what he called a “regional ontology” (p. 18).

Husserl’s transcendental ego has been widely critiqued. Researchers in phenomenology, like anyone else, are inevitably motivated and fallible. Nonetheless, phenomenological research (like qualitative research in general) has developed ways of compensating for and counteracting these limitations (see below, **Giorgi, van Manen, and contemporary qualitative research: an evaluation** section). Husserl’s versions of the reduction, particularly his notion of *Deskription* remain indispensable in contemporary phenomenology and research. Different methodologists have come up with various ways of understanding and labeling these techniques and their practical implications. Moreover, Husserl’s meticulous analyses of experiences from the first-person perspective still provide valuable insight into “general structures” of experience. For example, Husserl analyzes how the present moment of experience is ordered and “prefigured” through *expectations* based on past experience, and how these expectations can be *broken*:

Expectations [are] ... essentially susceptible to disappointment in all circumstances. Perception brings something new; that is its nature ... [And] it is certainly clear that it is perception that first decides, and that something new can be a slap in the face to all expectation. Think only of prefigured sequences of sensations like melodies.

Husserl (2012, p. 263).

My computer screen, for example, flickers and goes blank, revealing itself—against all expectation—as an inert, perhaps blemished or dusty black surface. Perception or lived experience has ways of surprising us like this and resisting the immediate realization of our intentions. Not only does this ever-present possibility help phenomenology guard against idealizations and solipsisms, Husserl’s notion of the disappointment of expectations forms the cornerstone of the phenomenological account of “learning-as-experience” (see **Pedagogical phenomenology: possible futures** Section, below).

Husserl’s work had an almost immediate international impact, with his insights developed and adapted, sometimes quite boldly, by others, including his own students. Martin Heidegger (1889–1976), the most (in)famous of these, transformed the phenomenology of his teacher in his most influential work, *Being and Time* (1928/1962). Here, Heidegger turns Husserl’s *pure, transcendental* project—focused on *essences*—into an *existential* one—concerned with *being*. Also, unlike Husserl, Heidegger does not begin *Being and Time* by invoking an act of solitary and sedentary contemplation. Instead, Heidegger opens with a critique of Descartes’ rational individualism and with the recognition that we are inescapably “thrown” into both world and time. This “thrownness,” this changing “mode of being,” is what Heidegger refers to as our “being-there” or *Da-sein*. Phenomenology comes into play when aspects of our “being” reveal themselves to us—moments when our own basic conditions come into view or are disclosed in particular ways: “Phenomenology” is thus the analysis of these conditions, of our *Dasein* (*Daseinsanalyse*). It “means to let that which shows itself be seen from itself in the very way in which it shows itself from itself” (Heidegger, 1928/1962, p. 58). The best example of this for Heidegger are moments of crisis where we confront our own death—moments which confirm that our being is time-bound, that being is time, or as Heidegger says “the meaning of *Dasein* is temporality” (p. 380).

Heidegger sees us as engaging with the disclosure of our own being (*Dasein*) through *interpretation*: we work to make sense of, to derive meaning from what we encounter in life. His is thus an *interpretive* or *hermeneutic phenomenology*. Heidegger also develops interpretations, analyses of *Dasein* that help to focus and guide phenomenological research: For example, he shows how we are beings who are *with* others (*Mitsein*) primarily through our common attunement or atmosphere (pp. 172–179). In doing so, he provides a basis for the phenomenological study of lived classroom atmospheres and school climates (e.g. see: Wolf, 2020).

Working from Husserl’s phenomenology and Heidegger’s *Being and Time*, Maurice Merleau-Ponty (1908–1961) developed an existential phenomenology of *embodied* existence. The preface to his *Phenomenology of Perception* (15 pages of required reading for any phenomenological researcher) presents a sympathetic critique of Husserl’s pure ego and his transcendental reduction, and offers this famous characterization of phenomenological “reflection.”

¹Eidetic, meaning “essential,” is taken from the Greek *eidos* (essence). Husserl characterizes phenomenology as a “new eidetic science” (1913/2001, p. 34). More narrowly, the eidetic reduction also refers to what Husserl calls imaginative “free variation” (2012, pp. 354–355), a process which consists of imagining different variations on the (type of) experience being studied.

Reflection does not withdraw from the world toward the unity of consciousness as the foundation of the world; rather, it steps back in order to see transcendences spring forth and it loosens the intentional threads that connect us to the world in order to make them appear; it alone is conscious of the world because it reveals the world as strange and paradoxical ... Perhaps the best formulation of the reduction is the one offered by Husserl's assistant Eugen Fink when he spoke of a "wonder" before the world.

Merleau-Ponty (1945/2012, p. lxxvii).

Merleau-Ponty rejects Husserl's transcendental ego and also notes that "the most important lesson which the reduction teaches us is the impossibility of a complete reduction" (p. lxxvii). Phenomenological engagement for Merleau-Ponty is thus not a retreat from the world, but rather, a matter of experiencing it in a different way. Phenomena are experienced as being released from their common practical or theoretical meanings and are thus allowed to "show themselves" in their relative generality, in ways perhaps strange and wonderful. Embodiment, as both the reference and the mode of our immersion in concrete phenomena, is our "general means of having a world" as well as being, as Merleau-Ponty puts it, an "eminently expressive space" (p. 174). Our "flesh" for Merleau-Ponty is the embodiment of this opposition (i.e. of both having a world and expressing ourselves in it) and of many other oppositions as well: The body sees and is seen; it can be both touching and touched, hearing and heard, sensed and sentient, matter and spirit (1964/1968, pp. 134–144). Merleau-Ponty's work enshrines the body and the space it inhabits in the form of what Heidegger calls an *existentiale*, one of the fundamental dimensions or conditions of our being. In addition to *time* as "the meaning of Dasein," these *existentiale* or conditions also include, as indicated above, our "being with," our *relationality* or Heidegger's *Mitsein*.

"Phenomenology's most important accomplishment," Merleau-Ponty concludes, is that it has "joined an extreme subjectivism with an extreme objectivism." The results of this include a kind of synthesis of objective and subjective that Merleau-Ponty refers to as "intersubjectivity" (2012, p. lxxvi) and "intercorporeity" (1964/1968, p. 141)—experienced conditions in which "perspectives intersect, perceptions confirm each other, and a sense appears ... the sense that shines forth at the intersection of my experiences with those of others ..." (2012, p. lxxiv).

From philosophy to educational research

The story of the adaptation of the phenomenologies of Husserl, Heidegger, and Merleau-Ponty to the domain of pedagogy and education is a long and varied one. It includes, for example, the work of the Dutch Utrecht School of phenomenology in medicine, psychology and human development as well as analyses of the future of education following Heidegger's later conception of the history of *Dasein* (*Daseinsgeschichte*; e.g. Fink, 1959/2022). However, given its brevity, this chapter considers only two exemplary works which explicitly link the phenomenological method together with the study of pedagogical experience. These texts—until recently unavailable in English—provide examples of methods and findings that not only exemplify a specifically "pedagogical phenomenology" but more generally illuminate the path leading from philosophical phenomenology to empirical research.

As early as 1914—while Husserl was still formulating his first philosophy—Aloys Fischer (1880–1937) adapted Husserl's skeptical reduction to constitute a specifically "*deskriptive Pädagogik*" (1914/2021; all quotes from foreign-language sources in this chapter are my own translation.) Fischer begins his text bearing this same title by discussing the difficult relationship of *theory* with everyday pedagogical *practice*: "Theory criticizes practice, wants to improve it, reform it, educate it," Fischer observes (1914/2021, p. 61). This is because "new proposals, criticism itself, are valued more highly than a solid understanding of *what already exists*" (p. 51, emphasis added). In keeping with Husserl's call to "return to the things themselves," Fischer outlines how phenomenology allows research in education to study precisely that which "already exists ... to secure the phenomenon of education as the starting point and guide for [understanding pedagogical] problems" (p. 52). Fischer then explains the bracketing process, Husserl's *Deskription*, in a way that remains instructive for researchers today:

If one wants to describe an actuality to be first investigated as it is, through the phenomenological reduction requiring the renunciation of preconceived interpretations, then it is necessary that one grasps and holds this phenomenon with unwavering fixation, that one lets the individual characteristics become clear and distinct—not even using a name or even a term [as a label]—until the structure of the object, its position to me, its parts and abstract moments, as far as all this is given ... has become transparent (p. 68).

Description, any text where a phenomenon is illustrated or characterized, must be studied in a slow, careful, even painstaking manner—involving multiple re-readings and interpretive efforts. The researcher must pause in order to "dwell with" (Giorgi, 2007, p. 132) this description, to become fixated on the phenomenon of concern and the ways it can be characterized—all while vigilantly refusing to ascribe it any easy label, explanation, or as Fischer says, any "genesis" or "origin" (Fischer, 1914/2021, p. 66). The point is to sustain and vary this examination until the specific phenomenon's aspects, structure, components and moments have become transparent (or today: as transparent as they can be). Fischer illustrates through a very brief but pointed example: An older boy mockingly imitates his father. His father, catching sight of this, shouts angrily: "Stop that!" (Fischer actually describes the father angrily slapping the child. This has been changed to keep the example in the pedagogical [rather than, for example, in the legal] domain.) Such a simple description, Fischer points out, can be interpreted *psychologically*, in terms of the affective dynamics of

the father's anger and the boy's mockery. One can also imagine it interpreted sociologically, in terms of different cultures of parenting (re: the father) and socialization (re: the boy) existing among different social groups.

But it can only be regarded *pedagogically*, Fischer emphasizes, if it is seen in terms of the influence of one party on another. Generally this is the influence of the one who "educates" and the one being educated or "brought up" (p. 58): The father influences the boy, for better or worse, by shouting "Stop that!" Although one's first instinct is likely to judge the father for his angry outburst, this reflex must be avoided at all costs. To simply evaluate a child's and especially an educator's actions as either good or bad is not *Deskription*. Instead, Fischer sees in this phenomenon something (still) widely ignored in educational discourse: This is a "deeper layer" of pedagogical influencing that underlies explicitly deliberate actions. It is influence that is *unintended, accidental, contingent* and *circumstantial*, influence that can never be planned in advance according to certain curricula or theories. Fischer's discovery of this unplannable dimension—which was enabled by refusing psychological and other theories, and by focusing on "the things themselves"—set the stage for further phenomenologically-oriented research of pedagogical phenomena. This includes the first book-length investigation of *Pedagogical Tact: A Study of a Contemporary Form* by [Jacob Muth \(1967/2021\)](#).

Otto Friedrich Bollnow (1903–1991), as a second example, was deeply influenced by Heidegger after attending Heidegger's seminars in the 1930s. Although Bollnow generally did not focus on questions of methodology in discussing the relation of phenomenology and pedagogy, in a talk to teachers given late in his life ([Bollnow, 1989/2022](#)), he outlines a number of descriptive and interpretive "exercises" to direct these teachers' own practical reflections. Like Fischer, Bollnow begins by denouncing the difficulty that so much abstruse and prescriptive theory has presented for teachers. He says he is instead seeking a kind of theory that is very different from technical approaches. Technological understandings of pedagogy, he adds, are developed only to optimize education as what he calls a type of "production." Bollnow's point is instead to develop theory that is *reflective* rather than prescriptive, that does not boldly attempt to *lead* practice, but to more humbly *follow* it. As with Fischer, the uncovering of practice for Bollnow begins with description as a form of "closer 'seeing,'" "a step out of one's usual way of dealing with things" (p. 6).

The next stage in Bollnow's exercises is hermeneutic in nature, involving understanding or interpretation—i.e. the "grasping [of] something as meaningful:"

Understanding is co-original with life itself. But where everything is always-already understood [i.e. in the natural attitude], no interpretation is required. Rather, interpretation becomes necessary only where our usual understanding of life is disrupted by something new, something that does not fit. It is here that reflection becomes necessary.

([Bollnow, 1989/2022](#), p. 5)

Bollnow explains that this sense-making "means: to grasp something initially incomprehensible as a meaningful part of a larger whole" (p. 3). Hermeneutics involves an iterative, circular movement from part to whole, from past to present—from understanding that is only partial to understanding that is ever more (but never fully) complete. Because Bollnow is dealing with specifically *pedagogical* interpretation, he emphasizes that the "whole to which all individual phenomena are to be connected as meaningful is the *human being*; more precisely, it is *the person to be educated*" (pp. 5–6, emphases in original). This insight and the challenge it represents for reflection and interpretation can be seen as tantamount to a specifically "pedagogical" reduction for a particularly *pedagogical* phenomenology. It outlines the interpretive stance to be adopted by the practitioner and also by the researcher once the phenomenon in question has been stripped of theoretical, practical and ethical interpretations and judgments.

What Bollnow sees as emerging from this careful reflection and reduction is a type of "theory" that is not about optimizing educational production, but about the reflective affirmation of the responsibility of the educator to the "whole person" of the student:

This reflection is the origin of educational theory. To put it another way, such theory does not spring from an uninterested "pure" desire to know. It arises from the needs of practice, because the emerging difficulties force the educator to reflect, and to reflect in the original sense of the word [meaning: reflected illumination] ... It develops and expands the horizon in which the educator understands [their] actions[, giving them] a more liberated perspective and greater sense of security in teaching.

([Bollnow, 1989/2022](#), pp. 3–10, emphasis in original)

Contemporary phenomenological methods in education represent ways of taking the results of this type of reflection and giving to them the status of formalized, trustworthy, qualitatively generalizable research findings.

Phenomenology and methodology: Amedeo Giorgi and Max van Manen

In English-speaking academia, particularly in North America, methods for undertaking phenomenological investigations—suitable for proposing, planning and undertaking empirical research—began to be formulated by researchers and methodologists in the 1970s and 80s. These include the "transcendental," "descriptive" and "scientific" phenomenologies of psychologists Clark Moustakas, Paul Colaizzi and Amedeo Giorgi (respectively) as well as the "hermeneutic" phenomenology of educationist Max van Manen.

The last two figures, Giorgi and van Manen, are the focus here. Giorgi's scientific psychological phenomenology exhibits an explicitly Husserlian influence, while van Manen's hermeneutic phenomenology, integrates a range of influences, but sees its ideal in the work of the aforementioned Utrecht School (e.g. van Manen and van Manen, 2021). Giorgi's method is developed in and for psychology; the early articulations of van Manen's phenomenology are explicitly related to pedagogy.

A critical evaluation of both Giorgi's and van Manen's techniques in the light of contemporary qualitative methodology shows their phenomenological methods to be ones whose data gathering is typically interview-based (involving small sample sizes typically half a dozen or more), whose data analysis and method is iterative, emergent and adaptable, whose results are developed recursively and presented thematically, and whose qualitative validity and generalizability rest like perhaps no other method on the researcher's and reader's self-awareness and reflexivity. Investigators adapting either of these methods would begin by asking what a particular kind (or object) of experience is like. In the case of Giorgi's method, this object or phenomenon would be regarded as a psychological one (e.g. jealousy or trauma), for the early van Manen, it would be pedagogical (e.g. teachable moments, caring for troubled youth).

Amadeo Giorgi: "scientific" psychological phenomenology

Amedeo Giorgi (1931-) describes his phenomenological approach—also known as the "descriptive phenomenological method in psychology"—as "follow[ing] Husserl as strictly as possible: it uses intuition, it is descriptive, it acknowledges eidetic findings and it employs the ... reduction along with imaginative variation" (2007, p. 65). Given that Husserl saw himself as developing a new "science" (*Wissenschaft*) in his phenomenology, Giorgi sees his own research method as similarly "scientific" in nature, as producing "factual" and "objective" interpretations or findings (2009, p. 80).² "The researcher begins by obtaining concrete descriptions of experiences from others who have lived through" them, Giorgi explains. Such a sample may involve three or more participants, and the descriptions they provide are seen as being reflective of "the natural attitude" (p. 96). Like other phenomenological research methods, Giorgi's emphasizes that "it is theoretically desirable to have participants remain naive with respect to phenomenological concepts and theories" (p. 99). This means that member checks are generally *not* a recommended technique for increasing accuracy and internal validity for phenomenology. Recorded descriptions are first transcribed and then read by the researcher as a whole. According to Giorgi, references to the first person in these records are to be consequently rendered in the third person. Transcripts are then marked (quasi-arbitrarily) according to "transitions in meaning" to identify individual "meaning units," used as a heuristic for the reductions that follow.

One reduction in particular constitutes "the heart of" Giorgi's "method." It is the "scientific phenomenological reduction" (p. xiv) or what Giorgi calls "description." This refers to

The transformation of participants natural attitude expressions into phenomenologically psychologically sensitive expressions ... The researcher starts interrogating each meaning unit to discover how to express in a more satisfactory way the psychological implications of the lifeworld description (2009, pp. 130–130, emphases in original).

These psychological meanings are gradually developed through successive interpretations and refinements of the data. To bring these meanings into view, Giorgi explains that the researcher must take on a "psychological perspective ... a generic, a-theoretical psychological attitude such as is often assumed by practicing clinicians and therapists every day" (p. 135). "The psychological attitude," Giorgi continues, "focuses ... on how individual human subjects present the world to themselves and how they act on the basis of that presentation" (p. 135)—with special emphasis on feeling, sensing and emotion. Through repeated interpretative reformulations, the researcher gradually develops more general and invariant meanings from the data. Table 1 shows the application of this process to a study of the phenomenon of jealousy, featuring one meaning unit in which Paula, a grade school student, observes

Table 1 The development of invariant meanings from third-person transcript data.

Meaning unit	First transformation	Second transformation
"She can't stand it. Why can't that be her? It is all Paula can do not to cry."	"Paula can barely watch the scene. She wonders why [she] cannot ... be the center of attention. It takes all of her power for Paula not to cry."	"Paula is near despair at that moment, and it takes all of her strength not to cry. She wishes with all her heart that she could be the center of such attention and she tends to emphasize the feeling of loss she feels vis-a-vis the admired one who has it all."

Adapted from Giorgi (2009, p. 164, used with permission).

²Ignoring a vast range of recent developments in the natural sciences—from quantum physics to the replication crisis—Giorgi states that "in the 21st century, advances in scientific knowledge are still pretty much determined by the same criteria as those over the last two centuries" (2007, p. 3). Contrary to Giorgi, then, phenomenology cannot (and should not) claim to be scientific in the same sense(s) that the natural sciences do.

one of her classmates as being the center of attention. Giorgi's scientific phenomenological reduction, perhaps ironically, often involves the *addition* of increasing contextual information to both draw out and heighten the subjective and, in this sense, psychological aspects of the transcribed "meaning unit."

The next step involves reviewing these transformations and working through further adaptations of Husserl's reductions: The researcher must bracket previous knowledge of the phenomenon and also undertake "free imaginative" or eidetic variation:

Free imaginative variation requires that one mentally remove an aspect of the phenomenon that is to be clarified in order to see whether the removal transforms what is presented in an essential way. If the given appears radically different because of the removal of a part, it [the part] is leaning toward being essential (p. 69).

In the case of Paula, the grade school student (Table 1), one could say that there are at least five essential elements: (1) Paula, (2) her feelings of despair in relation to (3) another who is at the center of (4) the admiring attention of (5) peers or others. Arguably, no one of these five elements can be removed without the phenomenon of "jealousy" losing its essential character. For example, if Paula shared in the admiration of others or was herself the center of attention, she would not be jealous; alternatively, if the attention in the scenario were negative, disapproving, Paula's feelings might turn to pity or *schadenfreude*. The phenomenon, as Giorgi puts it, should constitute a structure with neither too many nor too few elements (p. 103).

The ultimate goal of all of this stage is to develop a structure that "comprehend[s] all of the data, even though [it may be] supplied by different individuals [or participants]" as a single whole. In the fifth and final step, this essential structure "is then used to help clarify and interpret [further] the raw data of the research" (2014, p. 6). Analysis of these elements in various configurations, either possible or real, leads Giorgi to insights such as: Jealousy "produce[s] a reduced sense of self in the jealous person. It is as though good things cannot be shared" (p. 165). Or: "The most negative effect of the living of jealousy is actually with the relationship of the experiencer with herself" (p. 164).

Finally, although Giorgi originally developed his method in the field of psychology, he emphasizes that it is ultimately applicable to "any other human or social science The only difference is that one assumes the attitude of the discipline within which one is working: *pedagogical* if it is *pedagogy*, sociological if sociology, etc." (2012, p. 11, emphasis added). It is precisely such a pedagogical disciplinary "attitude" that is outlined by van Manen in his own *early* pedagogical phenomenology.

Max van Manen: hermeneutic pedagogical phenomenology

Throughout his career, van Manen (1942-) has consistently characterized his phenomenological method as a "hermeneutic phenomenology." The overview provided here begins with his most exhaustive articulation of this hermeneutic method, provided in *Phenomenology of Practice* (2014); then, it turns to his earlier discussions of a particularly *pedagogical* phenomenology in *Researching Lived Experience* (1997).

In his book from 2014, van Manen characterizes his approach as an "abstemious reflection" that has as its goal an explicitly foundational "primitive or originary contact with the primal concreteness of lived reality" (pp. 26–41). Following Husserl, van Manen sees this reflection as abstemious or methodical in that it involves the *epoché* or reduction in all of its original three facets. Van Manen introduces myriad methodological constructs to support this careful reflection, including multiple permutations of Husserl's "phenomenological attitude," and of Husserl's eidetic reduction (pp. 215–239).

If the key to Giorgi's method is the iterative and eidetic transformation of "meaning units," then the heart of van Manen's is *writing*. Descriptions and short accounts ("anecdotes" or "examples") for van Manen are first gleaned from interviews and more broadly "from literature and the arts, from life, and from the imagination" (2014, p. 248). These then serve as the basis for the work of "reflective writing" which involves careful and repeated revisions both of the researcher's interpretive texts as well as of the descriptive accounts themselves. According to van Manen, these descriptive accounts can be "heightened" and extended in ways similar to Giorgi's "transformations." In a key passage from *Phenomenology of Practice*, van Manen explains that such imaginative but simultaneously sober writing represents *the* central challenge of the phenomenological research process:

The difficulty of writing has ... to do with two things: First, writing itself is a reflective component of phenomenological method. Phenomenological writing is not just a process of writing up or writing down the results of a research project. To write is to reflect; to write is to research. ... Second ... writing a phenomenological text is a reflective process of attempting to recover and express the ways we experience our life as we live it—and ultimately to be able to act practically in our lives with greater thoughtfulness and tact (p. 20).

Over 15 years earlier, in his 1997 *Researching Lived Experience*, however, van Manen characterizes phenomenology and its goals rather differently: He makes it clear that writing is *not* simply a way to reflect about just *any* given phenomenon. He also emphasizes that phenomenology leads to something more than the expression of "the ways we experience our life." Instead, like Giorgi (2012), the early van Manen insists that to reflect on "a phenomenon always implies a particular interest, station or vantage point in life" (1997, p. 40). Speaking of a specifically *pedagogical* orientation or interest, van Manen adds: "We are not simply being pedagogues here and researchers there—we are researchers oriented to the world in a *pedagogic* way ... aim[ing] for the strongest *pedagogic*

interpretation of a certain phenomenon" (p. 151; emphases added). The early van Manen thus not only emphasizes the importance of a pedagogical stance, he also sees phenomenology as "producing" specifically "pedagogic understandings, interpretations, and formulations" (pp. 151–152). Additionally, van Manen clearly comes down on the *educator's* side of the "pedagogical difference" mentioned in this chapter's opening (see also: Prange, 2005, pp. 58–59). Pedagogical experience, in other words, has characteristics of its own that can be interpreted and analyzed, and this experience (at least for van Manen) is primarily the experience of the educator rather than the one being educated.

Phenomenological writing must not only be oriented to a particular interest or orientation for the early van Manen; its particular *pedagogical* orientation itself must also be "strong," "deep," and "rich." "A strong pedagogic orientation requires that one reads any situation in which an adult finds himself or herself with a child as a pedagogic situation, as an answer to the question of how we should be and act with children" (1997, p. 152). This question of how we should be and act itself represents a kind of "pedagogical reduction" like that identified earlier in Bollnow. The *richness* of this orientation is reflected through vividness and detail; it renders this orientation explicitly generative, capable (as quoted above) of "producing" specifically "pedagogic understandings, interpretations, and formulations" (pp. 151–152). A deep text is one that is ultimately inexhaustible (p. 153). For example, the description provided by Fischer, above, may be seen to display "depth" in its disturbing, layered resonance, while obviously *not* being particularly "rich" in the sense of the "concrete ... thick description" recommended by van Manen (p. 152).

The challenge in taking up a pedagogical perspective, of adopting a specifically pedagogical "attitude," van Manen explains, is that "pedagogy announces itself not as entity, behavior, feeling, or emotion but *through* them" (1997, p. 50, emphasis in original). Pedagogy is available to us through a general *mode*, via a set of *ways* in which we experience the world. It is expressed concretely in terms of the four fundamental dimensions or conditions of our being outlined above: (1) the body; (2) the space our body inhabits and in which it expresses itself (cf. Merleau-Ponty); (3) time (as the meaning of our *Dasein*, cf. Heidegger), and (4) relation which (as Heidegger maintains) can be disclosed through embodied *attunement*.

Van Manen outlines the last of these, the dimension of lived relation, specifically in terms of "the pedagogical relation." This is a special relationship of influence between the adult and the young person, undertaken for the sake of the younger and (ideally) entered into freely by the younger:

The pedagogical relation differs from ... everyday ... relations in that the pedagogue reflectively mobilizes his or her conscious desire and will to give direction and shape to [the pedagogue's] influence. But only when such intentions are responded to by the young person can the pedagogical relation come into existence (1991, pp. 73–74).

Van Manen also suggests that this relation is experienced by the educator or adult in terms of *care*: "I care for this child—for whom this child is and who this child may become" (1991, p. 74).

The tension between caring simultaneously "for whom this child is" and "may become" delineates the central characteristic of the educator's lived *time*: It is divided between the child's or young person's *current* happiness and struggles, and (generally, adult) hopes and plans for their *future*. This tension becomes explicit when educators find themselves saying things like "it's for your own good" or "no pain, no gain." Again taking the side of the educator, van Manen also emphasizes the experience of *hope* in this specifically pedagogical temporality:

the sense of lived time [is experienced] in the modality of hope which I cherish for my child's happiness and becoming. And we experience a sense of lived time in the child's desire to become someone himself or herself, to live for something and to create personal meaning in life (1997, p. 105).

Space and embodiment are ultimately inseparable in pedagogy. Like lived time, they are also marked by a dichotomous tension: As Jacob Muth (in his 1967 study on tact) describes it, this tension exists between "the educative help of the teacher" and "the possible self-help of the child" (2021, p. xx). The task of the educator is thus to carefully negotiate proximity and distance—whether embodied and literal or figurative and metaphorical—and to find "the correct middle point" within it: "to step back whenever possible, but remain available when things turn problematic" (van Manen, 1991, p. 162), to create a "space in which the young person is enabled to make decisions and act in his or her own way" (p. 162).

Experienced pedagogically, time, space, relation and embodiment show that the pedagogical attitude is often experienced precisely in terms of tensions and opposed possibilities: We are asked to intervene in the space between the child's present reality and their future possibility; to give them freedom but often by temporarily limiting their choices. Additionally, balancing the developing individuality of the child with collective expectations—or exposing the child to the world while still protecting them from it—highlight still further balancing acts for the educator. The adult lifeworld of pedagogy is consequently characterized by concern about straying too far on one side or another, as van Manen suggests:

[A]s soon as we gain a lived sense of the pedagogic quality of parenting and teaching, we start to question and doubt ourselves. Pedagogy is this questioning, this doubting. We wonder: Did I do the right thing? Why do some people teach or bring their children up in such a different manner? (1997, p. 147).

In other words, pedagogy is explicitly and inescapably *ethical*. Ethical questions, uncertainties and equivocations are central to its disciplinary, professional and practical perspective. Although a phenomenological inquiry, as noted above, must begin by bracketing precisely the ethical question of “balance” or of “doing the right thing,” questions of right or wrong, of better or worse, nevertheless return. However, in phenomenological inquiry, these questions come into focus only *after* the researcher has engaged in rigorous description or writing, *after* they have bracketed not only ethical judgments, but also theoretical explanations and simple practical solutions. These questions return only *after* engagement in eidetic variation in which the researcher (or practitioner) has explored multiple situations and their variation (e.g. of an educator’s intervention or non-intervention in a specific set of circumstances). Finally, it comes only *after* one has in this way developed a broader awareness of the contingency and complexity of pedagogical action that goes *beyond* any one situation.

As Bollnow puts it, description, reduction and hermeneutic reflection are all means through which “the educator ... develops and expands the horizon in which [he] understands his actions” (Bollnow, 1989/2022, p. 14). By “offering them alternatives,” the work of description, (eidetic) reduction and generalization all allow “educators to take full[er] responsibility for their actions,” Bollnow explains (p. 13). These alternatives provide the opportunity of understanding problems and phenomena not just as they arise in one situation or classroom, but in their broader pedagogical generality. It is in working with examples, in seeing the meaning of varied phenomena coalescing to form the (still ultimately unattainable) “whole” of the child or young person, that one engages with both the challenge and the value of a pedagogical phenomenology.

Bollnow further describes this heightened awareness of both the possibilities and generalities of pedagogy as a kind of “tentative theory” in the service of practice. Husserl before him suggested the term “regional ontology” for a similar type of understanding gained through patient phenomenological investigation. For Husserl and Bollnow, the theory or ontology of the domain in question is one always to be tested and reconsidered in the light of new experience and reflection. As hermeneutic knowledge cycles from present to future, from action to reflection, from pre-understanding to new understanding, it remains forever open to new experience and new insight, meaning that the task of research in pedagogical phenomenology (as for any phenomenology) is always ongoing.

Giorgi, van Manen, and contemporary qualitative research: an evaluation

Despite Giorgi and van Manen’s articulations of not just a method but also of a professional and practical standpoint for their psychology and pedagogy, their respective treatments of the validation of the results of such research remain incomplete. Both tend to look only to philosophical phenomenology for criteria; neither offers ways of validating research results that reflect the language of contemporary qualitative research. Moreover, both repeat Husserl’s error of understanding the potential of their respective methods in positivistic terms of researcher *certainty*. By insisting on the “factual” and “objective” nature of their findings, on “originary contact with the lived reality,” both Giorgi and van Manen (like Moustakas and others) appeal to positivistic certainties to justify their method, much in the same way that Husserl did over 100 years earlier. Such language is likely to introduce difficulties for those defending their research, since it contradicts an established post-positivist or interpretivist consensus in qualitative research. This consensus understands that in any science or methodical investigation (from particle physics through ethnography), all research is based on and can produce only probabilities and qualified findings, not final or objective truths or primal certainties.

Further, if (pedagogical) phenomenology is situated in a post-positivist frame, it can readily adopt common qualitative techniques for maintaining and enhancing validity, reliability and generalizability. Internal validity (or credibility) becomes a question of the quality of the broadly scientific rigor and of writing so valued by Giorgi and van Manen (respectively). It can be judged logically and hermeneutically within the research presentation itself by looking for unacknowledged contradictions, tensions or (in) consistencies in the data or in structures, themes and interpretations. External validity can be evaluated in relation to relevant, related or similar work, whether empirical or philosophical. Both the “phenomenological attitude” and the disciplinary, practical or professional positioning of the researcher in a specifically *pedagogical* phenomenology would also need to be explicitly stated and acknowledged in such research. This could occur in a statement of positionality, and also explored through reflective researcher journaling.

Both van Manen and Giorgi recommend working with others who share the researcher’s phenomenological attitude—if not also their pedagogical and psychological position. Giorgi suggests the involvement of a “critical other” in the process of transforming interview data—one who “should also be able to access the transformations lived through by the primary researcher” (2009, p. 105). Van Manen implies something similar in the form of “collaborative reflection” in a seminar or other small groups (2014, p. 312). Although van Manen and Giorgi do not put it this way, constructive, collaborative feedback on one’s phenomenological text can ultimately contribute to what has been called “naturalistic generalization”—“a reasoned judgment” on the part of informed readers regarding the extent to which findings might apply to other situations (Brinkmann and Kvale, 2009, pp. 260–265). To read another’s phenomenological text is to undertake reflection and research in a way that parallels writing such a text oneself and takes the form of the evaluation of the degree to which themes, structures, and other conclusions “generalize” to the critical reader’s own situation, experience and knowledge.

“Scientific” or “hermeneutic” phenomenological research can also simply be evaluated according to the most general criteria for qualitative research:

Hallmarks of high-quality qualitative research include producing a rich, substantive account with strong evidence for inferences and conclusions and then reporting the lived experiences of those observed and their perspectives on ... reality, while recognizing that these could be multiple and complex and that the researcher is intertwined in the portrayal of this experience.

Cohen and Crabtree (2008, p. 334).

The terms and criteria of qualitative research are much closer to those of phenomenological research than its own methodologists appear to recognize.

Before concluding this section, two additional approaches, designated here as “quasi-phenomenological” should also be mentioned: Jonathan A. Smith’s “Interpretative Phenomenological Analysis” (IPA) and Ference Marton’s “Phenomenography.” Both are psychologically oriented and are arguably based on Husserl’s notion of phenomenology as the study of the “contents of consciousness.” Both focus not so much on direct experience, but on mental constructions and “sense-making,” asking not what a *lived* experience or an *experienced* object is like, but rather about how accumulated experience is viewed, understood or “made sense of.” Together with Flowers and Larkin, Smith defines IPA as “a qualitative research approach committed to the examination of how people make sense of their major life experiences” (2009, p. 1). Phenomenography, on the other hand, is differentiated from phenomenology as follows: “whereas in a phenomenological study the phenomenon per se is investigated, in a phenomenographic study the researcher investigates how (a group of) people *view* or *understand* the phenomenon” (Larsson and Holmström, 2007, p. 62). Phenomenography will ask how a participant “understands” something like the steps in a complex task or a principle from a high school physics textbook. IPA, on the other hand, asks questions, for example, about how “gay men *think* about sex and sexuality” or “how ... HIV impact[s] one’s *sense of self*” (Smith et al., 2009, emphases added). Phenomenology would engage with these matters quite differently—asking after gay men’s concrete experience of sex or sexuality, or, say, about what one undergoes experientially in completing a difficult task.

Pedagogical phenomenology: possible futures

Phenomenological research, like the human phenomena it studies, is unavoidably historical in nature. Pedagogy itself, for example, has not always entailed a holistic focus “on the person to be educated;” nor is it limited today simply to the influence of the older generation on the younger. Populations continue to diversify both inside and outside of the classroom, ensuring that experiential perspectives do indeed become increasingly “multiple and complex,” as Cohen and Crabtree have put it. For example, with the young teaching their elders (e.g. about new technologies) and adults engaged in “life-long” learning, the phenomenological study specifically of the experience of *learning*—not just of the adult educator’s side of the “pedagogical difference”—has become commonplace.

Like the tension-laden lived time, space and relationality of the educator, the investigation of “learning-as-experience” also begins through reference to Husserl, Heidegger and Merleau-Ponty. Moments of learning are understood in terms of the broken expectations and disappointments that Husserl saw as unavoidable in experience: “perception brings something new ... something [which] can be a slap in the face to all expectation” (2012, p. 263). Such moments can move us from un-self-aware “pre-reflective” experience to experience that is *explicitly* reflective, meaning that we become differently aware of ourselves, of others, of what we are experiencing *as such*: “Only in experience, thus turning back on itself, which at the same time is a change in our capacity to experience, lies the actual educative power of experience” (Buck, 1989, p. 3). Experience of this kind arguably unfolds in any learning but is particularly characteristic of the “transformative” learning that occurs through crises and discontinuities (e.g. Mezirow, 1991; English, 2014). An educator’s role in this context is not simply to optimize, to constantly facilitate such learning, but to be sensitized to momentary, individualized manifestations, and to provide the patient support that such reflexive and destabilizing experiences require.

This understanding of “learning-as-experience” is just one example of a kind of broadly pedagogical phenomenology as it continues to expand and develop—both in response to new types of experiences in education, and also to new ideas in the philosophy of phenomenology itself. Over more than 100 years since Husserl’s early philosophy and Fischer’s descriptive method, phenomenology has shown itself to be especially suited to pedagogy in its various permutations, particularly to its intentional and interpersonal emphases and its intrinsic flexibility, generativity and subjectivity.

References

- Bollnow, O.F., 1989/2022. Theory and practice in education. *Phenomenology & Practice* 17 (1).
- Brinkmann, S., Kvale, S., 2009. *InterViews: Learning the craft of qualitative research interviewing*. SAGE.
- Buck, G., 1989/2019. *Lernen und Erfahrung*. Springer.
- Cohen, D.J., Crabtree, B.F., 2008. Evaluative criteria for qualitative research in health care: Controversies and recommendations. *Ann. Fam. Med.* 6 (4), 3312–3339. <https://doi.org/10.1370/afm.818>.
- Dall’Alba, G., 2009. *Exploring education through phenomenology: Diverse approaches*. Wiley-Blackwell.
- English, A., 2014. *Discontinuity in learning*. Cambridge University Press.
- Fink, E., 1959/2022. The questionableness of the modern educator. In: Friesen, N. (Ed.), *Tact and the pedagogical relation*. Peter Lang, pp. 117–136.

- Fischer, A., 1914/2021. Descriptive Pedagogy. https://www.academia.edu/49088505/Descriptive_Pedagogy_1914_.
- Friesen, N., Henriksson, C., Saevi, T. (Eds.), 2012. Hermeneutic phenomenology in education: Method and practice. SENSE.
- Giorgi, A., 2007. Concerning the phenomenological methods of Husserl and Heidegger. *Collection du Cirp* 1, 63–78. <http://www.cirp.uqam.ca/documents%20pdf/Collection%20vol.%201/5.Giorgi.pdf>.
- Giorgi, A., 2009. The descriptive phenomenological method in psychology: A modified Husserlian approach. Duquesne University Press.
- Giorgi, A., 2012. The descriptive phenomenological psychological method. *J. Phenomenol. Psychol.* 43 (1), 3–12. <https://doi.org/10.1163/156916212X632934>.
- Heidegger, M., 1928/1962. Being and time. Harper.
- Husserl, E., 1900/2001. Logical investigations, second ed.2 vols. Routledge.
- Husserl, E., 1913/2001. Ideas pertaining to a pure phenomenology. Springer.
- Husserl, E., 2012. Analyses concerning passive and active synthesis. Springer.
- Larsson, J., Holmström, I., 2007. Phenomenographic or phenomenological analysis: does it matter? *Int. J. Qual. Stud. Health Well-Being* 2 (1), 55–64. <https://doi.org/10.1080/17482620601068105>.
- Merleau-Ponty, M., 1945/2012. Phenomenology of perception. Routledge.
- Merleau-Ponty, M., 1964/1968. The visible and the invisible. Northwestern.
- Mezirow, J., 1991. Transformative dimensions of adult learning. Jossey-Bass.
- Muth, J., 1967/2021. Pedagogical tact (excerpts). In: Friesen, N. (Ed.), Tact and the Pedagogical Relation. Peter Lang, New York. pp. 85–113.
- Prange, K., 2005. Die Zeigestruktur der Erziehung. Ferdinand schöningh.
- Smith, S.J., 1998. Risk and our pedagogical relation to children. SUNY Press.
- Smith, J.A., Flowers, P., Larkin, M., 2009. Interpretative phenomenological analysis: Theory, method and research. SAGE.
- van Manen, 1991. The tact of teaching. SUNY Press.
- van Manen, M., 1997. Researching lived experience. Routledge.
- van Manen, M., 2014. Phenomenology of practice. Routledge.
- van Manen, M., van Manen, M., 2021. Classic writings for a phenomenology of practice. Routledge.
- Wolf, B., 2020. Atmospheres of Learning. Mimesis, Milan.

Qualitative research approaches and designs: discourse analysis

Lavinia Suciu, Politehnica University Timisoara, Timisoara, Romania

© 2023 Elsevier Ltd. All rights reserved.

This is an update of P. Nikander, S. Vehviläinen, Discourse Analysis, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 373–379, ISBN 9780080448947 <https://doi.org/10.1016/B978-0-08-044894-7.01521-9>.

Introduction	141
Discourse analysis as a qualitative research approach	142
The peculiarities of discourse analysis as discipline of study	144
Relevance of discourse analysis in terms of QUAL research	146
Conclusion	148
References	148
Books	148
Chapters in books	149
Articles in proceedings	149

Introduction

Discourse analysis is an effervescent field, which has undergone a rapid development, as a result of the convergence of influences coming from numerous disciplines: linguistics, anthropology, philosophy, and, more recently, communication, psychology, artificial intelligence. For this reason, the terms *discourse* and *discourse analysis* have different meanings, depending on the field in which they are used (Schiffrin et al., 2001). Recent and current studies reflect the interdisciplinary significance of this field of knowledge, attesting the coexistence of language sciences with the ones in the sociological, psychological, communicational area (Roventă-Frumușani, 2004; Dăncu, 2001; Schiffrin et al., 2001; Jaworski and Coupland, 1999; LeVine and Scollon, 2004). Our perspective on discourse analysis as a research method describes an epistemological framework for investigating discourse as a space of signification, the core of which is represented by linguistics (linguistic pragmatics) and complemented by theories derived from anthropology, sociology, psychology, communication (Suciu, 2005, 2014a,b, 2019).

The concept of *discourse* is shaped along different interpretive tendencies (such as, for example, the restriction to the oral aspect of communication, to face-to-face interaction as the equivalent of the concept of *speech*). From the *Dictionary of Linguistics and of Language Sciences* (1994), we note three uses of discourse, which are also signaled by Baylon (1991): speech, enunciation, rules for chaining the suites of sentences. Since the late 80s, the term *discourse* has incorporated the meaning of “speech in the situation”—in sociolinguistics and sometimes in pragmatics—and that of “text”—used in discourse analysis and in the grammar of the text (Baylon, 1991). From the various definitions attributed to the discourse, three categories have been highlighted where discourse is included: “(1) anything beyond the sentence, (2) language use, and (3) a broader range of social practice that includes non-linguistic and nonspecific instances of language” (Schiffrin et al., 2001: 1).

Since the above framing of discourse reveals its relationship to language, communication, meaning and context, the definition of the field of knowledge that deals with the study of discourse refers to a conceptual association where linguistic, sociological, rhetorical theories are found. Defined as the use of language for social purposes, discourse analysis updates the cohabitation of the epistemological paradigms of pragmatics (the new rhetoric, the theory of enunciation, the theory of language acts) and sociology (ethnomethodology, ethnography of communication, conversation analysis, sociolinguistics) (Schiffrin, 1994). Despite its heterogeneous aspect, discourse analysis finds its unity and consistency in certain points that the disciplines which contribute to its constitution have in common as follows:

- Conceptualization of the discourse as a collective construction;
- Intervention of social and cultural norms, which determine roles and relationships of the participants, as well as the content of the messages;
- Social and interactive character of language;
- Dynamics of the enunciation (Suciu, 2019).

We aim, in this paper, to provide a delimitative description of the discourse analysis, which brings together three perspectives: identifying its peculiarities as a qualitative research, the presentation of its specificity and the relevance of discourse analysis by highlighting its functioning in order to explain the current interaction in several fields. Therefore, in the first part, our approach focuses on defining discourse analysis as a qualitative research by comparison with the quantitative research. Also, the uses of discourse analysis as qualitative research in linguistics, rhetorical criticism, communication, and education are highlighted.

Focusing on the constituent theories and on the concepts used in discourse analysis, we aim to emphasize the role of linguistic analysis in understanding the social. From this perspective we draw attention to relevance of employing discourse analysis in communication sciences, educational settings and in cognitive psychology. Besides, considering discourse analysis as a qualitative

research in other fields, we aim to illustrate how it serves to explain the classroom interaction, so as to discover ways to increase the efficiency of the teaching-learning activity. The examples presented reveal how the discourse analysis helps to describe the interaction in school and to improve communication, the teacher-student relationship: "... insights from discourse analysis in schools can help to make them better." (Temple Adger: 503).

Discourse analysis as a qualitative research approach

Due to its characteristics, discourse analysis is considered a qualitative research approach and it is used in such fields as communication sciences, social psychology, language learning, translation (Schiffrin et al., 2001). Regarding its importance in the study of communication, for example, Tracy has noted: "An upshot of the disciplinary context within which discourse studies emerged is that discourse analysis in communication is conceived as a method of inquiry" (Schiffrin et al., 2001: 726). *The Dictionary of Qualitative Methods in the Human and Social Sciences* (Mucchielli, 2002) incorporates discourse analysis into the category of qualitative research, defined as "discursive and significant approach to reformulate, explain or theorize a testimony, an experience or a phenomenon" (Paillé, in Mucchielli, 2002: 33).

Emerged in the 60s in the Anglo-Saxon space, the term qualitative research designates "a methodological approach that resorts to an interviewing process or to a form of qualitative analysis" (Mucchielli, 2002: 55). Qualitative research establishes a relationship with reality, with the subjects of the research and with the scientific explanation. The roots of qualitative research may be identified in the founding disciplines belonging to the trunk of the social and human sciences, through which they are differentiated from the natural sciences.

The fundamental feature of the qualitative research is the comprehensive (subjectivist/interpretative) perspective, through which human actions are considered phenomena of the meaning. These phenomena can be understood through a specific effort, related both to the human nature of the researcher and to their own nature (through empathy). The interest of a qualitative research is in phenomena (observable events, existential experiences), rather than in the exact, precise language used by an individual in interactions. It provides a reformulation, an interpretation, a theorizing of the phenomena that are lived or observed. According to the qualitative research approach, a text, a discourse, a device, (generally a human creation/achievement) is regarded as a set of collective or individual actions or as intangible phenomena, as representations (e.g., cultural norms, ways of thinking), in order to extract the meaning for oneself and for others. The qualitative approach is not intended to attest a point of view, but to explain it or to argue it theoretically (a reasoned interpretation is proposed to the receiver, more than a proven one). In order to solve a research problem, qualitative approaches comprise mainly ideational operations, that is, techniques of data collection and processing, which are not quantifying, automatic operations. Qualitative researches provide a qualitative analysis of data, through which words are interpreted by means of words, by applying the resources of intelligence to the perception of meanings. The conclusion of this analysis consists of a story or a theory, not of a demonstration (Mucchielli, 2002).

Qualitative research is often presented in comparison with quantitative research. The latter uses quantitative methods of document analysis (generally texts), which are based on mathematical and statistical techniques (sampling, decomposition, numbering, encoding, comparison, linking) and is aimed at discovering the properties of the document. Identifying these properties allows for a commentary on how that document makes sense and creates effects in a particular context. In the qualitative approach, the utilized procedures are based on ideational (reasoning) operations performed on the elements of a document or on a set of documents and have as their objective the explanation of meaning (Mucchielli, 2002). The parallel reveals an opposition that, although considered artificial, is relevant because it emphasizes the distinction between the two epistemological and methodological poles, allowing the identification of the main features of qualitative research (Mucchielli, 2002: 56–57):

- It concerns comprehension, depth (as compared to quantitative research, which seeks control and amplitude and extension);
- The logic of discovery (unlike the quantitative research, which adopts a logic of verification);
- The importance of the context (the context is perceived: it is not posed or considered constant as in quantitative research);
- It is interested in local, circular and symbolic "causality" (as opposed to quantitative research that can establish "causal" relationships and correlations);
- It tries to understand the complexity of situations, in order to present it in detail (while in quantitative research complexity is sublimated or denied).

From a methodological point of view, the qualitative approach involves the use of one or more methods of data collection (e.g., observation, semi-direct interview, group interview, collection of organizational, personal, historical documents, etc., journal, artistic productions, videos, photos, etc.). In collecting data, the researcher is immersive in his/her research, establishes a personal and prolonged contact with the environment, with the people. In order to analyze the data, the investigator may choose among one or more methods (qualitative content analysis by theorizing, needs analysis, conversation analysis, etc.), but his/her involvement is also present at this stage, in that the research tool is represented by his/her own person (while in the natural sciences the instrument is external, for example, the microscope).

The object of discourse analysis, as a method of qualitative analysis is "the enunciative device that correlates a textual organization with a determined social place" (Charaudeau and Maingueneau, 2002: 43). Discourse analysis provides a link between language and social reality in order to explain, through comprehension, a human or social phenomenon. In analytical reflection, the role of context is decisive, primarily due to the integration of linguistic pragmatic theories, namely: the theory of enunciation

(Benveniste), and the theory of language acts (inaugurated by Austin and continued by Searle) and coupled with their interference with Perelman's new rhetoric, the analysis of conversation, and the ethnography of communication. Aiming at clarifying the use of language in concrete situations of communication, "for social, expressive and referential purposes" (Schiffrin, 1994: 339), discourse analysis reveals the functioning and effectiveness of discourse, in other words, the dynamics of interaction.

In discourse analysis, the function of language ceases to be limited to the transfer of information. It acquires the communication dimension, by taking into account the interhuman relationships. The connexion between the protagonists of a verbal exchange is part of the context and is inscribed in the language, a situation conceptualized by the subjectivity of language (Benveniste, 1966). The sender of a discourse uses the words not only with their referential function, of representing reality, but also in order to act on his/her interlocutor, which means changing, modifying an existing state of affairs, behaviors, ideas. Advanced by the pragmatic theory (language is an act), this conception prioritizes the intentional and actional function of language (Suciu, 2005, 2014a,b, 2019), allowing a complex description of social interaction. Participants create and transmit content based on their social role and status, on intention, on the social and cultural norms that work in that particular interaction. The discourse is thus perceived as a collective construction, whose semantic co-construction includes the following elements: the intention of the sender, his/her knowledge of the receiver and of his/her expectations, the socio-cultural framework of the interlocutors, their social context and identity, the relationship between them.

As interdisciplinary field, discourse analysis focuses on the comprehensive aspect, on the study of discursive practices of real communication situations and of the discursive genres generated by them and on the investigation of verbal interaction (Rovența-Frumușani, 2004). Discourse analysis incorporates different orientations—the most important being the theory of speech acts, the interactionist sociolinguistics, the ethnography of communication, the Gricean pragmatic approach, the analysis of conversation, the variationist linguistics—which have a common denominator although they use different research methods (Suciu, 2019). The stakes of this interdisciplinarity concern both the ability to understand the variety of discursive genres in modern and contemporary society and the ability to interact effectively with others, to diminish/annihilate barriers or social/political/cultural disadvantages (Rovența-Frumușani, 2004). Sharing the approach of researchers (Schiffrin, 1994; Rovența-Frumușani, 2004) who treat discourse analysis as an umbrella term, we consider that the analytical perspective proposed by discourse analysis is relevant because it:

- demonstrates how the linguistic fact is responsible for influencing/modifying the social fact;
- supports the role of linguistic analysis as an instrument of understanding the social, the relational fabric established in a concrete plan between individuals and/or groups having the status of social actors;
- reveals the functioning and effectiveness of a discourse and of the dynamics of interaction.

In the light of the above, we consider that discourse analysis is a useful qualitative research method in the study of written or spoken language in relation to the interactions taking place in the social context. The aim of the investigation is to decipher the meanings that escape immediate identification in texts/situations/processes. Depending on the field in which it is used as research method, discourse analysis has various connotations. The use of discourse analysis as a method of investigation in linguistics refers to the examination of a linguistic unit that goes beyond the boundaries of a sentence. In addition, it studies the way in which language is used socially. Discourse analysts in the field of communication are differentiated by the fact that they are concerned about treating the subject in everyday situations. In the field of communication, discourse analysis is the study of speech (or text) in context, where research reports use communication sequences and their analysis as central means of scientific argumentation (Tracy, 2001). The implementation of the method of discourse analysis in education aims to identify the specificity of the language in school, as it is related to the linguistic competence necessary for the student to obtain good results. The interest in the study of verbal exchanges with the help of discourse analysis also goes toward explaining the actions by which learning, as the main target of schools, is achieved (Temple Adger, 2001). In rhetorical criticism, unique political discourses and actions are studied within the theoretical framework offered by literary criticism, political and continental philosophy, history, film studies and so on (Tracy, 2001).

In order to reveal how the pragmatic theory in discourse analysis helps to identify the implicit meaning, we refer to an example that we have discussed on another occasion (Suciu, 2014a,b), leading to the discovery of the sense that is correct and adequate to the context in which it is issued. From the beginning, we should mention that, very often, in communication, the meaning of a message is not always limited to the meaning conferred by the sequence of words and phrases. In other words, the meaning of a message results also/especially from the manner in which it is uttered. The message consists of a sequence of sounds, words, phrases that have a certain meaning, forming the content of the communication or the information. It is accompanied in speech by a specific tone (serious, joking, derisive, etc.), by a mimicry, an attitude or a posture, which complete the literal meaning of the phrases, sometimes even modifying them. All these indicial elements have a role in establishing the register of interpretation of the message (whether it is a joke, a threat, an order, a suggestion, etc.). At the same time, they relate to the knowledge of the type of relationship between the participants in the verbal exchange and have to do with the utterance of the message.

According to the theory of utterance in pragmatics (Benveniste, 1966), the example *This Morning it has snowed* consists of the utterance, which represents the act/activity of a person who uses the language through an individual act of usage and a statement, which is the product of this activity. The statement *This morning it has snowed* is referential (a fact from reality is transposed into a linguistic form), and addressed by a sender to a receiver in a specific context (a certain moment, a certain place, there being a certain relationship between them). We note that the message contains the sender's underlying belief in what he/she says, but it is not expressed explicitly, by using the linguistic features (such as *I affirm, I am convinced*). The conviction, in this example, indicates

the attitude of the sender toward what he/she states and this conviction is contained in the meaning of the utterance, as we will discuss in the next section. If the statement is accompanied, for example, by the linguistic marker *be careful* and/or by the intonation corresponding to a warning, then the correct meaning of the message is one of warning. The message thus carries an indicator of the utterance and thus, it becomes:

This morning it has snowed, therefore you should be careful not to slip/leave home earlier than usual, since the traffic is hampered/put on your hat, etc.

Different features may be attached to the statement either explicitly, verbalized in the statement (*I say, I am convinced, I doubt, I do not think, I hope*, etc.), or non-linguistic and paralinguistic features (gestures, attitude, posture, mimicry, intonation). The indicators of the utterance perform the function of a code by which the statement acquires its complete significance. The precision of its meaning is given by the reference to the data provided by the context. In other words, in order for the receiver to get the exact and correct meaning of the message, he/she will act on the message through reasoning that involves reference to the context, restricting the semantic series of possibilities. For example, depending on the moment when it is issued, the indicator of the statement may be the warning (*you should leave home earlier, the traffic being hampered*), depending on the relationship of the participants, the message may be completed with the urge: *You'd better put on your cap* (if the message is sent by a mother to her child) etc.

Understanding the meaning of a message by using the linguistic features/markers, the paralinguistic features (such as intonation and tone of the voice) or the non-linguistic features (such as gesture) is easier than the case when they are missing (the case of written communication). Assigning a meaning to the message requires the identification of these features in order to deduce the intention of the sender. This purpose is fulfilled by some theories of pragmatics. Although these theories use different concepts, they cover the same object, namely the discursive principles/laws which, by being paired with the context data, help the receiver interpret the message and to reach its meaning.

We consider that, by appealing to theories from pragmatics and rhetoric, to sociological and anthropological theories, discourse analysis approach can provide a complex description and explanation of real life discourse and interaction. Thus, discourse analysis:

- addresses the comprehension of the communication situation in order to explain the meaning;
- goes into deeper layers of analysis, due to the interest in the "local" aspect;
- has the ability to reveal the complexity of situations, explaining them in detail;
- gives importance to the context in interpretation, which reflects the connection with the social aspect;
- determines the understanding of the social aspect, of the relationships established between individuals and/or groups as social actors;
- gives an interpretation that is updated in words;
- demonstrates flexibility and possibility of adaptation to the object of study because the research tool is intrinsic to the researcher;
- contributes to the improvement of the quality of interaction and to the efficiency of communication.

The peculiarities of discourse analysis as discipline of study

The following presentation focuses on highlighting the peculiarities of discourse analysis as a qualitative research, especially the correlation with reality, with the social, through the pragmatic conception that substantiates this field.

A first conception of discourse/communication is related to the customs in Ancient Greece and to the understanding of discourse as an instrument available to the Greek citizen in the form of a pleading. Delivered in public, discourse became a way to defend a personal interest or access a privileged social position. The art of communication (which was called *rhetoric*) developed as a means of persuading the audience, therefore, the interest in acquiring and studying it equaled the magnitude of the interest in the art and science of war, that had already been established at the time. Later, specialists in rhetoric appeared (Corax from Syracuse, Tisias, the Sophists) who developed elements of communication theory, representing communication methods and techniques imposed by practical requirements (to win litigations on wealth recovery, for example).

Plato conceived communication (and, implicitly, the discourse) as a sophisticated process involving not only the study of influencing and persuading techniques but also the study of knowledge, of the meaning of words or of human behavior and ways of life. Unlike Plato, for Aristotle rhetoric was a corpus of technical means to convince the audience and/or win a cause (Florescu, 1973). This definition of discourse implies the social aspect in addition to the communicative intention of the sender, both of them being specific to the pragmatic conception, that is predominant in discourse analysis, as well. Similar to ancient rhetoric, the pragmatic school of thought views discourse as a unit whose identity is shaped not only by its linguistic expression, but also by the assumptions produced by the receiver, relative to the speaker's (direct or underlying) intention and context. One of the points of interference between ancient and pragmatic rhetoric is the communicative intentionality of the speaker (Gross, 2000; Eco, 1996).

Discourse analysis is guided by the inherent junction of the language fact with the social fact, a foundation of the interactional theory which will be presented in what follows. Schiffrein (1994) says that discourse analysis involves six approaches that are considered dominant: the theory of speech acts, pragmatics, ethnomethodology, interactional sociolinguistics, ethnography of

communication, and variational sociolinguistics. The close connection between discourse and society has been supported over time by linguists, sociologists and anthropologists, philosophers, psychologists, mathematicians/computer scientists.

The interactional theory on communication claims the appeal to rules that do not belong to the linguistic code, but to the uses of language, updated in the verbal interaction, without a break from the linguistics of the language/code. Thus, the verbal interaction, viewed in a general way of functioning, is constructed and interpreted with the help of *rules* that are applied in a given contextual *framework*, on a material of heterogeneous semiotic nature (verbal, paraverbal, nonverbal units) (Kerbrat-Orecchioni, 1990). The interactionist theory is the premise for the development of a sociolinguistic research study on communication (ethnomethodology) and of pragmatic theory in linguistics. If the latter focuses on the study of language use, the socio-anthropological approach surveys the role that language plays in relation to society and culture (Suciu, 2019).

Defining language as a “guide to social reality”, in the first half of the 20th century, E. Sapir, updated a sociolinguistic conception of it (Sapir, apud Mounin, 1975: 82–97), of which we point out two significant aspects for discourse analysis. On the one hand, the value of language as a tool that mediates the relationship between the individual and the reality foreshadows the sociolinguistic theories of communication sciences, regarding the effect of communication on the individual and society. On the other hand, the introduction of the social dimension of language, which involves the psycho-social component, is later found in the semiotic and pragmatic theories of linguistics. Thus, the subjective and social stakes of language are converted, in discourse analysis, into issues such as: subjectivity of language, discursive genres and critical analysis of discourse, through the representative research of E. Benveniste, J. Austin, J. Searle, D. Hymes, T. van Dijk (Suciu, 2019).

The sociological perspective has as center of interest the role of language in culture and society and the act of communication as frame of reference. The study of language in its social context, sociolinguistics aims to highlight the relationship between the organization of the message, as a premise of linguistic analysis, and the social implication of this message (Suciu, 2014a). The study of language in sociocultural context branches off into three directions according to the importance of either the social aspect or the linguistic aspect: *the ethnography of communication* (where the social aspect is a priority), *variationist sociolinguistics* (the social aspect is integrated into the linguistic one) and *interactional sociolinguistics* (the social and the linguistic aspects have equal weight).

Pragmatics explain the relationship between the linguistic form and its user (the sender), how language is used in communication. “Usefully defined as the study of how utterances have meaning in situations” (Leech, 1989: 2), pragmatics became a central theory in discourse analysis, which refused to accept the limitation of the meaning in the abstract. Studying the language in situation, (i.e., communication) and aiming at analyzing concrete verbal exchanges as verbal interactions, by referring to participants, use and contexts, pragmatics becomes one of the basic paradigms on which discourse analysis develops alongside with the structural—functional paradigm, the social/socio-linguistic and the critical ones (Baylon, 1991). “Discourse analysis must not only explain why certain statements and not others have been preferred, but also it should account for the way those statements have mobilized forces and influenced social networks” (Maingueneau, 1991: 35). The interactionist perspective in pragmatics is based on three aspects related to language: subjectivity, the implicit and the social, exploited differently in pragmatic theories. The subjectivity of language, being the concern of the stylistic direction, has been reaffirmed by Benveniste (1966), by his theory of utterance (fr.: *l'énonciation*). The theory, from which the *linguistics of utterance* develops is linked to the preoccupation of linguists for the communicative intention of the speaker, named *force meaning* (Thomas, 1995). The utterance indicates how the speaker uses the language in order to work out the intended force and to act on the receiver. Thus, the theory emphasizes the relationship between the speaker and the receiver through their interaction (Suciu, 2005).

The *implicit* is a concept belonging to the cognitive orientation and invested in conversational analysis of H. P. Grice. The *conversational implicature* is defined by Yule (2006) as an additional unstated meaning that the hearer must assume the speaker to convey more than is being said. It deals with the determination of the utterance's meaning and force through the operation of some communication rules (*cooperative principle*) and through the contribution of both, the speaker and the receiver. To the same extent, the implicit meaning serves the speech acts theory (to determine the indirect speech act) outlining the direction of research in analytical philosophy or pragmatics itself, which studies language as an instrument of action. The link between the concept of *discourse* to that of utterance reveals two aspects: the intentionality, to which subjectivity is related, and the implicit, to which the social is correlated (Suciu, 2014a,b). The speaker's intention to produce a reaction in the interlocutor's attitude (*the utterance force*), through the issued discourse entails its conceptualization in the speech act theory. The philosophers of language, J. Austin and J. Searle, conceive language as an instrument of action: the speech act is an action performed by the use of an utterance to communicate (Austin, 1970; Searle, 1972). In the view of M. Pêcheux, discourse analysis offers the opportunity to study the immediate communicative situations through the appeal to the statement—utterance relationship and to the pertinence of the discourse situations, as well as to the possibility to study the social and cultural situations within the pragmatic framework by approaching them with linguistic criteria (Suciu, 2019).

Taking examples from advertising discourse or signature discourse of the organization (Suciu, 2005), we observe that the meaning of the utterance—closely linked to the sentence—indicate the way in which the sender uses the language by an individual act that bears his/her imprint. The funny tone, as a mark contained in the statement, which is explicitly expressed in some advertising messages, imprints the message with a relaxed, familiar atmosphere, thus a reduced social distance between the speaker and the receiver; the sentence will be understood in this orientation of the meaning.

Example: Pringles. Once you pop, you can't stop.

Below, there are two examples of a signature discourse of the organization expressed in a serious, even sober manner (an indicator of the utterance), which make us think about the sender's position of authority (he/she states something being entitled to do so).

Tiriac Bank: *Look ahead!*

Euromedia: *Show up with the best!*

Their interpretation is made on the basis of a common rule according to which a statement that communicate an order or an interrogation implicitly contains the idea of the sender's authority (the sender has the appropriate position to issue an order or to interrogate). The assigned sense of a message is oriented by the explanation of the utterance's meaning, through the intervention of discursive principles in the interpretation realized by the receiver, starting from the literal form of the statement (Ducrot and Schaeffer, 1995). For example, an order/request requires its fulfillment, a question requires for an answer, etc. In conclusion, the receiver's interpretation is based on the discursive principles from the discourse analysis theories and the context data.

Therefore, in order to study the discourse/communication, we consider that the inclusion of pragmatics is profitable for discourse analysis, being supported by the following aspects: (1) the function of some theories favorable to the analysis of a discourse (the interactionist theory, the implicit, the theory of the acts of language); (2) the openness toward the social environment; (3) the inclusion of the argumentative perspective.

Timeliness confronts discourse analysis specialists with a new challenge. The digitalization of communication has instituted new interactive, social and communicational practices, some of them being quantifiable (for example, the frequency of posts). These have not gone unnoticed in what discourse analysis means today. Numerous studies have been devoted to the technical artifact involved in virtual discourse (Baron, 1984; Herring, 2001; Webber, 2001; Kress, 2003; LeVine and Scollon, 2004; Suci, 2014a,b; Suci and Mocofan, 2015; Suci, 2021, 2018, 2019) or digital communication (Walther, 1996; Griffin, 2005; Kasdorf, 2003; Weasenforth, 2006; Frascara, 2012; Duarte, 2008, 2013). In a first stage, concerns about the virtual discourse were directed toward a certain discursive genre, materializing in descriptions of some virtual communication systems, in empirical descriptions of language in online interaction or in studies focused on the variety of discourse in this environment. Subsequently, the interest of research focused on the influence that situational and technical factors exert on communication, generating its variety and complexity (Suci, 2019).

Observations of issues related to the discourse of the digital society, such as the impact of technology on society in terms of users' linguistic or nonlinguistic behavior, interactional and social practices of individuals (users) have led to significant changes in the theoretical framework of approach and interpretation of current discourse. The examination of virtual discourse involves reconsidering the position of the sender (to the text and to the receiver), the receiver (to the text), as well as the inter-relational context in a given communication situation. The technological aspect does not leave its mark on verbal language alone. It is known that a so-called *code*, specific to the electronic environment, exists and operates and, through it, any action taken in this environment (a virtual act such as creating, posting/blogging, activating in a social network, etc.) acquires a meaning (Bădău, 2011).

While reconsidering the discursive approach, discourse analysis confirms its capacity to adapt to the forms of expression that are manifested at present and, at the same time, foreshadows new tendencies in investigation and interpretation. The interdisciplinarity of theories and methods of examination, specific to the field of discourse analysis, in necessary and productive complementarity to environmental factors, is likely to generate an integrated investigative system, which is a generous source for developing strategies for successful creation and interpretation of digital communication (Suci, 2014a).

Relevance of discourse analysis in terms of QUAL research

The role of discourse analysis as a qualitative research in the field of education is demonstrated by several accomplished research projects centered on this subject. It is not uncommon for our scientific preoccupations to be directed toward explaining the linguistic and nonverbal behaviors through which learning, as the main school objective, is achieved. We have focused our attention on the interaction in the academic environment, in order to find ways to make the teaching-learning activity more efficient. We have highlighted discursive strategies able to facilitate learning and lead to a greater efficiency of the teaching process. Since achievement of interaction in the educational institutions is mostly linguistic, we used discourse analysis as research method. Nevertheless, we sometimes considered appropriate to complete this method with elements/notions from the communication sciences, psychology, sociology, education. Here are some of the papers published on this topic: L. Suci, *The Role of Communication in Building the Pedagogical Relationship* (2014b), L. Suci, S., Şimon, *Informing Strategies: building the communication framework* (2015), L. Suci, *In search of a balanced relationship in communication. Discursive strategies that create resonance* (2017); L. Suci, *Practices of interaction between the real and virtual worlds* (2018), L. Suci, M. Mocofan, *Approaches to hate speech in cyberspace. A Metadiscursive Analysis* (2021), L. Suci, *A Look Inside Academic Communication During the Pandemic* (2021).

Usually, the school interaction, based on the transfer of specialized knowledge is characterized by the presence of two types of barriers: the cognitive one and the cultural one. This gap, due to cognitive and cultural barriers, generates an unequal positioning of the interlocutors in the interaction—it is obvious that the teacher, as the holder/possessor of the information, is placed on a higher position compared to the student, whose baggage of knowledge and culture is lower. Their positioning involves relational imbalance and thus interaction is defined by social distance (Meirieu, 2010).

The power of words/communication is a subject that reveals two conceptions which, without excluding each other, are differentiated by placing the emphasis on another aspect that the sending of a statement implies: the fulfillment of the preconditions of the transmission, respectively, its effect. In the view of sociologists (Bourdieu, 1991), the power of communication lies in its

exterior, namely in the achievement of external conditions (regarding status, role, framework, etc.) for a sender/a speaker to be able to communicate (can only work if social conditions external to the linguistic logic of the statement are met). In other words, the sender/the speaker must be legitimate to communicate. Defining the foundations of social power, social psychologists French and Reaven (1959) show that the legitimacy of a sender is based on the understanding on the part of the interaction participants that the sender has the right to demand obedience. Linguistic activity is, from this perspective, one of the various means available to someone to exercise his/her power, depending on his/her social status.

On the other hand, researchers in pragmatics (Austin, 1970; Searle, 1972) consider that the power of communication is linked to the power of words, emphasizing their effect on the receiver. They believe that through words we fulfill an action, they try to prove *how to do things with words*. Based on the concepts of intentionality and action, pragmatics states the fact that in the linguistic expression (in the word) lies the force by which a sender can act on a receiver (Suciu, 2014a,b). Pragmatic theory implies the condition of satisfying the extralinguistic premises of issuance, expressed in the form of the principle of legitimacy (one of the discursive principles). For a statement to be accredited worthy of being believed and thus, followed by effect, it must be expressed by someone in the right to do so, by an authority (Suciu, 2005).

The superior position of the sender, which underlies the principle of legitimacy, refers to the idea of responsibility, seriousness, trust, authority that is entitled to demand obedience and which must therefore be listened to. Sometimes the legitimate status of the sender, which coincides with the declaration of the fulfillment of the upstream of the discourse, is explicitly reflected at the linguistic level of the statement, by inserting some morpho-lexical elements (the relative superlative) or rhetorical ones (antithesis) with the function of arguments of the legitimacy and credibility of the locutor. Omnipotence, special vocation, power become arguments of the sender's legitimacy (Suciu, 2005, 2017).

In school interaction, the teacher enjoys the privileged position in interaction, i.e., the position of authority. This position is based on the legitimacy conferred by the knowledge he/she holds, while the teacher-student relationship is marked by social distance. The teacher's verbal and nonverbal behavior implicitly transmits the meaning of his/her superior position as in these examples: *First document yourself, then we shall talk about it; I know everything, you know nothing!* where the existing social distance is emphasized. In the paper *The Role of Communication in Building the Pedagogical Relationship* we have highlighted some discursive strategies may reduce the distance between the participants and re-balance the pedagogical relationship, for example strategies of indirection (*wouldn't it be better to ..., couldn't you ..., what about ...*), interrogatives strategies (*Do you really think that ...*), impersonal strategies (strategies generating an impersonal message) (*it is adequate to ...*) etc. It is important to emphasize that the implementation of these strategies does not aim at eliminating the social distance, which proves necessary in this type of interaction, the teacher's role being to guide the student in the sphere of knowledge (Suciu, 2014a,b).

In the same register of avoidance to offend the sender (the student, in our case) there are also strategies of positive and negative politeness contained by discourse analysis (Kerbrat-Orecchioni, 1991; Roventă-Frumușani, 2004), for example:

- assertion of a common territory: *Don't you think it's ...*
- explicit formulation of observations attesting the interest, empathy of the sender toward the receiver: *You have to be very tired; You're bored to ..., are you?*
- minimization of criticism, and of the refusal: *I'm a little disappointed with your last work; That's pretty farfetched*
- inclusion of the receiver in the action expressed in the statement: *Let's see your work*

We also emphasize the contribution of the discourse analysis to the investigation of communication through the referential works of Leech (1983, 2006), Brown and Levinson (1987), Kerbrat-Orecchioni (1992, 1994) regarding politeness as a pertinent system of analytical reporting in approaching verbal interaction in various contexts.

Being a face-to-face interaction, verbal behavior, is accompanied by the nonverbal one. Therefore, in order to reduce the teacher-student distance, the contribution of nonverbal strategies is an advantage. In the book *Bridges across the communication* (2017) we analyzed the statement of a specialist in education and of a psychotherapist specialized in human relationships, each addressing a non-specialized audience. I revealed the function performed by mimicry, especially by the language of the eyes, by gestures in creating that particular thread that weaves between the interlocutors in interaction. Contact is established through eye language, attention is mobilized, a proposal for closeness is transmitted. We have found that at the level of the statement, language, style and humor act complementarily, creating a relationship of proximity and familiarity between the sender and the audience, despite the social distance existing in that interaction due to the authority of the sender (the specialist, the expert). These factors are related to the sender's ability to elaborate the discourse by applying linguistic strategies of rapprochement. Also important is the way in which the sender relates to the audience in the transfer of the discourse and the manner in which he/she transmits it.

Timeliness, with the advent of new communication technologies, launches another challenge that the teacher has to face. The study of verbal interaction in the academic environment during the pandemic (2020–2022) highlights aspects related to linguistic and non-linguistic behaviors in the teacher-student interaction carried out online/remotely, such as: replacing the formal framework of communication with the informal one, posting inappropriate photos as the interface at the meeting for the course or seminar on the online platform, low feedback, low interactivity (up to the refusal of interaction), ignoring some explicitly expressed rules regarding the carrying out of activities, neglecting the teacher's moment of interpellation. Trying to explain these behaviors, we have revealed the importance of academic spatial landmarks (the imposing building of the university, the class/seminar room, the hallways, the library with the shelves full of books, the reading room, etc.) in order to set up the common cultural background. The spatial landmarks function as marks/insignia of the place (the university), of the university culture, which compete to the protagonists' sharing of the cultural background, ensuring the transgression of the cultural barrier (Meirieu, 2010) and transmit

meanings that help the student to create a real and correct image of the academic environment. Thus, the student will be ready to assume the appropriate role, updated in the congruence between the discourse and the context (Suciu, 2022).

In conclusion, the affective dimension (understood as regulation of the circulation of affections, cf. Meirieu, 2010), is responsible for the reconfiguration of the discourse with the help of the linguistic and non-linguistic languages that give rise to sympathy strategies. Under these conditions, it is possible to establish the relational balance between the participants in the interaction, which ensures the transmission of information. Without canceling the social distance and the authority of the teacher, the highlighted discursive strategies determine the success of the communication in educational context. The analysis of communication in educational settings attests to the contribution of discourse analysis to the explanation of the interaction in school, to the creation of *a common ground with the student/teacher*, using its insights: pragmatics, anthropology and sociology. The results of this investigation with the help of discourse analysis are relevant in terms of achieving more effective communication between the teacher and the student, with impact on an improved student's performance.

Conclusion

In our approach, we have set out on an incursion into the field of discourse analysis, trying to define it as a qualitative research, by bringing together three angles of approach. First of all, starting from the peculiarities that define a qualitative research method, we have identified them at the level of discourse analysis. This frame of reference outlines the descriptive picture of discourse analysis as a qualitative research approach, by delimiting it from quantitative research. At the same time, by referring to the different uses of discourse analysis in other fields (linguistics, rhetorical criticism, education), we have emphasized that this analytic perspective of discourse study is helping to improve the communication and the activity in school. In order to reveal the meaning in texts/situations/processes, the example discussed reflects the role of the discourse analysis in communication and education.

The detailed presentation on the peculiarities of discourse analysis is intended to complement and support its definition. This perspective highlights the contribution of the theories involved by the discourse analysis (pragmatics, sociological and anthropological theories) to give a description of the discourse and of the dynamics of interaction in everyday reality. Through the examples we have explained the functioning of the discourse analysis in finding the implicit meaning of the discourse.

The relevance of discourse analysis as a qualitative research is addressed by reference to education. The premise is the existence of a cognitive and cultural gap in the teacher-student interaction, coming from the legitimacy of the teacher and the different cultural belonging of the participants in the interaction. In order to demonstrate the possibility of overcoming the cognitive and cultural gap, we have revealed several discursive strategies aimed at reducing the social distance between the participants in the discourse and at rebalancing their relationship. The importance of applying the strategies lies in facilitating verbal exchanges in the educational environment and, therefore, increasing efficiency in the teaching-learning activity. We believe that defining discourse analysis through the three perspectives—which are interconnected and supporting each other—confers individuality to this qualitative research approach that may serve to enhance the effectiveness of communication in several professional contexts.

References

Books

- Austin, J., 1970. *Quand dire c'est faire*. Seuil, Paris.
- Bădău, H.M., 2011. *Tehnici de comunicare în social media*. Polirom, Iași.
- Baylon, C., 1991. *Sociolinguistique Société, langue et discours*. Nathan, Paris.
- Benveniste, E., 1966. *Problèmes de linguistique générale*. Gallimard, Paris, 1966 și 1993.
- Bourdieu, P., 1991. *Language and Symbolic Power*. Polity Press.
- Charaudeau, P., Maingueneau, D. (Eds.), 2002. *Dictionnaire de l'analyse du discours*. Seuil, Paris.
- Ducrot, O., Schaeffer, J.-M., 1995. *Nouveau dictionnaire encyclopédique des sciences du langage*. Seuil, Paris.
- Eco, U., 1996. *Limitele interpretării*. Pontica, Constanța.
- Florescu, V., 1973. *Retică și neoretică: geneză; evoluție; perspective*. Academia Română, București.
- French, J., Reaven, B., 1959. *The Bases of Social Power*. <https://pdfs.semanticscholar.org/f00c/7fd21a9774865ef7a3ddf871aa731c84e295.pdf>.
- Gross, A., 2000. *The Marriage of Rhetoric and Pragmatics*. www.eserver.org/audio/Gross.
- Kerbrat-Orecchioni, C., 1990. *Les interactions verbales (I)*. Armand Colin, Paris.
- Kerbrat-Orecchioni, C., 1991. *L'implicite*. Armand Colin, Paris.
- Maingueneau, D., 1991. *L'énonciation en linguistique française*. Hachette, Paris.
- Leech, G., 1989. *Principles of Pragmatics*. <https://www.pdfdrive.com/principles-of-pragmatics-longman-linguistics-library-e161016717.html>.
- Mounin, G., 1975. *La linguistique du XXe siècle*. Presses Universitaires de France, Paris.
- Mucchielli, A. (Ed.), 2002. *Dicționar al metodelor calitative în științele umane și sociale*. Polirom, Iași.
- Rovența-Frumușani, D., 2004. *Analiza discursului Ipoteze și ipostaze*. Tritonic, București.
- Schiffrin, D., 1994. *Approaches to Discourse: Language as Social Interaction*. Wiley-Blackwell.
- Schiffrin, D., Tannen, D., Hamilton, H.E. (Eds.), 2001. *The Handbook of Discourse Analysis*. Blackwell Publishers Ltd., Oxford.
- Searle, J., 1972. *Les actes du langage*. Hermann, Paris.
- Suciu, L., 2005. *Discursul-semnătură al Instituției Miza unei identități și premisei unei relații*. Orizonturi Universitare, Timișoara.
- Suciu, L., 2014a. *În căutarea sensului. De la analiza discursului la design-ul comunicării*. Casa Cărții de Știință, Orizonturi Universitare, Cluj, Timișoara.
- Suciu, L., 2017. *Bridges Across the Communication*. Lambert Academic Publishing.

Chapters in books

- Herring, S.C., 2001. Computer-mediated discourse. In: Schiffrin, D., Tannen, D., Hamilton, H.E. (Eds.), *The Handbook of Discourse Analysis*. Blackwell Publishers Ltd., Oxford, pp. 612–634.
- Meirieu, P., 2010. Formatorii și comunicarea. In: Cabin, P., Dortier, J.-F. (Eds.), *Comunicarea. Polirom, Iași*, pp. 198–202.
- Suciu, L., 2019. Introductory chapter: discourse and discourse analysis. A retrospective approach. In: Suciu, L. (Ed.), *Advances in Discourse Analysis*. IntechOpen, London, pp. 3–15.
- Temple Adger, C., 2001. Discourse in educational settings. In: Schiffrin, D., Tannen, D., Hamilton, H.E. (Eds.), *The Handbook of Discourse Analysis*. Blackwell Publishers Ltd., Oxford, pp. 503–518.
- Tracy, K., 2001. Discourse analysis in communication. In: Schiffrin, D., Tannen, D., Hamilton, H.E. (Eds.), *The Handbook of Discourse Analysis*. Blackwell Publishers Ltd., Oxford, pp. 725–750.
- Yule, G., 2006. Pragmatics. <https://www.pdfdrive.com/pragmatics-by-george-yule-e163862273.html>.
- Thomas, J.A., 1995. Meaning in Interaction: An Introduction to Pragmatics. <https://www.pdfdrive.com/meaning-in-interaction-an-introduction-to-pragmatics-e186528721.html>.

Articles in proceedings

- Suciu, L., 2014b. The role of communication in building the pedagogical relationship. *Procedia Soc. Behav. Sci.* 116, 4000–40004.
- Suciu, L., 2021. A look inside academic communication during the pandemic. In: Boldea, I. (Ed.), *The Shades of Globalisation. Identity and Dialogue in an Intercultural World*, vol. 8. Arhipelag XXI Press, Tirgu-Mures, pp. 81–90.

Qualitative evaluation

Michael Quinn Patton, Claremont Graduate University, Claremont, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction and chapter overview	150
The purpose of evaluation	150
The contribution of qualitative methods to evaluation	151
Qualitative evaluation examples	151
Individual case examples	151
Focus groups	152
Evaluation through observation	152
The purpose of qualitative evaluation	152
Different evaluation purposes	153
Connecting evaluation questions to uses and methods	153
Impact inquiries involving qualitative fieldwork	153
Unanticipated consequences	154
Framing practical qualitative evaluation questions to enhance use	154
Qualitative evaluation units of analysis	155
Purposeful sampling for qualitative evaluation	156
Sampling high-impact cases	156
Comparison-focused qualitative evaluations	156
Outlier sampling	156
Positive deviance	157
Typical case sampling	157
Key informants, key knowledgeable, and reputational sampling	157
Utilization-focused qualitative evaluations	158
Conclusion	158
References	158
Further reading	158

Introduction and chapter overview

Evaluation research examines whether interventions to change the world work, and if so, how and why. Evaluation can serve both accountability and learning purposes, can help improve programs and inform decision-making about the future of programs and other change interventions. Qualitative evaluations thus serve diverse evaluation purposes. Purpose is the controlling force in determining evaluation use. Decisions about design, data collection, analysis, and reporting all flow from evaluation purpose. Therefore, the first step in an evaluation is clarifying study purpose, intended uses, and intended users. This leads to consideration of how to match qualitative methods to different evaluation purposes and inquiries. Framing qualitative evaluation questions for impact affects fieldwork approaches, units of analysis, and purposeful sample selection.

This chapter begins by reviewing diverse evaluation purposes and the specific qualitative contributions to particular purposes using specific qualitative examples to illustrate key options. The chapter discusses how case studies, interviews, focus groups, site visits, observations, and document analysis can contribute to evaluation. Connecting evaluation questions to intended uses and appropriate methods is a central theme. A variety of qualitative evaluation questions aimed at different types of uses are illustrated with the coronavirus pandemic as the focus. A major issue in qualitative studies is how to select cases and deal with small samples. The answer is to be intentional about choosing among purposeful sampling strategies. The chapter concludes with attention to impact evaluations and inquiry into unanticipated consequences.

The purpose of evaluation

Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness and/or inform decisions about future programming. Policies, organizations, and personnel can also be evaluated. Evaluative research, quite broadly, can include any effort to enhance human effectiveness through systematic data-based inquiry. Human beings are engaged in all kinds of initiatives to make the world a better place. These efforts include assessing needs, formulating policies, passing laws, delivering programs, managing people and resources, providing therapy, developing communities, changing organizational culture, educating students, and educating the

general public. In these and other efforts to solve problems, provide services, and make the world a better place, the question of whether the people involved are accomplishing what they want to accomplish arises. When one examines and judges accomplishments and effectiveness, one is engaged in evaluation. When this examination of effectiveness is conducted systematically and empirically through careful data collection and thoughtful analysis, one is engaged in evaluation.

The contribution of qualitative methods to evaluation

Qualitative data for evaluation have traditionally come from three primary sources: (1) in-depth interviews and focus groups with program participants, program staff, and other key informants knowledgeable about the program; (2) program documents; and (3) observations of the program through site visits. In addition, qualitative evaluations are drawing on new data sources like social media, remote satellite scanning, information data bases ("Big Data"), photography and videography, and websites. Qualitative methods are often used in evaluations because they tell the program's story by capturing and communicating the participants' stories. Evaluation case studies have all the elements of a good story. They tell what happened when, to whom, and with what consequences. The purpose of such studies is to gather information and generate findings that are useful. Understanding the program's and participant's stories are useful to the extent that they illuminate the processes and outcomes of the program for those who must make decisions about the program. The primary criterion for judging such evaluations is the extent to which intended users actually use the findings for decision-making and program improvement. The methodological implication of this criterion is that the intended users must value the findings and find them credible. They must be interested in the stories, experiences, and perceptions of program participants beyond simply knowing how many came into the program, how many completed it, and how many did what afterward. Qualitative findings in evaluation illuminate the people behind the numbers and puts faces on the statistics, not to make hearts bleed, though that may occur, but to deepen understanding. That deepened understanding includes the local context where implementation happens to identify the factors and practices that shape implementation and outcomes in different ways. Direct observations allow comparisons of what people say happens in the program with what is actually observed in practice.

Specific examples of qualitative evaluations for educational initiatives will illustrate the contribution of qualitative data to evaluations.

Qualitative evaluation examples

The Technology for Literacy Center was a computer-based adult literacy program that operated out of a storefront facility in a lower-socioeconomic area of a large city. After 3 years of pilot operation, a major funding decision had to be made about whether to continue the program. Anticipating the funding decision, a year earlier local foundations and the public schools supported an evaluation to determine the overall outcomes and cost-effectiveness of the Literacy Center. The evaluation design included both quantitative and qualitative data.

The quantitative testing data showed great variation. The statistics on average achievement gains masked great differences among participants. The report concluded that while standardized testing showed substantial achievement test gains for the treatment group vs. the control group, the more important conclusion concerned the highly individualized nature of student progress. This finding highlights the kind of program where qualitative data are particularly helpful and appropriate. The Technology for Literacy Center had a highly individualized program in which learners proceeded at their own pace based on specific needs and interest. The students come in at very different levels, with a range of goals, participated in widely varying ways, and make very different gains. Average gain scores and average hours of instruction provide a parsimonious overview of aggregate progress, but such statistical data do little to help funders understand what the individual variation means. To get at the meaning of the program for individual participants the evaluation included case studies and qualitative data from interviews (Patton, 2015).

Individual case examples

One case study featured a 65-year-old black grandmother. She worked as a custodian and house cleaner and was proud of never having been on welfare. She was the primary breadwinner for a home with five children spanning three generations, including her oldest daughter's teenage children for whom she had cared since her daughter's unexpected death from hepatitis. During the week she seldom got more than 3 h of sleep each night. At the time of the case study, she had spent 15 months in the program and progressed from not reading at all (second-grade level) to being a regular library user (and testing a grade level higher than where she began). She developed an interest in black history and reported being particularly pleased at being able to read the Bible on her own.

A second contrasting case involved a 42-year-old Caucasian woman who dropped out of school in the tenth grade. She was a clerical office manager. She tested at twelfth-grade level on entry to the program. After 56 h of study over 17 days she received a high school degree making her a high school graduate.

These short excerpts from two contrasting cases illustrate the value of detailed, descriptive data in deepening our understanding of individual variation. Knowing that each woman progressed about one grade level on a standardized reading test is only a small part of a larger, much more complex picture.

Focus groups

To get the perspective of more adult students in the adult literacy program, the evaluator conducted focus group interviews. Students were asked to describe the program's outcomes in personal terms: "What difference has what you are learning made in your lives?" Here are some responses.

- "I love the newspaper now, and actually read it. Yeah, I love to pick up the newspaper now. I used to hate it. Now I love the newspaper."
- "I can follow sewing directions. I make a grocery list now, so I'm a better shopper. I don't forget things."
- "Helps me with my medicine. Now I can read the bottles and the directions! I was afraid to give the kids medicine before because I wasn't sure."
- "Just getting a drivers license was wonderful. I'm 50 years old. I can drive well, but I was scared to death of the written test. But after learning to read in this program, I passed the test."
- I can fill out applications now. You have to know how to fill out an application in this world. It used to be so embarrassing not to be able to fill out applications. At least my application is filled out right, even if I don't get the job, at least my application is filled out right.

These group interview excerpts provide qualitative insights into and perspectives on the individual, personal experiences of adults learning to read. These quotes also illustrate one type of what constitutes qualitative data: direct quotations from interviewees. It is important to note that focus groups generate data through interaction. A common question is when to use individual interviews vs. focus groups. Many evaluations include both. Focus groups are especially valuable for generating data through discussion, prompting conversations so participants can express different points of view and feed off each other to generate additional layers of meaning. But not all topics might be suitable for a focus group, for example, on sensitive topics, topics that might put some participants at risk.

Evaluation through observation

What people say is a major source of qualitative data, whether what they say is obtained verbally through an interview or in written form through document analysis or survey responses. There are limitations, however, to how much can be learned from what people say. To understand fully the complexities of many situations, direct participation in and observation of a program is important. Observational data, especially participant observation, permits the evaluation researcher to understand a program or treatment to an extent not entirely possible using only the insights of others obtained through interviews. Interviews and focus groups will capture perceptions—what people think, feel and say they do—but observations will capture what people do. Of course, not everything can be directly observed or experienced—and participant observation is a highly labor-intensive and, therefore, a relatively expensive research strategy.

The purpose of observational analysis is to describe the setting that was observed and what participants do. This means that observational data must have depth and detail. The data must be descriptive—sufficiently descriptive that a reader can understand what occurred and how it occurred. The observer's notes become the eyes, ears, and perceptual senses for the reader. The descriptions must be factual, accurate, and thorough without being cluttered by irrelevant minutiae and trivia. The basic criterion to apply to a systematic observation is the extent to which the observation permits the reader to enter the situation under study.

For example, an evaluation was conducted of a child development education program for first-time parents. The purpose of the program was to increase the skills, knowledge, and confidence of parents. The program was also aimed at providing a support group for parents. Site visits to the program allowed direct observation of parenting discussions. Descriptions of these sessions then became the primary data of the evaluation. In short, the evaluators were to be the eyes and ears of the legislature and the state program staff, permitting them to understand what was happening in various parent sessions. Descriptive data about the sessions also provided a mirror for the staff who conducted those sessions, a way of looking at what they were doing to see if that was what they wanted to be doing. In addition to direct observations, parents and staff were interviewed.

The purpose of qualitative evaluation

Qualitative inquiry for evaluation is utilization-focused. This means the purpose is to generate results that can be used to make the world, or some part of it, a better place, at least that is the hope and intent. Education evaluations aim to improve educational programs and processes, inform educational decisions, and enhance student outcomes. To be useful, evaluations must be geared to the information and decision-making needs of specific intended users. Evaluations to improve curriculum for teachers are different from evaluations for school board members to allocate funds effectively. The stated purpose of an evaluation expresses the desired evaluation uses and the goals to be achieved. Those who commission evaluations and those intended to use findings determine purpose, which makes determining desired evaluation use and impact a values-based political process. Design decisions flow from purpose. Subsequent decisions about data collection, analysis, and reporting flow from design decisions. Therefore, the first step in an evaluation inquiry process is clarifying desired uses by identifying the study's purpose and primary intended users. Who is the evaluation for? How is it intended to be used? What is the desired impact of the evaluation inquiry?

Different evaluation purposes

Evaluations study the extent to which interventions (programs) work. For interventions to work and be evaluated they must be implemented with high fidelity. To illustrate purpose and impact distinctions, I'll use examples from the coronavirus pandemic that spread throughout the world beginning in January 2020. The pandemic gave rise to many evaluation studies. For example, evaluations examined what affected whether people complied with social distancing advisories. To what extent was protective gear, like masks, used? How did schools, communities, organizations, families, cities, and even countries vary in how they dealt with Covid-19? Why did some people and groups resist vaccinations? What educational approaches and messages were effective in increasing vaccination coverage? Answers to these questions from interviews, observations, and analysis of hospital and school records generated findings that were used to improve programs and determine what approaches were effective and efficient.

Different kinds of evaluation studies serve different purposes.

- Summative evaluations determine if a model or approach works and should therefore be expanded to other places. For example, a summative evaluation of a new curriculum would be used to decide whether to adopt the curriculum for all students and recommend expanding the curriculum to other school districts.
- Formative evaluations aim to improve a policy or program as it is being implemented, for example, when the Coronavirus pandemic hit in 2020 schools and universities had to quickly move their courses online. Qualitative feedback from students and teachers was used to help improve online teaching of existing curriculum books and materials.
- Developmental evaluations provide feedback for generating and trying out new approaches, for example, curriculum had to be developed for online instruction during the pandemic and evaluations helped in that development including what kinds of support and assistance teachers needed who were teaching online for the first time. Programs to support parents in home-schooling had to be developed. Socioeconomic disparities in access to the Internet had to be documented and addressed. Qualitative interviews and observations supported that development. (Formative evaluation aims to improve an existing program or intervention model. Developmental evaluation supports development of innovative new approaches adapted to changed conditions.)
- Action research evaluation aims to understand and solve a specific problem as quickly as possible, for example, shortages of protective gear for hospital workers during the pandemic required adaptations of traditional practices with data from patients, nurses, doctors, and hospital administrators generating lessons that could be shared throughout the world.
- Accountability evaluations judge whether policies have been appropriately followed and whether funds have been properly allocated and used as intended.

No single evaluation can serve all of these different purposes and audiences equally well. With clarity about purpose and primary intended users, the evaluator can go on to make specific design, data-gathering, and analysis decisions to meet the priority purpose and to address the intended audience and users.

Connecting evaluation questions to uses and methods

Utilization-focused evaluation research is meant to distinguish what works from what doesn't work. To get a sense of what this means in real-world practice, [Exhibit 1](#) presents a variety of qualitative evaluation questions aimed at different types of uses illustrated by the coronavirus pandemic as the focus. For each question and intended use, an example of qualitative data is identified.

Impact inquiries involving qualitative fieldwork

Qualitative researchers and evaluators study impact through designs that involve going into the field where interventions (programs, initiatives and policies) are being implemented to find out through observations, interviews, and documents how well the implementation is unfolding and how the people targeted by those interventions are responding. Site visits typically involve a team of researchers and evaluators who spend time at program sites, like schools and universities, interviewing staff and participants, observing operations, and reviewing program documents. Each site visit leads to a report for that site and multiple site visits are synthesized to report aggregate and cumulative program impacts.

Site visits require sensitivity to context and culture. When new cases of polio emerged in a remote area of India, qualitative epidemiologists accompanied the revaccination team and interviewed people in the area to find out how and why some children had missed being vaccinated in the recent campaign. They learned that some Muslim mothers had resisted the vaccination and had hidden their children because they'd heard rumors of a Hindu plot to sterilize their boys. Understanding the resistance to vaccination or other health practices, and developing approaches to overcome that resistance, requires an in-depth understanding of the cultural, social, and political systems within a particular context. Context includes attention to and understanding of important nuances of culture, politics, economy, history, geography, resources, and institutions. Qualitative inquiry makes attention to context a priority both for data collection and for reporting findings. This means documenting diversity and the contextual factors that explain particular variations even while identifying cross-cutting patterns and themes.

Exhibit 1 Illustrative evaluation questions, intended uses, intended users, and qualitative data.

<i>Evaluation question</i>	<i>Intended evaluation use</i>	<i>Intended education users</i>	<i>Qualitative data</i>
1. What are variations in the symptoms of Covid-19?	1. Improve diagnoses and treatment of symptoms.	1. School nurses	1. Observations of symptoms that show up in teachers and students.
2. What are the factors that support wearing masks?	2. Improve compliance with mask-wearing recommendations.	2. Public health educators and advocates of mask-wearing.	2. Interviews with those who comply with mask mandates and those who resist following mask mandates.
3. What are differences in how different cultural, ethnic, racial, and religious communities are affected by Covid-19?	3-a. Target health and education messages to be meaningful and influential to different groups? 3-b. Develop and implement policies to reduce health disparities.	3-a. Leaders and health education workers in different cultural, ethnic, racial, and religious communities. 3-b. Policymakers and political actors.	3.a. Case studies of how different cultural, ethnic, racial, and religious communities react to Covid. 3.b. Interviews with policy makers about how they make public health decisions; case studies of decisions and their effects on different groups.
4. How do school closures affect students, parents, and teachers?	4-a. Improve online education. 4-b. Provide additional services and interventions as needed and appropriate to deal with the effects of Covid-19.	4. Educational leaders, parent groups, student advocates, teacher associations, and family social workers.	4. Focus groups with specific target groups affected by school closures.
5. What are the different policy approaches to the coronavirus pandemic in different communities, organizations, states, and countries?	5. Learn from natural variations in policy approaches to compare and contrast what works and doesn't work under varying conditions?	5. Public health officials in national and international agencies who provide advice about prevention, education, and treatment approaches.	Qualitative case studies of the cultural, societal, political, and economic factors that affect policy differences nationally. Generating lessons from cross-societal qualitative comparisons.

The questions, purposes, and intended users in Exhibit 1 are merely illustrative of the great variety of possible utilization-focused qualitative evaluations that are possible. The key point of Exhibit 1 is to spotlight the relationship between inquiry question, intended uses, intended users, and qualitative methods.

Responses to the coronavirus pandemic have varied greatly by culture, country, and socio-political context. Culturally responsive evaluations position and interpret impacts in the context of cultural understandings.

Unanticipated consequences

People running programs of all kinds establish objectives and implementation strategies aimed at achieving positive impacts. But actual impacts may be quite the opposite of those hoped for and planned. I've evaluated literacy programs aimed at helping high school students learn to read that made them hate reading. I've evaluated juvenile justice programs aimed at rescuing young people from a potential life of crime that propelled them inadvertently on a pathway to becoming hard-core criminals. I've studied interventions that intended to support welfare recipients become economically self-sufficient that contributed to multigenerational poverty. I've seen international agricultural development programs aimed at increasing food production and income from cash crops turn fragile soils into desert through the introduction of inappropriate technological and management schemes. None of these programs were run by uncaring or incompetent people. Quite the contrary. In each case, I found the program leaders and staff to be committed, motivated, hardworking, and deeply engaged in bringing about change but, in the end, ineffective. Indeed, they were not just ineffective but counterproductive, having the opposite of their intended effects.

Qualitative inquiry is especially valuable for identifying unintended consequences and side effects. If all a program evaluator looks at is whether the intended impacts are attained, especially using standard performance indicators such as reading tests, employment statistics, and health outcome data, then other, unintended effects will be missed. To find unanticipated impacts, evaluators have to go into the field where things are happening, observe what is really going on, interview program participants about what they're experiencing, and find out through open inquiry what impacts are occurring, both intended and unintended.

Framing practical qualitative evaluation questions to enhance use

Being practical means that the essential criteria for making design decisions are realistic, contextually responsive, and consequential. Being realistic means asking straightforward questions that generate useful and actionable answers. Being contextually responsive involves understanding the demands, opportunities, and constraints of the situation in which the evaluation will take place including attention to power dynamics. An evaluation that is consequential means that the findings are relevant, significant, and applicable to the purpose intended. This is expressed in utilization-focused evaluation as ensuring intended use by intended users.

A pragmatic qualitative evaluation design is based on six questions:

1. Who are the primary intended users of the findings?
2. What are their intended uses of the findings?
3. To what extent and in what ways can relevant impact questions be adequately answered through qualitative data?
4. Can the design be successfully carried out, taking into consideration such issues as access to information, time available, evaluators' skills, and money or other resources required for the evaluation?
5. To what extent and in what ways will the results be practically useable and useful given the intended purpose of the evaluation?
6. When are findings needed to inform decision-making? (Qualitative evaluations include rapid and real time inquiries, like action research, as well as longer term studies.)

With these questions as context, let's now look at some of the design issues that arise in qualitative evaluations.

Qualitative evaluation units of analysis

A design specifies the unit or units of analysis to be studied. Decisions about what cases to study—issues of both sampling strategies and sample size—depend on prior decisions about the appropriate units of analysis to study. Often individual people like teachers are the unit of analysis, or participants in programs, or students. This means that the primary focus of data collection will be on what is happening to individuals in a setting and how individuals are affected by the setting. Individual case studies and variation across individuals would focus the analysis.

Comparing groups of people in a program or across programs involves a different unit of analysis. One may be interested in comparing demographic groups (males compared with females, whites compared with African-Americans) or programmatic groups (dropouts vs. people who complete the program; people who do well vs. people who do poorly; people who experience group therapy vs. people who experience individual therapy). One or more groups are selected as the unit of analysis when there is some important characteristic that separates people into groups and when that characteristic has important implications for the program. The coronavirus pandemic exposed great disparities in sickness, access to health care, and death by socioeconomic status.

A different unit of analysis involves focusing on different parts of a program. Different classrooms within a school might be studied, making the classroom the unit of analysis. Outpatient and inpatient programs in a medical facility could be studied. The intake part of a program might be studied separately from the service delivery part of a program as separate units of analysis. Entire programs can become the unit of analysis. In state and national programs where there are a number of local sites, the appropriate unit of analysis may be local projects. The analytical focus in such multi-site studies is on variations among project sites more than on variations among individuals within projects.

Different units of analysis are not mutually exclusive. However, each unit of analysis implies a different kind of data collection, a different focus for the analysis of data, and a different level at which statements about findings and conclusions would be made. Neighborhoods can be units of analysis or communities, cities, states, cultures, and even nations in the case of international programs.

One of the strengths of qualitative analysis is looking at program units holistically. This means doing more than aggregating data from individuals to get overall program results. When a program, group, organization, or community is the unit of analysis, qualitative methods involve observations and description focused directly on that unit: the program, organization, or community becomes the case study focus not just the individual people in those settings.

Particular events, occurrences, or incidents may also be the focus of study (unit of analysis). For example, a quality assurance effort in a health or mental health program might focus only on those critical incidents in which a patient fails to receive quality treatment according established standards of quality. A criminal justice evaluation could focus on violent events or instances in which juveniles run away from treatment. A cultural study may focus on celebrations. Specific timepoints could be units of analysis, as the qualitative evaluation could be capturing perceptions and practices at pre-implementation, early implementation and post-implementation stages.

Choosing among all of the possible options for units of analysis and cases for study can become overwhelming, whether the decision is about which time periods to sample, which activities to observe, or which people to interview. The trick is to keep coming back to the criterion of usefulness. What data collected during what time period describing what activities will most likely illuminate the evaluation inquiry? What focus will be most useful given the purpose of the inquiry? There are no perfect qualitative evaluation designs, only more and less useful ones.

The key to making decisions about the appropriate unit of analysis and cases to study is to determine what the evaluation findings will be able to credibly address at the end of the study. Is it useful to have findings about individuals, families, groups, schools, or some other unit of analysis? For scholarly inquiries, disciplinary traditions provide guidance about relevant units of analysis. For evaluations, one has to determine what decision makers and primary intended users really need information about. Do they want findings about the different experiences of individuals in programs or do they want to know about variations in classroom processes at different sites? Or both?

Purposeful sampling for qualitative evaluation

Purposeful sampling involves electing information-rich cases to study, cases that by their nature and substance will illuminate the evaluation question being investigated. The term “purposeful sampling” emphasizes sample selection to serve a particular purpose. Qualitative inquiry typically focuses in depth on relatively small samples, even single cases ($n = 1$), selected for a quite specific purpose. The logic and power of purposeful sampling lies in selecting information-rich cases for study in depth. Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of the inquiry, thus the term purposeful sampling. Studying information-rich cases yields insights and in-depth understanding rather than empirical generalizations. For example, if the purpose of an evaluation is to increase the effectiveness of a program in reaching lower-socioeconomic groups, the evaluator may learn a great deal by studying in depth a small number of carefully selected poor families. Purposeful sampling focuses on selecting information-rich cases whose study will illuminate the questions under study.

Purposeful sampling is also called purposive sampling. There is generally no difference in meaning. It’s a matter of which term one prefers. My evaluation and applied research work involves close collaboration with non-researchers (community people) who have told me they found the term “purposive” too academic, off-putting, jargon-ish, and unclear. One said, “Who talks like that?” In contrast, “purposeful” they could understand. It means: for a purpose. No one I knew, including myself, had ever used “purposive” in common conversation. Ordinary people did know, understand, and sometimes used “purposeful.” My aim was to communicate clearly and be use-friendly, so I introduced “purposeful sampling” and ceased using “purposive.”

Case selection is the foundation of qualitative evaluation. The findings from an evaluation will be determined by the cases studied. The type of sample selected should follow from and support inquiry into the evaluation and impact questions being asked. The purpose of a purposeful sample is to focus case selection strategically in alignment with the inquiry’s purpose, primary questions, and data being collected. Here are a few examples of purposeful samples using, again, the coronavirus pandemic to illustrate how the sample is derived from the study’s purpose.

Sampling high-impact cases

High impact cases are studied and documented in depth because of the impacts illuminated and the significance of the case to a field, problem, or society. High impact cases have high visibility. In the United States, the philanthropic foundation *Atlantic Philanthropies* supported a campaign to overturn the juvenile death penalty leading up to the Supreme Court’s 2005 *Roper vs. Simmons* case, which did rule against the death penalty. Given the high visibility and high impact of the campaign, the case study identified critical factors that made the campaign exemplary and examined the question of the likelihood of the Court’s decision having been influenced by the campaign.

A more recent example of a high impact, high visibility case is the story of Columbia University economics professor Jeffrey Sachs, author of the influential and visionary best-selling treatise, *The End of Poverty* (2005). With \$100 million from philanthropic donors, he launched the Millennium Villages Project in 10 African countries with the promise to eliminate poverty in those villages with excellent schools, technical assistance, and low cost innovation in agriculture, health, water, energy, nutrition, and education. The successes and failures that ensued involve case studies of the village projects, but Jeffrey Sachs and his approach constitutes a single high visibility case (Munk, 2013).

In the early days of AIDS research, infections almost always resulted in death. However, a small number of people infected with HIV who did not develop AIDS became crucial outlier cases that provided important insights into directions researchers should take in combating AIDS. Qualitative case studies of these remarkably resilient AIDS survivors provided important insights into the challenges of living with AIDS for those on antiretroviral drugs once they became available. Qualitative studies of community prevention efforts generated life-saving lessons that helped turn back the AIDS epidemic.

Comparison-focused qualitative evaluations

We’ve been looking at the value of in-depth study of significant cases. Now we turn to multiple cases, which intrinsically invite comparisons. Comparison-focused sampling looks in-depth at the significant similarities and differences between cases and the factors that explain those differences. Rigorous qualitative inquiry ensures that evaluation comparisons are systematic and meaningful. Different kinds of comparisons serve different purposes thus, again, the importance of purposeful sampling decisions in design. Given the overall purpose of an impact inquiry, the specific question being investigated, and the data to be collected, what kind of comparative sample would be most appropriate? The next section reviews some major options for comparison-focused sampling.

Outlier sampling

Outlier sampling, also called extreme or deviant case sampling, involves selecting cases that are information-rich because they are unusual or special in some way, such as outstanding successes or notable failures. In statistical terms, extreme case sampling focuses on the endpoints of the bell-shaped curve normal distribution (outliers) that are often ignored (or even dropped) in aggregate data reporting. Schorr (1988) used this strategy in studying especially effective programs for families in poverty, published as the influential book *Within Our Reach*. Malcolm Gladwell’s study of *Outliers* (2008) exemplifies this sampling strategy. He posed the

intriguing question: why do some people succeed, living remarkably productive and impactful lives, while so many more never reach their potential? He did case studies of the lives of hugely successful outliers like Mozart, Bill Gates (founder of Microsoft), and the Beatles. He found that success involves more than intelligence and talent; these outliers benefitted from favorable timing, connections, unusual opportunities, and simple luck.

In evaluation, the logic of extreme case sampling is that lessons may be learned about unusual conditions or extreme outcomes that are relevant to improving more typical programs. In studying an international program with hundreds of local sites around the world, we can expect that many local programs are operating reasonably well, even quite well, and that other programs are having difficulties. We also know that most programs average in performance. This information comes from knowledgeable sources who have made site visits to enough programs to have a basic idea about what the variation is. With limited time and resources an evaluator might study one or more examples of really poor programs and one or more examples of really excellent programs. The evaluation focuses on understanding under what conditions programs get into trouble and under what conditions programs exemplify excellence.

Brinkerhoff's (2003) *Success Case Method* for program evaluation provides a systematic way to identify successful programs, compare them to less effective programs, and figure out what makes the difference. In a single program the same strategy may apply. Instead of studying some representative sample of people in the setting, the evaluator may focus on studying and understanding selected cases of special interest, for example, dropouts vs. outstanding successes. In an evaluation of the Caribbean Agricultural Extension Project we did case studies of the "outstanding extension agent" selected by peers in each of eight Caribbean countries to help the program develop curriculum and standards for improving extension practice. The sample was purposefully "biased," not to make the program look good, but rather to learn from those who were exemplars of good practice.

In essence, the logic of outlier sampling is that extreme cases may be information-rich cases precisely because, by being unusual, they can illuminate both the unusual and the typical. In proposing an outlier sample, as in all purposeful sampling designs, the researcher has an obligation to present the rationale and expected benefits of this strategy as well as to note its weakness (limited generalizability).

Positive deviance

Positive deviance is a special application of outlier sampling that focuses on finding people or communities that have solved a problem where the norm in the area is for the problem to remain unsolved. This approach emerged in the 1970s in nutrition when researchers discovered that some families in poverty had well-nourished children. The thing that distinguishes positive deviance from more general outlier sampling is that the search for positive deviants is specifically aimed at finding high impact solutions, it is action-oriented. (Pant and Odame, 2009; Pascale et al., 2010). Advocates and practitioners have formed the Positive Deviance Alliance to promote using this method. Here's an example.

Typical case sampling

This strategy involves selecting and studying several cases that are average to understand, illustrate, and highlight what is typical and normal. Sometimes an evaluation will focus on outliers, but sometimes the focus is on typical students. For example, in describing a culture, community, or program to people not familiar with the setting studied it can be helpful to provide a qualitative profile of "typical" cases. These cases are selected using statistical data about what is average, or with the cooperation of key informants, such as program staff or knowledgeable community members, who can help identify who and what are typical. Keep in mind that the purpose of a qualitative profile of one or more typical cases is to describe and illustrate what is typical to those unfamiliar with the setting—not to make generalized statements about the experiences of all participants. The sample is illustrative not definitive.

When entire programs, schools, universities, or communities are the unit of analysis, the processes and effects described for the typical program may be used to provide a frame of reference for case studies of "poor" or "excellent" sites. When the typical site sampling strategy is used, the site is specifically selected because it is not in any major way atypical, extreme, deviant, or intensely unusual. This strategy is often appropriate in sampling villages for community development studies in low-income countries. A study of a typical village illuminates key issues like school access for girls that must be considered in any development project aimed at that kind of village.

In summary, there are many purposeful sampling options. In evaluation and policy research, the interests and information needs of decision makers and intended users of the findings will shape the sampling strategy.

Key informants, key knowledgeable, and reputational sampling

Key informants are a prized group. These are people who are especially knowledgeable about a topic and are willing to share their knowledge. Key informant interviews were developed by ethnographers to help understand cultures other than their own. Because being an "informant" can have negative connotations, I prefer the term key knowledgeable. They inform our inquiry when we tap into their knowledge, experience, and expertise. Key informants are especially important sources on specialized issues. The evaluation of the implementation of the Paris Declaration Principles on development aid relied heavily on key informants in

governments and international agencies because the topic is highly specialized. A study of New Zealand's Prostitution Reform Act identified and interviewed as key informants "those who have had direct experience of working with or as sex workers".

Key knowledgeable may stand alone as a purposeful sampling strategy or be used in combination with other approaches. For example, identifying outliers or typical cases may begin with key informant interviews. Key knowledgeable provide valuable expertise on and insights into the root of problems, as in a study of barriers philanthropic foundations experience that limit their capacity to think and act strategically. Key informant interviews are widely used to identify trends and future directions as when Gopalakrishnan et al. (2013) interviewed more than a dozen foundation leaders, evaluation practitioners, and social sector thought leaders to identify "Next Generation Evaluation" characteristics and approaches. In essence, key knowledgeable interviews are among.

Utilization-focused qualitative evaluations

Utilization-focused evaluation is an approach to evaluation aimed at intended use by intended users (Patton, 2008, 2012). Within that overall framework, utilization-focused sampling involves selecting cases that will be relevant to the issues and decisions of concern to an identifiable group of stakeholders and intended users. Such sampling can employ any purposeful strategy but adds a requirement that cases selected for study will have credibility, relevance, and utility for primary intended users. The ideal way to meet these criteria is to actively engage primary intended users in design and methods decisions, especially sampling. Focusing on generating useful and actionable findings, a common result of utilization-focused qualitative inquiry is identification of factors that explain the differences between what works and what doesn't work. Those factors and explanations can be used to inform decision-making and support program, practice, and/or policy improvements. To accomplish this action-oriented result, cases must be selected with that kind of finding and use in mind. This means attending not only to the relevance, contextual diversity, and complexity, but having sufficient access to in-depth information about the cases to be able to decipher individual case dynamics and cross-case patterns that will support conclusions about causal factors and recommendations about actions to be taken more generally beyond the cases studied. That increases impact.

Having weighed the options and considered the alternatives, evaluators and primary stakeholders make their design and sampling decisions, sometimes painfully, but always with the recognition that there are no perfect designs. The sampling strategy must be selected to fit the purpose of the study, the resources available, the questions being asked, and the constraints being faced. This holds true for sampling strategy as well as sample size.

Conclusion

Which qualitative evaluation design is best? Which qualitative evaluation strategy will provide the most useful information to decision makers? No simple and universal answer to these questions is possible. The answer in each case will depend on the purpose of the study, the scholarly or evaluation audience for the study (what intended users want to know), the funds available, the political, organizational, and cultural context, the interests/abilities/perspectives of the researchers, and the intended impacts sought. The word "design" is both a noun and a verb, denoting both the product and the process of planning an impact-driven inquiry. Issues of context, resources, purpose, and hoped-for impact all come into play as the design emerges and is then implemented and the results used.

References

- Brinkerhoff, R.O., 2003. *The Success Case Method: Find Out Quickly What's Working and What's Not*. Berrett-Koehler, San Francisco, CA.
- Gladwell, M., 2008. *Outliers: the Story of Success*. Little, Brown, NY.
- Munk, N., 2013. *The Idealist: Jeffrey Sachs and the Quest to End Poverty*. Random House, New York, NY.
- Pant, L.P., Odame, H.H., 2009. The promise of positive deviants: bridging divides between scientific research and local practices in smallholder agriculture. *Knowl. Manag. Dev. J.* 5 (2), 138–150.
- Pascale, R., Sternin, J., Sternin, M., 2010. *The Power of Positive Deviance: How Unlikely Innovators Solve the World's toughest Problems*. Harvard Business Press, Cambridge, MA.
- Patton, M.Q., 2008. *Utilization-Focused Evaluation*, fourth ed. Sage Publications, London.
- Patton, M.Q., 2012. *Essentials of Utilization-Focused Evaluation*. London Sage.
- Patton, M.Q., 2015. *Qualitative Research and Evaluation Methods*, fourth ed. Sage, London.
- Schorr, L., 1988. *Within Our Reach: Breaking the Cycle of Disadvantage*. Doubleday, NY.

Further reading

- Flick, U., 2007a. *Designing Qualitative Research*. Sage, London.
- Patton, M.Q., Campbell, C.E., 2021. *Utilization-Focused Evaluation*, fifth ed. Sage Publications, Los Angeles.
- Patton, M.Q., 2015. *Qualitative Research and Evaluation Methods*, fourth ed. Sage, London.
- Stake, R.E., 2010. *Qualitative Research: Studying How Things Work*. Guilford, New York.

Arts-based research: inquiry through, in, and as art

Janinka Greenwood, University of Canterbury, Whangarei, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction: a photograph	159
My positionality	160
Researching through and in art	160
Arts-based qualitative research	161
Expanding existing forms of art to facilitate research	161
Art teachers researching academically	162
Using art to share findings of research	163
Using art to collect data	163
Using art as a different way of knowing	164
The research in arts-based research	165
The art in arts-based research	166
Knowledge-making through arts-based research	166
Conclusion	166
References	167

Introduction: a photograph

A young woman is sitting in long summer grass. She wears a floral dress and a Covid-dictated mask in matching fabric. A mirror in an ornate wooden frame stands—serendipitously—against a nearby tree. She reaches out to touch her image in the mirror—a girl who, although her mirror reflection, has nothing covering her face (Fig. 1).

The photograph is one of a series in an online exhibition in the 2020 INSEA conference. The artist, Michaela Petřek Linhartová, plays with images of mask wearers—some humorous, some fantastical, some macabre, and all somewhat unsettling—to explore practical and psychological effects of the first months of Covid lock-down.

It is possible to look at this photograph in a number of different ways. We might simply delight in the doubling of the masked and unmasked figure. We might delve into how the deceptive mirror image was created. We might focus on the incongruity in the semiotics—restrictive mask and the pretty girl, floral dress and burgeoning summer landscape. We might toss around *meanings* of the work, speculating about how the maskless mirror image reflects the yearning in the soul of the masked girl. We might consider the work as one item in the artist's deliberate inquiry into the impact of Covid on individuals in her community. We might see it and the rest of the series as direct evidence of the national mandated isolation that forced the artist to work within a context of tightly restricted contact and movement. And perhaps we might see other things.

This range of possibilities highlights the interconnections between art and research and the potential and challenges in arts-based research, and so points to my intentions in this chapter. I am interested in the fluidly morphing continuum that connects art and research, and in the ways arts and research play with elements of that continuum. I am also interested in the ways that the use of arts in research open up further possibilities in the broad and already varied field of qualitative research.



Fig. 1 Image in the mirror.

Qualitative research focuses on the particular, detailed and contextually situated. It acknowledges and even celebrates the partial nature of each contribution to knowledge, arguing the usefulness of discovering and responding to the specifics of people's experiences, understandings and interactions. It utilizes a wide range of tools and in recent decades processes of arts have become significant items in its repertoire. I explore how some researchers have used those tools.

There is already a substantive and growing body of writing about arts-based research (including Barone and Eisner, 2011; Leavy, 2017) and I have contributed to it (Greenwood, 2012, 2016, 2019). This writing will not offer a summary of or a commentary on that considerable array of writing although it will lightly reference some of it. Rather it focuses on the investigative elements that are inherent in the arts and, drawing on examples from my own work and that of my doctoral students as well as the published work of well-known practitioners, it examines a range of ways the research potential of art had been used in specific research projects. It discusses how art-based research interacts with research, art and knowledge.

A word about terminology. I need to emphasize that although arts-based research is a useful umbrella term there is no one arts-based methodology. I will use *arts-based* research to signify the umbrella term. When I refer to a specific process I will talk about an *art-based* approach. And because it is the domain I work with I would like to clarify drama/theater terminology. Theater is the parent art: as Schechner (1988) explains, it takes many forms and serves many functions. *Drama* is the term used by many British and Scandinavian practitioners to describe the use of facilitated theater strategies as processes for exploring situations and as a media for learning. Other Europeans and Latin Americans tend to prefer the term *theater* to cover this kind of work. Applied theater and process drama are terms that are also commonly used. When citing others' work I will use whatever term they used; in talking more generally I will use *drama* when I talk about specific applied investigative or learning strategies and *theater* when I talk about the wider art domain.

In this chapter I first explore ways research is embedded in the making of art and ways some artists have deliberately used their art to question and provoke investigation. Next I situate arts-based research within the broad qualitative paradigm. I then report some more formally academic uses of art in research and consider how these examples prompt further questioning of the nature and purpose of research and its relationship with learning. I suggest that while a broad range of both popularly accessible and professional arts may be used in research, certain elements are essential. Finally I highlight contributions arts-based research can make to our ways of knowing. First I acknowledge my own positionality in the discussion.

My positionality

Later in this discussion I will argue that an art-based approach to research is unlikely to be dispassionate and further that the subjectivities that come with arts-based research are themselves purposeful and interesting. Therefore it is useful to explain some of my own practice in the arts and research. I am a writer with particular interest in theater. I also work extensively using various drama/theater strategies for teaching and learning. As an academic researcher I use a wide spectrum of approaches. Almost all are qualitative; some are explicitly arts-based, others include historical and current case study, participatory action research and narrative inquiry. Looking back over my work I see vestiges of my interest in theater in all my research: I am particularly interested in and attentive to the people involved and their shifting perspectives, the impact of place, the unfolding of action, and the tensions that are played out. As a doctoral supervisor I have worked with emergent researchers from a many different countries and cultures, and again I find that my drama/theater practice urges me to consider the tensions they may face in navigating western academic expectations and what is meaningful and useful in their local and cultural context. My own experience makes me reluctant to draw too definite a line between art and research and even between arts-based research and that of other kinds.

My husband, Gerry, has been a visual artist who throughout his life explored wood, clay, bronze, pencil and paint. For many years he was a full-time potter producing both domestic ware and more sculptural pieces. As I lived alongside his processes of making I became very aware of the constant exploration of materials and of understandings of the world that impelled his work. I recognize that for him art was a persistent process of research.

Researching through and in art

Research through art has a much longer history than the popularization of the term arts-based research. Artists, it may be argued, seek to know the world through their making of art, and seek to offer others new insights into the world through what they present in their art. Novelists like Proust, Joyce and Woolf explore the phenomenon of time and its impact on human awareness and understanding. Perhaps their own intuitions and their strategies of inquiry evolve as they write. They are researching both *what* they write about and *how* they can write about it with authenticity. They offer their readers new knowledge: disturbing their existing schema about the nature of reality, provoking new insights and new questions.

Painters like Monet, van Gogh and Picasso probe the relationship between what their eyes see and what they experience because of their seeing. They also probe the ways they can manipulate paint to capture and communicate their exploration.

In his *Antigone* the dramatist Anouilh uses the lens of a well-known play by Sophocles to critically analyze and problematize moral dilemmas of resistance in occupied France. The epidemic transformation of ordinary intelligent citizens into short-sighted, thick-skinned and savage rhinoceros in Ionesco's play, *Rhinocéros*, serves as a critical lens to probe the collective hysteria

of people who not only readily accept the dictatorship of fascism but also willingly surrender their judgment and their moral accountability (Ionesco, 1960).

The artists mentioned above can readily be seen as deliberately using their art to probe the human condition and provoke conceptual reconsiderations. But art making is itself a form of research. Artists explore their materials and the possibilities of their interaction with them.

Each time he centers lumped clay on the wheel and teases it to open and flare out into the curve of a bowl or bottle, a potter is engaging in an inquiry into that piece of clay and into the impact of the movement and pressure of his hands. As he fires the kiln and builds the fire he is intensely engaging in experimentation—new for each firing because of weather outside and the placement of pieces inside—with how to spread and intensify the heat, how to reach 1300 °C uniformly through his kiln, how to ensure the oxygen is reduced sufficiently so that his black glazes gloss and break into red, and his whites pool in milky thickness and thin to flash shimmers of color.

So too the actor hunts for the physicality and the verbal nuances that will meld together the character, her own bodily reality and the evolving collective cast understandings of the intentions of the play. The writer fights through a jungle of possible words, tracking elusive ideas, resisting hijack by runaway familiar discourses, shifting blocks of development across pages, hunting for those words that have the right edge of meaning.

Making art, it can be asserted, is a continuous process of research.

Arts-based qualitative research

Broadly speaking, qualitative research enables us to understand and prompts us to interpret the human impact of situations and events. As I make the final edits to this chapter, the invasion of Ukraine has passed its hundredth day. Each day the media publishes positional maps and numerical data. This quantification allows us to appreciate the magnitude and the shifting parameters of the combat and of its effects. For the numbers to have meaning in terms of actual people and their lives, we need to know more: we need to know about the experiences of individuals and about the complexity of the social dynamics of the particular localized contexts in which people strive to make sense of their experiences. And we need to be offered lenses for interpreting these matters.

This is the role of qualitative research, an umbrella term for an increasing array of investigative methods that examine the situatedness and complexity of particular aspects of an issue. It also provides conceptual lenses that offer the ability to make critical interpretations of events and people and to frame the significance in particular ways (Denzin and Lincoln, 2018). Arts-based qualitative research offers the use of an art medium as a means of capturing situated instances of such complexity. The making of art allows probing, and often openly subjective, interpretation. As an example, in the war on Ukraine, personal stories of children who were eventually evacuated from Mariupol and photographs of the tense faces of firemen straddling burned out rafters in a bombed apartment-block humanize the statistics. At the same time, they offer the investigator's interpretation and also invite the audience's critical reaction. The examples that follow illustrate in more detail the possibilities of arts-based qualitative research further.

Expanding existing forms of art to facilitate research

There are artists who deliberately shape their art forms to engage others in investigation and analysis. I offer examples from the field with which I am most familiar. One such artist is Boal, a theater practitioner whose work has cut a path from stage-centered activist performance through a progression of theater forms for collective investigation and problem-solving to the creation of a community-groups-centered legislative theater. His early theater was shaped by a commitment to engage with social issues: in the context of a dictatorship in Brazil his theater company's project was to make art a means to fight for national liberation. Over time he developed what he later called the *Theater of the Oppressed* (Boal, 1979). His intention was to create forms of theater where audiences could become active in examining the oppressions they experience, communicate with each other to analyze the context and causes and to explore possible solutions. Central to the work were facilitator-led processes for keeping the focus *real*.

For example, *Image Theater* is a physical enactment process for groups to select an experienced oppression for examination, identify the power relation between oppressor and oppressed, brainstorm and critique ideal outcomes, conceptualize how the power imbalance might be shifted so that a positive outcome would be possible, and decide on a first step in a direction leading to a shift in power. The facilitator reminds participants that they cannot call on *magic*, that the only initiatives they can propose are ones they themselves can take and not those they may hope for from the oppressor. The first step has to be a practical one that can be taken in the current circumstances (Boal, 1979). Boal's process shares features with academic research strategies like participatory action research in that it has an evolutionary exploratory design and actively involves participants in defining the object of investigation and in carrying out the investigation. In discussing participatory action research as a social practice Kemmis and McTaggart (2005) highlight the importance of first developing communicative space that enables communicative action and consequently an exploratory space that supports exploratory action. Boal's process provides such spaces in which participants can communicate and explore through physicalization as well as in words.

Later in his life Boal entered formal politics and was a city councilor in Rio de Janeiro. He adapted his theater processes to a means of political engagement with his constituent communities. Very simply, his office staff involved a team of actors who would

engage theatrically with disadvantaged communities to select and explore social problems they wanted resolved and, following a performance in which they could intervene to change the course of action, vote on legislative changes that Boal would then bring to the City Council (Boal, 1998).

Two guiding beliefs are evident in Boal's approach. One is the value of collaborative inquiry. "The solution has to be discovered together," he says. "We bring a true question so that everyone can freely and democratically participate and intervene." (Boal, 1996, p. 25). The other is the power of the creative curiosity latent in all people. "We begin from the principle we are all artists, but handicapped artists because education handicaps us." He adds, "We are theater ... By observing myself in action, this dichotomization allows me to change my way" (Boal, 1996, pp. 34–35). His words assert that collaborative engagement in the creativity of art-making opens up and intensifies a common human capacity to critically investigate and to problem-solve.

Independently from Boal's work Heathcote also developed a project of using theater strategies, for which she preferred the term *drama*, for participant inquiry into social and ethical issues. Her use of teacher-in-role and the alternation of frames of action with deliberative reflection created an educational movement which is variously called *process drama*, *applied theater* and *drama as a learning medium*, and which has been copied and/or has been modified and adapted by numerous well-known theorizing practitioners (including O'Neill, 1995; O'Toole, 1992; Taylor, 2003; Schonmann, 2011) as well as by classroom teachers. While this form of work is designed to be investigative and analytical, its initial primary objective is a vitalization of teaching and learning process rather than research. However, the evolution of methodologies such as action research and reflective practice problematizes demarcation between learning and research. Much of the work carried out in arts-based research fits into a nexus between learning and research.

In New Zealand in the 1970s Wilson, a Māori sculptor and educator, developed a project, lasting almost two decades, that used art-making as a catalyst to challenge and change the monocultural bias of schooling (Greenwood and Wilson, 2006). Participating schools, represented by about twenty students and three or four teachers, would be brought to live for a week in a Māori space and work with the local Māori elders, artists and community to develop a mural, and sometimes other art forms, depicting a local story of place. To research their story and the traditional art forms for their mural the groups would need to learn from the local holders of knowledge. In this context research took an experiential form: the participants lived immersed in a space where Māori values and protocols were normal, listened to elders talk, helped in the kitchen, worked communally with artists who were expert exponents of their traditional forms. What was explored reached beyond the selected story and means of accomplishing the art work: it involved lived experience of different values and ways of seeing the world, and sometimes a revaluation of participants' own identity and societal role. To call such investigation research somewhat stretches commonly used definitions: the process was variable, and, while there was expectation of broad outcomes, the deeper discoveries were made at personal and small group levels and were neither predictable nor measurable. Much of the new knowledge would have been hard to put into words as it involved affect as much as cognition. The project invites reflection on the nature of research and the forms, popular and academic, that it might take, and about the kinds of knowledge research in and through the arts might generate.

Art teachers researching academically

In terms of how art forms are used there is a strong relationship between projects such as those of Boal and Wilson and those that formally construct themselves as academic research. The key difference perhaps is in terms of nominated focus, explicit, albeit sometimes evolving, methodological procedure and explicitly discussed outcomes. These factors are illustrated in the work of two post-graduate students.

Trudy Barrett's project (Barrett, 2014) took place in a secondary school art classroom where she needed to work within the prescribed curriculum but was given freedom to find her own approaches. She developed a progression of activities that each taught a particular skill and at the same time provoked students to incrementally examine and enrich their understandings of themselves and their situatedness. Barrett began with an explicit research question: "How can visual arts teachers use art making to help adolescent students to recognize, critically analyze, and interpret their senses of self, place, and community?" (p. 26). She details her methodological design and explains how she will collect and analyze her data and present her findings. She is explicit in identifying the historical, philosophical and epistemological contexts in which she has situated her study. She explains how she navigates the methodological duality of her work: the teaching methodology that is intended to provoke student inquiry and the research methodology that allows her to critically reflect on and evaluate the effect of her teaching and the visible and verbalized shifts in students' conceptualizations. She reports her findings verbally and in images of the artifacts.

Barrett's study engages with both verbally analytical interpretation and the less definable esthetic dimensions of her and her students' work. Montages of photographed artifacts complement her narrative, as does a series of manipulated images through which she reflected on the relationality of her role as insider and outsider, as teacher, researcher and artist, as a woman with her own history. She remains artist throughout her investigative journey at the same time as she adheres to the rigors of academic research. She offers us a basis on which we can critically discuss her findings, and the same time she lures us to further play with the presented images and accounts and explore their esthetic reverberations.

While it is generally agreed that an esthetic dimension is an intrinsic element of art, there is less agreement about what constitutes the esthetic. Minimally it can be described as an element that affects the senses and triggers subconscious responses as well as rational ones: it speaks of a dynamic interaction between art and viewer; perceptions vary historically, culturally and individually (Greenwood, 2010).

As a primary school teacher Toby Greenwood wanted to explore esthetic responses of 10–12 year old students (Greenwood, 2011). He explains that his own research question was: “How do students formulate and express their esthetic response to their own work and to that done by others?” which he gradually simplified to words that his participants were more likely to use: “What makes this particular piece of art good?” (p. 3). The study had two stages. In the first sixteen students from three schools were asked to bring two works they considered *good*, and explain their choice. Then they were shown a number of art pieces the researcher had selected and asked their opinion. In the second stage participants from one of the schools were invited to a workshop where they could work on any project they liked in a medium of their choice with the opportunity to consult with a professional artist, if they chose. His study found that, while the bases of their judgments varied, the students did actively make judgments about art works: some responded primarily to texture or color, some to implicit story, others to technique, still others to personal, family or cultural associations, and some to recognition of a famous name. The workshop component found commitment as well as variation in students’ ways of approaching art-making: some were concerned predominantly with technical process and others with imagery or narrative intention. The research report concludes with a discussion of the need for primary teachers, who tend to be generalists rather than subject specialists, to learn what students themselves think and value in order to achieve the goals of the curriculum and to develop the human right to access art that is asserted by the UNESCO *Road Map for the Arts* (UNESCO, 2006).

Using art to share findings of research

In many cases the use of art may be a final stage in the process of research: the sharing of findings. Sometimes the plan for an art based presentation is an integral part of the project. Verbatim theater (Soans, 2015; Brandt and Harcourt, 1994) and ethnodrama/ethnotheater (Saldaña, 2016; Sallis, 2003) are two performative events that are constructed out of careful research. The exact use of participants’ words in interviews is a feature of both. Verbatim theater presents itself primarily as a theatrical event and ethnodrama as research. The theater art form of both provides a platform for re-telling real people’s stories to highlight social problems and issues of social justice.

Sometimes researchers whose projects do not begin with the intention of making art might later turn to art as a means of sharing their findings with audiences outside the academy.

Hasnat (2017) studied the (dis)engagement of rural parents in Bangladesh in their children’s education. The data revealed not only a lack of engagement but also a major communicative gap between rural parents and teachers in local schools. His research project was not arts based and he submitted a formal written thesis. Later, however, he re-presented his findings as puppet theater. Two puppets were made from wooden spoons with the handle of one covered with a t-shirt and lungi and that of the other in a suit. The first represented the parents and was given dialog from verbatim data gathered from all the parents interviewed. Similarly the second, representing teachers, was given dialog from the teacher interviews. An overhead projector offered contextual images and occasional commentary. One reason for the performative presentation was to highlight the human aspects of what tends to be seen by policy makers, and some academics, as a statistical problem. An equally important motivation was to develop a means of sharing findings in a way that would be accessible to the people at the heart of the study and lead to further discussion by them. Many of the rural parents are not literate and although the teachers are literate they are not expected to read an academic report in English. The script was first published in a journal as an experimentation with art-based research (Greenwood and Hasnat, 2017) and then offered as live performance at a conference (Hasnat and Greenwood, 2018). The teachers in the conference audience actively debated the issues presented, relating them to their own, albeit different, contexts. Hasnat then planned to translate the script into Bangla on his return home and to present it not only to rural audiences, but also, if possible, to policy-makers who are unlikely to read any academic reports of research. So far Covid has interrupted this enterprise.

Tisi (2019) investigated how immigrant parents in Christchurch, New Zealand *teach* culture to their children. She considered both their strategies for maintaining the family’s original culture and the ways they positioned themselves and their children in terms of the culture they had immigrated into. Hers was also a written thesis, but she sought to bring her participants’ narratives to life by creating portraits of them and developing dramatic monologues from the interviews. After she had drafted the first monologue she trialled it in a live presentation at a conference. Teachers there suggested they would find such monologues useful for developing multicultural awareness. The feedback gave her confidence to continue with her plan. To preserve participants’ anonymity she photographed herself dressed in their style and surrounded by similar objects to those in their homes: the photographs were then turned into paintings. Each dramatic monologue used a careful selection of the participant’s verbatim words and she performed one of the monologues in her oral examination. She plans now to further develop a theatrical piece using the monologues.

These two examples highlight the role arts can play in making research findings accessible and interesting to those who are not academics but for whom the research is most relevant and potentially useful.

Using art to collect data

Arts can be useful in eliciting and collating data that may not be readily verbalized by participants. Photo voice (Wang and Burns, 1997) has become a widely used process that gives participants cameras and asks them for images that they find are meaningful in

relation to the topic being investigated. Alam (2016) asked teachers in his study to photograph what they considered effective practice and students to photograph what they valued most about their school. Both groups later talked about their images.

The taking of photographs allows a different kind of process than a simple interview: participants are given time to think about their response and they may be able to suggest something in a photograph that they cannot readily express in words. Often words are found later as participants explain and discuss their images, individually or in a group. The image continues as a presence during such discussion, inviting further speculation and perhaps new insights.

Other modes of representation might be used to elicit data, such as physical images, movement, stories, improvisations. In debriefing my students after their first school practicum I often use physical prompts. For example, I may ask them to make body sculptures of some aspect of their experience, offer interpretive readings of one another's images, and further refine their own if they want to clarify their meaning. Or I may place a chair in the center of an empty room, nominate a succession of significations of the chair (such as confidence in answering students' questions, managing unruly behavior) and ask them to place themselves at distance to the chair that would symbolize their feelings. They would then explain their positions to their nearest neighbors. I find that such activities allow participants to process their responses a little more than if they are simply questioned.

With those who cannot verbalize well, because of age, language difference, trauma or other constraints, the opportunity to offer visual or embodied information is particularly valuable, but often requires specialist expertise in interpretation.

Research may also investigate existing items of art. An early New Zealand play, *For love of Appin* (Mulgan, 1920) speaks of the split in the colonial psyche: a sense of still belonging to, but being far from, the homeland of Britain and the beginnings of realization that home could be in New Zealand. New Zealand identity is represented in the play in terms of the hard work of a farming life, the right to work one's own land, the importance of women as partners on the land, neighborly support and a sense of egalitarianism. At the same time that identity is presented as a homogenously monocultural one. There is no sign or mention of Māori or of any of the other immigrant groups that were significant at this time in the colony. So too the plays, poetry, sculptures, paintings of ancient Greece, medieval England, Franco's Spain offer researchers insights into the attitudes and perspectives of the time that are not necessarily recorded in other archives.

The current global proliferation of images of masks, including the photograph described in the introduction, offer data of how people around the world have accepted, rejected or subverted the restrictions of the Covid era.

Using art as a different way of knowing

In explaining that art opens up wider understandings of "what it means to know," Eisner (1998) crystallized and legitimized in the academic arena what artists have always intuitively appreciated. Humans are complex beings with body as well as mind: we experience sensually, viscerally, emotionally, and arguably spiritually, as well as intellectually. Making art is an embodied activity. It involves sensitivity to the multiple modalities of experience, evolving capacity to process aspects of experience through active physically engaged exploration and a drive to express discoveries sensually and sometimes kinetically as well as often posing conceptual challenges. Some researchers explicitly engage with these qualities of art in order to *know* in a different and richer way.

Exploration of visual ways of knowing is the basis of Barrett's project as described above: she sees the art class as opportunity for students to explore, stretch and sometimes transform the way they see themselves and their world, and she wants her research to *re-mark* the art classroom as a space for active inquiry. McNaughton (2019) explored the role of art gallery educators in New Zealand. In her rationale for the study she cites Arnheim (1969) who maintained that "sensory knowledge makes language possible, since our interface with reality is through our senses" (McNaughton, p. 49). Her participants' narratives highlighted the value of hands-on activities that provide multi-sensory experiences and problem-solving opportunities. McNaughton writes (p. 260): "Gallery educators were aware that various types of thinking could be woven together to develop complex understandings in art gallery learning ... Artworks were used to connect abstract ideas to concrete sensory data ... to make [students] wonder and trigger inquiry."

In my work I have used drama-based inquiry as a process that promotes discovery through both real and imagined experiences that allow participants to adopt roles that may be very different from their personal ones, to create possibilities and to navigate challenges, and to critically reflect on stages of the proceedings. I see it as a process that enables participants to try out—to embody—different perspectives and different contexts, to unlearn as well as to learn, to explore the fracture lines of meaning as well as to develop deeper, more complex understandings.

For example, as part of a scheme to train local people as teachers for a remote rural region in the north of New Zealand I was asked to teach an intensive course about the use of drama for learning (Greenwood, 1999, 2002). The course became not only an illustration of a methodology, but also a space for the participants to examine how they constructed their reality. The first sessions were used to teach a range of drama strategies. One was *Boss of My Story*. A participant shares a story that has particular personal meaning. She then chooses members of the group to take specific roles and act out her story. She has the right to stop the enactment at any stage and redirect the actors—it is her story and she is the boss. The participants were very careful with one another's stories but they were also willing to take small risks in enacting detail that might have been glossed over in the original narrative. The storytellers watched the action with emotional intensity and some did make corrections. When each storyteller debriefed after the enactment many said they were moved by the actors bringing their story to life and talked about new insights. Opportunity to look at their stories from outside as well as from inside seemed to fracture some relics of memory and allow fresh understandings. The next sessions moved to investigating local histories.

Participants selected the history of the Treaty of Waitangi as one topic. Briefly, the Treaty is the document of the agreement that in 1840 laid the foundation for colonization and that today still actively addresses, with contentions, issues of power, rights, self-determination and the two-cultural basis of New Zealand society. The Treaty is commemorated in an annual public holiday and in that year, as in many others, there were protests. As a group the participants were initially critical of the protests, but admitted they had scrappy knowledge about the events leading to the Treaty and about its subsequent history. As they researched and experimentally enacted key elements from their research individual and group understandings repeatedly shifted, splintered, and reshaped. For instance, one participant wrestled with notions of power and political activism. At one stage he questioned why the government was encouraging Treaty awareness: "Surely they realize the radicals are going to use the knowledge to bring the government down?" Later he re-conceptualized the issue: "I'm thinking where I'm wrong is that I see the government up there—that the government should be the ones making the changes—but it's the people who are making the changes." Still later he re-considered the contribution of those who protested: "It's like jumping the queue—if you wait, you stay at the back—if you do it outrageously, you'll get to the front" (2002, p. 327). In reporting the project I contended that "as the students grappled with understanding the Treaty ... and embodying their understandings in physical images they could share with each other, they were in fact writing their own inscription of New Zealand history. They were acquiring power as well as knowledge" (1999, p. 282). Looking back, I reflect how participants' images and enactments drew on the not-yet-conscious, the liminally understood and how it was only through the interaction of makers and audience that they were able to verbally explain what they had created and what it meant.

The discipline of a/r/tography (Irwin, 2004) is explicitly concerned with capturing the experience of exploring self and the world through art practice. Its orthography indicates the practitioner's fused role of artist, researcher and teacher as well as blurring distinctions between exploration through art-making and research as an academic discipline. The backslashes also highlight the notion of fracture: understandings are seen as emergent, not necessarily integrated, and resistant to tidy verbal encapsulation.

A/r/tography makes an explicit claim to be research, albeit with different ways of conceptualizing knowledge than has been generally accepted in academic contexts. It is construed as a research approach that celebrates multiple modalities of art and uses the making of art as lived inquiry, that is as a process of remaining continuously open to finding the evolving and *in-between* spaces in ideas, feelings and meanings as they emerge through the work. Accounts of a/r/tography are often couched in language that may seem esoteric to those outside the field, in part because its practitioners draw heavily on academic theorizations to situate it firmly as academic research, but also because of the difficulty in expressing concepts that are in their nature elusive and, while temporarily describable, not definable.

The intrusion of a/r/tography and other practices that use art as an investigative process further disturb already troubled boundaries of research.

The research in arts-based research

I have deliberately interchanged the words *research*, *inquiry*, *exploration* and *investigation*. All involve search and discovery, but while they have very slightly different everyday connotations they have different academic status.

The academy is already a site of contestations. In 2011 Denzin and Lincoln stated: "We are in a new age where messy, uncertain and multivocal texts, cultural criticism, and new experimental works will become more common, as will more reflexive forms of fieldwork, analysis and intertextual representation" (p. 1). They also noted that conservative and positivist political forces, with their power of funding, continue to privilege traditional forms of peer-reviewed evidence-based research. Within this contestation, some forms of art-based research are acknowledged, sometimes suspiciously, within the academy; others are not and are more likely to be classified as pedagogy.

The teacher training project I described above was part of my doctoral thesis and reports were published in academic journals. It is considered research: research about the use of drama. However the participants were also searching and discovering. They were learning about the use of drama, but they were also using the drama for other explorations and discoveries. Similarly, Barrett's doctoral work was recognized by her examiners as significant academic research. Her examiners judged that she made a new contribution to knowledge. Her student participants also found new knowledge. Their new knowledge was about personal, and sometimes group, understandings of themselves as idiosyncratic but emplaced individuals and their relationships with their significant communities. They were also discovering new knowledge about art: that it was something they could play with purposefully rather than a teacher's pre-ordained set of standards.

Distinctions between researching and learning are blurred. Action research has long challenged expectations that research would take place first and its findings would later be applied (perhaps) to professional action. Finley (2005) explores the nexus between research and learning in her work in projects that enable ordinary local communities to critically explore and resist the politically potent discourses that oppress their lives, a process she terms a people's pedagogy. "In my work," she writes, "I seek to create esthetic spaces in which to experience transformational performances at a visceral level ... In this genre of arts-based research, the researcher attempts to involve people as experts of their own lives and to create forums or outreach to venues outside the academy" (p. 444).

It is important to remember that life and its problems exist outside the academy as well as within it. Research and discovery are important to people as well as to universities. The spectrum of current arts-based approaches offer the possibility of relevance and accessibility to ordinary people as well as to scholars.

The art in arts-based research

Throughout this discussion I have talked about art as a vital creative and critical process. In daily usage, however, the word art has a very wide range of denotations. It may refer to works that are considered as high art and prized by galleries and collectors, works that indicate virtuoso technical skills, a subject that is taught in schools, street activities, reproductions, visual artifacts generally, or the wide spectrum of creative activities including performative and verbal works as well as visual ones. It may be used in phrases like *the art of parenting* or *the art of blackmail*. Academic definitions of art tend to be more restrictive. For example, my university, guided by the national Performance Based Research Fund (PBRF) that ranks and funds universities on their research outputs, frames art in terms of creative outputs that have been shared with audiences and are endorsed by peer review and citation. Interestingly the PBRF guidelines include art for advertising and business purposes, but not for learning purposes (Tertiary Education Commission, 2018). In contrast the UNESCO Road Map (2006) avoids definition of art and rather affirms that children and adults “need to progressively learn to understand, appreciate and experience artistic expressions by which fellow humans—often called artists—explore, and share insights on, various aspects of existence and coexistence” (p. 4).

Across the examples I have discussed art in the term arts-based also has a wide range of significations: the art within specific projects may align with high art, popular media, use of a particular tool that is used by art makers, and it may draw on any one of, or combination of, a range of creative visual, performative or even poetic processes. As I understand it, two qualities are needed for the art to be a vehicle for research: craft and vitality.

Craft involves both knowledge and skill. Each art process offers both opportunities and constraints. In drama, for instance, the craft of the facilitator involves balancing space for open exploration with provocations and challenges, using role to endow participants with role, creating fictitious frames that allow risk, building tension that sparks curiosity and further inquiry, juxtaposing action and critical reflection (Greenwood, 2012). It also involves incrementally developing participants’ craft so they are able to creatively use the strategies offered. In Hasnat’s puppet theater the theatrical elements are very simple; the craft is in his strategic selection and organization of the original data to create dramatic tension and yet retain the content of his findings. In a/r/tography there is craft is moderating and unifying the dialogue between the roles of artist, researcher and teacher which while interfused still speak from different perspectives.

Brook (1972) spoke of a *deadly* theater: it is theater that is perhaps technically skillful but replicative, boring and simply not alive. If art is to be a means of research it cannot be a deadly art. It has to pulse with the excitement as well as the attentiveness of inquiry. We may use procedures inherited from previous practitioners but we can use them in ways that are sensitive to the movement of the work and the interaction of the participants, just as the potter is sensitive to the clay and to the fire.

If arts-based research is to be an effective means of developing knowledge and knowing that goes beyond that produced by more traditional means, then the art within it needs to be more than a prefabricated method: it needs to be craftful and alive.

Knowledge-making through arts-based research

Alternative ways of knowing through the arts have been discussed throughout this chapter. Here I highlight five key contributions arts-based research can make to our understandings of knowledge and our repertoires of knowing.

As mentioned, discoveries through arts-based research may be emotional or visceral as well as conceptual. It has been repeatedly argued that western cultures neglect the holistic aspects of humanness by privileging the cerebral and verbal (Irwin, 2004; Leavy, 2017; Greenwood, 2016). Connecting participants with the sensation and perceptions of their bodies and inviting them to explore affect as well as ideas is one contribution.

Another is made by the way findings often exceed definition. Some projects do report clear and specific findings (and perhaps hint of other less specific possibilities). Others result in awakenings that may still be incomplete, fractured, ambiguous or latent. Specific knowledge is undoubtedly valuable. So is awareness of that which impacts on our lives but is not so readily definable. We live in a world of ecological crisis, political upheaval, brutal war and pandemic enclosure, a world where it is hard to untangle truth from misinformation. Capacity to admit and engage with the fractured and liminal helps us live with uncertainty.

Many of the discoveries through arts-based research are primarily personal, even if reports are published. They are forms of knowing that are held within the individual and by being internalized they present capacity for action. They are also subjective. Recognition and respect for the subjective, in ourselves and in others, offers resistance to hegemonic discourses.

Discoveries through arts-based approaches are often evolving, and may continue beyond the project itself. They highlight the non-static nature of knowledge and the importance of continuous learning and unlearning.

Dealing with the uncertainties of the present and the future requires development of action-orientated ways of knowing: including collaborative capacities, practical as well as theoretical criticality, and creativity in imagining and exploring alternatives. The group work and investigation strategies in arts-based projects foster growth of these capacities.

Conclusion

I began this discussion with a photograph, and I conclude with another.



Fig. 2 Exploration.

This one is from a retrospective exhibition of my husband's work. It shows a large rounded stoneware bowl with two figures who have climbed to its rim and peer inside. Their stance suggests the arduousness of the climb. They lean into the unmapped space with curiosity and uncertainty. The work is part of a sequence of Gerry's exploration of connections between people and artifacts (Greenwood, 2021/1983). It evokes a narrative, and it plays with form and space (Fig. 2).

For me, the work offers a metaphor for arts-based research. The enterprise is an embodied one. Processes of inquiry are supported by the strong structures of art traditions and contemporary art practices. Researchers stand on the rim of what is already known and what is yet to be explored. Light and shadow play over the space for exploration. There are latent possibilities created by the selected art strategies but the space is open. The results are not predicted. The quest is an adventure.

The umbrella of arts-based research covers an eclectic and protean range of approaches to research that make it difficult to circumscribe as a methodology. Each art and each application of an art opens up different opportunities for exploration and offers different constraints to guide the work. Nevertheless there are some common features in the diverse array. They offer a multi-faceted approach to investigation, invoking a mix of physical, emotional, spiritual and intellectual responses to experiences, and so offer alternative and additional kinds of knowing from those generated by more traditional qualitative and quantitative approaches. They may generate ideas and specific new understandings but they are also open to intuitions, complexities and liminalities. They have the potential to facilitate inquiry at the level and in the manner that are accessible and meaningful to their participants and share results of inquiry with audiences in ways that are accessible and meaningful to them. Importantly, they are grounded in art practices that have long histories of exacting attention to material, form, process and evolving meaning.

References

- Alam, S., 2016. Teachers, Collaboration, Praxis: A Case Study of a Participatory Action Research Project in a Rural School of Bangladesh. Doctoral Thesis, University of Canterbury. UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/11856>.
- Barone, T., Eisner, E., 2011. Arts Based Research. Sage.
- Barrett, T.-A., 2014. Re-marking Places: An A/r/tography Project Exploring Students' and Teachers' Senses of Self, Place and Community. Doctoral Thesis, University of Canterbury. UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/10040>.
- Boal, A., 1979. Theatre of the Oppressed. Pluto Press.
- Boal, A., 1996. Augusto boal. In: Delgad, M., Heritage, P. (Eds.), In Contact With the Gods? Directors Talk Theatre. Manchester University Press, pp. 15–34.
- Boal, A., 1998. Legislative Theatre: Using Performance to Make Politics. Routledge.
- Brandt, W., Harcourt, M., 1994. Verbatim. Victoria University Press.
- Brook, P., 1972. The Empty Space. Penguin Books, London.
- Denzin, N., Lincoln, Y., 2011. Introduction: the discipline and practice of qualitative research. In: Denzin, N., Lincoln, Y. (Eds.), The Sage Handbook of Qualitative Research, fourth ed. Sage.
- Denzin, N.K., Lincoln, Y.S., 2018. The Sage Handbook of Qualitative Research, fifth ed. Sage.
- Eisner, E., 1998. The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice. Merrill.
- Finley, S., 2005. Critical arts-based inquiry: the pedagogy and performance of a radical ethical aesthetic. In: Denzin, N.K., Lincoln, Y.S. (Eds.), The Sage Handbook of Qualitative Research, fourth ed. Sage, pp. 435–450.
- Greenwood, J., 1999. Journeys Into a Third Space: A Study of How Theatre Enables Us to Interpret the Emergent Space Between Cultures. Doctoral Thesis, Griffith University. Griffith Research Repository. <https://research-repository.griffith.edu.au>.
- Greenwood, G., 2021/1983. Exploration. Stoneware bowl with climbers. In: Living the Fire [Art exhibition. Reyburn House]. Whangarei, 8 June—4 July.
- Greenwood, J., 2002. Marking the unmarked: using drama cross-culturally in teacher education in New Zealand. Syst. Res. Behav. Sci. 19 (4), 323–328.
- Greenwood, J., 2010. Aesthetic learning, and learning through the aesthetic. In: Schonmann, S. (Ed.), Key Concepts in Theatre/Drama Education. Sense Publishers, pp. 47–52.

- Greenwood, T., 2011. Makes Art Good? A Case Study of Children's Aesthetic Responses to Art Works. Master's Thesis. University of Canterbury. UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/5751>.
- Greenwood, J., 2012. Arts-based research: weaving magic and meaning. *Int. J. Educ. Arts* 13. <http://www.ijea.org/v13i1/v13i1.pdf>.
- Greenwood, J., 2016. The limits of language: a case study of an arts-based research exploration. *N. Z. J. Res. Perform. Arts Educ.* 6, 88–100. <https://static1.squarespace.com/static/5ab9651cf2e6b1034679faab/t/5b051bb270a6adb9251135f0/1527061443118/Journal-2016-9-Greenwood.pdf>.
- Greenwood, J., 2019. Arts-based research. In: *Oxford Research Encyclopedia of Education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.29>.
- Greenwood, J., Hasnat, M.A., 2017. I am illiterate, they are educated; what can I say to them? A playful exploration of the impact of performative presentation of case study data. *N. Z. J. Res. Perform. Arts Educ.* 7, 108–123. <https://static1.squarespace.com/static/5ab9651cf2e6b1034679faab/t/5b03c284aa4a99b7ce243f30/1526973070843/Journal-2017-9-Greenwood-Hasnat.pdf>.
- Greenwood, J., Wilson, A.M., 2006. *Te Mauri Pakeaka: A Journey Into the Third Space*. Auckland University Press.
- Hasnat, M.A., 2017. Rural Parents' Engagement in Education in Bangladesh: Problems and Possibilities. Doctoral Thesis, University of Canterbury. UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/14615>.
- Hasnat, M.A., Greenwood, J., 2018. I am Illiterate, They are Educated; What Can I Say to Them? A Performative Presentation of Case Study Data. *International Drama in Education Research Institute Conference*, Auckland. <http://www.dialogueforcommunity.com/2018/07/mahammad-hasnat-and-janinka-greenwood-i-am-illiterate-they-are-educated-what-can-i-say-to-them-a-performative-presentation-of-case-study-data/>.
- Ionesco, E., 1960. *Rhinoceros and Other Plays*. D. Prouse, Trans. Grove Press, NY.
- Irwin, R.L., 2004. *A/r/tography: a metonymic metissage*. In: Irwin, R.L., de Cosson, A. (Eds.), *A/r/tography: Rendering Self Through Arts-Based Living Inquiry*. Pacific Educational Press, pp. 27–38.
- Kemmis, S., McTaggart, R., 2005. Participatory action research: communicative action and the public sphere. In: Denzin, N.K.S., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*, third ed. Sage, pp. 559–603.
- Leavy, P. (Ed.), 2017. *Handbook of Arts-Based Research*. Guilford Press.
- Linhartová, M., 2020. Masks for Every Day. INSEA Conference, Olomouc and on-line. <https://www.inseaconference.com/section-2>.
- McNaughton, E., 2019. Lighting Fires on the Beach: Learning in Art Galleries in New Zealand. Doctoral Thesis, University of Canterbury. UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/16931>.
- Mulgan, A., 1920. *Three Play of New Zealand*. Whitcombe & Tombs.
- O'Neill, C., 1995. *Drama Worlds: A Framework for Process Drama*. Heinemann.
- O'Toole, J., 1992. *The Process of Drama: Negotiating Art and Meaning*. Routledge.
- Saldaña, J., 2016. *Ethnotheatre: Research From Page to Stage*. Routledge.
- Sallis, R., 2003. Ethnographic performance in an all boys' school. *NJ* 27 (2), 65–78.
- Schönmann, S. (Ed.), 2011. *Key Concepts in Theatre/Drama Education*. Sense Publishers, pp. 47–52.
- Soans, R., 2015. *Crouch, Touch, Pause, Engage*. Out of Joint and National Theatre Wales production.
- Schechner, R., 1988. *Performance Theory*. Routledge.
- Taylor, P., 2003. *Applied Theatre: Creating Transformative Encounters in the Community*. Heinemann.
- Tertiary Education Commission, 2018. Performance-Based Research Fund Panel-Specific Guidelines for the 2018 Quality Evaluation. <https://www.tec.govt.nz/assets/Forms-templates-and-guides/PBRF/76cca23ea0/PBRF-2018-Panel-Specific-Guidelines-April-2018.pdf>.
- Tisi, S., 2019. Case Studies of South Asian Parents in Christchurch: How Do They Teach Culture to Their Children? Doctoral Thesis, University of Canterbury UC Research Repository. <https://ir.canterbury.ac.nz/handle/10092/17512>.
- UNESCO, 2006. Road map for arts education. In: *The World Conference on Arts Education: Building Creative Capacities for the 21st Century*. Lisbon, 6–9 March 2006.
- Wang, C., Burns, M., 1997. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ. Behav.* 24 (3), 369–387.

Responding to cultural diversity through action and community-based participatory approaches in qualitative research

Rosalind Beadle^a, Alfredo Ortiz Aragón^b, and Ernest Stringer^c, ^aFlinders University, Bedford Park, SA, Australia; ^bUniversity of the Incarnate Word, San Antonio, TX, United States; and ^cIndependent Scholar, Fremantle, West Australia, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	169
Situating community-based, participatory action research	169
Localized educational practices	170
Building community capacity	171
Using information technology and social media for community based action research	173
Culture, curriculum and pedagogy	174
Organizational and systems change: implementing holistic developmental processes	175
Ethics in participatory and action research	176
Conclusion	176
References	177

Introduction

In the current context schools, colleges and universities provide a meaningful and useful preparation for life in our complex modern societies, despite consistent decreases in funding and a narrowing focus of education. The lack of success in providing an adequate education for children from lower income and culturally different backgrounds continues as a central concern in most nations. Internationally, there is continuing and strong evidence of the failure of school systems to provide an adequate education for the children of poorer and culturally diverse families. An extended analysis by [Sarason \(1990\)](#) with the telling title, “The Predictable Failure of Educational Reform,” indicated that over a twenty year period there had been no progress in a key aspect of the US national reform agenda – closing the achievement gap of students of different social classes and ethnicities. [Cole’s \(2011\)](#) more recent study covering the following twenty years found fundamentally the same ineffective results. His analysis, “What’s Culture Got to Do With It?,” indicated that reform efforts were largely ineffective and restricted to improvements in ways of organizing schools, classrooms and teaching/learning practices that had emerged in the past. He advocated changes to the culture of the classroom on general pedagogical principles and the need for organizational reforms that would support those changes.

Within Australia, continuing reports of the Council of Australian Governments (COAG) over a fifteen years period reveal a failure to make any inroads on the scandalous difference between educational outcomes for Aboriginal children, as compared with non-Aboriginal children ([Close the Gap Campaign Steering Committee, 2016](#)). Despite significant investment in special programs and services, national tests indicate that Aboriginal children still perform at the very lowest levels of tests of academic achievement, despite occasional successes in a small number of locations. The continuing application of standard approaches to schooling for students whose culture and life-histories differ significantly from those of mainstream Australian students has failed to accomplish the desired ends of “special” programs.

This chapter presents processes of participatory, community-based action research as an important means to enable educators and educational institutions – schools, colleges, universities and so on – to modify and adapt their standard educational procedures to accommodate the realities of the diverse cultures that comprise the reality of modern societies. Following sections present the rationale for engaging in community-based approaches to teaching and learning, the value of building community capacity to enrich classroom learning activities and how current resources in information technology and social media can facilitate these developments. The chapter also suggests that changes involved in school procedures are likely to have an impact beyond classrooms, requiring concurrent negotiations with school authorities.

Final sections of the chapter describe how recent developments in research theory provide strong support for engaging participatory, community-based approaches to research in order to produce curriculum content and learning activities meaningful and purposeful for the educational and social needs of students in their particular locations. By extension, we also suggest that the processes of localized research and development are particularly relevant to contexts where cultural diversity requires modifications to standardized approaches to teaching and learning to accommodate the realities of the context.

Situating community-based, participatory action research

Community-based, participatory approaches to research derive from developments in qualitative inquiry that emerged relatively recently as a legitimate and powerful set of research methodologies. These types of investigation are articulated in a wide range

of literature, including textbooks, handbooks and encyclopedias describing the general theory and practice associated with qualitative and mixed methods research. A small sample includes Denzin and Lincoln (2019), Greenwood et al. (1993), Abma et al. (2018), Charmaz et al. (2019), Chevalier and Buckles (2019), Coghlan and Brydon-Miller (2014), Denzin and Giardina (2018), Heron (1996), Kemmis et al. (2014), Schubotz (2019) and Boden et al. (2015). General texts on research in education likewise include sections of action research, recent literature including those by Sutter (2011), Cresswell and Poth (2017), Johnson and Christensen (2019), Mills and Gay (2018), Boudah (2019) and Biesta (2020).

These and others provide a rich body of theory that embodies the philosophic and paradigmatic orientations associated with the diverse approaches to research now encompassed by qualitative research. In essence they challenge some of the assumptions embedded in earlier approaches to research, describing the fundamental differences required to establish the authority of a scientific approach to research. In doing so they reveal the substantial differences between the orientations of research in the human sciences and those of the physical sciences. The demarcation between them is not fixed, however, since positivistic approaches to research often provide information usefully applied to issues in the conduct of human affairs, while issues associated with research in the physical sciences are influenced by the social contexts in which research activities are engaged.

While the above citations relate to the general theory and practice of qualitative research, a significant corpus of literature related to action research in education continues to emerge. Though each takes a somewhat different approach, and some focus on first person action research which is essentially non-participative, a current sample of recent literature includes.

Stringer and Ortiz Aragón (2021), Mills (2017), Bradbury (2015), McNiff (2017), Mertler (2019a,b), Putman and Rock (2017), Hendricks (2016), Efron and Ravid (2019) and Sagor and Williams (2016).

Localized educational practices

As Cole (2011) and Sarason (1990) have indicated, effective change in educational outcomes are unlikely to be accomplished merely by focusing on classroom practices. These have a proven record of failure, pointing to the need not only in changes to standard procedures in teaching and learning, but also changes in the associated organizational arrangements. Although carefully articulated system-wide change processes provide the best context for engaging in truly effective educational practices, in many locations teachers and administrators have the capacity to instigate developmental teaching and learning processes that engage student energy and interest and provide the basis for highly educational learning experiences. Despite continuing pressures to centralize and control all aspects of schooling, participatory and community-based approaches to research have the potential to initiate more locally relevant educational programs and practices. For example, Barrett (2007) describes ways of conducting qualitative research in music education. Nah and Lee (2016) describe how action research was used to actualize children's participation in the development of outdoor play areas in an early childhood institution. Calhoun (1993) describes how action research can enhance the processes of educational leadership, while Hendricks (2016) describes the use of action research as reflective practice for school improvement. In similar vein Wells (1994) suggests the need to create communities of inquiry to change schools from within.

In our own spheres of work we have learned, with and from local people, how to modify and adapt learning processes to tie them more clearly to the contexts in which they are played out. Early in his teaching career, Ernie discovered that the Aboriginal students in his classroom, so silent and unresponsive in the classroom, came to life when he walked with them through the surrounding bush-land and learned from them all the wonders of the flora and fauna with which they were intimately connected. Using these as the basis for lessons in the classroom, his students became his teachers as they expounded on the knowledge with which they were so familiar (Stringer and Ortiz Aragón, 2021).

In a similar context (Stringer, 2015) observed a high school teacher, saddened by her unresponsive students, engaged them in a range of activities drawn from the local community. As part of her class program she made representations to the local police sergeant who sent officers to the school to speak to her class about their work. Students also visited the police station where they learned many features of police work, including processes for dealing with miscreants and the duty of care they had when an arrest was made. In the process students developed relationships with the police officers and came to better understand the nature of police work, including how it derived from legislated laws rather than the decisions of individual officers. The outcomes of this process were a better understanding of the operation of the law and a better relationship between police and youth in this small community. All these were encompassed within a carefully articulated program drawn from the state curriculum: (Fig. 1)

Wood and Govender (2013) report on the way a group of South African school leaders were able to promote sustainable development and improvement within the challenging contexts of their under-resourced school communities. Stringer (2008b, 2015) reports on a highly successful action research project in East Timor, where school authorities used participatory action research processes to engage local communities in the redevelopment of their school system when that small nation received independence after a long struggle with colonization.

The essential dimension of community-based participatory research is that it is not done *on* people, but *by* and *with* people who have previously been designated as *subjects*. In one highly successful research project Ernie assisted a neighborhood group in their local elementary school in a largely Hispanic suburb of Albuquerque in the US to engage in an action research project. Concerned about the quality of education provided for their children members of the neighborhood group sought to discover ways in which parents and teachers could improve the school experiences of their children. The neighborhood group recruited other parents and practiced interview techniques under direction of a university professor, then interview parents who came to school

WEEK 1 ICP2.1; ICP2.2; ICP1.3; ICP1.4; LS1.2; LS1.4; V2.1; R2.2; I1.2 Week 2 LS1.3; LS1.2; R2.3; R2.4; W2.3; ICP3.1; ICP1.4 Week 3 LS1.1; LS1.2; R2.1; R2.4; W2.1; W2.3; W3.1; ICP2.1; ICP2.2; ICP2.3; ICP2.4; C3.2 Week 4 C3.1; C3.2; C3.3; C2.3; LS1.1; LS1.3; LS1.4; W3.1; W2.1; W2.3	1. Brainstorm what students know about police. 2. Using the internet, research that different jobs that police can do, and what students would like to do. 3. Create a police file. Label and illustrate. 4. Develop questions to ask police when they visit the classroom. Write down questions and practice reading two questions each. 5. Ask questions when police visit classroom. 6. Discuss and record why people are arrested (in relation to non-Indigenous law). 7. Discuss and record what happens when Indigenous law is broken. 8. Use comparison chart to identify different values of Indigenous and non-Indigenous peoples. 9. Discuss why we need non-Indigenous laws – driving, break and enter, theft, assault, DV, rape, alcohol and drugs, etc. 10. Why do Indigenous people need Indigenous laws?	• Write up brainstorm for display. • Record info from net. • Police file • Written and verbal practice of questions. • Students questioning. • Write why people are arrested and what happens when Indigenous law is broken. • Comparison chart. • Complete Handout 1.
--	--	---

Fig. 1 High school class: theme – the police. N.B. Outcomes are drawn from the following Western Australian Curriculum areas: English: Listening and speaking (LS), Reading (R) Writing (W) and Viewing (V). Society and Environment: Investigation, communication and participation (ICP), Culture (C) and Natural and social systems (NSS).

for parent-teacher day. The results emerging from the participatory processes of analysis that followed provided a list of practical actions that teachers and parents could take to improve the children’s schooling (Stringer, 2008a).

Building community capacity

Much educational research continues to focus on classroom practices on the assumption that identifying “best practices” will provide the answer to enduring problems in education. Despite a continuing focus on the individual attributes of students, large scale studies dating back many decades indicate that despite the importance of classroom and school variables, the major part of variance in educational outcomes relates to family and community factors (Coleman, 1966, Alexander and Morgan, 2016). As this chapter previously has indicated, changes in educational or academic performance might best emerge from a closer alignment between the school, student families and the local community. Occasional events that invite families and community into the school are insufficient to accomplish the strategic activities required, however, and the following narrative describes the potential for developing intentional and long-term relationships with parents and community members. Not only do parents become increasingly aware of the nature of school life, but school staff build a better understanding of the life-worlds from which their students emerge.

These issues became evident to Beadle (2018) who discovered that a school breakfast program operated by a group of women in a remote Aboriginal community had multiple benefits within the classroom, the school and the community. In a context where test results of student academic performance in the local community were among the lowest in the nation, a group of local Aboriginal grandmothers, aunts and mothers initiated a solution for enhancing attendance and engagement in the school. They recognized that the school was not effectively addressing the needs of their children and their regular absences were a cause of concern,

We were worrying for them tjitji (kids). We wanted them to go to school and stay in class, not be running away all the time. It’s important for those kids to get learned, you know, grow up and get a job proper way, like the Breakfast Minyma

Breakfast Minyma (2012), pers. comm., cited in Beadle (2018:60).

Assisted by a developmental support worker a series of meetings were held between the women, local Council members and the school principal at which the idea of providing breakfast at the school took shape. Within a few weeks the project was launched, an efficiency that contrasts markedly with projects externally initiated and managed by non-Aboriginal “outsiders”. As well as having a positive impact on children’s attendance, by coming to the school each day to deliver breakfast the women became a strong presence at the school where almost all school staff were non-Indigenous “outsiders”.

The women’s daily presence meant that relationships could develop with school staff who previously had experienced very little interaction with their pupil’s family members. Most of the teachers had lived in cities, so the context in which they now worked and lived was alienating and unfamiliar, something that prevented relevant and effective teaching. The presence of the women acted as a “gateway” to the local community, enhancing the teachers’ understanding of community life and providing access to information that could assist their teaching and relationships with students (Beadle, 2018). This enabled them to reconfigure their programs in a way better suited to the needs of the children and gain a better grasp of the causes of the many challenges they faced in the classroom.



Fig. 2 A page from "mirlirtjarra Kuurl Mirrka Palyalpayi (Making Good Food at Warburton School)" by Olive Lawson.

The women became known as "The Breakfast Minyma" and acted as an informal Aboriginal reference group for the school principals. They were consulted about issues such as appropriate disciplinary measures, how to attract community members to school events, and determine solutions to matters such as school uniforms and school security. After the school breakfast had finished the women would sometimes accompany the children into their classrooms, an activity that had previously been well outside their comfort zone. As a result of these experiences some of the women were then invited to work as Aboriginal Education Officers (AEOs),¹ further enhancing the presence of local people in the school. The Minyma reflected positively on the impact of their presence in the classroom, and the way they had a natural authority over the kids, something the non-Aboriginal teachers sometimes struggled with.

When I am in the classroom [as an AEO (Aboriginal Education Officer)] I feel good, happy with the kids. It's important there is an AEO in the classroom. So you know, when the teacher can't tell the kids to sit down when they run around and make the teacher talk [get frustrated]. [But] the kids, they always listen to me, I tell them in language and they listen.

Breakfast Minyma (2012).

As the women established themselves as valued and productive workers and broadened their knowledge of day-to-day life at the school, their increasing confidence led them to explore additional ways of enhancing their children's educational experience. They engaged in collaborative planning with the teachers and ran a series of "lessons" framed around their own experiences as workers. These were delivered in the local language using techniques that were so successful that one teacher commented "I've never seen the class so attentive!" (Teacher 2011, pers. comm. cited in Beadle, 2018:170).

Because so few of the textbooks employed by the teachers included or acknowledged local events or cultural experiences, the women responded by developing a set of nearly thirty bilingual "story" booklets, using content derived from their experience as workers. In addition to making books about their work, they assisted other groups in the community to produce books that covered a variety of topics about community life. Feedback from teachers and the principal indicated that these were popular reading materials among the children. While the English text in the booklets reflected the literacy requirements determined by the school curriculum, the content and translations into the local language gave value to the children's remote desert lives and resonated with their cultural experiences. This contrasted with the externally produced "age-appropriate" readers with unfamiliar topics such as "Katie and her grandmother take the train to the beach" provided by the school. The importance of these books was demonstrated when the state Minister for Education launched their first book in Perth, her department offering financial support for the women to travel to the event (Fig. 2).

In addition to books, the Breakfast Minyma developed PowerPoints and short documentary films which they used during their lessons. They recognized the relevance of these learning materials for their children, one commenting that:

¹AEO's were equivalent to Teaching Assistants.

When we are showing them the photos and the movie the kids say: “Oh, that’s my karparli!” “Oh, that’s [child’s] Mummy!” It makes them feel proud looking at the pictures of what their families are doing. It’s good for the kids, to see the photos, watch the ladies. Like they want to learn, when they grow older, they see the photos, they see the family working there, they get happy the kids, *yuuu* (yes), and when they must finish school, [they will] get a job you know, like the ladies too

Breakfast Minyma (2012), pers. comm., cited in [Beadle \(2018:169\)](#).

In these and other ways the mothers and grandmothers began to understand more clearly how their children were expected to participate in a classroom shaped by a curriculum and social context controlled by the norms of the dominant culture. Observation of and feedback from their children, as well as their own student experiences in school had provided the women with a deep understanding of their children’s struggle to understand the rules, codes and social norms expected of them by their teachers ([Minutjukur and Osborne, 2014](#); [Tjitayi and Osborne, 2014](#); [Beadle, 2018](#)). As opportunities allowed, the women therefore oriented their school activities in a way that acknowledged the “deeply embedded axioms, epistemologies, ontologies and cosmologies” of their life-world ([Tjitayi and Osborne, 2014](#)). In the ways articulated above, they were able to make a positive contribution to their children’s schooling by delivering their programs in a way that observed the deeply engrained relational component of community life, conversing in their own language, and addressing behavioral issues -such as teasing - in a culturally appropriate way.

In these ways the women improved the children’s learning experiences by modeling their own ways of teaching in their family environments. In doing so the women provided what [Tjitayi and Osborne \(2014\)](#) refer to as “close family voices” that enabled the children to participate more confidently in classroom and school activities. Recognizing the familiarity and comfort their presence offered the children at the school, the women described their school-based workplace (The Family Place) as a “magnet”:

The Family Place is like a magnet pulling them kids. It’s really good that those kids hang around. At recess and lunch time the kids run to the Family Place to be with their families. That’s making them hang about at school instead of going home

Breakfast Minyma (2012), pers. comm. in [Beadle \(2018:90\)](#).

The status, pride and self-esteem the Minyma gained from their work led them to initiate solutions to address other concerns regarding the wellbeing of their families and community. Other projects and services to emerge included the provision of meals for elderly people in the community, an evening support program for teenage girls, and catering services for community and school events. In broadening their repertoire of activities, and thus the scope of their work in the community, the Minyma further consolidated their role as “workers”. In a context where Aboriginal voices are typically marginalized, the respect and authority they experienced as workers gave them the confidence to share their stories with a range of audiences at conferences and other events in the city. These stories helped dispel prominent “mainstream” narratives which rarely represent remote Aboriginal people as valued and productive workers.

The Minyma’s self-initiated activities within the school demonstrates the importance they attributed to the educational wellbeing of their children and their desire to incorporate local language and culture into their children’s learning. As is common in many other contexts where there is disparity between the cultures of the teachers and students, the prevalence of local people in this remote Aboriginal school helped dissipate what [Delpit \(1993\)](#) describes as a “culture of power” ([Tjitayi and Osborne, 2014](#), p. 29). In many ways this had a significant impact on the relationship between the school and the community, creating the context for greater interaction between the staff and family members who, over time, built the capacity of people in the community to enhance the education of their children.

Using information technology and social media for community based action research

The above narrative, in which a group of local Aboriginal women were able to communicate stories and presentations about their activities provides an example of the striking developments in information technology and social media that have emerged in recent years. [Kent et al. \(2021\)](#) review the extensive possibilities whereby social media applications can be used by educational practitioners to engage students, parents, community leaders and local people in their projects. A small sample of these uses of technology includes the following:

- [Annamalai \(2018\)](#) reports on the use of WhatsApp to investigate pre-service teachers’ online activity.
- [Arihastha and Basthomit \(2019\)](#) describe how project -based learning, process writing and Instagram were used to enhance student writing skills.
- [de Jager et al. \(2017\)](#) review the different ways of digital storytelling was used to enrich their research.
- [Latz \(2017\)](#) provides a practical guide to the use of Photovoice in education, while [Ha and Whittaker \(2016\)](#) provide a case study describing how this application was used to enable children with autism spectrum disorder to tell their stories.
- [Sirin et al. \(2018\)](#) describe digital game-based approaches to education for vulnerable Syrian refugee children.
- [Pereira et al. \(2019\)](#) describe the use of Second Life to develop clinical reasoning skills.
- [Saeed and Yang \(2008\)](#) describe how blogs, social bookmarks, and podcasts can be incorporated into unit teaching.

- Raj et al. (2019) suggest processes for using YouTube as a platform to learn social expression.

A strand of projects involving social media demonstrate how research can be used for purposes of social activism:

- Kunnath and Jackson (2019) describe the development of student critical consciousness, using Twitter as a tool to apply critical literacy in the English classroom.
- Mayorga (2014) describes movement toward digital, critical, participatory action research.
- Bogle (2019) describes how kids are using TikTok to politically organize, a high-profile example of this emerging during the run-up to the 2020 presidential elections in the US.
- Mitchell et al. (2017) also describe how participatory visual methodologies facilitate social change through community and policy dialogue.
- Trott (2019) describes the processes of collaborating with children for community-based climate change action.
- These activist-oriented approaches to educational development have associated theoretical framings, such as those presented by Orłowski (2019) who describes how Freirean conceptions of participatory inquiry may be used for teaching about social justice.

More broadly, Stringer (2015) and Stringer and Ortiz (2021) provide examples of highly successful community-based action research projects that used multi-media and performative presentations to achieve strong outcomes for Indigenous populations in Australia, Hispanic groups in the US and school development in East Timor. In these and many other contexts social media and information technology provide the means to strengthen links to local communities and to include stakeholding audiences – students, colleagues, administrators and community members – in rich and effective educational projects.

Culture, curriculum and pedagogy

The development of culturally appropriate education practices requires a deeper understanding of the nature of culture that would enable local schools to modify and adapt centrally devised curricula to the realities of everyday life of local families and communities. There is a sound theoretical basis for accepting the logic of this dynamic, since child development has long been recognized as being intimately associated with culture, the way children think and learn – their cognitive development – being primarily linked to the context of their cultural environment.

Piaget described how children build cognitive structures – mental maps, schema, or networked concepts – by responding to experiences within their environment: people, objects, language and culture (Piaget and Inhelder, 1969; Wadsworth, 2019). From Bruner's perspective, culture shapes the mind and provides the toolkit from which children construct their worlds and a conception of themselves (Bruner, 1974, 1997; Illeris, 2018). Perhaps more specifically, Vygotsky saw culture as the prime determinant of a child's cognitive development, and language and social interaction as key processes associated with learning (Vygotsky et al., 1978; Vygotsky and Kozulin, 2012; Moll, 2013). The culture that children learn provides the particular schema through which they are able to understand the world in which they live by engaging the enduring patterns of thinking, behaving, believing and valuing that are largely established in the first years of their lives.

Culture, for each individual in a social group, includes an embodied understanding of what *is* – concepts that define reality, what *can be* – systems of belief, what *should be* – systems of value, what *to do* – acceptable behaviors and actions, and *how to do it* – acknowledged methods or recipes for accomplishing a set of designated outcomes. Culture, however, not only provides the context by which people act out their day to day lives, but it is also a deeply embodied set of precepts through which people define who they are – their identity. This, in turn, provides them with the basis for defining their inner life—their concepts of self and identity—issues that cannot be taught as “lessons” but that are implicit in the ongoing interactions and activities embedded in their social lives.

At all levels of education, current procedures for establishing both curriculum content and teaching/learning processes tend to ignore these fundamental features of the students life-world. Curriculum development is treated as highly centralized activity by state, regional or national education authorities, with panels of experts and community leaders establishing content areas with an increasing emphasis on specifically designated pedagogical processes – direct instruction, testing regimes and so on. School systems establish parameters for the implementation of a state/regional curriculum and administrators and teachers at the local level work within the guidelines thus established (See Fig. 3). The result is an educational experience established according to the expertise and cultural life-world of a largely professional elite that often has very little understanding of the cultures that daily walk into the schools and classrooms in the form of the children, the living manifestation of the diverse cultures that comprise modern social contexts. In such circumstances the “cultural fit” of curriculum and pedagogy will be determined by the extent to which a student's cultural world is consonant with the cultural world of those responsible for developing the curriculum and the teachers and school administrators responsible for its implementation.

Further, however, those responsible for implementing these educational arrangements are subject to continuing pressures to comply with directives to organize teaching and learning according to issues dictated more by the financial and organizational needs of the institutions than the realities of student educational, social and cultural needs. Despite these restrictions, however, participatory and community-based action research processes have the potential to enable teachers and administrators to implement educational processes that are relevant, engaging and educationally productive. They hold promise of providing educational programs not only meaningful in terms of employment, but in service of the more human features of social life that value the identities of the children and provide the creative and artistic energy so essential to people's well-being (See Fig. 4).

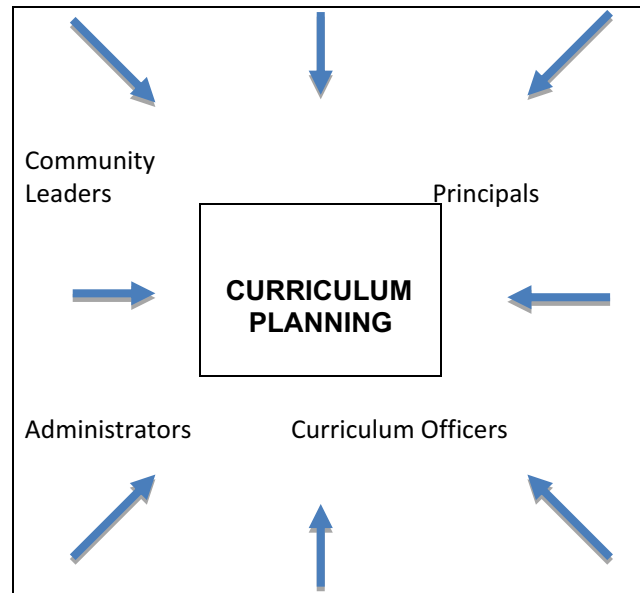


Fig. 3 Curriculum planning.

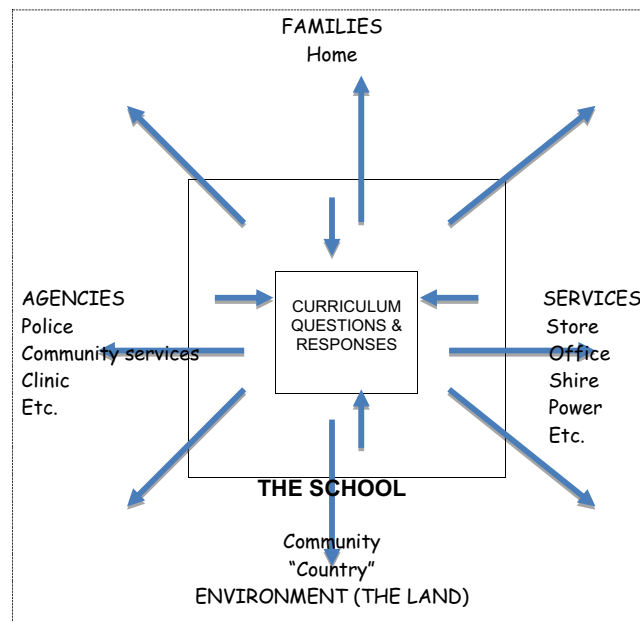


Fig. 4 Community-based curriculum planning.

As previous sections of this chapter suggest, participatory and community-based approaches to inquiry have the potential to inform administrators and curriculum leaders of the social and cultural make-up of the community, an essential aspect of a rich and comprehensive curriculum. This information is unlikely to be found in libraries or any other collection of texts, since the narratives associated with reports and descriptive materials are essentially limited, providing little of the empathetic understanding of the diverse life-worlds of the community that could provide the basis for a truly human and democratic curriculum.

Organizational and systems change: implementing holistic developmental processes

The issues raised in the opening section of this chapter have been amplified in recent decades by the gradual integration of a powerful ideology into the governance and operation of social systems across the world. A range of intellectual commentaries describe how neoliberal ideologies promote visions of social life determined by economic imperatives driven by market-based social

processes in which: “the techniques and values of accountancy have become a central organizational principle in the government and management of human conduct – and the new kinds of relationships, habits and practices this is creating.” (Denzin and Giar-dina, 2017). The language and processes of accountancy thus provide the basis for an audit culture in which the logics of the market structure life in the new neoliberal education system (Smart, 2010).

School systems likewise have been impacted by the application of neo-liberal ideologies. Within education, performance metrics on nationally standardized tests are used to calculate status in local and national rankings, providing the means to evaluate the well-being and worth of each school and each group of students. In some contexts, these metrics also provide the basis for determining not only the rank of the school, but also at the level of the individual, rewards in the form of promotion and/or salary. In some instances, failure to accomplish expected outcomes in school settings, may result in reduced funding for the school or reduced prospects/salary for individual teachers and administrators.

These processes, now well established in many parts of the world, have made little difference to the inequitable outcomes of existing education systems, despite sometimes massive efforts in the form of special programs and services. Ultimately the systems look for change in the pedagogical or organizational arrangements within schools, or place blame on teachers, entire schools, students or their families. What we suggest, however, is the need for rearrangements that more clearly fit the systems to the educational, social and cultural realities that make up the life-worlds of the students, rather than the reverse. This means, in essence, that those responsible for the operation of school systems invest some of their time and energy in understanding the social and cultural realities of people at all levels and groups within their locality in order to incorporate the strengths of their communities into the curriculum and the operation of the system.

Change, however, is something that cannot be achieved by focusing on classroom pedagogical practices, but needs to reach into all levels of a system (Harrison et al., 2019) rather than interpreted as something that can be added to existing arrangements, sometimes referred to as “spray-on” or “silver bullet” solutions.

Kelly and Westoby (2018) provide a powerful framework that describes the processes required to integrate changes through all levels of any social system:

- *Implicate level*, in which individuals reflect on changes that are needed in their own approaches to professional practice
- *Micro-level* which entails building purposeful relationships that enable people to interact productively with the diverse people they engage in their work and community contexts
- *Mezza-level* refers to effective ways of engaging participative processes with small groups
- *Macro-level* focuses on ways of forming and operating strong organizations that enable people to accomplish large-scale projects
- *Meta-level* development entails the integration of personal, group and organisational processes to accomplish large-scale movements at national and international scales, often accompanied by philosophical and ideological developments.

Though there is a tendency for people to state that “we’re already doing that”, developmental processes actually require carefully articulated human technologies, enabling practitioners and administrators to modify and adapt the centralized, hierarchical approaches to organization and development currently dominating many areas of social life. There is a general literature that supports these approaches to social, organizations and systems change. Checkland and Holwell (1997) describe the use of information systems to make sense of the field; Checkland (2000) reviews the use of soft systems methodologies over a thirty year period; Checkland and Poulter (2006) show how systems methodology might be used by professional practitioners, teachers and students; and Flood (2010) and Ison (2008) illuminate the relationship between action research and systems thinking. All point to the need for systematic professional and organizational development across all levels of a system.

Ethics in participatory and action research

Ethical issues have been particularly problematic in community-based, participatory and action research, since these approaches to research seem to contravene a number of issues previously thought to be basic to the uninvolved stance required of rigorous experimental and quasi-experimental research. A key challenge in participatory research related to the need for the development of appropriate protocols and practices for ethics, the need to balance participatory processes with the need for individual safety, security and confidentiality require carefully articulated ethical procedures. Furthermore, issues related to duty-of-care for research participants and the need for confidentiality are both seemingly problematic in action research, where sharing information and systematic interaction between participants often is essential. More recently academic researchers have presented ways in which academic ethical requirements can be accomplished systematically within an action research process. Banks and Brydon-Miller (2018) present a detailed analysis of issues related to ethics in participatory approaches to research, while Brydon-Miller et al. (2015) demonstrate how structured ethical reflection can widen the circle of participants and increase the power of an action research process.

Conclusion

Continuing movements in education to centralize decision-making and standardize curriculum and pedagogical arrangements have failed to produce improvements in educational outcomes for students from marginalized and culturally diverse social groups. Market-based educational arrangements continue to undermine the possibilities for more humanly productive education systems

that improve the social and cultural well-being of the diverse cultures they serve. Despite these restrictions we contend that the techniques and technologies of the current era provide many possibilities for local-level, participatory approaches to research that builds on the strengths of local communities and provides a more relevant, engaging and productive experience for all concerned – students, teachers, administrators, families and community members. These interactive approaches to education not only strengthen the educational intents of schooling, but also provide the means to develop the capacities of individuals, families and the broader community to deal with the complex web of issues that characterize the reality of today's social worlds. More particularly, they also provide the means to explore the possibilities of opening the rigidities and power-driven processes of many education systems to the creative and life-enhancing qualities inherent in the human spirit. Our experience suggests that even small projects have the potential to extend their power and reach to eventually be the seed from which significant changes emerge. In the words of an Australian songwriter, "From little things, big things grow."

References

- Abma, T., Banks, S., Cook, T., Dias, S., Madsen, W., Springett, J., Wright, M.T., 2018. *Participatory Research for Health and Social Well-Being*. Springer.
- Alexander, K., Morgan, S., 2016. The Coleman Report at Fifty: Its Legacy and Implications for Future Research on Equality of Opportunity. *Russell Sage Foundation Journal of the Social Sciences*. https://socweb.soc.jhu.edu/faculty/morgan/papers/Alexander_and_Morgan_2016.pdf.
- Annamalai, N., 2018. How do we know what is happening in WhatsApp: a case study investigating pre-service teachers' online activity. *Malays. J. Learn. Instr.* 15 (2), 207–225.
- Arihata, D., Basthomi, Y., 2019. Project-based learning, process writing, and Instagram: an attempt to enhance the students' writing skill. *J. Pendidikan Humaniora* 7 (3), 12606. Retrieved from: <http://journal.um.ac.id/index.php/jph/article/view/>.
- Banks, S., Brydon-Miller, M., 2018. Ethics in participatory research. In: Banks, S., Brydon-Miller, M. (Eds.). *Routledge*, pp. 1–30.
- Barrett, J.R., 2007. The researcher as instrument: learning to conduct qualitative research through analyzing and interpreting a choral rehearsal. *Music Educ. Res.* 9 (3), 417–433. <https://doi.org/10.1080/14613800701587795>.
- Beadle, R., 2018. *Remote Aboriginal Women and Meaningful Work: Key Dimensions for Ngaanyatjarra Women*. Ph.D. Dissertation. University of Melbourne.
- Biesta, G., 2020. *Educational Research: An Unorthodox Introduction*. Bloomsbury.
- Boden, R., Greenwood, D.J., Hall, B., Levin, M., Marshall, J., Wright, S., 2015. Action research in universities and higher education worldwide. In: Bradbury, H. (Ed.), *The SAGE Handbook of Action Research*. Sage, pp. 281–290. <https://doi.org/10.4135/9781473921290.n28>.
- Bogle, A., September 19, 2019. Free From Adults, Kidsare Using TikTok to Politically Organise. ABC News. Retrieved from: <https://www.abc.net.au/news/2019-09-19/tiktok-youth-led-climate-activism-schoolstrike/11520474>.
- Boudah, D., 2019. *Conducting Educational Research: Guide to Completing a Thesis, Dissertation or Action Research Project*. Sage.
- Bradbury, H., 2015. *The SAGE Handbook of Action Research*, third ed. Sage.
- Bruner, J., 1974. *Toward a Theory of Instruction*. Belknap Press.
- Bruner, J., 1997. *The Culture of Education*. Harvard University Press.
- Brydon-Miller, M., Rector Aranda, A., Stevens, D., 2015. Widening the circle: ethical reflection in action research and the practice of structured ethical reflection. In: Bradbury, H. (Ed.), *The SAGE Handbook of Action Research*, third ed. Sage, pp. 596–607.
- Calhoun, E., 1993. Action research: three approaches. *Educ. Leader* 51 (2), 62–65.
- Charnaz, K., Harris, S.R., Irvine, L., 2019. *The Social Self and Everyday Life: Understanding the World Through Symbolic Interactionism*. Wiley-Blackwell.
- Checkland, P., 2000. Soft systems methodology: a thirty year retrospective. *Syst. Res. Behav. Sci.* 17 (S1), S11–S58. [https://doi.org/10.1002/1099-1743\(200011\)17:1+<::AID-SRES374>3.0.CO;2-O](https://doi.org/10.1002/1099-1743(200011)17:1+<::AID-SRES374>3.0.CO;2-O).
- Checkland, P., Holwell, S., 1997. *Information, Systems and Information Systems: Making Sense of the Field*. Wiley.
- Checkland, P., Poulter, J., 2006. *Learning for Action: A Short Definitive Account of Soft Systems Methodology and its Use for Practitioner, Teachers, and Students*. Wiley.
- Chevalier, J.M., Buckles, D., 2019. *Participatory Action Research: Theory and Methods for Engaged Inquiry*, second ed. Routledge, Taylor & Francis.
- Close the Gap Campaign Steering Committee, 2016. *Close the Gap: Progress and Priorities Report 2019*. Close the Gap Campaign Steering Committee. www.humanrights.gov.au/social-justice/health/index.html.
- Coghlan, D., Brydon-Miller, M., 2014. *The SAGE Encyclopedia of Action Research*. Sage. <https://doi.org/10.4135/978146294406>.
- Cole, M., 2011. What's culture got to do with it? Educational research as a necessarily interdisciplinary enterprise. *Educ. Res.* <https://doi.org/10.3102/0013189X10380247>.
- Coleman, J., 1966. *Equality of Educational Opportunity*. U.S. Government Printing Office.
- Cresswell, J., Poth, C., 2017. *Qualitative inquiry and research design: choosing among five approaches*, fourth ed. Sage.
- de Jager, A., Fogarty, A., Tewson, A., Lenette, C., Boydell, K., 2017. Digital storytelling in research: a systematic review. *Qual. Rep.* 22 (10), 2548–2582.
- Delpit, L., 1993. The silenced dialogue: power and pedagogy in educating other people's children. In: Weis, L., Fine, M. (Eds.), *Beyond Silenced Voices: Class, Race, and Gender in United States Schools*. State University of New York Press.
- Denzin, N., Giardina, M. (Eds.), 2017. *Qualitative Inquiry in Neoliberal Times*. Routledge, NY.
- Denzin, N.K., Giardina, M.D. (Eds.), 2018. *Qualitative Inquiry in the Public Sphere*. Routledge.
- Denzin, N., Lincoln, Y., 2019. *Handbook of Qualitative Research*. Sage.
- Efron, S.E., Ravid, R., 2019. *Action Research in Education: A Practical Guide*, second ed. Guilford Press.
- Flood, R.L., 2010. The relationship of "systems thinking" to action research. *Syst. Pract. Action Res.* 23 (4), 269–284. <https://doi.org/10.1007/s11213-010-9169-1>.
- Greenwood, D.J., Whyte, W.F., Harkavy, I., 1993. Participatory action research as a process and as a goal. *Hum. Relat.* 46 (2), 175–192. <https://doi.org/10.1177/001872679304600203>.
- Ha, V.S., Whittaker, A., 2016. "Closer to my world": children with autism spectrum disorder tell their stories through PhotoVoice. *Global Publ. Health* 11 (5–6), 546–563. <https://doi.org/10.1080/17441692.2016.1165721>.
- Harrison, R., Blickern, C., Lamb, J., Kirk, S., Vassilev, I., 2019. Asset-based community development: narratives, practice, and conditions of possibility—a qualitative study with community practitioners. *Sage Open* 9 (1). <https://doi.org/10.1177/2158244018823081>, 2158244018823081.
- Hendricks, C., 2016. *Improving Schools Through Action Research: A Reflective Practice Approach*, fourth ed. Pearson.
- Heron, J., 1996. *Co-operative Inquiry: Research Into the Human Condition*. Sage.
- Illeris, K., 2018. *Contemporary Theories of Learning*, second ed. Routledge.
- Ison, R., 2008. Systems thinking and practice for action research. In: Reason, P., Bradbury, H. (Eds.), *The SAGE Handbook of Action Research*. Sage, pp. 139–158. <https://doi.org/10.4135/9781848607934.n15>.
- Johnson, R., Christensen, I., 2019. *Educational Research: Quantitative, Qualitative and Mixed Approaches*, seventh ed. Sage.
- Kelly, A., Westoby, P., 2018. *Participatory Development Practice: Using Traditional and Contemporary Frameworks*. Practical Action.
- Kemmis, S., McTaggart, R., Nixon, R., 2014. *The Action Research Planner: Doing Critical Participatory Action Research*. Springer.

- Kent, M., et al., 2021. Action research and digital media. In: Stringer, E., Ortiz Aragon, A. (Eds.), *Action research*, fifth ed. Sage.
- Kunnath, J.P., Jackson, A., 2019. Developing student critical consciousness: Twitter as a tool to apply critical literacy in the English classroom. *J. Media Lit. Educ* 11 (1), 52–74. <https://doi.org/10.23860/JMLE-2019-11-1-3>.
- Latz, A.O., 2017. *Photovoice Research in Education and Beyond: A Practical Guide From Theory to Exhibition*. Routledge, Taylor & Francis.
- Mayorga, E., 2014. Toward digital, critical, participatory action research: lessons from the #BarrioEd- Proj. *J. Interact. Technol. Pedagog.* (5) Retrieved from <https://jitp.commons.gc.cuny.edu/toward-digital-critical-participatory-action-research/>.
- McNiff, J., 2017. *Action Research: All You Need to Know*. Sage.
- Mertler, C. (Ed.), 2019a. *The Wiley Handbook of Action Research in Education*. Wiley-.
- Mertler, C., 2019b. *Action Research: Improving Schools and Empowering Educators*, sixth ed. Sage.
- Mills, G., 2017. *Action Research: A Guide for the Teacher Researcher*, sixth ed. Pearson.
- Mills, G., Gay, L., 2018. *Educational Research: Competencies for Analysis and Applications*, twelfth ed. Pearson.
- Minutjukur, M., Osborne, S., 2014. At the heart of learning: from cultural devastation to cultural re-invention. *AlterNative* 10 (1), 15–22, 1.
- Mitchell, C., de Lange, N., Moletsane, R., 2017. *Participatory Visual Methodologies: Social Change Through Community and Policy Dialogue*. Sage.
- Moll, L., 2013. *L.S. Vygotsky and Education*. Routledge.
- Nah, K.-O., Lee, S.-M., 2016. Actualizing children's participation in the development of outdoor play areas at an early childhood institution. *Action Res.* 14 (3), 335–351. <https://doi.org/10.1177/1476750315621610>.
- Orlowski, P., 2019. Freirean conceptions of participatory action research and teaching for social justice—same struggle, different fronts. *Can. J. Action Res.* 20 (1), 30–51. <https://doi.org/10.1111/j.1759-5436.2012.00321.x>.
- Pereira, M.M., Artemiou, E., McGonigle, D., Köster, L., Conan, A., Sithole, F., 2019. Second life and classroom environments: comparing small group teaching and learning in developing clinical reasoning process skills. *Med. Sci. Educ.* 29 (2), 431–437. <https://doi.org/10.1007/s40670-019-00706-4>.
- Piaget, J., Inhelder, B., 1969. *The Psychology of the Child*. Basic Books.
- Putman, S.M., Rock, T., 2017. *Using Strategic Inquiry to Improve Teaching and Learning*. Sage.
- Raj, S., Ann, W., Subramaniam, P., Yunus, M., 2019. Using YouTube as a platform to learn social expression. *Creativ. Educ.* 10 (2), 288–296.
- Saeed, N., Yang, Y., 2008. Incorporating blogs, social bookmarks, and podcasts into unit teaching. *Proc. Tenth Conf. Australas. Comput. Educ.* 78, 113–118.
- Sagor, R., Williams, C., 2016. *The action research guidebook: a process for pursuing equity and excellence in education*, third ed. Corwin.
- Sarason, S., 1990. *The Predictable Failure of Educational Reform*. Jossey Bass.
- Schubotz, D., 2019. *Participatory Research: Why and How to Involve People in Research*, first ed. Sage.
- Sirin, S., Plass, J.L., Homer, B.D., Vatanartiran, S., Tsai, T., 2018. Digital game-based education for Syrian refugee children: project hope. *Vulnerable Child. Youth Stud.* 13 (1), 7–18. <https://doi.org/10.1080/17450128.2017.1412551>.
- Smart, B., 2010. *Consumer Society: Critical Issues and Environmental Consequences*. Sage, Thousand Oaks, CA.
- Stringer, E., 2008a. *Action Research in Education*, second ed. Pearson.
- Stringer, E., 2008b. "This is so democratic." Action research and policy development in East Timor. In: Reason, P., Bradbury, H. (Eds.), *The SAGE Handbook of Action Research*, second ed. Sage.
- Stringer, E., 2015. Achieving equity in education. In: Bradbury, H. (Ed.), *The SAGE Handbook of Action Research*, third ed. Sage. <https://doi.org/10.4135/9781473921290.n35>.
- Stringer, E., Ortiz, A., 2021. *Action Research*, fifth ed. Sage.
- Sutter, W., 2011. *Introduction to Educational Research: A Critical Thinking Approach*, second ed. Sage.
- Tijitay, K., Osborne, S., 2014. At the heart of learning: holding knowledge in our spirit. *AlterNative* 10 (1), 23–32.
- Trott, C.D., 2019. Reshaping our world: collaborating with children for community-based Climate change action. *Action Res.* 17 (1), 42–62. <https://doi.org/10.1177/1476750319829209>.
- Vygotsky, L., Cole, M., John-Steiner, V., Scribner, S., Souberman, E., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Vygotsky, L., Kozulin, A., 2012. *Thought and Language*. MIT Press.
- Wadsworth, B., 2019. *Piaget's Theory of Cognitive and Affective Development: Foundations of Constructivism*, fifth ed. Pearson.
- Wells, C.G., 1994. *Changing Schools From Within: Creating Communities of Inquiry*. Heinemann.
- Wood, L., Govender, B., 2013. "You learn by going through the process." the perceptions of South African school leaders about action research. *Action Res.* 1, 2. June 2013. Sage.

Affirming methodologies—the Caribbean oral tradition of Ole Talk

Camille Nakhid^a and Anabel Fernández Santana^{b, a} Auckland University of Technology, Auckland, New Zealand; and ^b COMET - Te Hononga Akoranga, Auckland, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	179
Defining Affirming Methodologies	179
The significance of oral traditions, and as affirming research methodologies in education	180
Oral traditions in education research	180
Storytelling: embodying the oral traditions	181
Ubuntu storytelling	181
Pūrākau	182
Talanoa	182
Talking Circle	182
Yarning	182
Halaqah	182
Digital storytelling	183
Ole Talk as oral research method	183
Ole Talk—a Caribbean form of communication	183
Ole Talk in the research environment	183
Ole Talk in practice	184
Recording Ole Talk	184
Transcribing Ole Talk	184
Analyzing and reporting Ole Talk	186
Application of Ole Talk in research	186
Conclusion	186
References	187

Introduction

Oral traditions, including storytelling through folktales, legends, jokes, songs and chants, proverbs or wise sayings, and mythologies, have existed for millennia as vehicles for intergenerational transmission of culture and knowledge. Indigenous and local oral traditions are considered to be less prescriptive in their approach, embody a more relational epistemology, and balance the power dynamics between the researcher and those who share their knowledge. The researcher becomes a knowledge sharer as well as a knowledge gatherer and is likely to be grounded in the culture in which these traditions take place (Nakhid, 2021; Fernández, 2020). In recent years, oral traditions have been used to inform education research, providing an inclusive approach to engaging with academic communities where orality is a salient way of constructing and sharing knowledge. The use of oral traditions in research enables agency for those involved in this process (Nakhid, 2021). As such, oral traditions embed an Affirming Methodologies approach to research by affirming practices that are grounded in peoples' lived realities and worldviews.

An Affirming Methodologies framework centers local and indigenous practices of knowledge construction on the premise that they are likely to result in greater authenticity and accuracy. Nakhid (2021) and Nakhid-Chatoor et al. (2018) argue that local and indigenous research methodologies should present from within their own living experiences and realities—even as they hold an awareness of the impact of western and colonial research practices.

The first part of this paper defines the concept of Affirming Methodologies. Selected oral traditions are then conceptualized as affirming methodologies that ground research in the culture, lived experience and worldviews of those involved. Ole Talk, a Caribbean oral tradition that has been recently explored and utilized as a research method, is described to illustrate its potential for enabling culturally affirming knowledge construction and sharing.

Defining Affirming Methodologies

Affirming Methodologies is a theoretical, epistemological, and methodological concept which frames indigenous and local practices of knowledge construction as valuable in and of themselves. An affirming research methodology is centered on the strength and uniqueness of local and indigenous knowledges and emphasizes the value of culturally relevant and local ways of knowing (Nakhid and Farrugia, 2021). When conducting research with Indigenous, Black and People of Color, Affirming Methodologies enables

awareness and acknowledgment of how their histories, traditions, and local evolution have shaped their knowledges. The research approaches are grounded in practices that are culturally relevant and suited to their own cultures and the cultures of those with whom they work (Te Awakotuku, 1991; Castellano, 2014). Oral traditions and methods are seen as affirming methodologies when they elicit authentic knowledge that resonates with communities and are of value and relevance for the community in which the research takes place.

Indigenous communities have a right to select their preferred method of data collection (Drawson et al., 2017). Some existing research practices have been modified to meet the needs of communities incorporating “indigenous values, beliefs, and ways of knowing, facilitating research that is respectful, collaborative, and relational” (Wilson, 2001). Methodologies that are affirming utilize the specific modes of being and knowing of the communities where research operates, and their methods, techniques and tools respond to the continually evolving realities, practices, experiences, and aspirations of those communities.

Affirming Methodologies acknowledges and celebrates intersections and solidarities with other communities (Fernández Santana et al., 2019). With an affirming methodological approach to research, culturally relevant ways of knowing are presented as valuable within themselves and not in comparison to dominant Eurocentric approaches to research. Affirming Methodologies centers local voices previously excluded from mainstream research and allows for the building and developing of relationships among knowledge sharers (Nakhid, 2021).

The significance of oral traditions, and as affirming research methodologies in education

Oral traditions are classified by UNESCO as part of a country’s intangible cultural heritage, along with diverse practices that range from traditional dances to weaving and fishing (UNESCO, n.d.). The word *oral* gives the emphasis to oral tradition and oral narratives and, although it sometimes refers to speech and words, it needs to be applied cautiously so as not to exclude sign language and body language (Finnegan, 2003). Oral traditions are the memory of a people encompassing those aspects of life regarded as important within their culture. Tradition is a term frequently used by anthropologists, folklorists, and oral historians to refer to culture or an established way of doing things; the process of handing down practices, ideas or values; or to mark out a group’s identity. Oral traditions can be described as cultural heritage and practices that are shared and passed down intergenerationally through non-written communication (e.g. storytelling, conversation, sayings, songs, etc.). Oral practices are valued and important components of a culture and can be defined as traditions when they present a long-standing legitimacy in that community and are shared widely among its members (Finnegan, 2003).

Oral methodologies and methods are frequently preferred in research with indigenous and non-western communities, or those who have had their cultures and lives misrepresented by western mainstream research practices. Oral methodologies contribute to ensuring that a research practice is inclusive, respectful, and affirming of local and indigenous beliefs and communities’ ways of living and knowing. This has the potential to generate more authentic and locally grounded research results that can be accessed and used by the community where the knowledge was generated, thereby preserving ownership and agency. Mahuika (2012) argues that we need to navigate and analyze our world from our own knowledges and understandings if we are to have a true assessment of our realities.

The use of oral traditions as culturally affirming approaches to knowledge construction constitutes a shift away from mainstream research as follows: First, oral traditions are firmly grounded in the realities and communities from which they emerge. Consequently, when oral traditions are used in research, participants are likely to be familiar and comfortable with the modes of engagement. This creates an environment where subjects can draw on their cultural competencies and strengths to engage in knowledge construction as active agents.

Second, although they are dynamic and constantly evolving, oral traditions typically encompass practices and worldviews shared by their community. Thus, when oral traditions are used as a foundation to design and conduct research processes, the researcher is de-centered and is no longer the sole custodian of the “rules” of knowledge construction, but a subject building knowledge alongside other subjects.

Third, oral traditions are authentically interwoven in a culture or community. When research is designed around a community’s oral traditions, the phases of research flow organically. This avoids the need to “bridge” foreign engagement practices with local ways, a difficult and ethically questionable task for researchers, especially for those working in their own community employing culturally unrelated western mainstream research methodologies.

Oral traditions in education research

For Indigenous, Black and students of color, self-determination and identity affirmation are inherent in oral research methodologies and enable these students to politically express and assert themselves in what might be a hostile classroom environment (Mahuika, 2012). Oral research methods allow school communities to see the different ways that students perceive the world in which they live (Mahuika, 2012). They have the potential to be transformative for those students and staff typically marginalized by the school curricula, administrative systems, and regulations by opening the way for those in dominant positions to conscientize themselves and challenge the policies and processes in place that have constrained the opportunities of certain groups (Navabi, 2006).

Oral research methods give students the opportunity to use poems, legends, songs, and myths to share their experiences in a manner that more acutely expresses their academic life (Kargbo, 2008). Oral research methods allow for information to be validated by other students in a non-confrontational setting and give rich data that may not be possible to obtain through quantitative research methods (McAdoo, 1980). Oral research methods can also study interactions among students, and between students and teachers in group settings such as liming whereby ole talk can be used to address conflict and negotiate expectations (Nakhid, 2022).

For many of those failed by an oppressive education system, oral traditions have been their ways of passing on knowledge through the generations. The recognition for these students that comes from knowing that oral traditions illustrate a formidable and enduring mental capacity and fortitude, and that they are invaluable and sacred approaches to holding and sharing trusted and sustainable knowledge, is critical to how they value and advocate for their voices in the academy (Struthers and Peden-McAlpine, 2005). More so, oral traditions used in education research studies connect and strengthen solidarity among different marginalized groups. Critical race methodology sees oral research methods as a way that students of color can reveal and document their education journeys to benefit not only themselves but those students that come after, so that they are aware and knowledgeable of those and their systems that have attempted to marginalize and silence them (Solórzano and Yosso, 2002).

Storytelling: embodying the oral traditions

The use of oral traditions as research methodologies in education and other disciplines is underpinned by the notion that storytelling and conversation are powerful means of constructing, sharing, and preserving knowledge. Storytelling is seen as one of the more common research methods associated with oral traditions and has become increasingly valued and employed in contexts where Eurocentric ways of knowing are seen as culturally irrelevant or inappropriate. Storytelling is invaluable to how knowledge is sustained and protected within Indigenous communities. A story tells who we are and are “stories of origin and of ancestors, world view, values and knowledge for everyday survival” (Koukkanen, 2000, p. 421). Traditional storytelling is a “way of expressing Indigenous knowledge, culture, and oral traditions that privileges holistic interconnectedness, collaboration, reciprocity, spirituality, and humility” (Datta, 2018, p. 36). In Indigenous culture, stories are a common repository of knowledge and facilitate the process of knowing (Ware et al., 2018).

In recent years, Indigenous scholars in education have taken an interest in storytelling (Archibald, 2008; Kovach, 2009) and many indigenous research methods and methodologies involving storytelling have been documented and disseminated (Kovach, 2010). In research using storytelling, the relationship and what is created between the researcher and participant are important considerations (Dyll-Myklebust, 2014). Storytelling and conversations are seen as culturally affirming research methods in the way that they center local knowledge and ways of knowing.

Rodriguez (2010) has observed that storytelling can assist researchers build more meaningful relationships with students of color. These relationships can enable researchers to disrupt any racist messages that students of color might receive in white school settings and to share their own experiences of racism without minimizing but rather validating the students’ experiences. Counterstories, as Solórzano and Yosso (2002) suggest, can be used in education research with students of color to challenge racism, sexism, and classism and any tales of deficit that disadvantage the students.

Storytelling informs many of the oral research methods and methodologies currently used by local and indigenous communities, some of which are described below. Although this list is not exhaustive, the selection of these methodologies and methods is based on those characteristics similar in the practices across communities: authenticity, inclusiveness, shared connection, relationship, reciprocity, ownership, and equity.

Ubuntu storytelling

“Ubuntu storytelling is about engaging our relational selves” (Mucina, 2011, p. 1). Ubuntu, says Mucina, is a research method where we relate to each other through storytelling. The Ubuntu (people) use the art of oral storytelling because a “story can allow a culture to regenerate itself” (p. 1). Hall (1997) notes that the continual process of knowledge making is essential to living as a human being. Mucina (2011) is aware that indigenous Ubuntu knowledge production and governance are not exempt from oppressive and marginalizing practices, as is the case with the notion of indigenous knowledge itself. According to Dei (2000, p. 113), indigenous knowledges need to develop a critical stance toward “the production and validation of critical knowledge for decolonization purposes.”

During storytelling by Zulus, the audience become active participants who join in the singing and use facial expressions and gestures to coincide with what is taking place in the story. Ubuntu storytelling engages through interactive performance and both the storyteller and listener have specific functions to perform in the making of the story. Performing the story makes it whole in the Ubuntu structure and the repetitive nature of storytelling provides markers of important lessons for the listener.

Mucina (2011, p. 7) uses the Ubuntu structure of storytelling to make “Black people use their stories to talk to other Black people in complex, challenging and sometimes contradicting ways.” The narrative aspect of Mucina’s storytelling “takes the live performative part of storytelling and puts it into written form” so “that the story can be transported across the globe and read by a wider audience” (p. 8). In traditional Ubuntu orality, the storyteller does not analyze the story for the listener but leaves each listener

to analyze the story, knowing that each person will gain something different from it. Achebe (1988, p. 47) says that one of the “consequences of the transition from oral traditions to written forms of literature is the emergence of individual authorship.” That is, the story told by the storyteller no longer belongs to them once they have told it and will be re-composed in some form by the listener to whom it now belongs.

Pūrākau

Pūrākau, usually considered to be Māori myths and legends, was used as a research methodology by Lee (2009) to tell the stories of Māori teachers. Pūrākau is not commonly associated with academic writing or research methodology and has been “reconceptualised as a culturally responsive construct for narrative inquiry into Māori teachers” (Lee, 2009, p. 1). Pūrākau is a traditional form of Māori narrative containing “philosophical thought, epistemological constructs, cultural codes, and worldviews that are fundamental” to Māori identity (Lee, 2009, p. 1). It is a collection of traditional oral narratives, an anthology of literature that remains relevant today and can be “constructed in various forms, contexts and media” to better understand the experiences of Māori in a research context (Lee, 2009, p. 1).

Pūrākau as methodology “draws from and responds to the wider historical, social and political research contexts” and is part of the reclamation process of identifying the colonial influences of research on Māori cultural traditions (Lee, 2009, p. 1). Pūrākau existed for Māori prior to western-based treatment and offers *tangata whaiora* (mental health clients) a culturally specific and valid therapeutic process (Cherrington, 2003). As a kaupapa Māori methodology, it has been used in therapeutic clinical settings with Māori *tangata whaiora* and their *whānau* (family). Using pūrākau as therapy, *tangata whaiora* and their *whānau* have the opportunity to “retell the story using waiata, haka, poetry, drama, sculpting, painting, drawing, storytelling, and/or writing” Cherrington (2003, p. 118).

Talanoa

An oral research method unique to and employed by Pacific peoples is Talanoa (Suaalii-Sauni and Fulu-Aiolupotea, 2014). Talanoa is a “traditional way of developing and revisiting knowledge within Pacific cultures” through what is and was told (Vaka et al., 2016, p. 539; Fonua, 2005). It is an oral research method that enables a cultural synthesizing of participants’ stories in a shared setting (Suaalii-Sauni and Fulu-Aiolupotea, 2014). Talanoa, a Pacific word with similar meanings across many Pacific nations, can simply mean talking but can include cultural ways of talking including why, where, how we talk, and to whom (Vaka et al., 2016). Talanoa is influenced by the “social status, customs and rituals, language, personality and professionalism of the convenor and participants” (Vaka et al., 2016, p. 538). Within the Talanoa, the inclusivity of the participants is balanced with arriving at a consensus within the group. Stories from the heart are explored in a Talanoa, and meanings are unpacked in a friendly and informal approach (Halapua, 2002). Talanoa discussions can be wide-ranging without restrictions on the chosen topic. All perspectives and contributions are welcomed, and the diverse ways in which participants wish to talanoa are accepted.

Talking Circle

Talking Circles are a traditional pattern of intragroup communication specific to Native American communities. Talking Circles have been attributed to Woodland tribes in Midwest United States where it was used as a “form of parliamentary procedure” (Wolf and Rickard, 2003, p. 39). The circle itself is considered sacred and to represent the interconnectedness of all things (people, earth, moon, sun). Talking Circles incorporate symbolic traditions such as burning sage, using eagle feathers, or passing a sacred rock to symbolize a person’s turn to talk. Individuals are given the opportunity to express their thoughts on an issue in both large and small groups by going around the circle and acknowledging the value of each speaker until a collaborative consensus is reached (Sams and Childers, 1990 cited in Wolf and Rickard, 2003). Traditionally, many Native American communities have used the talking circle to bring people together to learn, listen, and teach (Wolf and Rickard, 2003). Talking Circles or Healing Circles are rooted in traditional storytelling and religious ceremonies. These circles have re-emerged among Native Americans to share information, offer support, or solve problems in a group setting and allow for conversations and stories from participants on uncomfortable subjects (Becker et al., 2006).

Yarning

Yarning describes a culture-specific type of conversation among indigenous people in the country that is called Australia (Fredericks et al., 2011). Yarning is more than a light exchange of words and pleasantries in casual conversation (Walker et al., 2014, p. 2). A yarn is both a process and an exchange; it encompasses elements of respect, protocol, and engagement in individuals’ relationships with each other. Yarning establishes relationality and determines accountability and is understood as a means to share stories and listen reflectively to each other (Russell-Mundine, 2007). Yarning has also been differentiated as to whether it is social yarning, therapeutic yarning, collaborative yarning or yarning for more formal research (Bessarab and Ng’andu, 2010).

Halaqah

Halaqah is a spiritual circle-time used by early Muslims and is an integral part of traditional Islamic education. It is a traditional reflexive practice in Islamic cultures with the Qur’an as the reference point. Halaqah is more than a group interview or focus group. It is purely oral with students and teacher sitting in a circle on the floor with the aim of generating group-based narratives (Ahmed, 2014). Halaqah is credited with empowering individuals and communities through a social-justice agenda (Zaimache, 2002).

Digital storytelling

Digital storytelling is a form of storytelling using modern technology. It is the “process of illustrating personal narratives and stories with photographs, artwork, music, voice-overlay, video clips, and text” (Cunsolo Willox et al., 2013, p. 132). Digitalization attempts to avoid the limitations of interview-based narrative research whereby participants’ narratives are processed through the lens of the researcher. The digital process removes the researcher from the data collection stage through to dissemination (Cunsolo Willox et al., 2013). Digital storytelling is an emerging method that has the capability to “preserve and promote indigenous oral wisdom, while engaging community members, developing capacities, and celebrating myriad stories, lived experiences, and lifeworlds” (Cunsolo Willox et al., 2013, p. 128).

Ole Talk as oral research method

Ole Talk—a Caribbean form of communication

The Caribbean is a cultural space marked by cultural diversity, and ethnic and social hybridization and intermixing. Caribbean oral traditions can be found woven into the fabric of Caribbean peoples’ everyday lives, or articulated in highly localized and structured ceremonies. Oral transmission of knowledge is anchored in the ancestral traditions of the Indigenous peoples of the Caribbean as well as the diverse and ethnically mixed groups in the region. In the Caribbean, oral knowledge is transmitted in many forms, ranging from song lyrics to public speaking,¹ or from storytelling to lunch time arguments.

Ole Talk is a culturally specific way of orally transmitting knowledge in Trinidad and Tobago. Ole Talk is found throughout the region albeit with different names,² and takes place across ethnic, class, and geographic boundaries (Fernández, 2020). However, despite their prevalence and familiarity, these practices have been generally overlooked in the literature on Caribbean orality which has focused instead on ancestral storytelling, Creole languages, and music genres like Reggae and Calypso (Cambridge, 2005; Reddie, 1998; Vété-Congolo, 2016).

Ole Talk is a living tradition in Caribbean culture and inheres certain competencies that relate to the Caribbean experience. Some of these include (a) the ability to participate simultaneously in several conversations, (b) managing overlapping talk that is not to be interpreted as interruption, (c) the negotiation of meaning based on humor and performance, (d) interpretation and use of body language and silence to transmit meaning, and (e) the freedom to flow in and out of conversations depending on personal interest.

Ole Talk has frequently been associated with gossip or equated with small talk. However, as has been observed, Ole Talk can be a space where meaning is negotiated, and social discourses are elucidated and contested (Fernández Santana et al., 2019). Significantly, it is also a critical space for collective sense-making, community building and networking (Fernández, 2020). The following section outlines key features of Ole Talk as a Caribbean oral tradition and discusses how Ole talk has been reclaimed as a culturally affirming research method in Caribbean contexts (Fernández, 2020; Nakhid et al., 2020).

Ole Talk in the research environment

In a research context, Ole Talk can be defined as an unstructured oral practice that frequently encompasses storytelling, humor and performance, as well as a multi-layered conversation where participants can engage in multiple dialogues at the same time or none at all. Ole Talk revolves around a research topic that participants have been invited to discuss. Once the researcher has stated the research topic, there are no constraints on the direction that the Ole Talk will take or on the time allowed. The researcher’s role is not that of a moderator, as in a focus group, but participates fully, asking and answering questions and sharing their own experiences (Fernández, 2020).

In education research, Ole Talk offers a valuable reference point for sharing and communicating knowledge about the classroom and the education environment. Ole Talk allows for student research participants from diverse ethnic, socioeconomic, and educational backgrounds to collaboratively make sense of their experiences, relying on shared cultural competencies. Ole Talk occurs naturally in most Caribbean educational environments, both inside and outside the classroom, and can be utilized to gain insights into students’ education experiences.

Ole talk generates a conversational environment in which any topic can be explored from different perspectives depending on what is relevant for each participant or group. The open nature of Ole Talk provides the flexibility for participants to interact and share knowledge on their own terms. For example, if the topic being discussed is not relevant or engaging, participants can choose to be silent or to begin a different conversation alongside the ongoing one. Conversations tend to flow spontaneously with often more than one conversation taking place at the same time and topics can change quickly. Participants frequently talk over each other especially if the ole talkers are engaged and invested in the topic. In most cases, this is not seen as an interruption but as a normal part of the interaction.

In Ole Talk, participants speak (or not) freely and the conversations continue easily, even during heated debates. It is expected and accepted that some people will contribute simply by listening. If participants are uncomfortable with the way in which the conversation is taking place, they can shift the discussion or remove themselves from it altogether. In Ole Talk, the attention given to speakers does not rely on procedure, ritual, or hierarchy, but on the general interest in the topic or conversation. Fernández

¹It is worthwhile to note the Dr. Eric Williams’ University of Woodford Square (Azeez, 1989) where a “speakers’ corner” remains active to this day.

²*dar muela* in Cuba and Dominican Republic; *par* in Jamaica.

(2020) observed that the attention given also depended on the speaker's ability to deliver an entertaining performance or argument, or when the speaker employed humor, teasing, storytelling, or gestures, or modulated the rhythm and tone of the conversation to build momentum. In these cases, listeners become the speaker's audience, and speakers employ any of the communicative strategies listed above to keep them engaged.

Ole Talk features culturally specific humorous practices including good-natured teasing called *relajo*³ or *choteo* in Spanish-speaking Caribbean countries, and *fatigue* in Trinidad and Tobago. The use of humor in Ole Talk for enjoyment and relaxation is not at odds with its significance in knowledge construction, sense-making, and reflection. It is a useful tool in getting research participants to relax, engage, and feel at ease with the research process and other participants. Humor is frequently used to challenge imposed discourses and meanings, not by directly confronting them, but by mocking and transforming them.

Shared humor can bind the group together and becomes a significant coping mechanism when discussing adverse experiences. The shared laughter can be taken as an expression of solidarity. Additionally, when dealing with complex issues, humor enables participants to shape and reshape opinions, intentions, and positions without them being taken as definitive or fixed. This creates an environment of permissibility in which diverse opinions on controversial issues can be explored at ease.

Ole Talk inheres important similarities with indigenous conversational methods which, as discussed by Kovach (2010).

... comprise a specific way of knowing based upon the oral tradition of sharing knowledge. It is akin to what different Indigenous researchers, the world over, identify as storytelling, yarning, talk story, re-storying, re-membling (p. 124).

With localized conversational methods such as Ole Talk, relationships are of vital importance. Through collaborative storytelling, both the researcher and the knowledge holder become engaged in a shared process of knowledge construction and the relationship builds and deepens as stories are shared.

Kovach (2010) has noted that, in many indigenous conversational methods, protocol is of high importance to ensure that activities are carried out in a manner that reflects tribal traditions. Although there are cultural expectations around Ole Talk, this mode of interaction is not reliant on a fixed format. The utilization of the cultural competencies of Ole Talk such as participating in several dialogs at once, using humor meaningfully, and the use of silence for communication is largely unprescribed and undirected. Additionally, the habitus that is embodied in Ole Talk is not particular to a single ethnic group but has emerged from the cultural integration of the many peoples of the Caribbean (Glissant, 1989).

Ole Talk is seen as a culturally affirming research method in research by and with Caribbean people as the strategies for knowledge construction emerge from Caribbean experiences and cultures. Researchers and participants are equally involved in the process of sense-making, are familiar with the underlying codes and dynamics of Ole talk, and share the power to shape knowledge. This differs from research approaches where the researcher has primary access to the structures that shape the knowledge (e.g. questions in an interview, time keeping in a group discussion, moderation in a focus group). The oral practices of Ole Talk, emerging from the practical and living realities of Caribbean communities, enable research that affirms the localized ways that these communities share knowledge and in which the results of the research are more likely to be effectively aligned with their own understandings and current lives.

Ole Talk in practice

The characteristics of Ole Talk, described above, frame the process of knowledge construction. When Ole Talk is used in research, other aspects of the research process need to take into account the fluid, unstructured and dynamic nature of Ole talk. The following section suggests ways in which certain steps of the research process could be carried out when using Ole Talk. These suggestions are not limited to Ole Talk and can be used with other oral traditions with the expectation that the integrity of their unique cultural groundings will be maintained.

Recording Ole Talk

Ole Talk continues as people move around from one group to the other, or gravitate toward shared food. This necessitates having more than one recording device in order to capture the many conversations. The use of multiple recorders, located in strategic places in the setting allows for the recording of multiple conversation groupings, and for people to shift between groups. Videos can be used, with permission, to capture non-verbal interactions in a way that is minimally intrusive. Recorder-free areas can be created to enable participants to be off record if they wish.

Transcribing Ole Talk

Ole Talk data captured in multiple recordings are likely to be complex and multi-layered. Transcription needs to capture this complexity to enable an accurate analysis of the knowledge shared. When transcribing Ole Talk, context needs to be captured as well as discourse. This includes prosodic information (intonation, gestures, and stress patterns), conversation overlap, humor,

³*Relajo* does not have an exact English equivalent, but it refers to a state of suspension of all seriousness, usually involving the use of well-intended back and forth mockery, double meaning, and humor. It denotes a mode of joking around and not taking anything seriously, and usually involves well-meaning teasing with shared laughter as a typical outcome. For example, a phrase like: “*Déjense de relajo!*” can be translated as “Stop joking around!”

Table 1 Examples of the use of oral traditions in education research.

Oral research method	Research studies in education	Focus of studies
Storytelling	Gallagher, K.M., 2011. In search of a theoretical basis for storytelling in education research: story as method. <i>Int. J. Res. Method Educ.</i> 34(1), 49–61. Collins, F., 1999. The use of traditional storytelling in education to the learning of literacy skills. <i>Early Child Dev. Care</i> 152(1), 77–108.	Utilizing the works of Hannah Arendt and Bertold Brecht, two theorists of great storytelling capability, the author considers how their methodological thinking might be productively imported. Investigates what storytelling offers to the formal education of primary age children.
Ole Talk	Ferguson, S., (forthcoming). <i>The Role of Community Environments on First-Generation Jamaican University Students' Persistence in Higher Education</i> . [Doctoral Dissertation, the University of Denver].	Employing ole talk in online (liming online) to understand the role of community environments for first-generation Jamaican university students in higher education journey.
Ubuntu	Quan-Baffour, K.P., Romm, N.R., 2014. Ubuntu-inspired training of adult literacy teachers as a route to generating “community” enterprises. <i>J. Lit. Res.</i> 46(4), 455–474. Mukuni, J., Tlou, J., 2018. The place of Ubuntu in global education. In: <i>Cases on Global Competencies for Educational Diplomacy in International Settings</i> . IGI Global, pp. 223–248.	The use of Ubuntu in the various learning settings in the training of adult literacy teachers. How global education can benefit from Ubuntu, and how Ubuntu aligns with the global competencies articulated in the Global Competence Matrix.
Pūrākau	Cliffe-Tautari, T., 2020. Using Purakau as a pedagogical strategy to explore Maori cultural identities. <i>Set: Res. Inf. Teach.</i> 1, 12–18.	Unpacks how pūrākau as a pedagogical strategy enables Māori learners to explore their own perceptions about being Māori.
Talanoa	Latu, M., 2009. <i>Talanoa: A Contribution to the Teaching and Learning of Tongan Primary School Children in New Zealand</i> (Doctoral Dissertation, Auckland University of Technology). Baice, T., Lealaiauloto, B., Meiklejohn-Whiu, S., Fonua, S.M., Allen, J.M., Matapo, J., et al., 2021. Responding to the call: talanoa, va-vā, early career network and enabling academic pathways at a university in New Zealand. <i>High. Educ. Res. Dev.</i> 40(1), 75–89.	Using Talanoa to explore alternative approaches to teaching and learning pedagogies of Tongan children in New Zealand. Using Talanoa to privilege the worldviews and stories of the under-representation of Pacific/Pasifika academics in universities across Aotearoa New Zealand.
Talking Circles	Hanson, A., Danyluk, P., 2022. Talking circles as indigenous pedagogy in online learning. <i>Teach. Teach. Educ.</i> 115, 103715. Schumacher, A., 2014. Talking circles for adolescent girls in an urban high school: a restorative practices program for building friendships and developing emotional literacy skills. <i>Sage Open</i> 4(4), 2158244014554204.	Using talking circles as indigenous pedagogy in online education for preservice teachers. The use of talking circles to investigate peers helping peers in an urban high school.
Yarning	Gruppette, M., Southgate, E., Ober, R., Cameron, L., John, F., Thunig, A., et al., 2018. Yarning the way: the role of indigenous paraprofessionals in guiding the post-school educational pathways of aboriginal and Torres Strait Islander youth. Flückiger, B., Diamond, P., Jones, W., 2012. Yarning space: leading literacy learning through family–school partnerships. <i>Australas. J. Early Child.</i> 37(3), 53–60.	Examining the role of indigenous education paraprofessionals in facilitating post-school education options and access for indigenous young people through yarning. The use of yarning to look at effective partnerships between parents, teachers and the community to improve young indigenous Australians’ participation in early childhood education and their literacy development.
Halaqah	Shirin, H.B., Islam, A.A., Islam, M.S., 2017. Development and validation of the study circle model in higher education: the application of structural equation modeling. <i>MOJEM</i> 2(3), 18–32. Munadia, A., 2016. <i>The Enhancement of Critical Thinking Using Halaqah Method Among Students of Dayah in Dayah Mudi Mesjid Raya Salamanga, Aceh</i> (Master’s Thesis, Gombak, Selangor: International Islamic University Malaysia, 2016).	The study circle (Halaqah) program to disseminate the knowledge and skills for overcoming the drawbacks and challenges among students, in one of the Islamic universities in Malaysia. This study investigated the enhancement of critical thinking skills using halaqah method among Dayah students at Dayah Mudi Mesjid Raya Samalanga Aceh, Indonesia. It examined the teachers’ perceptions and practices of critical thinking in the classroom.
Digital storytelling	Van Gils, F., 2005. Potential applications of digital storytelling in education. In: <i>3rd Twente Student Conference on IT</i> (vol. 7, No. 7). University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science Enschede. Alismail, H.A., 2015. Integrate digital storytelling in education. <i>J. Educ. Pract.</i> 6(9), 126–129.	Digital storytelling as offering a new way of to educate people. Explores how digital storytelling can be an effective tool for both teachers and students in order to support learning and skills.

and laughter. It is important to structure the transcribed material to the aims of the particular study so that the relevant information is emphasized. For example, transcription tables with columns can note when limers talk over each other or when they laugh. Different shading can be used to distinguish one conversation from another when multiple groups are talking at the same time (Table 1).

Talking over each other	Sam	You go now, scrap your elbow, your mother then come, take you up, rubs the alcohol, makes you holler ...
	Nina	And then cut your arse
Talking over each other	Sam	Are you going outside again? Sometimes I do get your point, people go too far, I agree.
	Nina	Tough love, tough love
	Helen	That's why you so?
Laughter	Renee	That's why I'm so tough
	Helen	I think you ... there is a difference between being ... articulate ... I think you are sensitive, I think you are. I know you are.

Analyzing and reporting Ole Talk

When Ole Talk is used as the primary source of data, it is recommended that data analysis and the writing of the report be carried out simultaneously. This allows the researcher to adapt the structure of the report to the narratives and ideas emerging from the Ole Talk. When writing up the study's findings, the researcher can thus go back to the recordings to contextualize participants' conversations and to make sense of the situation in which the conversation emerged, the reaction of other participants, and other relevant behaviors such as tone, use of humor, movement and position, and talking over each other.

Representing the rich interactional practices of Ole Talk in a written text requires simplifying the data. However, to preserve some of its complexity, writing strategies can be developed to include nonverbal and contextual information, and to maintain the integrity of participants' narratives. Oftentimes in academic text, textual quotes from participants are provided without description of the context in which they were said. In an Ole Talk interaction, the meaning of a phrase can change dramatically depending on tone, body language and context. Therefore, it is recommended to incorporate sufficient description to capture features such as nonverbal communication surrounding discourse, the position and movement of participants, the tone of their voices, and their gestures. Below is an example of how this can be done, with the description of context and non-verbal communication in italics.

Everyone around the table had been sharing their migrant journey. Gabe went back to when his parents arrived in the UK from the Caribbean.

"Okay, so in the sixties" Gabe started locating his narrative, "big, big, big mass migration from the Caribbean. They were invited, and said that everything was going to be ... they were gonna get a job, and blah, blah, blah and when they go over there ..." Gabe shook his head and rolled his eyes, looking at Anand.

"They were invited" confirmed Anand, nodding.

"They lied to them" Derek added, sadly.

MIRIAM'S LIME. OLE TALK FRAGMENT 6.

Additionally, Ole Talk fragments quoted in the text can be numbered (see above) and referenced throughout the text to enable the researcher and readers to go back and forth connecting different parts of participants' narratives.

Application of Ole Talk in research

Ole Talk is being used increasingly in research in Caribbean contexts and with Caribbean peoples. Liming and ole talk, developed as Liming Methodology, was employed by Fernandez (2020) to investigate how Caribbean Islanders living in Aotearoa articulated a Caribbean identity. Kerrigan (2016)'s study of "small goal" football in Trinidad relied on Ole Talk to understand the relationship between sport and masculinities in the country.

In ongoing research, Ferguson's (forthcoming) study uses online (liming online) to understand the role of community environments for first-generation Jamaican university students on their higher education journey, specifically focusing on students/graduates from under-resourced and under-served communities in Jamaica. Chapman (forthcoming) uses liming and ole talk to investigate how colonial policing systems have shaped the current paramilitary policing of black and brown bodies in Trinidad and Tobago.

Conclusion

In research, oral traditions have intrinsic value to the community where research is unfolding, and this value is not related to their compliance with or contestation of foreign paradigms. Importantly, oral traditions operate as affirming research methodologies when those who share knowledge are centered as agents in the research process in an intersubjective relationship with those who may hold a formal role as researchers.

This shift from traditional research hierarchies requires intentionality. Oral traditions are intrinsically fluid and grounded in the lived experience of the people who embody them. If this process of re-centering participants and their experience in the research

process is not intentionally ensured, mainstream research positionings can take over leading to the corruption and disenfranchisement of oral traditions to exploit the knowledges of communities. Thus, a determining factor in identifying oral traditions as affirming methodologies is their ability to reclaim and center local and indigenous ways of constructing and sharing knowledge.

Oral traditions are as diverse as the cultures and communities in which they are located. Storytelling is a profoundly important cultural resource in many Indigenous societies. The way in which stories are shared varies among the many peoples that use storytelling. For example, Ole Talk, Pūrākau, Talking/Sharing circles, Yarning, Talanoa, Halaqah, or Ubuntu or Digital storytelling enables the sharing of stories with varying degrees of structuration, formats, and modes of participation.

Oral traditions are central to communities that are diverse and highly transient such as Caribbean communities. Oral traditions keep people connected and make knowledge accessible to people in their day-to-day life. With research involving Indigenous, Black, and People of color education communities, oral traditions such as Ole Talk create a space for authentic engagement where participants can talk to and over each other, provide space for humor and reflection, and listen to histories that hurt and stories that heal.

In education research, the use of oral traditions facilitates a more authentic and participatory space for communities that rely on orality to explore their own learning processes (inside and outside formal learning institutions) as well as critique the factors that shape their educational outcomes. Research in education can use oral traditions within an Affirming Methodologies framework to elicit conversations about how learning takes place, in an environment that is safe, strengths-based and encourages participants to build knowledge that centers, and is connected authentically to who they are and how they see the world.

References

- Achebe, C., 1988. *Hopes and Impediments: Selected Essays, 1965–1987*. Heinemann Educational, London.
- Ahmed, F., 2014. Exploring halaqah as research method: a tentative approach to developing Islamic research principles within a critical “Indigenous” framework. *Int. J. Qual. Stud.* 27 (5), 561–583. <https://doi.org/10.1080/09518398.2013.805852>.
- Archibald, J.A., 2008. *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. University of British Columbia Press, Vancouver, Canada.
- Azeez, M., 1989. The legacy of Eric E. Williams: a selected bibliography. *Curr. Bibliogr. Afr. Aff.* 21 (3), 267–273. <https://doi.org/10.1177/001132558902100303>.
- Becker, S.A., Afonso, D.D., Beard, M.B.H., 2006. Talking circles: Northern Plains tribes American Indian women’s views of cancer as a health issue. *Publ. Health Nurs.* 23 (1), 27–36.
- Bessarab, D., Ng’andu, B., 2010. Yarning about yarning as a legitimate method in Indigenous research. *Int. J. Crit. Indig. Stud.* 3 (1), 37–50.
- Cambridge, V., 2005. Soundscapes: reflections on Caribbean oral and Aural traditions. In: *Soundscapes: Reflections on Caribbean Oral and Aural Traditions*. The University of the West Indies, Cave Hill, Barbados. Cave Hill.
- Castellano, M.B., 2014. Ethics of aboriginal research. In: Teays, W., Gordon, J.-S., Renteln, A.D. (Eds.), *Global Bioethics and Human Rights: Contemporary Issues*. Rowman & Littlefield, pp. 249–262. ProQuest Ebook Central. <https://ebookcentral.proquest.com/lib/aut/detail.action?docID=1676482>.
- Chapman, N., (forthcoming). “Are we stuck in a time warp?” Exploring Narratives of “Frame Case” Within the Context of “Neo-colonial Oppression” in 21st Century Trinidad and Tobago. PhD Thesis, University of the West Indies.
- Cherrington, L., 2003. The use of Māori mythology in clinical settings: training issues and needs. In: Nikora, L., Levy, M., Waitoki, M., Masters, B., Te Awekotuku, N., Etheredge, R. (Eds.), *Proceedings of the National Māori Graduates of Psychology Symposium, 2002: Making a Difference*. Māori & Psychology Research Unit, University of Waikato, Hamilton, New Zealand, pp. 117–120.
- Cunsolo Willox, A., Harper, S.L., Edge, V.L., 2013. Storytelling in a digital age: digital storytelling as an emerging narrative method for preserving and promoting Indigenous oral wisdom. *Qual. Res.* 13 (2), 127–147. <https://doi.org/10.1177/1468794112446105>.
- Datta, R., 2018. Traditional storytelling: an effective Indigenous research methodology and its implications for environmental research. *Alternative* 14 (1), 35–44. <https://doi.org/10.1177/1177180117741351>.
- Dei, G.J.S., 2000. Rethinking the role of Indigenous knowledges in the academy. *Int. J. Incl. Educ.* 4 (2), 111–132.
- Drawson, A.S., Toombs, E., Mushquash, C.J., 2017. Indigenous research methods: a systematic review. *Int. Indig. Policy J.* 8 (2).
- Dyll-Mylebust, L., 2014. Development narratives: the value of multiple voices and ontologies in Kalahari research. *Crit. Arts* 28 (3), 521–538. <https://doi.org/10.1080/02560046.2014.929214>.
- Ferguson, S., (forthcoming). *The Role of Community Environments on First-Generation Jamaican University Students’ Persistence in Higher Education*. Doctoral Dissertation, The University of Denver.
- Fernández Santana, A., Nakhid, C., Nakhid-Chatoor, M., Wilson-Scott, S., 2019. Liming and ole talk: foundations for and characteristics of a culturally relevant Caribbean methodology. *Caribb. Stud.* 47 (1), 99–124.
- Fernández, A., 2020. *Sharing Our Way: A Study of Caribbean Identity Using Liming as Culturally Affirming Research Methodology*. Doctoral Dissertation. Auckland University of Technology.
- Finnegan, R., 2003. *Oral Traditions and the Verbal Arts: A Guide to Research Practices*. Routledge.
- Fonua, P., 2005. *Talanoa, Talking From the Heart, Matangi Tonga*. Vava’u Press, Nuku’alofa, Tonga.
- Fredericks, B.L., Adams, K., Finlay, S.M., Fletcher, G., Andy, S., Briggs, L., et al., 2011. Engaging the practice of yarning in action research. *Action Learn. Action Res. J.* 17 (2), 7–19. <https://doi.org/10.1080/136031100284849>. Retrieved from.
- Glissant, E., 1989. *Caribbean Discourse. Selected Essays*. University Press of Virginia, London.
- Halapua, S., 2002. *Talanoa Process: The Case of Fiji*. East West Centre, Honolulu, Hawaii.
- Hall, S., 1997. *Representation: Cultural Representations and Signifying Practices*. Sage Publications, Thousand Oaks, CA.
- Kargbo, J.A., 2008. *Oral Traditions and Libraries*. Library review.
- Kerrigan, D., 2016. Languaculture and grassroots football: “small goal” in Trinidad. *Int. Rev. Sociol. Sport* 51 (6), 735–751.
- Koukanen, R., 2000. Towards an “Indigenous paradigm” from a Sami perspective. *Can. J. Native Stud.* 20 (2), 411–436.
- Kovach, M., 2009. *Indigenous Methodologies : Characteristics, Conversations, and Contexts*. University of Toronto Press, Toronto.
- Kovach, M., 2010. Conversational method in Indigenous research. *First Peoples Child Fam. Rev.* 5 (1), 40–48. Retrieved from. <http://journals.sfu.ca/Kovach>.
- Lee, J., 2009. Decolonising Māori narratives: Pūrākau as a method. *MAI Rev.* 2. Article 3, Page 1 of 12. <http://www.review.mai.ac.nz>.
- Mahuika, N., 2012. *‘Kōrero Tuku Iho’: Reconfiguring Oral History and Oral Tradition*. Doctoral Dissertation. University of Waikato.
- McAdoo, H., 1980. Oral history as a primary resource in educational research. *J. Negro Educ.* 49 (4), 414–422.
- Mucina, D.D., 2011. Story as research methodology. *Alternative* 7 (1), 1–14.

- Nakhid, C., Farrugia, C., 2021. Affirming methodologies in two African Diasporic contexts: the sharing of knowledge through liming and ole talk among Caribbean Islanders in Aotearoa New Zealand and the practice of sharing with Sydney-based Africans. *Peabody J. Educ.* 96 (2), 1–15.
- Nakhid, C., Mosca, J., Nakhid-Schuster, S., 2020. Liming as research methodology, ole talk as research method—a Caribbean methodology. *J. Educ. Dev. Caribb.* 18 (2), 1–18.
- Nakhid, C., 2021. A culturally affirming approach to research methodology through the Caribbean practice of liming and ole talk. In: Paula Groves Price. *Race and Education*. Oxford University Press, New York. https://doi.org/10.1093/acrefore/9780190264093.013.ORE_EDU-01630.R1.
- Nakhid, C., 2022. Liming and ole talk: a site of negotiation, contestation and relationships. In: *Decolonising Peace and Conflict Studies Through Indigenous Research*. Palgrave Macmillan, Singapore, pp. 89–103.
- Nakhid-Chatoor, M., Nakhid, C., Wilson, S., Fernandez Santana, A., 2018. Exploring liming and ole talk as a culturally relevant methodology for researching with Caribbean people. *Int. J. Qual. Methods* 17 (1). <https://doi.org/10.1177/1609406918813772>.
- Navabi, M., 2006. The power of oral tradition: critically resisting the colonial footprint. In: *Anti-colonialism and Education*. Brill Sense, pp. 175–192.
- Reddie, A., 1998. The oral tradition of African-Caribbean people as a resource for black Christian formation. *Br. J. Theol. Educ.* 10 (1), 16–25.
- Rodriguez, D., 2010. Storytelling in the field: race, method, and the empowerment of Latina college students. *Cult. Stud. Crit. Methodol.* 10 (6), 491–507.
- Russell-Mundine, G., 2007. Key factors for the successful development of Australian Indigenous entrepreneurship. *Turizam: Znanstveno-stručni časopis* 55 (4), 417–429. Retrieved from <http://hrcak.srce.hr/file/38857>.
- Sams, J., Childers, L., 1990. *Sacred Path Cards: The Discovery of the Self Through Native Teachings*. Harper, San Francisco.
- Solórzano, D.G., Yosso, T.J., 2002. Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8 (1), 23–44.
- Struthers, R., Peden-McAlpine, C., 2005. Phenomenological research among Canadian and United States Indigenous populations: oral tradition and quintessence of time. *Qual. Health Res.* 15 (9), 1264–1276.
- Suaalii-Sauni, T., Fulu-Aiolupotea, S.M., 2014. Decolonising Pacific research, building Pacific research communities and developing Pacific research tools: the case of the talanoa and the faafaletui in Samoa. *Asia Pac. Viewp.* 55 (3), 331–344.
- Te Awakotuku, N., 1991. He tikanga whakaaro: Research Ethics in the Maori Community. *Manatu Maori*. UNESCO, <https://ich.unesco.org/en/oral-traditions-and-expressions-00053>.
- Vaka, S., Brannelly, T., Huntington, A., 2016. Getting to the heart of the story: using Talanoa to explore Pacific mental health. *Issues Ment. Health Nurs.* 37 (8), 537–544.
- Vété-Congolo, H., 2016. *The Caribbean Oral Tradition: Literature, Performance, and Practice*. Palgrave Macmillan, Brunswick.
- Walker, M., Fredericks, B., Mills, K., Anderson, D., 2014. "Yarning" as a method for community based health research with Indigenous women: the Indigenous Women's Wellness Research Program. *Health Care Women Int.* 1–11. <https://doi.org/10.1080/07399332.2013.815754>.
- Ware, F., Breheny, M., Forster, M., 2018. Kaupapa Korero: a Māori cultural approach to narrative inquiry. *Alternative* 14 (1), 45–53. <https://doi.org/10.1177/1177180117744810>.
- Wilson, S., 2001. What is an Indigenous research methodology? *Can. J. Native Educ.* 25 (2), 175–179.
- Wolf, P.R., Rickard, J.A., 2003. Talking circles: a Native American approach to experiential learning. *J. Multicult. Counsel. Dev.* 31 (1), 39–43.
- Zaimache, S., 2002. Education in Islam: The Role of the Mosque. Retrieved from <http://www.muslimheritage.com/topics/default.cfm?TaxonomyTypeID=101&TaxonomySubTypeID=19&TaxonomyThirdLevelID=134&ArticleID=220>.

Semiotics in the context of qualitative research

Marcel Danesi, Department of Anthropology, University of Toronto, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	189
Defining and classifying signs	190
Signs and meaning	191
Structure	192
Post-structuralism	194
Semiotics in qualitative research	195
Concluding remarks	196
References	196
Further reading	197

Glossary

- Connotation** The extended or acquired meanings of a sign
- Denotation** The referential or intentional meaning of a sign
- Markedness theory** The theory which posits that one of the two elements in an opposition is the default (unmarked) form and the other the marked form (the one that stands out)
- Opposition** A relation that keeps signs distinct from each other through minimal contrasts (*night* vs. *day*)
- Paradigmatic structure** Any aspect of signs that keeps them distinct
- Semantic differential** Technique for identifying connotative meanings based on seven-point scales
- Semiosis** The process of producing and understanding signs
- Signification** The process of using signs for various psychological and social-cultural reasons
- Signified** The conceptual part of the sign; what the sign encodes and refers to
- Signifier** The physical part of a sign; also known as form
- Syntagmatic structure** The combinatory structure of signs (how they have been constructed)

Introduction

As the discipline studying and documenting signs, sign behavior, sign creation, and sign functions, semiotics is an ideal tool for conducting qualitative research in the cognitive and social sciences, since human cognition, behavior, and even perception is guided by sign structures, if not structured by them. A sign is defined as any physical form that stands for something other than itself in some specific context. A V-shape made with the raised index and third fingers is a perfect example of a sign, since it does not stand for raised fingers but for a series of referents ranging from the number two to abstract notions such as peace and victory. It all depends on who uses it and in what context it is used.

Today, semiotics is used mainly as a methodological tool to study such sign-based phenomena as body language, visual communication, social media, advertising, narratives, material culture (clothing, cuisine, etc.), rituals, the study of digital forms and systems—in a phrase, anything that is informed by sign-based activities or behavior. One of its modern-day founders, the Swiss philologist Ferdinand de Saussure (1857–1913), defined it as the science concerned with “the role of signs as part of social life” and “the laws governing them” (de Saussure, 1916/1958, p. 15).

A sign (a word, a visual image, a hand gesture, and so on) selects what is considered by people at specific moments in time as important to be known and recalled from the infinite variety of things that are in the world. It imprints knowledge into structured forms which filter out extraneous information. Although we create new signs to help us gain new knowledge and modify previous knowledge, by and large, we literally let our signs “do the thinking” for us, once we have acquired them in context. We are born into an already-fixed system of signs that will largely determine how we view the world around us and within us. This system sustains psychological life (based on nongenetic information) through expressive, communicative, and representational activities, forming a kind of symbiotic partnership with the biosphere to regulate human behavior and to shape cultural and social evolution. But it is not a deterministic partnership. Humans have the capacity to redesign aspects of semiotic life and thus completely refashion the sign structures they have created.

This chapter looks at the discipline of semiotics, starting with a historical synopsis of its origins and major developments. It then describes the two major theories of the sign, by Ferdinand de Saussure and Charles Sanders Peirce, which still form the foundation

of semiotic theory and practice. Saussure saw the sign as a binary structure—its physical form and its meaning—while Peirce saw it as a triadic structure—its physical form, its meaning, and its changing interpretation. The chapter then looks at Peirce's main categories of sign classification, the distinction between denotative and connotative meaning, the use of the semantic differential to evaluate the latter kind of meaning, the notions of oppositional structure and markedness, and their critique by the movement called post-structuralism. Finally, the applications and implications of semiotics to qualitative research are discussed, including an example of how it has been applied in the area of education.

Defining and classifying signs

The initial task of any science is to define its objects of investigation. This is also the case with semiotics, whose object of interest is *signs*—words, symbols, gestures, and so on, along with their social functions, their interpretations, and how they encode knowledge in specific ways. The first technical definition of the sign goes back to the ancient world, starting with the Greek physician Hippocrates (c. 460–370 BCE), who argued that the particular physical form that a bodily symptom takes – a *semeion* (mark) – constitutes a vital clue for finding its etiological source. He called the method of diagnosis *semeiotiké* (semeiotics), laying the foundations for the diagnosis, treatment, and prevention of disease, taking into account not only the biological sources of the physical symptomatic signs themselves, but also what they meant to patients and to the culture in which they were reared. After Hippocrates, the question arose at some point as to whether or not there was any connection between natural signs, such as symptoms, and conventional ones, such as words—a connection first discussed by the Stoics around 300 BCE, who argued that conventional signs were reflections of natural ones.

In the early medieval period, philosophers started referring to signs as being either natural (produced by the body or nature) or conventional (invented by humans). It was the early Christian church father, St. Augustine of Hippo (354–430 CE), who put forth an early typology of signs in his masterwork *De doctrina christiana* (Augustine, 1887), in which he describes natural signs (*signa naturalia*) as forms lacking intentionality and conventional ones (*signa data*) as forms produced by human intentions. The former include not only symptoms, but also plant coloration, animal signals, and the like; the latter include words, gestures and the various symbols that humans invent to serve their psychological, social, and communicative needs (Deely, 2001, pp. 24–56). The distinguishing feature of natural signs is that they have no human intent behind them. They occur naturally, outside of human volition. However, humans have an innate desire to interpret them in their own way. To do so, they invent conventional signs, which are products of intentions and needs.

St. Augustine's views lay largely unknown until the eleventh century, when interest in signs was rekindled by Persian scholars, who brought the works of Plato, Aristotle, and other Greek thinkers to medieval Europe. The result was the movement known as *Scholasticism*. Using Aristotelian philosophy as their model, the Scholastics asserted that signs allowed humans to capture truths, and thus to know the world through them. But within this movement there were some—the so-called *nominalists*—who argued that “truth” was a matter of subjective opinion and that signs captured, at best, only illusory and highly variable human versions of it. John Duns Scotus (c. 1265–1308) and William of Ockham (c. 1285–1349), for instance, stressed that signs only referred to other signs, rather than to actual things. The Scholastic theologian St. Thomas Aquinas (1225–1274) countered, however, that signs referred to real things, since they were derived from sense impressions. He also asserted that sacred signs revealed truths that were beyond rational comprehension and, therefore, had to be accepted on faith. In the same era, the English philosopher and scientist Roger Bacon (c. 1214–1292) developed one of the first comprehensive typologies of signs, claiming that, without a firm understanding of the role of signs in human understanding, discussing what truth is or is not would end up being a trivial matter. Bacon's proposal to study signs separately elicited little or no interest until 1690, when British philosopher John Locke (1632–1704) revived it in his *Essay Concerning Humane Understanding* (Locke, 1690). Locke used the term *semeiotics* in reference to a set of principles for examining signs and their meanings, not in reference to a distinct discipline or method of inquiry. It is not clear where Locke got the idea to use this Hippocratic term. The same term was used two decades earlier in 1670 by the English physician Henry Stubbe (1632–1676) to designate the interpretation of signs within medicine.

The idea of fashioning an autonomous science of sign study, however, did not crystallize until the nineteenth-century Swiss philologist, Ferdinand de Saussure (de Saussure and Baskin, 1916), put such a proposal forward, which he called *sémiologie* (semi-ology). As the following citation (de Saussure and Baskin, 1916, pp. 15–16) shows, Saussure suggested that the main goal of semi-ology (should it ever come into being) was to understand the social function of signs, as a branch of psychology, putting linguistics as itself a branch of semiology:

It is possible to conceive of a science which studies the role of signs as part of social life. It would form part of social psychology, and hence of general psychology. We shall call it *semiologie* (from the Greek *semeion*, “sign”). It would investigate the nature of signs and the laws governing them. Since it does not yet exist, one cannot say for certain that it will exist. But it has a right to exist, a place ready for it in advance. Linguistics is only one branch of this general science. The laws which semiology will discover will be laws applicable in linguistics, and linguistics will thus be assigned to a clearly defined place in the field of human knowledge.

Saussure saw the sign as a binary structure composed of two interrelated dimensions—a physical component (the sign itself) and a conceptual component (what it refers to and brings to mind). Saussure termed the former part, such as the phonemes that make

up the word *cat*, the *signifier*, and the concept or mental image that the sign elicits (the feline animal), the *signified* (literally meaning “that which is signified by the sign”). Saussure claimed, moreover, that there is no necessary sensory or experiential connection between the signifier and the signified. So, the actual phonemic structure of the word *cat* is an arbitrary one, with no perceivable linkage between the sounds that make up the word and the animal itself—that is, the word *cat* is not constructed to simulate or imitate the sounds a cat makes, as in onomatopoeic constructions. Any other signifier would have sufficed—as long as it conformed to the laws of word formation in the language (that is, the rules for combining phonemes to form legitimate words). While this aspect of Saussurean sign theory is disputed by semioticians today, the main idea behind it is still a relevant one—namely that the two parts of a sign imply each other conceptually. So, when we use or hear a word such as *tree* (the signifier), an image of the signified (a certain type of plant) crops up concomitantly in our minds—if we know English of course—even if a real tree is not present for us to observe. Now, if we do come across a tree in reality, then the word we use for it (*tree*) also comes to mind. In other words, when the word *tree* is used, an image of the arboreal plant occurs in the brain; vice versa when we see this arboreal plant the word *tree* is suggested concomitantly in the brain. This is why Saussure called word signifiers sound images and their signifieds mental images. The two are inextricably intertwined—one implying the other. Needless to say, wherever “trees” are not part of a culture’s reality (such as, for example, in nomadic desert cultures), no signifier for them is required in the first place (since there is no signified to be encoded).

The first modern-day technical classification of signs is due to the American pragmatist philosopher Charles S. Peirce (1839–1914), who identified 66 species of signs in total. Some of these provide detail as to how signs function cognitively. For example, as its name implies, a *qualisign* is a form that draws attention to some quality of its referent (the object it represents). In language, an adjective is a *qualisign* since it draws attention to the qualities (color, shape, size, etc.) of nouns. In other sign systems, *qualisigns* include colors (painting), harmonies and tones (music), etc. A *sinsign* is a form used to single out a particular object – a pointing finger and the words *here* and *there* are examples of *sinsigns*. A *legisign* is a form that designates something by convention (literally by law). *Legisigns* include symbols, emblems, such as those used on flags, and various other forms that are used according to specific social rules or conventions. However, there are three fundamental types of signs identified by Peirce that are now part and parcel of every semiotician’s methodological toolkit – *icons*, *indexes*, and *symbols*. Signs originating in the perception of some property in a referent are icons. So, iconic signs stand for their referents by resemblance, imitation, simulation, or emulation. Portraits are icons of human faces (and what they mean); onomatopoeic words are icons of sounds made by certain objects or actions (drip, bang, etc.); and so on. The meanings of icons can generally be figured out even by those who are not a part of the culture that use them. Signs that relate referents, sign users, and even other sign forms in some way are defined as indexes. The pointing finger is a perfect example of an index. When we point to something, we are in fact relating it to our location as pointers. If it is close by, we refer to it as *near* or *here*. If not, we refer to it as *far* or *there*. Finally, signs that stand for referents in historically based or conventional ways are symbols. The V-figure standing for peace has symbolic valence, that is, it is interpreted conventionally for peace as a result of some historically or culturally based event.

Signs and meaning

Semiotic inquiry is guided generally by two fundamental models of the sign – Saussurean and Peircean. As discussed briefly above, Saussure saw the sign as a binary form with two components –the signifier and the signified (de Saussure and Baskin, 1916, p. 101)—which are connected to each other in an arbitrary way. Peirce saw some signs as originating instead in perceptual modeling tendencies (icons). In contrast to Saussure, he claimed that iconicity is a basic tendency in sign-creation and sign-interpretation (which he called *semiosis*). Peirce put forward a triadic model of the sign – a structure consisting of the actual physical sign form, called the *representamen*, the thing to which it refers, termed the *object*, and the meanings that it elicits in real-world situations, called the *interpretant* (Peirce, 1931, vol. 2, pp. 247–248, 307). The interpretant is itself a derived mental sign, because it entails the further production of meanings arising from the context in which the sign is used. In our culture, a cat is considered to be a domestic companion, among other things; in others, it is viewed primarily as a sacred animal; and in some others, it is considered to be a killer of rodents. Thus, while the sign refers to virtually the same mammal in different cultural contexts (no matter what name is used), its interpretant varies considerably. Thus, while the sign *refers* to virtually the same mammal in different cultures (no matter what linguistic name is used), its interpretant varies considerably, constituting a source of supplementary semiosis.

Unlike Saussure, Peirce viewed conventional signs (such as words) as originating, by and large, in the attempt to reproduce perceivable properties in their objects. Therefore, the Peircean sign is a *motivated* (not *arbitrary*) structure, whose construction is guided originally by some sensory, affective, or historical impulse. Peirce termed this a *firstness* dimension, defined as the tendency to construct or understand signs—*first*—as simulations of some perceived property of their objects or as related by resemblance to other objects or signs. He called the actual process, *iconicity*. An obvious example of an iconic sign is the word *splash*, whose phonemic structure was clearly modeled to imitate the sound made by something striking or falling into a liquid. Saussure did admit that some signs were fashioned in this onomatopoeic way. But he maintained that this was the exception, not the rule. However, as Peirce suggested, our *first* tendency in semiosis is to construct signs iconically, even though we may no longer be aware of it. Many words possess a latent sound-imitative quality built right into their structure. Consider the word *duck*. The sound combination that makes it up is, to be sure, one of an infinite number of assemblages that can be envisioned in English. But the final/k/of that word hints at the kind of sound a “duck” is perceived to make, referenced much more directly in the word *quack*. This type of “iconic modeling” is intrinsic to sign creation, even if it is not always obvious. Since iconic signs are fashioned in specific historical

contexts, their physical forms are not universal, as might be expected. Peirce used the term *hypoicon* to acknowledge this context-based dimension of firstness. Nevertheless, because it is an experientially-based process the referent can be figured out even by those who are not part of the culture, if they are told how a sign simulates, resembles, or substitutes it. Thus, if someone who does not speak English is told that the expression *cock-a-doodle-do* stands for the sound a rooster is perceived to make, then that person would have little difficulty imagining the resemblance between the expression and its referent. Peirce called the tendency of relating referents and sign-users in some way, *secondness*. This engenders indexical signs, such as the pointing finger or words such as *here*, *there*, *up*, *down*, and so on. Finally, Pierce claimed that there exists a *thirdness* dimension in sign construction, which consists in learning and using the conventional signs (symbols) developed in specific cultural contexts. The Union Jack stands for Great Britain—it is interpreted historically, conventionally, and emotionally as a sign of nationhood.

A major problem for semiotics is to determine what meaning is and how it manifests itself through signs. To avoid ambiguity, the terms *semiosis* and *signification* are used by semioticians. The former, as mentioned, refers to the psychological process itself of producing and recognizing specific forms as signs (Fisch, 1978, pp. 32, 41); the latter refers to the mental images that crop up in the brain when a sign is used. Signification involves two main meaning-making processes – *reference* and *sense*. Reference is the process itself of identifying what the sign form captures; sense is what that form elicits psychologically, historically, and socially. Signs (such as words) may refer to the same (or similar) things, known as referents, but they have different senses. For example, the “long-eared, short-tailed, burrowing mammal of the family Leporidae” can be called *rabbit* or *hare* in English. Both words refer essentially to the same kind of mammal. However, there is a difference of sense between the two words – *hare* is the more appropriate term for describing the mammal if it is larger, has longer ears and legs, and does not burrow. Another difference is that a rabbit is often perceived to be a pet, while a hare is unlikely to be perceived as such (Danesi, 2007, p. 34).

For many semioticians, the terms *denotation* and *connotation* are preferred to reference and sense. Consider, again, the word *cat*. The word elicits an image of a “creature with four legs, whiskers, retractile claws”, etc. This is its denotative meaning. As this shows, the function of denotation is referential – that is, it allows users of the sign to determine if something real or imaginary under consideration is a cat or not. Denotative meaning thus divides the world of reference into “yes–no” domains – something is either a cat or it is not. All other meanings of the word are connotative – “He’s a real cool cat;” “She let the cat out of the bag;” and so on. These meanings are products of the historical associations forged between cats and socially significant concepts or processes. Connotation thus connects the world of reference to larger historical processes. Connotation is not an option, as some traditional philosophical and linguistic theories of meaning continue to sustain to this day; it is something we are inclined to extract from a sign. The numbers 7 and 13 have specific denotative meanings – each digit stands for a specific quantity of things. However, in various cultures, both invariably reverberate with connotative meanings such as fortune, destiny, mysticism, and so on. In 1957, the psychologists Osgood, Suci, and Tannenbaum showed how connotative meaning is built into everyday concepts by using a technique that they called the semantic differential. This consists in asking a series of evaluative questions to subjects about a particular concept – Is X good or bad? Should Y be weak or strong? etc. – which they are then told to rate on seven-point scales. The ratings are collected and analyzed statistically in order to sift out any general pattern that they might bear. Suppose that subjects are asked to rate the concept “ideal American president” in terms of the following scales: Should the president be: (1) Young or old? (2) Practical or idealistic? (3) Modern or traditional? (4) Male or female? and so on. A subject who feels that the president should be more youngish than oldish would place a mark toward the young end of the first scale; one who feels that a president should be traditional, would place a mark toward the traditional end of the scale; and so on. Research using the semantic differential has shown that connotation is invariably culture specific.

Research utilizing the semantic differential has also shown that the range of connotative variations is not a matter of pure subjectivity, but forms a socially-based pattern. Younger people may tend to rate the “ideal parent” as being flexible, older ones as “stern.” But this will vary according to culture, and in fact the semantic differential will flesh out the culturally-significant connotations of specific concepts. In other words, the connotations of concepts are constrained by psychological and cultural factors. A common critique of the technique, however, is that the scales used (*practical-versus-idealistic*, *flexible-versus-stern*, etc.) are themselves artifacts, utilized by the researchers to unconsciously guide subject choices along a certain path that is itself culturally-biased. In other words, they show what analysts want to show. But even if the scales are determined in advance, the results obtained may be unexpected ones and thus the whole technique would be legitimate as a form of randomized experimentation. In effect, the semantic differential remains a valid investigative tool into the nature of connotation, since it allows us to flesh out what something means in cultural terms (within limits).

Structure

In order to extract meaning from a sign, sign system, or sign assemblage (text), one must be able to recognize it as such in the first place. This means that it must have structure. Specifically, a sign (or sign assemblage such as a text) is recognizable as such if: (1) it is physically distinctive and (2) it is organized in a patterned way. The former implies *paradigmatic* and the latter *syntagmatic* structure. The phonic difference between *cat* and *rat*, for example, is a paradigmatic feature, allowing the native speaker to perceive the difference as a cue to meaning differentiation. Paradigmatic structure, as this simple example shows, involves some form of meaning-bearing differentiation in the physical makeup of a sign. In music, major and minor chords of the same key are perceivable as significantly different on account of a half-tone difference in the middle note of the chord; raising the index and middle fingers in a vertical orientation can mean victory, or peace (among other meanings), but aiming the same two fingers in a horizontal

way at someone would be interpreted instead as a threat; and so on. Now, the way in which the sounds of the word *cat* are put together is a further cue as to its meaning. Simply stated, the sounds of the word have been put together in line with English syllable structure. This is an example of syntagmatic structure. In music, a melody is recognizable as such only if the notes follow each other in a certain way (e.g., according to the rules of harmony); writing words in English involves putting the letters in a left-to-right form to stand for phonemes (more or less); and so on. In essence, a sign is perceivable as a meaning-bearing form if it bears paradigmatic (differential) and syntagmatic (combinatorial) structure in tandem.

Another main tenet of semiotics is that signs can vary in structural size, that is, in actual physical form. A sign can thus be something small, such as two fingers raised in a V-form; or it can be something much larger, such as a mathematical equation or a narrative. The interpretation of the sign is in all cases holistic, not discrete (decoded in terms of its constituent parts). If asked what $c^2 = a^2 + b^2$ means, a mathematician would say that it stands for the Pythagorean theorem, not for specific digits captured by the letters used (even if this is also true). If we ask someone who has just read a novel what it was all about, we would receive an answer that reveals a perception of the novel as a form containing a singular (holistic) message or purpose. The larger signs constructed by combining smaller ones (such as words in the novel) are called *texts* and the meanings we extract from them are called *messages*, rather than just signifieds. The term *text* embraces such things as conversations, letters, speeches, poems, myths, novels, television programs, paintings, fashion styles, scientific theories, mathematical equations, musical compositions, and so on. Texts are composite structures (signs made up of smaller signs) that are not interpreted in terms of their constituent parts, but holistically as single meaning-bearing structures.

Texts function primarily as representation devices, intended to relate, depict, portray, or reproduce something perceived, sensed, imagined, or felt as requiring signification by means of a complex (combinatorial) form. This implies knowledge of how signs cohere into systems or *codes*. To enter into a conversation, for example, one would need to know the language code involved – the system of sounds, words, syntactic rules, etc. that allows for conversation to occur. Language, dress, music, and gesture, among many other things, are examples of codes. These can be defined, more formally, as systems of signs that can be used over and over to encode and decode texts. There are many kinds of codes. For example, intellectual codes contain signs in them (numbers, words, symbols, etc.) that allow for representational activities of a logical, mathematical, scientific, or philosophical nature. Social codes (dress, gender, food, space, etc.) contain sign structures for making messages about oneself in socially appropriate ways and for regulating interpersonal activities. Food codes underlie how people interpret certain foods as signifiers of various rituals, meanings, etc. It is important to note that once a text is constructed, it takes on its own paradigmatic and syntagmatic properties – that is, it is differentiable from other texts in terms of the kinds of signs that constitute it (verbal, pictorial, etc.), and it is the result of specific syntagmatic properties associated with the code or codes utilized to construct it (language, narrative, and so on). A semiotic *code* can be defined as the meaning-assigning system that inheres between signs organized in a certain way, with a certain function, or bearing specific information. Technically, the definition of the sign is “something that stands for something else in a certain way.” A code is the term for the “stands for” relation, which is completed with “in a certain way.”

Some codes are codes within codes (or subcodes); others are meta-codes, that is “larger codes” that encompass these subcodes—although this distinction is rarely made in semiotics for practical purposes—it is a theoretical distinction. So, the term *code* can refer to a meta-code or a subcode, with each term used distinctively if necessary. An example of a larger all-encompassing meta-code is language. Each level within it has coded structure. For example, knowledge of words entails knowledge of (1) how they are formed phonemically; (2) how they cluster into phrases and sentences; (3) how they are inflected; which words (known as *function words*) can be used to relate the other words in a sentence to each other; and (4) how they bear denotative and connotative meanings when used in specific forms. Saussure called the language meta-code, *langue*, and the ways in which it is used or expressed, *parole*—rendered generically in English as *language* and *speech* respectively. Language is a mental code. Speech is how language is delivered, involving the intentional use of the organs of the vocal apparatus—the tongue, the teeth, the epiglottis, etc.—to articulate sounds and words. Language can also be delivered or expressed through other physical modes—writing, gesture, and so on. One can have language without speech (as do individuals with impaired vocal organs), because it exists in the mind. But one cannot have speech without language, because speech depends on the language code.

To identify forms as meaning-bearing structures, Saussure introduced the notion of *différence*. This was expanded by the so-called Prague Circle of linguists in the 1920s and 1930s into the technique known as opposition, which has been used across disciplines, from psychology to linguistics. For the sake of historical accuracy, it should be mentioned that this notion is, implicitly, an ancient one, being characteristic of myths (good vs. evil, youth vs. old age, etc.). The Prague Circle linguist Trubetzkoy (1936, 1968), however, was the first to apply it formally to the study of the structure of language. According to psychologist Ogden (1932, p. 18), an early promoter of opposition theory, opposition had general applicability because it offered “a new method of approach not only in the case of all those words which can best be defined in terms of their opposites, or of the oppositional scale on which they appear, but also to any word.” In the 1930s and 1940s, semioticians started noticing that oppositional structure was not confined to language. It cropped up in the analysis of nonverbal systems and codes as well – in mathematics, for example, fundamental oppositions include positive versus negative, odd versus even, and prime versus composite; in music, they include major versus minor and consonant versus dissonant; and so on.

The Prague Circle linguists also argued that there were orders of oppositions that went beyond the purely binary. In mathematics, for example, the addition-versus-subtraction opposition is a basic or first-order one, while the multiplication-versus-division opposition is a derived or second-order one – since multiplication is repeated addition and division repeated subtraction. French semiotician Algirdas J. Greimas later introduced the notion of the semiotic square to connect sets of oppositions (Greimas, 1987). Given a word such as *rich*, Greimas suggested that we determine its overall meaning by opposing it to its contradictory, *not rich*, its contrary,

poor, and its contradictory, *not poor*, in tandem. Moreover, as work with the semantic differential showed in the 1950s (mentioned above), there seem to be gradations within the binary oppositions themselves, which are due to culture-specific connotative processes. Anthropologist Lévi-Strauss (1958) further expanded opposition theory by showing that pairs of oppositions often cohere into sets forming recognizable units. In analyzing kinship systems, he found that the elementary unit of kinship was made up of a set of four oppositions: brother versus sister, husband versus wife, father versus son, and mother's brother versus sister's son. Lévi-Strauss suspected that similar sets, or orders, characterized oppositions in other cultural systems.

Although the idea goes right back to Aristotle and other ancient philosophers in various civilizations the cognitive importance of opposition was noticed by the psychologists Wilhelm Wundt (1832–1920) and Edward B. Titchener (1867–1927), who were the ones to term it opposition. Its use in semiotics and linguistics is due, as mentioned, to the Prague Circle linguists (Hjelmslev, 1939; Benveniste, 1946; Jakobson, 1939). The theory was criticized from the start as being useful but highly artificial and not really consistent with human psychology. However, already in the 1940s, Jakobson (1942) empirically showed that the theory of opposition predicts the sequence of phoneme acquisition in children, since he found that the sound oppositions that occur frequently are among the first ones learned by children, while those that are relatively rare are among the last ones to be acquired by children.

Another early criticism of opposition theory was the fact that it seemed to neglect the importance of a structural relation that is vital to understanding semiosis – associative structure. The study of such structure came, actually, to the forefront in the latter part of the 1970s, becoming a major trend within linguistics itself (Pollio et al., 1977; Lakoff and Johnson, 1980; Lakoff and Johnson, 1999; Fauconnier and Turner, 2002). American linguist George Lakoff and philosopher Mark Johnson set the ball rolling by claiming that a simple linguistic metaphor such as “My friend is a gorilla” is not an idiomatic exception to basic literal meaning, but a systematically formed utterance; it is a token of a cognitive associative structure that they called a *conceptual metaphor* (Lakoff and Johnson, 1980). This is why we can also say that John or Mary or whoever we want is a gorilla, snake, pig, puppy, and so on. Each specific linguistic metaphor (“John is a gorilla,” “Mary is a snake,” etc.) is an instantiation of a more general associative cognitive structure – *people are animals*. The question becomes: How are such concepts formed? They are formed through image schemata, as the two scholars argued (Lakoff, 1987; Johnson, 1987). The source for the people are animals conceptual metaphor seems to be an unconscious image that human personalities and animal behaviors are linked in some way. In other words, it is the consequence of an ontological imagistic opposition: *humans as animals*. This is a perfect example of how opposition manifests itself as an associative mechanism, not just a binary or multi-order one (as discussed above). In this case, the two poles in the opposition are not contrasted (as in night vs. day), but equated: humans as animals. This suggests that oppositional structure operates in a noncontrastive way at the level of figurative meaning.

Post-structuralism

The most severe critiques of opposition theory have revolved around the relative notion of *markedness* (Tiersma, 1982; Eckman et al., 1983; Andrews, 1990; Battistella, 1990). In oppositions such as night-versus-day, the question becomes: Which of the two poles is the more basic one (culturally or psychologically)? This is called the *unmarked* pole (the default one) and the other pole, the *marked* one (the one that stands out). It is easy to identify *day* as the unmarked form and *night* as its marked counterpart, since generally speaking we sleep at night and carry out conscious activities in the day. For example, we ask, “How many days are left before your birthday?” not “How many nights are left before your birthday?” Now, the problem is deciding which pole is marked and unmarked in a problematic opposition such as the male-versus-female one. The answer seems to vary according to the social context to which the opposition is applied. In patrilineal societies the unmarked form is male; but in matrilineal ones, such as the Iroquois one (Alpher, 1987), it appears to be female. Markedness, thus, seems to have social consequences. However, the late French semiotician-philosophers Michel Foucault (1926–1984) and Jacques Derrida (1930–2004) claimed that this whole approach to sign study was flawed, and was the unwitting source of social inequalities (Foucault, 1972; Derrida, 1976). Their refutations led to the movement known as post-structuralism, starting in the late 1950s and gaining prominence in the 1970s. In this movement, the oppositions identified by linguists and semioticians are seen as the result of an endemic logocentrism on the part of the analyst, not the result of some tendency present in the human brain. Saussure claimed that every sign was understandable in terms of its *différence* from other signs. In contrast to Saussure's idea of *différence*, Derrida coined the word *différance* (spelled with an “a,” but pronounced in the same way), to intentionally critique Saussurean theory. With this term, Derrida aimed to show that Saussure's so-called discoveries were unwitting, biased ones, because a science of language cannot succeed since it must unfold through language and thus partake of the slippage (as he called it) it discovers.

In effect, post-structuralists did not view markedness as a tool for investigating social-conceptual structure; they saw it, instead, as validating biases in that very structure. Derrida (1976, p. 237) went so far as to claim that our oppositions deconstruct themselves when analyzed reflectively, that is, they fall apart, revealing their biased origins:

In idealization, to an origin or to a “priority” seen as simple, intact, normal, pure, standard, self-identical, in order *then* to conceive of derivation, complication, deterioration, accident, etc. All metaphysicians have proceeded thus: good before evil, the positive before the negative, the simple before the complex, the essential before the accidental, the imitated before the imitation, etc. This is not just *one* metaphysical gesture among others; it is *the* metaphysical exigency.

But Derrida seems to have missed the fact that markedness is a theory, not a technique of mind manipulation. Markedness can actually be used to identify where the social stereotypes lie, not encode them. And this can, over time, actually lead to reversing markedness structure and the social inequalities it reflects. This has happened to the male-versus-female opposition in various modern societies. Such reversals exist across oppositions.

Post-structuralism made an impact on many fields of knowledge, not just semiotics and linguistics. As written language is the fundamental condition of knowledge-producing enterprises, such as science and philosophy, post-structuralists claim that these end up reflecting nothing more than the writing practices used to articulate them. However, in hindsight, there was (and is) nothing particularly radical in post-structuralism. Already in the 1920s, Jakobson and Trubetzkoy started probing the relativity of language oppositions in the light of their social and psychological functions. Basing their ideas in part on the work of German psychologist Karl Bühler (1879–1963), the Prague Circle linguists posited that language categories mirrored social ones. The goal of a true semiotic science, they claimed, was to investigate the isomorphism that manifested itself between signs and social systems. In other words, opposition theory was the very technique that identified social inequalities, not masked them.

Semiotics in qualitative research

Semiotics is well-suited as a tool for complementing or supplementing qualitative research in the cognitive sciences and education, providing a theoretical framework and *modus operandi* that aims to understand the connection among sign systems (language, symbolism, and so on) and specific forms of human behavior and understanding. The late semiotician Umberto Eco (1976, p. 74) provided five reasons for defining semiotics plausibly as a qualitative investigative tool:

1. It is an autonomous discipline that can bolster understanding in other disciplines, given that these are based on semiosis.
2. It has developed theories and practices that allow semioticians to determine how signs and sign systems are constructed and how they function psychologically, historically, and socially.
3. It generates hypotheses about human behavior and cognition.
4. It makes predictions about how societies and cultures will evolve.
5. Its findings can have an impact on the actual state of the world.

To these features, one can add the following relevant premises or principles that allow for the various applications of semiotics in diverse qualitative research paradigms (Danesi, 2007):

1. Signs and sign systems reveal similar properties of structure and function the world over.
2. Differences in the manifestations of these properties are due to the histories and traditions of particular cultures and societies.
3. Signs and sign systems entail culture-specific classifications of the world.
4. These influence the way people think, behave, and act.
5. Perceptions of “naturalness” are thus tied to sign systems; that is, what is often considered to be “common sense” turns out often to be “communal sense.”
6. Sign structures produce relativity effects.

The last principle implies that the signs we use constitute a guide for daily living. It does not claim, as some have maintained, that the meanings imprinted, say, in the symbols, vocabularies, and grammars of other cultures cannot be learned. Vocabularies bring out how language serves specific human needs in particular contexts and then doubles back on its users to guide their view of the world. Consider bodies of water as a simple, yet illustrative example. In English, we classify them as *lakes*, *oceans*, *rivers*, *streams*, *seas*, *creeks*, and so on. Clearly, bodies of water are important in English culture, perhaps because of the importance of navigation and other water-based activities in its history. People living in the desert have very few words for bodies of water, for obvious reasons. Because of the importance of water bodies in English (and other languages), the vocabulary is constructed to reflect this, including criteria such as size, which enters into the classificatory picture—*ocean-versus-lake*—as does width and length—*river-versus-stream*—among other oppositions. These sign structures shape how we come to perceive bodies of water and what they entail ontologically. Access to this ontological code is by acquiring the relevant vocabulary.

A sign selects what is considered important to be known and recalled from the infinite variety of things that are in the world—it imprints important information into a structured form which filters out unnecessary or subsidiary information, retaining within its signification core what is felt to be relevant. Although we create new signs to help us gain new knowledge and modify previous knowledge—that is what artists, scientists, writers, inventors, for instance, are always doing—by and large, we literally let our signs “do the understanding” for us, once we have acquired them in a rearing context. We are born into an already-fixed semiosphere that will largely determine how we view the world around us. Conducting research on human systems of meaning can clearly be informed practically by semiotic theory and method. We hardly realize the presence of unconscious semiotic mechanisms that manifest themselves in the signs we use (words and phrases used during conversations, in written texts, in kinesic and proxemic behaviors, and so on). The semiotic approach aims to understand how thought and feeling are intertwined with such verbal and nonverbal semiotic mechanisms.

One area where semiotics has been applied fruitfully is in education. A study by van der Schoot et al. (2009), for instance, examined the effects of the opposition consistent/inconsistent within a relevant arithmetic operation and markedness (the relational term being unmarked [“more than”] vs. marked [“less than”]) on word problem solving in a sample of 10–12 year old children

differing initially in problem-solving skill. The researchers found that less successful problem solvers will utilize a successful strategy only when the relational term is unmarked. In another significant and relative study, [Cho and Proctor \(2007\)](#) found that when classifying numbers as odd or even with left-right keypresses, performance was better with the mapping even-right than with the opposite mapping. Calling this a markedness association of response code (MARC) effect, the authors attributed it to compatibility between the linguistic markedness of stimulus and response codes. The MARC effect and its reversal are caused by a correspondence of the stimulus code designated as positive by the task rule with the positive-polarity right response code.

Semiotics is also expanding considerably in the context of changing technologies. For example, an area that is now utilizing the insights of semiotics more and more is in the study of social media discourse (see [Poulsen et al., 2018](#)). Hashtag literacy in particular is also becoming a relevant area of interest. A hashtag is a fascinating new kind of sign structure—it is a phrase beginning with the symbol “#” used as a way for tagging content so that interested users can locate it—that is, adding hashtags to a post allows users to find that post when searching for that topic. Hashtags can be used as well for finding users who are looking for topics on social media. Hashtags also have a “commentary” function. For example, a linguistic marker referencing the target of evaluation in a tweet not only renders the content searchable, but also constitutes a means to affiliate with values expressed in the tweet. This has led some linguists to call hashtags part of “searchable talk,” whose subtext is “Search for me and affiliate with my values” ([Zappavigna, 2019](#)). It also implies that hashtags can become ideological and politically subtle codes. In effect, digital systems now present new problems of semiotic analysis, expanding both semiotics itself and its potential applications.

Concluding remarks

The solution to the problem of what something *means* is not to examine meaning in any abstract way, but how it is encoded in specific ways through structures that have specific representational modalities—resemblance (iconicity), relation (indexicality), conventionality (symbolicity). As [Ogden and Richards \(1923\)](#) stressed in their classic study of this problem, meaning is to be understood not by analyzing isolated units, but by taking into account the forms that encode it. In this view, the real meaning unit is the sign.

Semiotics is ultimately a form of inquiry into how humans shape raw sensory information into knowledge-based categories through sign creation, that is, by using forms that stand for the categories. Signs are selections from the flux of information that is taken in by our senses allowing us to encode what we perceive as meaningful in it and, thus, to remember it. We are born into a system of sign use, called the semiosphere by the Estonian semiotician Juri Lotman (1922–1993), that will largely determine how we come to view the world around us ([Lotman, 1991](#)). The semiosphere, like the biosphere, regulates human behavior and shapes social evolution. However, although they can do little about the biosphere, humans have the ability to reshape the semiosphere any time they want. This ability to create new signs and sign systems is what distinguishes human semiosis from all other kinds of sign and signal systems in other species. Our textual resources, for instance, stimulate us to seek new meanings and new ways of seeing the world. These open up the mind, encourage creativity, and engender freedom of thought. As Charles [Peirce, 1931](#), vol. 2, p. 302) wrote so insightfully, it would seem that as a species we are inclined to “think only in signs”.

References

- Alpher, B., 1987. Feminine as the unmarked grammatical gender: buffalo girls are no fools. *Aust. J. Ling.* 7, 169–187.
- Andrews, E., 1990. *Markedness Theory*. Duke University Press, Durham, NC.
- Augustine, S., 1887. *De doctrina Christiana*. Patrologia latina, Migne Tomus XXXIV. Hachette, Paris.
- Battistella, E.L., 1990. *Markedness: The Evaluative Superstructure of Language*. State University of New York Press, Albany, NY.
- Benveniste, E., 1946. Structure des relations de personne dans le verbe. *Bull. Soc. Ling.* Paris 43, 225–236.
- Cho, Y.S., Proctor, R.W., 2007. When is an odd number not odd? Influence of task rule on the MARC effect for numeric classification. *J. Exp. Psychol. Learn. Mem. Cognit.* 33, 832–842.
- Danesi, M., 2007. *The Quest for Meaning: A Guide to Semiotic Theory and Practice*, second ed. University of Toronto Press, Toronto, ON.
- Deely, J., 2001. *Four Ages of Understanding: The First Postmodern Survey of Philosophy from Ancient Times to the Turn of the Twentieth Century*. University of Toronto Press, Toronto.
- Derrida, J., 1976. *Of Grammatology*. trans. In: Spivak, G.C. (Ed.). Johns Hopkins Press, Baltimore, MD.
- de Saussure, F., Baskin, W., 1916. trans. In: Bally, C., Sechehaye, A. (Eds.), *Course in General Linguistics*. McGraw-Hill, New York, NY. *Cours de linguistique générale*.
- Eckman, F.R., Moravcsik, E.A., Wirth, J.R. (Eds.), 1983. *Markedness*. Plenum, New York, NY.
- Eco, U., 1976. *A Theory of Semiotics*. Indiana University Press, Bloomington, IN.
- Fauconnier, G., Turner, M., 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. Basic Books, New York, NY.
- Fisch, M.H., 1978. Peirce's general theory of signs. In: Sebeok, T.A. (Ed.), *Sight, Sound, and Sense*. Indiana University Press, Bloomington, IN, pp. 31–70.
- Foucault, M., 1972. *The Archeology of Knowledge*. trans. In: Sheridan Smith, A.M. (Ed.). Pantheon, New York, NY.
- Greimas, A.J., 1987. *On Meaning: Selected Essays in Semiotic Theory*. trans. In: Perron, P., Collins, F. (Eds.). University of Minnesota Press, Minneapolis, MN.
- Hjelmslev, L., 1939. Note sur les oppositions supprimables. In: *Travaux de Cercle Linguistique de Prague*, vol. 8, pp. 51–57.
- Jakobson, R., 1939. Observations sur le classement phonologique des consonnes. In: *Proceedings of the Fourth International Congress of Phonetic Sciences*, pp. 31–41. Ghent.
- Jakobson, R., 1942. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Almqvist and Wiksell, Uppsala.
- Johnson, M., 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. University of Chicago Press, Chicago, IL.
- Lakoff, G., 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, Chicago, IL.
- Lakoff, G., Johnson, M., 1980. *Metaphors We Live by*. Chicago University Press, Chicago, IL.
- Lakoff, G., Johnson, M., 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, New York, NY.

- Lévi-Strauss, C., 1958. *Structural Anthropology*. Basic Books, New York, NY.
- Locke, J., 1690. *An Essay Concerning Humane Understanding*. Collins, London.
- Lotman, Y., 1991. *Universe of the Mind: A Semiotic Theory of Culture*. Indiana University Press, Bloomington, IN.
- Ogden, C.K., 1932. *Opposition: A Linguistic and Psychological Analysis*. Paul, Trench, and Trubner, London.
- Ogden, C.K., Richards, I.A., 1923. *The Meaning of Meaning*. Routledge and Kegan Paul, London.
- Osgood, C.E., Suci, G.J., Tannenbaum, P.H., 1957. *The Measurement of Meaning*. University of Illinois Press, Urbana, IL.
- Peirce, C.S., 1931–1958. In: Hartshorne, C., Weiss, P. (Eds.), *Collected Papers of Charles Sanders Peirce*, 1–8. Harvard University Press, Cambridge, MA.
- Pollio, H., Barlow, J., Fine, H., Pollio, M., 1977. *The Poetics of Growth: Figurative Language in Psychology, Psychotherapy, and Education*. Erlbaum, Hillsdale, NJ.
- Poulsen, V., Kvåle, G., Gunhild, van Leeuwen, T. (Eds.), 2018. *Social Media as Semiotic Technology*. Special Issue of *Social Semiotics*. Routledge, London.
- Tiersma, P.M., 1982. Local and general markedness. *Language* 58, 832–849.
- Trubetzkoy, N.S., 1936. *Essaie d'une théorie des oppositions phonologiques*. *J. Psychol.* 33, 5–18.
- Trubetzkoy, N.S., 1968. *Introduction to the Principles of Phonological Description*. Martinus Nijhoff, The Hague.
- Van der Schoot, M., Bakker Arkema, A.H., Horsley, T.M., van Lieshout, E., 2009. The consistency effect depends on markedness in less successful but not successful problem solvers: an eye movement study in primary school children. *Contemp. Educ. Psychol.* 34, 58–66.
- Zappavigna, M., 2019. *Searchable Talk: Hashtags and Social Media Metadiscourse*. Bloomsbury, London.

Further reading

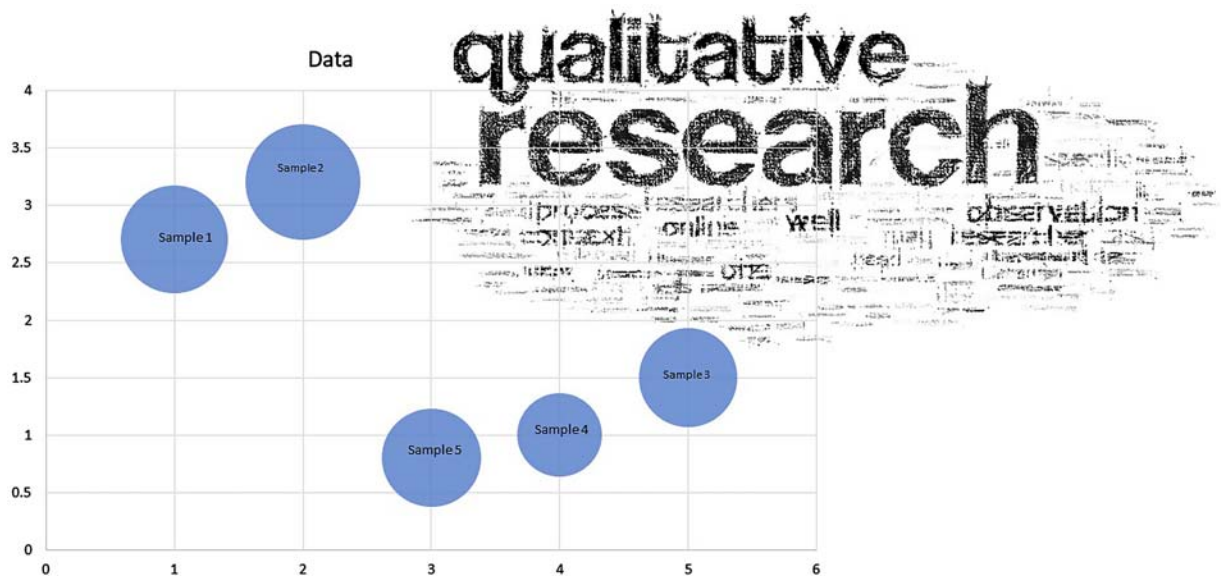
- Bühler, K., 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Fischer, Jena.
- Danesi, M., Santeramo, D., 2020. *Semiotics: The Science of Signs*. Lincom Europa, Munich.

Qualitative sampling and qualitative data

Mirka Koro^a, Nikki Fairchild^b, Angelo Benozzo^c, and Teija Löytönen^d, ^a University of Helsinki, Helsinki, Finland; ^b Research and Innovation, School of Education and Sociology, University of Portsmouth, Portsmouth, United Kingdom; ^c University of Turin, Turin, Italy; and ^d University of Arts Helsinki, Theatre Academy, Helsinki, Finland

© 2023 Elsevier Ltd. All rights reserved.

Sample 1: Data history in qualitative inquiry	199
Sample 2: Data reconceptualized	200
Sample 3: Data as being with	201
Sample 4: Sampling processes and strategies in conventional humanistic qualitative research	202
Sample 5: Alternatives to representational sampling strategies	203
Sample 6: Exercise in perhaps-cutting	204
Sample 7: Exercise in <i>ritornello</i>	205
Sample 8: (Political) drive of data and sampling	205
Sample 9: (Im)possible questions shaping data and sampling futures	208
References	208



The adjective “qualitative” shapes sampling and data significantly and in many different ways. For example, the definition of “qualitative” bounds, enables, and disables specific kinds of research practices and discourses. Schwandt (1997) defined “qualitative” as a “not-so-descriptive adjective attached to the varieties of social inquiry that have their intellectual roots in hermeneutics, phenomenological sociology, and the Verstehen tradition” (p. 129). He explained that qualitative connects with quality as “an inherent or phenomenal property or essential characteristics of something (object or experience)” (p. 130). “Qualitative” can also be a (research) choice, a decision to take a different ontological turn away from realism, to take multiple epistemological directions, and to constitute one’s (research) world beyond quantification and numeric values. A particular type of data, material, or object of inquiry does not make data qualitative or quantitative, but the ways data and processes are used and those theoretical assumptions that guide data and sampling practices can infect either qualitative or quantitative. Thus, “qualitative” in this chapter is conceptualized as an intentional move beyond positivism, predictability, and quantification, a shift away from independent, singular elements, and generalization, toward a focus on experiences, emotions, affects, materiality and matter, an emphasis on the quality of entities and relations. We would like to note that the ways in which qualitative research is practiced operate in a continuum. For example, sampling and data do not operate as binaries (e.g., positivist and post-positivist research, traditional and experimental data) but scholars work with these concepts within a continuum where theoretical orientations and scholarly preferences can be viewed as degrees.

We are less interested in how sampling and data could be labeled, categorized, and practiced in one unified way. The purpose of this chapter is not to tell scholars how to sample or define data. Instead, we aim to diversify sampling and data practices in order to illustrate their multiple connections to theories, processes, world views, and subjectivities these concepts and practices carry. The ways data and sampling are approached and used form important choices scholars make and data and sampling are always connected and influencing each other. Similarly, decisions researchers make about data and sampling impact how data and sampling work together theoretically and practically. Ideally, these decisions should align and support similar onto-epistemological goals and purposes.

Data and sampling processes' intertextuality and relationality (how data and sampling relate to other qualitative concepts and practices) can create links to histories, discourses, other scholars, diverse materials, with endless traces of conceptual and practical differences. The call for logical and rational approaches to sampling and data is often challenged by the affective, sometimes surprising relationalities that is produced on/with scholars. However, this challenge acknowledge the different characteristics and assumptions while at the same time acknowledging that these characteristics blend and come together different ways. We aim to illustrate what these thoughts/doings/entanglements/relationships might produce in/out/through/beyond readily available and normative views of sampling strategies and data. We are working within an interesting paradox: data and sample are always already here even when we do not consciously attend to them. For example, any act of thinking, sensing, collecting, organizing, relating, and living are samples, selections, choices, and as such always signs of partiality.

In the past ten years, the notion of data has been troubled by an increasing number of scholars (see e.g., Koro-Ljungberg and MacLure, 2013; Koro-Ljungberg et al., 2017; Taylor et al., 2018). However, similar problematization has not occurred with the concept of sampling. Thus, in this chapter we introduce data and sampling (forms) that do not have it/they-ness, singular objectivity, and/or a clearly defined past. Instead, we advocate for processes and research designs which are relational and collective and as such always plural and multiple. Maybe all of us (scholars and qualitative researchers) are active data-making subjects and objects, actors and actants in collective sampling networks and ontological structures which enable data emergence and functioning. For example, how would one explain data which stubbornly lingers, emerging at some point as text, images, sensations, movements, edible pieces of ink and paper, disturbing sounds, noises, and later traces of these data entanglements return in relations established in dark and wet Finnish forests, reminders of sun rays in the sky space of a southwest art museum, and in the faces of happy children playing in the playground? In so many ways, scholarly practices and inquiry as we presently sense them are inseparably entangled with our past experiences, enactments, and appearances of data and more.

This chapter has been written up as a sample itself. A sample as used here and radically reconceptualized later in this chapter is no longer about systematic, methodological, or random selection of data but it reminds us about processes of selecting, filtering, being aware, acknowledging, choosing, relating, exercising agency, and living in posthuman worlds. Sampling creates knowledges and realities similar to the texts, ideas, and experiments shared within this chapter. Throughout this chapter we sample what and how to write about sampling and data practices. We sample and share samples in the form of examples, theoretical connections, and literature we utilize. We also selectively filter our thoughts and claims thus creating a sample of sampling. Our first cluster of samples focus on data, second set of samples addresses sampling, then we offer few exercises for readers to engage with data and sampling at deeper and more personal level. Finally, we conclude with sample where we address politics, drive, and extensions of data-sampling leaving readers with provocative questions.

Instead of rejecting sampling altogether, we want to recognize that sampling might be seen as a continuous process of choices, selections, and decisions that are entangled with the process of inquiry. During a research process we sample/decide/select: what to study, who to listen to, what to ask, which phenomena to interrogate, what to read, what to write, where to go, how to approach ethics and rigor, and more. At the same time, *we (researchers) are sampled by*, because something in the process chooses us: it can be an utterance, a written sentence, an image, an object, an atmospheric event, a splinter of material and ... and ... and In the end, we sample something and we are sampled by *something*.

Our samples draw from diverse traditions, experimentations, past, becomings in a shifting and pulsating fashion. Samples shared here are meant to be read and experienced in any order, with or without skipping over some parts, returning back to previous samples, through returns, and/or to be read over each other. They are connected by a rhythm which brings readers back and forth, far and distant from the data and sampling processes, and through the familiar and unfamiliar as different forms of unexpected provocations emerge. We find it important to problematize some historical and traditional discourses and practices therefore some of our samples offer radical reimagining and different ways to think with the futures (of data and sample).

Sample 1: Data history in qualitative inquiry

The history of data and their conceptualizations in qualitative research often stems from quantitative, positivist conceptualizations of data and science. Even today some qualitative scholars refer to data being sought, selected, collected, located within acceptable practices and visible formats (see e.g., Braun et al., 2017; Taylor and Francis, 2013). Much has been written about data collection in qualitative inquiry (see e.g., Guest et al., 2013; LeCompte and Preissle, 1993; Weller and Romney, 1988). More recently, data and data "collection" has moved beyond traditional qualitative approaches such as ethnographical interviews, observations, and collection of artifacts (see e.g., Wolcott, 1999) to include various visual, virtual, performative, collective, and hybrid forms of data (see e.g., Braun et al., 2017; Koro-Ljungberg et al., 2017). Yet many studies and qualitative research practices have been influenced by quantitative method and its claims to rigor, validity, reliability, and "scientific" evidence. While earlier qualitative data have been

considered to “capture and represent the real as experienced and lived”, they have also been influenced by human judgment and subjective interpretations. In the past many qualitatively scholars have been concerned how to measure, sample, validate, and systematically order data. Often data and their qualities have been reduced to the materials collected to evidence arguments. For example, [Denzin’s \(1970\)](#) edited collection noted how the “problem of data analysis” largely centers on researchers’ abilities to create “sound causal explanations” (p. 3). Later in the 1970s, qualitative data were connected with ethics and ethical responsibilities of a researcher. Scholars were concerned with the questions that addressed who owned data materials and who had rights to data. At that point it was noted that data can be dangerous and harmful, and their generation needs to be monitored. In the first edition of the handbook of qualitative research ([Denzin and Lincoln, 1994](#)) data were referenced as a type of empirical material and discussed in chapters dealing with other issues such as data collection and analytical processes. [Gubrium \(1988\)](#) discussed data as field data, whereas [Fielding and Fielding \(1986\)](#) distinguished between qualitative (real and deep data) and quantitative data (hard data). [Weller and Romney \(1988\)](#), in turn, proposed that data collection must be systematic with data collection methods independent from data those methods produced. Additionally, [Wolcott \(1994\)](#) discussed the need to use the term data since physical and natural science demanded it. Otherwise, ethnographers would just do fieldwork without dedicating anything to “be data” per se ([Wolcott, 1994](#)). Wolcott proposed data are the “first evidence of your efforts, something you gather or collect or generate or create, depending on the tradition in which you are working” (p. 3). Later, being influenced by poststructuralism [St. Pierre \(1997\)](#) proposed that she would continue to use the word data, but their meaning has now shifted. Not only did [St. Pierre \(1997\)](#) problematize the “data signifier” but she also proposed alternative forms of data such emotional data, dream data, sensual data, and responsive data.

One might also argue that data are always situated and shaped by researchers’ episto-ontological frames. For example, data within positivism is often collected, gathered, and systematically documented. Data are treated as an easily identifiable set of materials and the object of inquiry. In comparison, from a constructivist perspective data could be seen as created and potentially co-constructed. Data could be located in the mind, perceptions, imaginations, shared ideas, collective projects and more. Critical theorists might approach data as something to be emancipated, transformed, generated through critique and within different practices and networks of power. For many post scholars data are to be re-thought, re-conceptualized, deconstructed, performed, questioned in a variety of ways where the idea of stable data and data identifiers are problematized and troubled. Some new materialist scholars may approach data as forces, flows, matter, and forms of living relationally. Response-ability ([Haraway, 2016](#)) could be used to refer to the ethical uses and practices of new materialist data. Finally, from a decolonial perspective data are always in some ways connected to the land and ownership of land. Data, of these kinds, are linked to colonial practices, colonized knowledges, and various subjects and objects of colonialization.

Sample 2: Data reconceptualized

When one thinks about more contemporary forms and processes of qualitative data something interesting is always happening. However, to define or bound this “something” with data might be only partially possible, sometimes impossible, and even undesirable. Sometimes qualitative researchers may find that data take multiple forms, and they cannot put data (object) aside. Despite carefully and well-designed studies scholars might find themselves with too much, too little, too broad and too detailed, too personal and too communal data. Due to these reasons some scholars (see e.g., [Koro-Ljungberg et al., 2017](#)) have moved beyond data as interviews, observations, or artifacts toward relational data, more-than-human data, experimental data, performative and embodied lifedata, data that are so entangled with life that they become multiple. Data such as these are always plural, live, and stimulate or enact processes that work through life, move beyond and between life and yet still always return to lives. Data of these kind exemplify multiple simultaneous relations and connections, connectivity in past-presence, and diverse timespaces. They are theoretical formed and ontologically situated. Data as multiple and lived are stubborn and through this stubbornness, can offer scholars possibilities to inquiry differently, and immanently, and experimental ways.

Experimental data and speculative data aim to generate different futures and methodological worlds which may build on radical imagination (see [Haiven and Khasnabish, 2014](#)). Furthermore, speculative fiction and data’s futuristic orientation shape the inquiry, imagination, and processes of many qualitative researchers. For many followers of process philosophy, data are always already relational, but they also form and constitute immanent (dis)orienting interactive effects and various diverse forms of relational potentiality and actualizations. For example, increasing numbers of post-qualitative scholars continue to work through radical imagination, performative, more-than-human data and their disorienting effects to rethink and recreate more equitable futures. Ecological data, eco-data, relational data, data assemblages, intra-active data, multispecies symbiotic assemblages highlight some possibilities of more-than-human entanglements ([Koro-Ljungberg et al., 2017](#)) and the importance of global ecosystems and networks.

Those scholars who would like to counter data that are already predefined and pre-categorized by previous representationalist processes can adopt strategies which build on the relational, affective, sensual, and embodied elements of living and experiencing. For example, Vannini’s conceptualizations of data related to island life did not attempt to classify types of islands or forms of island life, but “how one kind of islandness is done ... [through] embodied experiences that islandness demands and affords” ([Vannini and Taggart, 2012](#), p. 227). For Vannini and Taggart, “islandness” is a creative, inventive, and becoming verb rather than a fixed and stable representation noun of “an island” as already imagined and constructed. Similarly, [Pink’s \(2014\)](#) digital–visual–sensory data considered “how the temporalities of visual ethnography research might be engaged to think about the future” (p. 414). Furthermore, the work of Manning, Massumi, and Sense Lab (<http://sensemab.ca/wp2/>) exemplifies experimental, living, and collective

methodologies (Manning, 2016, 2020). Thinking inquiry as a thought in the act (Manning and Massumi, 2014), turns inquiry into research-creations exemplifying data's various forms. From this perspective, data function as events in the making, "data and sampling" processes are considered as pragmatic and "useless" (without predetermined intentions) while highlighting the world of sensing, neurodiversity, textures, living, and touch. Research-creations create "thinking-doing" and "forms and forces of intellectuality that cut across normative accounts of what it means to know" (Manning, 2016, p. 27). In Springgay and Truman's (2016) Stone Walks project, walking with rocks generated unexpected connections with more-than-human worlds and data's materiality noted how rocks may move and reproduce prompting speculative thought on inhuman animacies. Furthermore, in Stephens and Delamont's (2006) two-handed ethnography, data of a dancing body met language, rhetoric, and cultural understandings, producing bodily and embodied ethnographic data. Finally, inefficient mapping was taken up in Knight's (2016) project studying playgrounds and urban planning. Playgrounds and urban spaces were seen as data "sites brimming with activity, the intensities brought on by collisions and interceptions between aspects and agents" (Knight, 2016, p. 22) rather than bounded, pre-determined and pre-conceptualized architecture and architectural functions.

While engaging in inquiry, sleeping, taking showers, and cooking dinner we wonder how data are actively persisting, returning, and continuously moving us and others. Is it possible that data circles back and returns in their different formation and through emerging and multiplying events? Performative notions and practices of data are often used by critical, postmodern/poststructural, process oriented, and posthuman among other scholars to emphasize data's vitality, relationality, and generativity. Data that was once bitten, eaten, and digested as a ball of transcribed interview text can possibly shape how other scholars wash data and encounter data while walking their dogs and thinking about their inquiries and participants. Relational data bring together event, actors, systems and always more. They connect life and events surprising scholars again and again. Data bags and bag-data bring scholars together, join lipsticks, laptops, and notepads carried in conference bags, Emily Muller meets Ikea bag, hotels, and personal belongings hidden and displayed in hotel rooms and more (Benozzo and Koro-Ljungberg, 2017). Data linger and morph through events, relations, lives, and processes researchers encounter. Data live in conference spaces, hotel rooms, black holes, books, rain drops and more. Data meet the boundaries of established spaces, ghostly and silently creeping and affecting and being affected by same and different data and datalives. Data produce so much and more. In some ways trying to capture these productions through fixed language and bounded scholarly techniques can be frustrating and also perceived as useless (Manning, 2020; Fairchild et al., 2022).

Data's relatedness and stubborn entanglements have less to do with fixed researcher identities, scholarly attachments, or stable emotional commitments. Instead, stubborn data entanglements exemplify continuously changing forms, processes, and relational contexts that persistently shift and change inquiry while being responsive to relational differences. Why does stubborn data become this or that has no reason at all? Proactive affirmation of data's generative potential produces. "We" are data together. Living and relational data become a "subject" and "subjects" not an "object" of inquiry. Stubborn data can also shape the politics of locations through diverse assemblages including events of remembering, counter memories, cartographies, and transformations (see Braiddotti, 2013). Bennett (2010) discusses assemblages as "ad hoc groupings of diverse elements, of vibrant materials of all sorts. Assemblages are living, throbbing confederations that are able to function despite the persistent presence of energies that confound them from within" (pp. 23–24). For Bennett assemblages are stubborn, persistent yet emergent. "The effects generated by an assemblage are, rather, emergent properties, emergent in that their ability to make something happen" (p. 24). An open-ended collective assemblage of data has vitality. These data assemblages are also vibratory; mood, style, content, membership change, alter and produce endless data variations in their relationality. Entangled states of A and B are understood as a single data entity. The fate of data is entangled with the fate of researcher, theories, philosophies, inquiries at large, practices, senses and more. "Knowledge is not a matter of reflecting at a distance; rather it is an active and specific practice of engagement. To know is to become entangled; objectivity requires that one take responsibility for one's entanglements" (Barad, 2007, p. 453).

Sample 3: Data as being with

In working on the idea of *constitutive withness or generative entanglement* Pyyhtinen (2016) draws from Deleuze and Guattari. According to Pyyhtinen (2016) the concept of with-ness "removes the focus from individual actors to dynamic relations" (p. 24). Understood in this way data begin, produce and are produced, emerge in the in-between, in the middle of things since "[n]othing exists solely in and by itself, but entities are what they are only in and through their relations with others" (Pyyhtinen, 2016, p. 18). Every entity like data is only constituted in relations, individualization is a collective and relational process where data come to be. Individuation in the fielding of singular-plural; always more than one. Data are traversed by lines, threats and forces. Stubborn data might operate also as a relational verb or relation par excellence. "There is no substance to entities other than difference and event, the occasion of their associations. They are grounded in multiplicity; each singular being is a multiplicity in itself" (Pyyhtinen, 2016, p. 25). Following Deleuze and Guattari (1987) data multiple is not about adding or summing up but subtracting data, multiple = $n-1$. Relational data form ontological positionings and as such data do not function as a knowledge project but constituting possibilities for data's becoming. Data to the n th power could also manifest themselves as *mad*, experimental, agentic data. Maybe also exception data and ignorant data. In its insisting and continuous questioning and stubbornness ignorant data may stimulate and produce innocence and creativity. Ignorant data refuse recognition and they are incapable to follow established norms.

Qualitative data do not function as a thing-in-itself but possibly a force and relational energy field producing complex ecological and relational patterns. These patterns shape the practices, mundane everydayness of scholars and participants at the same time

moving beyond this everydayness toward particular and specific event time. New imaginings, becomings, and reconceptualization's of data have produced and continue to create different assemblages, ecological, and material systems shaping inquiry, researchers' relationality and their thinking-doing. Many qualitative researchers might have become more hesitant and unsure about data while at the same time entangled more than ever before with unpredictable, surprising, and returning data. In some ways data have become an unanswerable maybe even a hypothetical question rather than an answer. Stubborn data cannot be taken for granted but they return with new tricks, creations, and companions. Qualitative scholars might no longer ask definitional, singular, and categorical questions about what data are *but* how and what stubborn and entangled data produce and how might it transform. What are some possibilities of data?

Field data	Sensual data
Observation data	Invisible data
Archival data	Sound data
Visual data	Memory data
Indigenous data	Rhythm data
Missing data	Invisible data
Natural data	Hard data
Inaccurate data	Sense data
Meta data	Big data
Performance data	Thick data
Coded data	Virtual data
Raw data	Response data
Interview data	Ghost data
Land-data	Entangled data
Dream data	
Real data	Escaped data

Sample 4: Sampling processes and strategies in conventional humanistic qualitative research

The approaches of sampling in qualitative research have evolved significantly from the aims and processes that quantitative sampling signify. Many elements contribute to the differences found in sampling processes. For example, the definition of sampling varies, sampling units differ (from participants to textual elements, to sounds, place and space, matter and materiality, affect, and bodily sensations), scholars' approaches to saturation and timing of the sampling itself. [Guest et al. \(2013\)](#) referred to sample as a collection of a subset of units (people or otherwise) from a larger population related to a particular study. Oftentimes, scholarly conventions present among researchers highlights the importance of transparency and the need to provide an indication of quality and validity of research for scholars outside specific research traditions. [Marshall \(1996\)](#) suggests the aim of all quantitative sampling approaches is to "draw a representative sample from the population, so that the results of studying the sample can then be generalized back to the population" (p. 522) where participants are often selected randomly based on predetermined hypotheses and sample sizes. The trans-disciplinary nature of qualitative research sees sampling approaches that give richness, deeper insights and in-depth understandings of a particular phenomenon (see e.g., [Staller, 2021](#)). Instead of relying on the randomly selected representative samples of a particular group or population, qualitative sampling is often based on the purposive and flexible invitation of participants and the selection of diverse data (descriptions, narratives, documents, artifacts, visual, virtual, collective, and hybrid forms of data etc.) which might evolve even throughout the research process. Qualitative sampling requires specific attention to the research process and its evolution, ethical considerations, and study contexts, including the holistic socio-political and material elements of the study: the individuals, the temporal, spatial and situational elements. The following section describes the logic of sampling from a conventional (humanistic) qualitative research perspective. Sampling in conventional qualitative research is often concerned with representativeness and meeting of predetermined criteria. Sampling as an *a priori* process is also exemplified in dissertation proposals and Institutional Review Board processes. It has been argued that *a priori* sampling could form a useful yet tentative plan for more flexible and situated research approaches (see [Gentles et al., 2015](#)). Instead of proposing a stable understanding of qualitative sampling, we then problematize and reconceptualize sampling and its strategies. We aim to diversify the understanding of sampling and illustrate its multiple forms and the choices scholars make ongoingly during any qualitative research project.

This kind of sampling is often connected to post positivist, interpretivist, and sometimes critical traditions, in which the focus is on lived experiences and affects, on how people make sense of their (local) lives in diverse contexts, oftentimes in and through language. In many cases sampling is directly informed by and dependent on the methods selected such as qualitative interviews, participatory research approaches, and ethnographic methods. For example, interview methods and sampling within the context of interviews vary based on the disciplinary backgrounds of the researchers and the epistemological and ontological presuppositions of scholars ([Guba and Lincoln, 1994](#)). Sampling of this kind implies the need to indicate the *number and characteristics of the participants* interviewed and the reasons that led to choosing these participants. Sampling also involves decisions about timing

and when participants will be invited to join the study and research activities. In addition, scholars make decisions about the degree of their involvement in sampling processes. [Guest et al. \(2013\)](#) suggest that scholars should clarify how much control researchers have over sampling and recruitment and consider how flexible their study designs might be.

Sampling within conventional qualitative research is also concerned with numbers. Yet within the existing literature on qualitative interviews and on the selection/sampling of participants, a precise number of participants is rarely indicated. The study's purpose and the ontological, epistemological, methodological, and practical issues can guide the sampling process. [Brinkmann and Kvale \(2015\)](#) suggested 5–25 participants for a qualitative study depending on the aim of the research project. Some authors argue that if the research purpose is to provide a rich and dense account of an event or experience then even one interview may be appropriate ([Becker, 2012](#)). The number of participants may also depend on the balance between representativeness (in a broad sense) and the quality of the responses when obtaining sufficient information. If the purpose of a study is to establish commonalities or allow for comparison, then a larger sample of participants are likely to be appropriate. If one was to be concerned with data saturation, the number of participants should not be so small that it makes data saturation difficult and at the same time sample should not be too large to compromise in-depth and detailed data analysis.

Some researchers argue that ideally data collection through sampling should continue until saturation ([Morse, 1994](#)) or until informal redundancy ([Lincoln and Guba, 1985](#)). For some scholars the concept of saturation, originally derived from grounded theory ([Glaser and Strauss, 1967](#)), is considered the “gold standard” and saturation is reached when new participants or new data do not produce any additional theoretical or thematic insights (theoretical or thematic saturation). However, some scholars consider the adoption of saturation of sampling as a generic indicator of quality inappropriate and inaccurate across different theoretical orientations where rich, detailed, and in-depth data can be interpreted in multiple different ways.

Convenience sampling refers to the sample where participants are conveniently available for the researcher and are willing to take part in the research project (e.g., [Staller, 2021](#)). [Alvesson and Ashcraft \(2012\)](#) cautioned against overemphasizing the assumed richness of responses and relying too heavily on interviewees/participants that researchers have easy access to. While convenience sampling might be suitable sampling strategy for studies where research questions do not bound, characterize, or call for a unique set of participants, the outcomes may also decontextualize the phenomena being research by overly divorcing participants from the research aims. Even if the qualitative sampling of interviewees/participants or data did not need to represent a particular group or population to serve generalization, the convenience sample might limit insights gained from the data.

In addition to convenience sampling, purposive sampling strategies can serve the aims of a specific qualitative research project (see e.g., [Staller, 2021](#); [Marshall, 1996](#)). These include, studying a broad range of participants (maximum variation sample), outliers (deviant sample), participants who have specific experiences (critical case sample) or participants with special expertise or experience (key informant sample). Participants might also suggest others for the research project, which some scholars call a snowball sample. [Guest et al. \(2013\)](#) also propose homogenous sampling which highlights similarities among participants; intensity sampling that intensify the phenomenon under study; politically important sampling that includes politically sensitive cases; and criterion sampling builds on pre-existing criterion. The work of Joseph Maxwell cited in [Emmel \(2014\)](#) highlighted how realistic sampling balances ideas and evidence, theory and practice. Realistic sampling choices demonstrate different how the relationship between ideas and evidence are conceptualized. In addition, inductive, emergent or theoretical sampling (in grounded theory terms) allow for groups not specified in the original research design when data collection and analysis process evolves and more samples are needed ([Guest et al., 2013](#)). These different sampling procedures are expected to serve the research questions and produce interviewees, participants, individuals, groups, settings where and for whom the phenomenon being studied is most likely to occur.

Sample 5: Alternatives to representational sampling strategies

The idea of sampling as a form of representative reality can be troubled from postmodern, post-structural, post-qualitative and other post perspectives. The representational forms of sampling logic are based on a process that creates structures and categories, stasis out of movement and proliferation (see [MacLure, 2013a](#)). Even though structures and categories could be helpful to order and investigate the known and unknown, this kind of sampling easily reinforces simplified categories or hierarchies of representation, which are often based on identity, similarity, and opposition ([Löytönen, 2017](#)).

Representational uses of sampling may also be seen as a process of separation to demarcate the field, for example identifying individuals, newspaper and video, sites and places that might be relevant and give “information” about the phenomena under investigation. Sampling could imply a selection and process to define a perimeter/and edge/and end of the field and then it becomes a limitation. If sampling is a decision to separate, what are we leaving outside and aside? If sampling is taking a decision, it implies that we are dividing (to take a decision derives from the Latin *de caedere* = to cut out), but can this action be also violent, difficult to carry out, and full of contradictions? And what about what is left aside? Where is this potentially relevant excess gone?

Scholars who are interested in situated and fragmented knowledges, who are excited about experimental and artistic forms of inquiry, and those who desire to emphasize relationality and more-than-human approaches may find many alternatives. For example, one can think of sampling as a form of cutting together-apart, an agential cut ([Barad, 2007](#)). Cutting is never a separation, or it is never a complete separation (a final break which separate subject from object), because things are always already in relation.

Each sample cuts out a part of the world, excluding and including some things over other things, which is cutting together-apart or cutting apart-together. The researcher can choose which part to use from a recorded or transcribed sentence and which part can be cut. A research paper presents a specific selection of the world and could be different if other cuts/sentences were used. We can read the research paper through the technology of a computer which determines what we see. In this way agential cuts are inclusions and exclusions, we are responsible and accountable for the cuts we make, we practice our response-ability as an ethical sensibility and response to others (Haraway, 2016).

Barad's concepts of agential cut and cutting together-apart offers useful perspectives of inclusion and exclusion practices within qualitative research. According to Barad (2007), a cut is never a definitive separation, it is never a complete division (a kind of break that separates the subject from the object), because the data, the subjects, texts, images, videos are always and already in relationship with us, between them and more. Other scholars have created concepts that also open new potentialities in conceptualizing research-creation processes through non-hierarchical or non-categorical approaches. Some of these approaches could also be used in the context of sampling, for example concepts of assemblage (Deleuze and Guattari, 1987); entanglement (Barad, 2007); or actor network (Latour, 2007). As indicated with Barad's concepts of agential cut and cutting together-apart, these concepts might reframe how sampling is theorized and practiced, how things are always already relational and how they come together, or are brought together in different formations. Thus, different theoretical frameworks and concepts might offer scholars approaches beyond readily available and normative views of sampling strategies in qualitative research. The different ways of framing and theorizing sampling might then offer scholars possibilities to sample differently, or sample in a more nuanced and fluid way, and thus maybe move toward more relevant research creations in each specific project. We might call these kinds of sampling approaches the "logic of unholy mixture" as described by Lecerle (2002, p. 54; see also MacLure, 2013b).

What if qualitative sampling was understood as an assemblage, an entanglement, a collage, choices ranging from theoretical, philosophical, data, textual, visual, toward assemblages, and entanglements? Might this open up research practices for novel entanglements that allow thinking and practicing differently? These kinds of differentiating perspectives on sampling may allow thinking-feeling, movement-moving and more.

Purposive sampling	Open sampling
Convenience sampling	Awkward sampling
Representative sampling	Singularity sampling
Typical sampling	Deviant sampling
Intensity sampling	Loose sampling
Maximum variation sampling	Minimum variation sampling
Homogeneous sampling	Heterogeneous sampling
Revelatory case sampling	Opaque case sampling
Critical case sampling	Indiscriminate case sampling
Opportunistic sampling	Normative sampling
Confirming sampling	Disconfirming sampling
Snowball sampling	Waterfall sampling
Theoretical sampling	Poetic sampling

Sample 6: Exercise in perhaps-cutting

Experimental approaches to sampling need to be practiced, tried, and exercised. Let's begin!

"... [the] pastime of gathering very different objects and materials together and trying to classify them started in the sixteenth century, and this marked the beginning of the phenomenon of Wunderkammern or cabinets of curiosities which brought together various pieces from the world around us, a world deemed wonderful and full of amazing surprises. In the Wunderkammer, very different objects and materials – naturalia et artificialia – were juxtaposed, being placed alongside each other in daring combinations. Wunderkammern were places filled with natural (naturalia) and artificial (artificialia et mirabilia) things: optical instruments and games, mechanical toys, natural history specimens, maps, precious gems, anatomical limbs or anomalies, lenses, mirrors, strange objects, etc. They were places where things were accumulated and piled up without any clear order and where connections came into being without there having to be a logical reason for them.

Benozzo and Gherardi (2020, p. 151).

"... the wondrous is a meta-historical category that has been defined all along the eighteenth century, didactically first and foremost, as a form of knowledge, that is, a very special half-way stage, a kind of mental suspension that lies between ignorance and knowledge, which marks the end of ignorance and the beginning of knowledge.

Lugli (2006, p. 126).

Create your own Wunderkammern

Take half a day off
 choose a place that you love:
 a city, a village or ...
 wander around and gather or buy *naturalia, articialia et mirabilia*

After one week
 Take half a day off
 choose a place that you love:
 a wood or a garden or a river or ...
 wonder around and gather what you like (*naturalia, articialia et mirabilia*)

After another week
 Take half a day off
 choose a place that you love
 prepare your own Wunderkammer
 invite some friends
 introduce your Wunderkammern
 allow your friends to play with your *naturalia, artificialia et mirabilia*

Sample 7: Exercise in *ritornello*

Create your own ritornello

Gather your research journals or notebooks.
 Explore the journals/notebooks by looking for repetitions,
 repetitive phrases, repetitive expressions, wordings, visuals.
 Assemble these repetitions,
 bring them together
 in new combinations,
 maybe in a collage.
 Create your own *ritornello*, a *refrain*,
 or "little song"
 which might start to establish
 "the beginnings of order in chaos [Deleuze and Guattari \(1987, p. 323\)](#)).

Note: The ritornello, refrain, as noted by [MacLure \(2016\)](#) accomplishes holding and bringing togetherness to various heterogeneous elements.

Sample 8: (Political) drive of data and sampling

adjective

adjective: data-driven

1. determined by or dependent on the collection or analysis of data.

One could argue that (qualitative) data and sampling strategies in qualitative inquiry are always political and moving. As noted earlier qualitative research has often been seen as the less significant and important form of inquiry, with concerns raised over validity and reliability. Similarly, data and sampling strategies used by qualitative research are sometimes questioned as non-scientific, subject to bias with a lack of rigor ([Ercikan and Roth, 2006](#)). This has been coupled with a skeptical view of research that is focused on the "posts" (postmodernism, poststructuralism, postcolonialism, posthumanism, feminist materialism) where policy makers in particular do not always see the value of fluid and nuanced worldviews ([Wolgemuth et al., 2021](#)). In much of education research there has been a focus on randomized controlled trials and research that invokes a "what works" view of education ([Styles and Torgerson, 2018](#)). Generally, this has been a response to social policy desires to narrow educational attainment gaps to reduce inequality (see [Biesta, 2010](#)). However, it could be argued that this focus on data and sampling has driven aspects of education

policy to become more performative, more about metrics, linked to return on investment narratives (Heckman, 2000). Data monitors attendance (Bodén, 2015), governance (Williamson, 2016), attainment and assessment (O’Keefe, 2017; Kippin and Cairney, 2021), progress (Fairchild, 2020), employability (Molesworth et al., 2011) and drives policy decisions (Williamson, 2018). For many data become a currency that places a certain type of value on outcomes and, in turn, drives outcomes-based research and pedagogy (Gunn, 2018; Bradbury and Roberts-Holmes, 2018). Data and systematic and random sampling strategies, therefore, become political, part of the capitalist machine—an output, a mode of being, a performative tool—producing dividual subjectivities (Deleuze, 1992) and segmenting bodies (Deleuze and Guattari, 1987).

What might be possible if we let data drive ... where would that take us? (Figs. 1 and 2)



Fig. 1 Data driven gender-in-the-making.



Fig. 2 Data storytelling multiple positionality.

Data and sampling processes are not neutral; they hold the politics and mattering of research and research practice. The way they are **driven** can have significant ramifications for the outcomes of research and practice. If we start to problematize the notion of data and sampling and where they **drive** us we can re-think and reconsider what **data driven** research might look like. [Deleuze and Guattari \(1987\)](#) developed the term line of flight as a neologism that highlights the infinite possibilities available when we break from accepted and known norms of doing and being. Lines of flight can act as disrupting mechanisms, they can jolt or shock us, they can excite us and reveal different ways to think about phenomena.

Methods-as-affinity-groups were the means to generate data for this inquiry, however, I was left wondering what 'data' would be presented to me during the field work. I was left considering a number of points, such as 'What is data? What does data feel about these books descriptions? Does data want to be represented, coded and classified?

[Fairchild \(2017, p. 81\).](#)

When we start to rethink **data driven** research we can consider how data becomes a driver powered by forces of Deleuzo-Guattarian desire ([Deleuze and Guattari, 1987](#)) ... data become a desiring machines. Desire is conceptualized as a positive and active actualization of reality rather than as a psychological state of lack. Desire is a force or an intensity, an affect which can manipulate meaning and relations, inform and fabricate meaning and generate intensity. Desire is more than an individual need for something, it shapes and connects people, objects, things, matter, discourse and politics. It can reveal how data-bodies are made and remade and the physical, material and political affects that shape them. Forces of desire become machinic, they become self-replicating and a multiplicity—desire can only be understood by what it produces collectively and not by the individual components of the machine. Therefore, a desiring machine is a data-body (in the broadest sense—see [Baugh, 2010](#)) that connects with other bodies when assemblages coalesce and form around particular events producing a new body, experience or happening ([Deleuze and Guattari, 1983](#)). The concepts of lines of flight and desiring machines can reframe to nature and notion of **data driven** as a productive force or intensity that generates a multitude of new possible insights rather than the data being used to denote the fixity of "truth" or "certainty".

Data-as-a-desiring-machine ([Fairchild, 2017](#)) highlights the instability of the actuality of "data" and "sampling" that can be used to predict or fashion a response to a research question. Scholars have argued whether the traditional notions of data and sampling are even required in a post-positivist world ([Denzin, 2013](#)). If we start to question the form and function of research and what it might require, perhaps rather than **data driven** research we might stumble over data ([Brinkman, 2014](#)) and sampling, they might be a splinter or be a nuisance ([Benozzo et al., 2013](#)) or something we wonder at [MacLure \(2013a\)](#). Following these lines of flight can reveal the productive nature of data and sampling assemblages that do their own **driving** (see [Fairchild et al., 2022](#)). Reconsidering what **data driven** research might look like can highlight the influence of affects and sensations that are immanent, exciting, entangling, provocative, disturbing and politically informed. This can be exemplified in the work of those using "post" ontologies and epistemologies as a means to provide new and different understandings of research and intersectionality (see [Koro-Ljungberg et al., 2017](#); [Rantala, 2019](#); [Taylor et al., 2020](#)). Data-desiring-machines become situated, relational and located, affected and connected, where the interest is on what knowledge(s) **data driven** might produce rather than the "truths" it might tell.

Data driven desiring machines that can "engineer partial objects, flows and bodies, and that function of units of production" ([Deleuze and Guattari, 1983, p. 39](#)). Data become a force and flow, a line of flight, machinic in its connections with bodies, in constant motion, becoming, affecting, data-ing. Maybe the question we should be asking about **data driven** is:

What could happen if qualitative researchers were to shift their focus from acting on, manipulating, or analysing data to data's working on the researchers themselves or taking the initiative?

[Koro-Ljungberg \(2016, p. 54\).](#)

Fairchild et al.'s book assembles and produces events that problematize conference spaces highlighting the relationality and materiality between data and bodies. They encourage moving beyond linear progress narratives that become caught up in research practices, instead they call for researchers to:

undo the assumption – common when thinking of what has happened in 'the past' – that the workshop happenings (and the 'data' they produced and created) are lying in limbo waiting 'there' for some human to come back to it/them to resuscitate and make meaning out of it/them. We trouble these linear meaning-making logics.

[Fairchild et al. \(2022, p. 10\).](#)

When we follow lines of flight and become caught up in data-as-a-desiring-machine it becomes possible to re-envisage **data driving**, encouraging research to move beyond "what works" to "what is produced". Perhaps it is time for a quantum leap to move beyond representational logic where "the 'undoings' of methods-as-usual in post-qualitative inquiry promises ... to open a new empirical landscape for ... research investigations" ([Taylor, 2017, p. 322](#)). Maybe the examples in this chapter reveal of some of the possibilities for thinking and researching otherwise when we let **data drive**.

Sample 9: (Im)possible questions shaping data and sampling futures

Can we think about data and sampling without considering global pandemics and ecological crises?
 Can we live without data and sampling and what might that living look like?
 What might replace data?
 How might shifting data assemblage function in ways promoting response-ability?
 What potential implications will a shift from an object thinking (data and sample as objects) toward process (data and sample as processes) and subject (agency of data and sample)?
 How can relational data and sampling help us to move beyond the notion of a/the/one truth in qualitative inquiry?
 How can we live with and by multiplicities in data and sampling, and life?
 How might data and sampling help us in making propositions and understanding potentialities?

References

- Alvesson, M., Ashcraft, K.L., 2012. Interviews. In: Symon, G., Cassell, C. (Eds.), *Qualitative Organizational Research: Core Methods and Current Challenges*. Sage, pp. 239–257.
- Barad, K., 2007. Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press.
- Baugh, B., 2010. Body. In: Parr, A. (Ed.), *The Deleuze Dictionary*, Revised Edition. Edinburgh University Press, pp. 35–37.
- Becker, H.S., 2012. Expert voice. In: Baker, S.E., Edwards, R. (Eds.), *How Many Qualitative Interviews are Enough?* National Centre for Research Methods Review Discussion Paper, p. 15. Retrieved from: <http://eprints.ncrm.ac.uk/2273/>. (Accessed 26 September 2021).
- Bennett, J., 2010. *Vibrant Matter: A Political Ecology of Things*. Duke University Press, Durham.
- Benozzo, A., Gherardi, S., 2020. Working within the shadow: what do we do with “not-yet” data? *Qual. Res. Org. Manag. Int. J.* 15 (2), 145–159. <https://doi.org/10.1108/QROM-09-2018-1684>.
- Benozzo, A., Koro-Ljungberg, M., 2017. [Data within (data)-bag] diffracted. In: Koro-Ljungberg, M., Löytönen, T., Tesar, M. (Eds.), *Disrupting Data in Qualitative Inquiry: Entanglements With the Post-critical and Post-anthropocentric*. Peter Lang, New York, pp. 117–130.
- Benozzo, A., Bell, H., Koro-Ljungberg, M., 2013. Moving between nuisance, secrets, and splinters as data. *Cult. Stud. Crit. Methodol.* 13 (4), 309–315.
- Biesta, G.J.J., 2010. Why “what works” still won’t work: from evidence-based education to value-based education. *Stud. Philos. Educ.* 29, 491–503.
- Bodén, L., 2015. The presence of school absenteeism: exploring methodologies for researching the material-discursive practice of school absence registration. *Cult. Stud. Crit. Methodol.* 15 (3), 192–202.
- Bradbury, A., Roberts-Holmes, G., 2018. *The Datafication of Primary and Early Years Education*. Routledge.
- Braidotti, R., 2013. *The Posthuman*. Polity Press.
- Braun, V., Clarke, V., Gray, D. (Eds.), 2017. *Collecting Qualitative Data*. Cambridge University Press.
- Brinkman, S., 2014. Doing without data. *Qual. Inq.* 20 (6), 720–725. <https://doi.org/10.1177/1077800414530254>.
- Brinkmann, S., Kvale, S., 2015. *Interviews*, third ed. Sage.
- Deleuze, G., 1992. Postscripts on the society of control. October 59, 3–7. <https://www.jstor.org/stable/778828>.
- Deleuze, G., Guattari, F., 1983. *Anti-oedipus*. Trans. R. Hurley, M. Seem and H.R. Lane. Bloomsbury Academic.
- Deleuze, G., Guattari, F., 1987. A thousand plateaus: Capitalism and Schizophrenia. Bloomsbury. In: Massumi, B. (Ed.), *University of Minnesota Press*, Minnesota. Trans. Original work published 1980.
- Denzin, N.K., 1970. *Sociological Methods: A Sourcebook*. Aldine.
- Denzin, N., 2013. The death of data? *Cult. Stud. Crit. Methodol.* 13 (4), 353–356. <https://doi.org/10.1177/1532708613487882>.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 1994. *The SAGE Handbook of Qualitative Research*. Sage.
- Emmel, N., 2014. Sampling and Choosing Cases in Qualitative Research: A Realist Approach. SAGE.
- Ercikan, K., Roth, W.-M., 2006. What good is polarizing research into qualitative and quantitative? *Educ. Res.* 35 (5), 14–23. <https://doi.org/10.3102/0013189X035005014>.
- Fairchild, N., 2017. *Earthworm Disturbances: The Reimagining of Relations in Early Childhood Education and Care*. Doctoral Thesis. University of Chichester (Accredited by the University of Southampton). University of Portsmouth Repository. [https://researchportal.port.ac.uk/portal/en/theses/earthworm-disturbances\(e8a2c3eb-f794-4cf7-b59a-ba9a393b01fa\).html](https://researchportal.port.ac.uk/portal/en/theses/earthworm-disturbances(e8a2c3eb-f794-4cf7-b59a-ba9a393b01fa).html).
- Fairchild, N., 2020. Queering the data: the somatechnics of English early childhood education and care teachers. *Somatechnics* 10 (1), 52–72. <https://doi.org/10.3366/soma.2020.0300>.
- Fairchild, N., Taylor, C.A., Benozzo, A., Carey, N., Koro, M., Elmenhorst, C., 2022. *Knowledge Production in Material Spaces: Disturbing Conferences and Composing Events*. Routledge.
- Fielding, N., Fielding, J., 1986. *Linking Data*. Sage.
- Gentles, S., Charles, C., Ploeg, J., McKibbin, K., 2015. Sampling in qualitative research: insights from an overview of the methods literature. *Qual. Rep.* 20 (11), 1772–1789.
- Glaser, B., Strauss, A., 1967. *The Discovery of Grounded Theory*. Aldine.
- Guba, E., Lincoln, Y., 1994. Competing paradigms in qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, pp. 105–117.
- Gubrium, J., 1988. *Analyzing Field Reality*. Sage.
- Guest, G., Namey, E., Mitchell, M., 2013. Sampling in qualitative research. *Collecting Qualitative Data: A Field Manual for Applied Research*. Sage Publications, pp. 41–94. <https://doi.org/10.4135/97815063746>.
- Gunn, A., 2018. Metrics and methodologies for measuring teaching quality in higher education: developing the teaching excellence framework (TEF). *Educ. Rev.* 70 (2), 129–148. <https://doi.org/10.1080/00131911.2017.1410106>.
- Haiven, M., Khasnabish, A., 2014. *The Radical Imagination*. Fernwood Publishing.
- Haraway, D.J., 2016. *Staying With the Trouble: Making Kin in the Chthulucene*. Duke University Press.
- Heckman, J.J., 2000. Invest in the Very Young. Ounce of Prevention Fund in the University of Chicago Harris School of Public Policy Studies.
- Kippin, S., Cairney, P., 2021. The COVID-19 Exams Fiasco Across the UK: Four Nations and Two Windows of Opportunity. *British Politics*. <https://doi.org/10.1057/s41293-021-00162-y>.
- Knight, L., 2016. Playgrounds as sites of radical encounters: mapping material, affective, spatial, and pedagogical collisions. In: Snaza, N., Sonu, D., Truman, S.E., Zaliwska, Z. (Eds.), *Pedagogical Matters: New Materialisms and Curriculum Studies*. Peter Lang, pp. 13–28.
- Koro-Ljungberg, M., 2016. Reconceptualizing Qualitative Research. *Methodologies Without Methodology*. Sage.
- Koro-Ljungberg, M., MacLure, M., 2013. Provocations, re-un-visions, death and other possibilities of “data”. *Cult. Stud. Crit. Methodol.* 13 (4), 219–222.
- Koro-Ljungberg, M., Löytönen, T., Tesar, M. (Eds.), 2017. *Disrupting Data in Qualitative Inquiry*. Peter Lang.

- Latour, B., 2007. *Reassembling the Social: An Introduction to Actor Network Theory*. Oxford University Press.
- Lecerle, J.-J., 2002. *Deleuze and Language*. Palgrave Macmillan.
- LeCompte, M., Preissle, J., 1993. *Ethnography and Qualitative Design in Educational Research*, second ed. Academic Press.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. SAGE, Thousand Oaks, pp. 289–331.
- Löytönen, T., 2017. Educational development within higher arts education: an experimental move beyond fixed pedagogies. *Int. J. Acad. Dev.* 22 (3), 231–244.
- Lugli, A., 2006. *Arte e Meraviglia. Scritti Sparsi 1974–1995 (Arts and Wonder: Collected Writings 1974–1996)*. Umberto Allemandi & Company.
- MacLure, M., 2013a. The wonder of data. *Cult. Stud. Crit. Methodol.* 13 (4), 228–232. <https://doi.org/10.1177/1532708613487863>.
- MacLure, M., 2013b. Researching without representation? Language and materiality in postqualitative methodology. *Int. J. Qual. Stud. Educ.* 26 (6), 658–667. <https://doi.org/10.1080/09518398.2013.788755>.
- MacLure, M., 2016. The refrain of the a-grammatical child: finding another language in/for qualitative research. *Cult. Stud. Crit. Methodol.* 16 (2), 173–182. <https://doi.org/10.1177/1532708616639333>.
- Manning, E., 2016. *The Minor Gesture*. Duke University Press.
- Manning, E., 2020. *For a Pragmatics of the Useless*. Duke University Press.
- Manning, E., Massumi, B., 2014. *Thought in the Act. Passages in the Ecology of Experience*. University of Minnesota Press.
- Marshall, M.N., 1996. Sampling for qualitative research. *Fam. Pract.* 13 (6), 522–525.
- Molesworth, M., Scullion, R., Nixon, E. (Eds.), 2011. *The Marketisation of Higher Education and the Student as Consumer*. Routledge.
- Morse, J., 1994. Designing funded qualitative research. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook for Qualitative Research*. Sage, pp. 220–235.
- O’Keefe, C., 2017. Economizing education: assessment algorithms and calculative agencies. *E-Learn. Digital Media* 14 (3), 123–137. <https://doi.org/10.1177/2042753017732503>.
- Pink, S., 2014. Digital-visual-sensory design anthropology: ethnography, imagination and intervention. *Arts Humanit. High Educ.* 13, 412–427. <https://doi.org/10.1177/1474022214542353>.
- Pyyhtinen, O., 2016. *More-Than-Human Sociology: A New Sociological Imagination*. Palgrave Macmillan.
- Rantala, T., 2019. *Exploring Data Production in Motion: Fluidity and Feminist Poststructuralism*. Myers Education Press.
- Schwandt, T., 1997. *Qualitative Inquiry: A Dictionary of Terms*. Sage.
- Springgay, S., Truman, S.E., 2016. Stone walks: inhuman animacies and queer archives of feeling. *Discourse*. <https://doi.org/10.1080/01596306.2016.1226777>.
- St Pierre, E.A., 1997. Methodology in the fold and the irruption of transgressive data. *Qual. Stud. Educ.* 10 (20), 175–189.
- Staller, K.M., 2021. Big enough? Sampling in qualitative inquiry. *Qual. Soc. Work* 20 (4), 897–904. <https://doi.org/10.1177/14733250211024516>.
- Stephens, N., Delamont, S., 2006. Balancing the berimbau: embodied ethnographic understanding. *Qual. Inq.* 12, 316–339. <https://doi.org/10.1177/1077800405384370>.
- Styles, B., Torgerson, C., 2018. Randomised controlled trials (RCTs) in education research –methodological debates, questions, challenges. *Educ. Res.* 60 (3), 255–264. <https://doi.org/10.1080/00131881.2018.1500194>.
- Taylor, C.A., 2017. Rethinking the empirical in higher education: postqualitative inquiry as a less comfortable social science. *Int. J. Res. Method Educ.* 40 (3), 311–324. <https://doi.org/10.1080/1743727X.2016.1256984>.
- Taylor, B., Francis, K., 2013. *Qualitative Research in the Health Sciences: Methodologies, Methods and Processes*. Routledge.
- Taylor, C., Koro-Ljungberg, M., Fairchild, N., Benozzo, A., Carey, N., Cozza, M., Elmenhorst, C., 2018. Improvising bags choreographies: disturbing normative ways of doing research. *Qual. Inq.* 25 (1), 17–25.
- Taylor, C.A., Hughes, C., Ulmer, J. (Eds.), 2020. *Transdisciplinary Feminist Research: Innovations in Theory, Method and Practice*. Routledge.
- Vannini, P., Taggart, J., 2012. Doing islandness: a non-representational approach to an island’s sense of place. *Cult. Geogr.* 20, 225–242. <https://doi.org/10.1177/1474474011428098>.
- Weller, S., Romney, K., 1988. *Systematic Data Collection*. Sage.
- Williamson, B., 2016. Digital methodologies of education governance: Pearson plc and the remediation of methods. *Eur. Educ. Res. J.* 15 (1), 34–53. <https://doi.org/10.1177/1474904115612485>.
- Williamson, B., 2018. The hidden architecture of Higher Education: building a big data infrastructure for the “Smarter University”. *Int. J. Educ. Technol. Higher Educ.* 15 (12), 1–26. <https://doi.org/10.1186/s41239-018-0094-1>.
- Wolcott, H., 1994. *Transforming Qualitative Data: Description, Analysis, and Interpretation*. Sage.
- Wolcott, H., 1999. *Ethnography: A Way of Seeing*. AltaMira Press.
- Wolgemuth, J.R., Mam, T.M., Barko, T., Weaver-Hightower, M.B., 2021. Radical uncertainty is not enough: (in)justice matters of post-qualitative research. *Int. Rev. Qual. Res.* <https://doi.org/10.1177/19408447211012658>.

Qualitative interviews

Svend Brinkmann, Aalborg University, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Qualitative interviews in a conversational world	210
The history of qualitative interviews	211
Different forms of qualitative research interviews	212
Structure	212
Individual and group interviews	214
Interviewing using different media	215
Analytic approaches to interviewing	216
Conclusions	217
References	217

Qualitative interviewing has today become a key method in the human and social sciences, and also in many other corners of the scientific landscape such as education and the health sciences. Some have even argued that interviewing has become *the* central resource through which the social sciences—and society—engage with their issues of concern (Rapley, 2001). For as long as we know, human beings have used conversation as a central tool to obtain knowledge about others. People talk with others in order to learn about how they experience the world, how they think, act, feel and develop as individuals and in groups, and in recent decades such knowledge producing conversations have been refined and discussed as *interviews*.

In a classic text, Maccoby and Maccoby defined the interview as “a face-to-face verbal exchange, in which one person, the interviewer, attempts to elicit information or expressions of opinion or belief from another person or persons” (Maccoby and Maccoby, 1954, p. 449). This definition can be used as a very general starting point, but different schools of qualitative interviewing have interpreted, modified, and added to such a generic characterization in many different ways, as we shall see.¹

Qualitative interviews in a conversational world

Human beings are conversational creatures that live a dialogical life (Mulhall, 2007). Humankind is, in the words of philosopher Stephen Mulhall, “a kind of enacted conversation” (p. 58). From the earliest days of our lives, we are able to enter into proto-conversations with caregivers in ways that involve subtle forms of turn-taking and emotional communication. The dyads in which our earliest conversations occur are known to be prior to the child’s own sense of self. Of course, most of us do learn to talk privately to ourselves, and hide our emotional lives from others, but this is possible only because there was first an intersubjective communicative process with others. Our relationships with others—and also with ourselves—are thus conversational. In order to understand ourselves, we must use a language that was first acquired conversationally, and we try out our interpretations in dialog with others and the world. The human self exists only within what philosopher Charles Taylor has called “webs of interlocution” (Taylor, 1989, p. 36). Our very inquiring and interpreting selves are conversational at their core; they are constituted by the numerous relationships we have, and have had, with other people (Brinkmann, 2012).

Unsurprisingly, conversations are therefore a rich and indispensable source of knowledge about personal and social aspects of our lives. Since the late nineteenth century (in journalism) and the early twentieth century (in the social sciences), the conversational process of knowing has been conceptualized under the name of *interviewing*. The very term interview itself testifies to the dialogical and interactional nature of human life. An interview is literally an *inter-view*, an inter-change of views between two persons, conversing about a theme of mutual interest (Brinkmann and Kvale, 2015). *Con-versation* in its Latin root means “dwelling with someone” or “wandering together with,” and the root sense of *dia-logue* is that of talk (*logos*) that goes back and forth (*dia-*) between persons (Mannheim and Tedlock, 1995, p. 4).

Thus conceived, the concept of conversation in the human and social sciences should be thought of as much broader than simply another specific and technical research method. Certainly, conversations in the form of interviewing have been refined into a set of techniques, but they are also a mode of knowing and a fundamental ontology of persons. As philosopher Rom Harré has put it: “The primary human reality is persons in conversation” (Harré, 1983, p. 58). Cultures are constantly produced, reproduced, and revised in dialogs among their members (Mannheim and Tedlock, 1995, p. 2).

The pervasiveness of the conversation in human life is both a burden and a blessing for researchers who wish to employ qualitative interviews. On the one hand, qualitative interviewing appears as a very significant tool with which to understand central features of our conversational world. Contrary to widespread critiques of qualitative research that it is too subjective, one might

¹This article is a condensed and revised version of earlier work, notably Brinkmann (2013).

argue—given the picture of the conversational reality painted here—that qualitative interviewing is in fact the most *objective* method of inquiry when one is interested in qualitative features of human experience, talk, and interaction (at least if objectivity means being adequate to a subject matter), since qualitative interviews seems uniquely capable of grasping these features of our lives.

On the other hand, it is also a burden for qualitative interviewers that they use conversations to study a world, which is always already itself saturated with conversation. If Mulhall is right that humankind is a kind of enacted conversation, then the process of studying humans by the use of interviewing is analogous to fish wanting to study water. Fish surely “know” what water is in a practical, embodied sense, but it can be a great challenge to see and understand the obvious, that with which we are so familiar (Brinkmann, 2012). Similarly, some interview researchers might think that interviewing others for research purposes is easy and simple to do, because it employs a set of techniques that everyone masters in virtue of being capable of asking questions and recording the answers. This, however, is clearly an illusory simplicity, and many qualitative interviewers, even experienced ones, will recognize the frustrating experience of having conducted a large number of interviews (which is often the most fun part of a research project), but having ended up with a huge amount of data, in the form of perhaps thousands of pages of transcripts, and not knowing how to transform all this material into a solid, relevant, and thought-provoking analysis. Too much time is often spent on interviewing, while too little time is devoted to preparing for the interviews, and subsequently analyzing the empirical materials. And, to continue on this note, too little time is normally used to reflect on the role of interviewing as a knowledge-producing social practice in itself. Interviewing, in short, is often simply taken for granted.

The history of qualitative interviews

In one obvious sense, the use of conversations for knowledge producing purposes is likely as old as human language and communication. The fact that we can pose questions to others about things that we are unknowledgeable of is a core capability of the human species. It expands our intellectual powers enormously, since it enables us to share and distribute knowledge between us. Without this fundamental capability, it would be hard to imagine what human life would be like. It is furthermore a capacity that has developed into many different forms and ramifications in human societies. Already in 1924 could Emory Bogardus, an early American sociologist and founder of one of the first US sociology departments (at the University of Southern California) declare that interviewing “is as old as the human race” (Bogardus, 1924, p. 456), and he discussed similarities and differences between the ways that physicians, lawyers, priests, journalists, detectives, social workers, and psychiatrists conduct interviews with a remarkable sensitivity to the details of such different conversational practices.

In a more specific sense, and more essentially related to qualitative interviewing as a scientific human enterprise, conversations were used by Thucydides in ancient Greece as he interviewed participants from the Peloponnesian Wars to write the history of the wars. At roughly the same time, Socrates famously questioned—or we might even say *interviewed*—his fellow citizens in ancient Athens, and used the dialogs to develop knowledge about perennial human questions related to justice, truth, beauty, and the virtues. In recent years, some interview scholars have sought to rehabilitate a Socratic practice of interviewing, not least as an alternative to the often long monologues of phenomenological and narrative approaches to interviewing (see Dinkins, 2005, who unites Socrates with a hermeneutic approach to dialogical knowledge), and also in an attempt to think of interviews as practices that can create a knowledgeable citizenry and not merely chart common opinions and attitudes (Brinkmann, 2007). Such varieties of interviewing have come to be known as dialogic and confrontational (Roulston, 2010, p. 26), although care should be taken not to understand the latter term as implying some disrespectful practice of catching the interviewee in contradictions or something like that.

If we jump to more recent times, interviewing notably entered the human sciences with the advent of Sigmund Freud’s psychoanalysis around 1900. Freud is famous for his psychoanalytic theory of the unconscious, but it is significant that he developed this revolutionary theory (which, in many ways, changed the Western conceptions of humanity) through therapeutic conversations, or what he referred to as the talking-cure. Freud conducted several hundreds of interviews with patients that used the patients’ free associations as a conversational engine. The therapist/interviewer should display what Freud called an “even-hovering attention” and catch on to anything that emerged as important (Freud, 1963). Freud made clear that research and treatment go hand in hand in psychoanalysis, and Steinar Kvale has more recently pointed to the rich potentials of psychoanalytic conversations for qualitative interviewing even today (Kvale, 2003).

Many human and social scientists from the first half of the twentieth century were well versed in psychoanalytic theory, including those who pioneered in qualitative interviewing. Jean Piaget, the famous developmental researcher, had even received training as a psychoanalyst himself, but his approach to interviewing is also worth mentioning in its own right. Piaget’s (1930) theory of child development was based on his interviews with children (often his own) in natural settings, frequently in combination with different experimental tasks. He would typically let the children talk freely about the weight and size of objects, or, in relation to his research on moral development, about different moral problems (Piaget, 1932), and he would notice the manner in which their thoughts unfolded.

Raymond Lee, one of the few historians of interviewing, has charted how Piaget’s so-called “clinical” method of interviewing became an inspiration for Elton Mayo, who was responsible for one of the largest interview studies in history at the Hawthorne plant in Chicago in the 1920s (Lee, 2011). This study arose from a need to interpret the curious results of a number of practical experiments on the effects of changes in illumination on production at the plant: It seemed that work output improved when the lighting of the production rooms was increased, but also when it was decreased. This instigated an interview study with

more than 21,000 workers being interviewed for more than an hour each, leading to the whole human relations approach to work life (Roethlisberger and Dickson, 1939).

Not just Piaget, but also the humanistic psychologist Carl Rogers, was an important pioneer of interviewing. Like Freud, Rogers developed a conversational technique that was useful both in therapeutic contexts (so-called client-centered therapy), but also in research interviews, which he referred to as the “non-directive method as a technique for social research” (Rogers, 1945). As he explained, the goal of this kind of therapy/research was to sample the respondent’s attitudes toward herself: “Through the non-directive interview we have an unbiased method by which we may plumb these private thoughts and perceptions of the individual” (p. 282). In contrast to psychoanalytic practice, the respondent in client-centered therapy/research is a client rather than a patient, and the client is the expert. Although often framed in different terms, many contemporary interview researchers still conceptualize the research interview in line with Rogers’ humanistic, non-directive approach, valorizing the respondents’ private experiences, narratives, opinions, beliefs, and attitudes.

As the Rogerian nondirective approach to interviewing gained in popularity, early critiques of this technique also emerged. In the 1950s, the famous sociologist David Riesman and his colleague Mark Benney criticized it for its lack of interviewer involvement and they warned against the tendency to use the level of “rapport” (much emphasized by interviewers inspired by therapy) in an interview to judge its qualities concerning knowledge. They thought it was a prejudice “to assume the more rapport-filled and intimate the relation, the more ‘truth’ the respondent will vouchsafe” (Riesman and Benney, 1956, p. 10). In their eyes, rapport-filled interviews often spill over with “the flow of legend and cliché” (p. 11), since interviewees are likely to adapt their responses to what they take the interviewer expects from them (see also Lee, 2008, for an account of Riesman’s discussion of interviewing and its relevance even today).

The mid-20th century witnessed a number of other large interview studies that remain classics in the field, but which have also shaped public opinion about different issues such as *The Authoritarian Personality* by the well-known critical theorist Adorno and co-workers (Adorno et al., 1950), Kinsey’s *Sexual Behavior in the Human Male* (Kinsey et al., 1948), and Dichter’s *The Strategy of Desire* (1960) that communicates the results of an interview study about consumer motivation for buying a car. The latter was an important precursor to the market and consumer research that continues to be among the largest areas of qualitative interviewing in contemporary society.

Roulston (2010) has made a comprehensive list of some of the most recent varieties of interviewing and also of more traditional ones (here shortened a bit):

- *Neo-positivist* conceptions of the interview are still widespread and emphasize how the conversation can be used to reveal “the true self” of the interviewee (or the essence of her experiences), resulting in solid, trustworthy data that are only accessible through interviews with the interviewer having a non-interfering role.
- *Romantic* conceptions stress that the goal of interviewing is to obtain revelations and confessions of the interviewees by intimacy and rapport. These conceptions are somewhat close to neo-positivist ones, but put much more weight on the interviewer as an active and authentic midwife facilitating the “birth” of the inner psyche of the interviewee.
- *Constructionist* conceptions reject the romantic idea of authenticity and favor an idea of a subject that is locally produced in the situation. There is thus a focus on the situational practice of interviewing, and a distrust toward discourses of data as permanent “nuggets” to be “mined” by the interviewer. Instead, the interviewer is often portrayed as a “traveler”, involved in the co-construction of whatever happens in the conversation together with the interviewee (Brinkmann and Kvale, 2015).
- *Postmodern* and *transformative* conceptions stage interviews as dialogic and performative events that aim to bring new kinds of people and new worlds into being. The interview is depicted as a chance for people to get together and create new possibilities for action. Some transformative conceptions focus on potential decolonizing aspects of interviewing, seeking to subvert the colonizing tendencies that some see in standard interviewing, while other forms focus more on the researchers’ experience, sometimes expressed through autoethnography, an approach that seeks to unite ethnographical and autobiographical intentions (Ellis et al., 2011).

Different forms of qualitative research interviews

The semi-structured, face-to-face interview represents merely one form an interview may take, and there is today a huge variety of other forms. Each form has certain advantages and disadvantages that researchers and recipients of research alike should be aware of. I shall here describe how qualitative interviews may differ in terms of *structure*, the *number of participants* in each interview, and also different *media*. After this I will go on to look at two broad ways of theorizing interview talk.

Structure

It is quite common to make a distinction between structured, semi-structured, and unstructured interviews. This distinction, however, should be thought of as a continuum ranging from relatively structured to relatively unstructured formats. I use the word “relatively,” because, on the one end of the continuum, as Parker (2005) argues, there really is no such thing as a completely structured interview “because people always say things that spill beyond the structure, before the interview starts and when the recorder has been turned off” (p. 53). Utterances that “spill beyond the structure” are often quite important, and are even sometimes

the key to understanding the interviewee's answers to the pre-structured questions. One line of criticism against standardized survey interviewing actually concerns the fact that meanings and interpretive frames that go beyond the predetermined structure are left out with the risk that the researcher cannot understand what actually goes on in the interaction.

We might add to Parker's argument that there is also no such thing as a completely *unstructured* interview, since the interviewer always has an idea about what should take place in the conversation. Even some of the least structured interviews such as life history interviews that only have one question prepared in advance (e.g., "I would like you to tell me the story of your life. Please begin as far back as you can remember and include as many details as possible") provide structure to the conversation by framing it in accordance with certain specific conversational norms rather than others. Another way to put the matter is to say that there are no such things as non-leading questions. All questions lead the interviewee in certain directions, but it is generally preferable to lead participants only to talk about certain *themes*, rather than to specific *opinions* about these themes.

So, it is not possible to avoid structure, nor would it be desirable, but it is possible to provide a structure that is flexible enough for interviewees to be able to raise questions and concerns in their own words and from their own perspectives. This does not imply a total elimination of structure, but demands a careful preparation and reflection of how to involve interviewees actively, how to avoid flooding the conversation with social science categories, and how to provoke interviewees respectfully to bring contrasting perspectives to light (Parker, 2005, p. 63). In spite of this caveat—that neither completely structured nor completely unstructured interviews are possible—it may still be worthwhile to distinguish between more or less structure with semi-structured interviews somewhere in the middle as the standard approach to qualitative interviewing.

- *Structured interviews*: These are employed in surveys and are typically based on the same research logic as questionnaires: Standardized ways of asking questions are thought to lead to answers that can be compared across participants and possibly quantified. Interviewers are supposed to "read questions exactly as worded to every respondent and are trained never to provide information beyond what is scripted in the questionnaire" (Conrad and Schober, 2008, p. 173). Although structured interviews are useful for some purposes, they do not take advantage of the dialogical potentials for knowledge production that are inherent in human conversations. They are passive recordings of people's opinions and attitudes, and often reveal more about the cultural conventions of how one should answer specific questions than about the conversational production of social life itself. Since this article is concerned with qualitative interviewing, which characteristically has as a much more flexible format with the interviewer being able to follow up freely on interviewee answers, I shall leave the standardized and structured interview format behind and refer the interested reader to Conrad and Schober, who do in fact discuss creative ways of developing this traditional social science technique further (Conrad and Schober, 2008).
- *Unstructured interviews*: At the other end of the continuum lie interviews that have little preset structure. These are for example the life story interview seeking to highlight "the most important influences, experiences, circumstances, issues, themes, and lessons of a lifetime" (Atkinson, 2002, p. 125). What these are for an individual can only be known in the course of spending time with the interviewee, which means that the interviewer cannot prepare for a life story interview by devising a lot of specific questions, but must instead think about how to facilitate the telling of the life story. After the opening request for a narrative, the main role of the interviewer is to remain a listener, withholding desires to interrupt, and sporadically asking questions that may clarify the story. The life story interview is a variant of the more general genre of narrative interviewing about which Wengraf's (2001) *Qualitative Research Interviewing* gives a particularly thorough account, focusing on biographical-narrative depth-interviews. These need not concern the life story as a whole, but may address other, more specific, storied aspects of human lives, building on the insight of narratology that human beings experience and act in the world through narratives. With the more focused narrative interviews we get nearer to semi-structured interviews as the middle ground between structured and unstructured interviews.
- *Semi-structured interviews*: Interviews in the semi-structured format are sometimes equated with qualitative interviewing as such (Warren, 2002). They are probably also the most widespread ones in the human and social sciences and are sometimes the only format given attention in textbooks on qualitative research. Compared to structured interviews, semi-structured interviews can make better use of the knowledge producing potentials of dialogs by allowing much more leeway for following up on whatever angles are deemed important by the interviewee, and the interviewer has a greater chance of becoming visible as a knowledge producing participant in the process itself, rather than hiding behind a preset interview guide. And, compared to unstructured interviews, the interviewer has a greater saying in focusing the conversation on issues that he or she deems important in relation to the research project.

One definition of the semi-structured qualitative research interview reads: "It is defined as an interview with the purpose of obtaining descriptions of the life world of the interviewee in order to interpret the meaning of the described phenomena" (Brinkmann and Kvale, 2015, p. 6). The key words here are (1) purpose, (2) descriptions, (3) life world, and (4) interpretation of meaning, which I shall now unfold in turn:

- (1) *Purpose*: Unlike everyday conversations with friends or family members, qualitative interviews are not conducted for their own sake; they are not a goal in themselves, but are staged and conducted in order to serve the researcher's goal of producing knowledge (and there may be other, ulterior goals like obtaining a degree, furthering one's career, or positioning oneself in the field etc.). All sorts of motives may play a role in the staging of interviews, and good interview reports often contain a *reflexive* account and discussion of both individual and social aspects of such motives (does it matter, for example, if the interviewer is a woman, perhaps identifying as a feminist, interviewing other women?). Clearly, the fact that interviews are conversations

conducted for a purpose, which sets the agenda, raises a number of issues having to do with power and control that are important to reflect upon for epistemic reasons as well as for ethical ones.

- (2) *Descriptions*: In most interview studies, the goal is to obtain the interviewee's descriptions rather than reflections or theorizations. In line with a widespread phenomenological perspective, interviewers are normally seeking descriptions of *how* interviewees experience the world, its episodes and events, rather than speculations about *why* they have certain experiences. Good interview questions thus invite interviewees to give descriptions such as "Could you please describe a situation for me in which you became angry?", "What happened?", "How did you experience anger?", "How did it feel?" (of course, only one of these questions should be posed at a time), and good interviewers tend to avoid more abstract and reflective questions such as "What does anger mean to you?", "If I say 'anger', what do you think of then?", "Why do you think that you tend to feel angry?". Such questions *may* be productive in the conversation, but interviewers will normally defer them until more descriptive aspects have been covered.
- (3) *Life world*: The concept of the life world goes back to the founder of phenomenology, Edmund Husserl, who introduced it in 1936 in his book *The Crisis of the European Sciences* to refer to the intersubjectively shared and meaningful world in which humans conduct their lives and experience significant phenomena (Husserl, 1954). It is a pre-reflective and pre-theorized world in which anger, for example, is a meaningful human expression in response to having one's rights violated (or something similar) before it is a process occurring in the neurophysiological and endocrinological systems (and "before" should here be taken in a logical, rather than temporal, sense). If anger did not appear to human beings as a meaningful *experienced phenomenon* in their life world, there would be no reason to investigate it scientifically, for there would in a sense be nothing to investigate. In qualitative research in general, as in qualitative interviewing in particular, there is a primacy of the life world as experienced, since this is prior to the scientific theories we may formulate about it. This was well expressed by Maurice Merleau-Ponty, another famous phenomenologist, who built on the work of Husserl:

All my knowledge of the world, even my scientific knowledge, is gained from my own particular point of view, or from some experience of the world without which the symbols of science would be meaningless. The whole universe of science is built upon the world as directly experienced [i.e., the life world; my addition], and if we want to subject science itself to rigorous scrutiny and arrive at a precise assessment of its meaning and scope, we must begin by re-awakening the basic experiences of the world of which science is the second order expression.

Merleau-Ponty (1945, p. ix).

Objectifying sciences give us second-order understandings of the world, but qualitative research is meant to provide a first-order understanding through concrete description, e.g., as obtained by interviewing.

- (4) *Interpret the meaning*: Even if interviewers are generally interested in how people experience and act in the world prior to abstract theorizations, they must nonetheless often engage in interpretations of people's experiences and actions as described in interviews. One reason for this is that life world phenomena are rarely transparent and "monovocal," but are rather "polyvocal" and sometimes even contradictory, permitting multiple readings and interpretations. Who is to say what someone's description of anger signifies? Obviously, the person having experienced the anger should be listened to, but if there is one lesson to learn from twentieth century human science (ranging from psychoanalysis to poststructuralism) it is that we, as human subjects, do not have full authority concerning how to understand our lives, because we do not have—and can never have—full insight into the forces that have created us (Butler, 2005). We are, as Judith Butler has argued, authored by what precedes and exceeds us (p. 82), even when we are considered—as in qualitative interviews—to be authors of our own utterances. The interpretation of the meanings of the phenomena described by the interviewee can favorably be built into the conversation itself, since this will at least give the interviewee a chance to object to a certain interpretation, but it is a process that goes on throughout an interview project.

I have now introduced a working definition of the semi-structured qualitative research interview and emphasized four vital aspects: Such interviews are structured by the interviewer's *purpose* of obtaining knowledge; they revolve around *descriptions* provided by the interviewee; such descriptions are commonly about *life world phenomena* as experienced; and understanding the meaning of the descriptions involves some kind of *interpretation*. Although these aspects capture what is essential to a large number of qualitative interview studies now and in the past (and likely many of the future as well), all aspects can be challenged, and there are no principles of qualitative interviewing that are always proper and correct.

Individual and group interviews

It is not just the interviewer's agenda and research interests that structure the interaction in an interview. Unsurprisingly, the number of interviewees also plays an important role. As the history of interviewing testifies, the standard format of qualitative interviewing is with one person interviewing another person, but many other varieties exist that deserve to be mentioned.

- *Group interviews*: There is today an increasing use of group interviews. These have been in use since the 1920s, but became standard practice only after the 1950s, when market researchers in particular developed what they termed focus group interviews

to study consumer preferences. Today, focus groups dominate consumer research and are also often used in health, education, and evaluation research, but are in fact becoming increasingly common across many disciplines in the social sciences.

In focus groups, the interviewer is conceived as a “moderator,” who focuses the group discussion on specific themes of interest, and she or he will often use the group dynamic instrumentally to include a number of different perspectives on the given themes (Morgan, 2002). Often, group interviews are more dynamic and flexible in comparison with individual interviews, and they may be closer to everyday discussions. They can be used, for example, when the researcher is not so much interested in people’s descriptions of their experiences as in how participants discuss, argue, and justify their opinions and attitudes.

The standard size for a focus group is between six to ten participants, led by a moderator. Recently, qualitative researchers have also experimented with groups of only two participants (sometimes referred to as “the two-person interview,” although there are literally three people if one counts the interviewer), mainly because it makes the research process easier to handle than with larger groups, where people will often not show up. The moderator introduces the topics for discussion and facilitates the interchange. The point is not to reach consensus about the issues discussed, but to have different viewpoints articulated about an issue. Focus group interviews are well suited for exploratory studies in little known domains, or about newly emerging social phenomena, since the dynamic social interaction that results may provide more spontaneous expressions than in individual interviews.

- *Individual interviews:* Individual interviews with one interviewer and one interviewee may sometimes be less lively than group interviews, but they have a couple of other advantages: First, it is often easier for the interviewer in one-on-one interviews to lead the conversation in a direction that is useful in relation to the interviewer’s research interests. Second, when studying aspects of people’s lives that are personal, sensitive, or even taboo, it is preferable to use individual interviews that allow for more confidentiality and often make it easier for the interviewer to create an atmosphere of trust and discretion. And there are obviously also certain themes that simply just demand the presence of one person telling a story without being interrupted or gainsaid by other participants, such as in biographical research.

Interviewing using different media

It is noteworthy that the otherwise standardized format of “face-to-face interaction” was only named early in the twentieth century by the sociologist Charles Horton Cooley (see Briggs, 2007, p. 553), but was since constructed as “primordial, authentic, quintessentially human, and necessary” (p. 553). It is sometimes forgotten that also the face-to-face interview is a kind of situated interaction that is mediated by this particular social arrangement that has a history. Other well-known media employed in qualitative interviewing include the telephone and the internet, and we shall now take a brief look at differences between face-to-face, telephone, and internet interviews.

- *Face-to-face interviews:* In face-to-face interviews people are present not only as conversing minds, but as flesh and blood creatures that may laugh, cry, smile, tremble, and otherwise give away so much information in terms of gestures, body language, and facial expressions. Interviewers thus have the richest source of knowledge available here, but the challenge concerns how to use it productively. In most cases, how people look and act is forgotten once the transcript is made, and the researcher carries out her analyses using the stack of transcripts rather than the embodied interaction that took place. This is a problem especially when a research assistant, or someone else who was not the interviewer, transcribes the interview, because in that case, it is not possible to note and recall all the non-verbal signs and gestures that occurred. If possible, it is therefore preferable for the interviewer herself to transcribe the conversations, and it is optimal to do so relatively soon after the conversations are over (e.g., within a couple of days), since this guarantees better recollection of the body language, the atmosphere, and other such non-transcribable features of the interaction.
- *Telephone interviews:* According to Shuy (2002), the telephone interview has “swept the polling and survey industry in recent years and is now the dominant approach (p. 539)—often in a very structured format. In a research context, the use of telephone conversations was pioneered by conversation analysts, who were able to identify a number of common conversational mechanisms (related to turn-taking, adjacency pairs such as questions-answers etc.) from the rather constricted format that is possible over the telephone. The constricted format itself may have been productive in throwing light on certain core features of human talk.

Shuy emphasizes a number of advantages of telephone interviewing such as reduced interviewer effects (important in structured polling interviews, for example), better interviewer uniformity, greater standardization of questions, greater cost-efficiency, increased researcher safety (Shuy, 2002, p. 540), and—we might add—better opportunities for interviewing people who live far away from the interviewer. In qualitative interviewing, it is not possible (nor desirable) to avoid “interviewer effects,” because the interviewer herself is the research instrument, so only the latter couple of points are relevant in this context. However, Shuy also highlights some advantages of in-person interviewing as against telephone interviewing such as more accurate responses due to contextual naturalness, greater likelihood of self-generated answers, more symmetrical distribution of interactive power, greater effectiveness with complex issues, more thoughtful responses, and the fact that they are better in relation to sensitive questions (pp. 541–544). The large majority of interviews that are characterized as “qualitative” are conducted face-to-face, mainly because of these advantages listed by Shuy.

- *Internet interviews*: Email and chat interviews are varieties of internet interviewing with email interviewing normally implying an asynchronous interaction in time, with the interviewer writing a question and then waiting for a response, and chat interviews being synchronous or “real time” (Mann and Stewart, 2002). The latter can approach a conversational format that resembles face-to-face interviews with quick turn takings. When doing online ethnographies, e.g., in virtual realities on the internet, chat interviews are important (see Markham, 2005, on online ethnography). One advantage of email and chat interviews is that they are “self-transcribing” in the sense that the written text itself is the medium through which researcher and respondents express themselves, and the text is thus basically ready for analysis the minute it has been typed.

Disadvantages of such interview forms are related to the demanded skills of written communication. Not everyone is sufficiently skilled at writing to be able to express themselves in rich and detailed ways. Most research participants are more comfortable when talking, rather than writing, about their lives and experiences. However, as the psychiatrist Finn Skårderud has pointed out, there are some exceptions here, and Skårderud emphasizes in particular that internet conversations can be useful when communicating with people who have problematic relationships to their bodies (e.g., eating disorders). For such people, the physical presence of a problematic body can represent an unwanted disturbance (Skårderud, 2003).

In concluding on the different media of interviewing it should be emphasized that all interviews are mediated, even if only by the spoken words and the historical arrangement of questioning through face-to-face interaction, and there is no universally correct medium that will always guarantee success. Interviewers should choose their medium according to their knowledge interests, and should minimally reflect on the effects of communicating through one medium rather than another. That said, most of the themes that qualitative interviewers are interested in lend themselves more easily to face-to-face interviewing because of the trust, confidentiality, and contextual richness that this format enables.

Analytic approaches to interviewing

Before closing, I shall give a brief introduction to different analytic approaches to interviewing ordered by a simplified dichotomy, which should really be thought of as a continuum. I will introduce a distinction between interview talk as primarily descriptive phenomenological *reports* (concentrating on the “what” of communication) and interview talk as primarily discursive *accounts* (chiefly concerned with the “how” of talk). Phenomenological approaches to interviewing try to get as close as possible to precise descriptions of *what* people have experienced, while other analytical approaches (found, for example, in certain schools of discourse analysis and conversation analysis) focus on *how* people express themselves and give accounts occasioned by the situation in which they find themselves. The two approaches are contrasted in Table 1 below.

- *Interviews as research instruments*: Researchers working from the former perspective (corresponding to the left hand side of the Table) believe that interview data can reflect the interviewees’ reality outside the interview. They consequently seek to minimize the interviewer’s effects on coloring interviewees’ reports of their reality. The interview becomes a research instrument in the hands of interviewers, who are supposed to act receptively in order to interfere as little as possible with the interviewee reporting. The validity of the interviewees’ reports becomes a prime issue, when one approaches interviewing primarily as a research instrument. And since interviews normally concern things experienced in the past, this significantly involves considerations about human memory, and about how to enhance the trustworthiness of human recollections.

In one of the few publications to discuss the role of memory in interviewing, Thomsen and Brinkmann (2009) recommend that interviewers take the following points into account if they want to help interviewees’ improve the reporting of specific memories:

- Allow time for recall and assure the interviewee that this is normal
- Provide concrete cues, e.g., “the last time you were talking to a physician/nurse” rather than “a communication experience”
- Use typical content categories of specific memories to derive cues (i.e., ongoing activity, location, persons, other people’s affect and own affect)
- Ask for recent specific memories
- Use relevant extended time line and landmark events as contextual cues, i.e., “when you were working at x” to aid the recall of older memories

Table 1 Inspiration for slicing the cake of qualitative interviewing in this manner comes from Talmy (2010) and Rapley (2001), who builds on a distinction from Seale between interview-data-as-resource and interview-data-as-topic.

<i>Conception of interviewing</i>	<i>Research instrument</i>	<i>Social practice</i>
Conception of interview data	Reports, interview data as resource	Accounts, interview data as topics
Standard analytic focus	Lived experience—the “what”	Situated interaction—the “how”
Typical interviewer style	Receptive	Assertive
Main challenge	Validity of interviewee reports	Relevance of interviewee accounts
Paradigmatic background	Phenomenology, grounded theory etc.	Discourse analysis, conversation analysis etc.

- Ask the interviewee for a free and detailed narrative of the specific memory (adapted from [Thomsen and Brinkmann, 2009](#)).

Following such guidelines is meant to result in interviewee descriptions that are valid (which means that they are about what the researcher intends them to be about), and close to the “lived experience” of something, or what was called experiences of life world phenomena earlier in the article. Although phenomenology is one typical paradigm to frame interviews analytically as research instruments, many other paradigms do so as well, for example Grounded Theory, which has been developed since [Glaser and Strauss \(1967\)](#) with the intent of developing theoretical understandings of phenomena grounded in empirical materials with meticulous coding of data being a significant technique.

- *Interviewing as a social practice*: In contrast to those approaches that see interviewing as a research instrument designed to capture the “what” of what is reported as accurately as possible, others working from more constructionist and situated perspectives have much greater analytic focus on the “how” of interviewing. They view interviewing as a social practice; as a site for a specific kind of situated interaction, which means that interview data primarily reflect “a reality constructed by the interviewee and interviewer” ([Rapley, 2001](#), p. 304). The idea of obtaining valid reports that accurately reflect a reality outside the conversational situation is thus questioned, and the main challenge becomes instead how to explain the relevance of interview talk. That is, if what is said in an interview is a product of this social practice itself, how can it ever be relevant to conduct interviewing? Postmodern interviews, emphasizing performative and transformative aspects of interviewing, represent attempts to meet this challenge by arguing that if interviews do not concern a reality outside themselves, they can instead be used to perform or facilitate social change.

People subscribing to the right hand side of [Table 1](#) believe that interview talk should be conceived as accounts. Unlike reports, which refer to experiences in the interviewee’s past that can be articulated when prompted, accounts are answers that are “normatively oriented to and designed for the questions that occasion them” ([Talmy, 2010](#), p. 136). If interviewee talk is best understood as accounts, it must be seen as a kind of social action that has effects and *does* something in the situation of which it is a part. This perspective on interviewing is shared by some discourse analysts and conversation analysts, who limit themselves to analyzing interview talk as situated interaction.

Readers may wonder at this point if two approaches that are so different can co-exist. My own pragmatic answer is that they can, at least if the two approaches are not brought to an extreme: It is true that there are huge problems associated with viewing the interview as a site for pure reports of the past (we know too much about the constructive role of human memory, and of how the social practice of interviewing mediates what is said, to take this seriously). But it is certainly also true that there are problems associated with denying that we can refer more or less accurately to past experiences (and those who follow the right-hand side of [Table 1](#), and deny that data can be resources for understanding experiences of the past, still believe that their own communicative practice, materialized in their texts, are about matters outside this specific text).

So, taken to extremes, both approaches become absurd, and I believe that it is now time for the two (sometimes opposed) camps to learn from one another and realize that they need not exclude one another. In my view, some of the most interesting interview studies are those where analyses of the “what” and the “how” fertilize each other in productive ways.

Conclusions

In this article, I have given a broad introduction to qualitative interviewing. I have tried to demonstrate that the human world is a conversational reality in which interviewing takes a privileged position as a research method, at least in relation to a number of significant research questions that human and social scientists want to ask. Throughout the article, I have kept an eye to interviewing as a social practice, and we should be careful not to naturalize this kind of human interaction unreflectively, i.e., by viewing it as a particularly natural and unproblematic way of staging human relationships.

Furthermore, I introduced a number of distinctions that are relevant when mapping the field of qualitative interviewing, e.g., between different levels of structure, different numbers of participants, and different media of interviewing. I also provided a detailed presentation of semi-structured life world interviewing, as the standard form of qualitative interviewing today.

I finally gave particular attention to two broad analytic approaches to interviewing: On the one side, experience focused interviewing seeking to elicit accurate reports of what interviewees have experienced (in broad terms the phenomenological positions), and, on the other side, language and interaction focused interviewing (discourse oriented positions) that focus on the nature of interview interaction in its own right.

In my eyes, none of these approaches are superior *per se*, but they enable researchers to pose different kinds of questions to their materials. Too often, however, interviewers forget to make clear what kinds of questions they are interested in, and also forget to consider whether their practice of interviewing and their analytic focus enable them to answer their research questions satisfactorily.

References

- Adorno, T.W., Frenkel-Brunswick, E., Levinson, D.J., Sanford, R.N., 1950. *The Authoritarian Personality*. Norton, New York.
- Atkinson, P., 2002. The life story interview. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research: Context & Method*. Sage, Thousand Oaks, CA.

- Bogardus, E.M., 1924. Methods of interviewing. *J. Appl. Sociol.* 9, 456–467.
- Briggs, C., 2007. Anthropology, interviewing and communicability in contemporary society. *Curr. Anthropol.* 48, 551–567.
- Brinkmann, S., Kvale, S., 2015. *InterViews: Learning the Craft of Qualitative Research Interviewing*, third ed. Sage, Thousand Oaks, CA.
- Brinkmann, S., 2007. Could interviews be epistemic? An alternative to qualitative opinion-polling. *Qual. Inq.* 13, 1116–1138.
- Brinkmann, S., 2012. *Qualitative Inquiry in Everyday Life: Working With Everyday Life Materials*. Sage, London.
- Brinkmann, S., 2013. *Qualitative Interviewing*. Oxford University Press, Oxford.
- Butler, J., 2005. *Giving an Account of Oneself*. Fordham University Press, New York.
- Conrad, R.G., Schober, M., 2008. New frontiers in standardized survey interviewing. In: Hesse-Biber, S.N., Leavy, P. (Eds.), *Handbook of Emergent Methods*. The Guilford Press, London.
- Dichter, E., 1960. *The Strategy of Desire*. Doubleday, Garden City, NY.
- Dinkins, C.S., 2005. Shared inquiry: socratic-hermeneutic inter-viewing. In: Ironside, P. (Ed.), *Beyond Method: Philosophical Conversations in Healthcare Research and Scholarship*. University of Wisconsin Press, Madison, WI.
- Ellis, C., Adams, T.E., Bochner, A.P., 2011. Autoethnography: an overview. *Forum Qual. Soc. Res.* 12, 10. <http://www.qualitative-research.net/index.php/fqs/article/viewArticle/1589/3095>.
- Freud, S., 1963. *Therapy and Technique*. Collier, New York.
- Glaser, B.G., Strauss, A., 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Publishing Company, New York.
- Harré, R., 1983. *Personal Being*. Basil Blackwell, Oxford.
- Husserl, E., 1954. *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. Martinus Nijhoff, Haag.
- Kinsey, A.C., Pomeroy, W.B., Martin, C.E., 1948. *Sexual Behavior in the Human Male*. Saunders, Philadelphia.
- Kvale, S., 2003. The psychoanalytical interview as inspiration for qualitative research. In: Camic, P.M., Rhodes, J.E. (Eds.), *Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design*. American Psychological Association, Washington, DC, US.
- Lee, R.M., 2008. David Riesman and the sociology of the interview. *Socio. Q.* 49, 285–307.
- Lee, R.M., 2011. "The most important technique...": Carl Rogers, Hawthorne, and the rise and fall of nondirective interviewing in sociology. *J. Hist. Behav. Sci.* 47, 123–146.
- Maccoby, E.E., Maccoby, N., 1954. The interview: a tool of social science. In: Lindzey, G. (Ed.), *Handbook of Social Psychology*. Addison-Wesley, Cambridge, MA.
- Mann, C., Stewart, F., 2002. Internet interviewing. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research: Context & Method*. Sage, Thousand Oaks, CA.
- Mannheim, B., Tedlock, B., 1995. Introduction. In: Tedlock, B., Mannheim, B. (Eds.), *The Dialogic Emergence of Culture*. University of Illinois Press, Urbana, IL.
- Markham, A., 2005. The methods, politics, and ethics of representation in online ethnography. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, third ed. Sage, Thousand Oaks, CA.
- Merleau-Ponty, M., 1945. *Phenomenology of Perception*. This edition published 2002. Routledge, London.
- Morgan, D.L., 2002. Focus group interviewing. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research: Context & Method*. Sage, Thousand Oaks, CA.
- Mulhall, S., 2007. *The Conversation of Humanity*. University of Virginia Press, Charlottesville.
- Parker, I., 2005. *Qualitative Psychology: Introducing Radical Research*. Open University Press, Buckingham.
- Piaget, J., 1930. *The Child's Conception of the World*. Harcourt, Brace & World, New York.
- Piaget, J., 1932. *The Moral Judgment of the Child*. This edition published 1975. Routledge & Kegan Paul, London.
- Rapley, T.J., 2001. The art(fulness) of open-ended interviewing: some considerations on analyzing interviews. *Qual. Res.* 1, 303–323.
- Riesman, D.M., Benney, M., 1956. The sociology of the interview. *Midwest. Sociol.* 18, 3–15.
- Roethlisberger, F.J., Dickson, W.J., 1939. *Management and the Worker*. Wiley, New York.
- Rogers, C., 1945. The non-directive method as a technique for social research. *Am. J. Sociol.* 50, 279–283.
- Roulston, K., 2010. *Reflective Interviewing: A Guide to Theory and Practice*. Sage, Thousand Oaks, CA.
- Shuy, R.W., 2002. In-person versus telephone interviewing. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research: Context & Method*. Sage, Thousand Oaks, CA.
- Skårderud, F., 2003. Sh@me in cyberspace. Relationships without faces: the e-media and eating disorders. *Eur. Eat Disord. Rev.* 11, 155–169.
- Talmy, S., 2010. Qualitative interviews in applied linguistics: from research instrument to social practice. *Annu. Rev. Appl. Ling.* 30, 128–148.
- Taylor, C., 1989. *Sources of the Self*. Cambridge University Press, Cambridge.
- Thomsen, D.K., Brinkmann, S., 2009. An interviewer's guide to autobiographical memory: ways to elicit concrete experiences and to avoid pitfalls in interpreting them. *Qual. Res. Psychol.* 6, 294–312.
- Warren, C.A.B., 2002. Qualitative interviewing. In: Gubrium, J.F., Holstein, J.A. (Eds.), *Handbook of Interview Research: Context & Method*. Sage, Thousand Oaks, CA.
- Wengraf, T., 2001. *Qualitative Research Interviewing*. Sage, Thousand Oaks, CA.

Imagining and reimagining focus group research in education

Alyson Welker^a and George Kamberelis^b, ^a Department of English, Colorado State University, Fort Collins, CO, United States; and
^b Education Department, Western Colorado University, Gunnison, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction: our perspectives and biases	219
Retrospect: traditional uses of focus groups in social science research	220
Multiple functions of focus group research	222
Quasi-unique affordances of focus group research	222
Mitigating researcher's authority	223
Leveling power relations	223
Approximating the "natural"	224
Disclosing eclipsed or invisible connections and backfilling gaps in understanding	224
Drawing out complexity, nuance, and contradiction	225
Saturating understandings	225
Promoting political synergy	227
Summary	227
Figured worlds of focus group research	227
Figured World 1	229
Figured World 2	230
Figured World 3	230
Figured World 4	231
Figured World 5	231
Data collection and data analysis	232
Indexicality and its functions in focus group research	233
Analyzing collective conversations	233
Prospect: a social-cultural-political imaginal of focus group futures	234
"Mapping" the future of focus groups	236
Brief summaries of the studies	239
Study 1	239
Study 2	240
Study 3	241
Focus groups as tools of subversion	241
Widening the definition of focus groups: fabulative conversations, paratactic commoning, and collective self-inquiry	243
Using focus groups to promote and sustain global conversations	244
Final thoughts: focus groups as tools for interdisciplinary, critical, global explorations	246
References	248

Introduction: our perspectives and biases

We are not cutting people open. We are educators. Our work is about learning and teaching through social interaction and other communicative channels. We value pedagogies of reflective praxis that celebrate complexity and nuance. We value the transformative powers of both human and material agencies, and thus we value the use of focus groups to renew or transform education purposes, policies, and practices. Consequently, we are not going to write another traditional history of focus group research or another focus group research "cookbook." There are already many of those (e.g., [Lee, 2010](#); [Morgan and Hoffman, 2004](#)). Instead, we are going to do several things a bit differently in this chapter. First, we are going to lead with visuals and use text to support and unpack them. With rapidly changing communication practices afforded by digital technologies—graphic novels and informational texts, YouTubes, multimodal texts as normative, and data visualization dominating news stories in most media forms—this seems to make common 2.0 cultural sense. Second, whereas most handbook and encyclopedia chapters tend to be retrospective, we are taking a prospective stance in this chapter, concentrating more on where focus group research is going than where it has been. Finally, although we provide some history and practical advice on the use of focus groups in research, our arguments are largely conceptual and speculative ([Fig. 1](#)).

In the first part of this chapter, we review our previous conceptual work, which we believe is part of a radical methodological "break" or "turn" in focus group research. In the second part of the chapter, we argue that experimental uses of focus groups have induced an ontological/epistemological/axiological cascading effect that has all but forced us to reimagine social science research as productive in [Foucault's \(1977, 1988\)](#) sense of this term—as always producing new forms of life ([Wittgenstein,](#)



Fig. 1 Figured Worlds in focus groups timeline.

1953) or new ways of being in the world (Heidegger, 1962). Finally, we comment on the promise and potentials of focus group research after this methodological “break” or “turn.”

Looking at spectrums of various kinds, we see that qualitative researchers have pushed for focus group research to lead the way in blurring the lines between science and art, epistemology and ontology, fact and value; they have also pushed for research to have greater political teeth—not just to document the world but to change it strategically, usually in the interest of greater inclusion, equity, and social justice (Fig. 2).

Retrospect: traditional uses of focus groups in social science research

We like to think in/with the metaphor of spectrums because they help us see the wide range of variants or inflections of all social phenomena, as well as the many ways we might study them. On the traditional/historical end of the focus group spectrum, scholars have commonly imagined focus groups in relatively static terms. For example, they have proposed that the ideal size of focus groups is five to eight participants (e.g., Morgan and Hoffman, 2004). In this regard, they have argued that three or fewer participants often results in awkward pauses, long silences, and inertia in getting conversations off the ground. Conversely, groups larger than eight participants tend to leave some people voiceless and often results in tangent hopping. From this historical perspective and especially in market research and psychological research, it is common to stage only one focus group of about one to 3 h in duration. Finally, in most extant focus group research, the researcher/facilitator has typically chosen the topics for discussion and directed/controlled the flow of talk during discussions.

At the other end of this spectrum, dynamism, contingency, unpredictability, and transformation have been celebrated and practiced. From this perspective, scholars have foregrounded “the fundamental entanglement of the material and the nonmaterial, the human and the nonhuman” as “key to thinking and acting differently in the world” (Dimitriadis, 2016, p. 144). This move is fundamentally important given the decentering of the self in contemporary identity theory and the powerful insights from the sociology of scientific knowledge (especially Latour, 2007) about the distributed locus of agency in the ongoing production of material and social phenomena.

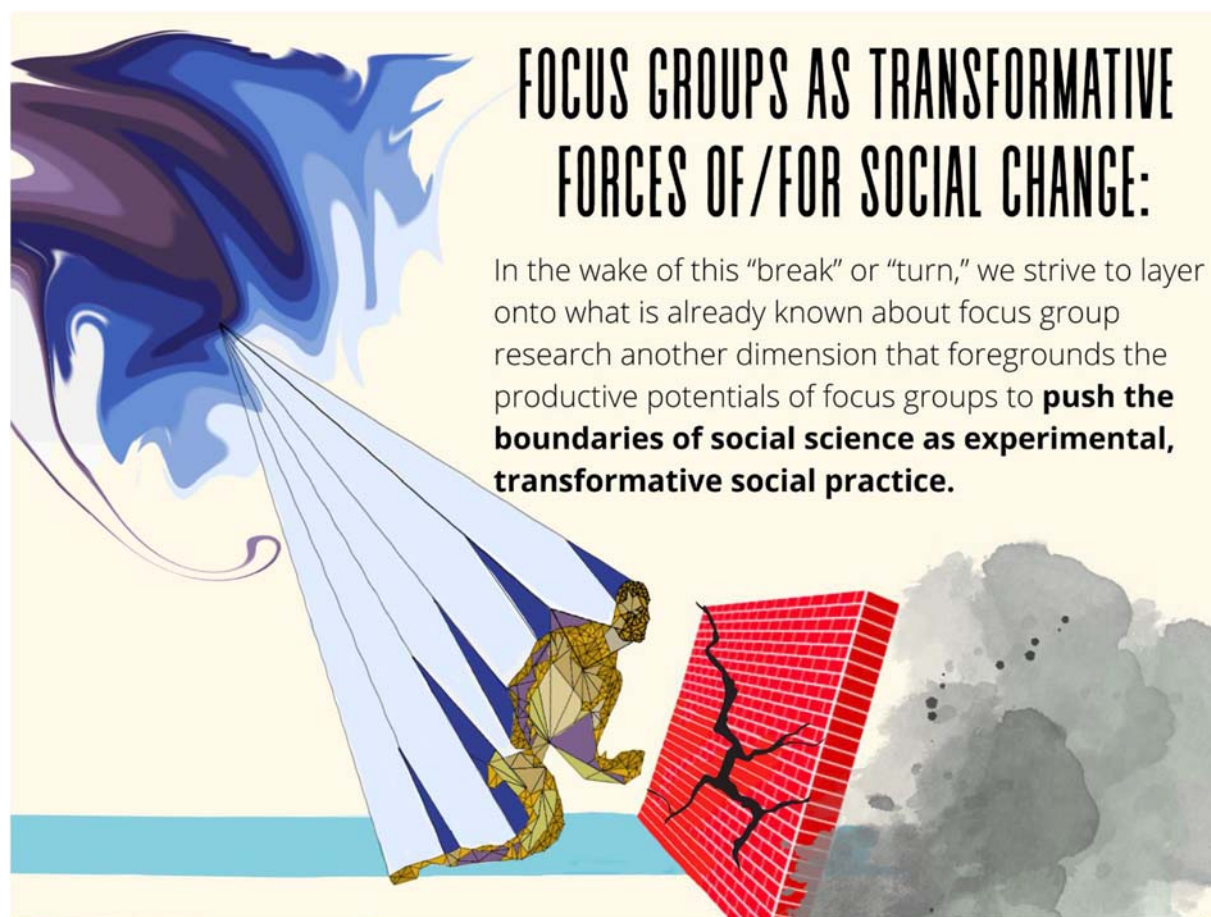


Fig. 2 Focus groups as transformative forces of/for social change.

Additionally, scholars frequently deploy Barad's (2007) construct of "intra-action" to work in this new ontological/epistemological/axiological space. Intra-action "signifies the mutual constitution of entangled agencies ... those distinct agencies do not precede, but rather emerge through, their intra-action" (Barad, 2007, p. 33). "Intra-action" thus replaces "interaction," rejecting the idea that individual entities exist separate from each other and only later enter into relations/relationships. Instead, intra-action views agency as a co-constitutive dynamism of forces (Barad, 2007, p. 141) in which humans, non-human entities, and discourses are never separate things but are constantly co-creating both each other and reality itself. From an intra-action perspective, objectivity becomes impossible because people and scientific instruments are always already part of the process that produces facts or outcomes (Fig. 3).

In our work, we started to notice that strategic assembly often intra-acted with the content, mood, tone, and outcomes of focus group conversations. Further, we found researchers such as Collins (2007), who studied sex workers in the Philippines, met their participants where they were already, where they felt safe (cafés, for example), in groups they had already formed, and within in activities they had already made their own. Collins did not schedule meetings, but instead was constantly on the lookout for when members got together on their own at different times on different days in different places, and she folded herself into their activities. Documenting their conversations and interactions, Collins was able to gather information that was more authentic—issuing from natural interactions relatively undetermined by the researcher. Although this approach is not always feasible or desirable, it speaks to the importance of strategic assembly in focus group research.

What about parties in conflict? In bringing heterosexual couples experiencing conflict together (or any social group—small or large—experiencing conflict), for example, what considerations would researchers make? Might they stage homogeneous focus groups for wives and husbands followed by heterogeneous ones? In these heterogeneous focus groups, would couples be the participants, or would the groups not include couples but just be cross-gendered groups? If couples were selected and arguments between couples emerged, how would facilitators handle this either within or beyond the scheduled focus groups?

More generally, what is the facilitator's role anyway? As focus group research has evolved, facilitators have increasingly been inclined to take a back seat and allow the groups to manage the topics and the flow of conversation. What about considering assigning the facilitative role to a research participant or to rotate that role across participants? Even when researchers take a back seat, though, how can they help participants go deeper without being intrusive or overly directive?

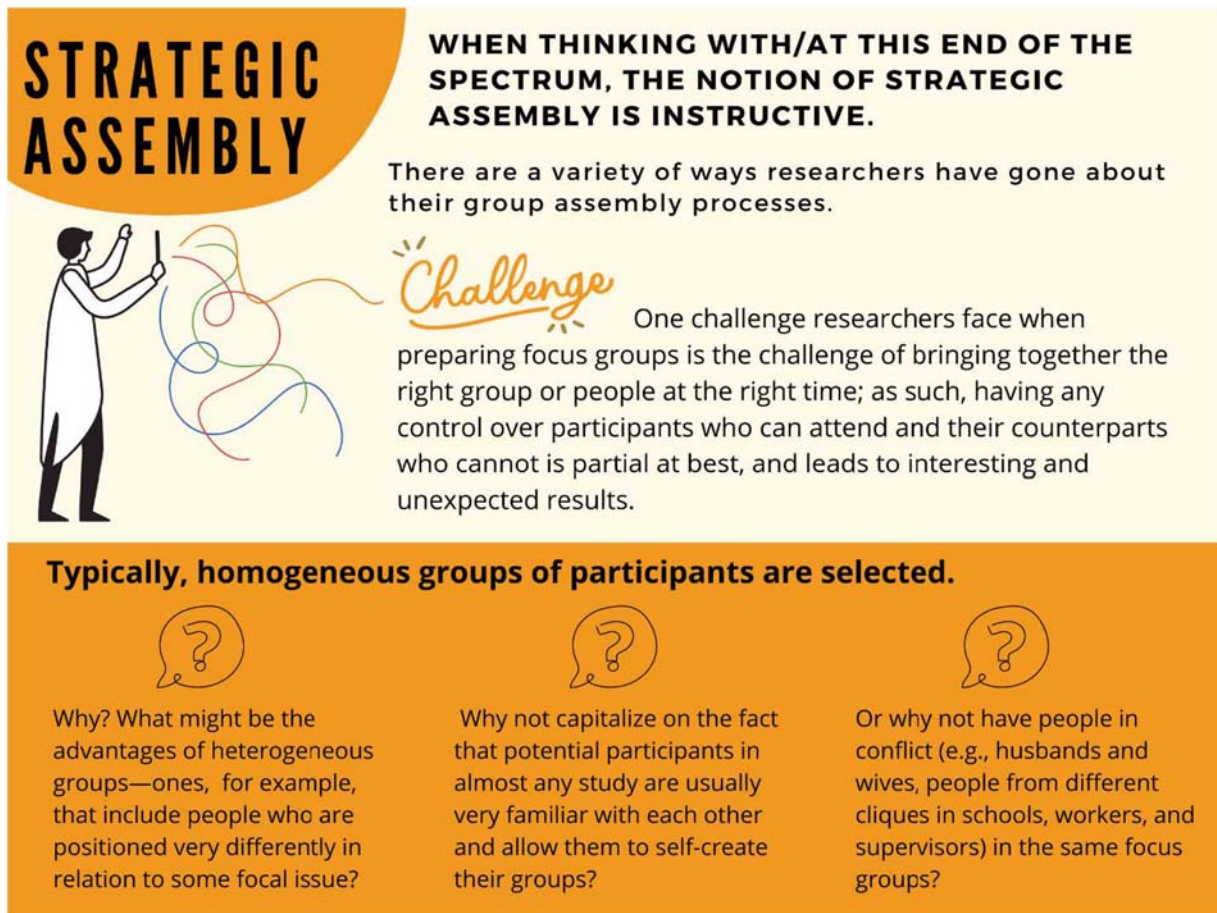


Fig. 3 Strategic assembly.

Multiple functions of focus group research

Kamberelis and Dimitriadis (2005) noted that focus groups have served three primary functions in qualitative research over the years: a pedagogical function, a political function, and an inquiry function. The *pedagogical* function of focus group work foregrounds what might be learned from collective conversations about issues profoundly familiar (in one way or another) to focus group participants. This function also highlights the dialogic nature of focus group interaction, as well as the possibility for transformative encounters (e.g., Freire, 1970; Morley, 1980). The goal of the *political* function of focus group work is advocacy and social change for social justice. This function alerts us to the sources of collective support that can be assembled in relation to social and political issues, as well as the dangers of naïve notions of collectivity (e.g., Behar, 1993; Dill, 1994; Espiritu, 1997; Madriz, 2000). The *inquiry* function of focus group work alerts us to deep epistemological issues and concerns around “the research act,” including the complex negotiations between “self” and “other” in inquiry (e.g., Lather and Smithies, 1997; Radway, 1984). Although one function is usually dominant in any given research endeavor, all three focus group functions are usually present to some degree, reflecting and refracting the content of focus group work in different ways (Fig. 4).

Although one function is usually dominant in any given research endeavor, all three functions are usually present to some degree, reflecting and refracting the content of focus group work in different ways. Moreover, these three functions are seldom related in simple, unproblematic ways. Political interventions, for example, do not necessarily emerge from inquiry, but in some situations, they might. Similarly, pedagogy is not always central to activist work, but it can be and often is. Within any given project, these different dimensions or functions of focus group work emerge and interact in distinct and often disjunctive ways, eventually resulting in some unique interactive stabilization.

Quasi-unique affordances of focus group research

All strategies for gathering information from research participants have different affordances—that is, any given strategy is likely to allow researchers to learn some things and not others or some things more than others. In this regard, Kamberelis and Dimitriadis (2011) noted that focus groups have some quasi-unique affordances compared to other data collection strategies. It is not so much

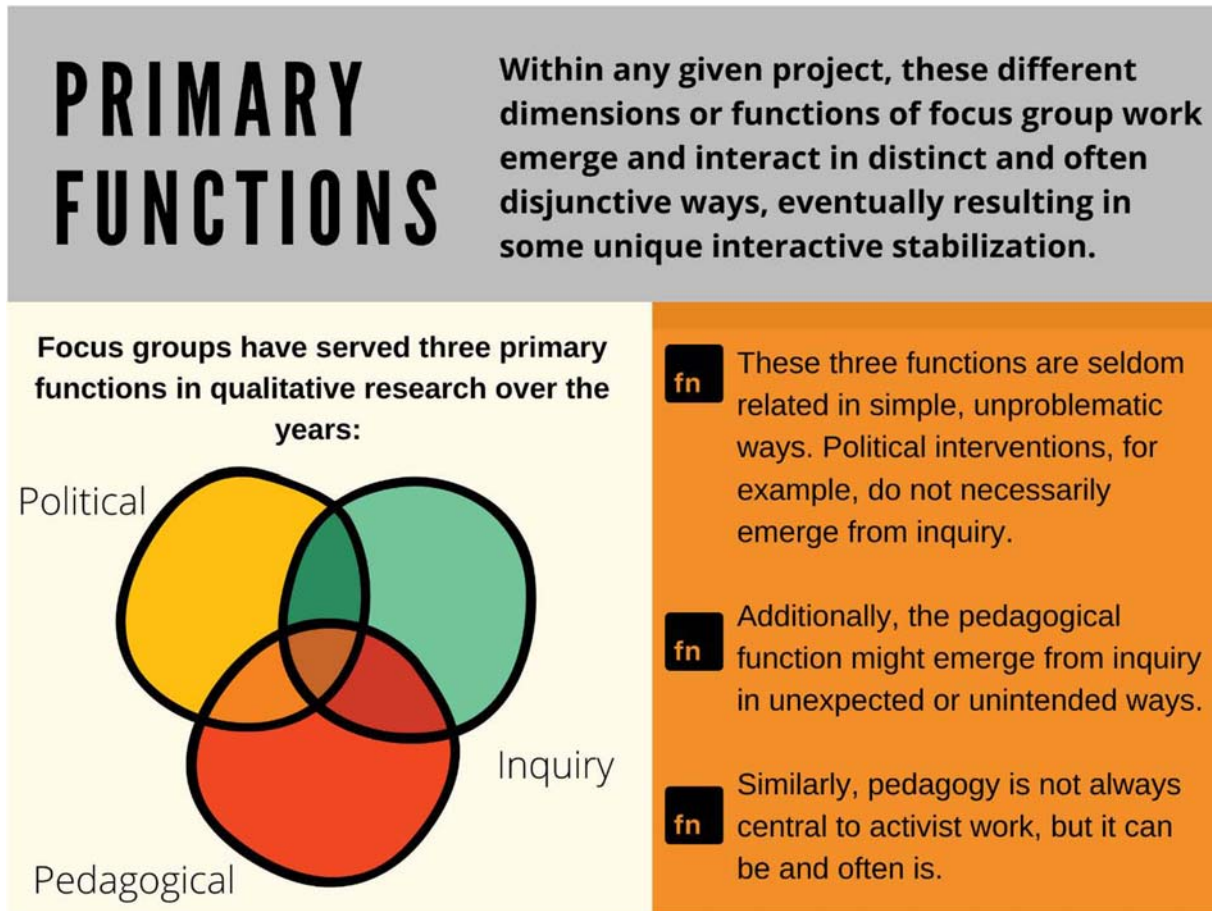


Fig. 4 Primary functions.

that they allow researchers to gather certain kinds of information that could not be gathered to some extent using other strategies; it is more that they are likely to allow researchers to generate more focused, richer, more complex, and more nuanced information, especially in relation to certain topics or domains of inquiry.

Some of the most important and powerful affordances of focus group work we have written about previously include (a) mitigating the researcher's authority, (b) leveling power relations among participants and researchers, (c) approximating the "natural," (d) backfilling gaps in understanding yielded by other data collection strategies, (e) drawing out depth, complexity, and nuance of the phenomena under study, (f) saturating understandings, and (g) promoting political synergy that can lead to social change.

Mitigating researcher's authority

Authority and control can lead to premature closure and weak standards of evidence. Focus groups offer the researcher unique opportunities to relinquish authoritative power, which often results in more and richer data, especially with marginalized populations and in relation to sensitive topics (Fig. 5).

A key exemplar of the power of mitigating the researcher's authority in focus group research is Hill's (2009) study of the ways complex personal and social issues were explored with high school students through the lens of rap and hip hop. Especially important here was Hill's extensive knowledge of rap and hip hop and how this knowledge paved the way for more equitable sharing between Hill and his research participants. We recommend reading this book to get a deeper sense of this affordance at play.

Leveling power relations

Focus groups also encourage leveling power relations among researchers and research participants, which usually allows researchers to witness something close to "natural" group dynamics, yielding data with considerable ecological validity (Fig. 6).

A key exemplar of leveling power relations among participants and researchers in focus groups is Duneier (1992). Duneier studied the lives and stories told by a group of African American (and some white) men who hung out together in a neighborhood eating establishment called "Valois" on the south side of Chicago in the late 1980s and early 1990s. Importantly, it took Duneier many months and many awkward quasi-conversations to finally build trusting relationships with his participants, after which



Fig. 5 Mitigating researcher authority.

hierarchical power relations began to erode and conversations became more authentic and more dynamic. We recommend reading this book to get a deeper sense of this affordance at play.

Approximating the “natural”

Although not entirely “natural,” focus groups can serve researcher purposes of approximating natural interactions when carefully designed. Focus groups have the potential to “create” naturalistic spaces for re-performing everyday knowledge and practice (Fig. 7).

Fine et al. (2004) *The Echoes of Brown Project* is an excellent exemplar of this affordance. This project involved working with scores of youths to investigate the “so-called” racial achievement gap from *their* perspective. When investigating this topic through typical academic means fell flat, Fine and her colleagues consulted with the youth about alternative investigative tools. Together and with community elders and local artists of various kinds, they adopted a more local, “oral history” approach to generating knowledge and a more “performative” approach to sharing what they learned—writing, directing, and producing a play of sorts “that brought together political history, personal experience, research, and knowledge gathered from generations living in the immediate and long shadow of Brown” (p. 33).

Disclosing eclipsed or invisible connections and backfilling gaps in understanding

Focus groups are particularly useful for revealing partially invisible aspects of phenomena under study and generating more complex, nuanced understandings of those aspects—either during focus group conversations themselves or by chasing down other data mentioned in these conversations (Fig. 8).

Martin’s (1995) book on the history of immunology in the United States is a great example of this affordance at work. By viewing immunology from the days of polio to the age of AIDS, through both the lens of scientific discourses of immunology and political-economic discourses of new capitalism allowed her to disclose perspectives that were counterintuitive and thus relatively invisible. And collecting data from many different kinds of people (e.g., scientists, political scientists, social activists, hospital patients) and in many different social spaces (e.g., laboratories, health care centers, neighborhood associations, support groups, block parties, community meetings) further helped Martin disclose and render complex, nuanced, and even polyphonic accounts of health



Fig. 6 Leveling power relations among participants and researchers.

and illness as a kind of social-cultural-racial survival of the fittest—an insightful characterization that required a complex theory-method assemblage to generate.

Drawing out complexity, nuance, and contradiction

Focus groups can be thought of as the “tip of the iceberg” when it comes to data collection. Conducting focus group research can help to remind readers to work against premature consolidations of data, thus signaling the limits of reflexivity (Fig. 9).

This affordance is especially visible in Fine and Weis’ (1998) work with poor and working-class Caucasian, African American, and Puerto Rican men and women in Buffalo, New York and Jersey City, New Jersey (See also Weis, 2005). These researchers were interested in how their research participants experienced and understood social problems like domestic abuse, the police, and equity in school. As it turns out, participants’ experiences and consequent understandings varied widely as a function of their intersectional identities (race-ethnicity-class-gender-etc.). Thus, these researchers’ findings about the focal issues of their study were more kaleidoscopic and indexical than focused and contained.

Saturating understandings

In many ways, saturation is the converse of indexicality, which we will discuss in the second half of this chapter (Fig. 10).

Fetene’s (2009) study of how young college students in Ethiopia understood HIV/AIDS is instructive here. Individual interviews suggested that government and media explanations about HIV/AIDS as rooted in an ignorance of safe sex practices were somehow off-base. Focus groups revealed that young people are acutely aware of such practices but actively resisted them for a variety of personal, social, and political reasons. Moreover, whereas his participants were reluctant to talk about sensitive topics related to sex and HIV/AIDS in one-on-one interviews, the floodgates opened when they are asked to discuss these same topics with peers in focus group conversations, which led to rich, relatively saturated understandings that challenged many status quo beliefs about childhood and adolescence, sex, condom use, and HIV/AIDS.

IF EVERYDAY INTERACTIONS THEMSELVES ARE PERFORMANCES, THEN FOCUS GROUPS CAN BE A WAY OF INVOKING AND ACCESSING THE VERY DYNAMICS OF THESE PERFORMANCES.



Approximating the "Natural"

Focus groups can afford a closer approximation to natural interactions than many other data collection strategies.



Focus groups have the potential to "create" naturalistic spaces for re-performing everyday knowledge and practice.

Fig. 7 Approximating the "natural".

DISCLOSING ECLIPSED OR INVISIBLE CONNECTIONS

BACKFILLING GAPS IN UNDERSTANDING

YIELDED BY OTHER DATA COLLECTION STRATEGIES



Researchers and their participants contribute to shaping the objects they study in the absence of extant models for contextualizing referents of research such as "the world system," "capitalism," "the state," "the nation," etc.



Focus groups can illustrate the myriad kinds of collective conversations that can emerge when interactions between researchers and research participants are not prefigured but emergent.

Fig. 8 Disclosing eclipsed or invisible connections.

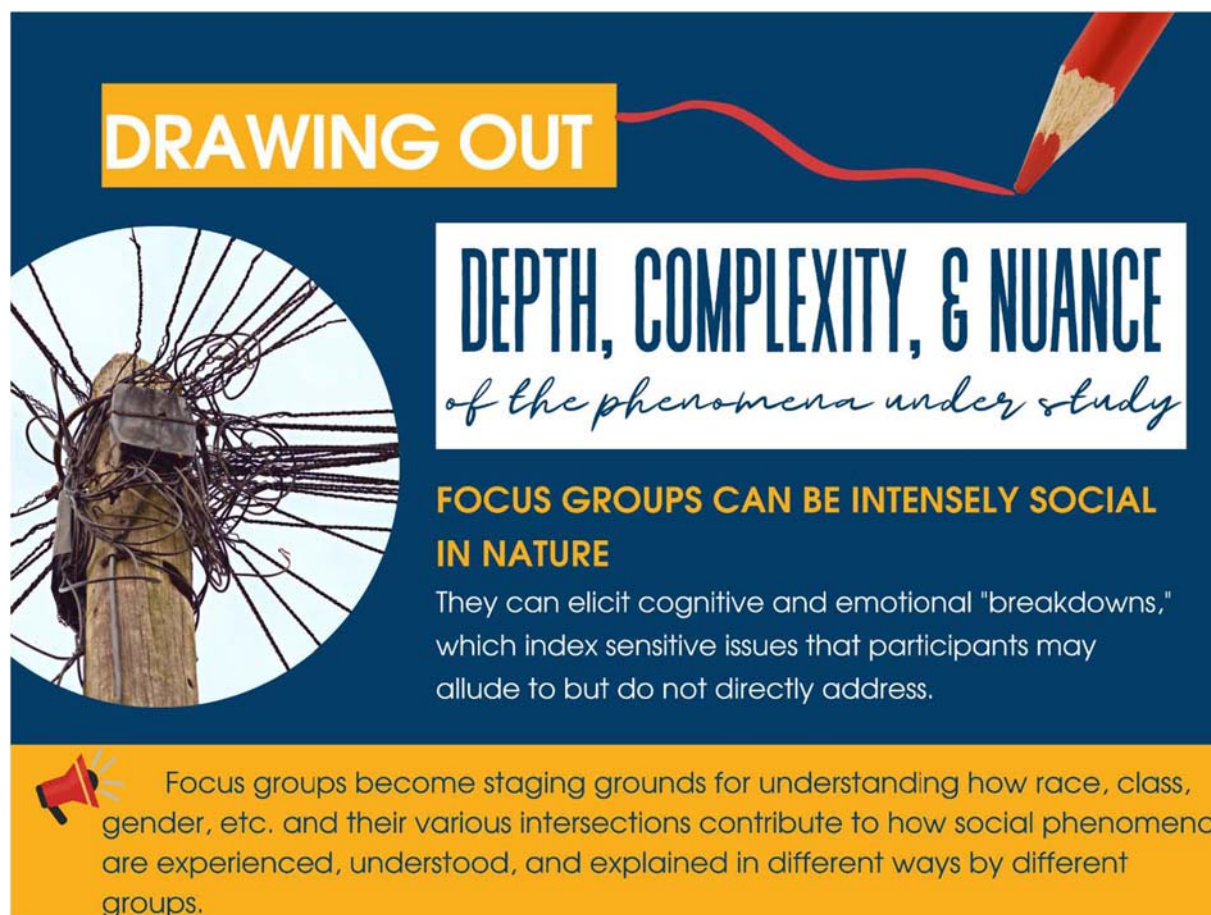


Fig. 9 Drawing out depth, complexity, and nuance of the phenomena under study.

Promoting political synergy

Whether they are designed to do so or not, focus groups often promote political synergy among participants that often leads to political activist and advocacy work (Fig. 11).

Tuck's (2012) research on students who dropped out of high school but enrolled in GED programs to "repatriate" their school experiences and recover some sense of dignity is an outstanding embodiment of this affordance. Importantly, what these students revealed in focus group conversations was taken up by their school districts to create policies and practices that were less alienating to these students. This process, however, revealed some fault lines that threatened institutional stability of school buildings and districts in non-trivial ways. Largely due to the integrity of Tuck's research design and process, however, these buildings and districts were able to manage these threats without abandoning the positive change efforts they had created for students.

Summary

We have described each of the key affordances identified by Kamberelis and Dimitriadis (2011) and provided an exemplary research study that illustrates each of them. Like the three primary functions of focus group research, however, most studies make use of multiple affordances either explicitly or tacitly, and when designing any new study, it is useful to think of facilitation strategies that are likely to draw out as many affordances as possible.

Figured worlds of focus group research

Using Holland et al. (1998) trope of the "figured world" Kamberelis et al. (2018) introduced the idea that the forms and functions of focus groups vary considerably depending upon the ontological-epistemological grounding of the research study. According to Holland et al. (1998), figured worlds are culturally constructed, socially produced horizons for understanding and acting that recognize particular kinds of actors, assign certain meanings to specific acts, and value some outcomes over others. Put another way, figured worlds are cultural imaginaries constituted by "people like us" who think, act, desire, and use the same social and cultural tools toward similar goals.

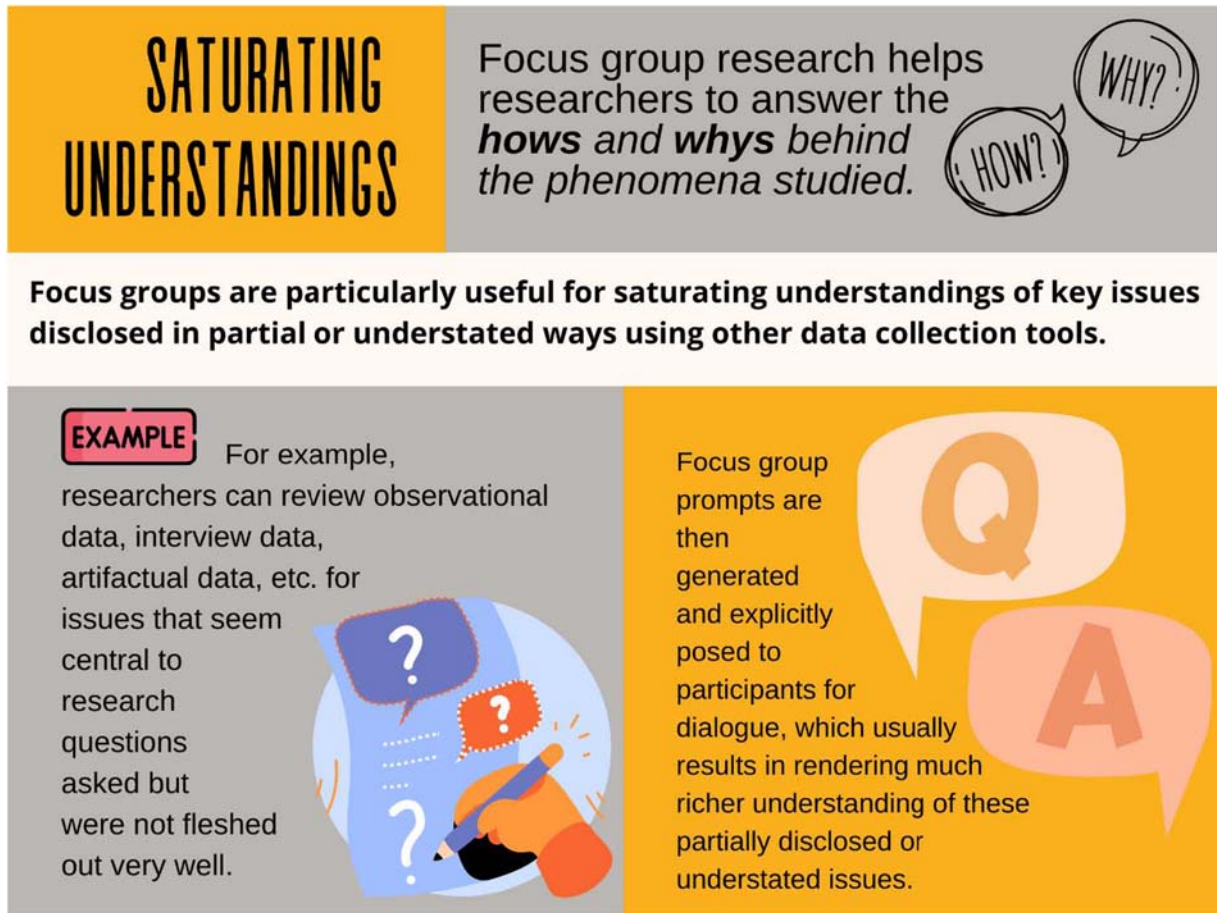


Fig. 10 Saturating understandings.

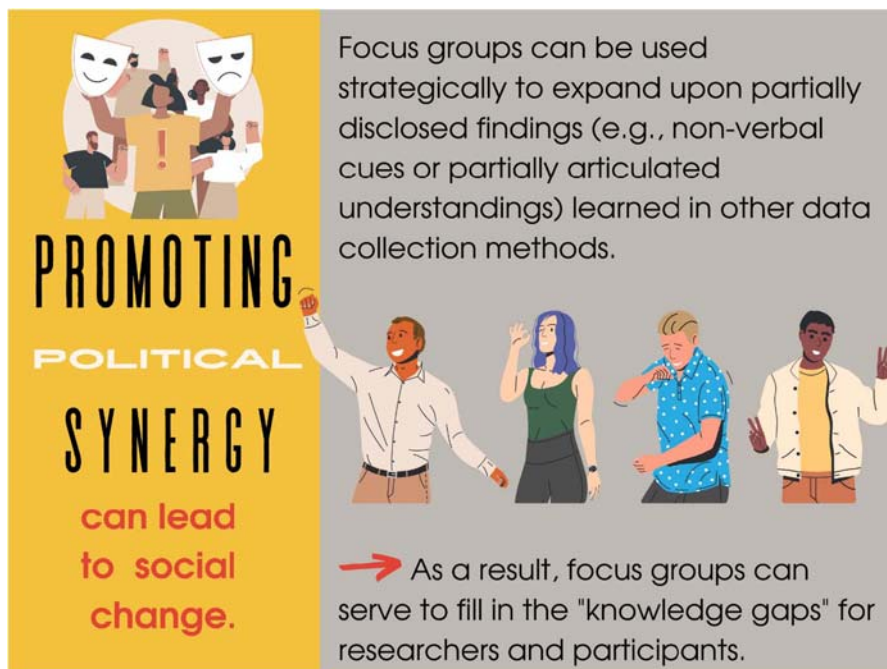


Fig. 11 Promoting political synergy that leads to social change.

George Kamberelis and Greg Dimitriadis have been cultivating the idea that there are multiple figured worlds of qualitative inquiry since their 2005 book *On Qualitative Inquiry*. In that book, they posited four chronotopes or figured worlds of qualitative inquiry. In Kamberelis et al. (2018), we added a fifth figured world of qualitative inquiry. Each figured world entails largely unique ways in which researchers think about the research process, use data collection strategies, and analyze, interpret, and explain information. Each figured world is complex and involves multiple, related assumptions. Chief among these are: (a) how knowledge and truth are construed, (b) what kinds of research questions are asked by whom and for what purposes, (c) how the relations between subjects and objects are thought about—including whether and how much agency different types of subjects and objects have, (d) whether and how reality is seen to be “brute data identifiable” (Taylor, 1979) or as culturally/socially/politically/economically produced (e.g., Foucault’s, 1977 “discourses”), (e) and how language and other cultural tools are thought about and used within each of them. Finally, as tropes or metaphors, figured worlds are more like topoi than surveyed plots separated off neatly in relation to each other. Additionally, each figured world leaks into, overlaps, faces off against, and/or slides into others.

Figured World 1

Figured World 1 is the figured world most of us leaned about in our high school science classes. It presumes a real world separate from and unaffected by humans. It also presumes that all knowledge of that world is discoverable. Through theory-making and hypothesis testing, the work of science is to represent that real world in ways that make it predictable and controllable (e.g., Boyle’s gas law, Piaget’s stage theory of child development, operant conditioning, $E = mc^2$). Like most modernist imperatives, the ultimate (and achievable) goal of the march of science is to discover and represent everything that is knowable about the objective world, which will presumably answer all human questions and end all human problems (Fig. 12).

Figured World 1 leans on facts and objectivism to provide quantifiable data for others to replicate. Such truths are meant to serve as “best practices” for the general population to follow and benefit from. Unlike the individual differences and affordances discussed in later figured worlds, Figured World 1 seeks statistical significance and verification.

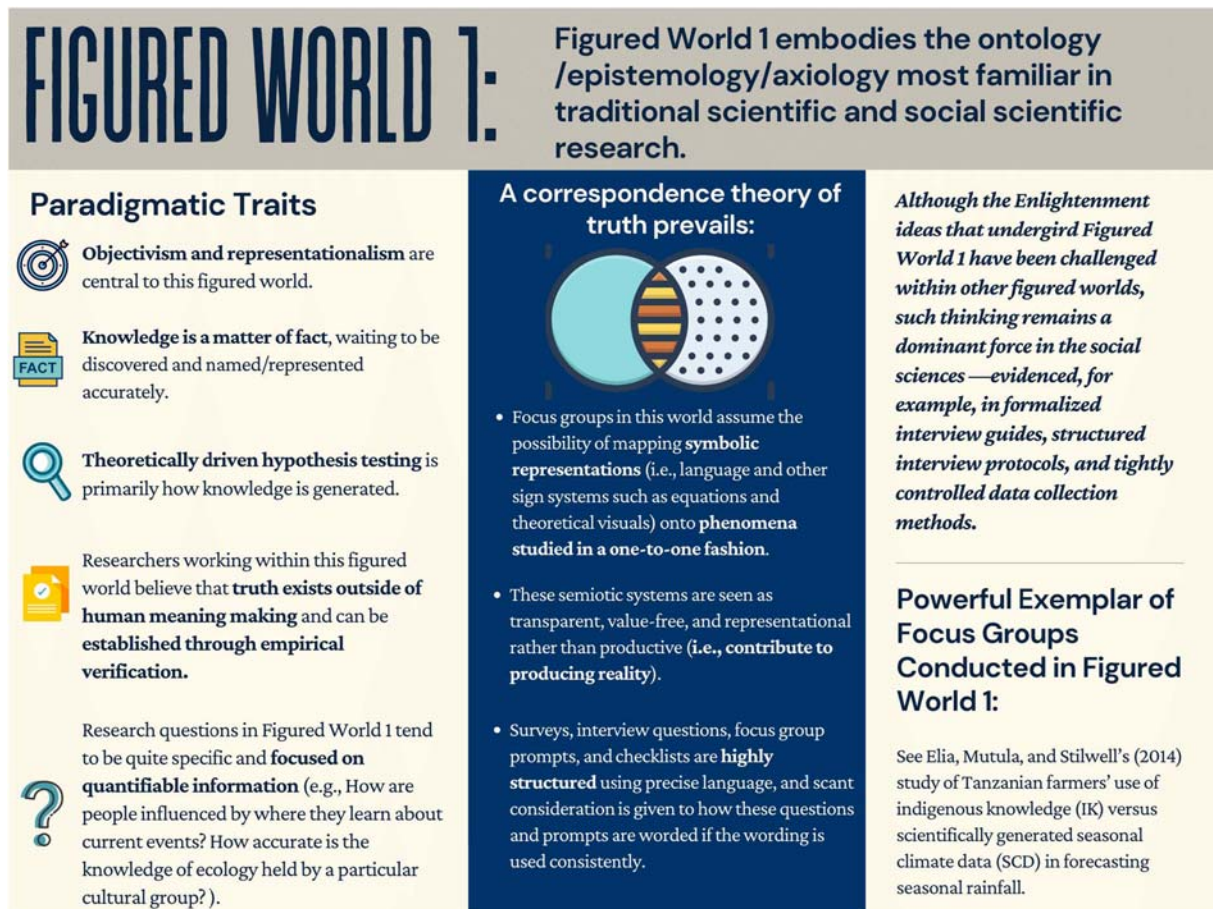


Fig. 12 Figured World 1.

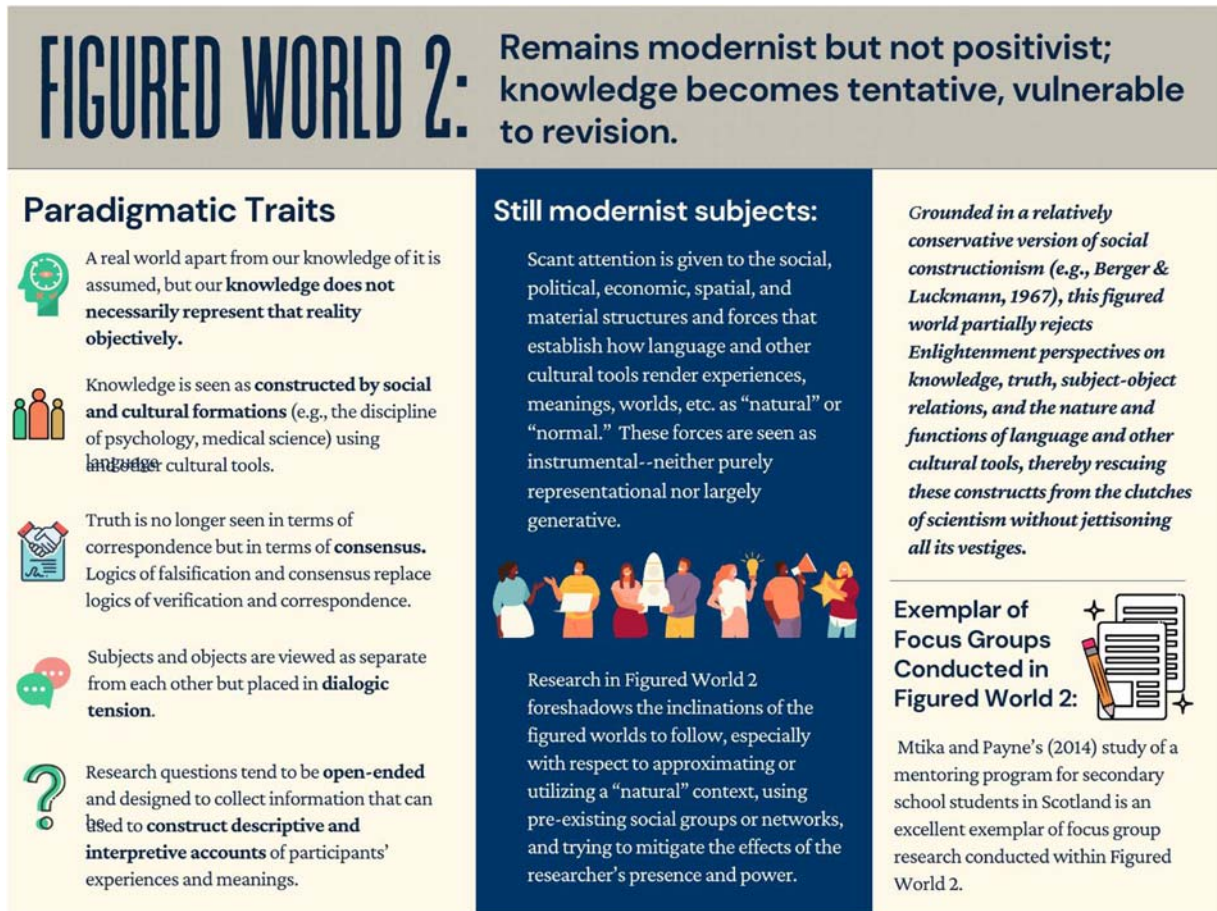


Fig. 13 Figured World 2.

Figured World 2

In Figured World 2, the constraints of scientism begin to loosen. Subjects and objects in Figured Word 2 are viewed as separate from each other but placed in dialogic tension, which is the hallmark of philosophical hermeneutics (the foundation upon which this figured world rests). Practices and meanings are understood in relation to the contexts in which they occur, and according to the principles of the hermeneutic circle where understanding parts of phenomena (a text, person, event, culture) always involves understanding the whole phenomenon (Fig. 13).

Because the goal is to understand people and their worlds, human subjects in Figured World 2 are still modernist subjects. These subjects have considerable agency to use the cultural (including scientific) tools at their disposal to construct themselves and to "discover" their worlds—even if they do not yet realize those worlds are always in the making. Unlike the more critical figured worlds discussed later in the chapter, the idea that knowledge and truth are effects of discourses (e.g., Foucault, 1977, 1988) is just emerging in Figured World 2.

Figured World 3

Figured World 3 is critical (e.g., Gramsci, 1971; Freire, 1970/1993) but not yet post-qualitative (e.g., Deleuze and Guattari, 1987; Rancière, 1991). In neo-Marxist terms, this is a matter of ideology critique, the goal of which "is to discern ... those 'ideologically frozen' relations of dependence that can be transformed only through critique. Thus, the critical approach is governed by the interest in emancipation, which Habermas also calls *self reflection*" (Ricoeur, 1992, p. 289). Habermas (1971) referred to this process as "emancipatory rationality," which is a mode of thinking/being that allows people to escape the lures of oppression through collective self-reflection. Social movements such as second-wave feminism and the civil rights movement exemplify how Habermas' rational, emancipatory, de/recolonization project has been historically enacted (Fig. 14).

Simply put, scholars working within Figured World 3 encourage people to engage in discursive practices designed to instigate change. However, the idea that researchers and research participants are also victims of language and discourse is not well developed, nor are other non-human phenomena seen as having agency. For instance, if electricity were introduced to a non-electrified geographic region, the people who brought the electricity and not electricity itself would be seen as agentic in people's lives. As we show later in the chapter, agency is viewed quite differently in Figured Worlds 4 and 5.

FIGURED WORLD 3:

Researchers become critical, co-investigators.

Paradigmatic Traits



This figured world is explicitly concerned with **questions of power**, especially how human activity is embedded in **hegemonic structures that reproduce existing structures of power**.



The goal is to **motivate social action that liberates people from the constraints** of extant ideologies about how they (should) think, feel, and act.



Researchers here assume that **surface-level meanings and actions hide embedded structural conflicts, contradictions, and fallacies** that maintain the status quo while hindering dialogic and democratic forms of life and their attendant practices.



The goal of all discursive and material activity in Figured World 3 is **praxis: the process of linking theory and practice to enhance the possibilities** of specific groups, thus improving the world



Reciprocal relationships must lead to common goals that express the **transformative possibilities of a dialogic community**.



Research questions are **emergent and open-ended**, typically inviting research participants to adapt these questions based on their own knowledge and experience before developing them further.

The path to knowledge:

- With emancipatory rationality as the path to knowledge, and with **knowledge seen to be socially/culturally/politically/economically constructed**, truth in Figured World 3 is a matter of consensus achieved through dialogue and debate in which reasoned argumentation, adequate evidence and warrants are all highly valued.
- **Discourses are treated more as objective facts** than as force fields that delimit what can or cannot be discussed, considered normal, and believed as true.
- Language and other cultural tools in this figured world are instrumental—used by people in dialogue to **deconstruct the conditions of possibility of their lives and to work towards different and better conditions**.

For the most part, the subject-object dualism central to Enlightenment thinking remains present (but is troubled). Individuals are still viewed as both rational and free, but these characteristics are impaired by a false consciousness induced by prevailing institutions of power. Therefore, rationality is no longer a property of the individual subject, but is instead best produced in unconstrained dialogue with others. This means that Figured World 3 is grounded in a social, rather than an individual, kind of idealism.

Notable Focus Group Work in Figured World 3:

See Baldissera, Bueno, and Hoga's (2012) action research on the sexuality of older adult women living in southwest Brazil.

Fig. 14 Figured World 3.

Figured World 4

Knowledge in Figured World 4 is viewed as an effect of power that is produced, reproduced, and transformed through discursive practices informed by discourses (Foucault, 1977, 1988). Therefore, researchers are (and should acknowledge) that they have been constructed within academic discourses that constitute filters through which they see and act in the world. Thus, getting in and out of the way of their participants becomes a key activity for researchers within these figured worlds, which illuminates the process by which voices of others are represented by juxtaposing narratives, counter-narratives, and metanarratives. Even more than in Figured World 3, agency in Figured World 4 is limited in that it is always already an effect of discourses (and not free will or individual intention). Because knowledge and subjects are both effects of discourse, agency has to do with deconstructing discourses and using them strategically (Fig. 15).

The characteristics and concepts guiding Figured World 4 require acknowledging that the research process itself is often riddled with contradictions. Constructs such as "giving voice" are neither "innocent" nor self-evident. Instead, they are named, troubled, and performed through talk, social interaction, and writing in ways that allow participants to re-imagine and enact possible worlds and possible selves.

Figured World 5

Figured World 5 is the most embryonic world on our heuristic landscape; it is thus "bricolage" par excellence—assembled from bits and pieces of philosophy and theory in intentional but not necessarily scientific ways. Yet that is part of the point, as truth is seen as both chimerical and an effect of hegemony. As in Figured World 4, language and discourse are seen as productive rather than representational. However, non-human and non-discursive forces— affect, esthetics, space, mobility, materiality, etc.—are also seen as agentic and productive, which means this work challenges the idea that the intelligible gets ordered and controlled primarily through language/discourse (Fig. 16).

Materiality is also a constitutive force in Figured World 5 as it involves the presence and organization of physical and/or digital materials that affect how social activity and work get done. Like materiality, space, spatial relations, and spatial mobility/immobility can be forms of agency in this figured world. Finally, Figured World 5 is fundamentally about the possible. Researchers and research

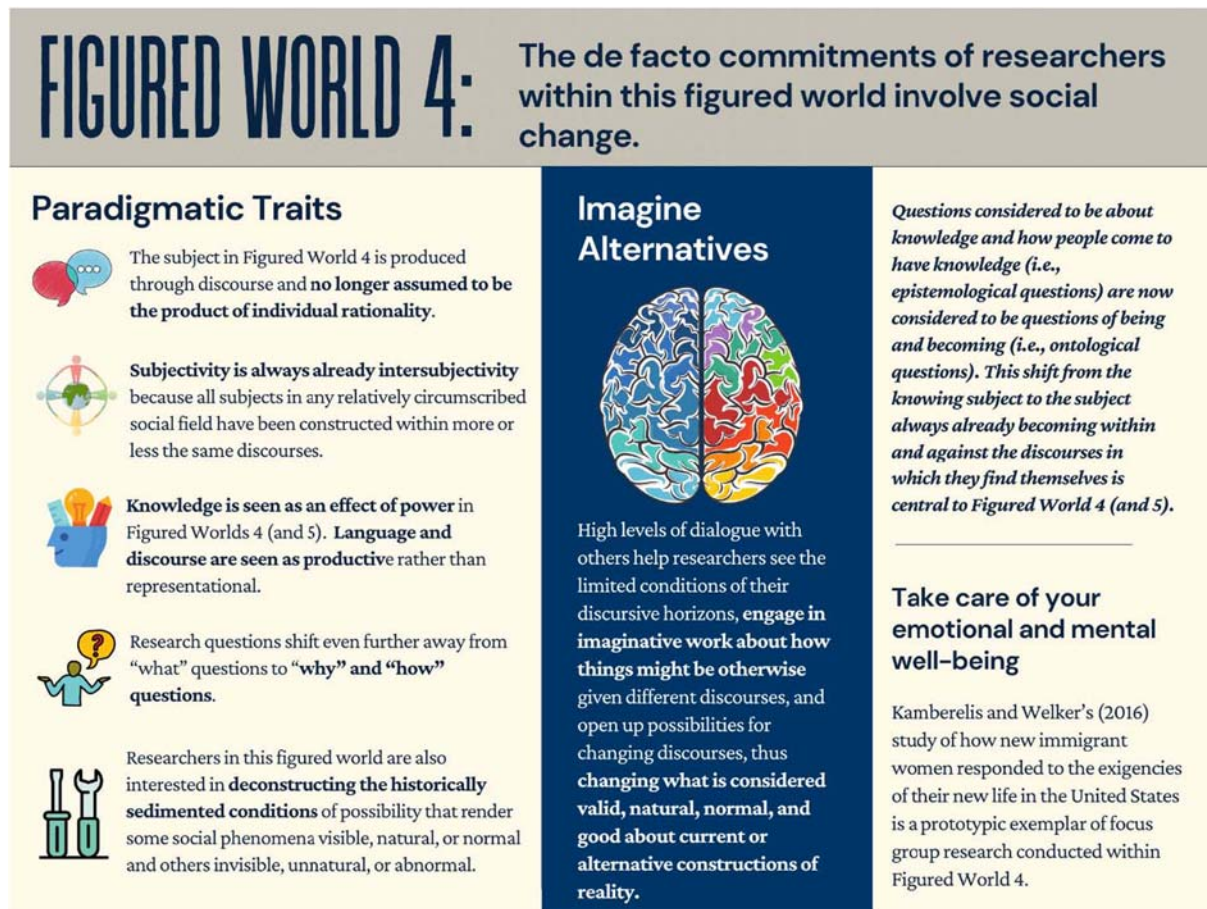


Fig. 15 Figured World 4.

participants typically work together both to deconstruct the worlds in which they live and to construct new worlds viewed as both possible and desirable.

Data collection and data analysis

Having introduced this “figured worlds” heuristic, it will not surprise readers that we encourage thinking about collecting and analyzing focus group data using a continuum model. In this section, we discuss the range of data collection and data analysis strategies common to focus group research located within different ontological/epistemological/axiological spaces (Fig. 17).

Collective conversations can be audio recorded and/or video recorded. Having video recordings of these conversations can provide a great deal more data for researchers to work with and learn from. In many of our studies, we have been able to see the meeting rooms where groups met, the facilitators stance and presentation, facial expressions, body language, what the weather was like that day, and even what participants were eating for lunch! All these visual elements provided insights that would have been lost had we only collected audio data.

Probably the most important factor in conducting focus groups and collecting focus group data is the positioning of the researcher. Three quasi-unique affordances of focus group research—(a) mitigating the researcher’s authority, (b) leveling power relations among participants and researchers, and (c) approximating the “natural” are instructive here, as are the various positions and positioning strategies implied in our discussion of figured worlds. Basically, focus groups conducted at the traditional end of the figured worlds continuum are typically tightly facilitated by the researcher with direct, highly focused questions, and conversations routinely unfold much like the initiation-response-follow-up (IRF) speech genre common in school classrooms. In contrast, focus groups conducted at the post-qualitative end of the figured worlds continuum are typically owned by and facilitated by the participants themselves after the researcher poses one or more open-ended prompts, and conversations are more likely to resemble those of coffee klatches or carnivalesque, multi channeled speech events (Bakhtin, 1981). Finally, focus groups conducted toward the middle of the figured worlds continuum fall somewhere between these two extremes.



Fig. 16 Figured World 5.

Although we understand the value of research located within the ontological/epistemological/axiological parameters of all figured worlds, because of the quasi-unique affordances of focus groups as data collection tools, we find the conversations toward the post-qualitative end of the figured worlds continuum to be most productive—both in terms of the richness of information gathered and the fact that more participant-owned, carnivalesque conversations are often transformative; they often disarticulate and rearticulate reality (Deleuze and Guattari, 1987)

Indexicality and its functions in focus group research

Based on experimental uses of focus groups in social science research during the past decade or so, we have added another quasi-unique affordance to the list we generated in Kamberelis and Dimitriadis (2011). Because focus group conversations are highly indexical (tips of icebergs), it is often advisable to chase down data indexed but not fully developed in focus group sessions. In a study of consuming and producing art (Welker and Kamberelis, 2021), for example, focus group conversations indexed (or pointed toward) artwork in the research site (a museum), art talks, students’ written reflections on exhibits in the museum, students’ art journaling, student-produced artwork, whole-class discussions about each students’ artwork, and much more. We collected and analyzed these data in relation to focus group data to produce our findings (Fig. 18).

What are all the things that we might learn about people, places, and naïve theories that are in maps such as this one? Indexicality is a fundamental research principle and should guide every step of the process. Imagining indexical potentials can begin before data collection and analysis even begins. We will talk more about how indexicality tends to play out in particular data collection and analysis activities in the second part of the chapter.

Analyzing collective conversations

How to analyze data from collective conversations, as well as other data indexed in these conversations is an important decision each researcher must make for any particular study. Just about any of the many common (e.g., grounded theory) and less common (e.g., rhizomatic analysis—a predilection grounded in the quasi-unique affordances of focus groups as information gathering tools,

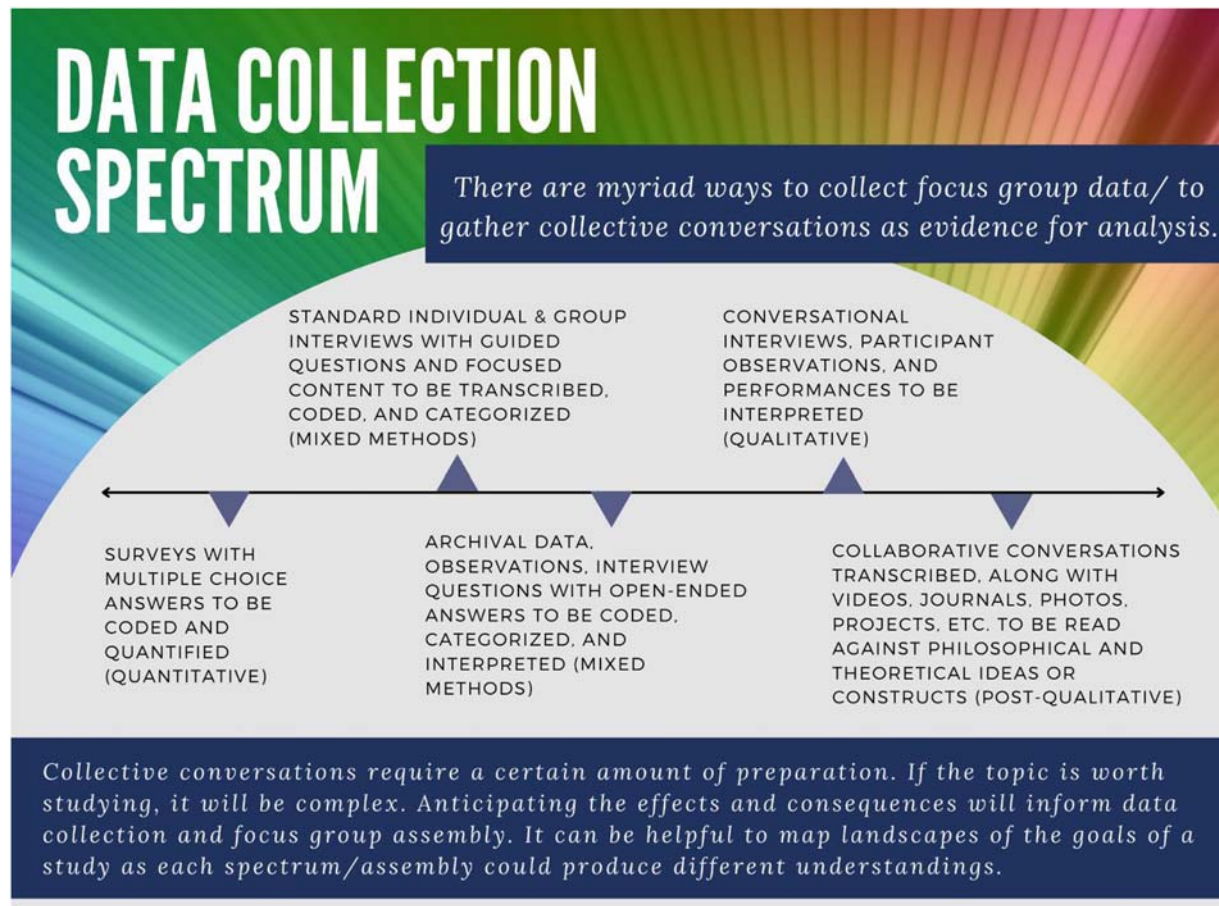


Fig. 17 Data collection strategies.

especially their highly indexical nature) qualitative data analysis strategies/systems may be used to analyze focus group data no matter what figured world a study is located within. However, just like certain theory-method assemblages are more coherent and make more sense than others, some figured world-data analysis assemblages seem more coherent, commensurable, and productive than others. In the remainder of this section, we will provide several examples and counterexamples of what we view as optimized assemblages. These examples will be glib; we will not unpack the various analysis strategies/systems; we discuss; we will, however, provide citations for resources we think are especially valuable/useful. That said, previous literature has focused on more positivist, post-positivist, and/or constructivist analysis strategies; however, we are excited about the unique and uncharted territories of more critical and post-qualitative analysis tactics used in relation to collective conversation data.

Focus groups as data have a “unique identity” that is not “naturally occurring” but can be orchestrated to be very close to naturally occurring by taking advantage of the quasi-unique affordances we discussed earlier in the chapter. This tends to occur less when studies are situated within traditional figured worlds and more when studies are situated within more critical/post-qualitative figured worlds (Fig. 19).

The potentials of research using focus groups have continued to expand considerably, especially during the last couple of decades. There are many ways to collect and analyze collective conversation data. We have already confessed our predilection for more critical, post-qualitative forms of data analysis such as situational analysis and rhizomatic analysis. Still, given our figured worlds heuristic, we encourage researchers to use whatever analytic tools and systems seem to fit best with the ontological/epistemological/axiological spaces they occupy and the kinds of data they collect. In the remainder of this chapter, we discuss very recent focus group research orientations and projects that seem to push beyond even the most critical/experimental types we have discussed so far and seem to foreshadow things to come in the ever-expanding world of focus group research.

Prospect: a social-cultural-political imaginal of focus group futures

Our understanding of the social-cultural-political imaginal futures of focus group research has been developed in conjunction with Denzin and Lincoln's (2018) tenth and eleventh moments of qualitative inquiry. The tenth moment (2010–2015) is characterized

INDEXICALITY

As educators, we find that supplementing collective conversations with additional forms of data allows us to understand more about individual participants as well as their interactions within the group as a whole.

Additional data often include:



Fig. 18 Indexicality.

as “the fractured, post-humanist present that battles managerialism in the audit-driven academy.” The eleventh moment (2016–) is characterized as “an uncertain, Utopian future where critical inquiry finds its voice in the public arena” (p. 9). We believe the new frontier of focus group research is located within these moments, which is important because decades of struggle by many scholars and practitioners have brought qualitative inquiry to a point where “politically informed action research, inquiry directed to praxis, and social change (Denzine and Lincoln, 2018, p. 890) are now deep and abiding commitments worth continuing to fight for. In relation to these commitments, scholars who use focus groups in their research have been among the key trailblazers and are likely to continue to be (Fig. 20).

Additionally, we utilized Deleuze and Guattari’s (1987) work to help illustrate the unique developments we encountered as we reviewed recent literature that included elements of collective conversations within their analyses, which resembled but were not defined specifically as focus groups. To work with and to better understand these similarities and differences, we leaned on Deleuze and Guattari’s (1987) key distinction between “tracing” and “mapping” that is instructive for understanding the “break” or “turn” we believe is happening with respect to using focus groups in contemporary social science research. Deleuze and Guattari espouse an ontology of becoming(s) rather than being. Reality is viewed as a continual process of flux or differentiation even though this fact is usually masked by powerful and pervasive illusory discourses of fixity, stability, and identity that have characterized most of western philosophy and theory since at least the Enlightenment. This ontology of becoming(s) enables (even urges) us to see things differently—in terms of what things might become rather than what they currently are. This ontology is characterized by its ability to engage productively with real movements of social change that open new forms of life both for individuals and for collectives. Based on this ontology, a “tracing” is a copy and operates according to “genetic” principles of reproduction based on an a priori deep structure and a faith in the discovery and representation of that structure. Tracings are typically based on phenomenological experience that is assumed to be essential, stable, and universal. Defined thus, the findings from most research projects are tracings. In this regard, our 2018 Handbook chapter, we also worked to trace focus group research through specifically defined representations; however, we found that we were in need of additional strategies as our understanding of terms and definitions continued to expand.

Deleuze and Guattari’s construct of “mapping” emerged as especially useful. In contrast to “tracing,” “mapping” charts open systems that are contingent, unpredictable, and productive:

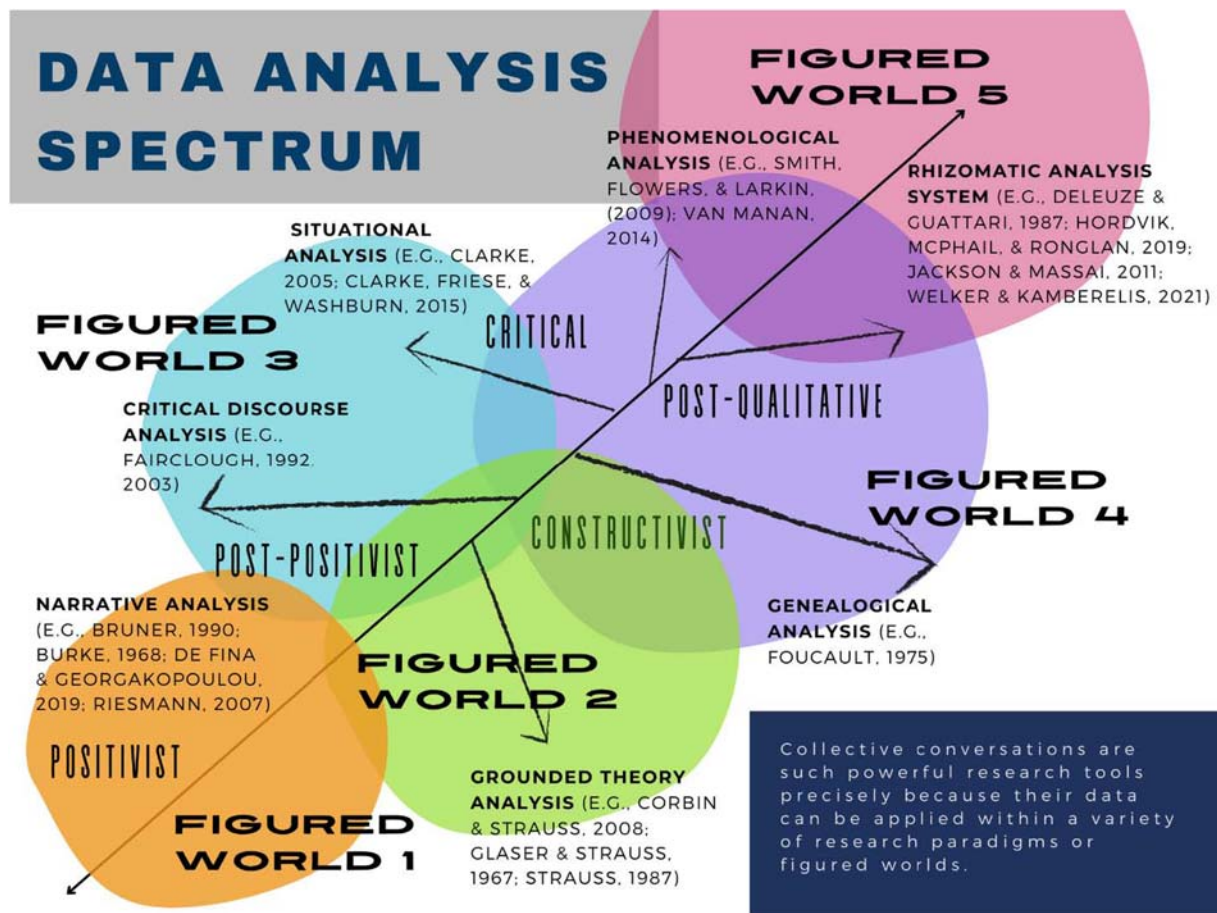


Fig. 19 Data analysis spectrum.

The map is open and connectable in all of its dimensions; it is detachable, reversible, susceptible to constant modification. It can be torn, reversed, adapted, to any kind of mounting, reworked by an individual, group, or social formation. A map has multiple entryways, as opposed to the tracing, which always "comes back to the same."

Deleuze and Guattari (1987, pp. 12–13).

"Mapping," when applied to focus groups conceived as collective conversations, allows for an adaptation of previous definitions and conceptions of focus group research, which we outline below.

Mapping alerts us to the need to look at discursive, social, and material formations in terms of their constitutive lines of force—their organization into lines of articulation and potential dissolution into lines of flight. This works against seduction by received models; it makes visible new possible organizations of reality; and it opens up new ways of organizing political resistance (praxis). Praxis-oriented research is research that offers up different ways of organizing reality so that new becoming(s) (both for individuals and for social formations) are possible.

In drawing maps, the researcher works at the surface, creating possible realities by producing new articulations of disparate phenomena and connecting the exteriority of objects to whatever forces or directions seem potentially related to them. As such, maps exceed both individual and collective experiences of what seems "naturally" real. Ultimately, mapping discloses potential organizations of reality rather than reproducing some prior organization of it.

"Mapping" the future of focus groups

So far in this chapter, we have engaged in "tracing"—sharing what is already known about the nature, functions, affordances, and uses of focus groups in social science research. In the remainder of this article, we shift to "mapping" focus groups futures. In this regard, we are not here to argue what focus groups are and what they are not. Nor are we here to claim that any belief/epistemology/discipline/paradigm/figured world, etc. is better than any other. Instead, we are here to share what we see as new lines of flight that

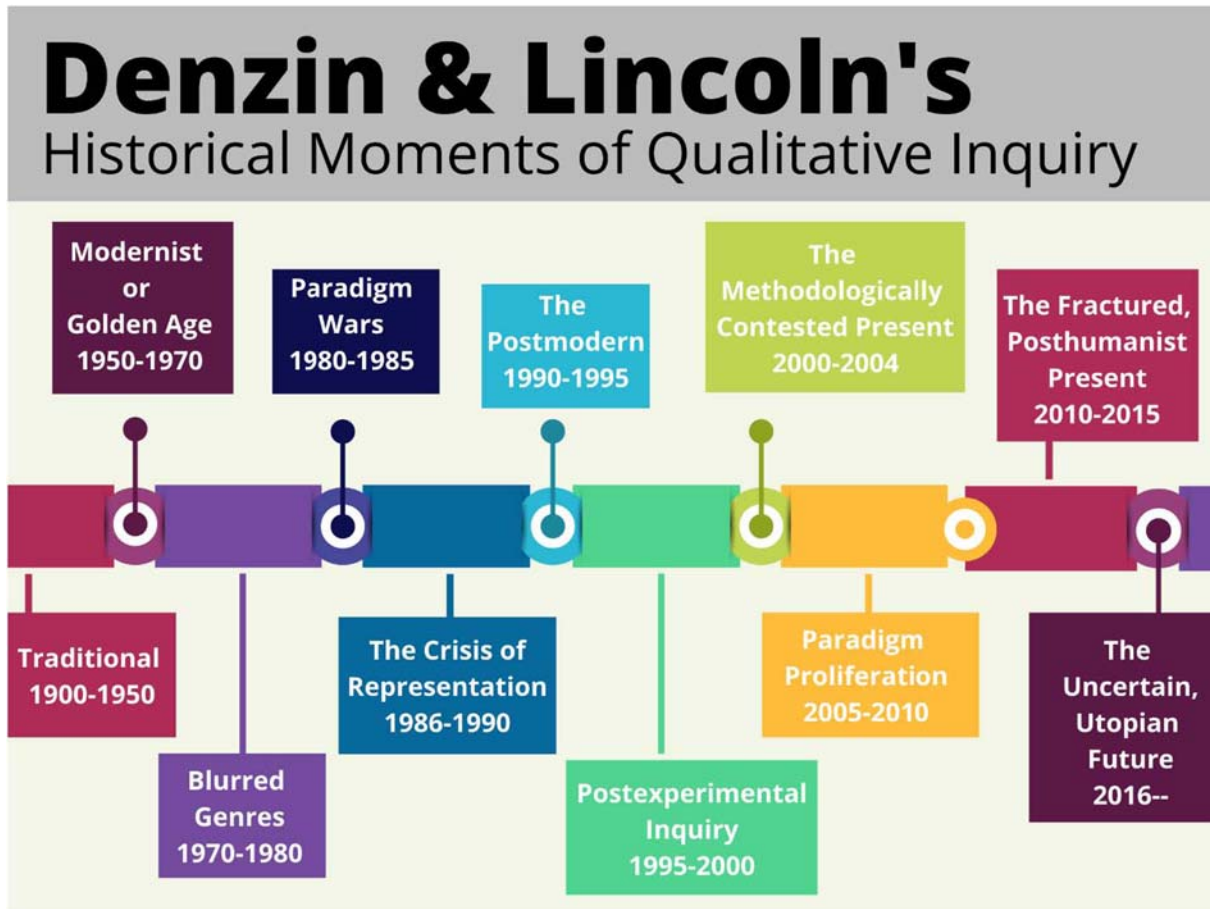


Fig. 20 Denzin & Lincoln's historical moments of qualitative inquiry.

can help us transverse/transcend the potentials of social science research in ways that play to Denzin and Lincoln's tenth and eleventh moment agendas.

Scholars using focus groups have previously tended to focus on the discursive or "talk data" produced by the transcripts of focus group conversations; however, we would like to redefine this discursive interaction to accommodate the subjective and performative intra-actions that create a renewed potential for asking and answering questions rendered all but invisible by traditional framings. Focus group research is now positioned to take us beyond patterned, rhythmic types of research (tracing in Deleuze & Guattari's lexicon) and into the unknown (mapping in Deleuze & Guattari's lexicon), which will shed light on how researchers might stretch how focus groups are conducted and what their purposes are. Such stretching includes cultivating more global conversations, especially given the power of such conversations to create confluences of peoples who may not know each other or have a shared agenda but what to come together to build new connections and new communities committed to co-constructing new knowledges with both scientific and political teeth. In other words, focus groups can be used to break free of the systemic, oppressive, and political failures of education (and other institutions and to discover/produce ways to think, act, and exist differently in the world ([Fig. 21](#))).

Many innovative, "post" focus group studies have been conducted during the past decade or so. We have selected three of them that have helped us expand our ideas about the potentials of focus group/collective conversation research ([Bass et al., 2020](#); [Kruger and Roux, 2017](#); [Yetiskin, 2020](#)). These studies showcase how focus groups are currently being used in innovative ways that utilize strategies and tools from the past while also integrating new and unique affordances that expand our previous notion of what it means to do focus group research. In this regard, some contemporary researchers don't even use the term "focus group," but, their work is clearly about collective, meaningful interactions among human beings around issues that are compelling to them ([Figs. 22 and 23](#)).

More importantly, these three studies provide exciting examples of the "becoming(s)" of focus group research that break from tradition to extend the questions asked and the meanings found within the evolving nature of educational research and social science research generally. Although the first and earliest study was conducted by educators within departments of education, the other two examples are more interdisciplinary, integrating ideas and practices from multiple social science and humanities disciplines. From these three studies (and other studies that have widened the parameters of focus group research), new affordances have been brought to light.

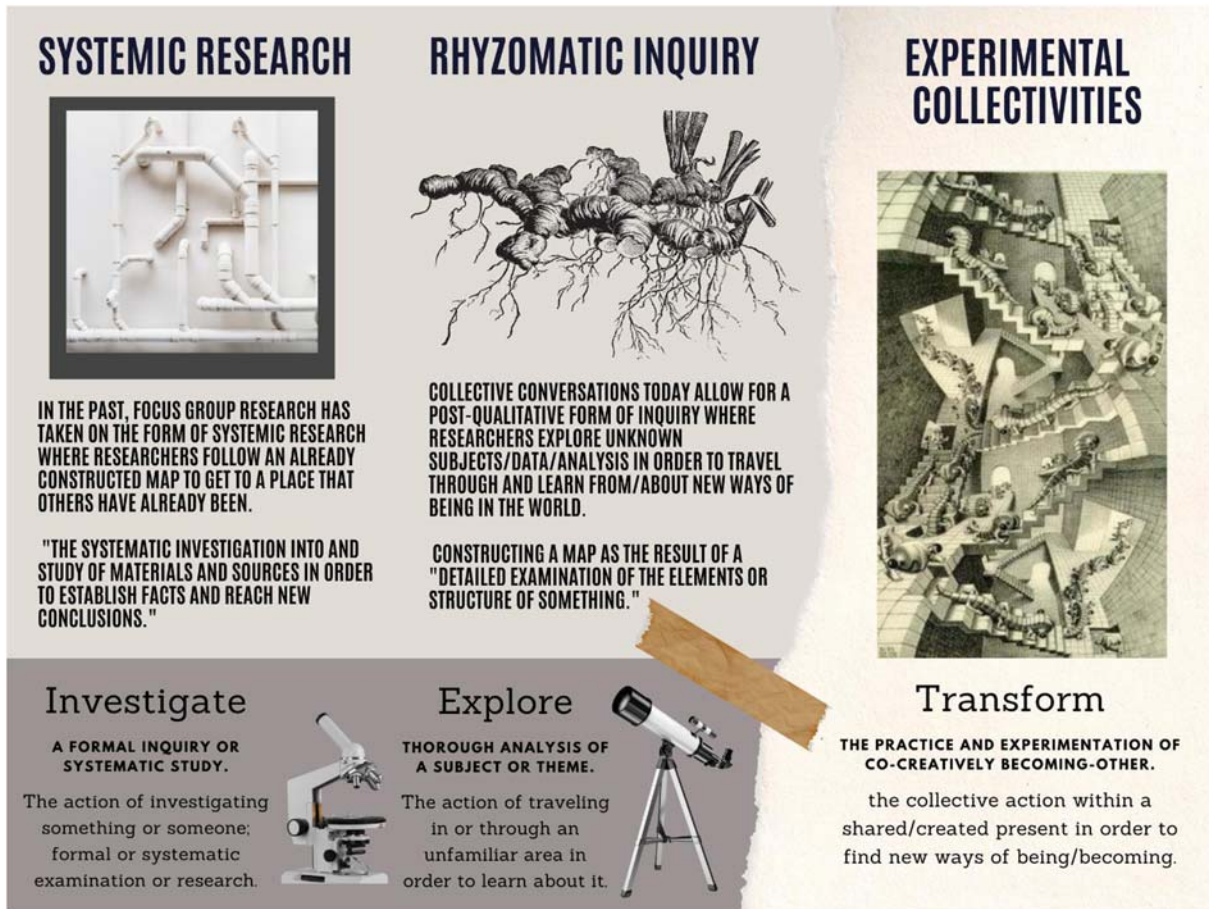


Fig. 21 Systemic research, rhizomatic inquiry, and experimental collectivities.

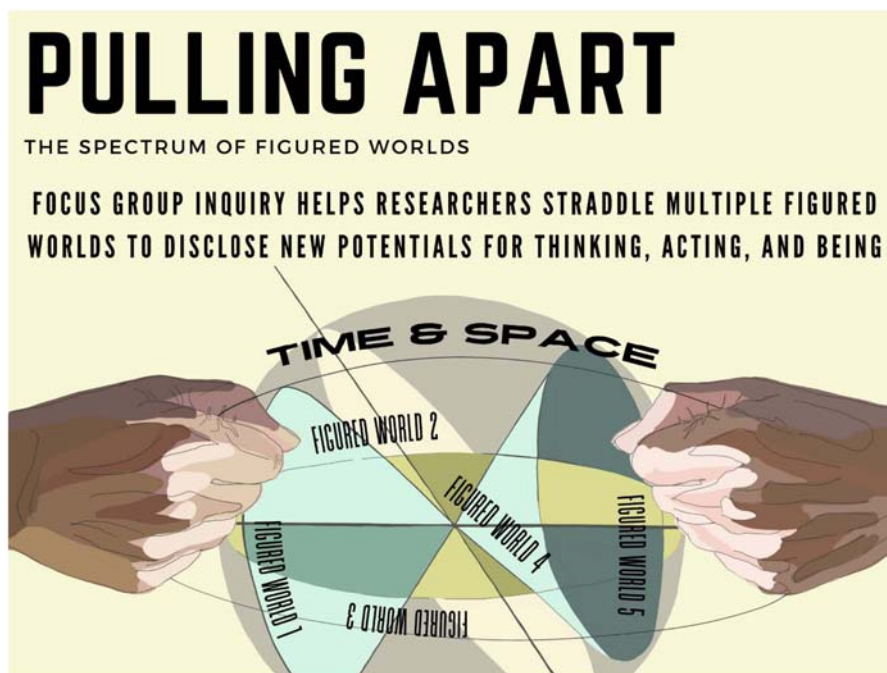


Fig. 22 Figured Worlds move across a new plane of time and space.

"FABULATION AS A PEDAGOGICAL POSSIBILITY: Working Towards a Politics of Affirmation"

Frans Kruger & Adré Le Roux



FINDINGS

Kruger & Le Roux's work is mostly theoretical. Although they discuss what this pedagogical shift has taught them, the researchers do not provide specific examples or empirical data. Additionally, it is important to note that Kruger and Le Roux do not use the term "focus group" in their work, but we find extensive parallels between fabulative conversations and focus groups oriented within post-qualitative worlds.



Kruger and Le Roux struggled to create social-justice-oriented pedagogies for their first-year, pre-service teachers. They observed that their past attempts at creating space for such conversations in the classroom often led to a perpetuation of a politics of negation.



The authors found that engaging students with fabulative conversations allowed them to experiment with a politics of affirmation, which provided opportunities for students to break away from intergenerational stories, hegemonic histories, and identity politics.



Further, these fabulative conversations were facilitated by instructors who intentionally devalued received truths and realities in favor of working collaboratively to create new perspectives and acknowledge the ever-present process of becoming.

Fig. 23 Frans Kruger & Andre Le Roux's research (study #1).

Brief summaries of the studies

Study 1

Kruger and Le Roux (2017) conducted a study in post-Apartheid South Africa with the aim of "enabling 1st-year pre-service [education] students to critically challenge their own identities in a diverse world by imagining the possibilities that exist for social responsiveness as agents of change" (p. 47). Kruger and Le Roux are clear about how important the historical context of their work was; their students held involuntary memories grounded in a hegemonic past. These memories varied depending on who was telling the stories of the past, what they had at stake, and whether the Apartheid was something that a person was harmed by or benefitted from.

Additionally, Kruger and Le Roux's student participants were living at a time when political and racial protests were occurring, such as the 2015 Rhodes Must Fall protest at the University of Cape Town. The national crisis indexed by such events prompted the need to create transformative spaces and thus redefine the moral responsibilities of university-level educators. The authors thus engaged in work to uncover opportunities for students and teachers "to break from the intergenerationally received stories and hegemonic histories through counter-actualizing regulation of difference through identity politics" and "to move from understanding social justice and pedagogical practices premised on a politics of identity and negation toward experimentation and a politics of affirmation" (Kruger and Le Roux, 2017, p. 49). Kruger and Le Roux's students were able to share their experiences, their stories, their families' stories, and stories they had heard from other students in the class who had both similar and different experiences and stories. This varied collection of stories created a wider, more expansive meta-narrative that was more inclusive and socially just. Key findings from this work are presented in the figure below:

In sum, Kruger and Le Roux staged collective conversations with a unique group of students in their education courses. The authors suggest that other educators can do this too, with similar and different topics of conversation depending upon the experiences, location, and historical moment each group exists within. Researchers collecting data in this way are not hoping for specific outcomes; instead, they are on the lookout for transformative potentials of data as they come into being.

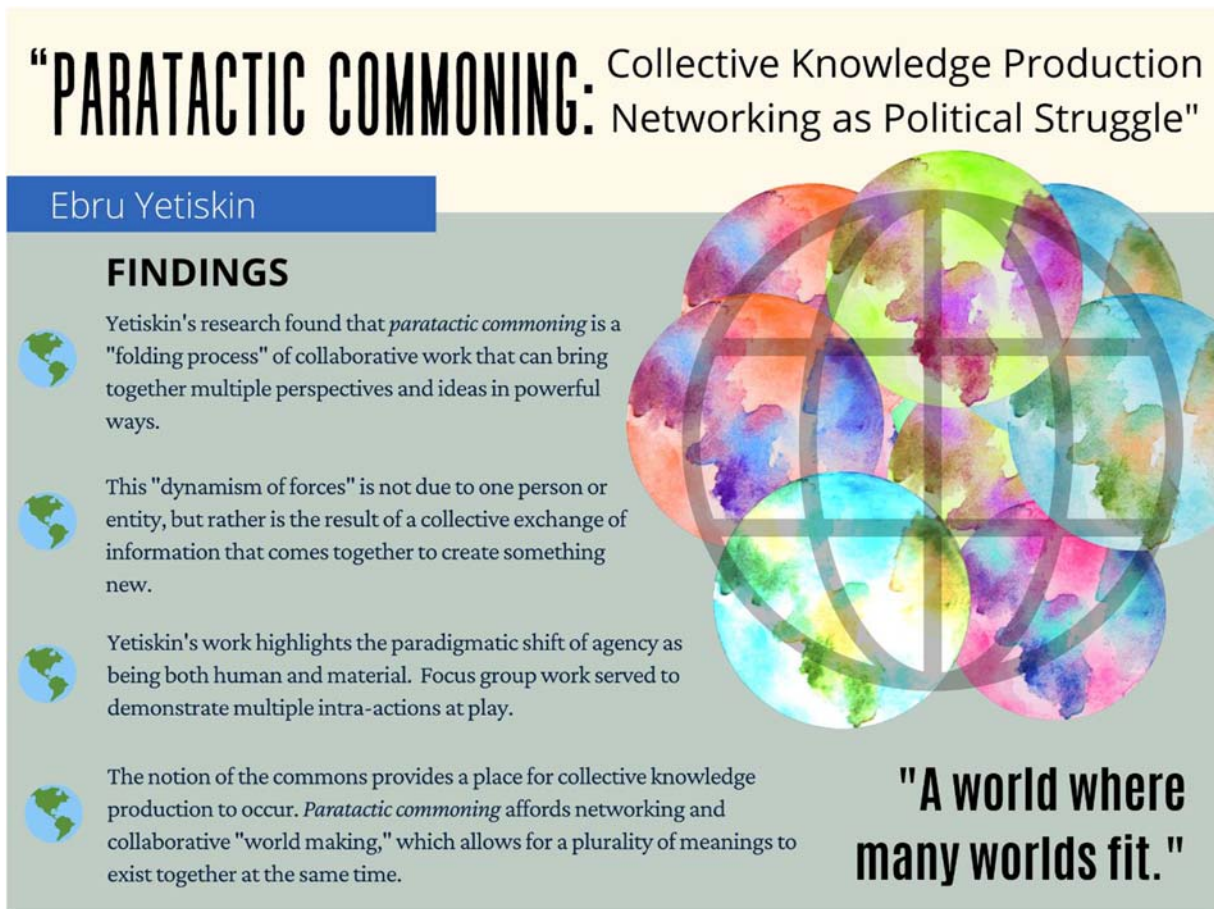


Fig. 24 Ebru Yetiskin's research (study #2).

Study 2

Yetiskin (2020) designed a study inspired by the civic data initiative: *Networks of Dispossession* (Adanali et al., 2013-present). This is an ongoing data-gathering initiative that works to collectively network and map exploitation of capital and power in Turkey. The construct that motivated this work is the "commons," which "are basic rights for accessing, using, and managing shared resources (i.e., data) in collaborative, distributed, and nonproprietary ways" (Yetiskin, 2020, p. 2). Using discourses about the Coronavirus pandemic as an example, Yetiskin demonstrated how the commons can be manipulated (even co-opted) by networks of the state and/or powerful corporations, which can lead to issues with transparency and accountability. In such cases, "power is ... centrally exerted by automating and regulating what *can* be accessed, transacted, processed, stored, understood and known in the public" (Yetiskin, 2020, p. 2). The primary purpose of Yetiskin's research was to learn how strategies of "paratactic commoning," may be used subversively to counteract this use of power. Paratactic commoning is a commoning practice where commoners collect and disperse public data as means for exploration, disclosure, and subversion to create and share alternate ways of knowing. Yetiskin uses focus groups to learn more about the potentials of paratactic commoning, which is similar to the general post-qualitative use of focus groups for exploring and redeploying tactics of power and control in ways that reveal and subvert such workings. "Thus, paratactic commoning can be practiced within proactive and subversive interventions of self-organized, self-reflexive, and self-aware networking assemblages that defend the rights to access, use and manage commons collectively and pluralistically" (Yetiskin, 2020, p. 2). In this sense, paratactic commoning becomes an on-going, online, asynchronous type of focus group where data may be extracted in invariable ways at various points across time for multiple research purposes by multiple researchers. The figure below was designed to outline some of the potential affordances of paratactic communing as a focus group variant researchers might enact in their current and future work (Fig. 24).

These subversive "commoning practices also take back control of redistributing power to take care of 'response-ability' of commons as well as commoners" (Yetiskin, 2020, p. 2). Driven by a commitment to global social change, Yetiskin pushed the usual limits of paratactic commoning further by distributing the ideas generated by her students on a larger scale (e.g., international conferences, peer-reviewed articles, etc.) These dimensions of her work "validated 'how' limited voices and partial sights in various parts of the world can be linked, mediated, and multiplied within skilled practices of commoners to co-produce situated knowledges and unexpected openings in contemporary political struggle" (Yetiskin, 2020, p. 5). Yetiskin imagines a world where local

groups engage in situational paratactic commoning as they work collectively to (re)distribute power, access, and information in more equitable ways. Further, Yetiskin hopes that small, paratactic commoning groups will work together with larger paratactic commoning groups to share information, help one another, and learn from each other, which would extend the global reach of such work.

Study 3

Bass et al. (2020) focused their research on migration scholars and the universities and other institutions that fund their work. These scholars noticed that faculty and students of humanities and philosophy viewed the “imperial eyes” of the academy as setting “the terms and limitations on interactions, locations, and relationality in research,” and they wondered whether this regime reduces “the agency of migrants” and produces “stratified configurations in the positionality of both migrants and researchers, and subsequently, exacerbate[s] [the] dynamics of exclusion and extraction” (p. 147). Influenced by scholars on the “frontlines” who resisted various manifestations of violence and exclusion, Bass, Córdoba, and Teunissen developed an approach to migration studies that “undermines the ongoing impact of colonialism and the normativity of institutionalized, hierarchical narratives that haunt academia” (p. 147). This study is particularly interesting in relation to Yetiskin’s research because the authors were working subversively in an attempt to “model a research design that takes tools from the research institution (e.g., funding, infrastructures, methodological frameworks) and uses them for liberatory purposes” (Bass et al., 2020, p. 148). Because these researchers received their funding from the European Union-funded Horizon 2020 Project, they were concerned their work might be used as “an instrument for migration management within the colonial context of European migration (Gutiérrez Rodríguez, 2018; Saucier and Woods, 2014)” (p. 148). To mitigate this threat, they “use[d] collective self-inquiry to undermine the academy’s imperial eyes, which produce stratified configurations in the positionality of migrants and researchers, while exacerbating hierarchies within the university” (Bass et al., 2020, p. 148). Collective self-inquiry is a matter of researchers with similar dispositions, interests, and commitments self-forming collective conversations both to critique/deconstruct their work and to refine their tactics for engaging in socially transformative work through shared dialog (Fig. 25).

Bass, Coroba, and Teunissen’s research demonstrates how focus group facilitators can fold and turn inside-out focus group activity so that researchers in collaboration with others studying both similar and different topics can become hyper-self-reflexive about their work. When using conversations for this purpose, we can see new affordances for focus group research—ones that expand the potentials for learning and being human in yet-to-be-imagined ways.

Focus groups as tools of subversion

Fabulative conversations, as described in Kruger and Le Roux’s (2017) work, are defined as imaginative spaces of creation and production. Such conversations allow participants to collectively invent their ideas of self and others as they acknowledge the continual becomings of individual and group identities. The authors explain that “group-subjects are transgressive because they subvert pre-established social relations, thus an identity based on involuntary memory and a hegemonic past” (Kruger and Le Roux, 2017, p. 53). These transgressive potentials are not possible without collective commitment and a facilitator who ontologically guides the participants through this work. Although Kruger and Le Roux do not call such fabulative conversations focus groups, their resemblance to focus groups can help us see how facilitators can guide the “ethico-onto-epistemology” (Barad, 2007) of a group to work past historical and social divides and attempt to create new realities.

Yetiskin’s (2020) work on potentials of paratactic commoning to interrupt and redirect flows of power and undermine strategies of prediction and control is another good example of how focus groups can be used as tools of subversion. Although Yetiskin does not refer to her work as focus group work, “paratactic commoning” bears a strong family resemblance to focus group research. This resemblance suggests that we may be helpful to expand our definitions of focus group work to include the various affordances and constraints that can be drawn via lines of flight across the spectrum of focused groups—collective conversations (Kamberelis and Dimitriadis, 2013) as tools for social transformation for increased diversity, equity, and inclusivity. Advocates of paratactic commoning argue that researchers must approach their topic as practice. This turn to practice can be used to challenge methodological symmetry when rethinking commoning across local settings perceived to be similar. Yet, ideas generated in one setting would not necessarily be generalized to others (except perhaps analogically), and such ideas would not preempt the analysis of how commoning is practiced differently in different socio-cultural settings. Paratactic commoning is thus a matter of co-producing “situated knowledges” (Haraway, 1988) and thus lived reality itself (Yetiskin, 2020, pp. 2–3).

What Yetiskin calls “tactical control of the commons” is similar to early focus group work such as Merton and Lazarsfeld (1968), where data were used to make predictions about citizens’ feelings toward war efforts. This research was intended to help researchers learn how best to persuade the masses to take a particular stance in relation to certain efforts related to the war. This use of power to influence public opinion is a type of control clearly has ethical implications that, taken to their extreme, result in fascism. Yetiskin’s paratactic commoning is one way of using focus groups to disclose and subvert this type of control.

Bass et al. (2020) highlighted the potentials of focus groups as tools of subversion as well. Although funded by the European Commission to conduct research on “migrants and their movements,” this group of doctoral students wanted to push against “imperial eyes” and the sedimented power of the academy. To do so, they appropriated and redeployed tools from the academy and related organizations (e.g., funding, infrastructures, methodological frameworks) for liberatory purposes—to undermine the academy’s imperial eyes (Fig. 26).

"(RE)SEARCHING WITH IMPERIAL EYES:

Collective Self-Inquiry as a Tool for Transformative Migration Studies"



Madeline Bass, Daniel Córdoba, & Peter Teunissen

FINDINGS

Using collective self-inquiry as a strategy, allowed these three researchers to look at their separate research projects reflexively and critically to address areas in need of improvement. From these reflections, researchers were able to see that, although their goals were clear, their actions needed to be further developed to push back against hidden imperial tendencies.



- Córdoba shared experiences of squatting with migrants as both an activist and later as a researcher in Barcelona. Córdoba collaborated with non-academics who expressed their distrust of the academy, which they viewed as serving capitalist agendas that were bound by domination and inequality.
- Córdoba's research aimed to create "beneficial and mutually enriching relationships" between the researcher and participants. Although Córdoba was unable to move the research beyond accepting differences and sharing storied experiences at the time, collective self-inquiry with Bass and Teunissen allowed for the generation of ideas for future transformative research.



- Teunissen's research involved traveling to the Greek/Turkish borderland in 2018 to demonstrate solidarity and support for the 'migrant/refugees' who were being detained. At the time, Teunissen was working with a group of volunteer activists and scholars to document the violence and deportation of Turkish citizens.
- Through this (re)searching and collective self-inquiry project, Teunissen was able to reflect on this past work to see that even though the goal of fighting against the violence at the border had been fulfilled, the research lacked a personal relationship with the migrants themselves.
- In the micro-focus groups that followed, Teunissen was able to see that forming personal relationships would have allowed the research to humanize migrants and refugees in ways that most research does not, as migrants are often rendered objects of research rather than subjects.



- In 2016, Bass taught 9th grade while completing graduate studies in sociology in a small village in Ethiopia. Although this was not an official research project, Bass was placed in a position of privilege when riding in a taxi with a local. Having both been removed from the same village due to violence in the area, Bass was able to work with, care for, and cry alongside this community member in a way that researchers should strive to do with their participants.
- Through re(search) reflections with Córdoba and Teunissen, Bass acknowledged that sharing feelings and stories with participants is an important first step, but never the last step in participatory action research.
- Action potentials such as "buying back an acre of settler land" could be enacted to build more socially just realities from these storied experiences.

Fig. 25 Bass, Coroba, & Teunissen's research (study #3).

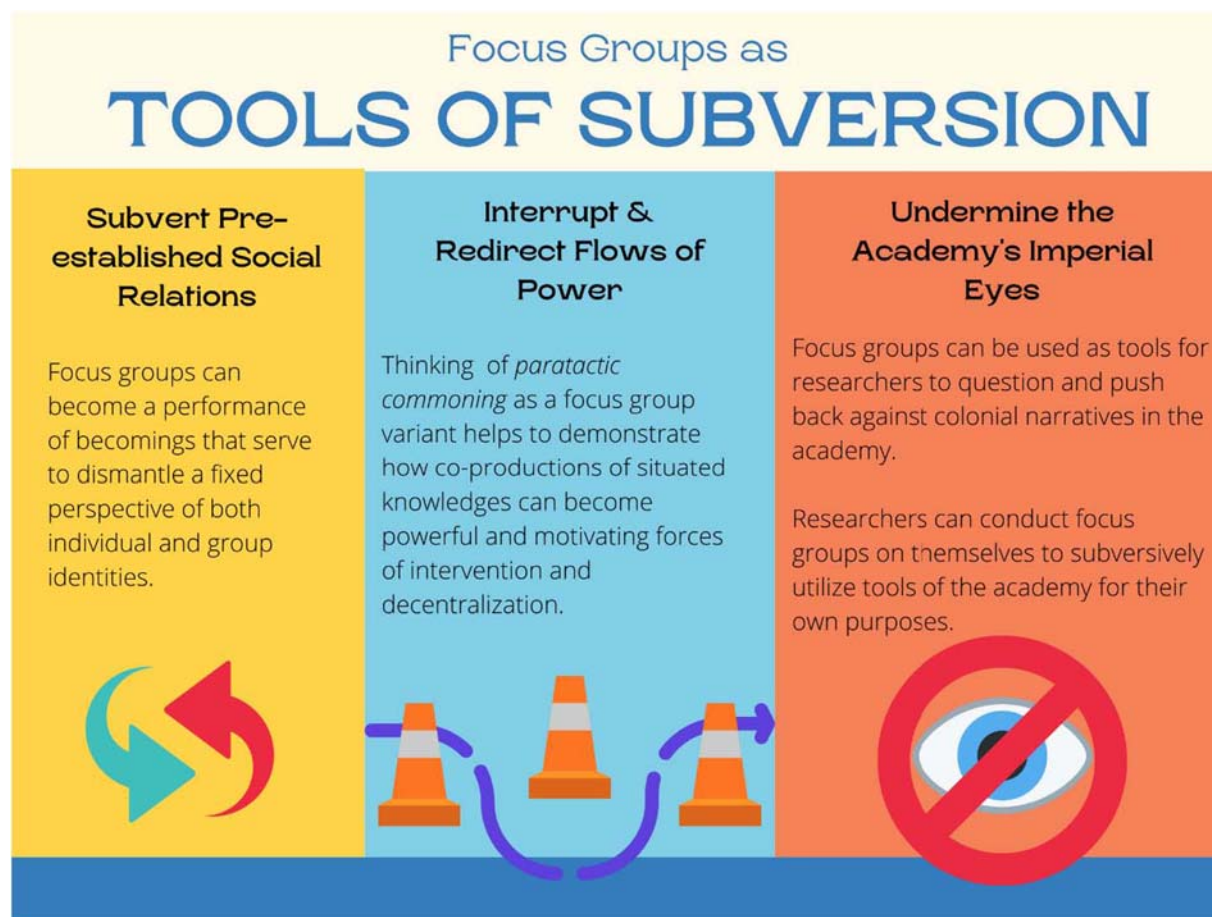


Fig. 26 Focus groups as tools of subversion.

Widening the definition of focus groups: fabulative conversations, paratactic commoning, and collective self-inquiry

Drawing on autoethnographic, duoethnographic, and multivocal ethnographic research practices, Kruger and Le Roux (2017) developed a tactic they adapted from Deleuze and Guattari that they call "pedagogical fabulation" to engage students in fabulative conversations to "provide the opportunity to disentangle from the 'gravitational pull of established identities and practices that unnecessarily limit the question of how life might be composed'" (Kruger and Le Roux, 2017, p. 49). The affordances of fabulation allow students to imagine futures that do not yet exist in a way one might write a fictional story, pushing the normative concept of the "real" into the background to allow collectives "to experiment with the emergence of different future-orientated socio-political collectivities that a different way of thinking about our pedagogical practices could make possible" (Kruger and Le Roux, 2017, p. 50).

Fabulative conversations create spaces for reassessing and negotiating the way things currently are and the various ways in which reality is always changing and becoming something new. "Transformative potential of social justice pedagogies infused with fabulation subsequently lies with the collective action of becoming-other through the creation of a shared present" (Kruger and Le Roux, 2017, p. 57). Fabulative conversations bring together unique groups of learners who may not otherwise have interacted with each other. Their repetitive sharing of experience presents unfathomable opportunities that would not otherwise have occurred.

Kruger and Le Roux feel strongly about the nature and place of identity politics in the classroom and "how the centrality of identity politics ... not only perpetuates the notion of identity as being stable, self-contained and constituted through the process of negotiation, but that it feeds into a continued positioning of ... students on different sides of historical and social divides" (Kruger and Le Roux, 2017, p. 53). These scholars also believe that fabulation in the classroom can allow for greater creativity regarding social justice pedagogies that aim to enable students see beyond polarities to construct new perspectives and possible becomings (Fig. 27).

In sum, Kruger and Le Roux's main purpose of bringing fabulation into the classroom was to help their students (and themselves) imagine and create worlds that had not previously existed. Although the authors acknowledge that "such a future cannot be predicted nor its justness assured" (p. 51), they still feel that it is worth trying and has the potential to create a better world than the one in which they/we currently live.



Fig. 27 Working toward a politics of affirmation.

Yetiskin's (2020) theory of paratactic commoning works to expand the notion of focus group research and the myriad intractions that constitute data when social collectives are involved. Traditionally, focus group research has been conducted locally and synchronously, but Yetiskin's concept of paratactic commoning allows us to think of focus groups as collaborating across time, across space (both physical and virtual), and without a designated facilitator. These groups still share interests and stakes as they work to expose flaws, create connections, and construct new knowledge that can help themselves and others (Fig. 28).

Occupying researcher, participant, and facilitator roles simultaneously, Bass, Córdoba, and Teunissen's "collective self-inquiry" focus group variant involves conversations about migratory experiences with the goal of exploring "the academy's potential for liberatory practices through the development of alternative methodologies which emphasize community, affinity, and shared knowledge" (p. 149), thus "challenging the normativity of institutionalized, colonial narratives, and the 'b/ordering' of movements(s) of people across former empires" (Bass et al., 2020, p. 148).

This "messy" work "serves to clarify not just the end result of collective self-inquiry as a practice, but also, in a very practical sense, to elucidate the ways in which the process was carried out and the passionate, sometimes messy backstory behind collective work" (Bass et al., 2020, p. 149). All in all, this work foregrounds the academy's potential for liberatory practices through the development of alternative methodologies that emphasize community, affinity, and shared knowledge (Fig. 29).

Using focus groups to promote and sustain global conversations

Kruger and Le Roux (2017) worked with their students to engage in both local and global conversations, and their examples of classroom pedagogies can benefit educators and focus group researchers alike. Additionally, their work can serve as an example for anyone looking to have productive and critical conversations across lines of difference and polarization. Kruger and Le Roux explain that to have such difficult conversations, a facilitator with a Deleuzian/post-qualitative ontological perspective can help to keep the conversation on the path of the imaginative and to avoid received truths or returning to places participants have already been. "Such a resistance, which is contingent and specific, is a hope for a better future" wherein "the counteracting of preestablished truths occurs through the invention of a collectivity that does not yet exist" (Kruger and Roux, 2017, p. 51).

Among other things, Yetiskin extends these ideas in the interest of generating a global politics of knowledge production and the basic rights to access and management of shared resources. In this regard, recall that paratactic commoning is a commoning practice that uses publicly available data collected by commoners to "explore, reuse, reveal and subvert" tactically obfuscated facts by creating alternate visions and sharing them globally through Internet outlets, media outlets, conference presentation outlets, and publication outlets that are part and parcel of contemporary network cultures. Thus, paratactic commoning can be practiced by "self-organized, self-reflexive, and self-aware networking assemblages" with proactive and subversive intentions to defend the rights to access as well as use and manage the commons collectively and pluralistically. Yetiskin hopes this work can lead to new



Fig. 28 "Post" focus groups vs. paratactic commoning.

knowledge and explorations through producing new types of evidence, which can lead to fact-checking across a variety of global resources and references.

Bass et al. (2020) research was also designed to promote social justice agendas globally. Recall that they recruited research participants and gathered data gathered from three different locations: Barcelona, Ethiopia, and the Greece/Turkey border. With widely



Fig. 29 (Re)searching: collective self-inquiry.



Fig. 30 Using focus groups to promote & sustain global conversations.

varying lived experiences, these researchers also conducted a meta-analysis of their past studies. This work demonstrates how researchers across the globe can engage together in collective conversations, collective struggle, collective problem solving, and collective activism. Additionally, because these authors chose to publish under a Creative Commons International License, not only did they engage in global conversations, but they did so in front of a global audience, which increased access to and extended the reach of their work (Fig. 30).

Final thoughts: focus groups as tools for interdisciplinary, critical, global explorations

We began this chapter promising a more prospective (rather than a more retrospective) account of how focus groups are/can be used in educational research. We think we delivered on that promise. That said, we want to reiterate that focus groups are amazingly versatile information gathering tools that may be deployed productively across the entire figured worlds spectrum we described, as well as beyond the boundaries of that spectrum, which we tried to demonstrate in the second half of this chapter. In this regard, even though we are inclined to privilege more experimental, transformative uses of focus groups, we in no way believe that others should share that inclination. Focus groups remain powerful data collection tools for all kinds of research—from studies grounded in positivist epistemologies to studies grounded in post-qualitative ontologies/epistemologies to whatever comes our way next.

In the remainder of this chapter, we would like to offer some final (mostly forward-looking) thoughts on what we see as current frontiers of focus group research, especially in relation to the new affordances we discussed. Indexicality is especially important in this regard because it underscores the power of integrating focus group findings with other kinds of findings (e.g., field notes or other means for capturing observations, individual interviews, artifacts of various kinds, archival data) to produce far richer, more complex, nuanced understandings of phenomena under study.

The other affordances we proposed here—tools of subversion, spaces for collective struggle, and promoters of global conversations—align beautifully with the commitments of Denzin and Lincoln's tenth and eleventh moments of qualitative inquiry. Recall that the tenth moment is post-humanist and committed to challenging the hierarchical power relations of the academy and beyond to imagine and produce future-orientated socio-political collectivities designed for social justice. Also recall that eleventh moment is

characterized as “an uncertain, utopian future where critical inquiry finds its voice in the public arena” (Denzin and Lincoln, 2018, p.9). Thus, Yetiskin’s “paratactic communing” Bass, Córdoba, and Teunisson’s “collective self-inquiry,” and Kruger and Le Roux’s “pedagogical fabulation” are all tactics that embody the commitments of the tenth and eleventh moments.

Paying attention to the language used by contemporary scholars who are pushing the boundaries of focus group research, three constructs emerge as crucial, especially in terms of moving from engaging in research that privileges “mapping” over “tracing,” all the while acknowledging that all research usually deploys both to some extent. These three constructs are “intersect,” “transverse,” and “transcend,” all of which are central to thinking about post-qualitative modes of research inspired by theorists like Deleuze and Guattari (1987) and Barad (2007). Additionally, they are central commitments of much contemporary qualitative research (and especially post-qualitative research) designed to produce social-cultural-economic-political realities that are more diverse, equitable, and socially just on both local and global scales (Figs. 31 and 32).

In this regard, foregrounding affordances (especially the four new ones we introduced) was strategic in relation to these constructs. Affordances usually co-exist in all research projects and how they interact with each other is usually quite complex—intersecting, transversing, and mutually reinforcing each other in ways that are often transcendent. Additionally, foregrounding the notion of affordances underscores the expansive and ecumenical definition of focus groups that we have tried to develop throughout this chapter. Whereas most discussions of focus groups have, in effect, defined them a priori, we have done something different. We have emphasized neither the “typical” form or size of focus groups nor the procedures for conducting them. Instead, we have emphasized their potentials: (a) how they can serve pedagogical, political, and inquiry functions, as well as combinations of these functions; (b) their potentials for unearthing rich data sets that most other data collection strategies do not afford (or afford as well), and (c) how they can transform the lives of participants and the environments in which they live and work. Finally, based on our review of the literature and our own research experience, we are confident that the affordances we have discussed here emerge much more often and in more powerful ways in focus groups compared to most (perhaps all) other data collection tools. Why focus groups draw out these affordances in unique and uniquely powerful ways seems related to how performative they are and how the performances they spawn recapitulate naturally occurring social activity.

We have also drawn attention to exciting theory and research produced in recent years that suggests reimagining focus group research looks something like this: Non-positivist theories of truth reinforce the idea that the march or science is always a tentative march. Thus, how we understand the nature and functions of focus groups is always changing. In 5 or 10 years from now we will understand them yet differently, maybe with more complexity or nuance, maybe within a new and unforeseen paradigm shift.

Through our examples and figures, we have tried to show that focus groups (by any name) are being used in exciting, transformative ways across the globe and that these uses are opening up new potentials for what research means, how we do it, and what effects it may have. In other words, we inhabit a historical moment when Deleuze and Guattari’s call for “mapping” rather than “tracing” resonates more powerfully than ever. This is brand-new territory; we are not sure how to traverse its topological landscape; we do not know what we will find around the next turn; but we remain committed to mapping it as best we can. As fellow explorers, we offer the following advice to you:

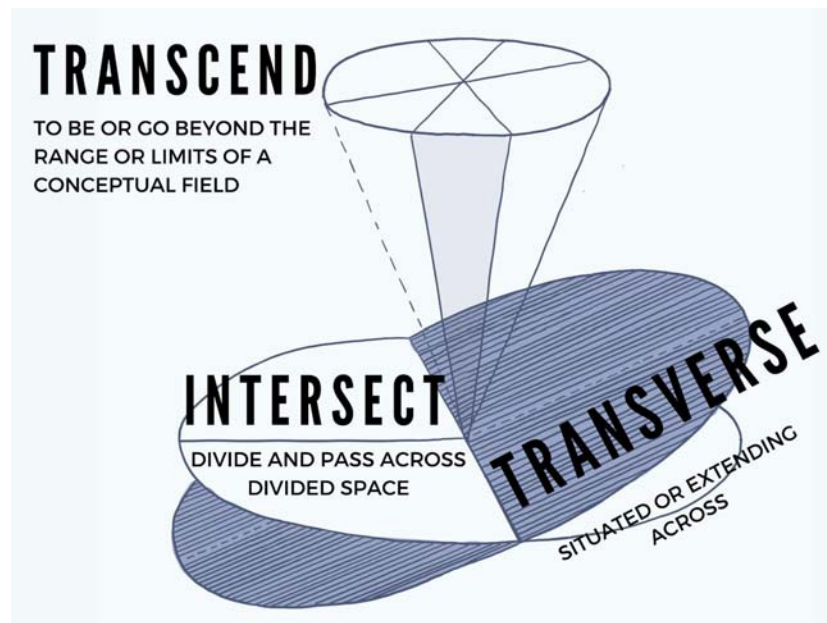


Fig. 31 Intersect, transverse, & transcend.

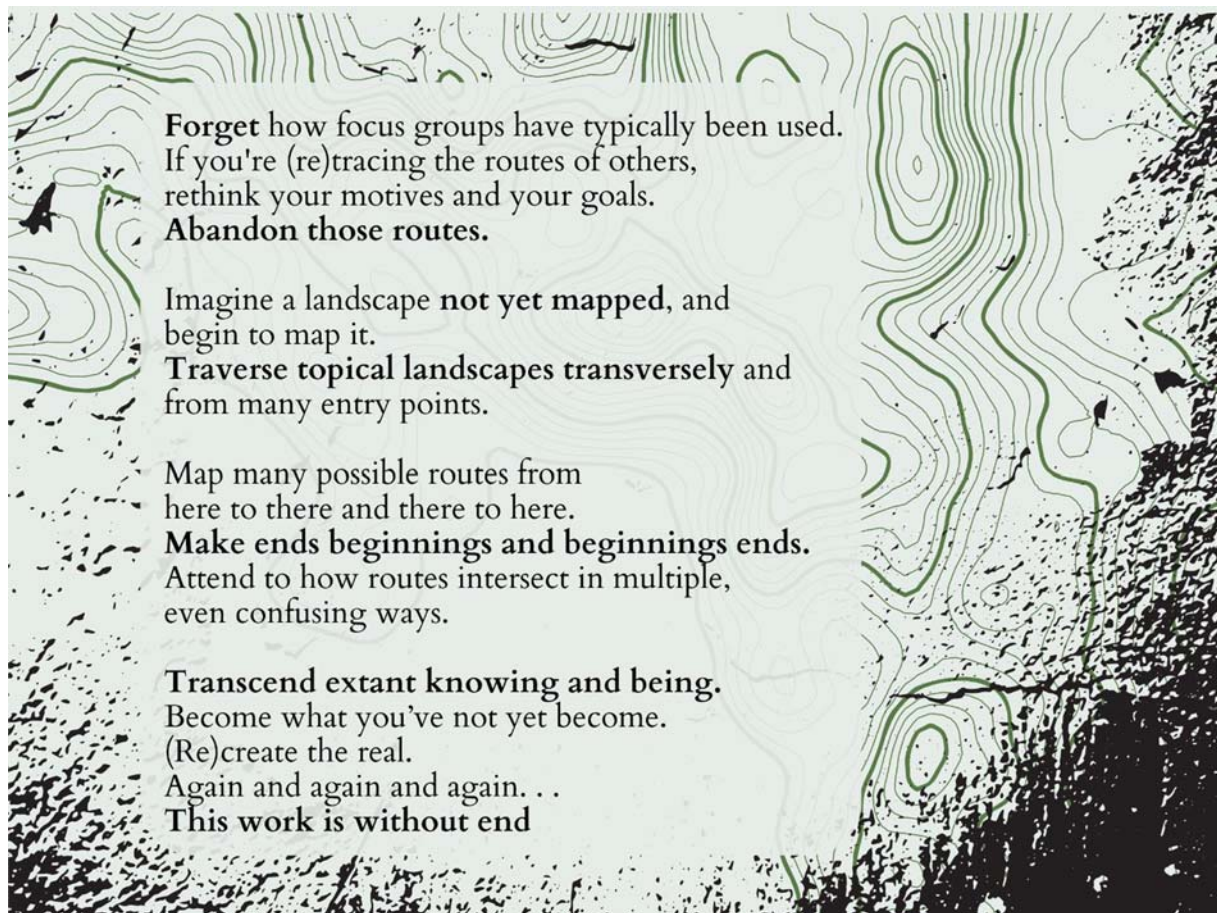


Fig. 32 Advice to fellow explorers.

References

- Adanalı, Y.A., Aldatmaz, A., Arıkan, B., Ince, E., Gurakar, E., Ustun, Z., Zıngıl, O., et al., 2013. Networks of Dispossession: Collective Data Compiling and Mapping on the Relations of Capital and Power in Turkey. Retrieved from: <http://mulksuzlestirme.org/index.en/>.
- Bakhtin, M.M., 1981. In: Holquist, M., Emerson, C., Holquist, M. (Eds.), *The Dialogical Imagination*. University of Texas Press, Austin, TX.
- Barad, K., 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, Durham, NC.
- Bass, M.J., Córdoba, D., Teunissen, P., 2020. (Re)searching with imperial eyes: collective self-inquiry as a tool for transformative migration studies. *Soc. Incl.* 8 (4), 147–156. <https://doi.org/10.17645/si.v8i4.3363>.
- Behar, R., 1993. *Translated Woman: Crossing the Border With Esperanza's Story*. Beacon Press, Boston.
- Collins, D., 2007. When sex work isn't "work": hospitality, gay life, and the production of desiring labor. *Tour. Stud.* 7 (2), 115–139. <https://doi.org/10.1177/1468797607083498>.
- Deleuze, G., Guattari, F., 1987. *A Thousand Plateaus* (B. Massumi, Trans.). University of Minnesota Press, Minneapolis.
- Denzin, N.K., Lincoln, Y.S., 2018. Introduction: the discipline and practice of qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, fifth ed. Sage Publications, Thousand Oaks, CA, pp. 1–26.
- Dill, B.T., 1994. Fictive kin, paper sons, and compadrazgo: women of color and the struggle for family survival. In: Zinn, M.B., Dill, B.T. (Eds.), *Women of Color in U.S. Society*. Temple University Press, Philadelphia, pp. 149–169.
- Dimitriadis, G., 2016. Reading qualitative inquiry through critical pedagogy: some reflections. *Int. Rev. Qual. Res.* 9 (2), 140–146. <https://doi.org/10.1525/irqr.2016.9.2.140>.
- Duneier, M., 1992. *Slim's Table: Race, Respectability, and Masculinity*. University of Chicago Press, Chicago.
- Espiritu, Y.L., 1997. *Asian Women and Men: Labor, Laws, and Love*. Sage Publications, Thousand Oaks, CA.
- Fetene, G.T., 2009. *Self-Reported Sexual Experiences, Sexual Conduct and Safer-Sex Practices of Ethiopian Undergraduate Male and Female Students in the Context of HIV/AIDS Pandemic*. Unpublished Ph.D. Dissertation. State University of New York at Buffalo.
- Fine, M., Weis, L., 1998. *The Unknown City: Lives of Poor and Working-Class Young Adults*. Beacon Press, Boston.
- Fine, M., Roberts, R.A., Torre, M., 2004. *Echoes of Brown*. Teachers College Press, New York.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison* (A. M. Sheridan-Smith, Trans.). Penguin, Harmondsworth, U.K.
- Foucault, M., 1988. *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*. Random House, New York.
- Freire, P., 1993. *Pedagogy of the Oppressed*. Continuum, New York (Original work published 1970).
- Gramsci, A., 1971. *The Prison Notebooks* (Q. Hoare & G. N. Smith, Eds. & Trans.). Lawrence & Wishart, London.
- Gutiérrez Rodríguez, E., 2018. The coloniality of migration and the "refugee crisis": on the asylum—migration nexus, the transatlantic white European settler colonialism-migration and racial capitalism. *Refuge* 34 (1), 16–28.
- Habermas, J., 1971. *Knowledge and Human Interests* (J. J. Shapiro, Trans.). Beacon Press, Boston.
- Heidegger, M., 1962. *Being and Time* (J. Macquarrie & E. Robinson, Trans.). Harper, San Francisco (Original work published in 1927).

- Haraway, Donna., 1988. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-600.
- Hill, M.L., 2009. *Beats, Rhymes, and Classroom Life: Hip-Hop Pedagogy and the Politics of Identity*. Teachers College Press, New York.
- Holland, D., Lachicotte, W.S., William, S., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge, MA.
- Kamberelis, G., Dimitriadis, G., 2005. On Qualitative Inquiry: Approaches to Language and Literacy Research. Teachers College Press, New York.
- Kamberelis, G., Dimitriadis, G., 2011. Focus groups: contingent articulations of pedagogy, politics, and inquiry. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, fourth ed. Sage Publications, Thousand Oaks, CA, pp. 545-561.
- Kamberelis, G., & Dimitriadis, G., 2013. Focus groups: From structured interviews to collective conversations. Routledge.
- Kamberelis, G., Dimitriadis, G., Welker, A., 2018. Focus group research and/in figured worlds. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, fifth ed. Sage Publications, Thousand Oaks, CA, pp. 692-716.
- Kruger, F., Le Roux, A., 2017. Fabulation as a pedagogical possibility: working towards a politics of affirmation. *Educ. Change* 21 (2), 45-61. <https://doi.org/10.14426/cristal.v8i2.267>.
- Lather, P., Smithies, C., 1997. *Troubling the Angels: Women Living With HIV/AIDS*. Westview Press, Boulder, CO.
- Latour, B., 2007. Turning around politics: a note on Gerard de Vries' paper. *Soc. Stud. Sci.* 37 (5), 811-820. <https://doi.org/10.1177/0306312707081222>.
- Lee, R.M., 2010. The secret life of focus groups: Robert Merton and the diffusion of a research method. *Am. Sociol.* 41 (2), 115-141. <https://doi.org/10.1007/s12108-010-9090-1>.
- Madriz, E., 2000. Focus groups in feminist research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, second ed. Sage Publications, Thousand Oaks, CA, pp. 835-850.
- Martin, J., 1995. The style and structure of cultures in organizations: three perspectives. *Organ. Sci.* 6 (2), 230-232. <https://doi.org/10.1287/orsc.6.2.230>.
- Merton, R.K., Lazarsfeld, P.F., 1968. *Studies in Radio and Film Propaganda*. In: Merton, R.K. (Ed.), *Social Theory and Social Structure*. Free Press, New York, pp. 563-582 (Original work published in 1943).
- Morgan, D.L., Hoffman, K., 2004. Focus Groups. *Approaches to Qualitative Research: A Reader on Theory and Practice*, pp. 263-285. <https://doi.org/10.4135/9781526416070.n16>.
- Morley, D., 1980. *The Nationwide Audience*. British Film Institute, London.
- Radway, J., 1984. *Reading the Romance: Women, Patriarchy, and Popular Literature*. University of North Carolina Press, Chapel Hill, NC.
- Ranciere, J., 1991. *The Ignorant Schoolmaster: Five Lessons in Intellectual Emancipation*. Stanford University Press, Stanford, CA.
- Ricoeur, P., 1992. Hermeneutics and the critique of ideology (K. Blarney & J. Thompson, Trans., pp. 270-307). In: *From Text to Action: Essays in Hermeneutics*, II. Northwestern University Press, Evanston, IL.
- Saucier, P.K., Woods, T.P., 2014. Ex aqua: the Mediterranean basin, Africans on the move and the politics of policing. *Theoria* 61 (141), 55-75. <https://doi.org/10.3167/th.2014.6114104>.
- Taylor, J.G., 1979. *From Modernization to Modes of Production: A Critique of the Sociologies of Development and Underdevelopment*. Springer, New York.
- Tuck, E., 2012. Repatriating the GED: Urban youth and the alternative to a high school diploma. *The High School Journal*, 4-18. <https://doi.org/10.1353/hsj.2012.0009>.
- Weis, L., 2005. *Class Reunion: The Remaking of the American White Working Class*. Routledge, New York.
- Welker, A., Kamberelis, G., 2021. Mapping trajectories: analyzing focus group data rhizomatically. In: Vanover, C., Saldana, J., Mihaus, P. (Eds.), *Analyzing and Interpreting Qualitative Data: After the Interview*. Sage Publications, Thousand Oaks, CA, pp. 183-206.
- Wittgenstein, L., 1953. *Philosophical Investigations*. MacMillan, New York.
- Yetiskin, E., 2020. Paratactic commoning: collective knowledge production networking as political struggle. *Popular Commun.* 18 (3), 216-232. <https://doi.org/10.1080/15405702.2020.1781860>.

Naturalistic observation

J. Schostak, Manchester Metropolitan University, Manchester, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Observation	250
Orienting	252
Participant observation	253
Joining	254
Ethics, dramatis personae, locations, and scenes of action	255
The familiar and the strange	255
Fieldwork as a basis for validity, generalization, reliability, and objectivity	257
Truth out of place	258
The significance of naturalistic observation	258
References	259

What is naturalistic about “naturalistic observation”? Although a method of social inquiry, located within contested academic discourses, it is also something we do all the time, in our everyday lives, without necessarily remarking upon it unless, of course, we are caught and accused of being “nosey”. “Are you looking at me?” can in certain situations be not so much a question as a threat; or, perhaps a seductive provocation. We have, in many circumstances to “keep our eyes open”, “have our wits about us” to avoid dangers. Or, simply, to “keep aware” to know what is going on, in order to participate, to contribute our views to a situation, to recognize others and be a part of a community. In all its forms, whether as something we do as a part of everyday living, or as part of a research project, “naturalistic observation” is highly complex, sensitive, and often problematic yet necessary to our social existence and identities. The question then is, how do we make sense of what we are continually observing? And how do we know what we think is going on, is actually going on?

Observation

Observation seems to imply that the something that is “going on” can be seen, that we are in effect looking at or for something “seeable” or in more generalizable terms, “perceptual” or indeed “intuitable” in ways that are describable. However, perhaps what is going on, is something we are not looking at, or cannot look at. Perhaps then, the first move in naturalistic observation is to reflect upon what is “natural” to our own habitual patterns of attention and inattention. For example, in the experiment by Chabris and Simons (2010) students were asked to count the number of times a basket ball was passed between players. This they duly did. Then they were asked if they had noticed the gorilla. Few did. Yet a person dressed in a gorilla suit had passed among the players as shown by a replay of the video taken. In attending to counting, they were unable to notice the “gorilla”. What we perceive to be as “the natural world out there”, the “reality”, is constructed by patterns of noticing and of not noticing.

This of course, is a laboratory or at least a laboratory-like experiment that disturbed or at least made visible something about the role of attention in “seeing”. However, there is a distinction to be called between the experimental method and the naturalistic method. The idea behind the laboratory is to refine and reduce complexity in order to permit the control of variables so that through the variation of one variable, while holding the others constant, the effects that result can be attributed to it. This reduction to measurable and controllable variables is often referred to as positivism. However, the natural environments of the world outside of the highly specialized settings of the laboratory are not in anyway so controllable. As Cohen very early on pointed out in the case of using statistical sampling in the social and psychological sciences:

In the end, the truth of a generalisation from a sample depends on the homogeneity of the group with respect to which we wish to generalise. A single experiment on a new substance, to test whether it is acid or alkaline, is much more convincing than the result of a questionnaire addressed to millions of army men to measure their intelligence. For the latter is not a simple quality of a uniformly repeatable pattern. In this respect the methods of social statistics are gross compared with refined analysis, so that when our analysis is thoroughgoing, as it generally is in physics, one or two samples are as good as a million. If what we are measuring is really homogeneous, one is sufficient. In the social field, therefore, statistics cannot take the place of analysis ...

Cohen (1944:134–5).

Analysis in the natural world, or perhaps it could be called the wild world, has complexity as its irreducible feature. How do we accomplish analysis in such naturalistic settings?

If such settings are to remain “naturalistic” then presumably they should not be “unnaturally”, or artificially disturbed. However, there is little, if anything that the human does that is not “artificial”. As Agamben (2009) put it in his exploration of the Foucaultian use of the term “dispositif” (usually, albeit inadequately translated as “apparatus”), the dispositif includes language, discourse, tools, forms of organization, material and cultural infrastructures—in short, anything made or used by humans. A dispositif can be generalized to encompass the material and symbolic resources that can be called upon to address issues or undertake activities in everyday life (Schostak and Schostak, 2013). This also effectively describes the field of naturalistic inquiry when applied to the study of human beings in their “natural” habitats undertaking their everyday lives. In the study of human activity, Weber for example argued that “meanings” and human action cannot simply be extracted from a study of so called objective “social facts” that are independent of their value, their connotations for lived experience. Thus, the influential “Chicago school” of symbolic interactionism in sociology drawing on the pragmatism of George Herbert Mead (1934) argued that one cannot know the meaning of an observed action until engaging with others in their ordinary everyday settings (e.g., Whyte, 1943) in order to explore their vocabularies of motives (Mills, 1940).

However, in being, or becoming a member, an ordinary participant in everyday life, how may one “know” that ones perceptions of what is going on and what members say is “true” or “real” is not in some way biased, distorted by the very acts of contributing to the accomplishment of everyday life? In short, is the act of becoming a member compromising the critical capacity of the observer? Of course, it can be argued that it is precisely the prejudices, the values, the meanings and so forth of everyday life that sustain a way of life that is being sought. However, if there are prejudices, this implies a contrast that enables a distinction to be drawn between the prejudiced and the “true”. How can one know the difference? What are the necessary principles and practices that enable such a distinction to be made? In essence, this is that kind of questioning that is behind the phenomenological turn in sociology. The philosophical method of Husserl was not to doubt ones beliefs, values and knowledge about the world but to suspend them - that is, take no position either for or against their “truth”, “existence” or “nature”—in order to “see” what remains after the suspension of what is consciously given that cannot be further “suspended”. For Husserl what could not be suspended was the very act of consciousness itself that is always directed toward an “object”—a structure he referred to as intentionality. This led to an understanding about the essential (that is “invariant”) structure of consciousness that in all possible variations it is always directed toward an “object” in relation to a “subject pole”. This insight would then, it was argued, enable the phenomenologist to describe the invariant structures of conscious acts of perceiving objects and of subjective organization. For example, in observing an object—say a tree—consciousness can be variously directed toward it: from one side of it to another, by getting up close and touching it, by smelling it, by trying to push it. Each act gives a bit more information about the structure of the object. By looking at other objects and following the same procedures one can see what is similar and different and thus identify what are the invariant features—that is, those that remain the same no matter the apparent variation—of a given object. For Husserl this phenomenological method revealed two realms, as it were, of conscious observation: the philosophical and the everyday. This latter he acknowledged was compatible with the descriptions of it by pragmatist perspectives drawing from William James, John Dewey, George Herbert Mead. Alfred Schutz (1976) applied Husserl’s philosophical insights along with those from pragmatism to develop a phenomenological sociology to explore the everyday realm, the paramount reality he called it. For him, the shortcoming of pragmatism was that it did not suspend their everyday attitude of taken for grantedness concerning the “truth” or “reality” of their perceptions. In a sense, as in Debord’s (1994) critique of contemporary society, there is the danger of the taken for granted observation or the spectacle dominating and indeed replacing the real so that what counts as “truth” seems indistinguishable from illusions, make believe or indeed lies—a condition many have called “post-truth”, particularly in reference to the Trump campaigns and presidency and the British campaign to leave the European Union (Schostak and Goodson, 2020).

Broadly, then, the objection is that observations of what is said and done are naive to the extent that the face value of what is being perceived and how it was being interpreted and justified is not subjected to a critical process of suspending belief in the truth, existence, and nature of what is being observed. Of course, everyday life includes objects of consciousness that do not have mathematical like qualities of precision and that cannot be measured as such. Thus, there are objects that are not arbitrary but cannot be fully grasped in their essential—invariant—structures, processes or nature. Tragesser (1977) called these “prehensions”. It is a useful term to apply in social contexts where what happens or what is perceived is not fully grasped but cannot just be “wished away” arbitrarily. To the extent that they are not arbitrary, they are “objective”, that is, independent of the whims and wishes of people that they were otherwise.

In practical terms this “objectivity” can be brought about by what Husserl called eidetic variation, or what sociologists call “triangulation”. Both involve processes of variation, the first through imagination, the second through perception. For example, a mathematician can imagine a triangle, its objectivity is underpinned not by its “natural occurrence” in the physical world but as an object of thought. This object is not arbitrary. Although the length of the sides can be varied to produce different kinds of triangles, it is not possible for example to bend a side, or for an additional side to be added without “breaking” the idea of the triangle and producing something different, say, a rectangle. This process of systematic variation reveals what is invariant, that is, what does not change, or is common to all possible variations as a criterion for a triangle to remain a triangle or a chair to remain a chair. Similarly, on a map, a position can be defined by correlating—or triangulating—its latitude and longitude or its position in a grid. Metaphorically, the validity, objectivity, and, indeed, reliability of an observation, an object, a category, a meaning, a state of affairs, or an event is increasingly established through correlating—or cross-checking—a multiplicity of viewpoints directed toward it over time and across different contexts. Thus in the social world where spectacles can be fabricated to produce a false sense of the real, what looks like a street of houses at first sight, might—on further inspection—turn out to be simply a facade constructed for a film set. By walking up to the facade, trying the doors, walking around the back of it, talking with people working at the site, and so on, its

nature can be assessed. To come to an appropriate judgment with regard to the validity of the street, more than one method and more than one perspective had to be employed. In social settings, naturalistic observers can talk with—or, more formally, interview—a range of people, in order to see whether they share understandings about given phenomena. These understandings can then be compared with what is seen during observations where the phenomena described in the interviews are supposed to occur and contrasted with other kinds of social phenomena. The more agreements there are between different methods and viewpoints on social phenomena, the stronger, it can be argued, the triangulation. Through a process of comparison and contrast distinctions can be made where one might say, “this” is the same as “that” but only has a family resemblance to, or at least apparently looks like some other “that”. In the latter case, there is a critical distinction to be made, a frontier to be maintained, as it were between “this” and “that”. These frontiers may be simply useful category boundaries creating taxonomies that enable people to navigate, manipulate and address attention to a complex dynamically changing world about. However, that analytic “utility” can become charged with values and be emotionally laced when for example distinctions between “good” and “bad”, “safe” and “dangerous”, “us” and “them” are made. This is a process of studying social phenomena that, is grounded in observations and accounts drawn from the experiences, the practices, the discourses of people as they go about their everyday lives. To that extent it is an essential feature of naturalistic observation.

Through our observations grounded in “natural” contexts, can we go beyond descriptions and start to offer analyses that develop theories and show causal relations? Mills (1940) for example explored the motives given by people to explain their actions. Explanations provide reasons for action, for “what is going on”. In that sense, people deploy everyday theorizations of why things happen and how to make things happen. Thus, unlike the mechanical “causes” of effects that laboratory experiments on material substances can identify, human behavior involves “wilful”, “intentional”, “reasoned” forms of behaviors as well as what may be called “unconscious drives” that are socially and culturally framed as distinct from “instincts” that are genetically inbuilt, programmed, that is, biologically driven independently of social learning. Thus in looking for the “causes” of human behavior we can observe and take note of the discourses, the narratives, the anecdotes employed in everyday conversation to justify behavior. Some of these behaviors have a certain regularity about them such as habits, routines, and customs that occur according to context and circumstances or perhaps vary according to different places, times and circumstances. Such regularities provide another source for theorization and prediction concerning the why and how of “what is going on” and what will happen next. Differences may be noticed due to the gender, age, social class, ethnicity or faith of the individuals involved in a given context. Indeed, people put a lot of effort into making life as predictable as possible—whether as bus timetables or the enforceable laws of a given society. Some individuals and classes of individuals may employ both lawful and unlawful power over others to coerce behaviors. The use of this power may be said to be justified according to the “natural laws” and “rights” that have their source in some religious, political, or legal order. However behaviors are justified, the naturalistic observer grounds explanations of them in the everyday life accounts given by the actors being observed. Glaser and Strauss (1967) systematized and popularized this approach as “grounded theory”. However, grounded theory has been criticized for its positivistic practices since it involves “positivistic” or statistical-like reductions of what is common in everyday life to ever more general categories or themes (Charmaz, 2006). In this process the lived particularity of individuals, their existential solitude and uniqueness is lost, repressed, or at least, submerged. Husserl (1970) tried to overcome such positivistic trends in his crisis of European sciences, a book that sought to critique the contemporary mathematization of social and material realities. It is a critique particularly relevant to the digitization of biological and social realities resulting in what Foucault called a bio politics where living systems are increasingly subject to surveillance and control (Zuboff, 2019).

In all this complexity, how does one avoid getting lost?

Orienting

It would be helpful to have a map. Taxonomies offer such a hope. Given the complexity and speed of change in social realities, however, the practice of making taxonomies does not precede but rather is an aspect integral to critical inquiry undertaken in the course of fieldwork. Through continual reflection on the nature of observation as one enters the field, distinctions can continually be made about the ways of seeing and not seeing what is going on “in the wild” or under “natural” rather than controlled situations and contexts. In doing this a first taxonomic strategic decision is made: whether to impose categories theoretically or to seek them as a process of engagement through observation on the ground. It is argued that the latter is the appropriate approach for “naturalistic” studies “in the wild”. Secondly, having discussed what counts as “naturalistic observation” the discussion turns to the role of the researcher: that is, how one positions oneself as researcher in relation to others in “naturally” occurring social settings. A second taxonomic distinction can then be made between observers as “outside” the scene or “inside”. Entering “inside” natural settings means entering as a participant however “neutrally” positioned as a “fly on the wall” researcher. Entering as a researcher then means making decisions that have taxonomic implications for the form of naturalistic observation to be undertaken, about how, even whether, to reveal the nature of the role being adopted. That is to say, naturalistic observation may involve participants being either “covert” or “overt” or at least “sly” in terms of revealing the purpose of their presence. The third step in the discussion involves recognizing the relation between “strangeness” and “familiarity” that is involved in entering social situations, scenes of action and social worlds as a researcher. Indeed, the discussion turns on the question of how to render the world of the familiar, the taken for granted, the “normal”, visible as socially constructed and fragile rather than as a “given” that is “natural”, “solid” and cannot be changed. Rendering the world “visible” is not simply a matter of making descriptions based upon observation, but of inquiring into what is not so easily observed: meanings. A further taxonomic

distinction can thus be made between those studies that seek description of what is observed as an “outsider” and those that go beyond descriptions in order to seek the meaning for participants of what is being observed. This latter involves engagement with the subjects of the research even to the extent of becoming actors who are inducted into the lived realities of those who are the subjects of the research. Thus, there are further taxonomic implications to be drawn between forms of naturalistic observation that either accept the descriptive givens of what is being observed without seeking to engage in action and those that seek to use the findings as evidence to support change and development as in forms of action research and evaluation. Any taxonomy detailing alternative approaches suggests there are choices to be made in the field. Each such choice has implications for the nature of the research methodology and methods employed, the data to be collected and the theoretical work that results. The approach to naturalistic observation adopted can either be sensitive and responsive to the particularities of the situations, circumstances and contexts faced or be imposed rather like a methodological straight jacket. The first, it is argued here, is appropriate to social research “in the wild”, whereas the latter is more appropriate where all variables are under control, as in laboratory experimentation. Each may reveal “what is really going on” in their different ways and for different purposes. But as we move out of laboratory-like settings, what is really going on becomes more difficult to notice. Each step further into the “wild” involves thinking critically and clearly about the kind of approach to naturalistic observation being adopted. It is not just a matter of selecting from a taxonomy of different approaches and “tools” of naturalistic observation but of critically assessing each step in the process of finding out “what is going on”. With each step into the wild, new relations with others are formed that require a critical re-assessment of one’s own position as “researcher”.

How, then does one critically position oneself as a naturalistic observer, in the wild, in an everyday setting?

Participant observation

Within the context of naturalistic forms of observation, participant observation provides a way of thinking about how to position oneself in the lives of others. Evolving from the writings of adventurers and missionaries and, later, the early anthropologists, it was adopted by sociologists who wanted to explore the meanings and experiences of people rather than measuring and counting (deWalt and deWalt, 2002). At present, participant observation is employed in any research, methodological perspective, or discipline in order to gain insight into the lives of others. In many ways, it is more than a method for the production of knowledge, understanding, and theory. It is a way of life which—for some—may last a lifetime. Yet, the very act of “naturalistic” observation is a conscious artifact as it were that disturbs the acts of observation of the observer. It is not normal to say “yes, I am a naturalistic observer, please take no notice of me and go about your business so that I can observe what it is”. To what extent then is concealment or deception necessary, echoing metaphorically, a camouflaged hut used by wildlife observers?

While we do not live in laboratory settings, we do live in settings that are variously manipulated whether by ourselves or by others. What is important in this for “naturalistic observation” is both the seeing and the not seeing. How to see like another in a given “naturalistic” situation, is also what not to see. It is also about how the others being observed “see” and also “not see” the naturalistic observer. How one engages as a participant then, that enables these strategic analytic distinctions is critical to the formation of analyses and theories while grounded in the life world of particular members. In some way, the members have to see the observer as someone who fits the natural setting otherwise the setting will be disturbed and the observations thereby compromised. Moreover, the observer has to see “more” than any of the individual members, otherwise connections between different contexts, circumstances and actors essential to building theory, particularly critical theory will be overlooked.

According to Spradley (1980: 58–62), there are different degrees of participation—complete, active, moderate, passive, and nonparticipation. As a passive participant, the observer is present but not engaged—more a bystander looking at a scene of action rather than being part of it. Rather than merely hanging around, loitering, or being a bystander, one can then try to act like the members by moving on from the stage of recording what went on to actually joining in—but not enough to be acceptable as a member. Going further than this involves entering the world of the other as an accepted member—like Whyte (1943) when he joined Doc’s gang in Cornerville. However, to be a complete participant observer, Doc would already have had to be a gang member who decided to undertake research while participating. The parallel would be, say, a teacher who decides to research the classroom—perhaps, focusing on their own practice in order to improve it as in Action Research.

In each case, participant observation involves being with others—whether living with people far from home whose language, customs, or behaviors are not known, as did Malinowski in his study of the Trobriand Islanders, or, perhaps, like Whyte (1943), where the participant observer goes to live a slum area, or join a criminal gang (Patrick, 1973) or study power elites (Mills, 2000). In these cases, there is a clear demarcation line between my home and their home, my way of life and their way of life, and where I go to visit strangers with the object of becoming familiar with their way of life, their view of the world. The demarcation becomes less obvious when what is observed is already a part of the researcher’s everyday life as, for example, in observing what happens in one’s life, one’s own community—say in a school or in one’s own place of work (as in action research) and the virtual worlds of globalized communications technologies. Whenever, wherever, and however people interact, share experiences, and, in some way, organize their lives in relation to each other, it becomes a possible focus for the participant observer. Understanding the nature and process of being a participant observer begins only at the moment of joining.

Joining

To know the world of others—to understand a particular group—involves joining them in some way, being with people in their everyday locations and activities, and learning how to do what they do in the ways that they do it, in the locations where they do it. Joining with, witnessing what happens, and acting alongside others makes participant observation more than just a method or technique for data collection. The participant observer is part of the data itself—that is, responsible for how this data is to be represented, analyzed, interpreted, and used, as in this example from a narrative account of fieldwork by [Schostak \(1980\)](#):

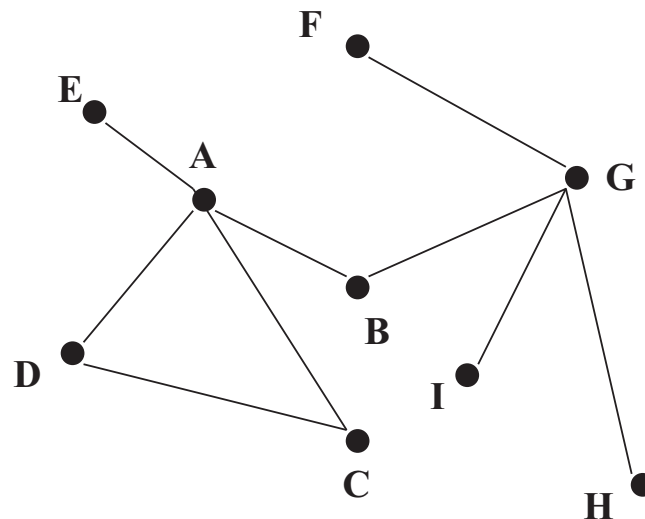


Fig. 1 An intentional network.

Entering the tutor room, a chemistry laboratory, under the gaze of Jacko's mates, I feel inverted—watched, examined. Where shall I sit?

Carol, a girl I know, smiles, invites me to Jacko's group.

The tutor, a teacher I know well enters. He ignores me.

"Jacko. All of you, come and sit at the front."

No one moves. I feel uneasy.

"Come on now!"

"But there's gas taps at that table."

"I don't care, I've told you to sit down here. Now do it."

Carol moves, the rest follow and I am last.

After the formalities of the tutor period—registration, reading a pupil news letter—the teacher approaches smiling. "So you're following Crawley around today?"

"Yes," I turn away as best I can.

"Good luck," the teacher retreats to the doorway where for the remaining time he blocks the escape of three girls who push against him.

Being there and living it, the researcher is a direct participant. It is an approach owing much to the sociologists who, during the first-half of the twentieth century, became known as the Chicago School. Their attention was on the city as a natural laboratory of human behavior ([Park and Burgess, 1925](#)). For example, Thomas studied the communities of Polish immigrants (1918)—and, Whyte, the Italian community (1943)—as the natural areas produced by human habitation and use. In any city, there are many such areas, such as the business district, the Red Light area, or the territories of particular gangs, as well as the secluded areas of the wealthy. In each case, the initial concern is to gain entrance. This, typically, involves finding a gatekeeper—that is, someone whose knowledge and authority can ease the passage into the world of others by hanging around sufficient times in a place where the target group tends to meet until someone begins a conversation. Such gatekeepers are easier to identify and contact in formal organizations—for example, the headteacher or principal of a school, the leader of a youth group or sports team, the head of a committee or director of a business. In other circumstances, greater secrecy may be vital as in the study of a criminal gang. To openly identify oneself might cost one's life, or, at least, physical health. Whether open or covert, participant observation requires developing strategies for meeting, joining, and engaging with others in their everyday scenes of action.

Ethics, *dramatis personae*, locations, and scenes of action

Joining with others means getting to know the members of scenes of action—the *dramatis personae*—getting to know their values, their reasons for action, and their ways of acting. For example, being with Jacko meant going where he went, meeting other people along the way—pupils and teachers—and, when engaging with these others, reflecting ethically on the decisions concerning what to say, how to act, and what and how to record. Observing Jacko's (A) actions toward his mates (C, D, and E) and say a teacher (B) as they approach or keep their distance, patterns can be observed and noted (Fig. 1):

Each individual acts in some way toward another—whether in friendship, hostility, or authority. In phenomenology, this directness toward another is called an intentional relation. The diagram, thus, describes an intentional network (Schostak, 2002) of people as they repeatedly engage with each other. In Fig. 1, Jacko's (A) attention is directed, in various ways, toward B, C, D, and E. B also has a particular relationship of mediation with G (say, a senior teacher, a parent, or a police officer, or a rival gang leader). This mediation may play out in events such as B telling G about A or misleading G about A's real intentions. Or, instead of people, the focus could be upon valued object B that A and G both desire hence leading to a conflict of interest and thus contributing to a study of the ways in which such conflicts are handled. The participant observer becomes a part of this structure, say G, as participants recognize and trust the researcher and as a valid part of the scene able to apply the ethics of the participants—that is, what is right, good, fair in their eyes.

Following Jacko throughout a particular day was part of a larger study of a school and its community (Schostak, 1983). The ethical principles for the conduct of the project had to be agreed. To do this, the participants needed to have an idea of the purposes of the project, what it involved and how data was going to be used. They needed to be assured that they could not be identified and that what they said and did would be treated confidentially so that others could not use the information to harm them. Such principles can be seen more formally in the guidelines provided by discipline-based organizations such as the American Educational Research Association or the British Sociological Association. Indeed, particular organizations—such as hospitals and universities—typically, have their own ethics committees that oversee research involving their institutions. In practice—for each organization—ethical protocols need to be tailored for the specific circumstances of projects. With a formal organization such as a school, general principles may be agreed in writing with its governing committee. Letters describing the research and assuring confidentiality and the anonymity of names in the use of data will be sent to parents as a basis for agreement to be involved in the research. Participants may be asked to sign letters that confirm their understanding of and agreement to the conditions of participation. However, ethical practice is not simply reducible to—nor does it stop with—satisfying the formal ethical procedures of committees.

The different circumstances of each meeting continually pose ethical issues. Not all of the people—the *dramatis personae*—involved in a given scene of action in the various locations of the project may be aware of the status of the participant observer as researcher rather than as an ordinary member of the scene. Indeed, Hilbert (1980) argues that there is always a degree of covertness in research, as aspects of the researcher's biography and later uses of the data remain obscure in the eyes of participants—whether through lack of understanding, familiarity with academic discourse, or disinterest. In the extreme circumstance of joining a gang, or being a football hooligan, it may be argued that not only may acts be seen that are criminal, but also, to gain trust as a bona fide member, illegal acts may have to be committed requiring a rethinking of what it means to act ethically (Ferdinand et al., 2007). Without trust, no access to the world of the members would be granted, thus no research into their life-world is possible.

The researcher needs to demonstrate a commitment to the reality of those being observed, however they define this reality as being true, natural, or right—it is what Quine (1948/1949) would call an ontological commitment. Commitment to what there is in the world is revealed through people's vocabulary, ways of talking, ways of justifying beliefs and actions, as well as in carrying out the kinds of acts expected of members.

In this way, a shared sense of reality is constructed that acts as the natural, taken-for-granted, familiar basis for undertaking social action. Typically, when this familiar world of shared ontological commitments is disturbed, its constructed features are revealed.

The familiar and the strange

Schutz (1964) described how being a stranger sensitizes the researcher to practices, knowledge, beliefs, and values that participants take for granted as being true, real, or right. In learning how to belong—to be a part of a scene of action—a distinction between two worlds is posed: that of the stranger and that of the members of a given group or society that the stranger attempts to enter. Participant observers, similar to strangers, can take nothing for granted—a standpoint paralleling the phenomenological attitude of suspending any position for or against the truth and reality of the world. Presupposed by this phenomenological attitude is a domain of taken-for-grantedness—or paramount reality that is placed in suspension and becomes the focus for the sociological phenomenology of Schutz. As with the symbolic interactionism (Mead, 1934; Blumer, 1969) of the Chicago School, to understand the everyday worlds of people and learn the meanings that actors gave to actions, objects, and the scenes of action in their everyday lives, participant observation was a natural strategy. For symbolic interactionists, it is not enough simply to observe behavior. Behavior has a symbolic dimension. Its meaning to an actor and its relevance to a social situation cannot be known without asking the actor(s) involved. The focus of the symbolic interactionist, then, is upon the symbolic nature of the actions and interactions of people who engage with each other in social settings. Mead (1934) is considered the father of symbolic interactionism, although the name was coined by Blumer (1969). Being a participant observer, however covertly, heightens awareness of the artificiality or socially constructed nature of the ways in which people behave in everyday life. Whether standing in bus queues, waiting at hotel

reception desks, or relaxing at home—what are the regular behaviors that occur? Why do they occur in that way? Could they be changed as in changing the rules of chess? Garfinkel (1967) demonstrated the fragile nature of social rules in his experiments where, for example, a student might go home and act like a stranger, thus transgressing the tacit rules of normal behavior—often, leading to angry reactions by family members. Less dramatically, when participating in Jacko's activities during the school day, the researcher—as a 32-year old—could not pass as a 14-year-old pupil. What becomes noticeable then, are the moves of inclusion as well as those of exclusion made by different members of the *dramatis personae* in the classroom, in the corridors, and out in the playground. Although Jacko had given permission to be followed, he always placed a distance between himself, his friends, and the researcher, although others invited the researcher to join them. Through such a process of seeking permissions, invitations to join, or distancing and excluding, social space becomes dynamically organized as boundaries between who is in, who is out, who is close, and who is distant. This negotiation of boundaries also took place between the researcher and the teachers and the placing of boundaries varied according to circumstances. Thus, the physical reality of streets, houses, and parks are continually mapped and remapped by the ways in which individuals project and negotiate boundaries around them over time. To see it requires distinguishing between “the scene of action, the human picture of that scene, and the human response to that picture working itself out upon the scene of action. It is like a play suggested to the actors by their own experience, in which the plot is transacted in the real lives of the actors, and not merely in their stage parts” (Lippmann, 1922). In this early statement can be seen the seeds of studying social action like a performance on a stage where the analysis of performance, talk and narrative structure, and a focus on biography becomes important rather than static descriptions of states of affairs (Goffman, 1959; Schostak, 2006).

The projection of a world and its boundaries is most obvious when there is a sense of disjunction or clash between alternative perceptions. This can be seen in the following transcript extract from an action research project involving creative practitioners (artists of various kinds), teachers, and pupils (funded by CapeUK Creative Partnerships for Education (CAPE) 2004–05; see also Schostak and Schostak (2008: 48–50)). The focus was on making representations of local communities—both real and imagined. The pupils—aged about 10 years—had just watched a video that had been made of their area:

Boy: ... if you looked at the surroundings they were like, you know “oh I wouldn't like to live there”. Cos if you looked at all the houses they were all boarded up and, all the graffiti and smashed bottles (...)

JFS: So would you have done a different story, you know for (place name)?

Boy: yeah.

JFS: What would you have shown?

Boy: Well I'd have shown the good bits, like the park and ... (laugh) hmmm, there's a point.

Girl: The good parks, the local parks are all trashed up so we don't get anywhere to play.

Immediately there is a clash between the picture in the head and the picture on the video. There are also the explicit and implicit viewpoints of the *dramatis personae*: the boy, the girl, the creative practitioners, the teachers, and the unnamed people who trashed the parks. The participant observer can elaborate the multiple pictures in the head by meeting, observing, and talking to representations of the different viewpoints. How these viewpoints relate to each other can be represented as intentional networks in Fig. 2 where the children are represented by group C who are able—through their reflection on their community—to imagine alternative visions of community that are, perhaps, seen to be enjoyed by group B. The depressing reality experienced by group C is represented by the trashed parks. However, group B never enter the children's location. For them, it is—for all practical purposes—invisible.

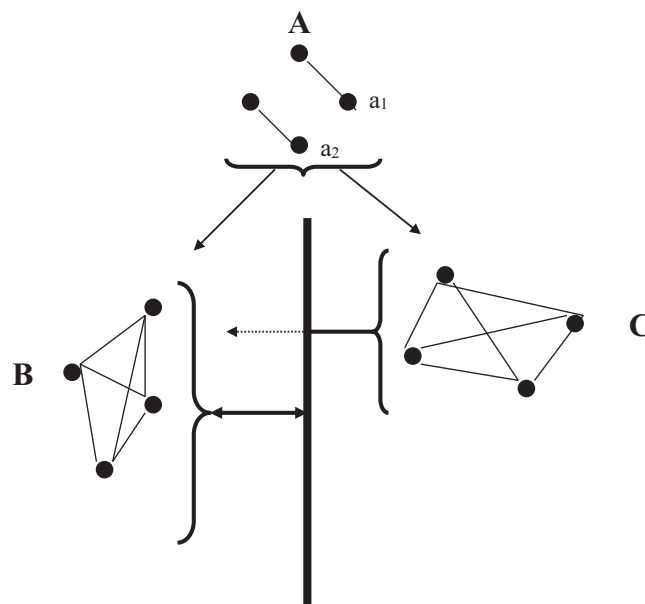


Fig. 2 Mirror groups (Schostak and Schostak, 2008: 172).

The central bar symbolizes a boundary dividing the two groups—B and C—and thus two ways of projecting a picture of the world. By participating in each group, the researcher A is able to build up pictures (a1, a2) of how each constructs their ways of seeing the world with their different ontological commitments: whether the hard-and-fast realities of the social organization of gang territories or the views and fears of locals residents who want to freely walk the streets, or the police and other figures of social and political authority who have quite different conception of how the world should be organized. If group B represents the dominant voice in the social, political, and economic organization of society, it will see itself as exemplars of good order. The boundary that delimits their picture of the world acts like a mirror reflecting back their own values (represented by the double arrow in the diagram). For B, C is rendered marginal at best—perhaps an underclass—invisible at worst. The “bar” can be explored as an infrastructure in itself, representing, say, communications technologies both to connect and to separate users into exploitable on-line communities (Zuboff, 2019).

Rendering the invisible visible means transgressing, perhaps deconstructing, the bar between the dominant and the powerless. The role of the participant observer is already transgressive. Between these multiple camps, the participant observers are always, to a degree, out of place because they are always between worlds. The danger is in slipping into a position of authority granted by some superior insight into what are the right ways of seeing and of telling the truth.

Fieldwork as a basis for validity, generalization, reliability, and objectivity

Fieldwork involves making records of what is seen, heard, and done in particular locations under particular conditions: making extensive notes, maps, and drawings; committing to memory what happened; recording sounds and images; collecting artifacts and documents, etc. It involves decisions regarding how obvious the process of recording will be. Note taking when being a member of a gang is unlikely to be perceived as natural but, as a student in a classroom (Hilbert, 1980), it is perfectly natural. Thus, depending upon circumstances, making notes may take place either during or following the scenes of action.

Even as a bystander, or recipient of stories, the participant observer becomes a member of the *dramatis personae* of a scene of action and thus as a member of the different *dramatis personae* that engage with each other in scenes of action across a range of locations and becomes a carrier of the local stories similar to the other members—each with their views. The multiplicity of views constructing stories of what happened, with whom, how, where, and why are retold for particular purposes and become vital elements in the sense of reality and drama of everyday life.

Each story retold connects to other stories, weaving each actor, the scenes of action, and the objects involved in the action into a reality, a sense of my place, my people, and my world where what happens becomes familiar, expected, routine, and ritualized and where the unusual stands out. Aspects of one story may recall other stories—stories that elaborate on the character of one of the actors, or add to a sense of pride, frustration, or tragedy whether it is “teachers always pick on you like that” or “that’s what always happens to people like us.” This connectedness—or intertextuality—composed of shared and sharable anecdotes, local histories, and biographies becomes the vital data of the participant observer (Schostak, 2006) who also contributes to the stories through various degrees of participation. This intertextuality, with its multiplicity of viewpoints scanning shared or disputed locations, territories, scenes of action, and events, contributes an intersubjective sense of generalizability, validity, reliability, and objectivity of observations and accounts described during fieldwork. More specifically—in phenomenological terms—across all the variations of viewpoint some relatively invariant or typical relations, actions, objects, and meanings attributed to them can be identified, thus leading to judgments as to the generalization of the observations and analyses made. The process is, typically, referred to as triangulation. Features seen with one group and in one scene of action can be compared and contrasted with other groups, scenes of action, and locations. With the aim of identifying the range and limits of variation across individuals, forms of social organization, and locations, the participant observer can seek out comparative and contrastive situations, people, and locations in order to test out emergent understandings, analyses, and theories in a process called theoretical sampling.

The participant observer, then, is continually learning, making analyses, and searching for opportunities to test out emergent understandings and theories concerning the social world being studied. As well as making field-notes to describe observations of people, events, and places, analytic memos can be written in order to formulate potential categories by which insiders describe the circumstances of a situation and provide motives for their action (cf. Mills, 1940) and theoretical memos on emergent theory that can be tested out through further observations (Schostak, 2002). The participant observer is interested not just in a single subject’s point of view but in how the sense of a common world, shared reality, and orchestrated behavior is brought about as an intersubjective reality in the lives of people.

When the researcher is able to represent the intersubjectively constructed world(s) of actors in ways that they recognize as their world, their definitions, and their way of doing things—whether through performance as a member or in spoken and written accounts—this achieves a degree of validation and reliability. More critically, participant observation by seeking out the multiple voices and perspectives and learning from their viewpoints provides a critical resource for researchers in their debates with regard to what counts as knowledge and its “truth”. Whether it is the view that there is the inevitability of a distortion—a violence even—in the interpretations of the researcher as shown by the critiques of feminists or by critical race theory or poststructuralism, the debate is continually renewed, shifted, and challenged by a return to the field.

Truth out of place

There is a sense of “ordinariness”, of the “routine”, the “predictability” underlying the idea of everyday life. However, this “everydayness” can be “naturally” interrupted, suspended, put out of joint by personal, local, regional, national, world crises. Where [Garfinkel \(1967\)](#) developed the phenomenological method of “suspending” or “bracketing” in his fieldwork practice as a method of causing trouble in order to lay bare the otherwise taken for granted social structures and practices that gave constancy to everyday life, such suspension can be bought about by accidents, natural disasters, economic crises, civil strife, technological revolutions, wars and epidemics that jolt or even render the “everyday” of yesterday as no longer valid, re-achievable.

Moreover, what such crises may make visible, are not only the lost familiarity of the everyday but the fundamental disagreements, injustices and indeed violence that have been covered over by its veneers. In these contexts, naturalistic observation addresses the struggles of people to make sense of their altered situation, to “return” to what was, or to call for a better world to be built from the ruins. In doing so, naturalistic observation can make clearer what matters to people in their lives, what is at stake in adopting one view or course of action than another. Indeed, it may well seem and be argued by many that the “truths” that had been established as a natural part of the social, economic, political, cultural order of life are “out of place” or no longer relevant for the emergent or desired realities of members.

The Covid 19 pandemic that led to the suspension of much of the world’s economic and social life in 2020 was just such an event. The “normal” practices of social contact and greeting were suspended by demands for “social distancing”. No more hugs, handshakes, kisses as rituals of greeting. Breath and surfaces—and the skin is such a surface—became potentially dangerous as conveyors of disease. Where the rich had access to large private secure spaces, for poor families their cramped spaces contributed to insecurity, anxiety, and increases in domestic violence. Masks, disinfectant and measures to avoid close contact when moving around public spaces and inside shops became a “new normal”. Suddenly, overlooked economic activities such as delivery drivers, transport workers, postal services, cashiers, farm laborers, refuse collectors were seen as indispensable in keeping society functioning. Nurses and medical staff were publicly applauded. These were called the front line workers, who largely work in what economists call foundational infrastructures (health, welfare, education, utilities such as gas, electricity, water, sewerage and indeed, banking as a system enabling market exchange) rather than those more superficial less vital economic activities where speculators make money, entertainers seek fame, and entrepreneurs sell their wares. These latter activities depend upon the existence of the foundational infrastructures ([Bowman et al., 2014](#); [FEC, 2018](#)). The Covid 19 crisis had rendered visible the “truth” of the different kinds of activity. Moreover, it occurred alongside a rise in right wing political popularism that offered nationalism as a solution to many complex historical inequalities and injustices as well as alongside an increasing awareness particularly among the young of impending environmental and climate crises. For the naturalistic observer of social life, the degrees of public recognition of the conjuncture of multiple crises creates opportunities for obtaining insights into the subjective and social lives of individuals as they try to make sense of their struggles under such circumstances.

Crisis has a particular significance in contemporary capitalism—it is, as [Friedman \(1962\)](#) argued, the occasion to capture political, social and economic power by having, as he put it, the only ideas “lying around” that provide a solution to the crisis. And capitalism, as many have pointed out, thrives on reinventing itself through crisis (e.g., [Harvey, 2010](#)). However, such ideas are always in contest with alternative ways of producing the “good society” and creating for “ordinary” people a sense of security or at least a hope for a better future. Indeed, it is this feature of everyday life where the “truth” or “facticity” of a given practice, or set of circumstances, can be publicly contested through debate from multiple view points that the early leaders of the French and American revolution saw as a way of embedding and guarding democratic forms of public life. The public, they argued, had eyes and ears everywhere and thus through debate could form opinions and knowledge with a validity, accuracy, objectivity and generalizability that no single person, or indeed, small group could do alone ([Rosanvallon, 2013](#); [Schostak et al., 2020](#); [Schostak and Goodson, 2020](#)). It is in this spirit that [Lippmann \(1939\)](#) argued for the necessity of the oppositional views in the development of “true” public knowledge and [Dewey \(1927\)](#) saw as necessary to the development of a democratic public. In this sense then, the practice of naturalistic observation through research is itself a constituent of and contributor to the more general practice of the development of an informed and informing public. Participant observation thus becomes part of the struggle both to inform and to create a public capable of determining “what is going on” in their everyday lives as a basis for their decision making in everyday life ([Schostak and Goodson, 2020](#); [Schostak et al., 2020](#)).

The significance of naturalistic observation

In summary, as Thomas (in [Thomas and Thomas, 1928](#): 571–572) put it: “If men define situations as real they are real in their consequences.” The task is to learn how situations are defined as a member, how reality is constructed as a member, what it means to be a member, or even to pass as a member of a particular group. This is not to say that the social realities of members are homogeneous. Indeed, where there are disputes or disagreements, these are managed or resolved according to the rules, rituals, habits, beliefs, values, and social positions of the members. For that reason—from an ethical and democratic standpoint—the naturalistic observer records the range of disagreements as well as agreements in order to produce a record that the French political philosopher Jacques Rancière calls being faithful to the disagreements (cf., [Schostak and Schostak, 2008, 2013](#)). This includes not only a representation of the world of those being studied but also extends to the public debates concerning what counts as knowledge and whose knowledge and perspectives are to be included in public debates that affect people’s lives. Naturalistic observation can be the basis for forming “ground up” research designs for descriptive case studies whether of particular forms of social organization

or of wider more complex communities. As a fieldwork method it can be employed in other more deliberately interventionist forms of research, such as action research that has the explicit aim of making changes, or in evaluations that explore the impacts of policy on practice. In each case, naturalistic observation seeks the experiences of people, exploring how through their activities, their discourses, their narratives they construct, interpret and validate their everyday lives with each other. By explicitly drawing upon multiple points of observation, the naturalistic observer can explore the constitutive practices of social life in order to reveal the ways in which material, social and symbolic objects are perceived and given significance in the lives of people. In this way, naturalistic observation can be publicly extended as a tool for the formation of the public so that potentially everyone has access and is a contributor to their own and public understanding about what counts as knowledge, justice, and the good society.

The only way of learning naturalistic observation is to do it, be there, and live it.

References

- Agamben, G., 2009. What is an apparatus? Trans. by David Kishik & Stefan Pedatella In: Agamben, G. (Ed.), *What is an Apparatus? And Other Essays*. Stanford University Press, Stanford.
- Blumer, H., 1969. *Symbolic Interactionism: Perspective and Method*. Prentice-Hall, Englewood Cliffs, NJ.
- Bowman, A., Ertürk, I.M., Froud, J., Sukhdev, J., Law, J., Leaver, A.M., Moran, M., Williams, K., 2014. *The End of the Experiment. From Competition to the Foundational Economy*. Manchester University Press, Manchester and New York.
- Chabris, C., Simons, D., 2010. *The Invisible Gorilla. And Other Ways Our Intuition Deceives Us*. Harper Collins, London.
- Charmaz, K., 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage, London.
- Cohen, M.R., 1944. *A Preface to Logic*. Dover, New York.
- Debord, G., 1994. *The Society of the Spectacle*. Translated by Donald Nicholson-Smith. Zone Books, New York.
- deWalt, K.M., deWalt, B.R., 2002. *Participant Observation. A Guide for Fieldworkers*. AltaMira Press, Walnut Creek.
- Dewey, J., 1927. *The Public and Its Problems*. Swallow Press, Ohio University Press, Athens.
- FEC, 2018. *Foundational economy. The Manchester Capitalism Book Series*. Manchester University Press, Manchester.
- Ferdinand, J.M., Pearson, G., Rowe, M., Worthington, F., 2007. A different kind of ethics. *Ethnography* 8 (4), 521–544.
- Friedman, M., 1962. *Capitalism and Freedom*, Preface, 1982 ed. University of Chicago Press, London.
- Garfinkel, H., 1967. *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs, NJ.
- Glaser, B.G., Strauss, A.L., 1967. The Discovery of Grounded Theory. In: *Strategies for Qualitative the Orchestration of Observation*. Aldine, Atherton, CA.
- Goffman, E., 1959. *The Presentation of Self in Everyday Life*. An Anchor Book, New York.
- Harvey, D., 2010. *The Enigma of Capital and the Crises of Capitalism*. Profile Books, London.
- Hilbert, R.A., 1980. Covert participant observation: on its nature and practice. *J. Contemp. Ethnogr.* 9, 51.
- Husserl, E., 1970. *The Crisis of European Sciences and Transcendental Phenomenology, An Introduction to Phenomenological Philosophy*. Translated with an introduction by David Carr. Northwestern University Press, Evanston.
- Lippmann, W., 1922. *Public Opinion*. Harcourt Brace, New York.
- Lippmann, W., 1939. The indispensable opposition. *Atl. Mon.* 86–89. <http://croker.harperthall.org/Must%20Know/Government/DemocracyLippmann.pdf>.
- Mead, G.H., 1934. *Mind, Self and Society*. University of Chicago Press, Chicago, IL.
- Mills, C.W., 1940. Situated actions and vocabularies of motive. *Am. Socio. Rev.* 5, 904–913.
- Mills, C.W., 2000. *The Power Elite*. With a new Afterword by Alan Wolfe, new ed. Oxford University Press, Oxford, New York. first published 1956.
- Park, R.E., Burgess, E.W., 1925. *The City*. University of Chicago Press, Chicago, IL.
- Patrick, J., 1973. *A Glasgow Gang Observed*, third ed. Neil Wilson Publishing, London, Methuen.
- Quine, W.V.O., 1948/1949. On what there is. *Rev. Metaphys.* 2, 21–38.
- Rosanvallon, P., 2013. *The Society of Equals*. Trans. Arthur Goldhammer. Harvard University Press, Cambridge Massachusetts and London, UK. First published 2011, Éditions du Seuil.
- Schostak, J.F., Clarke, M., Hammersley-Fletcher, L., 2020. *Paradoxes of Democracy, leadership and Education. Struggling for Social Justice in the twenty-first Century*. Routledge, New York, London.
- Schostak, J.F., Goodson, I., 2020. *Democracy, Education and Research*. In: *The Struggle for Public Life*. Routledge, London and New York.
- Schostak, J.F., Schostak, J.R., 2008. *Radical Research. Designing, Developing and Writing Research to Make a Difference*. Routledge, London.
- Schostak, J.F., Schostak, J.R., 2013. *Writing Research Critically and Radically—The Power to Make a Difference*. Routledge, London, New York.
- Schostak, J.F., 1980. *A Day With Jacko: An Experience of Fieldwork in School*. https://www.academia.edu/71001495/ethnography_participant_observation_a_day_with_Jacko. (Accessed 10 February 2020).
- Schostak, J.F., 1983. *Maladjusted Schooling: Deviance, Social Control and Individuality in Secondary Schooling*. Falmer, London, Philadelphia. Reprinted Routledge Library Editions 2012, paperback 2014.
- Schostak, J.F., 2002. *Understanding, Designing and Conducting Qualitative Research in Education. Framing the Project*. Open University, Buckingham.
- Schostak, J.F., 2006. *Interviewing and Representation in Qualitative Research Projects*. Open University, Buckingham.
- Schutz, A., 1964. The stranger: an essay in social psychology. In: Brodersen, A. (Ed.), *Collected Papers, Studies in Social Theory*, vol. II. Martinus Nijhoff, The Hague, pp. 91–105.
- Schutz, A., 1976. *The Phenomenology of the Social World*. Tr. G. Walsh and F. Lehnert. Heineman, London.
- Spradley, J.P., 1980. *Participant Observation*. Holt, Reinhart and Winston, New York.
- Thomas, W.I., Thomas, D., 1928. *The Child in America: Behavior Problems and Programs*. Alfred A. Knopf, New York.
- Tragesser, R.S., 1977. *Phenomenology and Logic*. Cornell University Press, Ithaca and London.
- Whyte, W.F., 1943. *Street Corner Society*. In: *The Social Structure of an Italian Slum*. University of Chicago Press, Chicago, IL.
- Zuboff, S., 2019. *The Age of Surveillance Capitalism*. In: *The Fight for a Human Future at the New Frontier of Power*. Profile Books, London.

Photovoice with children and youth

Helen Meyer^a, Laura Dell^b, and Randall Gibson^a, ^a University of Cincinnati, School of Education, Cincinnati, OH, United States; and
^b University of Cincinnati, College of Education, Criminal Justice & Human Services, Cincinnati, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	260
The history of photovoice	261
Photovoice as participatory action research	262
The photovoice process	262
Establishing the research framework	262
Taking pictures	263
Image selection and theme generation	263
Presentation	263
Children and youth photovoice research examples	264
Photovoice and young children	264
Preserve preschool: research setting and participants	264
Data collection and the photovoice process	264
Photography results	265
Theme #1	266
Theme #2	266
Theme #3	266
Theme #4	266
Theme #5	266
Children as co-collaborators in research	266
Photovoice and urban youth	267
The youth environmental club: research setting and participants	267
The photovoice process	267
Taking pictures	267
Understanding green spaces	267
Youth as co-researchers	268
Conclusion	268
References	269

Introduction

A primary role of qualitative research is to uncover the unspoken meanings of symbols and ideas held by communities or individuals, to elaborate the beliefs that guide the actions of people, and to document how communities and the individuals within them construct a consensus understanding of their shared world. Historically, qualitative researchers told the stories of individuals who did not hold the power or position to tell their own stories. More recently, critical researchers have shifted their work to give voice to marginalized and oppressed groups—silenced populations. Critical researchers break down the barriers silenced populations face as they seek to gain agency in their own stories (Clifford and Marcus, 1986; Fine, 1994; Shah, 2015).

Although children are frequent subjects of research, “much of the research on childhood has been conducted by adults observing children or by adults reflecting on childhood—both on their own childhoods and on the childhoods of today” (Dell, 2011, p. 5). Adult voices become the expert voices; adult interpretations replace children’s meanings; and adult reasoning gains consensus as the motivation behind children’s action. In this way, children’s voices are suppressed and children’s experiences are marginalized to a subculture of adult experience. In the same way critical theory pushed researchers to take a critical stance that empowers disenfranchised groups, children need their agency as active participants in studies of their lived experiences.

Participatory Action Research (PAR) is one mechanism available to close the gap between subject and researcher (Abma and Schrijver, 2020; Groundwater-Smith et al., 2015). PAR is an inclusive research methodology that requires children be involved in all stages of the process. It goes beyond simply asking about their ideas regarding a topic as part of data gathering or if the presentation of their ideas in the analysis sounds good, PAR puts children in charge of the research, allowing them to be the active agents of the exploration of their lived experiences (Abma and Schrijver, 2020). Photovoice is a particular form of PAR that combines the power of photography to capture, assess, and communicate a sophisticated view of community concerns and ideals that words

alone are difficult to express (Wang and Redwood-Jones, 2001). Carolyn Wang, Mary Ann Burris and their research colleagues developed and shared photovoice as a participatory research method in the late 1990s. They emphasized the importance of the collaborative processes of sharing pictures and the story of the individual's relationship to the image and the environment from which the picture arose.

In the last 20 years, photovoice's use with youth and children has increased. Photovoice with children serves the dual purposes of building on children's methodological skills and developing new pedagogical tools. The use of images and storytelling as a research method, builds on experiences even young children are familiar and engage in their daily lives through play, drawing, and reading with others. Pedagogically, photovoice may provide a new experience intentionally developing the children's voice, social action, and advocacy experience (McKenna et al., 2017).

The majority of photovoice studies with children and youth fall into one of three broad topic categories: health, poverty, and learning and schooling. However, many studies sit in the intersection of these categories. In 2006, Caroline Wang reviewed ten youth photovoice studies focused on public health issues facing youth, the majority of which addressed health issues in low-income urban environments, promoting critical dialog with the youth participants and policy makers. Shah's (2015) study encompassed poverty and schooling in her work with school age girls in India and the impact of traditional roles of women on the potential for empowerment through education. Other studies focus on inclusive schooling. Mackenzie and Talbott (2018) worked with elementary aged children parented by LGBTQ parents to create a more inclusive school environment. Multiple school learning examples include understanding the abilities of students with disabilities (Meyer and Kroeger, 2005; Agarwal et al., 2015), literacy struggles (Pollock, 2019) and disciplinary learning (Cook, 2015).

These works involving children and youth engaged in photovoice, demonstrate its potential to break the lock on adults interpreting children's experiences and give children back their voice. In this chapter, we focus on conducting photovoice with young children and urban youth, and its use to capture how the children and youth understand and value their experiences with nature. The importance of youth developing robust and positive views of nature must not be diminished. Educators and environmentalists are concerned that children are losing connection with nature, an impact that changes the children's physical and cognitive development and their adult environmental attitudes and behaviors (Dell, 2018). As we rush toward a worldwide climate crisis, we cannot afford any generations losing nature experiences that empower them to advocate for the environment and environmental change.

In the remainder of this chapter, we overview the history of photovoice and define it as a specialized form of PAR, separate from photo-documentary work. We then outline the photovoice method as described by Wang and Burris aligned to a condensed method we used for our work with children. This is followed by two exemplary photovoice studies with children. The first study was completed with pre-school age children (ages 3–6). This example highlights modifications to the photovoice methods for use with very young children and introduces what was learned when the children became co-researchers at their nature-based preschool. The second study was completed as part of an urban environmental club. The co-researcher youth in this study were middle school age children (ages 12–14) in an urban public school. Unlike the preschoolers in the first example, the youth in the second example had limited nature experiences, but we learn how their experiences in nature reflect their position in the community. Finally, we conclude with a discussion of the importance of photovoice to engage children as co-researchers of their world and experiences providing them with the agency to teach adults and advocate for a sustainable future.

The history of photovoice

Photovoice builds on Freire's (1970) critical examination of participants' local communities to derive educational opportunities through line drawings leading to text-based problem-posing education. In 1989, Rhoda Linton turned a feminist lens on Freire's problem-posing to focus specifically on issues facing women. Linton (1989) acknowledged women share knowledge and experience through a range of representations beyond text including, storytelling, cooking, arts, crafts, and political action. Each representation provides detail and texture to the lived experiences of women. Linton believed when women engaged in cooperative group activities, they began to understand their collective oppression within the status quo and identify public policies and actions leading to liberation. A characteristic of feminist research methods for Linton is the need for strategies that are open, inclusive, accessible, and creative in order to meet the varied educational backgrounds and lived experiences of women across the world; photovoice provides this research avenue.

Wang and Burris's (1997) article *Photovoice: Concept, Methodology, and Use for Participatory Needs Assessment* introduced photovoice as a research method. The methods arose from their work with women in rural China. They used photographic images first to address critiques of Freire's use of drawings by men as part of his cultural circles (Maguire, 1987), and second because the photographic images added simplicity to data gathering. "Because virtually anyone can learn to use a camera" (Wang and Burris, 1997, p. 370). As a research tool, a camera has many advantages. Participant researchers with cameras have access to settings and locations that external researchers may not readily gain access to or initially deem worthy of study. Wang and Burris found their photovoice participants took pictures of events and locations that were not typically accessible to or considered research settings, but were

integral aspects of the women's lives. An additional strength of photovoice they discuss is its adaptability—to the purpose of the research, the needs of the participants, the community setting, or issue under studied.

Photovoice as participatory action research

Photovoice is particularly relevant to Participatory Action Research (PAR). Similar to [Stringer's \(2022\)](#) community-based participatory action research (chapter 19 in this volume) the critical participatory goals are the same but in the case of photovoice, the photographic images create a powerful level of ownership for research participants. The use of a visual image empowers the participants to initiate the story, its meaning and centrality to their lived experience. With photovoice, the images engage the participants in the process of identifying and representing their experiences similar to Freirean problem posing's use of line drawings ([Meyer and Kroeger, 2005](#)). The shared images and stories honor the expertise of the individuals as actors in their own experiences and engage participants' in the development of and participation in democratic decision-making and public policy.

Photovoice is not about creating a beautiful photographic image, nor is it photo-documentary work. [Sontag \(1977\)](#) described image photography as a relationship between the photographer and their context and purpose for taking a picture. It is the viewer's role to form a relationship of their own with the photograph based on what the photographer has provided. In this way, a "photograph, like a riddle or problem to be solved, must be decoded" ([Meyer and Kroeger, 2005](#), p. 188). In photo-documentary work, the photographer is interpreting their subject's reality; selecting images and representations through the photographer's lens; creating the reality important to the photographer's narrative. The loudest voice in photo-documentary research is the voice of the researcher. In photovoice, the participant researcher fully owns this narrative. Their voices are activated by identifying what to photograph and selecting the photographs that represent their reality.

The photovoice process

In its early use, the photovoice process was defined as a nine-step strategy deeply embedded in community action and policy development ([Wang, 2006](#)). However, as it has expanded beyond its origins in health policy and, as technology has changed how we take pictures, some steps have become less prominent. [Table 1](#) shows Wang's original nine stage methods aligned with a condensed, four stage method ([Meyer and Kroeger, 2005](#)). A strength of photovoice is its adaptability based on the needs of the community of researchers, we found this consolidation of Wang's original stages ensured essential aspects were retained but other steps that might have been too difficult for young children or unnecessary with current technology were not included just to meet a methodological requirement.

Establishing the research framework

The activities of this first stage happen before any pictures are taking. Its purpose is to develop a cohesive research community who are aware of not only the purpose of the research, but their roles and responsibilities, its potential for growth, and its potential for conflict. [Wang \(2006\)](#) articulated four inter-related activities in this framing stage. Although some of these activities are technical, use of cameras or image storage, most of the activities involve honest discussions with the participants. Unlike more traditional research methods, rather than identifying research questions, in this stage the photovoice researchers discuss their lives and experiences to identify an appropriate audience that needs to learn from their life experiences. The identification of the audience for their research becomes a guiding idea in the next stage.

Table 1 Consolidated photovoice stages ([Meyer and Kroeger, 2005](#)) aligned to ([Wang, 2006](#)) stages.

<i>Photovoice stages (Wang, 2006)</i>	<i>Consolidated stages (Meyer and Kroeger, 2005)</i>
1. Select and recruit a target audience of policy makers or community leaders	Establishing a research framework
2. Recruit a group of photovoice participants	
3. Introduce the photovoice methodology to participants and facilitate a group discussion about cameras, power, and ethics	
4. Obtain informed consent	Taking pictures
5. Pose initial theme/s for taking photos	
6. Distribute cameras to participants and review how to use the camera	
7. Provide time for participants to take pictures	Image selection and theme generation
8. Meet to discuss photographs and identify themes	
9. Plan with participants a format to share photographs and stories with policy makers or community leaders	
	Presentation

Because the goal of photovoice research is to give voice to an oppressed group, the research participants need to be informed about political or policy issues their research may confront. Everyone needs to know how to act safely in a potentially charged environment. Additionally, Photovoice participants need to understand and discuss issues of power and ethics when taking pictures of other people or particular situations. During these framing activities Wang and Redwood-Jones (2001) suggest clear discussions about questions such as: how do you approach someone to take their picture; can you take pictures of people without their knowledge; or when is not safe to take a picture?; must be addressed. Finally, Wang and Redwood-Jones (2001) explicitly include discussion of informed consent. Photovoice researchers need to understand they are researchers and research participants who are consenting to being part of a research study that may cause upset the status quo.

Taking pictures

Once all the photovoice participants understand their role in the research process, the research transitions into the taking pictures stage. This stage includes deep discussions about the research theme and brainstorming ideas of images that might represent these themes, so once they begin to take pictures they are clear about the focus theme. This generation of a research theme and how the theme might be represented with images from the researchers lived experiences' substitutes for a formally articulated research question. After the research theme is understood, the cameras are distributed and basic photography skills practiced. The type of camera is an important consideration as ease of use needs to be balanced with safety, image security, and cost of printing. In the digital age, camera skills may focus more on centering objects, direction of lighting, or safe image storage than on the technical skills of winding on or loading film from when film-based cameras used. However, depending on the age and experience of the research participants, picture-taking practice may be required. Once this preparation is complete, a picture taking timeframe is agreed upon and the photovoice researchers begin documenting their experiences.

Image selection and theme generation

As a form of action research, this photovoice stage is likely to occur in iterative cycles in conjunction with the refining and revising of the research theme generated in stage two. This stage includes multiple cycles of "selecting photographs, contextualizing or storytelling and codifying issues, themes or theories—occurring in group discussion" (Wang, 2006, p. 151). When selecting photographs each participant picks a few of their images to share with the group. The contextualization of the selected image is guided by a series of questions termed as Wang termed as SHOWeD (Wang, 2006, p. 151).

- What do we See here?
- What's really Happening here?
- How does this relate to Our lives?
- Why does this situation, concern, or asset exist?
- What can we Do about it?

Each participant uses this five-question sequence as they share their images and their relationship to the research theme. The group listens for common ideas and themes arising from the collection of pictures from the group. These themes may lead to a new photo-taking round or the group may determine it is time to move to the final stage.

Presentation

This stage reconnects with stage one as the research participants prepare their findings to present to the audience selected in stage one. Participants determine the most impactful format for presenting their finding, as well as when and where the presentation should take place. Depending on the age of the researchers, subject matter, and audience, the presentations might include displays in a library or other public space, presentation to a parent group, a council meeting, or a formal research paper presentation at a public conference.

Although photovoice is adaptable, it is not ill defined or unstructured. Some processes and steps can be consolidated as we have done or some may need greater emphasis to meet the needs of the research participants, but each stage is important. If a research facilitator does not prepare new photovoice researchers with all information detailed in the steps, they are diminishing the photovoice researchers' role as expert arbitrators of their experiences. Photovoice research, "provides a compelling, critically reflective experience wherein students think and rethink their engagements with self and other" (Latz et al., 2015, p. 126). The next section looks specifically at the use of photovoice in educational settings with youth and young children.

Children and youth photovoice research examples

The two examples that follow demonstrate the power of photovoice to give voice to the lived experiences of children and the adaptability of the method to meaningfully engage researchers as young as 3 years of age. Both examples asked children to consider their relationship with the natural world. The specific foci of the two studies, their implementation and products are different, but from them we learn how children define nature, and what nature means to them and how children are the experts of their own lives.

Photovoice and young children

In this example, we discuss the application of photovoice methodologies with young children at a nature preschool. This research project positioned young children as an underrepresented group who may have their voices and experiences inadvertently marginalized by the dominant group (adults). One goal of this project was to determine if young children (ages 3–6) had the developmental skills and cognition to engage in a photovoice methodology (Wang and Burris, 1997) and what accommodations would need to be made with the methodology to address the development abilities of this age group. As discussed previously, photovoice has traditionally been utilized with adolescents and adults, but not with young children. The project also aimed to understand how children viewed their learning experience at a nature preschool. Nature preschools are a relatively new type of preschool where children have extended daily access to unstructured outdoor play in wild, natural spaces. The goal of these schools is to combine traditional early child education learning with environmental education strategies (Bailie, 2016).

Preserve preschool: research setting and participants

The research setting of this study was a private preschool (referred to as Preserve Preschool) located at a private nature preserve. Two licensed preschool teachers led the class of 15 children (11 girls and 4 boys), who ranged in age from 3 to 6. The children attended 3 days per week for half a day. While there was a traditional classroom building, the teachers aimed to spend at least 65% of the day outside. Families are told to dress the children appropriately for the weather each day, even if it is raining or snowing. Outdoor time included free play on a playscape adjacent to the classroom building and guided hikes to destinations around the nature preserve. The playscape was a natural space with some specifically designed elements added to engage exploration with outdoor materials, such as rocks, logs to jump on, a sand pit to dig in, and an earthen tunnel. Examples of hiking destinations included a creek, pond, trees for climbing, wildflower field, sledding hill, and a stand of fallen logs. Even though Preserve Preschool focuses on environmental education, it meets state preschool educational standards designed to prepare children for kindergarten.

All fifteen children in the classroom participated in the study; however, one child chose not to complete the study. Consent of parents and child assent were obtained. In order to develop trusting relationships with the children, the adult researcher volunteered in the classroom for several months before implementing the study. All children were of Caucasian descent and were assumed to be from middle class or upper middle-class homes, as this was a tuition based private preschool. Most of the families were also members of the nature preserve.

Since nature preschools are a relatively new and atypical type of preschool setting, the goal of the project was to use photovoice to understand children's experiences at a nature preschool from their perspective, instead of imposing the researcher's adult lens on their experience. The photovoice project was for each child to take photos of their school experience over several sessions and then select photos for a book that told the story of Preserve Preschool. The tools for the photovoice process were iPod Touches and hard-bound blank books to create their stories of preschool. The study was framed with children as co-collaborators in the research, as the children took the photos (data) and determined which photos were most important to the story of preschool and why (analysis). The next section goes into detail on the photovoice process.

Data collection and the photovoice process

The following outlines stages 1–4, as described earlier in this chapter, which were utilized in this study. The researcher made adaptations only when absolutely necessary to meet the developmental level of the children. We begin at stage 1 in the photovoice process, establishing the research framework. Since group discussion with very young children about a research theme would be challenging, the children were shown a sample photo book that told the story of a hot air balloon ride that the adult researcher went on. The children were told they would make their own photo story books that told the story of preschool using iPod Touch cameras (hereafter referred to as cameras).

Photovoice is traditionally used as a social justice tool by populations with reduced power. This project was designed to support continuous improvement in Preserve Preschool and a high-quality learning experience for the children. The adult researcher explained to the children there would be benefits to the children and for the adult researcher by creating their photo story books. The adult researcher went on to explain Preserve Preschool was a very special kind of school and the project would help other teachers and the adults, like the adult researcher to learn more about their school and what is important at this type of school. The benefit to the children would be, their teachers would better understand what they like and don't like about Preserve Preschool. Then the teachers could make good decisions about activities they, the children, would enjoy. The adult researcher also explained the benefit to other adult researcher would be to understand how children can be good researchers of their days.

Once the children understood what they would be photographing, the adult researcher moved to training the children in how to use the cameras and the basics of framing a photo. All of the children appeared to have previous experience using a camera app on smart phones, so the children learned how to use these cameras almost immediately and helped each other in learning how to use them. All of the cameras were protected in waterproof, shockproof Otterbox brand cases.

In stage 2 of the photovoice process, child researchers take photos to document a particular issue/experience. The children were encouraged to take photos of whatever they wanted to best tell their story of Preserve Preschool. The adult researcher carried the cameras for the children when they did not want to use them. This allowed children to get them when they wanted to take pictures, without feeling burdened to carry them for the whole hike or during a free-play period. The adult researcher would remind children at the beginning of each day of the goal for taking photos and that she was carrying the cameras. However, the child researchers were not required to use the cameras at any time. The children took photos over 4 days during mid to late spring. The number of pictures each child researcher took varied widely in number and frequency of use was also wide-ranging. The range of number of pictures the child researchers took varied from 16 to 786 photos.

Stage 3 of the photovoice process, image selection and theme generation, typically begins with the community members sharing their photos with each other and discussing the meaning of each photo as part of a participatory analysis. Due to the age of the participants, this was adapted to meet the interest level of the young child researchers. After the four photo taking sessions, the adult researcher printed two sets of all the photos taken by each child and allowed each child to look at their photos. One set of photos was for the child to take home and the other set was saved for creating the photo book. The most significant modification to the photovoice process was needed here, as it quickly became apparent that while the child researchers were very proud of their photos and wanted to show them to adults and other children in the classroom, the children were not interested in looking at or hearing about each other's photos. Thus, the adult researcher adapted this step by sitting with each child separately rather than as a group to discuss their photos.

The next part of stage 3 is the analysis of the photos. To do this, the adult researcher and each child utilized a protocol for conducting the participatory analysis. The three stages of a participatory analysis process are selecting, contextualizing, and codifying (Wang, 2006). To support the selecting portion of the analysis, each child was given a hardbound blank book (28 pages/14 sheets), colored pencils and a glue stick. With the adult researcher, each child researcher sorted through their photos to select the ones that best told the story of Preserve Preschool to include in their book. For the contextualization step, the child researchers wrote a caption for each photo. Some children wrote their own captions, while those who were pre-writers dictated their captions to an adult scribe. The children each decided how to organize the photos in the books and how many photos to use overall.

The final step of analysis was completed by the adult researcher only. This step met the project goal to disseminate information about nature preschool experiences by children to a wider academic audience. The decision to have the adult researcher complete this step without the children was made after attempting a more traditional photovoice group discussion format with the children. Although the researcher tried to engage the children in sharing of their photos, the researcher quickly found the children did not want to engage in a group discussion and observed this level of categorization and abstraction was too developmentally difficult for the young children. Most preschool children are only able to recognize simple patterns (Copley, 2010) and the codification step of the analysis required the capability to discriminate complex patterns over a significant number of photos, therefore the adult researcher completed this step alone. Additionally, since this step was for the benefit the educational research community more than for the Preserve Preschool community, it was determined it would not have been the best use of the children's time. The adult researcher wanted to respect the children's curiosity in the project and not attempt to force them into a step they were not interested in.

In the final stage of photovoice, the presentation of findings, the child researcher shared their final products with their wider community. At the end of year school picnic that included families, the children received their final photo books to share with their families and take home. The children were very proud of their work and were excited to be able to share their creations with the teachers and parents.

The final product of the story book had direct benefit to the children, a central goal of any photovoice project, as it helped parents and teachers to more fully comprehend the experiences and priorities of the children at Preserve Preschool. The teachers were able to use the information from the children's photo story books in the development of the school curriculum. One example of implemented by the teachers, was having children participate in photo documentation of learning projects. Photo documentation is regularly used in preschools as a way to show parents and other stakeholders what children are learning. Typically, teachers take the photos and develop the documentation/display panels. After seeing the final photo books, the teachers started to include the children in taking the photos and developing the displays. The photo books also confirmed for teachers the importance of the hike to the students and ensured that this remained a central portion of the daily schedule. Student evaluations are not typically used in preschools and the photo book project allowed for concrete child feedback regarding the Preserve Preschool experience.

Photography results

The children took between 16 and 786 photos each. This includes those used and not used in the photo books. The number of photos each child used in their final books also varied from 3 to 22 photos. This corresponds with final percentages of total photos taken per child to photos selected for the book ranged from 3% to 28%. There were no noticeable sex or age-related patterns in the number of photos taken or percent used in the final projects.

Based on the photo included, the child captions, and discussion with the adult researcher while creating the photo story books, the adult researcher identified five important themes about the children's experiences at Preserve Preschool. These themes carried across the photo books created by the children and were not isolated to a small number of children. Each theme will be discussed in depth in the next section.

Theme #1

Children valued the nature experience at Preserve Preschool. Plants and animals were the subject of 43% of the photo/caption pairs used in the books. Many children also used specific common names, which were not always completely accurate, for the plants and animals in the captions. Also, the captions typically indicated positive feelings they had for the plant or animal. In one example of a plant photo the child correctly identified the plant species and indicated his affinity for the plant. Henry, age 4 captioned his photo, "It's green. It's pretty. Mayapple."

Theme #2

Children valued the social experience at Preserve Preschool. Photos with other people in them accounted for 35% of the photos used in the books. The people included teachers, classmates, parents, volunteers, and the adult researcher. This centers the social experience as an essential component of Preserve Preschool for the children. Matilda, age 4 selected a photo of her teacher and captioned the photo, "Ms. [Teacher]. She always thinks about trees. She loves rocks. Because I love her." These data show that while the nature experience was the most valued part of Preserve Preschool for the children, the social aspect was also very important.

Theme #3

The children viewed the hike as the most important part of the day. Images taken during the hike accounted for 66% of the photos in the books. Photos taken during the hike included photos of plants, animals, ground, sky, and other people in the class. Additionally, some captions, such as Celeste's, age 6, directly referenced the hike in her self-written caption, "I LiKe hiceng" [I like hiking]

Theme #4

The children valued the hike more than free play in the playscape. While each day began with free play on the playscape, only one photo of the playscape appeared in the final books. Observation of the children by the adult researcher indicated that the children enjoyed playing on the playscape, but its absence from the books indicated that this part of the day was not central to the story of Preserve Preschool from the children's perspective.

Theme #5

Preserve Preschool is not about the indoor classroom experience. While the children spent one-third of their day inside the traditional preschool classroom, only 9% of their overall photos were of the inside space and only 9% of indoor photos were used in the final books. This indicates that the indoor time is not central to the story of Preserve Preschool.

The themes found in the child researcher photos and captions gives voice to the children's value of the aspects of learning that were unique at a nature preschool versus a traditional preschool. The total number of images and book inclusion of the outdoor photos in wild spaces demonstrate how valued nature was for the children. The children's inclusion of specific flora and fauna and the social aspects of being in nature demonstrates the children's ability to represent layers of their experiences. The children's pride in their knowledge and ability to communicate their expertise through their books to the community reflects the value of photovoice to engage and give voice to even young children.

Children as co-collaborators in research

The preschool children proved to be highly capable in participating as co-collaborators, data collectors and data analysts in a photovoice protocol. All 15 children took photos in this study and 14 completed the entire process. The child who dropped out of the study quit at the photo book stage because she "loved" all of her photos and was uninterested in selecting those that best told her story of preschool. The adaptations to the four stages of photovoice discussed above were minimal. Stage 1 was shortened and included parents in the consenting and understanding the overall research goals outcomes. The child researchers engaged in the research framework when the adult researcher's photo story was shared as an example of their task. Other adaptations to the traditional photovoice process with young children included the elimination of the group sharing and discussion of photos. The children were very proud of their personal work but were not developmentally ready to engage in the work as a group project. Similarly, in stage 3, the children's developmental skills related to classification did not allow for the children to be involved in the final stage of analysis. To preserve the voice of the children as much as possible the adult researcher kept to descriptive codes derived from the children's word captions.

The children's proficiency at taking photos framed by the research task and their ability to make and explain decisions during selection and contextualization, indicate photovoice is a methodology that should be considered when investigating research questions that affect the lives of young children. Children are members of our society who have little choice about their daily lives and experiences, in comparison to the independence enjoyed by adult members of the community. This research demonstrated that the experiences of this marginalized group can be better understood through photovoice protocols.

Photovoice and urban youth

In this example, the authors share the use of photovoice for conducting a Youth Participatory Action Research (YPAR) research study. As the name indicates, YPAR is a sub-genre of PAR that is conducted specifically with young adults—generally young adults in their teens. The range of YPAR research settings is as great as the locations where collections of teens are found. Although it may be difficult to believe teens need greater voice, their volume does not necessarily equate to decision-making power. The pedagogy of YPAR and photovoice shifts studies around to embrace the youth participant's point of view. YPAR does this by relying on the people who have the closest relationship to the oppressive situation (Caraballo et al., 2017). Situated in YPAR is the argument that youth are in the seat of power and should engage as experts regarding their daily lives in the research process (McIntyre, 2008).

The youth environmental club: research setting and participants

This photovoice project took place during an afterschool environmental club in an urban high school. The research was an extension of work club had been engaged in focused on a local watershed. The group of middle school students, 12–14 years of age, and their teacher club leader worked with an urban watershed group to help clean up a creek portion. During their watershed the students talked. On several occasions the conversations included the importance of green spaces. As these conversations were happening, I, the adult co-researcher, became intrigued by their descriptions of the green spaces in their immediate neighborhoods. Their conversations about green space became the starting point for our photovoice research study. Formally our research question became: How do local urban youth, with limited access to nature, view the green spaces in their communities? Taking a youth participatory action research photovoice design (YPAR), the youth participants and adult researcher began to explore. The example that follows details use and power of a youth photovoice project for this middle school teacher researcher.

The photovoice process

One difficulty in all PAR, youth and photovoice is detailing the iterative nature of the research process, which although intentional and organized, flows and shifts in subtle ways. However, for the purposes of this chapter, we have aligned the process to the stages above. Some of the research framing of Stage 1 was described above. The participants were part of an established interest group of youth and their club leader and the focus emerged from questions that arose as part of their on-going activities. As co-researchers, we set out on a study to learn what we thought of our green spaces. It is notable in this example, issues of power and safety did not happen in this first stage, but they do come up later and will demonstrate the importance of these stage setting discussions at the start, even when the youth are photographing their own environments. Finally, in the stage setting, the audience for sharing the research was left implicit; the focus arose collaboratively from the group and met the purposes of the group. Although an external audience for the results is easy to imagine, it is sufficient to say the groups' interest and educational experiences met this step.

Taking pictures

Although the youth had access to phone cameras, these personal cameras were not used. The logistics of downloading select pictures and printing these for use was impractical. Further, the ethics of having the youth share pictures from personal devices when one of the co-researchers, the club leader, was a mandatory reporter created an additional power differential, and for a variety of other equity and security reasons personal digital cameras were just not an appropriate choice. Therefore, disposable film cameras were distributed to each participant. This added cost in both the purchase of the cameras and developing of the film. Surprisingly however, many of the participants had never used a disposable camera. As stage 2 also requires, training the youth researchers to use cameras was essential. Additional training included a discussion about the ethics of taking photographs in public areas. We agreed to take pictures from multiple angles of each green space. However, no direction as to content for the photographs was given. This was intentional. As a co-researcher, I did not want to bias my youth researchers in their photograph choices. Photovoice should gain the point-of-view of the participants (Wang and Burris, 1997). Then, they set out to take photos of their neighborhood's green spaces.

Understanding green spaces

At the designated date, the youth researchers returned their cameras and the film was developed so Stage 3 could begin. The analysis and knowledge generating stage of the photovoice process was intense. I transcribed these recordings and shared them with youth researchers. The convergence of the photo sorts and conversation transcriptions created a better understanding of the participant's perspective.

Wang detailed the SHOWeD discussion method; the youth researchers modified the analysis to process. We used a previously learned school science sorting method. Our sorting process started by selecting and discussing individual photos. The discussions that happened during the card sorts were recorded. Based on the discussions, the youth researchers began placing the images into big themes. The themes came from the youths' discussions; the adult researcher only asked clarifying questions. As co-researchers, we discussed each theme and clarified them by placing photos into smaller categories. We repeated this process several times,

reframing our themes and categories. Eventually, the youth choose two significant themes (1) nature experiences in city parks and (2) nature experiences in large parks outside the city as their final organizing system. Categories within the themes included photos with fields, hills, trash, flowers, trees, people, safe space, and reverence for green spaces.

An important note, and one that became an organizing feature of our research, occurred when one young woman in the group went rogue and took photos of a large state park located several miles away from the city. This action was unexpected because the youth participants had all decided to take pictures of our neighborhood green spaces. However, she stated, "There are no safe green spaces in my neighborhood, and my parents and I go to this location at least once a week." For her, this was her accepted green space. As discussed previously, the value of YPAR is the shift to accepting the youths' lived experiences and not reinforcing existing oppressive forces.

What we learned from our photovoice analysis speaks to the youths' urban life context. We discovered that our language and expectations, represented by our photo perspectives, depend on the green space's (park) location. We found the parks away from the city were revered. The youth said they noticed the beauty in such large green spaces. They felt more inclined to go and play in nature; splash in the creek. One participant referred to this as the "Real woods." In contrast, the parks in our urban communities had these hidden rules. Don't climb the trees! Stay off the grass! People in the city seemed more likely to litter in these small green spaces. Over and over, participants spoke about not feeling safe in urban parks.

Youth as co-researchers

As mentioned earlier this project did not have a stage 4, sharing results, as the nature of a club activity tied to a school calendar made this too complicated. There was not an explicit audience to share this research project with, but the power of the findings was real. These youth tell us, urban spaces are not real green spaces. Real green spaces do not belong to urban youth. Urban green spaces are dirty and have rules. Real green spaces are places of beauty and freedom. These contrasts could and should be shared with policy makers—city, county and state parks boards. City planners and police departments need to own the reality that urban youth do not feel safe in city parks. However, in this instance the learning experience was closer to home, for the youth researchers and the adult co-researcher.

Using the YPAR photovoice method with this group of urban youth was a learning experience that far exceeded my expectation as a researcher. Wearing the dual hats of a practitioner and a researcher frequently became a blurred line. As a researcher, I tried to be a keen observer with a systematic approach to the given situation. However, the practitioner's side of me struggled with already knowing my co-researchers and wanting to overlay my formulated ideas about urban culture and green spaces. This photovoice allowed me to hear my co-researchers' voices and see the world from their point of view. It required my overruling my biases and allowing the youth to lead the discovery.

In this example, we see how urban youth begin to notice and define their green spaces, a simple circumstance. However, this moment should be seen as the beginning of their self-consciousness, a pivotal moment when the youth begin to see green spaces as a reflection of how they are valued in the community at large. According to Freire, the highest level of critical conscientiousness is to be aware of one's assumptions and make a change or keep the current situation (Freire, 1970). This experience enabled the youth to realize the inequities of their surroundings.

Conclusion

The power of photovoice is its ability to give voice to marginalized communities, whether the communities are the women in rural China, girls in Indian schools, children of LGBTQ parents, or the preschoolers and urban youth discussed in this chapter. The intellectual foundations of photovoice arise from critical and feminist theory that seek to expose the unspoken structures that empower some while marginalizing others. The beauty of photovoice is in the power of the image to illuminate the unseen and obscured structures that support an unequal status quo. For many photovoice participants it is only when they come together to share their personal images that societal patterns are exposed, and the participants can act on their new knowledge.

How we define marginalized or oppressed may vary but as a participatory research method, photovoice gives voice and power to each group. Young children at a private preschool in the future may not suffer as an outsider group. However, now, as young children it is adults who observe and interpret children's actions in the world and decide what they like best. What we learned from this study was the preschool children were capable researchers and actors in their world. We learned that they valued their time in nature, they found beauty in nature, and there was freedom in wild spaces. We also learned the children found nature to be a social space, one to share and enjoy with others. The nature preschool was a safe space for the children.

In contrast, urban youth are frequent subjects of critical research and are often involved in YPAR and photovoice. As this case demonstrates, photovoice was a powerful mechanism for introducing these youth to the hidden structures that define their world, as they, like the preschoolers, explored their relationship with nature. The urban youth also expressed the beauty of nature and the freedom to be found in wild spaces. In contrast to far away wild spaces, the urban parks they had regular access to were not safe, beautiful, or places to be free. Our urban youth researchers stepped into the critical realm with photovoice. With photovoice, they were able to see their community in a new way.

The power of photovoice as a qualitative research method is in its adaptability to meet the needs of the researchers, no matter their age, education, or photography skill. When Wang and Burris (1997) presented the photovoice method, they had experienced its power to empower rural women to define and set policy around their health needs. Since then photovoice has been used across the world and across ages. We believe there is potential for photovoice to engage children, youth, and adults to explore nature and the natural world as we seek to develop an equitable and sustainable future for all.

References

- Abma, T., Schrijver, J., 2020. "Are we famous or something?" Participatory health research with children using photovoice. *Educ. Action Res.* 28, 405–426.
- Agarwal, N., Moya, E., Yasui, N., Seymour, C., 2015. Participatory action research with college students with disabilities: photovoice for an inclusive campus. *J. Postsecond. Educ. Disabil.* 28, 243–250.
- Baillie, P.E., 2016. Nature preschools: the cross fertilization of early childhood and environmental education. In: Sobel, D. (Ed.), *Nature Preschools and Forest Kindergartens: The Handbook for Outdoor Learning*. Redleaf Press, St. Paul, MN, pp. 45–58.
- Caraballo, L., Lozenski, B., Lyiscott, J., Morrell, E., 2017. YPAR and critical epistemologies: rethinking education research. *Rev. Res. Educ.* 41, 311–336.
- Clifford, J., Marcus, G., 1986. *Writing Culture: The Poetics and Politics of Ethnography*. University of California Press, Berkeley, CA.
- Cook, K., 2015. Grappling with wicked problems: exploring photovoice as a decolonizing methodology in science education. *Cult. Stud. Sci. Educ.* 10, 581–592.
- Copley, J., 2010. *The Young Child and Mathematics*. The National Association for Education of Young Children (NAEYC), Washington, DC.
- Dell, L., 2011. *Critical Childhood Theory: A Paradigm Shift in Educational Research With Young Children*. Unpublished manuscript.
- Dell, L., 2018. *Nature Preschool Through the Eyes of Children*. PhD Thesis. University of Cincinnati, Cincinnati.
- Fine, M., 1994. Working the hyphens: reinventing self and other in qualitative research. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 70–82.
- Freire, P., 1970. *Pedagogy of the Oppressed*. The Seabury Press, New York.
- Groundwater-Smith, S., Dockett, S., Bottrell, D., 2015. *Participatory Research With Children and Young People*. Sage, Los Angeles.
- Latz, A., Phelps-Ward, R., Royer, D., Peters, T., 2015. Photovoice as methodology, pedagogy, and partnership building tool: a graduate and community college student collaboration. *J. Publ. Scholarsh. High. Educ.* 6, 124–141.
- Linton, R., 1989. Towards a feminist research method. In: Jaggar, A., Baroda, S. (Eds.), *Gender/Body/Knowledge*. Rutgers University Press, New Brunswick, NJ, pp. 273–292.
- Maguire, P., 1987. *Doing Participatory Research: A Feminist Approach*. Center for International Education Amherst, MA.
- Mackenzie, S., Talbott, A., 2018. Gender justice/gender through the eyes of children: a photovoice project with elementary school gender expansive and LGBTQ-parented children and their allies. *Sex. Educ.* 18, 655–671.
- McKenna, M., Greene, S., Burke, K., 2017. Teaching from the outside in: community-based pedagogy's potential for transformation. *Transformations* 27, 241–257.
- McIntyre, A., 2008. *Participatory Action Research Series 52: Qualitative Research Methods*. Sage, Los Angeles, CA.
- Meyer, H., Kroeger, S., 2005. Photovoice as an educational action research tool. *Qual. Res. J.* 5, 185–194.
- Pollock, S., 2019. Literacy difficulties: what are learnings' experiences? *Educ. Child Psychol.* 36, 101–114.
- Shah, P., 2015. Photovoice and the reimagination of girls' education in India. *Comp. Educ. Rev.* 59, 50–74.
- Sontag, S., 1977. *On Photography*. Farrar-Strauss- Giroux, New York.
- Stringer, E., 2022. *Action and Community-Based Participatory Approaches to in Qualitative Research*.
- Wang, C., 2006. Youth participation in photovoice as a strategy for community change. *J. Community Pract.* 14, 147–161.
- Wang, C., Burris, M.A., 1997. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ. Behav.* 24, 369–387.
- Wang, C., Redwood-Jones, Y., 2001. Photovoice ethics: perspectives from Flint photovoice. *Health Educ. Behav.* 28, 560–572.

Participatory photography: high school students' awareness of privileges in Nordic welfare states

Gunilla Holm, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

© 2023 Elsevier Ltd. All rights reserved.

Visual research and its place in research in education	270
Participatory photography	271
Participatory photography as a research method—a case study: high school students' awareness of privileges from living in a Nordic country	272
Methods and data	273
Students' perceived privileges regarding living in a Nordic country	273
Nature	274
Welfare state privileges	275
Equality	277
Freedom and safety	279
Carefree youth	281
Living in a bubble	281
Concluding points of interest in participatory photography	282
References	283

Visual research and its place in research in education

During the last decades there has been an explosion of visual images largely due to the rapid development and availability of mobile phones with cameras for producing photographs and videos. There are billions of photos and videos shared on social media every day. The production of photographs in research with the help of mobile phones or traditional cameras has been slower even though there has been an increase of research using visual data and visual research methods. In recent years there has been a plethora of books and articles on visual data and methods (e.g., Beeli-Zimmermann et al., 2020; Holm, 2018; Leavy, 2019; Mannay and Pauweis, 2019; Pienimäki, 2021; Rose, 2016; Moss and Pini, 2016). At universities courses on research methods now often include a lecture on visual methods as well, but few research studies in education are so far based on visual methods. This chapter briefly contextualizes the field of visual research in education and then focuses on the use of participatory photography as a qualitative method when conducting research in education.

There are two major kinds of qualitative approaches used in visual research in education namely still images and videos/films (Prosser, 2000). Both visual research approaches have their roots in anthropology and sociology (e.g., Collier and Collier, 1986; Flick, 2002; Harper, 2000). The most common use of videos in educational research is videos of interactions in classrooms made by a researcher. There are several big international video laboratories often collecting videos about teaching and learning in specific subjects like language and mathematics (see e.g., Clarke's International Center for Classroom Research <http://www.iccr.edu.au/> and Klette's Teaching and Learning Video Lab http://www.uv.uio.no/ils/om/tlv_lab/). A trend gaining in importance in video-based research in education is the micro-analysis of video data. Micro-analysis usually focuses on one (or a couple if several cameras are used) participant's views or a particular type of interaction, setting or content (Rusk et al., 2014). Participatory video is still not common qualitative method in educational research but is becoming more common with young people outside of school (see also Burns et al., 2021). Much of the participatory video research takes the form of video diaries (see e.g., Pini and Walkerdine, 2011).

The still images can be, for example, drawings, cartoons, paintings, constructed social media images, or photographs (see e.g., Farber et al., 1994; Jipson and Paley, 1997).

Photographs used in research can be pre-existing from sources such as archives (Grosvenor, 2007a, 2007b), family albums (Mitchell, 2011) or magazines (Holm, 1994) or produced by the researcher or the participants particularly for a research project (Holm, 2018). What kind of photographs and approaches are used depend on the interest and research approach of the researcher and the purpose of the research. The commonly used photo-voice focuses on social change and social justice, where the participants photograph, for example, problems or injustices in their communities with the goal of community action (Sutton-Brown, 2014; Wang, 2006; Wang and Burris, 1997). Some researchers prefer a more arts-based approach (Knowles and Cole, 2008). Holm, Sahlström and Zilliacus (2018, p. 312) describe arts-based visual research as a continuum since "At one end, there is a strictly art-centered approach that emphasizes intertwining of making visual arts and doing research. At the other end are researchers who use visual art and artifacts as fruitful data, and who see arts-based research as a methodological tool."

Visual ethnography in education can take many different shapes (see Holm, 2018 for an overview), but most commonly photography (or occasionally video) is used as a method alongside participant observations and interviews. The photographs in this case

constitute data in themselves or can be used in connection with other data. Especially participatory photography with students producing their own photos is becoming more common in ethnographies in education.

Another approach and the most frequently used approach, where the researcher or the participants produce the photos, is the photo-elicitation interview. Photo-elicitation interviews are sometimes used alone or as part of a larger study as one of several methods. Most often it is the participants taking the photographs. Rose (2016) provides a thorough and detailed description of photo-elicitation interviews as an approach and a method. She describes them as a way to obtain different kinds of insights and to explore everyday life as well as empower participants since they are the experts on their own lives. Even though photos are taken for the interview and the photos and interviews analyzed together, the focus tends to be on the interview. The reporting and publishing of results focus on the interviews and describing the photographs, but usually do not include many of the actual photographs.

Participatory photography is similar to photo-elicitation, but still different since the focus is more on participants as co-researchers and on the photographs as a data source in themselves alongside other possible data sources such as interviews, observations, and participants' writing. It is a method familiar to students of all ages with mobile phones and an easily available method for all researchers in education. Participatory photography as a visual research method is discussed in-depth in this chapter and the production, analysis, and reporting of photographs as research data are discussed through an example of a study using participatory photography as a qualitative research method.

Participatory photography

Participatory photography is a way to engage the participants in producing data that they think is relevant for the research task presented by the researcher. Hence, the interest lies on the participants' views, perceptions, and experiences regarding a particular topic or their lived realities. Participatory photography can be a way to engage particularly the withdrawn and marginalized students (Cappello, 2005; Darbyshire et al., 2005; Lodge, 2009). However, it is not a method for only a particular purpose or group of participants like photo-elicitation interviews that focus on community action. If the researcher and the research participants do not share a common verbal language, participant photography also creates a common visual language (Veintie and Holm, 2010). The power differences between the participants and the researcher do not disappear, but are somewhat evened out, and the participants become more like co-researchers. The researcher still decides on the overall task and does most of the analysis and reporting. However, some analysis and reporting/publishing can also be done together with the participants (Holm, 2008a). If the participants are students, they can, for example, do the reporting of the results to the entire school. At times, participants also take more charge of the research task by interpreting the task in ways not initially intended by the researcher, even if the task is clearly stated (Croghan et al., 2008). Sometimes a participatory photography project is produced from the beginning to the end collaboratively with the participants as in Holm's study (2008a) conducted with doctoral students about what it meant to be a doctoral student.

Participatory photography is rarely used as the only data production method. Having photographs as the only source of data is too susceptible to misinterpretations as the photographer's intentions are not always clear to the researcher from only the photograph. Non-figurative and symbolic photographs in particular can be difficult to interpret. To avoid misinterpretations, captions for the photographs usually clarify the photographer's intentions (Holm et al., 2015). The photographs in participatory photography are rarely snapshots but are often carefully constructed to communicate the photographer's views and perceptions about the task assigned by the researcher. Often, semi-structured and open-ended interviews are used to complement participatory photography because the two methods complement each other by providing partially different kinds of information about or different perspectives on the same issues (Holm et al., 2015). In interviews, the photographs can be used to facilitate or prompt discussion about certain issues more like in photo-elicitation studies, or the focus of the interview can be the same as in the participatory photography assignment, without explicitly using the photographs. In Holm's (1997) ethnographic study of pregnant and parenting teenage girls a photography project was done parallel with participant observations and interviews. Together they showed the complexity of the girls' lives. The photographs told a story about the carefree and happy girls they wanted to be while the interviews were about their difficult life circumstances. In addition, the participant observations told the story of a strictly controlled and managed daily school life with rote learning while caring for their babies. Thus if participatory photography is used in ethnographic studies, it can be combined with participant observations and documents or artifacts in addition to interviews providing possibilities for data triangulation (see also Holm, 2018 for further examples).

Participatory photography research in educational settings requires substantial planning partially due to that research permissions are usually required by municipalities, research institutions and school due to that young students are involved. In addition, consent forms from guardians and teachers as well as consent or assent forms from students depending on their age. Consent forms need to detail the different uses of the photographs like used for analysis, publishing, conference presentations and in teaching. Confidentiality issues are of high importance when using photographs in research. The permission process can be quite demanding if so-called sensitive issues like bullying or racism will be studied. Depending on the situation consents can be given on paper or digitally via smartphones. For participants where everybody does not have a smartphone, students should be able to borrow phones or cameras from the researcher. For projects with students as co-researchers/photographers it is necessary to collaborate with one or several teachers. Teachers can help especially with informing and obtaining consent from guardians. As in ethnographic studies, the researcher must obtain the trust of the participants and have to be careful in creating alliances and cooperation.

Participatory photography research needs to be introduced and explained to students and here too teachers are key to giving class time for introducing a project to students. The process with students giving consent or assent and also the asking of consent from those being photographed is of utmost importance. In general, ethical principles need to be discussed throughout the process and the respect for other participants' photographs need to be emphasized if photographs are shown to others.

A participatory photography project usually starts with an introduction of the process and the topic of the study by the researcher, then students can discuss the topic in small groups to stimulate students' thinking on the topic. Students usually need two-three weeks to compose their photographs. A longer period tends to mean that students forget about their photographs. Today the photos are in most cases taken with mobile phones, which means that it might be necessary to set a limit to the number of photographs that can be submitted depending on the number of participants.

If the teacher allows, it is useful to have a time in class when participants write the captions for each photo. The photos with captions can be collected in class on a memory stick or submitted electronically directly to a safe repository depending on the circumstances. If a trusting environment has been created in the classroom students can view others' photographs and together look for themes and patterns in the photographs. Hence, a first analysis can be done with the students. If possible, it is also good practice to return to the students and schools with the research results if they can benefit the schools and students involved. However, here confidentiality takes precedence. If several schools and several classes are involved, it is possible to exclude the photographs of the school to be visited if the topic is sensitive or if it is possible to recognize someone. If other kinds of data such as interviews or participant observations have been collected, they can all be combined in the final analysis and reporting. However, in research with participatory photography the photographs are not mere illustrations, documentation or supporting data, but data in their own right that are analyzed both separately from and together with other kinds of data.

There are many types of analysis that can be done of photographs. [Rose \(2016\)](#) provides a review of some methods like discourse analyses and a psychoanalytic approach. However, in educational research the photographs are usually analyzed in the same way as the interview and observational data. Qualitative content analysis, thematic analysis or discourse analysis are often used, and photographs are categorized, organized into themes or discourses like the verbal data.

Several things can influence what is photographed, such as the time of the year. Projects in wintertime result in very different photographs than those taken in summertime. Outdoor photos are fewer and darker during the winter ([Holm, 2018](#)). If it is not possible for the participant to make the photographs, they want due to the time of the year (for example, Christmas traditions in summertime), it is quite common that the participant decides to use older photos not produced specifically for the research project. Another aspect that influences the subject of a photograph is consent rules. There are usually fewer photographs taken that include people since consent is needed from those photographed. This in turn leads to more symbolic photographs that are difficult to interpret without captions. Thanks to the possibility to give consent digitally via mobile phones, students are now slightly more likely to photograph people.

Participatory photography as a research method—a case study: high school students' awareness of privileges from living in a Nordic country

In this chapter on visual research in education, I will use a recent study on high school students' thinking about living in a Nordic country as an example through which I explore the use of participatory photography as a research method. I will use the example study to bring up important and interesting issues in participatory photography. The purpose of the overall study was to explore, in a cross-cultural way, how high school students in the five Nordic capital regions view living in a Nordic welfare state. In this chapter, the focus is on one of the study's research questions: Are students aware of the privileges afforded them by living and going to school in a Nordic country and if so in what ways?

The study was conducted using participatory photography with complementary small group interviews exploring students' perceptions in depth ([Pink, 2012](#); [Rose, 2016](#)). For the methodology I use my previous research on participatory photography, arts-based visual research, and visual ethnography as a foundation ([Holm, 2018, 2020](#); [Holm et al., 2015](#), [Holm et al., 2018](#)). The focus in this chapter is specifically on exploring how young people can express themselves using participatory photography as a research method. The overall research purpose is quite abstract and not easily answered. It required the students to reflect upon their own lives and societies. The focus is on students' awareness of and views on their privileges as all the Nordic countries—Denmark, Finland, Iceland, Norway, and Sweden—emphasize equality and equity in education and all aspects of life, although very little research has been done on young people's perceptions and awareness of their privileged positions compared to other young people in the world.

Our theoretical framework for the study was youth identifications ([Ahmed, 1996, 1997](#)), as we were interested in how the students positioned themselves in their respective countries. We used the term 'identifications' instead of 'identities' to indicate that we considered 'identities' as constantly developing and changing in relation to a sense of belonging to cultures, communities, and personal relationships ([Anthias, 2011](#); [Hall, 1996](#)). We were interested in whether students' identifications extended beyond the national borders to a more transnational Nordic level. Hence, the underlying idea here is to show that participatory photography can be used in studies with theoretical frameworks and research questions that are seemingly unconnected to visual research. Photographs are students' constructions of their views and thoughts and thus closely connected to their identifications (see also [Croghan et al., 2008](#)).

Even though young people today have a lot of freedom and many possibilities to form their own identifications, they are also constrained by inequalities such as racism, classism, and sexism. Various individual abilities and structural constraints interact and influence how young people's identities develop ([Anthias, 2011](#); [Gewirtz and Crib, 2008](#)). Schools, with their relationships and

structural opportunities and constraints, are crucial for young people's sense of who they are. Therefore, we chose to do our study in high schools. Furthermore, we chose to explore young people's views in all the five Nordic welfare states as they are often considered quite similar, but very little is known about how high school students view their own realities. We were interested in whether students perceived a sense of belonging or community also with the other Nordic countries. Did they think in terms of the Nordic countries as 'we' and 'us', in terms of 'us' (their home country) and 'others' (the other Nordic countries), or not at all in terms of 'us' (their own identification with neither their own country or the other Nordic countries) (Reay, 2009).

Methods and data

The participatory photography project was conducted in collaboration with one or two teachers in each school and sometimes in connection with another class activity. Participatory photography generally works best with a person on-site assisting in organizing the process. In this case, it was necessary to collaborate with teachers, since the study was conducted in five different countries. The project was integrated into class instruction in sociology, art education, social sciences, and in Swedish and Danish language classes.

We visited each school two or three times. We obtained research permission from principals via email (and in some cases also from school districts). The principal contacted teachers interested in our study. We then visited the schools to present and set up the study, as well as obtain written consent from teachers and students. All consent forms were collected electronically. We also emphasized that the students themselves could withdraw from the project without any consequences if they wanted to do so at some point. On our first visit, we held a brainstorming session with students about their thinking about living in the Nordic countries and about what they associated with or perceived as important regarding living in a Nordic country. We also brought up some basic issues about photography like light and distance, even though all of them were familiar with mobile phone photography. We also informed them about their responsibility to obtain consent of the person photographed to include the photograph in our study. People photographed could also give their consent electronically via their phones. We emphasized that the people photographed had the right to withdraw their consent for a photograph if they wanted to. This requirement of consent led to students not photographing many people. However, there were photographs, mostly of other students, that had to be excluded from the study because consent had not been obtained. Even if a person has given permission to use a photograph of them in the research and publications it is always the researcher's responsibility to consider which photograph that can be published since not only faces, but also environments and buildings like schools are identifiable.

Students were grouped into small groups to more easily keep track of their consent forms and photographs, as well as for conducting the small group interviews later. They were given 1–3 weeks, depending on the schools' schedules, for constructing their photographs. Students could take as many photographs as they wanted about what it meant to them to live in a Nordic country, but could only upload ten photographs (including captions) to the project's platform to keep the number of photographs for analysis manageable. Teachers helped organize the conversion of photographs to PowerPoint slides, which was done to make it easier to add captions. After a preliminary reading and analysis of students' photographs and captions, we returned to the schools to conduct the interviews in small groups (3–5 students). The semi-structured interviews about what it meant living in a Nordic country lasted approximately half an hour. In the small group interviews, students could choose to discuss their photographs further or discuss other Nordic issues as a group. The small group interviews brought up new issues, showed that there were differing views, and provided more extensive descriptions about some issues related to the photographs (like the quote on page in section [Living in a Bubble](#)). After doing a thematic analysis (Braun and Clarke, 2021) of the photographs and the interviews we returned to the schools a third time to present the results to the participants, but in some schools many other students and staff also participated in the presentation of the results.

The data for this chapter consist of 700 photographs, taken in 2018, and the accompanying 33 small group interviews. A total of 178 upper secondary students from seven upper secondary schools in metropolitan regions of Copenhagen ($n = 33$), Helsinki (three schools, $n = 61$), Oslo ($n = 30$), Stockholm ($n = 24$) and Reykjavik ($n = 30$) participated in the study. The average age of the students was 17. The school language and home language were the same for most students, but in two schools there were many students with another ethnic/language background than Danish, Finnish, Icelandic, Norwegian or Swedish. There are three schools in Helsinki as Finland is a bilingual country, and thus one school was Finnish, another Swedish, and another Finnish school was included that represented more ethnic diversity.

There are a variety of methods that can be used to analyze visual images, but for photographs produced by participatory photography qualitative content analysis, discourse analysis or a thematic ethnographic approach (Holm, 2018; Rose, 2016) are common. In the Nordic study, a thematic ethnographic approach was used with the purpose of exploring whether major themes emerged. All photographs and captions were viewed and read several times (Keats, 2009) and then grouped into categories like 'nature' and 'equal rights'. It was necessary to read the captions together with the photographs to understand students' intentions with the photographs. The categories were then further analyzed. For example, the category nature was broken down to subcategories like clean nature, accessible for all, and general nature photographs. After many readings/viewings, privileges emerged as a theme stretching over multiple categories, like clean and accessible nature, rights, welfare state advantages like high-quality education, and infrastructure. A similar analysis was conducted of the recorded and transcribed group interviews. Thereafter, the themes from the photographs were compared to the themes of the interviews and combined when they overlapped. One of the major results of the study was that students in all five Nordic countries perceived that they were lucky to live in a Nordic country for various reasons that are detailed in section [Students' Perceived Privileges Regarding Living in a Nordic Country](#).

Students' perceived privileges regarding living in a Nordic country

There were six subthemes emerging under the theme of perceiving oneself as privileged due to living in a Nordic country. Even though there were variations between students from the different countries, the themes discussed below emerged in the

photographs from all the five Nordic countries. The subthemes are related to nature, welfare state privileges, equality, freedom and safety, carefree youth as well as a critique of what some of the consequences might be of living in a privileged bubble.

Nature

Interestingly, and somewhat surprisingly, the clean and beautiful nature, easily accessible for all, quickly emerged in the students' photographs as one of the important privileges of living in a Nordic country. Today's youth is often thought to be more occupied with and interested in mobile phones and social media, but the students in all countries were very much aware of how privileged they were by living in countries with a lot of clean and beautiful nature nearby (Fig. 1).

Students frequently pointed out that in the Nordic countries, people spend time outdoors regardless of the weather. With appropriate clothing, it does not matter if there is ice and snow, rain, or sunshine (Fig. 2). Students expressed an appreciation for having been taught to enjoy spending time in nature and the beauty of nature. Clean lakes and forests full of edible things were also regarded as part of being privileged with close access to nature (Fig. 3).



Fig. 1 Nature is our thing (I=Iceland).



Fig. 2 In the Nordic countries we have a lot of forests (or woods?) and nature where we happily spend time (N=Norway).



Fig. 3 It is so cool that we can pick berries and mushrooms in the woods since our nature is so clean (F=Finland).

Welfare state privileges

Many students brought up that living in a Nordic country meant living in a welfare state with all its advantages. The social security was brought up by students in all the countries. Most important was the high-quality and free education system. Education is free from early childhood education throughout tertiary education. "I think we've got one of the world's best school systems and we've also got free education, so that will give us every option in the world." (D = Denmark) Photographing high-quality education and

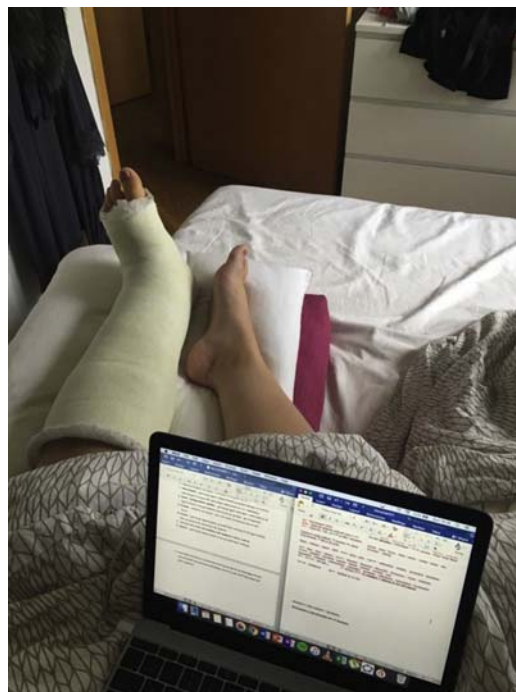


Fig. 4 I've had many many many many many surgeries and I'm still (in the same place) in the school system as my friends. I think we have the privilege that we can do that. Because not in every country you can. In every Nordic country, this is possible (I).

at the same time protecting the confidentiality of the school, students and teachers is difficult. Consequently, photographs of schools are not included in this chapter as they may enable identification of the schools.

With the help of a school system equipped with technology and a well-functioning healthcare system, students do not fall behind in school even if ill. Access to technology is also emphasized in other contexts as a privilege (Fig. 4).

Another student said, "Thanks to the Nordic welfare states, people have equal possibilities to live a good life and to choose the kind of life they want to live." (D = Denmark).

A well-functioning infrastructure including good public transportation was appreciated, especially because students in the Nordic countries are largely dependent on public transportation to get around (Fig. 5).

Especially in Denmark, but also in most of the other countries, bike lanes make biking the major alternative to public transportation for young people. Many students bike to school, as seen in Fig. 6.

Interestingly, the availability and priority of both healthy food and clean water were also connected to the privileges of the welfare state (Fig. 7).



Fig. 5 Good public transportation. The subway. The development of the infrastructure is good (S=Sweden).



Fig. 6 It is Nordic to bike everywhere and keep nature clean (D).



Fig. 7 Clean water.

Equality

Students in all countries perceived that equality is of high importance and were very proud of the fact that equality is an emphasized policy and characteristic of the Nordic welfare state. Gender equality was emphasized through photographs of fathers taking parental leaves to care for young children and also by expressing respect for women (Fig. 8).

As one student said:

"As soon as a man disrespects a woman, everyone is like "What the fuck did you say?!". Everyone is standing up and say: "This is not right!". And maybe you cannot see this in every country." (I)



Fig. 8 Gender equality is Nordic (man with a baby carriage) (D).

There were many photographs of rainbow flags, parades and even houses indicating equality for sexual minorities (Fig. 9).

A student in Iceland explained the support for the gay pride festival as “it is just like a family day, a family holiday, and everybody joins in to show our support and we support everything and everyone.”

Also, children’s rights and the rights of people with special needs were mentioned (Figs. 10 and 11).

Students pointed out that the Nordic countries have a Christian heritage, but that today there is freedom of religion (Fig. 12).



Fig. 9 Everybody is accepted (Fswe = Finland Swedish).



Fig. 10 Children have the same rights and value as adults (F).



Fig. 11 It is Nordic to be disability friendly and to accept people as they are (D).



Fig. 12 Freedom of religio (F).

Freedom and safety

A society emphasizing equality and equal rights is also connected to freedom and safety. Students in all countries talked about the freedom and safety in the Nordic countries, but the Icelandic students insisted on Iceland being the safest (Fig. 13). "It's a stress-free environment. We don't have an army, and the cops don't have guns, or they have guns now, but they have only once fired it ... not having to worry about some things that people in other countries have to worry about."



Fig. 13 Our environment gives us a sense of calm and security (F).

Another student pointed out that “It is calm and nice in the Nordic countries, sometimes so nice that we forget about the places in the world where it isn’t as nice.” (D)

A student from Finland said “I feel safe to live in Finland. You can walk on the street without being scared and you can freely express your opinion.” (F)

However, there were also some critiques of not taking advantage of the freedom, but instead everybody follows the familiar path with few going their own way (Fig. 14).



Fig. 14 We are all very similar in the Nordic countries. We are all walking in the same direction (D).



Fig. 15 Going to sunny places on vacation (N).

Carefree youth

In photographs and interviews, many students expressed that they could live a carefree youth. They know their education is free and that they can take a sabbatical year after finishing high school and still get back on track to education and employment later due to the social welfare system. After school, most of them socialize with friends or participate in hobbies. Only in Iceland did some students say they work after school to earn money for themselves. Norwegian and Danish students in particular described how they can afford to frequently go on vacations abroad (Fig. 15).

However, the carefree upbringing and a privileged life was also criticized by some students. As an example of the backside of a carefree youth the partying and drinking culture among young people in the Nordic countries was brought up and photographed by students in all countries.

Living in a bubble

The freedom and safety in the Nordic countries are questioned by some students as creating societies where homogeneity is emphasized and where the majority norm should fit everybody, even though the Nordic societies are at the same time argued to be open and where people can be themselves (Figs. 16 and 17).

Living in a safe and privileged bubble was brought forth and critically analyzed by students in Denmark, Iceland, Norway and the Swedish-speaking school in Finland. The privileges and sense of freedom and security can make people care less about others. Two students expressed the feeling of living in a bubble as follows:

"The Nordic countries are like bubbles of safety. There may be many tragedies outside the bubbles, but the bubbles remain secure." (D)

"I was thinking about the Nordic countries as one big bubble of security. Outside Scandinavia so many terrible things can happen, but it will remain secure inside our countries. And I feel like that's a big part of our identity that we don't have a lot to be scared of. But maybe for that reason, we are a bit like 'them and us', so it's like inside the bubble, we are taking care of each other, but we don't think as much about outside the bubble because we know it's secure inside the bubble, so we stay there. (D)

A student from the Swedish-speaking high school in Finland summed up their, and many other students', privileges in the following way:

"Being a nordbo (= person from the Nordic countries) is actually being part of the top one per thousand of the whole world, I'm set for life. I don't actually need to do anything that requires like a whole lot of anything, and I'm still going to live a significantly better life than 99.99999 percent of people. I was just born into this. I was just born into a little, nice, white family like this that lives in Finland, and I'm a white man, so I, like, have it all. I've been raised in a good family, a good environment, I've always had it well, good, I've gone to school for free, been educated, speak multiple languages fluently, like know multiple kinds of things that other people educate themselves their whole life for. I'm 19. I keep better track of, like, the whole world than the majority of the world, and that's *only* because I've been born into this. Nothing else. It's such a huge blessing, to live here. And I don't have any interest in moving away. Like sure, I want to travel the world, see what it's like in other countries. Maybe live somewhere else a few years. But I'm probably going to die here. I've already traveled enough to be of the opinion that nothing better exists. Like, you know, it's much more equal, safe, and good on average in Finland or in Sweden, Norway or Denmark, whatever, than in other countries. And like, I'm cool with that. I'm going to live here, actually, I don't have a reason to move away from here." (Fsw)



Fig. 16 People in the Nordic countries are relatively homogeneous. We are all alike and could as well come prepackaged in IKEA boxes. Build your own NORDBO! (Fswe).



Fig. 17 We live in our own small, privileged world (D).

Concluding points of interest in participatory photography

When analyzing photographs, it is important to look at what is missing since it can be equally important as what is included. In our study, for example, despite two ethnically diverse schools included in this study, there was only one photograph referring to ethnic or racial diversity, namely a photograph of two Somali girls with the text “Odd to belong to two worlds. Between two worlds.” In addition, even though there were photographs about the importance of equality, quality and equal access, there were no photographs about inequality, discrimination, prejudice, racism, or sexism. Equal opportunities were mentioned, but it was never questioned whether equal opportunities are the reality in the Nordic countries. For example, gender equality is photographed and mentioned, and gender equality is the official policy. However, generally men take shorter parental leaves than women do, and there is a substantial difference in women’s and men’s earnings. The reason for the lack of these kinds of photographs may be due to how the task was presented or that students did not feel comfortable with bringing up the more problematic aspects of living in the Nordic countries in the school context. It is also possible that it is difficult to photograph problems like prejudice and racism. The absence of a more critical view examining if privileges are afforded everybody living in a Nordic country independently of race, ethnicity, gender and social class would need to be investigated either by adjusting the participatory photography task by asking

students to focus specifically on inequalities or by more in-depth interviews. However, students in other studies have relied on symbolic photographs for things that are difficult to photograph, such as micro-aggressions against language minorities (see Holm, 2018; Holm et al., 2015). These kinds of symbolic photographs are usually discussed further in interviews, but the photographs tend to be a way to bring the difficult issue up for discussion. What is missing and absences are known to be difficult to photograph. Sometimes the photograph shows the presence of something or someone, while the caption explains that what is shown is missing. For example, in a study of doctoral students' daily life international students submitted photographs of their families and their children, which lead the researcher to interpret the photographs to mean that their families are important in their daily lives. However, the captions explained that they missed their families daily and the photograph of the families showed whom they missed (Holm, 2008a). In the Nordic study it is evident that students have photographed things that are currently relevant for them and their lives like bikes, buses, and health care but not, for example, the equal access to high-quality day-care.

In the interpretation of the photographs, there are several layers of interpretation that need to be considered. First, the photographs are by themselves the students' interpretations of the task given and of their realities. My interpretations as a researcher are influenced by my interpretation of the task given by me, as well as by my academic and personal background, which in this case includes my Nordic habitus, experience living in a Nordic country, and familiarity with all the Nordic countries and capitals. Important questions to ask are whether my Nordic habitus pushes me to interpret the students' photographs and captions as more realistic and truer or naïve. In what ways are students' photographs expressions of their national and Nordic habitus? Mannay (2010) sees a researcher's common habitus with the participants as decreasing the distance between them but also that there is a risk for making faulty assumptions about having a better understanding of the participants' photographs. For both Bourdieu (1990a) and Sweetman (2009), photographs work well in research to describe abstract notions and perceptions as well as what participants take for granted. The final interpretation is made by the viewer or reader of the photographs and texts, and their habitus may influence their interpretation. Hence, the photographer has no guarantee that the photograph is interpreted as intended by the researcher or the reader of a published photograph (Holm, 2020).

Photographs are one kind of qualitative data in a study and are valuable on their own and do not need to be translated into text, instead they interact with the text. The researchers' translation of photographs limits the viewers to the researcher's interpretation of the photographs. As part of the data, photographs can be included in the reporting of the results and in relation to other data (see Holm, 2008b; Pink, 2012), as shown in the examples from the Nordic study. Publishing photographs as part of the results has become more common thanks to the increase of online journals. Still, in most cases the photographs are published as very small images, which diminish their power to speak to the reader. Further, photographs printed in articles and chapters are almost always printed as black-and-white photographs, something that most of the time changes what and how they communicate. Colorful photographs, such as the bowl of colorful blueberries (Fig. 3), does not send an equally powerful message when the photograph is printed in shades of gray.

Most people encounter a lot of visual images, photographs, videos, games, emojis and other visual messages daily. Our communication is not only oral and written, but also visual. In order to connect with research participants, especially with young students, including a visual method is often advantageous. In participatory photography, the information emerging from the photographs is often surprising (see Holm, 1997; Holm et al., 2018). In this Nordic study, we were surprised that, although all the students attended schools in urban or suburban environments, there were surprisingly few photos from urban environments and very many from natural mountain, forest, or lake environments. Nature is, of course, also mentioned in interviews, but I as researcher would not have understood the extent of students' appreciation of and importance placed on a vast and clean accessible nature as a major privilege of living in a Nordic country. In conclusion, participatory photography as a qualitative research method brings forth the participants' views and knowledge by leaning on the participants' own data production.

References

- Ahmed, S., 1996. Identifications, gender and racial difference: moving beyond a psychoanalytical account of subjectivity. In: Sharma, R. (Ed.), *Representations of Gender, Democracy and Identity Politics in Relation to South Asia*. Indian Book Centre, Delhi, pp. 359–386.
- Ahmed, S., 1997. It's just a sun tan, isn't it?: auto-biography as an identificatory practice. In: Mirza, H. (Ed.), *Black British Feminism*. Routledge, London, pp. 153–167.
- Anthias, F., 2011. Intersections and translocations: new paradigms for thinking about cultural diversity and social identities. *Eur. Educ. Res. J.* 10, 204–216.
- Beeli-Zimmermann, S., Wannack, E., Staub, S., 2020. Video-based educational research: what happens after recording with two cameras? *Forum Qual. Soc. Res.* 21 (2). <https://doi.org/10.17169/fqs-21.2.3298>.
- Bourdieu, P., 1990a. *The Logic of Practice*. Stanford University Press, Stanford, CA.
- Bourdieu, P., 1990b. *Photography. The Middle-Brow Art*. Polity Press, Cambridge.
- Braun, V., Clarke, V., 2021. *Thematic Analysis. A Practical Guide*. SAGE.
- Burns, D., Howard, J., Ospine, S.M., 2021. *The SAGE Handbook of Participatory Research and Inquiry*. SAGE.
- Cappello, M., 2005. Photo interviews: eliciting data through conversations with children. *Field Methods* 17, 170–182.
- Collier Jr., J., Collier, M., 1986. *Visual Anthropology. Photography as a Research Method*. University of New Mexico Press, Albuquerque, NM.
- Croghan, R., Griffin, C., Hunter, J., Phoenix, A., 2008. Young people's constructions of self: notes on the use and analysis of the photo-elicitation methods. *Int. J. Soc. Res. Methodol.* 11, 345–356.
- Darbyshire, P., MacDougall, C., Schiller, W., 2005. Multiple methods in qualitative research with children: more insight or just more? *Qual. Res.* 5, 417–436.
- Farber, P., Provenzo Jr., E., Holm, G., 1994. *Schooling in the Light of Popular Culture*. SUNY Press, Albany, NY.
- Flick, U., 2002. *An Introduction To Qualitative Research*. SAGE, Thousand Oaks, CA.
- Gewirtz, S., Cribb, A., 2008. Taking identity seriously: dilemmas for education policy and practice. *Eur. Educ. Res. J.* 7, 39–49.
- Grosvenor, I., 2007a. From the 'eye of history' to 'a second gaze': the visual archive and the marginalized in the history of education. *Hist. Educ.* 36, 607–622.

- Grosvenor, I., 2007b. 'Seen but not heard': city childhoods from the past into the present. *Paedagog. Hist.* 43, 405–429.
- Hall, S., 1996. Who needs 'identity'? In: Hall, S., Gay, P. (Eds.), *Questions of Cultural Identity*. Sage Publications, London, pp. 1–17.
- Harper, D., 2000. An argument for visual sociology. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Collecting and Interpreting Qualitative Materials*. SAGE, Thousand Oaks, CA.
- Holm, G., 1994. Learning in style: the portrayal of schooling in *Seventeen Magazine*. In: Farber, P., Provenzo Jr., E.F., Holm, G. (Eds.), *Schooling in the Light of Popular Culture*. SUNY Press, Albany, pp. 59–79.
- Holm, G., 1997. Teenage motherhood: public posing and private thoughts. In: Jipson, J.A., Paley, N. (Eds.), *Daredevil Research: Re-creating Analytic Practice*. Peter Lang, New York, pp. 61–81.
- Holm, G., 2008a. Photography as a performance. *Forum Qual. Soc. Res.* 9, 1–21.
- Holm, G., 2008b. Visual research methods: where are we and where are we going? In: Hesse-Biber, S.N., Leavy, P. (Eds.), *Handbook of Emergent Methods*. Guilford Press, New York, pp. 325–341.
- Holm, G., 2018. Visual ethnography in education. In: Beach, D., Bagley, C., Silva, S.M.D. (Eds.), *The Wiley Handbook of Ethnography of Education*. Wiley Blackwell, Hoboken, NJ, pp. 325–354. <https://doi.org/10.1002/9781118933732.ch14>.
- Holm, G., 2020. Photography as a research method. In: Leavy, P. (Ed.), *The Oxford Handbook of Qualitative Research*. Oxford University Press, New York, pp. 569–600.
- Holm, G., Londen, M., Mansikka, J.-E., 2015. Interpreting visual (and verbal) data: teenagers' views on belonging to a language minority group in Finland. In: Smeyers, P., Bridges, D., Burbules, N.C., Griffiths, M. (Eds.), *International Handbook of Interpretation in Educational Research: Part One*. Springer, Dordrecht, pp. 753–782.
- Holm, G., Sahlström, F., Zilliacus, H., 2018. Arts-based visual research. In: Leavy, P. (Ed.), *Handbook of Arts-Based Research*. Guilford Press, New York, pp. 311–335.
- Jipson, J., Paley, N. (Eds.), 1997. *Daredevil Research. Re-creating Analytic Practice*. Peter Lang, New York.
- Keats, P.A., 2009. Multiple text analysis in narrative research. Visual, written, and spoken stories of experience. *Qual. Res.* 9, 181–195.
- Knowles, J., Cole, A. (Eds.), 2008. *Handbook of the Arts in Qualitative Research*. SAGE.
- Leavy, P. (Ed.), 2019. *Handbook of Arts-Based Research*. Guilford Press, New York.
- Lodge, C., 2009. About face: visual research involving children. *Education* 37, 3–13.
- Mannay, D., Pauweis, L. (Eds.), 2019. *Sage Handbook of Visual Research Methods*, second ed. SAGE.
- Manny, D., 2010. Making the familiar strange: can visual research methods render the familiar setting more perceptible? *Qual. Res.* 10, 91–111.
- Mitchell, C., 2011. *Doing Visual Research*. SAGE, London.
- Moss, J., Pini, B. (Eds.), 2016. *Visual Research Methods in Educational Research*. Palgrave Macmillan, London. <https://doi.org/10.1057/9781137447357>.
- Pienimäki, M., 2021. Participatory photography supporting the social inclusion of migrant youth. *J. Youth Stud.* 24, 1179–1198. <https://doi.org/10.1080/13676261.2020.1814227>.
- Pini, M., Walkerdine, V., 2011. Girls on film. Video diaries as 'autoethnographies'. In: Reavey, P. (Ed.), *Visual Methods in Psychology. Using and Interpreting Images in Qualitative Research*. Psychology Press, New York, pp. 139–152.
- Pink, S., 2012. *Doing Visual Ethnography*. SAGE, London.
- Prosser, J., 2000. The moral maze. In: Simmos, H., Usher, R. (Eds.), *Situated Ethics in Educational Research*. Routledge Falmer, London, pp. 116–132.
- Reay, D., 2009. Identity making in the classrooms. In: Wetherell, M., Mohanty, C.T. (Eds.), *The Sage Handbook of Identities*. Sage, London, pp. 277–294.
- Rose, G., 2016. *Visual Methodologies: An Introduction to Researching with Visual Materials*. Sage, London.
- Rusk, F., Pörn, M., Sahlström, F., Slotte-Lüttge, A.G., 2014. Perspectives on using video recordings in conversation analytical studies on learning instruction. *Int. J. Res. Method Educ.* 38, 29–55.
- Sutton-Brown, C.M., 2014. Photovoice: a methodological guide. *Photogr. Cult.* 7, 169–185. <https://doi.org/10.2752/175145214X13999922103165>.
- Sweetman, P., 2009. Revealing habitus, illuminating practice: Bourdieu, photography and visual methods. *Socio. Rev.* 57, 491–511.
- Veintie, T., Holm, G., 2010. The perceptions of knowledge and learning of Amazonian indigenous teacher education students. *Ethnogr. Educ.* 5, 325–343. <https://doi.org/10.1080/17457823.2010.511443>.
- Wang, C.C., 2006. Youth participation in photovoice as a strategy for community change. *J. Community Pract.* 14, 147–161. https://doi.org/10.1300/J125v14n01_09.
- Wang, C.C., Burris, M.A., 1997. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ. Behav.* 24, 369–387. <https://doi.org/10.1177/109019819702400309>.

Document analysis[☆]

Tim Rapley^a, Kevin Ardron^a, David Littlefair^a, Kirstin Mulholland^a, and K. Neil Jenkins^b, ^a Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom; and ^b School of Geography, Politics & Sociology, Newcastle University, Newcastle upon Tyne, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

This is an update of T. Rapley, K.N. Jenkins, Document Analysis, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010 Pages 380-385, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01522-0>.

Approaches to analyzing documents	285
Meta-synthesis: documents and creating (new) knowledge	286
Discourse analysis: documents create realities	287
Ethnography: documents, interactions, and organizations	288
Ethnomethodological ethnography: documents-in-interactions	289
Conclusion	290
References	290
Further reading	291

Documents are a seen-but-unnoticed part of our everyday lives. Whether physically held in your hand, viewed on a computer screen or scrolled through on a mobile phone, they are staple elements of all areas of our lives. They shape our thoughts, actions, talk and writing. We consume them in vast quantities, routinely refer to them in conversations, as well as produce them. However you conceive of them, they are essential to the formation of our contemporary societies and institutions.

Document-based research in qualitative educational research, as in all areas of the social sciences, is a relatively small-scale enterprise. The notable exception is historically orientated research, where—given the lack of living sources of evidence—the document is the potentially the only source of evidence. Also, the practice of “reviewing the literature” relies on collecting, analyzing and speaking about a range of documents. Commonly, this is often relegated to the front sections of articles and reports, prior to a primary focus on the analysis of interviews, conversations in focus groups or observations of the interactions of others, but rarely explained as a practice itself.

In this chapter, we will begin to explore some of the variety of ways that qualitative researchers have actively sought to engage with documents as a central part of their research practice. We will explore some more general issues about conducting research on and around documents. We will then focus on two main areas. Firstly, we look at research that focuses on the substantive content of documents—the relatively new area of systematically reviewing qualitative research articles—and the more discursive work that looks at how documents, be they newspaper articles or policy documents, creates specific versions of the world. Secondly, we will look at research that focuses on the situated use and creation of documents—how interview-based and ethnographic studies have also looked at the place of documents in the lives and routines of those they study and on the more ethnomethodological work, that seeks to describe in detail the practical ways that documents are used and created in specific instances of interaction.

Approaches to analyzing documents

We will be using the term documents to refer to both paper-based and digital texts. This includes both the written elements of texts alongside the extra-textual elements—images, photos, graphs, diagrams—that are routinely embedded in documents. The post-structural and post-modern turn in academia has led to an expansion of the meaning of the term “text”, to include buildings, bodies, clothing, alongside artifacts, devices and other aspects of material and technical culture. Simultaneously, interview transcripts, field notes and video-recordings are referred to and analyzed as “texts”, but we will not focus on either of these genres of work here. Instead, we are concentrating on four relatively distinct areas of work around documents (See [Table 1](#)). This table provides an “at a glance” comparison of the four analytic methods in terms of their general focus; a more descriptive account of their analytic focus; examples of their typical research questions; the data collected to answer these questions; the key methods involved in data

[☆] *Change History:* April 2021 T. Rapley, K. Ardron, D. Littlefair, K. Mulholland and K.N. Jenkins updated Abstract, Main Text, References, Further Readings and Table 1.

Table 1 Approaches to analyzing documents.

	<i>Meta-synthesis</i>	<i>Discourse analysis</i>	<i>Ethnography</i>	<i>Ethnomethodological ethnography</i>
<i>Focus</i>	<i>Content of documents</i>		<i>Documents-in-action</i>	
Analytic focus	Inter-textuality	Language & inter-textuality	Social interaction & organizations	Language & social interaction
Typical research questions	What is the evidence for a specific topic?	What versions of the world do documents create?	What part do documents play in the organization of a social institution?	How do documents shape (and are reflexively shaped by) ongoing social interaction?
Data	Academic outputs	Any	Any	Any
Method	Collect documents	Collect documents	Collect documents, field notes, formal and informal interviews	Collect documents, field notes, interviews, audio-visual data
Methodologies	Meta-ethnography, meta-synthesis, thematic synthesis Qualitative synthesis	Discourse analysis, critical discourse analysis, foucauldian discourse analysis, semiotics	Interactional ethnography, grounded theory, phenomenology, symbolic interactionism	Conversation analysis, ethnomethodology
Results	Synthesis of findings of a collection of documents.	How documents construct and reflect a specific version of an argument, truth, social identities or institutions.	Use, role and function of documents in specific organizational contexts.	Use, role and function of documents in specific interactional contexts.

collection; the methodologies often related to these approaches; and the type of results they each produce. These four analytic approaches will each be described individually in greater detail below.

Analytic work on documents can be loosely divided into two areas (see [Table 1](#), “Focus”):

1. Work that focuses on the actual textual and extra-textual content of documents (meta-synthesis and discourse analysis).
2. Work that focuses on the use, role and function of documents in interactional and organizational settings (ethnography and ethnomethodological ethnography).

The first focuses on the document as an object in its own right, as an “orphaned” or “docile” text, as a container of knowledge. Such work forgoes empirical observation of how people actually read, refer to or use the documents in question. The second area is primarily observational, seeking to understand how documents are active agents in organizational and/or interactional life. Obviously, the same topic can be approached in multiple ways and within the same project. So, with a research project on say, student evaluations of lecturers, you might focus on: the written comments students give (via discourse analysis of these evaluation forms); how results of these evaluations are fed-back on web-sites, internal memoranda or official reports (via discourse analysis of a range of types of documents); how lecturers introduce, distribute, collect and read such forms (via observation in classrooms and offices); what students, lecturers and managerial staff think is the value of such documents (via observation of lectures and meetings and interviews with people); the role of such evaluations in creating institutional change (via observation and interview); etcetera.

Whatever approach people take, and irrespective of the methodological traditions they follow, the practical work of analyzing data related to documents closely echoes that of all forms of qualitative research. So, the data—be it documents, interviews, field notes or audio-video recordings (see [Table 1](#), “Method”)—will be subject to some form of sampling, as it is collected it will be coded, and some version constant comparison method and (ideally) deviant case analysis undertaken, as well as some level of reflection on the role of the analysts in the process.

The analysis of documents is also related to the research tradition of content analysis (see [Krippendorff, 2018](#); [Weber, 1990](#)). In this, researchers gather specific documents, then establish a coding frame which is then applied to the documents to count the number of times particular words, phrases or themes are used. This enables descriptive and statistical findings to be established (see for example [Bozkurt et al., 2015](#) for a more statistical style of work). However, this is often seen by qualitative researchers as a relatively limited, albeit systematic approach, in that it overly predefines the key aspects of analysis alongside often removing words and phrases from their context (see [Paradis and Whitehead, 2015](#), who blend deductive and inductive approaches, as well as [Hartley and Morphew, 2008](#) for more emergent, descriptive, style of work). Some researchers use content analysis alongside other more discursive methodologies, in part to gain a general overview of their data, although this is not discussed here.

Meta-synthesis: documents and creating (new) knowledge

The social, medical and political sciences have seen the rise of “evidence-based” policy and practice over the last 30 years. Education research has been central in this trajectory, and this shift has not been without its critics (see, for example, [Hammersley, 2007](#)). The growth in systematic reviewing and meta-analysis has also been significant in prescribing a new format for the textual analysis of research (see [Table 1](#), first column). In relation to reviews of non-experimental research, [Noblit and Hare’s \(1988\)](#) work on “meta-ethnography” has been influential. They sought to combine the findings of six ethnographies on the impact of desegregation on

urban schools in the US in the late 1970s. They translated the relatively diverse findings of the case studies into a general theory about the impact of desegregation. Following this inter-textual work, a range of approaches to synthesizing the findings of qualitative research have emerged. They have been referred to under various terms including: metasynthesis (Sandelowski and Barroso, 2006), critical interpretative synthesis (Dixon-Woods et al., 2006a) and thematic synthesis (Thomas and Harden, 2008). They all rely on insights from Nobilt and Hare's (1988) original work, and all seek to move beyond simply aggregating and summarizing findings from qualitative research papers to actively create new conceptual models or theories and to provide evidence for future interventions.

As with quantitative systematic reviewing, this qualitative synthesis work involves searching, evaluation, data-extraction and presentation. However, given the non-numerical nature of the findings they have adapted each stage in the process. Importantly, the process echoes and draws on key ideas from how you conduct empirical, primary, qualitative research. For example, when searching the existing literature, you seek a broad range of heterogeneous studies, purposively sampling studies that cover a range of approaches and findings to the substantive topic of the review. Scruggs et al. (2007) work on co-teaching used a range of databases, reviewed the reference lists of included papers as well as those papers that had cited the work, alongside handsearching key journal in the field. You may additionally engage in further rounds of sampling to find further documents that challenge or extend your initial findings (see Dixon-Woods et al., 2006b). The evaluation of the quality of studies is also complex, in that there is no hierarchy of evidence available for qualitative research. In this way, inclusion and exclusion criteria and judgments about quality are more flexible with an initial focus on how relevant the article is to the topic and to the development of the synthesis, and only papers deemed to have major flaws are rejected at the outset. Scruggs et al. (2007) used such a broad approach of quality appraisal, rejecting a checklist approach. They report "employing ... considerations", that focused upon processes and outcomes, that:

[d]ata were systematically collected and recorded, multiple informants and/or data sources were obtained, disconfirming evidence was considered, and the conclusion was reasonable and appropriate based on the data collected.

Scruggs et al. (2007, 398).

They also note that all studies had already gone through peer review, be that through acceptance for a journal or thesis examination. Currently, we have a range of approaches to judging quality that draw upon and combine the key aspects of various quality guidelines that exist.

With data-extraction, the focus is on collecting data that enables the researcher to outline the key ideas, concepts and themes within the article. This also includes extracting verbatim quotes from interviews, field notes and other sources alongside summarizing key aspects. The final stage of synthesizing the data shifts the focus beyond just summarizing the extracted data, to producing explicit interpretations. Again, this process is similar to that undertaken in analysis of primary data in qualitative research. The extracted data is then coded in detail, coded sections are compared and contrasted, and you generate a coding frame which outlines the main themes and issues and relationship between them. These recurring "lines of argument" (Nobilt and Hare, 1988) are drawn from both your new conceptualization of the original data together with some of original concepts used in the articles. Ong et al.'s (2020) synthesis of work on women of color in undergraduate engineering education, that included both mixed methods and quantitative papers as well as qualitative work, used both deductive and inductive coding framework, that they developed through an iterative process. Their abductive analysis developed four analytic themes—interest, motivation, and self-efficacy; social pain; navigation; and support entities—with further, detailed dimensions to each area. For example, one of the four dimensions, "being the only one" of the theme "social pain", outlines the sense of isolation, loneliness, and anxiety that can emerge "from being the only one of their gender and/or race/ethnicity" (596) in an engineering classroom.

This style of working with documents is still relatively young, much of the work has been undertaken in relation to health-related topics, and currently we have a range of complementary methods and methodologies (e.g., scoping reviews, Arksey and O'Malley, 2005, and rapid reviews, Khangura et al., 2012 and see Gough et al., 2012, for a useful overview). What they do share is a concern to compare, combine and critically interpret the findings of different qualitative articles. In this way, they attempt to generate new micro- and meso- level conceptual models of the topic of the review.

Discourse analysis: documents create realities

Rather than see the documents as neutral sources of information to be searched, evaluated against some external quality criteria, and then the data abstracted from the context, another way that researchers analyze documents is to undertake a relatively close and detailed analysis of the language and meaning within the document (see Table 1, second column). This style of research, known as discourse analysis, focuses on how documents produce specific versions of reality. Within this approach, language (written or spoken) is never treated as a neutral, transparent, means of communication. Instead, language is understood as performative and functional. So, for example, describing something as a "systematic review" instead of as, say, a "review of the literature" or a child as "gifted and talented" over, say, "bright for their age" does specific work, and directs thoughts and actions. Researchers

are interested in the specific versions of the world that documents produce, the effects they have both on the potential reader and on the wider culture, society, and institutions.

Within this style of research, the focus is on how words are used. Methodologically, there are two main research trajectories – discourse analysis (Potter and Wetherell, 1987) and critical discourse analysis (see Fairclough, 2013; Rogers, 2011)—although in practice they often somewhat overlap. Whatever approach is taken, when studying documents in this way one is interested in the rhetorical work of the text, how the specific issues it raises are structured and organized, and how it constructs a specific version of the world.

You might try to understand how the document seeks to persuade someone about the authority of its understanding on an issue. So, for example, you might focus on the range of sources of knowledge and evidence a document draws on—so within a newspaper article you may have verbatim quotes from leading practitioners or academics, references to policy documents or academic articles, and report on conflicting sides of the argument. Such intertextual work helps build the factual status of the account. Moreover, you might focus on the forms and modes of language, knowledge and evidence that are employed. For example, Setyono and Widodo's (2019) work on the (re)presentation of multicultural values in an English as a foreign language (EFL) textbook, closely analyses parts of an Indonesian Ministry of Education and Culture-Endorsed EFL textbook. They show how both visual and verbal elements of the text—including the specific cartoon images used, as well stories and situations used as examples—seek to encourage readers to see and value pluralism. Centrally, such work is trying to make sense of how the document produces a specific argument, the positions it takes in relation to a topic alongside the other alternate or contradictory positions that are excluded or silenced.

Another central question asked of documents from work in this tradition is how are specific ideas, practices, or identities produced, sustained or negotiated within texts? Some researchers seek to understand and describe the historical trajectory of the contemporary ideas we all currently take for granted. People often work quite closely with some of the writings of the French philosopher Foucault (1977). One of Foucault's interests was in how specific discourses, specific formations of what he called power/knowledge (e.g., medical, psychological, or pedagogical discourses) produce, shape, and enable specific subjects (e.g., attention-deficit/hyperactivity disorder (ADHD)-diagnosed children) and specific actions (e.g., advising parents to get children tested; prescribing medications like Ritalin). For example, Rafalovich (2004) has focused on the conceptual antecedents of the contemporary diagnosis of ADHD. Drawing on articles and books from 1877 to 1929, Rafalovich looked at the trajectory of three concepts—"idiocy"; "imbecility"; and "encephalitis lethargica"—and was able to reveal the initial attempts by medicine to establish an organic basis for children's apparently immoral, or unconventional behavior.

Such histories of the present remind us that what we take for granted often has complicated and esoteric beginnings. This style of work may lead to us to the question why exactly we act in "just this way" or why certain groups of people have the knowledge and power, i.e., authority, to interpret, speak about and act upon others. Importantly, it seeks to map the trajectories of these discourses, to outline the specific strategies that led to our present state of social organization as well as those that were rejected or seen to fail. Obviously, given its historical nature, this style of work often uses documents to explore specific topics or questions. However, such histories may focus on something that has emerged in the recent past and may also rely on other forms of data—a multimethod approach. Therefore, in Rafalovich's (2004) study of ADHD, he also undertook interviews with parents, children, teachers, and clinicians. In addition, rather than focus in great detail on a small sample of texts these studies can instead focus on a broad range of texts as they try to show the styles of thought as they emerge, consolidate, and compete across and between texts often over a large time-scale. Importantly, such work asks: "what are the assumptions in specific documents?"

Ethnography: documents, interactions, and organizations

The ethnographic approach to documents (see Table 1, third column) can cover a broad spectrum of activities and sites, although it routinely focuses on the use, role, and function of documents in the context of schools. In general terms, such work seeks to combine a more micro-focus on moments of social interaction with and around documents in a range of classroom settings, a meso-focus on the organizational ecology of schools and a macro-focus on broader community, policy and political factors.

One of the most obvious roles documents play in education is reading, especially reading in class. Ethnographers are not really concerned with reading in terms of the cognitive processes involved and do not understand it as simply as a monadic activity—that is, that reading can only be understood as something people undertake in isolation or even individually. Rather, they focus on reading as a practice situated within and related to the broader contexts of classroom activities and organizational cultures. For example, the video-based ethnography of literacy by Castanheira et al. (2000) demonstrates the situated roles and relationships that people can establish around documents. They offer two contrasting cases. They show how in one moment, a teacher positions a specific workbook as the authoritative source of knowledge and how the student and their partner are positioned as potential resources for defining and completing the task. Whereas in a different lesson, a teacher directly engages with the student, asking questions, encouraging them, draws a graph on the student's paper, etc. In this context, the student and teacher collaborate on the problem at hand, thus the student is positioned as central in producing knowledge with the workbook positioned in a supporting role. Such ethnographic work also seeks to make sense of such situated reading work within the organizational ecology of education. Therefore, you may compare the use, role, and function of workbooks over time, over single classes, over different students, or over multiple subject areas; you may research the background of the teachers or how the workbooks get selected as well as the

schools or regional policies around such workbooks. In this way, such research focuses on the embeddedness of reading documents in relation to both students' classroom interactions and their broader educational and social worlds (Green and Meyer, 1991).

As education is bureaucratically organized—and routinely involves coordinating a diverse array of professionals, administrators, and technicians alongside learners over a range of activities, times, and spaces—documents also become central, a glue that binds these people and their work together. Mehan's (2000) ethnographic study of the identification, classification, and labeling of educationally disabled children shows us the central role that documents can play in the coordination of educational interactions, decision making, and organizational work. His study draws on a range of methods; these include observing, interviewing, video-recording meetings, and lessons alongside asking teachers to reflect on the videos of these lessons. He uses the sequential paper trail that is created by a student (in class work, examinations, and educational testing sessions with the school psychologist) and created about a student (in their personal school records, school appraisal team committee, and eligibility and placement committee meetings) to focus the fieldwork on the range of sites where bureaucratic assessment and classification work occurs in the special-education referral process.

Documents, in the form of student records, are continually created and discussed by a range of professionals in multiple sites and spaces. Centrally, these records come to represent and stand-on-behalf of the student. In meetings about students, discussions routinely draw on and highlight aspects of the paper trail created by and about the student. The decisions that emerge from these meetings are central in classifying the child and, therefore, are central in creating their specific educational trajectory. In his study, Mehan (2000) does not present us with the documents themselves, rather he focuses on the distributed talk and actions of those people who create and use them. He shows us how the documents created by and about students become objectified and how the technical talk and decisions they facilitate are given much higher status than other forms of evidence such as a parental voice.

Ethnomethodological ethnography: documents-in-interactions

Research on documents within the tradition of ethnomethodological ethnography—and we include conversation analysis within this area—focuses on describing in great detail the moment-by-moment social organization of interactions with, around, and about documents (see Table 1, fourth column). Like ethnographic work outlined in the last section, it is not so much interested in the content of documents, but in the organization of interactions where documents take some role. However, in general terms, there are two central differences. First, for ethnomethodological work, the immediate, here-and-now, to-and-fro of the participants' interactions is the central object of interest, description, and analysis. Broader contextual factors, like participants' backgrounds or organizational policies, are generally not drawn on as a resource *a priori* to explain why an interaction happens in a specific way. Instead, the focus is on what the participants in the interaction show themselves to actually do, exploring the (contextual) resources they draw on in just that moment of interaction. Second, ethnomethodological work increasingly relies on audio or video recordings of naturally occurring interactions. These form the centerpiece of analysis, are subjected to repeated viewings, and sequences of interaction—including verbal and nonverbal features—are transcribed in very fine detail, analyzed, and revised in conjunction with the original recordings.

An example of this method in practice is Heap's various analyses of audiotapes of young children reading in classrooms (e.g., Heap, 1985, 1991). Rather than assuming we already know what reading looks like, or when it occurs in classrooms—through some *a priori* theory or criteria—he focuses on the broad range of interactional activities that produce the ongoing activity as reading. For him, children learn what counts as reading criterially, by learning what counts as reading procedurally, in and through taking part in interactions where "good" reading is shown to be taking place. For example, Freebody and Freiberg (2001) show how in a moment of a parent-child interaction, where a child is reading a book aloud, the parent works to praise, instruct the child to sound the letters, correct soundings, etc. In this moment of reading a document, the central interactional task is reading out loud a written text, and "good" reading is orientated to and produced as correct word-saying. In other contexts, say a university seminar on a specific article, through debating and discussing, students learn what "good" academic reading involves—finding fault, contrasting with other documents, quoting and referencing from this and other documents, reading footnotes, following up references, etc. Such work focuses researchers' attention on unpacking, in detail, how specific tasks are undertaken.

Documents are not only paper-based, single-user, objects and ethnomethodological ethnographic research has focused on how mundane and digital technologies shape the world of educational activities. Such research has focused on how people interact, work with, and draw on such document-related technologies as white- and blackboards, overheads, computer-based presentations, and videos. Rendle-Short's (2006) study of computer-science seminar presentations focuses on the academic monologue. She shows us how presenters work to interact with the audience not only through their talk, but also through their gestures, gaze, and bodily movements. Using a collection of videotapes of seminars and transcribing in great detail the moment-by-moment verbal and nonverbal behavior, she focuses on how presenters coordinate a range of documents and technologies as they seek to engage with the audience. As presenters talk, they time their slides to focus, support, and supplement the issues they are raising. The text-based and visual images on the slides work to illustrate some aspects of their ongoing talk, and presenters through their hand gestures work to direct the audiences' attention to a specific aspect or issue. In this way, we begin to see how, what appears as a relatively simple task—giving a seminar—is saturated with a complex range of interactional work. The documents the presenters

work with, be it notes available to them, or slides available to all, are brought to life and shaped through the concerted, moment-by-moment, organization of talk, gesture, and technology.

Such ethnomethodological ethnographic research findings can seem focused on the quite obvious, in that it can show us what we take for granted. However, researchers in this tradition argue that it is only through an understanding of the details of what we actually do (the details of which can largely escape our notice)—rather than what people tell us they do, or what we think they do—that we can then develop theories or interventions that are actually relevant to current practice.

Conclusion

Given the ubiquity of paper and digital documents in contemporary social and organizational life, the range of potential types of documents open to qualitative analysis is immense. However, the primary sources for educational research are policy documents (be they national, regional, or local), textbooks, and internal organizational documents; the secondary sources are from the academic canon. Documents in any broader sense—be they blogs, diaries, magazines, newspapers, or websites—are underused and underanalyzed. As noted above, the lack of focus on documents is not restricted to educational research but is instead endemic to the social sciences. On the one hand, this is quite surprising, given the centrality of documents to everyday experiences, the readiness-to-hand of document-based “raw data”, and that it does not have to require large amounts of time in the field collecting data. However, ideas around hearing peoples’ voices, witnessing things with your own eyes, have been increasingly positioned as a central requisite of authentic and high-quality qualitative research. Nevertheless, in relation to critical discourse analysis style studies of education, Rogers et al. (2005) see the turn away from a focus solely on written documents toward a focus on talk as a positive direction, in which new theoretical, methodological, and analytic possibilities emerge. Notwithstanding these potentials for development alongside the existing tradition of ethnographic work focusing on the use, role, and function of documents-in-action, we are still left with a relatively limited corpus of work that focuses on the social and educational lives of documents. Documents shape, and are reflexively shaped by, our perceptions, interactions, institutions, policies, and society. They are central in the production, reproduction, and transformation of our educational landscapes. As such, they deserve a more sustained and systematic analytic focus.

References

- Arksey, H., O'Malley, L., 2005. Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol.* 8 (1), 19–32.
- Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogdu, E., Ucar, H., Guler, E., Sezgin, S., Karadeniz, A., Sen-Ersoy, N., Goksel-Canbek, N., Dincer, G.D., 2015. Trends in distance education research: a content analysis of journals 2009–2013. *Int. Rev. Res. Open Dist. Learn.* 16 (1), 330–363.
- Castanheira, M.L., Crawford, T., Dixon, C.N., Green, J.L., 2000. Interactional ethnography: an approach to studying the social construction of literate practices. *Ling. Educ.* 11 (4), 353–400.
- Dixon-Woods, M., Bonas, S., Booth, A., Jones, D.R., Miller, T., Sutton, A.J., Shaw, R.L., Smith, J.A., Young, B., 2006a. How can systematic reviews incorporate qualitative research? A critical perspective. *Qual. Res.* 6 (1), 27–44.
- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L.K., Riley, R., Sutton, A.J., 2006b. Conducting a critical interpretive review of the literature on access to healthcare by vulnerable groups. *BMC Med. Res. Methodol.* 6 (1), 1–13.
- Fairclough, N., 2013. *Critical Discourse Analysis: The Critical Study of Language*. Routledge.
- Foucault, M., 1977. *Discipline and Punish: The Birth of the Prison*. Pantheon.
- Freebody, P., Freiberg, J., 2001. Re-discovering practical reading activities in homes and schools. *J. Res. Read.* 24 (3), 222–234.
- Gough, D., Thomas, J., Oliver, S., 2012. Clarifying differences between review designs and methods. *Syst. Rev.* 1 (1), 1–9.
- Green, J.L., Meyer, L., 1991. The embeddedness of reading in classroom life: reading as a situated process. In: *Towards a Critical Sociology of Reading Pedagogy*, pp. 141–160.
- Hammersley, M. (Ed.), 2007. *Educational Research and Evidence-Based Practice*. Sage.
- Hartley, M., Morpew, C.C., 2008. What's being sold and to what end? A content analysis of college viewbooks. *J. High Educ.* 79 (6), 671–691.
- Heap, J.L., 1985. Discourse in the production of classroom knowledge: reading lessons. *Curric. Inq.* 15 (3), 245–279.
- Heap, J., 1991. A situated perspective on what counts as reading. In: *Towards a Critical Sociology of Reading Pedagogy*, pp. 103–139.
- Khangura, S., Konnyu, K., Cushman, R., Grimshaw, J., Moher, D., 2012. Evidence summaries: the evolution of a rapid review approach. *Syst. Rev.* 1 (1), 1–9.
- Krippendorff, K., 2018. *Content Analysis: An Introduction to Its Methodology*. Sage.
- Mehan, H., 2000. Beneath the skin and between the ears: a case study in the politics of representation. In: *Schooling the Symbolic Animal: Social and Cultural Dimensions of Education*, pp. 259–279.
- Noblit, G.W., Hare, R.D., 1988. *Meta-ethnography: Synthesizing Qualitative Studies*, vol. 11. Sage.
- Ong, M., Jaumot-Pascual, N., Ko, L.T., 2020. Research literature on women of color in undergraduate engineering education: a systematic thematic synthesis. *J. Eng. Educ.* 109 (3), 581–615.
- Paradis, E., Whitehead, C.R., 2015. Louder than words: power and conflict in interprofessional education articles, 1954–2013. *Med. Educ.* 49 (4), 399–407.
- Potter, J., Wetherell, M., 1987. *Discourse and Social Psychology: Beyond Attitudes and Behaviour*. Sage.
- Rafalovich, A., 2004. *Framing ADHD Children: A Critical Examination of the History, Discourse, and Everyday Experience of Attention Deficit/Hyperactivity Disorder*. Lexington books.
- Rendle-Short, J., 2006. *The Academic Presentation: Situated Talk in Action*. Ashgate Publishing, Ltd.
- Rogers, R. (Ed.), 2011. *An Introduction to Critical Discourse Analysis in Education*. Routledge.
- Rogers, R., Malancharuvil-Berkes, E., Mosley, M., Hui, D., Joseph, G.O.G., 2005. Critical discourse analysis in education: a review of the literature. *Rev. Educ. Res.* 75 (3), 365–416.
- Sandelowski, M., Barroso, J., 2006. *Handbook for Synthesizing Qualitative Research*. Springer publishing company.
- Scruggs, T.E., Mastropieri, M.A., McDuffie, K.A., 2007. Co-teaching in inclusive classrooms: a metasynthesis of qualitative research. *Except. Child.* 73 (4), 392–416.
- Setyono, B., Widodo, H.P., 2019. The representation of multicultural values in the Indonesian Ministry of Education and Culture-Endorsed EFL textbook: a critical discourse analysis. *Intercult. Educ.* 30 (4), 383–397.

Thomas, J., Harden, A., 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med. Res. Methodol.* 8 (1), 1–10.

Weber, R.P., 1990. *Basic Content Analysis*. Sage.

Further reading

Bowen, G.A., 2009. Document analysis as a qualitative research method. *Qual. Res. J.* 9 (2), 27–40.

Coffey, A., 2014. Analysing documents. In: *The SAGE Handbook of Qualitative Data Analysis*, pp. 367–380.

Grant, A., 2018. *Doing Excellent Social Research With Documents: Practical Examples and Guidance for Qualitative Researchers*. Routledge.

Livingston, E., 1995. *An Anthropology of Reading*. Indiana University Press.

Prior, L., 2003. *Using Documents in Social Research*. Sage.

Rapley, T., 2018. *Doing Conversation, Discourse and Document Analysis*, vol. 7. Sage.

Rapley, T., Rees, G., 2018. Collecting documents as data. In: *The SAGE Handbook of Qualitative Data Collection*, pp. 378–391.

Scott, J., 2014. *A Matter of Record: Documentary Sources in Social Research*. John Wiley & Sons.

Smith, D.E., 1990. *Texts, Facts and Femininity: Exploring the Relations of Ruling*. Routledge.

The ontological imperative in conducting qualitative research in online spaces

Hannah R. Gerber^{a,b}, ^a Sam Houston State University, Huntsville, TX, United States; and ^b University of South Africa, Pretoria, South Africa

© 2023 Elsevier Ltd. All rights reserved.

Introduction	292
Digital methods in qualitative research in online spaces	293
Technical outline of the ontological imperative in web-based qualitative research methods	294
Question 1: what digital tools, systems, and services are at play in my study? Who created them and why?	294
Question 2: what data do these digital tools, systems, and services render?	294
Question 3: what hidden limitations might there be to the data rendered via these digital tools, systems, and services?	294
Question 4: what are the epistemological implications of this ontological analysis?	294
Question 5: what are the axiological implications of this ontological analysis?	294
Empirical demonstration of ontological imperative in web-based qualitative research methods	295
Question 1: what digital tools, systems, and services are at play in my study? Who created them and why?	295
Question 2: what data do these digital tools, systems, and services render?	295
Question 3: what hidden limitations might there be to the data rendered via these digital tools, systems, and services?	296
Question 4: what are the epistemological implications of this ontological analysis?	298
Question 5: what are the axiological implications of this ontological analysis?	299
Ways to continue growing as a researcher in web-based qualitative research methods	299
Read up on digital methods and theories of interaction in digital spaces	299
Conduct software and app walkthroughs	300
Explore developer documentation for different spaces	300
Connect with colleagues interested in research in digital spaces	300
Final thoughts	300
References	301
Further reading	301
Relevant websites	301

Introduction

As of 2020, the global Internet penetration mark has exceeded 50%; there are now more than 4 billion active users who use the Internet on a daily basis for an average of 6 h a day (We Are Social, 2020). In fact, every minute there are over 40,000 searches on Google, 527,760 photos shared on Snapchat, 4.1 million videos streamed on YouTube, half a million Tweets sent, and approximately 46,740 photos shared (Marr, 2018). Annually, the global population spends an astounding combined 1 billion years online. By region, the highest Internet penetration rates remain in Northern Europe at an average of 94% and the lowest Internet penetration rates are in Mid-Africa, at an average of 12% Internet penetration (We Are Social, 2020). Although there is much progress to be made to bring the remainder of the world's population online, with the rapid transformation of digital innovation in methods of Internet connectivity in the coming years (e.g., launching of satellites, solar balloons, 5G, etc.), more and more individuals from far flung corners of the globe will gain access to the Internet. Society is living through an exciting shift in how meaning is made and how individuals communicate, thus opening up a wide world of opportunities for researchers to understand how people make meaning of their world within, around, and through digital spaces.

Because of the influx and rapid shift in human communication methods—as evidenced in the aforementioned statistics—it is imperative that researchers and scholars develop qualitative approaches tailored to understanding meaning making and communication in online spaces. Scholarly research must begin to better understand the digital worlds that people inhabit in order to make sense of the diverse ways that people make meaning in and across online spaces. One such theoretical framework that accounts for the nature of interconnectedness in human and machine communication activities, and begins to delve into the multifarious ways that people make meaning in and across online spaces, is *networked field sites* (Gerber et al., 2017). Networked field sites are “dynamic spaces that are user-driven, social, and collaborative” (Gerber et al., 2017, p. xix). One of the defining features of networked field sites is that of connectivity and the way that data flow among, across, and between users, sites, systems, and services. A person might post something on his or her blog, perhaps a picture, image, or story. Another individual might then share or post that blog link on his or her Twitter, Tumblr, Instagram, or Facebook. Other individuals across the network might modify, change, adapt, or repurpose that post for his or her own purposes. The original post, now modified, might make it back to the original person in his or her blog community, therefore shifting and reshaping the original intent of the post, but having met the needs of larger group across the network. That same individual might later delete the original post, or perhaps the social media site might

delete the user account or post, therefore also solidifying the temporal nature of how information exists across the network, however, although the original might be deleted, the modifications to the original still remain intact across the larger community. This is the reality of the way that information exists within networked field sites. It is temporal, porous, shifting, user-driven, social, and collaborative and as such requires certain methodological considerations when bringing into any study.

This chapter will introduce readers to the ontological imperative framework, a methodological framework that is comprised of five sets of questions that walk researchers through a critical examination of the data that are yielded within online spaces. In particular, this chapter will introduce the framework and will provide an empirical example of the framework used within a qualitative study. The chapter will conclude with practical suggestions for the reader to consider in their own growth as a scholar and future research projects.

Digital methods in qualitative research in online spaces

Because networked field sites are dynamic, complex, nuanced, and porous, it is imperative to think through methods that will allow researchers to capture the meaning-making processes and experiences that people are engaged in when taking part in activities in online spaces. This warrants an understanding and development of new methodological approaches to study the way people interact in online spaces. As Gerber et al. (2017) posited, developing new methodological approaches to understand socialization and meaning making in online spaces "... will be warranted as technologies continue to permeate public and private spaces and as movement among digital and nondigital text continues to become routine" (Gerber et al., 2017, p. 168). Scholars have been quick to point out that traditional methods of research (e.g., interviews, surveys, etc.) might not support understanding how people and information move among, across, and through digital platforms and online spaces of the networked field sites and such researchers have been quick to call for the development of new methods that can capture and trace the ways that people make meaning in online spaces (Bruns and Burgess, 2013; Gerber and Lynch, 2017; Rogers, 2013). Traditional methods (e.g., surveys, interviews, observations, etc.) that have been ported to online spaces are known as *digitized methods* whereas new methods used to study online interactions are known as *born digital methods*, or *natively digital methods* (Rogers, 2013). These methods exist only because online technologies allow such methods to exist. Natively digital methods, such as web scraping/crawling, calling an application programming interface (API), and agent-based modeling emerged from fields of computer science and data science. Although typically used for quantitative analyses within data science (i.e., with big data), these methods are important to acknowledge for how they can inform a qualitative understanding of the meaning making among people occurring within networked field sites. Regarding the use of digital methods to study interactions in online spaces Rogers (2013) stated

Ultimately, I propose a research practice that learns from the methods of online devices, repurposes them, and seeks to ground claims about cultural change and societal conditions in web data, introducing the term "online groundedness." The overall aim is to rework methods for Internet-related research, developing a novel strand of study, digital methods (p. 19).

Digital methods, or those that are born out of the study of how online systems and digital objects themselves operate—such as recommender systems, hyperlinks, APIs, algorithms—can provide a unique understanding of how to redesign and develop new methods for the study of meaning making, socialization, and culture in online spaces, including the networked field sites that a researcher might be engaged in studying. Simply taking traditional methods such as interviews, surveys, and observations and relying upon those methods as the sole methods to study online spaces means that a scholar has not acknowledged the operational facets of the networked field site that he or she is researching. A networked field site operates on binary code, is powered by algorithms, and driven by software. Acknowledging that the field site itself is unique, because of the networked system within which it operates, provides researchers a solid footing for engaging in a critical deconstruction of the very sites, spaces, systems, services, and people that they are studying.

Lynch and Gerber (2018) created a *methodological* framework, termed the *ontological imperative*, for researchers interested in studying back-end metadata (e.g., social media metadata, such as Twitter metadata) born out of networked field sites. The ontological imperative framework poses five sets of questions that allow the researcher to critically examine their data, in essence serving as a who, what, when, where, and why. The five sets of questions require the researcher to fully explore the strengths and weaknesses of their data so that any epistemological claims that they make can be grounded in truly understanding their data. Although originally the ontological imperative framework was developed to guide researchers as they delve into the black box of algorithms and social media metadata in mixed methods studies, the framework can be used to shape and frame qualitative research that includes networked field sites, primarily because qualitative studies take place in networked field sites that include online spaces, actors, and participants that are temporal, transitory, and porous. This temporal, transitory, and porous nature of all online spaces requires the same level of examination by researchers doing qualitative research as with scholars conducting mixed methods studies that employ social media metadata. Of particular note, the ontological imperative framework requires researchers to provide readers with a level of transparency for their research and their data. The following sections will outline and demonstrate uses of the ontological imperative within networked field sites.

Technical outline of the ontological imperative in web-based qualitative research methods

The ontological imperative framework is guided by five sets of questions that allow researchers to examine the data collected from the digital tools, systems, and services in their studies. These questions, when used as a methodological framework, provide context for the researcher's findings. Because the philosophical roots of ontology are concerned with the nature of being, it is important to engage in understanding the very nature of what it means *to be digital*. In the case of research that occurs within digital spaces and networked field sites, the very nature of what it means to be digital and networked, given the temporal, ephemeral, transitory, and porous nature of the digital and networked, must be examined at every level of a study. The five sets of guiding questions of the framework are as follows:

1. What digital tools, systems, and services are at play in my study? Who created them and why?
2. What data do these digital tools, systems, and services render?
3. What hidden limitations might there be to the data rendered via these digital tools, systems, and services?
4. What are the epistemological implications of this ontological analysis?
5. What are the axiological implications of this ontological analysis?

Question 1: what digital tools, systems, and services are at play in my study? Who created them and why?

This question delves into the nature that all digital technologies come laden with biases, purposes, and motivations of specific individuals who created them as well as those who are using them. This question provides a framing for researchers to critique and question the purposes behind tools, systems, and services that are part of the networked field site that are part of a study.

Question 2: what data do these digital tools, systems, and services render?

This question forces researchers to think about what sort of data are created and available within the networked field site under study. How much of the data are rendered by humans (e.g., text, chat, images) and how much of the data are rendered by machines (e.g., governed by an algorithm or system design). This question also forces researchers to contend with what data might not be made available by the system and to challenge why this is so and whether the data that they are able to collect is useful for the questions that they are asking.

Question 3: what hidden limitations might there be to the data rendered via these digital tools, systems, and services?

This question forces researchers to contend with what data might not be made available by the system and to challenge why and whether the data that they are able to collect is useful for the questions that they are asking. Furthermore, this question asks researchers to consider shifts in data and space use over time and in temporal contexts. Because digital tools, systems, and services are in a constant state of flux, one must continuously examine the shifts and changes in their data corpus. In networked field sites and online spaces, data are always incomplete. Companies choose which data to make available, which parts of the system can be accessed by the public, as well as how the online spaces, actors, and data can be accessed and by whom. This question forces researchers to think through what data might not be available and to challenge their own pre-conceived notions about shortcomings of their current data from both current and historical uses of the tools, systems, and services.

Question 4: what are the epistemological implications of this ontological analysis?

This question provides space for researchers to account for what they have come to learn and know about the very nature of the digital tools, systems, and services in their study in order to ensure validity and trustworthiness of their findings. It challenges researchers to make connections among the previous three questions and to consider if other questions, data, or purposes should be considered for the study.

Question 5: what are the axiological implications of this ontological analysis?

This question requires that the researcher consider the ethical concerns brought forward by the analysis of the aforementioned questions regarding the digital tools, systems, and services. What levels of concealment are needed to protect participants whose data were collected from within often public digital spaces and forums? Are levels of concealment on easily searchable names, places, sites, and spaces needed? What might be missing from the analysis that might need more careful consideration?

Empirical demonstration of ontological imperative in web-based qualitative research methods

This next section will walk through applying the five sets of questions from the ontological imperative framework to data from a 5-year ethnographic investigation that I have conducted of the networked field site of a competitive collegiate Esports team (Gerber, 2022). Esports, short for electronic sports, is competitive, team-based videogaming that takes place across multiple levels—from amateur to professional. In particular, for over 5 years the purpose of my research has been engaged in studying the literacies, epistemologies, and meaning-making practices experienced by individuals engaged in the culture of collegiate Esports. This research has required me to trace meaning-making experiences across and through both analog and digital spaces, including using traditional methods (e.g., interviews and observations), as well as natively digital methods (e.g., calling an API). In the remainder of the chapter I will walk through applying the ontological imperative framework to the digital tools, systems, and services within my study.

Question 1: what digital tools, systems, and services are at play in my study? Who created them and why?

In order to study the literacies of competitive collegiate Esports, I had to become familiar with the main tools, systems, and services used by my participants. The main digital tools, systems, and services that were used in my data corpus came from *League of Legends*, Twitter, and Discord. For the purposes of considerations of space in this chapter, I will focus the ontological imperative only on an analysis on *League of Legends*.

League of Legends is a multiplayer online battle arena (MOBA) developed and published by videogame company Riot and was released to the public in October 2009. *League of Legends* is a MOBA-style videogame (Multiplayer Online Battle Arena), which in particular, MOBAs are created for the purpose of television broadcasting of competitions. The game is often played with teams of five players versus other teams of five players, known as a 5v5 match. Players can join the game either by entering a waiting arena for matchmaking and allowing the system to assign them to ranked matchmaking or they can play on pre-selected teams of five. Players each have their own avatar, known as a champion, that is equipped with certain skills, traits, and attributes. Because of this character selection, each person plays a specific role, much like a traditional sport, such as basketball or football where players play a specific part of the field or court with specific rules for their position. On the financial side, Riot is a for-profit company and the game uses a freemium model, which means that although the game is free to play, there are items for purchase by players for specific upgrades, items, or features to enhance their character's abilities.

Question 2: what data do these digital tools, systems, and services render?

At the surface, the user interface of the game provides in-game statistics such as leaderboards listing kills, deaths, assists. It provides markers of in-the-moment in-game play such as progress bars that show amount of health left, amount of mana (magical power for casting spells), menu of items (e.g., swords, spells, potions, etc.) that can be equipped by the player at any moment in time in the game if needed, as well as charts of statistics that show amount of time lapsed, time to respawn, and icons for each of their team mates that provides details on status of the team mate. Additionally, there is a map that allows a player to see where other team mates and opponents are at any given time in a match. Lastly, there is a chat window that serves as a channel for team mates to discuss in-game strategy as they play. By examining the features that are viewable on the surface, it is apparent that there are a plethora of data to be collected. See Fig. 1. This first level of analysis to know what data someone might encounter in a digital space is known as surfacing the data (Gerber and Lynch, 2017).

The next level of analysis that a researcher must do, before fully determining what data will be helpful to paint a robust picture, is to delve into the back end of the digital tool, system, or service. Although this might seem to be a complex activity that requires a researcher to be a programmer, developer, or a computer scientist, the activity of looking to see what backend data are currently is available is actually easier than it might first appear. Generally, companies have a repository that lists what data that they are making available to developers via the API. This is known as Developer Documentation and nearly every digital space that allows transfer of information between competing sites and apps has Developer Documentation on their website. In the case of my research that included *League of Legends*, this was located on a page titled "League of Legends Developers Portal." A deep dive into the developer documentation in December of 2020 showed 22 specific API Keys, which each key allows an individual to procure specific data about specific access points of the game. For example, there is an API Key that is Titled "Account-V1" that allows an individual to access specific backend metadata about user accounts including backend data about user ID. Another specific API Key is titled "Tournament-V4" which allows an individual to access tournament codes, lobby events, and player specific data for these events. To understand which categories are important for a researcher to examine, it should be noted that the terminology used in API calls such as "Get," "Post," and "Put" are all related to the methods used in API communications. A researcher looking to retrieve data from a site should focus on the categories of the API Key titled "Get" versus "Post" or "Put." Figs. 2–4 show examples of the deep dive into the *League of Legends* Developer documentation.

Although the aforementioned walk-through might appear a bit cumbersome, it is important to note what data are available to developers and individuals. These data can often be accessed by researchers via their own API calls, providing insight into different ways to ask research questions about the participants under study or to think about different data that might be available to help answer a question. A researcher might not know how to program, but can often work with a computer programmer (e.g., a student,

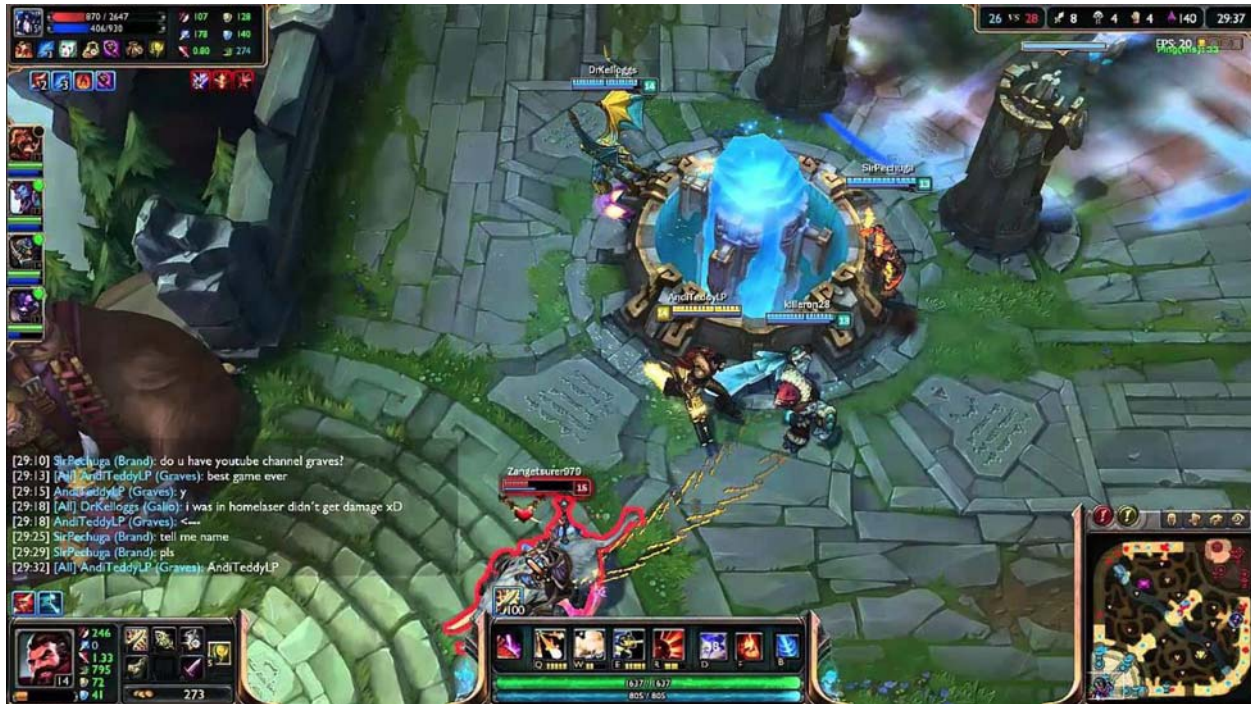


Fig. 1 Screenshot of data surfacing activity of league of legends

Note: This figure shows the various types of data available *on the surface*, when doing an analysis of League of Legends in-game screen. Various data points include in-game chat, status bars, in-game map, in-game statistics, team mate information, and available in-game equipment, potions, and weaponry.

faculty member, etc.) to conduct the API calls. Even if the researcher determines that they do not want to do the API call, it is powerful to know what data are being made available by the company, as this can lead to new insights about how to ask research questions based on data that the researcher might not even know is available (such as geolocations, time stamps, user metadata related to user networks, etc.).

Question 3: what hidden limitations might there be to the data rendered via these digital tools, systems, and services?

Question 2 of the framework encouraged researchers to engage in data surfacing and backend mining to understand what data were being made available. As noted, doing this sort of investigative work will allow a researcher to refine research questions, delve into data structures, and better understand how to study interactions with the tool, system, or service. However, as discussed previously, digital spaces are temporal, porous, and transitory. This means that what is available today might not be available tomorrow. Additionally, it should be noted that companies are not going to make available all of their data to developers or researchers. What one sees when they visit a company's Developer Documentation is only a partial glimpse into the data that are being collected from individuals who use that tool, system, or service.

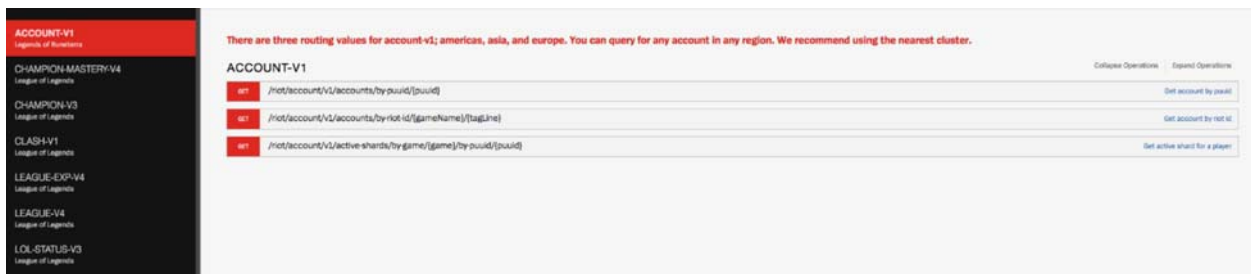


Fig. 2 API keys

Note: This figure shows a list of available API Keys on the left-hand side (note that the list is longer than shown), with three specific calls for that API key located in the middle of the Figure, as noted by the "Get" block. To the right of the "Get" block is the specific key followed by the far right which explains what sort of input criteria is needed to pull that information from the back end of the site.

There are three routing values for account-v1: americas, asia, and europe. You can query for any account in any region. We recommend using the nearest cluster.

ACCOUNT-V1

[Collapse Operations](#) [Expand Operations](#)

GET /riot/account/v1/accounts/by-puuid/{puuid} [Get account by puuid](#)

GET /riot/account/v1/accounts/by-riot-id/{gameName}/{tagLine} [Get account by riot id](#)

[Jump to Inputs](#)

RESPONSE CLASSES

Return value: AccountDto

AccountDto

NAME	DATA TYPE	DESCRIPTION
puuid	string	
gameName	string	This field may be excluded if the account doesn't have a gameName
tagLine	string	This field may be excluded if the account doesn't have a tagLine

RESPONSE ERRORS

HTTP STATUS CODE	REASON
400	Bad request
401	Unauthorized
403	Forbidden
404	Data not found
405	Method not allowed
415	Unsupported media type
429	Rate limit exceeded
500	Internal server error
502	Bad gateway

Fig. 3 Specific API key information

Note: This figure outlines the specific API call to get the backend data about specific accounts via the Riot ID.

As scholars, unless we work for the company for which we are studying, there is no way to know what other data are being collected, however, it is important to account for the fact that indeed other data might help to better answer that question and to think about what data a researcher would like to see, either on the surface via the user interface or on the back end. This type of critical questioning can help the researcher with the refinement of research questions. However, for purposes of transparency in the research process, when examining the available data (either through front-end surfacing or backend mining), it is important to note the temporal and transitory nature of data. As mentioned prior, the data that one collects today might not be available tomorrow, and it is important to track and trace shifts in data across time to understand the epistemological implications of one's research. The simplest way to understand how companies shift their access to data, or changes in types of data made or available or user interface structure is to think through the historical traces of the tool, system, or service and to use the Internet Wayback Machine (circa 2001) to do so.

The Internet Wayback Machine is an archive of most webpages dating back to 2001. As of this writing, the Wayback Machine has archived over 486 billion webpages. The Wayback Machine routinely crawls the Internet (using a web crawler) and archives and caches websites at specific intervals. This allows an individual to revisit an older version of a website to see what information was available at a key date in time. It is a useful tool for a researcher engaged in scholarship in networked field sites because it allows a researcher to note potential changes to available data, which might directly, or indirectly, influence how one arrived at his or her epistemological claims.

In order to understand my data on *League of Legends*, given that my study spanned 5 years, and to question shifts in my own participants' behaviors and literacy practices, I needed to examine some of the differences that have emerged on the site and in

The tournament-stub API is a stand in that simulates the behavior of the tournament API. Developers looking to apply for tournament API access should use the stub to mock their implementation before applying for a production key.

TOURNAMENT-V4

[Collapse Operations](#) [Expand Operations](#)

POST /ol/tournament/v4/codes [Create a tournament code for the given tournament.](#)

GET /ol/tournament/v4/codes/{tournamentCode} [Returns the tournament code DTO associated with a tournament code string.](#)

PUT /ol/tournament/v4/codes/{tournamentCode} [Update the pick type, map, spectator type, or allowed summoners for a code.](#)

GET /ol/tournament/v4/lobby-events/by-code/{tournamentCode} [Gets a list of lobby events by tournament code.](#)

POST /ol/tournament/v4/providers [Creates a tournament provider and returns its ID.](#)

POST /ol/tournament/v4/tournaments [Creates a tournament and returns its ID.](#)

Fig. 4 Specific tournament API keys

Note: This figure shows the types of calls that can be done with the Tournament-V4 API. Note calls to the API can be done that will "Post," "Get," and "Put." Researchers looking to retrieve back end data from a site should be concerned with API Keys that "Get."

the game over time. Using the Wayback Machine, I was able to access archived pages dating back to mid-2011 (see Fig. 5). Given that my study started in mid-2016, this year was the most relevant to start with for historical analysis. Historical access provided insight and reminder into parts of the game and community that are now defunct, such as access to the Tribunals, a community-driven system of removing toxic players from the game and access to developer documentation showing changes to API access and API data points.

Question 4: what are the epistemological implications of this ontological analysis?

A thorough analysis of the digital tools, systems, and services that were a part of a study allowed a deeper understanding of the production of data and the available data sources available. Question 4 of the framework is concerned with ensuring that researchers account for the nature of the digital data, products, or services that they use in their study in order to better support their epistemological claims. Primarily, this accounting allows for replicability of the study among other scholars. Replicability is the bastion of academic rigor, and although it is argued that qualitative inquiry is more difficult to replicate, if a researcher accounts for the processes of data procurement in digital tools, systems, and services, this provides a transparent audit trail for other scholars who wish to replicate the study.

In the case of my analysis of *League of Legends*, when examining the first three questions of the framework I developed a deeper understanding into the structure of the online space/networked field site (e.g., a for-profit company) which in turn influences the types of data made available for me as a researcher. This analysis allowed me to examine through front-end surfacing, back-end mining, and historical and chronological overview, insight into how my participants came to engage in the meaning-making processes and literacies practices within the *League of Legends* system. Understanding game updates, patches, and site upgrades,

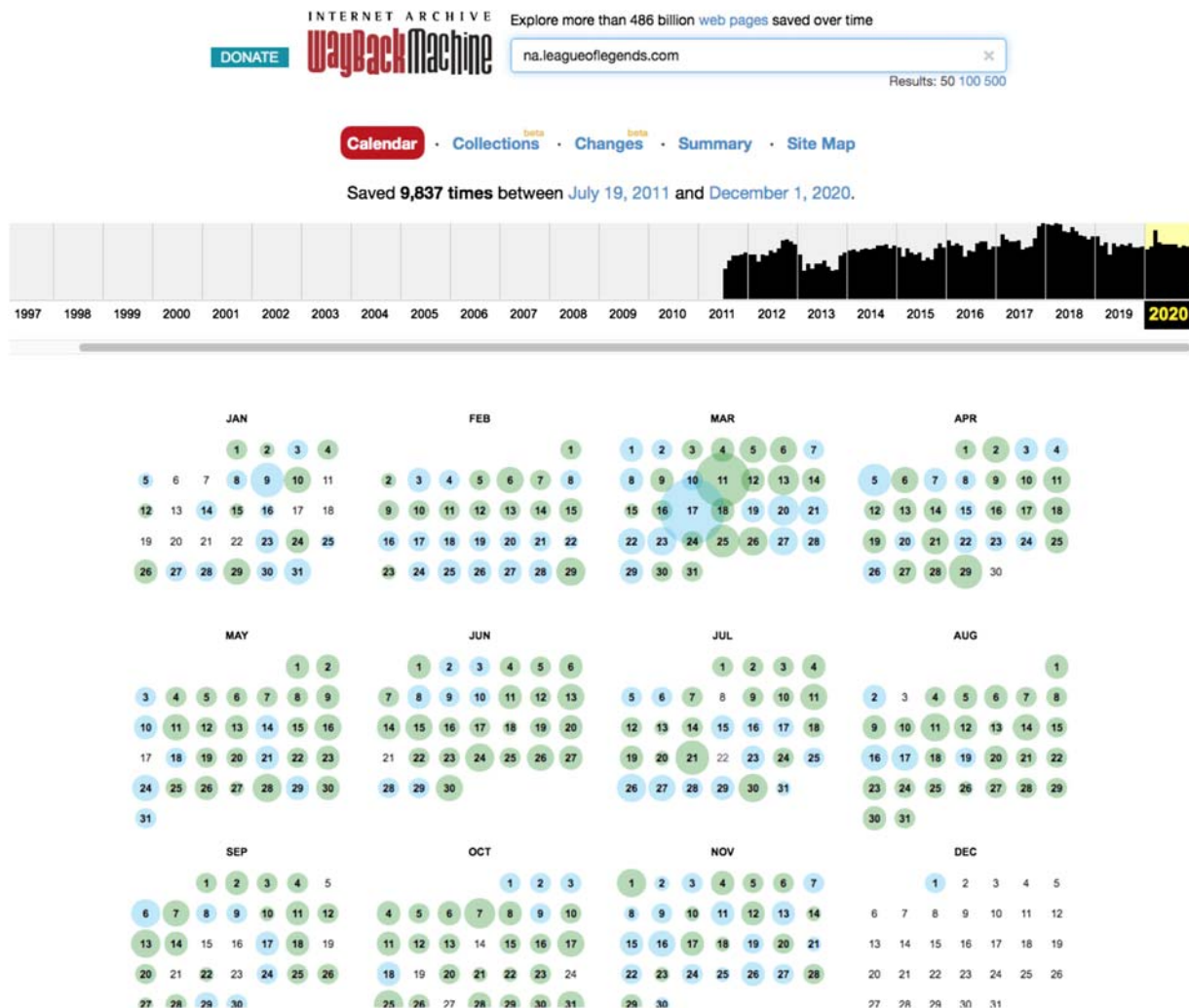


Fig. 5 Historical snapshot of the Internet Wayback Machine for league of legends

Note: This figure shows the historical archive of the *League of Legends* website dating back to 2011. The bubbles on each date are directly proportionate to how many snapshots, or archived crawls were captured on a given date.

which are all traceable through the aforementioned methods, allowed me to dig deeper into understanding the data that I had collected in order to trace the way that information has flowed among and between players and to in turn, reframe my research purpose as well as re-examine the types of data that I thought I originally needed to answer my research questions (e.g., looking at different metrics of in-game success as well as branching into different online social spaces and networked field sites, such as Discord).

The changes that have occurred in site design and purpose (i.e., the loss of the Tribunal space), the changes in the types of data visible on the surface and user interface, and/or the types of data made available in the back end via the API over time, all influenced the epistemological claims that I, as a researcher in this networked field site, was making. Accounting for these changes and shifts provided an audit trail for me to reflexively visit, which provided a more robust framing for the study's trustworthiness, which allowed me to more confidently state claims and posit implications about my study's findings.

Question 5: what are the axiological implications of this ontological analysis?

Although axiology deals with value, the focus of axiology in the ontological imperative is concerned primarily with ethics. Understanding the ethical implications of research that occurs within a networked field site is crucial. Whenever a researcher enters an online space or uses digital tools, systems, or services, with the main mission to study participants within the space, the researcher risks divulging sensitive participant information if due caution is not exercised.

For example, API data (e.g., back end system data) are generally considered public and scholars have argued that researchers should be able to freely use these data to help inform their studies (Manovich, 2012). However, data that at one point might be considered private data under password protection, might become publicly available if the company changes the Terms of Service or if the company is acquired by another company (Gerber et al., 2017). Because of this, researchers should think through ethical considerations as they engage in their ontological imperative analysis. Even if participants have signed informed consent, it is imperative that the researcher consider future harm that their research might have if data shifts, if acquisitions occur, and if Terms of Service changes make that data public. Guiding questions such as,

- To what extent is these data collection and analysis ethical?
- Do the Terms of Service allow for this use of the space and data collection from the space?
- What harm could these findings have on my participants if any of specifics of the study were suddenly made publicly available?

In order to mitigate risks, with my study of *League of Legends*, once I completed the ontological imperative analysis, I engaged in Bruckman's (2002) levels of concealment (see also Gerber et al., 2017). In particular, once I engaged in historical analysis through the Wayback Machine, when paired with my front-end surfacing, and back-end data mining, I determined the most appropriate level of concealment would be to engage in not just use of pseudonyms, but renaming any @mention, paraphrasing direct quotes from any game chat or messaging system, and concealing or changing specific mentions to sites or locations. Ensuring transparency in the methods used (i.e., by providing a walkthrough of the ontological imperative), as well as concealment of the actual data (i.e., by using levels of concealment of sensitive participant data), provided a deeper ethical treatment and consideration of my participants' data than simply reporting the data in its initial form and hoping that shifts in Terms of Service or data availability would not change in the future. These ethical considerations are crucial for all researchers who engage in research within networked field sites.

Ways to continue growing as a researcher in web-based qualitative research methods

The very nature of conducting research of networked field sites using the ontological imperative framework requires researchers to change the way that they think about research in online spaces and subsequently how they approach their research. Beyond simply accounting for the ontological imperative in the methods section of one's article, dissertation, or manuscript is acknowledging that one should be continuously learning about new methods of scholarship and inquiry. The following concrete examples offer suggestions for how researchers can continue to grow as scholars doing qualitative inquiry in online spaces and networked field sites.

Read up on digital methods and theories of interaction in digital spaces

Researchers interested in continuing to learn about qualitative research methods in online spaces should begin by digging into theories proposed in the arena of software studies and digital methods. Software studies provides a foundational theory for understanding networked field sites and research in online spaces. Furthermore, researchers interested in furthering their understanding of *digital methods* (e.g., natively digital, such as web crawling or API calls) versus *digitized* (e.g., online surveys) should invest in reading about digital methods. The following four books provide a primer into software studies and digital methods.

The first book, *Code/Space: Software and Everyday Life* (2011) by Rob Kitchin and Martin Dodge provides an overview of how software and computer code permeate every aspect of our lives, from the way that we grocery shop to how we decide where to go on vacation. Their text is part treatise and part manifesto for understanding how our lives are governed by software and what we can do to understand the social, economic, and political implications of living within a software space. Second, *Software Theory: A Cultural and Philosophical Study* (2015) by Federica Frabetti deconstructs software and code to show that at the center of the software that powers our lives, and central to the code that dominates software, are human languages. Software, and at its very base,

code, is indeed an incredibly human creation governed by rules of language, syntax, and grammar. These same rules are therefore then enacted with human biases and limitations. The third text, *Digital Methods* (2013) by Richard Rogers is a fantastic primer on the various digital methods available to researchers. Offering a historical overview of a variety of approaches such as web-crawling and API calls, Rogers noted "... [digital methods] seek to ground claims about cultural change and societal conditions in web data" (2013, p. 19). Finally, *Quantitative Ethnography* (2017) by David Williamson Shaffer is a must read that explores the qualitative approach of ethnography and pairs the approach with Big Data.

Conduct software and app walkthroughs

One activity that can help a researcher better understand the tools, services, and systems of a networked field site is to engage in a software or app walkthrough. A walkthrough is a process that allows a researcher to develop new eyes for the digital tools, services, systems, and spaces that he or she is studying. The process should progress from onboarding all the way through exit or leaving a system.

To do a thorough walkthrough, a researcher should ensure that they have screen capture software ready and that they are logged out of the tool, system, service, or space. This process requires that the researcher set up a new account for whatever tool, system, service, or space that they are studying in order to give them fresh eyes and a new perspective as they walk through the app or program. The researcher should document and annotate the entire process for later review. The process starts with the onboarding into the tool, system, or service which includes signing up for an account. As the researcher signs up for an account he or she should screen capture, annotate, and document what he or she is noticing, including issues, and/or ideas that he or she has about the user interface. Walking through the entire platform, the researcher should document and annotate the entire process. This walkthrough will help spark ideas for areas of later exploration and/or deeper understanding of the user interface and its impact on a study's participants.

Explore developer documentation for different spaces

Similar to the software and app walkthrough, researchers should also engage in exploring the developer documentation of any of the tools, systems, and services that are part of the networked field site that they are studying. By digging into the developer documentation, a researcher can develop a better understanding for what data are available, which can help to shape potential research questions and/or influence the research purpose. Similar to the software and app walkthrough, researchers should be documenting and annotating findings and insights, either via screen capture or recording software. These annotations can become part of later reflexive journaling and audit trails.

Connect with colleagues interested in research in digital spaces

As scholars engage in learning about conducting qualitative research in online spaces, it is important to join communities of researchers and scholars who are also engaged in learning about and developing new methods for studying online spaces. There are several existing organizations and/or groups focused on the study of methods for researching online spaces. The Association of Internet Researchers (AoIR), founded in 1999, is an international organization made up of researchers from across multiple disciplines who are dedicated to the critical and scholarly study of the Internet. AoIR hosts an annual conference as well as regular webinars and produces academic material and guides to aid scholars in their design of studies—both novice and experienced. For more information visit <https://aoir.org/>. Another organization, the Digital Methods Initiative (DMI), is based out of Amsterdam. DMI has created over 50 web-based applications and pre-built tools for working with back-end metadata from a variety of digital tools, systems, and services such as Google, Facebook, Twitter, Instagram, Amazon, etc. Furthermore, DMI hosts two events a year aimed at teaching people how to conduct research using digital methods. For more information visit <https://wiki.digitalmethods.net/Dmi/DmiAbout>.

Final thoughts

In developing new methods and approaches to studying online spaces, it is important that researchers concerned with studying networked field sites should strive for *the radical middle* (Onwuegbuzie, 2012).

... which should not be as passive and comfortable middle space wherein the status quo among quantitative and qualitative epistemologies is maintained, but rather a new theoretical and methodological space in which a socially just and productive co-existence among all research traditions is actively promoted, and in which mixed research is consciously local, dynamic, interactive, situated, contingent, fluid, strategic, and generative (p. 192).

Developing new methodological approaches for use within networked field sites, particularly when framed through the radical middle, means researchers must critically question every nuance and aspect of the online spaces that they and their participants

inhabit. This includes considering the flow of information and the fact that information is not static in these spaces, understanding that there is a middle ground between the online space of the physical and the digital that allow for the creation of that networked field site. This is the radical middle in studies that take place in online spaces. Although Onwuegbuzie's theory of the radical middle is primarily concerned with mixed methods research, we can learn a lot from how the radical middle can influence our understanding of qualitative approaches in online spaces. Learning to adopt new methods and approaches to understand interactions in online spaces is beyond simply adapting existing qualitative approaches. The radical middle philosophy provides a framing for why the ontological imperative is such a crucial framework for scholars to consider when conducting qualitative research in online spaces and networked field sites. Scholars must adopt a stance and frame of mind that allows them to interrogate the very tools, systems, and services in their study to look for inherent biases and limitations in the data and to expose these biases and limitations in order to ensure transparent scholarship that is not only replicable, but opens up a venue and space where research is subsequently, "... consciously local, dynamic, interactive, situated, contingent, fluid, strategic, and generative" (Onwuegbuzie, 2012, p. 192). Only then, once a researcher has completed an ontological imperative analysis, can a researcher's epistemological claims hold weight.

References

- Bruns, A., Burgess, J., 2013. Crisis communication in natural disasters: the queensland floods and christchurch earthquakes. In: Bruns, A., Mahrt, M., Weller, K., Burgess, J., Puschmann, C. (Eds.), *Twitter and Society* [Digital Formations, Volume 89]. Peter Lang Publishing, United States of America, pp. 373–384.
- Frabretti, F., 2015. *Software Theory: A Cultural and Philosophical Study*. Rowman and Littlefield, London.
- Gerber, H.R., 2022. The literacies of a competitive esports team: livestreaming, vods, and mods. *L1 Educ. Stud. Lang. Lit.* 21, 1–25. <https://doi.org/10.17239/L1ESLL-2021.21.03.02>.
- Gerber, H.R., Lynch, T.L., 2017. Into the meta: research methods for moving beyond social media surfacing. *TechTrends* 61, 263–272.
- Gerber, H.R., Abrams, S.S., Curwood, J.S., Magnifico, A.M., 2017. *Conducting Qualitative Research of Learning in Online Spaces*. SAGE Publishers, Thousand Oaks, CA.
- Internet Wayback Machine, 2022. <https://archive.org/web/web.php>. (Accessed 30 April 2022).
- Kitchin, R., Dodge, M., 2011. *Code/Space: Software and Everyday Life*. MIT Press, Cambridge, MA.
- Lynch, T.L., Gerber, H.R., 2018. The ontological imperative when researching the digital age. *Int. J. Mult. Res. Approaches* 10, 112–123.
- Manovich, L., 2012. Trending: The Promises and Challenges of Big Social Data. <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfd1e/section/65b49184-0ab4-4f29-859c-f7f02d875d8d>.
- Marr, B., 2018. How Much Data Do We Create Everyday? The Mind-Blowing Stats That Everyone Should Read. *Forbes*. <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#:~:text=The%20amount%20of%20data%20we,in%20the%20world%20was%20generated>.
- Onwuegbuzie, A.J., 2012. Putting mixed back into quantitative and qualitative research in educational research and beyond: moving toward the radical middle. *Int. J. Mult. Res. Approaches* 6, 192–219.
- Rogers, R., 2013. *Digital Methods*. MIT Press, Cambridge, MA.
- Schaffer, D.W., 2017. *Quantitative Ethnography*. Cathcart Press, Madison, WI.
- We Are Social, 2020. *Digital 2020 Global Digital Overview: Essential Insights Into How People Around the World Use the Internet, Mobile Devices, Social Media, and E-Commerce*. We Are Social Group Research Group. <https://wearesocial.com/digital-2020>.

Further reading

- Gerber, H.R., Lynch, T.L., Onwuegbuzie, A.J., (forthcoming). *Making Big Data Small: Integrating Digital Approaches for Social Science Researchers*. SAGE, Thousand Oaks, CA.
- Loukissas, Y.L., 2019. *All Data are Local: Thinking Critically in a Data-Driven Society*. MIT Press, Cambridge, MA.

Relevant websites

- Association of Internet Researchers, <https://aoir.org/>.
- Digital Methods Initiative, <https://wiki.digitalmethods.net/Dmi/DmiAbout>.

Qualitative research methods: approaches to qualitative data analysis

Paul Mihas, Howard W. Odum Institute for Research in Social Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	302
Overview of analytic practices and strategies	302
Coding practices	303
Memo-writing practices	304
Inductive thematic analysis	304
Qualitative research approaches	305
Case studies	306
Ethnographies	307
Grounded theories	308
Narrative analyses	309
Phenomenologies	310
Conclusion	311
References	311

Introduction

Qualitative analysis—the interpretive analysis of textual, graphical, or audio/video data—occupies a continuum from description to theory building, from addressing *what* questions to attempting to answer complex *how* and *why* inquiries. Though researchers might discuss analysis as a phase, it is best to anticipate analytic strategies even during design and data collection. For example, if we are interested in lived experience, discussed later in this chapter, we need to generate data that provide access to the nuanced texture of an experience as it is being lived. But we would also benefit from developing an analytic perspective as we collect data so that we can begin examining data as we collect them, such as writing memos on especially meaningful text segments.

Qualitative studies can be directed by a conceptual framework, suggesting a deductive thrust, or driven more by the data itself, inviting an exploratory, inductive lens. In either case, the topics we wrestle with can range from tangible to abstract. How we engage these topics varies dramatically based on our strategies for analysis. A descriptive study might identify and illustrate barriers and facilitators of a public health intervention whereas an inductive thematic study would raise topics, or combinations of topics, to a higher conceptual level. Theme-building allows researchers to make interpretive claims by weaving together evidence across data and data types and naming these dynamic links. A theoretical study would go even further to make explanatory claims based on evidence across data and data types, attempting to theorize relationships between individuals, structures, and processes and not simply describe a situation, event, or experience.

In addition to presenting thematic analysis as a strategy, this chapter also covers established qualitative traditions: case study, ethnography, grounded theory, narrative analysis, and phenomenology. Case study research involves identifying a bounded “case” and collecting multiple types of data to assess the case and its context, either descriptively or more thematically. Ethnographic research focuses on a cultural setting and uses rich description to understand how individuals participate in the same culture, perhaps in markedly different ways. Grounded theory encompasses several approaches, including objectivist and constructivist, that invite researchers to analyze process, contexts, structures, and consequences. Narrative analysis treats stories not only as possible representations of events, but as narrative events in themselves. Researchers using this approach analyze the form and content of narrative data and examine how these elements serve the storyteller and how the story shapes ontologies and epistemologies. Phenomenology is an approach designed to “open up” a phenomenon and make sense of its invariant structures, its identifiable essences across narrative accounts. In this approach, the focus is on the lived experiences of those deeply familiar with the phenomenon. Each of these traditions has its own investigative emphases and particular tools for analysis—specific approaches to coding and memo writing—and conventions for final products, such as diagrams, matrices, and condensed reports.

Overview of analytic practices and strategies

This section focuses on particular analytic practices—coding and memo writing—as well as a common analytic strategy, inductive thematic analysis. How analytic practices are adopted depends on the conceptual “compass” a researcher might bring to a study. There are numerous compasses or possible points of departure, depending on the current knowledge in the field as well as the particular focus of inquiry. The point of departure is the combined knowledge and objectives that guide the initial phase of inquiry.

There may be little established knowledge in an exploratory study or an extensive body of research that already speaks to a research question.

Researchers using a deductive approach might make use of an existing theoretical framework, such as activity theory, which sees activity as purposeful and transformative, or they might construct a conceptual framework based on a directed literature review. Particular fields, such as implementation science, offer specific guidance in the form of previously existing codebooks, such as the consolidated framework for implementation research (CFIR) (Damschroder et al., 2009), a collection of constructs reflecting the uptake of research findings into routine healthcare. When a theory or a conceptual framework is present in inquiry, these frameworks become a compass of sorts guiding the analysis. Researchers may then illustrate how their data align with or problematize the theory or conceptual scaffolding with which they began.

In the absence of a theoretical or conceptual framework, researchers are guided by a more data-driven approach, where the language of the data and the way that participants frame events, structures, and outcomes become topics that guide the analysis (Maietta et al., 2018; Saldaña, 2021). Though this approach might be considered more “exploratory,” it is not an unmoored one. Exploration is governed by research questions, by disciplinary understanding, and by a particular paradigmatic lens. For this reason, documenting one’s point of departure is imperative. It gives a researcher space to acknowledge how prior knowledge can be both a strength and a drawback. A researcher may need to “un-know” and un-learn what they have already learned in order to gain insight into the individuals’ experiential framing (Vagle, 2018). In an inductive approach, researchers experience the shifting sands of emergent ideas and build scaffolding for the study as they go. Entering the “unknown” privileges participant knowledge rather than relying on researcher knowledge or cumulative understanding based on pre-existing frameworks.

Reading textual data with an interpretive lens is itself a skill requiring an eye for nuance as well as holistic contiguity. In acknowledging the contiguity of an interview or focus group, we see how a participant or group has created a holistic account even if our questions were not attempting to generate holism. Attending to the participants’ framing can help us understand their complex experiences, governing values, or psychological landscape. For example, a participant who references religious beliefs throughout an interview even when the questions are not about religion can help us understand how a religion meaning system permeates their everyday experiences. In inhabiting data, we are invited not to be critical of participants but to understand events as they understand them—to analyze, not necessarily psycho-analyze—to consider the narrative territory of the interview in terms of ideas such as primacy (what comes first in a narrative), negation (negative language), emphasis (text treated with heightened regard), and repetition (repeated phrases) (Alexander, 1988). A narrative landscape, like a literal landscape, has different terrains, transitions, and elevations. We might notice places where participants center particular values, decenter other values, or demonstrate negotiated values (Daiute, 2014), such as librarians who are asked to catalog a title using categories that they find problematic because of their racist overtones. Prompts for attentive reading include: What is at stake here for the participant? What does the participant reveal beyond the “facts” of an incident? How does their way of speaking or constructing an account differ from how other participants in the study share their experiential accounts?

Coding practices

Many researchers have discussed possible coding practices for qualitative research (Charmaz, 2014; Miles et al., 2020; Saldaña, 2021; Strauss and Corbin, 2008). Codes are a vehicle for paying attention, a way for researchers to help establish consistency as they review and name pieces of data. If used intentionally and strategically, coding is a way of knowing and making meaning of copious amounts of data. In some studies, codes can be developed even before data collection, as would be the case if using the consolidated framework for implementation research (CFIR). Even without a framework, researchers might nonetheless develop an initial codebook based on the research questions, interview questions, and study aims.

In contrast to deductive coding, inductive coding is based on discerning and naming “surfacing” topics in close reading. For example, in vivo coding uses the participant’s own language to name a topic. In this practice, a concisely expressed belief, such as “life is empty without family,” becomes the name of the code itself, rather than introducing the researcher’s academic language that may lose the intensity of the quoted text. Another inductive coding practice, process coding, uses gerunds, or “ing” nouns, to capture actions and behaviors evident in the data (Charmaz, 2014; Saldaña, 2021). This practice helps researchers move closer to the participant experience—what they are doing or going through—rather than focusing on nominalizations that may ignore process. This coding technique can ultimately help us study what keeps a behavior or action in place, what reinforces it, and what destabilizes it.

We might also make a distinction between descriptive and interpretive coding. In descriptive coding, we take inventory of the tangible *who*, *what*, *where*, and *when* of data—emotions, settings, relationships, and timespans. However, focusing on description may result in a flat analysis without generating deeper insight. Interpretive coding, in its more conceptual nature, invites us to distinguish the implications and underlying meaning of what is being stated, moving closer to addressing *why* and *how*, but also reconceptualizing the *who*, *what*, *where*, and *when* in conceptual terms. So, we might rethink the codes *teachers* and *parents* in terms of *role models* or *dream makers*. Likewise, instead of coding for *frustration*, we might code for *managing emotions*, a more interpretive code that gets at an underlying process, a more meaningful topic with analytic leverage across transcripts.

Coding is an analytic act, but not a destination. Once we have coded data, more interpretive labor lies ahead in assessing what is “inside,” “outside,” and “in between” the codes. Assessing what is “inside” a code refers to evaluating the text segments contained within it to assess the data cumulatively regarding the spectrum of voices within, and properties of, the code. For example, we might find a range in the magnitude of the code, from low to high intensity. We might similarly assess what lies “outside” the code, that is,

what contextualizes or frames it. For example, we might discern that *managing emotions* is often coded within a particular time period or event, such as *turning points* in a career trajectory, contextualizing conflicting emotions within temporal markers. Finally, in investigating what happens “between” codes, we might see that two codes reveal more together than they do separately. For example, we might discern that *home life*, when assessed alongside *work life*, provides insight into a new concept, *actual self vs. ideal self*. These examples are reminders that though codes may seem like self-contained ideas, they become building blocks for further analytic inquiry when we look at the dynamic relationships between them (and contextualize codes within enveloping ones). We find other stories to tell and higher-level concepts that bridge seemingly unrelated topics (Maietta et al., 2021).

Memo-writing practices

Qualitative data collection can generate overwhelming amounts of textual data that require reliable practices to help us condense and synthesize our understanding. Coding is a practice that helps us consolidate this text into more manageable units and topics. Memo writing, in parallel, is another practice that can help us formally monitor and develop our evolving ideas as we move across data. The two practices differ in that codes tend to track individual, separable topics whereas memos function as a place to sustain complex inquiry, a space to draw connections between quotations, documents, or codes.

Types of memos include a *document reflection memo*, which treats an interview or other transcribed encounter as a contiguous whole that reveals consistent threads regarding values, beliefs, and attitudes or, of equal interest, contradictory threads regarding these topics (Mihás, 2022). Though we might have asked targeted questions in a structured or semi-structured interview guide, the landscape of an entire interview can help us understand the big picture, and as a result, we might become more aware of chronological trajectories, belief systems, or psychological states that take shape during an interview, insights we might otherwise miss in focusing solely on textual fragments.

In addition to coding text segments, *memos on key quotations* can help us better understand illustrative pieces of data. A key quotation is a text segment that is especially informative in that it reveals a way of thinking, behaving, or experiencing a phenomenon. When an opioid user says, “every ‘never’ I ever said came true,” we might pause to open up these words and explore what they mean for the participant and the larger research inquiry (Mars et al., 2014). This kind of reflective memo allows us to elaborate on what participant language is teaching us and how it helps provide insight into the intricate insider perspective. The content of these memos can also be used to develop codes.

Cross-document *comparison memos* move beyond document-level or segment-level quotations to synthesize meaning across data. By creating an intentional space to compare a particular topic—an experience, belief, or value—and its dimensions, this memo serves as a place of growing insight that also functions as pre-writing for a more formal paper or report. That is, when we have a hunch in one interview, we might intentionally write about it and refine the memo, and our proposition, as we review more data. In looking at the experiences of US parents whose children were soldiers in Iraq during the early years of the Iraq War (2003–2011), a comparison memo might show, for example, that mothers inhabit a limbo state of “what-if” thinking versus news of actual events (Collins, 2008).

Coding itself can also help us compare a topic across data, but it is sometimes difficult to code for an idea that permeates an entire transcript or to code for what is happening “between” text segments, that is, to code for the subtextual “rhyming” that happens when one segment is juxtaposed against another. A comparison memo can help us keep track of these slippery concepts that require elaboration as we find examples that offer complementary or contradictory perspectives.

Invariably, most qualitative findings show up as written claims. Memo writing as a practice can help us develop a writing voice that is congruent with our data and a way to interrogate evidence for a claim. In speaking with, not for, participants and braiding our voices with participant voices, we are more likely to stay close to the data and present co-constructed insights that serve both academic journals and communities from which the voices arose. For this reason, engaging in memos throughout the entire research lifecycle can help us invest in writing as a “way of knowing” (Richardson, 2000). This makes the task of final reports and papers more efficient because we have already developed the textured language of the study, rather than reducing writing to “write-up.”

Inductive thematic analysis

Thematic analysis, which can include both deductive and inductive practices, was developed across disciplines throughout the twentieth century, including work in physics by Gerald Holton (1973). In more recent years, psychologists Braun and Clarke (2006) outlined a set of theme-building procedure for the social sciences, but thematic analysis remains a loosely defined practice with many varied applications across fields of study and across a “range of theoretical and epistemological approaches,” such as case study and ethnography (Braun and Clarke, 2006, p. 6). Though some researchers argue that thematic analysis is a method in its own right, others, such as Morse and Cheek (2021) and Boyatzis (1998), consider thematic analysis a practice or a tool. As such, thematic analysis is not exclusively tied to a particular epistemology or paradigm. If we adopt a realist paradigm to assess the barriers of emergency medicine, we might construct themes that elucidate treatment barriers. If we instead adopt a constructivist paradigm, we might develop themes to capture social structures evident in an emergency setting, such as how built environment “acts” upon patients.

Though deductive approaches to thematic analysis, such as directed content analysis, are common (Bingham and Witkowsky, 2022; Hsieh and Shannon, 2005; Mayring, 2015), they are beyond the scope of this chapter. Instead, this section focuses on *inductive*

thematic analysis, whereby researchers do not begin with theoretical constructs or frameworks. They start only with provisional topics that work their way into an evolving codebook. During the initial review of data, called “open coding” or first-cycle coding, researchers generate numerous topics, both broad and specific (Charmaz, 2014; Saldaña, 2021). Open coding, often descriptive, invests in the language of the participant as well as the aims of the study. However, thematic analysis does not stay at the level of description or initial insight. Researchers move up the conceptual ladder from this vast array of ideas to develop conceptual clusters using evocative, condensed language.

Codes and themes occupy different semantic planes. A code is a container for a single topic, whereas a theme goes further in capturing dimension or meaning across multiple codes and in this way acts as an “argument” about the “phenomenon being examined” (Braun and Clark, 2006, p. 18). For this reason, topics such as *barriers* and *facilitators* are regarded as broad topics that do not adequately elevate ideas to a more compelling claim or theme, which tends to be more evocative, such as the *family obligations intensify transportation barriers*. In other words, themes pierce the surface of data to evoke a higher-level story or abstract concept. They go deeper in unearthing participants’ tacit assumptions and pervasive logics and in connecting seemingly different topics, such as family and transportation in the above example, and naming this dynamic link.

There is no single agreed upon strategy for determining themes, but researchers might create clusters of seemingly related codes and then name the conceptual glue that holds them together. Hence, in a study on experiencing chemotherapy, descriptive codes for *eating new foods*, *seeking new entertainment*, and *rethinking life choices* might be tied to a theme, *emerging self*, that functions as a conceptual bridge between or among codes. In studies using structured interviews, we often find ourselves with pre-existing topics. For example, in a study of indigenous people experiencing climate change in Alaska, we might have anticipated at the outset topics such as *physical health* and *mental health*. These are simply topics based on research questions. They are not themes. To construct a theme, we might review the quotations coded to each code and assess what we perhaps have not noticed in reviewing each code separately. We might discern that across these codes participants address issues of the *self as a marker of climate change*, a thematic bridge that connects the separate notions of physical and mental health.

Braun and Clarke’s thematic analysis method is an iterative process consisting of six steps: (1) becoming familiar with the data, (2) generating codes, (3) generating themes, (4) reviewing themes, (5) defining and naming themes, and (6) locating exemplars. Becoming familiar with data refers to data immersion and iterative cycles of reading, with each cycle generating further insight. Generating codes means coding for as many topics as possible and applying the code to a contextual segment, not just a phrase. Generating themes means sorting the codes into higher-level topics. Researchers might use tables, mind maps, or theme piles to cluster topics into these broader groupings. This phase ends with “candidate themes” and subthemes. Reviewing the themes means interrogating the candidate themes by revisiting the data coded to the component codes. We may determine that the data do not sufficiently support the theme or that there is too much variation across text segments to justify the theme. This may mean renaming the theme or making it a sub-theme of a broader construct. The next phase entails refining the names of themes and ensuring that they occupy the same semantic plane, i.e., that they are conceptually parallel. If the language of most of the themes is broad and abstract (e.g., collectivity and the communal self) but other ideas are more surface-level (e.g., keeping busy, reading a book, playing a game), this may require creating a higher-level construct to elevate the superficial topics to more abstract ones (e.g., suppression and distraction) (see Goodman, 2004). Producing the report means telling the complex story of the themes, first describing the meaning within each theme, with illustrative examples, and then perhaps looking across themes to discern connective takeaways or meta-themes.

Braun and Clarke (2006) state that a theme captures a prominent aspect of the data in a patterned way, “regardless of whether that theme captures the majority experience” (Scharp and Sanders, 2019, p. 1). Though identifying a pattern might entail noticing frequency, it is primarily about meaning making, not underscoring quantity. There is no “hard-and-fast answer to the question of what proportion of your data set needs to display evidence of the theme for it be considered a theme” (Braun and Clarke, 2006, p. 10). In other words, the identified themes address the research question, even if not every participant addresses the theme per se.

In a study of men managing body image and appearance, researchers identified four themes: *men value practicality*, *men should not care how they look*, *clothes are used to conceal or reveal*, and *clothes are used to fit a cultural ideal* (Frith and Gleeson, 2004). Each of these themes addresses the study’s overarching question—how men use clothing. Each theme tells a story based on these abstract concepts. The theme *men value practicality* underscores how men valued the functional nature of clothes, what is “necessary for daily living” (p. 43). Each theme is presented with quotations that add texture to the concept and provide evidence for how the researcher conceptually elevated numerous textual fragments, identified a crystalized pattern or distinguishable shape across different textual accounts. In this way, themes use the idiographic to construct the nomothetic. In some cases, researchers can use the participants’ own language to name the theme, as was the case in a study on whether patients tell their physicians that they are depressed, where themes included “My doctor just picked it up” and “They just check out your heart and things” (Wittink et al., 2006).

Qualitative research approaches

The next section of the chapter presents five established qualitative research approaches: case studies, ethnographies, grounded theories, narrative analyses, and phenomenologies.

Case studies

As others have noted, defining a “case study” is a “definitional morass” given that it can be considered a form of data collection, an approach to analysis, or a type of reporting or write-up, or all these (Gerring, 2007, p. 17; Savin-Baden and Howell Major, 2013). As an approach to qualitative research, a case is intended to have the following properties. First, it is bounded by time and place, such as the months following the rollout of a vaccine in a particular rural county, and has a “real-life,” naturalistic context. However, in the age of virtual spaces and online communities, the notion of “boundedness”—and culture itself—has become increasingly problematic given that individuals can participate in different cultures with complex boundaries simultaneously (Bartlett and Vavrus, 2017). Second, a case study employs methodological triangulation, the use of multiple data collection methods—interviews, focus groups, documents, observations, visual data, and online resources—to contribute to interpretive precision or a comprehensive understanding of an event, episode, organization, or individual. Third, a case has an identifiable context, whose boundaries with the case itself are sometimes difficult to discern, such as a school and the district in which it resides (Stake, 1995, 2006; Yin, 2018).

Case study analysis occurs at the level of a particular semantic plane, or level of abstraction. Analysis can focus on description, such as an account of school bullying at a particular high school, or move into more abstract conceptual territory, such as analyzing identity formation of the bully and victim. The level of abstraction depends on the research question, our intended audience, and whether our aim is rich description—the tangible *who*, *what*, *where*, and *when* of analysis—or also theory building—the *how* and *why* of conceptualizing.

A more explicitly comparative approach would be used in a multi- or comparative case study. Case comparisons can be homologous or heterologous (Bartlett and Vavrus, 2017) where homologous comparisons employ entities having a similar position or structure (such as two colleges with a large number of first-generation students) and heterologous comparisons assess entities that are categorically different or high-contrast (such as a retail store and a clinic) but are studied in order to understand a more abstract phenomenon, such as organizational death. In analysis, researchers would compare data and codes across cases (sites, episodes, or individuals) to investigate similarities and differences and provide evidence for building a logic model or other type of conceptual model, such as a process model of organizational death in a study of dying organizations (Sutton, 1987).

Analyzing a case study requires us to keep track of the complexity of the different units within our data collection, the case itself, and its context. The unit of data collection is often different from the unit or units of analysis (Yin, 2018). We might interview individuals or gather website mission statements to understand an organization, the ultimate unit of analysis. That is, our analysis must always remain attentive to the case itself—drawing conclusions at the case level—and avoid getting lost in the minutia of the units of data collection. Coding for the emotions of individuals in the above example would be of use only if it helps us understand the environment of the organization. Coding for components of participant life histories would be necessary only if the participants themselves were the cases, rather than the units of data collection. However, a particular kind of case study, the embedded case study, would allow us to study smaller units within the larger case, such as classrooms within a school.

Case studies have been characterized as variance oriented, interpretivist, and process oriented (Bartlett and Vavrus, 2017). Variance-oriented case studies, using a post-positivist framing, are concerned with variables and correlations among them (Maxwell, 2013) whereas interpretivist studies attempt to understand participants’ “sense-making” of events or phenomena. In contrast to these, process-oriented case studies tend to seek explanation based on an “analysis of how some situations and events influence others” (Maxwell, 2013, p. 29).

Researchers have outlined different approaches to analyzing case studies, including using a theoretical proposition, using an exploratory approach, developing a case description, and assessing rival explanations (Yin, 2018). Using a theoretical proposition does not necessarily mean a strict prediction involving a dependent and independent variable. Rather, it points to a kind of hunch that we seek to examine qualitatively with multiple data types. Analytic practices in this situation would necessitate linking data to proposition(s) through various strategies, including presenting a descriptive account of a site’s atmospherics or conditions (Yin, 2018). In a practice called pattern matching, we take an identifiable pattern from our descriptive account and compare it to an anticipated or predicted one. Another option, explanation building, involves making an initial theoretical statement, comparing initial findings with the original proposition, revising the original proposition, and repeating this process until one has settled on a refined proposition that matches the data. Yet another practice, the logic model, fleshes out the complex development of events over a period of time. Analytically, this may mean matching observed events with theoretically predicted ones. Across all these practices, analyzing qualitative case studies incorporates many of the strategies and practices employed in generic qualitative research—coding, memo writing, and diagrams.

Case study analysis makes use of methodological triangulation which entails comparing at least two different methods, such as interviews and observations, to assess whether researcher discoveries and participant points of view evident in one methodological practice also surface in the other, for example, whether there are conceptual echoes between interviews and naturalistic encounters (Yin, 2018) (For a discussion of other types of triangulation—data triangulation, investigator triangulation, theoretical triangulation, environmental triangulation—see Denzin and Lincoln, 1998). Triangulating methods allows us to assess convergence, complementarity, or divergence of these different data types, illustrating how evidence across methods supports a claim or complicates it, perhaps further dimensionalizing a topic (Farquhar et al., 2020). Convergence and complementarity strengthen the findings and precision of a case study and address the limitations of relying only on one type of (perhaps myopic) data. Equally valuable, divergence might point to how participants claim a type of behavior or belief that the researcher sees differently through a different method, such as in the context of a real-time observational setting. Though triangulation often aligns with a post-positivist lens,

with appeals to validity and reliability, it can be adopted in a more interpretivist approach, and treated as a kaleidoscope (Flick, 2018), enabling researchers to take on different interpretations of their study at different levels of abstraction (Farquhar et al., 2020).

Memo writing is a tool to systematically compare data types within a case study as well as findings across cases in a multi-case study. Researchers might also use a matrix with different data types on one axis and topics on another axis to further illustrate how a topic (or theme) shows up similarly or differently across these data types (e.g., Rafaeli, 1989).

Ethnographies

As others have noted, many approaches to ethnography emerged in the twentieth century, from researchers studying the internal logics of a holistic culture (Agar, 1980; Malinowski, 1922/1966) to attempts to analyze more interconnected “cultural scenes” (Spradley and McCurdy, 1972). Similarly, anthropologists moved from studying people “outside” their own culture (Park, 1925/1967) to looking at settings within their own culture, such as researchers in the US studying inner city Philadelphia (Anderson, 1999). Contemporary ethnography, especially in recent decades, also introduced efforts to incorporate one’s critical reflexivity into the analytic journey, inquiring how one’s positionality activates a particular ethnographic lens in both participant and non-participant observation and how it contributes to framing the final ethnography.

In entering the field—a school, a hospital, a refugee site—ethnographers encounter multiple ongoing stories, rather than a single cultural account, and concern themselves with verbal exchanges of various types: phatic communion, ordinary conversation, skilled conversation, personal narratives, and dialog (Goodall, 2000). Phatic communion refers to politeness rituals; ordinary conversation to discussion, gossip, or information exchange; skilled conversation to debate, negotiation, or interviews; personal narratives to individual or mutual self-disclosure; and dialog to “spontaneous mutuality” (pp. 103–104).

Ethnographic interviews, as a form of skilled conversation, take place after a researcher has been in a field long enough to be able to ask relevant questions. In analyzing these collected verbal exchanges, the researcher considers the temporal, spatial, and embodied nature of the encounter—body language, vocal inflections, and setting atmospherics. As such, a verbal exchange can be considered as much of an “event” as a cultural practice in the community, blurring the distinction between participant-observation and interview.

Ethnographic efforts function along a spectrum from *emic* (folk concepts) to *etic* (concepts of the ethnographer)—guided by theories of culture and shared meaning—but it is the inductive thrust, the surfacing folk concepts, that leads to the most unexpected findings. Field notes, memos regarding one’s impressions in the cultural spaces, “combine a high level of careful empirical detail with personal asides and impressions” (Mills and Morton, 2013, p. 83). As such, they are a cumulative space inviting the unforeseen to take shape. To refine analysis, ethnographic memos can be organized into types: observational, methodological, theoretical, emotional, and environmental (Gobo and Molle, 2017). In contrast, coding, an optional practice in ethnographic analyses, condenses copious accounts of the cultural group or cultural scenes—observations, interviews, collected artifacts—into meaningful topics. Codes used in ethnographic analyses might include *beliefs* and *values* as well as other established codes such as *routines* (also called “everyday practices”), *rituals*, *rules*, *roles*, *relationships* (Saldaña and Omasta, 2018) as well as *rich points* (speech acts that contain cultural knowledge), *turning points*, *habits of speech*, *myths*, and *actual practices* (or organizational realities) (Goodall, 2000, p. 113). Coding, and analysis itself, is iterative, with initial coding casting a wide net around data, generating many topics—both deductive and inductive—and subsequent analysis engaging more targeted coding. The researcher uses growing knowledge to better assess documented encounters in the field and identify patterns across data, such as power in particular cultural spaces (deCerteau, 1996). Memos and codes together offer two forms of “coherence”: contiguity and similarity (see Maxwell, 2012). Memos can be used to examine cultural spaces as contiguous wholes while coding can be used to identify components of culture (e.g., beliefs) and to assess similarity and dissimilarity within and across them.

Though an ethnography is not replicable in the traditional sense because it would be impossible for another researcher to have the same encounters with the same people in the field and generate the same data, data analysis can attempt to address replicability by developing a codebook. A codebook is a set of identified topics with operationalized definitions. Differentiating between similar codes, such as *sharing information* versus *knowledge building*, can create deeper insights into what shared meanings inhabit the community or group under inquiry. This does not mean that values per se are shared but that the culture as a whole activates a range of values of which others are aware, such as a country with polarized political views. An ethnography might go beyond the data at hand and draw connections between the study’s emerging findings and link them to existing theories on social structures, processes, and outcomes. Furthermore, ethnography is corroborative, relying on more than one data collection method (Wolcott, 2010) (See also discussion of triangulation in **Case studies** section of this chapter.)

Some researchers (St. Pierre and Jackson, 2014) advocate a “post-coding” approach to avoid overly fragmenting data into segments at the expense of a more cohesive understanding. Rather than coding, post-coding practices might include “focusing on occasions and instances” that provide more holistic meaning and writing about those instances as part of rich description (Harrison, 2018, p. 79).

As a final product, an ethnography is an overarching representation of a cultural space (or connected spaces) based on cumulative verbal exchanges, artifacts, and observations. The findings are typically shared in a narrative report that walks the reader through primary themes and illustrative accounts of various interconnected episodes, such as sense-making episodes, risk-taking episodes, face-saving episodes, and crises (Goodall, 2000, p 18; Van Maanen, 1988). In an ethnography of inner-city Philadelphia, anthropologist Elijah Anderson identified constructs, or themes, such as *campaigning for respect*, the *social shuffle*, *self-image based on*

"juice," and the *staging area*. These constructs are an example of applying "inference to a static cultural scene" (Goodall, 2000, p. 94). Agar refers to inferences as the "glue of coherence," what link "different pieces of knowledge and connect knowledge with the world" (Agar, 1980, p. 32). In other words, the evocative language of Anderson's themes provides an interpretive epistemological frame, combining theoretical exploration with evocative description. His ethnography not only identifies abstractions, it also walks the reader through textured accounts of particular youth, such as Tyree's story, which is based on an extended ethnographic interview. Tyree's story is "dramatized in places to represent vividly the intricacies of the code of the street," the unwritten rules of inner-city life, and to provide evidence for the constructed themes (Anderson, 1999, p. 80). Anderson's textured writing bring to mind Haraway's advice that the "only way to find a larger vision is to be somewhere in particular" (1988, p. 590). Synthesizing and reconciling the particular (and its many contradictions) with the holistic is a common objective of ethnographic writing. Visualization techniques for reporting findings include time-ordered displays (understanding causes and contexts of local causality), role ordered displays (showing who does what when), conceptually ordered displays, and agent-based modeling (Gobo and Molle, 2017; Ladner, 2014).

Grounded theories

Grounded theory, developed by sociologists Bernard Glaser and Anselm Strauss (1967), introduced an innovative approach to social and behavioral research that underscored an inductive entry point into initial data engagement, followed by a strategic shift to deduction and more targeted data collection. The ultimate goal of this approach was to theorize based on data—to answer largely "why" questions regarding actions, reactions, and interactions—to attempt to explain what activates, sustains, or destabilizes a complex process. Early grounded theory is sometimes called "objectivist" because it suggested an identifiable, objective theory "out there" in the world. Since then, other forms of grounded theory have been developed, including Straussian grounded theory, which focused on established constructs related to process, including contexts, conditions, antecedents, promoters, actions, and consequences (Corbin and Strauss, 2015; Strauss and Corbin, 2008). Another genre, constructivist grounded theory, developed by sociologist Kathy Charmaz, acknowledged the researcher's role in co-creating data and knowledge production. In contrast to researchers of formal grounded theory, advocates of constructivist grounded theory recognize "multiple realities" and see knowledge as situated within a particular setting or situation in an "ever-changing world" (Charmaz, 2014, p. 132).

Though grounded theory is inductive, researchers nonetheless begin with what have been called "sensitizing concepts," topics based on a discipline or literature review (Charmaz, 2014, p. 30), rather than an "empty head" (Dey, 1999, p. 251). However, these concepts do not direct our inquiry. We are simply aware of them as we begin our intimate reading of data and let the data direct our initial code construction. This phase, called "open coding," allows us to generate numerous ideas, capturing structures and processes as well as creating *in vivo* codes, topics named based on participants' distinctive language. Open coding then moves to focused coding, whereby researchers concentrate on those open codes with the most "analytical reach" (Charmaz, 2014). In interrogating both open codes and focused codes, grounded theorists move to higher levels of abstraction—developing categories—to theorize the data and address variation—how participants might engage a process, such as recovering from a house fire, differently based on shifting contexts and timespans (Charmaz, 2002; Stern and Kerry, 2009). This process might involve returning to the field to collect additional data concentrating on an emerging concept, a process called "theoretical sampling" (Charmaz, 2014; Glaser and Strauss, 1967; Morse, 2007).

Throughout the phases of coding, researchers review quotations coded to the same code or category, a practice called "constant comparison," which, as Glaser and Strauss (1967) note, "very soon starts to generate theoretical properties of the category" (p. 106). This not only allows for theoretical refinement but also encourages researchers to stay close to the data, not to overinterpret based on an underdeveloped understanding of a code. As a mode of rigor, constant comparison engenders a higher level of abstraction (Charmaz, 2014) and can be used in generic qualitative approaches as well. It is a means of precision and a way to avoid drift—losing one's operational sense of a code over time because one has lost track of how it has been applied.

Because grounded theorists are interested in *making sense* of a process—understanding the conditional or evolving contexts in which the process occurs or is transformed—their objectives go beyond practical description to instead grapple with social structures, how processes and individuals are situated, and how processes function in motion. Researchers do not simply move forward into abstract territory; they also go backward into the data itself. This is a shift from induction to deduction as researchers "test" the focused codes on more and more data. That is, the data, including data in the ongoing data collection, become territory that researchers examine with burgeoning hunches. Researchers view the emerging categories provisionally and seek evidence grounded in the data to support, refine, or retire the abstraction. They also seek negative instances and incorporate these occurrences into more overarching constructs. This conceptual elaboration increases the theory's credibility and reach. Researchers might also make use of diagrams throughout analysis to develop conceptual clusters and relationships among codes as they emerge (Keane, 2022; Soulliere et al., 2001).

In some forms of grounded theory, determining a central category or "core concept" is common (Strauss and Corbin, 2008). This summons researchers to conceptualize what they have been hearing across participants and gradually discerning across codes and memos and through constant comparison. As Charmaz (2014) says, a central category is one that "made explicit what I had sensed and understood but had not conceptualized" (p. 146) (Corbin and Strauss, 2015). It takes the foreground as the other categories become supporting entities in the background. In modified versions of grounded theory, there are several primary categories but not necessarily a central one.

Grounded theorists conclude their efforts when additional data confirm the properties of the categories; additional data do not develop the categories (Dey, 2007). This theoretical saturation signals a kind of closure, though the theory will develop as others examine it perhaps using other methods, such as surveys, scales, or other qualitative methods.

In reporting the results of a grounded theory, researchers create products that condense their identified or constructed categories, including perhaps conditional matrices and diagrams that map categories and dimensions (Corbin and Strauss, 2015; Kennedy-Lewis, 2014). A conditional matrix shows the variance of a category on each axis. For example, a matrix addressing survivors of hurricanes might show *perceived luck* on one axis—from high to low—and *trying to live normally* on the other—again from high to low. *Perceived luck* refers to the extent to which survivors claim to be lucky regarding their survivor status while *trying to live normally* captures the degree to which survivors attempt some semblance of a daily routine. The resulting matrix and discussion would show *how* and *why* participants fall in one of the four quadrants; it is one possible product of a study theorizing the process of surviving a hurricane and the variation in the survivors' experiences, but many other kinds of diagrams explicating theory are possible.

Narrative analyses

In narrative analysis, researchers move away from analyzing situations, individuals, and events per se to instead studying the constructed stories—the narrative accounts—of these entities. By becoming a textual event, a story is “detached from the moment it occurred and has assumed consequences of its own” (Moen, 2006; Ricoeur, 1981). The roots of narrative analysis are extensive and have led to numerous lenses for analyzing social encounters as revealed through textual data (see Bertaux, 1981; Clandinin, 2013; Kim, 2016; Polkinghorne, 1988; Riessman, 1993). In this long-established tradition, researchers collect storied data from participants—episodes that individuals present in terms of chronology and plot—in order to construct a meta-story, a higher-level account that incorporates patterns across stories (e.g., see Sarasa, 2015). In seeing data as entry points into storied accounts, researchers depart from more categorical ways of thinking about data and data collection (such as open-ended surveys). However, even in highly structured interview data, which often suppress participant stories, researchers might nonetheless analyze the data with storytelling in mind. Stories provide episodic knowledge—they reveal what participants know without our asking them what they know—and typically take longer to tell than simply sharing an opinion or attitude in more “efficient” forms of data collection, such as open-ended survey responses.

Approaches to analyzing narrative data are varied and have different foci. Analysis might focus on content (what is told), structure (how a story is constructed), or on the narrative's performative aspect (who is the intended audience) (Riessman, 2007). Engaging with narrative data may involve coding its components, such as *story type*, *form*, *genre*, *tone*, and *purpose*. More refined coding structures, such as those outlined by Labov (1972), include: *abstract*, *orientation*, *complicating action*, *evaluation*, *resolution*, and *coda*. *Abstract* encapsulates the reason the story is being told, *orientation* provides details of time and setting, *complicating action* indicates an event that changed the course of action, *evaluation* refers to the narrator's attitude about the event, *resolution* or *result* is the final outcome, and *coda* brings the audience back to the present time. Saldaña (2021) outlines additional analytic practices, including dramaturgical coding, motif coding, and metaphor coding. In dramaturgical analysis, we view life as performance and individuals as characters acting out a plot with an identifiable beginning, middle, and end. Attendant codes include *objectives*, *conflicts*, *tactics*, *attitudes*, *emotions*, and *subtexts* (see also Bogdan and Biklen, 2007). Another especially elaborate option is Thompson's Motif Index. A motif refers to particular details within a tale and may refer to a type of character, action, setting, or object (Thompson, 1932/2022). These include mythological motifs, such as quest motifs, and might be used to assess how certain present-day accounts are congruent with classic literary patterns, providing further insight into how individuals naturally draw from the well of storytelling devices to make claims and describe barriers to their fulfillment.

Despite these established coding systems, Clandinin and Connelly (2000) indicate that their approach is to “find a form to represent... storied lives in storied ways, not to represent stories lives as exemplars of formal categories” (p. 141). Hence, researchers might need to discern in the data those elements that may not align with a previously established code structure. One might also argue that social media is changing how human beings tell stories and these fragmented forms of storytelling—tweets, blogs, and Facebook posts—may require a special form of coding and attention to subtext.

Narrative as chronology, a textual account imbued with plot, invites us to consider how inciting events or a character's actions lead to subsequent events, suggesting a causal link under particular circumstances (Elliott, 2005; Polkinghorne, 1995). Analysis might involve developing a timeline, or an annal (Clandinin, 2013), a logical sequence underscoring the pivotal moments, twists-and-turns, and consequences of the events. After developing a narrative account of each participant, the second level of analysis would involve the “resonant threads or patterns” discernible across transcripts (Clandinin, 2013, p. 132). In one example, “narratives were temporally sequenced, telling of a journey from place to place through time, with key moments in the plot common across the stories” (Goodman, 2004, p. 6). Developing this meta-story involves identifying themes, or resonant threads, that tie together all stories. This re-storying of the stories could also use “cohering metaphors,” such as “rising from the ashes,” or other techniques to identify and name meaningful reverberations across accounts (Ladner, 2014).

Narrative analysis strives to find a meeting ground between the nomothetic and the idiographic. That is, in re-storying, narrative researchers look for broader themes that apply to participants across the data in addition to presenting evocative examples from distinctive personal experience. “In these individual stories, the unique way each participant interpreted his experiences came into view” (Goodman, 2004, p. 6). Quotations from individuals can foreground the unique encounters and language of individuals while the re-storying elevates the individual voices to narrative themes (e.g., Poveda et al., 2004). In Ahlsen et al.'s 2014 article on

men's and women's stories of chronic muscle pain, the researchers focus on a narrative account from one man and one woman, then demonstrate how each of these illustrate the gendered difference of chronic pain evident in the rest of the participants' stories, how women's stories "displayed selves that were actively trying to transcend their former identities and life conditions" while men's stories "tended to display selves that were actively working to find a solution to the problem within the context of medicine" (p. 6). Researchers might also use diagrams to illustrate individual chronologies and a composite diagram to illustrate a common trajectory and perhaps identifiable variations. See process mapping in [Miles et al. \(2020\)](#) and mapping individual narrative plots using space, time, mind in [Ryan \(2007\)](#).

Phenomenologies

Phenomenology is a philosophy, a perspective, and a qualitative research tradition ([Farrell, 2020](#)). As a research tradition, it is aimed at "opening up" our understanding of complex experiential accounts, moving beyond a simple description of participants' opinions, attitudes, and feelings "after the fact." A phenomenological approach is instead aimed at the "living now" ([Van Manen, 2006](#)), at how the "thing-in-itself-as-it-shows-itself in consciousness" (p. 63). In looking "through" what we normally look "at" ([Vagle, 2018](#)), researchers engaged in phenomenology attempt to see through assumptions that have shaped our understanding of the world ([Van Manen, 2006](#)). Phenomenology is engaged in pre-reflective experience, in how we live through encounters *in the moment* such as attempting to speak a new world language or how to program in a technical, computer language. Hence, phenomenological interviewing attempts to invoke the "living moment of immediate experience" ([Van Manen, 2006](#), p. 59)—something that highly structured interviews are not designed to do—rather than post-reflective theorizing, conceptualizing, and polemics. Phenomenology is guided by the concrete, direct contact with the world, rather than an objectification of it and is focused on the phenomenon more than on the psychological self of the individual. Various genres of phenomenology have been developed, including transcendental phenomenology ([Husserl, 1970](#)), personalistic phenomenology ([Scheler, 1970](#)), and ontological phenomenology ([Heidegger, 1962](#)).

Even though experiential accounts of a lived experience are not identical with the "living now" itself, we use open-ended interview data as an entry point into the experience and as content that we can use for analytic reflection. These accounts can come in the form of phenomenological interviews—conversations designed to invite participants to walk us through an experience—or in written accounts provided by participants. In either case, we are interested in the experience from the "inside," how individuals inhabited space, how things sounded, what stood out as vivid in the moment.

In analysis, researchers take on both description and interpretation, uncovering evocative moment-by-moment experience. However, creating standard procedures for phenomenological analysis is antithetical to the approach given by phenomenology. Each study has its own approaches to uncovering the phenomenon, to making its "essence" apparent. Nonetheless, this chapter will attempt to provide guidance on how a researcher might proceed to describe and interpret a phenomenon. The analytic labor of phenomenology involves reducing the experiences of the phenomenon to its identifiable structures across individuals. [Van Manen \(2006\)](#) offers different possibilities of reductive reflection: thematic reflection, linguistic reflection, etymological reflection, conceptual reflection, exegetical reflection, and hermeneutic interview reflection (p. 312). This list gives us options in our attempts to condense textual accounts into meaningful parallels across participants, opening up the phenomenon thematically and linguistically.

Rather than coding, developing taxonomies, or assessing the frequencies of topics, phenomenologists might instead first mark transcripts to identify "meaning units." These are evocative paragraphs that bring us into the living now. We might assign titles to these units to not only help organize them but to begin the process of condensing the data to its essential meanings. The title might be part of the participant's own language as a way of elevating particular evocative language. Next, a researcher might write a reflective memo on the meaning unit capturing the implications of the paragraph, then title the memo. [Van Manen \(2006\)](#) outlines an intermediate step whereby the original paragraph is turned into an "anecdote," (p. 29) with much of the same language as the original but with unnecessary detail removed. The anecdote is intended to focus on the details of the phenomenon. A reflective memo could then be written based on the anecdote, with an eye to reducing the datum to its phenomenological essential meanings. If writing this memo proves challenging, the researcher might, as Van Manen suggests, write a reflective sentence of each sentence of the original paragraph or anecdote to identify moment-by-moment significance. This invites the researcher to slow down and think more intentionally about the experience as it is unfolding. As we move through data and write more memos, we might also notice how participants engage time, physical settings, and other elemental structures of the phenomenon. More specifically, we might assess meaning at the level of the transcript, paragraph, or sentence, writing memos to keep track of our reductive reflection.

The product of a phenomenology is not theory but highly relevant description. Phenomenologists gather together perceptions and sensitivities across participants' open-ended interviews and synthesize these data into a structural description of the phenomenon. This description is less idiographic than it is summative of the common structures that hold together the phenomenon as a distinctive unfolding of moments. Hence, we might outline a phenomenological structure, organized by time or other common features. For example, a phenomenology of living through a hurricane might locate common moments, embedded in time, such as incremental awareness of danger and survivors' attempts to validate their fears by communicating with a neighbor or stranger (For other examples, see [Callister and Cox, 2006](#); [Eddles-Hirsch, 2015](#); [Mihás, 2019](#)). A product of a phenomenology might also include central features—a list of statements about the phenomenon that allows us inside of it

(Wertz, 2011). Some researchers might even use found poems or other artistic forms to capture the living now of the phenomenon as it is revealed across data.

Conclusion

Deciding on a qualitative research tradition and analytic approach depends on the research questions and the intended scope and analytical depth of the inquiry as well as on the receiving audience for the research. Research traditions can also be blended in innovative ways that fuse two or more approaches, such as using narrative inquiry to build a grounded theory (see Williams and Keady, 2012). Claiming a particular research tradition should come from attention to one's ontological and epistemological assumptions—the amniotic fluid of the inquiry—as well as from the research question itself. Ontological and epistemological considerations point to what will be considered “evidence” for a given study and whether textual data represents reality or generates its own narrative reality.

If the researcher is viewed as a reliable agent of logic (Vagle, 2021) who, with sufficient evidence, can discern an event or situation, then collecting qualitative data can sufficiently represent the reality they seek to uncover. But if one sees reality as a negotiated space co-constructed continuously, then this perspective alters one's stance in relation to the generated data and findings. This is why constructivist grounded theorists make no claims to stand outside the theory; they are aware that the interviews (or other data forms) are co-constructions, not a precise mirror of an objective reality. Narrative researchers, too, are aware that interviewees are telling a story *to the interviewer* in a manner than may differ from how they perform the story to others. In phenomenology, epistemological concerns also lead researchers to bracket their own experiences of the phenomenon and to un-learn what they thought they knew (Vagle, 2021). Epistemological paradigms also determine, in part, the language of the analysis—a constructivist researcher might talk about “constructing” themes whereas a pragmatic researcher might discuss “identifying” them. More importantly, knowledge might be discerned as derived from direct observation (discovered in a single reality “out there”), co-constructed with participants, or derived from a deep analytic dive into the intentionality of subject-object. Hence, one's reporting of what one knows, has constructed, or has found will be shaped by one's epistemological positioning. Researchers should consider these assumptions when embarking on analysis to not only understand how to share findings but to communicate what constitute and what problematize them.

References

- Agar, M.H., 1980. *Speaking of Ethnography*. SAGE.
- Ahlsten, B., Bondevik, H., Mengshoel, A.M., Solbraekke, K.N., 2014. Un)doing gender in a rehabilitation context: a narrative analysis of gender and self in stories of chronic muscle pain. *Disabil. Rehabil.* 36 (5), 359–366.
- Alexander, I., 1988. Personality, psychological assessment, and psychobiography. *J. Pers.* 56 (1), 265–294.
- Anderson, E., 1999. *Code of the Street*. W.W. Norton.
- Bartlett, L., Vavrus, F., 2017. *Rethinking Case Study Research*. Routledge.
- Bertaux, 1981. *Biography and Society: The Life History Approach in the Social Sciences*. SAGE.
- Bingham, A.J., Witkowski, P., 2022. Deductive and inductive approaches to qualitative data analysis. In: Vanover, C., Mihas, P., Saldaña, J. (Eds.), *Analyzing and Interpreting Qualitative Research: After the Interview*. SAGE, pp. 133–148.
- Bogdan, R.C., Biklen, S.K., 2007. *Qualitative Research for Education: An Introduction to Theories and Methods*, fifth ed. Pearson Education.
- Boyatzis, R.E., 1998. *Transforming Qualitative Information: Thematic Analysis and Code Development*. SAGE.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qual. Res. Psychol.* 3 (2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- Callister, L.C., Cox, A.H., 2006. Opening our hearts and minds: the meaning of international clinical nursing electives in the personal and professional lives of nurses. *Nurs. Health Sci.* 8 (2), 95–102. <https://doi.org/10.1111/j.1442-2018.2006.00259.x>.
- Charmaz, K., 2002. Stories and silences: disclosures and self in chronic illness. *Qual. Inq.* 8 (3), 302–328.
- Charmaz, K., 2014. *Constructing Grounded Theory*, second ed. SAGE.
- Clandinin, D.J., Connelly, F.M., 2000. *Narrative Inquiry: Experience and Story in Qualitative Research*. Jossey-Bass.
- Clandinin, D.J., 2013. *Engaging in Narrative Inquiry*. Left Coast Press.
- Collins, J., 2008. *For Love of a Soldier: Interviews With Military Families Taking Action Against the Iraq War*. Lexington Books.
- Corbin, J., Strauss, A., 2015. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE.
- Daiute, C., 2014. *Narrative Inquiry: A Dynamic Approach*. SAGE.
- Damschroder, L.J., Aron, D.C., Keith, R.E., Kirsh, S.R., Alexander, J.A., Lowery, J.C., 2009. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement. Sci.* 4 (50), 50. <https://doi.org/10.1186/1748-5908-4-50>.
- deCerteau, M., 1996. *The Practice of Everyday Life, Volume II: Living and Cooking*. University of California Press.
- Denzin, N., Lincoln, Y., 1998. *The Landscape of Qualitative Research*. SAGE.
- Dey, I., 1999. *Grounding Grounded Theory*. Academic Press.
- Dey, I., 2007. Grounding categories. In: Bryant, A., Charmaz, K. (Eds.), *The SAGE Handbook of Grounded Theory*. SAGE, pp. 167–190.
- Eddles-Hirsch, K., 2015. Phenomenology and educational research. *Int. J. Adv. Res.* 3 (8), 251–260.
- Elliot, J., 2005. *Using Narrative in Social Research: Qualitative and Quantitative Approaches*. SAGE.
- Farquhar, J., Michels, N., Robson, J., 2020. Triangulation in industrial qualitative case study research: widening the scope. *Ind. Market. Manag.* 87, 160–170.
- Farrell, E., 2020. Researching lived experiences in education: misunderstood or missed opportunity? *Int. J. Qual. Methods* 19, 1–8.
- Flick, U., 2018. *Doing Triangulation and Mixed Methods*. SAGE.
- Frith, H., Gleeson, K., 2004. Clothing and embodiment: men managing body image and appearance. *Psychol. Men Masc.* 5 (1), 40–48.

- Gerring, J., 2007. *Case Study Research: Principles and Practices*. Cambridge University Press.
- Glaser, B., Strauss, A., 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine.
- Gobo, G., Molle, A., 2017. *Doing Ethnography*, second ed. SAGE.
- Goodall, H.L., 2000. *Writing the New Ethnography*. AltaMira Press.
- Goodman, J.H., 2004. Coping with trauma and hardship among unaccompanied refugee youths from Sudan. *Qual. Health Res.* 14 (9), 1177–1196.
- Haraway, D., 1988. Situated knowledges: the science question in feminism and the privilege of partial perspectives. *Fem. Stud.* 14 (3), 575–599.
- Harrison, A.K., 2018. *Ethnography: Understanding Qualitative Research*. Oxford.
- Heidegger, M., 1962. *Being and Time*. Harper & Row.
- Holton, G., 1973. *Thematic Origins of Scientific Thought*. Harvard University Press.
- Hsieh, H.-F., Shannon, S.E., 2005. Three approaches to qualitative content analysis. *Qual. Health Res.* 15 (9), 1277–1288.
- Husserl, E., 1970. *The Idea of Phenomenology*. Martinus Nijhoff.
- Keane, E., 2022. Critical analytic memoing. In: Vanover, C., Mihas, P., Saldana, J. (Eds.), *Analyzing and Interpreting Qualitative Research: After the Interview*. SAGE, pp. 259–274.
- Kennedy-Lewis, B., 2014. Using Diagrams to Make Meaning in Grounded Theory Data Collection and Analysis. *SAGE Research Methods Datasets*. <http://methods.sagepub.com/Datasets>.
- Kim, J.-H., 2016. *Understanding Narrative Inquiry: The Crafting and Analysis of Stories as Research*. SAGE.
- Labov, W., 1972. The transformation of experience in narrative syntax. In: Labov, W. (Ed.), *Language in the Inner City: Studies in the Black Vernacular*. University of Pennsylvania Press.
- Ladner, S., 2014. *Practical Ethnography: A Guide to Doing Ethnography in the Private Sector*. Left Coast Press.
- Maietta, R., Hamilton, A., Swartout, K., Petruzzelli, J., 2018. *Qualitative Data Analysis Camp*. ResearchTalk, Carrboro, NC.
- Maietta, R., Mihas, P., Swartout, K., Petruzzelli, J., Hamilton, A., 2021. Sort and sift, think and shift: let the data be your guide: an applied approach to working with, learning from, and privileging qualitative data. *Qual. Rep.* 26 (6), 2045–2060. <https://doi.org/10.46743/2160-3715/2021.5013>.
- Malinowski, B., 1922/1966. *Argonauts of the Western Pacific*. Routledge & Kegan Paul.
- Mars, S.G., Bourgois, P., Karandinos, G., Montero, F., Ciccarone, D., 2014. "Every 'never' I ever said came true": transitions from opioid pills to heroin injecting. *Int. J. Drug Pol.* 25 (2), 257–266. <https://doi.org/10.1016/j.drugpo.2013.10.004>.
- Maxwell, J., 2012. *A Realist Approach to Qualitative Research*. SAGE.
- Maxwell, J., 2013. *Qualitative Research Design: An Interactive Approach*, third ed. SAGE.
- Mayring, P., 2015. Qualitative content analysis: theoretical background and procedures. In: Bikner-Ahsbals, A., Knipping, C., Presmeg, N. (Eds.), *Approaches to Qualitative Research in Mathematics Education: Advances in Mathematics Education*. Springer, Dordrecht, pp. 365–380.
- Mihas, P., 2019. Learn to Conduct a Phenomenological Analysis With Interview Data From the Southern Oral History Program. *SAGE Research Methods Datasets*. <http://methods.sagepub.com/Datasets>.
- Mihas, P., 2022. Memo writing strategies: analyzing the parts and the whole. In: Vanover, C., Mihas, P., Saldaña, J. (Eds.), *Analyzing and Interpreting Qualitative Research*. SAGE, pp. 243–258.
- Miles, M.B., Huberman, A.M., Saldaña, J., 2020. *Qualitative Data Analysis: A Methods Sourcebook*, fourth ed. SAGE.
- Mills, D., Morton, M., 2013. *Ethnography in Education*. BERA.
- Moen, T., 2006. Reflections on the narrative research approach. *Int. J. Qual. Methods* 5 (4), 56–69. <https://doi.org/10.1177/160940690600500405>.
- Morse, J., Cheek, J., 2021. Developing qualitatively driven mixed methods designs. In: *Short Course at International Congress of Qualitative Inquiry*. May 20.
- Morse, J.M., 2007. Sampling in grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The SAGE Handbook of Grounded Theory*. SAGE, pp. 229–244.
- Park, R.E., 1925/1967. The city: suggestions for the investigation of human behavior in the urban environment. In: Park, R.E., Burgess, E.W., McKenzie, R.D. (Eds.), *The City*. University of Chicago Press, pp. 1–46.
- Polkinghorne, D.E., 1988. *Narrative Knowing and the Human Sciences*. SUNY Press.
- Polkinghorne, D.E., 1995. Narrative configuration in qualitative analysis. *Int. J. Qual. Stud. Educ.* 8 (1), 5–23. <https://doi.org/10.1080/0951839950080103>.
- Poveda, D., Palomares-Valera, M., Cano, A., 2004. Putting school in its place: a narrative analysis of the educational memories of late adult and elder people. *Forum Qual. Soc. Res.* 5 (1).
- Rafaeli, A., 1989. When cashiers meet customers: an analysis of the role of supermarket cashiers. *Acad. Manag. J.* 32 (2), 245–273. <https://doi.org/10.2307/256362>.
- Richardson, L., 2000. Writing: a method of inquiry. In: Denzi, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, second ed. SAGE, pp. 923–948.
- Ricoeur, P., 1981. *Hermeneutics and the Human Sciences*. Cambridge University Press.
- Riessman, C.K., 1993. Narrative analysis. In: *Qualitative Research Methods Series*, vol. 30. SAGE.
- Riessman, C.K., 2007. *Narrative Methods for the Human Sciences*. SAGE.
- Ryan, M., 2007. Diagramming narrative. *Semiotica* 165, 11–40. <https://doi.org/10.1515/SEM.2007.030>.
- Saldaña, J., Omasta, M., 2018. *Qualitative Research: Analyzing Life*. SAGE.
- Saldaña, J., 2021. *The Coding Manual for Qualitative Researchers*, fourth ed. SAGE.
- Sarasa, M.C., 2015. Narrative research into the possibilities of classroom-generated stories in English teacher education. *Profile* 17 (1), 13–24. <https://doi.org/10.15446/profile.v17n1.43383>.
- Savin-Baden, M., Howell Major, C., 2013. *Qualitative Research: The Essential Guide to Theory and Practice*. Routledge.
- Schapp, K., Sanders, M.L., 2019. What is a theme? Teaching thematic analysis in qualitative communication research methods. *Commun. Teach.* 33 (2), 117–121. <https://doi.org/10.1080/17404622.2018.1536794>.
- Scheler, M., 1970. *The Nature of Sympathy*. Archon Books.
- Soulliere, D., Britt, D.W., Maines, D.R., 2001. Conceptual modeling as a toolbox for grounded theorists. *Socio. Q.* 42 (2), 253–269. <https://doi.org/10.1111/j.1533-8525.2001.tb00033.x>.
- Spradley, J.P., McCurdy, D., 1972. *The Cultural Experiences: Ethnography in Complex Society*. Science Research Associates.
- St Pierre, E.A., Jackson, A.Y., 2014. Qualitative data analysis after coding. *Qual. Inq.* 20, 715–719.
- Stake, R.E., 1995. *The Art of Case Study Research*. SAGE.
- Stake, R.E., 2006. *Multiple Case Study Analysis*. SAGE.
- Stern, P.N., Kerry, J., 2009. Restructuring life after home loss by fire. In: *Developing Grounded Theory: The Second Generation*. Routledge.
- Strauss, A., Corbin, J., 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE.
- Sutton, R.I., 1987. The process of organizational death: disbanding and reconnecting. *Adm. Sci. Q.* 32 (4), 542–569.
- Thompson, S., 1932/2022. *Tales Online*. https://sites.ualberta.ca/~urban/Projects/English/Content/Motif_Help.htm.
- Vagle, M.D., 2018. *Crafting Phenomenological Research*, second ed. Left Coast Press.
- Vagle, M.D., 2021. Crafting phenomenological research: how phenomena take shape in various contexts. In: *Short Course at the Qualitative Research Summer Intensive*. July 29–30.
- Van Maanen, J., 1988. *Tales of the Field: On Writing Ethnography*. University of Chicago Press.

- Van Manen, M., 2006. *Researching Lived Experiences: Human Science for an Action Sensitive Pedagogy*. Althouse.
- Wertz, F.J., 2011. A phenomenological psychological approach to trauma and resilience. In: Wertz, F.J., Charmaz, K., McCullen, L.M., Josselson, R., Anderson, R., McSpadden, E. (Eds.), *Five Ways of Doing Qualitative Analysis*. Guilford Press.
- Williams, S., Keady, J., 2012. Centre-stage diagrams: a new method to develop constructivist grounded theory—late stage Parkinson's disease as a case exemplar. *Qual. Res.* 12 (2), 218–238. <https://doi.org/10.1177/1468794111422034>.
- Wittink, M.N., Barg, F., Gallo, J., 2006. Unwritten rules of talking to doctors about depression: integrating qualitative and quantitative methods. *Ann. Fam. Med.* 4 (4), 302–309. <https://doi.org/10.1370/afm.558>.
- Wolcott, H.F., 2010. *Ethnography Lessons*. Left Coast Press.
- Yin, R., 2018. *Case Study Research and Applications: Design and Methods*. SAGE.

Qualitative content analysis

Philipp A.E. Mayring, Alps-Adria University, Association for Supporting Qualitative Research ASQ, Klagenfurt, Austria

© 2023 Elsevier Ltd. All rights reserved.

What is qualitative content analysis?	314
The relevance of qualitative content analysis for educational research	315
A short history of qualitative content analysis	315
Methodological backgrounds	316
Basic techniques	316
Inductive category formation	317
Deductive category assignment	318
Software support	319
Different versions of qualitative content analysis	320
Final remarks	321
References	321

What is qualitative content analysis?

The chapter gives an overview on theoretical background and specific techniques of qualitative content analysis. Its role in qualitative research is described, central elements of definition introduced, and basic characteristics discussed. The importance for educational research is outlined, and the historical and methodological background described. We present two basic procedures, inductive category formation and deductive category assignment, by step models and by an example from educational research. Differences to other approaches to qualitative content analysis are worked out. An open access software program especially for our techniques of qualitative content analysis is presented.

First, I want to give a definition and characterization of the approach. Qualitative content analysis (cf. Mayring, 2015b, 2020, 2021; Kuckartz, 2014) is a methodological approach to text analysis mainly in social sciences, which can be used in empirical research projects working with textual material like interview or group discussion transcripts, open observational protocols (field notes) or document analyses (files or webpages for example). In a systematic bibliometric analysis of the literature and research databases FORIS, SOLIS, Sociofile, Psynindex and MLA, Titscher et al. (2000) found open content analysis as the most frequent text analytical procedure with 39% from 1621 citations. Qualitative content analysis was followed by conversation analysis (21%), standardized content analysis (19%), grounded theory (12%), objective hermeneutics (5%), and ethnography (2%). Carrera-Fernandez et al. (2014) found that particularly in psychology (based on the Web of Science database) the qualitative content analysis was by far the first choice of analysis procedure, followed by grounded theory, discourse analysis and action research in their qualitative orientated journal article analysis.

What are the central characteristics of qualitative content analysis? I want to outline four points:

- The qualitative content analysis approach (the same in quantitative content analysis) is *category-based*, that is its distinguishing feature. Categories refer to aspects within the text, which put the meaning of those aspects in a short textual form. Text analysis is, therefore, restricted to the category system. Text contents that are not addressed by the categories or holistic impressions are not considered or would have to be addressed with a further content analytical conduct or other text analysis methods. In so far it is a selective procedure in contrast to a holistic text understanding.
- The qualitative content analysis procedure is *research question oriented*. Text analytical questions (possibly several) are derived from the main aims of the research project. These questions should be answered at the end of the analysis. This clearly distinguishes the qualitative content analysis from other more open, explorative approaches such as grounded theory.
- Qualitative content analysis is characterized by strict *rule orientation and systematic ongoing*. Process models enable the procedure to be described step-by-step, and this has proven itself in previous research processes. The specific rules for each technique are reviewed in a pilot study and should not be changed after that.
- Qualitative content analysis is an approach of *mixed methods*, standing between pure qualitative and pure quantitative ongoing. The first step to assign categories to text passages or to develop categories from text passages is a qualitative-interpretive operation, the comparing, accumulating text passages with similar categories, analyzing frequencies of categories means a quantitative ongoing (the second quantitative step is not obligatory in qualitative content analysis, but mostly obvious).
- A central quality criterium for qualitative content analysis consists in high *coder agreements*, which means that several procedures of categorization of the same material should lead to similar results. Intra-coder agreement, similar to reliability (retest reliability), means that the same content analyst restarts with coding after finishing all material. Inter-coder agreement means passing the material (or at least a selection) to a second person, a measure of objectivity of the analysis. We would not expect an identical result (e.g., in the wordings of inductive categories), but a similar solution.

We can summarize: Qualitative content analysis is a systematic category-based set of techniques for analyzing texts with strict theory-based rules, containing qualitative and quantitative steps of analysis.

The relevance of qualitative content analysis for educational research

Qualitative content analysis can contribute to substantial educational research for several reasons. First, we have to consider that texts represent an important part of empirical material, coming from the data collection process. A pure quantitative research strategy restricts itself to data collection methods like test, measurement, closed questionnaire and closed observation because they produce numerical data which can be analyzed statistically. But educational processes often can only be understood if participants are putting their inner condition, their emotions, cognitions, motivations into speech, communications with the researchers in an open-ended form. And all that leads to texts which must be analyzed.

Texts occur in educational research in different forms:

- Open-ended interviews represent the gold standard in qualitative research. A more or less elaborated interview guide contains the themes or topics, deduced from the research question(s), the interviewer introduces them in his or her own language into the interview and the interviewee can react to them freely and is not reduced to choosing from predefined answers. There are several forms of open-ended interviews, like focused interview or narrative interview. Valuable empirical material in educational research can come from interviewing teachers or parents about their educational aims or letting them report successful or problematic educational situations, it can come from asking students about their learning motivations or emotions at school.
- Group discussions or focus groups bring together several participants and present them a list of discussion topics. A moderator is leading the discussions, trying to foster the participation of all participants and to thematize contradictory statements. Group discussions are very effective in educational evaluation research.
- Open-ended questionnaires give the participants the possibility to reflect their answers unhasty. They are recommended for research questions which afford elaborated appraisals or detailed statements.
- Student learning diaries, instructed by guidelines or in a free form, are a rich source of insight in educational processes.
- Open observational studies or field studies imply a longer participation in an educational field (school, family, community). The researcher records his or her observations in a protocol or in field notes.
- Document analysis works with textual material from files, documentation systems, web pages or other forms of relevant fixed texts which the researcher has found (unobtrusive measurement). In schools exercise books or essays written by pupils can be an example.

In all those cases textual material is generated by data collection (in interview studies usually as transcripts). Qualitative content analysis represents one form of systematic analysis of those empirical materials.

A short history of qualitative content analysis

Systematic techniques of content analysis had been developed in the first half of 20th century in communication science to analyze the huge amounts of textual material coming from modern mass media like newspapers, journals, radio transcripts (Krippendorff, 2019). Therefore they were conceptualized quantitatively: "Content analysis is a research technique for the objective, systematic and quantitative description of the manifest content of communication" (Berelson, 1952, p. 18). Because of the focus on countable aspects of texts the analysis was limited to the manifest content and neglected latent meanings. A series of systematic, rule-guided and intersubjective verifiable procedures were developed (Gerbner et al., 1969):

- Simple frequency analysis defines certain text elements and counts them. Formal elements, headings or topics could be as well in the focus.
- Complex frequency analysis determines variables and defines indicators for them in the text material. Examples are variables like speech disturbance in psychotherapy protocols or ideological style in newspaper commentaries. The indicators are counted and summarized to variable values.
- Contingency analysis is testing for common occurrences of categories within defined text segments (sentence, paragraph, newspaper article). This could be in form of frequent word clouds or complex structures of associations (central term together with positive or negative evaluative terms).

In any case the categories must be defined in advance and the occurrences of categories in the text are registered automatically, usually by computer programs. However, quantitative content analysis leads to several problems, for example:

- Multiplicity of meaning (e.g., "madly" in the colloquial meaning, say, of "very"; or "madly" as pertaining to psychological disturbance)
- The nuances and connotations conferred on terms by the context
- Contextual modification of meaning (for instance in the case of "no anxiety," "little anxiety" and "a lot of anxiety," "anxiety" will be counted once in each case)

- The contextual relationship of the term counted (e.g., with “I am afraid of X” or “X is afraid of me,” “afraid” is counted once in each case)
- The problem of pro-forms (e.g., with “I didn’t notice any of that” the computer does not know what “of that” refers to)
- Dialect expressions (which occur in interview scripts regularly) need a great deal of re-working

Those problems led to the development of procedures of qualitative content analysis. The central idea is to maintain the logic of systematic text analysis, adapted as well for large text corpora, but to define the step of approaching categories to text passages as rule-guided interpretive procedure.

Methodological backgrounds

The content-analytical rules for the individual techniques are not arbitrary but have a solid theoretical foundation in the processes of everyday text analysis. In particular, regarding how these processes have been examined in cognitive psychology and psycholinguistics. For summary and inductive category formation, these are reductive operators (omission, generalization, construction, integration, selection, and bundling; Mandl, 1981), on which the gradual reduction of text segments is based. For explications, it is rather context theories from linguistics, whereas for deductive category applications, reference is made to the categorization theories from general psychology and language development research (Mayring, 2022). The result of this was that the exact wording for a human-readable general category requires an explicit definition (definition theory), a cognitive anchoring in typical examples for the category (prototype theory), and rules to demarcate the categories from one another (decision bound theory, Murphy, 2002). These coding guidelines are the basis for the three-part coding—definitions, anchor examples, and coding rules—applied in the procedure that I recommended. Thus, when trying to determine content-analytic rules, I try to use strategies that draw on the everyday handling of texts, a method that is common in qualitative research, for example, when linguistic approaches to storytelling in everyday life are employed in the rules for narrative interviews.

Basic techniques

The techniques of qualitative content analysis which we present here had been developed in 1980 (cf. Mayring, 2000, 2022) and refined since then into three groups of procedures:

Reducing procedures: The object of the analysis is to reduce the material such that the essential contents remain, in order to create through abstraction a comprehensive overview of the base material which is nevertheless still an image of it.

Explicating procedures: The object of the analysis is to provide additional material on individual doubtful text components (terms, sentences, ...) with a view to increasing understanding, explaining, interpreting the specific passage of text.

Structuring procedures: The object of the analysis is to filter out particular aspects of the material, to give a cross-section through the material according to pre-determined ordering criteria, or to assess the material according to certain criteria. In those procedures the categories are formulated in advance in the sense of a deductive category assignment.

These basic forms, however, must be further differentiated before an exact description of procedures is possible. In addition to the usual summaries, the same ongoing process is useful for inductive category formation; a criterion for the categories is defined and aspects of this criterion are stepwise gathered in the material. Forms of explication are possible which use the textual context for the elucidation of a specific text passage (narrow context analysis); however, the most common method of hermeneutical interpretation is to use further material beyond the textual context for explication (broad context analysis). With structuring, too, subgroups must be distinguished. The categories which are brought deductively to the material can consist of a list of aspects (nominal scale). Or the categories form an ordinal scale (e.g., more—less; positive—neutral—negative) and serve as a rating procedure for the text. In addition, some mixed procedures have been described (Mayring, 2021, 2022). One such is that in content structuring or theme analysis the material is deductively ordered to categories and within each category material an inductive process of category formation is performed. Type analysis is a similar procedure where categories in the first step have to meet a typologizing criterion (typical types, extreme types, frequent types, theoretical types). In category refinement a deductive category system is modified and supplemented with new categories in an inductive way. Parallel forms execute several procedures in one passage through the material.

Through this differentiation we arrive at ten distinct forms of analysis (cf. Mayring, 2020, 2021, 2022):

Reductive:

- (1) Summary: The whole text is reduced step by step to its core content.
- (2) Inductive category formation: Categories are formulated from the text, following a selection criterion and an abstraction level.

Explicating:

- (3) Narrow context analysis: An unclear text passage is explained by systematically collecting related content within the whole text.
- (4) Broad context analysis: An unclear text passage is explained by systematically collecting related content within the text and from additional material.

Structuring/Deductive:

- (5) Nominal categories: Predefined categories from an unordered list are assigned to the text.
- (6) Ordinal categories: Predefined categories from an ordinaly ordered list are assigned to the text.

Mixed:

- (7) Content structuring: Central themes are extracted from the text, and the referring text passages are summarized.
- (8) Type analysis: Typical aspects of the text are extracted, and the referring text passages are summarized.
- (9) Category refinement: A deductive category assignment is extended by adding new categories.
- (10) Parallel forms: Several inductive or deductive techniques are applied simultaneously to the text.

These procedures have been described extensively elsewhere (Mayring, 2021, 2022). We now demonstrate two central techniques of Qualitative Content Analysis: inductive category formation and deductive category assignment (structuring).

Inductive category formation

On the basis of summarizing qualitative content analysis, a technique for inductive category formation can be developed. We have underlined that category definition is a central step in content analysis. There are two possible procedures: deductive category definition tries to develop categories out of theoretical considerations, with theories or theoretical concepts used in a process of operationalization in direction of the material; inductive category formation develops categories directly out of the material. For qualitative content analysis the second is very fruitful. The ongoing inductive process has great importance within qualitative research. It aims at a true description without bias due to the preconceptions of the researcher, an understanding of the material in terms of the material. Inductive category formation is a central process within the approach of Grounded Theory (Strauss, 1987; Strauss and Corbin, 1990), which they call “open coding.” They developed many rules of thumb for open coding, and they recommended a systematic, line-by-line procedure. For content analysis, nevertheless, inductive category formation has to be more systematic. And it can use the same logic, the same reductive procedures, as in summarizing content analysis. The following process model (Fig. 1) will now be explained.

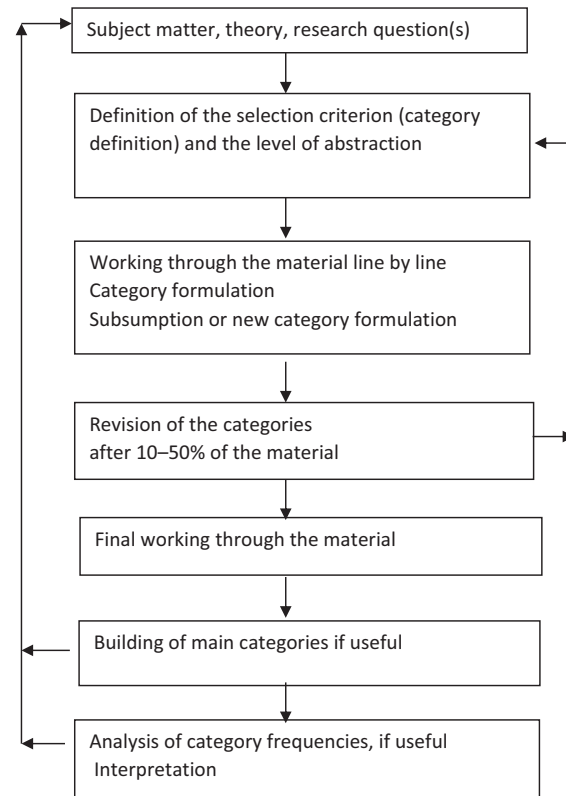


Fig. 1 Process model of inductive category formation.

Within the logic of content analysis, the level or theme of categories to be developed must be defined previously. There has to be a criterion for the selection process in category formation. This is a deductive element and is established within theoretical considerations about the subject matter and the aims of analysis. The second basic content analytical rule for inductive category formation is the establishment of the abstraction level. This comes from summarizing content analysis which reduces the material from one abstraction level to the next. If this level is not defined the categories remain chaotic, out of order. After those two rules are determined, the material is worked through line by line. The first time that material is found fitting the category definition, a category must be constructed. A term or short sentence which stands as near as possible to the material serves as category label. The next time a passage fitting the category definition is found it has to be checked whether it falls under the previous category, in which case it can be subsumed under this category (a reductive process); if not, a new category has to be formulated.

After working through a good deal of material (c. 10–50%) no new categories are to be found. This is the moment for a revision of the whole category system. It has to be checked whether the logic of categories is clear (e.g., no overlaps) and whether the level of abstraction is adequate to the subject matter and aims of analysis. Perhaps the category definition should be changed. If there are any changes in the category system, of course the complete material has to be worked through once again. After this analysis we have a set of categories to a specific topic, connected with specific passages in the material. The further analysis can go different ways: the whole system of categories can be interpreted in terms of the aims of the analysis and used theories; or the links between categories and passages in the material can be analyzed quantitatively (e.g., we can look at those categories occurring most frequently in the material).

For example, in a study on learning emotions (Glaeser-Zikuda and Mayring, 2003) we analyzed open-ended interviews and daily diaries on concrete learning experiences of 24 students of eighth grade. With inductive category formation we built up categories concerning positive learning experiences. We generalized those categories on a medium level of abstraction. Here are the most frequent categories:

C1: Happy about the interesting learning activities today	(21 occurrences)
C2: Happy to master the subject and having understood everything	(21 occurrences)
C3: Amazing subjects in the lesson (literature, poems)	(16 occurrences)
C4: Enjoyed the positive feedback by the teacher	(14 occurrences)
C5: Nice group work or partner work	(11 occurrences)
C6: Interesting problems (electricity) in the lesson	(3 occurrences)

Those inductive categories give an impression about positive emotions in learning processes. In a second step we found two main categories within this list: positive emotions about the learning processes (C1, C2, C4, C5) and positive emotions about the learning content (C3, C6). We then compared the occurrences of those main categories between the two groups of high and low achievers and found a correlation between positive emotions about learning processes and high classroom achievement. That means that it seems to be more important for good teaching to associate positive emotionality with successful learning processes than with learning content.

Deductive category assignment

This is the content-analytical method which is probably most often used. It has the goal of extracting a certain structure from the material. This structure is brought to bear on the material in the form of a category system. All text components addressed by the categories are then extracted from the material systematically. If one wishes to describe the structuring procedure quite generally, a few points are especially important: the fundamental structuring dimensions must be exactly determined; they must derive from the research question and must be theoretically based; these structuring dimensions can be further subdivided, split up into individual features or values; the dimensions and values are then brought together to form a category system.

The particular categorization of a given material component is something that must be determined precisely. A procedure for this, based on everyday life processes of categorization, has proven useful (cf. Ulich et al., 1985). Within developmental psychology (learning of categories in speech development) and within general psychology (categorization theories, cf. Murphy, 2002) it has been shown that categories are imagined in form of explicit definitions, prototypes and demarcation rules. So, a category can be defined best if all three determining approaches are used:

- **Definition of the categories** It is precisely determined which text components belong in a given category.
- **Anchor samples** Concrete passages belonging in particular categories are cited as typical examples to illustrate the character of those categories.
- **Coding rules** Where there are problems of delineation between categories, rules are formulated for the purpose of unambiguous assignment to a specific category.

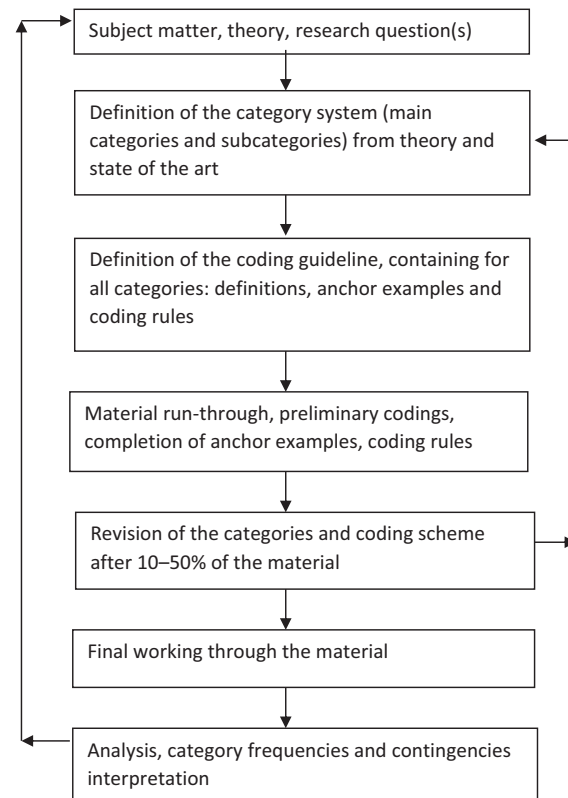


Fig. 2 Process model of deductive category application (structuring).

Test extracts are taken from the material to check whether the categories are at all applicable and whether the definitions, anchor samples and coding rules make category assignment possible. This trial run-through, like the main run-through proper, is subdivided into two steps of operation. First of all, the text passages in the material are marked in which the category concerned is addressed. These “points of discovery” can be marked by noting the category number in the margin of the text or through differently colored underlines in the text itself. In the second step the material thus marked is processed in accordance with the structuring intention (see below) and copied out of the text. As a rule, this trial run-through results in a revision and partial reformulation of the category system and its definitions. Now the main material run-through can finally begin, again split up into the two stages of marking the points of discovery and extracting and processing them. This general description of a structuring content analysis can be shown in a procedural model (Fig. 2).

To further explain the procedure for all techniques of Qualitative Content Analysis, rules of interpretation have been formulated. Those step-models and content analytical rules are explained in detail in Mayring (2021, 2022). Here we just demonstrate the idea of rule-orientated text analysis.

In the above-mentioned study on learning emotions (Glaeser-Zikuda and Mayring, 2003), we developed a category system with 3-point ordinal scales (much—some—no) for four central learning emotions: joy, interest, anxiety and boredom. We established a coding guideline containing definitions, anchor examples and coding rules for those 12 categories. Every student was coded one value (much—some—no) for the four emotions. We divided the sample into high and low achievers and compared the emotion results. In the material (interview, learning diary) of high achievers, significantly more joy was coded ($p < 0.05$; Mann-Whitney- $U = 42.00$), more interest ($p < 0.05$, Mann-Whitney- $U = 12.50$), but no significant difference in boredom was found.

Software support

In addition to existing approaches in computer-assisted qualitative analysis software (see e.g., Silver and Lewins, 2014), a purpose-made software was necessary. This was because the separate processing steps involved in qualitative content analysis are difficult to implement in conventional and commercial programs. For example, it is not easy to keep the central content-analytical rules (category definitions, levels of abstraction, coding guidelines) constantly visible alongside the analysis. This only works to a limited extent using the memo function in MAXQDA or ATLAS/ti, and that actually belongs to the grounded

theory methodology. A table notation, central for summary and coding guidelines, can only be achieved partially. For these reasons, we have developed our own software program QCAmapp (Fenzl and Mayring, 2017; Mayring, 2021), which offers the following advantages:

- Free use
- Interactively guiding users through the steps of content analysis
- Templates for the individual techniques such as summary, inductive category development, and deductive category application
- Templates for the individual analysis units that should be defined as well as content analysis rules
- Ongoing maintenance and further development of the program as a web application
- Interactive possibilities for raters, also for intercoder comparisons

The interactive nature of the program ensures that the essential steps of the qualitative content analysis are actually carried out. The program has been used in more than twenty thousand projects since 2013. In future versions, we plan to integrate video analysis into the program.

Different versions of qualitative content analysis

Finally, I want to compare the outlined approach with similar methodological developments (cf. Mayring, 2020). Within media analysis, Altheide (1996) has developed a procedure (“ethnographic content analysis”) working with deductive categories (codes), which were refined in the process of analysis. Then he summarizes the results for each category. This has similarities with our approach but is not so rule oriented as Qualitative Content Analysis. In the USA there exists an approach coming from quantitative content analysis which is called Codebook Analysis (Neuendorf, 2002). It is a deductive category application procedure, which defines all categories in the codebook and gives examples from the text. But this definition is not so systematic as the coding scheme (definitions, anchor examples and coding rules) in our procedure. In some ways similar is Thematic Text Analysis (Stone, 1997), which looks in the text for central themes, using theoretical preconceptions or empirical word frequencies and word contingencies. In both cases Qualitative Content Analysis defines the procedure more precisely. The related concept of Theme Analysis covers more free, phenomenological procedures (Meier et al., 2008). Some similarities can be found between Qualitative Content Analysis and text analysis following Berg (2004). He describes deductive (“analytic”) and inductive (“grounded”) categories which have to be defined explicitly, but it remains unclear how this has to be done.

The conceptual design of the qualitative content analysis by Kuckartz (2014) is an approach, which mixes systematic methods of qualitative content analysis from our conception with broader open interpretive procedures. As a first step, Kuckartz’s method begins with an initial processing of the text, which is hermeneutically interpretive and involves reading the complete text. While doing so, important points for the research question should be marked, yet this remains relatively vague. This should lead to case summaries and memos about the text. Then a profile matrix should be created, consisting of a table that, for example, should contain central topics and selected passages from each interviewee’s transcription. Only then does the analysis begin with basic content analysis methods, again requiring several text processing repetitions. This is only possible with smaller amounts of text and would, according to my calculations, require more than half a year to work on the text alone for an interview study with 30 people. In my opinion, this throws away some of the potentials of content analysis. Besides, the proposal actually represents a mixed form of different text analysis approaches that are not justified, since some steps are very openly interpretative and not very verifiable, and other steps are strictly guided and more verifiable. This could also be a problematic blending of different positions in scientific theory (constructivism with post-positivism). From my point of view, it also poses a contradiction because on the one hand, summaries should be created, which is a rather explorative process, but, on the other hand, a question is set out on which all analyses should be based. Furthermore, it is not clear which role the (very complex) summaries should play. Finally, basic methods of qualitative content analysis are presented that are based on my work, but only represent a selection of the repertoire and does not offer supporting argumentation for this selection. From the techniques I presented, Kuckartz (2014) selected three (content, type building, and scaling structuring, see **Basic techniques** section of this chapter); however, it is not clear why these three were preferred to the other techniques that I developed. Furthermore, when discussing case summaries, he does not refer to the content-analytical procedure described in **What is qualitative content analysis?** section and, for example, in my opinion, fails to take advantage of one of the most important inductive category formation techniques available.

The concept of qualitative content analysis by Schreier (2012) seems similarly problematic and confusing to me. Here the process is described in such a way that a higher-level category system (coding frame) is developed first, the categories are then defined and pilot-tested, and finally applied to the material in a further step. I think this is far too time-consuming in the case of inductive category development. Instead, I suggest both developing the categories and assigning them to the text in a single pass through the material. In the case of Schreier’s description of inductive category development, a detailed coding guide (category definition, anchor examples, and coding rules) is not necessary; I only designed it for deductive category applications. Theory-based definitions for each category are not feasible here since categories for the material are developed ad hoc. Anchoring examples are not necessary because the text passages were developed with the categories. Thus, they are marked already and represent the anchoring example. Only coding rules for setting boundaries were addressed, but that would not have been necessary

either, as these would have to be carried out for each inductive category. Usually, those categories are summarized during the analysis or hierarchically structured (main categories). Schreier then describes four different approaches for the development of inductive categories (data driven): summary, grounded theory coding, subsumption, and contrasting. This also seems somewhat implausible and of only limited use. Summarizing is usually far too complex for inductive category formation since all the material must be considered. I only regard summaries to be useful if the material is very homogeneous and the question is open and descriptive. In contrast, once a category definition and a defined level of abstraction have been established, inductive categorization continuously creates new categories on a previously defined topic. Grounded theory coding follows an open, exploratory, theory-building strategy that should be accompanied by further processes, in particular, case contrasting, theoretical sampling and memo writing. That differs from the content-analytical inductive category formation since a category definition is already specified as a selection criterion, and categories are formulated in the sense of a descriptive design (not exploratory, categories still have to be checked). Subsumption, the third inductive strategy according to Schreier, and does not in my understanding lead to inductive categories at all, but only assigns further text passages to previously developed inductive categories, if so desired, to reach category frequencies.

In comparison with those text analytical approaches, Qualitative Content Analysis seems to be most broad (describing a wide set of different procedures) and most exact (prescribing clear step models and analytical rules). So Steigleder (2008), after a praxis test of Qualitative Content Analysis, concludes that “it has proven its worth in many studies. With its different techniques of analysis and its methodological concept it is excellently adapted to analyze qualitatively collected material” (Steigleder, 2008, p. 197). But it should not be argued that Qualitative Content Analysis is the only legitimate text analysis procedure. It depends on the concrete research question and the quality of the material which procedure should be chosen. If use of the strict category relatedness and rule orientation of Qualitative Content Analysis neglects important deeper aspects of the material (e.g., repressions in the sense of psychoanalysis), then other procedures (e.g., psychoanalytical text interpretation) would be more adequate.

Final remarks

Qualitative Content Analysis enables a systematic analysis of textual material especially important for educational research. Transcripts from open ended data collection methods present insights in inner processes of learning and development.

Especially because of its intermediate position between qualitative and quantitative analysis, qualitative content analysis appears to be an important methodical approach. It represents a mixed methods analysis, combining qualitative steps (assignment of categories to text segments) and quantitative steps (analyzing category frequencies). It provides a series of concrete techniques (summary, inductive category formation, explication, deductive category assignments, mixed techniques) for different research questions approaching texts.

The central element of qualitative content analysis is its rigor, comprehending clear units of analysis, step models of techniques, and final checks for coding agreements. By this systematic, qualitative content analysis can meet scientific quality criteria, representing an alternative to more free and less controlled interpretation in text analysis.

References

- Altheide, D.L., 1996. *Qualitative Media Analysis*. Sage, Thousand Oaks, CA.
- Berelson, B., 1952. *Content Analysis in Communication Research*. Free Press, Glencoe, Ill.
- Berg, B.L., 2004. *Qualitative Research Methods for the Social Sciences*, fifth ed. Pearson, Boston, MA.
- Carrera-Fernandez, M., Guardia-Olmos, J., Pero-Ceborella, M., 2014. Qualitative methods of data analysis in psychology: an analysis of the literature. *Qual. Res.* 14, 20–36.
- Fenzl, T., Mayring, P., 2017. QCAMap: eine interaktive Webapplikation für Qualitative Inhaltsanalyse. *Zeitschrift für Soziologie der Erziehung und Sozialisation* 37 (3), 333–340.
- Gerbner, G., Holsti, O.R., Krippendorff, K., Paisley, W.J., Stone, P.J. (Eds.), 1969. *The Analysis of Communication Content*. Wiley, New York.
- Glaeser-Zikuda, M., Mayring, P., 2003. A qualitative oriented approach to learning emotions at school. In: Mayring, P., Rhoeneck, C.v. (Eds.), *Learning Emotions: The Influence of Affective Factors on Classroom Learning*. Lang, Frankfurt, pp. 103–126.
- Krippendorff, K., 2019. *Content Analysis. An Introduction to Its Methodology*, fourth ed. Sage, Thousand Oaks. Orig. 1980.
- Kuckartz, U., 2014. *Qualitative Text Analysis. A Guide to Methods, Practice & Using Software*. Sage, London.
- Mandl, H., 1981. *Zur Psychologie der Textverarbeitung. Ansätze, Befunde, Probleme*. Urban & Schwarzenberg, München.
- Mayring, P., 2000. Qualitative content analysis. *Forum Qual. Soc. Res.* 1 (2). Art. 20. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204>.
- Mayring, P., 2015. Qualitative content analysis. Theoretical background and procedures. In: Bikner-Ahsbals, A., Knipping, C., Presmeg, N. (Eds.), *Approaches to Qualitative Research in Mathematics Education. Examples of Methodology and Methods*. Springer, New York, pp. 365–380.
- Mayring, P., 2020. Qualitative content analysis: demarcation, varieties, developments [30 paragraphs]. *Forum Qual. Soc. Res.* 20 (3). <https://doi.org/10.17169/fqs-20.3.3343>. Art. 16.
- Mayring, P., 2021. *Qualitative Content Analysis—a Step-to-Step Guide*. Sage, London.
- Mayring, P., 2022. *Qualitative Inhaltsanalyse. Grundlagen und Techniken*, thirteenth ed., first ed. 1983. Beltz, Weinheim.
- Meier, A., Bovine, M., Meier, M., 2008. Theme-analysis: procedures and application for psychotherapy research. *Qual. Res. Psychol.* 5, 289–310.
- Murphy, G.L., 2002. *The Big Book of Concepts*. MIT Press, Cambridge.
- Neuendorf, K.A., 2002. *The Content Analysis Guidebook*. Sage, Thousand Oaks, CA.
- Schreier, M., 2012. *Qualitative Content Analysis in Practice*. Sage, London.
- Silver, C., Lewins, A., 2014. *Using Software in Qualitative Research: A Step-By-Step Guide*, second ed. Sage, London.
- Steigleder, S., 2008. *Die strukturierende qualitative Inhaltsanalyse im Praxistest. Eine konstruktive kritische Studie zur Auswertungsmethodik von Philipp Mayring*. Marburg, Tectum.

- Stone, P.J., 1997. Thematic text analysis: new agendas for analyzing text content. In: Roberts, C.W. (Ed.), *Text Analysis for the Social Sciences. Methods for Drawing Statistical Inferences From Texts and Transcripts*. Erlbaum, Mahwah, NJ, pp. 35–54.
- Strauss, A., Corbin, J., 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage, Newbury Park.
- Strauss, A.L., 1987. *Qualitative Analysis for Social Scientists*. University of Cambridge Press, Cambridge.
- Titscher, S., Meyer, M., Wodak, R., Vetter, E., 2000. *Methods of Text and Discourse Analysis*. Sage, London.
- Ulich, D., Haußer, K., Mayring, P., Strehmel, P., Kandler, M., Degenhardt, B., 1985. *Psychologie der Krisenbewältigung. Eine Längsschnittuntersuchung mit arbeitslosen Lehrern*. Beltz, Weinheim.

Rapid qualitative research

Georgia Chisnall^a, Stephanie Kumpunen^b, and Cecilia Vindrola-Padros^a, ^a Department of Targeted Intervention, University College London (UCL), London, United Kingdom; and ^b Department of Applied Health Research, University College London (UCL), London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Aims and structure of the chapter	324
Definitions of rapid qualitative research	324
Brief overview of rapid qualitative approaches and their history	325
Five key features of rapid qualitative research	325
The challenges of rapid qualitative research	330
Overcoming assumptions	330
Theoretical grounding and reflexivity	330
Breadth versus depth	330
Sampling and rapport	330
Team-based fieldwork and analysis	331
Funding	331
Ethics and governance	331
The future of rapid qualitative research	331
New approaches to rapid qualitative research	332
Conclusions	332
References	333

Qualitative data provide important insights into behavioral and perceptual factors often missed in quantitative research. In other words, qualitative data are often referred to as providing the “how”, while other typically quantitative data provide the “what” (Teti et al., 2020; Vindrola-Padros et al., 2020). A lesser discussed phenomenon is how the utility of such qualitative approaches is bound by timely delivery of study findings. Timeliness is a core feature which determines the extent to which research findings can be used to inform decision-making processes and be successfully translated into policy and practice (Grasso, 2003; McNall et al., 2004; McNall and Foster-Fishman, 2007; Vindrola-Padros et al., 2021; Vindrola-Padros, 2021a,b; Vindrola-Padros and Johnson, 2020). In some instances, research topics are particularly time-sensitive and can only be studied through rapidly implemented research (Johnson and Vindrola-Padros, 2017).

Understandably, this is the most common reason for enlisting rapid qualitative research methods, whether this is in the form of a study with a compressed timeframe, or integration of regular feedback loops within a longer-term project—both serve to present findings at strategic timepoints (Vindrola-Padros et al., 2021). Beyond timeliness, rapid qualitative methods can also enable mid-course program corrections, adaptation to changing circumstances due to flexible study designs, and communication with stakeholders (Vindrola-Padros et al., 2021). While time reduction tends to be the most frequently cited reason for the use of rapid techniques, a systematic review found five other reasons for adoption of such methods: to reduce cost, to increase the amount of collected data, to improve efficiency, to improve accuracy, and to obtain closer approximation to the narrated realities of research participants (for example, by analyzing findings directly from audio data to avoid transcription errors) (Vindrola-Padros and Johnson, 2020). Hence, long-term fieldwork might not be required for all research topics or contexts and can even limit the contributions of qualitative methods to inform changes in policy and practice (Beebe, 2014; Pink and Morgan, 2013; Abramowitz et al., 2015; Johnson and Vindrola-Padros, 2017).

Rapid qualitative research is becoming more widely used in contexts such as healthcare and education (Vindrola-Padros et al., 2021). Rapid data collection and analysis approaches are used to capture complexities surrounding service provision, the social and cultural factors shaping service use and delivery, and the nuanced practices of service provision in short time frames, which in turn, can be used to inform and shape implementation and future service design (Vindrola-Padros and Vindrola-Padros, 2018). More specifically, rapid qualitative data can be responsive to: local contexts for drawing on community resilience mechanisms; rumors and associated population-level behaviors; societal tensions which disproportionately affect marginalized populations; and organizational challenges of response efforts, highlighting associated gaps and omissions (Johnson and Vindrola-Padros, 2017).

While rapid qualitative methods can be used in a wide range of settings, topics and populations, there will always be questions and contexts that are not suitable for rapid research approaches (Vindrola-Padros and Johnson, 2020). Furthermore, even in contexts where rapid research may be beneficial there are a series of situations in which rapid qualitative studies should be avoided (Vindrola-Padros, 2021a; Vindrola-Padros and Johnson, 2020), such as: when researchers are being asked to cut corners which may violate the quality or ethical standards of research; when study findings may have (intended or unintended) negative

consequences; when researchers will not be able to access people or spaces required to carry out data collection; when time and resources are too constrained; when there is a lack of existing evidence to help guide the study; when there is a lack of willingness from researchers to engage with new methodologies or techniques; or when researchers lack the experience required to carry out rapid research.

Aims and structure of the chapter

This chapter is designed as an introduction to rapid qualitative research. Having briefly outlined the “when” and “why” of rapid qualitative research, this chapter progresses to defining rapid qualitative research before providing an overview of a wide range of rapid qualitative approaches and their history. The chapter then discusses five key features of rapid qualitative research. The challenges of rapid qualitative research are briefly put forth, while signposting to potential solutions. The chapter concludes with a discussion of the future of rapid qualitative research.

This chapter aims to equip the reader with knowledge on:

1. When and why rapid qualitative research methods are appropriate
2. How to define, or identify, a rapid qualitative approach
3. The various rapid qualitative approaches, their key features and origins
4. The challenges associated with rapid qualitative approaches

Definitions of rapid qualitative research

Rapid qualitative research is an umbrella term, analogous to “long-form qualitative research” which captures many study designs and approaches to collecting and analyzing data within short study timeframes. Common rapid qualitative research approaches include rapid ethnographic assessments (REAs), rapid assessment procedures (RAPs), rapid qualitative inquiry (RQI), rapid ethnographies (including quick, focused, and short-term ethnographies), and rapid evaluations (including rapid feedback evaluation, real-time evaluation, and rapid assessment). Each of these approaches have their own features, yet some authors have attempted to provide an overarching definition for this field.

McNall and Foster-Fishman (2007) reviewed a wide range of “Rapid Evaluation and Appraisal Methods” (REAM) and hypothesized that these shared the following features:

- Are conducted over a short timeframe (weeks or months)
- Often include participants or members from the community in the phases of study design or implementation
- Combine multiple research methods and triangulate data during data analysis
- Collect data and undertake analysis in parallel, using insights gained from analysis to iteratively shape the data collection process.

Beebe’s definition included some of the features outlined by McNall and Foster-Fishman (2007), but added focus on the insider’s perspective and the use of multiple data sources and triangulation during analysis:

... intensive, team-based qualitative inquiry with (a) a focus on the insider’s or emic perspective, (b) using multiple sources and triangulation, and (c) using iterative data analysis and additional data collection to quickly develop a preliminary understanding of a situation (2014, p. 3).

Rapid qualitative research has also been defined as empirical research that: lasts up to six months, (or longer if composed of multiple short and intensive stages of data collection and analysis); focuses on documenting aspects of the world through the eyes of others; integrates the subjectivities of the researcher as part of the research process (i.e., reflexivity); and engages with some form of social theory (Flick, 2018; Lapan et al., 2011; Vindrola-Padros, 2021a).

All of these definitions mention the intensive nature and short duration of the research, yet there is disagreement between pioneering authors in relation to what constitutes a study that is rapid in duration. Proposed study timeframes for rapid studies have included:

- A few days (e.g., 4–11 days) (c.f. study examples in Johnson and Vindrola-Padros, 2017)
- 2–4 weeks (but can be as quick as 5 days) (Beebe, 2014)
- 3 weeks (Pearson et al., 1989)
- 4–8 weeks (Long et al., 1988)
- 6 weeks (Scrimshaw et al., 1991)
- 8–12 weeks (Bentley et al., 1988)
- 12 weeks (Handwerker, 2001)
- 6 months (24 weeks) (Vindrola-Padros and Vindrola-Padros, 2018).

It is important to note, however, that study duration needs to be considered loosely as some rapid qualitative studies, particularly rapid evaluations, might have a longer duration, but include sequences of short periods of data collection and analysis and sharing of emerging findings in the form of feedback loops or cycles (McNall et al., 2004; Vindrola-Padros et al., 2021).

Brief overview of rapid qualitative approaches and their history

The origins of anthropology-informed rapid qualitative research involved quick surveys in rural settings in the 1970s, mainly in low- and middle-income countries (LMICs) (Chambers, 1994). They examined the local conditions, health seeking behaviors, and beliefs and attitudes of communities to inform the implementation of social or health interventions—seeking out insider or “emic” perspectives to improve existing programmes or services, or to design more effective and culturally appropriate interventions. The rapid timelines were in response to the short timelines and limited budgetary requirements associated with their work. Many studies employed a range of methods and some studies were guided by structured data collection templates or “field manuals” (e.g., Rapid Assessment Procedures, Rapid Ethnographic Assessments) (Scrimshaw and Hurtado, 1987; Bentley et al., 1988). Researchers often involved community members or people trained in fieldwork research, not only for the local knowledge and research capacity they brought to the team, but also because research teams had a desire to empower the communities they studied (Chambers, 1991).

In parallel to the projects described above, anthropologists also began partnering with multilateral and local government agencies to develop rapid evaluation methods to examine project impacts (McNall and Foster-Fishman, 2007). Other methods were then developed as rapid research moved into urban, high-income country (HIC) settings, in response to crisis issues such as the HIV/AIDS epidemic (Fitch et al., 2000).

This section of the chapter includes a description of commonly used rapid qualitative research approaches—grouped where significant overlap exists. An important point to note is that although these approaches share core features, their uniqueness provide researchers with a range of approaches based on preferences on the level of participation of local community, balance of pre-agreed structure, and whether the research is carried out as a lone researcher or in a team. We have presented this sub-section with an adapted table originally developed by the Rapid Research Evaluation Appraisal Lab (RREAL) team (see Vindrola-Padros, 2021b). **Table 1** is intended to allow researchers to use in combination with **Tables 2** and **3** to easily compare approaches and make choices about which approach would best suit a study’s aims, team structure and methods skills (see **Table 1**).

Rapid evaluations are different from other rapid qualitative approaches because they are specifically designed to assess the impact or document the implementation of specific interventions, services or programmes. While some may appear to have exploratory designs, they can be run as process or outcome evaluations. Three of the most common evaluation approaches include: rapid evaluation methods (REM), rapid feedback evaluation (RFE) and rapid cycle evaluation (RCE) (see **Table 2**) (Vindrola-Padros et al., 2020).

In summary, the rapid methods described vary primarily in relation to their aims, participatory nature, and level of pre-defined structure (vs. exploratory nature). The table below (**Table 3**) provides five continuums through which the methods described above can be better understood.

Five key features of rapid qualitative research

As mentioned above, the field of rapid qualitative research has a wide range of approaches, each with their rich history and unique purpose. Yet despite this variability, many of these approaches have characteristics in common, have been used in combination and have even cross-fertilized to create new types of designs. In our reviews of this field (Johnson and Vindrola-Padros, 2017; Vindrola-Padros and Vindrola-Padros, 2018; Vindrola-Padros et al., 2020), we have been able to identify key characteristics of rapid qualitative research that are shared across most approaches. While the characteristics might also describe some long-form qualitative research designs, they appear to be key features of all rapid research studies, and include:

1. The implementation of scoping studies or preparatory work
2. Reliance on team-based research
3. Carrying out data collection and analysis in parallel
4. Considering different types of analyses for different purposes
5. The development of feedback loops to share emerging findings

Scoping or preparatory stages are rapid explorations of potential study topics to define the scope and finalize the research questions as well as to engage with stakeholders before the study design is finalized. In many cases, the preparatory or scoping work, developed prior to the design of a rapid study, can take longer than the period of data collection. However, this stage is important in rapid qualitative research, as it is used to initially understand the area of study and identify the need for research, decide/refine research questions, identify suitable theoretical perspectives, review existing literature/collect information on the area, identify the best rapid research approach, select methods for data collection and analysis and design the dissemination plan/strategy.

Table 1 Overview of common rapid qualitative research approaches.

<i>Approach</i>	<i>Rapid rural appraisal</i>	<i>Participatory rural appraisal</i>	<i>Rapid ethnographic assessment (REA)</i>	<i>Rapid assessment, response and evaluation (RARE) model</i>	<i>Rapid appraisals and rapid assessment processes</i>	<i>Rapid qualitative inquiry ('team-based assessment')</i>	<i>Rapid ethnographies</i>	<i>Rapid assessment procedures (RAP)</i>
Aims	To diagnose/assess a situation/context. To be pragmatic, cost-effective, accurate and relevant/timely.	To diagnose/assess a situation/context. To lead to change/action and share empowering approaches.	To diagnose/assess local conditions to inform the design of effective interventions by exploring indigenous understandings of health issues.	To diagnose/assess a situation/context. To produce recommendations for interventions appropriate for local culture and area. To complement ongoing processes/monitoring rather than replace them. To determine effectiveness of RARE and the recommended intervention (once implemented). To generate local research capacity.	To diagnose/assess a situation/context. To produce recommendations for interventions appropriate for local culture and area with an emphasis on gaining rich data on insiders' perspectives.	To explore/seek understanding on 'why things are as they are' (helpful in investigating issues that are not yet well defined). To help make preliminary decisions about interventions or changes/additional research or monitoring.	To explore/seek understanding in rapid fieldwork periods.	To assess and evaluate programmes. To improve health program planning and evaluation by capturing the social and cultural factors that frame health behaviors. To empower community members.
Origins	Late 1970s, early 1980s in LMICs in agriculture/environmental studies.	1990s adaptation to RRA (also drawing on activist participatory research) with more community participation, covering studies in natural resources management, agriculture, poverty and social programmes, and health and food security in LMICs.	Late 1980s with a focus on public health in LMIC.	Late 1990s/early 2000s with a focus on health crises in the US.	Draws on REAs, ethnographic and case study research, action research and PAR. Developed in LMIC settings.	Child' method of RAP but framed within "systems thinking". Examples span disciplines and industries (e.g., agriculture, education, health). Originally based in LMICs but now being used to inform policy in many different settings.	Long-form ethnography (see Wall, 2014 ; Knoblauch 2005).	1980s in LMICs, driven by time constraints and the high costs associated with long-term evaluation.

Team	High-level of expertise and research skills. Community members involved in analysis and dissemination only.	Locals collect data (and researchers participate in community life).	Depends on individuals familiar with the local culture/area to act as researchers.	Ethnographer-led team including local working groups (e.g., local government or health agencies) with no previous research). Working groups trained and play a central role in the design and implementation of the study by reviewing available local data (surveillance, epidemiological, planning data, GIS maps) and identifying areas or at-risk groups.	Intensive teamwork, including members of the community/local development agencies as part of the team.	Team members (2–5) should be from different disciplines to contribute their own perceptions/theories to generate different ways of collecting/interpreting data. Insiders involved where appropriate.	Carried out by lone researchers or teams.	Multi-disciplinary research team including at least one member from the cultural group of interest (with no research experience required) who all design and implement studies. Development of materials to train team members.
Methods	Range of qualitative methods. Emphasis on making use of existing data.	Semi-structured interviews, group interviews, oral histories, transect walks, mapping and modeling of local conditions, timelines and trend change analysis, and seasonal calendars. Sources include secondary sources, key informants, local residents, and observations.	Defining characteristic is the structured field guide to suggesting the type of information required. More limited range of methods than PRAs and RAs, but multi-staged complementing open-ended approach.	RARE studies are often proposed by the community/area where the study will take place. Combines focus groups, key informant interviews, observations, mapping and rapid 'street intercept' assessment interviews. Uses RARE manuals and procedures adapted to context/topic. Findings feedback throughout. Findings are also shared regularly with key stakeholders throughout the project.	Used to collect data from key contacts or informants in the communities. Interview data are combined with focus groups, community walks and mapping, and surveys. When possible, existing data are used. Iterative data collection and analysis. Triangulation happens through research team interaction during analysis.	Relaxed, semi-structured interviewing (no formal questions written out) that encourages insiders storytelling. Sample focused on the poorer, less articulate, more upset, and those least like the members of the RQI team. Team interviewing used (i.e., at least two team members present at each interview). The interaction of team members captures different aspects of conversations, leading to triangulation of findings, and adds layers of interpretation to the research process. Data analyzed as researchers collect data, allowing near real-time sharing of findings/discovery of gaps before the fieldwork was over.	Similar to long-form ethnographic research (e.g., interviews, observations and documentary analysis).	RAP manual is crucial in guiding researchers without a background in disease-based research skills or qualitative research. Use several data collection methods to verify information through triangulation. Iterative data collection and analysis to facilitate continuous adjustment.

(Continued)

Table 1 Overview of common rapid qualitative research approaches.—cont'd

<i>Approach</i>	<i>Rapid rural appraisal</i>	<i>Participatory rural appraisal</i>	<i>Rapid ethnographic assessment (REA)</i>	<i>Rapid assessment, response and evaluation (RARE) model</i>	<i>Rapid appraisals and rapid assessment processes</i>	<i>Rapid qualitative inquiry ('team-based assessment')</i>	<i>Rapid ethnographies</i>	<i>Rapid assessment procedures (RAP)</i>
Key references	Chambers (1981), Heywood (1990), Manderson and Aaby (1992)	Chambers (1994), Chambers and Blackburn (1996), Chambers (2008)	Agyepong et al. (1995), Bentley et al. (1988), Culhane-Pera et al. (2015), Garko (2007), Goepp et al. (2008), Guerrero et al. (1999), Hundt et al. (2004), Kresno et al. (1993), Murray et al. (2016), Scott et al. (2014), Shaw et al. (2016), Williams et al. (1999), Kroeger and Sangaramoorthy (2020).	Brown et al. (2008), Fitch et al. (2000), Needle et al. (2003), Trotter et al. (2001).	Beebe (1995, 2005), Desmond et al. (2005), Garrett and Downen (2002), Kirsch (1995), Trotter and Singer (2005), Vincent et al. (2000).	Beebe (2001, 2014).	Ackerman et al. (2017), Chesluk and Holmboe (2010), Hussain et al. (2015), Saleem et al. (2015).	Fitch et al. (2000), Harris et al. (1997), Manderson and Aaby (1992), Scrimshaw and Hurtado (1987), Utarini et al. (2001).

Adapted from Vindrola-Padros et al. (2021).

Table 2 Types of rapid evaluation.

	<i>Rapid evaluation methods</i>	<i>Rapid feedback evaluation</i>	<i>Rapid cycle evaluation</i>
Aims & Origins	Aims: to evaluate and identify operational problems within a program/project and take action. Origins: developed by the World Health Organization to assess quality of health care services, identify operational problems, and assisting managers in taking corrective action.	Aims: to evaluate program/project/system and provide program managers with focused, timely evaluation conclusions. Origins: based on Joseph Wholey's experiences as an evaluator working with US federal program managers and his awareness of clients' needs for the timely delivery of findings to inform key programming decisions (McNall and Foster-Fishman, 2007).	Aims: to evaluate program/project/system, providing timely feedback to funding organization, program staff and care providers. To offer support for continuous quality improvement. Origins: influenced by PDSA cycles used in quality improvement.
Methods	Team: core team involving local community who design and implement evaluation. Active involvement and participation of all stakeholders (to increase engagement and use of findings). Methods: typically mixed-methods. Observation and survey-based diagnostic activities.	Team: team-based. Methods: quantitative or mixed methods. Data are continually collected (e.g., interviews, existing data), analyzed and used to inform action within a short time period.	Team: stable team (to ensure access and build relationships with sites). Methods: quantitative, qualitative or mixed methods.
Authors	Anker et al. (1993), Aspray et al. (2006), Chowdhury et al. (2004), Felisberto et al. (2008), Grant et al. (2011), Munday et al. (2018), Pearson (1989)	Bjornson-Benson et al. (1993), Hargreaves (2014), McNall et al. (2004), Sonnichsen (2000), Wholey (1983), Zakocs et al. (2015)	Keith et al. (2017), McNall and Foster-Fishman (2007), Schneeweiss et al. (2015), Shrank (2013), Skillman et al. (2019)

Adapted from Vindrola-Padros (2020).

Table 3 Questions guiding the selection of rapid qualitative approach.

<i>Questions to guide study design</i>	<i>Types of rapid qualitative research approaches</i>		
What is the aim of the study?	Diagnostic purposes: rapid appraisal, RRA	Exploration or to seek understanding: rapid ethnography, RQI	Evaluation: rapid evaluations, RAP
How participatory?	High degree of participation: PRA, RARE	Medium degree of participation: rapid appraisal, RQI	Low degree of participation: rapid ethnography, rapid evaluations
How structured?	Structured: RAP, REA	Somewhat structured: rapid evaluations, rapid appraisals, RQI, RRA, PRA	Unstructured: rapid ethnographies
Team or lone researcher?	Team-based: rapid appraisals, RQI, RARE, RAP, team-based rapid ethnographies		Lone researcher: rapid ethnographies and some types of rapid evaluations
When are findings needed? (one time-point or regular feedback?)	Regular feedback: rapid feedback or rapid cycle evaluations, rapid appraisals, RAP, RARE, rapid ethnographies		One-time feedback: REM, rapid ethnographies

Adapted from RREAL's Rapid qualitative research decision tool (available at <https://www.rapidresearchandevaluation.com/resources>).

Most of the rapid qualitative research approaches described above, except perhaps rapid ethnographies, rely on the use of teams of researchers. In the case of some approaches, such as RAPs, teams of field researchers are used to cover more ground, that is, carrying out more observations, interviews or visits to communities when compared to research carried out by individual researchers. Rapid appraisals and rapid qualitative inquiry conceptualize team-based research in a slightly different way, recognizing the importance of teams of researchers from different disciplines, and with levels of expertise and types of knowledge (including insiders and outsiders in community-based research). These teams carry out data collection together, including interviews and observations, and each member of the team contributes a different perspective to data collection and analysis, generating additional layers of insight or interpretation.

Data collection and analysis are seen as stages that need to be carried out in parallel. Thus, analysis begins with the first observation, interview or focus group and data are summarized on an ongoing basis. The summaries of findings in real-time mean that

these can be shared with stakeholders in a timely way. Some analyses can be focused on generating a high-level synthesis of findings if these are required quickly, but some rapid studies have also relied on more in-depth analysis, where researchers have more time to develop fuller data sets, engage with more complex theoretical frameworks and provide deeper insight into the findings. The type of analytical approach was selected based on the purpose of the analysis, whether this is to share findings that can be actioned quickly, inform the design of future studies, or make a contribution to academic research.

The type of data collected or analysis undertaken in rapid research is iteratively changing based on responses from stakeholders during project team feedback. While the main purpose of feedback loops is to share emerging findings so these can be used to inform decision-making processes, feedback from stakeholders can also be used to make changes in the study design. For instance, addressing areas that might have been missed or expanding the sampling framework to include additional points of view.

The challenges of rapid qualitative research

Delivering timely and actionable findings through implementation of rapid qualitative approaches remains a challenge and continues to lag behind the needs of evidence-based decision making (Glasgow et al., 2014; Riley et al., 2013; Vindrola-Padros et al., 2020, 2021). This section explores some of the challenges associated with rapid qualitative research which may account for this, and signposts to empirical as well as methodological papers that discuss the solutions and trade-offs.

Overcoming assumptions

One of the main challenges faced by researchers using rapid qualitative methods is overcoming the assumption that rapid research lacks rigor (Pink and Morgan, 2013; Vindrola-Padros, 2021a). Some authors warn against rapid methods for mainstream use (e.g., Fitch et al., 2000), while others warn of poor validity (e.g., Utarini et al., 2001). Alternatively, other research groups acknowledge these challenges but instead seek to develop strategies to overcome them, viewing implementation as the limitation rather than the study duration itself (Vindrola-Padros et al., 2020).

There is a need to distinguish between “rapid” and “rushed” research, rapid does not inherently make for a rushed or poorly implemented study (Beebe, 2005; Fitch et al., 2000; Manderson and Aaby, 1992; Vindrola-Padros, 2021a). As put by Robert Chambers, rapid approaches could be “... too over-sold, too rapidly adopted, badly done, and then discredited, to suffer an undeserved, premature burial as has occurred with other innovative research approaches” (Chambers, 1991). Unfortunately, these assumptions can influence how findings are viewed and used in practice, irrespective of the quality of their design and implementation (McNall and Foster-Fisher, 2007).

Theoretical grounding and reflexivity

Maintaining theoretical grounding during rapid qualitative research can be challenging due to time pressures and the applied nature of the research aims (Vindrola-Padros, 2021a; Vindrola-Padros and Johnson, 2020; Vougioukalou et al., 2019). This has resulted in some authors arguing that rapid qualitative research is atheoretical and lacks a critical angle (Cupit et al., 2018; Manderson and Aaby, 1992). However, there are many instances where knowledge can be produced in rapid timeframes which have strong theoretical roots (Pink and Morgan, 2013; Vindrola-Padros and Johnson, 2017).

Similarly, a recent review of rapid ethnographies revealed that not enough attention is paid by researchers to reflexivity, this being the authors’ critical analysis of their own epistemological influences, which interact with the production of knowledge and the ways in which the researcher’s presence might alter the practices being observed (Vindrola-Padros and Vindrola-Padros, 2018; Vindrola-Padros, 2021a). This has been reported in relation to team-based approaches, as typical of rapid qualitative research (Hesse-Biber, 2016; MacQueen and Guest, 2008). While this represents a significant limitation, it is one which is not limited to rapid research and reflects pitfalls in execution rather than an inherent limitation of the methods of rapid qualitative research (Vindrola-Padros, 2021a,b). For instance, a team-based reflexivity model has been developed and tested in groups undertaking rapid qualitative research (Rankl et al., 2021).

Breadth versus depth

Breadth and depth pose a challenge in rapid qualitative research, as it is not always feasible to capture all of the relevant social, cultural, political, or economic factors at hand, or to engage with all contradictions (Vindrola-Padros, 2021a,b; Bentley et al., 1988; Manderson and Aaby, 1992b). Scope, therefore, is a key factor to take into consideration. To prioritize the areas of study, a number of techniques (e.g., scoping studies) in rapid qualitative research have been developed to ensure the most salient issues are captured and, within reason, incorporate varied views (Vindrola-Padros, 2021a,b).

Sampling and rapport

Sampling strategies in rapid qualitative research could represent a challenge. Time constraints may mean that the sample reflects those most willing to engage with research and omits those with alternative views, or that sample sizes are small (Vindrola-

Padros and Vindrola-Padros, 2017). This said, reviews have found rapid research to justify sampling strategies and to seek some degree of representativeness (Vindrola-Padros and Vindrola-Padros, 2017). While a challenge, sampling in rapid qualitative research can capture a wide range of views through careful selection of study participants and implementation of sampling guides which set targets for the recruitment of key groups (Beebe, 2014; Handwerker, 2001).

Another challenge, due to the short duration of fieldwork, may include developing in-depth rapport with study participants. This said, just because a study is rapid does not mean that exposure to individuals may be less, as similar durations of contact may be similar at an increased intensity (Pink and Morgan, 2013). Where this is not the case, while time does of course contribute to the formation of strong relationships, relationships also stem from the researcher's ability to foster rapport independent of duration. Furthermore, rapport can be integrated into rapid qualitative research from pre-existing relationships, previous studies, or times outside of data collection periods (Vindrola-Padros, 2020).

Team-based fieldwork and analysis

Rapid qualitative research frequently demands the set-up of research teams in order to facilitate shorter research timeframes (Vindrola-Padros et al., 2020; Beebe, 2014; Brown et al., 2008). This can be challenging due to the various backgrounds within a research team, creating a collaborative working environment, and maintaining communication across a number of actors (Barry et al., 1999; Bikker et al., 2017). Furthermore, team-based approaches can delay research findings due to challenges with varying levels of experience, training needs, maintaining consistency, exploring everyone's perspectives, reaching consensus, collaborative writing, and settling authorship disagreements (Rankl et al., 2021).

While team dynamics in rapid qualitative research present several challenges, there is benefit in combined knowledge bases for improving rigor of data collection and analysis through the strength of collective interpretation (Rankl et al., 2021; Barry et al., 1999; Beebe, 2001, 2014; Armstrong and Lowndes, 2018). Furthermore, limitations can be problem-solved, such as maintaining consistency through the use of standardized tools (e.g., RREAL Sheets—previously referred to in publications as RAP sheets) (Bikker et al., 2017; Rankl et al., 2021).

Funding

Rapid qualitative research teams receiving funding per project may have challenges maintaining staff with full-time contracts. Hence, teams may have to combine rapid projects with long-term studies or conduct multiple rapid studies simultaneously in order to cover percentages of full-time employment or part-time contracts. Other, more traditional, means of obtaining funding via academic channels or grants may be possible, but rapid research teams will most likely use hybrid funding models where funding is sought from a number of options, including private consultancies.

Ethics and governance

Delays can often be produced by review of studies by research ethics committees (Driscoll et al., 2008; Iedema et al., 2012; Polito et al., 2014; Vindrola-Padros, 2021a). Reported challenges already exist in relation to qualitative research more broadly, but rapid qualitative research faces further challenges as an uncommon type of design which can result in misunderstandings and produce further delays with gaining approvals. Despite the implementation of a number of fast-track processes, ethical review continues to delay the set-up of rapid studies and deter research teams from engaging with rapid qualitative methods (Vindrola-Padros, 2021a).

In part, delays in gaining ethical approvals may be due to the deemed intrusiveness or burden for research participants, whether they be patients, healthcare workers, public health authorities or members of the public (Vindrola-Padros et al., 2020). While it is important to consider whether research should be carried out at a particular time and potential burden or harm, it could also be unethical not to carry out research and miss immediate, actionable data which could inform response efforts and preparedness strategies (Vindrola-Padros et al., 2020). Furthermore, steps can be taken to reduce burden (flexible interviewing times and shorted durations) and participants have reported that, despite burden, taking part in interviews can be a helpful, even therapeutic, process (Vindrola-Padros et al., 2020).

The future of rapid qualitative research

The rich history of rapid qualitative research spans decades, a period of time during which researchers have undertaken both methodological experimentation and adaptation, as well as reflected on their learning through various journal articles and practical handbooks. Standing back from the various methodological approaches, it is clear that the aims and approaches to data collection and analysis have been and continue to be combined or undergo cross-fertilization. Most of the adaptations to methodologies have been iterative, sometimes developed by the same authors as they refine their thinking and other times by other researchers who adapt methodologies to their own study design or to new disciplines or contexts resulting in a slightly different approach.

As the field of rapid qualitative research moves forward, it is important to recognize the changing nature of practice. New approaches are being produced regularly and this will continue to happen over time. This section includes a description of some of the emerging methods in healthcare and education, as well as other research areas.

New approaches to rapid qualitative research

In recent years, there has been significant experimentation among rapid ethnographies, which has involved combining elements from different rapid research approaches. These new developments or spin-offs include:

- *Focused rapid ethnographic evaluation (FREE)*: this is a modification of the focused ethnography approach and it involves the extensive use of fieldnotes instead of digital recordings (Patmon et al., 2016). FREE can be useful for analyzing human-computer interactions and the assessment of emerging technologies in the workplace (Patmon et al., 2016).
- *Short-term focused video ethnographic case study*: this approach combines video footage, observational field notes, and “video-cued” interviews to provide in-depth analysis of one case study or situation. Harte et al. (2016) have used it to explore the hospital design factors influencing childbirth supporters’ experiences.
- *RAPICE*: the Rapid Assessment Procedure Informed Clinical Ethnography (RAPICE) was developed in the field of clinical trials (Palinkas and Zatzick, 2019) to collect and use qualitative data to understand trial implementation processes and address barriers to recruitment and other issues. It combines some of the characteristics of RAPs (and other rapid qualitative methods that aim to assess and evaluate) with intensive participant observation normally used in clinical ethnography (Zatzick et al., 2011). Study findings are shared regularly with the trial team to inform changes in trial delivery (Palinkas and Zatzick, 2019).
- *Rapid fieldwork approach*: following trends in human-centered computing (HCC) to optimize the design and implementation of technology, Eden et al. (2019) propose the use of multiple focused field trips to sensitize research teams to the needs of stakeholders—however this is often done using extremely compressed timeframes of one to two-day field trips where researchers spend 2–4 h with participants (Eden et al., 2019).

Conclusions

In summary, “rapid qualitative research” is an umbrella term analogous to “long-form qualitative research”, with many approaches emerging over time. Rapid and long-form qualitative research have many common characteristics; however, rapid qualitative research approaches are unique because they near-unanimously rely on: a scoping phase of work, teams of researchers (rather than lone researcher designs), researchers carrying out data collection and analysis in parallel, implementing rapid techniques during analysis, and sharing emerging findings using regular feedback loops. The strengths of rapid qualitative research approaches are that they allow research teams to undertake meaningful qualitative research on short timelines, often for a smaller budget than traditional long-form qualitative research. On the other hand, the challenges of rapid qualitative research include overcoming the assumption that its approaches lack rigor and do not engage with theory, recognizing that the scope and pace of work will likely mean that it will not be possible to capture all of the relevant social, cultural, political factors at hand, and maintaining theoretical grounding despite the time pressures and the applied nature of the research aims.

In producing new rapid qualitative research, it is important to continue to acknowledge the historical roots of approaches, map where innovations are made and learn from these experiences. As Vindrola-Padros et al. (2021) argues,

if we disregard the rich history of rapid qualitative research, we risk making the same mistakes over and over and not learning from lessons already available in the literature. We also run the risk of applying these approaches in unintended ways, diminishing their engagement with theory, increasing their level of instrumentality and severing their ties with traditions in the social sciences. The future development of the field of rapid qualitative research will depend on our efforts to improve these approaches without losing sight of their history.

Vindrola-Padros (2021a,b).

It will also be important that researchers make explicit how theory and theoretical approaches from different disciplines have shaped all stages of research. Where further elicitation of these details is required, methods-focused papers or appendices to accompany the main findings of an academic paper will help explain the research.

Another outlined assumption is that rapid qualitative research is unable to engage with reflexivity, that is, the “researcher’s critical account of their ‘self-location’ (gender, class, ethnicity, and so on), interests, pre-assumptions and life experiences and how these shape their relationships with study participants and the research process itself” (Vindrola-Padros and Vindrola-Padros, 2018). Many publications reporting findings from rapid qualitative studies have not integrated this critical reflection into the text, but

as with issues related to theoretical engagement, there is no reason why time pressures should prevent researchers from considering reflexivity as an integral part of the research process (Vindrola-Padros, 2021a,b) and some potential models for doing this have already been published (Rankl et al., 2021).

References

- Abramowitz, S.A., McLean, K.E., McKune, S.L., Bardosh, K.L., Fallah, M., Monger, J., Tehoungue, K., Omidian, P.A., 2015. Community-centered responses to Ebola in urban Liberia: the view from below. *PLoS neglected tropical diseases* 9 (4), e0003706, 1609406920920962.
- Ackerman, S.L., Sarkar, U., Tieu, L., Handley, M.A., Schillinger, D., Hahn, K., et al., 2017. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in California. *J Am Med Inform Assoc* 24 (5), 903–912.
- Agyepong, I., Aryee, B., Dziku, H., Manderson, L., 1995. The Malaria Manual Guidelines for the Rapid Assessment of Social, Economic and Cultural Aspects of Malaria. WHO, Geneva.
- Anker, M., Guidotti, S., Orszeszy, S., Thuroax, M., 1993. Rapid evaluation methods (REM) of health service performance: Methodological observations. *Bulletin of the World Health Organization* 71 (1), 15–21.
- Armstrong, P., Lowndes, R., 2018. *Creative Teamwork: Developing Rapid, Site-Switching Ethnography*. Oxford University Press, Oxford, UK.
- Aspray, T.J., Nesbit, K., Cassidy, T.P., Hawthorne, G., 2006. Rapid assessment methods used for health-equity audit: Diabetes mellitus among frail British care-home residents. *Public Health* 120 (11), 1042–1051.
- Barry, C.A., Britten, N., Barber, N., Bradley, C., Stevenson, F., 1999. Using reflexivity to optimize teamwork in qualitative research. *Qualitative health research* 9 (1), 26–44.
- Beebe, J., 1995. Basic concepts and techniques of rapid appraisal. *Human Organ.* 54 (1), 42–51.
- Beebe, J., 2001. *Rapid Assessment Process: An Introduction*. AltaMira Press, Walnut Creek, CA.
- Beebe, J., 2005. Rapid assessment process. In: Kempf-Lenoard, K. (Ed.), *Encyclopedia of Social Measurement: Three Volume Set*, vol. 3. Academic Press/Elsevier, San Diego, CA, pp. 285–291.
- Beebe, J., 2014. *Rapid Qualitative Inquiry: A Field Guide to Team-Based Assessment*. Rowman and Littlefield, Lanham, MD.
- Bentley, M., et al., 1988. Rapid ethnographic assessment: applications in a Diarrhoea management program. *Soc. Sci. Med.* 27 (1), 107–116.
- Bikker, A.P., Atherton, H., Brant, H., Porqueddu, T., Campbell, J.L., Gibson, A., McKinstry, B., Salisbury, C., Zieband, S., 2017. Conducting a team-based multi-sited focused ethnography in primary care. *BMC medical research methodology* 17 (1), 1–9.
- Bjornson-Benson, W.M., Stibolt, T.B., Manske, K.A., Zavala, K.J., Youtsey, D.J., Buist, A.S., 1993. Monitoring recruitment effectiveness and cost in a clinical trial. *Controlled Clinical Trials* 14 (2 Suppl), 52S–67S.
- Chambers, R., 1981. Rapid rural appraisal: rationale and repertoire. *Publ. Adm. Dev.* 1 (2), 95–106.
- Chambers, R., 1991. Shortcut and participatory methods for gaining social information for projects. In: Cernea, M. (Ed.), *Putting People First: Sociological Variables in Rural Development*. Oxford University Press/World Bank, pp. 515–537.
- Chambers, R., 1994. Participatory rural appraisal (PRA): analysis of experience. *World Dev.* 22 (9), 1253–1268.
- Chambers, R., 2008. *Revolutions in Development Inquiry*. Earthscan.
- Chambers, R., Blackburn, J., 1996. *The Power of Participation*. Institute of Development Studies, University of Sussex.
- Chesluk, B.J., Holmboe, E.S., 2010. How teams work – or don't – in primary care: A field study on internal medicine practices. *Health Aff* 29 (5), 874–879.
- Chowdhury, S.N.M., Moni, D., 2004. A situation analysis of the menstrual regulation programme in Bangladesh. *Reproductive Health Matters* 12 (24 Suppl), 95–104.
- Culhane-pera, K.A., Sripetcharawut, S., Thawirichuchai, R., Yangyuenkun, W., Kunstadter, P., 2015. Afraid of delivering at the hospital or afraid of delivering at home: a qualitative study of Thai Hmong families' decision-making about maternity services. *Matern. Child Health J.* 19 (11), 2384–2392. <https://doi.org/10.1007/s10995-015-1757-3>.
- Cupit, C., Mackintosh, N., Armstrong, N., 2018. Using ethnography to study improving healthcare: reflections on the “ethnographic” label. *BMJ Qual. Saf.* 27, 258–260.
- Desmond, N., Allen, C., Cliff, S., Justine, B., Mzugu, J., Plummer, M., Ross, D.A., 2005. A typology of groups at risk of HIV/STI in a gold mining town in north-western Tanzania. *Soc. Sci. Med.* 60, 1739–1749. <https://doi.org/10.1016/j.socscimed.2004.08.027>.
- Driscoll, A., Currey, J., Worrall-Carter, L., Stewart, S., 2008. Ethical dilemmas of a large national multi-centre study in Australia: Time for some consistency. *Journal of Clinical Nursing* 17 (16), 2212–2220.
- Fitch, C., Rhodes, T., Stimson, G., 2000. Origins of an epidemic: the methodological and political emergence of rapid assessment. *Int. J. Drug Pol.* 11, 63–82.
- Flick, U., 2018. *An Introduction to Qualitative Research*. SAGE.
- Felisberto, E., Freese, E., Natal, S., Alves, C.K. de A., 2008. [A contribution to institutionalized health evaluation: a proposal for self-evaluation]. *Cadernos de saude publica* 24 (9), 2091–2102.
- Garko, B., 2007. Sexual and family planning practices and needs of people living with HIV/AIDS in Nigeria: a rapid ethnographic assessment. *Ann. Afr. Med.* 6 (3), 124–127. <https://doi.org/10.4103/1596-3519.55721>.
- Garrett, J.L., Downen, J., 2002. Strengthening rapid assessments in urban areas: lessons from Bangladesh and Tanzania. *Human Organ.* 61, 314–327.
- Glasgow, R., et al., 2014. Conducting rapid, relevant research lessons learned from the My Own Health report project. *American Journal of Preventive Medicine* 47 (2), 212–219.
- Goepp, J.G., Meykler, S., Mooney, N.E., Lyon, C., Raso, R., Julliard, K., 2008. Provider insights about palliative care barriers and facilitators: results of a rapid ethnographic assessment. *Am. J. Hospice Palliat. Med.* 25 (4), 309–314. <https://doi.org/10.1177/1049909108319265>.
- Grant, L., et al., 2011. Palliative care making a difference in rural Uganda, Kenya and Malawi: Three rapid evaluation field studies. *BMC Palliative Care* 10, 8.
- Grasso, P., 2003. What makes an evaluation useful? Reflections from experience in large organizations. *American Journal of Evaluation* 24 (4), 507–514.
- Guerrero, M., Morrow, R., Calva, J., Ortega-Gallegos, H., Weller, S., Ruiz-Palacios, G., Morrow, A.L., 1999. Rapid ethnographic assessment of breastfeeding practices in periurban Mexico City. *Bull. World Health Organ.* 77 (4), 323–330.
- Handwerker, P., 2001. *Quick Ethnography: A Guide to Rapid Multi-method Research*. Rowman Altamira, Lanham, MD.
- Hargreaves MB. Rapid evaluation methods from complex initiatives. HHS Office of the Secretary for Planning and Evaluation issue brief. Available from: http://aspe.hhs.gov/sp/reports/2014/EvalApproach/rs_EvalApproach.cfm.
- Harris, K., Jerome, N., Fawcett, S., 1997. Rapid assessment procedures: a review and critique. *Hum. Organ.* 56 (3), 375–378.
- Heywood, P., 1990. *W(h)ither Rapid (Rural) Appraisal Techniques in Nutrition*. Unpublished Manuscript, Brisbane.
- Hundt, G.L., Stuttford, M., Ngoma, B., Team, SASPI, 2004. The social diagnostics of stroke-like symptoms: healers, doctors and prophets in Agincourt, Limpopo Province, South Africa. *J. Biosoc. Sci.* 36 (4), 433–443. <https://doi.org/10.1017/S0021932004006662>.
- Hussain, R.S., McGarvey, S.T., Fruzzetti, L.M., 2015. Partition and poliomyelitis: An investigation of the polio disparity affecting Muslims during India's eradication program. *PLoS One* 10 (3) n/a.

- Iedema, R., Allen, S., Britton, K., Hor, S., 2012. Out of the frying pan? Streamlining the ethics review process of multisite qualitative research projects. *Australian Health Review* 37 (2), 137–139.
- Johnson, G.A., Vindrola-Padros, C., 2017. Rapid qualitative research methods during complex health emergencies: a systematic review of the literature. *Soc. Sci. Med.* 189, 63–75.
- Keith, R., 2017. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: A rapid-cycle evaluation approach to improving implementation. *Implementation Science* 12, 15.
- Kirsch, H., 1995. The use of rapid assessment procedures: focus groups and smallscale surveys for community programs. In: Kirsch, H. (Ed.), *Drug Lessons and Education Programs in Developing Countries*. Transaction Publishers.
- Knoblauch, H., 2005. Focused ethnography. *Forum Qual. Soc. Res.* 6 (3), Art. 44. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0503440>.
- Kresno, S., Harrison, G., Sutrisna, B., Reingold, A., 1993. Acute respiratory illnesses in children under five years in Indramayu, West Java, Indonesia: a rapid ethnographic assessment. *Med. Anthropol.* 15 (4), 425–434. <https://doi.org/10.1080/01459740.1994.9966103>.
- Lapan, S., Quartaroli, M.T., Riemer, F.J., 2011. *Qualitative Research: An Introduction to Methods and Designs*. John Wiley & Sons.
- Long, A., Scrimshaw, S.C.M., Hurtado, E., 1988. *Epilepsy Rapid Assessment Procedures (ERAP): Rapid Assessment Procedures for the Evaluation of Epilepsy Specific Beliefs, Attitudes and Behaviours*. Epilepsy Foundation of America, Landover.
- Manderson, L., Aaby, P., 1992a. An epidemic in the field? Rapid assessment procedures and health research. *Soc. Sci. Med.* 35 (7), 839–850.
- Manderson, L., Aaby, P., 1992b. Can rapid anthropological procedures be applied to tropical diseases? *Health Pol. Plann.* 7 (1), 46–55.
- McNall, M., Foster-Fishman, G., 2007. Methods of rapid evaluation, assessment and appraisal. *Am. J. Eval.* 28 (2), 151–168.
- McNall, M., Welch, V., Ruh, K., Mildner, C., Soto, T., 2004. The use of rapid-feedback evaluation methods to improve the retention rates of an HIV/AIDS healthcare intervention. *Eval. Progr. Plann.* 27, 287–294.
- Munday, D., et al., 2018. Rural palliative care in North India: Rapid evaluation of a program using a realist mixed method approach. *Indian Journal of Palliative Care* 24 (1), 3–8.
- Murray, J.K., DiStefano, A., Yang, J., Wood, M., 2016. Displacement and HIV: factors influencing antiretroviral therapy use by ethnic Shan migrants in Northern Thailand. *J. Assoc. Nurses AIDS Care* 27 (5), 709. <https://doi.org/10.1016/j.jana.2016.04.006>.
- Needle, R.H., Trotter, R.T., Singer, M., Bates, C., Bryan Page, J., Metzger, D., Marcelin, L.H., 2003. Rapid assessment of the HIV/AIDS crisis in racial and ethnic minority communities: an approach for timely community interventions. *Am. J. Publ. Health* 93 (6), 970–979. <https://doi.org/10.2105/AJPH.93.6.970>.
- Palinkas, L., Zatzick, D., 2019. Rapid assessment procedure informed clinical ethnography (RAPICE) in pragmatic clinical trials of mental health services implementation: methods and applied case study. *Adm. Pol. Ment. Health.* <https://doi.org/10.1007/s10488-018-0909-3>.
- Pearson, R., 1989. Rapid assessment procedures are changing the way UNICEF evaluates its projects. *Hygie* 8 (4), 23–25.
- Pink, S., Morgan, J., 2013. Short term ethnography: intense routes to knowing. *Symbolic Interact.* 36 (3), 351–361.
- Polito, C., Cribbs, S., Martin, G., O'Keefe, T., Herr, D., Rice, T., Sevransky, J., 2014. Navigating the Institutional Review Board approval process in a multicentre observational criticalcare study. *Critical Care Medicine* 42 (5), 1105–1109.
- Rankl, F., Johnson, G.A., Vindrola-Padros, C., 2021. Examining what we know in relation to how we know it: a team-based reflexivity model for rapid qualitative Health Research. *Qualitative Health Research* 31 (7), 1358–1370.
- Riley, W., et al., 2013. Rapid, responsive, relevant (R3) research: A call for a rapid learning health research enterprise. *Clinical and Translational Science* 2, 10.
- Saleem, J.J., Plew, W.R., Speir, R.C., Herout, J., Wilck, N.R., Ryan, D.M., et al., 2015. Understanding barriers and facilitators to the use of Clinical Information Systems for intensive care units and Anesthesia Record Keeping: A rapid ethnography. *International Journal of Medical Informatics* 84 (7), 500–511.
- Sangaramoorthy, T., Kroeger, K., 2020. *Rapid Ethnographic Assessments: A Practical Approach and Toolkit for Collaborative Community Research*. Routledge, New York, NY.
- Schneeweiss, S., Shrank, W.H., Ruhl, M., Maclure, M., 2015. Decision-making aligned with rapid-cycle evaluation in health care. *International Journal of Technology Assessment in Health Care* 31, 214–222.
- Scott, K., McMahon, S., Yumkella, F., Diaz, T., George, A., 2014. Navigating multiple options and social relationships in plural health systems: a qualitative study exploring healthcare seeking for sick children in Sierra Leone. *Health Pol. Plann.* 29 (3), 292–301. <https://doi.org/10.1093/heapol/czt016>.
- Scrimshaw, S., Hurtado, E., 1987. *Rapid Assessment Procedures for Nutrition and Primary Health Care*. Anthropological Approaches to Improving Programme Effectiveness. The United Nations University, Tokyo.
- Scrimshaw, S., Carballo, M., Ramos, L., Blair, B., 1991. The AIDS rapid anthropological assessment procedures: a tool for health education planning and evaluation. *Health Educ. Q.* 18 (1), 111–123. <https://doi.org/10.1177/109019819101800111>.
- Shaw, B., Amouzou, A., Miller, N.P., Tafesse, M., Bryce, J., Surkan, P.J., 2016. Access to integrated community case management of childhood illnesses services in rural Ethiopia: a qualitative study of the perspectives and experiences of caregivers. *Health Pol. Plann.* 31 (5), 656–666. <https://doi.org/10.1093/heapol/czv115>.
- Shrank, W., 2013. The Center for Medicare and Medicaid Innovation's blueprint for rapid-cycle evaluation of new care and payment models. *Health Aff.* 32 (4), 807–812.
- Skärås, M., 2018. Focused ethnographic research on teaching and learning in conflict zones: history education in south Sudan. *Forum Dev. Stud.* 45 (2), 217–238.
- Skillman, M., et al., 2019. A framework for rigorous qualitative research as a component of mixed method rapid-cycle evaluation. *Qualitative Health Research* 29 (2), 279–289.
- Sonnichsen, R., 2000. *High Impact Internal Evaluation: A Practitioner's Guide to Evaluation and Consulting Inside Organizations*. Sage, Thousand Oaks.
- Teti, M., Schatz, E., Liebenberg, L., 2020. Methods in the time of COVID-19: the vital role of qualitative inquiries. *International Journal of Qualitative Methods* 19.
- Trotter, R.T., Needle, R.H., Goosby, E., Bates, C., Singer, M., 2001. A methodological model for rapid assessment, response, and evaluation: the RARE program in public health. *Field Methods* 13 (2), 137–159. <https://doi.org/10.1177/1525822X0101300202>.
- Trotter, R.T., Singer, M., 2005. Rapid assessment strategies for public health: promise and problems. In: Trickett, E., Pequegnat, W. (Eds.), *Community Intervention and AIDS*. Oxford University Press.
- Utarini, A., Winkvist, A., Pelto, G., 2001. Appraising studies in health using rapid assessment procedures (RAP): eleven critical criteria. *Hum. Organ.* 60 (4), 390–400.
- Vincent, N., Allsop, S., Shoobridge, J., 2000. The use of rapid assessment methodology (RAM) for investigating illicit drug use: a South Australian experience. *Drug Alcohol Rev.* 19, 419–426. <https://doi.org/10.1080/713659426>.
- Vindrola-Padros, C., 2020. Can we Re-imagine research so it is timely, relevant and responsive? Comment on “experience of health leadership in partnering with University-based researchers in Canada: a call to ‘Re-imagine’ research”. *Int. J. Health Pol. Manag.* <https://doi.org/10.34172/IJHPM.2020.43>.
- Vindrola-Padros, C., 2021a. *Rapid Ethnographies: A Practical Guide*. University of Cambridge Press, Cambridge, UK.
- Vindrola-Padros, C., 2021b. *Doing Rapid Qualitative Research*. SAGE, London, UK.
- Vindrola-Padros, C., Johnson, G.A., 2020. Rapid techniques in qualitative research: a critical review of the literature. *Qual. Health Res.* 30 (10), 1596–1604. <https://doi.org/10.1177/1049732320921835>.
- Vindrola-Padros, C., Vindrola-Padros, B., 2018. Quick and dirty? A systematic review of the use of rapid ethnographies in healthcare organisation and delivery. *BMJ Qual. Saf.* 27 (4), 321–330.
- Vindrola-Padros, C., Chisnall, G., Cooper, S., Dowrick, A., Djellouli, N., Symmons, S.M., Johnson, G.A., 2020. Carrying out rapid qualitative research during a pandemic: emerging lessons from COVID-19. *Qual. Health Res.* 30 (14), 2192–2204. <https://doi.org/10.1177/1049732320951526>.
- Vindrola-Padros, C., Brage, E., Johnson, G.A., 2021. Rapid, responsive and relevant? A systematic review of rapid evaluations in healthcare. *Am. J. Eval.* (in press).

- Vougioukalou, S., Boaz, A., Gager, M., Locock, L., 2019. The contribution of ethnography to the evaluation of quality improvement in hospital settings: reflections on observing co-design in intensive care units and lung cancer pathways in the UK. *Anthropol. Med.* 26 (1), 18–32.
- Wall, S., 2014. Focused ethnography: a methodological adaptation for social research in emerging contexts [40 paragraphs]. *Forum Qual. Soc. Res.* 16 (1). Art. 1.
- Williams, H.A., Kachur, S.P., Nalwamba, N.C., Hightower, A., Simoonga, C., Mphande, P.C., 1999. A community perspective on the efficacy of malaria treatment options for children in Lundazi District, Zambia. *Trop. Med. Int. Health* 4 (10), 641–652. <https://doi.org/10.1046/j.1365-3156.1999.00471.x>.
- Wholey, J.S., 1983. *Evaluation and Effective Public Management*. Little Brown, Boston.
- Zakocs, R., Hill, J.A., Brown, P., Wheaton, J., Freire, K.E., 2015. The data-to-action framework: A rapid program improvement process. *Health Education & Behavior: The Official Publication of the Society for Public Health Education* 42 (4), 471–479.

Establishing quality in qualitative research: trustworthiness, validity, and a lack of consensus

Leia K. Cain^a, Rachel E. Williams^b, and Victoria Bradshaw^b, ^a Department of Evaluation, Statistics, and Methodology, University of Tennessee, Knoxville, TN, United States; and ^b Sport Psychology and Motor Behavior, University of Tennessee, Knoxville, TN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction—quality is complicated	336
The salad bar approach to considering quality in qualitative research	337
Determinants of quality in qualitative research	338
Authenticity	338
Centering ethical considerations	338
Consistency	340
Credibility	341
Critical subjectivity	341
Meaningful coherence	341
Member checking	341
Memoing	342
Positionality	342
Reciprocity	342
Reflexivity	343
Relational ethics	343
Resonance	344
Rich rigor	344
Sacredness	344
Sharing privilege	345
Significant contributions	345
Sincerity	345
Sites as arbiters of quality	346
Transgressive validities	346
Ironic validity	346
Rhizomatic validity	346
Paralogical validity	347
Voluptuous validity	347
Triangulation	347
Voice	347
Worthy topic	348
Conclusion	348
A trip to the salad bar	348
Implications for researchers in education	349
References	349

Introduction—quality is complicated

Qualitative scholars have long debated notions of quality within qualitative research, rarely arriving at a conclusive consensus. In fact, [Onwuegbuzie and Leech \(2005\)](#) reviewed literature and found that there were over 50 terms (such as authenticity, trustworthiness, and transparency, among others) used across qualitative publications at that time. We estimate that this list has grown, and were they to repeat their study, they would likely find even more. While this lack of consensus makes writing a chapter such as this daunting, it is imperative that we engage in these conversations in order to continually improve the craft of inquiry as well as our understanding of how scholars enact various forms of quality within qualitative research.

[Corbin and Strauss \(2008\)](#) summarized the uncertainty around describing *quality* within qualitative research that often leaves both new and established researchers feeling mentally debilitated, saying,

I feel paralyzed, unsure of where to begin, or what to write. As I search the literature, I find that evaluation is necessary but there is little consensus about what that evaluation should consist of. Are we judging for 'validity' or would it be better to use terms like 'rigor' ... 'trustworthiness' ... or 'goodness' ... or something called 'integrity' ... when referring to qualitative evaluation? (p. 297).

This lack of consensus can be traced back to earlier qualitative work as well. For example, many of our foundational scholars argued that the concept of *validity*, stemming from epistemological frameworks primarily found in quantitative traditions, is incompatible with values held by qualitative researchers (Kvale, 1995; Lather, 1993; Ravitch and Carl, 2019; Wolcott, 1990). Lather (1993), however, chose to specifically embrace the term *validity* in order to “both circulate and break with the signs that code it ... all of the baggage that it carries plus, in a doubled-movement, what it means to rupture validity as a regime of truth, to displace its historical inscription” (p. 674). Lather’s reclamation of the term validity marked a turning point in how many qualitative scholars conceptualize quality within their research.

Other scholars argue that there are three key concerns that prevent us from being able to present a unified framework for quality in qualitative research: (1) there is little consensus about what quality research looks like, (2) traditional positivist evaluation criteria, such as notions of validity and reliability within quantitative methods, continue to impact the way that we discuss quality in other methodological approaches, and (3) there are very specific culturally- and politically-defined notions about what constitutes as evidence (Cho and Trent, 2020; Trainor and Graue, 2013). These three complications can help us begin to explain why conversations around quality in qualitative research are complicated.

The factors that constitute good qualitative research are interpretable, mutable, and up for negotiation. These quality indicators should be engaged and critically examined; we encourage qualitative researchers to work within this more flexible framework of quality instead of seeking to uphold foundational traditions stemming from other methodological paradigms. Critically engaging with quality determinants turns our focus away from defining good research as the precise application of methods, as more traditionally used in quantitative methods, and instead toward process and outcomes from our work. However, our focus on process and outcomes does not mean that there are no criteria for determining the quality of qualitative methods—there certainly are!

The indicators of quality criteria that you ultimately choose will depend on your theoretical, ontological, epistemological, and methodological decisions and commitments. Ontological and epistemological commitments are particularly important to consider within qualitative research, as researchers will often choose to pursue quality indicators that fit within their beliefs. In other words, what you consider to be good, rigorous, and valid research stems from your assumptions about how knowledge is produced. For more information about how ontological, epistemological, and methodological considerations fit together and for an exploration of some frequently used paradigms, please see other chapters within this encyclopedia provided by Eileen Boswell and Wayne Babchuk (“Philosophical and Theoretical Underpinnings of Qualitative Research”) and Peggy Shannon-Baker (“Philosophical Underpinnings of Mixed Methods Research in Education”).

Within the following sections, we introduce the concept of what we call The Salad Bar Approach, detail 23 quality determinants that are often used across qualitative scholarship, and finally, summarize the chapter with an included demonstration of how some of the quality determinants can be used.

The salad bar approach to considering quality in qualitative research

You have a variety of choices when it comes to choosing which quality determinants to use within your qualitative project. In fact, there are arguments being made for new determinants all the time! It would be impossible to list and define every single existing quality construct due to practical concerns, such as page limitations, and timeliness concerns, as there are consistently new arguments taking place about what does/does not and should/should not exist in regard to measuring quality. Instead, we wish to share a metaphor that one of our authors, Leia, provides to students within her qualitative methodology courses: the salad bar approach.

Imagine that you are building a salad and that all of the available ingredients, toppings, and salad dressings are spread upon a table before you. You build your salad based on how you, personally, enjoy it; perhaps you include a fair amount of lettuce, a large number of other vegetables, such as zucchini, cucumbers, and carrots, a light sprinkling of cheese, and finally, a touch of ranch dressing. Another person walks up to the table where these ingredients are presented and builds their salad: a few handfuls of arugula, many onions, sunflower seeds, tomatoes, and a massive heaping of raspberry vinaigrette. While your salads differ in how you built them, they are both quality salads. If asked to describe how each ingredient decision was made, you would both be able to justify what you believe to contribute to a good salad.

Qualitative research is very much the same as these salads. Each qualitative researcher may work through the inquiry process in different ways, utilize diverse theoretical frameworks and paradigmatic beliefs, engage in various forms of data collection and analysis, and apply different forms of quality evidence. The most important thing to consider when judging each of these qualitative salads is to justify your composition. While it is often known that authors must provide a justification for readers as to how each action was taken throughout the research process fits their research question, it is less common knowledge that you are

also able to choose and should therefore justify which quality determinants best fit your theoretical, ontological, epistemological, and methodological commitments.

There are few limits across diverse representations of qualitative research—and this discussion of quality is no different. There are many types of quality determinants that you may choose to use in your qualitative work—you must simply ensure that you have justified your decisions and that they fit with the other elements of your study.

Determinants of quality in qualitative research

In this section, we present 23 determinants of quality in qualitative research. We define each construct and provide reflective questions for researchers seeking to utilize each particular determinant. We have also created [Table 1](#), which lists the quality determinants, key citations for each, similarly used terms, and paradigms with which each determinant is commonly paired. Again, we wish to reiterate that there is little consensus within the broad field of qualitative methodologies. We cannot impress upon you enough the vast importance of knowing what is valued within your field, as some of these constructs are more common and accepted than others within individual fields. Our entries, below, have been alphabetized for ease of locating specific determinants.

Authenticity

Scholars who seek authenticity are concerned with both engaging in a worthy topic and understanding the larger implications of their research (i.e. how the project potentially benefits society). [Guba and Lincoln \(1989\)](#) described five dimensions of authenticity: (1) fairness related to the research participants and stakeholders, (2) ontological authenticity, which refers to the alignment between the researcher's paradigmatic approach, (3) educative authenticity, meant to increase the audience's awareness of the issue, (4) catalytic authenticity, which refers to the study's potential for action, and (5) tactical authenticity, referring to change or empowerment as a result of the research project.

When considering fairness, it may be helpful to ask yourself:

- Am I doing what is in the best interest of my research project, my research participants, and the stakeholders?

When considering ontological authenticity, it may be helpful to ask yourself:

- Am I being authentic to my paradigmatic views?
- Am I appropriately employing my paradigms?

When considering educative authenticity, it may be helpful to ask yourself:

- Am I educating my readers clearly and accurately?

When considering catalytic and tactic authenticity, it may be helpful to ask yourself:

- What can my work catalyze people to do?

Centering ethical considerations

Qualitative scholars have long argued for emphasizing ethical considerations within your writing. You should consider the procedures you're using, participant well-being or burden, demonstrate cultural competency, respect relationships, and maintain other ethical boundaries set in place by the nature of the research study. [Miles et al. \(2019\)](#) posited that centering ethics allows research to come out of a place of respect, dignity, and honor to the research process and its participants. It is each researcher's responsibility to uphold ethical guidelines set by institutional review boards and other research governing bodies, yet the concept of centering ethical considerations calls for engagement in ethical thinking outside of the bounds set by those gatekeepers. Additionally, it is the researcher's responsibility to anticipate any ethically challenging situations that could arise throughout the research process and have contingency plans in place. One must remember to maintain ethics throughout project conception, data collection, analysis, and reporting results. These considerations should then be detailed within publications so that readers can understand the ethical thinking that the author(s) engaged in. When ethical considerations are clearly delineated within your work, you are able to bring attention to ethical decision-making throughout the study. It is your responsibility to report adverse events and consider these events in reviewing the conclusion of the study.

When considering centering ethical considerations, it may be helpful to ask yourself:

- What ethical concerns may arise during my study in regard to interacting with participants?
- What are the best ways for me to handle sensitive information shared with me by the participants?
- What steps should I take when working with protected or vulnerable populations (e.g., children, employees, incarcerated individuals)?
- How can I ensure that all members of my research team are on the same page regarding ethical decision-making?

Table 1 Table of quality determinants.

<i>Quality determinant (key citations)</i>	<i>Brief description</i>	<i>Similar terms</i>	<i>Common paradigms</i>
3.1 Authenticity (Lincoln and Guba, 1985)	Researchers must engage in a worthy topic and understanding the larger implications of their research.	N/A	Constructivism Interpretivism
3.2 Centering ethical considerations (Miles et al., 2019; Tracy, 2010)	Researchers must emphasize ethical decision-making within their work.	N/A	Critical paradigms Interpretivism Postpositivism
3.3 Consistency (Drisko, 1997; Lincoln and Guba, 1985; Reid, 1994)	Researchers must understand the literature that contextualizes their study and provide explicit connections between the literature and their findings.	• Confirmability (Johnson et al., 2020) • Dependability (Nassaji, 2020) • Reliability (Yilmaz, 2013)	Postpositivism
3.4 Credibility (Lincoln and Guba, 1985; Patton, 1999; Tracy, 2010)	Researchers must provide enough evidence to demonstrate that their interpretations of the data are meaningful.	• Internal validity (Meijer et al., 2002) • Trustworthiness (Little and Green, 2022)	Interpretivism Postpositivism
3.5 Critical subjectivity (Lincoln, 1995)	Researchers must deconstruct their beliefs, biases, subjectivities, experiences, identities, and more, to understand how their core beliefs affect their decisions throughout the study.	• Transformative subjectivity (Frieden, 1989)	Constructivism Critical paradigms Interpretivism
3.6 Meaningful coherence (Lather, 1993; Tracy, 2010)	Researchers must ensure that their research reports demonstrate significant alignment through their introduction, literature review, methodology, results, and discussion sections.	N/A	Interpretivism Postpositivism
3.7 Member checking (Lather, 1986; Lincoln and Guba, 1985)	Researchers must share their findings with participants in order to confirm the accuracy of their interpretation.	• Member reflections (Tracy, 2010)	Constructivism Critical paradigms Interpretivism Postpositivism
3.8 Memoing (Glaser, 1978)	Researchers must keep a record of their thoughts, feelings, emotional responses, and questions that arise throughout the inquiry process.	N/A	Interpretivism Postpositivism
3.9 Positionality (Lincoln, 1995)	Researchers must reflect on their identities in relation to their study's participants and data.	• Standpoint judgment (Nakata, 2015)	Constructivism Critical paradigms Interpretivism
3.10 Reciprocity (Lincoln, 1995; McLaughlin and Tierney, 1993)	Researchers must consider the bonds that exist between researcher and participant and work to create a relationship marked by trust, care, and mutual benefit.	N/A	Constructivist Critical paradigms Interpretivist
3.11 Reflexivity (Peshkin, 1988)	Researchers must reflect upon their paradigmatic commitments, their relationship to the study participants, and their beliefs about the research topic.	• Critical reflexivity (Glass, 2015) • Transformative reflexivity (Relano Pastor, 2011)	Constructivism Critical paradigms Interpretivism Postmodernism
3.12 Relational ethics (Adams et al., 2015)	Researchers must recognize and protect the connection between themselves, their participants, and the communities in which they live and work.	N/A	Constructivism Critical paradigms
3.13 Resonance (Schutz, 1967; Tracy, 2010)	Researchers must influence readers through evocative representations of their data.	N/A	Constructivism Interpretivism
3.14 Rich rigor (Golafshani, 2003; Tracy, 2010)	Researchers must provide a detailed, clear research design and methodology.	N/A	Interpretivism Postpositivism
3.15 Sacredness (Lincoln, 1995; Schaef, 1981)	Researchers must have a profound respect for human dignity and justice, and seek to turn their findings into action.	N/A	Constructivism Critical paradigms Feminism Interpretivism

(Continued)

Table 1 Table of quality determinants.—cont'd

<i>Quality determinant (key citations)</i>	<i>Brief description</i>	<i>Similar terms</i>	<i>Common paradigms</i>
3.16 Sharing privilege (Lather, 1993; Lincoln, 1995)	Researchers must examine how they can share with participants and communities the earned dignity, respect, prestige, and economic benefits from publishing their work.	N/A	Constructivism Critical paradigms Feminism Interpretivism
3.17 Significant contributions (Schwandt, 1996; Tracy, 1995)	Researchers must use their research to address theoretical, methodological, cultural, moral, and practical gaps within the literature.	N/A	Critical paradigms Postpositivism
3.18 Sincerity (Emerson et al., 1995; Seale, 1999)	Researchers must be transparent about the decisions made throughout their research.	N/A	Constructivism Interpretivism
3.19 Sites as arbiters of quality (Lincoln, 1995; Savage, 1988)	Researchers must pursue a situated and complex knowledge that embraces ways of knowing beyond their own in order to assist marginalized populations.	• Communitarian (Denzin, 2006) • Community as arbiter of quality (Boughton, 1997)	Constructivism Critical paradigms Interpretivism
3.20 Transgressive validities [ironic, rhizomatic, paralogical, and voluptuous validity] (Lather, 1993)	Researchers must reject traditional positivist conceptualizations of validity in order to proliferate form (ironic validity), understand how their participants exist within systemic networks (rhizomatic validity), foster differences and embrace contradictions (paralogical validity), and obtaining an abundance of data (voluptuous validity).	N/A	Feminism Postmodernism Poststructuralism
3.21 Triangulation (Denzin, 1978)	Researchers must utilize multiple data sources to obtain answers to their research question(s).	N/A	Constructivism Interpretivism Postpositivism
3.22 Voice (LeCompte, 1993; Lincoln, 1995)	Researchers must consider whose voice(s) are undervalued, and seek out those who are silenced within data collection.	• Polyvocality (Clifford, 1986)	Constructivist Critical theorist Interpretivist
3.23 Worthy topic (Guba and Lincoln, 1989; Tracy, 2010)	Researchers must choose topics that are significant to current literature, society, and personal events.	N/A	Postpositivist

Consistency

Scholars who choose to pursue consistency argue the importance of understanding the literature contextualizing your topic and sharing with readers a fluid story about those connections through the presentation of your findings. These scholars seek to make connections throughout the study's report, always linking back to the research purpose, as the purpose should drive all decisions that they make throughout the inquiry process (Noble and Smith, 2015; Lincoln and Guba, 1985), and ensure that they provide a rationale for these decisions that aligns with their research purpose. Further, Drisko (1997) suggested feedback sessions with the participants to confirm your interpretation of their data and the conclusions that you have drawn; this process is often referred to as member checking (described later in this list). It is critical to avoid being vague, leaving readers with too much of an opportunity to draw conclusions of their own. As the researcher, you are privy to the originally collected data and have the best sense (second to your participants!) of what to include in order to descriptively share findings with readers.

When considering consistency, it may be helpful to ask yourself:

- Are my research questions strong?
- Does each section within my writeup refer back to the purpose of my study?
- Do I make explicit connections between my research questions, the literature, my methodological decisions, and my findings?
- Does this clearly connect to the purpose of the study?
 - If your answer is "no," or unclear, you should approach your project in another way in order to maintain consistent alignment throughout the report.
 - If the answer is "yes," have a colleague read over your work to check for the consistency you aimed to maintain throughout the paper.

Credibility

Readers of qualitative research should trust your findings are credible. Credibility refers to your ability to use your data as proof of findings (Patton, 1999). When a researcher is credible, they present enough evidence regarding the data they have collected, analyzed, and shared to draw their conclusion. Researchers must carefully extrapolate pieces of data that best represent their conclusions while maintaining the nature of the participants' words. It is critical that the researcher draws conclusions from the data and *shows* rather than *tells* the readers how those conclusions were made, and that those conclusions are believable to the reader based on the evidence presented (Tracy, 1995).

To increase the credibility of your conclusions, participate in data-checking techniques such as triangulation, member checking, polyvocality, and reflexivity (see Tracy, 1995), which are detailed in this chapter. A common theme among these techniques is including your participants in the process of drawing conclusions based on the data. Once you have made conclusions about your findings, take time to ensure that (1) the information you shared was representative of participant experiences, (2) you included multiple voices throughout your findings (when appropriate), and (3) you have addressed various levels of relationship power dynamics. Further, you should indicate which credibility techniques you used to help your readers understand the quality of your findings.

When considering credibility, it may be helpful to ask yourself:

- Have I participated in one or more data-checking techniques?
- Have I described these techniques in my written report?

Critical subjectivity

Critical subjectivity involves the process of reflexively deconstructing our own beliefs, biases, subjectivities, experiences, identities, and more, in order to understand our core beliefs and then share those beliefs with our readers. Lincoln (1995) argued that "such reflexivity or subjectivity enables the researcher to begin to uncover dialectic relationships, array and discuss contradictions within the stories being recorded, and move with research participants toward action" (p. 283). It is important to understand not only the beliefs, experiences, and perspectives that each of us brings to our research but also how those beliefs affect our actions throughout the inquiry process.

In order to participate in critical subjectivity, you must first critically examine the lenses that you bring to a study. As mentioned above, these should include your own beliefs, biases, subjectivities, experiences, identities, and anything else relevant to the topic that you are studying. You should engage with this practice throughout the inquiry process, and include these considerations within your final written research report so that readers can understand your process and understand how you approached the study.

When considering critical subjectivity, it may be helpful to ask yourself:

- How do your beliefs about the creation of knowledge impact the way that you design your study?
- Which forms of evidence do you feel are worthy?
- How may your shared and unshared identities affect the way that you interact with participants?
- How may your beliefs affect the way that you analyze the data, and even which findings seem to jump off of the page to you?
- How might your lived experience shroud other findings from your immediate view?

Meaningful coherence

Scholars who wish to claim meaningful coherence should ensure that their research reports demonstrate significant alignment throughout the introduction, literature review, methodology, results, and discussion sections (Tracy, 2010). More specifically, authors should confirm that (1) there is a clear purpose statement at the onset of the research paper and that the research study achieves that stated purpose, (2) the literature review provides context and informs the research questions, (3) the methods address the stated research question(s) and are grounded in appropriate theoretical and paradigmatic approaches, and (4) the discussion section speaks to and addresses concerns brought to light in the literature review. Ultimately, a research study that is meaningfully coherent achieves the stated purpose and leaves readers feeling fulfilled in the pursuit of that purpose.

When considering meaningful coherence, it may be helpful to ask yourself:

- Have I ensured that I tell a single, cogent story of my study from beginning to end?
- Have I considered how other scholars from the existing literature may impact how I frame my study?
- Have I chosen a methodology that is best suited to answer my research questions?
- Are my research questions grounded in my theoretical and paradigmatic commitments?

Member checking

Researchers who choose to utilize member checking most often tend to share their findings with participants in order to determine "whether the participants recognize them as true or accurate" (Lindlof and Taylor, 2002, p. 242). This can be an incredibly useful

process, as participating in member checking can serve not only as validation of your interpretations, but also as new data or points of view by which to view your study (Tracy, 2010). This process also allows us to share power with the participants by allowing them to control how their stories are presented. We often participate in member checking by sending participants two files; (1) one with the full written findings section, so that they may see their data within the overall context of the study, and (2) another with just their portions pulled out in case that is easier for the participant to interact with.

When considering member checking, it may be helpful to ask yourself:

- Have I sought feedback from my participants about how I presented their data?
- Did I adjust or revise any portions of my findings that participants did not agree with?
- If there are revisions that I have chosen not to make, have I acknowledged these tensions within my written report?

Memoing

Many qualitative researchers utilize memoing throughout their study design, data collection, and analysis to keep a record of their thoughts, feelings, emotional responses, and questions that arise throughout the inquiry process (Birks et al., 2008). Researchers use memoing as a way to save relevant information about the data that may not be apparent or can be easily forgotten at a later time (Glaser, 1978). Further, by keeping records of decisions and meanings constructed from the data, the researcher can enhance their credibility.

You can keep a notepad of your continued thoughts and questions through each step of the research process. In fact, you can utilize memoing as a side conversation that you have with yourself or among your research team as you analyze your data. While you rarely include your notes within your final research publication, memoing can provide support and enhance your understanding of the research processes, allowing you to understand the data beyond the surface level.

When considering memoing, it may be helpful to ask yourself:

- How does the data make you feel?
- What questions do you still have for your participants or research team?
- Does your data stir any of your emotions or memories of personal experiences?

Positionality

Qualitative researchers often elaborate on positionality within their published work by reflecting upon their identities in relation to their study's participants and data. These scholars often do so because of the epistemological belief that we will never truly, completely, and fully understand truth outside of the context from which it was created. For example, Lincoln (1995) argued that "we can deduce that texts that claim whole and complete truth or that claim to present universal, grant metanarrative, or generalizable knowledge ... are themselves spacious, inauthentic, and misleading" (p. 280). Understanding how a researcher positions themselves within their work is imperative to grasp the meaning and quality of their work. Scholars who support positionality as an indicator of quality argue that appearing detached and claiming objectivity in your work detracts from its quality, and in fact places barriers on how much quality you can truly claim.

In order to utilize positionality within your own work, you must "come clean" (Lincoln, 1995, p. 280) about your own beliefs, identities, and paradigmatic commitments. In order for readers to understand how you have interpreted the data, they must first understand the beliefs that you have about the nature of truth and knowledge. You must also be willing to share your subjective, social, and cultural positions within the larger context of your work.

When considering positionality, it may be helpful to ask yourself:

- What identities do I have in relation to this study?
- How do those identities and beliefs affect the way that I approach and interact with my participants, design, analysis, and which findings I have chosen to report?
- Have I extended my engagement with my own identities beyond simply listing them and critically examined how they affect each aspect of the inquiry process?

Reciprocity

Scholars who embrace the idea of reciprocity argue that we must consider the bonds created between researcher and participant (Lincoln, 1995; McLaughlin and Tierney, 1993). That is to say that each individual participating in the research (which includes us as researchers!) must establish a relationship with each other "marked by a deep sense of trust, caring, and mutuality" (Lincoln, 1995, p. 284). This process can be considered one of purposeful, intense, deep sharing that "opens all lives party to the inquiry to examination" (Lincoln, 1995, pp. 283–284). Those who advocate for reciprocity as an indicator of quality ask how we can share ourselves in relation to the ways that our participants have shared themselves with us.

Indeed, the idea of reciprocity infers that we must be open with ourselves, and our experiences, as researchers both within our interactions with participants as well as throughout our published reports. We often task our participants with sharing their deepest

thoughts, beliefs, and experiences—we are seeking their vulnerability—however, we do not often share ourselves in those spaces as well. As authors of this chapter, we often feel that this ideology aligns closely with Roulston's (2010) conceptualization of romantic interviewing; we share of ourselves to build trust with the participant.

When considering reciprocity, it may be helpful to ask yourself:

- What parts of myself am I comfortable sharing with my participants?
- How can I make the interview space feel more like a conversation in which we both share about ourselves?

Reflexivity

Scholars who engage in reflexivity reflect on their own subjectivities and positionalities, deconstructing how those beliefs and identities shape and are shaped by each aspect of the research process. Reflexivity is the process of turning inward in an effort to recognize and take responsibility for the various ways in which one is situated within the research. Researchers should reflect upon their paradigmatic commitments, their relationship to the study participants, and their beliefs about the research topic.

Some questions to begin reflecting on include: What is your philosophy as a researcher? What is your relationship to the research topic? Answering these questions at the onset of the study allows you to understand how you think about and handle your own influence on the study and to recognize how your own subjectivities and positionalities may impact the creation of research questions. During the data collection process, engaging in reflexivity may include self-reflection in which you seek to understand your emotions and reactions during your interview. By engaging in this self-reflection, you may become attuned to the fact that you have emphasized certain things while shying away from others. During the final write-up, reflexivity may take the form of conscious editing, wherein you fully engage with the data and remain aware of the ways in which your sensitivities can drive what you deem as important. Some useful tools to practice reflexivity include, but are not limited to journaling, self-interviewing, talking with colleagues/peers/friends, researcher memos, video diaries, and poetry/art/music.

When considering reflexivity, it may be helpful to ask yourself the following questions based upon which step of the research process you are currently reflecting:

- Step 1: Finding your research question.
 - Why is this study important to you? Or why did you choose to pursue this study?
 - What do you already know about your topic?
 - What have you experienced?
 - What do you believe about it?
- Step 2: Reviewing the literature.
 - How do previous studies affect the way that you think about your research?
 - How do these other studies relate to your own study?
- Step 3: Figuring out your methods.
 - What do you believe about the nature of truth and reality?
 - What values do you hold related to the inquiry process?
 - How do these beliefs and values align with various research paradigms?
 - How do you feel about your methodological choices?
- Step 4: Collecting data.
 - What personal identities do you have that are related to this study?
 - Which identities do you share with your participants?
 - Which identities do you NOT share with your participants?
 - How do these common and different identities affect the way in which participants will interact with you? Or what kind of data they will choose to share or not share?
- Step 5: Analyzing data.
 - Will your identities lead you to look for specific things?
 - Which findings stand out to you more due to your own experiences?
- Step 6: Writing your findings.
 - How do you choose which of your findings to privilege over others?
 - How will you order your findings?
 - How will you practice transparency concerning your ethical concerns and reflexivity?

Relational ethics

A scholar who utilizes relational ethics “recognizes and values mutual respect, dignity, and connectedness between researcher and researched, and between researchers and the communities in which they live and work” (Adams et al., 2015). Qualitative researchers often form interpersonal bonds with their participants, whether purposefully or organically. When engaging in relational ethics, the researcher takes into consideration those interpersonal bonds and makes decisions in ways that seek to protect them. A researcher focused on relational ethics is often more concerned with creating mutual respect and engaging with their participants than they are

about taking advantage of their participants in an effort to churn out good research. In fact, they often have a sense of reciprocity in which the ultimate goal is to care for, respect, and do justice to and for their participants.

Your consideration of relational ethics may take a variety of forms depending on the type of research that you are engaging in and the ultimate goal of that research. However, all researchers incorporating relational ethics have one shared goal: that decisions be made within the context of the relationships being formed.

When considering relational ethics, it may be helpful to ask yourself:

- Have I considered the relationships I formed with participants while engaging in ethical reasoning?
- Have I ensured that I will not harm the growing mutual respect and care that I have built with my participants?
- Have I invited participants to collaborate during data analysis or within the writing of my findings?

Resonance

Researchers who seek to ensure that their study contains resonance seek to influence readers through evocative representations of the data. The researchers will use their participants' words to create a clear understanding of findings that are relevant and transferable to other areas of study and life (Tracy, 2010). Scholars who are particularly adept at utilizing resonance can often present their findings in such a way as to relate to readers from varied backgrounds.

If you choose to seek resonance within your work, make sure that you write in a way that is clear and concise so that a reader who is not an expert within your field can understand and connect to your study. Rather than leaving a reader to decide how your findings are meaningful, paint a vivid picture using explicit examples pulled from the data to represent your claims.

When considering resonance, it may be helpful to ask yourself:

- Have a *shown* instead of *told* the readers what my findings are by including many direct quotes from participants?
- Have I chosen quotes that are particularly emotional or powerful?
- Have I written my report in such a way that others outside of my field can understand my work?

Rich rigor

A scholar who seeks to claim that they are using rich rigor within their research ensures that they have provided a detailed, clear research design and methodology. From the beginning of the research design, the researcher identifies the purpose of the study and notes every decision made in alignment with the purpose of the study. The author's published methodological process is detailed, and the researcher has collected enough data to convincingly make conclusions (Golafshani, 2003). This is often achieved through data saturation, which refers to the process of collecting data until you no longer encounter new perspectives among your participants.

When considering rich rigor, it may be helpful to ask yourself:

- Did I clearly describe each step that I took in designing my study? In collecting data? In analyzing data?
- Have I presented enough data in my writing to convince readers that the conclusions I have drawn are meaningful and appropriate?
- How can I include other data sources to support my analyses?
- How have my methodological decisions contributed to my ability to create a data set that adequately represents the participants?

Sacredness

Feminist researchers have long argued for the concept of sacredness to be at the center of high-quality research (Lincoln, 1995; Schaefer, 1981). Lincoln (1995) described this concept as concerning researchers with "a profound concern for human dignity, justice, and interpersonal respect ... the collaborative and egalitarian aspects of the relationships created in the research-to-action continuum" (p. 284). Scholars who choose to embrace sacredness often seek to create relationships with participants based on mutual respect between researcher and participant and deep respect for each participant's personal dignity. Lincoln (1995) further argued that sacredness should be used to further explore the connection between the people with whom we are conducting research (fellow researchers, participants, gatekeepers, etc.) and ecological concerns. She additionally argued that if we do not hold sacred our very Earth and how our work affects it, we destroy not only our physical environment but also the human spirit by which we study.

In embracing the concept of sacredness we must choose to design a study that allows us to interact with participants in a way that respects them as unique individuals. For example, we often see this enacted as consistently checking in with participants to ensure that they are comfortable continuing to share their perspectives through multiple interviews, focus groups, observations, or phases within a study, comfortable with how we are portraying them in our final published reports (this can be achieved through member checking), and allowing them to dictate how data collection interactions occur. For example, you may choose unstructured or semi-structured interviews in order to allow your participants to take you on a journey of their choosing instead of sticking to a strict list of questions.

When considering sacredness, it may be helpful to ask yourself:

- What can I do within the interview/focus group/observation to make participants comfortable?
- Have I checked in with each participant to make sure that they are comfortable scheduling a follow-up interview?
- Did I give participants a choice as to where they would like to meet for an interview?
- Have I allowed participants to help shape our conversation within the interview space?

Sharing privilege

In seeking to share privilege we must grapple with the question of *who owns the data*? Do we, as the authors, own the stories that are shared with us that we then published? Or are data always owned by the participants who have shared it with us? Or, better yet, are data owned by the university that houses us as we publish in seeking tenure? [Lincoln \(1995\)](#) argued that “the somewhat dark side of research hides the fact that most of our research is written for ourselves and our own consumption, and it earns us the dignity, respect, prestige, and economic power in our own worlds that those about whom we write frequently don’t have” (p. 285). Sharing privilege is all about interrogating the ways in which we, as researchers, benefit from the lived experiences and stories shared with us by our participants, and then enacting action in a way that shares that privilege or gives back to those who have loaned us their stories.

There are many practical ways in which sharing privilege has been enacted within research studies. For example, [Lincoln \(1995\)](#) shared stories of an author who shared royalties from her book with her participant, even going as far as to contribute to her participant’s mortgage payment when in need, and [Lather \(2001\)](#), who sought to publish alongside her participants in order to ensure that they are given as much recognition for the work as she was. Sharing privilege can look like sharing financial gain you may experience with participants, sharing authorship, allowing participants to share in the decision-making process about what is and is not published, where it is published, and so on. One of our authors, Leia, shared privilege by asking the participant in a single-person case study where they would prefer that Leia donate money that she was awarded after the publication of the article. Originally, Leia had offered to send the award money to the participant, who instead indicated that they did not wish to receive it.

When considering sharing privilege, it may be helpful to ask yourself:

- How would I feel about inviting participants to write with me in exchange for including their name on my final publication?
- How can I share my data, findings, or compensation with organizations or non-profits that tackle issues related to my research?

Significant contributions

Though we do not determine statistical significance in qualitative analysis, we can determine the significant contribution qualitative research has on expanding the literature on a given topic. To be considered a significant contribution, the research question must first address a gap in the literature, whether that be theoretical, regarding professional practices, or overall cultural or moral views ([Tracy, 2010](#)). Our contributions as qualitative researchers can be significant when we address gaps within our existing knowledge and utilize our data to narrow those gaps. Though it is impossible to conduct a perfect study, as long as the research question, methods, and data-driven findings relate to the gap in our existing knowledge, a scholar can argue that they have made a significant contribution to the literature.

When considering significant contributions, it may be helpful to ask yourself:

- Have I thoroughly explored the literature about my topic in order to discover any existing gaps?
- Have I designed my study in such a way as to close or narrow those gaps?
- Are my chosen methods the best choices for exploring these concepts?
- Have I explicitly connected my findings and how they fit within the larger context of existing literature?

Sincerity

Similar to authenticity, a researcher’s sincerity is considered a good measure of quality when they are transparent in the decision-making process of their study ([Seale, 1999](#)). Sincere researchers seek to paint a picture for their readers, explaining the rationale behind the methods selected, how they were conducted in a way that promotes replicability, and the challenges that arose from selecting these methods. When researchers are sincere in their work, they maintain a great deal of reflexivity and recognize personal bias toward their own study and how that may have impacted the decision-making ([Emerson et al., 1995](#)). See our section on *Reflexivity* to identify personal biases and assumptions you bring into your research.

When considering sincerity, it may be helpful to ask yourself:

- Have I thoroughly described and justified why my methods are the best way to answer my research question(s)?
- Did I provide enough information to allow readers to understand what I accomplished using simple terminology that is descriptive enough for another researcher to replicate the steps that I took?
- Have I situated my work among key scholars within my field?

- Have I detailed all of the unexpected challenges or limitations that I encountered during my study?
- Did I include a reflexivity statement in my final report and connect it to any literature that I used to support my methodological decisions?

Sites as arbiters of quality

Scholars who embrace this concept argue that “research takes place in, and is addressed to, a community; it is also accurately labeled because of the desire of those who discuss such research to have it serve the purposes of the community in which it was carried out, rather than simply serving the community of knowledge producers and policymakers” (Lincoln, 1995, p. 280). In other words, knowledge is relational and situated within the communities in which it is created. To seek to objectively understand the world is to work within anti-communal, individualistic, and self-serving ways. By seeking to, instead, pursue a situated and complex knowledge that incorporates other ways of knowing beyond your own is to embrace “a complex intellectual dimension, exercised for the sake of assisting the marginalized on a journey toward greater participation in common life” (Savage, 1988, p. 13). Scholars who utilize sites as arbiters of quality should also be sure to explain how they have supported the communities with whom they conduct research; you should share how you are serving them, instead of allowing readers to assume that your work has occurred in a vacuum in which you, as the researcher, entered a space and left with only the knowledge that others assisted you in producing. Lincoln (1995) argued that “research is first and foremost a community project, not a project of the academic disciplines alone (or even primarily)” (p. 282).

When considering sites as arbiters of quality, it may be helpful to ask yourself:

- Have I devoted sufficient space in my written report to thoroughly describe the community or population that I am serving?
- How can I give back to the populations or communities that have shared their experiences and perspectives with me?
- Have I shared how I gave back to the communities within my written report?

Transgressive validities

Lather (1993) described her goals in creating transgressive validity as a “reconceptualized validity that is grounded in theorizing our practice ... fragmenting and colliding both hegemonic and oppositional codes, [her] goal is to reinscribe validity in a way that uses the anti foundational problematic to loosen the master code of positivism that continues to so shape even postpositivism” (p. 674). This concept was meant to disrupt the foundational, traditional, post/positivist concepts of reliability and validity while also adding newer conceptualizations of validity that correspond to other paradigmatic views such as poststructuralism. She argued that validity must be “a space of constructed visibility of the practices of methodology and a space of the incitement to see, an apparatus for observing the staging of the poses of methodology, a site that ‘gives to be seen’ the unthought in our thought” (Lather, 1993, p. 676). She divided this concept into four framings, which she also refers to as “scandalous categories” (Lather, 1993, p. 683), that we describe below. Transgressive validities are enacted in different ways, depending on which form you pursue.

Ironic validity

Lather’s (1993) conceptualization of ironic validity, which she also refers to as “validity as simulacra” (p. 677) espouses “a strategy of ironic validity proliferates forms, recognizing that they are rhetorical and without foundation, post-epistemic, lacking in epistemological support” (p. 677). A simpler way to engage in this concept is to consider that we can never truly know the original, absolute, objective truth (if one does, in fact, exist), and the research that we create thereby acts as simulations, simulacra, recreations, or reprints of a real, objective reality. We are providing recreated representations of what we learned while collecting data, but what we are producing is simulations of the truth within our own research.

We have recreated Lather’s ironic validity questions (see Lather, 1993, p. 685) to consider when reviewing your own work or the work of others for ironic validity:

- Have I foregrounded the insufficiencies of language in the production of meaning-making?
- Have I presented the concept of absolute truth, itself, as a problem?
- Have I acknowledged how my representations of the phenomena I am studying may be problematic or not endemic to the larger population or current time?
- Have I shared tensions that I may have found within the data as another dimension of the data?

Rhizomatic validity

Lather (1993) described rhizomatic validity, which she also referred to as Derridean rigor, as often used by researchers in order to undermine, unsettle, and center the marginalized others within a system; in doing so, the researcher generates new locally-determined norms to destabilize oppressive order (Lather, 1993). Scholars who seek to utilize rhizomatic validity concern themselves with understanding related contexts and networks of ideas and people-centered around their topic and/or participant(s). In examining rhizomatic connections and networks within your data, deeply interrogate how systemic and structural forces are at play within the world of your participants.

When considering rhizomatic validity, it may be helpful to ask yourself:

- What policies are affecting participants' lives?
- What other people are involved in their experiences?
- How might you be able to explore and present those connections and contexts?

Paralogical validity

According to [Lather \(1993\)](#), the goal of paralogical validity "is to foster differences and let contradictions remain in tension, as opposed to the recuperation of the other into the same that is always imposed at the end of a traditional philosophy" (p. 679). With the intention to embrace complexity, the researcher "fosters heterogeneity, refusing closure" (p. 679) during the data collection and analysis process as they continue to "search for instabilities" (p. 680). Lather also uses the terms Lyotardian paralogy and neo-pragmatic validity to describe paralogical validity. Often, researchers utilizing paralogical validity will prioritize sharing the differences inherent in their data versus seeking to obtain saturation.

When considering paralogical validity, it may be helpful to ask yourself:

- Have I looked for conflicting perspectives within the data?
- Have I explored any tensions inherent within the data?

Voluptuous validity

Voluptuous validity, or situated validity, is the fourth of [Lather's \(1993\)](#) transgressive validities. Lather uses this term to describe an abundance of data—by collecting as much data as possible, we are able to view our data from everywhere, which is directly in contrast with objectivism's "view from nowhere" (p. 682). She argued that acknowledging that we will never know *everything* and that our data is therefore perpetually incomplete and tentative, allows for the space for "others to enter, the joining of partial voices" (p. 682). Researchers, therefore, engage in voluptuous validity by assuming that they are incapable of knowing everything about a situation, and must continuously engage in both their topic and reflexive practice. In order to practice voluptuous validity, we must continually reflect upon ourselves, our decisions, and our perspectives about our data with the understanding that we will never completely understand our topic and/or participant(s). We may seek to continuously gain more information about our data.

When considering voluptuous validity, it may be helpful to ask yourself:

- Have I exhausted avenues of obtaining data regarding my topic?
- Have I ensured that my findings are written in such a way as to not portray an absolute or complete understanding of my topic?

Triangulation

Triangulation is the practice of using multiple data sources to obtain various answers and points of view to your research question. Typically selected prior to recruitment, the researchers intentionally identify and select specific individuals they aim to be represented in the study or multiple data sources that can be used in tandem to create a situated truth among them. Participants or data sources may be of similar or opposite views to provide a variety of context and depth to answer the research questions from multiple lenses. The order in which participants contribute to the study may be determined based on relational power between the participants, by order of participant influence, or completely at random. It's critical to provide opportunities for representation from multiple sources and individuals. [Denzin \(1978\)](#) suggests that when used appropriately, triangulation is a method to increase the overall credibility of your work.

When considering triangulation, it may be helpful to ask yourself:

- How can I counter-balance the voice of one participant with the voice of another? Or the findings from one source with another?
- Have I spent time with each data source? Each participant? Each document?
- Have I been intentional with my recruitment methods in order to ensure representative inclusion among the sources selected to represent the population or phenomenon that I am studying?
- Have I provided a rationale for spending more time on one data source as compared to another?

Voice

Voice refers to the process of weaving multiple voices throughout your research. [Lincoln \(1995\)](#) argued that we must consider "to who speaks, for whom, to whom, for what purpose ... who speaks for those who do not have access to the corridors of knowledge or the venues of the academic disciplines?" (p. 282). Further, [LeCompte \(1993\)](#) argued that qualitative researchers must "seek out the silenced because their perspectives are often counter-hegemonic" (p. 10). We must seek to involve ourselves deeply with our study participants in order to both accurately reflect their voices and understand the contexts from which they speak in order to ensure that our work does not silence or further marginalize them. We also see voice enacted through polyvocality, which refers to the process of including multiple voices and interpretations of your work. Allowing alternate voices to speak through your work, and therefore presenting nuanced representations of thought—and perhaps thereby embracing the tension created by lack of consensus—shows the strength of your work, while also resisting silence, disengagement, and marginalization ([Lincoln, 1995](#)).

It is one of the authors, Leia's, biggest pet peeves when others say that they want their work to "give voice to others." This determinant of quality, along with our team, argues that others already have a voice—we simply must choose to elevate those voices

through our research. One of the easiest ways to enact voice within your work is to simply include more words from your participants than from yourself within your findings section. Choosing to allow your participants to speak to your readers through gratuitous use of quotes enables your readers to understand your participants' experiences through their own words, while your interpretations serve as the glue that connects the quotes together. While it is true that we, as researchers, are responsible for interpreting the data, utilizing *voice* by including quotes from participants allows us to demonstrate that our findings are meaningful. When pursuing polyvocality, we typically track how often we have quoted each participant within our studies, ensuring that each individual's thoughts are represented throughout the report.

When considering voice, it may be helpful to ask yourself:

- Have I included many quotes in order to demonstrate my findings?
- Have I ensured that one participant is not quoted substantially more than others?
- How can I weave the voices of all participants in such a way as to elevate all of their voices?

Worthy topic

Researchers who seek to claim that their topic is *worthy* must consider the significance of the study as it relates to today's literature, society, and personal events. They strive to extend what has been previously known within the literature to newer findings, demonstrating that the traditional findings are still relevant in today's world. Rather than repeating studies or continuously testing theories in the same way, a worthy topic is one that leaves readers intrigued by the work that has been produced (Guba and Lincoln, 1989). Researchers must consider new perspectives and applications of established knowledge in order to add to the field. We may get hung up on the easy topics, resources, and methodologies that are popular, but it is the researcher who strives for a worthy topic to guide their research study that continues to push the boundaries in their field.

When considering if your topic is worthy, it may be helpful to ask yourself:

- Have I considered how this research may significantly contribute to a meaningful, relevant, topic in today's world?
- If I am studying an already well-researched topic, how does my work maintain the interest of readers? How does it add to the ongoing conversation?

Conclusion

Scholars have long debated how to determine quality within qualitative research without consensus. Trainor and Graue (2013) argued that these difficulties result from (1) disagreement as to what makes *good* qualitative research, (2) methodological training and practice primarily embracing quantitative, postpositivist commitments, and (3) the fact that qualitative researchers often seek to understand culturally-mediated and -defined notions of what constitutes as evidence. In fact, researchers often struggle to agree on even the most basic of vocabulary; for example, while some scholars have argued that the notion of *validity*, which stems from traditional quantitative practices, is incompatible with qualitative research (Cho and Trent, 2020; Kvale, 1995; Ravitch and Carl, 2019; Trainor and Graue, 2013; Wolcott, 1990), others have argued for embracing the term in order to redefine it (Lather, 1993). Therefore, it is each scholar's responsibility to choose and justify the quality determinants that fit their theoretical, ontological, epistemological, and methodological commitments. While this level of freedom may seem overwhelming, we argue that it is this very reprieve from traditional postpositivist practices that gives qualitative research powerful impact.

A trip to the salad bar

In section two, we discussed one of our authors, Leia's, metaphor of designing a qualitative project as constructing the perfect salad at the salad bar. Each researcher engages in this process differently, and therefore, even scholars who choose to use the same quality determinants (ingredients) will likely construct their study (salad) differently. For example, Cain and Velasco (2021) described their data quality efforts as follows;

We utilized triangulation (Tracy, 2010), member checking (Tracy, 2010), and reflexivity (Tracy, 2010). Instead of pursuing triangulation in a positivist sense to validate the collected data (Creswell, 2012), we pursued triangulation as a method to "understand the multiple perspectives available" (Glesne, 2011, p. 47). Each of us as the researchers shared our perspectives about Gray's story, and we worked together to negotiate our meanings and understandings in order to build a more credible meaning from the stories provided. Further, we pursued member-checking (Lather, 1986) by sharing our findings with Gray in order to ensure that we were presenting their story correctly, and that we had correctly understood the meanings and connections that they had wanted to portray. Finally, we the researchers shared a reflexivity journal document in Google Drive where we could write about our own thoughts and processes throughout data collection and analysis (pp. 363–364).

In this article, Cain and Velasco (2021) combined triangulation (determinant 3.21), member checking (determinant 3.7), and reflexivity (determinant 3.11). This fit the study perfectly, as it was carried out under an interpretivist paradigm and allowed Cain

and Velasco to consistently check on interpretations made from the data both with each other and with the participant. Other researchers may have utilized these determinants in alternative ways, or combined them with other determinants.

Implications for researchers in education

In this chapter, we sought to explore the complexity of quality in qualitative research and listed 23 widely-used quality determinants. Un/fortunately, we do not anticipate a consensus being accomplished in the field; qualitative methodologists often resist uniformity, choosing instead to embrace the gray areas inherent in the inquiry process—though this lack of clarity can at times feel overwhelming for learners. Previously, scholars have tried to create a unified framework of quality within qualitative research (e.g., Tracy, 2010), but have encountered resistance to their efforts (e.g., Albusaidi, 2019; Gordon and Patterson, 2013).

When engaging in qualitative research, it is imperative for researchers to critically engage with the entanglements they create among theoretical, ontological, epistemological, and methodological decisions and commitments to ensure that they are presenting an argument in which all of the individual pieces align. We hope that this chapter provides a starting point for further learning and exploration of quality in qualitative research.

References

- Adams, T.E., Jones, S.H., Ellis, C., 2015. *Autoethnography*. Oxford University Press.
- Albusaidi, S.S., 2019. Critiquing a qualitative study using Tracy's Big-Tent criteria. *Res. Soc. Sci. Technol.* 4 (1), 105–122. <https://doi.org/10.46303/ressat.04.01.6>.
- Birks, M., Chapman, Y., Francis, K., 2008. Memoing in qualitative research: probing data and processes. *J. Res. Nurs.* 13 (1), 68–75.
- Boughton, D., 1997. Reconsidering issues of assessment and achievement standards in art education: NAEA "Studies" Lecture. *Stud. Art Educ.* 38 (4), 199–213. <https://doi.org/10.2307/1320520>.
- Cain, L.K., Velasco, J.C., 2021. Stranded at the intersection of gender, sexuality, and autism: gray's story. *Disabil. Soc.* 36 (3), 358–375.
- Cho, J., Trent, A., 2020. Evaluating qualitative research. In: Leavy, P. (Ed.), *The Oxford Handbook of Qualitative Research*, second ed. Oxford University Press, pp. 1094–1122.
- Clifford, J., 1986. Introduction: partial truths. In: Clifford, J., Marcus, G. (Eds.), *Writing Culture: The Poetics and Politics of Ethnography*. University of California Press, pp. 1–26.
- Corbin, J., Strauss, A., 2008. *Criteria for Evaluation*, third ed. *Basics of Qualitative Research*, pp. 297–312.
- Creswell, J.W., 2012. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, fourth ed. Pearson.
- Denzin, N.K., 1978. *Sociological Methods: A Sourcebook*, second ed. McGraw Hill.
- Denzin, N., 2006. The politics and ethics of performance pedagogy: toward a pedagogy of hope. In: Madison, D.S., Hamera, J. (Eds.), *The Sage Handbook of Performance Studies*. Sage, pp. 325–338.
- Drisko, J.W., 1997. Strengthening qualitative studies and reports: standards to promote academic integrity. *J. Soc. Work. Educ.* 33 (1), 185–197.
- Emerson, R.M., Fretz, R., Shaw, L., 1995. *Writing Ethnographic Fieldnotes*. University of Chicago Press.
- Frieden, S., 1989. Transformative subjectivity in the writings of Christa Wolf. In: *The Personal Narratives Group* (Ed.), *Interpreting Women's Lives: Feminist Theory and Personal Narratives*. Indiana University Press, pp. 172–188.
- Glass, M.R., 2015. Teaching critical reflexivity in short-term international field courses: practices and problems. *J. Geogr. High Educ.* 39 (4), 554–567. <https://doi.org/10.1080/03098265.2015.1084610>.
- Glasser, B.G., 1978. *Theoretical Sensitivity Advances in the Methodology of Grounded Theory*. Sociology Press.
- Glesne, C., 2011. *Becoming Qualitative Researchers: An Introduction*, fourth ed. Pearson.
- Golafshani, N., 2003. Understanding reliability and validity in qualitative research. *Qual. Rep.* 8, 597–606.
- Gordon, J., Patterson, J.A., 2013. Response to Tracy's under the "Big Tent": establishing universal criteria for evaluating qualitative research. *Qual. Inq.* 19 (9), 689–695. <https://doi.org/10.1177/1077800413500934>.
- Guba, E.G., Lincoln, Y.S., 1989. *Fourth Generation Evaluation*. Sage.
- Johnson, J.L., Adkins, D., Chauvin, S., 2020. A review of the quality indicators of rigor in qualitative research. *Am. J. Pharmaceut. Educ.* 84 (1), 7120–7146. <https://doi.org/10.5688/ajpe7120>.
- Kvale, S., 1995. The social construction of validity. *Qual. Inq.* 1 (1), 19–40.
- Lather, P., 1986. Issues of validity in openly ideological research: between a rock and a soft place. *Interchange* 17, 63–84.
- Lather, P., 1993. Fertile obsession: validity after poststructuralism. *Sociol. Q.* 34, 673–693.
- Lather, P., 2001. Validity as an incitement to discourse: qualitative research and the crisis of legitimation. In: Richardson, V. (Ed.), *Handbook of Research on Teaching*. American Educational Research Association, pp. 241–250.
- LeCompte, M.D., 1993. A framework for hearing silence: what does telling stories mean when we are supposed to be doing science? In: McLaughlin, D., Tieme, W.D. (Eds.), *Naming Silenced Lives*. Routledge, pp. 9–28.
- Lindlof, T.R., Taylor, B.C., 2002. *Qualitative Communication Research Methods*, second ed. Sage.
- Lincoln, Y.S., 1995. Emerging criteria for quality in qualitative and interpretive research. *Qual. Inq.* 1, 275–289.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage.
- Little, D., Green, D.A., 2022. Credibility in educational development: trustworthiness, expertise, and identification. *High Educ. Res. Dev.* 41 (3), 804–819. <https://doi.org/10.1080/07294360.2020.1871325>.
- McLaughlin, D.E., Tieme, W.G., 1993. *Naming Silenced Lives: Personal Narratives and Processes of Educational Change*. Routledge.
- Meijer, P.C., Verloop, N., Beijard, D., 2002. Multi-method triangulation in a qualitative study on teachers' practical knowledge: an attempt to increase internal validity. *Qual. Quantity* 36 (2), 145–167. <https://doi.org/10.1023/A:1014984232147>.
- Miles, M.B., Huberman, A.M., Saldaña, J., 2019. *Qualitative Data Analysis: An Expanded Sourcebook*, fourth ed. Sage.
- Nakata, Y., 2015. Insider-outsider perspective: revisiting the conceptual framework of research methodology in language teacher education. *Int. J. Res. Method Educ.* 38 (2), 166–183. <https://doi.org/10.1080/1743727X.2014.923835>.
- Nassaji, H., 2020. Good qualitative research. *Lang. Teach. Res.* 24 (4), 427–431. <https://doi.org/10.1177/1362168820941288>.
- Noble, H., Smith, J., 2015. Issues of validity and reliability in qualitative research. *Evid. Base Nurs.* 18 (2), 34–35.
- Onwuegbuzie, A.J., Leech, N.L., 2005. On becoming a pragmatic researcher: the importance of combining quantitative and qualitative research methodologies. *Int. J. Soc. Res. Methodol.* 8 (5), 375–387.
- Patton, M.Q., 1999. Enhancing the quality and credibility of qualitative analysis. *Health Serv. Res.* 34 (5 Pt 2), 1189.
- Peshkin, A., 1988. In search of subjectivity—one's own. *Educ. Res.* 17 (7), 17–21.

- Ravitch, S.M., Carl, N.M., 2019. *Qualitative Research: Bridging the Conceptual, Theoretical, and Methodological*. Sage Publications.
- Reid, W., 1994. Reframing the epistemological debate. In: Sherman, E., Reid, W. (Eds.), *Qualitative Research in Social Work*. Columbia University Press, pp. 464–481.
- Relano Pastor, A.M., 2011. Crossing language and identity as a critical border ethnographer in southern California. *J. Lang. Ident. Educ.* 10 (3), 186–205. <https://doi.org/10.1080/15348458.2011.585309>.
- Roulston, K., 2010. *Reflective Interviewing: A Guide to Theory and Practice*. Sage.
- Savage, M.C., 1988. Can ethnographic narrative be a neighborly act? *Anthropol. Educ. Q.* 19, 3–19.
- Schaefer, A.W., 1981. *Women's Reality: An Emerging Female System in the White Male Society*. Winston Press.
- Schutz, A., 1967. *The Phenomenology of the Social World*. Northwestern University Press.
- Schwandt, T.A., 1996. Farewell to criteriology. *Qual. Inq.* 2 (1), 58–72.
- Seale, C., 1999. Quality in qualitative research. *Qual. Inq.* 5 (4), 465–478.
- Tracy, K., 1995. Action-implicative discourse analysis. *J. Lang. Soc. Psychol.* 14 (1–2), 195–215.
- Tracy, S.J., 2010. Qualitative quality: eight "big-tent" criteria for excellent qualitative research. *Qual. Inq.* 16 (10), 837–851.
- Trainor, A., Graue, E. (Eds.), 2013. *Reviewing Qualitative Research in the Social Sciences*. Routledge.
- Wolcott, H.F., 1990. Making a study "more ethnographic". *J. Contemp. Ethnogr.* 19 (1), 44–72.
- Yilmaz, K., 2013. Comparison of quantitative and qualitative research traditions: epistemological, theoretical, and methodological differences. *Eur. J. Educ.* 48 (2), 311–325. <https://doi.org/10.1111/ejed.12014>.

Triangulation: uses, abuses and recent developments

Alessandro Pratesi, Department of Social and Political Sciences, University of Florence, Florence, Italy

© 2023 Elsevier Ltd. All rights reserved.

Triangulation: origins and definitions	351
Further developments in the literature on triangulation	352
Triangulating theories, data, methods and analytical strategies: an American study on the phenomenology of care	353
Triangulating theories, data, methods and investigators: innovative UK research on care-related technologies	356
Concluding remarks and final recommendations	358
References	359

Triangulation: origins and definitions

In its basic formulations, the term triangulation designates the use of more than one perspective—in terms of method, theory, observer or data source—in the study of social phenomena, although over time more emphasis has been placed on the use of multiple methods of investigation and data sources (Bryman, 2016; Maxwell, 2016). Its use is aimed at producing more accurate, thorough and objective representations of the research subject; in other words, it is aimed at reinforcing the validity and/or the credibility of our research findings. Originally used as a heuristic tool in the context of military and geographical sciences to indicate the method for locating the position of points in a field through the combined use of three perspectives (Smith, 1975; Cox and Hassard, 2005; Konecki, 2008; Abdalla et al., 2018), triangulation eventually came to be widely employed in the social sciences through the work of several scholars in different fields (Campbell and Fiske, 1959; Webb et al., 1966; Smith, 1975; Denzin, 1970, 1978, 2006, 2001, 2012). Indeed, not all uses of multiple theories, methods, observers or data sources are commonly seen as triangulation. Similarly to other scholars (such as Hammersley and Atkinson, 2007; Heale and Forbes, 2013), Schwandt (2015, p. 307) describes the “central point” of triangulation as “to examine a conclusion (assertion, claim, etc.) from more than one vantage point”. This stance is consistent both with the original definition of triangulation in surveying and navigation (the use of multiple observations to determine the location of a particular object) and Denzin’s subsequent definitions (Denzin, 1970, 1978).

Without explicitly using the term triangulation, Campbell and Fiske (1959) introduced the concept to psychology, so as to empirically test the results obtained using different quantitative techniques (Abdalla et al., 2018). They used it as a synonym for convergent validity, that is, the degree to which different measurements related to the same or similar theoretical constructs are correlated. A few years later, Campbell and colleagues (Webb et al., 1966) coined the term *triangulation*, referring to the use of multiple methods to analyze and interpret the same data set and obtain valid results, in other words, providing a close representation of the phenomenon under investigation. Smith (1975) used the term triangulation as a metaphor to describe a research strategy borrowed from navigational or military practices: in the same way as navigators use maps and multiple reference points to locate an object’s exact position, social researchers can improve the accuracy of their findings by collecting different types of data on the same phenomenon. Finally, in 1978, Norman Denzin defined the process of triangulation as “the combination of methodologies in the study of the same phenomenon” (1978, p. 291). He described four types of triangulation: across different data sources, across different theoretical perspectives, across and among different researchers, and across data collection methods.

If disagreements on the initial use of triangulation within the social sciences seem to persist¹—for instance, according to Oppermann (2000), Webb et al. (1966) were among the first to introduce the term triangulation as a research approach in the social sciences, while Heesen et al. (2019) maintain that the first to advocate and apply “methodological triangulation” was the American sociologist Du Bois (1996/1899)—Denzin is unquestionably the scholar who has most widely employed, promoted and discussed triangulation in its comprehensive form and definition. In its early formulations, Denzin (1970, 1978) described triangulation as a combination of data, theories, investigators and methods to study the same phenomenon, whose aim is to overcome partial views and provide something very close to a comprehensive and reliable picture of the phenomenon.

Data triangulation refers to the collection of different sets of data in different periods of time and from different sources and places in order to obtain richer, deeper and more detailed accounts of the social phenomena under investigation. For example, a researcher interested in studying the changing nature of relationships, intimacies and families might collect several types of data (qualitative, quantitative or both) in different countries, from different subjects, at different points in time. *Theoretical triangulation* means using multiple theories or theoretical frameworks, sometimes from diverse disciplines or areas of knowledge, to interpret and analyze the same social phenomenon or set of data. For example, a researcher interested in studying the phenomenon of refugees and asylum seekers in Europe might use different theoretical frameworks drawing on theories on migration, citizenship, feminist theories, theories

¹See also Bryman (2007), Cho and Trent (2006), Creswell and Clark (2010), Ellingson (2009), Fielding (2012), Guba and Lincoln (2005), Moran-Ellis et al. (2006), Teddlie and Tashakkori (2011) and Yin (2015). According to Maxwell (2016), the use of the term triangulation was implicitly present even earlier, especially in the natural sciences.

on globalization, emotions, capitalism, and so on and so forth, from similar or different disciplinary perspectives (such as sociology, political sciences, economic sociology, anthropology, social geography, history, social psychology, etc.). *Investigator triangulation* involves the utilization of different researchers (an evaluation team) in the study of the same social phenomenon and the analysis of the same set of data. For example, a researcher who wants to study the long-term psycho-social effects of a global pandemic in a country might make use of a team of different researchers (possibly from different disciplinary fields) in order to achieve an in-depth, thorough interpretation of the data based on varied perspectives, considerations and analyses. *Methodological triangulation*, which is the most widely employed type of triangulation, refers to the use of multiple methods to study the same social phenomenon. For example, a researcher interested in studying the multiple implications and meanings of informal care with a phenomenological approach might make use of in-depth interviews, participant observation, questionnaires, diaries, video-recorded material, pre-existing databases, archives, etc. in order to obtain a more complete data set and a better understanding of the phenomenon.

According to Denzin's original definition (1978), the combination of different methods to study the same phenomenon and the convergence or agreement between the different methods increase the possibility that the research findings are not the result of a methodological artifact and reduce the shortcomings and limitations of a single method, thereby producing more credible conclusions. In other words, triangulation means analyzing the same phenomenon or research topic through more than one lens and from different angles and perspectives in order to shed new light on that phenomenon, reduce subjective and methodological biases, and increase the possibility of reaching more credible (or "valid") and replicable findings (Abdalla et al., 2018; Denzin and Lincoln, 2011; Flick, 2007; Maxwell, 2012; Patton, 2014). The usefulness of the four types of triangulation (data, theoretical, investigator and methodological triangulation) as a strategy for more thoroughness, comprehensiveness, accuracy and depth in any research is also reiterated in more recent publications (Denzin, 2006; Denzin and Lincoln, 2011). However, since its early formulations, different stances have been embraced and developed by different scholars with regard to the nature of triangulation (qualitative, quantitative methods or both?), its main goals (in terms of the validity, reliability or consistency of the results) and its relationship with different research paradigms (constructivism, positivism, postpositivism, and so on). So, let's now have a look at some of these more recent developments and debates in the literature on these subject matters.

Further developments in the literature on triangulation

Denzin himself highlights how "in its original forms (Denzin, 1970), triangulation referred only to the use of multiple forms of qualitative research methods, not the combination of quantitative and qualitative methods. It was also understood that each qualitative methodology—life story, case study, interviewing, ethnography, participant observation—rested on specific epistemological assumptions and each method had a complex disciplinary history. Hence, these interpretive methods could not be easily combined with one another" (Denzin, 2012, p. 82). In this article, not by chance entitled "Triangulation 2.0", Denzin (2012) revisits the multiple challenges of triangulation and proposes a sort of third way of conceptualizing mixed methods research based on critical interpretive methodologies. After illustrating some of the key controversies raised about mixed methods research—including disagreements over definitions, the real nature of a mixed methods study, debates on alleged irreconcilable differences between paradigms, contemporary preferences for postpositivism (Ellingson, 2009), and the added value of mixed methods—Denzin concludes by repeating what he thinks the "researcher-as-bricoleur-theorist" working between and within competing and overlapping perspectives and paradigms should do: act as a catalyst for positive social change through emancipatory action, and inspire readers and audiences toward this goal.

Can we retake the discourse surrounding this word [triangulation] and retrofit it to a postmodern world where meanings and politics are refracted off of the edges of crystals, not triangles? Can we train methodological bricoleurs who seek new third ways through and around obstacles to social justice in a neoliberal world? Can we create a discourse that does not play word games with itself, in methods-centric ways (Hesse-Biber, 2010)? Can we chart a way out of the present? Can we have a moratorium on mixed methods talk about designs and typologies and get back to the task at hand, which is changing the world?

Denzin (2012, p. 85).

For Denzin, triangulation is typically a qualitative tool. According to other scholars, on the other hand, triangulation involves the employment of both quantitative and qualitative methods (Konecki, 2008; Hammersley and Arkinson, 2007; Miles and Huberman, 1994; Moran et al., 2006; Silverman, 2001). Thus, the term triangulation assumes slightly or substantially different meanings depending on type and use. Konecki (2008) offers some methodological and epistemological considerations related to how to construct the "quality" of qualitative field research. The author clarifies how triangulation as a methodological tool is not only confined to positivist approaches, but can and should be used in field research in order to enhance the credibility, validity and objectivity of the research conclusions. His accent on a typically positivist lexicon, though, seems to be a bit at odds with his epistemological premises. In qualitative methodology, he claims, "the use of this procedure [triangulation] is often justified by the scientific requirements of gaining 'reliability' (it is an opportunity to repeat the observed behaviors together with the explanations, which means the accuracy of the 'measurement' or classification), validity (the answers to the question whether the explanations are suitable to present quantitative descriptions of people or situations), credibility (the degree of probability of occurrence of the observed phenomena), or the opportunity of generalizability" (Konecki, 2008, p. 23).

Heesen et al. (2019) compare and contrast methodological *pluralism* and methodological *purism* and define triangulation in terms of those methodological strategies used to strengthen the reliability of results obtained in the social sciences. Using examples taken from voting theory, the authors show how the use of multiple methods and convergence upon a single conclusion better supports the findings, increasing their validity and credibility. Their focus, therefore, is mostly on methodological triangulation, ignoring the other types. Reflecting on the origins of the term triangulation, its multiple connotations and its multiple potentialities, Opperman (2000), similarly to Denzin, highlights how the basic concepts underlying triangulation can be easily misunderstood. Often interpreted as an approach that combines both quantitative and qualitative data collection, triangulation—characterized by cross-linkages between and the systematic planning of these collection methods—is confused with multi-methodological approaches, in which there is limited cross-validation of the results.

Another common misconception is that the goal of triangulation is to achieve consistency across data sources or approaches (Guion et al., 2011; Patton, 2014). The belief that this research strategy provides convergence toward a single and coherent conclusion may be one of the most detrimental assumptions. Indeed, quite often the data produce inconsistent or even contradictory findings. Yet, these inconsistencies or contradictions should be seen as an opportunity to unravel the complexity of the phenomenon under inquiry and disclose a more in-depth understanding and interpretation of the data. Such an opportunity is clarified, among others, by Deacon et al. (1998), who illustrate an example of a multi-method investigation on the relationship between social scientists and the media in which the qualitative and quantitative findings appeared to contradict each other. Far from disconcerting or discouraging the authors, the clash between the two sets of findings allowed them to reflect upon and provide indications on how to explain and deal with these discrepancies. This case study shows the benefits as well the challenges of mixed-methods research. Accepting the internal validity of different sources of evidence and attempting to draw up alternative explanations induced the researchers to reassess their preliminary findings, while at the same time warning them that the issues may have been more complex than they had initially supposed and encouraging them to develop more convincing and robust explanations of the social processes being investigated.

Inconsistency and/or contradiction between data is actually explicitly posited by Mathison (1988) as one of the main purposes of triangulation. Her claim is that the researcher can analyze inconsistent or contradictory data to construct different, plausible explanations based upon multiple layers of evidence. Each of these layers can enable the researcher to produce richer, more in-depth descriptions and more credible explanations of the phenomenon being studied. Whichever conception of triangulation is employed in qualitative research, the end result is an increase in the *credibility* of the findings. The undisputable fact that triangulation adds depth to the data collected and analyzed is also reiterated by a recent publication (Fusch et al., 2018) which is specifically devoted to reconsidering triangulation in qualitative research. Criticizing scholars who, during the last decades, allegedly attempted to provide alternative approaches to triangulation but, in reality, merely proposed a restyled version of Denzin's approach, the authors discuss the continuing relevance of Denzin's approach to the triangulation of data, providing novice researchers and doctoral students with a series of useful recommendations on how to maximize its benefits. One of the best pieces of advice is that "researchers should identify the seminal (original) sources in their reading and critically read the seminal scholars' work to ensure that they are not creating more ambiguity. [...] Novice researchers should keep in mind that the triangulation of data can result in sometimes contradictory and inconsistent results; however, it is up to the researcher to make sense of them for the reader and to demonstrate the richness of the information gleaned from the data" (Fusch et al., 2018, p. 26).

To conclude this review of some of the main issues at stake when discussing the "used, abused and misinterpreted" term triangulation (Denzin, 2012, p. 85), it may be worth recalling a final, recent paper providing a systematic analysis of triangulation typologies and, through it, discuss the following dimensions in qualitative research: reliability, credibility, transferability and confirmability (Abdalla et al., 2018). Once again, this paper sets out to provide some guidance to researchers and this is why I mention it here. In particular, the authors illustrate some strategies on how to improve credibility in qualitative research and point to some advances in the discussions about the use of triangulation and its accomplishments in terms of research quality, essentially vis-à-vis rigor, scale, complexity, variety and depth in sociological research. It would be outside the scope of this chapter to comprehensively cover current and past debates on the uses and abuses of triangulation in the social sciences. Hence, in what follows, it may be useful to provide a couple of examples of triangulation, while making reference to Denzin's typology: the first example combines *theoretical, data and methodological triangulation*, and the second one includes an interesting and atypical example of *investigator triangulation*. Let us have a look at them both.

Triangulating theories, data, methods and analytical strategies: an American study on the phenomenology of care

The first example is based on a research project on family care, which I conducted a few years ago in Philadelphia (Pratesi, 2012, 2018). The overall aim of the research was to develop a 360-degree phenomenology of care by focusing on different kinds of care arrangements as they emerge in different types of family contexts and other forms of intimate relationships. Another aim of the research was to overcome the limitations of a traditional study on the gendered division of care work by focusing on a less investigated aspect of care: its *emotional implications* in terms of inclusion/exclusion. This in-depth, phenomenological analysis of the relational, interactive and emotional dimensions of informal care aimed to shed light upon the dynamics through which inequality is reproduced situationally, beyond the rigid and reifying categorizations of sex, sexuality and gender.

One of the research goals was to reflect upon the possibility of an embodied social theory of care; a theory aimed at reconciling the micro- and macro-aspects of social life by looking at the lived experience of care as an emotionally charged experience with important sociological and political implications. The analytic perspective I presented in this work provided a more detailed description of the person *who cares* and a broader phenomenological approach to the issues of care and care-related inequalities.

By looking at people's narratives about their *lived* and *felt* experience of care, this study offered a better understanding of the phenomenon of care, its multiple and not-so-visible purposes and implications, and, ultimately, its core nature.

Hence, the first type of triangulation I had to deal with was a *theoretical triangulation*: in practical terms, I had to handle at least three oceanic kinds of literature—on care, gender and emotions. Moreover, each of these fields of studies was further divided into several other areas and theoretical strands—the literature on care included the differences between formal and informal care, feminist theories on care, the ethics of care, etc.; gender included sexualities, feminist theories on gender and sexualities, theories on intersectionality, etc.; and the sociological theories on emotions are manifold. Therefore, it was not only necessary to use theoretical triangulation to come to terms with such complexity, but also to find a path to follow to prevent me losing my way. So, let's now have a look at what types of theories and theoretical frameworks allowed me to circumscribe and navigate the lived experience of care.

The theoretical background of this research considered the heuristic and epistemological potential of *phenomenology* for a critical understanding of *symbolic interactionism* and *ethnomethodology*. These traditions of inquiry seemed the most suited to obtaining an insight into the multiple ways through which emotional dynamics shape the caregivers' experience of care and produce different outcomes in terms of equality/inequality, inclusion/exclusion, belonging and entitlement to rights; in other words, in terms of citizenship. The *phenomenological perspective* allowed me to highlight the natural but not necessarily evident links between *symbolic interactionism*, *social constructionism* and *ethnomethodology*, the latter being more specifically identified with the approach of Sarah Fenstermaker and Candace West (2002). The authors conceptualize gender in terms of a fluid, contingent and situated accomplishment through which outcomes of inequality are constantly reproduced (Fenstermaker and West, 2002). They use the expression *doing gender* to suggest that the performative aspects of gender can be interpreted and understood as a way of *doing difference*. My argument is that a careful, phenomenological investigation of the processes involved in *doing care* can provide significant clues to interpret and visualize the concept of *doing citizenship*, in other words, the multiple ways in which inequalities and dynamics of inclusion/exclusion based on socially constructed categories result from ongoing interactional accomplishments. Care environments are places where dynamics of inclusion and exclusion are constantly produced; therefore, they are a strategic site for revealing the invisible, emotion-based interactional mechanisms through which social exclusion and inequality, on the one hand, but also social inclusion, status membership and citizenship, on the other, are constantly reproduced at the micro-level of analysis.

From the point of view of care theorists, the study drew on those approaches that suggest that examining care within specific historical and social contexts might be the most effective way to grasp a fuller understanding of its place and meaning in people's lives. *Care is a particular kind of work*, an activity directed toward identifying and meeting the needs or the well-being of certain others and challenging dichotomous thinking and dualisms which oppose *head* with *heart* and *rationality* with *emotion*. Hence, a different type of rationality is needed to understand care (Wærness, 1984), one compatible with but more substantive than the formal rationality by which orthodox economists analyze labor supply decisions. Care activities are different from but need to be integrated with other activities in both the economic and political spheres (Hochschild, 1983; Zelizer, 2005; Folbre and Nelson, 2000), and the increasing intertwining of *love* and *labor* calls for innovative research perspectives and theoretical formulations. From the point of view of the theories on emotions, the study drew on those approaches to the sociology of emotions which establish important links between the micro- and macro-levels of analysis by showing the emotional mechanisms through which the social bond, social behavior and social structures are interactionally and situationally reproduced. More specifically, Collins' *Interaction Ritual Chains* model was used (Collins, 2004).

In order to shed light on the links between the private, inner emotional processes related to care and their public, social and structural outcomes, I needed data collection methods that could capture and visualize the emotionally charged and lived experience of care in different types of caregivers; that is, methods that could be innovative, creative and reliable enough to allow me to claim the credibility of my findings and to induce the reader to endorse my view. Thus, from the methodological point of view, the research drew on a multi-method approach which involved: semi-structured in-depth interviews; weekly diaries on the emotional experience of care; participant observation; online discussion forums for members of parents' associations; ongoing conversations with the respondents beyond the interview context; key informant interviews; and secondary sources on informal care and parenthood collected from adoption agencies, local associations and online forums, journal and newspaper articles, and the web. Since I was the only researcher involved in the study, the entire data collection process was extremely time-consuming.

Being aimed at eliciting answers that could enable the emergence of the caregivers' "felt experience" of care and their stories, the interview format was extremely complex and included, among other research tools, a *thermometer of feelings*, an *existential ladder*, vignettes and several images of care (*photo elicitation*).² After the interview, out of all the respondents, a small sample accepted to

²The *thermometer of feelings* involved asking interviewees to qualify and weight a series of emotions related to their care responsibilities and activities. Interviewees were given a board with the drawing of a thermometer and twenty tags, representing twenty different emotional states, and then asked to locate the tags on a scale ranging from zero to one hundred while simultaneously commenting on their choices. Showing the interviewees a drawing of an *existential ladder* (a simple set of steps), I asked them to locate themselves on the ladder vis-à-vis their present, past and future situation, and then to describe what each of the steps represented for them. I then ask questions to understand whether and to what extent their care responsibilities affected their professional and existential careers in the past or present, and if there were alternative pathways they had to renounce. This apparently simple tool allowed me to collect a diachronic, multidimensional and holographic picture of their biographical pathways instead of a mere snapshot of their present situation. Moreover, it allowed me to get in-depth, fine-grained insights not only into the interviewees' existential and professional trajectories, but also into their inner emotional dynamics and motivations. *Photo elicitation* was another fundamental tool of the in-depth interview. At the end of the interview, I showed the participants a series of images of care in different contexts and situations (positive, neutral and negative) and asked them to choose those which best illustrated their *feelings* when thinking about their *lived* and *felt* experiences of care. These visual prompts stimulated rich accounts of the respondents' subjective interpretations and the emotional states they associated them with, and allowed me to access a deeper level of meaning and understanding. Through photo elicitation, the meanings of care became much more complex, at times even contradictory, but the result was a multi-layered, deeper ethnographic understanding of the phenomenon of care. For a more exhaustive description of these research tools, see Pratesi (2012, 2018).

keep a record of their daily care activities for three to four subsequent weeks. The *time sampling diary* (Brandstätter, 2001) is a research technique designed to gain insights into people's subjective experiences as they go about their daily activities, with a specific focus on the emotional dimension attached to such activities. The diary covered a 24 h period over two days (of choice) in three subsequent weeks and focused on the emotional states experienced by the respondents in different care contexts and moments. It focused on the type of care activity, time, place, other activities involved, their perception of resources/constraints and the eventual presence/absence of other persons. Even though the response rate for the diaries (22) was not sufficient for them to be attributed the same importance as the in-depth interviews in the analytical process, the richness and the quality of the data collected through this tool were invaluable.

The other form of triangulation I used in this study was *data triangulation*, to be intended here in terms of the collection of different sets of data, in different periods of time, and from different sources and places. Overall, the study involved meeting, interviewing and observing 80 informal caregivers, of whom 38 self-defined as heterosexual and 42 self-defined as gay and lesbian. The empirical part of the study lasted approximately two years, developing mostly between the winter of 2005 and the end of summer 2007, and involved several sources of data collected over time in different contexts. With the exception of a single father living in New York City, all the respondents lived in the urban and suburban areas of Philadelphia. The purposive sample included a variety of caregivers: partnered/married and single caregivers, male and female caregivers, heterosexual and gay/lesbian caregivers. On the other hand, the sample was relatively homogenous in terms of social class and type of care. The principal type of care involved in my study was childcare, but the sample included some cases of elderly care as the goal was to get a broad, exhaustive picture of the care phenomenon.

To construct the sample of married, cohabiting and single caregivers, I started from two local children's centers (a day care and a parent-infant center), informal networks, colleagues and two local LGBT associations. Getting in touch with gay/lesbian caregivers was more than difficult than anticipated; I therefore started to use several data collection channels involving: placing flyers in local schools, LGBT centers and activist organizations, gay-friendly cooperatives and a local gay bookstore; sending announcements via email through the mailing list of a local association of gay and lesbian parents and other local organizations; attending parents' events; creating new networks also through a yoga center. The purposive sample was eventually enlarged through snowball sampling for all the different kinds of caregivers. Secondary material and gray literature on public and private childcare centers, local organizations dealing with informal care (childcare and elderly care), local schools and associations, representatives of LGBT associations, LGBT parents associations, etc. were also collected together with many other sources of information which helped me to piece together a wide-ranging interpretive puzzle of the phenomenon of care with its multiple implications.

In terms of the analytical process that I used to interpret my data, I might perhaps mention a final form of triangulation which I shall call *analytical triangulation*. Indeed, the analysis of the substantial amount of qualitative data collected during the two years was guided by what Denzin calls *interpretive interactionism* (Denzin, 2001a) and what other scholars call *interpretive phenomenological analysis* (Smith, 2004). This analytical approach involves trying to understand the individuals' life experiences, what meanings individuals attribute to their experiences and how they make sense of them (Smith, 2004). These meanings were captured through the analysis of the interactions between the caregiver, the person cared for and a whole set of generalized others. Consequently, the analytical approach I adopted can be defined as *phenomenological*, *interactionist* and *interpretative*, in that it is concerned with the individuals' perceptions or accounts of a social phenomenon, focuses on people's subjective experiences and interpretations of the world, and views the analytical outcome as resulting from the interactions between the individuals' accounts and the researcher's frameworks of meaning.

The combined use of a *constant comparison* and *analytic induction* enabled the ongoing and systematic emergence of hypotheses directly from the data. Thus, although the research had a solid background based on a complex theoretical triangulation, as I mentioned, an emotion-based, micro-situated and relational theory of care was developed and deeply grounded in the qualitative data. As the data collection and analysis advanced, central theoretical concepts and emerging themes were identified and links were developed between those concepts and the collection of new qualitative data. The early stages of empirical research, especially when it also involves ethnographic work and observation, tend to be quite open, laborious and time-consuming, but also extremely productive. The subsequent stages of the research involve the constant corroboration, comparison and systematization of the emerging ideas, concepts, hypotheses and patterns. The whole process never really ends and could continue indefinitely: essentially, it ends when the researcher decides to or has to end it because of time and resource constraints.

Among other things, the research findings showed how same-sex families distanced themselves from hetero-normative definitions of family while simultaneously avoiding marginalizing definitions of "*parental citizenship*" based on hegemonic and heterosexist conceptions of sexuality and parenthood. Emotions make up the (missing) link between *doing care* at the micro-level of interactions and *doing or undoing difference* at the macro-level of social structures, beyond the rigid and reifying categorizations of sex, sexuality and gender; and different ways to *do care* and to *do gender* must be taken into account if we want to grasp a truly comprehensive and panoramic picture of the phenomenon of care and its *whatness*. It was one of the aims of the study (Pratesi, 2018) to highlight "the centrality of emotions to routine operations of social interaction" (Barbalet, 2001, p. 3) and their explanatory role in producing social change. Indeed, it revealed and clarified the emotional dynamics through which forms of relational social inclusion are produced at the micro-level of interactions even when forms of inequality and social exclusion persist at the

macro-level of social structures. More specifically, it clarified the interactional, emotional dynamics connecting parenthood, sexual orientation and citizenship and the complex ways they are shaped by and shape the structures in which they are embedded, producing social change. Ultimately, it clarified the extent to which empirical research can create contexts of learning and growth where the pleasure of knowledge is never merely cognitive or intellectual. By *doing care* we also *do citizenship, social inclusion and social change*: through emotions.

Triangulating theories, data, methods and investigators: innovative UK research on care-related technologies

Still related to “care”, but from a totally different perspective, the second example of triangulation concerns a pioneering project (2009–2011) which was conducted in the context of the UK government’s “ageing-in-place” policy aimed at enabling older people to live independently and safely within their homes and communities thanks to the use of smart, innovative technological solutions (Pratesi et al., 2012, 2013). The research aims and objectives included: identification of older people’s technological, psychological and social needs; definition of user requirements to inform an activity monitoring system for use in private homes and residential care settings; and analysis of the ways in which such systems impact on the everyday lives of older adults in different settings. The project was designed to develop an intelligent activity-monitoring system to support older and/or disabled people’s independence, safety and quality of life by charting activity/inactivity patterns in their living environments. What made this system “smart” and particularly innovative in comparison with other activity monitoring systems was its capability to learn patterns of behavior and over a period of time identify deviations from the norm. These deviations might indicate the presence of a problem and the necessity to intervene. Therefore, the innovative aspects of the system consisted of its capacity to go beyond mere safety and emergency issues (such as fall detection) and monitor both *health-* and *well-being-related* issues, through its ability to read and interpret *behavioral* deviations from the norm (such as not entering the kitchen). Being based on infra-red thermal images—which meant that neither the persons nor their specific activity could be identified—the system ensured complete respect of privacy. Nevertheless, at the same time, the system aimed to provide a much richer type of information than existing home monitoring systems that rely on sensors such as door switches and conventional movement detectors. Thus, for example, the system was also able to detect the presence or absence of other people in the house, the exact location of the person(s) within a specific room and if the person was sitting or lying down. In terms of design, it was intended to be similar to a smoke detector, therefore increasing privacy and avoiding the risk of potential stigmatizations connected to the use of assistive technologies.

The *theoretical triangulation* was based on extensive literature reviews on several disciplines and areas of study which included: aging and ageism; gerontology; gerontechnology; assistive technologies; telecare; user-driven, person-centered participatory research approaches; older adults as co-researchers; innovative research; behavioral symptoms of diseases; independent and safe aging at home, etc. The *methodological* and *data triangulation* were grounded on a number of different research tools, methods and sources of information, which included: secondary analysis of existing datasets from previous studies on aging and technology; face-to-face, in-depth interviews with older people who live alone at home; telephone interviews with professional stakeholders and health-care professionals; focus groups with older people and their formal and informal carers; workshops and quarterly meetings with the members of the advisory group; home trials of four to six months’ duration conducted on the prototype system; face-to-face interviews with the home trial participants and their informal carers; home trial participants’ diaries; vignettes and case studies.

The first phase of data collection yielded a considerable amount of qualitative data and resulted in the development of a series of use case scenarios³ which were designed to focus attention on the needs, motivations and perspectives of older people together with the potentialities and challenges of ambient-assisted living technology. On the basis of these qualitative data, a User Requirements Report fully grounded in the older people’s experience of the activity-monitoring technology was produced and a prototype infrared activity-monitoring system was developed by the team of technologists. Home trials aimed to assess the capacity, potentialities and acceptability of the activity-monitoring system within domestic settings. Data on the person’s activity in the house were collected via a number of infrared sensors and analyzed by the team of technologists. During the period of the home trials, pre- and post-home trial interviews were conducted with the participants and their informal carers and the participants were asked to keep a diary. The purpose of the diary was to provide a place for them to record information about their daily living activities. The information provided was then used to help improve the system and its response to events taking place in the home.

From the point of view of the *investigator triangulation*, we adopted a person-centered participatory approach to ensure that the needs, perspectives and preferences of older adults were taken into account throughout the process. The project was conducted in partnership with (1) a UK academic institution, represented by a research team of sociologists, (2) an industrial partner, represented by a team of engineers, and (3) an advisory group of older people, whose role was to provide systematic input into the research design and development, including data collection, analysis and dissemination. The advisory group was composed of six older adults from different cultural, social and occupational backgrounds. Participatory processes and/or user-driven approaches tend to include users (in this case, older people) simply as “test subjects” in the evaluation of already developed technologies. The innovative aspect of our more inclusive person-centered approach was the involvement of older people throughout the entire research

³Use case consists of a number of hypothetical scenarios, each representing specific situations and settings (domestic settings, institutional settings, etc.) and specific problems (a fall, a stroke, not being able to call for help, unusual, prolonged periods of activity or inactivity within the home, etc.), which were co-constructed with the advisory group of older adults and discussed by the three research teams. Each use case scenario was analyzed with the goal of figuring out both potential outcomes and strategies to deal with them.

process. To provide a sense of agency and equality in the research process, the members of the advisory group were considered co-researchers from the outset. As part of this participatory approach, they were actively involved in determining their level of participation and their role as co-researchers. All members of the advisory group agreed to review documents and relevant articles and provide input in the decision-making process. In addition, some expressed their desire to be involved in the data collection aspects of the research, creating the opportunity for the advisory group members to be directly involved in interviewing other older people and the potential to capture experiences which might not otherwise have emerged.

During the first year of the project, the members of the advisory group who opted for a more active involvement in the research process were provided with the necessary skills for interviewing and co-conducting focus groups via training workshops offered by the academic research team. Together with guidelines on how to conduct semi-structured, in-depth interviews, the training workshop provided advisory group members with the basic skills necessary to interview other older people within the community and to attend focus groups as co-researchers. Although the co-researchers felt that the training was stimulating and empowering, only three members of the advisory group felt that they had the necessary skills and confidence to be actively involved in the data collection and analysis. The benefits (and/or the shortcomings) of the participatory process were evidenced in the constant communication and interaction between the academic team and the advisory group of co-researchers throughout the project. From the outset, the members of the advisory group were provided with a broad range of possible forms of communication. This included: quarterly meetings with the research partners; face-to-face, Email and telephone contacts; the circulation and distribution of project-related reports, minutes and documents and information related to assistive technology, academic papers and journal articles, conference proceedings, seminars and outputs from other relevant events. As the research went on, these forms of communication were developed and became the primary means for transferring and sharing information between the academic team and the advisory group of older people. The members of the advisory group were encouraged to express and share their thoughts, feelings and opinions on events, meetings, research progress, status reports, latest news on assistive technologies, relevant websites, forthcoming conferences and articles, and to provide the academic team with regular feedback.

This approach facilitated the development of a continuous two-way dialog between the research team and the advisory group of older people, in a constant flow of information that was capitalized upon by the research team to reshape the scope of the research and refine its tools and methods. This process was enlightening for the research team, revealing issues that would not otherwise have been revealed, as the advisory group members were drawing on their own lived experiences. Their involvement provided crucial and critical insight into the potential impact of home monitoring technology on people's habits, behaviors and feelings; people's sense of privacy, autonomy, safety and independence; and finally, issues concerning personal identity, sense of belonging, the aging process and the symbolic meanings of home. In light of this continuous input, elements of both the sociological research process and technological development were reshaped or amended and the advisory group members reported that they felt empowered, more knowledgeable about activity-monitoring technology and able to have their voices heard and valued through this platform.

A key component of participatory research is including, acknowledging and making productive use of every participant's input (McCreadie and Tinker, 2005; Reed et al., 2004; Ross et al., 2005; Van den Hove, 2006; Warburton et al., 2009). As the project progressed, the comments and suggestions of the advisory group were incorporated into the research, with the expectation that the final outcome would be reflective of their experiences, values and needs. At this point, it is necessary to recognize and clearly state the trade-off between the system's technical capabilities (i.e., what *cannot* be designed) and requirement to incorporate user requirements into the system (i.e., what *can* be designed). This is as much about the users understanding the technical capabilities and limitations as the need to ensure that user requirements get put into the system design. The failure to consider this may well result in tokenistic and disingenuous participation and a system that is divorced from the everyday lives of older people.

The user-driven participatory approach we adopted also presented some challenges, and its management was not always successful. Participatory research is a continuous interactive process that must be constantly renewed and renegotiated. Undertaking participatory research requires the academic team to consider the requirements of all partners, co-researchers and stakeholders while at the same time maintaining high standards of research. This often requires working with each partner individually to reassert the benefits of taking this type of approach. Indeed, some of the main challenges we encountered in this project were maintaining high levels of collaboration, motivation and enthusiasm among the co-researchers while simultaneously facilitating the dialog between the different requirements of the advisory group of older people (co-researchers) and the team of technologists. At the beginning of the project, it was clear that the team of technologists were not necessarily very enthusiastic about involving the advisory group of older people in each and every stage of the project. On the other hand, the members of the advisory group were eager to be involved in the different levels of participation, that is, attending meetings, reviewing documentation, engaging in data collection and analysis, etc. As a result, flexibility became the byword in terms of role expectations while allowing the individuals themselves to determine and shape their level of involvement in the research. There was a fine line between ensuring that the older people made a positive contribution while sustaining their effort over the course of the research. A further important issue regarded the different work ethos, cultures and practices of the research partners involved. Whereas the academic research team had a consolidated tradition of conducting participatory research and a solid and grounded understanding of the ways in which participatory research must be conducted and whom it should benefit, for the technology partners the concept of participatory research was relatively unknown and alien to their traditional ways of thinking and working. Moreover, whereas the industrial partners were evidently and explicitly focused on the goal of producing a successful and marketable product, the academic team and the advisory group were concerned with the social, psychological and emotional effects of this type of care-related technological device and its potential impact on people's everyday lives. In this context, we assumed a bridging

role so as to mediate between the different research partners and ensure an authentic participatory process; nevertheless, the efficacy of this role was not always granted.

Regardless of these shortcomings, this example clearly shows how in certain contexts, *researcher triangulation* is not only valuable but necessary. The involvement of older people in research on ambient-assisted living technologies was key to achieving meaningful and ethically acceptable findings. Care-related technologies can create and expand opportunities for older people's empowerment, safety, independence and well-being. Nevertheless, they also involve important ethical issues concerning intimacy, privacy, power and control as well as the necessity to reflect on the psychological, emotional, social and cultural perspectives of all the social actors involved. Among other things, our findings showed how it was only by involving older people in the design and development processes of care-related technologies that we could attain productive and successful kinds of mutual inspiration between lay people and technologists. The effective inclusion of older people's perspectives, motivations and thoughts can be crucial in making care-related technologies more acceptable, useable and practical for them, but also in successfully achieving research aims and objectives. In order to be successful, the development of an activity/inactivity-monitoring system must be viewed less as a technological challenge and more as the creation of an integrated *socio-technical system* in which technology is meaningful and purposeful for all the people and organizations involved. In other words, the implementation of these technologies must be conducted in ways which are fully participatory and inclusive for older people, in order to ensure that the systems are designed both *for* and *with* them.

Concluding remarks and final recommendations

This chapter provides a discussion of the concept of *triangulation*, its early definitions and uses, and some of the most relevant theoretical and methodological controversies characterizing its development in the last decades. Even though extremely time-consuming, triangulation is a useful tool to use in qualitative (and quantitative) research so as to increase the validity, credibility and utility of our research findings. Together with this considerable benefit, triangulation can be useful to increase our confidence in data analysis and interpretation, facilitate our sociological imagination (Mills, 1959), foster innovative and creative ways to understand a phenomenon, reveal unexpected findings, challenge or integrate theories, and offer deeper and broader explanations of a phenomenon. Different scholars suggest different interpretations of triangulation, its advantages and potential, and its different types.

While recognizing the utility of further developments and integrations of this concept, this article considers Denzin's original typology an excellent starting point and orienting compass. Back in the 1970s, Denzin described triangulation as combining data, theories, investigators and methods to study the same phenomenon while aiming to overcome incomplete accounts and provide a comprehensive and reliable picture of the phenomenon. He described four types of triangulation: *data triangulation* (the collection of different sets of data in different periods of time and from different sources and places), *theoretical triangulation* (using multiple theories or theoretical frameworks, sometimes from diverse disciplines or areas of knowledge), *investigator triangulation* (the use of different researchers in the study) and *methodological triangulation* (the use of multiple methods to study the same phenomenon). Later on, when reassessing the multiple challenges of triangulation, the author suggests a sort of conceptual mediation aimed at reframing mixed methods research on the basis of critical interpretive methodologies (Denzin, 2012). Notwithstanding some of the key controversies and debates revolving around the nature of mixed methods research, Denzin reiterates his ontological and epistemological stance about the role of the "researcher-as-bricoleur-theorist" working between and within competing perspectives and acting as a catalyst for positive social change.

After having briefly outlined some of the most common uses, abuses and misconceptions of the term triangulation as well as some of the recent debates revolving around it, this article provides two examples of triangulation, making reference to Denzin's typology: the first one combining *theoretical*, *data* and *methodological triangulation*, and the second one including an example of *investigator triangulation*. Both examples refer to the use of triangulation in qualitative research and show how the ways we shape our research is reflected *in* and *by* the ways we shape our realities and how we construct our social worlds. By choosing one theory over another, one set of data over another, and one method over another, not only do we affect the research outcome, its solidity and credibility, but we also contribute to the construction of social reality. Our knowledge, our conceptions and our beliefs of what reality is—and this includes the products of our research—become embedded in the institutional fabric of society. This is why it is important to carefully, critically and reflexively ponder our theoretical and methodological choices, because, through such choices, we all—quite often unconsciously—help recreate dynamics of inclusion and exclusion.

The use of triangulation in qualitative research is a complex, time-consuming and onerous process. It requires the researcher to deal with the complexities and inconsistencies of research outcomes as well as to come to terms with the necessity to produce reliable, credible and solid results. Qualitative research relies upon a careful interpretation of themes, patterns, concepts, categories and narratives, while acknowledging and setting out potential biases and limitations in the context. Producing clear, coherent and convincing interpretations out of complex social realities is one of the problems shared by *all* researchers, not only qualitative ones. In the end, there is no research project that cannot be improved in its accuracy and credibility. As a consequence, we must be aware that what we find is always shaped by us (Marradi, 1988), that it is not the only truth that could be gathered from our data, and that the research methods require an ongoing critical and reflective consideration, the end products of which are never completely faultless. But equally, we should engage in designing research projects where what we construct cannot be seen as

arbitrary but as a truly *valid start* toward challenging conventional interpretations and expanding our current understanding of a phenomenon. The credibility of our findings can thus be solidly grounded both in the careful planning and implementation of the research design and in the multiple methods, theories, data, perspectives and creative tools we use to grasp the phenomenon we want to study. Regardless of what type of triangulation we employ and what level of credibility readers attach to our work, reflecting upon and using triangulation undeniably helps us to initiate a “dialogic relationship” with the “moral community” (Denzin, 2001, p. 43) that, as researchers, we all belong to.

Beyond the above-mentioned challenges and shortcomings, triangulation is an excellent alternative option for validating processes (Denzin and Lincoln, 2011; Denzin, 2012). We can use several strategies to increase the rigor, thoroughness and depth of our research, which include the collection of different theoretical and methodological perspectives, diversification of the empirical materials, and participation of various researchers—possibly with different perspectives—in a single study. This chapter aimed to illustrate some of the most recent debates revolving around the concept of triangulation and show examples of how to increase the quality of our data and interpretations as well as their credibility and reliability through the use of triangulation.

Any study based on a single-method theoretical approach or research technique is undoubtedly more inclined to errors and biases compared to other studies that use multiple ones. Together with the production of more exhaustive, valuable and reliable results, triangulation enables the researcher to dialog with both positivist and constructionist approaches. Ultimately, we should consider that the aim of any type of scientific knowledge is not merely to measure and assess data with the implicit assumption of discovering some universal truths, but rather to understand and grasp reality as something which is also (if not mostly) socially constructed. It is hoped that this chapter can help unravel current conceptual controversies on triangulation, bridge the gap between some of them and provide some suggestions for scholars and young researchers who want to advance existing knowledge and propel new findings in this direction.

References

- Abdalla, M.M., Oliveira, L.G.L., Azevedo, C.E.F., Gonzalez, R.K., 2018. Quality in qualitative organizational research: types of triangulation as a methodological alternative. *Administração: ensino e pesquisa* 19 (1), 66–98.
- Barbalet, J.M., 2001. Emotion, Social Theory, and Social Structure. A Macrosociological Approach. Cambridge University Press, New York.
- Brandstätter, H., 2001. Time sampling diary. An ecological approach to the study of emotions. In: Brandstätter, H., Elias, A. (Eds.), *Persons, Situations, and Emotions. An Ecological Approach*. Oxford University Press, New York, pp. 20–52.
- Bryman, A., 2007. Barriers to integrating quantitative and qualitative research. *J. Mix. Methods Res.* 1 (1), 8–22.
- Bryman, A., 2016. *Social Research Methods*. Oxford University Press.
- Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56 (2), 81.
- Cho, J., Trent, A., 2006. Validity in qualitative research revisited. *Qual. Res.* 6 (3), 319–340.
- Collins, R., 2004. *Interaction Ritual Chains*. Princeton University Press, Princeton, NJ; Oxford.
- Cox, J.W., Hassard, E.J., 2005. Triangulation in organizational research: a representation. *Organization* 12 (1), 109–133.
- Creswell, J.W., Clark, V.L.P., 2010. *Designing and Conducting Mixed Methods Research*, second ed. SAGE Publications, London.
- Deacon, D., Bryman, A., Fenton, N., 1998. Collision or collusion? A discussion and case study of the unplanned triangulation of quantitative and qualitative research methods. *Int. J. Soc. Res. Methodol.* 1 (1), 47–63.
- Denzin, N.K., 1970. *The Research Act*. Aldine, Chicago, IL.
- Denzin, N.K., 1978. *The Research Act: A Theoretical Introduction to Sociological Methods*, second ed. McGraw Hill, New York, NY.
- Denzin, N.K., 2001. The reflexive interview and a performative social science. *Qual. Res.* 1 (1), 23–46.
- Denzin, N., 2006. *Sociological Methods: A Sourcebook*, fifth ed. Aldine Transaction, New York, NY.
- Denzin, N.K., 2012. Triangulation 2.0. *J. Mix. Methods Res.* 6, 80–88. <https://doi.org/10.1177/1558689812437186>.
- Denzin, N.K., Lincoln, Y.S. (Eds.), 2011. *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA.
- Du Bois, W.E.B., 1996/1899. *The Philadelphia Negro: A Social Study*. University of Pennsylvania Press, Philadelphia.
- Ellingson, L.L., 2009. *Engaging Crystallization in Qualitative Research*. Sage, Thousand Oaks, CA.
- Fenstermaker, S., West, C., 2002. Doing difference revisited: problems, prospects, and the dialogue in feminist theory. In: Fenstermaker, S. (Ed.), *Doing Gender, Doing Difference: Inequality, Power, and Institutional Change*. Routledge, London, pp. 205–216.
- Fielding, N.G., 2012. Triangulation and mixed methods designs: data integration with new research technologies. *J. Mix. Methods Res.* 6, 124–136. <https://doi.org/10.1177/1558689812437101>.
- Flick, U., 2007. *Designing Qualitative Research*. Sage, London.
- Folbre, N., Nelson, J.A., 2000. For love or money—or both? *J. Econ. Perspect.* 14 (4), 123–140.
- Fusch, P., Fusch, G.E., Ness, L.R., 2018. Denzin’s paradigm shift: revisiting triangulation in qualitative research. *J. Soc. Change* 10 (1), 2.
- Guba, E., Lincoln, Y.S., 2005. Paradigmatic controversies, and emerging confluences. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, third ed. Sage, Thousand Oaks, CA, pp. 191–216.
- Guion, L.A., Diehl, D.C., McDonald, D., 2011. Triangulation: establishing the validity of qualitative studies. *Environ. Data Inf. Serv.* 2011 (8), 3.
- Hammersley, M., Atkinson, P., 2007. *Ethnography: Principles in Practice*, third ed. Routledge, London.
- Heale, R., Forbes, D., 2013. Understanding triangulation in research. *Evid. Base Nurs.* <https://doi.org/10.1136/eb-2013-101494>.
- Heesen, R., Bright, L.K., Zucker, A., 2019. Vindicating methodological triangulation. *Synthese* 196 (8), 3067–3081.
- Hesse-Biber, S.N., 2010. *Mixed Methods Research: Merging Theory With Method*. Guilford Press, New York, NY.
- Hochschild, A., 1983. *The Managed Heart. Commercialization of Human Feeling*. University of California Press, Berkeley.
- Konecki, K.T., 2008. Triangulation and dealing with the realness of qualitative research. *Qual. Sociol. Rev.* IV (3), 7–28.
- Marradi, A. (Ed.), 1988. *Costruire il dato: sulle tecniche di raccolta delle informazioni nelle scienze sociali*. Franco Angeli, Milano.
- Mathison, S., 1988. Why triangulate? *Educ. Res.* 17 (2), 13–17.
- Maxwell, J.A., 2012. *Qualitative Research Design: An Interactive Approach (Applied Social Research Methods)*, third ed. SAGE Publications, Thousand Oaks.
- Maxwell, J.A., 2016. Expanding the history and range of mixed methods research. *J. Mix. Methods Res.* 10 (1), 12–27. <https://doi.org/10.1177/1558689815571132>, 2016.

- McCreadie, C., Tinker, A., 2005. The acceptability of assistive technology to older people. *Ageing Soc.* 25, 91–110.
- Miles, B.M., Huberman, M.A., 1994. Data management and analysis methods. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 428–444.
- Mills, C.W., 1959. *The Sociological Imagination*. Oxford University Press, New York, NY.
- Moran-Ellis, J., Alexander, V.D., Cronin, A., Dickinson, M., Fielding, J., Sleney, J., Thomas, H., 2006. Triangulation and integration: processes, claims and implications. *Qual. Res.* 6 (1), 45–59.
- Oppermann, M., 2000. Triangulation—a methodological discussion. *Int. J. Tourism Res.* 2 (2), 141–145.
- Patton, M.Q., 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, fourth ed. SAGE Publications, Thousand Oaks.
- Pratesi, A., 2012. Exploring the emotional experience of same-sex parents by mixing creatively multiple qualitative methods. *Int. J. Qual. Methods* 11 (2), 82–101.
- Pratesi, A., 2018. *Doing Care, Doing Citizenship: Towards a Micro-situated and Emotion-Based Model of Social Inclusion*. Palgrave Macmillan/Springer, London. <https://doi.org/10.1007/978-3-319-63109-7>.
- Pratesi, A., Sixsmith, J., Woolrych, R., 2012. Participatory design for future care related technologies: lessons from the Smart Distress Monitor Project. In: Serrano-García, I., Jiménez, P., Resto-Olivo, J., Figueroa-Rodriguez, M. (Eds.), *International Community Psychology: Approaches to Contemporary Social Problems*, vol. II. Universidad Iberoamericana Puebla, pp. 183–196.
- Pratesi, A., Sixsmith, J., Woolrych, R., 2013. Genuine partnership and equitable research: working “with” older people for the development of a smart activity monitoring system. *Innov. J.* 19 (1), 1–17.
- Reed, J., Weiner, R., Cook, G., 2004. Partnership research with older people—moving towards making the rhetoric a reality. *Int. J. Older People Nurs.* 13a, 1–8.
- Ross, F., Donovan, S., Brearly, S., Victor, C., Cottey, M., Crowther, P., Clark, E., 2005. Involving older people in research: methodological issues. *Health Soc. Care Community* 13 (3), 268–275.
- Schwandt, T., 2015. *The SAGE Dictionary of Qualitative Inquiry*, fourth ed. Sage Publications, Thousand Oaks, CA.
- Silverman, D., 2001. *Interpreting Qualitative Data. Methods of Analysing Talk, Text and Interaction*. Sage, London, Thousand Oaks and New Delhi.
- Smith, H.W., 1975. *Strategies of Social Research: The Methodological Imagination*. Prentice Hall, Englewood Cliffs, NJ.
- Smith, J.A., 2004. Reflecting on the development of interpretative phenomenological analysis and its contribution to qualitative research in psychology. *Qual. Res. Psychol.* 1 (1), 39–54.
- Teddlie, C., Tashakkori, A., 2011. Mixed methods research: contemporary issues in an emerging field. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA, pp. 285–300.
- Van den Hove, S., 2006. Between consensus and compromise: acknowledging the negotiation dimension in participatory approaches. *Land Use Pol.* 23, 10–17.
- Warburton, J., Bartlett, H., Carroll, M., Kendig, H., 2009. Involving older people in community-based research: developing a guiding framework for researchers and community organisations. *Australas. J. Ageing* 28 (1), 41–45.
- Waerness, K., 1984. The rationality of care. *Econ. Ind. Democr.* 5, 185–211.
- Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, L., 1966. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Rand McNally, Chicago.
- Yin, R.K., 2015. *Qualitative Research From Start to Finish*, second ed. The Guilford Press.
- Zelizer, V.A., 2005. *The Purchase of Intimacy*. Princeton University Press, Princeton.

Mixed methods and multimethod research in education: six design decisions for collecting, combining, and developing datasets

Judith Schoonenboom, University of Vienna, Department of Education, Vienna, Austria

© 2023 Elsevier Ltd. All rights reserved.

Introduction	361
Multimethod and mixed methods research	362
The development of an integrated claim using multiple methods	363
The design decisions: an overview	364
Decisions related to collecting data	364
Decision 1: collecting qualitative and/or quantitative data	364
Decision 2: collecting specific types of data	365
Decisions 3 and 4: combining different types of data collection	365
Integrating outcomes of the analysis of separate datasets	365
To apply themes discovered in interviews for generalization in a quantitative study	366
To provide a quantitative overview as part of qualitative research	366
To further explore differences between quantitative groups or deviating individuals	366
To further explore any expected or unexpected quantitative results using qualitative data collection and analysis	366
Using quantitative data for development in a qualitative study (and vice versa)	367
To use quantitative data for sampling for qualitative data collection	367
To use qualitative data to illustrate quantitative findings	367
Decisions 5 and 6: further developing collected data	367
Quantitizing: converting qualitative data into quantitative data	367
The mixed dataset	368
The meta-inference: a real-life example of the development of an integrated claim in mixed methods research	368
Conclusion	369
References	370

Introduction

In some studies, researchers pursue one simple question. This can be a closed question for testing, such as “Does this drug perform better than a placebo?” Or it can be an open question for exploration, such as “Is there a Thyrotropin Releasing Factor?”, which was an important question in the biomedical sciences in the early 1960s (Latour and Woolgar, 1986, p. 147). In contrast, many educational researchers study complex problems in practice. They will investigate various aspects of the problem, considering it from different perspectives of participants such as schoolchildren, teachers, parents, and school leaders. Such educational studies pursue one or more complex questions, which result in one or more answers, inferences, or, to use the technical term of this chapter, *claims*, for each aspect and each different perspective. During such a complex study, those claims are connected into an ever-growing integrated claim, which at the end of the study results in the *final integrated claim*, in mixed methods research usually called the *meta-inference* (Creamer, 2018; Teddlie and Tashakkori, 2009).

One way in which different perspectives on a phenomenon can be obtained is by using multiple methods. Each method provides a unique perspective on the phenomenon; hence, each method can be used to answer a different research question. For instance, questions about teacher behavior can be answered by using video observations, whereas questions about their underlying reasons and motives can be answered through interviews (see Example 1 below). By combining the answers obtained through multiple methods, researchers can obtain a more detailed understanding than would be possible by using one method alone (Greene, 2007, 2015; Flick, 2017; Schoonenboom, 2019; Uprichard and Dawney, 2019).

This chapter discusses how a detailed understanding unfolds in a research study as a result of the various decisions that are made related to the use of multiple methods. Choosing to use multiple methods in a study is not born from one single decision at the beginning of the study, but instead a multi-faceted, multi-step decision that can be repeated throughout a study. The decision process contains three aspects, which together form the basis for this chapter's sections and its decision model (Fig. 1). The first aspect is the nature of the datasets involved, which can be qualitative datasets and/or quantitative datasets. Commonly, a study that includes one or more qualitative and one or more quantitative datasets is called a *mixed methods* study, whereas a study that includes either qualitative datasets or quantitative datasets, but not both, is called a *multimethod* study. The second aspect is the level of the decision, which is either the abstract level of qualitative versus quantitative datasets or the concrete level of specific types of datasets, such as questionnaire data or interview data. The third aspect is the type of decision. A distinction is made between decisions to *collect* data, to *combine* different forms of data collection and the decision to *further develop* collected datasets.

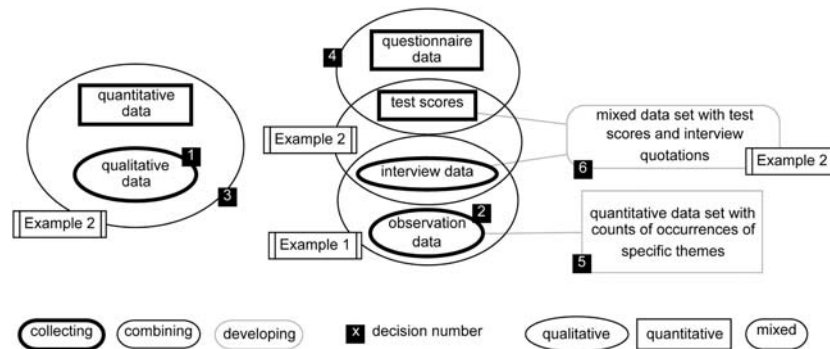


Fig. 1 Six decisions to collect, combine, and develop datasets in mixed methods and multimethod research. (1) Collecting qualitative or quantitative data; (2) Collecting specific kinds of data; (3) Combining qualitative and quantitative data collection; (4) Combining specific kinds of data collection; (5) Further developing qualitative data: quantizing; (6) Further developing qualitative and quantitative data: The mixed dataset.

Sections **Multimethod and mixed methods research** and **The development of an integrated claim using multiple methods** introduce the basic concepts and processes of research using multiple methods. Section **Multimethod and mixed methods research** discusses the difference between multimethod and mixed methods research. Section **The development of an integrated claim using multiple methods** demonstrates a simple multimethod example in which a final integrated claim is developed in two steps. Section **The design decisions: an overview** provides an overview of the design decisions related to the three aspects. In the following sections, decisions are discussed related to collecting data (Section **Decisions related to collecting data**), combining different forms of data collection (Section **Decisions 3 and 4: combining different types of data collection**), and further developing collected datasets (Section **Decisions 5 and 6: further developing collected data**). After the discussion of the decisions, Section **The meta-inference: a real-life example of the development of an integrated claim in mixed methods research** returns to the process that results from these decisions: the development of a final integrated claim, demonstrating how a meta-inference develops in a complex mixed methods example. Section **Conclusion** presents the conclusions.

Multimethod and mixed methods research

As noted in the introduction, research with multiple methods can be divided into *multimethod* research and *mixed methods* research. When two different methods are used, both of which result in distinct qualitative or quantitative datasets, this is known as *multimethod research* (for example, when quantitative test scores are combined with quantitative questionnaires or when qualitative interviews are combined with qualitative observations; see [Example 1](#) below). When two different methods are used—one of which results in one or more quantitative datasets, while the other results in one or more qualitative datasets—this is called *mixed methods research*. An example would be when test scores are combined with interviews (see [Example 2](#) below).

Mixed methods and multimethods research share important characteristics. For instance, obtaining a detailed understanding—typically called a deeper, comprehensive, or fuller understanding—has been recognized as an aim for mixed methods and multimethod research ([Greene, 2015](#); [Flick, 2017](#)). Similarly, including various perspectives is a characteristic of mixed methods research ([Mertens, 2007](#)) and multimethod research ([Flick, 2017](#)). Likewise, as will be explained in more detail later, abduction—the pattern of contradiction → explanation—is generally a characteristic of research with multiple methods ([Schoonenboom, 2022](#)), including but not limited to mixed methods research ([Morgan, 2007](#); [Wheeldon, 2010](#)) [See for an extension of these arguments to complex research generally, including mono-method research, [Denzin \(1978\)](#), [Reichert \(2014\)](#), [Sandelowski \(2014\)](#), and [Schoonenboom \(2022\)](#)].

Despite these fundamental resemblances, Section **Decisions related to collecting data** will show that some possible uses are unique to qualitative datasets, whereas others are unique to quantitative datasets. As a result, multimethod research can be used to answer a broader array of research questions than mono-method research, and mixed methods research can respond to an even broader array of research questions than multimethod research. Of course, this does not mean that conducting a mixed methods study is always preferable to multimethod research, just as it is not always appropriate to conduct a complex, multi-faceted study instead of investigating a singular question.

In this chapter, I use the term *research with multiple methods* as the overall term. I use the term *mixed methods* to refer to a study that includes both qualitative and quantitative methods and data, and I use the term *multimethod* for a study that includes the use of two or more qualitative methods or two or more quantitative methods, but not both types of methods. I am well aware that this is just one of the various ways in which the distinction between *multimethod* and *mixed methods* research has been drawn in the mixed methods literature (see [Box 1](#) below).

Box 1

In the mixed methods literature, the terms *mixed methods* and *multimethod* have various yet related meanings (for a historical and comprehensive overview, see Anguera et al., 2018). Different authors have emphasized different aspects of mixed methods and multimethod research, drawing different distinctions between them. Historically, the term *multimethod* came first: Brewer and Hunter (1989) used the term *multimethod* to refer to the use of more than one method in a study, including what I call mixed methods and multimethod research.

In the 1990s, the term *mixed methods* came to the fore, indicating the use of both quantitative and qualitative methods in one study. As a result, the term *multimethod* acquired a second meaning. Some authors continue to use the term *multimethod* to refer to any study in which multiple methods are used, including mixed methods studies (Hunter and Brewer, 2015; Plano Clark and Ivankova, 2016). Others specify a distinction between *mixed methods studies*, which involve quantitative and qualitative methods, and *multimethod studies*, which involve the use of two qualitative methods or two quantitative methods, but not both (Hesse-Biber, 2010; Creswell, 2014; Tashakkori et al., 2021).

A hallmark of mixed methods research is integration: bringing together quantitative and qualitative components—that is, the data and/or their results. As integration rose to prominence, the term *mixed methods* took on another meaning of its own. Some authors refer to a *mixed methods study* as one in which the qualitative and quantitative components are integrated (Creswell, 2014; Anguera et al., 2018), while others consider a study in which such integration lacks a *multimethod study*, even if both quantitative and qualitative data are used (Bazeley, 2018). For some authors, using qualitative and quantitative data is a requirement for a *mixed methods study*; for others, any study in which two or more methods are integrated can be called mixed methods research (Morse and Niehaus, 2009; Bazeley, 2018).

The development of an integrated claim using multiple methods

The aim of research using multiple methods is to obtain a more detailed understanding than would be possible using one method alone (Greene, 2007, 2015; Flick, 2017; Schoonenboom, 2019; Uprichard and Dawney, 2019). This detailed understanding becomes visible in an integrated claim, and this section demonstrates how an integrated claim is developed using multiple methods. Example 1 is a multimethod study by Finlay et al. (2015):

Example 1. Finlay et al. (2015) investigated the delivery of sexual health education to young people with intellectual disabilities. One of the study's foci was how educators anticipate problems of comprehension and how they act when a learner shows a lack of comprehension. Finlay et al. (2015) made use of two different qualitative methods. To investigate educators' behavior, they conducted and analyzed video observations; to learn more about the views that motivate their behavior, they conducted and analyzed semi-structured interviews. According to their discussion:

It seemed to be the case that the best evidence of comprehension was produced at first attempts, and rarely seemed convincing if it was the product of challenge and correction. In the examples here, the learner eventually offers the 'correct' response, and the leader then moves on to the next activity. A superficial resolution is achieved. We can understand this pragmatic approach in terms of both the models of understanding offered by the educators, whereby development of comprehension is nonlinear, and learners are only expected to pick up 'bits,' and by the immediate demands of the situation, in that there are other group members to consider and an agenda of activities to get through in the allotted time (p. 336).

As a result of incorporating the different datasets, the authors were able to develop a final integrated claim, as displayed in Table 1 (based on a format in Schoonenboom, 2022). Table 1 shows how an integrated claim develops. The authors used two qualitative methods of data collection, video observations and interviews, which resulted in two types of qualitative data: video recordings and interview transcripts, from which the authors derived Claims [1], [2] and [4], which were subsequently integrated into integrated claims (IC) [3] and [5].

Example 1 demonstrates the various characteristics of research with multiple methods. First, the development of the conclusion from start to finish involves the use of multiple methods, in this case, two qualitative methods, video observations and interviews; therefore, Example 1 is a multimethod study. Second, the conclusion—(IC) [5]—is not a simple claim; rather, it comprises three simple claims, [1], [2], and [4], which are all interconnected. Third, IC [5] is not created by simply combining claims [1], [2], and [4] in one step; instead, it is developed over several steps that build on each other (two steps in this case). In Step I, simple claims [1] and [2] are integrated into a complex claim, IC [3]; in Step II, IC [3] and simple claims [4] are integrated to form the conclusion, IC [5]. Fourth, each integration step can be assigned a label depending on the methods involved in deriving its claims. In this case, Step II is decisive for calling the research wholly multimethod because it integrates claims derived from two different qualitative datasets, observations, and interviews and is therefore labeled *multimethod*. Step I is labeled *mono-method* because it integrates claims derived from using only one type of data, the video observations.

Fifth, these two steps display a pattern. Step I results in a contradiction: "Evidence of comprehension produced at first attempts was convincing, *but* evidence of comprehension that was the result of challenge and correction was superficial." This contradiction is then resolved in Step II: "Evidence of comprehension produced at first attempts was convincing, *but* evidence of comprehension that was the result of challenge and correction was superficial, *because* educators did expect comprehension to slowly develop, and they wanted to give attention to all pupils." This pattern of contradiction, followed by an explanation, is a hallmark of research with multiple methods. It is generally known as *abduction* (Hoffmann, 1999; Reichertz, 2014) and is considered essential to mixed

Table 1 Development of a final integrated claim in Example 1.

Methods	Obtained claims	Integrated claim	Connection
I. Mono-method (qualitative)			
Video observations	→ [1] Evidence of comprehension produced at first attempts was convincing.	IC[3] Evidence of comprehension produced at first attempts was convincing, but evidence of comprehension that was the result of challenge and correction was superficial.	Contradiction
Video observations	→ [2] Evidence of comprehension that was the result of challenge and correction was superficial.		
II. Multimethod (qualitative)			
Video observations	→ IC[3] Evidence of comprehension produced at first attempts was convincing, but evidence of comprehension that was the result of challenge and correction was superficial.	Final integrated claim IC[5] Evidence of comprehension produced at first attempts was convincing, but evidence of comprehension that was the result of challenge and correction was superficial, because educators did expect comprehension to slowly develop, and they wanted to give attention to all pupils.	Explanation
Interviews	→ [4] Educators did expect comprehension to slowly develop, and they wanted to give attention to all pupils.		

Note. IC = integrated claim.

methods research (Morgan, 2007; Wheeldon, 2010). As one of the main rationales for conducting mixed methods research, it is also called *initiation* (Greene et al., 1989).

The design decisions: an overview

For claims to be derived and integrated, researchers make various decisions regarding the use of multiple methods in a study, as shown in Fig. 1. Fig. 1 shows that researchers make decisions on data collection on the abstract level of quantitative and qualitative data (Decisions 1 and 3) and at the level of specific data collection methods (Decisions 2 and 4). They make these decisions both related to the data to be collected separately (Decisions 1 and 2) and to their combination (Decisions 3 and 4). Finally, once data have been collected, researchers can decide to further develop their datasets into additional datasets (Decisions 5 and 6). These six decisions are discussed below.

Decisions related to collecting data

Decision 1: collecting qualitative and/or quantitative data

Depending upon their goals, mixed methods and multimethod researchers collect quantitative and/or qualitative datasets. The difference between quantitative and qualitative datasets is best captured by the labels *structured* and *unstructured* (Ahuja et al., 2020). Quantitative data are structured data that can be represented as numbers (Onwuegbuzie and Leech, 2019; Nzabonimpa, 2018; Sandelowski et al., 2009). They consist of variables and their values collected in a fixed structure, usually a table. In contrast, qualitative datasets are unstructured; they are non-numerical data of various types, such as text, video, audio, drawings, or material objects. As a result of this difference, qualitative and quantitative data partly differ in what researchers can use them for, which forms the basis for the decision on whether to include qualitative and/or quantitative data in their studies:

Possibilities of Qualitative Data

- Discover unknown themes and relations, including reasons, motives, and causal relations.
- Establish the influence of context on causal relations.
- Reveals patterns in the research data.
- Develop narratives of experiences and events in detail.

Possibilities of Quantitative Data

- Describe statistics at the group level and provides an overview.
- Uncover patterns among known variables.
- Provide quick diagnostics of individuals.
- Identify deviating individuals quickly.
- Establish overall effects (causal relations) in larger groups.
- Establish differences between groups.
- Establish the influences of known context variables on causal relations in larger groups.
- Construct narratives among known variables in larger groups.

As these lists make clear, both qualitative and quantitative data can be used to discover and/or test causal relations (Maxwell, 2019) and to establish the influence of context. In quantitative research, this is called moderation, and the influencing factor is termed a moderating factor (Hayes, 2017).

Both types of data can also be used to discover new patterns (Glaser, 2008) and construct narratives. Note, however, that newly discovered patterns in a quantitative dataset are necessarily based on the existing variables and their values—they are all the researcher has—whereas from a qualitative dataset, completely new “variables”, or “themes”, may arise.

Both types of data allow the construction of narratives. Using qualitative data, a researcher can construct detailed narratives that show how one event leads to another, what the motives of those involved are, and how the development of events is influenced by context. By coding the absence and presence of specific events, quantitative data allow us to determine which combinations of events, and hence which narratives, are likelier to occur. Various mixed methodologies transform qualitative data to construct such quantitative narratives, a transformation that will be explained in more detail in section **Quantitizing: converting qualitative data into quantitative data**. Example methodologies include Qualitative Comparative Analysis (QCA; Rihoux and Ragin, 2009) and multiple-case narratives (Abbott, 1992; for an example, see Abbott and Hrycak, 1990).

Quantitative data can be used for quick diagnostics and identification at the group and individual levels, based on the values of specific variables. This includes descriptive statistics—that is, a quick overview of a group with respect to a selection of known variables—as well as the calculation of effects and effect sizes. An important function of quantitative variables in mixed methods research is creating groups, allowing researchers to explore differences between experimental and control groups, for example.

Decision 1: What are your goals as a researcher? What type of data (quantitative or qualitative) could you use to reach your goals?

Decision 2: collecting specific types of data

The general properties of qualitative and quantitative data in the previous section can help researchers decide whether they want to include qualitative and/or quantitative data in their studies. Even more important is the decision on what specific data to collect and which specific methods to use. In **Example 1**, the qualitative data generated by classroom video observations allowed researchers to investigate pupils’ and teachers’ behaviors, but they could not tell them anything about the teachers’ ideas and motives behind these behaviors, whereas interviews allowed the researchers to investigate teachers’ ideas and motives, but were unable to show actual behavior. Both the methods and the generated datasets were qualitative, but the types of claims that researchers can draw from them are different.

Similarly, in multimethod quantitative research, differences in test scores can tell the researcher something about whether a classroom intervention had an effect, but they cannot tell the researcher about how learners experienced the intervention, while questionnaires can tell the researcher about how learners experienced the interventions, but they do not show the researcher their test results. In this case, the methods used are both quantitative, but the types of conclusions that researchers can draw from them are different.

Decision 2: What specific methods will you include for which specific purpose?

Decisions 3 and 4: combining different types of data collection

Having seen how qualitative and quantitative data can be applied to different uses, the next question is why researchers want to combine them in a single study. As stated above, the decision to combine methods can be situated both at the abstract level of combining qualitative and quantitative datasets (Decision 3 in **Fig. 1**) and at the level of the specific methods of data collection (Decision 4 in **Fig. 1**). This section will discuss Decisions 3 and 4 concurrently, using real-life examples from educational research. The real-life examples all involve mixed methods research and show the advantages of combining quantitative and qualitative datasets. In general, two reasons for combining can be distinguished: one is combining qualitative and quantitative data collection and data analysis (Section **Integrating outcomes of the analysis of separate datasets**); the other is using one type of data as an “add-on” to a study of the other type (Section **Using quantitative data for development in a qualitative study (and vice versa)**).

Integrating outcomes of the analysis of separate datasets

In a mixed methods study, qualitative data and quantitative data can be collected separately, and the resulting qualitative and quantitative datasets can each be subjected to their own types of data analysis. This is the “prototypical” mixed methods study: It is the type of study on which the three widely-used “core designs” of Creswell and Plano Clark (2018) are based. Such a prototypical mixed methods study can be conducted for various reasons. Some examples are presented below.

To apply themes discovered in interviews for generalization in a quantitative study

Mitchell Fuentes (2008) conducted in-depth life history interviews with an ethnically diverse convenience sample of 28 women with at least one experience of sexual abuse. The study revealed that depression was a main risk factor that connected sexual abuse to an increased risk of sexually transmitted infections. A questionnaire among a larger sample of women (with and without experience of abuse) largely confirmed these findings.

In this example, including qualitative data made it possible for Mitchell Fuentes to discover an unknown causal pattern between partly unknown risk factors. Subsequently, quantitative data collection widened the field of research and allowed the pattern to be tested. This would have been much more difficult to achieve without quantitative data collection. Equally, the larger sample afforded by the quantitative data collection allowed for a comparison between women with and without experience of sexual abuse. Related to the specific methods, both qualitative interviews and quantitative questionnaires allow people to express their experiences and opinions. Therefore, combining them made it possible to use quantitative questionnaires to generalize the qualitative findings.

To provide a quantitative overview as part of qualitative research

Varshney (2001) set out to better understand Hindu–Muslim violence in India. Although much of the resulting research is based on interviews with Hindus and Muslims, Varshney started with the question, “Where and when did Hindu–Muslim riots in India between 1950 and 1995 take place?” To answer this question, he used quantitized data derived from riot reports in the *Times of India*. By attaching the labels *place* and *time*, as well as derivatives such as *province* and *urban* vs. *rural*, Varshney was able to compile a geographical and chronological overview of Hindu–Muslim violence in this place. His findings highlighted how riots had hardly occurred at all in rural areas and that there were significant differences between cities.

In this example, the researcher could not have obtained such an overview from the interview data alone. Without the additional collection of data and its subsequent quantization, this overview would not have been possible. More specifically, whereas quantitized data from riot reports could provide an overview of where and how often riots took place, qualitative interviews could be used to obtain a detailed insight into the experiences and dynamics of Hindu–Muslim riots. In combination, quantitized data from riot reports allowed researchers to create a set of cities with different levels of riot-proneness from which to draw a heterogeneous sample for their interviews.

To further explore differences between quantitative groups or deviating individuals

Using data from a questionnaire among teaching staff at one university, Schoonenboom (2016) identified one department in which the digital learning environment Blackboard was used more often than in the other departments. Interviews with Blackboard organizers revealed that this was the only department where the use of Blackboard was compulsory and where Blackboard forums were prestructured to guide discussions among students.

Here, quantitative data collection and data analysis allowed the researcher to quickly identify variables from which (in this case) the teaching staff in one department deviated. Qualitative data about how Blackboard was organized was needed because organizing Blackboard was evidently a complex process, with many unknowns that could not be adequately captured by a pre-structured questionnaire. At a more specific level, quantitative questionnaires allowed a large group of teachers from different departments to indicate how often they used Blackboard, whereas qualitative interviews allowed Blackboard organizers to explain how they organized the use of Blackboard and why. In combination, the identification of a deviating department using the quantitative questionnaire allowed researchers to investigate differences between this department and other departments in their interviews.

To further explore any expected or unexpected quantitative results using qualitative data collection and analysis

Using a questionnaire, Ivankova and Stick (2007) discovered that the factors contributing to students’ persistence in higher education were *program*, *online learning environment*, *student support services*, *faculty*, and *self-motivation*. They conducted interviews with typical respondents to explore the results in greater depth. Four major themes emerged: (1) quality of academic experiences, (2) online learning environment, (3) support and assistance, and (4) student self-motivation.

By using quantitative data collection and data analysis, Ivankova and Stick (2007) were able to obtain a topical overview of a larger group of students and ascertain the extent to which known factors and their interactions contributed to their persistence in higher education. Through their use of quantitative data collection and analysis, they were able to discover themes related to persistence in higher education. At a more specific level, quantitative questionnaires allowed students to express how they valued the program, online learning environment, student support services, and the faculty while at the same time allowing researchers to correlate their judgments’ values. Qualitative interviews allowed the students to elaborate on these factors, thereby allowing the researchers to validate them and understand them in more depth. In combination, through their

qualitative interviews, they were able to refine their quantitative findings and discover the processes underpinning the initial questionnaire responses.

Using quantitative data for development in a qualitative study (and vice versa)

In addition to analysis, qualitative and quantitative datasets can also be used as an add-on to another study. This is a broad category of all uses of datasets in which data are not used to derive claims. This includes, but is not limited to, the use of a dataset of one type to develop a study with a dataset of the other, also called *development* (Greene et al., 1989).

I will discuss two uses of data as an add-on here. The first is the use of quantitative data and datasets for sampling. For example, researchers can use tests (e.g., standardized tests) to select a sample of participants for further research. They do so by selecting participants who show specific values for one or more variables.

To use quantitative data for sampling for qualitative data collection

Visser et al. (2018) used a questionnaire to identify students with high, middle, and low levels of procrastination and then selected students from these three groups for an interview.

Visser et al. (2018) used quantitative data to provide a quick diagnosis of individuals, and they used qualitative data to develop narratives of experiences and events of procrastination. At the level of the specific methods, questionnaires about procrastination allowed students to assess their different procrastination behaviors and experiences and allowed researchers to assess the level of procrastination of each student. Interviews allowed students to express their experiences with procrastination and researchers to discover patterns in the students' narratives. As a combination, the selection of students based on their quantitative level of procrastination allowed researchers to investigate the "worlds" of students with high, middle, and low levels of procrastination in more depth in their interviews.

In mixed methods research, quantitative data are often used for sampling. In the various examples above, quantitative data were used to select (a) students with different levels of procrastination for an interview (Visser et al., 2018); (b) cities in India with varying rates of riot violence for further exploration (Varshney, 2001); (c) departments in which to explore the organization of Blackboard (Schoonenboom, 2016); and (d) students to interview about their persistence in higher education (Ivanova and Stick, 2007).

To use qualitative data to illustrate quantitative findings

In this case, the function of the illustration is to further the reader's understanding by highlighting how the conclusion plays out in practice, in a specific real-life case. This is also called "putting 'meat on the bones' of 'dry' quantitative findings" (Bryman, 2006, p. 106).

Decisions 3 and 4: What are your motives or reasons for including both qualitative and quantitative data? What are your motives or reasons for including the specific types of data that you want to combine in your study (observations, interviews, test scores, questionnaire data, etc.)?

Decisions 5 and 6: further developing collected data

Quantitizing: converting qualitative data into quantitative data

When a mixed methods researcher uses both a quantitative dataset and a qualitative dataset, this does not necessarily mean that he or she carried out both qualitative and quantitative data collection. In fact, quantitative datasets can also be obtained by *quantitizing* collected qualitative data—that is, by creating a quantitative dataset from a qualitative one. This is often called "converting" qualitative data into quantitative data (Nzabonimpa, 2018; Onwuegbuzie and Leech, 2019).

The quantization procedure is as follows: Qualitative researchers often use codes to denote fragments of, for example, text data or video data. In qualitative data analysis, these codes can be used to find and analyze the fragments attached to them. The unit of a code with its fragment remains qualitative, unstructured data: The fragments are unstructured, and it is the fragments that are analyzed in qualitative data analysis, not the codes.

However, it is also possible to separate the codes from their corresponding fragments. In doing so, the codes are *quantitized*; they become quantitative data, which can then be stored in a list or table. Lists of quantitized codes can then be subjected to quantitative analysis, for example, by counting the number of specific codes (Sandelowski, 2014). Below is a real-life example (discussed in Bazeley, 2012) in which qualitative data are quantitized:

Kaplun (2012) studied the experiences of 188 families from a disadvantaged area in Australia with their child's transition to school. The researcher conducted semi-structured interviews with the parents before and after the transition. She asked them to describe their own experiences and those of their children, including activities they thought their child would enjoy and/or do well in at school. By counting, comparing, and classifying the activities mentioned in the interviews, Kaplun was able to show that the number of parents who said their child enjoyed or was good at a specific activity declined during the child's transition to school for arts-related activities such as dancing, singing, painting, and drawing, whereas it remained constant for academic activities.

Decision 5: Do you intend to quantitize your qualitative data? If so, for what purpose?

The mixed dataset

A second possibility for further developing collected datasets is by bringing the collected qualitative and quantitative data together into a mixed dataset. A mixed dataset has a structure similar to that of a quantitative dataset, namely a table. One or more columns of a mixed table will contain structured data (either numerical values or codes) derived either directly from quantitative data collection or by quantizing qualitative data. In addition, one or more columns will contain unstructured data, most typically text fragments. The process of creating a mixed dataset is also called *data consolidation* (Caracelli and Greene, 1993; Vogl, 2019).

In mixed data analysis, researchers analyze the structured and unstructured data jointly in a mixed dataset. Mixed datasets can be used to explore any types of differences and resemblances between qualitative and quantitative data. In a mixed dataset, researchers combine the possibilities of qualitative and quantitative data. For example, they can use the quantitative data to uncover patterns among the variables and then consider the qualitative data to see if these corroborate the patterns, reveal additional aspects, or provide additional insights.

The quantitative data in a mixed table are well suited to create groups among the qualitative data. This principle is used for data analysis in the case comparison table (Schoonenboom and Johnson, 2021), in which the values of the quantitative data are used to sort and re-sort the data, create groups, and inspect the qualitative data for each group separately, with the aim of identifying similarities within and differences between groups. As the analysis proceeds, more quantitative diagnostic codes are developed until subgroups that are discrete and make sense have been identified, which may potentially provide an explanation for the differences between groups. The example used by Schoonenboom and Johnson (2021) is presented here as Example 2.

Example 2. Schoonenboom and Johnson (2021) used quantitative and qualitative data published in Lee and Greene (2007), a study on the relationships between the language problems of international students and their grade point average (GPA) in their first semester at one university in the US. Schoonenboom and Johnson (2021) created a simple table containing one record for each student, along with their language score, GPA, and a quote from their interview, which could then be sorted and resorted to uncover patterns. One specific group of four students emerged. Despite their language problems, members of this group still obtained the highest GPA. The quotes show that three of the four students in this group referred to compensation strategies, while none of the students with language problems and less than the highest possible GPA did so. In this way, Schoonenboom and Johnson (2021) were able to formulate a hypothesis for further research: Language problems affect international students' GPA unless they deliberately use compensation strategies.

Whereas the quantitative data are used to create the groups initially, the qualitative data are then explored for often unidentified points of differences between the groups, as well as for identifying the reasons or causes for these differences. As a result, mixed data analysis is particularly suited to discovering and exploring moderation, as highlighted in Example 2: Schoonenboom and Johnson (2021) formulated an effect ("language problems affect international students' GPA") that is moderated by the factor *compensation strategies* ("unless they deliberately use compensation strategies").

Decision 6: Do you intend to bring your quantitative and qualitative data together in a mixed dataset? If so, for what purpose?

The meta-inference: a real-life example of the development of an integrated claim in mixed methods research

Section **The development of an integrated claim using multiple methods** demonstrated the development of a final integrated claim in a simple, multimethods example, Example 1. Now, at the end of the chapter, it is time to show the development of an integrated claim in a complex mixed methods example. Table 2 shows the integration process in Example 2.

Table 2 contains three steps: Step I is the main conclusion by Lee and Greene (2007), while Steps II and III stem from the reanalysis by Schoonenboom and Johnson (2021). Similar to Table 1, Lee and Greene's conclusions were more complex, but the majority have been omitted as they did not play a role in the reanalysis.

The integration process in mixed methods Example 2 is the same as in multimethod Example 1. First, its conclusion, integrated claim (IC) [8], is not a simple claim; rather, it contains five simple claims [1], [2], [4], [6], and [7] that are interconnected. Second, this conclusion is not created by combining the five simple claims in one step, but over several steps (three in this case) that build on each other. In Step I, claims [1] and [2] are integrated into a complex claim, IC [3]. In Step II, claims [2] and [4] are integrated into a complex claim, IC [5]. In Step III, claims [4], [6], and [7] are integrated to form the integrated inference, IC [8]. Third, these steps can be assigned a label depending on the number and types of methods involved in deriving their claims. In this case, all three steps are labeled *mixed methods* because all steps involve either a qualitative and a quantitative dataset (Step I) or a mixed dataset (Steps II and III). Fourth, these steps follow the same pattern of contradiction followed by an explanation. In Steps I and II, two different contradictions are developed, both of which are resolved in Step III: Some students had a maximum GPA despite less than perfect English skills *because* they possessed compensation strategies. This integrated inference explains both why English skills still affect

Table 2 Development of an integrated inference in [Example 2](#).

Methods	Obtained claims	Integrated claims	Connection
I. Mixed methods			
Quantitative questionnaire	→ [1] Nonsignificant correlations were found between test scores and GPA.	IC [3] Although nonsignificant correlations were found between test scores and GPA, qualitative findings indicated that English skills are an important factor affecting students' course performance (Lee and Greene, 2007).	Contradiction
Interviews	→ [2] English skills are an important factor affecting students' course performance.		
II. Mixed methods			
Interviews	→ [2] English skills are an important factor affecting students' course performance.	IC [5] Although English skills are an important factor affecting students' course performance, some students had a maximum GPA despite less than perfect English skills.	Contradiction
Quantitative data	→ [4] Some students had a maximum GPA despite less than perfect English skills.		
III. Mixed methods			
Quantitative data	→ [4] Some students had a maximum GPA despite less than perfect English skills.	Integrated inference IC [8] Some students had a maximum GPA despite less than perfect English skills because they possessed compensation strategies.	Explanation
Mixed data	→ [6] Three out of four of these students mentioned compensation strategies in their interviews.		
Mixed data	→ [7] None of the students with less than perfect English skills and a less than maximum GPA mentioned compensation strategies.		

GPA (contradiction of Step I) and why some students, despite less than perfect English skills, nonetheless obtained the highest GPA possible (contradiction of Step II).

Yet, at the same time, [Example 2](#) also highlights the advantages of combining qualitative and quantitative datasets. Claim [1] shows the benefits of including quantitative data. Analysis of this data in [1] showed that, for this sample, language problems did not affect the students' GPA (i.e., establishing overall effects—causal relations—with larger groups). This was an important conclusion, as it set the stage for both [Lee and Greene \(2007\)](#) and [Schoonenboom and Johnson \(2021\)](#) to ask the question, 'Why?' Without the quantitative data, this question would not have arisen. The quantitative data were then used by [Lee and Greene \(2007\)](#) to identify students whose performance had been affected by language problems as well as those for whom this was not the case (i.e., identifying deviating individuals quickly). Next, it was only through analysis of the qualitative data that [Lee and Greene \(2007\)](#) were able to show that language problems had an effect on individual students (i.e., discovering unknown themes and relations, including causal relations). The researchers were then able to use extreme case analysis to explain the lack of a quantitative effect, whereby variations in students' views of academic success and their perceived levels of background knowledge can help explain their low GPAs relative to their CEEPT scores ([Lee and Greene, 2007](#), p. 366).

At the beginning of their reanalysis, [Schoonenboom and Johnson \(2021\)](#) returned to the quantitative data. Whereas [Lee and Greene \(2007\)](#) mainly focused on individuals, [Schoonenboom and Johnson \(2021\)](#) identified a group of students who, despite having less than perfect English language skills, still managed to achieve the maximum GPA (i.e., establishing differences between groups). In [6], [Schoonenboom and Johnson \(2021\)](#) created a mixed dataset, showing that three of these four students mentioned compensation strategies in their interviews (i.e., identifying and exploring differences between subgroups). Finally, in [7], [Schoonenboom and Johnson \(2021\)](#) observed in their mixed dataset that none of the students with language problems and a less than perfect GPA mentioned such compensation strategies, which enabled them to formulate a hypothesis in [8] that some students had a maximum GPA despite less than perfect English skills *because* they possessed compensation strategies (i.e., uncovering context influences on causal relations between subgroups).

Conclusion

In this chapter, I describe mixed methods research and multimethod research as forms of research with multiple methods. Mixed methods and multimethod research are members of the same family: research that results in a final integrated claim developed from various claims and integrated claims. Mixed methods and multimethod research share the same structure and process of deriving claims through the analysis of empirical data and integrating these claims one by one. In this process, the abductive pattern of contradiction→explanation plays an important role, as seen in both multimethod [Example 1](#) and mixed methods [Example 2](#).

The implication of these commonalities is that abduction is not a specific characteristic of mixed methods research (see for a different view [Morgan, 2007](#)). Rather, as [Example 1](#) has shown, it is a characteristic of research with multiple methods more generally (for abduction in qualitative research, see [Reichert, 2014](#)). This also means that, rather than being one of the reasons for conducting mixed methods research, the contradiction→explanation pattern is an integral part of mixed methods research.

On the other hand, mixed methods research represents a special form of research with multiple methods, characterized by a combination of qualitative and quantitative data, which are fundamentally different. Structured quantitative data can be subjected to quantitative data analysis: Occurrences of specific values can be counted, effects can be calculated, and so on; quantitative data, on the contrary, cannot be used to discover completely new variables. Unstructured qualitative data can be subjected to qualitative data analysis, whereby more complex meanings are derived over several rounds of analysis. However, they cannot be counted or used for selection. By combining both types of data, the mixed dataset allows researchers to explore unknown differences (in the qualitative data) among emerging groups (identified by the quantitative data).

At the level of claim derivation, there is a clear overlap between what researchers can do with qualitative and quantitative data. Both types of data can be used to establish causation, construct narratives or discover moderation. Thus, instead of being different approaches at a higher level, *quantitative* and *qualitative approaches* refer to different types of datasets at a lower level. This shift of the fundamental differences between qualitative and quantitative—from the level of approaches to the level of data—has implications for how researchers should describe mixed methods research. Rather than saying that they used a qualitative and quantitative approach, researchers should describe what they do in terms of the use of qualitative, quantitative and mixed datasets and data analysis. For example:

- We used the quantitative dataset to draw a sample for the interviews, which generated a qualitative dataset, which we then analyzed thematically.
- We quantitized the qualitative interview data, thereby creating a quantitative dataset.
- In a mixed table analysis, we explored the differences between groups in how they responded.

In line with this, researchers should connect other aspects of mixed methods research to what they actually do with their data and not to quantitative or qualitative approaches in general. Validity criteria provide an excellent example of this. Researchers cannot say that they had to obtain a large sample size because they used a quantitative approach. This is because the sample size requirements for a quantitative dataset differ widely, depending on whether the dataset is used to establish an effect, to select individuals, to describe a sample or as part of a mixed dataset.

What about the implications for mixed methods design? Of the six design decisions, four could be considered strategies for mixed methods research because they involve a combination of qualitative and quantitative data: combining qualitative and quantitative data collection and data analysis; using data as an add-on; quantitizing qualitative data; and creating and analyzing mixed datasets. An important characteristic of these four strategies is that, in mixed methods research, they are, and should be, combined. Mixed methods researchers do not choose between these four strategies; they use each of them as and when needed to develop claims and to find and resolve contradictions (see [Schoonenboom, 2019](#)). Thus, in mixed methods research, quantitative data are often used not only for analysis, but for sampling as well. Equally, a mixed methods analysis is often conducted in addition to analyzing the qualitative and quantitative data separately, while qualitative data are often quantitized as well as analyzed ([Sandelowski, 2014](#)).

Significantly, this approach to mixed methods design differs in several respects from research design typologies, as exemplified by [Creswell and Plano Clark \(2018\)](#). Design typologies assume that mixed methods researchers choose between designs based on how the separate strands of qualitative and quantitative data collection and analysis are combined. In doing so, they hide the similarity of the integration process in research with multiple methods generally and the way in which, in good mixed methods research, several of the four strategies are combined. This is not to say that the authors are unaware of these four strategies, only that they are not visible in the design typologies.

Like the illness of atherosclerosis in philosopher [Mol's \(2002\)](#) brilliant book, the decision to include multiple methods in a study is complex and multiple. I hope that this chapter has inspired researchers to navigate this complexity and multiplicity in a productive way.

References

- Abbott, A., 1992. What do cases do? Some notes on activity in sociological analysis. In: Ragin, C.C., Becker, H.S. (Eds.), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge University Press, Cambridge, UK, pp. 53–82.
- Abbott, A., Hrycak, A., 1990. Measuring resemblance in sequence data: an optimal matching analysis of musicians' careers. *Am. J. Sociol.* 96, 144–185.
- Ahuja, R., Asthana, S., Ahuja, A., Agarwal, M., 2020. Data: its nature and modern data analytical tools. In: Gupta, D., Bhattacharyya, S., Khanna, A., Sagar, K. (Eds.), *Intelligent Data Analysis: From Data Gathering to Data Comprehension*. Wiley, Hoboken, NJ, pp. 17–42.
- Anguera, M.T., Blanco-Villaseñor, A., Losada, J.L., Sánchez-Algarra, P., Onwuegbuzie, A.J., 2018. Revisiting the difference between mixed methods and multimethods: is it all in the name? *Qual. Quantity* 52, 2757–2770.
- Bazeley, P., 2012. Integrative analysis strategies for mixed data sources. *Am. Behav. Sci.* 56, 814–828.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage, Los Angeles, CA.
- Brewer, J., Hunter, A., 1989. *Multimethod Research: A Synthesis of Styles*. Sage, Thousand Oaks, CA.
- Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Res.* 6, 97–113.
- Caracelli, V.J., Greene, J.C., 1993. Data analysis strategies for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 15, 195–207.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. Sage, Los Angeles, CA.
- Creswell, J.W., 2014. *A Concise Introduction to Mixed Methods Research*. Sage, Los Angeles, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*. Sage, Los Angeles, CA.
- Denzin, N.K., 1978. *The Research Act: A Theoretical Introduction to Sociological Methods*. McGraw-Hill, New York, NY.

- Finlay, W.M.L., Rohleder, P., Taylor, N., Culfear, H., 2015. 'Understanding' as a practical issue in sexual health education for people with intellectual disabilities: a study using two qualitative methods. *Health Psychol.* 34, 328–338.
- Flick, U., 2017. Mantras and myths: the disenchantment of mixed-methods research and revisiting triangulation as a perspective. *Qual. Inq.* 23, 46–57.
- Glaser, B., 2008. *Doing Quantitative Grounded Theory*. Sociology Press, Mill Valley, CA.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods research merger. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 606–615.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11, 255–274.
- Hayes, A.F., 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Press, New York, NY.
- Hesse-Biber, S., 2010. Qualitative approaches to mixed methods practice. *Qual. Inq.* 16, 455–468.
- Hoffmann, M., 1999. Problems with Peirce's concept of abduction. *Found. Sci.* 4, 271–305.
- Hunter, A., Brewer, J.D., 2015. Conundrums of multimethod research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 616–623.
- Ivankova, N.V., Stick, S.L., 2007. Students' persistence in a distributed doctoral program in educational leadership in higher education: a mixed methods study. *Res. High. Educ.* 48, 93–134.
- Kaplun, C., 2012. *The Starting School Study: The Transition to School Experiences of Families Living in Disadvantage*. Doctoral Dissertation. Charles Sturt University, Bathurst, Australia.
- Latour, B., Woolgar, S., 1986. *Laboratory Life: The Construction of Scientific Knowledge*. Princeton University Press, Princeton, NJ.
- Lee, Y.-J., Greene, J., 2007. The predictive validity of an ESL placement test: a mixed methods approach. *J. Mix. Methods Res.* 1, 366–389.
- Maxwell, J.A., 2019. Causality in qualitative research. In: Atkinson, P., Delamont, S., Cernat, A., Sakshaug, J.W., Williams, R.A. (Eds.), *SAGE Research Methods Foundations*. Sage, London, UK.
- Mertens, D.M., 2007. Transformative paradigm: mixed methods and social justice. *J. Mix. Methods Res.* 1, 212–225.
- Mitchell Fuentes, C.M., 2008. Pathways from interpersonal violence to sexually transmitted infections: a mixed-method study of diverse women. *J. Wom. Health* 17, 1591–1603.
- Mol, A., 2002. *The Body Multiple: Ontology in Medical Practice*. Duke University Press, Durham, NC.
- Morgan, D.L., 2007. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *J. Mix. Methods Res.* 1, 48–76.
- Morse, J.M., Niehaus, L., 2009. *Mixed Method Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Nzabonimpa, J.P., 2018. Quantitizing and qualitzing: (im-)possibilities in mixed methods research. *Methodol. Innov.* 11. <https://doi.org/10.1177/2059799118789021>.
- Onwuegbuzie, A.J., Leech, N.L., 2019. On qualitzing. *Int. J. Mult. Res. Approaches* 11, 98–131.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Practical Guide to the Field*. Sage, Los Angeles, CA.
- Reichert, J., 2014. Induction, deduction, abduction. In: Flick, U. (Ed.), *The SAGE Handbook of Qualitative Data Analysis*. Sage, Los Angeles, CA, pp. 123–135.
- Rihoux, B., Ragin, C.C., 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Sage, Los Angeles, CA.
- Sandelowski, M., 2014. Unmixing mixed-methods research. *Res. Nurs. Health* 37, 3–8.
- Sandelowski, M., Voils, C.I., Knaff, G., 2009. On quantizing. *J. Mix. Methods Res.* 3, 208–222.
- Schoonenboom, J., 2016. The multilevel mixed intact group analysis: a mixed method to seek, detect, describe, and explain differences among intact groups. *J. Mix. Methods Res.* 10, 129–146.
- Schoonenboom, J., 2019. Develop your case! How controversial cases, subcases, and moderated cases can guide you through mixed methods data analysis. *Front. Psychol.* 10 (1369).
- Schoonenboom, J., 2022. Developing the meta-inference in mixed methods research through integration of claims. In: Hitchcock, J.H., Onwuegbuzie, A.J. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge, London, UK, pp. 55–70.
- Schoonenboom, J., Johnson, R.B., 2021. The case comparison table: a joint display for constructing and sorting simple tables as mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York, NY, pp. 277–278.
- Teddlie, C.B., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Los Angeles, CA.
- Tashakkori, A.M., Johnson, R.B., Teddlie, C.B., 2021. *Foundations of Mixed Methods Research*. Sage, Los Angeles, CA.
- Uprichard, E., Dawney, L., 2019. Data diffraction: challenging data integration in mixed methods research. *J. Mix. Methods Res.* 13, 19–32.
- Varshney, A., 2001. Ethnic conflict and civil society: India and beyond. *World Polit.* 53, 362–398.
- Visser, L., Korthagen, F.A.J., Schoonenboom, J., 2018. Differences in learning characteristics between students with high, average, and low levels of academic procrastination: students' views on factors influencing their learning. *Front. Psychol.* 9 (808).
- Vogl, S., 2019. Integrating and consolidating data in mixed methods data analysis: examples from focus group data with children. *J. Mix. Methods Res.* 13, 536–554.
- Wheeldon, J., 2010. Mapping mixed methods research: methods, measures, and meaning. *J. Mix. Methods Res.* 4, 87–102.

Foundations of mixed methods research

Abbas Tashakkori^a and **Isadore Newman^{b,†}**, ^a Professor Emeritus, University of North Texas, Denton, TX, United States; and
^b University of Akron, Akron, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Purposes and utility of mixed-methods research	373
Research questions in mixed-methods research	374
Design of mixed-methods studies	374
Data-collection strategies in mixed-methods research	376
Data-analysis strategies in mixed-methods research	376
Integrated inferences in mixed-methods studies	376
Transferability of inferences in mixed-methods studies	377
Conclusions	377
References	378
Further reading	379

Glossary

Conversion mixed methods designs A type of mixed methods design in which the data of one strand of a study is the result of conversion of the data in another (for example, from qualitative to quantitative)

Design quality The degree to which all elements of investigator's research design are implemented appropriately and with fidelity

Hybrid mixed methods designs A type of mixed methods study with more than two strands/phases, utilizing more than one of the other three families of designs (parallel, sequential, conversion) in a dynamic manner

Inference (or research inference) A conclusion, understanding, or explanation derived from the results of data analysis in a research study

Inference quality The degree to which the interpretations and conclusions of a study are seen as valid, credible, acceptable, and more plausible than other possible explanations of the findings

Inference transferability The generalizability or applicability of conclusions made on the basis of the results of a study to other individuals, entities, settings, situations, time periods, or methods of data collection

Interpretive rigor The degree to which the conclusions made on in study are rooted in actual results and rule out alternative inferences based on the same findings

Mixed method research A type of research design in which the findings of qualitative and quantitative methods/approaches are integrated in order to gain a fuller understanding of the phenomenon under investigation

Parallel mixed methods designs A type of mixed methods design in which two or more strands of the study are implemented from beginning to the end relatively independently at the same time (concurrently) or with a time lag

Sequential mixed methods design A type of mixed methods design in which the questions and data of one strand comes from the completion of a previous one. For example, a qualitative strand leading to a second quantitative strand of the study

Transcendence Creating a higher-order cognitive or theoretical framework that incorporates and connects two seemingly inconstant sets of cognitions or research findings

Mixed-methods research systematically integrates quantitative and qualitative designs, data, or results of data analysis to answer complex research questions that might benefit from such a combination. One of its major strengths is that it empowers the researcher to use the most appropriate techniques for answering the research questions. Using an integrated approach has gained increasing recognition in the past three decades. Such recognition is evident in a relatively large number of books (e.g., [Bamberger, 2000](#); [Creamer, 2018](#); [Creswell and Plano Clark, 2018](#); [Hesse-Biber and Johnson, 2015](#); [Nastasi and Hitchcock, 2016](#); [Newman and Benz, 1998](#); [Plano-Clark and Ivankova, 2016](#); [Plowright, 2011](#); [Ridenour and Newman, 2008](#)). Also see *Further Reading* below), numerous articles in professional and academic journals, dissertations, conference presentations, and a dedicated journal: *Journal of Mixed Methods Research*.

[†]Deceased author.

Traditionally, quantitatively oriented researchers have worked within the postpositivist worldview and are predominantly interested in collecting numerical data with structured data-collection methods. They analyze their data using statistical techniques, and maintain a degree of objectivity in their interpretations. On the other hand, qualitatively oriented researchers emerged as a distinct reaction to what was seen, in the 1980s and 1990s as the dominant and excessive practices of quantitative research. Qualitative scholars follow a constructivist worldview and are predominantly interested in collecting narrative data using open-ended (unstructured, emergent) methods. They typically analyze their data via thematic analysis, and make holistic, value-based, and subjective inferences on the basis of the findings. Although qualitative research is often (perhaps wrongly) assumed to be exploratory, both qualitative and quantitative research may be exploratory/descriptive or explanatory/confirmatory. The dichotomizing based on this aspect of research is too simplistic to be valid (see Bergman, 2008; Tashakkori et al., 2021).

Negating the qualitative–quantitative dichotomy, mixed methods have emerged as a third alternative—allowing researchers to take relative stances on a continuum between the two. Mixed-methods researchers identify themselves as integrative and eclectic, work primarily within a pragmatic worldview (Johnson et al., 2017) to answer broad and complex research questions, collect both narrative and numerical data, employ both structured and emergent designs, analyze their data both via statistical and content analysis, and make meta-inferences as answers to their research questions by integrating the inferences gleaned from their qualitative and quantitative findings.

Mixed-methods researchers combine the characteristics of quantitative and qualitative traditions that yield answers to questions that neither approach alone can answer. This potentially enables the researchers to examine research problems from multiple perspectives and types of (qualitative and quantitative) evidence from multiple sources, leading to conclusions that may be complementary and/or confirmatory. Using an integrated approach allows one to generate new theoretical explanations or expand the current ones, as well as test these understandings in a single study, if needed.

Despite the general assumption that integrated methodology has emerged as an alternative to the strict dichotomy of the qualitative and quantitative approaches of the past century, in truth, it is well rooted in the eclectic approaches of social/behavioral science researchers of the twentieth century (see Maxwell, 2016; Peltó, 2017; Tashakkori et al., 2015; Waszack and Sines, 2003 for more details). Consistent with this inherent return-to-roots characteristic, the dominant tenet of mixed-methods research is that the research purpose and question (not the paradigm) dictates the method of inquiry.

Mixed-methods scholars do not ignore the possible impact of investigators' culture, worldview, and sociopolitical context on choosing and formulating research questions. These influences are often inherent in the purposes of the study, which, in turn, shape the research questions, and eventually lead the researcher's methods of study. Understandably, only some research purposes and questions are amenable to mixed methods. Many others might be answered more parsimoniously by utilizing either qualitative or quantitative designs/approaches, as appropriate.

Purposes and utility of mixed-methods research

Admittedly, one of the main reasons for the popularity of mixed methods is that it is more effective for answering some research questions. These are often the type of questions that are not readily answerable by either qualitative or quantitative designs/approaches, or the purposes underlying them are not reachable via qualitative or quantitative orientations, "mental models (Greene, 2007)," and techniques. At least seven reasons are often given for using such methods:

- Complementarity: Using both mixed methods for the purpose of integrating two expertly different but related answers to the research question, one gleaned from using a qualitative and the other from a quantitative approach.
- Completeness: Using mixed methods in order to obtain a fuller understanding of the phenomenon under investigation. The expectation is that such a more complete understanding will emerge if the inferences from qualitative and quantitative strands of a study are merged effectively (such an effective integration has been called "integrative efficacy" of mixed methods inferences).
- Development: Mixed methods is conducted with the explicit (preplanned or emergent) purpose of obtaining research questions, sampling framework, or data sources of a second (e.g., a qualitative) strand of the study from the first strand (e.g., quantitative).
- Expansion: Same as "developmental" discussed above, but with the purpose of expanding the answers already obtained in a previous strand of a study.
- Corroboration/Confirmation: Utilizing integrated methods in order to evaluate the credibility of inferences obtained from a (qualitative or quantitative) strand of a study. The research questions of the first strand are often exploratory, while the research questions of the second strand are often explanatory/confirmatory.
- Compensation: Utilizing mixed methods with the express purpose of compensating the weaknesses of one approach (e.g., data correction errors/biases) with the strengths of the other.
- Diversity: Conducting mixed methods with the hope/purpose of comparing and contrasting divergent pictures of the same phenomenon. (Dialectic approach in which the answers are pitted against each other; Greene, 2007).

Research questions in mixed-methods research

As we discussed earlier, the main purpose for utilizing a mixed-methods design is to answer research questions that are not effectively answerable by either a qualitative or quantitative approach alone. Such questions often have interrelated components—each following qualitative or quantitative orientations, such as questions including “what and how”, or “what and why”.

Mixed-methods questions are often seen in two non-mutually exclusive ways: (1) an over-arching question that potentially includes both qualitative and quantitative type of sub-questions, or (2) a collection of separate qualitative and quantitative type of questions, followed by a question with regard to the nature of integration (i.e., how do the answers to the two types of questions relate to each other). Both types of questions may be pre-planned, or may emerge during the course of a study, such as in emergent sequential designs discussed below (also see [Plano-Clark and Badiee, 2010](#)).

The first variation is more frequently seen in the literature. For example, one might ask: “What are the effects of teaching method X on achievement and self-perceptions of groups A and B?” Such a question might be followed by three sub-questions asking: (1) “Are groups A and B different in achievement and self-perceptions?” (2) “What are the perceptions and constructions of participants in groups A and B regarding teaching method X and its impact?” and (3) “How does teaching method X work differently in the two groups (if it indeed did)?” The sub-questions are answered in separate strands (phases) of the study. Sub-questions are components of the same question (rather than addressing related phenomena).

The second approach mentioned above includes explicitly separate qualitative- and quantitative-type of questions that are inter-related (e.g., “Does teaching method X impact the achievement and self-perceptions of students in groups A and B differently?” and “What are the perceptions of parents regarding teaching method X?”). They are followed by an explicit question about the nature of integration (for more information, also see [Tashakkori and Creswell, 2007](#)). Examples of such preplanned or emergent questions include: “do the findings of the two strands of the study agree/converge?” or “how do the two sets of answers contribute to a comprehensive understanding of the phenomenon under study?” It could be said that these later questions only address the nature/quality of integration in mixed methods studies.

Design of mixed-methods studies

Our framework for classifying mixed-methods designs is to divide them into four families that simultaneously include the sequence of procedures (parallel, sequential, conversion, hybrid) and sources of information needed to answer the research questions (s). Within each family of designs, there is variation on the basis of units of study (same/sub-sample, multiple samples, multilevel samples). There are two basic questions that the investigator has to ask and answer when making decisions about mixed methods design of their study:

“Which sequence of steps will provide a better opportunity to answer my research questions: Collecting and analyzing data sequentially (sequential mixed-methods study) or in a parallel manner (parallel mixed-methods study)?” In parallel designs, two sets of data (qualitative and quantitative) are collected and analyzed independently at the same time or with a time lag (see [Fig. 1](#)), in order to answer an overarching mixed methods research question. Despite the relative independence of multiple strands,

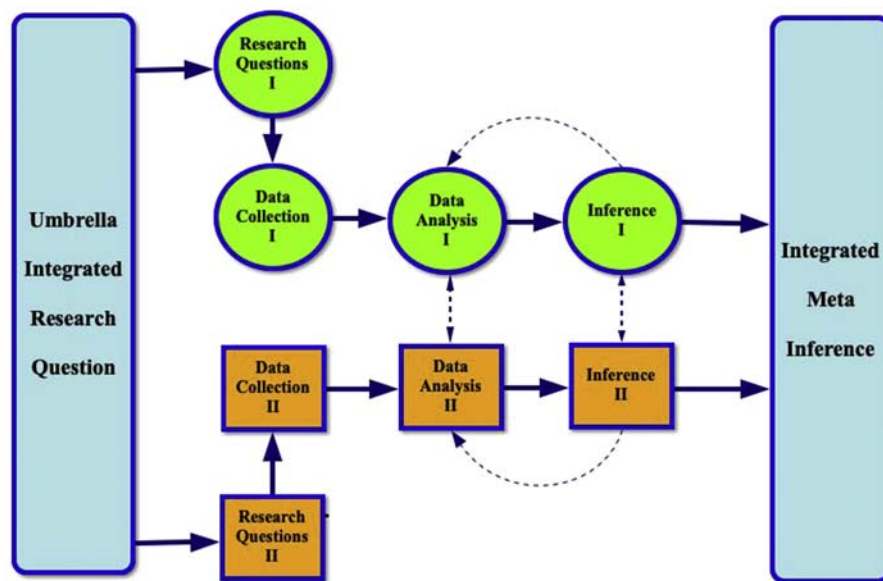


Fig. 1 Graphic presentation of parallel mixed-methods designs.

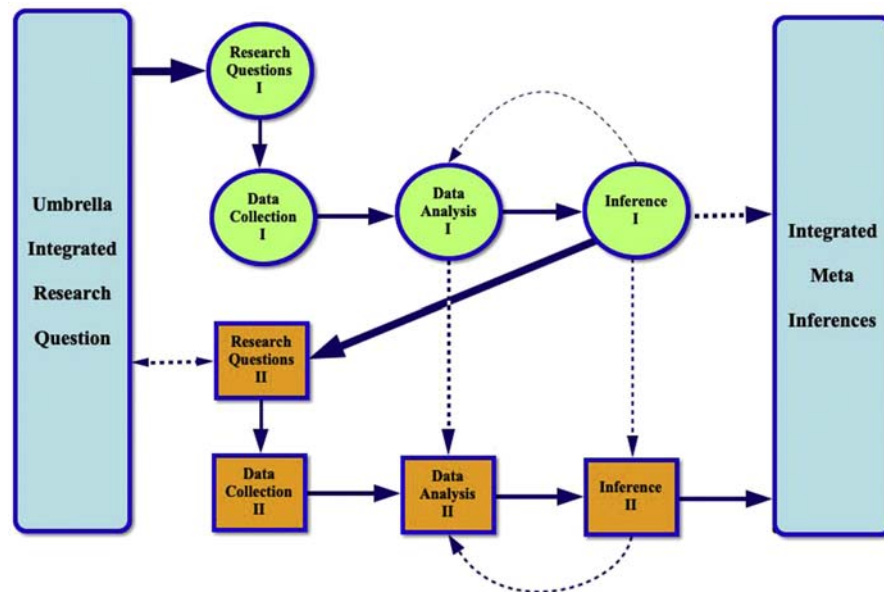


Fig. 2 Graphic presentation of a sequential mixed-methods design in which research questions of the second strand emerge from the findings of the first.

in these designs, investigators' knowledge of one strand (e.g., QUAL data collection or analysis) might lead to modification of another (e.g., QUAN data collection or analysis). In sequential designs, the second round of data collection and analysis is rooted in (initiated from, made possible by) the results of the first set (see Fig. 2), often in emergent manner.

An investigator might also ask "will I have an opportunity to convert my data from one type to another and reanalyze it in order to get more credible or alternative answers?" If the answer is affirmative, a conversion mixed-methods design is employed (see Fig. 3). For example, you may collect qualitative data (open-ended interviews, classroom observations, etc.), do a thematic analysis of the narrative, and make conclusions. In addition, you may also convert the obtained themes to numerical indicators that are analyzed statistically, leading to a new set of conclusions. In conversion mixed designs, the qualitative data are transformed (quantitized) into numerical indicators that may be analyzed separately using statistical techniques. Alternatively, in these studies, quantitative indicators (e.g., personality test results) might be transformed (qualitized) into qualitative data (e.g., narrative personality profiles) that are analyzed separately as if they originated from a separate sample (see Tashakkori et al., 2021 for more details). Obviously, the findings gleaned from both sets of data analysis are compared, contrasted, and integrated for better understanding of the phenomenon under study.

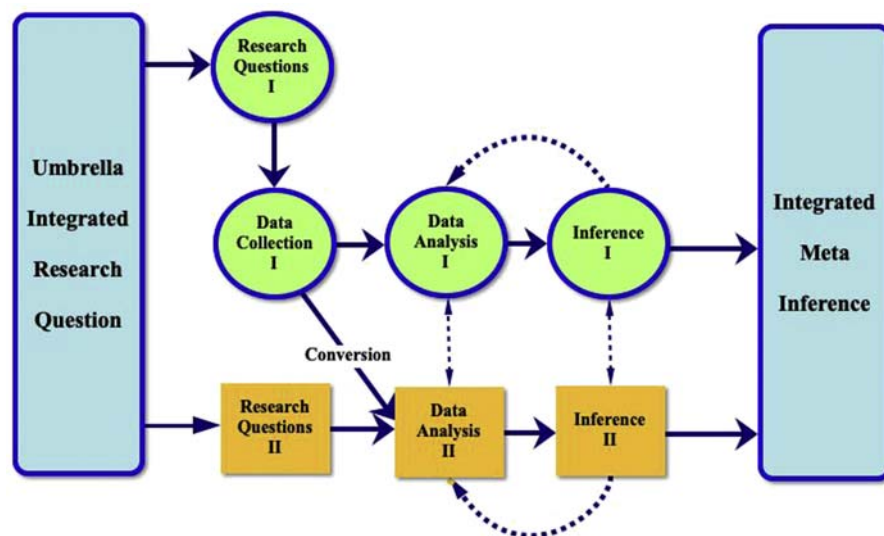


Fig. 3 Graphic presentation of conversion mixed-methods designs with multiple questions—one type of data that are also transformed and reanalyzed.

In practice, many mixed-methods studies use a combination of these three families in an iterative and dynamic manner, as needed. These studies have been labeled hybrid mixed methods studies (or, what in the past, was called fully integrated mixed-methods designs, [Teddle and Tashakkori, 2009](#)). This is the ideal design in mixed-methods studies because it allows the researcher to dynamically utilize qualitative and quantitative approaches in whatever combination/sequence needed, and use the insights from one type of data/process to modify/enhance the other within and across stages of the study. Hybrid mixed designs employ combinations of the other three families, with more than two strands.

In addition to the decisions about the sequence of strands and the possibility of data conversion, a mixed methods researcher must also make decisions about possible sources of data in order to find the best answers to the overarching research question. All qualitative and quantitative data might be obtained in the same sample. Alternatively, one type of data might be collected in a sub-sample of the other. This type of sampling (same, or sub-sample) leads to both type of data being available about some or all individuals in the study. Obviously, possibility of contamination (one data collection method creating reactivity in the sample) might be a disadvantage of this type of sampling (for example, focus groups following a survey study). Alternatively, qualitative and quantitative data may be collected in independent samples (i.e., two separate sub-samples). Obviously, in this type of sampling it is not possible to compare the qualitative and quantitative data/findings at the individual level. On the other hand, the possibility of either type of data collection impacting (i.e., contaminating) the other is also not present, potentially leading to more credible answers to the overarching research question.

Data-collection strategies in mixed-methods research

During the course of a mixed-methods study, an investigator collects a variety of narrative and numeric data, using a combination of qualitative (open-ended, naturalistic, and holistic) and quantitative (close-ended, structured, and variable-specific) approaches. It is possible to collect both qualitative and quantitative data in a single data-collection strategy. For example, in a mixed-methods study, observations may be recorded by taking extensive notes to record as many of the interactions as possible, while also using highly structured protocols yielding numerical data. Most observations include a combination of both unstructured (open-ended) recording and structured (closed-ended) items to record the behaviors/events. Similarly, interviews and questionnaires may be used to collect both types of data in a mixed-methods study. Open-ended interviews or questionnaires provide qualitative data, while structured (closed-ended) interviews and questionnaires provide quantitative data. Once again, most questionnaires and interviews in research do some of each. It should be noted that these mixed data are common in many studies that are not readily identified as mixed, such as in experiments, surveys, ethnographies, and case studies, but are not always analyzed fully (i.e., via both statistical and thematic analysis) that lead to pertinent inferences based on the results.

In mixed-methods studies, interviewers conduct funnel-sequenced interviews ([Teddle and Tashakkori, 2009](#)) that start from general questions/topics ("tell me about your school") and are gradually focused on more specific emerging or preplanned issues ("how do you like your math class?"). These interviews are conducted individually or in focus groups. Focus groups combine interviews, discussions, and observations of group interactions by the researcher(s), and, often, provide mixed data. Mixed questionnaires provide two types of data as well, although they are less dynamic (interactive) than interviews. Computerized questionnaires provide a greater potential for using both format.

In addition to using one method of data collection for obtaining both qualitative and quantitative data, mixed data might be obtained through multiple modes of data collection that each provide either qualitative or quantitative data (e.g., close-ended questionnaires among students and unstructured observations of teachers).

Data-analysis strategies in mixed-methods research

Highly consistent with the type of design, in mixed methods the narrative and numeric data may be analyzed relatively independently, or in an interdependent sequential manner. A detailed discussion of these analyses is beyond the scope of this article. Hence, we refer you to other sources (e.g., [Bazeley, 2018](#); [Creswell and Plano Clark, 2018](#); [Tashakkori et al., 2021](#)).

An example of data analysis in parallel mixed designs is when the investigator statistically compares an experimental and control group in a Randomized Control Trial (RCT) while also conducting a thematic analysis of post-experimental interviews (or open observation notes, if the individuals/groups were observed). An example for data analysis in sequential designs might be the formation of groups of people/settings on the basis of qualitative data-analysis results, and then comparing these groups statistically (analysis of variance, discriminant function analysis, etc.), using the available (or newly collected) quantitative data.

Integrated inferences in mixed-methods studies

As mentioned above, integration of the findings and conclusions are the primary objective of mixed methods studies (see [Plano-Clark, 2019](#)). Although a degree of agreement/consistency between the findings of qualitative and quantitative strands of a mixed-methods study might be desirable (leading to greater certainty about the conclusions), such consistency is not required. Integration in mixed methods denotes elaboration, complementarity, completeness, confirm/disconfirmation, and so forth. Lack of consistency

is examined carefully by mixed-methods researchers, to see if it might be a result of possible problems in data collection and analysis or the conclusions derived from the results. If such an examination does not reveal problems in the design, results, or interpretations, lack of agreement might indicate that the two sets of answers to the research question exist: The results reveal two different aspects of the same phenomenon, one set might provide the conditions for the applicability of the other, or there is a possibility that there are two plausible, but different, answers to the question (see [Erzberger and Kelle, 2003](#)). Lack of consistency (disagreement, discordance) might lead to *transcendence*, the creation of a new higher order theoretical explanation for the phenomenon or behavior.

Mixed methodologists have discussed the importance of developing quality criteria and audits that incorporate and expand the qualitative and quantitative concepts. For example, [Tashakkori and Teddlie \(1998\)](#) (see the section titled “Further reading”) have proposed using the term *inference quality* as a term to incorporate the term internal validity used by quantitative researchers and credibility used by qualitative researchers. Employing such an umbrella term gives mixed-methods researchers the opportunity to incorporate elements of quality assessment and strategies to deal with possible problems in making inferences (e.g., threats to quality) into a single integrative framework.

In such an integrative framework, qualitative concepts—such as prolonged engagement, peer debriefing, triangulation, audit trail, negative case analysis, and others—can be used to enhance the quality of inferences gleaned from quantitative findings. Similarly, concepts such as consistency between elements of the design, design quality, construct validity, investigator expectancy effects, and participant reactivity may be used as safeguards when making inferences on the basis of qualitative findings. In other words, using such an umbrella framework sensitizes the researcher to multiple perspectives when considering the quality/validity of their interpretations. A similar logic is applicable to evaluating and enhancing the quality of data that the findings and inferences are based on (i.e., data quality).

Within this integrative framework, conclusions that are made on the basis of the results of data analysis are evaluated in terms of (1) inference quality and, (2) inference transferability. We will discuss transferability (i.e., generalizability, external validity, transferability in QUAL research) below. Quality of inferences/conclusions are dependent on (a) quality of the research design that leads to the results of data analysis (i.e., “Design Quality”), and (b) quality of the conclusions that are made on the basis of such results (i.e., “Interpretive Rigor”).

Design quality is the degree to which: (I) the process of data collection and analysis is appropriate for answering the research question set forth (Design Suitability), (II) the process is implemented with fidelity (Implementation Fidelity), (III) the data are collected and analyzed appropriately (Analysis Adequacy), and (IV) all the design components are consistent with each other (Within-Design Consistency).

Interpretive rigor addresses the effectiveness of making interpretations and conclusions on the basis of the obtained results. At its basic form, it might be conceptualized as the degree to which alternative plausible explanations for the obtained results can be discounted or ruled out. Such rigor depends on meeting a number of criteria: (I) the degree to which the conclusions are consistent with the findings that they are based on (Interpretive Consistency), (II) there is agreement between interpreters such as peers, scientific community and/or the participants in the study (Interpretive Agreement), (III) the interpretations are consistent with the state of knowledge about the phenomenon (Theoretical Consistency), (IV) and each inference is distinctly more credible than other possible interpretations of the same results (Interpretive Distinctiveness). These standards might be applicable to any research project (QUAL or QUAN) or strand of a mixed methods study. Two other aspects of interpretive rigor are only applicable to mixed and multiple methods research: (e) the degree to which the results and inferences gleaned from multiple strands of the study are integrated as answers to the overarching question (Integrative Efficacy) and (f) the degree to which such final answers are effective answers to the umbrella integrated (mixed) research question (s) and purposes set forth (Interpretive Correspondence).

Please see [Tashakkori et al. \(2021\)](#) for details with regard to the components of these two quality criteria.

Transferability of inferences in mixed-methods studies

Inference transferability incorporates the qualitative concept of transferability and quantitative concept of generalizability or external validity. It is broadly defined as the degree of applicability or generalizability of conclusions of a study to other individuals or entities (population transferability), other settings or situations (ecological transferability), other time periods (temporal transferability), or other methods of observation/measurement (operational transferability).

We have differentiated the transferability of conclusions from the criteria about quality/rigor of such conclusion because any conclusion in a study might be transferable to a degree to similar other setting and people. The degree of transferability is dependent on the similarity of the “sending” and “receiving contexts.” Although the researcher tries to maximize the degree of transferability of the conclusions made in the study, the ultimate evaluator of that degree is the consumer of research. Providing in-depth details of the study assists such a consumer in making judgments about the degree of transferability.

Conclusions

Mixed methods research is arguably closer to our daily problem-solving efforts than the two approaches (QUAL, QUAN) that it attempts to integrate ([Tashakkori and Teddlie, 2010](#)). Mixed methodology puts each component of a research project (purposes,

questions, data, data analysis, design, inferences, etc.) on a continuum between alternative qualitative and quantitative approaches/conceptualizations. Mixed-methods research is often based on the premise that the research purpose/question—not the paradigm—dictates the method of inquiry.

An advantage of mixed-methods is that it encourages researchers to think in terms of multiple continua rather than a single categorical dichotomy of qualitative and quantitative approaches, employ multiple perspectives as needed, use multiple methods that are rooted in qualitative and quantitative traditions, and interpret the findings freely without the imposed constraints of either of the two broad approaches to research.

A challenge for the mixed-method researchers is that they must learn specific data collection, analysis, and interpretation techniques that have been traditionally identified as being qualitative or quantitative. They must also learn how to integrate understandings gleaned from these complementary or competing methods in order to reach meta-inferences that better inform other professionals and stakeholders. There is an underlying assumption that the more perspectives one can get of a particular situation, the more complete the understandings are likely to be (i.e., a gestalt, or whole insight, that is bigger than the sum of its component parts). The value of mixed methods to researchers may not solely be in empowering them to use all possible tools for answering their research question, but also in becoming better informed as a result of integrating the findings of these methods.

Perhaps one of the challenges facing mixed methods is linked to its aging. As a research approach with a community of scholars expanding in at least three decades, mixed methods is reaching a mature stage. Nevertheless, it is still struggling with basic issues of conceptualization/definition. It is not clear if mixed methods denotes research projects (a) that utilize two types of design (e.g., experimental and case study) to answer a research question, (b) include two type of data that are collected/converted and/or analyzed separately (e.g., numeric and narrative data), or (c) are mixed in methods of data analysis and inferences based on the results. The least controversial, perhaps, is conceptualizing mixed methods in terms of two or more types of designs/methodologies, at least one with qualitative approach/perspective (e.g., ex-post-facto, survey, quasi-experiment) and the other from qualitative (e.g., ethnography, case study, phenomenology). Conceptualizing mixed methods in terms of two types of data is the most common and also perhaps the most controversial. For example, is an experiment with open and structured (e.g., a “funnel-sequenced”) interview leading to two types of data that are analyzed accordingly a mixed design? (or perhaps, should it be called a mixed experimental design?) Similarly, should ethnographic studies with both types of data be called mixed? (e.g., see [Pelto, 2017](#) for examples). Unfortunately, definitions of mixed methods in the literature (especially those defining in terms of “data”) are fraught with ambiguity from this point of view, and need further consensus.

Another interesting dilemma facing the mixed method community goes back to its roots. Mixed methods emerged in 1980s as a third community negating the dichotomy of QUAL and QUAN approaches. In doing so, it was initially an attempt to bring the other two communities together, providing an umbrella. However, it is in danger of becoming a distinct category itself, expanding the dichotomy into a trichotomy of disparate communities. This may be seen in textbooks in research methodology, with distinct walls between sections that are devoted to QUAL, QUAN, and mixed methods, rather than teaching research as mixed (conceptualized in terms of a combination of QUAL and QUAN-types of designs). There are few exceptions to this trend in the textbooks that discuss independent research designs (ethnography, ex-post-facto, experiment, case study, etc.), each capable of collecting and analyzing numeric and narrative data in pre-planned or emergent ways. An alternative way to think about mixed methods is to conceptualize it in terms of continuity between many components of ANY research project (please see multi-dimensional view of mixed methods in [Tashakkori et al., 2021](#), p. 44). Such a conceptualization assumes that most research projects are mixed in some form and to some degree in one or more of their components (worldview, purpose, question, data, data collection methods, data analysis techniques, inference), each anchored in both (QUAL and QUAN) traditions/approaches.

References

- Bamberger, M. (Ed.), 2000. Integrating Quantitative and Qualitative Research in Development Projects. World Bank, Washington, DC.
- Bazeley, P., 2018. Integrating Analyses in Mixed Methods Research. Sage, Thousand Oaks, CA.
- Bergman, M.M., 2008. The straw man of qualitative–quantitative divide and their influence on mixed methods research. In: Bergman, M.M. (Ed.), *Advances in Mixed Methods Research: Theories and Applications*. Sage, London, pp. 11–21.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. Sage, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Erzberger, C., Kelle, U., 2003. Making inferences in mixed methods: the rules of integration. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 457–490.
- Greene, J.C., 2007. *Mixing Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Hesse-Biber, S., Johnson, R.B., 2015. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press Publications, New York, NY.
- Johnson, R.B., Onwuegbuzie, A.J., de Waal, C., Stefurak, T., Hildebrand, D., 2017. Unpacking pragmatism for mixed methods research: the philosophies of Peirce, James, Dewey, and Rorty. In: *The BERA/SAGE Handbook of Educational Research*. Sage, London, England, pp. 259–279.
- Maxwell, J.A., 2016. Expanding the history and range of mixed methods research. *J. Mix. Methods Res.* 10 (1), 12–27. <https://doi.org/10.1177/1558689815571132>.
- Nastasi, B.K., Hitchcock, J., 2016. *Mixed Methods Research and Culture-Specific Interventions: Program Design and Evaluation*. Sage, Thousand Oaks, CA.
- Newman, I., Benz, C.R., 1998. *Qualitative–Quantitative Research Methodology: Exploring the Interactive Continuum*. University of Illinois Press, Carbondale, IL.
- Pelto, P.J., 2017. *Mixed Methods in Ethnographic Research: Historical Perspectives*. Routledge, New York, NY.
- Plano Clark, V.L., Badiie, M., 2010. Research questions in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 275–304.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide for the Field*. Sage, Thousand Oaks, Los Angeles, CA.
- Plano Clark, V.L., 2019. Meaningful integration within mixed methods studies: identifying why, what, when, and how. *Contemp. Educ. Psychol.* 57, 106–111.

- Plowright, D., 2011. *Using Mixed Methods: Frameworks for an Integrated Methodology*. Sage, London, England.
- Ridenour, C.S., Newman, I., 2008. *Mixed Methods Research: Exploring the Interactive Continuum*. Southern Illinois University Press, Carbondale, IL.
- Tashakkori, A., Creswell, J., 2007. Exploring the nature of research questions in mixed methods research. *J. Mix. Methods Res.* 1 (3), 207–211.
- Tashakkori, A., Teddlie, C., 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Sage, Thousand, CA.
- Tashakkori, A., Teddlie, C., 2010. Putting the human back in “human research methodology”: the researcher in mixed methods research. *J. Mix. Methods Res.* 4 (4), 271–277.
- Tashakkori, A., Teddlie, C., Sines, M.C., 2015. Utilizing mixed methods in psychological research. In: Schinka, J.A., Velicer, W.F. (Eds.), *Handbook of Psychology: Research Methods in Psychology*, vol. 2. Wiley & Sons, Inc, New York, NY, pp. 428–450.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of mixed methods research: integrating quantitative and qualitative approaches in the social and behavioral sciences*, 2nd Edition. Sage, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Thousand Oaks, CA.
- Waszak, C., Sines, M., 2003. Mixed methods in psychological research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 557–576.

Further reading

- Bergman, M.M. (Ed.), 2008. *Advances in Mixed Methods Research: Theories and Applications*. Sage, London.
- Mertens, D.M., 2020. *Research and Evaluation in Education and Psychology: Integrating Diversity With Quantitative, Qualitative, and Mixed Methods*, fifth ed. Sage, Thousand Oaks, CA.
- Morse, J., 2018. Reframing rigor in qualitative inquiry. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*. Sage, Thousand Oaks, Los Angeles, CA, pp. 796–817.
- Tashakkori, A., Teddlie, C., 2003. *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA.
- Tashakkori, A., Teddlie, C. (Eds.), 2010. *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA.

Philosophical underpinnings of mixed methods research in education

Peggy Shannon-Baker^a, ^a Georgia Southern University, Statesboro, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	380
Philosophical paradigm	380
Elements of a philosophical paradigm	381
Ontology	382
Epistemology	382
Axiology	382
Methodology	382
Framing mixed methods research and philosophical paradigms	383
Philosophical paradigms in mixed methods research in education	383
Positivism and postpositivism	383
Critical realism	384
Constructivism	385
Pragmatism	385
Transformative(-emancipation)	385
Postmodernism	386
Dialectical pluralism	386
Yinyang philosophy	387
Indigenous philosophies	387
Conclusions	387
References	387

Introduction

Early in the formalization of mixed methods research as a field, Teddlie and Tashakkori (2003) identified that the “paradigmatic foundations” of mixed methods research was a key issue for the field. Nearly 2 decades later, this topic is still discussed (cf. Coates, 2020). Identifying one’s philosophical underpinnings is an important part of making the mixed methods research process more credible, transparent, and trustworthy (Howes, 2017). This discussion provides a justification for the decisions made during this process (Harrits, 2011; Howes, 2017). Articulating a clear philosophical paradigm can make one’s stance and research more accessible to audiences across disciplinary, national, and language boundaries (Fetters and Molina-Azorin, 2019).

The purpose of this chapter is to introduce the philosophical underpinnings of mixed methods research in educational research. First, I provide a definition for philosophical paradigm along with how it differs from other frameworks such as a methodological paradigm (e.g., qualitative research), theoretical framework, or conceptual framework. Next, I outline the four main elements to a philosophical paradigm: ontology, epistemology, axiology, and methodology. After defining mixed methods research, I then discuss several philosophical paradigms that are used in mixed methods research, describing each of the four elements within that paradigm and highlighting published examples of mixed methods research in education using that philosophy. The paradigms I describe in detail are positivism, postpositivism, critical realism, constructivism, pragmatism, transformative-emancipation, postmodernism, and dialectical pluralism. I also describe two meta-paradigms, yinyang philosophy and Indigenous philosophies, based on emerging research in mixed methods about these frameworks.

Philosophical paradigm

A *paradigm* refers to a community’s set of shared beliefs, values, and practices (Shannon-Baker, 2016). A *philosophical paradigm* refers to a set of beliefs, values, and practices about the nature of reality, how knowledge can be generated, and who generates it about that reality (Plano Clark and Ivankova, 2016; Shannon-Baker, 2016). Researchers’ philosophical paradigms are influenced by their education and training, discipline, and sociohistorical contexts (Creswell and Poth, 2018). Other researchers have referred to philosophical paradigms as *philosophical stances* (Greene, 2006; Greene and Hall, 2010; Maxwell and Mittapalli, 2010), *philosophical assumptions* (Plano Clark and Ivankova, 2016), or part of a *mental model* including philosophical assumptions and predispositions to particular methodological approaches (Greene, 2007).

A philosophical paradigm might be connected to but is different from a *methodological paradigm*, which indicates a set of values, practices, and expectations for rigor and quality that directly impact one's research methods. For example, qualitative research and quantitative research have been referred to as paradigms (e.g., Collins et al., 2006; Cooper and Hall, 2016; Corr et al., 2020). The *paradigm wars* reflect the conflict between researchers entrenched in philosophical and/or methodological paradigms and affiliated worldviews (Alise and Teddlie, 2010; Ponce et al., 2020; Shannon-Baker, 2016). Similarly, a philosophical paradigm differs from a *theoretical framework*, which identifies a set of assumptions about a particular topic or phenomenon (Plano Clark and Ivankova, 2016), and a *conceptual framework*, which defines key constructs related to a research study and how those constructs have been studied in the past.

A philosophical paradigm and research methodology might be aligned but these are neither inseparable nor synonymous. Certain research practices might align with a specific methodology and philosophical paradigm (e.g., strict structured observation of the frequency of events → quantitative research → postpositivism) whereas other methods or methodologies might align with multiple frameworks. Researchers can intentionally choose to combine, mix, or otherwise integrate multiple elements from several philosophical paradigms, referred to as using a *multiparadigmatic approach*, or use specific paradigmatic frames for a given research problem or setting (Greene, 2007; Reichardt and Cook, 1979; Shannon-Baker, 2016). Some scholars argue that philosophical paradigms are incommensurable with one another (Lincoln and Guba, 2011) whereas others critique this purist approach (Greene, 2007; Tashakkori et al., 2021). Additionally, others present these paradigms as more on a spectrum rather than strict, exclusionary and static categories (Hammersley, 1992; Shannon-Baker, 2016) where in some cases, a single paradigm could have multiple offshoots such as how there are multiple forms of pragmatism (Johnson et al., 2017).

Some researchers argue that they can conduct studies without influence from or specific attention to their philosophical paradigms, referred to as an *aparadigmatic* stance (Greene, 2007; Tashakkori et al., 2021). For example, this stance might be used (consciously or unconsciously) in relatively small and applied research studies where evaluating a specific context is seen as more important than attending to seemingly abstract philosophical notions (Greene, 2007; Tashakkori et al., 2021). However, some researchers have argued that philosophical assumptions, beliefs, and practices cannot be separated from one's research process (Shannon-Baker, 2016).

Educational researchers, when they explicitly identify their philosophical paradigm, may indicate this underpinning in a frameworks section (where theoretical and conceptual frameworks may also be described), in a methods section, and/or in a section about the researcher's *positionality*, which is where the researcher names their specific contexts, personal experiences, identities, personal and professional beliefs, predispositions, their role in the research process, and other aspects that might have influenced the study (Berger, 2015; Creswell and Poth, 2018). This positionality statement is an important place for researchers to demonstrate their critical reflexivity about their role and influence in a research study (Berger, 2015).

Elements of a philosophical paradigm

A philosophical research paradigm includes ontological, epistemological, axiological, and methodological components that are distinct but interrelated that then influence one's research methods and practices (see Fig. 1). Although these components might appear to be linear in influence, they are interrelated, depicted using dashed lines between each component and cyclical facing arrows.

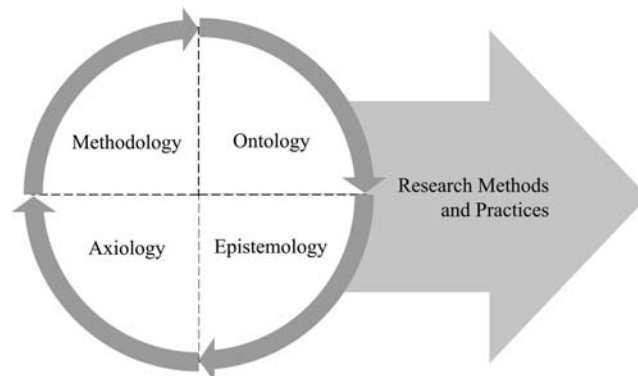


Fig. 1 The relationships between ontology, epistemology, axiology, and methodology, and their influence on research methods and practices within a given research philosophy.

Ontology

Ontology refers to one's beliefs, understandings, and assumptions about the nature of reality (Mason, 2002; Tashakkori et al., 2021). A key ontological question is "What is the form and nature of the social world" (Waring, 2017, p. 16). Considering one's ontology in the research entails asking what groups of people, institutions, and social processes are connected to the study (Mason, 2002). Ontology influences the research process through one's methods and how findings are represented (Alerby, 2000). For example, a researcher who believes that reality is complex and depends on an individual's perception would use methods that emphasize individual interpretations of their experiences (Alerby, 2000). Alerby (2000), for example, included a wide representation of children during sampling knowing that their individual experiences of their realities as children differs among them. Additionally, a philosophical paradigm that emphasizes multiple realities and how individuals construct their own perceptions of reality would emphasize representing findings from those multiple perspectives (Creswell and Poth, 2018).

Epistemology

What one understands about reality (ontology) influences our understandings about epistemology (Howes, 2017). *Epistemology* refers to how a person or group creates, obtains, and measures knowledge as well as the relationship between an individual/group and knowledge (Mason, 2002; Tashakkori et al., 2021). Epistemology identifies what is considered "knowledge" and what is considered "evidence" of knowledge (Creswell and Poth, 2018). A key epistemological question is "How can what is assumed to exist be known" (Waring, 2017, p. 16). One's epistemology also influences the research process. For example, epistemology influences how one frames research questions and hypotheses; if one's epistemology states that personal reflections about an experience are valid sources of knowledge, they can frame research questions based on those reflections (Creswell and Poth, 2018). Critiques from reviewers about whether a sample size is large enough or if quotes from participants provide rich details are critiques about evidence and therefore critiques indicating the reviewers' epistemological beliefs about evidence. In another example, Garcia and Mayorga (2018) described how Eurocentric and white epistemologies have heavily influenced the use of quantitative approaches as the only valid source of evidence about a phenomenon. They instead position their epistemologies as Chicana feminist and Latino scholars respectively as important foundations for their critique of how race and ethnicity has been measured in The Freshmen Survey used in higher education research. One's epistemology influences how one frames research questions and hypotheses, what data is collected and how, how it is analyzed, and how the methods and findings are represented.

Axiology

The axiology element of a philosophical paradigm blends how one understands reality (ontology) and how one creates and measures knowledge (epistemology). *Axiology* refers to the specific values one has about research, knowledge, and research methods and practices and the influence those values play in the research process (Tashakkori et al., 2021). Axiology recognizes that researchers make value-based judgments about research methods and ethics (Lincoln and Guba, 2011; Sandelowski, 2003). Researchers who are more aware of their philosophical paradigms can intentionally use their paradigms to enhance the authenticity of their work (Hesse-Biber, 2010; Lincoln et al., 2011).

Despite how some researchers overlook axiology (Hesse-Biber, 2010; Mertens et al., 2010), this element of philosophical paradigms influences much of the research process. For example, studies being conducted with humans and non-human animals need to account for research ethics in seeking institutional review board approval for their study (Biddle and Schafft, 2015). Axiology informs a researcher about which research questions are worth asking (Creswell and Poth, 2018). Camacho (2020), who studied academic migrants' experiences with advocating for better working conditions, emphasized the axiological importance of engaging in emancipatory research practices such as asking how one's data analysis procedures help to advance social justice initiatives when working with marginalized communities. The selection of specific quotes to represent one's themes represents value-based judgments about which quotes serve as good evidence (Denzin, 1989). Even whether to include a discussion of one's values and positionality in a study is influenced by their axiology, or their value for such discussions, as well as whether these contexts are seen as influencing the study.

Methodology

A research *methodology* refers to an overarching frame that guides values, practices, and decisions about research methods whereas *research methods* are the specific practices and procedures used in a study (Howes, 2017). A key question for one's methodology is "What procedure or logic should be followed" (Waring, 2017, p. 16) in the study. A key question for one's methods is "What techniques of data collect [and analysis] should be used" (Waring, 2017, p. 16). Research methods relate directly to one's research purpose and questions. Methodology is separate but related to one's philosophical paradigm. A given paradigm does not in most cases assume or align with a single research methodology; many philosophical paradigms align with multiple research methodologies and research methods.

Researchers can use their philosophical paradigm to evaluate the quality of their methods and interpretations (Howes, 2017). The process for engaging in one's philosophical paradigm has been presented as linear, using one's perspectives and philosophical assumptions to guide the research design process (e.g., Creswell and Poth, 2018), as well as ongoing and recursive (Howes, 2017). Howes (2017) for example traced how four domains of methodology impacted and connected to her work: philosophical paradigms, inquiry logics that inform the methodology, specific considerations that guide the methods used, and sociopolitical commitments to the context, researcher's own values, and research ethics.

Framing mixed methods research and philosophical paradigms

Mixed methods research is a research methodology that intentionally mixes, combines, and/or integrates more than one methodological approach in a single study (Shannon-Baker, 2016). Some mixed methods scholars argue that this mixing can also occur at the philosophical level (e.g., Corr et al., 2020; Greene, 2007). Mixed methods research has been widely applied in educational research worldwide (cf. Shannon-Baker, 2022).

Mixed methods research takes place within a socioecological framework where the researcher's own philosophical paradigm(s) are part of their personal contexts; these contexts influence the mixed methods research process (Plano Clark and Ivankova, 2016). These contexts can provide a kind of "feedback loop" to the mixed methods research process (Tashakkori et al., 2021, p. 23). Specific to education, the contexts that influence a mixed methods research study include but are not limited to: national curriculum; national or other standardized assessments; evaluation measures of schools, administrators, teachers, and students; teacher and leadership training requirements; shifting to online mediums in the wake of the COVID-19 pandemic and natural disasters; the form and function of schooling at the primary, secondary, postsecondary, and postgraduate levels; etc.

Mixed methods researchers can intentionally operationalize a philosophical paradigm that aligns with their planned research design (Creswell and Plano Clark, 2017). Understanding the differences and interconnections between one's philosophical paradigm and their research process is an important element of mixed methods research. Stakeholders may have different philosophical assumptions which impact their beliefs, values, and preferences for particular types of methods and data. As a result, mixed methods researchers can connect to these differences through the use of culturally appropriate methods and representations for specific stakeholders (cf. Shannon-Baker and Martinez, 2022). This is an important affordance to using mixed methods research: being able to speak to different stakeholders across diverse worldviews (Hesse-Biber, 2010).

Philosophical paradigms in mixed methods research in education

Mixed methods research has been conducted in relationship to many different philosophical paradigms, which is common for young research fields and approaches (Tashakkori et al., 2021). There are several philosophical paradigms that mixed methods researchers operationalize in educational research. These include positivism, postpositivism, critical realism, constructivism, pragmatism, postmodernism, and transformative-emancipation. Each has its own ontological, epistemological, and axiological beliefs that influence the methodological approaches used within that paradigm though many of these paradigms overlap with one another (see Fig. 2). Other philosophical paradigms, such as dialectical pluralism, yinyang philosophy, and various Indigenous philosophies, are based on beliefs that not only overlap with other philosophies but also describe the relationship between these frames.

Other publications additionally offer helpful comparisons of some of these philosophies (e.g., Plano Clark and Ivankova, 2016; Shannon-Baker, 2016; Tashakkori et al., 2021). It is important to note that although each paradigm is described separately below, there is overlap in the ontological, epistemological, axiological, and methodological features, which is why Fig. 2 portrays their comparisons as a spectrum rather than distinct, static categories. I represented dialectical pluralism, yinyang philosophy, and Indigenous philosophies as metaparadigms, forming a dotted box around the other philosophies, for how these paradigms describe relationships between paradigms; ontological, epistemological, and axiological beliefs; and methodological practices.

Positivism and postpositivism

Positivism is a philosophical paradigm that is centered on the ontological belief that reality is measurable and encompasses only what one can directly observe (Lincoln and Guba, 1985; Tashakkori et al., 2021). Axiologically, positivists tend to believe that research can and should be without value judgments, thereby emphasizing researcher objectivity (Tashakkori et al., 2021). This value centers the researcher as the one with power in the study (as opposed to the participants or others) (Hesse-Biber, 2017). The role of research in this paradigm is to observe and measure patterns in reality to test hypotheses or make predictions about reality using deductive reasoning and quantitative methods (Hesse-Biber, 2017; Tashakkori et al., 2021). They also use other methods that emphasize researcher objectivity through standardization of research practices, data collection, and analysis (Hesse-Biber, 2017).

Postpositivism is a more recently developed philosophical paradigm that is largely based in positivism but recognizes the impact of an individual's experiences and values on their perceptions of reality and therefore the research process (Plano Clark and Ivankova, 2016). Ontologically, postpositivists assume that there is a single reality that can be approximated by removing the researcher's influence on a study and creating more standardized research methods (Tashakkori et al., 2021). Postpositivism

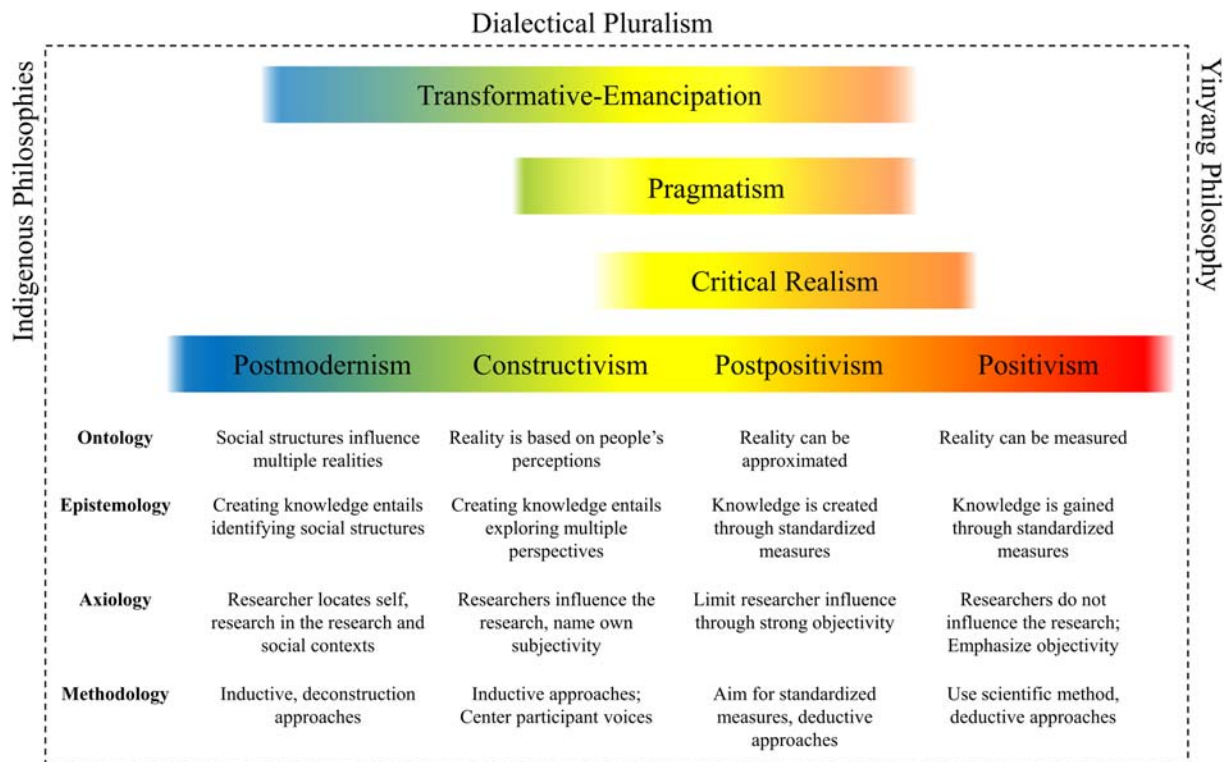


Fig. 2 A spectrum of ontological, epistemological, axiological, and methodological features of different philosophical paradigms.

emphasizes researcher objectivity and independence from the research process and participants. Researchers who hold this philosophical paradigm often seek out methods that remove them from the process, such as full observation and bracketing. This paradigm is often associated with quantitative research (Phillips and Burbules, 2000), but may also be applied in qualitative research (Creswell and Poth, 2018; Hesse-Biber, 2017). Mixed methods researchers with this paradigm “often prioritize the quantitative methods, ‘quantitize’ qualitative findings for statistical analysis, and emphasize causal inferences and generalizations” (Plano Clark and Ivankova, 2016, p. 198). For example, De Lisle (2013) combined critical realism with postpositivism to frame his evaluation of the continuous assessment programs in Trinidad and Tobago. These assessment programs are used to provide summative and formative assessments, policies, and procedures for in-school teachers. Postpositivism in particular was used to inform the quantitative phase of the evaluation.

Critical realism

Critical realism is aligned closely with the ontological beliefs of postpositivism that reality can be approximated (Guba, 1990; Maxwell and Mittapalli, 2010; Tashakkori et al., 2021). Critical realists’ ontology is based on the belief that reality includes what one perceives as well as “mental phenomena” (Plano Clark and Ivankova, 2016, p. 201). This ontology recognizes the influence of contexts thereby limiting complete objectivity (Shannon-Baker, 2016). Therefore, epistemologically, critical realists argue that we should study peoples’ interpretations of the world in order to understand their experiences (Danermark et al., 2002). Although some describe critical realism based more on its ontological differences with other frameworks (e.g., Tashakkori et al., 2021), some educational researchers and mixed methods scholars alike have adopted this paradigm in their work.

Mixed methods researchers use this stance to explore divergent findings across data sets or when compared to relevant literature, and recognize that a researcher’s values and ethics influence the research process (Plano Clark and Ivankova, 2016; Shannon-Baker, 2016). Critical realists emphasize the impact of contexts in the mixed methods research process, thereby aiming to include a wide representation of voices as well as accounting for the influence of contexts during interpretation (Plano Clark and Ivankova, 2016). In addition to De Lisle (2013) who combined critical realism with postpositivism, another example of applying a critical realist approach in mixed methods educational research comes from Ubha and Cahill (2014) who studied primary school children’s relationships with classmates, teachers, and other adults in the UK. They took a “critical realist position in order to use a ‘triangulation’ strategy” to bring together their various data sets including questionnaires, quantitative scales, and interviews (p. 278). In their case, it was critical realism that necessitated bringing together multiple data sets to examine children’s relationships.

Constructivism

Constructivism is a philosophical paradigm that ontologically emphasizes how an individual actively constructs their own notions of reality through their cognition (Lincoln and Guba, 1985; Schwandt, 1997) resulting in the existence of multiple realities. A researcher who believes in this paradigm emphasizes participants' own constructions, descriptions, and narrations of their lived experiences as well as the belief that knowledge is co-constructed between researcher and participant (Tashakkori et al., 2021). These researchers also recognize that they cannot be separated from the research process and instead see this as a positive influence on the research (Lincoln and Guba, 1985; Tashakkori et al., 2021). Some scholars who use constructivism have argued it is incompatible with other paradigms (e.g., Lincoln and Guba, 1985). However, this paradigm not only connects to other paradigms often used in mixed methods research (such as pragmatism and transformative-emancipation) but is also used on its own.

Methodologically, constructivism is often associated with qualitative research (Plano Clark and Ivankova, 2016; Tashakkori et al., 2021). Constructivist researchers emphasize inductive research methods (Tashakkori et al., 2021). Mixed methods researchers from this paradigm "often prioritize the qualitative methods, limit the research to a small number of information-rich cases, and emphasize each researcher's reflexivity and interpretations" (Plano Clark and Ivankova, 2016, p. 199). For example, Strayhorn (2015) studied the Black men students' readiness and preparation for a career in a science, technology, engineering, or mathematics field at Historically Black Colleges and Universities and predominantly white institutions in the USA. He used online surveys followed by individual interviews with students. Strayhorn identified that this study came from "an epistemological approach anchored in [a] pragmatic, constructivist tradition" (p. 50) that recognized the contextual influencing factors in these men's experiences. He stated that he used a constructivist approach because it aligned with his own ontology, epistemology, and his "own commitments to social justice" (p. 50). In another example, Smeda et al. (2014) applied constructivist learning principles, which emphasize students' active engagement in the construction of their own knowledge, in their multi-site case study in Australia on the impact of digital storytelling on students' engagement and learning. Teachers in the study were trained on digital storytelling, which centers the storyteller sharing their own experiences through digital media. Their use of digital storytelling in their teaching was then evaluated using a quantitative rubric, interviews, and observations. Smeda et al. found that using digital storytelling supported constructivist approaches to learning.

Pragmatism

At its core, *pragmatism* is a philosophical paradigm that aims to identify practical solutions to real world problems (Tashakkori et al., 2021). Pragmatic beliefs range on the spectrum between constructivism, critical realism, and postpositivism. Similar to constructivists, pragmatists ontologically contend that perspectives about reality are wide-ranging and situated in current understandings and experiences (Johnson and Onwuegbuzie, 2004; Plano Clark and Ivankova, 2016; Strayhorn, 2015). Epistemologically, they contend that researchers gain knowledge through a combination of independent observations and subjective interpretations with the aim of investigating the consequences or influence of one's understandings of reality (Johnson and Onwuegbuzie, 2004; Morgan, 2007; Plano Clark and Ivankova, 2016). Axiologically, pragmatists value the influence of subjectivities in the research process, but otherwise methodologically choose whether to embrace this influence or mitigate it (Johnson and Onwuegbuzie, 2004; Plano Clark and Ivankova, 2016). Methodologically, pragmatists' decisions in the research process are guided by the central goal of wanting to identify changes to address social problems and thus use any methods that best address the research purpose and questions (Tashakkori et al., 2021).

Johnson et al. (2017) demonstrated how classical pragmatists such as John Dewey and others align with particular equal, quantitative, or qualitative priority mixed methods designs.

Classical pragmatism aimed to find a "middle ground" between those who recognized the influence of philosophical paradigms and those who were skeptical of the influence (Johnson and Onwuegbuzie, 2004, p. 18). This belief persists in pragmatists' rejection of the incommensurability of qualitative and quantitative methods; pragmatic mixed methods researchers instead believe in the complementarity of these approaches (Tashakkori et al., 2021). For example, in their investigation of the impact of a four-week study abroad program in China for teachers' intercultural competence, He et al. (2017) applied pragmatism in rationalizing why mixed methods research was necessary. They argued that this paradigm was their foundation for selecting multiple methods "to corroborate and elaborate" on one another (p. 150). Strayhorn (2015) similarly applied pragmatism in his mixed methods educational research study. Pragmatism is often presented as a primary philosophical paradigm for mixed methods research (Johnson et al., 2007; Johnson and Onwuegbuzie, 2004; Tashakkori et al., 2021) because it "offers an epistemological justification [...] for mixing approaches and methods" (Johnson et al., 2007, p. 125). However, as this chapter demonstrates, mixed methods educational researchers apply many different philosophical paradigms in their work.

Transformative(-emancipation)

Transformative-emancipation is a philosophical paradigm that ultimately aims to use research to bring about positive social change to address injustices and inequity (Mertens, 2003, 2007). Although commonly referred to as the *transformative* paradigm, Mertens (2003) originally referred to this paradigm as "the transformative-emancipatory paradigm" (p. 135). Use of the full phrase (transformative-emancipation or transformative-emancipatory) keeps the focus of this paradigm directly on disrupting systems of oppression for the purpose of liberation from those systems (Shannon-Baker, 2016). Transformative-emancipatory researchers view reality as

subjective and socio-historically situated (Tashakkori et al., 2021). Multiple realities are influenced by social structures, especially those structures that perpetuate oppression and power differentials (Mertens, 2003; Plano Clark and Ivankova, 2016). As a result, researchers use their own values and subjectivities to meaningfully influence the research process (Plano Clark and Ivankova, 2016) as well as collaborators' perspectives and values. In this way, "objectivity is valued [for] providing a balanced and complete view" of an experience (Mertens, 2003, p. 141). Epistemologically, transformative-emancipatory researchers view knowledge creation as a process that takes place through collaboration with others (Plano Clark and Ivankova, 2016). Axiologically, this paradigm centers the lived experiences of marginalized communities and the use of research to bring about change to address social inequities (Mertens, 2003, 2007).

Transformative-emancipatory researchers emphasize working in collaboration with others throughout the research process by including research participants and other stakeholders in the research decision making (Mertens, 2003; Plano Clark and Ivankova, 2016; Tashakkori et al., 2021). For example, Garnett et al. (2019) combined transformative-emancipation with mixed methods and youth participatory research to explore the implementation of restorative practices in a school district in the USA. They found that this combination of methods with a transformative approach allowed researchers "to flatten hierarchies and actively engage marginalized youth to address the structural and programmatic inequities they experience in school" (p. 305). Mertens (2007) argued that combining mixed methods with a transformative-emancipatory approach is necessary to address the cultural complexities of social inequities. In another example, Garcia and Mayorga (2018) combined this paradigm with Critical Race Theory to examine how race and ethnicity are constructed in The Freshman Survey used in the USA to track access to higher education. Combining this paradigm and theoretical framework prompted them to ask critical questions about how race and ethnicity constructed through how questionnaires are framed, how data are disaggregated versus aggregated, and how samples are constructed in mixed methods educational research.

Postmodernism

The philosophical paradigm of *postmodernism* positions all knowledge and perspectives as equal and influenced by systems of power, but otherwise unstructured (Hesse-Biber and Kelly, 2010; Plano Clark and Ivankova, 2016). As a result, postmodernists reject the notion of objectivity in favor of socially situated and contextualized forms of knowledge (Hesse-Biber and Kelly, 2010). Similar to constructivism, postmodernists consider their own values and beliefs as an important influence on the research process but are no more important than participants' views or other stakeholders in the research process (Hesse-Biber and Kelly, 2010). Methodologically, postmodernist researchers often aim to expose and deconstruct systems of power through their research (Hutcheon, 1988; Plano Clark and Ivankova, 2016). Unlike other philosophical paradigms that might aim for social change, postmodernists do not aim to replace current systems of hierarchy (Rosenau, 1992). Instead, they aim to deconstruct texts, which can be policies, practices, institutions, or other systems, for their inherent power structures (Derrida, 1976; Hutcheon, 1988). This process of deconstruction involves a close reading of a text to identify the hierarchical system and expose inconsistencies, contradictions, and underlying assumptions (Derrida, 1976; Rosenau, 1992; Lyotard, 1984). In this way, postmodernists focus heavily on language, discourse, and *metanarratives*, or the discourses that govern one's beliefs and actions in a social system (Hesse-Biber and Kelly, 2010), while recognizing that the meaning of words and language is continuously shifting and socio-historically located (Derrida, 1976; Lyotard, 1984).

Postmodernism is particularly relevant for qualitatively oriented mixed methods researchers (Hesse-Biber and Kelly, 2010). Although mixed methods researchers have applied postmodern approaches in other fields (e.g., O'Loughlin et al., 2004), it is infrequently explicitly applied in mixed methods educational research. Educational researchers can combine postmodernism with mixed methods to deconstruct common myths about teaching and learning (Slaughter, 2001). Such researchers could ask research questions about what messages are embedded in policies and practices (Hesse-Biber and Kelly, 2010) in schooling, such as in dress code or bullying response policies, and how those messages impact students and teachers.

Dialectical pluralism

Dialectical pluralism has been referred to as a philosophical stance, paradigm, and metaparadigm (Greene and Caracelli, 2003; Greene and Hall, 2010; Johnson, 2017; Shannon-Baker, 2016; Tashakkori et al., 2021). I framed it here as a metaparadigm to showcase how it combines multiple philosophical paradigms, theoretical frameworks, and methodological approaches simultaneously (Johnson, 2017; Tashakkori et al., 2021). Dialectical pluralism emphasizes dialectical relationships among all frames within a research study, from methods to philosophical paradigms (Tashakkori et al., 2021). Dialectical pluralists believe in ontological pluralism, which posits that there are many ways to understand and perceive reality (Johnson, 2017). Knowledge is created through dialog between different perspectives in dialectical pluralism (Howes, 2017; Johnson, 2017; Plano Clark and Ivankova, 2016). Mixed methods researchers using this paradigm tend to emphasize dialog in data analysis, integration, interpretation, and writing (Plano Clark and Ivankova, 2016). For example, Shannon-Baker (2015) used this philosophical paradigm to inform the merging integration analysis of arts-informed, qualitative, and quantitative data in their study of students' experiences with culture shock during a study abroad program in a South American country for education majors. This paradigm also emphasized the importance of investigating divergence between the data sets during integration.

Yinyang philosophy

Mixed methods scholars have critiqued the Eurocentric discussion of philosophical paradigms in mixed methods research to show-case how yinyang philosophy from China aligns with mixed methods research thinking (Fetters et al., 2018; Fetters and Molina-Azorin, 2019). *Yinyang philosophy* is based on an ontology that embraces duality and the interconnectedness of elements. Within this philosophy, the researcher simultaneously holds that there is a singular and a dual reality, that multiple forms of knowledge are needed to understand a complex phenomenon, and that multiple methods need to be incorporated to achieve a balance of perspectives (Fetters, 2020; Fetters and Molina-Azorin, 2019). Although not an example specific to education, Knappertsbusch (2020) presents one of the few discussions applying yinyang philosophy to mixed methods research, describing how mixed methods integration aligns with this philosophy and the notion of *fractals*, which is a structure that contains smaller versions of the original. This philosophical paradigm offers a stance that emphasizes balance between perspectives and approaches that may be helpful for mixed methods educational researchers.

Indigenous philosophies

Indigenous researchers internationally have also critiqued the dominance of North American and Eurocentric perspectives on philosophical paradigms and mixed methods research (Chilisa and Tsheko, 2014; Martel et al., 2021). It is important to stress that specific philosophical beliefs and practices vary between different Indigenous communities. However, these philosophies generally share some similar characteristics. These characteristics include a recognition of Indigenous sovereignty; an embrace of the reciprocal relationship between humans, other animal species, the environment, the Land, and the spirit world; a focus on Indigenous voices and perspectives throughout the research process; and a disruption of colonial forms of knowledge and research (Chilisa and Tsheko, 2014; Martel et al., 2021; Smith, 2012). For example, Martel et al. (2021) argue that Māori worldviews, such as the metaphor of *he aua whiria* (braided river), support bicultural mixed methods research teams, like their own team comprised of Māori and New Zealand European researchers. Similar to the transformative-emancipation paradigm, their work describes an intertwining of multiple perspectives through the direct collaboration between Indigenous communities and non-Indigenous people (Martel et al., 2021).

Conclusions

As a mixed methods scholar, I fundamentally believe that philosophical paradigms influence all that we do, from designing a study to implementing and writing about it. The question is how that influence manifests and our intentionality around that influence. Exploring the differences and similarities among philosophical paradigms, like the discussion here introduces, provides mixed methods educational researchers one way to explore their own ontological, epistemological, and axiological beliefs and methodological practices. Seeking out manuscripts where researchers describe the influence of these beliefs on the research process will further elucidate the often intangible nuances to philosophical paradigms. How did their paradigm(s) influence the design, choice of sampling strategy, integration procedures, interactions with participants (or co-researchers), or how they presented their findings? Additionally, using the elements of philosophical paradigms to ask questions about the features of a study (published or one's own ongoing work) will further help to identify these hidden connections.

References

- Alberly, E., 2000. A way of visualising children's and young people's thoughts about the environment: a study of drawings. *Environ. Educ. Res.* 6, 205–222.
- Alise, M.A., Teddlie, C., 2010. A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *J. Mix. Methods Res.* 4, 103–126.
- Berger, R., 2015. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qual. Res.* 15, 219–234.
- Biddle, C., Schafft, K.A., 2015. Axiology and anomaly in the practice of mixed methods work: pragmatism, valuation, and the transformative paradigm. *J. Mix. Methods Res.* 9, 320–334.
- Camacho, S., 2020. From theory to practice: operationalizing transformative mixed methods with and for the studied population. *J. Mix. Methods Res.* 305–335.
- Chilisa, B., Tsheko, G.N., 2014. Mixed methods in Indigenous research: building relationships for sustainable intervention outcomes. *J. Mix. Methods Res.* 8, 222–233.
- Coates, A., 2020. The prevalence of philosophical assumptions described in mixed methods research in education. *J. Mix. Methods Res.* 1–19.
- Collins, K.M.T., Onwuegbuzie, A.J., Sutton, I.L., 2006. A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. *Learn. Disabil.* 4, 67–100.
- Cooper, J.N., Hall, J., 2016. Understanding Black male student athletes' experiences at a Historically Black College/University: a mixed methods approach. *J. Mix. Methods Res.* 10, 46–63.
- Corr, C., Snodgrass, M.R., Greene, J.C., Meadan, H., Santos, R.M., 2020. Mixed methods in early childhood special education research: purposes, challenges, and guidance. *J. Early Interv.* 42 (1), 20–30.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 4th ed. Sage.
- Danermark, B., Ekström, M., Jakobsen, L., Karlsson, J.C., 2002. *Explaining Society: Critical Realism in the Social Sciences*. Routledge, London.
- De Lisle, J., 2013. Exploring the value of integrated findings in a multiphase mixed methods evaluation of the continuous assessment program in the Republic of Trinidad and Tobago. *Int. J. Mult. Res. Approaches* 7, 27–49.
- Denzin, N.K., 1989. *Interpretive Biography*. Sage, Newbury Park, CA.

- Derrida, J., 1976. *Of Grammatology*. Johns Hopkins University Press, Baltimore, MD.
- Fetters, M.D., 2020. *The Mixed Methods Research Workbook: Activities for Designing, Implementing, and Publishing Projects*. Sage, Thousand Oaks, CA.
- Fetters, M., Wang, R.R., Johnson, R.B., 2018. Yinyang as a Paradigm for Mixed Methods Research. Paper presented at the Mixed Methods International Research Association Global Conference, Vienna, Austria. August 24, 2018.
- Fetters, M.D., Molina-Azorin, J.F., 2019. A call for expanding philosophical perspectives to create a more "worldly" field of mixed methods: the example of yinyang philosophy. *J. Mix. Methods Res.* 13, 15–18.
- Garcia, N.M., Mayorga, O.J., 2018. The threat of unexamined secondary data: a critical race transformative convergent mixed methods. *Race Ethn. Educ.* 21, 231–252.
- Garnett, B.R., Smith, L.C., Kervick, C.T., Ballysingh, T.A., Moore, M., Gonell, E., 2019. The emancipatory potential of transformative mixed methods designs: informing youth participatory action research and restorative practices within a district-wide school transformation project. *Int. J. Res. Method Educ.* 42, 305–316.
- Greene, J.C., 2006. Toward a methodology of mixed methods social inquiry. *Res. Sch.* 13, 93–98.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Greene, J.C., Caracelli, V.J., 2003. Making paradigmatic sense of mixed methods practice. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioural Research*. Sage, Thousand Oaks, CA, pp. 99–110.
- Greene, J.C., Hall, J.N., 2010. Dialectics and pragmatism: being of consequence. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 119–143.
- Guba, E.G., 1990. The alternative paradigm dialog. In: Guba, E.G. (Ed.), *The Paradigm Dialog*. Sage, Newbury Park, CA, pp. 17–28.
- Hammersley, M., 1992. The paradigm wars: reports from the front. *Br. J. Sociol. Educ.* 13, 131–143.
- Harrits, G.S., 2011. More than method? A discussion of paradigm differences within mixed methods research. *J. Mix. Methods Res.* 5, 150–166.
- He, Y., Lundgren, K., Pynes, P., 2017. Impact of short-term study abroad program: inservice teachers' development of intercultural competence and pedagogical beliefs. *Teach. Educ.* 66, 147–157.
- Hesse-Biber, S.N., 2010. *Mixed Methods Research: Merging Theory With Practice*. Guilford Press, New York, NY.
- Hesse-Biber, S.N., 2017. *The Practice of Qualitative Research*, third ed. Sage, Thousand Oaks, CA.
- Hesse-Biber, S.N., Kelly, C., 2010. Post-modernist approaches to mixed methods research. In: Hesse-Biber, S.N. (Ed.), *Mixed Methods Research: Merging Theory With Practice*. Guilford Press, New York, NY, pp. 154–173.
- Howes, L.M., 2017. Developing the methodology for an applied, interdisciplinary research project: documenting the journey toward philosophical clarity. *J. Mix. Methods Res.* 11, 450–468.
- Hutcheon, L.A., 1988. *Poetics of Postmodernism*. Routledge, London.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11, 156–173.
- Johnson, R.B., Anthony, J.O., Lisa, A.T., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1 (2), 112–133. <https://doi.org/10.1177/1558689806298224>.
- Johnson, R.B., de Waal, C., Stefurak, T., Hildebrand, D.L., 2017. Understanding the philosophical positions of classical and neopragmatists for mixed methods research. *Kölner Z. Soziol. Sozialpsychol.* 69, 63–86.
- Knappertsbusch, F., 2020. "Fractal heuristics" for mixed methods research: applying Abbott's "Fractal Distinctions" as a conceptual metaphor for method integration. *J. Mix. Methods Res.* 14, 456–472.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage, Newbury Park, CA.
- Lincoln, Y.S., Guba, E.G., 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *Handbook of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA, pp. 97–128.
- Lincoln, Y.S., Lynham, S.A., Guba, E.G., 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*, fourth ed. Sage, Thousand Oaks, CA, pp. 97–128.
- Lyotard, J.F., 1984. *The Postmodern Condition: A Report on Knowledge*. University of Minnesota Press, Minneapolis.
- Martel, R., Shepherd, M., Goodyear-Smith, F., 2021. *He awa whiria*—a "Braided River": an Indigenous Māori approach to mixed methods research. *J. Mix. Methods Res.* 1–17.
- Mason, J., 2002. *Qualitative Researching*, second ed. Sage, London, England.
- Maxwell, J.A., Mittapalli, K., 2010. Realism as a stance for mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 145–167.
- Mertens, D.M., 2003. Mixed methods and the politics of human research: the transformative-emancipatory perspective. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioural Research*. Sage, Thousand Oaks, CA, pp. 135–166.
- Mertens, D.M., 2007. Transformative paradigm: mixed methods and social justice. *J. Mix. Methods Res.* 1, 212–225.
- Mertens, D.M., Bledsoe, K.L., Sullivan, M., Wilson, A., 2010. Utilization of mixed methods for transformative purposes. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 193–215.
- Morgan, D.L., 2007. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *J. Mix. Methods Res.* 1, 48–76.
- O'Loughlin, J., Ó Tuathail, G.I., Kolossov, V., 2004. A "risky westward turn"? Putin's 9-11 script and ordinary Russians. *Eur. Asia Stud.* 56, 3–34.
- Phillips, D.C., Burbules, N.C., 2000. *Postpositivism and Educational Research*. Rowman & Littlefield, New York.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Ponce, O.A., Pagán-Maldonado, N., Gómez Galán, J., 2020. Philosophy of educational research: new epistemological, methodological and historical approach. *Int. J. Educ. Excellence* 6, 63–79.
- Reinhardt, C.S., Cook, T.D., 1979. Beyond qualitative versus quantitative methods. In: Cook, T.D., Reichardt, C.S. (Eds.), *Qualitative and Quantitative Methods in Evaluation Research*. Sage, Thousand Oaks, CA, pp. 7–32.
- Rosenau, P.M., 1992. *Post-modernism and the Social Sciences*. Princeton University Press, Princeton, N.J.
- Sandelowski, M., 2003. Tables or tableaux? The challenges of writing and reading mixed methods studies. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioural Research*. Sage, Thousand Oaks, CA, pp. 321–350.
- Schwandt, T., 1997. *Qualitative Inquiry: A Dictionary of Terms*. Sage, Thousand Oaks, CA.
- Shannon-Baker, P., 2015. "But I wanted to appear happy": how using arts-informed and mixed methods approaches complicate qualitatively driven research on culture shock. *Int. J. Qual. Methods* 14, 34–52.
- Shannon-Baker, P., 2016. Making paradigms meaningful in mixed methods research. *J. Mix. Methods Res.* 10, 319–334. <https://doi.org/10.1177/1558689815575861>.
- Shannon-Baker, P., 2022. Virtual special issue on "mixed methods designs, integration, and visual practices in educational research". *J. Mix. Methods Res.* 16 (2), 159–164. <https://doi.org/10.1177/15586898221083959>.
- Shannon-Baker, P., Martinez, C., 2022. Integrating arts-based approaches in mixed methods research with diverse stakeholders: a critically reflective review from two teacher-researchers. In: Hitchcock, J., Onwuegbuzie, A. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge, London, England, pp. 372–387.
- Slaughter, S., 2001. Problems in comparative higher education: political economy, political sociology and postmodernism. *High Educ.* 41, 389–412.
- Smeda, N., Dakich, E., Sharda, N., 2014. The effectiveness of digital storytelling in the classrooms: a comprehensive study. *Smart Learn. Environ.* 1, 1–21.
- Smith, L.T., 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books, London, England.

- Strayhorn, T.L., 2015. Factors influencing Black males' preparation for college and success in STEM majors: a mixed methods study. *West. J. Black Stud.* 39, 45–63.
- Tashakkori, A., Teddlie, C. (Eds.), 2003. *Handbook of Mixed Methods in Social and Behavioural Research*. Sage, Thousand Oaks, CA.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Thousand Oaks, CA.
- Ubha, N., Cahill, S., 2014. Building secure attachments for primary school children: a mixed methods study. *Educ. Psychol. Pract.* 30 (3), 272–292. <https://doi.org/10.1080/02667363.2014.920304>.
- Waring, M., 2017. Finding your theoretical position. In: Coe, R., Waring, M., Hedges, L.V., Arthur, J. (Eds.), *Research Methods & Methodologies in Education*. Sage, Thousand Oaks, CA, pp. 15–22.

Frameworks for conceptualizing mixed methods research

Vicki L. Plano Clark^a, Nataliya V. Ivankova^b, and Na Yang^c, ^a School of Education-Research Methods, University of Cincinnati, Cincinnati, OH, United States; ^b School of Health Professions, The University of Alabama at Birmingham, Birmingham, AL, United States; and ^c University of Cincinnati, School of Education-Second Language Acquisition, University of Cincinnati, Cincinnati, OH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	390
Definition and utility of conceptual frameworks	391
Different types of conceptual frameworks that describe mixed methods research	391
Frameworks describing the process of a mixed methods research study	392
Frameworks describing prototypical mixed methods research designs	393
Frameworks describing validity in mixed methods research	393
Frameworks describing the structure of the field of mixed methods research	394
Summary	394
Application of one framework for the field of mixed methods research	395
Overview of the socio-ecological framework for the field of mixed methods research	395
Application of the framework to organize the literature about mixed methods	396
Application of the framework to understand a published mixed methods study	396
Application of the framework to articulate one's positionality for mixed methods research	398
Application of the framework to plan teaching about mixed methods research	399
Concluding comments	400
References	400

Introduction

Mixed methods research involves the intentional combination of quantitative and qualitative research approaches to generate nuanced and valid insights about problems of interest (Johnson et al., 2007). Although there is a long history of researchers and evaluators combining quantitative and qualitative methods in their studies (Maxwell, 2016), interest in this approach has flourished in the 21st century across disciplines and countries. Educational scholars have been at the forefront of this interest by advocating for the value of mixing methods to address educational problems (e.g., Johnson and Onwuegbuzie, 2004; Mertens, 2015) and advancing considerations for how to design and implement quality mixed methods studies (e.g., Creamer, 2018; Creswell and Plano Clark, 2018; Greene, 2007; Tashakkori et al., 2021).

As interest in mixed methods research has grown within education and beyond, an entire field has emerged that is dedicated to the advancement of mixed methods research practice (Plano Clark and Ivankova, 2016; Tashakkori and Teddlie, 2003, 2010). This field encompasses an international community of scholars, an extensive body of literature, and a way of thinking that explicitly values different kinds of knowledge. Despite its relatively recent development, the field of mixed methods research is notably complex and dynamic (Fetters and Molina-Azorin, 2017; Hesse-Biber and Johnson, 2013; Poth, 2018). The field grapples with all dimensions of research from philosophical foundations to design and procedural considerations to team dynamics and funding priorities. Furthermore, the field intentionally combines different research approaches often associated with different assumptions, disciplines, and conventions. Educational scholars, researchers, and practitioners need to navigate this complexity to understand and effectively apply mixed methods research to important educational problems of interest.

One way that scholars address complexity is to develop conceptual frameworks that simplify and organize key elements of complex ideas. In response to the complexity of mixed methods research, methodologists have developed conceptual frameworks that identify key concepts that make up the field of mixed methods research and describe the relationships among them. These frameworks can be very helpful for individuals interested in navigating the different dimensions and perspectives associated with mixed methods research practice. In particular, researchers, methodologists, practitioners, teachers, and policymakers interested in educational issues can benefit from using conceptual frameworks to guide their understanding of mixed methods research as it applies to their interests.

Therefore, the aim of this chapter is to introduce and distinguish different types of frameworks for conceptualizing mixed methods research and to illustrate how readers can apply a socio-ecological framework for mixed methods research. We first define conceptual frameworks and briefly consider their general utility. Next, we differentiate types of frameworks that describe mixed methods research. From there, we examine the socio-ecological framework for the field of mixed methods research and consider how readers can apply this framework within their own educational research practice. We conclude with comments about future work for mixed methods research frameworks and recommendations for educators engaging with mixed methods research practice.

Definition and utility of conceptual frameworks

Before considering conceptual frameworks that describe mixed methods research, it is helpful to first briefly define what we mean by the term conceptual framework. In a conceptual framework, the word *concept* has been defined as assumptions, expectations, theories (Maxwell, 2012), points of view (Imenda, 2014), ideas, perceived facts, beliefs, mental pictures, perceptions, or theories (Rallis, 2018). These conceptions describe the nature of a complex topic by breaking it down into the most important ideas. The word *framework* in this term has been referred to as a system (Maxwell, 2012), a model (Plano Clark and Ivankova, 2016), a basic structure, or underlying organizational elements (Rallis, 2018). That is, a framework identifies the key concepts about a topic and describes how they are thought to work in the real world.

Based on the definitions of concept and framework, we define *conceptual framework* as an analytical model, map, structure, or organization of concepts that serves to represent and describe a complex idea, phenomenon, or system in a simplified way. Sometimes a conceptual framework is also referred to as a theoretical framework (Hughes, 2019), or a theoretical model (Plano Clark and Ivankova, 2016). These alternative terms often are used to refer to the researcher's assumptions about the substantive topic for a study and provide the theoretical foundation that informs decisions such as research questions and hypotheses and data collection forms (Grant and Osanloo, 2014; Plano Clark and Ivankova, 2016).

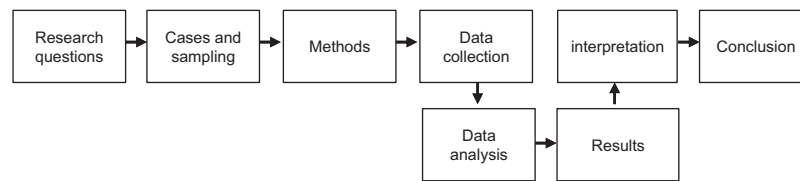
Conceptual frameworks have been used as a helpful tool for researchers and practitioners to describe their work's organization to different audiences, such as readers or students. As such, they can benefit different groups of people in education. For teachers, a conceptual framework can help them to organize their teaching concepts, assumptions, ideas, and examples into a logical mind map. Such maps can give them a clear picture of the relationship of different course components and help them to plan and articulate the flow of each class. For example, an instructor could use Bloom's taxonomy of six cognitive domains (e.g., remember, apply, evaluate) to organize the content in a course on curriculum design within the online learning environment. For researchers, a conceptual framework enables them to organize different theories, arguments, methods, procedures, and evidence in a clear and visual manner during the planning of their study. This framework can provide them the opportunity to check if any part of the research needs to be modified before being implemented. For example, Bustamante (2019) used the Technological Pedagogical Content Knowledge theoretical model to inform the decisions about and interpretations of the quantitative and qualitative methods in her study of the professional development of teachers of Spanish learning a new technology. A conceptual framework can also facilitate readers' comprehension of a research study's logic and design. Methodologists can use conceptual frameworks to organize complicated ideas about a research methodology into comprehensive or summative models to make sense of the field for others and to identify different methodological areas in need of further development. For example, Crotty (1998) advanced a four-level model of the elements that inform all research studies (i.e., epistemology, theoretical perspective, methodology, and methods) and used it to explain the diversity of approaches applied by researchers.

Different types of conceptual frameworks that describe mixed methods research

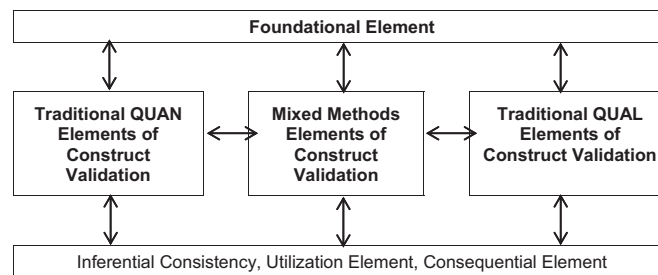
As the previous section highlighted, conceptual frameworks can be highly useful tools within any complex topic area or field of interest. Therefore, it is not surprising that methodologists working in the field of mixed methods have developed many different frameworks to conceptualize and describe mixed methods research. We consider *mixed methods conceptual frameworks* to be models, maps, structures, or organizations of concepts that represent and describe key aspects of mixed methods research considered by the field and relevant to the methodological considerations associated with this complex approach. Some mixed methods conceptual frameworks are very specific, focusing on one narrow topical area such as sampling (e.g., Collins et al., 2007) or validity (e.g., Delinger and Leech, 2007). Others are very broad, such as describing the domains of the field as a whole (e.g., Greene, 2008). Each framework has been thoughtfully crafted for a particular intent and has utility for different stakeholders to understand and engage with mixed methods research practice.

The availability of many different conceptual frameworks that describe mixed methods research has great potential benefit for individuals who use mixed methods research in their educational practice. However, the presence of so many different frameworks, developed by different scholars for different purposes, can also be confusing, especially for those who are new to this methodological approach. Therefore, before considering any particular framework in depth, it is helpful to first consider the different kinds of frameworks that exist for mixed methods research so that readers can recognize the variations and better select a framework that is appropriate for their own mixed methods research practice needs.

The prominent mixed methods conceptual frameworks fall into four types based on the framework's purpose. These types are: (a) frameworks describing the process of a mixed methods research study, (b) frameworks describing prototypical mixed methods research designs, (c) frameworks describing validity in mixed methods research, and (d) frameworks describing the structure of the field of mixed methods research. Each framework type is briefly introduced in the sections that follow and an example of each framework type is depicted in Fig. 1.

A Plowright's (2011) framework describing the process of a MMR study**B** Tashakkori, Johnson, and Teddlie's (2021) framework describing MMR design options

Four Families of MM Designs	• Parallel MM designs (e.g., QUAN + QUAL design) • Sequential MM designs (e.g., QUAN → QUAL design) • Conversion MM designs (e.g., QUAL data with QUAN and QUAL analyses) • Hybrid MM designs (e.g., QUAL → [QUAN + QUAL])
-----------------------------	---

C Dellinger and Leech's (2007) framework describing validity in MMR**D** Greene's (2008) framework describing the structure of the field of MMR

Domain 1: Philosophical Assumptions and Stances	Domain 2: Inquiry Logics (Methodology)
Domain 3: Guidelines for Practice	Domain 4: Sociopolitical Commitments

Fig. 1 Four examples of conceptual frameworks for mixed methods research (MMR) (A) Plowright's (2011) framework describing the process of a MMR study, (B) Tashakkori et al. (2021) framework describing MMR design options, (C) Dellinger and Leech's (2007) framework describing validity in MMR, (D) Greene's (2008) framework describing the structure of the field of MMR.

Frameworks describing the process of a mixed methods research study

When engaging with the mixed methods literature, readers are likely to come across conceptual frameworks that describe the dynamic process of doing a mixed methods research study (Creswell and Plano Clark, 2018). These frameworks highlight the different methodological decisions (or steps) involved when researchers are planning and conducting a study using mixed methods research and the interactions that occur among the decisions. These frameworks tend to conceptualize the research process as a dynamic system of choices and then describe how that process is adapted to mixed methods research. Many of these frameworks emphasize the research process by delineating major steps and how the steps fit together within a mixed methods study (e.g., Maxwell and Loomis, 2003; Plowright, 2011; Tashakkori et al., 2021).

There are many examples of frameworks describing the process of a mixed methods study in the literature. For example, Plowright (2011) introduced the Framework for an Integrated Methodology (FraIM) to discuss the process of mixed methods research. As depicted in Fig. 1A, he conceptualized the mixed methods research process involving eight connected steps. More recently, Tashakkori et al. (2021) used an input-process-outcome model to describe the stages of mixed methods research projects. In their model, the inputs include the study's background and purpose, the process includes data collection and analysis, and the outcomes include the results and conclusions. Other frameworks that describe the process of a mixed methods study are more conceptual in nature (e.g., Creamer, 2018; Hall and Howard, 2008; Long and Rodgers, 2017). For example, Creamer (2018) used a metaphor of an arch to describe the process of a mixed methods study. In her framework, the quantitative and qualitative components of the study are carefully constructed (like the two sides of an arch) to support integration and quality conclusions (represented by the keystone at the top of the arch).

Frameworks that describe the process of a mixed methods research study are very useful for understanding and applying mixed methods research within educational contexts. For educational researchers wanting to use mixed methods in their own research, these frameworks provide important guidance on how mixed methods decisions fit within the steps of the research process. For example, [Maxwell and Loomis' \(2003\)](#) interactive model helps researchers consider the linkages between their purpose, research questions, substantive conceptual framework, methods, and validity as they plan their mixed methods study. For educational practitioners and policy makers wanting to understand and learn from published mixed methods studies, these frameworks are also especially useful. Since mixed methods studies can be particularly complex to read because they often include two or more samples, types of data, and analytic procedures, a conceptual framework can help readers make sense of any mixed methods study in terms of its general process of research (such as its purpose, data collection and analysis, and results and conclusion). Readers interested in further discussion of these types of frameworks should see the chapter on dynamic approaches to designing mixed methods research by Maxwell.

Frameworks describing prototypical mixed methods research designs

The second type of framework was developed to delineate sets of different prototypical logics available for planning or comparing mixed methods research designs. These frameworks are often referred to as *design typologies* as they are made of different types of design logic ([Leech and Onwuegbuzie, 2007](#); [Teddle and Tashakkori, 2006](#)). As noted in [Fig. 1B](#), [Tashakkori et al. \(2021\)](#) provided an example of a design typology framework that includes four families of mixed methods designs (i.e., parallel, sequential, conversion, and hybrid) that describe different processes for implementing the quantitative and qualitative strands of a mixed methods study. As another example, [Creswell and Plano Clark \(2018\)](#) identified three mixed methods core designs in their framework based on the intent of the design (i.e., to explore, to explain, or to converge) and the sequence of qualitative and quantitative methods (e.g., sequential, concurrent). Each of these design prototypes can be depicted with a flow chart that demonstrates the timing of the application of each method (including data collection and analysis), and the point(s) when the results are mixed and interpreted. [Greene \(2007\)](#) categorized component design types (i.e., convergence, extension) based on research intent and integrated design types (i.e., iteration, blending, nesting or embedding, mixing for reasons of substance or values) based on mixing approach. [Morgan \(2014\)](#) highlighted the sequencing and priority options in his four core designs: qualitative input design, quantitative input design, qualitative follow-up design, and quantitative follow-up design. Interested readers can find a more in-depth discussion of design issues in the chapter on mixed methods research designs by Tajima.

Frameworks that describe typologies of different logics for mixed methods research can help researchers consider different possible options for their study and communicate their approach to others. In communicating with others, such frameworks can reflect the researchers' major emphasis and intent for using and integrating different types of databases ([Creswell and Plano Clark, 2018](#)). They also allow researchers from different fields to situate their studies within a specific domain of inquiry using mixed methods research ([Tashakkori and Teddle, 2010](#)). From a time-saving perspective, they also make researcher's work more efficient by providing different research approaches and sequence guidance for researchers to choose, to follow, and/or to modify based on their research purposes and resources.

Frameworks describing validity in mixed methods research

Frameworks describing validity aim to organize different methodological considerations necessary to assess the quality of mixed methods research. These frameworks have emerged in response to long-standing debates about what constitutes quality in mixed methods studies. Mixed methods validity remains one of the controversial issues in the mixed methods literature due to challenges associated with the need to appraise both the quality of the quantitative and qualitative study components and the validity of the derived mixed methods inferences ([Creswell, 2015](#); [Fàbregues and Molina-Azorin, 2017](#); [Teddle and Tashakkori, 2010](#)). A variety of the proposed quality standards and quality appraisal strategies can be also accounted for by a combination of contextual factors related to different quality standards in researchers' disciplines, prevalent philosophical paradigms and existing epistemological practices related to mixed methods research ([Collins, 2015](#); [Eckhardt and Devon, 2017](#); [Giddings and Grant, 2009](#); [Plano Clark and Ivankova, 2016](#)). The perspectives on how to assess the quality of a mixed methods study range from evaluating separately the methodological quality of the quantitative and qualitative study strands, to addressing quality concerns related to the integration of the methods in mixed methods designs, and to appraising the quality standards common to any research study ([Heyvaert et al., 2013](#)).

The existing frameworks combine these perspectives in unique ways offering different quality domains and evaluation criteria for appraising mixed methods studies. Despite using different approaches, these frameworks have an underlying common goal of inference quality, which is a criterion for considering quality of the meta-inferences as inductively and deductively derived conclusions from a mixed methods study ([Teddle and Tashakkori, 2003](#)). For example, [Onwuegbuzie and Johnson \(2006\)](#) described nine types of methods legitimization that researchers should consider to ensure inference quality during a mixed methods study process, whereas [Dellinger and Leech \(2007\)](#) approached mixed method quality from the perspective of construct validation of all study components to ensure inferential consistency of the conclusions (see [Fig. 1C](#)).

Since developing quality inferences from the integration of the quantitative and qualitative components is at the core of a mixed methods study, quality assessment is viewed as an ongoing activity spanning the whole study process from planning, design, and implementation to results interpretation and dissemination. Several validation frameworks provide a comprehensive set of assessment criteria that transcend all stages in a mixed methods study process. These include *Integrative Framework for Inference Quality* (Teddle and Tashakkori, 2009), *Comprehensive Framework for Assessing Quality of Mixed Methods Research* (O’Cathain, 2010), and *Critical Appraisal Framework for Quality in Mixed Methods Studies in Health Sciences* (Curry and Nunez-Smith, 2015). Some frameworks emerged that offer quality criteria related to specific characteristics of mixed methods designs, such as timing (concurrent or sequential) of the quantitative and qualitative components and their integration. For example, Creswell and Plano Clark (2018) described validity threats for core and complex designs, whereas Ivankova (2014) advanced quality criteria specifically related to sequential Quantitative → Qualitative design. Additionally, the Mixed Methods Appraisal Tool (MMAT) was developed to critically evaluate qualitative, quantitative, and mixed methods studies included within systematic mixed methods reviews (Hong et al., 2018; Pluye et al., 2009).

Despite different purposes and quality criteria, mixed methods validity frameworks have promised utility because they help researchers understand how to reduce inconsistencies between the inferences from the quantitative and qualitative study components and how to create credible meta-inferences from an entire study. Moreover, these frameworks recognize the complexity of the mixed methods research process and the critical role of integration in ensuring inference quality and inference transferability. When choosing a framework, researchers should consider its scope and focus, the feasibility of using the suggested evaluation criteria, and their own beliefs and related epistemological practices. For consumers of mixed methods research, validity frameworks offer tools to appraise the quality of the reported mixed methods studies in both the research process and outcomes. Readers interested in further discussion of these types of frameworks should see the chapter on quality in mixed methods research by Collins.

Frameworks describing the structure of the field of mixed methods research

A fourth type of mixed methods conceptual framework emerged as the field of mixed methods research has become more established and complex. This type includes frameworks developed to describe the complex structure of the field of mixed methods research as a whole. This type of framework specifies large conceptual domains of interest to the field and uses those domains to explain how the field of mixed methods is organized. Such frameworks serve to define the field of mixed methods and its primary topical areas of consideration. Several frameworks have delineated a small set of three to five core domains to describe the field, which have been referred to as “methodological domains” (Greene, 2008, p. 7), the “developing landscape” (Creswell, 2010, p. 46), and “overlapping components” (Teddle and Tashakkori, 2010, p. 3). For example, as depicted in Fig. 1D, Greene (2008) described the field of mixed methods research in terms of four large methodological domains: philosophical assumptions and stances, inquiry logics, guidelines for practice, and sociopolitical commitments. Plano Clark and Ivankova (2016) advanced a socio-ecological framework that identified components of the field and organized them into interconnected levels to depict how the components and levels interact with each other. More recently, Fetters (2018) conceptualized the field into six different equations that capture the “compelling issues” (p. 262) of interest to the field.

The previously described frameworks conceptualized mixed methods research in terms of decisions and options related to a specific aspect of mixed methods research practice. In contrast, the frameworks for describing the structure of the field are broader and more conceptual in nature. These frameworks serve to identify and make sense of the many domains and perspectives found within the field of mixed methods research. Due to their breadth, they can be especially useful for understanding, applying, and assessing the field of mixed methods research as a whole. For researchers who are new to mixed methods research, frameworks describing the field’s structure provide a useful introduction to the main conversations occurring in the field. Simply put, they are a quick way to learn what members of the field are talking and writing about so the ideas can be applied in one’s research applications. For consumers who are working to learn and understand mixed methods research, these frameworks facilitate recognizing the major topics of interest in the field and identifying how different sources from the literature fit among those major topics. Relatedly, these frameworks are also useful for making curricular decisions for instructors who are teaching about mixed methods research. Furthermore, methodologists interested in the advancement of mixed methods use these frameworks to assess the status of the field, including what is known and what remains unknown within major domains.

Summary

Each of the types of frameworks we have highlighted plays an important role in mixed methods research. Readers can benefit from identifying and applying all the types. The first step, however, is to be able to recognize that there are different kinds of frameworks and that each framework is developed for its own purpose and intent. From there, readers can apply selected frameworks in their mixed methods research practice in a multitude of ways. In the next section, we highlight one specific mixed methods conceptual framework and provide examples of different ways that it can be applied within one’s educational and research practice.

Application of one framework for the field of mixed methods research

Across the many mixed methods conceptual frameworks that are available (and continue to be developed), frameworks for describing the structure of the field of mixed methods research are unique in terms of their scope and applicability. Therefore, we have selected one specific framework to examine in more depth, namely the *socio-ecological framework for the field of mixed methods research* introduced by [Plano Clark and Ivankova \(2016\)](#). We provide an overview of this framework and then consider four different ways that it can be applied within educational contexts.

Overview of the socio-ecological framework for the field of mixed methods research

Recognizing the growing complexity of the mixed methods field and the dynamic relationships that exist among its different components, [Plano Clark and Ivankova \(2016\)](#) advanced a comprehensive framework that would help researchers and consumers navigate the mixed methods field. They described the mixed methods field from a socio-ecological perspective that recognizes the existence of the interwoven dynamic relationships between the various individual and environmental factors, such as personal, interpersonal, organizational, community and societal contexts ([Bronfenbrenner, 1979](#); [Stokols, 1996](#)). The framework aims to identify key mixed methods methodological considerations and contextual domains of interest to the field and explain how the considerations and contexts influence researchers' approaches to designing and implementing mixed methods studies.

The framework for the mixed methods field takes the form of a socio-ecological model of hierarchical interconnected levels and their components consisting of the mixed methods research process, mixed methods research content, and mixed methods research contexts (see [Fig. 2](#)). The mixed methods research process occupies the central position because it provides the foundation for any mixed methods study through the formulation of the research purpose and questions, selection of methods, and creation of inferences (see the central box in [Fig. 2](#)). The mixed methods research process is nested within five methodological domains, which form the mixed methods research content (see five overlapping inner circles in [Fig. 2](#)). These methodological domains represent the decisions that researchers have to make about the mixed methods research process and include the definitions of mixed methods research, the rationales for using mixed methods in a study, the mixed methods design logics, the criteria for assessing quality in mixed methods research, and mixed methods intersection with other approaches and designs. In their turn, these domains are shaped by three interconnected mixed methods research contexts: personal, interpersonal and social contexts (see three nested outer circles in [Fig. 2](#)).

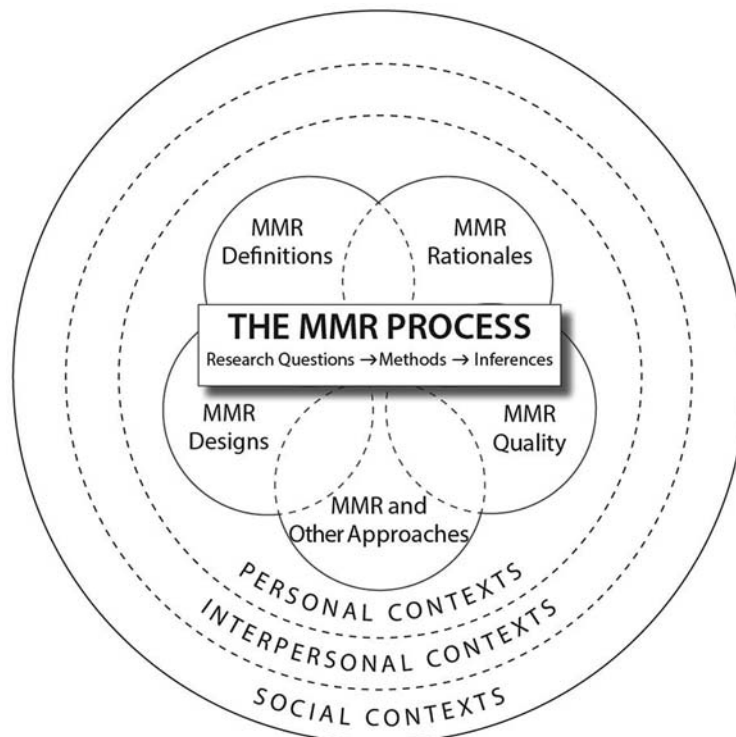


Fig. 2 [Plano Clark and Ivankova's \(2016\)](#) socio-ecological framework for the field of mixed methods research (MMR). Figure is reprinted with permission of Sage Publishing from [Plano Clark and Ivankova \(2016\)](#).

Mixed methods personal contexts include researchers' philosophical assumptions, theoretical models, and background knowledge. Interpersonal contexts are comprised of interpersonal relations and ethical considerations with study participants, research teams, and reviewers. Social contexts are institutional structures, disciplinary conventions, and societal priorities with regards to mixed methods research. These contexts jointly influence researchers' choices of methodological decisions related to five mixed methods content domains resulting in a variety of ways how researchers approach the research process in mixed methods studies.

The socio-ecological framework for mixed methods research has several advantages for educational research practice. This framework is the most comprehensive framework in the field of mixed methods research and can be applied in many ways to understand, describe, and explain different components of the mixed methods research practice. Moreover, the socio-ecological perspective provides the mechanism for applying mixed methods research to study complex educational issues by helping researchers position themselves regarding mixed methods research and make methodological decisions that align with the research questions of interest. The interactive and dynamic nature of the framework allows consumers of research to better understand the variety of perspectives on mixed methods across disciplines, research topics, and practices, thus helping them make sense of the methodological and empirical mixed methods writings. Its comprehensive nature also makes it useful for teaching and learning about mixed methods research. In the following sections, we illustrate how the socio-ecological framework for mixed methods research can be applied in educational research and teaching practice.

Application of the framework to organize the literature about mixed methods

The field of mixed methods has expanded into an extensive collection of literature that covers many topics written by many different scholars. There are thousands of articles and hundreds of books and chapters available to read that discuss mixed methods research. Some of the literature describes different applications of mixed methods research and some provides insights into the methodology. The extensive nature of this literature can quickly become overwhelming for readers, particularly those who are new to mixed methods research. It can be a challenge to make sense of individual readings and identify how they are similar to and differ from other readings.

One of the most useful features of any conceptual framework is that it identifies key domains for the phenomenon of interest, which can be applied to organize information. Education scholars and practitioners who are reading literature about mixed methods research can apply the domains from the socio-ecological framework for mixed methods to help them make sense of and organize that literature. For example, [Plano Clark and Ivankova \(2016\)](#) used the nine domains in the model (as identified in [Fig. 2](#)) to organize their book about mixed methods research by writing a chapter for each domain. Likewise, readers can use the nine domains (or just the three levels of process, content, and contexts) to classify the primary topics of their readings about mixed methods.

An example can illustrate this strategy. A reader interested in mixed methods research might encounter the following two articles: (a) *Dialectic Dialogue: Reflections on Adopting a Dialectic Stance* by [Cronenberg and Headley \(2019\)](#) and (b) *Multilevel Mixed Methods Research Designs: Advancing a Refined Definition* by [Headley and Plano Clark \(2020\)](#). These two articles happen to share a common author, but each fits into a different domain within the field of mixed methods research. First, educational researchers [Cronenberg and Headley \(2019\)](#) wrote about their use of a dialectical philosophical stance while conducting their mixed methods dissertations. Using the socio-ecological framework, readers could make sense of this article as fitting within the contextual domain of personal contexts for mixed methods research. Second, educational researchers [Headley and Plano Clark \(2020\)](#) wrote about the definition of multilevel mixed methods designs, using several examples from education to highlight the utility of such designs. Using the socio-ecological framework, readers could make sense of this article as fitting within the content domain of mixed methods research designs. Distinguishing these two articles by these larger mixed methods domains can be helpful for readers to distinguish the topics of these two methodological writings. As readers engage with additional literature about mixed methods, a helpful strategy is to develop folders for organizing the different sources into topical domains to facilitate understanding and retrieval of information later. This strategy can be facilitated by applying the domains identified within the socio-ecological framework to organize this literature.

Application of the framework to understand a published mixed methods study

Another challenge in mixed methods research is for readers to understand published reports of mixed methods studies. Because mixed methods approaches are well suited to address complex topics, they are becoming increasingly common for examining issues relevant to education. Therefore, no matter the particular educational area of interest, readers are likely to encounter reports of studies that used mixed methods approaches. Through the combination of quantitative and qualitative methods within the studies, the results and interpretations of these studies can provide readers with comprehensive and nuanced insights. However, because mixed methods studies include quantitative, qualitative, and mixed methods aspects, the published reports of these studies can be a challenge to understand. One way to address the challenge of understanding a published mixed methods study is for readers to use the socio-ecological framework to identify specific mixed methods elements occurring within the study and map out how they are related to each other.

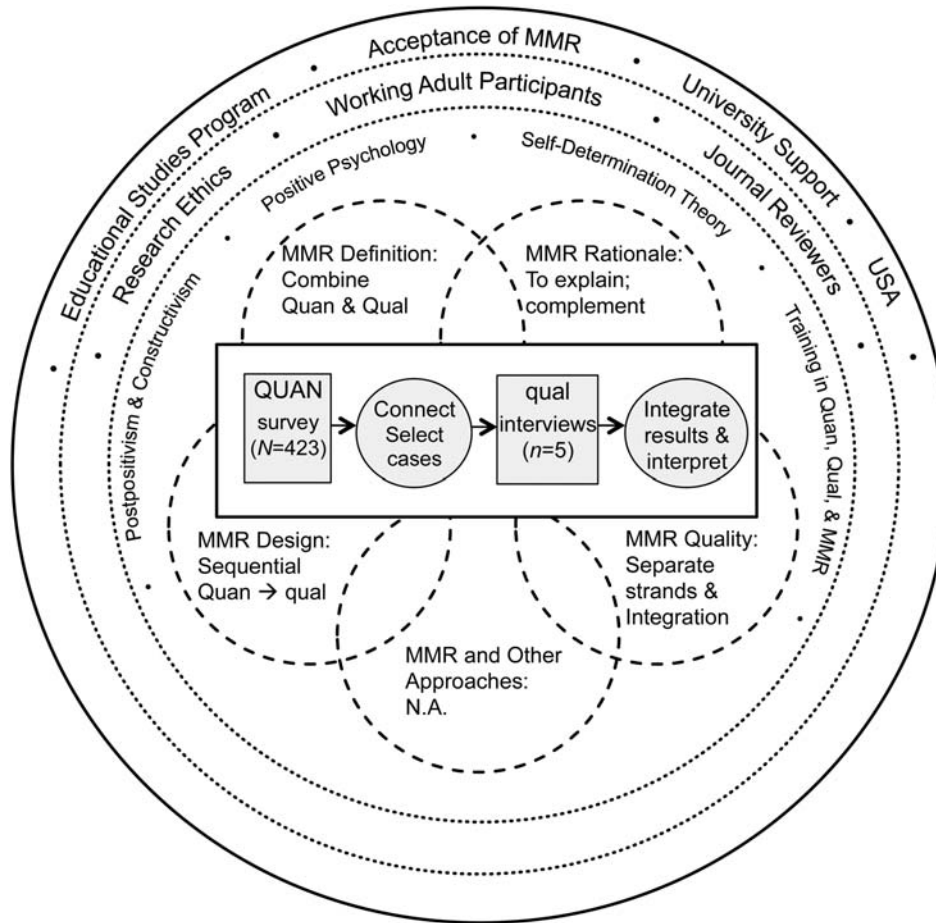


Fig. 3 Application of the socio-ecological framework for understanding a published mixed methods study. Figure based on Plano Clark and Ivankova's (2016) Socio-ecological Framework for Mixed Methods Research (MMR) and the study report published by Clark and Plano Clark (2019).

Fig. 3 is an example of how the socio-ecological framework can be applied to understand the design and implementation of a mixed methods study about grit in the context of career success (Clark and Plano Clark, 2019). The study's mixed methods research process is portrayed in the center of Fig. 3. Clark and Plano Clark designed the study using two sequential phases starting with a quantitative phase and then connecting to a qualitative phase. The uppercase letters indicate the priority of the quantitative phase for addressing the study's purpose. In the quantitative phase, the researchers gathered data from 423 working adults using an online survey. The researchers used statistical procedures to test the hypothesized positive relationship between grit and career success, and the results found no significant relationship. From there, they interviewed five individuals who were purposefully selected because they scored high in terms of both grit and career success to better understand the role of grit within their careers. The two sets of results were integrated in the end to provide novel insights into how grit manifests within the career context.

The five overlapping circles around the research process box in Fig. 3 highlight the mixed methods research content considerations that informed the mixed methods research process of Clark and Plano Clark's (2019) grit study. These considerations included defining mixed methods as the "intentional administration" (p. 94) of both quantitative and qualitative methods to understand a problem; having a rationale of using the qualitative phase to complement and explain the meaning of quantitative results; applying the logic of a sequential (QUAN → qual) mixed methods design; and using strategies to ensure the validity of the quantitative results and the credibility of the qualitative findings. The three outer circles found in Fig. 3 describe the interrelated mixed methods research contexts that shaped the study. For example, the authors' personal contexts included their alignment with the assumptions of positive psychology and training in mixed methods research. Interpersonal contexts included gaining approval for the study from the university's ethical review board and recruiting working adults to participate in the two phases of the study. Social contexts included disciplinary considerations from both education and psychology and the acceptance of mixed methods within the university program. By mapping out the mixed methods research process, content considerations, and contexts in this way, readers can have a good understanding of the major elements found within any mixed methods study.

Application of the framework to articulate one's positionality for mixed methods research

The previous section demonstrated how the socio-ecological framework for mixed methods research can be applied to identify the influential contexts and considerations found within a published mixed methods study. In addition to being able to describe others' mixed methods research, it is important for educational researchers to articulate their own positionality for mixed methods research. Much of the complexity found in mixed methods research can be attributed to the fact that scholars interested in mixed methods research come from different contexts and have different viewpoints about the considerations relevant to mixing methods. When engaging with mixed methods research practices, individuals benefit from articulating their own perspectives about mixed methods in addition to recognizing the perspectives of others.

The socio-ecological framework for mixed methods research provides a useful framework for describing one's mixed methods viewpoints. We have provided questions in [Table 1](#) that are aligned with the framework and may be helpful for individuals to answer as they articulate their own positionality. Starting from the outer-most layers depicted in [Fig. 2](#), individuals should recognize the social, interpersonal, and personal contexts that they bring to their thinking about the use of mixed methods research by answering the questions in the first section of [Table 1](#). Next, we encourage individuals to articulate their position relevant to the major context considerations for mixed methods research. The context consideration domains are depicted in the five overlapping circles within [Fig. 2](#) and corresponding questions appear in the middle section of [Table 1](#). Finally, individuals can describe their ideas about the mixed methods research process by answering the questions in the bottom section of [Table 1](#). For educational researchers, these answers might pertain to a specific study, but they can also be answered generally for anyone interested in educational issues informed by mixed methods research practices. Collectively, individuals can identify and describe their position within the field of mixed methods research and connect this position to their mixed methods research practice by addressing each of the domains found within the socio-ecological framework. [Sanders \(2021\)](#) used this approach when describing her positionality in the context of her dissertation study involving Black women faculty in higher education within the United States. As shown in [Fig. 4](#), she used the framework to articulate her perspectives about "ways Black women faculty and their lives are intertwined historically, interpersonally, and socially as it relates to their presence and work in higher education" (p. 18) along with her assumptions about the use of a mixed methods approach.

Table 1 Applying the socio-ecological framework for articulating one's positionality for mixed methods research.

Level	Domain	Example questions to answer
MMR contexts	Social contexts	• In what ways is MMR supported by your organization? • What is the status of MMR in your specific disciplinary area? • To what extent is MMR supported by relevant funding mechanisms?
	Interpersonal contexts	• How is MMR understood by your human subjects review board? • How will MMR be received by potential participants? • What different perspectives about research are represented within your partners and team members?
	Personal contexts	• How open are gatekeepers to the use of MMR? • What assumptions do you hold about the nature of knowledge and how MMR can contribute to knowledge? What philosophies help to describe these assumptions? • What assumptions do you hold about how your educational topic works? What theories help to explain these assumptions?
MMR content	MMR definitions	• What training and prior experiences inform your thinking about MMR?
	MMR rationales	• How do you define MMR? What literature informs this definition?
	MMR designs	• What are your reasons for using MMR? What literature informs this rationale? • What is your approach to thinking about MMR designs? What literature informs this design thinking?
	MMR and other approaches	• What other approaches to research are you using? How might this approach combine with MMR? What literature informs this combination?
MMR process	MMR quality	• What is your approach to assessing quality in MMR? What literature informs your quality standards?
	Process elements	• What research questions are of interest to you?
		• What quantitative and qualitative methods are best for addressing these questions? How can/will the two methods be used in combination? • What kinds of insights can be gained by combining the methods?

Note: MMR = mixed methods research.

Information based on [Plano Clark and Ivankova \(2016\)](#).

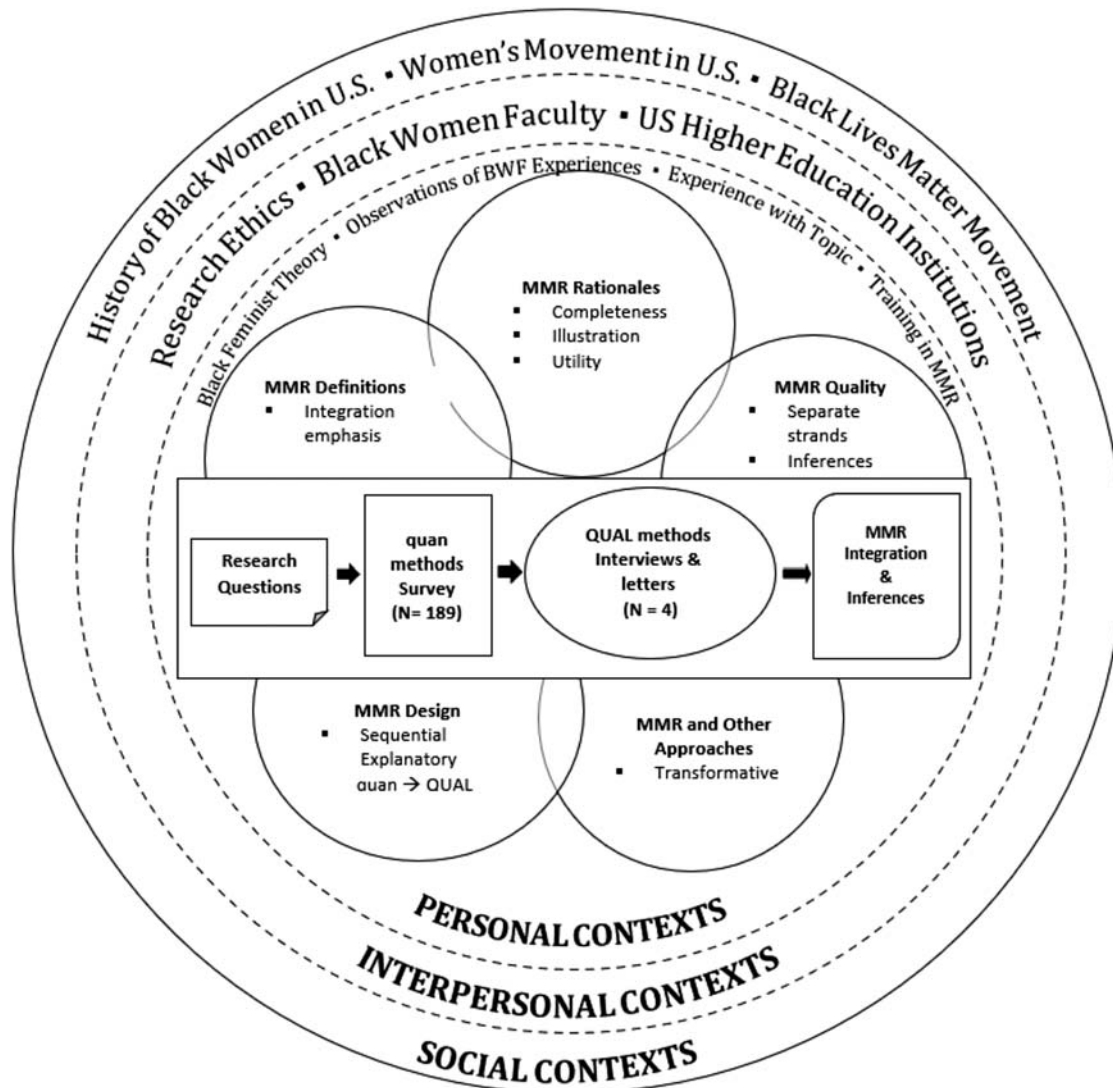


Fig. 4 Application of the socio-ecological framework for mixed methods research (MMR) to articulate a researcher's positionality. © Sanders (2021), figure is reproduced with permission of the author.

Application of the framework to plan teaching about mixed methods research

The growth of the mixed methods field in both the volume of mixed methods studies and their complexity creates significant challenges for learning about mixed methods research. These challenges are prompted by several factors. The field of mixed methods research is relatively new and embraces a variety of perspectives on what constitutes mixed methods, the rationales for using it, the choice of mixed methods designs, and the quality of the generated findings, to name a few. Moreover, researchers' use of mixed methods is often shaped by the epistemological approaches prevalent in their disciplines, which results in observed variations in how mixed methods studies are designed and conducted. Understanding and navigating the realm of different mixed methods perspectives and practices can be a daunting task for emergent scholars who range from consumers of mixed methods research to researchers planning to apply mixed methods in their studies.

The socio-ecological framework for mixed methods provides instructors with a pedagogical tool to conceptually organize learning about mixed methods around its methodological domains and influencing contexts. For example, [Ivankova and Plano Clark \(2018\)](#) used the framework to structure their Introduction to Mixed Methods Research course. The goal of the course is to provide students with foundational knowledge about mixed methods research and its fundamental principles and to introduce students to the major perspectives on the application of mixed methods. The course is organized around the framework domains starting with the process of mixed methods research and exploring how the different perspectives within each methodological domain shape researchers' approaches to applying mixed methods to research problems of interest. The influencing role of the three

contexts is then explored with a special emphasis on students' understanding of the dynamic interrelationship of all levels and components of the framework in their joint impact on the mixed methods research process. In addition to didactic knowledge, students learn to apply the framework domains to synthesize and critique different perspectives found in the literature about the methodological aspects of mixed methods research.

Concluding comments

In this chapter, we introduced different types of mixed methods conceptual frameworks and highlighted the socio-ecological framework for mixed methods research in educational practice. These frameworks are well-suited to enhance one's understanding of the complexity of mixed methods research in different ways. It should be noted that each type of conceptual framework illustrates an aspect (e.g., process, logic) of a mixed methods project or specific mixed methods topic (e.g., the structure of the field) that needs to be summarized, introduced, or clarified. Therefore, before adopting a conceptual framework, it is important to decide which aspect of mixed methods research needs to be elaborated. Remember that each framework is a tool, and each tool serves the best when being used in the right place and the right way.

In addition, as research and educational practice keep evolving, new aspects of mixed methods research will keep emerging, so will the conceptual frameworks. As researchers, instructors, and practitioners across many disciplines, including education, continue to expand their thinking about and application of mixed methods research, we expect that additional conceptual frameworks will be warranted. We can imagine that future frameworks may emerge to describe the increasing complexity of mixed methods research as it incorporates new data forms (e.g., social media, position tracking data), analytic approaches (e.g., machine learning, social network analysis), and stakeholders (e.g., policy makers, community members). Therefore, as much as we hope our chapter serves as an information tool box, we also hope it serves as an inspiring stepstone for our readers to step on to adopt, modify, and innovate conceptual frameworks for better conceptualizing mixed methods research for others.

Finally, we recommend a few resources as a quick-start for our readers who are interested to explore more about conceptual frameworks that describe mixed methods research.

Further readings:

1. For further elaboration and examples of conceptual frameworks describing the process of a mixed methods research study: see [Tashakkori et al. \(2021\)](#) and works cited in **Frameworks describing the process of a mixed methods research study** section.
2. For further elaboration and examples of conceptual frameworks describing prototypical mixed methods research designs: see [Creswell and Plano Clark \(2018\)](#) and works cited in **Frameworks describing prototypical mixed methods research designs** section.
3. For further elaboration and examples of conceptual frameworks describing validity in mixed methods research: see [Collins \(2015\)](#), [Fàbregues et al. \(2019\)](#), and works cited in **Frameworks describing validity in mixed methods research** section.
4. For further elaboration and examples of conceptual frameworks describing the structure of the field of mixed methods research: see [Greene \(2008\)](#), [Plano Clark and Ivankova \(2016\)](#), and works cited in **Frameworks describing the structure of the field of mixed methods research** section.
5. For the most current thinking about mixed methods research from a methodological perspective: see editorials and articles published in the *Journal of Mixed Methods Research* and the *International Journal of Multiple Research Approaches*.
6. For applications of mixed methods conceptual frameworks in complex educational contexts: see [Johnson and Onwuegbuzie \(2004\)](#), [Riazi \(2017\)](#), and [Poth \(2018\)](#).

References

- Bronfenbrenner, U., 1979. *The Ecology of Human Development*. Harvard University Press, Cambridge, MA.
- Bustamante, C., 2019. TPACK and teachers of Spanish: development of a theory-based joint display in a mixed methods research case study. *J. Mix. Methods Res.* 13, 163–178. <https://doi.org/10.1177/1558689817712119>. Available at. (Accessed 22 November 2021).
- Clark, R., Plano Clark, V.L., 2019. Grit within the context of career success: a mixed methods study. *Int. J. Appl. Posit. Psychol.* 4, 91–111.
- Collins, K.M.T., Onwuegbuzie, A.J., Jiao, Q.G., 2007. A mixed methods investigation of mixed methods sampling designs in social and health science research. *J. Mix. Methods Res.* 1, 267–294.
- Collins, K.M.T., 2015. Validity in multimethod and mixed research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 240–256.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. Sage Publishing, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE Publications, Thousand Oaks, CA.
- Creswell, J.W., 2010. Mapping the developing landscape of mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. SAGE Publications, Los Angeles, CA, pp. 45–68.
- Creswell, J.W., 2015. *A Concise Introduction to Mixed Methods Research*. SAGE Publications, Thousand Oaks, CA.
- Bronenberg, S., Headley, M.G., 2019. Dialectic dialogue: reflections on adopting a dialectic stance. *Int. J. Res. Method Educ.* 42, 267–287.
- Crotty, M., 1998. *The Foundations of Social Research: Meaning and Perspective in the Research Process*. SAGE Publication, Thousand Oaks, CA.
- Curry, L., Nunez-Smith, M., 2015. *Mixed Methods in Health Sciences Research: A Practical Primer*. SAGE Publications, Thousand Oaks, CA.
- Dellinger, A.B., Leech, N.L., 2007. Toward a unified validation framework in mixed methods research. *J. Mix. Methods Res.* 1, 309–332.
- Eckhardt, A.L., Devon, H.A., 2017. The MIXED framework: a novel approach to evaluating mixed-methods rigor. *Nurs. Inq.* 24, 1–13.

- Fàbregues, S., Molina-Azorin, J.F., 2017. Addressing quality in mixed methods research: a review and recommendations for a future agenda. *Qual. Quantity* 51, 2847–2863.
- Fàbregues, S., Paré, M.H., Meneses, J., 2019. Operationalizing and conceptualizing quality in mixed methods research: a multiple case study of the disciplines of education, nursing, psychology, and sociology. *J. Mix. Methods Res.* 13, 424–445.
- Letters, M.D., Molina-Azorin, J.F., 2017. The *Journal of Mixed Methods Research* starts a new decade: perspectives of past editors on the current state of the field and future directions. *J. Mix. Methods Res.* 11, 423–432.
- Letters, M.D., 2018. Six equations to help conceptualize the field of mixed methods. *J. Mix. Methods Res.* 12, 262–267.
- Giddings, L.S., Grant, B.M., 2009. From rigor to trustworthiness: validating mixed methods. In: Andrew, S., Halcomb, E.J. (Eds.), *Mixed Methods Research for Nursing and the Health Sciences*. Wiley-Blackwell, Chichester, UK, pp. 119–134.
- Grant, C., Osanloo, A., 2014. Understanding, selecting, and integrating a theoretical framework in dissertation research: creating the blueprint for your “house”. *Adm. Issues J.* 4, 12–26.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Greene, J.C., 2008. Is Mixed methods social inquiry a distinctive methodology? *J. Mix. Methods Res.* 2, 7–22.
- Hall, B., Howard, K., 2008. A synergistic approach: conducting mixed methods research with typological and systematic design considerations. *J. Mix. Methods Res.* 2, 248–269.
- Headley, M.G., Plano Clark, V.L., 2020. Multilevel mixed methods research designs: advancing a refined definition. *J. Mix. Methods Res.* 14, 145–163.
- Hesse-Biber, S., Johnson, R.B., 2013. Coming at things differently: future directions of possible engagement with mixed methods research. *J. Mix. Methods Res.* 7, 103–109.
- Heyvaert, M., Hannes, K., Maes, B., Onghena, P., 2013. Critical appraisal of mixed methods studies. *J. Mix. Methods Res.* 7, 302–327.
- Hong, Q.N., Gonzalez-Reyes, A., Pluye, P., 2018. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *J. Eval. Clin. Pract.* 24, 459–467.
- Hughes, S., 2019. Writing qualitative dissertations. *Oxf. Res. Encycl. Educ.* Available at: <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-737>. (Accessed 18 July 2020).
- Imenda, S., 2014. Is there a conceptual difference between theoretical and conceptual frameworks? *J. Soc. Sci.* 38, 185–195.
- Ivankova, N.V., Plano Clark, V.L., 2018. Teaching mixed methods research: using a socio-ecological framework as a pedagogical approach for addressing the complexity of the field. *Int. J. Soc. Res. Methodol.* 21, 409–424.
- Ivankova, N., 2014. Implementing quality criteria in designing and conducting a sequential QUAN→QUAL mixed methods study of student engagement with learning applied research methods online. *J. Mix. Methods Res.* 8, 25–51.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1, 112–133.
- Leech, N.L., Onwuegbuzie, A.J., 2007. A typology of mixed methods research designs. *Qual. Quantity* 43, 265–275.
- Long, H., Rodgers, C.R., 2017. Re-conceptualize mixed methods research: introducing a new conceptual framework. *Qual. Quantity* 51, 2813–2829.
- Maxwell, J.A., Loomis, D.M., 2003. Mixed methods design: an alternative approach. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications, Thousand Oaks, CA, pp. 241–271.
- Maxwell, J.A., 2012. *Qualitative Research Design: An Interactive Approach*. SAGE Publications, Los Angeles, CA.
- Maxwell, J.A., 2016. Expanding the history and range of mixed methods research. *J. Mix. Methods Res.* 10, 12–27.
- Mertens, D.M., 2015. Mixed methods and wicked problems. *J. Mix. Methods Res.* 9, 3–6.
- Morgan, D.L., 2014. *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. SAGE Publications, London.
- O’Cathain, A., 2010. Assessing the quality of mixed methods research: towards a comprehensive framework. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Thousand Oaks, CA, pp. 531–555.
- Onwuegbuzie, A.J., Johnson, R.B., 2006. The validity issue in mixed research. *Res. Sch.* 13, 48–63.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE Publications, Thousand Oaks, CA.
- Plowright, D., 2011. *Using Mixed Methods: Frameworks for an Integrated Methodology*. SAGE Publications, London.
- Pluye, P., Gagnon, M.-P., Griffiths, F., Johnson-Lafleur, J., 2009. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *Int. J. Nurs. Stud.* 46, 529–546.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. SAGE Publications, Thousand Oaks, CA.
- Rallis, S.F., 2018. Conceptual Framework. *Sage Research Methods*. Available at: <https://methods.sagepub.com/reference/the-sage-encyclopedia-of-educational-research-measurement-and-evaluation>. (Accessed 18 July 2020).
- Riazi, A.M., 2017. *Mixed Methods Research in Language Teaching and Learning*. Equinox Publishing, Bristol, CT.
- Sanders, K., 2021. *Microaggressions, Emotional Regulation, and Thriving in Higher Education: A Mixed Methods Study About Black Women Faculty* [Unpublished doctoral dissertation]. University of Cincinnati, Cincinnati, OH.
- Stokols, D., 1996. Translating social ecological theory into guidelines for community health promotion. *Am. J. Health Promot.* 10, 282–298.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. SAGE Publications, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2003. Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in the Behavioral and Social Sciences*. SAGE Publications, Thousand Oaks, CA, pp. 3–50.
- Teddlie, C., Tashakkori, A., 2006. A general typology of research designs featuring mixed methods. *Res. Sch.* 13, 12–28.
- Teddlie, C., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. SAGE Publications, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2010. Overview of contemporary issues in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. SAGE Publications, Thousand Oaks, CA, pp. 1–42.

Transformative mixed methods research

Donna M. Mertens, Gallaudet University, Rockville, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Transformative mixed methods research	402
Assumptions associated with transformative mixed methods	402
Application of transformative mixed methods design	404
Gender-based violence: heteronormativity and LGBTQ+	404
People of color	405
Refugees and immigrants	405
Disabilities and deafness	406
Indigenous populations	407
Implications for teacher research	408
Conclusions	409
References	409
Relevant website	410

Transformative mixed methods research

Educators face many challenges in their attempts to provide quality education to all learners. Although educators are not responsible for solving all the ills of society, they can be part of the solution by recognizing the historical inequities and current conditions that influence their abilities to provide needed educational services. For example, the Roma population in Europe are recognized as a population that has experienced multiple discriminations in the form of slavery, expulsions, persecutions, genocide, and criminalization (Flecha, 2014). A school in Albacete, a city in Spain, served the Roma population in a neighborhood that was conflict-ridden; teachers reported feeling threatened entering the school; and the media blamed the Roma. Flecha and his colleagues asked the question: How can educational research contribute to improved academic achievement, a safer environment, social inclusion, and empowerment? They adopted a transformative framework for their research because it is characterized by challenging existing power relationships, explicitly addressing the basis of discrimination, and creating data that support positive transformative change. The use of mixed methods within this framework allowed them to collect data that were responsive to the cultural context and inclusive of the voices of the full range of interested parties.

The transformative lens for research arose as a response to concerns raised by members of marginalized groups that they felt they were either not included appropriately or were misrepresented in research (Mertens, 1999, 2018, 2020a; Mertens and Wilson, 2019). Members of these communities have expressed their experience with research as either leaving them worse or no better off and have developed a high level of skepticism about the value of research. For researchers to contribute positively to change in communities, researchers need to critically examine the research assumptions about what makes for ethical research, the privileging of accepted versions of reality, the relationship between the researcher and participants, what is valued as knowledge, and the appropriateness of methodologies to support transformation. This is the core of the transformative approach to mixed methods. Following the explanation of the assumptions associated with a transformative approach to mixed methods, this chapter includes a description of multiple examples that illustrate its application in pursuit of increased justice for members of marginalized communities.

Assumptions associated with transformative mixed methods

The ethical assumptions that embody a transformative approach raise questions such as: What methods are needed in order to be culturally responsive to the full range of participants and those effected by the research? How can the research be designed to promote social justice and further human rights? How can the research explicitly deal with historical and current discrimination in order to reduce inequities? How can communities be engaged in ways that incorporate their strengths and resilience so that they are better off as a result of their participation (Mertens, 2020a)? Such questions provide opportunities for researchers to design studies that go beyond “giving voice” to members of marginalized communities to enable the needed transformative change and create conditions that lead to a more just and equitable society (Cook, 2014).

In 1903, W.E.B. Du Bois asked the question: “How does it feel to be the problem?” (1903/2003, p. 7) He was writing about the tendency in the United States to blame Black people for the inequitable conditions in which they live and the negative consequences of living in poverty, sometimes associated with high rates of violence. His words have a striking similarity to the example provided by Flecha’s (2014) study cited earlier: The Roma are the problem. Ontological assumptions relate to the understanding of the nature of reality. One version of reality that is sustained by an oppressive status quo is that members of the marginalized communities are the problem. The use of a transformative lens asks researchers to examine the social positionalities associated with different versions

of reality and to design research that reveals versions of reality that sustain an oppressive status quo, as well as to reveal those versions that contribute to a positive transformative future. For example, research designed to smooth the transition of youth involved in the criminal justice system back to school hit a snag when the researchers found that some of the school personnel were not supportive of the program (Drill et al., 2017). The school administrators held a version of reality that characterized the returning students as troublemakers and that the transition program was “dumping students off” (p. 38). The researchers needed to use a research lens that allowed them to collect data about this version of reality and its consequences, as well as make visible the version of reality from the transition team about the support services needed for successful transitions. A transformative, culturally responsive lens allowed them to document both versions of reality and follow through on the consequences of allowing an oppressive version to prevail.

The transformative epistemological assumption pertains to the relationship between the researcher and participants, as well as to the nature of knowledge. It raises the questions: Whose knowledge counts as valid and what kind of relationships allow for the respect of knowledge from marginalized communities? This assumption bears close relationship with the transformative ethical assumption regarding the need to be culturally responsive. Methods are needed to increase understanding of the history and context in order to get an accurate picture of the heterogeneity of the communities in the studies. The dimensions of diversity that are relevant are contextually dependent. For example, in many high poverty areas in India, girls are valued less than boys; this is used to justify denial of their rights to education. In Malaysia, many Deaf people are not allowed to attend school and thus read poorly or not at all (Wilson and Winiarczyk, 2014). The 2019–2020 pandemic has revealed gaps in access to education as many school systems changed to online instruction, with people from lower socio-economic status and students of color experiencing less access. A complicating factor is the ability of some parents to be home to support their children’s instruction. Many parents work at jobs in which they have to be physically present (e.g., grocery store clerks, nurses, postal carriers, sanitation workers). Oftentimes, parents with lower incomes both have to work to maintain their ability to provide the necessities of life. Worldwide, communities of color experience a larger gap in their ability to access online instruction (UNICEF, 2020). As this situation illustrates, attention to data that depicts the characteristics of relevance associated with power or marginalization is critical in order to support inclusion of all community members. The transformative epistemological assumption supports inclusion in culturally respectful ways of members of marginalized communities, as well as those who hold more power. It also supports valuing the knowledge and experience of all relevant community members and calls for the use of methods that are capable of accurately capturing the experiences of people who are marginalized.

The transformative assumptions of the nature of ethics, reality, knowledge and relationships guide methodological choices that can lead to respectful relationship building, contextual and historical understanding, and collection and use of data for transformative purposes. A cyclical design exemplifies one transformative mixed methods approach (See Fig. 1) (Mertens, 2018, 2020a; Bolinsson and Mertens, 2021). The first cycle includes identifying who needs to be included and how they can be included in respectful ways. This might include literature reviews, individual consultations, focus groups, review of extant data on demographic characteristics and school achievement, or group meetings to determine who is present, who is not represented yet, and strategies for working together. The second cycle would focus on a more thorough contextual analysis that could include collection of data from surveys, review of historical documents, community characteristics, and legal issues, as well as a more focused approach to extant data, e.g., on success rates transitioning from the criminal justice system to school. The third cycle would make use of the data collected in the first two cycles to determine what type of intervention might be appropriate to address the identified problems.

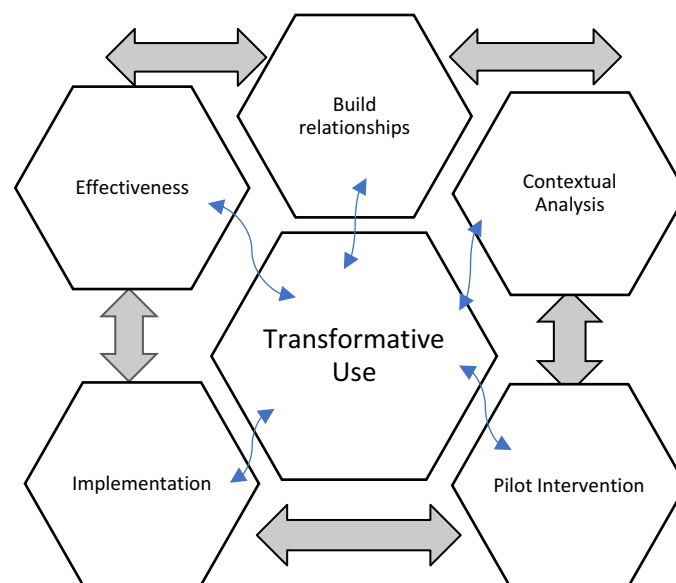


Fig. 1 Transformative mixed methods iterative cyclical approach.

Data collection during this phase would be intensive and would provide feedback to the program developers about adjustments that might be needed to improve recruitment, training, and the quality of the program implementation and results. If the data support moving forward, a fourth cycle could be used to study the implementation of an intervention on a wider scale. The fifth cycle would include collection of data on the effects of the intervention and the quality of the relationships that were formed during the first cycle. The final phase would be to determine the effectiveness of the intervention. The cycle is viewed as iterative as each stage informs the next and any stage can be re-visited based on what is learned in the prior stages. Use of the processes, experiences, and results of the research for transformative purposes can occur throughout the research process, e.g., to improve the intervention, expand dissemination, improve relationships, or change policy.

Mixed methods provide opportunities to gain understandings of the complexity of populations and contexts so that appropriate interventions can be designed and tested. The specific designs and methods used in a transformative mixed methods studies are inclusive of many of the designs and methods used in other paradigmatic frameworks (e.g., pragmatism). However, it does lead the researcher to ask different kinds of questions in the determination and implementation of a design. For example, if a researcher decides to use a randomized controlled trial as a part of a transformative mixed methods design, a transformative lens would lead to questions such as: “Who would easily access and benefit from this treatment? What additional investments would be required to make this treatment accessible to all? What cultural practices enable the adoption of this treatment, and what adjustments might be made to make receiving the treatment more culturally accepted?” (Mertens and Catsambas, 2022)."

Application of transformative mixed methods design

Transformative mixed methods have been used across the globe with many different populations. These studies have focused on people with disabilities, Deaf people, women and girls, members of the LGBTQ+ community, people of color, Indigenous populations, internally displaced persons, undocumented and refugee students, and children in high poverty areas and areas of conflict. The examples presented in this section illustrate the application of transformative mixed methods through an analysis of a number of studies with various marginalized populations in diverse locations.

Gender-based violence: heteronormativity and LGBTQ+

Sexual violence is a worldwide problem that occurs in schools and elsewhere (DeGue et al., 2012). Efforts to address this problem have occurred at all levels of education and have taken a number of different forms. Some programs educate girls and women on how to avoid threatening circumstances while others educate boys and men on the valuing of women and intervening when sexual violence is threatened or bragged about. Puigvert et al. (2019) conducted a transformative mixed methods study to address gender-based violence in Spanish universities that did not distinguish between heteronormative participants and members of the LGBTQ+ community. Their approach utilized a communicative methodology that aligns with the transformative approach in that it “entails the assessment of actual social improvements related to research outcomes by collecting data and engaging in dialog with manifold stakeholders (p. 362).” They conducted the study over several years. They described their methods as having two phases: first, they identified key informants and faculty and students, and established a safe environment for their participation. Then they examined contextual issues related to gender-based violence in universities. The second phase used the communicative method to evaluate the social impact of their study to reveal the transformations that occurred.

In the first phase of the study, the researchers began with a literature review of historical and current status of gender-based violence and interventions strategies for prevention and response. However, their approach incorporated transformative methods by including review of literature that specifically addressed wider issues of inequality in this context and the “mechanisms and policies regarding the role of the state, social movements, and other agents involved in fighting against violence in 10 European countries (p. 363).” They examined Spanish legislation related to gender-based violence and women’s rights, as well as historical attempts to address this issue in universities through mobilization and using the universities’ committee structures. They concluded that there was a culture of silence, people knew who the perpetrators were but no one dared to report them. They directly confronted issues of power and they collected data from social networks on the oppressive tactics of powerful university administrators and faculty to silence the researchers and their students by threatening their academic careers. They formed relationships with senior faculty as well as with students, nontenured faculty, and staff in order to be able to conduct dialogs that were safe and informative. They also conducted a survey of students to determine the extent of gender-based violence in Spanish universities and what steps could be taken to overcome this violence. This was followed by conducting semi-structured in-depth interviews with faculty and administrators to obtain a more thorough picture of the situation. Students participated in communicative dialog sessions in which they reflected on the results of the survey; they reported lack of recognition of gender-based violence, institution resistance to address this issue, and the absence of mechanisms to respond to this reality. All participants identified the need for safe spaces, support, and the need for a public debate to stimulate change.

The second phase of the study began with widespread dissemination of the results of the first phase through multiple methods, such as dialogs in diverse settings in the university, professional presentations, publications, and documentary film. The dissemination activities were complemented by conducting additional communicative dialogs with key informants (faculty and students) to obtain data about the repercussions of acting as social agents of change and the extent to which Spanish universities had become safer spaces.

The transformations that resulted from their study included generating new legislation, creating proper conditions for reporting abuse, and establishing solidarity with and among victims with the goal of achieving violence-free universities. The transformations

were enabled by engagement with a wider range of constituencies that resulted in mobilization of students around this cause, the establishment of peer-to-peer support networks, and solidarity between faculty and students committed to further action on gender rights. This formation of coalitions to be party to the research and to carry on with the transformative changes needed has become an important part of the transformative research process, not just in this study, but also in many projects funded by the US government (Wolfe et al., 2020).

Zea et al. (2014) provide an example of a transformative mixed methods study that focused on gay and bisexual men and transwomen in Colombia, South America. They described the process of building relationships with members of this community so that the participants felt safe and respected before any data collection began. They gathered qualitative data on the culture that was rife with hegemony for members of this community, as well as on resistance from political and church officials. They conducted qualitative life-history interviews with participants that covered such topics as: "childhood, adolescence, family life, sexual orientation development, gender identity, experiences of discrimination and stigma, experiences of internal displacement and violence, sexual risk, sexually transmitted infections, HIV status, and knowledge and attitudes toward circumcision" (p. 215). Based on the data from the life-history method, they constructed a quantitative survey that was administered in a safe space, a community center recently renovated for this population to safely gather. The results were shared interactively with activist groups, representatives of the Colombian government health sector, and international organizations working on HIV/AIDS in Columbia with the intent of changing policies for this marginalized community.

People of color

Inequalities associated with being a person of color are widespread; addressing these in an educational setting is fraught with challenges (Neubauer et al., 2020). Several researchers have used a transformative mixed methods approach to address inequalities based on racial or ethnic status such as Machimana et al. (2018) study of university engagement with rural communities in South Africa and Mertens' (2020b) study of the reduction of gang violence in school neighborhoods in Trinidad. The example used in this section expands on the study of the transformation of a school in Spain that was conducted by Flecha (2014) that was mentioned at the beginning of this chapter. The study is situated in an underprivileged neighborhood with many Roma families and was structured to include Roma at a grassroots level in a longitudinal 5-year study. The study began with discussions with Roma and non-Roma families, teachers, school administration, the researchers, policy makers and other stakeholders. They conducted observations and interviews with these constituencies to develop a picture of student and teacher experiences.

They developed a plan to shut down the dysfunctional school and open up a new one in the same building. The community members were involved in the development of the intervention through interactive dialogs in which researchers provided information about successful educational practices. These practices were adapted to fit the specific context based on community feedback. Community members initially expressed skepticism about the project, but through concerted effort, trusting relationships were built. A case study design was used that incorporated the collection of quantitative data on achievement and satisfaction, as well as qualitative data on the quality of the experiences and wider influences on health and employment. One important component of the study was the establishment of an advisory council. Membership on the council included members of the most vulnerable groups, those who had a higher education degree, and those who had experience overcoming inequalities through community participation and social and political involvement. Data were collected at regular meetings of the advisory council that guided decisions about the future stages of the research.

Every year for 4 years, the researchers collected quantitative data about school attendance and achievement as well as the quality of experiences through two questionnaires given to students and families. Leaders in the Roma community took responsibility for distributing the questionnaires to local people, especially those who were most isolated. Qualitative data were collected using four methods. Community activists, school administrators and professionals participated in standardized open-ended interviews to determine their professional experiences. Communicative daily stories were collected from family members and students that included personal reflections and perceptions of daily life through a conversational format. Focus groups were used that allowed groups to reflect on findings from the study. Observations were also conducted that allowed researchers to ask participants what they observed and experienced. The researchers' choice of observation spaces was informed by information provided by Roma leaders and community members who had valuable information about the school, neighborhood, history of socioeconomic problems, and relationships with the local administration. The researchers modified the instruments for quantitative data collection and the qualitative protocols every year of the study based on input from community members.

The data were analyzed to address this question: "How was it helping to reduce or prevent inequalities and marginalization and fostering social inclusion and empowerment?" (p. 250) Transformative changes were evident throughout the study. For example, a new school was developed and implemented; family participation in the school increased over the five years; student achievement increased in language and mathematics; families and students expressed higher levels of satisfaction with their progress. Active participation in decision making yielded more support for the schools, better experiences for students, increased health in family relationships, higher literacy rates for families overall, and further positive effects in terms of housing, health, employment, and social and political participation.

Refugees and immigrants

Immigrants cross borders, sometimes legally and sometimes illegally, often to escape intolerable conditions characterized by violence, poverty, and starvation. Sometimes undocumented immigrants are brought to the United States as children where they have lived the majority of their lives with a constant fear of being sent back to their birth country, even though they have no

familiarity with that country. In the United States, a great deal of turmoil emerged as the number of undocumented immigrants increased in the last decade, resulting in separation of families and deportation of many asylum seekers. [Barnhardt et al. \(2018\)](#) conducted a transformative mixed methods study that examined how campus administrators navigate the process of serving undocumented immigrant students in the Southeastern United States. They chose this approach because they saw a need to apply a research approach that contributes to understanding how the oppression of vulnerable populations is institutionalized in university settings.

The researchers began by collecting contextual data about immigration legislation and policy in the United States and how it is applied by local authorities. They documented the different versions of reality, some of which lead to oppression for undocumented students. They looked proactively at how transformative change could lead to changes in practices to be more inclusive and supportive. They focused on administrative processes because they were aware of the power structures that enable an oppressive status quo to continue. "Our unique focus on collective and organizational processes helps illuminate the multiple and competing ways in which institutional power and organizational power (or lack thereof) manifest when enacted by campus administrators and leaders operating in a particular localized cultural context (p. 414)." They continued: "transformation also emerges from disrupting structural and institutional oppression that can be addressed by understanding complex cultural realities and organizational responses to them (p. 415)."

The researchers began by forming a partnership with campus professionals at campuses across the Southeastern United States. The team included a lifelong activist in the undocumented student movement, campus administrators and educational researchers. This initial partnership identified a need to conduct a contextual analysis to identify the uniqueness of the local environments in terms of culture, political variables, power structures and economic status.

The researchers reviewed academic literature on organizational responses to contentious issues to gain insight into how an organization takes an ethical stand with regard to social responsibilities. They also analyzed federal and state policies to identify the barriers experienced by undocumented university students, such as banning such students from attending state or public universities, or lack of access to in-state tuition or obtain financial aid. In the United States, out-of-state tuition can be up to two to three times higher than in-state tuition, resulting in an increase of an average of \$8500 US ([National Center for Educational Statistics, 2018](#)). Such policies are reflective of public attitudes toward immigration in the United States that is revealed in local newspaper articles. The researchers used extant state-level data to identify the demographics, economic conditions, and political views about immigrants in the South Eastern United States as compared to the nation. They also used extant data from a higher education survey administered to 410 campus administrators.

They selected two 4-year campuses in different states in which to conduct case studies. They collected quantitative descriptive data about each campus, such as enrollment, race and ethnicity, amount of tuition, and percent of students receiving aid. For the qualitative part of the study, the researchers reviewed 215 newspaper articles published in newspapers in the two communities to collect data on how issues about undocumented students were reported. In one of the communities, a movement was started to pass a law to allow undocumented students to pay in-state tuition; the state's governor opposed the bill. The legislation was not passed and the universities were ordered to do an audit on all undocumented students and report on the tuition pricing. In the other case, a movement was started to allow an undocumented student to avoid deportation. The situation was widely reported and the student was allowed to stay because of her student status. However, the state legislator passed restrictive laws pertaining to admission and tuition pricing for undocumented students the next year.

The researchers included a stage of dissemination of findings for transformative purposes. They summarized the transformative effect as being the result of on-going engagement with campus practitioners:

Our assessment aided them in providing tangible patterns to draw on, and nuanced details about how the community privileges certain meanings that maintain oppressive realities. It was the integration of direct feedback from campus practitioners at all stages, from initiation of the research problem, to data collection, to analysis, and reanalysis that allowed us to shape the research process and the understanding of the issues that helped practitioners find a voice toward transformation (p. 431).

The practitioners reported that they had developed a sharper understanding of their complicit involvement in sustaining oppressive practices. The researchers continued their relationship with the campus collaborators through direct outreach and continued communication. They provided evidence to the campus practitioners and administrators about addressing power dynamics locally in order to be able to advocate for better practices in their institutions and professional organizations. Their collaborators developed an online repository of resources for campus leaders who wanted to advocate for the rights of immigrant and undocumented students. They have also produced policy briefs, trend reports, short factsheets, webinars, workshops, and syllabi to support critical examination of policies and avenues to support undocumented students.

Disabilities and deafness

The Secretary-General of the United Nations released a statement that strongly demands supportive inclusion of persons with disabilities and those who are Deaf in order to respect their right to receive an education ([UN General Assembly, 2020](#)). As noted earlier, students with disabilities and students who are Deaf are denied access to appropriate education in many parts of the world ([Wilson and Winiarczyk, 2014](#)). Deaf people are sometimes not provided with appropriate interpreters, a necessary service if they are to receive an education. Wilson and Winiarczyk conducted a transformative mixed methods study in Malaysia that included in-

depth interviews, focus groups, surveys, and naturally occurring observations. As both researchers are American, they employed the services of a Malaysian research facilitator who was fluent in Malaysian Sign Language and American Sign Language (ASL) (Winiarczyk is a Deaf researcher who uses ASL). They consulted as a team with the Malaysian Deaf Association and members of the Deaf Malaysian community on the appropriateness of the research objectives, questions, and data collection instruments and strategies to insure cultural responsiveness. They also used Participatory Rural Appraisal research techniques to facilitate group discussions. The resulting data revealed the reality of the Deaf Malaysians that was characterized by a lack of policy implementation to provide them with early childhood education, deaf education, and accessibility through supportive communication (e.g., trained interpreters). The Deaf community members used the data and the strategies they learned through participating in the research study to begin to advocate for education and accessibility in Malaysia.

The U.N. Secretary-General's statement (2020) identified the increasing use of online instruction due to the 2019–2020 Covid-19 pandemic as presenting additional challenges for persons with disabilities. Hyett et al. (2020) conducted a transformative mixed methods study to improve the quality of online instruction for students with disabilities at a mainstream school in Australia. The researchers identified one of the relevant contextual factors in this study as a major policy change in Australia in the form of the National Disability Insurance Scheme that provided funding to support people with significant disabilities in mainstream society.

The methods used included transcripts from six workshops, surveys (completed at the end of each workshop), document reviews, and individual interviews. The participants included teachers and their assistants, parents, and allied health professionals (occupational therapists and speech pathologists). The use of a codesign process to develop the intervention illustrates the use of a transformative lens in this study in that the researchers did not impose an intervention that they identified. Rather, the researchers developed relationships with educators, health professionals, and parents as equal partners in the research process. The knowledge that educators, health professionals, and parents gained from being in continuous and long-term contact with students with disabilities was valued and informed the development process. This knowledge was elicited by using a storytelling strategy during the workshops. The researchers emphasized "process, relationships, governance, and effective facilitation that prioritized power sharing and group dynamics" ((Greenhalgh et al., 2016) cited in Hyett et al., 2020, p. 2).

Representative educators, allied health professionals, and disability consultants (two parents of children with disabilities and two persons with disabilities) made up a governance council (advisory group) to guide the research team in decision making. The advisory group and the research team designed processes for engaging with these groups throughout the entire research process, from problem identification, understanding contextual issues, developing an intervention and processes for implementation, and collection of data throughout the study. The researchers chose not to include children with disabilities based on the rationale that they would not be the end users of the product. They further justified this decision based on the input they received from the persons with disabilities in the advisory group. The researchers specifically described how they addressed issues of power: "Strategies used to address power dynamics during this project were equal compensation, governance committee structure, group facilitation techniques, a self-report measure, and activities designed to promote equal participant contributions (p. 11)."

Transformations occurred for the participants during the codesign process in the form of increased understandings of the perspectives of those from the different groups, enhanced listening skills, and valuing of the codesign process. The online package for instruction was developed and approved through the workshop process, followed by a review by the advisory group. After making modifications based on input from these groups, the researchers posted the package to a website for external review by a wider professional community. After making additional revisions, the final online package was disseminated by posting to an open access website.

Additional examples of transformative mixed methods with populations who are Deaf or have a disability include: Hyett et al. (2020) used the transformative mixed methods approach to study strategies for improving inclusion of people with disabilities in educational programs in Australia. Munger and Mertens (2011) looked at ways to respectfully include adult people with cerebral palsy in decisions about their education in the United States. Wilson and Winiarczyk (2014) studied how to increase the educational supports for people who are Deaf in Malaysia, Chile and Argentina. Kutscher and Tuckwiller (2019) conducted a study to increase support for students with disabilities at the college level in the United States.

Indigenous populations

Indigenous scholars draw a distinction between the use of an Indigenous worldview and a transformative lens as frameworks for evaluation that share a common interest in positive transformation for those who experience discrimination and oppression (Chilisa, 2020; Chilisa and Mertens, 2021; Chouinard and Cram, 2020; Cram and Mertens, 2015, 2016; Mertens and Cram, 2016). An Indigenous perspective is rooted in the need to address decolonization as it relates to "sovereignty, land and resources, and freedom to live as Indigenous peoples" (p. 186) (LaFrance and Nichols, 2009, cited in Mertens and Cram, 2016). While both transformative and Indigenous researchers seek transformation to produce a more just society, Indigenous researchers see the importance of transformation in the form of repatriation of land and their relations to the land. This is a concept that is compatible with transformation for increased social, economic and environmental justice, but it is also quite specific to an Indigenous worldview. Both frameworks endorse the purpose of research as supporting change to increase the quality of lives for people who have experienced marginalization and discrimination. The Indigenous worldview adds to understanding the meaning of transformation by its focus on the ethics of developing relationships that are culturally respectful and that honor all those who came before, are here now, and will be here in the future. Framing research in this way raises this question: How would we design our research to incorporate this interpretation of ethics, develop ethical relationships, and ensure that the interventions that are implemented honor the past, present and future inhabitants? Chilisa (2020) and Chilisa and Mertens (2021) argue that the Indigenous framework needs to

be given priority in research with Indigenous communities. However, they also agree that the Indigenous framework can be combined with a transformative lens to create research that focuses on positive transformative change.

The HIV/AIDS epidemic has ravaged many communities, particularly in Southern Africa. Researchers from the University of Botswana and an American university partnered to address this problem through the development a program to prevent the spread of HIV/AIDS and other sexually transmitted diseases for adolescents in Botswana (Chilisa and Tsheko, 2014). The researchers framed their study from a postcolonial Indigenous perspective, deliberately chosen to avoid the dominance of Western approaches. They combined quantitative and qualitative data collection methods within a randomized controlled trial design. The research study began with forming relationships, including establishing three advisory groups: an external international expert research advisory group, a school and community professional and parent group, and a youth group. These groups discussed what it meant to decolonize the mind and research methodologies. The community advisory group was described as follows:

We formed community advisory boards to serve as community theorists that could bring to the research the cultural knowledge, values, and processes of knowledge creation that could serve to make the interventions we designed relevant and useful to the communities. Their main role was to decolonize the research process by bringing Botswana worldviews to the creation of knowledge to ensure that the research remained relevant and useful to them (p. 227).

Based on the advisory groups input, the researchers conducted preliminary data collections to understand the current status of HIV/AIDS and sexually transmitted diseases using both qualitative and quantitative methods. They used Indigenous qualitative methods of storytelling and examining local proverbs in local languages, as well as interviews, knowledge and attitudinal surveys, and extant data to obtain incidence data of disease disaggregated by age, gender, and tribe. The data revealed that the students had accurate knowledge of prevention and transmission of the diseases for the most part, however, they associated a sense of sadness and futility around the concept because they had lost so many loved ones and felt they might die too.

A second phase used Indigenous mixed methods to develop a culturally relevant survey to measure the beliefs, behaviors and attitudes toward risky behaviors. The quantitative survey was based on the qualitative data from interviews, but also on cultural knowledge that came through stories, myths, proverbs, songs, metaphors, and taboos. These methods resulted in the surfacing of perceptions of students about sex and sexuality that were not common in the literature, but were reflective of the cultural context. The community advisory board then reviewed the instrument to insure the acceptability of the content and language.

An intervention was then developed that reflected the need to address cultural context and emotional needs, rather than strictly focusing on knowledge of transmission and prevention (e.g., use of condoms, limiting number of partners). The intervention was situated in a school context and included 12 1 hour modules, and was led by trained youth. This intervention was then pilot tested with rich data collected via observations, interviews, and document reviews. Following revisions of the program based on the pilot data, the researchers employed a randomized experimental design to determine its effectiveness. The experimental group received the HIV/STI program while the control group received a health promotion intervention that addressed the risks of other diseases such as heart disease and diabetes. The researchers incorporated Indigenous mixed methods within the experimental framework, combining participatory action research with appreciative inquiry. "Indigenous conversational methods, namely, adolescents talking to their parents, talking circles and yarning, were used as data collection methods. One of the criticisms of participatory action research is that most of the approaches are problem focused aiming at discovering communities' unmet needs (p. 230)." The yarning activity was used on the last day to evaluate participants' views on the effectiveness of the intervention. It consists of a facilitator who holds a ball of yarn and sharing what they have learned; the ball is then thrown to the next person until each person has a chance to talk. The end result is a visual picture of a web of connections that shows how everyone is connected to everyone else. The participants reported being transformed by building relationships, developing positive identities, and creating images of hope and a brighter future. The curriculum was then shared more widely in the Botswana school system and the results were shared with policy makers to inform future educational policies. The importance of addressing not just the acquisition of knowledge but also the need to build confidence in a brighter future is key to making healthy behavioral choices.

Implications for teacher research

The paradigm shift in the research world toward a framework that is transformative and explicitly addresses issues of equity has implications for the preparation of teachers and for teacher research. Teachers in the United States are primarily White, hearing, CIS-gendered, able-bodied, middleclass women. This creates a challenge for the majority of the teachers in the workforce to understand the experiences of people of color and people with disabilities, those who are Deaf, members of the LGBTQ+ community, or students whose backgrounds and living conditions are not familiar to the teachers. Accrediting agencies are giving greater attention to issues of diversity and reducing disparities in educational achievement on the basis of race, gender, and disability (Beck, 2020). The goal of achievement of those impoverished and marginalized is not only an issue in the United States, many countries are exploring how to address educational disparities. It is one of the Sustainable Development Goals endorsed by the United Nations for achievement by 2035 (Chilisa and Mertens, 2021).

Beck (2020) recommends the use of a transformative lens in preparation and research as a mechanism for bridging the gap, especially in communities with a majority population that is Black, Latinx, or Indigenous. Her rationale is presented in the following quotation:

... transformative mixed methods research (Mertens, 2010) has the explicit purpose of responding to societal inequities, engaging culturally diverse groups, and advancing social justice. Mertens (2010) noted that this model “supports the use of a cyclical model in which community members are brought into the research process from the beginning and throughout the process in a variety of roles” (p. 472). Thus, this model privileges a variety of members in the research process ... Cultural competence is critical for researchers utilizing transformative mixed methods. This design begins with a qualitative component to learn about the community being studied and to develop trusting relationships between stakeholders. The focus in this design is on producing change.

Beck (2020, p. 385).

A transformative mixed methods research agenda in teacher education provides a framework for educational researchers to work with schools and community members to prepare teachers who are equipped to address the challenges of discrimination and oppression. Researchers can thus demonstrate the value they place on knowledge based on lived experience that is sometimes missing in research that is designed based on scholarly literature alone. They can design studies that address differences in power in the typical hierarchical structure of schools and universities. Development of interventions can be based on theory and practice vetted through the eyes of those with lived experience. This provides an opportunity for schools to engage in continuous improvement through the cyclical collection of data, review and interpretation, and revisions based on earlier stages of research.

Conclusions

Educators experience the diversity of students in their institutions and recognize that different approaches are needed in order to address disparities in achievement. The transformative lens provides insights into the basis for disparities through engagement with the full range of people impacted by educational services. It brings attention to the need to be inclusive in building relationships so that the knowledge from researchers, school personnel, students, parents and other community members can be used to inform understandings of the nature of the problem and of interventions that are rooted in the culture and context of the community. Transformative mixed methods research incorporates multiple sources of data collected through multiple culturally responsive strategies with the goal of using the process and the findings to increase social justice.

References

- Barnhardt, C.L., Reyes, K., Rodriguez, A.V., Ramos, M., 2018. A transformative mixed methods assessment of educational access and opportunity for undocumented college students in the Southeastern United States. *J. Mix. Methods Res.* 12 (4), 413–436.
- Beck, J.S., 2020. Investigating the third space: a new agenda for teacher education research. *J. Teach. Educ.* 7 (4), 379–391.
- Bolinson, C., Mertens, D.M., 2021. Transformative evaluation and impact investing: a fruitful marriage. In: De Orais Sarmento, E., Paul Herman, R. (Eds.), *Global Handbook of Impact Investing*. Wiley, Hoboken NJ, pp. 697–734.
- Chilisa, B., Mertens, D.M., 2021. Indigenous Made in Africa evaluation frameworks: addressing epistemic violence and contributing to social transformation. *Am. J. Eval.* 1–13.
- Chilisa, B., Tsheko, G.N., 2014. Mixed methods in indigenous research: building relationships for sustainable intervention outcomes. *J. Mix. Methods Res.* 8 (3), 222–233.
- Chilisa, B., 2020. *Indigenous Research Methodologies*, second ed. Sage, Thousand Oaks, CA.
- Chouinard, J.A., Cram, F., 2020. *Culturally Responsive Approaches to Evaluation*. Sage, Thousand Oaks, CA.
- Cook, J.R., 2014. Using evaluation to effect social change: looking through a community psychology lens. *Am. J. Eval.* 36 (1), 107–117.
- Cram, F., Mertens, D.M., 2015. Transformative and indigenous frameworks for multimethod and mixed methods research. In: Hesse Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford: UK, pp. 91–109.
- Cram, F., Mertens, D.M., 2016. Negotiating solidarity between indigenous and transformative paradigms in evaluation. *Eval. Matters* 2, 161–189.
- DeGue, S., Simon, T.R., Basile, K.C., Yee, S.L., Lang, K., Spivak, H., 2012. Moving forward by looking back: reflecting on a decade of CDC’s work in sexual violence prevention, 2000–2010. *J. Wom. Health* 21 (12), 1211–1218.
- Drill, K., Stack, E.E., Qureshi, C., Lahoff, R., 2017. Oregon’s African American/Black Student Success Plan. RMC Research, Portland, OR.
- Du Bois, W.E.B., 2003. *The Souls of Black Folks*. Barnes & Noble, New York, NY (Original work published 1903).
- Flecha, R., 2014. Using mixed methods from a communicative orientation: researching with grassroots Roma. *J. Mix. Methods Res.* 8 (3), 245–254.
- Greenhalgh, T., Jackson, C., Shaw, S., Janamian, T., 2016. Achieving research impact through co-creation in community-based health services: literature review and case study. *Milbank Q.* 94 (2), 392–429. <https://doi.org/10.1111/1468-0009.12197>.
- Hyett, N., Bagley, K., Iacono, T., McKinstry, C., Spong, J., Landry, O., 2020. Evaluation of a codesign method used to support inclusion of children with disabilities in mainstream schools. *Int. J. Qual. Methods* 19, 1–12.
- Kutscher, E.L., Tuckwiller, E.D., 2019. Persistence in higher education for students with disabilities: a mixed systematic review. *J. Divers. High. Educ.* 12 (2), 136–155. <https://doi.org/10.1037/dhe0000088>.
- LaFrance, J., Nichols, R., 2009. *Indigenous Evaluation Framework: Telling Our Story in Our Place and Time*. American Indian Higher Education Consortium, Alexandria, VA.
- Machimana, E.G., Sefotho, M.M., Ebersohn, L., 2018. What makes or breaks higher education community engagement in the South African rural school context: a multiple-partner perspective. *Educ. Citizen Soc. Justice* 13 (2), 177–196.
- Mertens, D.M., Catsambas, T., 2022. Ethical practice through a transformative lens and methodological implications for practice. In: van den Berg, R.D., Hawkins, P., Stame, N. (Eds.), *Evaluation in an Ethically Challenged World. Doing “Good,” “Bad,” or “No Harm”?*, pp. 165–187.
- Mertens, D.M., Cram, F., 2016. Integration tensions and possibilities: indigenous research and social transformation. *Int. Rev. Qual. Res.* 9 (2), 185–191.
- Mertens, D.M., Wilson, A.T., 2019. *Program Evaluation Theory and Practice*, second ed. Guilford Press, New York, NY.
- Mertens, D.M., 1999. Presidential address: inclusive evaluation: implications of transformative theory for evaluation. *Am. J. Eval.* 20 (1), 1–14.
- Mertens, D.M., 2010. Transformative mixed methods research. *Qual. Inq.* 16 (6), 469–474. <https://doi.org/10.1177/10778004>.
- Mertens, D.M., 2018. *Mixed Methods Design in Evaluation*. Sage, Thousand Oaks, CA.
- Mertens, D.M., 2020a. *Research and Evaluation in Education and Psychology*, fifth ed. Sage, Thousand Oaks, CA.
- Mertens, D.M., 2020b. Transformation as a goal of mixed methods research in the Caribbean. *J. Mix. Methods Res.* 1 (1), 16–28.

- Munger, K., Mertens, D.M., 2011. Conducting research with the disability community: a rights based approach. In: Rocco, T.S. (Ed.), *Challenging ableism, understanding disability, including adults with disabilities in workplaces and learning spaces*. Jossey Bass, San Francisco CA, pp. 23–34.
- National Center for Educational Statistics, 2018. Digest of Educational Statistics: 2018. Author, Washington, DC.
- Neubauer, L.C., McBride, D., Guajardo, A.D., Casillas, W.D., Hall, M.E. (Eds.), 2020. Examining Issues Facing Communities of Color: The Role of Evaluation to Incite Change. *New Directions in Evaluation*, 166, Summer 2020.
- Puigvert, L., Valls, R., Yeste, C.G., Aguilar, C., Merrill, B., 2019. Resistance to and transformationsof gender-based violence in Spanish universities: a communicative evaluation of social impact. *J. Mix. Methods Res.* 13 (3), 361–380. <https://doi.org/10.1177/1558689817731170>.
- U.N. General Assembly, 2020. Inclusive Development for and With Persons With Disabilities: Report of the Secretary General. U.N. General Assembly, New York, NY.
- UNICEF, 2020. Covid-19: Are Children Able to Continue Learning During School Closure? UNICEF, New York, NY.
- Wilson, A.T., Winiarczyk, R.E., 2014. Mixed methods research strategies with deaf people linguistic and cultural challenges addressed. *J. Mix. Methods Res.* 8 (3), 266–277.
- Wolfe, S.M., Price, A.W., Brown, K.K. (Eds.), 2020. Evaluating Community Coalitions and Collaboratives. *New Directions for Evaluation*, 165, Spring 2020.
- Zea, M.C., Aguilar-Pardo, M., Betancourt, F., Reisen, Gonzales, F., 2014. Mixed methods research with internally displaced Colombian gay and bisexual men and transwomen. *J. Mix. Methods Res.* 8 (3), 212–221.

Relevant website

Transformativeresearchandevaluation.com.

Integration in mixed method research

Elizabeth G. Creamer, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	411
Fully integrated mixed method research	411
Distinguishing between mixing and integration	412
Estimates of the frequency of meaningful integration in mixed method designs	412
Navigating the aftershock of the paradigm wars	412
Philosophical grounding that supports an interactive and iterative approach to integration	413
Procedures that promote integration	413
Framing a research problem in a multi-dimensional way	413
Integrating through case-based analysis	413
Integrating through joint displays	414
Integrating through a methodological or conceptual framework	414
Other design features that promote integration	414
Characterizing mixed priority research	415
Concluding thoughts	415
References	415

Introduction

Integration is central to any definition of mixed methods research. It is what distinguishes it from multimethod research where data from different sources are analyzed and sometime linked but not brought together in a meaningful way that impacts analytical insight. Integration involves strategies that allow for purposeful interaction between qualitative and quantitative data and/or analytical procedures. This chapter addresses both “what” and “how” questions about integration. It introduces different viewpoints about what is meant by the word “integration” as compared to “mixing.” The discussion about design and philosophical assumptions that facilitate its meaningful uses, and the identification of four integrative analytical strategies addresses the question of how integration occurs.

Both mixed methods and multimethod research share the motivation to more fully understand complex social phenomenon, such as about immigration or domestic and international terrorism in real-world contexts, and that this is best accomplished by accessing multiple different sources and types of data. Researchers using a multi-method approach often combine multiple qualitative approaches or multiple quantitative approaches (Plano Clark and Ivankova, 2016). Mixing is not achieved just at the technical level of by collecting both qualitative and quantitative data. Mixing can occur at multiple levels including by partnering more than one methodology, like mixed methods and case study, participatory action research, or grounded theory. It can also involve engaging different disciplinary perspectives or theoretical perspectives (Greene, 2015). “The possibility of mixing at *multiple levels* (emphasis hers) is integral to the characteristics of mixed method approach,” Jennifer Greene, an influential opinion leader in mixed methods, observed (2015, p. 607).

Fully integrated mixed method research

The logic of integration is most fully realized in a fully integrated mixed methods design (Fetters and Molina Azorin, 2017, 2019). Qualitative and quantitative strands of a research study are brought together at many points during the research process in a fully integrated mixed methods design. This is referred to as an embedded approach to integration (Fetters and Molina Azorin, 2017). Research done with a fully integrated approach are often difficult to classify with the conventional terminology used in mixed methods to describe the research design.

A fully integrated approach to mixed methods, for example, might be launched during the design phase with a research question that requires both qualitative and quantitative data to address. The same logic could be extended to sampling and could occur, for example, when both qualitative and quantitative data are used to identify a sample of participants. Similarly, both qualitative and quantitative data can be collected simultaneously, such as when a researcher engages a participant in co-constructing a map of his or her neighborhood. There are multiple strategies to integrate during analysis, including through joint displays that combine findings from quantitative and qualitative analytical procedures. At the final stage of the research process before reporting begins, findings from the different analytical techniques can be integrated to construct a figure that reflects a composite, holistic understanding of the phenomenon.

Distinguishing between mixing and integration

Some methodologists with expertise in mixed methods maintain that more attention needs to be devoted to distinguishing between the terms “mixing” and “integration” (e.g., [Fetters and Molina Azorin, 2017](#)). [Fetters and Molina Azorin \(2017\)](#) suggest that the term integration is often used too narrowly to refer to either merging or combining data. These authors suggest that the key to advancing the understanding of the term “integration” is to apply it to a fully integrated approach where qualitative and quantitative approaches interact in meaningful ways during all phases of the research process.

The word “integration” carries different meaning in the context of mixed method research than the word “mixing.” Integration refers to the type of purposeful engagement of different types of data or analytical procedures in ways that have an effect on the analytic density or conceptual multi-dimensionality. Analytic density refers to the potential to build the type of multi-dimensional conceptual understanding of social phenomenon that is essential to the iterative process of building knowledge in a scientific way ([Fielding, 2012](#)). No one approach or method is subordinated in this type of approach to mixed methods. The product is synergistic. That means that the outcome goes beyond what was extracted from the separate parts. In this case, it is much more difficult to disaggregate the findings during reporting into separate publications devoted to the qualitative findings and the quantitative results.

Estimates of the frequency of meaningful integration in mixed method designs

The potential to integrate findings from different analytical procedures in meaningful ways was rarely optimized when mixed methods first emerged as a distinct methodology. Some of this may be attributed to the social mindset in the 1980s during the so-called “paradigm wars” ([Howe, 1988](#)) that raged as qualitative approaches were beginning to gain legitimacy. [O’Cathain et al. \(2007\)](#) provided evidence that integration was not an every-day occurrence in mixed method research during the early days of mixed methods. Their systematic application of a rubric that tabulated at what phase in the research process mixing occurred, revealed the surprising finding that only a small percentage of studies (16% of 49) conceived as mixed methods, resulted in publications that provided evidence that the qualitative and quantitative strands were integrated. The practice of publishing qualitative and quantitative findings in separate articles complicates attempts to get a full picture of how integration occurs in a mixed methods study ([Archibald et al., 2015](#)).

It is challenging to reach an accurate estimate of how frequently integration occurs in mixed method publications. Drawing conclusions about integration from a single article is more feasible in a straightforward, two-phase design than it is in a more complex, multi-phase design where team members produce a much larger body of work. Some types of integration, such as how a theoretical framework was used as a guide to integration across research phases, cannot be gauged accurately at the level of a single study ([Maxwell et al., 2015](#)). Recognizing the complexity of the ways mixed methods are used in practice, particularly in a randomized control trial (RCT), [Creamer et al. \(2020\)](#) suggest that reaching a judgment about research quality in a multi-phase mixed method research project is better done by weighing more than a single article.

Navigating the aftershock of the paradigm wars

Discussion about philosophical paradigms has had a prominent role in the discourse about mixed methods since its inception as a distinct research methodology. The influence of postpositivist perspective in the mixed method discourse ([Denzin, 2010](#)) may account for the assumption that it is preferable to keep qualitative and quantitative analysis separate. The postpositivist paradigm dominated social science thought in the 1980s and 1990 when qualitative and quantitative research coexisted but as separate and competing paradigms ([Johnson et al., 2007](#)). The expression, incompatibility thesis, refers to the epistemological position that qualitative and quantitative data and analytical procedures should not be integrated prior to the final stage of interpretation because they emanate from antithetical philosophical positions about the nature of knowledge and legitimate ways of constructing it.

The idea that qualitative and quantitative approaches are incompatible paradigmatically still holds sway in some academic domains, like psychology and the health sciences, where experimental research is the gold standard. One way that this viewpoint is mirrored in the design of a mixed methods study is by keeping the qualitative and quantitative strands separate until they are brought together during the process of drawing conclusions. This further impacts research designs because it means that there are separate quantitative and qualitative research questions. The incompatibility thesis extends to those who hold fast to the idea that qualitative and quantitative data cannot be integrated because they reflect, as nursing educator [Morse \(2010\)](#) asserts, different levels of abstraction or interpretation.

A person writing about integration in mixed methods research for decades, Pat Bazeley, takes a strong position about designs that fail to engage qualitative and quantitative data in meaningful ways. “The strategy that is most destructive of coherent integration,” Bazeley writes, “is to work through a complex analysis of each of the separate sources or data types available *before* considering how they might be integrated” ([Bazeley, 2018](#), p. 283, emphasis hers). She joins others in challenging the incompatibility thesis by pointing out the similarities between research approaches with an empirical drive. Integration throughout the research cycle is paradigmatically supportable by the argument from constructivism that all research has a subjective, interpretive component.

Philosophical grounding that supports an interactive and iterative approach to integration

Mixed method approaches have been adapted as a research approach by investigators operating from paradigms with diverse epistemological and ontological assumptions about the nature of reality, what constitutes research, whether values play a legitimate role, and the types of tools that are needed to conduct it. Some writing about mixed methods affiliate with the lens of dialectical pluralism as a guide for investigators committed to the exploration of diverse perspectives (e.g., [Johnson, 2012](#)). The contribution of dissonance, paradox, incongruities, and gaps and silences that is part of the dialectical perspective is a defining characteristic of what Jennifer Greene famously referred to as a “mixed method way of thinking” (2007, p. 20). Unexpected findings can offer a powerful incentive to integrate findings from different sources of data in search of an explanation.

Procedures that promote integration

A number of analytical strategies promote meaningful integration, including by pairing with other methodologies, like grounded theory ([Creamer, 2021](#)). These include (1) framing a research problem in a multi-dimensional way, (2) integrating through case-based analysis, (3) integrating through visual displays, including joint displays, and (4) integrating through a methodological or conceptual framework.

Table 1 identifies four strategies for mixed method analysis and provides a definition for each.

Framing a research problem in a multi-dimensional way

Integration can be enhanced by conceiving social problems in a multi-dimensional way and by conceptualizing core constructs as nested within micro, meso, and macro levels of social structure. Micro level constructs relate to individuals’ beliefs, attitudes, interactions, motivations, and aptitudes. Meso level constructs could involve a community or organization. Macro level constructs reflect how individual perceptions or actions might more fully be understood within the broader context of the social and political environment. Using only one source of data necessarily represents a partial picture of a phenomenon. It is just one facet or dimension of a wider phenomenon.

Bi-dimensional thinking might pose a research question that considers the relationship between two variables. This could be the case, for example, in the exploration of the link between any type of intervention and anticipated outcomes. A more multi-layered, and thus theoretical, perspective is introduced by considering contextual factors or more local conditions that might mitigate the impact of the intervention or the flow in a temporal process.

Framing a research problem in a multi-dimensional way facilitates what could be called theoretical integration by creating an opportunity to link qualitatively and quantitatively derived constructs. In this context, connecting or integrating evidence occurs during analysis. It is not a mechanical or procedural matter, but a theoretical one that is generally accomplished with an eye to the literature and other relevant theoretical frameworks.

The purpose served is to gain a more holistic picture by exploring different aspects of the same phenomenon. This is referred to as complementarity ([Greene et al., 1989](#)). It is often the case in research involving interventions that quantitative data is used to measure outcomes and qualitative data to identify links between variables or to pinpoint an underlying social process ([Johnson and Schoonenboom, 2015](#)). There is nothing that prohibits that both qualitative and quantitative data could be used to corroborate the same construct or the links between them ([Bussing et al., 2012](#)).

Integrating through case-based analysis

Perhaps more so than any other procedure that I have identified, case-based analysis can provide the basis for integration at the multiple levels of the analysis that [Greene \(2015\)](#) referred to. In case-based analysis, narratives or visual displays are constructed that integrate data from multiple sources about the same unit of analysis. An individual, group, classroom, or location are examples

Table 1 Examples of mixed method analytical strategies that promote integration.

<i>Mixed analytical strategy</i>	<i>Definition</i>
Framing a research question	A research question that links qualitative and quantitative constructs to create a more multi-layered understanding of a social phenomenon.
Case-based analysis	Merges qualitative and quantitative data through the construction of a narrative or visual about each case for purposes of conducting a cross-case comparison.
Joint displays & other types of displays	Integrates qualitative and quantitative data or findings in a table, figure, matrix, or graph.
Integrated theoretical framework	Embeds multi-dimensionality in a figure or table by integrating qualitative findings and quantitative results, most often to offer an explanation for a social process.

of what can provide the unit of analysis for a case. The purpose of the construction of cases is formative in that they are used for the basis of further analysis through a cross-case comparison.

Case-based analysis create opportunities for integration at multiple levels. This includes integration of data from multiple sources into case narratives or profiles. It extends to integration that occurs through conducting cross-case comparisons with an abductive logic that occurs when preliminary inferences are integrated to generate tentative hypotheses which are subsequently explored further through an inductive or deductive logic. It also includes integration that occurs when an eye is kept on how emerging findings fit and do not fit with the literature. Internal validity is extended by the generation and testing of multiple, and on occasion, contradictory explanation that help to unsettle preconceptions, strengthen grounding, and distract from foreclosing prematurely on an obvious explanations.

An additional way that integration can occur through case-based analysis in a mixed method study is in the use of cases constructed with multiple sources of data to test an existing theory or to elaborate it. For example, support can be calibrated for the extent that each case supports different theoretical propositions. An investigator might turn to case profiles or memos that integrate data from multiple sources to explore if how an emerging theory might need to be refined to account for exceptions.

Integrating through joint displays

A figure or table that integrates qualitative and quantitative data is an integrative strategy that has probably received more attention than any other strategy in the mixed methods literature (Plano Clark and Sanders, 2015). A joint display is a table or figure that encourages integrative thinking during analysis by arraying and sometimes linking qualitative and quantitative data about the same constructs, research question, or themes.

A visual display that integrates qualitative and quantitative data can serve multiple roles in both developing and supporting findings. Their role is a formative one when the juxtaposition of findings from different sources of data generates new insight about a pattern or association. This can lead to a new round of theoretical propositions or themes that are pursued through further analysis and/or data collection even when the data at first seems contradictory.

An additional purpose for an integrative visual display is triangulation. This is when a table or figure provides enough data to support the credibility of conclusions. A well-crafted visual display can warrant a conclusion by showing the empirical support for it (Fetters and Guetterman, 2021). This type of visual display can document the quantitative and qualitative support for a hypotheses, theme, or construct. Its purpose is related to warranting because it shows the extent that key conclusions, inferences, themes, and/or hypotheses are corroborated by analysis from more than one source of data. This type of table generally reveals areas where findings triangulate, as well as areas where the data are not in accord (Creamer, 2021).

Integrating through a methodological or conceptual framework

It is an expectation in many academic disciplines in the human and behavioral sciences, such as psychology, that an investigator ground the design of a research study in a theoretical framework whose reliability has been documented. This is less so the case in applied fields like education and nursing. It is not uncommon for investigators in applied fields to launch a project with a conceptual framework produced from literature. Theoretical integration can be achieved through conceptual framework that combines components of more than one theoretical perspective. Over the life of a research project, a conceptual framework is generally subject to on-going refinement and revisions.

A conceptual framework can facilitate the integration of data and findings from multiple sources in a number of ways. One of its functions during the design of study is to single out a tentative set of core constructs that can be used initially to select appropriate instruments or to structure data collection instruments that are original to the project. A conceptual framework can provide a preliminary shell of a coding scheme to be applied qualitatively. A third way a conceptual framework is related to integration is to provide a preliminary explanation for incongruities that emerge between findings from different sources of data. When applied too prescriptively or used in a way that narrows the range of explanations considered, a conceptual framework can inhibit integration.

Other design features that promote integration

The design of a mixed method study is often represented in a figure that maps the sequence and timing of key phases of a research project. In addition to the analytical procedures described above, several features of design can promote integration in a mixed methods study. Most of these influence decisions that are made about the research design during the planning phase.

1. There is at least one research question that requires both qualitative and quantitative data to answer.
2. Both qualitative, quantitative, and sometimes visual or auditory, data are collected about core constructs.
3. Qualitative and quantitative data are linked in a single database (Bazeley, 2018; Creswell and Plano Clark, 2011).
4. Sampling is embedded. This means that there is both qualitative and quantitative data about the participants or a sub-set of the participants.
5. Comparable resources and expertise are invested to all of the methods.
6. If it is a team-based project, there is overlap in the methodological and content-oriented expertise that allows more than one person to play a part in analyzing the data from different strands or phases of the study.

It is often the case that selecting the most appropriate terminology to describe the design of a mixed methods study cannot be fully realized until the research is well underway.

Characterizing mixed priority research

It has been the practice in mixed methods reporting to label research as reflecting one of three priorities: quantitative, qualitative, or equal priority. Iterative and integrated mixed methods designs are often characterized as reflecting equal priority because neither method is subordinated or treated in a perfunctory way. The term “equal priority” is not entirely satisfactory in situations where the conclusions, like what might be presented in a composite theoretical model, communicates a situation where the outcomes of a project cannot be directly traced to a single method.

The expression mixed methods priority is a way to communicate that findings from each method contributed in a meaningful way to generating a cohesive and coherent explanation (Creamer, 2018). Mixed priority communicates that key findings emerged from a dialectical interplay between data and methods. It provides a shorthand to communicate that the results of a project were synergistic.

Concluding thoughts

Mixed methods is a dynamic and a rapidly evolving methodological approach that is embraced by diverse disciplinary communities and implemented in settings across the world. It is characterized by the remarkable variety of the ways it has been adapted to understand complex phenomena in diverse, multi-disciplinary settings. The topic of integration is at the center of contemporary scholarship about mixed methods where there is growing attention to its potential to be embedded at many points in the research process (Creamer, 2018; Fetter and Molina Azorin, 2019), including by mixing methods, perspectives, and theoretical and paradigmatic orientations.

The type of mixing used in a mixed method study and whether it occurs in one or more places during the research process has implications for how research methods and findings are reported and for how teams conduct their work. It is recommended that manuscripts include a section devoted to integration in the portions of a manuscript that describe the research procedures and the results. It is helpful, as well, to specify when mixing occurred in a flowchart or procedural diagram designed to documents how data collection and analysis occurred.

Research conducted with a distinct qualitative phase and a distinct quantitative one, accommodates settings where collaborators have infrequent interaction, work relatively independently because there is little overlap in expertise, and have responsibility for deliverables at different points in time. The more interactive and iterative approaches I have highlighted accommodate an applied purpose and are embedded in real-world agencies, like schools, hospitals, or community agencies that are in a constant state of flux. Some researchers turn to mixed methods when findings from different sources of data do not cohere in expected ways. An interactive approach is suited to settings where there is an opportunity for frequent and on-going interaction among collaborators. This is likely to be where there is enough overlap in both methodological and topic-related expertise, that multiple investigators contribute to the analysis. There is a high level of involvement with participants across many phases of the research project in participatory action research (Ivankova, 2015). It is not possible for a single approach to mixed methods to be suitable to all research questions and conditions.

References

- Archibald, M.M., Radil, A.I., Zhang, X., Hanson, W.E., 2015. Current mixed methods practices in qualitative research: a content analysis of leading journals. *Int. J. Qual. Methods* 14 (2), 1–33. <https://doi.org/10.1177/160940691501400205>.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage, Thousand Oaks, CA.
- Bussing, R., Koro-Ljungberg, M.K., Noguchi, K., Mason, D., Mayerson, G., Garvan, C.W., 2012. Willingness to use ADHD treatments: a mixed methods study of perceptions by adolescents, parents, health professionals, and teachers. *Soc. Sci. Med.* 74, 92–100. <https://doi.org/10.1016/j.socscimed.2011.10.009>.
- Creamer, E.G., Guetterman, T., Govia, I., Fetter, M., 2020. Challenging procedures used in systematic reviews by promoting a case-based approach to the analysis of qualitative methods in nursing trials. *Nurs. Inq.* <https://doi.org/10.1111/nin.12393>.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. Sage, Thousand Oaks, CA.
- Creamer, E.G., 2021. *Advancing Grounded Theory With Mixed Methods*. Routledge, London, UK.
- Creswell, J.W., Plano, C.V., 2011. *Designing and Conducting Mixed Methods Research*. Sage.
- Denzin, N.K., 2010. Moments, mixed methods, and paradigm dialogues. *Qual. Inq.* 16 (6), 419–427. <https://doi.org/10.1177/1077800410364608>.
- Fetter, M.D., Guetterman, T.C., 2021. Development of a joint display as mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Sage, pp. 259–276.
- Fetter, M.D., Molina-Azorin, J., 2017. The *Journal of Mixed Methods Research* starts a new decade: the mixed methods research integration trilogy and its dimensions. *J. Mix. Methods Res.* 11 (3), 291–307.
- Fetter, M.D., Molina-Azorin, J., 2019. Rebuttal-conceptualizing integration during both data collection and data interpretation phases: a response to David Morgan. *J. Mix. Methods Res.* 13 (1), 12–14. <https://doi.org/10.1177/1558689818780596a>.
- Fielding, N.G., 2012. Triangulation and mixed method designs: data integration with new research technologies. *J. Mix. Methods Res.* 6 (2), 124–136.

- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11 (3), 255–274. <https://doi.org/10.3102/01623737011003255>.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Wiley Publishers, San Francisco, CA.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods research merger. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Sage, pp. 606–615.
- Howe, K.R., 1988. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educ. Res.* 10–16. <https://doi.org/10.3102/0013189X017008010>.
- Ivankova, N., 2015. *Mixed Methods Applications in Action Research: From Methods to Community*. SAGE Publications, Thousand Oaks, CA.
- Johnson, R.B., Schoonenboom, J., 2015. Adding qualitative and mixed methods research to health intervention studies: interacting with differences. *Qual. Health Res.* 26 (5), 587–602. <https://doi.org/10.1177/1049732315617479>.
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1 (2), 112–133.
- Johnson, R.B., 2012. Guest editor's editorial: dialectical pluralism: a metaparadigm whose time has come. *Am. Behav. Sci.* 56, 751–754.
- Maxwell, J., Chmiel, M., Rogers, S.E., 2015. Designing integration in multimethod and mixed methods research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Sage, Oxford, UK, pp. 223–239.
- Morse, J.M., 2010. Simultaneous and sequential qualitative mixed method designs. *Qual. Inq.* 16 (6), 483–491. <https://doi.org/10.1177/1077800410364741>.
- O'Cathain, A., Murphy, E., Nicholl, J., 2007. Integration of publications as indicators of “yield” from mixed methods studies. *J. Mix. Methods Res.* 1 (2), 147–163.
- Plano Clark, V.L., Sanders, K., 2015. The use of visual displays in mixed methods research: strategies for effectively integrating quantitative and qualitative components of a study. In: McCrudden, M.T., Schraw, G., Buckendahl, C. (Eds.), *Use of visual displays in research and testing: coding, interpreting, and reporting data*. Information Age, Charlotte, NC, pp. 177–206.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. In: *Mixed Methods Research Series*, vol. 3. Sage.

Cultural studies in education: insights for the mixed methods researcher

Loraine D. Cook^a, Vimala Judy Kamalodeen^b, Ingrid E. Hunt-Anderson^a, Tashane Haynes-Brown^a, Leemoy Weaver^c, Steve R. Weaver^d, and Lamoine Samuel Lee^a, ^aSchool of Education, University of the West Indies, Mona, Jamaica; ^bSchool of Education, University of the West Indies, St. Augustine, Trinidad and Tobago; ^cSociology, Psychology and Social Work, University of the West Indies, Mona, Jamaica; and ^dThe UWI School of Nursing, Mona, Jamaica

© 2023 Elsevier Ltd. All rights reserved.

Overview of cultural studies in education	417
Cultural studies in education and mixed methods research	418
Philosophical foundations for mixed methods cultural studies in education	418
Mixed methods research and integration	419
Illustrative examples of cultural studies in education and MMR	419
A multicultural study in education using MMR	419
An intercultural study in education using MMR	420
A cross-cultural study in education using MMR	421
Relevance of paradigms in promoting cultural diversity and inclusion	422
The purposeful integration of quantitative and qualitative approaches in MMR	423
Conclusion	423
References	424

Overview of cultural studies in education

Cultural studies in education, considered an institutionalized discipline, is a plural field of enquiry that “draws from whatever fields are necessary to produce knowledge required for a particular project” (Grossberg et al., 1992, p. 2). The evolution of cultural studies in education started when British cultural studies first made its entry into the discipline of education in the United States during the late 1970s–1980s (Giroux and Shannon, 2013). As noted by Hytten (1999), educators in the discipline of cultural studies focus on two main areas, the traditions of pedagogy guided by critical theory (“critical theory challenges the status quo and strive for a balance and democratic society”, Asghar, 2013, p. 3123) and ethnic and diversity studies. Other dimensions of cultural studies in education include multiculturalism, interculturalism and cross-culturalism. Multi-cultural education involves recognizing and respecting diverse cultures and identities within the school system so that education can be more equitable, harmonious and fair for all (Gay, 1997; Bennett, 2015) whereas intercultural education emphasizes interpersonal relationships among different cultures (Chiriac and Panciuc, 2015). In other words, there is mutual exchange or interactions of ideas and cultural norms that result in deeper relationships in a multicultural society. Crosscultural studies involve persons from different ethnic groups and countries. Crosscultural studies in education allow for understanding human behavior through the comparison of different perspectives to enable the expansion of knowledge of a dominant culture (Papayiannis and Anastassiou–Hadjicharalambous, 2011).

Additionally, research into multiculturalism, interculturalism and cross cultural education can be guided by different paradigmatic stances, such as pragmatism, transformative and critical theory. Paradigmatic stances “are systems of beliefs and practices that influence how researchers select both the questions they study and methods they use to study them” (Morgan, 2007, p. 9). Pragmatism ascribes to multiple realities which enables research in cultural studies in education to be richer and comprehensive because of an interchange of different views on life, which may also productively stimulate new ways of thinking while conserving valuable aspects of tradition (Shannon–Baker, 2016; Shusterman, 2010). Critical educators challenge the dominant culture “to rethink notions of difference, privilege, and power. They have challenged the implicit social norms of those with power and called for us to uncover (so as to alter) the operations of the culture of power” (Hytten, 1999, p. 531). Transformative paradigm sensitizes researchers in education in both methodological and philosophical ways that recognize the construction of multiple realities. These realities are imbued by diverse socio-cultural impositions and human rights issues in those realities throughout the research process (Hunt–Anderson, 2021). The distinguishing feature of critical theory from transformative is that critical theory increases awareness of the differentials within power structures, while transformative theory leads to a change in beliefs and perspectives that ultimately change behaviors and activities of individuals (Wang et al., 2019).

Due to the variety of paradigms that are available to guide the cultural studies researcher across diverse contexts, we agree with Nelson et al. (1992) that there is no distinct methodology or set guidelines in how or what questions and answers need to be explored in educational contexts. A cultural study in education researcher may find that mixed methods research provides a range of philosophical and methodological approaches that are useful, this is because the approach can be framed from traditional to more critical lenses. Research in cultural studies in education can be richer in understanding and improving relationships among students of different cultures within education institutions, which may also productively stimulate new ways of thinking and relating while conserving core attributes of the multiple cultures represented in an educational context.

Mixed methods research (MMR) represents an approach with the potential to explore cultural issues or phenomena such as the nexus of context and education, capturing the uniqueness of educational issues within cultures (Bartholomew and Brown, 2012). Cultural phenomena are unpredictable and are characterized by combinations of multiple, and non-linear patterns and directions (Poth, 2018). While acknowledging that the definition of MMR has evolved over the years, Creswell and Plano Clark (2018) emphasize that MMR involves a process (not necessarily in the same order) of using specific mixed methods designs for a study, developing a philosophical and theoretical framework for the study, collecting and analyzing both quantitative and qualitative forms of data, and integrating these data forms and their results. The main purpose of integrating is to enhance understanding of a phenomenon or issue under investigation, and to this end, there is a growing interest among researchers in using MMR approaches in cultural studies in education.

In this chapter, we offer insights on how MMR designs are utilized in different domains of cultural studies in education—multicultural, intercultural, and cross-cultural, which enhanced greater understanding of the phenomenon being studied and the possibility of achieving positive socio-cultural change in the educational context. More specifically we discuss how integration of qualitative and quantitative approaches in a single study facilitates and enhances the research process in understanding issues in education within cultural contexts. In the chapter, we also examine the role of pragmatism as a fundamental MMR philosophical approach. The chapter also includes examples to provide insights of how researchers utilized and benefitted from using MMR designs in three domains of cultural studies in education (multicultural, intercultural, and cross cultural).

Cultural studies in education and mixed methods research

Discussions specific to methods in cultural studies in education have been largely under-described (Pickering, 2008). As such, methods that place emphases on understandings of the issues of power are relevant and appropriate for cultural studies researchers. Cultural studies in education are leaning toward more empirical forms of research as researchers seek evidence to analyze and interpret cultural theories and frameworks, pedagogies, and praxis (Pickering, 2008). There is a need to expand the debate in education where the field of cultural studies offer an opportunity to build on existing research designs and models by examining texts, stories, images and how one comes to understand the world they live in. Such understandings are facilitated by our theoretical underpinnings. For example, critical theoretical frameworks play a pivotal role in understanding cultural studies in education to uncover and discover issues that often challenge existing norms (Trifonas, 2020).

Several approaches and methods can be adopted in cultural studies in education for inquiry such as case study (Guetterman and Mitchell, 2016) or participatory action research (Goldberg, 2020) that utilizes deductive or inductive approaches or a combination of both, which is termed abductive. According to Mitchell (2018), abduction enables rigorous search of the data, testing ideas, checking and trying to understand differences in interview responses under various circumstances, in short, what occurs in practice and the reasons why. Although cultural studies in education has traditionally followed the mode of enquiry of an anthropologist seeking to develop an in-depth understanding of a culture through participation, mixed methods research allows a researcher to use abductive logic in their research and to work toward the “radical middle” (Onwuegbuzie, 2012).

Philosophical foundations for mixed methods cultural studies in education

Mixed methods researchers are encouraged to be cognizant of the philosophical assumptions that guide their work (Creamer, 2018a; Creswell and Plano Clark, 2018; Johnson and Onwuegbuzie, 2004; Shannon-Baker, 2016). The need for researchers to make their philosophical assumptions explicit is described by Creamer (2018a) as philosophical transparency, that is, when the authors discuss their work’s philosophical framework in details. A paradigm is often described as the researcher’s philosophical assumptions or worldview about the nature of reality and the researcher’s role in constructing what is agreed upon by a community of scholars (Creamer, 2018a,b). As mentioned earlier, pragmatism is most commonly associated with MMR (Creamer, 2018a,b). Pragmatism prefers action to philosophizing and “takes an explicitly value-oriented approach to research that is derived from cultural values” (Johnson and Onwuegbuzie, 2004, p. 18).

As we reviewed the literature for scholarly articles on cultural studies in education that utilized MMR approaches, we found only one of forty-three articles that identified and discussed a research paradigm in their study. Ungar and Liebenberg (2011), identified and discussed a research paradigm to frame their study; they declared their research paradigm as transformative. Our finding which is similar to that of Alise and Teddlie (2010) who, in their investigation of six hundred research articles, found only one that mentioned an underlying philosophical research assumption. This indicates that researchers have neglected to state their worldviews, thus, according to Creamer (2018a,b) lack philosophical transparency. Regardless of the disciplinary background, mixed methods researchers should at least understand the philosophical assumptions that underpin MMR and how these assumptions can have an impact on the way the study is designed and conducted. Throughout this chapter, we emphasize the importance of encouraging mixed methods researchers to be more explicit about their paradigm in their reporting of MMR investigations. As Shannon-Baker (2016) explains, “paradigms can help frame one’s approach to a research problem and offer suggestions for how to address it given certain beliefs about the world” (p. 321). Hence, the inclusion of the research paradigm will help to shape and authenticate an MMR study.

Mixed methods research and integration

In addition to exploring the paradigmatic underpinnings of MMR, it is important to discuss its defining attribute—integration. Integration is the point in the research process where the quantitative and qualitative research approaches interface (Creswell and Plano Clark, 2018). It is this attribute that distinguishes MMR from multi-methods research. It is recommended therefore, that mixed methods researchers illustrate how and where integration occurs in the research procedures (Creswell, 2022).

Although the interfacing of quantitative and qualitative research approaches may be written into a study's title, purpose statement and research questions (Creswell and Plano Clark, 2018), there are four basic methods of linking quantitative and qualitative datasets in mixed methods research: connecting, building, merging, and embedding (Creswell and Plano Clark, 2018; Fetters et al., 2013). Connecting is linking one database to the other through sampling. Usually, quantitative data are collected and analyzed in the first phase. Then, the results from the quantitative analysis with inclusion criteria allow the researcher to select a subgroup of participants for the qualitative phase (to collect data using techniques such as interviews, focus groups, observation, etc.). The researcher's intent for the qualitative phase is to expand or elaborate on the quantitative results. Building occurs when the qualitative data are collected and analyzed in the first phase. Then, the information garnered from the qualitative is used to develop an instrument, an intervention, or identify new quantitative variables. For integration through merging, the qualitative and quantitative data are collected and analyzed separately and then, the results from both stages are linked or merged to compare and contrast, in other words, to identify convergences and divergences (Creswell, 2022).

Integration in MMR is a purposeful interdependence between the methods within the two research approaches (quantitative and qualitative) (Creamer, 2018a). Creamer makes the distinction between a component (multi-methods) design and an integrated design. In the component design, integration of findings occurs only in the conclusions, while an integrated MMR design, according to Creamer (2018a) is, "one where there is an ongoing interaction between the data and/or methods at many points in the research process" (p. 2). This is a synergy between two research approaches (quantitative and qualitative) that can take place in the conceptual and methodological frameworks, presentation of results and inferential stages/sections of the research process; these sections are "brought together into a coherent whole" (Creamer, 2020, p. 2).

Illustrative examples of cultural studies in education and MMR

Given the above discussion, we offer an examination of published research in different cultural contexts to understand the benefits of using MMR designs in cultural studies in education. Thus, we present three examples from three cultural domains: multicultural, inter-cultural, and cross-cultural in a higher educational context, that illustrate how mixed methods research has been utilized in cultural studies in education. In the first example, we review a Finnish study by Acquah and Commins' (2013) who evaluated a multi-cultural course for exchange students. The second example is inter-cultural by Tajima (2019, 2020) who conducted a study on Japanese students in a short-term study-abroad program in Canada. The third example by Ungar and Liebenberg (2011) is based in Canada where they conducted a 3-year, cross-cultural pilot study of child and youth resilience across fourteen communities within eleven countries.

A multicultural study in education using MMR

Multicultural education provides a space for dialog and understanding of diverse cultural groups and their inter- and intra-relationships. Groups with the most power learn to listen to and collaborate with those with less power, particularly regarding education issues. It allows for the interrogation of power in defining education's purposes and processes in a diverse and unequal world (Gay, 1997; Bennett, 2015).

Acquah and Commins' (2013) study examined "pre-service teachers' knowledge of issues related to multiculturalism and diversity before and after taking a multicultural education course" (p. 445). The researchers used an MMR approach but did not explicitly provide a rationale nor did they declare the type of MMR design which would have provided some indication of the rationale. It was only in the reporting of the findings that it became evident that the intent for using an MMR was to ensure reliability through triangulation and to obtain a more expanded understanding of the quantitative results through the qualitative data analysis.

There were two research questions that they used to guide their study:

- (1) How culturally aware are pre-service teachers at the point of entry a multicultural education course and at the point of completion? and (2) Do pre-service teachers believe they know cultural diversity and teach minority students after taking a multicultural education course? (p. 4).

Data were collected from thirty-eight student teachers, thirty-five females and three males, registered for a multicultural education course at an urban university in Southwestern Finland during an autumn semester of an academic year. These students were from seventeen countries both within and outside of Europe, Finland, Cameroon, Nigeria, Pakistan, Greece, Iran, Ireland, Korea, Hungary, Romania, China, Germany Austria, Spain, Serbia, Taiwan and Poland. Data were collected using a survey questionnaire, pre- and post-test journals, reflections, focus groups, and a final evaluation. Participants were guaranteed anonymity and confidentiality.

Students were asked to keep a learning journal. These facilitated students' reflection "on their learning from each reading, lecture, and seminar, how they challenged their prior beliefs or reinforced them and how the new learning was going to influence their future practice as teachers" (Acquah and Commins, 2013, p. 450). The final evaluation involved students writing comments on the most and least helpful aspects of the course. The survey instrument required students to rate themselves on the items using a five-point scale for the first three items and 3-point scale for the last item.

In the article, quantitative results were first presented. The paired sample *t*-test result showed that preservice teachers were more knowledgeable about multicultural concepts at the end of the course. This was followed by the qualitative data that expanded the quantitative results. The qualitative details gave insights concerning how the course improved the student teachers' understanding of multicultural education and assisted the researchers in identifying the various aspects of the course that resulted in an improved understanding of multicultural education. This connection between the quantitative and qualitative datasets was not presented graphically. However, the researchers discussed the quantitative results and used several exemplary excerpts to substantiate the qualitative assertions concerning the improvements of the students' understanding of multicultural education. There was also some indication of a link between both datasets in their conclusions. We conclude from this article, that integration occurred at two obvious points which therefore validates the study as an MMR if we are to follow Creamer's (2018a) definition.

We also noted that these authors did not mention or discuss a research paradigm. As mentioned earlier, a research paradigm clarifies the worldview and assumptions that guide the research methods and strategies. In other words, the paradigm helps the audience to understand the underlining reasons for the methodological decisions made while conducting the study. Despite the lack of declaration, an MMR expert could infer that pragmatism was the underlining paradigm as the researcher appeared to be open minded to mixing different research approaches within the study (Shusterman, 2010). An alternate paradigm that could have guided the methodological decisions is critical theory since multiculturalism is concerned with alternate realities to that of the dominant culture a society. Critical theorists challenge the dominant culture to rethink notions of different privilege and power (Hytten, 1999, p. 531). Mixed methods researchers are being encouraged by international experts to be explicit about their research paradigm which can help to guide novice researchers in their MMR endeavors. While we must keep in mind the word count constraints of journal articles, at the very least, a paradigm(s) should be articulated concisely.

In terms of content, we felt that the information garnered from the MMR findings gave a more comprehensive understanding of the impact of multicultural education on the course participants. The qualitative findings extended the breadth of the quantitative results. For example, before engaging with the course materials, the students were required to write their definitions of key terms which would be covered in the course. In addition, the students indicated that they desired to acquire knowledge and skills to teach diverse students and broaden their understanding of multiculturalism. The participating students also wanted to increase their capacity to help children from minority groups integrate into school. At the end of the course, students responded to a survey indicating the extent to which the course goals were accomplished. Also, they were required to update their definitions of the key concepts in the course. Their definitions of the key concepts before and after the course were used to add validity to their responses on the survey and show how the students reframed their understanding of core concepts of multicultural based on the course. Through these data, the researchers were able to ascertain how the course content influenced the students' knowledge and skills. The information was also used to identify students (they were the minority) who did not benefit as the facilitators hoped. Also, the authors noted that a small number of students did not gain as much and still had "vague ideas about key terms." This was garnered from comparing both definitions of key concepts before and after the course. Therefore, this MMR study was equipped with multiple research tools that facilitated inclusiveness and unfolded different layers of students' understanding of the various concepts in multi-cultural education. As such, the combined results from the different layers provided a more comprehensive understanding of multi-cultural education which is the basis for conducting an MMR study.

An intercultural study in education using MMR

Intercultural education is the process through which "people become aware of other ways of thinking, living or behaving and how they perceive their experiences with others" (Merino and Avello, 2014, p. 285). Given racial and economic inequalities in educational outcomes, leaving no child behind must assume that intercultural sensitivity and competence are critical components to combating racism and discrimination in schools (Darling-Hammond, 2010). Intercultural competence is defined as "the ability to interact effectively with people of cultures other than one's own" (Byram, 2000, p. 297). The study by Tajima (2019, 2020) is an example of how MMR was used in exploring intercultural awareness in education.

Tajima published her study in two parts (2019 and 2020) describing her utilization of a mixed methods case study research design (MM + CSR) to investigate a study abroad (SA) program. She focused on twenty-nine Japanese students from two universities who took part in a four-week exchange program in British Columbia and Alberta, Canada. The program was organized through the university and earned the participants credits. They were females between 18 and 22 years and had limited opportunities to use English language outside the classroom. The SA program took place between semesters. The host university organized cultural activities and tours for participating students.

According to Tajima (2019), the purpose of the overall study was "to provide a critical understanding of the ... interconnection of Second Language (L2) proficiency, affective factors and language use in different SA contexts" (p. 515). She explored constructs of L2 learning, willingness to communicate (WTC), motivation, and anxiety in the quantitative scale, and intercultural awareness in the qualitative phase. Tajima (2020) cited definitions of intercultural awareness from Merino and Avello (2014) as "how people

become aware of other ways of thinking, living or behaving and how they perceive their experiences with others" (p. 515). Intercultural awareness is referred to as a "prerequisite" for intercultural communication competence.

The researcher used a mixed methods case study design (MM + CSR) which started with a quantitative first phase, followed by a qualitative second phase which involved cases that emerged from the quantitative data analysis. Integration of quantitative and qualitative components took place at the case selection and eventually in the data interpretation stage. Tajima adapted a MM + CSR design from Cook and Kamalodeen (2020). Fig. 1 shows the quantitative phase followed by the emergent 4 cases in the second phase.

Quantitative data were collected from four (4) tests administered pre-departure and post-return from the SA. These were (1) English proficiency—(Computerized Assessment System of English Communication: CASEC), Willingness to communicate, motivation and anxiety. The qualitative data consisted of pre-departure and post-return individual interviews and while-abroad focus group interviews. All interview sessions were conducted using learners' native language, Japanese, then the interview data were submitted to NVivo qualitative data analysis software. The results were translated into English for dissemination by the researcher. Tajima (2020) used NVivo for frequency counts and qualitative analyses of the cases.

Quantitative results were reported in Tajima (2019) as improvement in all constructs measured: English proficiency and affective factors, such as Willingness to Communicate, Motivation, and Language Anxiety after short-term SA. Based on English proficiency test data, four cases were generated from students' language proficiency levels and were utilized to give an understanding of how students' language proficiency influence their understanding and appreciation of the host culture. The proficiency levels were advanced (scores >1200), high-intermediate (scores between 1100 and 1200), intermediate (scores between 1000 and 1100), and low (score <999).

The study confirms that intercultural awareness was not related to English proficiency levels. The findings revealed that higher language proficiency learners did not have enhanced intercultural awareness. As such, students with lower English proficiency seemed more culturally aware. This became evident because the author used a MM + CRS design that allowed for different proficiency cases and investigating each case's typical response. The quantitative data were analyzed based on English proficiency tests and cases generated for an explanation using interviews, which helped to deepen understanding across categories of English proficiency and find explanations and nuances that can help connect second language acquisition and cultural awareness.

Although no paradigm was explicitly declared, Tajima (2019, 2020) used methods that were literature-supported and substantiated research procedures with appropriate citations and good judgment, indicative of a pragmatic worldview. As indicated earlier, pragmatism allows for a researcher in cultural studies in education to use a design they feel comfortable with and answers the research questions on intercultural competence. Language is a definable aspect of culture, and its words refer to internal meanings of its culture (Mazari and Derraz, 2015). The focus of language and perception in a foreign culture in Tajima's study (2019; 2020) shows the intricate role that language plays in cultural studies. We felt that the study would have been enhanced by an explicit declaration of the researcher's philosophical assumptions which, in turn, helps the reader to understand the mixed methods researcher's decisions about the design and integration within the study (Creamer, 2018a; Creswell and Creswell, 2018).

A cross-cultural study in education using MMR

Cross-cultural studies in education involves submerging students and/or teachers in a different culture from their own, for a comparison of cultures in order to understand variations in individual's behavior that is influenced by the cultural context (Papayiannis and Anastassiou-Hadjicharalambous, 2011; Husa, 2021). The study by Ungar and Liebenberg (2011) is an example of a cross-cultural study in education that utilized MMR approach.

Ungar and Liebenberg (2011) illustrated how the application of a transformative MMR design in the investigation of a global cross-cultural study of resilience among children and youth promoted cross-cultural sensitization and equity among diverse cultural identities and languages. The investigation also resulted in the validation of cross-cultural instrumentation and credibility

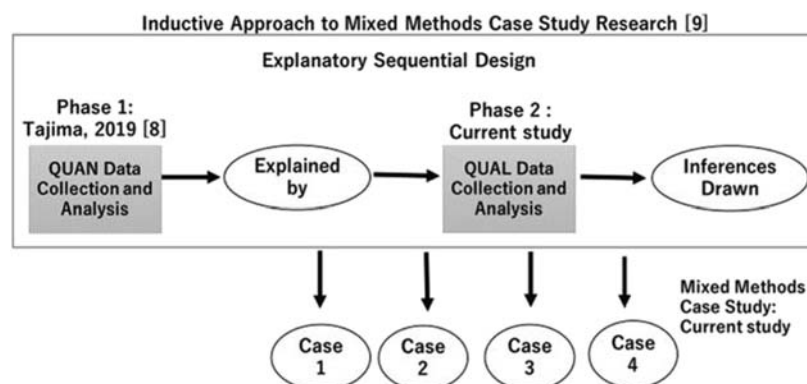


Fig. 1 Diagram showing the mixed methods case study research design (MM + CRS) of the Japanese Study Abroad program. From Tajima (2020).

of findings. The researchers utilized [Mertens's \(2003\)](#) transformative research paradigm to inform their mixed methods approach, which involved an advanced customized multiphase MMR design that involved sequential, concurrent and embedded designs ([Creswell and Plano Clark, 2018](#)). This design utilized non-linear approaches and methods ([Poth, 2018](#)) to uncover and consolidate the cross-cultural differences related to cultural identities, language and lifestyles among diverse groups of youth and their varied approaches to resilience and resourcefulness.

The researchers collaborated with a team of experts from the Resilience Research Center (RRC) in Canada, to conduct a 3-year, cross-cultural pilot study of child and youth resilience across fourteen communities within eleven countries: Canada, China, Palestine, Israel, Columbia, Russia, India, Florida, Gambia, Tanzania and South Africa. The overarching purpose of the study was to develop a universal Child and Youth Resilience Measure (CYRM-28) that would facilitate inclusion of both minority and majority populations. The researchers set out the following objectives: First, they aimed to create an instrument for the cross-cultural study of resilience that "could account for the psychosocial resources available to youth globally, making cross-cultural comparison of developmental outcomes associated with resilience possible" (p. 127). Second, to discern which internal and external assets most influence successful developmental outcomes across all the cultural groups included in the study. Third, to identify the "elements of a mixed methodology that was effective in the development of culturally sensitive psychological measures that avoided the exporting of concepts from Minority to Majority World contexts" ([Ungar and Liebenberg, 2011](#), p. 127). Their methods were guided by [Onwuegbuzie et al.'s \(2010\)](#) 10-step process for instrument development and construct validation. The aim was "to build a more culturally sensitive measure with face and item validity" (p. 129).

The researchers initiated the study by collecting qualitative data, across multiple sites through consultations with team members, feedback from local advisory committees, focus group discussions with youth and adults in each community, and informal conversations and assessments of youth. The qualitative data were used to develop an initial survey instrument (CYRM-58) for the pilot survey. This is an example of integration with the purpose of "building" ([Creswell and Plano Clark, 2018](#); [Fetters et al., 2013](#)). Subsequent individual interviews were conducted with eighty-nine youth (thirty-two boys, fifty-seven girls) across sites. These youth were purposely selected based on their ability to cope in the face of adversity (e.g., going to school, avoiding drugs and teen pregnancy, attachment to the caregiver).

The qualitative data were analyzed to derive thirty-two domains and four clusters (individual, relational, community and cultural) which were used to develop the CYRM-58. They used factor analysis to analyzed quantitative data which sought to "identify the structure of youth interpretations of the CYRM" (p. 135) which reduce the initial pilot instrument from fifty-eight to twenty-eight items that best represented the construct of resilience across all sites. The qualitative data were also utilized iteratively to investigate the validity of the instrument using students' voices for interpretation/explanation of the quantitative data.

Ungar and Liebenberg reported that the study's findings supported their hypothesis that "Global aspects of resilience (32 domains) can be identified through culturally diverse populations of youth that show unique patterns in how resilience is understood and manifested" (p. 142), and that the results underscore the hypothesized definition of resilience described in the study. The researchers also theorized that resilience is an interplay of individual and context that are reflected by the seven tensions associated with youth outcomes and resilience. These seven "tensions" include: access to material resources, relationships, identity, power and control, cultural adherence, social justice, and cohesion. In response to their third goal, Ungar and Liebenberg posited that through the application of a MMR design supported by the transformative paradigm, they were able to develop an informed, culturally sensitive yet holistic measure of resilience (CYRM-28).

We felt that this study was an appropriate example of the use of MMR in cultural studies in education. The researchers demonstrated integration of both datasets throughout the article, highlighted their research paradigm, stated their MMR goal, and reported using data collection methods that were guided by MMR expert researchers. These help the reader to get a more comprehensive understanding of their research findings. Also important in MMR studies is a rationale for utilizing such a method over others which these researchers provided. They stated:

The use of mixed methods also allowed us to compare the results of our quantitative findings with young people's descriptions of their experiences of complex interactions to nurture and maintain well-being within their challenging social ecologies (p. 128).

They also argued that using mixed method procedures helped to ensure validity of their instrument which would not have been gained had they used a single method.

Relevance of paradigms in promoting cultural diversity and inclusion

Insights gained from the use of MMR in the three selected illustrations in cultural studies in education are now presented relating to the role of paradigms and integration within each unique design. Of the three examples presented, only one declared a paradigmatic position. Ungar and Liebenberg utilized the transformative paradigmatic lens in addressing issues of power and social justice and promoting equity and "voice" across diverse cultural groups. Ungar and Liebenberg reported that this paradigm underlined the development of a culturally sensitized instrument that allowed for inclusion of cultural voices across both minority and majority world contexts and thus reducing cultural bias in the instrument for data collection. While mixed methods researchers such as [Creswell and Plano Clark \(2018\)](#) indicate that pragmatism is most commonly associated with MMR, other paradigms can frame cultural

studies in education to add value and guide the way the study may be designed and conducted. Consistent with the findings of [Alise and Teddle \(2010\)](#) and [Cook and Kamalodeen \(2020\)](#), researchers do not usually declare their paradigm. It is important for researchers to declare their paradigm for the audience as it explicates philosophical assumptions that underpin the methodological decisions of the study.

The purposeful integration of quantitative and qualitative approaches in MMR

Integration is a critical attribute of mixed methods research that distinguishes it from other research approaches. We used descriptions of integration by [Fetters et al. \(2013\)](#) as connecting, building, merging and embedding. Cultural studies in education benefits from the strength of integrating quantitative and qualitative research approaches, thus enabling understanding of the nuances of cultural constructs and contexts. Depending on the MMR design selected in each investigation, integration was enacted at various points in our three examples.

[Acquah and Commins \(2013\)](#) collected data using multiple methods (survey, pre-test and post-test, journals, reflections, final evaluations and focus groups). Integration occurred only in the reporting of the findings. It was clear from the study that the intent for collecting multiple forms of data was to ensure reliability through triangulation and to obtain an elaboration of the quantitative results using the qualitative analysis of data. The authors were therefore able to obtain a comprehensive understanding of the impact of the teaching strategies used in the course on the students' grasp of the principles of multicultural education. Because the quantitative and qualitative were collected simultaneously and one of the intents of the researchers was triangulation, we can infer that the integration using [Creswell and Plano Clark \(2018\)](#) typology was through merging.

[Tajima \(2019, 2020\)](#) initiated her study by analyzing pre- and post- test data to assess the relationship between English language proficiency and the willingness to communicate in English in a Study Abroad program among Japanese students. In the first phase of the study [Tajima \(2019\)](#), four cases of English language proficiency emerged relating to levels of cultural awareness of the host country. Integration occurred through connecting quantitative and qualitative approaches through the sampling frame ([Fetters et al., 2013](#)). The generated cases allowed for a deeper understanding of the meanings of intercultural awareness as it related to communications in English language among non-native speakers in non-native lands. It is clear that the intent of the researcher was development and explanation of the study's quantitative results using a combined mixed methods and case study research approach.

[Ungar and Liebenberg \(2011\)](#) conducted a complex, transformative MMR design to develop a culturally-sensitive global instrument of youth resilience (CRYM-28). To achieve that, the researchers enacted multiple phases of the study that involved the persistent collection, analyses and integration of qualitative and quantitative forms of data. The quantitative instrument was first developed from qualitative data gathered from fourteen sites across the varied participating countries. In addition, qualitative data were used to identify the youth samples for the quantitative pilot study. Following the outcomes from the pilot study, further interviews with the youth were conducted to improve and reduce the instrument. Here, integration was through building. The researchers further validated the instrument by comparing the qualitative construct of resilience with the quantitative data findings. Finally, regarding the interpretation of the quantitative data findings, the researchers then ascertained convergence with the qualitative data from another round of youth interviews. Hence, there was constant interfacing that resulted in intentional integration through building a refined and culturally sensitive universal measure of youth resilience. This study therefore provides insights into the value that MMR integration processes contribute to addressing cultural nuances and consolidating commonalities within cultural studies in education. It also illustrates the level of design flexibility that is afforded the MM researcher to adequately adapt to varied cultural contexts and satisfy the goal(s) of the research. Moreover, there are implications that repetitive rotations of MMR integration in cultural studies will lead to greater clarification of cultural concerns and accommodation of nuances.

Integration in MMR is a purposeful interdependence between the methods within the two research approaches (quantitative and qualitative). The discussions on utilizing paradigms and integration highlight their importance in designing and conducting MMR in cultural studies in education. The illustrations provide evidence that MMR is adaptable to diverse cultural contexts and allow core principles of MMR, once applied, to unravel various layers of the research problems or issues. MMR allows for flexibility in design and methods thus empowering the MM researcher to make optimum research decisions in their "pursuit of truth."

Conclusion

In this chapter, we discussed the use of the MMR approach to investigating issues or problems in cultural studies in education. Cultural studies as a discipline is concerned with the unique and shared ways of living among different communities and contexts. Like integrating multiple methods in MMR, cultural studies facilitate the integration of the various components (beliefs, values, and behaviors) in culture to understand individuals or communities. The integration of data sources facilitates "a more coherent and comprehensive view of what is shared across groups and unique to each other" ([Schrauf, 2018](#) p. 478). MMR facilitates the interrogation of the variances across and within cultures by integrating qualitative and quantitative methods. We shared selected examples to illustrate MMR's utilization in three distinct domains of cultural studies: multicultural, intercultural, and cross-cultural. The open-mindedness that resulted from understanding how culture and ethnicity affect our own beliefs, values, and behavior can lead educators to foster greater respect among cultures in the lecture hall of higher education. The emerging consensus from this paper is that "on the whole, culture can be best explored by employing mixed methods" ([Panda and Gupta, 2013](#) p. 293) because MMR

allows for multiple approaches and fluidity in methodology for examining complex cultural phenomena. Doing culturally relevant MMR, especially underlined by the utilization of appropriate philosophical paradigms in cultural studies in education, can influence the development of culturally sensitive measures and appropriate research methods, leading to socially just and equitable practice.

References

- Acquah, E., Commings, N., 2013. Pre-service teachers' beliefs and knowledge about multiculturalism. *Eur. J. Teach. Educ.* 36 (4), 445–463. <https://doi.org/10.1080/02619768.2013.787593>. Available from. (Accessed 27 July 2022).
- Alise, M.A., Teddlie, C., 2010. A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *J. Mix. Methods Res.* 4 (2), 103–126. <https://doi.org/10.1177/1558689809360805>. Available from. (Accessed 27 July 2022).
- Asghar, J., 2013. Critical paradigm: a preamble for novice researchers. *Life Sci. J.* 10 (4), 3121–3127 (ISSN:1097-8135). <http://www.lifesciencesite.com.41>.
- Bartholomew, T.T., Brown, J.R., 2012. Mixed methods, culture, and psychology: a review of mixed methods in culture-specific psychological research. *Int. Perspect. Psychol.* 1 (3), 177–190. <https://doi.org/10.1037/a0029219>. Available from. (Accessed 27 July 2022).
- Bennett, J.M., 2015. The Sage Encyclopedia of Intercultural Competence. In: Sage Publications.
- Byram, M., 2000. Routledge Encyclopaedia of Language Teaching and Learning. Routledge. <https://doi.org/10.4324/9780203219300>. Available from. (Accessed 27 July 2022).
- Chiriac, A., Panciuc, L., 2015. Intercultural education- objectives, values and perspectives. In: International Conference New Perspectives on Science Education. Available from. <https://conference.pixelonline.net/NPSE/files/npse/ed0004/FP/1426-NTST861-FPNPSE4.pdf#:~:text=Intercultural%20education%20refers%20to%20the%20environment%20%22>. (Accessed 27 July 2022).
- Cook, L.D., Kamalodeen, V.J., 2020. Combining mixed methods and case study research (MM+CSR) to give mixed methods case study designs. *Caribb. J. Mix. Methods Res.* 1 (1), 47–76. <https://doi.org/10.37234/CJMMR>. Available from. (Accessed 27 July 2022).
- Creamer, E.G., 2018a. An Introduction to Fully Integrated Mixed Methods Research. Sage Publications, Thousand Oaks.
- Creamer, E.G., 2018b. Enlarging the conceptualization of mixed method approaches to grounded theory with intervention research. *Am. Behav. Sci.* 62 (7), 919–934. <https://doi.org/10.1177/0002764218772642>. Available from. (Accessed 27 July 2022).
- Creamer, E.G., 2020. Why Integration? Why Now? *Caribbean J. Mix. Methods Res.* 1 (1), 1–15.
- Creswell, J.W., 2022. A Concise Introduction to Mixed Methods Research, second ed. Sage Publications, Thousand Oak CA.
- Creswell, J.W., Creswell, J.D., 2018. Qualitative, Quantitative and Mixed Methods Approaches, fifth ed. SAGE, Thousand Oaks CA.
- Creswell, J.W., Plano Clark, V.L., 2018. Designing and Conducting Mixed Methods Research, third ed. SAGE, Thousand Oaks CA.
- Darling-Hammond, L., 2010. Teacher education and the American future. *J. Teach. Educ.* 61 (1–2), 35–47. <https://doi.org/10.1177/0022487109348024>. Available from. (Accessed 27 July 2022).
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs: principles and practices. *Health Serv. Res.* 48 (6 Pt 2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>. Available from. (Accessed 27 July 2022).
- Gay, G., 1997. The relationship between multicultural and democratic education. *Soc. Stud.* 88 (1), 5–11. <https://doi.org/10.1080/00377999709603738>. Available from. (Accessed 27 July 2022).
- Giroux, H.A., Shannon, P., 2013. Education and Cultural Studies: Towards a Performative Practice. Routledge, New York.
- Goldberg, D.A., 2020. Resistance, action, and transformation in a participatory action research (PAR) project with homeless youth. In: Trifonas, P. (Ed.), *Handbook of Theory and Research in Cultural Studies and Education*. Springer International Handbooks of Education, Cham, Switzerland, pp. 313–331. https://doi.org/10.1007/978-3-030-01426-1_8-1. Available from. (Accessed 27 July 2022).
- Grossberg, L., Nelson, C., Treichler, P. (Eds.), 1992. *Cultural Studies*. Routledge.
- Guetterman, T.C., Mitchell, N., 2016. The role of leadership and culture in creating meaningful assessment: a mixed methods case study. *Innovat. High. Educ.* 41 (1), 43–57. <https://doi.org/10.1007/s10755-015-9330-y>.
- Hunt-Anderson, I.E., 2021. The cloak and dagger: a mixed study of covert bullying in Jamaican high schools. *Caribb. J. Mix. Methods Res.* 2 (1), 54–97. <https://doi.org/10.37234/CJMMR.2021.0201.A03>.
- Husa, J., 2021. Comparative law, literature and imagination: transplanting law into works of fiction. *Maastricht J. Eur. Comp. Law* 28 (3), 371–389. Available from. <https://journals.sagepub.com/doi/pdf/10.1177/1023263X21995337>. (Accessed 27 July 2022).
- Hyttén, K., 1999. The promise of cultural studies of education. *Educ. Theor.* 49 (4), 527–543. <https://doi.org/10.1111/j.1741-5446.1999.00527.x>. Available from. (Accessed 27 July 2022).
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26. <https://doi.org/10.3102/0013189X033007014>. Available from. (Accessed 27 July 2022).
- Mazari, A., Derraz, N., 2015. Language and culture. *Int. J. Humanit. Cult. Stud.* 2 (2), 350–359. <https://doi.org/10.1017/9781108582889.021>.
- Merino, E., Avello, P., 2014. Contrasting intercultural awareness at home and abroad. In: Pérez-Vi, C. (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Benjamins, Amsterdam, the Netherlands, p. 285.
- Mertens, D.M., 2003. Mixed methods and the politics of human research: the transformative-emancipatory perspective. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 135–164.
- Mitchell, A.J., 2018. A review of mixed methods, pragmatism and abduction techniques. *Electron. J. Bus. Res. Methods* 16 (3), 103–116.
- Morgan, D., 2007. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *J. Mix. Methods Res.* 1 (1), 48–76. <https://doi.org/10.1177/2345678906292462>. Available from. (Accessed 27 July 2022).
- Nelson, C., Treichler, P.A., Grossberg, L., 1992. Cultural studies: an introduction. In: Grossberg, L., Nelson, C., Treichler, P.A. (Eds.), *Cultural Studies*. Routledge, New York, pp. 1–19. Available from. https://pages.mtu.edu/7Ejdsack/readings/CSReadings/Nelson-Treichler_Grossberg_Cultural_Studies_an_introduction.pdf. (Accessed 27 July 2022).
- Onwuegbuzie, A.J., 2012. Introduction: putting the MIXED back into quantitative and qualitative research in educational research and beyond: moving toward the radical middle. *Int. J. Mult. Res. Approaches* 6 (3), 192–219. <https://doi.org/10.5172/mra.2012.6.3.192>. Available from. (Accessed 27 July 2022).
- Onwuegbuzie, A.J., Bustamante, R.M., Nelson, J.A., 2010. Mixed research as a tool for developing quantitative instruments. *J. Mix. Methods Res.* 4, 56–78. <https://doi.org/10.1177/1558689809355805>. Available from. (Accessed 27 July 2022).
- Pickering, M., 2008. *Research Methods for Cultural Studies*. Edinburgh University Press.
- Panda, A., Gupta, R., 2013. Using mixed methods approach in cross-cultural studies: lessons of a research experience. *Psychol. Stud.* 58 (3), 289–307. <https://doi.org/10.1007/s12646-013-0205-y>.
- Papayiannis, S., Anastassiou-Hadjicharalambous, X., 2011. Cross-cultural studies. In: Goldstein, S., Naglieri, J.A. (Eds.), *Encyclopedia of Child Behavior and Development*. Springer, Cham, Switzerland. https://doi.org/10.1007/978-0-387-79061-9_738. Available from. (Accessed 27 July 2022).
- Poth, C.N., 2018. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. Sage Publications, London.

- Schrauf, R.W., 2018. Mixed methods designs for making cross-cultural comparisons. *J. Mix. Methods Res.* 12 (4), 477–494. <https://doi.org/10.1177/1558689817743109>. Available from. (Accessed 27 July 2022).
- Shannon-Baker, P., 2016. Making paradigms meaningful in mixed methods research. *J. Mix. Methods Res.* 10 (4), 319–334. <https://doi.org/10.1177/1558689815575861>. Available from. (Accessed 27 July 2022).
- Shusterman, R., 2010. Pragmatism and cultural politics: from rortian textualism to somaesthetics. *New Lit. Hist.* 41 (1), 69–94. <https://doi.org/10.1353/nlh.0.0146>.
- Tajima, C., 2020. A mixed methods case study: Japanese sojourners' intercultural awareness. Part of the book series *Advances in Intelligent Systems and Computing*. In: Karwowski, W., Goonetilleke, R., Xiong, S., Goossens, R., Murata, A. (Eds.), *Advances in Physical, Social & Occupational Ergonomics (Conference Session AHFE 2020)*, vol. 1215. Springer. https://doi.org/10.1007/978-3-030-51549-2_68. Available from. (Accessed 27 July 2022).
- Tajima, C., 2019. Language and cultural learning in a short-term study abroad: an investigation of Japanese sojourners in Canada. In: Goossens, R., Murata, A. (Eds.), *Advances in Social and Occupational Ergonomics*. Springer, Cham, Switzerland, pp. 572–577.
- Trifonas, P.P., 2020. *Handbook of Theory and Research in Cultural Studies and Education*, first ed. Springer International Publishing, Cham, Switzerland. <https://doi.org/10.1007/978-3-319-56988-8>. Available from. (Accessed 27 July 2022).
- Ungar, M., Liebenberg, L., 2011. Assessing resilience across cultures using mixed methods: construction of the child and youth resilience measure. *J. Mix. Methods Res.* 5 (2), 126–149. <https://doi.org/10.1177/1558689811400607>. Available from. (Accessed 27 July 2022).
- Wang, V.X., Torrisi- Steele, G., Hansman, C.A., 2019. Critical theory and transformative learning: some insights. *J. Adult Cont. Educ.* 25 (2), 234–251. <https://doi.org/10.1177/1477971419850837>. Available from. (Accessed 27 July 2022).

Mixed methods research designs

Michael D. Fetters^a and **Chihiro Tajima^{a,b}**, ^a Mixed Methods Program, Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States; and ^b Gakushuin Women's College, Shinjuku, Tokyo, Japan

© 2023 Elsevier Ltd. All rights reserved.

Introduction	426
Comprehensiveness of typologies	426
Fixed and emergent designs	427
Prioritizing strands	427
Three categories of mixed methods designs	427
Core mixed methods research designs	427
Convergent mixed methods research designs	428
Longitudinal mixed methods research designs	430
Other variations in naming	430
Sequential mixed methods research designs	430
Explanatory sequential mixed methods research designs	430
Exploratory sequential mixed methods research designs	430
Scaffolded mixed methods research designs	430
Multistage mixed methods research designs	430
Analysis-centric mixed methods research designs	431
Philosophy-centric mixed methods research designs	431
Social justice-centric mixed methods research	432
Pragmatism-centric mixed methods research	433
Combined mixed methodology research designs	433
Theory-centric mixed methods research designs	434
Mono methodology mixed methods designs	435
Mono methodology mixed data collection designs	437
Mono methodology mixed data analysis designs	437
Variations in naming conventions	437
Conclusions	437
Caveats and controversies about the use of mixed methods designs	437
References	438

Introduction

In the field of mixed methods (MM), scholars began writing about MM designs and using typologies since the modern field of MM emerged in the late 1980s (Fetters, 2016). In our own MM projects, we have found design typologies to be frameworks useful for planning and guiding the integration of qualitative and quantitative data collection, analysis, and dissemination to achieve a whole greater than the sum of the individual parts (Fetters and Freshwater, 2015). Here we begin with relevant context in considering mixed methods research (MMR) designs and typologies. We address controversy about creating design typologies, about MM designs being fixed or emergent, and about weighting qualitative or quantitative strands in MMR designs.

Mixed methods research designs combine elements of both the qualitative and quantitative orientations and require creativity and flexibility in their construction as noted by Johnson and Onwuegbuzie (2004) (Tashakkori et al., 2021, p. 114). Authors of typologies include and emphasize differing dimensions especially beyond core designs, and researchers looking for guidance about naming conventions may find the differences among the published typologies to be confusing. Various design typologies fundamentally are similar for three core designs but entail differences in foci as the study procedures and dimensions addressed become more complex.

Comprehensiveness of typologies

Tashakkori et al. (2021) opine that scholars cannot create a comprehensive list of MM designs due to the many possibilities and combinations. We are of the opinion that there are commonly used designs that can be classified and used effectively by MM researchers. We hence value having MM design typologies but acknowledge the immaturity of the field and challenges of creating a complete inventory of MM designs.

Fixed and emergent designs

A contested area in the MM design discourse concerns the question of whether designs should be fixed or emergent. A fixed MM design indicates that the researcher has chosen a design to guide the conduct of the study. In contrast, an emergent MM design indicates that the researcher follows an iterative process in a program of research where the findings after completion of a phase then inform the design of the next phase. An initial qualitative phase might lead to a subsequent phase solely of quantitative data collection, or it could lead to a subsequent phase of both qualitative and quantitative data collection. [Fetters \(2020, p. 66\)](#) argues that the “either/or” dichotomy is arbitrary. Indeed, [Alwashmi et al. \(2019\)](#) describe a hybrid on the fixed and emergent continuum that they call an iterative convergent MM design.

Evidence for the continuum between fixed and emergent comes up frequently. Mixed methods researchers that have chosen in the planning stage a specific design frequently change their data collection procedures in response to emerging findings or change in circumstances. Similarly, a researcher embracing an emergent approach will likely have some “hunch” about what subsequent data collection will encompass. At a practical level, when writing a dissertation proposal or research proposal for funding, reviewers and sponsors of research are unlikely to accept a proposal that suggests the researcher will intentionally conduct subsequent research depending on the findings. Still, sponsors of research also acknowledge the impossibility of anticipating all eventual circumstances at the time a proposal is submitted. For example, many researchers had to change course in response to the COVID-19 pandemic as social distancing measures especially precluded direct observations or interviews. For example, [Bueddefeld et al. \(2021\)](#) changed their original design in response to the COVID-19, a term [Fetters and Molina Azorin \(2021\)](#) refer to as “agile mixed methods.”

Prioritizing strands

A third thread of contentious discussion derives from differences in opinion about handling the priority or emphasis of the qualitative or quantitative strand. Some methodologists opine that in MM studies, either the qualitative or quantitative methodology receives greater priority. For example, Morse characterizes qualitatively driven MM designs as studies with a core qualitative component and an inductive theoretical drive focused on discovery or exploration ([Morse and Niehaus, 2009; Morse, 2015](#)). In the Morse conceptualization of a qualitatively driven design, the secondary quantitative component would not stand on its own ([Morse, 2015](#)). [Hesse-Biber et al. \(2015\)](#) take a less restrictive position recognizing the strength of the secondary quantitative component which lies on a spectrum and could stand independently. [Johnson and Christensen \(2014, p. 671\)](#) characterize quantitatively driven as “a mixed research design in which the quantitative perspective or way of thinking is emphasized and some qualitative data are added to the study.” They also present “equal status” designs where neither component is considered the core driver of the study. [Hesse-Biber et al. \(2015, p. 5\)](#) provided a table that contrasts differences between a qualitatively driven and quantitatively driven approach based on ontology, epistemology, as well as the types of questions, data collected, analysis, and goal. Researchers in the field of MMR generally accept that studies can fall on a spectrum as shown in [Fig. 1](#). The spectrum illustrates that studies can vary in intensity with different degrees of emphasis when qualitatively driven, equivalently driven, and quantitatively driven.

Three categories of mixed methods designs

While consensus is lacking, methodologists in the field of MMR have created design typologies that can fit into the categories of core designs and scaffolded designs. We also include a third category called mono methodology designs. [Table 1](#) uses these three categories to compare the essential variations in MM design typologies through comparison of the methodologists and sources. The table further features the core designs found in common across authors, the array of scaffolded designs with both their degrees of difference and overlap, and the two types of mono methodology designs.

Core mixed methods research designs

Conceptually, MMR methodologists largely agree on three core designs. For the purposes of this chapter, we use the naming convention of [Creswell \(2021\)](#) and [Creswell and Plano Clark \(2018\)](#). These core designs involve (1) collection and analysis of qualitative and quantitative data synchronously—a convergent mixed methods design, (2) collection and analysis of quantitative data where results are used to inform subsequent qualitative data collection and analysis—an explanatory sequential mixed methods design, and (3) collection and analysis of qualitative data where results are used to inform subsequent collection and analysis of quantitative data—an exploratory sequential mixed methods design. These core designs can be used alone to conduct mixed methods

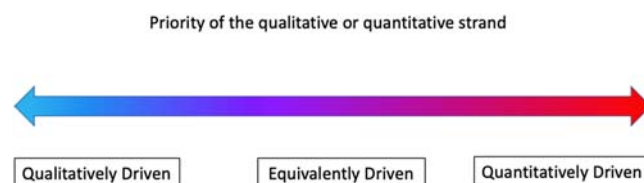


Fig. 1 Spectrum of methodological emphasis in mixed methods research designs.

Table 1 Key elements of mixed methods core designs and scaffolded designs according to selected authors.

Authors ^a	<i>Mixed methods core designs</i>		<i>Mixed methods scaffolded designs</i>				<i>Mono methodology designs</i>	
	<i>Core designs based on timing of strands</i>	<i>Multi-stage designs</i>	<i>Analysis-centric mixed methods research designs</i>	<i>Philosophy-centric mixed methods research designs</i>	<i>Combined methodology MM designs</i>	<i>Theory-centric mixed methods designs</i>	<i>Mixed data collection mono method designs</i>	<i>Mixed data analysis designs</i>
Creswell (2021)	3 core designs	Complex design	Not in typology	Complex design	Complex design	Complex design	Noted, but not in typology	Not in typology
Tashakkori et al. (2021)	Parallel designs Sequential designs	Hybrid designs	Conversion design	Not in typology	Not in typology	Not in typology	Not in typology	Not in typology
Fetters (2020)	3 core designs	Compound design and multi-stage	Not in typology	Scaffolded design	Scaffolded design	Scaffolded design	Not in typology	Not in typology
Creswell and Plano Clark (2018)	3 core designs	Complex design	Not in typology	Complex design	Complex design	Complex design	Mixed methods light	Not in typology
Plano Clark and Ivankova (2016)	3 basic designs	Intersected design	Not in typology	Intersected design	Intersected design	Intersected design	Not in typology	Not in typology
Morgan (2014)	4 basic designs	Not in typology						
Morse and Niehaus (2009)	4 simultaneous designs 4 sequential designs	Not in typology						

^aGreene (2007) did not provide a design typology.

studies or used together as “building blocks” to create designs that are even more expansive. For example, they can be integrated with other core designs, with a central philosophical approach, with theory, or with another methodology.

Table 2 illustrates the three core MM designs based on the timing of the collection of the qualitative and quantitative data strands. The table highlights current views and critical features about naming according to major MMR methodologists across multiple disciplines. The listing of the authors roughly occurs in reverse chronology according to the most recently available naming conventions. The following discussion is organized by the authors, and we describe the essential features of each design typology including key changes made by the authors in circumstances where authors changed their naming convention over time, and a critique about strengths and weaknesses for the naming convention.

While there is general agreement on these core designs, naming and typologies vary. Tashakkori et al. (2021) family of designs lumps the latter two into a single category of sequential designs. In our view, the purposes and details of sequential designs have differences of such magnitude that justifies treating them separately. Plano Clark and Ivankova (2016) use abbreviations for classifying the three types of designs, Quan + Qual for a convergent MMR design, Qual to Quan sequential research design, and a Quan to Qual sequential MMR design. Naming conventions have continued to evolve, and it would be prudent to consult the most up to date methodology books to determine naming of a research design in projects, and consider mention of historical analogues as well. In addition to the lack of consensus about naming, another source of confusion about core designs has stemmed from the same authors or groups of authors having changed their own naming conventions over time. In our understanding, this reflects an effort to optimize naming that is precise and intuitive. Ultimately, the differences illustrate evolution of methodologists’ views about elements critical for MMR design naming conventions.

Fig. 2 provides a procedural diagram for each of the three core designs. These follow the convention of Fetters (2020) where a rectangle depicts the quantitative data collection, an oval represents the qualitative strand, and diamond depicts integration of the findings of the two strands. Having portrayed the primary differences in designs and an approach for representing these designs, each design now merits further consideration.

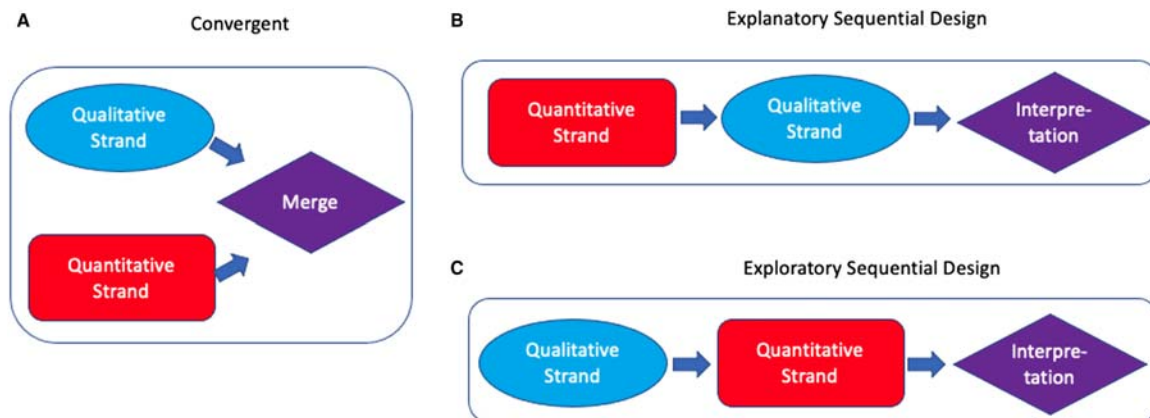
Convergent mixed methods research designs

The hallmark of a convergent design is the collection and analysis of qualitative and quantitative data synchronously, that is, roughly at the same time. While typically separation of the two forms of data collection and analysis is clear, researchers sometimes

Table 2 Core mixed methods designs typologies based on the relationship of the qualitative and quantitative strands.

	Naming convention	Qual and Quan strand relationship		
		Qual and Quan	Quan then Qual	Qual then Quan
Creswell (2021)	Core designs	Convergent	Explanatory sequential	Exploratory sequential
Tashakkori et al. (2021)	Two of the four families of mixed methods designs	Parallel designs	2 sequential	QUAL followed by QUAN
Fetters (2020)	Core designs	-QUAN and QUAL Convergent design	QUAN followed by QUAL Explanatory sequential design	QUAL followed by QUAN Exploratory sequential research design
Creswell and Plano Clark (2018)	Core designs	Convergent MMR design	Explanatory sequential design	Exploratory sequential design
Plano Clark and Ivankova (2016)	Basic mixed methods designs	Concurrent Quan + Qual	Sequential Quan → Qual	Sequential Qual → Quan
Morgan (2014)	Basic designs	Not addressed	QUAN to qual quan to QUAL	QUAL to quan qual to QUAN
Morse and Niehaus (2009)	Qualitatively and quantitatively driven designs	QUAL + quan quan + QUAL	QUAN → qual quan → QUAL	QUAL → quan qual → QUAN
Johnson and Christensen (2020)	Qualitatively/quantitatively driven from equal-emphasis or fully integrative designs ^a	QUAL + QUAN (equal status) QUAL + quan (qualitatively driven) QUAN + qual (quantitatively driven)	QUAN → QUAL (equal status) QUAN → qual(quantitatively driven) quan → QUAL (qualitatively driven)	QUAL → QUAN(equal status) QUAL → quan(qualitatively driven) qual → QUAN(quantitatively driven)

MMR = mixed methods research; Qual = qualitative; Quan = quantitative

^aFetters (2020, p. 212) refers to these as equivalently driven.**Fig. 2** Procedural diagrams for the Three Core Mixed Methods Designs. (A) Convergent. (B) Explanatory Sequential Design. (C) Exploratory Sequential Design.

collect the two strands sequentially, but there is no interfacing until both forms of data collection and analysis have been completed. This latter case would be considered synchronous as the timing relationship in practice is framed by when the two types of data interface.

The procedural diagram in Fig. 2A represents this design. Creswell refers to this as a convergent MM design (Creswell, 2021) though some authors such as Tashakkori et al. (2021) use the language of a concurrent MM design. Teddlie and Tashakkori (2009), previously referred to these as “concurrent parallel designs.” Plano Clark and Ivankova (2016) use Qual + Quan design to denote this type design. Older terminology has included triangulation design. At the time of publication, we perceive that the most up-to-date publications of empirical studies largely use convergent, or to a lesser extent, concurrent for naming. The criticism of the naming convention of convergent is that there is a possibility that the qualitative and quantitative findings may not converge. In response to this criticism, converging the findings does not mean the findings confirm each other, but rather that the two types of data are brought and considered together. In contrast, the criticism of the language of “concurrent” design comes from a nuance that the qualitative and quantitative data collection occurs independently without interfacing with each other. While neither convergent nor concurrent seems perfect as a naming convention, considering the push to emphasize the need for the strands to come together, the language of converging the findings seems most consistent. Converging further does not necessarily mean that the findings must confirm each other.

Longitudinal mixed methods research designs

A special category of convergent designs, the hallmark of longitudinal MMR designs is the collection of both qualitative and quantitative data at multiple time points. [Plano Clark et al. \(2015\)](#) suggested that to be considered longitudinal, these MMR designs should collect both types of data at more than one time point. While having received relatively little attention in most typologies, a growing number of researchers have used these designs. [Kamei et al. \(2021\)](#) utilized a fully longitudinal mixed methods case study design to examine the benefits of telenursing video consultations by examining longitudinal physical and psychological changes in a chronically ill elderly male.

Other variations in naming

As noted, in the rather recent past, the convergent design has been referred to by other names. The use of parallel mixed methods design, parallel concurrent, and parallel convergent mixed methods design, have largely fallen out of favor because of the nuance that parallel lines never join. Some researchers have referred to a triangulation mixed methods design, but this usage has been deemed less effective than the language of a convergent design given the intent of design is to compare rather than to confirm ([Creswell and Plano Clark, 2011](#)). In a more general sense, the place of triangulation in mixed methods methodology has generated considerable controversy, though details of that debate are beyond the scope this chapter's focus on designs ([Mertens and Hesse-Biber, 2012](#); [Fetters and Molina-Azorin, 2017a,b](#)).

Sequential mixed methods research designs

In contrast to the convergent design where in essence the qualitative and quantitative data are collected synchronously, in sequential designs, findings from the first strand can inform the data collection of the second strand. If a quantitative strand such as secondary dataset analysis precedes a second strand of qualitative interviews, findings from the secondary dataset analysis may help formulate questions to ask during the interviews ([Fig. 2B](#)). If a qualitative interview strand precedes the quantitative survey strand, findings from the interviews can be used to build items for the subsequent survey ([Fig. 2C](#)). As [Creswell \(2021\)](#) and [Creswell and Plano Clark \(2018\)](#) developed specific naming conventions that have become broadly recognized, these will be used in the following discussion of core sequential MM designs. As will be seen, the naming convention used by these authors can most easily be understood by considering the function of the qualitative strand.

Explanatory sequential mixed methods research designs

According to [Creswell \(2021\)](#) and [Creswell and Plano Clark \(2018\)](#) a quantitative strand followed by a qualitative strand is called an explanatory sequential mixed methods design. In this design, the initial quantitative strand is used to describe trends across a large population or used to make inferences about causality or correlations ([Fig. 2B](#)). These initial quantitative findings are then explained with a subsequent qualitative strand with a small sample of participants. [Plano Clark and Ivankova \(2016\)](#) refer to this as a Quan to Qual sequential MM design. An example of this design can be seen in a study by [Wong and Cooper \(2016\)](#). They conducted an explanatory sequential MM study by collecting quantitative data from students, followed by semi-structured interviews to explore and explain the students' online learning in Hong Kong.

Exploratory sequential mixed methods research designs

In contrast, a qualitative strand of data collection and analysis followed by a quantitative standard of data collection and analysis is called by [Creswell \(2021\)](#) and [Creswell and Plano Clark \(2018\)](#) an exploratory sequential MM design. In this design, the qualitative strand is used to explore initially a phenomenon before a second quantitative strand that is used to examine the findings across a larger population of participants ([Fig. 2C](#)). [Plano Clark and Ivankova \(2016\)](#) refer to this as a Qual to Quan sequential MM design. [Berman \(2017\)](#) conducted an exploratory sequential mixed methods study by collecting and analyzing qualitative data in the first phase, then conducting quantitative data collection and analysis in the second phase to understand data management behaviors and challenges of faculty at a university in the US.

Scaffolded mixed methods research designs

Scaffolded MMR designs denote research approaches that integrate the core designs with another core design, additional elements, or dimensions. An umbrella term accepted across MM methodologists to categorize this range of mixed methods designs that go beyond the core designs has been elusive. The lack of agreement stems at least in part to variations in the specific elements or dimensions methodologists have chosen to integrate into their design typologies. This expansive category includes multistage designs where two or more core MM designs are used sequentially, analysis-centric MM designs, philosophy-centric MM designs, combined methodology MM designs, and theory-centric MM designs. Each of these are reviewed in detail in the following.

Multistage mixed methods research designs

With multistage MM designs, researchers use two stages or more of data collection. A naming convention for this category is not found in the [Creswell \(2021\)](#) or [Creswell and Plano Clark \(2018\)](#) typologies. [Tashakkori et al. \(2021\)](#) refers to this family of designs

as hybrid designs. They use the language of hybrid to indicate combining two or more of what they call three basic families of designs (i.e., parallel MM designs, sequential MM designs, and conversion MM designs). [Fetters \(2020\)](#) and [Fetters and Kakai \(2021\)](#) use the language of compound designs to refer to combining two core designs together sequentially. The rationale for compound design is adapted from chemistry where two elements can be combined into compounds. While the resulting compound may have some similar properties to the elements from which it was made, the resulting compound will take on new characteristics or prospects not found in the individual elements. The language of multi-core designs has also surfaced in discussion with colleagues, though this language only expresses a multitude of core designs being used together while compound has the nuance of new properties.

The language of a multistage MM design represents the most generic of terms to indicate a study using multiple stages of research with various combinations of qualitative and quantitative strands. When multiple studies are used together, these may amount to mixed methods programs of research ([Crabtree et al., 2011](#)). Many multistage designs may have additional elements that can be in combination with a strongly emphasized analytic approach, philosophical stance, or theory. In such a case, we advise incorporating the language of the additional element of emphasis rather than just multistage design as this will be more specific and more comprehensible by readers.

Analysis-centric mixed methods research designs

The [Tashakkori et al. \(2021\)](#) family of designs typology includes a category of conversion MM designs. In their description of this design, they state, "In conversion MM designs, mixing occurs when one type of data is transformed and analyzed accordingly (thematic or statistical) to answer related aspects of the same research question." (p. 128). This family of designs most typically involves quantizing, where qualitative data are converted into numerical codes or categorical variables for analysis statistically with quantitative data from another strand. Hence, qualitative data converted or transformed into a quantitative or categorical form would then be combined and manipulated with other quantitative data for the analysis ([Sandelowski et al., 2009](#)). [Seltzer-Kelly et al. \(2012\)](#) utilized quantizing in a curriculum efficacy study of the diversity component of an undergraduate-level arts and humanities course. While it occurs far less frequently, qualitzing is the process of transforming quantitative data into data that can be analyzed qualitatively like a theme ([Onwuegbuzie and Leech, 2021](#)). [Weaver-Hightower \(2014\)](#) utilized a five-stage, embedded MM design for establishing influence on educational policy through an analysis of Australia's policy making on boys' education that featured both quantizing and qualitzing, within a larger qualitative study.

[Tashakkori et al. \(2021\)](#) are the only widely published methodologists who characterize an analysis-based approach to qualify as a design category. The lack of support by other methodologists of these as design categories seemingly stems from the point that data transformation is an analytic procedure that could be used in any of the other MM designs. Thus, in a study that has collected qualitative and quantitative data it would be possible to transform qualitative data into a quantitative form for analysis as quantitative data for a first analysis. Additionally, the researcher can transform the quantitative data into a qualitative form for analysis as qualitative data within the same study.

Take the case of a convergent design, or concurrent design as named by [Teddlie and Tashakkori \(2009\)](#) with roughly synchronous collection of both qualitative and quantitative data. Using the same qualitative and quantitative dataset, researchers hypothetically could conduct (1) a mixed data analysis without conversion, (2) a qualitative to quantitative data conversion for analysis quantitatively, and (3) a quantitative data to qualitative data conversion for analysis qualitatively. Under this hypothetical scenario, the authors would be using the same qualitative and quantitative datasets but with three different designs, including two different conversion designs. Rather than characterize the analytic approaches as specific designs, it would seem much more efficient and informative to refer to this as just a convergent design or concurrent design, which would be the same for all three situations, but be illustrative of three analytic strategies. This example supports the assertion that data transformation is an analysis strategy rather than a design category.

Philosophy-centric mixed methods research designs

We characterize philosophy-centric designs as approaches whereby researchers explicitly incorporate the world view throughout all dimensions of integrated qualitative and quantitative data collection, analysis, and dissemination to achieve an agenda set forth and framed by the world view at the outset of the study. Though extensive discussion falls beyond the scope of this chapter, historically, the field of MM experienced a period known as the paradigm wars. The essence of this debate could be distilled down to what many have called the incompatibility thesis, namely, that the post-positivist philosophical grounding of quantitative research differed from a constructivist philosophical grounding of qualitative research. After intensive debate and rhetoric, to date, seven major worldviews have been articulated in the literature that support integration of qualitative and quantitative strands as MMR. These seven include pragmatism, participatory/transformational, critical realism, postmodernism, yinyang, dialectical pluralism, and performative paradigm. The philosophical stance explicitly, or often implicitly, drives the choice of methodology and methods ([Greene, 2015](#)).

While philosophical assumptions or worldviews underlie all MM studies, the degree authors are explicit or implicit about this varies dramatically across disciplines. Mixed methods researchers could make their designs philosophy-centric through careful articulation of their assumptions in a variety of MMR dimensions. Still, as illustrated by [Coates \(2021\)](#), only a fraction of education researchers expressed explicitly their philosophical assumptions in published MM studies, and in many cases, these were mentioned

Table 3 Philosophy-centric mixed methods research approaches.

	<i>Design</i>	<i>Definition</i>	<i>References</i>
Social justice	Transformative	The transformative paradigm brings together many philosophical strands where social justice operates as a first principle in research endeavors.	Mertens (2003, 2009)
	CBPR	CBPR researchers consider the community as the unit of identity, build on strengths and resources of the community, seek equitable partnerships, foster <i>capacity building among all partners</i> , balance knowledge generation and intervention for the mutual benefit of all partners, leverage systems development using a cyclical and interactive process, and disseminate results to all partners in a long-term process with a commitment to sustainability.	Israel (2013)
Pragmatism	Action research	Action researchers engage in a participatory process with the intent of generating practical knowledge that can enhance human lives by integrating action, reflection, theory, and practice in their efforts to identify practical solutions for peoples' real-world problems and for their communities.	Reason and Bradbury (2008)

superficially without indicating how the assumptions influenced the study. Accordingly, the MM literature includes designs that explicitly drive the collection of MM data rather than justifying the collection of both strands of data. For the purposes of this chapter, these are referred to as philosophy-centric MMR designs. Creswell and Plano Clark (2018) include this type of design in their typology of complex designs. Plano Clark and Ivankova (2016) recognize this type of design as an intersected design. Fetter's (2020) categorizes this design under the rubric of a scaffolded design.

Table 3 illustrates two philosophy-centric approaches, namely social justice-centric designs and pragmatism-centric designs, as well as illustrative examples.

Social justice-centric mixed methods research

These studies can be characterized as approaches to inquiry that integrate qualitative and quantitative methods and are guided by a worldview that values and nurtures a commitment to collaborating with and incorporating as equals the voices of individuals from the population that is the focus of the inquiry. There is an explicit intent of promoting equity, fairness, and mitigation of discrimination. Researchers employing social justice-centric MM emphasize a particular ethic of inclusion relative to advocacy, equality, and participation of populations who are a central focus in the research as partners for determining the research undertaken, the methods to employ for the data collection, the framing of the analysis, and choices about dissemination.

Two social justice-centric MMR approaches have become widely recognized and utilized. These approaches include transformative MMR and community-based participatory research (CBPR). With both approaches, MM researchers make a commitment to including individuals from the study population into the research. Each has unique features as shown in Table 2. Rather than being specific in their data collection procedures, analysis, and dissemination, researchers using these philosophy-centric MM approaches utilize a variety of methods, qualitative and quantitative to achieve their goal of a more just society.

Transformative mixed methods research designs

Transformative MMR designs conceptually reflect approaches to the integration of qualitative and quantitative data collection, analysis, and dissemination where the researcher's choices are fundamentally driven by an explicit recognition that social injustice pervasively influences the phenomenon. With the intent of challenging social justice issues and emancipation, the research actively and intentionally chooses methods perceived as most useful for mitigating the injustice of the marginalized population of interest. Mertens (2015, p. 307) pioneered the MM transformative paradigm by contrasting the approach with post-positivism and constructivism. She emphasizes how MM procedures can empower researchers to address the complexity of issues of human rights and social transformation (2015, p. 306). Conducting a transformative MMR study involves using this transformative worldview with MM methodology. Mertens et al. (2010) stated:

The philosophical assumptions of the transformative paradigm led to consideration of approaches to research that reflect explicit recognition of values and knowledge of self and community that form a basis for methodological decisions (p. 195).

In short, the researcher has a philosophical commitment to addressing unjust sociopolitical problems and achieving social justice through sociopolitical solutions. For example, Camacho (2020) examined the workplace experiences of academic migrants, that is, international postdoctoral scholars who despite comprising a large percentage of the academic workforce experience low paying positions, limited employment stability, disparate degrees of rights and freedoms, and reputation fragility. She sought achievement of better work conditions for this population through her use of a transformative MM approach with systematic attention to criteria Mertens (2003, 2009) identified as essential for a MM transformative design.

Community-based participatory research mixed methods designs

Community-based participatory research using MMR represents a second philosophy-centric MMR design category. Like transformative MMR, CBPR using MMR signals a philosophical commitment to engaging as equals community members to achieve social justice. CBPR studies most typically focus on research about health, but also education. CBPR MMR designs refer to approaches to the integration of qualitative data collection, analysis, and dissemination where the research engages community representatives as partners throughout the entire research endeavor from identification of the research agenda, etc. While CBPR researchers frequently use a combination of qualitative and quantitative strands, CBPR studies might use solely qualitative or quantitative research. Arguably, CBPR represents a methodology as well as philosophical commitment. However, the central and explicitly driven emphasis on a social justice agenda at the community level favors categorization of this approach as a philosophy-centric approach rather than just a methodological approach being combined with MM methodology. Upadhyaya et al. (2015) conducted a CBPR project with doctoral students, community leaders, and faculty collaboration in teaching, learning, and research in all stages of the project on health research and education.

Integrating transformative mixed methods and community-based participatory research

A transformative lens can be used for the underlying philosophical grounding for a community-based participatory research study. For example, Maleku et al. (2021) used a transformative lens in their research with African refugees in the southern United States to examine the intersection of culture, financial stress, and financial self-efficacy, and then tested a culturally responsive solution to financial literacy based on community-identified priorities. In addition to scaffolding the study with the transformative lens using a CBPR approach, they employed an explanatory sequential MMR design to conduct their research in a cross-cultural context. This example illustrates how the two approaches are both compatible with each other as well as how the transformative MM component provided a lens that was implemented through use of a CBPR approach.

Pragmatism-centric mixed methods research

Pragmatism-centric MM studies represent another example of a philosophy-centric MM design. Pragmatism-centric studies integrate qualitative and quantitative methods and are guided by pragmatism (Morgan, 2014). These studies are focused on addressing practical, real-world problems, and identifying practical solutions for improving the lived reality of the population of interest. While there may be an underlying intent for achieving advocacy, equality, and participation of populations who are a central focus in the research as partners for determining the research undertaken, the researchers use of applied and practical approaches for conducting the data collection, analysis, and dissemination distinguish the approach from a social justice-centric intent.

Action research mixed methods

Action research mixed methods represents an excellent example of a pragmatism-centric research approach with increasing recognition and use. Action research is defined by Reason and Bradbury (2008) as:

... a participatory process concerned with developing practical knowing in the pursuit of worthwhile human purposes. It seeks to bring together action and reflection, theory, and practice, in participation with others, in the pursuit of practical solutions to issues of pressing concern to people, and more generally, the flourishing of individual persons and their communities (p. 4).

In the most recent edition of the SAGE Handbook of Action Research, Reason and Bradbury (2008) described action research as a “family of practices of living inquiry that aims, in a great variety of ways, to link practice and ideas in the service of human flourishing” (p. 1).

Participatory action research is one variant of many action research designs. Action researchers embrace a pragmatic approach with the intent of solving real-world problems. It has a practical focus, community-based orientation, participatory and collaborative nature, emphasis on empowerment, and values of reflection (Reason and Bradbury, 2008, p. 27/452).

Mixed methods researchers have harnessed the power of action research with CBPR and MM. For example, Garnett et al. (2019) used a transformative MM study design to examine the emancipatory potential of incorporating youth participatory action research and restorative practices implementation in a high school. In this study, the transformative MM component provided critical consciousness among students, families, staff, and administration in high schools and evaluation of restorative practices using a youth participatory action research approach (Garnett et al., 2019).

Combined mixed methodology research designs

The 2010s marked rapid uptake and advances in designs created from using mixed methodology combined with another methodology. Given the seemingly innumerable types of designs used by qualitative and quantitative researchers, this advance of combining MM methodology with another type of methodology naturally led to expansion of possibilities with MM methodology. In a special issue in the American Behavioral Scientist, Creamer and Schoonenboom (2018) proposed the language of inter-method

mixing for these designs. We have opted to use the language of combined MM methodology research designs as inter-method mixing refers to the process, and the language is not conducive to a compound noun for a name in a typology.

In the broader typologies of the authors featured in this chapter, Creswell (2021) and Creswell and Plano Clark (2018) include this combination of designs in their category of complex designs. While not defined as such by the authors, these designs would presumably fall under hybrid designs in the family of designs typology by Tashakkori et al. (2021). Fetters (2020) includes these combinations as scaffolded designs. Plano Clark and Ivankova (2016) include this combination as an intersected design. Some examples of qualitative methodologies integrated with MM include case study (Guetterman and Fetters, 2018; Bustamante, 2019), grounded theory (Shim et al., 2017, 2021), phenomenology (Martiny et al., 2021; Clark-Gordon et al., 2017), phenomenography (Roberts, 2003), narrative (Waller et al., 2021), art-based inquiry (Archibald, 2018; Archibald and Gerber, 2018), and ethnography (Kaczmarczyk and Salamońska, 2018; Berthod et al., 2017; Crede and Borrego, 2013). Some examples of quantitative methodologies integrated with MM include quasi-experimental designs and experimental designs (O'Cathain, 2018; Creamer et al., 2021).

In the process of integrating other methodologies with MM, methodologists have noted that the “added” methodology might be a component of a MM study, or that MM might be used carrying out a different methodology. Consequently, researchers have begun to divide these combined methodology designs according to (1) using the additional methodology as a component of a MM study, or (2) using MM procedures to conduct the alternative methodology. The first group involves MM using a specific qualitative methodology when conducting the qualitative component of a MM study. The second group involves MM to support the overall study. To understand this distinction, take for example the combining of case study methodology with mixed methodology. A MM study might employ the use of a case study for the qualitative component. Guetterman and Fetters (2018) denoted this as a mixed methods-case study (MM-CS). In contrast, a case study might employ the use of MM to conduct the study. Guetterman and Fetters (2018) referred to this as a case study-mixed methods (CS-MM) design.

While a broadly accepted naming convention is lacking, we suggest methodologists adopt a naming convention where both MM and the additional methodology names are found in the design name. We propose the order of naming should reflect first the component most reflective of the overarching purpose, and the second component to indicate the supportive component being used. Conceptually this could be understood by using the hyphen to signify the meaning of “using.” Hence, mixed methods-grounded theory (MM-GT) would denote MM using grounded theory for the qualitative component as illustrated in Fig. 3A, and grounded theory-mixed methods (GT-MM) would denote a grounded theory study using MM as shown in Fig. 3B. Grounded theory studies are typically qualitatively oriented. However, GT-MM incorporates both qualitative and quantitative data within an overarching grounded theory study.

Taken a step further, methodologists could be even more specific and descriptive. As noted in a Fig. 4A, a MM-GT study using a convergent MM design would be called a convergent MM-GT design. The quantitative strand and grounded theory qualitative strand findings are merged. As in Fig. 4B, a MM study using an explanatory sequential MM design could be called an explanatory sequential mixed methods-convergent design. In the first quantitative phase, quantitative data are collected and analyzed. In the qualitative second phase, the grounded theory methodology is used to build theory about quantitative results from the quantitative first phase. Finally, as in Fig. 4C, a mixed methods exploratory sequential MM design would thus be called an exploratory sequential MM-GT design. Within the qualitative first phase, theory building takes place via grounded theory methodology, while in the second quantitative phase, the theory is tested using a quantitative approach. At this point in the field of MMR, only a small number of authors have utilized some of these designs. While far from comprehensive, Table 4 provides multiple examples of combined mixed methodology research designs.

A detailed examination of all the studies in Table 4 that depicts many examples of combined methodology MM designs is beyond the scope of this chapter. That said, it is safe to say that the combined methodology examples provided are just a beginning to the variety expected to emerge.

For example, Feldon and Tolfel-Grehl (2018) introduce phenomenography MMR, where phenomenography involves framing personal conceptions as a necessary construct to related physical events experienced by people and the personal meaning of those experiences. They provide three examples of phenomenography used with the three core designs, an exploratory sequential MM design, an explanatory sequential MM design, and a convergent (concurrent) MM design. Mayoh and Onwuegbuzie (2015) provided five models for MM phenomenological research. They based the typology on purposes of the study and the priority within the study design.

Archibald and Gerber (2018) illustrated four categories of integration of arts-based research and MMR. They identified (1) communicative integration, namely, the use of arts to express MM findings; (2) data source integration where arts are used as qualitative or quantitative data integrated with the other type of data, (3) analytic integration where arts are used to enhance understanding of mixed data, and (4) conceptual integration where arts-based and MM approaches are integrated to illustrate constructs, concepts, theoretical, analytical, and methodological dimensions of a study.

Theory-centric mixed methods research designs

Researchers use theory in a variety of ways in their investigations. When theory underlies the approach of the MM study, researchers are using a theory-centric MM design. Implementation science uses many different theories to inform the data collection (Dam-schroder et al., 2009). Researchers guided by complexity theory also utilize precepts of the theory to consider the complex adaptive systems about which they conduct research. While not acknowledged in Tashakkori et al. (2021) typology, Creswell and Plano Clark

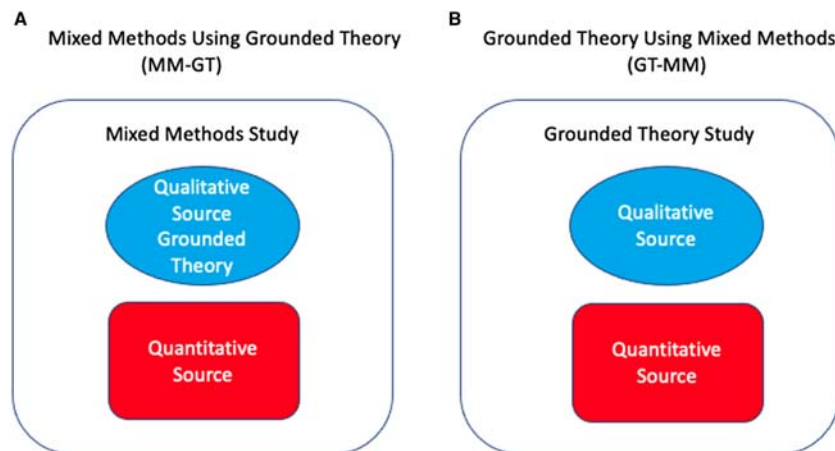


Fig. 3 Two approaches for Combining Mixed Methods and Grounded Theory. (A) Mixed Methods using Grounded Theory (MM-GT). (B) Grounded Theory using Mixed Methods (GT-MM).

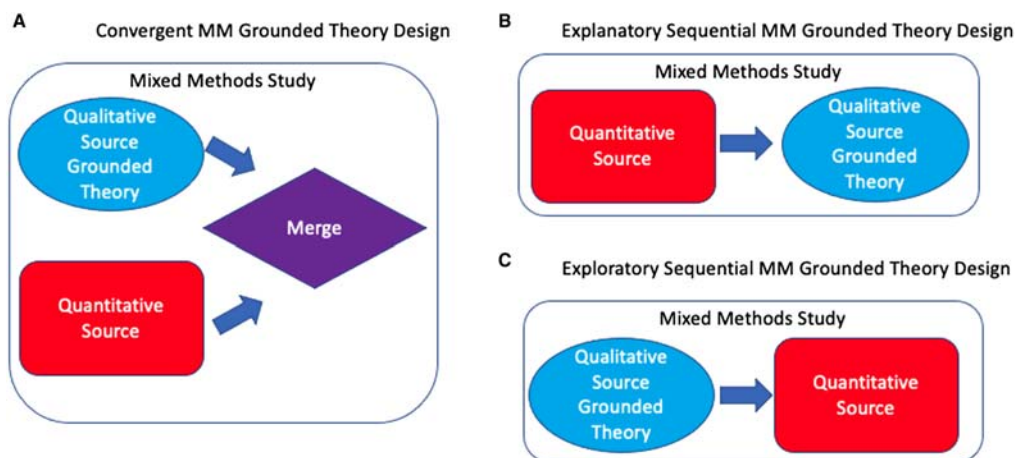


Fig. 4 Three Core Mixed Methods Designs using Grounded Theory (MM-GT). (A) Convergent MM Grounded Theory Design. (B) Explanatory Sequential MM Grounded Theory Design. (C) Exploratory Sequential MM Grounded Theory Design.

(2018) would include such designs in their category of complex designs. Plano Clark and Ivankova (2016) identify this as an intersected design while Fetters (2020) considers this a scaffolded design. In a widely cited paper, Peters et al. (2013) defined implementation research in the health sciences as "... the scientific inquiry into questions concerning implementation—the act of carrying an intention into effect, which in health research can be policies, programmes, or individual practices (collectively called interventions)."

In the implementation science field, researchers typically use theory to evaluate interventions. Chu et al. (2021) conducted a convergent MM process evaluation of a clustered-randomized controlled trial to determine the effects of a sodium reduction program in 120 rural villages in Northern China. The investigators used multiple qualitative and quantitative approaches guided by process evaluation theory to examine fidelity, delivery, reach, and receptivity of the intervention, as well as contextual factors to evaluate the clustered-randomized controlled trial.

Mono methodology mixed methods designs

As noted in Table 1, most MMR methodologists do not include mono methodology MM designs in their typologies. For the purposes of this chapter, mono methodology MM designs are defined as research strategies that utilize one approach to collection of both qualitative and quantitative data, or research strategies that collect data in one form that is analyzed using both qualitative and quantitative procedures. We accept this broader view since MM researchers commonly use these procedures to collect and analyze mixed data. Hence, these procedures use either "mixed methods" data collection, or "mixed methods" analysis. The data collection strategies and analytic strategies devised by MM researchers for integration are particularly apropos and can lead to richer interpretations.

Table 4 Designs combining mixed methods and other methodologies.

<i>Design variation</i>	<i>Description</i>	<i>Example</i>
Case Study-Mixed Methods (CS-MM)	Case study approach using mixed methods	Hall and Ryan (2011) conducted a case study that further used exploratory sequential mixed methods design where Phase 1 preliminary document reviews, interviews, and observations were used to build, then conduct a survey for a deep understanding of the survey findings.
Mixed Methods- Case Study (MM-CS)	Mixed methods study approach using case study as the qualitative strand	Onwuegbuzie and Leech (2010) conducted a mixed methods case study fully integrated throughout multiple dimensions initial quantitative assessment to examine the frequency authors of qualitative articles made generalizations in the journal <i>Qualitative Report</i> . They then conducted a subsequent qualitative assessment to assess the kinds of over-generalizations made in recommendations, and illustrating interpretive inconsistency based on the mixed methods analysis.
Grounded Theory- Mixed Methods (GT-MM)	Grounded theory approach using mixed methods	Shim et al. (2021) in Phase 1 constructed a theoretical model based on the literature through <i>meta-modeling</i> . In the second part of Phase 1, they used a classical GT model describing the mechanisms of dance movement therapy for resilience building in people with chronic pain based on in-depth interviews. We made every effort for this GT process to be independent of the initial formative model, so that the GT could freely emerge from the qualitative interview data because this is ... Shim et al. (2021) reported using an exploratory–confirmatory MM-GT design. The authors characterized their study as an exploratory–confirmatory MM-GT design.
Mixed Methods- Grounded Theory (MM-GT)	Mixed methods study using grounded theory for the qualitative strand	Howell Smith et al. (2020) utilized an exploratory sequential instrument development design study to develop and test a theoretical model. The first phase was a qualitative grounded theory to develop a theoretical model of constructs. The second, instrument development phase, developed a measurement instrument to test hypotheses related to the theory generated in the qualitative phase. The third phase was used to collect data from a larger sample. The fourth phase integrated the quantitative model-fit results with the grounded theory model to finalize an empirically tested theoretical model.
Phenomenology-mixed methods	Phenomenology using mixed methods	Martiny et al. (2021) discussed application of phenomenology-mixed methods in both experimental settings and in clinical settings. They provided three examples of applications in understanding the real-life experience of joint musical absorption by musicians, how to create and investigate positive and negative experiences of joint actions in cerebral palsy, and how to change and investigate prejudicial attitudes toward disability.
Mixed methods- phenomenology	Mixed methods study using phenomenology for the qualitative strand	Clark-Gordon et al. (2017) used an exploratory mixed-methods approach to explore college students' use of novel social media platform that was followed with a subsequent quantitative content analysis of postings to examine distribution of posting types.
Ethnography-mixed methods	Ethnographic study using mixed methods data collection	Kaczmarczyk and Salamóńska (2018) examined migration from Poland utilizing ethnosurvey methodology, an approach characterized as integration of ethnographic and survey approaches that iterates between quantitative and qualitative modes during design, data collection and analysis.
Mixed methods- ethnography	Mixed methods study using ethnography for the qualitative component	Crede and Borrego (2013) conducted an exploratory sequential mixed methods design where an initial ethnography was used as the qualitative approach and used the resulting themes to create constructs and survey questions for instrument was used to examine retention of graduate engineering students.
Narrative-mixed methods	Narrative research using mixed methods data collection	Waller et al. (2021) examined longitudinal data from the Fragile Families and Child Wellbeing study to create narrative case histories. They used a longitudinal mixed methods design to construct narrative case histories from survey data that they grounded in parents' qualitative accounts to achieve a better understanding of survey outcomes.
Mixed methods- phenomenography	Mixed methods study using phenomenology	Roberts (2003) conducted an explanatory sequential design with an initial mixed data collection survey about instructors online teaching practices followed by qualitative interviews using phenomenography.
Visual media-mixed methods	Photo elicitation MM study design	Peroff et al. (2020) utilized a photo elicitation mixed methods study design to assess residents' place attachment to the Guatemalan Maya Landscape. The authors integrated quantitative rankings of photos with qualitative interpretations of Guatemalan residents' perceptions of the meanings of the photos.

Mono methodology mixed data collection designs

Mixed data collection designs commonly used include survey research methodology and secondary dataset abstraction. Surveys commonly include a series of close-ended questions of categorical or continuous data as well as open response items including comments. The quantity and quality of qualitative data collected this way are often, but not always lacking in the true spirit of qualitative methods that seek thick and rich descriptions. Still, in many cases, the collected qualitative data prove invaluable as collecting the data allows respondents to explain choices, or raise critical issues not systematically examined by researchers in the close-ended items or questions. [Creswell and Plano Clark \(2018\)](#) have described this form of mixed data collection as “mixed methods light” as the qualitative component falls short compared to use of more intense procedures such as interviews or observations that would permit in-depth collection of qualitative data. Data abstraction methodology of secondary datasets represents another example of mixed data collection as this process involves retrieving both qualitative and quantitative data from an existing dataset.

Mono methodology mixed data analysis designs

Mixed data analysis designs involve converting quantitative data into qualitative data and/or converting qualitative data into quantitative data. Converting quantitative data refers to transforming quantitative data into narrative data to be qualitatively analyzed. This is often referred to as quantitizing ([Tashakkori et al., 2021](#)). Qualitative data can be converted into numerical codes or categorical variables to be statistically analyzed. This can be referred to as qualitizing ([Tashakkori et al., 2021](#)). The defining issue of such designs is the transformation of the originally collected data into the other format of data that is then followed by comparison of the qualitative and quantitative formats of data. For example, quantized data are compared with qualitative data. Video data, SMS text data, online forum data all serve as examples of data that are manipulated by the researcher from the original material, e.g., visual video data or audio data from videotaped encounters, into various forms of qualitative and quantitative data, e.g., transcripts or observations and time on topic or frequency counts of mentions of specific concepts, respectively.

For example, [Ruffin et al. \(2009\)](#) conducted research to understand the perspectives of individuals who were eligible for colon cancer screening, but had not been screened with the intent of creating a website chooser that would allow unscreened individuals to choose among acceptable screening approaches. The research team conducted focus group interviews, and for their analysis they coded and quantified the various responses that were elicited during the focus group interviews while also providing a series of quotes for each construct that represented the range of responses in a joint display. This mixed data analysis and comparison illustrated the most common issues vocalized as well variations in how participants expressed their opinions with representative quotations by different demographic groups. [Chang et al. \(2021\)](#) conducted mixed data analysis of SMS messages from adolescents to understand their perspectives of the evolving COVID-19 pandemic.

Variations in naming conventions

[Teddlie and Tashakkori \(2009\)](#) distinguished between mixed designs and quasi-mixed designs. According to their definition, quasi-mixed designs involve collection and analysis of both Quan and Qual data, but no mixing of the methods or the results. In our conceptualization of the two monomethodology designs, they are not quasi because there must be integration of qualitative and quantitative forms of data. Critics of categorizing the two monomethodology approaches under the mixed methodology umbrella of designs could argue that a MM study must include the collection and analysis of both qualitative and quantitative data. In the case of the monomethodology design, the concern relates to the quality of the collected data, especially the qualitative data. While conceding that the data collection is achieved through use of only one methodology, the kinds of sophisticated mixed data analyses that authors are using in studies of large qualitative datasets accentuates the relevance to MM approaches to analysis such as the use of joint displays as was the case in the research of [Ruffin et al. \(2009\)](#).

Conclusions

In this chapter we considered typologies of MM designs developed by MM methodologists across many published MM studies. In the broadest categorization, we identify core MM designs, scaffolded MM designs, and mono methodology MM designs. In the following, we synthesize contested areas about mixed methods designs.

Caveats and controversies about the use of mixed methods designs

When considering a particular design and design typology, one needs to carefully consider the specific dimensions that the author of the typology has considered. For example, in [Tashakkori et al. \(2021\)](#), the typology is organized by not only timing of different components, but also sampling and analysis. CBPR and transformative MM designs largely infer a social justice agenda where the description of designs used is in service of the philosophy. Likewise, in implementation research, theory guiding the evaluation is critical.

The scaffolded designs category builds and expands upon previous efforts to categorize diverse dimensions that we characterize by their nature of being focused on multiple stages, analysis, philosophy, methodology, and theory. Mixed methods researchers

continue to experiment with utilizing several of these different MMR designs together. For example, Ling and Pang (2021) conducted a multiphase quasi-experimental study using vignette-based MM and phenomenology in an interventional study examining financial education of early adolescent ethnic minority students. The authors used a philosophy-centric approach as a transformative lens provided the framework for their study. They further used a combined MM design as they employed through methodologies in their study, (1) a quantitative intervention, (2) a mono methodology mixed data collection survey, and (3) qualitative data collection guided by phenomenology. What remains unclear is how to name designs that incorporate several of these dimensions.

We have argued that mono methodology research qualifies as MMR as these designs call for methods for analyzing both qualitative and quantitative data. We acknowledge our stance as controversial. Plano Clark and Ivankova (2016) explains the importance of this distinction when reading published empirical mixed methods studies. Fetters and Molina (2017a) captured this notion of quasi-MM studies with the equation of $1 + 1 = 2$ where the sum of the two strands is cumulative, and contrasts with $1 + 1 = 3$ to indicate that study will produce a whole as a MM design that is greater than the sum of the individual components.

We did not conduct a systematic review to include all designs, nor a comprehensive history of design typologies. While we have provided several examples of how typologies have evolved, a full treatise on how all design typologies have evolved overtime extends beyond the scope of this chapter. Creswell and Plano Clark (2018) included information on how and why their designs and typology evolved overtime.

Despite the gains in maturity of the field of mixed methods research (Fetters and Molina-Azorin, 2021), many innovations have been and will continue to occur. We interpret changes as healthy if the thinking reflects a better conceptualization or expression. While acknowledging the importance of identifying the source of a concept or idea, forcing innovative thinking to fit with the use of the pre-existing language may be unhealthy for the field (Fetters and Molina-Azorin, 2017b). Still, once novel procedures become mapped, the field also benefits from consistency of naming conventions to facilitate comparisons.

In sum, while MM core designs are acknowledged across most methodologists, variations in naming = remain published typologies. Views diverge about MM designs categorizations beyond the core designs and are even less homogeneous than core designs. This chapter advances the literature through the organization and articulation of five types of scaffolded MM designs. These include multistage MM research designs, analysis-centric MM research designs, philosophy-centric MM research designs, combined methodology MM designs, and theory-centric MM designs. We advocate acceptance of two types mono methodology designs: mixed data collection mono method designs and mixed data analysis designs. The advancement of MM research lies in having rhetoric to describe research methods of their studies. Despite many advances in the field, opportunity abounds for further innovations in MMR design.

References

- Alwashmi, M.F., Hawboldt, J., Davis, E., Fetters, M.D., 2019. The iterative convergent design for mobile health usability testing: mixed methods approach. *JMIR Mhealth Uhealth* 7 (4), e11656. <https://doi.org/10.2196/11656>.
- Archibald, M.M., Gerber, N., 2018. Arts and mixed methods research: an innovative methodological merger. *Am. Behav. Sci.* 62 (7), 956–977. <https://doi.org/10.1177/0002764218772672>.
- Archibald, M., 2018. Integrating the arts and mixed methods research: a review and a way forward. *Int. J. Mult. Res. Approaches* 10, 1–14. <https://doi.org/10.29034/ijmra.v10n1a23>.
- Berman, E.A., 2017. An exploratory sequential mixed methods approach to understanding researchers' data management practices at UVM: integrated findings to develop research data services. *J. eScience Librariansh.* 6 (1), e1104. <https://doi.org/10.7191/jeslib.2017.1104>.
- Berthod, O., Grothe-Hammer, M., Sydow, J., 2017. Network ethnography: a mixed-method approach for the study of practices in interorganizational settings. *Organ. Res. Methods* 20 (2), 299–323. <https://doi.org/10.1177/1094428116633872>.
- Bueddefeld, J., Murphy, M., Ostrem, J., Halpenny, E., 2021. Methodological bricolage and COVID-19: an illustration from innovative, novel, and adaptive environmental behavior change research. *J. Mix. Methods Res.* 15 (3), 437–461. <https://doi.org/10.1177/15586898211019496>.
- Bustamante, C., 2019. TPACK and teachers of Spanish: development of a theory-based joint display in a mixed methods research case study. *J. Mix. Methods Res.* 13 (2), 163–178. <https://doi.org/10.1177/1558689817712119>.
- Camacho, S., 2020. From theory to practice: operationalizing transformative mixed methods with and for the studied population. *J. Mix. Methods Res.* 14 (3), 305–335. <https://doi.org/10.1177/1558689819872614>.
- Chang, T., DeJonckheere, M., Vydiswaran, V.G.V., Li, J., Buis, L.R., Guetterman, T.C., 2021. Accelerating mixed methods research with natural language processing of big text data. *J. Mix. Methods Res.* 15 (3), 398–412. <https://doi.org/10.1177/15586898211021196>.
- Chu, H., Zhang, J., Fetters, M.D., Niu, W., Li, H., Li, N., Yan, L.L., Wang, Y., Wu, Y., 2021. A mixed methods process evaluation of a clustered-randomized controlled trial to determine the effects of community-based dietary sodium reduction in rural China. *Front. Med.* 8, 646576. <https://doi.org/10.3389/fmed.2021.646576>.
- Clark-Gordon, C.V., Workman, K.E., Linvill, D.L., 2017. College students and yik yak: an exploratory mixed-methods study. *Soc. Media Soc.* 3 (2). <https://doi.org/10.1177/2056305117715696>.
- Coates, A., 2021. The prevalence of philosophical assumptions described in mixed methods research in education. *J. Mix. Methods Res.* 15 (2), 171–189. <https://doi.org/10.1177/1558689820958210>.
- Crabtree, B.F., Nutting, P.A., Miller, W.L., McDaniel, R.R., Stange, K.C., Jaen, C.R., Stewart, E., 2011. Primary care practice transformation is hard work: insights from a 15-year developmental program of research. *Med. Care* 49 (Suppl. 1), S28–S35. <https://doi.org/10.1097/MLR.0b013e3181cad65c>.
- Creamer, E.G., Schoonenboom, J., 2018. Inter-method mixing as a gateway to methodological innovation. *Am. Behav. Sci.* 62 (7), 879–886. <https://doi.org/10.1177/0002764218756917>.
- Creamer, E.G., Guetterman, T.C., Govia, I., Fetters, M.D., 2021. Challenging procedures used in systematic reviews by promoting a case-based approach to the analysis of qualitative methods in nursing trials. *Nurs. Inq.* 28, e12393. <https://doi.org/10.1111/nin.12393>.
- Crede, E., Borrego, M., 2013. From ethnography to items: a mixed methods approach to developing a survey to examine graduate engineering student retention. *J. Mix. Methods Res.* 7 (1), 62–80. <https://doi.org/10.1177/1558689812451792>.
- Creswell, J.W., Plano Clark, V.L., 2011. Analyzing and interpreting data in mixed methods research. In: *Designing and Conducting Mixed Methods Research*, second ed. Sage, Thousand Oaks, CA, pp. 203–250.

- Creswell, J., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., 2021. *A Concise Introduction to Mixed Methods Research*, second ed. Sage Publishing, Thousand Oaks, CA.
- Damschroder, L.J., Aron, D.C., Keith, R.E., et al., 2009. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement. Sci.* 4 (50). <https://doi.org/10.1186/1748-5908-4-50>.
- Feldon, D.F., Tofel-Grehl, C., 2018. Phenomenography as a foundation for mixed models research. *Am. Behav. Sci.* 62 (7), 887–899. <https://doi.org/10.1177/0002764218772640>.
- Fetters, M.D., Freshwater, D., 2015. The 1 + 1 = 3 integration challenge. *J. Mix. Methods Res.* 9 (2), 115–117. <https://doi.org/10.1177/1558689815581222>.
- Fetters, M.D., Kakai, H., 2021. *A Guide to Mixed Methods Research: A Treasure Hunt for Learning Research Design to Writing for Publication. 混合研究法の手引き——トレジャーハントで学ぶ研究デザインから論文の書き方まで*. Tomi Shobo, Mitaka, Tokyo, Japan, ISBN 978-4866161204.
- Fetters, M.D., Molina-Azorin, J.F., 2017a. The Journal of Mixed Methods Research starts a new decade: the mixed methods research integration trilogy and its dimensions. *J. Mix. Methods Res.* 11 (3), 291–307. <https://doi.org/10.1177/1558689817714066>.
- Fetters, M.D., Molina-Azorin, J.F., 2017b. The Journal of Mixed Methods Research starts a new decade: principles for bringing in the new and divesting of the old language of the field. *J. Mix. Methods Res.* 11 (1), 3–10. <https://doi.org/10.1177/1558689816682092>.
- Fetters, M.D., Molina-Azorin, J.F., 2021. Special issue on COVID-19 and novel mixed methods methodological approaches during catastrophic social changes. *J. Mix. Methods Res.* 15 (3), 295–303. <https://doi.org/10.1177/15586898211029100>.
- Fetters, M.D., 2016. "Haven't we always been doing mixed methods research?": lessons learned from the development of the horseless carriage. *J. Mix. Methods Res.* 10 (1), 3–11. <https://doi.org/10.1177/1558689815620883>.
- Fetters, M.D., 2020. *The Mixed Methods Research Workbook-Activities for Designing, Implementing, and Publishing Projects*. Sage, Thousand Oaks, CA.
- Garnett, B.R., Smith, L.C., Kervick, C.T., Ballysingh, T.A., Moore, M., Gonell, E., 2019. The emancipatory potential of transformative mixed methods designs: informing youth participatory action research and restorative practices within a district-wide school transformation project. *Int. J. Res. Method Educ.* 42 (3), 305–316. <https://doi.org/10.1080/1743727X.2019.1598355>.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods merger. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 606–615.
- Guetterman, T.C., Fetters, M.D., 2018. Two methodological approaches to the integration of mixed methods and case study designs: a systematic review. *Am. Behav. Sci.* 62 (7), 900–918. <https://doi.org/10.1177/0002764218772641>.
- Hall, J.N., Ryan, K.E., 2011. Educational accountability: a qualitatively driven mixed-methods approach. *Qual. Inq.* 17 (1), 105–115. <https://doi.org/10.1177/1077800410389761>.
- Hesse-Biber, S.N., Rodriguez, D., Frost, N.A., 2015. A qualitatively driven approach to multimethod and mixed methods research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford Press, Oxford, England.
- Howell Smith, M.C., Babchuk, W.A., Stevens, J., et al., 2020. Modeling the use of mixed methods—grounded theory: developing scales for a new measurement model. *J. Mix. Methods Res.* 14 (2), 184–206. <https://doi.org/10.1177/1558689819872599>.
- Israel, B., 2013. *Methods for Community-Based Participatory Research for Health*. John Wiley & Sons, San Francisco, CA.
- Johnson, R.B., Christensen, L., 2014. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, fifth ed. Sage Publications, Thousand Oaks, CA.
- Johnson, R.B., Christensen, L., 2020. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, seventh ed. Sage Publications, Thousand Oaks, CA.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26. <https://doi.org/10.3102/0013189X033007014>.
- Kaczmarczyk, P., Salamońska, J., 2018. Why we still need ethnoscapes in migration research: empirical evidence from Poland. *J. Mix. Methods Res.* 12 (4), 458–476. <https://doi.org/10.1177/1558689818796370>.
- Kamei, T., Kawada, A., Nishimura, N., Kakai, H., 2021. A fully longitudinal mixed methods case study design: an example based on the illness trajectory of a person with chronic conditions in home-monitoring-based telenursing during the COVID-19 pandemic. *J. Mix. Methods Res.* 15 (3), 413–436. <https://doi.org/10.1177/15586898211019512>.
- Ling, H.L., Pang, M.F., 2022. A vignette-based transformative multiphase mixed methods interventional study featuring Venn diagram joint displays: financial education with Hong Kong early adolescent ethnic minority students. *J. Mix. Methods Res.* 16 (1), 130–149. <https://doi.org/10.1177/1558689821989834>.
- Maleku, A., Kim, Y.K., Kagotho, N., Lim, Y., 2021. Expanding the transformative explanatory sequential mixed methods design archetype in a cross-cultural context: the polemics of African refugee livelihoods in places of resettlement. *J. Mix. Methods Res.* 15 (2), 212–239. <https://doi.org/10.1177/1558689820936378>.
- Martiny, K.M., Toro, J., Höfding, S., 2021. Framing a phenomenological mixed method: from inspiration to guidance. *Front. Psychol.* 12, 602081. <https://doi.org/10.3389/fpsyg.2021.602081>.
- Mayoh, J., Onwuegbuzie, A.J., 2015. Toward a conceptualization of mixed methods phenomenological research. *J. Mix. Methods Res.* 9 (1), 91–107. <https://doi.org/10.1177/1558689813505358>.
- Mertens, D.M., Hesse-Biber, S., 2012. Triangulation and mixed methods research: provocative positions. *J. Mix. Methods Res.* 6 (2), 75–79. <https://doi.org/10.1177/1558689812437100>.
- Mertens, D.M., Bledsoe, K.L., Sullivan, M., Wilson, A., 2010. Utilization of mixed methods for transformative purposes. In: Tashakkori, A., Teddlie, C. (Eds.), *Mixed Methods in Social & Behavioral Research*. Sage, Thousand Oaks, CA, pp. 193–214.
- Mertens, D.M., 2003. Mixed methods and the politics of human research: the transformative emancipatory perspective. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 135–164.
- Mertens, D.M., 2009. *Transformative Research and Evaluation*. Guilford Press, New York, NY.
- Mertens, D.M., 2015. *Research and Evaluation in Education and Psychology*. Sage, Thousand Oaks, CA, pp. 306–307.
- Morgan, D., 2014. *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. Sage, Thousand Oaks, CA.
- Morse, J.M., Niehaus, L., 2009. *Mixed Method Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Morse, J.M., 2015. Issues in qualitatively-driven mixed-method designs: walking through mixed method project. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford Press, Oxford, England.
- O'Cathain, A., 2018. *Practical Guide to Using Qualitative Research With Randomized Controlled Trials*, first ed. Oxford University Press.
- Onwuegbuzie, A.J., Leech, N.L., 2010. Generalization practices in qualitative research: a mixed methods case study. *Qual. Quantity* 44, 881–892. <https://doi.org/10.1007/s11135-009-9241-z>.
- Onwuegbuzie, A.J., Leech, N.L., 2021. Qualitizing data. In: Onwuegbuzie, T., Johnson, B. (Eds.), *Routledge Reviewer's Guide to Mixed Methods Analysis*, pp. 141–150.
- Peroff, D.M., Morais, D.B., Seekamp, E., Sills, E., Wallace, T., 2020. Assessing residents' place attachment to the Guatemalan Maya landscape through mixed methods photo elicitation. *J. Mix. Methods Res.* 14 (3), 379–402. <https://doi.org/10.1177/1558689819845800>.
- Peters, D.H., Adam, T., Alonge, O., Agyepong, I.A., Tran, N., 2013. Implementation research: what it is and how to do it. *BMJ* 347, f6753. <https://doi.org/10.1136/bmj.f6753>.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. Sage Publications, Inc, Thousand Oaks, CA. <https://www.doi.org/10.4135/9781483398341>.
- Plano Clark, V.L., Anderson, N., Wertz, J.A., Zhou, Y., Schumacher, K., Miaskowski, C., 2015. Conceptualizing longitudinal mixed methods designs: a methodological review of health sciences research. *J. Mix. Methods Res.* 239 (4), 297–319. <https://doi.org/10.1177/1558689814543563>.
- Reason, P., Bradbury, H., 2008. Introduction. In: Reason, P., Bradbury, H. (Eds.), *The SAGE Handbook of Action Research: Participative Inquiry and Practice*, second ed. Sage Publications, Thousand Oaks, CA, pp. 1–10.

- Roberts, G., 2003. Teaching using the web: conceptions and approaches from a phenomenographic perspective. *Instr. Sci.* 31, 127–150. <https://doi.org/10.1023/A:1022547619474>.
- Ruffin, M.T., Creswell, J.W., Jimbo, M., Fetters, M.D., 2009. Factors influencing choices for colorectal cancer screening among previously unscreened African and Caucasian Americans: findings from a triangulation mixed methods investigation. *J. Community Health* 34 (2), 79–89. <https://doi.org/10.1007/s10900-008-9133-5>.
- Sandelowski, M., Voils, C.I., Knaff, G., 2009. On quantitizing. *J. Mix. Methods Res.* 3 (3), 208–222. <https://doi.org/10.1177/1558689809334210>.
- Seltzer-Kelly, D., Westwood, S.J., Peña-Guzman, D.M., 2012. A methodological self-study of quantitizing: negotiating meaning and revealing multiplicity. *J. Mix. Methods Res.* 6 (4), 258–274. <https://doi.org/10.1177/1558689811425798>.
- Shim, M., Johnson, R.B., Gasson, S., et al., 2017. A model of dance/movement therapy for resilience building in people living with chronic pain. *Eur. J. Integr. Med.* 9 (1), 27–40. <https://doi.org/10.1016/j.eujim.2017.01.011>.
- Shim, M., Johnson, B., Bradt, J., Gasson, S., 2021. A mixed methods–grounded theory design for producing more refined theoretical models. *J. Mix. Methods Res.* 15 (1), 61–86. <https://doi.org/10.1177/1558689820932311>.
- Tashakkori, A., Johnson, B., Teddlie, C., 2021. *Foundations of Mixed Methods Research*, second ed. Sage, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2009. Integrating quantitative and qualitative approaches in the social and behavioral sciences. In: Bickman, L., Rog, D.J. (Eds.), *Foundations of Mixed Methods Research*, first ed. SAGE Publications, Inc, Thousand Oaks, CA.
- Upadhyaya, M., May, M., Highfield, L., 2015. Integrating classroom, community, mixed-methods research, and community-based participatory research to teach public health practice. *Publ. Health Rep.* 130 (3), 286–292. <https://doi.org/10.1177/003335491513000317>.
- Waller, M.R., Nepomnyaschy, L., Miller, D.P., Mingo, M., 2021. Using a narrative approach to analyze longitudinal mixed methods data. *J. Mix. Methods Res.* 15 (2), 261–283. <https://doi.org/10.1177/1558689820953237>.
- Weaver-Hightower, M.B., 2014. A mixed methods approach for identifying influence on public policy. *J. Mix. Methods Res.* 8 (2), 115–138. <https://doi.org/10.1177/1558689813490996>.
- Wong, S., Cooper, P., 2016. *Int. J. Cyber Soc. Educ.* 9 (2), 45–64. <https://doi.org/10.7903/ijcse.1475>.

Designing integrative mixed methods research

Joseph Alex Maxwell, George Mason University, Fairfax, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

An interactive concept of design	442
The components of a design	442
Goals: why are you doing this study?	442
Conceptual framework: what do you already know or believe?	442
Research questions: what, specifically, do you want to learn?	443
Methods: what will you actually do?	443
Validity and generalizability: how might you be wrong, and how will you address this?	443
The complementarity of differences	443
Combining methods for causal explanation	444
Combining methods for generalization	445
Internal generalization	445
External generalization	446
An illustrative study	446
Conclusions	448
References	448

Mixed methods research, the joint use and integration of qualitative and quantitative approaches, methods, and data, has a much longer history than is usually acknowledged in the self-identified “mixed methods” literature (Maxwell, 2016). In addition, mixed methods approaches are widely used in social science fields that are rarely recognized in this literature (Maxwell, 2018). This chapter will draw on both of these sources to address how to design integration in mixed methods research.

My argument in this chapter is based on an analysis of the differences in goals between quantitative and qualitative research, and the implications of these differences for research design and methods. In terms of goals, quantitative research is primarily focused on the effect of some variables or interventions on other variables or outcomes. Qualitative research, in contrast, aims more to understand the processes (including mental processes) involved in an individual or social setting, and how these are influenced by the context in which that individual or setting is located. These differences result in different strengths and limitations for each approach in dealing with two important issues in social research: the analysis of causation, and the generalizability of the results.

The concept of “design” also requires some discussion. There are various conceptions of “design” in the research literature, with major differences between quantitative and qualitative research. Experimental research has largely adopted a typological approach to design, creating multiple designs that vary in how subjects are assigned to conditions and when variables are measured. Non-experimental quantitative research lacks such an elaboration of typologies, but most such studies share an important feature of research design with experimental research: that the research parameters are established in advance of data collection, and can’t be altered during the study, in order to ensure the comparability of the data for aggregation and analysis. Robson (2011) referred to such inflexible approaches as “fixed” designs (Exploratory analysis of quantitative data, however, is flexible and widely used [Robson, 2011, pp. 419–420]).

Qualitative research, in contrast, has largely avoided typological thinking about design, since its main strength is the adaptability of its data collection and analysis procedures in response to unanticipated events or findings; Robson (2011) labeled these “flexible” designs. However, qualitative researchers have typically been much less explicit about the concept of design. Schwandt (2007) stated that

Designs or plans for conducting qualitative studies vary considerably; there are no prescribed designs as, for example, one might find in survey research, experimental, and quasi-experimental studies in the social sciences (p. 273).

Some qualitative textbooks (e.g., Glesne, 2016) don’t use the term “design” at all; others mention research design only in vague terms. Mason (2018) describes a more explicit approach to “qualitatively-driven research design,” addressing the intellectual puzzle, theoretical orientation, methodological approach and research strategy, data analysis, and ethical considerations (pp. 49–50), but with no attempt to identify “types” of designs.

Self-identified “mixed methods research” has largely followed a typological approach to design, identifying an array of design types that can be adopted by researchers, delineating the dimensions on which mixed methods studies can vary and identifying the possible and actual combinations of qualitative and quantitative methods. Caracelli and Greene (1997) identified two basic types of mixed method designs, which they called “component” and “integrated” designs. Component designs are ones in which “the

methods are implemented as discrete aspects of the overall inquiry and remain distinct throughout the inquiry”, while integrated designs involve “a greater integration of the different method types” (pp. 22–23). Within these broad categories, they describe seven subtypes, based largely on the goals for combining methods. More recently, [Creswell and Plano Clark \(2018\)](#), [Mertens \(2018\)](#), and [Tashakkori et al. \(2021\)](#) presented more elaborate typologies of design, involving the order in which the qualitative and quantitative methods are used, the relative dominance of the two approaches, and the degree of integration.

In contrast, researchers outside of this self-identified community, in fields such as political science, anthropology, linguistics, education, and many of the natural sciences ([Maxwell, 2016, 2018](#)), who were integrating quantitative and qualitative approaches and methods long before the term “mixed methods” was coined, have largely avoided typological thinking about design. In this, they are similar to qualitative research. Johnson, in what is apparently the first textbook on designing such integration (1978), did not explicitly define “research design,” but simply stated that “the purpose of research design is to use theoretical arguments to develop expectations about the world, and then to test them by collecting empirical data” (p. 2). [Gorard \(2017\)](#), primarily addressing quantitative research, defined research design as “the structure and organization of a research project” (2017, p. 203).

An interactive concept of design

In the present chapter, I use a broader conception of design than that common in self-identified mixed methods research ([Maxwell and Loomis, 2003](#); [Maxwell, 2013](#)). Webster’s Collegiate Dictionary (1984, p. 343) defined this conception of design as “an underlying scheme that governs functioning, developing, or unfolding” and “the arrangement of elements or details in a product or work of art.” This is the meaning of “design” in everyday use, in referring to the design of a building, a machine, or a work of art. Applied to research, this conception is of design as the actual relationships among the components of a study, not just a categorization of the study as a particular type of design, or as a fixed sequence of steps in conducting the study.

This approach to design focuses on the components of a study, and how these interact with one another in the planning and conduct of the research. For this reason, I originally labeled this an “interactive” approach to design, but the term “integrative” is equally appropriate, as the different components need to be integrated with one another in order to function effectively. This meaning includes not just the planned design of a study, but also the actual design of the research “on the ground,” which may differ from the planned design, and even from what the researcher claims or believes about the design, and may change during the course of a study.

This conception of design was originally developed for qualitative research ([Maxwell, 1996](#)), and was later applied to mixed methods design ([Maxwell and Loomis, 2003](#); [Maxwell et al., 2015](#)), but without emphasizing the implications of this design strategy for integrating qualitative and quantitative methods and approaches. The integration of the qualitative and quantitative aspects of each of the design components is central to the functioning of the design, and the differences between quantitative and qualitative research make integration in mixed methods research a rather different and more complex issue than for mono-method studies. [Greene \(2007\)](#) argued that “the interaction challenge in integrated designs is under theorized and understudied ... [but] constitutes the very heart of integrative mixed methods inquiry” (p. 125). Thus, my discussion of mixed methods design will focus importantly on these differences, and on the advantages, as well as the challenges, that they create for effective integration.

In what follows, I will first briefly describe the components of this design model, and how these interact with one another (for a more detailed discussion, see [Maxwell, 2013](#)). I will then present what I see as the most important differences between quantitative and qualitative approaches, and how integrating the two creates opportunities for deeper understandings of the phenomena studied than are possible with a mono-method approach.

The components of a design

This model of research design has five components: the researcher’s goals, conceptual framework, research questions, methods (including sampling/selection, research relationships, data collection, and analysis), and validity concerns (For a much more detailed discussion of these, see [Maxwell, 2013](#), and [Maxwell and Loomis, 2003](#)).

Goals: why are you doing this study?

The researcher’s goals include not only the explicit research goals of the study, but also the practical and personal goals that these may facilitate. These goals are often implicit, in which case they may unconsciously affect the research, and it is thus important to make them explicit.

Conceptual framework: what do you already know or believe?

The conceptual framework of a study (often called the “theoretical framework”) includes all of the concepts, theories, and beliefs that the researcher holds about the phenomena, settings, and individuals to be studied. As is true for goals, these are often implicit, and need to be made explicit. However, unlike the goals, this framework is generally something that needs to be constructed in planning the study. Its sources aren’t limited to “the literature” on the specific topic or research area; it can also be drawn from

literature in other fields or through communication with other researchers, or created through “thought experiments,” and modified during the course of the study.

Research questions: what, specifically, do you want to learn?

This component, the questions that the study is explicitly designed to answer, is the “hub” of my design model, the one that most directly connects to the other components. It is usually the component that is most explicitly formulated, although these questions may evolve during the study (It is different from the interview questions, which are part of the data collection methods).

Methods: what will you actually do?

I define this component broadly, to include all of the actions planned or taken in gaining access or gathering data during the study. It consists of four main parts: the relationships created with participants, or with “gatekeepers” that control access to participants; the selection of participants, times, and settings for data collection; the actual data collection strategies; and the data analysis procedures. While these may be planned in some detail in advance, in qualitative research they are typically not fully developed before the study begins, and may be modified substantially as the research proceeds.

Validity and generalizability: how might you be wrong, and how will you address this?

This component is often treated as part of methods. However, I see the key issue as the validity threat—a way that the researcher’s conclusion might be wrong—and focus on how that threat could be assessed and dealt with, which generally includes methods but is in principle much broader than this (Maxwell, 2013, chapter 6, 2017). I also include in this component the strategies for addressing these threats, and the potential generalizability of the conclusions. Qualitative researchers usually conceptualize the latter as transferability—something that is primarily the responsibility of the reader or potential user of the results, although the researcher can facilitate this transfer by detailed description of the setting and the processes that led to the results; I discuss this in more detail below.

One aspect of a research design—ethical issues—doesn’t have a single obvious place in this model. I see it as most closely aligned with goals, but it also could fit into validity, or be seen as a separate component of the design. Ravitch and Riggan (2017) seem to implicitly include it in the conceptual framework, which they define somewhat more broadly than I do.

All of these components interact with all of the other components. In addition, particular elements of a single component interact—for example, the researcher may have multiple goals, and any of these may have consequences for other goals, as well as for elements of other components. This is particularly true for mixed methods research; the qualitative and quantitative elements of any component may differ substantially, and successfully integrating these may become a major challenge.

The following diagram illustrates the relationships among these design components. Each component interacts with all of the others; the research questions are the component that most directly connects to all of the other components (Fig. 1).

Research design and research proposals.

One of the benefits of this model of research design is that it matches a widespread structure for a research proposal. This structure typically includes the following sections:

1. Why the study is worth doing (goals)
2. What we already know or believe about the topic (conceptual framework)
3. What specific questions the study will attempt to answer (research questions; these may be stated earlier in the proposal, but need to be justified by explaining the goals and conceptual framework)
4. What settings will be studied, what relationships will these require, and what data collection and analysis strategies will be used (methods)
5. Assumptions, limitations, and potential errors in interpretation (validity)

Developing the design for a study in this way makes it easier to create a proposal for a dissertation, or for funding for the project (for a more detailed presentation, see Maxwell, 2013, Chapter 7).

The complementarity of differences

The differences between qualitative and quantitative approaches are not simply a problem to be overcome; they are also the source of important advantages in combining the two. Many of these differences are complementary; the strengths of qualitative research often compensate for the limitations of quantitative research, and vice versa (Beach, 2021). I see these complementary differences as a key issue in integrating the two approaches. In what follows, I will focus on these differences, and how combining the two approaches creates a stronger and broader result than either approach can achieve by itself. I see two primary issues for which there are major advantages in combining the two approaches: causal explanation, and generalization.

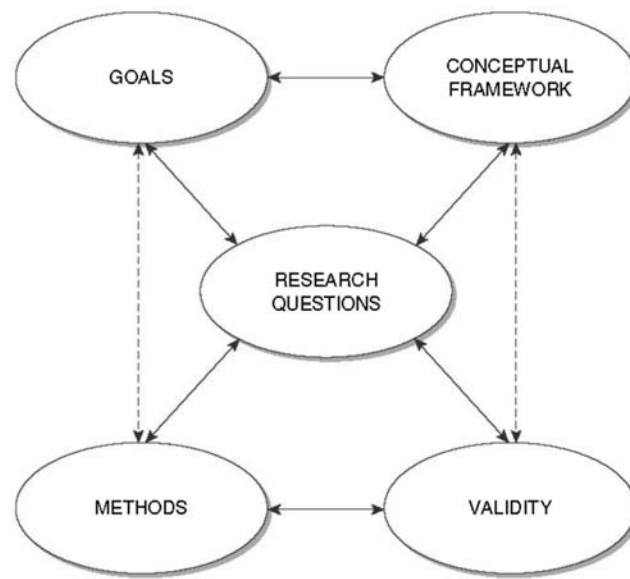


Fig. 1 Model of the relationships among research components.

Combining methods for causal explanation

I will begin with what I see as a major difference in the goals, and typical results, of the two approaches. Quantitative research is particularly good at showing *that* a particular treatment, event, or variable influenced another variable or contributed to a particular outcome. Qualitative research, in contrast, is particularly good at showing *how* a situation or event influenced another situation or event—providing an understanding of the processes that led to that outcome. For this reason, mixed methods research has important advantages in developing causal explanations of a situation or outcome.

Shadish, Cook, and Campbell, in their influential work on experimental and quasi-experimental methods (2002), stated that

the unique strength of experimentation is in describing the consequences attributable to deliberately varying a treatment. We call this **causal description**. In contrast, experiments do less well in clarifying the mechanisms through which and the conditions under which that causal relationship holds—what we call **causal explanation** (p. 9; emphasis in original).

They added that “qualitative methods provide an important avenue for discovering and exploring causal explanations” (p. 389), and described how, in one experiment, the actual cause of the result was “revealed only by the qualitative examination of the actual operation of groups” (p. 391).

This difference in the goals of the two approaches is closely connected to a difference in the view of causation associated with each approach, a part of the conceptual framework of a study’s design. Quantitative researchers often operate (implicitly or explicitly) with a positivist or “regularity” concept of causation—that causality is simply a consistent relationship between cause and effect, with nothing “behind” this that can be scientifically studied. Illari and Russo, in their book *Causality: Philosophical theory meets scientific practice* (2014), called such conceptions “difference-making” theories. These theories focus on the idea that causes make a regular, predictable difference in the things that they affect.

Qualitative researchers, in contrast, have often rejected the entire concept of causation, largely equating this with the positivist understanding of this concept. There have been two main versions of this rejection. The first of these argued that qualitative research deals with meaning and interpretation rather than causality, drawing on Max Weber’s distinction between understanding (*verstehen*) and (causal) explanation (*erklärung*). The second, and stronger, version, prominently advanced by Egon Guba and Yvonna Lincoln, was that the very concept of causation was invalid in social research: “There exist multiple, socially constructed realities ungoverned by natural laws, causal or otherwise” (Guba and Lincoln, 1989, p. 86), and “there are simply too many problems with causality to continue its use” (Lincoln and Guba, 1985, p. 159). This dispute over causation was a major feature of the so-called “paradigm wars” of the 1980s and 1990s, which polarized relationships between qualitative and quantitative researchers.

However, other qualitative researchers (e.g., Hammersley, 2012; Maxwell, 2012, chapter 3; Miles and Huberman, 1994) have accepted the legitimacy of causation, although a quite different understanding of causation from that held by most quantitative researchers. This conception emphasizes causal processes rather than regularities; it focuses on how a particular event or property of a situation affects other properties or subsequent events. This type of theory has been termed a realist, process, or mechanism theory; Illari and Russo (2014) called these “production accounts.”

The philosopher Nancy Cartwright, who has probably done the most to challenge traditional views of causation, stated that current approaches to causation

are not alternative, incompatible views about causation; they are rather views that fit different kinds of causal systems ... there is no single interesting characterizing feature of causation; hence no off-the-shelf or one-size-fits-all method for finding out about it, no "gold standard" for judging causal relations.

Cartwright (2007, p. 2).

These different conceptions of causation lead to, and draw on, different types of research questions and methods, and raise different potential validity threats to a study's conclusions. Integrating the two approaches to causation requires recognizing that each provides a major piece of a full causal explanation of some outcome or situation, and such integration is often essential to dealing with potential validity threats to that explanation.

This difference is closely connected to two types of data analysis, which I have labeled "categorizing" and "connecting" strategies (Maxwell and Miller, 2008). Categorizing strategies for analyzing data are intrinsic to quantitative analysis; data are aggregated by category (variable), allowing the researcher to identify the mean and variance of these data. Even non-numerical (nominal and ordinal measurement) data are aggregated in this way, though with different measures of central tendency and dispersion. Connections are then identified between variables; however, these provide only an aggregate measure of their association (e.g., correlation), having lost the local connections between particular events or beliefs in a specific time and context. This approach is well suited for identifying general causal patterns, consistent with a regularity understanding of causation.

Qualitative analysis, although it often employs coding of data to create categories, can also retain and analyze the unique connections among specific beliefs, actions, and events. Miles and Huberman (1984, p. 132, 1994, p. 228) used the term "local causality" for these relationships. Such local causal analysis may be essential for understanding the actual processes that led to a particular outcome. Thus, Becker (2009) argued that qualitative researchers

analyze directly observed processes, chains of events that produce the outcomes we want to understand ... Process analysis presents substantial difficulties for quantitative studies, which generally have to substitute such makeshifts as panel studies, population measurements at several selected times, etc. for realistic, more or less continuous empirical observation (p. 550).

Similarly, Goldenberg (1992), in a qualitative case study of the reading progress of two students and the effects of their teacher's behavior on this, stated that "if we see these dimensions as variables divorced from this context, we risk distorting the role that they actually play" (p. 540).

Combining methods for generalization

This difference in approaches to causation is also fundamental for generalizing causal (and other) conclusions, based on the specific individuals or settings for which data were collected. Such generalization can be either to the setting or population sampled (the "target" population), or to other populations or settings. Both of these are included in the quantitative concept of "external validity" (Cook and Campbell, 1979, p. 71; Shadish et al., 2002, pp. 53, 83–84), although Cook and Campbell clearly distinguished the two (pp. 71–73). However, these two types of generalization are quite different in the methods they require; I have labeled these "internal" and "external" generalization, respectively (Maxwell, 1992, 2017; Maxwell and Chmiel, 2014; see also Eisenhart, 2009).

Internal generalization

Internal generalization is generalization within the setting or population studied; for qualitative research, this type of generalization has been termed *petite generalization* (generalization within a case) by Erickson (1986), and *lower-order generalizability* (the generalizability of findings within the unit of analysis) by Brown-Saracino et al. (2008). Random sampling is the major quantitative strategy for internal generalization, although the "statistical significance" of such results is frequently misused or misinterpreted (Nuzzo, 2014; Amrhein et al., 2019).

Random sampling is rarely possible in qualitative studies, due to the large sample sizes required for this to be meaningful, but systematic non-random sampling can still be useful in making inferences from the individuals on whom data were collected, and the specific times and settings observed, to the larger social setting or population studied. Miles and Huberman argued that sampling broadly conceptualized (and not limited to random sampling) is a key issue for all research: "Knowing, then, that one cannot study everyone everywhere doing everything, even within a single case, how does one limit the parameters of a study?" (1984, p. 36) They stated that

you are not only sampling people, but also settings, events, and processes. It is important to line up these parameters with the research questions as well, and to consider whether your choices are doing a representative, time-efficient job of answering them (1984, p. 41).

In order to make claims about the internal generalizability of their findings, mixed methods (and qualitative) researchers need to address these issues—are the people, settings, activities, and times on which they have actually collected data adequately representative of the larger population or case that they are studying?

In addition, internal generalizability in a full sense needs to address not just whether the results capture the typical views, characteristics, or actions of the persons involved, but also the diversity of these within the case or population studied. [Ercikan \(2009\)](#) argued that quantitative studies “may use very large samples but focus on overall group results and have very little generalizability for key subgroups” (p. 211).

External generalization

Generalizing from the results for one population or setting to different populations or settings is a completely different matter. Random sampling can’t be used for this purpose, and quantitative researchers have few other tools for such generalization ([Cartwright, 2016](#)). Both quantitative and qualitative researchers have typically argued that “similarity” to the setting or population studied is the basis for generalization to other settings or populations. Shadish, Cook, and Campbell stated, as their first “principle of generalized causal inference,” that “scientists generalize by judging the apparent similarities between the things they studied and the targets of generalization” (2002, p. 353). Similarity was also invoked by Lincoln and Guba in their presentation of “transferability” (1985, pp. 124–198).

Cartwright and Hardie, in their book “Evidence-based Policy: A Practical Guide to Doing it Better” (2012), responded that

For us, the key question is how good a job this advice [similarity] does in getting you from “it worked there” to “it will work here.” The answer: you are lucky if it gets you anywhere (p. 46).

Similarity is both too vague (what features must be similar?) and too demanding (almost no two contexts are similar in all important respects), and provides no guidance in understanding how your specific context will affect the outcome of the policy (pp. 46–49).

However, the qualitative concept of “transferability” does not depend on similarity. Instead, it can draw on an understanding of the processes that resulted in the outcome in the original study, and how these processes are likely to operate in a specific new context. This approach was described in detail by Cartwright and Hardie, who emphasized the fundamental role of context in such generalization. Although they never used the term “qualitative” in presenting these issues, their book is to a substantial extent a detailed argument for the importance of qualitative methods and understandings in implementing any policy in a new context.

An essential issue in understanding these processes is meaning: the beliefs, values, and motives held by participants in the two contexts. These meanings are inherently involved in the processes involved, in both the setting studied and the new setting. As the psychologist [Bandura \(1986\)](#) argued, “What people think, believe, and feel affects how they behave. The natural and extrinsic effects of their actions, in turn, partly determine their thought patterns and affective reactions” (p. 25).

This is a second major strength of qualitative research for generalization; qualitative methods for data collection and analysis are necessary for an in-depth understanding of participants’ meanings. These meanings and processes inherently influence, and are influenced by, the context in which they occur, and qualitative methods are likewise essential for an understanding of this context and how it shapes the processes that lead to particular outcomes.

For this reason, transferability is primarily the responsibility of the reader or potential user of the original findings, who can best assess how these processes might operate in a specific new setting ([Lincoln and Guba, 1985](#), p. 298). However, the authors of the original study can facilitate such use by providing detailed information about the processes, meanings, and contexts involved in the original study.

An illustrative study

A classic mixed methods study, one that predates the coining of the term “mixed methods,” provides an example of how integrating the methods and findings of quantitative and qualitative methods and data led to a much deeper understanding than either method was able to achieve on its own. This study, an evaluation of a multi-site experimental federal housing-subsidy program, was described by [Trend \(1979\)](#) in a volume that contributed significantly to the development of mixed methods research ([Cook and Reichardt, 1979](#)).

Trend characterized the study as a “naturalistic experiment” (p. 69); it did not involve a control group. Extensive quantitative data were collected on agency activities, expenses, demographic characteristics of clients, and housing quality; survey researchers interviewed participants at scheduled interviews. In addition, each site had an observer (usually an anthropologist) who prepared a qualitative case study of that site, using field observations, interviews, and documents. The intent was that program outcomes would be determined through analysis of the quantitative data, while the case studies would provide a holistic picture of program process (Trend, p. 70).

In one site, composed of one urban and two rural subsites, the contracting agency was located 70 miles from the site, and had never been responsible for such a grass-roots program, but appeared aggressive and anxious to succeed, and intended to run a “lean

and mean" operation; few supportive services were offered. In the observer's view, the original plan never materialized. Despite delays in signing the contract, the agency was required to meet the original contract deadlines. Initial enrollment of families (which required the families to find acceptable housing) was slow in the urban site, which led the contracting agency to impose enrollment quotas and to order the curtailment of Black households, as these were overrepresented. The staff (most of whom had social service backgrounds) felt that the process was inhumane, and complained of overwork; the observer insisted that the program was hampered by managerial conflicts (with racial overtones) and incompetence, and that many families simply couldn't find acceptable housing.

However, when the quantitative outcome measures were finally reported, they showed that this site had performed quite well compared to other sites. It was demographically balanced and relatively cheap to run, and had achieved its outcome goals in terms of number of families who found acceptable housing. While neither side doubted "the facts" produced by the other, their interpretations of these facts differed radically. The agency conducting the evaluation sided with the quantitative results, and the observer was repeatedly told to rewrite his analysis to fit the quantitative conclusions.

After multiple unsuccessful attempts, Trend and the observer made a sustained effort to get at what had really been going on, using both the quantitative and qualitative data. They eventually came up with a coherent, process explanation for almost all of the data, one that went well beyond either the quantitative or the initial qualitative results, and that revealed serious shortcomings in both accounts. Separating the rural and urban subsites showed that the rural subsites were largely responsible for the positive outcome results. In addition, the fiscal requirements imposed by the granting agency had never been adequately explained to the staff at the urban subsite, leading to their frustration with the contracting agency's orders; the early departure of disgruntled employees helped lower costs (Much later, the housing quality measure used by the agency was shown to be flawed).

Trend concluded that multiple methods may not lead to an easy integration of findings, and that "unanimity may be the hallmark of work in which other avenues to explanation have been closed off prematurely" (p. 68). Had the discrepancy between the qualitative and quantitative accounts been discovered earlier, or the two approaches more closely integrated, it is possible that the observer would have been subtly or overtly coerced to make his conclusions fit the "hard" data (p. 84). Trend thus argued that "the proliferation of divergent explanations should be encouraged" (p. 68), but also that an effort be made to develop an account which does justice to all of the conflicting perspectives.

Important earlier studies that integrated qualitative and quantitative approaches, such as Blau, *The Dynamic of Bureaucracy* (1963), Festinger et al. *When Prophecy Fails* (1956), Becker et al. *Boys in White: Student Culture in Medical School* (1961), and Milgram, *Obedience to Authority* (1974), have received almost no attention in the mixed methods literature, despite being recognized as classics in their specific fields. None of these authors identified their study's design as a particular type, nor does typologizing their designs provide much useful additional understanding of what they did. Becker et al. (1961) described their design as follows:

In one sense, our study had no design. That is, we had no well-worked-out set of hypotheses to be tested, no data-gathering instruments purposely designed to secure information relevant to these hypotheses, no set of analytic procedures specified in advance. Insofar as the term "design" implies these features of elaborate prior planning, our study had none.

If we take the idea of design in a larger and looser sense, using it to identify those elements of order, system, and consistency our procedures did exhibit, our study had a design. We can say what this was by describing our original view of the problem, our theoretical and methodological commitments, and the way these affected our research and were affected by it as we proceeded (p. 17).

For a discussion of the designs of these and other early mixed methods studies, see Maxwell (2016, 2018) and Maxwell and Loomis (2003).

The study by Milgram is particularly explicit about the relative strengths of the two approaches, and how these were integrated. To understand why people might obey orders that clearly violated widely-held social values, Milgram designed an elaborate series of experiments in which participants were deceived into believing that they were part of a study of the effects of punishment on learning, and were then told to give increasingly severe (faked) "electrical shocks" to a supposed "subject" who was actually an accomplice of the researchers, and who feigned pain and eventual refusal to cooperate. The researchers employed numerous different experimental conditions designed to determine the effect of different variables on the degree of obedience (the dependent variable), and collected quantitative data about the level of shock that subjects administered (the main measure of obedience) in each of the different conditions (After publication, the study was criticized on ethical grounds).

However, Milgram also stated that

From each person in the experiment we derive one essential fact: whether he has obeyed or disobeyed. But it is foolish to see the subject only in this way. For he brings to the laboratory a full range of emotions, attitudes, and individual styles.... We need to focus on the individuals who took part in the study not only because this provides a personal dimension to the experiment but also because the quality of each person's experience gives us clues to the nature of the process of obedience (1974, p. 44).

The researchers covertly recorded the subjects' behavior during the experiment, and interviewed some subjects at length after the experiment was over to determine their reasons for compliance or refusal. The results chapters of the book combine quantitative tables and graphs with pictures of the experimental setup, observational notes, excerpts from recorded observations and interviews, and case studies of particular subjects' responses to the experimental situation, with extensive discussion of the social processes and subjective interpretations through which obedience and resistance to authority develop. And in discussing potential validity threats to the study's conclusions, Milgram used both the quantitative results from the experimental manipulations and qualitative data from the observations and interviews to rule out these threats.

Conclusions

The main lesson that I draw from these examples and analyses is that typological approaches to research design, which have dominated the self-identified "mixed methods" literature, are neither necessary, nor particularly helpful, in designing mixed methods studies. Such categories may be useful in describing a completed study, but not as a starting point for planning the study. I argue that creating a workable design needs to address the researcher's goals for the study, the conceptual framework for understanding the phenomena to be studied, the specific research questions to be answered, the methods to be used to answer these, and the potential validity threats to these answers and how these can be addressed. This is not a linear sequence; any of these components may influence any of the other components, and the design of a study needs to be an interactive and integrative process. The five components of my design model are also useful in developing a research proposal, since they cover the main issues that a proposal must address.

Integrating the quantitative and qualitative components of a design is often more challenging than in a monomethod study, but it also has important strengths for drawing conclusions. I've described two major issues for which quantitative and qualitative approaches have complementary strengths and limitations: the analysis of causation, and the facilitation of generalization. In addition, there are innumerable benefits that are specific to particular studies, as illustrated by the examples presented. In my view, the integration of methods is better facilitated by this design model than by typological conceptions of design.

References

- Amrhein, V., Greenland, S., McShane, B., et al., 2019. Scientists rise up against statistical significance. *Nature* 567, 305–307.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Beach, D., 2021. Evidential pluralism and evidence of mechanisms in the social sciences. *Synthese* 199 (3), April 2021.
- Becker, H.S., Geer, B., Hughes, E.C., Strauss, A.L., 1961. *Boys in White: Student Culture in Medical School*. University of Chicago Press, Chicago.
- Becker, H.S., 2009. How to find out how to do qualitative research. *Int. J. Commun.* 3, 545–553. Feature.
- Blau, P., 1963. *The Dynamics of Bureaucracy*, rev. ed. University of Chicago Press, Chicago.
- Brown-Saracino, J., Thurk, J., Fine, G.A., 2008. Beyond groups: seven pillars of peopled ethnography in organizations and communities. *Qual. Res.* 8 (5), 547–567.
- Caracelli, V.J., Greene, J.C., 1997. Crafting mixed-method evaluation designs. In: *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. New Directions for Evaluation, no. 74. San Francisco: Jossey-Bass, pp. 19–32.
- Cartwright, N., Hardie, J., 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford University Press, Oxford, UK.
- Cartwright, N., 2007. *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge University Press, Cambridge, UK.
- Cartwright, N., 2016. Where's the rigor when you need it? *Found. Trends Account.* 10 (2–4), 106–124. <https://doi.org/10.1561/14000000045?>
- Cook, T.D., Campbell, D.T., 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Houghton Mifflin, Boston.
- Cook, T.D., Reichardt, C.S., 1979. *Qualitative and Quantitative Methods in Evaluation Research*. Sage Publications, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage Publications, Thousand Oaks, CA.
- Eisenhart, M., 2009. Generalization from qualitative inquiry. In: Ercikan, K., Roth, W.-M. (Eds.), *Generalizing From Educational Research*. Routledge, New York, pp. 51–66.
- Ercikan, K., 2009. Limitations in sample-to-population generalizing. In: Ercikan, K., Roth, W.-M. (Eds.), *Generalizing From Educational Research*. Routledge, New York, pp. 211–234.
- Erickson, F., 1986. Qualitative methods. In: Wittrock, M. (Ed.), *Handbook of Research on Teaching*. Macmillan, New York.
- Festinger, L., Riecken, H.W., Schachter, S., 1956. *When Prophecy Fails*. University of Minnesota Press, Minneapolis.
- Glesne, C., 2016. *Becoming Qualitative Researchers: An Introduction*, fifth ed. Pearson, Boston, MA.
- Goldenberg, C., 1992. The limits of expectations: a case for case knowledge of teacher expectancy effects. *Am. Educ. Res. J.* 29, 517–544.
- Gorard, S., 2017. An introduction to the importance of research design. In: Wyse, D., Selwyn, N., Smith, E., Suter, L.E. (Eds.), *The BERA/SAGE Handbook of Educational Research*. SAGE Publications, London, pp. 203–212.
- Greene, J., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Guba, E., Lincoln, Y., 1989. *Fourth Generation Evaluation*. SAGE Publications, Thousand Oaks, CA.
- Hammersley, M., 2012. Qualitative causal analysis: grounded theorizing and the qualitative survey. In: Cooper, B., Glaeser, J., Gomm, R., Hammersley, M. (Eds.), *Challenging the Qualitative-Quantitative Divide. Explorations in Case-Focused Causal Analysis*. Continuum, London, UK, pp. 72–95.
- Illari, P., Russo, F., 2014. *Causality: Philosophical Theory Meets Scientific Practice*. Oxford University Press, New York, NY.
- Johnson, A.W., 1978. *Quantification in Cultural Anthropology: An Introduction to. Research Design*. Stanford University Press, Stanford, CA.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. SAGE Publications, Thousand Oaks, CA.
- Mason, J., 2018. *Qualitative Researching*, third ed. Sage Publications, Thousand Oaks, CA.
- Maxwell, J.A., Chmiele, M., 2014. Generalization in and from qualitative analysis. In: Flick, U. (Ed.), *SAGE Handbook of Qualitative Data Analysis*. SAGE Publications, London, UK, pp. 540–553.
- Maxwell, J.A., Loomis, D., 2003. Mixed methods design: an alternative approach. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage Publications, Thousand Oaks, CA, pp. 241–271.
- Maxwell, J.A., Miller, B., 2008. Categorizing and connecting strategies in qualitative data analysis. In: Hesse-Biber, S., Leavy, P. (Eds.), *Handbook of Emergent Methods*. Guilford Press, New York, pp. 461–477.

- Maxwell, J.A., Chmiel, M., Rogers, S., 2015. Designing integration in mixed method and multi-method research. In: Hesse-Biber, S., Johnson, B. (Eds.), *Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, UK, pp. 223–239.
- Maxwell, J.A., 1992. Understanding and validity in qualitative research. *Harv. Educ. Rev.* 62, 279–300 (Fall 1992).
- Maxwell, J.A., 1996. *Qualitative Research Design: An Interactive Approach*. Sage Publications, Thousand Oaks, CA.
- Maxwell, J.A., 2012. *A Realist Approach for Qualitative Research*. Sage Publications, Thousand Oaks, CA.
- Maxwell, J.A., 2013. *Qualitative Research Design: An Interactive Approach*, third ed. Sage Publications, Thousand Oaks, CA.
- Maxwell, J.A., 2016. Expanding the history and range of mixed methods research. *J. Mix. Methods Res.* 10 (1), 12–27. <https://doi.org/10.1177/1558689815571132>.
- Maxwell, J.A., 2017. The validity and reliability of research: a realist perspective. In: Wyse, D., Selwyn, N., Smith, E., Suter, L.E. (Eds.), *The BERA/SAGE Handbook of Educational Research*. SAGE Publications, London, pp. 116–140.
- Maxwell, J.A., 2018. The “silo problem” in mixed methods research. *Int. J. Mult. Res. Approaches* 10 (1), 317–327. <https://doi.org/10.29034/ijmra.v10n1a20>.
- Mertens, D., 2018. *Mixed Methods Design in Evaluation*. Sage Publications, Thousand Oaks, CA.
- Miles, M.B., Huberman, A.M., 1984. *Qualitative Data Analysis: A Sourcebook of New Methods*. Sage, Thousand Oaks, CA.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. Sage, Thousand Oaks, CA.
- Milgram, S., 1974. *Obedience to Authority: An Experimental View*. Harper & Row, New York, NY.
- Morse, J.M., Niehaus, L., 2009. *Mixed Methods Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Nuzzo, R., 2014. Scientific method: statistical errors. *Nature* 506, 150–152. <https://doi.org/10.1038/506150a>.
- Ravitch, S.M., Riggan, M., 2017. *Reason and Rigor: How Conceptual Frameworks Guide Research*, second ed. Sage, Thousand Oaks, CA.
- Robson, C., 2011. *Real World Research*, third ed. John Wiley & Sons, Chichester, UK.
- Schwandt, T.A., 2007. *The SAGE Dictionary of Qualitative Inquiry*, third ed. Sage Publications, Thousand Oaks, CA.
- Shadish, W.R., Cook, T.D., Campbell, D.T., 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin, Boston, MA.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research*, second ed. Sage Publications, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2006. A general typology of research design featuring mixed methods. *Res. Sch.* 13 (1), 12–28.
- Trend, M., 1979. On the reconciliation of qualitative and quantitative analyses: a case study. In: Cook, T.D., Reichardt, C.S. (Eds.), *Qualitative and Quantitative Methods in Program Evaluation*. Sage, Thousand Oaks, CA.
- Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison Wesley, ISBN 978-0-201-07616-5.

Ten mixed methods integration strategies for obtaining a detailed understanding

Judith Schoonenboom, University of Vienna, Department of Education, Vienna, Austria

© 2023 Elsevier Ltd. All rights reserved.

Introduction	450
Integration of claims	451
Ten integration strategies for mixed methods research	453
Strategy 1: incorporate the perspectives of all relevant stakeholders	454
Strategy 2: develop a range of data sets	455
Strategy 3: integrate claims step by step	456
Strategy 4: consider various types of integration	457
Strategy 5: use each integration point to reflect on the next step	457
Strategy 6: integrate qualitative and quantitative data into a single data set	457
Strategy 7: search for contradictions	459
Strategy 8: revisit and re-analyze the data	459
Strategy 9: analyze subgroups	459
Strategy 10: plan and re-plan the study	459
Conclusion	460
References	461

Introduction

Researchers in education often strive for a detailed understanding. They want to obtain an understanding that includes various groups of students, various stakeholders, and different aspects of a phenomenon. For instance, in [Example 3](#) below, the aim is to understand the effects of providing textbooks to primary school pupils with differing pretest scores. In [Example 2](#), the aim is to understand how children, parents, teachers, and school principals in one preschool experience and evaluate the school playground. In [Example 1](#), the aim is to understand teachers' attitudes to student-centered teaching (i.e., their beliefs), their actual teaching behavior, and how the two are related.

This kind of detailed understanding is essential in order to develop educational interventions that work. Before implementing a larger-scale intervention, it is important to identify and incorporate any group differences, conflicting stakeholder views, and inconsistencies between belief and behavior. Only then will it become possible to design the optimal intervention for all concerned ([Pawson and Tilley, 1997](#)).

This chapter describes ten research strategies that can help mixed methods researchers to develop a detailed understanding of the target phenomenon ([Greene, 2007, 2015](#); [Flick, 2017](#); [Schoonenboom, 2019](#); [Uprichard and Dawney, 2019](#)). In mixed methods studies, researchers derive conclusions from qualitative and quantitative data, which they combine using a process called *integration* ([Creswell and Plano Clark, 2018](#); [De Allegri et al., 2020](#); [Tashakkori et al., 2021](#)) as in the following real-world example.

Example 1

In a study by [Assen et al. \(2016\)](#), the target group were tutors at one Dutch university, who had been appointed to facilitate and support students' learning process. This role requires a student-centered approach to teaching rather than the more traditional teacher-centered style based on "knowledge transmission":

The purpose of this study was to explore the discrepancy between teacher beliefs and behavior in a Problem-Based Learning (PBL) environment. Using a survey and observations, this study demonstrated that tutors prefer learner-oriented beliefs, but in their teacher behavior they showed a more traditional approach to teaching. Analysis of semi-structured interviews indicated that this inconsistency could be attributed to the way in which problem-based learning is embedded in the curriculum, the confidence teachers have in the self-directed capabilities of students, and the self-confidence of teachers regarding their own facilitation skills.

[Assen et al. \(2016, p. 12\)](#).

In [Example 1](#), detailed understanding emerges from learning about the participating teachers' beliefs and behaviors, the discrepancy between those beliefs and behaviors, and how that discrepancy can be explained. The explanation can be characterized as

detailed because it provides three distinct reasons for the discrepancy by exploring various aspects of the tutors' teaching approach from several perspectives and comparing and connecting those aspects.

The study referred to in [Example 1](#) is a mixed methods study because the researchers used both qualitative and quantitative methods and integrated their findings ([Creswell and Plano Clark, 2018](#); [De Allegri et al., 2020](#); [Tashakkori et al., 2021](#)). On the quantitative side, they used the survey method, asking teachers closed questions to identify their beliefs. The survey provided quantitative, numerical, and structured data, enabling the teachers' answers to be represented as numerical values in tabular form. Based on these data, the researchers could calculate the extent to which the teachers' beliefs could be characterized as "student-centered." On the qualitative side, observations of the teachers' behaviors provided unstructured video data confirming that those behaviors were more teacher-centered than their declared beliefs. To explain this discrepancy between belief and behavior, the researchers used one additional qualitative method: semi-structured interviews. The interviews yielded audio data that were then transcribed. Using qualitative thematic analysis, the derived themes were used to explain the conflict between beliefs and behaviors. Here, the mixed methods approach led to a detailed understanding by comparing the quantitative survey results with the qualitative observations to make the discrepancy between beliefs and behaviors more visible.

Choosing a mixed methods approach for your study does not automatically ensure detailed understanding any more than a mono-method approach prevents such an understanding ([Symonds and Gorard, 2010](#); [Sandelowski, 2014](#); [Flick, 2017](#); [Gobo et al., 2022](#)). To gain a detailed understanding, you will have to conduct your study sensibly. The ten strategies described here support sensible decision making in developing research that can deepen understanding through *integration* or, more precisely, *integration of claims*. As such, the ten strategies complement other advice in the mixed methods literature on integration ([Fetters et al., 2013](#); [Maxwell et al., 2015](#); [Åkerblad et al., 2021](#); [Hitchcock and Onwuegbuzie, 2022](#)); and they complement advice on strategies for other aspects of mixed methods research, such as research design ([Creamer, 2017](#); [Creswell and Plano Clark, 2018](#); [Tashakkori et al., 2021](#)), data analysis ([Bazeley, 2012, 2018](#); [Onwuegbuzie and Johnson, 2021](#)), and obtaining validity or trustworthiness ([Long, 2017](#); [Fabregues et al., 2021](#)). After briefly explaining the term *integration of claims*, the ten strategies will be described in greater detail.

Integration of claims

The process of developing a detailed understanding is referred to as *integration of claims*. In its most basic form, a claim is a statement about an entity that may be true or false. One of the claims in [Example 1](#) is *Tutors prefer learner-oriented beliefs*. This is a *simple* claim because it makes one statement (*prefer learner-oriented beliefs*) about one entity (*tutors*). Claims can also be *complex*, where there is more than one entity-statement combination, connected by a word that indicates their relationship. For instance, [Example 1](#) makes the complex claim that *Tutors prefer learner-oriented beliefs, but in their teacher behavior they show a more traditional approach*. This complex claim consists of two simple claims: *Tutors prefer learner-oriented beliefs* and *Tutors' approach to teaching is more traditional*. These two simple claims are connected by the word *but*, expressing the mismatch between tutors' preference for student-centered teaching and their teaching behavior.

Developing a detailed understanding can be viewed as a linguistic process, in which simple claims are developed and connected to form complex claims. These are in turn integrated into ever more complex claims until the most complex resulting claims can be said to provide a detailed understanding.

What is the role of methods in integration? In my view, methods are not integrated with claims, nor are they integrated among themselves. In other words, methods remain outside the "integration box" and are used to derive simple claims, as shown by the arrows from methods to simple claims in [Fig. 1](#). This differs from the view more widely held by mixed methods scholars that researchers integrate methods or data (e.g., [Bryman, 2006](#); [Fetters et al., 2013](#); [Maxwell et al., 2015](#)). I contend that the term *mixed methods* is misleading, as it is the claims rather than the methods that are mixed ([Cartwright and Hardie, 2012](#); [Cartwright, 2018](#)) or better, *integrated*.

In line with [Cartwright \(2018\)](#) and [Cartwright and Hardie \(2012\)](#), I use the word *claim* rather than *inference* here, although the latter is more common in the mixed methods literature ([Teddle and Tashakkori, 2009](#); [Creamer, 2017](#); [Tashakkori et al., 2021](#)). I do so on the grounds that *claim* and *inference*, although interchangeable in many contexts, emphasize different aspects of a statement and connote different things. For instance, characterizing the statement *Tutors prefer learner-oriented beliefs* as an inference implies that someone inferred this statement from something else—in this case, from an analysis of the survey data—so emphasizing the connection between this statement and the method *Using a survey* (see [Fig. 1](#)). By focusing on the relationships between claims, I seek to de-emphasize the connection between methods and claims, and referring to claims as *inferences* is not helpful in this regard. In addition, as we will see, researchers often introduce various claims that are not their own inferences but are borrowed from the literature, making it misleading to refer to these as *inferences*.

The word *claim* emphasizes the status of a statement as an independent entity, regardless of the methods used to derive it. This is important when considering the issue of validity, which is a property of claims and not of methods; even correctly applied methods can produce a claim that is false ([Maxwell, 2013](#)). *Claim* also has another connotation that *inference* tends to underplay: the possibility of being false. A Google search (January 9th, 2022) returned many more hits for *false/invalid claim* (26,700,000) than for *true/valid claim* (1,570,000), but this ratio was reversed for the far less frequent term *inference*; *true/valid inference* returned 68,200 hits while *false/invalid inference* returned 51,500 hits. By implication, *claim* is much more strongly associated than *inference* with *possibly false*.

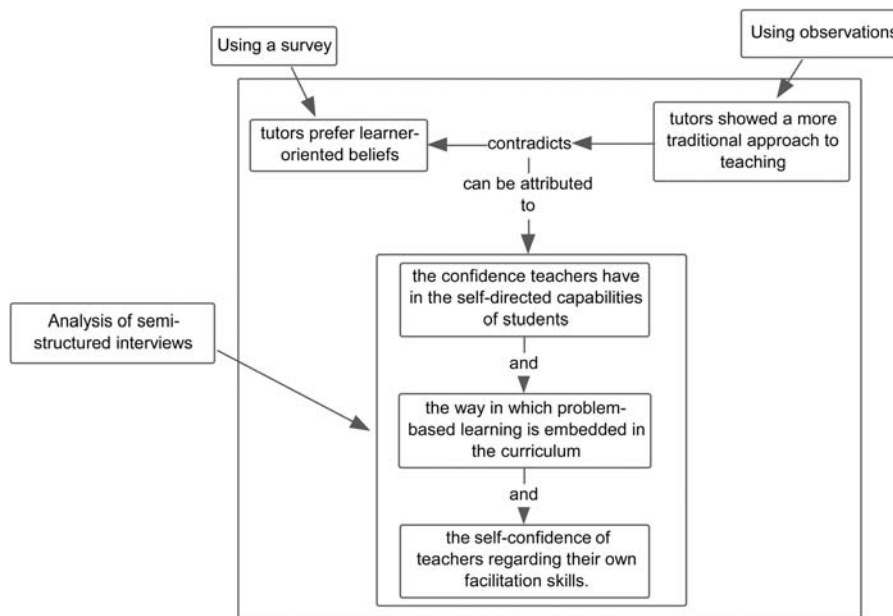


Fig. 1 Integration of claims in mixed methods research.

Taken together, these arguments point to different user expectations regarding the terms *claim* and *inference*. When I use the correct methods to *infer* something from my data, I do not expect the resulting inference to be false. In contrast, a derived claim—as an entity that has become independent from its methods—could possibly be false, given the strong connection between *claim* and *being false*. As mixed methods research pursues detailed understanding, it is important to be constantly reminded that our claims may be false, prompting further investigation. For that reason, I favor the term *claim* in the present context.

When researchers integrate claims in language, they do not necessarily “integrate” previously separate or “disintegrated” events or circumstances into one coherent whole. On the contrary, the integration underpinning the complex claim *Tutors prefer learner-oriented beliefs, but in their teacher behavior they showed a more traditional approach to teaching* in fact disintegrates a reality. Each of these simple claims describes a coherent reality, but together they create a disintegrated reality that contains an inherent contradiction (indicated by the word *but*). Thus, a linguistic integration can create either empirical coherence or empirical incoherence—an empirical “integration” or an empirical “diffraction” (Uprichard and Dawney, 2019). We could also characterize this process as *developing theory*: theory that consists of connected and integrated, simple claims and that stays close to the data from which it had been developed.

The integration process is summarized in Table 1 (Example 1).

As Table 1 shows, integration does not occur by combining simple claims into one complex claim all at once. Instead, the process of mixed methods research includes successive *integration points* (Guest, 2013); at each such point, two or more claims are brought together and linked by words like *but*, *and*, or *because*. In Table 1, at integration point I, the two simple claims [1] and [2] are brought together and connected by the word *but*, indicating that they are contradictory.

Each integration point is labeled to indicate its *integration type*; for example, integration point I is labeled *mixed methods* because it connects claims derived from quantitative surveys and qualitative observation. Integration point II is labeled *mono-method* because it integrates claims derived from one method only (semi-structured interviews). These different types of integration point serve to explain why Example 1 is a mixed methods study: because it contains at least one mixed methods integration point, in this case two (points I and III). In addition, these types explain why claim integration is not confined to mixed methods research (Symonds and Gorard, 2010; Sandelowski, 2014; Flick, 2017; Gobo et al., 2022). For example, even if the researchers in Example 1 had only conducted semi-structured interviews, they would still be able to develop and integrate claims using mono-method integration points like II. Thus, in a mono-method study too, claims can be developed and integrated to obtain a detailed understanding. In addition, claims can be developed and integrated using only different quantitative or only different qualitative methods (e.g., interviews and observations). Such integration is performed in *multi-method* integration points. We will meet two further types of integration point types later in the section *Consider various types of integration*.

The last and most complex integration point in a mixed methods study is the *meta-inference* (Teddlie and Tashakkori, 2009; Creamer, 2017; Tashakkori et al., 2021; Schoonenboom, 2022). *Meta* (Greek: beyond) refers to an inference “beyond” both quantitative and qualitative claims. The meta-inference, then, can be regarded as the study’s theory, when we reach a detailed understanding by connecting several claims in a way that makes sense. Often (though not always), the meta-inference resolves a contradiction by providing an explanation. In Example 1, there was a conflict between claims based on the survey and those based

Table 1 Integrating claims in Example 1.

Obtaining claims		Integrating claims	
Methods	Claims	Integrated claims	Connection
I. Mixed methods			
Survey	→ [1] Tutors prefer learner-oriented beliefs.	IC [3] Tutors prefer learner-oriented beliefs, but they show a more traditional approach to teaching.	Contradiction
Observations	→ [2] Tutors show a more traditional approach to teaching.		
II. Mono-method (qualitative)			
Semi-structured interviews	→ [4] Problem-based learning is embedded in the curriculum in a specific way.	IC [7] Problem-based learning is embedded in the curriculum in a specific way, and teachers have a certain amount of confidence in the self-directed capabilities of students and showed varying amounts of self-confidence regarding their own facilitation skills.	Juxtaposition
	→ [5] Teachers have a certain amount of confidence in the self-directed capabilities of students.		
	→ [6] Teachers showed varying amounts of self-confidence regarding their own facilitation skills.		
III. Mixed methods			
Survey & observations	→ IC [3] Tutors prefer learner-oriented beliefs, but they show a more traditional approach to teaching.	IC [8] Tutors prefer learner-oriented beliefs, but they show a more traditional approach to teaching because problem-based learning is embedded in the curriculum in a specific way and teachers have a certain amount of confidence in the self-directed capabilities of students and showed varying amounts of self-confidence regarding their own facilitation skills.	Explanation
Semi-structured interviews	→ IC [7] Problem-based learning is embedded in the curriculum in a specific way and teachers have a certain amount of confidence in the self-directed capabilities of students and showed varying amounts of self-confidence regarding their own facilitation skills.		

Note. IC = integrated claim.

Reprinted from Schoonenboom, J., 2022. Developing the meta-inference in mixed methods research through integration of claims. In: Hitchcock, J.H., Onwuegbuzie, A.J. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*, Routledge, p. 58, <https://doi.org/10.4324/9780429432828-6>. CC BY-NC-ND 4.0.

on the observations. In the meta-inference, the researchers resolved this contradiction by developing an explanation based on the claims from the interviews: that tutors' actions differed from their beliefs *because* of how their course was embedded in the curriculum and because of their estimation of their own skills and those of their students. In some mixed methods studies, however, a meta-inference creates a contradiction rather than resolving one (see Example 2 below).

Ten integration strategies for mixed methods research

Strategies for integrating claims form part of the mixed methods research process. Fig. 2 is based on a standard representation of that process (Ivankova et al., 2006; Morse and Niehaus, 2009; Teddlie and Tashakkori, 2009; Creswell and Plano Clark, 2018); (for an alternative representation, see Maxwell, 2013). According to this standard model, the study's goal and purpose guide all research activities, including data collection and analysis, deriving conclusions, and integrating qualitative and quantitative conclusions into a single meta-inference. There are three possible paths: collecting and analyzing qualitative and quantitative data at the same time (path a); collecting and analyzing quantitative data after qualitative data (path b); or collecting and analyzing qualitative data after quantitative data (path c).

The initial steps of the proposed alternative model (Fig. 3) are not dissimilar to the standard model of mixed methods research, in that the goal and research questions are used to plan the study. However, the proposed model is not confined to mixed methods research. A range of data sets is still collected and analyzed, but, as outlined in the Introduction, these data sets may include quantitative and qualitative data, qualitative data only, or quantitative data only, from which simple claims are derived. Fig. 3 represents only the first part of a study, ending at the first integration point; the result is not a meta-inference but a first integrated claim.

Beyond this first integrated claim, the research process continues to *Reflect on next step*, which is where the proposed model begins to deviate from the standard model (see Fig. 4). This reflection often creates backward loops, as the result typically involves taking a few backward steps in the research process. In Example 1, for instance, the researchers discovered a contradiction at integration point I (*Tutors prefer learner-oriented beliefs, but in their teacher behavior they show a more traditional approach*). The researchers' next step was to look for an explanation by re-analyzing the interviews; this step is indicated in Fig. 4 as a back arrow from *At each integration point, reflect on the next step* to *Search for explanation*, followed by integration of the contradiction and the explanation (path ex. 1).

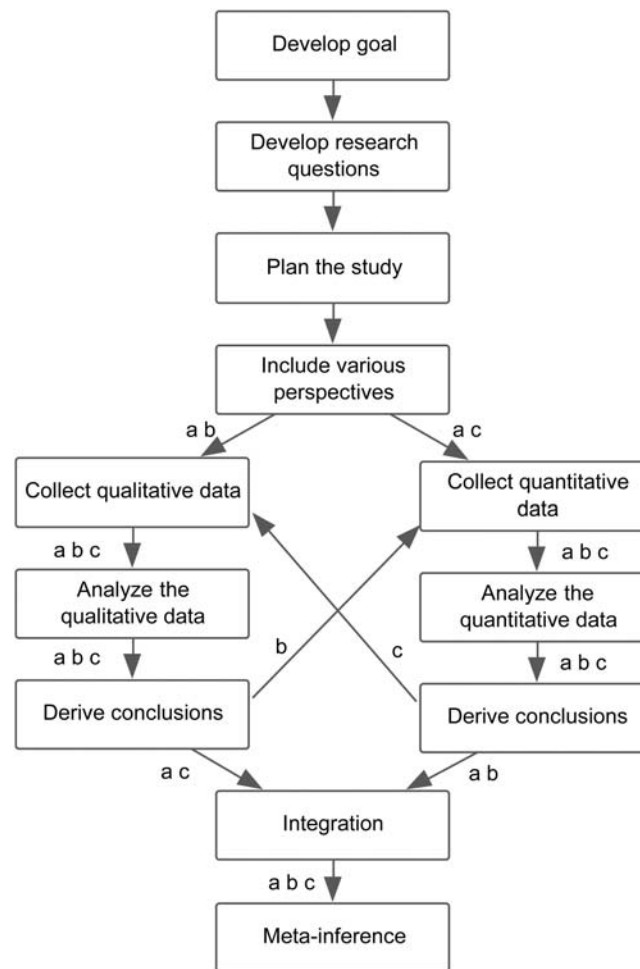


Fig. 2 A standard presentation of the mixed methods research process.

These backward loops have important implications for the process of mixed methods research. Mixed methods studies are guided by the research goal and research questions (Bryman, 2007; Ridenour and Newman, 2008; Teddlie and Tashakkori, 2009; Bazeley, 2018) and by each integration point. That is, the decision at each integration point may change the study's goal and research questions, and consequently the research plan. As shown in Fig. 4, *Adjust goal*, *Develop new research questions* and *Re-plan the study* are possible research activities, all steered by an integration point.

Fig. 4 also shows that integration points steer and initiate several integration strategies. For example, *Finding an explanation* becomes relevant as a next step only after *integrating claims* reveals a contradiction. Not all integration strategies assume previous integration (e.g., *Search for contradiction*), but other strategies typically arise after performing a first integration step (e.g., *Analyze subgroups*, *Search for an explanation*).

Of the 17 research activities or strategies in Fig. 4, this chapter discusses 10 (presented in *italics*). I selected these ten strategies because they relate directly to integration and because they seem most useful for developing a detailed understanding. Nine of these ten strategies relate to the integration of claims; the exception is *Integrate data sets where possible* (which obviously relates to data integration). Where relevant, I briefly discuss other research activities and strategies shown in Fig. 4.

Strategy 1: incorporate the perspectives of all relevant stakeholders

As claims can be integrated only if they exist, mixed methods researchers should include diverse perspectives on the target phenomenon at various levels of the research process (Greene, 2006; Mertens, 2007; Flick, 2017). The resulting data sets can be used to derive claims that may sometimes be contradictory. Elsewhere, I refer to this process as *diversification* (see Johnson (2017) and Johnson and Schoonenboom (2016) for extended lists of diversification strategies). Below, I discuss the two most important forms of diversification in mixed methods research: *including perspectives of various stakeholders* and *developing a range of data sets* (see Fig. 4). Example 2 illustrates the use of *including various stakeholder perspectives* in a mixed methods study.

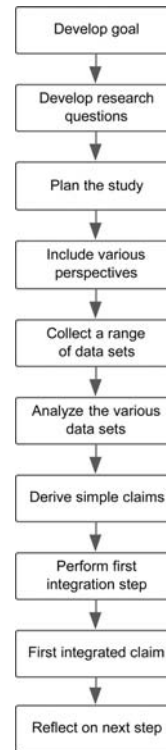


Fig. 3 Initial steps of proposed model of claim integration (to first integration point).

Example 2

A study by Clark and Moss (Clark, 2005, p. 14)

was a collaboration with Learning through Landscapes, a charity based in England that works to promote the use, development and maintenance of school playgrounds. [...] The study was based in a preschool for three- to four-year-olds. Twenty-eight children were involved, together with parents and practitioners. [...] The manager, practitioners and a group of parents wanted to take more advantage of the small outdoor space available to the preschool. This included a soft play surface, a small area of decking, a muddy bank and “boggy” ground where there was an underground spring.

According to Clark (2005, p. 16),

Observing the children revealed the house to be a key resource for them. The children confirmed this through their photographs, the tour, and their interviews. Parents also mentioned the house as an important space in the preschool. However, the interviews with practitioners showed that the house was a source of tension. They felt it was too small. The review with children, practitioners, and Learning through Landscapes recognised these opposing views and raised some possible solutions. The preschool has now turfed a new area for children to use to build their own temporary structures.

Here, the researchers collated the various claims of different stakeholders, including children, parents, practitioners, school principals, and Learning through Landscapes (the organization responsible for redeveloping the playground). Analysis of the interviews showed that some of these claims were contradictory, as practitioners differed from children and parents in their view of the playhouse, and this influenced the decision regarding further development of the playground. The study highlights the essential role of differing stakeholder perspectives (especially those of practitioners) in developing the playground in the best possible way.

Strategy 2: develop a range of data sets

Mixed methods researchers develop at least one quantitative and one qualitative data set, each providing a unique perspective on the target phenomenon. Researchers can build data sets by collecting primary data, using existing data sets, or transforming

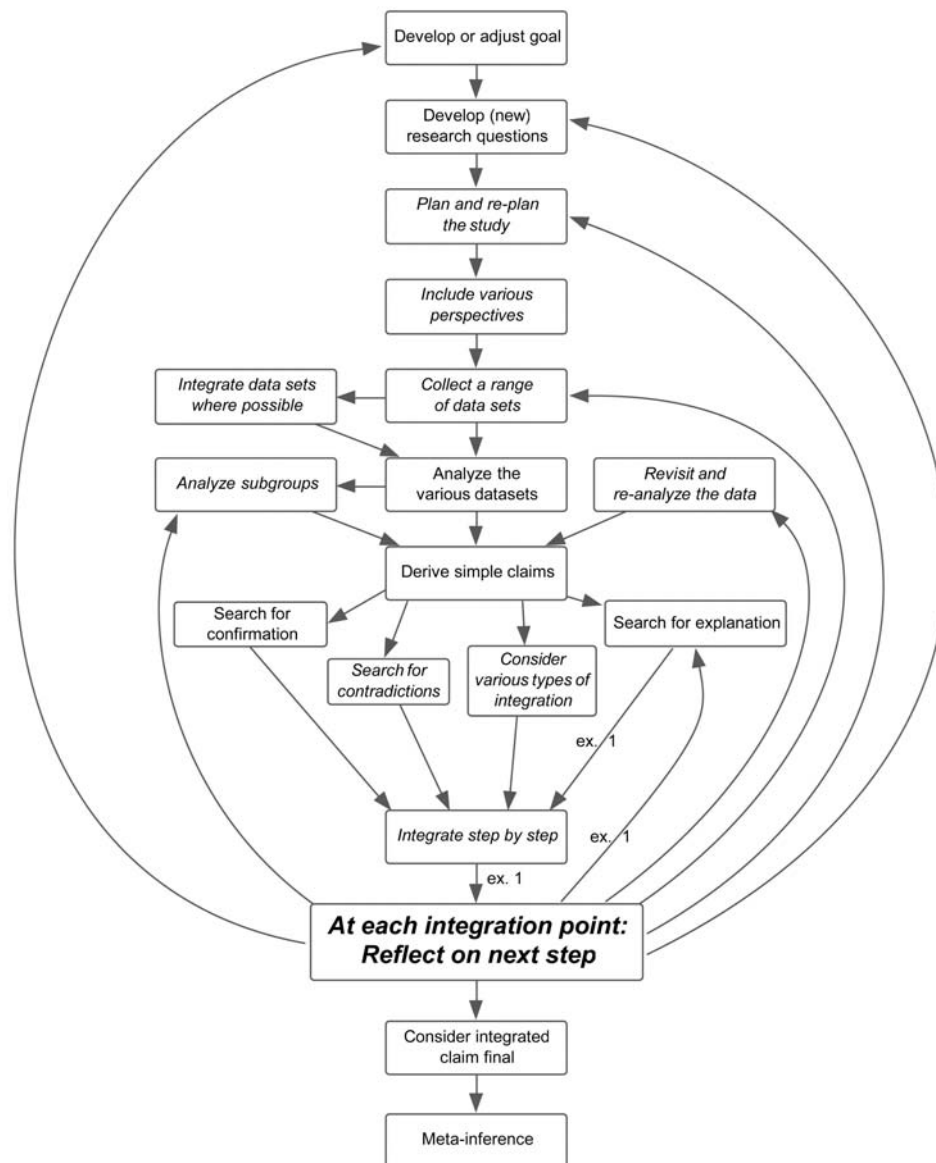


Fig. 4 Light bulb model of mixed methods integration with feedback loop. *Note.* **ex. 1** = **Example 1**. Strategies in *italics* are discussed in this article.

qualitative data into quantitative data in a process called *quantitizing* (Sandelowski et al., 2009; Bazeley, 2018; Nzabonimpa, 2018; Onwuegbuzie and Leech, 2019). The range of data sets in **Example 1** include a survey of tutors' beliefs about teaching, observations of tutors' behavior, and interviews. In **Example 2**, the data sets are even more varied, including photographs taken by the participating children, observations of children playing, reports of tours organized by children, and interviews with a range of stakeholders.

Strategy 3: integrate claims step by step

In a mixed methods study, claims are integrated step by step rather than all at once (De Allegri et al., 2020; Schoonenboom, 2022). In **Example 1** (see Table 1), separate claims were derived from a questionnaire, observations, and interviews. As a first integration step, the researchers identified a contradiction between claims based on the questionnaires and the observations: *Tutors preferred student-centered teaching, but their behavior was teacher-centered*. In a second step, they integrated the simple claims derived from the interviews. In a third and final integration step, they resolved the contradiction by deriving an explanation from the integrated interview claims.

Claims were also integrated step by step in [Example 2](#). The first step integrated claims from observations, children's photographs, tours, and interviews with children and parents confirming that the playhouse was an essential resource for the children. The second and final integration step identified a contradiction between these positive claims and the views of practitioners, who regarded the playhouse as a source of tension.

Strategy 4: consider various types of integration

Mixed methods researchers perform various types of integration; for instance, [Example 1](#) includes one mono-method and two mixed methods integration points (see [Table 1](#)). Another type is the multi-method integration point, which integrates claims derived from two qualitative methods or two quantitative methods. [Example 3](#) illustrates two further types of integration point.

Example 3

[Glewwe et al. \(2009\)](#) studied the effects of providing textbooks to primary school children in rural Kenya in schools where textbooks had not previously been used. The study started by comparing children's subject test scores in 50 intervention schools and 50 control schools. Next, in each intervention class, one child with a median score was asked to read aloud from their textbook and to answer some questions. Finally, the test scores of children with high pretest scores and other students were separately analyzed to determine whether providing textbooks had any identifiable effect:

A randomized evaluation in rural Kenya finds, contrary to the previous literature, that providing textbooks did not raise average test scores. Textbooks did increase the scores of the best students (those with high pretest scores) but had little effect on other students. Textbooks are written in English, most students' third language, and many students could not use them effectively.

[Glewwe et al. \(2009, p. 112\)](#).

[Example 3](#) is again a mixed methods study. In this case, researchers developed and integrated claims from test scores (quantitative data) and observations (qualitative data), but they also developed two integrated claims without referring to empirical data. Instead, they found that a claim from previous research (that *providing textbooks affects test scores*) contradicted their own finding (that *providing textbooks does not affect test scores*). This approach can be characterized as *Integration with previous research*.

Finally, the researchers in [Example 3](#) developed a claim on the basis of reasoning only—that is, without empirical verification—that *students with the highest pretest scores were able to read their textbooks*. Their reasoning was as follows. If providing textbooks had an effect among students with the highest pretest scores, these students must have been able to read their textbooks, which could not otherwise have had any effect, as the results for students with lower pretest scores suggested. Integration of the theoretically derived claim (that *students with the highest pretest scores were able to read their textbooks*) and the empirically derived claims about the other students can be characterized as *Integration with theoretically derived claims*.

Strategy 5: use each integration point to reflect on the next step

The primary strategy of mixed methods research can be described as follows: *Mixed methods researchers should reflect on their next step at each integration point*. In [Example 1](#), at integration point 1, the researchers identified a contradiction and decided to search their interviews for an explanation. In the absence of any such contradiction, they would probably have adopted a different perspective when analyzing the interviews.

In [Example 2](#), the crucial integration point is the point at which the positive opinions of children and parents are integrated with the practitioners' critique. Here again, a contradiction emerges. Instead of further claim integration, this contradiction steered the measures taken to develop the playground. To address the issue of playhouse overpopulation, the designers created a new area where the children could build their own temporary structures.

In [Example 3](#), a contradiction arose when integrating the claim from the quantitative study (*Providing textbooks had no effect*) with the claim from previous research (*Providing textbooks has an effect*). Reflecting on this contradiction prompted the researchers to search for an explanation by observing children, and they found that *most children were unable to read their textbooks*. Having established that textbooks had an effect on students with higher pretest scores, a second contradiction arose. In this instance, the researchers used reflection to arrive at an explanation: that *students with high pretest scores could read their textbooks*.

Strategy 6: integrate qualitative and quantitative data into a single data set

This strategy involves integrating data rather than claims. While qualitative and quantitative data can be visually integrated by presenting them side by side in a *joint display* ([Guetterman et al., 2015a,b](#); [Plano Clark and Sanders, 2015](#)), Strategy 6 goes one step further by bringing qualitative and quantitative data together in a new integrated data set. It is important to note that the mixed methods literature often uses the term "data integration" to refer to the use of different data sets in a mixed methods study ("integration within a study") without integration of the data sets themselves ([Irwin, 2008](#); [Fetters et al., 2013](#); [Maxwell et al., 2015](#)). In

Table 2 Reconstructed case comparison table of the data in Lee and Greene (2007).

ID	C	GPA	Quotation
0607	4	4.00	I do not have any language problems.
2020	3	4.00	I understand 80% of the lectures. Careful reading complements 20% of lack of understanding.
1315	3	4.00	My biggest problem is related with speaking in English. I will get a good grade because my mathematical background is strong.
0609	3	4.00	The professor is old so that his pronunciation is not clear. He usually handed out important contents and I could understand it by reading the textbook.
0620	2	4.00	I understand only 60–70% of the lectures. It has made my scores less than my expectation.
1310	3	3.89	I want to participate and argue the subject. But I cannot find myself enough. This does not affect my grade.
0624	2	3.80	It is easy to understand the lectures and participate in class discussions. The instructor speaks slowly.
2036	3	3.57	The major problem is speaking. I spoke once or twice during the whole semester.
0610	3	3.53	Listening is a problem. Lack of cultural knowledge interferes with understanding the concept.
0605	2	3.50	Because of my poor listening, I am struggling with catching up with my content courses.
0603	3	3.39	I still have some problem in speaking. This difficulty doesn't affect my ability to do well in all the courses I take.
2037	2	3.22	Lack of knowledge about idiomatic expressions prevents me from understanding questions on the homework assignment.
2025	3	3.18	I was not able to finish all the reading assignments every week.
2031	3	3.11	I have problems with my hearing and how to speak correctly.
0608	4	2.89	I understand almost 100% of the lectures. I understand the professor completely. [...] grades are not important. That is the least of my worries. [...]

ID = student ID; C = CEEPT score; GPA = student GPA; Quotation = utterance by student.

Adapted with permission from Schoonenboom, J., Johnson, R.B., 2021. The case comparison table: a joint display for constructing and sorting simple tables as mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*, Routledge, p. 278, <https://doi.org/10.4324/9780203729434-24>. CC BY-NC-ND 4.0.

contrast, the term *data integration* will be used here only when two data sets are integrated into one data set that contains both qualitative and quantitative data.

An integrated data set can be represented as a *case comparison table* (Schoonenboom and Johnson, 2021), as in **Example 4**.

Example 4

In a study of international graduate students at a large public university in the United States, Lee and Greene (2007) “investigated the relationships between graduate students’ placement test scores for English as a second language (ESL) and three measures of academic performance (grade point average [GPA], faculty evaluations, and student self-assessments)” (Lee and Greene, 2007, p. 366). In addition to Computerized Enhanced ESL Placement Test (CEEPT) scores and students’ GPA, qualitative data were obtained by interviewing 20 students. I re-analyzed the available data from the published article, and the results are presented below (Table 2) as a case comparison table (Schoonenboom and Johnson, 2021).

In a case comparison table, researchers use quantitative data to group and sort qualitative quotations until patterns begin to emerge. Table 2 includes qualitative and quantitative data for each student. The CEEPT and GPA scores are quantitative data, enabling us to count occurrences of individual values and to calculate means and standard deviations. The quotations are qualitative data; it would not make sense to count the number of occurrences of each quotation (which would be 1).

Table 2 shows that some students with an imperfect command of the English language nevertheless attained the highest possible GPA score. The four students with a CEEPT score of less than 4 and a GPA score of 4 are shown in bold. In their quotations, three of these students refer to strategies that compensate for their language problems: careful reading (ID 2020), a strong mathematical background (ID 1315), and reading the textbook (ID 0609). As none of these compensation strategies was mentioned by any of the less successful students, this formed the basis for my meta-inference and preliminary hypothesis: *Language problems affected students’ grade point average at one US university, with the exception of students who developed compensation strategies*.

Strategy 7: search for contradictions

Mixed methods researchers should search actively for contradictions (Greene, 2006; Johnson and Schoonenboom, 2016; Uprichard and Dawney, 2019) throughout the research process by comparing the claims derived from different data sets (e.g., **Example 1**: *Tutors prefer learner-oriented beliefs, but their approach to teaching is more traditional*). Researchers should also look for contradictions between study outcomes and previous research (e.g., **Example 3**: *Providing textbooks increased quantitative test scores elsewhere but not in Kenya*). A study may include more than one contradiction (e.g., **Example 3**: *Providing textbooks increased quantitative test scores among students with the highest pretest scores but not for others*). A second contradiction is often a refinement of the first (in this case, *Providing textbooks did not increase test scores*).

Where a contradiction emerges, researchers can deploy various strategies to resolve it by looking for an explanation (see **Fig. 4**)—for instance, by referring to the curriculum and teachers' estimations of their own capabilities and those of their students (**Example 1**); to differences in reading ability between students with high pretest scores and others (**Example 3**); and to differences between students who use language compensation strategies and those who do not (**Example 4**).

A contradiction can also be the outcome rather than the driver of a mixed methods study (Uprichard and Dawney, 2019), where the contradiction is not resolved but is used in a productive way. In **Example 2**, for instance, children and parents valued the playhouse, but practitioners felt it was a source of tension because it was too small. As the researchers did not explain this difference of opinion, the contradiction was unresolved. Instead, the designers used it to make a practical decision: "The review with children, practitioners, and Learning through Landscapes recognized these opposing views and raised some possible solutions. The preschool has now turfed a new area for children to use to build their own temporary structures" (Clark, 2005, p. 16).

Finally, the proposed light bulb model of mixed methods integration includes a strategy that complements the search for contradictions: *Searching for confirmation* (see **Fig. 4**), which was once considered the most important reason for conducting multi-method or mixed methods research (Campbell and Fiske, 1959; Denzin, 1978; Flick, 2017). For instance, in **Example 2**, the interviews with parents confirmed the children's view that the playhouse is an important resource for them. Today, searching for contradictions is considered more important (Greene, 2015; Johnson and Schoonenboom, 2016; Johnson, 2017), as this can lead to new insights (as the examples show).

Strategy 8: revisit and re-analyze the data

Mixed methods research is a step-by-step process. Researchers step forward by integrating and developing their claims, but they also step back. In a mixed methods study, researchers return to data already collected and analyzed and re-analyze those data to deepen understanding. By repeating this strategy wherever possible, researchers obtain an understanding that is as detailed as possible, given the data they had collected (Creswell and Plano Clark, 2018, Chapter 9).

In all four examples (most notably Examples 3 and 4), the researchers re-analyzed their data. In **Example 3**, analysis of the quantitative data had previously indicated that providing textbooks did not affect quantitative test scores because most children could not read their textbooks. However, the re-analysis revealed two distinct effects for two different groups; unlike most children, those with high pretest scores benefited from the textbooks because they could read them. In **Example 4**, I re-analyzed Lee and Greene's qualitative and quantitative data and used a case comparison table to develop a tentative hypothesis: *Some students with an imperfect command of the English language attained the highest possible GPA because they used compensation strategies*.

Strategy 9: analyze subgroups

Mixed methods researchers should search for subgroups and seek to identify any contradictions between them. Subgroup analysis often, but not always, involves re-analyzing parts of a data set previously treated as a single whole, as a special form of Strategy 8 (*Re-analyze the data*). However, subgroup analysis warrants discussion in its own right as an important integration strategy in mixed methods research.

In Examples 3 and 4, the researchers used this approach to explain the differences between subgroups. In **Example 3**, there were two subgroups: students with high pretest scores and all other students. In **Example 4**, the subgroups were students with language problems who attained the highest possible GPA and students with language problems and lower GPA scores. In quantitative research, this form of explanation involves a *moderating factor*, which prevents an adverse effect from occurring or facilitates a positive effect (Hayes, 2017). In **Example 3**, reading ability is a moderating factor because it explains why textbooks affect quantitative test scores. In **Example 4** the moderating factor is compensating strategies that mitigate the negative influence of language problems on GPA. In summary, mixed methods researchers should analyze subgroups and look for differences and moderating factors that might explain these differences.

Strategy 10: plan and re-plan the study

In the light bulb model, the steering effect of integration points has important implications for mixed methods design. The first steps of a mixed methods study can be planned using the standard model: developing an overall goal and an initial research question, collecting and analyzing a range of data sets, and deriving simple claims (see **Fig. 3**). After the first integration step, however, integration points begin to shape how the study should proceed. This means that researchers cannot plan all of their research activities at

the outset because the outcome of the first integration step cannot be known in advance. As the integration points guide further action and prompt re-analysis of the data or collection of additional data, initial planning is necessarily followed by re-planning at various points throughout a mixed methods study.

This re-planning is apparent in the examples presented here. In **Example 4**, re-analysis of the data in [Lee and Greene \(2007\)](#) was planned long after the authors had completed the original study. In **Example 1**, the researchers based their interviews on questionnaire data ([Assen et al., 2016](#)), designing the interview guidelines after the questionnaire data highlighted a contradiction: *Tutors prefer learner-oriented beliefs, but in their teacher behavior they show a more traditional approach*. This contradiction prompted the interviewers to search for an explanation. In **Example 3**, [Glewwe et al. \(2009\)](#) conducted classroom observations after identifying a first contradiction: *Providing textbooks raises test scores of primary school pupils elsewhere but not in rural Kenya*. This finding shaped the design of their observations, which sought to explain why providing textbooks appeared to have no effect.

Integration points do not fully determine the course of a study, as researchers can to some extent preplan what they will do after the first integration point. In the studies referred to here, researchers may have planned to conduct interviews (**Example 1**) or observations (**Example 4**) at the outset. However, the direction of these interviews and observations could not be planned until the researchers had performed their first integration step, which in this case revealed a contradiction. They may instead have decided to conduct the interviews and observations only after discovering a contradiction, although this is not clear from the published accounts. In any event, the researchers in Examples 1 and 4 could not continue *exactly* as planned after their first integration point because they had to explain a contradiction that was unknown at the beginning of the study. As a result, both studies necessarily involved feedback loops (see [Fig. 4](#)).

Conclusion

To gain a detailed understanding of the target phenomenon, mixed methods researchers proceed step by step, deriving claims from quantitative and qualitative data and then integrating those claims. In this chapter, I have highlighted ten strategies for integrating claims and data in mixed methods studies.

- Strategy 1: Include the perspectives of all relevant stakeholders
- Strategy 2: Develop a range of data sets
- Strategy 3: Integrate claims step by step
- Strategy 4: Consider various types of integration
- Strategy 5: Use each integration point to reflect on the next step
- Strategy 6: Integrate qualitative and quantitative data into a single data set
- Strategy 7: Search for contradictions
- Strategy 8: Revisit and re-analyze the data
- Strategy 9: Analyze subgroups
- Strategy 10: Plan and re-plan the study

According to the light bulb model of mixed methods integration ([Fig. 4](#)), these strategies are not mere guidelines or options but together optimize and deepen understanding by integrating different phases of the research process: conceptualization (*Include various perspectives*); data collection (*Develop a range a range of data sets; Integrate data sets where possible*); data analysis (*Analyze subgroups; Revisit and re-analyze the data*); integration (*Integrate claims step by step; Search for contradictions; Consider various types of integration*); and planning (*Reflect on the next step; Plan and re-plan the study*).

More importantly, these strategies can be used together because they build on each other. For example, integrated data sets can be used to perform subgroup analysis. If subgroup analysis reveals a contradiction, an explanation may be found by re-analyzing existing data or collecting new data. The process as a whole encompasses five integration strategies and may constitute only a small part of a larger study, which may in turn involve the use of many additional integration strategies.

We have known for some time that mixed methods research is emergent ([Poth, 2018; De Allegri et al., 2020](#)); in other words, the meta-inference is not developed at the end by incorporating and connecting the derived claims in a single move. The light bulb model fully acknowledges this emergent quality by incorporating points of integration that guide the research process. In practice, this means that one cannot and should not wait until data collection and analysis are complete before starting to develop the meta-inference.

A further implication for practice is that preplanning of a mixed methods study is necessarily limited because there is no way of anticipating where the first integration point will lead. It is important to plan your study as far as possible, and here the standard model ([Fig. 3](#)) can be helpful. It is, however, equally important to accept that you can never proceed exactly as planned, as you cannot know in advance what your findings will be or what claims will emerge, and for that reason, you cannot predict the outcome of your first integration step. This integrated claim will guide the next step of the research process, and any emergent contradictions can only be identified and resolved after integrating your claims.

References

- Åkerblad, L., Seppänen-Järvelä, R., Haapakoski, K., 2021. Integrative strategies in mixed methods research. *J. Mix. Methods Res.* 15, 152–170.
- Assen, J.H.E., Meijers, F., Otting, H., Poell, R.F., 2016. Explaining discrepancies between teacher beliefs and teacher interventions in a problem-based learning environment: a mixed methods study. *Teach. Educ.* 60, 12–23.
- Bazeley, P., 2012. Integrative analysis strategies for mixed data sources. *Am. Behav. Sci.* 56, 814–828.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage, Los Angeles, CA.
- Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Res.* 6, 97–113.
- Bryman, A., 2007. The research question in social research: what is its role? *Int. J. Soc. Res. Methodol.* 10, 5–20.
- Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105.
- Cartwright, N., Hardie, J., 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford University Press, New York, NY.
- Cartwright, N., 2018. First mix the claims; then mix the methods. Keynote presentation at the Third Global Conference of the Mixed Methods International Research Association. Vienna, Austria.
- Clark, A., 2005. Ways of seeing: using the Mosaic approach to listen to young children's perspectives. In: Clark, A., Kjørholt, A.T., Moss, P. (Eds.), *Beyond Listening: Children's Perspectives on Early Childhood Services*. Policy Press, Bristol, UK, pp. 29–49.
- Creamer, E.G., 2017. *An Introduction to Fully Integrated Mixed Methods Research*. Sage, Los Angeles, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*. Sage, Los Angeles, CA.
- De Allegri, M., Brenner, S., Kambala, C., Mazalale, J., Muula, A.S., Chinkhumba, J., Wilhelm, D., Lohmann, J., 2020. Exploiting the emergent nature of mixed methods designs: insights from a mixed methods impact evaluation in Malawi. *Health Pol. Plann.* 35, 102–106.
- Denzin, N.K., 1978. *The Research Act: A Theoretical Introduction to Sociological Methods*. McGraw-Hill, New York, NY.
- Fàbregues, S., Molina-Azorin, J.F., Fettes, M.D., 2021. Virtual special issue on "quality in mixed methods research". *J. Mix. Methods Res.* 15, 146–151.
- Fettes, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48, 2134–2156.
- Flick, U., 2017. Mantras and myths: the disenchantment of mixed-methods research and revisiting triangulation as a perspective. *Qual. Inq.* 23, 46–57.
- Glewwe, P., Kremer, M., Moulin, S., 2009. Many children left behind? Textbooks and test scores in Kenya. *Am. Econ. J. Appl. Econ.* 1 (1), 112–135.
- Gobo, G., Fielding, N.G., La Rocca, G., Van Der Vaart, W., 2022. *Merged Methods : A Rationale for Full Integration*. Sage, Los Angeles, CA.
- Greene, J.C., 2006. Toward a methodology of mixed methods social inquiry. *Res. Sch.* 13 (1), 93–98.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods research merger. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 606–615.
- Guest, G., 2013. Describing mixed methods research: an alternative to typologies. *J. Mix. Methods Res.* 7, 141–151.
- Guetterman, T., Creswell, J.W., Kuckartz, U., 2015a. Using joint displays and MAXQDA software to represent the results of mixed methods research. In: McCrudden, M.T., Schraw, G., Buckendahl, C.W. (Eds.), *Use of Visual Displays in Research and Testing*. Information Age Publishing, Charlotte, NC, pp. 145–175.
- Guetterman, T.C., Fettes, M.D., Creswell, J.W., 2015b. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13, 554–561.
- Hayes, A.F., 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford, New York, NY.
- Hitchcock, J.H., Onwuegbuzie, A.J. (Eds.), 2022. *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge, New York, NY.
- Irwin, S., 2008. Data analysis and interpretation. In: Hesse-Biber, S.N., Leavy, P. (Eds.), *Handbook of Emergent Methods*. Guilford, New York, NY, pp. 415–436.
- Ivankova, N.V., Creswell, J.W., Stick, S.L., 2006. Using mixed-methods sequential explanatory design: from theory to practice. *Field Methods* 18, 3–20.
- Johnson, R.B., Schoonenboom, J., 2016. Adding qualitative and mixed methods research to health intervention studies: interacting with differences. *Qual. Health Res.* 26, 587–602.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11, 156–173.
- Lee, Y.-J., Greene, J., 2007. The predictive validity of an ESL placement test: a mixed methods approach. *J. Mix. Methods Res.* 1, 366–389.
- Long, H., 2017. Validity in mixed methods research in education: the application of Habermas' critical theory. *Int. J. Res. Method Educ.* 40, 201–213.
- Maxwell, J.A., Chmiel, M., Rogers, S.E., 2015. Designing integration in multi-method and mixed methods research. In: Hesse-Biber, S., Johnson, B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research*. Oxford University Press, New York, NY, pp. 223–239.
- Maxwell, J.A., 2013. *Qualitative Research Design: An Interactive Approach*. Sage, Los Angeles, CA.
- Mertens, D.M., 2007. Transformative paradigm: mixed methods and social justice. *J. Mix. Methods Res.* 1, 212–225.
- Morse, J.M., Niehaus, L., 2009. *Mixed Method Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Nzabonimpa, J.P., 2018. Quantitizing and qualitzing (im-)possibilities in mixed methods research. *Methodol. Innov.* 11.
- Onwuegbuzie, A.J., Johnson, R.B. (Eds.), 2021. *The Routledge Reviewer's Guide for Mixed Methods Research Analysis*. Routledge, New York, NY.
- Onwuegbuzie, A.J., Leech, N.L., 2019. On qualitzing. *Int. J. Mult. Res. Approaches* 11, 98–131.
- Pawson, R., Tilley, N., 1997. *Realistic Evaluation*. Sage, London, UK.
- Plano Clark, V.L., Sanders, K., 2015. The use of visual displays in mixed methods: research strategies for effectively integrating the quantitative and qualitative components of a study. In: McCrudden, M.T., Schraw, G., Buckendahl, C.W. (Eds.), *Use of Visual Displays in Research and Testing*. Information Age Publishing, Charlotte, NC, pp. 177–206.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. Sage, Los Angeles.
- Ridenour, C.S., Newman, I., 2008. *Mixed Methods Research: Exploring the Interactive Continuum*. Southern Illinois University Press, Carbondale, IL.
- Sandelowski, M., Voils, C.I., Knafl, G., 2009. On quantitizing. *J. Mix. Methods Res.* 3, 208–222.
- Sandelowski, M., 2014. Unmixing mixed-methods research. *Res. Nurs. Health* 37, 3–8.
- Schoonenboom, J., Johnson, R.B., 2021. The case comparison table: a joint display for constructing and sorting simple tables as mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide for Mixed Methods Research Analysis*. Routledge, New York, NY, pp. 277–287.
- Schoonenboom, J., 2019. Develop your case! How controversial cases, subcases, and moderated cases can guide you through mixed methods data analysis. *Front. Psychol.* 10 (1369).
- Schoonenboom, J., 2022. Developing the meta-inference in mixed methods research through integration of claims. In: Hitchcock, J.H., Onwuegbuzie, A.J. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge, New York, NY, pp. 55–70.
- Symonds, J.E., Gorard, S., 2010. Death of mixed methods? Or the rebirth of research as a craft. *Eval. Res. Educ.* 23, 121–136.
- Tashakkori, A.M., Johnson, R.B., Teddlie, C.B., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Los Angeles, CA.
- Teddlie, C.B., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Los Angeles, CA.
- Uprichard, E., Dawney, L., 2019. Data diffraction: challenging data integration in mixed methods research. *J. Mix. Methods Res.* 13, 19–32.

Convergence and divergence in mixed methods research

Pierre Pluye^a and Quan Nha Hong^b, ^a Department of Family Medicine, McGill University, Montreal, QC, Canada; and ^b School of Rehabilitation, Université de Montréal, Montreal, QC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	462
Convergence and divergence of scientific worldviews	463
Problem	463
Common worldviews and their combinations	463
Three combinations without integration of worldviews	464
Three combinations with integration of worldviews	465
Proposed solution: recommendations for combining worldviews	465
Convergence and divergence in mixed studies reviews	466
Literature reviews of qualitative, quantitative and mixed methods studies	466
Synthesis designs in mixed studies reviews	466
Convergence and divergence in synthesis of QUAL and QUAN evidence	467
Convergence and divergence in primary MM research	468
Convergence	468
Divergence	472
Definition and examples	472
Three strategies for handling divergence	472
Conclusion	475
References	475

Introduction

In this chapter, convergence and divergence in mixed methods (MM) research derives from the integration of qualitative (QUAL) and quantitative (QUAN) methods. They can be made explicit (guided by a cognitive representation of the outcomes to be created) and reported in scientific publications. In accordance with [Greene \(2007\)](#), the authors “acknowledge and respect the value of divergence as generative of unanticipated insights and understandings” in MM research (p. 152). Convergence and divergence presented in this chapter refer to the integration of QUAL and QUAN methods at three dimensions: scientific paradigms and worldviews, mixed studies reviews, and data collection and analysis in primary research.

In the MM literature, seven dimensions of integration are commonly mentioned: (a) scientific paradigms, e.g., MM research team members’ scientific worldviews can be different and mixed; (b) literature reviews, e.g., mixed studies reviews include qualitative, quantitative and MM studies; (c) theories or conceptual frameworks; (d) methodologies; (e) research questions or objectives; (f) designs; and (g) data collection/analysis techniques ([Creswell and Plano Clark, 2018](#); [Pluye and Hong, 2014](#); [Tashakkori and Teddlie, 2010](#); [Johnson et al., 2007](#); [Fetters, 2020](#); [Pluye et al., 2022](#); [Hong et al., 2017](#)). MM must, at a minimum, integrate QUAL and QUAN methods through the research questions, the research design, and the data collection/analysis. Stated otherwise, the quality of MM studies that integrate QUAL and QUAN methods at all six dimensions are not necessarily better than those that integrate QUAL and QUAN methods involving at least three dimensions (question, design, and data collection/analysis).

MM integrate QUAL and QUAN methods in both empirical research and program evaluation ([Creswell and Plano Clark, 2018](#); [Tashakkori and Teddlie, 2010](#)). For example, MM allows for the development of practice-specific knowledge and training for medical education researchers, graduate students, and professionals ([Schifferdecker and Reed, 2009](#)). MM help to answer questions about the effectiveness of interventions and programs, the processes behind them, and the stakeholders’ experiences and perspectives on phenomena under study ([Pluye and Hong, 2014](#)). In this chapter, MM empirical research meets three conditions: (a) at least one QUAL method and one QUAN method are combined; (b) each method is used in a rigorous manner according to generally accepted criteria of the methodology or research tradition invoked; and (c) the methods are integrated at least through research question(s)/objectives(s), the MM design (be it planned or emergent), and the evidence ([Pluye and Hong, 2014](#); [Johnson et al., 2007](#); [Fetters, 2020](#); [Creswell and Plano Clark, 2018](#); [Tashakkori and Teddlie, 2010](#)). As mentioned by [Ivankova \(2015\)](#): “Stated simply, there is no true MM study without methods integration or mixing” (p. 21).

The purpose of this chapter is to provide students, teachers and researchers in education and related areas an analytical argument on the concepts of convergence and divergence in MM research. These concepts will be defined and illustrated using MM reference textbooks and articles, and MM education-related empirical studies. In the following section, they will be discussed in three dimensions: MM researchers’ scientific worldviews, mixed studies reviews, and MM primary research. This order has been chosen because worldviews influence reviews and primary research, and reviews influence primary research.

Convergence and divergence of scientific worldviews

MM research casts new light on the use of researchers' differing personal worldviews because researchers may experience worldview-related interpersonal tensions when working in MM teams (Mertens et al., 2016). While some tensions can be creative and constructive, worldview stereotypes and stigmatization can be very disruptive, especially for students (e.g., in thesis committee meetings). These challenges are rarely reported in MM publications, and few MM reference texts address how to handle worldview-related tensions and how to combine worldviews. The scarcity of solutions testifies to the need to provide comprehensive guidance for how to combine worldviews in MM research when doing so is constructive and desirable.

Thus, the purpose of this section is to propose a comprehensive understanding of the combinations of worldviews in MM research, which represent the convergence and divergence of MM research teams' worldviews. These combinations suggest recommendations to prevent and manage worldview-related tensions, and a decision tree for students to combine worldviews in MM projects. These recommendations can be especially useful in education where students and researchers use a variety of QUAL and QUAN methods, and adhere to a variety of worldviews.

Problem

In this chapter, worldviews are defined as individual sets of philosophical assumptions or beliefs that guide actions. Thus, combining worldviews means supporting the perspective or mental model of each member of MM research team or thesis committee (mental models used to read and interpret MM protocols, publications and studies). In contrast to the community-level concept of paradigm, which reflects the shared commitment of a community of scientists to a set of logics, models and exemplars (Kuhn, 1983; Hacking, 1981), the individual-level term worldview indicates that the focus is on one person's view of the world and sciences, which usually mirrors paradigms. As stated by Denzin (2010), "paradigms are human constructions, and define the shifting worldview of the researcher-as-bricoleur" (p. 421). When working in MM teams, each researcher's worldview reflects their preferred paradigm (consciously or unconsciously).

Multiple MM scholars claimed the pluralism of paradigms and related worldviews in MM research (Ghiara, 2020; Niglas, 2010; Teddlie and Tashakkori, 2003), and mentioned that team members' worldviews can be combined in a MM empirical studies (Ghiara, 2020; Wenger-Trayner et al., 2019; Harvey et al., 2020; Fàbregues et al., 2021). Typically, scientific paradigms and researchers' worldviews are underreported in MM studies (Greene, 2007; Coates, 2021; Alise and Teddlie, 2010), but that does not mean that they are ignored. It could simply mean that other things are considered more critical to include within the allotted word limit of scientific publications. In addition, some MM students and researchers may be unsure about their personal worldviews as one's worldview can evolve over time, especially during one's doctoral training (Baumard, 1997), and silence their uncertainties.

However, little MM reference texts address how to handle worldview-related tensions and combine worldviews beyond the dialectical approach and pragmatism (Johnson, 2017; Hesse-Biber, 2015). The absence of guidance on how to select a combination of worldviews is problematic (Pascal et al., 2018). Typically, reference texts provide a detailed example of how to perform only one combination (Ghiara, 2020; Harvey et al., 2020; Martel et al., 2021; McChesney and Aldridge, 2019; Schoonenboom, 2019; Wenger-Trayner et al., 2019) such as combining pragmatism and constructivism in a dialectical manner (performative paradigm) (Schoonenboom, 2019), whereas multiple combinations are possible.

Common worldviews and their combinations

In previous work (Pluye et al., 2022), a literature review identified and described five prevailing scientific contemporary worldviews in MM research: postpositivism, social constructivism, pragmatism, critical theory, and critical realism. Then, six possible combinations of worldviews were determined (a-paradigmatic, substantive theory, single worldview, multiple worldviews, complementary strengths, and dialectical pluralism). This does not imply that using different worldviews in the same study ensures their combination, nor that integration of worldviews is necessary, desirable, or better than no integration. It depends on the context of the study. Furthermore, one's worldview does not equate to one's choice of QUAL or QUAN or mixed methods; worldviews can influence methods, but they do not determine them. Stated otherwise, any method can be informed by one or more than one of contemporary paradigms or worldviews (Greene, 2007; Maxwell and Mittapalli, 2010). A table summarizing the five most common contemporary worldviews has been built for producing a "hand out" easy to read and retain for students (Table 1).

In mathematics, a combination is composed of one or more elements (each element being a combination of sub-elements, and so forth). Thus, a combination of worldviews is comprised of one or more worldviews, and a single worldview at the MM research team level can always be divided into variants. In other words, MM team members can apparently share the same worldview whereas they have different variants of this worldview. In previous work (Pluye et al., 2022), a literature review identified and described six combinations of worldviews at the MM research team level: three without integration of worldviews (a-paradigmatic, substantive theory, and single worldview) and three with integration (multiple worldviews, complementary strengths, and dialectical pluralism). These combinations are described below in terms of convergence and divergence of worldviews.

Table 1 The five most common contemporary worldviews in mixed methods research.

<i>Worldview</i>	<i>Ontology</i>	<i>Epistemology</i>	<i>Research purpose</i>	<i>Methodology</i>
Postpositivism	Evolutionary realism	Truthful scientific knowledge (but no truth)	Mainly to confirm hypotheses and generalize results of previous work, but sometimes to explore new ideas.	Mainly hypothetico-deductive, but potentially inductive. Biased data collection and analysis, e.g., influenced by researchers' values (influences at least partially controllable).
Social constructivism	Ontological relativism	Intersubjective knowledge	Mainly to explore new ideas (e.g., discover new scientific theories), but sometimes to confirm patterns and/or theories from previous work (e.g., transfer theories in other contexts).	Mainly inductive, but potentially deductive. Interpretation of data influenced by researchers' and participants' values.
Pragmatism	Action-related views of reality	Action-based knowledge	To explore and/or confirm action-oriented practical solutions and meaning-making.	Action-centered inductive and/or deductive data collection, analysis and solutions (what works).
Critical theory	Promotion of social justice via scientific understanding and explanations	Representation of the voice and views of the oppressed	To explore and/or confirm and reduce inequities. To enact individual empowerment and social changes for marginalized groups for instance.	Inductive and/or deductive participatory research methodologies and methods. Improvement of individual well-being and social justice. Research influenced by participants' values.
Critical realism	Ontological realism	Epistemological relativism	Mainly to confirm and/or test an a priori Middle-Range Theory (MRT), but sometimes to explore new ideas for building a new emerging MRT or improving an a priori MRT.	MRT-driven deduction, induction, abduction, and retroduction.

First published in Pluye et al. (2021): Reprinted with permission of the open access journal *McGill Family Medicine Studies*.

Three combinations without integration of worldviews

A-paradigmatic combination (convergence or divergence)

This combination suggests research methods, including MM, can be performed independently from worldviews. In this combination, paradigms (and personal worldviews) are “abstract conceptual ideas that can usefully inform” researchers’ methodologies broadly, i.e., they can help describe methods, but do not constrain them (Greene, 2007). Stated otherwise, the research purpose and specific research question(s) are more influential (compared to worldviews) for planning, conducting and reporting QUAL and QUAN methods, and their mixing. In this combination, there may be convergence or divergence of MM research team members’ worldviews because members set aside their worldviews. This means that there is no integration of worldviews pre-, per- and post-study. It does not mean that worldviews do not influence the MM study.

Substantive theory combination (convergence or divergence)

In this combination, MM can address research questions and mobilize concepts or theories that unify QUAL and QUAN methods, independent of the worldviews. For instance, a program theory can be used in a MM evaluation research, e.g., an ecological model or a logic model, independent of the research team members’ worldviews (Greene, 2007). The theoretical orientations are deemed more important than the philosophical underpinnings, e.g., a program theory becomes “the guiding framework for evaluation design and method choices, and the evaluation becomes theory-oriented” (Greene, 2007). In other words, the theory pertinent to a particular MM study is more influential (compared to worldviews) for planning, conducting and reporting QUAL and QUAN methods, and their mixing. Similar to the above combination, there may be convergence or divergence of MM research team members’ worldviews because members set aside their worldviews, which does not mean that worldviews do not influence the MM study.

Single worldview combination (convergence)

As shown in the above description of five common worldviews and in line with Greene (2007), several contemporary worldviews lend themselves to MM projects. There is convergence of MM research team members on one worldview for the project. This “single”

combination assumes all members share one similar worldview, or follow variants of the same worldview, and as such, reconciliation can be simple when there are worldview-related tensions between team members. Thus, there is no explicit integration of worldviews, but more than one variant can be at play.

Three combinations with integration of worldviews

Multiple worldviews combination (convergence)

In a MM study with this combination, multiple worldviews are present, but one is seen as particularly appropriate (Creswell et al., 2003; Teddlie and Tashakkori, 2010). Stated otherwise, although MM research team members are influenced by different worldviews that provide insight for the MM project, one worldview is seen as most influential and better aligned with the MM design, the methods for collecting and analyzing QUAL and QUAN data, and the way the results are reported. Researchers make their multiple worldviews explicit and provide a rationale for selecting one “best-fit” worldview for the entire MM project or sub-projects (Creswell and Plano Clark, 2011). Thus, MM research team members converge on one worldview for the project.

Complementary strengths combination (convergence)

In this combination, MM team members must assume that some worldviews are predominantly (but not necessarily) associated with QUAL methods, and must be kept separate from those predominantly (but not necessarily) associated with QUAN methods (Tashakkori and Teddlie, 2010; Greene, 2007). It follows, therefore, that methods for collecting and analyzing QUAL and QUAN data must be kept separate (e.g., a study may have a QUAN phase using one worldview followed by a QUAL phase using another worldview). Thus, QUAL and QUAN methods and results are reported separately, according to the corresponding worldviews, and their complementarity described. In a two-phase MM study, for instance, the integration of worldviews may take place where the QUAL and QUAN phases connect (a different worldview being used for each phase). In addition, pragmatism, critical theory and critical realism can help to integrate worldviews (Azzopardi and Nash, 2014; Christ, 2013; Shannon-Baker, 2016). Thus, MM research team members converge on the use of different worldviews depending on the component of the project.

Dialectical pluralism combination (divergence)

This combination assumes that the worldviews associated with QUAL and QUAN methods are different and interdependent (Johnson, 2008). Dialectical pluralism generates creative tensions leading to discovery and innovation. The similarities, differences and contradictions between the QUAL and QUAN methods and results must be described and discussed in the MM team meetings and publications. All worldviews have something to offer and the use of two or more different ones in a single MM study contributes to a greater understanding of the phenomenon under investigation (Greene and Caracelli, 2003). In a dialectical combination, “a reciprocal mutually respectful dialogic relationship between philosophical frameworks and methodological decisions” is warranted (Greene and Caracelli, 1997). MM team members can diverge on worldviews that influence the project, and must assume that (a) there are different ways to conceptualize reality, (b) knowledge is enhanced through respectful dialog among different conceptualizations, and (c) values of tolerance, acceptance, and equity are particularly important (Plano Clark and Ivankova, 2016). According to Shannon-Baker (2016), MM researchers must remain reflective throughout the dialectical inquiry; promote dialog among theories, data, and results; and integrate diverse viewpoints, particularly from tensions.

Proposed solution: recommendations for combining worldviews

This led to propose a conceptual framework, published elsewhere (Pluye et al., 2022), and recommendations to help students handle worldview-related tensions and combine worldviews in a MM project. The starting point is the description of worldviews at play and possible combinations. For example, MM research team members can read and discuss this chapter in a journal club, thereby eliciting and uncovering members’ worldviews. Similar to a compass, Table 1 can help the team to navigate the most common worldviews at play in their project.

Then, students and MM research team members can navigate the following four recommendations (R1, R2, R3 and R4), and select the most appropriate for planning and conducting their project (the most coherent with questions and objectives of the project). Research team members may decide to keep worldviews as indifferent kinds, avoid further debates, and adopt the a-paradigmatic combination, or the substantive theory combination. Or, should they share a worldview, they may elicit key characteristics of this worldview using the nominal group technique for instance (R1: move to single worldview). Indeed, R1 is the most straightforward recommendation. Then, it is also possible for team members to reach consensus on characteristics of a worldview that constitutes the “best fit” for their MM project (R2: move to multiple worldviews). Another option is that they discuss characteristics of their different worldviews, and use them for integrating QUAL and QUAN methods (R3: move to complementary strengths). Finally, they can define characteristics of their different worldviews, and select the dialectical approach (R4: move to dialectical pluralism).

This section adds to knowledge on worldviews in MM research at the research team level in terms of six possibilities how team members can mobilize and combine worldviews in their studies. Future educational research can empirically study how and why team members, e.g., a student and their mentors, combine worldviews in a MM study. This will provide useful practical examples and promote the explicit combination of worldviews (with convergence or divergence of worldviews) and the reporting of their consequences on planning, conducting and reporting MM research.

Convergence and divergence in mixed studies reviews

Convergence and divergence can be seen when conducting systematic mixed studies reviews, i.e., reviews combining qualitative, quantitative and mixed methods studies (Hong et al., 2017). This type of review is gaining in popularity due its potential to provide a more complete and deeper understanding of a complex phenomenon. The purpose of this section is to present examples of how convergence and divergence can be addressed when synthesizing evidence in mixed studies reviews.

Literature reviews of qualitative, quantitative and mixed methods studies

Literature reviews have been used by policy-makers, researchers and health service providers to inform decision-making (Bunn et al., 2015). Traditionally, reviews have given preference to QUAN evidence, e.g., systematic reviews of randomized controlled trials (RCTs) with meta-analysis to assess the effectiveness of educational interventions or programs. However, a focus on QUAN evidence is insufficient in areas where research is not dominated by RCTs (Goldsmith et al., 2007), as seen in educational research (Gough and Thomas, 2016). In addition, RCTs are not appropriate nor sufficient to address complex and multifaceted questions such as “Why does it work?”, “How does it work?”, and “What works for whom in what context?”. Such questions can be addressed in mixed studies reviews with the synthesis of QUAL and QUAN evidence (Dixon-Woods et al., 2004; Pluye and Hong, 2014; Pluye et al., 2016).

In education, Randolph (2009), among others, proposes guidance for students to plan, conduct and report qualitative and quantitative reviews with synthesis of QUAL and QUAN evidence, respectively. In contrast, there is a growing interest in synthesizing QUAL and QUAN evidence derived from qualitative, quantitative and mixed methods studies using a diversity of approaches and designs. This new type of review has been labeled with various terms such as integrative review (Whittemore and Knafl, 2005), mixed methods review (Harden and Thomas, 2005), mixed methods research synthesis (Heyvaert et al., 2016), mixed research synthesis (Sandelowski et al., 2006), and mixed studies review (Pluye et al., 2009a). These reviews can be used to provide (a) a deeper understanding of quantitative evidence, (b) a statistical generalization of findings from qualitative evidence, and (c) a corroboration or complementarity of knowledge obtained from quantitative and qualitative evidence (Pluye and Hong, 2014).

Synthesis designs in mixed studies reviews

A challenge when performing mixed studies reviews is the combination of findings from a diversity of studies (qualitative, quantitative and mixed methods). To address this challenge, a review of 459 systematic mixed studies reviews has been conducted and identified two main synthesis designs (Hong et al., 2017): convergent and sequential (described below). Based on where the integration occurred, three subtypes of convergent synthesis design were identified: data-based, results-based and parallel-results. The synthesis designs can provide guidance on how to integrate evidence, and report the added value and insight gained of this integration.

First, in the *data-based convergent synthesis design*, all included studies are analyzed using a same synthesis method and results are presented together. Since only one synthesis method is used for all evidence, data transformation is involved (e.g., qualitative data are transformed into numerical variables/values, or quantitative data are transformed into categories/themes). This transformation may reveal convergence or divergence between QUAL and QUAN data. This design usually addresses one review question. Among the systematic mixed studies reviews using this design, three main objectives are found. The first objective is to describe the findings of the included studies; the synthesis methods range from summarizing each study to grouping main findings. The second objective is to identify and define main concepts or themes using a synthesis method such as qualitative thematic synthesis (e.g., identify main barriers and facilitators to the implementation of a program, or types of impact). The third objective is to establish relationships between the concepts and themes identified from the included studies, or to build a conceptual framework (or theoretical model).

Second, in the *results-based convergent synthesis design*, the qualitative and quantitative evidence are analyzed and presented separately, but integrated using another synthesis method. The integration can consist of comparing or juxtaposing the findings of qualitative and quantitative evidence using tables and matrices, or reanalyzing evidence in light of the results of both syntheses. This comparison may reveal convergence or divergence between QUAL and QUAN results. For example, a review was interested in determining the interventions to optimize Type I diabetes care and management in school setting (Edwards et al., 2014). They performed (a) a synthesis of the QUAN studies on the intervention effectiveness, (b) a synthesis of studies of people’s viewpoints, and (c) an overarching synthesis matrix integrating the results of both syntheses. This comparison consisted of mapping the best practices against barriers and facilitators to see the extent to which the interventions were effective and addressed the barriers identified by young people and school personnel.

Third, in a parallel-results convergent design, QUAL and QUAN evidence are analyzed and presented separately. The integration occurs during the interpretation of results in the discussion section. This can reveal convergence and divergence between QUAL findings and QUAN results. For example, in a review on computerized virtual patients in health professional education, the authors studied two review questions: Q1. What is the learning effectiveness of virtual patients? and Q2. What virtual patients design features are associated with higher learning effectiveness? (Cook et al., 2010). The evidence for each question was reviewed separately (using meta-analysis for Q1 and thematic synthesis for Q2) and brought together in the discussion and recommendations.

Fourth, in a sequential design, there is complementarity between QUAL and QUAN evidence from included studies, i.e., a form of convergence. This design involves a two-phase approach where the data collection and analysis of one type of evidence occur after, and are informed by the collection and analysis of the other type. This design usually addresses one overall review question with sub-questions and both syntheses complemented each other. For example, in a review aiming to identify the impact of clinical information-retrieved technology on physicians, residents and medical students, a thematic qualitative synthesis of the QUAL studies (synthesis 1) led to identify types of impact of searches for information (Pluye et al., 2005). Then for each type of impact, a descriptive statistical synthesis of the QUAN studies was performed to identify the frequency of searches (synthesis 2 informed by results of synthesis 1).

Convergence and divergence in synthesis of QUAL and QUAN evidence

The science of systematic review grew from the need of making sense of convergences, divergences and inconsistencies between studies. For example, Sen. Walter Mondale, a former vice-president of the United States, provided the following statement during the American Education Research Association conference in 1971: "For every study, statistical or theoretical, that contains a proposed solution or recommendation there is always another, equally well documented, challenging the assumptions or conclusions of the first. No one seems to agree with anyone else's approach. But more distressing: no one seems to know what works. As a result I must confess, I stand with my colleagues confused and often disheartened." (cited in Light and Smith, 1971).

Since the 1970s, systematic reviews have been developed through formal and rigorous methods to better study the convergences and divergences between included studies (Hong and Pluye, 2018). For example, meta-analysis was one of the first methods developed by Gene Glass, a statistician in educational psychology, to statistically combine the results of several studies (Glass, 1976). Part of the evidence-based practice movement, it is recommended to conduct and use high quality systematic reviews and evidence synthesis methods to support educational practice (Davies, 2000).

Considering that the focus of this section is on the synthesis process in systematic mixed studies reviews, the literature on process studies especially in the fields of management provides insight into these concepts. In line with Langley et al. (1995), the convergent synthesis can be defined as a process of gradual and successive comparison and interpretation of QUAL and QUAN evidence derived from included qualitative, quantitative and mixed methods studies. Researchers are working forward in a non-linear manner guided by a cognitive representation of new data-based synthesis, or results-based synthesis, or interpretation of results to be created. Thus, they identify complementarities, corroborations, similarities, differences and contradictions, describe how the QUAN and QUAL evidence are integrated, and discuss the insight gained from this process. Furthermore, in line with Van de Ven (1992), a sequential synthesis design can be defined, according to a developmental perspective (phase-1 informing phase-2; and phase-2 building on the results of phase-1), as a change of focus at the level of data or synthesis over time, and as a cognitive transition into a new phase (e.g., from qualitative to quantitative, or from quantitative to qualitative evidence) where QUAL and QUAN evidence complement each other.

As mentioned, integration of QUAL and QUAN methods is an inherent component of MM research, specifically the integration of QUAL and QUAN evidence. In systematic mixed studies review-related syntheses, this integration requires to explore the relationships between results of included studies (convergences) and provide explanations of heterogeneity (divergences) (Thompson Coon et al., 2019). Based on different perspectives and different sources of evidence, convergences can reinforce the implications and recommendations issued from a review (Andrews, 2005). Yet, there can be divergences in priorities, values, language, meaning, agenda and time scale between researchers, practitioners and policy makers (Harlen and Deakin Crick, 2004). To address these divergences, there is a need for closer working collaboration between these parties (Harlen and Deakin Crick, 2004). For example, a mixed studies review on decisional needs assessment of patients with complex care needs used a participatory research approach in which multiple users were involved at the different steps the review process (Bujold et al., 2017).

Another source of divergences can be found in the heterogenous nature of included studies in mixed studies reviews. Indeed, the studies usually address different questions and perspectives, adopt different paradigms, and use different methodologies. Heterogeneity can also be found in terms of populations, interventions, outcomes and studies' quality. However, divergences are rarely reported; e.g., four of a sample of 59 mixed studies reviews (7%) mentioned some form of divergence, i.e., discrepancies between QUAL and QUAN evidence (Pluye et al., 2009a). This does not necessarily mean that divergences are ignored, but that they are not considered essential or worth of reporting. In contrast, this chapter argues that obtaining divergent results should not be viewed negatively because they provide a more comprehensive and nuanced understanding of the phenomenon under study (Thompson Coon et al., 2019).

The integration of QUAL and QUAN evidence requires to use an iterative process, and often go back to the data to better explore how they match and mismatch. For example, in a review on nonpharmacological school-based interventions for attention deficit hyperactivity disorder, an interweave synthesis approach was used to identify patterns and links between and across syntheses (Thompson Coon et al., 2019). This approach required to start once the reviewers developed an understanding of the initial findings, be immersed in the data, and use intersubjective questions. This review also involved discussion of the findings of the overarching synthesis with stakeholders.

Specifically in syntheses, a strategy to address convergences and divergences is to perform a sensitivity analysis, which consists of repeating the synthesis by excluding a set of particular studies to explore how the results obtained in the synthesis are modified (Deeks et al., 2019). For example, in a mixed studies review on factors influencing the implementation of mental health recovery into services, sensitivity analysis was conducted to determine whether the quality of the included studies influenced the findings of the synthesis (Piat et al., 2021). In doing so, reviewers verified whether some themes identified were solely based on lower-quality studies.

Convergences and divergences can also be explored using subgroup analysis that involves splitting the data into subsets of studies (such as particular geographical locations, types of interventions, types of populations, etc.) (Deeks et al., 2019). For example, in a review on healthy eating among children aged 4–19 years, the findings from the qualitative synthesis on the views of children were compared with results from controlled trials on the effectiveness of interventions to promote children's fruit and vegetable consumption (Thomas et al., 2004). This allowed to identify interventions that match the recommendations derived from the views and experiences of children on the consumption of fruits and vegetables. Subgroup analysis was then conducted to study whether interventions that match recommendations show bigger effect sizes.

Convergence and divergence in primary MM research

In primary MM research, convergence and divergence of QUAL and QUAN evidence (data or results) can occur no matter the type of MM design, and can derive from any type of integration. Nowadays, convergence and divergence in MM empirical research are well-known. In the MM literature, convergence is almost everywhere, whereas divergence is seldom mentioned. For example, 21 MM textbooks were reviewed in preparation for this chapter: nine presented only convergence of QUAL and QUAN evidence, and 12 allotted some space to divergence. In addition, a snowball technique was performed, e.g., using the Scopus database, and multiple reference articles were also reviewed. The analysis of these articles suggests that few MM empirical studies report divergence between QUAL and QUAN evidence, and little truly focus on divergence. Based on this sample of MM textbooks and reference articles, the definitions of convergence and divergence between QUAL and QUAN evidence in primary MM research are summarized in the following sections and Table 2, and illustrated using MM education-related empirical studies.

Convergence

In the MM literature, the concept “Convergence between QUAL and QUAN evidence” refers to two main sub-concepts: corroboration and complementarity. First, corroboration confirms or validates scientific data or results using parallel QUAL and QUAN evidence. This correspondence of evidence enhances their trustworthiness (evidence from QUAL methods supported by evidence from QUAN methods, and vice versa). Stated otherwise, the corroboration of QUAL and QUAN evidence from different sources, methods and investigators strongly reinforces a proposition about a phenomenon when there is theoretical and conceptual coherence, whatever the MM team researchers' worldviews and the methodologies at play. In corroborating evidence, the focus is on the verification of the quality of QUAL and QUAN methods, which is needed to confirm that the same conclusion from QUAL and QUAN evidence is trustworthy when the verification concludes that QUAL findings are credible, dependable and confirmable, while the QUAN results have internal validity, and are reliable and comparable. Specifically in MM case studies, corroboration refers to “confirming cases,” i.e., cases supporting a pattern of cases (e.g., a subset of cases presenting similarities in terms of QUAL and QUAN data). Furthermore, MM researchers and students should be cautious about a limitation of corroboration, that is the risk of spurious corroboration when QUAL and QUAN evidence share the same source of error, e.g., with respect to the sampling. In the literature, synonyms of corroboration are consensus, consistency, consonance, and similarity. Corroboration is sometimes associated with the trigonometric metaphor “Triangulation,” which has been highly controversial for the past three decades, and should be abandoned due to its vagueness and polysemy in scientific publications (Morgan, 2019).

In the following example, QUAL and QUAN evidence were obtained through separate data collection and analysis. Rosenkranz et al. (2015) sought to study what motivates medical students to do research. This MM study integrated a cross-sectional quantitative survey and a qualitative descriptive research, which were conducted in a concomitant parallel manner. The QUAN component consisted of 806 eligible students among which 579 (71.8%) completed the survey structured questionnaire. Survey results revealed that students who experienced a sense of achievement through research activities were more likely to perceive future research activities in a positive manner. In the QUAL component, all medical students were invited to interviews on prior research experiences and perceived barriers/facilitators to do research. The qualitative data analysis was informed by the Self-Determination Theory (three domains: autonomy, competence and relatedness). Then, QUAN results and QUAL findings were compared according to the theoretical dimensions and factors, both in a narrative and a tabulated format: “content or thematic areas were identified from both datasets, compared and contrasted; integration occurred at the point of final data interpretation, by identifying groupings

Table 2 Definitions of convergence and divergence between QUAL and QUAN evidence.

<i>Authors: research domain(s)</i>	<i>Convergence of QUAL and QUAN evidence</i>	<i>Divergence of QUAL and QUAN evidence</i>
Bazeley (2012), Bazeley (2018): multidisciplinary ^a	Complementarity of QUAL and QUAN evidence (better understanding, and generating stronger inferences, i.e., better supported by evidence, or more generalizable, or both). e.g., map complementary analysis (explore, sort by categories, compare or merge, describe and elaborate). Corroboration (confirming or validating using parallel evidence).	Dissonance of QUAL and QUAN evidence (different meanings). Varying intensity from inconsistencies, differences, disagreements, contradictions, conflicts, discords, or paradoxes. Three strategies for handling divergence: verify methods, reconcile divergences in a plausible way (context exploration, theory renewal, interpretive explanation), or initiate new research phase or project.
Caracelli and Greene (1993): evaluation studies	Corroboration and correspondence of evidence. Complementarity (overlapping distinct aspects of the phenomenon; enhancing, illustrating, or clarifying QUAN evidence with QUAL evidence, or vice versa).	Contradictions and paradoxes. One strategy for handling divergence: initiate new research phase or project (divergence recasting research questions).
Creamer (2018): multidisciplinary	Complementarity (gains of using QUAL and QUAN methods). Corroboration of evidence (enhancing validity).	
Creswell and Plano Clark (2018): multidisciplinary	Congruent evidence (comparison in a narrative or table format).	Discrepant or incongruent evidence (comparison in a narrative or table format). Conflicts, contradictions, discrepancies, and inconsistencies of evidence. Three strategies for handling divergence: verify validity of evidence (absence of methodological problems), reconcile divergences in a plausible way (reanalysis of datasets), or initiate new research phase or project.
Curry and Nunez-Smith (2015): health sciences	Coherent evidence. Complementary evidence. Spurious convergence (QUAL and QUAN evidence sharing same bias or source of error).	Conflicting, dissimilar, divergent, or inconsistent evidence (highlighting novel findings, discovery, insights and unanswered questions). Three strategies for handling divergence (citing Pluye et al., 2009b).
Erzberger and Kelle (2003): multidisciplinary	Corroboration of evidence: triangulation trigonometric metaphor (focus on verification or validation of methods). Complementary evidence: QUAL and QUAN evidence supplementing each other (adding breadth and depth to analysis, providing a more complete picture compared to QUAL or QUAN evidence alone). Potential spurious convergence (methods biased in the same way).	Inconsistent or contradictory evidence (pointing to new insights). Three strategies for handling divergence: verify validity of evidence (absence of methodological problems), reconcile divergences in a plausible way (reanalysis of datasets), or initiate new research phase or project.
Fetters and Molina-Azorin (2017): multidisciplinary	Corroboration (confirming same conclusion from QUAL and QUAN evidence). Complementarity (QUAL and QUAN methods providing non-conflicting evidence on different aspects of a phenomenon). Variant of complementary evidence: expansion (QUAL and QUAN evidence providing a broader overlapping understanding compared to QUAL or QUAN alone).	Discordance or conflicts. Three strategies for handling divergence (citing Pluye et al., 2009b).
Fetters (2020): multidisciplinary	Complementary evidence (evidence from QUAL methods enhanced by evidence from QUAN methods, or vice versa). Corroborating evidence (evidence from QUAL methods supported by evidence from QUAN methods, or vice versa).	Corroboration disconfirmed. Contradiction, discordance, conflict or paradox (based on the comparison and interpretation of evidence). Three strategies for handling divergence (citing Pluye et al., 2009b).

(Continued)

Table 2 Definitions of convergence and divergence between QUAL and QUAN evidence.—cont'd

<i>Authors: research domain(s)</i>	<i>Convergence of QUAL and QUAN evidence</i>	<i>Divergence of QUAL and QUAN evidence</i>
Greene (2007) and Greene (2015): social sciences	Complementarity (developing broader, deeper, and more comprehensive social understandings of complex social phenomena by using methods that tap into different facets or dimensions of these phenomena). Corroboration, consistency, consonance, or consensus (based on the comparison of QUAL and QUAN evidence).	Clashes, conflicts, divergence, dissonance, paradox, or puzzle (comparison of evidence generating new unanticipated perspectives, understandings and insights). One strategy for handling divergence: initiate new research phase or project (generative pathway).
Ivankova and Kawamura (2010): multidisciplinary Mathison (1988): education	Synergy (better understanding and gaining insights). Corroboration (QUAL and QUAN evidence from different sources, methods and investigators supporting a single proposition about a phenomenon).	Contradictory evidence (revealing the value of combining methods). Inconsistent evidence (disconfirming a single proposition, and presenting alternative propositions containing inconsistencies and ambiguities). Contradictory evidence (opposing views of the social phenomenon being studied). One strategy for handling divergence: reconcile divergences in a plausible way.
Moffatt et al. (2006): health sciences	Similarities. Complementarity (exploring QUAL and QUAN evidence separately: each addressing distinct, though related, aspects of a phenomenon).	Discrepancies. Discord. Of 6 strategies for handling divergence, 1 actually refers to convergence (complementarity), and 5 can be merged in the following 3 main categories: (a) VERIFY validity of evidence (absence of methodological problems); (b) RECONCILE divergences in a plausible way (reanalysis of datasets including the comparison of the participants from each component, and the exploration of whether the intervention under study worked as planned and whether the constructs and concepts under study in each component match, e.g., health outcomes); and (c) INITIATE new research phase or project (collecting further data to allow more comparisons).
Morgan (2014) and Morgan (2019): social sciences	Corroboration: convergent evidence (increasing credibility of similar evidence from different methods: "QUAL = QUAN"). Complementarity (covering different aspects of a phenomenon: "QUAL + QUAN").	Divergent evidence (raising doubts about whether the different methods studied a similar phenomenon). One strategy for handling divergence: initiate new research phase or project (generative pathway).
Nightingale (2009) and Nightingale (2016): environment	Complementarity (different methods with evidence complementing each other when paradigmatic and theoretical coherence). Corroboration (ensuring the robustness of methods when conceptual coherence—similar constructs).	Divergence and gaps (QUAL and QUAN evidence failing to make sense in relation to each other). Two strategies for handling divergence: verify validity of evidence (e.g., the absence of theoretical error), or reconcile divergence in a plausible way (each data set revealing a different understanding, e.g., referring to different paradigms).
Plano Clark and Ivankova (2016): multidisciplinary	Complementary evidence (synergistic responses to a question with additional coverage and enhanced effectiveness). Confirming evidence (verifying theories). Convergent evidence (answering similar questions with greater certainty and stronger inferences).	Divergent evidence (uncovering divergent results).

Table 2 Definitions of convergence and divergence between QUAL and QUAN evidence.—cont'd

<i>Authors: research domain(s)</i>	<i>Convergence of QUAL and QUAN evidence</i>	<i>Divergence of QUAL and QUAN evidence</i>
Pluye et al. (2009b): multidisciplinary		Three strategies for handling divergence: verify the quality of methods, reconcile divergences in a plausible way, or initiate new research phase or project.
Sammons (2010): education	Synergistic understanding of educational effectiveness (connection or comparison of evidence, each method informing the other in a reciprocal or dialectical manner).	
Teddlie and Tashakkori (2009) and Teddlie and Tashakkori (2010): social and behavioral sciences	Convergence of QUAL and QUAN evidence. Confirming cases (verifying patterns).	Dissimilarity or divergence of QUAL and QUAN evidence (providing insights into complex aspects of a phenomenon). Disconfirming cases (refuting patterns).

^aBazeley (2018) includes a chapter on inherently mixed integration strategies such as MM case study, social network analysis and qualitative comparative analysis (QCA), which are centered on convergence and divergence, e.g., within cases and across cases (intensive case review necessary when divergent cases).

of content areas, presenting both datasets in tables and generating narrative descriptions for each conceptual grouping" (p. 4). The comparison highlighted the corroboration of the QUAN results and QUAL findings, e.g., about intrinsic motivators associated with confidence in (competence) and clinical relevance of (relatedness) research activities. Integrated results led to propose a conceptual framework explaining the students' motivation to do research, and recommendations to redesign the curriculum for improving their motivation.

Second, complementarity between QUAL and QUAN evidence can generate more knowledge on the phenomenon under study compared to QUAL or QUAN evidence alone. It addresses distinct, though related, aspects of a phenomenon. For example, it can develop a broader and deeper understandings of complex phenomena by using methods that tap into different dimensions and facets of these phenomena, including through QUAL- and QUAN-informed theoretical models (Kawamura et al., 2009). It can also enhance, illustrate, and clarify QUAN evidence with QUAL evidence, and vice versa. Typically, complementarity leads to a more comprehensive understanding, or more generalizable inferences, or both. In other words, the complementarity of QUAL and QUAN evidence from different sources, methods and investigators adds breadth and depth to data analysis, and provides more complete results because QUAL and QUAN evidence supplement each other when there is theoretical and conceptual coherence, regardless of the MM team researchers' worldviews and the methodologies at play. In complementing evidence, the focus is on the synergy between methods as each method informs the other, and can gain more insights in a reciprocal manner, e.g., evidence from QUAL methods is enhanced by evidence from QUAN methods, and vice versa (synergistic responses to the research questions). In the literature, synonyms of complementarity are coherence, congruence and synergy. A variant of complementarity seems to be "expansion," which refers to QUAL and QUAN evidence providing a broader overlapping understanding compared to QUAL or QUAN evidence alone (Greene, 2007). This notion implies that QUAL and QUAN evidence can address both similar and distinct (interrelated) aspects of a phenomenon.

In the following example, QUAN and QUAL evidence were obtained in an interdependent manner. Bindiganavile Sridhar et al. (2013) studied the validity of the Information Assessment Method (IAM), which supports popular Continuing Medical Education (CME) programs. The IAM allows physicians to report the relevance, cognitive impact, intention to use and expected patient health benefits of CME information. Physicians read information, rated it using IAM, and earned CME credit for this individual reflective learning activity. The IAM items were based on literature reviews and qualitative research studies. Researchers aimed to assess the content validity of IAM items from the users' perspective using a MM validation study. Content validity is composed of relevance (are items appropriate to assess the construct?) and representativeness (do items represent all aspects of the construct?). QUAN and QUAL methods were used to evaluate the relevance and representativeness of IAM items, respectively. The QUAN component consisted in a longitudinal quantitative study, 40 family physicians completed 1767 IAM questionnaires (hereafter IAM ratings) over 320 days, which allowed the researchers to calculate a relevance ratio for each item, namely the number of items rated divided by the total number of ratings. The QUAL component was a qualitative multiple case study in which these physicians were interviewed to explain their ratings. A case was a CME information content rated by a physician using the IAM. "The integration of the quantitative and qualitative methods occurred at the stage of interpreting quantitative (relevance) and qualitative (representativeness) results" (p. 5). For each IAM item, QUAN and QUAL results were compared and interpreted (e.g., item relevant and representative) in a tabulated format (side-by-side comparison of the relevance ratio and the participants' perceived representativeness). Then, a decision was made, e.g., item not changed, modified or removed, by a multidisciplinary expert panel using a set of guiding principles. Overall, the content validity of 16 IAM items was supported, and these items were not changed. Nine other items were modified, and three new items were added. This led to a content validated version of the IAM questionnaire for CME programs.

Divergence

In the MM literature, the concept “Divergence between QUAL and QUAN evidence” refers to a continuum from inconsistencies to paradoxes, and three strategies for handling divergence (verification, reconciliation and initiation). These concepts and strategies will be described and illustrated using MM education-related empirical studies in the next sections. Specifically, Scopus identified 52 open access articles published in 2020 with “mixed method*” in the title and “education” in the title/abstract/keyword. Of those, 22 were excluded because they did not refer to complete empirical studies (e.g., protocols). Of the 30 remaining articles, three (10%) reported differences between QUAL and QUAN results and are presented below (Allison et al., 2020; Dashtestani, 2020; Mankelow et al., 2020).

Definition and examples

Divergence means that QUAL and QUAN evidence fail to make sense in relation to each other. Regarding the meanings of QUAL and QUAN evidence, divergence varies in intensity from inconsistencies (nuanced meanings) to ambiguities, discrepancies, discordances, dissimilarities, dissonances, incongruencies, differences, disagreements, contradictions, conflicts, discords, oppositions, clashes and paradoxes (antagonistic meanings). Divergence is usually reported after a comparison and interpretation of QUAL and QUAN evidence in a narrative, or table format, or both (or in a theoretical format (Kawamura et al., 2009)). It can also be sought by design, e.g., in building cases. It can highlight methodological issues, novel findings (pointing to new insights) and unanswered questions. When corroboration is the focus, divergence initially raises doubts about whether the QUAL and QUAN methods study a similar phenomenon. Then, it can be challenging (leading to modify the proposition about the phenomenon under study) or disconfirming (leading to a counter-proposition) when methodological and theoretical issues are discarded. For example, divergence can disconfirm cases and refute anticipated patterns. When complementarity is the focus, divergence can be initially puzzling, then generates a new unanticipated understanding of the different aspects of the phenomenon under study (showing the value of combining methods by providing insights into complex aspects of a phenomenon). Two examples are presented below.

First, Allison et al. (2020) assessed an educational health website on infections and antibiotics for school teachers and students. They combined a survey and an interpretive qualitative study (scenario-based observation of teachers). The QUAN results showed that almost all teachers liked the website (97.1% would use it themselves, and 98.1% would recommend it to others). In contrast, QUAL findings revealed some drawbacks such as difficulties to access the student’s section of the website and the resources for training teachers.

Second, Dashtestani (2020) assessed the teachers’ acceptance, learning achievements and satisfaction regarding online courses in higher education. They compared the perceptions of teachers and university faculty members using surveys and a qualitative descriptive study (semi-structured interviews). Overall, QUAL findings corroborated QUAN results. However, QUAN results showed that teachers were positive about online courses regarding all outcomes, whereas QUAL results revealed that online courses were seen (a) less rigorous compared to face-to-face courses, (b) limited by technical and Internet connection problems, and (c) lacking human interaction and practical content. In addition, teachers reported that some of them did not take online courses seriously, while certificates of online courses lacked credibility, e.g., low employers’ acceptance. Furthermore, students had low knowledge of the content from online courses, and the services they received did not match the tuition fees.

Three strategies for handling divergence

Table 2 shows that three main strategies for handling divergence of QUAL and QUAN evidence are proposed for MM students and researchers: verification, reconciliation and initiation. These strategies require (a) a critical appraisal of the methodological quality of the QUAL and QUAN methods, (b) the reanalysis and interpretation of the available QUAL and QUAN evidence, and (c) the collection and analysis of additional evidence when needed (Moffatt et al., 2006; Pluye et al., 2009b). The selection of the appropriate strategy is summarized in Fig. 1. Definitions and empirical illustrations of these strategies among educational research are described in the following sub-sections.

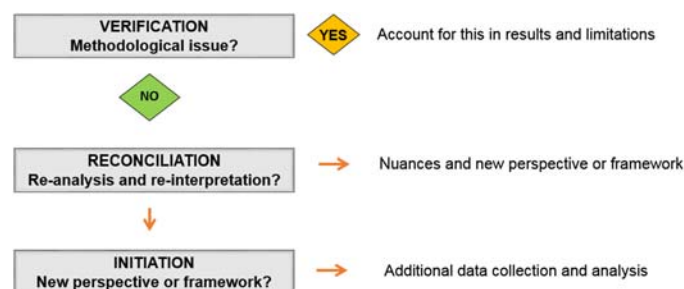


Fig. 1 Three strategies for handling divergence between QUAL and QUAN evidence.

Verification

Verification must be the first step to follow when there is divergence. For example, this strategy can lead to suggest that QUAL or QUAN evidence or both lacks trustworthiness, or that the results of the MM study are incomplete or inadequate, which might be the consequences of theoretical and conceptual incoherencies. In these situations, limitations must be reported in a transparent manner. Initial evidence must not be excluded, but reinterpreted (e.g., some QUAL or QUAN results can be reinterpreted as inconclusive due to methodological limitations). Rare empirical illustrations of verification were found among the MM empirical studies. It does not mean that it was absent or ignored, but that other information was worth reporting in the allotted word limit of publications for instance.

By way of illustration, [Parr et al. \(2021\)](#) studied teachers' motivations and ability to teach and their emotions when teaching. QUAL findings did not suggest a relation between perceived teaching abilities and teaching emotions. Yet, QUAN results suggested perceived teaching abilities were associated with greater teacher enjoyment and less anger. This led authors to verify potential methodological issues. Specifically, authors assessed the QUAN validity and QUAL credibility of their results, and the quality of their mixed methods "meta-inferences" (i.e., the integration of QUAL and QUAN results). They did this by "coming to consensus on the coding of all qualitative themes, ensuring that all quantitative measures had acceptable internal consistency reliability, and employing multiple mixed legitimization techniques" (p. 5). In turn, this led authors to find that this divergence might be associated with conceptual inconsistencies: QUAL interviews were broadly informed by a theory, whereas QUAN survey items represented each theory concept, dimension and facet (see quote below).

We did not prompt for specific motivations defined by FIT-Choice theory in qualitative interviews, but quantitative surveys had items for each FIT-Choice motivation. We took a broad approach in our interview protocol to allow for motivations not yet defined in the literature to emerge. However, an unintended consequence was that teacher interviewees talked relatively little about their perceived teaching abilities. It is plausible that the methodological constraints led to these discrepant findings (p. 11).

In another example, some divergences between QUAL and QUAN data were interpreted as inconclusive ([Pluye et al., 2013](#)). The contradictions led to reinterpret some data that initially suggested benefits of the intervention. This MM study was aimed to investigate family physicians' behavior with respect to searching for information in a specialized web database. The study integrated a longitudinal QUAN component and a QUAL multiple case study. Specifically, cases were critical searches for information by practicing physicians for specific patients. Critical searches had clear consequences in terms of patient health outcomes. The QUAN data collection documented physicians' searches for clinical information as well as their reflection on each search (physicians' ratings of the relevance, the cognitive impact, and the use of information). For each search, physicians completed the above-mentioned IAM questionnaire and earned CME credits. Guided by QUAN data, 39 physicians were interviewed about 365 rated searches. Interviews collected the physicians' explanation of the clinical situation, the objectives of the search, their use of this information, and the benefits for their patients. QUAL data included log reports, archives, and interview transcripts. Of 365 searches, 143 (39.2%) were forgotten or not used for a specific patient at the time of the interview. With respect to other searches, QUAN and QUAL data were merged into clinical stories (hereinafter called vignettes).

For each vignette, divergences between QUAL and QUAN data were identified as follows. Transcripts were analyzed by three researchers. The first researcher assigned extracts of transcripts to themes by going back and forth from textual data to themes, and synthesized QUAN and QUAL data into an initial vignette (highlighting potential divergences). The second researcher reviewed all data corresponding to the initial vignettes. Disagreements between these researchers regarding the interpretation of the data (summarized in the vignettes) were resolved by discussion and consensus. To ensure rigor in mixing data, they specifically discussed the divergences, and systematically documented strategies to solve these issues, e.g., they removed 34 (9.3%) unclear searches from the set of critical searches. In other words, they verified the trustworthiness of QUAL data, and classified these unclear searches as non-critical, i.e., searches without clear patient health outcomes. Then, the third researcher, a practicing physician, independently reviewed transcripts and vignettes corresponding to 188 (51.5%) critical searches with patient health benefits. He judged the trustworthiness of the clinical content, and the relationship between the use of information and patient health benefits. For each vignette, he answered the following two questions: Is this clinical story clear, convincing and coherent, from a medical perspective? Can these reported patient health outcomes be associated with these information hits and information use? No stories were deemed to be untrustworthy.

Reconciliation

Reconciliation is the second step to follow when there is divergence without methodological issues. This strategy must be performed in a transparent manner, and may occur when the divergence between QUAL and QUAN evidence can be reinterpreted in a sense-making plausible manner, e.g., following the reanalysis of available data in accordance with divergence-related insights. Such reanalysis can include an evaluation of the context and of the theoretical and conceptual coherence of methods. For example, reanalysis can consist of comparing participants from each method, determine whether the intervention under study worked as planned and whether outcomes match ([Moffatt et al., 2006](#)). Reconciliation can suggest nuances, e.g., through the integration of QUAL- and QUAN-informed theoretical models ([Kawamura et al., 2009](#)), and a new perspective or framework. However, QUAN editors

and peer-reviewers might call the “reanalysis” a “post-hoc analysis,” which is usually reported in the discussion section of scientific publications (not in results). In the following education-related example, the divergence between QUAL and QUAN data and results was reconciled.

Using a MM research study, [Huwendiek et al. \(2017\)](#) assessed 5th-year medical students’ acceptance, educational impact, and psychometric characteristics of two types of test used in pediatric exams: Context-Rich Single Best Answer questions (crSBA: patient vignette followed by one question with one correct answer) and Key-Feature Problems (KFP: patient vignette followed by more than one questions requiring free-text entry responses). QUAN data were exam results of 377 students and QUAL data were collected in eight focus groups. While QUAN results showed no difference in difficulty between the two types of test, QUAL findings revealed that KFP were perceived more difficult than crSBA questions. Based on the literature (see quote below), this divergence was reconciled as follows: “It may be hypothesized that if students truly believe that crSBA are easier, in the sense that these require less critical thinking, an examination that is purely crSBA-based could be perceived as less challenging” (p. 482). Interestingly, students who were given KFP performed better to the final practical test of clinical competences.

Our data indicate that KFPs with a long-menu response format provide a greater stimulus for the intense study of clinical reasoning than do crSBAs. This educational effect of KFPs has not been previously reported for clerkship students. A few studies, however, have demonstrated that the question format, notably the response format, can affect student learning. [One study] for instance, recently pointed out that the response format can indeed have an impact on the incidence of cueing, which again can impact perceived difficulty. In their study, they made a comparison between multiple choice-type and open-ended questions, presenting a first group of students with open-ended questions, followed by multiple choice questions, and a second group with the same questions in the reverse order. Irrespective of the format seen first, multiple choice scores resulted higher than those of the open-ended questions. The observed pattern suggests that it was cueing rather than memory for prior questions that led to increased multiple choice questions scores (p. 482).

Another example of reconciliation can be seen in the above-mentioned CME study ([Pluye et al., 2013](#)). Some divergences between QUAL and QUAN data were considered minor differences and easily reconciled, e.g., in the vignette on “Testing for multiple myeloma.” This vignette shows that on June 20, 2008, MD08 did a search at work, by themselves, and during an encounter with a patient (QUAN data—IAM responses). They retrieved one information hit about multiple myeloma and plasma cell malignancies. The reported search objectives were to address a clinical question and to look up something they forgot (QUAN data—IAM responses). “I was primarily just double-checking to see what sorts of tests I should be doing to either rule in or rule out the multiple myeloma as the diagnosis in this patient. [...] It was to make sure that there wasn’t something that I had forgotten in terms of specifically what to check in the urine” (QUAL data—interview). MD08 found relevant information in the specialized database that was provided by the research team, and used it for a specific patient (QUAN data—IAM response). They used this information to maintain and justify the diagnosis of the patient: “It was in terms of ordering the appropriate diagnostic test. [...] I would say that it [the information] probably justified it [the management] more than modified” (QUAL data—interview). No information-related patient health outcome was reported. There was a divergence between the cognitive impact of information reported in the IAM responses (rating “I recalled something”) vs. the interview: “*It was a question of confirming this thing about the urine part of the assessment, specifically.*” Whereas the former refers to a memory issue, the latter suggests that MD08 looked for information to confirm that they were doing the right thing. Researchers revised the interview and the archive (the informational content retrieved by MD08), and concluded that there may be something in the information content that confirmed MD08 was doing the right thing (urine test) while some details were recalled (a specific test). This reconciliation led researchers to include two types of cognitive impact in the final vignette (recall and confirmation).

Initiation

Initiation constitutes the third step to follow when there is divergence. This strategy begins with a new perspective or framework that emerges from QUAL and QUAN evidence. It consists of a generative process leading to collect additional data. Initiation requires two sub-steps: asking new research questions (or recasting questions), and collecting and analyzing new data to further examine from a new perspective or framework. By way of illustration, [Moffatt et al. \(2006\)](#) evaluated impacts of welfare rights advice on health and social outcomes among an aged population. While QUAL findings suggested many different impacts, QUAN results indicated no impact. Thus, methods were critically appraised, data reanalyzed, and additional data collected and analyzed.

[Rossman and Wilson \(1985\)](#) also used initiation. They aimed to evaluate the perceived usefulness of Regional Educational Service Agencies (RESAs) from the perspective of school administrators and teachers. They combined a survey with a qualitative research (interviews with school administrators and teachers, also the RESAs’ staff). QUAN results identified where and how frequently administrators sought information, whereas QUAL findings “revealed surprising variations in the ability of school administrators to select outside agencies for new information” (p. 638). This divergence led to conduct an additional survey “to see how frequent these modes of contact were for different categories of school personnel” (p. 637). These additional QUAN results showed that school “administrators received considerably more services from their RESA than did teachers” (p. 637). They provided details about differences between school administrators and teachers, which were not given by QUAL findings (see quote below).

In this example, general patterns of assistance seeking were described through analysis of the quantitative data. By turning to the qualitative data, a whole new perspective was revealed. Two very different approaches for using outside assistance agencies were discovered that have important implications and create new insights for the way these agencies recruit and retain school districts as clients (p. 639). (...) Initiation: The [second] survey data thus elaborated the interview data, providing a richness of detail about differences between teachers and administrators that the qualitative data alone could not provide (p. 637).

In another example, Taneja (2016) studied the effect of argumentation in science class on students' motivation. QUAL findings suggested teachers successfully facilitated argumentation to increase student engagement. In contrast, QUAN results suggested students' engagement (measured by the Student Motivation Toward Science Learning scale) showed improvement for only one learning strategy. There was no corroboration. Thus, they concluded that initiation could be a solution to this divergence. In future research, "expanding the time over which students respond to motivation survey [...], and probably collecting and analyzing changes in motivation within different contexts of argumentation [...] will provide greater validity" (p. 181). In this example, initiation was mentioned, but not performed.

In the above-mentioned CME study (Pluye et al., 2013), initiation was also used. Some divergences between QUAL and QUAN data led to initiate a new research cycle. Some QUAL data suggested that the IAM questionnaire item "I recalled something" was not always clear from the respondents' perspective. As shown in the vignette on "Testing for multiple myeloma," MD08 rated "I recalled something," and said that they were "to make sure that there wasn't something that I had forgotten" (interview). This, among other justifications, led us to conduct a mixed methods validation study for the delivery of educational material (ecological and logical content validation) (Badran et al., 2017). In the QUAN component, participants were 5596 physician members of the Canadian Medical Association who used the IAM (234,196 ratings collected in 2012). The relevance of IAM items with respect to their main construct was calculated using descriptive statistics (relevance ratio R). In the QUAL component, the representativeness of IAM items was evaluated by 15 family physicians (semi-structured interviews). QUAN and QUAL results were juxtaposed in a table, discussed with experts, and integrated. As a result, the item "I recalled something" was replaced with another item: "I was reminded of something I forgot."

Conclusion

The richness of MM research lies on its capacity of integrating qualitative and quantitative evidence. This integration requires to deal with convergence and divergence. This chapter provided definitions of what are convergence and divergence as well as illustrations of how they can be handled within three main dimensions: scientific worldviews, literature reviews and primary research. It is hoped that readers will find this chapter useful to address upstream challenges of integration in mixed methods research and will further pursue its conceptual development.

Specifically, implications for educational researchers are threefold. With regard to worldviews, our recommendations can help students and researchers handle worldview-related tensions that can occur between mixed methods research team members. We suggest that members describe worldviews at play, then select an appropriate combination of worldviews (a-paradigmatic, substantive theory, single worldview, multiple worldviews, complementary strengths, and dialectical pluralism). Regarding mixed studies reviews, we suggest using a synthesis design (either convergent synthesis designs or sequential synthesis design) and handle convergences and divergences using different strategies such as sensitivity analysis and subgroup analysis. Finally, with respect to divergences between QUAL and QUAN data and results in primary mixed methods research, we suggest students and researchers look for divergences, and apply the proposed strategies when needed (verification, reconciliation, and initiation).

References

- Alise, M.A., Teddlie, C., 2010. A continuation of the paradigm wars? prevalence rates of methodological approaches across the social/behavioral sciences. *J. Mix. Methods Res.* 4, 103–126.
- Allison, R., Hayes, C., Young, V., McNulty, C.A.M., 2020. Evaluation of an educational health website on infections and antibiotics in England: mixed methods, user-centered approach. *JMIR Form. Res.* 4, e14504.
- Andrews, R., 2005. The place of systematic reviews in education research. *Br. J. Educ. Stud.* 53, 399–416.
- Azzopardi, E., Nash, R., 2014. An examination of philosophical paradigms and a rationale for adopting a mixed methods approach. *Tourism Anal.* 19, 151–159.
- Badran, H., Pluye, P., Grad, R.M., 2017. When educational material is delivered: a mixed methods content validation study of the information assessment method. *JMIR Med. Educ.* 3, e6415.
- Baumard, P., 1997. Constructivisme et processus de la recherche : L'émergence d'une "posture" épistémologique chez le chercheur [Conference Paper]. In: *Colloque Constructivisme et Sciences de Gestion*. IAE, Cahiers de recherche LAREGO, Lille.
- Bazeley, P., 2012. Integrative analysis strategies for mixed data sources. *Am. Behav. Sci.* 56, 814–828.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage, Thousand Oaks, CA.
- Bindiganavile Sridhar, S., Pluye, P., Grad, R., 2013. In pursuit of a valid information assessment method for continuing education: a mixed methods study. *BMC Med. Educ.* 13, 137.
- Bujold, M., Pluye, P., Légaré, F., Haggerty, J., Gore, G.C., EL Sherif, R., Poitras, M.-E., Beaulieu, M.-C., Beaulieu, M.-D., Bush, P.L., 2017. Decisional needs assessment of patients with complex care needs in primary care: a participatory systematic mixed studies review protocol. *BMJ Open* 7, e016400.
- Bunn, F., Trivedi, D., Alderson, P., Hamilton, L., Martin, A., Pinkney, E., Iliffe, S., 2015. The impact of Cochrane reviews: a mixed-methods evaluation of outputs from Cochrane review groups supported by the National Institute for Health Research. *Health Technol. Assess.* 19, 1–100.

- Caracelli, V.J., Greene, J.C., 1993. Data analysis strategies for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 15, 195–207.
- Christ, T.W., 2013. The worldview matrix as a strategy when designing mixed methods research. *Int. J. Mult. Res. Approaches* 7, 110–118.
- Coates, A., 2021. The prevalence of philosophical assumptions described in mixed methods research in education. *J. Mix. Methods Res.* 15, 171–189.
- Cook, D.A., Erwin, P.J., Triola, M.M., 2010. Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Acad. Med.* 85, 1589–1602.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. SAGE, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2011. *Designing and Conducting Mixed Methods Research*. SAGE, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*. SAGE, Thousand Oaks, CA.
- Creswell, J.W., Tashakkori, A., Jensen, K.D., Shapley, K.L., 2003. Teaching mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Curry, L., Nunez-Smith, M., 2015. *Mixed Methods in Health Sciences Research: A Practical Primer*. SAGE, Thousand Oaks, CA.
- Dashtestani, R., 2020. Online courses in higher education in Iran: a stakeholder-based investigation into preservice teachers' acceptance, learning achievements, and satisfaction: a mixed-methods study. *Int. Rev. Res. Open Dist. Learn.* 21, 117–142.
- Davies, P., 2000. The relevance of systematic reviews to educational policy and practice. *Oxf. Rev. Educ.* 26, 365–378.
- Deeks, J.J., Higgins, J.P., Altman, D.G., Group, C.S.M., 2019. Analysing data and undertaking meta-analyses. In: Higgins, J.P.T., Thomas, J., Chandle, R.J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*, second ed. John Wiley & Sons, Chichester (UK).
- Denzin, N.K., 2010. Moments, mixed methods, and paradigm dialogues. *Qual. Inq.* 16, 419–427.
- Dixon-woods, M., Agarwal, S., Jones, D., Young, B., Sutton, A., 2004. *Integrative Approaches to Qualitative and Quantitative Evidence*. Health Development Agency, London, UK.
- Edwards, D., Noyes, J., Lowes, L., Haf spencer, L., Gregory, J.W., 2014. An ongoing struggle: a mixed-method systematic review of interventions, barriers and facilitators to achieving optimal self-care by children and young people with type 1 diabetes in educational settings. *BMC Pediatr.* 14, 228.
- Erzberger, C., Kelle, U., 2003. Making inferences in mixed methods: the rules of integration. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Fábregues, S., Escalante-Barrios, E.L., Molina-Azorin, J.F., Hong, Q.N., Verd, J.M., 2021. Taking a critical stance towards mixed methods research: a cross-disciplinary qualitative secondary analysis of researchers' views. *PLoS One* 16, e0252014.
- Fetters, M.D., 2020. *The Mixed Methods Research Workbook: Activities for Designing, Implementing, and Publishing Projects*. SAGE, Thousand Oaks, CA.
- Fetters, M.D., Molina-Azorin, J.F., 2017. The Journal of Mixed Methods Research starts a new decade: the mixed methods research integration trilogy and its dimensions. *J. Mix. Methods Res.* 11, 291–307.
- Giara, V., 2020. Disambiguating the role of paradigms in mixed methods research. *J. Mix. Methods Res.* 14, 11–25.
- Glass, G.V., 1976. Primary, secondary, and meta-analysis of research. *Educ. Res.* 5, 3–8.
- Goldsmith, M.R., Bankhead, C.R., Austoker, J., 2007. Synthesising quantitative and qualitative research in evidence-based patient information. *J. Epidemiol. Community Health* 61, 262–270.
- Gough, D., Thomas, J., 2016. Systematic reviews of research in education: aims, myths and multiple methods. *Rev. Educ.* 4, 84–102.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey Bass, San Francisco.
- Greene, J.C., 2015. Preserving distinctions within the multimethod and mixed methods research merger. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford.
- Greene, J.C., Caracelli, V.J., 1997. Defining and describing the paradigm issue in mixed-method evaluation. *N. Dir. Eval.* 5–17.
- Greene, J.C., Caracelli, V.J., 2003. Making paradigmatic sense of mixed methods practice. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Hacking, I., 1981. *Scientific Revolutions*. Oxford University Press, New York.
- Harden, A., Thomas, J., 2005. Methodological issues in combining diverse study types in systematic reviews. *Int. J. Soc. Res. Methodol.* 8, 257–271.
- Harlen, W., Deakin Crick, R., 2004. Opportunities and challenges of using systematic reviews of research for evidence-based policy in education. *Educ. Res.* 18, 54–71.
- Harvey, W.J., Michaud, M., Wilkinson, S., 2020. Mixed methods research in adapted physical education. In: Haegele, J.A., Hodge, S.R., Shapiro, D.R. (Eds.), *Routledge Handbook of Adapted Physical Education*. Routledge, London.
- Hesse-Biber, S., 2015. The problems and prospects in the teaching of mixed methods research. *Int. J. Soc. Res. Methodol.* 18, 463–477.
- Heyvaert, M., Hannes, K., Onghena, P., 2016. *Using Mixed Methods Research Synthesis for Literature Reviews: The Mixed Methods Research Synthesis Approach*. SAGE, Thousand Oaks, CA.
- Hong, Q.N., Pluye, P., 2018. Systematic reviews: a brief historical overview. *Educ. Inf.* 34, 261–276.
- Hong, Q.N., Pluye, P., Bujold, M., Wassef, M., 2017. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst. Rev.* 6, e61.
- Huwendiek, S., Reichert, F., Duncker, C., De leng, B.A., Van Der Vleuten, C.P.M., Muijtjens, A.M.M., Bosse, H.M., Haag, M., Hoffmann, G.F., Tönshoff, B., Dolmans, D., 2017. Electronic assessment of clinical reasoning in clerkships: a mixed-methods comparison of long-menu key-feature problems with context-rich single best answer questions. *Med. Teach.* 39, 476–485.
- Ivankova, N.V., 2015. *Mixed Methods Applications in Action Research*. SAGE, Thousand Oaks, CA.
- Ivankova, N.V., Kawamura, Y., 2010. Emerging trends in the utilization of integrated designs in the social, behavioral, and health sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. SAGE, Thousand Oaks, CA.
- Johnson, R.B., 2008. Living with tensions: the dialectic approach. *J. Mix. Methods Res.* 2, 203–207.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11, 156–173.
- Johnson, R.B., Onewuegbuzie, A., Turner, L., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1, 112–133.
- Kawamura, Y., Ivankova, N.V., Kohler, C.L., Perumean-Chaney, S., 2009. Utilizing mixed methods to assess parasocial interaction of an entertainment–education program audience. *Int. J. Mult. Res. Approaches* 3, 88–104.
- Kuhn, T., 1983. *La structure des révolutions scientifiques*. Flammarion, Paris.
- Langley, A., Mintzberg, H., Pitcher, P., Posada, E., Saint-Macary, J., 1995. Opening up decision making: the view from the black stool. *Organ. Sci.* 6, 260–279.
- Light, R.J., Smith, P.V., 1971. Accumulating evidence: procedures for resolving contradictions among different research studies. *Harv. Educ. Rev.* 41, 429–471.
- Mankelov, J., Ryan, C., Taylor, P., Martin, D., 2020. The effect of pain neurophysiology education on healthcare students' knowledge, attitudes and behaviours towards pain: a mixed-methods randomised controlled trial. *Musculoskelet. Sci. Pract.* 50, 102249.
- Martel, R., Shepherd, M., Goodyear-Smith, F., 2021. He awa whiria—a “Braided River”: an indigenous Māori approach to mixed methods research. *J. Mix. Methods Res.* <https://doi.org/10.1177/1558689820984028>. Preprint.
- Mathison, S., 1988. Why triangulate? *Educ. Res.* 17, 13–17.
- Maxwell, J.A., Mittapalli, K., 2010. Realism as a stance for mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Mcchesney, K., Aldridge, J., 2019. Weaving an interpretivist stance throughout mixed methods research. *Int. J. Res. Method Educ.* 42, 225–238.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N., Maxwell, J., Molina-Azorin, J.F., Niglas, K., 2016. Expanding thinking through a kaleidoscopic look into the future: implications of the mixed methods international research association's task force report on the future of mixed methods. *J. Mix. Methods Res.* 10, 221–227.

- Moffatt, S., White, M., Mackintosh, J., Howel, D., 2006. Using quantitative and qualitative data in health services research—what happens when mixed method findings conflict? *BMC Health Serv. Res.* 6, 1–10.
- Morgan, D.L., 2014. Integrating Qualitative and Quantitative Methods: A Pragmatic Approach. SAGE, Los Angeles.
- Morgan, D.L., 2019. Commentary—after triangulation, what next? *J. Mix. Methods Res.* 13, 6–11.
- Nightingale, A.J., 2009. Triangulation. In: Kitchin, R., Thrift, N. (Eds.), *International Encyclopedia of Human Geography*. Elsevier, Amsterdam.
- Nightingale, A.J., 2016. Adaptive scholarship and situated knowledges? Hybrid methodologies and plural epistemologies in climate change adaptation research. *Area* 48, 41–47.
- Niglas, K., 2010. The multidimensional model of research methodology: an integrated set of continua. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. SAGE, Thousand Oaks, CA.
- Parr, A., Gladstone, J., Rosenzweig, E., Wang, M.T., 2021. Why do I teach? A mixed-methods study of in-service teachers' motivations, autonomy-supportive instruction, and emotions. *Teach. Teach. Educ.* 98, e103228.
- Pascal, A., Aldebert, B., Rouziès, A., 2018. Les méthodes mixtes en systèmes d'information: Enjeux épistémologiques et méthodologiques. *Systèmes d'information et management* 23, 99–126.
- Piat, M., Wainwright, M., Sofouli, E., Vachon, B., Deslauriers, T., Préfontaine, C., Frati, F., 2021. Factors influencing the implementation of mental health recovery into services: a systematic mixed studies review. *Syst. Rev.* 10, 1–21.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE, Thousand Oaks, CA.
- Pluye, P., Grad, R.M., Dunikowski, L.G., Stephenson, R., 2005. Impact of clinical information-retrieval technology on physicians: a literature review of quantitative, qualitative and mixed methods studies. *Int. J. Med. Inf.* 74, 745–768.
- Pluye, P., Gagnon, M.P., Griffiths, F., Johnson-Lafleur, J., 2009a. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *Int. J. Nurs. Stud.* 46, 529–546.
- Pluye, P., Grad, R.M., Levine, A., Nicolau, B., 2009b. Understanding divergence of quantitative and qualitative data (or results) in mixed methods studies. *Int. J. Mult. Res. Approaches* 3, 58–72.
- Pluye, P., Grad, R.M., Johnson-Lafleur, J., Granikov, V., Shulha, M., Marlow, B., Ricarte, I., 2013. The Number Needed to Benefit from Information (NNBI): proposal from a mixed methods research study with practicing family physicians. *Ann. Fam. Med.* 11, 559–567.
- Pluye, P., Hong, Q.N., 2014. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. *Annu. Rev. Publ. Health* 35, 29–45.
- Pluye, P., Hong, Q.N., Bush, P.L., Vedel, I., 2016. The plurality of methods for reviews and syntheses: opening-up the definition of systematic reviews. *J. Clin. Epidemiol.* 73, 2–5.
- Pluye, P., EL Sherif, R., Hong, Q.N., Vedel, I., 2021. The pluralism of scientific worldviews in mixed methods research: where do I stand? [research report]. *McGill Fam. Med. Stud.* Online 16, e03.
- Pluye, P., Hong, Q.N., EL Sherif, R., Vedel, I., 2022. Combining scientific worldviews in mixed methods research. In: Ngulube, P. (Ed.), *Handbook of Research on Mixed Methods Research in Information Science*. IGI Global, Hershey, pp. 102–104.
- Randolph, J., 2009. A guide to writing the dissertation literature review. *Pract. Assess. Res. Eval.* 14, 13.
- Rosenkranz, S.K., Wang, S., HU, W., 2015. Motivating medical students to do research: a mixed methods study using Self-Determination Theory. *BMC Med. Educ.* 15, 95.
- Rossmann, G.B., Wilson, B.L., 1985. Numbers and words: combining quantitative and qualitative methods in a single large-scale evaluation study. *Eval. Rev.* 9, 627–643.
- Sammons, P., 2010. The contribution of mixed methods to recent research on educational effectiveness. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. SAGE, Thousand Oaks, CA.
- Sandelowski, M., Voils, C.I., Barroso, J., 2006. Defining and designing mixed research synthesis studies. *Res. Sch.* 13, 29–40.
- Schiffedercker, K.E., Reed, V.A., 2009. Using mixed methods research in medical education: basic guidelines for researchers. *Med. Educ.* 43, 637–644.
- Schoonenboom, J., 2019. A performative paradigm for mixed methods research. *J. Mix. Methods Res.* 13, 284–300.
- Shannon-Baker, P., 2016. Making paradigms meaningful in mixed methods research. *J. Mix. Methods Res.* 10, 319–334.
- Taneja, A., 2016. *Argumentation in Science Class: Its Planning, Practice, and Effect on Student Motivation*. PhD. Walden University.
- Tashakkori, A., Teddlie, C., 2010. *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2003. Major issues and controversies in the use of mixed methods in the social and behavioural sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. SAGE, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2010. Overview of contemporary issues in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, Thousand Oaks, CA.
- Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., Brunton, G., Kavanagh, J., 2004. Integrating qualitative research with trials in systematic reviews. *Br. Med. J.* 328, 1010–1012.
- Thompson Coon, J., Gwernan-Jones, R., Garside, R., Nunns, M., Shaw, L., Melendez-Torres, G., Moore, D., 2019. Developing methods for the overarching synthesis of quantitative and qualitative evidence: the interweave synthesis approach. *Res. Synth. Methods* 11, 507–521.
- Van De Ven, A.H., 1992. Suggestions for studying strategy process: a research note. *Strat. Manag. J.* 13, 169–188.
- Wenger-Trayner, B., Wenger-Trayner, E., Cameron, J., Eryigit-Madzwamuse, S., Hart, A., 2019. Boundaries and boundary objects: an evaluation framework for mixed methods research. *J. Mix. Methods Res.* 13, 321–338.
- Whittemore, R., Knafl, K., 2005. The integrative review: updated methodology. *J. Adv. Nurs.* 52, 546–553.

Thinking qualitatively in the context of mixed methods research: implications for study design

Brigitte Smit^a, Vanessa Scherman^b, and Linda Liebenberg^c, ^a The University of Johannesburg, Faculty of Education, South Africa; ^b Department of Psychology of Education, UNISA, Pretoria, South Africa; and ^c Faculty of Graduate Studies, Dalhousie University, Halifax, NS, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction and background	478
Theoretical implications of privileging qualitative methods in mixed methods research design	478
The role of qualitative research in conceptualizing a mixed method study	479
Methodological implications for the research design	480
Concluding thoughts	483
References	483
Further Reading	484
Relevant Websites – Optional	484

Introduction and background

Mixed methods research has been investigated from various perspectives by many authors. To summarize some of their definitions, mixed methods research is a deliberate blending of quantitative and qualitative methods, an ever-growing approach. Some authors perceive it to be a win-win approach with the best features of both worlds at the researcher's fingertips (Saldaña, 2015, p. 56). Saldaña and Omasta (2018, p. 159) build on this statement, affirming that the "intentional blending of qualitative and quantitative data collection and analyses for studies will benefit from the combined outcomes." Put differently, O'Reilly and Kiyimba (2015, p. 86) state that the fundamental claim of the mixed methods community is that the use of two or more methods within a single study increases the confidence in the findings, this is especially helpful within educational contexts where the school system itself is a subsystem within a much broader societal and indeed global system. In education we attempt to make sense of the system structures, and the interrelationships, as these are the key factors in the system, in understanding how these affect one another, sometimes in obvious but often non-obvious ways, change in the system is more likely (Senge, 2006). It is here that using mixed methods research becomes essential. The rationale of this type of research has been the ability to compensate for any weaknesses in a single research design, and when considering the education system this is an important consideration. Furthermore, integrating two different methods allows the researcher to obtain a more rounded and complete set of results, thus making the most of the complementary strengths of both quantitative and qualitative approaches. Therefore, mixed methods research aims not to replace either of the two approaches but to draw upon their strengths and minimize any limitations.

In this chapter, we privilege the notion of thinking qualitatively in the context of mixed methods research, which means, according to Saldaña (2015, p. 4–5), that researchers purposefully adopt different lenses, filters, and angles as they inquire into social life to discover new perceptions and cognitions about the phenomenon they are researching. Thinking qualitatively means acknowledging that not everyone thinks the same as researchers do and the ability to see things from the perspectives of others is essential in the search for understanding. Saldaña (2015) further argues that thinking qualitatively is a heightened awareness or metacognition, which means thinking about how the mind and heart work by analyzing data and reflecting on life (p. 4–5). Metacognition is not just about thinking how to think but knowing how to know. It is a hyper-awareness within social environments for observing sensory details and interpreting subtexts from people's words and actions. Thinking qualitatively is pondering the nuances of the data from data collection and data analysis. In this chapter, we first highlight the theoretical implications of privileging qualitative methods. We then discuss, from a methodological perspective, the importance of analysis components of mixed methods designs and how qualitative data can be transformed to add value to inferences drawn. The role of qualitative research in the conceptualizing phase of a mixed methods study is discussed, followed by a methodological discussion drawing on the analysis component of designing a mixed methods study. For each of these sections, an exemplar is used to highlight key aspects related to the theorizing of mixed methods studies. These exemplars were chosen for the purpose of highlighting how qualitative research can be highlighted and play a prominent role in design decision making.

Theoretical implications of privileging qualitative methods in mixed methods research design

If research is to be transformative, we need to reflect more critically on whose voice informs our research focus and research design (Mertens, 2017). There is a strong and established argument that without centering the voices of specific communities in the policies and programs that are directed at supporting and improving their health and well-being, the efforts made by policymakers to promote positive life outcomes will fail (Ford et al., 2003; Gaunle and Adhikari, 2010; Wallerstein and Duran, 2006). The same can be said in an educational context, where education plays an important role in the achievement of not only economic life

outcomes but also psychosocial life outcomes (Fernandez, 2019). Given the role of research in informing programs and policy in education, how research is conducted and how communities of interest are included is key, if not a critical, concern. When research findings more accurately reflect and account for the priorities of specific population groups, together with their lived realities, services and policies will be better able to support intended outcomes. Furthermore, when we integrate the perspectives of communities into what we research and how we do research, our research findings will be more closely aligned with their priorities and experiences (Kanáiaupuni, 2005; Krenichyn et al., 2007). In this way, the impact and success of programs and policies in education based on such research and related data will be significantly improved.

As researchers, we bring our substantive focus to our work. However, we often generate specific research questions and make related research design decisions based on our knowledge within this broad focus. This course of action does not guarantee that we account for the more immediate concerns facing school communities or particular population groups. Thought of differently, while we may be asking interesting and even good research questions, at best, these questions may be irrelevant given the lived experiences of those school communities we are working to understand. At worst, however, we may be generating research findings that reinforce discriminatory and oppressive practices (Lather, 1991; Mertens, 2021). Ignoring the self-determining nature of the individuals being studied raises questions about the ethics and validity of our work. As argued elsewhere (Liebenberg, 2009, p. 444), “If, ... the very basis of our research that is the questions asked in the research setting are based in existing “knowledge” formulated by dominant voices, how valid then is the data we analyze, and by extension, the findings of our research?” Strategically situating qualitative methods in research design and privileging the data from these methods better ensures that the research foundations are relevant to communities. This is especially true as demonstrated by Scherman and Liebenberg (2021) in how the instrument development process can be strengthened by privileging qualitative data in the process.

Similarly, and especially with quantitative research, we ordinarily draw on existing and privileged knowledge (i.e., published academic research) to make sense of research findings. Potts and Brown (2015, p. 19) explain that “knowledge is produced through the interactions of people, and as all people are socially and politically located ..., with biases, privileges, and differing entitlements, so too is the knowledge that is produced socially located and political.” Ironically, large-scale quantitative findings can be more richly explained by the lived experience brought to light by qualitative data (see, for example, the work of Munford and Sanders, <https://www.youthsay.co.nz/>). Asking school communities to comment on quantitative findings using qualitative approaches provides more accurate explanations of particular statistical results and what they mean to people and their daily lives. In addition to aiding in our understanding of statistical results, as researchers, we should also be mindful that while such results may provide crucial information on patterns in population education, health and behavior, for example, on their own, they cannot account for the complexity of people’s lives. Quantitative research results cannot necessarily explain why we see distinct patterns or what these patterns mean in real-life terms. The impact of such results is then significantly reduced, both in terms of policy and theory. Again, integration of qualitative approaches intended to illuminate the lived experience of quantitative results would contribute insights critical to effective policy and practice knowledge.

Collectively, research findings are of greater relevance to the populations of interest when qualitative research has been used intentionally to scaffold quantitative research design and explain quantitative results as has been illustrated by Scherman and Liebenberg (2021). Challenging dominant approaches to research that privilege quantitative research (Denzin, 2012), by amplifying the notion that knowledge is socially constructed, allows us to draw attention to the fact that while quantitative approaches can indeed provide important measurements of critical social issues, complex causes and understandings of these issues may remain unexamined (Potts and Brown, 2015). These methodological considerations reflect ontological and epistemological concerns regarding knowledge and how we create it (Potts and Brown, 2015). The argument is not whether we should be doing mixed methods research or, better, qualitative versus quantitative. The issue is, how do we integrate methods and acknowledge related data to produce research findings and theoretical models that support social change? These considerations are related to theoretical and social research concerns.

Hesse-Biber (2010b) reminds us that central to qualitative research is the assumption that reality is socially constructed and contextually informed. Relatedly, methods can be grouped into constructivist-interpretive, critical (Marxist, emancipatory), and feminist theory. Of importance to the argument we present in this chapter is that collectively, these approaches emphasize that knowledge is constructed, embedded in power structures that privilege some voices and marginalize others. These structures are best challenged by understanding lived experience. At its very heart, qualitative research intends to address issues of social injustice. Accordingly, theory development emerging from qualitative research needs to honor this intent. Intentionally privileging qualitative research methods within mixed methods studies should therefore align with these theories. Hesse-Biber (2010b) further explains that mixed methods studies informed by qualitative approaches enhance and achieve this in various ways. First, it augments the representativity and generalizability of findings stemming from the research and enhances the validity and reliability of these findings overall. Additionally, qualitative approaches can explore contradictions found in quantitative data, while quantitative approaches can test the validity of qualitative findings in a broader population. Finally, qualitatively driven mixed methods studies can enhance our understanding of the research question and related findings and provide convergence in findings using triangulation.

The role of qualitative research in conceptualizing a mixed method study

The previous section focused on the theoretical implications for the research design when privileging qualitative approaches for mixed methods research. In this section, the implications for practice and conceptualization are discussed, including a quick

tour of the use of analysis software in this context. Much has been written about mixed methods analysis and the use of computer-assisted qualitative data analysis software for mixed methods research (Smit, 2021). However, less is said about the design of mixed methods research and privileging qualitative approaches from the outset and incorporating analysis software.

A decade ago, Hesse-Biber (2010a) wrote prolifically about the focus on qualitative approaches to mixed methods research (p. 16) and argued for a move toward interdisciplinary scholarship, broader theoretical lenses, and social change. She discussed an essential distinction between qualitative methodologies and qualitative methods, explaining that qualitative methodologies include qualitative and quantitative methods (p. 17), which shapes the design decision. A few years later, Plano Clark and Ivankova (2016, pp. 198–202) offered a helpful table of prominent philosophies for mixed methods, giving a clear overview of the epistemological and ontological assumptions and mapping these philosophies and assumptions to accompanying scholarly writings.

Many scholars followed Hesse-Biber (2018), Onwuegbuzie and Johnson (2021), Plano-Clark and Ivankova (2016), which is not an exhaustive list of mixed methods scholars. Scholars such as Leavy (2014) and Tracy (2020) have also added educational value to the nuanced complexity of privileging qualitative research in mixed methods research. Ritchie et al. (2014, p. 45) in this regard remind us “when using qualitative and quantitative research in combination it is important to recognize that each offers a different way of knowing about the world, and it should not be expected that the evidence generated from the two approaches will replicate each other. Rather, the benefits of mixing methods may result from the ability of each method to add to a more nuanced picture of complex social phenomena.”

Considering a research design, Tracy (2020, p. 27) invites researchers to “ask what types of methods are best suited for the goals of the research project?” She elaborates, “many people consider methods to be akin to tools. Like a hammer is a better tool than a screwdriver for banging a nail into a wall, qualitative methodology is better than a quantitative methodology for richly describing a scene or understanding the stories people use to narrate their lives” (p. 27). The toolkit idea or using what works in the educational context has been a prominent thought in mixed methods research, especially in relation to a pragmatist point of view (Tashakkori and Teddlie, 1998; Teddlie and Tashakkori, 2003). Furthermore, fieldwork is especially worthwhile for understanding behaviors and interactions that people do not talk about – either because these phenomena are unconscious, hard to remember, embarrassing, or simply mundane. When people describe the unique value of qualitative, compared to quantitative research methods, they often point to the ability of qualitative research to provide insight into why and how questions. Such research questions seek meaning qualitatively and often pursue the implicit or the unspoken if people cannot put into words what is important to them or why they acted in a certain way (Tracy, 2020, p. 100).

The research question shapes the research design and methodology. The concepts of research design and methodology are often used interchangeably. For this chapter, we discern between these two constructs as follows: research designs can also be considered as approaches or strategies, for example, case study designs, action research, phenomenology, ethnography, or narrative inquiry. The research methodology section covers sampling, data collection, data analysis, and matters of rigor and quality. Depending on what the research question wishes to answer and the decision to employ a mixed methods study, the researcher needs to decide how and why qualitative approaches are privileged and in what way.

When the research rationale addresses complex or cultural activities where the daily processes and activities are explored (Nastasi and Hitchcock, 2016), qualitative approaches will suitably design mixed methods research. The strength of qualitative research lies in its holistic, contextual, co-constructed interpreted meaning, including the reinterpretation of quantitative data. These multiple interpretations offer theoretically compelling suggestions for the research design and methods (Motulsky, 2021). Accordingly, qualitative methods is an umbrella concept that covers interviews (group or one-on-one), participant observation (in person or online), and textual analysis (paper or electronic). Such methods can include research in the field, a focus-group room, an office, or a classroom. Qualitative methods need not include long-term immersion in a culture or require a holistic examination of all social practices. One of the most common ways qualitative research is understood is through comparison with critical features of the quantitative method, where quantitative research can transform data into numbers, including conversations, actions, media stories, facial twitches, or other social or physical activity (Tracy, 2020, p. 26).

Tracy (2020, p. 378) also offers a word of caution: “Qualitative researchers face challenges in terms of securing grants, as many governmental agencies assume that the only sound research is objective, scientific, and quantitative. And even though funding agencies are increasingly providing grants to mixed methods researcher teams, qualitative methods are still grossly underrepresented in funding dollars. Grants typically reward research that is framed as efficient, immediately useful, and easily communicated. This marginalizes qualitative research – much of which is long-term, inconclusive by design, philosophical, and artistic. Furthermore, writing grant proposals takes a lot of time and energy and usually requires multiple submissions before they are funded. It is important to weigh carefully the pros and cons of spending time on writing a grant proposal when that precious time might instead be devoted to conducting and publishing a smaller unfunded study.”

Methodological implications for the research design

Mixed methods research has established itself as a third methodological paradigm (Collins et al., 2012). Scholars are, however, still grappling with some issues. One issue is rethinking how qualitative research could be featured more prominently. As a result, a systematic understanding of the context becomes possible (Nastasi and Hitchcock, 2016). Also, qualitative richness in terms of meaning and understanding that people bring to phenomena cannot be replicated in quantitative research. Hesse-Biber (2018) as well as Morse and Cheek (2015) clearly state that while qualitative research brings richness to the mixed methods study, the focus

on methods means that empirical or rather quantitative methods are prioritized from the design phase. It should also be noted that the design decisions are then extended to all aspects of the project from the theoretical perspective that links a research problem and approach to the particular method or methods of data collection and analysis.

Furthermore, qualitative research is a powerful tool for theory and model development (Leech and Onwuegbuzie, 2007). In many research projects, these are vital elements, and within the repertoire of qualitative techniques, several choices will promote rigor and yield credible findings (Nastasi and Hitchcock, 2016). Our point of departure is, as described by Bazeley (2013, p. 1) that “analysis is laid on the foundation of our understanding of how the world works, what it is (ontology); and how we understand and learn about the world and especially about the world of people (epistemology) ... although these foundations may remain implicit rather than explicit, thinking about them sharpens and enriches ... our understanding.”

There are several analysis techniques within qualitative research which would be appropriate in terms of educational research, although the dominant approach in educational contexts would be categorizing and theming verbatim data. The process entails condensing data into a theme that has emerged from the categorization data. However, other techniques are used to extend the role of qualitative data in mixed methods designs. Of course, the aim of incorporation of such data speaks to what Miles et al. (2014, p. 11) call “richness and holism, with strong potential for revealing complexity.” Given that analysis is not as well expanded upon, a brief overview is provided of the seven techniques described by Leech and Onwuegbuzie (2007) as the different techniques can be applied within a mixed analysis. When thinking about mixed methods analysis, a parallel approach may come to mind where data from the qualitative component of the study and the quantitative approach are kept separate and inferences are brought together but the data is not merged or converted. Cross-over or mixed analysis implies that one data form is transformed into another, quantitative data transformed to qualitative data or where qualitative data is transferred to quantitative data (Onwuegbuzie and Johnson, 2021).

In terms of data analysis specifically, there are various approaches that can be used as part of what Tracy (2020) would refer to as the toolkit. Leech and Onwuegbuzie (2008, p. 590) describe additional analysis techniques associated with a specific type of data (see Fig. 1). The most prominent analysis techniques, highlighted by Leech and Onwuegbuzie (2008):

- The first data analysis technique is constant comparison analysis, like what researchers would understand by theming data. The transcripts are read and chunked into meaning segments. The segment is then coded and compared with other codes. Similar codes are grouped under a theme.
- The second analysis technique is keywords-in-context, where the words are analyzed by comparing what comes before and after keywords. Keywords are identified as words that are either used frequently or used unusually.
- The third technique is word counts, where the number of times a specific word is used is counted, this type of analysis is generally used throughout the research process.
- The fourth analysis technique is content analysis, where the number of times a code is used is counted. The codes are usually produced deductively.
- The fifth analysis technique is domain analysis. Domain analysis is used where the researcher would like to identify larger units of cultural knowledge, where symbols represent a way of communicating cultural meaning. Domain analysis is followed by taxonomic analysis, where a classification system is developed to understand the relationships among domains. The type of analysis described by Leech and Onwuegbuzie (2007) is componential analysis, where attributes associated with cultural symbols are identified.

In terms of mixed methods, the rationale for including a qualitative component would be to improve quantitative designs, optimize sample sizes, instrument fidelity, treatment integrity, or maximize interpretations based on quantitative data (Leech and Onwuegbuzie, 2007). In mixed analysis, where qualitative and quantitative analytical techniques are used within the same framework (Onwuegbuzie and Johnson, 2021), researchers require competence in qualitative and quantitative analytical techniques. For this reason, researchers should also understand the range of analytical techniques and what can be accomplished. Of course, mixed analysis and cross-over analysis, as described earlier, would be appropriate in any mixed method study but is also highlighted in the article by Scherman et al. (2018). Creamer (2018) identifies five analytical procedures which may be appropriate in an integrated study where mixed analysis is used:

1. Blending: A variable, category, and theme created in one type of data is tested using another type of data or is created by combining qualitative and quantitative data.
2. Converting: Qualitative data are converted to quantitative data or vice-a-versa.
3. Extreme case sampling: Identifying and selecting the extremes or outliers.
4. Cross-case comparison: Consolidating qualitative and quantitative data to test or expand qualitative or quantitative derived themes.
5. Meta-inference: Link, compare or contrasts inferences generated by qualitative and quantitative strands.

A practical example of mixed analysis where qualitative data are highlighted can be found in Scherman et al. (2018), where a mixed monotype analysis in the article where qualitative data are quantitized (Onwuegbuzie et al., 2007; Onwuegbuzie and Johnson, 2021). Scherman et al. (2018) further highlight cross-validation where qualitative and quantitative analyses were used to examine the exact dimension of the research problem. The goal was to illustrate how a framework for developing causal pathways can be built to identify possible ways of working with parents, based on the people’s experiences (Dixon-Woods et al., 2005).



Fig. 1 Different analysis techniques for data sources (Leech and Onwuegbuzie, 2008, p. 590)

Bazeley (2012) states that integration can range from discussions of the different phases identified in the research to synthesizing, combining, or converting data from one form to another. Thus, for the analysis described in the article by Scherman et al. (2018), individual models of expectations and a general model were postulated. The article by Scherman et al. (2018) further shows that while the individual models are explained and nuanced in the qualitative data, a general model from the qualitative data could be put forward using quantitative analysis. Qualitative and quantitative approaches were required to understand the phenomenon (Bazeley, 2012), making the most of the collected data.

The article by Scherman et al. (2018) also described a dual analytic approach, and data were analyzed in one form and then manipulated into another form (Bazeley, 2012). This involved quantifying the qualitative codings, which adds to detecting and describing patterns in the data. However, given that sample sizes are typically small, it was essential to use the appropriate statistical analysis. The article thus shows how nonparametric can be used when the sample sizes are too small to meet the requirements of parametric statistics.

An additional consideration which is important is that of methodological norms. Scherman et al. (2018) show how norms for qualitative and quantitative research can be adhered to (member checking for qualitative research, content-related validity, and statistical conclusion validity for the quantitative component). Several mixed methods legitimization types has to be made explicit. The inside-outside legitimization is highlighted in which the researchers present and appropriately use the insider's and observer's view to deepen understanding, description, and explanation. Conversion is another prominent legitimization type, which is the focus of the article by Scherman et al. (2018), in which qualitative data is quantified to yield meta-inferences. Multiple validities as mixed methods legitimization are also included and understood as the extent to which addressing both quantitative and qualitative legitimization yields high-quality meta-inferences (Collins et al., 2012).

It is vital to be cognizant that answering the research question requires an appropriate means for making sense of data. The primary processes for analyzing quantitative data are descriptive and inferential statistics, while qualitative data analysis generally

relies on coding (O'Dwyer and Bernauer, 2014, p. 35). Whether we analyze numbers of words, we need appropriate processes and tools to make meaning from these data, and this usually involves some data reduction using statistics and coding.

The last thought relates to using software for qualitative research in mixed methods research. A few scholars have written about ATLAS.ti, NVivo, MAXQDA, QDA Miner, and Dedoose, (see Onwuegbuzie and Johnson, 2021). While these programs were developed initially for qualitative data analysis, scholars have reimagined these programs, together with new add-ons by the developers to include mixed methods analyses. This is indeed an exciting development, and researchers can now benefit from this computer assisted qualitative data analysis software (CAQDAS) given the variety of data types that can be imported or added to the software. Smit (2021, p. 331) wrote explicitly about the use of ATLAS.ti 22 for mixed analysis, which now is in version 22 with even more features supporting mixed data analysis. Also, the work by Saldaña (2021) fits comfortably in the use of software, discussing qualitative and mixed methods coding strategies. Smit (2021, p. 333) also elaborates on the use of ATLAS.ti for the literature review, which is seen as qualitative text analysis, using qualitative data analysis strategies (Onwuegbuzie and Leech, 2021). Literature reviews are essential for qualitative and mixed methods studies, and software such as Mendeley and Endnote interface comfortably with ATLAS.ti. Hesse-Biber (2010a) has reminded us back in 2010 how the introduction of CAQDAS has enabled researchers to begin to transform qualitative and quantitative data to suit their methodological purposes, a design consideration, which we cannot ignore.

Concluding thoughts

Mixed methods research is complex by its very nature (Poth, 2018). The same can be said of education and education systems. Very often, single studies cannot account for the complexity of the education system but with mixed methods research, a fuller picture can be provided. Furthermore, by privileging qualitative methodology, the educational context can be studied in a more meaningful way. In this chapter, we have highlighted key issues related to theorizing and making methodological decisions where qualitative methods can be prioritized. Hesse-Biber (2010a, p. 24) stated eloquently, "mixed methods can be an exciting process that might open our eyes to new realities and levels of understanding, but such a research design must be utilized only with the certainty that it will aid in advancing our theoretical understanding and not detract from our existing methodologies." We concur that "a qualitative approach privileges qualitative methods, with quantitative methods component playing an auxiliary role in a mixed methods framework (Hesse-Biber, 2010a, p. 64).

References

- Bazeley, P., 2012. Integrative analysis strategies for mixed data sources. *Am. Behav. Sci.* 56 (6), 814–828. <https://doi.org/10.1177/0002764211426330>.
- Bazeley, P., 2013. *Qualitative Data Analysis: Practical Strategies*. Sage, London.
- Creamer, C.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. Sage, London.
- Collins, K.M.T., Onwuegbuzie, A.J., Johnson, R.B., 2012. Securing a place at the table: a review and extension of legitimization criteria for the conduct of mixed methods research. *Am. Behav. Sci.* 56 (6), 849–865.
- Denzin, N.K., 2012. Triangulation 2.0. *J. Mix. Methods Res.* 6, 80–88. <https://doi.org/10.1177/1558689812437186>.
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., Sutton, A., 2005. Synthesizing qualitative and quantitative evidence: a review of possible methods. *J. Health Serv. Res. Pol.* 10 (1), 45–53.
- Fernandez, E., 2019. Working towards better education for children in care: longitudinal analysis of the educational outcomes of a cohort of children in care in Australia. *Oxf. Rev. Educ.* 45 (4), 481–501. DOI: <https://doi.org/10.1080/03054985.2019.1612345>.
- Ford, N., Odallo, D., Chorlton, R., 2003. Communication from a human rights perspective: responding to the HIV/AIDS pandemic in eastern and southern Africa. *J. Health Commun.* 8, 519–612.
- Gaunle, S., Adhikari, P., 2010. *Youth in Policymaking*. Sangam Institute, Kathmandu.
- Hesse-Biber, S., 2010a. *Mixed Methods Research. Merging Theory with Practice*. Guilford, New York.
- Hesse-Biber, S., 2010b. Qualitative approaches to mixed methods practice. *Qual. Inq.* 16 (6), 455–468. <https://doi.org/10.1177/1077800410364611>.
- Hesse-Biber, S., 2018. Toward an understanding of a qualitatively driven mixed methods data collection and analysis: moving toward a theoretically centered mixed methods praxis. In: Flick, U. (Ed.), *The SAGE Handbook of Qualitative Data Collection*. SAGE Publications Ltd, pp. 545–563. <https://doi.org/10.4135/9781526416070.n35>.
- Kanāiaupuni, S.M., 2005. Ka'akalai Ku Kanaka: a call for strengths-based approaches from a Native Hawaiian perspective. *Educ. Res.* 34, 32–38.
- Krenichyn, K., Schaefer-McDaniel, N., Clark, H., Zeller-Berkman, S., 2007. Where are young people in youth program evaluation research? *Child. Youth Environ.* 17, 594–615.
- Lather, P., 1991. *Getting Smart: Feminist Research and Pedagogy With/in the Postmodern*. Routledge, New York.
- Leavy, P., 2014. *The Oxford Handbook of Qualitative Research*. Oxford University Press, New York.
- Leech, N., Onwuegbuzie, A.J., 2007. An array of qualitative data analysis tools: a call for data analysis triangulation. *Sch. Psychol. Q.* 22, 557–584.
- Leech, N., Onwuegbuzie, A.J., 2008. Qualitative data analysis: a compendium of techniques and a framework for selection for school psychology research and beyond. *Sch. Psychol. Q.* 23, 587–604.
- Liebenberg, L., 2009. The visual image as discussion point: increasing validity in boundary crossing research. *Qual. Res.* 9, 441–467. <https://doi.org/10.1177/1468794109337877>.
- Mertens, D., 2017. Transformative research: personal and societal. *Int. J. Transformative Res.* 4, 18–24. <https://doi.org/10.1515/ijtr-2017-0001>.
- Mertens, D.M., 2021. Transformative research methods to increase social impact for vulnerable groups and cultural minorities. *Int. J. Qual. Methods* 20, 1–9. <https://doi.org/10.1177/16094069211051563>.
- Miles, M.B., Huberman, A.M., Saldaña, J., 2014. *Qualitative Data Analysis. A Methods Sourcebook*, third ed. Sage, Thousand Oaks, CA.
- Morse, J., Cheek, J., 2015. Introducing qualitatively-driven mixed-method designs. *Qual. Health Res.* 25 (6), 731–733. <https://doi.org/10.1177/1049732315583299>.
- Motulsky, S.L., 2021. Is member checking the gold standard of quality in qualitative research? *Qual. Psychol.* 8 (3), 389–406. <https://doi.org/10.1037/qup0000215>.
- Nastasi, B.K., Hitchcock, J.H., 2016. *Mixed Methods Research and Culture-Specific Interventions: Program Design and Evaluation*. Sage, London.
- O'Dwyer, L.M., Bernauer, J.A., 2014. *Quantitative Research for the Qualitative Researcher*. Sage, London.

- Onwuegbuzie, A.J., Slate, J.R., Leech, N.L., Collins, K.M.T., 2007. Conducting mixed analyses: a general typology. *Int. J. Mult. Res. Approaches* 1, 4–17. <https://doi.org/10.5172/mra.455.1.1.4>.
- Onwuegbuzie, A.J., Johnson, R.B., 2021. Mapping the emerging landscape of mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York, pp. 1–24.
- Onwuegbuzie, A.J., Leech, N.L., 2021. Qualitizing data. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York, pp. 141–150.
- O'Reilly, M., Kiyimba, N., 2015. *Advanced Qualitative Research. A Guide to Using Theory*. Sage, Los Angeles.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE Publications. <https://doi.org/10.4135/9781483398341>.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research. A Practical Guide to Integrative Thinking with Complexity*. SAGE.
- Potts, K.L., Brown, L., 2015. Becoming an anti-oppressive researcher. In: Strega, S., Brown, L. (Eds.), *Research as Resistance: Revisiting Critical, Indigenous, and Anti-oppressive Approaches*, second ed. Canadian Scholar's Press, pp. 17–41.
- Ritchie, J., Lewis, J., McNaughton Nicholls, C., Ormston, R., 2014. *Qualitative Research Practice*, second ed. Sage, London.
- Saldaña, J., 2015. *Thinking Qualitatively. Methods of Mind*. Sage, Los Angeles.
- Saldaña, J., Omasta, M., 2018. *Qualitative Research. Analyzing Life*. Sage, Los Angeles.
- Saldaña, J., 2021. *The Coding Manual for Qualitative Researchers*, fourth ed. Sage, Los Angeles.
- Scherman, V., Liebenberg, L., 2021. Item response theory integrating qualitative data. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York, pp. 117–123.
- Scherman, V., Zimmerman, L., Smit, B., 2018. Mixed-method data analysis: an exploratory approach to strengthening inferences about relationships and affinities. *Int. J. Mult. Res. Approaches* 10, 57–76. <https://doi.org/10.29034/ijmra.v10n1a4>.
- Senge, P.M., 2006. *The Fifth Discipline: The Art and Practice of Learning Organisations*. Currency Publishers, New York.
- Smit, B., 2021. Introduction to ATLAS.ti for mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York, pp. 331–341.
- Tashakkori, A., Teddlie, C., 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Sage Publications, London.
- Teddlie, C., Tashakkori, A., 2003. Major issues and controversies in the use of mixed methods in the social and behavioural science. In: Tashakkori, A., Teddlie, C. (Eds.), *The Handbook of Mixed Methods in the Social and Behavioural Research*. Sage Publications, London, pp. 3–50.
- Tracy, S., 2020. *Qualitative Research Methods. Collecting Evidence, Crafting Analysis, Communicating Impact*, second ed. Wiley Blackwell, Hoboken.
- Wallerstein, N., Duran, B., 2006. Using community-based participatory research to address health disparities. *Health Promot. Pract.* 7, 312–323.

Further Reading

- Creswell, J.W., Creswell, J.D., 2018. *Research Design. Qualitative, Quantitative & Mixed Methods Approaches*, fifth ed. Sage, Los Angeles.
- Onwuegbuzie, A.J., Johnson, R.B. (Eds.), 2021. *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, New York.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research. A Guide to the Field*. Sage, Los Angeles.

Relevant Websites – Optional

- <https://sites.rmit.edu.au/graduateresearcherknowhow/becoming-a-research-writer/?s=03>.
- <https://researchdesignreview.com/2021/11/29/member-checking-importance-context/?s=03>.
- <https://mmira.wildapricot.org>.
- <https://journals.sagepub.com/home/mmr>.
- <https://ijmra.org/>.
- <https://us.sagepub.com/en-us/nam/international-journal-of-qualitative-methods/journal202499>.

Mixed methods sampling and data collection

Nancy L. Leech and Courtney Donovan, University of Colorado Denver, Denver, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	485
Sampling in mixed methods research studies	485
Types of sampling schemes	486
Probability sampling	486
Purposive sampling	486
Convenience sampling	486
Mixed methods sampling	486
Data collection in mixed methods research studies	487
One sample, one time point	488
Benefits	488
Limitations	488
One sample, different time points	488
Benefits	488
Limitations	488
Different samples, different time points	489
Benefits	489
Limitations	489
Conclusion	489
References	489

Introduction

Quantitative and qualitative methods can be mixed at any stage of the research process: during data collection, data analysis, and/or data interpretation (Greene, 2007; Tashakkori et al., 2021). In many mixed methods research studies, mixing typically occurs at the level of interpretation whereas mixing is less common during data collection and data analysis (Onwuegbuzie and Hitchcock, 2015; Onwuegbuzie et al., 2009). When planning a mixed methods research study, it is important to think through where the mixing will occur (Collins et al., 2006), and to ensure that the mixing assists in answering the research questions (Onwuegbuzie and Leech, 2006). Two places to mix qualitative and quantitative methods in a mixed methods study are during the sampling procedure and data collection.

In a social science research study, sampling is the process where cases/participants/units are drawn from a larger population to participate in the study (Johnson and Christensen, 2020). It is important that the group selected matches the population and/or includes participants who can provide information or insights to answer the research question(s). In a mixed methods study, there are a myriad of sampling schemes available. After sampling, these cases/participants/units are invited to participate in the research study and information is collected from them (i.e., data collection is conducted). Yin (2006) stresses that quantitative and qualitative data collection methods need to ensure the data overlap and/or complement one another. This is a major challenge in most mixed studies, but many accomplish this via conceptual frameworks that tie the over-arching concepts together (Wheeldon, 2010). To better understand sampling and data collection in mixed methods research studies, each are outlined below.

Sampling in mixed methods research studies

Sampling is defined as “the process of selecting a subset or sample unit from a larger group or population of interest” (Collins, 2010, p. 356). In mixed methods research studies sampling is an important step as it can impact the types of statistical analysis (i.e., quantitative analysis), strengths of the transferability of findings (i.e., qualitative analysis), and inferences (i.e., mixed analysis) that can be drawn (Collins et al., 2007; Teddlie and Yu, 2007). Sampling in mixed methods studies involves “two distinct yet interrelated decisions: decide on the strategy to select the participants (i.e., sampling scheme) and decide on the number of participants (i.e., sample size)” (Collins, 2010, p. 354). Thus, sampling in mixed methods studies can be complex, with one or multiple types of sampling designs and with one or multiple samples (Collins, 2010). The samples can be nested (e.g., having one large sample and then sampling a smaller sample from the original sample), sequential (e.g., having one sample for the first phase of the study and another sample for second), or parallel (e.g., having one sample for the quantitative phase and another sample for the qualitative phase with the data collected at the same time).

Types of sampling schemes

In many mixed methods research studies, more than one type of sampling is conducted (Teddle and Yu, 2007). Traditionally, depending on the type of design (qualitative or quantitative), different types of sampling are conducted. In general, from a quantitative perspective, sampling is conducted through probability sampling and from a qualitative perspective, sampling is accomplished through purposeful sampling. Yet, this differentiation is artificial (Bernard, 1995; Guest et al., 2006): there are four types of sampling available for mixed methods researchers: (a) probability, (b) purposive sampling, (c) convenience sampling, and (d) mixed methods sampling (Teddle and Yu, 2007). Understanding each of these can assist the mixed methods researcher in choosing sampling schemes, thus probability sampling, purposive sampling, convenience sampling, and mixed methods sampling will be briefly described with an example from the extant literature where the type of sampling has been used.

Probability sampling

Historically utilized in quantitative designs, probability samples focus on using random sampling to find a group of participants that represent the population of interest. By using a probability sample, theoretically every person in the population has an equal opportunity to be sampled, which then enables the research to make generalizations from the sample to the population. When conducting probability sampling, before conducting the sampling procedure, the researcher assesses the approximate number of needed participants/units/groups by conducting a power analysis. Sampling is then conducted utilizing random sampling or a variant of random sampling [i.e., stratified sampling, cluster sampling, and sampling using two-stage or multistage probability techniques (Johnson and Christensen, 2020)]. Stratified sampling is when the sample is divided into groups based on a specific variable (e.g., area of town where they live) and then samples are drawn from the groups. Cluster sampling is

a two-stage probability sampling procedure that is especially useful when the population is spread out geographically and/or when there is no single overall list of individuals in the accessible population; individual participants within a given cluster are usually geographically grouped together; basic strategy is to first select specific clusters from the list of all clusters using a probability sampling method, then select all or randomly select a specific proportion of participants from the selected clusters.

Gliner et al., p. 484.

Finally, sampling using two-stage or multistage probability techniques is conducted by sampling in multiple stages. For example, in a two-stage sampling framework clusters may be selected (e.g., areas of town to be utilized in the research study) and then in stage two participants would be sampled from the selected clusters. Jensen et al. (2020) explored company's performance and CEO's transformational leadership skills using a random sample to select participants. From a publicly-listed set of US and European companies, 42 CEOs were selected.

Purposive sampling

A second type of sampling is purposive sampling (Mile and Huberman, 1994). When conducting purposive sampling, the researcher searches for participants that can provide unique or a depth of information (Maxwell, 1997; Sandelowski, 1995). With purposive sampling, based on theory and research question(s), a criterion/specific characteristic or parameters are defined to limit who is searched for in the participants. Purposive sampling is typically nonrandom, but it can be conducted via random sampling. Lawrence (2020) provides a good example of purposive sampling where they investigated "influential family factors for youth with learning and mental health concerns" (p. 166). In this study, two participants were purposively selected for the qualitative strand of the study. Their goal was to interview middle schoolers due to this age "reflects increasing academic demands and related distress at school (Feurer and Andrews, 2009), as well as heightened developmental patterns of internalizing problems" (p. 171).

Convenience sampling

Probably the most common type of sampling, convenience sampling is when an accessible sample and/or volunteer sample is obtained (Teddle and Yu, 2007). When using a convenience sample, the researcher does not select or choose who will participate. For example, if every person walking by a store is asked to complete a survey, this is convenience sampling. Simon (2020) provides an excellent example of convenience sampling in their study to investigate imposter syndrome experienced by Black women in a doctoral program in science, technology, engineering, and mathematics (STEM). Convenience sampling was utilized in this study after completing interviews of participants, "To substantiate the data collected from the analysis of the interviews, an 89 min focus group was held with a convenient sample of five students" (p. 190).

Mixed methods sampling

There are five types of mixed methods sampling (i.e., basic, sequential, concurrent, multilevel, and combination of mixed methods sampling). When using a mixed methods sampling scheme, participants/units/cases are selected utilizing both purposive and probability sampling. By using both types of sampling, the external validity and transferability of the inferences are increased (Teddle and Yu, 2007). McGill et al. (2020) provides an example of using mixed methods sampling. First, McGill et al. (2020) used purposive sampling to identify advisors at a major university. After collecting data from the purposively sampled advisors, these authors

used random sampling to select 10 participants to further interview. Thus, McGill et al. (2020) utilized purposive sampling and then probability sampling as a mixed methods sampling technique.

The many types of sampling are only tools, and these tools can be used interchangeably in qualitative, quantitative, and/or mixed designs (Onwuegbuzie and Leech, 2007). Based on this belief and utilizing the sampling identified by Patton (1990) and Miles and Huberman (1994), Onwuegbuzie and Leech (2007) outlined 24 sampling schemes that are available for mixed methods researchers. Table 1 presents the sampling scheme developed by Onwuegbuzie and Leech (2007) with a brief description of each as presented by Onwuegbuzie and Collins (2007).

Data collection in mixed methods research studies

What is unique about mixed methods research designs is that they are not limited by the type of data that researchers choose to collect (e.g., surveys, interviews, or observational data). Instead, the confines of data collection are procedural issues regarding how many samples are needed and when data is collected (Hanson et al., 2005). To outline the types of data collection processes available for mixed methods researchers, the following are delineated along with benefits and limitations of each: one sample, one time point; one sample, different time points; and different samples, different time points.

Table 1 Major sampling schemes in mixed methods research.

<i>Sampling scheme</i>	<i>Description</i>
Simple	Every individual in the sampling frame (i.e., desired population) has an equal and independent chance of being chosen for the study.
Stratified	Sampling frame is divided into sub-sections comprising groups that are relatively homogeneous with respect to one or more characteristics and a random sample from each stratum is selected.
Cluster	Selecting intact groups representing clusters of individuals rather than choosing individuals one at a time.
Systematic	Choosing individuals from a list by selecting every k th sampling frame member, where k typifies the population divided by the preferred sample size.
Multi-stage random ^a	Choosing a sample from the random sampling schemes in multiple stages.
Maximum variation	Choosing settings, groups, and/or individuals to maximize the range of perspectives investigated in the study.
Homogeneous	Choosing settings, groups, and/or individuals based on similar or specific characteristics.
Critical case	Choosing settings, groups, and/or individuals based on specific characteristic(s) because their inclusion provides the researcher with compelling insight about a phenomenon of interest
Theory-based	Choosing settings, groups, and/or individuals because their inclusion helps the researcher to develop a theory.
Confirming/disconfirming	After beginning data collection, the researcher conducts subsequent analyses to verify or contradict initial results.
Snowball/chain	Participants are asked to recruit individuals to join the study.
Extreme case	Selecting outlying cases and conducting comparative analyses.
Typical case	Selecting and analyzing average or normal cases.
Intensity	Choosing settings, groups, and/or individuals because their experiences relative to the phenomena of interest are viewed as intense but not extreme.
Politically important case	Choosing settings, groups, and/or individuals to be included or excluded based on their political connections to the phenomena of interest.
Random purposeful	Selecting random cases from the sampling frame and randomly choosing a desired number of individuals to participate in the study.
Stratified purposeful	Sampling frame is divided into strata to obtain relatively homogeneous sub-groups and a purposeful sample is selected from each stratum.
Criterion	Choosing settings, groups, and/or individuals because they represent one or more criteria.
Opportunistic	Researcher selects a case based on specific characteristics (i.e., typical, negative, or extreme) to capitalize on developing events occurring during data collection.
Mixed purposeful	Choosing more than one sampling strategy and comparing the results emerging from both samples.
Convenience	Choosing settings, groups, and/or individuals that are conveniently available and willing to participate in the study.
Quota	Researcher identifies desired characteristics and quotas of sample members to be included in the study.
Multi-stage purposeful random	Choosing settings, groups, and/or individuals representing a sample in two or more stages. The first stage is random selection and the following stages are purposive selection of participants.
Multi-stage purposeful	Choosing settings, groups, and/or individuals representing a sample in two or more stages in which all stages reflect purposive sampling of participants.

^aRepresent random (i.e., probabilistic) sampling schemes. All other schemes are non-random.

Reprinted with permission.

One sample, one time point

Some mixed methods designs involve data that is collected from a single sample (i.e., the same people) at a single time point. Typical examples of this include: (1) Conducting a survey with quantitative and qualitative items and (2) Conducting an interview with a scale or survey items asked within the interview protocol or in a separate survey taken at the same time as the interview. In this design the mixing takes place by connecting the two datasets via a sampling framework (Fetters et al., 2013). Steed et al. (2021) provides from the extant literature an example of a single sample at a single time point: the authors conducted a concurrent equal status fully mixed-method study (Leech and Onwuegbuzie, 2009) in which the participants completed an online survey that included both qualitative and quantitative data. Thus, quantitative and qualitative data were collected from a single sample at the same time point.

Benefits

The major benefit to this method of data collection is in saving time and money (Leeuw and Hox, 2008). Collecting data all at once in one time point is simply faster and more efficient for the researcher. Almalki (2016) also notes this technique tends to yield less data making it easier for researchers to manage and merge, which creates an efficient mixed method design.

When both types of data are collected together it offers the researcher more opportunities to mix the data (Leeuw and Hox, 2008). Data can be analyzed together using one to inform the other. One type of data could be converted and used within analyses of the other via quantitizing or qualitzing the data. Quantitizing is defined as “the process of converting [qualitative] data into numbers that can be statistically analyzed” (Tashakkori et al., 2021, p. 35). On the other hand, qualitzing is defined as “the process whereby [quantitative] data are transformed into narrative data that can be analyzed qualitatively” (Tashakkori et al., 2021, p. 35). Data can also be analyzed separately and mixed in interpretation through methods like triangulation. This is also more flexible as it is not hindered by which data was collected first.

This design also tends to be easy for participants since data is collected all at once and in a manner that feels natural for participants to answer (Driscoll et al., 2007). Additionally, it has been considered good practice to ask personal information in face-to-face, private settings to build relationships and trust with participants (Gliner et al., 2017) and is standard practice to ask sensitive, confidential information in either interviews or via surveys (Leeuw and Hox, 2008). This design allows for both methods to be collected at once.

Limitations

In comparison to the other methods, this design does not include an opportunity to follow up on the learning from one set of data based on the other dataset. Another major limitation of this design is there are measurement differences between the qualitative and quantitative data that are not easily addressed. Meaning, it is challenging to treat the qualitative findings equally to the quantitative findings as they are measured differently (Yin, 2006). Finally, this design encourages smaller sample sizes as most research teams are not able to collect qualitative information on a large sample. This create validity challenges for studies who are trying to generalize findings to a large population (Ihantola and Kihn, 2011).

One sample, different time points

Many mixed methods designs use a single sample but collect one set of data at one time and then follow-up at a later date with a same sample or a smaller portion of the original sample. Two classic examples are: (1) conducting a survey then following up with interviews from participants drawn from the survey results, and (2) Conducting focus groups with an organization and then sending a follow up survey. Here again, mixing takes place via a sampling framework (Fetters et al., 2013) and also by allowing one dataset to inform the other in analysis and interpretation stages (Onwuegbuzie et al., 2007). McGill et al. (2020) is an example of one sample at different time points. In this sequential explanatory mixed methods study, the authors sampled 57 full time academic advisors who were participating in a professional development (PD) program and asked them to provide information on what they learned from the PD. After analyzing these data, the authors wanted to investigate further, so they randomly sampled 10 participants from the initial 57 academic advisors. These 10 participants were then interviewed to better understand how the PD affected their advising.

Benefits

The major benefit of these data collection design is that one set of data can be analyzed to inform the second set of data (i.e., sequential mixed methods designs). This also allows for the qualitative data to help explain and validate the quantitative findings or vice versa. When qualitative data is collected first it can inform quantitative tools and help generate stronger hypotheses for the quantitative analysis (Fetters et al., 2013).

Limitations

The foremost disadvantage of this design is the time needed to design and conduct separate studies (Driscoll et al., 2007). Recruitment of participants can be challenging as you are limited in the second data collection by the first sample (Almalki, 2016). Since the data is collected from the same sample, quantitative sample size can be limited by the qualitative data collection capabilities (e.g., you might collect 200 survey responses but not be able to conduct 100 matched interviews). Due to the sequential design,

there is potential for temporal validity concerns where the time between data collection can become a confounding variable (Leeuw and Hox, 2008). For example, if one data set was collected before a major, unexpected event occurred, will the second set of data still be valid based on the first one's findings?

Different samples, different time points

An incredibly popular method of data collection in mixed designs is to collect one sample at one time point that is used to inform data collected from second sample at a second time point. Sequential designs are a great example where mixing occurs in the analysis and/or interpretation stages as one dataset influences the next. Examples include: (1) Interviews and focus groups are used to create constructs and items for a larger survey, and (2) An anonymous survey is given with focus groups conducted after to explore/explain survey findings which are pulled from the sample population, but not necessarily the same sample. Lawrence (2020) provides a good example different samples at different time points. In this study, the quantitative phase was conducted first with a data from 24,915 parents. To better understand the quantitative results the qualitative phase was then conducted with two families who were purposively selected.

Benefits

Like the previous design above, the benefit to this design is having one dataset to help design and/or explain the other dataset's findings. Unique to this design is that the sample is not limited for the quantitative data collection by the sample size restrictions in the qualitative data collection which makes external validity arguments stronger (Ihantola and Kihn, 2011).

Limitations

Again a major limitation for sequential designs is the time and resources needed to conduct two studies. Temporal validity concerns occur here too depending on how much time occurs between sampling. A unique challenge to this design is directly linking the datasets (Driscoll et al., 2007) together in a logical manner so the studies are truly mixed and not simply parallel or separate studies (Yin, 2006). Along the same line is the challenge is the two datasets may not be equal, even if pulled from the same population, which raises validity and reliability concerns (Ihantola and Kihn, 2011).

Conclusion

In the process of conducting a mixed methods study, two important steps are sampling and data collection (Collins et al., 2006). Using appropriate sampling techniques in mixed methods studies has multiple implications for researchers in education and beyond. Utilizing appropriate and well-thought through sampling and data collections methods, whether quantitative, qualitative, or mixed, can increase the legitimation, generalizability, and/or trustworthiness of the results (Onwuegbuzie et al., 2010). Furthermore, being thoughtful regarding the data collection procedures can ensure the data collected will answer the research questions (Onwuegbuzie and Leech, 2006). Paying close attention to these steps can increase the rigor of mixed methods studies.

References

- Almalki, S., 2016. Integrating quantitative and qualitative data in mixed methods research—challenges and benefits. *J. Educ. Learn.* 5 (3), 288. <https://doi.org/10.5465/ambpp.2016.14453abstract>.
- Bernard, H.R., 1995. *Research Methods in Anthropology*. AltaMira, Walnut Creek, CA.
- Collins, K.M.T., Onwuegbuzie, A.J., Sutton, I.L., 2006. A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. *Learn. Disabil.* 4 (1), 67–100.
- Collins, K.M.T., Onwuegbuzie, A.J., Jiao, Q.G., 2007. A mixed methods investigation of mixed methods sampling designs in social and health science research. *J. Mix. Methods Res.* 1 (3), 267–294.
- Collins, K.M.T., 2010. Advanced sampling designs in mixed research: current practices and emerging trends in the social and behavioral sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 353–377.
- Driscoll, D.L., Appiah-Yeboah, A., Salib, P., Rupert, D.J., 2007. Merging qualitative and quantitative data in mixed methods research: how to and why not. *Ecol. Environ. Anthropol.* <https://digitalcommons.unl.edu/icwdmeea/18>.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs-principles and practices. *Health Serv. Res.* 48 (6pt2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>.
- Feurer, D.P., Andrews, J.J.W., 2009. School-related stress and depression in adolescents with and without learning disabilities: an exploratory study. *Alta. J. Educ. Res.* 55 (1), 98–108. <https://cjc-rc.ualgany.ca/index.php/ajer/article/view/55276>.
- Glaser, J.A., Morgan, G.A., Leech, N.L., 2017. *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis*, third ed. Taylor and Francis, New York, NY.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Wiley, San Francisco, CA.
- Guest, G., Bunce, A., Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18 (1), 59–82. <https://doi.org/10.1177/1525822X05279903>.
- Hanson, W.E., Creswell, J.W., Clark, V.L., Petska, K.S., Creswell, D.J., 2005. Mixed methods research designs in counseling psychology. *J. Counsel. Psychol.* 52, 224–235. <https://doi.org/10.1037/0022-0167.52.2.224>.
- Ihantola, E., Kihn, L., 2011. Threats to validity and reliability in mixed methods accounting research. *Qual. Res. Account. Manag.* 8 (1), 39–58. <https://doi.org/10.1108/1176609111124694>.
- Jensen, M., Potočník, K., Chaudhry, S., 2020. A mixed-methods study of CEO transformational leadership and firm performance. *Eur. Manag. J.* 38, 836–845.

- Johnson, R.B., Christensen, L., 2020. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, seventh ed. Sage, Thousand Oaks, CA.
- Lawrence, B.C., 2020. Bridging secondary survey data with in-depth case studies to advance understandings of youth learning and mental health concerns. *Int. J. Mult. Res. Approaches* 12, 166–184. <https://doi.org/10.29034/ijmra.v12n2a1>.
- Leech, N.L., Onwuegbuzie, A.J., 2009. A typology of mixed methods research designs. *Qual. Quant.* 43 (2), 265–275. <https://doi.org/10.1007/s11135-007-9105-3>.
- Leeuw, E.D., Hox, J., 2008. Mixing data collection methods: lessons from social survey research. In: Bergman, M.M. (Ed.), *Advances in Mixed Methods Research: Theories and Applications*. Sage, Thousand Oaks, CA, pp. 138–149.
- Maxwell, J., 1997. Designing a qualitative study. In: Bick, L., Rog, D.J. (Eds.), *Handbook of Applied Social Science Methods*. Sage, Thousand Oaks, CA, pp. 69–100.
- McGill, C.M., Heikkilä, M., Lazarowicz, T., 2020. Professional development, performance expectations and academic advisors' perceptions of relational skills: a sequential explanatory mixed methods study. *New Horiz. Adult Educ. Human Dev.* 32 (4), 50–69. <https://doi.org/10.1002/nha3.20296>.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, second ed. Sage, Thousand Oaks, CA.
- Onwuegbuzie, A.J., Collins, K.M.T., 2007. A typology of mixed methods sampling designs in social science research. *Qual. Rep.* 12 (2), 281–316.
- Onwuegbuzie, A.J., Hitchcock, J.H., 2015. Advanced mixed analysis approaches. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *Oxford Handbook of Multiple and Mixed Methods Research*. Oxford University Press, pp. 275–395. <https://doi.org/10.1093/oxfordhb/9780199933624.013.19>.
- Onwuegbuzie, A.J., Leech, N.L., 2006. Linking research questions to mixed methods data analysis procedures. *Qual. Rep.* 11 (3), 474–498.
- Onwuegbuzie, A.J., Leech, N.L., 2007. A call for qualitative power analyses: considerations in qualitative research. *Qual. Quant.* 41 (1), 105–121. <https://doi.org/10.1007/s11135-005-1098-1>.
- Onwuegbuzie, A.J., Slate, J.R., Leech, N.L., Collins, K.M.T., 2007. Conducting mixed analyses: a general typology. *Int. J. Mult. Res. Approaches* 1 (1), 4–17.
- Onwuegbuzie, A.J., Slate, J.R., Leech, N.L., Collins, K.M.T., 2009. Mixed data analysis: advanced integration techniques. *Int. J. Mult. Res. Approaches* 3, 13–33.
- Onwuegbuzie, A.J., Johnson, R.B., Collins, K.M.T., 2010. A framework for assessing legitimacy in mixed research: implications for the field of stress and coping. In: Collins, K.M.T., Onwuegbuzie, A.J., Jiao, Q.G. (Eds.), *Research on Stress and Coping in Education. Toward a Broader Understanding of Stress and Coping: Mixed Methods Approaches*. IAP Information Age Publishing, Scottsdale, AZ, pp. 3–29.
- Patton, M.Q., 1990. *Qualitative evaluation and research methods*, 2nd. Sage.
- Sandelowski, M., 1995. Sample size in qualitative research. *Res. Nurs. Health* 18 (2), 179–183. <https://doi.org/10.1002/nur.4770180211>.
- Simon, M., 2020. STEMming within a double minority: how the impostor syndrome affects black women Ph.D. students. *Int. J. Mult. Res. Approaches* 12, 185–201.
- Steed, E., Leech, N.L., Phan, N., Charlifue-Smith, R., 2021. Remote delivery of services for young children with disabilities during COVID-19 school closures in the US. *J. Early Interv.* 1–20. <https://doi.org/10.1177/10538151211037673>.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Teddlie, C., Yu, F., 2007. Mixed methods sampling: a typology with examples. *J. Mix. Methods Res.* 1 (1), 77–100. <https://doi.org/10.1177/1558689806292430>.
- Wheeldon, J., 2010. Mapping mixed methods research: methods, measures, and meaning. *J. Mix. Methods Res.* 4 (2), 87–102. <https://doi.org/10.1177/1558689809358755>.
- Yin, R.K., 2006. Mixed methods research: are the methods genuinely integrated or merely parallel? *Res. Sch.* 13 (1), 41–47.

Strategies to integrative mixed methods analysis

Susanne Vogl, University of Stuttgart, Department of Social Sciences, Stuttgart, Germany

© 2023 Elsevier Ltd. All rights reserved.

The integration challenge in mixed methods research	491
Mixed methods data analysis	491
Integrative mixed methods analysis	492
Typologies and classifications	493
Linking data	494
Data transformation	494
Data consolidation	496
Empirical example for integrative mixed data analysis	496
Conclusion: challenges and implications for future research	497
References	498

The integration challenge in mixed methods research

Mixed methods research (MMR) is defined as the process of combining different methodological approaches—usually labeled as qualitative and quantitative—within one overarching study. Many different types of combinations are possible at various steps in the research process and for a number of purposes. In simplistic terms, the key argument for MMR is that multiple investigative strands increase or improve the understanding of a phenomenon (Creamer, 2022). After debates on paradigms and compatibility in the early phases of the “mixed methods era” as well as the development of typologies for MMR designs, the focus has shifted toward more practical strategies of working with data (Baur et al., 2017; Bazeley, 2010). Within the MMR community, there is growing sensitivity regarding the practical and conceptual challenges in meeting the expectations of MMR to gain synergies and complementarities by using both qualitative and quantitative data methods. As expressed by Bryman (2007, p. 8), “the key issue is whether in a mixed methods project, the end product is more than the sum of the individual quantitative and qualitative parts.” This goal can only be achieved through integration.

Data integration means that different sources of data are used to construct a joint product through the application of various analytical strategies (Li et al., 2000). These strategies deal with elements in the source data in different ways, but each strategy aims for the same research objective. Integration asks about the relatedness of different data, or the degree of “mutual illumination” (Bryman, 2007, p. 8), and it occurs to the extent that different data elements become interdependent in the course of reaching a theoretical or research goal (Bazeley, 2010). The product of integration is more than the sum of its constituent parts (Fetters and Freshwater, 2015)—it is the creation of a new whole (e.g., Onwuegbuzie and Hitchcock, 2019). In contrast, *data combination* is an “additive process” in which different types of data are mixed, but the results can still be traced back to each of the original data sources.

Many researchers consider integration to be the hallmark of mixed methods research (Moseholm and Fetters, 2017), as well as “the heart and soul of mixed methods research” (Guetterman et al., 2020, p. 430). Despite integration having great potential in MMR, it often remains underdeveloped and undertheorized (Bazeley, 2010). Qualitative and quantitative analyses are frequently conducted separately, and the results are only integrated in the overall interpretation and meta-inferences. Thus, addressing the “integration challenge” (Fetters and Freshwater, 2015) is the key challenge for the future of MMR.

While integration can occur in all steps of a research design, it seems to be particularly relevant but also the most challenging during data analysis (Onwuegbuzie and Combs, 2010). In this contribution, I focus on data analysis in MMR, specifically with integration in mind. Integrative analysis can foster better, more comprehensive results and a deeper understanding of a phenomenon. However, methodological and conceptual challenges, such as different sample sizes, violation of assumptions for statistical tests, and integrating data, highlight the need for innovative strategies (Fakis et al., 2014; Niglas et al., 2008) to enable achieving the full potential of MMR. A scarcity of good examples and the lack of literature about the research process and techniques of integration (Woolley, 2009) make integrative data analysis strategies a fruitful area for further research.

In the following section, I will provide an overview of the specificities of mixed methods data analysis. Afterward, I will present different data integration strategies and illustrate them through the use of an empirical example. In the conclusion, I will summarize and reflect on data integration to show that the distinction between qualitative and quantitative data is oversimplified. This dichotomy can hamper integrative thinking and practice.

Mixed methods data analysis

Generally, data analysis fulfills the functions of reducing and organizing raw data to make the information manageable. Data analysis can reveal patterns of relations, connections, trends, and differences. *Mixed methods data analysis* covers qualitative and

quantitative analytic strategies under the same or a similar framework and entails the analysis of qualitative *and* quantitative data. The data could stem from sequential, parallel, or iterative processes in one or more phases of data collection (Onwuegbuzie and Hitchcock, 2015). Further, mixed methods analyses can also be performed sequentially, in parallel, or iteratively with different objectives and purposes (see for example Vogl, 2019). The analyses could consist of multiple steps or even strands. Usually, data from each strand are analyzed with “appropriate” methods (Creswell and Plano Clark, 2018). The interfaces (Creswell and Plano Clark, 2018) and the extent to which the qualitative and quantitative strands interact in the analysis depend on the research design and the research question.

Mixed methods analyses are not an end in themselves but must reflexively relate back to the research question. The research question determines how and to what extent different data are integrated during analysis and interpretation (Brannen and O’Connell, 2015). Researchers must have a clear vision of what “dynamic mixes they suggest or permit” (Sandelowski, 2000, p. 254) and to what end. The researcher’s vision shapes the combined use of qualitative and quantitative techniques and its purpose.

The *purposes of mixed methods analysis*, which are analogous to design and data collection in MMR, can include the following (Caracelli and Greene, 1993; Greene et al., 1989; Onwuegbuzie and Hitchcock, 2015; Schoonenboom and Johnson, 2017):

- *triangulation/corroboration/convergent validation*: results based on one method are confirmed by those gained through the use of another method (e.g., when qualitative typology building is compared with latent class modeling)
- *complementarity/analytic density*: qualitative and quantitative results are considered distinct but generate complementary insights that together create a bigger picture (e.g., when qualitative data add a more nuanced understanding of a phenomenon that is also assessed with quantitative measures)
- *development*: the analysis of one data type informs the generation of another (e.g., when a qualitative typology informs a classification for statistical analysis)
- *initiation*: the use of a first method sparks new hypotheses or research questions that can be pursued using a different method (e.g., when findings in qualitative data inform statistical modeling)
- *elaboration/expansion*: the use of one type of data analysis adds to the understanding gained by another (e.g., when quantitative data contain information that adds to qualitative data, such as some psychological measures complementing biographical analysis).

It is important to mention that these purposes are neither exclusive nor unalterably set at the beginning of a study; they can emerge as research progresses.

Integrative mixed methods analysis

After this brief introduction to MMR and mixed methods analysis, I will focus on integrative mixed methods analysis. Data integration in mixed analysis refers to finding relationships between different data such that a new whole is created. In the literature, different terms are used for relating various strands of mixed methods data: mixing, combining or connecting, and integrating and merging. Combining and connecting refers to an additive process in which the original source data can still be traced ($1 + 1 = 2$), while integrating and merging creates a new element to the data beyond the qualitative and quantitative sources ($1 + 1 = 3$; Fetters and Freshwater, 2015), or $1 + 1 = 1$ when the result of full integration cannot be distinguished into separate parts but becomes a new whole; e.g., Onwuegbuzie and Hitchcock, 2019). Data combination means that the types of data are largely independent, but for integration, the types of data are either highly or largely dependent on one another (van Velzen, 2016; Woolley, 2009). Mixing seems to be the most general and least consistently used term. Some researchers regard mixing as being synonymous with integration and merging (Johnson et al., 2017), while others view merging as a form of mixing (Collingridge, 2012). I suggest using the term mixing as an overarching concept that encompasses combination and merging/integration.

Integrated mixed methods data analysis not only involves the analysis of qualitative and quantitative strands—it also entails an analysis of the interaction between the two strands. This is consequential for the workload and required skills. Kuckartz (2014) draws an illustrative parallel between integrative analysis and families: The difference between having one child in a family and having two children is more than a simple doubling of the number of children; it also includes the interaction between the siblings, which may be anything but trivial (Kuckartz, 2014). Through integration, results have the potential to produce insights that would not be possible with a mere combination of data. The higher workload is “rewarded” with a higher level of “sophistication” in the results—at least this is the idea.

In terms of *purpose*, integrative data analysis strategies are particularly well suited for initiation, but also for expansion and development. If the purpose of a mixed methods study is triangulation—in the sense of researching the same topic from different perspectives in order to increase validity—analytical strategies have to be independent of each other (Caracelli and Greene, 1993; Onwuegbuzie and Combs, 2010). In this situation, integration is ill suited to the task. For triangulation, results from different data are compared and the conclusions seem more robust when all data types converge. This desired outcome implies that data have to be kept separate in order to compare the findings from different data for validity. However, although the purpose can be formulated ex ante, the realization of the purpose of mixing or integrating depends on the actual data; the purpose can emerge and change

in the process of a study. For example, although mixed analysis is intended to generate complementary results by researching different facets of the same phenomenon with different methods, the results might in practice contradict each other, precluding their integration into a bigger picture. Furthermore, when data analysis entails multiple steps or strands, different purposes can be pursued in each, depending on the respective strategy.

Generally, the starting points for integrative mixed methods analyses can be at one or more of the following four points in a project: from various data sources or elements in their more or less raw form; from data that have been preprocessed (e.g., undergone coding or indexing); during analyses, when different styles of analysis interact; or during writing when integration emerges from the writing process (Bazeley, 2018b). In the simplest form of integration, illustrative quotes from qualitative interviews complement or supplement results from statistical analyses. Although this approach can assist in communicating statistical results, “this type of integration strategy is quite limited” (Bazeley, 2012, p. 817). The highest level of integration is probably achieved when preprocessed data are integrated or when analytic styles inform and interact with each other.

Typologies and classifications

Parallel to the classification systems of mixed methods designs, there are also classifications of integrative mixed methods analysis. Commonly cited classifications include those by Creswell and Plano Clark (2018) and Caracelli and Greene (1993). Creswell and Plano Clark (2018, pp. 323–326) defined four basic integrative processes.

1. The first process involves *merging the results of analyses* of the quantitative and qualitative data sets to compare results in a narrative discussion, which means data as such are not merged, only results. Thus, integration happens more on an interpretation level than during analysis. This process assumes that analysis and interpretation are distinct steps, which may be more appropriate for quantitative research. In particular, integrative analysis is often an iterative process with analysis and interpretation going hand in hand. Thus, differentiating between data analysis and interpretation as separate steps in the research process, can be—depending on the methods employed—problematic because no data analysis is free of interpretation and ultimately all data are analyzed qualitatively “in so far as the act of analysis is an interpretation, and therefore of necessity a selective rendering, of the ‘sense’ of the available data” (Bazeley, 2018b, p. 57; see also Akremi et al., 2018; Baur and Knoblauch, 2018; Fielding and Fielding, 1986; Maxwell, 2010; Sandelowski et al., 2009).
2. The second process involves merging analyses through *data transformation*, which implies the merging of actual data. This requires one data type being transformed into another (e.g., text into numbers or vice versa).
3. The third process connects the analysis of quantitative data to the collection of qualitative data, with the latter being used to explain the former. This process represents a *sequential, explanatory model of integration*.
4. Finally, the fourth process connects the analysis of qualitative data to the collection of quantitative data, with the latter being used to build upon the former (Bazeley, 2018b, p. 62). In other words, the process is a *sequential exploratory mixed methods analysis*.

These four processes are not distinct; we notice that 1 and 2 imply strategies, whereas 3 and 4 focus on the sequence of the collection and analysis of data from different strands with no specific strategy for integration. Furthermore, one could argue that the first process is not an integrative analysis but rather a combination of interpretations. Overall, the degrees of integration and the relevance of analysis and interpretation for integration vary between these types.

Caracelli and Greene (1993) identify four integrative strategies:

1. *Data transformation*: Transformation strategies refer to the conversion of one type of data into another to allow for statistical or thematic analyses (Caracelli and Greene, 1993; Tashakkori and Teddlie, 2010), such as numerical coding or ranking of qualitative data to statistically analyze them together with quantitative data. Conversion strategies are also called “quantitizing” and “qualitizing.” Quantitizing means assigning numeric values to nonnumeric data, and qualitizing means converting quantitative data into narrative representations (Jang et al., 2008; Teddlie and Tashakkori, 2009; van Velzen, 2016). Transformation facilitates integration because different types of data can be analyzed jointly.
2. *Typology development*: The aim of developing a typology is to construct types or obtain substantive categories that allow the analysis of homogeneity within and heterogeneity between subgroupings of data. The typology is incorporated into the analysis of the respective other type of data. The typology development can be an iterative process. Results from one analysis can feed into the analysis of another form, which then feeds back into the reanalysis of the first data type. For example, dimensions identified by a factor analysis could be added to the coding of qualitative data, or typology developed with qualitative data could be included as a variable in statistical analyses and results mixed with the respective first type of data.
3. *Extreme case analysis*: Analyzing one type of data serves to identify extreme cases that can then be tested and refined with the other data type. Extreme cases are identified through analyzing one data type and further investigation of other types of data (Caracelli and Greene, 1993). The extreme cases can also guide further data collection. Thus, this approach also has the potential for iteration.
4. *Data consolidation/merging*: The joint use of both data types leads to consolidated variables or whole data sets in a qualitative or quantitative form that are comprehensively analyzed. This strategy is often considered the highest form of integration. It is well suited for initiation purposes when MMR serves the generation of new insights and perspectives.

A reviewing of this typology reveals that these four types are not necessarily distinct. Rather, they address the aim (2 and 3) and practice (1 and 4) of mixed methods analysis: Consolidation and transformation could serve for extreme case analysis or typology development. Furthermore, a practical question centers on how consolidation can be achieved. Presumably, it would contain some kind of transformation, making transformation a necessary but not sufficient step for consolidation. Furthermore, “transformation” and “merging” occur in both Creswell and Plano Clark’s and Caracelli and Greene’s typology. However, transformation was a distinct strategy for Caracelli and Greene, whereas Creswell and Plano Clark identified merging (type 4 in the Caracelli and Greene typology) as being achieved through transformation (Caracelli and Greene type 1).

Onwuegbuzie and Teddlie (2003) introduced a more process-oriented model for mixed methods analysis. They distinguished seven steps during data analysis: *data reduction*, *data display*, *data transformation*, *data correlation*, *data consolidation*, *data comparison*, and *data integration*. They extended the Caracelli and Greene typology, but again, processes are not distinct and they potentially overlap. Other approaches are more closely linked to research designs; for example, Moseholm and Fetters (2017) identified five types of data integration via merging in convergent designs. They introduced a directional dimension of merging and distinguished between *explanatory unidirectional*, *exploratory unidirectional*, *simultaneous bidirectional*, *explanatory bidirectional*, and *exploratory bidirectional*. This approach advances terminological discussions but has shortcomings regarding actual integration procedures.

Onwuegbuzie and Combs (2010) took a different approach; instead of a typology, they aimed for a conceptual framework for mixed (methods) analyses. They identified 13 aspects that can serve to delineate possibilities or decisions in mixed methods analysis. Their framework is derived from a content analysis of publications on typologies of mixed methods analyses, and it maps practices used in them. The aspects of the framework are (1) rationale/purpose for conducting mixed methods analysis; (2) philosophy underpinning the analysis; (3) number of data types that will be analyzed; (4) number of data analysis types that will be used; (5) time sequence of the mixed methods analysis; (6) level of interaction between quantitative and qualitative analyses; (7) priority of analytical components; (8) number of analytical phases; (9) link to other design components; (10) phases of the research process when all analysis decisions are made; (11) type of generalization; (12) analysis orientation; and (13) cross-over nature of analysis (Onwuegbuzie and Combs, 2010, p. 411). The advantage of this framework relative to a typology is that it is more flexible but outlines the important aspects of any mixed methods analysis.

In sum, despite a consensus on the value of integrating qualitative and quantitative approaches during data analysis, methodological reflection and practical applications remain open to further exploration. The scholarly debate lacks a common conceptual foundation, which leads to terminological confusion regarding types of designs—a confusion which, as suggested by Kelle (2007), applies to design terminologies in general. Typologies try to formalize approaches to integration and convey a sense of rigor and guidance but can also limit creative thinking. Therefore, examples of strategies for integrated mixed methods analysis seem more fruitful than overarching typologies.

I will now focus on three important integrative analysis strategies also mentioned in classifications mentioned above: data linkage, transformation, and consolidation. These three strategies entail different levels of integration. At the end of this section, I will use an example from my research that employed all three strategies (Vogl, 2019) to illustrate the implementation.

Linking data

Linking data refers to the combination of data (elements) through association, comparison, or relational analyses. Matching data for individual cases can serve comparative, additive, and/or relational (theory-building) purposes. Linking data thematically instead of by methods used to gather or analyze data is an effective means to add value to analyses and enhance integration and improve communication of results (Bazeley, 2016). Matching and comparing data within and across cases reveals the relevance of grouping variables and group patterns, and it can point to reasons for subgroup differences (Bazeley, 2018a). Group differences can be detected more easily because with linked, complementary data on a case level, “a more nuanced account of how the different data corroborate, illustrate, or elaborate each other is possible” (Bazeley, 2018b, p. 127). In addition, individual cases that diverge from the overall patterns are more easily detectable. Differences as such are still the same, but with linked data, they are easier to detect. Different sources can work together to constitute a complementary account for rich descriptions of social phenomena or to reveal differences and patterns across subgroups that might otherwise be obscured. A combined case data matrix can facilitate connections and contrasts for individual cases and across the whole sample to identify patterns but also to detect deviant cases (Bazeley, 2018b).

Data transformation

Data transformation refers to the conversion of one type of data into another to enable (a) applying a second type of analysis to the same set of data or (b) conducting a joint analysis of data from different origins. Teddlie and Tashakkori (2009) label the mixed methods design that utilizes these transformation approaches as “conversion design” (in contrast to sequential and parallel designs). The purposes of data transformation are extended understanding (Caracelli and Greene, 1993), complementarity, and initiation because consistencies and inconsistencies can be detected (Li et al., 2000, p. 126).

One approach to integrated mixed methods analyses are “cross-over analyses” or “cross-over-tracks analyses” (Li et al., 2000; Onwuegbuzie and Combs, 2010; Teddlie and Tashakkori, 2009). This is an inter-paradigmatic approach in which qualitative and quantitative data are analyzed concurrently and the analysis moves back and forth continuously. Cross-over-tracks analysis

intertwines findings from various methodological strands whereby they exchange information throughout the study. At least one type of analysis from one tradition is used to analyze data generated under the other tradition—that is, qualitative data are analyzed with quantitative strategies and vice versa.

Cross-over analysis requires some form of data transformation to make the data accessible for the analysis using a tradition other than the one under which the data were collected. The most common transformation strategy is the quantitizing of qualitative data. However, qualitzing is also possible in that quantitative data can be transformed into narratives and potentially analyzed together with qualitative data. Qualitized numeric data can offer further information or allow for cross-checking interpretations. For example, individual values on a scale can be the basis for creating narrative profiles or a typology (Bazeley, 2012; Sandelowski, 2000). Without calling it qualitzing, a sample description follows this logic. Profiles contain information about the most common characteristics or their average, and groups from within a population can be identified using cluster or factor analyses and described verbally (Teddlie and Tashakkori, 2009). In a longitudinal study, the narrative potential increases through event analyses and patterns, and sequences can guide the categorization of individuals or groups (Bazeley, 2012). A complex form of qualitzing can be found in Bourdieu's (2007) work. Based on quantitative indicators, he maps the social space and interprets this visualization qualitatively to describe social milieus. Nevertheless, qualitzing as the process of converting numerical data into data that can be analyzed qualitatively is rare, and the literature on this approach is scant (Onwuegbuzie and Leech, 2021).

Quantitizing qualitative data implies converting qualitative data into numerical data, and it is a more frequent transformation strategy. For example, an exploratory factor analysis determines thematic structures from qualitative analysis and reduces the dimensionality of qualitative data, allowing the thematic structure to be represented through a numeric coding scheme (Onwuegbuzie, 2003). In a next step, the relation and temporal adjacency of themes can be explored through multidimensional scaling and social network analysis (Christopoulos and Vogl, 2014). As another example, qualitative interview content can be transformed into numerical ratings representing evaluations of an issue. These ratings are integrated into a quantitative data set on personality measures to be analyzed together. In this way, the relations between personality traits and evaluations of an issue can be analyzed (Larner et al., 2018).

Quantitizing is a simplification that can help make large data sets more manageable, facilitate comparisons with other data sources, or allow statistical analysis (Sandelowski et al., 2009, p. 210). Thus, patterns and peculiarities that are hidden or not otherwise detectable—particularly in large data sets—can be detected by applying statistical techniques that add power and sensitivity to individual judgment (Bazeley, 2010). The most frequently used statistical procedures for quantitized data are the chi-squared test, Fisher test, Mann-Whitney U-test, bivariate correlation, or analyses of variance (Fakis et al., 2014). However, with these procedures, the complex relations inherent to qualitative data cannot be easily explained or interpreted. Multivariate analyses such as cluster analysis, multivariate regression analysis, multi-level analyses, or structural equation models would be better suited to identifying multifaceted connections in qualitative data (Fakis et al., 2014). Furthermore, nonparametric tests or multidimensional scaling and cluster or correspondence analysis (Greenacre and Blasius, 2006) pose fewer requirements on data and can facilitate exploratory analyses. Statistical techniques producing visual displays of relationships can assist in interpretive analysis—as can be seen in correspondence analysis and the mapping of social distinction (see for example Blasius and Friedrichs, 2008).

As with any data reduction process, quantitizing data is—like any categorization—a trade-off between retaining detail and managing a complex data set (Bazeley, 2018b). Some interpretive researchers condemn the use of numbers in qualitative research. Nevertheless, we frequently find quasi-quantifications such as many, often, and rarely in qualitative studies despite their limited precision and value. However, using counts of themes or codes to enhance the presentation of results or to gauge the strength of a phenomenon has a long tradition (Tashakkori and Teddlie, 2003). Ultimately, counting is a form of description and frequencies can indicate (comparative) relevance. Simple counts of themes or frequencies of codes can complement qualitative results. Quantifications cannot replace qualitative analysis—but they can be complementary or instigate new hypotheses about relationships advancing theoretical thinking.

However, numbers have a cultural and rhetorical appeal that can suggest greater generality for the conclusions than is justified. They “can be used rhetorically, to make a report appear more precise, rigorous, and scientific, without playing any real role in the logic of the study and thus misrepresenting the actual basis for the conclusions” (Maxwell, 2010, pp. 479–480). Furthermore, there is a danger of reducing evidence to the amount of evidence. In addition, compromises, judgments, and foundational assumptions in quantitizing qualitative data are typically glossed over. Sandelowski et al. (2009, p. 208) summarize the assumptions as qualitative and quantitative data constituting two kinds of data, quantitizing as a unidirectional process essentially different from qualitzing, and counting as an unambiguous process. But it is an active decision what and how to count and balancing numerical precision with narrative complexity requires compromises.

To counter these problems, researchers “must have a clear sense of why they want to quantitize at all. They must consider what the added value is of transforming qualitative into quantitative data” (Sandelowski et al., 2009, p. 219). Quantitizing should meet some criterion such as clear definition of codes and consistency in coding across the entire data set (e.g., inter-coder reliability can serve to assess consistency in coding). Furthermore, codes should be specified and illustrated with an example, particularly when a coding scheme develops in the course of a project, the data sources coded early in the project have to be reviewed for consistency, and the application of codes have to be evaluated for consistency (Bazeley, 2018b). Assuming these criteria are met, quantitizing facilitates more effective and reliable communication and analytical integrity by offsetting biased impressions (Bazeley, 2018b; Maxwell, 2010). With creative and reflexive use, quantitizing can assist in showing the complexity and multidimensionality of qualitative data (Sandelowski et al., 2009, p. 219).

Data consolidation

Data consolidation is probably the most sophisticated integration strategy. It goes beyond linking different sources of data given that it involves the creation of a new data set to support further analysis. The originally separate sources can no longer be distinguished. To achieve consolidation, qualitative and quantitative data have to be merged into a new data set. Thus, they have to take the same (or similar) format, which means some form of data transformation is required (Collingridge, 2012; Sandelowski et al., 2009; van Velzen, 2016). From this data set, it is not possible to identify the original qualitative or quantitative data; instead, genuinely new pieces of information are generated (Caracelli and Greene, 1993; Creswell and Plano Clark, 2018). Thus, qualitative, quantitative, and consolidated data can be analyzed, as well as the interaction between the three.

Consolidating data sources to create new composite variables is a generative strategy, often used to resolve puzzles raised by the data, but it can also create new questions (Bazeley, 2018b, p. 231). It often seems to be the case that consolidated data sets take a quantitative form. That is, qualitative data are transformed into numbers, and these numbers paired with the quantitative data constitute a new data set (e.g., Jang et al., 2008). However, data transformation can go both ways, quantitizing and qualitzing (Bazeley, 2018b; Vogl, 2017). Thus, a consolidated data set can also have a qualitative format.

Empirical example for integrative mixed data analysis

In a methodological research project, I studied the applicability of focus groups with children (Vogl, 2019). I conducted five focus groups with children aged 6–15 years in a rural part of southern Germany. Each focus group consisted of 8–10 participants of roughly the same age. The focus group guideline was structured by “task complexes” to facilitate comparisons. These task complexes primarily anticipated specific group processes. Some tasks were formulated to prompt the group to establish consensus (“collaboration tasks”), others to stimulate a discussion, and so on. Each focus group lasted for 45–60 min and was video recorded and transcribed in detail, including information about the addressee of communication (moderator, specific group members, the group as a whole). The analysis dealt with formal aspects, such as the progress of the conversation, group dynamics, role differentiation, and the abilities of the participating children.

Although raw data stemmed from “qualitative” data collection, a wide range of different types of information were available for analysis. Initially, I analyzed the data separately in three different strands: a qualitative, a quantitative, and a transformation strand. In the qualitative strand, utterances were analyzed regarding their type of content, for example, whether an utterance was a new aspect or a repetition or confirmation of something said before. Quantitative characteristics were word counts, number of “private” conversations in a subgroup, duration of pauses, and the number of overlapping and of incomprehensible utterances. In the transformation strand, I coded utterances with the interaction process analysis (IPA) coding scheme.

Within these strands, a group level and an individual level of analysis can be distinguished. At the group level, aggregated information was integrated to form *focus group profiles*. On the individual level, information by participant was compiled to create *participant profiles*. For the group level of analysis, I linked results from the qualitative, quantitative, and transformation strands into profiles per age group with focuses on participation and motivation and on collaboration/interaction. Original strands could still be identified in this step although they were integrated in a narrative description of the age groups. For the participant profiles, the qualitative, quantitative, and transformed information was consolidated by integrating verbal and numerical data. In essence, all results from the three strands were transformed into verbal or qualitative profiles to characterize each participant. With this consolidation into individual role profiles, age-specific types of behavior and underlying cognitive, social, and verbal skills could be defined. The results were case profiles as a new form of data, which were then compared both within and across age groups.

The *interaction analysis* in the transformation strand was the core of the study, but without additional *qualitative analysis*, participants’ role descriptions would have been too superficial. Without the additional *quantitative aspects*, age comparisons and the characterization of the course of the discussion would have been far more difficult. Furthermore, the number of words alone did not provide information on verbal or cognitive skills. Qualitative-descriptive aspects were crucial for interpreting the quantified information and giving a more detailed picture of verbal and communicative skills for children in focus group settings. Neither strand alone could have offered sufficient information to answer the research question. Qualitative, quantitative, and transformed information gave insights into partly overlapping but generally distinct aspects of the research problem. The *linked and consolidated data* offered yet another dimension of results. Linking and matching components of complementary data offered a more nuanced account of how the different data corroborate, illustrate, or elaborate on each other (Bazeley, 2018b).

The linked and consolidated data were used in a comparative analysis on the functioning of focus group discussions in different age groups and consequently facilitated meta-inferences on age-related specificities. Integrating different aspects paid tribute to the multidimensionality of the data produced in focus groups. At the same time, the extensive raw data became manageable without neglecting either the breadth or depth of the information, and systematic comparisons between age groups regarding structure and content of focus groups were facilitated. Furthermore, by merging different types of information and creating a new qualitative data set, credibility and intersubjectivity of results could be extended, thereby increasing the reach of results.

Conclusion: challenges and implications for future research

Although MMR is defined by an attempt to integrate qualitative and quantitative elements, integration is still underdeveloped and undertheorized in MMR and therefore represents a key challenge for the community. The aftermath of the paradigm wars hindered creativity in data integration. However, from a pragmatic point of view, neither philosophical nor epistemological underpinnings should guide research, but rather methods that are fit for the purpose of answering the research questions at hand. The research question and the underlying theoretical assumptions determine the processes needed for data collection and analyses. However, the link between theory and methods has to be maintained to avoid methods-centric approaches to MMR (Hesse-Biber and Johnson, 2015, p. xiii) and to resist a “mix and match syndrome” (Leininger, 1994, p. 103). Notably, MMR “should never be used because of the misguided assumptions that more is better, that it is the fashionable thing to do, or, most importantly, that qualitative research is incomplete without quantitative research (Morse, 1996)” (Sandelowski, 2000, p. 254).

MMR is not about picking methods from a toolbox with the labels qualitative and quantitative, but rather making knowledgeable choices about research designs and methods based on strengths and weaknesses (Pearce, 2012). Only a thoughtful, reflexive use of multiple methods can create added value. Yet, this should not gloss over the “spirit of adventure” (Greene, 2007) that MMR requires. Integration in MMR is often exploratory, iterative, and inherently messy. Furthermore, it requires a breadth of skills as well as imagination and the ability to envision what might be possible (Bazeley, 2006, 2012). The spirit of adventure can be a strength of MMR and promote a creative (yet reflected) use and mix of approaches—without being limited by typologies and terminologies.

Although typologies offer a framework to communicate research, MMR typologies are diverse and use various terms and terminologies. The lack of a uniform conceptual framework is one problem, while another is the limitation typologies can impose on innovation. Typologies (often) treat quantitative and qualitative analyses as unitary approaches despite the existence of many different ways of working with both numeric and nonnumeric data (Onwuegbuzie and Combs, 2010). The boundaries between these two families of methods are blurred and “now the emergence of alternative forms of data, can be considered too great for them to be seen as separate and coherent primary elements in typological models” (Bazeley, 2018b, p. 63).

The division into qualitative and quantitative data is not as straightforward as it may seem at first glance, and it is not helpful when conceptualizing mixed methods data analysis. After all, the distinction between qualitative and quantitative research is only a “straw man” (Bergman, 2008). The phenomena surrounding us are ontologically and epistemologically neutral and have both qualities and quantities. We can represent phenomena, including mental phenomena, in multiple ways using both numbers and words—or images or sounds—and indeed, in everyday life, we typically do so (Bazeley, 2018b). Thus, it becomes apparent that “qualitative and quantitative data are not so much different kinds of data as these data are experiences formed into, for example, words or numbers, respectively” (Sandelowski et al., 2009, p. 209). Quantitative and qualitative data are both created by means of the particular conceptual “lens” used by the researcher (Maxwell, 2010). Whether we use numbers, words, or images does not change the phenomenon being described or analyzed, although what we use can change how we think about the phenomenon and how we present it to others (Bazeley, 2018a, p. 336). This also means that data are always selectively constructed by researchers—not collected or generated.

For example, the data produced during interviewing or focus groups are neither qualitative nor quantitative per se. Verbal utterances would be described as qualitative data and the number of words or word share as quantitative data. However, are word counts the product of data transformation, or do they exist independently? Are we quantizing when we count words or interactions? Yes, we are, but in the same way, we would quantize verbal utterances in a survey interview when we assign a numerical value to the response. If quantizing is about converting verbal data into numerical data, then survey interviews quantize in a similar way as word counts or coding in qualitative analysis. The implication is that talking about data transformation suggests that the material underlying our research can be categorized as qualitative or quantitative. I would argue that the same material can be both. Meaning is inherent to numbers, just as numerical characteristics are inherent to verbal utterances. Data analysis—or, to be more precise, data interpretation—brings these different qualities to light. This also implies that researchers could approach the same material with different tools.

Furthermore, if data cannot be easily categorized as qualitative and quantitative, the separation of qualitative and quantitative methods is also questionable. Bergman (2008) perceives the perseverance of the labels qualitative and quantitative as being related to a considerable degree to “delineating and preserving identities and ideologies rather than to describe possibilities and limits of a rather heterogeneous group of data collection and analysis techniques” (p. 29). To conclude, the qualitative-quantitative distinction is unproductive (Miles and Huberman, 1994) and describes mental models rather than different types of data or methods (Maxwell, 2010). However, taking this thought to its logical conclusion, “If boundaries are porous, and quantitative plus qualitative is both disputed and inadequate as a distinguishing feature, is it possible to define a study as using mixed methods?” (Bazeley, 2018b, p. 7). Despite these issues, creating new terms and terminologies would add confusion and probably be similarly unsuccessful in delineating the territory (Bazeley, 2018b). Thus, we might prefer using generally recognized terms, while remaining aware of the limited value of the dichotomy in methods and data and taking account of the fractal distinction of methodologies (Knappertsbusch, 2020). Integrative thinking is not a terminology problem but mainly a practical one.

Furthermore, data integration is not only a challenge in MMR but in any study using more than one data type during the analysis—whatever data types are involved or from which source they stem. Integration can be based on the merging of qualitative and quantitative data, as well as other types of data or perspectives. Even integrating perspectives of different family members into a construction of the family as an interactional unit deals with some of these challenges (Vogl et al., 2019), as does the integration of multiple “qualitative” or “quantitative” data types in multi-methods research. Thus, strict application of terminologies limits

integration. A single, shared terminology is not absolutely necessary, but clear definitions can provide a solid conceptual framework. It is important to keep our general frameworks somewhat flexible to allow for “informed creativity” (Mertens et al., 2016, p. 3). More innovative and creative thinking in integrated data analyses and the publication of good-practice examples thereof could significantly advance the field. Of course, this innovation is not an end in itself but rather a means of promoting substantive insights in the field of education and advancing understanding and practice.

References

- Akreml, L., Baur, N., Knoblauch, H., Traue, B. (Eds.), 2018. *Grundlagentexte Methoden. Handbuch: Interpretativ forschen* (1. Auflage). Beltz Juventa, Weinheim, Basel.
- Baur, N., Knoblauch, H., 2018. Zur Interpretativität des Quantitativen: oder: Zur Kongruenz von qualitativer und quantitativer empirischer Sozialforschung. *Soziologie* 47 (4), 439–461.
- Baur, N., Kelle, U., Kuckartz, U., 2017. Mixed methods—Stand der Debatte und aktuelle Problemlagen. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie* 69 (S2), 1–37. <https://doi.org/10.1007/s11577-017-0450-5>.
- Bazeley, P., 2006. The contribution of computer software to integrating qualitative and quantitative data and analyses. *Res. Sch.* 13 (1), 64–74.
- Bazeley, P., 2010. Computer-assisted integration of mixed methods data sources and analyses. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Los Angeles, Calif., pp. 431–467.
- Bazeley, P., 2012. Integrative analysis strategies for mixed data sources. *Am. Behav. Sci.* 56 (6), 814–828. <https://doi.org/10.1177/0002764211426330>.
- Bazeley, P., 2016. Mixed or merged? Integration as the real challenge for mixed methods. *Qual. Res. Org. Manag. Int. J.* 11 (3), 189–194. <https://doi.org/10.1108/QROM-04-2016-1373>.
- Bazeley, P., 2018a. “Mixed methods in my bones”: transcending the qualitative-quantitative divide. *Int. J. Mult. Res. Approaches* 10 (1), 334–341. <https://doi.org/10.29034/ijmra.v10n1a22>.
- Bazeley, P., 2018b. *Integrating Analysis in Mixed Methods Research*. Sage, Los Angeles, California.
- Bergman, M.M., 2008. The straw men of the qualitative-quantitative divide and their influence on mixed methods research. In: Bergman, M.M. (Ed.), *Advances in Mixed Methods Research: Theories and Applications*. Sage, Los Angeles [Calif.], London, pp. 11–21.
- Blasius, J., Friedrichs, J., 2008. Lifestyles in distressed neighborhoods: a test of Bourdieu’s “taste of necessity” hypothesis. *Poetics* 36 (1), 24–44. <https://doi.org/10.1016/j.poetic.2007.12.001>.
- Bourdieu, P., 2007. *Distinction. A social critique of the judgement of taste*. Harvard University Press, Cambridge, Mass.
- Brannen, J., O’Connell, R., 2015. Data analysis I: overview of data analysis strategies. In: Hesse-Biber, S.N., Johnson, B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, New York, pp. 257–274.
- Bryman, A., 2007. Barriers to integrating quantitative and qualitative research. *J. Mix. Methods Res.* 1 (1), 8–22. <https://doi.org/10.1177/2345678906290531>.
- Caracelli, V.J., Greene, J.C., 1993. Data analysis strategies for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 15 (2), 195–207. <https://doi.org/10.3102/01623737015002195>.
- Christopoulos, D., Vogl, S., 2014. The motivation of social entrepreneurs: the roles, agendas and relations of altruistic economic actors. *J. Soc. Entrep.* 6 (1), 1–30. <https://doi.org/10.1080/19420676.2014.954254>.
- Collingridge, D.S., 2012. A primer on quantized data analysis and permutation testing. *J. Mix. Methods Res.* 7 (1), 81–97. <https://doi.org/10.1177/1558689812454457>.
- Creamer, E.G., 2022. *Advancing Grounded Theory With Mixed Methods*. Routledge, London, New York, NY.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Los Angeles.
- Fakis, A., Hilliam, R., Stoneley, H., Townend, M., 2014. Quantitative analysis of qualitative information from interviews: a systematic literature review. *J. Mix. Methods Res.* 8 (2), 139–161. <https://doi.org/10.1177/1558689813495111>.
- Fetters, M.D., Freshwater, D., 2015. The 1 + 1 = 3 integration challenge. *J. Mix. Methods Res.* 9 (2), 115–117. <https://doi.org/10.1177/1558689815581222>.
- Fielding, N., Fielding, J.L., 1986. *Linking Data*. Sage Beverly, Hills, CA.
- Greenacre, M., Blasius, J., 2006. *Multiple Correspondence Analysis and Related Methods*. Chapman and Hall/CRC.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*, first ed. Jossey-Bass, San Francisco, CA.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11 (3), 255–274.
- Guetterman, T.C., Molina-Azorin, J.F., Fetters, M.D., 2020. Virtual special issue on “integration in mixed methods research”. *J. Mix. Methods Res.* 14 (4), 430–435. <https://doi.org/10.1177/1558689820956401>.
- Hesse-Biber, S., Johnson, B., 2015. Preface. In: Hesse-Biber, S.N., Johnson, B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, New York, pp. xiii–xxii.
- Jang, E.E., McDougall, D.E., Pollon, D., Herbert, M., Russell, P., 2008. Integrative mixed methods data analytic strategies in research on school success in challenging circumstances. *J. Mix. Methods Res.* 2 (3), 221–247. <https://doi.org/10.1177/1558689808315323>.
- Johnson, R.E., Grove, A.L., Clarke, A., 2017. Pillar integration process: a joint display technique to integrate data in mixed methods research. *J. Mix. Methods Res.* 13 (1), 301–320. <https://doi.org/10.1177/1558689817743108>.
- Kelle, U., 2007. *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung: Theoretische Grundlagen und methodologische Konzepte* (1. Aufl.). VS, Verl. für Sozialwiss, Wiesbaden.
- Knappertsbusch, F., 2020. “Fractal heuristics” for mixed methods research: applying Abbott’s “fractal distinctions” as a conceptual metaphor for method integration. *J. Mix. Methods Res.* 14 (4), 456–472. <https://doi.org/10.1177/1558689819893573>.
- Kuckartz, U., 2014. *Mixed Methods*. Springer Fachmedien Wiesbaden, Wiesbaden. <https://doi.org/10.1007/978-3-531-93267-5>.
- Larner, M., Nagy, C., Halpern, R., 2018. *Inside the Black Box: Understanding Home Visiting Programs*. New Orleans.
- Leininger, M., 1994. Evaluation criteria and critique in qualitative research studies. In: Morse, J.M. (Ed.), *Critical Issues in Qualitative Research Methods*. Sage, Thousand Oaks, pp. 95–115.
- Li, S., Marquart, J.M., Zercher, C., 2000. Conceptual issues and analytic strategies in mixed-method studies of preschool inclusion. *J. Early Interv.* 23 (2), 116–132.
- Maxwell, J.A., 2010. Using numbers in qualitative research. *Qual. Inq.* 16 (6), 475–482. <https://doi.org/10.1177/1077800410364740>.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N., Maxwell, J.A., Molina-Azorin, J.F., Niglas, K., 2016. The Future of Mixed Methods: A Five Year Projection to 2020. Retrieved from: <https://mmira.wildapricot.org/resources/Documents/MMIRA%20task%20force%20report%20Jan2016%20final.pdf>.
- Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, second ed. Sage Publications, Thousand Oaks.
- Morse, J.M., 1996. Is Qualitative Research Complete? *Qual. Health Res.* 6, 3–5.
- Moseholm, E., Fetters, M.D., 2017. Conceptual models to guide integration during analysis in convergent mixed methods studies. *Methodol. Innov.* 10 (2), 1–11. <https://doi.org/10.1177/2059799117703118>.
- Niglas, K., Kaipainen, M., Kippar, J., 2008. Multi-perspective exploration as a tool for mixed methods research. In: Bergman, M.M. (Ed.), *Advances in Mixed Methods Research: Theories and Applications*. Sage, Los Angeles [Calif.], London, pp. 172–188.
- Onwuegbuzie, A.J., 2003. Effect sizes in qualitative research: a prolegomenon. *Qual. Quant.* 37, 393–409.

- Onwuegbuzie, A.J., Combs, J.P., 2010. Emergent data analysis techniques in mixed methods research: a synthesis. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Los Angeles, Calif., pp. 397–430.
- Onwuegbuzie, A.J., Hitchcock, J.H., 2015. Advanced mixed analysis approach. In: Hesse-Biber, S.N., Johnson, B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, New York, pp. 275–295.
- Onwuegbuzie, A.J., Hitchcock, J.H., 2019. Using mathematical formulae as proof for integrating mixed methods research and multiple methods research approaches: a call for multi-mixed methods and meta-methods in a mixed research 2.0 era. *Int. J. Mult. Res. Approaches* 11 (3), 213–234. <https://doi.org/10.29034/ijmra.v11n3editorial2>.
- Onwuegbuzie, A.J., Leech, N.L., 2021. Qualitizing data. In: Onwuegbuzie, A.J., Johnson, B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, London.
- Onwuegbuzie, A.J., Teddlie, C., 2003. A framework for analyzing data in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, pp. 351–383.
- Pearce, L.D., 2012. Mixed methods inquiry in sociology. *Am. Behav. Sci.* 56 (6), 829–848. <https://doi.org/10.1177/0002764211433798>.
- Sandelowski, M., 2000. Combining qualitative and quantitative exempling, data collection, and analysis techniques in mixed-method studies. *Res. Nurs. Health* 23, 246–255.
- Sandelowski, M., Voils, C.I., Knafl, G., 2009. On quantitizing. *J. Mix. Methods Res.* 3 (3), 208–222. <https://doi.org/10.1177/1558689809334210>.
- Schoonenboom, J., Johnson, R.B., 2017. How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie* 107–131. <https://doi.org/10.1007/s11577-017-0454-1>.
- Tashakkori, A., Teddlie, C. (Eds.), 2003. *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks.
- Tashakkori, A., Teddlie, C. (Eds.), 2010. *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Los Angeles, Calif.
- Teddlie, C., Tashakkori, A. (Eds.), 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Los Angeles.
- Van Velzen, J.H., 2016. Students general knowledge of the learning process: a mixed methods study illustrating integrated data collection and data consolidation. *J. Mix. Methods Res.* <https://doi.org/10.1177/1558689816651792>. Advance online publication.
- Vogl, S., 2017. Quantifizierung: Datentransformation von qualitativen Daten in quantitative Daten in Mixed-Methods-Studien. *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie* 69 (2), 287–312. <https://doi.org/10.1007/s11577-017-0461-2>.
- Vogl, S., 2019. Integrating and consolidating data in mixed methods data analysis: examples from focus group data with children. *J. Mix. Methods Res.* 13 (4), 536–554. <https://doi.org/10.1177/1558689818796364>.
- Vogl, S., Schmidt, E.-M., Zartler, U., 2019. Triangulating perspectives: ontology and epistemology in the analysis of qualitative multiple perspective interviews. *Int. J. Soc. Res. Methodol.* 22 (6), 611–624. <https://doi.org/10.1080/13645579.2019.1630901>.
- Woolley, C.M., 2009. Meeting the mixed methods challenge of integration in a sociological study of structure and agency. *J. Mix. Methods Res.* 3 (1), 7–25. <https://doi.org/10.1177/1558689808325774>.

Using software for mixed methods analysis

Udo Kuckartz^a and Stefan Rädiker^b, ^a Philipps University Marburg, Berlin, Germany; and ^b Methoden-Expertise.de, Verden, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	500
Different types of integration	501
Using software for results-based integration	501
Strategy 1: Connecting results with hyperlinks	502
Strategy 2: Side-by-side display of results	503
Using software for data-based integration	503
Strategy 3: Qualitative exploration of quantitative extreme cases	503
Strategy 4: Qualitative themes by quantitative groups	504
Strategy 5: Frequency of qualitative themes by quantitative groups	505
Strategy 6: Statistics by qualitative groups or types	505
Strategy 7: Case comparison table with mixed data	506
Strategy 8: Case maps with mixed data	506
Using software for transformation-based integration	507
Strategy 9: Quantitizing—transformation of qualitative into quantitative data	507
Strategy 10: Qualitizing—transformation of quantitative into qualitative data	508
Using software for sequential integration	508
Strategy 11: Developing a quantitative strand based on qualitative results (exploratory sequential design)	508
Strategy 12: Developing a qualitative strand based on quantitative results (explanatory sequential design)	509
Overview of strategies for integration and required software features	510
Summary and conclusions	510
References	511

Introduction

Using software in data analysis has long been taken for granted in both quantitative and qualitative social research. Software such as SPSS, Stata, or R in the field of statistics and NVivo, MAXQDA, or ATLAS.ti in the qualitative field are widely known and used by thousands of researchers. Mixed methods research has attracted widespread interest since the turn of the 21st century. Unlike for the quantitative and the qualitative research, no special software has been developed for mixed methods analysis, but qualitative data analysis programs (so-called QDA software¹ or CAQDAS²) have integrated more and more special functions for mixed data analysis over time. Which functions these are in detail can be found in the relevant publications: Frieze (2019) for ATLAS.ti, Jackson and Bazeley (2019) for NVivo, Kuckartz and Rädiker (2019) for MAXQDA, Salmons et al. (2020) for Dedoose and Provalis (2020) for QDA Miner. On the part of statistical software, mixed methods approaches have been largely ignored; no special functions for mixed-data analysis have been developed there in the last decade.

The central point of mixed methods data analysis is the integration of the data and the results of both method strands. This is reflected in the formula “1 + 1 = 3”, which is commonly used in the mixed methods research community (Fetters and Freshwater, 2015a). The formula means that through integration, the results of a mixed methods study amount to more than the individual results of the qualitative plus the quantitative strand (Bryman, 2006; Fetters et al., 2013). The situation in mixed methods data analysis can be compared with a two-factor analysis of variance in the field of statistics. Not only the individual effects of the two factors can be identified, but furthermore—and this is even more important—the interaction of the two factors.

A widely-used and common definition of mixed methods research is that qualitative and quantitative data are involved (Creswell and Plano Clark, 2018, p. 5; Fetters and Freshwater, 2015b, p. 208), while having several data from the same research paradigm is mostly referred as multi-methods research (Fetters and Molina-Azorin, 2017). Following this definition of mixed methods, a software that allows mixed methods analysis needs to support qualitative data analysis, quantitative data analysis, and above all the combined or mixed data analysis of both types of data. In the following, we focus on this issue of combining the two types of data, because integration is a major challenge in a mixed methods research process (Ivankova, 2014). Ideally, the software allows an interactive mode of working, that means to interactively access qualitative data through quantitative data and vice versa, while working with the original data or with results tables. For example, if you are working with a qualitative interview, coding and interpreting it, you also have access to all the quantitative data about the interviewee that was collected with a survey.

¹QDA is an acronym for Qualitative Data Analysis.

²CAQDAS is an acronym for Computer-Assisted Qualitative Data Analysis Software.

Different types of integration

The emphasis on integration as one of the keys to data analysis in a mixed methods project raises several questions: What type of integration is possible in a study? What type of integration is appropriate? And in particular, what factors influence the possibilities for integration in a project? From our perspective, there are three main interrelated factors that determine the integration in a mixed methods study (Kuckartz, 2017; Kuckartz and Rädiker, 2021):

First, the design. Creswell and Plano Clark (2018) proposed a distinction between three core mixed methods designs. The “convergent design” consists of one phase in which qualitative and quantitative data are collected and analyzed, while the primary goal of this design is to compare and/or combine the results. The “explanatory sequential design” consists of two phases. After collecting and analyzing quantitative data, the collection and analysis of qualitative data follows. And the “exploratory sequential design” starts with a qualitative phase which is followed by a quantitative phase. There are several other suggestions for naming and grouping mixed methods designs (e.g., Plano Clark and Ivankova, 2016; Tashakkori et al., 2021; Tashakkori and Teddlie, 1998), but only by focusing on “convergent/parallel” vs “sequential” as one of the main distinguishing criteria, it is obvious that integration in these design types has to follow different rules. In sequential designs, primary integration usually takes place *after* one of the strands has been analyzed to inform and shape the second phase. In convergent designs, integration can start during analysis phase, while in any mixed methods design, results of both strands can be combined and compared at the end.

Second, the purpose and the justification for selecting a mixed methods approach. If the purpose is greater validity and triangulation to evaluate the convergence of the results of the qualitative and the quantitative strand (Bryman, 2006; Greene et al., 1989), it might be misleading to start the integration during the analysis phase or use the results from one strand for data collection in the other. Instead, integration can usually begin only after both strands have been completed, thus the purpose determines the design and thereby the point in time and possible type of integration. But if the purpose is completeness and illustration (Bryman, 2006), it might be crucial to start the integration during data collection and combine data for individuals during analysis. This leads directly to the third determinant.

Third, the samples. Creating a table with qualitative and quantitative information for each case in a single row, as discussed in detail in Schoonenboom and Johnson (2021) and illustrated in many studies like the one of Leal et al. (2018), requires a sample that allows the matching of qualitative data and quantitative data for each person. If I have two samples with different people or if I collect qualitative and quantitative data from the same sample of people but cannot link both types of data due to privacy issues, the scope of integration possibilities is limited.

With these determinants in mind, we have developed twelve integration strategies which can be divided into four different types:

1. *Results-based integration:* The researchers combine, compare, and relate the results of the qualitative strand and the quantitative strand. Therefore, the analysis of both kinds of data has already been completed, at least partially, before the integration step.
2. *Data-based integration:* The researchers work with both kinds of data during the analysis phase. Data-based integration requires a “nested” or an “identical sample” (Onwuegbuzie and Collins, 2007), i.e., there are cases for whom qualitative and quantitative data have been collected. Furthermore, it must be possible to clearly assign the qualitative and quantitative data to the individual cases.
3. *Transformation-based integration* is a sub-part of data-based integration: qualitative data is transformed into quantitative data and, less common, quantitative data is transformed into qualitative data.
4. *Sequential integration* is used if the results of one strand inform the development of data collection instruments, the sampling strategy, or other aspects of the subsequent strand.

Figs. 1 and 2 locate these four types of integration in the course of mixed methods projects. Although types of integration can be combined within a project, usually, not all types of integration are carried out in the same mixed methods study. For example, data-based integration is typically used in convergent designs and is therefore rarely combined with sequence-oriented integration.

For all types of integration strategies, so-called “joint displays” play an important role. In a joint display, qualitative and quantitative data and/or results are arrayed together, and typically, a joint display consists of a table, conceptual graphic, or geographic map, but it also can follow a free-form (Fetters and Guetterman, 2021; Guetterman et al., 2015a,b, 2020). Joint displays serve as a tool for analysis and presentation of findings as well as for verifying and creating hypotheses. From our point of view, it is important that a software for mixed methods data analysis supports the creation of joint displays.

In the following sections, for each of the four types of integration, we will present appropriate strategies for combining qualitative and quantitative data and results and for creating joint displays with computer software. Each description of a strategy follows a defined structure, starting with a short description, followed by its application with computer software and additional notes. When we refer to software in the following, we mean QDA software or CAQDAS, because these are the most advanced in terms of mixed methods functionality. If we refer to other types of software, this is explicitly mentioned.

Using software for results-based integration

Results-based integration strategies can be conducted in any mixed methods study, because they rely on the findings from the qualitative and the quantitative strand. The typical goal of results-based strategies is to compare the qualitative and quantitative findings on specific research questions, topics, and themes. Of course, such a comparison only makes sense if both strands analyze the same

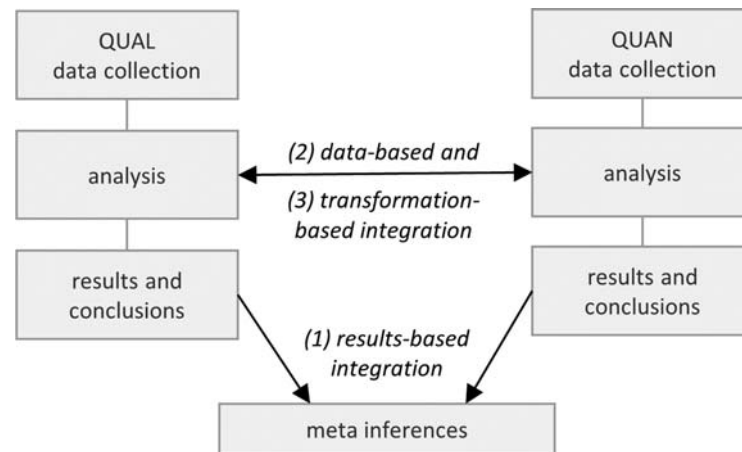


Fig. 1 Different types of integration in a convergent mixed methods design.

aspects of a phenomenon, thus, results-based integration strategies are common in a convergent design, particularly in those pursuing the purpose of triangulation. Results-based integration requires that the results of the qualitative and quantitative strand are available as written sub-reports or at least as compilations of partial results, graphs, and statistical tables (Kuckartz and Rädiker, 2019, p. 175). We describe two strategies for comparing and integrating the results using software: the first strategy connects qualitative and quantitative results by using hyperlinks and the second juxtaposes the results side-by-side in a table.

Strategy 1: Connecting results with hyperlinks

Description: Thematically related portions of the study results from both reports are hyperlinked so that when writing the final report, one can quickly find, relate, and summarize related results.

Implementation with software: To link related results of both strands, one can follow these steps:

1. Import both reports into the software and open them. If a second screen is available, one report should be placed there. Alternatively, both reports can be displayed in “tabs” as known from internet browsers or placed next to each other on a wide screen.
2. Start with reading one of the reports and search for relevant results and findings, for example, on willingness to pay fair prices on sustainable products (topic 1) and on actual behavior regarding sustainable consumption (topic 2). Using text search with relevant search items can help identifying sections. Highlight an identified section and mark it as the start of an internal link.
3. Go to the other report to identify and highlight a part related to the previously selected section; again, text search might be helpful. Set the target of the internal link to connect both sections. Now, if you click on the start or the target, the corresponding section in the other report will be focused and presented. You can add memos to the linked parts to record the addressed research question or topic.
4. When writing the integrative final report, use the memos for fast access to the thematically important links.

Notes: This is a universal type of integration that can be used whenever the results of a quantitative and a qualitative strand are to be compared. This integration strategy is primarily used in studies that work with different samples in the qualitative and quantitative strand, in other words, one does not have both types of data for each person. However, this strategy can also serve as a complement

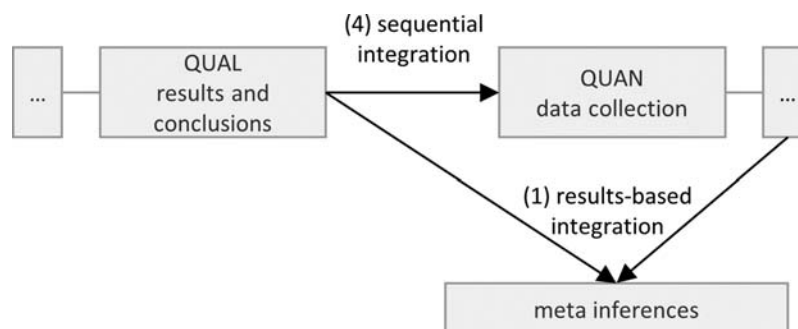


Fig. 2 Different types of integration in an exploratory sequential design.

for data-based integration strategies. Setting links during the analysis process is easy to accomplish, and when writing the research report later, it is possible to jump from one result to another very quickly.

Strategy 2: Side-by-side display of results

Description: The results of the qualitative and quantitative strand are juxtaposed and compared in a table. The table is structured by research questions or topics of interest (Table 1).

Implementation with software: Creating a side-by-side display with software is straight-forward because it relies on commonly used thematic coding techniques.

1. Create codes that reflect the research question and that can serve as the structure of rows, such as “Willingness to buy fair trade products” and “Actual behavior” in Table 1.
2. Work through the qualitative and quantitative sub-reports and apply the codes to relevant parts that contain substantial findings for the topic.
3. For both sub-reports, retrieve the coded segments for a topic and place the compiled results in the related cells of the side-by-side display. The cells of the table can contain the original texts, diagrams, and statistical tables from the sub-reports. This is no problem in the software where one can scroll even very large tables but, for integrating a side-by-side display into a report, the results are presented usually in a condensed and summarized form. The summaries can be written in the software and integrated in the table instead of the original sections.

Notes: If the software has no built-in function for creating side-by-side displays, you will need Word or another text-processing application to create the table. Creating a side-by-side display in the software can be an intermediate step for creating a more condensed form of this joint display in the final report. Often, an additional column containing the discrepancies and confirmations between both strands is included in the final joint display.

Using software for data-based integration

In data-based integration, qualitative and quantitative data are integrated during the data analysis phase, thus it is not a requirement that the results of both strands have been written down in a report. Integration and mixed analysis are carried out directly with the data and partial results during the data analysis phase. Data-based integration can be performed with all cases of a study for which both qualitative and quantitative data are available. The data must be clearly assignable to each case or group of cases: If qualitative interviews and quantitative questionnaires were used in a mixed methods study, it must be possible to assign the questionnaire data to each interviewee or at least a group of interviewees. For this reason, data-based integration is typically used in parallel designs, seldomly in sequential designs.

When it comes to data-based integration, particularly QDA software offering mixed methods features shows its full strength and actually none of the strategies no. 3 to no. 8 can be accomplished in a reasonable way without the use of QDA software. Most of the presented strategies rely on having the qualitative and quantitative data for each case included in the same project file to be able to use one of them to access cases (strategy no. 3), group cases (strategies no. 4, 5, 6), or present both kinds of data for each case in a table or diagram (strategies no. 7 and 8). Typically, for each qualitative case, a set of quantitative data is stored as so-called variables or attributes in the software.

Strategy 3: Qualitative exploration of quantitative extreme cases

Description: The quantitative data are used to identify “divergent and extreme cases” (Bazeley, 2018, p. 272) in order to further analyze their qualitative data (Creswell and Plano Clark, 2018, pp. 235–236). The quantitative data is the starting point for accessing and selecting qualitative cases for further investigation and for answering the questions *why are* and *what makes* these cases so special.

Implementation with software: This strategy consists of three steps, starting with (visual) analysis of quantitative data to identify special cases, storing of these cases, and their subsequent qualitative analysis.

1. Look for extreme cases by sorting variables by their values in the quantitative data table or by using charts, such as boxplots and scatterplots. Fig. 3 shows a scatterplot, a sorted column of one variable, and a boxplot with the extreme cases highlighted in each

Table 1 Schematic example of a side-by-side display.

	Qualitative strand	Quantitative strand
Willingness to buy fair trade products	Results, as (summarized) text, diagrams	Results, as (summarized) text, charts, statistical tables
Actual behavior	Results, as (summarized) text, diagrams	Results, as (summarized) text, charts, statistical tables

chart. In the scatterplot, people that have a very high value on one or both of the variables “environmental awareness” and “environmental behavior” are highlighted (please note that the graphics are not taken from a real study but created for illustration).

2. Store the identified cases or mark them as special cases, by adding a memo, changing their document color, or adding another variable. If groups of special cases are identified, they can be stored in a set of documents for joint exploration.
3. Analyze the identified cases by working through their qualitative data. If their data has already been coded, the code frequencies and code patterns can be included in the analysis.

Notes: It is also possible to use this strategy in the opposite direction. Qualitative extreme cases are identified, for example, by searching for interviews with a noticeable large number of coded segments regarding the purchase of sustainable products or with people who reject these products completely (Rädiker and Kuckartz, 2020, pp. 98–99). Subsequently, their quantitative data are analyzed and compared with those of the other study participants.

Strategy 4: Qualitative themes by quantitative groups

Description: In a tabular joint display, the qualitative data are compared by groups formed by quantitative data. The quantitative groups build the columns, the themes form the rows, and the cells contain the qualitative data (Table 2). If interviews have been conducted for the qualitative part, the cells can contain the original responses of the interviewees or a summary of the content written by the researchers.

Implementation with software: This strategy requires that the qualitative data have already been coded.

1. Select relevant codes that you want to compare by groups.
2. Select a quantitative variable as the differentiating criterion for grouping, e.g., a socio-demographic variable such as age groups or a test score of a person.
3. For each cell, retrieve the coded segments of the cases sharing the same variable value of their group and copy the segments into the table cells of the *qualitative themes by quantitative groups* joint display or use a built-in feature to create it automatically. If the content of the cells is very large, it might be helpful to summarize it before including it in the table.

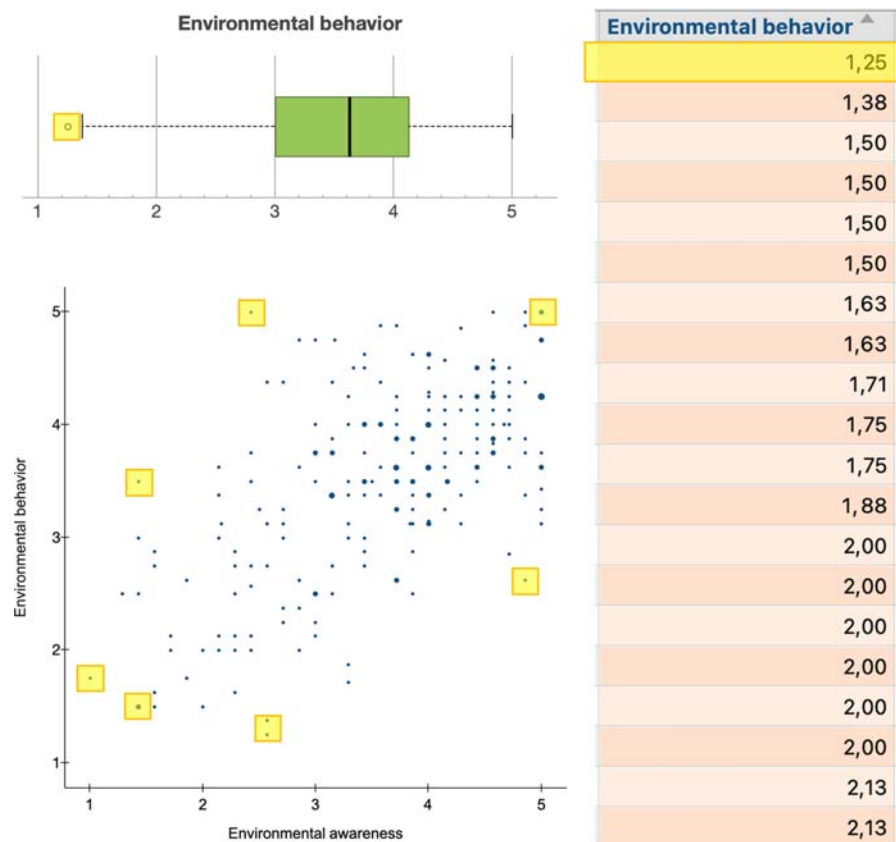


Fig. 3 Visually identifying divergent cases in a boxplot, a sorted list of variables values, or a scatterplot

Table 2 Qualitative themes by quantitative groups (joint display).

Qualitative themes	Awareness score: low	Awareness score: medium	Awareness score: high
Willingness to pay fair prices	Segments coded with “willingness ...” of persons with low score	Segments coded with “willingness ...” of persons with medium score	Segments coded with “willingness ...” of persons with high score
Actual behavior	Segments coded with “actual behavior” of persons with low score	Segments coded with “actual behavior” of persons with medium score	Segments coded with “actual behavior” of persons with high score

Notes: Similar to the side-by-side display of results, software allows one to work with larger tables during analysis, because they can be scrolled vertically and horizontally, but for printing and presentation it is necessary to create a condensed form of the table. When comparing several groups but only two or three themes, rows and columns can be switched to create more print-friendly tables with smaller width.

Strategy 5: Frequency of qualitative themes by quantitative groups

Description: In a tabular joint display, the qualitative data in form of code frequencies are compared by groups formed by quantitative data. This joint display is similar to the qualitative themes by quantitative groups in [Table 2](#), but the cells contain the frequencies of the coded themes in place of the texts or summaries.

Implementation with software: This strategy requires that the qualitative data have already been coded.

1. Select relevant codes that you want to compare by groups.
2. Select a quantitative variable as the differentiating criterion for grouping, e.g., a socio-demographic variable such as age groups or a scale or test value of a person.
3. Identify the number of coded segments or cases for which a theme has been coded and record this number in the cell of the results table or use a built-in feature to create it automatically. If the number of segments or cases is high enough, it might be helpful to use percentages instead of absolute frequencies. Thus, statements can be made such as “25% of the interviewees with a low awareness score expressed a strong willingness to pay fair prices for sustainable furniture.”

Notes: Quantification of qualitative data is not unproblematic and should be carried out after ensuring that it makes sense to analyze the code frequencies in this way (see also strategy 9 “quantitizing” below). Instead of working with the number of coded segments per theme, it is often a better and more reliable approach to work with the number of cases for which a theme has been coded, because the code frequency is influenced by many factors such as coding strategies and different data structures.

Strategy 6: Statistics by qualitative groups or types

Description: In a tabular joint display, quantitative statistics are compared for groups or types formed by the qualitative data. The groups or types are used as columns, while quantitative variables build the rows. For categorical variables such as the “Number of children”, the number of cases and a percentage value are presented in the cells, and for variables with interval scale such as “Awareness score”, the average mean and the standard deviation are shown ([Table 3](#)).

Implementation with software: This strategy requires that the qualitative data have already been analyzed and groups of cases have been identified or types have been built.

1. Select a code containing sub-codes that reflect the groups or types to be compared and convert it into a “qualitative” categorical variable to record the belonging to a group or type for each case.
2. Select the quantitative variables for the rows. Naturally, these should be variables for which one expects differences between the qualitative groups or types.
3. Select all cases of a group or type by using the grouping “qualitative” variable and calculate mean, standard deviation, number of cases and the respective percentage value for the selected row variables. Repeat this for all groups and types—or use the in-built feature of the software to create the results table in one step.

Notes: Highlighting the lowest and highest value per row helps to interpret the resulting joint table by allowing patterns to be identified, especially if a qualitative group holds most or all lowest or highest values.

Table 3 Statistics by qualitative groups or types (joint display).

Qualitative themes	Environmental ignoramuses	Environmental greenies	Environmental pragmatists
Awareness score	Mean (standard dev.)	Mean (standard dev.)	Mean (standard dev.)
Number of children	Number (%)	Number (%)	Number (%)

Strategy 7: Case comparison table with mixed data

Description: Data and analysis results from qualitative and quantitative strands are compiled in a tabular overview organized by cases. Each row of the table contains information about a case unit of the mixed methods study, e.g., a person, an institution, a family, or an organization. In the columns, coded segments or thematic summaries as well as variable values, scale values, or values calculated by statistical analyses (e.g., factor scores, cluster membership etc.) are displayed. Case comparison tables are commonly used to present research findings in publications or on conference posters (Schoonenboom and Johnson, 2021). They can also be used in other phases of a mixed methods project, such as to discover differences between cases or to find similarities and group cases accordingly (Bazeley, 2018).

The case comparison table in Table 4 was developed as part of an evaluation of a university statistics course (Kuckartz et al., 2008). It contains the data of four case units, namely the interviewees “I1”, “I2”, “I3”, and “I4”. The columns show four qualitative categories and the respective thematic summaries are listed for each case unit. The last two columns contain the individual values of two selected variables from the quantitative strand and one of these is used for sorting the table.

Implementation with software: Case comparison tables require that the qualitative data have been coded beforehand. It is also highly recommended to create thematic summaries per case unit for important categories. Summaries will make the table much more readable than full text segments. To create a case comparison table, proceed as follows:

1. Select the cases to be included in the table. The cases form the rows of the table. Usually, the case identifiers (and some demographic data) are displayed in the first column. It is also possible to group case units and then compare the groups in the table.
2. Select the codes whose coded segments or thematic summaries should be displayed in the columns.
3. Select the variables whose values should also be displayed in columns of the case comparison table.
4. Arrange the columns of the table in a meaningful order. Experiment with different ways to sort the table according to the contents of the columns.

Notes: Usually, the construction of a case comparison table goes through several stages (Schoonenboom and Johnson, 2021): additional codes or variables are added or removed from the table. Columns are sorted to list similar cases as close to each other as possible, and thematic integrative summaries are written or existing summaries are modified.

When constructing case comparison tables information overload should be avoided. This means that careful consideration should be given to whether the inclusion of a category or a variable really adds analytical value. If a study includes many cases, a reasonable selection should be made, or groups should be formed. Case overviews that extend over several pages should be avoided.

Case comparison tables have a twofold function; on the one hand they are an important tool of exploration and help to discover relationships and patterns in the data. Secondly, they can also serve as a means of presenting similarities and differences of cases. They can be used as a joint display in research reports and presentations to illustrate the results of a mixed methods study.

Strategy 8: Case maps with mixed data

Description: Visualizations such as the case map in Fig. 4 can show important characteristics of a case, e.g., of an interviewed person. In such a joint display, both qualitative and quantitative data are integrated.

Implementation with software: In principle, case maps may be created with any graphics software as well as with PowerPoint and similar presentation software. However, there are many advantages to using QDA software to create case maps, because this maintains a direct link to the original data. If one clicks on the corresponding symbols in the case map shown in Fig. 4, the underlying text passages are called up immediately.

Table 4 Case comparison table for four cases; standardized quantitative data marked with [x]; sorted by “Math grade at school”.

Literature usage	Study group	Expected exam result	Main attitude	Participation tutorial [x]	Math grade at school [x]
I3 No answer, probably no additional literature usage	Yes, for exam preparation	Pass! everything better is good	Fear pass!	Rarely	B
I2 Textbook completely worked through	Little, only once and again for exam preparation	No A, more a C	Interest ambitions	Very often	C+
I4 Textbook passed around in study group	Yes, during the whole semester	Pass! C would be nice	Interest	Rarely	C
I1 Only reader no further literature	No, only asking a friend during exam preparation	Pass!	Interest pass!	At least every 2nd week	D+

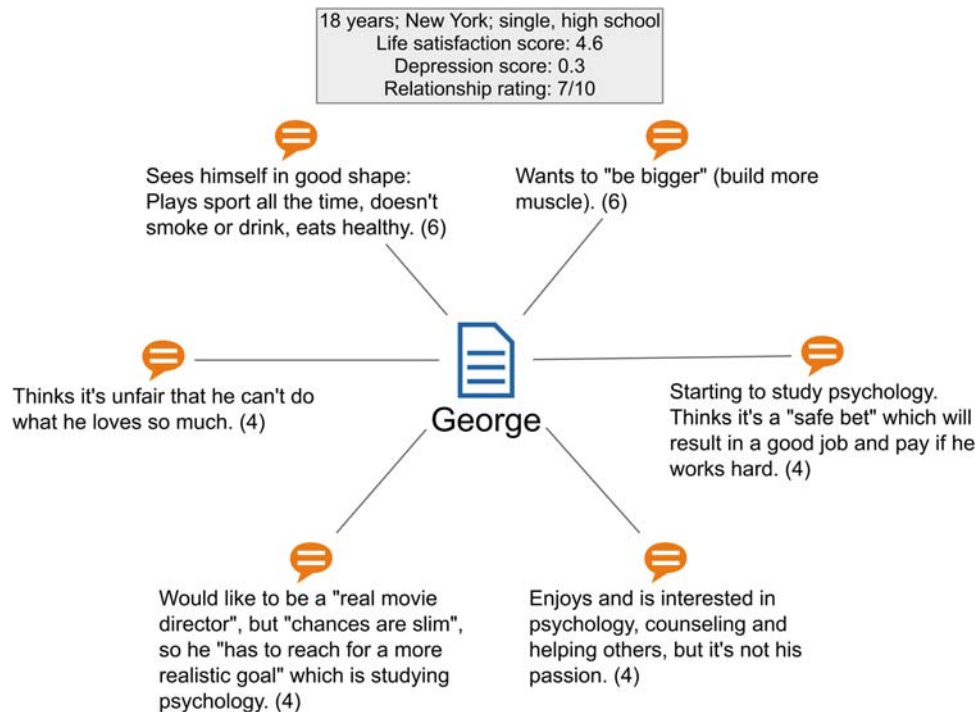


Fig. 4 Case map for the interviewee George

Notes: Case maps open up a variety of possibilities for creative social researchers; for example, photographs, images, and film clips can be incorporated and played when presenting research findings at conferences.

Using software for transformation-based integration

These types of integration strategy are based on the idea of performing the final data analysis with only one type of data, i.e., only as a qualitative analysis or only as a statistical analysis of quantitative data. For this purpose, a harmonization of the data type must take place, either a transformation of qualitative data into quantitative data ("quantitizing") or a transformation of quantitative data into qualitative data ("qualitizing"). On the one hand, both types of transformations are necessarily associated with information losses, but on the other hand, they also offer new opportunities to use approved methods of data analysis, e.g., cluster analysis, correlation, or regression analysis in the case of quantitizing.

Strategy 9: Quantitizing—transformation of qualitative into quantitative data

Description: The results of the qualitative component are transformed into quantitative variables using a methodologically controlled, usually automatic, transformation process. Through this "quantitizing", the results of both components are then available in quantitative form and the interrelationships can be examined using statistical methods.

Implementation with software: The transformation of the qualitative data can be done in several ways. A simple method is to record the frequency of application of a code per case as the value of a variable with a corresponding name. This method does not need to be described further, the software only needs to be able to perform the transformation automatically and store the data. More sophisticated is the transformation of a code with subcodes into a categorical variable. This strategy consists of three steps, which are described here in relation to a single category and its transformation:

1. *Code* qualitative data in such a way that a category is defined, and the subcategories capture the dimensions of this category. For example, the text passages of an interview that deal with the personal engagement of the research participants will be coded with the subcategories "low", "medium", "high" and "not mentioned".
2. These qualitative data are now *transformed* to create a quantitative variable "personal engagement" and the most frequently coded subcategory for each research participant is stored as a variable value.
3. Now *analyze* the data from the qualitative component, which was formed by means of transformation, for its interrelationships with the quantitative variables with the help of statistical methods. The condensed information obtained by the transformation (e.g., about the personal commitment per document) can be used for further analyses, e.g., to group research participants, to contrast them, or, for example, to compare the low-committed and the highly committed with each other.

Notes: Quantitizing strategies have been proposed for some time as a possible way of integrative analysis, e.g., with the aim of calculating relationships between qualitative and quantitative variables (Onwuegbuzie and Teddlie, 2003). The applicable statistical methods depend on whether the newly created variables fulfill the requirements of those methods. For example, depending on the appropriate level of measurement, parametric statistics or non-parametric statistics can be used. There are numerous examples in the literature of types of complex statistical analysis with quantified qualitative data. For example, van Haneghan (2021) reports about the usage of factor analysis, Cox et al. (2021) applied multiple linear regression, and Newman and Constantinides (2021) used structural equation modeling. Many more examples from research projects can be found in “The Routledge Reviewer’s Guide to Mixed Methods Analysis” edited by Onwuegbuzie and Johnson (2021).

Strategy 10: Qualitizing—transformation of quantitative into qualitative data

Description: Data from the quantitative study are transformed into qualitative data, usually narratives (Onwuegbuzie and Teddlie, 2003). In the simple case, the classification of a person on a scale or an index or the values for individual variables are expressed in words. For example, the information “132” of a variable “IQ” is transformed into the words “very intelligent”. Complex forms of qualitizing can create “comparative narrative profiles” based on complex statistical analyses such as cluster analysis or factor analysis (Onwuegbuzie and Leech, 2021).

Implementation with software: For qualitizing, there are no algorithmic procedures available as there are for quantitizing. This strategy consists of three steps, which are described here in relation to a single category and its transformation:

1. *Select* the variables, scales, or indices to be transformed. Under certain circumstances, it may be useful to first examine the interrelationships of several variables, e.g., the mutual correlation or, in the case of categorical variables, the correlations via cross-tabulation. Also, analysis procedures such as cluster analysis, factor analysis, or discriminant analysis can be performed first. Corresponding statistical calculations can be performed using statistical software or QDA software that supports these calculations.
2. *Transform* the numbers into words and sentences. For metric variables, it may be helpful to calculate mean and distribution measures. For ordinally scaled variables, calculate the median and quartiles or percentiles. This makes it easier to estimate the values of the persons or cases studied in relation to their position within the population of cases included in the study. The words and sentences formulated should remain close to the data and the wording should be respectful. For example, one should avoid referring to people with non-climate-responsive behaviors as “irresponsible citizens” or people who drive SUVs as “reckless horsepower fetishists”. The new qualitative data generated by the transformation should ideally be stored in the QDA software as a document memo for a case. It is also conceivable to insert these data as new texts in the project file and to code text passages with codes of the code system.
3. Now analyze in an integrated way by combining the data generated by qualitizing with the results of the qualitative component in a coherent text. This can be done by writing integrative memos. If the new data have been inserted and coded as their own text in the project file, joint displays can be created, which allow direct comparison of the results of the qualitative and quantitative strand in convergent designs.

Notes: Although the strategy of qualitizing is not chosen as often as the strategy of quantitizing in mixed methods research (Onwuegbuzie and Leech, 2021), it is by no means only applicable in special cases. Meanwhile, approaches of qualitizing are more frequently described (e.g., Vogl, 2019; Yousefi Nooraie et al., 2020) and also realized in mixed methods research (e.g., Headley and Plano Clark, 2020; Waller et al., 2021). This strategy appears to be particularly suitable when the qualitative strand is assigned greater relevance in the corresponding design of the mixed methods study. Qualitizing is also a promising option if publication in a qualitative research journal is intended.

Using software for sequential integration

At first glance, it looks as though researchers do not have to worry about integrating qualitative and quantitative strands when working with sequential mixed methods designs. While meta-studies on mixed methods research have often diagnosed the lack of integration of strands in convergent designs, a sequential design is not even thinkable without integration: study component A, conducted first, is by definition the precondition for study component B, which builds on its results. Integration is quasi-inherent to sequential designs, but the linking of the two components should be methodologically controlled and not just intuitive, and QDA software can play an important role in this integration process.

Strategy 11: Developing a quantitative strand based on qualitative results (exploratory sequential design)

Description: Designs in which a qualitative strand is conducted in the first step are also referred to as exploratory designs. The motive for such designs can be, for example, the need to extend or generalize the findings of the qualitative strand. The construction of a questionnaire for a subsequent survey is done in very close relation to the qualitative results. The source for each question and the reasons for asking this question with these answer specifications are recorded, so that the empirical foundation of the questions in the questionnaire is transparent at any time.

Table 5 Template for the development of a questionnaire in exploratory design.

<i>Source</i>	<i>Questionnaire</i>
Theme A: Findings of the qualitative strand	Question 1
Theme B: Findings of the qualitative strand	Question 2a to 2f
Quality of life scale according to WHO	Question 3a to 3f
Subcategories of category C of the qualitative strand	Question 4a to 4f
Demographic data	Question 5 to 9

Implementation with software: A direct development of a more or less immediately useable online questionnaire with the help of QDA software is not possible so far. No QDA software offers this function and probably will not do so in the future because, with software like Lime Survey, Qualtrics, or SurveyMonkey, highly specialized tools for the construction of online surveys are already available. However, QDA software can provide important support for the development of the quantitative strand, for example by recording on which results of the qualitative strand the questions of the newly constructed questionnaire are based or what the background for the construction of the questions is.

We recommend creating a correspondence table similar to **Table 5** for the design of the questionnaire; such a table can be created as a table document in the QDA software. The questionnaire to be constructed is gradually developed in the second column, while the source for each question is designated in the first column. The entries can be linked to the corresponding data or coded using the code system.

Of course, in a study with exploratory sequential design, not all questions included in the questionnaire have to be directly related to the results of the qualitative strand. It is standard practice to ask for demographic information such as age, gender, and place of residence. It may also be useful to collect supplementary quantitative information, for example, to include tests or well-established scales in the questionnaire.

Notes: Examples of using joint displays in exploratory designs can be found in [Ho et al. \(2021\)](#), [Munce et al. \(2021\)](#), and [Shim et al. \(2021\)](#).

Strategy 12: Developing a qualitative strand based on quantitative results (explanatory sequential design)

Description: Mixed methods designs, in which a quantitative strand is conducted in the first step, are also called explanatory designs. The motive for such designs, in which the qualitative strand builds on the quantitative strand, can be the desire to deepen the results of a survey or to clarify analysis results that cannot be explained with previous knowledge.

Implementation with software: How and in what way QDA software can support the design of a qualitative in-depth strand depends to a considerable degree on the planned method of qualitative data collection and the degree of openness.

If the qualitative strand is planned in a very open way, for example to better understand surprising results, one will only note very general keywords for narrative interviews or focus groups. In the software, memos should be used to record exactly why these findings should be explored in more depth. Such planning memos can be linked to the original data of the corresponding text passages of the quantitative research report.

If the subsequent qualitative strand is planned with a more structured approach, a detailed guideline for interviews or focus groups should be developed, the individual points of which are each linked to the corresponding results of the quantitative strand. Technically, this can be realized by creating the interview guide as a new document in the QDA software and linking each of the individual points to the results of the quantitative strand.

In an explanatory design, the integration does not only take place during the planning of the subsequent qualitative strand, but there is a second point of integration. After completion of the qualitative strand, the results of both project phases should be presented in a joint display. In this display, an overview should be given of which of the results of the quantitative strand can be better explained by the qualitative strand and which are still open. The joint display (**Table 6**) is similar to strategy 2 (side-by-side display of results).

Notes: Many suggestions for using joint displays in explanatory designs can be found in the literature, e.g., [Alexander et al. \(2021\)](#), [Ivankova et al. \(2006\)](#), and [Waller et al. \(2021\)](#).

Table 6 Template for a joint display in an explanatory mixed methods study.

<i>Quantitative strand</i>	<i>Findings of the qualitative strand</i>
Results that require explanation (based on the results of statistical analysis, e.g., correlations, regression analysis, factor analysis etc.)	Results clarified by the outcome of the qualitative strand. Results that could not be clarified even by the results of the qualitative strand.

Overview of strategies for integration and required software features

Most of the presented twelve strategies cannot be accomplished in a reasonable way without QDA software. At the time of writing this chapter, several competing software products are offered. Since QDA software, as well as the corresponding market, changes very quickly it does not seem reasonable to create a software comparison table, which necessarily becomes outdated quickly. The requirements that suitable software for mixed methods analysis must meet also depend to a large extent on the parameters of the specific project, for example, the types of data that have been collected, the size of the samples, the type of analysis method, etc. It makes a big difference if video data has been collected, because the software must of course be able to analyze it. It also matters if several types of data are combined, for example, if focus groups and narrative interviews were conducted, and documents are to be analyzed at the same time. Very important is what types of data can be handled (videos, paintings, photographs, Twitter data, YouTube comments, ...) and which file formats can be imported (docx, xlsx, pdf, png, mp3, mp4). Before starting the analysis, it is strongly recommended to check carefully whether the QDA software offers the required functionality. Some features should be met as basic requirements in any case:

- For all kinds of data-based integration, the QDA software needs to be able to import qualitative and quantitative data and connect them on a case level.
- QDA Software should be able to handle multiple qualitative data sources with different samples, such as focus groups, interviews, documents.
- Advanced features of searching and automatic coding should be provided.
- Options for linking different data should be available.
- In general, results tables should be interactive so that it is easy to jump from qualitative to quantitative data (and vice versa).
- At least a minimum of descriptive statistics functions should be available, as well as the ability to aggregate variables and build scales.
- Different types of joint displays for integrative analyses should be offered.

In [Table 7](#), all twelve integration strategies are listed once again. The second column indicates for each strategy which functionality the software must provide.

Most of the 12 strategies presented in this chapter and the corresponding joint displays are created in several cycles, i.e., experimentation is carried out until an optimal result is achieved. [Fetters and Guetterman \(2021\)](#), for example, have described this very impressively for the emergence of a Joint Display “Stakeholder perspectives on ethical advantages and disadvantages to adaptive clinical trials from the patients” that has been developed in seven iterations. To be able to realize such cyclical processes in a reasonable time, the software must be as interactive and flexible as possible.

Summary and conclusions

The community of mixed methods researchers agrees that integration is the crucial point of mixed methods approaches. Regardless of whether basic designs or fully integrated designs are involved, integration is generally unthinkable today without using QDA software. Software is needed that ideally allows the analysis of both qualitative and quantitative data. Only with such software can joint displays be created, in which qualitative and quantitative results are integrated. Of course, one can create such displays with Word, Excel, and PowerPoint, but the main advantage of QDA software is the interactivity of the joint displays ([Rädiker, 2018](#)).

Nevertheless, you should keep in mind that QDA software (CAQDAS) was originally created for qualitative data analysis. Special functions for mixed methods analysis were only developed over time as add-ons. This still results in some limitations at present. For example, the range of statistical analyses is not yet fully covered and there are no templates for established mixed methods designs. The guidance for typical analysis procedures in mixed methods research is therefore not yet optimal but is constantly being improved.

There are currently several QDA software packages on the market that fulfill the requirements listed in [Overview of strategies for integration and required software features](#) section to varying degrees. Since QDA software is constantly being developed further, one should check the current status of the packages and test available trial versions before planning a mixed methods analysis.

Ideally, software for mixed methods data analysis should have the ability to analyze both qualitative and quantitative data. Since many projects nowadays collect more than one type of data at a time (e.g., survey and qualitative interviews), the software should also be able to analyze many types of data simultaneously, such as surveys, qualitative interviews, focus groups, documents, videos. In addition, it must offer a wide range of possibilities for integrating both streams, and in this field joint displays are particularly important. Integration is a major challenge of mixed methods research and software is essential to meet this challenge.

Table 7 Twelve strategies of integration and necessary features in QDA software.

<i>Integration strategy</i>		<i>Software facilitates to ...</i>
Results-based integration		
1	Connecting results with hyperlinks	• Create internal links between parts of data and add notes or memos to them • Optional: compile links in a list for easy access
2	Side-by-side display of results	• Create a table with coded segments from qualitative and quantitative sub-report • Optional: write summaries of coded segments
Data-based integration		
3	Qualitative exploration of quantitative extreme cases	• Store quantitative data in variables for each case • Create charts for quantitative data, such as boxplots, scatterplots, histograms
4	Qualitative themes by quantitative groups	• Sort, select, and group cases by their variable values • Store quantitative data in variables for each case • Retrieve coded segments per sub-group that have been created by variable values and display them in a table
5	Frequency of qualitative themes by quantitative groups	• Optional: write summaries of coded segments • Store quantitative data in variables for each case • Calculate frequencies and percentages of cases or coded segments per sub-group that have been created by variable values and display them in a table
6	Statistics by qualitative groups/types	• Store quantitative data in variables for each case • Calculate mean and standard deviation for sub-groups and determine numbers of cases in a sub-group where a code has been applied; display these results in a table
7	Case comparison table with mixed data	• Store quantitative data in variables for each case • Create (interactive) tables containing quantitative variable values and qualitative summaries for each case
8	Case map with mixed data	• Store quantitative data in variables for each case • Create concept maps containing qualitative and quantitative elements that are linked to the original data
Transformation-based integration		
9	Quantitizing: transformation of qualitative into quantitative data	• Store quantitative data in variables for each case and to transform code frequencies into variable values
10	Qualitizing: transformation of quantitative into qualitative data	• Transform coded subcodes into categorical variables • Store quantitative data in variables for each case and compute statistical analysis like building reliable scales • Write integrative memos containing the results
Sequential integration		
11	Developing a quantitative strand based on qualitative results	• Create new document (questionnaire) • Link data and results to specific questions in the questionnaire • Create displays for the joint presentation of the findings
12	Developing a qualitative strand based on quantitative results	• Create new document (interview guide) • Link data and results to the interview guide • Create displays for the joint presentation of the findings

References

- Alexander, E., Eppler, M.J., Comi, A., 2021. Data integration: a real-time, participant-driven, and visually supported method. *J. Mix. Methods Res.* 15 (1), 87–113. <https://doi.org/10.1177/1558689820902294>.
- Bazeley, P., 2018. *Integrating Analyses for Mixed Methods Research*. SAGE.
- Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Inq.* 6, 97–113. <https://doi.org/10.1177/1468794106058877>.
- Cox, K., Lambert, R., Hitchcock, J.H., 2021. Multiple linear regression analysis with qualitative data that have been quantitized. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 77–87.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48 (6pt.2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>.

- Fetters, M.D., Freshwater, D., 2015a. The 1 + 1 = 3 integration challenge. *J. Mix. Methods Res.* 9 (2), 115–117. <https://doi.org/10.1177/1558689815581222>.
- Fetters, M.D., Freshwater, D., 2015b. Publishing a methodological mixed methods research article. *J. Mix. Methods Res.* 9 (3), 203–213. <https://doi.org/10.1177/1558689815594687>.
- Fetters, M.D., Gutterman, T.C., 2021. Development of a joint display as a mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 259–275.
- Fetters, M.D., Molina-Azorin, J.F., 2017. The journal of mixed methods research starts a new decade: principles for bringing in the new and divesting of the old language of the field. *J. Mix. Methods Res.* 11 (1), 3–10. <https://doi.org/10.1177/1558689816682092>.
- Friese, S., 2019. *Qualitative Data Analysis With ATLAS.ti*, third ed. SAGE.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11 (3), 255–274.
- Gutterman, T.C., Creswell, J.W., Kuckartz, U., 2015a. Using joint displays and MAXQDA software to represent the results of mixed methods research. In: McCrudden, M.T., Schraw, G.J., Buckendahl, C.W. (Eds.), *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*. Information Age Publishing, pp. 145–176.
- Gutterman, T.C., Fetters, M.D., Creswell, J.W., 2015b. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13 (6), 554–561. <https://doi.org/10.1370/afm.1865>.
- Gutterman, T.C., Molina-Azorin, J.F., Fetters, M.D., 2020. Virtual special issue on “integration in mixed methods research”. *J. Mix. Methods Res.* 14 (4), 430–435. <https://doi.org/10.1177/1558689820956401>.
- Headley, M.G., Plano Clark, V.L., 2020. Multilevel mixed methods research designs: advancing a refined definition. *J. Mix. Methods Res.* 14 (2), 145–163. <https://doi.org/10.1177/1558689819844417>.
- Ho, P., Chen, K., Shao, A., Bao, L., Ai, A., Tarfa, A., Brossard, D., Brown, L., Brauer, M., 2021. A mixed methods study of public perception of social distancing: integrating qualitative and computational analyses for text data. *J. Mix. Methods Res.* 15 (3), 374–397. <https://doi.org/10.1177/15586898211020862>.
- Ivankova, N.V., 2014. Implementing quality criteria in designing and conducting a sequential QUAN → QUAL mixed methods study of student engagement with learning applied research methods online. *J. Mix. Methods Res.* 8 (1), 25–51. <https://doi.org/10.1177/1558689813487945>.
- Ivankova, N.V., Creswell, J.W., Stick, S.L., 2006. Using mixed-methods sequential explanatory design: from theory to practice. *Field Methods* 18 (1), 3–20. <https://doi.org/10.1177/1525822X05282260>.
- Jackson, K., Bazeley, P., 2019. *Qualitative Data Analysis With NVivo*, third ed. SAGE.
- Kuckartz, U., 2017. Datenanalyse in der Mixed-Methods-Forschung: Strategien der Integration von qualitativen und quantitativen Daten und Ergebnissen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 69 (S2), 157–183. <https://doi.org/10.1007/S11577-017-0456-Z>.
- Kuckartz, U., Dresing, T., Rädiker, S., Stefer, C., 2008. *Qualitative Evaluation: Der Einstieg in die Praxis*, second ed. VS Verlag für Sozialwissenschaften.
- Kuckartz, U., Rädiker, S., 2019. *Analyzing Qualitative Data With MAXQDA: Text, Audio, and Video*. Springer Nature, Switzerland.
- Kuckartz, U., Rädiker, S., 2021. Using MAXQDA for mixed methods research. In: Johnson, R.B., Onwuegbuzie, A.J. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 305–318.
- Leal, I., Engebretson, J., Cohen, L., Fernandez-Esquer, M.E., Lopez, G., Wangyal, T., Chaoul, A., 2018. An exploration of the effects of Tibetan Yoga on patients' psychological well-being and experience of Lymphoma: an experimental embedded mixed methods study. *J. Mix. Methods Res.* 12 (1), 31–54. <https://doi.org/10.1177/1558689816645005>.
- Munce, S.E.P., Gutterman, T.C., Jaglal, S.B., 2021. Using the exploratory sequential design for complex intervention development: example of the development of a self-management program for spinal cord injury. *J. Mix. Methods Res.* 15 (1), 37–60. <https://doi.org/10.1177/1558689820901936>.
- Newman, D., Constantinides, S., 2021. Structural equation modeling with qualitative data that have been quantized. In: Johnson, R.B., Onwuegbuzie, A.J. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 89–97.
- Onwuegbuzie, A.J., Collins, K.M.T., 2007. A typology of mixed methods sampling designs in social science research. *TQR* 36.
- Onwuegbuzie, A.J., Johnson, B. (Eds.), 2021. *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge.
- Onwuegbuzie, A.J., Leech, N.L., 2021. Qualitizing data. In: Johnson, R.B., Onwuegbuzie, A.J. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 141–150.
- Onwuegbuzie, A.J., Teddlie, C., 2003. A framework for analyzing data in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. SAGE, pp. 351–383.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE.
- Provalis Research, 2020. QDA Miner 6. Qualitative and Mixed-Method Software. User's Guide.
- Rädiker, S., 2018. Interactive joint displays in MAXQDA for mixed methods data analysis. In: MMIRA International Conference. Vienna. (Accessed 22 August 2018).
- Rädiker, S., Kuckartz, U., 2020. *Focused Analysis of Qualitative Interviews With MAXQDA: Step by Step*. MAXQDA Press. <https://doi.org/10.36192/978-3-948768072>.
- Salmona, M., Kaczynski, D., Lieber, E., 2020. *Qualitative and Mixed Methods Data Analysis Using Dedoose: A Practical Approach for Research Across the Social Sciences*. SAGE.
- Schoonenboom, J., Johnson, R.B., 2021. The case comparison table: a joint display for constructing and sorting simple tables as mixed analysis. In: *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 277–290.
- Shim, M., Johnson, B., Bradt, J., Gasson, S., 2021. A mixed methods-grounded theory design for producing more refined theoretical models. *J. Mix. Methods Res.* 15 (1), 61–86. <https://doi.org/10.1177/1558689820932311>.
- Tashakkori, A.M., Johnson, R.B., Teddlie, C.B., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. SAGE.
- Tashakkori, A.M., Teddlie, C.B., 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. SAGE.
- Van Haneghan, J., 2021. Exploratory factor analysis of text. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 25–35.
- Vogl, S., 2019. Integrating and consolidating data in mixed methods data analysis: examples from focus group data with children. *J. Mix. Methods Res.* 13 (4), 536–554. <https://doi.org/10.1177/1558689818796364>.
- Waller, M.R., Nepomnyaschy, L., Miller, D.P., Mingo, M., 2021. Using a narrative approach to analyze longitudinal mixed methods data. *J. Mix. Methods Res.* 15 (2), 261–283. <https://doi.org/10.1177/1558689820953237>.
- Yousefi Nooraie, R., Sale, J.E.M., Marin, A., Ross, L.E., 2020. Social network analysis: an example of fusion between quantitative and qualitative methods. *J. Mix. Methods Res.* 14 (1), 110–124. <https://doi.org/10.1177/1558689818804060>.

Role of quality criteria in integrative mixed methods research

Kathleen M.T. Collins, Department of Curriculum & Instruction, University of Arkansas, Fayetteville, AR, United States

© 2023 Elsevier Ltd. All rights reserved.

Integration: an evolving concept	513
Research design	515
Sampling design	515
Validity design	516
Selective quality criteria frameworks	516
Onwuegbuzie and Johnson (2006)	516
Tashakkori et al. (2020)	516
Hall and Howard (2008)	517
Collins et al. (2012)	517
O’Cathain (2010)	517
Selective validity design concerns	517
Selective empirical examples of validity designs	518
Onwuegbuzie and Johnson (2006)	518
Onwuegbuzie and Johnson (2006)	518
Tashakkori et al. (2020)	519
O’Cathain (2010)	519
Selective aspects of designing quality criteria	519
Conclusion	520
References	520

High-quality research is characterized as justified in its selected approaches and procedurally transparent leading to outcomes that are viewed as viable, and finding that are viewed as applicable to stakeholders. Conceptually it is multidimensional construct comprising the transparency of the methodology, application of standards associated with self-selected intellectual communities, and construction of a validity design occurring at various junctures within the study (Collins, 2015). A strand, as a term defining the approaches (quantitative qualitative) of the study, was coined by Tashakkori et al. (2020). A strand comprises the following three stages of a study: (a) conceptualization stage, whereby the researcher forms the basis of the study, such as defining the purpose and the research questions; (b) experiential stage whereby the researcher makes the methodological and analytical decisions relative to implementing the study; and (c) the inferential stage, at which point the researcher forms conclusions per strand and integrates these conclusions leading to meta-inferences concerning the phenomena of interest. A meta-inference denotes an inference formulated based on the integration of inferences derived from the results obtained from qualitative and quantitative strands within a study (Tashakkori et al., 2020).

In this chapter, strand is the preferred term referring to the approaches implemented in an IMMR design. A critical aspect of high-quality research is design and implementation of quality criteria that will address concomitantly the rationale and purpose for integration, the research question addressed in the study, and the level of integration viable given the study’s design. Also in this chapter integration is the preferred term for mixing. The goal of this chapter is to prompt the design of high-quality research by discussing quality criteria within the context of integrative mixed methods research (IMMR), hereafter referred to as IMMR. The first purpose is to conceptualize integration as an evolving concept and to illustrate this evolution by presenting selective scholarly interpretations of integration within IMMR literature. The second purpose is to discuss the concept of inference as it relates to design decisions. The third purpose is to present selective approaches to designing quality criteria obtained from the published literature. I conclude this chapter with an outline of examples of empirical perspectives obtained from the published literature about how quality criteria are viewed and applied by researchers.

Integration: an evolving concept

Integration is used definitively as a descriptive term classifying an IMMR design; however, historically, there are varying interpretations about what constitutes integration when it occurs in an IMMR design. Greene’s (2007) interpretation is that the components, namely “samples, instruments, data sets, and analyses” “interact” “or have a conversation” and this conversation illustrates the “interact[ions]” occurring across the various stages of a study (p. 125, quotations in the original). Bryman (2007) interprets integration as a “negotiated account” occurring as the findings per strand “dialog” (p. 21). Bazeley and Kemp’s (2012) joint perspective centers on the interdependent relationship existing between data collected and the

subsequent analyses of these data leading to an integrated outcome. Collins' (2015) interpretive perspective centers on expanding the concept of integration as encompassing researchers' intertwined decisions pertaining to research, sampling, and validity designs, and the transparency of these decisions prompts the credibility of the study's outcomes to research consumers. Poth's (2018) perspective is that research practice is a non-linear endeavor, and it is designed and implemented under unpredictable conditions of complexity. In her view, successful integration requires researchers to adapt their practices of research to accommodate the encountered complexity. According to Poth (2018), the points of adaptation are occurring at the conceptualization stage when the researcher is (a) formulating the rationale and purpose for the study, (b) acknowledging the role of context in design, and (c) recognizing the strengths and limitations of the individual level or group members' levels of expertise impacting their design considerations.

Across these perspectives are similarities and dissimilarities relative to the focus that each researcher places on the concept of integration as it relates to IMMR. Greene (2007) and Bryman (2007) interpret integration as a conversation or dialogue that researchers self-construct as they view the contribution of each strand to the overall research outcome. In contrast, Bazeley and Kemp's (2012) focus on the codependent relationship between data collection and analytical activities, and it is the quality of this relationship that leads to integrated outcomes. Collins' (2015) focus also is on the interdependent relationship among design decisions, namely research, sampling, and validity designs, and it is the quality of these relationships that leads to integrated outcomes. Poth (2018) focus varies from the preceding interpretations. Poth's focus is on the complexity surrounding the study, and, in response to this complexity, researchers incorporate adaptations in their decisions about the study's design and implementation.

Researchers have indicated that there are complications surrounding the concept of integration. One potential complication is the assumption that there is a shared and consistent interpretation about the metaphors used to convey integration (Bazeley and Kemp, 2012). One example is the metaphor triangulation. Archibald (2016) has observed that researchers have expressed multiple interpretations of this term within the IMMR literature. Her observation has led her to conclude that the multiplicity of these interpretations downplays triangulation's utility as a term synonymous with integration, a point also made by Tashakkori et al. (2020). Bazeley and Kemp (2012) recommend that researchers specify *how* they are interpreting the term triangulation to illustrate how integration is occurring within a study. This leads to a second complication, which is the degree that the researcher's intentionality regarding integration is transparent to the research consumer. To address this complication, Fetter and Freshwater (2015) call for researchers to embed within their designs "intentional choices that can leverage integration" (p. 116). A third complication occurs when researchers' decisions regarding design and method selections are not transparent, subsequently, weakening the argument that "researchers [have] genuinely integrat[ed] [their findings] in such a way that the quantitative and qualitative components are mutually illuminating" (Bryman, 2007, p. 8). Maxwell et al. (2015) point out that the transparency issue associated with researchers' design and method decisions in terms of their relationship to integration "often are misleading" in the published literature (p. 225). To address these complications and other potential complications, researchers can review the literature to ascertain guidelines in the form of frameworks or typologies, with the understanding that these frameworks are not absolute.

An IMMR design has the potentiality to address concurrently exploratory and confirmatory research questions leading to results that when integrated allow the researcher to draw meta-inferences or conclusions not viable when using solely a quantitative or qualitative strand (Tashakkori and Teddlie, 2008). A meta-inference denotes an inference formulated based on the integration of inferences derived from the results obtained from qualitative and quantitative strands within a study (Tashakkori et al., 2020). There are varying types of inferences. Two types defined by Tashakkori and Teddlie (2008) are inference quality and data quality. Inference quality pertains to the process of interpretation when drawing conclusions and data quality contributes to this process by the quality of the data collected, analyzed, and interpreted, thereby leading to credible meta-inferences.

Influential aspects in IMMR design are forming a rationale and purpose for conducting an IMMR design. Collins et al. (2006b) developed an empirically based framework consisting of four rationales based on the results of their content analysis detailing how IMMR designs were implemented in the published special education literature. These four rationales followed by their formulated meaning in parentheses are (a) participant enrichment (participant recruitment), (b) instrument fidelity (assessment of the appropriateness of the instrument or the creation of one), (c) treatment integrity (assessment of the intervention's fidelity), and (d) significance enhancement (acquire diverse types of data leading to elevated interpretations). Accompanying these 4 rationales were 65 purposes identified in the literature and formulated by the authors. Examples of these purposes are in parentheses (a) participant enrichment (identify reasons for attrition, optimize sample selection by acquiring data leading to participant recruitment); (b) instrument fidelity (identifying potential biases in observational protocols, or instrument measurement error); (c) treatment integrity (identify characteristics of intervention implementation, such as consistency, barriers, and expeditors); and (d) significance enhancement (expansion of results across the quantitative and qualitative strands, and the development, modification, and testing of theory). The authors' intent underlying the design of this framework was to provide empirically based examples and author's suggestions for purposes, thereby providing the impetus to consider an IMMR design.

Research design

When conceptualizing a research design, there are prescribed design criteria that exist across many design typologies such as (a) level of integration, (b) time orientation of the ordering of the strands (concurrent vs. sequential), and (c) priority placed by researchers on the results when forming inferences and generalizations. Level of integration refers to the points in the study, whereby the researcher interconnects or integrates a decision or a result occurring within one strand of the study (i.e., quantitative) with a decision or a result occurring within the other strand (i.e., qualitative), and there can be multiple points of integration. Concurrent ordering of strands indicates that there is independence between the strands implemented and integration of results across the strands occurs at the conclusion stage. Sequential ordering of strands indicates that there is a dependency between strands, whereby data collected in one strand informs decisions implemented in a subsequent strand. When interpreting results per strand, the researcher decides as to how the analyzed data will be prioritized when drawing integrated conclusions. Factors that influence these decisions are the sample strategy used to collect these data per strand, sample size per strand, and the research question guiding the study's design. Examples of generalization choices are statistical generalization, internal statistical generalization (Onwuegbuzie and Leech, 2009), internal generalization (Maxwell, 2021), case-to-case transfer supported by theory (Creswell and Poth, 2017; Firestone, 1993), analytic or theoretical generalization (Yin, 2018), naturalistic generalization (Stake, 2005), and moderatum generalization (Payne and Williams, 2005). These criteria form the foundation of the design typologies of Onwuegbuzie and Leech (2009), Creswell and Plano Clark (2017), and Tashakkori et al. (2020). Guest (2013) makes the point that these typological design criteria focus on the study as a whole and are rudimentary because they cannot accommodate all types of design considerations occurring in more complex studies.

Other researchers offer an alternative interpretation of design. Guest (2013) recommends design criteria that involves the researcher focusing only on the "*point of interface*" occurring when data are integrated across the strands and to include two descriptive dimensions, namely timing and purpose of data integration (p. 146, italics in the original). Onwuegbuzie et al. (2013) introduce their conceptualization of design as context-specific, and it is based on their adaptation of Bronfenbrenner's (1979) ecological systems model. This model consists of four environmental contexts, and each context affecting differentially a person's development. Onwuegbuzie et al. (2013) adapted these four levels relative to the design of IMMR, namely (a) *micro-research studies* Level 1, that involves research focused on the individual or group within their current environmental context, (b) *meso-research studies* Level 2, that involves research focused on the individual or group selection of an environmental context that they choose to spend time, (c) *exo-research studies* Level 3, involves research focused on the individual or group choice of a context that they might be affected but it does not represent a context where the individual or group participates as an active member, and (d) *macro-research studies* Level 4, represents a cultural or societal context in which the individual or group participates as a member. In accordance with Onwuegbuzie et al.'s (2013) conceptualization, these levels are interconnected, and research design can be expanded to encompass more than one level.

Sampling design

A sampling design consists of the researcher's choice of sampling scheme and the determined sample size per strand. The choices of strategy and size reflect the interdependent relationship among research question, research design, and validity design. Additionally, these choices also are influenced by the practical aspect of balancing the collection of adequate data to address the research question and the available resources to implement the study. Using the existing literature as a source, Collins (2010) identified five criteria obtained from the published literature to consider when determining the sample design. Applying criterion 1 the researcher considers the dependent relationship between the sample per strand and the time orientation of the research design (concurrent or sequential). Applying criterion 2, the researcher indicates that the sampling unit per strand reflects the relationship between the strands. In this context, the sampling design could include an *identical* (same individuals participate in all strands of the study), *nested* (a subset of the sampling units in strand 1 is the sampling unit for strand 2), *multilevel* (the samples participating in the strands have a different level of engagement with the variable of interest [e.g., teachers, students, and the variable of interest is learning]), and *parallel* (different sampling units participate in the strands but the samples are drawn from the same population [e.g., different teachers participate in the strands]). Applying criterion 3, the researcher recognizes that the sampling unit selection is predicated on the relationship between the specific scheme selected and the type of generalization chosen by the researcher. Applying criterion 4, the researcher's selection of sampling units is predicated on choosing units that can produce multiple forms of data, such that the different data types can elevate the interpretation of the variable of interest. Applying criterion 5 the researcher is cognizant that the sample size can impact the degree of priority that is placed on the type and amount of data collected per strand and the integration of results across strands. This decision regarding priority is influenced by the type of generalization chosen by the researcher, sample design per strand, and the research question. Interpretive consistency refers to the relationship of the sampling design per approach, the emphasis placed by the researcher on the results of both strands, and the degree that integration leads to credible and defensible meta-inferences (Collins and Onwuegbuzie, 2013; Collins et al., 2006a). Interpretive consistency can be compromised when the design decisions pertaining to the sequencing of the strands (concurrent or sequential) leads to unintended order effects or the sample size per strand is not taken under consideration when forming the decision to place more value on the results of one strand (comprising a larger sample size) in contrast

to the results generated by another strand (comprising a smaller sample size), thereby making integration problematic (Onwuegbuzie and Johnson, 2006). Subsequently, the transparency of researchers' decisions pertaining to sampling design is a critical quality criterion in IMMR studies (Collins et al., 2012).

Validity design

A critical dimension of IMMR design is the rigor of the quality criteria embedded throughout the study (Onwuegbuzie and Johnson, 2006). The design, application, and evaluation of quality criteria that reflect standards of rigor identified by intellectual communities is referred to as a validity design (Collins, 2015). Crafting a validity design requires researchers to... "select and use the appropriate quality criteria for the particular research questions and for the particular research study" (Collins, 2015, p. 245). Constructing a validity design requires that researchers be cognizant of several points. The first point is the choice of language used to convey the dimensions of the design. Language descriptive of quality criteria varies across quantitative, qualitative, and IMMR. The choice of language is likely to be reflective of the researcher's engagement with specific intellectual communities. For example, rigor of the design in a quantitative strand is interpreted by researchers as validity, reliability, replication, and consistency (Campbell and Stanley, 1963). The language describing rigor in a qualitative strand centers on trustworthiness that comprises the following attributes: dependability, confirmability, credibility, and transferability (Lincoln and Guba, 1985). The choice of language will influence the degree that the validity design is viewed as credible by the research consumer at the study's dissemination stage. Second point is the alignment of the quality criteria and the researchers' intended generalizations of the integrated results is an important component of a validity design.

Selective quality criteria frameworks

There are several quality criteria frameworks available to researchers to assist in designing quality criteria for their IMMR study. The frameworks detailed in this section were selected because each framework presents a different perspective about the design of quality criteria. It is important to recognize that each framework and each criterion presented in the framework are not absolutes. Rather the frameworks and the quality criteria are guidelines for researchers to reflect upon when developing the specific quality criteria pertinent to their individual study.

Onwuegbuzie and Johnson (2006)

Integration within an IMMR design can blur the distinctive lines classifying quality criteria associated with a specific strand. To address this complication, researchers can design and apply legitimization criteria, a term coined by Onwuegbuzie and Johnson (2006), to distinguish quality criteria pertinent to the unique design considerations of an IMMR study. Actualizing this perspective into practice, the researcher applies legitimization criteria to address potential threats. Onwuegbuzie and Johnson (2006) identified nine potential design threats. The name of each threat is presented in the parentheses. First threat deals with the degree that the researcher's choices regarding sampling decisions relative to the sample design per strand facilitates formulation of credible meta-inferences (sample integration). Second threat notes the degree that there is clarity in the expressed views of the researcher and the participants in the forms of description, explanation of the data, and the conclusions (inside-outside). Third threat is the degree that the weakness or limitation of one strand in terms of providing adequate data is addressed by the design quality of the other strand (weakness minimization). The potential impact of an order effect occurring by the researcher's choice to implement one strand prior to the implementation of the other strand (sequential) is the fourth threat. Fifth threat is the degree of the quality of the data as a result of the transformation of data as in the cases of qualitzing and quantitizing (conversion). Qualitizing is a process whereby quantitative data are converted into data that can be analyzed using qualitative methods. Quantitizing is a process whereby qualitative data are converted into data that can be analyzed using quantitative methods (Onwuegbuzie and Teddlie, 2003). Sixth threat is the degree that a researcher's philosophical beliefs, such as "epistemological, ontological, axiological, methodological, and rhetorical" (p. 57) that underpin the researcher's decision to integrate results into a cohesive and credible meta-inference that supports their choice of design (paradigmatic mixing). The degree that the "meta-inferences made by the researcher reflect a mixed worldview based on the cognitive process of Gestalt switching and integration" (p. 57) (commensurability) is the seventh threat. The eighth threat pertains to the degree that the researcher has implemented appropriate quality criteria associated with quantitative, qualitative, and IMMR throughout the study, thereby leading to the formulation of credible meta-inferences (multiple validities). The ninth threat pertains to the degree that the research consumer values the meta-inferences formulated by the researcher when integration of the results occurs across the strands (political).

Tashakkori et al. (2020)

The integrative framework conceptualized by Tashakkori et al. (2020) is based on two interrelated quality standards. The first, design quality, focuses on the degree that the methods selected and the design of the strands pertaining to the "conceptualization, experimental methodological/analytical [and] inferential" decisions (p. 145) address appropriately the research question.

Underpinning design quality is design fidelity that addresses the rigor of the design accompanied by the appropriate methodological implementation of the design and procedural decisions. Design quality also is supported by design consistency and analytic adequacy. Design consistency reflects the degree that the strands are implemented in a “logical and seamless manner” and analytic adequacy is addressed by the execution of data analysis techniques that will address appropriately the research question (p. 301). Interpretive rigor, as a quality standard, examines the relationship between the inferences drawn by the researcher and the type of consistency noted in parentheses (a) “type, scope, and intensity” of the findings (interpretive consistency) (p. 301); (b) “theory and state of knowledge in the field” (theoretical consistency) (p. 301); (c) researcher’s conclusions also are considered viable by other researchers (interpretive agreement); (d) conclusions determined by the researcher are credible in contrast to other possible conclusions (interpretive distinctiveness); (e) researcher’s meta-inferences are supported by the conclusions (integrative efficacy); and (f) meta-inferences formed by the researcher are related to the purpose of the study and the research question, thereby supporting an IMMR design (interpretive correspondence).

Hall and Howard (2008)

Another example of an integrative framework is the synergistic framework designed by Hall and Howard (2008). Their synergistic approach to the design of quality criteria comprises four core principles. Researchers applying these four principles acknowledge that researchers’ philosophical assumptions and stances when applied to the design of quality criteria can support an IMMR design, application of quality criteria representing qualitative and quantitative traditions are viewed as equitably important in obtaining credible outcomes, integration of these multiple quality criteria can lead to credible outcomes that are potentially superior to and definitively different from outcomes derived from implementing a singular strand to design, and researchers can apply an inter-subjective approach toward interpreting claims, thereby balancing equitably claims about subjectivity and objectivity. The concept of synergy also underpins Fetter and Freshwater’s (2015) perspective on integration, and their call for researchers to identify the synergy that occurs when integrating across the qualitative and quantitative strands within an inquiry.

Collins et al. (2012)

Collins et al. (2012) incorporated Hall and Howard’s (2008) synergistic framework along with aspects of other published legitimization frameworks (e.g., Creswell and Plano Clark, 2017; Dellinger and Leech, 2007; O’Cathain et al., 2008; Onwuegbuzie and Johnson, 2006; Tashakkori et al., 2020) into a legitimization framework that is built on the premise that quality criteria in the context of an IMMR design is a “continuous, holistic, and synergistic process” (p. 855). Collins et al. (2012) introduced two additional quality criteria. These two criteria serve as their framework’s core elements. Philosophical clarity, as the first quality criterion, refers to the degree of awareness, on the part of the researcher, of their paradigmatic assumptions and equally important their awareness of the degree that these assumptions influence design decisions and create opportunities for obtaining credible conclusions and successful integration. The second quality criterion aligns with Collins et al.’s (2012) perspective that legitimization criteria are enhanced when researchers ... “dialogue with values [and] perspectives of multiple communities of practice” and concomitantly these two criteria ... “play continuous, iterative, holistic, and synergistic role in the legitimization process” (p. 849). In this context, process refers to researchers embracing the concept that the design of quality criteria occurs throughout the study and at the study’s conclusion, thereby leading to credible and defensible meta-inferences.

O’Cathain (2010)

O’Cathain (2010) developed a quality criteria framework consisting of aspects of the published literature. The framework was organized into the following sub-categories: (a) stage of the study, (b) domains of quality, (c) items within the domains, (d) definition of each item per domain, and (e) published source of domains and items. Quality domains identified by O’Cathain are (a) planning quality, (b) design quality, (c) data quality, (d) interpretive rigor, (e) inference transferability, (f) reporting quality, (g) synthesizability, and (h) utility.

Selective validity design concerns

Although constructing a validity design sounds intuitively logical to designing quality criteria, in the context of IMMR, O’Cathain (2010) queries the researcher’s dichotomous choice to design and to implement quality criteria tailored to the “methodological strand or to each method used” (p. 535, italics in the original). This choice is impacted by the decision to implement a sequential design involving interdependency between strands, the decision to implement multiple methods of data, and the decision to implement data conversions (quantitizing, qualitzing) (O’Cathain, 2010). A second concern is that integration of findings across strands blurs the distinction between strands and compromises the integrity of quality criteria designed per strand (O’Cathain, 2010). A third concern is the degree of the researcher’s degree of self-awareness of their philosophical stance and the influential impact it has on methodological decisions (Collins et al., 2012).

Selective empirical examples of validity designs

In this section, I outline selective examples of validity designs obtained from published studies. These examples were selected because the (a) the researcher was transparent in detailing their design decisions pertaining to research, sampling, and validity designs; (b) in the write-up of the study, they applied published quality criteria applicable for addressing potential threats to integration; and (c) the research centered on a topic relevant to the field of education. In the following outlined studies, the specific framework that guided researchers' decisions is noted as a header.

Onwuegbuzie and Johnson (2006)

Study 1. [Ivankova \(2014\)](#) applied a three-step procedure with a specific focus on procedures associated with implementing these steps and were accompanied by strategies that were designed to validate meta-inferences and to address validity threats in designing and implementing a sequential design. The goal of this study was to illustrate these procedures in practice, subsequently, they were applied in a study of graduate students' skill acquisition in applied research online courses. The ordering of the strands was quantitative followed by qualitative. The rationale for conducting a sequential design was that this type of design requires strategies crafted to address potential order effects of strands, which, subsequently, might impact the integration process.

The sample was a nested sample ([Onwuegbuzie and Collins, 2007](#)). In the quantitative strand, the sample comprised 84.5% ($n = 119$) of students who had completed at minimum of one of the four on-line research courses. In the qualitative strand, data collection consisted of telephone interviews containing 12 questions and selective probes that were designed to illicit more information per question. The sample was a purposive sample of 13 participants who had participated in the quantitative strand and were selected using a maximum variation sampling scheme. The resulting sample of students' interview responses were coded into the following categories: typical, extreme, unique, and the students represented a combination of varying levels of demographic status.

The validity design focused on implementing a three-step procedure. First step and second step were implemented by applying separate criteria aligned with specific research communities' standards and the type of data collected ([Bryman, 2006](#)). Specifically, step 1, the numerical data analyzed were questionnaire data and the analyses implemented were reliability coefficients, descriptive statistics, ANOVA, and multiple regression. In step 2, qualitative data were analyzed using a constant comparison method ([Glaser and Strauss, 1967](#)). Third step a contingency strand was applied, whereby quality criteria were designed based on the individual characteristic of the study's design ([Bryman, 2006](#)). Actualizing this step required application of strategies aligned to the design of a sequential IMMR design. The nested sample strategy (i.e., a subset of participants participated in both strands), thereby supporting sample integration, whereby the quality of the meta-inference is not compromised by introducing in one strand a sample that differs from the sample in the other strand ([Onwuegbuzie and Collins, 2007](#)). This study employed a sequential design that allowed [Ivankova \(2014\)](#) to assess the interaction of the trends noted in the results of the quantitative strand (strand 1, questionnaire data) and then explain these findings in the analysis of the data collected in the qualitative strand (strand 2, interview data). The integration of findings across the two strands led to an emerging design, whereby [Ivankova \(2014\)](#) expanded the number of questions in the telephone interview protocol and collected more data while the study was in progress. Triangulation for the purpose of corroboration of the qualitative findings occurred by examining three sources of data: interview data, follow-up emails, and the interviewers and Ivankova's reflective notes. Also, member checking was implemented to establish the credibility of the transcribed interviews and intercoder agreement on the results of the content analysis was measured on the themes (94%) and the codes (89%). In this study the three-step procedure supported the credibility of the meta-inferences formed from the conclusions integrated from both strands in a sequential design.

Onwuegbuzie and Johnson (2006)

Study 2. The goal of the study conducted by [Bustamante \(2019\)](#) was to assess professional development of teachers who were integrating Web 2.0 technologies into the Spanish curriculum in grades 7–12 in the context of online courses. Web 2.0 technologies that are included in the professional development are "blogs, wikis, podcasting, and poster, video, or comic generators" ([Bustamante, 2019](#), p. 165).

The rationale for conducting an IMMR design was significance enhancement ([Collins et al., 2006b](#)), whereby quantitative and qualitative strands are implemented to broaden the researcher's perspective on the topic of interest. Specifically, [Bustamante's \(2019\)](#) comparative purpose for conducting an IMMR design was to assess similarities and dissimilarities in the results of the data analyzed and to facilitate integration by implementing joint displays to visually merge data to facilitate interpretation and integration as recommended by [Guetterman et al. \(2015\)](#). Bustamante implemented joint displays to connect her results to a theoretical model, namely the Technological Pedagogical Content Knowledge (TRACK) model. This step is viewed as an "innovative" strand to design by Guetterman et al. ([Bustamante, 2019](#), p. 164, quotations in the original). The research design was a convergent design, whereby data in the forms of numerical and textual data were collected independently, and the results were integrated when forming meta-inferences. This design was selected because it allowed Bustamante to assess to what extent did the qualitative results confirm the quantitative results. Qualitative data comprised interviews, observation protocols, group discussions, documents, weekly journal entries, and final essays written by the participants detailing their experiences. Qualitative data were analyzed and coded to produce themes. Quantitative data, in the form of numerical questionnaire responses, were obtained at three points

in the study: pre-post-and follow-up tests. Quantitative data were analyzed using descriptive statistics, and ANOVA. The sample consisted of an identical sample (Onwuegbuzie and Collins, 2007). The quantitative strand consisted of 18 participants and 17 of the 18 participants were interviewed in the qualitative phase of the study, thereby addressing the threat of sample integration (Onwuegbuzie and Johnson, 2006).

Convergent quality criteria were implemented per strand (Bryman, 2006). Specifically, the analyzes were aligned to the type of data collected. A circular joint display was developed to facilitate data merging and integration. The joint display portrayed the following three components that were aligned to the TRACK theoretical model guiding Bustamante's decisions: (a) technology, (b) pedagogy, and (c) content. Bustamante (2019) also applied the list of potential validity threats identified by Creswell and Plano Clark (2017). Adherence to this list of their proposed strategies, Bustamante (2019) reported applying the following quality criteria strategies (a) implementing separate data collection procedures per strand and applying separate criteria aligned with type of data collected, (b) applying a joint display to illustrate visually the dissimilarities and similarities of the results leading to successful integration., Member checking also was implemented to confirm accuracy of the data collected and the degree that the results had face validity.

Tashakkori et al. (2020)

Study 3. Reid Kerrigan (2014) implemented a study to assess the potential impact of the application of data-driven decision making (DDDM) on the organizational capacity of community colleges. The rationale for implementing an IMMR design was significance enhancement (Collins et al., 2006b). Specifically, implementing qualitative and quantitative strands allowed Reid Kerrigan to ask confirmatory and exploratory questions, and as noted by Tashakkori et al. (2020) asking both types of questions within a single study is one of the defining characteristics of an IMMR design. The purpose for the collection of qualitative and quantitative data was that it permitted the development of an integrative analysis of quantitative and qualitative data, leading to an integrative data set, thereby providing a broadening of interpretation of a relationship between DDDM and organizational capacity of community colleges. The research design by Reid Kerrigan (2014) was a convergent parallel design implemented by applying a multiple case design. Application of a multiple case design allowed Reid Kerrigan to conduct a cross-case analysis. Application of a maximum variation sampling strategy permitted selection of colleges at two levels, top third and bottom third, in terms of their capacity for data use. The sampling design was a nested sampling design involving state selection ($n = 2$) by applying a critical case sampling scheme college selection ($n = 4$) by applying a maximum variation sampling scheme. Selection of administrators to participate was by invitation and faculty participants were randomly selected. The sample participating in the quantitative strand comprised full-time administrators ($n = 120$) and faculty ($n = 151$) employed at four colleges. Quantitative data consisted of collection and analysis of questionnaire data. The sample participating in the qualitative strand consisted of a subset of administrators and faculty who participated in the quantitative strand and who could provide "insight about data use" ($n = 38$) and who were selected using a criterion-based sampling scheme (p. 348). The validity design comprised reported reliability coefficients for questionnaire responses. A scheme to ensure consistency of the coding of the qualitative data was implemented, and inter-rater reliability to assess the level of consistency between coders was reported. Applying quality criteria recommended by Tashakkori et al. (2020), Reid Kerrigan (2014) addressed design suitability (convergent parallel design was appropriate to address the research questions), design fidelity (data collection and analyses procedures were aligned to ensure consistency with the design parameters), and analytic adequacy (detailed reporting supporting the adequacy of the data analysis to address the research questions).

O'Cathain (2010)

Study 4. Collins (2018) conducted a 4-phase analysis of the Reid Kerrigan (2014) study, Collins' intention was to detail the points of integration noted in the write-up of the study and to comment on the degree of transparency surrounding Reid Kerrigan's decisions relative to the design, method, analyses, and the points of integration in the write-up of the study. Collins (2018) concluded that Reid Kerrigan (2014) addressed the following transparency quality criteria identified by O'Cathain (2010): rationale and planning transparency, design transparency, data transparency, interpretive transparency, and reporting transparency.

Selective aspects of designing quality criteria

I conclude this section by presenting selective perspectives about the design of quality criteria obtained from the literature. Bryman's (2006) review of the literature identified three general perspectives expressed by researchers regarding quality criteria. The first viewpoint is to apply convergent criteria, and it involves applying these criteria to all strands. The second viewpoint is to apply separate criteria by employing quality standards embraced, for example, by the post-positivist research community (validity [internal external], reliability) and the constructivist research community (trustworthiness, transferability). These standards drive the construction of quality criteria to align the method of collection and the analysis of data collected per respective strand. The third viewpoint is to apply a contingency approach to designing quality criteria. Implementing this approach involves the researcher designing quality criteria based on the individual study; specifically, the design is centered on the level of integration and the degree that the findings obtained from one strand is prioritized appropriately when forming conclusions.

More recently, Fabregues et al. (2019) examined how quality criteria are conceptualized and procedurally implemented by researchers within an IMMR design. Their design was a multiple case design. Fabregues et al. (2019) selected a criterion sample of international researchers ($n = 44$) who represented the disciplines of education, nursing, psychology, and sociology. Participants were asked, via semi-structured interviews, for their opinions about how to conceptualize and implement procedurally quality criteria within an IMMR design. Data analysis consisted of a thematic analysis of the participants' responses and a within and cross-case analyses to detect potential patterns within the data set. Trustworthiness of the results was established by participants' member-checking the transcriptions of the collected data. Peer-debriefing interviews of the authors by a peer not associated with the study but who had familiarity with IMMR, were implemented. Peer debriefing interviews can serve to explore potential biases and to broaden researchers' understanding of the overall research process, and it is a form of an audit trail that is a recommended quality criterion by Onwuegbuzie et al. (2008). Additionally, data were collected to document methodological and analytical decisions and reflections made by Fabregues et al. (2019). These analyzed data also served as an audit trail. Results indicate that the quality criteria mentioned more than five times by the respondents were: (a) rationale for the IMMR study, (b) clarity of the purpose and the research questions, (c) clarity of the design in terms of phasing of the strands and the integration process, (d) clarity of the quality criteria reflected in the respective strands, (e) compatibility of the strands in terms of addressing the intent underpinning the study, (f) alignment of the inferences to the study's findings and inconsistencies noted, (g) transparency in the write-up of the study in terms of clarity and detail leading to credible and defensible inferences, and (h) the relevant IMMR literature was detailed in the write-up of the study.

Results indicated that published quality frameworks were favored by most of the participants. There were differences based on the participants' disciplines about the format of these frameworks. Educationalists and sociologists tended to favor frameworks that were "highly flexible [and] adaptable to different research contexts" (p. 435). In contrast, psychologists and nursing preferred a check-list format, specifically containing "criteria in the form of closed questions or scoring systems, highly structured, and that aimed at generating objective measurements of quality" (p. 435). The degree of familiarity with the frameworks was split. Half of the participants expressed familiarity with published frameworks, albeit a small number of frameworks, namely, the frameworks of Tashakkori and Teddlie (2008) and O'Cathain (2010). Advantages of frameworks from the perspective of those participants were their perceived utility in summarizing critical aspects of the IMMR literature and the practical component of detailing guidelines for implementing and appraising IMMR. Criticisms were centered on the degree that a framework can capture the complexity and dynamic interactions surrounding the IMMR study. As noted by Fabregues et al. (2019), there were points of convergence between the frameworks' content, and, subsequently, the literature on quality criteria. However, also noted by Fabregues et al. (2019) is that it is beyond the scope of this inquiry to determine the cause for this convergence. They ask the question Is it "the result of existing frameworks reflecting researchers' views on the quality IMMR or vice versa?" However, the convergence within this sample of 44 researchers suggests that the frameworks *potentially* impacted these researchers' views on quality criteria.

Conclusion

This chapter began with a definition of high-quality research. The goal of this chapter was to prompt the design of high-quality research by discussing quality criteria within the context of IMMR. The premise guiding the writing of this chapter was that the complexity of an IMMR design requires researchers to address issues of validity not required within a monomethod study. As a context for interpreting quality criteria in an IMMR design, I discussed the concept of integration by presenting selective and historically scholarly interpretations of integration within the IMMR and to emphasize the evolving nature of integration in the IMMR literature. I presented the concept of inference as it relates to design decisions within an IMMR by detailing design approaches leading to credible and defensible meta-inferences. Selective approaches to designing quality criteria, obtained from the published literature, presented in the form of selective frameworks were outlined to guide researchers' decisions when crafting quality criteria for a particular study. Also outlined, as exemplars of successful validity designs, were selective examples of empirical perspectives obtained from the published literature about how quality criteria are conceptualized and applied by researchers. I hope that the content of this chapter can inform researchers' future decisions in designing quality criteria for their studies.

References

- Archibald, M.M., 2016. Investigator triangulation: a collaborative strategy with potential for mixed methods research. *J. Mix. Methods Res.* 10, 228–250. <https://doi.org/10.1177/1558689815570092>.
- Bazeley, P., Kemp, L., 2012. Mosaics, triangles, and DNA: metaphors for integrated analysis in mixed methods research. *J. Mix. Methods Res.* 6, 55–72. <https://doi.org/10.1177/1558689811419514>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press, Cambridge, MA.
- Bryman, A., 2006. Paradigm peace and the implications for quality. *Int. J. Soc. Sci. Methodol.* 9, 111–126. <https://doi.org/10.1080/13645570701401644>.
- Bryman, A., 2007. Barriers to integrating quantitative and qualitative research. *J. Mix. Methods Res.* 1, 8–22. <https://doi.org/10.1177/2345678906290531>.
- Bustamante, C., 2019. TRACK and teachers of Spanish: development of a theory-based joint display in a mixed methods research case study. *J. Mix. Methods Res.* 13, 163–178. <https://doi.org/10.1177/1558689817712119>.
- Campbell, D.T., Stanley, J.C., 1963. *Experimental and Quasi-Experimental Designs for Research*. Rand McNally, Chicago, IL.
- Collins, K.M.T., 2010. Advanced sampling designs in mixed research: current practices and emerging trends in the social and behavioral sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 353–377.

- Collins, K.M.T., 2015. Validity in multimethod research and mixed research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 240–256.
- Collins, K.M.T., 2018. Crafting mixed research: interconnectivity of design decisions, choice of method, and integration [Special issue]. *Int. J. Mult. Res. Approaches* 10, 1–14.
- Collins, K.M.T., Onwuegbuzie, A.J., 2013. Establishing interpretive consistency when mixing strands: role of sampling designs in evaluations. In: Mertens, D.M., Hesse-Biber, S. (Eds.), *Mixed Methods and Credibility of Evidence in Evaluation. New Directions for Evaluations*, vol. 138, pp. 85–95. <https://doi.org/10.1002/ev.20060>.
- Collins, K.M.T., Onwuegbuzie, A.J., Jiao, Q.G., 2006a. Prevalence of mixed methods sampling designs in social science research. *Eval. Res. Educ.* 19 (2), 83–101. <https://doi.org/10.2167/eri421.0>.
- Collins, K.M.T., Onwuegbuzie, A.J., Johnson, R.B., 2012. Securing a place at the table: introducing legitimacy criteria for the conduct of mixed research [Special issue]. *Am. Behav. Sci.* 56, 849–865. <https://doi.org/10.1177/00027642114337999>.
- Collins, K.M.T., Onwuegbuzie, A.J., Sutton, I.L., 2006b. A model incorporating the rationale and purpose for conducting mixed methods-research in special education and beyond. *Learn. Disabil.* 4 (1), 67–100.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage, London.
- Creswell, J.W., Poth, C.N., 2017. *Qualitative Inquiry and Research Design Choosing Among Five Strands*, fourth ed. Sage, Thousand Oaks, CA.
- Dellinger, A.B., Leech, N.L., 2007. Toward a unified validation framework in mixed methods research. *J. Mix. Methods Res.* 4, 309–332. <https://doi.org/10.1177/1558689807306147>.
- Fabrigues, S., Pare, M., Menese, J., 2019. Operationalizing and conceptualizing quality in mixed methods research: a multiple case study of the disciplines of education, nursing, psychology, and sociology. *J. Mix. Methods Res.* 13, 424–445. <https://doi.org/10.1177/155868981987583333>.
- Fetters, M.D., Freshwater, D., 2015. The 1 + 1 = 3 integration challenge. *J. Mix. Methods Res.* 9, 115–117. <https://doi.org/10.1177/1558689815581222>.
- Firestone, W.A., 1993. Alternative arguments for generalizing from data, as applied to qualitative research. *Educ. Res.* 2 (4), 16–23.
- Glaser, B., Strauss, A., 1967. *The Discovery of Grounded Theory*. Aldine.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Guest, G., 2013. Describing mixed methods research: an alternative to typologies. *J. Mix. Methods Res.* 7, 141–151. <https://doi.org/10.1177/1558689812461179>.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13, 554–561.
- Hall, B., Howard, K., 2008. A synergistic strand: conducting mixed methods research with typological and systemic design considerations. *J. Mix. Methods Res.* 2, 248–269. <https://doi.org/10.1177/1558689808314622>.
- Ivankova, N.V., 2014. Implementing quality criteria in designing and conducting a sequential QUAN→QUAL mixed methods study of student engagement with learning applied research methods online. *J. Mix. Methods Res.* 8, 25–51. <https://doi.org/10.1177/1558689813487945>.
- Lincoln, Y.S., Guba, E.G., 1985. Emerging criteria for quality in quantitative and interpretive research. *Qual. Inq.* 1, 275–289.
- Maxwell, J.A., 2021. Why qualitative methods are necessary for generalization. *Qual. Psychol.* 8, 111–118. <https://doi.org/10.1037/qap0000173>.
- Maxwell, J.A., Chmiel, M., Rogers, S., 2015. Designing integration in multimethod and mixed methods research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, New York, NY, pp. 223–239.
- O’Cathain, A., 2010. Assessing the quality of mixed methods research: toward a comprehensive framework. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 531–555.
- O’Cathain, A., Murphy, E., Nicholl, J., 2008. The quality of mixed methods studies in health services research. *J. Health Serv. Res. Pol.* 13, 92–98. <https://doi.org/10.1258/jhsrp.2007.007074>.
- Onwuegbuzie, A.J., Collins, K.M.T., 2007. A typology of mixed methods sampling designs in social science research. *Qual. Rep.* 12, 281–316. <https://doi.org/10.46743/2160-3715/2007.1638>.
- Onwuegbuzie, A.J., Collins, K.M.T., Frels, R.K., 2013. Foreword: using Bronfenbrenner’s ecological systems theory to frame quantitative, qualitative, and mixed research. *Int. J. Mult. Res. Approaches* 7, 2–8.
- Onwuegbuzie, A.J., Leech, N.L., Collins, K.M.T., 2008. Interviewing the interpretive researcher: a method for addressing the crises of representation, legitimacy, and praxis. *Int. Inst. Qual. Methodol.* 7 (4), 1–17.
- Onwuegbuzie, A.J., Johnson, R.B., 2006. The validity issue in mixed research [Special issue]. *Res. Sch.* 13 (1), 48–63.
- Onwuegbuzie, A.J., Leech, N.L., 2009. Generalization practices in qualitative research: a mixed methods case study. *Qual. Quantity* 44, 881–892. <https://doi.org/10.1007/s11135-009-9241-z>.
- Onwuegbuzie, A.J., Teddlie, C., 2003. A framework for analyzing data in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 351–383.
- Payne, G., Williams, M., 2005. Generalization in qualitative research. *Sociology* 39, 295–314. <https://doi.org/10.1177/0038038505050540>.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. Sage, Thousand Oaks, CA.
- Reid Kerrigan, M., 2014. A framework for understanding community colleges’ organizational capacity for data use: a convergent parallel mixed methods study. *J. Mix. Methods Res.* 8, 341–362. <https://doi.org/10.1177/1558689814523518>.
- Stake, R.E., 2005. Qualitative case studies. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The Sage Handbook of Qualitative Research*. Sage, Thousand Oaks, CA, pp. 443–466.
- Tashakkori, A., Teddlie, C., 2008. Quality of inferences in mixed methods research: calling for an integrative framework. In: Bergman, M.M. (Ed.), *Advances in Mixed Methods Research*. Sage, London, pp. 101–119.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2020. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Yin, R.K., 2018. *Case Study Research: Design and Methods*, sixth ed. Sage, Thousand Oaks, CA.

Promising practices for addressing unique ethical dilemmas arising from integration in mixed methods research

Cheryl N. Poth^a, Elizabeth Creamer^b, and Leia K. Cain^c, ^a Department of Educational Psychology, University of Alberta, Center for Research in Applied Measurement and Evaluation, Edmonton, AB, Canada; ^b School of Education, Virginia Tech, Blacksburg, VA, United States; and ^c Department of Educational Psychology & Counseling, University of Tennessee, Evaluation, Statistics, and Methodology, Knoxville, TN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	522
Rationale for focus and description of team approach to this chapter	523
Challenges introduced by the inherent complexity in ethical reasoning within mixed methods research and teams	523
Ethical dilemmas arising from integration in evolving designs	525
Reasoned expectations of participation burdens	525
Comprehensive representation of integrated findings	526
Reasoned reporting of authorship contributions	526
Three promising practices for addressing ethical dilemmas in integration	526
Design integration procedures that affirm free and informed consent	526
Implement integration strategies that authentically acknowledge multifaceted findings	527
Communicate integration reports that fairly recognize contributions	528
Concluding thoughts	528
Acknowledgments	529
References	529
Further reading	530

Glossary

Collective reflexivity Engagement as a group in an active, ongoing process of critical reflection of how researchers co-construct knowledge from and during the research team process based on their identities, experiences, subjectivities, ontologies, and epistemologies

Ethical reasoning Decision-making in response to a moral dilemma based on a careful and thorough assessment of the different options in light of the facts, circumstances, and ethical issues (Poth, 2021)

Ethical review Consideration which intends to protect research subjects by minimizing the harms or risks to which they are exposed during research activities. The process often takes place within a community, an organization, or an institution

Informed consent A free and informed decision such as to participate in research. Individuals who provide informed consent must be legally competent and have enough decision-making capacity to consent to research

Mixed methods research An approach requiring the integration of both quantitative and qualitative data which assumes that their collective contribution mitigates inherent weaknesses in either type of data (Poth, 2018a)

Polyvocalism/polyvocality Incorporating more than one distinct authorial voice or perspective during reporting (Creamer, 2018)

Reflexive practice/reflexivity Engagement in an active, ongoing process of critical reflection of how the researcher constructs knowledge from and during the research process based on their identities, experiences, subjectivities, ontologies, and epistemologies (Cain et al., 2019)

Research ethics Application of ongoing thought and judgment of the fundamental ethical principles to the planning and implementation of research often guided by three principles: respect for persons, concern for welfare, and concern for justice.

Introduction

The conduct of ethical research is essential to protect participants, researchers, and society (Poth, 2021). Considering that ethical issues can arise at any time, researchers in educational settings need to be prepared for their ethical responsibility to respond appropriately. Ethical reasoning can offer important decision-making guidance in response to a moral dilemma based on a careful and thorough assessment of the different options in light of the facts, circumstances, and ethical issues (Anderson and Corneli, 2018). The need for ethical reasoning within a mixed methods research approach can be intensified by the requisite integration of qualitative and quantitative data and team configurations that are useful for tackling pressing societal issues in

education (e.g., Curry et al., 2013; Mertens et al., 2016; Poth, 2019). That descriptions of the unique ethical dilemmas faced by researchers and teams remain underdeveloped presents a key challenge for mixed methods researchers to engage in ethical reasoning. This chapter extends the assertion that unique ethical issues arise from integration in mixed methods research and leads us to offer promising practices to address that, “more so than using a single method design, mixed methods research yields the potential for complex and even amplified ethical problems” (Preissle et al., 2015, p. 157).

When investigating complex problems, mixed methods researchers often encounter dynamic or unstable research conditions which complicate the implementation of predetermined research plans. Key among the challenges for mixed methods researchers conducting research under conditions that are considered complex involves providing sufficient evidence in a proposal to satisfy the expectations of a supervisory committee as well as institutional, community, and organizational ethical review boards. The dynamic nature of a study means it is often impossible to fully predetermine the required elements such as participants, procedures, and expertise. We argue there is a pressing need to describe the unique ethical dilemmas related to integration that are encountered by mixed methods researchers and to offer practical guidance for how researchers and teams engage in ongoing ethical reasoning to realize the principles of ethical research under dynamic research conditions.

This chapter leverages our collective mixed methods research and team experiences to contribute practical ethical guidance in three ways: (1) by introducing the ongoing ethical reasoning necessary when engaging in the requisite integration of qualitative and quantitative research, (2) by illustrating why navigating three unique ethical dilemmas encountered by mixed methods researchers becomes intensified under dynamic conditions and team configurations, and (3) by describing three ethical practices including the use of reflexivity to guide mixed methods researchers managing the complexity inherent in ethical reasoning.

Rationale for focus and description of team approach to this chapter

To provide practical guidance reflecting the conditions mixed methods researchers are likely to encounter during the requisite integration of qualitative and quantitative research, we focus the promising practices on our experiences with dynamic conditions and team configurations. Our use of the term *teams* in this chapter includes the dissertation committees supporting graduate students, even though the literature on educational mixed methods research teams has not explicitly acknowledged this team configuration. Our own use of a team configuration for writing this chapter recognizes the enhanced capacity of mixed methods research teams for drawing upon members’ unique capacities for addressing more complex societal problems (e.g., Curry et al., 2013; Mertens et al., 2016; Poth, 2019).

To this work, Poth brings extensive expertise and experience as an educational mixed methods researcher diagnosing and managing dynamic mixed methods research conditions (Poth, 2018a,b), advancing mixed methods research team practices (Poth, 2012, 2019), and examining research ethics (Poth, 2021; Stadnick et al., 2021). Creamer brings extensive expertise and experience offering practical guidance for fully integrated mixed methods research (Creamer, 2018), intersections of different methods (Creamer, 2021), and reflexivity (Creamer, 2011). As an emerging mixed methods research scholar, Cain brings expertise about reflexive practice as a key component of ethical reasoning in mixed method research (Cain et al., 2019). Throughout the conceptualization and writing of this chapter, the team met regularly to move the chapter forward and create conditions for the emergence of new insights that would otherwise have been inaccessible by each researcher alone.

Challenges introduced by the inherent complexity in ethical reasoning within mixed methods research and teams

All research in education is guided by three principles for the conduct of ethical research: respect for persons, concern for welfare, and concern for justice (Anderson and Corneli, 2018; Poth, 2018a). The wording of these principles may vary by country, but their meaning remains the same. Fig. 1 provides a summary of the principles along with the questions researchers should ask themselves (Poth, 2021). The researcher is primarily responsible for the conduct of ethical review of their research and one way researchers can be supported in their efforts is to have an ethical review board review their research plan. Ethical reviews can take place in several ways, and more than one may be necessary: (1) a student’s supervisor and dissertation committee provide feedback on the research proposal (2) research undertaken through an institution (e.g., research center or university), likely needs approval from an institutional review board, and (3) research within a community or organization might need approval from those bodies. Consider that anticipating all the ethical issues that might arise unexpectedly in a well-articulated research proposal is unlikely so researchers must also be prepared to respond appropriately throughout the research.

Conducting mixed methods research in education can be challenging in addition to being rewarding and researchers need to be prepared to respond appropriately to ethical issues. There are often limitless possibilities for investigators with diverse methodological expertise to pursue research in education (Poth, 2019). According to the American Educational Research Association (AERA, 2021), education research involves:

[T]he scientific field of study that examines education and learning processes and the human attributes, interactions, organizations, and institutions that shape educational outcomes. Scholarship in the field seeks to describe, understand, and explain how learning takes place throughout a person’s life and

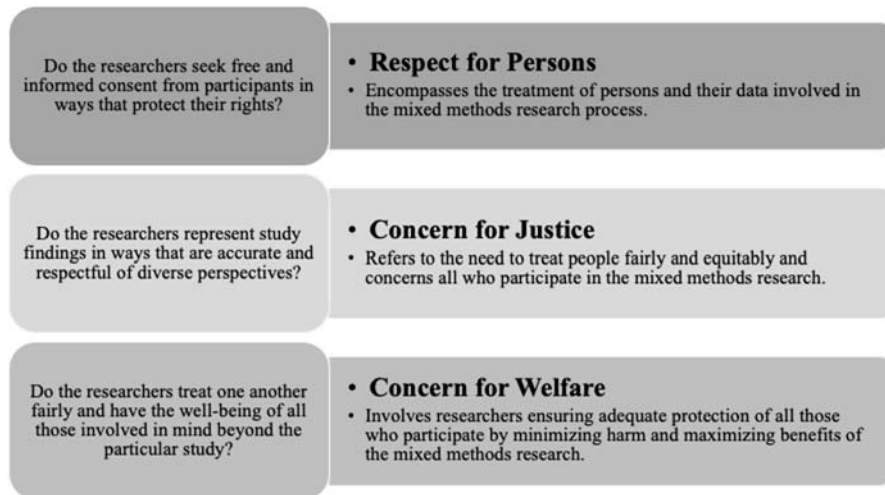


Fig. 1 Three principles for the conduct of ethical research.

how formal and informal contexts of education affect all forms of learning. Education research embraces the full spectrum of rigorous methods appropriate to the questions being asked and also drives the development of new tools and methods.

Research pursuing complex educational issues often benefit from team configurations drawing upon diverse disciplinary and methodological expertise.

Many of the inherent challenges when forming a mixed method research team with differing perspectives, experiences, and assumptions are well documented in the literature (Curry et al., 2013). These may become intensified under conditions of increased complexity and require new ways of interacting to leverage these differences productively (Poth, 2019). Many of our conventional approaches to methodological training restrict opportunities for meaningful interaction across disciplines and research approaches. They perpetuate the idea that research conditions are stable and that it is necessary to silo research methods. The training challenge noted by Hesse-Biber (2016, p. 650; emphasis in original) is that:

... researchers *do not practice interdisciplinarity well*. This is because, in part, they do not actively seek out ways to tap into the potential synergy of a team-based mixed methods project. They ethnocentrically do not see past their own comfort zone or horizon for theories, questions, and methods. I further argue that there is a *lack of conscious reflexivity* on the part of the research team; instead, the team often buys into the idea of 'inherent' synergy contained in these types of research configurations and designs. *Working in a group does not necessarily mean that you are working as a team.*

On teams, processes that allow for collective reflexivity can advance ethical practice by contributing to the methodological transparency that is necessary to share responsibility for ethical decisions that are made across the areas of responsibility of different team members.

Attending to ethical issues requires careful planning before conducting educational research, ethical actions during the mixed methods research process, and appropriate responses to emerging issues as the research process unfolds. Fig. 2 represents the iterative plan, act, and respond cycle to realize the principles of ethical research that is continually informed by ethical reasoning. Researchers must assume that ethical dilemmas can arise at any time throughout the research process and the conduct of ethical research requires ongoing thought and judgment specific to each mixed methods research project (Preissle et al., 2015). The need for ongoing ethical reasoning informed by reflexivity throughout the research is intensified in cross-disciplinary teams working in dynamic mixed methods research environments.

As a pillar of qualitative research, reflexivity is a practice that engages researchers in making explicit what they bring to study decisions and how they influence them. Glesne (2011) posited that reflexivity involves "critical reflection on how researcher, research participants, setting, and research procedures interact and influence each other" (p. 151). Popa and Guillermin (2017) applied the usefulness of these reflexive concepts to teams when they argued that "reflexivity can facilitate meaningful and problem-specific ways of combining methods across different disciplinary fields, types of expertise, and practices" (p. 19). We argue engaging in critical reflection about the many lenses educational researchers bring to a study and how they interact with topics, participants, data, and team members is essential for the conduct of ethical mixed methods research.



Fig. 2 Iterative plan, act, and respond cycle to realize guiding principles of ethical research.

Ethical dilemmas arising from integration in evolving designs

As a key characteristic of mixed methods research, the requisite integration of qualitative and quantitative research introduces unique ethical dilemmas in need of attention throughout the research process. Of paramount importance for meeting quality criteria for mixed methods research is to provide evidence of integration in the designs, procedures, and reports (Creamer, 2018; Fàbregues and Molina-Azorín, 2017; Onwuegbuzie and Poth, 2016; Plano Clark and Ivankova, 2016). For these reasons we have chosen to focus on three dilemmas arising from integration that we have ourselves experienced and observed others' experiencing with some frequency. For each, we describe the ways that changing research conditions and team configurations can intensify the unique ethical dilemmas and what is often considered an appropriate response.

Reasoned expectations of participation burdens

As is the case for all research, mixed methods researchers must respect participants' rights to know what is expected of them and to ensure that researchers seek free and informed consent throughout integration procedures. Providing full information about the mixed methods study to potential participants is key to demonstrating the ethical principle of *respect for persons* yet it can be especially challenging with sequential or longitudinal designs. Among the key challenges in predetermining research plans and consent procedures when research involves plans that build upon one another (i.e., sequential) or in designs where researchers may need to attend to the evolving needs of participants because the procedures occur over time (i.e., longitudinal). Common to both of these scenarios is that participants may be involved in multiple qualitative and quantitative data collections points in sequential or longitudinal research and that the protocols can evolve. For example, an interview protocol for a followup qualitative phase cannot be completely predetermined in a sequential design at the beginning because it is informed by the quantitative survey findings from the initial phase.

Similarly, protocols in longitudinal designs often evolve over the course of the study. Rare accounts of how researchers adjusted their procedures in response to their populations were described by Plano Clark et al. (2015) as part of an examination of longitudinal mixed methods designs from the health sciences. Whereas Wallen et al. (2012) replaced a long quantitative measure with a semi-structured qualitative questionnaire for palliative care participants, Symons et al. (2012) communicated through text messaging to be less intrusive to new parents. It is important that procedures be in place for participants to know and to have opportunities to reevaluate their consent to participate voluntarily once they become aware of the expectations, withdraw participation and data at any time, and clarify their understanding of participation with the researchers.

When we consider our experiences and observations, we note that attending to team members' and participants' vulnerability informs our ongoing ethical reasoning. Challenges can arise when working as a team where close coordination is required to develop a shared understanding of the consent processes individuals will be expected to follow across all phases of the study. Considering participant vulnerability and capacity is also essential to providing consent because research in education often involves working with vulnerable populations (such as school-aged children) within workplace settings where participants (e.g., teachers, principals) could be subject to coercive practices.

Comprehensive representation of integrated findings

The ethical *concern for justice* principle requires all researchers to report evidence and procedures honestly and in ways that respect diverse perspectives. Burch and Heinrich (2016) describe a fully integrated approach to mixed methods as including adaptability to the uncertainty and variability of complex designs as well as collaborators regularly and intensively discussing their data and findings. Ethical issues can arise from differing interpretations of the data. Unique to mixed methods research, reconciling divergence between the qualitative and quantitative phases can be problematic. Such challenges provide examples of how interpersonal context can introduce significant ethical challenges in mixed methods research (Plano Clark and Ivankova, 2016), especially after a project is underway (Hesse-Biber, 2010).

From our own experiences, we see the challenges of competing results causing tensions among researchers as relating to fairness with important consequences for how mixed methods researchers present their study outcomes. Some researchers have turned to publishing qualitative and quantitative findings in separate manuscripts which can mislead the reader and circumvent complexity by oversimplifying the outcomes of integration (Archibald et al., 2015). Searches for disconfirming evidence and exploration of alternative explanations about the same evidence is part of the process of warranting findings during reporting (AERA, 2006) and we argue these processes inform ongoing ethical reasoning in mixed methods research. The resulting tension poses a challenge for methodological transparency in mixed methods research (Evans et al., 2011; Hesse-Biber, 2010; Macfarlane, 2010; Popa and Guillermin, 2017) and for meeting expectations in reporting standards (Onwuegbuzie, 2012).

Reasoned reporting of authorship contributions

Establishing a fair way to designate contributions during reporting is a challenge encountered by all researchers. Attention is needed to meet the ethical principle of *concern for welfare* of all those involved in the research. This ethical issue can become pronounced when using mixed methods, when researchers actively integrate their methods during different phases of a research project (Hesse-Biber, 2010). Areas of responsibility and ownership of ideas are more readily determined in mixed methods projects when there is a clear division of labor that parallels areas of expertise.

When we consider our own research experiences, we note that attributing responsibility for ideas is more difficult to disentangle in a setting where there is a fluid division of labor or a high level of interaction, and that this is a dimension of ethical reasoning. This challenge is intensified in an integrated or dialectical approach to mixed methods that puts the emphasis on a back-and-forth between different methods throughout the research process (Creamer, 2021). For many researchers—and, in particular, those building their scholarly record in order to secure a job or funding—authorship, and even the order of authorship, can have high-stakes implications for professional well-being.

Three promising practices for addressing ethical dilemmas in integration

To advance the discussion of and illustrate real-world applications of ethical reasoning for managing the ethical dilemmas we have described, we offer three promising practices related to the design of integration procedures, the implementation of integration strategies, and the communication of integration outcomes. For each of the promising ethical practices, we describe the nature of the unique ethical dilemma intensified under dynamic mixed methods research conditions and team configurations. We also discuss the role that individual and collective reflexivity plays in managing the complexity inherent in ethical reasoning and offer examples of reflexive questions to guide mixed methods researchers in generating self-understanding and shared understandings among team members.

Design integration procedures that affirm free and informed consent

The requisite integration procedures within sequential, emergent, and evolving mixed methods research designs introduce the potential to intensify the time and expectation burdens placed on participants. Common across the designs are integration procedures that cannot be entirely predetermined because the follow-up phase of the research builds upon the results of the initial phase. Consequently, it can be impossible to provide fulsome information about the study procedures to participants at the onset as normally required to meet the standard of free and informed consent. With participants potentially vulnerable to increased time and expectation burdens introduced by some of the research designs, reflexivity offers the opportunity for researchers to individually and as a team to weigh their expectations of participants. This collective reflexive practice is noteworthy because it allows team members to develop a shared understanding of what might be asked of participants during the different phases of the research. These discussions have important consequences for developing consent procedures that inform participants about how initial consent will be sought and possibly reaffirmed later in the process. Of particular importance, engagement in individual and collective reflexivity enables researchers to interrogate procedural decision-making involved in the implementation of mixed methods research designs.

We have considered participant burden in our educational mixed methods research designs involving teams and changing conditions. Poth recalls having to adjust her team's expectations of school-based participants in a follow-up phase in response to the sudden shift to remote learning during the COVID-19 pandemic (Georgiou et al., 2021). Poth's ongoing discussions with

her interdisciplinary team to build a shared understanding of what could be reasonably expected of classroom-based participants helped communicate the impact of the shift to remote learning and to convey these changes to participants as the research unfolded.

Guidance for mixed methods researchers to attend and mitigate participant burden remains an understudied yet timely topic for further study. A survey involving 64 scientists conducting mixed methods health services research provides compelling evidence for this pressing need: Participant burden was the most frequently occurring ethical issue and among the most challenging one identified (Stadnick et al., 2021). It should also be noted that informed consent was the least reported ethical issue. Together with a limited focus of strategies that mitigate shortening data collection instruments, this evidence points to the pressing need for attention on consent procedures within mixed methods research. Similarly, experience tells us that, when working in teams, it can be a common practice for some researchers to be more involved in some aspects of the data collection and analysis involved in integration, and potentially not at all such aspects. A shared understanding of the expectations of how recurring consent procedures are implemented in practice is necessary to meet the standard for respect of persons as a guiding ethical principle. The following reflexive questions are offered to guide mixed methods researchers in affirming free and informed consent to generate self-understanding of their perspectives and build shared understandings among mixed methods research team members:

- How do my/our previous experiences influence my/our thinking about what research procedures I/we consider to be reasonable expectations for participants? How might my/our conceptualizations of “reasonable” differ from those among the research team members and participants?
- What would evidence of free and informed consent look like from participants throughout the research process? Do the research plans offer adequate protections for potentially vulnerable participants to provide consent?
- How do the research procedures provide adequate opportunities to ensure free and informed consent from participants during each data strand or phase? How might participants be affected by sequential, emerging, and evolving research designs and procedures and how can researchers keep participants adequately informed and reaffirm consent?

Implement integration strategies that authentically acknowledge multifaceted findings

The availability and use of several integration strategies make it possible to account for differences in findings or interpretations that emerge during mixed methods research without necessarily exceeding the word limits of a journal article. The differences can be important points for discussion and linked to various literature in a manuscript. Interestingly Pluye et al. (2009) describe that such points of divergence are frequently mentioned in reports, yet they rarely become the focus of analysis and integration. Common under dynamic and team conditions are integration outcomes that cannot be entirely predetermined because the data procedures may evolve and because of the contribution of interdisciplinary perspectives on the data that is generated. Consequently, it can be impossible to predict the nature of the integration outcomes. A table or figure can provide important opportunities for graphic techniques to acknowledge and represent variations in the strength of support for different constructs or paths that depict the relationships between them. Kawamura et al. (2009) offer an example of how textual symbols or shading can be used to signal elements of a figure where the qualitative and quantitative data are not consistent. An additional strategy is to introduce polyvocality (also recently called polyvocality) in the text (Creamer, 2011) involves acknowledging that team members could not reach consensus about the findings. These strategies have important consequences for developing integration strategies that acknowledge multifaceted findings. Of particular importance, engagement in individual and collective reflexivity enables researchers to interrogate what each individual brings to the integrated findings and the decision-making involved in the generation of meta-inferences or integrated insights.

In our own implementations of integration strategies throughout our educational mixed methods research, we have considered multifaceted findings in our integration strategies involving teams and dynamic conditions. Creamer recalls once having to navigate the political realities of a situation where research for a federally supported program to advance gender equity disclosed unpleasant realities about what had actually been accomplished. An additional example involves Poth having to broaden her team’s approach to integration in a study comparing teachers’ responses to items about personal responsibility for student motivation quantitatively and qualitatively because of unexpected divergence in findings and initial interpretations (Daniels et al., 2018). Having engaged in collective reflexivity with her interdisciplinary team to build a shared understanding of how the different disciplinary cultures and training among team members might influence their perspectives helped decisions about what and how to represent and discuss the multifaceted study findings as the research unfolded.

Dissemination of project results was the second most frequently occurring ethical issue reported by Stadnick et al. (2021), and the *most* challenging. Yet there was no discussion of effective strategies for addressing this. Similarly, our team’s experience and review of literature reveals that divergence as well as convergence in findings is to be expected and welcomed (Bazeley, 2018) and that authentic reporting is necessary to meet the ethical standard of *concern for justice* as a guiding ethical principle for ethical mixed methods research. The following reflexive questions are offered to guide mixed methods researchers in acknowledging multifaceted findings to generate self-understanding and shared understandings among mixed methods research team members:

- How are my/our decisions throughout the research process affecting the study findings? Have I/we included enough information so that a reader can understand and replicate the study?

- Have I/we shared enough information about my/our epistemological commitments and background experiences so that readers can understand which lenses I/we brought to the data interpretations? Have I/we given enough weight and value to each team member's epistemological views in our analysis and integration procedures?
- How are my/our decisions about which findings to include and how to organize the findings reflected in the final manuscript? Have I/we adequately protected the identities of our participants in our findings?

Communicate integration reports that fairly recognize contributions

The involvement of various researchers working under dynamic mixed methods research conditions introduces challenges for authorship when dissemination plans are emergent. The ethical principle guiding the *concern for welfare* extends to protecting the intellectual contributions of researchers in addition to protecting participants from harm and respecting the confidentiality or anonymity of the data they provide (Poth, 2021). Common under dynamic team conditions are opportunities for new ideas for publications and audiences to emerge during discussion of emerging results; yet, at the same time, additional challenges can be introduced when there is an undervaluing of individual contributions and power dynamics among team members. Specific complications for mixed methods research teams involve a lack of understanding of the time- and resource-intensive nature of the requisite integration and emergent processes (Curry et al., 2013; Sharp et al., 2012).

The contributions acknowledged by authorship often have to be renegotiated as a project unfolds because they cannot be entirely predicted at the onset of the reporting phase. It is important to highlight that authorship can be complicated and is not the only way to acknowledge contributions in mixed methods research. It is common to recognize editorial or minor contributions through a footnote or endnote. The practice of adding transparency to reporting through an endnote about the roles, perspectives, and contributions of team members is one strategy that helps a reader or audience member gauge the ownership for new ideas and original insight. Such approaches to communication have important consequences for fairly recognizing contributions. Of particular importance, engagement in individual and collective reflexivity enables researchers to interrogate what each individual brings to the work and authorship decision-making involved in the fair recognition of report contributions.

In our own communication of integration reports within our educational mixed methods research, we have considered fair recognition of authorship in our designs involving teams and dynamic conditions. Poth recalls having to adjust her educational team's authorship on a health-focused methodological contribution in the *Journal of Mixed Methods Research* to responsively involve a public health researcher in the manuscript (Poth et al., 2021). This means that teams should shift their practice to forming teams that can adapt to evolving understandings of the expertise that is needed as a project unfolds (Poth, 2019). Having engaged in collective reflexivity with her interdisciplinary team to build a shared understanding of how the team membership might be in flux and the impacts of that on authorship helped decisions about whom to involve as the research unfolded. The following reflexive questions are offered to guide mixed methods researchers in authentically recognizing contributions to generate self-understanding and shared understandings among team members:

- How do my/our previous publication experiences influence my/our thinking about what contributions I/we consider to be reasonable expectations for authorships? How might my/our conceptualizations of "reasonable expectations" differ from those among the research team members?
- To what extent and how was each team member involved in the design, implementation, analysis, and writing phases of the study? Have I/we considered the power dynamics that exist within the team and how they affect the work each person does?
- Has each team member been given an opportunity to weigh in on authorship order? How might my/our conceptualizations of authorship order differ from other team members'?

Concluding thoughts

Our impetus for this work was the lack of practical guidance for addressing the unique ethical dilemmas arising from integration and intensified by the complexity inherent to evolving mixed methods research designs and interdisciplinary team configurations. We welcomed the opportunity to describe promising ethical practices that addressed dilemmas we observed and experienced in our interactions with colleagues and graduate students in our varied roles as mixed methods researchers, instructors, team members, and dissertation supervisors. Our collaborative work on this chapter to advance three promising ethical practices leverages our collective expertise and experiences in diagnosing and managing dynamic mixed methods research conditions, research ethics, reflexivity, and research teams. We encourage others to do the same and to expand the descriptions of ethical dilemmas experienced beyond the three ethical practices described here and importantly, for others to learn from, the promising practices guiding successful navigation. A key takeaway points to the need for preparing mixed methods researchers to respond appropriately to ethical dilemmas as the research unfolds and to ethical reasoning informed by individual and collective reflexivity as central to the promising practices described in this chapter.

Acknowledgments

The chapter advances new insights and extends ideas we have previously published separately (e.g., Cain et al., 2019; Creamer, 2011, 2018, 2021; Poth, 2018a,b, 2019, 2021; Poth et al., 2021). We thank Adrienne Montgomerie for helpful comments on the draft manuscript. We are grateful for the opportunity to contribute to this encyclopedia and thank our colleagues from around the world who continue to influence our thinking as well as our Volume editor Dr. Nataliya Ivankova and the editors-in-chief for guiding this chapter through to publication.

References

- American Educational Research Association, 2006. Standards for reporting on empirical social science research in AERA publications. *Educ. Res.* 25 (6), 33–40.
- American Educational Research Association, 2021. What is Education Research. <https://www.aera.net/About-AERA/What-is-Education-Research>.
- Anderson, E.E., Corneli, A.L., 2018. 100 Questions (And Answers) About Research Ethics. Sage, Thousand Oaks, CA.
- Archibald, M.M., Radil, A.I., Zhang, X., Hanson, W.E., 2015. Current mixed methods practices in qualitative research: a content analysis of leading journals. *Int. J. Qual. Methods* 14 (2), 5–33.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage, London, UK.
- Burch, P., Heinrich, C.J., 2016. *Mixed Methods for Policy Research and Program Evaluation*. Sage, Thousand Oaks, CA.
- Cain, L.K., MacDonald, A., Coker, J., Velasco, J.C., West, G.D., 2019. Ethics and reflexivity in mixed methods research: an examination of current practices and a call for further discussion. *Int. J. Mult. Res. Approaches* 11, 144–155.
- Creamer, E.G., 2011. Experimenting with voice and reflexivity to produce multi-voiced texts. In: Conrad, C., Serlin, R.C. (Eds.), *The SAGE Handbook for Research in Education*, second ed. Sage Publications, Thousand Oaks, CA, pp. 367–380.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Method Research*. Sage, Thousand Oaks, CA.
- Creamer, E.G., 2021. *Advancing Grounded Theory With Mixed Methods*. Routledge/Taylor Francis, London, UK.
- Curry, L.A., Krumholz, H.M., O’Cathain, A., Plano Clark, V.L., Cherlin, E., Bradley, E.H., 2013. Mixed methods in biomedical and health services research. *Circ. Cardiovas. Qual. Outcomes* 6 (1), 119–123. <https://doi.org/10.1161/CIRCOUTCOMES.112.967885>.
- Daniels, L.M., Poth, C., Goegan, L., 2018. Enhancing our understanding of teachers’ personal responsibility for student motivation: a mixed methods study. *Front. Educ.* 3, 1–13.
- Evans, B.C., Coon, D.W., Ume, E., 2011. Use of theoretical frameworks as a pragmatic guide for mixed methods studies: a methodological necessity? *J. Mix. Methods Res.* 5, 276–292.
- Fàbregues, S., Molina-Azorín, J.F., 2017. Addressing quality in mixed methods research: a review and recommendations for a future agenda. *Qual. Quant.* 51, 2847–2863.
- Georgiou, G., Poth, C., Cruickshank, E., Kierstead, M., 2021. How Have Schools and Classrooms Responded to Learner Reading Needs During COVID 19? A Mixed Methods Case Study Examining Student Recording Scores and Instructional Responses. Department of Educational Psychology, University of Alberta.
- Glesne, C., 2011. *Becoming Qualitative Researchers: An Introduction*, fourth ed. Pearson, NY.
- Hesse-Biber, S.N., 2010. *Mixed Methods Research: Merging Theory With Practice*. Guildford Press, New York, NY.
- Hesse-Biber, S., 2016. Doing interdisciplinary mixed methods health care research: working the boundaries, tensions, and synergistic potential of team-based research. *Qual. Health Res.* 26 (5), 649–658.
- Kawamura, Y., Ivankova, N., Kohler, C., Purumean-Chaney, S., 2009. Utilizing mixed methods to assess parasocial interaction of an entertainment-education program audience. *Int. J. Mult. Res. Approaches* 3, 88–104.
- Macfarlane, B., 2010. Values and virtues in qualitative research. In: Savin-Baden, M., Howell Major, C. (Eds.), *New Approaches to Qualitative Research: Wisdom and Uncertainty*. Routledge, New York, NY, pp. 19–27.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N.G., Maxwell, J.A., Molina-Azor, J.F., Niglas, K., 2016. Expanding thinking through a kaleidoscopic look into the future: implications of the Mixed Methods International Research Association’s task force report on the future of mixed methods research. *J. Mix. Methods Res.* 10, 221–227.
- Onwuegbuzie, A., Poth, C., 2016. Editors’ afterword: toward evidence-based guidelines for reviewing mixed methods research manuscripts submitted to journals. *Int. J. Qual. Methods* 15, 1–13.
- Onwuegbuzie, A.J., 2012. Introduction: putting the mixed back into quantitative and qualitative research in educational research and beyond: moving towards the radical middle. *Int. J. Mult. Res. Approaches* 6, 192–219.
- Plano Clark, V.L., Anderson, N., Wertz, J.A., Zhou, Y., Schumacher, K., Miaskowski, C., 2015. Conceptualizing longitudinal mixed methods designs: a methodological review of health sciences research. *J. Mix. Methods Res.* 9, 297–319.
- Plano-Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Pluye, P., Grad, R.M., Levine, A., Nicolau, B., 2009. Understanding divergence of quantitative and qualitative data (or results) in mixed methods studies. *Int. J. Mult. Res. Approaches* 3, 58–72.
- Popa, F., Guillermin, M., 2017. Reflexive methodological pluralism: the case of environmental valuation. *J. Mix. Methods Res.* 11, 19–35. <https://doi.org/10.1177/1558689815610250>.
- Poth, C., Cain, L.K., Aquilina, A.M., 2021. Guiding Instructional Practices for Learner-Centred Mixed Methods Research: A Case Study Examining Significant Experiences. Department of Educational Psychology, University of Alberta.
- Poth, C., 2012. Exploring the role of mixed methods practitioner within educational research teams: a cross case comparison of the research planning process. *Int. J. Mult. Res. Approaches* 6, 315–332. <https://doi.org/10.5172/mra.2012.6.3.314>.
- Poth, C., 2018a. The curious case of complexity: implications for mixed methods practice. *Int. J. Mult. Res. Approaches* 10, 1–9.
- Poth, C., 2018b. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. Sage, London, UK.
- Poth, C., 2019. Realizing the integrative capacity of educational mixed methods research teams: using a complexity-sensitive strategy to boost innovation. *Int. J. Res. Method Educ.* 42 (3), 252–266. <https://doi.org/10.1080/1743727X.2019.1590813>.
- Poth, C., 2021. *Research Ethics*. Sage, London, UK.
- Preissle, J., Glover-Kudon, R., Rohan, E.A., Boehm, J.E., DeGroff, A., 2015. Putting ethics on the mixed methods map. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford, New York, NY, pp. 144–156.
- Sharp, J.C., Mobley, C., Hammon, C., Withington, S., Drew, S., Stringfield, C., Stipanovic, N., 2012. A mixed methods sampling methodology for a multisite case study. *J. Mix. Methods Res.* 6, 34–54.
- Stadnick, N., Poth, C., Maxwell, J.A., Guetterman, T.C., 2021. Strategies for promoting ethical integrity in mixed methods health services research. *Health Serv. Res.* 21, 577–586.
- Symon, A.G., Whitford, H., Dalzell, J., 2012. Infant feeding in Eastern Scotland: a longitudinal mixed methods evaluation of antenatal intentions and postnatal satisfaction—the Feeding Your Baby study. *Midwifery* 29, e49–e56.
- Wallen, G.R., Baker, K., Stolar, M., Miller-Davis, C., Ames, N., Yates, J., et al., 2012. Palliative care outcomes in surgical oncology patients with advanced malignancies: a mixed methods approach. *Qual. Life Res.* 21, 405–415.

Further reading

- Cain, L.K., MacDonald, A., Coker, J., Velasco, J.C., West, G.D., 2019. Ethics and reflexivity in mixed methods research: an examination of current practices and a call for further discussion. *Int. J. Mult. Res. Approaches* 11, 144–155.
- Creamer, E.G., 2011. Experimenting with voice and reflexivity to produce multi-voiced texts. In: Conrad, C., Serlin, R.C. (Eds.), *The SAGE Handbook for Research in Education*, second ed. Sage Publications, Thousand Oaks, CA, pp. 367–380.
- Stanick, N., Poth, C., Maxwell, J.A., Guetterman, T.C., 2021. Strategies for promoting ethical integrity in mixed methods health services research. *Health Serv. Res.* 21, 577–586.

Mixed methods and evaluation

Donna M. Mertens, Gallaudet University, Rockville, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Mixed methods in educational evaluation	531
Distinction between evaluation and research	531
Paradigm wars and current landscape of philosophical frameworks in evaluation	531
Designing mixed methods studies: different perspectives	532
Postpositivism and mixed methods	532
Constructivism and mixed methods evaluation	533
Pragmatism and mixed methods evaluation	534
Transformative mixed methods and evaluation	534
Indigenous mixed methods evaluation	536
Conclusion: mixed methods evaluation is multi-disciplinary and multi-perspectival	536
References	537

Mixed methods in educational evaluation

Program evaluation is a field that shares much common ground with applied research, however, it has developed as a separate discipline because of the political context in which evaluation is conducted (Mertens and Wilson, 2019). The purpose of evaluation is to inform decision making and to support sustainability of interventions, such as programs, policies, products, processes or plans (Mertens and Tarsilla, 2015). Evaluation in education is sometimes thought of as a method to determine the effectiveness of programs, however, evaluation can play a broader role when data are collected to identify important contextual factors (contextual analysis), need for improvement or changes in practice (needs assessment), or the quality of implementing an intervention (process evaluation). Because educational evaluation is conducted in the messy, political world of schools and agencies, many complex issues can arise and a wide range of people can be impacted by the interventions (known as stakeholders). The use of mixed methods in educational evaluation provides a means to address the complex, political, and human issues in this context.

Distinction between evaluation and research

Research and evaluation can look very similar, especially if the research is conducted to address a specific problem, such as effectiveness of an educational intervention. Both disciplines rely on the collection of systematic data to understand the phenomenon. The purpose of research has been defined as the collection of data to construct theory or create knowledge (Mertens, 2020). Thus, researchers can choose the context and focus of their studies based on their intellectual interests. The purpose of evaluation is to inform decision making and support sustainability of an intervention (Mertens and Wilson, 2019). Evaluators conduct their studies at the request of a stakeholder group (e.g., a funding agency, program director, policy maker, or community representatives) in contexts of resource scarcity, leading to competition for program support. The evaluator needs to be able to negotiate this political landscape through appropriate inclusion of stakeholders and provision of data viewed as credible by those stakeholder groups. To answer this challenge, many evaluators have turned to the use of mixed methods. The relationship between mixed methods research and evaluation in education is displayed in Fig. 1.

In this chapter, the evolution of the paradigms in educational evaluation is described. The assumptions of the five dominant paradigms (post-positivist, constructivist, pragmatic, transformative and Indigenous) in educational evaluation are presented. This is followed by examples of the design of mixed methods studies in educational evaluation for each paradigmatic perspective in order to demonstrate the impact of the evaluator's assumptions on methodological choices.

Paradigm wars and current landscape of philosophical frameworks in evaluation

Humans engage in evaluation activities everyday when they make decisions about what to eat, which movie to watch, or what to wear. However, the profession of evaluation was formally established in the United States in the 1960s because the government began to require evaluation of programs that were funded under Lyndon Johnson's Great Society initiatives. Many other countries have followed suit and evaluation is now a global practice. Evaluation's roots are very multi-disciplinary, reflecting eclectic methodological approaches. Evaluations conducted in its infancy years tended to rely on quantitative designs and methods that were informed by the postpositivist paradigm (Mertens and Wilson, 2019). However, qualitative methods were introduced to the evaluation community largely through the scholarly work of Guba and Lincoln (1989). The introduction of qualitative methods into this profession stirred up great controversy about the best methods for conducting evaluations; this period is known as the

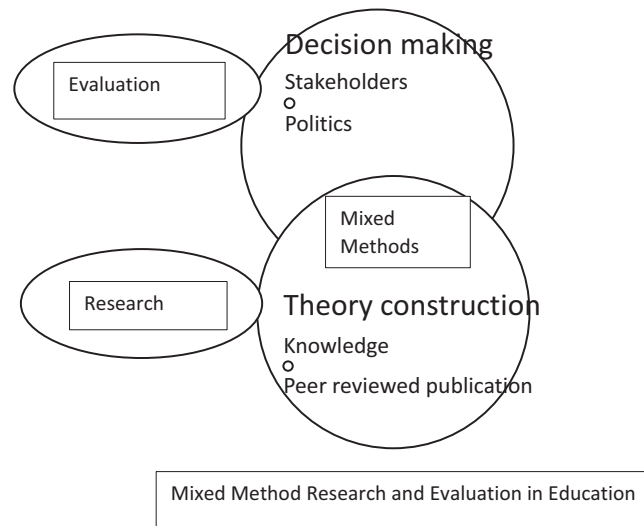


Fig. 1 Relationship between mixed methods research and evaluation in education.

“paradigm wars.” However, Guba and Lincoln were able to establish a way forward by explicating the concept of paradigm frameworks to make clear that decisions about quantitative or qualitative methods were based on fundamental assumptions about the nature of ethics and values (axiology), reality (ontology), value of knowledge and relationships with stakeholders (epistemology), and systematic inquiry (methodology). The assumptions thus drive the methodological decisions.

The evaluation community operates with five different paradigmatic frameworks: Postpositivism, constructivism, transformative, pragmatism, and Indigenous (Chilisa and Mertens, 2021; Chouinard and Cram, 2020; Mertens and Hesse-Biber, 2013; Mertens, 2018; Mertens and Wilson, 2019). Mixed methods can be applied to evaluations conducted within all of these paradigmatic frameworks, however, the assumptions associated with each lead to different decisions regarding the use of mixed methods. Briefly, postpositivists prioritize the use of quantitative methods with qualitative methods playing a supportive role (White, 2013). Constructivists prioritize qualitative methods with quantitative methods serving a supportive role (Hesse-Biber, 2010). Transformative evaluators rely on qualitative methods to build culturally respectful relationships across stakeholder groups and to understand the experiences of the full range of stakeholders in a culturally complex context (Mertens, 2018; Mertens and Wilson, 2019). These methods are combined with quantitative data to provide a fuller understanding of context and program effectiveness. Pragmatic evaluators use methods that they deem are necessary for answering the evaluation questions (Creswell and Plano Clark, 2018). Indigenous evaluators also prioritize building respectful relationships that are culturally responsive in order to make visible the legacy of colonialism and its continuing impact on the lives of Indigenous peoples (Chouinard and Cram, 2020). Quantitative methods can be included in Indigenous evaluations in order to understand context, track progress, and demonstrate effectiveness. The sections that follow provide examples of evaluations that are conducted using mixed methods within each of the five paradigmatic stances.

Designing mixed methods studies: different perspectives

Postpositivism and mixed methods

The postpositivist assumption of ontology, the nature of reality, is that there is one reality that exists and the evaluator needs to be able to measure that reality with a certain level of certainty. Thus, the choice for data collection tends to be a quantitative measure that has a known level of reliability and validity. The postpositivist epistemological assumption rests on the belief that evaluators need to be objective, defined in terms of not engaging with stakeholders in order to prevent personal biases from affecting the study. The preferred mixed methods design under this paradigm is a randomized control trial (RCT) when feasible and a quasi-experimental design if necessary. The quantitative data are prioritized and the qualitative data serve a complementary role.

White (2013) argues that postpositivism is the best framing for a mixed methods study because it prioritizes obtaining an answer to what he considers the most important question: Did the intervention work? He sees opportunities for integrating the collection of qualitative data into an RCT design in order to understand the appropriateness of recruitment strategies and to document the implementation process of the intervention. This question of fidelity of treatment, i.e., is the intervention being implemented as planned without making changes, is often answered through the collection of qualitative data (Mertens, 2020). This is necessary to ensure that the intervention that was implemented is the same intervention that was hypothesized as being an effective way to address a problem.

Edmunds et al. (2021) conducted a mixed methods study that was framed by the post-positivist paradigm. They used a RCT design to determine the comparative effectiveness of sections of a course that had the goal: “to increase the number of students, particularly students of color, completing online courses and improve the academic performance of those students, with the ultimate goal of increasing the percentage of students who remained in postsecondary education (p. 5).” The original plan for the research was to randomly assign students to course sections that were taught by instructors trained in the intervention and control instructors who taught as they normally would. In this part of the study, quantitative and qualitative data were collected by a survey with both open and closed ended questions that students completed. The researchers also collected quantitative data on Blackboard participation and outcome variables (e.g., course completion). Additional qualitative data were collected via interviews with staff.

The results of this first phase revealed: “The evaluation found that the intervention had overall (on average) statistically significant impacts on reducing withdrawals and positive, although not statistically significant, impacts on successful course completion (Edmunds et al. p. 2).” However, the researchers were intrigued by the finding that the results differed considerably when comparing the two experimental sections to each other. This finding led them to add a second phase to their study which provides further illustration of using a postpositivist framework to a mixed methods evaluation in education because they wanted to understand why the intervention seemed to work in one section and not the other. In addition to re-analyzing the data collected in the first phase, the researchers collected new quantitative data on the instructor characteristics and qualitative data through observations and interviews with staff. They found that students in Section A had significantly more log-ins for Blackboard than students in Section B. This analysis allowed them to have a comparative measure of exposure to the treatment in the two experimental sections. The observational data were collected by reviewing videos of the instructors’ uses of the experimental treatment strategies in their lectures. Section A instructor incorporated more of the treatment strategies than did Section B instructor; this exemplifies a mixed methods approach to process/implementation evaluation. This finding was supported by the interview data with the staff who reported differential use of the treatment strategies. The researchers acknowledged that a limitation of their study was that they did not interview the students, thus, they did not have the students’ perspectives of the course implementation and effectiveness. The researchers also discussed a very practical problem: they were not able to randomly assign instructors to the sections because one of the instructors in the experimental group had designed the intervention.

The logic of RCT is based on the prioritization of objectivity, control, standardized procedures and measurements, and maintaining distance between the program participants and the evaluators. Embedding a qualitative component into a study can create tensions because a truly interpretive framing would require an open-ended, emergent, close relationship between evaluators and participants. Logistical reasons can also prevent researchers from including a strong qualitative component, as Edmunds et al. (2021) noted, they had not planned to do the second stage of their study and they had very limited resources to carry out the second phase. Thus, the use of mixed methods under the postpositivist paradigm requires sacrifice of the assumptions of a constructivist approach to mixed methods. Mixed methods under the constructivist paradigm is specifically designed to address issues of context and participant experiences.

Constructivism and mixed methods evaluation

Evaluators can conduct their inquiries using only qualitative approaches, thus reflecting the core assumptions of the constructivist paradigm. However, evaluators have tended to combine the use of qualitative and quantitative methods in order to be responsive to the multiple stakeholders and to answer the different kinds of evaluation questions. Evaluators who work within the constructivist paradigm tend to design mixed methods studies that prioritize the qualitative methods and use quantitative data in a supportive role (Mertens and Hesse Biber, 2013). As constructivists, evaluators use more sophisticated and detailed qualitative designs, such as adaptations of ethnographic methods or case studies.

Loizzo et al. (2018) conducted a qualitatively driven mixed methods evaluation of a Massive Open Online Course (MOOC) in journalism for social change. The course was designed to increase journalism students’ knowledge of how to mobilize citizen journalists around social justice topics. The intervention was a 7-week MOOC that was offered at a university in the United States. The course was developed by the instructors who were interested in documenting the strategies they used to teach the course, as well as the successes and challenges they experienced. They wanted to know what the students identified as successes and challenges in the course and how it affected their attitudes toward constructivist journalism for social justice. The major focus of the study was on the experiential data for instructors and students, therefore, qualitative data collection was dominant. The evaluators used a mixed methods design that incorporated virtual ethnography and a quantitative survey. Virtual ethnography is used to study online communities to understand their cultures and interactions. The data collection methods consisted of overt observations of the MOOC on a weekly basis, collection of screen captures of artifacts used by the instructor, and review of documents such as syllabus, curriculum materials, student assignments, instructor announcements, and discussion posts. They also conducted interviews with the instructor and five of the learners two weeks after the end of the course. The interviews were conducted via email and Skype.

The quantitative portion of the study consisted of an email survey designed to assess the students’ attitudes. The responses were made using a 5-point scale. The results of the survey were analyzed and the means and standard deviations were reported. The qualitative data were analyzed using open codes through which categories emerged such as MOOC purpose, passionate and motivated instructor and students, and child welfare context. The evaluators presented the findings using constructivist practices of explaining the codes that emerged and providing exemplary quotations from the documents, interviews, and observations to support the relevance of the theme. They were also able to identify several challenges in learning through the MOOC format that provided opportunities for feedback to the instructor on modifications to improve the course. They integrated the survey findings into the

presentation of the qualitative themes to support the students' positive attitudes toward the course and applying the concepts that they learned.

Gubrium et al. (2016) provide another example of a constructivist mixed methods evaluation in their assessment of the effects of a digital storytelling intervention on the self-esteem, social support, and sexual attitudes of Puerto Rican girls. Digital story telling asks participants to create a short narrative account of a relevant life event using digital images, audio recording, music and text. The creation of the digital story is accomplished through small-group workshops. The process used in digital storytelling consists of individual work to write a script, collect images, and record. This is followed by a group process that is co-mediated by facilitators and work group members. Another group process phase occurs in the form of a story circle and story screening. The evaluation focused on the process and effect of these workshops held in a city with high health risks and poverty. Quantitative data were collected using established psychosocial measures. They reported that very little evidence of effectiveness was produced by the quantitative measures perhaps because these measures are not able to capture the effects of a process that involves writing, storytelling, visual representations, sound and voice, and esthetics. However, the qualitative data revealed vibrant positive changes and health benefits for the participants.

An ethnographic approach was used for the qualitative component of the study. Evaluators observed and assisted with each workshop; they created field notes; and the sessions were audiotaped. In addition, the evaluators conducted interviews with each participant. The documents for analysis included transcripts from the workshops, story circles and story screenings, interview transcripts, field notes, debrief meetings, and participant responses to a workshop evaluation. The evaluators concluded that the emphasis on the qualitative component of the study was appropriate given the cultural context and the demographics of the target population. They were able to demonstrate positive changes in self-esteem, social support and sexual attitudes through the words and images of the participants.

Pragmatism and mixed methods evaluation

Evaluators work in the real world; they encounter challenges that might preclude the use of a RCT or a rich qualitative approach. Pragmatism is characterized by the logic that evaluation methods should be chosen that will allow the evaluator to answer the prescribed questions. Thus, some questions can be answered with qualitative data, others with quantitative data, or at times, both types of data are needed. These evaluation questions in the pragmatic paradigm are generally generated by the funding agency in consultation with program staff and the evaluation team. Akerblad et al. (2020) used this lens to frame their mixed methods evaluation of vocational rehabilitation programs in Finland. The purpose of the study was to evaluate the "implementation and effectiveness of the rehabilitation model and its effects and benefits from the perspective of the service participants, workplace, occupational health services and rehabilitation service provider" (p. 7).

It should be noted that under this lens, the intervention was already decided by the rehabilitation specialists and the evaluation questions were set before the evaluation team entered the picture. The questions are derived from the purpose: How was the program implemented? How effective was the program for the different stakeholder groups? The evaluators developed a mixed methods design to answer these questions for their client. The data to answer the implementation question came from individual and group interviews with the human resource specialist, occupational health service providers and rehabilitation service providers. The data on effectiveness was gathered using quantitative questionnaires before and after the intervention; these were administered to service participants, supervisors and health service providers. The data from the various sources were analyzed separately before the quantitative and qualitative data were put into conversation with each other. The quantitative data were given priority and the qualitative data were interpreted within that framework. The results indicated that the supervisor played a critical role in insuring that the multiple actors collaborated effectively. The use of this mixed methods design answered the client's questions; it did not provide rich contextual information that is typically generated when a constructivist lens is used, as was described previously in this chapter. In the next section, the use of mixed methods from a transformative stance illustrates how prioritizing issues of justice and culture influences methodological choices.

Transformative mixed methods and evaluation

Postpositivist, pragmatic, and constructivist evaluators design their studies to align with their assumptions. They can, but are not compelled to, consider the issues of power and inclusion of members of marginalized communities. In the education sector, marginalized communities include people of color, Indigenous community members, sexual minorities, people who are disabled or Deaf, refugees, immigrants, and students from high poverty areas. Addressing issues of power, historical inequalities, cultural responsiveness, inclusion of the full range of stakeholders, and designing evaluations to support increased social, economic, and environmental justice are fundamental principles for transformative evaluators (Mertens, 2018, 2020; Mertens and Wilson, 2019). These are issues that educators know are important for sustainable change.

Transformative mixed methods evaluations utilize a cyclical approach that begins with a first phase of relationship building across the full range of stakeholders, with particular attention to groups who have historically been marginalized (See Transformative Mixed Methods, Fig. 1, Mertens, in this volume). Evaluators may need additional time to figure out the best way to work with the full range of stakeholders in order to demonstrate cultural responsiveness, consciously address issues of power, and provide safe environments for sharing knowledge and experience that comes from people with lived experience in the communities. This phase

of relationship building is a qualitative process that needs to be evaluated for inclusiveness and culturally responsive working strategies.

The relationship building phase leads to a phase for contextual analysis that includes collection of both qualitative and quantitative data in order to understand the history, political structures, economic conditions, heterogeneity in the populations, incidence data, and any other relevant indicators such as educational or employment characteristics. The results of these first two phases are used as the basis for the development or refinement of an intervention that can then be pilot tested using intensive collection of both quantitative and qualitative data on the strategies used for recruitment, training, implementation, barriers, challenges, strategies for overcoming barriers and challenges, and effectiveness of the piloted intervention. The data from this phase can be used to make refinements to any aspect of the intervention or to make decisions to try something quite different if the pilot is not successful. Once an intervention is viewed as appropriate in a given context, then it can be implemented on a wider scale with quantitative and qualitative data collected on variables similar to those used during the pilot testing. Qualitative data reveals the process of implementation and the quality of experiences. Quantitative data can reveal participation rates and effectiveness measures. Finally, the evaluation needs to include the use of the process and impact data for transformative purposes, e.g., refinement or change of intervention, changes in policies, and strategies to strengthen relationships.

Culturally responsive approaches in evaluation are commensurate with the transformative paradigm (Bledsoe and Donaldson, 2015). Drill et al. (2017) used a transformative mixed methods approach with a culturally responsive lens in their evaluation of a project titled "Bars to Bridges." The project was designed to support students of color who were in the juvenile justice system in their transition back into regular school. The multi-phase cyclical study began with building relationships with school personnel, juvenile justice agents, community organizations, families, and youth (See Fig. 2). The evaluators did not assume that they knew what evaluation questions to ask or how to collect data. In keeping with the transformative culturally responsive approach, they formed discussion groups with these stakeholders to formulate evaluation questions and develop data collection strategies. They conducted a contextual analysis that involved collection of data on relevant legislation (qualitative) and disproportional rates of discipline practices by race based on extant quantitative data. This first phase that included the collection of both quantitative and qualitative data (concurrent mixed methods) led them to identify diverse stakeholders who needed to be included and gave direction to their next data collection phase.

The project staff developed an intervention based on the data collected in the first phase of the evaluation. The second phase of the evaluation was primarily qualitative and included examining the quality of relationships with the students, reviewing program documents, and conducting interviews and focus groups with a conscious awareness of the importance of being culturally responsive with the full range of stakeholders. The data from the second phase documented the primary activities and the challenges and successes in implementing the intervention. The evaluators also collected data on the extent to which the stakeholders viewed the intervention and the evaluation as being culturally responsive. The evaluators collected quantitative and qualitative data concurrently in the third phase to establish the quality of the experience for the students, transition specialists in the schools, and the juvenile justice personnel. The quantitative data consisted of extant data on such variables as recidivism and school attendance. The evaluators consciously employed the process and results of their study for transformative purposes. They raised the need to address attitudinal barriers that prevented a smooth transition back into the schools, as well as other logistical barriers. They also met with state level policy makers to develop a state-wide protocol for school re-entry.

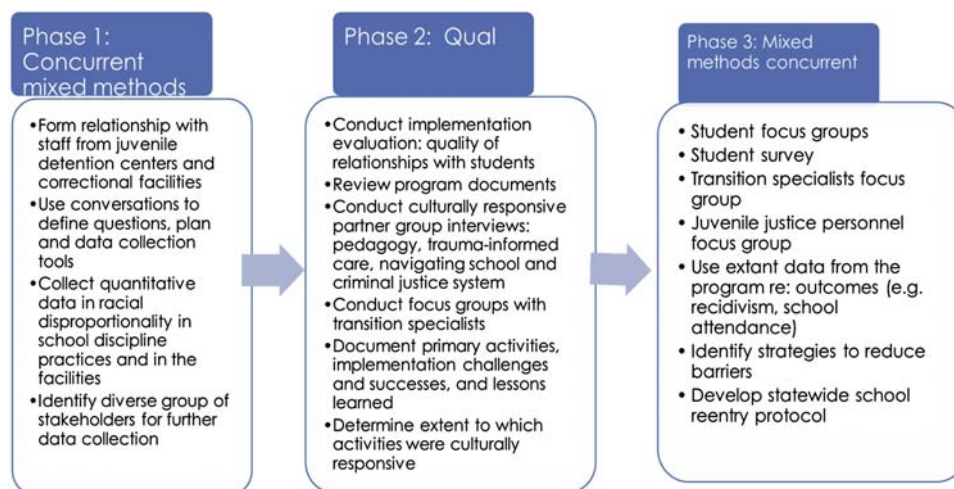


Fig. 2 Transformative mixed methods cyclical design of bridges to bars project (Drill et al., 2017).

Indigenous mixed methods evaluation

Indigenous and transformative evaluations share many commonalities; however, Indigenous scholars have identified important differences (Chilisa and Mertens, 2021; Cram and Mertens, 2015; Cram et al., 2013). One important difference is that Indigenous evaluators recognize the critical impact of a legacy of colonialism that contributes to current inequalities (Chilisa, 2020, 2015). All Indigenous scholars are not the same; uniqueness is associated with different Indigenous groups and also within Indigenous groups. Indigenous scholars from the United States (Lafrance and Nichols, 2010), Africa (Chilisa, 2020), and Aotearoa (New Zealand) (Cram and Mertens, 2015; Cram et al., 2018) provide different perspectives based on their countries of origin. However, they do share the demand that they have the authority to define problems, solutions, and evaluation strategies that are culturally appropriate for their own communities. Culturally responsive Indigenous evaluations can use many different methodologies, e.g., participatory, community-based, and tribal participatory methods can be used, as long as they are practiced through an Indigenous lens (Chouinard and Cram, 2020). An Indigenous lens requires a contextual analysis that reflects the history of colonialism, including genocide, racism, expropriation of land, and suppression of Indigenous knowledge.

Hamerton et al. (2012) conducted an evaluation of a program designed to encourage changes in behavior toward a healthier lifestyle in Aotearoa New Zealand. The program was located in communities that served the Maori population because quantitative data indicated that members of this community manifest greater health disparities than in the White communities and that programs that had been imposed on Maori had not been effective. The goal of the program was to support healthier eating and exercise choices based on information from the community itself. Thus, the Healthy Eating Healthy Action program was begun by providing funding to six Maori health agencies who worked with community members and health experts to develop interventions built on cultural values and practices that could be delivered with close connection to the communities. Although the study was implemented in health agencies, it has implications for educators who are concerned with addressing healthier life choices by their Indigenous students. The plans included activities to improve nutritional knowledge, increase physical activity, educate particular groups, and reduce obesity.

The evaluation used an Indigenous lens in the form of Kaupapa Maori, or by, for and with Maori (Cram et al., 2018); the evaluators were both Maori and non-Maori. The evaluation began by spending considerable time building relationships with the evaluators, community health agency staff, and other members of the Maori community. This meant they met face-to-face with the stakeholders, visited the programs, and listened carefully to their experiences. They conducted a contextual analysis that included collection of qualitative and quantitative data. The quantitative data included measures of physical activity and individual health indicators such as weight, blood pressure, and BMI. The qualitative component of the design consisted primarily of interviews and focus groups; they followed Maori custom of bringing healthy food to these data collection events. They consciously included opportunities for reciprocity in the evaluation by sharing information with the stakeholders about successful initiatives. The data collection instruments incorporated the 4-sided model of Maori health that includes body, mind and emotions, spirituality and extended family. Focus groups looked at not only individual benefits, but also benefits that accrued to families, community, and tribe.

The results indicated that the knowledge gained and behavioral changes made by older members of the community were valued by younger members. The communities began to support local gardens that gave them access to healthier foods. This resulted in healthier food being served in community centers as well as in schools. As the program was developed within communities and they were privy to the results as they emerged, the health agencies assumed greater control and responsibility for modifying the interventions as they gained insights. This is an indicator of sustainable change because of community ownership.

Conclusion: mixed methods evaluation is multi-disciplinary and multi-perspectival

Evaluators come from many different disciplines, such as education, psychology, sociology, health, social work, economics, and business. Thus, its history is replete with the use of different types of designs and data collection methods that are associated with these different disciplines. Many of the issues faced in the educational system require interdisciplinary solutions. Mixed methods approaches provide a model for working across disciplines for these types of problems. The growth of mixed methods enhances the evaluation community's ability to examine the assumptions that drove methodological decisions. The current status of having five paradigms that inform mixed methods evaluators is viewed as a rich opportunity to learn more about methodologies. Rather than having arguments about the incommensurability of different paradigms, scholars who situate themselves in each of these paradigms incorporate quantitative and qualitative methods that align with their assumptions. Evaluators in the postpositivist and pragmatic paradigms tend to accept that their responsibility is to collect data that is responsive to their client (funder). Evaluators in the constructivist paradigm are more inclined to ask about the range of stakeholders and how their methods can provide a rich picture of the program under study. Transformative and Indigenous evaluators are more likely to stress the importance of building relationships with the full range of stakeholders and incorporating quantitative and qualitative data into the design in order to understand the cultural complexity of the community of stakeholders. They are also more inclined to use the data from earlier stages as support for the development of an intervention that is culturally responsive and to answer evaluation questions that are of importance across the stakeholder groups. The conversations across paradigms within the context of the use of mixed methods provides fertile ground for growth in understandings of effective methodologies in educational evaluation.

References

- Akerblad, L., Seppanen-Jarvela, R., Haapakoski, K., 2020. Integrative strategies in mixed methods research. *J. Mix. Methods Res.* 1–19. <https://doi.org/10.1177/1558689820957125>.
- Bledsoe, K., Donaldson, S., 2015. Culturally responsive theory-driven evaluation. In: Hood, S., Hopson, R., Frierson (Eds.), *Continuing the Journey to Reposition Culture and Cultural Context in Evaluation Theory and Practice*. Information Age Publishing, Charlotte, NC, pp. 3–28.
- Chilisa, B., 2015. A synthesis paper on the made in Africa Evaluation concept. *Afr. Eval. Assoc.* <https://afrea.org/wp-content/uploads/2018/06/MAE2-Final-31st-august.pdf>. (Accessed 4 February 2020).
- Chilisa, B., 2020. *Indigenous Research Methodologies*, second ed. Sage, Thousand Oaks, CA.
- Chilisa, B., Mertens, D.M., 2021. Indigenous made in Africa evaluation frameworks: addressing epistemic violence and contributing to social transformation. *Am. J. Eval.* 42 (2), 241–253. <https://doi.org/10.1177/1098214020948601>.
- Chouinard, J.A., Cram, F., 2020. *Culturally Responsive Approaches to Evaluation*. Sage, Thousand Oaks, CA.
- Cram, F., Mertens, D.M., 2015. Transformative and Indigenous frameworks for multimethod and mixed methods research. In: Hesse Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, UK, pp. 91–109.
- Cram, F., Chilisa, B., Mertens, D.M., 2013. The journey begins. In: Mertens, D.M., Cram, F., Chilisa, B. (Eds.), *Indigenous Pathways Into Social Research*. Left Coast Press, Walnut Hills, CA, pp. 11–400.
- Cram, F., Pipi, K., Paipa, K., 2018. Kaupapa Maori evaluation in Aotearoa New Zealand. *N. Dir. Eval.* 159, 63–77.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*. SAGE, Thousand Oaks, CA.
- Drill, K., Stack, E.E., Qureshi, C., Lahoff, R., 2017. Oregon's African American/Black Student Success Plan. Oregon Department of Education, Office of Equity, Diversity, and Inclusion, Salem, OR.
- Edmunds, J.A., Gicheva, D., Thrift, B., Hull, M., 2021. Using mixed methods to explore variations in impact with RCTs: the case of Project COMPASS. *J. Mix. Methods Res.* 1–21. <https://doi.org/10.1177/15586898211033144>.
- Guba, E., Lincoln, Y.S., 1989. *Fourth Generation Evaluation*. Sage, Newbury Park, CA.
- Gubrium, A.C., Fiddian-Green, A., Lowe, S., DiFulvio, G. De, Toro-Mejias, L., 2016. Measuring down: evaluating digital storytelling as a process for narrative health promotion. *Qual. Health Res.* 26 (13), 1787–1801.
- Hamerton, H., Mercer, C., Riini, D., McPherson, B., Morrison, L., 2012. Evaluating Maori community initiatives to promote health eating, healthy action. *Health Promot. Int.* 29 (1), 60–69.
- Hesse Biber, S., 2010. Qualitative approaches to mixed methods practice. *Qual. Inq.* 16 (6), 455–468.
- Lafrance, J., Nichols, R., 2010. Reframing evaluation: defining an indigenous evaluation framework. *Can. J. Program Eval.* 23, 13–31.
- Loizzo, J., Watson, S.L., Watson, W.R., 2018. Examining instructor and learner experiences and attitude change in a journalism for social change massive open online course: a mixed methods case study. *J. Mass Commun. Educ.* 73 (4), 392–409.
- Mertens, D.M., Hesse Biber, S. (Eds.), 2013. *Mixed Methods and Credibility of Evidence in Evaluation*. New Directions for Evaluation, p. 138.
- Mertens, D.M., Tarsilla, M., 2015. Mixed methods evaluation. In: Hesse Biber, S., Johnson, B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford UK, pp. 426–446.
- Mertens, D.M., Wilson, A.T., 2019. *Program Evaluation Theory and Practice: A Comprehensive Guide*, second ed. Guilford, NY.
- Mertens, D.M., (in this volume) Transformative mixed methods research. In: Tierney, R., Rizvi, F., Erkican, K. (Eds.), *International Encyclopedia of Education*, 4th ed. In press.
- Mertens, D.M., 2018. *Mixed Methods Design in Evaluation*. Sage, Thousand Oaks, CA.
- Mertens, D.M., 2020. *Research and Evaluation in Education and Psychology: Integrating Diversity With Quantitative, Qualitative, and Mixed Methods*, fifth ed. Sage, Thousand Oaks, CA.
- White, H., 2013. The use of mixed methods in randomized control trials. In: Mertens, D.M., Hesse Biber, S. (Eds.), *Mixed Methods and Credibility of Evidence in Evaluation*. New Directions for Evaluation. Jossey-Bass, San Francisco, CA, pp. 61–73, 138.

Combining mixed methods and case study research

Vicki L. Plano Clark^a, Lori A. Foote^a, and Janet B. Walton^b, ^a School of Education, University of Cincinnati, Cincinnati, OH, United States; and ^b College of Education, NC State University, Raleigh, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	538
Defining mixed methods research and case study research	539
Mixed methods research defined	539
Case study research defined	540
Frameworks for combining mixed methods and case study research	540
Examining cases within a mixed methods study	540
Mixing methods within a case study	541
Intersecting the two approaches to form a mixed methods case study	542
Applying the three frameworks	542
Key decisions when combining mixed methods and case study research	543
Case study design and case selection	543
Mixed methods design and timing	544
The use of integration strategies to enhance case interpretation	545
Applying the key decisions	546
Applications of mixed methods case study in education	546
Future directions	547
Concluding comments	548
References	549

Glossary

Case study research a predominately qualitative approach to educational inquiry that is dedicated to developing in-depth understanding of one or more cases that represent bounded systems (Creswell and Poth, 2018; Yin, 2018)

Intersecting the intentional combination of the assumptions, designs, and methods of two distinct research approaches to form a hybrid approach to research (Plano Clark and Ivankova, 2016)

Mixed methods case study research an approach to educational inquiry that mindfully combines the assumptions, designs, and methods of mixed methods research and case study research in ways that integrate the qualitative and quantitative findings to generate enhanced knowledge of the case(s) under investigation (Guetterman and Fetters, 2018; Plano Clark et al., 2018)

Mixed methods research an approach to educational inquiry that combines at least one qualitative method (textual data) and one quantitative method (numeric data) within a single study such that the two types of data and analysis are brought together to respond to the study's aims (Johnson et al., 2007)

Introduction

There are many circumstances in education where it is beneficial to have a comprehensive understanding of the complexities of what happened in a particular situation. For example, teachers could benefit from knowing what happened when other teachers incorporated a new technology into their classroom during a semester. Student affairs officers could benefit from a description of a social justice program implemented over two years on a campus experiencing racial tensions. Educational policymakers could benefit from a comparison of what happened in different countries in the first three months following the occurrence of a health crisis. Each of these examples highlights a complex situation occurring within a system of people over some period of time where practical knowledge about what happened is needed.

To study complex real-world situations such as these, educational researchers have a wide range of research approaches from which to choose. Among the choices, some educational researchers are drawn to mixed methods research approaches, which provide a comprehensive understanding by combining quantitative and qualitative information about a complex situation (Creswell and Plano Clark, 2018; Johnson and Onwuegbuzie, 2004). At the same time, other educational researchers are drawn to case study research approaches, which provide an in-depth description and interpretation of a complex situation in one or more settings (Stake, 1995; Yin, 2018). Given the relevance of both mixed methods and case study approaches for examining the complex situations of interest in education, it is perhaps unsurprising that an increasing number of educational researchers choose to combine mixed methods research and case study research in their studies. With the increasing use of such combinations, educational

researchers, practitioners, policy makers, and other stakeholders interested in educational issues need to consider how these two approaches can be productively combined in research studies as well as understand how to identify and assess research that uses both mixed methods and case study research.

To address this need to understand and critically evaluate the combined application of mixed methods and case study research approaches to examine educational situations, this chapter discusses important assumptions and considerations involved when scholars combine the two research approaches in their studies. We begin by formally defining the two approaches separately before introducing frameworks that describe different perspectives for how the two approaches can be combined within educational research studies. We then present key decisions that go into designing a study combining mixed methods and case study research and that are useful for understanding and evaluating the study approach. From there, we highlight several examples of applications in education where mixed methods and case study were combined. We conclude by offering future directions for mixed methods case study in education and recommendations for educators interested in reading, reviewing, and conducting research using mixed methods and case study.

Defining mixed methods research and case study research

Before discussing the combined use of mixed methods research and case study research, it is essential that they each first be defined. This is no easy task as both of these methodologies have long histories influenced by scholars from different disciplines and with different perspectives about research. Here we limit the focus to definitions that are commonly accepted and applied within education and emphasize aspects most salient to considering how the two approaches can be combined in research studies. Readers interested in more complete treatments of the diversity of perspectives associated with these two methodologies are encouraged to examine other chapters in this volume (e.g., Schoonenboom, in press; Stake and Visse, in press) as well as the extensive literature available about mixed methods research and about case study research.

Mixed methods research defined

Mixed methods research is an approach to educational inquiry that combines two research components: one qualitative method (textual data) and one quantitative method (numeric data) within a single study such that the two types of data and analyses are brought together to respond to the study's aims (Johnson et al., 2007). In a mixed methods study, using one methodological approach simply will not answer all parts of the questions posed (Tashakkori et al., 2021). By carefully designing the qualitative and quantitative components of the study, investigators are able to intentionally integrate the various methods so that results maximize the response to research questions (Fetters and Freshwater, 2015). Integration occurs when researchers bring the different study components together into conversation with each other (such as by comparing qualitative themes to quantitative statistical results). Integration can be accomplished by connecting or merging the different components into a new whole so that new insights are generated from the combination (Creswell and Plano Clark, 2018).

Although education scholars agree that achieving integration is central to mixed methods research (Creswell and Plano Clark, 2018; Tashakkori et al., 2021), they often bring different assumptions about the nature of the social world and how they are able to learn about it through research. Several distinct sets of philosophical assumptions are associated with mixed methods research (Shannon-Baker, 2016). Educational researchers using mixed methods research often draw from ideas aligned with pragmatism (Johnson and Onwuegbuzie, 2004; Morgan, 2014), dialectical pluralism (Greene, 2007; Johnson, 2017), critical realism (Maxwell and Mittapalli, 2010), and transformative perspectives (Mertens, 2012). These different perspectives describe researchers' mental models for how knowledge is generated when combining different approaches and therefore shape what is being mixed and for what reasons in mixed methods studies.

In combination with philosophical assumptions, educational researchers invoke a mixed methods approach to their inquiries for several reasons (Greene et al., 1989). A widely cited rationale for using mixed methods is triangulation, where researchers directly compare the findings from the quantitative portion of a study (e.g., survey results) and the qualitative portion of a study (e.g., focus group findings). Triangulation provides assurance that the description of the phenomenon is stable across methodological components. Another rationale for using mixed methods is complementarity, which aims to produce elaborated or clarified results by using the strengths of the different methods to examine different facets of the phenomenon of interest. Complementarity allows the qualitative and quantitative components to each illuminate unique aspects of the examined phenomenon so that, when combined, they generate expanded understandings. A third rationale for using mixed methods is initiation, which intentionally seeks to locate areas where cross-method agreement is not present, resulting in what may appear to be, at first glance, contradictory findings. Initiation requires researchers to dig deeper into the analyses and probe additional perspectives to integrate the findings from the quantitative and qualitative parts into a meaningful interpretation. In short, reasons for invoking a mixed methods approach to research vary, but always aim to enhance knowledge of complex and multifaceted issues such as those found in education. Further discussion of the nuances in how mixed methods research is defined and the importance of different rationales are found elsewhere (for more on this, see Hesse-Biber and Johnson, 2015).

Case study research defined

Case study research is a predominantly qualitative approach to educational inquiry that is dedicated to developing in-depth understanding of one or more case(s) that represent bounded systems (Creswell and Poth, 2018; Yin, 2018). Case study researchers aim to understand how a phenomenon of interest (e.g., an activity, program, or event) exists in particular time(s) and place(s) (e.g., a classroom over the course of a year) (Creswell and Poth, 2018; Yin, 2018).

As with mixed methods research, researchers who work from a variety of philosophical positions are drawn to use case study research to address their aims and research questions. In education, commonly discussed perspectives associated with case study research include postpositivism, constructivism, and critical realism (Harvey, 2009; Schwandt and Gates, 2018; Yazan, 2015). Each of these perspectives gives a different emphasis to the application of case study research in education by placing importance on objectively describing a case, interpreting the meaning of the case, or identifying causal mechanisms acting within the case. Thus, the researcher's perspective informs how and why researchers select cases and the kinds of knowledge derived from the study of the cases.

Keeping these variations in mind, researchers using case study often invoke this approach for several reasons. First, case studies are selected for their flexibility of methods. Researchers employing a case study design may collect qualitative data from observations, interviews, and artifacts; however quantitative data, such as from surveys or test scores, are also well-suited for use in a case study (Luck et al., 2006; Yin, 2018). Moreover, researchers also have flexibility when choosing whether to investigate single or multiple cases (Merriam, 1998; Stake, 1995; Yin, 2018). Another reason for choosing a case study design is the ability to focus both on the phenomenon under investigation and on how it operates within the specific context of the case. A case study approach blurs the boundaries between phenomenon and context, allowing researchers to access understandings of the ways the context and phenomenon may interact with one another (Creswell and Poth, 2018; Merriam, 1998; Yin, 2018). A final reason for choosing case study methodology relates to the multiple ways cases themselves can be defined within a study. While cases can be defined by geographical location, they can also be conceptualized more abstractly (Ragin, 1992; Yin, 2018). As an example, a case can be concrete, such as a student or a school, but a case could also be more abstract, such as an incident or a group of schools deemed alike in some theoretical manner. Further discussion of other potential definitions of a case and the importance of describing what constitutes a case in a study can be found elsewhere (for more on this, see Ragin, 1992).

Frameworks for combining mixed methods and case study research

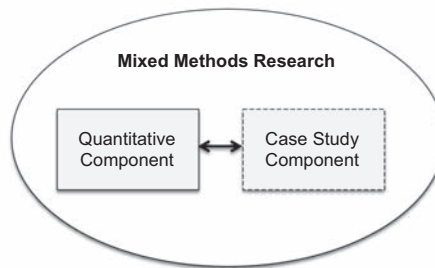
Although mixed methods research and case study can be defined as stand-alone and distinct inquiry approaches, there are many examples where researchers have combined the two in their studies to address their questions of interest. Researchers across disciplines including education are drawn to the ability of mixed methods approaches to provide comprehensive insights and the ability of case study approaches to provide rich descriptions of complex systems (Guetterman and Fetters, 2018). These two approaches can work well together, but the combinations can be challenging to plan, understand, and evaluate because of all the different methods that may be employed and the resulting variety of methodological considerations necessary. Here we highlight three frameworks useful for understanding different combinations found within education: (a) examining cases within a mixed methods study, (b) mixing methods within a case study, and (c) intersecting the two approaches to form a mixed methods case study. The frameworks differ in terms of the relative emphasis placed on the mixed methods approach and the case study approach for addressing the study's purpose. These differences are depicted in Fig. 1 and described in the following sections.

Examining cases within a mixed methods study

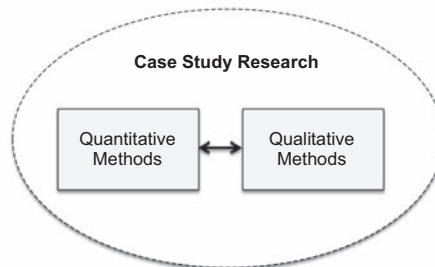
When researchers conduct and report mixed methods research studies, by definition these studies include both a quantitative component and a qualitative component (Creswell and Plano Clark, 2018). Therefore, one way that researchers combine mixed methods and case study research is to use a case study approach for the qualitative component of a mixed methods study. There are three different times when the qualitative case study component can occur in relation to the quantitative component: it can occur first and inform the quantitative component, it can occur at the same time as the quantitative component for comparison, or it can occur after the quantitative component to explain aspects of the quantitative results. In each of these possibilities, the case study component provides in-depth information about cases to complement the quantitative aspects of the mixed methods studies.

As depicted in diagram (A) of Fig. 1, educational applications that correspond to this framework are described as mixed methods studies where the researchers use a case study approach for the qualitative component of the study's mixed methods design. Yin (2018) referred to applications such as this as studies where "the larger (mixed methods) study encompasses the case study" (p. 235). More recently, mixed methods scholars Guetterman and Fetters (2018) similarly described this framework as having a "parent" mixed methods study that includes a "nested" case study (p. 901). As these descriptions convey, this framework emphasizes the use of mixed methods as the guiding design for the study with the use of case study for the qualitative component in support of the overall mixed methods study goals.

A Examining Cases Within a Mixed Methods Study



B Mixing Methods Within a Case Study



C Intersecting the Two Approaches to Form a Mixed Methods Case Study

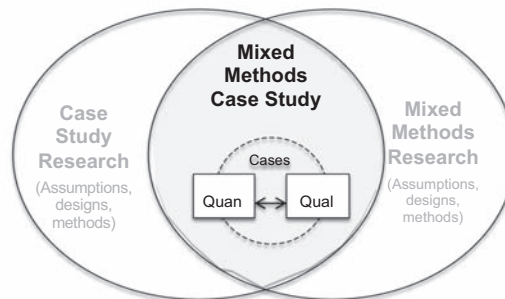


Fig. 1 Three different frameworks for combining mixed methods and case study research, (A) examining cases within a mixed methods study, (B) mixing methods within a case study, (C) intersecting the two approaches within a mixed methods case study.

Ivankova and Stick's (2007) mixed methods study of persistence in an online doctoral program is an example of a study that illustrates examining cases within a mixed methods study. Working from a pragmatic perspective, the authors used a mixed methods design that started with quantitative data collection and statistical analyses to identify internal and external factors that differentiated four groups of students defined by their level of persistence. Next, they conducted a qualitative multiple case study that examined the experiences of four typical students (one from each group) to explain the role of the identified factors. As such, this study used a sequential mixed methods design that incorporated a case study approach for the second, qualitative phase to provide an expanded understanding about the initial quantitative results.

Mixing methods within a case study

As depicted in diagram (B) of Fig. 1, the second framework for conceptualizing combinations of mixed methods and case study switches the emphasis between the two approaches from the previous framework. In this framework, the case study approach is the guiding methodology for the study and the researcher chooses to mix quantitative and qualitative methods to meet the case study objectives. Yin (2018) noted that researchers could use both quantitative and qualitative evidence to better describe the case(s) investigated (p. 63). Similarly, Guetterman and Fetters (2018) described this framework as having a "parent" case study that includes a "nested" mixed methods design (p. 901).

For many who are interested in the use of case study in education, mixing methods within a case study is simply an example of doing good case study research (Byrne, 2009; Yin, 2018). Australian scholars Luck et al. (2006) argued that case study was particularly well suited to provide a framework for combining quantitative and qualitative methods. They described the case study

methodology as being able to provide a bridge between the different methods. For example, scholars might use different methods (e.g., interviews, document analysis, surveys, and observations) to examine different stakeholders' experiences (e.g., administrators, teachers, students, and parents) and contexts of interest (e.g., demographics, performance, trends over time) to form in-depth descriptions of the cases. Such studies can be considered as case study research, but with a close examination of the details, it is apparent that the researchers combined both qualitative and quantitative data to enhance understanding of cases of interest.

Auslander et al. (2020) is an example of a study that illustrates mixing methods within a case study. The researchers designed a holistic, single-case study to examine teacher candidates' preparation for a newly mandated teacher assessment within the context of a university mathematics methods course. Data collected from all participants ($N = 51$) included pre- and post-course belief questionnaires containing open-ended qualitative and close-ended quantitative items related to pedagogy and teaching efficacy. Additionally, a portion ($n = 6$) of the participants were interviewed. Results of the quantitative and qualitative analysis were reported separately. No shift in teacher candidates' beliefs were detected in the quantitative analysis; the robust qualitative analysis described four themes that emerged from the interviews and open-ended questionnaire data. This study represents this framework since the emphasis was on describing the case with little attention given to mixed methods aspects such as integration of the different data and results.

Intersecting the two approaches to form a mixed methods case study

A third framework used to conceptualize the combination of mixed methods and case study research comes from the concept of intersecting. *Intersecting* is the intentional combination of the assumptions, designs, and methods of two distinct research approaches to form a hybrid approach to research (Plano Clark and Ivankova, 2016). From this perspective, researchers can intentionally combine the assumptions, designs, and methods of mixed methods research with the assumptions, designs, and methods of case study research to form a hybrid, mixed methods case study approach (Plano Clark et al., 2018; Walton et al., 2020). This conceptualization of mixed methods case study as the intersection of mixed methods research and case study research is depicted in diagram (C) of Fig. 1.

Building on these ideas, *mixed methods case study research* can be defined as an approach to educational inquiry that mindfully combines elements of each approach in ways that integrate the qualitative and quantitative findings to generate enhanced knowledge of the case(s) under investigation (Guetterman and Fetters, 2018; Plano Clark et al., 2018). As a hybrid approach, mixed methods case study can be considered as a complex research application with its own variants, procedures, and quality considerations drawing upon the many options and variations available for mixed methods research and case study research (Creswell and Plano Clark, 2018). In contrast to the previous two frameworks that gave a clear priority to one of the approaches (either mixed methods research or case study research) for providing the guiding methodology to a study's design, the intersecting perspective conceptualizes the two methodologies as equal partners in a study's design, which is depicted by the two larger overlapping circles in Fig. 1C, with negotiation between the different assumptions, designs, and methods.

Walton's (2014) mixed methods case study of a collaborative partnership to support STEM education is an example of a hybrid approach that emphasized both mixed methods and case study equally in the study design. The researcher's pragmatic approach pressed upon her the need for both mixed methods and case study thinking when approaching the examination of the partnership's structure and activities. The theoretical underpinnings of the study suggested the usefulness of a dual focus on the level of collaboration in the partnership and on partners' perceptions of the application of research-based collaboration practices. This theoretical framework led the researcher to a concurrent mixed methods approach to the study and drove the integration of findings from a quantitative measure of collaboration with findings from qualitative data gained from partner interviews. At the same time, the study was framed within a partnership that was deemed unique based upon its leadership structure. The uniqueness of the case led naturally to a case study approach that aimed to provide a rich description of the partnership and insight into the impacts of the leadership structure. The case description and interpretation, a dialogue between the broad view of collaboration and the fine-grained partner perspectives of partnership practices, had as its foundations the mixed methods integration strategies suggested by the study's theoretical framework.

Applying the three frameworks

Each of the three described frameworks provides a slightly different way of thinking about how case study research can be combined with mixed methods research in educational studies. They each tend to draw on different literature, use different terminology, and have study purposes and questions that emphasize different aspects of the studies. However, they also have many commonalities, such as the study of a case or cases, the use of both quantitative and qualitative methods, and the intention to learn from complex, real-world situations. Each of these frameworks is present and useful within education so it is helpful for reviewers and consumers to be aware of the different perspectives they may encounter when reading published reports of studies that combined mixed methods and case study research. Investigators planning studies will likely find that they are more drawn to one of these frameworks depending on how the combined approaches can best address the particular question of interest. As such, readers are encouraged not to think of these three frameworks as either-or types and certainly not to value one over others. Instead, each framework highlights a different way that these two useful approaches can be used together in research studies to advance understanding about important educational issues and systems.

Key decisions when combining mixed methods and case study research

All researchers, regardless of the types of studies they undertake, should plan those studies with careful consideration of issues such as their philosophical perspective and what research design best fits the research questions they seek to answer. Because of its utility and flexibility, a combined mixed methods case study approach is suitable for researchers who are associated with a variety of the philosophical perspectives discussed previously, including pragmatism, postpositivism, dialectical pluralism, critical realism, and transformative perspectives. For all of these perspectives, the mixed methods case study approach offers the ability to combine quantitative and qualitative information about cases to develop comprehensive and nuanced conclusions in response to specific research questions about complex phenomenon occurring within bounded systems.

Once researchers have chosen to use mixed methods case study to investigate research problems, they face decisions specific to combining case study and mixed methods approaches in a study. These decisions influence the orientation, depth, and breadth of a study's findings. Therefore, consumers of research should have a basic familiarity with key decision points as they read reports of mixed methods case study research for information to guide practice and policy. These decisions relate to key aspects of case study design, mixed methods design, and the use of integration strategies to enhance the understanding of cases.

Since the various decision points researchers navigate in conducting mixed methods case studies are best illustrated in the context of applied research, the sections that follow will rely on two hypothetical exemplar studies. In each of the hypothetical studies, the researcher aims to understand the phenomenon of challenges associated with online teaching in primary grades, however the researchers in the two studies each make key decisions that result in very different study designs and findings.

In the first hypothetical study, which we refer to as the "online bilingual school study," the researcher investigated instructional challenges faced by a fully online school using a dual-language immersion approach. This researcher collected data concurrently from teacher interviews and a survey administered in the same two-week period in which interviews were conducted. The study culminated in a case report that wove together individual (interview) and school-level (survey) challenges to online instruction in order to create a holistic picture of challenges within this school.

In the second hypothetical study, referred to herein as the "regional school study," the researcher conducted a study that also aimed to understand challenges in online teaching, but this researcher chose to conduct a multiple case study to examine a sample of schools taken from one geographic region. This researcher employed a survey to identify several schools that represented various demographic features and student outcomes and then used administrator interviews and student performance data from online courses at the selected schools to create reports for each case. These case reports depicted relationships between school-level resources (from the survey and interviews), school demographics (from the survey), and students' learning outcomes (from student performance data). The individual cases were then compared with the aim of understanding the differences in learning outcomes between schools with varying resources and demographic features.

The two hypothetical studies serve as a representation of just two of almost limitless ways in which mixed methods and case study can be combined. They illustrate the ways in which studies with similar purposes, both utilizing mixed methods case study approaches, can diverge in the scope and focus of their findings based on the researchers' choices. Building on the concepts already discussed, the following sections use the two hypothetical studies to illuminate three key sets of decisions regarding: (a) case study design and case selection, (b) mixed methods design and timing, and (c) the use of integration strategies to enhance case interpretation.

Case study design and case selection

The first key decisions we highlight for mixed methods case study approaches relate to the case study design and case selection. A researcher's choice of case study design in terms of the number of cases chosen for study (single case study or multiple case study) is a decision guided by the intent of the study. [Table 1](#) provides an overview of possible rationales for case selection in both single and multiple case studies along with the study purposes that guided case selection in each design. An examination of the two hypothetical studies reveals nuanced differences in study purposes and case selection rationales that led to the two different designs employed.

The researcher in the bilingual school study aimed to understand the challenges associated with online teaching in a case of particular interest. The researcher chose the case because of the unique setting of the case (i.e., online content being delivered in more than one language). The bilingual nature of instruction was a unique feature of this school that was intrinsically interesting to the researcher, and this study therefore aligns with [Stake's \(1995\)](#) conceptualization of an intrinsic case study. According to [Stake \(1995\)](#), an intrinsic case study is one that examines a single case that is interesting in itself, where the study is conducted in order to understand something about the particular case. Intrinsic case studies do not aim to create knowledge that can be applied across a variety of contexts.

The regional school study differs in the scope of its aims. Here, the researcher's purpose was to understand how learning outcomes in online courses are related to school context features. The purpose of this study was instrumental, meaning that the cases were studied in order to understand a phenomenon that transcended the individual cases ([Stake, 1995](#)). In this study, the interaction between school context and learning outcomes was the focus and, while the findings may not be broadly generalizable because of the small sample size of cases, other schools with similar contextual features might be guided by the study's findings.

Table 1 Case study designs: Case selection rationales and study purposes.

<i>Case type</i>	<i>Purpose of the study</i>	<i>Case selection process</i>
Single case – Intrinsic	To illuminate a single case as a particularly interesting instance of a phenomenon	Interest in the case as an individual situation may precede the study design; single cases may be critical, unusual, common, revelatory, or longitudinal
Single case – Instrumental	To explore a phenomenon within a bounded case to understand something beyond the case itself	Select a single case that is: critical, unusual, common, revelatory, or longitudinal; this decision relates to which cases provide the best opportunity to answer the research question
Multiple case – Collective (literal replication)	To examine a phenomenon in several cases that are similar to one another on a key dimension	After compiling a set of cases that meet the criteria for study inclusion, cases are reviewed for varying secondary characteristics. Cases are selected to broaden the opportunity to learn about the phenomenon of interest in its many manifestations and based on wider variety
Multiple case – Comparative (theoretical replication)	To compare and contrast cases based on their membership in distinct groups to determine how the phenomenon of interest exists in each of these groups	Cases from each group are selected to address rival explanations; for example, stratified or matched sampling across groups helps to control for other case features that may impact the comparison

Case study design ideas from [Stake \(1995\)](#) and [Yin \(2018\)](#).

The hypothetical studies highlight the role of case selection as a key decision point for researchers. The case study design, which includes the number and identity of cases selected, is closely tied to the purpose of the study in question. Determining the number of cases and which to select may be a process that interacts iteratively with formulating research purposes and questions. Researchers should first consider whether examining activities within a specific case or cases will provide insight into the research questions posed.

The study's purposes and research questions may call for multiple cases to be investigated and compared in order to understand differences across cases chosen for their varying contextual features. In some studies, the case selection process may represent a unique phase of the study and rely on quantitative methods to determine cases' eligibility for inclusion. For example, educational effectiveness researchers often group schools into comparative sets by their student outcomes while quantitatively controlling for demographic school features. These researchers use an initial quantitative phase to classify schools into sets for comparison, then make final case selections from these sets of schools (e.g., [Teddlie et al., 1985](#)).

The hypothetical examples and the example of educational effectiveness studies highlight some of the challenges associated with case study design, which include determining the number and selection of cases. Researchers conducting intrinsic case studies must provide a rationale for the in-depth investigation of the particular case ([Yin, 2018](#)) and must carefully determine the unit of analysis (e.g., a school, one class within a school, or a single grade level across several schools in a region) that will be investigated to address the study's aims ([Ishak and Bakar, 2014](#)). Likewise, researchers conducting instrumental case studies need to determine which cases will best provide information about the phenomenon that is the focus of the study, with the recognition that such decisions may necessitate a first phase of robust analysis of cases to inform well-reasoned case selection.

Mixed methods design and timing

A researcher choosing to use a mixed methods approach to a case study is faced with an array of possibilities for collecting, analyzing, and integrating quantitative and qualitative data and findings. Researchers' decisions about how and when to collect various types of data are driven by their purposes or rationales for incorporating multiple forms of data into their studies. [Greene et al. \(1989\)](#), [Bryman \(2006\)](#), and others have advanced sets of rationales for using mixed methods that include sampling or case selection, triangulation, complementarity, explanation, and instrument development. The researcher's rationale for collecting multiple types of data may be explicitly stated in the case study report. However, more often the rationale is implicit, and the reader of research studies is left to infer the rationale based upon the study's purposes and designs ([Heyvaert et al., 2013](#)).

Several typologies have been advanced to describe study designs where qualitative and quantitative methods are combined in a single study (e.g., [Creswell and Plano Clark, 2018](#); [Guest, 2013](#); [Tashakkori et al., 2021](#)). Researchers frequently use language that reflects the timing of data collection and analysis as well as the priority of qualitative and quantitative methods and data to describe their mixed methods designs. [Morse \(1991\)](#) advanced a notation system that relies on the abbreviated terms "quan" and "qual" to refer to quantitative or qualitative data, capitalizes one or the other term to indicate relative priority (e.g., QUAL to describe a study with a qualitative priority or QUAN to describe a study with a quantitative priority), and uses arrows (→) and plus symbols (+) to depict the timing and flow of data collection and analysis. [Creswell and Plano Clark \(2018\)](#) advanced a framework of three core

Table 2 Mixed methods designs: Timing and notation considerations.

Core mixed methods design	Timing of data collection and analysis	Possibilities for notation
Convergent	Qualitative and quantitative data are collected and analyzed simultaneously (or concurrently) to develop a comprehensive picture	(a) QUAL + quan (b) qual + QUAN (c) QUAL + QUAN
Explanatory sequential	Quantitative data is collected and analyzed in a first phase; first phase findings are used to inform a second phase in which qualitative data is collected and analyzed to explain quantitative findings	(a) QUAN → qual (b) quan → QUAL
Exploratory sequential	Qualitative data is collected and analyzed in a first phase; first phase findings are used to explore an issue and inform the design of a quantitative phase (e.g., a survey instrument) used to determine how the first phase findings relate to a different or larger population	(a) QUAL → quan (b) qual → QUAN

Mixed methods design ideas from [Creswell and Plano Clark \(2018\)](#) and [Morse \(1991\)](#).

mixed methods designs that focus on the timing of data collection and analysis: convergent design (simultaneous collection of qualitative and quantitative data), explanatory sequential design (a quantitative phase followed by a qualitative phase), and exploratory sequential design (a qualitative phase followed by a quantitative phase). [Table 2](#) presents these core designs and provides examples of how they may be described in literature using [Morse's \(1991\)](#) notation system.

Examining the two hypothetical studies through the lenses of mixed methods design and rationale reveals key decisions addressed by each researcher in terms of the mixed methods designs and timing employed. In the bilingual school study, the researcher collected and analyzed both interview and survey data simultaneously, but relied primarily on interview findings in the production of the case report, using quantitative survey findings to lend further insight into the qualitative findings. The researcher's rationale for using multiple data sources was the complementarity of the interview and survey constructs and the desire to use the two datasets to create a more complete picture of the challenges of online instruction in the bilingual school. Using [Morse's \(1991\)](#) notation system, the timing and mixed methods design of the study could be described as QUAL + quan.

In contrast, the regional school study used a sequential design with two distinct phases. First, the researcher defined four sets of context and outcomes criteria based on a guiding theoretical framework for the study, and employed a quantitative survey to select four school cases from the region, each of which met one of the criteria sets. In a second phase of the study, the researcher collected qualitative (interview) and quantitative (student performance) data at each of the selected schools, and generated a case report for each. The challenges to online instruction in each setting were then compared using the initial theoretical framework as a guide. The researcher's rationale for using multiple types of data included sampling (to inform the case selection) and complementarity (to gather complementary data to inform a broader and more nuanced understanding of the challenges to online instruction).

The design of the regional school study is not as easily described by any one of [Creswell and Plano Clark's \(2018\)](#) three core designs or [Morse's \(1991\)](#) notation system. From a timing perspective, the study utilized an initial quantitative phase followed by a second, convergent phase in which both quantitative and qualitative data were collected and analyzed simultaneously, a design that might be described using [Morse's \(1991\)](#) notation system as QUAN → [QUAL + QUAN]. The latter example points to the challenges associated with applying existing typologies to mixed methods case study designs due to the many options available for researchers for the key decisions.

The typologies and notation system discussed here are examples of commonly used frameworks for discussing mixed methods designs and are not exhaustive. While the savvy consumer of research may wish to become familiar with these conventions, he or she should remain aware of their limitations in depicting mixed methods case studies. As [Walton et al. \(2020\)](#) noted, conventions of timing and priority may fall short in attempts to describe mixed methods case studies. Likewise, [Creswell and Plano Clark \(2018\)](#) acknowledge these challenges, noting that "complex designs" (p. 102) that intersect mixed methods designs with other research approaches such as case study may not be fully described by any one of the core designs.

The use of integration strategies to enhance case interpretation

The third type of decision of interest in mixed methods case study are decisions about integration and interpretation. The use of mixed methods integration strategies to create linkages between qualitative and quantitative methods, datasets, and findings is a defining feature of mixed methods research (e.g., [Onwuegbuzie and Johnson, 2006](#); [Fetters, 2020](#)). In fact, methodologists have proposed that it is the use of such integration strategies that differentiates a mixed methods study from a study that, while it may collect multiple forms of data, holds data and findings in parallel without drawing relationships between them (e.g., [Plano Clark et al., 2018](#)). Mixed methods integration strategies drive data analysis procedures and the representation and interpretation of findings, and are typically closely linked to a study's purposes and to the mixed methods design employed ([Creswell and Plano Clark, 2018](#)).

There are numerous points within a study where qualitative and quantitative research approaches and procedures can interface, encompassing all phases of a study from the posing of research questions to the drawing of inferences from a study's findings (Fetters et al., 2013). Mixed methods integration strategies include text-based or tabular side-by-side comparisons of qualitative and quantitative data findings (Guetterman et al., 2015); the transformation of qualitative data to quantitative data (quantitizing) or vice versa (qualitizing) (Nzabonimpa, 2018); and using the findings of one strand of a study to inform a second, sequential strand (e.g., using qualitative data to identify important context features of sites to inform case selection) (Creswell and Plano Clark, 2018). Regardless of the strategy used, the ultimate goal of the integration strategies is to bring quantitative and qualitative findings into conversation with one another in order to provide new or more holistic insights into the issues under investigation.

The researchers in the hypothetical studies made various decisions to create conversations between the quantitative and qualitative databases and to make enhanced interpretations about the cases under study. In the bilingual school study, for example, the researcher compared and contrasted themes from teacher interviews with related findings from the survey constructs, presenting the findings from this process in a table where quantitative findings from a survey construct (e.g., ease of access to technology support) were displayed beside interview quotes reflecting either common or divergent findings. This joint display of combined findings served as a guide for case interpretation in an in-depth case report that outlined challenges to online teaching in this setting.

In the regional school study, the researcher had the opportunity to employ various integration strategies throughout the study. This researcher, for example, quantitized data from administrator interviews by counting codes for three constructs related to the theme of resource scarcity in the interview findings. The researcher then compared the relative frequency of these constructs within the interview data from each site to student outcome data in order to better understand how perceptions of resource scarcity related to student outcomes in online learning. The regional school study researcher also used findings from the first, quantitative, phase of the study in which demographic data for all schools in the region were collected as a contextual backdrop for the interpretation of individual cases.

Applying the key decisions

The decisions described in this section ultimately determine the focus and breadth of a study's findings and how these findings can be used. In order to understand the impact of these decisions on the use and further applications of a study's conclusions, consumers of research should analyze reports of research to understand how and why researchers combined these approaches and made decisions regarding case study and mixed methods elements of the study. Mixed methods case study researchers should therefore carefully consider decisions regarding how case study and mixed methods approaches are combined and should articulate their rationales for these choices in their study reports.

Applications of mixed methods case study in education

To illustrate the key decisions and considerations that have been discussed, in this section we present a brief examination of three existing education studies employing mixed methods case study. We have intentionally included different contexts and design elements to pique readers' interest as they consider the wide variety of possibilities available for designing their own mixed methods case study investigations.

In our first example, Duffin and Perry's (2019) mixed methods case study used a convergent design in which all data were collected simultaneously. These researchers aimed to understand how a Place-Based Ecology Education (PBEE) community, an intrinsically interesting case, contributed to changes in teacher practice. The PBEE community was an example of a unique collaborative approach to providing coordinated, impactful professional development to primary and secondary teachers situated in schools in the northeastern portion of the US. The researchers collected data concurrently using surveys and evaluator notes. Surveys generated both quantitative (Likert-scale numeric responses) and qualitative (open responses in participants' own words) data; additional qualitative data came from evaluators' notes. The researchers analyzed the quantitative data to determine whether changes in practice were statistically significant (t-tests) and to gauge the varying impact of different amounts of professional development (regression and cluster analysis). They separately analyzed the qualitative data to identify themes related to implementation and support needs. All results were then brought together for additional integrative analysis by displaying qualitative themes alongside quantitative professional development dosages and the stages of development reported. By merging the two types of data to produce a more complete understanding of the case, the researchers could "discover insights into best practices for PBEE professional development" (Duffin and Perry, 2019, p. 7).

Our second exemplar study is Wise et al.'s (2012) mixed methods case study, which aimed to capture individual students' experiences during asynchronous online interactions in an undergraduate course in Canada. Multiple cases were chosen to better understand the range of practices exhibited by students in online discussion boards. To account for the variety of activity, three student cases were selected based on a preliminary analysis indicating that these cases exhibited extreme online behaviors (e.g., in terms of frequency of visits to discussion board, length of session, frequency of self-reviews). The researchers collected quantitative and qualitative data concurrently as students interacted on an online discussion board. Data included quantitative log-file data measuring engagement time, determined reading versus skimming, and situated posts in timelines, while content of the posts served as the

Table 3 Example of a side-by-side joint display used to merge results.

Theme	Context 1 schools (schools in poorer, less-educated communities)	Context 2 schools (schools in wealthier, better-educated communities)
Assignment of students to classes	Heterogenous: Mix of all abilities in each classroom; to spread out the level of need; all types of need present in each classroom	Homogeneous: Students assigned to classes by ability (high with high, low with low); to allow teachers to modify instruction for one type of need; assignment of aides to support lowest achieving group
Assessment on math progress	Screening tests: 3 times a year; plus weekly monitoring (computer)	Screening tests: 2–3 times per year
Team meetings to review data	Weekly (teachers only)	2–3 times per year (teachers only or full diagnostic team)
Math supports for “strugglers” (without IEPs; Tier 2)	Provided by teacher	Provided by teacher, aides, and intervention specialists
Support for students with diagnosed math disabilities (Tier 3)	30 min per day with intervention specialist	Aide support during math class and 30 min per day with intervention specialist

Modified from Foote (2019, p. 186). Context 1 Schools and Context 2 Schools (seen in the second and third columns of this table) were two of the theoretical groups of school cases generated using quantitative methods (cluster analysis). The themes related to the operations of the schools; these were informed by the best-practices literature and qualitative differences present when comparing the two context types. This joint display allows readers to directly compare how operations in each context varied by theme. Ultimately, while schools in both contexts attempted to employ recommended best practices, needs and resources associated with each type of case resulted in different operations.

qualitative dataset related to the study’s theoretical framework. Integrating findings by creating joint displays of qualitative and quantitative findings allowed the investigators to generate narratives richly characterizing each student case and offering nuanced insights regarding conceptualizations of student online behaviors.

Our final example is Foote’s (2019) mixed methods case study aimed at understanding how elementary mathematics supports exist in different school contexts in the US. Using a sequential design, the researcher started by gathering quantitative data and used cluster analysis to generate groups of schools based on demographic features, resulting in seven unique sets of schools. Two of these groups of schools were selected for comparison based on maximum differences in contexts, and two individual school cases were selected from each selected group (for a total of four cases for study). Qualitative data were gathered for each case through interviews and classroom observations, then analyzed using thematic coding. All data were brought together in a comparative table (see Table 3) that displayed similarities and differences in schools’ methods for supporting mathematics learning. Results indicated that supports for students who struggled in mathematics varied by school context, providing a more nuanced view of mathematics supports in elementary schools.

As this examination of several mixed methods case studies illustrates, researchers using this combined mixed methods case study approach have the opportunity to integrate qualitative and quantitative methods to produce robust inferences related to single or multiple cases. In each educational example described, there is evidence that the researchers conceptualized how case study and mixed methods research elements could be fruitfully placed within an overarching design to respond to research aims and generate important insights about complex educational situations.

Future directions

Mixed methods case study has emerged as a valuable methodological option for educational researchers wanting to generate practical knowledge about what is happening within educational systems. This combined research approach has the potential to highlight examples of promising educational practices and suggests novel approaches to investigating important educational research questions and the local results of regional, national, and global policies. As the methodological thinking regarding mixed methods case study evolves, new ideas emerge for ways to conceptualize mixed methods case studies. In addition, the increasing prevalence of global issues impacting and interacting with the field of education suggest new areas of inquiry for researchers, new ways that researchers can conceptualize and bound cases for investigation, and the novel use of data and data analysis strategies within case studies.

As methodological understandings of mixed methods case study advance, new ways to conceptualize and discuss mixed methods case study are emerging (e.g., Guetterman and Fetters, 2018; Plano Clark et al., 2018). Due to the emphasis on description and the small number of cases examined in most studies, case study has long been perceived as a qualitative methodology with extremely limited potential for generalizability and has therefore been dismissed by some researchers seeking to influence public policy or generalize a study’s conclusions across an array of settings (Tsang, 2014). Mixed methods approaches to case study that creatively incorporate multiple sources of both qualitative and quantitative data may expand the scope of study findings beyond those traditionally associated with case study research, suggesting the potential to

legitimize case study among researchers and public policy makers. Similarly, due to many scholars' predilection for quantitative research, mixed methods research has long been perceived as an approach that undervalues qualitative research and employs weak qualitative methods (Giddings, 2006). Case study approaches to mixed methods research that creatively apply in-depth qualitative case study methods along with robust quantitative methods and integration strategies can enrich the findings generated by mixed methods research and advance the contribution of both mixed methods and case study research in the field of education.

Trends toward increasing globalization create broad-scale background conditions, or meta-contexts, in which national, regional and local contexts are nested. Meta-contexts such as that of the COVID-19 pandemic may shift researchers' thinking about topics relevant to educational research, the data sources that can be used in mixed methods case studies, and how systems are bounded and cases defined. Questions regarding the role of education in areas such as climate change and social injustice, and questions about how school policies affect public health may be addressed by leveraging the strengths of combined case study and mixed methods approaches. These meta-contexts, along with the increasing application of systems thinking in educational research (e.g., Mette and Riegel, 2018; Verhoeff et al., 2018), may also lead to new considerations of how cases are bounded and may present opportunities for multiple case study researchers to conceive of new ways to understand and describe case interdependence.

Research findings over the past several decades have highlighted the fact that educational outcomes are influenced by a number of contextual features (e.g., Chapman et al., 2012), suggesting that educational practitioners and policy makers should be especially mindful of the criteria and techniques they use to select cases for examination. The use of rigorous mixed methods approaches in case studies may result in the application of sophisticated methods such as cluster analysis (Foote, 2019) that have implications for the selection of and interpretation of cases. In addition, as large and less traditional data sets emerge, mixed methods case study researchers may grapple with whether or how to incorporate, for example, social media and biosocial data, into their studies.

Concluding comments

Mixed methods case study provides an agile framework with which to gain a deep understanding of educational issues and problems within a variety of contexts. The answer to the question "what happened here?" can be used to inform educators' practices, policy makers' decisions, and other researchers' research agendas. We conclude with recommendations for those conducting mixed methods case studies and those seeking to use the findings of these studies for practical purposes.

The understanding of the ways that case study and mixed methods approaches can be combined and the ways that the resulting studies can be described continues to evolve. Those interested in mixed methods case studies should therefore approach the conduct of and interpretation of such studies not with a rigid set of criteria with which to assess findings, but rather with an understanding of the considerations unique to the combined approach. Investigators creating research reports are encouraged to articulate clearly their studies' purposes, their rationales for the bounding and selection of cases, their choices of data sources, the timing of the quantitative and qualitative methods and the ways that they are combined, and their purposes for combining methods. Furthermore, producers of research should carefully consider how integration strategies can enhance their case interpretations and carefully frame their case reports in ways that allow consumers of research to assess the applicability of the findings to their own practices.

We recommend that consumers of research – graduate students, educators, policy makers, and others – carefully examine reports of mixed methods case studies for evidence of researchers' decisions about the points mentioned. Situating this information within the context of their own goals for reading the report may provide important information regarding the applicability of studies' findings to readers' specific needs. The following questions are provided as a guide for interested readers of mixed methods case study reports:

- What is the phenomenon being investigated in the study?
- Why and how were the cases chosen and how do they relate to the study's goals?
- What is the unit of study and how does investigating this unit answer the research questions posed?
- What is the purpose and timing of collecting and analyzing multiple forms of data?
- How do the different forms of data relate to the study's goals?
- What new insight is provided by the integration of the qualitative and quantitative data and methods?
- How does the researcher suggest that the findings of this study be used, and does the study's design warrant this use?

The potential of mixed methods case studies to provide comprehensive understandings of complex situations can be leveraged by both producers of and consumers of research to advance knowledge in the field of education, inform public policy, gain deep understanding of the implications of contextual factors for educational issues, and identify promising educational practices. As we grapple with changing meta-contexts and new areas of inquiry in education, this combined approach may increasingly appeal to researchers who seek to provide in-depth understandings of these complex situations.

References

- Auslander, S.S., Smith, S.Z., Smith, M.E., Myers, K., 2020. A case study of elementary teacher candidates' preparation for a high stakes teacher performance assessment. *J. Math. Teach. Educ.* 23, 269–291.
- Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Res.* 6, 97–113.
- Byrne, D., 2009. Case-based methods: why we need them: what they are, how to do them. In: Byrne, D., Ragin, C.C. (Eds.), *The SAGE Handbook of Case-Based Methods*, first ed. Sage, Los Angeles, CA, pp. 1–13.
- Chapman, C., Armstrong, P., Harris, A., Muijs, D., Reynolds, D., Sammons, P., 2012. Conclusion. In: Chapman, C. (Ed.), *School Effectiveness and Improvement Research, Policy, and Practice: Challenging the Orthodoxy*, first ed. Routledge, New York, pp. 230–235.
- Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, fourth ed. Sage, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Duffin, M., Perry, E.E., 2019. Regional collaboration for sustainability via place-based ecology education: a mixed-methods case study of the Upper Valley Teaching Place Collaborative. *Educ. Sci.* 9, 1–13.
- Fetters, M.D., 2020. *The Mixed Methods Research Workbook*. Sage, Thousand Oaks, CA.
- Fetters, M.D., Freshwater, D., 2015. Publishing a methodological mixed methods research article. *J. Mix. Methods Res.* 9, 203–213.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs – principles and practices. *Health Serv. Res.* 48, 2134–2156.
- Foote, L.A., 2019. *Planning for Success: A Mixed Methods Comparative Case Study Investigating Elementary Mathematics Supports across School-Dependency Profiles*. Doctoral Dissertation, University of Cincinnati. <https://www.proquest.com/pqdtglobal/docview/2244455350/B745EAA577D243EFPQ/1?accountid=2909>.
- Giddings, L.S., 2006. Mixed-methods research: positivism dressed in drag? *J. Res. Nurs.* 11, 195–203.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed method evaluation designs. *Educ. Eval. Pol. Anal.* 11, 255–274.
- Guest, G., 2013. Describing mixed methods research: an alternative to typologies. *J. Mix. Methods Res.* 7, 141–151.
- Guetterman, T.C., Fetters, M.D., 2018. Two methodological approaches to the integration of mixed methods and case study designs: a systematic review. *Am. Behav. Sci.* 62, 900–918.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015. *Ann. Fam. Med.* 13, 554–561.
- Harvey, D.L., 2009. Complexity and case. In: Byrne, D., Ragin, C.C. (Eds.), *The SAGE Handbook of Case-Based Methods*. Sage, Los Angeles, pp. 15–38.
- Hesse-Biber, S., Johnson, R.B. (Eds.), 2015. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, first ed. Oxford University Press, New York.
- Heyvaert, M., Hannes, K., Maes, B., Onghena, P., 2013. Critical appraisal of mixed methods studies. *J. Mix. Methods Res.* 7, 302–327.
- Ishak, N.M., Bakar, A.Y., 2014. Developing sampling frame for case study: challenges and conditions. *World J. Educ.* 4, 29–35.
- Ivankova, N.V., Stick, S.L., 2007. Students' persistence in a distributed doctoral program in educational leadership in higher education: a mixed methods study. *Res. High. Educ.* 48, 93–135.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11, 156–173.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1, 112–133.
- Luck, L., Jackson, D., Usher, K., 2006. Case study: a bridge across the paradigms. *Nurs. Inq.* 13, 103–109.
- Maxwell, J.A., Mittapalli, K., 2010. Realism as a stance for mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 145–167.
- Merriam, S.B., 1998. *Qualitative Research and Case Study Applications in Education*, second ed. Jossey-Bass, San Francisco.
- Mertens, D.M., 2012. Transformative mixed methods: addressing inequities. *Am. Behav. Sci.* 56, 802–813.
- Mette, I.M., Riegel, L., 2018. Supervision, systems thinking, and the impact of American school reform efforts on instructional leadership. *J. Cases Educ. Leadersh.* 21, 34–51.
- Morgan, D.L., 2014. *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. Sage, Thousand Oaks, CA.
- Morse, J.M., 1991. Approaches to qualitative-quantitative methodological triangulation. *Nurs. Res.* 40, 120–123.
- Nzabonimpa, J.P., 2018. Quantitizing and qualitzing (im-)possibilities in mixed methods research. *Methodol. Innov.* 11, 1–16.
- Onwuegbuzie, A.J., Johnson, R.B., 2006. The validity issue in mixed research. *Res. Sch.* 13, 48–63.
- Plano Clark, V.L., Foote, L.A., Walton, J.B., 2018. Intersecting mixed methods and case study research: design possibilities and challenges. *Int. J. Mult. Res. Approaches* 10, 14–29.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Ragin, C.C., 1992. Introduction: case of 'what is a case'. In: Ragin, C.C., Becker, H.S. (Eds.), *What Is a Case? Exploring the Foundations of Social Inquiry*. Cambridge University Press, Cambridge, pp. 1–18.
- Schwandt, T.A., Gates, E.F., 2018. Case study methodology. In: Denzin, N.K., Lincoln, Y.S. (Eds.), *The SAGE Handbook of Qualitative Research*, fifth ed. Sage, Thousand Oaks, CA, pp. 341–358.
- Shannon-Baker, P., 2016. Making paradigms meaningful in mixed methods research. *J. Mix. Methods Res.* 10, 319–334.
- Schoonenboom, J., 2023. Ten mixed methods integration strategies for obtaining a detailed understanding. In: Tierney, R. (Ed.), *International Encyclopedia of Education*, fourth edition, Elsevier.
- Stake, R., Visse, M., 2023. Case Study Research. In: Tierney, R. (Ed.), *International Encyclopedia of Education*, fourth edition, Elsevier.
- Stake, R.E., 1995. *The Art of Case Study Research*. Sage, Thousand Oaks, CA.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Teddlie, C., Stringfield, S., Desselle, S., 1985. Methods, history, selected findings and recommendations from the Louisiana School Effectiveness Study, 1980-85. *J. Classr. Interact.* 20, 22–30.
- Tsang, E.W., 2014. Generalizing from research findings. *Int. J. Manag. Rev.* 16, 369–383.
- Verhoeff, R.P., Knippels, M.P., Gilissen, M.G., Boersma, K.T., 2018. The theoretical nature of systems thinking: perspectives on systems thinking in biology education. *Front. Educ.* 3, 1–11.
- Walton, J.B., 2014. *Partners for Change: A Mixed Methods Case Study of an Intermediary-Led Partnership for STEM Education Reform* (Doctoral dissertation). University of Cincinnati. https://etd.ohiolink.edu/etd.send_file?accession=ucin1406810984&disposition=inline.
- Walton, J.B., Plano Clark, V.L., Foote, L.A., Johnson, C., 2020. Navigating intersecting roads in a mixed methods case study: a dissertation journey. *J. Mix. Methods Res.* 14, 436–455.
- Wise, A.F., Perera, N., Hsiao, y., Speer, J., Marbouti, F., 2012. Microanalytic case studies of individual participation patterns in an asynchronous online discussion in an undergraduate blended course. *Internet High Educ.* 15, 108–117.
- Yazan, B., 2015. Three approaches to case study methods in education: Yin, Merriam, and Stake. *Qual. Rep.* 20, 134–152.
- Yin, R.K., 2018. *Case Study Research and Applications: Design and Methods*, sixth ed. Sage, Los Angeles, CA.

Intersecting mixed methods with phenomenology

Joanne Mayoh^a, Jason Colditz^b, and Nollaig Frost^c, ^a Bournemouth University Business School, Bournemouth University, Poole, United Kingdom; ^b University of Pittsburgh School of Medicine, Pittsburgh, PA, United States; and ^c University College Cork, School of Applied Psychology, Ireland

© 2023 Elsevier Ltd. All rights reserved.

Introduction	550
Mixed methods research (MMR)	550
Exploring intersections with phenomenological approaches	551
The phenomenological method	551
The MMPR approach conceptualized	551
Exploring philosophical compatibility with the phenomenological approach	552
Mixing descriptive (eidetic) phenomenology	553
Illustrative example:	553
Mixing interpretive (hermeneutic) phenomenology	554
Illustrative example:	555
Strategies for mixing within MMPR studies	555
Concluding thoughts	556
References	557

Introduction

Researchers are gradually acknowledging the potential for methodological innovation that can be achieved through combining multiple methods within single research studies by utilizing mixed methods research (MMR) approaches. In recognition of this increasing popularity, and to address concerns regarding the uncritical adoption of MMR within research designs (Sale et al., 2002) there is a need for in-depth philosophical and methodological discussion with regards to how multiple methods can be combined on both a paradigmatic and practical level within research studies, as well as practical examples of these approaches being adopted successfully. Increasingly, methodologists are providing guidelines on how qualitative research traditions with established histories such as grounded theory (Johnson et al., 2010), performative methods (Schoonenboom, 2019), narrative (Waller et al., 2020) and phenomenology (Mayoh and Onwuegbuzie, 2015) can be sympathetically combined with quantitative methods to provide such methodological innovation.

In line with this movement and the recognition of the ongoing discussion that the phenomenological method works particularly well as a component of MMR designs (Mayoh and Onwuegbuzie, 2015), the purpose of the following chapter is to provide both a philosophical and practical discussion regarding the intersection between phenomenology and methods grounded in post-positivism within a single study, and showcase two practical examples of mixed methods research that incorporated phenomenology in a philosophically sensitive manner. In doing so, we present an overview of the ongoing conceptualization of Mixed Methods Phenomenological Research (MMPR) and help provide a framework for the use of phenomenology as a component within MMR. To provide a more comprehensive discussion, we deliberate different forms of phenomenological methodologies, namely descriptive (eidetic) and interpretive (hermeneutic), and the associated philosophical and methodological assumptions and compatibility with alternative methods. This engagement with the philosophy of mixing methods has utility for researchers in reinforcing justifications for using MMPR and in supporting defense of their methodological decision making when engaging with methodological purists. Following the philosophical discussion, to provide frameworks for researchers looking to utilize a MMPR design, practical strategies for mixing phenomenology with methods grounded in post-positivism are presented alongside current empirical examples of each form of MMPR. The illustrative examples illuminate key features of both descriptive and interpretive MMPR designs, in addition to providing a discussion of the benefits and challenges associated with adopting such an approach. In doing so they highlight different approaches to MMR, MMPR and Phenomenology itself. The final section “Concluding Thoughts” overviews the chapter’s purpose and content and addresses critical questions regarding future development and application of MMPR. It also overviews how MMPR expands current methodological practice and contributes to the advancement of MMR more generally as it provides a philosophical justification and practical strategies for those wishing to utilize such a design approach.

Mixed methods research (MMR)

Increasingly Mixed methods research (MMR) is being adopted within research practice, critically conceptualized, and recognized as the third methodological paradigm alongside Qualitative and Quantitative approaches. Whilst scholars argue that there may not be

an essentialist definition forthcoming for the term Mixed Methods Research (Johnson et al., 2007), dialog surrounding a practical definition remains an ongoing goal for the field. For the purposes of the current chapter, Mixed methods may be defined as “research in which the investigator collects and analyses data, integrates the findings and draws inferences using both qualitative and quantitative approaches or methods in a single study” (Tashakkori and Creswell, 2007, p. 4).

Exploring intersections with phenomenological approaches

The phenomenological method

The term phenomenology is wide-ranging and has been attributed to a historical philosophical movement, a disciplinary field within philosophy, and a range of associated practical research approaches. While phenomenological philosophy has been practiced in many forms for centuries, the movement gained serious momentum in the early 20th century following the publication of pivotal works from phenomenologists including Husserl, Heidegger, Sartre, and Merleau-Ponty. Although their interpretations of phenomenological philosophy differ, there is consensus in their shared belief that human experience is the starting point for philosophy (Todres and Holloway, 2004). This central experiential philosophy informs the phenomenological method in its many guises. Consequentially the central aim of all phenomenological inquiry is to develop a better understanding of individual’s lived experiences through the consciousness of the experienter (Giorgi, 2009). The traditions are united in studying subject’s perspectives of their own worlds, to describe in detail the content and structure of their consciousness and grasp the qualitative diversity of their experiences and to elucidate essential meanings (Kvale, 1996). Phenomenology is therefore focused on both the structure, or *essence* (Husserl, 1965) of human experience, as well as honoring the diverse interpretive variations. The theoretical justification for such an approach is that through uniquely taking a first-person point of view, human beings can be understood from *inside* their subjective experiences (Todres and Holloway, 2004). This is principally achieved by empowering participants to provide in-depth descriptions of their own lived experiences, which can be used as a form of qualitative data and presented as descriptions and/or interpreted. This unique focus rationalizes phenomenology as the ideal method for experiential work within the interpretive paradigm and highlights its potential in representing marginalized voices and reducing power imbalances that exists within traditional research approaches often grounded in positivist philosophy.

The unique strengths and methodological flexibility (Garza, 2007) of the phenomenological method subsequently has led to it increasingly being adopted as a component of mixed methods research approaches. Mixed methods may be defined as “research in which the investigator collects and analyses data, integrates the findings and draws inferences using both qualitative and quantitative approaches or methods in a single study” (Tashakkori and Creswell, 2007:4). This methodological cohesion through combination not only maximizes the key strengths of phenomenology in terms of the exploration of specific lived experiences and empowering potential, but also helps minimize the weaknesses of the phenomenological approach. For instance, while phenomenology has key strengths in exploring individual’s lived experiences through the consciousness of the experienter the tradition provides no way of identifying which experiential phenomenon is most pertinent to an individual or group, therefore being able to orientate the research to the most critical experiences to be studied (Mayoh and Onwuegbuzie, 2015). In addition, critics of phenomenology have highlighted its failure to address the social context of lived experience (Langridge and Ahern, 2003). The following discussion will provide an overview of how multiple forms of phenomenology can be mixed in a philosophically sensitive manner with methods grounded in post-positivism to maximize the strengths and minimize the weaknesses of the phenomenological approach.

The MMPR approach conceptualized

An early systematic review to identify empirical research that incorporated a phenomenological component within MMR (Mayoh and Onwuegbuzie, 2014) demonstrated that 24 articles were published between 2002 and 6 months into 2012. This reflects the early stages of MMPR adoption before formal conceptualization, which were predominantly conducted within the field of health research (70.8%, $n = 17$), with only four additional disciplines being cited: psychology (12.5%, $n = 3$); education (8.3%, $n = 2$); social care (4.2%, $n = 1$); and career development (4.2%, $n = 1$). This earliest recorded MMPR study according to the review was 2003, where only one study was recorded (Langridge and Ahern, 2003). Similarly, the review showed that only one article was published in the following year (Kent, 2004), however by 2011, 7 articles were published, demonstrating a growth in popularity of combining phenomenology as a component of MMR. The increasing practical application of phenomenological methods within MMR approaches led Mayoh and Onwuegbuzie (2015) to deliver a formal conceptualization of MMR that incorporates phenomenology—which they termed “MMPR”; p. 3). In this conceptualization paper they defined MMPR as “research that combines phenomenological methods with methods grounded in an alternative paradigm within a single study” (p. 21). The authors aimed to provide a philosophical justification for using MMPR alongside examples of MMPR in practice to underline several potential models that can practically be utilized in future research. The increasing practical adoption and theoretical discussion and conceptualization of MMPR shows that despite apprehension from advocates of the incompatibility thesis (see, Howe, 1988), and challenges associated with incorporating phenomenology into mixed methods research designs, phenomenological methods provide methodological compatibility with research methods grounded in alternative paradigms such as post-positivism. Since the formal conceptualization, the number of studies citing MMPR has grown with research being conducted within disciplines including child

abuse and neglect (McDonald et al., 2015), counseling psychology (Cobb et al., 2017) and the gamification of education (Kenny et al., 2017).

While the phenomenological method takes a multitude of forms, this chapter will focus specifically on how descriptive and interpretive phenomenology can be interfaced with methods grounded in an alternative paradigm as MMPR. This reflects how MMPR is frequently adopted within academic work, as an early systematic review of MMPR research demonstrated that research combining phenomenological methods with alternative approaches in a single study were predominantly conducted using descriptive (54.2%, $n = 13$) or interpretive (41.7%, $n = 10$) phenomenological approaches (Mayoh and Onwuegbuzie, 2014).

Exploring philosophical compatibility with the phenomenological approach

While historically critics of Mixed Methods Research (MMR) (see Howe, 1998) have adhered to the argument that the philosophical distinctions between qualitative and quantitative methods and accompanying paradigms are too significant to justify combination within a single study, mixed methods researchers have consistently negotiated such differences, making adjustments including using multiple methods within a single paradigm (Johnson et al., 2010) and exploring paradigmatic complementarity (Mayoh and Onwuegbuzie, 2015). These intersections are justified by the shared assumption within mixed methods research that practical research methods are not consistently tied to specific paradigms (Johnson and Onwuegbuzie, 2004). The growing practical adoption (Mayoh and Onwuegbuzie, 2014) and conceptualization of Mixed Methods Phenomenological Research (MMPR) demonstrates that despite the numerous challenges associated with using MMPR, particularly on a philosophical level, multiple phenomenological methods including descriptive (Eidetic), and interpretive (Hermeneutic) phenomenological traditions provide methodological compatibility with methods grounded in alternative paradigms. Table 1 below reviews postpositivist, and phenomenological (descriptive and interpretive) paradigms, and the philosophical assumptions associated with them.

Table 1 A comparison between the philosophical assumptions of postpositive and phenomenological inquiry.

	<i>Postpositivism</i>	<i>Phenomenology</i>	
		<i>Descriptive (eidetic)</i>	<i>Interpretive (hermeneutic)</i>
Ontology (nature of reality)	An objective reality or truth exists, however can never be interpreted fully.	There are multiple constructions of reality. Objectivity relates to the extent to which description is true to a phenomenon. Although experience is subjective, there are features to any lived experience that are common to all persons who have the experience.	There are multiple constructions of reality. Reality is constructed in unique ways depending context and personal frames of reference as individuals engage with the world.
Epistemology (knowledge)	Data are an approximation of objective reality. The ideal of objectivity can only be approximated through taking a critical approach.	Data are based on a subjective reality. The goal of the researcher is to achieve transcendental subjectivity.	Data are based on a subjective reality. These data are produced through the interaction between the participant and researcher, and the researcher's interpretations. There is a need for an orienting framework. People cannot abstract themselves from the world.
Axiology (role of values)	Researcher attempts to remain value free (values must be excluded).	Researcher acknowledges values and bias. The impact of the researcher on the inquiry is constantly assessed so that they do not influence the object of study.	The researcher's values impact the object of study. Expert knowledge on the part of the researcher is valuable.
Methodology (research strategies)	Multiple approaches that may be triangulated.	Methods that focus on obtaining participants descriptions of experience. Variations include Giorgi's (1985, 2009) descriptive phenomenological method in psychology.	Interpretive methods that move beyond the description of core concepts and essences to look for meanings. Variations include interpretative phenomenological analysis (IPA; Smith and Osborn, 2003).

Adapted from Mayoh and Onwuegbuzie (2015, p. 94).

Mixing descriptive (eidetic) phenomenology

Despite the clear paradigmatic tensions between post-positivist and phenomenological inquiry shown in [Table 1](#), the compatibility between methods grounded in these traditions appears more feasible when one considers that the founder of the descriptive (eidetic) method in phenomenology Edmund Husserl (1859–1938) was a mathematician as well as a philosopher. [Husserl \(1970\)](#) viewed phenomenology as a science of essence of consciousness and believed scientific inquiry could provide rich authentic descriptions of phenomena, advancing our understanding of human consciousness and experience ([Husserl, 1976](#); [Dowling, 2007](#)). While he was largely concerned with moving away from traditional forms of positivist inquiry ([Finlay, 2011](#)), Husserl sought to explain how researchers could overcome personal biases, that may stand in the way of achieving the state of pure consciousness by adopting the transcendental-phenomenological reduction, a condition of consciousness in which the researcher forsakes their own lived reality and are able to describe the phenomenon in its pure and universal sense. The very nature of the phenomenological reduction, and the descriptive phenomenologists' aim of recognizing and reducing the role of the researcher through a process called bracketing (i.e., *epoché*; [Husserl, 1931](#)) shows alignment to the post-positivist goal of researcher objectivity, rather than alternative qualitative and phenomenological methods that view qualitative data as co-constructed and inherently biased. Furthermore, [Table 1](#) demonstrates that while descriptive phenomenology acknowledges that human experience is subjective, there may be features to lived experience that are common. According to [Giorgi \(2009\)](#), one of the key characteristics of descriptive phenomenology is the acknowledgment that the essence or bare bones of what constitutes the phenomenon is articulated as the structure ([Holloway and Todres, 2003](#)). It describes the common themes or essential parts from within the experience that identify the phenomenon and can transcend the experiences of different individuals. In honoring that the structure of lived experiences can contain concrete details in addition to ceaseless diversity, the role of object is valued within descriptive phenomenological methods, aligning with the ontological assumptions of post-positivism. Subsequently, MMPR that incorporates a descriptive phenomenological component is increasingly being cited as the methodological approach within research studies.

Illustrative example: Colditz et al. (2019)

[Colditz et al. \(2019\)](#) present MMPR research that integrates a descriptive phenomenological element in observational research exploring narratives about Electronic Nicotine Delivery Systems (ENDS; i.e., vaping devices), using data from the Twitter social media platform on “World Vaping Day”. Given that this was a relatively discrete context, which generated a large amount of qualitatively dense data (5149 “tweet” messages), it was both feasible and beneficial to approach analysis using multiple complementary methods. To assess the quantitative breadth of data, studies on the Twitter platform often assess the frequency of informative phrases, keywords, or hashtags (e.g., #vapelife) to efficiently identify and quantify content of a priori research interest ([Colditz et al., 2018](#)). With respect to qualitative context of this work, researchers understood vaping as a contentious public health topic on the Twitter platform (e.g., politicization, health misinformation). As health scientists, the researchers decided it would be important to bracket out preconceived notions about the underlying health implications in order to focus on essential experiences of the World Vaping Day phenomenon. Bracketing is germane to descriptive phenomenology in health contexts and so this was chosen as a favorable approach to enhance rigor of the qualitative synthesis ([Lopez and Willis, 2004](#)). While the phenomenological and quantitative procedures were conducted independently, the data collection phases were informed by the concurrent analytical approach ([Creswell and Plano-Clark, 2017](#)), where the methods are given equal priority, analyzed independently, but interpreted together.

The authors developed a broad codebook containing themes that were relevant to ENDS propagation, focusing on positive and negative sentiment surrounding particular categorical themes. By categorically coding the data within this post-positivist framework, prevalence of identified themes could be quantified, and inter-rater agreement statistics calculated to demonstrate reliability of coding. In order to not limit the experiential data, and to honor the paradigmatic assumptions of the phenomenological element, observations relating to themes that didn't fit into the conceptual framework were collected to integrate parallel data analysis and synthesis, allowing for both structured (coded utterances in quantifiable categories) and unstructured data (notes, observations and examples). The unstructured data were then synthesized within a descriptive phenomenological research framework. In combining thematic prevalence estimates with descriptive methods for phenomenological contextualization, the authors were able to achieve both quantitative breadth and rich descriptive qualitative depth within their study. This also contributed to the methodological research narrative by outlining a strategy for utilizing descriptive phenomenology within a mixed methods study of primary social media data. In doing so, the authors were able to address a criticism of phenomenological work, that it can be alienating due to the dominance of phenomenological research that is prohibitively dense with philosophical theory to those less familiar with the concept ([Caelli, 2001](#)).

The descriptive MMPR approach made a unique contribution to the research narrative exploring the intersection between social media use and phenomenology that generally relies on interview data and fails to recognize the potential experiential depth of naturalistic primary online data ([Colditz et al., 2019](#)). While phenomenological purists may reject the idea that meaningful experiential data can be obtained through social media observation, the researchers in this study maintained a narrow scope of observation in order to explore a cohesive landscape of experiences within a specifically defined social context. In contemporary phenomenological approaches, such data might be obtained from interviews and thus the scope of data would be limited by the interview boundaries. Then horizons (i.e., units of meaning contextualized along the phenomenal landscape) would emerge from quotes and impressions

identified from interview transcripts. In the context of Colditz et al. (2019), the scope was limited by searching for any tweets containing ENDS-related keywords on the third annual “World Vaping Day”. As tweets are relatively short (280 text characters), Twitter users convey concise units of meaning within brief utterances and associated multimedia content (e.g., emoji, links, images and video). This provided observational researchers with a topically cohesive yet diverse landscape of horizons to synthesize. While results indicated that direct references to World Vaping Day were relatively rare in the data, it was clear that this online event was experienced differently among various interest groups (e.g., ENDS manufacturers, former tobacco smokers, health policy advocates). Overall, the sentiment of tweets mentioning ENDS was positive or enthusiastic (>3:1 ratio vs. negative sentiment), with the exception that policy-related tweets tended to be more negative (e.g., discontent over ENDS regulation). Phenomenological inquiry allowed for exploration of nuanced social patterns and themes in the data (e.g., #vapecommunity hashtag was primarily used by a niche of individuals promoting ENDS over smoking, #vapelite and #vapelyfe were more commonly used to directly promote ENDS products). Additional perspectives on the vaping phenomena included a sense of authenticity or “coolness” around individuals who used ENDS to quit smoking and those who were knowledgeable about complex aspects of ENDS devices (e.g., modifying components, “vape juice” flavors). This enthusiasm was contrasted by jokes ridiculing younger ENDS users, for using ENDS to appear cool rather than to quit smoking. The researchers also distinctly identified ENDS users’ animosity toward the “big tobacco” companies that entered the ENDS market then championed stricter US government regulations (running counter to the interests of ENDS users and smaller commercial manufacturers). Through this MMPR synthesis, the researchers came to understand World Vaping Day as a window into a more complex vaping culture, within a broader online landscape dominated by commercial interests, social comparisons, and—to a lesser extent—health-related discourse that researchers might otherwise anticipate and focus intently on.

In line with previous argument made in this chapter, Colditz & colleagues acknowledged the compatibility between phenomenology and mixed methods approaches. They recognized the descriptive method’s powerful ability to control for researcher bias through acknowledging and reducing the role of the researcher (bracketing), requiring researchers to recognize and monitor the relationship between their subjective experience (e.g., what topics might be of health importance) and objective reality (i.e., what was the essential experience of World Vaping Day). In doing so they state that this approach allowed the researcher to identify and reduce personal bias (Arslan and Yildirim, 2015) while recognizing both an ontological imperative to acknowledge the objective existence of a phenomenon with the epistemological assumption that individuals interpret phenomena subjectively. This was achieved through a process of synthesizing notes and observations into organized narratives that could be utilized to pursue didactic inquiry among coders and help uncover the structural aspects of vaping culture expressed in this context. The synthesis of qualitative data was made possible through the development of a structured codebook and coding process, complemented with annotations, quotes and narratives, and critical reflections through a phenomenological lens. These provided rich context that could contribute to and move beyond the conceptual framework provided by the coding categories. In converging the phenomenological synthesis into the categorical results narrative, the authors were able to move beyond the reporting of reductive statistics and key words and integrate deeper experiential themes that helped frame a more essential understanding of vaping culture in this context.

Observational social media research remains a relatively atypical application for MMPR. This particular example was chosen to demonstrate how descriptive phenomenology can be successfully incorporated into highly novel MMPR approaches. Thus, MMPR holds some promise in well-scoped observational research contexts, particularly for understanding social media phenomena where individual experiences align in complex ways. This may be considered a beneficial extension of the descriptive phenomenological framework—rather than a separate social media research framework—insomuch as researchers strive to maintain fidelity of the philosophical considerations and commonly accepted procedures for descriptive MMPR. These considerations of course remain important for contemporary research approaches that might rely on interview or self-report data.

Mixing interpretive (hermeneutic) phenomenology

While Husserl is oft credited as the founding father of descriptive phenomenology, it is his student Martin Heidegger (1889–1976) that is widely regarded as the innovator behind the interpretive phenomenological method. Heidegger (1962) contested that “the meaning of phenomenological description as a method lies in interpretation” (p. 61), viewing hermeneutic description and interpretation as a way of analyzing being. Unlike descriptive (eidetic) phenomenology, interpretive (hermeneutic) approaches put significantly less weight on the phenomenological attitude and bracketing to minimize the impact of researcher bias. Alternatively, they emphasize the co-constructed interpretations of the individual’s life world as well as the individual’s natural orientation to the phenomena (Van Manen, 1990). In this sense, interpretive phenomenology informed by Heidegger moves away from Husserl’s phenomenology as a science of essence of consciousness, rejecting the lingering notes of positivism, and moves into a less cognitive existential framework. Consequently, mixing interpretive phenomenological approaches with methods grounded in the postpositivist paradigm poses more significant challenges due to the competing axiological assumptions with post-positivism at a paradigmatic level (see Table 1). Despite this, scholars have suggested philosophically sensitive models for integrating interpretive phenomenological approaches into MMPR (Mayoh and Onwuegbuzie, 2015). In describing the interpretive method, Van Manen (1990) defines the process of orienting to the phenomenon within interpretive phenomenology, a stage which is conducted prior to focusing in on a specific phenomenological question involving focusing on identifying what lived experience would be the most relevant focus for phenomenological inquiry. Based on this practical feature of

interpretive methods, [Mayoh and Onwuegbuzie \(2015\)](#) suggest there is a strong rationale for the inclusion of a quantitative preliminary phase to orient toward the most valuable lived experience for phenomenological inquiry. In addition, we conclude there is a strong rationale for sequential, interpretive MMPR that utilizes an initial quantitative phase to give additional context, structure and focus to the secondary interpretive phase.

Illustrative example: Zamariola et al. (2019)

In their 2019 study, Zamariola et al. used a sequential MMPR approach adopting a preliminary quantitative phase for orientation, and a subsequent qualitative phase for context, to enrich an exploration into the relationship between interoception (the ability to perceive one's inner bodily feelings) and emotion regulation. The study contributed to existing literature in the field, which focused mainly on questionnaires or mood induction techniques in isolation and provided inconsistent results and low reliability of measures. The methodological approach encouraged rich, focused experiential descriptions from the participants' perspectives to better understand the lived experience of the relationship. Phase one of the study included a heartbeat counting task, and the administration of quantitative questionnaires designed to assess interoceptive sensibility, "the self-perceived sensitivity to internal bodily sensations" ([Pearson and Pfeifer, 2020](#), p. 3). This allowed the research team to develop subsamples of participants scoring low, average, and high interoceptive sensibility, so that the subsequent phenomenological phase could explore their perceived bodily awareness and uncover how this influenced emotional awareness. Participant quotes from semi-structured interviews provided evidence for the phenomenological themes and enabled enriched insight to the relationship and its lived experience.

The qualitative phase of this study utilized Interpretative Phenomenological Analysis (IPA) ([Smith, 1996](#)). This method enables exploration of how participants make sense of experience, closer to Husserl's focus on the contents of consciousness than on Heidegger's focus on description of lifeworld experiences. IPA recognizes that lived experience cannot be accessed directly and uses detailed analysis to follow a chain of connection in participant accounts from embodiment to language to emotion to cognition ([Frost, 2021](#)). By aiming to access mental phenomena, IPA enables greater understanding of how, for example, strategies to manage negative emotions are managed, making it an appropriate method for this study. It is tied to a double hermeneutic which recognizes that researchers seek to make sense of how participants make sense of their experiences. It is idiographic—focusing first on in-depth analysis of each participant account before moving to examination of similarities and differences across the accounts. Detailed insight to patterns of meanings and reflections on shared experiences are produced in the form of themes that give insight to strategies, abilities and complexities of experience are enacted and managed.

Three themes were discovered: Interoceptive Abilities, Emotional Abilities, and Influences on Interoceptive and Emotional Abilities. The Interoceptive Abilities theme described how all participants found the heartbeat counting task difficult to perform. Perception of heartbeats was challenging for all and those participants with low interoception had the greatest difficulty in perceiving cardiac signals. This supports existing studies that suggest that performance in this task relies largely on non-interoceptive processes and on general beliefs about heartbeat rates ([Desmedt et al., 2018](#); [Ring et al., 2015](#); [Zamariola et al., 2018](#)). The Emotional Abilities theme showed that some participants found negative emotions harder to identify, while others were able to recognize these emotions but had difficulty in regulating them. The use of the subjective phenomenological method enabled the complexity of recognizing and verbalizing emotions to be recognized. The third theme identified a range of strategies used by participants to manage negative emotions. These included talking to oneself, using food as comfort, and engaging in physical exercise, but also that the strategies provided varying degrees of relief. The study also found that participants with high interoception described seeking a balance between mind and body and reported being more self-focused and motivated to use techniques such as psychotherapy, mindfulness, and yoga to understand more about their inner life. Better emotion regulation was suggested as being more achievable by individuals with a positive and active attitude to managing negative emotions.

When considered alongside the quantitative phase of the study, the in-depth analysis of the interviews not only provided insight to the relationship between interoception and emotion regulation but also challenged the reliability of the heartbeat counting task. The study calls for further qualitatively orientated research to develop new techniques to assess interoceptive accuracy, and importantly, shows that interoception and emotion regulation are subjectively related.

Strategies for mixing within MMPR studies

Based on the theoretical discussion and illustrative examples, we support the notion that there is a strong justification for the inclusion of multiple forms of phenomenology within mixed methods studies. Specifically, the philosophical assumptions outlined in [Table 1](#) support our assertion that descriptive phenomenology has compatibility when utilized concurrently with methods rooted in the post-positivist paradigm, while there is a stronger justification for the combination of interpretive phenomenology sequentially within MMPR studies. In addition to this suggestion, the inductive nature of both forms of phenomenological inquiry encourage reasoning for the integration of an explanatory quantitative proceeding phase within sequential mixed method designs to test phenomenologically generated concepts ([Mayoh and Onwuegbuzie, 2015](#)). That being said, to increase the utility and scope

Table 2 Models for mixed methods phenomenological research.

<i>Model</i>	<i>Description</i>	<i>Purposes</i>
phen → QUAN or phen → QUAL	Preliminary phenomenological method with priority given to the method from an alternative paradigm.	The preliminary phenomenological stage is used to generate theory about the nature of lived experience that can feed into a larger QUAN QUAL study. This approach is especially valuable when the study requires an overall more deductive thrust.
quan → PHEN or qual → PHEN	Preliminary method from an alternative paradigm with priority given to the phenomenology	The preliminary quan/qual phase is used to orientate the predominant PHEN stage to the most relevant and interesting phenomenon. This model is valuable when the study requires an overall more inductive thrust.
QUAN → phen or QUAL → phen	Preliminary method from an alternative paradigm which is also given priority.	The secondary phen stage is used to explore unanticipated QUAN or QUAL findings. This model is especially valuable when the study requires an overall more deductive thrust.
PHEN → quan or PHEN → qual	Preliminary phenomenology which is also given priority.	The secondary quan/qual phase is used to help improve the utility and generalizability of phenomenological findings. This approach is valuable when the study requires an overall more inductive thrust.
quan + phen or qual + phen	Concurrent approach. Normally methods have equal priority (however, it is possible to give priority to a single method depending on the overall thrust of the study).	Within this model, phenomenological and complementary data are collected concurrently to cross-validate or to confirm findings.

Adapted from [Mayoh and Onwuegbuzie \(2015: 104\)](#).

of MMR as a methodological approach, future theoretical and practical work should not be limited to discussing and utilizing explicitly sequential or concurrent approaches based on the phenomenological approach adopted, as mixed methods researchers strive to develop more innovative and philosophically sound models for mixing. In contributing to this discussion, [Mayoh and Onwuegbuzie \(2015\)](#) propose several practical models for using MMR as a research design (see [Table 2](#)). In these models → or + is used to represent the sequence decision (→ for moving sequentially and + for working concurrently) and capital and lower-case letters are used to characterize the priority decision (QUAN or QUAL for principal method, quan or qual for complementary method) in line with previous MMR literature ([Morse, 1991](#); [Morgan, 1998](#)).

Concluding thoughts

The purpose of this chapter was to provide both a philosophical and practical discussion regarding the intersection between phenomenology and methods grounded in post-positivism within a single study, and to showcase two practical examples, one using descriptive (eidetic) and one interpretive (hermeneutic) phenomenology within mixed methods research. In doing so this chapter makes a unique contribution to the field of educational research and practice in providing a communicative representation of how phenomenology can be incorporated as a part of MMR studies. Furthermore, the humanizing potential of MMR provides great utility within educational research practice that oft deals with both human lived experiences and structure of learning organizations.

Specifically, in this chapter we focused on the integration of two forms of phenomenology that [Mayoh and Onwuegbuzie \(2014\)](#) have previously identified as the most often utilized within this form of research; descriptive (54.2%, $n = 13$) and interpretive (41.7%, $n = 10$) phenomenological approaches ([Mayoh and Onwuegbuzie, 2014](#)). However, it is important to note that alternative phenomenological methodological approaches, for example the dialogical phenomenological approach ([Stawarska, 2009](#)), have compatibility with post-positivism on a paradigmatic level, as axiologically they focus on empowering participants to take a leading role within the research process and adopt a research-facilitator role, thus minimizing the role of the researcher, and therefore, as with descriptive methods, helps eliminate individualist bias that honors subjectivity in some phenomenological approaches. Future theoretical work on MMR should focus on a more road conceptualization of phenomenological methods, and what alternative phenomenological approaches offer as components of mixed methods studies.

While the illustrative examples demonstrate great utility in integrating phenomenology within MMR studies, we advise all those considering MMR pay serious consideration to the cohesion between the multiple research methods, as well as thinking about the implications and potential associated with the priority and sequencing decisions, and how this relates to the unique philosophy of each approach within each methodological element.

References

- Arslan, P., Yildirim, S., 2015. Theoretical frameworks, methods, and procedures for conducting phenomenological studies in educational settings. *Turkish Online J. Qual. Inq.* 6 (1), 1–20.
- Caelli, K., 2001. Engaging with phenomenology: is it more of a challenge than it needs to be? *Qual. Health Res.* 11 (2), 273–281.
- Cobb, C.L., Meca, A., Xie, D., Schwartz, S.J., Moise, R.K., 2017. Perceptions of legal status: associations with psychosocial experiences among undocumented Latino/a immigrants. *J. Counsel. Psychol.* 64 (2), 167.
- Colditz, J.B., Chu, K., Emery, S.L., Larkin, C.R., James, A.E., Welling, J., Primack, B.A., 2018. Toward real-time infoveillance of Twitter health messages. *Am. J. Publ. Health* 108 (8), 1009–1014.
- Colditz, J.B., Welling, J., Smith, N., James, A.E., Primack, B.A., 2019. World Vaping Day: contextualizing vaping culture in online social media using a mixed methods approach. *J. Mix. Methods Res.* 13 (2), 196–215.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- Desmedt, O., Luminet, O., Corneille, O., 2018. The heartbeat counting task largely involves non-interceptive processes: evidence from both the original and an adapted counting task. *Biol. Psychol.* 138, 185–188.
- Dowling, M., 2007. From Husserl to van Manen: a review of different phenomenological approaches. *Int. J. Nurs. Stud.* 44 (1), 131–142.
- Finlay, L., 2011. *Phenomenology for Therapists: Researching the Lived World*. Wiley-Blackwell, Oxford.
- Frost, N., 2021. *Qualitative Research Methods in Psychology: Combining Core Approaches 2e*. McGraw-Hill Education, United Kingdom.
- Garza, G., 2007. Varieties of phenomenological research at the University of Dallas. *Qual. Res. Psychol.* 4, 313–342.
- Giorgi, A., 1985. *Sketch of a psychological phenomenological method*. Duquesne University Press, Pittsburgh, PA.
- Giorgi, A., 2009. *The Descriptive Phenomenological Method in Psychology: A Modified Husserlian Approach*. Duquesne University Press, Pittsburgh.
- Heidegger, M., 1962. *Being and Time*. Blackwell, Oxford, England.
- Holloway, I., Todres, L., 2003. The status of method: flexibility, consistency and coherence. *Qual. Res.* 3 (3), 345–357.
- Howe, K.R., 1988. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educ. Res.* 17 (8), 10–16.
- Howe, K.R., 1998. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educ. Res.* 17 (8), 10–16.
- Husserl, E., 1931. *Ideas (W.R. Boyce Gibson, Trans)*. George Allen & Unwin, London, England.
- Husserl, E., 1965. *Phenomenology and the crisis of philosophy*. Harper and Row, New York.
- Husserl, E., 1970. *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*. Northwestern University Press, Illinois.
- Husserl, E., 1976. *Logical Investigations (J. N. Findlay, Translation)*. Routledge, NY.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26.
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1 (2), 112–133.
- Johnson, R.B., McGowan, M.W., Turner, L.A., 2010. Grounded theory in practice: is it inherently a mixed method? *Res. Sch.* 17 (2), 65–78.
- Kenny, R.F., Gunter, G.A., Campbell, L.O., 2017. Teachers' assessment of the instructional efficacy of mobile apps: a formative case study. *J. Formative Des. Learn.* 1 (1), 56–63.
- Kent, B.C., 2004. Protection behaviour: a phenomenon affecting organ and tissue donation in the 21st century? *Int. J. Nurs. Stud.* 41, 273–284.
- Kvale, S., 1996. *InterViews: An Introduction to Qualitative Research Interviewing*. Sage, London.
- Langridge, M.E., Ahern, K., 2003. A case report on using mixed methods in qualitative research. *Collegian* 10 (4), 32–36.
- Lopez, K.A., Willis, D.G., 2004. Descriptive versus interpretive phenomenology: their contributions to nursing knowledge. *Qual. Health Res.* 14 (5), 726–735.
- Mayoh, J., Onwuegbuzie, A.J., 2014. Surveying the landscape of mixed methods phenomenological research. *Int. J. Mult. Res. Approaches* 8 (1), 2–14.
- Mayoh, J., Onwuegbuzie, A.J., 2015. Toward a conceptualization of mixed methods phenomenological research. *J. Mix. Methods Res.* 9 (1), 91–107.
- McDonald, S.E., Collins, E.A., Nicotera, N., Hageman, T.O., Ascione, F.R., Williams, J.H., Graham-Bermann, S.A., 2015. Children's experiences of companion animal maltreatment in households characterized by intimate partner violence. *Child Abuse Negl.* 50, 116–127.
- Morgan, D.L., 1998. Practical strategies for combining qualitative and quantitative methods: applications to health research. *Qual. Health Res.* 8 (3), 362–376.
- Morse, J.M., 1991. Approaches to qualitative-quantitative triangulation. *Nurs. Res.* 40, 120–123.
- Pearson, A., Pfeifer, G., 2020. Two measures of interoceptive sensibility and the relationship with introversion and neuroticism in an adult population. *Psychol. Rep.* 125 (1), 0033294120965461.
- Ring, C., Brener, J., Knapp, K., Mailloux, J., 2015. Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: a cautionary tale about interoceptive awareness. *Biol. Psychol.* 104, 193–198.
- Sale, J.E., Lohfeld, L.H., Brazil, K., 2002. Revisiting the quantitative-qualitative debate: implications for mixed-methods research. *Qual. Quantity* 36 (1), 43–53.
- Schoonenboom, J., 2019. A performative paradigm for mixed methods research. *J. Mix. Methods Res.* 13 (3), 284–300.
- Smith, J.A., 1996. Beyond the divide between cognition and discourse: using interpretative phenomenological analysis in health psychology. *Psychol. Health* 11 (2), 261–271.
- Smith, J.A., Osborn, M., 2003. Interpretative phenomenological analysis. In: Smith, J.A. (Ed.), *Qualitative psychology: a practical guide to research methods*. Sage, London, pp. 51–80.
- Stawarska, B., 2009. *Between You and I: dialogical phenomenology*. Ohio University Press, Athens: OH.
- Tashakkori, A., Creswell, J.W., 2007. Editorial: the new era of mixed methods. *J. Mix. Methods Res.* 1 (1), 3–7.
- Todres, L., Holloway, I., 2004. Descriptive phenomenology: life-world as evidence. In: Rapport, F. (Ed.), *New Qualitative Methodologies in Health and Social Care Research*. Routledge, Swansea, pp. 99–118.
- Van Manen, M., 1990. *Researching lived experience: human science for an action sensitive pedagogy*. State University of New York Press, Albany, NY.
- Waller, M.R., Nepomnyaschy, L., Miller, D.P., Mingo, M., 2020. Using a narrative approach to analyze longitudinal mixed methods data. *J. Mix. Methods Res.* 15 (2), 1558689820953237.
- Zamariola, G., Maurage, P., Luminet, O., Corneille, O., 2018. Interoceptive accuracy scores from the heartbeat counting task are problematic: evidence from simple bivariate correlations. *Biol. Psychol.* 104137, 12–17.
- Zamariola, G., Frost, N., Van Oost, A., Corneille, O., Luminet, O., 2019. Relationship between interoception and emotion regulation: new evidence from mixed methods. *J. Affect. Disord.* 246, 480–485.

Mixed methods and grounded theory

Elizabeth G. Creamer, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Purpose	558
Methodological compatibilities between mixed methods and grounded theory	559
The iterative exchange embedded in the constant comparative method	559
The analytical logic of abduction	559
Adaptability to complexity	560
Compatibility with multiple paradigms	560
Distinct contributions of each method	560
Two approaches to implementing a MM-GT study	561
Linking distinct qualitative and quantitative phases to develop and verify or elaborate a grounded theory	561
Interactive designs with integration through embedding	562
Gauging the quality of MM-GTM	563
Saturation	563
Methodological transparency	563
Utility and originality	563
Discussion	564
Design considerations	565
Implications for practice	565
Directions for future research	565
References	565

Grounded theory is one of several mainstream qualitative methodologies that have demonstrated adaptability to diverse approaches, including mixed methods. Leading voices in the development of grounded theory, including its founders, have long singled receptivity to using both qualitative and quantitative data with grounded theory. For example, in her framing of a constructivist approach to grounded theory, Charmaz extended the idea of its adaptability to mixed methods when she observed "an emerging grounded theory can indicate needing more than one type of data and can incorporate more than one type of analysis" (2014, p. 323). Similarly extending the idea to encompass not just data but analytical procedures, Walsh wrote: "A [grounded theory] may thus be generated using qualitative and/or quantitative data, methods, and techniques" (2015, p. 536). No one has accepted the challenge to flesh out a methodological rationale for partnering grounded theory with mixed methods.

At its most basic level, mixed methods research can be defined as the combination of qualitative and quantitative data or analytical procedures. It is distinguished from multi-method research in its conviction of the contribution of integrating data collected through different methods in purposeful ways. Multi-method is the preferred term for any combination within the same tradition, including data collected through multiple qualitative or multiple quantitative approaches.

Although often mistakenly applied as simply an approach to coding qualitative data, grounded theory is a methodology with its own distinct analytical logic. The methodological foundation of grounded theory encompasses a set of procedures or tools. In my teaching, I have found it useful to refer to studies that limit their use of grounded theory to a set of coding procedures (open, focused, axial) as "little GT." "Big GT" recognizes that the ultimate purpose of those tools is to build theoretical insight.

Mixed method grounded theory (MM-GTM) is a methodology that does not just add quantitative data to what is otherwise a qualitative approach. MM-GTM is a radical new way to conceive of grounded theory that embeds the dialectical logic of mixed methods in the constant comparative method and grounded theory procedures. Its aim is not prediction but to develop a theoretical explanation that conceptualizes the interrelationships between concepts and provides an explanation for why and how these occur.

Purpose

Methodologically oriented literature is just beginning to emerge about MM-GTM (i.e., Creamer, 2018b; Guetterman et al., 2017; Johnson et al., 2010; Johnson and Walsh, 2019; Walsh, 2014, 2015). These have concentrated on its usefulness (i.e., Walsh, 2015), ways to evaluate its quality (i.e., Guetterman et al., 2017), the variability in the way studies with this design have been implemented (i.e., Johnson and Walsh, 2019), and introduced the potential for a mixed methods approach to be embedded in the core set of procedures (i.e., Creamer, 2018b). With no lack of examples of MM-GT now in print, the time is right for the forward thinking required to advance a methodological rationale for partnering grounded theory and mixed methods.

The goal of this encyclopedia entry is to offer guidance for those striving to incorporate grounded theory in a mixed method project in ways that honor the integrity of each. The purpose of the entry is to present an argument for the methodological

compatibility between mixed methods and most approaches to grounded theory. It is also to demonstrate different ways it can be accomplished and identifies different reporting standards that are relevant to MM-GTM. It articulates the view of MM-GTM as a distinct methodology that includes an analytical logic and set of core analytical procedures that produce a synergistic outcome that is greater than the sum of its distinct parts.

Methodological compatibilities between mixed methods and grounded theory

“Methodological dynamism” is a term that has been used broadly to describe to how methods evolve and remain viable over time (Ralph et al., 2015). Both mixed methods and grounded theory have been characterized as methodologies that are adaptable to diverse circumstances and methodological priorities. Innovative applications of each continue to emerge. Highlighting the ways researchers adapt grounded theory to diverse circumstances, Seidel and Urquhart (2013, p. 237) observed: “Grounded theory method (GTM) is an evolving method that is subject to idiosyncratic interpretation and flexible deployment.” The way an investigator deploys a methodology is an expression of his or her ontological and epistemological perspective and the wider social and intellectual strains of thought that influence a time period (Ralph et al., 2015).

There are multiple areas of the overlap between the methodological assumptions of grounded theory and mixed methods that provide a rationale for their pairing. Areas of methodological compatibility extend to (a) the iterative nature of the constant comparative method, (b) abduction as an analytical logic, (c) the adaptability of both mixed methods and grounded theory to different ontological (the nature of reality) and epistemological (legitimate ways of constructing knowledge) paradigms, and (d) the way it accommodates complexity. The role of verification is one topic where the methodologies differ (Dey, 1999). Each of the areas of overlap are discussed in the next section.

The iterative exchange embedded in the constant comparative method

Although the idea has been neglected in the methodological literature about either grounded theory or mixed methods, one of the principal arguments for the compatibility between the two is based on an endorsement voiced by the founders of grounded theory, now more than fifty years ago. The constant comparative method is an approach to analysis conventionally framed to mean that data analysis and data collection co-occur. Analysis begins as the first unit of data is encountered.

In MM-GTM this necessarily means that in at least one phase of the research process, quantitative and qualitative analysis unfolds simultaneously. This involves an iterative process where at some point in the analytical process qualitative and quantitative data are *not* kept separate during analysis. This process is facilitated by merging data into a single spreadsheet or database.

MM-GTM is adaptable for use with an existing data set. This could be the case, for example, with a study involving social media posts or a body of documents like case notes, where analysis could be launched with a small subset of the data. In this scenario, theoretical sampling might involve dipping back into a data set to target the selection of cases to further elaborate the properties of a core construct.

The analytical logic of abduction

A second area where the methodological assumption of grounded theory and mixed methods overlap is about abduction as an analytical logic. Abduction is one of three approaches to reasoning that include induction (generalizing from the specific to the general) and deduction (generalizing from the general to the specific) (Locke et al., 2008). In MM-GTM, abduction is the process of alternating between an exploratory and a confirmatory stance to systematically generate and test a series of alternative explanations.

Several noted methodologists maintain that it is more accurate to describe the analytical process at the center of grounded theory as an abductive one rather than an inductive one (e.g., Bryant, 2017; Timmermans and Tavory, 2012). Co-editor with Kathy Charmaz of the influential *SAGE Handbook of Current Developments in Grounded Theory* (2019), UK academician Anthony Bryant associated abduction as an analytical logic with one of the co-founders of the methodology. He wrote that while in later writing Barney Glaser maintained his advocacy for the idea of pure induction, the more pragmatic one of the original pair, Anselm Strauss, gave some attention to abduction when he struck out on his own (Bryant, 2017). Bryant wrote: “There is now growing recognition that although the process of induction certainly has a role in grounded theory, it is abduction that needs to be understood as a far more important – indeed *essential* (italics his)- aspect” (2017, p. 269).

In building a case for the pragmatic foundation of grounded theory, Morgan (2020) extends the logic of abduction to theoretical sampling and brings to light its link to verification. He suggested that purpose of theoretical sampling is not only to flesh out the properties of categories but to explore and elaborate working hypotheses through an abductive process that alternates almost seamlessly between induction and deduction (Morgan, 2020). Theoretical sampling is not a one-time process in this approach.

Although this is at odds with a “classic” or “formal” approach to grounded theory, on-going engagement with the literature is at the center of the abductive process. Abduction occurs within the context of previous knowledge, including other theoretical frameworks (Kelle, 2019). Abduction “assumes extensive familiarity with existing theory at the onset and throughout every research step” (Timmermans and Tavory, 2012, p. 173). The active role of the literature throughout an abductive approach to analysis, means that

the analytical process is not as purely inductive as the founders first envisioned. Investigators often reach into unfamiliar bodies of literature to make sense of puzzling findings throughout the research process.

Adaptability to complexity

The complex nature of individual actions and interactions within multi-layered socio-cultural context is a third rationale for the compatibility between mixed methods and grounded theory (Guetterman et al., 2017; Johnson et al., 2010). “We believe that the human world is especially rich and complex, and that MM-GT can offer an important method for understanding complexity” Johnson et al. observed (2010, p. 72). Complexity requires a dynamic research design that can accommodate changes in the research environment and navigate empirical findings that do not always cohere in expected ways (Poth, 2018).

There are multiple indicators of a complex research design. Three signatures of a complex research design include (a) when the core construct is multi-dimensional and framed as embedded in a socio-cultural context, (b) data are collected from multiple levels of an organization, or (c) when participants are distributed across geographically diverse settings (Poth, 2018). In a school setting, for example, this might include teachers, parents, students, and administrators in multiple different settings. Diversity among collaborators is another dimension of complexity (Poth, 2018). Saturation of the properties of all categories can be an elusive standard to achieve in complex research designs.

Compatibility with multiple paradigms

A fourth way that the methodological grounding of mixed methods and grounded theory is compatible is that they are both adaptable to diverse epistemological perspectives. Both grounded theory and mixed methods have been adapted by investigators operating from diverse epistemological paradigms. Urquhart (2013, p. 36) observed, for example, that grounded theory can be “appropriated by researchers with different assumptions about knowledge and how it can be obtained.” Grounded theory has been approached from diverse paradigmatic perspectives, post-positivism, interpretivism, constructivism, and postmodernism (Sebastian, 2019).

As with grounded theory methodology, researchers have approached mixed methods from the perspective of many different paradigms (Creamer, 2018a). The original textbooks to appear about mixed methods that first emerged between the late 1980s and 2003 are characterized by Denzin (2010) as a mirror of the post-positivist and quantitative perspective that was dominant at the time. The postpositivist orientation of these influential early textbooks may explain why quantitative methods are prioritized in the first generation of research to carry the mixed methods label. During this period, there were few opportunities for training in qualitative methods, much less mixed methods. Lack of familiarity with qualitative methods and their variety, may explain the all too common practice of applying a quantitative logic to the analysis of qualitative data by relying on counting and quantification.

Morgan proposed that a pragmatic approach guides most social science research methodology, including mixed methods (2007) and grounded theory (2020). Without overlooking its link to a pragmatic grounding, others writing about mixed methods affiliate with the lens of dialectical pluralism as a guide for investigators committed to the exploration of diverse perspectives (e.g. Johnson, 2012) or a commitment to social justice that is associated transformative research agenda. Researchers joining interdisciplinary endeavors are likely to need competencies in more than one method and more than one paradigm.

Distinct contributions of each method

That each method makes a distinct contribution is a second part of the methodological rationale for MM-GTM. For grounded theory, this includes assurances that there will be a genuine qualitative or inductive phase. An additional contribution is that it offers a set of procedures for data collection and analysis that are absent from mixed methods. Each of these can offset an inclination to downplay the qualitative portion of MM-GTM. The unique contribution of mixed methods includes a push to be inclusive of a broad range of sources of data, a logic that justifies attention to dissonance, and more options for sampling, verification, and understanding exceptions.

The topics of verification or confirmation and the treatment of exceptions are two areas where it is possible to pinpoint what can be gained through the pairing of mixed methods and grounded theory. Leading voices in the different approaches to grounded theory, including its founders, do not award much attention to procedures that contribute to verification, seeing the purpose of grounded theory as developing theory rather than verifying it (Dey, 1999). The same ambivalence extends to exceptions and outliers. Charmaz acknowledges the role of confirming hunches and promising leads but resists the idea of systematic attempts at verification in grounded theory (Morgan, 2020). There is little justification for verification in constructivism because all truth claims are equally valid (Bryant, 2009). Similarly, dissonance is only prioritized as a source of original insight in postmodern approaches to grounded theory (Aparmian et al., 2017).

There is as much variability in the ways that MM-GT studies have been implemented as there is in other types studies. In the next section, we shift from more abstract, philosophical issues to considering different research designs where ways the two can work in tandem. The first approach is a classic one in that it considers studies that are launched, much like instrument development, to develop and test and theoretical framework. Integration often serves multiple purposes in both of these designs.

Two approaches to implementing a MM-GT study

There are a wide variety of ways that MM-GTM can be implemented (Johnson and Walsh, 2019). Referring to it as mixed grounded theory (MGT), Johnson and Walsh wrote: “There are perhaps an infinite number of versions of MGT to be explored and ‘invented’ by practicing researchers in the production of knowledge” (p. 523). One implication of this statement is that there is no formulaic template for implementing a MM-GTM study. The more widely used approaches range from those that sequester the analysis of qualitative and quantitative into separate phases to those that employ a more iterative design where there is an on-going exchange between the qualitative and quantitative approaches in more than one phase of the research. Each achieves integration in a different way. Two different approaches to implementing a MM-GTM study are described in the next section.

Linking distinct qualitative and quantitative phases to develop and verify or elaborate a grounded theory

The ability to generate a theory qualitatively and to test it quantitatively was the first purpose envisioned for MM-GT (i.e., Johnson et al., 2010). Both Johnson et al. (2010) and Walsh (2015) advocate for the potential of MM-GTM to generate and test formal grounded theory that awards comparable recognition of findings emerging from the analysis of qualitative and quantitative data. Any claim that data from the qualitative and quantitative phases are given equal weight, is mitigated, however, by an improbably small sample size for the qualitative phase.

A formal grounded theory is often expressed in a visual display or model that identifies the core phenomenon, contextual and more local individual and social conditions that influence the phenomenon, actions/strategies, and consequences or outcomes (Johnson et al., 2010). One way that mixing occurs is through the construction of a composite model that combines qualitatively and quantitative derived constructs (Johnson et al., 2010). This type of theoretical framework has utility in the evidence-based movement when it provides a theoretical justification for a new activity or program.

Table 1 summarizes a number of the design features of MM-GTM studies that use a research design with self-contained phases for the qualitative and quantitative data collection. These can occur sequentially or concurrently. The table summarizes some common trends, without suggesting that there is not substantial variety in the ways these are implemented. The table identifies purposes for this design, how mixing is achieved, and the type of priority that is often awarded for each phase, among other features.

Mixing in this approach to MM-GTM occurs through linking or connecting of qualitative and quantitative phases of data collection and analysis. In a sequential design, as for instrument development, the second phase is influenced by the one that precedes it, but it can be disaggregated for purposes of future publications. This type of design often reflects a quantitative priority. Some methodologists take the position that the multi-method label is most appropriate for a project undertaken with the purpose of developing an instrument where the qualitative data are mined exclusively for purposes of constructing an instrument or research protocol.

There are some advantages to a research design with separate qualitative and quantitative phases of data collection and analyses. One benefit is to time management when research unfolds in a predictable way that aligns more closely to an original plan than those with an emergent quality. The value-added of the approach might be described as additive in that the purpose and contribution of each method is readily identifiable. Additional advantages of this design apply to introducing efficiencies to a collaborative project. It accommodates a clear division of labor, enabling collaborators working at a distance to conduct the research in a relatively independent manner. The two-phase design makes it possible for distinct epistemological philosophies to co-exist.

Even in a two-phase design where data analysis and data collection are conducted relatively independently, it is not unusual for preliminary findings from the analysis of different sources of data to conflict. It is not uncommon in a research project involving the implementation of an intervention for the quantitative data to fail to provide statistical support for the intended outcomes, while the qualitative data paint a more nuanced picture.

Table 1 Linking distinct qualitative and quantitative phases to develop and verify or elaborate a grounded theory.

<i>Design feature</i>	<i>Characteristic of MM-GTM</i>
Type of Mixing	Linking
Principal Purpose(s)	Develop and test the applicability of a grounded theory to more than one setting; develop a theoretical basis for instrument development.
Design	Two-phase, sequential mixed method design
Priority	Equal or unequal
Expertize of Collaborators	Likely to be distinct
Level of Interaction among Collaborators	Can be minimal
Likely Sources of Dissonance	Quantitative data fail to support the effectiveness of an intervention.
Examples	Bussing et al. (2011), Howell Smith et al. (2019), Kawamura et al. (2009).

Interactive designs with integration through embedding

The display of variety is most evident in MM-GTM studies and projects that are implemented with an interactive design where integration occurs at multiple points in the research cycle and serves multiple purposes. These can be planned from the onset as multi-stage designs with multiple stages of data collection and may include different combinations of sequential, parallel or convergent approaches with exploratory and explanatory phases. These can also include pairing with other research methodologies, like case study or participatory community-based action research. These types of projects might be described as having a “mixed priority”, rather than an equal one, when key insights emerge from analytical procedures that bring different sources of together and when the sum is greater than the parts (Creamer, 2018b).

The idea of research design is generally thought to apply to how a research project is planned, but it has dual meanings in that it extends to how research is actually implemented as it adapts to a dynamic research environment or unexpected findings (Hunter and Brewer, 2015). It often happens that an initial plan for a tidy two phase, sequential mixed method design evolves to a more complex or messy one. This can occur, for example, when the outcomes of an intervention are not in line with what was predicated from prior knowledge. Similarly, researchers may need to make unplanned detour to collect more data or to conduct additional analysis to find an explanation for why the quantitative results fail to document the effectiveness of an intervention (Creamer, 2018b). Other factors can contribute to complexity, such as differences of opinion among collaborators or an unexpected inability to support a long-standing theoretical assumption. There are different labels for the type of integration between findings from different qualitative and quantitative research methods that characterize an interactive mixed method design. A type of integration referred to as “embedding” is particularly well suited to grounded theory because of its link to the back and forth between analysis and data collection that is part of the constant comparative method. The “hallmark” of embedding “is recurrently linking qualitative data collection to quantitative data collection at multiple points” (Fetters et al., 2013, p. 2141). In a fully integrated mixed method approach to grounded theory (FIMM-GTM), there is an active exchange between different sources of data at multiple stages of the research process (Creamer, 2018b). A fully integrated approach can extend to working with multiple sources of qualitative data. Collaboration is often characterized by a high level of interaction and some overlap in expertise (Creamer, 2020).

Table 2 provides an overview of some of the key design features of an interactive MM-GTM design.

We see more examples of this approach to MM-GTM than with the first, more compartmentalized approach where integration occurs through linking.

A long-standing and very likely irreconcilable debate about the role of the literature in developing theoretical insight with grounded theory procedures very likely at the root of why the potential for MM-GT to extend or elaborate an existing theory has been disregarded (Creamer, 2021). The failure of findings from different sources of data to cohere, for example, led Kaplan and Duchon (1988) to dig more deeply into group differences. Similar encounters with incongruity have led to unexpected insight about the multi-dimensionality of a construct, revealed constructs that elaborated conditions that influenced the achievement of an expected out, or to trace causation suggested by a time-ordered process (Creamer, 2020).

An iterative approach to MM-GTM that engages data from different sources at multiple points during data collection and analysis has both advantages and disadvantages. The principal advantage is an enhanced potential to tackle the complexity that is inevitable whenever research is conducted in a dynamic research setting, requires multiple layers of expertise, considers constructs like identity or sustainability that are multi-dimensional, and embeds understanding of a phenomenon research within the type of multi-layered socio-cultural context identified by Plano-Clark and Ivankova (2016).

Disadvantages of an iterative approach to integrating findings from different sources of data in MM-GTM relates to the unpredictability and complexity introduced by integration. An abductive process can translate to a commitment of time and resources in analytical detours that ultimately prove unproductive. Finding the appropriate terminology to communicate succinctly with a diverse interdisciplinary audience is an additional challenge.

Table 2 Interactive designs with integration occurring during analysis.

<i>Design feature</i>	<i>Characteristic of MM-GTM</i>
Type of Mixing	Embedding
Principal Purpose(s)	Elaborate existing theory, explain dissonant or unexpected findings
Design	Analysis of one source is influenced by findings from another and vice versa.
Priority	Mixed
Expertise of Collaborators	Likely to overlap
Level of Interaction among Collaborators	Often characterized by high levels of interaction during analysis.
Likely Sources of Dissonance	Unexpected findings that emerge from integration
Examples	Catallo et al. (2012), Kaplan and Duchon (1988), Walsh (2014), Wesely (2010), Westhues et al. (2008)

Gauging the quality of MM-GTM

An expansive view of criteria for weighing the merits of a MM-GTM include indicators related to the research methods (transparency, saturation) and the analytical and conceptual outputs (utility and originality) that apply to all research methods. Methodological transparency or precision about describing core procedures, including integration, is the principal criterion for judging quality in all but a few of the standards of reporting for mixed methods. On the other hand, saturation is positioned as a necessary condition of research quality in grounded theory. The application of saturation, transparency, utility, and originality as indicators of the quality of reporting are discussed next.

Saturation

Saturation is part of the core procedures in grounded theory. It is a signal that data collection is sufficient when no new properties of a category emerge that contribute to further analytical insight. The expression, “saturation,” is applied relative to theoretical categories. Reaching a judgment about saturation includes consideration of disconfirming cases. Saturation is accomplished when no new properties about a category emerge (Charmaz, 2014). “Saturation means more than a one-time check,” Charmaz wrote (2014, p. 215). Saturation is linked to theoretical sampling and complicates any attempt to establish sample size ahead of time. The more multi-dimensional the core construct and complex the model, the less likely that saturation can be achieved. An expectation to saturate categories and their properties can have an impact on a final composite model. It is very likely to produce a more streamlined or parsimonious final theoretical model.

Methodological transparency

With few exceptions, virtually all of the standards of quality that have been adapted for mixed methods emphasize transparency about the research procedures, including integration, and the utilization of appropriate terminology to label the mixed method design. In MM-GTM the expectation for methodological transparency extends to using appropriate terminology and an explanation about how core grounded theory analytical procedures, like theoretical sampling, were executed. The emphasis on methodological transparency is evident in the first analysis of a body of MM-GTM research articles where quality was assessed by the number of core grounded theory procedures that were identified and explained (e.g., Guetterman, et al., 2017), without attending to integration.

Transparency about research methods is strongly related to its utility for other researchers, but it is no guarantee that a study is rigorous or that the conclusions are warranted. In MM-GTM, evaluative criteria extend to finding a way to warrant the conclusions, including key themes or constructs and relationships depicted in a final, composite theoretical model. The process of warranting a conclusion refers to documenting the link between data and key findings. In mixed methods, findings are often warranted through a table or figure or table that incorporates qualitative and quantitative data and traces the link to key themes or hypothesis.

Utility and originality

Utility and originality have been put forward as generic criteria to reach a judgment about the quality of the theoretical insight produced through a systematic set of procedures (Agerfalk, 2014; Corley and Gioia, 2011). Originality is related to the novelty of the analytical insight and the extent it adds to what has previously been known about the phenomenon. The argument for an original contribution is something a researcher generally pursues in the discussion section of a manuscript, where it is embedded in the wider body of literature. Theoretical contribution must be discussed in relation to existing theory. In MM-GTM, a novel contribution might be the addition of a recognition of how contextual factors, like a marginalized status, can interrupt conventional patterns of association.

Pragmatically oriented researchers have been at the forefront of advocating utility as a key indicator of quality. The pragmatic view is that the key issue is research produces conceptual and theoretical insight that prove useful (Bryant, 2009). This includes its scientific usefulness in generating new research as well as utility in practice (Corley and Gioia, 2011). In practice-oriented fields, like management, information systems, nursing and other medical fields, and education, includes the synergy produced in generating new ideas for evidence-based practice.

There are some unique challenges to gauging quality in MM-GTM, particularly about integration. The common practice of reporting on the qualitative and quantitative phases of a research project in separate publications can camouflage the significance of integration and complicate an attempt to assess its contribution and the breadth of its application. This complicates the practice of drawing conclusions about how integration occurred based on a single publication rather than the wider body of related work. This has implications for the treatment of dissonance or paradox as well. It is not uncommon to find, for example, that in complex projects reported over a series of inter-related publications, that authors end one article with a recognition of dissonance or paradox that cannot be explained readily, only to pick up a in a few years later.

Discussion

The goal of this encyclopedia entry has been to add to the argument for the credibility of MM-GTM by detailing a methodological rationale for the pairing. A methodology is a rationale for a research approach. This entry builds a rationale for using mixed methods and grounded theory in tandem by, first, identifying commonalities in their methodological assumptions, and, secondly, by highlighting some areas where each method can make a distinct contribution. Commonalities extend to the analytic logics, the way that grounded theory procedures can be adapted to accommodate diverse sources of data, and a shared set of contextual factors including the recognition of complexity and diverse perspectives. The overlap in key methodological features, supplemented by the distinct contribution of each method, is the basis of my argument that when pursued with a dialectical orientation MM-GTM can be considered a distinct methodology.

Fig. 1 summarizes key areas of overlap between the methodological framework of grounded theory and mixed methods. Three areas are highlighted: analytical logic, analytic procedures, and contextual factors.

Analytical logic is the first dimension where there is a shared compatibility between grounded theory and mixed methods. Items appearing in this box in the figure include abduction, the constant comparative methods, and adaptability to diverse philosophical paradigms. The second box in **Fig. 1** recognizes analytical procedures that can accommodate many types of approaches, including mixed methods. This applies to theoretical sampling, developing a coding system, and the role of the literature and other theories in serving as sensitizing concepts and to explain unexpected findings that emerge during the coding process. The third box in the figure refers to contextual factors that are a factor in both mixed methods and grounded theory. These include an aim to contribute analytical insight, the resilience to adapt to complex research questions and a dynamic research environment, the application largely in applied fields, and the likelihood that the inquiry engages many layers of a sociocultural environment.

The methodological assumptions of most approaches to grounded theory and mixed methods might differ in terms of the appreciation for the contribution of dissonance to generating theoretical insight, as well recognition of the role of verification and the treatment of exceptions. Very little space is awarded to these topics in most textbooks about grounded theory, particularly the first

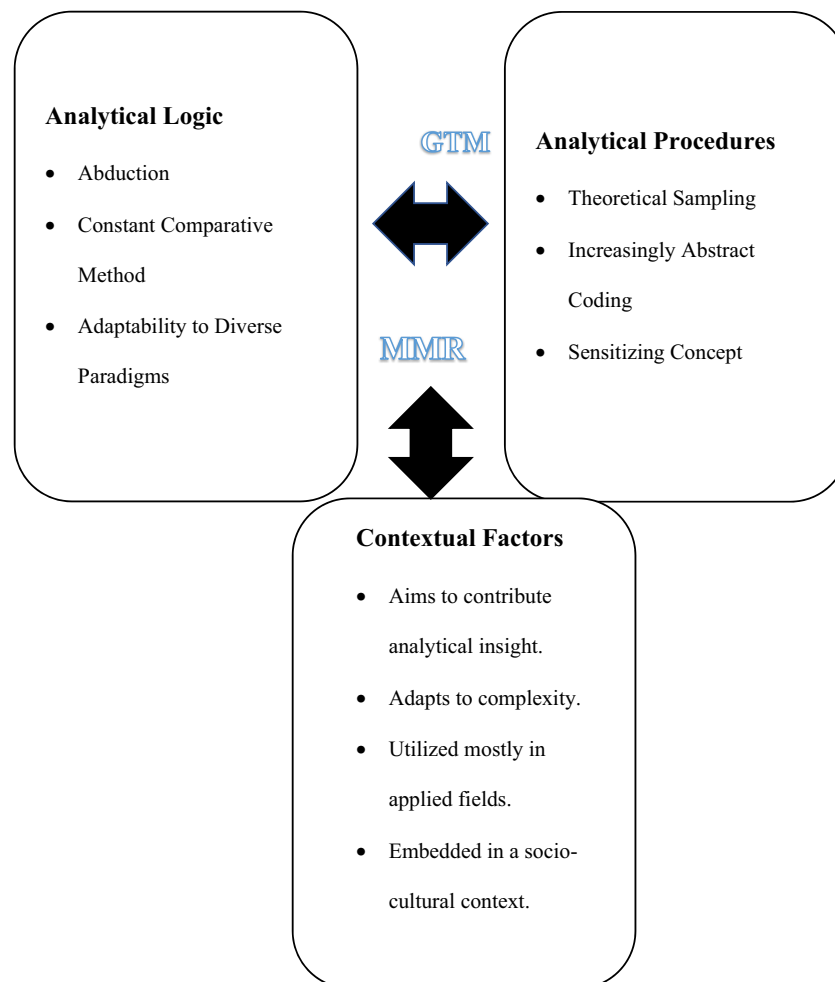


Fig. 1 Overlap between the analytic logic and procedures of grounded theory and mixed methods.

and second generation to appear. The act of engaging or integrating findings from different sources of data that occurs in mixed methods allows for confirmation, but at the same time the disconnects generate topics for further exploration.

Design considerations

The complexity of the type of research questions tackled in both grounded theory and mixed methods is one explanation for the remarkable variety in the way MM-GTM has been implemented in practice. Integration is accomplished in diverse ways and for diverse purposes. A two phase, sequential approach where the phases are linked can be used for multiple purposes, including for instrument development. It is a design that readily accommodates the classic purpose of developing and testing a theoretical framework that is developed through grounded methods. An interactive design is used more frequently with MM-GTM in part as a response to the emergent quality demanded when findings from different sources are engaged. The credibility of each of these approaches is based on utility, originality, and providing documentation that traces the links between the data and the results reported.

Implications for practice

The inclusion of grounded theory as at least one phase in a mixed method project has implications for supporting the methodological dynamism of both approaches. For proponents of grounded theory, the more expansive umbrella of mixed methods may promote further initiatives to partner with feminist and transformative approaches by allowing researchers to be creative about what sources of data can be useful and to include research that occurs in naturalistic settings. A robust sample size allows for a wider range of sampling strategies than is often found in a grounded theory study. For researchers with an inclination toward mixed methods, a component that retains the integrity of grounded theory will ensure that a genuine qualitative phase that involves an exploratory component is not overlooked.

Directions for future research

For readers with methodological interests, there is a wide range of topics related to MM-GTM that have not been explored in any detail in the literature. This includes on the topics of the (a) variety of sampling approaches that might be accommodated in MM-GTM, (b) ways that grounded theory has been used to explore contradictory findings, (c) its role in elaborating an existing theory, (d) how the procedures can be modified with a secondary analysis of data that has been collected already. Developing standards of quality unique to MM-GTM is yet another area worthy of attention.

References

- Agerfalk, P.J., 2014. Insufficient theoretical contribution: a conclusive rationale for rejection? *Eur. J. Inf. Syst.* 23, 593–599.
- Apramian, T., Cristancho, S., Watling, C., Lingard, L., 2017. (Re) grounding grounded theory: a close reading of theory in four schools of thought. *Qual. Res.* 17 (4), 359–376.
- Bryant, A., 2009. Grounded theory and pragmatism: the curious case of Anselm Strauss. *Forum Qual. Soc. Res.* 10 (3). Article 2. <https://www.qualitative-research.net/index.php/fqs/article/viewFile/1358/2851>.
- Bryant, A., 2017. *Grounded Theory and Grounded Visualization*. University Press, Oxford, UK.
- Bussing, R., Koro-Ljungberg, M.K., Noguchi, K., Mason, D., Mayerson, G., Garvan, C.W., 2011. Willingness to use ADHD treatments: a mixed methods study of perceptions by adolescents, parents, health professionals, and teachers. *Soc. Sci. Med.* 74, 92–100.
- Catallo, C., Ciliska, D., MacMillan, H.L., 2012. Minimizing the risk of intrusion: a grounded theory of intimate partner violence disclosure in emergency departments. *J. Adv. Nurs.* 69 (6), 1366–1376.
- Charmaz, K., 2014. *Constructing Grounded Theory*, second ed. Sage, Thousand Oaks, CA.
- Corley, K.G., Gioia, D., 2011. Building theory about theory building: what constitutes a theoretical contribution? *Acad. Manag. Rev.* 36 (1), 12–32.
- Creamer, E.G., 2018a. *An Introduction to Fully Integrated Mixed Methods Research*. SAGE, Thousand Oaks, CA.
- Creamer, E.G., 2018b. Enlarging the conceptualization of mixed method approaches to grounded theory with intervention research. *Am. Behav. Sci.* 62 (7), 919–934. <http://journals.sagepub.com/doi/10.1177/0002764218772642>.
- Creamer, E.G., 2020. Visualizing dynamic fully integrated mixed method designs. *Int. J. Mult. Res. Approaches* 12 (1), 1–13.
- Creamer, E.G., 2021. Extending the value-added of mixed methods in sustainability research. In: Factor, A., Ulhoi, J.P. (Eds.), *Sustainability and Small to Medium-Sized Enterprises: Lessons from Mixed Methods Research*. Routledge, Abingdon, UK.
- Denzin, N.K., 2010. Moments, mixed methods, and paradigm dialogs. *Qual. Inq.* 16 (6), 419–427.
- Dey, I., 1999. *Grounding Grounded Theory: Guidelines for Qualitative Inquiry*. Academic Press, San Diego, CA.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs-principles and practices. *Health Serv. Res.* <https://doi.org/10.1111/1475-6773.12117>.
- Guetterman, T.C., Babchuk, W.A., Howell Smith, M.C., Stevens, J., 2017. Contemporary approaches to mixed methods-grounded theory research: a field-based analysis. *J. Mix. Methods Res.* 13 (2), 179–195.
- Howell Smith, M.C., Babchuk, W.A., Stevens, J., Garrett, A.L., Wang, S.C., Guetterman, T.C., 2019. Modeling the use of mixed methods-grounded theory: developing scales a new measurement model. *J. Mix. Methods Res.* 14 (2), 184–206.
- Hunter, A., Brewer, J., 2015. Designing multi-method research. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Sage, Oxford, UK, pp. 185–205.
- Johnson, R.B., 2012. Guest editor's editorial: dialectical pluralism: a metaparadigm whose time has come. *Am. Behav. Sci.* 56, 751–754.
- Johnson, R.B., Walsh, I., 2019. Mixed grounded theory: merging grounded theory with mixed methods and multimethod research. In: Bryant, A., Charmaz, K. (Eds.), *The SAGE Handbook of Current Developments in Grounded Theory*. Sage, London, UK, pp. 517–531.

- Johnson, R.B., McGowan, M.W., Turner, L.A., 2010. Grounded theory in practice: Inherently a mixed method? *Res. Sch.* 17 (2), 65–78.
- Kaplan, B., Duchon, D., 1988. Combining qualitative and quantitative methods in information systems research: a case study. *Management Information Systems Quarterly* 12 (4), 571–586.
- Kawamura, Y., Ivankova, N., Kohler, C., Purumean-Chaney, S., 2009. Utilizing mixed methods to assess parasocial interaction of an entertainment-education program audience. *Int. J. Mult. Res. Approaches* 3, 88–104.
- Kelle, U., 2019. The status of theories and models in grounded theory. In: Bryant, A., Charmaz, K. (Eds.), *The SAGE Handbook of Current Developments in Grounded Theory*. Sage, London, UK, pp. 68–88.
- Locke, K., Golden-Biddle, K., Feldman, M.S., 2008. Making doubt generative: rethinking the role of doubt in the research process. *Organ. Sci.* 19 (6), 907–918.
- Morgan, D.L., 2020. Pragmatism as a basis for grounded theory. *Qual. Rep.* 25 (1), 64–73.
- Plano-Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Poth, C., 2018. Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking. Sage, Thousand Oaks, CA.
- Ralph, N., Birks, M., Chapman, Y., 2015. The methodological dynamism of grounded theory. *Int. J. Qual. Methods* 1–6.
- Sebastian, K., 2019. Distinguishing between types of grounded theory: classical, interpretive, and constructivist. *J. Soc. Thought* 3 (1), 1–9.
- Seidel, S., Urquhart, C., 2013. On emergence and forcing in information systems grounded theory studies: the case of Strauss and Corbin. *J. Inf. Technol.* 28, 237–260.
- Timmermans, S., Tavory, I., 2012. Theory construction in qualitative research: from grounded theory to abductive analysis. *Socio. Theor.* 30 (3), 167–186.
- Urquhart, K., 2013. *Grounded Theory for Qualitative Research: A Practical Guide*. Sage, London, UK.
- Walsh, I., 2014. A strategic path to study IT use through users' IT culture and IT needs: a mixed-method grounded theory. *J. Strat. Inf. Syst.* 23, 146–173.
- Walsh, I., 2015. Using quantitative data in mixed-design grounded theory studies: an enhanced path formal grounded theory in information systems. *Eur. J. Inf. Syst.* 24, 531–557.
- Wesely, P., 2010. Language learning motivation in early adolescents: using mixed methods research to explore contradiction. *J. Mix. Methods Res.* 4 (4), 295–312. <https://doi.org/10.1177%2F1558689810375816>.
- Westhues, A., Ochocka, J., Jacobson, N., Simich, L., Maiter, S., Janzen, R., Fleras, A., 2008. Developing theory from complexity: reflections on a collaborative mixed method participatory action research study. *Qual. Health Res.* 18 (5), 701–717.

Intersecting mixed methods with action and community-based participatory research approaches

Nataliya V. Ivankova^a and Nancy P. Wingo^b, ^a School of Health Professions, The University of Alabama at Birmingham, Birmingham, AL, United States; and ^b School of Nursing, The University of Alabama at Birmingham, Birmingham, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	567
Mixed methods research	568
Action research	568
Community-based participatory research	569
Common features of MMR, AR and CBPR approaches	569
Conceptual characteristics	569
Philosophical characteristics	569
Procedural characteristics	570
Advantages of intersecting MMR with AR and CBPR approaches in education	570
Challenges of intersecting MMR with AR and CBPR approaches in education	571
Methodological considerations of intersecting MMR with AR and CBPR	571
Levels of intersection	571
MMAR framework	572
MMAR study process model	573
Example of the MMAR process with MMR in the reconnaissance phase	573
Example of the MMAR process with MMR in the evaluation phase	573
Example of the MMAR process with MMR in both reconnaissance and evaluation phases	576
Applying MMAR framework to design studies that intersect MMR with AR and CBPR	576
Step 1: understanding the problem and community needs	576
Step 2: aligning information with the phases in the MMAR framework	576
Step 3: orienting to the MMAR study process and building community research capacity	578
Step 4: designing reconnaissance and/or evaluation phases of the MMAR study	578
Step 5: identifying the next steps in the MMAR study process	578
Conclusion	579
References	579

Glossary

Action research “A participatory process concerned with developing practical knowing in the pursuit of worthwhile human purposes. It seeks to bring together action and reflection, theory and practice, in participation with others, in the pursuit of practical solutions to issues of pressing concern to people” (Reason and Bradbury, 2008, p. 4)

Community-based participatory research “A systematic inquiry, with the collaboration of those affected by the issue being studied, for purposes of education and taking action or effecting change” (Green and Mercer, 2001, p. 1927)

Intersecting Intentional combination of two or more methodological approaches within a study design (Plano Clark and Ivankova, 2016)

Mixed methods research A research approach that emphasizes a meaningful integration of quantitative and qualitative approaches within a study or a program of inquiry with the purpose of generating more credible and persuasive conclusions about the research problem (Johnson et al., 2007; Tashakkori and Creswell, 2007)

Introduction

Mixed methods, action, and community-based participatory research approaches have long been applied in educational research. However, their meaningful intersection within a study provides enhanced opportunities to address complex educational problems by combining different methods and data sources, actively engaging stakeholders, and using cyclical approaches to finding solutions and bringing a positive change to a group or a community (Ivankova, 2015; Ivankova and Wingo, 2018; Johnson and Ivankova, 2019). A synergistic combination of multiple, action-oriented, and community-based methods allows researchers to explore the

course and outcomes of an issue, achieve more rigorous and transferable results to inform policy and practice, and deliver community-prioritized outcomes (DeJonckheere et al., 2019; Kim, 2014; O'Sullivan et al., 2017).

A methodological intersection occurs when two or more research approaches are intentionally joined within a study design. Plano Clark and Ivankova (2016) described intersecting mixed methods research (MMR) with other quantitative and qualitative research designs, methodological approaches, and theoretical frameworks by adding or embedding an MMR design to another approach, thus creating complex designs (Creswell and Plano Clark, 2018). Such intersection adds an MMR perspective to the research logic of another approach and design with consideration for their established research processes and practices.

Intersecting MMR with action research (AR) and community-based participatory research (CBPR) provides pragmatic opportunities for researchers to partner with stakeholders (e.g., administrators, faculty, staff, students, and parents) in education fields to identify and address areas in need of reform. Involving various stakeholders in research not only allows for a better understanding of their needs, but it can also promote solutions that are developed collaboratively rather than being imposed in a "top-down" approach, thereby enhancing the potential for stakeholder ownership (Ivankova, 2015). The systematic approach of the AR cycle also allows for iterative changes as needed. Ultimately, using MMR as part of AR and CBPR approaches can create policies or processes that can be sustained or revised in educational practice after the research has concluded (Ivankova and Wingo, 2018).

In the following sections we provide a brief overview of MMR, AR, and CBPR approaches. We discuss the common features of these approaches and the advantages of their meaningful intersection within a study context. We then introduce a mixed methods action research (MMAR) framework and its tools and discuss the procedural steps for applying the framework to design studies that intersect MMR with AR and CBPR.

Mixed methods research

Mixed methods researchers systematically collect quantitative and qualitative data and analyze them in an integrative manner in a single study (Tashakkori et al., 2021). An MMR approach can enhance the ability of researchers to conduct robust inquiry into a topic with complex research questions by using procedures that merge quantitative and qualitative data that have been collected concurrently or sequentially (Johnson and Onwuegbuzie, 2004). Integrating different types of data allows researchers to confirm knowledge using quantitative methods and generate knowledge using qualitative methods (Tashakkori et al., 2021). Researchers may choose to integrate quantitative and qualitative methods for various reasons, including exploring divergent views, enhancing credibility by confirming results, comparing and contrasting differences, expanding possibilities for discovery, and discovering pragmatic solutions to difficult problems in natural settings (Creswell and Plano Clark, 2018; Greene et al., 1989).

There are three core designs for MMR studies, depending on the timing or sequence of quantitative and qualitative data collection and analysis: (1) sequential QUAN → QUAL designs, in which quantitative data are first collected and analyzed, followed by qualitative data collection and analysis to expand on or explain results; (2) sequential QUAL → QUAN designs, in which qualitative data are first collected and analyzed, followed by quantitative data collection and analysis to generalize or confirm results; and (3) concurrent QUAN + QUAL designs, in which quantitative and qualitative data are collected and analyzed independently, with data integrated afterward to compare, contrast, and corroborate results (Creswell and Plano Clark, 2018; Plano Clark and Ivankova, 2016). These MMR designs may be used with other research approaches such as AR and CBPR, depending upon research questions and purposes (Ivankova, 2015). (See the cited sources for further description of these designs. You can also learn more about these designs by reading the chapter by Tajima and Feters in this volume.)

MMR has been increasingly popular among educators for decades. Creswell and Garrett (2008) traced a heightened interest in mixed methods research, in part, to the work of educators such as Jennifer Greene and colleagues in the late 1980s. As the use of MMR has increased, educators have also recognized the methodological advantages it provides in its ability to be combined with other research approaches such as AR and CBPR to address educational problems in practical ways (Ivankova and Wingo, 2018; Plano Clark and Ivankova, 2016).

Action research

AR seeks answers to questions that concern people in their groups or communities through a systematic approach that involves participants in all phases to develop "knowing in the pursuit of worthwhile human purposes" (Reason and Bradbury, 2008, p. 4). Due to its practical focus and collaborative nature, AR is often used as an umbrella term with reference to research that incorporates participatory and practice-oriented approaches (Duesbery and Twyman, 2020; Herr and Anderson, 2005). AR offers a process to address immediate needs and resolve problems in a specific setting in a more efficient way than traditional research (Greenwood and Levin, 2007). It also allows for inclusion of stakeholders as co-researchers in the research process, which can generate the potential for more effective solutions (McNiff and Whitehead, 2011). Rather than doing research on a group, researchers conduct the research together with group members, thereby promoting ownership of the research by the group participants (Brydon-Miller et al., 2003; McNiff and Whitehead, 2011). Participant empowerment is created throughout the research cycle, which ends with an intervention or action for improvement that also includes a mechanism to evaluate the intervention outcomes and their sustainability (Kemmis, 1982). By iteratively conducting research in this cyclical manner, educational researchers add to a body of knowledge that informs best practices for teaching in diverse settings (Hinchey, 2008; Stringer, 2021). AR also offers opportunities for collegial collaboration to accomplish goals that individual educators could not address alone (Sagor, 2000).

The process of AR allows researchers and stakeholders to combine theory with practice and action with reflection (Lewin, 1948; Reason and Bradbury, 2008). Kurt Lewin (1948) originally described AR as a four-step iterative process of reflecting, planning, acting, and observing, with a goal of solving real-life problems. Researchers begin by reflecting on possible solutions to a practical problem. They then gather facts about the problem (also known as *reconnaissance*) that allow them to plan how they might solve it. After implementing planned action steps to test their solution, they evaluate results, reflect on possible changes, and revise actions as needed. This cyclical process may be repeated continuously until positive changes are made to improve practice (Kemmis, 1982).

Community-based participatory research

CBPR is often viewed as a specific type of an AR approach that, as the term implies, is rooted in a community (Senge and Scharmer, 2001). The CBPR approach is often described as a “systematic inquiry, with the collaboration of those affected by the issue being studied, for purposes of education and taking action or effecting change” (Green and Mercer, 2001, p. 1927). Often used in public health and other healthcare fields, CBPR is based on the idea that understanding the culture and epistemology of a community is key to the inquiry process (DeJonckheere et al., 2019; Lucero et al., 2018).

To meet the needs of the community, CBPR incorporates a collaborative approach between researchers and community members at each stage of the research, seeking solutions to problems that impact everyday lives and empowering participants to enact social change (Israel et al., 2001; McNiff and Whitehead, 2011). This partnership can allow for perspectives of both “outsiders” and “insiders,” creating a trusting bond through shared power that can appeal to members of populations who might be skeptical about others proposing solutions to their problems (Lucero et al., 2018). Though CBPR can be “very complex, time-consuming, and resource intensive” (DeJonckheere et al., 2019, p. 497), it can give focus and depth to collaborative efforts to improve communities and create lasting impacts on the well-being of people who live in them. In the following section, we describe the common features of MMR, AR, and CBPR approaches and the advantages of intersecting them to address real-life educational problems.

Common features of MMR, AR and CBPR approaches

A number of shared conceptual, philosophical, and procedural characteristics make a meaningful intersection of MMR with AR and CBPR approaches possible and advantageous (Ivankova, 2015; Ivankova and Wingo, 2018; Ivankova and Johnson, 2021).

Conceptual characteristics

MMR, AR, and CBPR aim at providing a comprehensive exploration of the studied issue. MMR integrates quantitative and qualitative methods, data, and results, thus increasing the breadth and depth of understanding (Johnson et al., 2007). AR and CBPR rely on cyclical spirals of critical action and reflection with stakeholder engagement to design robust action plans grounded in more weighted solutions (McKernan, 1988).

Addressing complex issues requires dialectical thinking (Greene, 2007). MMR addresses a range of exploratory and confirmatory questions within a single study by creatively integrating available data and diverse perspectives (Tashakkori et al., 2021). The dialectical nature of AR and CBPR is reflected through the iterations of reflection, planning, acting, and observing (Lewin, 1948), while incorporating diverse stakeholders’ viewpoints in each stage of the research spiral.

Reflection is an important component of any research process. Due to a multi-strand nature of MMR, designing and conducting an MMR study requires reflection about a range of theoretical and procedural issues related to the research focus and design of each study phase and integration of the quantitative and qualitative components. Reflection is an inherent component of AR and CBPR as researchers critically reflect over the assessment of the identified problem, make decisions about the required intervention, and consider the action outcomes to inform further action steps (Mills, 2011).

MMR, AR, and CBPR are highly collaborative in nature. The breadth and scope of mixed methods studies requires different research strategies calling for different forms of expertise and a collaborative team approach (Shulha and Wilson, 2003). Collaboration is one of the principles of AR and CBPR, because it requires collegial interactions, active participation, and joint problem-solving by all stakeholders and at all stages in the study process (McNiff and Whitehead, 2011).

Philosophical characteristics

Pragmatism that rejects “either-or” thinking and views knowledge as both a process and a product of empirical investigation (Johnson and Gray, 2010) provides a major philosophical foundation for MMR, AR, and CBPR. Pragmatism rejects the quantitative and qualitative incompatibility thesis and helps justify the integration of quantitative and qualitative methods within an MMR study. The need to design and implement action plans that would provide effective solutions of practical problems also requires a “what works” approach in AR and CBPR studies.

The three approaches are often seen as value laden and applying a transformative or advocacy lens to advocate for social justice. Moreover, a transformative mixed methods approach is often viewed as embedded in participatory AR (Mertens et al., 2010), because it encourages community engagement to produce knowledge that reflects the power and social relationships in the society. AR and CBPR approaches are grounded in the practice of co-creating knowledge and empowering participants for change through an iterative process of action and learning (Herr and Anderson, 2005).

Balancing insider-outsider perspectives is common in MMR, AR, and CBPR. In MMR, researchers take an outsider observer role when they collect and analyze numeric or etic data. Alternatively, they take an insider role when they qualitatively explore and interpret perspectives and experiences of individuals, thus producing emic data. AR and CBPR often involve multiple stakeholders and collaborative partnerships, which helps preserve both the insider and outsider perspectives in the inquiry process (Herr and Anderson, 2005).

Procedural characteristics

MMR, AR, and CBPR are designed and conducted following a set of systematic procedures or steps, moving from the identification of the research problem and formulation of research questions to data collection, analysis, interpretation, and evaluation. Both quantitative and qualitative data are required when there is a need to address confirmatory and exploratory questions within a single MMR study. AR and CBPR traditionally rely on multiple sources of quantitative and qualitative evidence to find comprehensive solutions to complex practical issues (Koshy et al., 2011).

The three approaches follow a cyclical process consisting of clearly defined study phases. An MMR study often consists of multiple quantitative and qualitative strands that are implemented sequentially in an iterative manner, with one study strand informing another strand (Tashakkori et al., 2021). In AR and CBPR studies, one step leads to another and is repeated multiple times, forming a research spiral aimed to generate well-informed solutions (Herr and Anderson, 2005).

Advantages of intersecting MMR with AR and CBPR approaches in education

As we have discussed, intersecting MMR with AR and CBPR allows researchers to address practical issues in systematic ways that enhance the translation of research into practice. Table 1 outlines the advantages of combining these approaches in more detail. Studies intersecting MMR with AR and CBPR may be found across diverse fields of education, including elementary education (Kostos and Shin, 2010; Owen and Davies, 2020), secondary education (Bogar et al., 2017; Ollis and Harrison, 2016; O'Sullivan et al., 2017; Snyder et al., 2016; Zocher and Hougham, 2020) and higher education (Cabaroglu, 2014; Sampson, 2010; Tsai and Jao, 2020).

One advantage to using MMR with AR in educational studies is that it offers researchers the opportunity to better understand how the unique culture of a school may influence the implementation of pedagogy. For instance, one Montessori school in Western Australia that served children aged 18 months to 15 years used MMR in an AR project to determine how to introduce digital learning technologies to students without compromising the school's pedagogical and philosophical values (Owen and Davies, 2020). Researchers may also use MMR in AR projects to determine whether accepted educational practices will be effective with specific groups of students. Tsai and Jao (2020) conducted a study to examine how team-based learning (TBL) strategies influenced student learning and teacher course design in one radiologic technology course at a junior college in Taiwan. Quantitative data included student performance scores, while qualitative data included answers from open-ended surveys of students and a series of three teacher interviews (Tsai and Jao, 2020).

Table 1 Advantages of intersecting MMR with AR and CBPR.

Addressing a practical issue in a systematic and dialectic way through

- A combination of theory with practice and action with reflection
- Exploration of diverse views of stakeholders across a community
- Comprehensive assessment of a problem from different perspectives
- Development of a reliable and valid action/intervention plan tailored to diverse needs
- Rigorous evaluation of the action/intervention
- Opportunities for iterative changes if needed
- Credible and valid conclusions about action/intervention outcomes
- An evidence-based approach to action/intervention monitoring
- Enhanced action/intervention sustainability

Enhancing translation of research into practice by

- Cultivating an understanding of community culture and epistemology
- Enhancing stakeholder engagement at every phase in an action research cycle
- Creating a sense of ownership and empowerment among stakeholders
- Building bonds of trust among community members to enhance social justice
- Optimizing action/intervention outcomes
- Illuminating and assessing change over time
- Empowering young people to become change agents
- Supporting effectiveness of change action
- Promoting sustainability of meaningful change
- Enhancing transferability of action/intervention results to other contexts and community settings

Adapted from Ivankova and Wingo (2018) with permission of Sage Publishing

Intersecting MMR with CBPR can also be especially appropriate for educational research, as the involvement of stakeholders in CBPR research designs enhances opportunities to address advocacy issues for marginalized groups and promotes inclusion of multiple voices. For example, [Zocher and Hougham \(2020\)](#) used MMR in a CBPR study in an urban high school that explored how two different approaches to experiential projects enhanced students' environmental literacy and their abilities to advocate for changes in the community to overcome environmental oppression. Participants included students, staff, and alumni of the school; data were collected through observations, surveys, interviews, and document reviews ([Zocher and Hougham, 2020](#)). Another CBPR study in environmental education ([Bogar et al., 2017](#)) used MMR to assess urban youth's knowledge of and attitudes toward lead poisoning and other environmental health concerns by establishing a research team including a doctoral candidate, an education coordinator, and a group of young people ages 10–18 at a local youth organization.

In fact, a specific type of CBPR for educators to consider is “youth-led participatory action research” (YPAR), which empowers young people by engaging them at all stages of research to identify and solve problems that affect them ([Martinez et al., 2020](#); [Suleiman et al., 2021](#)). YPAR can be especially effective in helping young people develop leadership skills and providing opportunities for diverse and marginalized youth to address social justice issues ([Akorn, 2011](#); [Suleiman et al., 2021](#)).

Challenges of intersecting MMR with AR and CBPR approaches in education

The characteristics that make intersections of MMR with AR and CBPR so effective in educational research can also be challenging to implement. By its very nature, collaboration can be difficult, as it requires all stakeholders to be engaged in research tasks and appropriately responsive to feedback from other team members. Engaging stakeholders as co-researchers in MMR studies can be particularly challenging due to the complexity of mixed methods designs and the need for expertise in quantitative and qualitative data collection and analysis. Moreover, it requires a collaborative effort between researchers and community partners to construct joint meanings of the mixed methods data ([Shulha and Wilson, 2003](#)). Fostering co-learning and capacity building among all partners to enable a meaningful intersection of MMR with AR and CBPR is one of the principles of effective community-based action research studies ([Israel et al., 2001](#)).

Issues around trust and shared power among team members can also cause a breakdown in research progress ([Herr and Anderson, 2005](#)). The need to reflect takes time and patience, and the cyclical process of research designs that combine MMR with AR and CBPR may require extensive time and resources. Since the cyclical nature of MMAR projects is viewed as being “a continual improvement process” ([Hinchey, 2008](#), p. 5), educators are encouraged to take a reflective stance and critically evaluate the results from each cycle to determine whether the desired improvements have occurred or can be feasible.

Studies that are conducted with an advocacy lens almost inevitably face stakeholder discomfort or even power struggles, as they often call for dramatic changes within community culture. Considering contextual factors and carefully adapting effective interventions to specific community needs can help mitigate power structures and ensure the achieved results are tailored to the cultural practices and fit the worldviews of the target group ([Wingood and DiClemente, 2008](#)). Qualitative research approaches as part of a larger MMR study can help solicit stakeholders' views and their first-hand experiences with the issue to compare them with quantitative assessments of the action/intervention outcomes.

The challenges in combining MMR with AR and CBPR must be acknowledged as part of initial steps in study design. Collaboration among research teams can be strengthened if all members resolve to participate fully in the research process, communicate consistently, and provide honest and timely feedback. Setting expectations about flexible timelines may prevent unreasonable project deadlines and allow time for reflection and critical evaluation. Finally, for the project to drive meaningful change, research teams must understand the complexity of potential issues and commit to persevering despite obstacles they may encounter.

Methodological considerations of intersecting MMR with AR and CBPR

Levels of intersection

Ten shared characteristics of MMR, AR, and CBPR discussed above suggest an explicit blending of conceptual, philosophical, and methodological premises when intersecting MMR with AR and CBPR. Therefore, a meaningful intersection of MMR with AR and CBPR requires careful considerations at several levels, including epistemological, axiological, and methodological ([Ivankova and Johnson, 2021](#)).

At the epistemological level, researchers make decisions about how to generate the types of knowledge about the studied problem by capitalizing on the advantages of MMR, AR, and CBPR. For example, the decisions relate to the type of research questions and the integration of the quantitative and qualitative approaches to address these questions, the inclusion of stakeholders' perspectives and participants' voices, and working with communities to understand their needs and to identify meaningful solutions to a practical problem from their perspective ([Bradbury, 2015](#)). At the axiological level, researchers consider the values associated with the epistemological and methodological choices during the research process. When using MMR, researchers allow the qualitative and quantitative components to interact within a study by bringing stakeholder values into a dialog and interpretation of the quantitative results ([Greene, 2007](#); [Johnson, 2017](#)). AR and CBPR approaches promote collaborative and empowering research activities that encourage researcher reflexivity and power sharing between the researcher and researched ([Israel et al., 2001](#); [Kral and Allen, 2015](#)).

At the methodological level, MMR, AR, and CBPR intersect when researchers combine these approaches within a methodological framework to inform the planning, design, and implementation of MMAR/CBPR studies. The MMAR framework was first developed to provide a systematic approach to support decision-making when combining MMR with AR (Ivankova, 2015; Ivankova and Wingo, 2018). It was also applied to promote patient-centered research (Ivankova, 2017) and as a means of enhancing translation of knowledge into practice (Ivankova et al., 2018). In the following sections we describe the MMAR framework and its practical tool, the MMAR study process model, which aims to help researchers translate the MMAR framework into a practical application.

MMAR framework

The MMAR methodological framework depicted in Fig. 1 conceptually follows the cycle of AR methodological steps outlined by Lewin (1948) and captures the synergetic combination of the two approaches by intersecting MMR with AR in every step. The framework includes six iterative phases (diagnosing, reconnaissance, planning, acting, evaluation, monitoring) during which conceptual, procedural, and inferential aspects of MMR inform the steps in the AR cycle in different ways. Full MMR studies consisting of conceptualization, quantitative and qualitative data collection and analysis, and developing integrated inferences are conducted in the reconnaissance and evaluation phases using one of the core MMR designs (see Creswell and Plano Clark, 2018). MMR inferences from these phases inform the rest of the steps in the AR cycle. Consistent with the participatory principles of AR, each phase in the framework involves collaboration with study participants and key stakeholders, whose active involvement can provide support for the project's success in the community. In Fig. 1, solid arrows indicate the sequence of the phases in the AR process, while dashed arrows show other possible iterations of the research activities within an AR cycle.

During the diagnosing phase, the problem that requires solution in the workplace or community setting is identified, and the advantages of addressing it by using both MMR and AR are explored in collaboration with community partners. In the reconnaissance phase, an MMR study is conducted to generate more context-specific understanding of the community's needs and to help define the outcomes that require improvement. The MMR inferences from the reconnaissance phase are then used to plan a data-driven intervention and to inform its implementation during the acting phase. During the evaluation phase, an MMR design informs the collection, analysis, and integration of the quantitative and qualitative data to assess the intervention outcomes and to identify further steps. The MMR inferences drawn from the evaluation phase inform the intervention monitoring to help achieve a satisfactory resolution of the problem with the input of stakeholders and community. Monitoring helps identify whether the problem is successfully resolved or requires further diagnosis, reconnaissance, or re-planning of the intervention, which is shown by dashed arrows in Fig. 1.

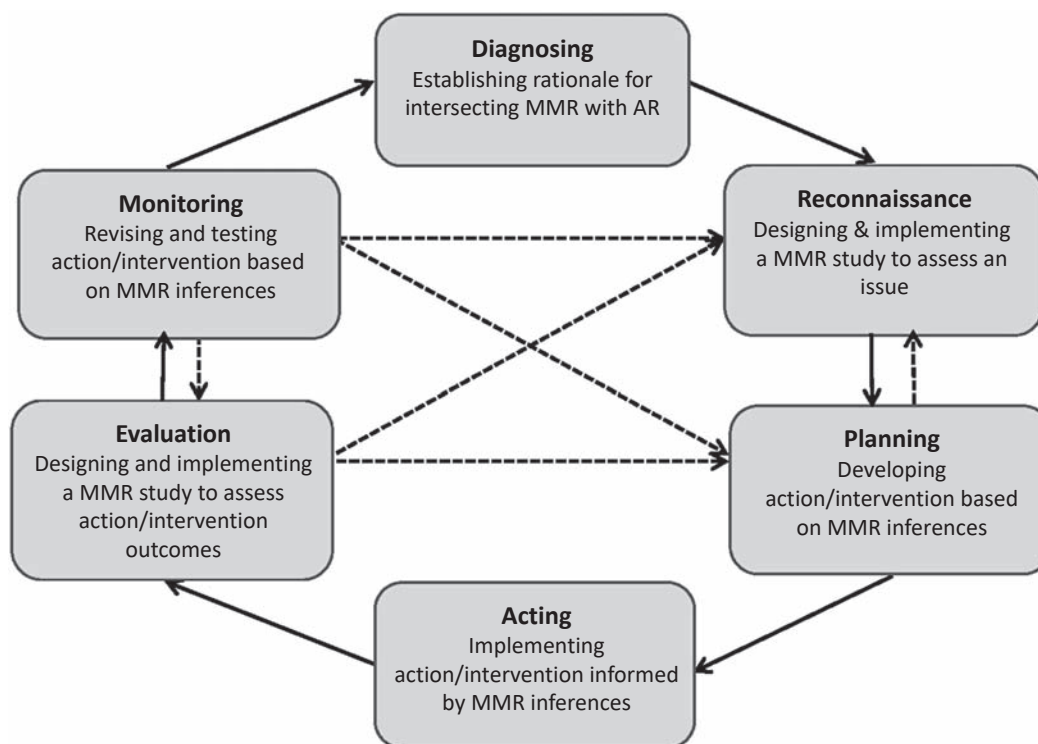


Fig. 1 Mixed methods action research framework. MMR: mixed methods research. Adapted from Ivankova (2015) with permission of Sage Publishing.

MMAR study process model

The MMAR study process model is a visual tool that helps apply the MMAR framework to conceptualize, design, and implement a study that intersects MMR with AR and CBPR to address a practical problem (Ivankova, 2015, 2017). The model unfolds the MMAR framework into an MMAR process to assist researchers with identifying the procedural steps within each phase of the AR cycle, using integrated MMR data and inferences, and soliciting the input of stakeholders and community partners while maintaining a focus on the overarching issue. By explicitly presenting methodological details for each phase, the model provides a roadmap for researchers to engage the key stakeholders in developing a plan to improve, monitor, and further adapt an effective intervention.

In the following sections, we illustrate the application of the MMAR framework and the MMAR study process model in education using three studies that intersected MMR with AR and CBPR approaches in different phases (reconnaissance, evaluation, and both phases) in the AR cycle. We briefly describe the study and the authors' rationale for intersecting MMR with AR and CBPR and discuss the MMAR process using diagrams depicting specific MMR and AR procedures during each phase. Solid arrows show the flow of the activities that were implemented in the study and reported in the paper, whereas dashed arrows suggest further steps related to monitoring of the implemented intervention.

Example of the MMAR process with MMR in the reconnaissance phase

Ivankova and Wingo (2018) described the MMAR process in Ollis and Harrison's (2016) AR study that used MMR in the reconnaissance phase to explore capacity for building a comprehensive sexuality education program using the health promoting school framework (HPS) in a five-campus, year 1–12 college in Victoria, Australia. The researchers chose AR because the spiral of "self-reflection, planning, implementation, monitoring and reflection on the data gathered, replanning in the light of reflection, and entering a new phase of action and monitoring, fitted with the aims of whole school change" (Ollis and Harrison, 2016, p. 141). Using MMR in the reconnaissance phase allowed researchers to employ quantitative methods to survey students, parents, teachers, and other key stakeholders to examine their attitudes and beliefs, while also using qualitative methods (participant observation and semi-structured interviews) to explore school culture and determine staff readiness to implement the program. Fig. 2 shows the MMAR process based on Ollis and Harrison's (2016) study.

The researchers used a concurrent (QUAN + QUAL) mixed methods design to determine critical factors to develop, implement, and potentially sustain the program. They surveyed year 9 students ($n = 34$), teachers ($n = 101$), parents and caregivers (no number reported), students at various levels (no number reported), and participants in professional development workshops ($n = 25$) to collect quantitative data. They also conducted semi-structured interviews with working group members and observations of working group meetings, teacher professional development and planning sessions, parent and community forums, and other school events. The researchers integrated these quantitative and qualitative data to plan and articulate the key aspects of the program (e.g., vision, strategic goals, program alignment, and stakeholder roles). They also implemented a professional development workshop series for teachers and conducted parent and community forums to create a support system for the program. The sexuality education program was successfully implemented, and although researchers did not report on the program evaluation and monitoring, which is marked by dashed arrows in Fig. 2, they emphasized that program assessment and sustainability would need to be researched in the future using quantitative and qualitative methods to better understand students' experiences and teachers' capacity.

Example of the MMAR process with MMR in the evaluation phase

Kostos and Shin (2010) applied the MMAR process to investigate the impact of math journaling instruction on second-grade students' communication of mathematical thinking by using MMR in the evaluation phase (Ivankova, 2015). An AR approach was chosen because it enabled the teacher-researcher "to utilize the insight that can only be obtained as an insider to the setting" (p. 226) and to capture the students' thinking process more closely. MMR was used in the evaluation phase "to provide a more in-depth look at how the students communicated their mathematical thinking when using math journals" (p. 226).

The study was conducted in a second grade mixed-ability classroom in a suburban school in Chicago. Using a concurrent (QUAN + QUAL) mixed methods design, the data were collected from 16 students, eight girls and eight boys, to assess whether a 5-week math journaling intervention had improved the second graders' communication of mathematical thinking. The quantitative data consisted of an identical math assessment administered to students before and after the intervention using the Illinois State Board of Education Mathematics Scoring Rubric. The qualitative data included students' math journal entries, interviews with eight randomly selected students, and the teacher-researcher's reflective journal. Fig. 3 shows the MMAR process based on Kostos and Shin's (2010) study.

Comparison of students' pre- and post-test scores, qualitative themes from students' math journals and interviews, and teachers' reflections allowed the researchers to obtain corroborating evidence that the use of math journals positively influenced the students' communication of mathematical thinking and use of math vocabulary. Based on these findings, the teacher-researcher decided to more systematically incorporate math journaling in daily math lessons. Solid arrows in Fig. 3 show the actual flow of the research activities in the study, while dashed arrows highlight potential future activities related to further evaluation and monitoring of the math journaling instruction or the need to conduct more literature review and students' needs assessment to inform effective changes in the instructional approach.

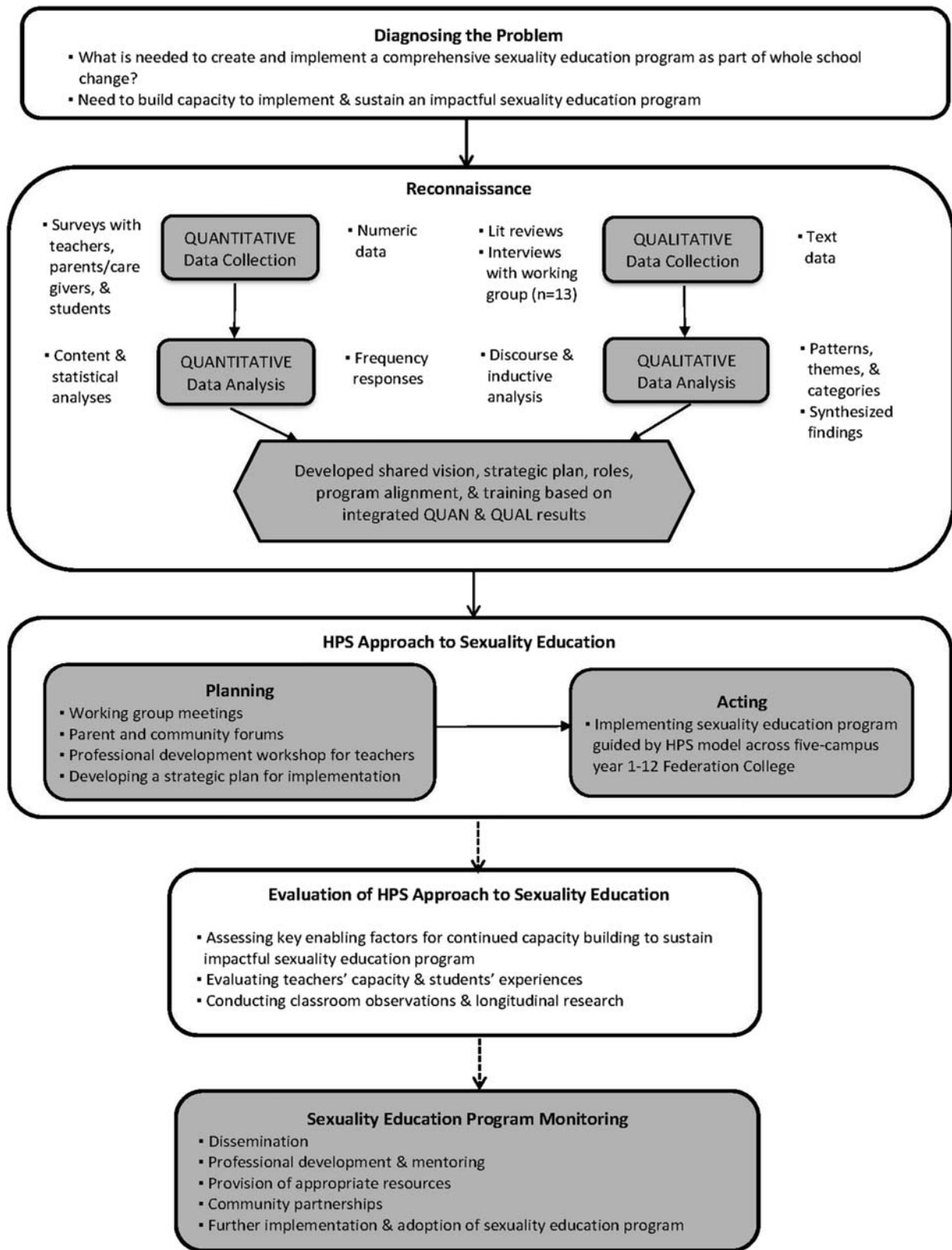


Fig. 2 MMAR process based on Ollis and Harrison's (2016) study. Reprinted from Ivankova and Wingo (2018) with permission of Sage Publishing.

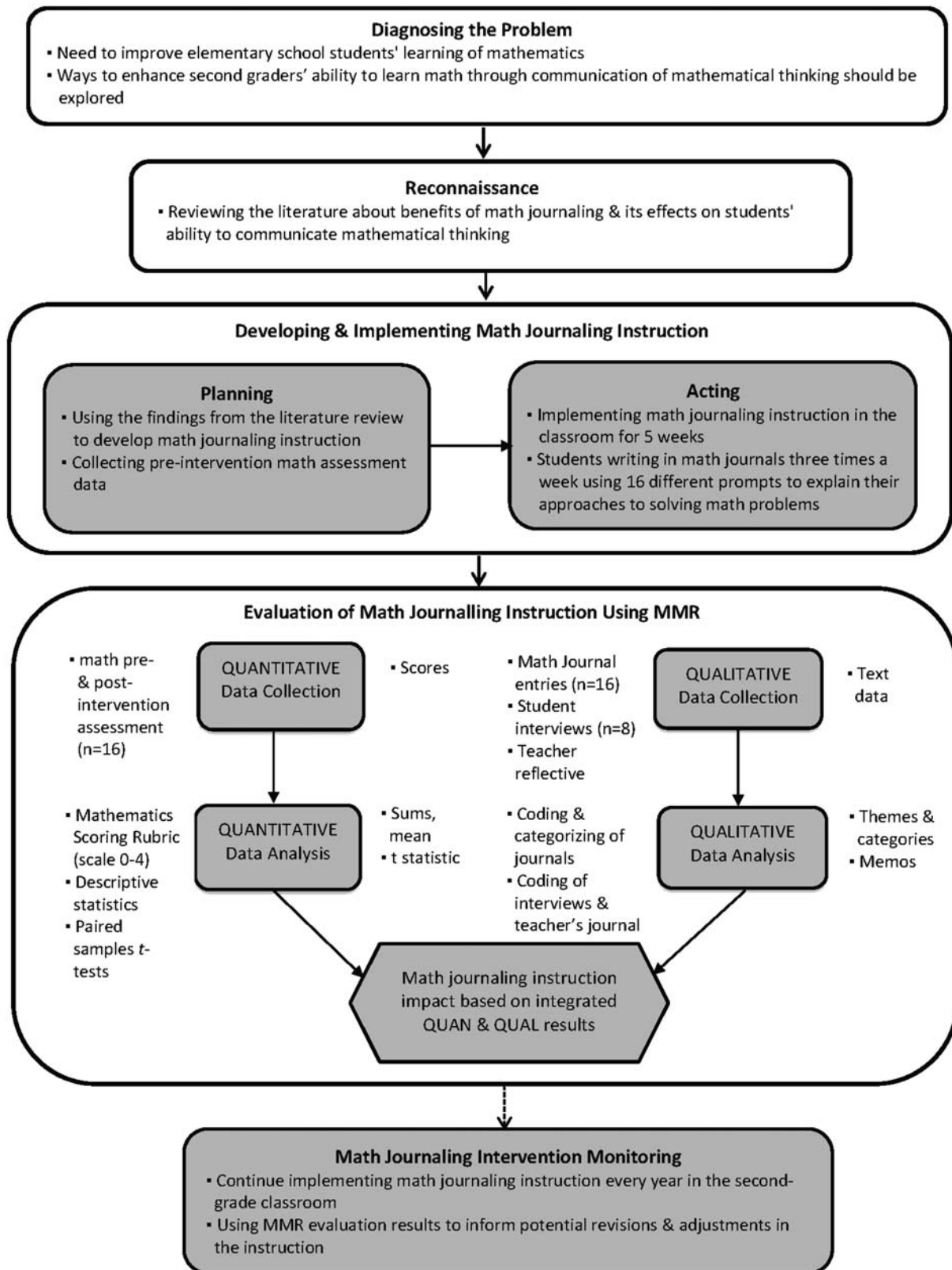


Fig. 3 MMAR process based on Kostos and Shin's (2010) study.

Example of the MMAR process with MMR in both reconnaissance and evaluation phases

Sampson (2010) applied the MMAR process in a semester-long study that used students' feedback to guide the choice of a lesson style in a college level class in which English was the language of instruction (Ivankova, 2015). AR was chosen for its participatory and collaborative nature and the fact that "the subjective views of learners are of most importance in enacting change" (Sampson, 2010, p. 286). MMR was used in both reconnaissance and evaluating phases for triangulation purposes to "most effectively inform the change-action" (p. 286).

Twenty-two first-year students in two Interpersonal Communication classes from a women's university in Japan took part in the study. During the reconnaissance phase, a concurrent (QUAN + QUAL) mixed methods design was used for collecting and analyzing the data, which consisted of a quantitative survey (the Lesson Style Questionnaire) and a narrative exercise (the Language Learning Autobiography), to identify students' preferred lesson styles. The integrated results from the analysis of the two data sets were used to develop task-based oriented lessons (TBL) to introduce more opportunities for using the English language in the classroom setting. The researcher employed a different sequential (QUAL → QUAN) mixed methods design to evaluate the instructional intervention. First, qualitative reflective narratives about the learning activities during each class from the students' journals were analyzed for themes to inform the development of the quantitative Learning Experience Questionnaire tailored to students' needs. The students completed the questionnaire at the end of the course. The quantitative results were then compared with the qualitative themes from the learning journal entries to obtain more informed conclusions about the effectiveness of the new lesson style and students' perceptions of their speaking and communication abilities in English. Fig. 4 shows the MMAR process based on Sampson's (2010) study.

Solid arrows in Fig. 4 show the actual flow of the research activities in the study, while dashed arrows suggest potential further evaluation and monitoring of the TBL intervention with the possibility for conducting more assessment of new students' preferences for EFL lesson style instruction. It is likely that the researcher will return to the reconnaissance phase when teaching another course to identify the lesson style preferences of a different cohort of Japanese college students. This cohort may be different from the one used in the study, and the previously developed intervention consisting of task-based lessons may not meet the needs of these students.

Applying MMAR framework to design studies that intersect MMR with AR and CBPR

The MMAR framework provides a methodological foundation for designing studies that intersect MMR with AR and CBPR approaches. The following procedural steps that align with the MMAR framework help translate the application of the framework in research practice (Ivankova and Johnson, 2021; Johnson and Ivankova, 2019). Table 2 summarizes the steps and their alignment with the phases in the MMAR framework.

Step 1: understanding the problem and community needs

During the first step, researchers collaborate with stakeholders to identify the problem that requires improvement using an MMAR approach and to understand what has been already done to address the issue. The diagnosing phase can also involve exploration of the previous activities and initial needs assessment by iteratively cycling through the diagnosing and reconnaissance phases. In Ollis and Harrison's (2016) study, researchers engaged all critical stakeholders including students, teachers, parents, and community members to help them understand how to design and implement a school comprehensive sexuality education program. Sampson (2010) relied on students' past English learning experiences and hopes for college speaking classes to motivate students to practice English in a classroom. Kostos and Shin (2010) conducted an extensive literature review and students' baseline math assessment to help students improve their math communication skills.

Step 2: aligning information with the phases in the MMAR framework

Based on the information about a practical problem gathered during Step 1 (diagnosing phase), researchers identify the point of entry into the AR cycle, which can be either developing an intervention tailored to stakeholders' and community needs or starting with the evaluation of the existing intervention or program. The entry point depends on whether this is a newly diagnosed issue or an ongoing problem with some actions already taken toward its resolution within the phases of the AR cycle. Using the MMAR framework, researchers map the activities to the MMAR process model to guide the research process and communicate the planned research activities to stakeholders for their support and input. In Ollis and Harrison's (2016) and Sampson's (2010) studies, researchers started with the reconnaissance phase because they wanted first to explore stakeholders' perceptions of and experiences with the problem before developing a plan for action. Included were not only students and teachers who would be directly affected by a sexuality education program, but also parents and other key stakeholders, whose perspectives were crucial for understanding the program's vision, mission, and potential resources. In Kostos and Shin's (2010) study, researchers started with implementing and evaluating the math journaling intervention because it was a reported effective pedagogical approach.

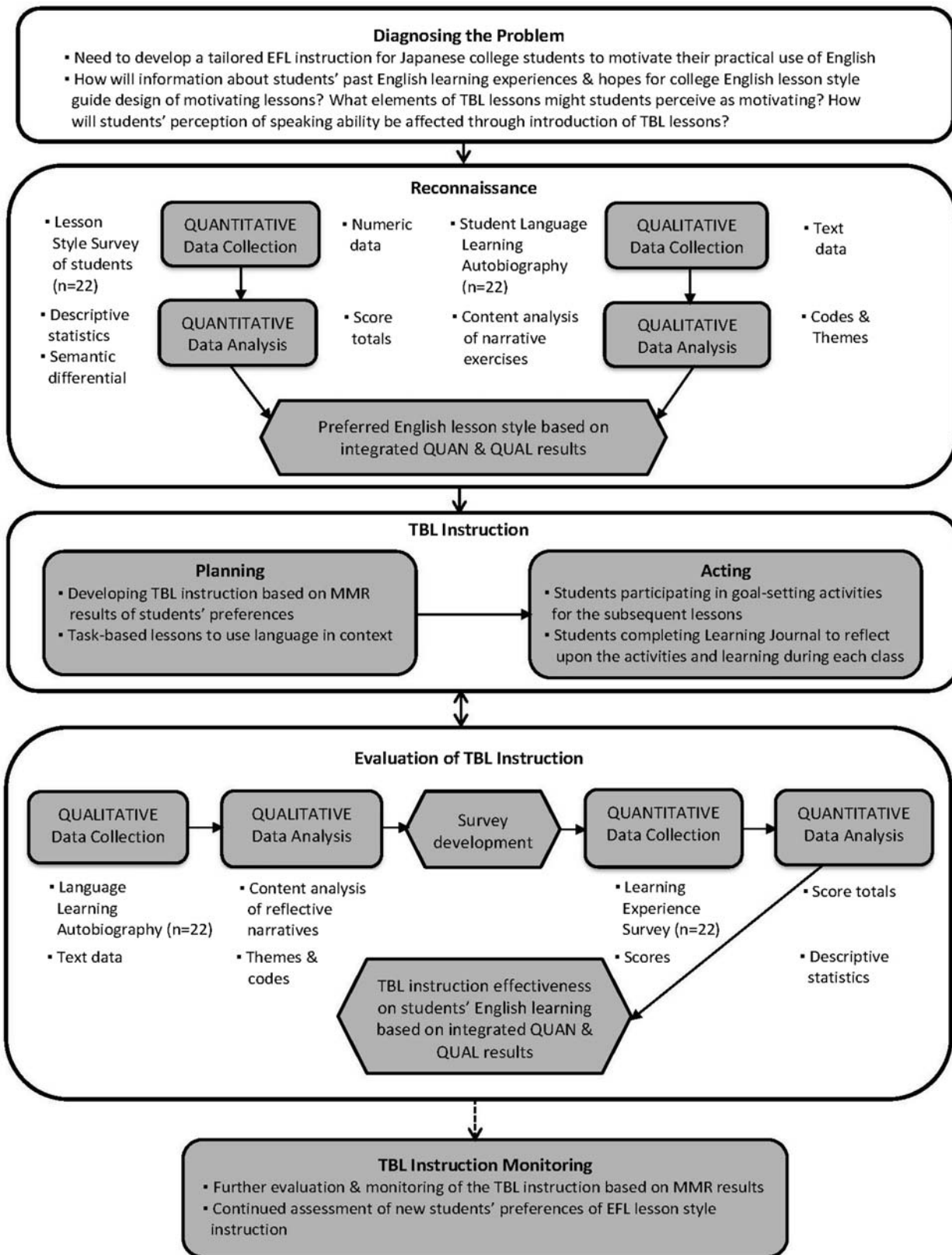


Fig. 4 MMAR process based on Sampson's (2010) study.

Table 2 Procedural steps for applying the MMAR framework to design studies that intersect MMR with AR and CBPR.

<i>Procedural steps</i>	<i>Alignment with the MMAR framework</i>
Step 1. Understanding the problem and community needs.	- Identifying the problem area that requires improvement using an MMAR approach. - Conducting needs assessment by iteratively cycling through the diagnosing and reconnaissance phases.
Step 2. Aligning information with the phases in the MMAR framework.	- Identifying the entry into the AR cycle depending on the state of the problem and community needs. - Framing prior activities within the phases of the AR cycle.
Step 3. Orienting to the MMAR study process and building community research capacity.	Unfolding the steps in the MMAR process and identifying preceding and subsequent phases in the MMAR study cycle.
Step 4. Designing reconnaissance and/or evaluation phases of the MMAR study.	Designing and implementing reconnaissance and/or evaluation phases using a MMR design to inform the development of the action/intervention plan and/or to evaluate intervention effectiveness and promote sustainability.
Step 5. Identifying the next steps in the MMAR study process.	Using the meta-inferences and deliverables that have been created during the reconnaissance and/or evaluation phases to identify the need for potential iterations between and among the phases.

MMR: mixed methods research; AR: action research; MMAR: mixed methods action research.

Adapted from [Ivankova and Johnson \(2021\)](#) with permission of Routledge.

Step 3: orienting to the MMAR study process and building community research capacity

Capitalizing on the collaborative nature of the AR and CBPR approaches, the next step involves engaging stakeholders and community partners with the research process in a meaningful and practical way. This includes educating all involved parties about the MMAR process while seeking their input into the phases of the MMAR study cycle and the components of the intervention plan. Using an MMAR study process model that visually depicts the evolving research activities makes it easier for researchers to work with community partners to reveal additional gaps in knowledge. In [Ollis and Harrison's \(2016\)](#) study, researchers worked with multiple community partners to identify the supporting factors for developing, implementing, and sustaining a sexuality education program. Training opportunities were provided to teachers, parents, and community members to help establish the infrastructure for the program implementation. [Sampson \(2010\)](#) engaged his students as co-researchers in developing the TBL instruction building on their previous language learning experiences. In [Kostos and Shin's \(2010\)](#) study, second graders actively participated in the research activities as part of their classroom duties.

Step 4: designing reconnaissance and/or evaluation phases of the MMAR study

During this step researchers design and conduct MMR studies in the reconnaissance and/or evaluation phases relying on available data, resources, and stakeholder participation to inform the subsequent phases in the MMAR study cycle. Decisions involve a choice of the MMR design, the timing of quantitative and qualitative data collection and analysis, and the planned outcomes from each phase. MMR inferences or integrated study conclusions ([Tashakkori et al., 2021](#)) are generated to provide information for intervention development, implementation, or monitoring. In [Ollis and Harrison's \(2016\)](#) study, the integrated results from the concurrent collection and analysis of the qualitative interviews with the members of the program's working group and quantitative surveys of teachers, students, and parents allowed researchers to better determine vision, strategic goals, program alignment, and roles of various stakeholders in the program implementation. In [Kostos and Shin's \(2010\)](#) study, the integrated quantitative and qualitative findings from students' math assessments, journals, and interviews provided evidence that math journaling enhanced student learning. In [Sampson's \(2010\)](#) study, the use of different MMR designs in the reconnaissance (concurrent) and evaluation (sequential) phases provided a comprehensive approach to developing and evaluating an instructional intervention resulting in improved language learning outcomes.

Step 5: identifying the next steps in the MMAR study process

Given the cyclical and iterative nature of an MMAR study that continues following the phases in the MMAR framework until the problem is positively resolved, the phases are repeated based on the outcomes from each phase. Researchers use the meta-inferences from the reconnaissance and/or evaluation phases to identify the next steps and potential iterations between and among the phases in collaboration with stakeholders. It took several years and multiple data sources to plan, develop, and implement the sexuality education program in [Ollis and Harrison's \(2016\)](#) study. The cyclical nature of the MMAR approach and continued capacity building "fitted with the aims of whole school change" (p. 141). In [Kostos and Shin's \(2010\)](#) study, although not explicitly stated, more evaluation of math teaching and learning strategies for second grade students would be conducted to inform potential revisions or adjustments in the intervention. In [Sampson's \(2010\)](#) study, it is likely that the researcher will return to the

reconnaissance phase when teaching another course to identify the lesson style preferences of a different cohort of Japanese college students. A new cohort may be different from the one used in the study, and the previously developed intervention may not meet these students' needs.

Conclusion

Educators often face challenges that require them to collaborate with diverse stakeholders to identify problems, design and implement solutions, and evaluate outcomes. Developing and evaluating educational interventions grounded in a systematic and comprehensive MMR assessment of the problem in collaboration with stakeholders helps to more efficiently reveal critical areas for action and find solutions that are meaningful and acceptable to all affected by the issue. Informed by the MMAR framework, the cyclical MMAR study process model provides a logical yet flexible structure to guide problem-solving, while the participatory nature of AR and CBPR approaches allows for the inclusion of multiple voices to advocate for stakeholder groups. Intersecting MMR with AR and CBPR approaches can provide robust, data-driven evidence to support decisions, reveal new insights, and drive meaningful and sustainable change in educational practice.

References

- Akom, A.A., 2011. Black emancipatory action research: integrating a theory of structural racialization into ethnographic and participatory action research methods. *Ethnogr. Educ.* 6 (1), 113–131.
- Bogar, S., Szabo, A., Woodruff, S., Johnson, S., 2017. Urban youth knowledge and attitudes regarding lead poisoning. *J. Community Health* 42, 1255–1266.
- Bradbury, H. (Ed.), 2015. *The SAGE Handbook of Action Research*, third ed. Sage, Thousand Oaks, CA.
- Brydon-Miller, M., Greenwood, D., Maguire, P., 2003. Why action research? *Action Res.* 1 (1), 9–28.
- Cabaroglu, N., 2014. Professional development through action research: impact on self-efficacy. *System* 44, 79–88.
- Creswell, J.W., Garrett, A.L., 2008. The “movement” of mixed methods research and the role of educators. *S. Afr. J. Educ.* 28, 321–333.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- DeJonckheere, M., Lindquist-Grantz, R., Toraman, S., Haddad, K., Vaughn, L.M., 2019. Intersection of mixed methods and community-based participatory research: a methodological review. *J. Mix. Methods Res.* 13 (4), 481–502.
- Duesbery, L., Twyman, T., 2020. 100 Questions (and Answers) About Action Research. Sage, Thousand Oaks, CA.
- Green, L.W., Mercer, S.L., 2001. Can public health researchers and agencies reconcile the push from funding bodies and the pull from communities? *Am. J. Public Health* 91 (12), 1926–1929.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Policy Anal.* 11 (3), 255–274.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.
- Greenwood, D., Levin, M., 2007. *Introduction to Action Research: Social Research for Social Change*, second ed. Sage, Thousand Oaks, CA.
- Herr, K., Anderson, G.L., 2005. *The Action Research Dissertation: A Guide for Students and Faculty*. Sage, Thousand Oaks, CA.
- Hinchey, P.H., 2008. *Action Research: Primer*. Peter Lang, New York, NY.
- Israel, B.A., Schulz, A.J., Parker, E.A., Becker, A.B., 2001. Community-based participatory research: policy recommendations for promoting a partnership approach in health research. *Educ. Health* 14 (2), 182–197.
- Ivankova, N.V., Johnson, S.L., 2021. Designing integrated mixed methods action research studies. In: Hitchcock, J., Onwuegbuzie, A. (Eds.), *Routledge Handbook for Advancing Integration in Mixed Methods Research*.
- Ivankova, N.V., Wingo, N., 2018. Applying mixed methods in action research: methodological potentials and advantages. *Am. Behav. Sci.* 62 (7), 978–997.
- Ivankova, N.V., Herbey, I., Roussel, L., 2018. Theory and practice of using mixed methods in translational research: a cross-disciplinary perspective. *Int. J. Mult. Res. Approaches* 10 (1). <http://ijmra.org/inaugural-special-issue/>.
- Ivankova, N.V., 2015. *Mixed Methods Applications in Action Research: From Methods to Community Action*. Sage, Thousand Oaks, CA.
- Ivankova, N.V., 2017. Applying mixed methods in community-based action research: a framework for engaging stakeholders with research as means for promoting patient-centeredness. *J. Res. Nurs.* 22 (4), 282–294. <https://doi.org/10.1177/1744987117699655>.
- Johnson, S.L., Ivankova, N.V., 2019. Intersecting mixed methods and action research to evaluate a community-based care coordination service: implications for conceptualizing and designing stakeholder-informed and context-specific health service research studies. *Int. J. Mult. Res. Approaches* 11 (1), 29–44. <https://doi.org/10.29034/ijmra.v11n1a1>.
- Johnson, B., Gray, R., 2010. A history of philosophical and theoretical issues for mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage Publications, Thousand Oaks, CA, pp. 69–94.
- Johnson, B., Onwuegbuzie, A., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26.
- Johnson, B., Onwuegbuzie, A., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1 (2), 112–133.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11 (2), 156–173.
- Kemmis, S. (Ed.), 1982. *The Action Research Reader*. Deakin University Press, Geelong, Australia.
- Kim, K.H., 2014. Community-involved learning to expand possibilities for vulnerable children: a critical communicative, Sen's capability, and action research approach. *J. Mix. Methods Res.* 8 (3), 308–316.
- Koshy, E., Koshy, V., Waterman, H., 2011. *Action Research in Healthcare*. Sage, Thousand Oaks, CA.
- Kostos, K., Shin, E., 2010. Using math journals to enhance second graders' communication of mathematical thinking. *Early Child. Educ. J.* 38, 223–231.
- Kral, M.J., Allen, J., 2015. Community-based participatory action research. In: Jason, L.A., Glenwick, D.S. (Eds.), *Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative, and Mixed Methods*. Oxford University Press, Oxford, UK, pp. 253–262.
- Lewin, G. (Ed.), 1948. *Resolving Social Conflicts: Selected Papers on Group Dynamics by Kurt Lewin*. Harper and Brothers Publishers, New York, NY.
- Lucero, J., Wallerstein, N., Duran, B., Alegria, M., Greene-Moton, E., Israel, B., et al., 2018. Development of a mixed methods investigation of process and outcomes of community-based participatory research. *J. Mix. Methods Res.* 12 (1), 55–74.
- Martinez, L.S., Yan, C.T., McClay, C., Varga, S., Zaff, J.F., 2020. Adult reflection on engaging youth of color in research and action: a case study from five U.S. cities. *J. Adolesc. Res.* 35 (6), 699–727.
- McKernan, J., 1988. The countenance of curriculum action research: traditional, collaborative, and emancipatory-critical conceptions. *J. Curric. Superv.* 3 (3), 173–200.
- McNiff, J., Whitehead, J., 2011. *All You Need to Know About Action Research*, second ed. Sage, London, UK.

- Mertens, D.M., Bledsoe, K.L., Sullivan, M., Wilson, A., 2010. Utilization of mixed methods for transformative purposes. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 193–214.
- Mills, G.E., 2011. *Action Research: A Guide for the Teacher Researcher*, fourth ed. Pearson Education, Inc, Boston, MA.
- Ollis, D., Harrison, L., 2016. Lessons in building capacity in sexuality education using the health promoting school framework: from planning to implementation. *Health Educ.* 16 (2), 138–153.
- O'Sullivan, K.C., Howden-Chapman, P., Sim, D., Stanley, J., Rowan, R.L., Clark, I.K.H., Morrison, L.L.A., The Waiopehu College Research Team, 2017. Cool? Young people investigate living in cold housing and fuel poverty. A mixed methods action research study. *Popul. Health* 3, 66–74.
- Owen, S., Davies, S., 2020. Maintaining an empowered school community: introducing digital technologies by building digital literacies at Beehive Montessori School. *Lond. Rev. Educ.* 18 (3), 356–372.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Reason, P., Bradbury, H., 2008. Introduction. In: Reason, P., Bradbury, H. (Eds.), *The SAGE Handbook of Action Research: Participative Inquiry and Practice*, second ed. Sage, Thousand Oaks, CA, pp. 1–10.
- Sagor, R., 2000. *Guiding School Improvement With Action Research*. Association of Supervision and Curriculum Development (ASCD), Alexandria, VA.
- Sampson, R.J., 2010. Student-negotiated lesson style. *RELC J.* 41 (3), 283–299.
- Senge, P., Scharmer, O., 2001. Community action research: learning as a community of practitioners, consultants and researchers. In: Reason, P., Bradbury, H. (Eds.), *Handbook of Action Research: Participative Inquiry and Practice*. Sage Publications, Thousand Oaks, CA, pp. 238–249.
- Shulha, L.M., Wilson, R.J., 2003. Collaborative mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 639–669.
- Snyder, C., Besozzi, D., Paska, L., Oppenlander, J., 2016. Is flipping worth the fuss: a mixed methods case study of screencasting in the social studies classroom. *Am. Second. Educ.* 45 (1), 28–45.
- Stringer, E.T., 2021. *Action Research*, fifth ed. Sage, Thousand Oaks, CA.
- Suleiman, A.B., Ballard, P.J., Hoyt, L.T., Ozer, E.J., 2021. Applying a developmental lens to youth-led participatory action research: a critical examination and integration of existing evidence. *Youth Soc.* 53 (1), 26–53.
- Tashakkori, A., Creswell, J., 2007. The new era of mixed methods. *J. Mix. Methods Res.* 1 (1), 3–8.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Tsai, M., Jao, J., 2020. Evaluation of the effectiveness of student learning and teacher instruction on team-based learning during quality control of diagnostic imaging. *Med. Educ. Online* 25 (1), 1732159.
- Wingood, G.M., DiClemente, R.J., 2008. The ADAPT-ITT model: a novel method of adapting evidence-based HIV Interventions. *J. Acquir. Immune Defic. Syndr.* 47 (Suppl. 1), S40–S46.
- Zocher, J.L., Hougham, R.J., 2020. Implementing ecopedagogy as an experiential approach to decolonizing science education. *J. Exp. Educ.* 43 (3), 232–247.

Mixed methods and qualitative comparative analysis

Heather Kane^{a,c} and Leila Kahwati^{b,c}, ^a School of Professional Advancement, Tulane University, New Orleans, LA, United States; ^b School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States; and ^c RTI international, North Carolina, United States

© 2023 Elsevier Ltd. All rights reserved.

History of QCA development	581
Method overview	581
QCA within mixed methods evaluation	582
Examples from the field	582
Process for using QCA	584
Considerations for use	584
Limitations	586
Concluding thoughts	586
References	586

Mixed methods research and evaluation are commonly used within the social sciences. Qualitative Comparative Analysis (QCA) is a method and analytic approach that can be used within a larger mixed methods evaluation. As a method, it entails a process and series of decisions based on theory, empirical information, and feasibility; as an analytic approach, it is used to identify complex patterns in data using non-statistical mathematical techniques (Schneider and Wagemann, 2012; Kahwati and Kane, 2020). In this encyclopedia chapter, we provide a brief introduction to the method, offer examples of how it has been used, and discuss considerations for its use.

History of QCA development

QCA was developed by a sociologist, Dr. Charles Ragin, in the early 1980s (Ragin, 1999, 2000, 2006). Ragin recognized the need for an approach within the social sciences that could assess complex causal relationships that are best formulated in terms of set relationships, rather than statistical relationships. For example, hypothesizing that school programs that provide disadvantaged students with mentors leads to increases in high school graduation rates can be expressed in terms of a set relationship. Stated as a set relationship, school programs with mentors are a subset of all schools with increased graduation rates. Similarly, he noted the need to define a systematic method to identify causally relevant commonalities across cases, which is a feature of cross-case comparison within qualitative methods (Ragin, 2000). Finally, because Ragin worked primarily in the comparative historical sociology and political science arena, the size of the samples available to researchers in this field were inherently restricted (e.g., numbers of countries, jurisdictions, events), which limited the use of traditional statistical methods. Since its initial development, many other social scientists have contributed to the further development of QCA methods. Increasing examples of its use in empirical research across a variety of academic fields including social and environmental policy research (Verma, 2012; Ide, 2015), management research (Fiss, 2011; Frambach et al., 2016), education (Kintz et al., 2015; Coburn et al., 2012), and healthcare and public health services research (Kahwati et al., 2016a, b; Kane et al., 2017; Thygeson et al., 2012) is also evident.

Method overview

QCA has several defining features (Schneider and Wagemann, 2012; Kahwati and Kane, 2020; Ragin, 2000, 2008). First, it entails a case-based approach, which means that cases are treated holistically and not deconstructed into component variables for analysis. Treating a case holistically means examining the complex interplay of factors that can lead to an outcome, rather than quantifying the degree to which any single variable is correlated with an outcome. Second, the method can integrate qualitative and quantitative data, though ultimately all data is transformed (or calibrated in QCA terms) into numeric values for analysis. Third, QCA involves using formal logic and set theory, a branch of mathematics that entails the examination of the relationships between sets, for cross-case comparison to identify set relationships between explanatory factors (*conditions* in QCA terminology) and outcomes. The two types of set relationships that QCA can identify if they are present in the data are *necessity* and *sufficiency*. A necessary condition (or combination of conditions) must be present (or present at a high level) for an outcome to occur but does not guarantee the outcome occurs. A sufficient condition indicates that where that condition (or combination of conditions) is present, the outcome is also present (Ragin, 2000, 2008). Finally, QCA is based on an underlying assumption of causal complexity within the phenomenon being studied. Three concepts characterize causal complexity: equifinality, conjunctural causation, and asymmetrical causation. These underlying assumptions are what distinguish QCA from probabilistic methods of testing for associations. **Box 1** illustrates these concepts using cookie recipes.

Box 1 Causal complexity explained: pathways to a good chocolate chip cookie (Kahwati and Kane, 2020)

The chocolate chip cookie, a favorite of many people was invented by American Chefs named Ruth Graves Wakefield and Sue Brides in 1938. It is made from flour, butter, white and brown sugar, vanilla, with little chips of chocolate mixed in. When preparing these cookies, both home and professional bakers want the desirable outcome of a “good tasting chocolate chip cookie.”

Many different chocolate chip cookie recipes are available. If one were to do an internet search and collect all the chocolate chip cookie recipes and then create a single recipe by taking the average of all the flour called for across the recipes, the average of all the butter called for, the average of all the sugars, and so on for all of the other ingredients, one could produce a single recipe, which would be a *unifinal* recipe. A cookie produced from this *unifinal* recipe represents the effect of using the average amounts of ingredients from across all of the various recipes. Some of the contributing recipes might not have been that good, and some of them might have been great. With the *unifinal* recipe, one has only that single recipe to use. Back to research for a moment, the probabilistic methods we use are based on assumptions of *unifinality*. In reality, many different recipes can result in a “good-tasting chocolate chip cookie;” there is not just one, single recipe that will lead to this outcome. This is *equifinality*, the concept that more than one combination of factors (or ingredients) can lead to an outcome.

For *conjunctural causation*, the cookie ingredients lend insight. Many recipes for a good-tasting cookie call for the use of both baking soda (bicarbonate) and baking powder (bicarbonate plus an acid like cream of tartar). A recipe with only baking powder leads to a cookie that is too tangy (acidic) and that is too pale. A recipe with only baking soda does not rise well and has a bitter taste, but a recipe with both baking soda and baking powder (in conjunction) leads to a chemical reaction that provides the perfect lift, the right amount of browning, and a balance between acid and base for a good-tasting cookie. This is a *conjunctural causation*: a factor may not be causally related to an outcome on its own but only related to the outcome in the presence of another factor.

Finally, the fat in the cookie recipe is useful for illustrating *asymmetrical causation*. All good-tasting cookies have some fat in the recipe. Using real butter leads to good-tasting cookies, but not using butter does not necessarily mean a cookie will not taste good. Many cookie recipes that result in a “good-tasting chocolate chip cookie” use margarine, shortening, or some alternative fat. This is *asymmetric causation*; even though the presence of an explanatory factor produces the outcome, one cannot assume that the absence of the factor leads to the *non-occurrence* of the outcome.

QCA is particularly well-suited to addressing configural research questions. Configural research questions are questions formulated to ask about the relationship of combinations of factors (i.e., conditions) with an outcome. A configural research question asks, “What combinations of conditions or factors are necessary or sufficient for an outcome?” For example, “Which combinations of new teacher training, supports, and policies are found among school districts with high rates of teacher retention?” All three concepts related to causal complexity are represented by configural questions.

QCA within mixed methods evaluation

Mixed methods study designs or evaluations include both qualitative and quantitative methods to complement each other for addressing specific research questions (Curry and Nunez-Smith, 2015; Plano Clark and Ivankova, 2016). Various mixed methods designs exist, and the specific design investigators select depends on the research questions. Because QCA is not a purely qualitative or quantitative method, it may not be traditionally included in the arsenal of methods to consider as part of a mixed methods design. However, it can be a useful approach to include in a mixed methods design when one of more of the research aims include configural research questions (Kahwati and Kane, 2020).

QCA can be used in such designs by (1) integrating the qualitative and quantitative data into a single analysis, (2) providing an analytic step between sequenced data collection, or by (3) supplementing a secondary point of integration between the qualitative and quantitative methods (Kahwati and Kane, 2020). Findings from the qualitative strand of a mixed methods study (e.g., case studies) can also be used to enhance the findings of a QCA by offering rich details about how the patterns identified by the QCA are exemplified in the cases. First, when developing conditions or an outcome for a QCA in a mixed methods study, one can transform qualitative data into numeric data; additionally, one can also apply substantive knowledge to measured quantitative variables. Second, QCA can be a part of a mixed methods study and, if needed, be used at multiple points in a mixed methods study. For example, a researcher could develop conditions based on qualitative data and use QCA to create typologies of cases that could be included in a statistical model. Alternately, the quantitative data could also be transformed into conditions for typology development through QCA, and then the researcher could use the different typologies to purposively select cases for qualitative data collection. Third, QCA can offer an additional analysis to support the qualitative and quantitative strands; a mixed methods study could have a research question with a configural aim, which would rely on QCA and merged or embedded integration.

Examples from the field

In this section we offer high-level summaries of published examples of QCA used to address configural research questions in the field of education research and closely related social science fields. These examples illustrate the use of QCA with different units of analysis as cases: countries, organizations, schools. Cases can also be individuals or defined based on discrete events (e.g., mass school shootings, labor strikes) or conceptually-bounded phenomenon (e.g., legislation, guidelines, programs, or standards).

- Assmann and colleagues used QCA to explore combinations of country-specific conditions related to differing rates of young people (age 15 to 29) not engaged in the labor force or in the education system (NEETs) (Assmann and Broschinski, 2021). The authors ran several models based on the reason that the young people were not engaged (this chapter provides one model as an example). Model 1 examined the factors related to high rates of NEETs who are actively seeking employment but cannot find work.
 - *Configural Question*: What combinations of country-specific conditions are necessary or sufficient for having high rates of NEETs who are actively seeking employment but cannot find work?
 - *Cases*: 26 European countries
 - *Conditions*: Gross Domestic Product, employment protection legislation for regular workers, employment protection legislation for temporary workers, expenditures on active labor market policies (activation policies), and work-based learning in the curriculum (vocational education).
 - *Outcomes*: Having a high rate of young people not engaged in the labor force or in the education system because they are unemployed but are actively seeking work.
 - *Findings*: The authors found 3 sufficient combinations for having a high rate of unemployed NEETs. They were (1) having high employment protections combined with a strong economic downturn, (2) having “high employment protection for permanent and temporary worker and the lack of activation policies, and (3) High employment protection for regular worker and the lack of activation policies and vocational education” (pp. 107–108).
 - *Conclusions*: The authors suggest that findings from this analysis offer some hypotheses to evaluate in future research and also highlight the need for country-specific measures to reduce the rate of disengaged young persons.
- Van Kessel and colleagues used QCA to identify combinations of conditions related to inclusive education for persons with autism among European Union countries (van Kessel et al., 2021). The conditions evaluated stem from the Disability Rights in Education Model, which provided the conceptual framework for the analysis.
 - *Configural Question*: What EU member state policies are necessary or sufficient for having a high level of access for children with special education needs (SEN)?
 - *Cases*: 20 EU member states
 - *Conditions*: having a definition of special education needs, having a policy for the right to education for children with SEN, having supported teaching staff, having supportive services for children with SEN, having inclusive education policies, having a high level of parental involvement, having individualized learning outcomes, and having mixed mainstream classes
 - *Outcome*: Having access to education for children with SEN
 - *Findings*: Mixed classrooms, that is, classrooms that serve children with and without special needs, and support services for children with special needs were necessary components for inclusive education. Necessary components are those that are “must have” but alone may not be sufficient for achieving inclusive education. Five combinations of conditions were identified as sufficient for inclusive education; sufficient combinations are those that when present result in the outcome of inclusive education. The combination with the most empirical relevance was having policies focused on inclusive education combined with a high level of parental involvement and no established definition for special education needs.
 - *Conclusions*: The authors suggest that achieving inclusive education is a complex process with key policy areas that can be used in future studies as the development of inclusive education evolves.
- Dill and colleagues used QCA to explore combinations of organizational partnership characteristics and their relationship to frontline healthcare workers’ perceptions of career ladder quality and rewards after completing a workforce education and training program (Goicolea et al., 2015). The program relied on partnerships between healthcare organizations, educational institutions, and other workforce-improvement organizations to establish career ladders and professional growth opportunities for frontline workers. The study team conducted an analysis for each outcome (one is presented here).
 - *Configural Question*: What dimensions of healthcare organization and educational partnerships are necessary or sufficient for achieving frontline worker outcomes in career ladder programs?
 - *Cases*: 11 career development programs involving healthcare organization and educational organization partnership
 - *Conditions*: Management support/leadership in the employer and educational organization (i.e., having a manager leader, education leader, and frontline management support), having a long history with of partnership between the organizations, having a significant community need for the program, and having supportive organizational policies and practices in the employer and educational organization
 - *Outcomes*: Career self-efficacy, program satisfaction, work change/promotion, and college credit/credential provided
 - *Findings*: For career self-efficacy, the team identified the combination of having management support from an educational leader in combination with management support from a frontline manager and having supportive educational-institution policies as a sufficient combination, as well as having management support from an educational leader in combination with the organizational partnership having a long-standing history of collaboration and having supportive educational-institution policies.
 - *Conclusions*: This finding pointed to the importance of leadership support, as leaders could encourage frontline worker participation and engagement in the career ladder programs, while also having supportive policies and history of partnership, that supported successful implementation of the program.

- Bandaranayake used QCA to identify combinations of factors that might explain low versus high performance among secondary schools in Victoria, Australia (Bandaranayake, 2016). The factors evaluated included geographic location, a socio-educational index representing family backgrounds in terms of parent occupation and education, average student attitude and parental satisfaction, proportion of families with language other than English spoken at home, and average per student government funding.
 - *Configural Question*: “Why do different groups of government schools continue to perform well below the average and at inferior levels while another group of government schools performs consecutively and persistently well above the average and at superior levels? What are the conditions or causal complexities that govern polarization and continuation of educational inequalities?” (p. 589)
 - *Cases*: 49 secondary schools in Victoria, Australia
 - *Conditions*: High level of community advantage, high level of having language background other than English at home, student positive attitude toward school, parents have high satisfaction with school, and high level of state funding for school
 - *Outcomes*: High achievement of Victorian Certificate of Education (VCE) performance and low achievement of VCE performance
- *Findings*: The author identified 3 combinations of factors associated with high-performing schools; the combination with the highest empirical relevance included a high socioeducational index in combination with location in a metro area in combination with positive student attitudes (i.e., engagement and enjoyment of school). The author identified 5 combinations of factors associated with low-performing schools; the combination with the most empirical relevance included a low socio-educational index combined with low parental satisfaction.
- *Conclusions*: The author suggests that no single factor alone results in high or low performing school, and that the pathway to high performance is not simply the opposite of the pathway to low performance—a nice example of asymmetric causality. Further because some of the factors explored are immutable by schools themselves, inequalities across schools are likely to persist without attention to measures to prevent polarization and close gaps.

Process for using QCA

QCA involves an iterative process (see Fig. 1) and starts with having a configural research question. Ideally, this configural question is drawn from existing theory, scholarship, or conceptual frameworks, as theory and substantive knowledge inform the next steps of case, condition, and outcome selection. (Considerations for case, condition, and outcome selection are described below.) In this step, the researcher defines what cases are included and how the conditions and outcome will be specified in the analysis. After identifying the cases, conditions, and outcome, the researcher then calibrates the values for the conditions and outcome (calibration is discussed below). Analysis of the conditions and outcomes occurs next; analysis includes identifying individual and combinations of necessary and sufficient conditions for the outcome (i.e., identifying causal complexity) and conducting robustness tests. At this stage, the researcher may need to revisit earlier decisions about case selection or definitions of conditions and outcomes, or the researcher may proceed with exploring the solutions using a range of techniques, such as process tracing or case studies.

Considerations for use

In this section we describe some things researchers should consider when deciding whether QCA is an appropriate method for use in their research or evaluation. These include theoretical, empirical, and practical considerations.

First and foremost, QCA is designed to address configural research questions. QCA may be a good fit to address configural research questions that are part of a larger mixed methods study. QCA is not a method for statistical hypothesis testing. Further some research questions may be better suited for qualitative research designs. When implementing QCA in a mixed method design, researchers should consider where and how a QCA fits into the overall study design. In general, QCA serves as a complementary method to qualitative and quantitative strands of a mixed method study. If the QCA is meant to integrate qualitative and quantitative data strands, then planning for not only the data sources needed, but also the timing of the analysis will ensure that adequate analysis time is established in the project timeline.

In QCA, the case is the unit of analysis. As such, QCA is ideally suited for use with between 10 and 50 cases so that researchers have enough familiarity with each case to guide important analytic decisions (Rihoux and Ragin, 2009). However, the sample or population size alone should not drive the decision to use QCA. Although QCA is challenging to use with fewer than 10 cases, it can be used in large-N analyses, though consideration as to whether it is a good fit for the research questions is important. The cases included in a QCA should be similar enough to be comparable to each other; yet, they should not be exactly the same on all dimensions of interest. Moreover, case selection depends upon the researcher's theoretical framework and research question; a practical consideration informing case selection (as well as condition and outcome selection) is the feasibility of obtaining data (see below). Importantly, to evaluate which combinations of conditions are found among cases with an outcome, it is important to have cases with and without the outcome in the analysis (Schneider and Wagemann, 2012; Kahwati and Kane, 2020).

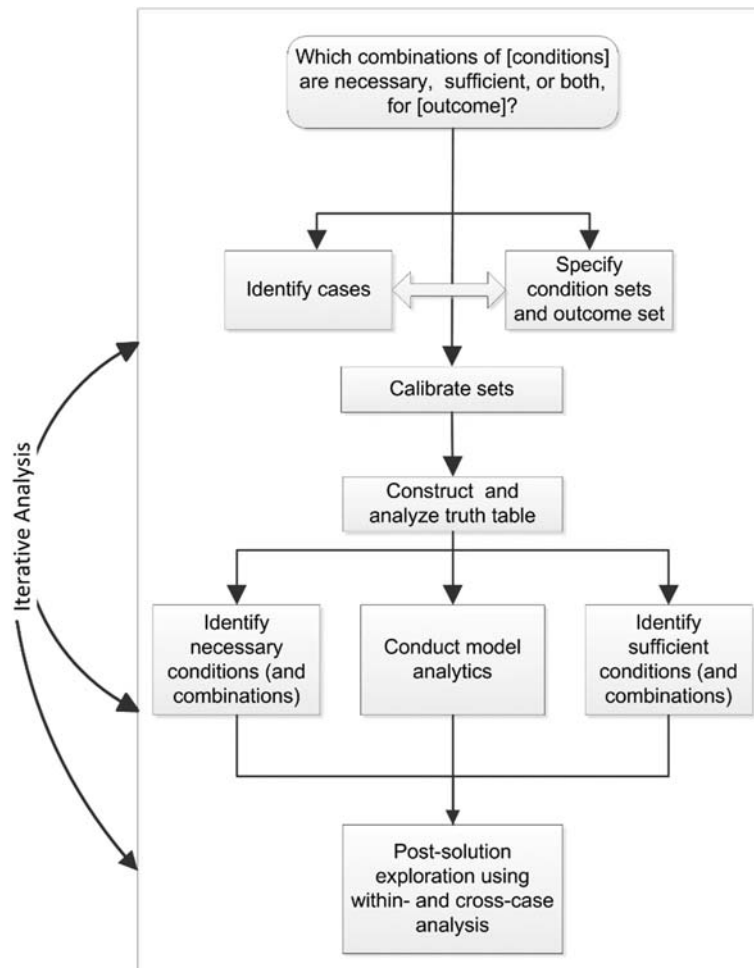


Fig. 1 Process for conducting a QCA Figure adapted from Kahwati LC, Kane HL. Qualitative comparative analysis in mixed methods research and evaluation. Vol 6: Sage Publications, 2020.

The availability and suitability of available data is a consideration. As with any method, if appropriate data sources are not available or have limited validity, then a QCA may not be feasible. When conducting a prospective study or evaluation, planning thoughtfully for a QCA at the outset of a data collection will facilitate use of the method, as researchers can ensure that relevant data are collected to support defining and assessing conditions and outcomes and that data sources are triangulated. However, because outcomes and potential variation is not known at the outset, one may find that all cases in the study achieved (or did not achieve) the outcome, which would limit use of QCA. When trying to implement a QCA after data collection has occurred, the researcher may face challenges in defining conditions and outcomes based solely on what information is readily available (Kahwati and Kane, 2020).

QCA uses calibration, rather than measurement, to assess conditions and outcomes (Schneider and Wagemann, 2012; Kahwati and Kane, 2020; Ragin, 2000, 2008). Calibration establishes the level of membership that a case has in a condition set or outcome set; it applies substantive and conceptual meaning to a measured value. Measurement of variables, in contrast, captures the degree to which a case has the property of the variable (Schneider and Wagemann, 2012). Ideally, calibration uses external standards to determine the level of membership (Schneider and Wagemann, 2010). For example, in the US secondary school system an "A" (excellent grade) is often determined by an average aggregated achievement level of 90% or more on evaluated work. The 90% or higher is a well-recognized standard for a high-level of achievement, and an individual scoring 90% or higher would have a high level of membership in the condition/outcome set of "high performance." A measured variable, for instance, would use the observed values of the achievement scores and apply measures of central tendency (i.e., mean, median, mode). Sometimes, however, external standards are not readily available. When that occurs, expert or substantive knowledge in the field can support decision making about level of membership in a set. Measures of central tendency (mean, median, mode) are not ideal for calibration because those are sample dependent (Schneider and Wagemann, 2010).

Use of software is essential when conducting a QCA. Several software programs are available; the two most commonly used are R's QCA and SetMethods modules and fsQCA (Ragin et al., 2008); each has benefits and limitations. Designed by Dusa (QCA

module (QCA)) and Oana and Schneider (SetMethods module (Oana and Schneider, 2018)), the R programs are updated regularly to stay abreast of ongoing enhancements to the method; however, R has a learning curve and can be frustrating to new users who are not familiar to working within the R environment. Designed by Ragin, fsQCA has a user-friendly interface, but has not been consistently updated and lacks some of the more advanced features contained in the R packages.

Like quantitative methods, a QCA may require model refinements and iterations, and, like qualitative, post-model analysis depends on delving into cases to make sense of the findings. Researchers, therefore, must plan use of time and financial resources to ensure that the QCA is implemented thoughtfully, that the model is thoroughly tested, and that the qualitative assessment of cases is comprehensive.

Because QCA is young and evolving as a method, many researchers and peer reviewers are unfamiliar with the method and may provide feedback on scholarly proposals or evaluations using QCA that are not relevant or appropriate. For example, more quantitatively-oriented researchers may comment on use of statistical controls for a model and question whether results are produced by chance; neither comment is relevant for QCA. Other reviewers may comment on QCA terminology (e.g., necessity, consistency) because they interpret the terms from the common usage perspective. Thus, when submitting a proposal or a manuscript for publication, one should strive to minimize use of jargon, offer explanatory analogies, use illustrative visuals (e.g., Venn diagrams), or create technical appendices and glossaries to aid reviewers' understanding.

Limitations

Although QCA offers a formalized and systematic approach to cross-case comparison, debate still exists in the research and evaluation communities about its use (Collier, 2014; Sartori, 2014). More quantitatively-oriented researchers question the robustness of QCA findings (Baumgartner and Thiem, 2015; Lucas and Szatrowski, 2014). However, QCA results are not meant to be statistically generalizable and can still describe patterns in data from the sample at hand. More qualitatively-oriented researchers question QCA's commitment to in-depth analysis of cases, which is the heart of the qualitative tradition (Thomann and Maggetti, 2017). Though some applications of QCA have not invested in the in-depth analysis of cases, "empirical intimacy" or deep knowledge of cases has been a part of the vision for QCA and is central to small- and intermediate-N applications of QCA (Rihoux and Ragin, 2009).

Concluding thoughts

QCA is still a relatively young method that is evolving. QCA experts continue to engage in ongoing dialogue with QCA's critics, making improvements to address limitations, and specifying new applications for QCA. Best practices associated with the use of QCA will continue to be refined in the upcoming years.

References

- Assmann, M.-L., Broschinski, S., 2021. Mapping young NEETs across Europe: exploring the institutional configurations promoting youth disengagement from education and employment. *J. Youth Stud.* 1–23.
- Bandaranayake, B., 2016. Polarisation of high-performing and low-performing secondary schools in Victoria, Australia: an analysis of causal complexities. *Aust. Educ. Res.* 43 (5), 587–606.
- Baumgartner, M., Thiem, A., 2015. Identifying complex causal dependencies in configurational data with coincidence analysis. *R Journal* 7 (1).
- Coburn, C., Russell, J., Kaufman, J., Stein, M., 2012. Supporting sustainability: teachers' advice networks and ambitious instructional reform. *Am. J. Educ.* 119 (1), 137–182.
- Collier, D., 2014. Problematic tools: introduction to symposium on set theory in social science. *Qualitative & Multi-Method Research* 12 (1), 2–9.
- Curry, L., Nunez-Smith, M., 2015. *Mixed Methods in Health Sciences Research: A Practical Primer*. Sage, Los Angeles, CA.
- Fiss, P.C., 2011. Building better causal theories: a fuzzy set approach to typologies in organizational research. *Acad. Manag. J.* 54 (2), 393–420.
- Frambach, R.T., Fiss, P.C., Ingenbleek, P.T.M., 2016. How important is customer orientation for firm performance? A fuzzy set analysis of orientations, strategies, and environments. *J. Bus. Res.* 69 (4), 1428–1436.
- Goicolea, I., Vives-Cases, C., Hurtig, A., et al., 2015. Mechanisms that trigger a good health-care response to intimate partner violence in Spain. Combining realist evaluation and qualitative comparative analysis approaches. *PLoS One* 10 (8), e0135167.
- Ide, T., 2015. Why do conflicts over scarce renewable resources turn violent? A qualitative comparative analysis. *Glob. Environ. Change* 33, 61–70.
- Kahwati, L.C., Kane, H.L., 2020. *Qualitative Comparative Analysis in Mixed Methods Research and Evaluation*, vol. 6. Sage Publications.
- Kahwati, L., Jacobs, S., Kane, H., Lewis, M., Viswanathan, M., Golin, C., 2016a. Using qualitative comparative analysis in a systematic review of a complex intervention. *Syst. Rev.* 5 (1), 82.
- Kahwati, L., Viswanathan, M., Golin, C., Kane, H., Lewis, M., Jacobs, S., 2016b. Identifying configurations of behavior change techniques in effective medication adherence interventions: a qualitative comparative analysis. *Syst. Rev.* 5 (1), 83.
- Kane, H., Hinnant, L., Day, K., et al., 2017. Pathways to program success: a qualitative comparative analysis (QCA) of communities putting prevention to work case study programs. *J. Publ. Health Manag. Pract.* 23 (2), 104–111.
- Kintz, T., Lane, J., Gotwals, A., Cisterna, D., 2015. Professional development at the local level: necessary and sufficient conditions for critical collegiality. *Teach. Teach. Educ.* 51, 121–136.
- Lucas, S., Szatrowski, A., 2014. Qualitative comparative analysis in critical perspective. *Socio. Methodol.* 44 (1), 1–79.
- Oana, I.-E., Schneider, C.Q., 2018. SetMethods: an add-on R package for advanced QCA. *R J* 10 (1), 507.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research. A Guide to the Field*. Sage, Los Angeles, CA.

- QCA: Qualitative Comparative Analysis [computer Program]. Version R Package Version 2.62017.
- Ragin, C.C., Strand, S.I., Rubinson, C., 2008. User's guide to fuzzy-set/qualitative comparative analysis. University of Arizona 87.
- Ragin, C., 1999. Using qualitative comparative analysis to study causal complexity. *Health Serv. Res.* 34 (5 Pt 2), 1225–1239.
- Ragin, C., 2000. *Fuzzy-Set Social Science*. University of Chicago Press.
- Ragin, C., 2006. Set relations in social research: Evaluating their consistency and coverage. *Polit. Anal.* 14, 291–310.
- Ragin, C., 2008. *Redesigning Social Inquiry: Fuzzy Sets and beyond*. University of Chicago Press, Chicago, IL.
- Rihoux, B., Ragin, C., 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Sage, Thousand Oaks, CA.
- Sartori, G., 2014. Logic and set theory: a note of dissent. *Qualitative & Multi-Method Research* 12 (1), 14–15.
- Schneider, C., Wagemann, C., 2010. Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comp. Sociol.* 9 (3), 397–418.
- Schneider, C., Wagemann, C., 2012. *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. Cambridge University Press, Cambridge, United Kingdom.
- Thomann, E., Maggetti, M., 2017. Designing research with qualitative comparative analysis (QCA) approaches, challenges, and tools. *Socio. Methods Res.* 0049124117729700.
- Thygeson, M., Solberg, L., Asche, S., Fontaine, P., Gregory Pawlson, L., Scholle, S., 2012. Using fuzzy set qualitative comparative analysis (fs/QCA) to explore the relationship between medical "homeness" and quality. *Health Serv. Res.* 47 (1 Pt 1), 22–45.
- van Kessel, R., Hrzic, R., Cassidy, S., et al., 2021. Inclusive education in the European Union: a fuzzy-set qualitative comparative analysis of education policy for autism. *Soc. Work. Publ. Health* 36 (2), 286–299.
- Verma, P.O., 2012. *Linking Policy Process to Environmental Impacts in the United States Environmental Protection Agency's National Estuary Program: A Comparative Case Study Analysis of the Science-Policy Interface*.

Intersecting mixed methods with translational research: implications for educational research and practice

Nataliya V. Ivankova^a, Jami L. Anderson^a, Ivan I. Herbey^b, Linda Roussel^c, and Daniel Kim^d, ^a School of Health Professions, The University of Alabama at Birmingham, Birmingham AL, USA; ^b School of Medicine, The University of Alabama at Birmingham; ^c Cizik School of Nursing at UTHealth, The University of Texas Health Science Center; and ^d College of Business, Idaho State University, Pocatello, ID

© 2023 Elsevier Ltd. All rights reserved.

Introduction	589
Translational research and translational science	589
Methodological components of the translational research process	590
Evidence-based practice	590
Adaptation	591
Dissemination and implementation	592
Community-based participatory research	593
Action research	593
Conceptual model for intersecting mixed methods with translational research	594
Advantages and challenges of intersecting mixed methods with translational research: implications for educational research and practice	596
Conclusion	597
Acknowledgment	597
References	597

Glossary

Action research: “A participatory process concerned with developing practical knowing in the pursuit of worthwhile human purposes. It seeks to bring together action and reflection, theory and practice, in participation with others, in the pursuit of practical solutions to issues of pressing concern to people” (Reason and Bradbury, 2008, p. 4)

Adaptation: The degree to which an innovation is undergoing adjustments or modifications in the process of its adoption and implementation in real-life settings (Rogers, 1995).

Community-based participatory research: “A systematic inquiry, with the collaboration of those affected by the issue being studied, for purposes of education and taking action or effecting change” (Green and Mercer, 2001, p. 1927).

Dissemination and implementation research: Active approaches and planned research strategies to promote distribution and use of evidence-based interventions designed for specific audiences or populations (Rabin et al., 2008, National Institutes of Health, 2021).

Evidence-based practice: Objective, balanced, and responsible use of relevant research and best current data to guide scientifically based practice and policy decisions, to improve outcomes for consumers (Crime and Justice Institute at Community Resources for Justice, 2009).

Intersecting: Intentional combination of two or more methodological approaches within a study design (Plano Clark and Ivankova, 2016).

Knowledge translation: “A dynamic and iterative process that includes the synthesis, dissemination, exchange and ethically sound application of knowledge” in real-world settings (Straus et al., 2009, p. 165)

Translational research: Research aimed at moving knowledge and discovery obtained from the basic sciences to its application in clinical and community settings (Khoury et al., 2007).

Translational research continuum: Phases or stages in translational process corresponding to four knowledge translation concepts: translating knowledge from the basic sciences to humans (T1 research); translating the research findings to develop new treatments and programs (T2 research); translating evidence-based guidelines into everyday practice (T3 research); and translation to the community and real-world settings (T4 research) (Fort et al., 2017).

Translational science: The field of investigation which proposes to understand the scientific and operational principles which underscore each step and phase of the translational process (Collins, 2011).

Introduction

In recent years, there has been a tremendous growth in the use of mixed methods research across multiple disciplines (Coyle et al., 2016; Creswell and Plano Clark, 2018; Tashakkori et al., 2021). Mixed methods research is an approach where quantitative and qualitative methods are intentionally mixed during data collection, analysis, and interpretation within a study process (Creswell and Plano Clark, 2018). Thoughtful and intentional mixing of methods, data and results within one study allows researchers to examine the research landscape more widely from various lenses and viewpoints (Shorten and Smith, 2017). Moreover, a mixed methods approach allows investigators to explore complex issues and generate more complete and robust conclusions due to its notable methodological flexibility to intersect with other research approaches and designs (Creswell and Plano Clark, 2018; Nastasi and Hitchcock, 2016; Plano Clark and Ivankova, 2016).

Mixed methods can optimize the knowledge translation process and facilitate adoption of research findings into practice through adaptation, dissemination, implementation, and evaluation of complex interventions when intersected or meaningfully combined with translational research (Ivankova et al., 2018). Knowledge translation is defined as a multiphase and “dynamic process including the synthesis, dissemination, exchange, and ethically sound application of knowledge” into real-world settings through the use of translational research (Straus et al., 2009, p. 165). While there are different definitions of translational research, the most common definition is a type of research examining and gathering information aimed at moving knowledge and discovery obtained from the basic sciences to its application in clinical and community settings (Mace and Critchfield, 2010; Khoury et al., 2007). The focus of translational research is on translating and transforming knowledge. Knowledge synthesis is a critical component in the translational research process. Effective knowledge translation requires multiple methods and transdisciplinary approaches to generate results that are applicable in real time and settings (Khoury et al., 2010; Woods and Magyary, 2010). Using multiple methods and a team science approach helps researchers to incorporate best practices and produce solutions that have the greatest effect on people’s lives (Taylor et al., 2016).

The critical role of mixed methods in translating knowledge has been well-recognized (Creswell et al., 2018; Green et al., 2015; Ivankova et al., 2018; Meister, 2018; Palinkas and Cooper, 2017). Some noted advantages of using mixed methods within the field of translational research include: (1) integrating qualitative stakeholder engagement methods with quantitative outcome-oriented approaches; (2) examining the issue at individual and system levels and from multiple perspectives; and (3) providing an integrative thinking approach to addressing translation research questions (Martin, 2007). By intersecting mixed methods with translational research, educational researchers can increase the breadth and depth of understanding the pertinent issues related to knowledge translation and adoption of the evidenced-based practices in educational settings (Johnson et al., 2007).

This chapter offers a perspective on intersecting mixed methods with translational research. In the following sections we describe translational research and its integral components – evidence-based practice, adaptation, dissemination and implementation, and community-based participatory research and action research – and discuss how their intersection with mixed methods research provides a more methodologically robust approach to facilitate the knowledge translation process. We conclude the chapter with discussing advantages and challenges of using mixed methods in translational research and their implications for educational research and practice.

Translational research and translational science

Translational research emerged in the medical field stimulated by concerns over the extended time it takes to incorporate scientific discoveries in treatments, practices, and health policies (Wethington and Burgio, 2015). Translational research gained increased awareness after the National Institutes of Health (NIH) prioritized it in 2006 to bridge the gap between research and practice. Since then, it has been widely used across disciplines including education. A goal of translational research is to identify, support and promote sustainable adoption of evidenced-based practices. Using translational science, researchers integrate evidence-based practices, clinical experiences, and the individuals’ preferences to produce outcomes that impact population well-being (Taylor et al., 2016).

Translational research and translational science have been used interchangeably to describe the process of moving knowledge and discovery gained from the basic sciences to its application in professional and community settings. However, these terms differ in several ways (Austin, 2021). Zoellner and Porter (2017) describe translational research as a bidirectional process which purports to advance scientific discovery for patient and community benefit and involves multidisciplinary integration among basic, clinical, practice, population, and policy-based research. Translation is a practical multi-step process and is the focus of translational science. The latter is defined as the field of investigation which proposes to understand the scientific and operational principles which underscore each step and phase of the translational process (Collins, 2011). The intent of translational science is to advance the process of translation with results which can have a broader impact by utilizing effective research methods including mixed methods research.

Knowledge translation is a complex process consisting of multiple stages. The translational research continuum most commonly embraces four knowledge translation concepts: translating knowledge from the basic sciences (i.e. bench research including disciplines such as biochemistry, microbiology, or pharmacology) to humans, referred to as T1 research; translating research findings to develop new evidence-based treatments, programs and guidelines, referred to as T2 research; translating evidence-based guidelines into professional practice, referred to as T3 research; and translating evidence-based practices for adoption in the community and real-world settings, referred to as T4 research (Fort et al., 2017). While each stage builds upon and informs other stages, the

knowledge translation process is dynamic and iterative and includes synthesis, adaptation and dissemination of knowledge with continuous input from stakeholders (Callard et al., 2011; Straus et al., 2009).

Methodological components of the translational research process

Translational research is a multifaceted approach and incorporates different methodologies and approaches that contribute to the process of translating research findings into practice. The following five approaches, further referred to as the methodological components of the translational research process — evidence-based practice (EBP), adaptation, dissemination and implementation (D&I), community-based participatory research (CBPR), and action research — are integral to the knowledge translation process as each plays a specific role in facilitating translation of research findings to real-world applications (Ivankova et al., 2018). However, while pursuing different goals, they interact and inform each other in a joint effort to promote and facilitate the translation process.

In addition to having unique methodological characteristics, these five components employ mixed methods research by meaningfully integrating quantitative and qualitative methods to enhance and achieve knowledge translation goals. In the following sections, we describe these components and discuss how their intersection with mixed methods research enhances the knowledge translation process. We illustrate this intersection with examples of published studies that used mixed methods in translational research. We also draw on the perspectives of translational science researchers who apply mixed methods in their research practice and present these perspectives throughout the chapter. We collected those perspectives as part of the study aimed at understanding how mixed methods research can optimize translational science. **Box 1** provides a study overview including the study purpose, methods, and synopsis of the results.

Evidence-based practice

EBP is rooted in medicine with Archie Cochrane, a British epidemiologist, rebuking the medical community for their failure to critically examine evidence (Cochrane, 1972). Fundamental tenets of evidence-based medicine (EBM) include strategic awareness and critical appraisal of evidence to inform trustworthy and accurate decision making. EBP has been described as practice approach based on the best available evidence, patient/client preferences, and clinical judgment (Melnyk and Fineout-Overholt, 2019). In 2005, the Institute of Medicine (IOM) identified EBP as one of the core competencies for healthcare professionals. EBP is foundational to translational research and science.

EBP has been expanded from EBM to apply to other health and social science disciplines including education. The concept of evidence-based learning and teaching is complex and multifaceted. It is concerned with connecting learning and teaching to evidence-based methods, strategies and processes, through nurturing a systemic institutional culture invested in continuous improvement of student learning (Othie, 2018; Cranney and McDonald, 2012). A participant from the study we introduced in described the importance of EBP as the following:

Basically, the activities used to move innovations, bench science discoveries, policy innovations, you know, any sort of new significant event, ideally that is evidence-based into practice. Again, that kind of encompasses both bench science discoveries

Box 1 Study Overview

Title: Optimizing Translational Science: Role and Value of Mixed Methods Research.

Purpose: To understand how mixed methods research can optimize translational science by exploring the views and real-world experiences of translational researchers who apply mixed methods along the knowledge translation continuum.

Sample: Using purposeful and snowball sampling, we recruited 35 translational science researchers from two countries: US (n = 30) and Canada (n = 5).

Data Collection: Semi-structured interviews were conducted and recorded using Zoom web conferencing. The interview questions focused on participants' experiences with and views on applying mixed methods across the knowledge translation continuum from T1 to T4, and how mixed methods can optimize translational research through its different domains (evidence-based practice, adaptation, dissemination and implementation, community-based participatory research and action research). We also explored participants' perspectives on what researchers should be aware of when implementing integrated mixed methods studies in the translational research context. We also conducted follow up member checking interviews with 80% of the participants via Zoom and email.

Data Analysis: Interview transcripts were analyzed using inductive thematic and qualitative content analysis approaches with NVivo 12.6 Plus by several coders with regular input from the research team.

Results: Participants represented 10 disciplines including 8 participants from the education field. The majority (n = 31) served as faculty at academic institutions, three represented non-profit research organizations, and one was a program officer from the National Institutes of Health. Two-thirds were in senior ranks of full or associate professor (n = 22). Participants indicated that there are considerable advantages to using mixed methods in translational research; namely, the rigorous methodology of mixed methods, the potential to conduct practically focused outcomes research, and relevance of mixed methods across the translational spectrum. Participants also noted multiple challenges of incorporating mixed methods into the translational research spectrum. In particular, the time required to complete mixed methods research and the resources to ensure rigorous methodologies were often noted amongst participants as challenges of incorporating mixed methods into the translational research spectrum.

that eventually get translated into practical interventions; as well as, how ... we get those interventions to be adopted and utilized by practitioners and patients and so forth.

According to [Wisdom and Creswell \(2013\)](#), using a mixed methods approach can provide researchers with a deeper understanding of interactions, connections, or contradictions between qualitative and quantitative data. Incorporating the participants' voices, in which they share experiences of the process, can facilitate diverse means of exploration, enhancing the evidence, and providing opportunity for questions to be answered more thoroughly. This is essential to EBP. However, a robust process of using mixed methods can also add to the complexity of carrying out research that often requires more time, resources, and research training. Use of a multidisciplinary research team is critical to all aspects including sample selection, data analysis and synthesis or integration ([Wisdom and Creswell, 2013](#)). Our study participant explained:

... You need to have the personal experiences of people as well as the measurable quantitative outcomes. To really understand the real-world application, we need qualitative data, and we also need the quantitative data. And being able to put the two together using mixed methods designs ... would give us more insight than just using the quantitative results or the qualitative results.

[Aarons et al. \(2012\)](#) used mixed methods to examine the impact of EBP implementation and fidelity monitoring on staff retention in the context of an effectiveness trial of a statewide child welfare system change and EBP. Data used in the study were collected across a three-year period as part of a larger five-year longitudinal mixed methods project examining organizational factors influencing statewide implementation of an EBP throughout a network of private nonprofit community based mental health and social service provider organizations. The researchers described the unique aspect of their study as 2×2 experimental designs, in which the EBP versus services as usual (SAU) was crossed with the level of fidelity monitoring (monitored vs. non-monitored). There were 21 teams of home-based service providers operating in six regions covering an entire state, with approximately one-quarter of the teams operating in each study condition. For quantitative analyses, researchers examined both the independent effects of EBP and fidelity monitoring (framed as ongoing consultation for service providers) on provider turnover as well as the interaction of EBP and monitoring conditions on provider turnover.

Data were collected longitudinally using sequential and concurrent mixed methods designs. Quantitative data consisted of bi-annual web-based surveys of 153 home-based service providers and qualitative data included semi-structured interviews and focus groups with multiple stakeholders at three different intervals over a 3-year period. The results were interpreted looking for convergence and complementarity in both data sets and using expansion to explain unexpected results in one data set by the results from an alternative data set. Overall findings noted that EBP implementation has the potential to improve workforce outcomes as well as clinical outcomes. EBP implementation generally entails several costs including training, materials, fidelity monitoring and clinicians time away from clinical duties. Their findings suggested that policy makers and organization leaders consider these benefits, in addition to those of EBP utilization, in making decisions about EBP implementation.

Adaptation

Adaptation is defined as "the degree to which an innovation is undergoing adjustments or modifications in the process of its adoption and implementation in real-life settings" ([Rogers, 1995](#), p. 149). The concept of adaptation was advanced by [Rogers \(1995\)](#) as part of his Diffusion of Innovations Theory. This theory attempts to explain how a novel product, practice, or idea get adopted by different groups of people and focuses on assessing different aspects of the innovation, such as its complexity, relative advantage, compatibility with existing interventions, evidence of the positive effect and the readiness to adopt the innovation.

Since the innovation is not easily accepted by different people, adaptations of an innovation can fall under the following reasons ([McKleroy et al., 2006](#)), including the need to: (1) simplify a complex innovation; (2) address a lack of knowledge about the innovation; (3) focus in on a specific problem or expanding to other problems; (4) increase stakeholders' ownership of the innovation; (5) address the needs of more diverse population; (6) choose an appropriate form of innovation delivery; (7) respond to a change initiated by agency; and (8) fit the available resources. While modifications of the innovation during the adaptation process are necessary to increase the innovation uptake by different people, it is important to identify and preserve core components accounting for the quality and integrity of the intervention. Moreover, any adaptation process requires contextual knowledge of the target population's experience with the issue to be responsive to cultural practices and fit the worldviews of individuals involved ([Wingood and DiClemente, 2008](#); [Resnicow et al., 2000](#)).

Adaptation plays a crucial role in translating evidence-based research into practice, including education ([Chambers and Norton, 2016](#); [Brownson et al., 2012](#)). More specifically, adaptation can improve the fit of a certain intervention or educational program to a specific group of people by taking political, environmental, sociocultural, and economic factors into account ([Allen et al., 2012](#); [Chambers and Norton, 2016](#)). For example, groups in charge of the delivery of the innovation, characteristics of a specific population, unique characteristics of an organizational setting and resources are all factors that can immensely affect how the innovation gets adopted by a particular group in a particular community ([Cohen et al., 2008](#)).

Mixed methods research plays an integral role in the adaptation process because, for an adapted innovation to be successfully integrated and implemented, it must be viewed in multiple ways. One of our study participants observed, "A key issue for adaptation is understanding the context that you are trying to adapt to. I think mixed methods allows researchers to gather a better understanding of a particular context. I don't see how you can do an adequate job of that without conducting mixed methods research because you can't get those multiple viewpoints in any other way." Mixed methods research is particularly important when considering the culture and preferences of the population of interest. A study participant explained it in the following way: "Mixed methods can inform the content of the program, can inform cultural adaptations of both content and the delivery, can inform

adaptations that may be required to increase the feasibility of the program for the target population. It may also provide insight into evaluation outcomes that may be important to the stakeholders who you would like for them to continue to do the program. It can also give you ideas about how to improve your program."

Baydala et al. (2014) used mixed methods research to examine the cultural adaptation of the school-based drug and alcohol abuse prevention program for an Aboriginal community in Central Alberta, Canada. In their study, researchers worked with community members to utilize the Life Skills Training Program, which emphasized cultural perspectives, community values, and measurable short- and long-term outcomes. Several stakeholders such as community representatives, school personnel, and the Elders were included in the adaptation process. Stakeholders had weekly meetings to discuss the adaptation process, ensuring the process reflected the culture of the community and the school curriculum. Each stakeholder brought different perspectives in the meeting. For example, the Elders brought in traditional knowledge on appropriate language to utilize, cultural inclusion and practices; whereas, community members oversaw recording information and assisting the Elders commute back and forth from their residences.

The program was provided to elementary and junior high students at the Alexis Nakota Sioux Nation School and a mixed methods research approach was used to evaluate the impact of the program. A longitudinal survey of students was conducted to examine program fidelity and effect on the community. Seventeen students from grade 3 participated in the survey. Additionally, university researchers conducted several interviews throughout a 2-year span with various stakeholders to gather their perspectives on the effectiveness of the program. Lastly, data from the program implementation were collected and analyzed quantitatively to integrate with the qualitative interview findings. The integrated results demonstrated a positive effect of the program adaptation and implementation.

Dissemination and implementation

D&I research are active approaches and planned research strategies to promote distribution and use of evidence-based interventions designed for specific audiences or populations (Rabin et al., 2008; National Institutes of Health, 2021). Dissemination research can be defined as "an active approach of spreading evidence-based interventions to the target audience via determined channels using planned strategies" (Rabin et al., 2008, p. 118). Alternatively, implementation research is a "systematic inquiry regarding innovations enacted in controlled settings or in ordinary practice, the factors that influence innovation enactment, and the relationships between innovation, influential factors, and outcomes" (Century and Cassata, 2016, p. 170). As there is considerable overlap between the two concepts, many investigators and educators utilize the term 'D&I research' as an umbrella term for research focusing on both dissemination and implementation of an intervention (Rabin et al., 2008; National Institutes of Health, 2021).

D&I research is complex and requires an understanding of D&I research terminologies, theories, designs, analytic approaches, and practice-based considerations (Padek et al., 2015). Engaging multiple stakeholders, ensuring representation throughout the community, and building common goals are often required for successful D&I of an intervention (de Visser et al., 2020; Knoepke et al., 2019). Considering these complexities, specific and mentored training is often required to learn D&I approaches as these researchers use varying combinations of quantitative, qualitative, and mixed methods research to achieve intervention dissemination and implementation (Padek et al., 2018). As one of our participants noted, "When I started to work with translation and dissemination research, I realized that without qualitative methods, the quantitative data alone was not rich enough to understand processes." Notably, educators and training developers have created a list of educational competencies for D&I researchers and have outlined critical training components for mentored research in D&I to support these endeavors (Padek et al., 2015, 2018). Additionally, researchers developing a tool supporting systematic approaches for designing implementation research used multiple methods including qualitative methods, quantitative methods, and literature reviews to consolidate components of effective implementation research into a feasible tool and supplementary guide for conducting complex D&I research (Hull et al., 2019).

While the processes of D&I research are complex, the value of D&I approaches using rigorous research methods are highlighted in successful implementation and sustainability of education programs and interventions (Moir, 2018). For example, a study in Uganda found that a comprehensive extracurricular school-based AIDS education program was poorly implemented and had very little effect on improving students' knowledge about HIV and AIDS even after the program had been in place over a year (Kinsman et al., 2001). Employing a mixed methods evaluation, in which quantitative data were collected from 20 schools participating in the intervention and qualitative data were collected from focus groups among students within 5 of these schools, evaluators concluded that the program was not successful due to limitations in classroom time, fear of controversy, and inadequate training among teachers to implement the program. The use of mixed methods in this study was beneficial in illuminating barriers to successful implementation of the HIV education program; the investigators also noted that successful future implementation efforts were contingent on incorporating support for the educators and developing training programs to ensure the programs are implemented as intended.

In addition to identifying barriers to successful implementation of interventions or programs at a local level, mixed methods research can support implementation and evaluation of programs with wider dissemination efforts. As one of our participants noted about translational research, "It's really a global perspective, you have technology and you have society and want to have them work together, and you need to have ways of answering questions that address both the natural phenomena and the constraints we live in." For example, investigators developed, tested, and refined a harm reduction educational program to reduce HIV and hepatitis C transmission among people who inject drugs in France (Roux et al., 2021). The investigators disseminated the program to four other countries including Bulgaria, Greece, Portugal, and Romania to test the effectiveness of the program after adaptations were made for

different settings. Utilizing a multi-phased mixed methods implementation study design following the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) Qualitative Evaluation for Systematic Translation (QuEST) framework, the investigators were able to identify perceived benefits of the program, barriers to implementation, and the value of implementing educational interventions for people in Europe who inject drugs.

Community-based participatory research

CBPR is defined as a “systematic inquiry, with the collaboration of those affected by the issue being studied, for purposes of education and taking action or effecting change” (Green and Mercer, 2001, p. 1927). As a collaborative research approach, CBPR equitably involves all community members including recipients of interventions, practitioners, service agencies, community leaders, policy-makers, and other stakeholders in all phases of the research process (Isreal et al., 2001). CBPR is based on the idea that understanding the culture and epistemology of a community is key to the inquiry process (DeJonckheere et al., 2019; Lucero et al., 2018); thus, requiring partnerships between researchers and community representatives to design and implement studies reflecting community values and needs. Additionally, engagement of stakeholders in CBPR research designs enhances opportunities to address advocacy issues for marginalized groups and promotes inclusion of multiple voices (Ivankova and Wingo, 2022 (in press)). The CBPR process relies on nine core principles (Isreal et al., 1998, 2001), which were originally developed for public health but have been extensively applied by researchers across disciplines:

- recognizing community as a unit of identity;
- building on strengths and resources within the community;
- facilitating collaborative, equitable involvement of all partners in all phases of the research;
- integrating knowledge and action for mutual benefit of all partners;
- promoting a co-learning and empowering process that attends to social inequalities;
- involving a cyclical and iterative process;
- addressing health from both positive and ecological perspectives;
- disseminating findings and knowledge gained to all partners; and
- involving a long-term commitment by all partners.

Additionally, the likelihood that research findings grounded in evidence-based practices will be accepted and adopted by communities increases by fostering equitable partnerships, providing capacity building, promoting co-learning, and engaging community representatives from multiple settings in shared decision making. A study participant observed, “... CBPR is clearly an appropriate methodology for engaging communities to embrace and try to promote uptake for adoption of an intervention, because they’re the key stakeholders that are going to be able to say, why or why not this will work, and when. It’s why it will work, but how, here’s how it’s got to be delivered here.”

There are multiple advantages of intersecting mixed methods with CBPR in the context of translational research (DeJonckheere et al., 2019, Ivankova and Wingo, 2022 (in press)). Integrating quantitative and qualitative methods in a CBPR study design allows investigators to develop, implement, evaluate, and promote interventions that are effective and transferable to other communities and practices. A study participant explained: “... we look at theory and research to help us develop culturally informed, applicable research questions. And then how we develop these interventions that we can modify based on the cultural, community and contextual needs, and then implement those and then translate those into research and practice and policy.” By capitalizing on qualitative methods as part of mixed methods designs with CBPR, researchers can develop a better understanding of community culture and epistemology and enhance community members’ engagement with the intervention process and results.

Bogar et al. (2017) used mixed methods in a CBPR study to assess urban youth’s knowledge of and attitudes toward lead poisoning and other environmental health concerns. The researchers created a CBPR partnership with a doctoral candidate, an education coordinator, and youth advisory council consisting of eight young people ages 10–18 at a local youth organization. This partnership helped with a mixed methods study design, interpretation of the quantitative, qualitative, and mixed results and dissemination of the study findings. Engagement with young members of the community as co-researchers allowed investigators to identify and explore environmental health issues specific to the community, and to develop strategies for increasing environmental health literacy among urban youth attending after school community organization programming.

Data were collected through a survey of 105 youths and 12 focus groups with 64 young people. The survey aimed to assess knowledge and attitudes toward lead poisoning and focus groups further contextualized youth’s perceptions and understanding of the environmental concerns to inform tailored prevention education programs for young people. Integrated conclusions emphasized the importance of considering a local context and considering relevance of an educational intervention based on the social and economic factors in a community.

Action research

The concept of action research was originally introduced by Kurt Lewin to describe a research approach underscoring a practical focus, community-based orientation, change action, collaboration between participants and researchers, participant empowerment and importance of reflection (Lewin, 1948). As a systematic and cyclical approach that incorporates both quantitative and qualitative methods, action research aims to identify an action or intervention that will generate improvement and produce meaningful

change for those who have a vested interest in the issue (Hacker, 2013; Ivankova, 2015; Koshy et al., 2011; McNiff and Whitehead, 2011).

Action research consists of four iterative stages of reflecting, planning, acting, and observing and includes six methodological steps: identifying a problem, reconnaissance or fact finding, planning action, acting, evaluating action, and monitoring action (Lewin, 1948). Action or intervention is central to the idea of action research and requires a spiral of action cycles aimed at developing, implementing and evaluating action/intervention (Kemmis, 1982). Observing and reflecting on the effects of action/intervention and using this information for planning subsequent action cycles in collaboration with stakeholders provides an evidence-based approach to “jump-start” the process of translation and adoption of generated evidence into practice (Hacker, 2013, p. 103). A study participant observed, “From a translational research point of view you might be operating with a wonderful advantage because ... folks start to recognize the problem that they want to work on, that they wanna effect change, versus one where you’re going in and telling them that there’s a difficulty.” Engaging stakeholders in the process of change action through partnerships and active participation throughout the study process provides opportunities for creating more tailored and meaningful solutions that have better chances to be adopted in real-world settings. A study participant explained,

It’s going forward with a research question and making certain that all aspects are understood. And it can come from early input from stakeholders in terms of qualitative input, what their priorities are, what outcomes are important to them, all the way through sitting down and saying here’s what we found. Does it resonate with you? Do you see something else in the results that we can draw conclusions that would give us actionable items on?

A mixed methods approach has long been applied in action research due to its focus on finding comprehensive solutions to complex practical issues requiring the use of multiple sources of quantitative and qualitative evidence (Creswell, 2012; Ivankova, 2015; Koshy et al., 2011; Mills, 2011). The synergistic combination of mixed methods and action research allows investigators to address practical issues in a systematic and dialectic way by using multiple methods, drawing on multiple stakeholders’ perspectives, and producing more credible and valid results (Ivankova and Wingo, 2018).

Breimaier et al. (2015) used a participatory action research approach combined with mixed methods to identify effective strategies for the implementation of evidence-based guidelines for preventing falls among older patients by nurses in an Austrian teaching hospital. In Breimaier et al.’s (2015) study, a mixed methods approach was used to inform all phases in the action research cycle. In the diagnosing phase, researchers identified the problem that required exploration by using both mixed methods and action research to enhance understanding and buy-in of nurses to adopt and actively implement the tailored fall prevention guidelines for their patients. During the reconnaissance phase, quantitative survey and qualitative interview and focus group data were collected and analyzed to assess nurses’ knowledge and attitudes to the guidelines implementation and to explore potential facilitators and barriers both at the individual and organization levels. The quantitative and qualitative results were combined and compared to inform the selection of the implementation strategies tailored to nurses’ and patients’ needs and to develop the content of the educational intervention that would promote patient safety by reducing patient falls and their negative outcomes.

The intervention consisted of educational meetings with nurses, inclusion of local opinion leaders, compilation of informational folders, audit and adaptation of nursing record systems, and improvements in physical structure, facilities and equipment. The results of the educational intervention were evaluated twice using a concurrent mixed methods design at mid-term and at the project completion. The same quantitative survey instrument was used, and the same number of nurses were interviewed as in the reconnaissance phase to establish consistency in the evaluation of the intervention outcomes. The integrated evaluation results provided evidence about the effectiveness of the tailored implementation strategies, revealed potential facilitators and barriers both at the individual and organization levels, and outlined directions for further dissemination, implementation, and monitoring of the evidence-based generated practices. As a result of the study, nurses better understood the importance of creating a safer environment for patients and felt self-confident and empowered to explain the guidelines to their patients.

Conceptual model for intersecting mixed methods with translational research

As we have shown, the five methodological components of the translational research process play an important role in and provide a unique contribution to the process of translating research findings into practice. When meaningfully combined or intersected with mixed methods research, these components jointly expand the scope of knowledge translation and add rigor to the assessment of the translational research process and outcomes, thereby providing a way of integrative thinking for addressing complex problems.

Indeed, like Rothman’s causal pies in disease epidemiology (Rothman and Greenland, 2005), these translational research components are not mutually exclusive. Successful dissemination and implementation of research into practice depends on the nature of provided evidence, the contextual environment, and the mechanism by which the process is implemented in real-life settings (Kitson et al., 1998); whereas community and stakeholder engagement is critical for ensuring the evidence-based innovation is accepted and adopted. The process of innovation adaptation is integrated into the dissemination and implementation process, relying on stakeholders’ involvement and active participation in all stages using community-based participatory research methods (Allen et al., 2012). Action research, with its experiential learning and dynamic nature, focuses on the issues relevant to the community and provides a foundation for community-based participatory research (Faridi et al., 2007). The knowledge translation process is often iterative and involves cycles of observing, reflecting, planning, and acting to ensure that evidence-based generated findings are reflective of stakeholders’ views and experiences and are aimed at benefiting those affected by the issue (McNiff and Whitehead, 2011).

Ivankova et al. (2018) have advanced a conceptual model for intersecting mixed methods with translational research, which underscores the interaction of EBP, adaptation, D&I, CBPR, and action research with mixed methods research and with each other (Fig. 1). Mixed methods research takes a central position because it intersects with all five components by providing methodological support through a meaningful integration of the quantitative and qualitative approaches during the study design and implementation.

For example, Aarons et al. (2012) intersecting mixed methods with EBP to identify factors influencing staff turnover during a large-scale EBP intervention implementation in a state child welfare system. Using multiple types of quantitative and qualitative data in a sequential design, researchers were able to obtain convergent and complementary results to better understand the implementation context and the EBP fit in the organization; thereby enhancing knowledge translation. According to researchers, both quantitative and qualitative data were necessary to examine the EBP implementation process and its effects on staff turnover, while adding increased validity to the integrated study conclusions.

Baydala et al. (2014) used mixed methods research to evaluate the effectiveness of the culturally adapted school-based drug and alcohol abuse prevention program for Aboriginal youth. Using a CBPR approach, researchers engaged community members in the process of cultural program adaptation emphasizing community values, cultural perspectives, and measurable short- and long-term outcomes. To assess the program delivery and impact, qualitative data were collected over several years and included focus groups with school and community stakeholders and individual interviews with the elder members of the community. Quantitative data included pre- and post-assessment of students' health knowledge, attitudes, drug refusal skills, and self-esteem in the context of a sequential longitudinal cohort design. Intersecting mixed methods with a contextually responsive cultural adaptation approach allowed the researchers to conclude that culturally adapted EBP programs can have a positive effect on Aboriginal youngsters.

Kinsman et al. (2001) employed mixed methods research to conduct a large-scale evaluation of the implementation of a comprehensive extracurricular school-based AIDS education program in Uganda. Using a national sample of 20 schools and involving teachers in helping with data collection, researchers were able to obtain reliable information about the outcomes of the HIV educational intervention. The qualitative results from student interviews at the completion of the intervention provided corroborating evidence to the quantitative assessment of the intervention effectiveness and helped identify implementation barriers for further action. The mixed methods evaluation results indicated the need for more active teacher involvement with the program curriculum, potentially through CBPR and action research.

Bogar et al. (2017) intersected mixed methods with the CBPR approach by engaging young people with the mixed methods study design, study implementation and interpretation and dissemination of the results. Members of the Youth Advisory Council were instrumental in developing quantitative and qualitative data collection instruments, administering surveys, and facilitating focus groups capitalizing on their experiences and knowledge of the local community. As a result, young co-researchers were

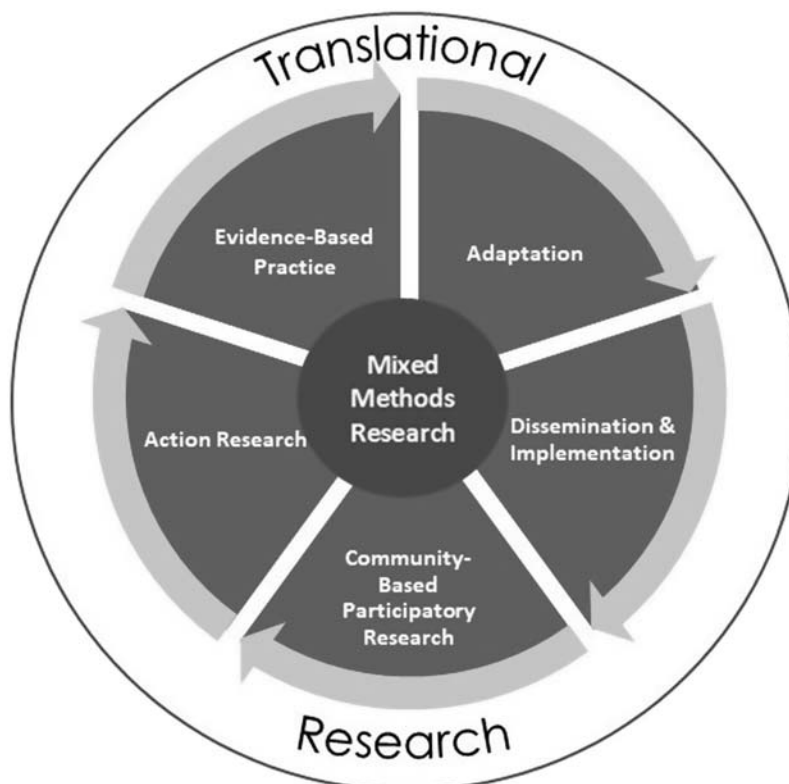


Fig. 1 Intersecting mixed methods with translational research framework. Adapted from Ivankova et al. (2019) with permission of Dialectical Publishing.

able to identify environmental health issues unique to the young people in this community, which enhanced knowledge translation and prioritization among urban youth.

In Breimaier et al.'s (2015) study, mixed methods research informed all phases in the action research cycle. Both quantitative and qualitative data were collected and analyzed to assess nurses' knowledge and attitudes to fall-prevention guidelines implementation and to explore potential facilitators and barriers both at the individual and organization levels. The integrated results informed adaptation of the implementation strategies and the content of education sessions. Evaluation of the educational intervention also employed a mixed methods design and yielded rigorous insights about the implementation outcomes and directions for further dissemination, implementation, and monitoring of the evidence-based generated practices.

Advantages and challenges of intersecting mixed methods with translational research: implications for educational research and practice

There are multiple advantages of intersecting mixed methods with translational research. An obvious advantage is that mixed methods research can be used within all aspects of the knowledge translation process. All our study participants believed that mixed methods can be effectively applied in all stages of the knowledge translation continuum, from T1 to T4. At T1 stage, researchers should think about potential utility of their study results by considering stakeholder's perspectives and the quantitative indicators. At T2, evidence-based guidelines are formed by meaningfully combining evidence from multiple sources including quantitative and qualitative data. During T3 and T4 stages, employing mixed methods provides a holistic perspective by adding greater depth and scope to complex processes and allowing better understanding of intervention contexts and users by investigating their needs and values.

More specifically, mixed methods research enhances the knowledge translation process in the following ways. As a rigorous research methodology, mixed methods provides methodological support to EBP, adaptation, D&I, CBPR, and action research by enabling a more systematic approach to exploring, evaluating, disseminating, and implementing an innovation (Ivankova et al., 2018). A meaningful integration of multiple methods within a translational research project generates more valid inferences and informs the need and direction for innovation adaptation, dissemination, and implementation, thereby promoting transferability of evidence-based practices in real-world settings and creating a solid ground for sustainability of the innovation in a community of practice. As one of our participants observed about the role of mixed methods in bridging the gap between research and practice, "Mixed methods, if anything, helps connect researchers to the real world." This is particularly true in educational research aimed at correctly assessing students' knowledge and developing teaching and learning strategies that will be reflective of diverse students' needs.

A pragmatic nature of mixed methods research makes it advantageous in illuminating, assessing and supporting change over time without sacrificing the credibility and validity standards (Ivankova and Wingo, 2018; Perry, 2009). Using qualitative stakeholder-oriented methods, mixed methods research allows for creating contextual understandings of what may impact the process of knowledge translation by exploring stakeholders' perspectives, needs and experiences and integrating them with the quantitative outcomes data. Doing so, helps identify venues for engaging stakeholders and creating partnership to optimize adoption of research findings in practice similar to Kinsman et al.'s (2001) study that engaged teachers as co-researchers to evaluate the implementation of a comprehensive extracurricular school-based AIDS education program.

While there are considerable advantages, intersecting mixed methods with translational research is not a straightforward process and may involve several challenges for researchers. Conducting mixed methods translational research requires depth and breadth of expertise in quantitative, qualitative, mixed methods, and translational research methods; experience in data collection, management, and analysis of large volumes of data; facilities to conduct these types of research; and sufficient time to complete all phases of the study (Halcomb, 2019; Fudge et al., 2016; Wooten et al., 2014). As the need for individual expertise within each discipline is often noted in mixed methods research to maintain rigorous research methods and discipline specific standards, identifying collaborators with individual expertise and establishing discipline-specific team leaders may also be challenging in mixed methods translational research (Fudge et al., 2016; Roberge-Dao et al., 2019).

Beyond discipline-specific expertise, challenges persist in identifying investigators with personal characteristics that may support the intersection of mixed methods research with translational research continuum. In our study regarding the role and value of mixed methods to optimize translational research we also explored perspectives about the individual-level characteristics of mixed methods translational thinkers. We asked our participants to describe characteristics of researchers they considered to be mixed methods translational thinkers. Several characteristics emerged as common for mixed methods translational thinkers including: thinking in the context of the intervention to recognize the setting and potential value for stakeholders; an interest in the process of translation as well as the outcome; an interest in streamlining the timeframe from innovation development to adoption; and a willingness to explore alternative methods to support the translational research continuum. For example, one participant noted, "I'm learning every day, I am more and more open to creative study designs that are good for real world use." While additional research is needed to more fully explore the emerging concept of mixed methods translational thinking, individual characteristics that foster these perspectives may support the intersection of mixed methods and translational research.

Other challenges that have been noted by our study participants relate to the need for considering community priorities and engaging stakeholders with mixed methods translational research projects, requiring expertise in and understanding of mixed methods research. In educational settings, developing students' mixed methods skills and engaging them with translational research while carefully considering their interests and needs can help promote evidence-based practices aimed at educational improvements

and develop the sense of ownership for innovative learning and teaching practices. Time and resources to conduct mixed methods studies, due to its complexity and scope within a translational research continuum, are of critical importance and are particularly challenging when designing and implementing different components of the knowledge translation process.

Conclusion

Intersecting mixed methods with translational research and its integral components offers methodological advantages for translational research and practice and provides a methodologically robust foundation for optimizing the knowledge translation process. This intersection expands the scope of EBP, adaptation, D&I, CBPR, and action research and adds rigor to the assessment of the translational research process and outcomes; thereby providing a way of integrative thinking for addressing complex problems (Martin, 2007). The flexibility of mixed methods research to intersect with each of these components and its ability to facilitate different stages in the knowledge translation process, makes it advantageous for application in educational research and practice, creating and evaluating the learning environments that meet the needs of all students.

Acknowledgment

This work was supported by 2018–19 The University of Alabama at Birmingham Faculty Development Grant Program (FDGP).

References

- Aarons, G.A., Fettes, D.L., Sommerfeld, D.H., Palinkas, L.A., 2012. Mixed methods for implementation research: application to evidence-based practice implementation and staff turnover in community-based organizations providing child welfare services. *Child. Maltreat.* 17, 67–79.
- Allen, J.D., Linnan, L.A., Emmons, K.M., 2012. Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. In: Brownson, R.C., Colditz, G.A., Proctor, E.K. (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice*. Oxford University Press, Oxford, England.
- Austin, C.P., 2021. Opportunities and challenges in translational science. *Clin. Transl. Sci.* 14, 1629–1647.
- Baydala, L., Fletcher, F., Worrell, S., Kajner, T., Letendre, S., Letendre, L., Rasmussen, C., 2014. Partnership, knowledge translation, and substance abuse prevention with a First Nations Community. *Progress in Community Health Partnerships: Research, Education, and Action* 8, 145–155.
- Bogar, S., Szabo, A., Woodruff, S., Johnson, S., 2017. Urban youth knowledge and attitudes regarding lead poisoning. *J. Community Health* 42, 1255–1266.
- Breimaier, H.E., Halfens, R.J., Lohrmann, C., 2015. Effectiveness of multifaceted and tailored strategies to implement a fall-prevention guideline into acute care nursing practice: a before-and-after, mixed-methods study using a participatory action research approach. *BMC Nurs.* 14, 18.
- Brownson, R.C., Colditz, G.A., Proctor, E.K. (Eds.), 2012. *Dissemination and Implementation Research in Health: Translating Science to Practice*. Oxford University Press, Oxford, England.
- Callard, F., Rose, D., Wykes, T., 2011. Close to the bench as well as the bedside: involving service users in all phases of translational research. *Health Expect.* 15, 389–400.
- Century, J., Cassata, A., 2016. Implementation research: finding common ground on what, how, why, where, and who. *Rev. Res. Educ.* 40, 169–215.
- Chambers, D.A., Norton, W.E., 2016. The adaptome: advancing the science of intervention adaptation. *Am. J. Prev. Med.* 51, S124–S131.
- Cochrane, A.L., 1972. Effectiveness and Efficiency: Random Reflections on Health Services. Nuffield Trust, London, UK.
- Cohen, D.J., Crabtree, B.F., Etz, R.S., Balasubramanian, B.A., Donahue, K.E., Leviton, L.C., Clark, E.C., Isaacson, N.F., Stange, K.C., Green, L.W., 2008. Fidelity versus flexibility: translating evidence-based research into practice. *Am. J. Prev. Med.* 35, S381–S389.
- Collins, F.S., 2011. Reengineering translational science: the time is right. *Sci. Transl. Med.* 3, 1–6.
- Coyle, C.E., Schulman-Green, D., Feder, S., Toraman, S., Prust, M.L., Plano Clark, V.L., Curry, L., 2016. Federal funding for mixed methods research in the health sciences in the United States: recent trends. *J. Mix. Methods Res.* 12, 305–324.
- Cranney, J., McDonald, F., 2012. Evidence-based learning. In: Seel, N.M. (Ed.), *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA.
- Creswell, J.W., 2012. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, fourth ed. Merrill Prentice Hall, Upper Saddle River, NJ.
- Creswell, J.W., Klassen, A.C., Plano Clark, V.L., Smith, K.C., FOR THE OFFICE OF BEHAVIORAL AND SOCIAL SCIENCES RESEARCH, 2018. *Best Practices for Mixed Methods Research in the Health Sciences*. National Institutes of Health, Bethesda, MD.
- Creswell, J.W., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Los Angeles.
- Crime And Justice Institute At Community Resources For Justice, 2009. *Implementing Evidence-Based Policy and Practice in Community Corrections*, second ed. National Institute of Corrections, Washington, DC.
- de Visser, R.O., Graber, R., Abraham, C., Hart, A., Memon, A., 2020. Resilience-based alcohol education: developing an intervention, evaluating feasibility and barriers to implementation using mixed-methods. *Health Educ. Res.* 35, 123–133.
- DeJonckheere, M., Lindquist-Grant, R., Toraman, S., Haddad, K., Vaughn, L.M., 2019. Intersection of mixed methods and community-based participatory research: a methodological review. *J. Mix. Methods Res.* 13, 481–502.
- Faircl, Z., Grunbaum, J.A., Gray, B.S., Franks, A., Simoes, E., 2007. Community-based participatory research: necessary next steps. *Prev. Chronic Dis.* 4.
- Fort, D.G., Herr, T.M., Shaw, P.L., Gutzman, K.E., Starren, J.B., 2017. Mapping the evolving definitions of translational research. *J. Clin. Transl. Sci.* 1, 60–66.
- Fudge, N., Sadler, E., Fisher, H.R., Maher, J., Wolfe, C.D.A., McKeivitt, C., 2016. Optimising translational research opportunities: a systematic review and narrative synthesis of basic and clinician scientists' perspectives of factors which enable or hinder translational research. *PLoS One* 11, e0160475.
- Green, C.A., Duan, N., Gibbons, R.D., Hoagwood, K.E., Palinkas, L.A., Wisdom, J.P., 2015. Approaches to mixed methods dissemination and implementation research: methods, strengths, caveats, and opportunities. *Adm. Pol. Ment. Health* 42, 508–523.
- Green, L.W., Mercer, S.L., 2001. Can public health researchers and agencies reconcile the push from funding bodies and the pull from communities? *Am. J. Public Health* 91, 1926–1929.
- Hacker, K., 2013. *Community-Based Participatory Research*. Sage, Thousand Oaks, Ca.
- Halcomb, E.J., 2019. Mixed methods research: the issues beyond combining methods. *J. Adv. Nurs.* 75, 499–501.
- Hull, L., Goulding, L., Khadjesari, Z., Davis, R., Healey, A., Bakolis, I., Sevdalis, N., 2019. Designing high-quality implementation research: development, application, feasibility, and preliminary evaluation of the implementation science research development (ImpRes) tool and guide. *Implement. Sci.* 14, 80.

- Isreal, B.A., Schulz, A.J., Parker, E.A., Becker, A.B., 1998. Review of community-based research: assessing partnership approaches to improve public health. *Annu. Rev. Publ. Health* 19, 173–202.
- Isreal, B.A., Schulz, A.J., Parker, E.A., Becker, A.B., 2001. Community-based participatory research: policy recommendations for promoting a partnership approach in health research. *Educ. Health* 14, 182–197.
- Ivankova, N.V., 2015. *Mixed Methods Applications in Action Research: From Methods to Community Action*. Sage, Thousand Oaks, Ca.
- Ivankova, N.V., Herbey, I., Rousset, L., 2018. Theory and practice of using mixed methods in translational research: a cross-disciplinary perspective. *Int. J. Mult. Res. Approaches* 10.
- Ivankova, N.V., Wingo, N., 2018. Applying mixed methods in action research: methodological potential and advantages. *Am. Behav. Sci.* 62, 978–997.
- Ivankova, N.V., Wingo, N., 2022. Intersecting mixed methods with action and community-based participatory research approaches. *International Encyclopedia of Education*, fourth ed. Elsevier, Oxford, UK (in press).
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1, 112–133.
- Kemmis, S.E., 1982. *The Action Research Reader*. Deakin University Press, Geelong, Australia.
- Khoury, M.J., Gwinn, M., Ioannidis, P.A., 2010. The emergence of translational epidemiology: from scientific discovery to population health impact. *Am. J. Epidemiol.* 172, 517–524.
- Khoury, M.J., Gwinn, M., Yoon, P.W., Dowling, N., Moore, C.A., Bradley, L., 2007. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries in health care and disease prevention? *Genet. Med.* 9, 665–674.
- Kinsman, J., Nakiyingi, J., Kamali, A., Carpenter, L., Quigley, M., Pool, R., Whitworth, J., 2001. Evaluation of a comprehensive school-based AIDS education programme in rural Masaka, Uganda. *Health Educ. Res.* 16, 85–100.
- Kitson, A., Harvey, G., McCormack, B., 1998. Enabling the implementation of evidence based practice: a conceptual framework. *Qual. Health Care* 7, 149–158.
- Knoepke, J., Ingle, M.P., Matlock, D.D., Brownson, R.C., Glasgow, R.E., 2019. Dissemination and stakeholder engagement practices among dissemination and implementation scientists: results from an online survey. *PLoS One* 14, e0216971.
- Koshy, E., Koshy, V., Waterman, H., 2011. *Action Research in Healthcare*. Sage, Thousand Oaks, Ca.
- Lewin, G., 1948. *Resolving Social Conflicts: Selected Papers on Group Dynamics* by Kurt Lewin. Harper and Brothers Publishers, New York, Ny.
- Lucero, J., Wallerstein, N., Duran, B., Alegria, M., GREENE-Moton, E., Isreal, B., Kastelic, S., Magarati, M., Oetzel, J., Pearson, C., Schulz, A., Villegas, M., White Hat, E.R., 2018. Development of a mixed methods investigation of process and outcomes of community-based participatory research. *J. Mix. Methods Res.* 12, 55–74.
- Mace, F.C., Critchfield, T.S., 2010. Translational research in behavior analysis: historical traditions and imperatives for the future. *J. Exp. Anal. Behav.* 93, 293–312.
- Martin, R., 2007. How successful leaders think. *Harv. Bus. Rev.* 85, 60–67.
- Mckleroy, V.S., Galbraith, J.S., Cummings, B., Jones, P., Harshbarger, C., Collins, C., Gelaude, D., Carey, J.W., ADAPT, T.E.A.M., 2006. Adapting evidence-based behavioral interventions for new settings and target populations. *AIDS Educ. Prev.* 18, 59–73.
- McNiff, J., Whitehead, J., 2011. *All You Need to Know about Action Research*, second ed. Sage, London, England.
- Meister, L., 2018. On methodology: how mixed methods research can contribute to translation studies. *Transl. Stud.* 11, 66–83.
- Melnik, B.M., Fineout-Overholt, E., 2019. *Evidence-Based Practice in Nursing and Healthcare: A Guide to Best Practice*, fourth ed. Lippincott Williams & Wilkins, Philadelphia, Pa.
- Mills, G.E., 2011. *Action Research: A Guide for the Teacher Researcher*, fourth ed. Pearson Education, Inc, Boston, Ma.
- Moir, T., 2018. Why is implementation science important for intervention design and evaluation within educational settings? *Front. Educ.* 3, 61.
- Nastasi, B., Hitchcock, J., 2016. *Mixed Mixed Research and Culture-specific Interventions: Program Development and Evaluation*. Sage, Thousand Oaks, Ca.
- National Institutes Of Health, 2021. *Dissemination and Implementation Research in Health (R01, Clinical Trial Optional)*. PAR-19-274.
- Otchie, W.O., 2018. Evidence-based teaching in higher education. *International Perspectives on University Teaching and Learning Symposium*. Auburn University, Orlando, FL.
- Padek, M., Colditz, G., Dobbins, M., Koscielnik, N., Proctor, E.K., Sales, A.E., Brownson, R.C., 2015. Developing educational competencies for dissemination and implementation research training programs: an exploratory analysis using card sorts. *Implement. Sci.* 10, 114.
- Padek, M., Mir, N., Jacob, R.R., Chambers, D.A., Dobbins, M., Emmons, K.M., Kerner, J., Kumanyika, S., Pfund, C., Proctor, E.K., Stange, K.C., Brownson, R.C., 2018. Training scholars in dissemination and implementation research for cancer prevention and control: a mentored approach. *Implement. Sci.* 13, 18.
- Palinkas, L.A., Cooper, B.R., 2017. Mixed methods evaluation in dissemination and implementation science. In: Brownson, R.C., Colditz, G.A., Proctor, E.K. (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice*, second ed. Oxford University Press, Oxford, UK.
- Perry, J., 2009. A combined social action, mixed methods approach to vocational guidance efficacy research. *Int. J. Educ. Vocat. Guid.* 9, 111–123.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE, Los Angeles, California.
- Rabin, B.A., Brownson, R.C., Haire-Joshu, D., Kreuter, M.W., Weaver, N.L., 2008. A glossary for dissemination and implementation research in health. *J. Publ. Health Manag. Pract.* 14, 117–123.
- Reason, P., Bradbury, H., 2008. Introduction. In: Reason, P., Bradbury, H.E. (Eds.), *The SAGE Handbook of Action Research: Participative Inquiry and Practice*, second ed. Sage, Thousand Oaks, Ca.
- Resnicow, K., Soler, R., Braithwaite, R.L., Ahluwalia, J.S., Butler, J., 2000. Cultural sensitivity in substance use prevention. *J. Community Psychol.* 28, 271–290.
- Roberge-Dao, J., Yardley, B., Menon, A., Halle, M.C., Maman, J., Ahmed, S., Thomas, A., 2019. A mixed methods approach to understanding partnership experiences and outcomes of projects from an integrated knowledge translation funding model in rehabilitation. *BMC Health Serv. Res.* 19, 230.
- Rogers, E.M., 1995. *Diffusions of Innovations*, fourth ed. The Free Press, New York, Ny.
- Rothman, K.J., Greenland, S., 2005. Causation and causal inference in epidemiology. *Am. J. Public Health* 95, S144–S150.
- Roux, P., Donadille, C.C., Magen, C., Schatz, E., Stranz, R., Curado, A., Tsiakou, T., Verdes, L., Aleksova, A., Carrieri, P., Mezaache, S., Ali, B.C., EUROSIDER STUDY GROUP, 2021. Implementation and evaluation of an educational intervention for safe injection in people who inject drugs in Europe: a multi-country mixed-methods study. *Int. J. Drug Pol.* 81, 102992.
- Shorten, A., Smith, J., 2017. Mixed methods research: expanding the evidence base. *Evid. Base Nurs.* 20, 74–75.
- Straus, S.E., Tetree, J., Graham, I., 2009. Defining knowledge translation. *CMAJ (Can. Med. Assoc. J.)* 181 (3–4), 165–168.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, Ca.
- Taylor, M.V., Priefer, B.A., ALT-White, A.C., 2016. Evidence-based practice: embracing integration. *Nurs. Outlook* 64, 575–582.
- Wethington, E., Burgo, L.D., 2015. Translational research on caregiving: missing links in the translational process. In: Gaugler, J.E., Kane, R.L. (Eds.), *Family Caregiving in the New Normal*. Academic Press, London, England.
- Wingood, G.M., Diclemente, R.J., 2008. The ADAPT-ITT model: a novel method of adapting evidence-based HIV interventions. *J. Acquir. Immune Defic. Syndr.* 47, S40–S46.
- Wisdom, J., Creswell, J.W., 2013. *Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis while Studying Patient-Centered Medical Home Models*. BMJ Publishing Group, Rockville, Maryland.
- Woods, N.F., Magyary, D.L., 2010. Translational research: why nursing's interdisciplinary collaboration is essential. *Res. Theor. Nurs. Pract.* 24, 9–24.
- Wooten, K.C., Rose, R.M., Ostir, G.V., Calhoun, W.J., Ameredes, B.T., Brasier, A.R., 2014. Assessing and evaluating multidisciplinary translational teams: a mixed methods approach. *Eval. Health Prof.* 37, 33–49.
- Zoellner, J.M., Porter, K.J., 2017. Translational research: concepts and methods in dissemination and implementation research. In: Coulston, A.M., Boushey, C.J., Ferruzzi, M.G., Delahanty, L.M. (Eds.), *Nutrition in the Prevention and Treatment of Disease*, fourth ed. Academic Press, Cambridge, MA.

Mixed methods in education RCTs

Howard White^a and Ting Yang^b, ^aEvaluation and Evidence Synthesis, Global Development Network, Delhi, New Delhi, India; and
^bIndependent Evaluation Consultant, Beijing, China

© 2023 Elsevier Ltd. All rights reserved.

Introduction	599
Terminology	600
Evaluation type and design	600
Mixed methods research?	600
Advantages and disadvantages of mixed methods with RCT	601
Formative research and evaluation	602
Formative research	602
Formative research can address the following questions	603
Formative evaluation	603
Efficacy and effectiveness trials	605
Conclusions	606
References	607

Introduction

Randomized Controlled Trials (RCTs) have become an increasingly popular approach in social science research. RCTs are commonly used to evaluate the effectiveness of interventions intended to change behavior and improve wellbeing. The core of the approach is that a set of individuals, households, schools, communities, firms, or any other unit of interest are randomly assigned to either receive the intervention (the treatment group) or to not do so (the control group, which may receive usual services, be put on a wait list, an alternative intervention or nothing at all).

The number of published RCTs has grown rapidly across social science disciplines: see, for example, trends in the rise of social work RCTs (Thyer, 2015), of education interventions (Connolly et al., 2018), and in the field of international development (White, 2019). Connolly et al. (2018) identified 1017 education RCTs, of which over three-quarters were in the period 2007–2016. And, notably, the Nobel Prize for Economics was awarded in 2019 for the use of RCTs to evaluate development interventions in low- and middle-income countries. Whilst RCTs have grown in popularity, they are also subject to a number of criticisms; see, for example, summaries of such criticisms in Connolly et al. (2018) and White (2013). Many of these criticisms may be overcome through the combination of RCTs with other methods as part of a mixed methods approach.

“Mixed methods” is often taken to mean the combination of quantitative and qualitative approaches. In this chapter, we use the term more broadly to mean the combination of any two different research or evaluation approaches; and in the context of RCTs, this means combining any non-RCT research or evaluation approach with an RCT. For example, a representative survey may be undertaken to assess the prevalence of an issue such as substance abuse amongst adolescents, combined with focus groups of adolescents to understand pathways into use. These are, respectively, quantitative and qualitative formative research which would inform intervention design.

In line with these definitions and with reference to common practice in health research, but less common in other social sciences, we recommend preceding an RCT (efficacy and effectiveness studies) of an education intervention with both formative research and formative evaluation, which may use both qualitative analysis and descriptive quantitative analysis. Without such initial steps, an RCT may be a waste of time and financial resources as we may not need an RCT to discover the reasons why the intervention does not work. Formative research may discover that the proposed intervention would be tackling the wrong problem or issue. Formative evaluation may find that the intervention as planned is infeasible for any one of a number of reasons.

Furthermore, if an intervention appears feasible following the formative evaluation, then an efficacy trial should be undertaken. Efficacy trials test whether the intervention has an effect when implemented with high fidelity to the intended target population. If the intervention is found to work in the efficacy study, then it should be taken to scale and subject to an effectiveness trial. Effectiveness trials test if the intervention has an effect when carried out under actual field conditions – so implementation fidelity is likely to be weaker and targeting to intended beneficiaries imperfect. Both the efficacy and effectiveness trials should have integrated mixed methods designs, which combine both impact evaluation (e.g. RCT) and process evaluation.

As such, mixed methods design is embedded in the approach following these four stages. Each stage is further elaborated in this chapter with examples particularly from education settings.

This chapter begins with a short discussion of terminology (Section **Terminology**), followed by a discussion of the advantages and disadvantages of using mixed methods with RCTs (Section **Advantages and disadvantages of mixed methods with RCT**). Section **Formative research and evaluation** describes formative research and evaluation, and Section **Efficacy and Effectiveness trials** describes the use of mixed methods in efficacy and effectiveness trials. Section **Conclusions** draws out our conclusions.

Terminology

Evaluation type and design

Terminology is not used consistently in the evaluation literature, so we begin by defining some terms.

First, we distinguish evaluation type and design. Regarding types of evaluation, Scriven proposed a dichotomy of formative and summative evaluation as the two types of evaluation. He defined these respectively as: “Formative evaluation is conducted during the development or improvement of a program or product” (1981, p. 63), and “Summative evaluation of a program (etc) is conducted after completion” (1981, p. 150). A second dichotomy of type distinguishes process evaluations and impact evaluations. As pointed out by Scriven (1996), this is different from the formative/summative dichotomy since both summative and formative evaluations may involve a process evaluation. Process evaluations assess the management and implementation of an intervention, and impact evaluations assess what difference the intervention has made to outcomes. In this definition, process evaluations rely on either quantitative or qualitative factual analysis, whereas the core of impact evaluation is a counterfactual analysis (that is comparing what happened with the intervention to the hypothetical case of the same group not receiving the intervention), though the latter is contested by some (see White, 2010).

Second, we refer to the evaluation design as all aspects of how the evaluation will be undertaken. That is (1) data collection, which covers both sampling (from whom data are collected) and data collection instruments and approaches (how data are collected), and (2) methods, which are how the data will be analyzed.

An RCT is a specific type of impact evaluation. A complete description of the design would include both sampling strategy, method of random assignment (e.g., cluster, simple, or matched pair, and how units are randomly assigned) and the way in which the data will be analyzed, e.g., t-tests of double difference of the means of outcomes, or instrumental variable regression with assignment as the instrument. A mixed methods RCT will include also a process evaluation, for which both sampling strategy and analysis will be different from the quantitative counterfactual analysis of effects.

Mixed methods research?

Mixed methods research is commonly taken to mean the combination of quantitative and qualitative approaches. There are two issues here. “Approaches” is a bit vague, and people often conflate type of data, or data collection, with methods (White, 2022). Neither focus groups or a sample survey are an analytical method. Our definition of design above clarifies this issue. So, a qualitative evaluation design should, for example, describe how focus group members are recruited, how the data from the group are recorded, and how they will be analyzed. For example, Gittelsohn et al.’s (2006) formative research to inform physical activity amongst girls in secondary schools states that the study covered six schools with purposive sampling for 80 individual interviews, and 11 focus groups with boys, and 13 groups with 100 girls in seventh and eighth grade. The focus groups covered a list of 44 open-ended questions, with the session lasting 45–60 min. A matrix-based analysis of themes from the focus groups was used, assisted by NUD*IST qualitative analysis software.

Second, we see the quant/qual distinction as too restricted. A further useful distinction, especially in the context of RCTs and other impact evaluations, is between factual and counterfactual analysis. Factual analysis, that is, simply what happened, can be used to address many questions of both research and policy interest. Factual analysis has no need of a control group. This can be illustrated with reference to three examples: the frequently neglected issues of targeting, understanding attrition and problem analysis.

First, take the example of targeting in after-school programmes. These programmes are generally universal. But one intent may be to reach children from disadvantaged backgrounds, for purposes of remedial education, for social and life skills development or as a diversionary activity from potential engagement with gangs or criminal behavior. Targeting can consist of both quantitative and qualitative factual data and analysis. Using quantitative data on participation we can calculate targeting errors: Type I error is the proportion of the target group who don’t participate, and Type II is the proportion of those attending who are not in the target group. A thematic analysis of qualitative data, collected from one-on-one interviews, focus groups or other less structured forms of data collection, can identify the reasons for targeting errors. Reasons for Type I error may include that there is a fee for attending, the fee may be waived but parents do not know that or are reluctant to apply for it (stigma, suspicion of authority, skepticism etc), or the child is enrolled but does not attend. It may also be the case that the school, or those running the aftercare, exclude children displaying disruptive behavior. Type II error can occur if the after-school program is a cost-effective means of childcare for working parents, so those who are better off may patronize it, and indeed influence content and activities to be less appealing to the target group.

The second example of factual analysis is understanding attrition in an RCT. Attrition is often unjustly neglected, with studies reporting what is called the “treatment of treated” which is the effect of those fully exposed to the intervention, which ignores those who dropped out or who didn’t show up in the first place. But, as argued by Gorard et al. (2017), the findings of many RCTs may be unsound as attrition reintroduces possible selection bias. The Institute of Education Sciences (IES) publishes acceptable thresholds for attrition, which show that even low levels of differential attrition between treatment and control undermine the validity of RCTs (IES, 2014). But, whilst many researchers may see attrition, as a technical inconvenience which may undermine the internal validity of their effect estimates, for the program implementer it is a failure which reduces the impact of their program, and so the program’s cost effectiveness and likely sustainability. Hence, researcher and implementer alike have a common interest in understanding and reducing attrition. Attrition is likely if the incentives for participation are not sufficient. This comes down to a simple cost-benefit

assessment on the part of the participant, including the assessment of the likelihood of detection and punishment if participation is mandatory.

Attrition is usually more complicated than simply participated or didn't participate. We can usually distinguish at least the following stages: get identified as a participant, referred or recruited, attend, maintain attendance, complete. The notion of the "funnel of attrition" captures the fact that children may drop out at any stage (see [White, 2014](#), on the funnel of attrition). Suppose 100 children are intended to participate in a program. Some of these notional 100 will not be identified as they may be street children, from currently homeless or nomadic families, be out of school or home-schooled, or being temporarily excluded. Suppose five per cent of potential participants are lost for these reasons. Not showing up is usually the reason for the largest drop in participation. This depends on the nature of the program, especially if it is mandatory or voluntary, in school hours on school premises or out of hours and elsewhere. For teenagers "turf" issues can affect their willingness to attend activities in off site venues. Another issue is parental consent which is increasingly required around the world: parents may not give it, or they may not get the letter home in the first place. Say 35% are lost at this stage, so the total number of children attending is 60, already less than two-thirds of the original target group.¹ But many of these 60 may drop out or only attend irregularly, so perhaps one-third are exposed to the full treatment. As in the case of targeting, the quantitative statement of attrition rates is best complemented by qualitative analysis to understand individual motivations which explain attrition.

Finally, is the often-neglected assessment of the intervention's problem analysis. Weak problem analysis, which is common in the absence of formative research, may lead to misdiagnosis and misleading explanations and understanding of project performance ([Dixon and Bamberger, 2022](#)). Using the same example of after-school programmes, with the targeted groups of students, additional factual analysis would help identify and verify if real problems have been targeted and appropriate solutions are provided by the proposed intervention. For example, if the real reason for not attending after-school programmes and related activities is due to that some parents are not even aware of them, then proposing measures such as increasing incentives (e.g., provision of vouchers) for participating may not directly help. The problem-based, tailored solution would be first exploring ways of conveying the messages to parents. Formative research is one way of addressing the gap in problem analysis by identifying priority issues, outcomes and key constraints to the success of any intervention. For example, by combining formative research (e.g., through interviews or focus groups with parents in this case) and a quantitative survey during the re-visit of the problem analysis would help figure out the diagnostic issue so that timely adjustment could be made.

Advantages and disadvantages of mixed methods with RCT

Evaluation should be issues-led not methods-led. Different evaluation types and designs are most appropriate depending on the evaluation purpose and questions. The choice of appropriate evaluation design depends on several factors such as the nature of the intervention, evaluation purpose and scope, context, and a range of operational and policy issues (; [Stern et al., 2012](#); [White, 2010](#)). Mixed methods designs can be usefully added to RCTs, as RCTs are the most appropriate design for addressing the question of causal effects, but that is not the only question of interest to both researchers and practitioners.

Mixed methods bring in a range of benefits. In addition to enabling triangulation by providing complementary data and findings drawn from these data, mixed methods design allows analysis of each individual program component instead of using one unified design ("black box") for largely different components ([White, 2014](#); [Yang, 2018](#)). And most importantly, mixed methods deepen the explanatory power of a study – unpacking the causal mechanisms² of a program to understand how and why the program works and the intermediary outcomes connecting inputs, outputs and outcomes to achieve the desired impact or goal [Creswell et al. \(2009\)](#), [Drabble et al. \(2014\)](#), [Edmunds et al. \(2021\)](#), [Plano Clark et al. \(2013\)](#), [Song et al. \(2010\)](#), [White \(2008, 2013\)](#), and [Yang \(2018\)](#).

For example, in education research and evaluation, the use of RCTs has drawn critiques regarding its feasibility in the education sector, its neglect of context and experiences and the tendency to make over-simplistic universal causal claims ([Connolly et al., 2018](#)). These criticisms can be largely countered with a mixed methods approach – embedding RCTs in a broader evaluation design, which combines the counterfactual analysis of attribution with a factual analysis across the causal chain, or adding process evaluation to RCTs under an overarching theory-based evaluation design. However, mixed methods design requires the right technical expertise and will add to the total evaluation budget.

There are various ways of mixing or combining methods. [Carvalho and White \(1997\)](#) proposed three ways of employing mixed methods in below [Table 1](#): that is (1) integrating, where the qualitative work informs the design or interpretation of the quantitative component or vice versa; (2) confirming, refuting, enriching, or explaining where qualitative findings help validate and add color to a quantitative finding, or conflict with it which then needs further exploration (for example, we suggested above that qualitative methods are best suited for understanding targeting success and failure, and attrition); and (3) merging the findings to draw out a single set of implications for policy, practice and research. The choice of specific way of mixing and matching, combining or

¹The funnel of attrition continues to operate on outcomes along the causal chain to learning outcomes. Children need to show up to class, pay attention, understand the material, and apply it to internalize ("learn") it. Only a small percentage make it to the last stage.

²A program mechanism focus needs to unpack what particular aspects or components of a program make it work, or indicates how both macro and micro processes construct and influence the program ([Pawson and Tilley, 1997](#), p. 66).

Table 1 Different ways to combine quantitative and qualitative methods.

<i>Typology of mixing methods</i>	<i>Specific ways to operationalize the use of mixed methods</i>
(1) Integrating methodologies	• Using quantitative baseline survey data to determine the individuals/communities to be studied through the qualitative approach • Using the quantitative baseline survey to inform the design the interview guide of the qualitative survey by identifying issues of interest • Using qualitative work to determine beneficiary selection for the intervention and possible stratification of the RCT (quantitative) sample • Using qualitative work to inform the design of the quantitative survey questionnaire, e.g. outcome selection and definition • Using qualitative work to pre-test the quantitative survey questionnaire • And/or using qualitative analyses to refine, for example, measurement tools e.g. an index of deprivation
(2) Confirming, refuting, enriching, and explaining the findings of one approach with those of the other	• “Confirming” or “refuting” are achieved by verifying quantitative results through the qualitative approach. • “Enriching” is achieved by using qualitative work to identify issues or obtain information on variables not obtained by quantitative surveys. • “Examining” refers to generating hypothesis from qualitative work for testing through the quantitative approach. • “Explaining” involves using qualitative work to understand results from quantitative data, in particular the lower parts of the causal chain.
(3) Merging the findings of the two approaches into one set of policy recommendations	Analyzing the information provided both by the quantitative approach as well as the qualitative approach to derive one set of implications for policy, practice or research.

Source: Adapted from [Carvalho and White \(1997, p. v\)](#).

integrating methods is determined by specific context, research and evaluation questions, purposes and scope. Examples of these uses are given in [Table 1](#).

This typology also applies to the use of mixed methods with RCTs in various ways. Different types of quantitative and qualitative methods or factual and counterfactual analysis, can be combined as shown as ways of operationalizing in the right column of [Table 1](#) to achieve the ends (e.g., aim or purpose of mixing methods) as shown in the left column of [Table 1](#).

Formative research and evaluation

Formative research and evaluation are two initial stages which are often unduly neglected. Formative research is undertaken to help shape an intervention which is planned but not yet fully, or at all, designed. Formative evaluation is an initial field testing of the feasibility of the intervention in the intended setting. As emphasized above, terminology is not standard, so what we would call formative evaluation is sometimes called formative research, and what we would call summative process evaluation is sometimes labeled as formative evaluation.³

Formative research

Formative research thus takes place before any intervention starts as the role of the research is to inform intervention design. It is by its nature, factual not counterfactual, and may use both quantitative and qualitative data. Formative research may be used to identify both priority issues or outcomes, and the most important constraints to improving those outcomes which should be addressed

³We found examples of each of these in our search for case to use in this chapter.

by any intervention if it is to succeed. Using formative research prior to undertaking an RCT is, by definition, mixing methods across the whole research design, even if the formative research is purely qualitative and the RCT purely quantitative.

Sherry (2013) compares the substantial formative research behind the successful educational television program *Sesame Street* with the lack of such research, and subsequent millions of dollars wasted, in developing educational video games. For example, comedy writers were hired to write the *Sesame Street* scripts as it was found that writers could be taught to teach, but teachers could not be taught to write suitable scripts. *Sesame Street* has subsequently been found to improve children's learning outcomes in a number of countries around the world (e.g., Ball and Bogatz, 1970, on the first season in the United States, and Diaz-Guerrero et al., 1976, on Mexico, and Fluent Public Opinion and Market Research, 2008, on Kosovo).

Another example comes from the US National Heart, Lung, and Blood Institute's Trial of Activity for Adolescent Girls (TAAG), a multi-center, multicomponent, randomized field trial of 36 middle schools, of a multicomponent intervention linking schools with community organizations to reduce the age-related decline in physical activity in middle school girls. The design of TAAG involved extensive formative research with a range of methods including survey (e.g., with schools, parents and community agency), in-depth interviews, focus groups and activity checklist (Gittelsohn et al., 2006). For example, focus groups were conducted at the six TAAG field sites during the formative research to help identify audience segments, which were labeled athletic, preppy, quiet, rebel, smart, tough, and other. When responding to the question about barriers to participating in physical activity, for instance, focus groups felt preppy girls were concerned about appearance, injury and time, which helped generate messages in relation to themes featuring fashionable girls and social opportunities. In contrast, quiet girls were more concerned about embarrassment, shyness, uncertainty about what they may be supposed to do or believe that they don't fit in. The appropriate messages for this group stress comfort and ability to be active either alone or with small groups of friends (Staten et al., 2006).

Formative research can address the following questions

- What are the priority outcomes? These priorities can be determined through an examination of current strategy, engagement with decision-makers and implementing agency, and consultation with the intended beneficiaries. Other than the beneficiary consultation (which might use a sample survey), these exercises typically have small samples, so that quantitative approaches are not appropriate. Qualitative thematic analysis will be most appropriate. Different stakeholders may well have different priorities, so some transparency is required in the decision-making. This transparency is also good research practice as qualitative studies are often poor at describing their methodology. Who was consulted, why these people and how were they selected? How was the analysis undertaken in a way to ensure that the reported findings are an accurate representation?
- Given the desired outcome, what are the binding constraints on achieving that outcome? Interventions will not achieve their intended goal if the binding constraints remain untackled. If the goal is to improve learning outcomes for disadvantaged children, are the worst performing children not attending school or being excluded, so that issue need to be addressed? Is teacher absenteeism or lack of learning materials a problem? Related formative research questions are about barriers to children attending school and barriers to effective learning.
- For example, the policy issue of concern may be increasing access to tertiary education to students from disadvantaged backgrounds. Obvious policy solutions may seem to be positive discrimination or quotas in university places, targeted grants for fees and living expenses, scholarship programmes or providing information about the benefits of tertiary education to children and their parents. But formative research could show that the issue is ensuring that children from disadvantaged backgrounds complete secondary education. Hence, although the policy goal is to increase participation of children from disadvantaged backgrounds in tertiary education, the appropriate interventions may be at secondary, not tertiary, level. Indeed, taking the long view, the most cost-effective intervention may be early childhood development (ECD) programmes such as parenting classes to encourage parents to provide the stimulation to a child to support cognitive development.
- Another important question is to assess the existing provision for what a project plans to do. For a computer-assisted learning project, what technology do children already have access to? If considering ECD interventions, what are existing childcare provision and practice?
- What is the feasibility of undertaking the intervention in the planned context? This may involve discussions with teachers and teacher unions, NGOs/charities, parents and children themselves, as well as an examination of context and existing interventions.
- Establishing the target group: who are the groups that would benefit most from the intervention and are receptive to it?
- Formative research may establish priority issues. What do communities care about? Is it clean water, jobs, quality of schooling or something else? These issues could be addressed through key informant interviews with community leaders, townhall-style community meetings or quantitative opinion surveys. If using quantitative surveys some initial qualitative work is advised to ensure that the full range of relevant options are included in the questionnaire.

Formative evaluation

Interventions commonly fail for one of three reasons: low acceptability in the target population, technical failures, and poor implementation. The first two of these can be anticipated by a good formative evaluation, which will usually be qualitative in nature. Formative evaluation is simply trying the program out on a very small scale. This is not efficacy testing – that comes next. This stage is simply to see if the intervention makes sense. Large scale evaluations costing large amounts of money often identify reasons for

failure which could have been anticipated by a good formative evaluation, such as textbooks at too advanced a level, providing computers to schools without electricity or advising teachers to arrange their desks in circles not rows when the schools have no desks. As with formative research, undertaking formative evaluation prior to the RCT is a mixed methods approach, even if the formative evaluation is purely qualitative.

Low acceptability may come for many reasons. Parents do not send their children to after-school activities as they are needed to collect water, do domestic chores and look after younger siblings. Girls refuse to attend special sessions for those designated as being at high risk of teenage pregnancy because of the stigma - the lesson of labeling theory is that you should avoid labeling, whereas such programmes may do the opposite. Children may resist after-school activities which are simply seen as "more school".

Technical issues are most obvious with technology: providing computers in schools without also providing appropriate software has been a notable cause of failure. Of course, providing computers to schools without electricity is a more obvious, though thankfully less common one. But it can apply to all teaching equipment and materials, such as telling teachers to put the children's desks in circles and posters on the walls when they have neither desks nor posters - nor, even, walls (White and Masset, 2003). Textbooks can be at the wrong level (too high or too low) or simply in the wrong language.

An example of formative evaluation comes from an educational assessment technology innovation Assessment Tools for Teaching and Learning (asTTle) educational software, in New Zealand (Hattie et al., 2006).⁴ The study aimed to gather insights, feedback and related effects with the deployment of the innovative education technology, particularly from the end-users' perspective (e.g., teachers, students, and school leaders), and to learn about areas for future improvement. Study components included teacher and student evaluation of the test materials, the accuracy and added value of the reports, the reactions and effects of the professional development, and the evaluation of the utility of the software. Multiple methods were combined during the evaluation including questionnaire surveys designed for teachers to seek their feedback on the asTTle test questions, tasks, and instructions; student surveys to gather their views about the test materials at the end of a trial test; focus groups and additional survey questionnaires to examine the accuracy of understanding of teachers in relation to asTTle reports; telephone interviews, questionnaire surveys, and field visits undertaken to look at the effectiveness of the software and its professional development resources and processes. For example, in terms of the understanding of asTTle reports, evaluations in primary schools found that teachers who best understood the asTTle results made more positive and fewer negative comments. Those who made the most negative comments appeared to be due to their poor understanding of the reports. This finding led the program to include additional assistance needed by teachers for correctly understanding and using asTTle. The use of mixed methods (both quantitative and qualitative) in this case showcases the typology in Table 1 that information obtained from different tools helps enrich one another and their findings can get merged to generate learnings and (program) policy implications.

A further example is a study commissioned by the William and Flora Hewlett Foundation who engaged edBridge Partners to conduct a formative evaluation of open education networks in North America in both K-12 and higher education, and their role and efforts in developing an inclusive and responsive open education field (edBridge, 2021). The evaluation sought to examine how well the open education networks are engaging various stakeholders, other actors and communities and serving their needs, and so to provide network leaders and organizers with evidence to inform how they can adjust and sustain their networks and the open education field. A range of mainly qualitative methods were used to gather the perceptions and beliefs of open education network leaders. Approaches used included landscape analysis (e.g., involving higher education and K-12 sector open education fields to identify networks and key actors and to examine the intersection between networks and sectors); surveys (e.g., with various groups of network leaders such as those of regional networks: states, provinces, and school districts), in-depth interviews with key leaders, and a series of facilitated dialogues (edBridge, 2021). For example, in terms of the effectiveness in engaging and serving the stakeholders, when responding to the question "how would you rate your network in its ability to serve its audience or stakeholders?" 77% of respondents rated their networks as Good or Excellent. Factors underlying positive responses were strong relationships, established systems and mechanisms which were found helpful for leaders to expand reach and best serve stakeholders. Perceptions about limitations were more about limited resources or the need for more intentional and focused efforts in networks. Accordingly, relevant recommendations were raised for future endeavor such as expanding outreach to funders and seeking collaborative opportunities (Ibid). This example adopts mixed methods in the way that findings from both quantitative and qualitative studies are merged to gather evidence-based recommendations, as illustrated in Table 1.

The Center for Evidence and Implementation (CEI) conducted a formative evaluation of a EduCarer Aide Training (ECAT) project by Presbyterian Community Services (PCS)⁵ in Singapore (CEI, 2022). The ECAT project is a capability-building initiative focusing on special education school graduates - helping build their skills for employment in pre-school settings and increasing their confidence and self-esteem through providing center-based modular training and on-site (e.g., classroom) coaching in pre-schools. The formative evaluation sought to review existing implementation mechanisms and examine preliminary effectiveness of the project to gather learning and improvement and to determine the possibility of scale-up. CEI revisited the project logic with the program team, reviewed and advised on the improvement of the curriculum and also conducted semi-structured interviews with the trainees to gather insights and information on the project performance. The formative evaluation revealed that the project

⁴A software that automates assessment capacity not by delivering interactive on-screen test questions, tasks, or portfolios as suggested by several educational technologists, but by integrating computer-aided test assembly with computer analysis of test performance into a teacher-controlled, curriculum-based system that graphically reports student achievement against norms, objectives, and standards. asTTle is a combination of computer-based and computer-assisted assessment with networking capability (Hattie et al., 2006, pp. 4-5).

⁵A social service agency registered with the National Council of Social Service in Singapore.

had good early potential to be scaled up to the national level and that there were possibilities in terms of building skills of people with disabilities for better employment opportunities and quality of life (CEI, 2022). This case applies a range of methods with a qualitative focus, findings gathered from each tool help enrich and triangulate one another.

Efficacy and effectiveness trials

Efficacy is testing if a program works when done in the target population and high implementation fidelity. Factual monitoring data makes sure things going according to plan. Qualitative process evaluation data identify implementation issues. An efficacy trial also of course measures the effect of the intervention. The non-RCT component of an efficacy may repeat some of the formative evaluation questions, but also uses process evaluation to better understand the causal chain.

Effectiveness tests the program at scale as it will be implemented in “real life”. For example, in efficacy trials the research team may be involved in implementation but that won’t be true at effectiveness. Hence at effectiveness stage, implementation issues still need to be evaluated. It is often the case that successful programmes at the pilot stage fail when taken to scale, the cause being poor implementation fidelity. A well know example is a study of contract teaching in Kenya, which found that the intervention improved student test scores when implemented at scale by NGOs but not when implemented by government agencies (Bold et al., 2013).

The Education Endowment Foundation (EEF) in the UK engaged the Sheffield Institute of Education to review its scale-up of EEF efficacy trials to effectiveness trials (the efficacy–effectiveness continuum) completed by 2019 (some of the trials adopted RCTs) to provide insights and support for the EEF, developers and delivery partners in the design and management of the scale-up of trials. The study involved a quantitative comparison of the effects at the efficacy and effectiveness stages and analyses of qualitative interviews with program team members and leaders to identify influencing factors – models, structures, processes and program features that either support or hinder intervention scale-up. A process evaluation was embedded here following the efficacy and effectiveness trials. For instance, the quantitative study found the mean unweighted effect size for seven interventions fell from 0.25 for the efficacy trials to 0.01 for the effectiveness trials. This change may result from the differences between the two trials in their design, target group, size of study and the geographical coverage (e.g., effectiveness trials were larger studies based in primary schools). The qualitative analysis found that facilitators for successful implementation at scale were clearly established structures and processes (including quality assurance of training and implementation), access to in-depth expertise on the intervention and its implementation, highly skilled trainers with experiences in children learning, teaching relevant subjects and working in school settings. These findings implied that a more systematic and evidence-based approach may be needed to bring in more clarity on necessary changes or inputs for intervention scale-up (Maxwell et al., 2021). Linking back to Table 1, this case exemplifies the way that mixed methods (e.g., RCTs and non-RCT methods) get applied and their findings merged to generate one set of policy implications.

The Learning Together intervention on bullying and aggression in English secondary schools (Bonell et al., 2018) serves a good example in terms of using mixed methods in RCTs in an educational setting – it combines and embeds economic and process evaluations to examine different key aspects and fulfill different information needs. The intervention aimed to modify the school environment through reducing bullying and aggression and improving student health and wellbeing across various domains. Following a pilot trial in eight schools to demonstrate feasibility and acceptability among participants, this inclusive cluster randomized trial was undertaken in 40 schools in England together with process and economic evaluations. The process evaluation intended to assess factors in relation to implementation, context, acceptability and employed qualitative methods such as case studies (six intervention schools were selected for in-depth assessment) and focus groups with students and staff. The economic evaluation applied a cost-consequence analysis including all main outcomes and evaluated incremental effects at 24 months and 36 months since randomization, with costs relating to use of education, police, NHS resources (e.g., through student survey questionnaires) and staff time (through staff survey questionnaire), and payment for facilitators and trainers. This application of mixed methods in education RCTs helps provide a more complete picture and evidence chain and add explanatory power especially in the complex educational context. Results showed positive effects despite the variable fidelity to this whole-school intervention – it had small but significant effects on bullying but no effect on aggression. This intervention and its findings had important implications for public health policy – a single and very low-cost intervention could bring a set of important public health-related outcomes – and for the evidence base on the feasibility and effectiveness of using whole-school approach to preventing bullying and aggression and promoting student health in broader high-income, middle-income, and low-income settings. The use of mixed methods in education RCTs in this case corresponds to the typology in Table 1 – enriching or supplementing data sets with different methods and merging results to provide evidence-based implications for policy and research.

The Continuing Professional Development (CPD) programmes in the UK were initiated to support teachers responsible for teaching science in primary schools with subject-specific knowledge and pedagogical content knowledge, as well as leadership training (e.g., leading the science provision in schools). An evaluative study (Bennett et al., 2019) was undertaken to generate evidence on the program impacts and to inform the decision-making of the possible national roll-out. The evaluation used mixed methods approach by integrating a RCT (conducted with three groups including a full treatment group receiving the full CPD program, a partial treatment group with the short CPD program, and a control group) with a series of case studies with a subset of approximately one-third of the schools in each of the three RCT groups. Under the case studies, the evaluation team conducted site visits to interview key respondents such as science specialists, teaching colleagues and management staff to gather their respective beliefs, perceptions and perspectives on the program, and to observe some of the teaching practices in classrooms. This mixed methods RCT approach helped triangulate and validate results from different data sources and provide a more complete picture of

the reality. For example, findings from the quantitative data from the RCT through the assessment of teachers' subject knowledge did not show statistically significant impact on teachers' performance after receiving the full or the partial training programmes; whereas the qualitative evidence indicated many teachers participating in the full intervention perceived improved subject knowledge and increased confidence in teaching science. This was corroborated by evidence found during lesson observations – changes in teachers' practice and etc. This example of mixed methods RCT showcases the way how data from different methods are “confirming” or “refuting” one another as indicated in [Table 1](#). Specifically, results from qualitative studies refute to some extent those from the quantitative process in this case.

The eSchool 360 model implemented by Impact Network in Zambia delivers low-cost education to children in rural communities through a holistic solution with the aim of improving educational outcomes by incorporating three key components: a standardized e-learning curriculum and pedagogy, ongoing teacher training and professional development, and community ownership. The American Institutes for Research has been engaged in the evaluation of this model by designing and implementing a mixed methods cluster-randomized controlled trial (cluster-RCT) – an impact evaluation through the cluster-RCT (63 eligible schools have been randomly assigned to either the treatment or control group), a cost-effectiveness analysis (assessment of the costs of the model using the ingredients method which needs to specify all ingredients necessary to replicate the model and then collect data on the unit costs of these ingredients), and a process evaluation (assessing the fidelity of implementation). Specifically, the qualitative process evaluation complements the impact evaluation using RCT and explores informal patterns, unanticipated interactions and unforeseen areas of interest in addition to formal activities and anticipated results; it also helps pursue in-depth understanding of why or why not the program shows impact (e.g., facilitators and barriers) as well as contextual factors. Qualitative methods such as key informant interviews, focus group discussions and classroom observations are employed to collect relevant data ([Hoop et al., 2020](#)). The results of this mixed methods cluster-RCT showed statistically significant intention-to-treat effects in the primary cognitive skills outcomes such as the early grade reading assessment, early grade mathematics assessment, the Zambian Achievement Test, and Oral Vocabulary assessment scores. The findings from the qualitative evaluation process verified those from quantitative studies – respondents perceived highly of the quality of education and improvements in literacy at IN schools (though less reported improvements in numeracy). It was also revealed that the high level of implementation fidelity should have contributed to the positive effects found in learning outcomes, and that there remains considerable room for further improvements and evaluation of longer-term effectiveness of the intervention. This case serves an example of using findings from the qualitative evaluation process to verify and confirm those from quantitative studies.

All of these examples show how qualitative methods can help strengthen external validity. A single RCT cannot be the basis for a universal causal claim. But a good understanding of contextual factors which condition success or failure of an intervention can help us understand the settings to which findings may be transferred, and so when, where and how we may try a transfer an intervention which has been tested elsewhere. The principle of transferability of evidence findings underpins the evidence revolution and the work of What Works Centers, such as the Education Endowment Foundation ([White, 2019](#)), so further research to strengthen our confidence in transferability is an important area of work.

Conclusions

RCTs are criticized for a black box approach, unable to disentangle causal pathways. The black box may in fact be a blessing not a curse ([White, 2014](#)). It is incorrectly argued that the black box means RCTs cannot evaluate complex interventions with multiple, interacting causal pathways. Their strength is that they can make valid statements about causality in precisely these circumstances. However, in general it is better to go beyond “bare bones RCTs” to answer a broader range of questions of interest to policy makers and practitioners.

However, in this chapter we have also argued a different point to the usual emphasis on combining qualitative approaches into a quantitative RCT. We argue that when planning an RCT prior stages of non-RCT research are necessary – and that such an approach is by definition a mixed methods approach. Formative research helps identify priority outcomes and the issues to be addressed by any intervention intended to support the achievement of those outcomes. Formative evaluation assesses the feasibility of the proposed intervention, most usually through qualitative fieldwork. This approach is common in health research, and our paper illustrates examples of its use in the education field.

Hence, we propose not just incorporating non-RCT methods into RCTs evaluation designs, but undertaking non-RCT studies, that is formative research and evaluation, before moving onto RCT-based efficacy and effectiveness trials. Funders should make such an analysis a pre-condition for support for an RCT.

And, we also make the case, as is now quite commonly argued, that the RCT itself should combine both quantitative and qualitative methods. As is widely accepted, evaluations should be issues-driven not methods-driven. Hence a range of methods – quantitative and qualitative, factual and counterfactual – should be used in evaluations. RCTs are the most appropriate approach for answering questions about effectiveness. Complementing RCTs with process evaluation will shed light on implementation issues, as well as understand contextual factors to help understand the extent to which findings are transferable. Too many RCTs provide too little information on the intervention and context, which make them less useful than they would otherwise be. Using mixed methods brings more of this detail to the study.

There are various ways of using mixing methods as exemplified in [Table 1](#), which show an iteration between quantitative and qualitative. Leading intervention and RCT design by formative research and evaluation is an example of qualitative research

informing quantitative research. But the baseline survey for the RCT may also identify issues warranting further qualitative inquiry, and the RCTs findings almost certainly will. From a policy perspective, qualitative insights on say barriers to participation can help inform program targeting and outreach.

In summary, this paper calls for researchers designing RCTs to consider the appropriate role of qualitative research at all stages of the study, and for qualitative researchers to see RCTs as complementary to their own research, with the two best working together. Commissioners should seek to commission more truly mixed methods RCTs, and journal editors be open to studies presenting both quantitative and qualitative findings in a single paper.

References

- Ball, S., Bogatz, G.A., 1970. The First Year of Sesame Street: An Evaluation. Final Report, Volume III of V Volumes (PR-70-15; ETS Program Report). Children's Television Workshop, New York.
- Bennett, J., Hanley, P., Abrahams, I., Elliott, L., Diepen, M.T., 2019. Mixed methods, mixed outcomes? Combining an RCT and case studies to research the impact of a training programme for primary school science teachers. *Int. J. Sci. Educ.* 41 (4), 490–509.
- Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., Sandefur, J., 2013. Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education. Center for Global Development. <https://doi.org/10.2139/ssrn.2241240>. Working Chapter No. 321, Available at: <https://ssrn.com/abstract=2241240>.
- Bonell, C., Allen, E., Warren, E., McGowan, J., Bevilacqua, L., Jamal, F., Legood, R., Wiggins, M., Opondo, C., Mathiot, A., Sturgess, J., Fletcher, A., Sadique, Z., Elbourne, D., Christie, D., Bond, L., Scott, S., Viner, R.M., 2018. Effects of the learning together intervention on bullying and aggression in English secondary schools (inclusive): a cluster randomised controlled trial. *Lancet* 392, 2452–2464.
- Carvalho, S., White, H., 1997. Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis: The Practice and the Potential. World Bank technical chapter No.366. World Bank.
- CEI, 2022. Formative Evaluation in Singapore on Skilling Program for Youth With Special Needs. Centre for Evidence and Implementation (CEI). Available at: <https://www.ceiglobal.org/our-work-original/building-capacity-through-implementation-science/formative-evaluation-singapore-skilling-program-youth-specia>.
- Connolly, P., Keenan, C., Urbanska, K., 2018. The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educ. Res.* 60 (3), 276–291.
- Creswell, J.W., Fettes, M.D., Plano Clark, V.L., Morales, A., 2009. Mixed methods intervention trials. In: Andrew, S., Halcomb, E.J. (Eds.), *Mixed Methods Research for Nursing and the Health Sciences*. Wiley-Blackwell, Hoboken, NJ, pp. 161–180.
- Diaz-Guerrero, R., Reyes-Lagunes, I., Witze, D.B., Holtzman, W.H., 1976. Plaza Sesamo in Mexico. *J. Commun.* 26, 145.
- Dixon, V., Bamberger, M., 2022. Incorporating Process Evaluation into Impact Evaluation: What, Why and How. Working (Chapter 50), 3ie. International Initiative for Impact Evaluation.
- Drabble, S.J., O' Cathain, A., Thomas, K.J., Rudolph, A., Hewison, J., 2014. Describing qualitative research undertaken with randomised controlled trials in grant proposals: a documentary analysis. *BMC Med. Res. Methodol.* 14 (1), 24–35. <https://doi.org/10.1186/1471-2288-14-24>.
- edBridge, 2021. Formative Evaluation of Open Education Networks. Final report, edBridge Partners. Available at: <https://hewlett.org/wp-content/uploads/2021/10/EdBridge-Open-Education-Networks-Evaluation.pdf>.
- Edmunds, J.A., Gicheva, D., Thrift, B., Hull, M., 2021. Using mixed methods to explore variations in impact within RCTs: the case of project compass. *J. Mix. Methods Res.* 15586898211033144.
- Gittelsohn, J., Steckler, A., Johnson, C.C., Pratt, C., Grieser, M., Pickrel, J., Stone, E.J., Conway, T., Coombs, D., Staten, L.K., 2006. Formative research in school and community-based health programs and studies: "state of the art" and the TAAG approach. *Health Educ. Behav.* 33 (1), 25–39.
- Gorard, S., See, B.H., Siddiqui, N., 2017. *The Trials of Evidence-Based Education*. Routledge, Abingdon and Oxon.
- Hattie, J.A., Brown, G.T.L., Ward, L., Irving, S.E., Keegan, P.J., 2006. Formative evaluation of an educational assessment technology innovation: developers' insights into assessment tools for teaching and learning (asTTle). *Journal of MultiDisciplinary Evaluation* (5).
- Hoop, T., Ring, H., Siwach, G., Dias, P., Tembo, G., Rothbard, V., Toungui, A., 2020. Midline Report for the Mixed-Methods Cluster-Randomized Controlled Trial of Impact Network's eSchool 360 Model in Rural Zambia. American Institutes for Research. Available at: <https://www.air.org/search?search=/resource%20midline%20report%20mixed%20methods%20cluster%20randomized%20controlled%20trial%20impact%20network%20s%20eschool>.
- IES, 2014. Assessing Attrition Bias – Addendum. Institute of Education Sciences (IES). Available at: https://ies.ed.gov/hcee/wwc/Docs/ReferenceResources/wwc_attrition_v3.0.pdf.
- Maxwell, B., Stiell, B., Stevens, A., Demack, S., Coldwell, M., Wolstenholme, C., Reaney-Wood, S., Lortie-Forgues, H., 2021. Review: Scale-Up of EEF Efficacy Trials to Effectiveness Trials. Education Endowment Foundation (EEF).
- Pawson, R., Tilley, N., 1997. *Realistic Evaluation*. Sage, London.
- Plano Clark, V.L., Schumacher, K., West, C., Edrington, J., Dunn, L.B., Harzstark, A., Melisko, M., Rabow, M.W., Swift, P.S., Miaskowski, C., 2013. Practices for embedding an interpretive qualitative approach within a randomized clinical trial. *J. Mix. Methods Res.* 7 (3), 219–242.
- Scriven, M., 1981. *Evaluation Thesaurus*, third ed. Edgepress, California.
- Scriven, M., 1996. Types of evaluation and types of evaluator. *Eval. Pract.* 17 (2), 151–161. <https://doi.org/10.1177/109821409601700207>.
- Sherry, J.L., 2013. Formative research for STEM educational games: lessons from the children's television workshop. *Z. für Psychol.* 221 (2), 90–97. <https://doi.org/10.1027/2151-2604/a000134>.
- Song, M.-K., Sandelowski, M., Happ, M.B., 2010. Current practices and emerging trends in conducting mixed methods intervention studies in the health sciences. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 725–747.
- Staten, L.K., Birnbaum, A.S., Jobe, J.B., Elder, J.P., 2006. A typology of middle school girls: audience segmentation related to physical activity. *Health Educ. Behav.* 33 (1), 66–80.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B., 2012. Broadening the Range of Designs and Methods for Impact Evaluations. DFID Working (Chapter 38). Department for International Development (DFID).
- Thyer, B.A., 2015. A bibliography of randomized controlled experiments in social work (1949–2013): Solvitur Ambulando. *Res. Soc. Work. Pract.* 25 (7), 753–793.
- White, H., 2008. Of probits and participation: the use of mixed methods in quantitative impact evaluation. *IDS Bull.* 39 (1), 98–109.
- White, H., 2010. A contribution to current debates in impact evaluation. *Evaluation* 16 (2), 153–164.
- White, H., 2013. An introduction to the use of randomised control trials to evaluate development interventions. *J. Dev. Effect.* 5 (1), 30–49.
- White, H., 2014. Current challenges in impact evaluation. *Eur. J. Dev. Res.* 26, 18–30. <https://doi.org/10.1057/ejdr.2013.45>.
- White, H., 2019. The twenty-first century experimenting society: the four waves of the evidence revolution. *Palgrave Communications* 5, 47. <https://doi.org/10.1057/s41599-019-0253-6>.
- White, H., 2022. The Unfinished Evidence Revolution: Riding the Five Waves. CEDIL Methods Working Paper 8. Centre of Excellence for Development Impact and Learning (CEDIL).
- White, H., Masset, E., 2003. *Books, Buildings and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana*. World Bank, Washington DC.
- Yang, T., 2018. *Evaluating Development Projects: Exploring a Synthesis Model of the Logical Framework Approach and Outcome Mapping*. PhD thesis. Institute of Development Studies, University of Sussex.

Mixed methods and the Delphi method

Mehdi Taghipoorreynh, University of Applied Science and Technology (UAST), Tehran, Iran

© 2023 Elsevier Ltd. All rights reserved.

Introduction	608
Historical roots and development of the Delphi method	608
Guiding theory and philosophy of the Delphi	608
Types of Delphi methods	609
The forecasting Delphi method	609
The policy Delphi method	609
The Scenario Delphi	609
The value of Delphi technique	609
Characteristic of Delphi variants	610
Number of rounds	610
Delphi sample size	610
Quantitative, qualitative and mixed methods Delphi	610
The value of Delphi as part of mixed methods research	611
Improving integration in mixed methods research	611
Enhancing mixed methods rigor	612
Issue identification/prioritization	612
Conclusion	612
References	613

Introduction

The Delphi method entails use of a group technique that aims to obtain the most valid and reliable consensus from the panel of skillful and knowledgeable individuals by using a series of questionnaires within a controlled feedback process (Dalkey and Helmer, 1963). Delphi studies have been used in educational settings in predicting trends, standards and in forming guidelines (Green, 2014). Delphi Method works well for educator in learning experiences to prepare students for their future careers, collecting data from students and alumni regarding the curriculum and funding as they might not like to speak in focus groups, and, finally, in setting universities-wide educational goals (Green, 2014). This book chapter gives an overview of the Delphi Method and its value as a part of the mixed methods research.

Historical roots and development of the Delphi method

The Delphi method was first introduced by the RAND Corporation in the 1950s to prepare forecasts of future technological capabilities of potential interest to the military (Brown, 1968; Hsu and Sandford, 2007). The method gained popularity in the 1960s and 1970s and continues to yield solutions dealing with various problems in several diverse fields, for example, corporate management (Linstone and Turoff, 2002), information systems (Schmidt et al., 2001), E-commerce (Addison, 2003), marketing (Hayes, 2007), internet banking (Bradley and Stewart, 2003), health research (De Meyrick, 2003), and tourism research (Kaynak and Marandu, 2006).

The Delphi method has been modified over time to serve a variety of purposes such as defining a concept, setting priorities, and identifying best policies (Okoli and Pawlowski, 2004; Yousuf, 2007). Given these modifications, it is instructive to know there is a variant called Consensus Delphi, which follows the original format of the classical Delphi; there is also the Disagreement Delphi which is designed to generate discussion around topics that do not initially meet consensus (Hasson et al., 2000). These modifications¹ and different subvariants led Delphi Method to find its way into sociology, psychology, and political science.

Guiding theory and philosophy of the Delphi

Pragmatism was informed by John Dewey, who believed that social science research should focus on research objectives and problems to reflect real-world practices and decision-making (Kirk and Reid, 2002). Pragmatism is a philosophical approach that entails

¹Additional discussion of variants is offered in the "Types of Delphi Methods" subsection of this chapter.

assess in theories and beliefs based on their practical applications. The approach is based on the belief that the best research methods are the ones that help to answer the research question most effectively (Dalkey and Helmer, 1963). Pragmatism is evident in the Delphi method in the following ways: (a) The Delphi method allows researchers to use diverse approaches and valuing both subjective and objective knowledge, (b) the method is designed for practical and empirical research, (c) its findings must induce a direct influence on policy and informing practice (Alder and Ziglio, 1996).

Types of Delphi methods

Although many researchers classified the Delphi technique as a forecasting procedure, it has many other applications (Linstone and Turoff, 1975). For example, in its current application, the Delphi method was implemented to attain expert opinion, knowledge, and consensus relating to the content validity, and readability of a pool of survey and/or instrument items (Okoli and Pawlowski, 2004). There are several types of Delphi methods recognized widely by the literature:

The forecasting Delphi method

Perhaps the most basic and common Delphic exercise is the Forecasting Delphi Method, which uses numbers to estimate dates, amounts, and values. A typical example relates to Ament's (1970) long-term predictive that solicited estimates from a panel of scientists about scientific breakthroughs in biological and physical technologies, world population growth, space advances, and new weapons systems. Ament repeated this study 5 years later and found relative consistency of the forecast, shift to later median dates of many of the space forecasts and earlier dates of several biological forecasts and similar forecasting behavior as a function of median time in the future.

The policy Delphi method

Policy Delphis, unlike forecasting Delphis, produces qualitative rather than quantitative data. Policy Delphi's goal is not explicitly seeking consensus as it is to explore all the principal pro and con arguments on a particular policy issue (Linstone and Turoff, 1975). It has been used in a variety of situations, such as environmental management (Wright, 2006), social policy (Marchau and van der Heijden, 1998), and health care (De Meyrick, 2003).

The Scenario Delphi

A Scenario Delphi can be defined as a systematic narrative or description of a series of events that may occur in the future. Scenarios are forward-looking and action-oriented and often combine quantitative and qualitative data and information that describes actions over time. Each participant produces two views of the future—the probable and the preferred. One of the effective way to systematically explore varieties of possibilities in the future is the Scenario Delphi Method (Rowe and Bolger, 1991). This is because the agreements and disagreements of experts shed light on the issue under discussion and can yield alternative solutions. The Scenario Delphi typically combines two methods: (a) cluster analysis to categorize quantitative data and (b) thematic analysis to identify and interpret themes in the qualitative data (Tapio et al., 2011).

The value of Delphi technique

One of the main advantages and characteristics of the Delphi method is it affords an opportunity to gather expert input under anonymous circumstances, which can reduce the impacts of peer pressure and dominant individuals or leaders which often is a concern when using group-based techniques to synthesize and collect data (Dalkey, 1967). The anonymity issue is comforted by geographic distance and dispersion of the experts as well as the use of new technology and online communication such as social media, skype, and email to exchange and solicit data. Therefore, certain downsides related to group dynamics such as coercion or manipulation to adopt a certain perspective can be minimized.

Other merits inherent in using the Delphi Method are the ability to provide anonymity to subjects, manage a controlled feedback procedure, and the flexibility to use a variety of statistical methods for analyzing and data interpretation (Ludlow, 1975). These features in the Delphi method variant are used to minimize the shortcomings of conventional means of pooling opinions achieved from a group communication such as noise, impacts of dominant person, and group pressure for conformity (Dalkey, 1967). Noise often occurs in a group interaction that distorts the information and data and deals with personal and/or group interests instead of concentrating on solving a problem (Dalkey, 1967). Lastly, the ability to use techniques and statistical analyses tools is an exercise that further minimizes the potential of group pressure for conformity (Dalkey, 1967). Specifically, statistical analysis helps ensure agreement and consensus generated by each respondent participating in a Delphi research endeavor are well reflected in the final round (Dalkey, 1967). That is, each respondent should experience limited stress, either perceived or real, to adopt and conform to another subject's feedback that perhaps originate from obedience to custom and social values or standing within a profession.

Characteristic of Delphi variants

Number of rounds

Rounds numbers depends upon the reasonable consensus level and sufficient data and information (Delbecq et al., 1975; Ludwig, 1997). Some Delphi studies have achieved consensus after three (Brungs and Jamieson, 2005), two rounds (Robinson et al., 2005), and some others in one round (Kuo and Yu, 1999). Prior studies suggested that two or three rounds should be enough for most exploratory research projects (Skulmoski et al., 2007). Linstone and Turoff (2002) indicated a typical Delphi study will have four phases including (a) exploring, (b) understanding, (c) reasoning for differences and (d) evaluation. They believed three rounds are typically enough to achieve these goals in the general Delphi process.

The trustworthiness of the result of the Delphi method lies in its ability to draw on expertise (Miller, 2001). This can be achieved by the purposeful selection of careful selection of qualified experts for inclusion in the panel (Okoli and Pawlowski, 2004). While the term “expert” can be defined as a person with high experience and knowledge level in a specific area, the literature on the Delphi method provides many diverse ways that experts can be defined. Bowling (1997) warned that emphasis on the term *expert* may draw an individual’s attention to a person who has knowledge on the topic and is not viewed as “professionally expert.” One can however be an expert without having enough knowledge (Crisp et al., 1999). Also, selecting experts only based on their experiences may not make them an expert, as they may not possess some necessary knowledge required (Baker et al., 2006).

Although the term “expert” is not clearly defined, two key approaches have been adapted in the field to identify an expert: (a) self-assessment, and (b) sample criteria. Self-assessment relates to the evaluation of the self as an expert. The reliability of the self-assessment approach came under criticism due to inappropriate self-rating results among experts. As Pill (1971) mentioned, different people have very different ways of rating their expertise.

The second approach concentrates based on certain criteria, normally derived from the objects of the study. Numerous studies used this approach to assess the suitability of experts. For example, Simon (1965) considered the level of knowledge and experiences as criteria. For others, “expertness” is assumed based on membership of a particular organization (Snyder-Halpern, 2002) or group (Campbell et al., 2000). Therefore, the criteria for choosing experts are dependent on the purpose of the study. As noted by Goldschmidt (1975) the designer of the Delphi study must ensure that selected experts will produce much more meaningful responses than any others who fill out the questionnaire. However, there is some agreement that the expert’s “capacity” and “willingness” to participate is critical for any Delphi studies (Goodman, 1987).

Delphi sample size

In most Delphi studies, the criteria for deciding needed sample sizes are not informed by standard statistical procedures (Powel, 2003). A key point here is sample sizes should be informed by the coverage a group of experts have for a given field, the actual expertise of participants, and then consideration of agreement rates. The Delphi Method entails use of a group decision mechanism which requires group consensus toward an object and does not require the statistical sample to represent any population (Okoli and Pawlowski, 2004). If a subgroup of experts disagrees with another group, then sample size might need to be increased to better understand the phenomena of interest. Hence, sample size interacts with sample selection. Moreover, in Delphi Methodology, one is usually attempting to find experts; there will often be few such experts so randomly sampling will, typically, not be applicable (Beretta, 1996). Instead, the representativeness of the sample relates to the evaluated quality of panel members, rather than the number of experts. As such, more often, non-probability sampling techniques, including, convenience, purposeful and snowballing sampling has been implemented.

The literature on the Delphi technique does not clearly identify the sample size of the research (Adler and Ziglio, 1996). There are various opinions on the number of participants ranging from 7 to 15 (Skulmoski et al., 2007), 7 to 12 (Philips, 2000), and 8 to 12 (Cavalli-Sforza and Ortolano, 1984). The type of the sample became under close attention as criteria for sample size (Skulmoski et al., 2007). Some researcher believed a smaller group of expert panels are best fit for homogeneous groups and a larger group of experts are best suited for heterogeneous sample (Briedenhann and Butts, 2006). Since the process itself leads to attrition, Donohoe and Needham (2009) suggested starting with a larger initial panel can be useful to overcome attrition. As Garson (2014) suggested exploratory studies can use smaller panels. Martino (1972) linked the reliability of the technique with a panel of as few as 15 members. He suggested that 5 to 10 carefully selected experts could be used for a homogeneous population.

Quantitative, qualitative and mixed methods Delphi

There are many varieties of the Delphi method that range from having a quantitative orientation to a qualitative one. In the 1950s and 1960s, Delphi studies were conducted primarily by an industry to forecast dates, technological developments, and to estimate unknown parameters (Skulmoski et al., 2007) such as automated language translator, reliable whether forecast, USSR orbital rendezvous and so on. Participants generally were asked to make quantitative assessments of the issues raised. The basic assumption at hand was that a statistical aggregate of experts and skillful individuals’ judgments is more accurate than the judgment of random individuals. Despite the increased interest in using Delphi for forecasting purposes, the evaluation of the accuracy of a quantitative assessment, being a judgment method, has been problematic due to the possible influence of person- and situation-specific biases (Woudenberg, 1991). Researchers began to describe Delphi as “art instead of science.” (Murray, 1979).

In the 1970s, qualitatively-oriented Delphi approaches gradually gained increasing acceptance because of two reasons. First, the usage of the Delphi technique for forecasting has gained less attention, and focus was shifted to the communicational and educational possibilities of Delphi (Goldschmidt, 1975). Second, the evaluation of the accuracy of a quantitative assessment, being a judgment method, has been far from unproblematic due to the possible influence of person- and situation-specific biases (Woudenberg, 1991). The philosophical assumption was that the judgments cannot be measured quantitatively or equated to measurements. This led researchers to think that the Delphi method may not be fit for only quantitative application (Weaver, 1971). Later, the qualitatively-oriented Delphi found its way into the areas of public health and (Alder and Ziglio, 1996; Linstone and Turoff, 1975), and management and organizational development as a mechanism for improving the decision-making process (Hartman and Baldwin, 1995).

The broad qualitative Delphi approach was altered so that it provided several flexible techniques and analyses to interpret the judgments for consensus (Woudenberg, 1991). It can be applied in diverse situations requiring complex thinking and collaboration and to a wide range of complex problems; however, like a quantitatively oriented Delphi application, the reliability of the qualitative Delphi method was also an issue due to similar problem of person- and situation-specific biases (Woudenberg, 1991).

To overcome the reliability and accuracy concerns of any Delphi study, scholars suggested that both qualitative and quantitative perspectives should be used together, and this recommendation has been largely heeded to date (Hasson and Keeney, 2011). Nowadays, researchers believe Delphi is a mixed-methods technique in nature. As is mentioned by Critcher and Goldstone (1998), Delphi studies straddle the divide between qualitative and quantitative methodologies. There is a preponderance of studies that show the usefulness of integrating qualitative and quantitative research in a Delphi process. One mixed-methods Delphi study is Tapio et al. (2011) work involving insightful analyses of the strengths, weaknesses, and pitfalls encountered when combining qualitative and quantitative information in a Delphi process by studying seven disaggregate policy Delphi material conducted in Finland from 1999 to 2008. More recently, Taghipoorreynah and de Run (2020) illustrated the value of mixed methods Delphi in advancing the validity and reliability of an indigenous cultural values instrument. The Delphi study began with a qualitative opening round wherein the experts were asked to respond to open-ended questions concerning the indigenous culture. After generating the data input, the emphasis changed to quantitative analysis and the collection of data to achieve a consensus of the experts' opinions on the emerged values and items. At the same time, the experts were asked to respond to other open-ended questions for clustering purposes. The proposed value dimensions did not load on any quantitative or qualitative values instruments before. In each case, the mixed methods Delphi study has validated researchers' findings through integration and has enabled the understanding of a phenomenon in greater depth. This potential value makes the mixed methods Delphi a suitable tool to produce a holistic picture and to capture the internal organization of complex abstract entities.

The value of Delphi as part of mixed methods research

Improving integration in mixed methods research

The traditional quantitative vs. qualitative debates seem to have ceased recently. Instead, scholars have accepted both approaches' merits (Johnson and Onwuegbuzie, 2004). The quantitative approach assesses causes of objective knowledge or social phenomena, guided by the principle that factors, variables, and concepts should be measured by using objective research methods, rather than being inferred subjectively by using intuition, reflection, or observation (Easterby-Smith et al., 1995). The qualitative approach to research emphasizes the subjective aspects of human activity by concentrating on the meaning, rather than the measurement of objective knowledge or social phenomena (Easterby-Smith et al., 1995). On this basis, the quantitative and qualitative approaches complete each other. But the main question remains as to how researchers can use these two approaches together to gain better results instead of using them separately.

To have a holistic picture, the mixed-methods approach, in which quantitative and qualitative methods are used together, has become a more popular research approach since the 1990s (Creswell, 2008). Various concepts such as multi-method, integrated, and combined research, have been used for labeling such a methodological approach (Creswell, 2008). The main idea of the approach is to mix methods together and specify how to elaborate quantitative and qualitative elements in data collection and analyses, and explicitly explain how quantitative and qualitative aspects relate to each other (Johnson et al., 2007).

Despite the increasing interest in using mixed methods, however, the practice of integrating quantitative and qualitative methods has faced many problems. Most of the justifications for using mixed methods have not been sufficiently considered (Onwuegbuzie et al., 2011). Further, conversely, inconsistencies between qualitative and quantitative data cause unforeseen problems in interpreting research findings (Bryman, 2006).

Delphi is one of the effective methods that allows overcoming the problems in integrating quantitative and qualitative methods (Tapio et al., 2011). One well-known study in determining the value of Delphi as part of mixed methods research is that of the effectiveness of the enforcement of new international regulations on ship emissions by Bloor et al. (2015). Onwuegbuzie et al. (2011) indicated the usefulness of Delphi method as a tool for developing quantitative instruments in mixed methods research. Taghipoorreynah and de Run (2020) illustrated how using Delphi in mixed methods research can increase the sensitivity and reliability of the quantitative instruments. Torres-Alzate et al. (2020) conducted a three-round modified Delphi study using a mixed methods research approach to reach consensus among experts on global health competencies for bacca-aureate nursing students in the USA. In the first round, they revised the original list of competencies. In the second round, they conducted two surveys for the validation of the list and at the last round they conducted a survey to rate the extent to which they thought the competencies

obtained in Round Two were essential for baccalaureate nursing education in the United States. In all cases, researchers did not seek to rehabilitate Delphi as a stand-alone research method, but rather considers the possible limited role of Delphi Groups within multimethod research projects. Particularly they proposed using Delphi in the initial phase of the research can potentially serve several useful functions.

Enhancing mixed methods rigor

Given the claim that the mixed methods research combines and integrate both qualitative and quantitative methods, another approach is needed to measure rigor instead of adopting strategies that qualitative and quantitative researchers use to ensure credibility (Engels and Kennedy, 2007). According to Hasson and Keeney (2011) the term trustworthiness is more appropriate than validity and reliability to measure the appropriateness and effectiveness of a mixed methods research. Onwuegbuzie and Johnson (2006) recommended that validity in mixed research be termed legitimization to use a bilingual nomenclature that can be used by both quantitative and qualitative researchers. They indicated nine types of legitimations including sample integration legitimization, insider-outsider legitimization, weakness minimization legitimization, sequential legitimization, conversion legitimization, paradigmatic mixing legitimization, commensurability legitimization, multiple validities legitimization, and political legitimization.

Delphi enhances the trustworthiness of a mixed methods research by increasing the credibility, dependability, transferability, and confirmability (Lincoln and Guba, 1985). Delphi improve the credibility of a mixed method research by member check and feedback given to the panelist (Engels and Kennedy, 2007). It increases dependability by carefully selection and including a range and representative sample of panelists (Cornick, 2006). Delphi enhances conformability by concentrating on the accuracy at all stages of the Delphi study such as the appropriateness of the technique to address the problem selected, choice of panel members, data collection method and procedures, identification of justifiable agreement and consensus levels (Powell, 2003). It establishes transferability through the use of verification of the applicability of Delphi results (Kennedy, 2004).

Delphi also enhances legitimization of the mixed methods research through improving construct and content validity. Okoli and Pawlowski (2004) found that the Delphi method contributes to construct validity by careful research design and by pretesting. The philosophy behind Delphi technique is that group thinking can enhance content validity. The experts' knowledge and the number of successful rounds can increase content validity (Goodman, 1987). Hong et al. (2019) found that e-Delphi improved the content validity of the mixed methods appraisal tool.

Issue identification/prioritization

Using the Delphi method in the early stages of mixed methods research can help identify or prioritize the issues under study. There are several mixed methods studies in the field of Information System (IS) that have used this feature in the early stages of research to identify key managerial concerns in IS projects. For example, Lai and Chung (2002) identified a prioritized list of international data communications activities essential for multinational corporations in information exchange management to control and implement global business strategies. Scale developers also use Delphi at the early stage of research. For instance, Taghipoorreyneh and de Run (2020) used Delphi at the beginning of the mixed methods research to identify a list of values to develop instruments responsive to indigenous culture. Another example relates to Onwuegbuzie et al. (2010) work in developing quantitative instrument using mixed methods research.

Conclusion

This chapter gives researchers invaluable resources to use the Delphi method in the mixed methods research for application in educational research. The study indicates that Delphi method is the mixed methods in nature and using it as a part of the mixed methods research enhances research projects integration and rigor. Although the use of Delphi in mixed methods research may seem appealing to the educational researchers, there are some problems and limitations that need to be addressed. (1) Delphis are time-consuming. We all know that new technologies have enabled fast communication. But even if new methods are implemented in conducting the Delphi, each round can take several weeks. Because receiving information and data from experts, analyzing, and categorizing them, sending them back to receive feedback, and finally, achieving agreement requires a lot of time. This issue can be a more serious problem in mixed methods research because these methods are inherently time-consuming. (2) The technique is highly depending on to design characteristic, particularly the clarity of the research question and the level of participants' expertise. The respondents may misunderstand the brief written inputs of the moderator or research questions. (3) The experts panel might be too homogeneous or heterogeneous, therefore producing a skewed data set. (4) There is no specific criterion for measuring the level of agreement. The minimal consensus does not necessarily bring good results. (5) Delphi studies also suffer high rates of attrition. It is an iterative process which can be time consuming and therefore experts may not fully engage in the process. To reduce attrition, it is not uncommon to offer monetary incentives. (6) Finally, results are difficult to analyze and needs skill in written communication to get feedback from participants.

References

- Addison, T., 2003. E-commerce project development risks: evidence from a Delphi survey. *Int. J. Inf. Manag.* 23 (1), 25–40.
- Alder, M., Ziglio, E., 1996. *Gazing Into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*. Jessica Kingsley, London, England.
- Ament, R.H., 1970. Comparison of Delphi forecasting studies in 1964 and 1969. *Futures* 2 (1), 35–44.
- Baker, D.P., Day, R., Salas, E., 2006. Teamwork as an essential component of high-reliability organizations. *Health Serv. Res.* 41 (4), 1576–1598.
- Beretta, R., 1996. A critical review of the Delphi technique. *Nurse Res.* 3 (4), 79–89.
- Bloor, M., Sampson, H., Baker, S., Dahlgren, K., 2015. Useful but no Oracle: reflections on the use of a Delphi Group in a multimethods policy research study. *Qual. Res.* 15, 57–70.
- Bowling, A., 1997. *Research Methods in Health*. Open University Press, Buckingham.
- Bradley, L., Stewart, K., 2003. A Delphi study of Internet banking. *Market. Intell. Plann.* 21 (5), 272–281.
- Briedenhann, J., Butts, S., 2006. Challenges and shortcomings: use of the Delphi technique in rural tourism project evaluation. *Curr. Issues Tourism* 9 (2), 171–190.
- Brown, B.B., 1968. *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*. Rand Corp Santa Monica CA.
- Brungs, A., Jamieson, R., 2005. Identification of legal issues for computer forensics. *Inf. Syst. Manag.* 22 (2), 57–66.
- Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Res.* 6 (1), 97–113.
- Campbell, S.M., Cantrill, J.A., Roberts, D., 2000. Prescribing indicators for UK general practice: Delphi consultation study. *BMJ* 321, 425–428.
- Cavalli-Sforza, V., Ortolano, L., 1984. Delphi forecasts of land-use—transportation interactions. *J. Transport. Eng.* 35 (4), 324–339.
- Cornick, P., 2006. Nitric oxide education survey—use of a Delphi survey to produce guidelines for training neonatal nurses to work with inhaled nitric oxide. *J. Neonatal Nurs.* 12 (2), 62–68.
- Creswell, J., 2008. Mixed methods research. In: Given, L.M. (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE, Thousand Oaks, pp. 526–529.
- Crisp, J., Pelletier, D., Duffield, C., Nagy, S., Adams, A., 1999. It's all in a name: when is a Delphi study not a Delphi Study? *Aust. J. Adv. Nurs.* 16 (3), 32–37.
- Critcher, C., Gladstone, B., 1998. Utilizing the Delphi technique in policy discussion: a case study of a privatized utility in Britain. *Publ. Adm.* 76, 431–449.
- Dalkey, N.C., 1967. *Delphi*. The RAND Corporation, Santa Monica, CA.
- Dalkey, N.C., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. *Manag. Sci.* 9 (3), 458–467.
- De Meyrick, J., 2003. The Delphi method and health research. *Health Educ.* 103 (1), 7–16.
- Delbecq, A.L., Van de Ven, A.H., Gustafson, D.H., 1975. *Group Techniques for Program Planning*. Scott-Foresman and Co, Glenview, IL.
- Donohoe, H.M., Needham, R.D., 2009. Moving best practice forward: Delphi characteristics, advantages, potential problems and solutions. *Int. J. Tourism Res.* 11 (5), 415–438.
- Easterby-Smith, M., Thorpe, R., Lowe, A., 1995. *Management Research an Introduction*. SAGE publications.
- Engels, T.C., Kennedy, H.P., 2007. Enhancing a Delphi study on family-focused prevention. *Technol. Forecast. Soc. Change* 74 (4), 433–451.
- Garson, G.D., 2014. *The Delphi Method in Quantitative Research (Statistical Associates Blue Book Series)* [Kindle Version]. Retrieved from: [Amazon.com](https://www.amazon.com).
- Goldschmidt, P.G., 1975. Scientific inquiry or political critique? Remarks on Delphi assessment, expert opinion, forecasting, and group process by H. Sackman. *Technol. Forecast. Soc. Change* 7 (2), 195–213.
- Goodman, C.M., 1987. The Delphi technique: a critique. *J. Adv. Nurs.* 12 (6), 729–734.
- Green, R.A., 2014. The Delphi technique in educational research. *Sage Open* 4 (2), 1–8.
- Hartman, F., Baldwin, A., 1995. Using technology to improve the Delphi method. *J. Comput. Civ. Eng.* 9 (4), 244–249.
- Hasson, F., Keeney, S., 2011. Enhancing rigor in the Delphi technique research. *Technol. Forecast. Soc. Change* 78 (9), 1695–1704.
- Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* 32 (4), 1008–1015.
- Hayes, T., 2007. Delphi study of the future of marketing of higher education. *J. Bus. Res.* 60 (9), 927–931.
- Hong, Q.N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., et al., 2019. Improving the content validity of the mixed methods appraisal tool: a modified e-Delphi study. *J. Clin. Epidemiol.* 111, 49–59.
- Hsu, C., Sandford, B., 2007. The Delphi technique: making sense of consensus. *Pract. Assess. Res. Eval.* 12 (10), 1–8.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26.
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A., 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1 (2), 112–133.
- Kaynak, E., Marandu, E.E., 2006. Tourism market potential analysis in Botswana: a Delphi study. *J. Trav. Res.* 45 (2), 227–237.
- Kennedy, H.P., 2004. Enhancing Delphi research: methods and results. *J. Adv. Nurs.* 45 (5), 504–511.
- Kirk, S.A., Reid, W.J., 2002. *Science and Social Work*. Columbia University Press, New York, NY.
- Kuo, N.W., Yu, Y.H., 1999. An evaluation system for national park selection in Taiwan. *J. Environ. Plann. Manag.* 42 (5), 735–743.
- Lai, V.S., Chung, W., 2002. Managing international data communications. *Commun. ACM* 45 (3), 89–93.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage Publications, Newbury Park, CA, USA.
- Linstone, H.A., Turoff, M., 1975. *The Delphi Method: Techniques and Applications*. Addison-Wesley, Reading, MA.
- Linstone, H.A., Turoff, M., 2002. *The Delphi Method: Techniques and Applications E-Book*. Jersey Institute of Technology, Newark, New Jersey.
- Ludlow, J., 1975. Ill. B. 2. Delphi Inquiries and Knowledge Utilization. *The Delphi method: Techniques and applications* 102.
- Ludwig, B., 1997. Predicting the future: have you considered using the Delphi methodology? *J. Ext.* 35 (5).
- Marchau, V.A.W.J., Van der Heijden, 1998. Policy aspects of driver support systems implementation: results of an international Delphi study. *Transp. Policy* 5 (4), 249–258.
- Martino, J.P., 1972. *Technological Forecasting for Decision Making*. American Elsevier, New York, NY.
- Miller, G., 2001. The development of indicators for sustainable tourism: results of a Delphi survey of tourism researchers. *Tourism Manag.* 22 (4), 351–362.
- Murray, T.I., 1979. Delphi methodologies: a review and critique. *Urban Syst.* 4, 153–158.
- Okoli, C., Pawlowski, S.D., 2004. The Delphi method as a research tool: an example, design considerations and applications. *Inf. Manag.* 42 (1), 15–29.
- Onwuegbuzie, A.J., Johnson, R.B., 2006. The validity issue in mixed research. *Res. Sch.* 13 (1), 48–63.
- Onwuegbuzie, A.J., Johnson, R.B., Collins, K.M., 2011. Assessing legitimacy in mixed research: a new framework. *Qual. Quant.* 45 (6), 1253–1271.
- Onwuegbuzie, A.J., Rebecca, M.B., Judith, A.N., 2010. Mixed research as a tool for developing quantitative instruments. *J. Mix. Methods Res.* 4 (1), 56–78.
- Phillips, R., 2000. New Applications for the Delphi Technique, 2. Annual Pfeiffer and Company, pp. 191–196.
- Pill, J., 1971. The Delphi method: substance, context, a critique and an annotated bibliography. *Soc. Econ. Plann. Sci.* 5 (1), 57–71.
- Powell, C., 2003. The Delphi technique: myths and realities. *J. Adv. Nurs.* 41 (4), 376–382.
- Robinson, Q.M., Collins, C.J., Oreg, S., 2005. The effects of recruitment message specificity on applicant attraction to organizations. *J. Bus. Psychol.* 19 (3), 319–340.
- Rowe, G., Bolger, F.G., 1991. Delphi: a re-evaluation of research and theory. *Technol. Forecast. Soc. Change* 39, 235–251.
- Schmidt, R.C., Lyytinen, K., Keil, M., Cule, P., 2001. Identifying software project risks: an international Delphi study. *J. Manag. Inf. Syst.* 17 (4), 5–36.
- Simon, H.A., 1965. *The Shape of Automation*. Harper and Row, New York, NY.
- Skulmoski, G.J., Hartman, F.T., Krahn, J., 2007. The Delphi method for graduate research. *J. Inf. Technol. Educ.* 6, 1–21.
- Snyder-Halpern, R., 2002. Indicators of organizational readiness for clinical information technology systems innovation: a Delphi study. *Int. J. Med. Inf.* 63 (3), 179–204.
- Taghipoorreynah, M., de Run, E.C., 2020. Using mixed methods research as a tool for developing an indigenous cultural values instrument in Malaysia. *J. Mix. Methods Res.* 14 (3), 403–424.

- Tapio, P., Paloniemi, R., Varho, V., Vinnari, M., 2011. The unholy marriage? Integrating qualitative and quantitative information in Delphi processes. *Technol. Forecast. Soc. Change* 78 (9), 1616–1628.
- Torres-Alzate, H.M., Wilson, L.L., Harper, D.C., Ivankova, N.V., Heaton, K., Shirey, M.R., 2020. Essential global health competencies for baccalaureate nursing students in the United States: a mixed methods Delphi study. *J. Adv. Nurs.* 76 (2), 725–740.
- Weaver, W.T., 1971. The Delphi forecasting method. *Phi Delta Kappan* 52, 267–271.
- Woudenberg, F., 1991. An evaluation of Delphi. *Technol. Forecast. Soc. Change* 40 (2), 131–150.
- Wright, T.S.A., 2006. Giving “teeth” to an environmental policy: a Delphi study at Dalhousie University. *J. Clean. Prod.* 14 (9–11), 761–768.
- Yousuf, M.I., 2007. Using experts’ opinions through Delphi technique. *Pract. Assess. Res. Eval.* 12 (4), 1–8.

Mixed methods research design practices to address complexity in education

Cheryl N. Poth^a and **Emma P. Bullock^b**, ^a Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, AB, Canada; and ^b Department of Mathematics and Statistics, College of Science and Engineering Technology, Sam Houston State University, Huntsville, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	615
Situating complexity science in education and mixed methods research	616
Illustrating four mixed methods research design decision transformations	617
Representations of the nested structure of CAS informs initial integration decisions	617
Attention to interactions of the CAS informs data source and sampling decisions	620
Adaptations to emergent CAS conditions inform integration procedural decisions	621
Transparency of accounts of CAS outcomes informs credible interpretation decisions	623
Implications for complexity-informed mixed methods research in education design practices	624
Concluding thoughts	624
Acknowledgments	625
References	625

Glossary

Complex adaptive system A complex system with individual agents that are not fixed (i.e., individual agents that learn and adapt in response to interactions with other agents of the complex system) that defy simplistic analyses of cause and effect, where the whole is greater than the sum of its parts (Bullock, 2017; Weaver, 1948)

Complexity lens A view of how to see, understand, and influence research in which we assume nonlinearity and which is reflective of the principles of complexity science

Complexity science A collective of theories and conceptual tools used to study complex systems

Complex systems Phenomena which have the following characteristics nested structure(s) (i.e., composed of and/or comprising other complex systems), self-organizing units (i.e., spontaneously arising structures), bottom-up emergence (i.e., often surprising qualities and abilities that arise due to individual agent interactions and not over-arching governing structures), ambiguously bounded (i.e., they are open in that they interact with their surroundings), organizationally closed (i.e., they are enduringly stable even while dynamically changing), and far-from equilibrium (i.e., they do not operate in balance, rather they are continuously evolving in reaction to disturbances within and without) (Davis and Sumara, 2006). Of the two types of complex systems (i.e., complex physical systems and complex adaptive systems) the latter is relevant to the current article

Integrated findings Results that are produced by the requisite integration of qualitative and quantitative research in mixed methods designs (Poth, 2018b)

Mixed insights Also often referred to as meta-inferences, represent the interpretation of the requisite integration of qualitative and quantitative research in mixed methods designs (Poth et al., 2020)

Mixing purpose The rationale for integrating qualitative and quantitative research in mixed methods designs

Mixed methods research Research distinguished by the requisite integration of quantitative and qualitative research, and assumes that their collective contribution mitigates inherent weaknesses in either research (Poth, 2018a)

Introduction

Recognizing complexity in our educational research settings requires new ways of designing mixed methods studies. New approaches to design have arisen to address the challenges researchers face with navigating dynamic conditions in their education research settings (Koopmans and Stamovlasis, 2016). A typical response by researchers working under conditions of complexity involves attempts to reduce, control, or simply ignore its effects (Poth, 2018a). Researchers benefit from design approaches that help them conceptualize the focus of their educational research (i.e., the phenomena they are studying) to more accurately represent the dynamic realities they will need to account for in their designs. There is a pressing need for such methodological advances if we are to adequately equip researchers for studying complex educational problems using mixed methods designs. This article aims to do just that by describing and illustrating four mixed methods research design practices to address complexity in educational studies.

Mixed methods research requires the integration of aspects of both quantitative and qualitative research (e.g., data, perspectives, findings) and assumes that their collective contribution mitigates inherent weaknesses in either type of data (Poth, 2018a). For general introductions, see also the articles specific to using mixed methods in educational research (Onwuegbuzie and Marlette, Section 11) and mixed methods designs (Tajima, Section 11). The usefulness of mixed methods research for solving complex societal issues is well established (e.g., Mertens, 2015; Mertens et al., 2016; Poth, 2018a) yet researchers continue to lack practical guidance for addressing complexity and its “messiness” in educational settings. According to Plano Clark and Ivankova, the concept of messiness “recognizes the inherent complex, dynamic, and undetermined nature of mixed methods research practice” (2016, pp. 276–277). By providing practical design guidance in four practice areas, we begin to equip researchers in education with the necessary understandings and skills to make explicit their complexity-informed design decisions.

Key to the novel contribution of this article is the use of complexity science for informing the mixed methods research designs of educational settings as complex adaptive systems (CASs). Complexity science is a collective of theories and conceptual tools used to study complex systems. CAS is a complex system with individual agents that are not fixed (i.e., individual agents that learn and adapt in response to interactions with other agents of the complex system) that defy simplistic analyses of cause and effect, and where the whole is greater than the sum of its parts (Bullock, 2017; Weaver, 1948). Researchers in education need help navigating the changeable conditions that are characteristic of studies of CAS. To that end, this article offers important guidance for helping researchers manage their mixed methods research designs to assimilate emerging understandings of the educational CAS studied.

The impetus for this article was the lessons that emerged from our joint reflection upon our experiences and their potential usefulness for others. In this work, we were building upon the initial descriptions of six adaptive design practices by Poth (2018a), the conceptualization of educational research teams as CAS (Poth, 2019), and Bullock’s innovative practice transformations for addressing complexity in education (Bullock, 2017). The details of these previous works are not important because just like CAS, our own thinking and practices have since evolved. Importantly, while we are not alone in our use of complexity science to frame educational phenomena as CAS (Davis and Samara, 2006; Jacobson et al., 2019; Koopmans and Stamovlasis, 2016), its nascent application to mixed methods research holds great promise and warrants attention (Kallemeyn et al., 2020). Our novel contribution in this article involves providing practical design guidance for mixed methods researchers to address complexity in education in a way that is accessible to a range of expertise in mixed methods research (i.e., from newcomers to established experts).

This article shows you how to transform mixed methods research design practices to address complexity in education by focusing on four areas of design decisions and their unique contributions to making explicit the point(s) of integration, defining the perspectives and contexts captured by the study, type and sequencing of data procedures, and accounts of how the outcomes were derived. To help researchers bridge theory and practice, we will interweave discussions of two published mixed methods study examples from our own work. Key theoretical and practical implications of the design conclude the article with guidance for managing the complexity inherent in educational settings and study designs. We begin with an orientation to situate the reader to complexity theory in education and mixed methods.

Situating complexity science in education and mixed methods research

Many of society’s most pressing and pernicious social problems involve multiple, interacting influences with no known solution or established methods and expertise for studying the interrelated and dynamic contexts in which they take place. Thus, it is logical that novel solutions to such societal problems are likely to be derived from interdisciplinary perspectives and methodological innovations. A growing body of literature explores the intersections and yet unrealized potential of complexity science, educational settings, and mixed methods research (e.g., Kallemeyn et al., 2020; Poth, 2018a,b). As a theoretical framework, complexity science can provide important guidance to study CAS encountered by educational researchers. Studying educational CAS requires us to reconsider our assumptions about the stability of the research conditions and transform our research design practices accordingly. Complexity science as an ontology or theoretical frame of reference for research on, or within, educational settings offers an alternative to understanding phenomena by examining its parts by instead focusing on organic and holistic systems and their interactions (Lewin, 1993). Complexity-informed research and clinical practice transformations are occurring across diverse disciplines (e.g., business, evaluation, and health).

Using a complexity lens to inform mixed methods practice offers researchers a better way to make sense of the realities of educational settings and transform their approach to research design, making it more realistic. Among the key benefits for mixed methods researchers of complexity-informed approaches is their increased capacity to appropriately respond and adapt design practices to evolving conditions (Poth, 2018a,b). Understandably, this thinking permeates the mixed methods research process from initial conceptualization of the design to data collection, analysis, and integration as well as the reporting of study outcomes.

Kallemeyn et al. (2020) conceptualized four ways of thinking about complexity theory and its implications for mixed methods research: (1) describing complexity theory as an alternative paradigm to scientific inquiry, (2) applying complexity theory to the study of complex systems, (3) using complexity theory as a theoretical framework, and (4) treating the field of mixed methods as a complex adaptive system itself. We agree with our colleagues that these conceptualizations of complexity theory and mixed methods research point to a “mutually informing relationship between study phenomena and methods” (Kallemeyn et al., 2020, p. 289). Furthermore, similar to Byrne and Callaghan (2014), we recognize that the state of the field is such that there are not yet well-defined and well-prescribed methods for conducting research through a complexity-informed lens. It is with this article that we contribute important methodological advancements for approaching the design of studies in education informed by complexity science to more accurately study educational institutions and settings as the interrelated and dynamic contexts that they are.

Illustrating four mixed methods research design decision transformations

To learn how to approach the design of CAS informed by complexity science can be challenging, yet examining real world dilemmas and opportunities can provide a logical starting place. In the following sections, we use two published mixed methods research studies from our own programs of research to illustrate the contributions of four complexity-informed design decisions to the mixed methods research process. Through the lens of complexity theory Bullock (2017) study sought to explain the role the school leader plays in students' mathematics achievement by integrating survey data and focus group interviews. Through applying a systems perspective Poth et al. (2020) conceptualized the doctoral course in which the learners were enrolled as a CAS to generate a more complete understanding of the impactful experiences learners attributed to their competency gains by integrating existing qualitative and quantitative data from course artifacts.

We noted in our reflection that both the commonalities and differences between our studies were worthy of attention: First, both studies took place in educational settings and benefitted from conceptualizing the interrelated and dynamic contexts as CAS. Second, each of the studies drew upon different data source combinations to generate novel understandings that could not be accessible by either qualitative or quantitative research alone. Third, as we each grappled to make explicit our design adaptations, we came to use the metaphors of braiding or twisting to describe how our decisions were interwoven in often unexpected, non-linear, and changeable ways. It is our hope that by reading about two distinct studies, researchers can imagine the infinite possibilities that exist for their own work.

We organize our descriptions of our unique complexity-informed approaches by four interdependent design decisions visually represented in Fig. 1: Initial design decisions, evolving data decisions, ongoing analytical decisions, and authentic reporting decisions. Each of the four descriptions of the mixed methods research design decisions makes explicit its unique contribution to the mixed methods research process, the opportunities of CAS conceptualization to the mixed methods research process, and the practical guidance afforded by the mixed methods research design practice transformation for researchers (see Table 1 for summary).

Representations of the nested structure of CAS informs initial integration decisions

The need for integration, often represented by a mixing purpose statement, is essential to determine if mixed methods research is a suitable approach for the study (Creswell and Plano Clark, 2018; Poth, 2018a). The nested structure of the CAS can help researchers relate the point(s) of integration appropriate for addressing the study's mixing purpose by identifying the unit of analysis. We operationalize a unit of analysis as the focus for the requisite integration for a mixed methods study. When a researcher conceptualizes their study context as a CAS, they intensify the need for appropriately identifying the unit of analysis within the

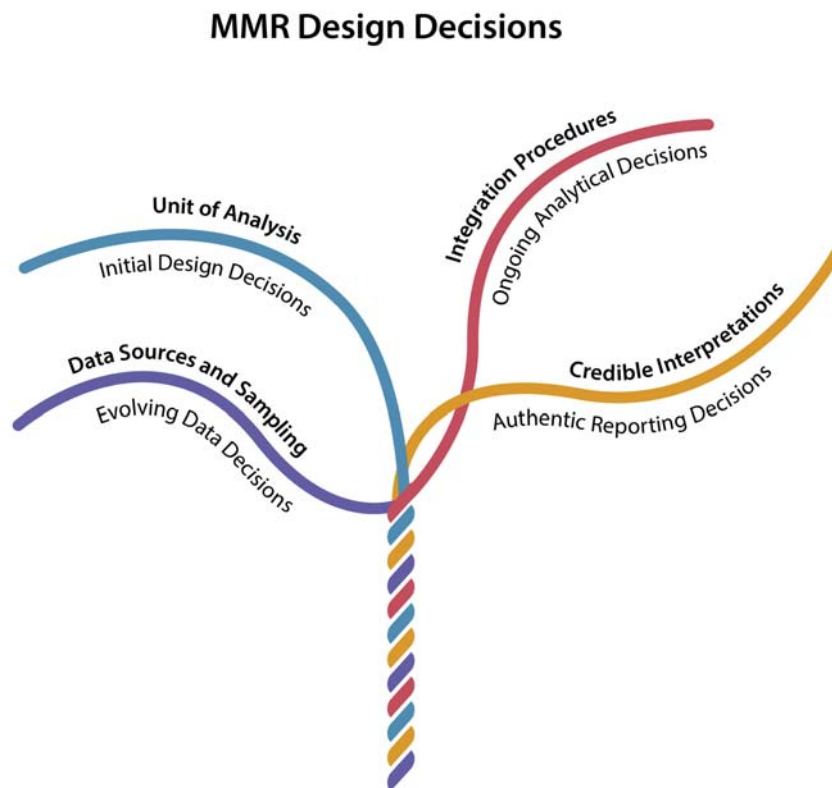


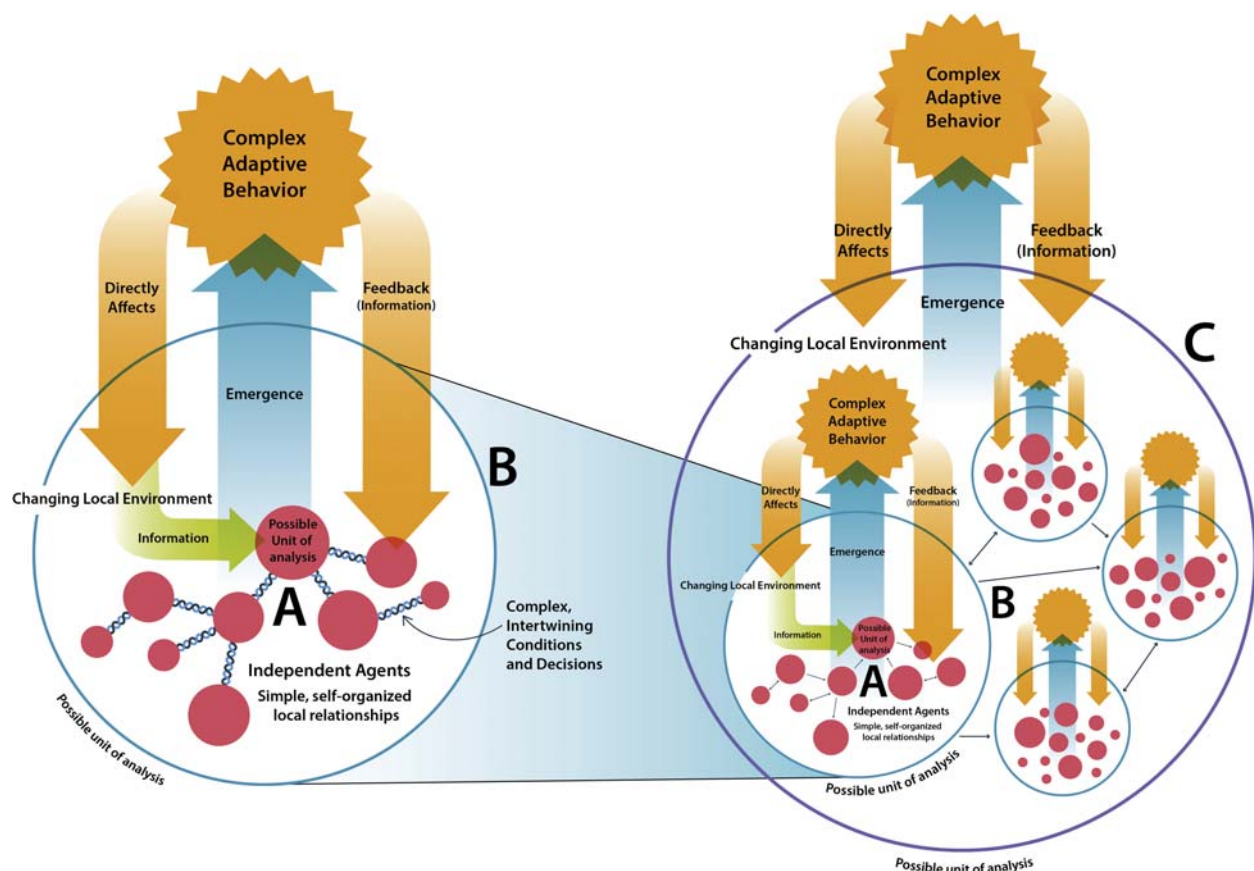
Fig. 1 The braiding of four key design decisions when conceptualizing educational study contexts as CASs.

Table 1 Four key mixed methods research design decisions and complexity-informed practice transformations.

<i>Mixed methods research design decisions</i>	<i>Unique contribution to the mixed methods research process</i>	<i>Opportunities of CAS conceptualization to the mixed methods research process</i>	<i>Guiding mixed methods research design practice transformations</i>
Initial design	Identification of unit of analysis	Relates point(s) of integration to guide the study	Represent the nested CAS structures
Evolving data	Selection of data sources and sampling	Defines the perspectives and contexts captured by the study	Attend to the dynamic CAS interactions
Ongoing analytical	Description of integration procedures	Describes the type and sequencing of data procedures	Adapt to emergent CAS conditions
Authentic reporting	Credible interpretations	Explains how outcomes were derived	Convey transparent accounts of CAS studies

nested and interdependent systems. Sound near impossible? How is this done? Several options exist for units of analysis within CAS found in educational institutions and settings and distinguishing among the embedded levels can be challenging. For example, at the level of a school, the agents consist of stakeholders (e.g., administration, teachers, parents, students) which are influenced by—and in turn, influence—a set of initial conditions (e.g., demographics, school setting, teacher quality, supports for teachers, supports for students, school schedules, data usage). These agents are all dynamically and continuously changing and adapting as different stakeholders move in and out of the system or exercise their individual agency, interact with, and influence each other. This system of the school, as a whole, undergoes an emergent process in which phenomena like school culture, academic learning, and traditions constantly evolve, moving in measurable trajectories which can follow predictable patterns as groups within the system continuously self-form and interact with each other.

To illustrate what we mean by nested structures in educational settings, while an individual school may represent the whole of our CAS, we know that there are influences on that nested system from larger CASs such as districts and states (or border regions). Within the nested structures, the individual school is an embedded complex agent (Jacobson et al., 2019) and is both influencing to and influenced by smaller CASs within an individual school such as grade level cohorts, individual classrooms, and even individual students (Pennings and Mainhard, 2016; van Vondel et al., 2016). Fig. 2 represents this nested structure to help researchers

**Fig. 2** A visual representation of nested structures within CAS for guiding the selection of the unit of analysis for the study.

distinguish among the larger and smaller CASs to identify their unit of analysis for the initial mixed methods research design decisions. The zoomed-out graphic on the right (labeled C) represents the larger CAS such as a school or institution, whereas the smaller zoomed-in graphic on the left (labeled B) represents the nested systems such as an individual class or an individual stakeholder. It is also possible that the larger CAS is part of an even larger CAS, such as schools that are part of the same district. By distinguishing the nested structures involved in your study helps identify the unit of analysis you intend to study. It may be that researchers are trying to decide which nested system to focus on which is where the interdependencies among the decisions exist—perhaps looking at the data sources (see next section) can help researchers. We recommend that newcomers to mixed methods research focus on a single unit of analysis whereas more experienced researchers may decide to focus on more than one.

You might now be asking, “How does identifying the study’s unit of analysis within the context of a CAS help inform the initial design decisions relating points of integration?” If we look at our two studies; we see the nested structures for each of our studies in our visual representations similar to Fig. 2. For Bullock’s (2017) study in Fig. 3, we see that the smaller CAS (labeled B) represents the principal’s school as nested within the larger CAS (labeled C) representing the school district. If you remember the study purpose was to explain the role the school leader plays which suggests the point of integration for this study was school principals as the unit of analysis. Conceptualizing the study as a CAS allows the identification of the unit of analysis at the individual level of a school principal (labeled A) within the context of their school’s complex initial conditions (labeled B) and intertwined with the various stakeholders within the school district system (labeled C) to more authentically predict and explain the emergent property of school-wide student mathematics achievement. With the selection of the individual principals as the unit of analysis, Bullock could focus on collecting both quantitative and qualitative data sources to generate a more complete understanding of the influences on, and by, the principal within the CAS of their particular schools to explain the principals’ role in promoting student mathematics achievement. Together these emerging understandings of the unit of analysis of the school principal helped Bullock to consider the data collection sequencing as appropriate for her explanatory-focused mixing purpose of the role of the school leader.

In contrast, in Poth’s et al. (2020) study in Fig. 4, we see that the smaller CAS (labeled B) represents the doctoral class as nested within the larger CAS (labeled C) representing the university. If you remember the study purpose was to generate a more complete understanding of the competency gains learners in the class reported through the integration of descriptions of impactful course experiences which suggests the point of integration for this study was the class as the unit of analysis. Conceptualizing the study as a CAS allows the identification of the individual learners (labeled A) as nested within the unit of analysis at the group level of the class (labeled B) and intertwined with the various stakeholders within the university system (labeled C) to more authentically

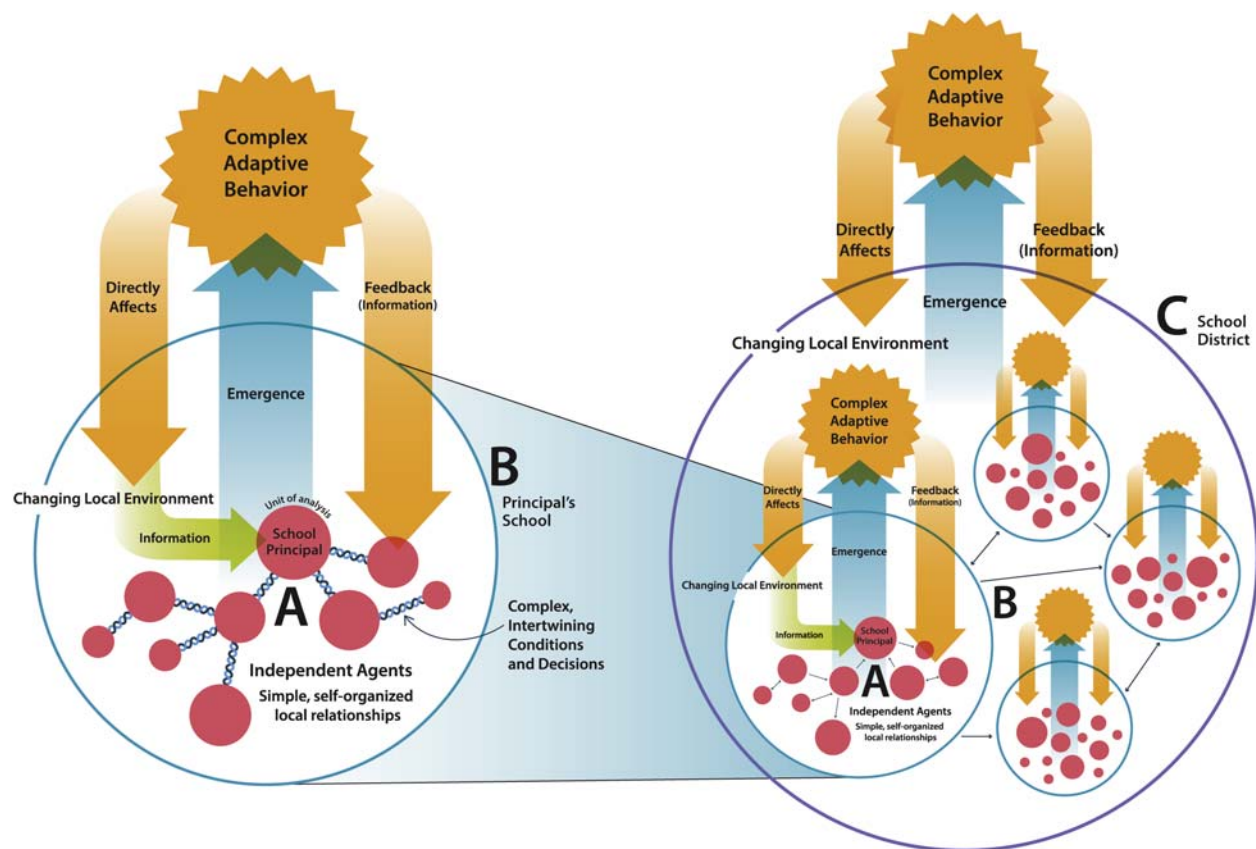


Fig. 3 A visual representation of nested structures within CAS for Bullock (2017).

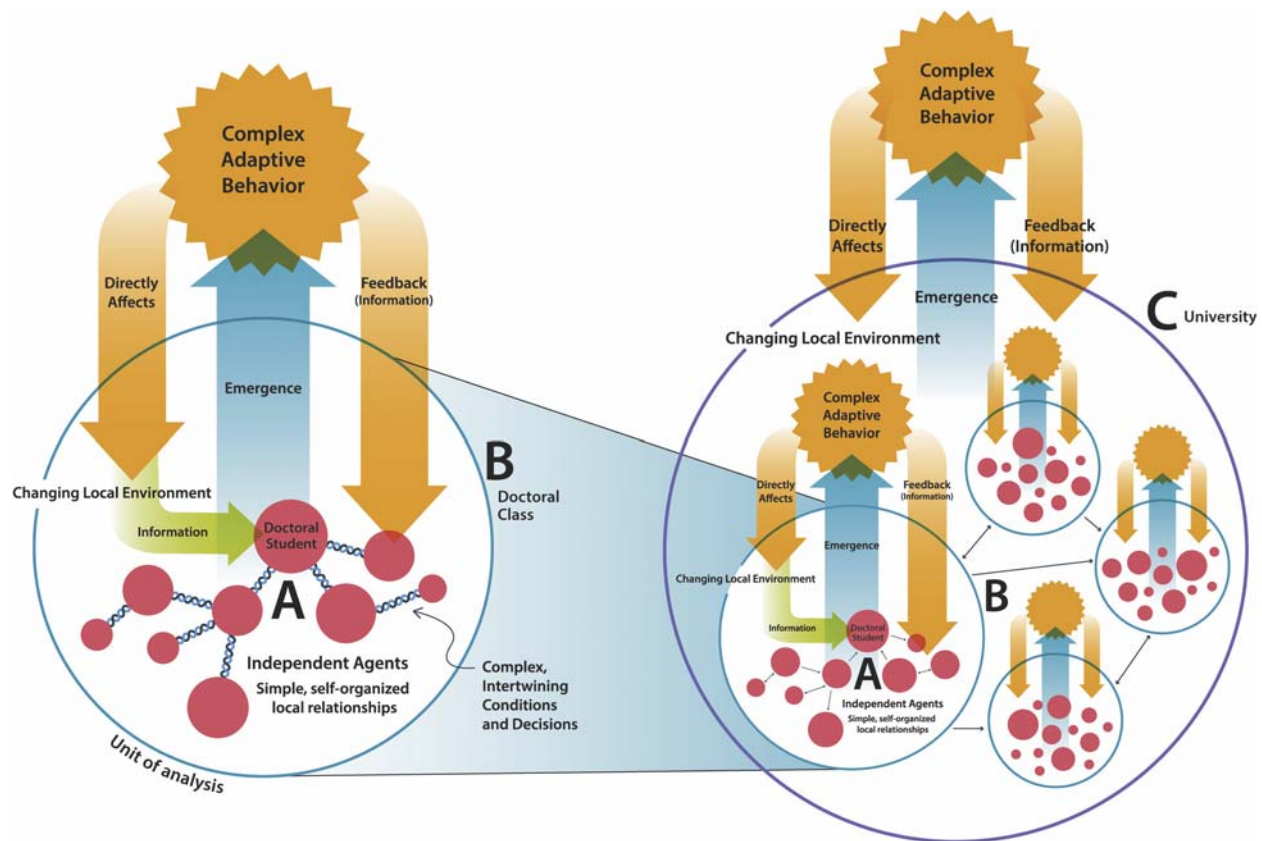


Fig. 4 A visual representation of nested structures within CAS for Poth et al. (2020).

predict and explain the emergent property of increased gains in competency. With the selection of the group of learners as the unit of analysis, Poth could focus on collecting both quantitative and qualitative data sources to generate a more complete understanding of the influences on, and by, the individual learners within the CAS of their particular course to generate understandings of the impactful experiences the group of learners attributed to their competency gains. Together these emerging understandings of the unit of analysis of the group of learners helped Poth to consider the use of a case study as appropriate for her completeness-focused mixing purpose.

As we reflected on our unit of analysis decisions, we found that recognizing how we focused our studies as situated in the broader CAS structure was helpful. Yet we have also come to realize that pre-determining the unit of analysis within a CAS can be especially challenging because of the fractal and dynamically changing nature of CASs. Common to both our approaches was the decision to represent the people involved as influencing and being influenced by their surrounding conditions. For Bullock, it was the influences of the many different stakeholders within their specific schools as well as the district, state, and national contexts in which the school was embedded. For Poth et al. (2020), it was the influences of the many contributors to the group of learners in the class, as well as the instructional, institutional, and societal contexts in which the course was embedded. Interestingly, we realized that our familiarity with our study contexts from past instructional experiences meant that neither of us experienced challenges in identifying our unit of analysis at the beginning of our studies. When studying CAS using mixed methods research in educational settings, we found transforming our practice to represent the nested structures helped identify possible units of analysis for guiding where the points of integration might take place. Thus it becomes clear, but only rather implicit, that the choice of unit of analysis for the study is intertwined with data source decisions (Fig. 1) because the perspectives that have an influence on the unit of analysis are key data sources.

Attention to interactions of the CAS informs data source and sampling decisions

The justification of data sources is essential for planning mixed methods research and has important implications for defining the perspectives and contexts captured by the study (Poth, 2018a). We use sampling strategies involved in qualitative and quantitative research as well as ones that address relationships across the research samples, sometimes referred to as the mixed sampling strategy (Onwuegbuzie and Leech, 2007). For an overview of mixed methods sampling and data collection see also the article offered by Leech and Donovan (this Section). When a researcher conceptualizes their study context as CAS, they highlight the need to consider

the interactions among the multiple and interrelated levels influencing the selected unit of analysis because the behaviors of the systems may not yet have been studied or even identified. Determining appropriate data sources and sampling plans for the interactions within a unit of analysis is challenging because individuals can be associated with multiple groups within a level or even across levels. This can lead to groupings of individuals which are not independent, with important sampling and statistical implications.

Within educational settings, while data are collected for individuals, we know that the individual is both influencing and being influenced by the larger system as well as by the individual's previous experiences. Collecting data at one time point or perspective may or may not be sufficient to accurately represent the system's behavior and address the study need for integration. Fig. 2 presents possible interactions within systems. The zoomed-out graphic on the right (labeled C) could represent the larger CAS such as a school or institution, whereas the smaller zoomed-in graphic on the left (labeled B) could represent the individual class or an individual stakeholder. As before, it is also possible that the school is part of an even larger CAS, such as a district or educational system which also requires data collection. Focusing attention on the interactions provides researchers with important information for determining appropriate data sources and sampling.

You might now be asking "How does attention to the interactions help identify appropriate data sources to inform sampling"? If we look at our two studies, we see that in Bullock's (2017) study, attention to the possible interactions involved in the CAS required multiple levels of data sources to be collected over several months and that researchers build upon emerging understandings of the unit of analysis. The initial data collected involved a survey via census sampling, with questions for the school leader (labeled A in Fig. 3) addressing a total of 193 possible characteristic variables. Quantitative analytical techniques allowed Bullock to recognize characteristics of school principals which predict students' mathematics achievement (and whether these variables were independent, or not) and determine the most important relationships between those characteristics and the dependent variable. Bullock then utilized maximal variation sampling to select appropriate focus groups to qualitatively explain the influential stakeholder relationships for the school leaders' decisions and actions involving those most important characteristics across CASs (those labeled C and higher levels beyond in Fig. 3). Integrating the two data sources generated novel descriptions of emergent properties that were otherwise obscured; these related to how school leaders' decisions and actions were associated with students' mathematical achievement (labeled B in Fig. 3).

In contrast, while Poth et al.'s (2020) study focusing attention on the possible interactions involved in the CAS also required data from multiple levels, the data sources and sampling were vastly different. Of particular importance to Poth's study was capturing both the individual (labeled A in Fig. 4) and the group of learners (labeled B in Fig. 4) as well as their collective interactions and outcomes in a way that went beyond the simple recollection. Her use of individual and group course artifacts—including pre- and post-course ratings and narratives generated throughout the class, along with instructional documentation and an end-of-course focus group—generated numerous data points contributing to the rich case description which would have been inaccessible by either qualitative or quantitative data sources alone. All those involved in the study were invited to participate; this yielded matched data from the group of 11 learners enrolled in the class.

As we reflected on our data sourcing decisions in both studies, we discovered that recognizing the focal interactions and assessing the extent of independence of agents for our study was helpful in determining appropriate sampling. Yet we have also come to realize that pre-determining the interactions within a CAS and selecting sampling strategies can be especially challenging with sequential designs because of the emergent nature of the interactions. Common to both studies' approaches were multilevel data sources, sampling suitable for our mixing purposes, and understandings that emerged over time. Bullock had tried to determine overlapping group memberships of her unit of analysis. In her reporting, she justified the independence, or lack thereof, of data sources and selected sampling dependent upon the outcome. For Poth, a convenience sample was appropriate for including all those involved in the course. When studying CAS using mixed methods research in educational settings, we found transforming our practice to attend to the dynamic CAS interactions helped identify possible data sources and samples for capturing the various contexts in which interactions occurred. Thus it becomes clear, but only rather implicit, that the choice of data sources and study samples is intertwined with integration decisions (Fig. 1) because the samples that have an influence on the integration are key to generating relevant insights addressing the mixing purpose.

Adaptations to emergent CAS conditions inform integration procedural decisions

Adaptations to integration and analytical procedures are required in mixed methods research, yet rarely are rationales provided in sufficient detail to inform future researchers' appropriate responses to emerging study conditions (Poth, 2018a). What adaptations are made and why has important implications for when and how integration will take place to realistically represent the complexity inherent in education and, ultimately, the complexity in point(s) of integration and procedural descriptions. When conducting research, we conceive of emergent conditions as those that are nonlinear and changeable in unpredictable ways as well as those that may be linear and more predictable. Points of integration are where we find the "*purposeful interdependence* between different sources, methods, or approaches" (Bazeley, 2018, p. 7, italics in original). When researchers conceptualize their studies as CAS, they intensify the need to adapt to emergent conditions in ways that allow integration to generate understandings of the whole that are greater than the sum of its parts (Fetters and Freshwater, 2015). The integration of qualitative and quantitative data provides opportunities through which we can more accurately understand the emerging properties of the phenomena.

Complexity science guides educational researchers in considering the dynamic and nonlinear aspects of CAS in addition to accounting for many possible independent and not-so-independent variables which impact the dependent variable. Byrne and

Callaghan's (2014) argument about the important distinctions between restricted complexity and general complexity provide important guidance. Whereas "restricted complexity" results from interactions at simpler levels only and thus allowing for known quantitative modeling techniques, "general complexity" results from nested interactions that allows for quantitative modeling of individual agency. What is noteworthy about their work, which remains in its infancy, is that the "general complexity" represents what Bryne and Callgahan describe "interpenetrating systems with causal powers running in all directions" (2014, p. 45). If we adopt this perspective of "general complexity," those studying CASs must carefully consider the most appropriate integration and procedural decisions and adapt their initial research design appropriately for emergent conditions.

You might now be asking, "How does adaptation to emergent conditions support integration and inform appropriate procedural designs?" As we reflected upon our two studies, Bullock noted that she had paused her study after the initial quantitative analysis to refine her qualitative data collection procedures and rethink the points of integration. Importantly she described having shifted from an originally conceived linear explanatory sequential design (See Fig. 5) to one that was implemented to incorporate opportunities for further integrations (see Fig. 6). The implemented design is considered less linear than the original design because the bidirectional arrows represent opportunities she took to revisit phases in the research to assimilate emerging understandings of the study CAS. Bullock justified these adaptations as a way to better understand the influences on, and from, the school leader—a way to account for the overlap between individual agency in decision making and stakeholder influences.

While the adaptations in Poth et al.'s (2020) study also involved rethinking the points of integration, she noted that the refinements had been focused on the case analysis procedures. She described the need to expand the scope of integration; whereas she had initially planned to integrate on only one construct of instructional experiences. However after attending to the emerging understandings, the integration was instead focused on three constructs. Poth justified these adaptations as a way to generate more nuanced descriptions of the outcomes emerging not only among learners but also emerging in their engagement with the course instructor and guest speakers. Thus the points of integration in this case became focused on the instructional, relational, and personalized experiences of the learners. Working through a complexity lens allows researchers to acknowledge the messiness inherent in studying CAS and to adapt their procedures as new information comes to light.

As we reflected on our integration decisions in both studies, we found it helpful to recognize what understandings emerged while studying CASs. Yet we have also come to realize that predetermining suitable integration procedures can be especially challenging because of the evolving nature of the surrounding conditions. Common to both study's approaches was a point at which researchers stop and reconsider initial procedural design decisions in light of emerging understandings. For Bullock, it was the stakeholder influences and their resulting impacts on the school leader's decisions that warranted further attention. For Poth, it was the instructor and guest speaker influences and their resulting impacts on class learning outcomes that warranted further analysis. Interestingly, we realized both from our own experiences and from those of our colleagues and students that procedural adaptations are much more common than study accounts portray. Research accounts must carefully consider and describe why and how adaptations are justified by the researcher. When studying CAS using mixed methods research in educational settings, we found transforming our practice to adapt to emerging CAS conditions helped generate descriptions of integration procedures that were more realistic. Thus, it becomes clear, but only rather implicit, that the description of integration procedures is intertwined with credible interpretations decisions (Fig. 1) because the procedures that have an influence on the integration are key to methodological transparency.

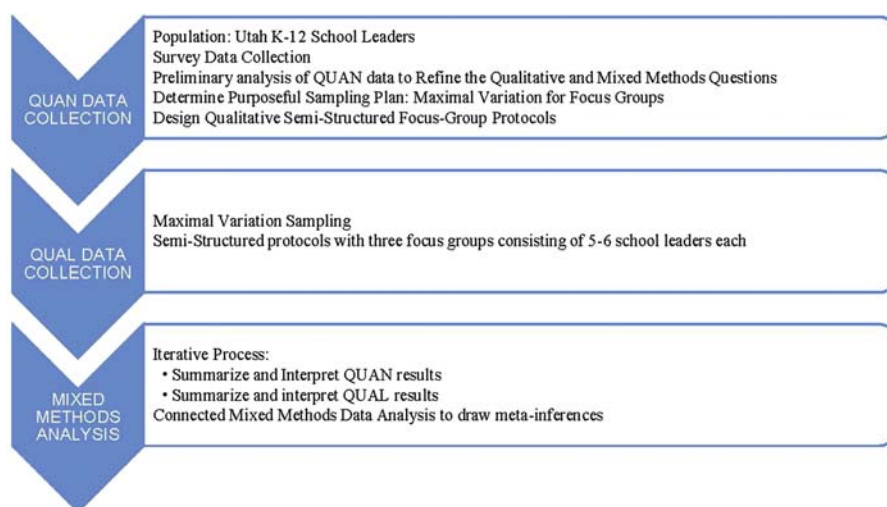


Fig. 5 Originally conceived linear explanatory sequential design for Bullock (2017).

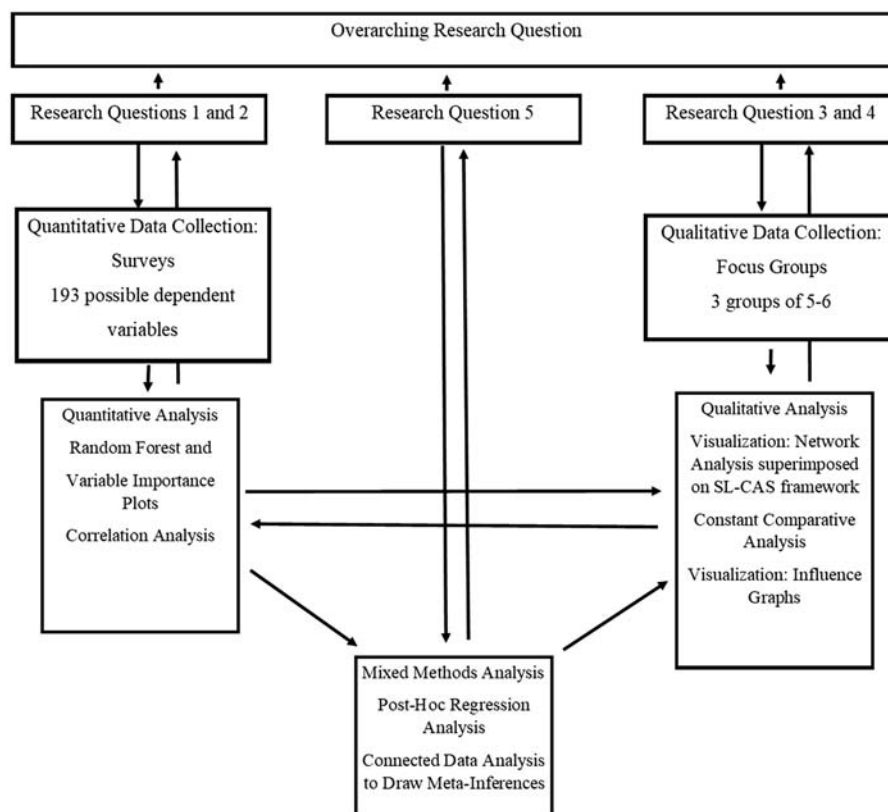


Fig. 6 Implemented less linear explanatory sequential design for Bullock (2017).

Transparency of accounts of CAS outcomes informs credible interpretation decisions

A key responsibility for any mixed methods research account is to demonstrate methodological transparency to the extent possible. For mixed methods research it is particularly important to present evidence of its key characteristic of integration. In practice, we consider evidence of methodological transparency to involve descriptions of the study's need for integration, procedures, and insights generated from the integrated findings. Methodological transparency has emerged as a key quality indicator in mixed methods research (Cain et al., 2019; Levitt et al., 2018; Onwuegbuzie and Poth, 2016). For a primer on quality in mixed methods research see also the article by Collins (this Section).

When studying educational CAS, predetermining study outcomes is challenging because of the lack of certainty and the inter-connecting and changing nature of CAS. Thus, when we study CAS in educational settings we must authentically report our methods in ways that provide evidence of methodological transparency specific to mixed methods research. Every effort to document and justify our integration and analytical decisions is as important as reporting our findings and discussing their implications, for our findings and implications depend on those decisions.

For complexity-informed mixed methods insights to be considered credible beyond the local, we have to recognize that the localization of our findings constraints how we may talk about causality or transferability. In CAS, causation can be best understood through the comparison of differences across cases which allows researchers to use these differences as a way of guiding action to achieve outcomes (Byrne and Callaghan, 2014). This has huge implications for mixed methods researchers as we provide authentic accounts of how we operate under conditions of complexity (Poth, 2018a; Poth et al., 2020). Researchers commonly report their procedures as if no adaptations took place. The result is underreporting of adaptations that researchers make. Detailing the messiness in adapting research plans and reports is appropriate and necessary when studying CAS because researchers need to recognize and respond to "the inherent complex, dynamic, and undetermined nature of mixed methods research practice" (Plano Clark and Ivankova, 2016, pp. 276–277).

You might now be asking, "how does the transparency of accounts of CAS outcomes inform the authentic reporting decisions?" In our two studies; we see that Bullock's (2017) study used visualizations in addition to narrative descriptions to convey her interpretation processes. These visualizations shed light on how she mixed quantitative and qualitative analyses and what the analyses contributed to the overall mixing purpose. Similarly, Poth et al. (2021) used a procedural figure to convey the processes involved in the initial separate analysis of qualitative and quantitative data and then how those findings were integrated using a qualitative-dominant cross-over mixed analysis approach to generate the four mixed insights for the study. A qualitative-dominant cross-over mixed analysis involved using the qualitative themes as the organizational framework on which to integrate the

quantitative data (Onwuegbuzie and Hitchcock, 2015). A procedural visual helped make explicit the interpretation processes to provide evidence of reasoned insights.

As we reflected on both our research studies, we found that being explicit about how our mixed insights had been generated from the study of CAS was helpful to contribute evidence of our reasoning. Yet we have also come to realize that pre-determining the necessary evidence of integrity for the study of CAS can be especially challenging because of the yet-to-be-known emergent study properties and outcomes. Common to both study's approaches was representation of the processes involved in the inference-generation procedures as influencing and being influenced by the data interpretations. Bullock was explicit about the iterative influences of the multilevel data samples on the emergent analytical procedures which led to emergent insights. Poth was explicit about how the insights relating the collective experiences which learners attributed to their course outcomes were different from what could have been measured from individuals' data alone. Interestingly, both realized that their comfort with research messiness from previous research experiences meant that neither assumed the interpretation processes would be straightforward. When studying CAS using mixed methods research in educational settings, we transform our practice by identifying possible interpretations and seeking purposeful evidence of integrity to inform the study's results decisions.

Implications for complexity-informed mixed methods research in education design practices

To support mixed methods researchers in managing the complexity inherent in educational settings and problems, we now advance key theoretical and practical implications of the design guidance. Together, the discussion of complexity-informed mixed methods research design practice transformations, the generation of practical insights for navigating the challenges faced by researchers studying CAS, and the lessons gleaned from a reflection of two authentic accounts realizes an important aspect of the methodological advancements described in the Future of Mixed Methods Task Force report (Mertens et al., 2016).

Our discussions and illustrations of complexity-informed mixed methods research design practice transformations advances our theoretical understandings of the realworld applications of the four design decisions summarized in Table 2. In so doing we recognize the key contributions of four designs decisions which serve as indicators of complexity-informed mixed methods research quality (e.g., Fàbregues and Molina-Azorín, 2017; Creamer, 2017; Onwuegbuzie and Poth, 2016). Our reflection upon our experiences has yielded important insights about the challenges researchers must navigate when studying CASs. In doing so, we extend Kallenmeyer's theoretical conceptualization of a complexity lens onto mixed methods research design practice. The comparison of two published studies provided both illustrative examples of vastly different studies of CASs in education and essential lessons for others to be guided by and extend in the future. Our accounts also advance the use of authentic accounts of complexity-informed mixed methods research practices for others to learn from (Poth, 2018a).

Concluding thoughts

Our impetus for this work was the lack of practical guidance we could find as we were engaging in addressing complexity in education in our own studies. Only later could we articulate the source of our challenges as the complexity inherent to our studies and how we had transformed our practices to manage. Over the last few years we have been encouraged to confirm that the work we

Table 2 Summary of four key mixed methods research design decisions and complexity-informed practice transformations.

<i>Mixed methods research design decision and unique contribution</i>	<i>Key design practice transformations</i>	<i>Complexity-informed mixed methods research design practices</i>
Initial design decisions involve identifying the unit of analysis	Represent the nested structure of the CAS to focus the study on the appropriate point(s) of integration	Distinguish among the embedded levels of the CAS Select a suitable unit of analysis for the study's need for integration
Evolving data decisions involve selecting data sources and sampling	Attend to the dynamic CAS interactions to define the appropriate perspectives and contexts captured by the study	Determine appropriate data sources for the unit of analysis in the CAS Distinguish suitable data sources for the study's interactions
Ongoing analytical decisions involve describing the integration procedures	Adapt the procedural decisions to respond to emerging CAS conditions and respond to the study's mixing purpose	Describe appropriate integration for the multilevel data collected within the CAS Distinguish suitable adaptations for the study's integration procedures
Authentic reporting decisions involve accounting for credible interpretations	Generate transparent accounts of how the interpretations were generated from the findings	Report appropriate evidence of integrity for the emergent conditions of the CAS Distinguish suitable evidence of validity for the study interpretations

engaged in was innovative and we share our lessons learned here for the benefit of readers. We encourage others to do the same. Our collaborative work on this article reflects the forward thinking necessary for managing the complex conditions surrounding many pressing societal issues and was informed by our own research experiences and the experiences of others around us and as authors, reviewers, and readers of mixed methods research literature.

Acknowledgments

The article advances new insights and extends ideas we have previously published separately (e.g., Bullock, 2017, and Poth, 2018a,b, 2019). We thank Adrienne Montgomerie and an anonymous reviewer for helpful comments on the draft manuscript. We are grateful for the opportunity to contribute to this encyclopedia and thank our colleagues from around the world who continue to influence our thinking as well as our Volume editor Dr. Nataliya Ivankova and the editors-in-chief for guiding this article through to publication.

References

- Bazeley, P., 2018. Integrating Analyses in Mixed Methods Research. Sage, London.
- Bullock, E.P., 2017. An Explanatory Sequential Mixed Methods Study of the School Leaders' Role in Students' Mathematics Achievement Through the Lens of Complexity Theory (Order No. 10288685). Doctoral thesis. Utah State University. Available at: <https://digitalcommons.usu.edu/etd/6096/>. (Accessed 14 December 2020).
- Byrne, D., Callaghan, G., 2014. Complexity Theory and the Social Sciences: The State of the Art. Routledge, New York.
- Cain, L.K., MacDonald, A.L., Coker, J.M., Velasco, J.C., West, G.D., 2019. Ethics and reflexivity in mixed methods research: an examination of current practices and a call for further discussion. *Int. J. Mult. Res. Approaches* 11, 1–12.
- Creamer, E., 2017. An Introduction to Fully Integrated Mixed Methods Research. Sage, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. Designing and Conducting Mixed Methods Research, 3rd Ed. Sage, Thousand Oaks, CA.
- Davis, B., Sumara, D., 2006. Complexity and Education: Inquiries Into Learning, Teaching, and Research. Lawrence Erlbaum Associates, Inc, Mahwah, New Jersey.
- Fàbregues, S., Molina-Azorín, J.F., 2017. Addressing quality in mixed methods research: a review and recommendations for future agenda. *Qual. Quantity* 51, 2847–2863.
- Fetters, M., Freshwater, D., 2015. The 1 + 1 = 3 integration challenge. *J. Mix. Methods Res.* 9, 115–117.
- Jacobson, M.J., Levin, J.A., Kapur, M., 2019. Education as a complex system: conceptual and methodological implications. *Educ. Res.* 48, 112–119.
- Kallemeyen, L.M., Hall, J.N., Gates, E., 2020. Exploring the relevance of complexity theory for mixed methods research. *J. Mix. Methods Res.* 14, 288–304.
- Koopmans, K., Stamovlasis, D., 2016. Complex Dynamical Systems in Education: Concepts, Methods and Applications. Springer, Switzerland.
- Levitt, H.M., Bamberger, M., Creswell, J.W., Frost, D.M., Josselson, R., Suárez-Orozco, C., 2018. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: Publications and Communications Board Task Force Report. *Am. Psychol.* 73 (1), 26–46.
- Levin, R., 1993. Complexity: Life at the End of Chaos. University of Chicago Press, Chicago, IL.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N.G., Maxwell, J.A., Molina-Azor, J.F., Niglas, K., 2016. Expanding thinking through a kaleidoscopic look into the future: implications of the Mixed Methods International Research Association's Task Force Report on the future of mixed methods research. *J. Mix. Methods Res.* 10, 221–227.
- Mertens, D.M., 2015. Mixed methods and wicked problems. *J. Mix. Methods Res.* 9, 1–6.
- Onwuegbuzie, A.J., Hitchcock, J.H., 2015. Advanced mixed analysis approaches. In: Hesse-Biber, S.N., Burke Johnson, R. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University, Oxford, pp. 275–295.
- Onwuegbuzie, A.J., Leech, N., 2007. Sampling designs in qualitative research: making the sampling process more public. *Qual. Rep.* 12, 238–254.
- Onwuegbuzie, A.J., Poth, C., 2016. Editors' afterword: toward evidence-based guidelines for reviewing mixed methods research manuscripts submitted to journals. *Int. J. Qual. Methods* 15, 1–13.
- Pennings, H.J., Mainhard, T., 2016. Analyzing teacher–student interactions with state space grids. In: *Complex Dynamical Systems in Education*. Springer, Switzerland, pp. 233–271.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks.
- Poth, C., Searle, M., Aquilina, A., Ge, J., Elder, A., 2020. Assessing competency-based evaluation course impacts: a mixed methods case study. *Eval. Progr. Plann.* 79, 1–17.
- Poth, C., Bulut, O., Aquilina, A., Otto, S.J.G., 2021. Using data mining for rapid complex case study descriptions: examples of public health briefings during the onset of the COVID-19 pandemic. *J. Mix. Methods Res.* 15, 1–26.
- Poth, C., 2018a. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. Sage, London.
- Poth, C., 2018b. The curious case of complexity: implications for mixed methods research practices. *Int. J. Mult. Res. Approaches* 10, 403–411.
- Poth, C., 2019. Realizing the integrative capacity of educational mixed methods research teams: using a complexity-sensitive strategy to boost innovation. *Int. J. Res. Method Educ.* 42, 252–266.
- van Vondel, S., Steenbeek, H., van Dijk, M., van Geert, P., 2016. "Looking at" educational interventions: surplus value of a complex dynamic systems approach to study the effectiveness of a science and technology educational intervention. In: Matthijs, S., Dimitrios, S. (Eds.), *Koopmans[SciEd1], Complex Dynamical Systems in Education*. Springer, Switzerland, pp. 203–232.
- Weaver, W., 1948. Science and complexity. *Am. Sci.* 36, 536–544.

Conducting large-scale mixed methods studies in education

Katrin Niglas, Rain Mikser, Krista Loogma, Meril Ümarik, Inge Timoštsuk, and Tiia Õun, Tallinn University, Tallinn, Estonia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	626
Challenges in conducting large-scale MM studies in education – the literature review	627
Reflections on an exemplary large-scale MM study	628
Background and relevance of the TePP project	629
The central line of MM design in the TePP project	629
Added value and challenges of MM design – some examples with discussion	631
Concluding remarks	633
References	634

Introduction

The implementation of mixed methods research practices in various academic fields has gained vast popularity in recent decades. It is argued that mixed methods research has a great potential to contribute in uncovering the challenges and offering solutions to important and complex societal problems (Mertens et al., 2016). In the field of education, the universities and funding agencies have encouraged research that addresses broad and enduring problems of education in an increasingly diverse and unequal society. Such large-scale research is best served by research teams drawn from different disciplines and combining a wide range of research designs and methods (Weis et al., 2019). Commonly this kind of research practices are called as “multimethod”, “multiple methods” or “mixed methods” research (Niglas, forthcoming). We choose to follow the call of the MMIRA task force welcoming flexible and open methodological framework and using “mixed methods” (MM) as the general term for the tradition (Mertens et al., 2016), because the concept of MM is widely used in the broad interdisciplinary field of education.

Recognizing that often MM is about combining qualitative and quantitative methods and methodologies, we purposefully adhere to the broad definition of MM suggested in the abovementioned report: “use of more than one method, methodology, approach, theoretical or paradigmatic framework” and “integration of results from those different components” (Mertens et al., 2016a, p. 4). Our stance, looking at MM more broadly than specific tool used to collect information, is congruent with the ways the group of influential scholars in education have conceptualized this methodological tradition:

We thought of “mixed methods” as encompassing broad uses of research concepts, theories, designs and methods drawn from a wide range of disciplines, intellectual traditions, and research paradigm; as dynamic and integrated uses of methods that complement one another as they unfold over time and across levels and scales of a system; as guides to more comprehensive studies; as ways to collect, analyse, and integrate data creatively; and as ways to interpret and disseminate results to a wide range of audiences.

Weis et al. (2019).

We recognize the important contribution that the authors around the globe have made in developing the scholarship and awareness of MM research designs (see for example Niglas, forthcoming for concise overview). However, the field is still in the phase of fast development and there is a need for further insights from specific disciplinary and methodological point of views. We take special focus on large-scale MM studies as this seems to be the gap in the current methodological discourse.¹ Therefore, our modest contribution in the first part of this chapter provides a concise overview of the literature on the challenges that the teams of the large-scale MM studies in the field of education are facing and need to unriddle.

In the second part of the chapter we build on the results and experiences from our recent large-scale MM research project focusing on the vital influences that the teachers’ professionalism and national educational trajectories have on the global policy transformations and related educational reforms.² Our exemplary study focusses on Estonia, but helps to shed some light to the broader context of fundamental education change that followed the regime change in post-socialist countries. By tackling how global initiatives are modified, redirected and potentially transformed across different spatial and temporal contexts and across various societal

¹The concept of a *large-scale empirical study* is most often used in connection with survey research where the number of the participants or subjects, representing the population, is high. Typically surveys with a sample size ca 1000+ could be considered large-scale. In the MM context the concept of large-scale study has more dimensions and the sample size does not always play a pivotal role. Instead, the complexity of the MM design (incl the number of combined methods and the depth of integration) can determine the need for multiple and/or multilevel data collection phases, prolonged intensive study period and a large research team with diverse competences and experiences.

²See <https://etis.ee/Portal/Projects/Display/54ef685f-4b28-497a-b409-25f8d2f8d008?lang=ENG>.

levels our research helps to understand the reasons of the widespread failure of educational change initiatives and finding ways to better support teachers' professionalism.

While providing pool of internationally relevant and thought-provoking results we, as a team of the large-scale MM research project, had to find solutions to a number of methodological and practical challenges. In an earlier paper we reflected on the issues of the team dynamics in the context of non-hierarchical shared leadership model that we implemented (Niglas et al., 2018). Our further insights provided in this chapter demonstrate, that the integration of the data and findings from different stages of the sequential MM study and especially MM publishing are the central issues that need careful attention by the large-scale MM research teams.

Challenges in conducting large-scale MM studies in education – the literature review

The mixed methods research community has provided important scholarly discussions and produced a vast pool of theoretical as well as practical knowledge to guide designing and implementing of MM studies (Niglas et al., 2018). However, limited attention has been paid to the ways in which MM can be thoughtfully and rigorously employed to tackle the broad and enduring issues in the field of education. To address this limitation the US based Mixed Methods Working Group (MMWG) sponsored by Spencer Foundation was convened in 2014 with the aim to provide guidelines for broad-based MM research projects that address significant research questions in education by capitalizing on the complementary strengths of different methods (Weis et al., 2019).

The MMWG emphasized the need to move beyond the “qualitative-quantitative binary” when designing educational studies. This means one should keep in mind that methods are not independent of the disciplinary and theoretical traditions and choose the research methods that best serve the research questions at issue (see also Bowers et al., 2013; Creswell et al., 2011; May and Burke, 2010). Working group saw fruitful combinations, integration and iteration as the key features of effective MM research. Using various methods and techniques like experiments, surveys, ethnographic observations, in-depth interviews, etc. is a well-established practice in education research. When combined effectively, the complementarity of the methods can help to deepen understanding of the studied phenomenon, but also challenge one another illuminating taken-for-granted assumptions or initiating new insights (Weis et al., 2019).

The need to integrate, coordinate, iterate, or juxtapose the data and findings from different methods within a large-scale educational research project demands a high level of team interaction and shared attitude of methodological and theoretical flexibility. Constituting and managing the MM research team is posing a challenge that the leaders of large-scale studies or programs of study need to face and successfully resolve (Bowers et al., 2013; Creswell et al., 2011; May and Burke, 2010). Researchers from the field of health sciences have provided valuable suggestions for establishing practices that help to solve emerging difficulties when managing MM teams (e.g., Bowers et al., 2013; Hemmings et al., 2013; Robins et al., 2008).

To unravel the potential conflict among differing epistemological standpoints or colliding worldviews, disciplinary- and paradigm-level borders need to be converged before a MM team can effectively operate (see Johnson, 2017 for the useful discussion). As maximizing the qualitative contribution has been experienced often as a challenge for MM teams, qualitative experts should be involved in the design as well as the execution of the MM study as early as possible (Robins et al., 2008). On the other hand, it is important to realize that researchers bring their home group (i.e. organizational, professional, and other representational groups) identities to the new group (Curry et al., 2012; Hemmings et al., 2013). Thus, the convergence and operation as one team inevitably involves working outside of one's comfort zone (May and Burke, 2010).

Successful leaders for MM teams should be able to overcome and bridge differences and create an effective working environment. This can be achieved by keeping an open dialogue going within the team and by taking into account team members' responsibilities outside the project while being concrete about their roles and duties inside the team and also promoting collaborative relationships among team members (Creswell et al., 2011; Curry et al., 2012). Bowers et al. (2013) argued that in MM research study, a non-hierarchical, shared leadership model might be needed for implementing the full potential of the team. In this model, all activities (e.g. resource allocation and sharing, project and measure design, data collection, analysis, interpretation, dissemination) require collaboration by and active engagement of every member of the research team. To meet the challenges that they faced in an exemplary THRIVE project, the Coordinating Center and various regular meetings were established in order to introduce relevant literature and theoretical-methodological approaches to the members of the team with different disciplinary and methodological backgrounds, as well as to support the commitment of the whole team.

Practical challenges related to the team dynamics were thoroughly tackled in the chapter reflecting on the experiences of the research team with a non-hierarchical shared leadership model conducting a large-scale MM program of study in education (Niglas et al., 2018). The analysis demonstrated that team members brought into the project different home group identities and experiences even though they came from the same broad discipline of education. This, in combination with a multilevel leadership model, seemed to result in the varied interpretations of the conceptual and methodological aspects of the study within different work-groups. We learned that a fine balance needs to be found between the autonomy of the work-groups and centrally facilitated construction of a shared conceptual framework.

Also, ongoing communication and collaboration among all team members (within workgroups and across workgroups) as well as the establishment of clear communication patterns to produce the functioning networks are important for project success. Communication should include learning about the competencies of all members of the project and also understanding and adopting of project leadership roles and activities. Successful leadership of large research teams requires an awareness of the consequences

of practical issues such as being paid directly from the project or not, agreeing about the allocation of project work, balancing the duties and commitments the team members might have outside the project, and so forth, because these issues significantly impact the commitment of team members (Niglas et al., 2018).

May and Burke (2010, p. 3) emphasize that issues around analyzing, interpreting, and integrating different data may pose a key problem even if the “creative, intellectually-challenging-yet-supportive team” has been formed. The MMWG pointed out the challenge of supporting the professional identity building for MM researchers, referring to the researchers who have the training and ability to work on MM teams, but assumed that up to date majority of scholars are most likely to identify themselves as aligned with particular method (Weis et al., 2019). Therefore, MM projects require educating team members about MM research on theoretical and practical levels, continual dialogue about the added value of combining approaches, and developing understanding and agreement about how to produce integration through the use of the MM design. It is important that members of the project are motivated to engage in the challenge and possibilities for lifelong learning and professional development (Niglas et al., 2018). Thus, although recognizing that it is not necessary or even possible for every member of the team to hold expertise in all methods employed in a specific research project, it is important to make clear the added value of the MM approach and set out the clear division of labor in terms of when and by whom the mixing should take place (May and Burke, 2010).

Additional challenge in large-scale MM research is related to funding proposals and publishing the research results. Editors of education research journals and funding agency program officers find it difficult to both advise researchers who wish to submit work that draws on multiple methods and identify reviewers with appropriate expertise to evaluate the submissions (Weis et al., 2019). On the other hand, there is evidence that editors of educational journals increasingly encourage presenting internationally based cross-disciplinary MM research that can inform and enrich our understandings of complex phenomena related to learning and teaching (Reio, 2013; Niglas, forthcoming).

To provide means for enhancing MM practices and publication editors of the Journal of Mixed Methods Research (JMMR) and other prominent MM authors have provided numerous guidelines on quality issues in MM research and MM manuscripts (e.g., Creswell, 2007; O’Cathain, 2010; Dahlber et al., 2010; Mertens, 2011; Fetter and Freshwater, 2015; Reio and Werner, 2017; Baim-Lance et al., 2020; Cheek, 2021). Recent concise overview of the development of quality criteria and frameworks can be found in the editorial for the virtual special issue of JMMR on “Quality in Mixed Methods Research” by Fabregues et al. (2021).

It is commonly emphasized that MM publications pose a great challenge to the research teams because the volume of diverse data and analytical complexities make it difficult to meet the publication norms in terms of article length and type accepted by most academic journals (Creswell et al., 2011; Baim-Lance et al., 2020; Reio and Werner, 2017). In the early insights, the challenge of integration of qualitative and quantitative components in writing up and the need for guidance and templates for reporting of MM studies was strongly emphasized (e.g., Bryman, 2007; Creswell and Tashakkori, 2007; O’Cathain et al., 2009). The burning questions of the kind - *Should QUAN and QUAL be reported as separate studies and complemented by the section or separate publication where they are integrated?* or *Should QUAL and QUAN always be integrated throughout the writing in the context of MM studies?* – were raised and different point-of-views provided.

The simple suggestion put forward is to organize the results of a MM study into three publications QUAN, QUAL and MM article (e.g., Creswell, 2015, p. 92). While commonly noticed that the structure of the MM paper depends on the design of given empirical MM study, the need for providing a rationale for using MM design and specific evidence of integration of qualitative and quantitative components is strongly emphasized (Fetter and Freshwater, 2015; Fabregues et al., 2021).

The above-mentioned simple model of publishing can be challenged from several viewpoints. The most radical reasoning is represented by the suggestion that true MM research “cannot be divided into two articles for publication”, because in doing so the most important aspect of the analysis will be lost (Sims, 2011). Quite in the same vein, Julianne Cheek has recently argued in her editorial to the journal of Qualitative Health Research that “severing a mixed methods study into several publications is antithetical to the purpose and principles of mixed method design”. She specifically points out the problem of publishing qualitative part of the study with a clause “this work is part of a larger study” leaving the audience to wonder what were the exact design details of this larger study and if this piece is actually a qualitative study in its’ own right or just a “supplementary component of a mixed-methods study”? As a solution, to meet the page limit requirements one could provide online supplements, give a reference to the more voluminous publication like monograph or doctoral dissertation where to obtain detailed information, etc.

However, we argue, that when conducting large-scale MM programs of study neither of the publishing models described above seem to work for many practical reasons. In the large-scale studies running up to 5 years (even if purely quantitative or qualitative), the volume of data and complexities of interpretation, but also the expectations of funders concerning the dissemination, often dictate publishing the results of the study in a number of separate papers and volumes in parallel with implementing the project. Thus, the MM team has to find ways for fragmented publication of the results of the large-scale study so that not only the integrity of its’ sub-parts will be convincingly presented but also the added value of the MM design becomes tangible.

Reflections on an exemplary large-scale MM study

Our MM research project called Teachers’ Professionalism and Professionalism in Changing Context (TePP) aimed to scrutinize how global policy transformations and teachers’ life histories have affected teachers’ professionalism in Estonia. Central concepts in the theoretical framework of the TePP project were professionalism, professionality (i.e., personal concept of professionalism), and

refraction, which elucidates how teachers' professionalism and national educational trajectories vitally influence the fate of educational reform and restructuring initiatives (see [de Groot et al., 2013](#); [Goodson and Lindblad, 2010](#)).

Background and relevance of the TePP project

Contradictory demands on contemporary educational systems and standardized top-down reforms often erode teachers' motivations. Therefore, stakeholders continuously need to negotiate the meaning of professionalism of teachers to understand the expectations for teachers as professionals and to implement best practices to support them. In conditions of the continuously diversifying student population external stakeholders, and parents in particular, have become ever more demanding and are increasingly positioned as consumers in relation to teachers. First and foremost, these increasing demands concern the academic achievements of the students, in which terms the parents' expectations to their children's outcomes have considerably risen ([Mikser et al., 2020a](#)).

The lack of knowledge and attention in that respect sheds some light to the reasons of the widespread failure of educational change initiatives ([Goodson et al., 2006](#); [Sahlberg, 2010](#); [Wallace, 2003](#)). Thus, the broader problem that shaped our research focus was that previous studies and the international policy discourse of change have insufficiently succeeded in explaining actual changes in teachers' professionalism and professionalism, and in education and social sphere more generally.

Additionally, we acknowledged that rather little attention has been paid to the context of fundamental education change that followed the regime change in post-socialist countries. A significant inadequacy in interpreting educational change has been a scarce attention to how global initiatives are modified, redirected and potentially transformed across different spatial and temporal contexts and across various societal levels (e.g. global, national, local and individual). We argue that when global policy initiatives are implemented and their effects discussed, non-dominant social and political contexts are particularly prone to being oversimplified and misinterpreted. This tendency, in turn, evokes avoidance and resistance against global reform initiatives in these contexts, particularly at sub-national levels ([Goodson and Mikser, 2022](#)).

The reception and implementation of global educational reform initiatives is a multilevel, reciprocal and complicated process, where initiatives from one level are rarely if ever uncritically accepted at another level. More often, these initiatives are ignored, redirected, resisted or, in some cases, rejected altogether. In this process, teachers deserve particular attention, because the success or the failure of large-scale educational reform initiatives ultimately depends on each of them ([Mikser and Goodson, 2018](#)).

Oversimplifications and misinterpretations occur on various dimensions. In the *temporal* dimension, changes are too often analyzed from an unreasonably short time perspective, whereby persistences and continuities, however influential, are unjustifiably ignored. For example, certain mentalities and practices of socialist-experienced teachers are routinely misattributed to the authoritarian influences of the Soviet period, whereas in fact these practices can result from much older traditions of certain educational schools of thought or of a country under observation. Distortions on the *spatial* dimension typically occur when cardinal points and other geographical concepts are expanded and misapplied to mark, amplify and perpetuate geopolitical discriminations. [Wolff \(1994\)](#) has thoroughly demonstrated that the East-West geopolitical axis in Europe is far from being a neutral construction, nor was it invented only during the Cold War period. From the Enlightenment, the construction of "Eastern Europe" served as Western Europe's first model of underdevelopment, which the later historical periods have accommodated and perpetuated. Derivatives such as "Eastern European" are used to mark people as well as their mentalities and modes of behavior, in order to legitimize the application of neoliberal policies in these socio-historical contexts ([Goodson and Mikser, 2022](#)). The same practices are also applied to other non-dominant contexts in global East and South ([Tabulawa, 2003](#)).

Finally, distortions on what we call the *representational* dimension occur when the interpretation of changes by representatives of one societal level (e.g. state level) disproportionately dominates over those of other societal levels. State policy representatives, local authorities, school leaders, individual teachers and other interest groups can interpret reform initiatives in different and even contrasting ways. [Goodson and Mikser \(2022, p. 12\)](#) explain:

The conflict typically appears between the systemic narrative that includes the storyline of major transitions as described in state policy documents, research and media outlets, and the personal and professional narratives which appear in the life histories of individual professionals.

The systemic narrative is not neutral, since its composers necessarily tend to select and interpret the documentary sources according to their own reform agenda. As [Kitaevich \(2014\)](#) demonstrates in the Georgian context, state narrative is heavily overrepresented at the expense of that of individual educational professionals. This representational conflict, in turn, can reveal the hidden aspirations of neoliberal policies, in which case the stigma of "(post-)socialist" is used by policy makers to subject teachers to the new, arguably liberal democratic requirements of standardization and accountability ([Silova and Brehm, 2013](#)).

The central line of MM design in the TePP project

Institutional grant awarded to our team by the Estonian Research Foundation enabled to carry out large-scale research project with duration of 5 years (2014–2019) led by principal investigator and 6 senior staff members. During the project 24 other researchers and more than 10 Ph.D. candidates were involved in specific research activities for shorter periods. Active dissemination of the results was expected by the funders throughout the project. In the final report ca 45 publications were listed, 5 defended Ph.D.

dissertations and more than 10 defended master theses were related to the project. Additionally, several publications have been published and doctoral dissertations defended after the official end of the project. Although the core team was formed from the members of the same faculty, the co-operation with international partners enriched the project and provided opportunities to gain new insights.

The aims and research questions of this large-scale study were complex and multifaceted demanding integration of macro-level processes with personal micro-level conceptualizations and practices of teachers. Therefore, our early decision was to integrate systematic policy analysis, large-scale quantitative survey, and in-depth qualitative interviews with teachers, academic professionals and education policy makers from various historical contexts (who, in our terms, covered different levels at the representational dimension). Sequential MM design was chosen for integrating qualitative and quantitative methods, so that methodological decisions during later stages of the study could be informed by the results from the earlier stages, but also to allow the results of the earlier stages to be explained and enriched by the analysis in the later stages. Referring to the widely used typology by Jennifer Greene (2007) the main purposes for mixing were development and complementarity (i.e. seeking elaboration, illustration, enhancement, and clarification of the findings from one analytical strand with results from the other analytical strand).

Three educational sectors – general education, early childhood education, and vocational education – were represented by the leading members of the team. Some of us had stronger backgrounds and experience in qualitative research and others in quantitative research, but everybody welcomed the MM approach and suggested to share the workload not by the quantitative and qualitative stages of the study, but to divide the team into three workgroups according to educational sectors mentioned above. This resulted in a two-level shared leadership model where the role of the principal investigator was seen most importantly as the conceptual leader and coordinator of the team (Niglas et al., 2018). All workgroups implemented the same general MMR design, but it was agreed that some details like the sampling principles and data collection instruments might vary according to the contextual needs of the particular educational sector.

The TePP study was divided into four stages (see Fig. 1). It started with document analysis and secondary analysis of existing empirical evidence with the aim of mapping the historical background and changes in theoretical, ideological, and normative

Integration for complementarity	Stage of the study and main output (=>)	Integration for development
	Stage 1: Historical and background analysis	
	1.1 policy analysis	
	1.2 newspaper analysis	
	1.3 semi-structured interviews	
	1.4 secondary empirical sources	
Integrate for interpreting survey results (S2) and work life narratives (S3)	=> <i>systemic narratives</i>	Use to inform survey development and narrative interviews on S2
	Stage 2: Survey	
Integrate for interpreting systemic narratives (S1) and work life narratives (S3)	=> <i>typology of teachers</i>	Use to inform sampling as well as conduct and analysis of interviews on S3
	Stage 3: Narrative Life history analysis	
	3.1 Narrative life story interviews	
	3.2 Semi structured interviews	
Integrate for interpreting systemic narratives of public discussion (S1) and (S2) typology of teachers	=> <i>work-life narratives</i>	Integrate the results of S1-S3 for informing scenario building on S4
	Stage 4: Merged analysis	
	4.1 Integrated analysis of the results of stages 1-3	
	4.2 Scenario building	
	=> <i>Scenarios for supporting teacher professionalism and professionalism in Estonia</i>	

Fig. 1 Sequential mixed methods design of the TePP project.

concepts of teachers' professionalism. To complement policy analysis, number of expert-interviews were carried out with experienced teachers and educational administrators (see, for example Mikser and Goodson, 2018; Loogma et al., 2019; Loogma, 2016).

On the next stage, we implemented large-scale survey with representative samples of 653 preschool teachers, 897 general education teachers and 501 vocational education teachers to test models of professionalism to identify possible typologies of teachers by understanding their perceptions of professionalism and identifying the factors they attributed to the changes of professionalism. We also intended to shed some light on the implications that different concepts of professionalism might have on teachers' work and teaching practices (see, for example Sirk et al., 2017, 2020; Mikser et al., 2018; Näkk and Timoštšuk, 2021).

The third stage of the project involved two-step narrative life history interviews with purposeful samples of 15 preschool teachers, 63 general education teachers and 24 vocational education teachers. The main aim on this stage was to provide detailed understanding about the paths that teachers have followed to develop certain views on professionalism and about the perceived changes and influences forming teachers' professionalism (see, for example Goodson and Ümarik, 2019; Ümarik and Goodson, 2020; Mikser and Goodson 2020; Näkk and Timoštšuk, 2021).

In the final stage, which has continued after the formal end of the project, we have closely interrogated the juxtapositions between teacher professionalism and teachers' professionalism. The meta-analysis integrating the main results of the MM program was theoretically framed. For example, social ecology approach was applied for interpreting the results from vocational education (Loogma, 2020), while the results indicating shift toward and tensions about learner-oriented teaching in general education were conceptualized following motivational theories (Timoštšuk, 2020). Initially, the scenarios for future development of teachers' professionalism were planned to be constructed, to explain and clarify possible policies and choices for actions. However, due to the limited empirical data with the focus on the future trends, the scenarios were not provided as the main outcome of the project.

In the light of the above it can be recognized that from the methodological perspective, our aspirations were ambitious, yet we believe that learning from the experiences of our exemplary large-scale MM study is potentially useful for further investigations on the patterns of implementation of educational reform initiatives in different contexts.

Added value and challenges of MM design – some examples with discussion

The first stage of the TePP project incorporating analysis of several data sources (see Fig. 1) produced vast pool of results which were integrated on the later stages both for instrument development and complementary interpretation of empirical outcomes, but also for enhancing the theoretical conceptualization of the *refraction*, that was central notion in our project.

Development of the survey instrument and the choice of indicators was informed and validated by the results of the documentary and secondary analysis carried out during the first stage of the project. Quite traditionally, the secondary analysis of existing quantitative and qualitative studies helped to reduce the amount of new data we needed to collect by the survey instrument. On the other hand *systemic narratives*, constructed on the basis of content analysis of policy documents (incl. relevant newspaper articles) and data from expert interviews (Loogma, 2016; Loogma et al., 2019; Tinn, 2020), helped to compose the periodic list of the reforms and related (policy) changes, which were used as a backbone in the survey instrument.

The general education workgroup defined *professional learning* as an important additional focus, conceptualizing it as a constant journey of becoming a better teacher. Using social learning theory (Wenger, 1998, 2010) as a methodological framework, the semi-structured interviews with teachers' about their learning experiences focusing on their professional practice, meaningful experiences, teachers' identity, and belonging were carried out. The contradiction between teachers' liberal values (such as appreciation of the learner's individuality and humanity) and educational performance and measurement orientation was found (Timoštšuk et al., 2018). This triggered the workgroup to add an extra section into the survey instrument exploring the relationship between teacher work motivation and their teaching practices.

The importance of the input from the first stage of the study into development of the survey instrument as well as providing historical and political context for interpreting the survey results has been emphasized by the senior staff members, however, we reflect that this MM aspect was not very clearly displayed in publications where the survey results were presented (see Mikser and Goodson, 2018; Sirk et al., 2017, 2019, 2020; Timoštšuk et al., 2018). Recognizing the varying challenges (like traditional forms of publishing in quantitative-biased journals or time-pressure in preparing the publications within the large-scale MM project) hindering the just presentation of the input from the prior qualitative phase of the study, we encourage finding ways for clearly demonstrating the added value by incorporating the MM aspect even if it has had only facilitating role at given point.

The survey results, enriched by the input from the first stage of the research project, enabled reliable and generalizable comparisons of the factors that teachers attributed to their professionalism in different education sectors in Estonia. For example, the collaboration was indicated as the key factor, determining the professionalism of vocational education teachers (Sirk et al., 2017). At the macro/institutional level, the problem of low status of VET has historical roots, strengthen by the signals from labor market (Loogma et al., 2019). This macro factor has consequences of the characteristics of the VET students and therefore, to VET teachers professionalism (Sirk et al., 2016). The relation between policy changes and VET teachers' professionalism are not unambiguous. Results concerning the VET teachers' professionalism demonstrated the differentiation of VET teachers' community on the basis of their scope of collaboration and as well, identify the collaboration as determining factor of professionalism (Sirk et al., 2017). The differentiation of teachers manifests also in teachers' attitudes toward reform changes.

The results of historical and policy analysis and conceptual apparatus built during the first stage of the project gave also important input for analyzing spatial and temporal refraction in the survey results of preschool and general school teachers. The results from preschool teachers' survey indicated an apparent inconsistency in teachers' views on changes in pre-service preparation

curricula concerning then balance between theoretical and practical parts of the study. This echoed our earlier results, and enforced our understanding of neoliberal policies as keeping different stakeholders (e.g. practitioners and policy makers) apart from each other, which results in “policy evasion” by practitioners (Mikser and Goodson, 2018; Mikser et al., 2019).

As the main output of the survey study, the typology of teachers, based on differences in perceptions on professionalism, helped to determine the implications that different concepts of professionalism have on teachers’ work and teaching practices, as well as to explore the on-going relationship between concepts of professionalism and professionalism. For instance, different clusters of Estonian preschool teachers had considerably different views and expectations of educational reforms as reflected by the survey results. However, all clusters alike tended to place low value on further academization of preschool teaching profession, both in terms of the university-based training and participation in the research activity (Mikser et al., 2019).

On the other hand, the general education workgroup found that teachers’ self-reported perceptions on autonomy support and structured teaching did not vary significantly. Overall, teachers prioritized autonomy support and intrinsic motivation in learning (Timošćuk and Näkk, 2020). This seemed to be to some extent in contradiction with adjacent results and therefore additional data collection through classroom observations and students’ self-reports was introduced for validation purposes. The results revealed that teachers’ self-reports about their autonomy-supportive and structured teaching practices followed their actual teaching. Nevertheless, the context of teaching (room settings, students background and experiences) significantly influenced students’ learning engagement (Näkk and Timošćuk, 2019). This example illustrates the added value that allowing openness in conducting MM projects may bring about. The core design of our MM study was enhanced by emergent design elements (Schoonenboom and Johnsson, 2017) bringing in another purpose for mixing quantitative and qualitative approaches – namely triangulation (Greene, 2007).

For the third stage of the project, it was planned that teachers from different educational sectors and age cohorts would be selected for narrative *life history interviews* based on the typology of teachers constructed as a result of the survey study. This plan was implemented by the vocational education workgroup (Ümarik and Goodson, 2020; Goodson and Ümarik, 2019; Loogma, 2018), while for the other workgroups the plan was altered due to the practical reasons like the need to start collecting the interviews earlier than the analysis of the quantitative data was finished, being not able to identify and connect with the selected participants as most of them preferred to answer the survey anonymously, etc. However, the careful purposeful selection of the participants for the life history interviews based on the age cohorts and other contextual factors like subject taught, educational background, ensured that the opinions of different generations were represented and longer-term cultural trajectories present within each context. This enabled to draw indirect connections and comparisons with emerged clusters of teachers allowing to bring in some integrated interpretations (e.g., Mikser et al., 2020; Näkk and Timošćuk, 2019).

In the life history interview method, it is essential that in the first interview the open unstructured style is practiced, which allows the person to tell the life story that she wants to tell. All life history work, by definition, has to be the life story as experienced and represented by the storyteller. The subsequent interviews with the same respondents are more reciprocal and structured, allowing the interviewer to cross-question the major points of interest, the contradictory items and the unfinished utterances (Mikser and Goodson, 2018). In our TePP project life history interview questions for the second round were developed on the basis of all previous stages — documentary analysis, survey results, and preliminary results of the first round of interviews that asked for interviewees’ personal interpretations of their life-course and work.

While the survey results were integrated on the third stage of the study serving the development purposes, the results from the life history interviews were used to help to understand more deeply the refractions from the systemic narratives and other tendencies that the survey illuminated. The survey with preschool teachers revealed controversies in the respondents’ views on changes in teachers’ professionalism. On the one hand, the global tendency of raising preschool teachers’ formal qualification requirements was generally appreciated as a symbolic credential of their status as teachers (Mikser et al., 2020). This appreciation, however, was intertwined with preschool teachers’ widespread alienation from “theoretical” studies, possibly because this is traditionally not associated with their professional identity and, secondly, because theory is perceived as being presented in political and ideological terms and detached from teachers’ daily practice (Mikser et al., 2018). Moreover, although remarkably different attitudes toward professionalism existed *within* preschool teachers as a professional group, nearly all the respondents alike tended to be averse toward academic career (including Ph.D. education) and research activity. The global trend of raising the requirements for preschool teachers’ professional preparation is thus refracted in the Estonian preschool teachers’ mindsets that on the one hand rely upon the Nordic, homelike and non-academic tradition of preschool and on the other hand on the Soviet model, where teachers’ work was tightly regulated and requirements for professional preparation were therefore low. Our life history interviews with preschool teachers confirmed that although teachers generally perceived their work as having become more difficult over the course of their professional career, the rising qualification requirements and the widespread curriculum autonomy officially assigned to teachers has a valuable and alleviating psychological impact on them (Mikser et al., 2020).

The life history interviews with teachers from schools of secondary education revealed historical and cultural refractions from the systemic narrative in a different manner. Whereas preschool teachers’ concerns were largely related to achieving the political and public status as teachers (and not caretakers), the school teachers’ narratives expressed hesitation and incredulity toward the policy rhetoric which on the one hand grants them with professional freedom and autonomy, but on the other hand subjects them increasingly to external accountability and output control (Mikser and Goodson, 2018, 2020; Tinn and Ümarik, 2021).

The analysis of teachers’ life history interviews on the third stage of the project yielded a *re-periodization* of the systemic narrative, which we consider significant added value from MM approach. Instead of the overwhelming progress rhetoric that is apparent in the systemic narrative of the transition period (i.e. “things always become better”), many experienced teachers perceived the 1990s as “a window of opportunity” between the oppressive circumstances of the Soviet system on the one hand and the current neoliberal

education policies on the other hand (Mikser and Goodson, 2020). Furthermore, although the experiences from the Soviet period are often described as espousing outdated mentalities and practices and lacking individual initiative, creativity, responsibility and tolerance, the experienced teachers' narratives demonstrated insightful and nuanced understandings of the difficulties in balancing between the oppressive external circumstances and their own values, beliefs and practices. Contrary to the narrative, widely spread in parallel with recent reforms, professionalism in the Soviet circumstances required independence, creativity, responsibility and a highly critical mindset (Mikser and Goodson, 2020).

Thus, the juxtaposition of the (state-enhanced) *systemic narrative* (developed on the first stage of the project), survey results and teachers' *work life narratives* provided basis for truly integrated insights into the impact of neoliberal educational reforms on teachers' life and work in Estonia (Tinn, 2020). The interview results contributed to explaining the results of the quantitative survey study and added new dimensions to the historical grid by showing how teachers' personal life paths had shaped their perceptions of educational changes and societal demands on their professionalism. Most importantly, these results indicated that some negative consequences of the reforms as perceived by teachers, were the effect of refraction rather than directly caused by the initial aims of the reforms (Mikser and Goodson, 2020). Drawing parallels with earlier studies we hope that our results provide valuable contribution for explaining refractions in educational reforms also in the other former socialist European countries.

As mentioned earlier, to explain and clarify possible policies and choices for action, the scenarios for future development of teachers' professionalism were planned to be constructed as the final stage of the TePP project. However, we realized that the rich empirical data we had collected was mostly reflecting on the past events and experiences. The latter in interplay with the practicalities of the large-scale MM study and complexities of the integrated analysis, putting considerable time-pressure on the team, led to the decision to renounce the development of the scenarios. However, the vocational education workgroup tackled the challenge of constructing some provisional scenarios (see Loogma et al., 2020) building on the selected results from the first stage of the project giving some input for identifying trends of the development of teachers' professionalism in the future.

The intrinsic interest toward further conceptual development of the notion of refraction and related methodology has motivated the team to continue their joint research and publication activities after the formal end of the TePP project (e.g., Mikser et al., 2020). For example, Mikser et al. (2020) have further elaborated and combined the data from survey and life history interviews from Estonian preschool teachers. In this chapter they analyze and discuss *symbolic capital* as applied to the professionalization of Estonian preschool teachers, building on conceptual elaboration of teacher professionalism as *symbolic capital* based on the previous work of Bourdieu, Schinkel and Noordegraaf (e.g., Schinkel and Noordegraaf, 2011).

Concluding remarks

Being inspired by earlier findings of Goodson and his colleagues (Goodson, 2010; Goodson and Rudd, 2016), pointing out that different historical periods and country- or region-specific trajectories generated significant refractions from global, homogenizing neoliberal educational reform initiatives, the large-scale MM research project TePP was initiated in Tallinn University with the focus on the educational reforms the Estonian teachers had experienced over several decades including transition from socialist period. The results from our project detected refractions from the conventional perceptions of the mentalities of socialist-experienced teachers. This finding supports the

argument that the global initiatives are not straightforwardly adopted in various contexts. Instead, they are "modified, redirected and potentially transformed, providing spatial, temporal and representational variations" Goodson and Mikser (2022, p. 2). As, on the other hand, the global initiatives can not and are not completely ignored or rejected the findings from the TePP study are relevant in the other educational contexts, and especially in the context of former socialist European countries.

We consider that MM research design, that we purposefully implemented played an important role on multiple levels. It helped us to get deeper and multifaceted understanding of teachers' perceptions on their professionalism and its' interplay with educational reforms initiated in the context of global policy transformations. We have noticed that in studies on curriculum policy official documents and systemic guidelines are often considered as the natural starting point against which a certain curricular aspect is to be measured. The latter approach has serious limitations in terms of understanding the reality of schooling or illuminating mechanisms by which existing patterns of power and governance are transformed and maintained:

Politically conservative and quietist, such research leaves both the managerial class and teachers playing on their own grounds, so leaving politicians and managers space to develop and to impose narratives on a population which is experiencing something very different from what the narrative is telling them.

Mikser and Goodson (2020).

Our research strategy allowed to juxtapose top-down systemic narratives covering different educational levels with teachers' work-life narratives and their attitudes toward educational changes and is therefore "emancipating in the sense of demonstrating how myths are made by governments right through the major changes in political systems and are then peddled as systemic narratives" (Mikser and Goodson, 2020). Being able to reveal such aspects attributes certain political importance to the MM research design implemented in our project.

The meta-analysis presented in the first part of the chapter and our experiences from an exemplary large-scale MM study demonstrate that while the approach is potentially very fruitful the research team has to meet number of challenges to fully exploit and raise all the prospects of this promising methodological approach. Therefore, when planning a large-scale MM study, it is important to pay duly attention to the challenges highlighted in this concise chapter and to take special care that time for screening possible other obstacles would be allocated early in the project so that the teamwork and methodological expertise of the team could be adequately facilitated.

It is not surprising that in parallel with challenges in managing the team dynamics, the integration of the findings from different stages of the project needs special attention. However, our experiences indicate that the publication appears to be even more serious stumbling block. Our self-critical glance to the publication practices within TePP project indicate that the divergent presentation of the sub-parts of the study and findings does not effectively convey the general methodological framework and added value of the implemented MM strategy of the large-scale project even if the integrated results are interwoven into the discussion.

Anecdotal evidence can be pointed out where the modest attempts from the team to propose an MM article presenting the results of some parts of the large-scale study was rejected by the editors and substituted by the authors with the paper presenting only a certain quantitative or qualitative stage of the larger MM project. At the same time, we recognize that the number of editors calling for MM papers and chapters is increasing and ultimately the worth of MM publications depends on the quality of MM inquiry that is presented and the proficiency of the MM team to link the publications so that the role of any separately published part will be adequately understood and interpreted within the larger MM framework. To conclude we argue that specific issues concerning publishing in the framework of large-scale MM programs of study are not thoroughly discussed and need further examination.

References

- Baim-Lance, A., Onwuegbuzie, A., Wisdom, J., 2020. Project management principles for optimizing publication productivity of mixed methods studies. *Qual. Rep.* 25 (3), 646–661.
- Bowers, B., Cohen, L.W., Elliot, A.E., Grabowski, D.C., Fishman, N.W., Sharkey, S.S., Zimmerman, S., Horn, S.D., Kemper, P., 2013. Creating and supporting a mixed methods health services research team. *Health Serv. Res.* 48 (6 Pt 2), 2157–2180.
- Bryman, A., 2007. Barriers to integrating quantitative and qualitative research. *J. Mix. Methods Res.* 1 (1).
- Cheek, J., 2021. Maintaining the integrity of qualitatively driven mixed methods: avoiding the “this work is part of a larger study” syndrome. *Qual. Health Res.* 31 (6), 1015–1018.
- Creswell, J.W., 2015. *A Concise Introduction to Mixed Methods Research*. Sage, Thousand Oaks: CA.
- Creswell, J.W., Tashakkori, A., 2007. Editorial: developing publishable mixed methods manuscripts. *J. Mix. Methods Res.* 1 (2), 107–111.
- Creswell, J.W., Klassen, A.C., Clark, V.L.P., Smith, K.C., 2011. *Best Practices for Mixed Methods Research in the Health Sciences*. Office of Behavioral and Social Sciences Research, NIH.
- Curry, L.A., O’Cathain, A., Clark, V.L.P., Aroni, R., Fethers, M., Berg, D., 2012. The role of group dynamics in mixed methods health sciences research teams. *J. Mix. Methods Res.* 6 (1), 5–20.
- Dahlberg, B., Wittink, M., Gallo, J., 2010. Funding and publishing integrated studies: writing effective mixed methods manuscripts and grant proposals. In: *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications, Inc, pp. 775–802.
- de Groot, I., Goodson, I., Veugelers, W., 2013. Bridging the gap between young people’s civic selves and theoretical positions on democratic citizenship education. *Curr. Teach.* 27 (1), 21–41.
- Fàbregues, S., Molina-Azorin, J.F., Fethers, M.D., 2021. Virtual special issue on “quality in mixed methods research”. *J. Mix. Methods Res.* 15 (2), 146–151.
- Fethers, M.D., Freshwater, D., 2015. Publishing a methodological mixed methods research article. *J. Mix. Methods Res.* 9 (3), 203–213.
- Goodson, I., 2010. Times of educational change: towards an understanding of patterns of historical and cultural refraction. *J. Educ. Pol.* 25 (6), 767–775.
- Goodson, I.F., Lindblad, S., 2010. *Professional Knowledge and Educational Restructuring in Europe*. Sense Publishers, Rotterdam Boston and Taipei.
- Goodson, I., Mikser, R., 2022. Historical and cultural refractions in recent education transitions: the example of former socialist European countries. *Br. J. Educ. Stud.* 1–18.
- Goodson, I.F., Rudd, T., 2016. Restructuring, reform and refraction: complexities of response to imposed social change. *Educ. Pract. Theor.* 38 (2), 5–21.
- Goodson, I., Ümarik, M., 2019. Changing policy contexts and teachers’ work life narratives: the case of Estonian vocational teachers. *Teach. Teach.* 25 (5), 589–602.
- Goodson, I., Moore, S., Hargreaves, A., 2006. Teacher Nostalgia and the sustainability of reform: the generation and degeneration of teachers’ missions, memory, and meaning. *Educ. Adm. Q.* 42 (1), 42–61.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco: CA.
- Hemmings, A., Beckett, G., Kennerly, S., Yap, T., 2013. Building a community of research practice: intragroup team social dynamics in interdisciplinary mixed methods. *J. Mix. Methods Res.* 7 (3), 261–273.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11 (2), 156–173.
- Kitavich, E., 2014. History that splinters: education reforms and memory politics in the Republic of Georgia. *Southeast Eur. Black Sea. Stud.* 14 (2), 319–338.
- Loogma, K., 2016. Europeanization in VET policy as a process of reshaping the educational space. *Int. J. Res. Vocat. Educ. Train.* 3 (1), 16–28.
- Loogma, K., 2018. Critical incidences and events in vocational Teachers’ life careers. In: Lazaro Moreno Herrera, M.T. (Ed.), *Vocational Education & Training - the World of Work and Teacher Education*. Premiss, Stockholm, pp. 402–420.
- Loogma, K., 2020. Kutsehariduse sotsiaalne ökoloogia. In: Toim, Haridusmõte, M.H. (Ed.). TLÜ Kirjastus, Tallinn, pp. 236–261.
- Loogma, K., Erss, M., Ümarik, M., Aasa, M., 2020. Õpetaja professionaalsuse võimalikud tulevikustsenaariumid aastaks 2035. Eesti Haridusteaduste Ajakiri = Estonian J. Educ. 8 (1).
- Loogma, K., Ümarik, M., Sirk, M., Liivik, R., 2019. How history matters: the emergence and persistence of structural conflict between academic and vocational education: the case of Post-Soviet Estonia. *J. Educ. Change* 20, 105–135.
- May, V., Burke, H., 2010. Practical Considerations for Leading and Working on a Mixed Methods Project. Toolkit #11. <http://www.socialsciences.manchester.ac.uk/realities/resources/toolkits/mixed-method-working/>.
- Mertens, D.M., 2011. Publishing mixed methods research. *J. Mix. Methods Res.* 5 (1), 3–6.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N., Maxwell, J., Molina-Azorin, J.F., Niglas, K., 2016. Expanding thinking through a kaleidoscopic look into the future: implications of the mixed methods international research association’s task force report on the future of mixed methods. *J. Mix. Methods Res.* 10 (3), 221–227.
- Mertens, D.M., Bazeley, P., Bowleg, L., Fielding, N., Maxwell, J., Molina-Azorin, J.F., Niglas, K., 2016a. The Future of Mixed Methods: A Five Year Projection to 2020. MMIRA.
- Mikser, R., Goodson, I., 2018. The concept of refraction and the narrative approach to exploring multi-level social reform initiatives: conceptual and methodological issues. In: Vihaelemm, A.M.P. (Ed.), *The Routledge International Handbook of European Social Transformation*. Routledge, Taylor & Francis Ltd, Abingdon; New York, pp. 84–97.
- Mikser, R., Goodson, I., 2020. Narratives of education and curriculum transition in the formerly socialist European countries: the example of Estonia. In: McCulloch, G.G.-G. (Ed.), *Transnational Perspectives on Curriculum History*. Routledge, London, pp. 41–62.

- Mikser, R., Tuul, M., Veisson, M., Goodson, I., 2018. The place of theory and practice in preschool teachers' pre-service education: Estonian preschool teachers' evaluations in conditions of changing professionalism. In: Chris Pascal, T.B. (Ed.), *Early Childhood Education and Change in Diverse Cultural Contexts*. Routledge.
- Mikser, R., Niglas, K., Tuul, M., Veisson, M., Goodson, I., 2019. Academia undermining professionalism? The Estonian preschool teachers' views and expectations of teacher professionalism. In: Phillips, S., Garvis, S. (Eds.), *Teachers' and Families' Perspectives in Early Childhood Education and Care: Early Childhood Education and Care in the 21st Century*, Vol. II. Taylor & Francis Ltd. Routledge, Abingdon, Oxfordshire, pp. 58–71.
- Mikser, R., Veisson, M., Tuul, M., Õun, T., Kööp, K., 2020. Lasteaiaõpetajate hinnangud ja selgitused oma töö raskuse muutumisele: professionaliseerumine kui sümboolne kapital. *Eesti Haridusteaduste Ajakiri = Estonian J. Educ.* 8 (1), 128–155.
- Mikser, R., Õun, T., Tuul, M., Kukk, A., Hussar, D., 2020a. Novice preschool teachers' professional skills as assessed by preschool principals and the novice teachers themselves. *J. Early Child. Teach. Educ.* 42 (4), 437–454.
- Näkk, A.-M., Timoštšuk, I., 2019. The dynamics of learning engagement and its relationship with teachers' classroom practices in primary school. *Education* 47 (1), 89–100, 3–13.
- Näkk, A.-M., Timoštšuk, I., 2021. In pursuit of primary teachers' work motivation amid increased external neoliberal pressure in education. *Teach. Dev.* 25 (5), 585–603.
- Niglas, K., Ümarik, M., Tinn, M., Goodson, I., Mikser, R., Veisson, M., Loogma, K., Ugaste, A., 2018. Planning and conducting a large-scale mixed methods research study: research team experiences and recommendations. *Int. J. Mult. Res. Approaches* 10 (1), 283–295.
- Niglas, K. (forthcoming). Mixed methods designs: international influences on historical developments and future. In C. Poth *The SAGE Handbook of Mixed Methods Research Designs*.
- O'Cathain, A., 2010. Assessing the quality of mixed methods research: toward a comprehensive framework. In: *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications, Inc, pp. 531–556.
- O'Cathain, A., Nicholl, J., Murphy, E., 2009. Structural issues affecting mixed methods studies in health research: a qualitative study. *BMC Med. Res. Methodol.* 9, Art No.82.
- Reio, T.G., 2013. Exploring the links between adult education and human resource development: learning, risk-taking, and democratic discourse. *New Horiz Adult Educ. & Hum. Res. Dev.* 25 (4), 4–11.
- Reio, T., Werner, J., 2017. Publishing mixed methods research: thoughts and recommendations concerning rigor. *Hum. Resour. Dev. Q.* 28 (4), 439–449.
- Robins, C.S., Ware, N.C., dosReis, S., Willging, C.E., Chung, J.Y., Lewis-Fernández, R., 2008. Dialogues on mixed methods and mental health services research: anticipating challenges, building solutions. *Psychiatr. Serv.* 59, 727–731.
- Sahlberg, P., 2010. Rethinking accountability in a knowledge society. *J. Educ. Change* 11, 45–61.
- Schinkel, W., Noordegraaf, M., 2011. Professionalism as symbolic capital: materials for a Bourdieusian theory of professionalism. *Comp. Sociol.* 10, 67–96.
- Schoonenboom, J., Johnson, R.B., 2017. How to construct a mixed methods research design. *Köln Z Soziol* 69, 107–131.
- Silova, I., Brehm, W.C., 2013. The shifting boundaries of teacher professionalism: education privatization(s) in the post-socialist education space. In: Seddon, J.O.T. (Ed.), *Educators, Professionalism and Politics: Global Transitions, National Spaces, and Professional Projects*. Routledge, New York, pp. 55–74.
- Sims, W., 2011. Forum. *J. Res. Music Educ.* 59 (3), 227–228.
- Sirk, M., Liivik, R., Loogma, K., 2016. Changes in the professionalism of vocational teachers as viewed through the experiences of long-serving vocational teachers in Estonia. *Empir. Res. Vocat. Educ. Train.* 8 (13), 1–26.
- Sirk, M., Ümarik, M., Loogma, K., Niglas, K., 2017. Koostöö kutseõpetaja professionaalsust määra tegurina. *Eesti Haridusteaduste Ajakiri = Estonian J. Educ.* 5 (2), 80–105.
- Sirk, M., Liivik, R., Ümarik, M., Loogma, K., 2019. Kutseõpetajate professionaalse mälu roll õpilastega seotud muutuste kogemisel. *Eesti Haridusteaduste Ajakiri = Estonian J. Educ.* 7 (1), 101–127.
- Sirk, M., Ümarik, M., Loogma, K., 2020. The necessity of vocational educational changes as perceived by Estonian vocational teachers. *J. Vocat. Educ. Train.* 20 (5), 1–21.
- Tabulawa, R., 2003. International Aid Agencies, Learner-Centred Pedagogy and Political Democratisation: A Critique.
- Timoštšuk, I., 2020. Õppijakeskne lähenemine õppimisele. In: (Toim), M.H. (Ed.), *Haridusmõte*. TLÜ Kirjastus, Tallinn, pp. 318–330.
- Timoštšuk, I., Näkk, A.-M., 2020. Balancing teaching style in the engagement of primary students in science lessons: the examples of four teachers. *Education* 3–13, 960–971.
- Timoštšuk, I., Ugaste, A., Mets-Alunurm, K., 2018. Õpetajate õppimiskogemused neoliberaalsete haridusmuutuste taustal. *Eesti Haridusteaduste Ajakiri = Estonian J. Educ.* 6 (1), 77–101.
- Tinn, M., 2020. Murrangute uurimine – neoliberaalsete haridusreformide mõju õpetajate tööelule Eestis. *Eesti Haridusteaduste Ajakiri* 8 (1), 156–182.
- Tinn, M., Ümarik, M., 2021. Looking through Teachers' eyes - investigating teacher agency. *Br. J. Educ. Stud.* 1–17.
- Ümarik, M., Goodson, I.F., 2020. Nostalgia in the narratives of vocational teachers as a way of understanding responses to change. *Crit. Stud. Educ.* 61 (4), 448–463.
- Wallace, M., 2003. Managing the unmanageable?: coping with complex educational change. *Educ. Manag. Adm.* 31 (1), 9–29.
- Weis, L., Eisenhart, M., Duncan, G.J., Albro, E., Bueschel, A.C., Cobb, P., Eccles, J., Mendenhall, R., Moss, P., Penuel, W., Ream, R.K., Rumbaut, R.G., Sloane, F., Weisner, T.S., Wilson, J., 2019. Mixed methods for studies that address broad and enduring issues in education research. *Teach. Coll. Rec.* 121 (10), 1–16.
- Wenger, E., 1998. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.
- Wenger, E., 2010. Communities of practice and social learning systems: the career of a concept. In: Chris, B. (Ed.), *Social Learning Systems and Communities of Practice*. Springer, London, pp. 179–198.
- Wolff, L., 1994. *Inventing Eastern Europe. The Map of Civilization on the Mind of the Enlightenment*.

Advancing education research through mixed methods with existing data

Daphne C. Watkins^a and Natasha C. Johnson^{b, a} University of Michigan, School of Social Work, Ann Arbor, MI, United States; and
^b Columbia University School of Social Work, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	637
What are existing data in mixed methods?	637
Advantages and disadvantages of using existing data in mixed methods	638
Existing quantitative data in mixed methods	639
Gauge the quantitative “breadth” of a topic, then proceed qualitatively	639
Existing qualitative data in mixed methods	639
Gauge the qualitative “depth” of a topic, then proceed quantitatively	640
The knowledge-level continuum in mixed methods	640
Applying the knowledge-level continuum to mixed methods with existing data	641
Three questions when choosing existing data in mixed methods research	642
Question 1: How will you define the terms in your research question?	642
Question 2: How do you plan to find the existing data?	642
Question 3: Do the existing data help answer the research question?	643
Conclusion	643
References	643

Glossary

Conceptual framework A collection of concepts and empirical findings from the literature, used to show relationships among ideas and how they relate to research.

Descriptive research The second level of the knowledge-level continuum that uncovers descriptive information on the topic of interest using qualitative and/or quantitative methods.

Existing data Data already collected by you or someone else that can be used in a new research project. Also referred to as “secondary data.”

Existing qualitative data Text, visual, or recorded (e.g., audio, video) data collected by you or someone else that can be used to address a new research question.

Existing quantitative data Numerical data collected by you or someone else that can be used to address a new research question.

Explanatory research The highest level of the knowledge-level continuum and is conducted using quantitative methods of inquiry because previous data exists on the topic of interest.

Exploratory research The first level of the knowledge-level continuum; includes research that begins to fill the substantial gap in the existing literature related to your topic of interest.

Knowledge-level continuum A framework for describing the progression in the levels of scientific inquiry. If there is little to no pre-existing knowledge on your topic of interest, an exploratory level of investigation should be used. If some pre-existing knowledge exists, a descriptive level should be used. If there is a robust amount of pre-existing knowledge, an explanatory level should be used.

Mixed methods The rigorous and epistemological application and integration of qualitative and quantitative research approaches to draw interpretations based on the combined strengths of both approaches to influence research, practice, and policy.

Mixed methods with existing data Identifying, evaluating, and incorporating one or more existing (i.e., secondary) qualitative or quantitative data sources in a project.

Primary data Data collected in real-time to achieve the goals outlined in a research project, also called “original” or “raw” data.

Secondary data Data already collected by you or someone else that can be used in a new research project. Also referred to as “existing data.”

Theory An explanation of the “how” or “why” of phenomena, typically referencing existing facts. A theory is usually based on an existing framework and is frequently tested.

Introduction

Funders and publishers are instrumental in establishing the guidelines and processes we use for sharing data (Chawinga and Zinn, 2020). Education researchers—particularly those who use existing data (or “secondary” data) sources in mixed methods—can benefit from the recent increase in accessible data, as the advantages of using existing data in education research are far and wide (Logan, 2020; Smith, 2011). One of the most fundamental benefits of existing data is its ability to extend the knowledge about a research topic above and beyond the current state of the science. For example, a growing body of literature and data sources considers the impact of the global COVID-19 pandemic on students’ learning and educational experiences (Dorn et al., 2020; Grooms and Childs, 2021). Existing data sources that include COVID-19-related items or information about student experiences during the global pandemic can be used to extend the field’s understanding of educational disparities and shifting academic norms. Existing data provide an opportunity to think critically about a research question outside its initial parameters, using data collected by you or someone else. Furthermore, having a chance to work with existing data gives researchers an insider view of a complete dataset, its strengths and weaknesses, and its ability to address a research question associated with their research interests. An insider view of a project can also teach budding researchers how to lead their own research projects.

This chapter describes how researchers can use existing data in mixed methods projects to advance education research. We begin by defining existing data for mixed methods using an education research context. We then cover the advantages and disadvantages of using existing data broadly and in mixed methods projects. Next, we present the knowledge-level continuum as a way to frame the exploratory, descriptive, and explanatory research conducted in educational settings. Then, we use examples from education research to describe existing qualitative and quantitative data in mixed methods and how to “mix” these data to answer research questions. Easy-to-follow instructions do not always accompany existing data sets. Therefore, sifting through data records and making sense of the available information associated with existing data sets is an essential part of the process. We end with a discussion of how existing data can support the deepening of knowledge and the breadth of educational topics researchers and practitioners can explore.

What are existing data in mixed methods?

Mixed methods are the rigorous and epistemological application and integration of qualitative and quantitative research approaches to draw interpretations based on the combined strengths of both approaches to influence research, practice, and policy (Plano Clark and Ivankova, 2016; Creswell, 2015; Watkins and Gioia, 2015). John Creswell (2015) defines mixed methods as:

An approach to research in the social, behavioral, and health sciences in which the investigator gathers both quantitative (closed-ended) and qualitative (open-ended) data, integrates the two, and then draws interpretations based on the combined strengths of both sets of data to understand research problems (p. 2).

Previous definitions of mixed methods clearly distinguish between quantitative and qualitative data. However, they rarely note whether these data can come from existing data sources or are newly collected data specific to each mixed methods project (Watkins, 2022). Given the importance of this distinction, we insist that one might use at least two types of data in their mixed methods: primary and secondary data. *Primary data* are collected in real-time to achieve the goals outlined in a research project. Other names for primary data are “original” or “raw” data. *Secondary data*, or “existing” data, are the data that have already been collected by you or someone else that can be used in a research project. Various data sources exist, ranging from large, publicly available data (e.g., US Census data) to smaller, privately-owned data (e.g., local social service agency data of the client population). Whether they include publicly available case studies, data points, papers, artifacts, or electronic documents, existing data are valuable in education research (Logan, 2020). Scholars have gone to great lengths to define mixed methods studies apart from qualitative and quantitative single-method studies (Bazeley, 2018; Creswell and Plano Clark, 2018; Hesse-Biber, 2010; Mertens, 2009, 2017; Morse, 2009). Mixed methods have also been described across various disciplines (Curry and Nunez-Smith, 2015; Johnson and Christense, 2013; O’Cathain, 2010) and professions (Andrew and Halcomb, 2009; Haight and Bidwell, 2016; Johnston, 2012; Magee et al., 2006; Sheperis et al., 2016; Watkins and Gioia, 2015). Yet, sparse resources provide clear guidelines for what mixed methods with existing data involve.

Existing data can be either qualitative or quantitative. Smith’s (2008) book *Using Existing Data in Education and Social Research* provides a concise account for using existing qualitative data in social science research. The author affirms that existing data “... can be numeric or non-numeric. Non-numeric, or qualitative secondary data can include data retrieved second hand from interviews, ethnographic accounts, documents, photographs, or conversations” (pp. 4–5). Though most existing data sources tend to be quantitatively oriented, Smith reminds us of the possibilities of existing qualitative data in our education and social research endeavors. Let’s consider an example pertinent to education research. Charlie is a graduate student interested in examining K-12 students’ sense of belonging within the virtual learning environment since the COVID-19 pandemic. Currently, the knowledge base on virtual learning since the start of the pandemic is still relatively new, so there is little to no data on students’ sense of belonging in the widespread but new virtual learning environment. Charlie’s first step is to finalize the research question, then search for existing

data sources to address the research question. What if Charlie helped collect data on a similar topic for their professor when COVID first began? Does this still count as “existing data?”

Adding a time element to the definitions of primary and existing data is critical when distinguishing between the two because you can collect primary data now that you use for secondary purposes later (Watkins, 2022). There is no hard and fast rule to the length of time that must pass before primary data can be considered “existing” data. The only real distinction involves whether someone has a fundamental research question they want to answer and whether the data to address that question exists. For example, if Charlie collected data 2 years ago and those data can help answer a different research question now, Charlie can use those data to answer the new, different research question. Whether the new research question is related or unrelated to the original study, Charlie can use those existing data for a secondary purpose.

Education research can benefit from combining mixed methods with existing data methods (Fig. 1). Mixed methods involve collecting and analyzing both quantitative and qualitative data to address a research question. Whereas existing data methods involve evaluating quantitative and qualitative data that have been previously collected and deciding how to use them for a secondary purpose. Bringing the two methods together results in mixed methods with existing data. Hence, doing *mixed methods research with existing data* involves identifying, evaluating, and incorporating one or more existing data sources into a mixed methods project (Watkins, 2022). Incorporating existing data into mixed methods takes the definition of mixed methods to a new level, as mixed methods are ways to collect, analyze, and integrate qualitative and quantitative data in rigorous and theoretically sound ways to encompass the breadth and depth of a topic (Creswell, 2015; Guetterman et al., 2015; Tashakkori and Teddlie, 2010).

Using existing data in mixed methods also expands the purpose of secondary data analysis, which is to further analyze an existing data source by addressing a research question like (or distinct from) that for which the original data were collected (Hewson, 2006). Increasing our use of existing data sources adds to the unique possibilities for the future of mixed methods in education research. Previous studies suggest existing data are typically used at the beginning (or for the first phase) of a mixed methods study (Hewson, 2006; Johnston, 2014). However, as qualitative and quantitative data sources become more robust and sophisticated, the potential for implementing entire mixed methods studies using education data previously collected for a different purpose is promising.

Using data collected by you or someone else to answer your research questions is common in education research and academic settings (Davis-Kean and Jager, 2017; Hewson, 2006; Panchenko and Samoilova, 2020). It can maximize a project’s resources (e.g., time, resources, and energy) beyond those used for its original purpose. The purpose statement, research questions, and theory should guide the decision to use existing data in a mixed methods project. Further, applying or generating theory in mixed methods with existing data will vary depending on whether the research prioritizes the quantitative or qualitative phase of the study. For instance, let’s revisit the education research example presented above. Suppose Charlie wants to investigate students’ sense of belonging within the virtual learning space, and there is little pre-existing knowledge about students’ sense of belonging specific to virtual learning. In that case, Charlie must prioritize a research approach that allows them to gather more comprehensive insight into how students describe their connectedness and belonging to peers and educators within the virtual space. Charlie might use mixed methods with existing data for this study. Moreover, Charlie should also review the advantages and disadvantages of using mixed methods with existing data before proceeding.

Advantages and disadvantages of using existing data in mixed methods

There are several advantages to using existing data in mixed methods. For example, using existing data to address a mixed methods question may save time, money, and the resources necessary to collect primary data. Also, researchers might decide to use existing data because they have access to a high-quality and robust data source. Students are drawn to existing data for their mixed methods because the data can extend their current scholarship. Deep diving into someone else’s data can expand a researcher’s training and scope. Though one’s education, training, and skills are unique to them and allow them to view the world through a particular lens, this uniqueness might also limit the researcher’s contribution to research. In short, using existing data can enhance a researcher’s scope and abilities.

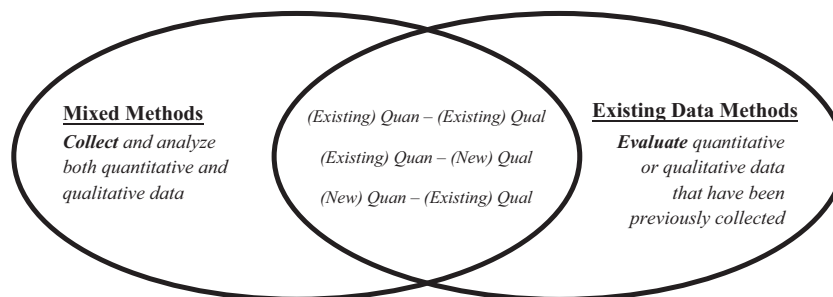


Fig. 1 The merging of methods: mixed methods with existing data.

Despite the advantages of using existing data in mixed methods, there are also some challenges. For example, some researchers might find using someone else's data mysterious and unclear. Budding researchers may also question their authenticity because they did not undergo all the steps in the research process (e.g., idea conceptualization, IRB application, recruitment, instrument development, data collection, data analysis, reporting, etc.) to achieve their project goals. Then there is the matter of missing information. Sometimes existing data files are incomplete, leaving researchers who want to use the data for secondary purposes confused about the procedures undertaken by the original research team and uncertain about the best way to repurpose the data to answer a new research question. Despite the challenges associated with existing data, they can be valuable for education research and can help streamline the path to answering the research questions. However, it is best to start at the beginning of scientific inquiry to address a research topic to do this well. This process involves deepening your understanding of your topic and the ways mixed methods with existing data can enhance exploration in an educational context. To do this, researchers must examine the topic, and the possibility of using existing data for a mixed methods study in the context of the knowledge-level continuum.

Existing quantitative data in mixed methods

Though the popularity of existing qualitative data has increased (Beck, 2019; Hughes and Tarrant, 2020; Largan and Morris, 2019), historically, education researchers who have used existing data in their studies have been guided by the quantitative data. In other words, the quantitative data are the dominant or prioritized phase in the mixed methods. Theory will guide the variables used if the researcher prioritizes the quantitative study component in mixed methods research. In sequential designs, prioritized phases occur before other phases of the study. Suppose the prioritized quantitative step involves using an existing dataset. In that case, the theory used by the original investigators will guide your understanding and selection of the pre-existing variables in those data to help answer your research questions.

Existing quantitative data can come in many forms, from thousands of individual high school student cases to a dozen or so school administrative cases in your office building. Existing quantitative data are everywhere and include population-based studies, which use methods to survey all people within a specific population (e.g., US census). Other examples of existing quantitative data are cohort-based and longitudinal surveys, such as the National Survey of American Life, conducted from 2001 to 2003 (Jackson et al., 2004). Though previously collected sources exist, researchers still need to request access to the data or investigate if the data are publicly available. A Google search of "data repository" would help you start your journey to identify publicly available existing quantitative data.

Gauge the quantitative "breadth" of a topic, then proceed qualitatively

Existing quantitative data in a mixed methods study can help you gauge the breadth of a topic before you proceed with the qualitative inquiry. Quantitative inquiry uses deductive reasoning to test and confirm (or refute) a theory (Grinnell and Unrau, 2018). This process usually begins with a theory and then establishes a plan to confirm or refute that theory. Given the sheer number of cases that can be observed using quantitative methods, there are apparent advantages to analyzing existing quantitative data. One clear advantage is gauging the quantitative breadth of a topic before deciding how to proceed qualitatively. For example, let's assume you are a data analyst at Booker T. Washington High School. Your assistant principal has asked you to explore possibilities with the school's confidential student survey data. You decide to use these existing quantitative data to understand what adolescents think about the social-emotional learning services offered by Booker T. Washington High School. After you get descriptive data about students' feelings about social-emotional services, this information could be strengthened by a small, qualitative study in which you assemble groups of students who represent the demographics reflected in your existing quantitative analysis and then ask them to expound on your quantitative findings with new qualitative responses.

Existing qualitative data in mixed methods

In qualitative research, inductive reasoning identifies patterns that generate theory. Compared to mixed methods with a dominant quantitative component, theory plays a different role with a dominant qualitative component. Rather than guide the selection of variables to test the theory, when using existing data in mixed methods, the data generate theory, or in most cases, a conceptual framework that can be tested with quantitative data. In this context, a *theory* explains the "how" or "why" of phenomena, typically referencing existing facts (Kimmons, 2022). A theory is usually based on an existing framework and is frequently tested. On the other hand, a *conceptual framework* is a collection of concepts and empirical findings from the literature used to show relationships among ideas and how they relate to research. For the purposes of this chapter, conceptual frameworks are "preliminary theories" you hope to evolve into testable explanations for the how and why of a particular phenomenon. When beginning mixed methods with a dominant qualitative component, you might consider building a preliminary conceptual framework (to test as a theory in a subsequent study). Whether you plan to generate theory (e.g., qualitative analysis) or test theory (e.g., quantitative analysis), the theory has a role in mixed methods that should be embedded throughout your research.

Existing qualitative data (e.g., text, image, etc.) provide insight and help address some of your unanswered qualitative research questions. Though not as popular as their quantitative counterpart, existing qualitative data has gained some traction from previous scholars (Bishop, 2007, 2009; Fielding and Fielding, 2000; Hammersley, 2009; Hinds et al., 1997). In particular, and given the increased respect, rigor, and credibility of qualitative data in recent years, more scholars consider its use in an existing analysis

capacity. In fact, in 2019 and 2020 alone, three books were published on secondary analysis of qualitative data (Beck, 2019; Hughes and Tarrant, 2020; Largan and Morris, 2019). Despite this progress, a comprehensive appreciation for existing qualitative data lags that of existing quantitative data. This seems to be due to the plethora of resources and guidelines for using existing quantitative data compared to the sparse resources available for using existing qualitative data. Some examples of existing qualitative data may include video recordings, audio recordings, and text documents (i.e., transcripts) of previous interviews with students.

Gauge the qualitative “depth” of a topic, then proceed quantitatively

Qualitative research uses inductive reasoning, beginning with observations and ending with theory. Therefore, if information from an existing qualitative data source aligns with the information needed to address your research question, you should consider using these existing qualitative data as a source for your mixed methods. There is an implied depth associated with qualitative research; therefore, you can immerse yourself in the data, conduct a rigorous qualitative analysis, then generate a conceptual framework (i.e., preliminary theory) that connects relevant concepts about your topic of interest. Deep inquiry into existing qualitative data can serve as a springboard for building a new quantitative study component into your mixed methods project.

Let’s consider another example where your mentor owns a qualitative data source on educational outcomes of court-appointed youth, and you are interested in exploring these data further. You and your mentor study this topic, but he collected focus group data three years ago and recently received a Department of Education grant to which he needs to turn his attention for the next 5 years. He offers the data to you to help advance your scholarship in this area. Using your mentor’s existing qualitative data can help you gauge the depth of the educational outcomes of court-appointed youth so that you can know how you want to proceed quantitatively. Namely, you could analyze the focus group data to generate a list of concepts and themes for which a conceptual framework can be created. After which, you could collect new quantitative data (including items that measure the concepts and themes generated from your qualitative analysis) to generate hypotheses and test your conceptual framework. These ideas are ways to enhance the use of your mentor’s existing qualitative data by repurposing them so that you have some direction in your research.

Both quantitative and qualitative data have advantages and disadvantages, but sometimes researchers are reluctant to use qualitative data as an existing data source. This is unfortunate given the plethora of possibilities for open-ended inquiries into the human experience. For example, there are benefits to assessing social and cultural contexts, connections, relationships, and subjective processes that result from social phenomena (Watkins and Gioia, 2015). Furthermore, uncovering the factors that influence decision-making and the opinions of research participants will strengthen your research, not weaken it.

The knowledge-level continuum in mixed methods

To understand how mixed methods with existing data can be helpful in education research, start by making decisions about your study based on the current research. This framing of education research can help because it addresses two fundamental questions when using existing data: (1) What is the existing *knowledge* on the topic? (2) What are the existing *data* on the topic? But before we ask these questions, we should first ask, what is the knowledge-level continuum and how is it helpful in guiding mixed methods with existing data in education research?

The *knowledge-level continuum* is a framework (Fig. 2) for describing the current state of science in an area of inquiry and the progression of science therein (Watkins, 2022). It outlines existing knowledge, and the ways existing knowledge merges with the organization, creation, and dissemination of new knowledge (Watkins, 2022). Previously, the knowledge-level continuum has been used in other disciplines, such as information technology (Vassallo, 1999) and social work (Grinnell and Unrau, 2018; Watkins and Gioia, 2015). Research inquiry can fall anywhere along the continuum depending on how much is already known about the topic. The knowledge-level continuum encourages researchers to think about research as being a direct by-product of the information about a topic and the purpose of the study (Grinnell and Unrau, 2018; Watkins, 2022). In short, if there is little to no pre-existing knowledge of your topic of study, an exploratory level of inquiry should be used. If some pre-existing knowledge exists, a descriptive level of investigation should be used. If there is a robust amount of pre-existing knowledge, an explanatory level of inquiry should be used (Watkins, 2022).

Let’s revisit the previous example with Charlie to understand how the knowledge-level continuum can be applied to education research. Charlie is interested in examining students’ sense of belonging in virtual learning spaces. Specifically, Charlie is interested in reviewing the factors that support and limit students’ sense of belonging in the new virtual learning environment. Let’s say Charlie completes a thorough literature search, looking through dozens of peer-reviewed article databases. Unfortunately, very little published research examines students’ sense of belonging within virtual learning spaces for K-12 students.

Additionally, of the very little information published on the topic, Charlie found it to be of poor quality, leading to challenges regarding Charlie’s trust in the rigor and quality of the research. Given the limited amount of pre-existing research and the poor quality of the research, if Charlie were initiating a research project in this area, they would use an exploratory level of inquiry. *Exploratory research* is the first level of the knowledge-level continuum. It includes research that begins to fill the substantial gap in the existing literature related to the topic of interest. Exploratory research can be conducted using qualitative methods of inquiry. For the example described, Charlie may begin with interviewing all students at a specific school who participated in virtual learning and then asking about their sense of belonging.

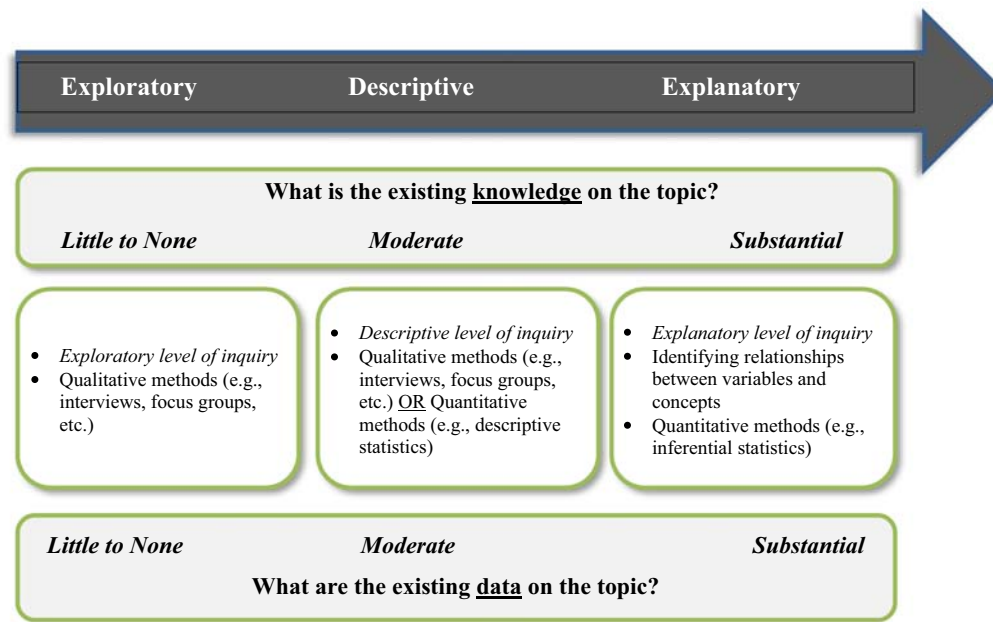


Fig. 2 The knowledge-level continuum as a framework for mixed methods with existing data.

Considering the same example above, let's say Charlie continued their search for peer-reviewed articles that demonstrate students' sense of belonging in virtual learning spaces and cannot locate articles. However, Charlie's faculty advisor sends them eight peer-reviewed articles she found while working on a grant. As Charlie reviews the articles, they are excited to see the articles synthesize a lot of the information they hoped to find. The articles also describe preliminary findings on college students participating in online classes and their sense of belonging. However, none of the articles include results for K-12 students and their sense of belonging in virtual learning spaces during the unprecedented time of the COVID-19 pandemic. In short, there are still some gaps in research that focus on this topic.

In reviewing the knowledge-level continuum, Charlie acknowledges that while *some* information exists on the topic of interest, the existing literature does not extend to cover the specific topic Charlie was hoping to investigate. Given this, Charlie decides that a descriptive level of inquiry would be best. *Descriptive research* is the second level of research within the knowledge-level continuum. It involves expounding on the pre-existing knowledge by describing the sample and concepts of interest in more detail. Descriptive research can be conducted using either qualitative or quantitative methods of inquiry. In the case above, Charlie may use an existing sense of belonging psychometric tool to measure students' sense of belonging in the virtual learning environment.

If the existing literature included substantial knowledge of sense of belonging, specifically for K-12 students that participated in virtual learning during the unprecedented 2020–2021 academic year, then Charlie would use an explanatory level of inquiry for the research. *Explanatory research* is the highest level of the knowledge-level continuum and is conducted using quantitative methods of inquiry because previous data exists on a topic. If Charlie conducted an exploratory study (i.e., first level of the knowledge level continuum) to understand the students' sense of belonging in virtual learning environments and then conducted a descriptive study (i.e., secondary level of the knowledge continuum) to collect more detailed information about students and educators' sociodemographic background and characteristics, then existing data exists on the topic! Therefore, an explanatory study can help Charlie understand the associations between concepts identified in the exploratory and descriptive studies on this topic. Explanatory research studies produce inferential statistics that Charlie can generalize to the larger population.

Applying the knowledge-level continuum to mixed methods with existing data

The knowledge-level continuum can be used as a framework for mixed methods research with existing data because it both informs, and is informed by, how much is already known about the topic of interest. Let's consider an example where a researcher, Marcus, is trying to decide whether to use a convergent, exploratory sequential, or explanatory sequential mixed methods design for a study with existing data (Creswell, 2015; Watkins, 2022). In this example, Marcus wants to use existing quantitative data on SAT scores for Deferred Action for Childhood Arrivals (DACA) students who attend ivy league colleges and universities in the United States.

The first step is for Marcus to search for previous knowledge and data on the topic. He might decide to check Google Scholar, Psych Info, EBSCOhost, and Science Direct for articles that include keywords in the title and abstract. Then, he might finalize his research question (*How do SAT scores for DACA students who attend ivy league colleges and universities compare to SAT scores for other*

DACA students at other colleges and universities at their colleges and universities?). After the research question is finalized, Marcus should consider whether the question is suitable for integrating his existing data with a second, qualitative data source. Marcus must determine which mixed methods design (Fig. 3) will help address his research question. He can do this by examining the research question, the current state of science (i.e., previous knowledge and data on the topic), and the existing data on the subject (i.e., existing data on DACA students' SAT scores).

If Marcus finds little to no existing research on the topic, he has the option to conduct an exploratory sequential design. However, given the variables in his existing dataset, he could likely build a convergent design where the quantitative phase involves analyzing the existing data on SAT scores for DACA students at ivy league colleges and universities. His qualitative phase could include online interviews with a select number of DACA students at ivy league colleges and universities across the United States. Though his existing quantitative data are more descriptive, and he could also conduct an explanatory sequential design, he would prefer to do a convergent mixed methods study where the new qualitative findings can be merged with the existing quantitative data results.

Three questions when choosing existing data in mixed methods research

Before incorporating existing data in your mixed methods research, you must first identify the most appropriate existing data for your project. You may be asking yourself: How do I choose the best existing data for my mixed methods project? Below, we provide three questions you should ask yourself as you review potential data sources.

Question 1: How will you define the terms in your research question?

The first question you should ask yourself will help outline the plan of action for your mixed methods project. When looking at your research question, what does each word mean? In other words, how will you define the terms in your research question for your mixed methods with existing data? Essentially, what you need to uncover is what is under investigation in your study. As you choose specific terminology for your research question, be sure to determine how you will define and measure these terms. The words you use to develop your research question will outline an action plan for your mixed methods with existing data. In the earlier example, Charlie can perform a thorough literature review of students' sense of belonging, then choose terms for their research question that signal the study's exploratory nature. Therefore, Charlie's research question could be *What are the barriers and facilitators to students' sense of belonging in virtual learning spaces?* Looking closely at this research question, we would advise Charlie to highlight words that need a definition and are directly connected to the success of their study. For example, terms such as "barriers," "facilitators," "students," "sense of belonging," and "virtual learning spaces" all need clear definitions. They also need units of measurement as the success of Charlie's mixed methods is contingent upon a deeper understanding of and the ability to measure these terms.

Question 2: How do you plan to find the existing data?

There may be field-specific data repositories for existing data; however, you must actively seek out information to help locate existing data sources regardless of your field of study. The next question you ask yourself should point you toward existing data that will help answer your research question. Begin by asking your mentors, advisors, professors, and classmates for recommendations for data sources related to your well-defined research question. Additionally, use institutional resources, such as academic librarians, to strategize the methods used to search the web and online databases. It is important to note that the more defined your research question is, including your concepts and variables, the easier it is for others to assist with your search. You want to develop a plan or strategy to locate the data at this stage. Which search terms will you use? What strategies (e.g., filters, Boolean connectors) will you use to search across data repositories? What data depositories can you access? In Charlie's case, they could retrieve existing

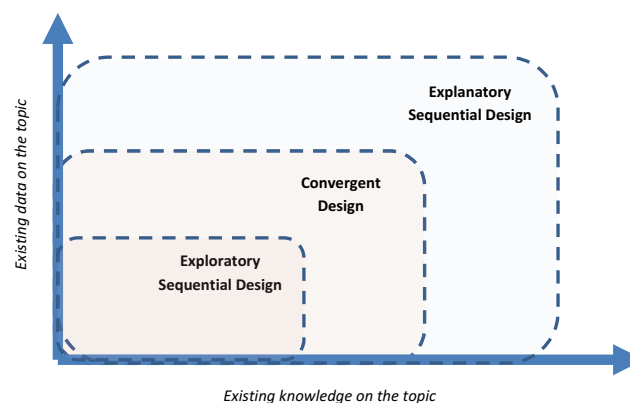


Fig. 3 Choosing mixed methods designs that inform, and are informed by, existing data and the knowledge-level continuum.

data from their professor, so thankfully, Charlie did not need to look very far to locate existing data. However, it will be necessary for Charlie to closely examine the appropriateness of their professor's existing data to ensure that the research question (*What are the barriers and facilitators to students' sense of belonging in virtual learning spaces?*) can be addressed. This leads us to the final question.

Question 3: Do the existing data help answer the research question?

Now that you have defined the terms in your research question and developed a plan for locating existing data, you must address the final question regarding the appropriateness of the existing data for your mixed methods project. Evaluate the existing data to determine if it has adequate information to answer your research question (Kiecolt and Nathan, 1985; Smith, 2008, 2011). This final question aims to ensure you have thoroughly assessed the fit of the existing data for answering the research question *before* incorporating it into your mixed methods. Go through the existing data, supplemental materials, and any previously published reports on the data so you can learn about the primary study, from conceptualization to evaluation. So Charlie will need to conduct a rigorous review of the variables (if quantitative), concepts reviewed (if qualitative), and all the professor's supplemental materials on file. If there are materials the professor does not have, Charlie should try to locate these materials to help them develop a comprehensive knowledge base about the data. Charlie should be able to answer questions such as: What was the original purpose for the data? How were participants recruited? How were relevant concepts measured? How was missing data handled?

In sum, there are three questions to address when choosing to include existing quantitative or qualitative data in your mixed methods study. This strategic process begins with you asking yourself questions that will help you explicitly define your research question. Next, you must develop a plan to locate existing data that includes the variables or concepts captured by your research question. Finally, once you locate the existing data, you must assess if they meet the needs of your research question. These steps provide a roadmap and an action plan for selecting existing data.

Conclusion

This chapter provides an overview of how mixed methods with existing data can advance education research. The knowledge-level continuum can serve as a framework for your mixed methods study. It describes the research process for all studies, depending on how much is already known about the topic of interest. The knowledge-level continuum is a direct by-product of the existing information on a topic, the purpose of the research, and your anticipated next steps. Using existing data in mixed methods includes identifying, evaluating, and incorporating one or more existing data sources to serve as one or more data phases of your mixed methods. It acknowledges not only the purpose of mixed methods but also the purpose of existing data analysis, which is to further the investigation of an existing data source by addressing a research question similar to (or distinct from) that for which the original data were collected.

Using existing data is advantageous in mixed methods because it allows you to expand your knowledge and understanding of a topic, gauge the qualitative depth of a topic so you know how to proceed quantitatively, and gauge the quantitative breadth of a topic so you know how to proceed qualitatively. Almost any research question in education research would benefit from an existing data source. Access to an existing study's protocols, guidelines, training materials, codebooks, and field notes can make you feel like you were part of the original research team. Therefore, consider locating an existing data source before agreeing to collect data for a new mixed methods project.

References

- Andrew, S., Halcomb, E.J. (Eds.), 2009. *Mixed Methods Research for Nursing and Health Sciences*. Wiley-Blackwell Publishing, Chichester, UK.
- Bazeley, P., 2018. *Integrating Analyses in Mixed Methods Research*. Sage Publications, Thousand Oaks, CA.
- Beck, C.T., 2019. *Secondary Qualitative Data Analysis in the Health and Social Sciences*. Routledge, London, UK.
- Bishop, E., 2009. Ethical sharing and reuse of qualitative data. *Aust. J. Soc. Issues* 44 (3), 255–272.
- Bishop, L., 2007. A reflexive account of reusing qualitative data: beyond primary/secondary dualism. *Socio. Res. Online* 12 (3), 43–56.
- Chawinga, W.D., Zinn, S., 2020. Research data management at an African medical university: implications for academic librarianship. *J. Acad. Librariansh* 46 (4), 102161.
- Creswell, J.W., 2015. *A Concise Introduction to Mixed Methods Research*. Sage Publications, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage Publications, Thousand Oaks, CA.
- Curry, L., Nunez-Smith, M., 2015. *Mixed Methods in Health Sciences Research: A Practical Primer*. Sage Publications, Thousand Oaks, CA.
- Davis-Kean, P., Jager, J., 2017. Using secondary data analysis for educational research. In: Wyse, D., Selwyn, N., Smith, E., Suter, L.E. (Eds.), *The BERA/SAGE Handbook of Educational Research: 2 Volume Set*, first ed. Sage Publications, London, UK, pp. 505–522.
- Dorn, E., Hancock, B., Sarakatsannis, J., Viruleg, E., 2020. COVID-19 and Student Learning in the United States: The Hurt Could Last a Lifetime. McKinsey & Company, p. 1.
- Fielding, N.G., Fielding, J.L., 2000. Resistance and adaptation to criminal identity: using secondary analysis to evaluate classic studies of crime and deviance. *Sociology* 34 (4), 671–689.
- Grinnell Jr., R.M., Unrau, Y.A., 2018. *Social Work Research and Evaluation: Foundations of Evidence-Based Practice*, eleventh ed. Oxford University Press, New York, NY.
- Grooms, A.A., Childs, J., 2021. "We need to do better by kids": changing routines in U.S. schools in response to COVID-19 school closures. *J. Educ. Stud. Placed A. T. Risk* 26 (2), 135–156.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13, 554–561.
- Haight, W., Bidwell, L.N., 2016. *Mixed Methods Research for Social Work: Integrating Methodologies to Strengthen Practice and Policy*. Oxford University Press, Oxford, UK.

- Hammersley, M., 2009. Can we re-use qualitative data via secondary analysis? Notes on some terminological and substantive issues. *Socio. Res. Online* 15 (1), 47–53.
- Hesse-Biber, S.N., 2010. *Mixed Methods Research: Merging Theory With Practice*. Guilford Press, New York, NY.
- Hewson, C., 2006. Secondary analysis. In: Jupp, V. (Ed.), *The Sage Dictionary of Research Methods*, first ed. Sage Publications, London, pp. 274–275.
- Hinds, P., Vogel, R., Clarke-Steffen, L., 1997. The possibilities and pitfalls of doing secondary analysis of a qualitative data set. *Qual. Health Res.* 7 (3), 408–424.
- Hughes, K., Tarrant, A., 2020. An introduction to qualitative secondary analysis. In: Hughes, K., Tarrant, A. (Eds.), *Qualitative Secondary Analysis*, first ed. Sage Publications, London, UK, pp. 3–18.
- Jackson, J.S., Torres, M., Caldwell, C.H., Neighbors, H.W., Nesse, R.M., Taylor, R.J., Trierweiler, S.J., Williams, D.R., 2004. The national survey of American life: a study of racial, ethnic, and cultural influences on mental disorders and mental health. *Int. J. Methods Psychiatr. Res.* 13 (4), 196–207.
- Johnson, R.B., Christensen, L.B., 2013. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. Sage Publications, Thousand Oaks, CA.
- Johnston, M.P., 2012. School librarians as technology integration leaders: enablers and barriers to leadership enactment. *Sch. Libr. Res.* 15 (1), 691–693.
- Johnston, M.P., 2014. Secondary method: a method of which the time has come. *Qual. Quan. Methods Librar.* 3 (3), 619–626.
- Kiecolt, K.J., Nathan, L.E., 1985. *Paper Series on Quantitative Applications in the Social Sciences: Secondary Analysis of Survey Data*. Sage Publications, Newbury Park, CA.
- Kimmons, R., 2022. *Education Research Across Multiple Paradigms*. EdTech Books.
- Largan, C., Morris, T., 2019. *Qualitative Secondary Research: A Step-by-Step Guide*. Sage Publications, London, UK.
- Logan, T., 2020. A practical, iterative framework for secondary data analysis in educational research. *Aust. Educ. Res.* 47 (3), 129–148.
- Magee, T., Lee, S.M., Giuliano, K.K., Munro, B., 2006. Generating new knowledge from existing data: the use of large data sets for nursing research. *Nurs. Res.* 55 (2), S50–S56.
- Mertens, D., 2009. *Transformative Research and Evaluation*. Guilford Press, New York.
- Mertens, D.M., 2017. *Evaluation in Practice Series: Mixed Methods Design in Evaluation*. Sage Publications, Thousand Oaks, CA.
- Morse, J.M., 2009. Mixing qualitative methods. *Qual. Health Res.* 19 (11), 1523–1524.
- O’Cathain, A., 2010. Assessing the quality of mixed methods research: toward a comprehensive framework. In: Tashakkori, A., Teddlie, C. (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Thousand Oaks, CA, pp. 531–558.
- Panchenko, L., Samoilova, N., 2020. Secondary Data Analysis in Educational Research: Opportunities for Ph.D. Students, pp. 1–7. *SHS Web of Conferences* 75.
- Plano Clark, V.L., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Sheperis, C., Young, J., Daniels, M., 2016. *Counseling Research: Quantitative, Qualitative, and Mixed Methods*, second ed. Pearson, Boston, MA.
- Smith, E., 2008. *Using Secondary Data in Educational and Social Research*. McGraw-Hill Education, New York, NY.
- Smith, E., 2011. Special issue on using secondary data in educational research. *Int. J. Res. Method Educ.* 34, 219–221.
- Tashakkori, A., Teddlie, C. (Eds.), 2010. *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage Publications, Thousand Oaks, CA, pp. 581–611.
- Vassallo, P., 1999. The knowledge continuum - organizing for research and scholarly communication. *Internet Res.* 9 (3), 232–242.
- Watkins, D.C., 2022. *Secondary Data in Mixed Methods Research*. Sage Publications, Thousand Oaks, CA.
- Watkins, D.C., Gioia, D., 2015. *Pocket Guides to Social Work Research Methods Series: Mixed Methods Research*. Oxford University Press, New York.

Using mixed methods research with Geographic Information Systems (GIS)

Nazgol Bagheri, University of Texas at San Antonio, San Antonio, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	645
Defining GIS and QGIS	645
Using QGIS as a mixed methods approach in education	647
GIS as a teaching tool in the classroom	648
GIS as a research tool outside the classroom	648
An application of QGIS: Women's perceptions and experiences in Tehran's public spaces	649
Advantages and limitations of using QGIS as a mixed methods research approach	650
Conclusion	652
Acknowledgments	652
References	653

Introduction

In this chapter, I introduce Geographic Information Systems (GIS) and its empowering applications in research. GIS is a computer-based tool for mapping, measuring, monitoring, modeling, and managing spatially referenced data. GIS can be considered a database system and a set of spatial visualization and analytic operations for working with such a database. Quantitative and qualitative data can be combined in GIS to produce a fuller understanding of our world. This integration of qualitative data in GIS is called Qualitative GIS (QGIS). In a sense, QGIS may be thought of as a higher-order map in the context of mixed methods research (MMR), as it enables us to overlay different layers of data.

Similar to early proponents of mixed methods research in social sciences, critical and feminist geographers who have advocated for QGIS applications (for early QGIS works see: Pickles, 1995; Schuurman, 2000; Sheppard, 2001; Elwood, 2002; Pavlovskaya, 2006; Kwan, 2004) challenge the dualism separating quantitative and qualitative methods while highlighting the importance of both human (soft) and physical (hard) sciences (Maxwell, 2010; Fielding, 2012). In other words, as Fielding and Cisneros-Puebla (2009) suggest “spatially oriented qualitative social science and qualitatively oriented geography” (p. 349)—known as QGIS in GISciences—are more recent, innovative examples of combining technology with ethnography in mixed methods research handovers. QGIS emphasizes integrating multiple forms of data to create a bigger picture and greater insight into what is happening (and why) in the study area.

This chapter starts by revisiting the definitions of GIS and QGIS. I will then contextualize the application of GIS and QGIS in education, both inside and outside the classroom. Through several examples of classroom activities, I will illustrate how GIS can be used by teachers inside the classroom for promoting interactive learning skills. In addition, I will review a few research case studies in which educational scholars, school administrators, and public policy makers have used GIS to address school districts' academic performance and social (in)justice issues. Next, in order to better illustrate the integration of different types of data in GIS, I will rely on my long-term methodological research combining GIS and ethnographic methods focusing on women's social perceptions and experiences of public spaces in Tehran, Iran. My goal here is to acknowledge the methodological opportunities of using mixed methods approaches such as QGIS by both human and physical sciences researchers. I will examine the advantages and limitations of using QGIS as a mixed methods approach, highlighting the challenges I came across throughout my research project in Tehran. I end the chapter by contending that QGIS is an invitation to scholars in social sciences (e.g., human geography and anthropology) to benefit from GIS spatial visualization and technical analysis tools. It is also a recommendation for hard science scholars (e.g., environmental sciences, geology and physical geographers) to challenge the inherently quantitative, positivist framework of understanding space and to adopt a more mixed method approach in their research.

Defining GIS and QGIS

GIS according to the National Center for Geographic Information and Analysis (see <http://www.ncgia.ucsb.edu/>), is a system of hardware, software, and procedures designed to support the capture, management, editing, analysis, modeling, and displaying of geographically referenced data for solving complex planning and management problems. Key components of GIS include computer hardware (e.g., desktop computers, laptops), specialized software (e.g., ESRI ArcGIS, Online GIS, Free and Open Source GIS web and mobile apps), geographically referenced data (e.g., demographic, social, economic, political, racial composition, household median income), and people who are involved in all phases of making and reading maps. GIS is used to map, measure, monitor, model, and manage spatially referenced data. Overlay tool, or the concept of stacking layers containing different kinds of data and comparing them with each other based on where things are located is the foundational concept of spatial analysis. For

a more visual understanding of a GIS map, as I often say in the first week of teaching Introduction to GIS, it resembles “a multi-layered sandwich made of data,” as shown in Fig. 1. In GIS each dataset is managed as a layer and can be graphically combined with other layers. By combining layers using spatial analysis and visualization operations, GIS enables us to explore critically important questions and possibly find answers to those questions. GIS basically offers a systematic way of overlaying all sorts of datasets together; then, through this superimposed composition of layers, spatial relationship among different layers, spatial patterns, and future trends can be identified, explained, and predicted.

QGIS refers to an array of methodological efforts to incorporate into GIS more qualitative data than was traditionally included. QGIS can also be placed among those post-positivist, poststructuralist, and non-representational approaches in contemporary cultural geography that strive to reconcile “social-cultural and spatial-analytical geographies” (Kwan, 2004, p. 756). Although many feminist and critical geographers (see Pratt, 1989; Sheppard, 2001; Moss, 2002; Kwan, 2002, 2004; Cope and Elwood, 2009) have already challenged the dualism separating quantitative and qualitative geography, innovative mixed methods approaches such as QGIS (that includes geo-ethnography, feminist data visualization, grounded-visualization, and geo-narrative) are still in their early stages. An increasing number of researchers integrated GIS-based spatial visualization and analysis into qualitative methods, such as ethnography. Photographs, sketches, mental maps, social behavior maps, audio and video records, field notes, interviews, and narratives gathered through qualitative methods can be linked to spatially referenced data in GIS software and offer new ways of representing individuals’ lived experiences (Al-Kodmany, 2002; Aitken and Kwan, 2009). For example, McLafferty (2002) narrated a story of how women affected by and concerned with breast cancer used GIS to explore the relationship between environmental hazards and the geographical distribution of disease in suburban Long Island, New York. QGIS were used to map the environmental factors that were correlated to increasing prevalence of breast cancer in their neighborhood as well as qualitative data collected during interviews with the affected women and their community members. McLafferty (2002) concluded that GIS was “more than a computer-based technology” and could not only serve as “a vehicle for women’s empowerment” (McLafferty, 2002, p. 268) but empower other underrepresented populations as well. In another study asserting the value of grounded theory in urban cultural research, Knigge and Cope (2006, 2009) applied what they call “Grounded Visualization” as an integrated method that enabled them to explore spatially referenced data in close relationship to ethnographic and multi-media data for the study of community gardens and the meanings of vacancy in Buffalo, New York.

Cieri (2003) combined 3D GIS and qualitative data to study lesbian public spaces in Philadelphia. What made her successful in such a critical endeavor, however, was her application of mixed methods approach, similar to Knigge and Cope’s grounded visualization. As illustrated from the natural resources management, Nightingale (2003) claimed to go beyond the limitations of partial and situated knowledges imposed on feminist geography by using QGIS in tracing forest change in Nepal. Through her study of gender and individual access to urban employment opportunity in Columbus, Ohio, Kwan (2002), introduced the concept of

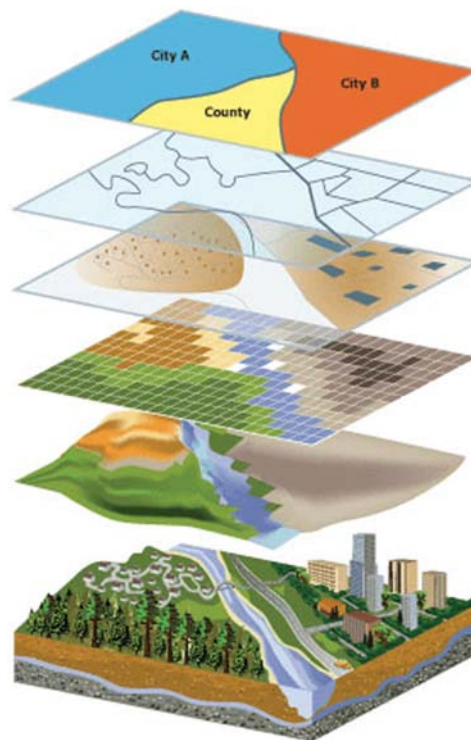


Fig. 1 Overlaying different data layers on a GIS map. Source: <https://www.esri.com/arcgis-blog/products/product/analytics/how-to-perform-spatial-analysis/>.

“Feminist Visualization”. Kwan and Ding (2008) and Kwan (2008) were excellent examples of emotional geographic studies in which the “Geo-narrative” method—integrating GIS’s spatial analysis and narratives—was used to examine the effects of anti-Muslim hate violence on Muslim women’s lives in Columbus, Ohio, in the post-September 11 period. In another study, Watts (2010) reconciled on-the-ground qualitative data with point-distribution GIS data to investigate participants’ socio-spatial behaviors in Los Angeles’s public protests. Jones and Evans (2012) encouraged mobility scholars to take advantage of QGIS application to better examine the role of space in data creation.

Coupling GIS techniques with ethnographic methods was also referred to as “Geo-ethnography” by Matthews et al. (2005) in their ongoing study of low-income families and their children and welfare reform in Boston, Chicago, and San Antonio. While examining job accessibility among working poor members in Columbus, Ohio, and feelings of safety among Lesbian, Gay, Bisexual, Transgender and Queer people in St. Louis, Missouri, Boschmann and Cubbon (2014) digitized and transformed qualitative sketch maps using GIS. Sketch maps are subjective, hand-drawn outlines of people’s observation and perceptions and are not from exact survey measurements. They concluded that, while such a process is time consuming, it can achieve QGIS’ objective of “collecting unique spatial data of individual experiences, visualizing socio-spatial processes, breaking down particular barriers of positionality in research, and developing new uses of GIS” (Boschmann and Cubbon, 2014, p. 3).

QGIS can be considered an umbrella term for the mixed methods approaches described above. “The qualitative” can be associated with the qualitative nature of either the data gathered (e.g., interviews or sketches) or the analytic methods used in the research (e.g., content analysis, triangulation, and grounded theory) (Jung and Elwood, 2010, p. 67). Despite different approaches of integrating multiple forms of data, this review of literature highlighted the common theme of creating a bigger picture and greater insight into the current research in this area. These approaches can also empower political research methods, through which the stories of subaltern and/or marginalized groups including women, children, the disabled, LGBT, racial, and religious minorities can be heard and included in the research process. Often, these methods are used in community-based and participatory planning projects where a deeper understanding of context and the groups who are being represented is essential. QGIS—like any other mixed methods approach—can improve our understanding of space from rigid to fluid, through including ethnographic data such as people’s experiences. In this approach, research becomes truly an open-ended emergent process and partial, contextual, and situated (Haraway, 1991) rather than the set of predictable determined steps often advocated by GIS scientists. In what follows, I discuss the use of QGIS as a mixed methods approach in education and present a case study of using QGIS in a research project conducted in Tehran, Iran, an examination of women’s social perceptions and experiences of public spaces, and their symbolic interpretations of modern and traditional public spaces developed by urban planners. This case study is not intended to present a full account of women’s conceptualization of modern and traditional public spaces in Tehran, but rather to illustrate how QGIS can be applied as a mixed methods approach.

Using QGIS as a mixed methods approach in education

Research over the past two decades has emphasized the value of story mapping and geo-narrating in scientific communications for non-expert audiences as well as teaching science to the students in all levels of educational system (see Drennon, 2005; National Research Council, 2005; Bodzin and Anastasio, 2006; Artvinli, 2010; McCloughlin, and Matthews, 2012; Hong, 2014; Kerski, 2015; Patterson and Bickel, 2016; Cope et al., 2018; Chrysi and Mark, 2019; McCrudden and Marchand, 2020.).

To highlight GIS application in classroom, Cope et al. (2018) developed an ESRI Story Map called “Soil Forming Factors: Topography” for teaching an introductory soil science course at Clemson University in South Carolina. The objective of this story map was to introduce students to topography as one of the five soil-forming factors. Elevation, slope, aspect, and landscape position were interrelated elements of topography as a soil-forming factor that were traditionally taught using topographic maps or instruments in the field. In this innovative usage of ESRI Story Map, the research team introduced students to these elements via interactive web maps. The team assessed the performance of the newly developed Story map and students’ learning from Story map based on their responses to a quiz. Specific learning objectives were stated on each story map and throughout the learning materials. Highlighting the benefit of mixed methods, students used already-collected qualitative and quantities data in their mapping. The research team found out that overall, students were receptive about the ESRI Story map as a stand-alone teaching tool or in combination with PowerPoint slides. They suggested that their new Story Map could be further improved by incorporating interactive exercises.

In another example of using web-based GIS in classroom, fifth graders at the Zilker Elementary school in Austin, Texas engaged in a hands-on experiential online mapping project using free version of Social Explorer platform to explore the neighborhoods they live in. Students were able to switch among different data layers such as population density, racial composition, and the average age of buildings. Students were able to observe the formerly invisible associations, patterns, and tendencies in their local communities. In addition, GIS enabled students to construct reports, tables, and graphs of the given data. The educators concluded that the interactive GIS assignment was useful in creating awareness and sharing knowledge regarding the environment, communities, and social resources. Both these examples confirmed GIS can be used as an effective teaching tool in the physical and social sciences education. However, in both cases, students and teachers suggested more data to be added to their maps, particularly qualitative data. Additional qualitative and perhaps intangible data layers were important for students to understand the concepts they were learning in a more holistic, engaging, and realistic manner. This integration of qualitative data in GIS is what advocates of mixed methods research call QGIS.

With my background in academics and a major in non-education, I readily advocate for the applications of GIS in the current mixed methods approaches in educational research. Spatial visualization and analytic techniques are powerful and empowering in physical and human sciences. GIS methods provide further insights into several aspects of education and offer new opportunities for teachers to apply learner-centered methods in an engaging manner. The applications of GIS in and outside the classroom is summarized as follows:

GIS as a teaching tool in the classroom

GIS—particularly its more-user-friendly Environmental Systems Research Institute (ESRI) Story Map tool—has been considered as a research tool in the classroom for students and teachers studying diverse concepts in a variety of disciplines, not only in geography, but in history, mathematics, language arts, environmental studies, chemistry, biology, civics, and many more. According to ESRI (see <https://www.esri.com/arcgis-blog/products/uncategorized/uncategorized/story-maps-101/>), “Story map uses geography as a means of organizing and presenting information. They tell the story of a place, event, issue, trend, or pattern in a geographic context. They combine interactive maps with other rich content—text, photos, illustrations, video, and audio—within intuitive user experiences.” In other words, Story maps are web applications that allow students and teachers to combine web maps with narrative text, images, and multimedia, including video and audio. These applications are designed to be attractive and useable by anyone, which makes them great for education and outreach to the general public or to a particular audience.

This application happens at all levels including primary and secondary grades and continues to college and university courses. The GIS maps or Story Maps created by students offer both problem-based learning (PBL) and inquiry-based learning (IBL). “GIS as one of the 21st Century communication tools will assist in nurturing students’ abilities in investigating, evaluating, integrating, creating and analyzing issues and information at various scales and locations.” (Kibetu and Jackson, 2019: p. 202). Using mixed methods such as Story Map enables students to learn environmental analysis and problem-solving skills in a multi-layered and interactive visual way. This helps students think critically about the spatial relations among different layers, use real data, and connect them to their own communities. Because of its multi-layered, interactive spatial nature, the GIS application by students in the classroom can increase students’ curiosity, enhance learning capacity, and advance map literacy. In addition, GIS application enables teachers to emphasize concepts in different and relatable ways and adapt learning experiences to meet the needs of all learners.

GIS as a research tool outside the classroom

GIS can be used as a research tool by decision-makers, school administrators, counselors, and educators outside the classroom. The city may use GIS to redefine or adjust school district boundaries with respect to changing urban demographics. Schools may use GIS to plan access to facilities and to design new facilities and resources (e.g., technology lab, room allocations, and science lab). In addition, GIS visualization and analysis may help administrators to understand the relationship among students’ performances, the school district’s social and demographic status, and students’ access (or the lack thereof) to resources. These spatial understandings may decrease the long-standing inequity and accessibility gaps among different neighborhoods of a city. School administrators and teachers may use GIS to plan recruiting students, arrange extracurricular and after-hours activities, and manage alumni networks. On the other hand, GIS can also help the students, teachers, and school staff to become more spatially aware of the school (e.g., safe rooms, exits, or escape routes) and prepare for an emergency. GIS offers administrators a way to visualize and manage everything from monitoring campus safety, mapping campus buildings, cable and other infrastructure, routing school buses, planning where and when to close schools and open new ones, and strategizing recruitment efforts. GIS provides educational policy-makers with tools to examine spatial patterns and relations among educational achievement, school performance, racial and class inequalities and where to target new programs (see Taylor, 2007; Miller, 2012; Wei et al., 2018; Smith et al., 2019).

Traditionally, educational research is driven by qualitative designs such as ethnography, phenomenology, grounded theory, narrative, and case study. However, GIS visualization tools can be used to overlay different layers of data, including school rankings, students’ educational performance, school districts’ demographic characteristics, the socioeconomic status of the neighborhoods, levels of parents’ education, assigned governmental funding. Applying GIS spatial analysis tools, one can also illustrate any potential spatial clustering, patterns, and relationship among the above-mentioned factors. Recent studies in mapping social inequalities have highlighted the strong relationship between schools’ locations and the relative advantage or disadvantage of the neighborhoods in which schools are located. Research on neighborhood context on educational outcomes can benefit from GIS applications in order to better understand educational inequality. In addition, qualitative data such as interviews and/or participatory mapping by students, teachers, and parents can add a nuance dimension to GIS maps.

Given the advantages of using GIS in educational settings, there is a rationale in employing GIS mixed methods approach in exploring socio-cultural and geopolitical issues of a region or a nation. In the case study that follows later in this chapter, I rely on one innovative application of QGIS in geographic research in a non-western context. I draw on the methodological component of my ongoing geographical research on women’s social perceptions of and experiences in Tehran’s public spaces, the process that is highly gendered and classed. Photographs, sketches, and narratives collected from women in Tehran, Iran, were linked to GIS data gathered from the Iranian Census Organization, National GIS Database, and Tehran Municipality. This case study contextualizes QGIS applications in perhaps an unfamiliar cultural and political setting and therefore facilitates examining its limitations. The intensive literature review and findings from the case study illustrate not only the opportunities QGIS introduces to researchers

but also the boundaries of the technology at its current level. Through this case study based on QGIS application, I aim to invite more educators and educational scholars to take advantage of powerful and empowering qualities of GIS.

An application of QGIS: Women's perceptions and experiences in Tehran's public spaces

This section is written from a feminist geographer's perspective in which mapping is seen as a subjective, geopolitical tool reinforcing global and local hierarchies. Therefore, while GIS can open new avenues of knowledge to researchers by aiding in analyzing the spatial and social contexts of research subjects, it still emphasizes the significance of using ethnographic methods. GIS provides educators, researchers, and policymakers an empowering decision-making tool that can be used in educational administration and policy, and teaching. However, the integration of GIS tools with qualitative data becomes even more essential in educational research as we study different human subjects (e.g., students and teachers), their identities, perceptions, relationships, access to resources (or the lack thereof), and finally the art and science of teaching and learning.

To emphasize the differences between social construction of public spaces in non-Western and Western contexts and to understand Muslim women's socio-spatial behaviors in public spaces, I studied women's social perceptions and experiences of public spaces in Tehran, Iran (see Bagheri, 2013, 2014a,b, 2020). One of the main reasons I chose to focus on female users was the societal norms and my own positionality. As a female researcher, I found it more feasible to approach and talk with women than men. In addition, having grown up in Tehran and being familiar with local symbolic politics, I approached women in public spaces as a researcher, collected and interpreted their narratives. For instance, I chose to wear a typical academic/student hijab (*Maghnae*) instead of a more relaxed everyday hijab (*Roosari*) during my field visits. My research assistant (a fellow graduate student at Tehran University) and I approached women by introducing ourselves as scholars and users of public spaces. We explained how our research findings could benefit women like our interviewees through designing and managing public spaces.

The goal of this fieldwork was to illustrate how modern public spaces (e.g., shopping malls that were often located in Tehran's northern neighborhoods) and traditional public spaces (e.g., bazar and squares that were often located in Tehran's southern neighborhoods) in Tehran were used by women. However, the women's narrative did not recognize this often-used binary categorization of modern and traditional public spaces by urban professionals. Modern and privatized public spaces in North American contexts have been criticized for their highly commercialized atmosphere, controlled behavior and design, and segregation of certain racial and class groups (see Soja, 1989; Sorkin, 1992; Crawford, 1992; Banerjee and Loukaitou-Sideris, 1992; Zukin, 2003). Critics suggested that the privatization and modernization of public spaces which are standard designs have reduced them to more exclusive places and negatively influenced the social life happening in them. I selected two public spaces in Tehran with contrasting locations, functions, and architectural designs to examine these criticisms in a non-Western context. While there was not a clear boundary between traditional and modern public spaces according to the women with whom I interacted, the selected public spaces' contrasting locations in Tehran (one in the very northern and one in the south-central part of the city), their socio-political history, as well as their architectural designs and functions, allowed me to represent and comparatively study a more modern and a more traditional public space.

The case study pursued a mixed methods approach with a rich ethnographic quality and took advantage of socio-spatial data visualization and analysis in GIS. A mixed methods approach was driven from the fieldwork and became useful because of the nature of research questions. A qualitative-quantitative dichotomy seemed inadequate in understanding the complex relationship between women's social perceptions and experiences of public spaces, and the design and location of such places. Even in my very first round of participant observation, I noticed that both qualitative and quantitative characteristics of public spaces influenced women's presence and preferences. I collected and analyzed several layers of qualitative and quantitative data (see Fig. 2).

I used diverse methods of data gathering that were based on the socio-spatial experiences of women's everyday lives. I followed Lofland and Lofland's (1994) suggestion and began fieldwork visits with short-term covert participant observation. This gradual, "bottom-up" process let the research interests and questions grow from the field itself, rather than from theory. First, I collected photographs (qualitative data) to illustrate visual details such as women's style of hijab and make-up, and created spatial behavior maps to track women's numbers, activities, and locations as well as characteristics that are readily observable, such as approximate age, sex, and whether the individual was alone or in a group. As another qualitative data, I also drew architectural sketches to capture



Fig. 2 An application of QGIS in mapping women in Tehran's public spaces. Maps created by the author.

a sense of place. In addition, I conducted in-depth semi-structured interviews with 83 women who were using two seemingly contrasting, modern and traditional, selected public spaces in Tehran.

The women talked about their feelings, experiences, and preferences in those public spaces. I examined different ethno-cultural and symbolic meanings women attached to these public places. In contrast to the predominant critique of modern public spaces in the literature, I observed that Iranian women enjoyed a unique sense of freedom and equity with men in Tehran's modern and semi-private public spaces. Women narrated how they felt more comfortable doing certain activities such as holding hands with their boyfriends, smoking, talking louder, or laughing in modern public spaces when compared to more traditional public spaces. Women's narratives and hijab details gathered during participant observation also suggested that women had more freedom in their hijab and dressing styles in more modern public spaces; many interviewees emphasized that they would put on more make-up and fashionable, colorful clothes in such places. However, women's responses indicated favorable opinion of both public spaces, although for different reasons. Women's narrative and experiences suggested that they enjoyed a sense of freedom in the more modern public space while appreciating a sense of nostalgia and connection to their Iranian-Islamic identity in the more traditional one, such as Marvi Meidancheh (see Fig. 3). Fig. 3 illustrates an integration of qualitative data and quantitative data in a mixed methods approach, that is, an overlay of women's selected narratives, photographs, sketches, and dot density (produced in GIS) of women vs. men users in Marvi public space.

Advantages and limitations of using QGIS as a mixed methods research approach

While QGIS offers a valuable approach to many studies that deal with the concept of space and socio-cultural production of space, it is neither easy nor complete. Similar to any other MMR method, it is important to be aware of the potential limitations of QGIS, particularly due to the rapid pace of GIS software capabilities and accessibilities. In what follows, I offer a critical examination of the limits of QGIS in the case study while briefly reviewing its merits. Benefits include data management, discovery of spatiality of social processes, and an emphasis on the importance of ethnographic data in cultural geographic research. Limitations consist of potential data exclusion and data scale incompatibility.

In the above-explained research project, I used GIS to integrate qualitative and quantitative data gathered in the field. I first digitized the spatial behavior maps and transferred the population counts to GIS maps. Using Overlay, arguably the most important GIS

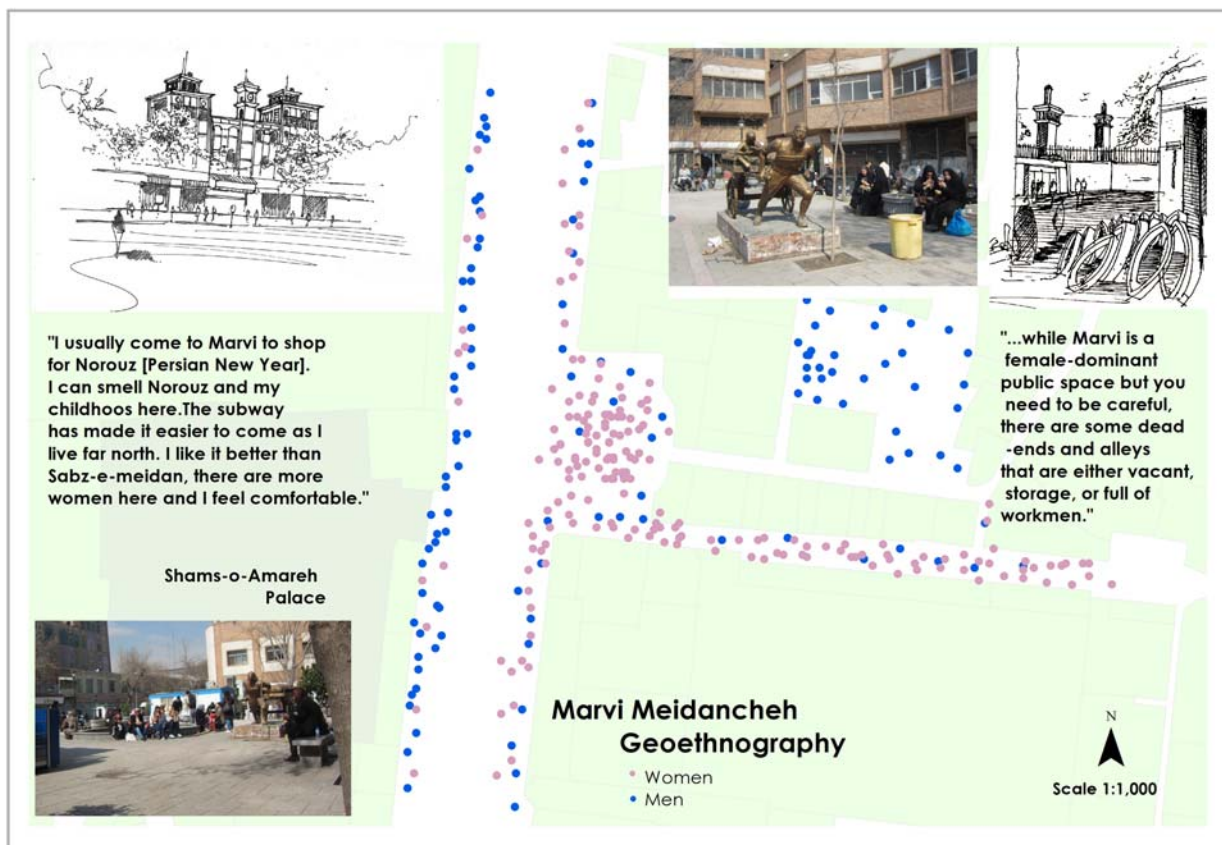


Fig. 3 A Mixed Methods Approach: Combining qualitative and quantitative data in GIS. Map created by the author.

analysis function, I examined the relationship between different demographic and socio-economic data layers gathered from the Iranian Census Organization, National GIS Database, Tehran Municipality, and population counts by gender on maps. I used the Kernel Density Estimation tool in GIS to illustrate female and male density and spatial distribution on maps. The maps showed interesting spatial patterns of female and male users in selected public spaces; however, they did not explain why such patterns existed. For instance, both the geo-visualization and statistical spatial analysis done within GIS software suggested that women in the more modern public space were less segregated from men compared to those in the more traditional one. In the absence of any regulations regarding male and female users' segregation in any of those public spaces, why did GIS maps suggest such spatial segregation? The GIS results also suggested the density of female users were relatively similar in both public spaces, particularly on weekends when the places witnessed their highest percentage of female users. How did the differences in design, location, and more importantly, attached symbolic meanings influence women's social perceptions and experiences of public spaces between these two public spaces? The supplementary women's narratives and field notes were useful in explaining these findings. Through talking with women, I examined their ways and values, and it became clear that "culture ... is at the heart of everything" (Mitchell, 2004, p. 155). Women attached different symbolic and ethno-cultural meanings to the selected public spaces that influenced their presence and behaviors in places.

Meanwhile, I created a geodatabase to record the location of each interview and to store the interviewee's characteristics (such as age, education, and home location). Using data management tools such as Query and Attribute Table in ArcGIS allowed me to be more organized in the gathering phase. For instance, I digitized the location of each interview to a Point Feature and used the hyperlink tool to connect the point to qualitative data I gathered in the field (e.g., key points of interviews and field notes). Watts (2010) found it difficult to "map" sounds of riots in Los Angeles, however, for me it seemed feasible to map narratives as I was concerned more with the women's narratives, rather than the background sounds in public spaces. However, I encountered different challenges in mapping my narratives. For example, women named different places, similar or contrasting to the one they were in, to further explain their feelings and preferences. This raised several questions: Was mapping narratives to the specific interview location enough? Would it be feasible to include the names of other places referenced in the interviews on GIS maps without actually being there or studying them? Would that be a useful or even readable map? What kinds of data gathered during field observation and interviewing could be transformed into GIS? Basically, the question was to what extent the rich and complex qualitative data, such as interview transcriptions, were transferrable into quantitative GIS codes and symbols. Examining possible answers to these types of questions, Jung and Elwood (2010), recognized what can be deemed an inevitable disadvantage of QGIS, potential data exclusion, and argued that "a host of other potentially significant details are lost" (p. 68) in the transformation process.

A similar challenge appeared when we considered the meaning of visual and spatial representation, one of the most exciting prospects of QGIS reported in earlier celebratory works on bridging the qualitative and quantitative domains. To this end, hyper-linking was enough for storing data but not for visually representing it. To represent ethnographic data, I included summarized and chosen data from each interview on maps. That means I had to choose which data was more important to be included on maps. This emphasized that GIS is not an independent tool; rather, it becomes part of the storytelling itself. In other words, in the process of "embedding" the individuals' lived experiences, the GIS user's agency (his or her personal characteristics and research goals) and the specific spatial analysis tools used affected the reality being represented and consequently the research results. While GIS can incorporate diverse forms of data, this case study illustrated the two worlds of complex reality and GIS representation cannot be seamlessly overlapped. Here, I questioned Krygier and Wood's claim (2005) that "you can map just about any data you can collect from the environment" (p. 51). I suggest that you may map almost all types of data, but not necessarily in an inclusive, efficient, and/or meaningful way. There remain some realms of human experiences and stories that are not "mappable," in a meaningful manner, at least not with current GIS technology. For example, women showed specific facial expressions and other body languages when they talked about their feelings in public spaces. How women felt while we were talking, why they attached certain symbolic meanings to those public spaces, or how such meanings changed their presence and behaviors are just few examples of qualitative data that cannot be mapped in GIS. This highlighted the importance of ethnographic data in cultural geographic research.

The second limitation I experienced with QGIS was data scale incompatibility. It is argued that QGIS allowed researchers to link pre-existing datasets (e.g., from censuses) to one's specific dataset (Bell and Reed, 2004; Jung and Elwood, 2010). However, this linkage can be meaningful or meaningless considering the difference among dataset scales. In my case study, the data gathered during participant observation and interviews were space and time sensitive and were based on individuals' opinions while the background GIS data (e.g., Census) were prepared based on official urban planning, social, and/or demographic data that were gathered and normalized at either *mahaleh* (neighborhood), *mantagheh* (districts), or the city of Tehran. I found it difficult and sometimes impossible to analytically link qualitative data to quantitative data layers in GIS considering their different scales. For example, while there could be a relationship between the interviewee's hijab style and her age, education, and median household income; there was not any meaningful relationship between her hijab with the similar data aggregated at the level of the district in which I interviewed her, or even at the level of the district in which she lived. Multiple GIS representations of different data can lead to wrong conclusions, in this case, to a wrong correlation between women's hijab and the quantitative attributes of the district where the interviews were conducted (particularly, for non-local scholars that may not be familiar with the symbolic politics of hijab in Iran). This data scale incompatibility can cause a major challenge in the data analysis phase because GIS application can both facilitate and hinder efforts to represent complex social and spatial realities. While Knigge and Cope (2009) found it useful in their case study, linking GIS and ethnographic data gathered at different scales was not always helpful or even meaningful in my case study. Therefore, I suggest caution in overlapping often quantitative census, urban planning, and land use data that are gathered by the

census organizations, city and district governments, or other state officials with qualitative data collected by a researcher for scholarly purposes. Depending on one's research questions, this capability of QGIS can become an advantage or disadvantage.

I am by no means suggesting GIS and particularly Qualitative GIS applications are not useful in understanding complex social and spatial phenomena, patterns, or processes. On the contrary, I acknowledge the advantages QGIS can offer to researchers, particularly in public participatory planning projects by giving voice to under-represented groups. In my case study, the geo-visualization capability of QGIS clarified that the specific geographical locations of selected public spaces (one in northern and one in south-central Tehran) played important roles in the symbolic meanings and ethno-cultural values women attached to them. In addition, users' spatial patterns on GIS maps also pointed out new aspects of public spaces and led me to modify my interview questions for the next round of field visits. For example, I did not question women about how they felt about their movements in a place, with regard to male users, in the first round of my interviews. But after I observed a spatial segregation in the more traditional public space on GIS maps, I added new relevant questions to my interviews. Indeed, QGIS enhanced and in some cases expedited my data interpretation and analysis, and allowed me to see more, but not necessarily "all." In the next section, I aim to contextualize the application of QGIS in mixed methods research in education, both inside and outside the classroom.

Conclusion

In this chapter, I introduced QGIS as a potential mixed method approach in social sciences. By combining qualitative data often gathered in detailed, contextualized, ethnographic manners and quantitative data often collected systematic, standard, and generalizable ways, one can potentially bring the subjective to the objective in understanding what (and why) is happening and potentially see more. QGIS as an innovative mixed methods approach emphasized the geographical and spatial dimension of phenomenon and processes. Hence, QGIS can offer both educational researchers and educators a powerful and empowering tool. Through an example of QGIS in my examination of women's social perceptions and experiences of public spaces in Tehran, Iran. In this research, I explained how I combined quantitative data such as socioeconomic, political, and religious status of neighborhoods, the geographic location of selected public spaces, the number of women using selected public spaces, with qualitative data including the photographs of women's hijab, architectural sketches of selected public spaces, women's narratives collected during open-ended interviews and my notes during participant observations in the field. These data were transformed to GIS platform, overlaid and analyzed. While the neighborhoods' characteristics and geographic location of the selected public spaces affected women's social perceptions and experiences, the qualitative data (e.g., women's narratives) explained why such spatial patterns were seen on GIS maps. Iranian women were using both traditional and modern public spaces for different reasons; GIS maps accompanied by ethnographic data emphasized the spatiality of social behaviors, in this case, women's social perceptions and experiences in selected Tehran's public spaces. In addition to advantages of using QGIS, I pointed out two potential disadvantages of combining data in GIS, that is, the scale incapability and potential data exclusion. There are ongoing progress to address these disadvantages in QGIS, but we have not yet fully overcome of the GIS limitations in visualizing qualitative data.

I encourage scholars from various disciplines interested in spatial visualization and analysis to consider the limitations of these methods while enjoying its empowering advantages.

"Digital technology [including GIS] can ... reduce critical reflection on the enhanced slippage occurring between map, experience, and context," (Jackson, 2008, p. 346) however, I believe that scholars' recent endeavors in thinking out of the box and acting innovatively to make GIS more accessible to diverse participants are admirable. Similar to the reflections King et al. (1994) made on combining qualitative and quantitative methods, my QGIS application has highlighted the challenges of either technologically impossible or epistemologically meaningless "quantifying" of the qualitative data.

This chapter contends GIS as "a powerful mediator of spatial knowledge and a powerful social actor," (Pickles, 2006, p. 763) expanding our epistemological and methodological horizons and elevating our geographical imaginations. QGIS as an innovative, inclusive mixed methods approach invites us to reimagine the complex, fluid, and diverse nature of human sciences. The application of QGIS encourages scholars to reengage with the political dimensions of knowledge production (e.g., knowledge by whom and for whom). Wood (1992) argued GIS is often used to promote distinctive interests or convince others; however, he emphasizes the interest can be anyone's, including researchers' and researched groups'. In sum, mixed methods approaches including QGIS can give voice to those under-represented groups that are often the subjects of studies conducted by feminist/critical geographers. Without any doubt, QGIS, compared to traditional applications of GIS, promotes more and crucial engagement with the people who are being studied than using solely GIS or solely ethnography.

Acknowledgments

I started writing this paper during the COVID-19 pandemic; meanwhile, I lost my dear friend and graduate student, Mr. Eliberto Vela to this virus. Several people have cared for and supported me during the last couples of years, I am so grateful for their unconditional love and acceptance. Also, I cannot thank the Elsevier staff members of the International Encyclopedia of Education enough for believing in me to finish the project. Dr. Nataliya Ivankova and Ms. Paula Davies were patient, understanding, persistent and encouraging. This contribution would not be possible without Dr. Ivankova's constructive comments and dedication to the subject.

References

- Aitken, S., Kwan, M.-P., 2009. GIS as qualitative research: knowledge, participatory politics, and cartographies of affect. In: DeLyser, D., et al. (Eds.), *The Handbook for Qualitative Methods in Geography*. Sage, London, pp. 287–303.
- AL-Komany, Kh., 2002. GIS and the Artist: Shaping the Image of a Neighborhood Through Participatory Environmental Design. In: *Community Participation and Geographical Information Systems*. CRC Press, pp. 346–355.
- Artvinli, E., 2010. The contribution of geographic information systems (GIS) to geography education and secondary school students' attitudes related to GIS. *Educ. Sci. Theor. Pract.* 10 (3), 1277–1292.
- Bagheri, N., 2013. *Modernizing the Public Space: Gender Identities, Multiple Modernities, and Space Politics in Tehran*. PhD Dissertation. University of Missouri.
- Bagheri, N., 2014a. Mapping women in Tehran's public spaces: a geovisualization perspective. *Gen. Place Cult.* 21 (10), 1285–1301. <https://doi.org/10.1080/0966369X.2013.817972>.
- Bagheri, N., 2014b. What qualitative GIS maps tell and don't tell: insights from mapping women in Tehran's public spaces. *J. Cult. Geogr.* 31 (2), 166–178. <https://doi.org/10.1080/08873631.2014.906848>.
- Bagheri, N., 2020. QGIS in feminist geographic research: its merits and limits. In: Datta, A., Hopkins, P., Johnston, L., Olson, E., Silva, J.M. (Eds.), *Gender and feminist geographies*. Routledge Press, pp. 455–466.
- Banerjee, T., Loukaitou-Sideris, A., 1992. *Private Production of Downtown Public Open Space: Experiences of Los Angeles and San Francisco*. University of Southern California, Los Angeles, CA.
- Bell, S., Reed, M., 2004. Adapting to the machine: integrating GIS to qualitative research. *Cartographica* 39 (1), 55–66. <https://doi.org/10.3138/Y413-1G62-6HG-OL3Q>.
- Bodzin, A.M., Anastasio, D., 2006. Using web-based GIS for earth and environmental systems education. *J. Geosci. Educ.* 54 (3), 297–300.
- Boschmann, E., Cubbon, E., 2014. Sketch maps and qualitative GIS: using cartographies of individual spatial narratives in geographic research. *Prof. Geogr.* 66 (2), 236–248. <https://doi.org/10.1080/00330124.2013.781490>.
- Chrysi, R., Mark, F., 2019. Mixed methods research in inquiry based instruction: an integrative review. *Int. J. Res. Method Educ.* 42 (3), 288–304. <https://doi.org/10.1080/1743727X.2019.1598356>.
- Cieri, M., 2003. Between being and looking: queer tourism promotion and lesbian social space in greater Philadelphia. *ACME* 2 (2), 147–166.
- Cope, M., Elwood, S. (Eds.), 2009. *Qualitative GIS: A Mixed Methods Approach*. Sage, Los Angeles, CA.
- Cope, M., E.A., Post, C.J., Schlautman, M.A., Carbajales-Dale, P., 2018. Developing and evaluating an ESRI story map as an educational tool. *J. Nat. Resour. Life Sci. Educ.* 47 (1), 1–9. <https://doi.org/10.4195/nse2018.04.0008>.
- Crawford, M., 1992. The world in a shopping mall. In: Sorkin, M. (Ed.), *Variations on a Theme Park: The New American City and the End of Public Space*. Hills and Wang, New York, NY, pp. 3–30.
- Drennon, C., 2005. Teaching geographic information systems in a problem-based learning environment. *J. Geogr. High. Educ.* 29 (3), 385–402.
- Elwood, S., 2002. GIS use in community planning : a multidimensional analysis of empowerment. *Environ. Plann.* 34 (5), 905–922.
- Fielding, N., 2012. Triangulation and mixed methods designs: data integration with new research technologies. *J. Mix. Methods Res.* 6 (2), 124–136.
- Fielding, N., Cisneros-Puebla, C.A., 2009. CAQDAS-GIS convergence: toward a new integrated mixed method research practice? *J. Mix. Methods Res.* 3 (4), 349–370.
- Haraway, D., 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge, New York, NY.
- Hong, J.E., 2014. Promoting teacher adoption of GIS using teacher-centered and teacher-friendly design. *J. Geogr.* 113 (4), 139–150. <https://doi.org/10.1080/00221341.2013.872171>.
- Jackson, S., 2008. The city from thirty thousand feet: embodiment, creativity, and the use of geographic information systems as urban planning tools. *Technol. Cult.* 49 (2), 325–346. <https://doi.org/10.1353/tech.0.0039>.
- Jones, P., Evans, J., 2012. The spatial transcript: analyzing mobilities through qualitative GIS. *Area* 44 (1), 92–99. <https://doi.org/10.1111/j.1475-4762.2011.01058.x>.
- Jung, J.K., Elwood, S., 2010. Extending the qualitative capabilities of GIS: computer-aided qualitative GIS. *Trans. GIS* 14 (1), 63–87. <https://doi.org/10.1111/j.1467-9671.2009.01182.x>.
- Kerski, J.J., 2015. Geo-awareness, geo-enablement, geotechnologies, citizen science, and storytelling: geography on the world stage. *Geogr. Compass* 9, 14–26. <https://doi.org/10.1111/gec3.12193>.
- Kibetu, K., Jackson, C., 2019. Emerging technologies in teaching, research, and learning: GIS technology and methods in education. In: *Technology-Supported Teaching and Research Methods for Educators*, pp. 201–217.
- King, A., Keohane, R., Verba, S., 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ.
- Knigge, L., Cope, M., 2006. Grounded visualization: integrating the analysis of qualitative and quantitative data through grounded theory and visualization. *Environ. Plan. A* 38 (11), 2021–2037. <https://doi.org/10.1068/a37327>.
- Knigge, L., Cope, M., 2009. Grounded visualization and scale: a recursive analysis of community spaces. In: Cope, M., Elwood, S. (Eds.), *Qualitative GIS*. Sage, Los Angeles, CA, pp. 95–104.
- Krygier, J., Wood, D., 2005. *Making Maps: A Visual Guide to Map Design for GIS*. Guilford Press, New York, NY.
- Kwan, M.P., 2002. Feminist visualization: re-envisioning GIS as a method in feminist geographic research. *Ann. Assoc. Am. Geogr.* 94 (2), 645–661.
- Kwan, M.P., 2004. Beyond difference: from canonical geography to hybrid geographies. *Ann. Assoc. Am. Geogr.* 94 (4), 756–763. <https://doi.org/10.1111/j.1467-8306.2004.00432.x>.
- Kwan, M.P., 2008. From oral histories to visual narratives: re-presenting the post- September 11 experiences of the Muslim women in the United States. *Soc. Cult. Geogr.* 9 (6), 653–669. <https://doi.org/10.1080/14649360802292462>.
- Kwan, M.P., Ding, G., 2008. Geo-narrative: extending geographic information systems for narrative analysis in qualitative and mixed-method research. *Prof. Geogr.* 60 (4), 443–465. <https://doi.org/10.1080/00330120802211752>.
- Lofland, J., Lofland, L., 1994. *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Wadsworth, Belmont, CA.
- Matthews, S., Detwiler, J.E., Burton, L.M., 2005. Geo-ethnography: coupling geographic information analysis techniques with ethnographic methods in urban research. *Cartographica* 40 (4), 75–90. <https://doi.org/10.3138/2288-1450-W061-R664>.
- Maxwell, J.A., 2010. Using numbers in qualitative research. *Qual. Inq.* 16, 475–482.
- McCloughlin, T.J.J., Matthews, P.S.C., 2012. Repertory grid analysis and concept mapping: problems and issues. *Probl. Educ. 21st Century* 48 (1), 91–106. <https://doi.org/10.33225/pec/12.48.91>.
- McCrudden, M., Marchand, G., 2020. Multilevel mixed methods research and educational psychology. *Educ. Psychol.* 55 (4), 197–207. <https://doi.org/10.1080/00461520.2020.179315>.
- McLafferty, S.L., 2002. Mapping women's worlds: knowledge, power and the bounds of GIS. *Gen. Place Cult.* 9 (3), 263–269. <https://doi.org/10.1080/096636902200003879>.
- Miller, P.M., 2012. Mapping educational opportunity zones: a geospatial analysis of neighborhood block groups. *Urban Rev.* 44 (2), 189–218.
- Mitchell, D., 2004. Cultural geography at the millennium: a call to (intellectual) arms. *J. Cult. Geogr.* 22 (1), 155–157. <https://doi.org/10.1080/08873630409478254>.
- Moss, P. (Ed.), 2002. *Feminist Geography in Practice: Research and Methods*. Blackwell, Oxford, U.K.
- National Research Council, 2005. *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum*. National Academies Press, Washington, DC. http://dels.nas.edu/resources/staticassets/materials-based-on-reports/reports-in-brief/learning_to_think_spatially_final3.pdf. (Accessed 20 October 2017).

- Nightingale, A., 2003. A feminist in the forest: situated knowledges and mixing methods in natural resource management. *ACME* 2 (1), 77–90.
- Patterson, J., Bickel, A., 2016. Communicating Local Relevance of Ocean Observations: Integrating Real-time Ocean Sensor Data Visualizations, Online Communications, and Ocean Issues to Engage Public Audiences. Monterey.
- Pavlovskaya, M.E., 2006. Theorizing with GIS: a tool for critical geographies? *Environ. Plan. A* 38, 2003–2020.
- Pickles, J., 1995. *Ground Truth: The Social Implications of Geographical Information Systems*. Guilford, New York.
- Pickles, J., 2006. Commentary: ground truth 1995–2005. *Trans. GIS* 10, 763–772.
- Pratt, G., 1989. Quantitative techniques and humanistic-historical materialist perspectives. In: Audrey, K., Suzanne, M. (Eds.), *Remaking Human Geography*, 101–15. Unwin Hyman, Boston.
- Schuurman, N., 2000. Trouble in the heartland: GIS and its critics in the 1990s. *Prog. Hum. Geogr.* 24, 569–590.
- Sheppard, E., 2001. Quantitative geography: representations, practices, and possibilities. *Environ. Plann. Soc. Space* 19 (5), 535–554. <https://doi.org/10.1068/d307>.
- Smith, Parr, N., Muhidin, S., 2019. Mapping schools' NAPLAN results: a spatial inequality of school outcomes in Australia. *Geogr. Res.* 57 (2), 133–150. <https://doi.org/10.1111/1745-5871.12317>.
- Soja, E., 1989. *Postmodern Geographies: The Reassertion of Space in Critical Social Theory*. Verso Books, New York, NY.
- Sorkin, M., 1992. *Variations on a Theme Park: The New American City and the End of Public Space*. Hills and Wang, New York, NY.
- Taylor, C., 2007. Geographical information systems (GIS) and school choice: the use of spatial research tools in studying educational policy. In: Gulson, K.N., Symes, C. (Eds.), *Spatial Theories of Education: Policy and Geography Matters*. Routledge, London and New York, pp. 7–104.
- Watts, P., 2010. Mapping narratives: the 1992 Los Angeles riots as a case study for narrative-based geo-visualization. *J. Cult. Geogr.* 27 (2), 203–227.
- Wei, Y.D., Xiaoa, W., Simon, B., Liu, B., Ni, Y., 2018. Neighborhood, race and educational inequality. *Cities* 73, 1–13.
- Wood, D., 1992. *The Power of Maps*. The Guilford Press, New York, NY.
- Zukin, S., 2003. *Point of Purchase: How Shopping Changed American Culture*. Routledge, New York, NY.

Mixed methods research synthesis for literature reviews

Patrick Onghena, KU Leuven, Faculty of Psychology and Educational Sciences, Leuven, Belgium

© 2023 Elsevier Ltd. All rights reserved.

Introduction	655
The need for literature reviews	655
The need for systematic literature reviews	656
The need for mixed methods to conduct a systematic literature review	656
A pitfall for mixed methods research synthesis: ontology and epistemology	657
A step-by-step guide for mixed methods research synthesis	658
Writing a protocol and selecting a design	658
Selecting a sampling strategy, a search strategy, and eligibility criteria	658
Critically appraising the primary-level studies	659
Extracting the data	659
Analyzing and synthesizing the data	659
Writing the report	659
Mixed methods research synthesis in educational research	660
Mixed methods research synthesis in educational research: an example	661
Writing a protocol and selecting a design	661
Selecting a sampling strategy, a search strategy, and eligibility criteria	661
Critically appraising the primary-level studies	661
Extracting the data	662
Analyzing and synthesizing the data	662
Writing the report	662
Conclusion	662
Appendix 1: Mixed methods research synthesis in educational research	663
References	664

Introduction

A mixed methods research synthesis is defined as a systematic literature review that provides a synthesis of existing knowledge and research evidence on a certain phenomenon by applying the principles of mixed methods research at the meta-level (Heyvaert et al., 2017; Heyvaert et al., 2011, 2013b). These principles are based on the premise that many complex real-world phenomena have both qualitative and quantitative features and that an integration of methods for synthesizing qualitative research and for summarizing quantitative research has an added value as compared to a mono-method approach (Harden and Thomas, 2005; Pluye and Hong, 2014; Sandelowski et al., 2006). Because education as a field of scientific inquiry is concerned with complex real-world phenomena and because a large diversity of scientific disciplines is involved –going from philosophy of education to educational neuroscience– it is expected that mixed methods research synthesis might be well suited as an approach to systematically review educational research.

This article discusses the need for literature reviews, the need to perform these literature reviews in a systematical way, and the need to include both quantitative and qualitative material. In addition, ontological and epistemological issues are presented as a potential pitfall for a mixed methods research synthesis. The article concludes with an overview of the different steps and decisions that have to be taken in a mixed methods research synthesis, the application of mixed methods research synthesis in education, and a worked-out example from educational research.

The need for literature reviews

The increasing number of scientific articles and books makes literature reviews indispensable (Snyder, 2019). When policy makers, practitioners, and researchers want to learn about the state of the art regarding a particular phenomenon or a specific research problem, it is way more efficient to read a topical and authoritative literature review than to be swamped by all separate studies published on the phenomenon or problem. For policy and practice, time is a solicitude, and looking into the scientific literature should be functional for an evidence-based recommendation, a short-term decision, or a prompt (plan for) intervention or action. Literature reviews are ideally suited to fit that purpose. For researchers, reading a literature review provides a first orientation regarding a research problem and can help to avoid reinventing the wheel. Of course, researchers are not only consumers of literature reviews, but also active producers. Conducting a new literature reviews is often described as a first step in the empirical cycle or as a research method on its own for building a comprehensive conceptual model and theory (Hart, 2018).

Although scientific literature reviews can already be found in the 18th century, it is only in the 20th century that scientists started developing methods to improve the reliability and the validity of those reviews (Chalmers et al., 2002). This development was triggered by acknowledging that the sheer number of articles and books entails the risk that a review author, possibly guided by his own predilection or by his eagerness to write a coherent narrative, is highly selective in his reading and digesting of the literature. By being selective, a review author might be prone to draw imprecise or one-sided conclusions, and might instill a biased perspective in the readers of the literature review (Clarke and Chalmers, 2018). Consequently, the need for literature reviews resulted in a need for *systematic* literature reviews.

The need for systematic literature reviews

Lasserson et al. (2021) define a systematic (literature) review as an attempt “to collate all the empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question” (section-1-1 para. 2). The methods for such a literature review are described as “highly structured” (hence systematic), “transparent”, and “reproducible”, and “are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made” (section-1-1 para. 2). The reference to “more reliable” in this quotation implies a comparand and is left implicit, but it is clear that a traditional narrative or unsystematic review without pre-specified eligibility criteria is targeted.

Meta-analysis, a term coined by the educational researcher Gene Glass in the 1970s, is a well-established part of such a systematic review and is concerned with summarizing the quantitative information that is available in the articles and books that deal with the systematic review’s research question (Booth et al., 2016; Gough et al., 2017). The “meta” in the term refers to the fact that the analysis usually involves a statistical analysis of the statistical analyses that have already been conducted by the authors of the primary-level studies.¹ In that way, “meta-analysis” is a statistical analysis at the meta-level (Borenstein et al., 2009).

However, not all parts of answers to research questions or solutions to problems are quantitative or quantifiable. Some parts might be intrinsically qualitative, for example evidence and arguments with respect to feasibility, appropriateness, and meaning (Pearson et al., 2005). This poses a challenge for the systematic reviewers: on the one hand they have the highly formalized meta-analytical toolbox at their disposal and on the other hand they might have collected tons of qualitative material about the theoretical background, the history, ethics, politics, and practical implications of the phenomenon of interest. What to do? Ignore the qualitative material? Perform a traditional narrative review of the qualitative material and glue it to the meta-analysis? Fortunately, in recent years, we witnessed the development of several qualitative meta-synthesis methods for summarizing qualitative primary-level studies and for generating new insights and understanding from interrelated qualitative research findings, such as formal grounded theory (Eaves, 2001; Kearney, 1998, 2001), meta-ethnography (Britten et al., 2002; Noblit and Hare, 1988), thematic synthesis (Thomas and Harden, 2008), and meta-aggregative synthesis (Hannes and Lockwood, 2011). Such qualitative meta-synthesis methods can be combined with quantitative meta-analysis methods to arrive at more comprehensive systematic reviews.

The need for mixed methods to conduct a systematic literature review

This combination of qualitative meta-synthesis methods and quantitative meta-analysis methods, however, poses a next challenge for the systematic reviewers: how to combine these qualitative and quantitative methods. Is a juxtaposition sufficient? Or should they aim for an integration? The answer to these questions can be found by going back to the original phenomenon and the research problem of interest, but above all, these questions suggest that insights from mixed methods research can also guide us with respect to performing comprehensive systematic reviews. To wit, the juxtaposition-integration question is at the core of mixed methods research methodology (Bryman, 2007; Creswell et al., 2003; Plano Clark and Ivankova, 2016). This brings us to the application of mixed methods to conduct a systematic literature review: a so-called “mixed methods research synthesis” (Heyvaert et al., 2017; Pluye and Hong, 2014; Sandelowski et al., 2012).

The need for mixed methods research synthesis will be felt most if the phenomenon or the research problem of interest is complex, consists of multiple components and layers, and if quantitative and qualitative aspects can be identified *a priori* (Ferguson et al., 2020; Hong et al., 2020; Noyes et al., 2019). Ideally, the use of mixed methods research synthesis is preplanned rather than incidental. However, it cannot be ruled out that after conducting a meta-analysis, the statistical results for certain moderators are difficult to interpret, there is concern about relevant qualitative empirical evidence that was not included in the meta-analysis, or there are still other qualitative issues that have to be resolved. In such circumstances, a mixed methods research synthesis comes as an after-thought, but is nonetheless worthwhile to consider. Furthermore, in our experience, meta-analysis is rarely a mere mathematical exercise that produces effect size estimates and confidence intervals. Usually there is uncertainty beyond the width of the

¹In order to make a distinction between a review study and the studies on which the review study is based, the latter are commonly called “primary-level studies”. The review study is called a “secondary-level study” or a “study at the meta-level”.

confidence intervals, there are contextual factors that have to be taken into account to qualify the interpretations, and there are qualitative considerations of generalizability that have to be discussed to put the results in perspective. Consequently, some notion of mixed methods research synthesis might be beneficial for every meta-analyst and every systematic reviewer.

In addition to this general benefit, there are four conditions that are particularly conducive for engaging in a mixed methods research synthesis (Heyvaert et al., 2017):

1. If the phenomenon or the research problem of interest does not only consist of multiple components and layers, but if the focus is also on how these components and layers are related, and how these components and layers interact,
2. If the mixed methods research synthesis is carried out in the context of evaluating an educational, behavioral, or medical intervention, and the focus is not only on the effectiveness of the intervention, but also on its feasibility, appropriateness, and meaningfulness,
3. If some of the primary-level studies are already mixed; in that case, those studies can be smoothly integrated with the quantitative and qualitative studies; otherwise, those studies have to be left aside or artificially dissected in a quantitative and a qualitative part, and
4. If qualitative syntheses and quantitative meta-analyses on a topic are already available; in that case the mixed methods research synthesis can act as the scene where the two strands of evidence are combined, possibly updated by new primary-level studies, either confrontationally or complementary.

Notwithstanding the benefits and the opportunities, several hazards for mixed methods research synthesis can also be identified (see e.g., Heyvaert et al., 2017). These hazards are in one way or another related to the major pitfall: ontology and epistemology.

A pitfall for mixed methods research synthesis: ontology and epistemology

Ontology refers to the philosophical study of “what is”; a theory of reality. Epistemology refers to the philosophical study of “what can be known”; a theory of knowledge (Hathcoat et al., 2018). In scientific research, ontology and epistemology play a key role in the research questions being asked, the kind of methods and techniques being applied, and the kind of conclusions and interpretations considered valid or worthwhile. Especially researchers in the qualitative tradition emphasize the relevance of this perspective on the world and on knowledge in the research process, and they recommend that researchers be explicit about their ontological and epistemological stance (Ravitch and Carl, 2019; Savin-Baden and Major, 2013).

Simplifying a complex philosophical debate, ontological positions toward research are often located on a continuum with on the one end the realist position and on the other end the idealist position. Researchers taking a realist position consider reality as singular, concrete, external, and existing independent of the researcher, whereas researchers taking an idealist position consider reality as a social construction, that is at least partially, a product of human perception and imagination, implying the existence of multiple realities (Morgan and Smircich, 1980).

Epistemological positions may be closely related to these ontological positions. On the one hand, the post-positivist position holds that empirical primary-level studies can be valid sources of knowledge about the external world and that it is possible to accumulate this knowledge by verification and falsification. On the other hand, the interpretative position holds that phenomenological insight about the multiple realities can be obtained by hermeneutic methods and critical reflection (Moon and Blackman, 2014).

The crux for mixed methods research is now that quantitative researchers lean toward the realist post-positivist position and that qualitative researchers lean toward the idealist interpretative position (Sale et al., 2002; Savin-Baden and Major, 2013). This makes the endeavor of a mixed methods research synthesis sometimes a hard nut to crack. Not only at the level of the primary studies incompatible world and incommensurable knowledge views might be at stake, but also at the meta-level some common ground has to be found.

The way out can be found by relying on one or more of the following four interrelated approaches:

1. The first approach is that the realist post-positivist position and the idealist interpretative position, as described above, are extremes on a continuum of positions (Howe, 1988; Newman and Benz, 1998). If researchers position themselves more in-between, then combination and integration become more reasonable and expedient. At the same time, this first approach implies that a mixed methods research synthesis has to remain sensitive to the positions in the primary-level studies, and to the position taken in the mixed methods research synthesis itself.
2. The second approach is that the positions do not have to be muddled if the separation and articulation of the positions is paramount. The mixed methods research synthesis can give dominance to a certain position (with other positions being auxiliary) or can work in sequences (Heyvaert et al., 2011, 2013b; Johnson and Onwuegbuzie, 2004).
3. The third approach comprises a philosophical third way out of the realist-idealist duality by embracing pragmatism (Biesta, 2010; Cersosimo, 2019; Feilzer, 2010). In pragmatism, as a philosophical tradition going back to the writings of Peirce (1878), James (1907), and Dewey (1925), knowledge about the world is defined by our actions. In this tradition, the meaning of statements and hypotheses have to be clarified by tracing their practical consequences – their implications for experience in specific situations (Legg and Hookway, 2020). In this way, pragmatism bypasses the contentious issues of the exact nature of reality and knowledge, and puts an emphasis on the consequences of our actions, possibly but not

necessarily based on our knowledge and world views (Feilzer, 2010). Pragmatism considers knowledge as a tool for prediction, action, and problem-solving rather than something that aims to merely describe, represent, or mirror one or multiple realities (Cersosimo, 2019).

4. The fourth approach refers to the research team that is conducting the mixed methods research synthesis (Curry and Nunez-Smith, 2015; Shulha and Wilson, 2003; O'Cathain et al., 2008). This team can be as diverse as the research problem requires, including statisticians, qualitative methodologists, and substantive experts. Such a research team can capitalize on the complementary background and skills of the team members, but perhaps more importantly, such a research team can prepare a thorough integration by engaging in a dialog about the purpose of the mixed methods research synthesis, the research question it proposes to answer, and the material that can be included. The in-between position of the first approach or the pragmatism in the third approach may offer a common language and make this dialog possible. On the one hand, this dialog assumes sensitivity and respect for the different positions and perspectives of the team members, and on the other hand, it can prevent the mixed methods research synthesis of degenerating in a clumped hodgepodge.

One or more of the previous four approaches can be used to position oneself as a systematic reviewer conducting a mixed methods research synthesis. A step-by-step guide for conducting the mixed methods research synthesis itself is presented and illustrated in the next sections.

A step-by-step guide for mixed methods research synthesis

The actual procedure for conducting a mixed methods research can be described in six discrete steps, which are similar to the step-wise approach of an empirical study: (1) writing a protocol and selecting a design, (2) selecting a sampling strategy, a search strategy, and eligibility criteria, (3) critically appraising the primary-level studies, (4) extracting the data, (5) analyzing and synthesizing the data, and (6) writing the report. In an ideal scenario, these steps are followed sequentially, but in practice there can be overlap or back and forth between steps. In this section, we focus on the essence of each step. For more details, the reader is referred to Heyvaert et al. (2017).

Writing a protocol and selecting a design

This protocol documents all methodological and substantive choices that are needed throughout the mixed methods research synthesis process. The protocol contains the selected mixed methods research synthesis design and a description of the choices made for the next steps. General guidelines for writing such a protocol can be found in Shamseer et al. (2015).

With respect to the designs, a basic distinction can be made between a segregated design, an integrated design, and a contingent design (Sandelowski et al., 2006). In a segregated mixed methods research synthesis design, the qualitative meta-synthesis and the quantitative meta-analysis are conducted separately and independently, and only afterward an attempt is made to synthesize both strands of evidence. In an integrated mixed methods research synthesis design, qualitative and quantitative studies are directly combined, for example by transforming the qualitative data into quantitative data and/or by transforming the quantitative data into qualitative data. In a contingent mixed methods research synthesis design, qualitative meta-synthesis methods and quantitative meta-analysis methods are used in cycles until saturation is reached and a comprehensive research synthesis can be presented that addresses the review objectives.

An example of a mixed methods research synthesis using a segregated design can be found in Heyvaert et al. (2014, 2015), who performed a systematic review of restraint interventions for challenging behavior among persons with intellectual disabilities. In the quantitative strand, they focused on the effectiveness of the interventions and in the qualitative strand, they focused on the experiences of the persons with intellectual disabilities, bringing the two strands together in the discussion section. Examples of mixed methods research syntheses using a contingent design can be found in Noyes et al. (2019), who presented three case studies of guidelines developed by the World Health Organization, which incorporated quantitative and qualitative evidence in sequential and iterative cycles. These guidelines only became established after the research team reached consensus based on the available quantitative and qualitative evidence. An example of an integrated mixed methods research synthesis design is presented in **Mixed methods research synthesis in educational research: an example** section.

Selecting a sampling strategy, a search strategy, and eligibility criteria

The sampling strategy may be exhaustive, selective, or purposeful, depending on the review objectives and the available resources (Booth, 2006). The search strategy stipulates which bibliographic databases, gray literature databases, journals, reference lists, and citation indexes are consulted, and which key words were used to perform the search. It may also involve contacting experts and authors for unpublished data. The eligibility criteria refer to the inclusion and exclusion criteria that are needed to filter out irrelevant studies and to retain only the primary-level studies that are relevant for answering the review questions.

Critically appraising the primary-level studies

Critical appraisal refers to the assessment of the methodological quality of the included primary-level studies. Because these studies can be quantitative, qualitative, or mixed, the critical appraisal is performed best for each of the evidence strands separately and should take into account that mixed methods studies are more than just the sum of their individual qualitative and quantitative strands of evidence, and thus require specific criteria (Heyvaert et al., 2013a). Furthermore, a decision is needed about a way to deal with the differences in methodological quality between the studies. Finally, because critical appraisal is heavily based on expert judgment, differences between the members of the review team should be identified and discussed.

Extracting the data

Data collection in a mixed methods research synthesis consists of extracting the relevant data from the included primary-level studies (Cooper, 2010). These data include (1) formal characteristics of each study, such as the authors, the publication date, and the publication source, (2) methodological characteristics of each study, such as a description of the participants, the design, and the instruments, and (3) substantial characteristics that are related to the research questions and the theoretical framework of the mixed methods research synthesis, such as the type of intervention that is studied, the main results, and the conclusions drawn in the primary-level studies.

It is recommended that a preliminary data extraction form and a coding guide are developed and that the extraction form and the coding guide are piloted for a couple of primary-level studies before extracting the data from the complete set. Also in this step, identifying and discussing differences in extraction between the members of the review team is mandatory.

Analyzing and synthesizing the data

Data analysis follows the logic of the selected design in the first step (Heyvaert et al., 2017). In a segregated mixed methods research synthesis, the quantitative studies (and the quantitative part of the mixed methods studies) are summarized using meta-analytical tools and the qualitative studies (and the qualitative part of the mixed methods studies) are synthesized using formal grounded theory, meta-ethnography, thematic synthesis, or meta-aggregative synthesis. Both strands of evidence are brought together by discussing how they complement each other and whether the potentially complementary findings can be arranged into a line of argument, into a theory that posits relationships among concepts, or into a narrative that posits a temporal ordering of events (Sandelowski et al., 2006).

In an integrated mixed methods research synthesis, techniques can be used for “qualitizing” the quantitative findings and for “quantitizing” the qualitative findings. For example, percentages addressed in questionnaires may be translated into textual data (e.g., “the majority of the participants”, “some participants”, “only a limited number of participants agreed on a particular topic”) to facilitate the creation of a narrative storyline or qualitative data, or the verbal counts in a qualitative study may be converted into numbers (e.g., the frequency of a particular finding, category, or theme). Moreover, three mixed synthesis approaches can be used for directly integrating qualitative, quantitative, and mixed primary-level studies: narrative summary (Borman and Grigg, 2009), critical interpretive synthesis (Dixon-Woods et al., 2006a; Dixon-Woods et al., 2006b), and realist synthesis (Pawson, 2002, 2006).

A contingent mixed methods synthesis may involve consecutive qualitative and quantitative synthesis approaches. However, it is also possible that within one, some, or all of the strands, a qualitizing, quantitizing, or mixed synthesis approach is used. An alternative approach is sequential Bayesian synthesis to draw on both qualitative and quantitative evidence. In such a sequential Bayesian synthesis, findings from the qualitative studies are used for eliciting prior probabilities for each of the research hypotheses included in the mixed methods research synthesis. Thereafter the findings from the quantitative studies on the same topic are used to update these prior probabilities, yielding posterior probabilities for each of the research hypotheses. If “quantitizing” and “qualitizing” techniques are added to this approach, then several sequences and cycles of updates are possible (see e.g., Crandell et al., 2011; Roberts et al., 2002; Voils et al., 2009).

Writing the report

Writing and disseminating the mixed methods research synthesis report is the final step in the process. If the synthesis report is intended for organizations promoting and supporting literature reviews, such as The Campbell Collaboration or The Cochrane Collaboration, then it is recommended to use their standard templates (Heyvaert et al., 2017). Additionally, templates have been developed for reporting on specific data synthesis approaches, such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009), ENhancing Transparency in REporting the synthesis of Qualitative research (ENTREQ; Tong et al., 2012), the RAMESES publication standards for realist syntheses (Wong et al., 2013), and the Joanna Briggs Institute’s standards for reporting on meta-aggregative syntheses (Lockwood et al., 2015).

Table 1 Number of mixed methods research synthesis articles indexed in the web of science for each publication year.

Year	07	08	09	10	11	12	13	14	15	16	17	18	19	20
Number	2	1	2	4	6	13	18	23	30	45	47	82	92	149

Note. Search was conducted on September 28, 2021. The search string was TS = ("mixed methods research synthesis") OR TS = ("mixed methods systematic review") OR TS = ("mixed method systematic review") OR TS = ("mixed methods synthesis") OR TS = ("mixed method synthesis") OR TS = ("mixed research synthesis") OR TS = ("mixed studies review") OR TS = ("mixed study review"). Topic Searches (TS) include title, abstract, author keywords, and Keywords Plus. Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.

Mixed methods research synthesis in educational research

In order to explore the use of mixed methods research synthesis –as defined and discussed in the previous sections– in international peer reviewed journal articles, and in articles devoted to educational research in particular, we conducted a Web of Science search for the terms “mixed methods research synthesis”, “mixed method research synthesis”, “mixed methods systematic review”, “mixed method systematic review”, “mixed methods synthesis”, “mixed method synthesis”, “mixed research synthesis”, “mixed studies review”, and “mixed study review”, without any publication date restriction. We found a total of 514 unique articles, with the earliest publication date being 2007. The number of articles in each publication year from 2007 until 2020 is shown in **Table 1**.

Two results in **Table 1** are striking. First, mixed methods research syntheses, identified as such or with similar terms, do not appear in the scientific literature before 2007. Given that the advocacy of [Harden and Thomas \(2005\)](#) and [Sandelowski et al. \(2006\)](#) stemmed from a few years before, this result is understandable. Although it does not necessarily imply that (precursors of) mixed methods research syntheses are absent from the pre-2007 literature, it gives some indication that pre-2007 research syntheses did not bother to refer to the issues that are involved in mixing methods. Second, we see a steep increase in the number of articles between 2007 and 2020 referring to mixed methods research synthesis, with an increase of more than 60% from 2019 to 2020. Obviously, the popularity of this approach is rising and a plateau is not in sight.

Table 2 shows the number of articles referring to mixed methods research synthesis for different science categories in the Web of Science. The most frequently indexed category is “Nursing” with 103 out of 514 records (20%). The acclaim of mixed methods research synthesis in nursing is not surprising because nursing involves many complex real-world problems that have both qualitative and quantitative features. Furthermore, some of the leading methodologists in the field (e.g., [Margarete Sandelowski](#)) are associated with schools of nursing. As can be seen in **Table 2**, most other categories also have a link with health, and a minority with social sciences and multidisciplinary fields. “Education, educational research” shares the 14th, 15th, and 16th position with “Gerontology” and “Psychology, multidisciplinary” and is mentioned in only 13 records (2.5%). Further down the list (not in the Table), we encounter “Education, scientific disciplines” with 6 mentions (1.2%), “Psychology, educational” with 4 mentions (0.8%), and “Education, special” with 1 mention (0.2%).

Table 2 Number and percentage of mixed methods research synthesis records for each of the web of science categories.

<i>Web of science categories</i>	<i>Records</i>	<i>% of 514</i>
Nursing	103	20.0
Public environmental occupational health	93	18.1
Medicine general internal	82	16.0
Health care sciences services	62	12.1
Health policy services	30	5.8
Psychiatry	24	4.7
Rehabilitation	23	4.5
Social sciences interdisciplinary	22	4.3
Oncology	17	3.3
Pediatrics	17	3.3
Multidisciplinary sciences	14	2.7
Social sciences biomedical	14	2.7
Education educational research	13	2.5
Gerontology	13	2.5
Psychology multidisciplinary	13	2.5
Psychology clinical	12	2.3
Clinical neurology	11	2.1
Nutrition dietetics	11	2.1
Obstetrics gynecology	11	2.1

Note. Only the 20 most prevalent categories are included in the table. Percentages do not sum to 100% because records can be listed in more than one category.

The 22 unique records (4.3%) that are in an educational category are included in [Mixed methods research synthesis in educational research](#) Appendix. This relatively small number of articles referring to mixed methods research synthesis in educational research suggests that this approach is not reaching its full potential in this research field. Evidently, education is different from nursing, but they share the complexity, the societal relevance, and the importance of both qualitative and quantitative features. Moreover, a closer look at [Mixed methods research synthesis in educational research](#) Appendix learns that most mixed methods research syntheses in educational research even have a connection with health, exemplified by the fact that 14 of those 22 articles were published in medical education journals (*BMC Medical Education*, *Health Education Journal*, *Journal of Education and Health Promotion*, *Medical Teacher*, *Nurse Education Today*, and *Sex Education: Sexuality, Society and Learning*), school health journals (*Journal of American College Health*, *School Mental Health*, *The Journal of School Health*), or in journals on disabilities (*Journal of Applied Research in Intellectual Disabilities* and *Research in Developmental Disabilities*). It is difficult to fully understand this difference in the prevalence of mixed methods research synthesis between health-related research and educational research. Possibly the evidence-based medicine movement is a crucial factor, with educational research being less focused on summarizing evidence about effectiveness and safety, which is most prominent in the systematic review and meta-analysis literature.

Mixed methods research synthesis in educational research: an example

In this final section, mixed methods research synthesis is illustrated with an example from educational research found in our search: a study on sense-making for beginning special educators ([Mathews et al., 2017](#)). The example demonstrates that mixed methods research synthesis does not only deal with effectiveness and safety questions in health-related research fields, but that it also can be harnessed for more descriptive research questions in educational research. In this example, we follow the six mixed methods research synthesis steps mentioned before.

Writing a protocol and selecting a design

[Mathews et al. \(2017\)](#) posited the following research questions: “As represented in the literature, which elements of schools as organizations –including stressors and supports– are most influential in beginning special educator’s sense-making process? How do these elements promote or hinder beginning special educator’s development of professional identity and their response to the unique complexities of their professional role?” (p. 26). These research questions contain causal language, but the literature study is mainly about perceptions of “influence”, “promotion”, and “hindrance”.

For their theoretical framework, [Mathews et al. \(2017\)](#) drew on institutional and sense-making theories. Institutional theory refers to a theory about the ways in which organizational structures, norms, practices, and patterns of social relationships are connected to the broader social and cultural environment. Sense-making theory is used to focus on individual processes through which the beginning special educators construct their own understanding of policies, norms, and cultural messages, and how they construct meaning regarding their role, responsibility, and identity as a special education teacher. With respect to design, the authors used an integrated mixed methods research synthesis design in which qualitative and quantitative studies were directly combined using their theoretical framework. Their framework enabled them to transform the quantitative data into qualitative data.

Selecting a sampling strategy, a search strategy, and eligibility criteria

Their sampling strategy was purposeful. They conducted an electronic search of EBSCO and PsycNet databases and did a hand search of two journals that featured prominently in the electronic search, *Teacher Education and Special Education* and *Exceptional Children*. Furthermore, they scrutinized the references of and the citations to every selected article. For the search strategy, they used key words such as “teacher induction”, “teacher mentoring”, “special educat*”, “special education teacher”, “novice”, and “begin*” in various combinations, and they limited their search to articles published from 2002 onwards because they wanted to capture the experience of beginning special educators since the United States “No Child Left Behind” act was implemented. Given their research questions and for practical purposes, they only included studies that (1) examined special education teachers as a distinct group in sampling and results, (2) addressed beginning special educators in the induction period or mentoring programs, (3) were conducted in primary or secondary education settings in the United States, and (4) were available in English in a peer-reviewed journal.

Critically appraising the primary-level studies

The selected articles were critically appraised by using a modified version of the Mixed Methods Appraisal Tool (MMAT; [Pluye and Hong, 2014](#)). A minimum score of 25% on the MMAT was set by the research team for inclusion. Based on this criterion, three of the 21 articles were excluded. Furthermore, the MMAT scores were also used in the analysis (see Step 5); claims made in studies with a MMAT score smaller than 75% always required the support of studies with a MMAT score larger than 75%. The MMAT scores were double-coded for 35% of articles and an inter-rater agreement of 96% was obtained.

Extracting the data

The coding guide consisted of codes for the quantitative data with respect to the stressors and supports, the role of the stressors and supports in the professional experiences of the beginning special educator, and the extent to which factors such as mentoring and professional support helped beginning special educators navigate those stressors. Quantitative coding was performed for *a priori* specified variables deduced from the literature on teacher attrition in special education. The inter-rater agreement for the codes of the quantitative data was 75%, but the research team came to a consensus on all codes after a careful deliberation. The codes for the qualitative data emerged inductively during the analysis. Those codes specified how the stressors and supports, identified by the quantitative data, shaped the professional identity of the beginning special educator. After discussion and refinement with the research team, three main themes and corresponding codes emerged: affirmation, interruption, and identity construction. "Affirmation" referred to experiences that confirmed what was expected or anticipated, "interruption" referred to experiences that differed from what was expected or anticipated, and "identity construction" referred to experiences that defined, clarified, or potentially shaped the beginning special educator's professional role, responsibilities, and position within the organization.

Analyzing and synthesizing the data

A total of 9 quantitative, 8 qualitative, and 1 mixed methods study were analyzed and synthesized. All quantitative studies were surveys. The qualitative studies contained interviews, observations, journals or written reflections, questionnaires, teacher evaluation documents, and mentoring time charts. The mixed methods study combined interviews, observations, and a survey. During the analysis, the authors looked for emerging themes, relationships among those themes, and disconfirming evidence. They refrained from an actuarial approach and grouped the emerging themes with a focus on examining the influence of institutional elements on beginning special educator's sense-making.

Writing the report

For writing their report, Mathews et al. (2017) did not use a specific template, but rather presented their findings in a regular journal article format. In doing so, they used their theoretical framework to organize their report, present the results, and discuss the implications. The main findings from their mixed methods research synthesis were that (1) beginning special educators needed support to access resources and understand local school norms and culture, (2) rising student needs were associated with rising needs for support on the part of the beginning student educator, (3) school climate, including acceptance of students with disabilities and collective responsibility for student needs, is associated with a perception of relationships as collaborative, (4) principal leadership was instrumental in establishing school norms that supported a climate accepting of students with disabilities and in clarifying the role of special education teachers, and (5) the professional identities and roles of the beginning special educators were highly subjective and shaped less by fixed, regulated structures and more by local, dynamic interpretations of school regulations, school norms, and school culture.

Conclusion

Mixed methods research synthesis has become a versatile and popular method to systematically review research literature consisting of both qualitative and quantitative evidence for complex real-world phenomena. After initial methodological proposals and a few scarce applications in the first decade of this century, the number of published mixed methods research syntheses has risen to about 150 in the year 2020 alone and is expected to even surpass this yearly number in the years to come. This growing popularity is supported by the publication of step-by-step user's guides, such as in the handbook by Heyvaert et al. (2017) and a number of instructive journal articles that outlined the method, such as the articles by Heyvaert et al. (2013b), Pluye and Hong (2014), and Sandelowski et al. (2012).

This popularity is not a historical accident. As was argued in this article, mixed methods research synthesis is the method of choice for a systematic review if the research problem of interest is complex, consists of multiple components and layers, and if quantitative, qualitative, and mixed method studies have to be combined to tackle the research problem. Because systematic reviewing is more and more applied to complex research problems with important societal relevance and because qualitative, quantitative, and mixed method studies about similar phenomena are increasingly available and valued, it is expected that a mixed methods approach to research synthesis would be on the rise.

Surprisingly, however, in our exploration of the research literature, we only found 22 mixed methods research syntheses in the field of educational research out of a total of 514 applications in all fields between 2007 and 2020. This small number is surprising because educational research commonly has to deal with complex research problems that have huge societal relevance and because often qualitative, quantitative, and mixed method studies should be brought to bear to find a balanced, convincing, and sustainable solution. In this sense, educational research would be an ideal breeding ground for applications and methodological advances in mixed methods research synthesis.

With this article in the fourth edition of the *International Encyclopedia of Education*, we hope to have planted the seed for countering the underrepresentation of educational research in the mixed methods research synthesis literature. In this article, we have introduced this new method of systematic reviewing to a new generation of educational researchers, together with the benefits and pitfalls, with an accessible step-by-step summary, and with a worked-out example from educational research. After planting this seed, we hope that in the fifth edition a more rosy **Table 2** for educational research can be presented.

Appendix 1: Mixed methods research synthesis in educational research

1. Alley, K.M., Peterson, B.J., 2017. Ideas as a springboard for writing in K-8 classrooms. *Writing Instruction to Support Literacy Success* (Literacy Research, Practice and Evaluation, vol. 7), pp 65–93. Emerald Group Publishing Limited.
2. Connell, C.J., Endacott, R., Jackman, J.A., Kiprillis, N.R., Sparkes, L.M., Cooper, S.J., 2016. The effectiveness of education in the recognition and management of deteriorating patients: A systematic review. *Nurse Educ. Today* 44, 133–145.
3. Galbraith, K., Ward, A., Heneghan, C., 2017. A real-world approach to Evidence-Based Medicine in general practice: A competency framework derived from a systematic review and Delphi process. *BMC Med. Educ.* 17, 78.
4. Heyvaert, M., Saenen, L., Maes, B., Onghena, P., 2014. Systematic review of restraint interventions for challenging behaviour among persons with intellectual disabilities: Focus on effectiveness in single-case experiments. *J. Appl. Res. Intell. Disabil.* 27, 493–510.
5. Heyvaert, M., Saenen, L., Maes, B., Onghena, P., 2015. Systematic review of restraint interventions for challenging behaviour among persons with intellectual disabilities: Focus on experiences. *J. Appl. Res. Intell. Disabil.* 28, 61–80.
6. Hong, Q.N., Fabregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.C., Vedel, I., Pluye, P., 2018. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ. Inf.* 34, 285–291.
7. Jones, M., Hutt, P., Eastwood, S., Singh, S., 2013. Impact of an intercalated BSc on medical student performance and careers: A BEME systematic review: BEME Guide No. 28. *Med. Teach.* 35, E1493–E1510.
8. Kasymova, S., 2020. Human papillomavirus (HPV) and HPV vaccine knowledge, the intention to vaccinate, and HPV vaccination uptake among male college students. *J. Am. Coll. Health*. Advance online publication.
9. Karbasi, C., Pacheco, E., Bull, C., Evanson, A., Chaboyer, W., 2018. Registered nurses’ provision of end-of-life care to hospitalised adults: A mixed studies review. *Nurse Educ. Today* 71, 60–74.
10. Kariippanon, K.E., Cliff, D.P., Ellis, Y.G., Ucci, M., Okely, A.D., Parrish, A.M., 2021. School flexible learning spaces, student movement behavior and educational outcomes among adolescents: A mixed-methods systematic review. *J. Sch. Health* 91, 133–145.
11. Lindsay, S., Rezai, M., Kolne, K., Osten, V., 2019. Outcomes of gender-sensitivity educational interventions for healthcare providers: A systematic review. *Health Educ. J.* 78, 958–976.
12. Litsou, K., Byron, P., McKee, A., Ingham, R., 2020. Learning from pornography: Results of a mixed methods systematic review. *Sex Educ.: Sex., Soc. Learn.* Advance online publication.
13. Mathews, H.M., Rodgers, W.J., Youngs, P., 2017. Sense-making for beginning special educators: A systematic mixed studies review. *Teac. Educ.* 67, 23–36.
14. McGarty, A.M., Melville, C.A., 2018. Parental perceptions of facilitators and barriers to physical activity for children with intellectual disabilities: A mixed methods systematic review. *Res. Dev. Disabil.* 73, 40–57.
15. Nye, E., Melendez-Torres, G.J., Gardner, F., 2019. Mixed methods systematic review on effectiveness and experiences of the Incredible Years Teacher Classroom Management programme. *Rev. Educ.* 7, 631–669.
16. Nye, E., Melendez-Torres, G.J., Gardner, F., 2019. Context and implications document for: Mixed methods systematic review on effectiveness and experiences of the Incredible Years Teacher Classroom Management programme. *Rev. Educ.* 7, 670–674.
17. O’Mara-Eves, A., Thomas, J., 2016. Ongoing developments in meta-analytic and quantitative synthesis methods: Broadening the types of research questions that can be addressed. *Rev. Educ.* 4, 5–27.
18. Pluye, P., Hong, Q.N., Granikov, V., Vedel, I., 2018. The wiki toolkit for planning, conducting and reporting mixed studies reviews. *Educ. Inf.* 34, 277–283.
19. Sandu, C., Panisoara, G., Lazar, I., Fat, S., 2018. A tool for specialists to explore the quality of IT assistance in speech therapy. *Proceedings of the 11th annual International Conference of Education, Research and Innovation (ICERI2018)*, 5167–5172.
20. Simons, M.A.G., Koordeman, R., Willems, A.P.A.M., Hermsen, M., Rooijackers, L.M., Otten, R., 2020. Factors facilitating or hindering meaningful staff-client interactions in people with intellectual disabilities and challenging behaviour: A systematic mixed studies review using thematic synthesis. *J. Appl. Res. Intell. Disabil.* Advance online publication.
21. Swaminathan, N., Ravichandran, L., Ramachandran, S., Milanese, S., 2020. Blended learning and health professional education: Protocol for a mixed-method systematic review. *J. Educ. Health Promot.* 9, 46.
22. Tougas, A.M., Rassy, J., Frenette-Bergeron, E., Marcil, K., 2019. Lost in transition: A systematic mixed studies review of problems and needs associated with school reintegration after psychiatric hospitalization *School Ment. Health* 11, 629–649.

References

- Biesta, G., 2010. Pragmatism and the philosophical foundations of mixed methods research. In: Tashakkori, A., Teddlé, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 95–118.
- Booth, A., Sutton, A., Papaioannou, D., 2016. *Systematic Approaches to a Successful Literature Review*, second ed. Sage, London, UK.
- Booth, A., 2006. Brimful of STARLITE: toward standards for reporting literature searches. *J. Med. Libr. Assoc.* 94, 421–429.
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. *Introduction to Meta-analysis*. Wiley, Chichester, UK.
- Borman, G.D., Grigg, J.A., 2009. Visual and narrative interpretation. In: Cooper, H., Hedges, L.V., Valentine, J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*, second ed. Russell Sage Foundation, New York, NY, pp. 497–519.
- Britten, N., Campbell, R., Pope, C., Donovan, J., Morgan, M., Pill, R., 2002. Using meta ethnography to synthesise qualitative research: a worked example. *J. Health Serv. Res. Pol.* 7, 209–215.
- Bryman, A., 2007. Barriers to integrating quantitative and qualitative research. *J. Mix. Methods Res.* 1, 8–22.
- Cersosimo, G., 2019. Pragmatism. In: Atkinson, P., Delamont, S., Cernat, A., Sakshaug, J.W., Williams, R.A. (Eds.), *SAGE Research Methods Foundations*. <https://doi.org/10.4135/9781526421036837556>. (Accessed 28 November 2020).
- Chalmers, I., Hedges, L.V., Cooper, H., 2002. A brief history of research synthesis. *Eval. Health Prof.* 25, 12–37.
- Clarke, M., Chalmers, I., 2018. Reflections on the history of systematic reviews. *BMJ Evid. Based Med.* 23, 121–122.
- Cooper, H.M., 2010. Gathering information from studies. In: Cooper, H.M. (Ed.), *Research Synthesis and Meta-Analysis: A Step-by-step Approach*, fourth ed. Sage, London, UK, pp. 84–144.
- Crandell, J.L., Voils, C.I., Chang, Y., Sandelowski, M., 2011. Bayesian data augmentation methods for the synthesis of qualitative and quantitative research findings. *Qual. Quantity* 45, 653–669.
- Creswell, J.W., Plano Clark, V.L., Gutmann, M.L., Hanson, W.E., 2003. Advanced mixed methods research designs. In: Tashakkori, A., Teddlé, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 209–240.
- Curry, L., Nunez-Smith, M., 2015. Managing mixed methods teams. In: Curry, L., Nunez-Smith, M. (Eds.), *Mixed Methods in Health Sciences Research*. Sage, Thousand Oaks, CA, pp. 259–280.
- Dewey, J., 1925. *Experience and Nature*. Kessinger, Whitefish, MT.
- Dixon-Woods, M., Bonas, S., Booth, A., Jones, D.R., Miller, T., Sutton, A., Shaw, R.L., Smith, J.A., Young, B., 2006a. How can systematic reviews incorporate qualitative research? A critical perspective. *Qual. Res.* 6, 27–44.
- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L., Riley, R., Sutton, A.J., 2006b. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med. Res. Methodol.* 6, 35.
- Eaves, Y.D., 2001. A synthesis technique for grounded theory data analysis. *J. Adv. Nurs.* 35, 654–663.
- Feilzer, M.Y., 2010. Doing mixed methods research pragmatically: implications for the rediscovery of pragmatism as a research paradigm. *J. Mix. Methods Res.* 4, 6–16.
- Ferguson, S.L., Kerrigan, M.R., Hovey, K.A., 2020. Leveraging the opportunities of mixed methods in research synthesis: key decisions in systematic mixed studies review methodology. *Res. Synth. Methods* 11, 580–593.
- Gough, D., Oliver, S., Thomas, J. (Eds.), 2017. *An Introduction to Systematic Reviews*, second ed. Sage, London, UK.
- Hannes, K., Lockwood, C., 2011. Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregative approach to qualitative evidence synthesis. *J. Adv. Nurs.* 67, 1632–1642.
- Harden, A., Thomas, J., 2005. Methodological issues in combining diverse study types in systematic reviews. *Int. J. Soc. Res. Methodol.* 8, 257–271.
- Hart, C., 2018. *Doing a Literature Review: Releasing the Research Imagination*. Sage, London, UK.
- Hathcoat, J.D., Meixner, C., Nicholas, M.C., 2018. Ontology and epistemology. In: Liampittong, P. (Ed.), *Handbook of Research Methods in Health Social Sciences*. Springer, Singapore, pp. 1–19.
- Heyvaert, M., Maes, B., Onghena, P., 2011. Applying mixed methods research at the synthesis level: an overview. *Res. Sch.* 18, 12–24.
- Heyvaert, M., Hannes, K., Maes, B., Onghena, P., 2013a. Quality assessment of mixed methods studies. *J. Mix. Methods Res.* 7, 302–327.
- Heyvaert, M., Maes, B., Onghena, P., 2013b. Mixed methods research synthesis: definition, framework, and potential. *Qual. Quantity* 47, 659–676.
- Heyvaert, M., Saenen, L., Maes, B., Onghena, P., 2014. Systematic review of restraint interventions for challenging behaviour among persons with intellectual disabilities: focus on effectiveness in single-case experiments. *J. Appl. Res. Intellect. Disabil.* 27, 493–510.
- Heyvaert, M., Saenen, L., Maes, B., Onghena, P., 2015. Systematic review of restraint interventions for challenging behaviour among persons with intellectual disabilities: focus on experiences. *J. Appl. Res. Intellect. Disabil.* 28, 61–80.
- Heyvaert, M., Hannes, K., Onghena, P., 2017. *Using Mixed Methods Research Synthesis for Literature Reviews*. Sage, Thousand Oaks, CA.
- Hong, Q.N., Rees, R., Sutcliffe, K., Thomas, J., 2020. Variations of mixed methods reviews approaches: a case study. *Res. Synth. Methods* 11, 795–811.
- Howe, K.R., 1988. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educ. Res.* 17, 10–16.
- James, W., 1907/2003. *Pragmatism*. Barnes and Noble, New York, NY.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Kearney, M.H., 1998. Truthful self-nurturing: a grounded formal theory of women's addiction recovery. *Qual. Health Res.* 8, 495–512.
- Kearney, M.H., 2001. Enduring love: a grounded formal theory of women's experience of domestic violence. *Res. Nurs. Health* 24, 270–282.
- Lasserson, T.J., Thomas, J., Higgins, J.P.T., 2021. Chapter 1: starting a review. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.2 (updated February 2021). Available from: <https://training.cochrane.org/handbook/current/chapter-01>.
- Legg, C., Hookway, C., 2020. Pragmatism. In: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2020/entries/pragmatism/>.
- Lockwood, C., Munn, Z., Porritt, K., 2015. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. *Int. J. Evid. Base. Healthc.* 13, 179–187.
- Mathews, H.M., Rodgers, W.J., Youngs, P., 2017. Sense-making for beginning special educators: a systematic mixed studies review. *Teach. Teach. Educ.* 67, 23–36.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151, 264–269.
- Moon, K., Blackman, D., 2014. A guide to understanding social science research for natural scientists. *Conserv. Biol.* 28, 1167–1177.
- Morgan, G., Smircich, L., 1980. The case for qualitative research. *Acad. Manag. Rev.* 5, 491–500.
- Newman, I., Benz, C.R., 1998. *Qualitative-quantitative Research Methodology: Exploring the Interactive Continuum*. Southern Illinois University Press, Carbondale, IL.
- Noblit, G.W., Hare, R.D., 1988. *Meta-ethnography: Synthesising Qualitative Studies*. Sage, London, UK.
- Noyes, J., Booth, A., Moore, G., Flemming, K., Tunçalp, Ö., Shakibazadeh, E., 2019. Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Global Health* 4 (Suppl. 1), e000893.
- O'Cathain, A., Murphy, E., Nicholl, J., 2008. Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research. *Qual. Health Res.* 18, 1574–1585.
- Pawson, R., 2002. Evidence-based policy: the promise of 'realist synthesis'. *Evaluation* 8, 340–358.
- Pawson, R., 2006. *Evidence-based Policy: A Realist Perspective*. Sage, London, UK.
- Pearson, A., Wiechula, R., Court, A., Lockwood, C., 2005. The JBI model of evidence-based healthcare. *Int. J. Evid. Base. Healthc.* 3, 207–215.

- Peirce, C.S., 1878/1992. How to make our ideas clear. In: Houser, N., Kloesel, C. (Eds.), *The Essential Peirce*, vol. 1. Indiana University Press, Bloomington IN, pp. 124–141.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Pluye, P., Hong, Q.N., 2014. Combining the power of stories and numbers: mixed methods research and mixed studies reviews. *Annu. Rev. Publ. Health* 35, 29–45.
- Ravitch, S.M., Carl, N.M., 2019. *Qualitative Research: Bridging the Conceptual, Theoretical and Methodological*, second ed. Sage, Thousand Oaks, CA.
- Roberts, K.A., Dixon-Woods, M., Fitzpatrick, R., Abrams, K.R., Jones, D.R., 2002. Factors affecting uptake of childhood immunisation: a Bayesian synthesis of qualitative and quantitative evidence. *Lancet* 360, 1596–1599.
- Sale, J.E., Lohfeld, L.H., Brazil, K., 2002. Revisiting the quantitative-qualitative debate: implications for mixed-methods research. *Qual. Quantity* 36, 43–53.
- Sandelowski, M., Voils, C.I., Barroso, J., 2006. Defining and designing mixed research synthesis studies. *Res. Sch.* 13, 29–40.
- Sandelowski, M., Voils, C.I., Leeman, J., Crandell, J.L., 2012. Mapping the mixed methods-mixed research synthesis terrain. *J. Mix. Methods Res.* 6, 317–331.
- Savin-Baden, M., Major, C.H., 2013. *Qualitative Research: The Essential Guide to Theory and Practice*. Routledge, New York, NY.
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., The PRISMA-P Group, 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 349, g7647.
- Shulha, L.M., Wilson, R.J., 2003. Collaborative mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 639–669.
- Snyder, H., 2019. Literature review as a research methodology: an overview and guidelines. *J. Bus. Res.* 104, 333–339.
- Thomas, J., Harden, A., 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med. Res. Methodol.* 8, 45.
- Tong, A., Flemming, K., McInnes, E., Oliver, S., Craig, J., 2012. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Med. Res. Methodol.* 12, 181.
- Voils, C.I., Hasselblad, V., Crandell, J., Chang, Y., Lee, E., Sandelowski, M., 2009. A Bayesian method for the synthesis of evidence from qualitative and quantitative reports: the example of antiretroviral medication adherence. *J. Health Serv. Res. Pol.* 14, 226–233.
- Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., Pawson, R., 2013. RAMESES publication standards: realist syntheses. *BMC Med.* 11, 21.

Mixed methods in survey research

John H. Hitchcock and Allison Crean Davis, Westat, Rockville, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	666
Mixed methods survey research	666
Mixed methods surveys and understanding individual respondents	667
Using ecological systems theory to conceptualize a survey project	667
Approaches to using mixed methods in education survey research	668
Understanding mixed methods survey research using the NSLES example	668
Introducing the context of our survey project	668
Mixed methods design decisions applied to NSLES	670
Applying EST can help the surveyor think more completely about the survey context	671
Mixed methods survey design options	672
Data analyses and interpretation	674
Conclusion	676
References	676

Introduction

A mixed methods approach to survey work in the education sector can yield benefits when it comes to instrument design (i.e., item writing, item placement, developing directions, a scoring scheme, and so on) and when conducting a full survey project. For purposes of this chapter, we define a survey project as a survey endeavor that includes all facets of survey work, including research question conceptualization, instrument design, framing a sample, drawing a sample, performing analyses, interpreting findings within the survey context, reporting, and dissemination of findings. This chapter describes general survey concepts and discusses how mixed methods are being used in a survey project we (the chapter authors) are currently leading called the *National Summer Learning and Enrichment Study* (NSLES). Discussion of mixed methods used in the NSLES offers readers an example meant to help describe some of the cost-benefit decisions researchers make when engaged in survey work with a mixed methods approach, particularly with respect to a projects' timeline and budget. To provide our conclusions up front, we see value in mixed methods survey designs, particularly one design where qualitative inquiry allows one to understand the research context in detail, prior to writing a survey instrument and delivering it. However, this design can be time consuming, so we opted for a mixed methods approach that allows for near simultaneous use of qualitative and quantitative inquiry, which was desirable because our example entailed a fast turnaround study.

Mixed methods survey research

Survey research is generally seen as a quantitative approach wherein an instrument is used to gather information from a large group of people. Much of survey item writing entails reducing respondent opinions to a numerically coded responses, and many survey sampling efforts will seek a sample that is statistically representative of some population of interest. Response samples are typically assessed for nonresponse bias, and statistical weighting and data imputation are often used to build confidence in results (see for e.g., Enders, 2010). These quantitative design approaches and statistical techniques yield data collection efficiencies and allow researchers to assess the validity of their findings.

Quantitative survey findings can respond broadly to critical research questions (e.g., for the NSLES: *What kinds of summer programming did school districts launch in 2021?*) and can provide insight into if, and how, the phenomenon in question varies across respondent characteristics (e.g., *Did districts with prior experience in summer programming provide a broader range of offerings than those that were launching summer programming for the first time?*). But if limited to quantitative methods, findings from surveys may lack sufficient detail and depth to fully explain what is happening, why, and under what conditions. This is because people are complicated (think about the different contexts in which people live, think, and behave) and their views are not easily boiled down to neat numerical scales and categorical choices. Because of these complexities, strictly quantitatively-oriented survey projects are vulnerable to multiple sources of error (reviewed below) that have the potential to threaten the validity of findings (see Dillman et al., 2014).

Mixed methods thinking can help limit these concerns. This type of thinking is an "orientation toward social inquiry that actively invites us to participate in dialog about multiple ways of seeing and hearing, multiple ways of making sense of the social world, and multiple standpoints on what is important and to be valued and cherished" (Greene, 2008, p. 20). Survey researchers who exercise mixed methods thinking hold the strengths and limitations of survey instruments front-and-center as they conceptualize how an instrument is designed, administered, analyzed, and how they report what they learn from it. They consider if and how integrating

qualitative inquiry within a survey (or via complementary tools) can add detail and depth to findings that stretch beyond what quantitative scales can obtain alone. By introducing qualitative inquiry into their work, researchers can find additional ways of capturing respondents' inherently multi-dimensional and nuanced opinions and life experiences, enable a deeper understanding of context, and mitigate survey error.

Mixed methods surveys and understanding individual respondents

When responding to a survey, individuals bring their prior assumptions about a topic (both explicit and tacit), their perceptions of themselves, their emotions related to specific topics, and their reasoning to the assignment. This can be true even if dealing with highly specific and technical matters. If, for example, a survey queries about staffing educational programs, it is likely respondents will have views informed by their prior school experiences. If a survey includes questions about how schools allocate limited funding, again most respondents will likely have underlying tacit assumptions that will influence how they respond. Tacit assumptions are personal opinions and ideas that may guide respondent behavior without their full awareness or, perhaps, without thorough self-examination (LeCompte, 2000). In addition to thinking about respondent assumptions, it might help to conceptualize respondent behavior; that is, how one interacts with a survey. Survey behavior can include whether one chooses to respond to the survey in whole, ignore some items, remain engaged during the response process (e.g., reading questions carefully and thinking through answers as opposed to offering rote responses to several items just to get through the experience), and so on. Tying these concepts together, if survey researchers are to expect reasonably high response rates, and thoughtful and honest responses (i.e., desirable survey behavior), it is critical to have some sense of the tacit assumptions respondents will likely carry.

Another challenge in survey work is that it can be difficult for a survey researcher to know what to even query in a survey, or to have confidence that a set of items adequately cover a topic because human beings are complex. After all, people are intersectional; that is, they exist within dynamic social contexts that shape their identities (American Psychological Association [APA], 2017). It is also important for survey researchers to develop a sense of the context in which the survey project will be conducted. A mix of quantitative and qualitative information can help account for contextual variation, which informs and influences individuals, the ways in which they experience and perceive the world, and how they might respond to a survey.

Using ecological systems theory to conceptualize a survey project

Bronfenbrenner's (1979) ecological systems theory (EST) provides a basis for thinking about context at different levels, its dynamic nature, and how it is co-constructed by people who live and operate within it. Furthermore EST, summarized in Table 1, can be used by researchers when making methodological decisions (Onwuegbuzie et al., 2013), particularly when considering how to blend quantitative and qualitative items in a way that informs findings and helps audiences consider the impact of programs and practices on individuals across a range of situational influences.

For example, this chapter was written in the midst of a recent and significant contextual shift: the global COVID-19 pandemic and this strongly influenced the NSLES. During this time, schools abruptly closed their doors, students began remote learning at

Table 1 Summary of Bronfenbrenner's ecological systems theory, focusing on system levels.

<i>System level</i>	<i>Description as applied to education research</i>
Microsystem	The level at which survey respondents interact with and influence their immediate environment; this represents context a person can co-construct (e.g., a student influencing the classroom context).
Mesosystem	This is the level wherein microsystems interact (e.g., a student's parents and teachers working together entails interaction between classroom and family microsystems).
Exosystem	There are typically multiple exosystems that influence microsystems but are not directly influenced by research participants (e.g., the school or district a student attends).
Macrosystem	This level represents broader cultural and societal factors (how a society values the teaching profession, funds schools, etc.).
Chronosystem	This level refers to time and changes brought on by development and life events.

scale and with little to no preparation, schools then began to reopen, and state and district educational agencies received federal stimulus funds to, among other things, accelerate student learning to make up for lost time. In addition to educational issues, health and safety issues, have and continue to weigh heavily on people's minds, but there is also controversy about matters such as vaccine mandates and whether children should wear masks in schools. Such context will influence how people respond to most surveys pertaining to education topics. The survey we are working on addresses urgent questions related to how educators are responding to this challenge within the United States, and we assume respondents will be influenced by their views around how schools should operate during the COVID-19 pandemic.

A key point here is that although a survey must be designed to inform research questions, it should also account for context, intersectionality, tacit assumptions, and individual differences throughout its construction and/or with complementary methodology (e.g., interviews). Mixed methods research can accomplish such survey design by blending quantitative inquiry with a qualitative mindset throughout different stages of a survey project is emblematic of [Greene's \(2007\)](#) "mixed methods way of thinking" (p. 20). Survey researchers with both quantitative *and* qualitative information can have a more holistic understanding of research questions.

Approaches to using mixed methods in education survey research

Qualitative data collection can be integrated within a survey instrument or can be accomplished through supporting tools within a larger survey project. For example, there is an option to obtain narrative responses via survey items, or a researcher can ask respondents to provide representative documents (e.g., written summer programming plans), both of which might be subject to qualitative analyses. In addition, survey efforts could be supplemented with qualitative phases (e.g., interviews, focus groups, document analyses) and it is possible for a project to be weighted toward qualitative work. At a more fundamental level, survey data can be seen as having both quantitative *and* qualitative features ([Sandelowski et al., 2009](#)). This is because responding to Likert-type items can be seen as a way to quantize a respondent's opinion and, conversely, one could qualitize a numerical response by asking a respondent to elaborate in a follow-up interview ([Onwuegbuzie and Leech, 2021](#)). Hence, survey work (as in instrument design and when conducting a full survey project) can be highly aligned with mixed methods thinking and design. An implication of this view of mixed methods survey work is researchers can use both qualitative and quantitative inquiry in flexible ways to conduct a survey project.

Understanding mixed methods survey research using the NSLES example

This chapter describes how mixed methods survey ideas can be pursued by describing how the authors applied them to a current survey project being carried out at a national level, within the United States. This chapter does not present survey findings; indeed, these findings are arguably irrelevant to our purpose in writing this chapter. Rather, we present mixed methods design thinking that went into conceptualizing our project so that readers might consider these matters when they engage in their own mixed methods survey design work. Some, but not all of the mixed methods ideas presented in this chapter have been used in the actual survey effort, and we describe how we arrived at a range of these methodological decisions. For example, due to the time-sensitivity of the topic and a strong appetite for rapid feedback for the education sector, we had to conceptualize and field our survey within a matter of weeks. Yet several techniques described in the remainder of this chapter take up time and resources that were not consistent with this study's demands. We therefore direct the reader to published studies that used survey techniques that we did not use in our survey project. Readers will need to perform their own cost-benefit analyses for their own survey projects. In our project, we are confident in our decisions to not use some mixed methods techniques because we think that we have a good grasp of the survey context at hand. Through this presentation, we hope to clarify the kinds to cost-benefit analyses that are familiar to many social scientists.

Introducing the context of our survey project

Fig. 1 presents our use of EST to understand the survey context. This figure uses EST levels from **Table 1** and provides a visual depiction of how these systems interrelate in the context our national survey project, which focused on understanding school-district summer programming. The broadest EST level is the macrosystem, and we conceptualize this as US macro-culture, federal policy, and national response to the pandemic. Within the macrosystem are exosystems, which we conceptualized as school-districts and the community a district served. Within a given exosystem, we conceptualized mesosystems that capture the interaction between summer programming and other district and community initiatives. We conceptualized specific summer programs as microsystems. Finally, EST reminded us to consider the chronosystem, or change over time. The implication of the figure is it provided us with guidance around item writing (e.g., we asked about micro-level summer program characteristics, and how they were informed by mesosystem-level community partnerships) and interpreting responses within the macrosystem and across time.

The survey project the chapter co-authors are leading is designed to promote an understanding of the types of summer programming school districts offer in 2021. There are several challenges and opportunities surrounding this project. To begin, at the exosystem and chronosystem levels, the United States is as of this writing struggling with social and economic consequences of the COVID-19 pandemic. Schools have experienced long periods of unplanned, emergency remote instruction and there remains

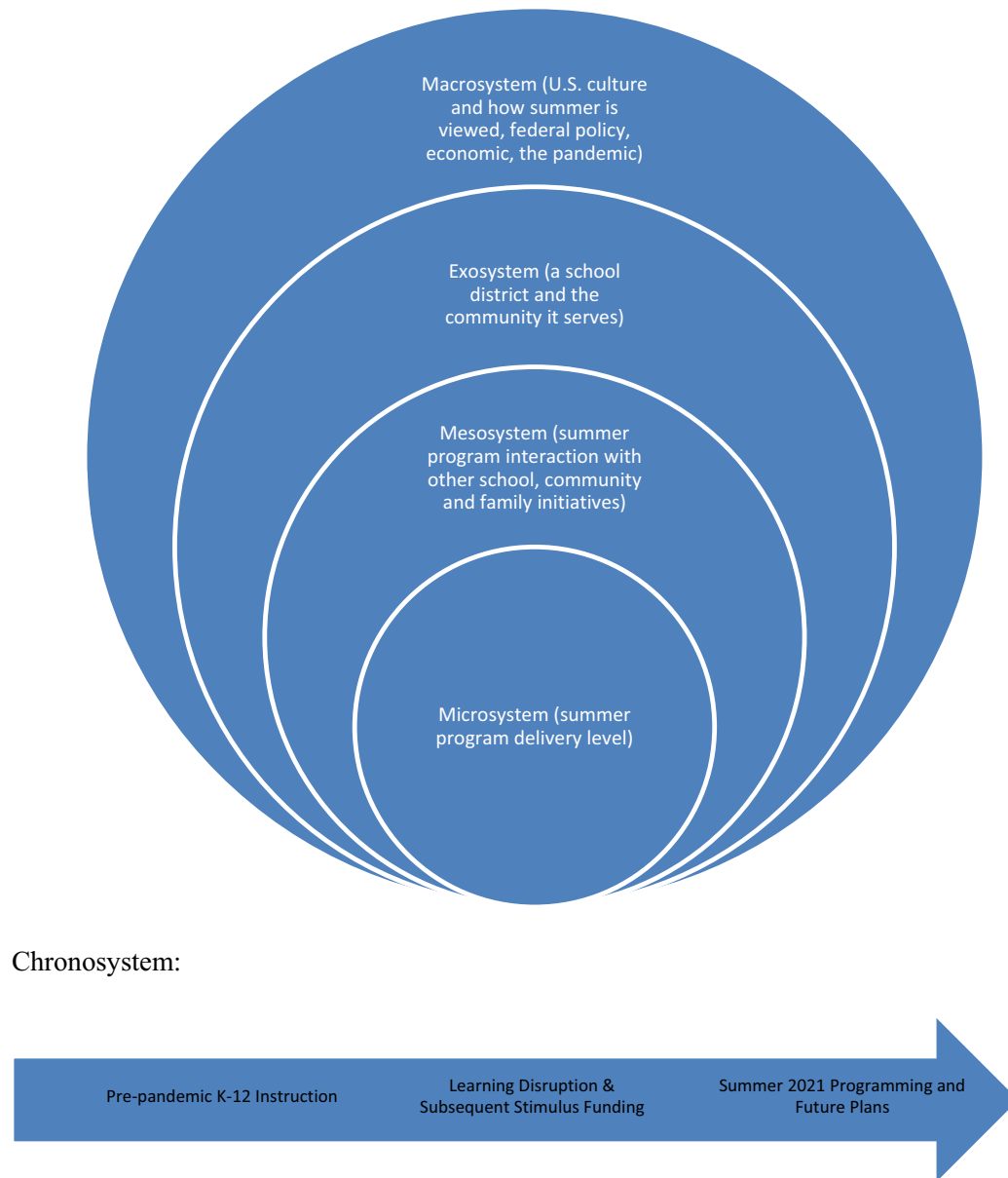


Fig. 1 Applying EST to understand the context of a survey project.

deep concern that interruptions to normal learning processes will reap long-term developmental challenges for the nation's children, especially among more vulnerable groups. Consider for example that the New York Times (a prominent newspaper in the country) worked with Stanford University to analyze kindergarten enrollment data from 33 states, covering 10,000 schools. The analyses showed a 20% enrollment drop among these children; furthermore, this drop in enrollment increased by 28% in communities close to or below the poverty line (Goldstein and Parlapiano, 2021). The implication of these enrollment declines is that tens of thousands of students began the first grade in 2021 without the benefit of a year in kindergarten.

Long-term consequences from this event are to be expected. Of course, one might wonder if some on-line instruction at least ameliorated the problem; however, Engzell et al. (2021) offer some troubling insights. These researchers evaluated the effect of school closures on primary school children's academic performance in The Netherlands by exploiting the fact that national examinations were conducted shortly before and after an 8-week lockdown was initiated to combat virus spread. Engzell et al. compared the academic progress of these children across these eight weeks to three prior cohorts for which there are testing data during the same point in the school year. The main finding is that learning loss for the cohort that experienced the pandemic induced school closure was equivalent to about one-fifth of a school year (or about eight weeks, the length of the school closure time). In contrast, prior cohorts demonstrated growth in student performance during their respective timeframe during the school year. In short, the study found that emergency remote instruction was not effective. It seems reasonable to assume this form of instruction was also not

effective in other countries. It is for example widely documented that in March of 2020 the United States abruptly engaged in “shut-downs” and school buildings were closed quickly enough that teachers had almost no time to plan for needed instructional changes. This event was problematic enough for all students and their families, but whereas some households could at least afford things like broad-band Internet access and tutoring, others did not. Empirically speaking, [Lewis et al. \(2021\)](#) found that in the United States, students in Grades 3–8 demonstrated learning gains in 2020–2021 but not as much compared to learning typically observed in pre-pandemic trends. Furthermore, students in minority groups performed lower than their majority-group peers, and students attending high-poverty schools performed worse than students attending schools in communities with more resources.

One macrosystem level approach toward addressing widespread learning disruption in the United States is the *American Rescue Plan Act (ARP)*, which is part of a major financial stimulus package that among other goals supported schooling. Among the various provisions in the plan, states were given in March of 2021 funding to begin new forms of summer enrichment and learning programming. ARP support should continue through 2024. It is hoped that summer programming and other out-of-school time enrichment experiences will serve to ameliorate the pandemic-related learning disruptions, especially among more vulnerable children. This provides critical context for the survey project presented in this chapter that is informed in part by mixed methods thinking.

It should not be surprising that we think it a challenge to understand the summer programming offerings delivered across a diverse nation and how to generate findings that school district personnel might use. There is a summer learning evidence base (e.g., [McCombs et al., 2021](#); [Peterson and Vandell, 2021](#); [Schwartz et al., 2018](#)) but plans and strategies have been changing at an extraordinary pace to address this ever-evolving pandemic context. One key challenge is timing, and the authors of this chapter are leading a survey effort to understand the summer programming offered by school districts just months after passage of the ARP. The effort entails obtaining a nationally representative sample of districts and presenting findings in time to help inform decisions for Summer 2022, which creates the sense of urgency underlying this project that in turn shortened the project timeframe. The project, the NSLES, is also a mixed methods endeavor because parallel efforts are underway to analyze state and district documents that describe summer programming. Furthermore, the project is carrying out an extensive set of interviews that will contextualize and extend survey findings.

Mixed methods design decisions applied to NSLES

The remainder of this chapter is structured by (a) summarizing widely used and longstanding ideas in survey design and instrument development coupled with formally mixing in qualitative design elements to yield a mixed methods survey project and (b) describing how these ideas were and were not applied to the NSLES. Survey projects, and the context in which a survey is delivered can begin to be understood by thinking through four sources of survey error ([Dillman et al., 2014](#)), and the degree to which such error will invalidate findings. These sources of error are summarized in [Table 2](#).

As noted in [Table 2](#), the four sources of error include coverage, measurement, nonresponse, and sampling error.

Table 2 Sources of survey error.

Source of error	Description
Coverage	Researchers develop a sampling frame from which to draw a sample. The degree to which a given frame does not adequately cover (i.e., capture) the target population of interest yields a form of error. For example, the NSLES target population is all U.S. public school districts. ^a The degree to which the sampling frame NSLES used does not cover all U.S. school districts is the degree to which the researchers assume there is error undermining estimates observed from a randomly drawn sample. Fortunately, NSLES was able to use the National Assessment of Education Progress school district frame, which is maintained by the National Center of Education Statistics housed in the U.S. Department of Education, ^b meaning that the sampling frame has excellent coverage, so this form of error does not represent a serious threat to the validity of the study findings.
Sampling	This form of error is dictated by sample size; it deals with precision, or the width of a confidence interval around a sample estimate that should be described because observations are based on sample statistics and not population parameters. Sampling error decreases as sample sizes increase. NSLES drew a random sample of 550 districts. Dillman et al. (2014) offer a sampling formula (p. 56) for anyone wishing to calculate precision estimates. Suffice it to say for this chapter that a 60% response rate will yield adequate sample sizes assuming a 5% confidence level.
Measurement	This form of error deals with psychometric properties and the degree to which survey items properly cover the constructs of interest. NSLES needed to field an instrument approximately six months after passage of the ARP and no time was available for prior qualitative inquiry and pilot testing (details below). The survey was developed with the aid of a technical working group (TWG) with considerable content expertise. The implications of NSLES timing for mixed methods survey design are reviewed at several points in this chapter. Broadly speaking, prior qualitative inquiry can help item writers be aware of and try to account for matters of context, respondent intersectionality, and tacit assumptions that will influence survey behavior.
Nonresponse	Survey findings are undermined when randomly sampled participants do not respond. Statistical techniques such as nonresponse weighting and data imputation can mitigate this form of error. Strong contextual knowledge can however promote response rates; mixed methods implications are described later in this chapter.

^aThere are approximately 13,452 school districts in the United States ([National Center for Education Statistics, 2020, n.d.](#)).

^bSee: <https://nces.ed.gov/nationsreportcard/>.

Coverage error deals with the degree to which a sampling frame from which a sample is drawn matches up to the actual population of interest. These frames are critical when researchers plan to draw random samples as a basis for arguing that a sample is representative of a target population. In the NSLES, we specified all 13,000+ US public school districts as the target population and drew a stratified random sample of 550 of these districts from the National Assessment of Education Progress data frame. There are however instances where a researcher might opt for purposive sampling (Patton, 2015). By purposive sampling, we mean taking a non-random approach toward identifying respondents (Patton, 2015). Researchers might adopt one of approximately 40 different purposive sampling strategies such as criterion-based sampling (e.g., ensuring the selected sample meets some pre-specific criterion like district size in terms of number of students served), or maximum variation sampling (e.g., ensuring a sample of districts varies in size). Purposive sampling can be useful when an adequate sampling frame is not available, the population of interest is quite small (e.g., when conducting a case study in a school and one surveys teachers), or when the survey focus is not on statistical representation but instead exploring and describing phenomena (e.g., barriers teachers experienced when implementing a new teaching technique in a school). Furthermore, surveys sometimes are used in some qualitative investigations such as ethnographies (e.g., Schensul et al., 1999) and are not necessarily predicated on statistical generalization. In such studies precision estimation is rendered moot although it is still worthwhile to think through the degree to which survey findings generalize within the study context (Onwuegbuzie and Johnson, 2004). In terms of the NSLES, statistical generalization is a goal, so coverage error is evaluated in Table 2.

Sampling error connects to mixed methods survey design thinking because it deals with quantitatively-oriented considerations around sample size and precision of sample estimates. This is however primarily a quantitative concern, so this chapter does not offer detail; readers instead only need to understand that (a) there is an inverse relationship between sample size and sampling error (i.e., sampling error decreases as samples get larger) and (b) even though sampling error decreases as sample size increases, there is a point of diminishing returns such that very large sample sizes (n of 500 or more) are often not needed for reasonable precision (Dillman et al., 2014).

Mixed methods researchers have interconnected measurement error, psychometrics, and qualitative inquiry when it comes to instrument development (see for e.g., Creswell and Plano Clark, 2017; David, 2022; David et al., 2018; Newman et al., 2013; Onwuegbuzie et al., 2010). A fundamental connection between psychometrics and mixed methods can be understood via the social exchange framework, and the aforementioned EST. Social exchange, a concept from the survey literature, posits that the likelihood a person will respond to a survey is informed by the cognitive demand of the instrument and the perceived social benefit of responding to it (Blau, 1964; Dillman et al., 2014). And, germane to this chapter, we suggest that perceived social benefit can be seen as a tacit assumption. In terms of an exchange, as the perceived value of the work increases, so will the effort respondents put into providing survey answers. This means questions need to be interesting and reflect the values of the target population, and it must be evident that the results of the survey will yield a positive social impact if surveyors wish to see respondents persevere through a long list of items or take seriously items that are challenging (as in cognitively challenging, emotionally taxing, or both). Conversely, if a survey entails just a few simple items, then response rates might be high even in absence of a perceived positive social benefit because the effort to respond will be minimal.

Applying EST can help the surveyor think more completely about the survey context

Considering the microsystem level, one might be cognizant about items querying about respondents' capacity to influence an environment, or just as well, any lack of agency they might experience. Again, items need to be interesting and perceived to be of value, and querying people about what they can and cannot do is usually a good start to compelling item writing. The microsystem context also informs a potential respondent's decision to begin a survey and answer all items (i.e., survey behavior), which gets into the matter of nonresponse error. Understanding the lives of target respondents can inform decisions about providing incentives to respond and positioning a survey request that honors respondents' circumstances. The NSLES for example is surveying district superintendents and their delegates. Awareness of the superintendent role as it typically plays out in the United States and how US schools do not always accept emailed survey links means that the NSLES team put considerable resources into multimodal forms of outreach and multimodal ways for respondents to send back data. Personalized letters announcing the survey with descriptions of how study findings might inform Summer 2022 planning, coupled with offering respondents the opportunity to communicate by phone, Email, paper mail, and weblinks are all built into the system because of the broad understanding of superintendent's daily lives. The meso-system level helps surveyors be cognizant that (a) summer programming generally entails community partnerships and (b) interactions of staff occur at all levels of a district. To the first point, considerable attention was put into writing items about district partnering; to the second point, the survey was designed to be shareable in the event a superintendent does not know some given detail about summer programming. Finally, at the exosystem level, one needs to think about the context of a district's capacity to participate in research from both a policy level and at a technology level. For example, from a policy perspective, the NSLES randomly sampled some districts with formal Institutional Review Boards that require the survey team to apply for a response request. An example of being aware of technology context is the team understands that some districts have strict rules around whether emails with hyperlinks can pass through firewalls and there are any number of security and data privacy considerations that can undermine response rates. Awareness of these concerns highlight the importance of context.

Having made these points, the NSLES was able to field a survey without relying on specific mixed methods techniques. Recall from Fig. 1 that the macro context at hand is the NSLES was funded after the Spring 2021 passage of the ARP to study how this federal act influenced Summer 2021 programming. This means the survey had to be conceptualized, written, and tested all in a matter of weeks while also drawing a random sample of districts. More development time could be taken on the back end,

but social exchange in this case is largely predicated on the message that survey responses will be analyzed and disseminated in time to help inform future summer programming. NSLES does however have some unusual assets that made project leaders comfortable in foregoing several mixed methods development techniques. The first is the contracting organization fielding the survey, Westat, has staff with hard won contextual knowledge around district-level surveying, having just completed other survey projects that entailed asking questions of superintendents. A second asset is the project is supported by a panel of experts with considerable summer learning and educational leadership content knowledge and research methodology knowledge. This panel will inform mixed methods analyses that NSLES will pursue (described later in this chapter).

We list several techniques that ideally would have been used in NSLES instrument development had there been more time, representing a trade-off referenced at the outset of this chapter. See [Table 3](#). The techniques in this table include conducting pre-survey interviews, focus groups, and document analyses to learn more about the research context prior to instrument design, developing a Table of Specifications, using Delphi methods, cognitive interviews, pilots and psychometric analyses. Note that citations for follow-up study by interested readers are offered in the table.

Use of any or all of these methods assumes there is not already a survey that is adequate for addressing a set of research questions at hand. Furthermore, use of these techniques assumes researchers enter survey development with a strong command of relevant literatures and therefore understand prior survey efforts, constructs of interest, and how constructs interrelate with each other.

Mixed methods survey design options

Fortunately, there are established mixed methods research designs that can simplify the use of techniques described in [Table 3](#). One such design is the *Exploratory Sequential Design* (see for e.g., [Creswell and Plano Clark, 2017](#)). Creswell & Plano Clark refer to this design as one of their prototypical designs, meaning they can be altered for different purposes. [Fig. 2](#) provides a basic application of this design to survey work.

The figure shows that qualitative data collection and analyses are performed at the outset of the survey project. This can be a highly effective way to understand contextual concerns at all levels captured by EST, likely tacit assumptions held by potential respondents, and some sense of their intersectionality. Furthermore, strong qualitative inquiry can go a long way toward informing research questions to be addressed by a survey, identifying specific populations of interest, considerations in sampling approaches, messaging about a survey, developing survey directions, how to write items, optimal ways to group items within an instrument, survey length respondents will tolerate, and effective ways to disseminate findings. In short, we cannot think of any part of a survey project that would not benefit from prior qualitative inquiry. As mentioned above, this is true even when performing statistical procedures for purposes of psychometrics, missing data imputation, and non-response bias and weighting analyses. This is because statistical decisions and interpretation of analyses can all be better informed by the type of insights typically won from prior qualitative work. We wish we could have applied an exploratory sequential design to NSLES. The fact that we did not apply this design points to the key challenge that [Creswell and Plano Clark \(2017\)](#) recognized; the design does entail considerable time to implement in part because of the potential to develop a new instrument. The NSLES team recognized that several prior qualitatively oriented project phases could not be executed in the available time, so the exploratory sequential design was not used. We do hope that if we are able to pursue future survey work then time will be available to take full advantage to this powerful approach to surveying.

Another prototypical sequential option that could be deployed is the *Explanatory Sequential Design*, also described by [Creswell and Plano Clark \(2017\)](#). See also [Ivankova et al. \(2006\)](#) for an extensive discussion and insights about this design; [Gesser-Edelsburg et al. \(2020\)](#) offer an example of how it has been used. The basic design is depicted in [Fig. 3](#).

The explanatory sequential design can be used in a survey project if the researchers conduct follow-up qualitative inquiry with survey responders to gain more in-depth understanding of phenomena and potentially refine the survey if it will be used again. One could for example purposefully sample respondents who were outliers, as well as respondents who offered more typical survey responses, and ask them questions around why they responded the way that they did, if they have more to say about the phenomena of interest, and so on. This provides clear advantages, but as with the exploratory sequential design, [Creswell and Plano Clark \(2017\)](#) note that it will generally take a relatively long time to deploy.

Incidentally, if time pressures are not a serious concern, then a variant of multiphase sequencing can yield a powerful form of mixed methods inquiry used to address survey-instrument development. Imagine being able to start a research project with a qualitative phase, then a quantitative one, followed up by yet more qualitative inquiry and continuing the process until research objectives are met. The [David et al. \(2018\)](#) study referenced in [Table 3](#) was able to use a multiphase design (interviews were conducted to inform item writing, a pilot sample produced responses subject to item response theory analyses, more interviews were conducted to work out item refinement, psychometric analyses were again conducted). If this sounds like a lengthy process, it was. Like most research endeavors, time (and budget for funded work) needs to be carefully considered whenever sequencing is deployed.

One mixed methods design that can be more time efficient is the *Convergent Parallel Design*. In this approach, the qualitative and quantitative research strands are mostly independent, and findings are mixed during analyses. See [Fig. 4](#).

The NSLES used this approach; the qualitative strand entails document analyses of state level plans to use ARP funding to support summer programming, analyses of district websites that describe summer programming, and interviews with state and district officials. The quantitative strand is the survey of a nationally representative sample of school districts. A key justification for using this design is it minimizes the weaknesses inherent in each component approach. NSLES stakeholders want a statistically representative sample of districts to describe Summer 2021 programming; this will provide a unique set of findings in that we will

Table 3 Mixed methods survey development techniques.

<i>Technique</i>	<i>Description</i>
Pre-survey stakeholder exploratory interviews and/or focus groups	NSLES did not use this approach because of time limitations but if an additional month or two of development time were available the study team would have interviewed a few superintendents experienced with summer program delivery. There are varied approaches for how to conduct interviews and numerous considerations for how extensive this effort should be. Interviews can be exceedingly valuable for better understanding phenomena of interest and how to write survey items that properly cover constructs to be covered by a survey. Our own personal experiences lead us to believe that informal interviewing is quite common, even among survey researchers who are not especially familiar with qualitative or mixed methods research (our position is that if interviews are used, there is merit in taking full advantage of quantitative inquiry to generate item writing). On the other end of the continuum, sometimes extensive interviewing is needed to inform item writing, such as when dealing with abstract constructs (David et al., 2018 for example develop a measure of trust, which is an idea that is hard to concretize in a survey). Hitchcock et al. (2005) conducted extensive interviews prior to item writing to account for cultural factors; a key need because the survey was fielded in Sri Lanka.
Document analyses	NSLES was able to engage in brief document reviews by examining available state ARP plans that described summer programming (full qualitative analyses were not completed until after the survey was fielded); the project director also had prior knowledge of best practice guidance around summer programming. Documents can be thought of as any physical or written artifact (e.g., scientific reports, newspaper articles, marketing materials, websites, video, audio recordings, personal diaries; see Patton, 2015) that will help researchers understand target phenomena. Analyses generally entail developing a coding protocol that is designed to extract relevant data for a multitude of coding options (see for e.g., Saldaña, 2021).
Mixed methods Table of Specifications (TOS)	The NSLES did not develop a formal TOS because of time limitations but did vet survey items with its expert panel. A TOS is a tool or checklist used to help researchers be sure that the totality of items adequately covers (measures) the phenomena of interest. Newman et al. (2013) developed a mixed methods TOS procedure using techniques like audit trails, expert debriefing, and peer reviews to develop transparent assessments of content validity.
Delphi methods	These methods generally entail multiple rounds of surveying or interviewing stakeholders/experts about some phenomena of interest to obtain feedback or make predictions. Usually, anonymity is afforded in the process because after a first round of feedback participants can see the judgments of others to refine their thinking. Anonymity allows for later changes in opinion without letting others know who changed their mind, and influence is won through force of argument and not the name or position of the person advancing an idea. There are several Delphi Method variants (Patton, 2018); when applied to survey development they can be used to brainstorm items or sometimes judge existing items for purposes of refinement (e.g., Tapio et al., 2011). In the context of NSLES, a Delphi study with the expert panel would have been an excellent exercise, if time allowed.
Cognitive interviews	This process is sometimes referred to think-aloud protocols or cognitive labs. There are different approaches; a common technique is to ask a proxy survey respondent to take a survey and react to it verbally to help researchers identify if there are confusing wording in directions or items, identify mismatches between item stems and response options, and so on. Willis (2005) authored a book on using interviews to improve survey design. NSLES would have ideally conducted cognitive interviews with school superintendents.
Piloting	The benefits of piloting should be obvious. If time allows, identify 30 or so proxy sample members and ask them to take the survey and provide feedback on any and all aspects of the tool. Johanson and Brooks (2010) found that a sample of about 30 participants will be adequate for obtaining Cronbach alpha estimates, which can be used to assess internal consistency of any scales and if any items reduce the size of these estimates (which would mean they might be problematic and should be deleted or rewritten). Hitchcock et al. (2015) provides other examples that demonstrate the benefits of piloting. The NSLES would have used piloting to refine the survey if not for time constraints.
Mixed methods psychometrics	Methods for assessing reliability and various forms of validity can be facilitated by relying on qualitative techniques and mixing them into psychometric work. David et al. (2018) for example conducted a series of focus groups to understand how to write items about trust, piloted an initial survey, assessed item fit using item response theory analyses, and conducted follow-up interviews to refine items that were determined to be problematic. This process was used across multiple iterations until the survey was finalized. Hitchcock et al. (2005) demonstrate how to combine exploratory factor analysis and qualitative thematic analysis as part of a mixed methods construct validation exercise. NSLES will over time be able to conduct some of these analyses after descriptive findings are made available to the public.

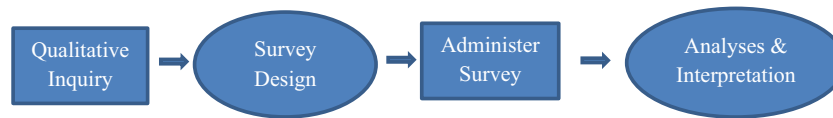


Fig. 2 The basic exploratory sequential design applied to survey work.

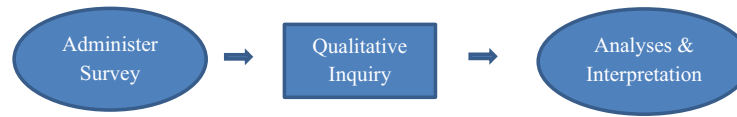


Fig. 3 The basic explanatory sequential design applied to survey work.

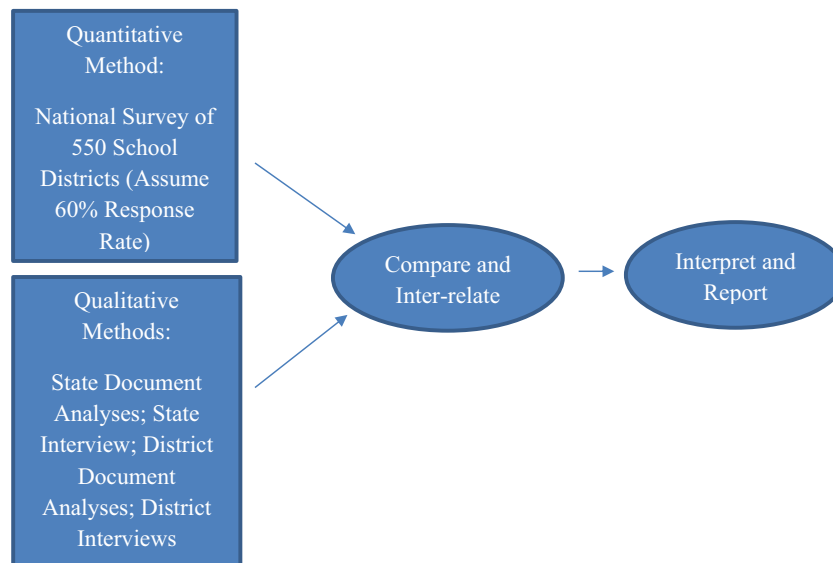


Fig. 4 The NSLES convergent parallel design.

be able to generalize results to a national level and understand how federal investment was deployed to combat mass learning disruptions in primary and secondary schooling. But of course, surveys are not designed for highly in-depth and exploratory inquiry. Hence, the interviews and document analyses of a smaller purposive sample will provide us with more in-depth understanding.

Data analyses and interpretation

Given the variety in design options there is no one way to pursue analyses in mixed methods survey projects. Three broad pieces of advice are: (a) to follow best practices in component analyses (i.e., quantitative analyses and qualitative analyses), (b) be sure to account for all data sources during analyses, and (c) find optimal ways to mix information when interpreting findings. In the case of NSLES, our survey analyses largely will be descriptive in nature because this fits the nature of the research questions (e.g., what percent of school districts in the United States offered summer programming in 2021? What percent of districts plan to offer summer programming in 2022?). The analyses will however entail some exploratory prediction modeling. When conducting analyses of survey data best practices are to evaluate and account for nonresponse bias, perform outlier analyses (i.e., the accuracy of and understand any extreme responses noted in the sample) and clearly describe the degree of precision through presenting confidence intervals when deriving sample estimates. Any prediction modeling will account for modeling assumptions and our audience will be informed of project limitations and threats to the validity of findings pertaining to our understanding of Summer 2021 programming in the United States.

As of this writing all qualitative data collection is pending. Our qualitative coding generally will follow [Saldaña's \(2016\)](#) straightforward process of (a) identifying all available raw data (interview data, document data), (b) collating datum points into codes, (c) categorizing codes, (d) developing themes and concepts from the codes, and (d) generating findings and assertions. Coding procedures will be vetted with our expert panel. Since we will adopt an exploratory position, we will reserve the right to revisit analyses through different forms of second-cycle coding. To explain, Saldaña identified multiple styles of coding wherein analysts can focus on identifying, for example, data patterns, change over time (longitudinal coding; respondents might describe programming

Table 4 Summary of credibility techniques to be applied to NSLES.

<i>Credibility technique</i>	<i>Description</i>
Member checks	We will ask some interview respondents to check the accuracy of our interpretations and conclusions. Did we capture their intended meaning? Member checks can also support document analyses. We will consider the question with some respondents: did we capture and correctly interpret key details about summer programming?
Triangulation	This entails searching for and interpreting consistency across different data sources (i.e., survey, documents, interviews), as well as any inconsistency (e.g., if a survey respondent claims a district offered little summer programming but interviewees describe robust programs, then we need to address the inconsistency to develop credible findings).
Negative case analyses	These analyses entail a purposeful search for discrepant evidence to challenge and test initial assumptions and findings.
Peer debriefing	We will leverage our expert panel to review findings and provide critical feedback.
Audit trail	We will keep records of all research activities (data collection, dates for key decisions, document choices to not pursue some line of inquiry but identify it as being worthy of future study) to substantiate claims about findings.

prior to the ARP, Summer 2021 programming, and future plans), and coding that helps us develop a theory that summarizes how district and state personnel think about summer programming as an intervention for addressing disruption in student learning.

Following best practices, we will deploy different credibility techniques (see for e.g., [Brantlinger et al., 2005](#); [Nastasi and Schensul, 2005](#); [Patton, 2015](#)) including member checks, triangulation, negative case analyses, and so on. The techniques the NSLES will use are summarized in [Table 4](#).

Note that [Table 4](#) does not list all credibility techniques used in qualitative and mixed methods research; it instead highlights key approaches we plan to use in NSLES as a part of our intended qualitative analyses.

Our mixed analyses will at a minimum rely on joint display of findings and profile analyses. A joint display is an analytic tool (i.e., developing a joint display is an analytic process) that presents both quantitative data (for NSLES, survey findings) and qualitative data (NSLES themes and assertions) to allow for direct comparison, merging, and mixed triangulation (see for e.g., [Creswell and Plano Clark, 2017](#); [Fetters, 2020](#); [Guetterman et al., 2015](#); [Haynes-Brown and Fetters, 2021](#)). We cannot as of this writing provide a visual example of our planned joint display because data have yet to be collected and we will adopt an exploratory analytic approach, so we do not yet know what story we will tell. [Haynes-Brown and Fetters \(2021\)](#) however offer straightforward guidance on joint display development. In summary, Haynes-Brown & Fetters would counsel us to consider: (a) how the joint display supports our analyses as we compare, contrast, and merge findings from different data sources, (b) ways to present a complex array in a user-friendly manner so that we can describe findings, (c) incorporate theory and understanding of context to explain our interpretations, and (d) select the most descriptive of available text data.

Profile analyses ([Tashakkori and Teddlie, 1998](#)) will exploit the fact that we will purposefully select some districts that responded to our survey and meet maximum variation criteria pertaining to Summer 2021 offerings. Specifically for purposes of profile analyses, we will purposefully select districts that:

- (a) offered extensive programming and there is evidence the district has advanced experienced with summer learning;
- (b) offered typical forms of summer programming and the district is not a particular outlier in the dataset in terms of experience; and
- (c) offered minimal programming and has a history of doing so.

We will analyze select districts in these profiles and compare and contrast their Summer 2021 circumstances to identify lessons learned from advanced programming, and to understand barriers to programming. We hope our analyses will yield insights as to how these barriers might be addressed in future summers with the support of ARP funds.

The combined use of joint displays and profile analyses that are supported by high quality component analyses should position us to generate defensible findings and varied forms of user-friendly reports. At this point it helps to revisit the idea that mixed methods survey work can and should take full advantage of contextual knowledge to inform report dissemination as one way to honor social exchange. We plan to disseminate information as fast as possible through venues supported by the U.S. Department

of Education and websites visited by school superintendents in the United States (e.g., the School Superintendent Association) and topically aligned venues such as the National Summer Learning Association. NSLES procedures and findings will eventually be presented in peer-reviewed journal outlets, but we know our context and audience as a function of the project being led by a person with deep content expertise and an expert panel that understands the summer learning and enrichment field. This knowledge compels us to present findings with rapidity and widely to promote the greatest possible reach of this work. Other mixed methods survey projects might have the benefit of more time at the outset of an endeavor and/or after inquiry is completed but might also not have access to a broad swath of experts who enter a project with a strong understanding of context. In these cases, adopting an exploratory sequential design will almost certainly lead to a successful project because this design allows researchers to better understand the research context through initial qualitative inquiry.

Conclusion

A key take home point from this chapter is mixed methods thinking can interweave through all stages of a survey project, from initial conceptualization, formulation of research questions, specifying a target population and sampling frame, sampling decisions, survey design and item writing, analyses and interpretation, reporting and, finally dissemination. Importantly, variants of the advice and ideas presented in this chapter can be found in survey resources provided by major publishers and have informed the field across multiple book editions (e.g., Dillman et al., 2014; Fowler, 2013; Groves et al., 2009). However, key contributions of a mixed methods way of thinking come from taking full advantage of formal qualitative inquiry and mixed analysis approaches that can help one generate instruments and understand findings from quantitative, qualitative, and mixed lenses. We have emphasized the importance of using qualitative inquiry to understand the context in which a survey will be fielded and to help account for respondent intersectionality, people's tacit assumptions, and context that in turn will influence survey behavior. In the example we provided, the NSLES, initial qualitative inquiry was not possible because of time limitations, which is something to consider whenever thinking about a design with sequenced phases. Although NSLES is supported by staff with deep content expertise who are embedded in national summer learning ecosystems, and operational staff who have repeatedly surveyed school districts for national projects, we still would have adopted an exploratory sequential design had time allowed. Doing so would have given us the chance to interview stakeholders and documents before item writing and we could have further deployed techniques described in Table 3 of this chapter.

Despite this limitation of not being able to use mixed methods procedures for survey-instrument design, the NSLES will take advantage of mixed methods procedures using qualitative data strands (i.e., document/website analyses and interviews). *The key take-home point to appreciate here is a mixed methods survey project can exploit disparate methods to minimize weaknesses in component design approaches.* Surveys can be delivered to large sample sizes and utilize sampling technology to help researchers have some assurance that findings are generalizable to a target population. This might be especially important for large-scale endeavors such as mobilizing national efforts to address a pandemic-related learning disruption. It is also true, however, that at least compared to qualitative inquiry, surveys are weak with respect toward deeply exploring topics. It is hard to imagine how even the most sophisticated of surveys will include all needed items, written in an optimal way. Interviews and document analyses will offer much better chances to explore phenomena. But of course, the weakness of these approaches is they are intensive and time consuming. We cannot interview staff from all randomly sampled districts in time to offer the kind of information we think the field will need to inform Summer 2022 programming. Hence, although the NSLES is an imperfect example of a mixed methods survey project, readers will hopefully see the power that mixed methods can bring to inquiry through this concept of minimizing weaknesses of component designs. This weakness minimization is not addressed just by collecting different forms of data, the notion of mixing is crucial during analyses. In the case of NSLES, the combination of joint displays and profile analyses will create in-depth insights and elaborations around parameter estimates that should generalize to all public-school districts in the United States. Hence, even though mixed methods surveys can be time consuming, they need not be under the right circumstances and hopefully readers of this chapter will be inspired to deploy some of the design ideas described here when they pursue their own survey work.

References

- American Psychological Association, 2017. Multicultural Guidelines: An Ecological Approach to Context, Identity, and Intersectionality. Retrieved from. <http://www.apa.org/about/policy/multicultural-guidelines.pdf>.
- Blau, P.M., 1964. Exchange and Power in Social Life. Wiley.
- Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., Richardson, V., 2005. Qualitative studies in special education. *Except. Child.* 71 (2), 195–207. <https://doi.org/10.1177/00140290507100205>.
- Bronfenbrenner, U., 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage.
- David, S.L., 2022. Rethinking integration in the Mixed IRT Framework: advancing research in athletic training. In: Hitchcock, J., Onwuegbuzie, A. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge.
- David, S.L., Hitchcock, J.H., Ragan, B., Brooks, G., Starkey, C., 2018. Mixing interviews and Rasch modeling: demonstrating a procedure used to develop an instrument that measures trust. *J. Mix. Meth. Res.* 12 (1), 75–94. <https://doi.org/10.1177/1558689815624586>.
- Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. *Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method*, fourth ed. John Wiley & Sons Inc.
- Enders, C.K., 2010. *Applied Missing Data Analysis*. Guilford Press.

- Engzell, P., Frey, A., Verhagen, M.D., 2021. Learning loss due to school closures during the COVID-19 pandemic. *Proc. Natl. Acad. Sci. U. S. A.* 118 (17). <https://doi.org/10.1073/pnas.2022376118> e2022376118.
- Fetters, M., 2020. Performing fundamental steps of mixed methods research data analysis. In: *The Mixed Methods Research Workbook*. Sage.
- Fowler, F.J., 2013. *Survey Research Methods*, fifth ed. Sage. <https://us.sagepub.com/en-us/nam/book/survey-research-methods-4>.
- Gesser-Edelsburg, A., Cohen, R., Shahbazi, N.A.E., Hijazi, 2020. A mixed-methods sequential explanatory design comparison between COVID-19 infection control guidelines' applicability and their protective value as perceived by Israeli healthcare workers, and healthcare executives' response. *Antimicrob. Resist. Infect. Control* 9, 148–155. <https://doi.org/10.1186/s13756-020-00812-8>.
- Goldstein, D., Parlapiano, A., August 7, 2021. The kindergarten exodus. *The New York Times*. <https://www.nytimes.com/2021/08/07/us/covid-kindergarten-enrollment.html>.
- Greene, J.C., 2007. *Mixed Methods in Social Inquiry*. Jossey-Bass.
- Greene, J.C., 2008. Is mixed methods social inquiry a distinctive methodology? *J. Mix. Methods Res.* 2 (1), 7–22. <https://doi.org/10.1177/1558689807309969>.
- Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R., 2009. *Survey Methodology*, second ed. Wiley.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13 (6), 554–561.
- Haynes-Brown, T.K., Fetters, M.D., 2021. Using joint display as an analytic process: an illustration using bar graphs joint displays from a mixed methods study of how beliefs shape secondary school teachers' use of technology. *Int. J. Qual. Meth.* <https://doi.org/10.1177/1609406921993286>.
- Hitchcock, J.H., Nastasi, B.K., Dai, D., Newman, J., Jayasena, A., Bernstein-Moore, R., Sarkar, S., Varjas, K., 2005. Illustrating a mixed-method approach for validating culturally specific constructs. *J. Sch. Psychol.* 43, 259–278. <https://doi.org/10.1016/j.jsp.2005.04.007>.
- Hitchcock, J.H., Onwuegbuzie, A.J., Khoshaim, 2015. Examining the consequential validity of standardized examinations via public perceptions: a review of mixed methods survey design considerations. *Int. J. Mult. Res. Approaches*. <https://doi.org/10.1080/18340806.2015.1076757>.
- Ivankova, N.V., Creswell, J.W., Stick, S.L., 2006. Using mixed-methods sequential explanatory design: from theory to practice. *Field Meth.* 18 (1), 3–20. <https://doi.org/10.1177/1525822X05282260>.
- Johanson, G.A., Brooks, G.P., 2010. Initial scale development: sample size for pilot studies. *Educ. Psychol. Meas.* 70, 394–400. <https://doi.org/10.1177/0013164409355692>.
- LeCompte, M.D., 2000. Analyzing qualitative data. *Theor. Pract.* 39 (3), 146–154. https://doi.org/10.1207/s15430421tp3903_5.
- Lewis, K., Kuhfeld, M., Ruzek, E., McEachin, A., 2021. Learning During COVID-19: Reading and Math Achievement in the 2020–21 School Year. NWEA. [Learning-during-COVID-19-Reading-and-math-achievement-in-the-2020-2021-school-year-research-brief-1.pdf](https://www.nwea.org/learning-during-covid-19-reading-and-math-achievement-in-the-2020-2021-school-year-research-brief-1.pdf) (nwea.org).
- McCombs, J.S., Augustine, C.H., Unlu, F., Zioli-Guest, K.M., Naftel, S., Gomez, C.J., Marsh, T., Akinniranye, G., Todd, I., 2021. Investing in Successful Summer Programs: A Review of Evidence Under the Every Student Succeeds Act. RAND Corporation. Available. https://www.rand.org/pubs/research_reports/RR2836.html.
- Nastasi, B.K., Schensul, S.L., 2005. Contributions of qualitative research to the validity of intervention research. *J. Sch. Psychol.* 42, 177–195. <https://doi.org/10.1016/j.jsp.2005.04.003>.
- National Center for Education Statistics, 2020. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d20/tables/dt20_214.10.asp.
- National Center for Education Statistics, (n.d.). National Assessment for Education Progress. https://nces.ed.gov/nationsreportcard/dw/sample_design/.
- Newman, I., Lim, J., Pineda, F., 2013. Content validity using a mixed methods approach: its applications and development through the use of a table of specifications methodology. *J. Mix. Methods Res.* 6 (3), 243–260. <https://doi.org/10.1177/1558689813476922>.
- Onwuegbuzie, A.J., Bustamante, R.M., Nelson, J.A., 2010. Mixed research as a tool for developing quantitative instruments. *J. Mix. Methods Res.* 4 (1), 56–78. <https://doi.org/10.1177/1558689809355805>.
- Onwuegbuzie, A.J., Collins, K.M.T., Frels, R.K., 2013. Foreword: using Bronfenbrenner's ecological systems theory to frame quantitative, qualitative, and mixed research. *Int. J. Mult. Res. Approaches* 7, 2–8. <https://doi.org/10.5172/mra.2013.7.1.2>.
- Onwuegbuzie, A.J., Johnson, R.B., 2004. Mixed method and mixed model research. In: Johnson, R.B., Christensen, L.B. (Eds.), *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. Allyn & Bacon, pp. 408–431.
- Onwuegbuzie, A.J., Leech, N.L., 2021. Qualitizing data. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 141–150.
- Patton, M.Q., 2015. *Qualitative Research and Evaluation Methods*, fourth ed. Sage.
- Patton, M.Q., 2018. Building a community of practice through adaptation and integration of multiple methods and diverse approaches to evaluative inquiry. *Int. J. Mult. Res. Approaches* 10 (1), 183–198. <https://doi.org/10.29034/ijmra.v10n1a12>.
- Peterson, T.K., Vandell, D.L., 2021. The Evidence Base for Summer Enrichment and Comprehensive Afterschool Opportunities. Collaborative Communications. Available. <https://spark.adobe.com/page/synw23DwBdPms/>.
- Saldaña, J., 2016. *The Coding Manual for Qualitative Researchers*, third ed. Sage.
- Saldaña, J., 2021. Coding techniques for quantitative and mixed data. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 151–160.
- Sandelowski, M., Voils, C.I., Knaf, G., 2009. On quantizing. *J. Mix. Methods Res.* 3, 208–222. <https://doi.org/10.1177/1558689809334210>.
- Schensul, S., Schensul, J., LeCompte, M., 1999. *Essential Ethnographic Methods*. Altamira Press.
- Schwartz, H.L., McCombs, J.S., Augustine, C.H., Leschitz, J.T., 2018. Getting to Work on Summer Learning: Recommended Practices for Summer Success, second ed. RAND Corporation. Available. <https://www.wallacefoundation.org/knowledge-center/Documents/Getting-to-Work-on-Summer-Learning-2nd-ed.pdf>.
- Tapio, P., Paloniemi, R., Varho, V., Vinnari, M., 2011. The unholy marriage? Integrating qualitative and quantitative information in Delphi process. *Technol. Forecast. Soc. Change* 78, 1616–1628. <https://doi.org/10.1016/j.techfore.2011.03.016>.
- Tashakkori, A., Teddlie, C., 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. In: *Applied Social Research Methods Series*, vol. 46. Sage.
- Willis, G.B., 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Sage.

Mixed methods in culturally tailored interventions

Bonnie Kaul Nastasi, Tulane University, New Orleans, LA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	678
Foundations	678
Conceptual	678
Methodological	679
Social-cultural-political	679
Culturally tailored interventions	680
Using mixed methods in PCSIM	681
Facilitating participatory action research for capacity building	681
Examples of MMR for culturally tailored interventions	682
Conclusion	683
References	684

Introduction

This discussion of mixed methods research to inform culturally tailored interventions is situated within several developments in education and psychology: *Conceptually* within an ecological systems approach to program development and evaluation; *methodologically* within implementation science, evidence-based practice, and translational research; *socially-cultural-politically* within efforts to address cultural diversity in science and practice. Addressing cultural diversity is especially critical to development of interventions within the international realm, given the limitations of current research evidence based predominantly on populations from North America and Western Europe (Arnett, 2008). Moreover, the limitations of current evidence have implications for addressing diverse cultural needs within domestic contexts (e.g., US communities). Mixed methods hold promise for understanding the role of “culture and context when one attempts to influence, or assess, the behavior, thought processes, and perceptions of others” (Nastasi and Hitchcock, 2016, p. 2, italics in original). In this chapter, we (a) explore the conceptual, methodological, and social-cultural-political foundations for culturally tailored interventions; (b) examine the role of mixed methods in development and evaluation of culturally tailored interventions, using a participatory approach; (c) provide examples from international literature; and (d) conclude with implications for educational research and practice.

Foundations

Conceptual

Bronfenbrenner’s (1998, 1999) ecological systems theory (EST) provides a framework for addressing *culture* and *context* in development and evaluation of culturally tailored interventions. Briefly, culture refers to “a dynamic system of meanings, knowledge and actions that provides actors collectively, interpersonally, and individually with community-legitimized strategies to construct, reflect upon, and reconstruct their world and experience, and guide behavior” (Nastasi et al., 2015, p. 96).

Context, or the *microsystem* within EST, refers to the immediate setting in which experiences occur (e.g., classroom, playground). The microsystem is embedded in a larger *exosystem* (e.g., the school), which is embedded within a social-cultural-political-economic system, referred to as the *macrosystem* (e.g., the larger community/society). Understanding or influencing experiences within a given microsystem (e.g., classroom) also necessitates consideration of other microsystems relevant to the individuals (peer group, family, neighborhood, etc.). In addition, the experiences of the teacher and other students within the classroom are relevant to understanding classroom dynamics. The interaction of the components of the ecosystem (micro-, exo-, macro-) and across ecosystems (e.g., school, family) is referred to as the *mesosystem*. Moreover, functioning in the immediate context is influenced by past experiences (i.e., the *chronosystem*). Finally, the interactions within an ecosystem are dynamic, such that the individuals within a microsystem are active agents and influence the system.

Applying EST to intervention research and development requires a mixed methods formative inquiry step to explore the influential contextual and cultural factors within the ecology. For example, to facilitate school connectedness, the researchers must first examine the perspectives of all key stakeholders (students, teachers, school administrators, other school staff, parents/caregivers) about relationships, feelings of safety and danger, conditions or interactions which facilitate or inhibit healthy interactions. This inquiry should include data relevant to culture, by gathering from all stakeholders, the values, beliefs, norms, and language related to key constructs. In addition, researchers would need to explore school and district policies that govern teacher-student and student-student interactions, such as policies related to discipline practices, bullying, social-emotional learning, peer relationships, etc. Data can be gathered using observations, interviews, focus groups, survey instruments, secondary data, and records review. Finally, researchers would need to examine points of conflict or agreement across stakeholder groups or context related to positive school relationships and safety. These data could then guide planning of interventions to facilitate school connectedness.

The EST can guide examination of the immediate context and factors that influence interactions in that environment. This understanding is essential to development and evaluation of interventions, especially those that attempt to address culture and context (i.e., are culturally tailored). The complexity of EST necessitates research and evaluation methodology that can account for experiences of individuals, influential cultural and contextual factors, and the dynamic nature of human and systemic interactions. In this chapter, we explore the role of mixed methods in development and evaluation of interventions embedded within ecosystems, in this case specific to educational settings (classrooms, schools, community).

Methodological

The impetus for considering the complexity of intervention development and evaluation is the recognition of the limitations of existing approaches that rely on research conducted in highly controlled settings (e.g., randomized controlled trials, RCTs), often without consideration of multiple factors within the immediate context and larger ecosystem. These limitations became apparent as practitioners attempted to implement RCTs and met challenges in adapting the interventions to contextual and cultural variations. Recognition of the limitations of RCTs led to developments known as evidence-based practice, implementation science, and translational research.

The call for *evidence-based practice (EBP)* came from psychology (APA, 2006) and medicine (Institute of Medicine, 2001). Briefly, EBPs require the integration of existing intervention research, the expertise and experience of the implementer (e.g., clinician, teacher), and consideration of culture and context (see Nastasi and Hitchcock, 2016, for further discussion).

Implementation science (IS) extended the study of interventions within highly controlled settings in recognition of the importance of contextual and cultural factors on implementation of evidence-based interventions by researchers and interventionists from multiple disciplines (psychology, education, medicine, public health). IS requires the study of myriad factors that influence development and evaluation of interventions and can inform evidence-based practice (Eccles and Mittman, 2006; Fixsen et al., 2005; Moullin et al., 2019).

Similar to implementation science, *translational research* is concerned with the implementation of evidence-based interventions in natural settings, and requires the study of factors that influence acceptability, integrity, ecological validity, adaptation, and effectiveness as practitioners attempt to apply interventions supported by existing research. Translational research thus involves the examination of intervention implementation *in vivo*, as practitioners attempt to implement and adapt evidence-based interventions and requires ecological or systems thinking and participatory approaches (Silver et al., 2021; Scarpa and Trickett, 2022).

EST provides a conceptual framework for engaging in EPB, implementation science, and translational research. That is, the study of the ecosystem in which an intervention occurs can lead to the identification and monitoring of influential contextual and cultural factors, and thus help to facilitate effective interventions in natural settings. Quantitative and qualitative methods and their integration through mixed methods research (MMR) are necessary to capture the complexity of the ecosystem to inform development of culturally tailored interventions (Nastasi and Hitchcock, 2016). For example, the formative inquiry related to school connectedness described above might yield data about student perceptions of teachers' discipline procedures as increasing levels of stress and inhibiting coping, thus leading to more aggressive behaviors toward peers. Furthermore, teachers might indicate that the school's discipline policies inhibit developing caring relationships with students and preclude efforts to foster student coping with strong emotions. Data may also reveal high levels of bullying with minimal adult intervention, and lack of opportunities to develop better peer relationships. Data may also reveal inequity related to punishment of students based on race and gender, with black males most likely to receive severe punishment. Further data collection might reveal perceptions of black male students as more aggressive, despite observational data to the contrary. These findings suggest need to address racial and cultural beliefs, to develop better understanding of how to interpret behavior within a diverse population, and to institute efforts to foster better teacher-student relationship with all students. Bell and colleagues (Bell et al., 2017) provide an example (detailed in a later section) of using formative MMR data to develop a culturally and contextually tailored intervention in one school.

Social-cultural-political

Coinciding with calls for better understanding of implementation of science in natural settings was social-political recognition of the importance of addressing diversity within the population at global, national, and local levels. Recognition of the limitations of science based on Western cultures of North America and Western Europe with populations that were predominantly White, middle class, well educated, and represented only 5% of the world's population (Arnett, 2008), led to calls for research and practice that account for diversity and recognition of the need for culturally specific or culturally tailored interventions. Of course, developing such interventions requires first the understanding of the specific culture of the target group within their ecology, and subsequently the monitoring of cultural and contextual factors that can influence effectiveness. Qualitative research methods are ideally suited for in-depth study of culture and context (Nastasi and Hitchcock, 2016). Combined with quantitative methods, researchers can further track changes and outcomes of interventions. In this chapter, we examine the potential role of MMR for engaging in translational research, thereby addressing limitations of existing intervention research and facilitating culturally tailored evidence-based practice.

Culturally tailored interventions

The integration of ecological systems theory, implementation science, and cultural diversity, for the purpose of developing culturally tailored interventions, requires complex mixed methods research designs. *Culturally tailored interventions*¹ refer to interventions that are designed to address the needs of specific cultural group(s) embedded within cultural context(s); for example, designing or adapting an intervention to the needs of a marginalized group (e.g., low-income African American students) within a specific school context (e.g., elementary school in urban neighborhood). *Tailoring* the intervention requires understanding of the experiences of the students within their ecology (e.g., classroom within school located within community and larger society). Understanding the ecology necessitates exploration of cultural norms, values, beliefs, and practices of the classroom and larger school context through observation, interviewing and examination of cultural artifacts. Thus, before the tailoring can occur interventionists must first study the culture of the target population within the target ecology, that is, beginning with study of the ecological system from perspective of stakeholders.

The study of culture is depicted as the second phase (learning the culture) in the *Participatory Culture Specific* (cf. *Culturally Tailored*) *Intervention Model* (PCSIM; Fig. 1; Nastasi et al., 2004). The PCSIM extends a more traditional approach to intervention research by including in-depth research that leads to design, adaptation, and evaluation of an intervention tailored to the cultural group and context. The tailoring is achieved through use of mixed methods research to inform the development of the intervention. Furthermore, the process requires engaging key stakeholders as partners in the research process.

As depicted in Fig. 1, PCSIM begins with examining existing theory, research, practice, and policy (Phase 1); followed by formative MMR to learn the culture and context (Phase 2) and forming partnerships with stakeholders (Phase 3). Once the goal or problem is identified (Phase 4), researchers in partnership with stakeholders conduct formative MMR (Phase 5) to gather data about key constructs of interest (e.g., student mental health, peer relationships) and to identify influential factors (e.g., stressors and supports, cultural differences among peers, lack of school mental health programming). A critical component of Phase 5 is understanding the key constructs and influential factors from a cultural and contextual perspective, through gathering perspectives of stakeholders and studying the ecology. The data collection leads to development of a local model or theory (Phase 6) that will guide program design (Phase 7) and adaptation (Phase 8; e.g., culturally tailored version of existing program). The local theory/model will



Fig. 1 Participatory Culture Specific (cf. culturally tailored) Intervention Model (PCSIM)

The 10-phase model for development of culturally tailored interventions reflects a dynamic and recursive process of mixed methods research, conducted in partnership with key stakeholders, to inform program design, implementation, and evaluation. The goal of PCSIM is culturally and contextually tailored interventions that are acceptable, socially valid, and sustainable. The *light blue* boxes depict the phases of traditional program development; the *dark blue* boxes indicate the additional phases within PCSIM. From Nastasi, Moore, Varjas (2004, p. 54). Copyright 2004 by the American Psychological Association. Adapted with permission. The use of APA information does not imply endorsement by APA.

¹Also referred to as culturally specific, culturally relevant, culturally sensitive, culturally appropriate (Nastasi et al., 2015).

likely reflect existing theory, research, practice and policy, but be tailored to the local culture and context. MMR data collection continues throughout program implementation to facilitate monitoring, further adaptation, and assessing impact; and at program conclusion to examine outcomes (Phase 9). The final phase (10) is focused on ensuring that the system has capacity to continue program efforts and to engage in the process of program development (capacity building), and on translation of the program to other contexts. Although Phase 10 explicitly addresses capacity building, the researchers with stakeholders will likely focus on building capacity throughout the PCSIM process.

Traditional program design is reflected in the light blue phases (1, 4, 7, 9) in the model. Thus, typically interventionists examine existing theory, research, practice, policy (Phase 1) to identify evidence-based interventions (EBIs) often endorsed by a policy body as validated for use in school settings, for the purpose of addressing the identified goal or problem (Phase 4). The interventionists then design, implement, and evaluate the intervention (Phases 7 & 9) based on the identified EBI, often in a standardized or manualized fashion. The chosen EBI may or may not have been validated for the target population and context, and thus is not necessarily culturally tailored. The addition of the dark blue phases (2, 3, 5, 6, 10) facilitates the development, evaluation, and translation of culturally tailored interventions.

One might conceptualize PCSIM as a participatory mixed methods action research process. In the subsequent sections, we examine the use of mixed methods with PCSIM and discuss the role of participation to facilitate an action research process.

Using mixed methods in PCSIM

As depicted in Fig. 1, PCSIM is characterized by an iterative process of *research* $\leftarrow$ $\rightarrow$ *intervention* that occurs at specific phases and across the cycle from conceptualization to translation. Thus, research is intended to inform program design, monitor program implementation, facilitate program adaptation, and evaluate program impact. Nastasi et al. (2010) have characterized the research cycle depicted in Fig. 1 as a synergistic and fully integrated mixed methods design. That is, qualitative and quantitative methods are used in each phase of the cycle, with outcomes of each phase reflecting the mixing or integration of quantitative and qualitative findings. So, for example, researchers might consider both qualitative and quantitative findings from existing research to inform their worldview and conduct mixed methods research to understand the cultural context (Phase 1 & 2). The study of culture and context (Phase 2) using qualitative and quantitative methods is especially critical in understanding the ecological system. For example, in the context of a school, this might involve participatory observations, interviews with key informants (e.g., administrators, lead teachers), focus groups with key stakeholder groups (e.g., teachers, parents, students), and examination of school record data and policies. From this process, one is prepared to form partnerships, identify system needs, and formulate goals (Phases 3 & 4). Mixed methods are again needed for conducting formative research (Phase 5) to fully understand the system needs and contributing factors. Researchers can engage in mixed methods inquiry using qualitative methods (e.g., participant observations, interviews, focus groups) and quantitative methods (e.g., surveys, school record data, community census data) to define key constructs (e.g., school success, student performance, psychological well-being) and identify contributing factors (e.g., school climate, economic factors, poverty, community violence, etc.). Consistent with an ecological approach, research would extend beyond the school to include, for example, community and national data to depict population needs and contributing factors. The integration of the MMR data can then inform development of a model for culturally tailored interventions relevant to the school population and ecosystem (Phase 6) that guides program design (Phase 7). Phase 7 can involve selection and adaption of evidence-based interventions or creation of new culturally and contextually relevant interventions. During implementation, researchers engage in formative evaluation using mixed methods to monitor program acceptability (e.g., for students, parents, teachers, administrators), social validity (cf. cultural fit), integrity (cf. fidelity), and short-term impact (Phases 8 & 9). This formative evaluation may lead to further program adaptation and system support (e.g., teacher coaching, team problem-solving, additional resources), and mixed methods can help to track implementation of adaptations and their impact on program success. Summative evaluation (Phase 9) addresses program outcomes following program completion, and can address impact on participants (e.g., students, teachers) and the system (e.g., changes in policies and procedures), as well as examine the links between process (e.g., implementation practices) and outcomes. Phase 10 addresses program sustainability and capacity building which are dependent the participation of key stakeholders as partners in the PCSIM process. Mixed methods can be used to assess factors relevant to sustainability and capacity, such as stakeholder commitment, resource allocation, infrastructure, staff competencies, and policies.

Facilitating participatory action research for capacity building

Development of culturally tailored interventions is likely to necessitate active participation of key stakeholders; for example, in the case of school interventions—administrators, teachers, parents, students, and perhaps community members. Achieving an in-depth understanding of culture requires capturing the voices and experiences of key stakeholders, and preferably their engagement as partners. Using participatory approaches such as participatory action research (PAR; Friedman and Antal, 2005; Greenwood et al., 1993; Maister et al., 2008), transformative mixed methods (e.g., Mertens, 2007), and PCSIM, researchers can effectively engage stakeholders as partners to create meaningful and sustainable change. Furthermore, capacity building is dependent not only on partnerships but also on facilitating professional development of staff and building necessary infrastructure that may require changes in school policy and procedures and acquisition of additional resources.

Mixed methods research can contribute to sustainable change by first incorporating sustainability as a goal, and secondly, using integrated qualitative and quantitative methods to examine system and stakeholder needs and resources, understand system

dynamics that influence decision making and change, represent voices of all stakeholders, and inform data-based decision making. So, for example, researchers might meet with partners (e.g., representatives from respective stakeholder groups) to discuss long-term goals for the system in the context of the current short-term intervention goals, and then engage in the PCSIM cycle of research $\leftarrow$ $\rightarrow$ intervention to design and implement capacity building plans. This might include training and/or hiring new staff, forming partnerships with community agencies, seeking support from government agencies, instituting new policies and practices, etc. To complete the cycle, MM would be used to evaluate implementation of plans for sustainability and capacity building.

What is necessary to facilitate PAR for systems change? First, identifying partners with vested interests and resources, and being inclusive in forming partnerships by inviting representatives of all groups that will be affected (e.g., students, parents, community members as well as school staff). Second, as part of the PCSIM process, ensuring that partners develop capacity to engage in a continual process of monitoring and change. Thus, the goal extends beyond implementing a culturally tailored intervention to include building system capacity to continue similar processes in the future. Third, facilitating collaborative problem-solving and decision-making among stakeholders; for example, through training of staff and other stakeholders, creating a team process, and promoting culture of collaboration in the organization. Finally, identifying and fostering stakeholder strengths to facilitate empowerment. Creating sustainable systems change requires that stakeholders have the capacity to continue the process of change as context and culture vary or evolve over time. In the next section, we examine examples of the transformative process through participatory MMR for development of culturally tailored interventions.

Examples of MMR for culturally tailored interventions

Examples of mixed methods research and development can be found in international literature. In this section we highlight research and development approaches using participatory MM, consistent with PCSIM, to develop culturally tailored interventions with populations both inside and outside of the United States. The studies summarized here represent school- and community-based work conducted in Botswana, Canada, Hong Kong, Sri Lanka, the US, Korea, and South Africa.

Chilisa and Tsheko (2014) used an indigenous mixed methods approach to develop sexual risk prevention programs for adolescents in Botswana. Program developers engaged in a process consistent with PCSIM, first using research to learn the culture and establish partnerships, then conducting formative and evaluative research using “indigenous mixed methods” (p. 223) as the basis for designing and evaluating interventions that were culturally tailored. A unique aspect was the use of indigenous mixed methods relevant to oral societies, described as “metaphorical sayings, stories, songs, and proverbs serve as the literature, words of wisdom, commentaries on events and behaviors, and socialization instruments for youth and adults” (p. 226). Research methods included qualitative, a combination of indigenous methods (stories, songs, proverbs, etc.) and semi-structured and structured interviews; and quantitative, a culturally tailored (indigenous) survey instrument to measure sexual attitudes, beliefs, and behaviors. The survey instrument was developed based on qualitative data. The initial data collection phase served as basis for designing and evaluating the indigenous intervention, which incorporated indigenous methods (e.g., talking circles), and was co-constructed with stakeholders. Central to research and development within this project were the partnership with community stakeholders, use of indigenous methods for communication and intervention, and inclusion of adolescent voices.

Bird-Naytowhow et al. (2017) describe a participatory community research partnership with Plains Cree and M’etis youth from one urban Canadian environment of Saskatoon, Saskatchewan. The project engaged youth, local elders, and community collaborators as partners in identification of sources of resilience that would inform program development for positive youth development. A basic assumption of the project was that “resources, knowledge, and capabilities required to support the health and wellness of Indigenous youth are already present within inner-city contexts and the young people themselves and thus need only to be collaboratively identified, learned about, and fostered” (p. 2). Youth were primary partners throughout the project. Adopting the view of research as a *ceremonies of relationship* process, researchers integrated photovoice and talking circles as context for interviews with youth. At the conclusion of the project, the youth developed a community exhibit and engaged in interviews with local media as ways to disseminate findings. This project illustrates how researchers can use Indigenous knowledge to facilitate meaningful research for a community. Drawing on local ways of knowing and communicating, researchers can build relationships with local stakeholders and identify culturally relevant (tailored) strategies for effecting social change. The authors described their approach as an example of *two-eyed way of seeing*, combining Indigenous and Western ways of knowing (cf. combining emic and etic perspectives), an approach that is applicable to conducting research with other non-Western populations.

Ling and Pang (2022) conducted a transformative MMR study with Grade 7 ethnic minority (non-Chinese) students from a Hong Kong secondary school. The goal of the study was “to improve the financial literacy of ethnic minority students through financial education, as a step toward ending intergenerational poverty” (p. 133). Researchers used a 3-phase transformative MMR design that included: Phase 1, MMR to facilitate relationship building and contextual analysis, which informed intervention design; Phase 2, intervention implementation with formative qualitative evaluation; and Phase 3, MMR summative/outcome evaluation. Data collection in Phase 1 consisted of surveys and semi-structured interviews; these data informed development of vignettes to be used in the intervention. Phase 2 data consisted of audio-recordings of lessons and written work for qualitative analyses. Phase 3 data were derived from qualitative and quantitative questions of post-evaluation survey. This study illustrates the use of student voices for informing development of educational curriculum, using vignettes based on real, hypothetical, and phenomenologically oriented stories to enhance financial education of ethnic minority early adolescents.

Nastasi and Jayasena (2014) engaged in a multi-year project in Sri Lanka to develop culturally tailored instruments and interventions for promoting psychological well-being of school-age children and adolescents. Using MM across the phases of PCSIM, the

researchers collaborated with 18 schools in Central Province of Sri Lanka to conduct formative ethnographic research (e.g., interviews and focus groups with students, teachers, and school administrators) that provided the basis for development of self-report measures of perceived competence and stress and coping, and development of a culturally tailored intervention program that was implemented and evaluated in one school. The intervention program was later modified to address long-term psychological well-being needs of students, parents and teachers living in the Southern Province who were affected by the 2014 tsunami. The Nastasi and Jayasena article also describes the levels of collaboration that included not only relationships with ministries of education and schools, but also the partnership of the two researchers.

Bell et al. (2017) used a participatory action research approach to develop universal mental health programming in an urban school in the Southern part of the US. The school's action research team (principal, teachers, parents, students, and organizational consultants) participated in decision making throughout the process of formative research, goal identification, and intervention design, implementation, and evaluation. Mixed methods used for formative and evaluative research purposes included focus groups, interviews, observations, and surveys. The integration of qualitative and quantitative data with interpretation by the school action research team informed program design, monitoring, and evaluation. This study illustrates the integration of stakeholder voices with MM data to inform system change.

Kim (2014) used a mixed methods action research approach to understand the context of poverty among Korean children, including relevant policy and values, and to develop, implement, and evaluate a community-based intervention to facilitate transformative life-long learning among students living in poverty. The research was participatory and included voices of elementary and secondary students, parents, instructors, administrators, and community members. The project involved three phases. In Phase 1, researchers conducted a contextual analysis, with qualitative methods to examine legislation related to eliminating generational poverty in Korean and history of local child community centers (CCCs), and quantitative analyses to examine demographics of children served by these centers. Phase 2 involved a 2-year case study of a local CCC; researchers conducted focus groups with adult (center staff, parents, community members) and child (elementary and middle school students) participants in CCC to examine their perspectives on the inhibiting and transformative factors related to perpetuation of inequality and social injustice. Focus group data contributed to development of adult and child surveys to confirm findings with a larger sample. The researchers then conducted a workshop with parents and CCC staff to interpret findings related to inhibiting and facilitating factors and to inform development of transformative programming for children. In Phase 3, mixed methods action research phase, researchers partnered with CCC staff and parents to develop, implement, and evaluate two transformative lifelong learning programs for children. The authors described the process as "community-involved, collaborative and [characterized by] communicative interactions among the participants" (p. 314). This study is an excellent example of a systematic approach to action research, using MMR and participation of key stakeholders throughout the process of program development and evaluation.

Bridwell-Mitchell (2012) used a two-stage sequential, mixed methods design (Qual→Quan) design to explore the role of macro-level societal institutional factors as influencers on teachers' reactions to school reform and their subsequent instructional decisions at the micro-level of classroom (cf. ecological systems theory, EST). Findings from their research, conducted in the north-eastern region of the United States, provided support for EST applied to education, particularly teacher practices. That is, teacher decisions about classroom practices, at the microsystem level, are influenced by both external societal (macrosystem) level factors and internal institutional (exosystem) level factors, necessitating the consideration of the ecological system when trying to understand how school reforms influence classroom practices.

Kapofu (2019) engaged in focused ethnography using mixed methods (observations, interviews with teachers, school artifacts of incidents/events) to understand classroom culture as interpreted by teachers within one school in South Africa. Using semi-structured interviews, the researcher explored teachers' cultural assumptions about pedagogy, power, relationships, and social identity, to understand teaching practices. The study highlighted the critical relationship of the teacher's cultural assumptions as they guide the teaching-learning environment.

These illustrations highlight the integration of MMR with a participatory process for the purposes of (a) understanding culture and context as a necessary step to facilitating culturally tailored interventions; (b) engaging stakeholders in intervention development and evaluation to ensure representation of their voices in decision making; (c) understanding key constructs and influential factors embedded in the cultural context; (d) creating and evaluating culturally tailored interventions; and (e) ensuring capacity building and sustainability. As illustrated in Fig. 1, a participatory action research process extends traditional practices related to selecting, implementing and evaluating evidence-based interventions (depicted in light blue) by incorporating MMR to facilitate learning the culture and forming partnerships; conducting formative research to inform culturally tailored intervention model; conducting evaluation research to facilitate implementation, adaptation, and effectiveness; and facilitate sustainability and capacity building (depicted in dark blue). Existing research, illustrated herein, demonstrates the feasibility and effectiveness of processes such as PCSIM.

Conclusion

Participatory MMR situated within an ecological systems approach holds promise for facilitating implementation science and translational research to address cultural diversity in science and practice. Engaging stakeholders as partners in the process of research and development can help to ensure cultural relevance, ownership, and sustainability. Using MM to study the target culture and context and key variables of interest can help to inform development of a culturally and contextually tailored intervention. Furthermore,

using MM and stakeholder participation in design, implementation and evaluation of interventions can help to facilitate ecological fit, acceptability, integrity, and effectiveness. Finally, MMR in collaboration with stakeholders can facilitate identification of the necessary policies, structures and processes for building capacity within the system to sustain culturally tailored interventions. As highlighted herein, processes such as PCSIM can be helpful not only for advancing implementation science and translational research but also for facilitating transformations in existing practices through the integration of MM with participatory approaches.

References

- American Psychological Association (APA) Presidential Task Force on Evidence-Based Practice, 2006. Evidence-based practice in psychology. *Am. Psychol.* 61 (4), 271–285. <https://doi.org/10.1037/0003-066X.61.4.271>.
- Arnett, J.J., 2008. The neglected 95%: why American psychology needs to be less American. *Am. Psychol.* 63, 602–614. <https://doi.org/10.1037/0003-066X.63.7.602>.
- Bell, P.B., Larrazolo, H.L., Nastasi, B.K., 2017. Promoting universal psychological well-being in an urban U.S. public school using a culture-specific, participatory action research approach to consultation. *Int. J. Sch. Educ. Psychol.* 5 (3), 178–191. <https://doi.org/10.1080/21683603.2016.1276815>.
- Bird-Naytowhow, K., Hatala, A.R., Pearl, T., Judge, A., Sjoblom, E., 2017. Ceremonies of relationship: engaging urban indigenous youth in community-based research. *Int. J. Qual. Methods* 16, 1–14. <https://doi.org/10.1177/1609406917707899>.
- Bridwell-Mitchell, E.N., 2012. The rationalizing logics of public school reform: how cultural institutions matter for classroom instruction. *J. Mix. Methods Res.* 7 (2), 173–196. <https://doi.org/10.1177/1558689812468792>.
- Bronfenbrenner, U., 1989. Ecological systems theory. In: Vasta, R. (Ed.), *Annals of Child Development*, vol. 6. JAI Press, Greenwich, CT, pp. 187–249.
- Bronfenbrenner, U., 1999. Environments in developmental perspective: theoretical and operational models. In: Friedman, S.L., Wachs, T.D. (Eds.), *Measuring Environment Across the Life Span: Emerging Methods and Concepts*. American Psychological Association, Washington, DC, pp. 3–28.
- Chilisa, B., Tshelo, G.N., 2014. Mixed methods in indigenous research: building relationships for sustainable intervention outcomes. *J. Mix. Methods Res.* 8 (3), 222–233. <https://doi.org/10.1177/1558689814527878>.
- Eccles, M.P., Mittman, B.S., 2006. Welcome to implementation science. *Implement. Sci.* 1, 1. <https://doi.org/10.1186/1748-5908-1-1>.
- Fixsen, D.L., Naoom, S., Illase, K.A., Friedman, R.M., Wallace, F., 2005. *Implementation Research: A Synthesis of the Literature*. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231, Tampa, FL. <http://nirn.fmhi.usf.edu>.
- Friedman, V.J., Antal, A.B., 2005. Negotiating reality: a theory of action approach to intercultural competence. *Manag. Learn.* 36 (1), 69–86. <https://doi.org/10.1177/1350507605049904>.
- Greenwood, D.J., Whyte, W.F., Harkavy, I., 1993. Participatory action research as a process and as a goal. *Hum. Relat.* 46, 175–192.
- Institute of Medicine, 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academies, Washington, DC.
- Kapofu, L.K., 2019. Teacher culture and emergent context in two desegregated science classrooms in South Africa: a focused ethnography. *S. Afr. J. Educ.* 39 (2), 8. <https://doi.org/10.15700/saje.v39n2a1581>. Art. #1581.
- Kim, K.H., 2014. Community-involved learning to expand possibilities for vulnerable children: a critical communicative, Sen's capability, and action research approach. *J. Mix. Methods Res.* 8 (3), 308–316. <https://doi.org/10.1177/1558689814527877>.
- Ling, H.L., Pang, M.F., 2022. A vignette-based transformative multiphase mixed methods interventional study featuring Venn diagram joint displays: financial education with Hong Kong early adolescent ethnic minority students. *J. Mix. Methods Res.* 16 (1), 130–149. <https://doi.org/10.1177/1558689821989834>.
- Maiter, S., Simich, L., Jacobson, N., Wise, J., 2008. Reciprocity: an ethic for community-based participatory action research. *Action Res.* 6 (3), 305–325. <https://doi.org/10.1177/1476750307083720>.
- Mertens, D.M., 2007. Transformative paradigm: mixed methods and social justice. *J. Mix. Methods Res.* 1 (3), 212–225.
- Moulin, J.C., Dickson, K.S., Stadnick, N.A., Rabin, B., Arons, G.A., 2019. Systematic review of the exploration, preparation, implementation, sustainment (EPIS) framework. *Implement. Sci.* 14. <https://doi.org/10.1186/s13012-018-0842-6>. Article 1.
- Nastasi, B.K., Hitchcock, J., 2016. *Mixed Methods Research and Culture-Specific Interventions: Program Design and Evaluation*. In: *The New Mixed Methods Research Series*. Sage, Thousand Oaks, CA.
- Nastasi, B.K., Hitchcock, J.H., Brown, L.M., 2010. An inclusive framework for conceptualizing mixed methods design typologies: moving toward fully integrated synergistic research models. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 305–338.
- Nastasi, B.K., Jayasena, A., 2014. International partnership for promoting psychological well-being in Sri Lankan schools. *J. Educ. Psychol. Consult.* 24 (4), 265–282. <https://doi.org/10.1080/10474412.2014.929965>.
- Nastasi, B.K., Moore, R.B., Varjas, K.M., 2004. *School-based Mental Health Services: Creating Comprehensive and Culturally Specific Programs*. American Psychological Association, Washington, DC.
- Nastasi, B.K., Schensul, J.J., Schensul, S.L., Mekki-Berrada, A., Pelto, B., Maitra, S., Verma, R.K., Saggurti, N., 2015. A model for translating ethnography and theory into culturally constructed clinical practice. *Cult. Med. Psychiatr.* 39, 92–109. <https://doi.org/10.1007/s11013-014-9404-9>.
- Silver, H.C., Zinsser, K.M., Zulauf-McCurdy, C.A., Lowe-Fotos, A., Orr, L., Estlund, M., 2021. Transforming early childhood discipline policy through collaborative research and evaluation. *Transl. Issues Psychol. Sci.* 7 (1), 21–34. <https://doi.org/10.1037/tps0000279>.
- Scarpa, M.P., Trickett, E.J., 2022. Translating ecology: similarities and differences in the ecological images of Bronfenbrenner and Kelly. *Transl. Issues Psychol. Sci.* 8 (2), 185–196. <https://doi.org/10.1037/tps0000315>.

Mixed methods and social network analysis

Dominik E. Froehlich, University of Vienna, Vienna, Austria

© 2023 Elsevier Ltd. All rights reserved.

Introduction	685
Defining mixed methods social network analysis	685
Why MMSNA? A brief methodological history of (mixed methods) social network analysis	687
Qualitative beginnings	687
Increasing adoption of quantitative social network analysis	687
Critique and integration of qualitative and quantitative methods	687
MMSNA in educational research	687
Prevalence of MMSNA in education research	688
Methods integrated in MMSNA research designs in education research	688
Current debates	688
Data collection and research designs	688
Ethical considerations	688
Human capital limitations	689
Practical impact	690
Outlook and conclusions	690
References	690

Introduction

The idea of mixed methods, that is integrating distinct qualities of data and different methods of analysis in one study (Creswell and Plano-Clark, 2018; Tashakkori et al., 2020), has been extensively discussed in the domain of educational research (Johnson and Onwuegbuzie, 2004). In this chapter, we will contextualize the idea of mixed methods within the field of social network analysis, a methodological and theoretical framework that highlights the role of social structures and relationships (Borgatti et al., 2009). In other words, we are intersecting the discussions about mixed methods and social network analysis to arrive at Mixed Methods Social Network Analysis (Froehlich and Brouwer, 2021; MMSNA; Froehlich et al., 2020c).

The chapter is structured as follows. First, we will define MMSNA and describe its two “components,” mixed methods and social network analysis. We then briefly discuss methodological advancements in social network analysis from a historical perspective to understand the necessity for MMSNA to study educational phenomena. We then review the current academic discourse around MMSNA. For that, we discuss recent systematic reviews and then zero in on a selection of topics that receive a lot of attention in current methodological texts about MMSNA. We end the chapter with a conclusion and outlook.

Defining mixed methods social network analysis

For rest of this chapter, it will be useful to have a clear understanding of the nature of what we call MMSNA. Therefore, we will now define this methodological approach. We first turn to the individual parts that make up MMSNA: social network analysis and mixed methods.

Social network analysis (Borgatti et al., 2009; Wasserman and Faust, 1994) describes an array of methods and associated theories (Fuhse, 2020) that put an emphasis on social relations and social structures (Wasserman and Faust, 1994). Conceptually, social network data is specific in that cases are (assumed to be) interdependent and at least two pieces of information should be available. First, information about the actors of a network, which are more technically referred to as nodes. Second, information about the relationships (ties) between these actors. These two pieces of information are then used to construct a new entity, the network, which then often becomes the central object of the study. The network is then, in its simplest form, composed of the actors (e.g., pupils, teachers, family members, etc.) connected through various types of relationships (e.g., friendship, collaboration, etc.). The visual representation of such a network may also be called a sociogram; a tabular presentation could be a relationship or adjacency matrix or an edgelist (see Fig. 1; Wasserman and Faust, 1994). In this chapter, we cannot focus more on the basics of social network analysis—which are, of course, a prerequisite to make sense of MMSNA—but other accessible sources exist for the interested reader (Längler et al., 2020; Mejeh, 2020; Scott, 2007).

Mixed methods research can be defined as “research in which the investigator collects and analyses data, integrates the findings and draws inferences using both qualitative and quantitative approaches or methods in a single study” (Tashakkori and Creswell, 2007, p. 4). While variations of this definition exist, for the scope of this chapter it is sufficient to focus on the integration of different

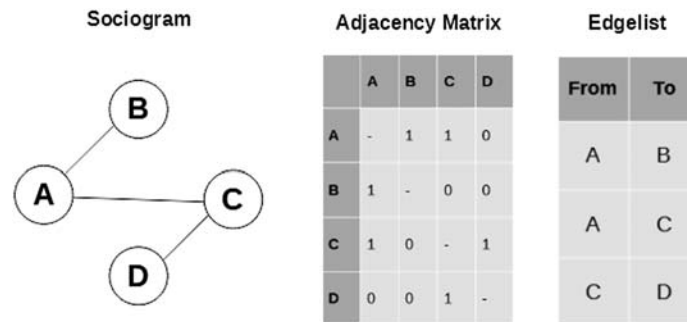


Fig. 1 Representations of a network.

natures of data or vastly different approaches of looking at the data to explore a phenomenon. We just noted “different natures of data” here and not the common qualitative/quantitative distinction, as some social network researchers argue against the usage of the terms “qualitative” and “quantitative” data in the context of social network analysis. This is because the main feature of network data is more about its structural or relational properties. Social network data not only contains information about cases’ attributes (like most statistical datasets that assume independence of cases), but also includes information about how cases are dependent on each other. Indeed, social network researchers often merge two datasets into one to perform network analysis; one describing a set of cases (often referred to as a “nodes table,” which is akin to a conventional statistical dataset) and the other describing the relationships (for instance in the form of an edgelist, which we will discuss below). Since academic training often emphasizes statistical methods, readers may wrongly assume statistical datasets when reading about quantitative data. Nevertheless, in order to improve the clarity of writing, we will refer to qualitative and quantitative methods as a shorthand for qualitative and quantitative structural or relational methods.

In the previous paragraphs, we have discussed SNA and mixed methods as individual elements of MMSNA. But how do the given definitions inform a definition of MMSNA itself? There have been especially two discussions in the literature that should be noted here. *MMSNA in the broadest sense* describes any research design that includes both definitions given above. Note that we interpret this to mean adding mixed methods-thinking to social network analysis, rather than the other way around. This means that not only different natures of data or analytical procedures are included, but that at least one element of the research design features a relational or structural component (Froehlich et al., 2020c). As a simple illustration, we might think of a mono-method quantitative network study, to which we add a qualitative pre-study to develop the measurement instrument (see Schoonenboom et al., 2018 for development as a purpose of mixing methods). This research design includes quantitative and qualitative data and methods and hence can be labeled a mixed methods research design. Also, it features social network analysis. Applying the broad definition of MMSNA, we could, therefore, call it a MMSNA research design.

In a narrower sense that we call mixing *within* social network analysis, we require the use of qualitative and quantitative relational/structural data and/or qualitative and quantitative methods of relational/structural analysis. The example from above would not meet this requirement, as the qualitative pre-study presumably is not about any relational phenomena. As an illustration of a MMSNA design in the stricter sense, we might think of a mono-method qualitative interview-study, in which we seek to explore the social relationships the interviewer is entangled in. Again, we now add another strand of research to further develop this qualitative strand of research: We add a quantitative strand of research to inform the selection of interview partners. We do so by conducting a short network survey in a given organization, from which we extract network metrics that may help us with the sampling process. In sum, the difference to the first example is that both strands of research include a relational aspects.

Here, it is important to note that hardly any qualitative procedures originally developed for social network analysis exist (Franke and Wald, 2006). While many tools (e.g., VennMaker by Kronenwett and Adolphs, 2021) and original qualitative social network research methods (e.g., qualitative structural analysis by Herz et al., 2015) have been developed since the referenced statement by Franke and Wald (2006), this narrow definition would still be *too* narrow to be useful. For this chapter, we stick to the broader definition to be more inclusive.

Next to the narrow and broad definitions given above, it is important to note that SNA has sometimes been described as an inherently mixed method (cf. Onwuegbuzie and Hitchcock, 2015). As argued before, this description makes sense if compared to more “traditional” quantitative methods, as social network analysis adds a new quality to the data through its interdependent cases and mixes attribute data (e.g., the gender and age of pupils) with relational data (e.g., friendship relationships between the pupils). But for the scope of this chapter, this definition is not useful, as it would equate MMSNA with social network analysis, in general. As we point out in the next section, the term “social network analysis” without any further description as “qualitative” or “mixed” often means quantitative social network analysis in the literature. For example, see the definition of Freeman (2004), that explicitly includes the use of mathematical/computational tools as one of four defining factors of social network research.

Why MMSNA? A brief methodological history of (mixed methods) social network analysis

The necessity of MMSNA and the gaps that it fills in contemporary social network research is best understood if taking a brief look at the history of social network analysis. Simplifying the seminal article about the history of social network analysis by Freeman (2004; also see Onwuegbuzie, 2020), we can see three phases, which can be mapped to a predominant methodological approach: qualitative, quantitative, mixed.

Qualitative beginnings

The first phase spans most of the 20th century, when early social network methods were being proposed by Moreno and his colleagues (Moreno, 1934). While these methods of “sociometry” have been predominantly of qualitative nature at this time, it is important to note that social network analysis was not very often applied in absolute terms. Here, the sociogram first appeared as a way to illustrate *abstractions* of relationships to make sense of them and the network they make up (for the role of abstractions, the interested reader may also refer to the Euler’s solution to the “Seven Bridges of Königsberg” (Euler, 1956), which is often regarded as the very first application of social network analysis in 1736). As a sidenote, visualization could be seen as another feature that is prominent both in mixed methods and social network analysis (Shannon-Baker and Edwards, 2018; Shannon-Baker and Hilpert, 2020).

Increasing adoption of quantitative social network analysis

The application of social network methods picked up pace toward the end of the 20th century, when masses of quantitative, relational data became more readily available through advances in technology. Especially the advent of the Internet further increased the volume of studies produced based on social network analysis (e.g., Twitter networks; Brantner and Pfeffer, 2019). In this time, the qualitative roots of social network analysis seem to have been largely forgotten and quantitative methods based on graph theory in mathematics (Barnes, 1983) became the new (and largely unquestioned) standard. With this new method at hand, a host of new research questions could be answered, which further increased the prominence of social network research. It is worth pointing out that the development of this new field of research was largely driven by technological and mathematical advancements and data availability, other questions, such as having solid theories about social networks, were mostly dealt with at a later stage. This also led to the then common perception that social network analysis is void of theory. This perception persisted also in later decades. But with a plethora of original network theories available (Borgatti and Halgin, 2011; Fuhse, 2020; Monge and Contractor, 2003), this statement can nowadays hardly be accepted as true.

Importantly, this was also a time for social network analysis to be increasingly used across academic disciplines. Indeed, there hardly is an academic faculty in which social network analysis and its general and abstract language cannot be applied. In education science, social network analysis was readily adopted to inform a new branch of (social) learning research of, for instance, pupils (e.g., Pearson and Michell, 2000), teachers (e.g., Moolenaar, 2012), or institutions (e.g., Daly and Finnigan, 2010). The adoption of social network analysis in the domain of education science can be sketched through a quick search in the ERIC database: While only 37 social network-related publications were identified for the year 2003, this number rises to more than 400 only a decade later.

Critique and integration of qualitative and quantitative methods

With the rise in popularity across disciplines, criticism against the highly technical and quantitative nature of social network analysis rose, too (Froehlich et al., 2020d). This criticism was first voiced in sociology and was then carried over to the field of education. These are also the two fields in which edited volumes about MMSNA emerged; first in sociology (Domínguez and Hollstein, 2014) and later in education (Froehlich et al., 2020c). Next to MMSNA, also standalone qualitative research designs were increasingly propagated (Herz et al., 2015; Hollstein and Straus, 2006). The general claim was that the quantitative methods cannot adequately capture the complexity of the research questions used in these fields. For example, quantitative social network analysis can give us a lot of insight into how the social structures of pupils in a classroom looks like. But it does tell us little about, for example, how network configurations fluctuate across time, what content is shared between two actors (and why it is shared), what the qualitative consequences of certain positions in the network are for individual pupils, or whether the classroom per se is actually the right network to study for some pupils, who may have found their friends in other contexts. Put succinctly, Crossley’s (2010) concern is “that the quantitative tools of [social network analysis] process the hurly burly of social life in such a way as to create a very abstract, formal and structural mapping [... that] is too abstract, overly formal and insufficiently attentive to inter-agency and process” (p. 2).

MMSNA in educational research

In this section, we dive into the current state of the art of MMSNA research within education science. Specifically, there are three main questions that we seek to answer:

1. How prevalent is MMSNA in empirical educational research?
2. What methods are being integrated in MMSNA designs?
3. What are the methodological topics that are being discussed within this domain?

Prevalence of MMSNA in education research

A recent systematic review by [Froehlich et al. \(2020a,b,c,d,e\)](#) investigated the usage of MMSNA in education science, highlighting not only the prevalence of application but also its purpose for answering the research question, design features, and the quality of mixing. In the end, 52 studies were included in the review. The prevalent contexts of use were K-12 (23 studies, 44%), higher education (18 studies, 35%) and school administration (three studies, 6%). The main *purpose of mixing* consisted in follow-ups (32 studies, 62%), followed by the purpose of comparison (21 studies, 40%) or a combination of these purposes (eight studies, 15%). Four studies (8%) used mixing to develop one method from the other, and another four studies (8%) did not include reasons for mixing. Regarding the *timing of mixing*, eighteen studies (35%) had parallel designs, of which fifteen used methods that were dependent on each other. 34 studies (66%) had a sequential design, with only six using dependent methods.

Many of the MMSNA studies used network maps (also called sociograms) as a way to illustrate the findings, give feedback to participants, or as a basis to develop interventions. Prominent stages where methods were integrated were the sampling stage, the analysis of results, and the discussion/conclusion part. The review also included an assessment of the quality of mixing present in the reviewed studies based on the Mixed Methods Appraisal Tool ([Hong et al., 2018](#)). Common shortcomings were not explicating the reason for using quantitative and qualitative data and/or methods, the lack of integration of outputs (“meta-interference”), and the often missing elaboration of how and why qualitative and quantitative procedures come to different conclusions.

Methods integrated in MMSNA research designs in education research

Another type of systematic review conducted by [Froehlich \(2020b\)](#) used social network analysis to study how methods of data collection and analysis were connected to each other in educational MMSNA studies. Forty-four papers were included in the review. In these publications 32 different methods were mentioned and 287 relationships were coded, which by the aggregation of the ties (e.g., if two studies were coded to show a sequence of X to Y, this would result in a weight of “2”) resulted in 98 ties total. Based on the visualization of the coded method data as nodes and their connections, there are clear separations between methods that are (1) used very often, like the calculation of basic network metrics, sociometric surveys, and semi-structured interviews, (2) applied with moderate frequency, including group interviews and quantitative content analysis, or (3) hardly integrated with each other, such as video analysis, multilevel modeling or exponential-family random graph models.

These methods were then grouped using modularity algorithms (i.e., algorithms highlighting subgroups of nodes), which identified three subgroups: One subgroup of largely quantitatively oriented methods, such as surveys or communication logs for data collection and calculation of network metrics and correlations for data analysis, one subgroup of largely qualitatively oriented methods, such as semi-structured interviews and observation for data collection and thematic analyses and grounded theory for data analysis, and one subgroup that contained both quantitative and qualitative but rarely used methods, such as qualitative and quantitative content analysis or exponential-family random graph models.

Current debates

In this section, we focus on a selected number of topics that are discussed in the current debate about MMSNA.

Data collection and research designs

Social network analysis has strict data requirements ([Wasserman and Faust, 1994](#)) and there is a long-standing debate about data collection related questions such as measurement validity and the boundaries of a network (e.g., [Burt et al., 2012](#); [Längler and Froehlich, 2022](#); [Laumann et al., 1983](#); [Marsden, 1990, 2003](#)).

Introducing mixed methods to social network analysis has been relevant for both of these exemplary debates, as mixing has been used to inform the crucial boundary decision ([Smith, 2016](#)) or make more valid measurements through including multiple perspectives during data collection ([Längler et al., 2020](#)). The debate about how to make valid measurements also triggered a formalization of specific techniques of relational data collection, for example interviews ([Van Waes and Van den Bossche, 2020](#)) or ethnographies ([Sarazin, 2020](#)). Also, it is a useful approach to measuring and analyzing networks at multiple units of analysis ([Bolibar, 2015](#); [Froehlich et al., 2020b](#)); after all, quantitative methods of social network analysis do have clear strengths when it comes to observing macro-level constructs such as a network’s structure, while qualitative methods of social network analysis may be more useful to investigate individual network members’ views of and actions in the network.

In terms of research designs, there is a notable lack of “prototypical” MMSNA designs that could be used as a point of departure especially for novice MMSNA researchers. One exception is [Froehlich and Van der Wilt’s \(2020\)](#) and [Froehlich’s \(2020a\)](#) attempt to turn [Herz et al.’s \(2015\)](#) qualitative structural analysis into a MMSNA design. Otherwise, a design that could be described as an explanatory MMSNA research design, in which a sociographic survey is followed by in-depth interviews, is a common occurrence in the field of education ([Froehlich, 2020b](#)).

Ethical considerations

Ethical issues have received a lot of attention in social network research (e.g., [Borgatti and Molina, 2003](#); [Hogan, 2021](#); [Kadushin, 2005](#)), as there are several concerns with social network data, for example, in the direction of anonymity and the disclosure of

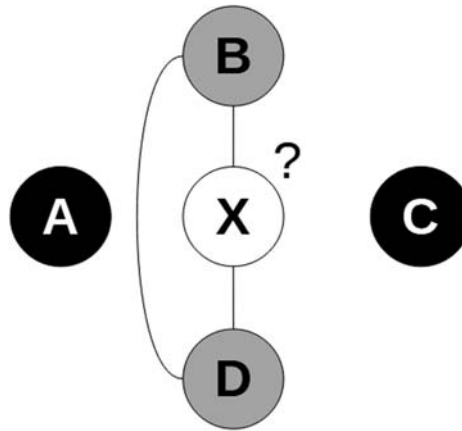


Fig. 2 Ethical concerns about non-participant data.

information about non-participants. Basically, the problem is that often precise inferences can be made from an individuals' network to the individual. **Fig. 2** shows an exemplary network, which could be about sensitive information such as political or sexual orientation. In this example, A and C have disclosed information about themselves (e.g., a certain political or sexual orientation) and B and D did the same. On the relational level, we only note connections of B and D to our focal actor, X, who did not participate in this fictional study. Even in this simple example, the inference that X will be more similar to B and D (assuming that the relationship describes something positive, such as friendship or frequent information exchange) and less similar to A and C seems likely.

While these concerns are applicable to any social network analysis, MMSNA amplifies these challenges, as even more in-depth information about (non-)participants are being collected. Hence this topic received quite some attention in the work about MMSNA, for example by [Korir et al. \(2020\)](#) or [Onwuegbuzie \(2020\)](#). Unfortunately, limited options exist for a resolution. The researchers could ask for consent also from those individuals that did not participate in the data collection, but who were mentioned by others. But even if consent is not given, it may be a difficult decision to leave out a person from analysis, as this could have a large effect on the overall structure of a network. In the worst case, this could render a dataset worthless for analysis. This dilemma emphasizes the importance of transparently communicating upfront and discussing the ethical challenges with independent reviewers.

Human capital limitations

Another problem that is frequently highlighted in mixed methods research concerns the human capital necessary to execute high-quality mixed methods research (cf. [Onwuegbuzie and Hitchcock, 2015](#)). Compared to mono-method studies, mixed methods studies are more demanding in terms of methodological knowledge. This is because capabilities in qualitative and quantitative methods need to be available, as well as knowledge about how to integrate both successfully. Arguably, this problem is exacerbated in MMSNA for three reasons. First, quantitative social network methods tend to be rather complex and require in-depth knowledge about quantitative procedures. Also, as outlined above, researchers usually do not receive a general training in relational methods, statistical procedures applicable within the field of social network analysis, such as quadratic assignment procedures ([Dekker et al., 2003](#); [Krackhardt, 1992](#)), exponential-family random graph models ([Robins et al., 2007](#)), or stochastic actor-based modeling ([Brouwer and Froehlich, 2020](#)) are usually not discussed in introductory statistics courses ([Froehlich et al., 2020a](#)). At the same time, these highly quantitative procedures often cannot simply be left out, as they contribute valuable structural information about networks. Also, as the field is very much driven by quantitative analyses ([Hollstein, 2014](#)), inclusion of these procedures makes it easier to integrate own findings in to the existing body of knowledge. Here, further standardizing and automating some of the quantitative procedures may alleviate the issue ([Froehlich et al., 2020a,b,c,d,e](#)).

Second, methods of qualitative network research are quite under-developed. While there has been some discussion about why qualitative social network research is important ([Decuyper, 2020](#); [Hollstein and Straus, 2006](#)), much less attention was given to the specification of qualitative research procedures. Hence, individual researchers need to be comfortable navigating this uncertainty by inventing their own procedures on the fly. Also, due to the recency of the last wave of qualitative and mixed methods network research compared to quantitative social network analysis, the (software) tools are often less refined, accessible, or stable than their quantitative counterparts. For example, published lists about software for social network research predominantly focus on quantitative applications (e.g., [Huisman and van Duijn, 2005](#)).

Third, while some publications break down the discussion of how mixed methods might be applied within the context of social network research—as reviewed in this chapter—it needs to be recognized that this strand of the literature is still in its infancy. Here, too, researchers are mostly required to build the path as they go, which again increases the burden for MMSNA researchers.

Practical impact

One emphasis in the educational mixed methods literature is its potential to not only have impact in academia, but to transform practice, too (Mertens, 2007, 2008). This is echoed also in the literature about MMSNA; which has been highlighted as a method highly conducive to deriving practical implications. For example, Palonen and Froehlich (2020) discuss the merits of MMSNA for organizational analysis. One advantage that social network analysis has over other methods of analysis is that it easily provides a novel (relational) perspective to stakeholders, while the visual nature of many (mixed methods) social network studies (Shannon-Baker and Hilpert, 2020) make the data very accessible also to laypersons.

We might also view organizational analysis using social network analysis as an important predecessor of MMSNA, as organizational analysis usually does not stop with a purely quantitative assessment. Instead, quantitative results are contextualized through interviews and workshops with stakeholders, which adds qualitative data (cf. the highly influential managerial book about organizational network analysis by Cross and Parker, 2004). However, we also need to point out that the (quite profound) ethical concerns outlined above are even more pronounced when using the data for other purposes than research (Borgatti and Molina, 2003).

Outlook and conclusions

As presented in the brief historical overview of the evolution of social network research, MMSNA as a concept is a relatively recent idea. While there are some empirical applications in the field of learning and education research (e.g., as included in the systematic reviews referenced above), the methodological discussion is just at the beginning. In this chapter, we have outlined a few strands of the discussion, namely data collection and research designs, ethical considerations, human capital limitations, and the direct impact on practice that MMSNA often seeks to make.

While these discussions are important starting points, they clearly are not conclusive at this point in time. Instead, the debate about how to best fit the idea of mixed methods into the context of social network analysis needs to be further refined. We believe that this may not be only useful for further applications of MMSNA, but it may also lead researchers to address some underdeveloped strands of the general social network literature, for example, reflecting on the very nature of networks, which in quantitative mono-methods studies was often reduced to the perspective of the researchers (thus raising validity concerns).

Another promising avenue for further research about MMSNA has less to do with mixing but with the development of qualitative relational research methods. While publications about social network analysis or publications using social network analysis have enjoyed exponential growth over the past decades, this is only true for quantitative (mostly mono-method) studies. A more coherent set of qualitative data collection procedures and qualitative analytical strategies would help to build a more varied base for future MMSNA designs. This claim can be substantiated by the review of Froehlich (2020b), which has shown that hardly any MMSNA study exists in the field of education that does not rely on some form of questioning (survey or interview).

References

- Barnes, J., 1983. Graph theory in network analysis. *Soc. Netw.* 5 (2), 235–244. [https://doi.org/10.1016/0378-8733\(83\)90026-6](https://doi.org/10.1016/0378-8733(83)90026-6).
- Bolibar, M., 2015. Macro, meso, micro: broadening the “social” of social network analysis with a mixed methods approach. *Qual. Quant.* 3, 2217–2236. <https://doi.org/10.1007/s11135-015-0259-0>.
- Borgatti, S.P., Halgin, D.S., 2011. On network theory. *Organ. Sci.* 22, 1168–1181. <https://doi.org/10.1287/orsc.1110.0641>.
- Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G., 2009. Network analysis in the social sciences. *Science* 323 (5916), 892–895. <https://doi.org/10.1126/science.1165821>.
- Borgatti, S.P., Molina, J.L., 2003. Ethical and strategic issues in organizational social network analysis. *J. Appl. Behav. Sci.* 39 (3), 337–349. <https://doi.org/10.1177/0021886303258111>.
- Brantner, C., Pfeffer, J., 2019. Content analysis of Twitter: big data, big studies. In: Eldridge, S.A., Franklin, B. (Eds.), *The Routledge Handbook to Developments in Digital Journalism Studies*. Routledge, Abingdon, UK, pp. 79–92.
- Brouwer, J., Froehlich, D.E., 2020. Co-evolution models of longitudinally measured interactions. In: Huber, M., Froehlich, D.E. (Eds.), *Analyzing Group Interactions: A Guidebook for Qualitative, Quantitative and Mixed Methods*. Routledge, London.
- Burt, R.S., Meltzer, D.O., Seid, M., Borgert, A., Chung, J.W., Colletti, R.B., et al., 2012. What's in a name generator? Choosing the right name generators for social network surveys in healthcare quality and safety research. *BMJ Qual. Saf.* 21 (12), 992–1000. <https://doi.org/10.1136/bmjqs-2011-000521>.
- Creswell, J.W., Plano-Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE Publications, Thousand Oaks, CA.
- Cross, R.L., Parker, A., 2004. *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*. Harvard Business Review Press, Boston, Mass.
- Crossley, N., 2010. The social world of the network. Combining qualitative and quantitative elements in social network analysis. *Sociologica* (1). <https://doi.org/10.2383/32049>.
- Daly, A.J., Finnigan, K.S., 2010. A bridge between worlds: understanding network structure to understand change strategy. *J. Educ. Change* 11 (2), 111–138. <https://doi.org/10.1007/s10833-009-9102-5>.
- Decuyper, M., 2020. Visual network analysis: a qualitative method for researching sociomaterial practice. *Qual. Res.* 20 (1), 73–90. <https://doi.org/10.1177/1468794118816613>.
- Dekker, D., Krackhardt, D., Snijders, T.A.B., 2003. Multicollinearity Robust QAP for Multiple-Regression. Nijmegen.
- Dominguez, S., Hollstein, B. (Eds.), 2014. *Mixed Methods Social Networks Research: Design and Applications*. Cambridge University Press, New York, NY. Retrieved from. <https://doi.org/10.1017/CBO9781139227193>.
- Euler, L., 1956. The seven bridges of Königsberg. *World Math.* 1, 573–580.
- Franke, K., Wald, A., 2006. Möglichkeiten der Triangulation quantitativer und qualitativer Methoden in der Netzwerkanalyse [Opportunities for triangulating quantitative and qualitative methods in social network analysis]. In: Hollstein, B., Straus, F. (Eds.), *Qualitative Netzwerkanalyse: Konzepte, Methoden, Anwendungen [Qualitative Network Analysis: Concepts, Methods, Applications]*. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 153–175.
- Freeman, L.C., 2004. *The Development of Social Network Analysis*. Empirical Press, Vancouver. Retrieved from. <http://aris.ss.uci.edu/in/book.pdf>.

- Frøehlich, D.E., 2020a. Exploring social relationships in "a mixed way": mixed structural analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 126–138.
- Frøehlich, D.E., 2020b. Mapping mixed methods approaches to social network analysis in learning and education. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London.
- Frøehlich, D.E., Brouwer, J., 2021. Social network analysis as mixed analysis. In: Onwuegbuzie, A.J., Johnson, R.B. (Eds.), *Routledge Reviewer's Guide for Mixed Methods Analysis*. Routledge.
- Frøehlich, D.E., van der Wilt, F., 2020. Mixed structural group analysis. In: Huber, M., Frøehlich, D.E. (Eds.), *Analyzing Group Interactions: A Guidebook for Qualitative, Quantitative and Mixed Methods*. Routledge, London.
- Frøehlich, D.E., Mamas, C., Schneider, H.W., 2020a. Automation and the journey to mixed methods social network analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London.
- Frøehlich, D.E., Mejeu, M., Galey, S., Schoonenboom, J., 2020b. Integrating units of analysis: applying mixed methods social network analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 38–48.
- Frøehlich, D.E., Rehm, M., Rienties, B.C., 2020c. Mixed methods social network analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 1–10.
- Frøehlich, D.E., Rehm, M., Rienties, B., 2020d. Reviewing mixed methods approaches using social network analysis for learning and education. In: Peña-Ayala, A. (Ed.), *Educational Networking*. Springer International Publishing, Cham, pp. 43–75. https://doi.org/10.1007/978-3-030-29973-6_2.
- Frøehlich, D.E., Van Waes, S., Schäfer, H., 2020e. Linking quantitative and qualitative network approaches: a review of mixed methods social network analysis in education research. *Rev. Res. Educ.* 44 (1), 244–268. <https://doi.org/10.3102/0091732X20903311>.
- Fuhse, J., 2020. Theories of social networks. In: Light, R., Moody, J. (Eds.), *The Oxford Handbook of Social Networks*. Oxford University Press, Oxford, England.
- Herz, A., Peters, L., Truschkat, I., 2015. How to do qualitative structural analysis: the qualitative interpretation of network maps and narrative interviews. *Forum Qual. Soc. Res.* 16 (1). <https://doi.org/10.17169/fqs-16.1.2092>.
- Hogan, B., 2021. Networks are a Lens for Power: A Commentary on the Recent Advances in the Ethics of Social Networks Special Issue. *Social Networks*. <https://doi.org/10.1016/j.socnet.2020.12.003>.
- Hollstein, B., 2014. Mixed methods social networks research: an introduction. In: Domínguez, S., Hollstein, B. (Eds.), *Mixed Methods Social Networks Research: Design and Applications*. Cambridge University Press, New York City, NY, pp. 3–34.
- Hollstein, B., Straus, F. (Eds.), 2006. *Qualitative Netzwerkanalyse [Qualitative Network Analysis]*, first ed. VS Verlag für Sozialwissenschaften, Wiesbaden. Retrieved from. <https://doi.org/10.1007/978-3-531-90074-2>.
- Hong, Q.N., Gonzalez-Reyes, A., Pluye, P., 2018. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *J. Eval. Clin. Pract.* 24 (3), 459–467. <https://doi.org/10.1111/jep.12884>.
- Huisman, M., van Duijn, M.A.J., 2005. Software for social network analysis. In: Carrington, P.J., Scott, J., Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. Cambridge University Press, Cambridge, MA, pp. 270–316. <https://doi.org/10.1017/CBO9780511811395.013>.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33 (7), 14–26. <https://doi.org/10.3102/0013189X033007014>.
- Kadushin, C., 2005. Who benefits from network analysis: ethics of social network research. *Soc. Netw.* 27 (2), 139–153.
- Korir, M., Mittelmeier, J., Rienties, B.C., 2020. Is mixed methods social network analysis ethical? In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 206–218.
- Krackhardt, D., 1992. A caveat on the use of the quadratic assignment procedure. *J. Quant. Anthropol.* 3 (4), 279–296.
- Kronenwett, Adolphs, 2021. VennMaker I Akteurszentrierte Darstellung und Analyse sozialer Netzwerke. Retrieved from. <https://www.vennmaker.com/>. (Accessed 27 April 2021).
- Längler, M., Frøehlich, D.E., 2022. Name generator. In: Frey, B.B. (Ed.), *The SAGE Encyclopaedia of Research Design*, second ed. SAGE Publications, Los Angeles, CA.
- Längler, M., Brouwer, J., Gruber, H., 2020. Data collection for mixed method approaches in social network analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 25–37.
- Laumann, E.O., Marsden, P.V., Prensky, D., 1983. The boundary specification problem in network analysis. In: Burt, R.S., Minor, E. (Eds.), *Applied Network Analysis: A Methodological Introduction*. SAGE Publications, Beverly Hills, CA, pp. 18–87.
- Marsden, P.V., 1990. Network data and measurement. *Annu. Rev. Sociol.* 16 (1990), 435–463.
- Marsden, P.V., 2003. Interviewer effects in measuring network size using a single name generator. *Soc. Netw.* 25 (1), 1–16. [https://doi.org/10.1016/S0378-8733\(02\)00009-6](https://doi.org/10.1016/S0378-8733(02)00009-6).
- Mejeu, M., 2020. Cross-sectional social network analysis. In: Huber, M., Frøehlich, D.E. (Eds.), *Analyzing Group Interactions: A Guidebook for Qualitative, Quantitative and Mixed Methods*. Routledge, London, pp. 85–95.
- Mertens, D.M., 2007. Transformative paradigm: mixed methods and social justice. *J. Mix. Methods Res.* 1 (3), 212–225.
- Mertens, D.M., 2008. *Transformative Research and Evaluation*. Guilford Press.
- Monge, P.R., Contractor, N.S., 2003. *Theories of Communication Networks*. Oxford University Press, New York, NY. Retrieved from. <https://doi.org/10.1093/os019780195160369.001.0001>.
- Moolenaar, N.M., 2012. A social network perspective on teacher collaboration in schools: theory, methodology, and applications. *Am. J. Educ.* 119 (1), 7–39. <https://doi.org/10.1086/667715>.
- Moreno, J.L., 1934. *Who Shall Survive?: A New Approach to the Problem of Human Interrelations*. Nervous and Mental Disease Publishing Co, Washington, DC, US, pp. xvi–441. <https://doi.org/10.1037/10648-000>.
- Onwuegbuzie, A.J., 2020. The PRICE of mixed methods social network analysis. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 245–262.
- Onwuegbuzie, A.J., Hitchcock, J.H., 2015. Advanced mixed analysis approaches. In: Hesse-Biber, S., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, NY, pp. 275–295.
- Palonen, T., Frøehlich, D.E., 2020. Mixed methods social network analysis to drive organizational development. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London.
- Pearson, M., Michell, L., 2000. Smoke rings: social network analysis of friendship groups, smoking and drug-taking. *Drugs Educ. Prev. Pol.* 7 (1), 21–37. <https://doi.org/10.1080/dep.7.1.21.37>.
- Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007. An introduction to exponential random graph (p*) models for social networks. *Soc. Netw.* 29 (2), 173–191. <https://doi.org/10.1016/j.socnet.2006.08.002>.
- Sarazin, M., 2020. Ethnographic mixed methods social network analysis studies: opportunities and challenges. In: Frøehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Approaches to Social Network Analysis*. Routledge, London, UK, pp. 72–86.
- Schoonenboom, J., Johnson, R.B., Frøehlich, D.E., 2018. Combining multiple purposes of mixing within a mixed methods research design. *Int. J. Mult. Res. Approaches* 10 (1), 271–282. <https://doi.org/10.29034/ijmra.v10n1a17>.
- Scott, J., 2007. *Social Network Analysis: A Handbook*, second ed. Sage Publications, London.
- Shannon-Baker, P., Edwards, C., 2018. The affordances and challenges to incorporating visual methods in mixed methods research. *Am. Behav. Sci.* 62 (7), 935–955. <https://doi.org/10.1177/0002764218772671>.

- Shannon-Baker, P., Hilpert, J.C., 2020. Visual methods and representations in mixed methods (and) social network research: a discussion. In: Froehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London.
- Smith, S.S., 2016. A three-step approach to exploring ambiguous networks. *J. Mix. Methods Res.* 10 (4), 367–383. <https://doi.org/10.1177/1558689815575855>.
- Tashakkori, A., Creswell, J.W., 2007. Editorial: the new era of mixed methods. *J. Mix. Methods Res.* 1 (1), 3–7. <https://doi.org/10.1177/2345678906293042>.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2020. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. SAGE Publications, Thousand Oaks, CA.
- Van Waes, S., Van den Bossche, P., 2020. Around and around: the concentric circles method as powerful tool to collect mixed method network data. In: Froehlich, D.E., Rehm, M., Rienties, B.C. (Eds.), *Mixed Methods Social Network Analysis: Theories and Methodologies in Learning and Education*. Routledge, London, pp. 159–174. Retrieved from. <https://doi.org/10.4324/9780429056826>.
- Wasserman, S., Faust, K., 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press, New York, NY. Retrieved from. <https://doi.org/10.1017/CBO9780511815478>.

Mixed methods and web-based technologies for data collection

Janet Salmons, Research Community Manager SAGE Methodspace, SAGE Publishing, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	693
Collecting data online	694
Collecting qualitative and/or quantitative data online	694
Mapping potential ways to mix method	695
Defining ways to mix methods and milieux	696
Priority	696
Timing	696
Milieu	697
Analyzing mixed online and hybrid studies	697
Fully-online mixed methods studies	698
Hybrid on-and off-line mixed methods studies	699
Learning from examples	700
Summary: online mixed methods researchers have many options	700
References	700

Glossary

Big data A set of data that is too large and complex for processing by traditional database management or computational tools

Big qual Data sets containing either secondary qualitative data or primary data with at least 100 participants, analyzed by teams of researchers, often funded by a government agency or private foundation, and conducted either as a stand-alone project or in conjunction with a large quantitative study (Brower et al., 2019)

Elicited data Data collected by posing questions or prompts to human participants. These methods include interviews, focus groups, or surveys

Enacted methods Data collected through experiential activities such as creative or arts-based methods, participant observations, or simulations

Extant data Data that exists without intervention by the researcher

Methods Procedures or techniques used to implement the sampling, data collection, or data analysis steps within a research study (Plano Clark and Ivankova, 2016)

Methodology The process of research from formulating questions to drawing conclusions in a study (Plano Clark and Ivankova, 2016)

Mixed methods research A term used for the process of research when researchers integrate quantitative methods of data collection and analysis and qualitative methods of data collection and/or analysis (Plano Clark and Ivankova, 2016)

Mixed-reality world A term used to describe living with online and technology-assisted experiences in conjunction with life in the physical world

Multimethod research A term used for the process of research when researchers integrate multiple quantitative approaches, multiple qualitative approaches, or multiple quantitative and qualitative approaches (Plano Clark and Ivankova, 2016)

Research design A thoughtfully constructed link between the purposes of a research study and the strategies used to implement it (Creamer, 2018)

Research paradigm The collection of assumptions and values about the nature of reality and knowledge that provide the foundation for a research study (Plano Clark and Ivankova, 2016)

Secondary data Data that have already been collected for some other purpose. In this chapter, secondary chapter is referred to as extant data

Introduction

We need mixed methods to study a mixed-reality world. Activities associated with independent and classroom learning, reading and study, discussion and collaboration have changed profoundly with widespread adoption of the internet and related information and communications technologies (ICTs). To understand the experiences and dilemmas of digitally-oriented modern life, we need to re-think methods for planning studies and collecting data.

Quantitative and qualitative researchers view this challenge of studying technology-infused lived experiences differently. The ever-growing use of online communication and social media—and the resulting big data—excites quantitative researchers who can scrape and download secondary data rich with potential relationships and correlations. Quantitative researchers are glad to access diverse participants for online surveys or experiments. Similarly, the ability to use online communication and social media to create deep exchanges with research participants across the globe—the humans on the other side of the monitor or device—excites qualitative researchers. Nevertheless, the line between qualitative and quantitative methods is blurring. Qualitative approaches, termed *Big Qual*, are being used to analyze large bodies of content previously associated with quantitative studies (Brower et al., 2019). Meanwhile some quantitative researchers are realizing that “experts with deep knowledge of a subject are needed to ask the right questions and to recognize the shortcomings of statistical models”. Sometimes in the midst of a sea of data, the quantitative researcher needs to stop and ask individual participants, “why?” When researchers mix qualitative and quantitative methods, online and in-person approaches, they have the ability to study the rich complexity of contemporary life with research that is big in scope and deep in meaning. This chapter focuses on emerging ways to mix qualitative and quantitative methods in studies that make use of the internet, applications on devices, or other ICTs for data collection.

Collecting data online

We will broadly define mixed methods as the “intentional integration of quantitative and qualitative research approaches to best address a research problem” (Plano Clark and Ivankova, 2016). In short, integration of methods is not limited to the collection of different types of data. In some cases, it is not the nature of the data itself that signifies whether the study is deemed qualitative or quantitative, but the way the data are analyzed and used within the study. For example, in mixed methods studies stories collected with qualitative approaches can be coded and analyzed with statistical methods, and narrative stories can be told based on statistical findings drawn from quantitative studies. While mixing methods means thinking and re-thinking all aspects of research from initial formulation of questions through analysis and reports of findings, this chapter centers attention primarily on data collection.

For the purpose of this chapter, we recognize that the source and setting of the study are noteworthy when we consider ways to use the digital environment in data collection for mixed methods research, and will use this definition:

Conducting mixed methods online means the researcher intentionally selects and integrates quantitative and qualitative data collected on the internet, through applications or platforms, or using other information and communications technologies (ICTs) to best address a research problem.

Collecting qualitative and/or quantitative data online

Clearly, the simplistic conceptions of qualitative data as words and quantitative data as numbers are inadequate when considering the wide variety of source materials and interactions that are possible in the digital world. One way to categorize data collected is by the type and degree of relationship between the researcher and the source of the data. As illustrated in Fig. 1, researchers have three options for using qualitative or quantitative methods to collect or generate data that correspond to close or distant relationships between the researcher and source(s) of data.

The first option is to use extant data, that is, data that exist without intervention by the researcher. Extant data were created for another purpose but is available to the researcher. The researcher can find extant data on the public web, in libraries or archives, private collections, organizational or government records. It can take many forms, from casual posts to well-researched scientific reports, from personal diaries to large-scale datasets. Some extant data are readily available, other types might require permissions to access and use for research purposes.

The advantage of extant data is obvious: the materials already exist. Some are in written form that means transcriptions are not necessary, and media posted in many libraries include transcriptions. The popularity of user-generated posts and comments means all kinds of perspectives can be discovered on open sites, social media, and online groups or communities. The disadvantages are that the materials might not be fully relevant to the research problems, might be incomplete, or lack context.

The second type of data are elicited from participants. Whether using a quantitative instrument for an online survey to be completed by hundreds of respondents or interviewing twenty participants on a videoconferencing platform, these methods invite responses to researchers’ questions. The advantage of this kind of data is the focused relevance to the research problem, and greater degree of control by the researcher than is possible when using extant data. The disadvantage is that participants must be found, recruited, and they must consent to contribute to the study.

The third type of data collection, enacted methods, takes a different tack with participants. While still eliciting responses, rather than asking questions, the researcher creates some way to generate data through some kind of activity. Quantitative researchers conduct experiments, and qualitative researchers use arts-based, creative, or participatory methods. Some studies using enacted methods consider participants as collaborative partners or co-researchers. The advantage of methods in this category is that they can produce rich, substantial data that can produce substantive findings. The disadvantage is that general, enacted methods expect

Typology of Online Methods (Salmons, 2022)		Data collected online from:	Researcher & Participant
Extant	Studies using existing materials developed without the researcher's influence.	• Blogs, websites, or posts on social media. • Published literature, reports, or documents. • Archived materials, records, film or video. • Databases, datasets or Big Data. • Unobtrusive observation.	No direct contact with individual participants.
Elicited	Studies using data elicited <i>from</i> participants in response to the researcher's questions/prompts.	• Interviews, 1-1 or group. • Focus groups. • Written interviews, questionnaires, or surveys. • Participant observation of events or communities.	Interaction between researcher and one or more participants
Enacted	Studies using data generated <i>with</i> participants during the study.	• Vignette, scenario, or problem-based interviews. • Games, simulations or role plays. • Arts-based or creative methods. • Experiments or Field experiments. • Participatory Action Research.	Multiple interactions and collaboration involving researcher one or more participants

Fig. 1 The typology of online methods (Salmons, 2022).

a greater commitment from participants than do elicited methods and might involve a multi-stage process that lasts a period of time.

Researchers can mix qualitative and quantitative data collection approaches across the extant, elicited, and extant types of data collection. Researchers might use one method to gain insights or understanding about the problem, and another method to inquire about how people experience the problem. These three open-access articles offer examples that mix qualitative and quantitative approaches for extant, elicited, and/or enacted data collection:

- **Qualitative Elicited and Quantitative Enacted Data Collection** The first example is: "College Students and Yik Yak: An Exploratory Mixed-Methods Study" (Clark-Gordon et al., 2017). In this study researchers began with a qualitative, elicited stage by interviewing a group of participants about their use of a social media platform. Researchers asked participants to respond and self-report their experiences on a smartphone application. In the second quantitative stage, they used extant data in the form of posts ($n = 3905$) for a content analysis.
- **Quantitative Elicited, Quantitative Extant, and Qualitative Elicited Data Collection** The second example is: "International Platforms, International Prejudice in the Platformization of Crafting" (Close and Wang, 2020). In this study researchers began by eliciting data with an online survey. They conducted a discourse analysis of extant data in the form of online posts, then elicited data with qualitative interviews.
- **Quantitative Extant, Qualitative Elicited, and Qualitative Extant Data Collection** The third examples is: "Ghosts of white methods? The challenges of Big Data research in exploring racism in digital context" (Nikunen, 2021). After analyzing extant big data, the researchers elicited data by interviewing the founder and participants of the project under investigation and using an ethnographic qualitative approach to study a selection of postings.

Mapping potential ways to mix method

To illustrate some of the many ways researchers can mix methods to collect or generate different types of data online, Figs. 2–5 show four hypothetical research designs.

Fig. 2 shows stages in an online mixed methods study that with extant quantitative methods using datasets to gain large-scale sense of trends, and interviews to elicit stories and personal experiences from a small set of participants.

Fig. 3 shows stages in an online mixed methods study that uses qualitative methods to elicit insights and find background about trends in extant data, and a quantitative survey to elicit responses that will make it possible to evaluate the extent to which those insights are common to a larger population.

Fig. 4 shows stages in an online mixed methods study that uses quantitative survey methods to elicit responses that will help to identify trends, and arts-based qualitative enacted methods to generate visual representations of participants' experiences.



Fig. 2 Quantitative extant and qualitative elicited methods.

Fig. 5 shows stages in a hypothetical online mixed methods study that uses quantitative methods to elicit responses with a survey, experimental quantitative enacted methods to measure types of participant experiences, and qualitative interviews to elicit verbal explanations.

Defining ways to mix methods and milieux

Mixed methods researchers have classified some common design types and created a notation system as a concise way of signaling which one is being used. The two main classification factors are *priority* and *timing*, and we will add a third: *milieu*. The milieu refers to the setting where the data is collected. For the purpose of this chapter, research activities occurring in the same physical location are referred to as *in-person*. Research activities that use technology tools are referred to as *online*. See Table 1 for a description of option.

Priority

Priority refers to the degree of dominance of one method over another. Creswell et al. (2003) explain:

The mixed methods researcher can give equal priority to both quantitative and qualitative research, emphasize qualitative more, or emphasize quantitative more. This emphasis may result from practical constraints of data collection, the need to understand one form of data before proceeding to the next, or the audience preference for either quantitative or qualitative research. In most cases, the decision probably rests on the comfort level of the researcher with one approach as opposed to the other.

Creswell et al. (2003).

It is important to communicate the emphasis or balance of methods to the reader. In the notation system developed by Morse (1991), the dominant method used is shown with all capital letters: *QUAL* or *QUAN*. In a study where both methods are equally important, both are capitalized; if one is more important, that term is capitalized while the supportive or secondary type of data is shown in lower case: *qual* or *quan* (Morse, 1991). For example, if I am conducting a small number of qualitative interviews to identify issues I want to explore widely in a large-scale quantitative survey, I might describe the interview part as *qual*, because it is minor stage of the study, and the survey part as *QUAN*, because it is a more significant part of the research. Or, if the interviews and survey have equal importance, I would indicate that by labeling the stages as *QUAL* and *QUAN*.

Timing

The second factor is timing. Morgan identified two core assumptions: that in addition to degrees of priority for qualitative or quantitative methods, the designs varied in terms of a sequence of collecting quantitative and qualitative data (Morgan, 2007). Options are:



Fig. 3 Qualitative extant and quantitative elicited methods.



Fig. 4 Quantitative elicited methods and qualitative enacted methods.

- Concurrent, or simultaneous, meaning both kinds of research are occurring at the same time. This is indicated with a +sign. For example, if I am conducting qualitative interviews while waiting for responses from a quantitative survey, I can describe two types of data collection as concurrent.
- Sequential, meaning the researcher collects and analyzes one kind of data before another. The next phase builds on the results of the first phase. This is indicated with an arrow →. For example, if I am conducting qualitative interviews to identify issues I want to explore widely in a subsequent quantitative survey, I can describe two types of data collection as sequential.

Milieu

The milieu is an additional factor that is important to indicate when classifying a study that includes a setting important to the design of the research. A study could have all parts of the study in one setting or take a hybrid approach. All parts of the study could be conducted online or using an app or other technology tool, or with a hybrid approach using some data collected online and some collected in-person.

The examples illustrated in Figs. 2–5 offer just a few combinations of data collection types with qualitative and quantitative data collection that occurs in a digital milieu. The surveys could be conducted with a tool such as Qualtrics or Survey Monkey, and potentially use a panel of respondents recruited by those companies. Private, proprietary applications or platforms could be used for the enacted research events. Interviews could be conducted using videoconferencing tools such as Zoom™, Skype™, or Teams™, or researchers could collect data by sending written questions by email.

Researchers could also mix online data collection activities with face-to-face approaches. As shown in Fig. 6, some stages could collect data from a larger population online, while other stages involve local participants in-person and online. The potential for combinations means these methods is limited only by the researcher's imagination and energy!

For the purpose of this chapter, we will signal that the data were collected online by prefacing the method with an "o," and if the data were collected in-person with a "p." We will use the same capitalization convention, with the dominant form noted with a capital letter. Using the same example, if I am conducting online qualitative interviews and an online quantitative survey, I can describe two types of data collection as *o-qual* and *o-quan*. Alternatively, if I am conducting qualitative interviews in-person and an online quantitative survey, I can describe two types of data collection as *p-qual* and *o-quan*. Again, use capitalization to show equal or different levels of priority within the study. Table 1 illustrates online and hybrid milieu mixed methods design options for collecting data.

Analyzing mixed online and hybrid studies

How do researchers mix methods and combine different research paradigms when some or all of the data are collected online? Let's examine some published studies to see what approaches are used, and why the researchers chose to use mixed methods. These examples demonstrate how the methodological concepts introduced in this chapter have been used in contemporary research.

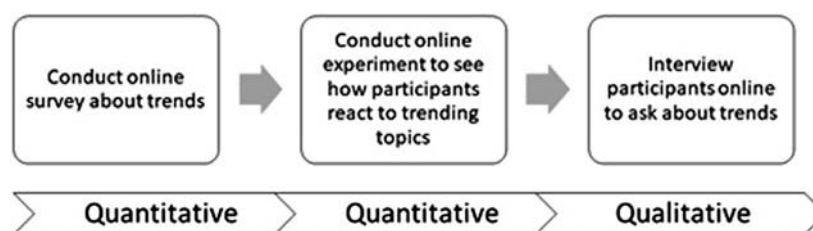


Fig. 5 Quantitative elicited methods and qualitative enacted methods.

Table 1 Communicating online and hybrid milieu mixed methods designs.

<i>Design notation</i>	<i>Explanation</i>
QUAL + QUAN	Qualitative and quantitative parts of the study have equal weight, and both parts of the study are conducted in the same time frame.
O-QUAL + O-QUAN	Qualitative and quantitative parts of the study have equal weight, and both kinds of data are collected online in the same time frame.
O-QUAL + o-quan	Qualitative and quantitative data are collected online in the same time frame. The qualitative part of this study is more significant than the quantitative part.
o-qual + O-QUAN	Qualitative and quantitative data are collected online in the same time frame. The quantitative part of this study is more significant than the qualitative part.
O-QUAL + P-QUAN	Qualitative and quantitative parts of the study have equal weight, and both kinds of data are collected in the same time frame. The qualitative data is collected online; the quantitative data is collected in-person.
O-QUAL + p-quan	Qualitative and quantitative data are collected in the same time frame. The qualitative data is collected online; the quantitative data is collected in-person. The qualitative part of this study is more significant than the quantitative part.
o-qual + P-QUAN	Qualitative and quantitative data are collected in the same time frame. The qualitative data is collected online; the quantitative data is collected in-person. The quantitative part of this study is more significant than the qualitative part.
QUAL → QUAN or QUAN → QUAL	Qualitative and quantitative parts of the study have equal weight, but one type data is collected before the other.
O-QUAL → O-QUAN or O-QUAN → O-QUAL	Qualitative and quantitative parts of the study have equal weight, and both are collected online, but one type data is collected before the other.
O-QUAL → p-quan	Qualitative data are collected before quantitative data. Qualitative data take priority and are collected online. Quantitative data are collected in-person.
O-QUAL → o-quan	Qualitative data are collected before quantitative data. Qualitative data take priority. Both types of data are collected online.
P-QUAL → o-quan	Qualitative data are collected before quantitative data. Qualitative data take priority and are collected in-person. Quantitative data are collected online.
O-QUAN → p-qual	Quantitative data are collected before qualitative data. Quantitative data take priority and are collected online. Qualitative data are collected in-person.
O-QUAN → o-qual	Quantitative data are collected before qualitative data. Quantitative data take priority. Both types of data are collected online.

Fully-online mixed methods studies

The first example describes a mixed method study with all data collection carried out online. The article, “Student’s experiences with online teaching following COVID-19 lockdown: A mixed methods explorative study” (Almendingen et al., 2021) reports on a study that collected data from two questionnaires and an online focus group. Questionnaires were sent to participants in a class on weeks two and twelve of the online class, followed by interviews. The researchers described it as a “a mixed methods cross-sectional design, where quantitative and qualitative methods complemented each other” (p. 3). They also pointed out that “[s]ince the study is exploratory, we had not planned the data collection in order to test hypotheses. The study seeks to provide a snapshot in time of an evolving situation” the sudden shift to online learning (p. 14). So while this was a mixed methods study, the researchers used descriptive analysis of the quantitative side, and did not complete a full statistical analysis of the data.

Here is the sequence of data collection stages:

1. The researchers adapted an existing instrument with Likert-scale and open-ended queries. They used quantitative and qualitative questions in the same online questionnaire to elicit data. Respondents could complete it using a browser on a computer, mobile phone or tablet.

**Fig. 6** Mixing methods and milieux.

2. The researchers followed with a second questionnaire to elicit more detailed participant responses.
3. The researchers conducted focus group interviews online in the third phase. The researchers used qualitative methods to elicit data online using a videoconferencing platform. Interviews were not recorded to preserve anonymity, so the moderator's notes served as data.

The data analysis stages progressed the same way, with the quantitative data studied before the qualitative data was coded and summarized. Given this sequence and the discussion of the study where both types of data were equally important, data collection stages could be notated as O-QUAN → O-QUAN → O-QUAL.

Another fully-online study is reported in the article, "Harnessing Technology for Research During the COVID-19 Pandemic: A Mixed Methods Diary Study Protocol" (Saltzman et al., 2021). This article describes a study based on data collected from qualitative diary entries and a quantitative survey. Participants were students in a Master of Social Work program, so it was not necessary to go through a laborious recruitment process.

Here is the sequence of data collection stages:

1. Researchers elicited responses from student participants using an online survey. These data were used "to supplement and further elucidate the themes that emerged in the longitudinal phenomenological data from the virtual diary entries" (p. 4).
2. The researchers set up an enacted data collection process for a longitudinal 8-week series of video diaries. Once the research team received the survey responses, they arranged eight weekly Zoom™ meetings for each participant. Each week, researchers sent prompts to elicit Participants logged into Zoom™ and recorded their video diary entries. If they entered text in the chat function, their words were used without edits from the research team. The platform generated an automated transcript for each entry.

Data analysis followed the same pattern, with a descriptive analysis of data from the survey, followed by a multi-stage thematic analysis of the qualitative diaries. Given that the study drew from phenomenological methodology and a statistical analysis of the quantitative data was not conducted, it is apparent that the qualitative side took priority. To notate this study, o-quan → O-QUAL seems appropriate.

Hybrid on-and off-line mixed methods studies

It is not always possible to conduct all phases of a mixed methods study online. Hybrid research allows for the advantages of both research modes. For example, the article, "Questioning Assumptions About Online Tutoring" (Wisniewski et al., 2020) describes a "comparative study of face-to-face and synchronous audio-video online tutorials that collected data from writing tutorials, writers' postsession surveys, and interviews with writers" (p. 261). The setting for this study was an online platform the university writing center used for tutorials, which meant it was a familiar tool for participating students. Data were collected by recording online and face-to-face tutorial sessions, collecting postsession writer surveys, collecting tutor logs, and conducting interviews with writers and tutors.

This study was primarily conducted online, with online data collection and use of technology tools for data analytics. One stage, recording tutorial sessions, occurred in-person and while the article does not detail the interview settings it appears that they were conducted in person.

The sequence went as follows:

1. Interviews to pin down existing beliefs about and attitudes toward face-to-face and online tutoring. The researchers used qualitative methods to elicit data.
2. Recording online and face-to-face tutorial sessions. The researchers used qualitative enacted methods with recorded events.
3. Anonymous survey. The researchers used quantitative methods to elicit data.
4. Interviews. The researchers used qualitative methods to elicit data.

The data analysis stage was also sequential:

1. Coding session transcripts to generate categories of instruction within the tutorials (qualitative)
2. Coding tutorial topical content (qualitative)
3. After establishing codes and counts, the researchers created an SPSS dataset and used it to analyze mean differences between variables in the two online and face-to-face sessions (quantitative).

Given that three of four data collection methods were qualitative, and two out of three data analysis stages used qualitative methods, we can surmise that the capitalization for QUAL and quan reflects the researchers' priorities. The sequence of on- and off-line stages of data collection could be notated as: O-QUAL → P-QUAL → o-quan → O-QUAL. The stages of data analysis could be notated as: O-QUAL → o-quan.

The article, "Stress and family relationships among college student parents: A mixed methods study" (Dotterer et al., 2020) explores research using a concurrent mixed methods design. Participants completed an online survey and in-person semi-structured interviews. Demographic and financial income information were also collected about the participants.

While the researchers described the study as "concurrent," they describe research stages as:

1. Completion of an online survey. The researchers used quantitative methods to elicit data.

Table 2 Published examples of online and hybrid mixed methods designs.

<i>Design notation</i>	<i>Citation</i>
Data collection: O-QUAL → P-QUAL → o-quan → O-QUAL data analysis: O-QUAL → o-quan	Wisniewski et al., 2020
Data collection and analysis: O-QUAN + P-QUAL.	Dotterer et al., 2020
Data collection and analysis: O-QUAN → O-QUAN → O-QUAL.	Almendingen et al., 2021
Data collection and analysis: o-quan → O-QUAL	Saltzman et al., 2021

2. In-depth interview conducted in-person with a small subsample of the population. The researchers used qualitative methods to elicit data.

The analytic process similarly progressed in two stages:

1. Descriptive analysis and regression analysis using SPSS with survey data (quantitative)
2. Thematic analysis using interview data (qualitative).

Based on the information provided in this article, this study balanced the qualitative and quantitative priorities in both data collection and analysis so the QUAL and QUAN would be accurate. The study also mixed an online and an in-person stage of data collection. The notation that summarizes the research as described in this article is: O-QUAN + P-QUAL.

Learning from examples

The styles used in this small collection of studies are summarized in [Table 2](#).

Given the time frame for the research described in these four studies, at the advent of the Covid-19 pandemic, the researchers took advantage of the opportunity to collect some or all of the data online. Given this situation, their rationales for carrying out their respective studies could be described as convenient. They had limited options for collecting data face-to-face. That said, the researchers discovered the value of using technology tools, such as automated transcriptions or the ability to reach participants in relaxed settings. [Table 2](#) shows research examples for online and hybrid milieu mixed methods designs.

Summary: online mixed methods researchers have many options

This chapter introduced a range of possibilities for combining qualitative and quantitative methods and for using technology. By mixing qualitative and quantitative data at various stages of the study, researchers can dig deep and explore broadly. Researchers also need to explain complex studies in ways readers can understand. The Typology of Online Methods ([Salmons, 2022](#)) and notational system articulated by Morse in 1991 and modified here to recognize the online or in-person setting, provide ways to succinctly communicate web-based research designs.

Online research offers researchers the ability to visit libraries and archives to discover historical and contemporary materials, conduct interviews, observations, questionnaires, and surveys without leaving their offices. Researchers can reach across the globe or collect important local perspectives. They can tap into culturally-rich ways people experience some event by scraping massive numbers of posts to social media sites or ask someone about their singularly unique experience.

With so many options, you will need to be able to make design choices appropriate to the research purpose and intentionally determine which types to use. Begin by learning how each paradigm works, and the strengths and limitations of each. Then look at how you can bring the most useful mode of inquiry into your own research plans.

References

- Almendingen, K., Morseth, M.S., Gjølstad, E., Brevik, A., Tørris, C., 2021. Student's experiences with online teaching following COVID-19 lockdown: a mixed methods explorative study. *PLoS One* 16, 1–16.
- Brower, R.L., Jones, T.B., Osborne-Lampkin, L.T., Hu, S., Park-Gaghan, T.J., 2019. Big qual: defining and debating qualitative inquiry for large data sets. *Int. J. Qual. Methods* 18, 1609406919880692.
- Clark-Gordon, C.V., Workman, K.E., Linvill, D.L., 2017. College students and Yik Yak: an exploratory mixed-methods study. *Social Media + Society* 3, 2056305117715696.
- Close, S., Wang, C., 2020. International platforms, international prejudice in the platformization of crafting. *Social Media + Society* 6, 2056305120940691.
- Creamer, E.G., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. SAGE Publications, Inc., Thousand Oaks, California.
- Creswell, J.W., Plano Clark, V., Gutmann, M.L., Hanson, W.E., 2003. An expanded typology for classifying mixed methods research into designs. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, pp. 209–240.
- Dotterer, A.M., Juhasz, A.C., Murphy, K.N., Park, S., Boyce, L.K., 2020. Stress and family relationships among college student parents: a mixed methods study. *J. Soc. Pers. Relat.* 38, 888–911.
- Morgan, D.L., 2007. Combining qualitative and quantitative methods Paradigms lost and pragmatism regained: methodological implications. *J. Mix. Methods Res.* 1, 48–76.

- Morse, J.M., 1991. Approaches to qualitative-quantitative methodological triangulation. *Nurs. Res.* 40, 120–123.
- Nikunen, K., 2021. Ghosts of white methods? The challenges of Big Data research in exploring racism in digital context. *Big Data Soc.* 8, 20539517211048964.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE Publications, Inc., Thousand Oaks, California.
- Salmons, J., 2022. *Doing Qualitative Research Online*. SAGE Publications, London.
- Saltzman, L.Y., Terzis, L.D., Hansel, T.C., Blakey, J.M., Logan, D., Bordnick, P.S., 2021. Harnessing technology for research during the COVID-19 pandemic: a mixed methods diary study Protocol. *Int. J. Qual. Methods* 20, 1609406920986043.
- Wisniewski, C., Regidor, M.C., Chason, L., Groundwater, E., Kranek, A., Mayne, D., Middleton, L., 2020. Questioning assumptions about online tutoring: a mixed-method study of face-to-face and synchronous online writing center tutorials. *Writ. Cent. J.* 38, 261–296.

Integrating art and artistic practices in mixed methods research in education

Lyric Smith and Peggy Shannon-Baker*, Georgia Southern University, Statesboro, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	702
Methodology, design, or methods: bringing arts-related and mixed methods research together	702
Framing various forms of arts-related research	703
Framing mixed methods research	704
Defining arts-based and arts-informed mixed methods research	704
Integrating art and artistic practices into mixed methods studies in education	705
Embedding art within other methodological approaches: examples of AIMMR	705
Centering artistic practices in a mixed methods study: examples of ABMMR	707
Challenges to using and integrating arts-related and mixed methods research	707
Challenges faced in both arts-related and mixed methods approaches	707
Level of training	707
Clear discussion of methods/practices	708
Support for art in a mixed methods study	708
Using clear terminology	708
Challenges from the tensions between arts-based research and mixed methods research	708
Aim of the research	709
Evaluating the research	709
Conclusions	709
Acknowledgment	709
References	709

Introduction

Some knowings cannot be conveyed through language.

Elizabeth Ellsworth (2005, p. 156).

As educational researchers study increasingly complex phenomena and educational contexts (Poth, 2018), they need methods that are expansive enough to explore those complexities. Mixed methods studies that integrate art (e.g., dance, painting, photography, creative writing) and artistic practices (e.g., keeping a visual journal, experimenting with art techniques) can meet this need. The purpose of this chapter is to introduce how and why educational researchers integrate art and artistic practices into mixed methods research (MMR) studies. We first define several key terms including arts-based, arts-informed, and other arts-related research approaches and MMR. We then highlight several studies to demonstrate how educational researchers have used the arts in different ways and their rationales for including/integrating art. Additionally, we identify several challenges facing researchers who aim to integrate art in MMR including strategies to address each.

Methodology, design, or methods: bringing arts-related and mixed methods research together

It is important to first differentiate between research methodology, research design, and research methods. *Research methodology* refers to the overarching framework of values and practices related to the research designs and methods that a researcher uses. Examples of common research methodologies include qualitative research, quantitative research, arts-based research, and MMR. Although connected to philosophical paradigms such as constructivism, pragmatism, postmodernism, etc., researchers can use a specific research methodology with multiple paradigms (cf. Shannon-Baker's chapter on philosophical paradigms in this volume). Within a research methodology or across multiple methodologies are the *research designs*, or the specific frames that guide and organize specific research methods. Example qualitative research designs include case study, ethnography, critical discourse analysis, etc. Example mixed methods designs include sequential explanatory or exploratory, concurrent/convergent parallel, etc. (Creswell and Plano Clark, 2017; DeCuir-Gunby and Schutz, 2017). Within a research design are *research methods*, or the specific practices, techniques, and procedures to collect, analyze, and interpret data. For example, a case study design might use research methods such as

*Corresponding author

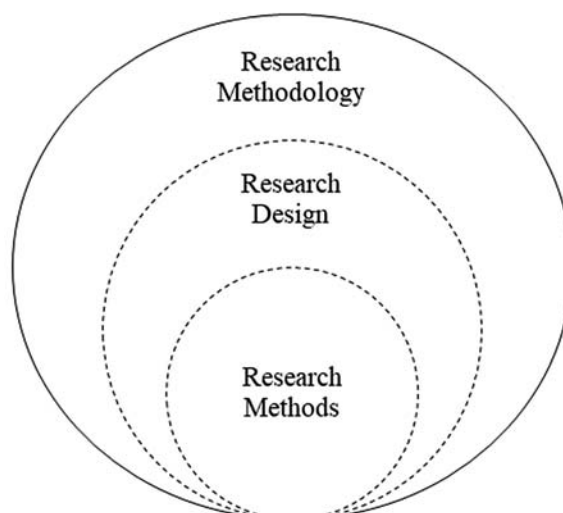


Fig. 1 Venn Diagram showing the relationship between research methodology, research design, and research methods.

individual interviews, focus group interviews, and deductive coding of the data to compare across cases. **Fig. 1** displays how these three frames are interconnected. This figure depicts research methodology providing the overarching frame that encompasses both research design and research methods. Although research methodologies might determine particular research designs and research methods, these are not static categories. Some methods and designs cut across multiple methodologies, signified by the dashed arrows.

In the following sections, we provide a foundation for defining and characterizing the following: various forms of arts-related research including arts-based and arts-informed research, MMR, and both arts-based mixed methods research and arts-informed mixed methods research.

Framing various forms of arts-related research

We use the term *arts-related research* to capture the wide range of research forms that use art and artistic practices, where art refers to the product or material created (e.g., a painting, performance, poem) and artistic practice refers to the process of creating and interpreting art (e.g., painting, performing, writing poetry). Although some researchers have dated the use of arts-related research approaches in social sciences back to the 1970s with the work of Elliott Eisner and Maxine Greene (Sinner et al., 2006), art was considered research long before. For example, Pablo Picasso stated, “I never made a painting as a work of art, it’s all research” (as cited in McNiff, 2008, p. 29). Here Picasso is referring to the process of engaging in artistic practices including studying and practicing techniques in an art form, experimenting with different forms and format, and studying the topic that will be portrayed in the art.

Arts-based research has been characterized as a “set of methodological tools” (Leavy, 2009, p. ix), as living inquiry (Irwin and Springgay, 2008) and a “dynamic process of inquiry” (Sinner et al., 2006, p. 1229). For the purpose of this chapter, we conceptualize *arts-based research* (ABR) as a research methodology where artistic practices are deeply integrated in and central to the entire research process (Austin and Forinash, 2005; Leavy, 2009; Sinner et al., 2006; Slattery, 2003). In ABR, some practitioners differentiate between practices and methods: “Practices may involve a range of activities at the center of research, such as customary approaches to art-making, creative rituals, original performances [...]. In contrast, *methods* suggest a more traditional research orientation involving” sampling techniques and data collection procedures and products (Sinner et al., 2006, p. 1229). These artistic practices might include but are not limited to the creation of artistic pieces—literary, visual, and performative—as data, analysis, and/or sharing research results; intensive reflection on the researcher’s and participants’ subjectivities; and heightened concern about form, representation, and creation of art pieces. ABR emphasizes inquiry, ambiguity, open-endedness, and fluidity (Eisner, 2008; Ledger and Edwards, 2011), the practices, processes, and products of research (Sinner et al., 2006), and is often participatory in nature (Estrella and Forinash, 2007). Researchers might use arts-based research for a variety of reasons. They may seek to use “alternative formats to evoke or provoke understandings that traditional research formats cannot provide” (Sinner et al., 2006, p. 1225). ABR may also be used to promote agency within communities by offering alternative forms of communication to explore and share ideas and experiences (Lincoln, 1995; Shannon-Baker and Martinez, 2022). Based on the research objectives, arts-based research can attempt to operate within, alongside, and/or against traditional research methodologies and methods. For more introduction to other forms of arts-based research, see the chapters in this volume by Hinekura Smith and Janinika Greenwood.

Similar to arts-based research, *arts-based educational research* (ABER) reflects the use of arts-based approaches specific to educational research. (Barone and Eisner, 2006), former student and teacher considered among the founders of ABER, state that ABER should specifically consider the format of the final write-up of the research—that it should be “evocative, contextual and vernacular” (p. 97). Examples of ABER include short stories, oral history, auto/biographies, educational criticism, literary essays, theater, and poetry. These writing-based art forms invite readers’ interpretations, utilize metaphors and thick description, and attract diverse

readers. Researchers who use ABER carefully design and plan their use of form and format with the research purpose and audience in mind. Although much of the criteria for evaluating ABER is similar to ABR (e.g., generating new questions, revealing what was previously unknown or little discussed), “incisiveness” is one criterion that differs. According to Barone and Eisner (2006), this criterion refers to the ability of the research to get to the “heart” of the issue, addressing what is educationally significant. For more introduction to ABER, see Largo’s chapter in this volume.

Arts-informed research is distinguished from arts-based research in that it is a research design where art and/or artistic practices are not central to the research process but instead are used to influence another methodological approach or method (e.g., Shannon-Baker, 2015a). Art is incorporated to enhance or supplement the other approach or method such as an interview or survey (Cole and Knowles, 2008). Researchers use arts-informed research to find alternative representations for information and make research studies more accessible. Arts-informed research shares many of the same concerns for integrity and quality as arts-based researchers do, such as the importance of researcher’s self-reflection throughout the process and being open to the creative process in using art. Arts-informed research could address Ardra Cole’s (2004) call for “words plump and dripping with life juice, compelling and evocative images, representations that drew readers and viewers in to experience the research ‘text’” (p. 16). However, arts-informed inquiry differs in its concerns for accessibility for the audience. Whereas arts-based research can create art for the sake of scholarship, or more standalone art pieces based in extensive artistic training (Cahnmann-Taylor, 2007), arts-informed research does not produce art for this purpose. Instead, these researchers are more concerned with the advancement of knowledge than aesthetically advanced art (Cole, 2004; Cole and Knowles, 2008). As a result, some researchers may begin with arts-informed approaches as they develop their technical skills in the artistic practices they use.

Framing mixed methods research

MMR is a philosophically grounded research methodology that meaningfully mixes or integrates two or more methodologies into a single study to enhance, deepen, or complicate one’s inquiry beyond what would be possible through using a single methodology (Shannon-Baker, 2015a; Shannon-Baker and Edwards, 2018). Notably, we do not limit MMR to the mixing of just qualitative and quantitative methodologies as others in the field do (e.g., Johnson and Christensen, 2008; Tashakkori et al., 2021). Instead, we frame MMR as the integration of “two or more methodologies” to be open to methodologies beyond the qualitative-quantitative binary such as arts-based research. For a broader introduction to other ways of defining MMR, see the chapter in this volume by Plano Clark et al. *Integration* refers to connecting, building, merging, or mixing data and/or findings in a mixed methods study (Creswell and Plano Clark, 2017). Through connecting, an initial set of data can be used to identify who to recruit for subsequent data collection, such as using survey findings to recruit for an in-depth qualitative phase. Through building, a researcher uses the findings from one approach to inform another, such as using interview data to create instruments for a second data collection phase. Through merging, a researcher comparatively analyses multiple data sets to explore similarities and differences in the data as well as to corroborate research findings between data sets (Creswell and Plano Clark, 2017). *Data transformation* is a type of integration where one form of data is converted into a different type of data, such as changing interview data into visual or performing art (e.g., Bagley and Castro-Salazar, 2012). One common form of data transformation is *quantitizing* where another data form is transformed into quantitative data to be statistically analyzed (e.g., Akkus, 2013). Integration is possible through the complementarity thesis, which believes that multiple approaches can be integrated to build upon and balance their strengths and limitations (Tashakkori et al., 2021). “If one relies on just one lens of inquiry to view any educational (or other) phenomena, one can never know the extent to which the lens itself distorts what is seen” (Haney et al., 2004, p. 268). MMR is used to investigate complex phenomena and/or contexts that would not be fully honored, captured, or explored through a single methodological approach (Poth, 2018).

Defining arts-based and arts-informed mixed methods research

Arts-based mixed methods research (ABMMR) brings together ABR and MMR to create a philosophically grounded research methodology that meaningfully mixes or integrates arts-based research with at least one other methodology (e.g., quantitative research). Based on the complementarity thesis, this interface builds on the strengths and characteristics of the ABR methodology: centering artistic practices, taking into account form and representation in the art, embracing open-endedness, and engaging in participatory practices (Eisner, 2008; Estrella and Forinash, 2007; Sinner et al., 2006). It also builds upon the strengths and characteristics of the MMR methodology: remaining philosophically grounded, intentionally mixing/integrating multiple methodologies, and bridging multiple methodologies to address complex phenomena and research contexts (Creswell and Plano Clark, 2017; Morse and Niehaus, 2009; Poth, 2018).

We differentiate between ABMMR and *arts-informed mixed methods research* (AIMMR), which is an MMR design that incorporates an art form or artistic practice embedded within another methodological approach in order to supplement or otherwise aid that approach (e.g., Haney et al., 2004; Edwards and Creamer, 2018; O’Connell, 2013). Fig. 2 is one way we envision the differences between AIMMR and ABMMR. On the left side of the figure, AIMMR is portrayed as “air” or arts-informed research as a smaller component embedded in the qualitative strand of a mixed methods study (where the cross sign and double-headed arrow indicate that the mixed methods design could be concurrent or sequential). On the right side of the figure, ABMMR is portrayed where the “ABR” or arts-based research component provides the overarching frame for the whole study. Here qualitative and/or quantitative methods are embedded within this ABR frame. The relationship then for the arts-based research framework to qualitative or

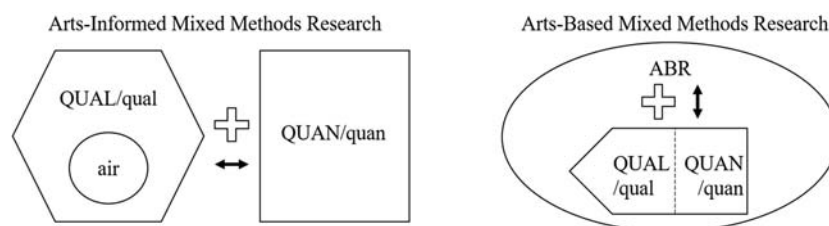


Fig. 2 Visual comparison between arts-informed mixed methods research and arts-based mixed methods research.

quantitative methods could either be concurrent (cross sign) or sequential (double-headed arrow). For visual consistency, we used hexagonal shapes for qualitative research, rectangular shapes for quantitative research, and rounded shapes for the arts-related components.

Researchers may choose an AIMMR approach instead of ABMMR while developing their artistic skills and techniques, when they do not have the resources available to use arts-based approaches, and/or in instances where disciplinary or other evaluative processes are not as supportive as arts-based approaches. For example, Shannon-Baker (2015a) used an AIMMR approach in their study about culture shock among students participating in a study abroad program. In this study, students' self-portraits were embedded in the qualitative strand of a concurrent study that combined qualitative written reflections and a quantitative scale. This study was used as a pilot study for Shannon-Baker's (2015b) dissertation that used ABMMR. Whereas the self-portrait was a smaller component of the pilot study, Shannon-Baker expanded the study to include three self-portraits over the duration of the program to center their descriptive importance in the study. Students also analyzed their own arts-based data from the portraits along with other data. Shannon-Baker used these methodological changes to position the arts-based component as its own methodological strand alongside qualitative and quantitative research, making it a three-strand concurrent mixed methods study.

It is important to note that mixed methods studies about arts-based interventions are not necessarily examples of either AIMMR or ABMMR (e.g., Coppola et al., 2017; Gowda et al., 2018). Making art or artistic practices the topic of a study does not necessarily mean that researchers use that art as part of the methods in their study. Catching this distinction requires a close reading of the reported methods to see if/how art and artistic practices are used in the research methods to determine if a study uses AIMMR or ABMMR approaches.

To compare arts-based, arts-informed, mixed methods, arts-based mixed methods, and arts-informed mixed methods research, see Table 1. This table compares these methodologies across four dimensions: definition, characteristics, rationales for using, and challenges to using.

Integrating art and artistic practices into mixed methods studies in education

Educational researchers integrate art and artistic practices in various ways in their mixed methods studies. To showcase some of these possibilities, we highlight several examples in this section including what art and/or artistic practices were integrated and why.

Embedding art within other methodological approaches: examples of AIMMR

Educational researchers can use an AIMMR approach to integrate art within another methodology. In an example study that resembles a convergent mixed methods design (where both qualitative and quantitative data are collected relatively simultaneously without influencing one another; Creswell and Plano Clark, 2017), Ubha and Cahill (2014) embedded drawings in their qualitative interviews. They studied attachment theory, which looks at how we build relationships with others especially between children and adults, in a primary school context. They implemented a 10-week intervention for five children who had been identified by a teacher as being "passive and withdrawn" from others (p. 6). They evaluated the intervention based on pre- and post-intervention data from quantitative questionnaires and qualitative interviews with the children that included a drawing activity. Ubha and Cahill used the Kinetic Family Drawing activity using the prompt "draw everyone in the family doing something" (p. 8) because drawings are "fun and non-threatening; it also does not require literacy or verbal skills" and because drawings help in "generating and refining hypotheses" (p. 8). This is an example of AIMMR for how the drawings primarily served the qualitative interview as a *visual elicitation method*, or a method that uses visuals (already created or newly created by participants) to promote deeper conversations.

In another AIMMR example, Syeeda (2015) used a mixed methods case study approach (which combines MMR with a case study design that focuses data collection on a specific case; Cook and Kamalodeen, 2020; see also Plano Clark et al.'s chapter on MMR and case study in this volume) to explore secondary students' attitudes toward math and whether their demographics (e.g., age, gender, linguistic background) related to their attitudes. In this example, she included open-ended items in a questionnaire that allowed students to either write or draw their perceptions about math and its impact on their lives. Syeeda included drawings in the study as an "alternative medium of communication [for students] to exhibit their feelings and thoughts" (p. 234). This questionnaire was the second phase of the study that followed initial qualitative focus groups. She found that less than 10% of the 187 students who filled out the questionnaire chose to draw, possibly due to the short amount of in-person time they had to fill out the questionnaire.

Table 1 Comparing arts-based and arts-informed research, mixed methods research, and arts-based and arts-informed mixed methods research.

	<i>Arts-based research (ABR)</i>	<i>Arts-informed research (AIR)</i>	<i>Mixed methods research (MMR)</i>	<i>Arts-based mixed methods research (ABMMR)</i>	<i>Arts-informed mixed methods research (AIMMR)</i>
Definition	A research methodology that centers art and artistic practices throughout the inquiry	A research design that uses art and/or artistic practices to supplement another research methodology or method	A philosophically grounded research methodology that meaningfully integrates two or more methodologies	A philosophically grounded research methodology that centers art and artistic practices from ABR while meaningfully integrating another methodology	A mixed methods research design that uses art and/or artistic practices to supplement another research methodology or method
Characteristics	Create and use art throughout the study; engage in reflexivity; emphasize esthetics; concern about form, representation, processes, and practices; embrace ambiguity and openness	Create and/or use art; engage in reflexivity; concern for the accessibility of art; emphasize art for the purpose of advancing knowledge	Embrace the complementarity thesis; meaningfully integrate data and/or methods through mixing, connecting, and/or building; focus on pragmatic applications	Embrace the complementarity thesis; create and use art throughout the study; engage in reflexivity; meaningfully integrate data and/or methods through mixing, connecting, and/or building; embrace ambiguity and openness with a focus on pragmatic applications where possible	Embrace the complementarity thesis; create and/or use art; engage in reflexivity; concern for the accessibility of art; meaningfully integrate data and/or methods through mixing, connecting, and/or building; emphasize art for the purpose of advancing knowledge
Rationales for using	Provide alternative forms of representation and communication; promote agency especially in marginalized communities/groups	Provide alternative forms of representation and communication especially with those who have limited verbal/language skills; promote agency especially in marginalized communities/groups; use when lacking training, resources, and/or support for ABR	Enhance, deepen, and/or complicate the phenomena being studied; address research contexts and/or design complexities	Provide alternative forms of representation and communication; promote agency especially in marginalized communities/groups; enhance, deepen, and/or complicate the phenomena being studied; address research contexts and/or design complexities; connect art elements and research topic	Enhance, deepen, and/or complicate the phenomena being studied; provide alternative forms of representation and communication especially with those who have limited verbal/language skills; promote agency especially in marginalized communities/groups; use when lacking training, resources, and/or support for ABMMR
Challenges to using	Evaluation of art and artistic practices; amount of aesthetic and arts-based training; mystifying qualities of art; lack of or minimal support for using art and artistic practices; lack of clear terminology in published work	Evaluation of art and artistic practices; amount of art-related training; mystifying qualities of art; lack of or minimal support for using art and/or artistic practices; lack of clear terminology in published work	Time and resources needed to conduct; amount of methodological training in each area used; lack of or minimal support for using multiple approaches in a single study; lack of clear terminology in published work	Amount of training in ABR, MMR, and other approaches used; clarity of methods/practices in published work; lack of or minimal support for ABR and MMR approaches; lack of clear terminology in published work; tension between embracing openness and ambiguity versus pragmatism; lack of consistent evaluation criteria	Training in AIR, MMR, and other approaches; clarity of methods/practices; lack of or minimal support for AIR and MMR approaches; lack of clear terminology in published work; lack of consistent evaluation criteria

(25 min) or because, according to Syyseda, students become “less enthusiastic about drawing as they grow and move up in school” (p. 247). In this case, although Syyseda’s focus in the article is on the importance of using visual data in mixed methods studies in education, making the drawings optional means this study might be better identified as an arts-informed mixed methods study.

Centering artistic practices in a mixed methods study: examples of ABMMR

Educational researchers can also center the artmaking and artistic practices in their mixed methods studies. From nursing education, Ridgway et al. (2018) used a longitudinal mixed methods design to study nurses’ attitudes toward and perceptions of older people during their undergraduate nursing program. Three times per year for three years, the researchers collected data from a survey on attitudes toward older people and a drawing of a person at 75 years of age. The researchers also used a scale on attitudinal change and the drawings as a form of visual elicitation during an interview. These data were collected as part of the students’ coursework and clinical field experiences that included typical content in “aging and disease processes [...] in accordance with the European Union requirements” (p. 64). Although Ridgway et al. used the drawings as a form of visual elicitation, we characterize this study as an arts-based mixed methods study because of the central role that the drawings played in understanding the nursing students’ perceptions.

Ridgway et al. (2018) used drawings and visual elicitation due to the prevalence of visual tropes around aging such as wrinkles, gray hair, and mobility aids. Visuals can help articulate ideas that are difficult to put into words and can help to identify implicit biases (Rose, 2016). Ridgway et al. (2018) analyzed the drawings by identifying broad categories from the visual and/or the accompanying written description. These findings were then integrated through comparison analyses with quantitative findings from the attitudinal change scale and the students’ scores on an attitudinal scale toward older people. These integration analyses identified meta-themes across the three data sets. Findings from analyzing the visual elicitation discussions during the interviews were then used to supplement these meta-themes. Combining visual methods with the quantitative methods “provided a clearer picture” of the students’ attitudes (p. 67). The drawings in particular “provided a powerful discourse relating to the participants’ social constructs of aging” (p. 67).

From the field of dance education (preparing future dance teachers), Wilson and Moffett (2017) evaluated the use of an arts-based research project called “Same Story, Different Countries” that explored racism in the USA and South Africa in a mixed methods study that combined qualitative and dance as an embodied form of arts-based “data collection, analysis, and representation” (p. 139). They combined these approaches for the purpose of triangulation and to address the research purpose through multiple perspectives. Over the 15 months of the project, they collected and analyzed qualitative data included group discussions, reflective writing, document analyses, final project transcriptions, and a final open-ended survey four months later. The arts-based data included “movement exploration, improvisation, performance tasks, embodied storytelling and movement workshops” (p. 139) as well as the final project performance. The research culminated in a 2-h performance that included dance, poetry, music, stories, and visual projections. This study represents the collaborative potential of ABMMR not just across countries but also across professions, combining “choreographers, researchers, educators, visual artists, musicians, poets, and postsecondary dance students” (p. 136). They also found that the arts-based project was a “powerful tool [...] for creating valuable, integrated learning spaces between dance as art and social justice learning, without sacrificing one for the other” (p. 146). They integrated the arts-based approach of dance because it allowed them to integrate “multiple domains of dance learning, knowledge, and practice, teacher and student identities, and connect academics and practitioners across disciplines and communities” between the two countries (p. 146). They argued that integrating arts-based approaches can broaden understandings and possibilities for their field.

Challenges to using and integrating arts-related and mixed methods research

There are many challenges to using arts-related mixed methods approaches for educational researchers to consider. Whereas some challenges speak to both the challenges faced in arts-related research and MMR individually, other challenges come from the potential tensions between integrating these two methodologies. However, we present some strategies for addressing these challenges.

Challenges faced in both arts-related and mixed methods approaches

Level of training

An important challenge in both arts-related and MMR approaches is the issue of training. How much training in the arts is needed? How much training is needed in the other methodologies being used? What about training in mixed methods integration in particular? Some argue for requiring arts-based coursework before using these methods in a study (Piirto, 2002). Art can mystify people (Barone and Eisner, 2006) including those who are asked to make the art as well as the audience for the art. Educational researchers using either AIMMR or ABMMR need to consider their own and the audience’s familiarity and/or comfortability with the artistic practices they plan to use (e.g., Bagnoli, 2009). Researchers can also consider how the design, implementation, and sharing of a study that uses art can help demystify the art aspect (Rolling and Bey, 2012).

Educational researchers may need to spend more time learning and practicing the methods, techniques, skills, and practices they plan to use in an AIMMR or ABMMR study than had they chosen a single methodological approach. However, such educational work need not necessarily be done within a fine arts canon or in formal educational programs. Such stipulations invariably define

whose art practices are legitimate or not (Shannon-Baker and Martinez, 2022). Limited training, at least in the arts-based realm, could be considered a positive aspect; it could limit the bias of the artist in conforming to the expectations of the medium (McNiff, 2008). Without being restricted to the use of specific techniques, arts-related researchers with limited training or those open to veering from method-based norms might be able to develop new methodologies and techniques. However, to more fully consider the ways to methodologically and conceptually integrate art with other methodological approaches, educational researchers need to have some knowledge in the art genre and practices they chose, the other methodological approaches being used, and integration strategies. It is important to note that this challenge faces all mixed methods researchers regardless of whether they plan to use arts-related approaches; mixed methods researchers need to have an in-depth understanding of any of the methodological approaches used to envision the possibilities they have for integrating those approaches.

Clear discussion of methods/practices

Presenting a clear discussion of the methods and/or practices used in an AIMMR or ABMMR study is another challenge facing educational researchers. Grant applications and journals with strict word or page count requirements leave little room to elaborate on all steps in the study. However, without painting a clear picture of the methods, it is difficult for the reader to transfer or replicate the study elsewhere and/or evaluate it (Shannon-Baker and Martinez, 2022). Research diagrams, or visuals that communicate the process and timing for various elements in the research process, can be one helpful way to clearly communicate one's methods in situations where word or page count is limited. Clear diagrams can help identify where and how art and artistic practices are integrated in the study. Additionally, researchers should include explicit rationales behind why particular methods were used (Piantanida et al., 2003). Not only clearly identifying the process but why the researcher made certain methodological decisions provides a series of rationales that can clarify their process to the audience and aid others considering similar methods/practices.

Support for art in a mixed methods study

Whether from supervisors, funders, or publishing outlets, getting support for doing and publishing ABMMR studies can be a challenge. This is because educational researchers who want to use this approach will face others' misconceptions and concerns about both approaches—arts-based research and MMR. Some may perceive arts-based research to lack in rigor, quality, and/or validity (Shannon-Baker and Edwards, 2018). Others have concerns about MMR due to the perceived amount of time and resources needed to conduct such studies rather than using a single methodological approach. These perceptions are due in part to misunderstanding and limited knowledge and/or experience in these methodologies. Similar to the ways in which qualitative and quantitative methodologies have different criteria for judging rigor, quality, and/or validity (or something adjacent), so too do arts-based and MMR methodologies.

To address this challenge, we recommend educational researchers ensure they have a strong grounding in the literature. This literature should help provide rationales and support for choosing an AIMMR or ABMMR approach, how it has been used in the past (or specific methods/practices within the study have been used), and why this approach is best suited for the research purpose and questions. Providing this literature, whether through citations in one's writing or sharing with supervisors, can help address misconceptions and limited knowledge about arts-based, arts-informed, and MMR approaches (and their combinations). When choosing supervisors or research team members, consider what perspective (methodological and otherwise) each brings to the whole of the project. When identifying grant or publishing outlets, consider if they have supported/published works that use arts-based and/or mixed methods approaches.

Using clear terminology

Terminology used is an issue faced in both ABR and MMR approaches (Shannon-Baker and Edwards, 2018). Is the researcher using "art-based," "arts-informed," something else? Is the mixing "mixed methods," "mixed-methods," "multi-method," something else? Using terminology clearly and intentionally helps to align one's research approaches to certain ways of thinking within the field. Also, using clear terminology helps other researchers find studies relevant to their own work (Shannon-Baker and Edwards, 2018). This challenge also addresses how the researcher is theorizing their methods because that connection to other researchers presents one's theoretical and methodological roots for their thinking that then influence how a study develops.

To address this challenge of selecting and using clear nomenclature, we recommend researchers design, implement, and write about their AIMMR and ABMMR studies with a strong grounding in the methodological literature. This starts with presenting clear definitions for both the arts-related and mixed methods approaches being used as we have modeled here. This challenge is made more difficult by the philosophical tension between the embrace of ambiguity in ABR (Eisner, 2008) opposed to the emphasis in MMR for explicit and thorough planning (Creswell and Plano Clark, 2017). In recognition of this tension, researchers may need to be more fluid in their conceptualization of these two approaches or consider adopting a more open-ended definition for MMR (e.g., Greene et al., 1989; Shannon-Baker and Edwards, 2018).

Challenges from the tensions between arts-based research and mixed methods research

Specific to educational researchers who intend to use ABMMR (rather than AIMMR), there are particular challenges that arise from the philosophical and methodological tensions between arts-based research and MMR.

Aim of the research

One tension between arts-based research and MMR is when considering the aim of research. In general, arts-based research tends to inspire more questions than answering questions or providing practical results (Eisner, 2008; Piantanida et al., 2003; Slattery, 2003), although some arts-based researchers have more recently called for more practical applications of this approach (e.g., Eaves, 2014). Conversely, many MMR studies are guided by a pragmatic approach to the research, using the integration of multiple approaches to identify and evaluate practical solutions to everyday problems (Biddle and Schafft, 2015; Johnson and Onwuegbuzie, 2004; Morgan, 2007). This pragmatic approach is particularly present in much of the practitioner-oriented or -initiated educational research. As a result, educational researchers will need to negotiate these philosophical tensions when meaningfully integrating arts-based and mixed methods approaches (Archibald and Gerber, 2018). ABMMR approaches will be especially beneficial if the aim of the study is to explore and/or share counter-narratives from people who have been marginalized (e.g., Bagley and Castro-Salazar, 2012).

Evaluating the research

Another tension between arts-based research and mixed methods research in ABMMR is how to evaluate such studies. Within MMR, several scholars have put forward frameworks that evaluate the clarity of methods particularly during integration and strength of the methodological grounding in defining mixed methods and supporting research decisions (e.g., O’Cathain, 2010; Plano Clark and Ivankova, 2016; Tashakkori et al., 2021). Within arts-based research, there is some debate about the impact of and criteria in evaluating studies. There is some concern that without clear and strong evaluative criteria, arts-based research will glorify “inferior” examples of art (Piirto, 2002, p. 431). Some have argued against firm evaluative criteria for fear that they amount to “methodological policing” (Finley, 2003, p. 284), trying to constrain arts-related approaches within the evaluative criteria established by and for more traditional approaches to research. “If we embrace rigor, if we engage in hegemonic control of the beautiful in research, then we run the risk of missing the opportunity of the dance,” alluding to her epiphanic experience viewing a dance performance (p. 292). Finley argues that trying to validate the arts within a post-positivist framework of science may be futile because these scholars are already biased against it.

Integrating arts-based research and mixed methods research in an ABMMR study will compound these issues. We recommend researchers and their supervisors, reviewers, and evaluators set evaluation criteria that are more contextual to the given study. For example, these criteria can be based on the goal and perspectives/frameworks used in the research – philosophical, theoretical, methodological, and contextual (Leavy, 2017); when mixing arts-based and mixed methods, the evaluation criteria necessarily need to integrate criteria that reflects the methodological approaches used in the study. Some example criteria could be identifying the usefulness for the community or context of the research, how voices were used or represented in the work, the role of the researcher and their relationships with participants (or co-researchers), and how the art evoked empathy, passion, and multiple meanings (Finley, 2003). Additionally, if a researcher provides a clear definition for ABMMR, this definition then indirectly identifies the criteria to evaluate that specific study, such as how the art and artistic practices were centered in the research process, how and when the art was integrated with other methodological approaches, and the researcher’s intentionality behind the art and integration strategies.

Conclusions

In this chapter, we outlined the similarities and differences between arts-based research, arts-informed research, and other arts-related approaches as well as between arts-informed mixed methods research (AIMMR) and arts-based mixed methods research (ABMMR). We showcased several examples of AIMMR and ABMMR in action and the reasons why researchers chose to integrate art and artistic practices with mixed methods. We also discussed the challenges facing educational researchers who use these approaches related to training, clarity of the methods or practices used, support for both the art/artistic practices and mixed methods elements in a study, clear terminology, and the philosophical tensions between arts-based research and MMR related to the overall aim of the research and evaluation. This chapter is intended to provide an introduction into the uses, benefits, and challenges to incorporating art and artistic practices in mixed methods studies in educational research.

Acknowledgment

The authors would like to thank Cherie D. Edwards for her feedback on an earlier draft of this chapter.

References

- Akkus, H., 2013. Pre-service secondary science teachers’ images about themselves as science teachers. *J. Baltic Sci. Educ.* 12, 249–260.
- Archibald, M.M., Gerber, N., 2018. Arts and mixed methods research: an innovative methodological merger. *Am. Behav. Sci.* 62, 956–977.
- Austin, D., Forinash, M., 2005. Arts-based research. In: Wheeler, B.L. (Ed.), *Music Therapy Research*, second ed. Barcelona, Gilsum, NH, pp. 458–471.
- Bagley, C., Castro-Salazar, R., 2012. Critical arts-based research in education: performing undocumented historias. *Br. Educ. Res. J.* 38, 239–260.

- Bagnoli, A., 2009. Beyond the standard interview: the use of graphic elicitation and arts-based methods. *Qual. Res.* 9, 547–570.
- Barone, T., Eisner, E., 2006. Arts-based educational research. In: Green, J.L., Camilli, G., Elmore, P.B. (Eds.), *Handbook of Complementary Methods in Education Research*. Routledge, London, pp. 95–109.
- Biddle, C., Schafft, K.A., 2015. Axiology and anomaly in the practice of mixed methods work: pragmatism, valuation, and the transformative paradigm. *J. Mix. Methods Res.* 9, 320–334.
- Cahnmann-Taylor, M., 2007. Arts-based research: histories and new directions. In: Cahnmann-Taylor, M., Siegesmund, R. (Eds.), *Arts-based Research in Education*. Routledge, London, pp. 3–15.
- Cole, A.L., 2004. Provoked by art. In: Cole, A.L., Neilsen, L., Knowles, J.G., Luciani, T. (Eds.), *Provoked by Art: Theorizing Arts-Informed Research*. Backalong Books, Halifax, Nova Scotia, pp. 11–17.
- Cole, A.L., Knowles, J.G., 2008. Arts-informed research. In: Knowles, J.G., Cole, A.L. (Eds.), *Handbook of the Arts in Qualitative Research*. Sage, Thousand Oaks, CA, pp. 55–70.
- Cook, L.D., Kamalodeen, V.J., 2020. Combining mixed methods and case study research (MM+CSR) to give mixed methods case study designs. *Caribbean J. Mixed Methods Res.* 1, 47–76.
- Coppola, S., Miao, A.F., Allmendinger, C., Zhang, W., 2017. Art in occupational therapy education: an exploratory mixed-methods study of an arts-based module. *The Open J. Occup. Therapy* 5, 1–18.
- Creswell, J.W., Plano Clark, V.L., 2017. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Thousand Oaks, CA.
- DeCuir-Gunby, J.T., Schutz, P.A., 2017. *Developing a Mixed Methods Proposal: A Practical Guide for Beginning Researchers*. Sage, Thousand Oaks, CA.
- Eaves, S., 2014. From art for arts sake to art as means of knowing: a rationale for advancing arts-based methods in research, practice and pedagogy. *Electron. J. Bus. Res. Methods* 12, 147–160.
- Edwards, C.D., Creamer, E.G., 2018. Cultural transition, individuality, and the arts: examining individuality in acculturation through a mixed methods approach to arts-informed research. *Int. J. Educ. Social Sci. Res.* 1, 30–40.
- Eisner, E., 2006. Does arts-based research have a future? Inaugural lecture for the first European conference on arts-based research Belfast, Northern Ireland, June 2005. *Stud. Art Educ.* 48, 9–18.
- Eisner, E., 2008. Persistent tensions in arts-based research. In: Cahnmann-Taylor, M., Siegesmund, R. (Eds.), *Arts-based Research in Education: Foundations for Practice*. Routledge, London, pp. 16–27.
- Ellsworth, E., 2005. *Places of Learning: Media, Architecture, Pedagogy*. Routledge, New York.
- Estrella, K., Forinash, M., 2007. Narrative inquiry and arts-based inquiry: multinarrative perspectives. *J. Humanist. Psychol.* 47, 376–383.
- Finley, S., 2003. Arts-based inquiry in QI: seven years from crisis to guerrilla warfare. *Qual. Inq.* 9, 281–296.
- Gowda, D., Dubroff, R., Williams, A., Swan-Sein, A., Capello, C., 2018. Art as sanctuary: a four-year mixed-methods evaluation of a visual art course addressing uncertainty through reflection. *Acad. Med.* 93, S8–S13.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11, 255–274.
- Haney, W., Russell, M., Bebell, D., 2004. Drawing on education: using drawings to document schooling and support change. *Harv. Educ. Rev.* 74, 241–272.
- Irwin, R.L., Springgay, S., 2008. A/r/tography as practice-based research. In: Springgay, S., Irwin, R.L., Leggo, C., Gouzouasis, P. (Eds.), *Being with A/r/tography*. Sense Publishers, Rotterdam, The Netherlands, pp. xix–xxxiii.
- Johnson, R.B., Christensen, L., 2008. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, third ed. Sage, Thousand Oaks, CA.
- Johnson, R.B., Onwuegbuzie, A.J., 2004. Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26.
- Leavy, P., 2009. *Method Meets Art: Arts-Based Research Practice*. Guilford Press, New York.
- Leavy, P., 2017. *Research Design: Quantitative, Qualitative, Mixed Methods, Arts-Based, and Community-Based Participatory Research Approaches*. Guilford Publications, New York, NY.
- Ledger, A., Edwards, J., 2011. Arts-based research practices in music therapy research: existing and potential developments. *Arts Psychother.* 38, 312–317.
- Lincoln, Y.S., 1995. Emerging criteria for quality in qualitative and interpretive research. *Qual. Inq.* 1, 275–289.
- McNiff, S., 2008. Art-based research. In: Knowles, J.G., Cole, A.L. (Eds.), *Handbook of the Arts in Qualitative Research*. Sage, Thousand Oaks, CA, pp. 29–40.
- Morgan, D.L., 2007. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *J. Mix. Methods Res.* 1, 48–76.
- Morse, J.M., Niehaus, L., 2009. *Mixed Method Design: Principles and Procedures*. Left Coast Press, Walnut Creek, Calif.
- O’Cathain, A., 2010. Assessing the quality of mixed methods research: toward a comprehensive framework. In: Tashakkori, A., Teddlie, C. (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 531–555.
- O’Connell, R., 2013. The use of visual methods with children in a mixed methods study of family food practices. *Int. J. Soc. Res. Methodol.* 16, 31–46.
- Piantanida, M., McMahon, P.L., Garman, N.B., 2003. Sculpting the contours of arts-based educational research within a discourse community. *Qual. Inq.* 9, 182–191.
- Piirto, J., 2002. The question of quality and qualifications: writing inferior poems as qualitative research. *Int. J. Qual. Stud. Educ.* 15, 431–445.
- Plano Clark, V.L., Ivankova, N.V., 2016. *Mixed Methods Research: A Guide to the Field*. Sage, Thousand Oaks, CA.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research*. Sage, Thousand Oaks, CA.
- Ridgway, V., Mason-Whitehead, E., McIntosh-Scott, A., 2018. Visual perceptions of ageing: a longitudinal mixed methods study of UK undergraduate student nurses’ attitudes and perceptions towards older people. *Nurse Educ. Pract.* 33, 63–69.
- Rolling Jr, J. H., & Bey, S. (Interviewees). (2012). Interview on Arts-Based Research [Interview transcript]. Retrieved from YouTube: <https://www.youtube.com/watch?v=zd3ezkSPzAE>.
- Rose, G., 2016. *Visual Methodologies: An Introduction to Researching with Visual Materials*. Sage, London, UK.
- Shannon-Baker, P., 2015a. “But I wanted to appear happy”: how using arts-informed and mixed methods approaches complicate qualitatively driven research on culture shock. *Int. J. Qual. Methods* 14, 34–52.
- Shannon-Baker, P., 2015b. *Microaggressions, Self-Segregation, and Gender Performance: Exploring Undergraduate Students’ Culture Shock in a Study Abroad Program* (Doctoral Dissertation). University of Cincinnati, Cincinnati, OH.
- Shannon-Baker, P., 2022. Philosophical underpinnings of mixed methods research in education. In: Tierney, R., Rizvi, F., Smith, G. (Eds.), *International Encyclopedia of Education*, fourth ed. Elsevier, Oxford (forthcoming).
- Shannon-Baker, P., Edwards, C., 2018. The affordances and challenges to incorporating visual methods in mixed methods research. *Am. Behav. Sci.* 62, 935–955.
- Shannon-Baker, P., Martinez, C., 2022. Integrating arts-based approaches in mixed methods research with diverse stakeholders: a critically reflective review from two teacher-researchers. In: Hitchcock, J., Onwuegbuzie, A. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge (forthcoming).
- Sinner, A., Leggo, C., Irwin, R.L., Grauer, K., 2006. Arts-based educational research dissertations: reviewing the practices of new scholars. *Can. J. Educ.* 29, 1223–1270.
- Slattery, P., 2003. Troubling the contours of arts-based educational research. *Qual. Inq.* 9, 192–197.
- Syeda, F., 2015. A picture is worth a thousand words: examining learners’ illustrations to understand attitudes towards mathematics. *Exchange: Interdisciplin. Res. J.* 2, 234–248.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, second ed. Sage, Thousand Oaks, CA.
- Ubha, N., Cahill, S., 2014. Building secure attachments for primary school children: a mixed methods study. *Educ. Psychol. Pract.* 30, 272–292.
- Wilson, L., Moffett, A.-T., 2017. Building bridges for dance through arts-based research. *Res. Dance Educ.* 18, 135–149.

Developing qualitative research proposals

Marcus B. Weaver-Hightower, Virginia Tech, Blacksburg, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Before starting writing	711
What makes a proposal truly qualitative?	712
The format	712
Writing it: what to include in your qualitative proposal	713
Beginnings	713
The middle, part I: setting up your design	714
Literature	714
Theoretical framework	715
Methodology	715
The middle, part II: your design	715
Research questions	716
Sites and/or participants	716
Data collection	716
Data analysis	717
Validity and reliability	717
Ethics	718
Delimitations: what are you <i>not</i> doing?	718
Pilot studies and previous data (if any)	718
Researcher disclosures	718
Logistics	719
Endings	719
Revise, revise, share, revise, share, revise	719
Conclusion	720
References	720

Qualitative research proposals primarily aim to excite readers about possibilities. Possibilities of new understandings about thorny, complex issues. Possibilities of justice for those long denied it. Possibilities of new insight into the many ways that humans live and believe. To be clear, qualitative proposals don't answer one's questions; the study does that. Proposals, rather, offer the possibilities of answers an academic field needs to know.

Exciting readers about a study's possibilities doesn't happen accidentally. Instead, planning a solid study and carefully writing to readers' interests generate that excitement—or at least their eventual approval of the proposal. In this chapter, then, I provide planning and writing advice for those hoping to develop a qualitative research proposal full of possibilities.

In this chapter I provide advice for writing qualitative research proposals, whether for your dissertation, a grant proposal, or any other purpose. My advice covers the things you should do before writing the proposal, like analyzing your audience, understanding what makes a qualitative proposal unique from other approaches, and some format concerns. I then discuss writing the proposal itself, including all the parts that make up a standard qualitative proposal—from title to conclusion and everything between. Finally I exhort readers to revise many times and seek feedback to make the proposal maximally clear, engaging, and persuasive.

One caveat for my advice: While I've got about 20 years' experience writing and reading qualitative proposals of all types, my way isn't the only way. (For instance, not everyone would accept my using contractions like *isn't*.) Your country, your institution, or your advisor may have specific traditions, rules, or preferences to follow. So in addition to considering my advice and the authors cited here, I urge you to speak directly with those who will review your proposal, which I say more about next.

Before starting writing

You must discover and keep in mind two things as you begin writing: (a) the audience for the proposal and (b) what that audience wants to know (Weaver-Hightower, 2019, Chapter 12). Naturally, many scholarly activities require proposals, including theses and dissertations, internal and external grants, published books, sabbaticals, evaluation projects, and human subjects and ethics reviews. Slightly different audiences read each of these, and readers' expectations stem from what you are asking for—permission to do the study, time off (or on), or funding to make it all happen. You must convince them that their time and treasure will be well spent on you. Yes, the quality of your ideas—their significance, potential implications, originality, and so on—ought to be the deciding factor, but don't overlook the necessity to tailor your proposal to speak to readers' needs and concerns. You can't please everyone's

tastes, but you can address why, for example, you've used a theorist you know one of your committee members hates, or you can carefully explain why you need money for in-person travel to a particular archive rather than just using their online catalog.

Sometimes you know your specific reader; sometimes you don't. For a dissertation proposal, you likely know most of your committee members and your advisor, so you might have a good sense of what they like, how they themselves write and research, and how they work with students. If you don't know a particular member that well—say, a member from outside the department or a replacement for a retired member—take them to a beverage-based meeting (you buy, or at least offer!) and get to know them. You might similarly know the people reading a sabbatical proposal, but get to know anyone you don't already, and explain what you're hoping to do before it comes to a vote.

For most grant proposals or ethics reviews, on the other hand, you are unlikely to know the readers in advance. You therefore can't write to their specific desires or avoid their pet peeves. You can, however, do your homework on what the organization wants from a proposal or what they have funded previously. Phone the grants officer and ask for advice. Most websites also now contain clear directions and winning proposal examples. Make sure to read such materials and follow directions exactly. If, for instance, the directions say your bio should be two pages, don't let even one line spill onto a third page.

You can also safely write to a "general audience" of scholars when writing a proposal for unknown readers. Assume that general audiences share three important characteristics: intelligence and curiosity, local knowledge, and significant skepticism when they read (Weaver-Hightower, 2019). For readers like that, write without "talking down" to them. No need to define *education* when used in a standard sense, but don't assume that everyone knows a specific philosopher's definition of education, either. General readers' local knowledge means you cannot assume they know about current events or the specifics of another country you researched. Better to write as if a general reader lives elsewhere, meaning that you'll have to explain things. Readers' skepticism, finally, requires you to defend your ideas and anticipate their "yes, buts" throughout. Scholars read analytically, against the grain, looking for alternative explanations. Make sure your proposal honors that by giving reasons and evidence for why your theoretical framework and methods will work best.

What makes a proposal truly qualitative?

Speaking of frameworks and methods, it behooves us to consider what makes a qualitative proposal particularly qualitative. Just doing interviews is not enough. Analyzing a handful of open-ended questions on an otherwise quantitative survey is not enough. Coding one's data is not enough. Doing a qualitative study and a proposal for it requires that you embark from a qualitative mindset with qualitative goals.

Maxwell (2013, Chapter 2) outlines a great set of qualitative mindset goals—"intellectual" rather than "practical" to emphasize that these are not the what-I-will-do-with-the-knowledge goals. In short, a study fits qualitative paradigms if it seeks to understand (a) the meanings participants hold about the events and objects around them, (b) the particular contexts participants face and how those influence events, and (c) the processes taking place. It also fits if the study intends to develop (d) ground-up theories for unanticipated phenomena and influences or (e) causal explanations about phenomena based on context or belief rather than, say, demographic categories. One distinguishes qualitative and quantitative studies based on "the distinction between thinking of the world in terms of variables and correlations," as a quantitative study does, versus "events and processes" on which qualitative studies focus (Maxwell, 2010: 477). It might seem like an obvious thing, but ensure that you are actually doing a qualitative study, with a mindset focused on beliefs, contexts, processes, and events, before you propose a qualitative study. If you're just borrowing a qualitative method or two, but your analysis has a quantitative focus, like code frequencies, correlating demographic factors, and so on—basically "quantitizing" your qualitatively collected data (e.g., Sandelowski et al., 2009)—you might reconsider calling your study "quantitative" or "mixed methods."

The format

Also ask your advisor or grant officers about the proposal format. Even though we currently enjoy a detente in the "paradigm wars" (Gage, 1989) and qualitative methodologists have established scores of uniquely qualitative approaches to valid research, quantitative research's dominance still burdens proposal formats for many grant applicants and graduate students. Nowhere is this truer than the dreaded first-three-chapters dissertation proposal. In many academic departments, tradition requires the proposal be the first three chapters of a five-chapter dissertation. The student writes the introductory chapter, including the problem statement, rationale for the study, and overview; a literature review chapter focused on all other studies that have investigated the topic; and a methods chapter, which lays out the exact procedures to be undertaken. This format works well for quantitative studies that simply require one to then implement their survey or their controlled experiment, write what statistics were generated in the fourth chapter, and discuss whether their hypothesis was confirmed and possibilities for why in the fifth chapter.

Though sensible for quantitative studies, in my experience the first-three-chapter proposal format usually works terribly for qualitative studies. It misunderstands qualitative research's nature. Qualitative research doesn't start from a hypothesis that one will simply confirm or reject. Sure, we have hunches about what we will find, but most will have those hunches blown to smithereens quite early in a study. Questions—even entire topics—emerge and change when qualitative researchers encounter the field's realities. One's proposal can hardly be expected to anticipate the literature that will illuminate those changes. Methods may have to shift, too, as new questions emerge and access to people and places shifts. The collision with reality thus creates two possible problems for those who have written a three-chapter proposal. They can either scrap most of their hard work on those chapters to reflect the actual study, or, if they leave the three chapters as is, they end up with a kind of schizophrenic study whose topic, literature, and

methods don't match their conclusions and themes. I get why people like the three-chapter format: it's comfortable and familiar, and they think, "Hey, you're more than halfway done with your dissertation when you propose!" It's just not true for high-quality qualitative research, though. Using that format adds extra work for qualitative researchers.

Obviously, I alone can't make your department shift its approach to proposals if they love the three-chapter model. Show this chapter to your advisor or the department chair, though, and see if you might do something more appropriate to qualitative research. I agree with [Maxwell \(2013: 127\)](#), who suggests a 5000-word (roughly 20 pages double-spaced) proposal works well for a qualitative study. In some countries, like the UK or some European countries, something more like 6 or 7 pages are preferred, but I personally would find it hard to be confident of the student's direction at so short a length. On the other end, I would suggest nothing over 40 pages, because then you're not focused enough, not investing the time and thought needed to be concise, or just showing off how much you've read! A singular, coherent prospectus with plenty of headings works better than "chapters" in so short a length. That length also provides plenty of room for the required information I suggest below, but it doesn't over-invest you in literatures and procedures that become irrelevant when your study design inevitably "emerges" in new directions.

A quick note on format for those considering an article-based dissertation, consisting of three or more "publishable" articles (if that is an option at your university). The proposal naturally needs to fit that unique kind of dissertation. Most of what I say in this chapter still applies, but you must describe the setup of all three (or four or however many) articles *individually*, including literature review, proposed research questions, methods, ethics, and so on. This might take much more word count than describing just one study. The article-based dissertation proposal also requires a section—perhaps two, one at the beginning and one at the end—that explains how the various articles relate to one another and how, together, they build toward contributing to a larger research problem. As with other approaches, ask your advisor for guidelines and examples of such proposals.

Writing it: what to include in your qualitative proposal

Whatever their limitations, first-three-chapter proposals include all the basic requirements for any proposal, including the 20- to 40-page dissertation proposal format I recommend: an introduction, a review of some related literature, and the intended methodology and procedures (see also [Marshall et al., 2022: 80](#)). In this section I discuss each of these general parts so that you know what to include.

First, though, let's review your proposal's two main goals: "convincing a supervisor and committee [or grant reviewer or editor] that the topic is worth researching and the candidate has the wherewithal to carry it out" ([Kilbourn, 2006: 529](#)). Those goals might sound simple, but meeting them requires telling readers key things in a coherent way, showing how your design elements interrelate seamlessly. Still, while easier said than done, keep those goals in mind as you decide what to include.

Don't treat the following as a fixed recipe, for any proposal must be tailored to the unique needs of the study. Nor do you have to compose in this order, start to finish; write sections as you feel prepared for them or as the ideas pop into your head. Your final proposal should probably present them in roughly this order, though—unless you have compelling reasons not to—because readers expect information to flow this way.

Beginnings

Go ahead, have an interesting title. Nothing too gimmicky nor obscure ([Wolcott, 2009: 123–124](#)), but your title can usually be more interesting at the proposal stage than at others. Your title creates your first impression, so make it the impression you want. Serious and grave? Engaging? Fun? Innovative? Your title conveys a great deal. This can include methodology, if a key aspect. [Creswell and Plano Clark \(2018: 146\)](#), for example, suggest explicitly including "mixed methods" in the title of any mixed methods study. I like doing this for any specialized qualitative approach (autoethnography, critical ethnography, transcendental phenomenology, and so on; many are covered in chapters in this volume), because it primes readers to assess the proposal with those approaches in mind. Also make sure the title conveys the study's qualitative nature; avoid "factors," "measuring," and other quantitative terms in favor of "beliefs" and "exploring." If you know the specificities of location and participants already, include those to avoid overgeneralization. For instance, "The Construction of Identity in Authors Theater: A Case Study of one Fourth Grade Classroom in the Midwest," is much better with the subtitle than without because it doesn't overpromise the amount of data and diversity you will have. Just do consider it a working title, subject to frequent change as your ideas evolve. Reflecting later on this titular evolution can provide insights into your project's development ([Peshkin, 1985](#)).

Once you get the title right (for right now), focus on the proposal's most important paragraph: the first. My best advice is get right to the point! Make the very first sentence plunge headlong into your topic or aim. Rather than trite gimmicks like "Merriam-Webster's dictionary defines ..." or burying readers in 200 years of history leading up to the topic, hit readers explicitly with the subject. Consider starting with a sentence like "The purpose of this study (chapter, monograph, article) is ..." ([Wolcott, 2009: 13](#)). Readers will be wondering this most urgently, so why not give it immediately? I'm decidedly not saying the opening has to be boring or clinical. Qualitative research gives you the opportunity to introduce observations and quotations, representative events from the field, quotes from participants, or piquing anecdotes that catapult readers into the topic. Just make such examples clearly related so that readers aren't still wondering what your subject is five, six, seven or more pages into your proposal.

Consider the following introduction to an example dissertation proposal:

Educators desperately need alternative methods of distributing and receiving professional development. Research reveals that “good professional development” should be designed around teacher interest, connected to teacher’s current knowledge, and supportive and nurturing of teachers’ intellectual capacity (Corcoran, 1995; Little, 1993). In other words, professional development should be learner-centered. However, providing “good professional development” for teachers is extremely challenging for local school districts. Teacher interest, knowledge, and intellectual capacity vary widely within and among schools. Offering appropriate professional development requires a large investment in time and money. This type of investment is often neither a priority nor a possibility. Therefore, despite research, school districts typically offer professional development opportunities that are irrelevant and uninteresting for. (Quoted in Maxwell, 2013: 181)

Key in on the first sentence. The topic, professional development, appears in the very first sentence, and in a prime position for emphasis—the sentence’s end. At a svelte 11 words, the sentence gets straight to the point with no flourishes. Teachers also “desperately” need professional development to be high quality, according to the author, establishing the importance and urgency right away. It’s not “stylish,” but this opening works well.

This sample introduction also highlights the next thing you need to accomplish: establishing the research problem. (If the word *problem* sounds like deficit thinking, replace it with *statement* or something more to your liking.) Clearly state what needs the reader’s attention, like inequitable educational outcomes for particular groups, an incorrect way of conceptualizing a phenomenon, or gaps in the field’s knowledge. For the sample above the problem is clear: professional development is “irrelevant and uninteresting,” and something “desperately” needs to be done about it. Make your own problem statement similarly clear and explicit. No beating around the bush or burying it under an avalanche of literature and jargon. Readers want to know your topic’s significance, not just to you and your advancement, but to everyone. As Kilbourn (2006) notes, a key challenge for the proposal writer involves “the need to construct an argument for doing the research in the first place.” Kilbourn also points out that the research should be based on “genuine inquiry” rather than seeking to prove what you already think—and readers can often tell if you’ve already made up your mind about the topic. Why do we—scholars, practitioners, or society in general—need to know the knowledge you plan to produce? What do we already know, and what are you adding to what we know? What usefulness does your study have? Not that it must have immediate application, but it should at least pique the curiosity of more than just you. (This is what a book publisher will be looking for in a prospectus, by the way, to get the all-important answer of “Who will buy this?” [Allen, 2016, Chapter 6].)

The middle, part I: setting up your design

The middle, overall, is the veggie patty in the qualitative proposal sandwich. It’s the “meat” between the beginning and the ending. Honestly, readers care most about this information. While the beginning establishes the problem and need for studying it, the middle—by far the longest part—establishes the study design. Of the two major questions—is this worth doing and are you the one to do it—the middle of the proposal best answers the latter. By discussing what has come before (theory and literature), what you’ll do to answer your new questions (methods), and the logistics (timelines and any budget), you give readers insight into your training, reasonableness, thoughtfulness, and skill.

For convenience I have separated the middle into two parts, the setup for the all-important design and then the design itself. This first section, the setup for your design, largely revolves around what others have said and done, including the literature review, the theoretical framework, and the methodology. You will do much citing of sources in this section, while in the design the focus shifts to your study plans.

Literature

What do we as a field know about the topic already, and what don’t we know? Those basic questions guide a literature review. It can be tempting to spend vast time and word count reviewing literature, particularly as reading and citing others often dominate graduate education and feedback from peers and professors. Resist the feeling that you must have read every article and book ever written on your topic. Don’t drown readers in a forty-page literature review, no matter how much you fear your committee are only testing how much you’ve read. I like how Vagle (2018) calls it a “partial review of the literature” (142). For a qualitative proposal, particularly the 20–40-page version I suggested, just spend a few pages reviewing the literature that supports what you anticipate right now, at the point of entering the field. You can also forecast other bodies of scholarly literature that may become important as you proceed and the design emerges.

For the literature you do include, focus on themes and broad strokes. Only focus at length on individual studies when they have influence, historical value, or close resemblances to your study. Few things bore readers more than the “furniture catalog” literature review (Haywood and Wragg, 1982), where the author summarizes one study after another with little attention to the relationships among studies. Writing instead about themes in the literature, citing the studies as groups or examples, shows that you have digested and synthesized the knowledge in the field rather than just having collected a bunch of studies from a keyword search. Be selective about what studies get put center stage in your proposal.

Make sure to also include studies that contradict your approach or beliefs—and as a reader I always assume some exist. Exploring and taking stands on disputes and tensions puts you in dialogue with others in your field, ultimately giving you an opportunity to advance knowledge by critically assessing the field’s knowledge base. Not only might readers already know of opposing studies and

wonder why you don't address the contradictions, but including them emphasizes your fair mindedness and honesty, another sign that you are the right person for the job ahead.

Theoretical framework

How will you approach theory? Do you have a particular theory you are trying to “test” in your qualitative research? Some researchers do. They apply, say, a model of culturally relevant pedagogy (e.g., [Ladson-Billings, 2021](#)) to their research site to see how the model does or doesn't explain the teaching going on there. Some turn their noses up at *a priori* theory in qualitative research, but I personally believe it has its place. No one really begins a study with a truly blank slate.

Some researchers, not dissimilarly, use the theory to “think with” ([Jackson and Mazzei, 2012](#)), meaning that a theory provides sensitizing concepts that frame their view of the research, a particular way of viewing a phenomenon. So something like New Literacy Studies, for example, views literacy as practices and events shaped by social context and relationships, rather than literacy as fixed in one's ability to decode language (e.g., [Barton et al., 2000](#)). That's not a theory to be tested, but a perspective that influences how one sees the world and sways one's approach to research design.

Of course, some qualitative researchers intend to create theory, such as through grounded theory approaches (e.g., [Charmaz, 2006](#); [Glaser and Strauss, 1967](#)). In these cases, relevant theory either doesn't exist, or the researcher believes that available theories are not adequate or plausible. Even if creating theory, you must still discuss your relationship to theory in your qualitative proposal. What theories have you considered, and why have you rejected them? Do any come close and just need to be tweaked? How will you go about constructing your own theory while sticking closely to participants' meanings and experiences?

Methodology

Speaking of theory, your qualitative proposal should address the intellectual traditions of knowledge production that guide your procedural choices—that is, your methodology. What general approach will you take (methodology) and what specific procedures (methods) will you use? How will you use “evidence and argument” ([Kilbourn, 2006](#)) to make claims about your topic?

Scores of specific methodological approaches huddle under the qualitative umbrella. [Creswell \(2013\)](#) has his classic five: ethnography, case study, phenomenology, grounded theory, and narrative research. Many more are available, like hermeneutics, discourse analysis, content analysis, ethnomethodology, life history, conversation analysis, portraiture, autoethnography, arts-based research, duoethnography, photovoice, task analysis, and poetic inquiry (many, again, covered in other chapters in this volume). You can even just choose “generic” ([Caelli et al., 2003](#)) or “pragmatic” qualitative research ([Savin-Baden and Major, 2013](#), Chapter 11) rather than struggling over which specialty fits. (In my experience, students tend to lose more sleep over which specialty they're doing than over their actual findings.) Regardless, each approach has its own perspective on what counts as evidence, how one collects and analyzes evidence, and even how one should write ([Weaver-Hightower, 2019](#), Chapter 11). Whatever your chosen methodology, describe for readers its basic tenets and views of the social world, the procedures typically involved, and cite the methodologists you rely on.

A quick caution: while increasingly acceptable to use “alternative” methodologies, both in conducting and representing qualitative research (autoethnographies, participatory methods, comics, theater, etc.), know your audience! Have they approved or funded work like this before? Are they knowledgeable about or open to non-traditional research? Not that you should compromise a design that fits your research questions, but know that extra steps and some convincing might be involved.

The middle, part II: your design

Experienced readers—and assume yours are—anticipate seeing certain sections in a qualitative design description, usually in the following order:

- Research question(s)
- Site(s) and/or participant(s)
- Data collection
- Data analysis
- Validity and reliability
- Ethics
- Logistics (timeline, budget, data management)

Include each in your proposal. Most of all, plan these elements out in as much detail as you possibly can.

“But wait,” you might be thinking, “qualitative designs must be emergent! You said so yourself.” Absolutely true. Qualitative proposals have an important contradiction at their heart: you must have a detailed plan even while knowing it will likely change considerably. That's why [Sandelowski and Barroso \(2003: 781\)](#) call qualitative proposals “exercises in imaginative rehearsal.” You must anticipate how your study might go, showing your expertise and willingness to adapt through your detailed design. No one wants to give you money or unleash you on participants if they think you will “wing it” in the field. They want to feel confident that you know enough to at least construct a plausible opening approach.

So how does one both plan in detail and plan for emergence? Plot the ideal approach but frequently use subtle statements like “My initial plan to analyze the data is to ...” or “Emerging evidence might suggest other means of ...” to indicate your awareness of

likely change. Explicitly convey to readers how qualitative designs, by their nature, must be open to emerging realities. Look at how [Sandelowski and Barroso \(2003\)](#) phrased it in their successful National Institute for Nursing Research grant:

As is the case with any qualitative project, the design will be emergent, or highly dependent on the ongoing results of this experimentation as the study progresses. Although we separate data collection, analysis, and interpretation here, the better to communicate our research plans, these processes typically proceed together and strongly influence each other in qualitative studies. Design dictates what the quantitative researcher will do; in contrast, what the qualitative researcher does determines the design. (796)

They then cite other studies that can guide them moving forward, even when the path is unclear. Citing other researchers and their decision processes convey to readers that your plan's gaps aren't anomalous or failings, but rather accepted means of working qualitatively.

So emergence cannot be an excuse not to plan. Even if you don't know yet, for example, how many participants you will interview, at least "explain the *basis* on which you'll make your decision" ([Maxwell, 2013](#): 125; emphasis original). That basis could be things like how many people respond to your recruiting, the richness of the data you're getting, or how many instructors use the practices you're studying. Explaining your thinking will give your audience certainty that you can confidently handle it when things inevitably change.

Research questions

Your research questions are vitally important. They also will likely change a good deal across your project. Because you will confront a reality that might not align with your expectations, some questions might need to change or might not be answerable at all. I myself had ten research questions for my dissertation proposal, but I was only able to answer six and a half of them. It is fine to start with broad questions, knowing that you'll winnow them down and sharpen them as your study progresses.

When composing your question or questions—and no one I know suggests more than three or four, much less ten!—make sure that each demonstrates six characteristics. Each question should (a) be concise, (b) ask a singular question, (c) have a qualitative purpose, (d) reflect your study population, (e) fit the theoretical or methodological orientation, and (f) avoid judgment, evaluation, and practical goals ([Weaver-Hightower, 2019](#): 175–176). Take for example two questions comparing qualitative approaches to the same topic:

Pragmatic qualitative research: What do online teachers see as the incentives and barriers of teaching online?
Grounded theory: What explains (how can we theorize) instructor online teaching behaviour? ([Savin-Baden and Major, 2013](#): 109)

Each asks only one question—though one might quibble over *incentives* and *barriers* creating separate questions in the first—and neither has superfluous words. Each is qualitative in that it examines social behavior and meanings. Each reflects its methodology's tenets, as explicitly labeled. Even though the first is "pragmatic," it doesn't seek to judge or evaluate an existing program, just to solicit what participants believe. Finally, these questions don't name a specific site or participants, but in a proposal you don't always know this information yet. If you do know that information, rephrase these with something like "What do online teachers at Eastern Seaboard University see as the incentives and barriers of teaching online?"

Sites and/or participants

Speaking of your study site and participants, describe these if you know them, or describe how you will determine these after the study or grant gains approval. Provide a rationale for your sites and participants; why can they, particularly, answer your research questions? Explain any reasons for excluding some participants, like only wanting a certain age group or geographic region, say.

Once you have explained why your participants are the right ones, explain how you will (or already did) find them and secure their participation, both at the beginning and throughout the study. Such issues should include any planned recruiting efforts, how you will negotiate entry, how you intend to gain and maintain trust, your role (participant, observer, reciprocity agreements, etc.), and plans for exiting the site responsibly.

Data collection

Data collection (or *generation*, if you prefer; e.g., [Garnham, 2008](#))—alongside data analysis, discussed in the next section—will likely receive the most scrutiny from readers. These provide a sense of how well you understand the topic, the research process, how humans behave, ethics, and how to solicit knowledge. If you propose to collect data you can't access qualitatively—say, how something happens in the brain—or suggest that participants will unflinchingly answer questions that are deeply private or embarrassing, readers will doubt your abilities. So first make sure that you can get the data you need, particularly that your data collection methods match your research questions.

It might just be my personal bias, but I tend to be most skeptical of projects that solely interview participants. Certainly some questions researchers can best answer with interviews, like constructing life histories (e.g., [Seidman, 2006](#)), and many methodologists have well developed methods for interviewing (e.g., [Roulston, 2010](#)). Many, many research questions, though, could be

answered much better by including other methods, as well. I suspect that interviews seem easiest, whether logistically or analytically, because of the directness of asking a question and getting an answer straight from participants. Interviews can be complex and raise many concerns, though (see, e.g., chapters in [Holstein and Gubrium, 2003](#)). Interviews may certainly tell you what participants believe explicitly—or, at least what they will share with an outsider—but interviews may not be as insightful about processes and behaviors.

If your research questions ask about processes and behaviors, consider observation, whether as participant or not. Consider collecting artifacts, as well, including texts participants create or use, media, art, and so on. These have meanings one can glean from how, when, where, and why people use them. My point is simply that observational, textual, and artifactual data can be as or more useful than interviewing, and some readers will wonder why you don't include them if your research questions suggest them. Don't forget, too, that these other methods can be used in combination with interviewing to create more integrated, trustworthy findings—a form of methodological triangulation ([Denzin, 1970](#), Chapter 12).

As with other methods sections, make sure your data collection description includes significant detail. If you plan to conduct interviews, include at least some sample questions to show readers you'll ask the right things. How many participants will you interview, for how long, how many times, and how long between? Will interview questions vary by participants, say by age or role? If you're observing, specify what kinds of events you'll observe. What will you be looking for and how will you know it when you see it? The latter can be quite important, as concepts like *motivation*, *relationship*, or *emotion* can be difficult to spot. Just as with interviews, how many observations will you conduct, how many times, for how long, and how long between? If you are working with artifacts, like documents, where will you find them, how will you get permission to have them, how many do you need, and so on? Again, readers will generally understand that details are hard to know in advance and are subject to change, but giving these details shows you have thought your project through and can anticipate both changes and their bases.

Data analysis

In my experience, qualitative proposal writers fall down most often on describing data analysis. Many data analysis processes—given qualitative research's nature, largely cognitive processes—can be challenging to describe even when researchers have already done them. Indeed, most writers struggle to get any more specific than “themes emerged” ([Dickie, 2003](#); [Humble and Radina, 2019](#)). To be fair, most quantitative researchers couldn't explain how their statistical tests work; they just report what the computer spits out. Anticipating precisely how you will analyze data before collecting any and when you expect your design to change can sound fantastical to many qualitative researchers, particularly novices who might never have done a large-scale analysis.

Solutions are available, though. First, look at how methodologists in your chosen qualitative approach suggest analyzing data. Perhaps, for example, you like [Moustakas' \(1994\)](#) approach to phenomenology; outline his specific analysis steps and suggest how your data will likely fit them. If you have some sample data—perhaps from a pilot study—consider demonstrating the process. Second, look at other authors' data analysis strategies in their published work. If you are enamored with [Strauss and Corbin's \(1990\)](#) axial coding process, for example, look at how researchers like [Morrow and Smith \(1995\)](#) implemented it, and imagine your data being analyzed in that way. Third, qualitative researchers have access to many inclusive, cross-approach means of analysis and coding ([Miles et al., 2014](#); [Saldaña, 2021](#)). Here, too, imagine the data you'll have and project how you might make connections between your data and the research questions. On this last point about matching the data with research questions, I have often found it helpful, both for planning and for communicating with readers, to create a table with the research questions in one column and the data sources used to answer each in another column.

Making a table of data–research question correspondences can also help ensure you've accomplished another key task: accounting for all the data sources you propose to collect. If you have, for example, said that you plan on doing ten observations of a classroom in addition to interviews with four teachers, don't only describe analyzing the interviews. Readers will rightly wonder why you're doing the observations if you overlook analysis of them.

Some might wonder about the role of data analysis software in a qualitative proposal. I don't suggest writing much about it at the proposal stage, honestly. After all, most qualitative software has similar features and uses, and rather than doing the analysis for you like statistics packages do, they mostly just manage data. If you do discuss software, name it and describe how you'll use (or perhaps overcome) its features to fit your unique data and analysis processes. As [Woolf and Silver \(2018\)](#) rightly note, one has to “translate” the emergent processes of qualitative research into the cut-and-dried operations of software. If you aren't at this point yet, though, don't feel the need to even mention software.

Again, crafting the details of an emergent process can prove challenging, but embrace the challenge. To do less, to essentially say “I'll figure it out once the data is collected,” risks readers thinking that you don't understand qualitative research or that you'll waste your participants' time and good will or waste your grantor's money, all without getting valid results.

Validity and reliability

While data analysis gets discussed poorly in qualitative proposals, validity and reliability—*trustworthiness* ([Lincoln and Guba, 1985](#)) or *quality criteria* ([Tracy, 2010](#)) or myriad other terms for good work, if you prefer—often don't get mentioned at all. Or, too often, just the names of procedures are tossed around as if such “boilerplate” lists are “magic charms that are intended to drive away evil” ([Maxwell, 2013](#): 123). Just do the technique and you've checked the box. Presto! Validity accomplished. That too—common view fundamentally misunderstands validity. As [Maxwell's \(2013\)](#) chapter title succinctly puts it, validity seeks to answer “How might you be wrong?” You can't rule out being wrong with lists of techniques; you rule it out by the evidence techniques generate (121).

Proposal readers want details of your validity and reliability plans, just as they do with other sections. Regarding reliability, how might you use multiple coders and establish interrater reliability (e.g., [Hruschka et al., 2004](#)), if applicable? How might you coordinate teams of researchers to ensure they keep their focus and interpretations mostly aligned? Regarding validity, which of the common techniques of trustworthiness—such as member checking, triangulation, peer debriefing, prolonged engagement, negative case analysis, and more ([Lincoln and Guba, 1985](#))—would effectively double check that your analysis really answers your research questions? Consider adding a third column to the table of research questions and data sources suggested above, this time including the procedures and evidence you might use to assert the validity of each question's answer.

Ethics

Discuss your study's ethics—the right and wrong of how you'll treat participants—and don't be dismissive of your responsibilities by only relying on Institutional Review Board (IRB) approval or suggesting that nothing could possibly go wrong. I've never been part of a study, even as a committee member, that didn't have an ethical dilemma arise—usually small but sometimes large. Discuss the ethical problems you can anticipate (or have confronted already while planning the project) and how you might manage dilemmas you can't anticipate.

Though I can't cover every ethical situation that could arise, some questions are frequent. Consider how your presence might impact the people and places you conduct research with. Are you too close to the participants already? Are some of them particularly vulnerable? Could your research uncover things they wouldn't want you to know or share? Might you endanger them, their reputations, or their livelihoods? How might you handle power imbalances? Could any confusions arise about your role in the setting? Do you have conflicts of interest that might make you unable to be honest? What cultural misunderstandings or tensions might you have to work through? Will anyone not directly participating be affected by your study? How might you write about participants in ways that are respectful and fair? All researchers face questions like these. No one can remove all ethical challenges, but they can be honest about them and put in place techniques to check biases, like reflective memoing, peer review, and debriefing sessions.

Delimitations: what are you not doing?

You can't do everything in any single study. Leave some questions for your second and third books! Make sure to tell readers what you are not doing in your study, your delimitations or boundaries. You might be researching the behavior of Egyptian students at school, for example, but for various reasons, despite its potential value, you won't be watching their behavior at home because access would be too difficult. Maybe you're studying new professors' teaching, but you're not going to include anyone from the college of education because they usually have teaching experience already. In your proposal, make it clear what you are doing, but make it just as clear what you won't do and why. Drawing boundaries around what is and isn't your project and providing a rationale make you look thoughtful and purposeful. Ignoring such discussion makes you look forgetful or capricious.

Pilot studies and previous data (if any)

Let readers know what you have done on the project already. If you did a pilot study, briefly overview what you did that relates to the current study. You don't have to spend pages and pages of your proposal explaining the pilot study's every detail. Just include a few paragraphs about what you learned and tried in the pilot and what you plan to change in the proposed research as a result.

Also mention if you have already begun working on the current project. Have you contacted some of the gatekeepers? Have you enrolled participants? Do you have IRB approval? Are you planning on reusing some previously collected data? If your proposal is for a dissertation, disclosing previous work to your committee is critical because they may want to view any material already collected to ensure that it meets standards.

Researcher disclosures

In qualitative studies, researchers are the "instruments of the research" (e.g., [Lincoln and Guba, 1985](#): 192–195); their cognition and emotions filter what they see, hear, and interpret. Thus readers need to know about you and how you will perform as that instrument. Who are you, and what experiences do you bring that will shape your views? What biases might you bring? What reactivity might you provoke in the field with participants? Why are you, particularly, equipped to do this?

In this section of your proposal, explore your subjectivities, biases, training, and expertise. Disclose demographics that could reasonably be assumed to relate directly to your study. If your project focuses on race, for example, share yours and how that might influence you. The same goes for identities like sexuality, gender, social class, disabilities, and so on. If you hold political, religious, or moral stances that might influence how you think about the people and topic, confess to it. Talk about how you have prepared for your research. If you have taken doctoral courses on the topic, trained in advanced methods, published in the area already, or worked on related grants, let readers know. They will have more confidence that you can successfully complete the study.

Though often focused on checking biases, a disclosures section is not meant to be all negative. Consider positives that might result from your subjectivity, too. Might being a man, for example, give you privileged access to the group of boys you want to study? Might you understand your study of first year teachers better from having spent the last few years supervising student teachers? Don't oversell your identities as making things surefire or easy, but be thoughtful about the advantages you might have.

Use your section on researcher disclosures for developing methods, as well. How will you prevent any biases from skewing your interpretations? What insights can you only get because of who you are? [Maxwell \(2013\)](#) suggests memoing about your researcher identity. Debriefing with peers or supervisors can keep you honest and sensitive, too. It's OK to have identities and stances; you are

human, after all. It's not OK to hide these or pretend that they don't matter. Share your plan for keeping them in check or capitalizing on them.

Logistics

You can also show your skill and thoughtfulness—in other words, that you can accomplish the study—by providing a realistic timeline, budget, data management plan, possible “deliverables” for a grant, and, for a dissertation or book proposal, a tentative outline of chapters.

What steps will you take and how long will it take you to do them? Breaking down the parts of the study and crafting a timeline tests your ability to be realistic. Grant agencies are more likely to hold you to this timeline than are dissertation committees, but whichever you're writing, give yourself ample time to collect enough data, think through it all, and then write the study up in an engaging and coherent way. It amazes me how many students think they'll do ambitious fieldwork in a semester and then write and defend a dissertation in a semester. Those who actually accomplish such feats are as rare as mega-million lottery winners. High-quality qualitative research takes time, frequently faces obstacles (the “emergent” part!), and doesn't enjoy the straightforward writing templates that other methods might. Tamp down your hunger to finish your degree or get that tenure-assuring project done, and plan in extra time for every stage. One great approach involves creating a table listing in detail your study's phases or steps, and give a tentative completion date beside each. And by detail I don't mean a step as broad as “data collection”; I mean breaking large phases like that into individual tasks, like “Complete first set of observations in all schools,” or “Interview first five principals.” Such detail quickly exposes places where your timeline might be too aggressive.

If you are asking for money, tell readers what it's for in as much detail as possible, and focus your ask on how that money forwards the project (rather than your career or your convenience). Don't be overly frugal (“I'll get all my supplies from the dollar store!”), but don't be extravagant, either. Outline your professional needs and give a solid rationale. Seek help early and often from people who are good with budgets if you're not.

Endings

Goodall (2008), writing about narrative research projects, calls endings “the pay off.” I think that description works well for proposals in any research approach. Give readers a sense of closure when you write your conclusion. No new information! Only give, reminders of what they have learned already in the proposal. Remind them, after that long trip through your detailed research design, that your project has important potential implications for policy, practice, research, and/or theory. Reinforce that your literature review suggests the need for your study. Emphasize that your chosen theoretical framework and the methodology and methods you proposed make it doable. Such summarizing helps readers see the whole sweep of the project in one concise package. The repetition of the information also helps them remember it. My motto for conclusions is “Leave 'em informed! Repeat, repeat.”

Too often I feel writers rush their conclusions, so they don't get all the “pay off” that conclusions make possible. Perhaps they are tired of writing by that point. Perhaps the deadline is at hand. Perhaps they just don't know what tone to set, so they simply stop at the end of the research design. Hopefully you won't fall prey to such temptations or bad habits, though. A strong ending can leave your reader feeling satisfied and confident, and that mental state bodes well for gaining approval for your proposal. To write your own ending, look for examples you like, get others to read yours, and revise frequently.

Revise, revise, share, revise, share, revise

My final bit of new advice concerns making your prose great—or at least better. Few writers can crank out perfect sentences and paragraphs, with style and concision and coherence, on the first try. Many attempt the one-shot composition, but few succeed. Many also—let's admit it—don't bother trying to make their writing clear or stylish, particularly in the sciences and social sciences. As Becker (2008) notes, writers in these fields get away with bad writing all the time. They confuse complicated, jargon-filled prose for smart ideas. Maybe they just know so much about the topic that they forget how to write well (Warren et al., 2021). In a proposal, though, writing well—clearly, concisely, grammatically, and with detail—brings pronounced advantages. Doing so enhances your image as someone capable of doing the work. I mean, if you can't be bothered to at least correct the misspellings your word processor points out for you, why would I trust you to research carefully?

The path to good writing goes through revision, and lots of it. Read your proposal over and over, making changes to make your ideas and words clearer. Take out unnecessary words that slow readers down. Explain potentially unfamiliar concepts. Add details and examples. Revision relies on many processes, which I detail at length in *How to Write Qualitative Research* (Weaver-Hightower, 2019), but make sure to leave sufficient time to do the many “waves” of revision I suggest. Revision should take longer than the first draft!

Also leave time for the “share” parts of this section's title. Receiving feedback as you develop the proposal and before you submit it is to your advantage. Get multiple rounds of feedback from your advisor, your peers, your colleagues—heck, from your grandma—and make sure to revise based on that feedback before the next round of readers. The more you give real readers an opportunity to catch mistakes, ask questions, and make suggestions, the more likely your proposal will be clear and compelling to the final readers who decide your fate.

Conclusion

I hope that this chapter, as I said of proposals writ large at the beginning, excites you about the possibilities for your own proposal. Yes, proposals are complex documents that take time to construct and refine, but they are not inscrutable. Identify your readers and what they need to know. Find out from those readers what format they expect, then follow it explicitly. Write an interesting descriptive title and make the opening hook readers by engage the topic right away. Review the relevant literature, but keep it a partial review given that new topics might arise as your study evolves. Then lay out your research design in as much detail as possible, including your larger methodological approach, your research questions, sites and participants, data collection strategies, data analysis approach, validity and reliability plans, ethics considerations, and reasonable logistics about timelines, budgets, and so on. Put a satisfying ending on it. Then—*voilà!*—your proposal is complete. All you have to do then is multiple rounds of revisions and sharing to make sure that readers recognize the brilliance of your ideas and the promise of your study. Naturally all of that is easier said than done.

The doing now lies before you. Take comfort from the knowledge that thousands upon thousands of scholars have written successful qualitative proposals; they have found a way through and you will, too. Hopefully, with my advice throughout this chapter, the way will be at least a little clearer. Good luck on your proposal!

References

- Barton, D., Hamilton, M., Ivanic, R. (Eds.), 2000. *Situated Literacies: Reading and Writing in Context*. Routledge, London.
- Becker, H.S., 2008. *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*, second ed. University of Chicago Press, Chicago.
- Caelli, K., Ray, L., Mill, J., 2003. "Clear as mud": toward greater clarity in generic qualitative research. *Int. J. Qual. Methods* 2 (2), 1–13.
- Charmaz, K., 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage, London.
- Creswell, J.W., 2013. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. Sage, Los Angeles.
- Denzin, N.K. (Ed.), 1970. *Sociological Methods: A Sourcebook*. Aldine, Chicago.
- Dickie, V.A., 2003. Data analysis in qualitative research: a plea for sharing the magic and the effort. *Am. J. Occup. Ther.* 57, 49–56.
- Gage, N.L., 1989. The paradigm wars and their aftermath: a "historical" sketch of research on teaching since 1989. *Educ. Res.* 18 (7), 4–10.
- Garnham, B., 2008. Data generation. In: Given, L.M. (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods*. Sage, Thousand Oaks, CA, pp. 192–193.
- Glaser, B.G., Strauss, A., 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine, Chicago.
- Goodall Jr., H.L., 2008. *Writing Qualitative Inquiry: Self, Stories, and Academic Life*. Left Coast Press, Walnut Creek, CA.
- Haywood, P., Wragg, E.C., 1982. *Evaluating the Literature*. Redguide 2. Nottingham, England: University of Nottingham.
- Holstein, J.A., Gubrium, J.F. (Eds.), 2003. *Inside Interviewing: New Lenses, New Concerns*. Sage, Thousand Oaks, CA.
- Hruschka, D.J., Schwartz, D., St John, D.C., Picone-Decaro, E., Jenkins, R.A., Carey, J.W., 2004. Reliability in coding open-ended data: lessons learned from HIV behavioral research. *Field Methods* 16 (3), 307–331.
- Humble, Á., Radina, M.E. (Eds.), 2019. *How Qualitative Data Analysis Happens: Moving Beyond "Themes Emerged"*. Routledge, New York.
- Jackson, A.Y., Mazzei, L.A., 2012. Thinking With Theory in Qualitative Research: Viewing Data Across Multiple Perspectives. Routledge, Abingdon, England.
- Kilbourn, B., 2006. The qualitative doctoral dissertation proposal. *Teach. Coll. Rec.* 108 (4), 529–576.
- Ladson-Billings, G., 2021. *Culturally Relevant Pedagogy: Asking a Different Question*. Teachers College Press, New York.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*. Sage, Beverly Hills, CA.
- Marshall, C., Rossman, G.B., Blanco, G.L., 2022. *Designing Qualitative Research*, seventh ed. Sage, Thousand Oaks, CA.
- Maxwell, J.A., 2010. Using numbers in qualitative research. *Qual. Inq.* 16 (6), 475–482.
- Maxwell, J.A., 2013. *Qualitative Research Design: An Interactive Approach*, third ed. Sage, Los Angeles.
- Miles, M.B., Huberman, A.M., Saldaña, J., 2014. *Qualitative Data Analysis: A Methods Sourcebook*, third ed. Sage, Los Angeles.
- Morrow, S.L., Smith, M.L., 1995. Constructions of survival and coping by women who have survived childhood sexual abuse. *J. Counsel. Psychol.* 42 (1), 24–33.
- Moustakas, C., 1994. *Phenomenological Research Methods*. Sage, Thousand Oaks, CA.
- Peshkin, A., 1985. From title to title: the evolution of perspective in naturalistic inquiry. *Anthropol. Educ. Q.* 16 (3), 214–224.
- Roulston, K., 2010. *Reflective Interviewing: A Guide to Theory and Practice*. Sage, Los Angeles.
- Saldaña, J., 2021. *The Coding Manual for Qualitative Researchers*, fourth ed. Sage, Los Angeles.
- Sandelowski, M., Barroso, J., 2003. Writing the proposal for a qualitative research methodology project. *Qual. Health Res.* 13 (6), 781–820.
- Sandelowski, M., Voils, C.I., Knaff, G., 2009. On quantizing. *J. Mix. Methods Res.* 3 (3), 208–222.
- Savin-Baden, M., Major, C.H., 2013. *Qualitative Research: The Essential Guide to Theory and Practice*. Routledge, London.
- Seidman, I., 2006. *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*, third ed. Teachers College Press, New York.
- Strauss, A., Corbin, J., 1990. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage, Newbury Park, CA.
- Tracy, S.J., 2010. Qualitative quality: eight "big-tent" criteria for excellent qualitative research. *Qual. Inq.* 16 (10), 837–851.
- Vagle, M.D., 2018. *Crafting Phenomenological Research*, second ed. Routledge, New York.
- Warren, N.L., Farmer, M., Gu, T., Warren, C., 2021. Marketing ideas: how to write research articles that readers understand and cite. *J. Market.* 85 (5), 42–57.
- Weaver-Hightower, M.B., 2019. *How to Write Qualitative Research*. Routledge, New York.
- Wolcott, H.F., 2009. *Writing Up Qualitative Research*, third ed. Sage, Thousand Oaks, CA.
- Woolf, N.H., Silver, C., 2018. *Qualitative Analysis Using MAXQDA: The Five-Level QDA Method*. Routledge, New York.

Writing up qualitative research

James Drisko, Smith College, Northampton, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Getting started	721
Forms, formats, purposes and standards	722
Recently developed methods—qualitative inquiry	722
Text based reports	723
Approaches and methods/techniques	724
Quality checklists and guidelines	724
Finding the right journal or publisher	725
Submitting and being reviewed	725
Reviewing	726
Conclusions	727
References	727

Writing up qualitative research and inquiry can be difficult. We write to communicate but we also learn a lot during the writing process. Qualitative research is a large, diverse, and growing family of research approaches (Denzin and Lincoln, 2017). These many approaches draw on different western and Indigenous epistemologies, may or may not seek to be scientific, are applied for different purposes, serve the knowledge needs of different audiences, and must maintain ethical standards. These qualitative approaches may apply any of over 100 named methods or generate new ones to fit specific knowledge needs, and must persuade readers that the study's analysis, conclusions, and any recommendations are consistent with and limited to the study's delivered methods. Often the data generated by a study may require much more space to communicate in depth than is afforded in the 20-page manuscript limit set by many journals. Some forms of qualitative inquiry take the form of "hung" photographic exhibits, of creatively altered books, of spoken plays, and of autobiographical stories. Yet all are important parts of the large qualitative research community. Writing up qualitative research draws on and integrates many sources.

Of course, we do not write for ourselves. We write to inform, or to inspire, to bear witness, or to motivate others. These other people are our unseen and unknown audiences. Students, even at Ph.D. level, often have difficulty imagining that others will actually read their work. As researchers, we seek to learn, understand, and innovate. As qualitative research writers, we re-shape the processes of our own learning and find the best ways to inform others about what we have done and learned. Rarely is it best to simply restate our personal learning paths to get ideas across to others. Wolcott (2009) describes what we produce as a monologue. Readers do not tell authors why they chose to read this work, where to begin, or what parts are puzzling. As writers, we must imagine our readers and speak to them. We must inform our readers. Authors monopolize the conversation. Authors have to imagine an audience and then convince them that we know what we are talking about. It's not easy.

Across disciplines and professions, authors of qualitative manuscripts frequently report that their works have been unfairly turned down by journal reviewers (Pratt, 2009). These authors, among which I included myself, sometime believe that their work has been "slammed" by reviewers who are not knowledgeable about qualitative approaches and methods, or who have applied inappropriate standards in their reviews. It is also true that we have failed to reach, and to adequately educate, this crucial audience. I believe such hostile reviews are the result of the minimal way in which qualitative approaches and methods are taught and mentored at graduate level in most disciplines and professions. This is further compounded by a lack of consensus standards for reviews fitting the author's chosen approach. In my profession—social work—I found that qualitative research is most often taught in the United States at master's level in only one class session, usually squeezed between sampling and statistics (Drisko, 2008). No assignments asking for active involvement in qualitative research are found. At doctoral level, the typical Ph.D. program includes just one course in qualitative research. These are either an overview course or a course in a single named approach (Drisko et al., 2015). Minimally taught and mentored, many able students are simply not educated adequately to become expert qualitative researchers. For example, Emmons and Moravcsik (2020) view inadequate education in qualitative research as a current disciplinary crisis within political science. Similarly, there are few reviewers with extensive and varied training, wide-ranging qualitative research experience, and methodological expertise in qualitative research across professions. Qualitative audiences may have to be taught core research approaches as we report our research results (Stenius et al., 2017). We have to up our game.

Getting started

Getting started in writing up qualitative research can be a challenge for newcomers and veteran writers alike. The good news is that even false starts create fodder for later edits—once they are written! Common concerns are the pressure to write and the shame or guilt that emerges when we are not writing. The most useful response is to sit and write: a title, a statement of purpose, an

outline, ideas about what you want to study, or ideas about what you do *not* want to study. Clarifying the topic and delimiting what the study will include can be terrific starting points. Get going; put pen to paper or click those keys. Of course, it won't be perfect at first. But it's a start!

Having a steady base helps one be an effective academic writer (Sword, 2017). Sword studied qualitatively 100 academics and found that behavioral, artisanal, social, and emotional factors and habits each strongly impact writing output. By rating one's own comfort on each factor, one can identify how to build greater support and comfort in writing. Habits on these four dimensions evolve over time and change in prominence. Both developing writing rituals and letting them go can enhance one's writing. As qualitative writers we differ both one from another and we each evolve over time and experience. Yet we can increase our comfort and output. Experienced qualitative authors review and reflect upon how, when, and where they write.

In leading professional writing seminars, I often ask people *when* during the day they write most successfully. Many of us have a time of day in which we are most academically productive. Try to protect these times and write. I also ask *where* one writes. After laughter about famous writers' rooms and desks, it turns out we favor a wide range of locations and settings: indoors, outdoors, in tiny carrels, or in coffee shops, with noise, or in quiet. Some of us use outlines, others do not. The most interesting question was to what do you like to read? Response ranged from novels, to religious matter, to poetry, and on to graphic novels. Academic articles were never mentioned. Knowing what impacts your productivity can help you learn about why, and how, good writing 'works.' Identifying and respecting the conditions that help you write efficiently can save time and create comfort. Each writer is unique. Self-reflection on one's qualitative writing is important.

For those who use outlines, a detailed outline is a great way to start. Kellogg (1988) found that both written and "mental" outlining led to higher quality final documents and reduced attention overload. Outlines allow researchers to delimit what will be in the piece, and what will not. This information is vital for finding participants, and for informing colleagues, committees, and institutional review boards. Using a named research approach, having expertise and experience in data collection and data analysis, and planning out how to use 20 pages of double spaced text focus and guide our qualitative writing. The outline helps us to work in parts or steps, section by section. This allows us to focus in depth on one part of the work. Inevitably, how the sections integrate may need revision or expansion as the work progresses. Some material may prove unimportant and warrants removal. Working section by section can also help identify progress.

For those who do not use outlines, free writing may be a useful technique (Becker, 1986). Just write, with no changes or edits, for 10 minutes. Just let go. This is a great technique for exploring your own reactions in qualitative research and to set down for review what you notice, question, and doubt. If writing reflects thinking, then writing an introduction or a study overview gets us started in both realms. Set a timeline. Review the draft a day later and look for ways to improve its clarity and its organization. What works? What's missing? Are key terms defined? Is there a flow, a clear organization, developing? Does the writing have "grab" to draw in readers (Gilgun, 2005)? One can build incrementally on such a small start with additional sessions of free writing. These are directed to specific sections of the developing manuscript. As always, edits and reorganization will likely be needed as the manuscript expands. As Wolcott (2009) states "Until you have a rough draft of what you have to report, there is no chance of improving it" (p. 162). Of course, you also have no way to get detailed feedback from others unless they can sit with, read, consider, and respond to a written draft. Starting qualitative writing may be challenging but it is a vital first step. And—it builds confidence.

As the draft is edited and refined, qualitative researchers must look carefully for the internal consistency of the overall, enacted, research approach (e.g., grounded theory, phenomenology). How the approach, epistemology, data collection and data analysis methods fit together become key areas for review and consideration. Quality of content and clarity of delivery are both key.

Forms, formats, purposes and standards

Qualitative research and inquiry are presented and published in a wide range of forms and formats. Several recent developments expand on the traditional text-based qualitative report. Standards for such evolving works are not yet fixed: there are neither consensus guidelines for qualitative researchers across approaches, nor across disciplines and professions. Those guidelines that have been developed have been heavily questioned and critiqued. They often do not fit well with some qualitative research approaches (Kross and Giust, 2019). Qualitative inquiry using approaches like autoethnography and performance methods are not compatible with many Realist/Post-positivist oriented guidelines and publications standards (Drisko, 2013). There are, however, some general criteria for what makes a strong text-based qualitative study.

Recently developed methods—qualitative inquiry

While the focus of this chapter is on writing up qualitative research, several forms of qualitative research and inquiry are presented in forms other than text. Further, not all inquiry seeks to be scientific, and some qualitative approaches purposefully emphasize the aesthetic aspects of communication. Several recently developed methods (e.g., participatory action research, performance inquiry) seek to empower others, to democratize knowledge production, to broaden the ways in which qualitative research is presented, and to expand how it is evaluated. As an example, photovoice is a method in which "participants" are given cameras by researchers to document a (usually) researcher defined aspect of life about which participants are knowledgeable (Wang and Burris, 1997). One well known photovoice study documented the lives of the homeless through framed pictures with short captions. The study was "hung" in public settings such as libraries with the purpose of helping housed folks to better understand the experiences of the

homeless. The presentation's location was chosen to provide access to likely "consumers" of the study in an easily accessible public setting. While building public understanding and reducing stigma may have been a key research goal, the aesthetic dimensions of the presentations also impact how it is viewed and understood.

Similarly, "performance pieces" are short spoken or acted plays that *show* the audience the message, linking words with inflections, feelings, tones, and structures that engage the audience and shape understanding of the material (Gergen and Gergen, 2011; Haseman, 2006). Human interactions may also be viewed as performance (Denzin, 2001). There may be a complete or near complete text from which the performance is generated, but the presentations are embodied, enacted, live, and are not mainly intended as text. Here too aesthetic aspects of the performance both convey its message and link to everyday ways of taking in new information. Each of these qualitative research approaches may also be designed to shift consumer's attention: to shift attitudes or perspectives, to make viewers uncomfortable, to make the familiar seem unfamiliar. For example, the *Performance Research* journal offers a "dynamic space of performance that produces inspiring conversations, unlikely connections, and curious confluences" (2020, para 1). Guinier and Smith (2001) also argue that artistic performance may be used to influence others and to highlight power dynamics. Here research and the arts intersect and show mutual influence. Criteria for the "success" of such qualitative work may not be consistent with some journal reporting standards.

Autoethnography is another innovative form and format in which a researcher composed story links intimate personal experiences and observations to the social structures in which these experiences occur (Goldschmidt, 1977). The structure and design of the story, its aesthetic dimensions, are vital to its impact and success at informing and moving the consumer (Bochner and Ellis, 2016). As personal stories, authenticity is required, voice can be direct and strong, the imperfections of our understandings of others is side-stepped, and the legitimacy of telling one's own story is affirmed. Evaluative standards for these recent innovations in research approach have yet to be established, and may emphasize the artistic and aesthetic over the scientific.

Text based reports

Traditional forms of text based qualitative research have no required format other than the expectations set by researchers, journals, and committees. Different approaches to qualitative research will, however, also require specific sets of information to orient and guide the reader. Some of these methods and related terminology are specific to named research approaches. Write ups must inform the reader/consumer about what was done and how it was completed. Authors must guide the reader through the research and show what has been done and found.

A clear statement of the study question is a typical place to start (Drisko, 1997, 2013). All good research starts with a question of merit and worth. What will be studied, what it adds to knowledge, and its practical implications should all be mapped out briefly. It may be the case that the research centers on exploration of unknown territory, yet still a focusing question or topic is vital. The question orients researchers and readers alike.

Next, state the epistemology that orients the study (Drisko, 2013). Many qualitative publications fail to state their epistemological premises, forcing the reader to guess what is intended. Epistemologies have implications that may lead to omissions or inconsistencies in later stages of the report. For example, a constructivist study drawing on a convenience sample of 12 may close with recommendations and generalizations that would be more appropriate to a realist approach and a larger and more varied probability sample. The ways of knowing that orient a study should be understood by the researcher, clearly stated, and linked to appropriate and limited knowledge claims.

The researcher should also be clear about the intended audience of the report. Forms of writing often differ when created for the public or a specific group rather than other professionals from the same discipline. Again, focusing on an intended audience helps focus the researcher and the writing up of the report.

Ethical issues arise in most qualitative research with human participants. All research reports involving newly collected data from people should include a statement about providing informed consent and other ethical issues. Such a statement may be a short sentence. For example, "after review and approval of the study questions and approach by the XXX institutional review board ..." shows appropriate professional conduct and that the author has taken steps to consider and to protect human research participants. There are study questions that may be challenging for institutional review boards (IRBs) (i.e., the tea room trade study; some studies of drug use). It is also true that IRBs may not fully understand some forms of qualitative research (Van Den Hoonaard, 2002). A clear description of the study purposes and methods—whatever they are—helps inform and often to educate IRB members and readers alike. The participants in our qualitative research deserve our care and protection, which should be stated in all publications (Drisko, 2013; Lahman, 2018).

Doing qualitative research may also generate personal ethical challenges that are not often addressed by biomedical ethics models (van den Hoonaard and van den Hoonaard, 2013). These ethical challenges may also be unfamiliar to many ethics review boards (van den Hoonaard, 2002). Many such differences exist for qualitative researchers: The process of entering existing human groups and settings, making choices about payment or other incentives, staying in and participating within groups over time (or not), and in leave taking, can each pose interpersonal and potentially ethical challenges (Wolcott, 2002). Including diverse participants may raise the issue of "othering" (Lahman, 2018) while the depth of our information needs may also raise the issue of intrusion. How to ask difficult questions of vulnerable populations requires self-reflection and thoughtfulness. Asking difficult questions requires ongoing attention to the interests and experience of the participant. Such ethical and practical concerns should be explored in the published report, and may generate new questions for additional research.

Approaches and methods/techniques

Unfortunately, too many qualitative reports offer minimal detail on their approaches and methods. The term “approach” is used here to indicate a named qualitative research model, such as phenomenology, ethnography, or grounded theory. Qualitative research approaches may specify sampling methods or data collection methods, but all specify—to some level of detail—the research methods to be employed and how they jointly build knowledge. Sampling, data collection, and data analysis should be consistent with the selected epistemological premises of the study. That is, the overall study approach should be internally consistent with the study epistemology. For example, open, discriminate, and selective sampling should be reported in a realist premised grounded theory study drawing on [Glaser and Strauss’ \(1967\)](#) approach.

Methods section in qualitative reports need not be long but should address several topics. Is the sampling plan and its rationale clearly stated? Will the plan include possibly contradictory cases or evidence? Are the data collection methods clearly stated and consistent with any named research approach and methods? Is the method of data analysis consistent with any named research approach and methods? Are codes, themes, and interpretations clearly linked to reader available data? (Beware of reports that only offer a few sound bites summarizing complex coding processes!) Are the data and analysis credible? Is it clear how the researchers applied codes to varied material? Confirmable or disconfirmable? Meaningful in context? Saturated and complete? How is contradictory data understood? Are any claim of transferability or generalization consistent with the study epistemology, sample, reported data, and analysis ([Drisko, 2013](#))? The goal of the writer is to explain to the reader how the study was done. This is essential to establishing the credibility and verisimilitude of the study. A clear statement of methods helps to persuade the reader that the results of the work are meaningful.

Good qualitative studies also require a statement of the author’s personal location and pre-existing views about the topic. This gives the consumer an insight into the author’s potential biases and interests, supporting critical thinking by the reader. Such statements can also clarify the researcher’s intentions and assumptions. Reflection on the topic, the study, and one’s changing views each may help the reader understand the report more fully. [Stenius et al. \(2017\)](#) point out that each of these quality criteria also serve as reasons that qualitative manuscripts may not accepted for publication by journal referees or jurors. [Drisko \(2005\)](#) and [Pratt \(2009\)](#) note that even well described methods may be misunderstood or incorrectly portrayed by journal reviewers. Care in ensuring rigorous manuscripts is vital to their acceptance and publication. Clear explanations of methods may educate both reviewers and future readers.

[Humphrey and Simpson \(2012\)](#) argue that writing up collected qualitative data builds confidence and demonstrates beginning mastery. They call completion of a qualitative data analysis a “threshold” step in becoming an academic writer. Humphrey and Simpson call for attention to writing up qualitative data as promoting independent and autonomous research writing skills. They note that feeling “stuck” is a common, almost routine, step in becoming a qualitative researcher and writer. Instructors, mentors, and peers all can contribute to becoming unstuck and helping the researcher in “taking control of the thesis as a textual synthesis of data, theory” (p. 744).

Writing up qualitative research may be difficult, but it is vital to communicating ideas, methods, and results to others. Researchers often identify new results and implications as they work to create a cohesive and complete report: qualitative research is an iterative or cyclical process of thinking, planning, collecting data, analyzing data, and writing. Different parts of a project may be configured or re-configured differently to better impact selected audiences, and to address different information needs and learning styles. Providing a “road map” for readers that show how the study was conceptualized and completed is vital.

Quality checklists and guidelines

Academics, professional organizations, and commercial publishers have developed several formats and checklists to promote quality in qualitative research. These generally reflect an effort to establish scientific principles for the conduct of qualitative research. They also tend to be heavily and solely premised on post-positivist or realist epistemologies. Further, most tend to minimize the role of researcher self-reflection and reflexivity, seeking an objective account. [Calderón Gómez](#) argues that

The “scientific format” required by the majority of journals determines the representation of the research process by fragmenting it and reducing it to previously designed spaces. Similarly, the author is often asked to use an impersonal linguistic style that does not represent his or her involvement in the research process.

[Gómez and Carlos \(2009, para 34\).](#)

Given the growing diversity of purposes, audiences, approaches, and methods within qualitative research and inquiry, it is unclear if these formats fully and usefully reflect writing in today’s qualitative research landscape ([Barbour, 2001](#)).

On the other hand, checklists can provide some guidance for researchers, students, and educators. They show, in detail, expectations for realist qualitative studies ([Turk and Kalarchian, 2014](#)). They may not, however, include all components of a qualitative study deemed important to the participants, researchers, or future readers. The formats also point to content for educational curricula in qualitative research, though hardly comprehensively. They focus on mainstream accounts.

Most standards for studies and manuscripts offer a set of general guidelines for authors. In education, [American Educational Research Association \(2006\)](#) offered standards for all empirical publications, including qualitative studies. The 42 main standards are realist in orientation; epistemology is not mentioned in the document. The AERA standards do include attention to research ethics and a strong focus on establishing the evidence that supports each conclusion. Similarly, in medicine, [O'Brien et al. \(2014\)](#) created a set of 21 broad standards for reporting qualitative research, synthesizing the recommendations of others across professions. In 2018 the American Psychological Association published its recommended "JARS" or "journal article reporting standards" for qualitative research ([Levitt et al., 2018](#)). The detailed, four page, list seeks to address a range of study types and writing styles. It does not offer much guidance for studies not premised on science, though it mentions narrative reporting styles. Epistemology is not mentioned as an important orienting choice for researchers, nor is self-reflection mentioned, though author reflexivity is included in terms of a structured way to enhance methodological integrity. Another checklist is [Tong et al. \(2007\)](#) Consolidated Criteria for Reporting Qualitative Research (COREQ). This checklist draws attention to three areas (research team and reflexivity, study design and data analysis, and reporting) and covers many details in its 33 items. Some journals may require submission of the COREQ or other checklists as part of the submission process. However, JARS and COREQ may not fit well with more contemporary qualitative inquiry efforts. Such checklists are useful guides for some qualitative research projects but others may find them limiting or confining.

All checklists for qualitative studies and reports focus attention to topics that are likely concerns of reviewers, and in this way can be useful to authors. On the other hand, checklists may fail to address all forms of contemporary qualitative research and inquiry ([Kross and Giust, 2019](#)). For example, [Cheek \(2008\)](#) notes that standardization of qualitative research may force authors to conform to views that may not match their purposes, research approaches, or collected data. Checklists may be a valuable tool for beginners, but may artificially and unnecessarily limit innovation in qualitative research. Standard checklists also fail to address the growing interest in qualitative inquiry, in which scientific standards may not be the researcher's orientation.

Finding the right journal or publisher

It is pivotal to effective qualitative research writing to determine just who is your intended audience. By this I mean to whom do you seek to inform, persuade, or give testimony? A key component of traditional academic publishing is determining the best few journals to which to submit. This is often based on the journal's impact factor. Journals serve as gatekeepers to these readers as well as providing a service of organizing and presenting the research. They also match readers and research writers.

To identify a target journal for a study, it is often helpful to reflect on the journals that were most useful in framing the current research literature and methods. These are often good targets for your publication as they publish similar topics using similar research approaches and methods. Journal mission statements may also clarify the kinds of topics and methods being sought for publication. Journal web sites provide details about the length, format, and other requirements for submissions. Reviewing these statements can further clarify if a journal fits a specific research manuscript.

There are many online resources to aid in selecting a target journal for one's work. Simple searches can reveal lists of professional journals along with their impact factors and other useful information. For example, [ERIC \(2020\)](#) offers a huge array of education publications ranging from journals to conference proceedings. The ERIC list may be searched selectively or comprehensively, and can be set to search only for peer reviewed publications, or only for those available free in full text format. While comprehensive, it is easy to be overwhelmed by so many choices. [Scimago Lab \(2020\)](#) lists international journals in education by several journal and individual impact factors. Sometimes the best target journal will be outside one's home country. Online tools such as [Elsevier's \(2020\) Journal Finder](#) may also narrow a list of likely target journals. All such lists and tools may guide us, but the final choice of target journal requires additional critical thinking.

For longer works, a monograph length publication may better capture the detail and complexity of the research. Monographs are typically reports that are well over 20 double-spaced pages, and may be much longer. Some publishers will publish monographs on specific topics; this is an option that qualitative researchers should explore for extended works. A monograph, according to Case (2009) is "a large, specialized work of scholarship that treats a narrow topic in great detail," and that focuses on "establishing facts or narrative in a set of fields in which facts and narratives are often hard to establish" (p. 4). Monographs target very specific and specialized audiences. Alternatively, some longer works may be divided into two-part article publications. These are often published in the same issue or may be published in two successive issues. For example, [Shields and King \(2001a,b\)](#) reported a large qualitative study addressing hospital services in four countries in two articles published in successive journal issues. Similarly, [Petty et al. \(2012a,b\)](#) used a two-part article to first explore the philosophical underpinnings of qualitative research. They next linked this discussion to an empirical analysis of specific named qualitative research approaches in the second article.

Submitting and being reviewed

Once a manuscript is submitted for publication, it is typically reviewed by two or three anonymous referees or jurors. These reviewers are typically successful academics with depth knowledge in the specialty areas of the target journal. Sometimes reviewers are selected from researchers who have written extensively on a topic or who have particular expertise in the topic area ([Allen, 2016](#)). The identities of authors and reviewers are kept anonymous from each other during the review process to limit bias. Reviewers are

rarely publicly identified in ways that link them to specific reviews. They are, however, sometimes acknowledged by the authors and named after acceptance when they have made particularly valuable comments or additions to the manuscript. Bear in mind that reviewers are asked to serve as critical readers who appraise both the strengths and the limitations of manuscripts. Of course, reviewers ideally review manuscripts using approaches and methods which they have some familiarity. Key issues include clarity of the research question or topic, a reasonable literature review, and internal consistency among epistemology, research approach and methods employed. Staying close to the data or evidence is always wise.

Reviewers are not solely serving as educators for authors; they must instead ensure that published works are sound, clear, and rigorous. Reviewers are usually asked to return completed reviews within three weeks or less of assignment. This is done to keep overall publication turnaround time to a minimum. Reviewing is an academic “service”, meaning it is an unpaid role necessary for the professions continuation but is given little reward toward tenure and promotion (Staller, 2013). It is a process vital to academic publishing.

Once all reviews are submitted, the editor-in-chief makes a publication decision on the manuscript. Decision options generally include (1) accept as is, (2) minor revisions needed, (3) major revisions needed, (4) reject, or (5) the manuscript should be directed to another journal which better fits its content and focus. Full accepts are rare, while revision requests are common. The journal editor will send a decision letter or email to the authors. The letter states the decision and if revisions are recommended, includes statements and suggestions from the reviewers. Some journals set a deadline for revisions to be completed and the manuscript resubmitted. Quick revision is best.

In responding to recommendations for revisions, the authors should list every critical comment from each reviewer and then show specifically how the concern has been addressed. A literal numbered list of comments by each reviewer and author replies is useful. Such a list documents the authors’ responses and that all revisions were made. Be sure to address all stated concerns. However, if the reviewer’s concern seems inaccurate, or would lead to other problems, it is fine to explain why the content was left as is. Of course, editors and reviewers may not agree with the author’s logic.

Once the authors submit the revised manuscript, the journal editor sends the referees the revised manuscript. Another timed cycle of reviews by the same referees begins. At the end of the second review, the authors are presented again with a decision on the revised manuscript.

From acceptance to publication is a process with no fixed time frame. Publications lag may be short or quite long. Some journals have significant print backlogs, and in turn offer “online first” publication in which the article is made available online electrically before it is assigned to a specific print volume, issue, and final pagination.

Reviewing

Reviewing qualitative research and inquiry manuscripts is another important role in the professions for which education and preparation is limited. The literature on conducting manuscript reviews is also quite small. It seems that grading student papers is taken as an adequate introduction to reviewing professional publications; though the content and methods of academic manuscripts are often novel and innovative, perhaps even unfamiliar to their reviewers. The challenge of a manuscript review is quite unlike grading an academic assignment one has personally created on familiar content. A benefit of doing journal and book proposal reviews is that one is exposed to the latest issues, and sometimes to cutting-edge ideas and methods. At times reviewers recuse themselves from reviews when they can determine the identity of the author; or when the reviewer may be a colleague or competitor. To limit bias, reviews are intentionally “blind” with neither the authors’ nor the reviewers’ names revealed during the process.

In doing a review, I find that a quick read through of the full manuscript is best followed later by a slower and closer look in which I make specific comments linked to text (Drisko, 1999). Staller (2013) cites Lee’s (1995) list of items to consider when reviewing a manuscript. Lee begins by summarizing the manuscript as read. I also summarize “what I’ve read” at the start of my own reviews. This process helps identify omissions of content (i.e., no named method, no mention of an IRB review) and makes clear to the authors that one has read the manuscript. Following this overall summary, reviewers should provide detailed comments that are each specifically linked to content by page and line numbers (The pdf versions of manuscripts created during the electronic submission process typically adds page and line numbering). Strengths and concerns should be specifically identified. Actionable comments are very useful (Lee, 1995). Reviewers seek clarity of topic or question, details on the selected qualitative research approach, and details on the sample employed and a short rationale on its selection and merits. Reviewers look for a “road map” that guides them through the material in a useful manner. A statement of how human research participants were protected during the study and the write up is always expected—and is oddly often neglected. Reviewers also look at the quality and clarity of the writing, its direct linkage to the collected data, and some reflection of the strengths and limitations of the methods employed are expected. Conclusions and recommendations must stay close to the reported data.

Lee (1995) also calls for reviewers to be kind and thoughtful of the authors. It is good to be frank but not to be devaluing. The worst reviews are short, negative comments that do not guide the authors to improve their work or to learn needed content. Reviewers should also state how their expertise is limited if it impacts their ability to complete the review and point to needed revisions. I had a manuscript turned down by a reviewer who explicitly questioned what “triangulation” was in the content of a study that purposefully compared and contrasted (triangulated) the views of adoptive parents, child welfare workers, and therapists. The reviewer demonstrated that they were not qualified to do a review of this manuscript, but they did not recuse and instead rejected (slammed?) the study apparently based on limited knowledge. Reviewers hopefully maintain a helpful and supportive approach

even while pointing out the limitations of a manuscript. Providing specific guidance on areas of concern found in a manuscript is a valuable professional service—to the authors and to future readers.

Authors frequently find reviewers' comments to be conflicting or inconsistent (Staller, 2013). Such differences may be due to differences in expertise or experience among reviewers; differences that are purposefully sought by journal editors. The goal is to point out potential concerns in a manuscript to guide its improvement. Author, reviewers, and readers may bring different perspectives and even different epistemologies to their reads of a manuscript. Through the review process a dialog is begun among these people, ideally improving the final version of the article. Staller (2013) notes that inconsistent reviews may suggest that the authors must make their thesis clearer, or to more fully and transparently invite and guide readers through the material and conclusions.

As a reviewer, I find it quick and easy to read through well organized and clearly written manuscripts. This is in large part because the author is effectively guiding the reader through the material. Good headings and transitions between paragraphs and sections also help build flow. In contrast, I can spend hours reading through materials that do not seem clear as I attempt to develop a narrative line of my own to integrate the material—a story line which will become a key part of my review content. Authors need to provide a transparent road map through the literature, the relevant empirical studies, and the relevant research approach and methods. Where these are lacking, inconsistent reviews may reflect different “reads” of the material by reviewers.

Poor spelling, changing verb tenses, omitted words, and even shifts in the overall “voice” of a manuscript also suggest more care is needed before submission. One of the most difficult aspects of professional writing is proof reading. Reading each sentence aloud, slowly and carefully, can help identify problems in structure and grammar. It is a challenging process that takes time and cannot be done successfully if rushed. Another option is to ask a skilled colleague to carefully review the manuscript. This often sets up a later return of the favor for the colleague at another time. Graduate students often comment that this process is very different from grading where specific content is sought. Organization, clarity, and flow of ideas is instead the main focus of late-stage manuscript preening.

Conclusions

Writing up qualitative research takes attention and effort because there are so many qualitative research approaches and so many methods of qualitative inquiry. These may be scientific or not, may draw on many different ways of knowing such as Indigenous methods or participatory action research in which power differentials between researchers and participants are altered, and autoethnographies that offer the views of the author in social context with attention to aesthetics. Yet we learn by writing and the thinking it requires. We may find new dimensions of content, refine our knowledge of concepts and theories, and get clearer on our ideas as we write. Thinking, writing, and learning all merge and intersect in professional writing.

Offering our ideas publicly benefits our readers who learn from us, and can benefit authors in supporting their academic careers. Writers also learn about topics of interest. Outside the academy, offering our ideas to others provides a forum for feedback that can help us expand, correct, deepen, and clarify what we know. Writing up qualitative research informs others. What others can learn from our research work includes content, epistemologies and ontologies, values, approaches and methods, and how we organize and write up our qualitative research work. With feedback from others and self-reflection, we learn how effectively our written work has educated others, what our strengths are as researchers and writers, and which areas need continued attention to improve our effectiveness as researcher and communicators. Developing excellence in qualitative research writing is best understood as a career long endeavor.

References

- Allen, M., 2016. *Essentials of Publishing Qualitative Research*. Routledge.
- American Educational Research Association, 2006. Standards for reporting on empirical social science research in AERA publications. *Educ. Res.* 35 (6), 33–40.
- Barbour, R., 2001. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ* 322, 1115–1117.
- Becker, H., 1986. *Writing for Social Scientists: How to Start and Finish Your Thesis*. University of Chicago Press.
- Bochner, A., Ellis, C., 2016. *Evocative Autoethnography: Writing Lives and Telling Stories*. Routledge.
- Cheek, J., 2008. A fine line: positioning qualitative inquiry in the wake of the politics of evidence. *Int. Rev. Qual. Res.* 1 (1), 19–32.
- Denzin, N., 2001. The reflexive interview and a performative social science. *Qual. Res.* 1 (1), 23–46.
- Denzin, N., Lincoln, Y., 2017. *The Sage Handbook of Qualitative Research*, fifth ed. Sage Publications.
- Drisko, J., 1997. Strengthening qualitative studies and reports: standards to enhance academic integrity. *J. Soc. Work. Educ.* 33, 187–197.
- Drisko, J., 1999. Reviewing a manuscript: the reviewer's experience and standards. *Fam. Soc.* 80 (4), 417–419.
- Drisko, J., 2005. Writing up qualitative research. *Fam. Soc.* 86 (4), 589–593.
- Drisko, J., 2008. How is qualitative research taught at the Master's level? *J. Soc. Work. Educ.* 48, 85–101.
- Drisko, J., 2013. Standards for qualitative studies and reports. In: Fortune, R., Reid, W., Miller, R. (Eds.), *Qualitative Research in Social Work*, second ed. Columbia University Press, pp. 3–34.
- Drisko, J., Hunnicutt, C., Berenson, L., 2015. A national content analysis of PhD program objectives, structures and curricula: do programs address the full range of social work's needs? *J. Teach. Soc. Work* 35 (1–2), 14–28.
- Elsevier, 2020. Find Journals online search tool. <https://journalfinder.elsevier.com/>.
- Emmons, C.V., Moravcsik, A., 2020. Graduate qualitative methods training in political science: a disciplinary crisis. *Polit. Sci. Polit.* 53 (2), 258–264.
- ERIC, 2020. Search Educational Resources online search tool. <https://eric.ed.gov/?journals>.
- Gergen, M., Gergen, K., 2011. Performative social science and psychology. *Hist. Soc. Res.* 36 (4), 291–299.

- Gilgun, J.F., 2005. "Grab" and good science: writing up the results of qualitative research. *Qual. Health Res.* 15 (2), 256–262.
- Glaser, B., Strauss, A., 1967. *The Discovery of Grounded Theory*. Sociology Press.
- Goldschmidt, W., 1977. Anthropology and the coming crisis: an autoethnographic appraisal. *Anthropologist* 79 (2), 293–308.
- Gómez, C., Carlos, C., 2009. Assessing the quality of qualitative health research: criteria, process and writing [49 paragraphs]. *Forum Qual. Soc. Res.* 10 (2). Art. 17. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0902178>.
- Guinier, L., Smith, A.D., 2001. Rethinking power, rethinking theater: a conversation between Lani Guinier and Anna Deavere Smith. *Theater* 31 (3), 31–45.
- Haseman, B., 2006. A manifesto for performative research. *Media Int. Aust.* 118 (1), 98–106.
- Humphrey, R., Simpson, B., 2012. Writes of passage: writing up qualitative data as a threshold concept in doctoral research. *Teach. High. Educ.* 17 (6), 735–746.
- Kellogg, R.T., 1988. Attentional overload and writing performance: effects of rough draft and outline strategies. *J. Exp. Psychol. Learn. Mem. Cognit.* 14 (2), 355–365.
- Kross, J., Giust, A., 2019. Elements of research questions in relation to qualitative inquiry. *Qual. Rep.* 24 (1), 24–30.
- Lahman, M., 2018. *Ethics in Social Science Research: Becoming Culturally Competent*. Sage Publications.
- Lee, A.S., 1995. Reviewing a manuscript for publication. *J. Oper. Manag.* 13, 87–92.
- Levitt, H., Bamberg, M., Creswell, J., Frost, D., Josselson, R., Suárez-Orozco, C., 2018. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology. *Am. Psychol.* 73 (1), 26–46.
- O'Brien, B., Harris, I., Beckman, T., Reed, D., Cook, D., 2014. Standards for reporting qualitative research. *Acad. Med.* 89 (9), 1245–1251.
- Performance Research, 2020. *The Journal*. Retrieved from: <https://www.performance-research.org/>.
- Petty, N., Thomson, O., Stew, G., 2012a. Ready for a paradigm shift? Part 1: introducing the philosophy of qualitative research. *Man. Ther.* 17 (4), 267–274.
- Petty, N., Thomson, O., Stew, G., 2012b. Ready for a paradigm shift? Part 2: introducing qualitative research methodologies and methods. *Man. Ther.* 17 (5), 378–384.
- Pratt, M., 2009. For the lack of a boilerplate: tips on writing up (and reviewing) qualitative research. *Acad. Manag. J.* 52 (5), 856–862.
- Scimago Labs, 2020. Scimago Journal and Country Rank for Education. Online search tool. <https://www.scimagojr.com/journalrank.php?category=3304>.
- Shields, L., King, S., 2001a. Qualitative analysis of the care of children in hospital in four countries—part 1. *J. Pediatr. Nurs.* 16 (2), 137–145.
- Shields, L., King, S., 2001b. Qualitative analysis of the care of children in hospital in four countries—part 2. *J. Pediatr. Nurs.* 16 (3), 206–213.
- Staller, K., 2013. Writing and reading peer reviews. *Qual. Soc. Work* 12 (6), 715–721.
- Stenius, K., Mäkelä, K., Miovský, M., Gabrhelík, R., 2017. How to write publishable qualitative research. In: Babor, T., Stenius, K., Pates, R., Miovský, M., O'Reilly, J., Candon, P. (Eds.), *Publishing Addiction Science: A Guide for the Perplexed*. Ubiquity Press, pp. 155–172.
- Sword, H., 2017. *Air & Light & Time & Space: How Successful Academics Write*. Harvard University Press.
- Tong, A., Sainsbury, P., Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int. J. Qual. Health Care* 19 (6), 349–357.
- Turk, M., Kalarchian, M., 2014. What makes a good qualitative research article? *Bariatr. Surg. Pract. Patient Care* 9 (1), 26–28.
- van den Hoonaard, W. (Ed.), 2002. *Walking the Tightrope: Ethical Issues for Qualitative Researchers*. University of Toronto Press.
- van den Hoonaard, W., van den Hoonaard, D., 2013. *Essentials of Thinking Ethically in Qualitative Research*. Left Coast Press.
- Wang, C., Burris, M.A., 1997. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ. Behav.* 24 (3), 369–387. <https://doi.org/10.1177/109019819702400309>.
- Wolcott, H., 2002. *Sneaky Kid and Its Aftermath: Ethics and Intimacy in Fieldwork*. AltaMira Press.
- Wolcott, H., 2009. *Writing Up Qualitative Research*. Sage Publications.

Mixed methods research proposal

Jessica T. DeCuir-Gunby^a and **Whitney N. McCoy^b**, ^a University of Southern California, Los Angeles, CA, United States; and ^b Duke University, Durham, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

What is special about a mixed methods research proposal?	729
Components of a mixed methods research proposal	730
Introduction	731
Background	731
Research aims	731
Conceptual framework/inquiry worldview	731
Literature review	731
Methods	732
Research questions/hypotheses	732
Mixed methods designs	732
Mixed methods procedural diagram	732
Participants/context/sampling	733
Data sources	733
Data collection procedures	733
Data analysis strategies	733
Data integration	734
Data validation and reliability	734
References and appendices	734
Institutional Review Board documents	734
Questionnaires/interview questions	735
Timeline	735
Budget	735
Issues to consider when developing a mixed methods research proposal	735
Title	735
Abstract	735
Data management	736
Extenuating circumstances	737
Technology usage	737
Conclusion	737
References	737

Mixed methods research (MMR) involves collecting and analyzing both quantitative and qualitative data within the same research study. More specifically, MMR is considered “research in which the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry” (Tashakkori and Creswell, 2007, p. 4). Although MMR was conceived as a distinct research method in the 1970s/1980s, it was not until the early 2000s did it begin to grow in popularity (Creswell and Plano Clark, 2018). Now MMR can be seen as a third distinct research method to accompany quantitative and qualitative approaches. MMR is now a common research approach used in a variety of academic disciplines including the health sciences, education, and the social sciences.

Regardless of the academic discipline, in order to engage in a good MMR study, it is necessary to have a research plan, to begin with a research proposal. The purpose of this chapter is to discuss the importance of the mixed methods research proposal (MMRP), how it is used, and its various components. We will begin our discussion by describing the goals of the MMRP. We then provide details regarding the different sections or components of the MMRP. Last, we discuss challenges or issues associated with writing a MMRP. We also provide resources on how to write research proposals and how to conduct mixed methods research for a research study and/or a funding agency.

What is special about a mixed methods research proposal?

Researchers cannot undertake a study without having a plan. Therefore, every research study requires a detailed plan of action. In the case of MMR, the plan is called a mixed methods research proposal (MMRP). The MMRP is a detailed research plan that describes the steps that are required and needed to undertake a MMR study. The MMRP outlines everything that will be done within the study,

Table 1 Components of the mixed methods proposal.

<i>Mixed methods proposal sections</i>	<i>Mixed methods proposal sub-sections</i>
Title page	
Abstract	
Introduction	Background Conceptual framework/inquiry worldview Purpose/research aims
Literature review	
Methods	Research questions/hypotheses Mixed methods design Research question/hypothesis Participants/context/sampling Data sources Data collection procedures Data analysis strategies Data integration Data validation and reliability
References	
Appendices	Mixed methods procedural diagram IRB documents Questionnaires/interview questions Timeline Budget

as well as provides parameters for what will not be done within the study. The MMRP is tentative in that it is subject to change and should be seen more as a flexible contract between the researcher and their stakeholders. However, researchers should follow the proposal as closely as possible and deviate only when extremely necessary.

According to [DeCuir-Gunby and Schutz \(2017\)](#), the MMRP has three major sections (introduction, literature review, and methods) and two minor sections (references and appendices). See [Table 1](#) for a listing of the proposal components. The three major sections are the most important sections of the proposal. The introduction section provides an overview of the entire proposal by focusing on the research aims and purpose. It also provides the conceptual framework or inquiry worldview that guides the study. The literature review helps to situate the study into a larger context by making connections to the extant literature. The methods section focuses on the actual data collection and analysis process, including a discussion of data integration, a major component of mixed methods research. The two minor sections, the reference section and the appendices, are both needed as support for the three major sections. The reference section lists all of the citations used to create the document. The appendices are a collection of documents that are relevant to understanding the proposed study and help to provide the needed context for the study.

There is no standard way or a set of common rules to develop a MMRP. According to [Hesse-Biber \(2010\)](#), a MMR study can be qualitative-dominant (emphasis on the qualitative aspects), quantitative-dominant (emphasis on the quantitative aspects), or mixed (emphasizing both the qualitative and quantitative components equally), which in turn will influence the manner in which a MMRP is developed. There is also variation depending upon academic discipline as well as research proposal type including dissertations/theses, grants, and general research studies. However, there are components that should be included in all MMRPs. The remaining discussion will focus on the components that should be included in all MMRPs, with distinctions being made for research proposal type when necessary.

Components of a mixed methods research proposal

As previously mentioned, there are three major components (introduction, literature review, and methods) and two minor components (references and appendices) of a MMRP. However, within these components, there are several sub-components that are central to developing a MMRP. The explanation of the introduction includes a discussion of the research background, research aims, and the conceptual framework/inquiry worldview. The literature review discusses how the research literature is explored and connected to the research purpose. The methods section involves an explanation of research questions/hypotheses, mixed methods designs, sampling/participants/context, data collection procedures, data analysis procedures, data integration, and data validation. For this discussion, the references and appendices are explained in a combined section. The importance of the references are explored as well as what is to be included in the appendices such as the mixed methods procedural diagram, IRB protocol, questionnaires/interview protocols, timelines, and budgets.

Introduction

The introduction provides necessary background information and gives readers an overall sense of the research topic. This section of the proposal is important because it allows readers to explore the problem that needs to be solved by the research, the current state of the topic, issues or controversies surrounding the topic, the goals of the research, potential policy implications of the work, and how the researcher's perspective will inform/guide the study. While this section is typically one of the shorter sections of the proposal, a concise, succinct approach is needed to effectively convey the researcher's stance on the topic. Thus, to begin this section, the writer should situate the study, provide empirical evidence and references regarding where the study stems from, provide an overview of what will be shared throughout the proposal, and most importantly engage the reader in a way that if they are unfamiliar with the topic, they can easily understand the content that is being discussed.

Background

Framing a background or problem statement allows knowledge gaps and issues to provide context for why the study is necessary. This should be stated clearly, and should be a focused statement on the constructs that will be addressed within the research study. This section can highlight statistics on the topic of interest, briefings on the topic, history of where the field stands on the subject matter, and a clear delineation on where future research or the current proposal aims to add to the current literature in the field.

Research aims

It is essential that the proposal have a clear description of the goals or what the researcher hopes to discover within the proposed study. While the goals of the study may change, knowing how to focus the study will help the researcher address the key concepts. Along with including a rationale for the study, the researcher should specify why this work is different from other work that has been previously conducted, discuss the significance of how the approach will generate new knowledge, and explain general understandings regarding the topic. MMRP aims align with the components of the design of the study. For example, in a sequential design, the researcher could focus on the quantitative aim before moving to the qualitative aim. Also, the aims can be qualitative, quantitative, or mixed methods focused. Some common mistakes made within MMRPs include having too many aims, being vague, or developing unachievable goals. When developing research aims, make sure they are measurable and can be realistically achieved with the given time frame. These aims will guide the study as they provide further context regarding what the researcher intends to achieve by carrying out the study in its entirety.

Conceptual framework/inquiry worldview

The conceptual or theoretical framework is a broad set of ideals used to guide and understand the development of the research process. Oftentimes in research, one theory cannot sufficiently address all of the issues or frame the problem that needs to be solved. A theoretical framework is based on existing theories and models that ground the study. A conceptual framework is a display that demonstrates how constructs are connected based on what the researcher intends to observe in the study. Thus, instead of just using a theoretical framework, a conceptual framework allows for multiple theories to situate the research study. [Creswell and Plano Clark \(2018\)](#) suggest using the theory to guide all aspects of the research such as questions, variables, analysis, and integration. By using concepts, relevant literature, and models relevant to the problem, the framework helps guide the researcher to evaluate the MMR in a more critical manner (see [Shannon-Baker, 2022](#)).

In addition to the theoretical perspectives that will guide the research, the researcher's individual experiences and beliefs, also known as the inquiry worldview ([DeCuir-Gunby and Schutz, 2017](#)), will guide the direction of the proposal. Being able to identify how the researcher's multiple identities (e.g., race, gender, religion, socioeconomic status, geographic location) contribute to the process in which they engage with data collection and data analysis, as well as how the writing is influenced is essential. The interpretation of research allows readers to make meaning based on the results that are reported, which is related to the subjectivity and voice of the researcher. While some may interpret their own inquiry worldview as problematic, this is the true nature of how research evolves, so it is vital that the researcher remains aware and conscious of personal biases and share them with the reader within the proposed research.

Literature review

The literature review provides a foundation for the entire proposal. Within this section, it is important to highlight the gaps within the existing literature as well as define the constructs, theories, and other information that is needed for the reader to understand the proposed study. While the researcher may not have read everything regarding the content area, this section should demonstrate that the researcher is an expert when it comes to the problem that is being addressed. To complete the literature review, the researcher should engage in a critical analysis of how the literature has evolved over time. By reviewing seminal pieces related to the topic as

well as recent developments within the past 5–10 years, the work should discuss what scholars have written about the topic that is being studied, the theories utilized, and how this connects to the approach that the researcher will use. As indicated by [Xiao and Watson \(2019\)](#), the inclusion criteria used within search engines can assist in the searching of articles that are relevant to the research pertinent to the study. The literature review should provide a clear explanation of why specific theoretical frameworks, constructs, and/or methodologies should be utilized to address the particular issue and questions that guide the study.

Within the literature review, the researcher should demonstrate the individuality of the research problem that is being presented. After sharing the big picture of the gaps in the literature, the review should situate the current topic among the larger discipline work. It should also discuss why using mixed methods research may provide stronger evidence for your research. While developing the literature review, using a reference manager or spreadsheet may assist with the organization of ideas such as the specific citations, research questions that are addressed, the approach/theoretical framework, measures utilized, analysis, key findings, and other notes that can guide the study. This synthesis can provide guidance for how the researcher should approach the study methodologically. Moreover, the literature review should not be a list of all the work that has been done, rather it should provide a more conversational tone of the vast or dearth of literature that is present, allow the researcher to justify why the work is needed, and state how the research can benefit society in the future.

Methods

The next section, the methods section, is perhaps the most important component of the MMRP. As described by [DeCuir-Gunby and Schutz \(2017\)](#), it is the “heart of the mixed methods research plan” (p. 107). The methods section explains the specific data that will be collected, how the data will be collected and analyzed, as well as the validation and reliability process. Because this is a mixed methods proposal, the methods section also discusses aspects that are specific to mixed methods research including the mixed methods design to be used and data integration plans. The methods section is comprehensive and extremely detailed. It is necessary that the researcher painstakingly describes all steps of the research process for readers within this section of the proposal. It is important that the subcomponents of the methods sections adhere to the particular mixed methods design chosen. For instance, if using one of the sequential designs, it may be necessary to discuss the different subcomponents according to the design phases. Ultimately, the subcomponents of the methods section are all necessary for a MMRP although the order may differ based upon academic discipline or personal writing preferences.

Research questions/hypotheses

The first section of the methods section is generally the discussion of the research questions/hypotheses. Research questions are the questions that are guiding the research investigation. Hypotheses are the specific relationships that are to be tested within a study. It should be noted that qualitative studies do not focus on hypotheses while quantitative studies generally feature research questions or hypotheses. In a mixed methods study, research questions can be quantitative ([Onwuegbuzie and Leech, 2006](#)), qualitative ([Agee, 2009](#)), or mixed methods ([Clark and Badiee, 2010](#)). Research questions are important because they help shape the components of the proposal. Although we recommend including the research questions within the research methods section, many researchers often include the research questions within the introduction section, particularly in dissertation proposals.

Mixed methods designs

Different research questions are best answered by using specific mixed methods research designs. In mixed methods research, there are five basic or core designs: explanatory sequential (e.g., qualitative data explaining quantitative data), exploratory sequential (e.g., quantitative data building upon the qualitative findings), convergent parallel (e.g., quantitative and qualitative data collected and analyzed separately but findings merged), embedded (e.g., qualitative data embedded within quantitative data), and multiphase (e.g., the alternating of quantitative and qualitative phases) ([Creswell and Plano Clark, 2018](#)). There are also more complex, advanced mixed methods designs that can be used such as case study mixed methods designs ([Cook and Kamalodeen, 2020](#); [Guetterman and Fetters, 2018](#)), action research mixed methods designs ([Ivankova and Wingo, 2018](#)), transformative mixed methods designs ([Mertens, 2012](#)), scale development designs ([Taghipoorreynah and deRun, 2020](#)), mixed methods phenomenological research ([Mayoh and Onwuegbuzie, 2015](#)), and fully integrated mixed methods designs ([Creamer, 2017](#)), among many others. For more details regarding mixed methods designs, see [Creswell and Plano Clark \(2018\)](#). Regardless of the chosen design, it is imperative that the designs are explained in detail. The mixed methods design will dictate how the data are collected and analyzed, as well as provide a framework for how the final write up will be organized or even disseminated. There can also be variability across disciplines in how designs are implemented.

Mixed methods procedural diagram

The MMRP should include a visual of the mixed methods design. The complexity of mixed methods research can make it difficult to articulate and describe what is needed to execute the research. A visual or graphic, also known as a procedural diagram, can be utilized to easily communicate the sequence of the design, as well as the procedures, and products that will be produced

throughout. [Ivankova et al. \(2006\)](#) suggest ten rules to drawing a procedural diagram for mixed methods research. The procedural diagram should have a name/title that is reflective of the research design that will be used, can be positioned in a vertical or horizontal manner, and boxes should include all stages of the research for quantitative and qualitative data collection, analysis, and integration. Arrows and proper nomenclature will allow the reader to understand the focus of the design, the approach, and what type of research is more dominant. Providing a descriptive procedure for data collection and analysis, as well as what will result from the data collection (e.g., artifacts, transcripts, survey data), will help connect the sequence of events with the expected products that will guide the results. The model will serve as a guideline as to how the research will be completed, making it easier to share the direction of the research with others.

Participants/context/sampling

In addition to explaining the design, it is important to discuss who will be involved with the data collection process. More specifically, the methods section involves a detailed discussion of the participant descriptions, the context involved in the study, and the sampling methods. It is imperative that the researcher understands the study's population of interest. Providing a description of the qualities and demographic characteristics that the participants will possess is necessary to help the readers better understand how the research questions will be addressed. Once the participants are determined, it will be essential to choose the appropriate context(s). The context is the environment in which the participants live, work, and/or experience the phenomena of interest. Sometimes the context is just as important as the actual individual participants. Providing a description of the research context provides more background information regarding the participants in the study. As such, sampling in a research study is extremely important. The sampling method is the strategy used to gain access to the data in the study, particularly any participants that will be actively involved. According to [Collins et al. \(2007\)](#), in mixed methods there are four major types of sampling: identical sampling (the same participants are used in the quantitative and qualitative phases), parallel sampling (different participants from the same population are used in the quantitative and qualitative phases), nested sampling (a subset of participants in one phase participates in the other phase of the study), and multilevel sampling (different people from the population are used for the various phases of the study). Describing the sampling process is necessary for helping to determine the manner in which the data are collected.

Data sources

Within a methods section, it is necessary to list and describe in detail the data that will be collected or used in the study. Some studies rely upon existing data sets (e.g., previously collected national longitudinal data set) or artifacts (e.g., historical documents) while others may require the researcher collecting new data sources. Common quantitative data sources include survey instruments, tests (e.g., psychological, physiological, academic), and observation protocols. Qualitative instruments often include interview protocols (individual or focus group) or observation protocols. Regardless of the type of data source, the researcher should fully describe the instrument including the purpose, the structure, and any reliability/validity information that is available. Also, when describing the data sources, it is important to distinguish between both the quantitative and qualitative data as well as any data sources that could be analyzed in a manner that results in both quantitative and qualitative data.

Data collection procedures

It is necessary to explain everything about the data collection process. As stated by [DeCuir-Gunby and Schutz \(2017\)](#), this includes the *when*, *where*, *how* and *why*. The researcher must explain *when* in terms of the time of day, month, year, etc. *Where* involves explaining the context (e.g., participants' homes, on campus, at a hospital) of the data collection. The *how* gets into the explicit details. It involves providing all of the steps that are needed to collect the data. How will the researcher get access to the existing data or the participants? What will be the role of the participants? What will the participants do? How long will the tasks take? Then the researcher will need to address the *why* in terms of explaining how all of this will be done and why these steps are necessary.

Data analysis strategies

After the data are collected it will be analyzed. Thus, it is imperative that the researcher have a thoughtful plan for data analysis. The data analysis plan should involve three major stages: data cleaning/organization, data exploration, and final data analysis. There should be a detailed plan for how the data will be cleaned and organized once it is collected. This will include how any software will be used for data management. Data cleaning/organization involves searching and replacing missing data as well as creating code books for both quantitative and qualitative analysis. Data cleaning/organization may be one of the most undervalued yet most essential steps in a research study. The manner in which data are cleaned and organized has the potential to impact the quality of data analysis. As such it is recommended that the researcher invest substantial time in this stage. After the data has been thoroughly cleaned and organized, it is important for the researcher to explore the data noting trends. In terms of quantitative data this will involve descriptive statistics. For qualitative data this may involve reading interviews and writing memos regarding potential themes. Exploratory data analysis is helpful in terms of determining the next best steps for analysis. It's possible that based upon the trends found in the data that new data analysis strategies may emerge, or changes may have to be made regarding the planned data analysis strategies. The last stage is the final analysis stage. This is where the researcher describes in detail how the data will be

analyzed. In terms of the quantitative data, specific statistical procedures will be described. For the qualitative data, the coding process will be described including how codes will be created and applied to the data. It also involves describing how the inquiry worldview will be used within the data analysis process. Researchers can also engage in mixed methods data analysis which involves the mixing of quantitative and qualitative data concurrently or sequentially (Onwuegbuzie and Teddlie, 2003).

Data integration

A core component of mixed methods research is the integration of the quantitative and qualitative data. The MMRP must describe how the quantitative and qualitative data or findings will be mixed or integrated. Data integration is important because it illustrates what was found in the study and why the collection of both quantitative and qualitative data were important for your study. Data integration can occur in the form of merging (combining the quantitative and qualitative data or findings), explaining (using one form of data to explain another), building (using one form of data to build upon another), or embedding (one form of data are embedded within another form of data) (Creswell, 2014). Data integration can occur in any stage of the research process, but most researchers engage in data integration during the data interpretation and reporting stages. According to Fetters et al. (2013), the common data integration approaches are narrative (integrating the quantitative and qualitative findings in a narrative account), data transformation (quantifying qualitative data or qualitzing quantitative data), and joint display (using visualizations to represent the data) (see Guetterman, 2022).

Data validation and reliability

The last subcomponent of the methods section discusses the data validation process, or legitimation (Onwuegbuzie et al., 2011). Data validation involves discussing the process for determining whether or not the claims being made are accurate while reliability addresses whether or not claims are consistent. In qualitative research this is often described as trustworthiness and credibility (Lincoln and Guba, 1986). Within mixed methods research, there needs to be a plan for how validity/trustworthiness and reliability/credibility will be addressed in terms of the separate quantitative and qualitative components as well as the mixed methods study as a whole. Mixed methods validation is concerned with inference making, specifically the metainferences that are being made from the combining of the quantitative and qualitative data (Tashakkori et al., 2021). It is also important to consider the consequences of the study and how the findings will be used (Dellinger and Leech, 2007).

References and appendices

The last section of the research proposal includes the references and the appendices. The references and appendices include working sources that guide or inform the research that will be conducted. Citations that are included in the proposal should be presented in alphabetical order according to the publication manual that guides the researcher's academic discipline (e.g., MLA, APA, Chicago). Additionally, some citations should be referred to in numeric order, based on the discipline (e.g., Vancouver, AMA, IEEE). The appendix section, on the other hand, allows the researcher to share any relevant information that could not be added within the proposal narrative. The information included in the appendix section depends upon the purpose of the proposal and what stakeholders may need to review the proposal. The appendices should include information such as the mixed methods procedural diagram, the Institutional Review Board (IRB) research protocol, and other documentation that is used for the prospective research.

Institutional Review Board documents

Some, not all, researchers like to include copies of their IRB application. An IRB reviews the methods that will be used to execute research to ensure that it is ethical. To protect participants, the IRB board monitors the safety of the research design by examining the materials (i.e., interview protocol, experimental design, recruitment flyers, data management), as well as reviewing updates and changes throughout the research, such as participants that withdraw from the study, protocol deviations, and amendments. IRB and ethics boards in numerous countries may have varying foci and regulations. DeCuir-Gunby and Schutz (2017) mentions meeting with IRB staff and discussing the IRB process with others as a way to learn about the process before encountering it. When the IRB is submitted, all protocols and surveys must be included, as well as information regarding how data will be shared with other researchers and how anonymity will be maintained among participants. Potential risks and a plan of action will also need to be detailed should one of these risks present itself to a participant. This information is often included in a MMRP, particularly in dissertations. An IRB board consisting of a committee with various backgrounds will convene and decide whether the study should be approved based on the risks and benefits of the study. It is important to note that should any changes come after initial IRB approval, these should be communicated immediately to the IRB board. Questionnaires, interview questions, and the associated protocols may need to be amended throughout the phases of the study because of changes to the research staff, funding sources, study sample, data findings, data storage, risks associated with the study, or changes in confidentiality. MMRP that are sequential may build on the protocols throughout the duration of the study based on data collection and analysis of the quantitative or qualitative data.

Questionnaires/interview questions

The MMRP should include copies of all questionnaires, tests, and interview protocols to be used in the study. To serve as a reminder to the researcher/interviewer before preparing to disseminate questionnaires or conducting interviews, a protocol is a guide to determining the key components and information that should be collected from the research participants. The interview protocol, specifically, may begin with a script of what to say before the interview begins or pertinent information that is needed before completing a questionnaire. This should address issues of confidentiality (in some instances participants can disclose a pseudonym to be used throughout the interview) and also remind the participant of the informed consent that they have/may need to sign. To make the participants more comfortable, the researcher(s) should also introduce themselves by including information such as their name, background/position, and why they are conducting the study. Clear, concise language should be used to alleviate any concerns from the participant. In addition, at the end of the questionnaire/interview, a closing script should provide information regarding how to contact the researcher/interviewer should the participant need to clarify anything throughout the duration of the study.

Timeline

Although not required, many MMRPs often include timelines. Timelines are particularly helpful for grants and dissertation proposals. Developing a schedule or timeline will provide a realistic plan based on institutional or organizational deadlines. By goal-setting, the researcher can understand all the elements and tasks that must be completed throughout the duration of the research. While planning, the researcher should account for elements such as IRB submission and approval, study design execution, and purchasing supplies. However, many of these tasks will take longer than anticipated. One should ask, how long will it take to organize survey data, clean and transcribe transcripts, run analyses and code data, as well as integrate the findings before actually writing up the findings of the study? Therefore, making a plan allows the researcher to effectively engage in time management and to execute the research proposal in an efficient manner.

Budget

Whether you are applying for funding or your research is self-funded, a budget can assist the researcher in planning the research in detail and allow the researcher to assess the use of funds throughout the duration of the study. Including a budget is a requirement for grant proposals but generally helpful for dissertations as well. To develop a budget, first list all the activities needed to complete the research (e.g., research equipment, study materials, transcription services, data analysis software costs, participant compensation). Consulting with the university or institutional regulations regarding budget spending will allow the researcher to examine that they are using the funds as outlined by the guidelines. While it may seem like paying participants or using two transcription services is allowable, every organization has specific rules about fund usage. After doing this and estimating (as closely as possible) what the intended costs are, utilizing a spreadsheet to make cost estimates, considering the scope of the research, will provide a method to calculate costs. Lastly, for each item be sure to justify why the expense is essential to the research.

Issues to consider when developing a mixed methods research proposal

When developing a MMRP, the nature of the research can allow for various issues to arise when planning, developing, and executing the research. The amount of skill level needed to produce the proposal requires researchers to consider multiple facets of the process. The execution and writing delivery of the proposal, data management, and extenuating circumstances can all contribute to the success or demise of the proposal. In order to address these potential issues, it is important to consider supporting and guiding resources (see [Table 2](#)).

Title

It is important to create a title that captures the essence of the proposal. Specifically, [Creswell and Plano Clark \(2018\)](#) suggest that titles for mixed methods proposals should include the participants and location of the study, the phrase “mixed methods,” as well as reflect the type of mixed methods design being used. What is most important is that the researcher captures the main topics/constructs and populations of interests. When creating a title it is important for the researcher to keep in consideration the academic discipline’s publication manual requirements for the title’s word limits.

Abstract

This component of the proposal provides the essence of the paper in brevity, before going into further detail. It allows readers to decide whether or not the research is of interest to them. By using approximately 100–350 words, the informative abstract should introduce the research problem and indicate why the research problem is important. This can be done by providing an

Table 2 Resources for writing a mixed methods proposal.

<i>Topic</i>	<i>Resource</i>
Mixed methods proposals	DeCuir-Gunby, J.T., Schutz, P.A., 2017. <i>Developing a Mixed Methods Proposal: A Practical Guide for Beginning Researchers</i> . SAGE Publications, Thousand Oaks, CA.
Publishing mixed methods	Fetters, M.D., 2019. <i>The Mixed Methods Research Workbook: Activities for Designing, Implementing, and Publishing Projects</i> , vol. 7. SAGE Publications, Thousand Oaks, CA.
Designing and conducting mixed methods	Creswell, J.W., Clark, V.L.P., 2017. <i>Designing and Conducting Mixed Methods Research</i> . SAGE Publications, Thousand Oaks, CA. Douglas, J., Conducting intersectionality-informed mixed-methods research to explore cigarette smoking among young African-Caribbean women in the UK. In: <i>Third Regional Mixed Methods Research Conference</i> . The University of the West Indies, Trinidad and Tobago. Fàbregues, S., Escalante-Barrios, E.L., Molina-Azorin, J.F., Hong, Q.N., Verd, J.M., 2021. Taking a critical stance toward mixed methods research: a cross-disciplinary qualitative secondary analysis of researchers' views. <i>PLoS One</i> 15(7), 1–21. Mannell, J., Davis, K., Akter, K., Jennings, H., Morrison, J., Kuddus, A., Fottrell, E., 2021. Visual participatory analysis: a qualitative method for engaging participants in interpreting the results of randomized controlled trials of health interventions. <i>J. Mixed Methods Res.</i> 15(1), 18–36. Poth, C.N., 2018. <i>Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity</i> . SAGE Publications. Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. <i>Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences</i> . SAGE Publications, Thousand Oaks, CA, Incorporated. Scherman, V., Liebenberg, L., 2021. Item response theory integrating qualitative data. In: <i>The Routledge Reviewer's Guide to Mixed Methods Analysis</i> . Routledge, New York, NY, pp. 117–124.
Writing literature reviews	Ridley, D., 2012. <i>The Literature Review: A Step-by-Step Guide for Students</i> . SAGE Publications, Thousand Oaks, CA.
Developing mixed methods research questions	Onwuegbuzie, A.J., Leech, N.L., 2006. Linking research questions to mixed methods data analysis procedures. <i>Qual. Rep.</i> 11(3), 474–498. Clark, V.L., Badiee, M., 2010. Research questions in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), <i>Handbook of Mixed Methods in Social and Behavioral Research</i> , second ed. Sage, Thousand Oaks, CA, pp. 275–304.
Writing dissertation and grant proposals	Locke, L.F., Spirduso, W.W., Silverman, S.J., 2013. <i>Proposals That Work: A Guide for Planning Dissertations and Grant Proposals</i> . SAGE Publications, Thousand Oaks, CA.
Developing references in APA style	American Psychological Association, 2020. <i>Publication Manual of the American Psychological Association</i> , seventh ed. Washington, DC.

introductory sentence with historical background, identifying the purpose, theoretical framework, study methodology, procedures, and key findings or predicted findings. To conclude the abstract, share recommendations and implications that capture the nature of how the work can impact society. The researcher should be concise, clear, and not include references in this section. In some instances, abstracts may ask for less than 100 words (i.e., descriptive abstract), which will require the researcher to provide less detail and highlight only key components of the work. Writing this section toward the end of the proposal development may be easier considering that the researcher has gained clarity on what is needed to execute the study in its entirety (DeCuir-Gunby and Schutz, 2017).

Data management

Throughout the research process, there will be a host of materials that will need to be organized, stored, and shared with others. The daily management of these resources require consistent conventions that will allow the researcher to easily access and archive the materials since MMRP usually contain a large amount of qualitative and quantitative data. Therefore, it is necessary to understand how to be strategic regarding the organization of data, how to integrate data to fully answer the aims of your study, and how to leverage the resources needed to access and store data (see Niglas et al., 2022, Chapter 11063). Data can include anything such

as audio recordings, surveys, physical artifacts, images, etc., and effectively organizing these materials will save the researcher time, stress, and in the long run, reduce the amount of errors that can lead to poor data analysis and validity, as well as risk the anonymity of the participants' information. When considering how data will be managed, exploring the university or organizational processes and policies is vital. Each entity may have different rules and regulations regarding what type of data management system must be used. For example, the researcher may want to use Google Drive, but the university utilizes a password protected drop box program to protect their data. Depending on the amount of data, the kind of data, privacy, and the roles of other researchers on the research team, the data management system that is utilized should consider all of these factors.

Extenuating circumstances

Unprecedented occasions may push researchers to seek guidance outside of the typical IRB guidelines to conduct research in an ethical manner. During the COVID-19 pandemic, while research is a major function of many universities and other institutions, it is critical that researchers follow the strict guidelines provided by state mandates to maintain personal safety and the safety of human subjects. Thus, in order to continue to transform the research community, researchers should look to collaborate based on financial and personal constraints and provide good leadership during difficult times to build a better research culture. While research is important, it should not risk the well-being of those that are being studied. Researchers should consider ways in which the study can be conducted without harming those that are being studied and evaluate the effectiveness of their research design within the parameters of extenuating circumstances.

Technology usage

When conducting research, many researchers engage in technology usage. Whether it be a software program for data analysis, voice recorders, a PowerPoint presentation, or a tool that is used to conduct the research, all of these are utilized for the completion of the study (see Kuckartz de, 2022). Technological advances allow researchers to conduct the study in an effective, timely manner, but the researcher should always consider what process will occur when technology fails or is inaccessible. Technology has varying costs so budgeting the cost of software programs based on the amount of time they will be utilized, as well as investigating what resources can be accessed at a reduced or no cost should be considered. In addition to costs, creating a system to effectively collect and retain technology (e.g., voice recorders, iPads) that participants may use is vital, so labeling them in a way that is easily understood and providing protective covers may be necessary. The other consideration is what to do when technology fails. This is something that no one wants to deal with, but we know that it can happen and often cannot be prevented. Thus, checking the technology before it is used is important and having a backup is always good. Taking preventative measures such as keeping paper copies of surveys, interviews, and backing up data throughout the research process can assist with technological issues throughout the proposal and executions phases.

Conclusion

Writing a MMRP can seem like a daunting task. It requires the understanding of mixed methods research approaches, the proposal writing process, as well as how to merge the two approaches. Like with any research proposal, a MMRP requires a lot of time and planning. Depending upon the purpose of the proposal, it may even require a team approach. Following the recommended guidelines for writing a MMRP is necessary for a successful proposal development process. Regardless, executing a well-planned and detailed MMRP will enable a comprehensive and successful research study.

References

- Agee, J., 2009. Developing qualitative research questions: a reflective process. *Int. J. Qual. Stud. Educ.* 22 (4), 431–447.
- Clark, V.L., Badiie, M., 2010. Research questions in mixed methods research. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, second ed. Sage, Thousand Oaks, CA, pp. 275–304.
- Collins, K.M., Onwuegbuzie, A.J., Jiao, Q.G., 2007. A mixed methods investigation of mixed methods sampling designs in social and health science research. *J. Mix. Methods Res.* 1 (3), 267–294.
- Creswell, J.W., 2014. *A Concise Introduction to Mixed Methods Research*. SAGE publications, Thousand Oaks, CA.
- Creswell, J., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE, Thousand Oaks, CA.
- DeCuir-Gunby, J.T., Schutz, P.A., 2017. Developing a Mixed Methods Proposal: A Practical Guide for Beginning Researchers. Sage, Thousand Oaks, CA.
- Dellinger, A.B., Leech, N.L., 2007. Toward a unified validation framework in mixed methods research. *J. Mix. Methods Res.* 1 (4), 309–332.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48 (6), 2134–2156.
- Guetterman, T., 2022. Joint displays in mixed methods research. In: Ivankova, N. (Ed.), *International Encyclopedia of Education*, fourth ed. (Chapter 11076).
- Guetterman, T.C., Fetters, M.D., 2018. Two methodological approaches to the integration of mixed methods and case study designs: a systematic review. *Am. Behav. Sci.* 62 (7), 900–918.
- Hesse-Biber, S.N., 2010. *Mixed Methods Research: Merging Theory With Practice*. Guilford Press, NY.
- Ivankova, N., Wingo, N., 2018. Applying mixed methods in action research: methodological potentials and advantages. *Am. Behav. Sci.* 62 (7), 978–997.
- Ivankova, N.V., Creswell, J.W., Stick, S.L., 2006. Using mixed-methods sequential explanatory design: from theory to practice. *Field Methods* 18 (1), 3–20.

- Kuckartz de, U., 2022. Mixed methods software analysis. In: Ivankova, N. (Ed.), *International Encyclopedia of Education*, fourth ed. (Chapter 11049).
- Lincoln, Y.S., Guba, E.G., 1986. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *N. Dir. Progr. Eval.* 1986 (30), 73–84.
- Mayoh, J., Onwuegbuzie, A.J., 2015. Toward a conceptualization of mixed methods phenomenological research. *J. Mix. Methods Res.* 9 (1), 91–107.
- Mertens, D.M., 2012. Transformative mixed methods: addressing inequities. *Am. Behav. Sci.* 56 (6), 802–813.
- Niglas, K., Mikser, R., Loogma, K., Umarik, M., Timostuk, I., Oun, T., 2022. Conducting large-scale mixed methods studies in education. In: Ivankova, N. (Ed.), *International Encyclopedia of Education*, fourth ed. (Chapter 11063).
- Onwuegbuzie, A.J., Leech, N.L., 2006. Linking research questions to mixed methods data analysis procedures. *Qual. Rep.* 11 (3), 474–498.
- Onwuegbuzie, A.J., Teddlie, C., 2003. A framework for analyzing data in mixed methods research. In: *Handbook of Mixed Methods in Social and Behavioral Research*, vol. 2, pp. 397–430 (1).
- Onwuegbuzie, A.J., Johnson, R.B., Collins, K.M., 2011. Assessing legitimation in mixed research: a new Framework. *Qual. Quant.* 45 (6), 1253–1271.
- Shannon-Baker, P., 2022. Philosophical underpinnings of mixed methods research in education. In: Ivankova, N. (Ed.), *International Encyclopedia of Education*, fourth ed. (Chapter 11037).
- Taghipoorreynah, M., de Run, E.C., 2020. Using mixed methods research as a tool for developing an indigenous cultural values instrument in Malaysia. *J. Mix. Methods Res.* 14 (3), 403–424.
- Tashakkori, A., Creswell, J.W., 2007. The new era of mixed methods. *J. Mix. Methods Res.* 1, 3–7.
- Tashakkori, A., Johnson, R.B., Teddlie, C., 2021. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. SAGE Publications, Incorporated, Thousand Oaks, CA.
- Xiao, Y., Watson, M., 2019. Guidance on conducting a systematic literature review. *J. Plann. Educ. Res.* 39 (1), 93–112.

Reporting mixed methods research studies

June E. Gothberg^a and **Alycia J. Sterenberg Mahon^b**, ^a Cornell University, Battle Creek, MI, United States; and ^b Western Michigan University, Kalamazoo, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Reporting mixed methods research studies	739
The context of mixed methods reporting	740
A step-by-step approach to reporting mixed methods research	741
Framing your paper	741
Titling your paper	742
Abstract	742
Keywords	742
Voice	742
Introduction	743
Problem statement, research questions, and rationale for mixing	743
Methods	745
Research design overview	745
Integrating mixing	746
Sampling	746
Data collection	747
Analysis	748
Findings and results	749
Qualitative findings	750
Quantitative findings	750
Integrated mixed methods findings	750
Discussion and interpretation	750
Responding to the reporting limitations for mixed methods	751
Final thoughts for reporting mixed methods	751
References	751

Reporting mixed methods research studies

Conducting mixed methods research (MMR) can capitalize on the strengths of both quantitative methods and qualitative methods to yield “stronger and more credible evidence” that can help “exclude or minimize potential alternative explanations” and provide “enough information to explain divergent aspects of the phenomenon” (Plano Clark and Ivankova, 2016, p. 9). Often in education, mixed methods are used to explore complex situations that involve multiple and often hierarchical interacting systems that are brimming with institutional uncertainties tied to *wicked problems* (Levin et al., 2012; Rittel and Webber, 1973). For example, research has shown that school attendance plays a vital role in a child’s educational outcomes. Attendance is often studied quantitatively to identify the number of absences for individual students. This information alone does little to improve attendance. Other approaches that include integrating quantitative and qualitative data are needed for solutions to be found. On this issue, Childs and Lofton (2021) advised that

Individually, the root causes of chronic absenteeism are not only in-school and out-of-school factors that contribute to chronic absenteeism, but also aid in the framing of chronic absenteeism as a wicked problem ... As a multifaceted problem, with causes and implications across multiple facets of society, chronic absenteeism should be tackled like other wicked problems ... (p. 214).

As shown, this type of multi-layered multifaceted *wicked problem* is ripe for a mixed-methods research approach. Table 1 shows additional examples of wicked educational problems explored using MMR approaches. The issues shown would be difficult to understand using solely a mono-method approach. For more information, see Onwuegbuzie and Mallette’s chapter on using mixed methods in educational research and Poth and Bullock’s chapter on designing complex mixed methods in this volume.

As mixed methods researchers in the field of education, we need to disseminate our research in ways that readers can fully understand and act. In her final remarks as editor for the Journal of Mixed Methods Research (JMMR), Mertens (2015) admonished that “business as usual will not lead to effective use of research to address wicked problems” and called for researchers to accept the challenge by “using mixed methods to respond” (p. 5). Craddock Lee (2021) shared that “mixed-method results can help bridge the gap for decision makers” improving research-to-practice (p. 99).

Table 1 Research studies using MMR to understand wicked educational problems.

<i>Author and year</i>	<i>Wicked problem</i>	<i>MMR approach</i>
Fredricks et al. (2018)	Gender-based engagement in STEM	Sequential exploratory
Garwood et al. (2018)	Teacher burnout	Sequential explanatory
Hunt-Anderson (2020)	Covert bullying	Transformative
Rouse et al. (2020)	Pregnancy and parenting teens	Concurrent
Theunissen et al. (2012)	Early dropout risk	Sequential explanatory

The purpose of this chapter is to build on the knowledge shared in previous chapters and to share best methods for reporting MMR. We will consider the guidelines from the [American Psychological Association Publication Manual \(2019\)](#) for which many journals and social science programs require, along with recommendations of the best practices highlighted in the seminal and current literature focusing on the critical aspects of reporting for the field of education. This chapter will assist the reader in the complex task of reporting quantitative, qualitative, and mixed methods strands within a single manuscript as well as offer recommendations for generating separate papers for each. We take a pragmatic step-by-step approach that encourages educational researchers to base their reporting choices on the relatedness of the methods and methodologies used, ensuring contextual relevance, and grounded in relevant theory. We include sections to provide guidance for framing your paper; selecting the title, abstract, and keywords; determining voice; writing each section from the introduction to the conclusions broken down into quantitative, qualitative, and mixed methods integration; and how to split into multiple papers. While we are limited to one chapter, we refer you to chapters offering more details on specific topics in this book and in other publications.

Deep consideration must be given to what ensures quality for mixed methods reporting. The standard of introduction, methods, results, and discussion grew from the development of reporting guidelines ([Altman, 2015](#)) but does not address the complexities of integrating quantitative and qualitative approaches. [Levitt et al. \(2018\)](#) found that “the thoughtful and robust use of mixed methods requires meeting the standards of both quantitative and qualitative research methodology in the design, implementation, and reporting stages” (p. 26). An example of the many typologies and designs supported for MMR is shown in [Table 2](#). Due to the number of possible purposes for using mixed methods, it is not possible in this space to address each one individually. In this chapter we share a basic approach to writing up an MMR study. Generating persuasive rationales, processes, and outcomes that use high quality criteria for mixed methods reporting is key ([Fàbregues and Molina-Azorín, 2017](#); [Poth, 2018](#)). [Wharton’s \(2017\)](#) study showed the most difficult part of research is dissemination. [Leech et al. \(2011\)](#) suggested having a framework to guide the writing process would “help to increase the rigor of mixed methods studies” and publications (p. 11). For more information on frameworks, see Plano Clark, Ivankova, & Yang’s chapter in this volume. Several others have suggested criteria for reporting mixed methods from the seminal typology of [Onwuegbuzie and Collins \(2007\)](#), the framework of Good Reporting of a Mixed Methods Study (GRAMMS, [O’Cathain et al., 2008](#)), recent guidelines for publishing in the JMMR ([Fetters and Freshwater, 2015](#)), the Mixed Methods Appraisal Tool (MMAT; [Hong et al., 2018](#)), and the American Psychological Association Journal Article Reporting Standards for MMR ([Levitt et al., 2018](#)). In addition, [Plano Clark and Ivankova \(2016\)](#) developed a socio-ecological framework to address the complexity and intertwining of domains to guide the entire research process including dissemination. Whichever you choose MMR reporting “must be defensible and transparent” ([Poth, 2018](#), p. 239). A well-written report will be one that another researcher can gather enough detail to reproduce the design, analysis, mixing, and reporting of the integration for each phase ([Reischer and Cowan, 2020](#)).

The context of mixed methods reporting

MMR uses the process of integrating quantitative and qualitative research to address a study’s purpose or research questions more completely ([Plano Clark and Ivankova, 2016](#)). To be considered MMR, a study must include (a) at least one quantitative and one qualitative method that are combined, which are (b) used rigorously and in accordance with the accepted criteria of the methodology and method used, where (c) the combination of the methods is carried out at minimum through a mixed methods design, and (d) the quantitative and qualitative methods are integrated in the study ([Creswell and Plano Clark, 2018](#)). [Hong et al. \(2018\)](#) caution that some research being reported should not be considered MMR including (a) those using a quantitative method that includes the collection and analysis of qualitative data but does not refer to the qualitative approach or method, (b) research using a qualitative method that includes the collection and analysis of quantitative data but does not refer to the quantitative approach or method, (c) a combination of multiple quantitative approaches only, (d) a combination of qualitative approaches only, (e) or the juxtaposition of qualitative and quantitative approaches without the integration of the two. Examples might include a quantitative survey with open-ended questions ([Creamer, 2018](#)), a qualitative interview that collects quantitative demographic data, or two

Table 2 Common examples of mixed method typologies and designs.

<i>Typology</i>	<i>Explanation</i>
Convergent parallel ^a	Quan and qual performed separate; brought together at interpretation
Explanatory sequential ^a	Quan first then qual used to explain quan results
Exploratory sequential ^a	Qual first then quan used to test or generalize
Embedded ^a	A strand of other type is added to enhance the overall design
Transformative ^a	Critical theory, shapes interaction, priority, timing, and mixing
Multiphase ^a	More than two phases
Parallel mixed ^b	One has two or more parallel quantitative and qualitative strands
Sequential mixed ^b	Strands occur across chronological phases
Conversion mixed ^b	Mixing occurs when one type of data is transformed to the other type and then analyzed
Multilevel mixed ^b	Mixing occurs across multiple levels of analysis
Fully integrated mixed ^b	Mixing occurs in an interactive manner at all stages of the study
QUAL + QUAN ^c	Equal-status concurrent design
QUAL + quan ^c	Qualitatively driven concurrent design
QUAN + qual ^c	Quantitatively driven concurrent design
QUAL → QUAN ^c	Equal-status sequential design
QUAN → QUAL ^c	Equal-status sequential design
QUAL → quan ^c	Qualitatively driven sequential design
Qual → QUAN ^c	Quantitatively driven sequential design
QUAN → qual ^c	Quantitatively driven sequential design
Quan → QUAL ^c	Qualitatively driven sequential design

^aCreswell and Plano Clark (2018).^bTashakkori and Teddlie (2009).^cJohnson and Christensen (2019).

separate studies conducted by the same researchers that were never intended to be integrated. Despite decades of setting mixed methods standards, these wrongly labeled studies can still be found in the literature and professional conferences.

A step-by-step approach to reporting mixed methods research

Framing your paper

When framing your paper consider that “the credibility of a mixed methods study rests almost wholly on the way it is reported” (Creamer, 2018, p. 280). Levitt (2020) advised that you need to clearly show the reader “from the start that both qualitative and quantitative methods are in use” (p. 112). Moreover, the use of mixed methods should appear in each section of your paper from the abstract to the implications. Further, the points of integration need to be explained to identify where the two are mixed or connected in some way (Schoonenboom and Johnson, 2017). Thus, you should provide a clear description the goals for the study’s quantitative approach, qualitative approach, and why there is a need for integration. Most of the standards and frameworks related to reporting MMR (e.g., APA-JARS Mixed, GRAMMS, MMAT, etc.) suggest whether strands are written about separately or interwoven, there should be continuity throughout the report. The approach should give the reader confidence in the integrity of your study and ability to replicate your research. Johnson and Christensen (2019) share that you need sufficient quality to achieve *multiple validation legitimation*, in other words meeting the relevant individual and combined quantitative, qualitative, and mixed methods validity needed. For more information, see Delmas and Giles’ chapter on qualitative approaches and Drisko’s chapter on quality reporting for qualitative research in this volume.

It may benefit the writer and ultimately the reader to capture your approach in a figure to visualize the order in which you will write (Plano Clark and Ivankova, 2016). Consistency throughout your writing is the key to translating your research to the reader. Typically, the order will be based on the sequence of use in sequential designs or in integrated or complex designs the importance of the approach or the purpose for mixing methods (Creswell and Creswell, 2018). Importance is usually determined by dominance, priority, weighting, emphasis, or status (Walker and Baxter, 2019). For instance, if one were conducting a mixed methods sequential exploratory designed study (Creswell and Plano Clark, 2018) where interviews inform the creation of a survey which is then implemented, the report would begin with the qualitative descriptions in each of the sections. If a study is run concurrently, but the

importance is on the quantitative findings, then the quantitative descriptions would lead each section of the report. Unpacking this, your study may use one of the commonly used designs shown in [Table 2](#) (i.e., [Creswell and Plano Clark, 2018](#); [Johnson and Christensen, 2019](#); [Tashakkori and Teddlie, 2009](#)).

For simplicity's sake, we leave it to the reader to determine the order of presenting their study. For the rest of this chapter, we will share how your chosen approach can be communicated consistently in each section of your paper.

Titling your paper

Imagine yourself in a brick-and-mortar bookstore or library. Slowly you walk down the aisles hunting the next book you will read. You pull out a few books, take a look at the covers, maybe you open to the table of contents or skim the first chapter. One appeals to you, so you drop it into your basket and continue your walk down the aisle. How did you choose what to read? What was the reason you decided to pull the book off the shelf? Usually, there is something about the title that appeals to you. You probably looked at the cover and either decided to buy it or put it back on the shelf. In most cases though, it was the title that first piqued your interest.

A title is often the most underappreciated component of the research reporting process. However, titles are the initial part of a written document that captures a potential reader's attention, whereby the reader determines whether a text is worth reading ([Zhi-jun and Jinfen, 2019](#); [Hallock and Bennett, 2021](#)). To prevent a paper from being dismissed, an article must have a proper title ([Soler, 2007](#)). A recent study showed that titles are a crucial decision-making element of grant applications, journal articles, reports, and documents, but they are often overlooked and almost never taught in the academic setting ([Anderson and August, 2020](#)). Titles assist readers to know the subject, methodology, and sometimes the context of your research. Often journals will have a word or character count for the title with a 15–20-word average. Titles should “indicate the topic and scope of the study,” and “be self-explanatory to readers” ([Swales and Leeder, 2012](#), p. 379) in “the fewest possible words” ([Day, 1989](#), p. 15). When crafting your title, you should include the methodology or design in the title. For example, “Urban high schools addressing the dropout crisis: A sequential explanatory mixed methods study.” This simple, 12-word title illuminates the topic and design indicating that the quantitative data shows school dropout rates are high, that the study focused on urban schools, and qualitative data was collected to explain the phenomenon. The title may not illuminate the complexity of the research, or the number of methods used, but it does give the reader a basic understanding of what to expect. It is also necessary to include an author's note on the title page acknowledging any funding sources, contributors, conflicts of interest, or special circumstances.

Abstract

The abstract is your reader's first glimpse of your study, and you should keep them in mind at the forefront of your writing ([Creswell and Plano Clark, 2018](#); [Leech and Onwuegbuzie, 2011](#)). A helpful metaphor put forth by [Leech \(2012\)](#) is to think about drafting your report as you would a cooking recipe with all the necessary ingredients and the process of the actions taken by the researcher, clearly delineated. This applies to the abstract as well. In your abstract, you should succinctly (a) state the problem, questions, and hypothesis or phenomenon of the study, (b) describe the participants and other data with pertinent characteristics, (c) indicate the mixed methods design, including how intersecting approaches to inquiry are combined, (d) the analytic approach, (e) the main findings, and (f) implications and significance. When the viewer finishes reading your abstract, they should be able to have a basic understanding of your study.

Keywords

In addition to the title, [Rostami et al. \(2014\)](#) found having different words in the keywords than in the title was a predictor of the number of citations an article received. When considering your keywords, two are fundamental in MMR reporting. One keyword should be included that describes the type of mixed methods design. Another keyword should be included that describes the problem addressed. The remaining keywords should be relevant to the reader understanding the intent and context of the study. Editor-in-chief, [Villar \(2020\)](#) advises that authors should make sure they have suitable keywords so the paper will be among the first to appear when a future author wishes to make a citation. Keywords are also one of the many ways search engines provide accurate results. Therefore, choosing keywords carefully are not only necessary to make your work more discoverable, but are also valuable to journals as this helps increase readership, citations, and their Impact Factor ([SAGE Publishing, n.d.](#)).

Voice

When disseminating MMR, it's beneficial to engage the reader in your text, while maintaining a scholarly presence. When writing up mixed methods, many researchers are unsure what approach they should take regarding the author's voice. Typically, qualitative writing is written in first person and quantitative in neutral or objective third person language. The question becomes should you write your paper in one voice or separate the paper between a qualitative first person voice and quantitative third person voice. [Zhou and Hall \(2018\)](#) share that traditionally entire mixed methods reports have used third person only and argue for more inclusion of the first person. The standard use of the third person writing style effectively writes out the presence of the researcher and renders them invisible ([Lazard and McAvoy, 2020](#)). [Zhou and Hall \(2018\)](#) go on to explain that the mixing of first person and third person voices can be beneficial so that

work is not tied one-dimensionally in its descriptions, in its dissemination of results, but is open to new ways of perceiving the world ... [incorporating] ideas of multiple voices ... conjoining voices ... to juxtapose different perspectives and demonstrate the possibilities of mixing voices in a single text (p. 354).

No standard has been set on whether to use first person only, third person only, or a blending of the two. [McCurdy \(2021\)](#) exemplifies the dual voices in her transformative mixed methods dissertation exploring science identities as she explains first the quantitative approach in the more formal third person voice, then uses personal pronouns as she describes the qualitative approach.

Logistic regression was conducted to determine whether students' SIS and SRS predict their decisions to pursue science professions after college ... Based upon the transformative mixed-methods research design, I integrated the quantitative and qualitative components of the research findings for research question (pp. 95–100).

[McCurdy \(2021\)](#) continues this pattern throughout the methods and reporting sections. When deciding which voice to include in your writing, we recommend that you consider which voice will engage your reader in your text and create the desire to continue reading.

Introduction

The introduction section plays a key role in engaging the reader in the background of your research. It serves as a roadmap for your research study, one that a reader can easily follow and understand. The introduction should show the importance of your topic by reviewing, critiquing, and synthesizing the applicable literature to identify key issues; theoretical frameworks; clarify barriers, knowledge gaps, or practical needs; and frame the problem in its context. In particular, the introduction needs to convey the barriers in the literature that suggest a need for both qualitative and quantitative data ([Levitt et al., 2018](#)). It “foreshadows the need for a mixed methods approach by outlining the literature illustrating the complexities” ([Wu et al., 2019](#), p. 910). Fundamentally, after reading the introduction, the reader should be able to understand the topic, why it is important, and why you adopted a particular method or approach. For more information, see Onghena's chapter on research synthesis for literature reviews in this volume.

As you write your mixed methods paper, it may be helpful to visualize an hourglass that is broad at the top, narrows slightly, and then broadens at the bottom, representing the broad introduction, narrower methods and results, and broad discussion (See [Fig. 1](#)). With this in mind, the paper should then begin with the more general aspects of the topic, then narrow toward the more specific information needed to place it in context and flow fluidly to the purpose and rationale for the study including the “need and value for the use of a mixed methods research” approach ([Fetters, 2020](#) p. 258). You should concisely cover what's already known about your topic and then show the gap of what is unknown, explaining notable concepts needed for the reader to understand your rationale. Keep in mind that the overarching goal of your introduction is to make your readers want to read your paper. All of your writing should tie back into this introduction section. Therefore, we propose writing your mixed methods paper much like an artisan weaves yarn in and out to create a beautiful tapestry. Just like the tapestry integrates a variety of colors, so should the mixed methods researcher weave the colorful strands of qualitative, quantitative, and integrated knowledge (see [Fig. 2](#)).

Problem statement, research questions, and rationale for mixing

MMR studies raise a level of complexity and potential incongruencies that will need to be carefully navigated regarding paradigms and procedures (see Philosophical underpinnings of MMR, Shannon-Baker, in this volume). As in all research, but in mixed methods particularly, the research question is the foundation and focus of all methodological decisions and determines the combination or integration of approaches ([Ponce and Pagán-Maldonado, 2015](#)). Thus, the goal of the research questions is to build on and mutually complement the findings of each other. Buchholz (2019) explained that the

components should provide insights into a superordinate (integrated) research question, which should therefore be formulated in such a way that the components can also focus on corresponding sub-questions and provide insights that can be related to the overarching question, that is, they can be integrated (p. 134).

Thus, a mixed methods study should be driven to answer the research questions.

Typically, a mixed methods study will address three strands of questioning—a quantitative question, a qualitative question, and an overarching question that can only be answered by weaving the strands together to integrate results. Perhaps a researcher wanted to study the use of a teacher professional development series for promoting certain instructional skills. The overall research question may be “*How did the professional development intervention affect the competencies of the teachers?*”. This overarching question may be approached in a mixed methods manner using both quantitative and qualitative inquiry to identify and understand teacher competencies. Quantitatively, the researcher may want to ask to what extent did teachers implement the intervention with fidelity, was

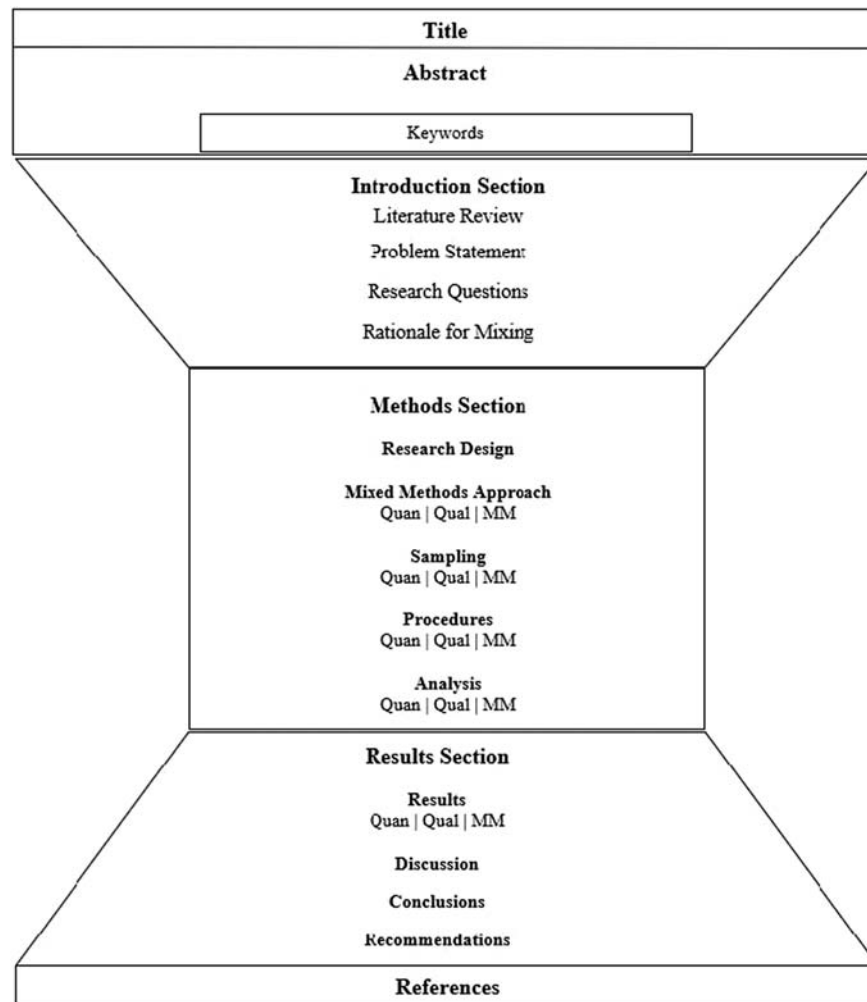


Fig. 1 The basic structure of a mixed methods report.

there a significant difference between student outcomes for teachers who received the professional development compared to the teachers who did not? Qualitatively, the researcher may want to know what did teachers perceive as barriers and facilitators to



Fig. 2 Woven components of a mixed methods report.

implementation? By using mixed methods, the researcher is in a unique position to use the results to complement each other and thereby understand a broader spectrum of the professional development's impact.

As described in section **Framing your paper**, the strands should be ordered to reflect the type of mixed methods design used. You should provide a thorough description as to how the approaches to inquiry were combined regarding the objectives and mixed methods rationale and how the results were obtained from using the mixed methods design type where mixing or integration occurred (Levitt et al., 2018).

Methods

The methods section is your opportunity to share your step-by-step plan for implementation from methodological considerations to sampling and participant selection. Ideally, this section “provides sufficient information to guide other researchers to replicate the study, assess the outcomes, and compare the findings with other studies” (Jalongo and Saracho, 2016, p. 141) by the readers’ basic questions of who, what, where, when, how, and why. You should explain and defend the methodologies and methods in your data collection and analysis processes. What were the reasons for choosing each approach and why is integration necessary? A good rationale for your methods typically is as heavily cited as the knowledge base in the literature review section. Whether you choose a separate or interwoven approach, the methods section should address the integrity of the qualitative methods, quantitative methods, and integrated approach. For more information, see Leech and Donovan’s chapter on MMR sampling and data collection in this volume.

Research design overview

The research design overview provides a road map for the study methods, which leads to a sharper understanding of the data collection approaches and helps the reader to interpret the results more accurately (Ghasemi et al., 2019). The problem your research is addressing as well as your research questions will drive the design of your mixed methods approach. Your report should be laid out according to the mixed methods design selected. While Table 2 listed different typographies and approaches, Table 3 gives simple examples of the order in which you might report your research design based on Hollstein’s (2014) categorization of four overarching groups. Many mixed methods studies are much more complex and require multiple rounds of data collection, analysis, and integration requiring a more advanced approach to reporting. While authors have reported there are as many as 40 types of mixed methods designs, we focus on four simplified categories from which you may build a (a) sequential exploratory, (b) sequential explanatory, (c) conversion, or (d) fully integrated design. Hollstein (2014) explains these designs as.

A sequential exploratory design starts with a qualitative phase, which is then followed by the quantitative phase ... A sequential explanatory design, in contrast, starts with the collection and analysis of quantitative data, which is then followed by a qualitative strand ... Conversion designs are designs that involve the transformation of data of one type into data of the other type for purposes of analysis ... The fully integrated mixed design is a specific kind of multistrand parallel design ... that most closely meshes and integrates qualitative with quantitative approaches (pp. 10–13).

Creswell and Plano Clark (2018) share that if the objective is to explore a problem posed by quantitative data followed with qualitative data, the design will require a sequential exploratory quan → qual design. If this is your design, the structure of the report and sections that follow should start with quantitative information, followed by qualitative information, and finish with the mixed integration. If the objective is to explain a phenomenon posed by qualitative data with quantitative data, the design will require a sequential explanatory qual → quan design (Creswell and Plano Clark, 2018). If this is your design, the sequence of your reporting should be the qualitative information first, followed by quantitative, ending with the mixed integration. If you are reporting convergent, concurrent, parallel, embedded, nested, or other complex designs, the importance of the quantitative

Table 3 Examples of methods order based on MM research design.

Design	Sequential exploratory	Sequential explanatory	Conversion	Fully integrated
Methods section	Research design Quan: - Participants and sampling - Procedures - Analysis Qual: - Participants and sampling - Procedures - Analysis Mixed integration	Research design Qual: - Participants and sampling - Procedures - Analysis Quan: - Participants and sampling - Procedures - Analysis Mixed integration	Research design Can determine the order-based priority and weighting. Mixed integration	Research design Mixed integration - Participants and sampling - Procedures - Analysis

approach versus the qualitative approach should guide your decision for the sequence of reporting with the most important approach leading each section. For more information, see Tajima and Fetter's chapter on MMR designs, Merten's chapter on transformative designs, Creamer's chapter on integration, and Maxwell's chapter on designing integrative MMR in this volume.

Notation

Much as music has a specialized notation system for musicians, a notation system was developed to convey the MMR process. In this system, the four letters *quan* indicates the quantitative approaches, while *qual* indicates the qualitative approaches (see Creswell and Plano Clark, 2018). Upper and lower case (e.g., QUAN vs. quan) indicate priority and importance with upper case indicating priority and lower case indicating supplementary. Symbols indicate the flow of the research process with an arrow sign (e.g., →) indicating the sequence of the quan and qual, a plus sign (e.g., +) indicating quan and qual were conducted concurrently, parentheses (e.g., ()) indicating either quan or qual was embedded, and brackets (e.g., []) indicating there were a series of steps or studies within studies. Finally, an equal sign (e.g., =) indicates the purpose of the study. Adhering to this system assists the reader to understand your research purpose and process.

Conceptual models

As an author, you should be aware of the potential difficulty readers may have following the path of complex designs. Many readers find visuals helpful for understanding the processes, contexts, and procedures for the study and can help explain how things were woven together. Ghasemi et al. (2019) found the use of appropriate tables and figures helped authors to summarize large amounts of complex information of the study procedures. It is common practice in mixed methods to include a figure in a publication that maps the steps taken and conceptualizes the process of data collection and analysis (Creamer, 2018). A conceptual model can assist with framing and showing the reader a blueprint of the overall research design (Baran, 2020), highlight the point or points of integration (Schoonenboom and Johnson, 2017), and share the context in which a study is conducted (Plano Clark and Ivankova, 2016). For example, Fig. 3 shows Kattan's (2022) socio-ecological framework used to guide his complex mixed methods study on teacher evaluation in Saudi Arabia.

In complex designs, Nair (2020) advises the inclusion of innovative figures to show the timing, interventional frequency, and follow up of each step in your study. Furthermore, a side-by-side implementation table of qualitative and quantitative sources could include type and sources of data, location and timing of data collection, and from whom. For more information see McCrudden's chapter on visual displays in this volume.

Integrating mixing

Mixed methods studies have several layers of complexity (Poth, 2018). The complexity of a study not only depends on the number of components, but also on the extent to which they depend on each other (Schoonenboom and Johnson, 2017). As mentioned previously, the integration or weaving of both qualitative and quantitative should be carefully navigated. Therefore, for this section of the methods, you need to share your plan for integration. If you collected data from multiple sources, you need to clearly identify the sources of qualitative and quantitative data (e.g., participants, text), their characteristics, and the relationship between the data-sets. For the reader to understand the mixing, provide information on sequencing, weighting, and be clear on whether the data were collected simultaneously or sequentially.

Sampling

Qualitative sampling

There are five areas critical for ensuring the integrity and trustworthiness of the qualitative approach (a) researcher description, (b) participants, (c) researcher relationship with participants, (d) participant recruitment and selection, and (e) any other details specific to the qualitative approach chosen. You should include a description of your researcher position including your experiences, worldviews, potential bias regarding the phenomena and participants being studied, along with your plan for how you managed these influences. Referred to as reflexivity, this allows your research claims and conclusions drawn to be understood and evaluated in context (Lazard and McAvoy, 2020). Reflexivity is a philosophically driven methodology of self-reflection and introspection and is not a one-and-done activity; rather, it should be embedded in the entire research process from inception to dissemination to mitigate any impacts of bias on the study.

Regarding participants, you need to fully describe the number and characteristics of participants and other data sources paying special attention to any perspective influences on the data collected. The researcher-participant relationship, if any, needs to be disclosed. Additionally, you should include a detailed description of all recruitment processes and any protocols that you created for data collection. The reader should understand the type of sampling used, whether participants were provided any incentives or compensation, how the number of participants was determined, what ethical considerations were made, and how consent was attained. For more information, see Koro et al.'s chapter on qualitative sampling in this volume.

Quantitative sampling

There are four topics to report for quantitative approaches (a) sampling procedures and participant selection, (b) sample size, power, and precision, (c) measures and covariates, and (d) other details specific to the quantitative approach chosen. Participant recruitment should include the inclusion and exclusion criteria and any restrictions based on demographics, culture, or contextual considerations. Further, the procedures used to select participants should be described (e.g., systematic sampling, self-selection,

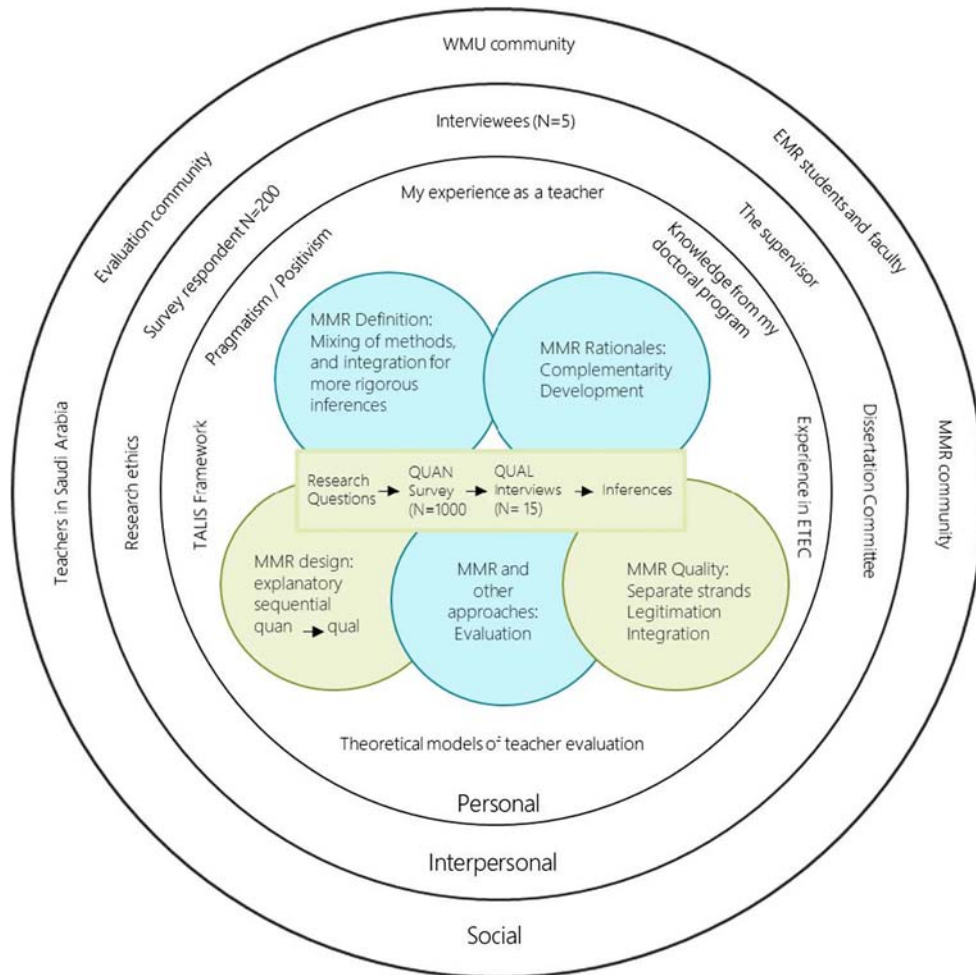


Fig. 3 Kattan's (2022) Socio-ecological framework of a mixed methods study of teacher evaluations in Saudi Arabia. Used with author permission.

purposive, etc.), along with the setting and locations for data collection, any agreements or incentives, and how you maintained and monitored ethical and safety measures. For sample size, power, and precision, include your intended sample size, how the sample size was determined, and how and if the sample size was achieved including the theoretical inferences for population-level power analysis or localized purposive sampling (Reischer and Cowan, 2020). All the primary and secondary measures and covariates collected should be reported including those not used. Finally, you need to provide the details specific to the quantitative approaches chosen (e.g., matching criteria, unit of assignment, unit of randomization, etc.).

Data collection

Next, full details of your data collection methods need to be provided. Leech and Onwuegbuzie (2011) counsel authors to provide detailed information about all data collection instruments and protocols used for both the qualitative and quantitative phases including

the developer of the instruments (with appropriate citations); the format of the instruments; when, how, and why they were administered; the context and focus of data collection; the duration of data collection; and information about the quality of the data collected (e.g., score reliability, score validity, interrater reliability) (p. 64).

For this section of the report, you ought to share your plan for data collection with enough detail for reproducibility (Reischer and Cowan, 2020). Moreover, as studies evolve, you may make changes that might confuse the reader. Reischer and Cowan (2020) advise researchers to "report details of methodology and decisions made throughout the research process; do not assume shared understanding with readers" (p. 4). A researcher's transparency increases the validity, credibility, and trustworthiness of the study.

Qualitative procedures

When reporting on qualitative data collection, there are five areas to address the (a) form of data collected, (b) origins and evolution of the protocols, (c) extensiveness of engagement, (d) how data were recorded and transformed, and (e) how the researcher managed their reflexivity. Qualitative procedures tend to be more flexible and subjective, so it is essential to explain the approach taken and the choices made. Qualitative data may be collected in a variety of ways including observations, interviews or focus groups, dairies, or journals, and even art forms, each requiring a different type of protocol (Creswell and Plano Clark, 2018). These could be adopted or adapted from a pre-existing source or created specifically for the study. It is imperative that any changes made during the data collection process are documented. In addition, the extensiveness of engagement, how data were recorded, and how data were transformed play a major role in understanding the integrity of the research results.

As transparency of your researcher position is important, details also need to be provided with how you bracketed your position throughout the data collection and data analysis process. Bryman (2016) found that researchers need to conduct a critical examination of their beliefs and assumptions, including the way those can potentially affect interactions and relationships with the people and content under study. If a researcher waits until the data analysis stage, there is the potential impact on the integrity of data collected. Shufutinsky (2020) suggested a researcher needs to self-bracket presuppositions and beliefs throughout the entire research process, so it doesn't open the potential for "previous experience, knowledge, and assumptions to take root and create presupposition in the early stages of the research, including in the design of the study" (p. 53).

Quantitative procedures

When reporting quantitative data collection, the areas an author needs to address vary depending on the quantitative approach. Seven basic procedures to consider are (a) the methods used to collect the data, (b) quality of the measurements, (c) instrumentation, (d) masking, (e) psychometrics, (f) conditions and design, and (g) data diagnostics. For valid and generalizable results, a description of your methods should provide enough detail for another researcher to replicate your study. Essential is an in-depth explanation of how concepts were operationalized, variables were measured, and detailed information on the instruments, tools, procedures, and materials used for data collection. Ghasemi et al. (2019) explained that the level of detail needed will depend on the type of quantitative procedures and the audience. They share that the author should include what the reader needs for understanding how the key findings were derived.

Mixed procedures

If integration of qualitative and quantitative data occurred during the data collection process, the author needs to provide a description of the integrated procedures. In most instances, researchers don't approach integration until the analytic phase of the study (Levitt et al., 2018). When the qualitative and quantitative components co-occur, authors may use their discretion in presenting the procedures, but are encouraged to do so in a way that presents a logical progression of the report narrative as well as including an audit trail to show the steps to data collection, analysis, and how it informs the overall study purpose (Merriam and Tisdell, 2016). It may be helpful to your reader to refer them back to the notation and conceptual model sections explained previously.

Analysis

The data analysis section can be one of the most difficult and one of the most rewarding experiences of MMR. The data analysis should link back to the study purpose, objective, and aims (Fetters, 2020). In the analysis section, your focus should be to assist the reader to understand your process for analyzing qualitative data, quantitative data, and the integration of the two. As you explain the analysis, help the reader to see how you wove the strands together, pulling new colors into your tapestry; remind your reader how they refer to the colors you integrated from the introduction section. Visual displays such as tables and figures are helpful to document both the process of and product of the analytical procedures. The structure of the data analysis will determine the structure for the reporting. If you took a separative approach and conducted the qualitative and quantitative data collection prior to integrating the analysis, the reporting will be separate as well. If you took an iterative and interactive approach thereby allowing the data to interact during data collection, then reporting of the data analysis will be integrated. Since most of the published standards highlight the independent analysis of qualitative and quantitative data followed by mixing, and to provide continuity to the process we have described so far, we continue the approach for reporting from the perspective of qualitative, quantitative, and then mixed. This segregated approach lends itself well to the joint displays recommended by many in the field. For more information, see Guetterman's chapter on joint displays in this volume.

Qualitative analysis

When writing up the procedures for qualitative analysis, you should share your data analytic strategies with enough detail to be fully comprehended by the reader. Levitt et al. (2018) emphasize that authors need to provide rationales to illuminate analytic choices in relation to the study goals. As mentioned in **Analysis**, this is an opportunity for you to weave aspects of the introduction section back into your paper as you discuss the qualitative analysis. A detailed description should be provided of the process of your analysis using the principles of transparency, including a discussion of the procedures (e.g., coding, thematic analysis). If not described elsewhere, this section should include coder selection, training, and collaboration. Next, include information on whether coding was conducted in an inductive or deductive manner, (e.g., coding categories emerged during the analyses or developed in advance). In addition, the units of analysis should be included (e.g., entire transcript, unit, text) and how you formed those units. If an analytic

scheme was used, include an illustration along with how and when the analytic scheme was determined. Finally, indicate what tools were used during the data analysis process, including any software.

Traditional for reporting qualitative data analytic procedures is the need to include a section on trustworthiness (see Cain et al. in this volume). With this in mind, it is critical that you share how you maintained methodological integrity and consistency. It should be clear to the reader that the claims you made from the analysis are warranted, grounded in the evidence, contextual, insightful, and meaningful. Share how you supplemented by checks added to data analysis to support the claims you made. Moreover, a description should be included on how you bracketed and managed your biases and perspectives. The reflexivity process is predicated on the broad questions of what the research process is and how you are influencing it from inception, through data collection, into analysis, mean-making, and dissemination (Lazard and McAvoy, 2020). When you acknowledge your personal bias and how you bracketed for them, it establishes a level of confidence in the reader and enhances the trustworthiness of your study (Daniel, 2018). Finally, you need to provide a description of how the study remained consistent or share any alterations you made including any changes due to ethical issues. For more information, see Miha's chapter on approaches to qualitative analysis in this volume.

Quantitative analysis

When reporting on the analysis of quantitative data, the areas you need to address vary depending on the quantitative approach. Fetter's (2020) suggested seven overarching steps to data analysis regardless of the approach (a) enter, clean, and address gaps or deficiencies, (b) frame the analysis in accordance with the study's purpose, (c) conduct a preliminary descriptive analysis, (d) use an organizational structure to report the findings, (e) check for inconsistencies, (f) organize findings for dissemination, and (g) interpret findings and write up results. The APA JARS-Quant includes 12 tables with details for the variety of quantitative approaches and recommends reporting the analytic strategy with sub-areas in each of the 12 tables.

Mixed analysis

Mixed methods data analysis strategies are "at the heart of methodological discussions about mixed methods" (Vogl, 2019, p. 551). In mixed methods, value accrues from both qualitative findings and quantitative results, and the integration of the two in a thoughtful way leads to greater mining of the data and enhanced insights (Levitt et al., 2018). Integration is the process of connecting, combining, and merging the qualitative and quantitative data for generating inferences and the mixed meta inferences of the conclusions drawn after the integration (Creswell and Plano Clark, 2018). Though there are instances where you may focus in on the specific strands from the colors of qualitative or quantitative, this is the opportunity to step back and look how these are being integrated to create your tapestry (see Fig. 4). For more information, see Vogl's Chapter on approaches to MMR analysis and Kuckartz de's chapter on using analysis software in this volume.

Integrating data offers the possibilities of answers that would not be accomplished by qualitative or quantitative alone, nor if these approaches were merely overlaid but not integrated (Oliveira et al., 2018). This underscores the importance for connecting the analysis back with the research question and goals and aligning this section with the need for justifications for the use of mixed method research studies.

Findings and results

Your results section should be concise but thorough and follow the order of the design referenced throughout the paper (Fetter's and Freshwater, 2015). You should begin with a sentence or two overview about the study and discuss only those findings that relate to the research questions and the purpose of the study. Here you should emphasize the major findings in a balanced progression, report both negative and positive findings, organize the data in tabular or illustration format, and provide associations, variances,

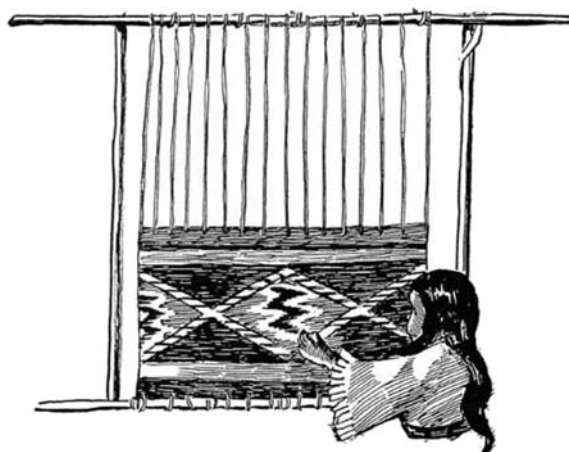


Fig. 4 Native American weaving an intricate pattern into a tapestry.

and magnitude of the findings with adequate interpretation and transparency (Meo and Eldawlatly, 2019). “Failure to adequately, transparently, and rigorously report findings has been described as a form of scientific failure” with “unacknowledged biases, inability to replicate due to missing information, or lack of clarity around context” that could potentially lead to preventing evidence to be put into practice (Wharton, 2017, p. 487). Worse yet, studies show that these missing components could be “dangerously misleading to policy decisions” (Kac and Hirst, 2011, p. 1872).

Qualitative findings

What you include for the qualitative findings will depend greatly on the approach taken in your study. Each qualitative methodological approach has its own style of writing, or storytelling (Schinke et al., 2020). You should present research findings in a way that is compatible with the study design and demonstrate the credibility or trustworthiness of your study. Daniel (2018) explained that trustworthiness requires establishing that findings are credible, relevant, and congruent thereby giving the reader confidence that “the findings reflect participants views outside of researchers’ personal opinions” (p. 265). Often, synthesizing illustrations (e.g., diagrams, tables, models) are useful in organizing and conveying findings. You may decide your study requires separate illustrations of your qualitative findings or you may choose a side-by-side visual to show the qualitative, quantitative, and mixing.

Quantitative findings

Much of what should be included in the quantitative findings will depend inherently on the quantitative approach taken in your study (e.g., the statistical output, missing data, model estimation, variance and covariance, etc.). Annesley (2010) shared that a valuable concept for you to remember is that data and results differ. Data are usually presented in tables and figures and results are statements in the main text that summarize or explain what the data show. In general, though, many researchers take the opportunity to show their quantitative findings with visual displays. These can assist readers to grasp, comprehend, and remember information (Jalongo and Saracho, 2016). The most effective visuals are self-explanatory allowing the reader to see trends, relationships, outcomes, categories, or general experimental parameters (Annesley, 2010). You should also display data in an order that makes logical sense to the viewer (Evergreen and Emery, 2016). However, close attention should be made to ensure the data type and visualization are compatible. Müller et al. (2012) warn that badly chosen visualizations can compromise the truth of academic reporting.

Integrated mixed methods findings

By the end of your report, the data analyzed from the qualitative and quantitative approaches should be integrated in such a way as to provide a fuller understanding of the phenomenon under study. Creswell and Tashakkori (2007) seminally shared that “integration might be in the form of comparing, contrasting, building on, or embedding one type of conclusion with the other” (p. 109). Leech et al. (2011) emphasize the importance of the context stating it’s essential for the content “to be delineated for the reader so that it is clear how the quantitative and qualitative findings relate and the extent to which meta-inferences can be made” (p. 21). Much like our tapestry, Fetters and Freshwater (2015) share that the results “may be structured and described using a weaving pattern where the quantitative and qualitative findings presented together thematically with a weaving back and forth of quantitative and qualitative data by theme or domain” (p. 212).

In their editorial on publishing methodological MMR, Fetters and Freshwater (2015) shared three ways for demonstrating integration of illustrative quantitative and qualitative results: narrative, data transformation, or joint display presentation. A narrative format is simply a text-based account of your findings and may be shown separately or integrated. In data transformation, qualitative data are transformed into quantitative data or the opposite with quantitative data transformed into qualitative. A joint display allows the quantitative and qualitative data to be observed in a side-by-side display in such a way that helps to make mixed interpretations that would not otherwise be known if the researcher conducted a qualitative-only or quantitative-only study (see example in Table 4). You should avoid discussing or interpreting the results until later in the discussion and conclusion section.

Discussion and interpretation

The discussion piece is one of the most valuable components of your paper, it provides clarifications on synthesis of your findings, any issues, and includes an interpretation of the results and recommendations for practice. A good discussion section leaves the reader with an understanding of “why the researcher took the time to do such a complicated study and can help to pull together findings, which may otherwise appear contradictory or unrelated” (Bronstein and Kovacs, 2013, p. 359). Your interpretations should go “beyond first-level results or literal description of observed behaviors, patterns, and themes from analysis” (Johnson et al., 2020). Further, in discussing your mixed methods study, your primary challenge is to provide a complete and thorough exploration of how specific results relate to each other, contribute to answering the research questions, and achieve the primary purpose of the research endeavor. If you have presented your study with Figs. 2–4 in mind, the discussion section should relate the findings with that to the introduction. It should take the strands of qualitative and quantitative data to weave together a beautiful tapestry of research questions, study design, data collection and analysis, and interpretation of results. While readers may come to the same conclusions as you after reading the findings, a beautifully woven discussion can assist the reader to visualize and more fully understand the delicate interlacing of the data. This is especially crucial when the findings are complicated or contradictory (Bronstein and Kovacs, 2013). If done well, you will have reported out in such a way that is both understandable to the reader and has been woven

Table 4 Example of a joint display for a mixed methods study on high school dropout.

Theme	Quan results	Qual results	Mixed interpretation
Dropout mindset	98% of students who dropped out from XYZ school in 2021 tested at below proficiency in at least two academic areas in 4th grade	Students shared they struggled academically even in elementary school. They never felt they could “catch up” “so what’s the point?”	Dropout mindset begins early
Dropout absenteeism	98% of students who dropped out from XYZ school in 2021 were absent more than 10 days	Students shared many reasons for being absent from “I feel like others think I am just lazy but I am not” and “my dad works 3rd and I couldn’t leave the house I’d miss the bus and didn’t have a ride to school” to “my mom needed me to stay home with my baby sister so she could work” to “I had to get a job to help pay the bills”	Dropouts feel they are perceived as lazy or stupid, but often they aren’t in school because of other more pressing needs or obligations.

together like a beautiful tapestry. The final sections of your paper should contain an ending that gives the reader the utmost satisfaction and that their time reading your manuscript was well spent (Meo and Eldawlaty, 2019).

Responding to the reporting limitations for mixed methods

A pervasive issue with reporting MMR is the limited space given to authors in many dissemination paths. For example, many academic journals enforce strict word limits, which often are insufficient for the descriptions required in mixed methods studies (Brown et al., 2015). Bazeley (2015) found that journals “often have 3000–5000-word limits on the length of articles” (p. 301). She goes on to recommend the following strategies to deal with this problem (a) publish a methodological article first to describe the approach and data sources, then draw on this for reference in later articles, (b) provide an overview paper first and then a *wrap up* paper drawing on the mixed methods findings at the end, (c) focus separate papers on specific aspects of the study bringing together all relevant sources for this issue, or (d) seek out journals that allow for longer articles. Another approach that Brown et al. (2015) share is to take advantage of the option for including supplementary materials online to include more details regarding methods and rigor. Creswell (2015) recommended generating three separate papers from the mixed methods study: quantitative, qualitative, and *overall* or adding a methodological paper discussing unique mixed methods procedures making sure to observe the sequence in publishing separate papers and providing cross-references from one paper to another to establish consistency and reduce the length. Careful self-examination of your design, purpose, and context is essential before deciding your approach.

Final thoughts for reporting mixed methods

With a trend of increasing use of MMR by education researchers to investigate complex issues within the field (Crawford and Tan, 2019), there is a need for integrity in reporting MMR. While there is debate in the field as to the specific structure or framework to use when reporting mixed methods studies, we have outlined an approach to assist the writer to include the critical components of mixed methods reporting. Researchers, such as Creamer (2018) have suggested full integration with the qualitative, quantitative, and mixed methods woven together throughout reporting. There is a concern in the mixed methods field that the pressure to conform to a specified set of designs has led to too little attention on its products (Guest, 2013), where most “designs show the qualitative and quantitative phases as quite self-contained” (Creamer, 2018, p. 297). There is a small but growing number of researchers calling for more integrated methods throughout the research process. Whether you create reports where each approach is highlighted separately or interwoven throughout, the items outlined in this chapter are imperative to ensure trustworthiness and reproducibility of your study. As mixed methods researchers in the field of education, the focus needs to be on the dissemination of our research in ways that readers can fully understand and act, especially regarding wicked problems. With such attention to detail, we have the potential to improve our research processes, educational environments, and ultimately outcomes.

References

- Altman, D.G., 2015. Making research articles fit for purpose: structured reporting of key methods and findings. *Trials* 16. <https://doi.org/10.1186/s13063-015-0575-7>.
- American Psychological Association, seventh ed., 2019. *Publication Manual of the American Psychological Association*, Washington, DC.
- Anderson, O., August, E., 2020. A disconnect between classroom and professional writing: we are not teaching public health students to write effective titles. *J. Publ. Health* 42, 412–415. <https://doi.org/10.1093/pubmed/fdaa034>.
- Annesley, T.M., 2010. Put your best figure forward: line graphs and scattergrams. *Clin. Chem.* 56, 1229–1233.

- Baran, M.L., 2020. Mixed methods research design. In: Baran, M.L., Jones, J. (Eds.), *Applied Social Science Approaches to Mixed Methods Research*. IGI Global, pp. 26–52. <http://doi:10.4018/978-1-7998-1025-4.ch002>.
- Bazeley, P., 2015. Writing up multimethod and mixed methods research for diverse audiences. In: Hesse-Biber, S.N., Johnson, R.B. (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. Oxford University Press, Oxford, UK, pp. 296–313. <https://doi.org/10.1093/oxfordhb/9780199933624.001.0001>.
- Bronstein, L.R., Kovacs, P.J., 2013. Writing a mixed methods report in social work research. *Res. Soc. Work. Pract.* 23, 354–360. <https://doi.org/10.1177/1049731512471564>.
- Brown, K.M., Elliott, S.J., Leatherdale, S.T., Robertson-Wilson, J., 2015. Searching for rigour in the reporting of mixed methods population health research: a methodological review. *Health Educ. Res.* 30, 811–839. <https://doi.org/10.1093/her/cyv046>.
- Bryman, A., 2016. *Social Research Methods*, fifth ed. Oxford University Press, Oxford, UK.
- Childs, J., Lofton, R., 2021. Masking attendance: how education policy distracts from the wicked problem(s) of chronic absenteeism. *Educ. Pol.* 35, 213–234. <https://doi.org/10.1177/0895904820986771>.
- Craddock Lee, S.J., 2021. Mixed methods: capturing complexity in family medicine research. *Ann. Fam. Med.* 19, 98–99. <https://doi.org/10.1370/afm.2682>.
- Crawford, R., Tan, H., 2019. Responding to evidence-based practice: an examination of mixed methods research in teacher education. *Aust. Educ. Res.* 46, 775–797.
- Creamer, E., 2018. *An Introduction to Fully Integrated Mixed Methods Research*. SAGE Publications, Inc., Thousand Oaks, CA.
- Creswell, J.W., Creswell, J.D., 2018. Mixed methods procedures. In: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, fifth ed. SAGE Publications, Inc., Thousand Oaks, CA, pp. 213–246.
- Creswell, J.W., Plano Clark, V.L., 2018. *Designing and Conducting Mixed Methods Research*, third ed. SAGE Publications, Inc., Thousand Oaks, CA.
- Creswell, J.W., Tashakkori, A., 2007. Editorial: developing publishable mixed methods manuscripts. *J. Mix. Methods Res.* 1, 107–111. <https://doi.org/10.1177/1558689806298644>.
- Daniel, B.K., 2018. Empirical verification of the “TACT” framework for teaching rigour in qualitative research methodology. *Qual. Res. J.* 18, 262–275. <https://doi.org/10.1108/QRJ-D-17-00012>.
- Day, R.A., 1989. *How to Write and Publish a Scientific Paper*, seventh ed. Cambridge University Press, Cambridge, UK.
- Evergreen, S., Emery, A.K., 2016. The Data Visualization Checklist. https://stephanieevergreen.com/wp-content/uploads/2016/10/DataVizChecklist_May2016.pdf.
- Fábregues, S., Molina-Azorín, J.F., 2017. Addressing quality in mixed methods research: a review and recommendations for a future agenda. *Qual. Quant.* 51, 2847–2863. <https://doi.org/10.1007/s1135-016-0449-4>.
- Fetters, M.D., Freshwater, D., 2015. Publishing a methodological mixed methods research article. *J. Mix. Methods Res.* 9, 203–209. <https://doi.org/10.1177/1558689815594687>.
- Fetters, M.D., 2020. *The Mixed Methods Research Workbook: Activities for Designing, Implementing, and Publishing Projects*. SAGE Publications, Inc., Thousand Oaks, CA.
- Fredricks, J.A., Hofkens, T., Wang, M.T., Mortenson, E., Scott, P., 2018. Supporting girls’ and boys’ engagement in math and science learning: a mixed methods study. *J. Res. Sci. Teach.* 55, 271–298. <https://doi.org/10.1002/tea.21419>.
- Garwood, J.D., Werts, M.G., Varghese, C., Gosey, L., 2018. Mixed-methods analysis of rural special educators’ role stressors, behavior management, and burnout. *Rural Spec. Educ. Q.* 37, 30–43. <https://doi.org/10.1177/8756870517745270>.
- Ghasemi, A., Bahadoran, Z., Zadeh-Vakili, A., Montazeri, S.A., Hosseiniapanah, F., 2019. The principles of biomedical scientific writing: materials and methods. *Int. J. Endocrinol. Metab.* 17, e88155. <https://doi.org/10.5812/ijem.88155>.
- Guest, G., 2013. Describing mixed methods research: an alternative to typologies. *J. Mix. Methods Res.* 7, 141–151.
- Hallock, R.M., Bennett, T.N., 2021. I’ll read that!: what title elements attract readers to an article? *Teach. Psychol.* 48, 26–31. <https://doi.org/10.1177/0098628320959948>.
- Hong, Q.N., Fábregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.C., 2018. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. *Educ. Inf.* 34, 285–291. <https://doi.org/10.3233/EFI-180221>.
- Hollstein, B., 2014. Mixed methods social networks research: an introduction. In: Dominguez, S., Hollstein, B. (Eds.), *Mixed Methods Social Networks Research: Design and Applications*. Cambridge University Press, Cambridge, UK, pp. 3–25.
- Hunt-Anderson, I.E., 2020. The cloak and dagger: a mixed study of covert bullying in Jamaican high schools. *Caribb. J. Mixed Methods Res.* 54.
- Jalongo, M.R., Saracho, O.N., 2016. *Writing for Publication*. Springer International Publishing Switzerland, Birkhäuser Verlag AG.
- Johnson, R.B., Christensen, L., 2019. *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. SAGE Publications, Thousand Oaks, CA.
- Johnson, J.L., Adkins, D., Chauvin, S., 2020. A review of the quality indicators of rigor in qualitative research. *Am. J. Pharm. Educ.* 84, 138–146. <https://doi.org/10.5688/ajpe7120>.
- Kac, G., Hirst, A., 2011. Enhanced quality and transparency of health research reporting can lead to improvements in public health policy decision-making: help from the EQUATOR Network. *Cad. Saude Publica* 27, 1872–1873.
- Kattan, S., 2022. *Teacher Evaluation in Saudi Arabia Relative to National and International Teacher Evaluation Standards and Best Practices*. Ph.D. Dissertation. Western Michigan University. <https://scholarworks.wmich.edu/dissertations/>.
- Lazard, L., McAvoy, J., 2020. Doing reflexivity in psychological research: what’s the point? What’s the practice? *Qual. Res. Psychol.* 17, 159–177. <https://doi.org/10.1080/14780887.2017.1400144>.
- Leech, N.L., 2012. Writing mixed research reports. *Am. Behav. Sci.* 56, 866–881.
- Leech, N.L., Onwuegbuzie, A.J., 2011. Guidelines for conducting and reporting mixed research in the field of counseling and beyond. *J. Counsel. Dev.* 88, 61–69. <https://doi.org/10.1002/j.1556-6678.2010.tb00151.x>.
- Leech, N.L., Onwuegbuzie, A.J., Combs, J.P., 2011. Writing publishable mixed research articles: guidelines for emerging scholars in the health sciences and beyond. *Int. J. Mult. Res. Approaches* 5, 7–24. <https://doi.org/10.5172/mra.2011.5.1.7>.
- Levin, K., Cashore, B., Bernstein, S., Auld, G., 2012. Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change. *Pol. Sci.* 45, 123–152.
- Levitt, H.M., 2020. Reporting mixed methods research: bridging reporting standards. In: Levitt, H.M. (Ed.), *Reporting Qualitative Research in Psychology: How to Meet APA Style Journal Article Reporting Standards*. American Psychological Association, pp. 111–133. <https://doi.org/10.1037/0000179-009>.
- Levitt, H.M., Bamberg, M., Creswell, J.W., Frost, D.M., Josselson, R., Suárez-Orozco, C., 2018. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: the APA publications and communications board task force report. *Am. Psychol.* 73, 26–46. <https://doi.org/10.1037/amp0000151>.
- McCurdy, R., 2021. *Male Students of Color in STEM through the Lens of Intersectionality: A Transformative Mixed-Methods Exploration of Their Science Identities, Relevant Science Learning Experiences, and Decisions to Pursue Science Professions*. Ph.D. Dissertation. University of Central Florida. <https://stars.library.ucf.edu/etd2020/732/>.
- Meo, S.A., Eldawlaty, A.A., 2019. Pathophysiology of a scientific paper. *Saudi J. Anaesth.* 13 (5), 9–11. https://doi.org/10.4103/sja.SJA_48_19.
- Merriam, S.B., Tisdell, E.J., 2016. *Qualitative Research: A Guide to Design and Implementation*, third ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Mertens, D.M., 2015. Mixed methods and wicked problems. *J. Mix. Methods Res.* 9, 3–6. <https://doi.org/10.1177/1558689814562944>.
- Müller, H., van Biljon, J.A., Renaud, K.V., July 2012. Information visualization in research reporting: guidelines for representing quantitative data. In: Southern African Computer Lecturers’ Association, Thaba Nchu, pp. 13–19. In: <http://www.sacla.org.za/wp-content/uploads/2020/03/Proceedings-SACLA-2012.pdf#page=18>.
- Nair, B., 2020. Pitfalls in article submissions for publication. *Indian Dermatol. Online J.* 11, 937–943. https://doi.org/10.4103/idoj.IDOJ_658_20.
- O’Cathain, A., Murphy, E., Nicholl, J., 2008. The quality of mixed methods studies in health services research. *J. Health Serv. Res. Pol.* 13, 92–98. <https://doi.org/10.1258/jhsrp.2007.007074>.
- Oliveira, J.L.C., Magalhães, A.M.M., Matsuda, L.M., 2018. Mixed methods in nursing research: application possibilities according to Creswell. *Texto Contexto Enferm.* 27, 1–8. <https://doi.org/10.1590/15174522-018004221>.

- Onwuegbuzie, A.J., Collins, K.M., 2007. A typology of mixed methods sampling designs in social science research. *Qual. Rep.* 14, 281–316. <https://doi.org/10.46743/2160-3715/2007.1638>.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research: A Guide to the Field*. SAGE Publications, Inc., Thousand Oaks, CA. <https://www.doi.org/10.4135/9781483398341>.
- Poth, C.N., 2018. *Innovation in Mixed Methods Research: A Practical Guide to Integrative Thinking With Complexity*. SAGE Publications, Inc., Thousand Oaks, CA.
- Ponce, O.A., Pagán-Maldonado, N., 2015. Mixed methods research in education: capturing the complexity of the profession. *Int. J. Educ. Excel.* 1, 111–135. <https://doi.org/10.18562/ijee.2015.0005>.
- Reischer, H.N., Cowan, H.R., 2020. Quantity over quality? Reproducible psychological science from a mixed methods perspective. *Collabra* 6, 26. <https://doi.org/10.1525/collabra.284>.
- Rittel, H.W.J., Webber, M.M., 1973. Dilemmas in a general theory of planning. *Pol. Sci.* 4, 155–169. <https://doi.org/10.1007/BF01405730>.
- Rostami, F., Mohammadpoorasl, A., Hajizadeh, M., 2014. The effect of characteristics of title on citation rates of articles. *Scientometrics* 98, 2007–2010. <https://doi.org/10.1007/s11192-013-1118-1>.
- Rouse, H.L., Hurt, T.R., Melby, J.N., Bartel, M., McCurdy, B., McKnight, E., Zhao, F., Behrer, C., Weems, C.F., 2020. Pregnancy and parenting among youth transitioning from foster care: a mixed methods study. *Child Youth Care Forum* 50, 167–197. <https://doi.org/10.1007/s10566-020-09567-0>.
- Schinke, R.J., Mellalieu, S., Ntoumanis, N., Kavussanu, M., Standage, M., Strauss, B., Papaioannou, A., 2020. Getting published: suggestions and strategies from editors of sport and exercise psychology journals. *J. Appl. Sport Psychol.* 32, 1–14. <https://doi.org/10.1080/10413200.2020.1741725>.
- SAGE Publishing, (n.d.). Help Readers Find Your Article. Retrieved June 4th, 2021, from <https://us.sagepub.com/en-us/nam/help-readers-find-your-article>.
- Schoonenboom, J., Johnson, R.B., 2017. How to construct a mixed methods research design. *Kolner Z. Soz. Sozpsychol.* 69, 107–131. <https://doi.org/10.1007/s11577-017-0454-1>.
- Shufutinsky, A., 2020. Employing use of self for transparency, rigor, trustworthiness, and credibility in qualitative organizational research methods. *Organ. Dev. Rev.* 52, 50–58. <https://doi.org/10.1111/1467-8578>.
- Soler, V., 2007. Writing titles in science: an exploratory study. *Engl. Specif. Purp.* 26, 90–102. <https://doi.org/10.1016/j.esp.2006.08.001>.
- Swales, J.M., Leeder, C., 2012. A reception study of the articles published in English for specific purposes from 1990–1999. *Engl. Specif. Purp.* 31, 137–146. <https://doi.org/10.1016/j.esp.2011.07.003>.
- Tashakkori, A., Teddlie, C., 2009. Integrating qualitative and quantitative approaches to research. In: *The SAGE Handbook of Applied Social Research Methods*, vol. 2, pp. 283–317.
- Theunissen, M.J., Verdonk, P., Feron, F., Bosma, H., 2012. The early identification of risk factors on the pathway to school dropout in the SIODO study: a sequential mixed-methods study. *BMC Publ. Health* 12, 1–8. <https://doi.org/10.1186/1471-2458-12-1033>.
- Villar, R., 2020. How to write that paper. *J. Hip Preserv. Surg.* 7, 1–3. <https://doi.org/10.1093/jhps/hnaa010>.
- Vogl, S., 2019. Integrating and consolidating data in mixed methods data analysis: examples from focus group data with children. *J. Mix. Methods Res.* 13, 536–554. <https://doi.org/10.1177/1558689818796364>.
- Walker, C., Baxter, J., 2019. Method sequence and dominance in mixed methods research: a case study of the social acceptance of wind energy literature. *Int. J. Qual. Methods* 18. <https://doi.org/10.1177/1609406919834379>.
- Wharton, T., 2017. Rigor, transparency, and reporting social science research: why guidelines don't have to kill your story. *Res. Soc. Work. Pract.* 27, 487–493. <https://doi.org/10.1177/1049731515622264>.
- Wu, Y.P., Deatrick, J.A., McQuaid, E.L., Thompson, D., 2019. A primer on mixed methods for pediatric researchers. *J. Pediatr. Psychol.* 44, 905–913. <https://doi.org/10.1093/jpepsy/jsz052>.
- Zhijun, L.I., Jinfen, X.U., 2019. The evolution of research article titles: the case of journal of pragmatics 1978–2018. *Scientometrics* 121, 1619–1634. <https://doi.org/10.1007/s11192-019-03244-3>.
- Zhou, X., Hall, J.N., 2018. Mixed methods papers in first-person and third-person: writing voices in dialogue. *J. Mix. Methods Res.* 12, 344–357. <https://doi.org/10.1177/1558689816652755>.

Joint displays in mixed methods research

Timothy C. Guetterman, University of Michigan, Ann Arbor, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

Approaches to integration	755
Integrative analysis of qualitative and quantitative data or results to generate new insights	755
Sequential integration in which qualitative or quantitative results inform the methods of the other strand	755
Data transformation to convert one type of data to another	756
Hybrid approaches	756
Overarching guidance for using joint displays	756
Common intents for using joint displays	757
Comparing qualitative and quantitative data or results for a more complete understanding through joint displays	757
Examining patterns of qualitative and quantitative results through joint displays	757
Connecting to identify samples through joint displays	758
Building data collection instruments through joint displays	758
Explaining variation in outcomes through joint displays	759
Summary of different intents of joint displays	760
Summary	760
References	761

Joint displays provide a visual method of facilitating the integration of qualitative and quantitative research (Guetterman, Fetters et al., 2015b). Joint displays have emerged as one strategy to report integrated results in mixed methods research. As presented in publications, theses, and dissertations, and other reports, joint displays typically consist of a table or a figure that presents qualitative and quantitative data, methods, or results. Recently, researchers have stated to use innovative joint displays that include graphs, charts, images, and other figures (Guetterman et al., 2021). Ideally, joint displays depict the outcome of integration, such as the development of meta-inferences, the generation of follow-up studies, or the building of an instrument. However, joint displays are increasingly being used as a method to integrate qualitative and quantitative research, rather than merely as a results presentation (Fetters and Guetterman, 2021).

Numerous types of joint displays have been identified in the literature. The most common type is a side-by-side display that juxtaposes qualitative and quantitative results along with meta-inferences from the integration of the two. Another type is a statistics-by-themes joint display that arrays themes as rows and statistical results as columns. While articles and chapters have cataloged different types or styles of joint displays (Guetterman, Fetters et al., 2015b; Plano Clark and Sanders, 2015; Younas et al., 2020) and presented their usage by design (Guetterman, Creswell et al., 2015a), little has been written about selecting a joint display based on a research question, purpose, or aim. For clarity, the term *intent* will be used to generally refer to research questions, purposes, or aims. The intent might be thought of as what is meant to be achieved through the integration of qualitative and quantitative research. Although intents and designs are highly interrelated, linking joint displays to their intent overcomes the fact that researchers use different design typologies. Uncoupling joint displays from specific research designs can considerably broaden their usage and unlock potential.

Mixed methods research is conducted in educational disciplines, such as teacher education, educational psychology, and cognition and learning. Moreover, scholars conducting educational research in various professions are increasingly using mixed methods approaches. For example, mixed methods research has been applied to engineering education (Crede and Borrego, 2013; Howell Smith et al., 2019), health professions education (McLaughlin et al., 2016; Schifferdecker and Reed, 2009), and accounting education (Lamprecht and Guetterman, 2019) to name a few. Research on education in the professions and other fields often layers on additional complexity and makes the need to integrate qualitative and quantitative research even more compelling. The methods of integration become even more salient. Critiques of mixed methods research in many of these fields have identified weaknesses with integration specifically (Institute of Education Sciences Technical Working Group, 2015; Plano Clark, 2019). Failing to achieve meaningful integration is so important because it limits researchers achieving the full potential of mixed methods research. Joint displays provide on potential solution for achieving integration. Next, I will examine some common strategies for integration in mixed methods research before discussing how the strategies can be applied through joint displays. However, relatively little literature has addressed the use of joint displays for reporting integration in educational mixed methods research. Even less literature has discussed the use of joint displays for mixed methods analysis in educational research.

To summarize, joint displays are emerging as a method of integrative analysis in addition to a way to report results, yet relatively little literature has linked joint displays to their intent, and few studies have examined joint displays from the perspective of educational research, which often intersects multiple interdisciplinary fields. Therefore, the purpose of this chapter is to discuss and advocate for the use of joint displays as an integration method in mixed methods educational research. This chapter presents templates for different types of joint displays as a starting point and cites examples of joint from educational research whenever possible. To achieve that purpose, this chapter will situate different types of joint displays based on the intent for integrating qualitative and

quantitative research. Before discussing joint displays, it may be helpful to review several major approaches to integration in mixed methods research.

Approaches to integration

Integration is the defining feature of mixed methods research that distinguishes it from monomethod and multimethod approaches, by most definitions. Integration refers to the interface of qualitative and quantitative research or where the two come together (Creswell and Plano Clark, 2018). Integration leads to something new: (1) new insights called meta-inferences that researchers identify through integrative analysis, (2) new follow-up phases of research that are systematically informed by a strand of research (e.g., results from a quantitative phase inform a subsequent qualitative phase), or (3) new data collection instruments systematically developed by a strand of research (e.g., results from a qualitative exploratory phase inform quantitative instrument development).

Integration can occur at multiple levels of research, such as integrating multiple philosophical worldviews (Johnson, 2017), intersecting more than one research design (Plano Clark and Ivankova, 2016), and thinking from a mixed methods perspective to ensure integration at all aspects in fully integrated mixed methods research (Creamer, 2018). For the purpose of this chapter and discussing joint displays, I will focus on methods of integration and discuss several of these methods below.

Mixed methods researchers have discussed and employed numerous methods for integration. These methods fall into several broad categories: (1) integrative analysis of qualitative and quantitative data or results to generate new insights, (2) sequential integration in which qualitative or quantitative results inform the methods of the other strand, (3) data transformation to convert one type of data to another for further analysis, or (4) hybrid approaches such as applying both qualitative and quantitative analysis methods to the same dataset and integrating results. Researchers often use multiple methods of integration in a single mixed methods study. For additional reading on integration, Schoonenboom presents different strategies for integration (see 11,045. Ten mixed methods integration strategies for obtaining a detailed understanding), Vogl reviews analysis of mixed methods data (11,048. Strategies to integrative mixed methods analysis), and Kuckartz discusses the use of software for mixed methods analysis (11,049. Mixed methods analysis software).

Integrative analysis of qualitative and quantitative data or results to generate new insights

One of the prevalent methods of integrative analysis is the analysis of qualitative and quantitative data or results together to generate meta-inferences. One type of this integrative analysis is merging to compare qualitative and quantitative results in similar domains or concepts, to assess whether they are convergent, divergent, or expand understanding and to identify new meta-inferences from the integrative analysis (Fetters et al., 2013). Pluye and Hong (2010, 11046. Convergence and divergence in mixed methods research) discuss ways to handle convergence and divergence in mixed methods research.

One specific type of this process is *following a thread* that consist of conducting separate analysis of each strand's data, identifying an analytic domain (i.e., the thread) based on the literature or research question to examine in each dataset, developing a multi-faceted understanding based on findings in each dataset, and iteratively conducting analysis and synthesizing threads. For example, White and colleagues examined the roles of racial identity, science identity, science self-efficacy beliefs, and science achievement among African American students attending historically Black colleges and universities (White et al., 2019). They used a convergent mixed methods design in which qualitative and quantitative data were analyzed separately and then merged for a more complete understanding. Guided by Critical Race Theory, a science identity model, and self-efficacy theory, they merged findings to examine whether the qualitative results corroborated the quantitative results, and when they were corroborated, used the qualitative data to expand on results. For example they explained the significant relationship between science identity and self-efficacy variables "as an interaction between recognition and mastery experiences (White et al., 2019, p. 67).

Another method of integration is to interrelate qualitative and quantitative results, such as examining statistical patterns by theme or examining thematic patterns by levels of a variable. For example, a study examining differential effects of an educational intervention to improve communication examined thematic differences for the high, middle, and low performing learners (Guetterman et al., 2019). A key meta-inference from integrative analysis was that for the group with the worst quantitative outcomes at post-test, their qualitative data indicated less motivation to learn more and less interest in learning about communication.

Sequential integration in which qualitative or quantitative results inform the methods of the other strand

Connecting is a method of integration through which the results of one strand of research inform the sampling of the other (Fetters et al., 2013). Connecting is a sequential type of integration because the results of one phase are needed to inform the next. Qualitative results could be used to help plan quantitative probability sampling such as identifying what characteristics would be important in a stratified sample. More common is beginning quantitatively and using the results to identify a purposeful sample for the qualitative follow-up. For example, researchers might use quantitative results to find individuals or sites that meet characteristics (e.g., differential performance levels of students, teachers, schools, or other systems), to find outliers, or to identify participants based on intervention outcomes. McCrudden and McTigue (2019) used quantitative results about high school students beliefs about climate change to identify participants to invite for interviews.

In *building* integration methods, the results of one strand of research inform the data collection of the other strand. Building is also a sequential type of integration. In a design that begins qualitatively, the results could be used to systematically build a quantitatively instrument, to adapt an instrument ensuring contextual relevance, or to identify variables for a follow-up quantitative phase. Alternatively, when a mixed methods design begins with a quantitative phase, those results could be used systematically to build an interview guide or focus group guide. For example, researchers studied teachers' integration of cultural content and heritage language into their classrooms and Latino students math achievement using an explanatory sequential design (Matthews and López, 2019). They wrote questions the follow up qualitative phase based on the quantitative survey items to further expand on the quantitative findings and learn how they implemented specific classroom practices.

Data transformation to convert one type of data to another

Data transformation is the concept of converting one type of data into another and integrating the transformed data. Quantitative data could be transformed to qualitative by adding a description or a narrative account to quantitative results like clusters, factors, or profiles (Bazeley, 2018). Transforming qualitative data into quantitative is more common and involves converting qualitative data or results into counts or binary variables (Bazeley, 2018). For example, a binary variable can denote whether a code, category, or theme was present. The now quantized data can be analyzed with simple or advanced statistics such as structural equation modeling (Newman and Constantinides, 2021) or other techniques. Hitchcock et al. (2021) provided an example of examining the implementation of a formative assessment system in kindergarten classrooms by a US state education agency. They integrated quantitative school data with quantized focus group data that identified whether schools had implementation problems and analyzed the data using hierarchical linear modeling to identify predictors of poor implementation (Hitchcock et al., 2021).

Hybrid approaches

Researchers have identified numerous hybrid approaches, also called inherently mixed methods approaches to analysis. Examples of these hybrids include conducting ethnographic analysis, case study analysis, and natural language processing of text data. An inherently integrative approach is ethnographic decision models that are used to understand decision-making processes. In one method, qualitative methods are used to gather decision criteria and then to develop a structured elicitation instrument to develop a preliminary model (Ryan and Bernard, 2021). The model is further developed and tested through mixed methods research. In another hybrid approach, case study typically involves collecting multiple forms of data, linking data sources, and analyzing data to better understand the case. When the data involve a combination of qualitative and quantitative and their integrative analyses, the case-oriented analysis is inherently a mixed methods analysis (Bazeley, 2018). Finally, natural language processing applied to a corpus of text can be an integrative analysis when augmented with human coding (Guetterman et al., 2018). In this method of integration, the largely quantitative analysis identifying phrases and assigning meaning is compared with traditional human coding of the text to develop a better understanding of text such as open-ended survey responses, text message survey responses, or written records (Chang et al., 2021).

In this brief review of integration methods, only select methods that are commonly applied were discussed. Learning other approaches to integration may be helpful in addressing research questions. For each of the approaches discussed, the development joint displays is a way to visually facilitate the integration of qualitative and quantitative research. The remainder of this chapter will focus of guidance for using joint displays generally and advice of selecting a joint display based on its intent.

Overarching guidance for using joint displays

Before discussing specific joint displays, this section provides broad guidance for their selection and use. First, I encourage you to think about the research question to be answered, the aim to be addressed, or the specific integration need to be conveyed through a joint display. In other words, what is the intent of the joint display? Later in the chapter, I will cover information about intents for a mixed methods joint display to help select one or more that may be most appropriate for your study. Consistency between the joint display (i.e., what type, what is included) and its intent is important for a reader.

Second, carefully consider what you want to communicate to the reader. The joint display can be used to display integrated results only or to walk the reader through integration steps, such as the development of meta-inferences or the systematic development of an instrument from qualitative findings. In addition, you might consider how to best communicate to the reader. For example, are numbers best displayed or would a chart or graph more easily communicate results (Guetterman et al. 2021)? Perhaps, grayscale or color coding related qualitative and quantitative results would assist with readability.

Third, developing a joint display is a process and creating the joint display itself might be your method of integration (Fetters and Guetterman, 2021). Termed joint display analysis, the process is a method to break down the relatively complex cognitive task of integration. Joint display analysis involves aligning related qualitative and quantitative concepts, considering what to include (e.g., text in cells of table or figures), and developing meta inference. Because the development often involves continual revision and refinement, the process can be incredibly helpful for conducting integrative analysis.

Fourth, it might be helpful to think broadly and innovatively about joint display development. Joint displays have been described as a data display or a way to show the integration of qualitative and quantitative data. I intentionally use a broader

definition of joint displays as a visual means to represent and facilitate the integration of qualitative and quantitative research. The rationale for this broad definition is because joint displays can be used to indicate how data are integrated, how results are integrated, or how results are integrated to inform methods of the other strand of research. The chapter provides examples of these different types of integration in the discussion of joint displays by intent below.

Common intents for using joint displays

Other articles have examined joint displays by mixed methods designs; this chapter takes a different approach and examines five intents for using a joint display and pair each intent with a template joint display. The reason for presenting joint displays by intent is because types of joint displays can often be used across multiple designs—they are not mutually exclusive. Here, I discuss using joint displays for (1) comparing qualitative and quantitative results, (2) relating qualitative and quantitative results, (3) connecting to identify samples, (4) building data collection instruments, and (5) explaining variation in outcomes. The list of intents is not exhaustive but represents common uses for joint displays. The following sections include a template as a structure for the joint display. For presentation purposes, the templates include only three rows or domains, but it is important to include as many are needed to address the intent. In addition, the cells of the first row or column (in italics) contain guidance for what to include.

Comparing qualitative and quantitative data or results for a more complete understanding through joint displays

Joint display can be used to facilitate the comparison of qualitative and quantitative results. This structure for a joint display is applicable to most any design, including common designs such as a convergent, explanatory sequential, and exploratory sequential in addition to more complex designs such as case study to conduct a case-based analysis (Creamer, 2021). Table 1 provides a template for a side-by-side joint display that juxtaposes qualitative with quantitative results along with meta-inferences developed through integrative analysis. The process of developing this joint display also constitutes integrative analysis and has occasionally been termed *joint display analysis* (Fetters and Guetterman, 2021). To facilitate a comparison of qualitative and quantitative results, the joint display is organized by domain. The domain could be a concept from a theoretical model, a qualitatively derived domain, or a quantitative construct or variable. The idea of this template is to link related qualitative and quantitative results and examine those results one domain at a time to develop meta-inferences. The template provides a framework for conducting that analysis. The cells of the joint display could contain a summary of related qualitative, quantitative, and mixed methods (i.e., meta-inferences) results. Sometimes, cells also include illustrative quotes, statistics, images, or graphs such as bar charts, column graphs, boxplots, or scatterplots (Guetterman et al., 2021). Through a side-by-side joint display, the researcher has a framework for integration, but also has a way to communicate the comparison of results and meta-inferences to the reader.

An excellent example of a side-by-side joint display is the study by Haynes Brown in which she displayed quantitative results next to classroom observation data. The design was an explanatory sequential case study that began with questionnaires and lead to follow-up qualitative observations to better explain results. By looking across, they attempted to “determine whether the case study findings for each teacher confirmed their initial case classification based on the self-report data from the questionnaire” (Haynes-Brown and Fetters, 2021, p. 11). Another example is Howell Smith et al.’s (2019) joint display that shows the iterative development of a grounded theory model of interest in engineering PhDs. The joint display is oriented vertically rather than side-by-side so the reader can easily compare refinement of the visual model through qualitative and quantitative phases of research.

Examining patterns of qualitative and quantitative results through joint displays

A joint display can also be used to examine patterns such interrelationships of the qualitative and quantitative results by examining themes-by-statistics or statistics-by-themes. This type of joint display looks like a crosstabulation or contingency table. One option to is to examine statistics for themes to determine whether any patterns are present. Table 2 provides a template for this joint display in which columns are themes and quantitative constructs or variables are rows. For example, researchers might select a theme, and look at a statistical profile for the individuals for whom that theme was present. This method works best when themes are not

Table 1 Side-by-side joint display template.

<i>Domain</i>	<i>Qualitative results</i>	<i>Quantitative results</i>	<i>Meta-inferences</i>
Domain 1	<i>Summary of themes/subthemes, perhaps illustrative quotes</i>	<i>Summary of results, perhaps statistical results such as coefficients, confidence intervals, p values, effect sizes</i>	<i>Summarize meta-inference, perhaps indicate convergent or divergence of qualitative and quantitative</i>
Domain 2			
Domain 3			

Table 2 Themes-by-statistics joint display template.

	<i>Theme 1 or type 1</i>	<i>Theme 2 or type 3</i>	<i>Theme 3 or type 3</i>	<i>Meta-inferences</i>
Construct 1 or variable 1	<i>Summary of theme 1 related to construct 1, perhaps illustrative quotes</i>	<i>Summary of theme 2 related to construct 1, perhaps illustrative quotes</i>	<i>Summary of theme 3 related to construct 1, perhaps illustrative quotes</i>	<i>Meta-inference 1: Summary of meta-inference across themes or types</i>
Construct 2 or variable 2				<i>Meta-inference 2</i>
Construct 3 or variable 3				<i>Meta-inference 3</i>
	<i>Meta-inference A: Summary of meta-inference across constructs or variables</i>	<i>Meta-inference B</i>	<i>Meta-inference C</i>	

universally present for all participants or sites. The final row and final column of this template are to include meta-inferences looking across themes or across variables. An extension of this approach is to qualitatively develop a typology of individuals and to examine statistical differences or similarities among the types. Figure could alternatively include a type for each column or be transposed with variables as columns and themes as rows.

As an example of this type of joint display, researchers examined uncertainty during critical incidents for individuals in medical residency training (Hamui-Sutton et al., 2015). They developed a qualitatively typology of five types of uncertainty (Technical, Communicational, Conceptual, Systemic, Ethical) and examined quantitative variables for each. Although they did not use the term, they essentially developed a joint display as a matrix with types of uncertainty as rows and as columns for the mean level of perceived uncertainty and mean level of perceived stress measured on a five-point scale.

Connecting to identify samples through joint displays

A joint display can be used to describe the systematic selection of samples based on the results of one phase of research. This structure can apply to any sequential type of mixed methods design in which one phase informs sampling decisions for the next phase. Table 3 presents a sampling joint display template that depicts the relationship between the results of one phase and the sample of the next (McCrudden et al., 2021). The joint display is organized by domain, which could alternatively be a concept from a theoretical model, a variable, or a key result. The next column is quantitative results, if phase came first, followed by a column to describe the qualitative sample. For instance, this column could be used to describe demographic characteristics to consider or the plan to sample those who improved in an educational intervention and those who did not. The final column might be used to indicate a rationale for the sampling decision—strengthening the tight connection while integrating from the quantitative phase to the qualitative phase. For representing how qualitative results informed a subsequent quantitative sample, having the qualitative column to the left and quantitative to the right will better reflect the flow of procedures in the design.

As an example of this joint display, McCrudden and McTigue (2019) included a joint display that indicated how they used quantitative data to sort individuals into two groups for subsequent purposeful sampling. The focus of the explanatory sequential mixed methods study was to examine bias in beliefs about climate change among high school students. They calculated a bias score to identify students with more-objective versus less-objective reasoning. The joint display sorted these two groups and provided their rating of arguments for different types of climate change from which they purposefully sampled participants for follow-up interviews to better understand the quantitative ratings (McCrudden and McTigue, 2019).

Building data collection instruments through joint displays

A joint display is also used to depict the systematic development of data collection instruments from a phase of research. For example, from quantitative results, a joint display could identify specific questions in a semi-structured interview to explain results. An innovation on this type of display developed by McCrudden and McTigue (2019) includes a rationale for the question as the

Table 3 Sampling joint display template.

<i>Domain</i>	<i>Quantitative results</i>	<i>Qualitative sample</i>	<i>Rationale for sample</i>
Domain 1	<i>Statistical summary for domain 1, perhaps coefficients, confidence interval, p values, or effect sizes</i>	<i>Description of sampling for domain 1, such as purposive sampling criteria, sample size</i>	<i>Reason for sampling decision</i>
Domain 2			
Domain 3			

final column. Alternatively, a joint display could depict how a survey was developed or adapted based on results from an initial qualitative exploratory phase.

Tables 4 and 5, respectively, provided templates for showing how quantitative results inform interview questions and how qualitative results can inform an instrument. **Table 4** provides a potential structure for an Interview Questions joint display to show how quantitative results informed qualitative interview questions. This joint display is also organized by domain, which could alternatively be a construct from a theoretical model, a variable, an aim, a research question, or another relevant field. The template begins with quantitative results, which come first temporally followed by a column for the relevant interview question. The quantitative results and interview questions do not have to be a one-to-one relationship. For example, a particular result might have no related questions or multiple follow-up interview questions. The point of this joint display is to systematically write interview questions (or focus group questions) based on quantitative results to help address research questions. As previously noted, the final column provides a rationale for the question, a recommendation by [McCrudden and McTigue \(2019\)](#). They provided an exemplar of this type of joint display from their study of beliefs about climate change in which they linked specific follow-up questions based on students' responses to a rating scale along with what they hoped to elicit through the interview question.

Table 5, on the other hand, presents a template for linking qualitative results to quantitative items being developed based on those results. Again, following the sequential order of the research procedures, this joint display has qualitative results in the first two columns followed by the quantitative items developed in the next column. The link between codes and items does not have to be one-to-one and should be based on the intent of the instrument developed. Researchers used this type of joint display to link constructs and subconstructs from a grounded theory model and quotes from participants to scale items to measure interest in an engineering PhD ([Howell Smith et al., 2019](#)). The joint display provides an example of including quotes that then inform item writing.

Table 6 provides a variation when the intent is to adapt an existing instrument or instruments. This type of joint display systematically indicates how qualitative results inform adaptations. For example, if the desire is to use mixed methods to adapt an instrument to ensure cultural or contextual relevance, this joint display provides a method to review existing items and show what items are revised, added, or removed. The final column includes a rationale for the specific adaptations made.

Explaining variation in outcomes through joint displays

A joint display can be used to explain variation in outcomes. The template in **Table 7** is similar to the statistics-by-themes joint display but transposed in this template. Developing this type of joint display facilitates the analysis of patterns for a theme across levels of a variable or patterns for levels of a variable across themes. The cells might include a summary of the result for a particular theme and level of the variable. The final column and row of the template includes a space for meta-inferences. This type of joint display facilitated analysis in a study examining differential effects of an intervention to improve communication skills using practice-oriented training with virtual humans ([Guetterman et al., 2019](#)). The joint display allowed us to examine questions such as, what did we learn about the theme of motivation across low, middle, and high performers in the intervention and what is a thematic profile of a high performer? The published joint display in [Guetterman et al. \(2019\)](#) did not include meta-inferences, but developing the joint display facilitated our analysis and identification of meta-inferences that are included in the prose instead.

Table 4 Interview questions joint display template.

Domain	Quantitative result	Qualitative interview question	Rationale for question
Domain 1	Statistical summary for domain 1, perhaps coefficients, confidence interval, p values, or effect sizes	List specific interview questions or focus group questions for domain 1	Reason for adding particular question—what will the question elicit to help explain quantitative results
Domain 2			
Domain 3			

Table 5 Instrument development joint display template.

Theme or category	Qualitative code and illustrative quote	Quantitative item
Theme 1/Category 1	Specific codes or subthemes for theme 1/category 1 and perhaps illustrative quotes	List specific quantitative items for theme 1/category 1
Theme 2/Category 2		
Theme 3/Category 3		

Table 6 Adapting an instrument joint display template.

<i>Initial item</i>	<i>Qualitative theme, code, or illustrative quote</i>	<i>Revised quantitative item</i>	<i>Reason for revision, addition, or deletion</i>
Item 1	<i>Summary of theme or code related to particular item and perhaps an illustrative quote too</i>	<i>List specific revised item or items</i>	<i>Reasons for revising the item</i>
Item 2			
Item 3			

Table 7 Explaining variation joint display template.

	<i>High</i>	<i>Middle</i>	<i>Low</i>	<i>Meta-inferences</i>
Theme 1	<i>Summary of theme 1 for high performers, perhaps illustrative quotes</i>	<i>Summary of theme 1 for middle performers, perhaps illustrative quotes</i>	<i>Summary of theme 1 for lower performers, perhaps illustrative quotes</i>	Meta-inferences: <i>Summary of meta-inference for theme 1 across performance levels</i>
Theme 2				Meta-inferences
Theme 3	Meta-inferences: <i>Summary of meta-inference for higher performers across themes</i>	Meta-inferences	Meta-inferences	Meta-inferences

Table 8 Matching integration intents to research questions to address through joint displays.

<i>Intent</i>	<i>Research question</i>
Comparing qualitative and quantitative results	• How do qualitative and quantitative results compare?
Relating qualitative and quantitative results	• What patterns emerge examining statistics by themes or themes-by-statistics?
Connecting to identify samples	• What sample will best help to explain or further elucidate quantitative results?
Building data collection instruments	• How should an instrument be developed or adapted based on qualitative results gathered to ensure contextual relevance?
	• What interview questions will help to explain or understand quantitative results?
Explaining variation in outcomes	• What do qualitative results explain about quantitative results?
	• What distinguishes higher versus lower performers?
	• How does it add to understanding

Summary of different intents of joint displays

Common intents of joint displays are comparing results, relating qualitative and quantitative results, connecting to identify samples, building data collection instruments, explaining variation in outcomes. This lists only represents five common intents and is not intended to be all-inclusive. While much has been focused on using a joint display as a presentation mechanism, researchers have found the process of developing a joint display to be useful as a mixed methods integration method (Fetters and Guetterman, 2021). I recommend thinking about the intent of the joint display—what do you hope to do and what do you want to communicate. Ideally, the joint display is matched to aims or research questions of the study. To think through options, Table 8 summarizes the common intents of joint displays and suggests potential mixed methods research questions that could be addressed.

Summary

This chapter reflects a contemporary trend in thinking about joint displays as more than just a way to present mixed methods results by advocating for the use of joint displays as an integration method in mixed methods educational research. Linking joint displays to specific intents further uncouples joint displays from design typologies. While I do not believe that a mixed methods study must

include a joint display, they provide one concrete method for integrating. Certainly, other methods of integration are valuable and valid.

Before closing, I will discuss several questions that I often receive about joint displays. First, can I include more than one joint display? Yes, many articles include more than one joint display, subject to space guidelines of journals. In addition, mixed methods dissertations often include more than one joint display. For example, a joint display might be included that links qualitative and quantitative data collection question in addition to a joint display that presents results side-by-side or relates statistics-by-themes. Moreover, researchers sometimes need to split side-by-side or statistics-by-themes joint display into multiple displays when there are many domains of interest. A second common question is: do I need to have qualitative and quantitative data in the joint display? In my opinion, both qualitative and quantitative research needs to be represented in the joint display. What is represented could be data, results, or methods (e.g., sampling decisions), but both qualitative and quantitative aspects need to be included. A third common question is: What if I do not have related quantitative (or qualitative) data to link to the other type? This issue is more difficult to overcome. However, it may be possible to present results side-by-side along with meta-inferences if the qualitative and quantitative samples are comparable. Regardless, the issue should be acknowledged as a limitation. Fourth and finally, a common question is: do meta-inferences need to be in the joint display? For some joint displays (e.g., connecting results to sampling decisions), the goal was not to identify meta-inferences, but to generate new samples or data collection instruments. When the goal is to develop meta-inferences, their inclusion is ideal, but could make the joint display too dense. If it is not possible to include meta-inferences in the joint display, it is imperative to include these insights in the text.

In conclusion, the use of joint displays has increased dramatically in the past 5–7 years. Recent joint display innovations include the use of figures, images, graphs, or other graphics (Bustamante, 2017; Guetterman et al., 2021). I recommend researchers use these templates and others only as a starting point to develop new and innovative joint displays. Future methodological scholarship should focus on additional innovations and the insight that can be gleaned from using joint displays as an integration and analysis method. Software can be extremely useful for integration and developing joint displays (Kuckartz and Rädiker, 2022). As new features are added, researchers will have more tools for creating innovative joint displays.

References

- Bazeley, P., 2018. Integrating Analyses in Mixed Methods Research. Sage.
- Bustamante, C., 2017. TPack and teachers of Spanish: development of a theory-based joint display in a mixed methods research case study. *J. Mix. Methods Res.* 13 (2), 163–178. <https://doi.org/10.1177/1558689817712119>.
- Chang, T., DeJonckheere, M., Vydiswaran, V.G.V., Li, J., Buis, L.R., Guetterman, T.C., 2021. Accelerating mixed methods research with natural language processing of big text data. *J. Mix. Methods Res.* 15 (3), 398–412. <https://doi.org/10.1177/15586898211021196>.
- Creamer, E.G., 2018. An Introduction to Fully Integrated Mixed Methods Research. Sage.
- Creamer, E.G., 2021. Leveraging an integrated visual display for case-based analysis in mixed method research. *Int. J. Qual. Methods* 20. <https://doi.org/10.1177/16094069211059000>, 16094069211059000.
- Crede, E., Borrego, M., 2013. From ethnography to items: a mixed methods approach to developing a survey to examine graduate engineering student retention. *J. Mix. Methods Res.* 7 (1), 62–80. <https://doi.org/10.1177/1558689812451792>.
- Creswell, J.W., Plano Clark, V.L., 2018. Designing and Conducting Mixed Methods Research, third ed. Sage.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48, 2134–2156. <https://doi.org/10.1111/1475-6773.12117>.
- Fetters, M.D., Guetterman, T.C., 2021. Development of a joint display as a mixed analysis. In: *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, London, pp. 259–275.
- Guetterman, T.C., Chang, T., DeJonckheere, M., Basu, T., Scruggs, E., Vydiswaran, V.V., 2018. Augmenting qualitative text analysis with natural language processing: methodological study. *J. Med. Internet Res.* 20 (6), e231. <https://doi.org/10.2196/jmir.9702>.
- Guetterman, T.C., Creswell, J.W., Kuckartz, U., 2015a. Using joint displays and MAXQDA software to represent the results of mixed methods research. In: McCrudden, M., Schraw, G., Buckendahl, C. (Eds.), *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*. Information Age Publishing, pp. 145–176.
- Guetterman, T.C., Fàbregues, S., Sakakibara, R., 2021. Visuals in joint displays to represent integration in mixed methods research: a methodological review. *Methods in Psychology* 5, 100080.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015b. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13 (6), 554–561. <https://doi.org/10.1370/afm.1865>.
- Guetterman, T.C., Sakakibara, R., Baireddy, S., Kron, F.W., Scerbo, M.W., Cleary, J.F., Fetters, M.D., 2019. Medical students' experiences and outcomes using a virtual human simulation to improve communication skills: mixed methods study. *J. Med. Internet Res.* 21 (11), e15459.
- Hamui-Sutton, A., Vives-Varela, T., Gutiérrez-Barreto, S., Leenen, I., Sánchez-Mendiola, M., 2015. A typology of uncertainty derived from an analysis of critical incidents in medical residents: a mixed methods study. *BMC Med. Educ.* 15 (1), 198. <https://doi.org/10.1186/s12909-015-0459-2>.
- Haynes-Brown, T.K., Fetters, M.D., 2021. Using joint display as an analytic process: an illustration using bar graphs joint displays from a mixed methods study of how beliefs shape secondary school teachers' use of technology. *Int. J. Qual. Methods* 20. <https://doi.org/10.1177/1609406921993286>, 1609406921993286.
- Hitchcock, J.H., Lambert, R., Holcomb, T.S., 2021. Hierarchical linear modeling with qualitative data that have been quantized. In: *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 99–108.
- Howell Smith, M.C., Babchuk, W.A., Stevens, J., Garrett, A.L., Wang, S.C., Guetterman, T.C., 2019. Modeling the use of mixed methods—grounded theory: developing scales for a new measurement model. *J. Mix. Methods Res.* <https://doi.org/10.1177/1558689819872599>. Advance online publication.
- Institute of Education Sciences Technical Working Group, 2015. Mixed Methods in Education Research. <https://ies.ed.gov/ncer/whatsnew/techworkinggroup/pdf/MixedMethodsTWG.pdf>.
- Johnson, R.B., 2017. Dialectical pluralism: a metaparadigm whose time has come. *J. Mix. Methods Res.* 11 (2), 156–173. <https://doi.org/10.1177/1558689815607692>.
- Kuckartz, U., Rädiker, S., 2022. Using Maxqda for integration in mixed methods research. In: Hitchcock, J.H., Onwuegbuzie, A.J. (Eds.), *The Routledge Handbook for Advancing Integration in Mixed Methods Research*. Routledge, pp. 540–562.
- Lamprecht, C., Guetterman, T.C., 2019. Mixed methods in accounting: a field based analysis. *Meditari Account. Res.* 27 (6), 921–938. <https://doi.org/10.1108/MEDAR-11-2018-0403>.

- Matthews, J.S., López, F., 2019. Speaking their language: the role of cultural content integration and heritage language for academic achievement among Latino children. *Contemp. Educ. Psychol.* 57, 72–86.
- McCrudden, M.T., Marchand, G., Schutz, P.A., 2021. Joint displays for mixed methods research in psychology. *Methods in Psychology* 5, 100067. <https://doi.org/10.1016/j.metip.2021.100067>.
- McCrudden, M.T., McTigue, E.M., 2019. Implementing integration in an explanatory sequential mixed methods study of belief bias about climate change with high school students. *J. Mix. Methods Res.* 13 (3), 381–400. <https://doi.org/10.1177/1558689818762576>.
- McLaughlin, J.E., Bush, A.A., Zeeman, J.M., 2016. Mixed methods: expanding research methodologies in pharmacy education. *Currents in Pharmacy Teaching and Learning* 8 (5), 715–721.
- Newman, D., Constantinides, S., 2021. Structural equation modeling with qualitative data that have been quantitized. In: *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 89–98.
- Plano Clark, V.L., 2019. Meaningful integration within mixed methods studies: identifying why, what, when, and how. *Mixed Methods Research in Educational Psychology* 57, 106–111. <https://doi.org/10.1016/j.cedpsych.2019.01.007>.
- Plano Clark, V.L., Ivankova, N.V., 2016. C, vol. 3. Sage.
- Plano Clark, V.L., Sanders, K., 2015. The use of visual displays in mixed methods research. In: McCrudden, M., Schraw, G., Buckendahl, C. (Eds.), *Use of Visual Displays in Research and Testing*. Information Age Publishing, pp. 177–206.
- Ryan, G.W., Bernard, H.R., 2021. Ethnographic decision models with qualitative data: a thoroughly mixed method 1. In: *The Routledge Reviewer's Guide to Mixed Methods Analysis*. Routledge, pp. 173–184.
- Schifferdecker, K.E., Reed, V.A., 2009. Using mixed methods research in medical education: Basic guidelines for researchers. *Med. Educ.* 43 (7), 637–644. <https://doi.org/10.1111/j.1365-2923.2009.03386.x>.
- White, A.M., DeCuir-Gunby, J.T., Kim, S., 2019. A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of African American students at HBCUs. *Contemp. Educ. Psychol.* 57, 54–71.
- Younas, A., Pedersen, M., Durante, A., 2020. Characteristics of joint displays illustrating data integration in mixed-methods nursing studies. *J. Adv. Nurs.* 76 (2), 676–686. <https://doi.org/10.1111/jan.14264>.

Visual displays in mixed methods research: integration of methods

Matthew T. McCrudden, Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Visual displays in mixed methods research: integration of methods	763
Visual displays and communicating the relationship between research questions and data collection methods	763
Visual displays and communicating a mixed methods study design procedure	764
Guideline 1: generate a title for the visual display	764
Guideline 2: use either a horizontal or vertical layout for the visual display	766
Guideline 3: draw boxes for the various quantitative and qualitative stages	766
Guideline 4: use letter case to designate priority	766
Guideline 5: use arrows to show the flow	767
Guideline 6: specify procedures for data collection and analysis	767
Guideline 7: specify expected products or outcomes	767
Guideline 8: use concise language	768
Guideline 9: make your visual display simple	768
Guideline 10: keep you display to one page	768
Visual displays and integration of methods	768
Connecting	768
Building	769
Merging	769
Embedding	771
Visual displays and integration of findings	771
Conclusion	772
References	772

Visual displays in mixed methods research: integration of methods

Researchers who use mixed methods regularly include visual displays in their research articles. A visual display is a graphic representation of information (Schraw et al., 2013). Visual displays can be used to show descriptive statistics from a questionnaire, inferential statistics from an experiment, themes from qualitative interviews, and the integration of quantitative and qualitative findings to generate insights beyond a single set of data. In each of these instances, the visual display is used to show the findings from a study after the data have been analyzed. Displaying the findings from an inquiry is an effective way of communicating this information to readers. That said, researchers who conduct mixed methods research can include visual displays for several purposes such as providing an overview of a study's research process, demonstrating integration through methods, and integrating the findings from the quantitative and qualitative strands to draw new insights (Guetterman et al., 2015a, 2015b; Ivankova et al., 2006; Plano Clark and Sanders, 2015). This article provides an overview of guidelines for the design of visual displays in mixed methods research studies for different purposes. I present ideas for using visual displays (1) to show the relationship between research questions and data collection methods, (2) to communicate a mixed methods study design procedure, (3) to demonstrate integration of methods, and (4) to demonstrate integration of findings.

Visual displays and communicating the relationship between research questions and data collection methods

There are numerous mixing purposes for conducting mixed methods research (Greene et al., 1989; Newman et al., 2003). Mixed methods research involves integrating quantitative and qualitative strands in a study (Teddle and Tashakkori, 2009). A strand is "a component of a mixed methods study that encompasses the basic process of conducting quantitative or qualitative research: posing a question, collecting and analyzing data, and interpreting results" (Plano Clark and Ivankova, 2016, p. 119). Researchers use research questions to focus a study's purpose on specific statements that can be addressed empirically. In mixed methods studies, researchers can pose three main types of questions: quantitative, qualitative, and mixed methods research questions (Creswell and Plano Clark, 2018). Researchers can show an overarching mixed methods purpose and pose quantitative questions for the quantitative strand and qualitative questions for the qualitative strand. Or they can pose mixed methods research questions that signal the integration of the two strands. A discussion about the framing of research questions is beyond the scope of this article (please see Creswell and Plano Clark, 2018 for a discussion about the framing of research questions in a mixed methods study). The point being raised here is that researchers often pose multiple research questions in a mixed methods

Table 1 The relationship between research questions and data collection methods.

Research questions	Data collection methods
How do students act when they do not understand?	<i>Video recorded observations:</i> 1-hr habituation observation followed by five 1-hr observations of each class. <i>Running records:</i> For each observation, the first author took a running record of classroom events while the teacher introduced and concluded the lesson. <i>Participant observation:</i> The first author roved the class and spoke with students as they attempted practice questions. <i>Work samples:</i> Collect and photocopy students' work during lessons to determine who had struggled with the math they were learning. <i>Semi-structured teacher interviews:</i> 1-hr interview with each teacher to discuss focus students' behavior.
What is the tacit reasoning that explains students' particular patterns of help-seeking?	<i>Stimulated recall student interviews:</i> Students described aspects of their help-seeking behavior that were not directly observed (e.g., attending remedial lessons). <i>Stimulated recall interviews:</i> 30-min interviews in which students ($N = 18$) received verbal descriptions of their class behavior and were asked what they were thinking and feeling at those times.
What are the consequences of students' help-seeking/avoidance for their academic outcomes?	<i>Achievement data:</i> Collect math achievement data for the school year for all students ($N = 18$) who were interviewed. <i>Work samples:</i> Photocopy all student workbooks at the completion of each observation.

Note: Modified from Peeters et al. (2020, p. 537).

inquiry and different data collection methods are used to address these research questions. As such, both readers and the researchers themselves (when in the planning stages of a study) benefit from seeing the relationship between research questions and data collection methods.

Data collection methods are the specific means used to gather evidence to address research questions. In a mixed methods study, it is common for there to be more than one research question and more than one data collection method. When researchers are planning a study, mapping the relationship between research questions and data collection methods helps readers evaluate the extent to which the research questions and data collection methods are aligned. That is, whether the methods are appropriate for addressing the questions. For the reader, research questions and data collection methods are likely to be spread across multiple pages of an article, making it more difficult to see how the different methods will be used to address the research questions. Thus, centralizing the research questions and data collection methods in a visual display is beneficial to the researchers and the readers alike.

The visual display in **Table 1** provides an illustrative example based on Peeters et al. (2020). The research questions are provided in the left-hand column, and the adjoining data collection methods are provided in the right-hand column. Lines are included to make it easier to see how each specific method is aligned with the respective research question. Notice how the researchers provide concise descriptions of the data collection methods. They include key information in the descriptions so that readers can understand and potentially envision the enactment of each method, but they do not provide too much detail. They used the text of the article to provide more detailed descriptions of the specific methods. Thus, the description in the visual display is meant to provide a brief description of the data collection methods, not to replace the detailed description in the text of the article.

Visual displays and communicating a mixed methods study design procedure

Visual displays can play a key role in providing an overview of a study. A display can succinctly summarize the study design procedure in one place and illustrate the adage that "a picture is worth a thousand words". A mixed methods study includes two or more research strands. It can be easier for readers to comprehend the strands of a study by seeing the study design procedure. Mixed methods researchers emphasize the value of using a visual display to communicate the design procedures of a study (Creswell et al., 2003; Creswell and Guetterman, 2019; Morse 1991; Tashakkori and Teddlie 1998). Based on Morse's (1991) notation system and recommendations from mixed methods researchers (Creswell et al., 2003; Creswell and Guetterman, 2019; Tashakkori and Teddlie 1998), Ivankova et al. (2006) developed 10 guidelines for creating a visual display of a mixed methods study design procedure (see **Table 2**). Next, I will describe each of these guidelines in more detail and provide examples to illustrate them.

Guideline 1: generate a title for the visual display

This guideline signals to the reader what is being shown in the display. Displays can be used to for different purposes. Thus, signaling that the display shows the design procedures is a straightforward and efficient way of helping readers know the purpose of the display. To facilitate comprehension, it is beneficial to signal not only that the display shows the procedures, but also the type

Table 2 Guidelines for creating a visual display of a mixed methods study design procedure.

1. Generate a title for the visual display.
2. Use either a horizontal or vertical layout for the visual display.
3. Draw boxes for quantitative and qualitative stages of data collection, data analysis, and interpretation of the study results.
4. Use capitalized or lowercase letters to designate priority of quantitative and qualitative data collection and analysis.
5. Use single-headed arrows to show the flow of procedures in the design.
6. Specify procedures for each quantitative and qualitative data collection and analysis stage.
7. Specify expected products or outcomes of each quantitative and qualitative data collection and analysis procedure.
8. Use concise language to describe procedures and products.
9. Make your visual display simple.
10. Size your visual display to a one-page limit.

Note: Modified from Ivankova et al. (2006, p. 12).

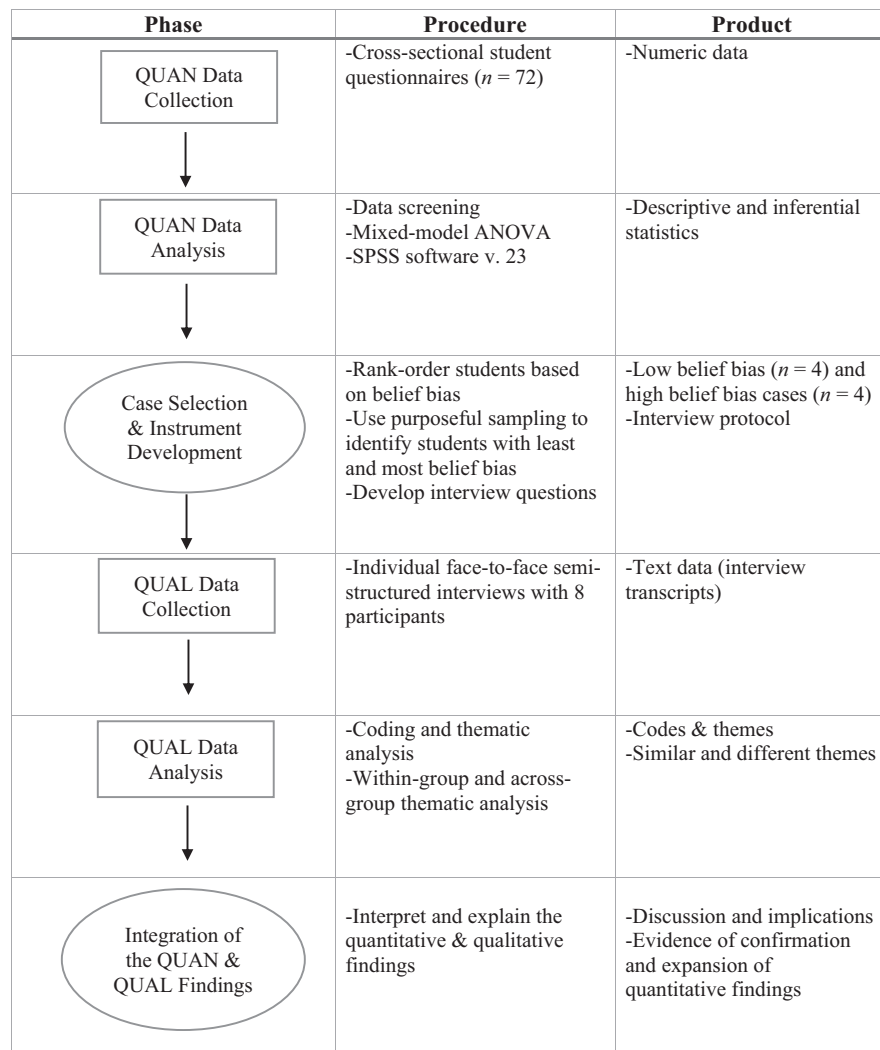
Phase	Procedure	Product
 QUAN Data Collection	-Cross-sectional student questionnaires ($n = 72$)	-Numeric data
 QUAN Data Analysis	-Data screening -Mixed-model ANOVA -SPSS software v. 23	-Descriptive and inferential statistics
 Case Selection & Instrument Development	-Rank-order students based on belief bias -Use purposeful sampling to identify students with least and most belief bias -Develop interview questions	-Low belief bias ($n = 4$) and high belief bias cases ($n = 4$) -Interview protocol
 QUAL Data Collection	-Individual face-to-face semi-structured interviews with 8 participants	-Text data (interview transcripts)
 QUAL Data Analysis	-Coding and thematic analysis -Within-group and across-group thematic analysis	-Codes & themes -Similar and different themes
 Integration of the QUAN & QUAL Findings	-Interpret and explain the quantitative & qualitative findings	-Discussion and implications -Evidence of confirmation and expansion of quantitative findings

Fig. 1 Example vertical visual display of the Sequential Quantitative → Qualitative Mixed Methods Study Design Procedure described in McCrudden and Barnes (2016).

of design being used in the study. This gives the reader a frame of reference for understanding how the various parts of the study fit together. For instance, if the study uses a sequential Quantitative → Qualitative design, use a title such as, “Visual Display of the Sequential Quantitative → Qualitative Mixed Methods Study Design Procedure”, then the reader can anticipate that the display will first show a quantitative strand followed by a qualitative strand. Fig. 1 is an example of a visual display that shows the study design procedure from McCrudden and Barnes (2016).

Guideline 2: use either a horizontal or vertical layout for the visual display

The decision to use either a horizontal or vertical layout is influenced by the complexity of the study design and space available. One suggestion is to create a display with a horizontal layout and another display with a vertical layout and decide whether one layout more effectively communicates the study design procedures.

A sequence is a type of visual display that affords temporal inferences through spatial organization, commonly with other conventions such as arrows, to signal earlier and later steps/events (McCrudden and Rapp, 2017; McCrudden and Van Meter, 2021). A study's design procedure unfolds over time; thus, a sequence is an optimal visual display for conveying a study's design. Consider Fig. 1. As can be seen in the column labeled "phase", the steps taken are provided in boxes and ovals that are connected by arrows to show the temporal sequence of the study design procedures.

Further, in many instances, a sequence can be extended into a matrix. A matrix is a type of visual display that affords relational inferences through spatial contiguity (McCrudden and Rapp, 2017; McCrudden and Van Meter, 2021). The organization of the overall visual display affects the layout of the matrix. With a vertical display, the sequence can be placed in the left column of the display and flows from top to bottom (see Fig. 1). The middle column provides the procedures, the right column provides the products, and below each of those column headings are the details about the respective topics. With a horizontal display, the sequence is placed across the top row of the visual display and flows from left to right as can be seen in Fig. 2, which includes the same information as Fig. 1 but with a horizontal layout. Below the sequence are the next two rows labeled procedures and products, and the details about each are provided in the cells of the display.

The layout of the visual display is highly dependent on the study design. The examples in Figs. 1 and 2 are for a sequential Quantitative → Qualitative design. A visual display of the study design procedures for a concurrent Quantitative + qualitative design will have a different layout. For instance, Fig. 3 is an adapted visual display from White et al. (2019), who used a concurrent Quantitative + qualitative design. The examples shown here illustrate have one quantitative strand and one qualitative strand. It is important to keep in mind that designs with more strands may have different formats.

Guideline 3: draw boxes for the various quantitative and qualitative stages

The third guideline is to use boxes for the components of the quantitative and qualitative strands (e.g., data collection and data analysis), and ovals for points of interface between the strands. For instance, in Fig. 1, "QUAN Data Collection" appears in a box, whereas "Case Selection and Instrument Development" appears in an oval. Points of interface refer to specific points in the study design procedure that involve integration between strands, such as through purposeful sampling (Morse and Niehaus, 2009). In the next section, examples of integration between methods for the different strands will be provided. The overall visual display of the study design procedure can signal these points of integration and provide the reader with a zoom out view of the procedure, whereas the detailed displays can provide a zoom in view of more specific aspects of the procedure. There is nothing magical about using boxes or ovals but following some sort of consistent conventions makes interpretation of visual displays more user-friendly.

Guideline 4: use letter case to designate priority

The fourth guideline is to use capitalized or lowercase letters to designate the priority of quantitative and qualitative data collection and analysis. Priority refers to "the relative importance of the quantitative and qualitative methods for answering a mixed methods study's research questions" (Plano Clark and Ivankova, 2016, p. 33). The priority of the methods in a study is signaled with

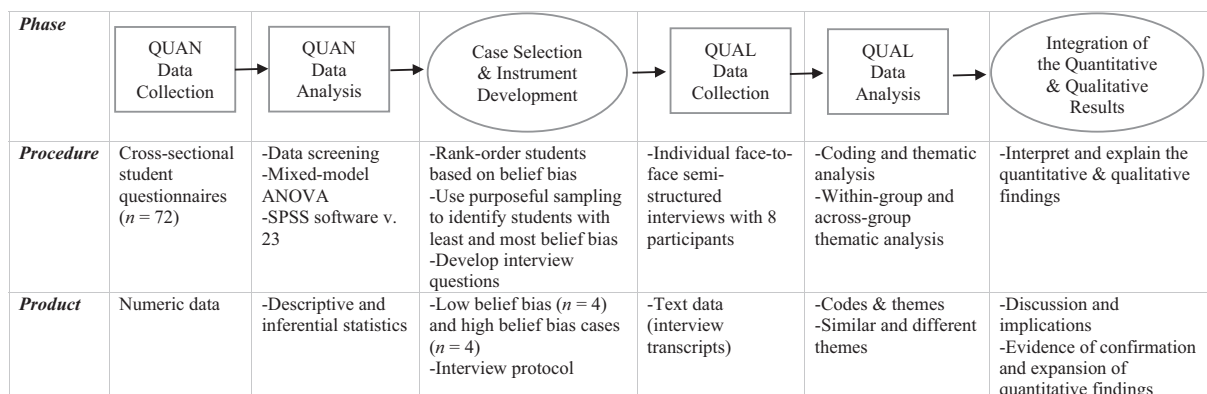


Fig. 2 Example horizontal visual display of the Sequential Quantitative → Qualitative Mixed Methods Study Design Procedure described in McCrudden and Barnes (2016).

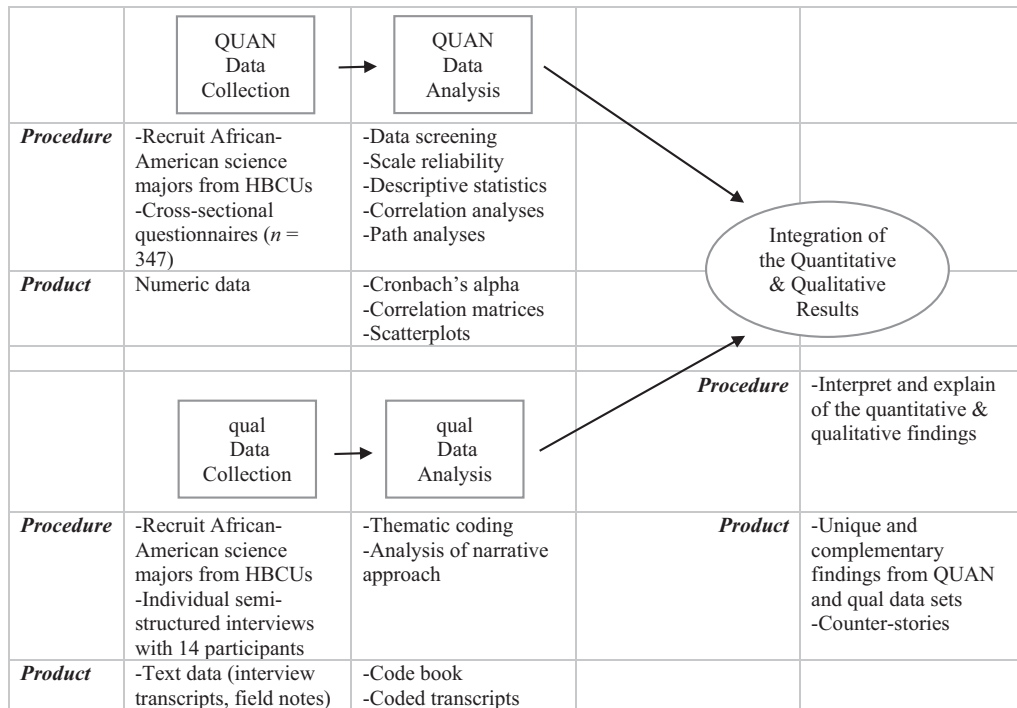


Fig. 3 Example visual display of concurrent Quantitative + Qualitative Study Design Procedure modified from White et al. (2019, p. 59).

uppercase and lowercase letters (Morse, 1991). When one or more of the methods are given greater priority, uppercase letters are used (i.e., QUAN, QUAL). When one or more methods are secondary, lowercase letters are used (i.e., quan, qual). In Fig. 1, the use of uppercase letters for both strands signals that both the quantitative and qualitative strands are given priority, whereas the use of uppercase and lowercase letters in Fig. 3 signals that the quantitative strand is given priority over the qualitative strand.

Guideline 5: use arrows to show the flow

Single-headed arrows can be used to show the flow or sequence of procedures in the design. Whether the visual display is vertical (Fig. 1) or horizontal (Figs. 2 and 3), the use of single-headed arrows signals the temporal sequence of the study design procedures. This helps the reader understand the study's flow and how the earlier and later steps/events are interrelated. In Figs. 1 and 2, the arrows signal the sequential flow of the entire study. In Fig. 3, the arrows signal (a) the sequential flow of each strand, (b) the concurrent flow of each strand, and (c) the integration of the two strands at the end.

Guideline 6: specify procedures for data collection and analysis

The sixth guideline is to specify the procedures for the quantitative and qualitative data collection and analysis stages. In a mixed methods research study, researchers use several procedures to collect the data and several approaches to analyze the data. To complicate matters, the text description of data collection and analysis for different strands may be described across disparate pages of the article. Summarizing the procedures for data collection and analysis for each strand of the study in a visual display provides the reader with a succinct summary of these steps. Specifying the procedures of data collection and analysis go hand-in-hand with expected products or outcomes. In the next section, I will briefly describe expected products and then provide an example of how procedures and products can be presented in a visual display.

Guideline 7: specify expected products or outcomes

The seventh guideline is to specify expected products or outcomes of each quantitative and qualitative data collection and analysis procedure. The sixth guideline specifies how data will be obtained and the analytic approaches to be used. The seventh guideline specifies the product of each data collection procedure and the product of each data analytic approach. In the quantitative strand, the product of data collection will be data from structured data-collection instruments that be in the form of variables, and the product of data analysis will commonly be descriptive and/or inferential statistics. In the qualitative strand, the product of data collection includes, for example, interviews, observations, and field notes, and the product of data analysis can include words, categories, themes, and images.

To illustrate the procedures and products for data collection and analysis, consider [Fig. 1](#). In the quantitative phase, data collection in the quantitative phase involved the use of a cross-sectional survey of 72 participants (procedure column) and resulted in numeric data (product column). Data analysis in the quantitative phase involved data screening, a brief description of the specific data analysis technique used to analyze the data (mixed model ANOVA), and the software used to analyze the data (SPSS software volume 23). The point of interface between the quantitative and qualitative strands involved case selection and instrument development. The procedures included a rank-ordering of participants based on their quantitative responses, which was used to develop criteria to purposefully sample participants for the follow-up qualitative strand (case selection). This led to the selection of eight participants for follow-up interviews. Further, quantitative data were reviewed and used to develop interview questions, which were used to create the interview protocol (instrument development). The procedure and product for the qualitative strand, their integration, and their respective procedures and products can be seen in [Fig. 1](#). Thus, the including the phase, procedure, and product columns in a visual display can provide readers a succinct overview of your mixed methods study design.

Guideline 8: use concise language

The eighth guideline is to use concise language for describing procedures and products. Readers will better understand a visual display if extraneous or unnecessary words are not included ([Mayer and Fiorella, 2014](#)). When too much information is provided, a display can become too complicated or confusing. In [Fig. 1](#), next to the oval for “Case Selection & Instrument Development”, concise language is used to describe the purposeful sampling procedure and for the development of the interview protocol. These study features can be described elsewhere in greater detail. Thus, eliminating unnecessary words helps readers focus on the important information needed to understand the study design procedures.

Guideline 9: make your visual display simple

As indicated in the eighth guideline, too much information in a display, not following general conventions, or general lack of clarity can defeat the purpose of the visual display, which is to provide a clear and concise overview of the study design procedure. For designs with two strands, it is more straightforward to develop a visual display. However, not all designs consist of only two strands. If there are more than two strands in your study is the case, you should consider drafting several different displays or using different layouts and deciding which one best addresses your purposes. Further, consider asking a co-author or colleague to review and interpret this display and use this feedback to decide whether and what changes might be needed.

Guideline 10: keep you display to one page

The 10th and final guideline is to limit your visual display to one page or less. You do not need to include every detail of the study in one visual display. You can show more detailed aspects of the study in a separate visual display. For instance, in a study design procedure, you may briefly describe the purposeful sampling criteria, and then in a separate visual display or in text you could provide a more detailed description of the purposeful sampling procedure and a description of the specific participants.

Visual displays and integration of methods

Researchers can integrate the data collection and data analysis strands of a mixed methods study in several ways. [Fetters et al. \(2013\)](#) identified four ways (connecting, building, merging, & embedding) in which researchers can link quantitative and qualitative data collection and analysis in a mixed methods study. Importantly, each of these approaches can be shown with a visual display.

Connecting

The first approach to integration through methods involves linking one database to the other through sampling (connecting). For instance, suppose researchers collect and analyze quantitative data from an experiment and want to gain a better understanding of some quantitative findings with qualitative interviews. The researchers could use the quantitative findings to inform the sampling frame for the qualitative interviews. For instance, in [McCrudden and Barnes \(2016\)](#), high school students indicated their beliefs about whether humans affect the climate and then rated equally compelling strong arguments for and against human influence on climate and equally compelling weak arguments. Their belief ratings were used to determine whether the arguments were consistent or inconsistent with their beliefs. They used these data to compute bias scores for each participant, defined as the difference between summed ratings for the belief-consistent and belief-inconsistent arguments (a low score would indicate low bias as ratings would be similar independently of whether the argument was consistent or inconsistent with the student’s beliefs). Based on these scores, they identified a subsample of students to participate in qualitative interviews and provided a visual display to describe and contrast the two distinct subgroups (see [Table 3](#)).

Table 3 Participant selection display using ratings for belief-consistent and belief-inconsistent arguments by strength.

		<i>More-objective group (n = 8)</i>			<i>Less-objective group (n = 8)</i>		
		<i>Belief-consistent arguments</i>	<i>Belief-inconsistent arguments</i>	<i>Difference score</i>	<i>Belief-consistent arguments</i>	<i>Belief-inconsistent arguments</i>	<i>Difference score</i>
		<i>M (SE)</i>	<i>M (SE)</i>		<i>M (SE)</i>	<i>M (SE)</i>	
Argument strength	Strong	6.25 (0.60)	5.69 (0.53)	0.56	6.61 (0.82)	3.06 (0.88)	3.55
	Weak	3.44 (0.82)	3.56 (0.82)	-0.12	5.33 (0.61)	2.39 (0.74)	2.94

Modified from McCrudden and Barnes (2016, p. 285).

As seen in **Table 3**, the top rows provide labels for each of the subgroups (more-objective and less-objective groups), their average ratings for belief-consistent and belief-inconsistent arguments for the strong and weak arguments, and their difference scores. The table shows differences between the two groups and provides a rationale or justification for their inclusion in the follow-up qualitative strand of the study. For instance, the difference scores for the strong arguments for the students in the more-objective group are much lower than the difference scores for the strong arguments from the students in the less-objective group (0.56 vs. 3.55), which suggests that students in the less-objective group use more biased reasoning. The same pattern holds true for the weak arguments (-0.12 vs. 2.94). Thus, the researchers could use these data to purposefully sample participants for the follow-up qualitative interviews to gain a deeper understanding of the different reasoning patterns students on both groups used to evaluate the arguments. The visual display gives the researcher a basis for identifying participants for the interviews and it makes the participant selection procedure clearer to the reader.

Building

The second approach involves using one database to inform the data collection of the other database (building). In other words, data from a previous strand is used to inform data collection for a subsequent strand. For instance, suppose researchers collect and analyze quantitative survey data and want to gain a better understanding of some quantitative finding with qualitative interviews. The researchers could use the quantitative findings to inform the development of the qualitative interview protocol. [Cunningham-Erves et al. \(2018\)](#) used a sequential Quantitative → Qualitative design to investigate Black mother's intention to vaccinate daughters against Human Papillomavirus (HPV) given the cervical cancer disparity between Black and non-Hispanic White women. In the quantitative strand, Black mothers of daughters ($n = 237$) completed a survey and nine of these women participated in follow-up qualitative interviews that were based on the survey results (see **Table 4**). Based on the quantitative findings, the researchers identified relevant constructs from the survey and provided sample survey items along with corresponding items from the qualitative interview to gain a deeper understanding of the constructs.

This approach can also be used in a sequential Qualitative → Quantitative study design. For instance, suppose researchers collect and analyze qualitative focus-group interview data and want to test the generalizability of the themes identified from the qualitative data analysis. The researchers could use the themes from the qualitative interviews to develop items for a quantitative survey, which could be shown in a visual display. [Kumar et al. \(2019\)](#) used a sequential Qualitative → Quantitative study design to investigate features of culturally responsive learning environments with students from across 12 middle schools in two nearby school districts. In the qualitative strand, they conducted 57 focus-group interviews with students ($n = 333$) from different cultural backgrounds about their interactions and experiences in their schools. They used the themes generated from the qualitative data analysis to develop a quantitative questionnaire and then tested its validity with students ($n = 2894$) who had similar backgrounds as students from the qualitative strand. The researchers linked the themes from the qualitative strand to the factors and survey items in the quantitative strand as shown in **Table 5**, which includes an example excerpt from one of their theme/factor pairs. The left-hand column provides an excerpt from the qualitative strand; specifically, one of the themes identified in the data analysis and illustrative quotes to substantiate the theme. The right-hand column provides an excerpt from the quantitative strand; specifically, one of the factors and the corresponding survey items. Including the qualitative and quantitative strands in the visual display enables readers to understand and trace the relationship between the two strands.

Merging

The third approach involves bringing together the databases together for analysis (merging). With this approach, researchers develop/identify quantitative and qualitative measures around a common topic before collecting data. After collecting data, each data set is analyzed independently. Next, the researcher brings together the findings from each strand together to find convergence, divergence, or extension. For instance, when using a concurrent design, researchers might collect data with a quantitative survey and a qualitative interview. Elements of the survey and interview protocol can be aligned to show how they are linked before data

Table 4 Visual display of developing interview questions from relevant survey results.

Table used by [Cunningham-Erves et al. \(2018\)](#).

Reprinted from Cunningham-Erves, Forbes, Ivankova, Mayo-Gamble, Kelly-Taylor, and Deakings, Black mother's intention to vaccinate daughters against HPV: A mixed methods approach to identify opportunities for targeted communication. *Gynecologic Oncology*, 149(3), 506–12 (2018), with permission from Elsevier.

<i>Applied theory</i>	<i>Construct</i>	<i>Sample survey (quan) item</i>	<i>Sample interview (qual) protocol item</i>
TRA	Distal variable (knowledge)	Only women get HPV.	What do you know about HPV and cervical cancer?
TRA	Distal variable (personal experience)	Have you, or anyone close to you, ever had HPV, an abnormal Pap smear (pap test), genital warts, and/or cervical cancer?	What personal experiences have you had with HPV?
TRA	Subjective norms	If your doctor recommends the HPV vaccine, how likely is it that your child will get it?	How have other people you know influenced your decision to vaccinate/not vaccinate your daughter against HPV?
TRA	Distal variable (culture)	I often think about how my actions today will affect my health when I am older.	What role, if any, has your cultural experience played in your decision to vaccinate/not vaccinate your daughter against HPV?
TRA	Perceived barriers	I don't trust the information on the HPV vaccine provided by the physician.	What challenges, if any did you/do you face when deciding to vaccinate/not vaccinate your daughter now?
TRA	Perceived benefits	The HPV vaccine is a good way to protect my daughter's health.	In your opinion, what benefits, if any, would there be to vaccinate/not vaccinate your daughter against HPV?

Note: TRA stands for Theory of Reasoned Action.

Table 5 Visual display for building by linking databases in sequential Qualitative → Quantitative study design.

<i>Qualitative strand</i>	<i>Quantitative strand</i>
Theme: Perceptions of teachers as prejudiced and culturally insensitive versus respectful and culturally responsive Quotes: • "They should talk to you with some respect that we talk to them with. Some teachers don't have any respect for you even though you wouldn't disrespect." • "We have like our own whole caste system going on that the teachers have like completely no idea about." • Female 1: "They don't say anything about it (political issues). Anytime someone tries to bring it up, it's inappropriate. It's too political. Shouldn't talk about that here. It's not for school." • Female 2: "See both sides. Don't make it that you're talking about one side and not the other. Like all the teachers ..."	Factor: Promoting cultural openness and positive intergroup relationships Survey items: 1. Listen with an open mind to what students in the different groups are saying. 2. Help students in each of the groups understand how students from the other group feel.

Modified from [Kumar et al. \(2019, p. 92\)](#).

collection (see [Table 6](#)). [Ogilvie and McCrudden \(2017\)](#) used a concurrent Quantitative + Qualitative design to investigate the social validity of a behavioral intervention for children with autism. Parents completed a quantitative questionnaire and participated in qualitative interviews to investigate their perceptions of the social validity of the behavioral intervention for their children. [Table 6](#) consists of three columns and includes an excerpt from the materials from [Ogilvie and McCrudden \(2017\)](#). The left column provides the factor or category targeted by the specific quantitative survey items (middle column) and the specific qualitative interview prompts (right column).

Table 6 Visual display for merging by linking databases for analysis in a concurrent Quantitative + Qualitative design.

Category	Quantitative survey items	Qualitative interview prompts
Understanding	- How clear is your understanding of this treatment?	- What aspects of this treatment do you understand well? - What aspects of this treatment do you not understand? - How can we make it more understandable in the future?

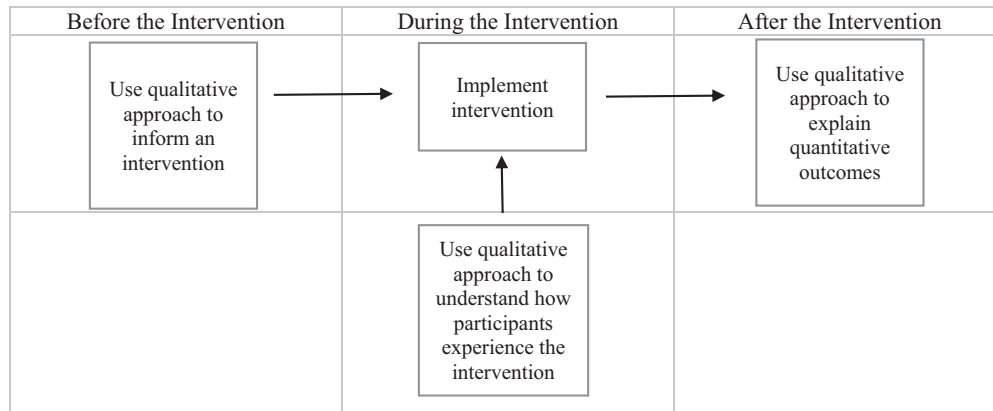
Modified from Ogilvie and McCrudden (2017, p. 2908).

Embedding

The fourth approach involves linking data collection and analysis at multiple points throughout a study (embedding). Embedding “is especially important in interventional advanced designs, but it can also occur in other designs” (Fetters et al., 2013, p. 2141) and it can involve different combinations of integration through methods (connecting, building, merging), although embedding can be used for a variety of mixed methods research inquiries. For an interventional advanced design, a qualitative data can be embedded at various points in the design, including before, during, and/or after the intervention (see Fig. 4). Creswell and Plano Clark (2018) provide detailed examples of different ways to use visual displays to show advanced designs.

Visual displays and integration of findings

Researchers can use visual displays to integrate the quantitative and qualitative findings of a mixed methods study. Visual displays can be used to communicate important findings from the respective quantitative and qualitative strands and to support inferences about the integration of the strands via spatial arrangement. An example is an integrated results matrix (see Table 7) based on McCrudden and Barnes (2016). This type of display facilitates comparison of the quantitative and qualitative results by organizing the results from the respective strands. In Table 7, the second column of the display (Quantitative

**Fig. 4** General visual display for embedding qualitative approaches in an interventional design. Modified from Guetterman et al. (2015a, p. 169).**Table 7** Integrated results matrix.

Group	Quantitative findings	Qualitative findings	Meta-inference
More-objective	Rated strong arguments more favorably than weak arguments, independently of belief-consistency.	Uniform application of normative evaluation criteria at a deep level.	Holding a belief did not necessarily lead to biased reasoning; rather, biased reasoning occurred when individuals applied a more critical standard of evaluation to belief-inconsistent arguments.
Less-objective	Rated belief-consistent arguments more favorably than belief-inconsistent arguments, independently of strength.	Belief-based application of normative evaluation criteria (more critical of belief-inconsistent arguments) and use of ideas that were irrelevant to the argument's validity.	

Modified from McCrudden and Barnes (2016, p. 292).

findings) shows important quantitative data and a summary for each group; the third column (Qualitative findings) provides a summary for of the qualitative themes for each group. The last column on the right provides the meta-inference, which is a conclusion a researcher generates by integrating the inferences obtained from the quantitative and qualitative strands of a mixed methods study (Teddle and Tashakkori, 2009). There are several ways to develop joint displays to convey findings from a mixed methods research study. Guetterman et al. (2015a, p. 158) list nine guidelines for creating joint displays and Guetterman et al. (2015b) provide illustrative examples of joint displays that have appeared in journals that publish state-of-the-art mixed methods research.

Conclusion

Visual displays can give readers a conceptual scaffold for understanding information (McCrudden et al., 2015; McCrudden and Rapp, 2017). Researchers can use visual displays in mixed methods research in ways that effectively communicate key aspects of a study. This review highlighted four purposes of visual displays in mixed methods research, including to show the relationship between research questions and data collection methods, to communicate a mixed methods study design procedure, to demonstrate integration of methods, and to demonstrate integration of findings. Each of these purposes pertains to the methodological features of a mixed methods study. For instance, an effective visual display that shows a mixed methods study design procedure succinctly describes the steps taken to collect, analyze, and interpret data. However, there are other purposes that are not covered here, such as integration of paradigms. Further, the examples provided here represent a handful of different ways the visual displays can be used in mixed methods research. Other formats, such as the use of geocoding, Venn diagrams, graphs, are but a few possibilities. I hope that the examples here are useful and readers are encouraged to consider additional possibilities including different formats and purposes. In sum, considering the purposes and formats of visual displays can guide researchers to identify effective ways of communicating information and ideas about mixed methods research.

References

- Creswell, J.W., Guetterman, T.C., 2019. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Approaches to Research*, sixth ed. Merrill/Pearson Education, Upper Saddle River, NJ.
- Creswell, J.W., Plano Clark, V., 2018. *Designing and Conducting Mixed Methods Research*. Sage Publications, Inc, Thousand Oaks, CA.
- Creswell, J.W., Plano Clark, V., Gutmann, M., Hanson, W., 2003. Advanced mixed methods research designs. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook on Mixed Methods in the Behavioral and Social Sciences*. Sage, Thousand Oaks, CA, pp. 209–240.
- Cunningham-Erves, J., Forbes, L., Ivankova, N., Mayo-Gamble, T., Kelly-Taylor, K., Deakings, J., 2018. Black mother's intention to vaccinate daughters against HPV: a mixed methods approach to identify opportunities for targeted communication. *Gynecol. Oncol.* 149 (3), 506–512. <https://doi.org/10.1016/j.ygyno.2018.03.047>.
- Fetters, M.D., Curry, L.A., Creswell, J.W., 2013. Achieving integration in mixed methods designs-principles and practices. *Health Serv. Res.* 48 (6 Pt 2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>.
- Greene, J.C., Caracelli, V.J., Graham, W.F., 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Pol. Anal.* 11, 255–274. <https://doi.org/10.3102/01623737011003255>.
- Guetterman, T., Creswell, J.W., Kuckartz, U., 2015a. Using joint displays and MAXQ software to represent the results of mixed methods research. In: McCrudden, M.T., Schraw, G., Buckendahl, C. (Eds.), *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*. Information Age Publishing, Charlotte, NC, pp. 145–175.
- Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015b. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann. Fam. Med.* 13 (6), 554–561. <https://doi.org/10.1370/afm.1865>.
- Ivankova, N.V., Creswell, J.W., Stick, S.L., 2006. Using mixed methods sequential explanatory design: from theory to practice. *Field Methods* 18 (1), 3–20. <https://doi.org/10.1177/1525822X05282260>.
- Kumar, R., Karabenick, S., Warnke, J., Hany, S., Seay, N., 2019. Culturally inclusive and responsive curricular learning environments (CIRCLEs): an exploratory sequential mixed-methods approach. *Contemp. Educ. Psychol.* 57, 87–105. <https://doi.org/10.1016/j.cedpsych.2018.10.005>.
- Mayer, R.E., Fiorella, L., 2014. Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, New York, pp. 279–315.
- McCrudden, M.T., Barnes, A., 2016. Differences in student reasoning about belief-relevant arguments: a mixed methods study. *Metacogn. Learn.* 11 (3), 275–303. <https://doi.org/10.1007/s11409-015-9148-0>.
- McCrudden, M.T., Rapp, D.N., 2017. How visual displays affect cognitive processing. *Educ. Psychol. Rev.* 29 (3), 623–639. <https://doi.org/10.1007/s10648-015-9342-2>.
- McCrudden, M.T., Van Meter, P., 2021. Multimedia learning with visual displays. In: Mayer, R., Fiorella, L. (Eds.), *Cambridge Handbook of Multimedia Learning*, third ed. Cambridge University Press, Cambridge, UK, pp. 512–522.
- McCrudden, M.T., Schraw, G., Buckendahl, C. (Eds.), 2015. *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*. Information Age Publishing, Charlotte, NC.
- Morse, J.M., Niehaus, L., 2009. *Mixed Method Design: Principles and Procedures*. Left Coast Press, Walnut Creek, CA.
- Morse, J.M., 1991. Approaches to qualitative-quantitative methodological triangulation. *Nurs. Res.* 40, 120–123.
- Newman, I., Ridenour, C.S., Newman, C., DeMarco, G.M.P., 2003. A typology of research purposes and its relationship to mixed methods. In: Tashakkori, A., Teddlie, C. (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Sage, Thousand Oaks, CA, pp. 167–188.
- Ogilvie, E., McCrudden, M.T., 2017. Evaluating the social validity of the Early Start Denver Model: a convergent mixed methods study. *J. Autism Dev. Disord.* 47 (9), 2899–2910. <https://doi.org/10.1007/s10803-017-3214-1>.
- Peeters, A., Robinson, V., Rubie-Davies, C., 2020. Theories in use that explain adolescent help seeking and avoidance in mathematics. *J. Educ. Psychol.* 112 (3), 533–550. <https://doi.org/10.1037/edu0000423>.
- Plano Clark, V., Ivankova, N., 2016. *Mixed Methods Research. A Guide to the Field*. Sage, Thousand Oaks, CA. <https://doi.org/10.4135/9781483398341>.

- Plano Clark, V.L., Sanders, K., 2015. The use of visual displays in mixed methods research: strategies for effectively integrating the quantitative and qualitative components of a study. In: McCrudden, M.T., Schraw, G., Buckendahl, C. (Eds.), *Use of Visual Displays in Research and Testing: Coding, Interpreting, and Reporting Data*. Information Age Publishing, Charlotte, NC, pp. 177–206.
- Schraw, G., McCrudden, M.T., Robinson, D., 2013. Visual displays and learning: theoretical and practical considerations. In: Schraw, G., McCrudden, M.T., Robinson, D. (Eds.), *Learning through Visual Displays*. Information Age Publishing, Charlotte, NC, pp. 3–17.
- Tashakkori, A., Teddlie, C., 1998. *Mixed methodology: combining qualitative and quantitative approaches*. In: *Applied Social Research Methods Series*, vol. 46. Sage, Thousand Oaks, CA.
- Teddlie, C., Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Thousand Oaks, CA. ISBN: 9780761930129.
- White, A.M., DeCuir-Gunby, J.T., Kim, S., 2019. A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of African American students at HBCUs. *Contemp. Educ. Psychol.* 57, 54–71. <https://doi.org/10.1016/j.cedpsych.2018.11.006>.

FOURTH EDITION

International Encyclopedia of Education

Assessment, Evaluation
and Accountability

VOLUME
13

Editors-In-Chief

Robert J. Tierney
Fazal Rizvi
Kadriye Ercikan

Volume Editor

Therese Hopfenbeck



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 13

Assessment, Evaluation and Accountability

VOLUME 13 EDITOR

Therese Hopfenbeck



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

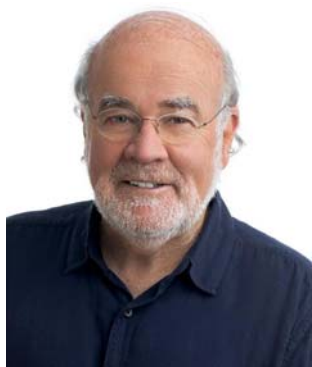
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia–Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "universities" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



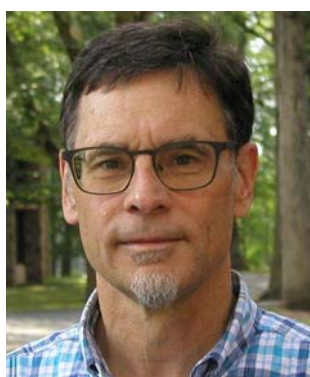
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 13: Assessment, Evaluation and Accountability

Therese N Hopfenbeck, University of Melbourne, Melbourne, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction: An Overview of Educational Assessment, Evaluation, and Accountability

The present section documents recent trends and historic perspectives within educational assessment, evaluation, and accountability studies. The field contains areas on classroom-based assessments, both summative and formative assessment, evaluation studies, alternative assessments, accountability measures, international large-scale studies, and standardized testing. Much of the educational assessment field has historically focused on what teachers do in the classroom, including teachers' instructions and assessment practices, previously referred to as *classroom evaluation* (Natriello, 1987). A related term, *classroom-based assessments*, has also been used, with the main purpose to establish practices which would enhance students' learning and provide feedback on what steps to take in instruction (Black and Wiliam, 1998, 2009). This line of research has also looked at different forms of feedback processes both in classrooms and for students in higher education, with the purpose of documenting what is seen as high-quality feedback within formative assessment processes (Nicol and Macfarlane-Dick, 2006).

The recent decades have witnessed a push for outcome data collected through large-scale studies in order to monitor and evaluate education systems nationally, in school districts and at school level, and to be able to compare outcomes across countries, districts, and schools. The increasing focus upon international comparison and pressure on holding teachers, school leaders, and district leaders accountable for students' performance, has become one of the most researched fields within the educational assessment, as reflected in the many chapters in this section.

Further, there has been a call for increasing teachers' assessment literacy and to better understand and use information from both classroom-based assessments and standardized testing programs, i.e., national and international assessments. Some have even argued that teachers need to play a more active role in high-stakes summative assessment because even the best external tests are bound to be of limited validity (Black, 2013, p. 171). Understanding the field of educational assessment therefore involves knowledge of how classroom-based assessment and standardized testing work together, both in terms of practice and policy. This section on *Educational Assessment, Evaluation and Accountability* provides such knowledge.

The current section contains of eight parts; (1) Assessment and Learning, (2) Educational Assessment—Assessment Activities, (3) Educational Assessment—Policy and Practice, (4) Educational Accountability—Purpose, Models, and Data use, (5) Philosophical and Theoretical Underpinnings of Assessment, Evaluation, and Accountability, (6) Educational Evaluation—Contemporary Themes in Evaluation, (7) Educational Evaluation—Key concepts and practices, and (8) Educational Accountability—Global trends and future directions.

The first four parts (1, 2, 3, and 4) have chapters outlining what historically was seen as one of the most important areas for teachers, *classroom-based assessment*, what every teacher should know and be able to do, to improve students' learning. The first part of this overview therefore provides chapters covering what generally has been viewed as necessary skills for teachers with respect to assessment literacy. It includes chapters outlining the theoretical and conceptual work around formative assessment, also known as "assessment for learning," with examples of how students can enhance their learning through self-assessments, peer-assessments, self-regulation, and other more dialogical practices (Allal, 2023; James, 2023; Topping, 2023). Other chapters also outline theoretically and empirically how different forms of feedback can support students' learning, when used timely and with high quality (Ajjawi et al., 2023; Gamløy and Vattøy, 2023). The section further includes chapters on different forms of assessments, and outlines alternative ways of assessing both student outcomes and teachers' instruction and teaching (McGrane, 2023).

Classroom-based assessment research has focused upon the differences in formative and summative assessment, and how such practices can enhance or prevent students' learning. While formative assessment practices have often been characterized as low-stakes, as these assessment practices usually had no high-stake consequences for students or teachers, summative assessment practices are usually seen as high-stakes. Standardized external test results are intended to determine not only the individual student's future possibilities for university places and future careers but aggregated test results may also be used to rank and compare schools in districts, states, or countries. Several of the chapters in this section discuss the tension between high-stakes and low-stakes assessment and the balance between formative and summative assessment practices.

The current section further includes new empirical studies on the experiences from the Global South, where large-scale implementation of assessment for learning practices have offered new knowledge on the possibilities and limitations of these practices. These chapters challenge the status quo of what has been seen as ideals for assessment practices as they question the notion that western practices can be implemented in the Global South. They even make a call for rethinking what is seen as "best practices" of educational assessment (Halai et al., 2023; Klinger and Nicholas, 2023).

From the teacher's perspective, some of the challenges with classroom-based assessment are linked to national and international assessments. In the final four parts (5, 6, 7, and 8) of this section, we find chapters on evaluation studies and future global trends. Globally, we have seen a trend where international large-scale studies provide data to governments and policy-makers, from country and districts at an aggregated level (OECD, 1999; IEA, 2021). While Sherer and Nilsen (2023) argues that the first international

large-scale studies, conducted by the international organization IEA, had more of a research focus rather than policy impact, the OECD studies, such as PISA, have had more of a policy impact and have also received more media attention (Addey et al., 2017).

The primary intention of these datasets has been to inform policy-makers about trends over time, and monitor systems on how they perform. The second section provides an overview of different perspectives from international large-scale studies, the use of such data, and how different accountability measures are now put in place to ensure students' performance are monitored. Some of the chapters provide a critical perspective on these measures, as they document how large-scale standardized assessment can benefit educational stakeholders with valuable data about students' performance when used well, but simultaneously, data can be misused by policy-makers if set targets are not met (Ozga, 2023). Some chapters further explore and question the validity of such data, and to what extent different stakeholders have the necessary knowledge to interpret, use, and act upon the data provided. The second part of this section therefore also includes chapters on validity and assessment, with important discussions on what is needed for education systems to become fair, valid, and reliable. It includes chapters on what is known as intelligent accountability, with examples on how different stakeholders can use assessment data to improve achievement and learning on all levels within an education system.

While international large-scale assessment studies (ILSAs) have in some of the literature been described as low-stake tests as they compare countries rather than evaluate schools, teachers, or students, the use of these rankings as described in some chapters in this section clearly demonstrate the increasingly high-stakes role of ILSAs as policy-makers have used results from the studies to make calls for policy changes in the education systems.

Some chapters do warn that the field of educational assessment runs the risk of being rather ethnocentric, with groups of scholars working in different countries pursuing isolated agendas. One of the challenges for the field of educational assessment will therefore be to bridge the gap between different research groups, to share knowledge and disseminate results and data in a way which enhances transparency. It will also include working with researchers across the globe and acknowledge the importance of different cultures and contexts to better understand how to develop and interpret different assessments. Indeed, Jude (2023) suggests that the increasing number of immigrant students globally demonstrates that one of the most pressing challenges for the future is to secure equality through education policies. With references to the UN sustainable development goals, which states "equality for all and quality in education" (United Nations, 2015), Jude (2023) outlines the importance of using ILSA studies to monitor the progress across countries in order to create education systems that can support a more heterogeneous population and help students develop intercultural skills.

Educational assessment continues to drive what is taught in classrooms, and hence what students learn. The coming decades will determine the balance between classroom-based assessment and ILSAs, their use and interpretations. Those who are in charge of what counts in educational assessment have the power to influence what students will study and learn. The chapters in this section make a powerful call for more intelligent use of assessment data, broader research collaboration, increased focus upon equality for all students and the further development of assessment practices in the classroom which support students' agency.

References

- Addey, C., Sellar, S., Steiner-Khamsi, G., Lingard, B., Verger, A., 2017. The rise of international large-scale assessments and rationales for participation. *Compare* 47, 434–452.
- Ajjawi, R., Joanna, J.T., Dawson, P., 2023. Feedback for learning. In: *International Encyclopedia of Education*, fourth ed.
- Allal, L., 2023. Assessment and the regulation of learning. In: *International Encyclopedia of Education*, fourth ed.
- Black, P., William, D., 1998. Assessment and classroom learning assessment in education. *Princ. Policy Pract.* 5 (1), 7–75.
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21 (1), 5–31.
- Black, P., 2013. Formative and summative aspects of assessment: theoretical and research foundations in the context of pedagogy. In: McMillan, J.H. (Ed.), *Sage Handbook of Research on Classroom Assessment*. Sage, Thousand Oaks, CA, pp. 167–178.
- James, M.E., 2023. Assessing and learning to learn. In: *International Encyclopedia of Education*, fourth ed.
- Jude, N., 2023. Educational assessment and the rise of immigration students?. In: *International Encyclopedia of Education*, fourth ed.
- Halai, A., Sarungi, V., Hopfenbeck, T.N., 2023. Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania. In: *International Encyclopedia of Education*, fourth ed.
- IEA, 2021. TIMSS: Trends in International Mathematics and Science Study [Online]. Available. <https://www.iea.nl/studies/iea/timss>.
- Klinger, D., Nicholas, W., 2023. Building teachers' capacity in assessment for learning in rural East Africa. In: *International Encyclopedia of Education*, fourth ed.
- McGrane, J., 2023. Comparative judgement. In: *International Encyclopedia of Education*, fourth ed.
- Natriello, G., 1987. The impact of evaluation processes on students. *Educ. Psychol.* 22, 155–175.
- Nicol, D.J., Macfarlane-Dick, D., 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. Higher Educ.* 31 (2), 199–218.
- OECD, 1999. *Measuring Student Knowledge and Skills a New Framework for Assessment*. OECD, Paris.
- Ozga, J., 2023. Accountability in an age of international testing. In: *International Encyclopedia of Education*, fourth ed.
- Sherer, R., Nilsen, T., 2023. Accountability for the future? International testing and future directions. In: *International Encyclopedia of Education*, fourth ed.
- Topping, K.J., 2023. Peer and self-assessment. In: *International Encyclopedia of Education*, fourth ed.
- United Nations, 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development* [online]. Available at. https://www.unfpa.org/sites/default/files/resource-pdf/Resolution_A_RES_70_1_EN.pdf.

Further Reading

- Gamlem, S., Vattøy, K.-M., 2023. Feedback and classroom practices. In: *International Encyclopedia of Education*, fourth ed.

CONTRIBUTORS TO VOLUME 13

Rola Ajjawi

*Center for Research in Assessment and Digital Learning,
Deakin University, Melbourne, VIC, Australia*

Linda Allal

University of Geneva, Geneva, Switzerland

Gavin TL Brown

*Faculty of Education & Social Work, The University of
Auckland, Auckland, New Zealand; and Institutionen
för Tillämpad Utbildningsvetenskap [Department of
Applied Educational Science], Umeå Universitet,
Umeå, Sweden*

Hugh Burkhardt

*School of Education, University of Nottingham,
Nottingham, United Kingdom*

Evi Charalambous

*Department of Education, University of Cyprus, Nicosia,
Cyprus*

Gregory J Cizek

*Educational Measurement and Evaluation, School of
Education, University of North Carolina at Chapel Hill,
Chapel Hill, NC, United States*

Phillip Dawson

*Center for Research in Assessment and Digital Learning,
Deakin University, Melbourne, VIC, Australia*

Christopher DeLuca

*Faculty of Education, Queen's University, Kingston, ON,
Canada*

Elizabeth A Dickson

*School of Law, Faculty of Business and Law, QUT,
Brisbane, QLD, Australia*

Jannette Elwood

*Faculty of Arts, Humanities and Social Sciences, Queen's
University Belfast, Belfast, Northern Ireland, United
Kingdom*

Ola Erstad

*Department of Education, University of Oslo, Oslo,
Norway*

Siv M Gamlem

*Faculty of Humanities and Teacher Education, Volda
University College, Volda, Norway*

Melissa Gholson

ETS, Princeton, NJ, United States

Samuel Greiff

*Department of Behavioral and Cognitive Sciences,
University of Luxembourg, Luxembourg*

Sue Grey

*Graduate School of Education, University of Exeter,
Exeter, United Kingdom*

Anjum Halai

*Institute for Educational Development, Aga Khan
University, Karachi, Pakistan*

Areum Han

*Department of Behavioral and Cognitive Sciences,
University of Luxembourg, Luxembourg*

Amy Hanna

*School of Education, University of Strathclyde, Glasgow,
Scotland, United Kingdom*

Michael A Harnar

*Western Michigan University, Kalamazoo, MI, United
States*

Therese N Hopfenbeck

University of Melbourne, Melbourne, Australia

Mary E James

*University of Cambridge, Faculty of Education,
Cambridge, United Kingdom*

J Joy Cumming

*Institute for Learning Sciences and Teacher Education,
Australian Catholic University, Brisbane, QLD,
Australia*

Nina Jude

*Ruprecht-Karls-Universität Heidelberg,
Institut für Bildungswissenschaft, Heidelberg,
Germany*

Lale Khorramdel

TIMSS & PIRLS International Study Center at Boston College, Chestnut Hill, MA, United States

Don A Klinger

Te Wānanga Toi Tangata, Division of Education, University of Waikato, Hamilton, New Zealand

Florian Krieger

Department of Rehabilitation Sciences, TU Dortmund University, Dortmund, Germany

Leonidas Kyriakides

Department of Education, University of Cyprus, Nicosia, Cyprus

Sølvi Lillejord

SLATE, The Center for the Science of Learning and Technology, University of Bergen, Bergen, Norway; and Department of Education, Oxford University, Oxford, United Kingdom

Slki N Lim

Learning Sciences and Psychological Studies, School of Education, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Anastasiya A Lipnevich

Queens College and the Graduate Center, The City University of New York, New York, NY, United States

Scott F Marion

National Center for the Improvement of Educational Assessment

Graham S Maxwell

Institute for Learning Sciences and Teacher Education, Australian Catholic University, QLD, Australia

Joshua A McGrane

Assessment and Evaluation Research Centre (AERC), Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC, Australia

Paul Morris

Professor of Comparative Education IOE, Department of Education Practice and Society, University College London, London, United Kingdom

Paul E Newton

Office of Qualifications and Examinations Regulation, Coventry, United Kingdom

Trude Nilsen

University of Oslo, Faculty of Educational Sciences, Department of Teacher Education and School Research, Oslo, Norway

Jenny Ozga

University of Oxford, Department of Education, Oxford, United Kingdom

Tryfonas Papadamou

Department of Education, University of Cyprus, Nicosia, Cyprus

Rosalina Rawi

National Institute of Education, Asian Languages and Cultures Academic Group, Singapore

David Rutkowski

Indiana University, Bloomington, IN, United States; and Center for Educational Measurement, University of Oslo, Oslo, Norway

Leslie Rutkowski

Indiana University, Bloomington, IN, United States; and Center for Educational Measurement, University of Oslo, Oslo, Norway

Christopher M Saldaña

School of Education, University of Wisconsin, Madison, WI, United States

Veronica Sarungi

Aga Khan Education Services, Dar-es-Salaam, Tanzania

Ronny Scherer

Center for Educational Measurement at the University of Oslo (CEMO), Faculty of Educational Sciences, University of Oslo, Oslo, Norway

Alan H Schoenfeld

School of Education, University of California Berkeley, Berkeley, CA, United States

Lorrie A Shepard

School of Education, University of Colorado, Boulder, CO, United States

Fazilat Siddiq

Unit for Digitalization and Educational Quality, University of South-Eastern Norway, Drammen, Norway

Jeffery K Smith

University of Otago, Dunedin, New Zealand

Gordon Stobart

University College London Institute of Education, London, United Kingdom

Sue Swaffield

University of Cambridge Faculty of Education, Cambridge, United Kingdom

Joanna Tai

Center for Research in Assessment and Digital Learning, Deakin University, Melbourne, VIC, Australia

Sally M Thomas

School of Education, University of Bristol, Bristol, United Kingdom

Zachary D Tilton
Western Michigan University, Kalamazoo, MI, United States

Keith J Topping
University of Dundee, Dundee, United Kingdom

Fabienne van der Kleij
School of Education, University of Queensland, Brisbane, Queensland, Australia

Kim-Daniel Vattøy
Faculty of Humanities and Teacher Education, Volda University College, Volda, Norway

Nicholas Wachira
Aga Khan Education Services, Dar-es-Salaam, Tanzania

Lori A Wingate
Western Michigan University, Kalamazoo, MI, United States

Philip H Winne
Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

E Caroline Wylie
ETS, Princeton, NJ, United States

Christian Ydesen
Department of Culture and Learning, Aalborg University, Aalborg, Denmark

This page intentionally left blank

CONTENTS OF VOLUME 13

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 13: Assessment, Evaluation and Accountability</i>	<i>xix</i>
<i>Contributors to Volume 13</i>	<i>xxi</i>

ASSESSMENT AND LEARNING

Formative assessment: an overview of history, theory and application <i>Gregory J Cizek and Slki N Lim</i>	1
Assessing and learning, and learning to learn <i>Mary E James</i>	10
Assessment for learning <i>Sue Swaffield and Roszalina Rawi</i>	21
Assessment and the regulation of learning <i>Linda Allal</i>	35
Feedback for learning <i>Rola Ajjawi, Joanna Tai, and Phillip Dawson</i>	41
Assessing learning processes <i>Philip H Winne</i>	50
Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania <i>Anjum Halai, Veronica Sarungi, and Therese N Hopfenbeck</i>	63

EDUCATIONAL ASSESSMENT – ASSESSMENT ACTIVITIES

Comparative judgment <i>Joshua A McGrane</i>	73
Peer- and self-assessment <i>Keith J Topping</i>	79

Feedback and classroom practice	89
<i>Siv M Gamlem and Kim-Daniel Vattøy</i>	
Alternative assessment	96
<i>Gordon Stobart</i>	

EDUCATIONAL ASSESSMENT - POLICY AND PRACTICE

Educational assessment – trends in research, policy and practice	103
<i>Christian Ydesen</i>	
U.S. federal assessment policy reforms in the twenty first century and their impacts	108
<i>E Caroline Wylie and Melissa Gholson</i>	
Assessment reform and students' voices	119
<i>Jannette Elwood and Amy Hanna</i>	
Measuring what we cannot see: assessing extreme populations in international large-scale assessments	129
<i>Leslie Rutkowski and David Rutkowski</i>	
Taking purpose into account when designing, using and evaluating educational assessments	136
<i>Paul E Newton</i>	
Implementation of Assessment for Learning	147
<i>Fabienne van der Kleij and Christopher DeLuca</i>	

EDUCATIONAL ACCOUNTABILITY - PURPOSE, MODELS AND DATA USE

A more intelligent accountability—future directions	155
<i>Sølvi Lillejord</i>	
Educational accountability: the use and abuse of assessment data	163
<i>Christopher M Saldaña and Lorrie A Shepard</i>	
Social, educational and legal perspectives on assessment and accountability measures in inclusive education for students with disability	171
<i>J Joy Cumming and Elizabeth A Dickson</i>	
Educational accountability, international large scale assessments and PISA: powerful influences based on weak foundations	181
<i>Sue Grey and Paul Morris</i>	

PHILOSOPHICAL AND THEORETICAL UNDERPINNINGS OF ASSESSMENT, EVALUATION AND ACCOUNTABILITY

Assessment, evaluation, and accountability: a brief introduction	192
<i>Anastasiya A Lipnevich, Lale Khorramdel, and Jeffery K Smith</i>	
Building teachers' capacity in assessment for learning in rural East Africa	202
<i>Nicholas Wachira and Don A Klinger</i>	
Validity for through-year assessments	212
<i>Scott F Marion</i>	
Assessment and mathematical literacy: a brief introduction	220
<i>Hugh Burkhardt and Alan H Schoenfeld</i>	

EDUCATIONAL EVALUATION – CONTEMPORARY THEMES IN EVALUATION

Assessment of collaborative problem-solving: past achievements and current challenges <i>Areum Han, Florian Krieger, and Samuel Greiff</i>	234
Educational assessment of 21st century skills—novel initiatives, yet a lack of systemic transformation <i>Ola Erstad and Fazilat Siddiq</i>	245
Educational assessment and the rise of immigrant students <i>Nina Jude</i>	256

EDUCATIONAL EVALUATION - KEY CONCEPTS AND PRACTICES

Metaevaluation: key concepts explained <i>Lori A Wingate, Zachary D Tilton, and Michael A Harnar</i>	263
Evaluating schools as professional learning communities <i>Sally M Thomas</i>	275
International trends on teacher evaluation <i>Leonidas Kyriakides, Evi Charalambous, and Tryfonas Papadamou</i>	287

EDUCATIONAL ACCOUNTABILITY – GLOBAL TRENDS AND FUTURE DIRECTIONS

Leading school-based assessment for educational improvement: rethinking accountability <i>Gavin TL Brown</i>	298
Accountability in an age of international testing <i>Jenny Ozga</i>	306
Assessment and accountability global trends and future directions in 21st century competencies <i>Graham S Maxwell</i>	313
Accountability for the future? International testing and future directions <i>Ronny Scherer and Trude Nilsen</i>	324

This page intentionally left blank

Formative assessment: an overview of history, theory and application

Gregory J. Cizek^a and Siki N. Lim^b, ^a Educational Measurement and Evaluation, School of Education, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States; and ^b Learning Sciences and Psychological Studies, School of Education, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	1
Foundational concepts	1
Inference	1
Test	2
Assessment	2
Origins of formative assessment	3
Current concepts, confusions, and controversies	3
Distinguishing characteristic 1—primary purpose	4
Distinguishing characteristic 2—timing and feedback	4
Distinguishing characteristic 3—student involvement	4
Formative assessment: a definition	5
Recent research on formative assessment	5
Critical features of effective formative assessment	5
The relationship between formative assessment and self-regulated learning	5
Discipline-specific formative assessment	6
Technology-enhanced formative assessment	7
Conclusions and future of formative assessment	7
References	8

Introduction

Formative assessment has become increasingly popular as strategy for improving teaching and learning since the popularization of that concept deriving from the publication of a research review in an academic journal (Black and Wiliam, 1998a), a summary of that work in a magazine for educators and policy makers (Black and Wiliam, 1998b), and an update to that work (Black and Wiliam, 2018). It is likely that no other educational intervention in the last 20 years has spawned greater research interest, commercial development, and practitioner appeal as has formative assessment.

Ironically, however, it is also likely there is no other concept introduced over the last two decades for which there is greater disagreement as to definition, greater disputes about efficacy, or greater deliberation about implementation.

The following sections of this entry provide the framework necessary for understanding and using formative assessment. First, a section called **Foundational concepts** lays the groundwork for the specialized practice of formative assessment by defining concepts from which it derives: inference, testing, and assessment. The next section traces the historical roots of formative assessment. Drawing on these foundations, section **Current concepts, confusions, and controversies** presents a contemporary, consensus definition of the concept, followed by a brief overview of some lingering confusions and controversies surrounding formative assessment. Section **Recent research on formative assessment** summarizes some of the most recent research on the practice of formative assessment, focusing exclusively on findings published since Black and Wiliam's (1998a) seminal review. A final section suggests conclusions and next steps for formative assessment research and practice.

Foundational concepts

As used in the field of education, formative assessment comprises a set of specialized procedures for gathering information about learning. To be sure, the preceding is not an adequate definition of formative assessment and a more complete definition will be provided in a subsequent section. In order to fully introduce a definition of formative assessment, however, it is important to locate that concept among a constellation of relevant, superordinate concepts within the broader field of educational measurement. In the following paragraphs, three related concepts are defined: inference, testing, and assessment.

Inference

A first essential concept that underlies all educational measurement is inference. The term *inference* can be defined as the process by which a tentative conclusion is reached based on incomplete, but pertinent information. Whenever any information is collected and conclusions drawn about student learning, inference takes place. This is the case because, typically, it is only practical to gather

a small sample of what might be observed regarding some characteristic relevant to learning in schools (e.g., mathematics problem solving, laboratory skills, essay writing ability, and so on). A teacher may seek to understand a student's proficiency at solving math problems, may want to evaluate a student's safety in the lab, or may need to assign grades in English composition. However, although it might be desirable for the teacher to observe the student solving dozens of math problems, conducting numerous lab experiments, or writing many essays, there is limited time for either the student to generate that information or the teacher to review it. Nonetheless, decisions about math homework, lab safety, and English Comp placements must still be made. Thus, it is nearly always the case that educators must come to those conclusions or make those decisions based on limited samples of information. The act of going from that smaller set of observations to a conclusion about a student's standing on some characteristic is called inference. The measurement specialist Benjamin Wright (1998) described inference this way:

I don't want to know which questions you answered correctly. I want to know how much ... you know. I need to leap from what I know and don't want to what I want but can't know. That's called inference.

In education, because complete information about student learning is never available, inference is a ubiquitous activity. From at least one perspective, it is also a desirable one; that is, gathering and correctly interpreting all possible information about a student would be prohibitive in terms of cost, time, and burden for all involved, so the practice of inference-making is an efficient process. Nonetheless, because they are based on limited samples of information, inferences should always be considered tentative and subject to reconsideration when additional information relevant to the conclusion becomes available.

Test

The second concept relevant to consideration of formative assessment is testing. Testing is a process often used by educators for gathering information—and making inferences—about student learning. A *test* is the tool used in that process for collecting a systematic sample of information about the characteristic of interest, such as a students' knowledge, skill, ability, interest, or other attribute which cannot be directly measured.

There are nearly an infinite number of configurations that would qualify as a test. A slate of multiple-choice questions administered under strict time and location requirements is only one test format—and certainly not the most informative for every information need. Surveys, performance tasks, essays, interviews, class discussions, direct observations—all of these and more qualify as tests to the extent that they are conducted and scored in systematic ways that foster accuracy and dependability of the conclusions based on their results.

Finally, because tests typically capture only a small portion of what could be gathered, and because the inferences based on test results can lead to consequential decisions about students, it is essential that the sample of information elicited by a test is carefully structured—whether the test is informal (e.g., the observation of clinical technique) or formal (e.g., a standardized college admissions test). Although beyond the scope of this entry, comprehensive information on the procedures and guidelines that characterize careful assessment development are treated extensively elsewhere (see, e.g., Lane et al., 2015).

Assessment

The final foundational concept presented here is that of assessment. Although often used synonymously with “test”, assessment is far broader in meaning. Prior to its use in education contexts, the term assessment was correctly applied in many other fields, including finance, medicine, and elsewhere. For example, when a retirement advisor conducts a financial assessment for a client, he or she examines a diverse set of variables (e.g., personal savings, stock holdings, IRAs and other investments, along with liabilities, retirement goals and timeline) to provide an overall summary and plan for the client. In the field of medicine, health care workers in an emergency department must rapidly perform an assessment of incoming patients. The assessments may take the form of measurements of blood pressure, temperature, respirations, reflexes, blood tests, x-rays, and other diagnostic procedures. The aim of gathering these disparate pieces of information (i.e., the results of various tests), is to produce a synthesis of the information in support of a conclusion about the present (i.e., financial position, diagnosis) and path for the future (i.e., retirement plan, treatment).

In short, *assessment* is the process of collecting and synthesizing many samples of information—that is, many tests—toward a specific purpose. In the context of education, Cizek (1997, p. 10) has defined assessment as “the planned process of gathering and synthesizing information relevant to the purposes of: discovering and documenting students' strengths and weaknesses; planning and enhancing instruction; or evaluating and making decisions about students” (Cizek, 1997, p. 10).

The term “assessment” is used accurately in education to describe the work of Individualized Educational Planning (IEP) team meetings. Such meetings might involve the teacher of a student relating observations of the student's in-class achievement and behavior, the parents of the student providing insights on the student's at home experiences, the school psychologist explaining the results of cognitive testing that was performed, and so on. Each of these contributions can be thought of as providing “test” information, because each is a sample of information about a student. The assessment involves aggregating all of the diverse sources of information, arriving at some tentative conclusions about what is happening for the student and developing tentative plans regarding appropriate placements, interventions, and supports.

Although each term has a specific, technical meaning that warranted consideration as conceptual background, it is likely that the terms “test” and “assessment” have become so pervasively used as synonyms that insistence technical accuracy would serve no

purpose. Indeed, even as regards the topic of this entry, although it would often be more conceptually accurate to refer to a “formative test”, the more familiar “formative assessment” label will be adopted.

Origins of formative assessment

The origin of the term formative assessment in education contexts has been traced by Cizek (2010), who observed that the roots of formative assessment can be traced to a chapter by Michael Scriven that appeared in a monograph series produced by the American Educational Research Association (AERA). In that source, Scriven (1967) coined the term, *formative evaluation*. Given the volume’s focus on curriculum evaluation, it is unsurprising that Scriven’s use of the term was in the context of program evaluation—in particular, evaluation of the effectiveness of school-based curricula and other educational program features and their effects on student achievement.

The concept of formative evaluation gained greater attention when it was popularized via the influential volume by Benjamin Bloom and his associates entitled the *Handbook of Formative and Summative Evaluation of Student Learning* (Bloom et al., 1971). Although that volume is perhaps most widely-known for its treatment of taxonomies of educational objectives introduced previously by Bloom (1956), it can also be recognized as providing one of the first explications of two concepts—formative and summative evaluation. Bloom et al. (1971) defined the term “summative” in way that subsumed, but went beyond Scriven’s comparatively narrower focus on evaluation of programs, extending the use of the term to the assessment of learning and to individual students:

We have chosen the term “summative evaluation” to indicate the type of evaluation used *at the end* of a term, course, or program for purposes of grading, certification, evaluation of progress, or research on the effectiveness of a curriculum, course of study, or educational plan [T]he essential characteristic of summative evaluation is that a judgment is made about the student, teacher, or curriculum with regard to the effectiveness of learning or instruction *after* the learning or instruction has taken place (p. 117, emphasis added).

Bloom et al. (1971) then distinguished summative evaluations—arguably the most prominent assessments at that time—from evaluations intended to serve a different purpose. According to Bloom et al. “formative evaluation is for us the use of systematic evaluation *in the process* of curriculum construction, teaching and learning *for the purpose of improving* any of these three processes” (1971, p. 118, emphasis added). This essential difference in the nature of summative and formative activities has been expressed pithily by Stake who noted that: “When the cook tastes the soup, that’s formative: When the guests taste the soup, that’s summative” (quoted in Scriven, 1991, p. 169).

Although some residual connection between the concept of “formative” and the field of program evaluation is still apparent in Bloom’s work, it is equally clear that Bloom’s expanded definition foreshadowed the linkages and uses of a formative approach for the assessment of individual students and for improving teaching and learning. According to Grant and Gareis (2014), an evolution in usage from *formative evaluation* to *formative assessment* began in the 1980s as educational researchers shifted focus from program evaluation to student learning and achievement. Importantly, the early distinction made by Bloom between summative (connoting an activity occurring “at the end” for purposes of assigning final grades and so on) and formative (connoting activity “in the process” for purposes of changing or improving the process) is now well-established. And, since the 1980s, formative assessment in educational settings has flourished and become a specialized field of academic study and school-based practice.

Current concepts, confusions, and controversies

Clear, consensus conceptualizations of formative assessment have been somewhat elusive. A recent feature on formative assessment in *Education Week* began by noting that “If you ask five teachers what formative assessment is, you’re likely to get five different answers” (Gerwertz, 2015, p. S2). Differences in definition are not only held by educators, but also evident in scholarly and professional writing on the topic. For example, in their book on formative assessment, Cizek et al. (2019) provide a table listing 10 varying definitions of formative assessment published in the professional literature since 2006. In a similar vein, a recent commentary on formative assessment theory and practice noted that:

The literature base on using formative assessment for instructional and intervention decisions is formidable, but the history of the practice of formative assessment is spotty. Even with the pressures of high-stakes accountability, its definition is fuzzy, its adoption is inconsistent, and the prognosis for future use is questionable.

Dorn (2010, p. 325).

One particularly troubling aspect of popular notions of student formative assessment is that the term has been used to describe a diverse array of products and practices ranging from commercially-produced, multiple-choice item banks from which teacher can create customized quizzes, benchmark tests, and drill materials, to guidelines for student creation of rubrics to evaluate drafts of their writing (and for assessing and providing feedback on the writing of their peers).

Accordingly, one difficulty in defining formative assessment is that whereas the preceding activities *might* have a formative aspect, they might also be conducted in such a way as to provide little, if any, “in process” information for adjusting or improving learning or instruction by teachers, and typically scant information that is truly formative for students. Although several assessment specialists have warned against accepting anything *labeled* “formative” as actually fulfilling a formative purpose (see, e.g., [Popham, 2006](#); [Shepard, 2005](#)), the concern persists. Thus, clarity about what formative assessment is—as well as what it is not—is essential. The following sections focus on formative assessment for students. Three distinguishing features of formative assessment are described and a consensus definition of the term is provided.

Distinguishing characteristic 1—primary purpose

A first distinguishing feature of contemporary definitions of formative assessment is the centrality of purpose. Whereas both summative and formative assessments share the aim of gathering information about student learning, the *reason* for gathering the information and how it will be used are key considerations in distinguishing between summative and formative activities. A summative assessment is any system of information gathering that is administered at the end of some unit of instruction (e.g., unit, semester, school year) and has as its primary purpose to categorize, label, or otherwise produce an evaluative conclusion about the merit, value, achievement or performance of a student or system. That is, the primary purpose of summative assessment is obtain a measurement of achievement to be used in decision making, assigning grades, awarding or denying a diploma, allocating compensation, or similar actions. In contrast, formative assessment describes information gathering that occurs “midstream” or in the course of some unit of instruction and has as its primary at least one of the following: to identify the test taker’s strengths and weaknesses, to assist educators in the planning of subsequent instruction, to aid students in guiding their own learning, revising their work, gaining self-evaluation skills, or fostering increased autonomy and responsibility for their learning. This distinction in primary purpose has, in other works, been described as a difference of assessment *of* learning versus assessment *for* learning ([Assessment Reform Group, 1999](#); [Stiggins, 2005](#)).

Although this distinction between summative and formative assessment regarding primary purpose is essential, it is not exclusive. That is, it is possible for assessments with primarily a summative purpose to serve a secondary, formative purpose; likewise, it is possible for assessments intended primarily to provide formative information to serve a secondary, summative purpose.¹

Distinguishing characteristic 2—timing and feedback

As is perhaps apparent from the distinction made previously between assessment *for* learning and assessment *of* learning, the second distinguishing difference between formative and summative assessment involves the timing of the activity. Formative assessment occurs during instructional activities. Feedback from formative assessment activities occurs during learning and points to areas of success, progress toward goals, and areas for improvement. Summative assessment occurs at the end of an instructional period (e.g., class, unit, semester), after instruction has taken place. Whereas corrective feedback is a hallmark of formative assessment, summative assessment may not be accompanied by any corrective feedback on performance other than an evaluative indicator such as a grade, award, or placement.

Distinguishing characteristic 3—student involvement

The final distinguishing characteristic of note regarding formative assessment is that of student involvement. This characteristic is perhaps the least well understood, accepted, or enacted, or all the features of formative assessment—and yet it is the characteristic that may be primarily responsible for the efficacy of the practice.

By “student involvement” we refer to several related aspects. First, effective formative assessment entails empowering students to articulate their own learning goals or, at minimum, aiding them in understanding the goals set for them, where they stand currently in relation to those goals, and what next steps will enable them to make progress toward the goals. Effective formative assessment provides explicit assistance to students in developing meta-cognitive skills and self-regulation of their own learning (see e.g., [Zimmerman, 1986](#)).

Second, effective formative assessment engages students via self-assessment. At minimum, formative assessment affords students opportunities to understand the evaluative norms applied to work in a discipline, to internalize those norms, and to apply them accurately to the work they produce ([Andrade, 2010, 2019](#)). At times, this kind of evaluation may be broadened to include peer evaluation, where trusted, critical colleagues who also have a good grasp of the evaluative norms can provide feedback on each other’s work. And, self-assessment may be a particularly powerful formative assessment strategy when students actively take part in the development of the specific evaluation guidelines (e.g., grading frameworks, scoring rubrics, etc.) that will be applied to their work. A good deal of research has taken place on this topic across diverse areas such as writing and the performing and creative arts (see, e.g., [Andrade, 2019](#); [Andrade et al., 2008](#)).

¹Although, see [Cizek et al. \(2019\)](#) for cautions on using formative assessments for summative purposes.

Formative assessment: a definition

Given these distinguishing characteristics, it is possible to derive a definition of formative assessment that is broad enough to cover assessment information for use by teachers and by students, and to incorporate the elements most likely to produce the hoped-for efficacy of formative assessment practice. The following definition is modified only slightly from that provided in [Cizek et al. \(2019\)](#):

Formative assessment comprises a diverse array of generic and discipline-specific information gathering practices, ideally nested within a planned assessment system, the primary aim of which is to support teachers' and students' inferences about learning and conceptual organization, strengths, and diagnosis of weaknesses and areas for improvement. Formative assessment yields information that educators can use in instructional planning and students can use in deepening their understandings, improving their achievement, taking responsibility for, and self-regulating their learning.

The preceding definition not only addresses the three distinguishing characteristics noted earlier, but foreshadows an emerging characteristic of formative assessment: the notion that formative assessment is not comprised exclusively of generic practices, but also may take different forms depending on the specific disciplinary area in which the information gathering takes place. A number of scholars have begun to explore this topic (see e.g., [Andrade et al., 2019a,b](#); [Heritage and Wiley, 2020](#)). The specific aspects of formative assessment that may differ across disciplines is clearly an area of current and likely future research.

Recent research on formative assessment

In general, studies have found fairly consistent positive effects of formative assessment on student achievement ([Graham et al., 2015](#); [Kingston and Nash, 2011](#); [Klute et al., 2017](#); [Weurlander et al., 2012](#); [Wiliam, 2011](#)). But what accounts for the positive effects? Over the past decade, the three distinguishing characteristics of formative assessment have received increasing attention, as researchers have sought and gained refined notions of the mechanism of action for formative assessment. This section provides a review across the most recent 10 years of empirical studies that have explored the efficacy of formative assessment on student learning with special attention to the distinguishing characteristics identified in the previous section. In the following sections, four areas of contemporary research on formative assessment are reviewed: (1) critical features of effective formative assessment; (2) the relationship between formative assessment and self-regulated learning, (3) discipline-specific aspects of formative assessment; and (4) technology-enhanced formative assessment.

Critical features of effective formative assessment

A number of research reviews and opinion papers have discussed the efficacy of formative assessment. Although a wealth of information already exists regarding the features of effective formative assessment generally, knowledge about what constitutes effective formative assessment continues to expand. Over the past decade, new research evidence has been emerging about several features that have gained empirical support for the design and implementation of effective formative assessment. These include: clearly defined purpose, design, and format of formative assessment in different classroom contexts ([Lane et al., 2019](#)); the use of explicit criteria for students to use in planning, monitoring, and evaluating own learning ([Panadero and Romero, 2014](#); [Jonsson, 2014](#)); the instructor's role in modeling self-assessment and the use of rubrics and descriptive feedback ([Panadero and Jonsson, 2013](#); [Panadero et al., 2016](#)); and the use of positive feedback under learner-centered assessment methods ([Pereira et al., 2016](#)). Finally, [Black and Wiliam \(2018\)](#) substantially extended their earlier work on formative assessment by locating it within a broader theoretical perspective. (See also a history and introduction to their work by [Hopfenbeck, 2018](#)). Black and Wiliam propose a model in which the design of learning activities and assessments is influenced by three factors: (1) relevant educational theories (e.g., learning theory, pedagogical theory); (2) the specific academic discipline under consideration; and (3) broader educational, political, cultural, and social contexts. They conclude by identifying and discussing issues of policy, practice, and implementation related to formative assessment.

The relationship between formative assessment and self-regulated learning

A recent development has been the exploration of linkages between formative assessment and self-regulated learning (SRL). The theoretical association between these concepts has been addressed by a number of researchers ([Andrade and Brookhart, 2020](#); [Clark, 2012](#); [Panadero et al., 2018](#)). Most recently, a special issue *Assessment in Education: Principles, Policy & Practice* brought together not only the emerging theoretical work, but also included some contemporary empirical studies (see [Brandmo et al., 2020](#)).

There is an abundance of literature exploring the theoretical relationship between formative assessment and SRL compared to the quantity of empirical studies of that relationship (see, e.g., [Panadero et al., 2016, 2017](#)). As formative assessment has also sometimes been called *assessment for learning*, studies using that concept were also included and analyzed.

An overall relationship between formative assessment and SRL was detected: Formative assessment and SRL are closely related, as both formative assessment and SRL "are oriented toward assessing own performance" ([Panadero et al., 2017](#), p. 77). The picture

is somewhat blurry, however, as studies have incorporated different types/tools/components of formative assessment (e.g., feedback, self-assessment, peer-assessment, the use of rubrics, scripts, etc.) to study their associations with the primary features of SRL (e.g., self-evaluation, self-monitoring, cyclical process). Although the research base is only beginning to form, there would seem to be reason for optimism about the relationship between formative assessment that includes an SRL emphasis. For example, a study by [Perry et al. \(2020\)](#) investigating writing performance of third graders found that students in classrooms with higher emphases on SRL and formative assessment demonstrated more sophisticated self-assessment, greater interest, higher levels of self-regulation, and higher quality writing than students in classrooms with lower emphasis on SRL and formative assessment.

A number of researchers have investigated the effectiveness of feedback as an instructional process through literature and meta-analytic work over a period of time ([Black and Wiliam, 1998a](#); [Hattie and Timperley, 2007](#); [Lipnevich et al., 2014](#); [Wiliam, 2011](#)). Overall, feedback—one of the distinguishing characteristics of formative assessment—has often been connected to SRL for measuring student achievement and improvement in learning, and numerous recent studies have provided empirical evidence of the effective use of feedback on student learning (see, e.g., [Evans, 2013](#); [Hudesman et al., 2013](#); [Lipnevich et al., 2014](#); [Pereira et al., 2016](#)). The mechanism of action for this effect was explored by [Orsmond et al. \(2013\)](#) who provided evidence that self-regulation of learning fosters effective learning by engaging students in their own learning processes. They also found that feedback, as an interactive process between learners and teachers, encourages students to make use of feedback for their self-regulation and improvement of learning.

Another distinguishing characteristic of formative assessment, *self-assessment*, has been studied by several researchers. Self-assessment facilitated by rubrics ([Panadero and Jonsson, 2013](#)) and scripts ([Panadero et al., 2012](#)) has been found to be effective for learning SRL skills (e.g., self-reflection, forethought), which subsequently leads to improvement in students' learning ([Panadero et al., 2016](#)). Another study conducted by [Panadero and Romero \(2014\)](#) concluded that self-assessment had statistically significant effects on all three phases of SRL—*forethought*, *performance*, and *self-reflection*. Therefore, it has been argued that promoting students' self-assessment should yield positive effects on student learning, as the process of self-assessment encourages students to plan, monitor, and evaluate their own learning ([Brown and Harris, 2013](#); [Jonsson, 2014](#)).

Discipline-specific formative assessment

The importance of accurately measuring the effect of formative assessment on student learning improvement has been raised by several researchers. For example, the meta-analysis of formative assessment conducted by [Kingston and Nash \(2011\)](#) was critiqued on methodological grounds (see [McMillan et al., 2013](#)), including the fact that the features of formative assessment are often non-transferable from one discipline to another. Increasingly over the last decade, researchers have delved into the differential approaches to formative assessment that account for its effects in specific disciplines (see [Andrade et al., 2019a,b](#); [Bennett, 2011](#); [Coffey et al., 2011](#); [Heritage and Wylie, 2020](#); [Hudesman et al., 2013](#); [Lane et al., 2019](#); [McMillan et al., 2013](#)).

Research on discipline-specific approaches to formative assessment is still developing, so it is too early to identify major trends. Further, most discipline-specific formative assessment research has taken place primarily in the K-12 contexts of science, mathematics, and writing. Studies from two representative disciplines are presented in the following paragraphs.

Science, as a discipline, has been considered to be difficult for both students and teachers as it requires complex and abstract concepts, problem-solving skills, and reconstruction of diverse concepts and links ([Miller, 1991](#)). In science education, helping students to clarify their current knowledge, develop conceptual change, and construct scientifically-accurate ideas are seen as essential components of science learning. Hence, formative assessment that incorporates focused feedback has been considered as one of the crucial elements of teaching and learning in science education ([Bell and Cowie, 2001](#)). However, science teachers have often encountered obstacles such as lack of understanding about formative assessment processes that can be effectively used as instructional practices and lack of science content knowledge that hinders teachers in helping students minimize the gap between new concepts and their prior knowledge ([Coffey et al., 2011](#); [Sabel et al., 2015](#)). As a result, when using formative assessment, science teachers often struggled with providing appropriate guidance and scaffolding to students because they often failed to communicate with students about students' reasoning and misconceptions. To resolve these issues, science education researchers have focused on designing formative assessment activities that can be embedded in the curriculum, and helping teachers improve communication with students ([Sabel et al., 2015](#)). They concluded that embedding formative assessment practices in curriculum with focused feedback can improve the quality of communication between teachers and students in terms of detecting and revising misconceptions, understanding and bridging new concepts and prior knowledge, and demonstrating problem-solving skills around abstract ideas ([Anderson et al., 2014](#); [Sabel et al., 2015](#)). Other researchers have found similar results: If embedded properly in the curriculum and if provided adequate support to students, formative assessment can enhance students' acquisition of science content knowledge and improve learning progressions ([Decristan et al., 2015](#); [Yin et al., 2013](#)).

A similar trend to that observed in science has been found in the domain of mathematics. Lucariello and colleagues conducted a study that administered formative assessment to middle school students to diagnose their misconceptions when learning algebra ([Lucariello et al., 2014](#)). Their formative assessment successfully identified students' misconceptions and how those misconceptions and reasoning errors arose. Their findings enabled teachers to gain insightful understandings about students' misunderstandings which can be later led to developing differentiated instruction for their students. Furthermore, several other studies in mathematics have revealed that diagnosing misconceptions alone is likely insufficient for facilitating the greatest learning gains from formative assessment; specific teacher behaviors that build students' skills in self-regulating their own learning are critical ([Hudesman et al., 2013](#)).

Finally, recent research on formative assessment in the disciplines of science and mathematics has revealed another key element of effective formative assessment: teachers' content knowledge. This element comes into play as educators strive to comprehend students' conceptual understandings, abstract thinking processes, and problem-solving skills, and as teachers are required to craft specific, supportive, constructive, and content-grounded feedback to students.

As a result, across disciplines, recent work in the area of formative assessment has highlighted an increasing need for professional development training for teachers to enhance their understandings about student cognition, learning progressions, and self-regulation, and to enhance their skill in designing and conducting appropriate formative assessments (Hondrich et al., 2016; Lane et al., 2019; Sabel et al., 2015).

Technology-enhanced formative assessment

As technology has enabled shifts in pedagogy from face-to-face settings to online and blended learning, increasing attention is being paid to online and remote assessment. Consequently, in conjunction with increased emphasis on well-designed technology-supported learning environments, some research has begun to focus on the design and effects of technology-enhanced formative assessment (Hwang and Chang, 2011).

The conclusion that formative assessment provides rich information about student learning, progress, and achievement, and the importance of instructional feedback on students' problem-based learning in diverse learning settings (e.g., blended settings, online learning, virtual learning environments) is a key finding in research reviews provided by Spector et al. (2016) and Russell (2010). For example, Russell identified ways in which technology-enhanced formative assessment can be supportive and effective toward students' achievement of their learning goals by: monitoring and reviewing student learning progress; identifying weaknesses, strengths, and misconceptions; providing rapid and timely feedback; and tracing learning progress and collecting information for guiding future learning.

In addition, individual empirical studies have investigated the use of technology-enhanced formative assessment and found that these four functions independently and collectively impacted student achievements under various contexts. For example, Dalby and Swan (2019) applied iPad-based formative assessment to year 10 students in mathematics classes. Results indicated that the use of iPad actually benefitted both teachers and students by providing timely feedback, by facilitating tracking of student learning, and by quickly identifying common misconceptions.

Beyond the direct impact of technology-enhanced formative assessment on academic achievement, studies have shown the indirect effects of technology-enhanced formative assessment on factors associated with academic improvement such as student attitude toward learning goals, classroom behavior, and other variables. One example can be found from the study conducted by Tempelaar et al. (2018). These authors investigated relationships between trace data generated by formative assessment using e-tutorials and learning dispositions (a term that refers to individual learner characteristics including affective, behavioral, and cognitive that influence every learning process). The authors reported that the formative assessments contributed to better academic performance of undergraduate students in mathematics and statistics classes. The e-tutorials constituted rapid and important feedback to individual student, provided procedural records of student learning, and anticipated learning for further improvement. In a similar vein, Chen et al. (2019) conducted a study that implemented tablet-based formative assessment system to improve student academic performance by enhancing positive classroom behaviors. The authors concluded that the tablet-based classroom behavior management system (CBMS) positively influenced student's positive classroom behaviors across both short- and long-term periods. The use of tablet-based CBMS also resulted in improved learning outcomes by providing frequent and real-time feedback along with the recorded behavior ratings on a daily basis. The authors suggest that this was made possible because the technology significantly reduced the administrative load of the teachers and enhanced student engagement in the classroom activities.

Conclusions and future of formative assessment

This history of formative assessment can be traced to the introduction by Scriven (1967) of the concept of formative evaluation. In the late 1990s, and after decades of implementation of summative assessment to leverage achievement gains, formative assessment began to receive increased attention because of its potential power to improve teaching and learning.

At present, however, formative assessment is a concept that does not have a universally agreed-upon meaning. In this chapter, we reviewed some key elements of effective formative assessment, including:

- * encouraging student responsibility for their learning;
- * communicating clear, specific learning goals;
- * providing timely, detailed, and constructive feedback to students;
- * instructing students in how to conduct accurate self-assessment of progress toward learning goals; and
- * teacher use of assessment information to enhance instructional planning and practices.

As is evident from this list, authentic formative assessment comprises more than merely more frequent, shorter, or lower-stakes summative assessments such as quizzes, classroom tests, or homework assignments. Rather, formative assessment that has the greatest potential to increase learning is one part of, and implemented within, an educational curriculum, instruction, and assessment ecosystem in which all aspects have distinct purposes and are thoroughly integrated. Although the earliest estimates of the efficacy of

formative assessment may have been overstated, research on formative assessment—if implemented with fidelity with respect to its key elements—has demonstrated that it can be a powerful classroom-based intervention.

Current research on formative assessment has largely focused on four topics including: its relationship to self-regulated learning, formative assessment practices specific to individual disciplines, crucial features that account for the efficacy of formative assessment, and the role of evolving technologies. Overall, the themes evident in current research suggest that formative assessment should be designed to encourage teachers and students to actively communicate about the processes of teaching and learning. For teachers, recognizing whether they are accurately diagnosing students' understanding and misconceptions, and whether they are providing appropriate feedback and instructional guidance to students should be prioritized. For students, taking the lead of their learning and initiating improvement of their learning should be prioritized. By doing so, formative assessment can be most useful and beneficial to both teachers and students.

For the future, it is likely that there will be continuing research emphasis on these four topics. However, in addition, it seems appropriate to recommend that not only additional academic study of formative assessment should continue, but that a new effort must be undertaken. Toward the goals of addressing lingering confusion about the nature of formative assessment, providing educators with effective tools for enhancing learning and instruction, and enhancing the utility of assessment information for guiding educational policy and practice, it is recommended that:

- (1) coursework in assessment be required as part of educator preparation programs;
- (2) pre-service preparation experiences for educators (in particular, teachers and principals) with substantial emphasis on effective formative assessment practices should be developed, tested, and included within assessment-focused coursework or in discipline-specific pedagogical training;
- (3) professional development materials covering contemporary best practices in formative assessment should be developed, tested, and provided for in-service educators; and
- (4) educational leaders at state and local levels should review and redesign, as needed, coherent curriculum/assessment/instruction ecosystems where all components (especially diagnostic, formative, and summative components) provide non-redundant, relevant, and actionable information in the service of classroom decision making and educational policy making.

Formative assessment was initially introduced as a strategy for teachers to gain information about student learning and to use that information to refine instruction. Surely that is still a valuable goal to be pursued. However, ultimately, it is likely the greatest potential of formative assessment lies in its power to change students' own abilities to establish, monitor, and accurately evaluate their learning. To the extent that they are provided with the metacognitive and self-regulation skills that are a key part of formative assessment practice, they will not only be well positioned for greater achievement in their elementary and secondary school years, but well prepared with the self-guided, lifelong-learning orientation that will serve them in future educational and career contexts.

References

- Anderson, J.L., Ellis, J.P., Jones, A.M., 2014. Understanding early elementary children's conceptual knowledge of plant structure and function through drawings. *CBE Life Sci. Educ.* 13 (3), 375–386.
- Andrade, H.L., Brookhart, S.M., 2020. Classroom assessment as the co-regulation of learning. *Assess. Educ. Princ. Pol. Pract.* 27 (4), 350–372. <https://doi.org/10.1080/0969594X.2019.1571992>.
- Andrade, H., Du, Y., Wang, X., 2008. Putting rubrics to the test: the effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educ. Meas.* 27, 3–13. <https://doi.org/10.1111/j.1745-3992.2008.00118.x>.
- Andrade, H.L., Bennett, R.E., Cizek, G.J., 2019a. *Handbook of Formative Assessment in the Disciplines*. Routledge, New York, NY.
- Andrade, H.L., Heffernan, J., Palma, M.E., 2019b. Formative assessment in the arts. In: Andrade, H.L., Bennett, R.E., Cizek, G.J. (Eds.), *Handbook of Formative Assessment in the Disciplines*. Routledge, New York, pp. 126–145.
- Andrade, H., 2010. Students as the definitive source of formative assessment: academic self-assessment and the self-regulation of learning. In: Cizek, G.J., Andrade, H. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, NY, pp. 90–105.
- Andrade, H., 2019. A critical review of research on student self-assessment. *Front. Educ.* 4 (87), 1–13. <https://doi.org/10.3389/feduc.2019.00087>.
- Assessment Reform Group, 1999. *Assessment for Learning: Beyond the Black Box*. School of Education, Cambridge University, Cambridge, UK.
- Bell, B., Cowie, B., 2001. The characteristics of formative assessment in science education. *Sci. Educ.* 85 (5), 536–553.
- Bennett, R.E., 2011. Formative assessment: a critical review. *Assess. Educ. Princ. Pol. Pract.* 18 (1), 5–25. <https://doi.org/10.1080/03075071003731135>.
- Black, P., William, D., 1998a. Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5 (1), 7–74.
- Black, P., William, D., 1998b. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan* 80 (2), 139–144, 146–148.
- Black, P., William, D., 2018. Classroom assessment and pedagogy. *Assess. Educ. Princ. Pol. Pract.* 25 (6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>.
- Bloom, B.S. (Ed.), 1956. *Taxonomy of educational objectives: The classification of educational goals*. McKay, New York.
- Bloom, B.S., Hastings, J.T., Madaus, G.F. (Eds.), 1971. *Handbook of formative and summative evaluation of student learning*. McGraw-Hill, New York.
- Brandmo, C., Panadero, E., Hopfenbeck, T.N., 2020. Bridging classroom assessment and self-regulated learning. *Assess. Educ. Princ. Pol. Pract.* 27 (4), 319–331. <https://doi.org/10.1080/0969594X.2020.1803589>.
- Brown, G.T.L., Harris, L.R., 2013. Student self-assessment. In: McMillan, J. (Ed.), *The SAGE Handbook of Research on Classroom Assessment*. SAGE, Thousand Oaks, CA, pp. 367–393.
- Chen, W., Gu, X., Wong, L.H., 2019. To click or not to click: effectiveness of rating classroom behaviors on academic achievement with tablets. *Br. J. Educ. Technol.* 50 (1), 440–455.
- Cizek, G.J., 1997. Learning, achievement, and assessment: Constructs at a crossroads. In: Phye, G.D. (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment*. Academic Press, New York, NY, pp. 1–33.
- Cizek, G.J., Andrade, H.L., Bennett, R.E., 2019. Formative assessment: history, definition, and progress. In: Andrade, H., Bennett, R., Cizek, G.J. (Eds.), *Handbook of Formative Assessment in the Disciplines*. Routledge/Taylor and Francis, New York, pp. 1–19.

- Cizek, G.J., 2010. An introduction to formative assessment: history, characteristics, and challenges. In: Andrade, H., Cizek, G. (Eds.), *Handbook of Formative Assessment*. Taylor and Francis, New York, NY, pp. 3–17.
- Clark, I., 2012. Formative assessment: assessment is for self-regulated learning. *Educ. Psychol. Rev.* 24 (2), 205–249. <https://doi.org/10.1007/s10648-011-9191-6>.
- Coffey, J.E., Hammer, D., Levin, D.M., Grant, T., 2011. The missing disciplinary substance of formative assessment. *J. Res. Sci. Teach.* 48 (10), 1109–1136.
- Dalby, D., Swan, M., 2019. Using digital technology to enhance formative assessment in mathematics classrooms. *Br. J. Educ. Technol.* 50 (2), 832–845.
- Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., et al., 2015. Embedded formative assessment and classroom process quality: how do they interact in promoting science understanding? *Am. Educ. Res. J.* 52 (6), 1133–1159.
- Dorn, S., 2010. The political dilemmas of formative assessment. *Except. Child.* 76 (3), 325–337.
- Evans, C., 2013. Making sense of assessment feedback in higher education. *Rev. Educ. Res.* 83 (1), 70–120.
- Gewertz, C., 2015. Searching for clarity on formative assessment. *Educ. Week* 35 (12), S2.
- Graham, S., Hebert, M., Harris, K.R., 2015. Formative assessment and writing: a meta-analysis. *Elem. Sch. J.* 115 (4), 523–547.
- Grant, L.W., Gareis, C.R., 2014. Formative assessment. In: *Oxford Bibliographies in Education*. <https://doi.org/10.1093/obo/9780199756810-0062>.
- Hattie, J., Timperley, H.S., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Heritage, M., Wylie, E.C., 2020. *Formative Assessment in the Disciplines: Framing a Continuum of Professional Learning*. Harvard University Press, Cambridge, MA.
- Hondrich, A.L., Hertel, S., Adl-Amini, K., Klieme, E., 2016. Implementing curriculum-embedded formative assessment in primary school science classrooms. *Assess. Educ. Princ. Pol. Pract.* 23 (3), 353–376.
- Hopfenbeck, T.N., 2018. Classroom assessment, pedagogy and learning—twenty years after Black and Wiliam 1998. *Assess. Educ. Princ. Pol. Pract.* 25 (6), 545–550. <https://doi.org/10.1080/0969594X.2018.1553695>.
- Hudson, J., Crosby, S., Flugman, B., Issac, S., Everson, H., Clay, D.B., 2013. Using formative assessment and metacognition to improve student achievement. *J. Dev. Educ.* 37 (1), 2–13.
- Hwang, G.J., Chang, H.F., 2011. A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Comput. Educ.* 56 (4), 1023–1031.
- Jonsson, A., 2014. Rubrics as a way of providing transparency in assessment. *Assess. Eval. High Educ.* 39 (7), 840–852. <https://doi.org/10.1080/02602938.2013.875117>.
- Kingston, N., Nash, B., 2011. Formative assessment: a meta-analysis and a call for research. *Educ. Meas.* 30 (4), 28–37.
- Klute, M., Aphthor, H., Harlacher, J., Reale, M., 2017. *Formative Assessment and Elementary School Student Academic Achievement: A Review of the Evidence (REL 2017–259)*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central, Washington, DC. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Lane, S., Raymond, M.R., Haladyna, T.M., 2015. *Handbook of Test Development*, second ed. Taylor and Francis, New York, NY.
- Lane, R., Parrila, R., Bower, M., Bull, R., Cavanagh, M., Forbes, A., Jones, T., Leaper, D., Khosronejad, M., Pellicano, L., Powell, S., Ryan, M., Skrebneva, I., 2019. *Literature Review: Formative Assessment Evidence and Practice*. Australian Institute for Teaching and School Leadership (AITSL), Melbourne.
- Lipnevich, A.A., McCallen, L.N., Miles, K.P., Smith, J.K., 2014. Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instr. Sci.* 42 (4), 539–559.
- Lucariello, J., Tine, M.T., Ganley, C.M., 2014. A formative assessment of students' algebraic variable misconceptions. *J. Math. Behav.* 33, 30–41.
- McMillan, J.H., Venable, J.C., Varier, D., 2013. Studies of the effect of formative assessment on student achievement: so much more is needed. *Pract. Assess. Res. Eval.* 18 (1), 1–15. <https://doi.org/10.7275/tmwm-7792>.
- Millar, R., 1991. Why is science hard to learn? *J. Comput. Assist. Learn.* 7 (2), 66–74. <https://doi.org/10.1111/j.1365-2729.1991.tb00229.x>.
- Orsmond, P., Maw, S.J., Park, J.R., Gomez, S., Crook, A.C., 2013. Moving feedback forward: theory to practice. *Assess. Eval. High Educ.* 38 (2), 240–252.
- Panadero, E., Jonsson, A., 2013. The use of scoring rubrics for formative assessment purposes revisited: a review. *Educ. Res. Rev.* 19, 129–144.
- Panadero, E., Romero, M., 2014. To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assess. Educ. Princ. Pol. Pract.* 21 (2), 133–148.
- Panadero, E., Tapia, J.A., Huertas, J.A., 2012. Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learn. Individ. Differ.* 22 (6), 806–813.
- Panadero, E., Jonsson, A., Stribos, J.W., 2016. Scaffolding self-regulated learning through self-assessment and peer assessment: guidelines for classroom implementation. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer International Publishing, Cham, Switzerland, pp. 311–326.
- Panadero, E., Jonsson, A., Botella, J., 2017. Effects of self-assessment on self-regulated learning and self-efficacy: four meta-analyses. *Educ. Res. Rev.* 22, 74–98.
- Panadero, E., Andrade, H., Brookhart, S., 2018. Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *Aust. Educ. Res.* 45 (1), 13–31.
- Pereira, D., Flores, M.A., Simão, A.M.V., Barros, A., 2016. Effectiveness and relevance of feedback in higher education: a study of undergraduate students. *Stud. Educ. Eval.* 49, 7–14.
- Perry, N.E., Lisaingo, S., Yee, N., Parent, N., Wan, X., Muis, K., 2020. Collaborating with teachers to design and implement assessments for self-regulated learning in the context of authentic classroom writing tasks. *Assess. Educ. Princ. Pol. Pract.* 27 (4), 416–443. <https://doi.org/10.1080/0969594X.2020.1801576>.
- Popham, W.J., 2006. Phony formative assessments: buyer beware! *Educ. Leader* 64 (3), 86–87.
- Russell, M., 2010. Technology-aided formative assessment and learning: new developments and applications. In: Andrade, H.L., Cizek, G.J. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, NY, pp. 125–138.
- Sabel, J.L., Forbes, C.T., Zangori, L., 2015. Promoting prospective elementary teachers' learning to use formative assessment for life science instruction. *J. Sci. Teach. Educ.* 26 (4), 419–445.
- Scriven, M., 1967. The methodology of evaluation. In: Tyler, R.W., Gagne, R.M., Scriven, M. (Eds.), *Perspectives of Curriculum Evaluation*. Rand McNally, Chicago, IL, pp. 39–83.
- Scriven, M. (Ed.), 1991. *Evaluation Thesaurus*, fourth ed. SAGE, Newbury Park, CA.
- Shepard, L.A., 2005. *Formative Assessment: Caveat Emptor*. Presentation at the 2005 ETS Invitational Conference on the Future of Assessment: Shaping Teaching and Learning. New York, NY. Retrieved from http://www.cpre.org/ccii/images/stories/ccii_pdfs/shepard%20formative%20assessment%20caveat%20emptor.pdf.
- Spector, J.M., Ifenthaler, D., Sampson, D., Yang, L., Mukama, E., Warusavitarana, A., Lokuge Dona, K., Eichhorn, K., Fluck, A., Huang, R., Bridges, S., Lu, J., Ren, Y., Gui, X., Deneen, C.C., San Diego, J., Gibson, D.C., 2016. Technology enhanced formative assessment for 21st century learning. *Educ. Technol. Soc.* 19 (3), 58–72.
- Stiggins, R.J., 2005. *Student-Involved Assessment for Learning*. Prentice Hall, New York, NY.
- Tempelaar, D., Rienties, B., Mittelmeier, J., Nguyen, Q., 2018. Student profiling in a dispositional learning analytics application using formative assessment. *Comput. Hum. Behav.* 78, 408–420.
- Weurlander, M., Söderberg, M., Scheja, M., Hult, H., Wernerson, A., 2012. Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assess. Eval. High Educ.* 37 (6), 747–760.
- Wiliam, D., 2011. What is assessment for learning? *Stud. Educ. Eval.* 37 (1), 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>.
- Wright, B.D., 1998. *Introduction to the Rasch model [video]*. University of Chicago, IL: MESA Press.
- Yin, Y., Tomita, M.K., Shavelson, R.J., 2013. Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *Int. J. Sci. Educ.* 36 (4), 531–552. <https://doi.org/10.1080/09500693.2013.787556>.
- Zimmerman, B.J., 1986. Becoming a self-regulated learner: which are the key subprocesses? *Contemp. Educ. Psychol.* 11, 307–313.

Assessing and learning, and learning to learn

Mary E. James, University of Cambridge, Faculty of Education, Cambridge, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	10
Theories of learning and implications for assessment	10
Learning as acquiring what is taught, and implications for assessment	11
Folk views of learning	11
Behaviorist views of learning	11
Implications for assessment	11
Learning as individual sense-making, and implications for assessment	12
Cognitive constructivist views of learning	12
Implications for assessment	13
Learning as building knowledge through social interaction, and implications for assessment	14
Sociocultural/situated views of learning	14
Implications for assessment	14
Issues of progression and generalizability in sociocultural approaches	15
Can different approaches to learning and assessment be combined? The importance of curriculum	16
Learning (how) to learn and implications for assessment	17
References	19

Introduction

This contribution is particularly concerned with the verbs “to assess” and “to learn” and the relationships between the processes they entail in theory and practice. The focus is first on learning and then on the implications for assessment practice. The rationale for this is that education, at whatever level, from early years to adult learning, needs to start from a clear conception of aims and purposes, and then proceed to create curricula and learning opportunities coherent with those goals. Ways of assessing the achievement of those goals (assessment of learning) or ways of assessing that contribute to learning (assessment for learning) also need to serve these educational purposes. Therefore, educational aims, curricula, pedagogical practices, conceptions of learning, and assessment all need to be aligned. Consideration of what learning is, in different contexts, is important for making good choices about assessment tools and practices.

The following sections identify a number of different views of learning and then explore their implications for assessment, taking account of the importance of the variety of aims and learning objectives in different sectors, domains and subject areas. The concept of “learning how to learn” is then explored, together with recent expressions in assessment practice.

The main sources for these inquiries are a series of publications on these themes by the author [James \(2006, 2008, 2012a\)](#) including an entry in the third edition of this encyclopedia ([James, 2010](#)).

Theories of learning and implications for assessment

The first point to make is that many assessments, such as ability and aptitude tests, are not underpinned by explicit theories of learning at all. They were developed to measure mental qualities, such as intelligence, which were believed to be innate traits. They were designed to predict capacity to learn and future success but were not thought to be the products of learning themselves. Indeed they were often designed to screen out the effects of learning, in order to identify candidates with potential rather than privileged access to resources, although it is debatable whether this can ever be wholly achieved. The uses of such assessments, now as in the past, are largely concerned with selection e.g., choosing among candidates those most likely to fulfill the requirements for courses, jobs or promotions, or those who would benefit from special needs support. This approach to assessment is sometimes referred to as “differential”. The technology associated with it is “psychometrics”—literally the “measurement of mind”. As the term suggests it draws heavily on psychology as the foundation discipline.

Psychometrics and cognitive sciences play important parts in the development of assessment and measurement, but other disciplinary perspectives have a role too. Approaches derived from social theory, anthropology, philosophy, and sociology also make a contribution, especially in cultures that value community, as in the East, over individualism, as in the West. For instance, in Russia there has long been debate about the relative importance of “upbringing”, a social good, and “education”, as an individual good. [Anna Sfard’s \(1998\)](#) seminal paper on metaphors of education also makes a key distinction between learning as (individual) acquisition and learning as (social) participation. But, crucially, she points to the dangers of choosing to adhere to just one of these positions.

These themes play out in the following discussions of learning and assessment, and learning to learn.

Chris Watkins (2003) differentiates three views of learning: learning as being taught (equivalent to Sfard's acquisition metaphor); learning as individual sense-making; and learning as building knowledge as part of doing things with others (equivalent to Sfard's participation metaphor). These three views can also be categorized according to their theoretical underpinnings, most usually as behaviorist, cognitive constructivist, and sociocultural or situated. Each of these three views of learning has different implications for assessment practice.

Learning as acquiring what is taught, and implications for assessment

Two different sets of assumptions underpin this view of learning but both share a preoccupation with what is taught, and how well knowledge, both declarative/propositional (knowing that) and procedural (knowing how), is absorbed by learners. Learning is regarded largely as the acquisition of what the teacher or trainer wants to teach and little attention is paid to whether, or how, this knowledge might be modified in the process. Thus learners are regarded as rather passive and success is judged by the extent to which they acquire the knowledge and reproduce the behaviors that the teacher intended.

Folk views of learning

The first set of assumptions is not based on any evidence-informed theory of learning, but rather a folk psychology that regards the brain or mind (the two are not always clearly distinguished) as a kind of vessel into which information is poured. Learning is assumed to have occurred when this knowledge is retained. Perhaps the most explicit expression of this view is found in Charles Dickens' novel, *Hard Times*, in which Gradgrind describes children's minds as like little pitchers to be filled full of facts. Although this is a caricature, there are still a lot of people who hold something like this view.

Of course, learning does involve memory. Eric Kandel (2006), who won a Nobel Prize for his research on memory, emphasizes how much memory matters to our sense of being human. Remembering things is fundamental. What is more questionable is that learners can, or should, always remember things in exactly the same way as taught by their teachers, and be assessed on this. We need not be against the memorization of facts and information, and the development of routines, if we nevertheless want to argue for more complex and nuanced views of learning and assessment.

Behaviorist views of learning

A second set of assumptions that underpin a view of learning as acquisition of what is taught is derived from research evidence within a tradition known as behaviorism. This perspective was derived from a concern to develop a proper science of human behavior based on observation and experiment. Mental processes are not directly observable so they were excluded from investigation.

The basic tenet of behaviorism is that learning can be viewed as conditioned response to external stimuli. In other words, if a particular stimulus produces a response, the recurrence of the stimulus, or something closely associated with it, will cause the response to be repeated. In Pavlov's dogs, the presentation of food caused them to salivate but, over time, after repeated occasions when the food was presented while a bell was rung, the ringing of a bell alone caused them to salivate. Applied to human learning, this theory has been used successfully in situations where a particular behavioral response is sought. Rewards "condition" desired responses, and punishments, or more commonly the withholding of rewards, extinguish undesired responses. Repetition of the stimulus makes the response habitual; hence the value of repeated practice. Eventually the response is so automatic that it can be produced at speed.

These ideas, applied to human learning, sound somewhat old fashioned today (although they are still the basis of successful training of dogs). However, while theory has moved on, practice stemming from it has persisted, and sometimes for good reason. There are times when it is legitimate to expect people to learn to behave in particular ways, instantly, and without worrying too much about the whys and wherefores. As every parent knows, there are times when a repeated instruction, the sharp rebuke, or non-specific praise elicits a behavior that keeps a child safe—from falling under a car, for instance. Furthermore, in a world where there is so much to know and understand, individuals have to make strategic judgments about what they want to understand in depth and what they will be content to know only "on the surface" or as a conditioned response. And this may vary from person to person and context to context.

As with folk psychology, rote learning is associated with the behaviorist approach but there are other features as well. For example, behaviorism assumes that a complex skill can be taught by being broken down into components, each of which can be taught separately then re-assembled. Another assumption is that if an idea or practice is common to many contexts, it is most economical to teach it separately so that it can then be applied in many situations. Moreover, it is considered best to learn basic facts and skills first, then build more sophisticated knowledge and practice on this foundation.

Implications for assessment

The assumptions and beliefs about learning, outlined above, lead to equivalent assumptions about the kinds of assessment practice that are most appropriate. James (2008), labeled this "first generation assessment":

- The learning of the individual is assessed; performance that is influenced by others is regarded as cheating.
- The focus is on performance under test conditions.

- Performance tends to be defined as demonstration of ability to recall facts and information, or to demonstrate skills.
- Tests or tasks are specially constructed for the purpose of assessment and are scheduled at the end of a course or sequence of learning; they are regarded as separate from learning.
- Preparation for tests may involve practice, for example by using past papers, but results on these will not count toward a score or grade.
- Tests are time-limited, on the assumption that speed of completion of the tasks indicates that the learning is well established.
- No access to materials is allowed other than those needed to do the immediate task because it is expected that knowledge can be recalled from memory, or skills performed.
- Questions or tasks for completion are organized in a hierarchy of difficulty, starting with the least difficult, often in the form of short response or multiple-choice items.
- Responses or performances are assessed as correct or incorrect and results from a test are aggregated to give a score enabling comparison with other learners (norm-referenced), against a standard (criterion-referenced) or against the learner's own previous performance (ipsative referenced). Sometimes these referencing systems are combined e.g., norm with criterion referencing, as in most summative assessments for the award of leaving certificates or credentials.
- It may be possible to infer areas for improvement (where incorrect responses were made) from returned marked scripts, if these are available, but rarely from global scores. Global scores only indicate that a learner has done well or should try to do better. However, tightly structured, interactive computer-based assessments can direct test-takers back to an easier level of performance, during an assessment, if they fail at a particular level. This is common in apps for foreign language learning.

The remarkable thing about this list is how familiar it has become. Indeed, many lay people, including the media, and some professionals and policy makers, take it for granted that these are characteristics of all good tests and assessments.

In their key text, [Pellegrino et al. \(2001\)](#) note:

Most current tests, and indeed many aspects of the science of educational measurement, have theoretical roots in the differential and behaviorist traditions. The more recent perspectives – the cognitive and situative – are not well reflected in traditional assessments but have influenced several recent innovations in the design and use of educational assessments (p. 60).

It is these alternative approaches, more congruent with cognitive and situated/sociocultural views about learning, that are discussed next.

Learning as individual sense-making, and implications for assessment

This second cluster of theories of learning also focus on the individual and the acquisition of knowledge and skill, but take a more active view of learning, recognizing that what is learned is rarely identical to what is taught. An emphasis on individual meaning making indicates a particular interest in cognition: the working of the mind as a function of the brain. Assessments, from this viewpoint, go beyond testing factual recall to tap into deeper understandings.

Cognitive constructivist views of learning

Behaviorists' denial of the importance of mental processes in learning provoked a reaction from psychologists, linguists, cognitive scientists, computer scientists and latterly, neuroscientists, who were quick to point out that human learning must, at the very least, involve thinking. All argued that learning is determined largely by what goes on in people's heads. More significantly, however, they argued that learning involves people in making sense of the world by building mental models of how the world works so that they can interpret—or make sense—of new information. Thus learning is not simply absorbing information but an active process of meaning making.

Recently, there has been much talk of personalized learning, however, according to cognitive constructivism, learning is by definition personalized because people will always make sense of their experiences in unique ways. No one person's mental models will be quite the same as those of others because they have been shaped by different prior experiences. These models provide the lenses through which we view new experiences and lead us to pay great attention to some features and perhaps ignore others. Generally this is useful because selecting phenomena, or aspects of phenomena, for our attention helps us to manage the sheer quantity of incoming sensory data. Of course, such frameworks can also blind us to the significance of new experiences—even lead us to adopt prejudiced views. In this context the task of instruction is not simply to impart information, or to provide practice for skill development, but to help learners evaluate established mental maps and adapt, extend or renew them so that they expand perception and deepen understanding.

The emphasis within this perspective is on the development of conceptual understanding, which is particularly important in the learning of academic subjects. Thus concepts, and the relationships among them, are important because these form the basis of our mental models (schema). The need to build such models is familiar in schools, from the use of spider diagrams to map topics in elementary schools, to concept mapping as a revision practice in high schools. The nodes in these diagrams are usually nouns and the links are verbs.

Such models are built using a range of mental processes. Acquiring knowledge is one such process but at a higher (or deeper) level are processes such as comprehension, application, analysis, synthesis and evaluation—categories derived from Benjamin Bloom's (1956) influential classification of cognitive objectives. The application of concepts in solving problems is particularly valued for its capacity to reveal understanding and to lead to the construction of new knowledge. Failures to solve problems can often be traced to misunderstandings of concepts or their relationships. For example, misunderstandings or misunderstanding of the nature and role of photosynthesis in plants can have profound consequences for solving problems of climate change.

The task for education, according to this view, is to help novices acquire the ways of thinking possessed by experts, notably the ways they organize knowledge so that it is more retrievable and useful. Characteristically experts solve problems more quickly, with "less search", because their models help them to know what is salient. They are also capable of reflecting upon and controlling their mental processes through the exercise of metacognitive processes—thinking about thinking.

Implications for assessment

Some of the implications for assessment practice, that flow from cognitive constructivist views of learning, are similar to those for the assessment of learning viewed as the acquisition of knowledge, but there are also significant differences in approach and emphasis. James (2008), labeled this "second generation assessment":

- The learning of the individual is assessed.
- But, the focus is on problem solving and understanding.
- Performance tends to be defined as demonstration of ability to apply conceptual frameworks to find solutions to problems, and to demonstrate cognitive skills.
- Tests and examinations are specially constructed for the purpose of assessment and are often scheduled at the end of a course or sequence of learning.
- Tasks may be extended in order to allow those being assessed to display the breadth and depth of their understanding, or to solve complex problems. Essays, concept maps, think-aloud protocols, open-ended assignments, projects, coursework or theses may be the preferred form of response.
- Assessments are, nevertheless, often time-limited, on the assumption that speed of completion of a task indicates that acquisition of conceptual models is secure and promotes understanding and problem solving with less search.
- Some limited access to materials, such as the text of novels or poetry in literature examinations, may be permitted because the assessment is less a test of memory than an assessment of understanding.
- Responses to the tasks or problems are assessed according to specified criteria which involve judgment on the part of the assessor because there will rarely be one correct answer. Judges therefore need to be trained to ensure that their assessments are in agreement with those of others. They may be given access to marking schemes or scoring rubrics, and exemplars (or model answers) for this purpose. This will probably also require the setting up of moderation or standardization procedures so that assessors, examiners or verifiers can share and discuss their judgments of a selection of work, thus developing a conception of appropriate standards in a progression from novice to expert.
- The notion of a progression often implies a ladder of scores, levels or grades based on some norm-referenced assumptions.
- Another assumption is that improvement comes from efforts to close the gap between novice and expert performance. This need not be interpreted as a linear progression but could be regarded as development toward a horizon of possibilities.
- Areas for improvement can be inferred from misunderstandings or failures to find viable solutions to problems. These may indicate flawed conceptual understanding or poor frameworks that need to be revisited, unpicked and reconstructed. This implies that certain ideas may need to be unlearned or re-learned in formative classroom processes, or that new cognitive skills need to be developed.

This approach to assessment remains focused on the acquisition and processing of knowledge, as something possessed by individuals. For this reason, some tests or examinations combine elements of both behaviorist and constructivist approaches. Examples might be tests that begin with short answer questions based on factual recall, progressing to more extended, problem-solving tasks or essay-type questions.

Pellegrino et al. (2001) provide many examples from mathematics and science of how the cognitive constructivist perspective has been incorporated into the recent design of assessments. Particular emphasis is given to ways of detecting how people organize information in schema in long term memory, how they use this knowledge to solve problems, how they develop metacognitive skills, and how they engage in communicative practices appropriate to subject domains. They advise that assessment methods can be expanded beyond traditional tests by drawing on the ways in which academic research is conducted to investigate learning; computerized systems are mentioned including intelligent tutoring, time-reaction monitoring, and simulations. However, the social dimension is largely interpreted as a context for individual learning, rather than integral and inseparable.

A third approach to assessment practice (below) makes a more radical shift—away from seeing learning as acquisition of knowledge and/or understanding to learning as participation in social practices.

Learning as building knowledge through social interaction, and implications for assessment

A third view of learning is often called the sociocultural, or situated. It is sometimes regarded as a new development. However, [Bredo \(1997\)](#) traces its intellectual origins back to the work of William James, John Dewey and George Herbert Mead at the beginning of the twentieth century. Its origins therefore pre-date both behaviorism and cognitive constructivism.

Sociocultural/situated views of learning

A key principle of sociocultural theory is that learning involves both thought and action in context and is particularly influenced by the interactions among these phenomena. This may explain why it is often seen as more complex than behaviorism, which concentrates on behavior, or cognitivism that focuses on thought processes. The interactionist approach of Dewey and Mead viewed human development as a transaction between the individual and the environment, between actor and structure. This had something in common with the development of cultural psychology in Russia, associated with [Vygotsky \(1978\)](#), who was himself influenced by William James, and possibly met Dewey. This transatlantic conversation continued as Vygotsky's thinking subsequently influenced theorists such as [Bruner \(1996\)](#) in the US and [Engeström \(1999\)](#) in Finland. Other key theorists who regard individual learning as situated in the social environment include Barbara [Rogoff \(1990\)](#), Jean Lave and Etienne Wenger ([Lave and Wenger, 1991](#); [Wenger, 1998](#)) who draw on anthropological work to characterize learning as cognitive apprenticeship in communities of practice.

According to this perspective, thinking is conducted through actions that alter the situation and the situation changes the thinking; the two constantly interact. Especially important is the notion that learning is a mediated activity in which cultural tools and artifacts have a crucial role. These can be physical artifacts such as books and equipment but they can be symbolic tools such as language and sign systems. Since language, which is central to our capacity to think, is developed in relationships between people, social relationships are necessary for learning. Learning is by definition a social and collaborative activity in which people build knowledge and develop thinking together.

Learning involves participation and what is learned is not the property of an individual but distributed within the social group. For example, an ability to use language skills is not solely an indication of individual intelligence but of the intelligence of the community that developed the language, which the individual then puts to use. The collective knowledge of the community is internalized by the individual. Similarly, as an individual creates new knowledge, for example, a new way of using a tool, then he or she will externalize that knowledge in communicating it to others who will put it to use and then internalize it. In this way, knowledge is created and shared in expansive learning cycles.

These ideas have implications for instruction. They imply that the teacher needs to create an environment in which people can be stimulated to think and act in authentic tasks, like apprentices, beyond their current level of competence, but in what Vygotsky calls their zone of proximal development, see below. Access to, and use of, an appropriate range of tools is an important aspect of such an expansive learning environment. It is important to find activities that a learner can complete with assistance but not alone so that the more expert other, in some cases the teacher but often a peer, can scaffold their learning and remove the scaffold when they can cope on their own. Tasks need to be collaborative and students need to be involved both in the generation of problems and of solutions. Teachers and students become a community and jointly solve problems, so all develop their skill and understanding.

According to this view, the most valued outcomes of learning are therefore engaged participation in ways that others in the community of practice find beneficial. It follows that it is problematic to judge an individual as having acquired knowledge in general terms, abstracted from action settings.

Implications for assessment

These sociocultural theories provide powerful descriptions and explanations of learning but their implications for assessment, especially in school contexts, have not been fully worked out. This is not really surprising because the key theorists have focused on learning, rather than assessment. However, if assessment is to be aligned with learning then appropriate linkages need to be found. The pointers that follow are therefore somewhat speculative. [James \(2008\)](#), labeled these "third generation assessment":

- If learning cannot be separated from the actions in which it is embodied, then assessment too must be situated. This implies that assessment is carried out alongside learning, not after learning.
- Assessment alongside learning implies that it needs to be done by the community rather than by external assessors. There is a role for self-assessment, peer-assessment and teacher assessment.
- Assessment of group learning is as important as the learning of the individual.
- "In vivo" studies of complex, situated problem solving are the most appropriate form for assessments to take because learning is expressed in participation in authentic, real-world activities or projects.
- The focus is on how well people exercise agency in their use of the resources or tools (intellectual, human, material) available to them to formulate problems, work productively and evaluate their efforts. This is a proper justification for coursework assignments and encourages access to source materials because it is the ways in which these are used that is of most significance.
- Learning outcomes can be captured and reported through various forms of recording, including narrative accounts and audio and visual media.

- The portfolio has an important role because it has the potential to capture assessment as inquiry. However, attempts to grade portfolios according to scoring rubrics, as has happened in the USA, is out of alignment with the sociocultural perspective because it reduces the possibilities to assessment as procedure or even assessment as measurement (Serafini, 2000).
- Judgment needs to be holistic and qualitative, to be consistent with a sociocultural or situated approach, not atomized and quantified as in measurement approaches.

This list of characteristics points to possibilities for forms of assessment in sympathy with valued instruction and learning practices in schools, colleges and workplaces. It is especially suited to the assessment of collaborative group work on extended projects. This could make it especially attractive to employers who increasingly claim they are interested in recruits who can demonstrate their capability to work in teams to find creative solutions to complex problems. However, it raises a number of questions about how progression is conceptualized and recognized, and how achievements can be generalized in trustworthy ways when large numbers of students are involved and when those who are interested in the outcomes of such learning cannot participate in the activities that generate them.

Issues of progression and generalizability in sociocultural approaches

The following discussion is also tentative because these issues have not been resolved. It draws substantially on discussions in James (2010, 2017).

On progression, Vygotsky and his followers have something interesting to say.

The structures of grades, scales and attainment levels have accustomed us to regard progression as step-by-step and linear. According to Vygotsky, the mastery of tools of the mind takes place in the zone of proximal development (ZPD). As Grigorenko (1998, p. 210) points out:

The word 'zone' refers to the nonlinearity of children's development. In this zone, the child might move forward or backward, to the left or to the right. ... The main characteristic of the ZPD is its sensitivity to the child's individuality – its responsiveness to the unique profile of each child's skills.

The point to be made here is the value in plotting learners' development as a profile within a zone and encouraging them to expand, deepen and enrich their knowledge, skills and understanding.

Vygotsky's view that what learners can do at lower levels should not limit opportunities for their development of higher functions was almost certainly influenced by his experience of working at a boarding school for children who were both deaf and blind. He claimed that education targeted on the formation of higher mental processes could help people overcome, or compensate for, deficits because it equipped them with adaptive strategies that might be unique for each person.

Although sociocultural approaches can make claims for greater authenticity and validity, they are a long way from providing convincing assurances about the dependability of assessment results. A particular challenge is the question of whether, within a sociocultural theoretical framework, it is possible to make generalizations about a learner's mastery beyond the specific context of the assessment, however extended and rich this assessment activity might be. Baird et al. (2017) point out that psychometricians design assessments with multiple items to sample the domain of interest in representative ways using the now sophisticated statistical techniques available to them. These approaches command respect and generate confidence although there is a tendency for much of the debate surrounding them to shift to technical issues of measurement, such as the strength of correlations among items designed to represent constructs. The nature of the constructs themselves is sometimes scrutinized less well, e.g., what constitutes "reading". The danger is that the items can come to define the construct, rather than the other way round, with the consequence that teachers feel pressured to teach their students how to answer test items, which become the de facto curriculum.

A sociocultural approach, in which rich assessment tasks are integrated with the curriculum in such a way that they can be valued as learning tasks themselves, might be a solution to this problem. The act of generalization could then become integral to the assessment task if it draws together many relevant parts of the domain, which may be multidisciplinary, and requires students to demonstrate, evaluate and apply their accumulated knowledge and skill to a substantial problem. An example, of this approach in schools might be the "exit exhibitions" promoted by the US Coalition of Essential Schools founded bySizer (1992). In post-graduate higher education, mastery of the chosen field is demonstrated in the doctoral thesis or dissertation with oral examination. In craft training, in the past, the "masterpiece" was the piece of work designed to demonstrate the highest levels of expertise of the apprentice or journeyman wishing to become a member of the guild. In all these cases, the basic premise is that an individual who is capable of identifying and organizing the knowledge salient to the problem, and devising, executing, and communicating a solution which is judged worthy by more expert others, is judged to have mastered the more general field at the required level. In all these contexts, mastery is the key concept.

The roots of all these assessment practices predate psychometrics by several centuries. However, the confidence placed in them depends on the strength of the arrangements for fairly judging performances and outcomes. A justified concern is that judgment can be biased by prejudice, favoritism, nepotism or corruption. These are extreme examples but the danger of unwitting bias when teachers are assessors of their own students is real (Bennett, 2011). The strength of psychometric methods of assessment is that they have effective tools for detecting bias and strengthening the reliability and generalizability of inferences drawn from results. These tools cannot easily be used within sociocultural approaches that promote the use of authentic assessment tasks, the outcomes of which cannot be wholly pre-specified. However, efforts to eliminate bias, to ground judgment securely in evidence, and to draw

conclusions about mastery that have more general applicability are crucial. These challenges might be met by using established ethnographic methods for assuring quality of inferences from evidence. Among these, triangulation and reflexivity are especially important. Reflexivity requires assessors to think very carefully about why their judgments might be wrong and actively seek evidence that might, in Popperian terms, falsify them. Triangulation can be achieved both by comparing different pieces of evidence of mastery, e.g., in the judging of a portfolio of work or exhibition as is characteristic of visual and performing arts courses, and by making arrangements for more than one assessor to judge performances and outcomes with a view to coming to an agreement on quality. Systems of moderation are examples of this.

High quality teacher moderation has been developed for large-scale assessment (Maxwell, 2010) although generally such systems are regarded as costly in terms of money and time. Nevertheless, apprenticeship models, from which some sociocultural ideas derive, may offer solutions because underpinning such models is the concept of the community of practice that is the guardian and arbiter of developing standards. In other words, validation of standards by a community of experts, or, at least, “more expert others”, may be a way of assuring quality. This raises a question, however, of whether assessors or judges should work from standard sets of criteria (rubrics) to form generalizable judgments, or whether it is more appropriate, in the context of assessing authentic tasks, to work back from expert judgment to the criteria of quality that inform them. Lawrence Stenhouse, who was opposed to the objectives model of curriculum development, argued that complex performances can be critically evaluated but not pre-specified (Stenhouse and Verma, 1980; see also James, 2012b).

More work could be done to develop assessments that are coherent with sociocultural perspectives on learning, and to find solutions to evident problems. It is not necessary to displace psychometric approaches entirely, but it may be important to find ways to complement them in order to provide more complete accounts and evaluations of learning and achievement.

Can different approaches to learning and assessment be combined? The importance of curriculum

The previous sections have attempted to show what congruence between assessment practice and beliefs about learning might look like. If assessment practice is out of line, effective learning can be undermined because assessment wash back is so powerful, especially in high stakes settings. A key question for teachers, trainers and assessors, is whether they can blend approaches or whether they need to adopt a particular approach, to the exclusion of others. Theoreticians often take the view that theories should take adequate account of all relevant phenomena to provide as complete a solution as possible. However, principled and pragmatic compromises can be reached. The overarching principle is fitness for purpose: always a key criterion for choice of assessments.

The nature of the subject domain might encourage one approach in preference to another. For example, science and mathematics, with hierarchically ordered conceptual structures may lend themselves to constructivist approaches better than broader fields of study, such as the expressive arts, with contested or multiple criteria of what counts as quality learning. The latter may sit better with sociocultural approaches that value context, process and community. However, it would be wrong to claim that the relationship between particular subject domains and particular learning theories, and their associated assessment practices, are exclusive. Almost every subject domain has a number of different aspects that might require different approaches. Take reading and writing in mother tongue for example. In terms of functional literacy, a behaviorist approach might be adequate, but when it comes to understanding and interpreting text, in terms of concepts such as plot, character, setting and metaphor, a cognitive constructivist approach may be needed. Creative writing will demand something different again. Novice writers will need to learn how to develop expertise in a community of practice (see James, 2006, for a developed example). This implies that it is important to scrutinize the variety of curriculum aims and/or desired outcomes within a subject domain, or field of study, before adopting a particular approach to learning, instruction and assessment.

A grounded analysis by James and Brown (2005) of 30 projects in the UK's Teaching and Learning Research Program (2000–2011) identified seven categories of valued outcomes. None were exclusive to any particular subject domain or education sector.

- *Attainments*—often implying mastery of information, facts, specific rules or mental procedures associated with particular tasks. They are frequently assessed in traditional ways through tests and examinations.
- *Understanding*—of ideas, concepts, processes. These imply more abstract, conceptual ways of thinking and the capability to make reasoned judgments about meanings and inference.
- *Cognitive and creative*—imaginative construction of meaning, arts or performance. These focus on the capability to create or discover new knowledge, whether personal, cognitive, esthetic or practical.
- *Using*—how to practice, manipulate, behave, engage in process or systems. This implies the development of known practical and technological skills and being able to apply them in appropriate contexts.
- *Higher order learning*—advanced thinking, reasoning, metacognition. These concepts of learning transcend other learning outcomes. They place emphasis on more sophisticated overarching cognitive processes.
- *Dispositions*—attitudes, perceptions, motivations. These reflect the affective conditions within learners that are assumed to be necessary to effective learning in specific contexts.
- *Membership, inclusion, self-worth*—the learners' affinity toward, readiness to participate in, and sense of worthwhile contribution to, the group where the learning takes place. It is concerned with establishing social and substantive identities.

The first question to ask is, what curriculum aims and outcomes are we seeking to achieve? Then we can ask, what approaches to learning and assessment are most appropriate?

There is also evidence that learning theory is shifting from former paradigmatic stances. Many cognitive constructivists have taken on board the importance of the social dimension of learning: hence the increasing use of the term “social constructivism”. And sociocultural researchers are involved in a “discursive shift” to recognize the cognitive potential in explaining how we learn new practices (Edwards, 2005). This suggests possibilities for synthesis whereby a more complete theory can emerge from blending and bonding of key elements of previous theories. Such synthesis may lead to a new alignment. The possibility for a more complete and inclusive theory of learning to guide practice and assessment seems to be a goal worth pursuing.

Learning (how) to learn and implications for assessment

The importance of “learning to learn”, or “learning how to learn”, has been claimed for a long time but more recently because it is felt to be especially important for lifelong learning in the 21st century. The crucial resource is thought to be the ability of people to respond flexibly and creatively to demands for new knowledge, skills and dispositions in continuously changing environmental, social and economic contexts, especially in a digitally networked world. In the first decade of the 21st century, a number of projects and initiatives sought to elaborate the concept and to investigate the roles that assessment might have. Many of these initiatives were generated in Europe although Learning to Learn was also the title chosen by Hong Kong for its holistic reform of curriculum and assessment in the same decade (James, 2021).

In 2006, the European Parliament and Education Council defined the concept learning to learn as:

Learning to learn is the ability to pursue and persist in learning, to organize one's own learning, including through effective management of time and information, both individually and in groups. This competence includes awareness of one's learning process and needs, identifying available opportunities, and the ability to overcome obstacles in order to learn successfully. This competence means gaining, processing and assimilating new knowledge and skill as well as seeking and making use of guidance. Learning to learn engages learners to build on prior learning and life experiences in order to use and apply knowledge and skills in a variety of contexts: at home, at work, in education and training. Motivation and confidence are crucial to an individual's competence.

European Union Education Council (2006, para. 5, annex).

This definition attempts to avoid reducing the concept to study skills that can be taught, learned and assessed, often as a separate course or program, in the hope and expectation that they will then be applied to learning in the formal curriculum. It also attempts to embrace the range of perspectives on learning that an expert group of researchers and stakeholders brought to the discussions that informed the Council in 2005. The intention was to produce an indicator system that might be used across countries and regions for mainly diagnostic purposes i.e. to identify areas for intervention. A suite of measures of learning to learn capability, drawing mainly on scales previously developed at the Universities of Helsinki and Bristol were proposed as a pilot in eight countries. It is difficult to find evidence of the outcomes of this initiative. It seems that, in the end, the European Union decided to accept the non-content specific measures, such as problem-solving, in the OECD's Program for International Student Assessment (PISA), as a proxy for learning to learn in Europe.

The whole field of learning (how) to learn research and scholarship is too vast to deal with adequately here. For the purposes of this chapter, and in light of the discussion in previous sections, the most pertinent feature is the extent to which the debate is framed in terms of learning theories. The report from the European Commission Joint Research Center (Hoskins and Fredriksson, 2008), on the deliberations of the expert group, has a section on the paradigm debates. Most of the projects examined for helpful insights claimed that they drew on both cognitivist and sociocultural theories, although their work had different emphases that indicated different approaches in practice. Three examples illustrate this. All were represented on the European Commission's expert group.

The University of Helsinki projects (Hautamäki, 2014) acknowledge the influence of both Piaget and Vygotsky in the development, begun in 1996, of a set of scales to assess learning to learn (LTL) capability. This is defined as cognitive competences and the attitudes and beliefs that support the effective use of them. These skills are assessed by pencil-and-paper or computer-based group tests comprised of cognitive tasks and self-report questionnaires. The cognitive component is measured by tasks that are related to curricular content but they require the application of “higher-order” thinking skills. Thus the Finnish LTL scales cover reading comprehension, mathematical thinking skills and more general thinking and reasoning skills, including problem solving. They were developed by psychologists, behavioral scientists and psychometricians and were designed to assess, in cross-sectional and longitudinal studies, the effectiveness of education at system level. They were to provide evidence-informed explanations of how differences between groups of students, classes, schools, districts and countries develop and what kinds of factors are related to different trajectories of development. For all these reasons the scales were of particular interest to the European Commission in its hope of developing a learning to learn indicator for use in member countries. The Helsinki tools were not designed to be diagnostic at individual level; however, the focus has been shifting toward the use of assessment results in the development of classroom practices (Vainikainen, 2014).

The University of Bristol project has similarities but also marked differences. The metaphor of learning as a journey is implicit in Helsinki's developmental approach but it is quite explicit in the work at Bristol. Moreover their definition of learning how to learn “involves the person who is learning, and requires motivation, a sense of direction and desire, and a sense of agency and self-regulation” (Deakin Crick, 2007, p. 138). Thus the focus is on attitudes, values and dispositions and what the Bristol team call

“learning power”. This refers to seven core dimensions of the capacity to learn how to learn, which emerged from a factor analytic study of nearly two thousand learners (Deakin Crick et al., 2004). These factors are: changing and learning; critical curiosity; meaning making; dependence and fragility; creativity; relationships/interdependence; strategic awareness. This analysis gave rise to the design of a self-report 72-item questionnaire—the Effective Lifelong Learning Inventory (ELLI). This elicits information from the learner about how they report themselves on each of these dimensions at a particular point in time. The results can be fed back to the individual learner in profiles represented as spider diagrams, or aggregated for classes and schools in pie charts.

The ELLI learning profile is described by its authors, as a form of dynamic assessment with an interactionist (sociocultural) orientation. According to Deakin Crick (2007, p. 150)

The dynamic assessment of students’ learning power serves a number of pedagogical purposes. First, it reflects back to the learner what they say about themselves in relation to their personal power to learn. Second, it reflects back to the teacher data about individuals, and groups, which can be used for diagnosing what is needed to move forward in the development of self-awareness, ownership and responsibility for learning. Third, the dimensions of learning power provide scaffolding for the ways in which students encounter the formal content of the curriculum. All of these operate together

Deakin Crick et al. (2015) later reanalyzed 15 years of data from ELLI and concluded that a more parsimonious and robust model was possible. This analysis examined the relationships among the previous dimensions of learning power and identified the overarching importance of developing resilient agency in learning—an embodied and relational process. This locates it firmly with a sociocultural frame.

Whereas the Helsinki scales were first designed to be assessments “of” learning to learn, the Bristol inventory was intentionally assessments “for” learning to learn. However, because both projects produced diagnostic assessment instruments they are open to both summative and formative uses.

The Cambridge based project (James et al., 2007), which also involved researchers from King’s College London, the Open University and Reading University, had a very different focus, although within a sociocultural theoretical approach. This project chose the phrase “learning how to learn” (LHTL) in order to draw attention to the development of learning practices. This focus was consistent with the project’s starting point in assessment for learning: a label for a group of formative practices that had been shown to help learners to improve learning and achievement (Black and William, 1998). Four clusters of classroom practices were thought to be particularly important:

- helping learners to understand learning aims and objectives and what counts as high-quality learning;
- helping learners to build on strengths and overcome barriers in their previous learning;
- helping learners to know how to act on constructive feedback and how to develop the motivation to do so;
- helping learners to use the help of others, including through peer-assessment, to enhance their understanding and take responsibility for their learning.

The central aims of the project were to investigate what teachers can do to help students to learn how to learn, and to investigate how this can be supported and managed in schools and across networks, including electronic networks. A suite of tools was developed (James et al., 2006) but these tools are mostly in the form of professional development activities. However, they include a survey instrument for investigating teachers values and practices, which was a key instrument used in the research.

As in Helsinki and Bristol, clarification of the concept of learning how to learn was a first project task (Black et al., 2006). The Cambridge/King’s/OU/Reading team was multidisciplinary and they explored philosophy as well as psychology for definitions. They noted that, as early as 1976, Robert Dearden, a philosopher of education, put forward the idea that:

Learning how to learn is at one stage further removed from any direct specific content of learning. It might therefore reasonably be called ‘second-order learning’. There could be many such comparably second-order activities, such as deliberating how to deliberate, investigating how to investigate, thinking out how to think things out, and so on.

Dearden (1976, p 70).

However, this begs the question of what should be given priority in the whole range of potential second-order learning. Dearden’s leading candidate is promoting learning autonomy. The project came to regard this as the principal aim or goal of the activity of LHTL, which is promoted by assessment for learning (AFL) practices. The Cambridge project concluded that what is needed is that these ways of thinking and practicing, principally “making learning explicit” and “promoting learning autonomy”, which emerged from a factor analysis of the teachers’ survey, should permeate the whole curriculum and be integrated into every subject, and most lessons, so that students have opportunities to reflect critically and constructively on their learning while they are engaged in learning curriculum content. This approach acknowledges that there might well be different LHTL practices across the various subjects of the curriculum. It is unlikely, therefore, to be a unitary and wholly generalizable cognitive or affective competence.

This project also came to the view that it is important not to see learning only in terms of individual minds, but also to recognize the importance of the social element in learning. Even such processes as metacognition and self-regulation are developed through discourse, social interaction and collaborative activities. Adult support serves to mediate such metacognition, hence the focus in this project on working with teachers and schools to develop and support classroom practice.

As a consequence of this emphasis on LHTL practices integrated into the formal curriculum and mediated by others, the Cambridge project did not develop a measure to assess LHTL competence, although it did research LHTL practice through surveys, interviews and classroom observations. The success of these practices was assessed by the extent to which they appeared to have an impact on first-order learning i.e., learning of formal curriculum content in subject areas as measured by national tests and examinations.

The Helsinki researchers were also interested in the associations between patterns of results from their instruments and achievements as measured by assessments such as PISA. It was thought that the former might be explanatory of the latter.

Two points arise from these three projects. First, all appear to combine cognitivist and sociocultural perspectives on learning although each places more weight on one than the other. Secondly, they differ in the extent to which they focus on assessment “of” learning to learn (through the development of measurements) or on assessment “for” learning to learn (through the development of classroom practices). However, all have a diagnostic aspect, and the tools they each develop can be, and in some cases have been, used for both formative and summative purposes. According to Harlen and James (1997) what unifies formative and summative assessment processes is the extent to which they serve shared aims, goals and criteria for evaluating outcomes, whether these are framed in terms of powerful knowledge in subject domains (first-order learning) or the clusters of competences and practices termed learning (how) to learn (second-order learning).

This last observation reflects the claim made at the beginning of this chapter: that assessment needs to be congruent with our understandings about learning and that both need to be aligned with the educational aims and purposes that are valued and promoted. Therefore clarity about all these elements, and their relationships, is important.

References

- Baird, J., Andrich, D., Hopfenbeck, T.N., Stobart, G., 2017. Assessment and learning: fields apart? *Assess Educ. Princ. Pol. Pract.* 24, 317–350.
- Bennett, R., 2011. Formative assessment: a critical review. *Assess Educ. Princ. Pol. Pract.* 18, 5–25.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5 (1), 5–75.
- Black, P., McCormick, R., James, M., Pedder, D., 2006. Assessment for learning and learning how to learn: a theoretical inquiry. *Res. Pap. Educ.* 21 (2), 119–132.
- Bloom, B.S. (Ed.), 1956. *Taxonomy of Educational Objectives, the Classification of Educational Goals—Handbook I: Cognitive Domain*. McKay, New York.
- Bredo, E., 1997. The social construction of learning. In: Phye, G.D. (Ed.), *Handbook of Academic Learning: Construction of Knowledge*. Academic Press, San Diego, CA, pp. 3–45.
- Bruner, J., 1996. *The Culture of Education*. Harvard University Press, Cambridge, MA.
- Deakin Crick, R., 2007. Learning how to learn: the dynamic assessment of learning power. *Curric. J.* 18 (2), 135–153.
- Deakin Crick, R., Broadfoot, P., Claxton, G., 2004. Developing an effective lifelong learning inventory: the ELLI project. *Assess Educ. Princ. Pol. Pract.* 11, 248–272.
- Deakin Crick, R., Huang, S., Shafi, A.A., Goldspink, C., 2015. Developing resilient agency in learning: the internal structure of learning power. *Br. J. Educ. Stud.* 63 (2), 121–160.
- Dearden, R.F., 1976. *Problems in Primary Education*. Routledge, London.
- Edwards, A., 2005. Let's get beyond community and practice: the many meanings of learning by participating. *Curric. J.* 16 (1), 49–65.
- Engeström, Y., 1999. Activity theory and individual and social transformation. In: Engeström, Y., Miettinen, R., Punamäki, R.-L. (Eds.), *Perspectives on Activity Theory*. Cambridge University Press, Cambridge, pp. 19–38.
- European Union Education Council, 2006. Recommendation of the European Parliament and the Council of 18 December 2006 on Key Competencies for Lifelong Learning. Official Journal of the European Union, Brussels.
- Grigorenko, E., 1998. Mastering tools of the mind in school. Trying out Vygotsky's ideas in classrooms. In: Sternberg, R., Williams, W. (Eds.), *Intelligence, Instruction and Assessment: Theory Into Practice*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 201–231.
- Harlen, W., James, M., 1997. Assessment and learning: differences and relationships between formative and summative assessment. *Assess Educ. Princ. Pol. Pract.* 4 (3), 365–379.
- Hautamäki, J., Kupiainen, S., 2014. Learning to learn in Finland. Theory and policy, research and practice. In: Deakin Crick, R., Stringher, C., Ren, K. (Eds.), *Learning to Learn. International Perspectives From Theory and Practice*. Routledge, London.
- Hoskins, B., Fredriksson, U., 2008. *Learning to Learn: What is It and Can It Be Measured?* JRC Scientific and Technical Reports. Office for Official Publications of the European Communities, Luxembourg.
- James, M., 2006. Assessment, teaching and theories of learning. In: Gardner, J. (Ed.), *Assessment and Learning*. Sage, London, pp. 47–60.
- James, M., 2008. Assessment and learning. In: Swaffield, S. (Ed.), *Unlocking Assessment*. Routledge, London, pp. 20–35.
- James, M., 2010. Educational assessment: overview. In: Baker, E., McGaw, B., Peterson, P. (Eds.), *International Encyclopedia of Education*, third ed., vol. 3. Elsevier, Oxford, pp. 161–171.
- James, M., 2012a. Assessment in harmony with our understanding of learning: problems and possibilities. In: Gardner, J. (Ed.), *Assessment and Learning*, second ed. Sage, London, pp. 187–205.
- James, M., 2012b. An alternative to the objectives model: the process model for the design and development of curriculum. In: Norris, N., Elliott, J. (Eds.), *Curriculum, Pedagogy and Educational Research: The Work of Lawrence Stenhouse*. Routledge, Abingdon, pp. 61–83.
- James, M., 2017. Re/Viewing assessment: changing lenses to refocus on learning. *Assess Educ. Princ. Pol. Pract.* 24 (3), 404–414.
- James, M., 2021. A decade of reform in Hong Kong. In: McLaughlin, C., Ruby, A. (Eds.), *Implementing Educational Reform: Cases and Challenges*. Cambridge University Press, Cambridge.
- James, M., Brown, S., 2005. Grasping the TLRP nettle: preliminary analysis and some enduring issues surrounding the improvement of learning outcomes. *Curric. J.* 16 (1), 7–30.
- James, M., Black, P., Carmichael, P., Conner, C., Dudley, P., Fox, A., Frost, D., Honour, L., MacBeath, J., McCormick, R., Marshall, B., Pedder, D., Procter, R., Swaffield, S., William, D., 2006. *Learning How to Learn: Tools for Schools*. Routledge, Abingdon.
- James, M., McCormick, R., Black, P., Carmichael, P., Drummond, M.-J., Fox, A., MacBeath, J., Marshall, B., Pedder, D., Procter, R., Swaffield, S., Swann, J., William, D., 2007. *Improving Learning How to Learn—Classrooms, Schools and Networks*. Routledge, Abingdon.
- Kandel, E., 2006. *In Search of Memory: The Emergence of a New Science of Mind*. Norton, New York.
- Lave, J., Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- Maxwell, G.S., 2010. Moderation of students' work by teachers. In: Baker, E., McGaw, B., Peterson, P. (Eds.), *International Encyclopedia of Education*, vol. 3. Elsevier, Oxford, pp. 457–463.
- Pellegrino, J.W., Chudowsky, N., Glaser, R. (Eds.), 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academy Press, Washington, D.C.

- Rogoff, B., 1990. *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press, New York.
- Serafini, F., 2000. Three paradigms of assessment: measurement, procedure, and inquiry. *Read. Teach.* 54 (4), 384–393.
- Stard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27 (2), 4–13.
- Sizer, T., 1992. *Horace's School: Redesigning the American High School*. Houghton Mifflin, New York.
- Stenhouse, L., Verma, G., 1980. *Educational Procedures and Attitudinal Objectives: A Paradox*. CARE, University of East Anglia (mimeograph), Norwich.
- Vainikainen, M.-P., 2014. Finnish Primary School Pupils' Performance in Learning to Learn Assessments: A Longitudinal Perspective on Educational Equity. Research Report 360. University of Helsinki, Helsinki.
- Vygotsky, L.S., 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Watkins, C., 2003. *Learning: A Sense-Maker's Guid*. Association of Teachers and Lecturers, London.
- Wenger, E., 1998. *Communities of Practice*. Cambridge University Press, Cambridge.

Assessment for learning

Sue Swaffield^a and Roszalina Rawi^b, ^a University of Cambridge Faculty of Education, Cambridge, United Kingdom; and ^b National Institute of Education, Asian Languages and Cultures Academic Group, Singapore

© 2023 Elsevier Ltd. All rights reserved.

Introduction	21
Assessment	22
Learning	22
Assessment for learning	23
History	23
Definitions and key characteristics	24
Student agency	24
Focus on learning	25
Evidence, dialog, feedback, reflection	25
In the moment, simultaneously with teaching and learning	25
Principles and strategies	26
Ten principles	26
Three principles	26
Strategies	27
Distinguishing assessment for learning from formative assessment	28
Formative assessment	28
Assessment for learning and formative assessment	29
Formative assessment and assessment for learning principles	29
Commentaries, critiques and counter-arguments	29
Reasserting the distinctiveness of assessment for learning	30
Assessment for learning and cultural values	30
Assessment for learning principles and classroom learning culture	30
Making learning explicit, behaving virtuously	30
Promoting learner autonomy, respecting hierarchy	31
Resistance to reform due to continued good performance	31
Reconciling AfL principles and East Asian classroom assessment practices	32
Leadership support	32
Increasing teachers' awareness	32
Conclusion	32
References	33

Introduction

The assessment considered in this article focuses on and directly supports student learning rather than, for example, assessment that prioritizes qualification and accountability. Words and phrases can carry multiple meanings, and may be variously interpreted by different people. We use the term “assessment for learning” (AfL) intentionally, and set out our understanding of the concept below, for as Paul Newton urged, “To avoid getting ourselves confused, and to avoid confusing others, we need to use the language of assessment with greater precision” (Newton, 2007, p. 157). Semantics and conceptualizations make a difference to individual, collective, political, and societal understanding, discourse, and policy—and most importantly for students, for practice. Aiming for clarity is particularly pertinent in an international encyclopedia that seeks to reach across countries, cultures and languages.

The article falls into two main parts. Firstly, a decade on from the publication of the article “Getting to the heart of authentic assessment for learning” (Swaffield, 2011) we revisit arguments therein, examine points made by citing authors who have engaged with the issues, and refer to other relevant publications. This is a manifestation of the intellectual dialog that the author invited (Swaffield, 2011, p. 434). It was anticipated that such work might result in future development or change in thinking; the exercise has led to what is essentially a reaffirmation of her 2011 position, as well as a re-emphasis of key points and further clarifications.

The conceptual discussion of part 1 begins with brief considerations of assessment and of learning (both within educational contexts and particularly as relevant to AfL), as a prelude to the substantial discussion of assessment for learning, particularly as related to learning and including distinguishing AfL from formative assessment. In this first part of the article we draw on literature predominantly from Europe, North America, Australia and New Zealand.

The second part further extends the international scope of the article by focusing on AfL and learning in East Asian heritage cultures. Research across Asia has shown the importance of cultural change for successful assessment reform (Cagasan et al.,

2020; Leong et al., 2018; Wicking, 2020). We discuss AfL practice in the context of East Asian classrooms, specifically in Vietnam, Japan, Taiwan, Singapore, Brunei and the Philippines. We explore tensions and conflicts between the principles of AfL and the developmental trends that affect learner dispositions across these countries, cultures and languages.

Assessment

As is clear from other articles in this volume educational assessment is a multi-faceted concept that can be thought about in many different ways, for example: its purpose, function and use; timescale, participants, focus and scope; method and instrumentation; impact, outcome and consequence. Assessment involves evidence, interpretation and judgment. Many of these elements are key to the discussion of AfL below. The etymological root of “assessment” in the Latin word “assidere,” meaning “to sit beside,” resonates closely with the practices we understand as AfL and that are embedded in the on-going everyday dynamic of teaching and learning.

This integration and the range of practices (such as feedback) associated with AfL have led some (e.g., Brown, 2019) to suggest that what is being referred to should just be regarded as part of good teaching, and that AfL should not be termed “assessment” at all. To do this though would imply that assessment refers only to approaches such as testing and examinations for grading, selection and accreditation. The principle “Assessment for learning should be recognized as central to classroom practice” (ARG, 2002, p. 2) reflects the fact that every day learners demonstrate their knowledge, understanding and skills - evidence that is noted and interpreted as the basis for decision-making. Similarly Dylan Wiliam (2011a, p. 11) argued that “The value of using the term assessment in phrases like assessment for learning and formative assessment is that it is illuminating to draw attention to the fact that the processes under consideration can be thought of as assessment processes.” Indeed, a teacher in the course of a lesson eliciting evidence by asking questions, interpreting the responses, and making decisions about the next question or activity, is undoubtedly assessing. Our view is that this should be recognized and named as such, thus reinforcing the understanding that assessment is a broad concept encompassing much more than testing.

Acknowledging AfL as assessment also highlights the important role of students in assessment. Rick Stiggins, reflecting on his own learning about classroom assessment from Anne Davies and Ruth Sutton, highlighted “the power of student involvement in the ongoing classroom assessment process as a way to develop in students a sense of control over their academic well-being—to help them maintain the confidence that success is within reach if they keep striving” (Stiggins, 2016, p. vi). Self-assessment and peer support are central to students’ autonomous regulation of their learning.

Learning

Learning, like assessment, is a concept with multiple meanings and interpretations, and that is also explored in depth in other articles in this volume, notably Mary James’ (2023). Drawing on earlier work (e.g., James, 2008) she examines relationships between processes of assessing and processes of learning, demonstrating that different theories of learning have various implications for assessment. She identifies and discusses three views of learning using Chris Watkins’ (2003) framing (learning as acquiring what is taught; learning as individual sense-making; and learning as building knowledge through social interaction), incorporates Anna Sfard’s (1998) two metaphors for learning of acquisition and participation, and matches them to underlying theories (behaviorist; cognitive constructivist; and sociocultural). Key points that are particularly relevant for this discussion of AfL are that: (1) assessment practices appropriate for behaviorist views of learning prevail, and many people regard them as good assessment per se; (2) cognitive constructivist perspectives, while also focusing on individuals’ knowledge acquisition and processing, might additionally involve problem-solving; and (3) sociocultural learning theories suggest collaborative group work on extended projects, but the assessment of such practices has yet to be fully developed, and the issues arising (such as the trustworthy attribution of achievement and the recognition of progression) resolved. James argues that relationships between learning theories, assessment practices and indeed subject domains, are not exclusive, and stresses the overarching principle that assessment must be fit for purpose. Her final section considers learning (how) to learn through three projects including the four-university Learning How to Learn (LHTL) project that she led (James et al., 2007). All three projects combined cognitivist and sociocultural perspectives. The LHTL project was particularly related to AfL, and will be returned to later.

In the previous edition of this encyclopedia Paul Black (2010) offered another perspective on learning particularly relevant to AfL. He summarized “established principles for effective learning” as:

- “Start from a learner’s existing understanding
- Involve the learner actively in the learning process
- Develop the learner’s understanding of the aims and criteria for effective learning
- Promote social learning i.e., learning through discussion”. Black (2010, pp. 361–362).

Black pointed out that while these principles concern cognitive dimensions of learning, Ruth Butler (1988) and Carol Dweck (2000) have highlighted the importance of affective and motivational dimensions. Their work underpins other principles about learning especially concerning the nature and focus of feedback, advocating sensitive and constructive comments that relate to the progress and achievement of the task and how to improve still further. All these points are also germane to the discussion of AfL below.

Assessment for learning

History

Our understanding of AfL is informed by research published predominantly since 1998, along with important antecedents. The early history and underpinnings of AfL have been detailed elsewhere (e.g., [Swaffield, 2011](#); [Wiliam, 2011a,b](#); [Baird et al., 2017](#)); of significant note are:

1967 Michael Scriven's coining of the term "formative evaluation"

1969 Benjamin Bloom describing brief classroom tests to aid learning as formative evaluation

1984 England's and Wales' statement of policy for Records of Achievement

1986 Lynn Fuchs' and Douglas Fuchs' review of the effects of formative evaluation

1986 Harry Black's coining of the term "assessment for learning"

1987 Gary Natriello's review of the effect of evaluation on students

1988 Terence Crook's review of the impact of classroom practices on students

1992 Mary James' conference paper entitled "Assessment for learning"

1995 Ruth Sutton's book entitled "Assessment for Learning"

1998 Paul Black's and Dylan Wiliam's review of classroom formative assessment.

[Black and Wiliam's \(1998a\)](#) review has been enormously influential and spawned a huge increase of activity in the field, encompassing research, practice and policy. Its impact was probably due not only to the finding that substantial learning gains are possible, but also to a number of other factors. The review had been commissioned by a group of UK academics, called the Assessment Reform Group (ARG), who continued to support and advance the work. The review's main findings were widely disseminated: alongside their journal article the authors wrote a summary ([Black and Wiliam, 1998b](#)) that rapidly sold tens of thousands of copies. This was the first of a number of A5 sized pamphlets published by King's College London (the "Inside the black box" series) and by the University of Cambridge (e.g., [ARG, 1999](#)) making academic research on AfL accessible by distilling key messages for teachers, school leaders, advisers/consultants and policy makers. In England AAIA (The Association for Assessment Inspectors and Advisers, later renamed the Association for Achievement and Improvement through Assessment) worked closely with teachers and school leaders to develop AfL practice. ARG members were active and influential throughout the UK and beyond, while in 2001 Sutton and Stiggins initiated what was to become a series of international symposia [2001 Chester, UK; 2005 Portland, USA; 2009 Dunedin, New Zealand; 2011 Bergen and Solstrand, Norway; 2014 Fredericton, Canada; 2016 Brisbane, Australia] and a network moving the field forward (see [Laveault and Allal, 2016](#)).

Of the many empirical research projects that built on Black and Wiliam's review, two were directly related and are noteworthy. The King's Medway Oxfordshire Formative Assessment Project (KMOFAP) involved Black, Wiliam and colleagues working closely with 48 teachers from six senior-age schools. Together they explored the translation of ideas from the 1998 review into classroom practice ([Black et al., 2002, 2003](#)). These teaching and learning approaches were also found to improve test and examination results ([Wiliam et al., 2004](#)). The teachers' commitment, creativity and professional realism generated many effective and contextualized teaching and learning activities. The collaborative process of knowledge creation was also the basis for the researchers' wider

theorization about putting AfL into practice including the centrality of teachers' attitudes and beliefs, school leadership and professional learning.

The other closely related research was the Learning How To Learn project previously mentioned (James et al., 2007). It was another development and research project, though much more "light touch" in terms of intervention, and much larger in scope and size, than KMOFAP. LHTL involved researchers from four universities including members of the KMOFAP team, 40 schools, and approximately 1500 teachers. Using AfL as a starting point the project investigated how LHTL could be promoted in classrooms, schools and networks. One of the many research tools was a questionnaire that investigated teachers' classroom practices and values revealing salient factors later framed as principles.

Alongside and complementing funded research projects, many proponents of AfL have extensive and long-term engagement with practitioners, directly developing teachers' skills and understanding while growing the knowledge base through analysis and publication. Particularly prominent either side of the north Atlantic are Shirley Clarke, Dylan Wiliam, and Anne Davies and colleagues.

Definitions and key characteristics

Sue Swaffield (2011) discussed two definitions of AfL, both developed by experts in the field through extended dialog drawing on research. The first:

Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

ARG (2002, pp. 2–3).

was widely adopted and continues to be used with fidelity. However, its misinterpretation by some policy makers and practitioners as frequent use of summative tests led to the 31 attendees at the third international symposium drawing up a position paper on AfL (Klenowski, 2009). Central to this statement was a revised definition, coherent with and complementing the ARG's (2002) definition, to make clear the central focus on learning by students...

Assessment for Learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning."

Klenowski (2009, p. 264). line breaks added to the original.

The accompanying commentary elaborated each of the five phrases (emphasized by layout above) to elucidate the definition. These two key definitions of AfL, together with their originators' arguments and discussion by other authors, portray practice with a number of features that are fundamental characteristics of our understanding of AfL. Students are the supreme focus; their learning is the primary aim of AfL, and their active participation essential.

Student agency

Centrally and crucially, students are key actors in AfL. As Stiggins (2002, p. 761) asserted "assessment for learning must involve students in the process." The importance of active student involvement in assessment has long been recognized (for example DES/Welsh Office, 1984; Broadfoot, 1986; Sutton, 1992, 1995; Perrenoud, 1998), and for much longer and even more internationally in relation to learning in general (for example Confucius, Socrates, John Dewey, Maria Montessori, A.S. Neill, and Lev Vygotsky). Early exponents of AfL practice, notable Ruth Sutton (1992, 1995) and Shirley Clarke (1998) from England, and Anne Davies (2020) from Canada, learned the criticality of student involvement through personal experiences in their own classrooms and with other teachers. Research strengthened their convictions.

Students can be actively involved in their learning in many ways, and it is their self-regulation that is essential to the discussion here. Christian Brandmo et al. (2020) (drawing on Schunk and Greene, 2018) describe self-regulated learning (SRL) as "the process where the learner activates and sustains their cognition, motivation, behaviors and feelings toward the attainment of a learning goal" (Brandmo et al., 2020, p. 332). This quotation is from their editorial for a journal special issue on "Bridging classroom assessment and self-regulated learning" (Assessment in Education: Principles, Policy & Practice, Volume 27 Number 4). The authors of the articles in the special issue drew on rich and diverse backgrounds to make substantial conceptual contributions (notably Allal, 2020; Andrade and Brookhart, 2020; Chen and Bonner, 2020; Greene, 2020), and reported empirical studies.

Linda Allal's (2020) model is actually of the over-arching concept of co-regulation (following Perrenoud, 1998), which has self-regulation at its core, nested within peer interaction, teacher interventions and interactions with students, all encompassed by the regulating structure of the teaching/learning situation. At each layer a range of "tools" (including specific assessment procedures and instruments) is operating. The different levels of regulation are interconnected, therefore tools need to be designed, used and

understood coherently. For example, for a student to learn from feedback (from whatever source) time needs to be built in for the feedback to be considered and acted upon, and—as Ron Berger’s (2003) work so powerfully illustrates—a culture to have been created in which feedback is regarded as positively helpful rather than a label or determinant of potential. Heidi Andrade and Brookhart (2020) also take a co-regulation perspective on classroom assessment, and their organizing framework (built on Pintrich and Zusho, 2002) for self- and co-regulation of learning considers four phases (spanning forethought to reflection) and four areas of learning including cognition and affect. Their work “suggests that classroom assessment plays a vital role in the regulation of learning, as well as the converse: the regulation of learning requires assessment evidence” (Andrade and Brookhart, 2020, p. 366). This iterative process is also evident in a third framework (Chen and Bonner, 2020)—developed from Barry Zimmerman’s (2002) cyclical model of SRL—that emphasizes students as active agents at all stages of classroom assessment/self-regulated learning.

Self- and co-regulation of learning are predicated on a number of other key characteristics of AfL. These relate to the what, how, when, where and who of AfL, and all involve students as active participants.

Focus on learning

“The primary aim of Assessment for Learning (AfL) is to contribute to learning itself” (Klenowski, 2009, p. 263)—in contrast to focusing directly on test performance. The 2009 position paper went on to explain the logic of that assertion: “when true learning has occurred, it will manifest itself in performance. The converse does not hold: mere performance on a test does not necessarily mean that learning has occurred” (Klenowski, 2009, p. 263). As Manu Kapur (2016, p. 289) argues “Learning and performance are not always commensurable. Conditions that maximize performance in the initial learning may not maximize learning in the longer term.”

As part of everyday practice (that includes the development of good learning habits) AfL is concerned with progress on specific short-term learning intentions. Cumulatively these curriculum outcomes contribute to the attainment of long term learning goals that are often expressed as standards expected by the end of one or more years of study. It is only fairly recently (corresponding approximately with the development of AfL practices, and in England with the introduction of a national curriculum from 1988) that it has become common practice to share with students what they should be learning and not just what they should be doing. The importance of doing so was demonstrated by John Frederiksen and White (1997) in one of the eight cornerstone examples of research used by Black and Wiliam (1998a): put simply, students improve more when they understand what is required. Learning intentions and their associated success criteria (Clarke, 2005, 2014) can and should take different forms including in their degree of openness. Tightly specified objectives for particular pieces of knowledge or a narrow skill help students to work toward deeper understanding and build expertise, but they have also been associated with and then criticized for narrowing learning—what Harry Torrance (2012) terms “conformative assessment.” However, as Swaffield (2011) pointed out, Elliott Eisner’s (2002) categorization of objectives includes not only these “curriculum objectives” but also “problem solving objectives” and “expressive outcomes.” The latter two are crucial for encouraging exploratory and divergent learning, and for creativity. Students come to understand what they are aiming for and to evaluate their efforts through descriptions of qualities and exemplification, and through process-focused success criteria that separate the objective from the context—a crucial insight from Clarke (2005).

When AfL practice includes the sophisticated use of learning intentions and success criteria, students are not only learning “the curriculum” they are at the same time learning how to learn—in other words, developing and improving lifelong capability for learning. The LHTL project (James et al., 2007) saw AfL as a means of developing learning, with learner autonomy key to both. Thus the learning to which AfL contributes is subject specific, cross-curricular, and generic.

Evidence, dialog, feedback, reflection

Along with the centrality of student agency, and a particular conception of learning, AfL is also characterized by evidence, dialog, feedback and reflection including judgment. Although similar groups of activity are often presented as circular with one leading to another, in AfL their interrelationship is much more complex. For example, evidence may be elicited through dialog as well as being a subject of dialog; reflection on evidence, particularly against learning intentions and criteria, may occur during and following dialog and lead to feedback; dialog may be used to clarify feedback and its implications. The whole process is rooted in cognitive and socio-cultural theories of learning.

Furthermore, it is only when evidence, dialog, feedback and reflection are used and enacted in particular ways that they are characteristic of AfL. Students need to be actively involved, and the focus kept on learning as discussed above. It is the fine details that make all the difference (see particularly Clarke (2014, 2021), Davies (2020), Davies et al. (2012), Marshall and Drummond (2006), and Wiliam (2011b)).

In the moment, simultaneously with teaching and learning

Some aspects of the everyday practice of AfL need to be planned carefully in advance. For example, the teacher being clear about the exact learning intention, the context in which it will be taught (e.g., to know ways of controlling drought, in the context of savannah grassland; listing the different causes of draught, and explaining how these could be reduced (Clarke, 2005)) and pertinent success criteria, and how all this will be shared with and involve the students in the lesson.

Other aspects are contingent and require moment-by-moment decision-making and enactment. Gordon [Stobart \(2014, p. 116\)](#) uses an example of a classroom exchange to illustrate how the right diagnosis and therefore decision depended on the teacher deciding to ask a key additional question and on instantaneous analysis:

Teacher: How much is 7 take away 4?

6 year-old: 2

Teacher: How did you get that answer?

6 year-old: I knew that 7 take away 4 is 2 because I knew 4 plus 2 is 7. And if 4 plus 2 is 7, then 7 take away 2 must be 4.

[Ginsburg \(1997\)](#).

The teacher could have very easily focused on the incorrect answer “2”, or on the different calculation posed at the end of the explanation (“7 take away 2”), but instead recognized the 6 year-old’s sophisticated mathematical understanding as well as weaknesses in basic computation. If the teacher had been relying solely on written work it is highly unlikely that the child’s reasoning would have ever been made explicit. This is why skilled questioning contingent on evidence (of knowledge, understanding and misconceptions) that is often only available in the moment, and dialogic moves that are central to exploring ideas with others (“interthinking” ([Mercer, 2000](#))), play such important roles in AfL.

Principles and strategies

Ten principles

Alongside the 2002 definition, the ARG published ten principles of AfL:

- AfL should be part of effective planning of teaching and learning
- AfL should focus on how students learn
- AfL should be recognized as central to classroom practice
- AfL should be regarded as a key professional skill for teachers
- AfL should be sensitive and constructive because any assessment has an emotional impact
- AfL should take account of the importance of learner motivation
- AfL should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed
- Learners should receive constructive guidance about how to improve
- AfL develops learners’ capacity for self-assessment so they can become reflective and self-managing
- AfL should recognize the full range of achievements for all learners. [ARG \(2002, 2–3\)](#).

As with the definition, the principles were research-based and developed through consultation with a wide variety of individuals and associations. They encompass the key characteristics discussed above, and align with the conception of AfL being explicated in this article. Each principle was accompanied by an explanatory paragraph providing a succinct rationale and a very brief overview of what AfL practice entails, usefully pointing to the complexity of AfL. For example, the fourth principle—AfL as a key professional skill for teachers—is elaborated thus: “Teachers require the professional knowledge and skills to: plan for assessment; observe learning; analyze and interpret evidence of learning; give feedback to learners and support learners in self assessment. Teachers should be supported in developing these skills through initial and continuing professional development” ([ARG, 2002, p. 2](#)). Each element of this elaboration concerns huge bodies of theory, research and practice: for example, one of the very many publications on feedback in educational settings runs to over 600 pages ([Lipnevich and Smith, 2018](#)). This, and the difficulty of remembering ten principles, perhaps places limitations on them as a practical tool for guidance in the immediacy of day-to-day teaching and learning.

Three principles

It is much easier to remember a small number of statements, which is one of the advantages of the three principles that arose from the LHTL project. A survey of staff from 32 schools (junior and senior age ranges) elicited reported practices and values about classroom assessment ([James and Pedder, 2006; James et al., 2007](#)). Factor analysis of data from teachers led to the identification of three groups:

Factor 1: *Making learning explicit*—Eliciting, clarifying and responding to evidence of learning; working with students to develop a positive learning orientation.

Factor 2: *Promoting learning autonomy*—A widening of scope for students to take on greater independence over their learning objectives and the assessment of their own and each other's work.

Factor 3: *Performance orientation*—A concern to help students comply with performance goals prescribed by the curriculum through closed questioning and measured by marks and grades.

James and Pedder (2006, pp. 122–123), James et al. (2007, p. 67).

These three dimensions “provide a sound basis for identifying the fundamental principles that underpin the use of assessment in the classroom” (James and Pedder, 2006, p. 129). “Learner autonomy” is probably a more accurate representation of factor two since it is the students who develop autonomy. Reframing the third factor actively to align with the first two principles, and positively to promote AfL practice, gives a formulation for the third principle of “Focusing on learning” rather than focusing on grades or marks as indicators of performance. The three resulting principles—

Making learning explicit

Promoting learner autonomy

Focusing on learning

(mnemonic LEAF) reflect much other relevant research as the 30 items of the questionnaire were derived from literature—not only about assessment practice per se, but also about underlying principles of learning as well as related processes, contexts and roles (James and Pedder, 2006). The wording of the principles emphasizes that AfL is all about learning—learning portrayed as overt, evidenced, discussed, broader than prescribed curriculum goals or performance measured by marks and grades, and positively engaged with by students who develop greater responsibility for their learning.

The three principles are a framework for thinking about AfL that “holds the promise of providing teachers with a basis for making informed decisions in unpredictable circumstances” (James and Pedder, 2006, p. 129). We have used and promoted these three principles, finding them to be powerful tools for analysis (see the discussion of AfL in East Asian countries below), research, policy, practice and professional development. They provide a common language and coherence for discussions of AfL across varied contexts, as demonstrated by teachers of different subjects working with students aged from 4 to 16 years (Swaffield et al., 2016).

Strategies

The discussion of characteristics of AfL above, and the lists of principles, give a sense of the breadth and complexity of AfL practice. Over the years key strategies central to LfL practice have been depicted in various ways (Black et al., 2003, p. 22; James et al., 2007, pp. 8–12; Wiliam, 2011, p. 46; Davies et al., 2012, p. xi; Clarke, 2014, pp. 6–7). The number of categories vary (between four and 12 in the references just cited), reflecting the breadth, detail and approach to grouping. While authors use different language, and emphasize different aspects, the following list represents commonly accepted strategies of AfL—which importantly are interlinked and not mutually exclusive:

Creating a culture conducive to everyone learning

Clarifying and sharing learning intentions, co-creating success criteria

Developing dialogic talk and worthwhile questioning

Eliciting and evaluating evidence of learning

Engaging in self assessment and peer support

Giving feedback that promotes learning

The first strategy, creating the culture, is to a large extent an outcome of the enactment of other strategies, but listing it separately draws attention to the necessary conscious nurturing of conditions conducive to learning. Every one of the strategies lend themselves to reflecting the three principles of making learning explicit, promoting learner autonomy, and focusing on learning. Moreover, the principles should be used as a check on putative AfL practice and a guide to “authentic” AfL (Swaffield, 2011). This is especially important as the strategies are realized (or not) in multiple contexts through an almost infinite range of techniques whose fine details make all the difference to the quality and quantity of students’ learning. Whichever aspects of AfL are being planned, carried out, or observed, the answers to the questions “Is this making learning explicit? Is it promoting learner autonomy? Is it focusing on learning (rather than marks or grades)?” must be indisputably “Yes, Yes, Yes.” “If practices fail to serve the underlying principles ... then they cease to be assessment for learning” (James et al., 2007, p. 215). The argument that practices are only AfL if they match certain principles underpins differences between AfL and formative assessment.

Distinguishing assessment for learning from formative assessment

The terms “Assessment for Learning” and “Formative Assessment” are often used synonymously, both in published academic work and in practice-orientated documents and discourse. Some consistently refer to formative assessment, others have used both terms over time while a few, not least ourselves, maintain a deliberate and differentiated use. Swaffield (2011) discussed points of difference and argued for explicitly upholding the distinction between AfL and formative assessment.

We have set out our understanding of AfL in sections **Introduction** and **Assessment for learning** above. Crucial determinants are captured by the two linked seminal definitions and the three fundamental principles. The earlier set of ten principles, key characteristics, and strategies also discussed above further elaborate the richness and complexity of AfL. In arguing that there is a distinction between AfL and formative assessment it is necessary to clarify our understanding and use of the term formative assessment.

Formative assessment

According to the Merriam-Webster dictionary, “formative” is an adjective meaning “giving or capable of giving form.” Therefore formative assessment is assessment that forms or shapes something. It often refers to teachers using assessment to inform their teaching, though the formative use of assessment data can also be much broader. Importantly, formative assessment does not—and should not—refer to a particular assessment instrument, but to evidence, interpretation and judgment (as in section **Assessment** above) that is used to make decisions about future action.

For many years following Scriven’s (1967) seminal discussion of formative and summative evaluation, the term “formative assessment” was most commonly found alongside its partner “summative assessment,” referring to two broad purposes of assessment. Formative assessment denotes using evidence and consequent judgments about students’ current learning to inform future teaching and learning. It is contrasted with summative assessment which is concerned with summarizing students’ attainment—often expressed by grades, scores or levels and used particularly for certification, selection, and monitoring. Despite appearing a neat dichotomy and there being very important conceptual differences concerning their prime purposes, they are interrelated in multi-various and sophisticated ways (Harlen and James, 1997), including through the formative use of summative tests (Black et al., 2003).

There are many ways in which assessment data can be used formatively; Wiliam’s (2011a, pp. 10–11, 2011b, pp. 40–43) seven scenarios of the use of evidence of student learning to guide future teaching and learning-related actions illustrate the scope. The first three refer to educators operating at a local or district level, using test scores to make decisions about professional learning, improving teaching, or the provision of remedial classes. The other four feature individual teachers making decisions about their next teaching based on purposefully elicited evidence of their students’ current understanding. The teacher-focused scenarios provide clear evidence of the strategy “engineering effective classroom discussions, activities and tasks that elicit evidence of learning” (Wiliam, 2011a, p. 12, 2011b, p. 46) and of assessment being used formatively. The students demonstrated their knowledge and understanding by handing in quiz papers, completing exit passes, holding up cards to indicate their responses to the teacher’s questions, or showing a sketch on a whiteboard.

Together the scenarios demonstrate that assessment data can be used formatively by a wide range of educational decision-makers not least teachers, to inform the next teaching moment of a lesson, the focus of the next lesson(s), additional instruction during the

year, or teaching in future years. The beneficiaries may be the students who have provided the evidence of learning, or indeed many other students in different contexts.

In summary, formative assessment—the formative use of assessment information—is an educational purpose and a function with a broad scope.

Assessment for learning and formative assessment

Formative assessment and assessment for learning principles

It has been argued above and elsewhere (James et al., 2007; Swaffield, 2011) that practices deemed to be assessment for learning must serve certain principles—specifically making learning explicit, promoting learner autonomy, and focusing on learning.

It is not clear to what extent the examples of using assessment data formatively referred to above involved making learning explicit. There was certainly no mention of clarifying, sharing or understanding learning intentions and success criteria, and indeed doing so would rarely be seen as a natural part of the analysis of either test scores or students' responses to individual questions. The students' role in all cases was to provide evidence of their learning; they were not expected or given the opportunity to take the initiative or to support others. None of the scenarios would have fulfilled the AfL principle of promoting learner autonomy. The teachers' judgments and decisions would undoubtedly have helped the students toward better mastery of the specific learning intentions, but the assessment does not appear to serve the double duty (Boud, 2000) of also helping the students become better learners. The longer time-span examples of using assessment data formatively are particularly problematic for the involvement of students and their development as learners. If information about learning gained from assessment "is used for long-term curriculum improvement, it will not help the learning of the students currently involved. It might be formative for the teacher, but not for the students" (Black et al., 2003, p. 122).

None of these observations should detract from the value of the formative use of assessment data. Teaching, in both the long and short term, should be adapted to students' learning needs—the proactive and retroactive regulation of learning (Allal and Lopez, 2005). They do though illustrate that formative assessment (as understood in section above: **Formative assessment**) does not necessarily involve the students beyond being data generators, nor does it necessarily make them better, self-regulating, learners.

Commentaries, critiques and counter-arguments

Some of the extensive educational assessment literature relates directly to the conceptualizations and interrelationship of AfL and formative assessment. One line of argument concerns their purpose and function, with near universal acceptance that both are intended to inform, and understood to improve, teaching and learning. Function should actually match purpose—assessment information must actually be used. Wiliam (2011a,b), building on his and Black's previous extensive work, stressed the use of evidence for decisions as a defining feature of formative assessment.

The critiques by Harry Torrance (2012) and Shuichi Ninomiya (2016) (who undertook her work during a sabbatical hosted by Torrance) relate particularly to assessment criteria, high-stakes performativity pressures and student autonomy. Their concerns that convergent criteria, conformative assessment procedures and an emphasis on coaching and practice to "close the gap," inhibit the development of independent and critical learners are well founded. However, as has been stated (Swaffield, 2011 and section **Focus on learning** above) explicit criteria do not have to be narrow, and Torrance's and Ninomiya's criticisms do not apply to AfL guided by the three principles. Indeed, their advocacy for developing student understanding and autonomy are served by making explicit the type of learning that is valued, and by promoting learner autonomy; their warning against an emphasis on coaching and practice matches the third principle of focusing on learning rather than performance. The authors conclude that the problems they discuss could be ameliorated by attending to the relationships and interactions between teachers and students—echoing one of Swaffield's (2011) key characteristics of authentic AfL.

Chris Cookson (2018) contrasted three pairs of assessment concepts including formative assessment and assessment for learning. He commented on a number of authors, and found himself broadly in agreement with Swaffield (2011). The reservations he did raise stemmed from his translation into a table of six significant statements about AfL and formative assessment (made by Swaffield after ten pages of explication and discussion). In so doing he interpreted elements within each statement as being mutually exclusive, rather than the more nuanced understanding particularly about emphasis and scope as in the original. Cookson's examination of other authors' difficulties in differentiating assessment and evaluation led him to a preference for the later term. He concluded by proposing replacing the term assessment for learning with "evaluation for learning" as "most suitable as a representative phrase of using testing and the insights derived therefrom, to advance learning and teaching" (p. 939); this despite his apparent acceptance that AfL is much broader than testing. Another critic with testing very much on his mind is Gavin Brown (2019), whose position in his self-declared opinion piece was that AfL cannot be considered as assessment since it cannot meet the technical requirements of tests, particularly in terms of validity, reliability and verification. As argued in 1.1 we (along with many others) do not equate AfL solely or primarily with tests.

Randy Bennett's (2011, p. 6) discussion of "the definitional issue" starts by considering "the distinction between summative and formative roles," before focusing on conceptions of formative assessment as instrumentation (a position often adopted by test publishers) or as a process favoring qualitative insights above scores. His position was that while the prime purpose of formative assessment is to suggest modifications to teaching, it should also be able to give an impression of what students know and can do.

He made a parallel argument for a primary and a secondary purpose of summative assessment—the inverse of those for formative assessment. For Bennett, the terms assessment for/of learning suggested a dichotomous oversimplification, which supported his rejection of the term assessment for learning that he saw as a synonym for formative assessment.

Joanne Baird and colleagues' (Baird et al., 2017) consideration of assessment, learning theories and their relationship included AfL as one of two cases for detailed examination. Their stance was to

... treat Assessment for Learning as offering some distinct emphases within formative assessment approaches. While there is a common function (to directly improve learning), what distinguishes AfL is the focus on student self-regulation and autonomy and on viewing it as an informal and continuous process (p. 336).

These differentiating features resonate with the three AfL principles advocated by James and Pedder (2006) and Swaffield (2011). The focus on learners' self-regulation is also integral to the internationally agreed AfL definition (Klenowski, 2009).

Reasserting the distinctiveness of assessment for learning

The discussion above leads to a reassertion of Swaffield's (2011) argument for the distinctive conceptualization of AfL.

In summary, as agreed by 31 international experts, "assessment for learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialog, demonstration and observation in ways that enhance ongoing learning" (Klenowski, 2009, p. 264). AfL is underpinned by the three principles of making learning explicit, promoting learner autonomy, and focusing on learning. It is a multi-faceted process, with the purpose and function of supporting learning. Students are the supreme focus; their learning is the primary aim, and their active participation—including the practice and development of the regulation of learning—essential. The beneficiaries of AfL practice are the very students involved, working together in a particular learning context; the decisions are acted on immediately or in the near future. AfL enhances at least two types of learning: it aids students' mastery of the curriculum material being studied; what is more, it helps participants become better learners. AfL is itself a broad educative learning process.

AfL viewed holistically is a manifestation of the formative use of assessment information—that is formative assessment. However, some strategies integral to AfL, for example the sharing of learning intentions, are not in themselves formative assessment. Moreover, many of the ways in which assessment is used to inform future teaching and learning are not AfL (see sections above, particularly **Formative assessment** and **Formative assessment and assessment for learning principles**).

Crucially, AfL (a process) and formative assessment (a purpose and function) are fundamentally and categorically different.

We now extend the discussion of AfL to global contexts culturally different from those where much of the development work and research discussed so far were located.

Assessment for learning and cultural values

Research over the last two decades has revealed issues concerning the implementation of AfL in different countries. Philip Hallinger (2010) attributed cultural factors as reasons behind the obstacles to education reform in Asian countries where the process of implementing "western educational innovations" has been anything but smooth.

While there are cultural differences both among and within countries, the seminal work of Geert Hofstede revealed some broad cultural trends that can influence learner dispositions, namely: (1) virtuous behavior toward others consists of maintaining respect, harmony and not questioning authority, (2) the stability of society is based upon hierarchy and (3) the importance of acquiring skills and education, working hard, being patient and persevering (Hofstede and Bonde, 1988).

A study of the influence of cultural context on the implementation of AfL as an assessment reform in Japan found that while practices such as retaining students' focus on learning and conducting peer lesson observations are common in Japanese school culture, other practices such as those that value students' spontaneity and creativity are sometimes discouraged as language and social customs often emphasize distance (Armito et al., 2015). In Japanese schools, personal effort and silent endeavor are favored over verbosity and group discussions, and there is a long history of undifferentiated group instruction and rote learning. Similarly, researchers considering Vietnam (Ho et al., 2016) opined that emphasis has to be placed on AfL strategies that are situated and culturally relevant, arguing that Vietnamese values such as respect for harmony and hierarchy can either encourage or suppress AfL practices.

Assessment for learning principles and classroom learning culture

In order to explore further the influence of cultural values on AfL practice, we now draw on some regional specific research to discuss the alignment between broadly Confucian cultural trends and the three fundamental principles of AfL (James et al., 2007).

Making learning explicit, behaving virtuously

In the Philippines, students rarely ask questions of their teachers (Cagasan et al., 2020). Power distance is a possible explanation, with students occupying a lower status role than the teacher in the Philippine classroom. The consequence of this is that the student

does not see it as appropriate to question the teacher, which limits the opportunities for the teacher to provide the students with more insight into what is being learned.

The same notion of respect for elders, humility and courtesy, is also entrenched in the Malay culture and in some respect enhances the teaching and learning context within the Malay Language (ML) classroom. According to Roszalina Rawi (2017), the attitudes of ML students in Singapore were different in ML classes compared to other subjects. ML teachers interviewed shared that their students were more positive and respectful in their ML lessons, where even students who had behavioral issues in other classes behaved well. This could be related to the affinity that ML teachers enjoy with their students. Rapport is an important element in the context of classroom learning as it develops a positive learning orientation and makes learning explicit. Learning is facilitated in ML classrooms because of the close bond that stems from the cultural context of ML education itself, and correlates with ML teachers' sense of moral responsibility (Rawi, 2017).

However, for teachers to elicit, clarify and respond to evidence of learning, students need to be vocal and actively question in class. It is unusual for students in an ML classroom to question the authoritative elder figure. Students also feel awkward praising themselves even if they do well, as humility is strongly encouraged in the Malay culture (Rawi, 2017). The sense of moral responsibility that Malay teachers have toward their students encourages them to overcome such cultural norms that inhibit assessment reform. Within the Malay community this sense of moral responsibility is taken seriously. The Malay culture imbues the role of an educator with a larger responsibility than what is normally expected and extends it beyond the formal remit. ML teachers seek to uphold the trust and moral responsibilities expected of them to develop their students' learning potential. Rawi (2017) reported that ML teachers focus on the benefits that their students can derive from AfL, and that the teachers were committed to continued effort to develop their AfL practices because they believed the change would be in the best interest of their students.

Promoting learner autonomy, respecting hierarchy

There are cultural clashes between Western-developed educational reforms that emphasize independent thinking and Asian cultural values that emphasize respect for authority and conformity (Ho et al., 2016). Differences between the AfL principle of promoting learner autonomy and facets of Asian culture are evident. While Vietnamese tradition emphasized harmony and "face saving" (not embarrassing others), these were barriers to developing students' awareness of their current level of learning as they could not be honest and critical. In Vietnam, where respect for harmony and hierarchy are highly valued, the idea of AfL, which changes the power relationship between teachers and learners, may be problematic (Ho et al., 2016). When carrying out self and peer assessment, and reporting to the lecturer in the front of the class, Vietnamese university students training to be teachers seemed to be shy and humble (Tien et al., 2020). They were observed to speak in low-pitched voices, to look obediently toward the lecturer and other students, waiting for their reactions. This is the antithesis of AfL strategies such as active questioning and peer assessment where students need to be vocal, critical and unafraid to challenge the opinions of authoritative figures.

Indeed, within the Asian community the notions of having utmost respect for elders, being humble and soft-spoken are inherent. Western modernity and influences are often regarded with suspicion and thought to encroach on traditional heritage values. When AfL is introduced, heritage language teachers may struggle to understand practices such as allowing students to speak for themselves and to decide their own learning objectives, or to promote peer assessment. These are not processes which teachers and students are used to within the traditional heritage language classroom culture (Rawi, 2017). Hence while AfL may be better received and practiced in other subjects such as English or mathematics, it is crucial to note that cultural processes are deeply embedded in specific learning contexts and a "one size fits all" approach to systemic reform simply cannot work (Chan and Law, 2011; Lynch, 2001).

Resistance to reform due to continued good performance

Despite efforts for reform, the existing norm for assessment practices is very much performance-oriented (Wicking, 2020). In Asian societies standardized assessment concentrates on a competitive examination culture (Leong et al., 2018; Tien et al., 2020). In Singapore for example, pupils are prepared to answer national summative test items correctly through rote learning and repetitive practices. While they are scoring high marks in national examinations, fundamentally they might not have learned much other than how to answer exam questions well (Rawi, 2017). The absence of an AfL culture in schools is an impediment to the development of assessment practices that focus on learning and not performance. AfL was introduced into the Brunei education curriculum alongside school-based assessment (SBAfL) but Bruneian teachers, who were more used to summative assessment procedures, expressed confusion and uncertainties in what SBAfL entails and as a result became less confident in integrating assessment reforms in their lessons (Haji Botty and Shahrill, 2015; Rashid and Jaidin, 2014). Recent formative assessment reform in Japan has emphasized criteria compliance and convergent assessment to reach targets (Ninomiya, 2016). When getting good results is what matters most, and teachers continue with assessment practices which ensure such results, students too become conditioned to expect learning to occur under such circumstances (Sadler, 2007). Research into testing programs shows that the increase in test scores over a period of time is most likely because teachers and pupils have become used to the tests, and not due to increased learning (Harlen and Deakin Crick, 2003).

In Asian classrooms, teachers' beliefs about learning are closely associated with their students' abilities to answer questions and do well in a competitive examination culture (Tien et al., 2020). While teachers consider their close rapport with students, and the freedom to be flexible with the scheme of work as factors that contribute positively to learning, they still hesitate to fully embrace AfL because of exam pressures (Rawi, 2017). Research into the assessment practices in Hong Kong and Taiwanese classrooms showed that there were many gaps between the intentions and outcomes of assessment reforms toward AfL (Berry, 2011; Pan, 2008).

Arimoto and colleagues highlighted a conclusion drawn from the 2014 International Symposium on Classroom Assessment and Afl, which is that “assessment is a matter of culture” (Arimoto et al., 2015, p. 42). Unique Japanese pedagogical techniques such as “neriage” whole-class discussions are deemed as being aligned to Afl as they focus on learning. Although the Japanese ethic of valuing personal effort meant that students seek assistance from their peers less actively than their Western counterparts (Arimoto et al., 2015), moments of peer support still made a positive difference in the learning process for Japanese students. According to a recent study by Wicking (2020) Japanese students in higher learning institutions believed that “peers are able to more or less help each other during assessments” (p. 186) with very little rivalry among them. A majority of these Japanese students engaged in assessment tasks with a sense of camaraderie and mutual support from their peers.

Reconciling Afl principles and East Asian classroom assessment practices

Creating cultures where there are opportunities for teachers to learn in context is a core premise for successful reform. According to Richard Elmore (2004):

The problem [is that] there is almost no opportunity for teachers to engage in continuous and sustained learning about their practice in the settings in which they actually work ... Cultures do not change by mandate; they change by the specific displacement of existing norms, structures, and processes by others; the process of cultural change depends fundamentally on modelling the new values and behaviour that you expect to displace the existing ones (p. 11).

Cultural disconnects between the intentions of those who develop education innovations that have been globally accepted and the perspectives of the local practitioner implementing them have resulted in the stalling of certain reforms (Carnoy and Rhoten, 2002; Pan, 2008). While building rapport with students could indeed be important for meaningful learning, it is necessary for teachers to mediate this aspect of classroom culture with Afl strategies so that students are given the confidence to be autonomous learners. Opportunities need to be made available for teachers to learn how to practice Afl in the context of the East Asian classroom where previously (and perhaps still currently) the existing norm for assessment practices is very much performance-oriented (Rawi, 2017).

Leadership support

The leadership support required for educational reform is associated with the provision of opportunities for teachers to put their learning in context and build their capacity of knowledge (Elmore, 2004; Fullan, 2007). It is necessary for school leaders to create a learning culture for Afl and provide opportunities to engage in learning about Afl within the settings where they actually work. Such a situation helps to develop systematic knowledge about Afl by changing the prevailing culture of assessment within East Asian schools.

Increasing teachers' awareness

Rawi (2017) proposed a new understanding concerning the influence of cultural values in the classroom on Afl implementation. If Afl is to be fully embraced, the challenge lies in reconciling tensions between Afl strategies and learning cultures. In order to enact Afl strategies such as active questioning and peer assessment in classrooms, students need to be vocal, critical and unafraid to challenge the opinions of authoritative figures. These are not character traits inherent in the Asian culture that tends to promote humility, being soft-spoken and respecting elders. As argued by Ho et al. (2016) there are cultural clashes between Western-developed educational reforms that emphasize independent thinking, and Asian cultural values that emphasize respect for authority and conformity.

Teachers need to be aware of such cultural disconnects and mitigate the situation by being consistent and persistent with their Afl practices in order to enculture Afl into their classrooms. Teachers' constant practice of Afl strategies will act as semiotic mediators that produce “qualitative transformations” (Moll, 2014, p. 31) in both the teachers themselves and the milieu in which they are situated (Holland et al., 1998). Indeed, as mentioned by Lynch (2001) a one-size-fits-all approach to systemic reform simply cannot work when cultural processes are so deeply embedded in learning contexts. Chan and Law's (2011) notion of the cultural space within which education reforms evolve is also relevant here. The cultural space of teaching and learning is embedded with traditional values that have to be pragmatically regarded when moderating an initiative such as Afl. Teachers have to make explicit that it is still possible for students to be critical yet respectful when voicing their opinions, and to create opportunities for students to familiarize themselves with such practices by including them in the cultural space of classrooms.

Consistency and persistence are required. As unusual as it might be for Afl strategies such as active questioning and clarifying learning intentions to be implemented in Asian classroom settings, with frequent and systematic practice Afl can be encultured into such classrooms. While East Asian education systems are traditionally seen as intensely competitive, research such as Wicking's (2020) proves that increasing exposure to collaborative learning and group work can result in a change to traditional classroom assessment culture.

Conclusion

Assessment integrated into everyday practice should directly improve learning. Afl (as conceptualized in this article) meets this purpose by developing students as better learners, as well as providing teachers with information to make decisions so that the

content and process of their teaching meets students' learning needs. Learners are active participants in AfL, which is a sophisticated set of practices typified by interrelating strategies that can employ multiple techniques. AfL is best briefly summarized by the definition drawn up by an international group of experts (Klenowski, 2009), along with three principles derived from a multi-university research project (James et al., 2007).

Together the three principles of making learning explicit, promoting learner autonomy and focusing on learning are defining criteria for AfL. As well as being an essential check of whether practices are actually AfL, they can guide planning and implementation. The principles also provide a common language for discussing AfL across contexts (for example subjects, age ranges and countries) where AfL practice may superficially appear quite different. As an analytical tool the principles shed light on the challenges and affordances of implementing AfL in different cultures. Cultural differences exist at scales ranging from the transnational to individual classrooms in one school: particular views and values about the nature of learning and about the teacher-learner relationship can be problematic for AfL practice wherever they occur.

The process of AfL is an example of formative assessment—by which we mean the formative use of assessment. Many valuable ways of using assessment information formatively are not AfL, and some essential elements of AfL are not in themselves formative assessment. The terms are not synonymous.

Many academics and practitioners chose to call what we describe as assessment for learning “formative assessment.” In some respects this does not matter: students' experience and their learning are the most important. But what happens in classrooms is influenced by terminology, and especially by understanding. If educational leaders and teachers (quite understandably) interpret formative assessment as adapting teaching on the basis of student performance, all the agency resides with the teacher, and students will not be developing self-regulating skills and dispositions.

Information technology is an enabler for some ways of using assessment data formatively. Technology-driven learning and testing systems can be designed to incorporate the formative use of assessment information, efficiently and effectively improving students' results on test scores, but the danger is this is at the expense of developing better learners and promoting learner autonomy. Artificial intelligence holds the promise of ever more adaptive learning, but could never replace the knowledge about an individual student that informs a teacher's in-the-moment decisions during a learning dialog. AfL is essentially human, relational and dialogic.

Concerns about the normalization of formative assessment practices that do not adhere to the Klenowski (2009) definition and the three principles (James et al., 2007) of AfL still hold, particularly when it is at the expense of authentic assessment for learning (Swaffield, 2011). Learners and their teachers are the losers, while it may be that test and software developers profit substantially.

References

- Allal, L., 2020. Assessment and the co-regulation of learning in the classroom. *Assess. Educ. Princ. Pol. Pract.* 27, 332–349.
- Allal, L., Lopez, L.M., 2005. Formative assessment of learning: a review of publication in French. In: Organisation for Economic Co-operation and Development (OECD), *Formative Assessment: Improving Learning in Secondary Classrooms*. OECD publishing, Paris, pp. 241–264.
- Andrade, H.L., Brookhart, S.M., 2020. Classroom assessment as the co-regulation of learning. *Assess. Educ. Princ. Pol. Pract.* 27, 350–372.
- ARG, 1999. *Assessment for Learning: Beyond the Black Box*. University of Cambridge School of Education, Cambridge.
- ARG, 2002. *Assessment for Learning: 10 Principles*. University of Cambridge Faculty of Education, Cambridge.
- Arimoto, M., Clark, I., Yamamoto, S., Shinkawa, M., 2015. Cultural perspectives on classroom assessment: a path toward the “Japanese assessment for learning network”. *Annu. Bull.* 1, 41–62. Graduate School of Education, Tohoku University.
- Baird, J., Andrich, D., Hopfenbeck, T.N., Stobart, G., 2017. Assessment and learning: fields apart? *Assess. Educ. Princ. Pol. Pract.* 24, 317–350.
- Bennett, R.E., 2011. Formative assessment: a critical review. *Assess. Educ. Princ. Pol. Pract.* 18, 5–25.
- Berger, R., 2003. *An Ethic of Excellence: Building a Culture of Craftsmanship With Students*. Heinemann, Portsmouth, NH.
- Berry, R., 2011. Educational assessment in mainland China, Hong Kong and Taiwan. In: Berry, R., Adamson, B. (Eds.), *Assessment Reform in Education: Policy and Practice*. Springer, Dordrecht, pp. 49–61.
- Black, P., 2010. Formative assessment. In: Baker, E., McGaw, B., Peterson, P. (Eds.), *International Encyclopedia of Education*, third ed., vol. 3. Elsevier, Oxford, pp. 359–364.
- Black, P., William, D., 1998a. Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5, 5–74.
- Black, P., William, D., 1998b. *Inside the Black Box: Raising Standards Through Classroom Assessment*. School of Education, King's College, London.
- Black, P., Harrison, C., Lee, C., Marshall, B., William, D., 2002. *Working Inside the Black Box: Assessment for Learning in the Classroom*. Department of Education and Professional Studies, King's College, London.
- Black, P., Harrison, C., Lee, C., Marshall, B., William, D., 2003. *Assessment for Learning: Putting It Into Practice*. Open University Press, Maidenhead.
- Boud, D., 2000. Sustainable assessment: rethinking assessment for the learning society. *Stud. High. Educ.* 22, 151–167.
- Brandmo, C., Panadero, E., Hopfenbeck, T.N., 2020. Bridging classroom assessment and self-regulated learning. *Assess. Educ. Princ. Pol. Pract.* 27, 319–331.
- Broadfoot, P., 1986. Records of achievement: achieving a record? *Stud. Educ. Eval.* 12, 313–323.
- Brown, G., 2019. Is assessment for learning really assessment? *Front. Educ.* 4, 64.
- Butler, R., 1988. Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *Br. J. Educ. Psychol.* 58, 1–14.
- Cagasan, L., Care, E., Robertson, P., Luo, R., 2020. Developing a formative assessment protocol to examine formative assessment practices in the Philippines. *Educ. Assess.* 25, 259–275.
- Carnoy, M., Rhoten, D., 2002. What does globalization mean for educational change? a comparative approach. *Comp. Educ. Rev.* 46, 1–9.
- Chan, S., Law, M., 2011. Taking education seriously as reform. *Cult. Stud.* 25, 25–41.
- Chen, P.P., Bonner, S.M., 2020. A framework for classroom assessment, learning, and self-regulation. *Assess. Educ. Princ. Pol. Pract.* 27, 373–393.
- Clarke, S., 1998. *Targeting Assessment in the Primary Classroom: Strategies for Planning, Assessment, Pupil Feedback and Target Setting*. Hodder and Stoughton Education, London.
- Clarke, S., 2005. *Formative Assessment in the Secondary Classroom*. Hodder Murray, London.
- Clarke, S., 2014. *Outstanding Formative Assessment: Culture and Practice*. Hodder Education, London.

- Clarke, S., 2021. Unlocking Learning Intentions and Success Criteria: Shifting From Product to Process Across the Disciplines. Corwin, Thousand Oaks, CA.
- Cookson, C., 2018. Assessment terms half a century in the making and unmaking: from conceptual ingenuity to definitional anarchy. *Assess. Eval. High. Educ.* 43, 930–942.
- Davies, A., 2020. *Making Classroom Assessment Work*, fourth ed. Connections Publishing, Courtenay, BC.
- Davies, A., Herbst, S., Parrott Reynolds, B., 2012. *Leading the Way to Assessment for Learning: A Practical Guide*, second ed. connect2learning, Courtenay, BC.
- DES/Welsh Office, 1984. *Records of Achievement: A Statement of Policy*. HMSO, London.
- Dweck, C.S., 2000. *Self-Theories: Their Role in Motivation, Personality and Development*. Taylor and Francis, London.
- Eisner, E., 2002. *The Arts and the Creation of Mind*. Yale University Press, New Haven, CT.
- Elmore, R., 2004. *School Reform From the Inside Out: Policy, Practice, and Performance*. Harvard University Press, Cambridge, MA.
- Frederiksen, J.R., White, B.J., 1997. *Reflective Assessment of Students' Research Within an Inquiry-Based Middle School Science Curriculum*. Paper Presented at the Annual Meeting of the AERA Chicago.
- Fullan, M., 2007. Change theory as a force for school improvement. In: Burger, J., Webber, C., Klinck, P. (Eds.), *Intelligent Leadership: Constructs for Thinking Education Leaders*, vol. 6. Springer, Dordrecht, pp. 27–39.
- Ginsburg, H.P., 1997. *Entering the Child's Mind*. Cambridge University Press, Cambridge.
- Greene, J.A., 2020. Building upon synergies among self-regulated learning and formative assessment research and practice. *Assess. Educ. Princ. Policy Pract.* 27 (4), 463–476.
- Haji Botty, H.M.R., Shahril, M., 2015. A critical review of the school-based assessment in Brunei Darussalam. *Int. J. Educ. Pract.* 3, 17–27.
- Hallinger, P., 2010. Making education reform happen: is there an "Asian" way? *Sch. Leader. Manag.* 30, 401–418.
- Harlen, W., Deakin Crick, R., 2003. Testing and motivation for learning. *Assess. Educ. Princ. Pol. Pract.* 10, 169–207.
- Harlen, W., James, M., 1997. Assessment and learning: differences and relationships between formative and summative assessment. *Assess. Educ. Princ. Pol. Pract.* 4, 365–380.
- Ho, N., Adie, L., Klenowski, V., 2016. The prospects of assessment for learning in higher education in Asian settings: the case of Vietnam. *Assess. Matters* 10, 5–26.
- Hofstede, G., Bond, M.H., 1988. The Confucius connection: from cultural roots to economic growth. *Organ. Dynam.* 16, 5–21.
- Holland, D., Lachicotte Jr., W., Skinner, D., Cain, C., 1998. *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge, MA.
- James, M., 1992. *Assessment for Learning*. Paper Presented at the Annual Conference of the Association for Supervision and Curriculum Development Held at New Orleans, LA.
- James, M., 2008. *Assessment and learning*. In: Swaffield, S. (Ed.), *Unlocking Assessment: Understanding for Reflection and Application*. Routledge, London, pp. 20–35.
- James, M., 2023. *Assessing and learning, and learning to learn*. In: Tierney, R., Rizvi, F., Ericikan, K. (Eds.), *International Encyclopedia of Education*, fourth ed. Elsevier, Oxford (In the volume on Assessment and Accountability, edited by Therese Hopfenbeck.).
- James, M., Pedder, D., 2006. Beyond method: assessment and learning practices and values. *Curric. J.* 17, 109–138.
- James, M., McCormick, R., Black, P., et al., 2007. *Improving Learning How to Learn: Classrooms, Schools and Networks*. Routledge, Abingdon.
- Kapur, M., 2016. Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educ. Psychol.* 5, 289–299.
- Klenowski, V., 2009. Assessment for learning revisited: an Asia-Pacific perspective. *Assess. Educ. Princ. Pol. Pract.* 16, 263–268.
- Laveault, D., Allal, L. (Eds.), 2016. *Assessment for Learning: Meeting the Challenge of Implementation*. Springer, Rotterdam.
- Leong, W.S., Ismail, H., Costa, J.S., Tan, H.B., 2018. Assessment for learning research in East Asian countries. *Stud. Educ. Eval.* 59, 270–277.
- Lipnevich, A.A., Smith, J.K. (Eds.), 2018. *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press, Cambridge.
- Lynch, S., 2001. "Science for all" is not equal to "one size fits all": linguistic and cultural diversity and science education reform. *J. Res. Sci. Teach.* 38, 622–627.
- Marshall, B., Drummond, M.-J., 2006. How teachers engage with assessment for learning: lessons from the classroom. *Res. Pap. Educ.* 21, 133–149.
- Mercer, N., 2000. *Words and Minds: How We Use Language to Think Together*. Routledge, London.
- Moll, L.C., 2014. *L.S. Vygotsky and Education*. Routledge, Abingdon.
- Newton, P.E., 2007. Clarifying the purposes of educational assessment. *Assess. Educ. Princ. Pol. Pract.* 14, 149–170.
- Ninomiya, S., 2016. The possibilities and limitations of assessment for learning: exploring the theory of formative assessment and the notion of "closing the learning gap". *Educ. Stud. Japan* 10, 79–91.
- Pan, H.L., 2008. School leadership for organizational learning and empowerment: the Taiwan perspective. In: MacBeath, J., Cheng, Y.C. (Eds.), *Leadership for Learning: International Perspectives*. Sense Publishers, Rotterdam, pp. 255–267.
- Perrenoud, P., 1998. From formative evaluation to a controlled regulation of learning processes: towards a wider conceptual field. *Assess. Educ. Princ. Pol. Pract.* 5, 85–102.
- Pintrich, P., Zusho, A., 2002. The development of academic self-regulation: the role of cognitive and motivational factors. In: Eccles, J., Wigfield, A. (Eds.), *Development of achievement motivation*. Academic Press, pp. 249–284.
- Rashid, R.A., Jaidin, J.H., 2014. Exploring primary school teachers' conceptions of "assessment for learning". *Int. Educ. Stud.* 7, 69–83.
- Rawi, R., 2017. *Assessment for Learning in Malay Language Classrooms: From Zahir to Batin*. University of Cambridge, Cambridge.
- Sadler, R., 2007. Perils in the meticulous specification of goals and assessment criteria. *Assess. Educ. Princ. Pol. Pract.* 14, 387–392.
- Schunk, D.H., Greene, J.A., 2018. Historical, contemporary, and future perspectives on self-regulated learning and performance. In: Schunk, D.H., Greene, J.A. (Eds.), *Handbook of Self-Regulation of Learning and Performance*, second ed. Routledge, Abingdon, pp. 1–15.
- Scriven, M., 1967. The methodology of evaluation. In: Tyler, R., Gagne, R., Scriven, M. (Eds.), *Perspectives on Curriculum Evaluation*. AERA monograph series – Curriculum evaluation. Rand McNally and Co, pp. 39–83.
- Stard, A., 1998. On two metaphors for learning and the dangers of choosing just one. *Educ. Res.* 27, 4–13.
- Stiggins, R., 2002. Assessment crisis: the absence of assessment FOR learning. *Phi Delta Kappan* 83 (10), 758–765.
- Stiggins, R., 2016. Foreword. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer, Rotterdam, pp. vi–viii.
- Stobart, G., 2014. *The Expert Learner: Challenging the Myth of Ability*. Open University Press, Maidenhead.
- Sutton, R., 1992. *Assessment: A Framework for Teachers*. Routledge, London.
- Sutton, R., 1995. *Assessment for Learning*. RS Publications, Salford, UK.
- Swaffield, S., 2011. Getting to the heart of authentic assessment for learning. *Assess. Educ. Princ. Pol. Pract.* 18, 433–449.
- Swaffield, S., Rawi, R., O'Shea, A., 2016. Developing assessment for learning practice in a school cluster: primary and secondary teachers learning together. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation, the Enabling Power of Assessment*. Springer, Rotterdam, pp. 119–217.
- Tien, N.H., Anh, D.T., Luong, M.V., Ngoc, N.M., Vuong, N.T., 2020. Formative assessment in the teacher education in Vietnam. *J. Hunan Univ. Nat. Sci.* 47, 8.
- Torrance, H., 2012. Formative assessment at the crossroads: conformance, deformative and transformative assessment. *Oxf. Rev. Educ.* 38, 323–342.
- Watkins, C., 2003. *Learning: A Sense-Maker's Guide*. Association of Teachers and Lecturers, London.
- Wicking, P., 2020. Formative assessment of students from a Confucian heritage culture: insights from Japan. *Assess. Eval. High. Educ.* 45, 180–192.
- William, D., 2011a. What is assessment for learning? *Stud. Educ. Eval.* 37, 3–14.
- William, D., 2011b. *Embedded Formative Assessment*. Solution Tree Press, Bloomington, IN.
- William, D., Lee, C., Harrison, C., Black, P., 2004. Teachers developing assessment for learning: impact on student achievement. *Assess. Educ. Princ. Pol. Pract.* 11, 49–65.
- Zimmerman, B.J., 2002. Becoming a self-regulated learner: an overview. *Theor. Pract.* 42, 64–70.

Assessment and the regulation of learning[☆]

Linda Allal, University of Geneva, Geneva, Switzerland

© 2023 Elsevier Ltd. All rights reserved.

This is an update of L. Allal, *Assessment and the Regulation of Learning*, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, *International Encyclopedia of Education* (Third Edition), Elsevier, 2010, Pages 348-352, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00362-6>.

Regulation as the function of formative assessment	35
Implications of theories of learning for formative assessment	36
Formative assessment and external regulation of learning	37
Formative assessment and self-regulation of learning	37
Formative assessment and co-regulation of learning	38
Continuity between formative and summative assessment	39
References	39

Regulation as the function of formative assessment

The function of formative assessment, as practiced in schools and classrooms, is to ensure the regulation of teaching and learning so that students can attain valued learning outcomes. This conception of regulation refers to a dynamic process of adaptation of teachers' and students' actions, rather than to the usage of the term regulation to designate rules that specify prescribed actions (e.g., school system regulations regarding the grading of student work).

Regulation as the function of formative assessment in the classroom has been a key topic in French-language publications since the late 1970s, as described in the reviews in English by Allal and Mottier Lopez (2005) and by Perrenoud (1998). The landmark review by Black and Wiliam (1989) of English-language publications associated with formative assessment highlighted the concept of feedback and its uses, but also referenced research on self-regulated learning which has since gained prominence in the classroom assessment literature.

Formative assessment allows regulation of classroom learning through activities, both formal and informal, that are orchestrated by teachers to assess the learning of students under their pedagogical responsibility and to promote students' involvement in assessment of their own learning. Orchestration of formative assessment by teachers includes the ways teachers use assessment tools provided by the school system, the assessment activities that teachers create and put into practice in their classrooms, and the strategies that teachers adopt to encourage student involvement in assessment. The concept of orchestration highlights the key role of teachers in creating the conditions for formative assessment but also recognizes the role of students as active participants whose actions ultimately determine the effects of formative assessment. Formative assessment, when well designed and implemented, ensures the regulation of teaching and learning through four processes in which teachers and students engage: goal setting, monitoring of progress toward the goal, interpretation of feedback derived from monitoring, adaptation of goal-directed actions and/or adjustment of the goal itself. These processes do not always unfold in a strictly sequential manner. For example, interpretation of feedback may lead to continued monitoring, or to clarification of the goal, rather than to adaptation of instructional activities.

In describing the way formative assessment regulates teaching and learning, Black and Wiliam (2009) proposed a distinction between synchronous contingencies of regulation in which the processes of regulation are closely linked during an activity (e.g., a teacher makes real-time adjustments during a lesson, or a student does so while carrying out a task) and asynchronous contingencies in which the processes are distributed over several points in time (e.g., a teacher uses information from student homework to plan a new lesson, or a student uses feedback from a test to adjust the way of preparing for a subsequent test). Allal (1988, 2020) described three modalities of regulation (interactive, retroactive, proactive) resulting from formative assessment that differ in terms of the temporality of their contingencies. Interactive regulation is based on the student's interactions with teachers, peers, instructional materials, and assessment tools during an on-going learning activity. It entails synchronous contingencies between the processes of goal setting, monitoring, interpretation, and adaptation. Retroactive and proactive regulation involve asynchronous contingencies. In retroactive regulation, formative assessment identifies objectives a student failed to master during a prior phase of instruction and, on this basis, proposes remedial activities so that the student can return to these objectives and overcome difficulties initially encountered. In proactive regulation, information from formative assessment is used to plan new learning activities rather than the correction or improvement of the outcomes of activities already accomplished.

It is generally accepted that formative assessment contributes to the regulation of student learning in an optimal manner when the contingencies are synchronous and interactive: in particular, when formative assessment (1) is integrated within on-going teaching and learning activities; (2) leads to differentiation of activities according to individual students' knowledge, skills, interests, and needs; (3) encourages active student involvement. It is nevertheless recognized that in ordinary classroom conditions, it is

[☆]Change History: May 2021. L. Allal updated the text and the reference list.

difficult to ensure optimal interactive regulation for every student in each learning activity and that asynchronous (retroactive and proactive) modalities of regulation are therefore necessary and useful. This means that the regulation of learning based on formative assessment can include, for example, the following situations: (1) the teacher uses information from an assessment to adapt the remainder of an instructional sequence (adds an extra lesson or a new activity, etc.) for an entire class, but without any individual differentiation; (2) the teacher uses information from an assessment to adapt a lesson plan for future cohorts of students, but this regulation does not benefit the current cohort. Student self-regulation can also include asynchronous actions, for example, a student's decision that it would be useful to return to and review material presented in lessons occurring some time ago, or a decision to not revise an essay, but to take into account peer feedback when drafting a future essay.

Although the regulation of student learning is not the primary function of other forms of educational assessment, all forms of assessment nevertheless influence—and therefore indirectly regulate—the processes of teaching and learning in the classroom. Summative assessment, aimed at the reporting and certification of student achievements with respect to external stakeholders (parents, administrators, potential employers, etc.) influences students' motivation to engage in further learning and their views of themselves as learners. Predictive assessment which determines admission to educational programs or to alternative tracks within a system affects students' access to educational resources and consequently their learning outcomes. Large-scale assessments organized by the education system provide data that guide policy, school curriculum and teacher education, all of which indirectly impact student learning. In short, all forms of educational assessment participate to some degree in the regulation of learning. While the implementation of formative assessment is influenced by the actions and decisions of policymakers, teacher educators and school leaders, teachers nevertheless remain the key initiators of formative assessment practices that regulate teaching and student learning in the classroom. In addition, because teachers are responsible for carrying out summative assessments prescribed by the school system, they are constantly faced with the challenge of coordinating formative and summative assessment.

This article examines in more detail four aspects of the relations between formative assessment and the regulation of student learning:

- the implications of different theories of learning for the conception of formative assessment,
- approaches to formative assessment that emphasize external regulation resulting from teacher actions and the structure of learning environments,
- practices of formative assessment emphasizing students' active engagement in processes of self-regulation,
- conceptions of formative assessment as a process of co-regulation resulting from the joint influence of contextual factors and student self-regulation.

The final section of the article discusses the question of continuity between formative and summative assessment in the classroom and the consequences for the regulation of student learning.

Implications of theories of learning for formative assessment

All theories of learning propose a mechanism of regulation that ensures adaptation of the learner's behavior and thought processes. Well-known mechanisms include reinforcement in behaviorist theory, equilibration in Piaget's constructivism, feedback devices in cognitive models, and social mediation in sociocultural and social constructivist approaches. Although there are important differences among these theories, they all consider processes of regulation to be a central feature of learning.

Formative assessment in the classroom was first proposed by Bloom (1968) as a key component of his model of mastery learning formulated in a neo-behaviorist perspective emphasizing external sources of regulation—namely, formative tests, corrective feedback, and remedial resources provided to students by the teacher. As other theories of learning came into prominence, they led to re-conceptualizations of formative assessment. Drawing on constructivist and cognitive theories of learning, Allal (1979) proposed that formative assessment be based on well-designed tasks that reveal salient aspects of students' reasoning and conceptual understanding, coupled with classroom observations and interactions (teacher-student interactions, peer interactions) that allow monitoring of learners' progression, as well as interactive regulation integrated in teaching/learning activities.

Two theoretical developments helped to further define processes of regulation to be taken into account in formative assessment. One was the work on metacognitive regulation of learning, notably by Brown (1987) who analyzed the processes of self-regulation in a Piagetian perspective and of "other-regulation" in a Vygotskian perspective. The second was the elaboration of the social cognitive theory of learning by Bandura (1986) who emphasized the role self-efficacy beliefs in the regulation of the learner's task engagement and motivation. Bandura's theory had an important influence on the models of self-regulated learning which have become increasingly prominent in the formative assessment literature (see the chronological review by Panadero et al. (2018a)).

A seminal paper by Shepard (2000) traced the long-standing influence of behaviorist and neo-behaviorist learning theories on educational assessment and the emergence of a transformed paradigm of classroom assessment based on cognitive and constructivist learning theories, as well as on Vygotskian concepts of social mediation and scaffolding. Other researchers developed in more detail the implications of sociocultural theories of learning and of theories of situated cognition for classroom assessment (Allal and Pelgrims, 2000; Dann, 2014; Gipps, 1999; Hickey, 2015; James, 2008). One important contribution of this work is the recognition of the interdependence between other-regulation (by a teacher, a mentor, a peer) and the self-regulation of the learner. This means that scaffolding should not be seen as something done *to* students but as something done *with* students. Another major

contribution, particularly from the situated cognition perspective, is the idea that formative assessment concerns not only the regulation of learning but also the formation of students' identity as learners and as participants in a learning community.

In a special journal issue on the relations between learning and assessment, the lead article by Baird et al. (2017) discussed why learning and assessment sometimes appear to be "fields apart" and why new forms of synergy are needed for classroom assessment and for large-scale assessment. Multiple commentaries on this article provided a range of perspectives on learning and assessment.

Formative assessment and external regulation of learning

The distinction between formative and summative evaluation was proposed by Scriven (1967) in an article on the methodology of curriculum and program evaluation conducted by researchers and school administrators. In the mastery learning model developed by Bloom (1968) and Bloom et al. (1971), this distinction was transposed to the assessment of student learning by classroom teachers. Mastery learning was formulated in a neo-behaviorist perspective emphasizing external sources of regulation that guide and reinforce student learning, namely:

- the communication of instructional objectives, defined by curriculum specialists and teachers, at the beginning of each instructional unit;
- the use of formative tests at the end of each instructional unit to identify the objectives attained and those not attained by each student;
- for the objectives not attained, the organization of remedial activities for the students in question: allocation of additional time and assignment of appropriate correctives (prescriptive feedback) and alternative resources designed to promote mastery.

Criterion-referenced summative tests, covering the objectives of several units, are subsequently administered to determine students' final level of achievement in a domain.

In this approach, the teacher is responsible for finding effective means of instruction and assessment to ensure the regulation of student learning. The student's role is to execute diligently the tasks (formative tests, remedial activities) prescribed by the teacher. The mastery learning model recognizes that internal processes (students' prior knowledge and motivation) are important, but its focus is on retroactive regulation from sources in the learning environment.

Although the mastery learning model has been criticized by proponents of other theories of learning, it is important to recognize two major contributions of this model. The first is its radical break with the traditional notion that success and failure in school learning are determined by purely internal dispositions of the learner (aptitude, ability, intrinsic motivation). A second contribution is the idea that formative assessment provides a profile of information for each student (indication of objectives attained and not attained) and that the resulting adaptation of instruction is differentiated according to the profile of student outcomes. The linkage between formative assessment and the differentiation of instruction remains an enduring heritage of the mastery learning model.

External forms of regulation can of course be implemented by other means than by a teacher, for example, by embedding formative assessment tasks in curriculum materials. The development of computer-based learning and assessment environments has led to more powerful ways of providing sources of external regulation to be used by teachers and by students (Bennett, 2015; Tomasik et al., 2018). These environments enlarge the scope of mastery learning by the inclusion of features (e.g., generation of items as a function of the student's responses and errors) that allow synchronous interactive regulation, as well as periodic retroactive regulation. Another feature of these environments is the possibility of data aggregation to ensure a formative assessment continuum from adaptations of teaching and learning in the classroom to system-wide adaptations of curriculum and policy. In upper secondary and higher education, computer-based tools sometimes function in a stand-alone way without any teacher intervention, but in classroom settings with younger students, the teacher usually has an important role in the orchestration of how these tools of formative assessment are deployed and used by students.

Formative assessment and self-regulation of learning

In constructivist, cognitive and social cognitive theories of learning, external sources of regulation can enhance or inhibit processes of self-regulation, but learning occurs only when self-regulatory mechanisms intervene in the transformation of existing knowledge and skills. Student self-regulation, which had been discussed episodically in the formative assessment literature, became the primary focus of the chapter by Andrade (2010) entitled "Students as the definitive source of formative assessment." She defined the relation between self-regulation and formative assessment by associating the three recursive phases of the Zimmerman (2000) model of self-regulated learning with the questions from the Hattie and Timperley (2007) review of feedback and its effectiveness. This led to a definition of student self-regulation in formative assessment as follows:

- Forethought: "Where am I going?"
- Performance and control: "How am I doing?"
- Reflection: "What's next?"

Classroom practices of formative assessment, she argued, should create the conditions which allow students to actively engage in these self-regulatory processes.

Research on self-regulated learning has been a very productive field of investigation leading to multiple models (reviewed by Panadero, 2017) with key insights for the conceptualization of formative assessment. In addition to the models describing recursive phases of self-regulation, the dual process model of Boekaerts (2011) is of particular interest because it postulates two simultaneous directions of self-regulation: a growth pathway linked to (meta)cognitive regulation and a well-being pathway linked to motivational regulation. In some cases, formative assessment practices aimed at providing positive support for (meta)cognitive regulation of learning can have detrimental effects on motivation. For example, detailed feedback on improvements and revisions that a student is expected to carry out can sometimes trigger a self-protective coping strategy leading to inaction. The opposite could also occur: a formative assessment practice could foster student motivation (e.g., an attractive activity with little domain-content structure) but not ensure progress in the acquisition of more advanced skills and cognitive competence.

Student involvement in formative assessment can be enhanced in several ways in the classroom: by whole-class discussions between teachers and students about learning goals and assessment criteria, by students' active engagement in self-assessment procedures, and by students' participation in various forms of peer assessment. A review of research on student self-assessment by Andrade (2019) emphasized the importance of finding ways to integrate cognitive and affective mechanisms of self-assessment and self-regulation. Referring to research on self-assessment and on peer assessment, Panadero et al. (2016) proposed a series of guidelines for teachers who wish to create conditions that encourage active student involvement in formative assessment. A "roadmap" for fusing self-regulated learning and formative assessment was subsequently formulated by Panadero et al. (2018a) based on their review of 33 publications that provide insights and empirical data concerning relations between self-regulation and formative assessment. In the introduction to a special journal issue, Brandmo et al. (2020) discussed the convergent aims of formative assessment and self-regulated learning, as well as factors that may limit implementation of high-quality practices in the classroom.

A major contribution of the above publications is their specification of the processes of self-regulation in which students need to engage for learning to effectively take place and the assessment practices that are likely to foster these processes. For example, goal setting initiated by the teacher, using curriculum materials that define assessment standards, can through whole-class discussions and analysis of exemplars enhance student understanding of standards and also incorporate students' suggestions about how standards can be exemplified in a formative assessment tool. Although approaches emphasizing student self-regulation recognize and promote sources of external regulation that can support and enhance self-regulation, they consider external regulation as a means that will ultimately allow students to become autonomous self-regulated learners.

Formative assessment and co-regulation of learning

There are several different conceptions of co-regulation and its relationship to formative assessment (Allal, 2020). Hadwin and Oshige (2011) proposed an enlargement of the field of self-regulated learning to include three categories: self-regulation, which entails active monitoring and regulating of one's own learning; co-regulation, which is considered as a transitional process in the learner's appropriation of self-regulation strategies through interaction with a more capable other (teacher, more advanced peer); socially shared regulation, which refers to co-constructed regulation by multiple participants (of equivalent status) who ensure the progression of their shared activity. For other researchers (Volet et al., 2009), co-regulation refers to any socially mediated regulation, whether with a more capable other (teacher or more advanced peer) or with peers of equivalent status. With a focus more specifically on classroom assessment, Andrade and Brookhart (2020, p. 350) defined co-regulation as "regulation by others" resulting from teachers' interventions, from other students, from instructional materials, and from assessment practices. They enumerated for each phase of the Pintrich and Zusho (2002) model of self-regulated learning (forethought, monitoring, control, reaction and reflection) the attributes of self-regulation and the corresponding attributes of co-regulation in four areas: cognition, motivation, behavior, context. Their discussion provided many indicators for developing the regulation of student learning through classroom assessment, but largely retained the self-regulated learning premise that co-regulation is a means for helping students become autonomous self-regulated learners.

A model of co-regulation based on the perspective of situated cognition and learning has led to a different conception of its relationship to formative assessment (Allal, 2007, 2020). In this model, co-regulation is defined as the *joint* influence on student learning of the learner's processes of self-regulation and of the sources of regulation in the learning environment: namely, the structure of the teaching/learning situation; the teacher's interventions and interactions with students; the interactions between students; the tools (materials, artifacts, instruments) used for instruction and for assessment. In this perspective, co-regulation is not one type of regulation among others but an overarching construct that bridges the internal-external regulation distinction. Moreover, co-regulation is not a transitory process or a support system to allow students to become autonomous self-regulated learners. Rather it is a general function of instruction and assessment in classroom communities where students learn to participate in increasingly complex and diversified forms of co-regulation.

In this model of co-regulation, students' interactions with teachers and peers are the dynamic force moving learning forward but tools, and in particular assessment tools, have several important functions. (1) Tools ensure linkages between the different sources of external regulation (situation, teacher, peer) and the learner's self-regulation. For example, the teacher may select an assessment rubric provided by the curriculum materials and then conduct a whole-class discussion during which more specific success criteria are defined and illustrated; the resulting extended rubric may then be used by students both for peer assessment and for self-assessment. (2) Tools can amplify the effects of interactive co-regulation by making the goals more explicit and the monitoring

of progress more systematic than would be the case in informal interactions without assessment tools. (3) Tools often allow recording of traces of assessment that can be used for deferred regulation. For example, when a group of students assesses its collaborative writing activity using a checklist, this information can be subsequently used by the teacher to prepare a new classroom activity designed to help students overcome previous difficulties.

An article by [Panadera et al. \(2018b\)](#) proposed an extension of the Allal definition of co-regulation to encompass not only the effects on the student (assesse) but also the effects on the teacher or peer assessor. An interesting topic for future research would be to document how assessors, particularly teachers, learn to improve their interventions through their participation in formative assessment interactions with students.

Continuity between formative and summative assessment

Formative assessment coexists in the classroom with the summative assessment procedures that are prescribed by the school system. The development of continuity and coherence between these two functions of assessment is in the interest of both teachers and their students. A starting point is the alignment of formative and summative assessment activities and tools with well-defined curriculum goals and descriptions of learning progressions. In addition, it is important that teachers develop a classroom assessment culture that allows students to understand the linkages between different assessment activities and the primary purpose, or in some cases multiple purposes, of each activity.

The information from summative assessments (e.g., grades in report cards) influences student motivation and engagement in learning ([Brookhart, 2001](#)): the effect is often positive for high-achieving students but can be detrimental for other students. Despite the differences in the basic purposes of formative and summative assessment, and the negative effects of some forms of summative assessment, the question can be raised as to their possible synergy in promoting student learning ([Harlen, 2005](#)). Can summative assessment actively support student learning? Can it contribute to what the UK Assessment Reform Group ([ARG, 1999](#)) termed “assessment for learning”?

It has been argued by [Black and Wiliam \(2009\)](#) that any given assessment activity or instrument, depending on how the resulting information is used, could have a formative and/or summative function. [Bennett \(2011\)](#) has similarly claimed that well-designed assessment instruments or systems can serve both functions, albeit with primacy often given to one function or the other. These views provide openings for continuity and possible synergy between formative and summative assessment. [Laveault and Allal \(2016\)](#) have proposed that the ARG concept of assessment for learning, which is often treated simply as a synonym for formative assessment, can in fact encompass both formative assessment practices and practices situated in an interface between formative and summative assessment. This interface includes:

- use of information from summative assessments to implement the regulation of teaching and learning activities (e.g., results from an end-of-term test included in students’ report cards are also used to organize remedial activities for some students);
- use of information from formative assessments to enrich summative assessment reports (e.g., observations from formative assessment activities are integrated in comments about students’ grades in their report cards in order to provide a more complete picture of student competencies);
- assessment activities that include phases or components with a formative function and other phases or components with a summative function (e.g., in portfolio-based or project-based assessment, teachers provide feedback and advice about student plans and progress, but subsequently formulate a summative appraisal of the final product).

Although assessment activities with a strictly formative function are generally the most powerful means for ensuring the regulation of teaching and learning, it is important to recognize the role of activities in the formative-summative interface which, if well-conceived and conducted, may also contribute to the progression of student learning.

There remain, nevertheless, many forms of summative assessment used for reporting that are inappropriate for the regulation of student learning in a productive way. Examples would include norm-referenced summative assessments, assessments that summarize achievement by a single outcome measure rather than a profile of outcomes, and summative assessments that are so frequent and time consuming that there is little space for the adaptation of teaching and learning.

To conclude, many challenges still exist, of which two will be mentioned. The first is the adoption by school systems of summative assessment practices that—at the least—are not detrimental to student learning and—at best—provide positive support for the regulation of teaching and learning. The second is the promotion of school-based teacher professional development in the area of classroom assessment so that teachers can share their assessment practices and ways of coordinating formative and summative assessment, and thereby develop a collaborative assessment culture across classrooms.

References

- Allal, L., 1979. Stratégies d’évaluation formative: conceptions psycho-pédagogiques et modalités d’application. In: Allal, L., Cardinet, J., Perrenoud, P. (Eds.), *L’Evaluation Formative dans un Enseignement Différencié*. Peter Lang, Bern, Switzerland, pp. 130–156.
- Allal, L., 1988. Vers un élargissement de la pédagogie de maîtrise: processus de régulation interactive, rétroactive et proactive. In: Huberman, M. (Ed.), *Assurer la Réussite des Apprentissages Scolaires? Les Propositions de la Pédagogie de Maîtrise*. Delachaux & Niestlé, Neuchâtel, Switzerland, pp. 86–126.

- Allal, L., 2007. Régulation des apprentissages: orientations conceptuelles pour la recherche et la pratique en éducation. In: Allal, L., Mottier Lopez, L. (Eds.), *Régulation des Apprentissages en Situation Scolaire et en Formation*. De Boeck, Brussels, Belgium, pp. 7–23.
- Allal, L., 2020. Assessment and the co-regulation of learning in the classroom. *Assess Educ. Princ. Pol. Pract.* 27, 332–349. <https://doi.org/10.1080/0969594X.2019.1609411>.
- Allal, L., Mottier Lopez, L., 2005. Formative assessment of learning: a review of publications in French. In: OECD (Ed.), *Formative Assessment: Improving Learning in Secondary Classrooms*. OECD, Paris, France, pp. 241–264.
- Allal, L., Pelgrims Ducrey, G., 2000. Assessment *of* – or *in* – the zone of proximal development. *Learn. Instruct.* 10, 137–152.
- Andrade, H.L., 2010. Students as the definitive source of formative assessment: academic self-assessment and the self-regulation of learning. In: Andrade, H.L., Cizek, G.J. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, NY, pp. 90–105.
- Andrade, H.L., 2019. A critical review of research on student self-assessment. *Front. Educ.* 4 (87). <https://doi.org/10.3389/educ.2019.00087>.
- Andrade, H., Brookhart, S.M., 2020. Classroom assessment as the co-regulation of learning. *Assess Educ. Princ. Pol. Pract.* 27, 350–372. <https://doi.org/10.1080/0969594X.2019.1571992>.
- Assessment Reform Group (ARG), 1999. *Assessment for Learning: Beyond the Black Box*. University of Cambridge School of Education, Cambridge, UK.
- Baird, J.-A., Andrich, D., Hopfenbeck, T.N., Stobart, G., 2017. Assessment and learning: fields apart? *Assess Educ. Princ. Pol. Pract.* 24, 317–350.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Bennett, R.E., 2011. Formative assessment: a critical review. *Assess Educ. Princ. Pol. Pract.* 18, 5–25.
- Bennett, R.E., 2015. The changing nature of educational assessment. *Rev. Res. Educ.* 39, 370–407. <https://doi.org/10.3102/0091732X14554179>.
- Black, P., William, D., 1989. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5, 7–74. <https://doi.org/10.1080/0969594890050102>.
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21 (5). <https://doi.org/10.1007/s11092-008-9068-5>.
- Bloom, B.S., 1968. Learning for mastery. *Evaluation Comment (UCLA/CSEIP)* 1 (2), 1–12.
- Bloom, B.S., Hastings, J.T., Madaus, G.F., 1971. *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill, New York, NY.
- Boekaerts, M., 2011. Emotions, emotion regulation, and self-regulation in learning. In: Zimmerman, B.J., Schunk, D.H. (Eds.), *Handbook of Self-Regulation and Performance*. Routledge, New York, NY, pp. 408–425.
- Brandmo, C., Panadero, E., Hopfenbeck, T.N., 2020. Bridging classroom assessment and self-regulated learning. *Assess. Educ.: Princ. Policy Pract.* 27, 319–331. <https://doi.org/10.1080/0969594X.2020.1803589>.
- Brookhart, S.M., 2001. Successful students' formative and summative use of assessment information. *Assess Educ. Princ. Pol. Pract.* 8, 153–169.
- Brown, A.L., 1987. Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In: Weinert, F.E., Kluwe, R.H. (Eds.), *Metacognition, Motivation, and Understanding*. Lawrence Erlbaum, Hillsdale, NJ, pp. 65–116.
- Dann, R., 2014. Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assess Educ. Princ. Pol. Pract.* 21, 149–166. <https://doi.org/10.1080/0969594X.2014.898128>.
- Gipps, C., 1999. Socio-cultural aspects of assessment. *Rev. Res. Educ.* 24, 355–392.
- Hadwin, A., Oshige, M., 2011. Self-regulation, coregulation, and socially shared regulation: exploring perspectives of social in self-regulated learning theory. *Teach. Coll. Rec.* 113, 240–264.
- Harlen, W., 2005. Teachers' summative assessment practices and assessment for learning: tensions and synergies. *Curric. J.* 16, 207–223.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Hickey, D.T., 2015. A situative response to the conundrum of formative assessment. *Assess Educ. Princ. Pol. Pract.* 22, 202–223.
- James, M., 2008. Assessment and learning. In: Swaffield, S. (Ed.), *Unlocking Assessment: Understanding for Reflection and Application*. Routledge, Abingdon, UK, pp. 20–35.
- Laveault, D., Allal, L., 2016. Implementing assessment for learning: theoretical and practical issues. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer International, Cham, Switzerland, pp. 1–18.
- Panadero, E., 2017. A review of self-regulated learning: six models and four directions for research. *Front. Psychol.* 8 (422). <https://doi.org/10.3389/fpsyg.2017.00422>.
- Panadero, E., Andrade, H., Brookhart, S., 2018a. Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *Aust. Educ. Res.* 45, 13–31. <https://doi.org/10.1007/s13384-018-0258-y>.
- Panadero, E., Broadbent, J., Boud, D., Lodge, J.M., 2018b. Using formative assessment to influence self- and co-regulated learning: the role of evaluative judgement. *Eur. J. Psychol. Educ.* 34, 535–557. <https://doi.org/10.1007/s10212-018-0407-8>.
- Panadero, E., Jonsson, A., Strijbos, J.W., 2016. Scaffolding self-regulated learning through self-assessment and peer assessment: guidelines for classroom implementation. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer International, Cham, Switzerland, pp. 311–326.
- Perrenoud, P., 1998. From formative evaluation to a controlled regulation of learning: towards a wider conceptual field. *Assess Educ. Princ. Pol. Pract.* 5, 85–102.
- Pintrich, P., Zusho, A., 2002. The development of academic self-regulation: the role of cognitive and motivational factors. In: Eccles, J., Wigfield, A. (Eds.), *Development of Achievement Motivation*. Academic Press, San Diego, CA, pp. 249–284.
- Scriven, M., 1967. The methodology of evaluation. In: Tyler, R.W., Gagné, R.M., Scriven, M. (Eds.), *Perspectives of Curriculum Evaluation*, vol. 1. Rand McNally, Chicago, IL, pp. 39–83.
- Shepard, L., 2000. The role of assessment in a learning culture. *Educ. Res.* 29 (7), 4–14.
- Tomasik, M.J., Berger, S., Moser, U., 2018. On the development of a computer-based tool for formative student: epistemological, methodological and practical issues. *Front. Psychol.* 9 (2245). <https://doi.org/10.3389/fpsyg.2018.02245>.
- Volet, S., Vaurus, M., Salonen, P., 2009. Self- and social regulation in learning contexts: an integrative perspective. *Educ. Psychol.* 44, 215–226.
- Zimmerman, B., 2000. Attaining self-regulation: a social cognitive perspective. In: Boekaerts, M., Pintrich, P., Zeidner, M. (Eds.), *Handbook of Self-Regulation*. Academic Press, San Diego, CA, pp. 13–39.

Feedback for learning

Rola Ajjawi, Joanna Tai, and Phillip Dawson, Center for Research in Assessment and Digital Learning, Deakin University, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

Conceptualizations of feedback	41
Information transmission	41
Constructivist	42
Sociocultural	42
Trends in feedback research	43
Deliberate and iterative feedback designs	43
Feedback literacy	43
Technology and feedback	44
Trust, relationships and feedback	45
Emerging areas	46
Supporting inclusion and diversity through feedback	46
Internal comparisons	46
Authentic feedback	46
Conclusion	47
References	47

Feedback can powerfully influence learning. Meta-analyses consistently show that feedback has a statistically significant effect on performance, recently revised from a high (0.7) to medium effect (0.48) (Wisniewski et al., 2020). However, across these reviews and others (Evans, 2013), there is significant heterogeneity in the data indicating diversity in feedback practices and effects with almost a third of interventions reportedly having negative effects on learners (Kluger and DeNisi, 1996). Here we grapple with this diversity of feedback interventions through highlighting differences in conceptualization of feedback and how it has influenced practice and research. We then discuss four contemporary trends in feedback research which show promise for improved cognitive (e.g. achievement of learning outcomes, development of evaluative judgment) and affective (e.g. motivation) effects on learners. Finally, we present several emerging ideas from feedback research.

Conceptualizations of feedback

First, we map significant differences, or shifts, in conceptualizations of feedback in research and how these have influenced practice. This is by no means the only way to categorize the field, nor is it meant to be systematic or exhaustive. Instead, this synthesis of significant shifts in conceptualization is a result of years of scholarship in the field and our interpretations and frames of reference. Further, research continues strongly in each of these categories to this day, so while some researchers have “shifted” others maintain important scholarship within these domains.

Information transmission

The “traditional” approach to feedback is one of information transmission. Here feedback is conceptualized as information provided by an agent regarding aspects of one’s performance (Hattie and Timperley, 2007). Research in this tradition focuses on what the teacher does and the form and content of the message (i.e. the feedback). Feedback is thus an “input”. The source of this input may be a significant other, peer, book, self, or teacher, with the majority of research focusing on teacher delivery of feedback information often in written form and often in response to submitted work. Context is presumed mostly irrelevant and therefore advice stemming from such a perspective focus on the message and might include a focus on specificity, timeliness, balanced comments, etc. These criteria serve to help us judge the quality of feedback.

Recent work in this tradition of feedback as information does acknowledge the role of the learner such as Lipnevich and Smith (2018), who defined feedback as any information about a performance that learners can use to improve their performance or learning. However, that feedback actually improves performance is either assumed or hoped for (Boud and Molloy, 2013). Feedback may be *given* at the level of the task, process of how to do a task, regulatory or metacognitive process level, and/or the self or personal level (Hattie and Timperley, 2007). The language of this tradition involves delivery, giving, and metaphors such as feedback as a gift, are common. Students report feedback as something done to them rather than with them; and they are passive recipients if they are acknowledged at all (Boud and Molloy, 2013).

Critique of this conception of feedback abounds, briefly, it includes that information (generally) ignores the role of the learner in the feedback, and there is no certainty about whether following the provision of information it has been read, understood or if the

learner can or is willing to act on it. There is also a lack of attention to further specific opportunities where the student may apply this information. Jonsson (2013) identified five reasons for students not to engage with feedback as information: (1) students do not find it to be useful, (2) it does not meet their needs, (3) they find it difficult to decipher, (4) it serves to justify the grade rather than promote learning, and (5) the lack of strategies for improvement. Proponents of feedback as information might argue that with the “right” information these problems may be addressed. However, those who advocate for more constructivist perspectives on feedback would argue that the focus on information and its delivery is too narrow to achieve the purpose of feedback for learning. Rather than focusing on what the teacher does, there is a need to focus on what the learner does.

Constructivist

The absence of the student in the previous discourse, led some scholars to argue for a more constructivist approach that foregrounds the learner. Definitions of feedback thus shift from an input to a “process” whereby learners obtain information about their work in order to appreciate the similarities and differences between the appropriate standards for any given work, and the qualities of the work itself, in order to generate improved work (Boud and Molloy, 2013). Alternatively, feedback has been defined as “dialogic processes whereby learners make sense of information from various sources and use it to enhance their work or learning strategies” (Carless, 2016). Language in this discourse talks of students engaging with feedback processes or being active recipients of feedback information (Winstone et al., 2017). The teacher’s role includes delivery of information and creating the conditions for students to interpret and use feedback information through deliberate feedback designs. Feedback is thus a *process* of meaning making and action by the learner. The analytical research lens is therefore trained to foreground the learner and examine their responses as they interpret the information in context.

Given the learner’s active and interpretive stance, emotions become more relevant and influence how feedback is interpreted and acted on. Although significant swathes of literature, align negative feedback comments with negative emotions and lack of action (deactivating) and positive feedback with positive emotions and action (activating), some show more nuance. For example, Goetz et al. (2020) work on emotions, drawing from Pekrun (2006), show that positive emotions (e.g. relief) may be deactivating and negative emotions (e.g. anxiety) may be activating. There are also more complicated individual and context interactions, for example a high achieving student who reaches their goal of a pass may choose not to make further revisions despite positive comments (Ajjawi et al., 2022a) or a student who engages with negative comments because they recognize and trust the teacher’s intentions (Telio et al., 2016). Students’ perceptions of feedback and its effects are thus context-dependent, subject-dependent, and individual-dependent (Van Der Kleij and Adie, 2020).

Students are provided greater agency, and pedagogical design as a mechanism for prompting student sense making and use of feedback is also in greater focus. The extreme of this shift might argue that feedback is only so if it makes a difference in student thinking, action or work. In other words, feedback should be judged by its effects on learners (Henderson et al., 2019). Feedback designs touted here include nested tasks, sequenced tasks, formative tasks and dialogue. Feedback “no longer needs to be something that comes just at the end of an assignment but needs to be put into a context and integrated into students’ overall learning” (Orsmond et al., 2005). The recognition of the potential value of peers within feedback processes has also been recognized (Nicol et al., 2014). However, the extent to which the student plays an active role in the feedback process varies even within conceptualizations of feedback that consider the critical role of the student (Van der Kleij et al., 2019).

Critique of the constructivist conceptualization is that feedback should not be judged by its effect as this introduces many variables including student capability, agency, volition, and interest, alongside the design of opportunities to demonstrate an effect. Further, constructivist conceptualizations of feedback that foreground learners’ interpretation blurs the divide between teaching and assessment (Pokorny and Pickford, 2010). From a practical perspective, greater effort and time is needed in designing assessment to promote feedback loops. In addition, students do not have free will to act as they choose, instead their agency is bounded and mediated by social, relational, political and cultural factors and these factors strongly influence feedback practices.

Sociocultural

The third shift is one of feedback as a process of co-construction. Co-construction rests on principles of shared understanding and working (Reusser and Pauli, 2015). Here too feedback is a “process”, but one that is entwined within the cultures which produce it. Core to this conceptualization is the notion of “feedback practices”, that is, the processes, relations and interactions that take place when students and teachers are engaged in feedback interactions (Esterhazy, 2019). Here context is not merely a neutral container that individuals make sense of and control, instead learning cultures/environments fundamentally imbue and shape feedback practices and the ways in which teachers, learners, peers and others respond can be productive or detrimental to learning. Issues of power, relationality, socio-historical factors are enmeshed with what individuals bring, from which feedback unfolds.

Sociocultural theories of feedback conceptualize it as meaning-making trajectory students move along by orienting toward and elaborating on both task-specific and general-knowledge content (Esterhazy and Damşa, 2019). Meaning-making processes are supported by tools and capacities that are available to the students. One way to think of this is that if learning environments are rich with performance relevant information and scaffolding, learning necessarily ensues. Another definition elaborates feedback as “a dynamic and co-constructive interaction in the context of a safe and mutually respectful relationship for the purpose of challenging a learner’s (and educator’s) ways of thinking, acting or being to support growth” (Ajjawi and Regehr, 2019). The relational

dimension of feedback is also accounted for, a recognition that there should be trust, and that this is dynamic and fragile but it influences how teachers and learners interact and enact feedback practices. Emotions are not merely reactions to feedback but influencing of feedback behaviours and indicative of the learning environment (Ajjawi et al., 2022b).

The unit of analysis in this discourse can be the feedback culture, or teacher-student-peer groups, none are prioritized but could be foregrounded at different times with the understanding that they influence each other. Culture can refer to “behaviors, values, norms, beliefs, assumptions and attitudes that are found within a particular organization or context” (Bearman et al., 2021). Research shows that feedback cultures differ between for example school and higher education (Tian and Lowe, 2013), university and workplace settings (Urquhart et al., 2018), and between different professions/disciplines (Watling et al., 2013) or ethnic groups (Tian and Lowe, 2013). Feedback practices common to a particular culture might feel foreign, personal and devastating in another.

Critique of sociocultural approaches to feedback include that, “culture” is an amorphous term, despite regular appeals to culture change, it does need to be approached from a specific theoretical perspective such as through organizational or practice theories (Watling et al., 2020). In addition, defining feedback so narrowly excludes from analysis much of what may be understood elsewhere as feedback. Privileging human relationships places less emphasis on the feedback information learners receive from material sources, both traditional formal feedback (e.g. MCQ feedback) and cues from the environment.

Trends in feedback research

We now synthesize four contemporary trends in feedback that research suggests might improve the effectiveness of feedback interventions on student learning. These four categories cross-cut the above conceptualizations but do undeniably privilege the learner and the learning environment.

Deliberate and iterative feedback designs

The landmark paper that highlighted the importance of feedback design was by Boud and Molloy (2013). They described the need to create opportunities for learners to apply feedback information and to show that performance had indeed changed. The recognition of a need to be able to detect a change in performance was also evident in earlier definitions of feedback: “information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way” (Ramaprasad, 1983).

Findings from a recent realist review suggest that feedback often has a limited impact because of the lack of action and practice that accompanies it, hence the need for cognitive scaffolding through task design (Ajjawi et al., 2022a). An important feature of deliberate feedback design is iteration, that there are multiple/repeat practice opportunities, not just one-off demonstrations of effects (a deliberate closing of the feedback loop). Three design features of impactful learner-centered feedback are: that feedback processes are designed to promote learning, comments invite feedback dialogue, and comments provide actionable information for future performance (Ryan et al., 2020). Designs that promote learning include formative assessment opportunities which can lead to improved performance for those students who are organized and motivated enough to submit (approximately 20–50% of students). This prompted Beaumont et al. (2016) to make the formative task compulsory. A suitable compromise appears to be to make the activities count (i.e., not optional) through sequenced task design. This equalizes opportunities for feedback across students of varying levels of competence.

Students were more likely to use feedback information when tasks were sequenced (Steen-Utheim and Hopfenbeck, 2019). Other examples of tasks that prompt students to become more active in the feedback process include: the use of proformas such as interactive cover pages (Bloxham and Campbell, 2010; Daniel et al., 2015), reflective tasks to process information (Hughes et al., 2015; Todd and McIlroy, 2014), action planning (Adcroft and Willis, 2013), and feedback dialogue (Ajjawi and Boud, 2018). Gamlem and Vattøy (2023) also highlight the importance of dialogue between a student’s internal feedback and feedback provided by external agents. Feedback loops and relevant education about feedback has also been shown to be useful in relation to peer feedback (Mercader et al., 2020). Recent research shows improvements in student engagement with feedback as a result of ipsative design processes (Malecka et al., 2021).

Students are reported to use feedback to identify the teachers’ expectations rather than for learning (Orsmond and Merry, 2011). The best designs can be undermined if students simply believe they are going through a process to satisfy a teacher. They can also find it difficult to transfer learning from feedback comments to other tasks with similar learning outcomes. Iterative feedback designs signal to students that the purpose of feedback is to support learning rather than merely justifying the grade. Hence, scaffolded feedback designs can help students and teachers to reach a common understanding as to the purpose of feedback. Overall, the evidence suggests that there are benefits to trying to make students more active in the feedback process but an awareness of the extent of engagement with feedback across the continuum of ability must be considered in design decisions (Ajjawi et al., 2022a).

Feedback literacy

There has been a growing realization that individuals require capabilities for effective feedback, regardless of which feedback paradigm is adopted, termed feedback literacies. The requisite capabilities were initially introduced by Sutton (2012) who applied an academic literacies approach, which argues for literacy as a social practice undertaken within contexts—more than just to read and

write. The focus of this work was on the promise of feedback literacy to contribute to overall student development in higher education, in how they know, who they are becoming, and how they might act as a result of feedback. Carless and Boud (2018) reframed the concept of feedback literacy as “an understanding of what feedback is and how it can be managed effectively; capacities and dispositions to make productive use of feedback; and appreciation of the roles of teachers and themselves in these processes”. They identified four inter-related features which students needed to do for effective uptake of feedback: appreciating feedback, making judgments, managing affect, and taking action. This framework was expanded upon by Molloy et al. (2020) in an empirical study, to denote more fine-grained behaviors and actions that feedback literate students might exhibit. These behaviors include attitudinal aspects, “commits to feedback as improvement”, “appreciates feedback as an active process”, and “acknowledges feedback as a reciprocal process” in addition to categories that aligned with the previously identified features of student feedback literacy. This framework identified that being a producer of feedback information for others—such as peers—is a key part of feedback literacy.

Researchers have subsequently applied different perspectives to the concept of feedback literacy, to develop an ecological model where the role of environment and context are acknowledged (Chong, 2021), to draw out sociomaterial implications which place further emphasis on the entangled nature of students and teachers with their context-bound environments (Gravett, 2020; Tai et al., 2021), and understand how feedback literacies vary in different disciplinary contexts (Winstone et al., 2020). This work suggests that both students and teachers must be able to detect, understand, and adapt to the various feedback situations they find themselves in, rather than having just one set of feedback capabilities to apply consistently across those different situations. One aspect of feedback literacy has been likened to a related concept in the organizational studies literature, feedback-seeking behavior (Joughin et al., 2021) which, as the nomenclature implies, focuses on the individual within workplace contexts, and what attitudes and dispositions are required to successfully elicit information about their performance. While student feedback literacy was the initial focus of work within education (e.g. Sutton, 2012; Carless and Boud, 2018; Molloy et al., 2020), teachers’ feedback literacy is also increasingly researched, both in their roles as academics (Gravett et al., 2020), and in how they interact with their students (Chan and Luo, 2021; van Heerden, 2020; Xu and Carless, 2017; Carless and Winstone, 2020). Teacher feedback literacy matters as they have considerable influence in designing feedback processes and in setting expectations of students about feedback (Boud and Dawson, 2021).

Across all the work that has been conducted so far on feedback literacy, the key implication is that successful feedback—whether it be the creation of performance-relevant information, or the participation in co-construction of understanding in relation to performance and standards—requires the development of capabilities, and this in turn requires appropriate opportunities for learning. We cannot expect students or teachers in any situation to instantly participate well in feedback for learning without a clear conception of what feedback is for, orientation and practice. Therefore, attention must be paid to the sequencing of learning opportunities about feedback, in addition to crafting the feedback opportunities themselves.

Technology and feedback

There is an expansive research literature on technology and feedback (Dawson et al., 2018), which can be characterized in different ways. One is to focus on the particular types of technology in use. In a recent systematic review, Dawson et al. (2018) identified the main types of feedback technologies as: digital recordings, “bug in ear” technologies (e.g. a Bluetooth headset) that allow real-time feedback dialogues with someone who is not physically located with the learner, automated feedback, and “intelligent tutoring systems” which build a cognitive model of the student and provide feedback information in response to student performance. They found evidence of efficacy for all of these approaches, though caution that ultimately it is the underlying feedback design that will likely have the most impact. Another approach is to consider who is generating the feedback; in that same systematic review the main generators of feedback information were students, their peers, their teachers, and computers. In this section we take a different approach, by considering the main affordances of feedback technology: richer feedback communication through digital media; logistical enablement of more sophisticated feedback designs; and the generation of additional feedback information.

Media switches (such as written to audio or video feedback) are one of the most researched areas at the intersection of technology and feedback. In these approaches, rather than delivering feedback information to students through written comments, alternative media are used. Three media switches that are heavily researched are when educators shift to (1) audio, (2) video or (3) screen cast. Reviews of these different modes concur that students tend to like feedback more when it is delivered by these digital media (Killingback et al., 2019; Mahoney et al., 2018; Bahula and Kay, 2021). In addition, there is evidence of affective and/or relational benefits, as these modes allow richer verbal and non-verbal communication like gesture and tone of voice. Other benefits include greater efficiency for some educators (Henderson and Phillips, 2015), though this comes only once some technical hurdles have been overcome, and greater perceived detail and usefulness of the feedback for students (Ryan et al., 2019).

There is one gap in the feedback media switch literature: robust, replicated evidence that media switches are associated with improvements to learning. This may be partially because many studies have largely focused on student perceptions and experiences (Mahoney et al., 2018; Killingback et al., 2019). More generally in education, when educational designs are held constant but the mode of instruction is changed, educational outcomes tend not to get better (Clark et al., 2005); this has been dubbed the “No Significant Difference Phenomenon” (Russell, 2013). If this effect applies to feedback as well, it suggests that simply implementing the same feedback design through a different mode may not yield significant improvements to learning. However, changes in mode may prompt changes in design, which might lead to improvements to learning.

Technology can also be used to support the logistics of feedback designs that would otherwise be infeasible. For example, bug in ear technologies allow access to outside expertise, and online peer review platforms support the scalability of peer feedback

approaches to thousands of students (Dawson et al., 2018). This logistical enablement also extends to supporting online equivalents of face-to-face learning activities that incorporate feedback, such as discussion (e.g. discussion boards) and collaboration (e.g. wikis). As with feedback media switches, the effectiveness of these approaches in supporting improvements to student learning is under-explored in the literature (Dawson et al., 2018). But as these approaches support the implementation of feedback designs that may usually be infeasible in mass higher education—for example, the 360° feedback approaches used in business contexts (Corrin and Bakharia, 2020)—they are not just doing the same thing, but online. However, as they are new designs requiring more from learners, they also bring challenges that are not just technical. For example, sophisticated online peer feedback designs require a range of non-technical capabilities and attitudes of learners and educators, in addition to capability with technology (Tai and Adachi, 2020).

In addition to being a medium or enabler of feedback, technology can also itself be a provider of feedback information for students. Dawson et al. (2018) identify five types of computer-to-student feedback. At the most rudimentary level, computers can provide pre-programmed feedback information on multiple choice questions, which provides significant support in terms of correcting student misunderstandings. This pre-programmed feedback can also be implemented through a second group of technologies, student response systems, which include approaches like apps or “clickers” that are used by students to respond to in-class questions. The next three types of computer-to-student feedback are more sophisticated as they do not rely as much on simplistic pre-programmed responses to a choice by the student, and instead consider broader work a student has done. Automatic writing evaluation software provides guidance to students on their writing. This category includes consumer tools such as Grammarly which provides grammar guidance, as well as education-specific tools such as AcaWriter which provides task-specific guidance to students based on the rhetorical moves the student makes in their work (Knight et al., 2020). These tools are often also used in the fourth category, Computer Assisted Language Learning. The final category, intelligent tutoring systems, are a cognitivist approach that develops a model about the student based on their performance in different tasks, often multiple choice questions, and tailors future instruction and feedback to that model (Dawson et al., 2018). A general theme across these five categories is that while they can be resource-intensive to set up, they can provide useful supplementary inputs to students at scale. And as with the other categories considered, the overarching feedback design used will heavily influence their effectiveness and uptake by students.

The literature on technology and feedback for students covered thus far is almost exclusively concerned with feedback that is either provided by teachers or designed by them. A future avenue for feedback research may be to investigate the feedback processes students independently undertake. This may include, for example, consulting with peers online, paying for someone to review their work online, or uploading their work into a text-matching tool to identify any potential plagiarism.

Trust, relationships and feedback

Increasingly issues of trust, relationships, care and psychological safety are being recognized as important for motivating students to seek and engage with feedback to change behavior. In 2009, Carless (2009) observed:

Feedback requires lecturers and students to enter into a relationship of trust in which the former try to provide helpful comments that the latter attempt to use. For formative feedback to flourish it is necessary for students to be willing to reveal their own partial conceptions: in other words to invest trust in the teacher. Conversely, “faking good” (Gibbs, 2006, 26) occurs when students present themselves as knowing more than they actually do, for fear that revealing their weaknesses may be used against them (p. 82).

Teachers showing care for their students, influences their self-efficacy and effort regulation and strengthens student perceptions of trust (Price et al., 2011; Steen-Utheim and Hopfenbeck, 2019). Trusting relationships between teachers and students also influences whether students seek feedback and encourages students to talk to their tutor (Crimmins et al., 2016). “Developing positive learning relationships with personal tutors at the point of assessment feedback can encourage a sense of achievement and success at a time when students may feel most vulnerable to low self-esteem” (Cramp, 2011). Telio et al. (2016) suggest that the strength of the relationship as perceived by the student, influences how they interpret the feedback information and whether they use it. Learners were found to make sophisticated judgements about the credibility of the source of the feedback information including whether they had the student’s best intentions at heart. Students are unlikely to act on information provided by someone they perceive as untrustworthy (Telio et al., 2016).

The detrimental impact of a lack of perceived relatedness can result in students’ loss of self-efficacy, feedback avoidance and inability to use feedback comments (Ajjawi et al., 2022a). For example, Hughes et al. (2015) found that where there was minimal relationship building between students and teachers, the role of the tutor for seeking clarification was unclear. Seeing the tutor was equated with failing or having personal problems, rather than improving learning. Further, due to lack of trust, students did not want to reveal weaknesses in seeking clarification (Hughes et al., 2015). How we create learning environments of trust within our constrained and massified university “classrooms” is a fruitful area for research.

Emerging areas

The four trends identified so far (deliberate designs, feedback literacy, technology and trust) all have significant swathes of, for the most part, favorable research. In this section we highlight emerging (nascent) areas of feedback scholarship that might shape future practices and research agendas.

Supporting inclusion and diversity through feedback

We are increasingly seeing students from diverse backgrounds and with diverse characteristics admitted into and participating in higher education. Feedback is likely to be a key ingredient for the success and retention of these students. However, generally students who would most benefit from feedback often do not seek it out (Adcroft and Willis, 2013), and diverse students might also have different expectations for when and how feedback occurs. Adopting the four key messages synthesized in this chapter is likely to support diverse students, acknowledging that different approaches might be needed for different students. For example, research shows that high achieving versus low achieving students approach and use feedback information differently (Ajjawi et al., 2022a). Being explicit about deliberate and iterative feedback designs will help orient students to the goals and processes of feedback. Supporting the development of student feedback literacy capabilities will be important for engagement in feedback processes. Drawing on technology to facilitate multimedia communication can improve accessibility, and may also better account for neurodiverse students, in alignment with Universal Design for Learning (CAST, 2018). Finally, developing trusting relationships will ensure safe spaces for students who might otherwise still feel stigmatized and othered (O'Shea et al., 2016). Overall, opportunities for dialogic and interactive processes will ensure that feedback can be adapted and tailored to students' need, monitoring improvement, and ultimately supporting them to meet their learning goals. Here we clearly see feedback positioned as an opportunity to tailor the curriculum to specific students and their needs.

Internal comparisons

Further extending the idea of student-centeredness in feedback, Nicol (2021) focuses on the "internal feedback" that students generate when they compare their knowledge or performance against reference information. He argues that this process of making inner comparisons is the underlying mechanism at play in all feedback processes. In this understanding, when a student receives comments about their work from a teacher, the actual learning occurs when the student uses this information and compares it to their own knowledge and/or the work they have done. Nicol's model of feedback builds on predominantly cognitive understandings of comparison. He argues that most feedback research focuses on teacher comments, but that if we view the learning process of feedback as one of inner comparisons, we should instead broaden our focus to all information that students might use to make these comparisons.

The implications of Nicol's model for feedback practices are many. This conception of feedback as "internal" moves the field beyond typical assumptions of what constitutes feedback mentioned earlier. Viewing feedback as primarily a process of internal comparison-making and knowledge generation means that educators' roles shift from a focus on providing information to one of designing the circumstances that enable students to make comparisons between their own work and some reference information. This reference information can come from a variety of sources: teachers, peers, rubrics, exemplars, or anywhere students might glean an understanding of quality. In addition to helping students to make comparisons, Nicol also urges educators to help students make their comparisons explicit, deliberate, and mindful, for example, through requiring students to write about their comparisons. Internal feedback may also be activated through external dialogues (Gamlem and Vattøy, 2023).

Authentic feedback

When students graduate, they will continue to encounter feedback throughout their professional lives. This feedback can tend to be different from the feedback they encountered in their studies. In education, feedback information might predominantly take a written, summative form, with sanitized comments carefully crafted to avoid hurt feelings. In contrast, in workplace settings feedback information might be an angry phone call from an upset customer; frequent dialogues with a peer; or a constant stream of comments on social media. The notion of authentic feedback (Dawson et al., 2021) attempts to bridge these two worlds by considering how the feedback practices of a discipline or profession might be incorporated into education. There are five dimensions to authentic feedback in Dawson, Carless and Lee's (2021) framework: realism, which is the extent to which the educational feedback processes represent those in the target discipline or profession; cognitive challenge, which includes aspects such as problem solving and critical thinking; affective challenge, which recognizes that students need to be prepared to face the emotional challenges of workplace feedback; evaluative judgment, which involves students' understandings of quality; and enactment of feedback, which considers the obligations professional employees might have to make use of feedback. Dawson et al. argue that through incorporating these dimensions into feedback practices, educators may better prepare students for feedback in their future working lives.

Implementing authentic feedback involves analyzing the feedback practices used in a discipline or profession and considering how those might be simulated in an education setting. Doing so comes with challenges, the most significant of which relates to scaffolding. For example, first-year university students would likely find the level of affective challenge present in some workplace feedback practices to be far too high, and so they would need inauthentic supports. Authenticity is not inherently a good thing for

feedback; not all workplace feedback approaches are desirable for learning. However, by the end of a degree, students will need to be able to engage in workplace feedback without supports. A gradual incorporation of authentic feedback throughout a degree may be one way to prepare students for their future feedback lives.

An example of authentic and disciplinary feedback research conducted with software engineers, shows feedback as messages that come from the computer, and the iterative debugging processes that students undertake (Esterhazy, 2018). This is very different from what is considered as feedback in clinical placements for allied health professions students, where feedback takes place between mentors and students and is intended to promote psychological safety (Penman et al., 2021). Quinlan and Pitt (2021) summarize this by identifying two key dimensions of disciplinary feedback practices: sources and timings. The self, objects, disciplinary colleagues, and knowledge or service users can all contribute sources of feedback information and ways to interact with feedback. These contextual considerations merit investigation and may contribute further to our understanding of how and why feedback is effective, or ineffective.

Conclusion

In this chapter we have traced the many shifts and turns in conceptualization of feedback research. As can be seen multiple conceptions of feedback abound as research into feedback continues to proliferate. The multiple notions of feedback and its multiple designs may explain the heterogeneity identified in the research surrounding effects. Being clear about which conceptualization of feedback you are situated in matters. We also urge feedback researchers to engage with questions that seek to better understand when feedback works (or doesn't work), for whom, how and under which circumstances – rather than blunt “does it work?” questions. While we agree with Wiliam (2018) and Sadler (2010) that there are “no magic formulas” for the kinds of feedback that are likely to be helpful in any given situation; we do synthesize the following strategies for feedback for learning. That feedback should: (1) position students as active learners, (2) design deliberate and scaffolded iterative processes, and (3) occur within environments of trust and care. Embedding these strategies requires investment at institution, course, staff and student levels. An investment we suggest is worth the effort.

References

- Adcroft, A., Willis, R., 2013. Do those who benefit the most need it the least? A four-year experiment in enquiry-based feedback. *Assess Eval. High Educ.* 38, 803–815.
- Ajjawi, R., Boud, D., 2018. Examining the nature and effects of feedback dialogue. *Assess Eval. High Educ.* 43, 1106–1119.
- Ajjawi, R., Kent, F., Broadbent, J., Tai, J.H.-M., Bearman, M., Boud, D., 2022a. Feedback that works: a realist review of feedback interventions for written tasks. *Stud. High Educ.* 47, 1343–1356.
- Ajjawi, R., Olson, R.E., McNaughton, N., 2022b. Emotion as reflexive practice: a new discourse for feedback practice and research. *Med Educ.* 56 (5), 480–488.
- Ajjawi, R., Regehr, G., 2019. When I say ... feedback. *Med. Educ.* 53, 652–654.
- Bahula, T., Kay, R., 2021. Exploring student perceptions of video-based feedback in higher education: a systematic review of the literature. *J. Higher Educ. Theor. Pract.* 21.
- Bearman, M., Mahoney, P., Tai, J., Castanelli, D., Watling, C., 2021. Invoking culture in medical education research: a critical review and metaphor analysis. *Med. Educ.* 55, 903–911.
- Beaumont, C., Moscrop, C., Canning, S., 2016. Easing the transition from school to HE: scaffolding the development of self-regulated learning through a dialogic approach to feedback. *J. Furth. High. Educ.* 40, 331–350.
- Bloxham, S., Campbell, L., 2010. Generating dialogue in assessment feedback: exploring the use of interactive cover sheets. *Assess Eval. High Educ.* 35, 291–300.
- Boud, D., Dawson, P., 2021. What feedback literate teachers do: an empirically-derived competency framework. *Assess Eval. High Educ.* 1–14.
- Boud, D., Molloy, E., 2013. Rethinking models of feedback for learning: the challenge of design. *Assess Eval. High Educ.* 38, 698–712.
- Carless, D., 2009. Trust, distrust and their impact on assessment reform. *Assess Eval. High Educ.* 34, 79–89.
- Carless, D., 2016. Feedback as dialogue. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Singapore, Singapore.
- Carless, D., Boud, D., 2018. The development of student feedback literacy: enabling uptake of feedback. *Assess Eval. High Educ.* 43, 1315–1325.
- Carless, D., Winstone, N., 2020. Teacher feedback literacy and its interplay with student feedback literacy. *Teach. High. Educ.* 1–14.
- CAST, 2018. Universal Design for Learning Guidelines Version 2.2. <http://udlguidelines.cast.org>.
- Chan, C.K.Y., Luo, J., 2021. Exploring teacher perceptions of different types of ‘feedback practices’ in higher education: implications for teacher feedback literacy. *Assess Eval. High Educ.* 47, 61–76.
- Chong, S.W., 2021. Reconsidering student feedback literacy from an ecological perspective. *Assess Eval. High Educ.* 46, 92–104.
- Clark, R.C., Nguyen, F., Sweller, J., 2005. *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. Pfeiffer, San Francisco.
- Corrin, L., Bakharia, A., 2020. Re-Imagining peer assessment in self-paced online learning environments. In: Bearman, M., Dawson, P., Ajjawi, R., Tai, J., Boud, D. (Eds.), *Re-imagining University Assessment in a Digital World*. Springer.
- Cramp, A., 2011. Developing first-year engagement with written feedback. *Act. Learn. High. Educ.* 12, 113–124.
- Crimmins, G., Nash, G., Oprea, F., Liebergreen, M., Turley, J., Bond, R., Dayton, J., 2016. A written, reflective and dialogic strategy for assessment feedback that can enhance student/teacher relationships. *Assess Eval. High Educ.* 41, 141–153.
- Daniel, F., Gaze, C.M., Braasch, J.L.G., 2015. Writing cover letters that address instructor feedback improves final papers in a research methods course. *Teach. Psychol.* 42, 64–68.
- Dawson, P., Carless, D., Lee, P.P.W., 2021. Authentic feedback: supporting learners to engage in disciplinary feedback practices. *Assess Eval. High Educ.* 46, 286–296.
- Dawson, P., Henderson, M., Ryan, T., Mahoney, P., Boud, D., Phillips, M., Molloy, E., 2018. Technology and feedback design. In: Spector, M.J., Lockee, B.B., Childress, M.D. (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*. Springer International Publishing, Cham.
- Esterhazy, R., 2018. What matters for productive feedback? Disciplinary practices and their relational dynamics. *Assess Eval. High Educ.* 43, 1302–1314.
- Esterhazy, R., 2019. Re-conceptualizing feedback through a sociocultural lens. In: Henderson, M., Ajjawi, R., Boud, D., Molloy, E. (Eds.), *The Impact of Feedback in Higher Education: Improving Assessment Outcomes for Learners*. Springer International Publishing, Cham.

- Esterhazy, R., Damşa, C., 2019. Unpacking the feedback process: an analysis of undergraduate students' interactional meaning-making of feedback comments. *Stud. High Educ.* 44, 260–274.
- Evans, C., 2013. Making sense of assessment feedback in higher education. *Rev. Educ. Res.* 83, 70–120.
- Gamlem, S., Vattøy, K., 2023. Feedback and classroom practice. In: Tierney, R.J., Rizvi, F., Ercikan, K. (Eds.), *International Encyclopedia of Education*. 4th edition, Elsevier, In press.
- Gibbs, G., 2006. How assessment frames student learning. In: Bryan, C., Clegg, K. (Eds.), *Innovative Assessment in Higher Education*. Routledge, London, pp. 23–36.
- Goetz, T., Keller, M.M., Lüdtke, O., Nett, U.E., Lipnevich, A.A., 2020. The dynamics of real-time classroom emotions: appraisals mediate the relation between students' perceptions of teaching and their emotions. *J. Educ. Psychol.* 112, 1243–1260.
- Gravett, K., 2020. Feedback literacies as sociomaterial practice. *Crit. Stud. Educ.* 63, 261–274.
- Gravett, K., Kinchin, I.M., Winstone, N.E., Balloo, K., Heron, M., Hosein, A., Lygo-Baker, S., Medland, E., 2020. The development of academics' feedback literacy: experiences of learning from critical feedback via scholarly peer review. *Assess Eval. High Educ.* 45, 651–665.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Henderson, M., Ajjawi, R., Boud, D., Molloy, E., 2019. Why focus on feedback impact? In: Henderson, M., Ajjawi, R., Boud, D., Molloy, E. (Eds.), *The Impact of Feedback in Higher Education: Improving Assessment Outcomes for Learners*. Springer International Publishing, Cham.
- Henderson, M., Phillips, M., 2015. Video-based feedback on student assessment: scarcely personal. *Australas. J. Educ. Technol.* 31, 51–66.
- Hughes, G., Smith, H., Creese, B., 2015. Not seeing the wood for the trees: developing a feedback analysis tool to explore feed forward in modularised programmes. *Assess Eval. High Educ.* 40, 1079–1094.
- Jonsson, A., 2013. Facilitating productive use of feedback in higher education. *Act. Learn. High. Educ.* 14, 63–76.
- Joughin, G., Boud, D., Dawson, P., Tai, J., 2021. What can higher education learn from feedback seeking behaviour in organisations? Implications for feedback literacy. *Assess Eval. High Educ.* 46, 80–91.
- Killingback, C., Ahmed, O., Williams, J., 2019. 'It was all in your voice' – tertiary student perceptions of alternative feedback modes (audio, video, podcast, and screencast): a qualitative literature review. *Nurse Educ. Today* 72, 32–39.
- Kluger, A.N., Denisi, A., 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.* 119, 254–284.
- Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sandor, A., Kitto, K., 2020. AcaWriter: a learning analytics tool for formative feedback on academic writing. *J. Writ. Res.* 12, 141–186.
- Lipnevich, A.A., Smith, J.K., 2018. *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press, Cambridge.
- Mahoney, P., Macfarlane, S., Ajjawi, R., 2018. A qualitative synthesis of video feedback in higher education. *Teach. High. Educ.* 24, 157–179.
- Malecka, B., Ajjawi, R., Boud, D., Tai, J., 2021. An empirical study of student action from ipsative design of feedback processes. *Assess Eval. High Educ.* 1–15.
- Mercader, C., Ion, G., Díaz-Vicario, A., 2020. Factors influencing students' peer feedback uptake: instructional design matters. *Assess Eval. High Educ.* 45, 1169–1180.
- Molloy, E., Boud, D., Henderson, M., 2020. Developing a learning-centred framework for feedback literacy. *Assess Eval. High Educ.* 45, 527–540.
- Nicol, D., 2021. The power of internal feedback: exploiting natural comparison processes. *Assess Eval. High Educ.* 46, 756–778.
- Nicol, D., Thomson, A., Breslin, C., 2014. Rethinking feedback practices in higher education: a peer review perspective. *Assess Eval. High Educ.* 39, 102–122.
- O'shea, S., Lysaght, P., Roberts, J., Harwood, V., 2016. Shifting the blame in higher education – social inclusion and deficit discourses. *High Educ. Res. Dev.* 35, 322–336.
- Orsmond, P., Merry, S., 2011. Feedback alignment: effective and ineffective links between tutors' and students' understanding of coursework feedback. *Assess Eval. High Educ.* 36, 125–136.
- Orsmond, P., Merry, S., Reiling, K., 2005. Biology students' utilization of tutors' formative feedback: a qualitative interview study. *Assess Eval. High Educ.* 30, 369–386.
- Pekrun, R., 2006. The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341.
- Penman, M., Tai, J., Thompson, T., Thomson, K., 2021. Feedback practices as part of signature pedagogy for clinical placements. *Assess Educ. Princ. Pol. Pract.* 28, 151–169.
- Pokorny, H., Pickford, P., 2010. Complexity, cues and relationships: student perceptions of feedback. *Act. Learn. High. Educ.* 11, 21–30.
- Price, M., Handley, K., Millar, J., 2011. Feedback: focusing attention on engagement. *Stud. High Educ.* 36, 879–896.
- Quinlan, K.M., Pitt, E., 2021. Towards signature assessment and feedback practices: a taxonomy of discipline-specific elements of assessment for learning. *Assess. Educ.: Princ. Policy Pract.* 28 (2), 191–207.
- Ramaprasad, A., 1983. On the definition of feedback. *Behav. Sci.* 28, 4–13.
- Reusser, K., Pauli, C., 2015. Co-constructivism in educational theory and practice. In: Wright, J.D. (Ed.), *International Encyclopedia of the Social & Behavioral Sciences*, second ed. Elsevier, Oxford.
- Russell, T., 2013. The No Significant Difference Phenomenon [Online]. International Distance Education Certification Center. Available: <http://www.nosignificantdifference.org/>. (Accessed 12 July 2013).
- Ryan, T., Henderson, M., Phillips, M., 2019. Feedback modes matter: comparing student perceptions of digital and non-digital feedback modes in higher education. *Br. J. Educ. Technol.* 50, 1507–1523.
- Ryan, T., Henderson, M., Ryan, K., Kennedy, G., 2020. Designing learner-centred text-based feedback: a rapid review and qualitative synthesis. *Assess Eval. High Educ.* 46, 894–912.
- Sadler, D.R., 2010. Beyond feedback: developing student capability in complex appraisal. *Assess Eval. High Educ.* 35, 535–550.
- Steen-Utheim, A., Hopfenbeck, T.N., 2019. To do or not to do with feedback. A study of undergraduate students' engagement and use of feedback within a portfolio assessment design. *Assess Eval. High Educ.* 44, 80–96.
- Sutton, P., 2012. Conceptualizing feedback literacy: knowing, being, and acting. *Innovat. Educ. Teach. Int.* 49, 31–40.
- Tai, J., Adachi, C., 2020. The future of self and peer assessment: are technology or people the key? In: Bearman, M., Dawson, P., Ajjawi, R., Tai, J., Boud, D. (Eds.), *Re-imagining University Assessment in a Digital World*. Springer.
- Tai, J., Bearman, M., Gravett, K., Molloy, E., 2021. Exploring the notion of teacher feedback literacies through the theory of practice architectures. *Assess Eval. High Educ.* 1–13.
- Telio, S., Regehr, G., Ajjawi, R., 2016. Feedback and the educational alliance: examining credibility judgements and their consequences. *Med. Educ.* 50, 933–942.
- Tian, M., Lowe, J., 2013. The role of feedback in cross-cultural learning: a case study of Chinese taught postgraduate students in a UK university. *Assess Eval. High Educ.* 38, 580–598.
- Todd, V.J., McIlroy, D., 2014. Application of formalised developmental feedback for feed-forward to foster student ownership of the learning process. *Psychol. Learn. Teach.* 13, 137–143.
- Urquhart, L.M., Ker, J.S., Rees, C.E., 2018. Exploring the influence of context on feedback at medical school: a video-ethnography study. *Adv. Health Sci. Educ.* 23, 159–186.
- van der Kleij, F., Adie, L., 2020. Towards effective feedback: an investigation of teachers' and students' perceptions of oral feedback in classroom practice. *Assess Educ. Princ. Pol. Pract.* 27, 252–270.
- van der Kleij, F.M., Adie, L.E., Cumming, J.J., 2019. A meta-review of the student role in feedback. *Int. J. Educ. Res.* 98, 303–323.
- van Heerden, M., 2020. (How) do written comments feed-forward? A translation device for developing tutors' feedback-giving literacy. *Innovat. Educ. Teach. Int.* 58, 555–564.
- Watling, C., Driessen, E., van der Vleuten, C.P.M., Vanstone, M., Lingard, L., 2013. Beyond individualism: professional culture and its influence on feedback. *Med. Educ.* 47, 585–594.
- Watling, C.J., Ajjawi, R., Bearman, M., 2020. Approaching culture in medical education: three perspectives. *Med. Educ.* 54, 289–295.

- Wiliam, D., 2018. Feedback at the heart of - but definitely not all of - formative assessment. In: Lipnevich, A.A., Smith, J.K. (Eds.), *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press, Cambridge, UK.
- Winstone, N.E., Balloo, K., Carless, D., 2020. Discipline-specific feedback literacies: a framework for curriculum design. *High Educ.* 83, 57–77.
- Winstone, N.E., Nash, R.A., Parker, M., Rowntree, J., 2017. Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes. *Educ. Psychol.* 52, 17–37.
- Wisniewski, B., Zierer, K., Hattie, J., 2020. The power of feedback revisited: a meta-analysis of educational feedback research. *Front. Psychol.* 10.
- Xu, Y., Carless, D., 2017. 'Only true friends could be cruelly honest': cognitive scaffolding and social-affective support in teacher feedback literacy. *Assess Eval. High Educ.* 42, 1082–1094.

Assessing learning processes

Philip H. Winne, Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

© 2023 Elsevier Ltd. All rights reserved.

Learning is more than what learners know	51
States, events and processes	51
Operational definitions	52
Operationally defining cognitive operations and processes	53
Aspects shared among theories of learning	53
Basic cognitive operations	53
Operationally defining a learning process	54
Complexities attending operational definitions for learning events and processes	55
Agency enacts motivation	55
Exercising agency is self-regulated learning	55
Conditions triggering operations are keys to modeling learning processes	55
Conventional instruments for measuring learning as a process	55
Observations	56
Distal self report: surveys	56
Proximal self report: think aloud	56
Proxies: multimodal measurements	56
Software logs and trace data	57
nStudy: tracing learning processes using fine-grained data	57
Process measurements	59
Assessments of learning events	59
Assessments of learning processes	60
Psychometric issues	60
Conclusions	61
References	61

Glossary

Agency A learner's inherent capability to choose what and how to learn

Cognitive operation Manipulating information in working memory to create a particular product

COPEs A model describing main factors in a learning event: conditions that have bearing on learning, operations performed to learn, product(s) generated by operations, an evaluation of the product, standards used in evaluating the product

Declarative information Information formatted as a statement of a fact, belief or step in a procedure

Event A segment of time within which a behavior or operation happens

IF-THEN procedure A model of a learning tactic; IF particular conditions are present, THEN perform a particular learning tactic

Learning What is known; a process for adding to and modifying knowledge, beliefs and preferences

Learning strategy A pattern of learning tactics organized to guide choosing the next learning tactic based on the state produced by previously executed learning processes

Learning tactic The basic unit in a learning event whereby a particular cognitive operation is applied to particular information

nStudy A software system designed to record trace data and develop learning analytics that support learners and advance research on learning

Operational definition An explicit, unambiguous and thorough specification of conditions for reliably recording information or observing a process

Procedural information Information describing an action. When procedural information is enacted, information is transformed into a procedure

Process A sequence of events organized to achieve a goal

Self-regulated learning Exercising choice to select and adapt learning tactics and learning strategies in ways a learner believes has benefits according to the learner's standards

SMART A set of basic cognitive operations used in learning: searching, monitoring, assembling, rehearsing, translating

State A particular condition of the learner or a system at a point in time.

Trace data Information recorded about learning events according to an operational definition that provides a valid account of a theorized learning event or process

Approaches to assessing learning focus on methods for gathering, analyzing and interpreting information about two broad clusters: a learner's knowledge, and psychological factors—predominantly motivation, beliefs and affect—that moderate whether, how and what a learner learns. Assessments of knowledge and other individual differences describe the state of a learner at one point in time. For example: Does the learner know the capitol of Canada? Is the learner currently motivated to research global warming? If offered a choice, would the learner choose to study chemistry or history? Other entries in this encyclopedia comprehensively examine approaches to these conventional assessments.

Learning is also a process. Unlike states of knowledge and current beliefs or preferences, processes are dynamic. They unfold across time. In some ways, assessing learning as a process parallels assessing observable skills. Was a piano student's keyboarding of a melody correct and expressive? Was a second-language learner's use of idioms appropriate when asking for directions to a landmark? In cases like these, process are directly observable. In contrast, assessing learning as a process requires inferring latent cognitive and motivational factors.

A first step toward generating high-quality assessments is understanding what is to be assessed. The next four sections build a foundation for conceptualizing processes involved in learning. Following sections examine issues relating to assessing learning as a process.

Learning is more than what learners know

Because knowledge is a social construction, identifying what a learners knows is nuanced. Some knowledge is factual in the sense information is widely agreed to be valid or true. An example is: Water freezes at 0 °C. Other information a learner knows may be relative to a particular social frame of reference, a belief or a preference. Examples are: My university is top ranked. Statistics are always biased. Astronomy is interesting.

Instruments used in assessing knowledge conventionally present statements or tasks on paper or a computer screen. Several formats are common. Selection items include multiple-choice and matching exercises. Supply formats include fill-in-the-blank and essay items. Beliefs and preferences are typically assessed by describing a situation or presenting a claim, then inviting the respondent to select from a provided scale how strongly, how often or how similarly a description applies to them. In these cases, responses are interpreted to reflect the magnitude of an attribute at a point in time and in the context of instructions (e.g., Solve each problem to the best of your ability) or a brief description of a situation (In your economics course this semester, ...).

Assessments developed from learners' responses to these kinds of instruments represent information about the state of a learner in a particular context at a particular time. The range of states is wide. Examples include: Can a learner identify a graph of a parabola? Can the learner write the quadratic formula? Is the learner able to compute roots of a quadratic expression? Can the learner plot a parabola? Is the learner interested to investigate applications of parabolic forms in astronomy and the automobile industry? Does the learner estimate they can explain why graphs of parabolas differ?

Assessments of knowledge sometimes are based on observations of learners as they perform a task. Teachers may ask, "Please show me how you did that." A golf coach may watch a player's driving backswing. Violin teachers inspect their student's bowing. Records of behaviors made by video or audio machines often have a role in accounts of learners' states of knowledge.

In contrast to these approaches to assessing states of knowledge (or belief or preference), what kinds of information can serve in assessing how states arise or and change? What issues bear on using these kinds of information to assess learning as a process? What happens while a learner is learning? Which learning activities are effective?

A first step toward answers to these kinds of questions about assessing learning as a process is to clarify basic concepts.

States, events and processes

Data gathered using conventional assessment instruments primarily take form as factual propositions that cognitive psychologists label declarative information. Declarative information can be expressed in various formats including: definitions (Clock: an instrument measuring time.), a rule that applies if a condition is present (e.g., If temperature falls below 0 °C, water will freeze.), a list (e.g., Vowels in the English alphabet are a, e, i, o and u.), a table (Mendeleev's periodic table of chemical elements), a graph (e.g., the normal curve), and symbolic expressions (e.g., a molecule of water, H₂O; the formula for momentum, $p = mv$). Topics represented by declarative information are as broad as human experience: school subjects, interests, preferences, and motivations.

Some declarative information can be enacted. A learner can demonstrate knowledge about straight lines by plotting a graph of $y = 4x - 3$, debugging lines of computer code, finding a bus stop by following directions, safely lighting a Bunsen burner or dancing a samba. These expressions of knowledge represent a second kind of information, procedural information. Procedures almost always have two representations. In declarative form, procedures are a list of steps sometimes augmented by conditions for whether a step is taken (e.g., If the light is red, stop. If the light is green, go). When steps are enacted, the declarative form of a procedure is transformed into action.

Enacting one step in a procedure animates knowledge and creates an event. A series of events forms a process. Two key features differentiate declarative knowledge per se from events and processes. First, when a learner expresses declarative knowledge by making a selection from a computer menu, writing or typing information, that event renders observable what a learner knows, believes or prefers. Second, events can be characterized by features not applicable to declarative knowledge. One example is lag, the time between a prior event and a focal event. For instance, how long after reading an instruction to plot a parabola using a given formula does a learner make the first mark on a Cartesian grid?

Another important feature differentiates declarative information from a process. A single expression of declarative knowledge can be created by different processes. Is 11 a prime number? A learner may directly recall the answer, "Yes." Another learner may recall the definition of a prime number (a number divisible with no remainder only by 1 and by itself), test whether 11 can be divided evenly by each number between 2 and 10 and, tallying each division has no remainder, then answer "Yes." A third learner may use a search engine on a smartphone to answer, "Yes." The first two learners display the declarative knowledge through different processes. The third learner appears to know 11 is a prime number, but that may be an invalid inference. The third learner knows how to search for information but, even after answering "Yes" may not have learned 11 is a prime number. While each of these three learners expresses the same answer, the processes generating answers were very different.

To summarize, assessments of declarative and procedural knowledge involve presenting a learner with a cue and providing an opportunity for the learner to make a declaration or perform a process. Common formats for expressing declarative information are words and larger structures (sentences, essays), symbols (e.g., Na for the chemical element sodium) and structured arrangements of symbols (e.g., equations), ratings (e.g., a number selected from a provided set of options ordered on a scale) and visualizations (e.g., graphs, diagrams). Assessments of processes almost always can include declarative forms of expression. But, because a process has qualities beyond declarative expressions of knowledge, instruments for recording events and structured organizations of events making up processes typically allow gathering information about additional characteristics, such as temporal attributes.

Operational definitions

What events and patterns of events can account for learning? Answers to this question are the topics of many and varied theories of learning. Theories guide designs for instruments used to gather data, methods for analyzing data gathered and guidelines for interpreting results of analyses. Bridging from theory to an interpretation about learning as a process spans these waypoints of instrumentation, analysis and inference. A first pillar supporting that bridge answers a challenging question: How is a learning event operationally defined?

An operational definition is an explicit, unambiguous and thorough specification of conditions for reliably recording a datum. That is, an operational definition specifies how to measure something. For example, to measure a construct like effort to learn, consider this simplified case. Factors within square brackets are conditions that can be varied.

- One unit of nonsense is a consonant-vowel-consonant pseudoword (Examples are yat, bek, kib.)
- A pseudoword is displayed on a computer screen for $[d]$ seconds after which it disappears.
- Immediately after a pseudoword disappears, the computer screen displays: As rapidly as you can, please [do task t].
- At a lag of $[l]$ seconds after disappearance of the pseudoword, recall of a pseudoword is prompted with this instruction: "Please [do task r]."
- Effort will be measured by [measurement m].
- To assess effort, the measure of effort will be compared to [standard s].

Operational definitions that set a stage to measure effort can vary. One variation is the display interval, d . Another variation alters the intervening task, for example, tapping an index finger as fast as possible versus recalling words of four letters ending with the letter r (door, moor, soar ...). The recall task also can be varied, for example, "Please spell the pseudoword" or "Please spell the pseudoword backward" (yat → tay, bek → keb). The scale for measuring can vary. For example, the measurement might ask the learner to judge effort on a 5-point range anchored by "very easy" to "very difficult." For this measurement, the standard is whatever the learner chooses. Alternatively, instruments gauging dilation of the learner's pupil many measure this physiological indicator of effort (van der Wel and van Steenbergen, 2018). Standards for pupil dilation might be compared to a reference measurement of dilation when the learner views actual words, e.g., "cat" and "tan."

Each variation creates a different operational definition which reflects theoretically important differences. Theories suggest: Longer display times increase opportunity to rehearse information in working memory which increases recall. More cognitively complex intervening tasks interfere with rehearsing the pseudoword during the lag interval. Recalling the pseudoword

backward adds an additional cognitive operation beyond simply recalling the pseudoword. According to theory, each variation differently affects a learner's perception of effort or could affect the measurement of pupil dilation. A further question is how strongly the learner's subjective rating of effort correlates with effort indicated by pupil dilation, an objective measurement of effort.

This simplistic example of assessing effort to learn illustrates central issues relating to every operational definition of learning. First, what task will the learner complete and what are properties of the product(s) of the measurement of that product are proposed as evidence of learning? The example illustrates two measurements: verbatim recall and, for lack of a better label, transformed recall (recalling backwards). Both are objectively scoreable as correct or incorrect. Second, assuming unobserved cognitive events and processes are responsible for creating achievements, what does theory identify as those events and how does theory describe those events are patterned to form learning processes? Finally, interference can stymie events and obstruct or distort processes. What instrumentation will be used to observe and, if possible, quantify factors affecting whether events occur and how processes unfold? Answering these critical questions provides guidance for investigating the reliability of measurements that underlie assessments of learning as a process and the validity of interpretations about learning as a process.

Operationally defining cognitive operations and processes

A useful survey of the landscape of theories of learning is the serial *Annual Review of Psychology*, first published in 1950. Distilling the massive body of scholarship about learning as a process cannot be done justice here. With that caveat, this section introduces a typology of cognitive operations used here to set a stage for considering how to assess learning as a process.

Aspects shared among theories of learning

Excepting for the family of stalwartly behaviorist theories, which reject the scientific value of theorizing about unobservable mechanisms, cognitively-oriented theories of learning share, in varying expressions, three basic axioms. First, when learners learn, unobservable mental operations manipulate information. For example, the operation of rehearsing reinstates in working memory a practically identical copy of whatever original information is rehearsed.

Theories in the cognitive family share an assumption that learners perform operations in a hypothetical "place" in the mind called working memory. Often, but not always, operating on information in working memory increases the probability it will be stored in a second hypothetical region of memory called long-term memory. Importantly, not all information stored in long-term memory is easily or, perhaps, ever retrievable. Cognitively manipulating information in working memory does not guarantee it will be learned.

Second, no matter the topic or the format (declarative or procedural) of information, every unit of information is related to other units of information. For example, a student of chemistry knows the symbols H_2O represent one molecule of water; and the H, the O and the subscript "2" signify a water molecule consists of two hydrogen items and one oxygen atom. Information that cannot be related to knowledge stored in long-term memory is nonsense. Nonsense can be learned but not understood until it is related in some way to knowledge.

Third, an inherent property of the human cognitive system is, when one unit of information is active in working memory, other units of information related to it automatically become activated although likely with varying "strengths" of activation. The chemistry student just mentioned likely realizes the bond between the hydrogen and oxygen atoms is covalent although that fact is not evident in the symbols H_2O . Other knowledge previously learned may have less probability of being realized, perhaps that water is practically incompressible and it can have a form called heavy water, 2H_2O .

Learning happens when (a) particular units of information in working memory (b) are operated on (c) under particular conditions. One key ingredient when assessing learning as a process is operationally defining operations learners use to learn, the topic of the next section.

Basic cognitive operations

Theories of learning in the cognitive family posit cognitive operations are mechanisms that manipulate information. Because theories vary widely, the "basic" set of five operations presented here is heuristic. While this set of five operations is characteristic of theories in general, a particular operation described here may slightly misrepresent properties that make a theory distinctive. The five cognitive operations are: searching, monitoring, assembling, rehearsing and translating. They are conveniently grouped using the first letter of each operation to assemble an acronym, SMART.

Searching takes advantage of the spreading activation property of human memory by which knowledge in long-term memory that is not presently active in working memory becomes active in working memory. Usually, this is labeled recall or retrieval. To search, a learner focuses attention on one or several units of information in working memory. For example, what animal is a pet, has fur and purrs when petted? Individually, each of these standards for filtering knowledge in long-term memory is associated to a wide variety of knowledge. Attending to them as a set reduces members of that large collection to a particular product: cat.

Monitoring is an operation that compares a focal bit of information to one or multiple standards. Monitoring has two inputs: the standards and a bit of information to be examined using the standards. For example, is $y = ax + b$ a linear equation? This equation is inspected to judge whether it exhibits two properties: Is one quantity (y) equated to another (the sum of ax and b)? Does the variable quantity (x) have an exponent of 1? The product of those comparisons is a judgment: The expression is a linear equation.

Assembling creates meaningful links among bits of information to form a larger, more complex structure of information. Inputs are the bits to be assembled. The product is the new, organized information consisting of the inputs and their meaningful relationship(s). Some assemblies are simple: A sentence has a subject and a verb. Other assemblies are more complex. SMART is an assembly. A procedure to safely light a Bunsen burner can be assembled from individual steps placed in a particular serial order. A conceptual schema can assemble characteristics or properties describing a large collection of instances that vary in particular features. For example, a schema for an undergraduate student includes characteristics, or slots, such as: age, academic major, and institution of enrollment. Each slot in this schema can be filled with a unique value to describe a particular undergraduate student, a 20-year old studying educational psychology at Simon Fraser University or a 35-year old studying accounting from an online university.

Rehearsing, as noted earlier, is a cognitive operation that reinstates a practically identical copy of original information in working memory. Input is a unit of information. Output is the copy of that information.

Translating is the operation that changes characteristics of an input to have a different form. Translations preserve critical characteristics of the input and usually introduce new some new characteristics. A paraphrased sentence is a translation of the original. Because the paraphrased sentence uses different words to convey basically the same meaning, the new sentence also introduces new meaning. For example, there are nuanced differences if "This movie is disturbing." is translated "This film frightened me." Translations can change the format of information. The linear equation symbolized as $y = x + 3$ can be translated as a graph. The graph explicitly shows how the line rises at a 45 degree angle, a property not explicitly conveyed by the symbolic form (but, perhaps, readily recalled by an expert in graphing linear expressions). A procedure written as a set of instructions can be translated into a performance, as you may have experienced when assembling furniture from instructions. Listening to a ballad played by a skilled violinist invites a range of different interpretations than does reading the ballad's score.

Operationally defining a learning process

To learn, a learner applies one or a strategically arranged set of cognitive operations to information. Because cognitive operations cannot be directly observed, a first step toward assessing learning as a process is to operationally define cognitive operations. The operational definition is a bridge joining a theory about how learning happens to measurements representing learning events.

Perhaps the most readily observed event in learning is marking information. Learners everywhere highlight or underline text, circle a region in a diagram or chart, and draw an arrow pointing to a notable feature in a table or graph. To select particular information for marking, it must be distinguished from other information. We can infer from such markings the learner monitored information while studying. When particular information satisfied standards—this is important or surprising or merits further research—a mark was applied to permanently record the product of monitoring. Presumably, making this mark helps to find this special information again. Perhaps highlighted information is marked for review before exams.

Selective marking of information is an operational definition for the cognitive operation of monitoring. Input to the operation is information within the learner's purview. Output is a subset of that information that is marked.

What is missing from this operational definition is clarity about standards the learner used to monitor information for marking. Suppose after highlighting some text on the page of a textbook, a learner pens a ? in the margin near the selected text. A plausible inference is the learner experienced some confusion about the meaning or importance of information provided by the selected text. The ? is a signal about standards the learner used to monitor information. Some learners have elaborate systems for tagging selections they mark. For example, different colors used for highlighting may distinguish definitions, useful examples, and material needing additional attention to be understood. Symbols can represent standards such as confused (?), understood (✓), surprising (!) and important (*).

Cognitive operations can have multiple operational definitions, and some may involve a complex of activity and even appear counterintuitive. Consider this case: A learner using a laptop is studying a webpage about asteroids. Among other information, the page briefly mentions a category of asteroids called trojans that orbit at the L4 and L5 Lagrange points. The learner selects and copies the text "Lagrange points," pastes that selection into a search box, and presses the enter key to initiate a search of the internet for more information about Lagrange points. What do these measures imply about cognitive operations used in learning? What do they operationally define?

Selecting the text "Lagrange points" supports an inference the learner monitored this information. At this point, standards used to monitor are not apparent. The next observable event was the learner's search of the internet for more information about Lagrange points. This supports an inference about a latent process. The learner searched long-term memory for knowledge about Lagrange points, monitored what was retrieved using standards about how much was known and with what certainty, and judged prior knowledge was lacking. Because software can record these events—selecting, copying, pasting into a search box—and the exact text selected from the source and pasted into the search box—Lagrange points—it is possible to infer the what information the learner was processing.

While events comprising a learning process are measured, the precise standards the learner had in mind when monitoring are not well supported. To reveal those, the software might provide the learner with a tool to tag selections. Suppose this learner tagged the selected text “investigate.” Adding this event to the operational definition extends grounds for inferring learner’s plan: seek more information about Lagrange points to assemble with information the current webpage presents and what was already known about Lagrange points. Altogether, this pattern of measurements is an operational definition for a “small-scale” strategy for learning. Moreover, because the learner was willing to do this work, it can be inferred there is motivation to extend knowledge by searching the internet.

Complexities attending operational definitions for learning events and processes

Before describing operational definitions for particular cognitive and motivational events, it is useful to consider three issues.

Agency enacts motivation

Underlying almost all theories of learning and motivation is a subtle presumption of agency. Agency means learners have choices about how they will behave. They govern how they engage in learning by choices they make. In deciding how to engage cognitively, learners balance what they perceive are relative costs and benefits of enacting a particular choice about how to learn. Embedded in this choice are two basic motivational constructs (Bandura, 1997). An outcome expectation is the learner’s best forecast about what the product will be if a particular operation is enacted and the consequences of choosing that product as a goal. For example, does an argumentative essay sprinkled with many examples conform to requirements of the assignment? Will the instructor praise an essay with those features? An efficacy expectation is the learner’s best prediction about the extent to which (a) they have skill(s) and (b) environmental conditions allow successfully engaging processes to create that product. Building on the preceding example, a learner judges if they are skilled at creating examples.

Assuming agency in learning leads to an important corollary. How learners behave can reveal what they value—their goals—and what they are willing to do to reach goals. Thus, choices learners make about learning processes reflect motivation. Operational definitions for learning events and learning processes also are operational definitions about motivation.

Exercising agency is self-regulated learning

Because learners choose information on which they will operate and operations they will apply, the locus of learning processes is the learner. Assuming learners are rational—they try to manage learning processes to reach their goals and benefit their interests—learners are regulating their learning every time they exercise agency. Combining these propositions leads to a conclusion that learners are ubiquitously self-regulating learners (Winne, 1995a, 2018). Assessing learning as a process therefore equates to assessing self-regulated learning.

Conditions triggering operations are keys to modeling learning processes

The earlier analysis of a monitoring event hints at a subtle property of cognition which operational definitions can distort. When human beings are awake, they appear to be thinking constantly. New thoughts are triggered by changes in a learner’s environment and spreading activation that retrieves knowledge used to make sense of each instant of experience. As a result, working memory is provided a constant flow of new information from the environment, refreshed knowledge activated from long-term memory, and updates to and new assemblies involving these sources of information.

In contrast to this continuous flow of thinking, operational definitions refer to a segment of experience. Each segment corresponding to an operationally defined event is like a “dot” on a continuously updated timeline of learning processes. To connect these dots, it is helpful to conceptualize a segment as a contingency: *If* conditions in a segment have a particular profile established in the immediate past, *Then* the learner chooses operations for the next segment taking into account those conditions. This representation of thinking as *If-Then* allows modeling cognition as a flow of successive segments whereby history shapes the future.

A critical issue when designing operational definitions is whether an observer’s account of conditions correspond to conditions the learner perceives. The preceding example of monitoring is illustrative. If the operational definition did not provide a software tool the learner could use to tag a selection, an observer might misinterpret why the learner was searching the internet. By supplementing the operational definition with instrumentation the learner could use to tag the selection, the learner can reduce misinterpretations by an observer.

Conventional instruments for measuring learning as a process

Other entries in this encyclopedia comprehensively describe instrumentation, measurement protocols and analyses of measurements gathered by conventional instruments—surveys, observations of performance, and think aloud—to describe and assess learning. The next sections briefly recap material those entries provide a contrast to focusing on assessing learning as a process using computer software.

Observations

Observation systems typically provide the observer with a list behaviors. Each is labeled, e.g., “paying attention” or briefly described, e.g., “The student attends to teacher writing on the chalkboard.” Within a specific time period or within the boundaries of a learning activity, such as a lesson or class period, the observer tallies whether the behavior occurred or counts occurrences. Accounts describing variations in conditions surrounding the behavior may supplement these tallies.

After the observation period ends, a profile of learning processes can be developed from these records to indicate presence or absence of a behavior, or a count of occurrences.

Topics observed span the spectrum of labels describing cognitive, metacognitive, motivational and emotional events. Examples include: paying attention, mind wandering, judging understanding, interest and frustration. Observations may be time-stamped to allow sequencing events across the timeline of the observation period.

Distal self report: surveys

Learners can describe properties of learning processes in surveys administered on paper or by computer. Commonly, the survey begins with a description about a context the learner is to recall or imagine. Examples are; “In this course ...” or “When you study for a test” Relative to that context, items on the survey describe a learning process in behavioral terms, e.g., underlining or re-viewing a flashcard; or as a cognitive process, e.g., monitoring understanding or assembling ideas as a summary. Response formats for describing the learning process may be yes or no, or scales for rating frequency with which the process is typically applied, how important the process is in the given context, or how validly the practice describes the learner’s usual approach to learning in that context.

Almost all surveys present a set of related items, a subscale, the designer of the survey considers to represent a single learning process. Individual items on a subscale differ in what the survey designer considers superficial or incidental properties of context. For example, rehearsing with flashcards is considered the same cognitive operation as rehearsing by re-reading a paragraph in a textbook then covering it with a piece of paper. Responses to items in a subscale are summed or averaged to describe the extent of the learning process.

Proximal self report: think aloud

A think aloud protocol is a common technique used to gather data about learning processes. A learner is instructed to talk about what they think about as they engage with a learning task. Normally, learners are not coached about terms to use or what a researcher or assessor hopes or expects the learner will describe. In some research studies, participants first watch a model perform a think aloud activity or have an opportunity to try out thinking aloud in a practice context. During the learning task when the think aloud protocol is in force, it is often the case learners need prompting to continue thinking aloud while working on their task.

The learner’s uttered thoughts are recorded and transcribed. In some assessments, qualitative methods are used to find categories of learning processes, and reports are provided about the presence and relative prevalence of events and processes. In other assessments, transcripts are analyzed using a predefined system describing learning processes. Typically, instances of particular learning processes are tallied. In some cases, a timeline is developed to illustrate transitions across learning processes and to search for patterns of transitions.

Because learners rarely are experts in learning science or psychology, their descriptions usually are much less precise than experts’ wording on observation schedules and survey items. Consequently, researchers sometimes must infer what a particular learner’s phrasing means and generalize over differences in learners’ descriptions to categorize processes.

Proxies: multimodal measurements

Several technologies have gained popularity for gathering proxy indicators of learning processes. Gaze tracking equipment measures which distinct regions, called areas of interest, a learner visually inspects in an information display, the duration of each inspection and paths traversed across content as gaze travels from one area to another. Sustained inspection is interpreted as focusing attention although it is not possible to more specifically characterize how information is operated on once attended to. Paths of a learner’s gaze across areas of interest track how a learner refocuses attention and present similar difficulty in inferring operations more specifically. For example, oscillating between two areas of interest may represent monitoring information in one area of interest using information in another as a standard. Or, it may represent translating information expressed in one medium, such as text or an equation, presented in different format, such as diagram or a graph.

Measurements of a learner’s physiological states have been used to indicate intensity of cognitive processing at a point in time. Examples are heart rate, blood pressure and pupil dilation. While intensity of processing is inferred, the particular processes being engaged are not directly identifiable.

Face recognition technologies measure the configuration of several loci on a learner’s face. To develop norms, norming groups are shown images strongly expected to produce specific emotions responses. For example, an image of a large dog growling viciously is expected to produce a fear response. To infer the particular emotion a learner is experiencing in a particular context, these normed configurations are compared to a particular learner’s.

All these forms of multimodal data are time stamped. When aligned with other records, e.g., a video of screen activity on a computer, the information a learner was attending to or manipulating can be placed into correspondence with a multimodal measurement or a pattern of multimodal data. Also, a timeline can be assembled describing durations and points of transitions across constructs inferred from modal data.

Software logs and trace data

Computer software can record five basic events for constructing a timestamped record of learning processes. Four of these events are generated by a learner interacting with features of the software: clicking a feature in the software's interface, such as a button or an option in a menu list; dragging a cursor across text or diagonally across a region in an image (picture, diagram, graph) to frame it; typing text or symbols into a field; and simply moving the cursor across the display. Combinations of these events are common, e.g., clicking and holding the handle in a scroll bar, then sliding it to scroll. These events are meaningful if learners understand text labeling a feature in the software's interface (e.g., words such as *Open*, *Paste*, *Sidebar*), read and understand just-in-case documentation or just-in-time popup tooltips, and correctly interpret icons and their functions, e.g., clicking ↺ replaces information currently displayed with new information according to a developer's or author's design. The number of actions executed by the software can be extended when the learner depresses one or more modifier keys such as ⌘.

A fifth, wide-ranging class of events a computer can record are actions the software performs. For example, a tooltip may appear if the learner hovers the cursor over a feature. This is equivalent to a click event triggered by when the cursor is positioned for a specified time interval. Wikipedia uses this method to display the lead paragraph of hyperlinked pages when a learner hovers over blue styled text that is a signal for a hyperlink.

By carefully engineering software functions, labeling features and designing icons in the software interface, single events and patterns of events can trace learning processes. If data logged about events have a strong coupling to a theoretical construct, the resulting measurements are trace data (Winne, 2020). For example, in the nStudy software system (Winne et al., 2019), learners can select text or a region in a diagram, then create a tag to catalog that information. Tags can describe a type of information—e.g., example, explanation, equation—or a task—e.g., investigate, review, analyze. Tags are efficient means for retrieving information in a particular category. When a learner tags in nStudy, the software records the information selected, the tag the learner applied to that selection, and a timestamp for this event.

Tagging events trace a learning process. The click-and-drag event signals the learner monitored the information selected. The tag applied traces standards used in monitoring, e.g., information was classified as an example or meriting analysis. Assuming the learner also knows how to use tags to retrieve types of information or list reminders, tagging also signals planning about future engagement as well as motivation to apply that plan. When the tag names a task, such as review or investigate, there are strong theoretical grounds for inferring the learner has a plan for a future learning activity. For example, the tag "review" allows inferring the learner forecasts the tagged information will be important for some future task and it will not be fully or accurately recalled at that time. By tagging information, the learner offloads storage of information to the software.

nStudy: tracing learning processes using fine-grained data

To illustrate software instrumentation for gathering data about learning processes, several scenarios illustrate a learner's learning activities while using the nStudy software system (Winne et al., 2019). Fig. 1 shows two screenshots of nStudy's STUDY VIEW in a sequence spanning a few seconds of time. Primary learning activities in the STUDY VIEW are viewing and annotating information to be learned or, if a source was previously annotated, the STUDY VIEW supports review.

In the first panel, the learner is selecting information about outgassing in a comet's tail. When the learner lifts their finger from the trackpad (or stops depressing the button on a mouse), nStudy automatically displayed the left-side contextual menu listing operations the learner can apply to this selection. Not shown in the panel is a learner click on the menu option Create note. That click causes nStudy to display the right-side of the contextual menu where the learner has clicked Explore. That action causes nStudy to open the window shown in panel 2. The window is a blank form for creating an annotation. The note form is a scaffold listing slots in a schema for annotating the selected information in a way that explores its features.

This sequence of actions traces several learning events comprising a learning process. In panel 1, nStudy logs the selection of text as a trace of the learner monitoring information in this webpage. If information matches standards, THEN select it. From reading nStudy's documentation or experience with the software, the learner knows selecting the Create note option in nStudy's contextual menu will open the next menu where options for choosing a note form are available. When the learner chooses the schema for assembling information available in an Explore note, this traces standards the learner used when monitoring information selected in the webpage.

These measurements also trace motivated behavior. The learner has made specific choices among options nStudy provides. Choosing traces motivation. Performing these activities requires effort by the learner. That the learner was willing to spend effort to use tools nStudy provides traces the learner's judgment about a positive incentive associated with this learning process.

Panel 2 shows in the upper right corner of the note form the title the learner applied to this Explore note: "Tails of comets." nStudy automatically inserted text the learner selected as a hyperlink in this note. This is not a trace of rehearsing as the software,

The figure consists of two screenshots of the nStudy interface, showing a learner's study view. Both screenshots display the Wikipedia article "Comet" in the background. The interface includes a sidebar on the left with search and navigation options, and a main content area on the right where a note is being created.

Top Screenshot: The note titled "Comet" is being created. The text of the note is: "This article is about the astronomical objects. For other uses, see *Comet (disambiguation)*. A comet is an icy, small Solar System body that, when passing close to the Sun, warms and begins to release gases, a process that is called outgassing. This produces a visible atmosphere or coma, and sometimes also a tail. These phenomena are due to the effects of solar radiation and the solar wind acting upon the comet. Comet nuclei range from a few hundred meters to tens of kilometers across and are composed of loose collections of ice, dust, and small rocky particles. The coma may be up to 15 times Earth's diameter, while the tail may stretch for millions of kilometers (1 astronomical unit). If sufficiently bright, a comet may be seen from Earth without the aid of a telescope and may subtend an arc of 30° (60 Moons) across the sky. Comets have been observed and recorded since ancient times by many cultures. Comets usually have highly eccentric elliptical orbits, and they have a wide range of orbital periods, ranging from several years to potentially several millions of years. Short-period comets originate in the Kuiper belt or its associated scattered disk, and lie beyond the orbit of Neptune. Long-period comets are thought to originate in the Oort cloud, a spherical cloud of icy bodies extending from outside the Kuiper belt to halfway to the nearest star.^[1] Long-period comets are set in motion towards the Sun from the Oort cloud by gravitational perturbations caused by passing stars and the galactic tide. Hyperbolic comets may pass once through the inner Solar System before being flung to interstellar space. The appearance of a comet is called an apparition. Comets are distinguished from asteroids by the presence of an extended, visible atmosphere or coma, and sometimes also a tail." The note is titled "Comet" and is created by "nStudy".

Bottom Screenshot: The note titled "Tails of comets" is being created. The text of the note is: "outgassing. This produces a visible atmosphere or coma, and sometimes also a tail." The note is titled "Tails of comets" and is created by "nStudy". The note includes a "Project" field with the value "Project", a "Topic" field with the value "Outgassing", a "Describe what needs more research about this topic" field with the value "Parts", a "Types" field with the value "Are there different kinds, forms, categories?", a "Cause-effect relations" field with the value "Are there particular contexts/places/conditions?", and a "Context" field with the value "What are you curious about?".

Fig. 1 nStudy's Study View: a learner creates a note.

not the learner, reproduced this information. Data gathered using eye tracking gear could support such an interpretation. Other fields in the Explore note form identify slots in a schema suggesting how to explore the selected information. Light gray text in some text fields is replacement text. It is a scaffold to prompt the learner about topics or kinds of information that might be entered into each field. Replacement text disappears when the learner begins to type in a field. The note form as shown is empty except for the title and the reproduced selection. The learner is poised to assemble information with the selection by entering information into the note's fields.

At the right of the field labeled Parts is a button, . Clicking this button clones the field, adding an additional field to assemble more information in this Explore note. At the right of several other fields in this Explore note form is a button, . The learner knows clicking this button will display a message crafted by the instructional designer who authored this note form. For example, the message could be scaffolding about (a) topics or kinds of information to enter in a field and (b) an incentive, such as improving knowledge, when the learner assembles information using the note form's schema. If the learner clicks one of these buttons, the log of that event traces: (a) the learner searched knowledge about how to fill in a field, (b) monitored the product of search, judging it did not yield a satisfactory product, and (c) was motivated to seek scaffolding.

At some later time, if the learner viewed a previously filled-in note, clicking the selection which triggered making the note in the first instance will open (or switch to) the window from which that information was sourced and scroll to the selected text. This click event traces a monitoring event under conditions of the note being open plus, perhaps, other open windows displaying other information. Before clicking the hyperlinked text in this note, the learner searched knowledge in memory about outgassing. The product of that search was monitored relative to standards: Was retrieved information sufficient? Clicking the hyperlink to reinstate the source and the selection traces the learner monitoring current knowledge was not sufficient, as well as motivation to re-examine that source.

Process measurements

Records describing timestamped occurrences of learning events and patterns of events forming learning processes can be displayed in two basic formats. One is a timeline. Timelines begin with a starting event, such as a login event or a page open event, and display succeeding events on the timeline marked by the times of their occurrence. Timelines are easily interpreted to describe how events transition from one to another and the pace of learning activities. A second format shows flow of events as a graphical pattern. Graphs of learning processes are elaborated in the next section.

Assessments of learning events

Assessment has three elements: a measurement, a standard (or set of standards) to which a measurement is compared, and a judgment whether a difference(s) between the measurement and the standard(s) is (are) within allowable variation from a standard, called tolerance. Qualities of the measurement and methods that generate it, methods used to quantify differences between a standard and the measurement, and the choice of a standard and its reference framework—criterion, norm, or ipsative—each have implications for interpreting an assessment.

Other entries in this encyclopedia consider matters relating to assessments derived from observations, surveys and think aloud protocols. In those cases, although the topic assessed often will be something other than learning processes, there are few reasons or empirical grounds to expect assessments of learning processes should be understood differently because the topic of the assessment is learning processes.

Norms and specifics of reference systems can differ importantly depending on the topic assessed. Regarding normative assessments of learning processes, national norms are provided for the *Learning and Strategies Study Inventory* (LASSI; Weinstein et al., 2016), a paper-and-pencil survey for students at and above grade 6 (age 12–13). Items on the LASSI describe a variety of operations learners can apply when they study and prepare for tests. The LASSI norms are expressed as percentile rank scores for each subscale, e.g., information processing, selecting main ideas and self testing. Weinstein et al. recommend the 75th percentile as minimum score for assessing a student as proficient in a skill or competency described by a subscale. A score less than the 50th percentile is recommended as a criterion for identifying a student who may benefit from instruction or assistance in the processes described by a subscale. While there are many surveys about learning processes, norms like those for the LASSI are scarce.

Criterion-referenced assessments of learning events have been made in some research studies in which learners were trained in learning processes. For example, participants in a study investigating a reading comprehension strategy may be trained in processes comprising the strategy. After training, participants may be administered a paper-and-pencil test about the strategy or asked to demonstrate the strategy so it can be observed. To assess whether a participant in the research has developed (sufficient) skill in the strategy, the presence and sometimes qualities of component learning processes are compared to a criterion created for the experiment. Criterion-referenced assessment in this case takes a step toward providing evidence of participants' capabilities to apply the reading strategy when they read materials at a later stage in the research study. Because criteria are unique to each research study, comparing assessments across studies is difficult.

Norms and criteria have not been created for assessments based on think aloud protocols.

Norms for physiological data used in some multimodal assessments are well established owing to extensive data gathered in medical research. However, norms have not yet been related to learning processes. This can challenge the validity of interpretations of assessments unless multimodal data can be triangulated with data from normed instruments. For example, pupil dilation, elevated blood oxygen and a configuration of facial features normed to indicate surprise can converge on an assessment the learner is experiencing a surprise state. Other data, for example, trace data, would be needed to correlate that state of arousal to a cognitive operation such as monitoring or translating.

Assessments of learning processes

A learning tactic can be described as an IF-THEN sequence in which particular conditions set a stage for the learner to apply a particular operation or set of operations. For example, IF text is italicized, some learners THEN highlight it, a trace of monitoring where the standard relies on the author's interpretation of important or special information. When learning tactics are assembled into a regular pattern, a learning strategy is formed. For example, having highlighted an italicized term, IF the learner uses it to search knowledge and IF the product of search is monitored to be slight or uncertain, THEN the learner may break away to look for additional information to remedy that unsatisfying state. But, IF knowledge retrieved is monitored as ample and well-founded, THEN the learner may mark it for review and proceed with studying. The defining characteristic of a learning process that is strategic is the presence of a decision point at which one or another tactic is activated based on the status of a monitoring event.

Fig. 2 shows the flow across events in a fabricated learning process. Letters A, B ... represent distinct IF-THEN learning tactics. In this graph, one pattern of tactics observed has this sequence: A B C A B D For example, the learner selects text (A: IF text is italicized THEN select and scan for important disciplinary terms), executes a search for a term appearing in the selection (B), then annotates the selection (C) using information returned by the search of memory. The next event recorded is another selection of text (A). The learner again searches for terms in that selection (B). This time, results of the search are monitored and a different product results: information is not sufficient. So the learner tags the selection (D) to create a reminder to ask a classmate for information the search did not yield. In the lowest branch of Fig. 2, a different choice is made after the learner selects text (A). In this process, the learner monitors the selected information merits review and tags it such (E), then adds an entry in a calendar tool (F) to review items tagged for review on a particular date.

Graph representations of tactics and strategies can be assessed for several properties using a criterion referenced perspective. A tactic and a process can be assessed for its presence. However, note that a learner's failure to enact a theoretically expected tactic is ambiguous. The learner may not know conditions (IFs) that should be monitored or may have overlooked conditions. Regarding the operations (THENS), the operation(s) may not be available to the learner or the learner may not have assembled a theoretically expected operation with its triggering condition(s) to create a functioning IF-THEN production.

Assessing strategies is more complex because of potential variations in configurations of a strategy's components. Methods borrowed from graph theory (Winne et al., 1994) and process mining (e.g., see Saint et al., 2021) are available to analyze the degree to which patterns match a criterion, and to describe learning processes in terms such as their complexity and regularity. These assessments can be criterion-referenced in the sense of quantifying conformance to a canonical model for a process. Assessments also can be ipsative when a learner's current learning process is compared to earlier forms. At this time, there are no normative comparisons available to assess learner's tactics and strategies.

Psychometric issues

Reliability and validity merit attention in all assessments. Reliability is the degree to which a measurement obtained under particular conditions will replicate if that measurement is made again using the same operational definition applied under apparently identical conditions. The concept of reliability can be extended to generalizability when focus is the degree to which a measurement replicates with variation in the operational definition and/or conditions for obtaining the measurement. Validity is a claim about the accuracy and justifiability of an inference made or action taken with reference to a measurement. As is well known, reliability is a necessary but not sufficient condition for developing valid inferences.

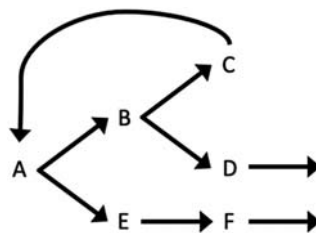


Fig. 2 Graph representation of a learning strategy.

When assessments of learning as a process are operationalized via responses to surveys and multimodal data, following well-established guidelines generates quite reliable measurements. Other entries in this encyclopedia examine those cases. Reliability of measures of learning processes grounded in think aloud data is addressed by examining the extent to which different coders or raters generate corresponding measurements. As documented elsewhere, reliability is advantaged when operational definitions are carefully constructed and coders are thoroughly trained to apply those operational definitions.

Reliability of records marking learning events and processes when learners use software is practically perfect. It is exceedingly rare that the kinds of interactions with software previously identified are not recorded when a learner engages with software or that events are recorded when they have not happened. Nonetheless, the question of reliability of trace data is complicated. Like all human beings, learners are not perfectly reliable. They may unpredictably overlook conditions (Ifs) in which case a learning operation (Then) is not recorded. As well, they may erratically misinterpret conditions (Ifs) leading them to perform an operation (Then) inconsistently with respect to conditions. And, self-regulating learners are continually striving to develop satisfying learning processes (Winne, 1995a,b), so they purposively vary If-Then events as they explore new approaches to learning. It is not yet clear how reliability should be modeled with regard to trace data.

A critical characteristic of distal and proximal self-report data is important when considering the validity of inferences about learning processes documented by these methods. Learners' memories are imperfect and their recall can be inestimably biased by the instructions provided about how to report learning events and processes. The validity of inferences and actions referenced to self-reports about learning events and processes should explicitly recognize these measures are what learners believe they do. Beliefs are not necessarily accurate accounts about what learners do.

Regarding multimodal measurements, approaches to examining validity of inferences parallel other kinds of proxy measurements, e.g., a teacher's judgment of a pupil's knowledge or interest. These issues are thoroughly addressed elsewhere in this encyclopedia.

The extent to which trace data support valid inferences about learning events and processes is directly proportional to judgments about the degree to which the operational definition of an event or process can be judged apt in representing a proposition in the parent theory.

Conclusions

Learning is a process unfolding over a series of events as a learner searches for information in long-term memory or in the external learning environment, and operates on results those searches provide. Operations create knowledge or revise what the learner knows, believes and prefers. To model learning as a process comprehensively requires identifying conditions that initiate operations on information, kinds of operations performed, and characteristics of the product(s) each operation produces. In transitioning to the next event in a learning process, learners decide what they will do in next learning event by monitoring characteristics of operations applied, such as effort applied and pace, and properties of products those operations create, such as clarity, comprehensiveness and certitude. This self-regulated learning is modeled by the COPEs set of facets. Each facet and relationships among them are the major topics of theories of learning.

Every instrument used to gather data for measuring learning processes inherently filters records of one or several of the COPEs facets. Learners, observers and software logs all selectively and purposefully choose information to observe and measure according to a theory or perceptual model of what is important and how observations should be quantified (scaled) and analyzed. In this sense, "raw data" is an oxymoron (Gitelman, 2013).

Trace data have gained traction for measuring learning as a process. These data operationally define a specific learning process by what a learner does and the information on which a learner operates in a learning event. Operational definitions of traces avoid filtering that can distort COPEs as represented by learners' responses to survey items, think aloud protocols, observers' records of learners' activity, and multimodal data. As well, trace data are advantaged by software's near perfect reliability in capturing the population of learning events as opposed to samples gathered when other instruments are used. Assessments of learning grounded in trace data are well suited to ipsative and criterion-referenced assessment. However, there is currently a near total lack of norms when learning processes are assessed using trace data.

References

- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. Freeman, New York, NY.
- Gitelman, L. (Ed.), 2013. "Raw Data" is an Oxymoron. The MIT Press, Cambridge, MA.
- Saint, H., Fan, Y., Singh, S., Gašević, D., Pardo, A., 2021. Using Process Mining to Analyse Self-Regulated Learning: A Systematic Analysis of Four Algorithms. LAK21, Irvine, CA, USA.
- van der Wel, P., van Steenbergen, H., 2018. Pupil dilation as an index of effort in cognitive control tasks: a review. *Psychon. Bull. Rev.* 25, 2005–2015. <https://doi.org/10.3758/s13423-018-1432-y>.
- Weinstein, C.E., Palmer, D.R., Acee, T.W., 2016. *LASSI User's Manual: For Those Administering the Learning and Study Strategies Inventory*, third ed. H & H Publishing Company, Dunedin, FL.
- Winne, P.H., 1995a. Inherent details in self-regulated learning. *Educ. Psychol.* 30, 173–187.
- Winne, P.H., 1995b. Self regulation is ubiquitous but its forms vary with knowledge. *Educ. Psychol.* 30, 223–228.

- Winne, P.H., 2018. Cognition and metacognition within self-regulated learning. In: Schunk, D., Greene, J. (Eds.), *Handbook of Self-Regulation of Learning and Performance*, second ed. Routledge, New York, NY, pp. 36–48.
- Winne, P.H., 2020. Construct and consequential validity for learning analytics based on trace data. *Comput. Hum. Behav.* 112. <https://doi.org/10.1016/j.chb.2020.106457>.
- Winne, P.H., Gupta, L., Nesbit, J.C., 1994. Exploring individual differences in studying strategies using graph theoretic statistics. *Alberta J. Educ. Res.* 40, 177–193.
- Winne, P.H., Teng, K., Chang, D., Lin, M.P.-C., Marzouk, Z., Nesbit, J.C., Patzak, A., Raković, M., Samadi, D., Vytasek, J., 2019. nStudy: software for learning analytics about processes for self-regulated learning. *J. Learn. Anal.* 6, 95–106.

Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania

Anjum Halai^a, Veronica Sarungi^b, and Therese N. Hopfenbeck^c, ^a Institute for Educational Development, Aga Khan University, Karachi, Pakistan; ^b Aga Khan Education Services, Dar-es-Salaam, Tanzania; and ^c University of Melbourne, Melbourne, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction to the background	63
The study	64
Participants	64
Teacher development	64
Methods	64
Literature	65
Assessment for learning	65
Teaching and learning	66
Data and results	66
Teachers' understanding AfL within normative context of practice	66
Seeking evidence of learning and feedback: tensions in whole class and group learning	67
Use of chalkboard	67
Group work	69
Nature of feedback	69
Conclusions and recommendations	70
Acknowledgment	71
References	71

Introduction to the background

Tanzania including mainland and Zanzibar islands is a country of more than 50 million (National Bureau of Statistics, 2017). Rapid urbanization is a demographic trend in Tanzania and the increase in the urban population is much higher in proportion to the region's rural population. Currently, the proportion of the country's urban population grows at a rate of approximately 5% per year as compared to the national average growth rate of 2.7% (UNDP, 2015). Dar es Salaam is a rapidly expanding city and, in spite of its higher HDI score, it has within it huge disparities with acute poverty in unplanned dwellings (Lugalla and Mbwambo, 1999; UNDP, 2015). In a study of the plight of young children and youth in cities in Tanzania, (United Nations Children's Fund (UNICEF), 2012, p. 64) maintains that Dar es Salaam has one of the highest proportions of children living in unplanned settlements in sub-Saharan Africa.

According to the 1995 Education and Training Policy that applied to the standard four classes in 2017, when the research was carried out, formal education in Tanzania constituted 2 years of pre-primary education, 7 years of primary education, 4 years of junior secondary (ordinary level), 2 years of senior secondary (advanced level). Since then, an education policy 2014–15 has been introduced (United Republic of Tanzania, 2014).

The country has made strides in providing access to primary education primarily due to strong policy commitment to education since its independence in 1960, where successive governments have seen education as necessary for development. However, the quality of learning processes and outcomes is low (Uwezo Tanzania, 2013, p. 43).

Recognizing that the teacher is a significant determinant of education quality preservice teacher education was a requirement for joining the profession. All primary school teachers were required to hold a pre-service teacher qualification called Grade A teachers' certificate. An outcomes based curriculum had been introduced in the country in 2005 and updated in 2013 (Ministry of Education and Culture, 2005; Taasisi ya Elimu Tanzania, 2013). According to it teachers were expected to follow prescribed syllabus plans in each subject including mathematics syllabus. Lesson plans were written in the structure provided by the Tanzania Institute of Education (Taasisi ya Elimu Tanzania, 2008).

The teachers' college curriculum was developed by the Tanzania Institute of Education, which was also responsible for curriculum design for pre-primary, primary and secondary levels. The pre-service teacher curriculum included participatory and interactive learning as one of its objective in order to ensure that learner-centered approaches were employed by teachers. Some of the strategies suggested at pre-service teacher education and also in primary school syllabus included discussions and group work by children. The pre-service curriculum in discussing teaching of young learners noted that assessment was important for teaching and two major categories were discussed namely, continuous assessment and final assessment. The latter was conducted at the end of primary education and the responsibility of the National Examination Council of Tanzania. Continuous assessment includes all assessment that a teacher would conduct while teaching in order to determine if learners had mastered a certain sub-topic before

moving on. Ways suggested to achieve this type of assessment included tests, projects, and questions. While elements of assessment for learning may be discerned from these curricular guidelines, the concept was not used explicitly, for details see [Taasisi ya Elimu Tanzania \(2013\)](#).

The study

Assessment for learning in Africa was a 3-year (2016–2019) project carried out in selected disadvantaged school districts in South Africa and Tanzania. It aimed to generate knowledge about how to develop and sustain teacher capacity in use of assessment for improving learning in challenging educational settings.

In Tanzania the study included quantitative data from baseline and end line tests of students' performance in a specially designed mathematics test administered to more than 500 students from a purposively selected informal urban settlement in Dar es Salaam. Alongside, a teacher development program was offered to all the mathematics teachers in six selected schools as described in the forthcoming sections. This paper reports from the teacher development component in Tanzania (see also [Halai et al., 2018](#)).

Participants

The selected informal urban settlement had a total of 159 schools of which 111 were public or government owned ([Prime Minister's Office Regional Administration and Local Government, 2015](#)). Of these six were selected meeting the following criteria: (a) in low socio economic settings in urban slums; (b) with large class size class, exceeding the recommended 40:1 pupil to teacher ratio ([Taasisi ya Elimu Tanzania, 2013](#), p. 10) finally, although not a specific selection criteria for schools those working with a Teachers' College in the vicinity and serving the settlement were preferred. An assumption being that involving the teacher college in the teacher development workshop would add to the sustainability of the AfL practice introduced.

The six selected schools had class size in the range of 85–215. They were in the vicinity of a Teachers' College whose tutors were invited to all the workshops.

A brief description is provided about the demographic profile of the Mathematics teachers from the six schools showed that: All participating teachers had received pre-service teacher education; none had attended professional development in the past 3 years that was related to teaching and learning strategies; the years of teaching experience ranged from 5 to 24; the age range of teachers was from early twenties to late forties.

Teacher development

The teacher development component comprised of eight workshops aimed to: explore teachers' perspectives about learning, teaching and AfL; introduce teachers to AfL principles and practice; and engage teachers in reflection on issues arising for AfL. Thirty eight mathematics teachers and teacher educators including the mathematics teachers from the six selected schools participated in the workshops.

Additionally, twelve (12) mathematics teachers from the six selected schools were followed up with lesson observation and post-observation meetings with their mentors, analysis and evaluation of the lesson to understand issues in implementing AfL in their classroom.

Methods

A range of methods were used to collect data. Classroom observation was recorded in a specifically developed observation-protocol geared to observe key dimensions of AfL. Fieldnotes were also used and teachers were asked to write their evaluation of lessons. Interviews were undertaken with the teachers, pupils and head teachers to capture their views about teaching, learning and assessment of learning. Teachers' interviews were meant to probe issues around AfL with specific reference to issues arising during the lesson observations. At the end of the project teachers were asked to write a narrative report addressing questions such as: (i) what is your understanding of assessment for learning? (ii) How did participation in the workshop(s) shift your thinking about assessment for learning? (iii) When you used assessment for learning what were the benefits? (please give concrete examples from your practice) (iv) When you used assessment for learning what were the challenges? (please give concrete examples from your practice). Researchers involved in the process were local, trained mathematics educators, grounded in the context of primary education in the country.

Table 1 summarizes collection methods and the data generated.

Framework of analysis was mainly drawn from the works of [Wiliam \(2006\)](#) and [Hopfenbeck \(2015\)](#) as discussed below. It had the following main features:

- Students should understand what they are going to learn and what is expected of them.
- Students should get feedback on the quality of their learning and achievement
- Students should get feedback on how to take their learning forward

Table 1 Summary of methods and data generated.

<i>Activity</i>	<i>Number</i>	<i>Data generated</i>
Lessons observed (grade 4)	48	Observation schedule, fieldnotes, lesson plans, artifacts
Post-observation visits	48	Mentors' notes
Workshops	08	Workshop plans and reports
Teachers' reflection	48	Teachers' lesson evaluation, teachers' written narratives on AfL
Teacher interviews	12	Interview transcripts

- Students should be engaged in their own learning by assessing their own work and their own learning.

Research team across Tanzania and UK developed a coding scheme that was drawn from the four key principles of AfL as noted above.

Literature

Assessment for learning

Assessment in education is typically seen with a focus on outcomes in high stakes testing. Assessment of learning from such an evaluator position typically occurs at the end of a teaching unit or at the end of an academic year and is summative in nature (Brookhart, 2001). On the other hand assessment for learning is formative in nature as it is essentially concerned with how assessment can take forward the process of learning (also see Cizek and Lim, 2022).

In their large scale survey covering 404 teachers across Canada and United States to look at teachers' approaches to classroom assessment in relation to existing accountability and standards-based education DeLuca et al. (2018) maintain that more experienced teachers use more varied approaches as compared to early career teachers and go on to suggest that professional learning opportunities should be given to teachers depending upon their career stage and previous assessment education.

It could be argued that assessment for learning is an integral part of the teaching-learning process and is subsumed in it. However, learning is a complex process and depends on a number of contributory variables that play out differently for different learners. A consequence is that the same teaching sequence in a lesson may result in different learning outcomes for different students. An important consideration in the process of learning and teaching is learners' previous knowledge and experience. (Wiliam, 2011) points out that it is the process of assessment for learning that could help take into account the individual learners' needs. It is therefore important to ensure that the strategies, techniques and methods for AfL are focused upon explicitly in the teaching learning process.

In their seminal work Black and Wiliam (1998, p. 2) looked into the black box of classroom to look at formative assessment in the course of teaching and learning in the classroom and maintained that assessment becomes formative in nature when, "evidence is actually used to adapt the teaching to meet the needs of the students." Wiliam proposes five key strategies that underpin good practice in assessment for learning:

Clarifying and understanding learning intentions and criteria for success

Engineering effective classroom discussions, questions and tasks that elicit evidence of learning

Providing feedback that moves learners forward

Activating students as instructional resources for each other, and

Activating students as owners of their own learning.

Wiliam (2006, p. 8).

Along similar lines but in the context of Norway, Hopfenbeck maintains that in the Norway Education Act the main purpose of assessment is for learning based on the following principles:

- (1) Students should be able to understand what they are going to learn and what is expected of them.
- (2) Students should get feedback that informs them about the quality of their work and their level of achievement.
- (3) Students should be advised on how to improve their learning outcome.
- (4) Students should be involved in their own learning process and in self-assessment.

Hopfenbeck (2015, p. 45).

A significant element of the above principles of assessment for learning is that the onus of learning is on the students and the teachers' role is to create a facilitative environment for students' learning. This is a significant shift that would impact the dynamics of teaching and learning in the classroom (Hopfenbeck, 2018).

Teaching and learning

Kajoro (2012, 2016) confirms that teaching practice in the Tanzanian classroom is largely informed by theoretical perspectives that position the teacher in the role of a transmitter and the student in that of a passive recipient. Along similar lines, Hardman et al. (2009) profiled patterns of interaction in public schools in sub-Saharan Africa and maintained that classroom discourse in sub-Saharan African classrooms was found to be highly ritualized creating a semblance of curriculum coverage, knowledge and understanding with the most prevalent methods of teaching being teacher explanation, punctuated by a question and answer approach, choral responses, pupils copying from the chalkboard, written exercises and teachers marking pupil work and so teacher-led recitation, with its constant demand for pupil participation, ensured teachers maintained tight control of the learning environment. Halai (2016, p. 42) notes that "Classrooms in sub-Saharan Africa are often over-crowded, under-resourced, with teacher centered patterns of interaction that provide little scope for learners to discuss ideas through discussion with each other or with the teacher."

Consequently, for an Afl environment as described in the preceding section the role of teachers and students in the classroom would have to be renegotiated with concomitant shifts in the classroom pedagogic practices to empower learners to take charge of their learning. Tabulawa (2013) cautions that change in teachers' practice must take into account teachers' underpinning beliefs about the purpose of education or else the change in practice will be superficial. In an in-depth critique of reforms in teachers' pedagogic practice, he holds that imported concepts and ideas of classroom reform such as student-centered classroom' do not change teachers (and students) philosophical outlook and therefore fail to improve students' learning.

To conclude, key principles of assessment for learning support a pedagogic practice that empower learners to take their learning forward and questions the fundamental one-way didactic interactions between the teacher and the learner.

Data and results

Several key findings emerge from the data and analysis. Teachers understood assessment for learning within the normative frameworks of an outcomes based curriculum, mandated lesson progressions and examination. Second, in large classes centralized approaches to seeking evidence of learning and provision of feedback appeared to be more effective as compared to group work, which was used largely as a management strategy. Third, nature of feedback remained general and focused on the right answer.

Teachers' understanding Afl within normative context of practice

Teachers' perspective of learning was complex and reflected the understanding of learning within the constraints and possibilities of a normative context.

Assessment for learning is about knowing how one has successfully understood what he/she was required to learn and what is supposed to be improved in areas that were not well covered.

Teacher WR, School-6, Sep 11, 2017.

Assessment for learning is a process of making the pupil the major actor and major participant in the learning activity. Assessment for learning also makes the teacher not to be the chief speaker; the pupil should be the major actor and participant in his/her own learning. Thus assessment for learning does not target classes with fewer children but also assists in classes with big numbers of pupils. Assessment for learning requires the teacher to provide feedback. The feedback can be oral or written. This enables the pupil to get good responses and wide understanding in his/her learning.

Teacher TM, School-2 September 11, 2017.

The quotes above appear to put the onus of understanding on the learner and recognize that the pupil is the major actor. This is a significant insight as it reflects an understanding of teaching where the teacher is not the chief speaker. However in the reality of the classroom teachers faced certain challenges within normative requirements of outcomes based curriculum, structured lesson plans, lesson progressions and students' achievement to be evaluated in yearly examination. Content of the lesson was drawn from the

syllabus and teachers were expected to follow the sequence of progression as mandated and prepare students for the national examination:

In fact, in this new curriculum it is a lot of changes, first I desire that anything I teach should relate with coming national examinations. Also to increase the performance of children, so that's why I'm still struggling here and there to find materials, because I do not know how the examinations will look like, because we first started a new curriculum, so I get challenged, and I wish to teach what relate with the examinations in future.

Teacher SK, School-5 April 20, 2018.

Using ample time to teach a given topic so that students understand [–]. This causes me to lag behind the school plan of finishing the syllabus. Using time from other teachers teaching-period for enabling students understand better thus causing other teachers to complain that I have used part of their teaching time.

Teacher NC, School-5, September 11, 2017.

40 minutes are not enough to reach 130 students and to provide them with feedback that one has planned where and when needed[–]. The demand from the quality assurers that requires teachers to be consistent with the scheme of work (topics) that the teachers has planned to implement in a given academic year.

Teacher WR, School-6, September 11, 2017.

In the first place, my school has big numbers of students. My class has 1,099 children divided into six (6) streams namely, A, B, C, D, E and F; every stream has more than 180 students. To say the truth, it becomes very difficult for me to reach every child in the time (40 minutes) allotted for the lesson. This is the most serious challenge I face in my school. Secondly, based on the big number of students, there are some students who cannot read or write. This is another challenge in using assessment for learning.

Teacher MNj, School-4, October 24, 2017.

The above interview excerpts illustrate that teachers' perceived completion of syllabus and preparation for examination as the main purpose of their teaching. They noted the structure of a 40 min teaching period and prescribed syllabus plan posed several challenges. Teachers feel the pressure of coverage of syllabus even if their professional judgment suggests slowing down. Additional issues arise due to lack of teaching materials such as textbooks. A consequence of the requirement of strict adherence to progression was that teachers were unable to plan lessons geared to their students learning needs. Besides time, lack of materials was also a challenge that was observed and also noted by the teachers.

Seeking evidence of learning and feedback: tensions in whole class and group learning

To seek evidence of students' learning and provide feedback teachers employed three main approaches that were use of the chalkboard, group work and questioning. In practice these strategies were in the course of lessons that largely followed a three-phased structure. In the first phase the teacher introduced the topic, shared the objectives of the lesson often making reference to the previous lesson. The second phase comprised of the main body of the lesson where new knowledge was introduced and reinforced. The teacher explained a mathematical procedure or the concept using the chalkboard and this was further reinforced through group-work or other strategies. The third phase comprised of individual tasks. The teacher would write additional questions or tasks to be done in exercise books and the purpose was largely to consolidate the understanding of new knowledge. Of course there were variations and the three stages overlapped.

The three-phased lesson delivery as described above covered more or less the five elements in the prescribed lesson plan structure provided as follows: (1) Introduction; (2) New knowledge; (3) Reinforcement of new knowledge; (4) Reflection; and (5) Consolidation (Taasisi ya Elimu Tanzania, 2008). However, analysis and reflection on learning was rarely if ever undertaken as an explicit part of the lesson. In some lessons it could be seen as part of the process of the questions asked by the teachers.

Use of chalkboard

A large chalkboard along the width of one wall was found in each classroom. Teachers used the chalkboard creatively for a variety of purposes. To check the homework assignment if all students had got the questions right:

Before moving to new lesson, on the following day, we make correction on the chalkboard, given that there are some who got all the questions right, some got few. We look at every question and ask those who got the question wrong to raise their hands, you check their number, then you ask those who got the same question right to explain [...] One explains the answer and we move to other questions. Therefore, corrections and exercises enable us to understand if they understand or not.

Teacher SK, School-5, November 9, 2017.

Teachers wrote on the chalkboard exercises taken from the textbook, as many students did not have the textbooks.

Shortage of text books. I have to write questions on the blackboard every day after teaching. For instance, if I teach for 30 minutes, only 10 minutes are left for me to write the homework on the blackboard so that the students also write the questions in their exercise books. Because of this, students do not complete the work; at times, they write fewer questions on the ground that they were cleaned off by the teacher who came for the following period as he/she also needs to use the blackboard and this happens many times.

Teacher RK, School-3, September 11, 2017.

In case they assigned individual or group tasks to the students, they were invited to present their work on the chalkboard. To accommodate the demands of the large class size, in some classes two or three students would be invited simultaneously to present their work on the chalkboard divided into columns to let each students work be represented separately. The whole class was then invited to vote whether or not their peer's answer was correct but a critique of the answers was rarely undertaken.

For example, lesson observation data when simultaneous groups presented during a lesson on fractions:

Group work: The class has 8 permanent groups for discussions; two groups have eleven members while each of the other six groups has ten members.

Each group is provided with a question

Group 1: $\frac{3}{11} + \frac{4}{11}$ Group 5: $\frac{2}{9} + \frac{3}{9}$

Group 2: $\frac{4}{15} + \frac{7}{15}$ Group 6: $\frac{3}{8} + \frac{4}{8}$

Group 3: $\frac{6}{14} + \frac{5}{14}$ Group 7: $\frac{9}{30} + \frac{14}{30}$

Group 4: $\frac{2}{20} + \frac{13}{20}$ Group 8: $\frac{10}{25} + \frac{10}{25}$

The teacher goes around each group to check the progress. All the groups present their works, one presenter from each group. The teacher divides the chalkboard in to 8 sections, named the eight groups and invited each group representative to present.

All presenters present their work at the same time. After every group has written its work on the chalkboard, the teacher leads the whole class to judge the works, errors are identified and pupils are invited to the chalkboard to correct errors. For example, looking at group six's work,

$\frac{3}{8} + \frac{4}{8} = \frac{7}{8}$, some pupils claim that the work is incomplete, then the teacher invites one pupil to show the complete work on chalkboard, the pupils writes this;

showing that $\frac{3+4}{8}$ was missing in the first place

Field Observation, Teacher EO, School-1, July 14, 2017.

Data from another lesson observation highlights the difference in oral and written responses:

Generally, the questions asked point the whole class and then those raised up their hands are picked (by calling their names) to give answers. After the answers have been given, again the teacher asks the whole class to say if the answer is correct or not. If the answer is not correct, then another pupil is chosen to answer. In most cases, (orally) the incorrect responses are not discussed, what is searched is the right answer. However, when pupils are given writing tasks, the teacher is able to check and discuss with the pupils about the wrong answers, asks for clarification before asking pupils to re-do basing on teacher's comments.

Field Observation, Teacher WR, July 11, 2017.

Group work

Use of group work was observed in most of the lessons. Teachers employed group work to manage their class, draw upon more knowledgeable pupils as a resource to teach others, and also as a management strategy:

As for me I have a big challenge right now as I see pupils do not understand especially the current topic [...] I have decided that we must resume the groups because the ones who are good in numeracy will help others in their respective groups, not to mix things there are those talented pupil, they get right answers for all questions, and as the teacher you real feel happy and you see that the pupil have real understood.

Teacher IM, School-6, March 21, 2018.

By dividing students with different levels of understanding into groups and assigning them with work. Before they make presentations, they assist each other to reach consensus.

Teacher MI, School-4, September 11, 2017.

[U]sing participatory methods, dividing the students into groups they can help each other and find out that the success criteria has been easily elicited.

Teacher NC, School-5 September 11, 2017.

Teachers also used group work as a management strategy:

I have 25 students in each group, and in each group there is leadership; the chairperson and secretary, this has helped to control truancy, because it is easier to identify the absent student in a group than in the whole group of 1074, I also use the groups to build self confidence for students, it is easier for a student to present in a group of 25, and get an experience to present for the whole class, and in their presentation they alternate between the students with high ability and those with low, this helps a lot.

Teacher MNj, School-4, March 2, 2018.

While group work might have been a good strategy to manage large classes, it did raise other issues for the environment of teaching and learning and promoting learning as one field observer noted:

Due to large number of pupils, noise and other distractions were generated, since the classroom was small and did not allow many movements for pupils to attend to the activities given. As a result, the teacher spent more time in trying to stop the distractions, but the pupils did not stop until the teacher went outside and came back with a stick, she threatened to beat them. Seeing the stick, most of the students stopped making noise, however, few continued, not until the teacher called them by name.

Field Observation, Teacher IM, School-6, July 11, 2017.

Nature of feedback

Teachers often used questions to help students move forward with mathematical procedures. For example a number of multiplication tasks involving carry-over with only one digit for the multiplicand as indicated in the data:

312	144
x 5	x4
-----	-----
1560	576

Students were invited in turn to the chalkboard to present their work and the teacher asked questions in order to make explicit the process of thinking when multiplying.

T: 2 times 5 equals 10, how much do we take in head? Students chorus: 1

The above process of question and answer went on until the multiplication was complete. However, if a student made a mistake such as providing wrong multiplication facts (e.g., $4 \times 4 = 12$) she corrected the mistake but gave a general kind of feedback "she made a mistake because she does not know the tables" (*amekosea kwa sababu hajui tebo*).

Field Observation, Teacher WR, School-6, February 24, 2017.

Another example of general feedback was from a lesson on areas of rectangles:

Generally, the feedback is given by praising (words such as "Very good," "You have tried") the correct responses but for the incorrect response, areas where the mistakes occurs and how to correct are not effectively spoken to a particular pupil, instead, the teacher, either asks the whole class to prove or disprove the answer before asking another pupil to attend the same question. For example, a teacher draws the rectangle on chalkboard, 40 cm and 20 cm for length and width respectively. The she asks- "What a kind of the figure is that?", the whole class answer—"it is a rectangle."

Then the teacher asks another question- "Why?" and she picks one pupil to answer, pupil response—"because all sides are equal." The teacher says "no" and then she appoints another pupil to answer the same question.

Field Observation, Teacher IM, School-6, July 5, 2017.

There were some cases where individual pupils' mistake was used to provide a general feedback to the class by asking peers to correct the process. For example:

The third example (6 hours 25 minutes + 1 hour 35 minutes) and selected a volunteer to solve the question. The pupil made a mistake in solving and got 7 hours 00 minutes—instead of 8 hours 00 minutes—the pupil converted but forgot to "carry forward." The teacher said "I agree with [pupil's name]—anyone with a different view." Some pupils put up their hand and from those selected one comes and corrects the process. Teacher then explains the misconception to the whole class.

Field notes, Teacher TM, School-2, September 25, 2017.

Besides the feedback being very general in nature an issue was that it often did not provide the students with a guidance on how to take their learning forward. Teachers' response to questions about the nature and purpose of feedback are as follows:

For example a person answers correctly, they can clap hands for him/her, if a person answers incorrectly, One you select another one who can answer correctly; if all have failed I can explain and assign them to work in groups, they discuss within five minutes, the answer will have been found.

Teacher MNy, School-1, March 21, 2018.

First of all, if the response is incorrect I would tell her—"thanks for trying" then I would ask another child to come and assist, and if the second child misses I would ask another one, I would continue this until I get someone who would give the correct answer, [...]even if the correct answer is given [...]I would conclude by analyzing the correct answer, some of the responses may be not exactly the correct answer but close to the answer, however some of the responses are out of the topic, at times children laugh at these kind of responses, this consumes lesson time and takes time to bring them back to normal, at this stage I am force to be strict so as to maintain discipline.

Teacher WR, School-6, October 12, 2017.

The illustrative examples above show that teachers probed for feedback to elicit evidence of students' learning. However, the extent to which the information received was used to provide feedback to the students' about their learning remained a question. For instance, in the case of multiplication of money, multiplication with carry over (e.g., "holding in your head"), not knowing the tables for multiplication facts, it was not apparent if students understood the place-value of digits. Teachers identified what was the mistake or the wrong answer but there was little evidence of probing why the students had provided the wrong answer.

Conclusions and recommendations

Three main conclusions could be drawn from the data and findings in the preceding section. First, teachers predominantly saw the purpose of their teaching and their role as teachers to "cover the syllabus" and prepare students for the examinations. Second, teachers' expressed nuanced understanding of assessment for learning, the function of feedback and its implications for changes in the role of students and teachers. However, there was gap in their espoused theory and practice in relation to AfL, potentially due to several structural barriers. These include, fixed lesson structures (time, lesson plan, sequence, progression), highly prescriptive curricula, extraordinarily large classes, teachers' inadequate academic and professional preparation. Third, teachers' practice showed creative approaches to manage the large classes, provide students an opportunity to participate in the classroom and to use the work on the chalkboard as a method to seek feedback on what students' understood.

Teachers' understanding of the purpose of their practice as noted above brought to surface the inherent tension in the principles underpinning the AfL framework as introduced to them and the inherent beliefs about their practice. For example, Students taking ownership, "Activating students as owners of their own learning" (William, 2006, p. 8), or Hopfenbeck, "Students should be

involved in their own learning process and in self-assessment" (2015, p. 45) suggests that students would play an active role in the classroom and shape the dynamics through their questioning and nature of participation.

However, the observed classroom dynamics were strongly led by the teacher. Students did not initiate questions and their responses to teacher's questions remained brief. Several studies have noted that interventions in the classroom such as "Afl" are ideologically laden and cannot succeed unless the underlying beliefs and purposes of education are taken into account. For example in his in-depth work set in sub-Saharan Africa, Richard Tabulawa, looked at, why pedagogic reforms fail in sub-Saharan Africa. His maintains that aid agencies largely from Euro-western contexts, made an effort to globalize "student-centered pedagogies" from a purely technical view of education. However, pedagogic reform is theory-laden and cannot succeed through a mere technical restructuring of the education process without taking into account the historical social, political and cultural context of education (Tabulawa, 2013). Along similar lines Barrett (2007) highlights the limitations of analytical frameworks developed in western contexts applied in under-resourced education systems and proposes, a more nuanced understanding of primary school teachers' classroom practice one that would take account of teachers working with an eclectic mix of techniques and ideas. Elsewhere, in Halai and Dylan (2011) several education researchers note with clarity the challenges and dilemmas when people from one culture introduce or research education reform in the context of a very different society.

It was also noted that the teachers espoused theoretically a nuanced and complex understanding of Afl. They recognized the importance of "to understand what they(students) understand" (Teacher SK, school 5, November 9, 2017). Concomitantly, their practice also reflected an attempt to seek feedback on what the students understood or not. However, as noted earlier, the dynamics were teacher directed. Questions were posed by the teacher and almost never initiated by the students. The large class size and other normative conditions might be a factor in unidirectional classroom interaction. For example, assessment for learning would assume a differentiated pace of learning in the classroom. However, an over loaded curricula and deep concern over students' performance in examination constrained any flexibility in the teaching process. The pace of the lessons was maintained for coverage and not necessarily learning. An alternate explanation could be that the socio-cultural mores in the wider society expected children to be deferential toward their elders and not raise questions.

Finally, the practice sheds light on specific aspects that could be addressed in targeted teacher development interventions. For example, the quality and nature of feedback provided by the teachers showed that feedback was largely for "positive reinforcement" and when provided on the content of learning it was general in nature. However, for feedback to be meaningful, it must provide information on the quality of learning and achievement. It reflects the need for specific focus on Afl in teacher preparation and continuing teacher development.

To conclude, teachers largely saw their role and purpose to prepare students for success in examination. In order to progress the curriculum delivery, they employed different strategies to seek evidence of students' learning within the constraints of large class size, limited resources and a normative policy framework. However, there was little evidence of students taking ownership and charge of their learning. The paper illustrates well the tensions in introducing external frameworks of teaching and learning in socio-cultural contexts where there is a potential conflict in the underlying principles of the framework with the beliefs and practices of the teachers. It raises questions for policy and practice in improvement of mathematics teaching and learning in Tanzania and other low-income countries.

Acknowledgment

We are grateful to the field researchers in Tanzania namely, Allan Fumbo, Evodius Jackson, Naomi Katunzi, Esther Kibga, Frida Kiliwa, Rehema Mdoe, Eristine Munisi, and Kipelo Obedi for their work in collecting and recording data from the field. Grant No: ES/N010515/1 funded by ESRC UK.

References

- Barrett, A.M., 2007. Beyond the polarization of pedagogy: models of classroom practice in Tanzanian primary schools. *Comp. Educ.* 43, 273–294.
- Black, P., William, D., 1998. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan* 92, 81–90.
- Brookhart, S.M., 2001. Successful students' formative and summative uses of assessment information, assessment in education: principles. *Pol. Pract.* 8 (2), 153–169. <https://doi.org/10.1080/09695940123775>.
- Cizek, G.J., Lim, S., 2022. In: Tierney, R., Rizvi, F., Ercikan, K., Smith, G. (Eds.), *International Encyclopaedia of Education*, vol. IV. Elsevier.
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., Luhanga, U., 2018. Teachers' approaches to classroom assessment: a large-scale survey. *Assess. Educ. Princ. Pol. Pract.* 25, 355–375. <https://doi.org/10.1080/0969594X.2016.1244514>.
- Halai, A., 2016. Teaching and learning mathematics: insights from classrooms in East Africa. In: Halai, A., Tennant, G. (Eds.), *Mathematics Education in East Africa: Towards Harmonization and Enhancement of Education Quality*. Springer International Publishers, Cham, pp. 35–46. <https://doi.org/10.1007/978-3-319-27258-0>. Book.
- Halai, A., William, D. (Eds.), 2011. *Research Methodologies in the South*. Oxford University Press Pakistan, Karachi. ISBN978-0-19-906356-7.
- Halai, A., Sarungi, V., Hopfenbeck, T., 2018. Assessment for learning in Africa: insights from classrooms in Tanzania. In: Halai, A., Mtenzi, F. (Eds.), *Proceedings of the AFRICME 5 Conference*, pp. 230–235. In: <https://www.mathunion.org/fileadmin/ICMI/AFRICME/AFRICME5/AFRICME%205%20Proceedings%20A4.pdf>.
- Hardman, F., Abd-Kadir, J., Agg, C., Migwi, J., Ndambuku, J., Smith, F., 2009. Changing pedagogical practice in Kenyan primary schools: the impact of school-based training. *Comp. Educ.* 45, 65–86.
- Hopfenbeck, T.N., 2018. Assessment and learning in an uncertain world. *Assess. Educ. Princ. Pol. Pract.* 25 (4), 351–354.
- Hopfenbeck, T.N., Flórez Petour, M.T., Tolo, A., 2015. Balancing tensions in educational policy reforms: large-scale implementation of assessment for learning in Norway. *Assess. Educ. Princ. Pol. Pract.* <https://doi.org/10.1080/0969594X.2014.996524>.
- Kajoro, P., 2012. Conventionalization of numerals in Tanzania. *Math. Teach.* 227, 15–21.

- Kajoro, P.M., 2016. Transition of the medium of instruction from English to Kiswahili in Tanzanian primary schools: challenges from the mathematics classroom. In: Halai, A., Clarkson, P. (Eds.), *Teaching and Learning Mathematics in Multilingual Classrooms*. SENSE Publishers, Rotterdam, pp. 73–85.
- Lugalla, J.L., Mbwambo, J.K., 1999. Street children and street life in urban Tanzania: the culture of surviving and its implications for children's health. *Int. J. Urban Reg. Res.* 23, 329–344.
- Ministry of Education and Culture, 2005. *Mathematics Syllabus for Primary Schools: Standard I–VII*. Tanzania Institute for Education, Dar-es-Salaam.
- National Bureau of Statistics, 2017. 2016 Tanzania in Figures. Viewed 14 June 2021. https://www.nbs.go.tz/nbs/takwimu/references/Tanzania_in_Figures_2016.pdf.
- Prime Minister's Office Regional Administration and Local Government. Enrolment of pupils in primary schools, 2015 [data file] Viewed 16 May 2016. <http://opendata.go.tz/dataset/primary-school-enrolment-by-sex-and-age/resource/96e8299e-a2ac-4396-922e-bc224f683c1a>.
- Taasisi ya Elimu Tanzania, 2008. *Mwongozo wa kufundisha somo la hisabati kwa shule za msingi Tanzania*. Dar-es-Salaam: TET. Note: This is the Teachers' Guide to Accompany the Mathematics Syllabus. Only available in Swahili (Unlike Syllabus Which is Available in Both English and Swahili Versions).
- Taasisi ya Elimu Tanzania (TET), 2013. *Mtaala wa elimu ya msingi Tanzania*. Viewed 14 June 2021. <http://196.41.32.93/MTAALA%20WA%20ELIMU%20YA%20MSINGI.pdf>.
- Tabulawa, R., 2013. *Teaching and Learning in Context: Why Pedagogical Reforms Fail in Sub-Saharan Africa*. Council for the Development of Social Science Research in Africa, Dakar.
- Tanzania, U., 2013. *Are Our Children Learning?: Annual Learning Assessment Report 2012*. Viewed 14 June 2021. <https://twaweza.org/wp-content/uploads/2021/02/UwezoTZ2013-FINAL-EN.pdf>.
- United Nations Children's Fund (UNICEF), 2012. *Cities and Children: The Challenge of Urbanization in Tanzania*. Viewed 14 June 2021. http://www.unicef.org/infobycountry/files/Cities_and_Children_-_FINAL.pdf.
- United Nations Development Programme (UNDP), 2015. *Tanzania Human Development Report 2014: Human Development for Economic Transformation*. Viewed 14 June 2021. <https://www.thdr.or.tz/docs/THDR2014-Main.pdf>.
- United Republic of Tanzania, 2014. *Sera ya elimu na mafunzo*. Viewed 14 June 2021. <http://www.moe.go.tz/sw/publications/policy-sera>.
- William, D., 2006. Keynote Address. Cambridge Assessment Network Conference. www.dylanwilliam.org/Dylan_Williams.../Cambridge%20AfL%20keynote. (Accessed 3 December 2016).
- William, D., 2011. What is assessment for learning? *Stud. Educ. Evaluation* 37 (1), 3–14.

Comparative judgment

Joshua A. McGrane, Assessment and Evaluation Research Centre (AERC), Melbourne Graduate School of Education, The University of Melbourne, Melbourne, VIC, Australia

© 2023 Elsevier Ltd. All rights reserved.

A brief history of comparative judgment	73
Educational applications of comparative judgment	75
Reliability and validity of comparative judgment	76
Summary	77
References	77

Comparative judgment, also referred to as paired or pairwise comparisons, is a choice methodology that has come to prominence in educational assessment over the past 30 years, although it has a longer history in the human sciences more broadly. It is commonly argued to be a superior method for assessing complex educational constructs (e.g., extended writing skill) that are better evaluated by holistic judgments of quality rather than by rating according to a pre-determined criterion or criteria (e.g., six traits of writing rubric) or by reducing the construct to a conventionally testable form (e.g., a multiple-choice test). Moreover, it has been used as a method to evaluate and equate the comparability of different assessment tasks for a given construct across time and contexts, as well as more broadly with respect to psychological constructs such as attitudes, values, and preferences.

In essence, comparative judgment works by having a group of judges make a series of judgments on pairs of artifacts (e.g., essays, portfolios, videos, exams, attitude statements) regarding which one in the pair represents more of the construct, whether that be, for example, a higher quality of performance, a more difficult task, or a more favorable opinion. These pairwise judgments, both within and across judges, may then be scaled using a psychometric model to represent the ranking and relative differences of the different artifacts in terms of the construct.

In the following sections, I will provide a brief history of the method including its scaling, a brief overview of existing applications in educational assessment, and a discussion of the evidence for its reliability and validity in these applications. Along the way, I will provide a selection of relevant references so the interested reader may delve deeper into the methodology and its applications.

A brief history of comparative judgment

Comparative judgment was pioneered by [Thurstone \(1927a\)](#) in the field of psychophysics where it was used to elicit people's pairwise judgments regarding a physical property of a set of artifacts, e.g., which one of a pair of cylindrical weights is heavier,¹ to evaluate how people's subjective experiences of the relationships between the magnitudes of the property relate to the actual magnitudes. From this line of research, they observed that, among other things, people are better at discriminating between the physical magnitudes of the artifacts as they become more divergent, e.g., one cylindrical weight is much heavier than the other, or to put it another way, people will tend to make more errors regarding which of the two cylinders is heavier when they are similar in weight. Further, when a judge is presented with the same pair of artifacts over time, they will not always make the same judgment of which one possesses more of the property and these "switches" in judgment will occur more frequently when each of the artifacts are similar in magnitude, e.g., there will be more switches in judgment for pairs of cylinders that are more similar in weight relative to other pairs.

From these indeterminacies in the judgments, we can see how this method may be used to not only infer the perceived ranking of the artifacts in terms of the relevant attribute, e.g., how many times each cylinder was judged to be heavier than the other cylinders, but also the relative differences of the perceived magnitudes of this attribute across the artifacts, e.g., the proportion of times the judgments of the relative weight of two cylinders are switched whereby the closer this proportion is to half, the closer the two cylinders' weights are perceived to be. [Thurstone's \(1927a\)](#) seminal contribution was to formalize these observations into a model that he loftily termed the "Law of Comparative Judgment".

For a more technical yet accessible introduction to the Law of Comparative Judgment, see [Bramley \(2007\)](#), but in general terms, it involves three key assumptions. Firstly, each time a judge encounters an artifact, this interaction sets off what Thurstone termed a "discriminal process"² regarding the construct of interest, which results in a psychological magnitude for that attribute, e.g., the perceived weight of the cylinder. Secondly, this discriminial process involves what Thurstone termed a psychological continuum,

¹Note that the artifacts are presented in an experimentally controlled manner such that they are effectively identical in terms of other relevant attributes such as their size and shape.

²Thurstone invented this term to reflect his intended neutrality with respect to the actual physiological and/or psychological processes involved in the judgment process.

which is a linear continuum of the psychological magnitudes of the attribute across all artifacts, e.g., a continuum of the perceived weights of all cylinders. Finally, this discriminative process is assumed to vary across judgments whereby the perceived magnitude of the attribute will likely change each time a judge encounters the artifact, e.g., the judge will likely perceive a cylinder to be a different weight every time the same artifact is encountered, which Thurstone referred to as the “discriminative dispersion”. Moreover, this variation is assumed to be normally distributed (e.g., bell-curve shaped) on the psychological continuum, meaning that smaller differences in perceived magnitude across encounters with an artifact will be more frequent than larger differences.

Based on these assumptions, according to the Law of Comparative Judgment, the outcome of a paired comparison between two artifacts depends on both the distance between the modes (i.e., the artifact’s most frequently perceived magnitude; the top of the bell-curve) of their discriminative dispersions and the degree of overlap in these dispersions on the psychological continuum. Thus, if two artifacts have a very small modal distance and high degree of overlap in their dispersions, then there will be many switches in the judgments of which one instantiates more of the construct. In contrast, if the modal distance is large and the dispersal overlap is small, there will be very few switches. Consequently, given other simplifying assumptions, including that the discriminative dispersions are all equal and all judgments are independent, the set of comparative judgments can be used to scale the relative perceived magnitudes in terms of the proportion of times each artifact is judged to have more of the construct than the other artifacts. However, it should be noted that [Thurstone \(1927a\)](#) considered this simplest case of his model (Case 5) to be only applicable for “coarse scaling” for most applications in psychology and education. Interestingly, despite this caution, almost all contemporary applications of comparative judgment in educational research apply a scaling model known as the Bradley-Terry-Luce model ([Bradley and Terry, 1952](#); [Luce, 1959](#)), which for all intents and purposes is identical to Case 5 of the Law of Comparative Judgment ([Andrich, 1978](#)).³

[Thurstone’s \(1927b, 1928, 1959\)](#) next seminal insight was that the method of paired comparisons and his scaling “law” could also be applied to constructs for which there is no obvious physical correlate, i.e., the kinds of constructs that psychological and educational assessment routinely deal with, including social attitudes and values, and constructs more educational in nature, such as quality of handwriting and English composition (i.e., writing). He and his collaborators were particularly interested in the former two and set about compiling and scaling sets of attitude statements, e.g., sets of statements that cover the full range of possible attitudes toward capital punishment, and value exemplars, e.g., statements of different crimes that cover the range of seriousness, to create attitudes and value scales, which [Thurstone \(1928\)](#) argued could be the “yardsticks” for the measurement of these constructs.

In short, taking social attitudes as an example, the approach initially involves a group of judges making paired comparisons on large sets of attitude statements with respect to a social issue to judge which of each pair represented more of the construct, e.g., a more favorable attitude toward capital punishment. These judgments are then scaled using the Law of Comparative Judgment and distilled down to a smaller set by discarding statements that were too divergent in their discriminative dispersions compared to other statements and by selecting a subset that were clearly ordered and represented approximately equal intervals across the full range of the scale.⁴ These distilled sets of statements could then be used to assess people’s attitudes toward the social issue by asking them to agree/disagree with each of the statements and then assigning them the median scale value of the statements they agreed with as a “measure” of their attitude.⁵

Although Thurstone and his colleagues quickly compiled dozens of these attitudinal “yardsticks”, they had little long-term impact on social research and this approach almost completely disappeared from the research literature within a decade. While the reasons for this are multi-faceted, they are largely attributable to [Thurstone \(1934\)](#) quickly turning his attention to other methods in the burgeoning field of intelligence research, and crucially, [Likert’s \(1932\)](#) publication of his method of summated ratings. This method enabled attitude researchers to forgo what has been described as the “laborious and time-consuming” initial step in Thurstone’s approach of using comparative judgments (or other methods) to create the attitude scale, and instead people were given sets of pre-selected positively and negatively phrased statements about an issue and were asked to indicate their approval of them on a 5-point scale⁶ ([Edwards and Kenney, 1946](#)). Their attitude toward the issue was then “measured” by simply reverse scoring their responses to the negatively phrased statements and then summing up their ratings across all statements. [Likert \(1932\)](#) justified this approach in terms of its parsimony while also being comparably reliable to, and highly correlated with, corresponding “measures” from Thurstone’s approach.

This charge of the laboriousness of the comparative judgment approach remains its greatest criticism even today and largely led to the method being placed on the shelf in mainstream psychological and educational research for around six decades. The greatest exception to this trend was in the methodological fields concerned with human preference and choice (e.g., [Guttman, 1946](#); [Bradley and Terry, 1952](#); [Luce, 1959](#); [Coombs, 1964](#); [Andrich, 1978, 1989](#); [Michell, 1994](#)). However, in recent decades, the method has

³The Bradley-Terry-Luce model only differs from Case 5 of the Law of Comparative Judgment through its use of the logistic function, which made the mathematical calculations more straightforward prior to digital computing, but the scale scores from the two models would be almost perfectly correlated.

⁴Thurstone also applied other methods for this purpose including the method of equal-appearing intervals, which involves judges rating the statements in terms of their favorability rather than pairwise comparing them ([Thurstone and Chave, 1929](#)).

⁵In addition to an attitude measure, this approach also provides information on the individuals’ latitude of acceptance, i.e., the range of statements they agree with, their latitude of rejection, i.e., the range of statements they disagree with, and their range of non-commitment, i.e., the range of statements to which they neither agree nor disagree ([Thurstone, 1959](#); [Sherif and Hovland, 1961](#)).

⁶Despite the widespread use of the term “Likert scale” to describe a wide variety of different rating scales in terms of both number of points and category descriptors, the original 5-point [Likert \(1932\)](#) scale included the category descriptors, 1—Strongly Disapprove, 2—Disapprove, 3—Undecided, 4—Approve, 5—Strongly Approve.

undergone a renaissance in educational assessment, as it has been argued to be a reliable and valid solution for various difficulties in the field and technological advancements have helped streamline its application.

Educational applications of comparative judgment

The starting place for the revival of comparative judgment in educational research was in the context of the England's upper secondary school examination system and specifically in terms of maintaining high-stakes examination standards over time, e.g., ensuring that a "C grade" on the A-level mathematics examination in 2022 represents the same standard of achievement as the equivalent grade in previous years, as such standard maintaining is known to be affected by subjectivity and unreliability in expert judgments (Benton, 2021; Benton and Elliot, 2016; Black and Bramley, 2008; Bramley, 2005, 2007; Bramley et al., 1998; Bramley and Gill, 2010; Leech et al., 2022; Pollitt and Murray, 1996; Seery et al., 2022). Moreover, England's examination system is unique by international standards in that these high-stakes examinations are created by several different organizations, known as examination boards, which further complicates the process of ensuring comparable examination standards, as they must also be maintained across these boards. Consequently, this system has pushed research into different robust methods for maintaining these standards.

In this context, the application of comparative judgment has typically involved experts making paired comparisons of exemplar examination scripts of different grade levels from different examination years. Judges were required to make holistic judgments of which script in the various pairs demonstrated the higher standard of achievement. These judgments were then scaled to evaluate whether the ranking of the scripts mapped on to the assigned grades across years, as well as to evaluate whether scripts assigned the same grade across years substantially diverged in their perceived quality. Generally, the findings of this research have provided support for the use of comparative judgment, either in its own rights or as a complementary approach to the current official methods, to enhance the reliability and validity of expert judgments in standard maintaining (Curcin et al., 2019).

The other main area of educational research that has pushed the resurgence of comparative judgment is writing assessment, or what is sometimes referred to as essay scoring (Steedle and Ferrara, 2016). Writing assessment has been a perennial issue for education systems, as writing is a fundamental skill for students to learn, but it is typically highly resource intensive to assess reliably and validly. As a result, systems often historically either excluded it from their assessments, attempted to reduce it to its constituent parts (e.g., knowledge of grammar and punctuation) that are more readily amenable to conventional testing, attempted to standardize it using generic assessment rubrics such as the six-traits or writing rubric, or simply just acquiesced to the limitations of teacher-based assessments of writing.

Comparative judgment has been argued to be a panacea for issues related to writing assessment for several reasons. Firstly, as the method entails judgments of relative performance, it does not require assessors to rate the performances into discrete categories and so alleviates the various rater and rubric effects (e.g., rater accuracy, rater severity, central tendency bias) that are known to bias rating and rubric-based writing assessment (Humphry and Heldsinger, 2014; Wolfe, 2004). Secondly, given that comparative judgment does not require scoring pieces of writing in terms of discrete categories, it allows for more discriminating and potentially more nuanced judgments of the whole performance, and may also act as a deterrent to formulaic approaches to writing that "tick off" the criteria (Humphry and Heldsinger, 2014, 2019a). Further, McGrane et al. (2018) demonstrated how higher-level criterion judgments may be integrated into a comparative judgment framework to provide both criterion-level and combined scales to further alleviate the need for rating written performances according to discrete categories.

To date, comparative judgment has proliferated across many writing assessment contexts, including large-scale primary and secondary-level assessment (e.g., McGrane et al., 2018; Pollitt, 2012; Wheadon et al., 2020), primary-level classroom-based assessment (e.g., Heldsinger and Humphry, 2010, 2013; Humphry and Heldsinger, 2019b), higher education contexts (e.g., Lesterhuis et al., 2022; van Daal et al., 2019), English as an additional language writing contexts (e.g., Şahin, 2021; Sims et al., 2020), as well as for the longitudinal equating of a standardized writing assessment (Humphry and McGrane, 2015).

Inventive technological developments have been coupled to this proliferation of research using comparative judgment for writing assessment to help alleviate the laborious nature of the paired comparisons method (Humphry and Heldsinger, 2019b; Goossens and De Maeyer, 2017; Pollitt, 2012; Wheadon et al., 2020). These developments have primarily involved developing systems that distribute the pairwise judgments across a pool of judges in such a way that every judge only compares a small subset of them while ensuring each written performance is involved in enough comparisons for them to obtain reliable scale scores. Pollitt (2012) and colleagues add an "adaptive" component to this distribution whereby the written performances are judged and scaled in rounds, and only closely scaled pairs of performances are compared in subsequent rounds to further increase the efficiency of the judgments. Wheadon et al. (2020) allow for the "anchoring" of the scale across location and time through the inclusion of pre-scaled performances, which can also be used to moderate the judgments to ensure teachers' judgments are not biased toward their own students. Humphry and Heldsinger (2019b) adopt a two-step method, like Thurstone's "yardstick" approach, whereby in the first step, comparative judgments are used to create a scale and select a set of exemplar performances that cover the full of range of performance, and in the second step, their system has judges assess new performances by comparing them to this set of pre-scaled exemplars rather than in a pairwise manner (see also Heldsinger and Humphry, 2010, 2013; McGrane et al., 2018).

The development of such systems has additionally enabled the proliferation of comparative judgment research in areas beyond examination standards and writing assessment. These areas include using paired comparisons for predicting and comparing the

item difficulty of examinations at different education levels (e.g., Attali et al., 2014; Green, 1985; Holmes et al., 2018), for assessing students' early literacy development (e.g., Humphry et al., 2017; Pollitt and Murray, 1996), for assessing mathematical problem solving in secondary students (e.g., Jones and Inglis, 2015; Jones et al., 2015), for creative arts assessment of secondary students including digitized portfolios (e.g., Tarricone and Newhouse, 2016a,b, 2017), for design assessment of secondary and higher education students including the integration of formative assessment (e.g., Bartholomew et al., 2022, 2019; Kimbell, 2012; Seery et al., 2019), as well as competency assessment for teacher performance (e.g., Wyatt-Smith et al., 2020). As applications of comparative judgment in educational assessment grow ever broader to encompass students of all ages across various disciplines and assessments of varying stakes, the reliability and validity of the method becomes an increasingly important issue.

Reliability and validity of comparative judgment

In the context of comparative judgment, reliability is typically conceived of as the ratio of the variance in scale scores that is attributable to true differences in the construct being judged to the total variance of the scores, which includes careless judgments and other sources of uncertainty (Bramley et al., 1998; Humphry and McGrane, 2015; Verhavert et al., 2018). The typical selling point of the paired comparison method is its high level of reliability, as much of the above cited research achieved reliability indices above 0.90, indicating near perfect reliability of the scale scores. These impressive levels of reliability are typically attributed to several factors in the literature, including the highly discriminating nature of the method where judges are forced to choose between each of the artifacts rather than matching them to discrete categories, which may increase the construct-relevant variance in the scale scores (Humphry and Heldsinger, 2019b). Moreover, the relative nature of the judgments has been argued to be cognitively simpler for assessors than making categorizations when making judgments about complex constructs, which in turn enhances the internal consistency of the judgments (Pollitt, 2012).

Benton and Gallagher (2018) put forward an alternative view that comparative judgment's high level of reliability is simply a by-product of its requirement for multiple judgments on each artifact, whereas other assessment approaches such as providing a rating typically only involve a single or small number of judgments for each artifact. This conjecture is consistent with other empirical findings that have demonstrated a high association between the reliability of the comparative judgment scale scores and the number of comparisons per artifact (Crompvoets et al., 2020; Goossens and De Maeyer, 2017; Verhavert et al., 2019, 2022). A recent meta-analysis by Verhavert et al. (2019) showed that between 10 and 14 comparisons per artifact were necessary to achieve a reliability level of 0.70 and between 26 and 37 to achieve reliability levels of 0.90. More conservatively, Crompvoets et al. (2022) recommend a minimum of 41 comparisons per artifact to achieve an adequate and unbiased estimate of scale score reliability. Verhavert et al. (2019) additionally argue that the required number of comparisons will be affected by the experience levels of the judges and the specific construct being assessed.

There has recently been some controversy in the literature regarding the of comparative judgment and particularly the adaptive variant briefly described above. Bramley and Vitello (2019) show that while the use of an adaptive approach increases the efficiency of the judgment process, it also spuriously inflates the reliability statistic that is typically used to evaluate the method; in their simulation this inflation is quite dramatic from moderate to extremely high levels of implied reliability. This finding is in some ways unsurprising, as the adaptive algorithm violates Case 5 of the Law of Comparative Judgment's simplifying assumption of the independence of judgments, and the same assumption is inherent in the Bradley-Terry-Luce model. This assumption is violated by the adaptive selection algorithm because the pairs of artifacts being judged in subsequent rounds from the first depend on the pairings in earlier rounds. In the broader psychometric context, this is referred to as a violation of local independence and these violations are known to spuriously inflate the variance of scale scores, which in turn biases the reliability statistics for the scale (Andrich and Kreiner, 2010). Consequently, researchers have begun to explore other ways to increase the efficiency of the judgment process through different approaches to adaptively selecting pairs in a manner that does not inflate the scale reliability (e.g., De Vriendt et al., 2022).

In comparison to the issue of reliability, the validity of comparative judgment has received much less attention, although there has recently been a proliferation of research focused on this issue (e.g., Chambers and Cunningham, 2022; Humphry and Heldsinger, 2019a; Lesterhuis et al., 2022; Mentzer et al., 2021; van Daal et al., 2017, 2019). Typically, the method has been found to have moderate to high levels of concurrent validity with other more established methods like rubric-based ratings (e.g., Attali et al., 2014; Heldsinger and Humphry, 2010, 2013; McGrane et al., 2018). Similarly, it has been found to have moderate to high levels of criterion validity with future exam performance (e.g., Jones and Inglis, 2015; Kimbell, 2012). The most recent comparative judgment validation studies have tended to evaluate whether judges attend to construct-relevant features when comparing artifacts as well as which specific features they attend to when making these comparisons.

For example, van Daal et al. (2019) collected judges' explanations for their comparative decisions of a set of essays. After careful coding of responses, they discerned three main categories of explanation: explanations related to the assessed construct and to the construct description, explanations related to the assessed construct but unrelated to the construct description, and explanations referring to the general quality of the essays. They found that around two-thirds of the explanations referred to both the construct and construct descriptions, and very few only referred to general quality. Similarly, Humphry and Heldsinger (2019a) asked judges to indicate the extent to which ten different aspects of writing were relevant to each paired comparison they made for a set of essays of varying quality. They showed there was a sensible relationship between the aspects attended to and the level of writing being compared, i.e., judges paid attention to more basic aspects of writing like punctuation and paragraphing when comparing lower

quality writing, and higher-level aspects like ideas and structuring when comparing higher quality writing. Moreover, they provide evidence that the holistic nature of paired comparisons allowed judges to make more nuanced assessments than the rubric-derived aspects would have alone. Both studies provide positive evidence for the validity of comparative judgment and, specifically, that such judgments are made based on construct-relevant features and may even be more sensitive to construct-relevant variance than either construct descriptors or assessment rubrics are able to represent. However, to date, most of this validation research has been conducted in the context of writing assessment and so caution must be shown when drawing broader conclusions about the method.

One aspect of the validity of comparative judgment that has received almost no attention in the literature is the appropriateness of the scaling model. As discussed earlier, [Thurstone \(1927a\)](#) only considered Case 5 of his model (and by extension the Bradley-Terry-Luce model) appropriate for “coarse scaling” due its simplifying assumptions being unrealistic in most applications, yet almost every article cited in this chapter uses a variant of this model. This is even more concerning when you consider that Thurstone argued that this case of his model is more appropriate in contexts where differences in the construct are reasonably homogeneous, like when comparing the weights of cylinders, yet in the educational context, comparative judgment is often lauded as a method for scaling complex and heterogeneous constructs. The earlier example of the way violations of the local independence assumption of the model confound interpretations of the reliability of the scale scores is cautionary in terms of simply taking the scaling model assumptions for granted. Future validation research should, therefore, examine the implications of these assumptions, including how much they “coarsen” the construct of interest, and whether other models that have been developed for paired comparisons may be more appropriate (e.g., [Böckenholt, 2001](#)).

Summary

Comparative judgment has re-emerged as a mainstream approach in educational assessment over the past 30 years, and particularly in contexts where complex constructs are assessed and where expert judgment and holistic evaluation are valued to capture the nuances of the construct, including examination standard maintenance, and the assessment of writing, design, and the creative arts. Technological advancements have furthered this re-emergence by streamlining the otherwise laborious method through the easy distribution of subsets of pairs across judges, rather than requiring each judge to make each comparison, and using selective algorithms to prioritize the most informative pairwise judgments. A review of existing research has shown comparative judgment to be a highly reliable assessment method for a wide range of constructs. Moreover, it has been shown to have good validity evidence in terms of concurrent and criterion validity, as well as being highly sensitive to construct-relevant variance, although more research is required in this space. Given the proliferation of comparative judgment research in educational assessment in just the past decade, the next 30 years should prove to be even more fruitful in terms of the scope of its educational applications and further enhancement of its reliability and validity across these applications.

References

- Andrich, D., 1978. Relationships between the Thurstone and Rasch approaches to item scaling. *Appl. Psychol. Meas.* 2 (3), 451–462.
- Andrich, D., 1989. A probabilistic IRT model for unfolding preference data. *Appl. Psychol. Meas.* 13 (2), 193–216.
- Andrich, D., Kreiner, S., 2010. Quantifying response dependence between two dichotomous items using the Rasch model. *Appl. Psychol. Meas.* 34 (3), 181–192.
- Attali, Y., Saldivia, L., Jackson, C., Schuppan, F., Wanamaker, W., 2014. Estimating Item Difficulty With Comparative Judgments. Research Report. ETS RR-14-39. ETS Research Report Series.
- Bartholomew, S.R., Strimel, G.J., Yoshikawa, E., 2019. Using adaptive comparative judgment for student formative feedback and learning during a middle school design project. *Int. J. Technol. Des. Educ.* 29 (2), 363–385.
- Bartholomew, S.R., Mentzer, N., Jones, M., Sherman, D., Baniya, S., 2022. Learning by evaluating (LbE) through adaptive comparative judgment. *Int. J. Technol. Des. Educ.* 32 (2), 1191–1205.
- Benton, T., 2021. Comparative judgement for linking two existing scales. *Front. Educ.* 6. <https://doi.org/10.3389/feduc.2021.775203>.
- Benton, T., Elliot, G., 2016. The reliability of setting grade boundaries using comparative judgement. *Res. Pap. Educ.* 31 (3), 352–376.
- Benton, T., Gallagher, T., 2018. Is Comparative Judgment Just a Quick Form of Multiple Marking?, vol. 26. Research Matters: A Cambridge Assessment Publication, pp. 22–28.
- Black, B., Bramley, T., 2008. Investigating a judgemental rank-ordering method for maintaining standards in UK examinations. *Res. Pap. Educ.* 23 (3), 357–373.
- Böckenholt, U., 2001. Hierarchical modeling of paired comparison data. *Psychol. Methods* 6 (1), 49–66.
- Bradley, R.A., Terry, M.E., 1952. Rank analysis of incomplete block designs, I. The method of paired comparisons. *Biometrika* 39, 324–345.
- Bramley, T., 2005. A rank-ordering method for equating tests by expert judgment. *J. Appl. Meas.* 6 (2), 202–223.
- Bramley, T., 2007. Paired comparison methods. In: Newton, P.E., Baird, J., Goldstein, H., Patrick, H., Tymms, P. (Eds.), *Techniques for Monitoring the Comparability of Examination Standards*. Qualifications and Curriculum Authority, London, pp. 246–294.
- Bramley, T., Gill, T., 2010. Evaluating the rank-ordering method for standard maintaining. *Res. Pap. Educ.* 25 (3), 293–317.
- Bramley, T., Vitello, S., 2019. The effect of adaptivity on the reliability coefficient in adaptive comparative judgement. *Assess. Educ. Princ. Pol. Pract.* 26 (1), 43–58.
- Bramley, T., Bell, J.F., Pollitt, A., 1998. Assessing changes in standards over time using Thurstone paired comparisons. *Educ. Res. Perspect.* 25 (2), 1–23.
- Chambers, L., Cunningham, E., 2022. Exploring the validity of comparative judgement: do judges attend to construct-irrelevant features? *Front. Educ.* 7, 802392. <https://doi.org/10.3389/feduc.2022.802392>.
- Coombs, C.H., 1964. *Theory of Data*. J. Wiley, New York.
- Crompvoets, E.A.V., Béguin, A.A., Sijtsma, K., 2020. Adaptive pairwise comparison for educational measurement. *J. Educ. Behav. Stat.* 45 (3), 316–338.
- Crompvoets, E.A.V., Béguin, A.A., Sijtsma, K., 2022. On the bias and stability of the results of comparative judgment. *Front. Educ.* 6, 788202. <https://doi.org/10.3389/feduc.2021.788202>.

- Curcin, M., Howard, E., Sully, K., Black, B., 2019. Improving Awarding: 2018/2019 Pilots. Research and Analysis (Ofqual). Office of Qualifications and Examinations Regulation, London.
- De Vrindt, M., Van den Noortgate, W., Debeer, D., 2022. Text mining to alleviate the cold-start problem of adaptive comparative judgments. *Front. Educ.* 7, 854378. <https://doi.org/10.3389/educ.2022.854378>.
- Edwards, A.L., Kenney, K.C., 1946. A comparison of the Thurstone and Likert techniques of attitude scale construction. *J. Appl. Psychol.* 30 (1), 72–83.
- Goossens, M., De Maeyer, S., 2017. How to obtain efficient high reliabilities in assessing texts: rubrics vs comparative judgement. In: *International Conference on Technology Enhanced Assessment*. Springer, pp. 13–25.
- Green, K., 1985. The paired comparison method in educational research. *Educ. Psychol. Meas.* 45 (3), 567–575.
- Guttman, L., 1946. An approach for quantifying paired comparisons and rank order. *Ann. Math. Stat.* 17 (2), 144–163.
- Heldsinger, S., Humphry, S., 2010. Using the method of pairwise comparison to obtain reliable teacher assessments. *Aust. Educ. Res.* 37 (2), 1–19.
- Heldsinger, S.A., Humphry, S.M., 2013. Using calibrated exemplars in the teacher-assessment of writing: an empirical study. *Educ. Res.* 55 (3), 219–235.
- Holmes, S.D., Meadows, M., Stockford, I., He, Q., 2018. Investigating the comparability of examination difficulty using comparative judgement and rasch modelling. *Int. J. Test.* 18 (4), 366–391.
- Humphry, S.M., Heldsinger, S.A., 2014. Common structural design features of rubrics may represent a threat to validity. *Educ. Res.* 43 (5), 253–263.
- Humphry, S., Heldsinger, S., 2019a. Raters' perceptions of assessment criteria relevance. *Assess. Writ.* 41, 1–13.
- Humphry, S.M., Heldsinger, S., 2019b. A two-stage method for classroom assessments of essay writing. *J. Educ. Meas.* 56 (3), 505–520.
- Humphry, S.M., McGrane, J.A., 2015. Equating a large-scale writing assessment using pairwise comparisons of performances. *Aust. Educ. Res.* 42 (4), 443–460.
- Humphry, S., Heldsinger, S., Dawkins, S., 2017. A two-stage assessment method for assessing oral language in early childhood. *Aust. J. Educ.* 61 (2), 124–140.
- Jones, I., Inglis, M., 2015. The problem of assessing problem solving: can comparative judgement help? *Educ. Stud. Math.* 89 (3), 337–355.
- Jones, I., Swan, M., Pollitt, A., 2015. Assessing mathematical problem solving using comparative judgement. *Int. J. Sci. Math. Educ.* 13 (1), 151–177.
- Kimbell, R., 2012. Evolving project e-scape for national assessment. *Int. J. Technol. Des. Educ.* 22 (2), 135–155.
- Leech, T., Gill, T., Hughes, S., Benton, T., 2022. The accuracy and validity of the simplified pairs method of comparative judgement in highly structured papers. *Front. Educ.* 7, 803040. <https://doi.org/10.3389/educ.2022.803040>.
- Lesterhuis, M., Bouwer, R., Van Daal, T., Donche, V., De Maeyer, S., 2022. Validity of comparative judgment scores: how assessors evaluate aspects of text quality when comparing argumentative texts. *Front. Educ.* 7, 823895. <https://doi.org/10.3389/educ.2022.823895>.
- Likert, R., 1932. A technique for the measurement of attitudes. *Arch. Psychol.* 22, 5–55.
- Luce, R.D., 1959. *Individual Choice Behaviours: A Theoretical Analysis*. J. Wiley, New York.
- McGrane, J.A., Humphry, S.M., Heldsinger, S., 2018. Applying a Thurstonian, two-stage method in the standardized assessment of writing. *Appl. Meas. Educ.* 31 (4), 297–311.
- Mentzer, N., Lee, W., Bartholomew, S.R., 2021. Examining the validity of adaptive comparative judgment for peer evaluation in a design thinking course. *Front. Educ.* 6, 772832. <https://doi.org/10.3389/educ.2021.772832>.
- Michell, J., 1994. Measuring dimensions of belief by unidimensional unfolding. *J. Math. Psychol.* 38 (2), 244–273.
- Pollitt, A., 2012. The method of adaptive comparative judgement. *Assess. Educ. Princ. Pol. Pract.* 19 (3), 281–300.
- Pollitt, A., Murray, N.L., 1996. What raters really pay attention to. In: Milanovic, M., Saville, N. (Eds.), *Studies in Language Testing 3: Performance Testing, Cognition and Assessment*. Cambridge University Press, Cambridge.
- Şahin, 2021. Şahin, A., 2021. Feasibility of using comparative judgement and student judges to assess writing performance of English language learners. *J. Pedagog. Res.* 5 (4), 140–154.
- Seery, N., Buckley, J., Delahunty, T., Canty, D., 2019. Integrating learners into the assessment process using adaptive comparative judgement with an ipsative approach to identifying competence based gains relative to student ability levels. *Int. J. Technol. Des. Educ.* 29 (4), 701–715.
- Seery, N., Kimbell, R., Buckley, J., 2022. Using teachers' judgments of quality to establish performance standards in technology education across schools, communities, and nations. *Front. Educ.* 7, 806894. <https://doi.org/10.3389/educ.2022.806894>.
- Sherif, M., Hovland, C., 1961. *Social Judgment*. Yale University Press, New Haven.
- Sims, M.E., Cox, T.L., Eckstein, G.T., Hartshorn, K.J., Wilcox, M.P., Hart, J.M., 2020. Rubric rating with MFRM versus randomly distributed comparative judgment: a comparison of two approaches to second-language writing assessment. *Educ. Meas.* 39 (4), 30–40.
- Steedle, J.T., Ferrara, S., 2016. Evaluating comparative judgment as an approach to essay scoring. *Appl. Meas. Educ.* 29 (3), 211–223.
- Tarricone, P., Newhouse, C.P., 2016a. Using comparative judgement and online technologies in the assessment and measurement of creative performance and capability. *Int. J. Educ. Technol. High. Educ.* 13 (1).
- Tarricone, P., Newhouse, C.P., 2016b. A study of the use of pairwise comparison in the context of social online moderation. *Aust. Educ. Res.* 43 (3), 273–288.
- Tarricone, P., Newhouse, C.P., 2017. An investigation of the reliability of using comparative judgment to score creative products. *Educ. Assess.* 22 (4), 220–230.
- Thurstone, L.L., 1927a. A law of comparative judgment. *Psychol. Rev.* 34 (4), 273–286.
- Thurstone, L.L., 1927b. The method of paired comparisons for social values. *J. Abnorm. Soc. Psychol.* 21, 384–400.
- Thurstone, L.L., 1928. Attitudes can be measured. *Am. J. Sociol.* 33, 529–554.
- Thurstone, L.L., 1934. The vectors of mind. *Psychol. Rev.* 41 (1), 1–32.
- Thurstone, L.L., 1959. *The Measurement of Values*. University of Chicago Press, Chicago.
- Thurstone, L.L., Chave, E.J., 1929. *The Measurement of Attitude*. University of Chicago Press, Chicago.
- van Daal, T., Lesterhuis, M., Coertjens, L., van de Kamp, M.-T., Donche, V., De Maeyer, S., 2017. The complexity of assessing student work using comparative judgment: the moderating role of decision accuracy. *Front. Educ.* 2, 44. <https://doi.org/10.3389/educ.2017.00044>.
- van Daal, T., Lesterhuis, M., Donche, V., De Maeyer, S., Coertjens, L., 2019. Validity of comparative judgement to assess academic writing: examining implications of its holistic character and building on a shared consensus. *Assess. Educ. Princ. Pol. Pract.* 26 (1), 59–74.
- Verhaver, S., De Maeyer, S., Donche, V., Coertjens, L., 2018. Scale separation reliability: what does it mean in the context of comparative judgment? *Appl. Psychol. Meas.* 42 (6), 428–445.
- Verhaver, S., Bouwer, R., Donche, V., De Maeyer, S., 2019. A meta-analysis on the reliability of comparative judgement. *Assess. Educ. Princ. Pol. Pract.* 26 (5), 541–562.
- Verhaver, S., Furlong, A., Bouwer, R., 2022. The accuracy and efficiency of a reference-based adaptive selection algorithm for comparative judgment. *Front. Educ.* 6, 785919. <https://doi.org/10.3389/educ.2021.785919>.
- Wheadon, C., Barmby, P., Christodoulou, D., Henderson, B., 2020. A comparative judgement approach to the large-scale assessment of primary writing in England. *Assess. Educ. Princ. Pol. Pract.* 27 (1), 46–64.
- Wolfe, E.W., 2004. Identifying rater effects using latent trait models. *Psychol. Sci.* 46, 35–51.
- Wyatt-Smith, C., Humphry, S., Adie, L., Colbert, P., 2020. The application of pairwise comparisons to form scaled exemplars as a basis for setting and exemplifying standards in teacher education. *Assess. Educ. Princ. Pol. Pract.* 27 (1), 65–86.

Peer- and self-assessment

Keith J. Topping, University of Dundee, Dundee, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	79
Definitions	79
Peer assessment	79
Self-assessment	79
Theory	80
Theory of peer assessment	80
Theory of self-assessment	80
Reliability	80
Online or face-to-face	80
Peer assessment	80
Literature reviews	80
Elementary school	81
High school	82
Higher education	82
Self-assessment	83
Literature reviews	83
Primary school	84
Secondary school	84
Higher education	85
Discussion and conclusion	85
Discussion	85
Conclusion	86
References	87

Introduction

Peer- and self-assessment have been widely used for some time, not only in elementary (primary) schools and high (secondary) schools and in higher education, but also in a wide range of work-place scenarios. Here we will consider peer assessment before self-assessment, but in both cases confine ourselves to formal educational environments. As the contexts are so different, elementary (primary) school, high (secondary) school and higher education will be considered separately. First, however, literature reviews or meta-analyses in each sector will be explored, although unfortunately these quite often mix studies from different contexts together without any further thought. After the literature reviews, a few specimen studies in each of primary, secondary and higher education will be discussed. The Discussion and Conclusion trace the threads of similarity between peer assessment and self-assessment.

Definitions

Peer assessment

A widely quoted definition of peer assessment refers to an arrangement for learners to consider and specify the level, value or quality of a product or performance of other equal-status learners (O'Donnell and Topping, 1998). However, other similar terms (synonyms) are in the literature (e.g., peer grading/marking—giving a score to a peer product/performance; peer feedback—peers giving elaborated feedback; peer evaluation—more usually in workplaces regarding skill and knowledge; or peer review—more usually in academia regarding assessment of written papers). Several previous studies compare two or three types of peer assessment, but the variety in types goes far beyond that. O'Donnell and Topping (1998) described a typology of relevant variables. Gielen et al. (2011) offered a more developed inventory. Further developments came from Topping (2018), outlining 43 variables.

Self-assessment

Self-assessment can be defined as the involvement of learners in making judgments of merit or worth about their achievements and the outcomes of their learning—identifying standards and/or criteria to apply to their work and making judgments about the extent to which they have met those criteria and standards (Boud, 2013). Brown and Harris (2013) defined self-assessment as an “evaluative act carried out by the student concerning his or her own work and academic abilities” (p. 368). Panadero et al. (2016) defined it as a “wide variety of mechanisms and techniques through which students assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products” (p. 804). Andrade (2019) offered: “Self-assessment is the act of monitoring one’s

processes and products in order to make adjustments that deepen learning and enhance performance” (p. 10). As with peer assessment, clear criteria are crucial, and these may be developed with students in discussion or simply inflicted upon them (the latter option yielding worse results).

Theory

Theory of peer assessment

Peer assessment is not just for managing assessment burdens for teachers, but more importantly a mechanism for learning, particularly when incorporating elaborated feedback. For the assessor, the intellectual demands of reflecting, making a balanced assessment, formulating and delivering feedback can all lead to learning gains (Yu, 2011). For the assessee, the intellectual demands of receiving and evaluating the feedback, deciding what aspects to implement and what not, and reflecting on other issues prompted by the feedback but not contained within it can all lead to learning gains (Li et al., 2012). A more comprehensive and integrated theoretical model was proposed by Topping (2018), encompassing: organization and engagement; cognitive conflict; scaffolding and error management; communication; affect; re-tuning and inter-subjectivity; practice and generalization; reinforcement; meta-cognition, self-regulation and self-efficacy; and levels of learning. This indicated many of the processes which may or may not occur during PA.

Theory of self-assessment

By contrast, self-assessment has been relatively under-theorized, although Boud (2013) considers it at some length. Perhaps the most useful theoretical analysis comes from Panadero et al. (2019), who discuss the role of developing evaluative judgments in self-assessment and relate this to two models of self-regulated learning (those of Zimmerman and Winne). Zimmerman’s model mentions motivation, attention focusing and self-efficacy, while Winne’s model mentions motivational issues but is more cognitive. Andrade (2019) comments that formative self-assessment can promote knowledge and skill development, but that this is not surprising, given that it involves many of the processes known to support learning, including practice, feedback, revision, and especially the intellectually demanding work of making complex, criterion-referenced judgments.

Reliability

It is often assumed that teacher “expert” assessment should be the criterion for validity, but several studies show that peer assessment is more reliable and effective than teacher assessment, although teacher assessment is not very reliable (e.g., Johnson, 2013). The relationship with teacher assessment is better described as “consistency” rather than “accuracy”. There are similar effects at primary school, secondary school and in higher education. Likewise, for self-assessment, Ross (2006) found that it produced consistent results across items, tasks, and short time periods and contributed to higher student achievement and improved behavior, while corresponding only in part to teacher assessments.

Online or face-to-face

Self-assessment can of course be an endeavor conducted in isolation, although arranging for students to buddy with others to reflect jointly is a useful practice. Peer assessment is by definition more complex, and particularly when applied to large classes in higher education, it can become difficult to manage. Having peer assessors and assessees meet online at a time of their choosing (whether synchronously or asynchronously) gets round these problems. Another advantage of online peer assessment is that it can easily be made anonymous, so students will not know who is assessing them, and this may encourage their responses to be somewhat more critical than they would otherwise be, at least in the early stages. Developing trust with someone you may never have met and who remains anonymous can be a challenge, so often social and trust-building activities are built into initial training, and/or anonymous peer assessment is used only for the first stages of the project (Castle and McGuire, 2010). There are also software systems for managing peer assessment on a large scale, which allocate assessors to assessees, collect the assessments, and average out assessments in small groups of students who mutually assess.

Peer assessment

Literature reviews

There have been many literature reviews on peer assessment which have all been positive, from the earliest reviews (e.g., Topping, 1998, on peer grades and feedback; Falchikov and Goldfinch, 2000, on peer grades) to the latest meta-analyses. A systematic literature review on peer assessment was reported by Van Gennip et al. (2009). Fifteen studies 1990–2008 dealt with effects on achievement. However, only one of these studies included students from a school (a secondary school); the remainder consisting of university students. Peer assessment had positive effects. The authors developed four underlying constructs: psychological safety, value diversity, interdependence and trust. Psychological safety was defined as a belief that it is safe to take interpersonal risks in a group of people or the extent of confidence in the rest of the group. Value diversity referred to differences in opinion about what a team’s task, goal or mission should be—it should be low for peer work to be effective. Interdependence has been long

studied, but needed to be perceived by the participants rather than assumed by teaching staff. Several studies noted that students felt uncomfortable criticizing one another's work, or found it difficult to rate their peers, at least initially. [Tillema et al. \(2011\)](#) considered what quality criteria were specifically relevant to peer assessment. One hundred and thirty-two studies of peer assessment were analyzed. Where emphasis was placed on authenticity and future learning needs across the lifespan, peer assessment showed positive effects in terms of generalizability, and particularly utility in contexts beyond the present institution.

[Li et al. \(2016\)](#) meta-analyzed studies on peer assessment in digital platforms (mainly in universities) from 1999 to 2015, finding an average correlation between peer and teacher ratings of 0.63—moderately strong. In further meta-analyses, [Li et al. \(2020\)](#) found an overall Effect Size of 0.29 in 58 studies (Effect Size [ES] is a quantitative measure of the strength of the relationship between two variables which is independent of any measures and free of any units—the larger the ES the stronger the relationship). There were significant moderator variables of Training & Online/Digital (Moderator Variables are third order variables that affect the size or nature of the relationship between an independent and dependent variable). [Double et al. \(2020\)](#) found an overall ES of 0.31 in 54 studies, but no significant moderator variables. There are issues here about replicability – were the studies and analyses the same in the two papers?

[Zheng et al. \(2020\)](#) meta-analyzed 37 studies of computerized peer assessment from 1998 to 2018. Technology-facilitated peer assessment had a significant and medium mean ES (0.58) on learning achievement. The use of extra supporting strategies had a similar ES (0.54). Moderator variables such as training for assessors, duration and grouping types were related to effect sizes. A review of fifteen studies of peer assessment in teacher education from 2002 to 2020 was offered by [Cabello and Topping \(2020\)](#). Various types of peer assessment were used but overall peer assessment was found to be effective. Theoretical perspectives were reviewed and practical implications for decision-makers highlighted.

Elementary school

[Lin et al. \(2006\)](#) reported the successful use of a web-based portfolio for peer assessment for visual art education in elementary school. The peer assessment component helped the students' learning and the web-based portfolios helped teachers assess student learning progress. [Harrison et al. \(2015\)](#) found that teacher assessment worked against the self-reliance development of students and used peer assessment with 523 students. During self- and peer-assessment, students developed skills as critical and creative thinkers, effective communicators and collaborative team workers, becoming more personally productive and effective. Their self-awareness and self-reflection increased significantly. All of these were essential components of self-control and self-direction. [Boon \(2016\)](#) studied children's uptake of feedback during peer assessment in primary school writing. Findings showed that pupils made better use of feedback if: it was task-involving and useful; there was sufficient time given for them to act on it and discuss it with their peers; and they were asked to reflect on how it had been used to improve the quality of the assessed work.

Two studies examined the efficacy of different methods of applying peer assessment. [Yang et al. \(2005\)](#) developed a web-based interactive writing environment in a 2-year longitudinal study. Peer assessment was used to provide constructive comments to foster students' ability to review and criticize other writers' essays, to enable students to review their own essay, and to encourage students to improve their writing skills. Students who participated submitted many essays, interacted with other students online, and reviewed other essays, improving their writing skills. By contrast, [Chin and Teou \(2009\)](#) used concept cartoons with two parallel experimental classes of younger (9–10-year-old) and older (10–11-year-old) students. The cartoons presented opposing viewpoints about scientific ideas to stimulate talk and argumentation among students in small peer assessment groups. The dialogic talk and interactive argumentation of the students provided diagnostic feedback about students' misconceptions about scientific principles to the teacher, and was helpful in moving students toward better understanding.

Students who score high on academic achievement are more accepted, less rejected and disliked by peers, viewed by teachers as less deviant, and engage in more positive interactions than those who score low (e.g., [Malecki and Elliott, 2002](#)), which may suggest that able students would make the best peer assessors. However, [Bryan's \(2005\)](#) research demonstrated that certain types of social skills interventions (particularly those focused on self-perceptions, self-attributions and locus of control) had consistent positive effects on academic achievement, as did peer assessment, through which it was possible to raise the self-esteem and social-connectedness of rejected children and also raise their academic achievement.

[Rockhill and Asher \(1992\)](#) examined the types of behavior that distinguished between low-accepted children and their better-accepted classmates. Third-through fifth-graders ($n = 881$) in five elementary schools completed a sociometric scale and a peer nomination measure. The same behaviors were important for boys and girls. Both aggressive and withdrawn children received lower peer ratings for prosocial behavior, and these were less popular as peer assessment partners. [Yurdabakan \(2011\)](#) studied 46 fourth grade (aged 9–10) participants (28 female, 18 male) in a social sciences course. Students scored their same and opposite sexes with respect to their contribution to group work and their learning levels. The consistency between female student and teacher scores was higher than male student and teacher scores (the teacher was male). Peer assessment has been also used successfully with special needs children (e.g., [Scruggs and Mastropieri, 1998](#)), with students as young as grade four (9 years old). Importantly, there were gains from functioning as either assessor or assessee. Studies in this category also have implications for matching students involved in peer assessment.

Turning to more recent research, [Colognesi et al. \(2020\)](#) sought to understand which form of peer assessment was most effective in helping pupils progress with their speaking skills: when the peers discuss their feedback in subgroups or when they provide it individually in writing. Their total sample was 42 pupils aged 10–11 years from the same school. The discussed oral feedback condition was the most effective, and included more suggestions for action. [Hwang and Chang \(2020\)](#) developed a bi-directional

peer-assessment-based concept mapping system, which engaged students in assessing peers' concept maps and evaluating peers' ratings and comments. They tested this with 101 fifth graders from four classes, one condition adopting the bi-directional peer-assessment approach, while the control group adopted the conventional peer assessment approach for concept mapping. The bi-directional system increased learning achievement and the completeness of concept mapping, and also enhanced their critical thinking and the feedback quality of the peer assessment. [Chang et al. \(2020\)](#) developed a peer assessment approach incorporating virtual reality (VR) into design activities with fifth-grade natural science students. All the students were asked to design a VR project. From two conditions, one with peer assessment and the other with teacher feedback, data was collected on learning achievements, learning motivation, self-efficacy, critical thinking, creativity, and cognitive load before and after the activities. The peer assessment group showed higher learning effectiveness, and higher self-efficacy and critical thinking tendencies.

High school

In Taiwan, [Sung et al. \(2005\)](#) used progressively focused self- and peer assessment procedures with students developing designs for new websites. Two classes of 76 14–15-year-olds of above average ability assessed random websites produced by their peers. Self-assessment preceded peer assessment. The quality of the students' performance improved only after the peer assessment activities. [Tseng and Tsai \(2007\)](#) found that 184 10th grade 16-year-old students significantly improved their projects for a computer course by engaging in successive rounds of peer assessment activities. The study also related the type of peer assessment feedback (reinforcing, didactic, corrective or suggestive) to subsequent performance in the student projects. Reinforcing peer feedback was useful in the development of better projects. Suggestive feedback was helpful at the beginning of peer assessment. However, didactic feedback, and to an extent corrective feedback, were negatively correlated with student achievement. Similarly, [Chang and Chou \(2011\)](#) examined the effects of reflection quality in peer assessment with 45 14-year-old students during a web-based portfolio process. The immediate influence of reflection quality on learning outcomes was small, but positive and statistically significant. Follow-up found reflection quality significantly related to scores on an achievement test, work, and attitude outcomes.

A Belgian team ([Gielen et al., 2010](#)) examined the effectiveness of certain characteristics of peer assessment feedback in a quasi-experimental, repeated measures study of 43 13-year-old students. Written assignments showed that receiving justified comments in feedback was related to improvements in writing performance, but this effect diminished for students with higher pre-test performance. The effect of accuracy of feedback was less than the effect of justification. [Sung et al. \(2010\)](#) explored peer assessment in individual music performance with 116 seventh graders. Then 110 eighth graders had student-constructed web pages subject to peer assessments. Reliability and validity increased with the number of raters in both studies. Low- and high-achieving students tended to over- and underestimate the quality of their work, respectively.

[Lu and Law \(2012\)](#) studied 181 high school students engaged in online peer assessment. Peers graded and gave feedback, which was analyzed. Lu and Law found that provision by student assessors of feedback that identified problems and gave suggestions was a significant predictor of the performance of the assessors themselves, and that positive affective feedback was related to the performance of assesseees. [Hsia et al. \(2016\)](#) used a web-based peer-assessment method for performing arts activities, applied to 163 junior high students (experimental group and control group). Controls learned with a web-based streaming video-supported environment. Peer assessment using a rubric significantly improved the students' performance, self-efficacy and motivation. Performance ratings were highly related to the students' self-efficacy. Students who learned with peer assessment were significantly more satisfied with the learning activity than controls.

More recent research has tended to focus on more developed kinds of information technology. Thus, [Chien et al. \(2020\)](#) used a spherical video-based virtual reality environment to situate English as Foreign Language students in authentic English-speaking contexts, and coupled this with peer assessment to have students provide feedback to peers and also reflect on their own performance. Comparing peer assessment with non-peer assessment conditions showed the former was superior in terms of English speaking, learning motivation and critical thinking skills, as well as reduced anxiety. Praise feedback was positively correlated with performance.

Higher education

Peer assessment in teacher education was investigated by [Sluijsmans et al. \(2002\)](#). Ninety-three pre-service teachers were randomly assigned to control and experimental groups. Experimental groups were trained in defining performance criteria, giving feedback and writing assessment reports, through peer assessment tasks embedded in a course. Analysis of assessment reports written by students showed the experimental group surpassed the control group in quality of assessment skill. Experimental students also scored significantly higher on end products of the course. [Strijbos et al. \(2010\)](#) worked with graduate students in the context of peer assessment of academic writing, investigating the impact of various types of feedback and the sender's competence level. Students were assigned to four experimental groups and a control group. Experimental groups received a scenario with either Concise General (CGF) or Elaborated Specific (ESF) feedback by a high or low competence peer. ESF by a high competence peer was perceived as more adequate, but led to more negative affect. Overall, CGF groups outperformed ESF groups. Sender's perceived competence clearly affected student perceptions and should be taken into account.

[Cho and MacArthur \(2010\)](#) worked with undergraduates who received feedback from a single expert, or a single peer, or multiple peers. Students receiving feedback from multiple peers received more feedback of all types (directive, non-directive, and praise). Students in the multiple peers group revised their work better than other groups, largely as a result of non-directive feedback.

This study clearly showed that students performed better with feedback from their peers rather than with feedback by an expert. New technology was used by [Chen \(2010\)](#), who developed a Mobile Assessment Participation System (MAPS) using Personal Digital Assistants (PDAs), which enabled more flexible assessment arrangements, more efficient use of time, and more opportunities for student reflection. Thirty-seven teacher-education students participated, using MAPS over two rounds of assessment for self- and peer assessment. Scores obtained from the assessment activities confirmed the benefits of MAPS. However, the students were still concerned over the objectivity of peer assessment and the difficulty of providing constructive feedback.

University freshmen studying English as a Foreign Language were involved by [Van Steendam et al. \(2010\)](#) in either observing or practicing higher-order peer feedback, in the hope that subsequent emulation would be more efficient for the quality of revision. If emulation took place individually, then observation and practice were equally effective for strategy acquisition. However, for dyadic emulation to be productive, it needed to be preceded by observation. Thus, the effectiveness of each instructional format was influenced by whether the students worked individually or in dyads. [Jones and Alcock \(2014\)](#) had 1st year mathematics undergraduates ($N = 194$) judge pairs of scripts against one another in the absence of assessment criteria, i.e., pairwise comparative judgment. The students performed well as peer assessors despite the absence of assessment criteria.

Developing critical thinking in a context which did not favor this among Hong Kong undergraduates was carried out by [Zhan \(2020\)](#). Elements of the program included: online peer assessment training, provision of guiding questions, peer feedback requirements, anonymity, student evaluation of the usefulness of peer feedback and summative use of online peer assessment. E-journal entries revealed a significant improvement in critical thinking in the experimental groups. This study highlights the need to adopt a culturally embedded approach to design online peer assessment for cultivating critical thinking in non-western contexts. [Hsia and Hwang \(2020\)](#) used a mobile technology-assisted peer assessment approach over 15 weeks in a university dance choreography class of 266 students, divided into experimental and control groups and involving two rounds of peer assessment. The results showed that mobile peer assessment could significantly enhance the innovative competence of dance routines and dance skills. In addition, the students felt mobile peer assessment could assist them in understanding accurate evaluation criteria, reflecting in a more objective way, and broadening aspects of appreciation.

[Lai et al. \(2020\)](#) worked with 100 4th-year nursing students in Taiwan divided into experimental and control groups, implementing a web-based video-annotation system supporting two rounds of online peer-assessment activities in a nursing communication training scenario. The control group added comments using YouTube, which also has video-annotation capabilities. Effects on communication skills and professional attitudes were considerable—the video-annotation tool notably enhanced nursing students' communication performance. Students using the tool provided more suggestions than those who did not use it. An online peer assessment learning platform involving viewing of videos of others' performance was integrated into a musical theater course in a university in Taiwan by [Chen et al. \(2020\)](#). There were three peer assessment modes, namely peer scoring, peer commenting, and peer scoring and commenting. Three classes of students each used one of the three different peer assessment modes. The students who learned with the peer scoring and commenting mode improved their performance; in addition, they provided more critical feedback and read the rubrics more, watching example videos, watching peers' work, and reading peers' feedback. Students who learned with the peer scoring mode evaluated more peers' works, but showed few behaviors regarding viewing others' comments or reading the rubrics.

Self-assessment

Literature reviews

An early systematic review was conducted by [Sebba et al. \(2008\)](#), who found that self-assessment improved pupil attainment across a range of subjects (9 out of 15 studies) and pupil self-esteem also increased (7 out of 9 studies). There was increased engagement with learning, especially goal setting, clarifying objectives, taking responsibility for learning, and/or increased confidence (17 out of 20 studies). Teachers needed to be able to foster an interdependent relationship between teacher and students. It seemed to be important to involve students in "co-designing" the criteria for evaluation. Seven studies identified the need for students to receive some training in self-assessment.

A subsequent meta-analysis found the impact of self-assessment on achievement to have an ES of 0.40–0.45 ([Brown and Harris, 2013](#)), but with great variability. The authors argued that it was a form of self-regulation, a principle that was to become increasingly popular in the future. A further meta-analysis gave strength to the influence of self-assessment interventions on self-regulated learning with ESs of 0.23, 0.65 and 0.43 on three different measures of self-regulated learning and also an ES of 0.73 for impact of self-assessment on self-efficacy, a crucial self-regulatory feature ([Panadero et al., 2017](#)).

[Andrade \(2019\)](#) reviewed 76 empirical studies 2013–2018. Of these, 44 focused on the consistency of student self-assessment with other judgments, 25 investigated the relationship between self-assessment and achievement, 15 explored students' perceptions and 12 focused on the association between self-assessment and self-regulated learning. Of the studies of achievement, 20 focused on formative self-assessment and all of these demonstrated a positive association with learning. Treatment groups outperformed their comparison or control groups in everything from writing to dart-throwing, map-making, speaking English, and exams in a wide variety of disciplines. However, none of these papers discriminated between primary, secondary and higher education, and participants were from Grade 1 through to pre-service teachers and nurses.

Primary school

Self-assessment has formed a large part of the Scottish Government's push toward formative assessment, particularly in the context of the "Assessment for Learning" initiative. [Kirtan et al. \(2007\)](#) reported on an evaluation in 12 Scottish primary schools. Data from pupils, teachers and parents showed pupils taking more responsibility for their learning, contributing to improved motivation, confidence and classroom achievement, especially for lower attainers. Teachers and pupils in all schools provided evidence of real change in teaching and learning practices, involving changes in concentration, attainment, quality of work and learning skills. Self-assessment was associated with enjoyment and increased motivation of pupils, whereas pupils' attitudes to national summative tests indicated fear and loathing. The initiative also moved some teachers from a teacher-centered pedagogy to one which was more student-centered.

In the USA, [Andrade et al. \(2008\)](#) scaffolded self-assessment by having 116 grade 3 and 4 students read a model written assignment, generate a list of assessment criteria and use a rubric to self-assess first drafts. Evaluation considered time spent writing, prior rubric use and previous achievement. A comparison group generated a list of criteria for an effective story/essay, and reviewed first drafts, but did not self-assess. The experimental group gained significantly more than the comparison group, and the more the rubric was used, the better were the outcomes. Using a model and rubric seemed to be effective in helping students produce better writing. A further study by [Andrade et al. \(2010\)](#) investigated 162 students using a model essay, generating criteria and self-assessing first drafts with a rubric in relation to writing assignments. The comparison condition again involved generating a list of criteria and reviewing first drafts. Findings included main effects of treatment, grade level, writing time, and previous achievement on total essay scores, as well as main effects on scores for every criterion on the scoring rubric. Using a rubric helped students produce more effective writing.

Turning to Ireland, [Keane and Griffin \(2018\)](#) sought to investigate whether 85 children's academic self-assessments became more accurate in line with increased age and higher prior literacy attainment. Children from Grades 2, 5 and 6 wrote an English essay then self-assessed their work using rubrics. Children's self-assessments became significantly more accurate in line with increased age. Strong correlations also emerged between higher prior literacy attainment and children's accuracy in self-assessment. Primary school children with low literacy attainment displayed some difficulty making accurate self-assessments of their academic work in literacy, but this improved with age.

How might self-assessment work in a country where highly teacher-directed education was the norm? [Yan et al. \(2020a\)](#) report on a study with 1425 students in Hong Kong. Self-efficacy and perceived controllability were statistically significant predictors on intention to self-assess, while self-efficacy and intention had a significant influence on self-assessment practice. Psychological safety also had a significant impact on both self-assessment intention and practice. The role of self-efficacy here is clearly important, and is closely connected with that of psychological safety. Taking on individual responsibility in a situation where that responsibility has formerly rested completely with the teacher obviously requires considerable self-confidence—which teachers have to foster if self-assessment is to be effective with all students. Also recently, [Falk et al. \(2020\)](#) found that children from high socioeconomic status (SES) families were more accurate in their self-assessment compared to children from low SES families. The children were then engaged in a mentoring program designed to enhance their self-assessment skills, which had a positive effect on the accuracy of self-assessment, and the mentoring program was most effective for low-SES children.

Secondary school

[Kostons et al. \(2012\)](#) modeled engaging in self-assessment and task selection with 80 students. A second experiment with 90 students showed that providing examples and practice further enhanced student performance at self-assessment. Self-assessment and task-selection skills could play an important role in self-regulated learning and training these skills could significantly increase student performance. The effects of two different self-assessment supports—rubrics and scripts—on self-regulation, learning and self-efficacy were investigated with 120 geography students by [Panadero et al. \(2012\)](#). Self-regulation was measured through questionnaires and thinking aloud protocols. The results showed that the use of rubrics and scripts increased learning over the control group, and that scripts enhanced self-regulation more than rubrics. [Fastre et al. \(2012\)](#) worked with 68 nursing and care students involved in self-assessment. One group had all possible assessment criteria while the other had the most relevant criteria highlighted. Students in the highlighted group outperformed the others, including in generating more points of improvement, but they experienced higher mental effort in self-assessing their performance.

In Taiwan, [Chang et al. \(2013\)](#) explored the reliability and validity of Web-based portfolio self-assessment with 72 students enrolled in a computer applications course. They created learning portfolios, viewed peers' work, and twice performed self-assessments on the web-based portfolio assessment system. The results showed that: (1) there was high consistency between students' two assessment results; (2) self-assessment and teacher-assessment results were highly consistent; (3) the results of self-assessment and end-of-course examination were highly consistent. This implies that web-based portfolio self-assessment accurately reflects learning achievements, and is a reliable and valid assessment method. In Indonesia, [Ratminingsih et al. \(2018\)](#) used self-assessment in teaching English as a Foreign Language, investigating the effect of self-assessment on students' independence and writing competence. The effects of self-assessment on writing texts in three separate genres was explored, while a control group received a traditional teacher assessment. The outcome measure was a writing competence test, and student opinions were also sampled. On both measures a positive effect was evident. Teachers were recommended to promote the use of self-assessment to improve student independence and writing competence.

Groenendijk et al. (2020) used a visual rubric for assessing production and reception activities in art education. Students appreciated the rubric's visual form. It was not found that students, or boys in particular, overestimated their own performance in art education. Yan et al. (2020b) were similarly innovative, using a self-assessment diary intervention with 74 12-year-old students to explore effects on academic achievement, self-regulation and motivation. After each homework assignment, the intervention condition used the diary while the control condition did not. Self-assessment diaries significantly enhanced students' academic achievement, self-efficacy, and intrinsic value. Students with lower past achievement benefited more than other students from the intervention.

Higher education

Mort and Hansen (2010) used video review to heighten 1st-year undergraduate students' ability to self-assess communication skills. Students participated in a video-recorded counseling-simulation exercise, completing self-assessment before and after viewing their video. The largest portion of the students overestimated their skills in comparison to faculty member's evaluations, but video review brought about an increase in the self-assessment scores for nearly two-thirds (62%) of the students. It was argued that there should be opportunities for students to develop self-assessment skills early, and this should be reinforced throughout the curriculum. The gender perspective in student self-assessment was investigated by Torres-Guijarro and Bengoechea (2017) in three different courses across several academic years in an engineering school in Spain. Previous research demonstrated sex differences in self-image and self-confidence, both closely related to self-esteem and same- and other-sex esteem. Thus, gender effects might be expected in self-assessment, which is closely linked to self-image and self-confidence. The results indeed showed that women judged themselves too harshly.

Self-assessment is said to be related both to academic achievement and self-regulated learning (SRL). With 63 students enrolled in a 1-year teacher education master's program, Yan (2020) investigated the characteristics of self-assessment practices at different SRL phases and their relationship with academic achievement. The participants conducted four self-assessment actions at the SRL Preparatory, Performance and Appraisal phases. Final scores on the assignment showed that self-assessment was a fundamental skill for SRL and occurred at each SRL phase with different patterns. Self-reflection at the Performance phase was found to influence feedback seeking at the Appraisal phase. Self-directed feedback seeking was the strongest and most positive predictor of academic achievement. A study with 1st-year undergraduates on a chemistry course in biotechnology self-assessing technical laboratory skills, supported by explicit performance standards, was conducted by Lau (2020). Two interventions, a self-assessment (SA) checklist and a learner-recorded video, examined how these impacted appraisal abilities. All the participants self-assessed titration competencies against a checklist, and about half were supported by video replay. Self-assessment with video showed a significant main effect, and these students were more targeted in their critique responses.

The effects of self-assessment and indirect teacher feedback on self-regulated learning (SRL) in the English as a Second Language context was investigated by Vasu et al. (2020). Three intact classes were randomly assigned to experimental and control groups, who completed the Self-Regulation Strategy Inventory Self-Report. Qualitative data were derived from the SRL microanalytic protocol. Both self-assessment and indirect teacher feedback helped learners become more self-regulated. However, self-assessment reduced maladaptive behaviors more than indirect teacher feedback. It was concluded that self-assessment is an effective instructional practice that promotes SRL in terms of goal setting, strategy planning, strategy use, attribution and adaptive behavior.

Discussion and conclusion

Discussion

Turning now to draw together the threads that have emerged from this brief survey, we must bear in mind that the papers selected in each of the three later categories (primary, secondary, higher) were merely a selection, and do not pretend to encompass all the factors that might be relevant. Nonetheless, we will find a number of commonalities. Looking first at the introduction section, for both peer assessment and self-assessment definitions were given and the underlying theory briefly addressed. The concept of reliability was addressed and it was noted that correlation with instructor scores was better referred to as "consistency". Some of the differences between face-to-face and digital peer and self-assessment were considered, with advantages and disadvantages found on both sides.

Considering the narrative and systematic reviews and meta-analyses, we see that overall peer assessment and self-assessment appear to be more or less equally effective. Both emphasize the positive effect of training. The peer assessment reviews emphasized psychological safety, value diversity, interdependence and generalizability, and the effect of duration of project. The self-assessment reviews emphasized the importance of co-designing criteria, self-efficacy, the association with self-regulation and the wide variety of subjects to which self-assessment was applied. In fact, however, all these factors could be applied equally to either peer assessment or self-assessment. If you wanted to do both, an issue here is which would you do first. Peer assessment, particularly with more than one assessor, is likely to give a much broader view of the value of your products, and may sensitize you to criteria which would benefit later self-assessment. On the other hand, self-assessment is less challenging, and may provide a less threatening introduction to some of the key concepts.

This commonality continues when we consider the primary (elementary) education sector, where we find self-efficacy reported on in both peer assessment and self-assessment. The important of confidence and pupils feeling enabled to apply themselves in

their own way cannot be over-estimated. There seems to be more emphasis on technology in peer assessment, with several mentions of web-based and sometimes virtual reality activity. In primary there is considerable mention of the effects of peer assessment on critical thinking, while in self-assessment there is considerable emphasis on the effect of rubrics. Again, with a larger sample of papers we would likely see more cross-over between peer and self-assessment.

Considering primary peer assessment only, we see mention of the importance of pupils becoming creative thinkers, effective communicators and collaborative team workers. Feedback is more effective if there is time to reflect and revise. Social skills interventions can raise self-esteem and improve academic achievement. Peer assessment has worked with concept cartoons and concept maps. Oral feedback seems effective than written feedback. Peer assessment can be used effectively with children with special educational needs. Females tend to show more consistency with teacher assessments than do males.

From a primary self-assessment viewpoint, a model essay coupled with criteria was effective, and self-assessment capability improved with increased age and with better literacy performance. A mentoring (training) program was very effective for low-socio-economic-status children. In a highly teacher-directed context, psychological safety was very important.

Turning now to secondary education, we find that for both peer assessment and self-assessment rubrics are very important, together again with self-efficacy. Application to e-portfolios was mentioned in both cases. Writing competence was also said to have improved under both interventions.

Under peer assessment, the effectiveness of reinforcing and suggestive feedback was mentioned more than once, and critical thinking also appeared again. Peer assessment in music and art were highlighted, emphasizing relevance to subjects other than core subjects. Virtual reality was used. Comments needed to be justified and reliability went up with the number of graders. Peer assessment led to lower anxiety.

Under self-assessment, self-regulation appeared again. Modeling, examples and practice featured, as did a few selected criteria rather than many. High correlation of self-assessment to subsequent formal examinations was mentioned. Novel means of self-assessment via visual rubrics and self-assessment diaries were reported.

Finally, we turn to higher education. For both peer assessment and self-assessment, video was frequently used to add value to the process. Both interventions also emphasized the use of explicit assessment criteria. Teaching English as a Foreign Language featured for both interventions, and work with pre-service teachers appeared twice under peer assessment.

For peer assessment, again practice and critical thinking were mentioned. Again, use in artistic subjects featured—one in dance choreography and one in musical theater. There was mention of training in co-defining criteria and giving feedback, the role of observation and practice, that general was better than specific feedback, and that feedback from multiple peers was more reliable than from a single peer. In terms of technology, PDAs and annotation of videos featured. Comparisons between peers were found effective in the absence of criteria, while grading with commenting was more effective than grading alone.

Self-assessment operated in counseling skills and engineering (very different subjects), and in the latter subject women were found to have lower self-efficacy (although that might be subject-specific). Self-regulated learning was mentioned frequently again. Self-assessment was as effective as teacher feedback but reduced maladaptive behavior more.

Many teachers in all sectors of education use both peer assessment and self-assessment. Returning to the interesting question posed earlier in this Discussion: in what order should they be used? My own opinion is that face-to-face peer assessment should be used before digital peer assessment wherever that is possible, that peer assessment should be used before self-assessment, and that self-assessment alone should be used before self-assessment with video. Of course, this is just an opinion. A research investigation which compared the effects of different orders of involvement in these interventions would be highly desirable (albeit large and expensive). Naturally, this would then need doing in each sector of education separately, and also in different subject areas – a massive undertaking.

Conclusion

Overall, then, peer assessment and self-assessment both appear equally effective. Both are used in a wide variety of subjects, including music, art, education, English as a Foreign Language, counseling and engineering. Many initiatives deployed technology, including PDAs and virtual reality. Two factors were emphasized most: (1) the need to co-design explicit criteria with students, and (2) the important role of self-efficacy in both interventions. Both were thought to lead to self-regulation of learning, although this was more prominent with self-assessment. Beyond this, there was much emphasis on the importance of training, which could include modeling or observation and should include practice. Rubrics were frequently mentioned (closely connected to co-designed criteria). Both methods could lead to improved critical thinking. The number of assessors was important, several being better than one (particularly in peer assessment). Video was often added to self-assessment in various ways to improve accuracy. Finally, psychological safety was important. Trust had to be developed between students and teachers and between students and students, and this is obviously linked to the issue of self-efficacy.

Although the literature is now quite extensive, neither peer assessment nor self-assessment are as widely used as one might expect, given that they have the advantage that they appear to transfer the burden of assessment at least partly to the student and thus lighten the load of the teacher. Some schools have a whole-school policy on peer learning, which may well include peer assessment and self-assessment, but these are in the minority. In the university sector, while peer assessment and self-assessment might feature in many universities in a few departments or subjects, but whole-university approaches are much less common. Now that we are more aware of the conditions necessary to make these interventions work, we are in a stronger position to increase their take-up.

References

- Andrade, H.L., 2019. A critical review of research on student self-assessment. *Front. Educ.* 4 (87), 1–13. <https://www.frontiersin.org/article/10.3389/feduc.2019.00087>.
- Andrade, H.L., Du, Y., Mycek, K., 2010. Rubric-referenced self-assessment and middle school students' writing. *Assess. Educ. Princ. Pol. Pract.* 17 (2), 199–214.
- Andrade, H.L., Du, Y., Wang, X.L., 2008. Putting rubrics to the test: the effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educ. Meas.* 27 (2), 3–13.
- Boon, S.I., 2016. Increasing the uptake of peer feedback in primary school writing: findings from an action research enquiry, 3–13 *Education* 44 (2), 212–225.
- Boud, D., 2013. *Enhancing Learning Through Self-Assessment*. RoutledgeFalmer, London & New York.
- Brown, G.T.L., Harris, L.R., 2013. Student self-assessment. In: McMillan, J.H. (Ed.), *The SAGE Handbook of Research on Classroom Assessment*. Sage, Thousand Oaks, CA, pp. 367–393.
- Bryan, T., 2005. Science-based advances in the social domain of learning disabilities. *Learn. Disabil. Q.* 28, 119–121.
- Cabello, V.M., Topping, K.J., 2020. Peer assessment of teacher performance: what works in teacher education? *Int. J. Cognit. Res. Sci. Eng. Educ.* 8 (2), 121–132.
- Castle, S.R., McGuire, C.J., 2010. An analysis of student self-assessment of online, blended, and face-to-face learning environments: implications for sustainable education delivery. *Int. Educ. Stud.* 3 (3), 36–40.
- Chang, C.C., Chou, P.N., 2011. Effects of reflection category and reflection quality on learning outcomes during web-based portfolio assessment process: a case study of high school students in computer application course. *Turkish Online J. Educ. Technol.* 10 (3), 101–114.
- Chang, S.C., Hsu, T.C., Jong, M.S.Y., 2020. Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Comput. Educ.* 146, 103758. <https://doi.org/10.1016/j.compedu.2019.103758>.
- Chang, C.C., Liang, C.Y., Chen, Y.H., 2013. Is learner self-assessment reliable and valid in a web-based portfolio environment for high school students? *Comput. Educ.* 60 (1), 325–334.
- Chen, C.H., 2010. The implementation and evaluation of a mobile self- and peer-assessment system. *Comput. Educ.* 55 (1), 229–236. <https://doi.org/10.1016/j.compedu.2010.01.008>.
- Chen, I.C., Hwang, G.J., Lai, C.L., Wang, W.C., 2020. From design to reflection: effects of peer-scoring and comments on students' behavioral patterns and learning outcomes in musical theater performance. *Comput. Educ.* 150, 103856. <https://doi.org/10.1016/j.compedu.2020.103856>.
- Chien, S.Y., Hwang, G.J., Jong, M.S.Y., 2020. Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-speaking performance and learning perceptions. *Comput. Educ.* 146, 103751. <https://doi.org/10.1016/j.compedu.2019.103751>.
- Chin, C., Teou, L.Y., 2009. Using concept cartoons in formative assessment: scaffolding students' argumentation. *Int. J. Sci. Educ.* 31 (10), 1307–1332. <https://doi.org/10.1080/09500690801953179>.
- Cho, K., MacArthur, C., 2010. Student revision with peer and expert reviewing. *Learn. Instr.* 20 (4), 328–338.
- Colognesi, S., Vassart, C., Blondeau, B., Coertjens, L., 2020. Formative peer assessment to enhance primary school pupils' oral skills: comparison of written feedback without discussion or oral feedback during a discussion. *Stud. Educ. Eval.* 67, 100917. <https://doi.org/10.1016/j.stueduc.2020.100917>.
- Double, K.S., McGrane, J.A., Hopfenbeck, T.N., 2020. The impact of peer assessment on academic performance: a meta-analysis of control group studies. *Educ. Psychol. Rev.* 32, 481–509. <https://doi.org/10.1007/s10648-019-09510-3>.
- Falchikov, N., Goldfinch, J., 2000. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Rev. Educ. Res.* 70 (3), 287–322.
- Falk, A., Kosse, F., Schildberg-Hörisch, H., Zimmermann, F., 2020. Self-Assessment: The Role of the Social Environment. CESifo Working Paper No. 8308. Retrieved from SSRN. <https://ssrn.com/abstract=3618831>.
- Fastre, G.M.J., van der Klink, M.R., Sluijsmans, D., van Merriënboer, J.J.G., 2012. Drawing students' attention to relevant assessment criteria: effects on self-assessment skills and performance. *J. Vocat. Educ. Train.* 64 (2), 185–198.
- Gielen, S., Dochy, F., Onghena, P., 2011. An inventory of peer assessment diversity. *Assess. Eval. High Educ.* 36 (2), 137–155. <https://doi.org/10.1080/02602930903221444>.
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., Struyven, K., 2010. Improving the effectiveness of peer feedback for learning. *Learn. Instr.* 20 (4), 304–315.
- Groenendijk, T., Kárpáti, A., Haanstra, F., 2020. Self-assessment in art education through a visual rubric. *Int. J. Art Des. Educ.* 39 (1), 153–175. <https://doi.org/10.1111/jade.12233>.
- Harrison, K., O'Hara, J., McNamara, G., 2015. Re-thinking assessment: self- and peer-assessment as drivers of self-direction in learning. *Eurasian J. Educ. Res.* 60, 75–88.
- Hsia, L.H., Hwang, G.J., 2020. Enhancing students' choreography and reflection in university dance courses: a mobile technology-assisted peer assessment approach. *Br. Educ. Res. J.* <https://doi.org/10.1111/bjjet.12986>.
- Hsia, L.H., Huang, I., Hwang, G.J., 2016. A web-based peer-assessment approach to improving junior high school students' performance, self-efficacy and motivation in performing arts courses. *Br. J. Educ. Technol.* 47 (4), 185–632.
- Hwang, G.J., Chang, S.C., 2020. Facilitating knowledge construction in mobile learning contexts: a bi-directional peer-assessment approach. *Br. J. Educ. Technol.* <https://doi.org/10.1111/bjjet.13001>. Online first.
- Johnson, S., 2013. On the reliability of high-stakes teacher assessment. *Res. Pap. Educ.* 28 (1), 91–105.
- Jones, I., Alcock, L., 2014. Peer assessment without assessment criteria. *Stud. High Educ.* 39 (10), 1774–1787. <https://doi.org/10.1080/03075079.2013.821974>.
- Keane, L., Griffin, C.P., 2018. Assessing self-assessment: can age and prior literacy attainment predict the accuracy of children's self-assessments in literacy? *Ir. Educ. Stud.* 37 (1), 127–147.
- Kirton, A., Hallam, S., Peffers, J., Robertson, P., Stobart, G., 2007. Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools. *Br. Educ. Res. J.* 33 (4), 605–627.
- Kostons, D., van Gog, T., Paas, F., 2012. Training self-assessment and task-selection skills: a cognitive approach to improving self-regulated learning. *Learn. Instr.* 22 (2), 121–132.
- Lai, C.Y., Chen, L.J., Yen, Y.C., Lin, K.Y., 2020. Impact of video annotation on undergraduate nursing students' communication performance and commenting behaviour during an online peer-assessment activity. *Australas. J. Educ. Technol.* 36 (2), 71–88. <https://doi.org/10.14742/ajet.4341>.
- Lau, P.N., 2020. Enhancing formative and self-assessment with video playback to improve critique skills in a titration laboratory. *Chem. Educ. Res. Pract.* 21, 178–188. <https://doi.org/10.1039/C9RP00056A>.
- Li, L., Liu, X.Y., Zhou, Y.C., 2012. Give and take: a re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *Br. J. Educ. Technol.* 43 (3), 376–384.
- Li, H., Xiong, Y., Hunter, C.V., Guo, X., Tywoniu, R., 2020. Does peer assessment promote student learning? A meta-analysis. *Assess. Eval. High Educ.* 45 (2), 193–211.
- Li, H., Xiong, Y., Zang, X., Kornhaber, M.L., Lyu, Y., Chung, K.S., Suen, H.K., 2016. Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings. *Assess. Eval. High Educ.* 41 (2), 245–264.
- Lin, K.C., Yang, S.H., Hung, J.C., Wang, D.M., 2006. Web-based appreciation and peer-assessment for visual-art education. *Int. J. Dist. Educ. Technol.* 4 (4), 5–14.
- Lu, J., Law, N., 2012. Online peer assessment: effects of cognitive and affective feedback. *Instr. Sci.* 40 (2), 257–275. <https://doi.org/10.1007/s11251-011-9177-2>.
- Malecki, C.K., Elliott, C.N., 2002. Children's social behaviors as predictors of academic achievement: a longitudinal analysis. *Sch. Psychol. Q.* 17, 1–23.
- Mort, J.R., Hansen, D.J., 2010. First-year pharmacy students' self-assessment of communication skills and the impact of video review. *Am. J. Pharmaceut. Educ.* 74 (5), 78. <https://doi.org/10.5688/aj740578>.
- O'Donnell, A.M., Topping, K.J., 1998. Peers assessing peers: possibilities and problems. In: Topping, K., Ehly, S. (Eds.), *Peer-Assisted Learning*. Lawrence Erlbaum, Mahwah, NJ.
- Panadero, E., Broadbent, J., Boud, D., et al., 2019. Using formative assessment to influence self- and co-regulated learning: the role of evaluative judgement. *Eur. J. Psychol. Educ.* 34, 535–557. <https://doi.org/10.1007/s10212-018-0407-8>.

- Panadero, E., Brown, G.L., Strijbos, J.W., 2016. The future of student self-assessment: a review of known unknowns and potential directions. *Educ. Psychol. Rev.* 28, 803–830. <https://doi.org/10.1007/s10648-015-9350-2>.
- Panadero, E., Jonsson, A., Botella, J., 2017. Effects of self-assessment on self-regulated learning and self-efficacy: four meta-analyses. *Educ. Res. Rev.* 22, 74–98.
- Panadero, E., Tapia, J.A., Huertas, J.A., 2012. Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learn. Indiv. Differ.* 22 (6), 806–813.
- Ratminingsih, N.M., Marhaeni, A.A.I.N., Vigayanti, L.P.D., 2018. Self-assessment: the effect on students' independence and writing competence. *Int. J. Instr.* 11 (3), 277–290.
- Rockhill, C.M., Asher, S.R., 1992. Peer Assessment of the Behavioral Characteristics of Poorly Accepted Boys and Girls. Educational Resources Information Center Document Reproduction Service ED346372.
- Ross, J.A., 2006. The reliability, validity, and utility of self-assessment. *Pract. Assess. Res. Eval.* 11 (10), 1–13.
- Scruggs, T.E., Mastropieri, M.A., 1998. Tutoring and students with special needs. In: Topping, K.J., Ehly, S. (Eds.), *Peer-Assisted Learning*. Lawrence Erlbaum, Mahwah, NJ.
- Sebba, J., Crick, R.D., Yu, G.X., Lawson, H., Harlen, W., Durant, K., 2008. Systematic Review of Research Evidence of the Impact on Students in Secondary Schools of Self and Peer Assessment. Report no. 1614. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.
- Sluijsmans, D.M.A., Brand-Gruwel, S., van Merriënboer, J.J.G., 2002. Peer assessment training in teacher education: effects on performance and perceptions. *Assess. Eval. High Educ.* 27 (5), 443–454. <https://doi.org/10.1080/0260293022000009311>.
- Strijbos, J.W., Narciss, S., Dünnebier, K., 2010. Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learn. Instr.* 20 (4), 291–303.
- Sung, Y.T., Chang, K.E., Chang, T.H., Yu, W.C., 2010. How many heads are better than one? The reliability and validity of teenagers' self- and peer assessments. *J. Adolesc.* 33 (1), 135–145.
- Sung, Y.T., Chang, K.E., Chiou, S.K., Hou, H.T., 2005. The design and application of a web-based self- and peer-assessment system. *Comput. Educ.* 45 (2), 187–202.
- Tillema, H., Leenknicht, M., Segers, M., 2011. Assessing assessment quality: criteria for quality assurance in design of (peer) assessment for learning – a review of research studies. *Stud. Educ. Eval.* 37 (1), 25–34.
- Topping, K.J., 1998. Peer assessment between students in college and university. *Rev. Educ. Res.* 68 (3), 249–276.
- Topping, K.J., 2018. In: MacMillan, J.H. (Ed.), *Using Peer Assessment to Inspire Reflection and Learning*. Student Assessment for Educators Series. Routledge, New York & London. ISBN: 978-0-8153-6765-9 (pbk). www.routledge.com/9780815367659.
- Torres-Guijarro, S., Bengoechea, M., 2017. Gender differential in self-assessment: a fact neglected in higher education peer and self-assessment techniques. *High Educ. Res. Dev.* 36 (5), 1072–1084. <https://doi.org/10.1080/07294360.2016.1264372>.
- Tseng, S.C., Tsai, C.C., 2007. On-line peer assessment and the role of peer feedback: a study of a high school computer course. *Comput. Educ.* 49 (4), 1161–1174.
- Van Gennip, N.A.E., Segers, M., Tillema, H.M., 2009. Peer assessment for learning from a social perspective: the influence of interpersonal variables and structural features. *Educ. Res. Rev.* 4 (1), 41–54.
- Van Steendam, E., Rijlaarsdam, G., Sercu, L., Van den Bergh, H., 2010. The effect of instruction type and dyadic or individual emulation on the quality of revision in ESL. *Learn. Instr.* 20 (4), 316–327.
- Vasu, K.A.P., Fung, Y.M., Nimehchisalem, V., Rashid, S.M., 2020. Self-regulated learning development in undergraduate ESL writing classrooms: teacher feedback versus self-assessment. *Regional Lang. Centre J.* <https://doi.org/10.1177/0033688220957782>.
- Yan, Z., 2020. Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assess. Eval. High Educ.* 45 (2), 224–238. <https://doi.org/10.1080/02602938.2019.1629390>.
- Yan, Z., Brown, G.T.L., Lee, J.C.K., Qiu, X.I., 2020a. Student self-assessment: why do they do it? *Educ. Psychol.* 40 (4), 509–532.
- Yan, Z., Chiu, M.M., Ko, P.Y., 2020b. Effects of self-assessment diaries on academic achievement, self-regulation, and motivation. *Assess. Educ. Princ. Pol. Pract.* 27 (5), 562–583. <https://doi.org/10.1080/0969594X.2020.1827221>.
- Yang, J.C., Ko, H.W., Chung, I.L., 2005. Web-based interactive writing environment: development and evaluation. *Educ. Technol. Soc.* 8 (2), 214–229.
- Yu, F.Y., 2011. Multiple peer-assessment modes to augment online student question-generation processes. *Comput. Educ.* 56 (2), 484–494.
- Yurdabakan, I., 2011. The investigation of peer assessment in primary school cooperative learning groups with respect to gender. *Education* 39 (2), 153–169, 3–13.
- Zhan, Y., 2020. What matters in design? Cultivating undergraduates' critical thinking through online peer assessment in a Confucian heritage context. *Assess. Eval. High Educ.* 46 (4), 615–630. <https://doi.org/10.1080/02602938.2020.1804826>.
- Zheng, L.Q., Zhang, X., Cui, P.P., 2020. The role of technology-facilitated peer assessment and supporting strategies: a meta-analysis. *Assess. Eval. High Educ.* 45 (3), 372–386. <https://doi.org/10.1080/02602938.2019.1644603>.

Feedback and classroom practice

Siv M. Gamlem and Kim-Daniel Vattøy, Faculty of Humanities and Teacher Education, Volda University College, Volda, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	89
Classroom feedback informed by responsive pedagogy	90
A framework for classroom feedback practice as responsive pedagogy	91
Goals	91
Feedback strategy	91
Feedback content	91
Feedback dialog	92
Self-efficacy	92
Emotions	92
A model of effective classroom feedback informed by responsive pedagogy	93
Conclusion	94
References	94

Glossary

Classroom feedback practice How feedback is exercised in classroom contexts based on the beliefs teachers and students hold about feedback

Feedback agency A learner's adaptive or maladaptive agency in relation to internal feedback generated by the self and/or external feedback provided by others

Feedback dialog Internal or external feedback provided and generated through dialogic interactions

Responsive pedagogy The recursive dialog between a student's internal feedback and feedback provided by external agents (e.g., teachers, peers, parents, or digital resources)

Self-regulated learning How a learner metacognitively, motivationally, emotionally, and behaviorally regulates one's learning processes

Self-efficacy A learner's beliefs in one's capabilities to exercise influence and execute actions to attain desired goals

Introduction

This conceptual paper proposes a framework for classroom feedback practice as responsive pedagogy, based on key trends in the international research literature. Responsive pedagogy is defined as the recursive dialog between a student's internal feedback and feedback provided by external agents (e.g., teachers, peers, parents, or digital resources). Feedback is a powerful influencer of student learning when it clarifies understanding, bolsters motivation, and supports improvement (Ajjawi et al., 2022; Black and Wiliam, 1998a; Gamlem and Smith, 2013; Hattie and Timperley, 2007; Orsmond et al., 2005; Shute, 2008). Several studies have highlighted the importance of dialogic feedback for students' learning and achievement (e.g., Askew and Lodge, 2000; Carless, 2020; Gamlem and Smith, 2013). However, empirical studies have found an absence of dialogic feedback interactions in observed lessons (e.g., Gamlem and Smith, 2013). Classroom interactions are too complex for research to tell teachers what to do, but research can inform and support teachers' feedback practices. In addition, teachers exercise professional judgment in feedback interactions, which are realized through unplanned pedagogical moments of classroom interactions (Smith, 2015; Wyatt-Smith and Klenowski, 2013). When feedback is used in a formative process, students become more active and engaged in regulating their learning (Andrade, 2010; Brandmo et al., 2020; Gamlem et al., 2019; Vattøy and Gamlem, 2020).

Classroom feedback practices have resisted change due to a focus on one-way information delivery to students (Boud and Molloy, 2013; Gamlem, 2015a). A transmission model of feedback is argued for as questionable as it assumes a passive student role (Ajjawi et al., 2022; Gamlem and Smith, 2013; van der Kleij et al., 2019). Furthermore, feedback has often been interpreted by teachers and students as information provided by an agent—for example, a teacher, peer, or self—about aspects of performance or understanding (Hattie and Timperley, 2007). Since students need support and information on how to improve, this approach is insufficient for the implementation of powerful feedback processes (Andrade and Brookhart, 2020). Goals, dialogs, emotions, student agency, and opportunities to make sense of comments are essential in feedback practices (Harris et al., 2018; Panadero and Lipnevich, 2022; Winstone and Carless, 2020; Winstone et al., 2017). Responsive pedagogy emphasizes the importance of dialogic feedback interactions in classroom contexts, which ultimately open the door for more equitable feedback dialogs with students seeking feedback from and providing it to teachers (Smith et al., 2016; Vattøy and Smith, 2019).

Feedback in classrooms has often taken the form of written comments or grades on an assessment task; this is an approach which often focuses on justifying an assessed grade and controlling that the assignments are completed (Black and Wiliam, 1998a; Gamlem and Smith, 2013; Gamlem 2015a). One of the most significant limitations of this accountability model of feedback is that it does little in supporting the agency of students in improving their work (Askew and Lodge, 2000; Boud and Molloy, 2013). Research on responsive pedagogy has identified a hidden accountability system that impedes effective feedback dialogs due to a systemic focus on examinations and testing of students (Vattøy, 2020; Vattøy et al., 2020). Empirical studies have documented that teachers and student express different views as to the usefulness of feedback (Carless, 2006; Gamlem, 2015a). In many situations teachers' feedback in classrooms is neither recognized by students nor perceived as useful (Gamlem and Smith, 2013; van der Kleij and Adie, 2020; Vattøy et al., 2020). Indeed, much of the literature has documented students' barriers, discomfort, and irritation with feedback provided by a teacher (e.g., Winstone et al., 2017).

Conceptualized within a learning process, classroom feedback is not just a way of accounting for students' achievements—it is also integral to teaching and learning, where teachers and students have important roles to play (Andrade and Brookhart, 2020; Gamlem, 2015a; Gamlem and Smith, 2013; Henderson et al., 2019). The importance of students actively engaging with feedback processes has been emphasized in studies that have focused on how students perceive classroom feedback (Gamlem and Smith, 2013; Lui and Andrade, 2022; Winstone et al., 2017). In addition, Winstone et al. (2017, p. 17) argued for students' proactive receipt of feedback and emphasized "the fundamental contribution and responsibility of the learner." If feedback is not received by students, or not perceived as intended by the teacher or peer, it is unlikely that the feedback will achieve its intended effect of supporting student learning (Gamlem and Smith, 2013; Perrenoud, 1998; Sadler, 2010; van der Kleij and Adie, 2020). A study by van der Kleij and Adie (2020) provides evidence that feedback perceptions—and thus feedback effectiveness—are context-dependent, subject-dependent, and individual-dependent. In this view, the effectiveness of classroom feedback is dependent on whether students perceive it as effective. As stated by Gamlem (2015a), feedback processes might be hard to manage. The emerging concept of teacher feedback literacy has highlighted the teacher's role as a designer of classroom settings that promote students' feedback literacy (Carless and Winstone, 2020). Teacher feedback literacy consists of designing effective feedback processes, responding to the relational aspects of feedback and navigating tensions between the different functions of feedback. The development of student and teacher feedback literacy benefits from situations when teachers and students share goals, strategies, and responsibilities (Carless and Winstone, 2020). In contrast, a focus on feedback as "telling" might explain why students do not recognize all forms of feedback they are exposed to, such as peer feedback or oral feedback (Ajjawi et al., 2022; McLean et al., 2015; Sadler, 2010).

Classroom feedback informed by responsive pedagogy

Recently, there has been a call to embed assessment and feedback in the wider field of pedagogy (Black and Wiliam, 2018; Smith, 2015). This is a realization that learning-oriented assessment is not merely an ambitious goal but a profound shift in thinking and habits (Earl, 2013; Wiliam, 2020). Responsive pedagogy embeds classroom feedback in pedagogy and draws attention to the relevance of feedback dialogs in classroom settings (Smith, 2015; Smith et al., 2016; Gamlem et al., 2019; Vattøy and Gamlem, 2020). A dialogic model of feedback has been connected to a more active student role (van der Kleij et al., 2019). In responsive pedagogy, a key concern is activating students as agents in their learning processes through feedback dialogs (Smith et al., 2016; Vattøy and Gamlem, 2020).

A goal in responsive pedagogy is to capitalize on teacher-student interactions, emphasizing students' internal feedback dialog, to strengthen self-efficacy and self-regulated learning (Smith et al., 2016; Vattøy and Gamlem, 2020). This approach aligns with research by Nicol (2021), where "internal feedback" is argued for as an approach to extending the idea of student-centeredness in feedback. Nicol argues that a process of making inner comparisons is the underlying mechanism at play in all feedback processes (Ajjawi et al., 2022; Nicol, 2021). The interplay between internal and external feedback has highlighted the importance of students' metacognitive processes and self-monitoring for their self-regulated learning (Butler and Winne, 1995). The student's internal feedback is understood through the three phases of self-regulation: forethought, performance, and self-reflection (Zimmerman et al., 2017). These three phases can be exemplified by the fact that when a student is given an assignment, the student will begin an internal dialog with themselves about what the assignment is about and their competence in relation to goal. The student forms a pre-understanding of his or her personal knowledge needed to complete the assignment (forethought phase). Furthermore, the student will search for and decide what strategies are needed to achieve the learning objectives. In the performance phase the student actively monitors his or her learning processes. This is the student's internal dialog that the teacher, in his or her interaction with the student, tries to understand and respond to through a recurring feedback dialog that can support the student's learning. These dialogs can provide clarity and deeper understanding of the outcome of an assignment based on learning objectives and assessment criteria (self-reflection phase).

The recursive dialog between internal and external feedback highlights the importance of teacher scaffolding in students' self-regulation processes (Butler and Winne, 1995; Gamlem et al., 2019; Smith et al., 2016). Self-regulated learning has been defined as "an active, constructive process whereby students set goals for their learning and then attempt to monitor, regulate and control their cognition, motivation and behavior" (Pintrich, 2000, p. 453). The active self-monitoring process in responsive pedagogy is interconnected with external feedback from teachers, peers, or parents (Smith et al., 2016). The concept of feedback agency is relevant in this context, as students exercise adaptive or maladaptive feedback agency in feedback processes (Harris et al., 2018; Vattøy et al., 2020). The key concern for the teacher is to tap into the internal feedback dialog of the student in unplanned pedagogical

moments to scaffold the student's learning processes (Smith et al., 2016). In recent years, several researchers have pointed to the notion of feedback as the crucial interaction between teachers and students carried out for the purpose of fostering student learning (e.g., Gamlem and Smith, 2013; Hawe et al., 2008; Shute, 2008; Vattøy and Gamlem, 2020).

A framework for classroom feedback practice as responsive pedagogy

The proposed framework for classroom feedback practice as responsive pedagogy comprises six dimensions: goals, feedback strategy, feedback content, feedback dialog, self-efficacy, and emotions. The framework explains how classroom feedback functions as a guide for teachers' and students' feedback processes. Feedback agency is central for understanding how these six dimensions interact and influence one another. These dimensions are explained in detail below.

Goals

The connection between learning goals and self-regulated learning is central in responsive pedagogy (Gamlem et al., 2019; Gamlem and Munthe, 2014). Students exercise feedback agency by internalizing or rejecting external goals (Butler and Winne, 1995; Harris et al., 2018; Lui and Andrade, 2022). Goal orientation is linked to self-regulated learning as students set goals for their learning and aim to regulate their cognition, motivation, and behavior (Brandmo et al., 2020; Pintrich, 2000). Self-regulated learning can be described as the process where the learner activates and sustains their behaviors, cognition, motivation, and feelings toward the attainment of a learning goal (Schunk and Green, 2018).

External goal orientation is defined as "students' comprehension and activation or rejection of external learning goals set by the teacher" (Vattøy and Smith, 2019, p. 261). A study that examined dimensions of responsive pedagogy found that external goal orientation was the strongest predictor of students' perceptions of teachers' feedback practice (Vattøy and Smith, 2019). Students also felt more confident in their skills and were able to regulate when they knew the goals related to the teachers' feedback. However, classroom studies that have examined the role of feedback have indicated that teachers struggle to facilitate clear learning goals, which might impede students' own goal setting (e.g., Gamlem and Munthe, 2014; Vattøy and Gamlem, 2020). Attention to learning goals when engaging in learning activities and achievement tasks is important for the teacher and student (Brandmo et al., 2020; Gamlem and Munthe, 2014; Lui and Andrade, 2022). Without clear learning intentions and goals, students may have problems knowing in which direction to move (Hattie and Timperley, 2007; Sadler, 1998, 2010). Furthermore, feedback should be linked with learning goals (Gamlem and Smith, 2013). Feedback dialogs are important for students to identify criteria and goals (Sadler, 2010). Feedback that supports students' self-regulated learning can get students to set their own goals related to the task they are given, to plan how they can solve the task, and to check whether they are on the right track (Andrade, 2010; Andrade and Brookhart, 2020; Sadler, 2010). This aligns with self-regulated learning, which is defined in responsive pedagogy by inspiration from the work of Pintrich and Zuscho (2002): an active constructive process in which students set goals for their learning and monitor, regulate, and control their cognition, motivation, emotions, and behavior, guided and limited by their goals and contextual characteristics of the environment.

Feedback strategy

The development of feedback strategies is a key concern in responsive pedagogy, with the emphasis on how students gain strategic knowledge and skills in seeking feedback from teachers and providing feedback to their peers (Smith et al., 2016; Vattøy and Smith, 2019). Feedback strategies can vary in timing (when given, how often), amount (how many points made, how much about each point), mode (oral, written, visual/demonstration), and audience (individual, group, class, teacher, student), (Brookhart, 2008; Panadero and Lipnevich, 2022; Gamlem and Smith, 2013). Students choose strategies when working independently and engaging in peer feedback dialogs. Feedback as a process in which students and teachers make sense of information about their performance and use it to enhance the quality of their work or learning strategies is key (Boud and Molloy, 2013; Carless, 2020).

Working with feedback does not come naturally to all students, and researchers have stated that teachers should teach students how to use received feedback and how to seek feedback to improve their work and increase performance (Ajjawi et al., 2022; Andrade, 2010; Gamlem and Munthe, 2014; Henderson et al., 2019; Sadler, 2010). Students are often unused to or prevented by constraints of the classroom context from seeking feedback, or when they do, they may be unable to judge the relevance of the comments (Gamlem and Smith, 2013; Winstone et al., 2017). Students need to be trained in how to interpret feedback, how to make connections between the feedback and the characteristics of the work they produce, and how they can improve their work in the future by seeking and using feedback (Black and Wiliam, 2009; Henderson et al., 2019; Sadler, 2010).

Feedback content

Feedback content is essential for students' understanding of performance and how to improve (Black and Wiliam, 1998a; Panadero and Lipnevich, 2022; Perrenoud, 1998; Shute, 2008). Students who are not provided with strategic feedback or with clear explanations of their specific errors or what made their work superb might have difficulties understanding why they are performing poorly or excellently and what they need to do to improve (Andrade and Brookhart, 2020; Black and Wiliam, 1998b).

Feedback content varies in focus (task, process, self-regulation, self), comparison (criterion-referenced, norm-referenced, self-referenced), function (description, judgment), valence (positive, negative), clarity (clear/unclear to a recipient), specificity (specific, general), and tone (implication, climate), (Brookhart, 2008; Gamlem and Smith, 2013; Hattie and Timperley, 2007; Panadero and Lipnevich, 2022).

However, the effectiveness of feedback content is dependent on both teachers and students. Empirical research has documented teachers' difficulties enhancing students' deeper understanding of content in feedback interactions (e.g., Vattøy and Gamlem, 2020). Feedback quality and feedback quantity have been identified as two important predictors of students' use of feedback (Vattøy et al., 2020). Henderson et al. (2019) emphasized the importance of both teachers and students understanding that feedback is a process, and comments given by a teacher are only one possible part of the process. Critically, students need to be able to make sense of feedback and then act upon it (Carless and Boud, 2018; Lui and Andrade, 2022; Sadler, 2010). As explained by Sadler (2010), feedback as the act of "telling" is not effective in facilitating student learning because the connection between feedback, students' work, and future learning relies on clear and unambiguous interpretations of that feedback, and often the key messages remain invisible to students. An implication is that classroom feedback needs to be constructed so that it can be easily understood and acted upon (Ajjawi et al., 2022; Gamlem and Smith, 2013; Henderson et al., 2019; Winstone et al., 2017). The focus has thus shifted from the teacher to the teacher and student(s) working in partnership (Ajjawi et al., 2022; Carless, 2020; Hawe et al., 2008).

Feedback dialog

The emphasis on feedback dialogs is important for students' active participation in giving, seeking, and receiving feedback to promote learning. The backdrop for the development of feedback dialogs in responsive pedagogy has sociocultural roots in the notion of a Vygotskyan dialog between internal and external feedback (Gamlem, 2015b; Smith, 2015; Vygotsky, 1978). Vygotsky's belief that language is fundamental to the process of learning and dependent upon the complex interplay of thought and language in shaping meaning is at the heart of how classroom talk and feedback promote learning, as conceptualized in responsive pedagogy (Smith et al., 2016; Vygotsky, 1962). Central to Vygotsky's sociocultural theory of learning is the idea that conceptual knowledge first appears between people on an interpsychological level and then inside the learner on an intrapsychological level (Vygotsky, 1978). The notion of a teacher assisting student performance in the zone of proximal development also suggests that teachers can guide discourse on the interpsychological level to support student learning (Gamlem, 2015b; Mercer and Littleton, 2007). This recognizes the importance of teacher-student discourse in the classroom, which may be considered a form of scaffolding, although the latter was originally conceived in the context of tutoring a single individual in problem-solving (Wood et al., 1976). In responsive pedagogy, teachers tap into and capitalize on students' internal feedback dialogs about emotions, self-efficacy, and self-regulation processes via the external feedback that occurs recursively (Gamlem et al., 2019; Vattøy, 2020; Vattøy and Gamlem, 2020). Carless (2020) emphasizes the importance of partnerships between teachers and students in feedback dialogs, where teachers have as much to learn as students. For example, teacher-student feedback dialogs give a teacher a chance to see whether feedback content was appropriate and whether his or her teaching prepared the students adequately for the learning goals.

Self-efficacy

Self-efficacy is defined as a learner's beliefs in his or her capabilities to exercise, influence, and execute actions to attain desired goals (Bandura, 1977; Zimmerman et al., 2017). This concerns the learner's ability to organize and carry out actions required to achieve a given result (Bandura, 1997; Lui and Andrade, 2022). Self-efficacy as a concept originates from Bandura's (1977) social-cognitive theory, involving the individual's development and exercise of self-control over desired outcome. Zimmerman (2000, p. 86) stated that "self-efficacy beliefs have shown convergent validity in influencing such key indices of academic motivation as choice of activities, persistence, level of effort, and emotional reactions." In classroom contexts, students' beliefs about their efficacy to manage academic tasks can also influence them emotionally by decreasing anxiety, depression, and stress (Bandura, 1997; Gamlem et al., 2019; Pekrun et al., 2002). An intervention study focused on responsive pedagogy found short-term effects on students' self-efficacy, motivation, and self-regulation (Gamlem et al., 2019). Such results seem to indicate the relevance of training teachers specifically in terms of facilitating feedback dialogs that promote students' beliefs in their capabilities. Yet, unrealistic expectations about grades and a summative accountability system have been identified as negative factors working against teachers' work in supporting students' self-efficacy (Vattøy, 2020).

Emotions

Emotions facilitate or impede students' self-regulation of learning (Ajjawi et al., 2022; Pekrun et al., 2002). Winstone and Carless (2020, p. 150) state that "the relationship between emotions and feedback is complex," since praise can elicit positive emotions, and at the same time it can undermine the motivation to invest future effort. Pekrun (2006) found that emotions differ not only according to valence (positive/negative) but also students' activation potential. Students activate emotions by initiating action and effort or deactivate them by inhibiting action and effort. Positive activating emotions such as enjoyment of learning may generally enhance academic motivation, whereas negative emotions may be detrimental (Harley et al., 2019; Panadero and Lipnevich, 2022;

Pekrun et al., 2002). Both negative (e.g., shame) and positive (e.g., relief) emotions can lead to outcomes that are either aligned or misaligned with students' academic goals, depending on how they deal with these emotions (Ajjawi et al., 2022; Harley et al., 2019). Emotional aspects (e.g., enjoyment or anxiety) have been found to be significantly associated with other dimensions of responsive pedagogy, such as self-efficacy and self-regulation (Gamlem et al., 2019). Enjoyment has also been found to significantly predict students' external goal orientation, self-regulation, and perceived usefulness of teachers' feedback practice (Vattøy and Smith, 2019). Emotions and feedback are connected, as students who invest a high degree of energy and effort but receive disapproval from their teacher might lose motivation (Gamlem and Smith, 2013; Panadero and Lipnevich, 2022). Emotional aspects such as mistrust and negativity have been found to cause maladaptive feedback agency in students (Gamlem and Smith, 2013; Vattøy et al., 2020). Emotions might become more relevant and influence how feedback is interpreted and acted on, due to students' active and interpretive stance (Ajjawi et al., 2022).

A model of effective classroom feedback informed by responsive pedagogy

Fig. 1 visualizes the reciprocal relationship between the dimensions that influence effective classroom feedback informed by responsive pedagogy: goals, strategy, feedback content, feedback dialog, self-efficacy, and emotions. The arrows visualize how each dimension affects and is affected by other dimensions. Feedback agency is at the core of understanding the model, as students exercise adaptive or maladaptive influence on teaching and learning processes.

Effective classroom feedback practice as responsive pedagogy is dialogic. The student starts with, for example, goals, selects strategies, and monitors his or her processes during the provision and generation of feedback content. Students and teachers engage in recursive feedback dialogs in which internal feedback is externalized and becomes clearer to the student. The clarity of the preceding phases causes the student to have increased self-efficacy, supported by the teacher. A positive learning experience and feedback interaction causes positive emotions and attainment. However, these processes might transpire differently (e.g., emotional reactions as points of departure).

Effective classroom feedback involves continually challenging and improving one's practice (Gamlem, 2015a; Henderson et al., 2019; Vattøy et al., 2020). This requires a degree of self-reflection on the part of teachers and students, along with an inquiry mindset (Smith, 2015). The proposed model of effective classroom feedback as responsive pedagogy might increase teachers' valuation of students' experiences of feedback. Although classroom interactions are inherently complex and sometimes bewildering, the model of effective classroom practice might support teachers' professional judgment processes. The model also contributes to understanding the dimensions that students are affected by in their feedback interactions, which consequently affect their learning. Understanding students' encounters with classroom feedback might contribute to understanding students' learning trajectories and support teachers in providing support in feedback processes. Recognizing students' feedback agency implies understanding that students exercise adaptive or maladaptive influences in feedback dialogs (Harris et al., 2018; Vattøy et al., 2020). Thus, a key concern in classroom feedback practice is supporting students' feedback agency processes (as visualized in the six dimensions) to strengthen their learning.

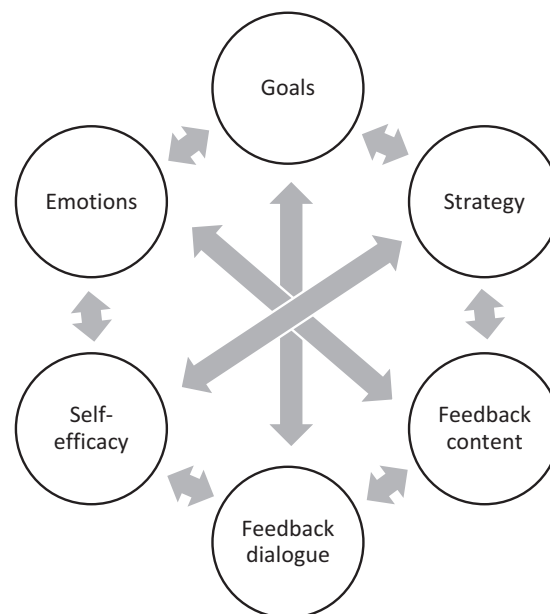


Fig. 1 A model of classroom feedback agency processes informed by responsive pedagogy.

Conclusion

This paper proposed a framework for classroom feedback practice as responsive pedagogy comprising six interacting dimensions: goals, strategy, feedback content, feedback dialog, self-efficacy, and emotions. Furthermore, a model was introduced to visualize how the dimensions interact and influence one another. A fundamental aspect of the model is how students' feedback agency processes are acknowledged and actively worked with in classroom feedback dialogs. A key concern is how teachers support students in setting goals and selecting strategies and how they monitor their learning processes during the work with feedback content. Teachers and students engage in recursive feedback dialogs that activate internal feedback through external dialogs. Teachers support students' self-efficacy and beliefs in their capabilities, recognizing and caring for students' emotions in feedback processes. Since students exercise feedback agency, the interacting dimensions might alter the way the dimensions affect feedback agency processes. We argue that responsive pedagogy seems to be an approach to supporting students' internal feedback processes via teacher-student feedback dialogs in classroom practice.

References

- Andrade, H., 2010. Students as the definitive source of formative assessment. In: Andrade, H., Cizek, G. (Eds.), *Handbook of Formative Assessment*. Routledge, New York, pp. 90–105.
- Andrade, H., Brookhart, S., 2020. Classroom assessment as the co-regulation of learning. *Assess. Educ. Princ. Pol. Pract.* 27 (4), 350–372. <https://doi.org/10.1080/0969594X.2019.1571992> [online]. Available at. (Accessed 1 July 2022).
- Askew, S., Lodge, C., 2000. Gifts, ping-pong and loops—linking feedback and learning. In: Askew, S. (Ed.), *Feedback for Learning*. Routledge Falmer, London, pp. 1–18.
- Ajjawi, R., Tai, J., Dawson, P., 2022. Feedback for Learning. Encyclopedia, EDU4 Section: Assessment and Accountability [Cross-ref].
- Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84 (2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191> [online]. Available at. (Accessed 5 July 2020).
- Bandura, A., 1997. *Self-efficacy: The Exercise of Control*. Freeman, New York.
- Black, P., William, D., 1998a. Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5 (1), 7–74. <https://doi.org/10.1080/0969595980050102> [online]. Available at. (Accessed 5 July 2020).
- Black, P., William, D., 1998b. Inside the black box. raising standards through classroom assessment. *Phi Delta Kappan* 80 (2), 139–148. <https://doi.org/10.1177/003172171009200119> [online]. Available at. (Accessed 5 July 2020).
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21 (5), 5–31. <https://doi.org/10.1007/s11092-008-9068-5> [online]. Available at. (Accessed 5 July 2020).
- Black, P., William, D., 2018. "Classroom assessment and pedagogy", assessment in education: principles. *Pol. Pract.* 25 (6). <https://doi.org/10.1080/0969594X.2018.1441807> [online]. Available at. (Accessed 5 July 2020).
- Boud, D., Molloy, E., 2013. Rethinking models of feedback for learning: the challenge of design. *Assess. Eval. High. Educ.* 38 (6). <https://doi.org/10.1080/02602938.2012.691462> [online]. Available at. (Accessed 5 July 2020).
- Brandmo, C., Panadero, E., Hopfenbeck, T., 2020. Bridging classroom assessment and self-regulated learning. *Assess. Educ. Princ. Pol. Pract.* 27 (4), 319–331. <https://doi.org/10.1080/0969594X.2020.1803589> [online]. Available at. (Accessed 1 July 2022).
- Brookhart, S., 2008. *How to Give Effective Feedback to Your Students?* Association for Supervision and Curriculum Development, Alexandria, Virginia.
- Butler, D.L., Winnie, P.H., 1995. Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.* 65 (3), 245–281. <https://doi.org/10.2307/1170684> [online]. Available at. (Accessed 5 July 2020).
- Carless, D., 2006. Differing perceptions in the feedback process. *Stud. High. Educ.* 31 (2). <https://doi.org/10.1080/03075070600572132> [online]. Available at. (Accessed 5 July 2020).
- Carless, D., 2020. Longitudinal perspectives on students' experiences of feedback: a need for teacher–student partnerships. *High. Educ. Res. Dev.* 39 (3), 425–438. <https://doi.org/10.1080/07294360.2019.1684455> [online]. Available at. (Accessed 5 July 2020).
- Carless, D., Boud, D., 2018. The development of student feedback literacy: enabling uptake of feedback. *Assess. Eval. High. Educ.* 43 (8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354> [online]. Available at. (Accessed 5 July 2020).
- Carless, D., Winstone, N., 2020. Teacher feedback literacy and its interplay with student feedback literacy. *Teach. High. Educ.* <https://doi.org/10.1080/13562517.2020.1782372> [online]. Available at. (Accessed 5 July 2020).
- Earl, L.M., 2013. *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*, second ed. Corwin, Thousand Oaks, CA.
- Gamlem, S.M., 2015a. Feedback to support learning: changes in teachers' practice and beliefs. *Teach. Dev.* 19 (4). <https://doi.org/10.1080/13664530.2015.1060254> [online]. Available at. (Accessed 5 July 2020).
- Gamlem, S.M., 2015b. *Tilbakemelding for læring og utvikling. Feedback to Support Learning and Development*. Gyldendal Akademisk forlag, Oslo.
- Gamlem, S.M., Munthe, E., 2014. Mapping the quality of feedback to support students' learning in lower secondary classrooms. *Camb. J. Educ.* 44 (1). <https://doi.org/10.1080/0305764X.2013.855171> [online]. Available at. (Accessed 5 July 2020).
- Gamlem, S.M., Smith, K., 2013. "Student perceptions of classroom feedback", assessment in education: principles. *Pol. Pract.* 20 (2). <https://doi.org/10.1080/0969594X.2012.749212> [online]. Available at. (Accessed 5 July 2020).
- Gamlem, S.M., Kvinge, L.M., Smith, K., Engelsen, K.S., 2019. "Developing teachers' responsive pedagogy in mathematics, does it lead to short-term effects on student learning?". *Cogent Educ.* 6 (1). <https://doi.org/10.1080/2331186X.2019.1676568> [online]. Available at. (Accessed 5 July 2020).
- Harley, J.M., Pekrun, R., Taxer, J.L., Gross, J.J., 2019. Emotion regulation in achievement situations: an integrated model. *Educ. Psychol.* 54 (2). <https://doi.org/10.1080/00461520.2019.1587297> [online]. Available at. (Accessed 5 July 2020).
- Harris, L.R., Brown, G.T.L., Dargusch, J., 2018. Not playing the game: student assessment resistance as a form of agency. *Aust. Educ. Res.* 45 (1). <https://doi.org/10.1007/s13384-018-0264-0> [online]. Available at. (Accessed 5 July 2020).
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77 (1), 81–112. <https://doi.org/10.3102/003465430298487> [online]. Available at. (Accessed 5 July 2020).
- Haue, E., Dixon, H., Watson, E., 2008. Oral feedback in the context of written language. *Aust. J. Lang. Lit.* 31 (1), 43–58 [online]. Available at. <https://www.questia.com/library/journal/1G1-195070320/oral-feedback-in-the-context-of-written-language>. (Accessed 5 July 2020).
- Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E., Mahoney, P., 2019. Conditions that enable effective feedback. *High. Educ. Res. Dev.* 38 (7). <https://doi.org/10.1080/07294360.2019.1657807> [online]. Available at. (Accessed 5 July 2020).
- Lui, A.M., Andrade, H.L., 2022. The next black box of formative assessment: a model of the internal mechanisms of feedback processing. *Front. Educ.* 7, 751548. <https://doi.org/10.3389/educ.2022.751548> [online]. Available at. (Accessed 4 July 2022).

- McLean, A.J., Bond, C.H., Nicholson, H.D., 2015. An anatomy of feedback: a phenomenographic investigation of undergraduate students' conceptions of feedback. *Stud. High. Educ.* 40 (5), 921–932. <https://doi.org/10.1080/03075079.2013.855718> [online]. Available at. (Accessed 5 July 2020).
- Mercer, N., Littleton, K., 2007. *Dialogue and the Development of Children's Thinking. A Sociocultural Approach*. Routledge, New York.
- Nicol, D., 2021. The power of internal feedback: exploiting natural comparison processes. *Assess. Eval. High. Educ.* 46 (5), 756–778. <https://doi.org/10.1080/02602938.2020.1823314> [online]. Available at. (Accessed 1 July 2022).
- Orsmond, P., Merry, S., Reiling, K., 2005. Biology students' utilization of tutors' formative feedback: a qualitative interview study. *Assess. Eval. High. Educ.* 30 (4), 369–389. <https://doi.org/10.1080/02602930500099177> [online]. Available at. (Accessed 5 July 2020).
- Panadero, E., Lipnevich, A.A., 2022. A review of feedback models and typologies: towards an integrative model of feedback elements. *Educ. Res. Rev.* 35. <https://doi.org/10.1016/j.edurev.2021.100416> [online]. Available at. (Accessed 1 July 2022).
- Pekrun, R., 2006. The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18 (4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9> [online]. Available at. (Accessed 5 July 2020).
- Pekrun, R., Goetz, T., Titz, W., Perry, R.P., 2002. Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ. Psychol.* 37 (2), 91–105. https://doi.org/10.1207/S15326985EP3702_4 [online]. Available at. (Accessed 5 July 2020).
- Perrenoud, P., 1998. From formative evaluation to a controlled regulation of learning processes: towards a wider conceptual field. *Assess. Educ. Princ. Pol. Pract.* 5 (1), 85–102. <https://doi.org/10.1080/0969595980050105> [online]. Available at. (Accessed 5 July 2020).
- Pintrich, P., 2000. The role of goal orientation in self-regulated learning. In: Boekaerts, M., Pintrich, P., Zeidner, M. (Eds.), *Handbook of Self-regulation*. Academic Press, San Diego, CA, pp. 451–502.
- Pintrich, P.R., Zusho, A., 2002. The development of academic self-regulation: the role of cognitive and motivational factors. In: Wigfield, A., Eccles, J.S. (Eds.), *Development of Achievement Motivation*. Academic Press, San Diego, CA, pp. 249–284.
- Sadler, D.R., 1998. "Formative assessment: revisiting the territory", assessment in education: principles. *Pol. Pract.* 5 (1). <https://doi.org/10.1080/0969595980050104> [online]. Available at. (Accessed 5 July 2020).
- Sadler, D.R., 2010. Beyond feedback: developing student capability in complex appraisal. *Assess. Eval. High. Educ.* 35 (5). <https://doi.org/10.1080/02602930903541015> [online]. Available at. (Accessed 5 July 2020).
- Schunk, D.H., Green, J.A., 2018. Historical, contemporary, and future perspectives on self-regulated learning and performance. In: Schunk, D.H., Green, J.A. (Eds.), *Handbook of Self-regulation of Learning and Performance*, second ed. Routledge, pp. 1–15.
- Shute, V., 2008. Focus on formative feedback. *Rev. Educ. Res.* 78 (1), 153–189.
- Smith, K., 2015. Assessment for learning: a pedagogical tool. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, vol. 2. Sage, London, pp. 740–755.
- Smith, K., Gamlem, S., Sandal, A.K., Engelsen, K.S., 2016. Educating for the future: a conceptual framework of responsive pedagogy. *Cogent Educ.* 3 (1). <https://doi.org/10.1080/2331186X.2016.1227021> [online]. Available at. (Accessed 5 July 2020).
- van der Kleij, F.M., Adie, L., 2020. Towards effective feedback: an investigation of teachers' and students' perceptions of oral feedback in classroom practice. *Assess. Educ. Princ. Pol. Pract.* 27 (3), 252–270. <https://doi.org/10.1080/0969594X.2020.1748871> [online]. Available at. (Accessed 5 July 2020).
- van der Kleij, F.M., Adie, L.E., Cumming, J.J., 2019. A meta-review of the student role in feedback. *Int. J. Educ. Res.* 98, 303–323. <https://doi.org/10.1016/j.ijer.2019.09.005> [online]. Available at. (Accessed 5 July 2020).
- Vattøy, K.-D., 2020. Teachers' beliefs about feedback practice as related to student self-regulation, self-efficacy, and language skills in teaching English as a foreign language. *Stud. Educ. Eval.* 64. <https://doi.org/10.1016/j.stueduc.2019.100828> [online]. Available at. (Accessed 5 July 2020).
- Vattøy, K.-D., Gamlem, S.M., 2020. Teacher–student interactions and feedback in English as a foreign language classrooms. *Camb. J. Educ.* 50 (3). <https://doi.org/10.1080/0305764X.2019.1707512> [online]. Available at. (Accessed 5 July 2020).
- Vattøy, K.-D., Smith, K., 2019. Students' perceptions of teachers' feedback practice in teaching English as a foreign language. *Teach. Teach. Educ.* 85, 260–268. <https://doi.org/10.1016/j.tate.2019.06.024> [online]. Available at. (Accessed 5 July 2020).
- Vattøy, K.-D., Gamlem, S.M., Rogne, W.M., 2020. Examining students' feedback engagement and assessment experiences: a mixed study. *Stud. High. Educ.* <https://doi.org/10.1080/03075079.2020.1723523> [online]. Available at. (Accessed 5 July 2020).
- Vygotsky, L.S., 1962. *Thought and Language*. The Massachusetts Institute of Technology Press, Cambridge, Massachusetts.
- Vygotsky, L., 1978. *Mind in Society. The Development of Higher Psychological Processes*. Harvard University Press, Cambridge.
- William, D., 2020. Foreword. In: Heritage, M., Harrison, C. (Eds.), *The Power of Assessment for Learning: Twenty Years of Research and Practice in UK and US Classrooms*. Corwin, London, pp. vii–ix.
- Winstone, N., Carless, D., 2020. *Designing Effective Feedback Processes in Higher Education*. Routledge, New York.
- Winstone, N.E., Nash, R.A., Parker, M., Rowntree, J., 2017. Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes. *Educ. Psychol.* 52 (1), 17–37. <https://doi.org/10.1080/00461520.2016.1207538> [online]. Available at. (Accessed 5 July 2020).
- Wood, D., Bruner, J., Ross, G., 1976. The role of tutoring in problem solving. *JCPP* 17, 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x> [online]. Available at. (Accessed 5 July 2020).
- Wyatt-Smith, C., Klenowski, V., 2013. Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assess. Educ. Princ. Pol. Pract.* 20 (1). <https://doi.org/10.1080/0969594X.2012.725030> [online]. Available at. (Accessed 5 July 2020).
- Zimmerman, B.J., 2000. Self-efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25 (1). <https://doi.org/10.1006/ceps.1999.1016> [online]. Available at. (Accessed 5 July 2020).
- Zimmerman, B.J., Schunk, D.H., DiBenedetto, M.K., 2017. The role of self-efficacy and related beliefs in self-regulation of learning and performance. In: Elliot, A.J., Dweck, C.S., Yeager, D.S. (Eds.), *Handbook of Competence and Motivation: Theory and Application*, second ed. The Guilford Press, New York, pp. 313–333.

Alternative assessment

Gordon Stobart, University College London Institute of Education, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

G. Stobart, C. Gipps, Alternative Assessment, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 202–208, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00307-9>.

Introduction	96
Level 1: Alternative formats	97
Level 2: Alternative forms of assessment	97
Performance assessment	97
Competency assessment in the workplace	97
Realistic simulation	98
Alternative forms of classroom assessment	98
Extended assignments/projects	98
Assessment by teachers for external purposes	98
Portfolio assessment	99
Theoretical underpinnings	99
Summary	100
Level 3: Alternative uses of assessment	100
What is the purpose of formative assessment?	100
Feedback	100
Diagnostic assessment	101
Dynamic assessment	101
Summary	101
References	102
Further reading	102
Relevant websites	102

Glossary

Dynamic assessment A form of diagnostic assessment which looks at what progress a learner can make when offered adult support in a one-to-one situation

Educational assessment An approach which focuses on an individual achievement relative to him or herself rather than to others, places little emphasis on standardization and seeks to help the individual progress. It is contrasted with psychometric testing

Formative assessment Assessment used to inform teaching and learning so that further learning takes place

International Baccalaureate (IB) A high status international diploma offered in three languages around the world

Performance assessment Judging what students can do in terms of the actual tasks and end performances that are the goals of instruction (Used interchangeably with “authentic assessments” and “direct assessments”).

Psychometric testing Assessment which focuses on standardization of tasks and scoring, on norms and distributions

Realistic simulation Seeking to represent the intended real-world performance, for example, a flight simulator for trainee pilots

Social constructivist Cognitive and sociocultural theories of learning which emphasize the importance of students actively making meaning in their learning and the role of the social context

Summative assessment Assessment which “sums up” where learners have got to in their learning

Systemically valid testing Assessment which leads to curriculum and instructional changes that foster the cognitive skills that the test measures

Introduction

Alternative assessment is an elastic term which stretches from doing the same thing in a different format to describing radically different ways of approaching assessment. These are alternatives to conventional “pencil-and-paper” standardized tests and examinations. Alternative assessments include performance, authentic, formative, and diagnostic assessments. These terms overlap and

reflect alternative approaches to both learning and assessment. A function of this contribution is to clarify, and question, the way the term alternative assessment is used.

Alternative assessments can be classified into three levels: alternative formats of assessment, alternative forms of assessment; and alternative uses of assessment. While there are considerable overlaps among these, the distinctions may be useful in discussing the extent of the differences from conventional assessment. Examples are considered from each of these levels.

Level 1: Alternative formats

The alternative label is usually applied at this level to new or different delivery of an assessment—for example, transferring a pencil-and-paper test to a computer screen. The test content and demands remain much the same while the means of delivery changes. It may be the case that as the computer testing becomes more sophisticated, it begins to make new demands and might move to the next level—for example, using multimodal stimulus materials or forms of computer-adaptive testing.

Where a particular form of assessment is dominant—for example, multiple-choice testing in the USA—then, more open-ended tests requiring constructed answers have also been claimed to be alternative, in the sense that they are different. It may be that, once again, little else changes in the test demands and accompanying pedagogy. We would challenge whether this kind of claim should automatically be called alternative, since it is essentially business-as-usual with varied assessment formats. Little changes in what is assessed and in classroom teaching and learning.

Continuous Assessment by teachers, for example the Grade Point Average (GPA) in the North American education system, may fall into this category. These assessments are less standardized than external tests but often may cover the same curriculum as the end of unit/course test. However, where the classroom teaching and learning is more imaginative and goes beyond this narrowed curriculum and includes, for example, projects, debates and group work it may best be considered as Level 2.

Level 2: Alternative forms of assessment

Hanson has highlighted the “fabricating quality of tests” in which “the likelihood that someone will be able to do something, as determined by the tests, becomes more important than one’s actually doing it” (1994: 288). At this level, changes in assessment are underpinned by the intention to provide a different approach to assessment which focuses on performance of the skills themselves. This can be illustrated through developments in performance and authentic assessments—terms which are used interchangeably here. These assessments emphasize direct, or as direct as possible, demonstration of the skill being assessed and minimize the use of tests as a proxy for actual performance. This is a long-standing tradition in many occupational areas, and there have been moves to apply this within more general education. At a conceptual level these approaches to assessment have been encouraged by a shift toward more constructivist theories of learning. These encourage a more holistic and “sense-making” approach to learning and its assessment in contrast to the “knowledge-in-bits” of most multiple-choice tests which behaviorist views of learning have encouraged.

Along with these conceptual shifts in approaches to learning, the drive to develop alternative models of assessment has received support from the fact that traditional types of test, for example, multiple-choice standardized tests in the US and formal written examinations in the UK, have had unintended and negative effects on teaching and on the curriculum. This is particularly the case where test results play a central role in educational accountability systems, with test results used to evaluate school performance by local authorities and government. The “high stakes” use of tests in countries such as the United States, England and Chile have narrowed the curriculum as teachers “teach the test”. “Under intense political pressure, test scores are likely to go up without a corresponding improvement in student learning. In fact, distortions in what and how students are taught may actually decrease students’ conceptual understanding” (Shepard, 2000: 13).

Performance assessment

Shepard and Bleim (1995) have argued for the equivalence of some of the key terms: “We use the terms performance assessments, authentic assessments and direct assessments interchangeably, the idea being to judge what students can do in terms of the actual tasks and end performances that are the goals of instruction” (p. 25). While this equivalence may not satisfy all, it provides us with a broad base from which to work. The spirit of this approach is to offer assessments which relate as closely as possible to the real-life knowledge and skills required. How close this match is may depend on the context. The philosophy behind this approach is straightforward: to be assessed as having a skill or competency we need to demonstrate it.

Two well-established forms of performance assessment, both of which have their roots in occupational and vocational training, illustrate this emphasis on practical demonstrations of competence. With pencil-and-paper testing so dominant in the education sector, these serve as a reminder of other models of assessment which have a long-established pedigree.

Competency assessment in the workplace

The most direct correspondence between learning a skill is competency assessment in the workplace. We would expect our trainee doctors, plumbers, and pilots to have demonstrated competency in real-life situations. It is the same with musicians and drama

students; they will be assessed on the basis of actual performances. A key feature of these assessments is that they rely on expert human judgment. Competency assessment was embodied in the medieval guild system, in which persons would be apprenticed and would then, after learning their trade for a number of years, demonstrate the skills in their craft through their products. It was the quality of these products that determined their admittance to the guild. This is a distinctly different form of assessment from psychometric approaches which seek to minimize individual judgment and in which competency is often reduced to achieving the cut-score on multiple-choice tests.

Realistic simulation

This approach also seeks to align the assessment closely with the intended real-world performance, but may take place in a context, for example in a college, which relies on simulation. For example, medical training may involve direct simulation, using actors who mimic patients' symptoms. Similarly, trainee pilots will use flight simulators and trainee mechanics work on workshop engines. Another form of realistic simulation may be working under highly controlled and supervised conditions such as special catering events for trainee chefs, or conducting scientific experiments in school science laboratories. We consider these alternative assessments because they represent a model of learning-by-doing which differs from the dominant educational model in which tests become a proxy for the required skills and practical knowledge.

While we categorized early computer-based testing as little more than changing the format of the tests (paper to screen), many of the more sophisticated applications are able to offer realistic simulations through multimedia stimulus materials. We also recognize that adaptive testing, which responds to the student's choices and decisions, offers an alternative when, for example, it works to capture the student's own understanding of a problem.

Alternative forms of classroom assessment

The classroom rarely offers such clear-cut alternative assessment practices. There are, however, opportunities to move away from the dominant testing model and in doing so offer a richer curriculum and more in-depth learning. Some of these may be loosely described as performance assessments since they focus on what learners can demonstrate in less constrained settings than conventional standardized tests. These are often summative in purpose, though they contain many opportunities for formative assessment (see below).

Extended assignments/projects

In standardized tests, even the inclusion of constructed (open-ended) answers offer the candidates little room to express their own understandings since they generally require convergent answers which are anticipated by the mark scheme. An alternative approach, which sits well with constructivist principles, is to provide the learner with opportunities to create a substantial piece of work in which they have choice over what they do. At classroom level, this may be projects which explore local issues—for example, history projects, or which feature the students' own designs in technology. This may also incorporate group activities which would encourage the collaborative aspects of learning.

A useful example of how individual projects can be managed is provided by the extended essay in the International Baccalaureate (IB)—a high-status international diploma for 16–18-year-olds offered in three languages around the world. The extended essay is an independent, self-directed piece of research, culminating in a 4000-word paper which contributes to the diploma. It involves an in-depth study of a topic of interest within a chosen subject, with the emphasis placed on the research process: formulating an appropriate research question; engaging in a personal exploration of the topic; communicating ideas; and developing an argument. The assessment of this process is by the students' own teachers, who have received professional development in marking the essays. There is considerable external confidence in the dependability of IB assessments, especially among universities for whom the extended essay is a good preparation for similar work at degree level.

Assessment by teachers for external purposes

Teachers regularly assess their students' classroom learning for both formative and summative purposes. The interest here is in those assessments which are used as part of external certification or for high-stakes selection purposes. Such assessments can be considered as alternative assessments when they both replace conventional standardized testing and encourage a broader approach to teaching and learning.

Much classroom assessment may not meet these requirements since it mimics standardized tests and is little more than frequent summative testing which is used for reporting or management purposes rather than as formatively contributing to students' further learning. Continuous assessments such as the grade point average (GPA) in US classrooms can be seen as a means of keeping learning and effort on track—a form of classroom control, rather than encouraging the principles of educational assessment.

A useful model for this broader approach to teaching and learning which encourages deeper understanding is [Frederiksen and Collins' \(1989\)](#) principles of systemically valid testing. At the heart of this approach is the principle that:

A systemically valid test is one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure. Evidence for systemic validity would be an improvement in those skills after the test has been in place within the education system for a period of time (p. 27).

They proposed standards for assessment which included:

Directness. This involves assessing the cognitive skill of interest—the emphasis being on the authenticity of the extended tasks. They call for sets of tasks which are ecologically valid, which are representative of the way knowledge and skills are used in real-world contexts, and for which the primary traits have been identified.

Scope. This considers the range of skills needed to do well in the tasks.

Reliability. This seeks effective ways of assessing which, at the same time, fosters learning.

Transparency. This is the concern that those being assessed are clear about how they are being judged. This should enable learners to “assess themselves and others with almost the same reliability as the actual test evaluators achieve” (p. 30).

This theoretical framework can be applied to a variety of alternative assessments such as assessment by teachers which contribute to credentials as well as extended assignments and portfolio-based assessment.

An example of systemically valid assessment by teachers is that found in Queensland, Australia, which abolished examinations over 30 years ago and replaced them with criteria-based subject-achievement assessments by teachers. This offers an alternative to selection examinations and encouraging a richer curriculum in which teachers and students are actively involved in the assessments. The system involves extensive moderation by over 400 district review panels, composed mainly of teachers. These panels review samples of work across subjects and schools and provide advice to schools. The process is seen as an effective form of professional development in assessment. However, increased accountability pressures within the Australian system have recently obliged Queensland to introduce some external testing in addition to the Queensland Core Skills Test (cognitive skills based on the common curriculum elements) which is taken by 17-year-olds wishing to enter university.

Portfolio assessment

At one level, portfolios may be regarded simply as collections of pieces of student work which are kept as a record. However, these collections can also be used as a basis for summative assessment, and so portfolios are often described as a method of assessment.

A portfolio used for assessment and learning purposes involves documentation of achievements, self-evaluation, process artifacts, and analyses of learning experiences. It is, therefore, significantly more than a collection of assignments (Klenowski, 2000: 219).

Using collections of student work for assessment allows a wide range of approaches to marking—from traditional marks and grades assigned by teachers, through detailed written comments by teachers (with no grades), teacher–student discussion, and negotiation of the assessment, to student self-evaluation. As with other forms of alternative assessment, the proponents of portfolio assessment present a rationale based on the enhancement of learning. First, assessment that celebrates achievement on a self-referenced or criterion-referenced basis builds students’ confidence and ultimately enhances their performance. These can be contrasted with the situation in which assessment draws attention to failure through the use of grades and marks, particularly when norm-referenced. Second, learning is enhanced if students understand clearly what they are trying to achieve and are provided with personal and substantive feedback in trying to reach their target. Third, collecting, sorting, and annotating evidence to demonstrate achievement in the form of a portfolio of work is a valuable aid in developing the skills of self-evaluation, as well as for communicating standards in a relatively unambiguous and accessible way.

Advocates see portfolio assessment as the embodiment of good social constructivist principles. While the validity of using these for summative purposes at a local level is strong, there have been far more challenges when portfolios have been used in relation to wide-scale accountability testing. Portfolio experiments in the USA, in the states of Vermont and Kentucky, led them to be incorporated into accountability measures, which then raised questions of reliability. For high-stakes purposes, portfolio assessments remain vulnerable to reliability problems relating to sampling, standardization, and scoring. In high-stakes assessment contexts, this will often leave them exposed to criticisms of bias and unreliability, particularly where it is the student’s own teacher who is making the assessment.

Theoretical underpinnings

The philosophy behind these alternative approaches distinguishes educational assessment from psychometric assessment. For present purposes, the key issues are that, in the psychometric model, the focus on reliability requires the standardization of administration and tasks as well as of scoring. The preoccupation of psychometrics with norms and norm-referenced testing stemmed from tests focused on aptitude, selection, and prediction. An individual’s performance was reported in relation to that of his/her peers, so that performance was seen in relative, rather than in absolute, terms. A recent debate (William and Baird, 2017) has questioned in what way psychometric approaches actually contribute to learning.

Alternative assessments can be seen as theoretically grounded in those learning theories which emphasize the importance of individuals constructing meaning for themselves within a social context and the importance of broader-based assessment.

Gipps (2011), for example, has argued that educational, as opposed to psychometric, assessment deals with the individual’s achievement relative to him or herself rather than to others; takes place in relatively uncontrolled conditions and so does not produce well-behaved data; looks for best—rather than typical—performances; is most effective when rules and regulations characteristic of standardized testing are relaxed; and embodies a constructive outlook on assessment, where the aim is to help rather than sentence the individual.

Shepard (2000) developed a conceptual framework for this understanding of assessment based on social constructivist theories of learning. She summarizes this view of learning in the following principles: (1) intellectual abilities are socially and culturally

developed; (2) learners construct knowledge and understandings within a social context; (3) new learning is shaped by prior knowledge and cultural perspectives; (4) intelligent thought involves metacognition or self-monitoring of learning and thinking; (5) deep understanding is principled and supports transfer; and (6) cognitive performance depends on dispositions and personal identity.

Summary

This section has reviewed examples of alternative assessments which seek to directly measure the desired skills and competencies rather than through the proxy of written tests. This is an approach historically grounded in occupational and vocational approaches to learning represented by the guild system. In educational settings, alternative approaches to assessment are based on different models of learning. Social constructivist approaches encourage students to make meaning and to understand the principles behind what they are doing. This leads to forms of assessment such as portfolios and extended project work. Classroom assessment by teachers will need to be interrogated to see if it is offering such opportunities.

Level 3: Alternative uses of assessment

The assessments considered so far have been primarily used for summative purposes, summing up where students have got in their learning. This section considers an alternative uses of assessment: the formative and the diagnostic. Formative assessment serves different purposes since its main intention is to assist the learning process rather than to measure what has been learned. Diagnostic assessment can be seen as a subset of this in seeking to find out what is known in order to help with further learning. There are considerable overlaps with the previous level: summative assessments can be used formatively if the results are used to guide further learning; performance and educational assessments will often involve the kind of self-regulation and feedback which characterize formative assessment.

What is the purpose of formative assessment?

The intention of formative assessment is to directly inform learning and teaching processes so that further learning takes place. This can be contrasted with summative assessment, in which the intention is to provide a summing up of where learners have got to in their learning. At the heart of formative assessment is the assumption that assessment, broadly understood as information gathering, can help identify what learners know, so that further teaching can build on this. It also assumes that learners will play an active part in the process and will increasingly be able to monitor and regulate their own learning. A key mechanism in this process is feedback, since this can help the learner bridge the gap between current and desired performance. My preference is to call this approach *assessment for learning* as this makes the intention clear. A useful definition is:

The process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go to and how best to get there.

Assessment Reform Group (2002: 2–3).

This is essentially about day-to-day classroom assessment and depends on the quality of classroom interactions. It supports a pedagogy in which teachers seek to negotiate, and make explicit, learning intentions and success criteria (“where they need to go to”). It also involves finding out where learners are in their learning. This may be by diagnostic assessment—for example, miscue analysis in reading or investigating errors on tests. On a day-to-day basis, it occurs through classroom interaction—for example, questioning and dialog.

One practice which illustrates how assessment for learning can modify approaches to teaching and learning is that of “wait-time”. When teachers ask questions, they expect an almost immediate reply, the research suggests less than a second’s wait (Rowe, 1987). If effective learning is about the student-constructing meaning, what chance is there for this when an immediate answer is expected? Also, what kind of question only takes a second to process before answering? Wait-time, therefore, encourages the teacher to ask richer questions and allow students more time to process the answers (and there maybe more than one correct answer). This approach also encourages students to discuss briefly the answer with each other before the teacher asks for responses. This may also lead to following up on unexpected or incorrect answers since these misconceptions may help in understanding where learners are in their learning.

Wait-time is just one small example of how formative assessment may change the quality of classroom interaction. This is no longer the teacher using surface-level questions to check quickly whether students know the correct answers, it involves a more deliberate attempt to gauge their understanding.

Feedback

One of the key components of assessment for learning is feedback. Its role is to help move learners from where they are toward the desired standard—the “and how to get there” of the definition. What is known about feedback is that it does not always work as intended. In their meta-analysis of research on feedback, the psychologists Kluger and DeNisi (1996: 275–277) concluded that:

In over one third of cases Feedback Interventions reduced performance ... we believe that researchers and practitioners alike confuse their feelings that feedback is desirable with the question of whether Feedback Intervention benefits performance.

Whether feedback has positive learning effects depends on many interacting factors: motivation, the complexity of the task, the expertise of the learner, and the level and quality of the feedback. This makes it highly situational: the same feedback given to two learners could have opposite effects. Simply telling novice learners that they are wrong may set back learning, telling engaged experts the same may be enough to get them to increase their efforts and change their strategy.

Assessment for learning encourages feedback which focuses on the task rather than on the learner, and which has the goal of learners who become increasingly self-regulating. This involves being sufficiently aware of the desired standard that they can begin to give themselves, and their peers, feedback. This leads to alternative thinking about classroom practices, for example, questioning whether praise is an effective form of feedback (how does praise provide information to close the gap in learning?). It may also question routine marking and grading practices. For example, if the feedback to a piece of work is a mark (7/10) how does that provide information to move learning forward? Most teachers may put a comment alongside this—but research suggests that this is largely ignored, it is the mark that counts.

This leads to classroom practice with an emphasis on developing self-regulated learners who receive and use informative feedback. It leads to less emphasis on marks and grades, and more encouragement for comment-only marking.

The claim here is that formative assessment represents an alternative purpose for assessment: to contribute directly to the learning process. In terms of the approach to learning, it overlaps with the social constructivist views that we considered in the previous section and aligns well with approaches such as performance assessment and portfolios.

Diagnostic assessment

A further example of different purposes of assessment is its use for diagnosis. Genuine diagnostic testing can be seen as a subset of formative assessment (where the learners are in their learning). We distinguish this from diagnostic testing which is used for managerial purposes—for example, to allocate students to groups. Diagnostic testing can involve informal analysis—for example, of a pupil's reading skills and limitations—or the use of test results to identify what has not been understood. This test-and-remediate approach would often be called formative assessment in the US. In this, the formative element is in the remediation work following a test, occupying, for example, 1 week in a 6-week module. This approach to formative assessment differs from assessment for learning, which treats it as a continuous everyday classroom practice.

Dynamic assessment

An alternative approach to diagnostic assessment is provided by dynamic assessment. This has drawn on the sociocultural theories of Vygotsky (1978), who pointed to the importance of tools and aids in learning. What this diagnostic approach seeks to find out is what progress a learner can make when offered adult support in a one-to-one situation. It is therefore not so much what learners know, but how effectively they learn when assisted. In the traditional examination and test, the student is denied the use of external tools. As Lunt (1994) explains:

Dynamic assessment procedures ... involve a dynamic interactional exploration of a learner's learning and thinking process and aim to investigate a learner's strategies for learning and ways in which these may be extended or enhanced. Since it offers individuals an opportunity to learn, dynamic assessment has the potential to show important information about individual strategies and processes of learning and, therefore, to offer potentially useful suggestions about teaching (p. 152).

This approach—like formative assessment and educational assessment—represents a challenge to approaches to assessment found in the psychometric tradition. They are alternative not because they do the same thing differently but because they attempt to achieve something different. This difference is rooted in current understandings of how we learn and in a search for ways of aligning assessment with approaches to teaching and learning which emphasize the importance of using assessment to support good-quality learning. This requires a different relationship between assessor and learner, with learners playing an active part in their learning.

Summary

Alternative Assessment is an umbrella term which shelters a range of approaches to assessment which are outside the conventional formats of standardized tests. These have been grouped into three broad, and overlapping, categories which progressively indicate their distance from the typical psychometric model of test development. The first level involves alternative *formats*, in which test content is presented in alternative ways, for example by computer rather than on paper. The second level sees a move toward different assessment tasks or different ways of presenting evidence. Performance assessments involving demonstration of a skill

either directly (e.g., musical performance) or through simulation (e.g., pilot training) would represent this level. Alternative methods of assessing students' work through portfolios, extended essays or coursework are also part of this level. The third level involves alternative *purposes* for the assessments. While conventional tests are typically used for summative purposes, "summing up" where a student is at that moment, the purpose of formative assessment ("assessment for learning") is to feed directly, and more informally, into the learning process.

This three-level classification is intended to bring into clearer focus how assessments can be considered as "alternative".

References

- Assessment Reform Group, 2002. *Assessment for Learning: 10 Principles*. University of Cambridge, Assessment Reform Group, UK.
- Fredericksen, J.R., Collins, A., 1989. A systems approach to educational testing. *Educ. Res.* 18 (9), 27–32.
- Gipps, C., 2011. *Beyond Testing: Towards a Theory of Educational Assessment*. Falmer, London.
- Klenowski, V., 2000. Portfolios: promoting teaching. *Assess Educ. Princ. Pol. Pract.* 7, 215–234.
- Kluger, A.V., DeNisi, A., 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.* 119, 252–284.
- Lunt, I., 1994. The practice of assessment. In: Daniels, H. (Ed.), *Charting the Agenda*. Routledge, New York, pp. 145–170.
- Rowe, M.B., 1987. Wait time: slowing down may be a way of speeding up. *J. Teach. Educ.* 1, 37–43.
- Shepard, L.A., 2000. The role of classroom Assessment in teaching and learning. In: Richardson, V. (Ed.), *Handbook of Research on Teaching*. American Educational Research Association, Washington, DC.
- Shepard, L.A., Bleim, C.L., 1995. Parents' thinking about standardized tests and performance assessments. *Educ. Res.* 24 (8), 25–32.
- Vygotsky, L., 1978. *Mind in Society*. Harvard University Press, Cambridge, MA.
- William, D., Baird, J. (Eds.), 2017. *Assessment and Learning, Assessment in Education, Principles, Policy and Practice*, vol. 24, p. 3.

Further reading

- Bennett, R., 1998. *Reinventing Assessment*. Educational Testing, Princeton, NJ.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5, 7–74.
- Gipps, C., 1999. Socio-cultural aspects of assessment. *Rev. Res. Educ.* 24, 357–394.
- Gipps, C., Stobart, G., 2003. Alternative assessment. In: Kellaghan, T., Stufflebeam, D.L. (Eds.), *International Handbook of Educational Evaluation*. Kluwer, Norwell, MA, pp. 549–576.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Koretz, D., 1998. Large-scale portfolio assessments in the US: evidence pertaining to the quality of measurement. *Assess Educ. Princ. Pol. Pract.* 5, 309–334.
- Sadler, R., 1989. Formative assessment and the design of instructional systems. *Instr. Sci.* 18, 119–144.
- Shepard, L.A., 2000. The role of assessment in a learning culture. *Educ. Res.* 29 (7), 4–14.
- William, D., 2011. *Embedded Formative Assessment*. Bloomington. Solution Tree Press.

Relevant websites

- www.aiaa.org.uk/.../category/assessment-reform-group, Assessment Reform Group.
- <http://www.cse.ucla.edu>, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Educational assessment – trends in research, policy and practice

Christian Ydesen, Department of Culture and Learning, Aalborg University, Aalborg, Denmark

© 2023 Elsevier Ltd. All rights reserved.

Introduction – unraveling assessment as a phenomenon in education	103
Scholarly debates in educational assessment	104
Research approaches to educational assessment	104
Concluding discussion: implications in terms of policy and practice	105
References	106

Introduction – unraveling assessment as a phenomenon in education

Assessment is an inherent component of education. Ever since the dawn of educational activities, educators – and various stakeholders – have taken an interest in knowing whether students have actually learnt what they were supposed to have learnt (Broadfoot, 2007; Wiliam, 2010; Stobart, 2008). Some historical examples of education assessment practices are the Chinese civil servant examination instituted during the Qin Dynasty around 200 BC (Wang, 2013) and Roman students' cramming for the Law of the Twelve Tables from 450 BC in the Late Roman Republic (Cornell, 2012).

But assessment in education also has a long history of teachers and schools being assessed and held to account for their performance (Dorn and Ydesen, 2014). Over the last 60 years we have seen a move from indicators of input (curriculum) to indicators of output (achievement) and various versions of standalone schools, directly responsible for their own progress. This development is reflected in a shift in emphasis from streaming of students according to assessment of their abilities to the assessment of schools based on student and teacher data. In our own time, we have witnessed the global dissemination of an assessment agenda which places students' learning outcomes and educational standards at the forefront of policymakers' agendas (Hill and Kumar, 2009; Smith, 2016).

To the extent that we can discern future trajectories, we begin to see new forms of educational assessment taking shape, where real-time monitoring and artificial intelligence induced recommendations based on cross-references of performance scores and data about social background and biological capacities in individuals and cohorts are gaining ground (Williamson, 2018). In this sense, it is possible to speak of a development in educational assessment from the measurement of past progress (diagnostic assessment) into governing in relation to future projections of progress (prognostic assessment).¹

The history of educational assessment is however highly irregular (McArthur, 1987). This irregularity arises because assessment is not an unequivocal phenomenon. While educational assessment may be broadly defined as the judgment of performance, it has multiple inseparable and closely interwoven layers arising in relation to different stakeholders, instruments and purposes (Richardson, 2022; Stobart, 2008; Madaus et al., 2009). For instance, assessment may be used as a pedagogical, evaluation or governance tool, namely for the implementation or monitoring of policy and/or a device of accountability (Newton, 2007). While these uses are not necessarily mutually exclusive, they do add to the complexity surrounding educational assessment where various agendas, priorities, and ideals crisscross any development and enactment of educational assessment policies and practices.

In a broader sense, it can be argued that all societies to a certain extent are dependent on some form of assessment system to establish criteria of human worth in terms of knowledge, skills, or competencies. In his grand, classic survey *Prestige, Class, and Mobility*, sociologist Kaare Svalastoga (1914–97) stated that achievement factors rooted in the passing of some formal or informal assessment are predominant among a society's status determinants (Svalastoga, 1959, see also Weber, 2009). Through this lens, educational assessment is closely associated with the societal, and bureaucratic, division of labor as a provider of formal certifications, but also serves certain ideals and ambitions within both society and individuals, such as the strive for meritocracy and status.

Both these bureaucratic and idealistic layers of educational assessment interact with an economic approach to education whereby education is seen as a key production factor in the pursuit of economic growth (Spring, 2015; Ydesen and Bomholt, 2020). In the language of human capital theory, education assessment systems act as gatekeepers to a hierarchy of education qualifications and thereafter training and employment opportunities. Credentialism holds increasing value in the labor market, i.e. the strive for elite status in a market of professions (Collins, 1990). The idea is that education in general and credentialism in particular will lead to increasing productivity benefitting both the individual and society (Brown et al., 2020).

At the same time, however, the same bureaucratic and idealistic layers also intertwine with citizenship education where ideals about democracy, participation, human growth and enlightenment take pride of place. Viewed in a historical perspective, the building of education systems in general, and assessment systems in particular, has been closely wedded to constructions of

¹It should be duly mentioned that the architects of IQ testing in the interwar years worked from the assumption that IQ testing would be an objective scientific tool for engaging in social engineering activities with the promise of optimizing society and putting everyone in their "right place" (Chapmann, 1988; Ydesen, 2011). In this sense, there was a distinct prognostic component associated with IQ testing.

nationhood, states and ultimately nation-states through the fabrication of the model national citizen (Tröhler et al., 2011; Tröhler, 2020; Popkewitz and Lindblad, 2004).

Following an exploration of some of the most pregnant scholarly debates in education assessment, it is the purpose of this chapter to unpack some of these complex layers, and their interwovenness, while unraveling the assumptions, preconditions as well as the possibilities and limitations of the possible research paths for exploring the phenomenon of education assessment. Finally, the chapter outlines some of the implications for policy and practice in a concluding discussion.

Scholarly debates in educational assessment

Educational assessment is a research topic which sometimes creates polarizations and acrimonious debates in the scholarly field (Baird and Elliott, 2018; Garsdal and Ydesen, 2009). The contentious nature of researching educational assessment can be explained as a result of at least three interconnected factors which have to do with the practical enactment of assessment and assessment research.

The first is that a wide array of research fields are engaged with research on educational assessment from psychometrics, psychology and sociology to history and political science. These fields often subscribe to different philosophies of science, and they have diverse traditions in terms of methodology, theory and even the purpose of research. These research fields often also publish their research in different journals which means that much research on education assessment “runs” in parallel paradigmatic strands which tend to reproduce the polarization.

The second factor is that something is generally “at stake” in educational assessment. Different stakeholders – for instance, students, parents, teachers, school leaders, local authorities, national authorities, edu-businesses and global actors – have different interests and priorities when it comes to educational assessment and that has an impact on the field in terms of *which research* is used for *which purposes* and, even, sometimes, which types of studies get funded.

The third factor relates to the different understandings of the nature and purpose of education and therefore educational research (Lagemann, 2009). According to Biesta (2015), there is a division between education researchers “who approach education first and foremost in technological terms – that is, as an activity or practice that is ultimately understood as being governed by cause–effect relationships (...) – and those who see education first and foremost as a human event of communication, meaning making and interpretation in which questions of cause and effect actually have no place.” (p. 12; see also Lagemann, 2009). To a large extent we can recognize the economic – i.e. an applied – approach versus the citizenship – i.e. a *Bildung* – approach to education mentioned above in Biesta’s distinction.

On a more philosophical level, the contentious nature of research on educational assessment surfaces in three key issues. The first is whether educational assessment can be viewed as an isolated phenomenon or whether it must be viewed in connection with other aspects of education (Baird et al., 2017; Hegarty, 2020). This issue is very clear if we look at research critical to assessment practices where a clear line between endogenous and exogenous critiques is visible. Endogenous critiques claim that problems with education assessment can be remedied if better methodological tools and procedures are used. This kind of critique often comes from within the ranks of psychometrics and psychology (e.g. Allerup and Kjeldsen, 2017). Exogenous critiques on the other hand argue that education assessment tools and practices must be viewed in their context and that they are essentially flawed, skewed and beyond repair (e.g. Au, 2009). This kind of critique often comes from sociology and various sub-disciplines of history (e.g. Danziger, 1990; Simon, 1953).

Another related issue is whether educational assessment is a neutral or normative undertaking. In other words, whether assessment tools provide a neutral and objective snapshot of a student’s proficiency level or whether assessment tools are essentially filled with values and normativity by default. For instance, certain studies point to the existence of class assumptions about merit and ability (e.g. Stobart, 2008) whereas others highlight the objectively beneficial contributions of assessment to education in a wider sense (e.g. Phelps and Gottfredson, 2009; Roediger et al., 2011).

A third issue pertains to the purposes of assessment. Here it is commonplace to find conceptual distinctions between summative and formative types of assessment whereby the former is defined as end-of-process evaluations and the latter as in-process evaluations. Newton (2007) has problematized the formative versus summative distinction arguing that there are many distinct assessment purposes and not just two broad categories. Thus, this dichotomous vocabulary runs the risk of blurring and reducing our understanding of how assessment actually works in practice.

Given these issues and the contentious nature of educational assessment, it is possible to use a key issue as an illustrative case for teasing out the assumptions and preconditions associated with research approaches in educational assessment which serves as a foundation for understanding implications for policy and practice in the final section of the chapter. It is the purpose of the next section to explore how different research approaches have dealt with the question of how to balance human diversity and standardization within an educational assessment system.

Research approaches to educational assessment

It is often said that research findings and knowledge creation depend on the questions, assumptions and approach taken at the outset to the study. Research in educational assessment is no exception. As already implied above, it is possible to discern two

overall different methodological paradigms in educational assessment research: (i) an empirical positivist approach based on findings drawn from measurement data and (ii) a sociologically oriented approach incorporating the human, cultural, historical, conceptual, and technical dimensions of assessment (Andreasen et al., 2013).

The “positivist” paradigm is often guided by an applied, “what-works”, and all-other-things-being-equal approach to research in educational assessment. The research questions occupying this paradigm are typically concerned with how assessment data can provide insights into how students learn, how learning processes can be optimized, and what works for teachers. It often pursues the ideal of educational contexts as laboratories from which data can be extracted and improve our practices and policies. Some of the fields most keenly engaged in this paradigm are education economics, psychometrics, behaviourist psychology, and quantitative-statistical parts of evaluation research and political science.

The “sociological” paradigm is generally guided by an ambition to contextually decipher the meanings, workings and implications of educational assessment and sometimes even to de-naturalize the natural in order to gain new insights. It is in this tradition that we find a number of research agendas, such as critical policy analyses, and the ambition to unravel power compositions in education contexts. Some typical research questions pursued in this paradigm explore constructions of normality and deviance in education, the reproduction of inequalities, and the cultural and/or material factors affecting assessment performance and practice. Some of the fields most keenly engaged in this paradigm are critical sociology of education, critical policy research, science and technology studies (STS), new materialism and international political economy.

Turning to attempts to balance human diversity in the classroom with standardization of assessment practices, it could be claimed that the challenge lies in the fact that assessment technologies are often used to categorize, rank and standardize students in relation to predetermined categories and/or learning goals. Education assessment practices, to a very large extent, determine which students are actually “good enough” to contribute in “the right way” to educational activities and demonstrate “visible” learning progression. On the other hand, human diversity means that idiosyncrasies, individual preferences, interests, talents, attributes and abilities make it difficult to standardize the educational performance of individual human beings. Some students will thrive in certain learning environments and excel in certain forms of assessment, whereas others will not. As emphasized in UNESCO’s Global Education Monitoring report 2020 on inclusive education, all students are potentially at risk of education exclusion, depending on the context and content of the learning environment (UNESCO, 2020).

In the “positivist” paradigm, the ambition and the belief are that the human diversity-standardization issue can be overcome by means of an ontology of objective domains such as particular skills, attributes and abilities. It means that assessment technologies generate an extrapolation of a student’s constructs – e.g. intelligence, competence, ability, functional literacy – in relation to the domain being assessed. As pointed out by Madaus et al. (2009), the domains for many high-stakes tests “are defined by the standards or curriculum frameworks established by state departments of education.” (p. 39). Since domains are fixed targets, it makes sense to assess students in relation to these domains. It requires the construction of valid and reliable assessment tools. When these tools have been developed, the assessment results will reflect student diversity and the assessment results yield useful data for the construction of hypotheses and interventions that may help to alleviate politically defined problems in education like closing the attainment gap and optimizing learning environments. In this paradigm, assessment technologies are part of the solution.

In the “sociological” paradigm, domains, constructs and student abilities are moving targets and, as I have already implied, the focus is fundamentally different. Assessment technologies are seen as having the potential to be a significant co-constructor of differences between pupils, and thus also for the construction of the way in which these differences are understood by both teachers and other actors in education, such as parents, headteachers, etc. (Feniger et al., 2016; Kelly et al., 2018). These are the so-called backwash effects (Madaus, 1988). From an STS perspective, Addey and Gorur (2020) argue that the preparation and implementation of assessment technologies often rests on a significant reduction of complexity. They argue that the process constantly considers how reality can be reconciled with the assessment design. This creates a design space containing certain technologies and tools, but also political ideas, understandings and priorities, which cannot in any way be described as objective and which affect understandings of normality and deviation, as well as good and bad results. The “sociological” paradigm often demonstrates how educational assessments can lead to negative practices, such as selective admission, off-rolling, strict discipline policies, student reassignment and greater focus and time given to those most likely to succeed. Through this lens, assessment technologies are often seen as part of the problem.

Concluding discussion: implications in terms of policy and practice

It hardly comes as a surprise that the policy and practice implications associated with these two paradigms are very different. While one paradigm seems to suggest that data, numbers, and results constitute a possible and feasible way to run and develop an education system, the other paradigm seems to promote a much more contextual understanding. Here structural explanations and the unraveling of suboptimalities, like counterproductive mixes of performance indicators, and contra-finalities, like plagiarism, off-rolling and practices of gaming the system, take center stage (Baird and Elliott, 2018). It even adds to the complexity that policy and practice also must take into account the layers of educational assessment arising in relation to different stakeholders, instruments and purposes mentioned above. There are no quick fixes in educational assessment.

In order to provide policymakers and practitioners with an appropriate level of agency allowing them to create good and relevant solutions in the context in which they operate, it is essential that they are aware of the assumptions, outlooks, limitations and explanatory power provided by different types of assessment research. No research is value free and devoid of normativity even

though it may be cloaked as such. Stakeholders need to be empowered to judge whether the policies and practices which develop from various types of assessment research are relevant in the given context and under which conditions the research findings are appropriate. At the training level this requires a certain level of data and assessment literacy among stakeholders and practitioners, but it also requires awareness of the mechanisms, agendas and interests promoting different types of assessment research and assessment data. It is essentially about the politics of policy and practice (Ozga, 2020). The development, implementation and enactment of educational assessment never takes place in splendid isolation and education research in a wider sense has a big task in creating awareness among stakeholders.

In his musings on policy implications, Newton (2007) suggests that the design of an educational assessment system must lay bare and make transparent to all stakeholders those purposes for which assessment results are fit and those for which they are unfit. A good assessment technology in one area may be harmful in another area. Newton argues that “stakeholders should be deprived of ignorance as an excuse for misuse.” (p. 168). Another useful approach to working with assessment implications in terms of policy and practice comes from Stobart (2008) who argues that the three best questions to ask of any educational assessment are (p. 14):

1. What is the principal purpose of this assessment?
2. Is the form of the assessment fit-for-purpose?
3. Does it achieve its purpose?

On a deeper level we might even ask how educational assessment came to be seen as a “problem” for governments and other experts to address. Asking this sort of question connects with the question of how it is possible to establish fairness and credibility in an assessment system (Nisbet and Shaw, 2019). The issue here is to balance ideals about assessment technologies being “free from bias, and ideally fit all test-takers, independent of their gender, race or socio-economic background” (Hopfenbeck, 2019, 537) and trustworthy so that they are able to fill the function that society expects of them. This question cuts across the two paradigms. But while the “positivist” paradigm tends to emphasize how assessment systems can offer fairness and credibility – typically from a technical-methodological angle where validity, reliability and transparency are the core ideals –, the “sociological” paradigm tends to focus on the deprivations of assessment by focusing on the hidden curriculum, the reproduction of social inequalities, governance issues and the role of political interest.

In an attempt to strike a balance between these perspectives, UNESCO emphasizes how “Assessment should be formative and allow students to demonstrate learning in a variety of ways.” (UNESCO, 2020, p. 21). It means that it must be possible to identify and recognize all students’ progress and achievement in order for all students to be able to demonstrate what they can do without any exclusionary consequences.

Nevertheless, it is important to take into account the inertia of educational assessment systems. The logics and styles of reasoning surrounding and driving educational assessment policies and practices tend to reproduce themselves. For instance, in 2013, following critiques of the OECD Program for International Student Assessment (PISA) by Professors of Statistics Svend Kreiner and Karl Bang Christensen (Kreiner and Christensen, 2014), the Danish Minister of Education, Christine Antorini, stated that international ranking of education systems based on PISA-data was meaningless. But this important reservation had already been forgotten by the time of the next PISA-round. The inertia may be explained by a “fear of falling behind” among PISA participating nations (Krejsler, 2019). It is a strategic development narrative about world order dictating that nation-states will lose the global competition between knowledge economies, if they fail to optimize their human capital, i.e., produce “employable” or “career-ready” adults who participate in the economy (see also Grek, 2009).

Another explanation comes from an increasing trend among policy makers to intervene in education practices in the interests of policy objectives such as efficiency, reform, modernization, globalization and competition. Pasi Sahlberg has described this development as the Global Education Reform Movement (GERM). GERM is an education reform approach that broadly follows the tenets of postwar neoliberalism and especially 1980s New Public Management (Fuller and Stevenson, 2019, see also Hill and Kumar, 2009; Verger et al., 2018). It is structured around a common set of policy ideas, including standards-based management, performance evaluation, accountability demands and evidence-based policies, all of which thrive on assessment data.

Such inertia in educational assessment again calls for greater awareness among stakeholders as argued above. Allowing a senseless and unreflective reproduction of educational assessment policies and practices runs the risk of creating systems unfit for human lives. Instead, decision-makers at various levels of education systems must take the lead and make wise decisions by asking questions such as: What should be the consequences for students associated with different forms of assessment? What is the level of flexibility in the system? How can everyone achieve positive destinations and recognition? Such questions require also reflections on what degree of standardization is actually required in different assessment tools, how educational assessments should be designed given the culture and values of the context, and what needs to be controlled and what can be left to trust (Lillejord, 2020).

References

- Addy, C., Gorur, R., 2020. Translating PISA, translating the world. *Comp. Educ.* 1–18. <https://doi.org/10.1080/03050068.2020.1771873>.
- Allerup, P.N., Kjeldsen, C.C., 2017. Standard setting in Denmark: challenges through computer-based adaptive testing. In: Blömeke, I.S., Gustafsson, J.-E. (Eds.), *Standard Setting in Education: The Nordic Countries in an International Perspective*. Springer, pp. 101–121 (Methodology of educational measurement and assessment).
- Andreasen, K.E., Rasmussen, A., Ydesen, C., 2013. In: Ainsworth, J. (Ed.), *Standardized Testing, Sociology of Education: An A-to-Z Guide*, vol. 2. Sage Publications, pp. 737–740.
- Au, W., 2009. *Unequal by Design: High-Stakes Testing and the Standardization of Inequality*. Routledge.

- Baird, J., Elliott, V., 2018. Metrics in education—control and corruption. *Oxf. Rev. Educ.* 44 (5), 533–544. <https://doi.org/10.1080/03054985.2018.1504858>.
- Baird, J., Andrich, D., Hopfenbeck, T.N., Stobart, G., 2017. Metrology of education. *Assess Educ. Princ. Pol. Pract.* 24 (3), 463–470. <https://doi.org/10.1080/0969594X.2017.1337628>.
- Biesta, G., 2015. On the two cultures of educational research, and how we might move ahead: reconsidering the ontology, axiology and praxeology of education. *Eur. Educ. Res. J.* 14 (1), 11–22. <https://doi.org/10.1177/1474904114565162>.
- Broadfoot, P., 2007. *Ance Introduction to Assessment*. Continuum.
- Brown, P., Lauder, H., Cheung, S.Y., 2020. *The Death of Human Capital?* Oxford University Press.
- Chapman, P.D., 1988. *Schools as Sorters – Lewis M. Terman, Applied Psychology, and the Intelligence Testing Movement, 1890-1930*. New York University Press, New York & London.
- Collins, R., 1990. Market closure and the conflict theory of the professions. In: Burrage, M., Torstendahl, R. (Eds.), *Professions in Theory and History: Rethinking the Study of the Professions*. SAGE Publications, London, pp. 24–43.
- Cornell, T., 2012. The Beginnings of Rome: Italy and Rome from the Bronze Age to the Punic Wars (c.1000–264 BC). In: *The Routledge History of the Ancient World*. Taylor & Francis.
- Danziger, K., 1990. *Constructing the Subject, Historical Origins of Psychological Research*. Cambridge University Press, Cambridge.
- Dorn, S., Ydesen, C., 2014. Towards a comparative and international history of school testing and accountability. *Educ. Pol. Anal. Arch.* 22 (120), 1–11. <https://doi.org/10.14507/epaa.v22.1913>.
- Feniger, Y., Israeli, M., Yehuda, S., 2016. The power of numbers: the adoption and consequences of national low-stakes standardised tests in Israel. *Glob. Soc. Educ.* 14 (2), 183–202. <https://doi.org/10.1080/14767724.2015.1010438>.
- Fuller, K., Stevenson, H., 2019. Global education reform: understanding the movement. *Educ. Rev.* 71 (1), 1–4. <https://doi.org/10.1080/00131911.2019.1532718>.
- Garsdal, J., Ydesen, C., 2009. The Debate on Educational and Psychological Testing in the United States: An Essay Review and Outsider Perspective with Some Philosophical Musings. Review of Madaus, George; Russell, Michael; and Higgi. *Education Review/Reseñas Educativas*. <http://edrev.asu.edu/edrev/index.php/ER/article/viewFile/1398/69>.
- Greck, S., 2009. Governing by numbers: the PISA “effect” in Europe. *J. Educ. Pol.* 24 (1), 23–37. <https://doi.org/10.1080/02680930802412669>.
- Hegarty, S., 2020. Inclusion and Learning Assessment: Policy and Practice. Background paper prepared for the 2020 Global Education Monitoring Report. <https://unesdoc.unesco.org/ark:/48223/pf00000373661>.
- Hill, D., Kumar, R. (Eds.), 2009. *Global Neoliberalism and Education and its Consequences*. Routledge, New York.
- Hopfenbeck, T.N., 2019. Does a test have to be fair to be valid? *Assess Educ. Princ. Pol. Pract.* 26 (5), 537–540. <https://doi.org/10.1080/0969594X.2019.1666229>.
- Kaestle, C., 2012. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, Princeton, NJ. http://www.gordoncommission.org/rsc/pdf/kaestle_testing_policy_us_historical_perspective.pdf.
- Kelly, P., Andreasen, K.E., Kousholt, K., McNess, E., Ydesen, C., 2018. Education governance and standardised tests in Denmark and England. *J. Educ. Pol.* 33 (6), 739–758. <https://doi.org/10.1080/02680939.2017.1360518>.
- Kreiner, S., Christensen, K.B., 2014. Analyses of model fit and robustness: a new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika* 79 (2), 210–231. <https://doi.org/10.1007/s11336-013-9347-z>.
- Krejsler, J.B., 2019. How a European “Fear of falling behind” discourse co-produces global standards: exploring the inbound and outbound performativity of the transnational turn in European education policy. In: Ydesen, C. (Ed.), *The OECD’s Historical Rise in Education: The Formation of a Global Governing Complex*. Springer International Publishing, pp. 245–267.
- Lagemann, E.C., 2009. *An Elusive Science the Troubling History of Education Research*. The University of Chicago Press.
- Lillejord, S., 2020. From “unintelligent” to intelligent accountability. *J. Educ. Change* 21 (1), 1–18. <https://doi.org/10.1007/s10833-020-09379-y>.
- Madaus, G.F., 1988. The distortion of teaching and testing: high-stakes testing and instruction. *Peabody J. Educ.* 65 (3), 29–46.
- McArthur, D.L., 1987. Educational assessment: a brief history. In: McArthur, D.L. (Ed.), *Alternative Approaches to the Assessment of Achievement*. Springer Netherlands, pp. 1–20. https://doi.org/10.1007/978-94-009-3257-9_1.
- Madaus, G.F., Russell, M.K., Higgins, J., 2009. *The Paradoxes of High Stakes Testing: How They Affect Students, Their Parents, Teachers, Principals, Schools, and Society*. Information Age Pub.
- Newton, P.E., 2007. Clarifying the purposes of educational assessment. *Assess Educ. Princ. Pol. Pract.* 14 (2), 149–170. <https://doi.org/10.1080/09695940701478321>.
- Nisbet, I., Shaw, S.D., 2019. Fair assessment viewed through the lenses of measurement theory. *Assess Educ. Princ. Pol. Pract.* <https://doi.org/10.1080/0969594X.2019.1586643>.
- Ozga, J., 2020. The politics of accountability. *J. Educ. Change* 21 (1), 19–35. <https://doi.org/10.1007/s10833-019-09354-2>.
- Phelps, R.P., Gottfredson, L.S., 2009. Summary and discussion. In: Phelps, R.P. (Ed.), *Correcting Fallacies about Educational and Psychological Testing*. American Psychological Association, Washington, pp. 247–255.
- Popkewitz, T.S., Lindblad, S., 2004. Historicizing the future: educational reform, systems of reason, and the making of children who are the future citizens. *J. Educ. Change* 5 (3), 229–247. <https://doi.org/10.1023/B:JEDU.0000041042.53119.f5>.
- Richardson, M., 2022. *Rebuilding Public Confidence in Educational Assessment*. UCL Press. <https://doi.org/10.14324/111.9781787357242>.
- Roediger III, H.L., Putnam, A.L., Smith, M.A., 2011. Chapter one - ten benefits of testing and their applications to educational practice. In: Mestre, J.P., Ross, B.H. (Eds.), *Psychology of Learning and Motivation*. Academic Press, pp. 1–36. <https://doi.org/10.1016/B978-0-12-387691-1.00001-6>.
- Simon, B., 1953. *Intelligence Testing and the Comprehensive School*. Lawrence & Wishart Ltd, London.
- Smith, W.C. (Ed.), 2016. *The Global Testing Culture: Shaping Education Policy, Perceptions, and Practice*. Symposium Books, Oxford.
- Spring, J., 2015. *Economization of Education: Human Capital, Global Corporations, Skills-Based Schooling*. Routledge, New York, London.
- Stobart, G., 2008. *Testing Times: The Uses and Abuses of Assessment*. Routledge.
- Svalastoga, K., 1959. *Prestige, Class and Mobility*. Gyldendal, Copenhagen.
- Tröhler, D., Popkewitz, T., Labaree, D.F., 2011. Introduction: children, citizens, and promised lands: comparative history of political cultures and schooling in the long 19th century. In: Tröhler, D., Popkewitz, T., Labaree, D.F. (Eds.), *Schooling and the Making of Citizens in the Long Nineteenth Century: Comparative Visions*. Routledge, New York, NY, pp. 1–27.
- Tröhler, D., 2020. National literacies, or modern education and the art of fabricating national minds. *J. Curric. Stud.* 52 (5), 620–635. <https://doi.org/10.1080/00220272.2020.1786727>.
- UNESCO, 2020. *Global Education Monitoring Report 2020: Inclusion and Education: All Means All*. UNESCO, Paris.
- Verger, A., Parcerisa, L., Fontdevila, C., 2018. The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms. *Educ. Rev.* 71 (1), 5–30. <https://doi.org/10.1080/00131911.2019.1522045>.
- Wang, R., 2013. *The Chinese Imperial Examination System: An Annotated Bibliography*. Rowman & Littlefield.
- Weber, M., 2009. *From Max Weber: Essays in Sociology*. Routledge, New York.
- William, D., 2010. Standardized testing and school accountability. *Educ. Psychol.* 45 (2), 107–122. <https://doi.org/10.1080/00461521003703060>.
- Williamson, B., 2018. Intimate data infrastructure : emerging comparative methods of predictive analytics and psycho-informatics. In: Gorur, R., Sellar, S., Steiner-Khamisi, G. (Eds.), *World Yearbook of Education 2019*. Routledge, Abingdon, pp. 59–75.
- Ydesen, C., 2011. *The Rise of High-Stakes Educational Testing in Denmark, 1920-1970*. Peter Lang Verlag, Frankfurt a.M.
- Ydesen, C., Bomholt, A., 2020. Accountability implications of the OECD’s economic approach to education: a historical case analysis. *J. Educ. Change* 21 (1), 37–57. <https://doi.org/10.1007/s10833-019-09355-1>.

U.S. federal assessment policy reforms in the twenty first century and their impacts

E. Caroline Wylie and Melissa Gholson, ETS, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

The assessment landscape in the US	108
The No Child Left Behind Act of 2001	109
NCLB and special populations	109
Every Student Succeeds Act of 2015	110
Positive impacts of federal assessment policies	111
Benefits for special populations	112
Negative impacts of federal assessment policies	113
Negative impacts on special populations	113
Impact of accountability assessments on teacher evaluations	114
Impact of accountability assessments on the broader assessment system	115
Conclusion	116
References	116

Education has the potential to serve as a driver for enabling social and economic development and reducing systemic inequality. Governmental policy makers must establish educational priorities in response to complex and critical societal challenges in an ever-changing landscape. Education policy makers have focused on policy reform efforts to improve the quality of education and supply a workforce for economic development (Barone and Ruggera, 2018; Sahlberg, 2016). Market system priorities are often reflected by competing political priorities entangled into national educational policies. National assessments are viewed as mechanisms for achieving broad policy goals, such as improving the labor market, promoting equity, and learning outcomes, or having an informed citizenry (Sahlberg, 2016), although such policies may not prioritize equitable educational outcomes for all learners (Sant, 2019).

Globally, educational reform has been articulated by “three policy principles: accountability, standards and decentralization” (Verger et al., 2019, p. 7). Decentralization is the use of national assessment results to control state, regional and local authorities often reflected by the tensions between national and local priorities. Reviews of international assessment policy and practice often highlight a tension between formative and summative purposes of assessment (also known as assessment *for* and assessment *of* learning), and the ways in which assessment policies can have unintended consequences (Birenbaum et al., 2015; Högberg and Lindgren, 2020; Verger and Parcerisa, 2018; Ydesen and Bomholt, 2019). However, some countries have managed to maximize the positive benefits of external accountability assessment through sampling approaches and careful attention to assessment timing (Heritage and Wylie, 2013; Volante et al., 2020). Every country’s accountability system has unique features and the nuances that come from understanding the context in detail are important. We examine educational policy and accountability assessment practices in the United States (US) during the first two decades of the 21st century in detail as a case study of how policy can achieve some of its initial goals as well as the unintended consequences that can emerge.

The assessment landscape in the US

During the first two decades of this century in the US there have been two significant acts of federal legislation, the *No Child Left Behind* Act of 2001 (NCLB; Public Law No. 107-110, §115, Stat. 1425, 2002) and its successor, the *Every Student Succeeds* Act of 2015 (ESSA; Public Law No. 114-95, §114 Stat. 1177, 2015). These legislations represent educational law for public schools in the US and provide the requirements for how state-level departments of education establish standards, administer annual standardized assessments, and implement accountability requirements. Both Acts resulted in considerable changes to how accountability assessments were conducted across the nation. NCLB mandated that summative assessments be administered to all students annually beginning at third grade, with sanctions imposed on schools that failed to meet established achievement targets. Under ESSA some of the sanctions were reduced, but state-level accountability assessments remained as a central component of the legislation. In sections that follow we examine these two Acts in more detail and the impacts that they had on teaching and learning in the US.

The reauthorization of the Elementary and Secondary Education Act in 2001 under the NCLB required states receiving Title I funding¹ to participate in National Assessment of Educational Progress (NAEP) reading and mathematics assessments. The

¹Title I funding comes from the federal level Department of Education to local educational agencies for children from low-income families to help ensure that all children meet challenging state academic standards.

NAEP is administered by the National Center for Education Statistics (NCES) within the U.S. Department of Education and the Institute of Education Sciences (IES). NAEP is a congressionally mandated project administration of assessments, generally on a 4-year cycle, in a variety of subjects. Results from the mathematics, reading and science assessment are perhaps the most often cited, but other content areas include civics, geography, US history, and technology and engineering literacy. Representative samples of students are assessed in each administration, and the results have provided longitudinal measures of what students know and can do at the national and state level. However, the NAEP assessments are not used as state accountability measures nor do they have sanctions associated with them, like the state-required assessments that followed the NCLB and ESSA legislations.

In the US, the regulation and governance of education systems is enacted at the state and territory levels. Rather than using sets of national standards, the 1990s saw each state create their own standards, and revise them after NCLB was enacted. Thus, accountability assessments are also state-specific rather than nationally administered. The closest that the US has come to national standards has been the Common Core State Standards (CCSS) in mathematics and English language arts, released in 2010 and adopted by 45 states and the District of Columbia. Although the CCSS initially were viewed as strong standards that almost all states could support, adoption of them became a political issue and since then 24 states have revised their standards (Coe and Daugherty, 2017; Evers, 2015). The introduction of the Next Generation Science Standards in 2013 have been adopted more slowly, with 19 states to date having signed on. During NCLB, a significant move away from the model of every state developing its own accountability assessments was initiated in 2009 through a federal grant competition (Race to the Top) that included encouragement of the adoption of the CCSS and support for the development of common state assessments by requiring consortia to establish a common cut score for all consortium states (Hain and Piper, 2016).

In addition to state autonomy, many school districts also function largely independently, described as “local control” which means that state departments of education can make recommendations, but not mandate decisions about curriculum or assessment at the local level. As a result of this devolved approach to education, while almost all students within a school district will take several state-determined accountability assessments each year, most assessments (e.g., classroom summative, interim, diagnostic) that students complete will be determined locally, either at the district or school level.

The No Child Left Behind Act of 2001

One motivation that informed the *No Child Left Behind* Act of 2001 was the need to close achievement gaps among groups of students, and the very title of the Act was intended to serve as a clarion call to act to meet the needs of students of color, and students from economically disadvantaged families, English Learners (ELs) and students with disabilities. The Act was intended to create increased accountability for assessment results, greater flexibility at the state and local levels to reduce bureaucracy, greater choice for parents of children who are attending failing schools, support for early grade reading, and a focus on improving teacher quality.

During the 1990s states had put in place accountability testing of students in at least fourth and eighth grade (U.S. Department of Education, National Center for Education Statistics, 2003). NCLB mandated that states had to adopt challenging academic standards and that student progress and achievement must be assessed annually in reading and mathematics from grade three through eight and once in high school. The data from these assessments, including results disaggregated by race, gender, and other criteria, were to be made publicly available to support comparisons across schools and districts.

One of the least realistic requirements of NCLB was that all students demonstrate proficiency on state standards by the year 2013–2014, as measured by the state assessments, and that schools make improvements toward that goal each year. While this was a federal mandate, each state had the freedom to decide what assessments to use and where to establish the bar for proficiency. School assessment results were then tracked to determine whether students were making progress toward achieving proficiency goals. Failure to meet these goals for 2 years in a row, described as “not making adequate yearly progress (AYP),” resulted in an escalating series of school sanctions, up to school closure. Accountability assessment results were reported by various subgroups (major racial/ethnic groups, students with disabilities, economically disadvantaged and ELs) which drew attention to schools that had overall strong student performance, but where the average results masked significantly lower results for subgroups of students (Hakuta, 2011). However, one challenge with the requirement of 100% of students demonstrating proficiency was that progress had to be made for each reportable subgroup of students. Schools with the most diverse student bodies would have to meet the AYP for every subgroup which penalized schools based on individual school demographics. In the next section we focus on students with disabilities and ELs given how NCLB regulations interacted with other federal policies.

NCLB and special populations

NCLB aimed to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at a minimum, proficiency on challenging state academic achievement standards on state academic assessments. One federal policy that intersects with NCLB’s assessment and accountability provisions is the Individuals with Disabilities Education Act (IDEA, 2004), originally enacted in 1975 as the Education for All Handicapped Children Act assured that a free and appropriate education is made available for all students with disabilities and established, to the maximum extent appropriate, that students be educated in the least restrictive environment to the extent possible, thus raising expectations for students with disabilities.

IDEA (2004) requires all students with disabilities receiving special education and related services have an Individualized Education Program (IEP). Each student’s IEP outlines the specially designed instruction necessary to provide full participation and access to the general curriculum and how students participate in state assessment (NCLB, 2002). Students with disabilities are included on

state assessments, with or without accommodations (i.e., changes made to the assessment itself or additional supports or services that allow students with disabilities to demonstrate their true knowledge, skills, and understandings on the assessment without being hindered by issues of accessibility), and students with the most significant cognitive disabilities may need alternate assessments (NCLB, 2002). NCLB regulations allowed alternate assessments to be based on either alternate academic achievement standards (AAAS) or modified academic achievement standards (MAAS). For accountability NCLB regulated states to the number of students enrolled in the tested grades for alternate assessment based on AAAS at one percent and alternate assessment based on MAAS at two percent (U.S. Department of Education, 2007). Alternate assessments based on the MAAS were later eliminated based on research that demonstrated that students with disabilities who struggle with reading and mathematics can achieve at grade-level standards if provided “appropriate instruction, services and supports” (U.S. Department of Education, 2015). IDEA (2004) requires every state to establish guidelines for accommodations to make sure that students use only accommodations that produce a score that is valid for school accountability purposes.

NCLB mandated the inclusion of all students in the state accountability system. In response to NCLB, students with disabilities with IEPs were expected to reach the same level of performance as all other students participating in the general curriculum. IDEA requires every state to establish guidelines for accommodations to make sure that students use only accommodations that produce a score that is valid for school accountability purposes. IDEA (2004) and accountability requirements served as a mechanism for change in raising expectations, gauging the progress of students with disabilities, and improving opportunities for access to the general classroom which have improved dramatically in the years since *No Child Left Behind* was implemented, as will be described below (Cole, 2006; Hardman and Dawson, 2008).

The law required the inclusion of English language learners in state assessment mandates and accountability systems; accordingly, these students were expected to make annual progress as measured on standardized tests of both English language proficiency and academic content. Under NCLB states are required to establish English language proficiency (ELP) standards and to use a single ELP test to assess students’ progress in and mastery of these standards in four domains: reading, writing, speaking, and listening (Fast et al., 2004). Prior to NCLB, English language proficiency (ELP) assessments were of poor quality in terms of content and psychometric properties (Abedi, 2008). NCLB regulations increased national awareness of ELs as a diverse and growing population, while also disproportionately labeling them as failing under NCLB accountability policies.

NCLB defined ELs as Limited English Proficient (LEP). States were required to include English language proficiency in their accountability systems including requirements for monitoring ELP progress and achievement through Annual Measurable Achievement Objectives (AMAOs). NCLB led to the termination of the Bilingual Education Act, which was replaced with the Title III,² the English Acquisition, Language Enhancement and Academic Achievement Act, which further emphasized English learning and removed the term bilingual from federal law. While states developed ELP standards, the mandate allowed local entities to choose the method of instruction was to be determined at the local level. Policies for improving English language acquisition did not address how to support the inclusion of ELs in standards-based instruction and failed to recognize the national emphasis on high stakes testing often led to teaching methods were ineffective for meeting the needs of ELs, such as a focus on test preparation (Menken, 2008). ELs disproportionately attend high poverty schools with limited resources and were unprepared to meet AYP expectations (Fry, 2008).

Every Student Succeeds Act of 2015

NCLB was in place for 14 years, despite significant criticism elaborated on below, to be replaced by the *Every Student Succeeds Act* (2015). Under ESSA, greater flexibility was granted for State and local education leaders to do what is best for children, while protecting historically underserved students. The state retains authority and flexibility to set goals within the federal framework (Saultz et al., 2017). The ESSA is clear that the role of the federal government would not “influence or incentivize” educational policy and specifically stated that authority would be up to individual states including the adoption of academic standards (ESSA, 2015, p. 130).

Specific areas of flexibility under ESSA include exempting eighth grade students from the mathematics assessments if the student takes an end of course assessment administered in high school for accountability purposes (ESSA, 2015). A State may permit its districts to administer a selected, nationally recognized high school academic assessment in place of the academic statewide assessment recognized by higher institutions of education for the purposes of entrance of placement in courses in postsecondary education or training programs. To encourage innovations in assessments systems, states may also apply to the Department of Education to pilot new approaches in some of its districts instead of using statewide assessments.

In some ways the ESSA did not make significant changes in that states must still conduct annual assessments in mathematics and reading from grades three to eight and once in high school, along with testing science once per grade band, in elementary, middle, and high school. Students with the most significant cognitive disabilities may be tested on alternate achievement standards (AAAS) aligned to grade-level and state standards. ESSA provides that states may have one percent of accountability testing to include students identified as having the most significant cognitive disabilities. ESSA, for the first-time required states to submit their state English language proficiency assessments, which includes alternate English language proficiency assessments for ELs with the most

²Title III funding is provided by the Department of Education to local educational agencies to help ensure that English learners (ELs) attain English language proficiency and meet state academic standards.

significant disabilities to the U.S. Department of Education for federal peer review³ (U.S. Department of Education, 2018). Flexibility is given whether to include recently arrived ELs who have been enrolled less than twelve months into state testing requirements. Under ESSA, reporting on state assessments has expanded significantly from the NCLB categories to also include children of military parents, foster children, homeless students, migrant students, gender categories, school quality, climate, and safety data, teacher equity data and per student expenditure data.

Under ESSA, funding flexibility has increased to provide resources to states and school districts to establish, implement, and sustain high-quality instruction designed to ensure that ELs develop both English language proficiency and content proficiency in math, English, and science. Additionally, states could create individual statewide accountability systems by developing plans which identify schools that are underperforming and how to support schools in getting them on track. States now establish achievement targets, and rather than federal penalties for struggling schools, schools will receive additional funding and support. States have significant discretion with respect to the funds that it must reserve to support school improvement activities in schools needing comprehensive or targeted support and improvement and for the State's statewide system of technical assistance and support. The State can distribute funds on a formula or a competitive basis, which may have impact on underperforming schools. Federal regulations also require that State accountability measures must include four academic factors: test scores for reading and math, English language proficiency test scores, high school graduation and a state-chosen academic measure for grade and middle school. States may choose a factor that affects school quality such as, kindergarten readiness, access and completion of advance coursework, college readiness, school climate or absenteeism.

Positive impacts of federal assessment policies

NCLB and ESSA have had several positive impacts on teacher qualifications, parent information, curricular alignment, student learning, and particularly on special populations, although as will be described below some of these positive impacts were relatively minor.

One requirement of the NCLB legislation was that core academic classes were to be taught by "highly qualified" teachers which was defined as having a bachelor's degree, being fully certified, and demonstrating their knowledge and skills in the subjects they teach through sufficient subject-matter coursework, passing a state test, or meeting other state criteria. The Center on Education Policy conducted a nationally representative survey of 349 school districts, between November 2006 and February 2007 (McMurrer, 2007), with one aspect of the survey focusing on changes to teacher qualifications and the perceived impact on student learning outcomes. One contradiction in the results is that while most districts were moving toward compliance, the majority of state and school district officials did not believe that the changes were having any impact on either teacher effectiveness or student learning outcomes (McMurrer, 2007). Less than ten percent of states or districts indicated that the law's requirement had improved student learning or teacher effectiveness to a great extent. Most districts indicated that they would be in compliance with the requirements by the 2006-07 school year, although some doubted that they would ever be given persistent staffing challenges. The lack of perceived impact of these requirements stemmed from the view that the NCLB definition of a highly qualified teacher was too narrowly focused on content knowledge rather than additional pedagogical competencies such as the ability to differentiate instruction to meet student needs (McMurrer, 2007). While there the proportion of teachers' meeting the standard of "highly qualified" improved, it was not a highly valued change given the focus on qualifications rather than a focus on teacher learning of critical teaching practices.

A second impact from increased use of accountability assessments and disaggregated reporting was that parents and other stakeholders had greater access to information about student learning outcomes, which could be compared to outcomes of other students within the school, district, and state. It is notable that across several different studies African American and Hispanic parents placed a greater value on standardized tests and their reports than White parents (e.g., Brewer et al., 2014; Phi Delta Kappa and Gallup, 2015; Tompson et al., 2013). Tompson et al. (2013) found that not only were parents of color more likely than White parents to think that it was important or extremely important to regularly assess whether students are meeting statewide expectations, similar differences were found when survey respondents were grouped by income with lower income parents responding more favorably to the importance of assessing whether students are meeting statewide expectations. Their report does not attribute reasons for the different attitudes among racial/ethnic or income groups, but in part it is likely a recognition that if disparities are not identified, they cannot be remediated.

A third area of practice affected by NCLB that was English language arts (ELA) and mathematics curriculum alignment with standards and with state assessments. A study by the Center on Education Policy, using a representative survey of 349 school districts, indicated that almost all the surveyed districts had required curricula for ELA and mathematics for elementary, middle, and high schools, with the majority reporting that it was aligned or very well aligned with the state assessments (McMurrer, 2008a). The study included case study reports from districts that provided examples of how districts had adopted standards, reviewed, and revised curricula, and created supporting documentation to help teachers (McMurrer, 2008a). A later study (Polikoff, 2012), using data

³The Peer Review process was first introduced as part of NCLB and required state departments of education to submit a report with documentation about the quality of their accountability assessments in mathematics and ELA based on peer review standards to a national panel of experts.

from the Surveys of Enacted Curriculum (SEC) found evidence of small to moderate changes in alignment between standards and reported classroom instruction, with the largest and most consistent increases in mathematics.

The fourth area of impact of NCLM was on student outcomes. In a study using NAEP data to examine the impact of the introduction of accountability systems on student outcomes pre-NCLB, [Hanushek and Raymond \(2005\)](#) concluded that the combination of reporting accountability assessment results together with school sanctions had an impact on student learning outcomes. One of the most thorough examinations of the impact of NCLB on student learning outcomes is the emphasis on data collection and examination of results from multiple data sources to gather evidence for accountability ([Dee et al., 2010](#)). To examine the impact on student learning outcomes the authors used data from the National Assessment of Educational Progress (NAEP) to examine national trends over time and to compare results for states that had previously more NCLB-like accountability requirements with those that did not. They also drew on data from international assessments, Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). From these various studies they concluded that NCLB had a positive impact on student performance in mathematics, particularly at lower-grade levels, and for traditionally disadvantaged populations, although there was not a comparable effect on reading ([Dee et al., 2010](#)). Another study ([Lauen and Gaddis, 2012](#)) examined the impact of NCLB specifically on the performance of subgroup performance in schools where that subgroup failed to meet AYP the previous year. They reported similar findings, although found improvements in both mathematics and reading, with small improvements for Black and poor students and greater improvements for Hispanic students. Moreover, this effect was stronger for schools where the subgroups were further from proficiency than those that were closer.

Benefits for special populations

In a descriptive study of twelve states [Harr-Robins et al. \(2015\)](#) examined the inclusion of students with disabilities into accountability systems under NCLB. Schools in each of the twelve states' accountability system identified as "consistently accountable" were compared to schools identified as "never accountable." They found that in schools identified as "consistently accountable" the number of students with disabilities moving from self-contained settings to regular classrooms increased in elementary schools (15.8%) and middle schools at (16.7%). Schools identified as consistently accountable were found to be more likely to implement highly effective practices such as, adopting reading and mathematics programs, supplying more instructional time and support, using assistive technology, and implementing positive behavior support programs ([Harr-Robins et al., 2015](#), p. 8).

ESSA (2015) requires that all students, including students with disabilities, have access to an appropriate education that will prepare all students for college and career readiness ([Tomasello and Brand, 2018](#)). Under ESSA, supports for students with disabilities have included support for universal design as a technical requirement for designing state assessments. The expansion of personalized learning to support individual needs of learners and the identification of projects and programs to support special populations and districts using the flexibility in funding ([Cook-Harvey et al., 2016](#)). Under ESSA, flexibility in funding provided professional development to support educators in meeting the needs of students with disabilities and ELs, led to class size reduction and increased mechanisms to engage parents, families, and communities.

As a result of ESSA, which emphasized college and career readiness for all students, other regulations were written such as the Strengthening Career and Technical Education for the 21st Century Act (2018) which sought to reform the public workforce system. These regulations emphasize the importance of the education-to-workforce pipeline to provide individuals with the academic, technical, and employability skills they need to be successful in the workforce. These regulations support students with disabilities on their path to college and career readiness, postsecondary education, and employment which can lead to improving independence, self-sufficiency, and opportunities for competitive integrated employment ([U.S. Department of Education, 2014](#); WIOA (Workforce Innovation and Opportunity Act), 2014).

ESSA encouraged states to include college and career readiness indicators in their accountability systems. States track and report on such indicators which has also benefitted students with disabilities ([Tomasello and Brand, 2018](#)). ESSA flexibility provides additional resources for supports and services for students with disabilities to pursue their transition goals. Transition planning from high school to postsecondary education is required under [IDEA \(2004\)](#). ESSA flexibility in funding supports local educational agencies to provide transition services, which improves postsecondary outcomes for students with disabilities. While flexibility in funding and inclusion into state systems are important, it is reliant upon state and local leadership to make wise decisions about how programs are delivered to create career and technical educational opportunities for students with disabilities.

ESSA changed the way schools identify, serve, test and report information about students who do not speak English. ESSA establishes consistency in the identification of ELs and requires a uniform process for the identification of ELs to be created at the state-level. Under NCLB, district level approaches established their own processes and resulted in a variety of criteria for identification and exiting services. ESSA provided flexibility in funding to support EL (English Language) programs. Under the new reporting requirements, EL (English Learner) subgroups will be disaggregated and ELs with disabilities will be separated from the overall EL population. This disaggregation represents a crucial step in supporting investigations to understand the effectiveness of different supports and instructional models for ELs ([Burr et al., 2015](#)).

Under ESSA, federal assessment policies offer flexibility for local priorities. For example, the Seal of Biliteracy was not a direct positive impact of ESSA, but rather more indirect. Many advocates for bilingual education have criticized ESSA regulations for promoting monolingual deficit perspectives, which elevates English language acquisition and fails to value multilingualism ([Kim, 2020](#); [Harper et al., 2008](#)). A grassroots effort which values diversity and the multilingual language skills of many ELs has

brought about a movement for a Seal of Biliteracy (TESOL, 2015). The Seal of Biliteracy is a language policy initiative in 32 states that is an award affixed to high school graduates' transcripts and diplomas, to promote and foster students "bilingualism and biliteracy" for dual language learners in K-12 schools (Heineke et al., 2018.)

Negative impacts of federal assessment policies

As identified in the previous sections, NCLB and ESSA had several positive impacts on classroom practice in terms of alignment of standards and curriculum and classroom practice, although improvement in teacher qualifications was not viewed as impactful and impacts on student learning were modest at best and certainly not what were needed to reach the goal of one hundred percent student proficiency by 2013–14. Further, taken to extremes the benefits of alignment also resulted in distortions to teaching and learning. In this section we review these negative impacts.

One of the documented negative impacts that resulted from perceived or actual pressures resulting from pressures around accountability assessments included reducing or eliminating time spent on subjects that were not part of the accountability testing program, which essentially meant everything except mathematics and ELA, particularly in elementary schools. The nationally representative survey conducted by the Center on Education Policy conducted (McMurrer, 2008b) reported large shifts in instructional time in favor of English language arts and mathematics and decreases in time for studies, science, art and music, physical education, recess, or lunch. In their sample of districts, 58% and 45% reported increasing the amount of time in elementary schools for ELA and mathematics, respectively. For those districts that reported increased time, it amounted to an extra 141 minutes for ELA and 89 minutes for mathematics per week.

One study used surveys and case studies to collect data from teachers about their instructional responses to NCLB (Hamilton and Berends, 2006; Hamilton and Stecher, 2006). The researchers identified evidence of teachers making efforts to identify more effective instructional methods, they also identified large majorities of teachers reallocating time to spend more time on tested standards, using item types common to the state assessments and teaching test-taking strategies. They also found evidence of principals encouraging teachers to focus their attention on students close to meeting standards. This later finding was also identified by Nichols and Berliner (2008) noting that teachers in some schools were encouraged to focus on students who were closest to reaching proficiency and to ignore those who were further from meeting proficiency levels. At the most extreme end was the Atlanta Public School system cheating scandal that involved both principals and teachers changing student responses on NCLB assessments (Vogell, 2011).

Darling-Hammond (2007) described some of negative impacts on schools that resulted from the passing of the NCLB Act, drawing attention to evidence from states that had already imposed sanctions for schools not meeting standards prior to NCLB being enacted, suggesting that at least some of these negative consequences should have been anticipated. For example, schools that were rated as failing found it difficult to both retain their existing teachers and recruit new teachers, a pattern that had been previously observed in Florida prior to NCLB (DeVise, 1999). Second, schools that were not meeting AYP could potentially lose federal funding, had to fund transportation of students who wished to transfer out to other schools, and had to use twenty percent of their Title I funds for student services such as tutoring or after-school programs, resulting in failing schools having less money to spend on their students.

Negative impacts on special populations

The focus on accountability under NCLB illuminated achievement gaps for underserved populations (Reardon, 2011). The unrealistic expectations of one hundred percent student proficiency carried mandated sanctions which were ineffective in supporting goals of equal opportunity (Koretz, 2008). District administrators often blamed the assessment performance of students with disabilities as being the only factor that kept a school from reaching AYP. The failure of the special education subgroup to make AYP occurred mainly because the subgroup is expected to maintain the same proficiency levels as their general education peers, ignored the reality that students with disabilities are placed in special education because they have unique needs and require specialized instruction as documented by an IEP and therefore need more time to demonstrate learning. The students with disabilities subgroup began at a lower level of proficiency than any other subgroup and were unrealistically expected to achieve proficiency levels which required them to grow at a faster rate than their general education peers (Eckes and Swando, 2008).

In addition to a backlash against special education students and programs perceived as responsible for a school's failure to meet AYP, NCLB assessment sanctions led to a focus on state test performance results which resulted in a narrowing of the curriculum. A focus on accountability led to a greater emphasis on time spent remediating on standardized test content at the expense of other instructional priorities, students with disabilities received more time in pull out services for reading and mathematics basic skills and limited opportunities to enroll in vocational and career programs needed for post high school success (Olson, 2004). Accountability policies that are focused on reporting achievement gaps are unlikely by themselves to sufficiently strengthen academic outcomes of students with disabilities and improve their access to the general education curriculum (Gilmour et al., 2019).

High stakes state assessments and accountability systems place enormous political pressure locally resulting in many states sought methods such as lowering cut scores to facilitate the illusion of proficiency gains for students with disabilities and economic disadvantage (Koretz, 2008). Under NCLB there was a failure to set a consistent policy about the number of students needed in a subgroup for results to be reportable for that subgroup. As a result, states had significant variation in the rules for establishing

the number of students required for reportable subgroups and how students were included in the accountability system, which became a mechanism to manipulate proficiency numbers. Another mechanism to inflate proficiency rates is the process by which states establish cut scores on tests, setting expectations low and reporting results which mask a lack of progress in learning (Ho, 2008).

ESSA's limitations of federal authority and the increased authority of states to determine failing schools may not be any improvement over NCLB for students with disabilities and other vulnerable populations (Saatcioglu et al., 2021). Special education advocates have been concerned about the lack of federal guidance for state accountability under ESSA and are concerned whether state and local districts will lower expectations for students with disabilities and access to the general classroom will become more restrictive for students with disabilities (Adler-Greene, 2019; Edgerton, 2019; Gilmour et al., 2019).

NCLB assessment and accountability requirements for ELs exposed issues in identification, greater risk of disability representation, restrictive language policies, lack of quality programming and limited opportunity to learn (Gándara and Hopkins, 2010). Research on opportunity to learn for both ELs and students with disabilities has documented differential opportunities to receive quality instruction and services commensurate to non-disadvantaged peers (Burr et al., 2015; Kurz, 2011; Lopez et al., 2016; Pullin and Haertel, 2008). A lack of instructional time and content coverage has a direct impact on assessment results. An important yet often overlooked indicator of ELs' opportunity to learn at the secondary level is access to and completion of rigorous, college preparatory coursework (Callahan and Shifrer, 2016). Such indicators get at the heart of equity and systematically work to address disparities in achievement. While state tracking of EL progress may be included in some state longitudinal data systems, many data systems are not yet developed to track the nuances of EL data that might be important to capture opportunity to learn such as advanced coursework, years in program, reclassified ELs and dually identified ELs with disabilities.

Many proponents in the field of English learning viewed reporting mandates as unfair since accountability tests are primarily administered in English, and language proficiency impacts the performance of ELs. Unsurprisingly, ELs have consistently underperformed on these tests. Menken (2010) highlighted the impact on performance and the consequences of the mandates on English language learners (ELs) for participation in both English language proficiency and academic content assessments. Menken's position is that by subgroup definition ELs have not mastered the English language, and language proficiency impacts the performance of ELs on those assessments. Menken (2010) found linguistically complex English words, likely unknown by ELs, dominate high stakes content assessments. ELs who have been in the US for 1 year take the same English language arts tests as those taken by native English speakers. Menken further elaborated that performance data then documents inferior performance, which further disenfranchises ELs. Schools serving ELs are disproportionately penalized under NCLB as accountability requirements that expect ELs to make annual progress as measured on standardized tests of both English language proficiency and academic content only serve to further maintain the achievement gap.

ESSA provides flexibility in funding which can be used to address the recruitment and professional development needs of educators supporting ELs. ELs are an increasingly growing population, and the provision of services is complex. There is a growing need for professional development and a shortage of educators to serve ELs and provide quality programs (Arroyo-Romano, 2016; Wigan et al., 2020; Satcher et al., 2019). Under ESSA, Title III funding has increased to provide resources to states and school districts to establish, implement, and sustain high-quality instruction designed to ensure that ELs develop both English language proficiency and content proficiency in math, English, and science. Many states have not yet addressed teacher shortages and have little capacity to provide the professional development that would provide quality services.

Under ESSA, key decision-making is up to individual states. In this era of increased flexibility, it is possible for states to move from a governmental locus of controlled accountability to one of local authority (Edgerton, 2019). State leaders have expressed concern about how to meet requirements for English language learners and ELs with disabilities (Tomasello and Brand, 2018). How states define underperforming schools is critical and many states are establishing low criteria (Ushomirsky et al., 2017). Equity issues are deeply concerning for students in historically underserved communities, and especially for ELs, students with economic disadvantage, students of color, and students with disabilities who have the most to lose if the spirit of the legislation under ESSA is not achieved.

Impact of accountability assessments on teacher evaluations

In addition to impacting students through an increase in accountability testing, NCLB also shifted key aspects of teacher evaluations using student assessment data as a major component of those evaluations although it was not a direct requirement of the legislation. Rather, one of the requirements of the NCLB legislation was for every classroom to have a "highly qualified teacher." In 2002, the National Governors Association (NGA) published six strategies for policy makers to move their states toward achieving this particular NCLB goal (Goldrick, 2002). Policymakers were encouraged to consider taking legislative action to:

- define teaching quality;
- focus evaluation policy on improving teaching practices;
- incorporate student learning into evaluation;
- create professional accountability;
- train evaluators; and
- broaden participation in evaluation design (Goldrick, 2002, p. 2).

The use of student learning data as part of teacher evaluation was not a widespread practice in 2002 when it was recommended by the NGA and was initially slow to gain traction. In a study of state statutes and regulations, Hazi and Racinski (2009) reported on how states had begun to implement these NGA strategies. At the time of their analysis, all but eight states had adopted at least one of the NGA strategies, but the strategy of incorporating student achievement into teacher evaluation ratings had been adopted by only 12 states. However, this initiative received increased attention from the Obama administration (2009–2017) and the then Secretary of Education, Arnie Duncan (2009–2015). Thus, 2 years later the number of states that had passed legislation that would incorporate student achievement data into teacher and principal evaluations had jumped to 23 (Piro et al., 2011). This accelerated focus on the use of student achievement data in teacher evaluations was in large part due to a requirement for states applying for Race to the Top funding. State departments could apply for Race to the Top funding in order to support reform efforts focused on improving their assessments, turn-around of failing schools, supporting teacher improvement efforts and teacher evaluation processes. Eighteen states and the District of Columbia were awarded Race to the Top funding in three phases from June 2010 to December 2011. The use of student test scores as part of teacher evaluation was further incentivized by the requirement to include it as part of a state request for states seeking waivers from NCLB requirements (Close et al., 2018).

By the 2016-17 school year all but four states had implemented teacher evaluation reform, with the majority using classroom observation data as the biggest contributor to the overall evaluation, but with student test score data playing an increasingly significant role as a component of teachers' summative evaluation (Steinberg and Donaldson, 2016). Teacher evaluation, and specifically the use of student test scores, was still a highly controversial topic. Supporters of the process argue that it makes sense to include a measure of student learning as part of teacher evaluation (Goldhaber, 2002; Rivkin et al., 2005) while opponents argue that the measurement challenges render its use invalid (Baker et al., 2010; Braun, 2015; Papay, 2011).

The simultaneous introduction of new accountability assessments to evaluate student progress against newly adopted Common Core State Standards (or close variants of those standards) and the use of these data to evaluate teachers led to a very contentious reception for those assessments. The use of student test score data was often included in teacher evaluation systems using student value-added models (VAMs) which are complex, statistical models designed to measure how much value a teacher adds to (or detracts from) students' growth based on large-scale standardized achievement tests over each school year (Haertel, 2013). There were many highly publicized reports of teachers who were viewed by parents, administrators and peers as exemplary but were penalized unfairly because of student test results (Fairbanks, 2015). Further pressure on evaluation systems came from lawsuits at the state and federal level which claimed that these systems violated teachers' constitutional rights, especially in systems that used student value-added data models as significant factors in high stakes decisions about teachers (Pullin, 2013; Sawchuck, 2015).

NCLB reforms heralded the widespread use of student test scores in teacher evaluations, holding the view that improving teacher quality could erase the achievement gap (Haertel, 2013). Almost 15 years later, federal influence on teacher evaluation ended with the passage of the ESSA in late 2015, ending the federal definition of a "qualified teacher" and leaving it up to states to define. In addition, it became a state-level responsibility to establish criteria for teacher evaluations. An examination by Close et al. (2019) of these changes indicated that the use of student assessment results in value-added models (VAMs) as a component of teacher evaluation was becoming less popular although still used by a sizable number of states: 15 states (29%) reported use of VAMs in summative evaluations, two states (4%) used them formatively, and 10 (20%) states left decisions about their use to local school districts. One state was in the process of developing their system at the time of the analysis, and the remaining 23 states (45%) do not use VAMs. An in-depth evaluation by Mette et al. (2020) one year later however suggests that more states have shifted decisions to local districts about whether to include VAMs. Given how contested was the use of VAMs, there will likely be few teachers or school- or district-based leaders who will mourn their loss.

Impact of accountability assessments on the broader assessment system

One documented impact of increased attention paid to accountability assessments in terms of both school and teacher sanctions was the emergence of a new form of assessment, frequently called interim or benchmark assessments (Marsh and Robyn, 2006; Perie et al., 2009). These assessments are often purchased from a vendor and completed by the majority of students within a grade level so that comparisons can be made across classes or schools within a district. These assessments can be administered for instructional, evaluative, or predictive purposes (Perie et al., 2009). As schools were paying increased attention on meeting AYP, these periodic assessments provided a way for teachers and principals to have access to more frequent information about student learning. Some of these assessments are designed to provide common end-of-unit evidence of student learning that can be used to inform additional learning supports for specific students. However, some interim assessments are essentially mini-accountability assessments, used to predict student performance on the end-of-year assessments. The breadth of content coverage of these assessments reduces their likely instructional value and formative impact on student learning (Nabors Oláh, 2022). Paul Black observed the US assessment system over several decades and noted in the earlier nineties that "[h]igh-stakes testing can dominate classroom work and so distort teaching that the conditions for good formative assessment do not exist" (Black, 1993, p. 52). Almost 20 years later as NCLB was in full swing "These current dominant forms [of accountability] have undermined what should be the normal and healthy relationship between formative and summative assessment" (Black et al., 2011, p. 72).

The seminal report *Knowing What Students Know* (National Research Council, 2001) described a vision for assessments in which summative and formative assessment contributed various kinds of information for different purposes. The authors recommended that an assessment system should possess three essential characteristics to support student learning: comprehensiveness, coherence,

and continuity. These initial characteristics were subsequently augmented by Chattergoon and Marion (2016) to include efficiency and utility. However, creating a balanced assessment system that reflects a common understanding of student learning across components (comprehensive), that can provide evidence of student learning over time (coherent), that provides multiple opportunities for students to demonstrate learning (continuous), and that provides enough information for stakeholders without being bloated with redundant assessment (balancing efficiency with utility) is no small challenge. As previous sections indicated, the accountability assessments arising from NCLB and ESSA created challenges for teachers, principals and district leaders that resulted in assessment systems failing to meet all the criteria laid out for a balanced system. As noted earlier, teachers felt it necessary to change, some might say distort, what they were teaching to account for the end-of-year accountability assessments (McMurrer, 2008b), and concerns were raised by formative assessment scholars such as Heritage (2010) that too much focus on accountability assessment resulted in insufficient use of formative assessment.

Conclusion

The 2020-21 COVID-19 crisis introduced major upheaval in the US education system, which resulted in the suspension of the accountability assessment process for a year, so it is not yet readily apparent whether ESSA goals are to be realized. Challenges of access to broadband technology for digital learning have occurred for many students living in rural communities. The loss of teaching time has further disadvantaged students with learning and language needs and the impact of interrupted schooling is unlikely to be fully understood for several years. Accountability assessments for the 2020-21 school year were administered but the results were not used for accountability purposes given the huge variations, even within states and districts, in student opportunity to learn and technology access. As a result, there is currently much discussion about the implications for state accountability assessments moving forward. In the short term, state and district leaders are having to rethink what assessment data will be most effective to inform teachers about student understanding and how to support their learning moving forward. In several states this has resulted in a new appreciation for or understanding of the role of formative assessment to support student learning (Rathke, K., July 12, 2021, personal communication; Farley, D., July 12, 2021, personal communication).

As accountability assessments are being reconceptualized, it is important to recognize the complexity of educational systems. Desired changes within a system are best achieved when the intended purposes and goals of assessment are made explicit. One way that states might go about addressing these complexities is to develop a theory of action.⁴ The National Council on Measurement in Education (NCME) developed a position paper on theories of action, noting that they should include explicit identification of the assessment components and associated supports, the intended outcomes and the causal mechanisms expected to lead to those outcomes. Ideally the theory of action should also anticipate the ways that improper use of the assessment could lead to unintended, negative outcomes (NCME, 2018). The process of developing a theory of action supports transparency in goals and outcomes to help stakeholders know if they are successful, and for which learners. It also invites discussion about whether the action mechanisms are plausible in terms of leading to the desired outcomes. A review such as this could have created opportunities to raise questions about whether the NCLB focus on *teacher* quality, as described by qualifications, rather than *teaching* quality or sanctions on poorly performing schools, rather than additional supports, would result in the desired closing of achievement gaps. While hindsight is always clearer than foresight, we advocate for the development and critical review of theories of action as one mechanism that can support a system perspective on accountability programs.

References

- Abedi, J., 2008. Classification system for English language learners: issues and recommendations. *Educ. Meas.: Issues Pract* 27 (3), 17–31. <https://doi.org/10.1111/j.1745-3992.2008.00125.x>.
- Adler-Greene, L., 2019. Every student succeeds act: are schools making sure every student succeeds? *Touro Law Rev.* 35 (1). Article 4. Available at: <https://digitalcommons.tourolaw.edu/lawreview/vol35/iss1/4>.
- Arroyo-Romano, J.E., 2016. Bilingual education candidates' challenges meeting the Spanish language/bilingual certification exam and the impact on teacher shortages in the state of Texas, USA. *J. Latinos Educ.* 15 (4), 275–286.
- Baker, E.L., Barton, P.E., Darling-Hammond, L., Haertel, E., Ladd, H.F., Linn, R.L., et al., 2010. Problems With the Use of Student Test Scores to Evaluate Teachers. EPI Briefing.
- Barone, C., Ruggera, L., 2018. Educational equalization stalled? Trends in inequality of educational opportunity between 1930 and 1980 across 26 European nations. *Eur. Soc.* 20 (1), 1–25. Paper# 278. Economic Policy Institute.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., 2015. International trends in the implementation of assessment for learning: implications for policy and practice. *Pol. Futures Educ.* 13 (1), 117–140.
- Black, P., 1993. Formative and summative assessment by teachers. *Stud. Sci. Educ.* 21 (1), 49–97. <https://doi.org/10.1080/03057269308560014>.
- Black, P., Wilson, M., Yao, S.Y., 2011. Road maps for learning: a guide to the navigation of learning progressions. *Measurement* 9 (2–3), 71–123.
- Braun, H., 2015. The value in value added depends on the ecology. *Educ. Res.* 44 (2), 127–131.
- Brewer, C., Knoepfel, R.C., Lindle, J.C., 2014. Consequential validity of accountability policy public understanding of assessments. *Educ. Pol.* 29, 711–745.

⁴A theory of action illustrates the claims made the impact of an assessment system through a diagram that connect system components to action mechanisms and then to intermediate and long-term outcomes for various stakeholders. A review of supporting literature for those claims made in the diagram can be one way to test their plausibility.

- Burr, E., Haas, E., Ferriere, K., 2015. Identifying and Supporting English Learner Students With Learning Disabilities: Key Issues in the Literature and State Practice (REL 2015–086). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West, Washington, D.C. Retrieved from: <http://ies.ed.gov/ncee/edlabs>.
- Callahan, R.M., Shiffrer, D., 2016. Equitable access for secondary English learner students: course taking as evidence of EL program effectiveness. *Educ. Adm. Q.* 52 (3), 463–496.
- Chattergoon, R., Marion, S.F., 2016. Not as easy as it sounds: designing a balanced assessment system. *State Educ. Stand.* 16 (1), 6–9.
- Close, K., Amrein-Beardsley, A., Collins, C., 2018. State-Level Assessments and Teacher Evaluation Systems After the Passage of the Every Student Succeeds Act: Some Steps in the Right Direction. National Education Policy Center.
- Close, K., Amrein-Beardsley, A., Collins, C., 2019. Mapping America's teacher evaluation plans under ESSA. *Phi Delta Kappan* 101 (2), 22–26.
- Coe, T., Daughtery, M., 2017. Strong Standards: A Review of Changes to State Standards Since the Common Core. Achieve, Inc.
- Cole, C., 2006. Closing the Achievement Gap Series: Part III. What is the Impact of NCLB on the Inclusion of Students With Disabilities? Education Policy Brief, vol. 4. Center for Evaluation and Education Policy, Indiana University. Number 11, Fall 2006.
- Cook-Harvey, C.M., Darling-Hammond, L., Lam, L., Mercer, C., Roc, M., 2016. Equity and ESSA: Leveraging Educational Opportunity Through the Every Student Succeeds Act. Learning Policy Institute, Palo Alto, CA.
- Darling-Hammond, L., 2007. Race, inequality and educational accountability: the irony of "No Child Left Behind". *Race Ethn. Educ.* 10 (3), 245–260.
- Dee, T.S., Jacob, B.A., Hoxby, C.M., Ladd, H.F., 2010. The impact of No Child Left Behind on students, teachers, and schools [with comments and discussion]. *Brook. Pap. Econ. Act.* 149–207.
- DeVise, D., 1999. A+ Plan Prompts Teacher Exodus in Broward County, vol. 5. Miami Herald.
- Eckes, S.E., Swando, J., 2008. Special education subgroups under NCLB: issues to consider. *Teach. Coll. Rec.* 111 (11), 2479–2504. November.
- Edgerton, A.K., 2019. The essence of ESSA: more control at the district level? *Phi Delta Kappan* 101 (2), 14–17.
- ESSA, 2015. Every Student Succeeds Act of 2015. Pub. L. No. 114-95 § 114 Stat. 1177 (2015–2016).
- Evers, W.M., 2015. Federal Overreach and Common Core. Pioneer Institute Public Policy Research. White Paper No.133.
- Fairbanks, A., 2015. Will Test-Based Teacher Evaluations Derail the Common Core? Hechinger Report. <https://hechingerreport.org/will-test-based-teacher-evaluations-derail-common-core/>.
- Fast, M., Ferrara, S., Conrad, D., 2004. Current Efforts in Developing English Language Proficiency Measures as Required by NCLB: Description of an 18-State Collaboration. American Institute for Research, Washington, D.C.
- Fry, R., 2008. The Role of Schools in the English Language Learner Achievement Gap. Pew Hispanic Center, Washington D.C.
- Gándara, P., Hopkins, M., 2010. English Learners and Restrictive Language Policies. Columbia University, Teachers College, New York.
- Gilmour, A.F., Fuchs, D., Wehby, J.H., 2019. Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Except. Child.* 85 (3), 329–346.
- Goldhaber, D., 2002. The mystery of good teaching. *Educ. Next* 2 (1), 50–55.
- Goldrick, L., 2002. Improving Teacher Evaluation to Improve Teaching Quality: Issue Brief. NGA Center for Best Practices, Education Policy Studies Division, Washington D.C.
- Haertel, E.H., 2013. Reliability and Validity of Inferences About Teachers Based on Student Test Scores. Education Testing Service, Princeton, NJ. Retrieved from: <http://www.ets.org/Media/Research/pdf/PICANG14.pdf>.
- Hain, B.A., Piper, C., 2016. PARCC as a case study in understanding the design of large-scale assessment in the era of common core state standards. In: Jiao, H., Lissitz, R.W. (Eds.), *The Next Generation of Testing: Common Core Standards, Smarter-Balanced, PARCC, and the Nationwide Testing Movement*. Information Age Publishing, Charlotte, N.C., pp. 29–48.
- Hakuta, K., 2011. Educating language minority students and affirming their equal rights: research and practical perspectives. *Educ. Res.* 40 (4), 163–174. <https://doi.org/10.3102/0013189X11404943>.
- Hamilton, L.S., Berends, M., 2006. Instructional Practices Related to Standards and Assessments. RAND, Santa Monica, CA.
- Hamilton, L.S., Stecher, B.M., 2006. Measuring Instructional Responses to Standards-Based Accountability. RAND, Santa Monica, CA.
- Hanushek, E.A., Raymond, M.E., 2005. Does school accountability lead to improved student performance? *J. Pol. Anal. Manag.* 24 (2), 297–327.
- Hardman, M.L., Dawson, S., 2008. The impact of federal public policy on curriculum and instruction for students with disabilities in the general classroom. *Prev. Sch. Fail.* 52 (2), 5–11.
- Harper, C.A., De Jong, E.J., Platt, E.J., 2008. Marginalizing English as a second language teacher expertise: the exclusionary consequence of No Child Left Behind. *Lang. Pol.* 7 (3), 267–284.
- Harr-Robins, J., Song, M., Garet, M., Danielson, L., 2015. School Practices and Accountability for Students With Disabilities (NCEE 2015–4006). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.
- Hazi, H.M., Rucinski, D.A., 2009. Teacher evaluation as a policy target for improved student learning: a fifty-state review of statute and regulatory action since NCLB. *Educ. Pol. Anal. Arch.* 17, 5.
- Heineke, A.J., Davin, K.J., Bedford, A., 2018. The seal of biliteracy: considering equity and access for English learners. [commentary]. *Educ. Pol. Anal. Arch.* 26 (99). <https://doi.org/10.14507/epaa.26.3825>.
- Heritage, M., 2010. Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity? Council of Chief State School Officers.
- Heritage, M., Wylie, E.C., 2013. Policy support for formative assessment. In: Heritage, M. (Ed.), *Formative Assessment Practice: A Process of Inquiry and Action*. Harvard Education Press, Cambridge, MA.
- Ho, A.D., 2008. The problem with "proficiency": limitations of statistics and policy under No Child Left Behind. *Educ. Res.* 37 (6), 351–360. <https://doi.org/10.31022F0013189X08323842>.
- Högberg, B., Lindgren, J., 2020. Outcome-based accountability regimes in OECD countries: a global policy model? *Comp. Educ.* 1–21.
- Individuals With Disabilities Education Act 20 U.S.C. §1400, 2004.
- Kim, A.I., 2020. (De) legitimization of monolingual ideologies in a US teachers' online forum. *Int. J. Biling. Educ. Biling.* 1–12.
- Koretz, D., 2008. *Measuring Up: What Educational Testing Really Tells Us*. Harvard Education Press, Cambridge, MA.
- Kurz, A., 2011. Access to what should be taught and will be tested: students' opportunity to learn the intended curriculum. In: Elliott, S.N., Kettler, R.J., Beddow, P.A., Kurz, A. (Eds.), *The Handbook of Accessible Achievement Tests for All Students: Bridging the Gaps Between Research, Practice, and Policy*. Springer, New York, NY.
- Lauen, D.L., Gaddis, S.M., 2012. Shining a light or fumbling in the dark? The effects of NCLB's subgroup-specific accountability on student achievement. *Educ. Eval. Pol. Anal.* 34 (2), 185–208.
- Lopez, A.A., Pooler, E., Linquanti, R., 2016. Key issues and opportunities in the initial identification and classification of English learners. *ETS Res. Rep. Ser.* 2016 (1), 1–10.
- Marsh, J.A., Robyn, A.E., 2006. School and District Improvement Efforts in Response to the No Child Left Behind Act. RAND, Santa Monica, CA.
- McMurrer, J., 2007. Implementing the No Child Left Behind Teacher Requirements. Center for Education Policy, Washington, D.C.
- McMurrer, J., 2008a. NCLB Year 5: Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era. Center on Education Policy.
- McMurrer, J., 2008b. Instructional Time in Elementary Schools: A Closer Look at Changes for Specific Subjects. Center on Education Policy.
- Menken, K., 2008. English learners left behind: standardized testing as language policy. *Multilingual Matters* 65.
- Menken, K., 2010. NCLB and English language learners: challenges and consequences. *Theory Pract.* 49 (2), 121–128.
- Mette, I., Aguilar, I., Wiczorek, D., 2020. A thirty-state analysis of teacher supervision and evaluation systems in the ESSA era. *J. Educ. Superv.* 3 (2), 105.
- Nabors Oláh, L., 2022. Interim and Benchmark Assessment. Routledge. <https://doi.org/10.4324/9781138609877-REE46-1>.

- National Research Council. 2001. *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- NCME. 2018. Position Statement on Theories of Action for Testing Programs. Online at: https://higherlogicdownload.s3.amazonaws.com/NCME/c53581e4-9882-4137-987b-4475f6cb502a/UploadedImages/Documents/NCME_Position_Paper_on_Theories_of_Action_-_Final_July_2018.pdf.
- Nichols, S.L., Berliner, D.C., 2008. Testing the joy out of learning. *Educ. Leader*. 65 (6), 14–18.
- No Child Left Behind Act of 2001 (NCLB) (2002) Pub. L. No. 107–110, 115 Stat. 1425.
- Olson, L., 2004. All means all. *Educ. Week* 23 (17), 44–50.
- Papay, J.P., 2011. Different tests, different answers: the stability of teacher value-added estimates across outcome measures. *Am. Educ. Res. J.* 48 (1), 163–193.
- Perie, M., Marion, S., Gong, B., 2009. Moving toward a comprehensive assessment system: a framework for considering interim assessments. *Educ. Meas.* 28 (3), 5–13.
- Phi Delta Kappa, Gallup, 2015. The 47th PDK/Gallup Poll of the public's attitudes toward the public schools: testing doesn't measure up for Americans. *Phi Delta Kappan* 97 (1). Retrieved from: <http://www.fsba.org/wp-content/uploads/2014/01/PDK-Gallup-Poll-2015.pdf>.
- Piro, J., Wiemers, R., Shutt, T., 2011. Using student achievement data in teacher and principal evaluations: a policy study. *Int. J. Educ. Leader. Prep.* 6 (4), n4.
- Polikoff, M.S., 2012. Instructional alignment under No Child Left Behind. *Am. J. Educ.* 118 (3), 341–368.
- Pullin, D., 2013. Legal issues in the use of student test scores and value-added models (VAM) to determine educational quality. *Educ. Pol. Anal. Arch.* 21 (6), 1–27. Retrieved from: <https://epaa.asu.edu/ojs/article/view/1160>.
- Pullin, D.C., Haertel, E.H., 2008. Assessment through the lens of “opportunity to learn”. In: Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J. (Eds.), *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press, New York, pp. 17–41.
- Reardon, S.F., 2011. The widening academic achievement gap between the rich and the poor: new evidence and possible explanations. In: Duncan, G.J., Murnane, R.J. (Eds.), *Wither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation, New York, NY, pp. 91–115.
- Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools, and academic achievement. *Econometrica* 73, 417–458.
- Saatcioglu, A., Skrtic, T.M., Kingston, N.M., 2021. High-stakes accountability in social and political context: skill gains and losses in the No Child Left Behind era. *Socio. Inq.* 91 (1), 60–113.
- Sahlberg, P., 2016. The global educational reform movement and its impact on schooling. In: *The Handbook of Global Education Policy*, pp. 128–144.
- Sant, E., 2019. Democratic education: a theoretical review (2006–2017). *Rev. Educ. Res.* 89 (5), 655–696.
- Saultz, A., Fusarelli, L.D., McEachin, A., 2017. The Every Student Succeeds Act, the decline of the federal role in education policy, and the curbing of executive authority. *Publius* 47 (3), 426–444.
- Sawchuck, S., October 6, 2015. Teacher Evaluation Heads to the Courts. *Education Week*.
- Steinberg, M.P., Donaldson, M.L., 2016. The new educational accountability: understanding the landscape of teacher evaluation in the post-NCLB era. *Educ. Finance Pol.* 11 (3), 340–359.
- Strengthening Career and Technical Education for the 21st Century Act (2018) Pub. L. No. 115-224.
- Sutcher, L., Darling-Hammond, L., Carver-Thomas, D., 2019. Understanding teacher shortages: an analysis of teacher supply and demand in the United States. *Educ. Pol. Anal. Arch.* 27 (35).
- TESOL, 2015. Seal of Biliteracy Guidelines. The American Council on the Teaching of Foreign Languages (ACTFL), the National Association of Bilingual Education (NABE), the National Council of State Supervisors for Languages (NCSSFL), and TESOL International Association. Retrieved on May 7 at: https://www.tesol.org/docs/default-source/advocacy/seal-of-biliteracy-approved-guidelines-final-3-10-2015.pdf?sfvrsn=3da414dc_2.
- Tomasello, J., Brand, B., 2018. How ESSA and IDEA Can Support College and Career Readiness for Students With Disabilities: Considerations for States. *College and Career Readiness and Success Center*.
- Tompson, T., Benz, J., Agiesta, J., 2013. Parents' Attitudes on the Quality of Education in the United States. Retrieved from: <http://www.apnorc.org/projects/Pages/parents-attitudes-on-the-quality-of-education-in-the-united-states.aspx>.
- U.S. Department of Education, 2007. Final Regulations of Modified Academic Achievement Standards. <https://www2.ed.gov/policy/speced/guid/modachieve-summary.html>.
- U.S. Department of Education, 2014. Dear Colleague Letter to Chief State School Officer. Retrieved from: <https://www2.ed.gov/about/offices/list/osep/osep-rda/050914rda-lette-to-chiefs-final.pdf>.
- U.S. Department of Education, 2015. Improving the Academic Achievement of the Disadvantaged; Assistance to States for the Education of Children With Disabilities, Final Regulations. *Federal Register* 80 (162), 50773–50785. <https://www.federalregister.gov/documents/2015/08/21/2015-20736/improving-the-academic-achievement-of-the-disadvantaged-assistance-to-states-for-the-education-of>.
- U.S. Department of Education, 2018. A State's Guide to U.S. Department of Education's Assessment Peer Review Guidance Process. U.S. Department of Education, Washington, D.C. Retrieved June 1, 2021, from: <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>.
- U.S. Department of Education, National Center for Education Statistics, 2003. In: Hurst, D., Tan, A., Meek, A., Sellers, J. (Eds.), *Overview and Inventory of State Education Reforms: 1990 to 2000*, NCES 2003–020. Project Officer: Edith McArthur, Washington, D.C.
- Ushomirsky, N., Smith, A., Bommelle, S., 2017. Trends in state ESSA plans: equity advocates still have work to do. *Education Trust*.
- Verger, A., Parcerisa, L., 2018. Test-based accountability and the rise of regulatory governance in education: a review of global drivers. In: Wilkins, A., Olmedo, A. (Eds.), *Education Governance and Social Theory: Interdisciplinary Approaches to Research*. Bloomsbury, London, UK, pp. 139–158.
- Verger, A., Parcerisa, L., Fontdevila, C., 2019. The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms. *Educ. Rev.* 71 (1), 5–30. <https://doi.org/10.1080/00131911.2019.1522045>.
- Vogell, H., July 26, 2011. Investigation in APS cheating finds unethical behavior across every level. *Atlanta J. Constitution*. <http://www.ajc.com/news/news/local/investigation-into-aps-cheating-finds-unethical-be/nQJHG/>.
- Volante, L., DeLuca, C., Adie, L., Baker, E., Harju-Luukkainen, H., Heritage, M., Schneider, C., Stobart, G., Tan, K., Wyatt-Smith, C., 2020. Synergy and tension between large-scale and classroom assessment: international trends. *Educ. Meas.* 39 (4), 21–29.
- Wiggin, G., Smith, D., Watson-Vandiver, M.J., 2020. The national teacher shortage, urban education and the cognitive sociology of labor. *Urban Rev.* 1–33.
- Workforce Innovation and Opportunity Act of 2014, P. L. 113-128, 29 U.S.C. Sec. 3101, et. seq.
- Ydesen, C., Bomholt, A., 2019. Accountability implications of the OECD's economic approach to education: a historical case analysis. *J. Educ. Change* 1–21.

Assessment reform and students' voices

Jannette Elwood^a and Amy Hanna^b, ^a Faculty of Arts, Humanities and Social Sciences, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom; and ^b School of Education, University of Strathclyde, Glasgow, Scotland, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	119
Policy reform	120
Education policy reform	120
Assessment policy reform	120
Student voices: research and practice	122
Key messages from research in assessment reform and students' voices	123
Student voices on assessment – reform, policy and experiences	124
Learning from listening to young people	124
Young people as assessment policy actors	125
System level issues	125
Equality (and ethical) implications at the system level	125
Discussion and reflections	126
References	127

Introduction

Educational assessment is often positioned as having three dimensions: policy, process and practice. These dimensions are not mutually exclusive but co-exist in relation to how assessment is considered within educational settings (the doing of assessment), but also how it is used systematically to be a major driver of educational change (the use of assessment). Assessment under these three dimensions is continually modified, adapted and adjusted to suit both pedagogical initiatives to enhance learning (see Mary James within this same publication) but also to advance political objectives. Assessment is used to drive policy imperatives within education more generally as a way of allocating resources (based on assessment outcomes) or cementing ideological directives through mandatory assessment of defined curricula and associated subject hierarchies. The consequences of these modifications, adaptations and adjustments to local, national and international educational assessment policy, processes and practices show assessment to be a much more complex socio-political phenomena than systems of neutral, technical evaluations of students' learning. Thus, for decades, across many jurisdictions and across many governments, assessment has been placed firmly within the education policy reform arena with all the subsequent turmoil that high profile policy focus can bring.

Another area of educational reform that has garnered more attention over recent years is that of "student voice" (Bourke and Loveridge, 2018). Cook-Sather (2006, 2018) has detailed the evolution of student voice in educational research that has, over time, repositioned the worth of students' unique perspectives on teaching, learning and schooling in general. Fielding (2001) suggests that the student voice movement can be seen as a partnership that has developed along a trajectory that has moved from students as a data source in early studies to students as active respondents, co-enquirers and co-knowledge creators. However, there has been considerable critique and debate about the capturing of student voice to inform educational policy and change practice. Lundy (2007) has argued that the inclusion of student voice is about more than just good practice; it is about children's rights. She is particular in arguing that fully embracing children and young people as key stakeholders in their educational experience and provision is rarely positioned by government in terms of children's rights. Under Article 12 of the United Nations Convention on the Rights of the Child (UNCRC), not only must governments consult children and young people in all matters that affect them (including education), but they must also support children's contribution to decision-making as well as listen to their views in constructive ways (Lundy 2007).

This chapter brings together these two seemingly disparate areas of educational interest, assessment reform and student voice. Rarely are the two areas considered as mutually enhancing, yet as we will show, there is a growing interest in how the practices of consulting young people about assessment matters and having them participate in the deliberations around national assessment policies can bring to the fore more radical ways of thinking about one of the most important activities in the educational lives of students. We begin by considering the policy landscape in education and assessment, together with research on student voice before discussing key messages and reflections that emerge when these seemingly disparate areas of educational interest are brought together.

Policy reform

Education policy reform

For many decades educational researchers have been investigating an ever-changing landscape of education policy reform (Taylor et al., 1997) that has seen shifts in how education policy and its reform has been considered and understood. Changes in definitions, approaches and theoretical analyses are well-documented and show how thinking has changed in this field as education policy has moved from a local responsibility to a national and global enterprise (Lingard et al., 2016). Ball and Bowe (2020, p. 97) inform us that for the last 40 years, state education systems have been “subjected to incessant and fundamental reform” as governments around the world have navigated new social, economic and technological challenges while also looking to exert more authority in what happens in schools and in schooling.

The field of education policy is a highly contested one, yet the meaning of policy in the everyday education arena is often a taken-for-granted understanding of what policy is. Ozga (2000) stresses that there is no fixed, single definition of policy and how it is understood depends significantly on the perspective of the researcher, policy maker, or educational stakeholder. Some understand policy as it relates to the actions of governments documented in official texts that are handed down to local legislators, school governors, schools, teachers, etc. who are expected to implement these actions for the required educational outcomes. However, Ozga (2000, p. 1, 2) takes a more “diffuse” view of policy. She defines policy as “a process rather than a product, involving negotiation, contestation or struggle between different groups who may lie outside the formal machinery of official policy making ... [policy] is struggled over, not delivered in tablets of stone to a grateful or quiescent population”. Ball et al. (2012, p. 3) take our understanding further to consider policy as a process of enactment. Under Ball, Maguire and Braun’s understanding, policy is taken as “things” (legislation and national strategies) but also “discursive processes that are complexly configured, contextually mediated, and institutionally rendered. Policy is done by and done to teachers; they are actors and subjects, subject to, and objects of, policy”. Both Ozga (2000) and (Ball et al., 2012) show that many studies that focus on how educational policies are disseminated describe ways in which they are implemented, usually from governments to educational institutions in “top-down” or “bottom-up” style models. This is often seen as the way in which educational policy works, with further local support, in a cascading manner, interpreting what is intended by the policy in order for schools and teachers to embrace and integrate the change into their own practice. Yet, in contrast, a different positioning considers policy as a series of enactments; “as dynamic and non-linear aspects of the whole complex that makes up the policy process” (Ball et al., 2012, p. 6) in which stakeholders (teachers, students, senior managers, etc.) act as policy actors who mediate and “produce” policy through interpretation and adaptation within local contexts and settings.

Researchers who work more predominantly in the field of childhood studies and public policy domains conceptualize students as public policy actors and they call for young people to have their views taken into account in policy decisions that affect them directly. They argue that young people need to be afforded greater opportunities to participate in political life at the public level (Arnott, 2008; Hinton et al., 2008). A common underpinning of these calls for greater involvement of children in public policy is their right to participation under article 12 of the UNCRC (see more detail below in Section 3) which asserts children’s right to participate in decisions which affect them directly (United Nations, 2001; see also Lundy, 2007). While the UNCRC has been influential and has driven much research in childhood and participation, the drive for children to be recognized as public policy actors is not only made on a legal basis. As users of public policy, children are affected by reforms to public institutions and policies, and so, it is argued, they should be given the opportunity to participate in decisions regarding change and development (Tisdall, 2008). Underlying this is a shift in perceptions of young people as having the capacity to give valid insights on institutions and services (Arnott, 2008).

In the field of Education as a form of public policy, students are often missing from deliberations on policy reform and how they enact policy through their own interpretations as well as mediating those of their teachers. In this chapter, and the alignment of policy reform, especially in the area of assessment, we take the position that Ozga (2000) and Ball et al. (2012) as well as others, offer us on policy: fluid processes that are contextualized and understood within the environments in which they are situated where students are (and should be) at the heart of policy enactment. Student voice research has shown the value of consultation and participation of students in their own educational development and experience, as well as the improvement of school development activities and the inclusivity of schools that consult all stakeholders about key activities (see Baroutsis et al., 2016; Charteris and Sardon, 2019a, 2019b; Mitra, 2018; Flutter and Rudduck, 2004 for a clear articulation of the workings of student voice work in education). However, students have been less involved in what they themselves call “higher level” policy reforms either in school settings or with governments directly; assessment has been identified as one such “high level” policy arena where students and their voices are absent. More recent research (Elwood, 2012, 2013; Elwood et al., 2017) has taken students as key policy actors in the assessment arena and as such has much to tell policy makers and assessment developers about fair and relevant assessment. It is to this assessment policy reform we now turn.

Assessment policy reform

Isaacs and Lamprianou (2018) remind us that over the last 30 years, governments have increasingly been using assessment policy reforms as mechanisms for improving school outcomes locally, as well as reinforcing their social and economic positions within world-wide settings. National assessment outcomes at both primary and secondary levels and educational policy frameworks that set benchmarks for raising standards in schools are invariably linked to narratives about educating children to face a rapidly changing world as well as improving the economic preparedness of young people as they transition from education to work.

Thus, assessment systems, once a relatively everyday way to gauge children's and young peoples' educational achievements and provide them with educational outcome profiles to make choices about the next stage, have become a major driver for policy makers to influence policy. Most importantly, they have significantly changed pedagogical practice as well as the adaptation of centrally derived curricula in schools.

At local levels across many jurisdictions, we have seen a "take-over" of assessment systems from locally devised programmes, supported by teachers and test development agencies in collaboration and direct communication with schools, to centralized systems of large-scale assessments. These are formally administered at designated key stages, the outcomes of which are used to evaluate not only the students taking these assessments, but teachers, schools and local educational authorities, often to the detriment of schools and their staff if they do not meet set national benchmarks of performance. In the UK, [Baird and Lee-Kelley \(2009\)](#) have described this "take over" of assessment systems and increased government involvement in assessment provision as a significant policy frontline, with assessment policy reform often implemented at speed to enable governments to evidence measurable impacts on education to the electorate within their terms of office. Thus the "role of national assessments can become dominated by a culture of delivery – in particular, the fast delivery of government policies as politicians look for quick wins between elections" ([Isaacs and Lamprianou, 2018; Martens and Jakobi, 2010](#)).

In relation to summative assessments and end of school examinations, there are real risks posed to the reliability and validity of these systems with this culture of fast delivery ([Cadwallader and Tremain, 2013](#)). Limited time frames for development, and lack of piloting and evaluation, allow instability to creep into new systems, often resulting in students and schools left with only post-hoc appeals on final awarded grades when systems do not work as intended. [Cadwallader and Tremain \(2013\)](#) have suggested that many of the problems experienced with more recent assessment policy reform can be understood within the amplified, politicized context of reform programmes and the use of assessment outcomes to fulfill multiple purposes (e.g. individual grades for students, performance measures for teachers and schools). This deliberate and increased use of assessment policy to drive educational change and national assessment systems used to evaluate more things than students, are all activities that have fallen foul of a key fundamental principle with regard to the use (and/or misuse) of assessments and their social impact: "the more any quantitative social indicator is used for social decision making the more likely it will be to distort and corrupt the social processes it is intended to monitor" ([Madaus, 1988](#), p. 89). Once assessment systems become high stakes, even those that might be reasonable monitors of educational success can lose dependability and credibility ([Koretz, 2008](#)).

For years, many assessment researchers have been increasingly concerned with the negative washback that such a reliance on summative assessment has had on the improvement of students' learning and teachers' assessment practice. A global response to this negative washback over recent years has included concerted attempts by educational researchers and practitioners to influence governments away from a reliance on summative assessments and tests and to look to more formative assessment policies to drive change and improve teachers' own assessment practices ([Black and Wiliam, 1998a, 2003; Hopfenbeck and Stobart, 2015](#)). [Florez-Petour \(2015, p. 3\)](#) shows clearly how the policy aspects of formative assessment (colloquially termed *Assessment for Learning*) have been globally imported, piloted, researched and evaluated in a wide variety of contexts; so much so that she suggests that the movement has "the characteristics of a research epidemic". Indeed, many governments have reviewed their national curriculum and associated assessment systems in light of formative assessment developments and looked to the large-scale implementation of assessment for learning programmes to improve learning, teaching and educational standards (see [Hopfenbeck and Stobart, 2015](#) for a special edition on this issue). However, in spite of significant support and almost "evangelical" adoption by practitioners and researchers of the tenets of formative assessment, [Florez-Petour \(2015\)](#) goes on to show that besides its widely recognized positive effects on students and teachers, major questions have seldom been asked about the true efficacy of formative assessment at national levels. She argues that the wide extent of policy borrowing across many contexts that has occurred with frameworks of formative assessment has had minimal impact on the pervasive presence of summative assessment systems. The success of this movement in driving change and improving learning at national levels has at best been "partial" ([Black, 2015, p. 161](#)) with many small-scale interventions successful at the classroom and local levels through increasing the validity of student evaluation but failing to translate into national systems that are robust and reliable, and able to dislodge long established systems of summative assessment.

In broader, more international global perspectives, any consideration of assessment policy and reform need to consider large-scale international assessments and the impact their outcomes can have on nations' policy formation and implementation. At international levels, the rise of the OECD's program for international student assessment (PISA) has had major influence on national governments' assessment policy discussions and directives with direct international policy borrowing becoming ever more prominent as a direct consequence of nations' short comings on PISA tests. [Sellar and Lingard \(2014\)](#) suggest that the OECD has an ever-growing importance and influence in global governance of education through its testing programmes that can direct policy and practice through "soft power", particularly through its unique mechanisms of peer pressure and review. They go on to argue that the "soft power" exerted by the OECD relies on the organizations' leadership drive to both "define the problem" but also to "offer the solution" ([Martens and Jakobi, 2010](#)) through policy evaluation and data generation: "the OECD has therefore gained an important status in several stages of national policy-making, ranging from agenda setting to policy formulation and implementation" ([Martens and Jakobi, 2010, p. 175 in Sellar and Lingard, 2014, p. 929](#)).

With so much continuous policy reform within the assessment arena, there have recently been more concerted efforts from researchers interested in both assessment and students' voices to foreground the importance of listening more systematically to students' experiences as an essential component of understanding the impact of national assessment systems and policies (especially those at school-leaving age) on young people (e.g. [Baird et al., 2011; Banks and Smyth, 2015; Chamberlain, 2012; Daly](#)

et al., 2012; Elwood, 2012; Elwood et al., 2017). Some researchers have started to present the issue of assessment and assessment reform in relation to impact on students as a rights issue (see Elwood and Lundy, 2010); a link that while more obvious now than perhaps in the past, is still rarely made. Thus, there have been calls for students' rights to be taken seriously in light of potential negative washback of large-scale assessment systems, as well as significant power imbalances between test-takers and highly influential global testing organizations (Shohamy, 2001). As the next section of this chapter outlines, the UNCRC designates children as rights holders with rights entitlements, and as such demands that "the best interests of the child" guide decisions regarding public policy actions that affect children and young people directly (Elwood and Lundy, 2010, p. 335). With regard to assessment, Elwood and Lundy (2010) argue that those devising national testing systems should abide by the UNCRC's core principles and take positive action to ensure that young people's views are taken into account in assessment decisions as a matter of course.

Student voices: research and practice

Student voice as a discipline emerged in the early 1990s as educators noted the silence of young people about matters affecting them, at a time when the UNCRC was being widely ratified by countries around the world (see Noyes, 2005; Lundy, 2007). Article 12 of the UNCRC states that "States Parties shall assure to the child who is capable of forming his or her views the right to express those views freely in all matters affecting the child, the views of the child being given due weight in accordance with the age and maturity of the child". In practice, the inclusion of young people's voices in decision making has repositioned student perspectives on teaching and learning more generally (Cook-Sather, 2006), perhaps because "student voice" itself is repositioned as a rights entitlement using the UNCRC, not merely a benevolent idea extended to young people when deemed appropriate by adults. Student voice is widely recognized as a medium for examining the relationship between pedagogy and curriculum (Baroutsis et al., 2016) and as an instrument in assessment for learning (Thomson and Gunter, 2006). Academic and practitioner work has proliferated as early champions of student voice argued for systemic inclusion of students' perspectives in education (see, for example, Fielding, 2001, 2004; Flutter and Ruddock, 2004; Arnot et al., 2004; Cook-Sather, 2002).

Cook-Sather (2006) suggests that the word "voice" arose from the silence of student voices in educational policy and practice. "Student voice" does not indicate simply words spoken by students, however, but includes many ways in which young people express their views and experiences (Robinson and Taylor, 2007). Whilst there is no fixed definition of student voice, its adaptation in a range of contexts and purposes is premised on three convictions: young people have unique perspectives on their education, including teaching and learning; their insights, views and experiences warrant the attention and response of adults; and young people should be afforded opportunities to shape their education (Cook-Sather, 2006). Student voice is thereby founded on students' ability to exert influence in their learning context, to transform learning experiences, and to expand learning (Cook-Sather, 2020).

Proponents of student voice work are amongst the strongest advocates of the need to conceptualize young people as actors rather than the acted upon, or subjects of intervention and assessment (Cook-Sather, 2010). For Cook-Sather (2006), the convictions of rights and respect which underlie student voice demand a rethink of our assumptions about who can and should be an authority on educational practice. Such student voice approaches, however, require a shift in mindset and a sharing of power (Cook-Sather, 2020). Consultation and participation, inherent in student voice practices, are an enactment of redefining the status of young people in education (Ruddock and Fielding, 2006), and article 12 of the UNCRC has played a major part in such changes in mindset because the recognition of children's participation rights demands a transformation of children from the status of passive recipients to respect as active agents with rights entitlements.

Lundy (2007) observes that the implementation of article 12, however, depends on cooperation of adults who may be skeptical about children's capacity to have meaningful input in decision-making, belief that giving children a say will undermine adult authority, and concerns about the time required to comply with article 12. By conceptualizing article 12 in a four-part model: space; voice; audience; and influence, Lundy (2007) addresses some critiques of student voice (see, for example, Robinson and Taylor, 2007) as being tokenistic in some forms by presenting a model of consulting young people in decision-making processes. Yet, Lundy (2018) re-evaluates the criticism of tokenism in child participation, reminding us that tokenism does not justify the stance that it is better not to consult young people at all, than to involve them in a manner that may be tokenistic. Whilst tokenistic consultation with young people may have no impact on decision-making (Tisdall, 2015), Lundy (2018) argues that children also have a right to freedom of expression under article 13 of the UNCRC, regardless of whether these views are given due weight, and States Parties should take measures to respect and fulfill this right.

Discussions and critiques of tokenism in student voice highlight that it is not without its shortcomings. Matters of voice are embedded in structures and relations of power, and student voice is often understood as a challenge to such power structures and relations (Cook-Sather, 2006; Ruddock and Fielding, 2006; Taylor and Robinson, 2009), which can be met with resistance by adults. Fielding (2004) cites Alcoff (1991/1992, p. 12) who asserts that "who is speaking, who is spoken of, and who listens is a result, as well as an act, of political struggle". It is perhaps pertinent to clarify, therefore, that how student voice is defined, what forms it takes, who listens, and how it is interpreted depends in part on the structural context in which this relationship exists (Cook-Sather, 2006). For most students, the opportunity to voice their opinions and experiences rests on the premise of adult educators and policy makers asking them to do so.

The right to have a say, when linked to the right to have that say matter, means that to speak is to have power (Cook-Sather, 2014). Nevertheless, the term "voice" implies a singular, unitary voice but such an assumption is misleading (Robinson and Taylor,

2007); in fact, Arnot and Reay (2004) and Reay (2006) suggest that instead student voice is a cacophony of competing voices. Limitations of some student voice work reflect the tendency by some to perceive the experiences of certain young people as representative of all others (Bragg, 2007; Thomson, 2011). Such assumptions about a homogenous "student voice" deny diversity in students' needs and opinions and serve to occlude the hierarchies of power at work across different student groupings (Cook-Sather, 2006). Any consultation should therefore be reflexively aware of the multiplicity of voices, and the various power disparities that may inhere within and between these groupings as well as in comparison with adults. Student voice, consequently, demands an understanding of equity; what Cook-Sather (2006) argues is founded in rights and respect. Within education, curriculum and pedagogy are the medium and mode of participation for both teachers and students (Lundy and Cook-Sather, 2016), and students' right to express their views freely in matters affecting them, and have them given due weight, is denied if student voices are absent from learning processes, or only allowed in response to adult decisions (Lundy and Cook-Sather, 2016).

Despite this, consideration of children's rights specifically in relation to educational assessment has been neglected (Elwood, 2012; Elwood and Lundy, 2010), and studies within the field of assessment policy formation, qualifications and examination development that have considered students' involvement as critical stakeholders have been more limited in number (Elwood and Lundy, 2010). Indeed, Elwood (2012, p. 499) argues that the qualifications arena seems to be one of the "last bastions of education that has fostered good practice of involving students in policy formation and in consultation around decisions that affect them". Over the years, calls for changes in mind-sets, structures and power relations to support respectful dialogue have emphasized the potential for student voice to enhance understanding of the impacts and effects of educational policies and practices as they are played out (Cook-Sather, 2002; Fielding, 2001). Nonetheless, in some cases, student voice is co-opted into managerial agendas as an imperative of accountability rather than recognition of young people's subjectivity (Fielding, 2001). Indeed, some forms of student voice are contrived through neoliberal perspectives, where young people may be consulted about their learning to raise standards and enhance outcomes for comparative data (Fielding, 2004).

Percy-Smith (2006, p. 172) argues that "the honeymoon period for young people's participation and the celebration of their voices has now passed" and highlights the lack of attention paid to the "wider social, organizational and systemic contexts within which young people participate" – contexts which have become somewhat stark considering assessment decisions during the Covid-19 pandemic. Whilst Bourke and Loveridge (2014, p. 143) argue that "student voice, when used effectively, inherently becomes a pedagogical tool for effective teaching, learning, assessment and social practices within the classroom", evidence of the impact of student voice, and article 12 of the Convention, in the context of assessment continues to be difficult to find (Elwood and Lundy, 2010). Where it has been considered systematically and formally in educational research in assessment reform and student voice, key messages have arisen that are pertinent to the development of assessment systems for fair and just evaluation of student learning. We explore some of these key messages next.

Key messages from research in assessment reform and students' voices

In this section we explore some key messages that have emerged from research which has considered the engagement of students in consultation about their views and experiences of assessment reform. These messages have emerged from studies that were carried out within the jurisdictions of the UK – particularly England, Wales and Northern Ireland (Baird et al., 2011; Barrance and Elwood 2018a, 2018b; Elwood, 2012, 2013a) – as well as the Republic of Ireland (Baird et al. 2015, 2016; Elwood et al., 2017). We consider the key messages collated from these studies about student's experiences of examinations and assessment at major junctures in their education, in "high-stakes" (for them) situations, and in different educational policy contexts. We also present key messages to show the articulation of students' views and perspectives on high-stakes examinations systems about which they are rarely consulted and to offer their perspectives as a contribution to overall discussions about the improvement of assessment policy reform more generally.

In line with Ball et al. (2012) and Cook-Sather (2006) discussed above, students in these studies were considered as social and political agents interacting and mediating assessment policy and practice. These studies set out to change the perception that students do not have something of value to say about assessment reform. The studies collated students' views on the most significant assessment phases of their education, specifically examinations at the end of compulsory schooling (GCSEs and A levels in England, Wales and Northern Ireland for those aged 16 and 18 respectively; Leaving Certificate (LC) in Ireland for students aged 17/18). The authors are aware that there are other studies internationally that have also considered students' views and experiences of assessment (Brown and Hirschfeld, 2008; DeLuca et al., 2018; Hodgson and Spours, 2003, 2005), but the studies referred to in the following sections positioned students deliberately as "social policy actors", significant players in the mediation of national assessment systems not just subjects in their implementation and completion.

As indicated above, students have long been consulted about their learning. In educational practice, consulting with students and having them directly influence their learning and educational experience has been part of school improvement programmes across many jurisdictions and systems (Bourke and Loveridge, 2018). The idea that students should be involved in discussions regarding assessment is not new; it is a central principle of formative assessment, fostering students' engagement and active participation in their learning (Hayward, 2013). However, what is new from this research is the strength of concern from students about their lack of involvement in higher level national decisions regarding changes to key summative assessments that determine their life chances and future opportunities. As we have argued above, the same cannot be said for including students directly within policy making around educational assessment, or national and international assessment systems generally. In this respect, we argue that

students are policy actors, but they are also the “acted upon”, students shape educational assessment policies by their interaction with them within their own contexts but they are also being shaped by assessment policy through assessment opportunities afforded to them and, ultimately, the outcomes they receive.

Student voices on assessment – reform, policy and experiences

From the research studies carried out within UK and Ireland (referenced above), several key themes or messages emerged that allowed a much fuller understanding of the lived experiences of students as they navigate through national assessments and end of compulsory schooling examinations. These messages include students as assessment policy actors, and the importance of listening to young people, to system level issues including equality and ethics. There were similarities between students’ experiences across all the jurisdictions – an interesting point in itself. What these studies show is that students are rarely consulted about what they think about major assessment reform, but when they are, they indicate that assessment is a powerful motivator for them in providing clear indicators of areas of strength and expertise allowing for evidence to support decisions about next steps. However, it can also dominate the final years of their educational careers and create pressure and anxiety generally in their lives that extends well beyond the school gates.

Learning from listening to young people

Assessment and examinations (also referred to in the UK as qualifications) are seen as some of the most important aspects of students’ education, dominating their lives at the end of compulsory schooling but recognized as the educational currency needed for the next phases of their lives post school. A key message from the research studies cited above is the balance between the purpose of assessment policy and exam systems, and over-assessment which impinges on young people’s mental health and detracts from other aspects of their lives as young people. As [Elwood \(2012, 2013a\)](#) and [Elwood et al. \(2017\)](#) indicated, good assessment outcomes are very much linked to aspirations and goals for access to tertiary education, jobs, apprenticeships and whatever comes next, but they also need to be maximized in present day contexts of high competition for jobs and top university places as well as difficult and fluid economic climates, not just locally but throughout the world. Students show clear indications that assessments and, specifically, examinations, were necessary and that they understand the point of such systems, especially in providing fair and commonly understood mechanisms for selection into the next phases of study or employment. The research studies discussed were obviously carried out in those jurisdictions that have a long history of national examinations systems at key stages of education which are also clearly embedded in the culture of education in these countries.

In recent years assessment has been commonly aligned with concerns about students’ well-being and mental health ([Putwain et al., 2015](#)). Under such considerations, there have been calls for national assessments and examinations to be removed in favor of more teacher-based and continuous assessment ([Roome and Soan, 2019](#)). However, for those students who live within systems of high-stakes examinations, not all assessment is damaging or unwanted. In some jurisdictions ([Elwood 2012, 2013](#); [Barrance and Elwood, 2018b](#)) students have indicated that examinations can be a good experience if they are well-prepared and that they can influence learning in affirmative ways, with students being allowed to orchestrate, and take responsibility for, their own success, beyond that decided by their teachers. Such systems of national assessment also provide opportunities for students to evidence dealing successfully with pressurized situations, as well as showing that effort spent in learning and preparing for examination sessions can improve outcomes significantly (and vice versa) ([Baird et al., 2016](#)).

However, assessments and examinations can also be some of the most daunting activities that students have to participate in through their school life. At major crunch points in students’ school careers when examinations and national assessments take place, they can take over all other activities. Moreover, overt (and hidden) messaging from teachers and school leaders about doing well can dominate students’ interactions in schools with negative scenarios of what might transpire if students’ full potentials are not reached through assessment outcomes. Research has also indicated concerns about the amount of testing and assessment that young people encounter in their school careers, with this having increased significantly, especially within post-primary settings, and observation that students do not go for very long within a school year without formal assessments and/or tests. This over assessment has particularly been compounded by those assessment systems that are modular in structure, with students taking a range of formal assessments at fixed times in the year as well as availing of re-sit opportunities ([Pinot de Moira et al., 2020](#)). Indeed, the UK has been criticized in the past for over assessing its school children and having some of the most assessed children in Europe ([House of Commons, 2008](#)). Recent calls for new benchmarks on standards for primary school children in England continue to reflect policies that are outcomes driven ([Weale, 2022](#)).

The over assessment of students in UK schools is a real cause for concern, especially when it starts to interfere with extracurricular activities and comes to dominate what is taught and how it is taught (teaching to the test) in order for schools, rather than students, to be as successful as they can be. The dominating context of assessments and national examinations at particular stages of education for students can limit their extra-curricular activities, impacting on their opportunities to become “rounded individuals” and therefore more attractive to educational institutions looking for “balanced individuals” and “involved citizens” ([Elwood, 2012, p. 505](#)). So, while students see the value of assessments and qualifications, they have shown that often assessment policy reform initiatives do not consider the wider picture of the amount of assessment that can take place when particular models are introduced. Some reforms actually increase the amount of formal assessment that takes place with the rest of students’ school activities suffering as a consequence ([Baird et al., 2011](#)).

Young people as assessment policy actors

Like other educational agendas, assessment reform is no less political. In fact, it is constantly used to drive political agendas (Isaacs and Lamprianou, 2018) and, as we have argued above, not always with children's best interests at the forefront of decision making. There is an important message, therefore, in consulting youth action groups as key assessment policy actors. Often, students see themselves with very little political clout as they do not have the right to vote, and assessment reform, like other reforms, is something that is done to them rather than with them by governments who they may or may not have voted for if enfranchised. Political movements over recent years have called for the lowering of the voting age for young people (Runciman, 2021). Indeed, in the devolved nations of the UK (England, Northern Ireland, Scotland and Wales) there is variation in the articulation of such a move, but there have been changes in the degree to which children are consulted about local-national issues and in them having more of a say in devolved politics. Students' movements across nations are becoming more organized in relation to many school issues, and assessment is high amongst their concerns (see Youth Assemblies in NI, Scotland and Wales for examples). These movements take the form of student parliaments and/or assemblies, students' unions as well as action groups that focus on issues that are of importance to students and determined by them. The COVID-19 pandemic impacted significantly on education systems around the world (Peleg et al., 2021) with subsequent major changes to assessment systems and delivery of end of compulsory school outcomes. Subsequently, more organized student groupings have emerged through and post-pandemic that are engaged more directly with politicians, showing that governments should be interested in students' experiences and should be asking them about major changes envisaged for education in post pandemic societies.

Students in the studies carried out in the UK and Ireland (ibid.) between 2010 and 2017, were clear even then that they should be asked for their views as they are the ones most affected by change, acknowledging that often assessment policy reform is enacted and experienced very differently for different students within localized contexts which also has implications of equality of opportunities afforded to students. Students within their local contexts form learner and assessment identities in relation to how policies are enacted within the local settings of schools: a student is "governed by others and at the same time is governor of him/herself" (Ball, 2015, p. 310). As Elwood (2012, 2013a) has previously argued, this positions students as both national subjects in assessment policy regimes but also as local policy actors, capable of influencing their own lives and the development of their educational identities through localized assessment processes and regulations.

System level issues

A major consideration for students, and key message when asked about policy reform in assessment, is the constant change to systems without consultation with them. Significant assessment reforms, across many jurisdictions are continually made by governments without the deliberate participation of students on stakeholder committees or focus groups that are often established at times of consultation. As discussed above, this is despite governments being under the auspices of Article 12 UNCRC (Lundy, 2007); they are duty bound to consult with young people in all matters that affect them directly in policy directives, and educational assessment should be no exception (Elwood and Lundy, 2010). The main concern about not being consulted about system level changes is that usually the change is brought in "live", i.e., it is changed while students are in the system and often without piloting, so that systems are "tried out" on students, yet it is their grades/results that are affected if the system is flawed. There is often unclear articulations of the ideologies and epistemologies that underpin assessment reform, with students being impacted by the whims of politicians and policy makers to implement their own versions of what they think good assessment should be.

Other key system level factors that students consider as problematic are issues of comparability, with questions about the year-on-year, inter-cohort value of qualifications. Often assessment reform can bring about change for the good, but there are questions to address with various stakeholders about what these new assessments are equivalent to, and how they relate to those that have gone before. Students, who are on the receiving end of taking new qualifications, have argued that there is often no safety-net for them, that they end up with qualifications that are either misunderstood, or not understood as valuable as those qualifications that have gone before because they are new and untested. There is a clear sense that assessment reform can often destabilize well-established assessment systems, because of the desires of politicians to change systems to suit immediate political needs. Students are aware that there are degrees of uncertainty built into new systems, but it is often hard for teachers and students to understand what is expected of new systems, syllabuses and courses, to the extent that policy enactment of new systems at local levels introduces an extra layer of uncertainty and instability that can impact disastrously on students by providing outcomes significantly out of sync with those expected.

Equality (and ethical) implications at the system level

A further important key message that emerges from students is that of fair assessment and equality in policy formation. Equality is a dimension of assessment that has been closely considered within the field, more so perhaps than epistemologies of rights and/or childhood studies (Barrance 2020; Elwood, 2005; Nisbet and Shaw 2020). In the studies that have listened to students' views and experiences, fair assessment was a constant factor raised by students in any considerations for change of systems and/or practices in the assessment arena. Systems of national assessment should encompass different types of assessment techniques so all students can show themselves to good effect. Systems that solely rely on end of courses examinations, it is argued, suit only those students who thrive under pressure or who are good at that type of assessment. Other ways to show learning and competence (such as coursework, practicals, etc.) should always be part of national assessment systems.

Issues around equality are also very closely aligned to issues of ethics and ethical standards in assessment policy reform, yet the ethical implications of what is done to students in the name of reform is rarely studied (Elwood, 2013b; Hamp-Lyons 1997; Oates,

2007; Shohamy, 2001). It is not always clear whether decisions that are made about changes to assessment or the implementation of new assessment systems are considered in the context of whether these decisions bring about good things and in ethical terms, do no harm to those on the receiving end of the changes (Warnock, 1998). Oates (2007) has outlined a range of ethical issues that, he suggests, are associated with assessment policy changes and reform. Of major concern to any system change, he argues, is whether there is misuse of assessment systems for political ends and any significant instability of established systems through mass change with resultant consequences (especially if negative) to the integrity of assessment systems. Key messages from the research with students tended to reinforce those concerns considered by Oates (2007) in that students were aware of the risks involved (to them) of one choice of assessment system and its constituent parts over another, a long side a general lack of a clear rationale, on the part of policy makers shared with students, as to why these choices are made. The adverse consequences of "live" changes to assessment systems becomes another significant source of inequalities in schooling experienced by children and young people, over which they have no control.

Discussion and reflections

This chapter has brought together debates about assessment policy reform and students' voices. Considered separately, they are significant areas of research with their own literatures and underpinning theories. However, these two areas of interest are rarely considered in the same research frame. Yet, as we have shown, assessment is a key aspect of education for young people and their experiences of "doing" assessment and "being" assessed often occur at stages in their educational careers when assessment becomes "most intense and most critical in its impact on lives" (Lumby and Foskett, 2005, p. 89). Listening to students and consulting them on their views about one of the most important aspects of education they encounter has shown to provide insights and perspectives that have much to inform those who are responsible for devising assessment policy as well as those who develop the assessments themselves.

The rise of the student voice movement in education has made the inclusion of students' views in research policy and practice more commonplace (Bourke and Loveridge, 2018), and mainstreamed the acceptance of having students' reflections on their own learning integrated into school level policies, so much so that the possibilities of including student voice in change agendas in school has shown great promise for radical change. Cook-Sather (2002, p. 10) has suggested that through listening to students more systematically and deliberately, "students feel more engaged but are also more inclined to take responsibility for their education [as] it is no longer something being done to them but rather something they do". However, the involvement of children and young people in educational research, policy and practice is so much more complex than some of the work in the field might otherwise suggest. To this end, students across many jurisdictions often indicate that they feel included to a certain degree in shaping certain aspects of their educational experience locally (Bourke and Loveridge, 2018), what (Fleming, 2013) has termed the "private spheres" of institutional decision-making, but feel excluded from higher-level decision making at regional and national levels and assessment policy reform is a prime example of a higher level sphere of decision making from which students feel excluded. As we have indicated, assessment reform can be a "high risk" endeavor, often controversially used to change national systems to advance political agendas. In these high-risk contexts, very little consultation ends up taking place, or is conducted but ignored, to suit political ends. It is in these contexts that students become disempowered in influencing any big decisions and their participation only tolerated in low risk (to adults) domains (Elwood, 2013a) once bigger decisions have been made.

What the research on assessment reform and students' voice evidences is that young people with the capacity to make considered judgements do not automatically opt for what may be perceived as the easiest option. It suggests that many of the arguments made against consulting young people on assessment reform are based on unfair assumptions about what young people will say or the value of their contributions to high-risk policy domains such as assessment systems and their reform. Indeed, Cook-Sather (2006) talks of a "cultural shift" that needs to take place that authenticates the processes of consultation with, and participation of, young people. We are reminded of the need for this "cultural shift" if authentic engagement with students is to succeed in assessment reform. A cultural shift would see the implementation of processes that normalize more informed dialogue between students and policy makers in assessment to realign the balance of power and ultimately challenge, in constructive ways, the sharing of authority about assessment issues (Cook-Sather, 2002; Elwood, 2012).

Elwood (2013a) has argued previously that for direct influence of students' views on assessment reform to become routine, the cultural shift needs to also occur not least within test and assessment development agencies and ministries of education responsible for national assessment systems. This would mean the embedding of consultation with, and participation of, young people within these organizations' decision-making processes as well as their governance structures, to actively support the diverse voices of students' being heard on the range of their activities as well as them being listened to around deliberations on system change and informed of any subsequent actions. The continued absence of student voice in assessment reform either within these organizations or within policy-making domains perpetuates the notion of students being "voiceless" in some of the major educational policy areas that impact on them. Mitchell et al. (1997, p. 853) also talk about the salience of stakeholders, in that they possess "one or more of three relationship attributes: power, legitimacy, and urgency". It would seem that the salience of young people as powerful or legitimate or urgent stakeholders in assessment reform, under Mitchell et al.'s definition, is limited; their voices are not routinely heard or acted upon and their power to influence is restricted. Elwood (2013a) has further suggested that while policy makers continuously reform systems in pursuit of ideological and political agendas, students become "policy victims", suffering all the adverse consequences of unreliable and possibly unethical systems.

In this chapter, we have focused on what Bourke and Loveridge (2018) have called the “promises and possibilities of including student voice in change agendas”. We are convinced that embedding students’ voices in change agendas for assessment reform will ultimately realize better and fairer quality educational provision for all students across ages and stages of schooling, as well as across diverse groups of students. As Elwood (2013a, p. 109) suggests “by authorizing student voice in this area, policy-makers and [assessment] developers must be prepared to share authority about ... assessment matters with students and to rethink who is authoritative about these aspects of educational life”. Only then will students be positioned as consultants within assessment regimes in order that their perspectives can improve our knowledge of fundamental assessment practices and provide for us a fuller understanding of those elements that will promote high quality, fair assessment.

References

- Alcoff, L., 1991/1992. The problem of speaking for others. *Cult. Critiq.* 20, 5–32.
- Arnot, M., 2008. Public policy, governance and participation in the UK: a space for children? *Int. J. Child Right.* 16, 355–367.
- Arnot, M., Reay, D., 2004. The framing of pedagogic encounters: regulating the social order of classroom learning. In: Muller, J., Davies, B., Morais, A. (Eds.), *Reading Bernstein, Researching Bernstein*. Routledge; Falmer, London, pp. 137–152.
- Arnot, M., McIntyre, D., Pedder, D., Reay, D., 2004. *Consultation in the Classroom: Developing Dialogue About Teaching and Learning*. Pearson, Cambridge, UK.
- Baird, J., Caro, D., Hopfenbeck, T.N., 2016. Student perceptions of predictability of examination requirements and relationship with outcomes in high-stakes tests in Ireland. *Ir. Educ. Stud.* 35 (4), 361–379.
- Baird, J., Elwood, J., Duffy, G., Feiler, A., O’Boyle, A., Rose, J., Stobart, G., McWhirter, A., 2011. 14–19 Centre Research Study: Educational Reforms in Schools and Colleges in England. QCDA, London.
- Baird, J., Hopfenbeck, T.N., Elwood, J., Caro, D., Ahmed, A., 2015. Predictability in the Irish Leaving Certificate. Oxford University and Queen’s University, Oxford and Belfast.
- Baird, J., Lee-Kelley, L., 2009. The dearth of managerialism in implementation of national examinations policy. *J. Educ. Pol.* 24, 55–81.
- Ball, S.J., 2015. What is policy? 21 years later: reflections on the possibilities of policy research. *Discourse – Studies in the cultural politics of Education* 36 (3), 306–313.
- Ball, S.J., Bowe, R., 2020. The neoliberalization of the state, the processes of “fragmentation”, and research implications of the new political terrain of English schooling. In: Brown, A., Wisby, E. (Eds.), *Knowledge, Policy and Practice in Education and the Struggle for Social Justice: Essays Inspired by the Work of Geoff Whitty*. UCL Press, London, pp. 97–114.
- Ball, S., Maguire, M., Braun, A., 2012. *How schools do policy: policy enactments in secondary schools*. Routledge.
- Banks, J., Smyth, E., 2015. “Your whole life depends on it”: academic stress and high-stakes testing in Ireland. *J. Youth Stud.* 18, 598–616.
- Baroutsis, A., McGregor, G., Mills, M., 2016. Pedagogic voice: student voice in teaching and engagement pedagogies. *Pedagog. Cult. Soc.* 24 (1), 123–140.
- Barrance, R., 2020. Tiering in the GCSE: a children’s rights perspective. *Br. Educ. Res. J.* 46 (6), 1210–1231.
- Barrance, R., Elwood, J., 2018a. Inequalities and the curriculum: young people’s views on choice and fairness through their experiences of curriculum as examination specifications at GCSE. *Oxf. Rev. Educ.* 44 (1), 19–36.
- Barrance, R., Elwood, J., 2018b. National assessment policy reform 14–16 and its consequences for young people: student views and experiences of GCSE reform in Northern Ireland and Wales. *Assess Educ. Princ. Pol. Pract.* 25 (3), 252–271.
- Black, P., 2015. Formative assessment – an optimistic but incomplete vision. *Assess Educ. Princ. Pol. Pract.* 22 (1), 161–177.
- Black, P.J., William, D., 1998a. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5, 7–74.
- Black, P., William, D., 2003. “In praise of educational research”: formative assessment. *Br. Educ. Res. J.* 29, 623–637.
- Bourke, R., Loveridge, J., 2014. Student voice(s) as a pedagogical tool. In: St George, A., Brown, S., O’Neill, J. (Eds.), *Facing the Big Questions in Teaching: Purpose, Power and Learning*, second ed. Cengage Learning, Melbourne, VIC.
- Bourke, R., Loveridge, J., 2018. Using student voice to challenge understandings of educational research, policy and practice. In: Bourke, R., Loveridge, J. (Eds.), *Radical Collegiality through Student Voice*. Springer Pvt Ltd, Singapore.
- Bragg, S., 2007. “Student voice” and governmentality: the production of enterprising subjects? *Discourse* 28 (3), 343–358.
- Brown, G.T.L., Hirschfeld, G.H.F., 2008. Students’ conceptions of assessment: links to outcomes. *Assess Educ. Princ. Pol. Pract.* 15 (1), 3–17.
- Cadwallader, S., Tremain, K., 2013. How Policy Formation and Implementation Interacts With Risks to High Stakes Qualifications. Retrieved from CERP. https://cerp.aqa.org.uk/sites/default/files/pdf_upload/CERP_RP_SMC_25102013.pdf.
- Chamberlain, S., 2012. Public perceptions of reliability. In: Opposs, D., He, Q. (Eds.), *Ofqual’s Reliability Compendium* Coventry: Office of Qualifications and Examinations Regulation, pp. 691–725.
- Charteris, J., Smardon, D., 2019a. Democratic contribution or information for reform? Prevailing and emerging discourses of student voice. *Aust. J. Teach. Educ.* 44 (6), 1–18.
- Charteris, J., Smardon, D., 2019b. Student voice in learning: instrumentalism and tokenism or opportunity for altering the status and positioning of students? *Pedagog. Cult. Soc.* 27 (2), 305–323.
- Cook-Sather, A., 2002. Authorizing student perspectives: toward trust, dialogue, and change in education. *Educ. Res.* 31 (4), 3–14.
- Cook-Sather, A., 2006. Sound presence, and power: “student voice” in educational research and reform. *Curric. Inq.* 36 (4), 359–390.
- Cook-Sather, A., 2010. Students as learners and teachers: taking responsibility, transforming education and redefining accountability. *Curric. Inq.* 40 (4), 555–575.
- Cook-Sather, A., 2014. The trajectory of student voice in educational research. *N. Z. J. Educ. Stud.* 49 (2), 131–148.
- Cook-Sather, A., 2018. Tracing the evolution of student voice in educational research. In: Bourke, R., Loveridge, J. (Eds.), *Radical Collegiality Through Student Voice: Educational Experience, Policy and Practice*. Springer Singapore Pvt Limited.
- Cook-Sather, A., 2020. Student voice across contexts: fostering student agency in today’s schools. *Theory Into Pract.* 59 (2), 182–191.
- Daly, A., Baird, J., Chamberlain, S., Meadows, M., 2012. Assessment reform: students’ and teachers’ responses to the introduction of stretch and challenge at A-level. *Curric. J.* 23, 139–155.
- DeLuca, C., Chapman-Chin, A., LaPointe-McEwan, D., Klinger, D., 2018. Student perspectives on assessment for learning. *Curric. J.* 29 (1), 77–94.
- Elwood, J., 2005. Gender and achievement: what have exams got to do with it? *Oxf. Rev. Educ.* 31 (3), 373.
- Elwood, J., 2012. Qualifications, examinations and assessment: views and perspectives of students in the 14–19 phase on policy and practice. *Camb. J. Educ.* 42 (4), 497–512.
- Elwood, J., 2013a. The place(s) of student voice in 14–19 education policy reform: reflections on consultation and participation. *Lond. Rev. Educ.* 11 (2), 97–111.
- Elwood, J., 2013b. Educational assessment policy and practice - a matter of ethics. *Assess Educ. Princ. Pol. Pract.* 20 (2), 205–220.
- Elwood, J., Hopfenbeck, T., Baird, J., 2017. Predictability in high-stakes examinations: students’ perspectives on a perennial assessment dilemma. *Res. Pap. Educ.* 32 (1), 17.
- Elwood, J., Lundy, L., 2010. Revisioning assessment through a children’s rights approach: implications for policy, process and practice. *Res. Pap. Educ.* 25 (3), 335–353.
- Fielding, M., 2001. Students as radical agents of change. *J. Educ. Change* 2 (3), 123–141.
- Fielding, M., 2004. Transformative approaches to student voice: theoretical underpinnings, recalcitrant realities. *Br. Educ. Res. J.* 30 (2), 295–311.
- Fleming, J., 2013. Young people’s participation – where next? *Child. Soc.* 27, 484–495.

- Flórez-Petour, M.T., 2015. Systems, ideologies and history: a three-dimensional absence in the study of assessment reform processes. *Assess Educ. Princ. Pol. Pract.* 22 (1), 3–26.
- Flutter, J., Ruddock, J., 2004. *Consulting Pupils: What's in it for Schools?* Routledge; Falmer, London.
- Hamp-Lyons, L., 1997. Ethics in language testing. In: Clapham, C., Corson, D. (Eds.), *Encyclopedia of Language and Education, Language Testing and Assessment*, 7. Kluwer Academic, Dordrecht, pp. 323–333.
- Hayward, L., 2013. From why to why not? The conundrum of including learners' perspectives: a response to this special issue. *Lond. Rev. Educ.* 11, 184–189.
- Hinton, R., Tisdall, E.K., Gallagher, M., Easley, S., 2008. Children's and young people's participation in public decision-making. *Int. J. Child Right.* 16, 281–284.
- Hodgson, A., Spours, K., 2003. Beyond a Levels: Curriculum 2000 and the Reform of 14–19 Qualifications. Kogan Page, London.
- Hodgson, A., Spours, K., 2005. The learner experience of curriculum 2000: implications for the reform of 14–19 education in England. *J. Educ. Policy* 20, 101–118.
- Hopfenbeck, T.N., Stobart, G., 2015. Large-scale implementation of assessment for learning. *Assess Educ. Princ. Pol. Pract.* 22 (1), 1–2.
- House of Common, 2008. Children, Schools, and Families Committee (HoC) 2008. *Testing and Assessment: Third report of session 2007–08 – Volume I*, HC 169–I. The Stationery Office, London.
- Isaacs, T., Lamprianou, I., 2018. International assessment policy reform: nothing new under the sun. *Assess Educ. Princ. Pol. Pract.* 25 (3), 227–229.
- Koretz, D., 2008. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, Cambridge, MA.
- Lingard, B., Martino, W., Reza-Rashti, G., Sellar, S., 2016. *Globalising Educational Accountabilities*. Routledge, Abingdon and Oxon.
- Lumby, J., Foskett, N., 2005. *14–19 Education: Policy, Leadership and Learning*. Sage, London.
- Lundy, L., 2007. "Voice" is not enough: conceptualising article 12 of the United Nations convention on the rights of the child. *Br. Educ. Res. J.* 33 (6), 927–942.
- Lundy, L., 2018. In defence of tokenism? Implementing children's right to participate in collective decision-making. *Childhood* 25 (3), 340–354.
- Lundy, L., Cook-Sather, A., 2016. Children's rights and student voice. In: Wyse, D., Hayward, L., Pandya, J. (Eds.), *The Sage Handbook of Curriculum, Pedagogy and Assessment*. Sage, London.
- Madaus, G., 1988. The influence of testing on the curriculum. In: Tanner, L.N. (Ed.), *Critical Issues in Curriculum*, 87th Yearbook of NSSE Part 1. University of Chicago Press, Chicago, IL, pp. 83–121.
- Martens, K., Jakobi, A.J., 2010. Expanding and intensifying governance: the OECD in education policy. In: Martens, K., Jakobi, A.P. (Eds.), *Mechanisms of OECD Governance: International Incentives for National Policy-Making?* Oxford University Press, Oxford, pp. 163–179.
- Mitchell, R.K., Bradley, R., Agle, B.R., Wood, D.J., 1997. Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. *Acad. Manag. Rev.* 22 (4), 853–886.
- Mitra, D., 2018. Student voice in secondary schools: the possibility for deeper change. *J. Educ. Adm.* 56 (5), 473–487.
- Nisbet, I., Shaw, S., 2020. *Is Assessment Fair?* Sage, London.
- Noyes, A., 2005. Pupil voice: purpose, power and the possibilities for democratic schooling. *Br. Educ. Res. J.* 31 (4), 532–540.
- Oates, T., 2007. Protecting the innocent – the need for ethical frameworks within mass educational innovation. In: Saunders, L. (Ed.), *Educational Research and Policy Making*. Routledge, London, pp. 146–174.
- Ozga, J., 2000. *Policy Research in Educational Settings: Contested Terrain*. Open University Press, Buckingham, UK.
- Peleg, N., Lundy, L., Stalford, H., 2021. COVID-19 and children's rights: space for reflection, tracing the problems, and facing the future. *Int. J. Child Right.* 29, 255–259.
- Percy-Smith, B., 2006. From consultation to social learning in community participation with young people. *Child. Youth Environ.* 16 (2), 153–179.
- Pinot de Moira, A., Meadows, M., Baird, J., 2020. The SES equity gap and the reform from modular to linear GCSE mathematics. *Br. Educ. Res. J.* 46 (4), 421–436.
- Putwain, D., Daly, A., Chamberlain, S., Sadreddini, S., 2015. Academically buoyant students are less anxious about and perform better in high-stakes examinations. *Br. J. Educ. Psychol.* 85 (3), 247–263.
- Reay, D., 2006. I'm not seen as one of the clever children: consulting primary school pupils about the social conditions of learning. *Educ. Rev.* 58, 171–181.
- Robinson, C., Taylor, C., 2007. Theorizing student voice: values and perspectives. *Improv. Sch.* 10 (1), 5–17.
- Roome, T., Soan, C., 2019. GCSE exam stress: student perceptions of the effects on wellbeing and performance. *Pastor. Care Educ.* 37 (4), 297–315.
- Ruddock, J., Fielding, M., 2006. Student voice and the perils of popularity. *Educ. Rev.* 58, 219–231.
- Runciman, D., 2021. *Votes for Children! Why We Should Lower the Voting Age to Six*. The Guardian Online.
- Sellar, S., Lingard, B., 2014. The OECD and the expansion of PISA: new global modes of governance in education. *Br. Educ. Res. J.* 40 (6), 917–936.
- Shohamy, E., 2001. *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. Pearson, London.
- Taylor, C., Robinson, C., 2009. Student voice: theorizing power and participation. *Pedagog. Cult. Soc.* 17 (2), 161–175.
- Taylor, S., Henry, M., Lingard, B., Rizvi, F., 1997. *Educational Policy and the Politics of Change*. Routledge, Oxford.
- Tisdall, E.K., 2015. Addressing the challenges of children and young people's participation: considering time and space. In: Gal, T., Duramy, B. (Eds.), *International Perspectives and Empirical Findings on Child Participation: From Social Exclusion to Child-Inclusive Policies*. Oxford University Press, Oxford, pp. 381–404.
- Tisdall, E., 2008. Is the honeymoon over? Children and young people's participation in public decision-making. *Int. J. Child Right.* 16, 419–429.
- Thomson, P., 2011. Coming to terms with "voice". In: Czerniawski, G., Kidd, W. (Eds.), *The International Handbook of Student Voice*. Emerald, Bingley, pp. pp19–30.
- Thomson, P., Gunter, H., 2006. From "consulting pupils" to "pupils as researchers": a situated case narrative. *Br. Educ. Res. J.* 32 (6), 839–856.
- Weale, S., 2022. *Plans for English School's Include National Behaviour Survey*. <https://www.theguardian.com/education/2022/mar/28/plans-for-englands-schools-include-national-behaviour-survey>. (Accessed 29 March 2022).
- United Nations, 2001. UN Committee on the Rights of the Child (CRC). General comment No. 1. Article 29 (1), The aims of education. CRC/GC/2001/1. Geneva, United Nations.
- Warnock, M., 1998. *An Intelligent Person's Guide to Ethics*. Gerald Duckworth, London.

Measuring what we cannot see: assessing extreme populations in international large-scale assessments

Leslie Rutkowski^{a,b} and David Rutkowski^{a,b}, ^a Indiana University, Bloomington, IN, United States; and ^b Center for Educational Measurement, University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

A brief overview of item response theory	131
Extreme proficiency distributions	132
Discussion	133
Acknowledgments	134
References	135

International large-scale assessments (ILSAs) of educational achievement such as the Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment (PISA) have established themselves as an important part of the educational research and policy landscape. For example, achievement rankings that compare performance of participating countries are fodder for newspaper headlines and are discussed by national, state, and local politicians in most participating countries. The reforms that have been attributed to these studies continue to have a far and lasting reach (Sellar et al., 2017). Reforms directly attributable to PISA in Germany alone included nation-wide education standards and a four-billion euros expansion of all-day schooling, according to the German Federal Ministry of Education and Research. In the past decade, based on PISA results, the Organization for Economic Co-operation and Development (OECD) has strongly advocated for the revamping of the US educational system, arguing that the adoption of the Common Core State Standards will most likely result in increased PISA scores (OECD, 2013; PBS, 2013). More recently, Estonia has embraced educational reforms spurred PISA aspirations, with associated achievement that places this small Baltic country as the new “reference country” in the European Union (Volante and Klingler, 2021).

With steady growth in participation since its inception in 2000, PISA represents the largest ILSA in terms of participating educational systems.¹ For example, 43 educational systems participated in the inaugural PISA cycle in 2000. By contrast, the most recent completed PISA cycle included 85 educational systems in 2022 (see also Sherer and Nilsen, 2023 this encyclopedia). Although PISA is normally on a three year cycle, due to the Covid-19 pandemic, the usual PISA cycle was delayed by one year. Commensurate with PISA’s growth is a marked increase in participant diversity, where differences are due to language, culture, geography, and economic development. For example, Table 1 summarizes the number of PISA participating educational systems, students’ average performance, and the gross domestic product per capita (GDP/capita) in each PISA cycle by OECD members and non-OECD systems, termed *partner* systems. Notable from Table 1 is that among OECD countries, for whom PISA was originally designed, average GDP/capita was US \$38,649 in 2017 (the most recent year reported). This is nearly three times the average of partner systems, who had an average GDP/capita of US \$13,330 in the same year.

Table 1 OECD and partner educational systems that participated in PISA.

OECD countries	2000 ²	2003	2006	2009	2012	2015	2018
Number of participating countries	28	30	30	34	34	35	40
Mathematics score	N/A	499	494	495	494	490	489
Reading score	492	494	489	493	496	493	487
Science score	N/A	N/A	498	501	501	493	489
GDP per capita (USD)	22,880	27,814	37,005	36,900	40,509	37,133	38,649
Partner countries	2000	2003	2006	2009	2012	2015	2018
Number of participating countries	15	11	27	31	31	37	39
Mathematics score	N/A	447	437	437	451	433	431
Reading score	416	440	414	431	448	430	423
Science score	N/A	N/A	429	437	449	439	430
GDP per capita (USD)	10,288	14,138	14,506	14,823	22,233	15,196	13,330

¹Because ILSA participants can be non-national jurisdictions (e.g., cities, areas with contested status like Palestine, and parts of countries like the Flanders region of Belgium), we refer to all participating systems, national and otherwise, as *participants*.

²The original PISA 2000 administration included 27 OECD countries and 5 partner countries. The numbers reflected here include additional countries that took part in a second administration of PISA 2000 in 2001 and 2002.

Such a heterogeneous collection of participating educational systems poses challenges in terms of deciding what should be measured and how to measure it in a comparable way. To that end, carefully developed assessment frameworks for PISA and similar studies (e.g., TIMSS) guide this process. From developing an internationally recognized definition of the content domains (International Association for the Evaluation of Educational Achievement, 2013; OECD, 2017a) to the complex process of instrument translation (OECD, 2014), cross-cultural considerations loom large. But regardless of the care with which a study such as PISA or TIMSS is designed, a cross-cultural measurement paradox exists: “The larger the cross-cultural distance between groups, the more likely cross-cultural differences will be observed, but the more likely these differences may be influenced by uncontrolled variables” (van de Vijver and Matsumoto, 2011, p. 3). In the context of ILSAs, this issue is particularly germane, given recent additions to PISA, which include Cambodia and Guatemala. With an influx of educational systems, ILSA achievement heterogeneity has also increased. In PISA 2000, for example, eight countries scored one standard deviation below the OECD average of 500 ($SD = 100$) in mathematics literacy while in PISA 2018, that number grew to 12 participating systems at least a standard deviation below the historic OECD average (OECD, 2019a). Similar findings are present in TIMSS. For example, in 1995 three countries fell one standard deviation below the TIMSS international average in mathematics at grade eight, compared to six countries in 2019 (Mullis et al., 2020).

The OECD, aware for some time of the widening proficiency levels between participants, has attempted to make various accommodations for lower performing countries. For example, in 2009 the OECD included sets of easy booklets into its design to help better assess lower achieving countries (OECD, 2010). Enthusiasm for this innovation was substantial, with 20 educational systems opting for this option (OECD, 2012). However, as shown in Rutkowski et al. (2018) the approach of including easy booklets did little to improve measurement when compared to administering a common test for all participants. In other words, although the OECD attempted to accommodate for low performing countries the accommodation did not seem to improve measurement for low performers.

To develop an assessment more appropriate for lower performing countries, in 2013 the OECD initiated PISA for Development (PISA-D) with a stated purpose “to make the assessment more accessible and relevant to low-to-middle-income countries” (OECD, 2016, p. 2). Designed as a one-off assessment project resulting in scores comparable to PISA, nine countries are members of the PISA-D project including: Bhutan, Cambodia, Ecuador, Guatemala, Honduras, Panama, Paraguay, Senegal and Zambia. Despite efforts to better align the assessment to what students know and can do in these countries, recent research showed that there remains substantial mismatch between the difficulty of the items on PISA-D and participating students (Rutkowski and Rutkowski, 2021). An example of this discrepancy is represented in Fig. 1. The top panel shows the achievement distributions of the participating PISA-D countries whereas the bottom panel shows the distribution of items in terms of their difficulty. A good examinee-item match would be represented by an achievement distribution which is situated over many of the items. In contrast, Fig. 1 shows that

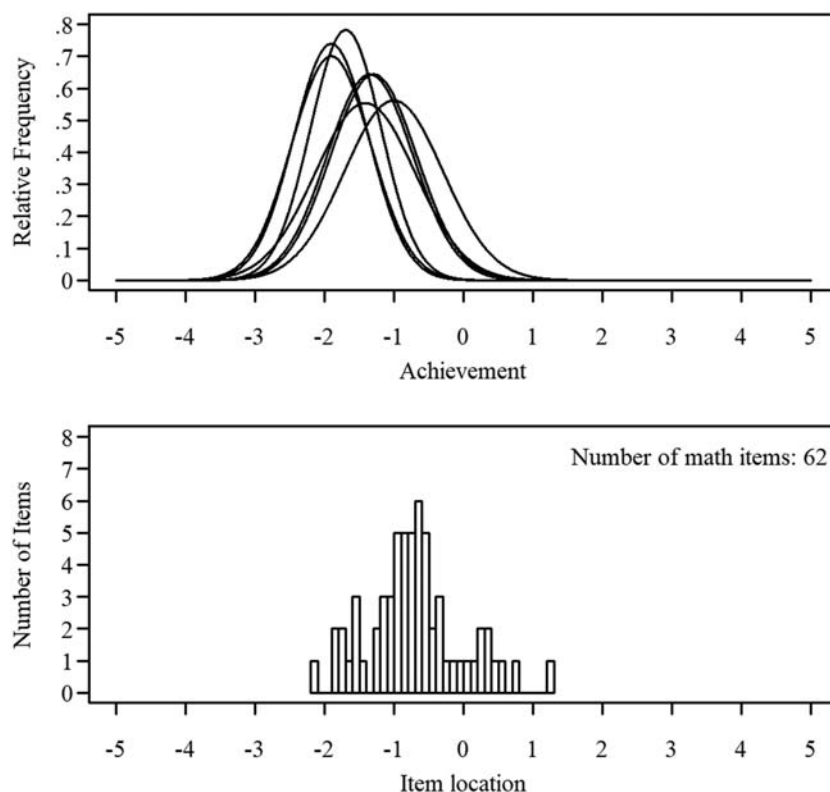


Fig. 1 PISA-D proficiency distributions contrasted against item difficulty distribution.

for all but the highest achieving country, there is little overlap between achievement distributions and item difficulty. This result implies that in most PISA-D countries, there are few, if any items that measure these populations. A theoretical consequence of this issue is that measurement precision is reduced in these countries (Hambleton and Swaminathan, 1985). In other words, we have less certainty about what these students know and can do. What is not documented in the literature are limits for *accurately* measuring populations that are markedly different from populations for which the test was designed. Understanding whether and to what degree the current models are able to capture achievement in very low performing countries is an important matter since, as noted above, new comers to PISA are characterized by particularly low performance on this assessment.

A brief overview of item response theory

A comprehensive overview of the models and methods used in PISA is well outside the scope of this chapter; however, to orient our conversation we provide a brief overview of item response theory (IRT), which forms the basis of our subsequent analysis. For an introduction to IRT, readers are invited to consult Embretson and Reise (2000) or Hambleton and Swaminathan (1985). In IRT, test items and examinee proficiency occupy the same scale or metric, which is usually assumed to follow a standard normal distribution with mean = 0 and standard deviation = 1. Then, a probability model is used to describe the relationship between items and examinees, where the probability of a correct response to an item j , $P(x_j = 1)$, depends on an examinee i 's proficiency level (θ_i). One equation that describes this relationship is known as a two-parameter logistic (2PL) model (Birnbaum, 1968) and is given as:

$$P(x_j = 1 | \theta_i, a_j, b_j) = \frac{1}{1 + e^{a_j(\theta_i - b_j)}}, \quad (1)$$

where parameters a_j and b_j are characteristics of item j . In particular, a_j indicates the degree to which item j can discriminate among students of different proficiency levels. Higher values indicate that the item discriminates well. The parameter b_j is the item difficulty, which locates item j along the proficiency continuum and can be interpreted as the proficiency value that corresponds to a 50% chance of a correct answer. Higher values indicate a more difficult item and lower values indicate an easier item. The model states that students with higher levels of θ will have a higher probability of correctly answering items. Fig. 2 illustrates a curve, known as an *item response function* (IRF), which graphically represents the model in Eq. (1). This IRF corresponds to an item with a difficulty parameter of 0 and a discrimination parameter of 1. The x-axis represents both item difficulty and examinee proficiency. Here, we can see that as θ increases, the probability of a correct answer also increases.

Implicit in this representation is that test takers, in fact, are measured by items on the assessment. In other words, on a given assessment, test takers should have some probability of a correct answer on many of the items. Unfortunately, however, this is not the case for the lowest performers. As Rutkowski and Rutkowski (2021) show the probability of a correct response on a typical PISA-D item is just 0.11 for an average student and just 0.06 for lower performers in Zambia. Importantly, for a multiple choice item with the standard four response options, these probabilities are well below chance.

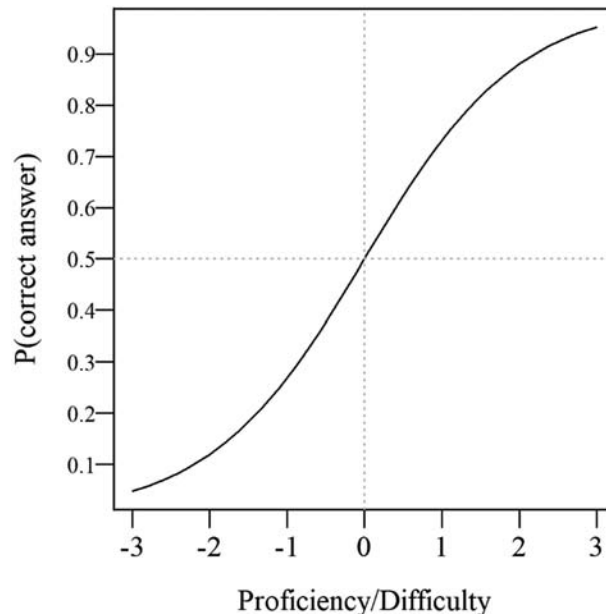


Fig. 2 Item response function for $b = 0$, $a = 1$.

Extreme proficiency distributions

To understand whether the models used in ILSAs like PISA are able to accurately measure extreme proficiency distributions such as those observed in recent cycles of PISA and PISA-D, we conducted a Monte Carlo simulation study. Although our primary interest is on low performers, we also incorporated in our design a population of high performers. As the basis for our simulation, we considered an extreme proficiency mean to be well outside of what has been historically observed in PISA (or PISA-D). In mathematics, Zambia was the lowest performing country in PISA-D with an average score of 258 (OECD, 2018). As an aside, this score also suggests that a typical Zambian student is measured by a single math item. Given that this score is about 2.5 standard deviations below the historic PISA mean of 500 ($SD = 100$), we selected as our extreme achievement distributions the equivalent of 200 and 800. On an IRT scale, which typically uses a standard normal distribution, this translates to the lowest and highest populations having achievement distributions of $N(-3, 1)$ and $N(3, 1)$, respectively. To represent a typically performing population, we also simulated a population of examinees with an achievement distribution of $N(0, 1)$. This implies mean proficiency differences of six standard deviations, with the extreme groups located three standard deviations above or below the middle group. By comparison, in PISA 2018, the lowest achiever, the Dominican Republic, had a mean in science of 336. The highest achiever, the participating regions in China, achieved a mean of 590. Given a historic standard deviation of 100, this suggests a span of 2.54 standard deviations. To that end, our simulation represents an extreme case in terms of cross-country achievement differences.

For our simulation, we used PISA 2015 science item parameters (OECD, 2017b). Although there are both dichotomous (right/wrong) and polytomous (partial credit) items in PISA, we simplified our simulation design and treated all parameters as though they resulted from dichotomous items, by eliminating the additional step parameters associated with a polytomous item. To mimic the PISA design (Rutkowski et al., 2014), we assigned items to 12 blocks of non-overlapping items, which were assembled into 36

Table 2 Simulated incomplete block design.

Form	Item block											
	1	2	3	4	5	6	7	8	9	10	11	12
1	X						X					
2	X									X		
3		X						X				
4			X						X			
5			X									X
6				X			X					
7				X						X		
8					X						X	
9						X						X
10						X	X					
11	X							X				
12					X			X				
13		X							X			
14						X			X			
15			X							X		
16		X									X	
17				X							X	
18					X							X
19							X	X				
20							X		X			
21							X				X	
22								X		X		
23								X				X
24								X	X			
25									X		X	
26							X			X		
27									X	X		
28										X		X
29								X			X	
30									X	X		
31							X					X
32									X			X
33											X	X
34		X		X								
35	X				X							
36			X			X						

different forms. This design included two blocks of items per form. For example, form 1 featured block 1 and 7; form 2 featured block 1 and 10. This design ensured that forms were linked and that each simulated examinee responded to as many as 35 items. Our simulated assessment design is represented in Table 2, where “X” indicates that a block of items occurs in a form.

From the previously specified population proficiency distributions, we sampled 5000 examinees for each group, giving a total sample size of 15,000. Test forms were randomly assigned to examinees and simulated responses were recorded. The simulation and item response routing was completed in R (R Core Team, 2021) using the *lsasim* package (Matta et al., 2018). We then treated item parameters as fixed values in a latent regression model where the group served as a covariate in the model (Mislevy et al., 1992; von Davier and Sinharay, 2014). Using the group with a mean of zero as the reference group, the resulting regression coefficients represent the estimated achievement means for the other two groups. Population achievement distributions were estimated using the TAM package (Robitzsch et al., 2017). We repeated this process 500 times to achieve a stable solution.

The results of the simulation produced achievement means of -3.006 ($SD = 0.0096$) and 3.007 ($SD = 0.0087$) for the low and high performing groups, respectively. This result is essentially perfect recovery of the generating population proficiency values. To give some context to these results, Fig. 3 shows the distribution of item difficulty for these items, with the bulk centered around a difficulty value of zero.

To further contextualize the results, we calculated the probability of a correct response for a *typical* examinee in the low group—one who has a proficiency value of -3 . On the easiest item (difficulty value $= -2.403$), the probability of a correct response was calculated as 0.15. The probability fell markedly for an item of typical difficulty (difficulty value $= 0.054$) and was calculated as 0.0005. Finally, the most difficult item (difficulty value $= 3.327$) produced a probability of correct response equal to 0.0001, or nearly a 1 in 10,000 chance of a correct answer. Notably, these probabilities are for a middle performing examinee in this group—half will have slightly higher or lower probabilities, depending on their proficiency value. Nevertheless, for most of the examinees in the low group (and to a large extent the high group), most are measured by few or no items. Recall, too, that the assessment design was such that each examinee took just 2 of 12 possible blocks of items, meaning that the likelihood of being measured by no items for a given examinee in either of these groups was high.

Despite poor or non-existent measurement for groups at the extreme ends of the proficiency distribution, the estimation method recovered accurately the underlying proficiency distributions. To our thinking, this points to a strange phenomenon that requires serious attention. The statistical machinery used to estimate population-level achievement distributions can locate a group on the achievement continuum, regardless of how poorly that group is measured. Taking a mechanical perspective on this, the response vectors for the low achieving population are mostly zeros overall and, for many individual examinees, there will be response vectors that are entirely zero. This suggests that we learn *nothing* about what these students know and can do. We only know that they don’t know the material on the assessment.

Discussion

Historically, the OECD developed PISA for economically developed, industrialized member countries. Indeed, the very first group of PISA participants included 28 OECD members and just four partners. In the second round of PISA 2000, conducted several months later, the lowest performer was Peru, a low-income country at that time. Indeed, early PISA documentation and procedures reflected the OECD orientation. For example, the 2000 framework was explicit in that the instruments should reflect OECD member country educational and cultural contexts and the target age of 15 was selected as one where most OECD students were still in compulsory education (OECD, 1999). Further, until the 2015 cycle, the methods used to estimate item parameters used information *only* from OECD countries completely omitting partner countries (OECD, 2014). Only recently was a plan developed to allow partner countries to attain an “associate” membership tier that would allow for full voting rights (Schneider, 2019); however, in 2018 most partner countries still held “observer” status (OECD, 2020) with no right to vote on design, implementation, or other

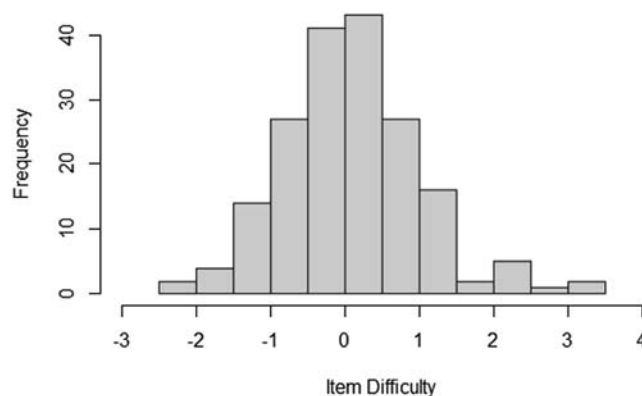


Fig. 3 Item difficulty distribution.

PISA matters. Currently, the expert groups that are responsible for overseeing the development and technical implementation of PISA are heavily represented by researchers, civil servants, and contractors from OECD member countries (OECD, 2020). PISA architects are aware of the privileged position of OECD countries and the possible disadvantage that this confers on partner countries. To that end, the OECD has embraced efforts to more inclusively involve all participants in the process. Nevertheless, the evolution is slow and vestiges of the OECD-centered perspective remain.

In the realm of PISA innovation, practicalities also prevail. For example, the PISA for Development assessment contained no new material to measure participating countries. Rather, the assessment comprised previous PISA content, supplemented with existing material from other international assessments (OECD, 2019c). To that end, the main survey contained 50–60% of trend items from PISA 2015, with the remainder of material drawn from existing surveys (2019b). Although about 60% of the item pool measured the two lowest of eight achievement levels in PISA, previous research showed that most students were poorly measured or not measured at all by PISA-D (D. Rutkowski and Rutkowski, 2021). Nevertheless, our simulation showed that, from a modeling perspective, the statistical machinery used to estimate achievement in PISA and other ILSAs is able to recover the true underlying proficiency distributions for even a very extreme populations.

In our simple simulation we found that current models and methods assure statistical validity on PISA and related ILSAs in that few, if any, test questions need to measure participants to get an accurate estimate of aggregated achievement distributions. Nevertheless, we believe that the statistical results should be decoupled from the validity conversation around our findings. Although it is possible to locate an educational system along an achievement continuum, regardless of the extent to which participants are measured, we are not reasonably able to say what this population knows and can do in a given domain. We return to PISA-D for another example: in Cambodia, Guatemala, Honduras, Paraguay, and Senegal, median math achievement was at level 1b. Students at this level of achievement can follow instructions in simple text and can sometimes perform the first step of a two-step math problem (OECD, 2018). Importantly, just 11 items on the math portion of PISA-D measure at this level or below. Zambia's median math achievement was at level 1c, which is measured by only two items. Students at this level can follow a single, clearly prescribed instruction and perform a single step, such as reading a single value from a chart or table. We can hardly argue that educational outcomes are poor in PISA-D countries given that most participants are below minimum levels of competency. Yet we clearly are not tapping into the knowledge and skills in these populations by measuring with so few items. The reading domain on PISA-D was far better targeted to the tested populations. This portion included a domain called *reading components*, which includes sub-skills or building blocks of reading literacy. A substantial portion of the PISA-D reading items came from the reading components domain 88 of 154 total items and 75% of students received a cluster of reading component items. Importantly, nearly all of these items were located at level 1a or lower, allowing for better measurement at the lower end of the achievement continuum (OECD, 2019c).

With expanding participation in international assessments, where should testing organizations go from here? We highlight several possibilities, which are currently in various stages of implementation. First, computerized assessment platforms offer a promising solution to many of the issues outlined in the current manuscript. Improvements are offered by way of adaptive testing, which allows for an assessment to better target the achievement levels of participants. PISA 2018 introduced an adaptive design to the reading domain and PISA 2022 will introduce an adaptive design in math. Although the OECD plans to develop items to enhance measurement at the low end of the continuum for 2021 (now 2022, due to the pandemic), this innovation is only planned for the computer-based version of PISA (OECD, 2019b). Unfortunately, many of the lowest performing countries are expected to participate in the paper-and-pencil platform, meaning that these countries will not be able to take advantage of specially designed items. Nevertheless, countries have an option to take part in what is effectively a continuation of PISA-D, administered via paper-and-pencil in parallel with the computer-based version of the assessment, in which most countries are encouraged to participate. A second possibility, should the OECD continue to offer this PISA-D option, would be to include content that is analogous to reading components for math and science. The IEA has recently implemented a design feature called a “group adaptive design” where booklets of easier or more difficult items are distributed in varying proportions across countries, depending on observed or expected achievement (Mullis and Martin, 2019). The results of this innovation are not yet published; however, this might also offer an opportunity for improving measurement in extreme populations.

Given the promise of computer-based testing, we conclude with a call for testing organizations such as the OECD and IEA to move toward universal computer-based testing. Of course, this is no small feat. Available resources limit computer availability in less economically developed countries and some middle-income countries refuse the move to a computerized platform. Providing necessary technology during the assessment window could be done by aid organizations or charitable foundations. And gently but firmly compelling technological holdouts is a delicate but important task that should be taken up by organizational leadership.

Acknowledgments

This project was partially funded by a grant from the Norwegian Research Council, FINNUT program, # 255246.

References

- Birnbaum, A., 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord, F., Novick, M.R. (Eds.), *Stat. Theo. Mental Test Scores*. Addison-Wesley Publishing, pp. 397–424.
- Core Team, R., 2021. R: A Language and Environment for Statistical Computing (4.1.1) [Computer software]. R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Embretson, S.E., Reise, S.P., 2000. *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates, Inc.
- Hambleton, R.K., Swaminathan, H., 1985. *Item Response Theory: Principles and Applications*. Kluwer.
- International Association for the Evaluation of Educational Achievement, 2013. *TIMSS 2015 Assessment Frameworks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Matta, T., Rutkowski, L., Rutkowski, D., Liaw, Y.-L., 2018. Isasim: an R package for simulating large-scale assessment data. *Large-Scale Assess. Educ.* 6 (15). <https://link.springer.com/article/10.1186/s40536-018-0068-8>.
- Mislevy, R.J., Beaton, A.E., Kaplan, B., Sheehan, K.M., 1992. Estimating population characteristics from sparse matrix samples of item responses. *J. Educ. Meas.* 29 (2), 133–161.
- Mullis, I.V.S., Martin, M.O., 2019. PIRLS 2021 group adaptive design. In: *PIRLS 2021 Assessment Frameworks*. TIMSS & PIRLS International Study Center, Boston College. <http://pirls2021.org/frameworks/home/assessment-design-framework/group-adaptive-design/>.
- Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D.L., Fishbein, B., 2020. *TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/timss2019/international-results/>.
- OECD, 1999. *Measuring Student Knowledge and Skills: A New Framework for Assessment*. OECD Publishing. <https://www.oecd.org/education/school/programme-for-international-student-assessment-pisa/33692793.pdf>.
- OECD, 2010. *PISA 2009 Assessment Framework*. OECD Publishing. <https://doi.org/10.1787/9789264062658-en>.
- OECD, 2012. *PISA 2009 Technical Report*. OECD Publishing. <http://www.oecd.org/edu/preschoolandschool/programme-for-international-student-assessment-pisa/pisa2009technicalreport.htm>.
- OECD, 2013. *Lessons from PISA 2012 for the United States*. OECD Publishing. http://www.oecd-ilibrary.org/education/strong-performers-and-successful-reformers-in-education_2220363x.
- OECD, 2014. *PISA 2012 Technical Report*. OECD Publishing.
- OECD, 2016. *PISA for Development*.
- OECD, 2017a. *PISA 2015 Assessment and Analytical Framework*. Organisation for Economic Co-operation and Development. <https://doi.org/10.1787/9789264281820-en>.
- OECD, 2017b. *PISA 2015 Technical Report. Draft*. OECD Publishing.
- OECD, 2018. *PISA for Development: Results in Focus, vol 91*. OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-for-development_c094b186-en.
- OECD, 2019a. *PISA 2018 Results: Combined Executive Summaries*. OECD Publishing. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf.
- OECD, 2019b. *PISA 2021 Assessment and Analytical Framework*. OECD Publishing.
- OECD, 2019c. *PISA for Development 2018 Technical Report*. OECD Publishing. <https://www.oecd.org/pisa/pisa-for-development/pisaforddevelopment2018technicalreport/>.
- OECD, 2020. Annex J. In: *PISA 2018 Technical Report*. OECD. <https://www.oecd.org/pisa/data/pisa2018technicalreport/>.
- PBS, 2013. How Does American Education Measure up to Schools Around the Globe? http://www.pbs.org/newshour/bb/education-july-dec13-pisa_12-03/.
- Robitzsch, A., Kiefer, T., Wu, M., 2017. TAM Package.
- Rutkowski, D., Rutkowski, L., 2021. Running the wrong race? The case of PISA for Development. *Comp. Educ. Rev.* 65 (1), 147–165. <https://doi.org/10.1086/712409>.
- Rutkowski, L., Gonzalez, E., von Davier, M., Zhou, Y., 2014. Assessment design for international large-scale assessment. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. Chapman & Hall/CRC Press, pp. 75–95.
- Rutkowski, D., Rutkowski, L., Liaw, Y.-L., 2018. Measuring widening proficiency differences in international assessments: are current approaches enough? *Educ. Meas.* 37 (4), 40–48. <https://doi.org/10.1111/emip.12225>.
- Schneider, M., 2019. *PISA Is a Unique Resource for Testing Educational Attainment of 15-Year-Olds in 78 Countries: Adding 40 More Would Be a Mistake*. Pamphlets, Speeches. Institute of Education Sciences (IES), part of the U.S. Department of Education (ED). <https://ies.ed.gov/director/remarks/pisa-2019.asp>.
- Sellar, S., Thompson, G., Rutkowski, D., 2017. *The Global Education Race: Taking the Measure of PISA and International Testing*. Brush Education. <http://books.google.com/books?hl=en&lr=&id=cWdEDgAAQBAJ&oi=fnd&pg=PR7&dq=info:l-p5XsebA6MJ:scholar.google.com&ots=7JuDLYToml&sig=RB5RhscO2I7jh6GpLBAmP5nGiP0>.
- Sherer, R., Nilsen, T., 2023. *Accountability for the Future? International Testing and Future Directions*, fourth ed. International Encyclopedia of Education.
- van de Vijver, F.J.R., Matsumoto, D., 2011. Introduction to the methodological issues associated with cross-cultural research. In: Matsumoto, D., van de Vijver, F.J.R. (Eds.), *Cross-cultural Research Methods in Psychology*. Cambridge University Press.
- Volante, L., Klinger, D.A., 2021. PISA and education reform in Europe: cases of policy inertia, avoidance, and refraction. *Eur. Educ.* 53 (1), 45–56. <https://doi.org/10.1080/10564934.2021.1987839>.
- von Davier, M., Sinharay, S., 2014. Analytics in international large-scale assessments: item response theory and population models. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, pp. 155–174.

Taking purpose into account when designing, using and evaluating educational assessments

Paul E. Newton, Office of Qualifications and Examinations Regulation, Coventry, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Tagline: P.E. Newton, The Multiple Purposes of Assessment, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 392–396, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.00355-9>.

Purpose	136
Listing purposes	136
Analyzing purposes	137
Differentiating between purposes	138
Differentiating within purposes	139
Classifying purposes	139
Purposes and validity	140
Purposes and evaluation	141
Multiple perspectives on purposes	143
Taking multiple purposes into account	144
References	145

Purpose

Assessment in education—as it is practised within schools and colleges all over the world—is so commonplace, so deeply embedded, and often so ritualized, that it is easy to take it for granted. Indeed, sometimes an assessment can become so embedded and so ritualized that the reason for its very existence becomes obscured. Unfortunately, when we fail to pay sufficient attention to the purpose, or purposes, that an educational assessment is supposed to be serving, bad things can happen. The present chapter concerns the very many different purposes that drive assessment in education, and explains why they need to be carefully taken into account whenever an educational assessment is designed, used or evaluated.

Listing purposes

Examinations are used for many purposes, and many of the difficulties and evils connected with their use arise from the attempt of those who frame the Examinations, either to effect too many things by one instrument; as, for instance, to pick out the most proficient, to reject the dunces, and to give an impetus to a certain kind of study all by the same series of papers; or from their not having a clear view of what they are aiming at.

Latham (1886, p. 2).

Although much has been written about the purposes of educational assessments, there is no cohesive literature on the topic, in the sense of an established body of accumulated wisdom. Instead, the corpus of work on purposes, such as it is, tends to be fragmented, with little in the way of cross-referencing. Many of the interesting papers that I have read on assessment purposes I have merely stumbled across. This makes the topic quite hard to review!

Papers on assessment purposes often feature a particular construction—the “long list”—which helps to illustrate the fact that there are a surprising number of them, while also suggesting (in line with Latham) that problems can arise when the same assessment is used to achieve too many different purposes. The earliest “long list” that I have discovered comes from a paper entitled *Preliminary Steps in Initiating a Testing Program*. In it, [Stevenson \(1922\)](#) identified 26 purposes, which he described as merely a “partial list of uses for standardized tests” (p. 156). The list began:

1. For diagnosing common and individual errors as a basis for remedial instruction.

2. For prognosis of pupils’ abilities.

3. For classification of pupils according to ability.

4. To determine whether sufficient time is being spent on a given subject.

Stevenson (1922, p. 156).

To cut to the chase, the list ended:

23. As a means of interesting children in school work and improvement. (Especially is this true if charts or graphic records of their own work are kept.)

24. As a protection to the teacher or principal from the irate parents. (Parents who are not satisfied with the teacher's marks may be shown definitely how their child stands in relation to the other children of the room.)

25. To show the relative standing of different grades or buildings within a system or a comparison of these with other cities.

26. To show where emphasis on different subjects should be placed.

Stevenson (1922, p. 156).

Lists of purposes—some shorter, some longer—are a feature of many of the papers that I have stumbled across over the years, including:

- Odell (1928)—15 purposes of “measurements”
- Tyler (1942)—6 purposes of “evaluation”
- Hieronymus (1972)—10 purposes of reading assessment “instruments”
- Shepard (1979)—7 purposes of “assessment and evaluation”
- Haladyna et al. (1991)—29 uses of “standardized test information”
- Baker (2000)—10 purposes of “tests”
- Mansell (2007)—20 uses to which “test and exam results” contribute
- Cedefop (2010)—40 purposes and functions of “qualifications”

Having prepared a few long lists myself (e.g., Newton, 2007a,b), I can vouch for their sobering rhetorical force. They are very useful tool for helping stakeholders to appreciate why an assessment that has been designed to be used for one particular purpose may not necessarily be equally suitable to be used for a quite different one.

Analyzing purposes

One of the easiest mistakes to make, when considering what the purpose of an educational assessment might be, is to foreclose analysis too soon, typically by remaining purely at the *measurement level*. In other words, it is tempting to assume that the fundamental purpose of any educational assessment is to measure. (I use this term in a fairly liberal, encompassing sense, allowing it to be interpreted: either strongly or weakly; literally or metaphorically; macroscopically or microscopically.) If so, then all that really needs to be established is the domain of learning that needs to be measured, and the approach that will be taken to measuring it.

Imagine, for instance, an assessment that a teacher had decided to administer, following a year of teaching maths to a class of 13-year-olds. If we were to ask her what her purpose was, in administering the assessment, she might reply (matter-of-factly) that it was simply to measure how much maths each of her students had learned that year, which is why she had designed a test to sample all of the content that had been taught. This would represent a simple measurement-level analysis.

As it happens, even this response would require a little more unpacking. In particular, did the teacher specifically want to measure how much maths each student had learned *that year*? If so, then this would actually require a progress calculation of some sort—to compute the difference between each learner's level of attainment at the beginning of the year and their level of attainment at the end of the year—which might require two separate tests, both of which would need to cover all of the content of that year's course, or perhaps just the same test administered twice. Alternatively, was the teacher simply interested to know how much mathematical knowledge and skill each student had acquired *by the end of her course of instruction*? In that case, a single,

end-of-year maths test might suffice. On the other hand, that test might need to cover more than just the content that she had taught, *that year*, perhaps also including content that had been taught in previous years. This would mean measuring students in terms of a broader attainment domain, related to progression across the school curriculum for maths.

In fact, to fully understand why this teacher felt the need to administer an end-of-year test, we would need to extend our analysis beyond the measurement level, to the *decision level*. The primary question, here, is what she actually intended to *do* with the information that the test would provide, i.e., the information captured by the test result. The bottom line is that we assess in order to generate information, and that information will subsequently be used to make a decision (generally speaking).

It is important to understand this, because different kinds of decision-making may require access to different kinds of information, which may require different kinds of assessment evidence. In other words, in order to know how to design any particular educational assessment, we will need to work backwards from the decision that ultimately needs to be made, and the kind of information that would be needed in order to make it.

Imagine, now, that our maths teacher had wanted to decide which of her students was the most deserving of a reward for having worked hard that year. If so, then she might conclude that this purpose could best be served by measuring progress in attainment in maths; defined purely in relation to the content of the course that she had taught that year. This would enable her to identify the most deserving student in terms of the magnitude of the difference between their level of attainment at the beginning of the year (based on one test) and their level of attainment at the end of the year (based on a second version of the same test).

If, instead, she wanted to make a quite different kind of decision—for example, to decide which of her students had the most potential to benefit from studying an optional advanced level maths course in the following year—then she might conclude that she needed to measure a quite different aspect of attainment. This might require just a single, end-of-course test of overall level of attainment in maths; this time defined relative to the broader school curriculum, and perhaps extending to content that had not yet been taught.

Differentiating between purposes

Results from educational assessments are used for all sorts of purposes, to make all sorts of decisions. This is a long list that I prepared for the last edition of this encyclopedia (Newton, 2010):

1. To decide whether individual students are making sufficient progress in attainment in relation to expectations or targets; and, potentially, to allocate rewards or sanctions (student monitoring)
2. To identify students' proximal learning needs, to guide subsequent teaching interventions (formative)
3. To judge the social or personal value of students' achievements (social evaluation)
4. To clarify the type and extent of students' learning difficulties, or needs, in light of well-established criteria (diagnosis)
5. To determine whether students meet eligibility criteria for special educational provision (provision eligibility)
6. To identify students who differ significantly from their peers, for further assessment (screening)
7. To segregate students into homogeneous groups, on the basis of aptitudes or attainments, to make the instructional process more straightforward (segregation)
8. To identify the most suitable courses, or vocations for students to pursue, given their aptitudes (guidance)
9. To identify the general educational needs of students who transfer to new schools (transfer)
10. To locate students with respect to their position in a specified learning sequence, to identify the level of course which most closely reflects it (placement)
11. To decide whether students are sufficiently qualified for a job, course, or role in life—that is, whether they are equipped to succeed in it—and whether to enroll them or to appoint them to it (qualification)
12. To predict which students—all of whom might, in principle, be sufficiently qualified—will be the most successful in a job, course, or role in life, and to select between them (selection)
13. To provide legal evidence—the license—of minimum competence to practice a specialist activity, to warrant stakeholder trust in the practitioner (licensing)
14. To provide evidence—the certificate—of higher competence to practice a specialist activity, or subset thereof, to warrant stakeholder trust in the practitioner (certification)
15. To identify the most desirable school for a child to attend (school choice)
16. To decide whether institutional performance—relating to individual teachers, classes or schools—is rising or falling in relation to expectations or targets; and, potentially, to allocate rewards or sanctions (institution monitoring)
17. To identify institutional needs and, consequently, to allocate resources (resource allocation)
18. To identify institutional failure and, consequently, to justify intervention (organizational intervention)
19. To evaluate the success of educational programs or initiatives, nationally or locally (program evaluation)
20. To decide whether system performance—relating to individual regions or the nation—is rising or falling in relation to expectations or targets; and, potentially, to allocate rewards or sanctions (system monitoring)
21. To guide decisions on comparability of examination standards for later assessments on the basis of cohort performance in earlier ones (comparability)
22. To “quality adjust” education output indicators (national accounting).

This is not an exhaustive list, but it does help to illustrate the many different kinds of decision that can be based upon educational assessment results. Again, the important point is that different kinds of information may be required to make different kinds of decision, which may necessitate quite different assessment design choices. Thus, [Millman and Greene \(1989\)](#) illustrated how assessments designed for program evaluation might differ from those designed for other purposes, such as student monitoring; by targeting learning objectives beyond those specific to the course; by employing item formats unfamiliar to those following the course; by allocating different items to different students, to sample learning objectives thoroughly; and by scoring by learning objective, rather than at an overall level. This kind of assessment would not even report at the level of individual students, let alone allow a user to monitor student progress in relation to a specific course of instruction.

Similarly, the assessment evidence required to decide whether a student is progressing at a suitable rate in chemistry (student monitoring) will be quite different from the assessment evidence required to identify their day-to-day learning needs in chemistry (formative). The former is likely to involve a high-level inference, concerning the student's overall level of attainment; whereas the latter is likely to involve many low-level inferences, concerning the nature of the student's knowledge or skill, in relation to specific concepts or practices. Clearly, evidence that is only elicited at a high level, for the purpose of monitoring, cannot subsequently be used to draw low-level formative inferences. And, although it might be possible, in theory, to re-use evidence elicited at a low level (for formative purposes) to draw high-level inferences (for monitoring, selection, or suchlike) it is a moot point as to whether this is really viable, in practice.

The purpose of this section is to reinforce the key message of this chapter: there are many different purposes for which educational assessments are designed; and an assessment designed to serve one purpose cannot necessarily be used effectively to serve others.

Differentiating within purposes

The list of 22 presented above is best understood, not as a list of discrete purposes, but as a list of categories of purpose. Within each of these categories, it would be possible to identify distinct sub-purposes; and these distinct sub-purposes—even within a single purpose category—might implicate quite different kinds of decision, which might require quite different kinds of information, which might necessitate quite different assessment design choices. Again, program evaluation provides a good case in point. As Cronbach explained many decades ago:

If you wish only to know how well a curriculum is achieving its objectives, you fit the test to the curriculum; but if you wish to know how well the curriculum is serving the national interest, you measure all the outcomes that might be worth striving for.

[Cronbach \(1964, p. 243\).](#)

Similar distinctions could be made in relation to many other categories of purpose, where different sub-purposes are associated with different inferences. This explains why [Nitko \(1989\)](#) was able to identify five different categories of diagnostic test, each requiring a different approach to test design. It also explains why tests used for system monitoring are often designed very differently from one country to the next; for example, New Zealand's National Education Monitoring Project versus the USA's National Assessment of Educational Progress (see [Newton, 2008](#)). In New Zealand, the NEMP was intended to be of particular use to teachers and curriculum designers; and, consequently, it was designed to report at the level of individual tasks and small clusters of tasks. In the USA, although originally reporting at the level of individual tasks, the long-term trend NAEP was specifically redesigned to be of use to policy makers and stakeholders, subsequently reporting only at the overall subject level.

Even the distinction between high stakes and low stakes decision-making will necessitate different design choices. Assessments that have high stakes for individuals—such as qualification, selection or licensing—will need to be designed with high levels of security in mind, at all stages of the assessment lifecycle, e.g., from identity checking at a test center, to the disposal of completed test scripts ([Millman and Greene, 1989](#); [Schmeiser and Welch, 2006](#)). Such considerations simply do not arise when results are used only for low stakes purposes, where cheating does not present a significant threat to validity.

Classifying purposes

It is fairly easy to develop schemes for classifying assessment purposes. For instance, those up to purpose 14 (above) relate primarily to individuals and are based upon individual results; whereas those from 15 to 22 relate to larger entities (such as schools, cohorts or nations) and are based upon aggregated results. Equally, though, the same 22 categories of purposes could be classified in an entirely different manner, as illustrated by the following scheme:

1. Personal decisions, made by students and parents (e.g., social evaluation, guidance, school choice)
2. Educational decisions, made by teachers and psychologists (e.g., formative, diagnosis, screening, placement, transfer, segregation)
3. Administrative decisions, made by gatekeepers (e.g., provision eligibility, qualification, selection)

4. Managerial decisions, made by those supervising and directing performance (e.g., student monitoring, institution monitoring, system monitoring, resource allocation, organizational intervention)
5. Endorsed decisions, made by consumers (e.g., licensing, certification)
6. Technical decisions, made by technicians and analysts (e.g., comparability, national accounting)
7. Evaluative decisions, made by researchers (e.g., program evaluation).

Scholars who write about assessment purposes often introduce a new classification scheme to the literature (e.g., [Odell, 1928](#); [Cook, 1951](#); [Findley, 1963](#); [Bloom et al., 1971](#); [Millman and Greene, 1989](#); [Nitko, 1989](#); [Hopkins et al., 1990](#); [Mehrens and Lehmann, 1991](#); [Black, 1998](#); [Pellegrino et al., 2001](#); [Schmeiser and Welch, 2006](#); [Gronlund, 2006](#); [Nitko and Brookhart, 2007](#); [Thorndike and Thorndike-Christ, 2010](#); [Baker, 2013](#); [Haertel, 2013a](#)). This observation, in itself, would seem to cast some doubt over the utility of the exercise. In an ideal world, a classification scheme might help the reader to understand how certain kinds of purpose are associated with certain forms of assessment. But the reality is simply not that straightforward, and the many classification schemes in existence seem to provide little more than relatively arbitrary and contestable groupings which lack obvious utility.

Both [Millman and Greene \(1989\)](#) and [Schmeiser and Welch \(2006\)](#) have noted the lack of a universally accepted scheme, with the implication that it would be useful to have one. Yet, classification schemes are not necessarily always a good thing. Questionable classifications can mislead policy makers, users and members of the public into drawing inappropriate inferences, by implying that sub-purposes within each category or class are fundamentally equivalent, in a functional sense. This can mislead stakeholders into assuming that similarly grouped purposes can straightforwardly be supported by the same assessment instrument. Conversely, emphasizing differences between, and even within, purposes helps to foreground the absolute necessity of tailoring assessment design to a precise statement of intended purpose(s). This is another reason to recommend the long list, which resists further classification.

Purposes and validity

Validity is about assessing the right thing in the right way to produce accurate and useful assessment results ([Newton, 2017a](#)). Expressed like this, it is clear that validity and purpose are intimately entwined, which is evident in the definition of validity that appears in the most recent (sixth) edition of the North American *Standards for Educational and Psychological Testing*:

Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests.

[AERA, et al. \(2014, p. 11\)](#).

The precise formulation of this definition—in terms of interpretations for uses—is significant. It states that validity concerns the suitability of information provided by assessment results for the use to which it will be put, i.e., the decision that needs to be taken. Although this interpretation-for-use formulation has only featured in the last two editions of the *Standards*, its fundamentally pragmatic approach, which links validity directly to purpose, has been evident since the very first edition.

Ironically, this important principle that validity can only be judged in relation to purpose—more specifically, the way in which this principle was presented in the first few editions of the *Standards*—seems to have been responsible for entrenching what is now generally accepted to be a particularly unhelpful construction of validity. This is the idea that there are distinct kinds of validity, including content validity, criterion validity, and construct validity ([Newton and Shaw, 2013](#)).

Table 1 is adapted from [Anastasi \(1976\)](#), who was attempting to illustrate the validation of a single arithmetic test in relation to multiple purposes. The idea that she was trying to illustrate is that different ways of interpreting test results (in terms of different kinds of attributes) may suggest different approaches to validation. Thus, a test of achievement (i.e., attainment) will tend to be validated logically, by analyzing its content; whereas a test of aptitude will tend to be validated empirically, by correlating its outcomes against a future criterion measure. Back in the 1960s and 1970s, the former tended to be described as establishing content validity, and the latter tended to be described as establishing criterion validity.

Table 1 Using a single test for multiple purposes.

Testing purpose	Illustrative question	Validity type
Test of achievement in elementary school arithmetic	How much basic arithmetic has Jo learned?	Content
Test of aptitude for high school maths	How effectively will Jan learn higher maths (in the future)?	Criterion
Test of logical reasoning	How intellectually gifted is Jean?	Construct

Adaptation of Table 13, from [Anastasi \(1976, p. 158\)](#).

Unfortunately, this over-simplified characterization of validity theory and validation practice became increasingly associated with inadequate validation research, and unjustified validity conclusions (see [Newton and Shaw, 2014](#)). Scholars including [Messick \(1980\)](#) and [Guion \(1980\)](#) argued forcefully against what they described as a “trinitarian doctrine” of validity—which distinguished three different kinds of validity—and argued instead that all validity ought to be conceptualized as construct validity. In turn, construct validation ought to be conceptualized as a scientific program of research, requiring the integration of evidence and analysis from a variety of logical and empirical techniques. This proposal to unify validity theory and validation practice was welcomed, and is now generally accepted.

This shift toward a unified conception of validity does nothing to undermine the link between validity and purpose. However, the link tends now to be framed in terms of different interpretations for different uses, rather than in terms of different kinds of validity. Notice that the point of column 1 of Anastasi’s table was also to foreground the different interpretations that needed to be validated, in relation to different uses. So, the underlying principle of relating validity to use (i.e., decision-making) has remained the same. (Although Anastasi labeled column 1 “testing purpose” these purposes were actually left implicit, while the interpretations were foregrounded.)

The logic of validation, implicit in Anastasi’s table, would go like this. First of all, results from the arithmetic test would need to be validated as measures of achievement in elementary school arithmetic. Of course, as the test was specifically designed to measure this attribute, we might expect the validation program to be able to establish that it had a high degree of validity, in relation to this interpretation of results, and therefore in relation to uses premised upon an interpretation of this sort. Subsequently, if results from the same arithmetic test were to be interpreted in terms of a more general aptitude for mathematics, for a different kind of decision-making, then that would require additional validation research. This might include evidence of the degree to which structurally similar elementary school arithmetic tests (administered in the past) had been able to predict subsequent achievement in high school maths. Naturally, though, as these elementary school tests were designed to measure arithmetic, rather than aptitude, *per se*, we might expect a somewhat lower degree of validity in relation to this interpretation (and use) of results. The same would be true for an interpretation of results in terms of logical reasoning ability; for which we might expect an even lower degree of validity.

The key insight from this pragmatic approach to validity and validation is that, it is not *the test* that is validated, but plausible *interpretations* of results from the test, and therefore the *uses* that are premised upon such interpretations. Consequently, different interpretations (and uses) of results from a single test may be judged to have different degrees of validity. Although, in recent years, there has been a renewed emphasis upon interpretations within validity theory, this is certainly not a new way of thinking. The link between validity and purpose has been recognized for as long as validity has been theorized, for instance:

If a test is “valid,” it is valid for a *given purpose*, with a given group of pupils, and it is valid only to the degree that it accomplishes that specific purpose for that specific group. It is meaningless to speak of any given test as being valid or invalid apart from any consideration of the purpose it is intended to serve or of the group to which it is to be given. [...] If a test is to be made most valid for one purpose, its validity for other purposes must often be sacrificed. An “all-purpose” test can perform none of its many purposes effectively.

[Lindquist \(1936, pp. 21–22\).](#)

Purposes and evaluation

The pragmatic approach to conceptualizing validity has remained popular since validity was introduced to the educational measurement lexicon during the 1920s ([Newton and Shaw, 2014](#)). More recently, though, [Messick’s \(1989\)](#) argument in favor of a unitary conception of validity—which focused a great deal of attention upon interpretations, uses, meanings, and consequences—initiated what has become known as the Great Validity Debate ([Crocker, 1997](#)), which rumbles on to the present day ([Newton and Baird, 2016](#)). As this debate developed, the pragmatic approach was challenged from opposing directions.

From one direction, it was argued that the pragmatic approach to defining validity—in terms of interpretations, and especially in terms of uses—is too broad. Instead, validity should be defined more scientifically, by focusing purely upon the capacity of a test to measure. This should exclude the evaluation of test use, which is an unnecessary distraction from proper validation ([Scriven, 2002](#); [Borsboom et al., 2004](#)). However, from precisely the opposite direction, it was argued that the pragmatic approach to defining validity is too narrow. Instead, validity should be defined more ethically, by extending its focus to include even unintended consequences from testing. This should emphasize the evaluation of test impact, which is a necessary component of validation ([Linn, 1997](#); [Moss, 1998](#)). The key features of these three approaches are summarized in [Table 2](#); the first row illustrates the intended focus of validation research; the second row illustrates the ultimate validation question.

It is not possible to identify a logically correct answer to the question concerning how the word “validity” ought to be defined ([Newton and Shaw, 2016](#)). This is really a consequential matter, which turns on the consequences of defining validity one way rather than any other. Although this is a challenging, and an as-yet-unresolved debate, it has been really useful in illustrating, and emphasizing, the many different facets of assessment policy and practice that need to be evaluated, whether under the banner of validity, or not.

Table 2 Alternative approaches to defining validity.

<i>Essentially scientific approach</i>	<i>Fundamentally pragmatic approach</i>	<i>Ultimately ethical approach</i>
Technical quality of measurement.	Technical quality of measurement (for specified purposes). Practical utility of measurement-based decisions (for specified purposes).	Technical quality of measurement (for specified purposes). Practical utility of measurement-based decisions (for specified purposes). Social value of the (full set of) consequences associated with the testing program.
How effectively do test results measure what they are supposed to measure?	How suitable are test results for the uses to which they are supposed to be put?	How acceptable is it to implement the testing policy?

In particular, debate over the ethical approach to defining validity has emphasized that, beyond the measurement level and the decision level, lies another level of analysis: the *impact level*. In order to evaluate the overall acceptability of an assessment policy, we need to look way beyond the technical quality of the assessment procedure. We need also to consider all of the consequences that the assessment gives rise to, whether anticipated or unanticipated, intended or unintended.

The idea of an intended impact brings us right back to the purpose of the assessment. It reminds us, not simply that there might be multiple intended purposes for administering any particular assessment, but also that these might include entirely different kinds of intended purpose. Earlier on, we discussed purpose superficially in terms of measurement, then more deeply in terms of decisions. Now, we need to discuss purpose independently in terms of impacts.

Obviously, the intention underlying any measurement-based decision-making is to ensure substantial positive consequences, with minimal negative ones. In other words, good quality decisions, based upon good quality information, ought to reap rewards. As such, the impacts arising from measurement-based decision-making will need to be evaluated: first, in their own right, to check that the anticipated cost-benefit payoff has been delivered; but also, second, because of the relevance of those impacts to the (pragmatic) validity argument. In other words, if the actual consequences of using the educational assessment are in line with the anticipated ones, then this helps to validate the measurement-based decision-making procedure from a purely technical perspective, as well as being reassuring from an ethical one.

In recent years, we have come to recognize (anew) the importance of purposes for educational assessments that are entirely independent of measurement-based decision-making. These intended impacts are presumed to arise directly from the administration of the assessment, or from its anticipation; and they are presumed to arise even if the assessment happens to lack validity from a technical perspective. In other words, these consequences operate independently of the quality of measurement-based decision-making. Such impacts include incentivizing and messaging effects (Haertel, 2013a; Ho, 2014). They have been the focus of considerable attention among scholars working on theories of action related to intended impacts (e.g., Bennett et al., 2011; Carpentieri, 2013; Haertel and Herman, 2005; Lane et al., 1998). Incidentally, when validity is defined straightforwardly in relation to what an assessment practice claims to do (Shepard, 1993), this formally extends the scope of validation research to include evaluations of this sort, in line with the ethical approach described above (Haertel, 2013b).

This chapter began by noting that bad things can happen when we fail to pay sufficient attention to the purpose, or purposes, that an educational assessment is supposed to be serving. These bad things include the following. First, when results from an educational assessment are used for a purpose for which it was not specifically designed, the information captured by those results may be sub-optimal for that secondary use, and this will impact negatively upon the accuracy and utility of the decisions that the results will be used to make.

This might happen if the effectiveness of a primary school were to be judged purely on the basis of aggregated test results, from attainment tests in maths and English, sat by successive (graduating) cohorts of students. It is quite possible that results from those tests could be valid for measuring attainment in maths and English, for individual students, at the same time as being invalid for measuring the effectiveness of the school. After all, there is far more to an effective school than the ability to teach maths and English. Moreover, without using those aggregated results to calculate a progress measure, they would not be adjusted to account for the quality of the school's intake. Clearly, if parents were to decide the most effective school for their children based *only* upon evidence like this, then very many incorrect decisions would be made. The likelihood of errors like these ought to be evaluated and publicized, if decisions like these (based on secondary uses of results) are actually being taken; and especially if these are officially sanctioned uses.

Second, the broader consequences of using results for secondary purposes must also be evaluated. For instance, if parents actually were to use aggregated results in maths and English tests as a basis for school choice, then those results would end up having very high stakes for schools. This would create perverse incentives for schools to do everything in their power to ensure that each graduating cohort scored as highly as possible on the maths and English tests. This might lead to teaching time being directed away from subjects that were not tested; or perhaps to certain important aspects of the curriculum not being taught at all. More worrying still, to the extent that it might be possible to improve scores in maths and English by nefarious means, e.g., by using a variety of questionable drilling or coaching techniques, this could negatively impact the quality of teaching and learning of maths and English.

In short, the greatest risk is not that results from a secondary use might be less suitable for that secondary use than for the primary use (which would be true); nor that impacts arising from the secondary use—e.g., reallocation of curriculum time to maths and English—might make those results even less suitable for that secondary use (which might also be true); nor even that impacts arising from the secondary use—e.g., reallocation of teaching time away from those elements of maths and English that featured least frequently in the tests—might render those results less suitable for the primary purpose as well (which might also be true). The most serious concern is that impacts arising from the secondary use—e.g., a classroom culture that is focused upon attempting to game the tests wherever possible—might come to undermine the pedagogical process itself, meaning that students end up learning less than they would have done, if the secondary use had never been mooted. The bad things that can happen when we fail to pay sufficient attention to the many different purposes of educational assessments—especially accountability related purposes—are eloquently discussed by Koretz (2008, 2017).

Multiple perspectives on purposes

Debates that arise when designing qualifications are useful for illustrating the complexity of policy and practice in relation to assessment purposes. In particular, they emphasize how all three levels of analysis—measurement, decision, and impact—will generally need to be taken into account, simultaneously, when planning how to satisfy a variety of assessment purposes. The issues at stake are partly technical (concerning effective assessment design) but they are also educational, economic, and social. It is not simply that there are multiple purposes for assessment in education, nor multiple categories of purpose, there are also quite different perspectives on assessment purposes. This becomes especially apparent when viewed through the lens of qualification design, because qualification purposes are discussed in all sorts of different ways (Cedefop, 2010), revealing a plethora of ways of thinking about what they need to do.

Three perspectives are especially important: the information perspective (focused on decisions); the engagement perspective (focused on impacts); and the expertise perspective (focused on learning outcomes). To illustrate these perspectives, imagine that a state or country is considering introducing a new suite of vocational qualifications, for 14- to 16-year-olds who choose not to opt for the traditional “academic” route. If different stakeholders were to be asked what they needed these qualifications to deliver, i.e., what the purpose(s) of these qualifications ought to be, how might they answer? It is quite possible that their answers might have much in common. Yet, it is also possible that different stakeholders might have quite different priorities.

An employer, for instance, might prioritize the need for a qualification to provide high quality *information* concerning the candidates who apply for the jobs that they will be advertising in the future. They might have had bad experiences, in the past, hiring ostensibly qualified candidates, who turned out not to have been competent after all. Employers like these need qualifications that they can rely upon, so as not to make costly mistakes when hiring employees. This necessitates high quality assessment information, derived from high validity educational assessments.

A learner, on the other hand, might prioritize the need for the new qualifications to provide high value learning outcomes, to provide them with the right kind of *expertise*, which effectively equips them for both life and work. Indeed, if learners envisaged a world in which they were likely to have to adapt to a variety of different roles throughout their career, switching jobs and possibly also sectors, then this might dispose them toward qualifications based upon broad, transferrable, learning outcomes. Conversely, were an individual employer to offer their thoughts on this perspective, they might argue instead for qualifications based upon narrow, job-specific, learning outcomes. A construction company might therefore prefer the country to be producing a steady stream of fully qualified carpenters, bricklayers, electricians, and so on; certainly not vocational all-rounders, on whom they would need to spend considerable sums of money on additional training to prepare them for specific job roles. As such, even when debating a single purpose perspective, different stakeholders might champion different requirements.

Finally, a policy maker, perhaps keen to tackle decades of high educational drop-out and failure rates, might prioritize a quite different purpose perspective, wishing to maximize the *engagement* of students and teachers with the processes of learning and teaching. For the sake of argument, this policy maker might be fairly ambivalent about the expertise perspective, and the exact nature of the learning outcomes that these new vocational qualifications ought to deliver. Indeed, prioritizing the engagement perspective, they might argue for a guiding principle which states that learning outcomes must simply be motivating for learners, to maximize the likelihood that they do not turn off and drop out. This might lead them to recommend a system that is characterized by choice from a wide variety of optional units, each one targeting a body of content that has been chosen specifically because it has proven to be interesting to teach and to learn (almost regardless of its value in terms of equipping learners for life and work).

It is worth noting that assessment professionals tend to contribute to debates such as these from the information perspective. Assessment is quintessentially a process of deriving information about individuals from evidence of their behavior, so this should come as no surprise. The information perspective establishes the foundational paradigm for assessment in education. Equally, though, assessment professionals should not be surprised when education professionals who work within alternative paradigms prioritize different perspectives. This can be particularly challenging when those other professionals also have a legitimate stake in assessment. Thus, assessment professionals may find it confusing or frustrating that teachers often use assessment as much to reward learners as to represent their attainments, if not more so (Brookhart, 1993, 1994). What these teachers are doing, in fact, is prioritizing the engagement perspective alongside, or perhaps even above, the information perspective. In other words, they are using assessment to facilitate greater engagement with teaching and learning, and not just to measure.

Of course, the “threat” of an end-of-course assessment has always been assumed to serve an important function in motivating students to learn (Latham, 1886). In other words, the engagement perspective has always been fundamental to assessment design. Problems arise, however, when the engagement perspective recommends assessment design decisions that compromise the ability of the assessment to generate high quality information. Even then, it is still possible that a small hit on the technical quality of the assessment might be deemed acceptable, if there were a sufficiently large payoff in terms of teaching and learning.

Taking multiple purposes into account

I began writing on assessment purposes during the early 2000s. Testing had become endemic, in England, by then—involving the vast majority of 7-year-olds, 11-year-olds, and 13-year-olds, each year—and results from national curriculum tests were being used for all sorts of purposes, spread across at least 16 different purpose categories (Newton, 2012a). A House of Commons enquiry report, *Testing and Assessment*, concluded that:

The evidence we have received strongly favours the view that national tests do not serve all of the purposes for which they are, in fact used. The fact that the results of these tests are used for so many purposes, with high-stakes attached to the outcomes, creates tensions in the system leading to undesirable consequences, including distortion of the education experience of many children. [...] In short, we consider that the current national testing system is being applied to serve too many purposes.

HCCSFC (2008a, p. 20).

In response, the Government stated that it did not accept the argument that the tests served too many purposes; and its Office for Standards in Education added that one of the core principles for efficient use of data in government should be “collect once, use more than once” which is how they characterized the situation of testing in England (HCCSFC, 2008b).

Assessment professionals often express concerns over the use of results for multiple purposes, citing maxims like this one, from the current edition of the *Standards*, which echoes the same warning from Lindquist, some 7 decades earlier:

Most educational tests will serve one purpose better than others; and the more purposes an educational test is purported to serve, the less likely it is to serve any of these purposes effectively.

AERA et al. (2014, p. 188).

As we have already seen, there are good reasons for expressing such concerns. Results from assessments that have been designed to serve one purpose may be completely unfit to be used for another purpose; owing to design choices that increase validity in relation to one use of results, but decrease validity in relation to others. Moreover, using results for purposes that were not originally anticipated may not only beg the validity question, but may also result in negative educational and social consequences. The current edition of the *Standards* provides a number of useful recommendations related to the multiple purposes of assessments, including:

The rationale for a test, recommended uses of the test, support for such uses, and information that assists in score interpretation should be documented. When particular misuses of a test can be reasonably anticipated, cautions against such misuses should be specified.

AERA et al. (2014, Standard 7.1, p. 125).

In educational settings, when a test is designed or used to serve multiple purposes, evidence of validity, reliability/precision, and fairness should be provided for each intended use.

AERA et al. (2014, Standard 12.2, p. 195).

An index that is constructed by manipulating and combining test scores should be subjected to the same validity, reliability, and fairness investigations that are expected for the test scores that underlie the index.

AERA et al. (2014, p. 210).

With renewed interest in validity theory in recent years, closer attention has been paid to the challenges of taking multiple purposes into account as a routine component of validation practice, and new tools for thinking about these challenges have been proposed (e.g., Kane, 2006; Zumbo and Forer, 2011; Newton, 2012b; Koch and DeLuca, 2012; Koch, 2013; Ho, 2014).

Taking multiple purposes into account as a routine component of assessment design presents even greater challenges (Newton, 2017b). There will always be trade-offs to discuss, and compromises to be made, and answers will almost never be straightforward. Conversations like these inevitably take assessment professionals beyond their comfort zone, requiring them to engage with

professionals who may think very differently, and value different outcomes. As complex and frustrating as such discussions can be, it is critical that all relevant concerns are aired, from all possible perspectives, for design choices to be made as defensibly as possible.

It seems fair to say that assessment professionals have a tendency to gravitate toward purpose purism, which idealizes the principle that assessment design ought to be driven by a single assessment purpose; after all, that really is the best approach to take, from a purely technical perspective (cf. Baker, 2000). However, a broader, more realistic outlook is captured by the idea of purpose pluralism, which idealizes the principle that assessment design should be driven by a multiplicity of assessment purposes simultaneously (Newton, 2017b). This reminds us that there is far more to assessment in education than can be captured by the “simple” idea of measuring learners.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. *Standards for Educational and Psychological Testing*. American Educational Research Association, Washington, DC.
- Anastasi, A., 1976. *Psychological Testing*, fourth ed. Macmillan Publishing Co. Inc, New York.
- Baker, E.L., 2000. *Understanding Educational Quality: Where Validity Meets Technology*. William H. Angoff Memorial Lecture Series. Educational Testing Service, Princeton, NJ.
- Baker, E.L., 2013. The chimera of validity. *Teach. Coll. Rec.* 115 (9), 1–26.
- Bennett, R.E., Kane, M., Bridgeman, B., 2011. *Theory of Action and Validity Argument in the Context of Through-Course Summative Assessment*. Educational Testing Service, Princeton, NJ.
- Black, P.J., 1998. *Testing: Friend or Foe? The Theory and Practice of Assessment and Testing*. Falmer Press, London.
- Bloom, B.S., Hastings, J.T., Madaus, G.F., 1971. *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill, New York.
- Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2004. The concept of validity. *Psychol. Rev.* 111 (4), 1061–1071.
- Brookhart, S.M., 1993. Teachers' grading practices: meaning and values. *J. Educ. Meas.* 30 (2), 123–142.
- Brookhart, S.M., 1994. Teachers' grading: practice and theory. *Appl. Meas. Educ.* 7 (4), 279–301.
- Carpentieri, J.D., 2013. Evidence, evaluation and the “tyranny of effect size”: a proposal to more accurately measure programme impacts in adult and family literacy. *Eur. J. Educ.* 48 (4), 543–556.
- Cedefop, 2010. *Changing Qualifications*. Publications Office of the European Union, Luxembourg.
- Cook, W.W., 1951. The functions of measurement in the facilitation of learning. In: Lindquist, E.F. (Ed.), *Educational Measurement*. American Council on Education, Washington, DC, pp. 3–46.
- Crocker, L., 1997. Editorial: the great validity debate. *Educ. Meas.* 16 (2), 4–4.
- Cronbach, L.J., 1964. Evaluation for course improvement. In: Heath, R.W. (Ed.), *New Curricula*. Harper and Row, New York, pp. 231–248.
- Findley, W.G., 1963. Purposes of school testing programs and their efficient development. In: Findley, W.G. (Ed.), *The Impact and Improvement of School Testing Programs*. The Sixty-Second Yearbook of the National Society for the Study of Education. Part II. National Society for the Study of Education, University of Chicago Press, Chicago, IL, pp. 1–37.
- Gronlund, N.E., 2006. *Assessment of Student Achievement*, eighth ed. Pearson/Allyn and Bacon, Boston, MA.
- Guion, R.M., 1980. On Trinitarian doctrines of validity. *Prof. Psychol.* 11 (3), 385–398.
- Haertel, E.H., Herman, J.L., 2005. A historical perspective on validity arguments for accountability testing. In: Herman, J.L., Haertel, E.H. (Eds.), *Uses and Misuses of Data for Educational Accountability and Improvement*. The 104th Yearbook of the National Society for the Study of Education, Part 2. Blackwell Publishing, Malden, MA, pp. 1–29.
- Haertel, E.H., 2013a. How is testing supposed to improve schooling? *Measurement* 11 (1–2), 1–18.
- Haertel, E.H., 2013b. Expanding views of interpretation/use arguments. *Measurement* 11 (1–2), 68–70.
- Haladyna, T.M., Nolen, S.B., Haas, N.S., 1991. Raising standardized achievement test scores and the origins of test score pollution. *Educ. Res.* 20 (5), 2–7.
- Hieronymus, A.N., 1972. Evaluation and reading: perspective '72. *Read. Teach* 26 (3), 264–267.
- Ho, A.D., 2014. Variety and drift in the functions and purposes of assessment in K–12 education. *Teach. Coll. Rec.* 116 (11), 1–18.
- Hopkins, K.D., Stanley, J.C., Hopkins, B.R., 1990. *Educational and Psychological Measurement and Evaluation*, seventh ed. Allyn and Bacon, Needham Heights, MA.
- House of Commons Children, Schools and Families Committee, 2008a. *Testing and Assessment*. Third Report of Session 2007–08, vol. I. HC vols. 169–I. The Stationery Office Limited, London.
- House of Commons Children, Schools and Families Committee, 2008b. *Testing and Assessment: Government and Ofsted Responses to the Committee's Third Report of Session 2007–08*. Fifth Special Report of Session 2007–08. HC 1003. The Stationery Office Limited, London.
- Kane, M.T., 2006. Validation. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger Publishers, Westport, CT, pp. 17–64.
- Koch, M.J., 2013. The multiple-use of accountability assessments: implications for the process of validation. *Educ. Meas.* 32 (4), 2–15.
- Koch, M.J., DeLuca, C., 2012. Rethinking validation in complex high-stakes assessment contexts. *Assess. Educ. Princ. Pol. Pract.* 19 (1), 99–116.
- Koretz, D., 2008. *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press, Cambridge, MA.
- Koretz, D., 2017. *The Testing Charade: Pretending to Make Schools Better*. The University of Chicago Press, Chicago, IL.
- Lane, S., Parke, C.S., Stone, C.A., 1998. A framework for evaluating the consequences of assessment programs. *Educ. Meas.* 17 (2), 24–28.
- Latham, H., 1886. *On the Action of Examinations Considered as a Means of Selection*. Willard Small, Boston, MA.
- Lindquist, E.F., 1936. The theory of test construction. In: Hawkes, H.E., Lindquist, E.F., Mann, C.R. (Eds.), *The Construction and Use of Achievement Examinations: A Manual for Secondary School Teachers*. The Riverside Press, Cambridge, MA, pp. 17–106.
- Linn, R.L., 1997. Evaluating the validity of assessments: the consequences of use. *Educ. Meas.* 16 (2), 14–16.
- Mansell, W., 2007. *Education by Numbers: The Tyranny of Testing*. Politico's Publishing, London.
- Mehrens, W.A., Lehmann, I.J., 1991. *Measurement and Evaluation in Education and Psychology*, fourth ed. Holt, Rinehart and Winston, Fort Worth, TX.
- Messick, S., 1980. Test validity and the ethics of assessment. *Am. Psychol.* 35 (11), 1012–1027.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education/Macmillan, New York, pp. 13–103.
- Millman, J., Greene, J., 1989. The specification and development of tests of achievement and ability. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education/Macmillan, New York, pp. 335–366.
- Moss, P.A., 1998. The role of consequences in validity theory. *Educ. Meas.* 17 (2), 6–12.
- Newton, P.E., 2007a. *Evaluating Assessment Systems*. Paper 1. Submission to the Education and Skills Select Committee Inquiry Into Testing and Assessment. Qualifications and Curriculum Authority, London.
- Newton, P.E., 2007b. Clarifying the purposes of educational assessment. *Assess. Educ. Princ. Pol. Pract.* 14 (2), 149–170.
- Newton, P.E., 2008. *Monitoring National Attainment Standards: A Collection of Working Papers*. Ofqual/08/3916. Office of the Qualifications and Examinations Regulator, London.
- Newton, P.E., 2010. Educational assessment—concepts and issues: the multiple purposes of assessment. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, Oxford, pp. 392–396.

- Newton, P.E., 2012a. Validity, purpose and the recycling of results from educational assessments. In: Gardner, J. (Ed.), *Assessment and Learning*, second ed. SAGE, London, pp. 264–276.
- Newton, P.E., 2012b. Clarifying the consensus definition of validity. *Measurement* 10 (1–2), 1–29.
- Newton, P.E., 2017a. An Approach to Understanding Validation Arguments. Ofqual/17/6293. Office of Qualifications and Examinations Regulation, Coventry.
- Newton, P.E., 2017b. There is more to educational measurement than measuring: the importance of embracing purpose pluralism. *Educ. Meas.* 36 (2), 5–15.
- Newton, P.E., Baird, J., 2016. Editorial: the great validity debate. *Assess. Educ. Princ. Pol. Pract.* 23 (2), 173–177.
- Newton, P.E., Shaw, S.D., 2013. Standards for talking and thinking about validity. *Psychol. Methods* 18 (3), 301–319.
- Newton, P.E., Shaw, S.D., 2014. *Validity in Educational and Psychological Assessment*. SAGE, London.
- Newton, P.E., Shaw, S.D., 2016. Disagreement over the best way to use the word “validity” and options for reaching consensus. *Assess. Educ. Princ. Pol. Pract.* 23 (2), 178–197.
- Nitko, A.J., 1989. Designing tests that are integrated with instruction. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education/Macmillan, New York, pp. 444–474.
- Nitko, A.J., Brookhart, S.M., 2007. *Educational Assessment of Students*, fifth ed. Pearson Education, Upper Saddle River, NJ.
- Odell, C.W., 1928. *A Glossary of Three Hundred Terms Used in Educational Measurement and Research*. University of Illinois, Urbana, IL.
- Pellegrino, J.W., Chudowsky, N., Glaser, R., 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academy Press, Washington, DC.
- Schmeiser, C.B., Welch, C.J., 2006. Test development. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger Publishers, Westport, CT, pp. 307–353.
- Scriven, M., 2002. Assessing six assumptions in assessment. In: Braun, H.I., Jackson, D.N., Wiley, D.E. (Eds.), *The Role of Constructs in Psychological and Educational Measurement*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 255–275.
- Shepard, L.A., 1979. Purposes of assessment. *Stud. Educ. Eval.* 5 (1), 13–26.
- Shepard, L.A., 1993. Evaluating test validity. *Rev. Res. Educ.* 19, 405–450.
- Stevenson, P.R., 1922. Preliminary steps in initiating a testing program. *Educ. Res. Bull.* 1 (17), 155–159.
- Thorndike, R.M., Thorndike-Christ, T., 2010. *Measurement and Evaluation in Psychology and Education*, eighth ed. Pearson Education Inc, Boston.
- Tyler, R.W., 1942. General statement on evaluation. *J. Educ. Res.* 35 (7), 492–501.
- Zumbo, B.D., Forer, B., 2011. Testing and measurement from a multilevel view: psychometrics and validation. In: Bovaird, J.A., Geisinger, K.F., Buckendahl, C.W. (Eds.), *High-Stakes Testing in Education: Science and Practice in K-12 Settings*. American Psychological Association, Washington, DC, pp. 177–190.

Implementation of Assessment for Learning

Fabienne van der Kleij^a and Christopher DeLuca^b, ^a School of Education, University of Queensland, Brisbane, Queensland, Australia; and ^b Faculty of Education, Queen's University, Kingston, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Introduction	147
AfL: what is it?	147
Implementing AfL: what we know	148
The edge of implementation: two cases of large-scale implementation	148
The Australian case	149
The Canadian case	150
Learning from the cases	152
Future implementation	152
Acknowledgment	153
References	153
Further Reading	154

Introduction

This article is about the implementation of Assessment for Learning (AfL). First, we will discuss definitions of AfL. We will then provide a brief literature review to highlight the potential of AfL, as well as common pitfalls in implementation based on previous research. This literature review will close by providing an analysis of cutting-edge studies focused on successful large-scale implementation of AfL across contexts. Specifically, we will turn to discussing findings from two recent projects conducted in Australian and Canadian schools respectively. This discussion will illustrate some of the most recent evidence and ongoing issues in implementing AfL. Our interest in this chapter is to consider next steps for AfL implementation: What is needed for wider implementation? What systemic pitfalls should be avoided? What gaps in theory exist that limit current AfL uptake? Accordingly, this article concludes with a discussion on potential ways forward in facilitating implementation of AfL in schools.

AfL: what is it?

The term Assessment for Learning gained popularity after the Assessment Reform Group (ARG) proposed it in 1999 and 2002. The ARG was a UK-based group of assessment researchers which aimed to provide evidence-informed recommendations for educational policy and practice in relation to classroom assessment, teaching and learning. Their work built on the seminal [Black and Wiliam \(1998a,b\)](#) review, which showed substantial potential for classroom assessment and feedback (i.e., formative assessment) to enhance student learning outcomes. The ARG claimed that while these research findings were commonly accepted, “the development of practice in this area will need a concerted policy-making push” ([ARG, 1999](#), p. 1). Further, they identified that the term “formative assessment” was often misinterpreted in a way that did little to support student learning. To overcome this issue, they coined a new term for using assessment to promote learning: *Assessment for Learning* (AfL).

The ARG defined AfL as: “the process of seeking and interpreting evidence for use by learners and their teachers to decide where learners are in their learning where they need to go and how best to get there” ([ARG, 2002](#), p. 2). AfL was described as being in direct contrast to *assessment of learning*, which focuses on grading and reporting. To facilitate translation of research evidence to policy and practice, the ARG published a leaflet in 2002 which provided a definition of AfL and outlined 10 evidence-informed principles for classroom practice. The 10 principles state that AfL: (1) is part of effective planning; (2) focuses on how students learn; (3) is central to classroom practice; (4) is a key professional skill; (5) is sensitive and constructive; (6) fosters motivation; (7) promotes understanding of goals and criteria; (8) helps learners know how to improve; (9) develops the capacity for self-assessment; and (10) recognizes all educational achievement (p. 2). Each of these principles was seen as essential to successful implementation of AfL.

In more recent literature, AfL is often described as focusing on a set of five interrelated key strategies (e.g., [Heitink et al., 2016](#); [Wiliam and Thompson, 2008](#)):

- Clarifying and sharing learning goals and criteria for success;
- Gathering evidence of student learning by using informal and formal assessment techniques, such as questioning;
- Providing feedback to students to move learning forward;
- Enhancing student autonomy and self-regulation by engaging them in peer and self-assessment.

When used in combination, these strategies have significant potential to enhance student learning outcomes, motivation and autonomy by (1) making learners aware of learning goals and what quality work looks like, (2) encouraging students to generate feedback about their own and/or their peer's progress in relation to these goals, and (3) assisting students in taking action based on

feedback (Hawe and Dixon, 2014; Hawe and Parr, 2014; James and McCormick, 2009; Sadler, 1989; Wiliam and Thompson, 2008). As students become more autonomous and self-regulating, they become increasingly less reliant on teacher feedback, resulting in more sustainable classroom feedback practices (Hawe and Parr, 2014).

A key point of discussion in the literature is the nature of formative assessment versus AfL. In fact, there is still much discussion about the meaning of both terms (e.g., Wiliam, 2011, 2017). Many have used these concepts as synonymously (e.g., Brookhart, 2019). Yet others have argued that these concepts overlap but also have several distinct features (Swaffield, 2011; Van der Kleij et al., 2015). Although many assessments can be used formatively, AfL is about the day-to-day classroom assessment processes that directly promote student learning. This means that students are the direct beneficiaries of AfL, which is not necessarily the case in formative assessment (Swaffield, 2011). Moreover, they are active agents within the assessment process (i.e., users of assessment information) to benefit their own and others' learning progress. Accordingly, AfL has been termed a "participatory pedagogy" as it draws heavily on social-constructivist principles. In AfL, students learn within communities of practice by actively using feedback (and other assessment strategies) from interactions with themselves, peers, and teachers to develop their knowledge, abilities and skills (Willis, 2010).

Implementing AfL: what we know

Over the past three decades, AfL has been widely promoted in research, policy and practice across the globe. Despite widespread promotion and substantial efforts to translate and disseminate research evidence to policy and practice, many challenges in AfL implementation have been reported. We will briefly discuss predominant challenges reported in the literature.

Although the 2002 ARG definition may seem straightforward, it has come to be interpreted in a variety of ways. One issue is that "the way in which the words are interpreted and made manifest in educational policy and practice often reveal misunderstanding of the principles, and distortion of the practices, that the original ideals sought to promote" (Klenowski, 2009, p. 263). AfL has often been interpreted in a simplistic manner, what has come to be known as implementation of the letter rather than the spirit of AfL (DeLuca et al., 2019; James and McCormick, 2009; Marshall and Drummond, 2006), resulting in practices that have little resemblance to the evidence-informed strategies the ARG sought to promote. For example, Swaffield (2011) reflects on the erroneous interpretation of the original AfL definition as "at what level is the student, what is the [sic] next in the sequence and which learning objectives need to be targeted" (p. 439). This interpretation has resulted in practices with a rigid focus on pre-determined learning outcomes, as opposed to valuing all learning outcomes as per the original AfL definition. This issue is no doubt related to the competing and at times contradictory requirements of AfL versus assessment for accountability (Baird et al., 2014; Cumming et al., 2019).

A second linked challenge observed in the research is the compatibility of AfL with the dominant summative mode of assessment operating in many school contexts. Tied to accountability regimes and the longstanding grading traditions in education, summative assessment often emphasizes tested achievement rather than more comprehensive learning. Several studies have pointed to the challenge of implementing AfL within a summative orientated assessment culture (Baird et al., 2014; Cumming et al., 2019; Shepard, 2000). In instances where AfL prevails, it is often through slow, intentional, and school wide commitments to AfL pedagogies and mandates, where feedback rather than grades is prioritized as central to education (Andersson and Palm, 2017; Hopfenbeck et al., 2015; Marshall and Drummond, 2006). Research evidence suggests that school-wide reform efforts are needed to change the predominant assessment culture in a school in a way that challenges entrenched assessment values, routines and priorities among all stakeholders, including students and parents.

Successful efforts toward AfL implementation—those that promote a spirit orientation of AfL in classrooms—show significant commitment to teacher professional development (Heitink et al., 2016; Hopfenbeck et al., 2015; Lee et al., 2019). Consistent across these studies is that teachers tend to (a) learn collaborative and reflect on their assessment practices in groups; and (b) have administrative support for their professional learning in assessment.

When the spirit of AfL is implemented across classrooms and schools, there is an authentic and embedded use of strategies to promote student learning, autonomy and agency (James and McCormick, 2009; Marshall and Drummond, 2006). Assessment becomes the mode of operandum for teaching and learning. Students ask for feedback and require it to move forward in their learning. Student work (i.e., assessments and learning progression) are highly visible to students, parents, and the learning community to support a continuous conversation about growth and progression. Ultimately, teachers who implement the spirit of AfL lead others in their school both by modeling effective implementation strategies but also through their ongoing professional learning. They continue to ask questions and seek feedback on their own practice as they internalize the principles of AfL to continuously improve their practice (DeLuca et al., 2019). These findings correspond with findings from the broader school effectiveness literature, which highlight the criticality of teacher self-reflection, peer learning and continuous professional learning (Hopkins, 2013; Louis and Lee, 2016).

The edge of implementation: two cases of large-scale implementation

This section describes two cases of large-scale implementation of AfL. This discussion will illustrate some of the most recent evidence and ongoing issues in implementing AfL. The cases will be described individually, followed by a synthesis of what can be learned from these cases.

The Australian case

Assessment for Learning has been endorsed at the national and state/territory levels over the past decades. Despite widespread official endorsement, research suggests that AfL implementation across Australian schools is variable (Birenbaum et al., 2015; Cumming et al., 2019; Van der Kleij et al., 2018). Specifically, researchers have expressed concerns over consistency of practical implementation support. In addition, there are concerns over the nature of AfL implementation in light of accountability pressures. It is against this background that the ARC Linkage project “Improving student outcomes: Coaching teachers in the power of feedback” project (2017–2019) was initiated.

The project examined the impact of a school-wide professional learning intervention in feedback. The intervention was underpinned by a student-centered feedback model (SCFM; Brooks et al., 2021c). Thirteen Queensland state primary schools were involved as partners in the three-year project, which was contextualized in English writing. Each project year focused on a different student Year level, starting with Year 5 (students aged 10–11) and ending with Year 3 (students aged 8–9). The project involved school leaders (principals, deputy principals, and heads of curriculum or leaders in coaching roles) and a sample of at least three teachers in each school per project year.

Within each project year, the intervention took place through externally facilitated school-based collaborative sessions over a six-month period. In these sessions, the external facilitators and school leaders guided teachers through different aspects of the SCFM. The sessions involved discussion of research evidence, discussion of practice in light of this evidence, as well as practical activities, firmly grounded in curricular and assessment planning. Strategies were designed to be effective yet as simple as possible for ease of implementation. Teachers were encouraged to trial strategies in their classrooms using an approach that suited their particular context. The collaborative sessions were followed by scheduled implementation-focused discussions between school leaders and teachers. The external facilitators provided continued on-demand support to school leaders via phone or email in between sessions. The intention was for school leaders to learn alongside teachers while gradually increasing their capability to lead the professional learning within their school.

Although the project name suggests a primary focus on feedback, the intervention comprehensively reflected the spirit of AfL (James and McCormick, 2009; Marshall and Drummond, 2006). The SCFM (Brooks et al., 2021c) which underpinned the project foregrounds the active role of students in realizing the potential of feedback to improve learning outcomes. It places a key focus on enhancing capacity for student autonomy in assessment through self-regulated learning. Students are at the heart of the SCFM, which is made up of two parts, detailed next.

(1) The model’s inner wheel (Brooks et al., 2021c) focused on clarifying success (feeding up), checking in on progress (feeding back) and promoting improvement (feeding forward); based on Hattie and Timperley (2007). These processes capture all strategies typically associated with AfL (Heitink et al., 2016; Wiliam and Thompson, 2008). Intervention strategies for clarifying success included co-construction of success criteria by teachers and students and providing models demonstrating different levels of quality. Checking in on progress involved formal and informal assessment. Importantly, these processes were tightly connected with clarifying success, which enabled identification of key focal quality features. Promoting improvement focused on identifying and enacting strategies and next steps for improvement. All aspects of the feedback cycle involved teachers as well as students, through scaffolded peer- and self-assessment and feedback.

(2) The model’s outer wheel focused on the broader conditions for effective feedback, such as the school culture and teacher feedback perceptions (Brooks et al., 2021c). For example, the intervention focused on shifting school leader and teacher perceptions of feedback as corrective information that teachers provide to students, to jointly constructed notions of what success looks like, how students are progressing, and what next steps are to promote improvement. In addition, the intervention aimed to bring about a shift in the schools’ learning cultures by placing a core focus on continuous improvement and feedback as a powerful learning tool.

Project evaluation drew on various qualitative and quantitative data sources, including focus group and individual interviews with school leaders, teachers, and students, as well as classroom observations, field notes, teacher and student self-report surveys and student assessment scores. Each project year had a slightly different evaluation focus to enable comprehensive analysis of evidence and triangulation of findings. It is beyond the scope of this article to discuss all findings in depth (a full list of papers based on this project is available from the first author upon request). Rather, key findings pertaining to AfL implementation will be highlighted here.

Overall, the results showed substantial potential for this type of professional learning. Although shifting perceptions of feedback from transmissive to student-centered took a considerable amount of time, many school leaders identified that the impact of the professional learning on classroom practice had been “massive”. For example, school leaders reported that the intervention had resulted in stronger alignment of teaching and feedback to student learning needs as aligned with curricular assessment standards. Many teachers reported that the intervention had substantially changed their classroom practices. One teacher reflected that they now used strategies that were “much more effective; and really simple and easy to use.” Although teachers were initially hesitant to open up their classrooms for peer observation, such observation was highlighted as one of the most powerful levers for collaboratively bringing about changes in classroom practice.

Findings across all project years highlighted the criticality of clarifying and co-constructing of assessment standards and criteria. Through the intervention, teachers were challenged to jointly construct more in-depth understandings of curriculum standards and success criteria (Brooks et al., 2021c). For example, school leaders and teachers highlighted the value of collaborative discussions focused on the curriculum and pedagogy (Brooks et al., 2021b). Another school leader reported that teachers developed a better understanding of the nature of learning progression in English writing, which resulted in teachers better tailoring their lessons to

student needs: “checking-in in formative ways, along the way; and then tweaking and refining and moving their teaching more regularly.” Having a better understanding of curriculum requirements also resulted in stronger teacher agency and more meaningful curriculum implementation.

Further, in-depth teacher understanding of curriculum standards and success criteria was foundational to nurturing student agency and self-regulated learning, through processes such as co-constructing success criteria with students (Brooks et al., 2021c). Through the intervention, teachers were encouraged to co-construct the most important success criteria with students. One teacher reflected: “we have cut down the success criteria and made it really—like, instead of having all those huge success criteria for them, it’s two/three things, ‘This is what’s important’”. Student understanding of success criteria was foundational to activating them in the assessment process, through the use of models and peer and self-assessment. These enhanced understandings also fostered self-regulated learning, enabling students to identify how they were progressing and what next steps they needed to take to improve, providing a motivational mechanism. For example, one student reflected on clarity of criteria for the A-E curriculum achievement levels: “A, B and C’. And they have got certain things in them. And if you meet those things, that’s the [sic] great ... and trying to get it better every time.”

Another key finding based on data from the second intervention year was students in the intervention group made substantially larger achievement gains than those in the comparison group (Brooks et al., 2021a). These findings add important insights to the evidence-base on AfL, demonstrating the benefits of a genuinely student-centered approach. In addition, these findings provide support for the notion that AfL can be beneficial even for young students. When given adequate scaffolding tools and support, students as young as eight were capable of self-regulating their learning and effectively using self- and peer assessment (Brooks et al., 2021c). For example, one Year 3 teacher reflected: “I honestly did not think that they would give very good feedback to each other. I didn’t think they would; and they have.” Seeing the impact of student-centered practices on student learning propelled teacher buy-in and commitment to the project.

Although the project overall was impactful, the extent to which the intervention resulted in practices that can be classified as the spirit of AfL varied across schools and classrooms. There are many possible reasons for this. The main reasons are discussed next.

Teachers had different levels of knowledge and experience prior to the intervention, with some more open to trialing student-centered strategies than others. Of note, activation of students is dependent on teacher willingness and capacity to step away from the traditional model of the teacher as an expert who tells students what to do. A teacher reflected: “I usually try and give them leading questions. So I might say to them, ‘How do you feel when you are reading this?’, ‘How would someone else feel when they were reading this?’ I sort of tend to try and make them more active, rather than me just saying, ‘You need to go and work on this.’” Further, some teachers reported explicitly communicating expectations for the learning culture to students, for example stating, “that they are in charge of their learning and that it’s not just me standing at the front and telling them what to do.”

Most school leaders participated across the three project years, resulting in a substantial expansion of their knowledge of effective feedback. However, the extent to which the intervention resulted in school leader understanding of fundamental ideas of AfL was variable; some leaders continued to focus on strategy compliance rather than fostering genuinely student-centered classroom practices. Further, there was considerable variation in school leader capacity and confidence to co-lead the professional learning, with some identifying a perceived lack of leadership skills.

A core focus on collegial support and collaboration appeared to have fostered implementation, with some differences observed in collaborative practice across schools. These targeted opportunities to collaborate were made possible by school leader commitment to the project. Interestingly, teacher and school leaders in two school pairs reported having benefited from inter-school collaboration. A related finding was that having developed a common language to talk about processes surrounding learning at the classroom and school levels appeared to have strengthened implementation.

At the conclusion of the project, schools appeared to be at different stages of scaling up AfL practices to become school-wide and embedded across the curriculum, with each school taking a unique approach. Schools where AfL became a core focus of collaborative planning processes appeared to have most success in gradually scaling up implementation. Participants perceived that the intervention’s focus on evidence-informed practice as embedded within the writing aspect of the English curriculum aided implementation. Yet, it was up to school leaders and teachers to broaden implementation to other classrooms and curriculum areas; one of the key remaining challenges at the conclusion of the project. Interestingly, some teachers reported that students became highly accustomed to the new student-centered practices, resulting in changed student pedagogical expectations. A such, a sense of students holding teachers accountable for broader implementation of AfL emerged.

The Canadian case

The second case of large-scale AfL implementation is from Ontario, Canada where two school districts participated in a three-year Ministry of Education endorsed professional learning program aimed at building systemic capacity in AfL, and specifically teachers’ adoption of AfL strategies across elementary and secondary classrooms. Within the Ontario context, there is a robust policy framework that articulates assessment as a core element of teaching and learning. Within the guiding document, *Growing Success: Assessment, Evaluation and Reporting in Ontario Schools* (Ontario Ministry of Education, 2010), assessment is described as primarily serving the improvement of student learning with definitions and practices articulate for assessment of, for, and as learning.

Within this specific professional learning case study, teachers and school leaders, were focused on deepening their understandings and pedagogical commitments to AfL as articulated in the *Growing Success* (2010) document and as supported by the *Assessment Reform Group’s* (2002) initial definition and further operationalized through *Wiliam and Thompson’s* (2008) five

assessment-based pedagogies as outlined above (see the *What is AfL?* section). Within this broad AfL framework, teachers and school leaders pursued specific learning goals that were relevant to their instructional and school contexts, to engage in cycles of inquiry.

The two partnered school districts, one public and one Catholic, were in the same geographic region and involved a total of 12 schools. Throughout the three-year professional learning project, a total of 12 school principals and 88 teachers were involved (51 from one district and 37 from another). The professional development project was purposefully designed to leverage several key aspects known in the literature to support teacher learning, detailed next.

First, teacher learning was supported by a variety of planned and resourced methods including whole group learning sessions that provided structured learning and reflection time, smaller group learning sessions that allowed for in-school monitoring and support of teacher learning, access to school district assessment coach to support teachers' individual AfL planning and implementation, and structured teacher reflections to enable tracking of performance and growth. In-school teacher learning was also supported by an instructional rounds model where teachers would observe one another implement AfL strategies and provide supportive feedback for improvement. The process and impact of this strategy in this professional learning program is further detailed in [DeLuca et al. \(2015\)](#).

Second, teacher learning of AfL principles was paralleled by administrator (i.e., school principal) learning of AfL. While teachers were learning to implement AfL in their classrooms, their school principals were learning to implement AfL as strategy to support their teachers' development (i.e., as a strategy for leadership and staff development). This parallel learning was intended to not only lend administrator support to the initiative and to teachers' professional development was also intended to take "formative assessment schoolwide" by establishing meaningful learning partnerships between school leaders and teachers to form a genuine learning community with and across schools in the two districts ([Moss and Brookhart, 2009](#), p. 137).

Finally, a key advantage of the program was that it was supported by the Ministry of Education and the two school districts for three years, providing prolonged exposure to the various resources, professional learning structures, and learning goals of teachers, school leaders, and school districts. As previous research has repeatedly shown, AfL is not a practice that can be cultivated and internalized quickly, either among teachers or their students. As such, the three years allowed for teachers to develop skills, reflect on their progress and setbacks, and see the impact of their efforts on students and school assessment cultures. Specifically, data were collected through a variety of qualitative and quantitative measures to trace teacher learning in AfL over the three years. Instructional rounds sessions and classroom instruction were observed to determine how classroom assessment practices changed for 18 focal teachers. Semi-structured interviews were also conducted with these teachers repeatedly to detail their changing conceptions of assessment and learning progression. All teachers were invited to complete yearly surveys that focused on their professional development in assessment and changes to their assessment practice. These surveys were coupled with student surveys on classroom assessment practices (see [DeLuca et al., 2018](#)). In a companion paper, we detail the professional learning continua teachers experienced throughout the three-year project to demonstrate the growth potential in such a model (see [DeLuca et al., 2019](#)).

In considering the systemic challenge of AfL implementation, as posed in this article, from this case we draw out three important themes. The first of these themes is professional accountability. Embedded in this case were multiple points and people to whom teachers were accountable for their learning, including themselves and their students. Defined by [LaPointe-McEwan \(2018, p. 115\)](#) as "constructive in nature" professional accountability is about owning one's learning goals and narrating a learning journey to key stakeholders to fuel continued development through feedback and encouragement. Throughout this project, teachers, perhaps more so than school principals, were able to demonstrate their learning in action through the instructional rounds process and reflect on their progress in small and large group learning sessions. The district assessment coach also served as a professional accountability partner to sustain teacher learning throughout the project. As the project progressed teachers commented that students also began to expect AfL activities and practices, asking for peer, self, and teacher feedback; thereby serving as yet another source of professional accountability.

The second theme that contributed to program success was systemic commitment to AfL learning. While teachers were learning alongside their peers, school leaders (i.e., principals), and district assessment coaches, Ministry of Education staff also participated in large group learning sessions, typically as co-facilitators but also as co-learners. As such, from ministry to teacher there was continuous learning about AfL across various educational roles within the system, which provided fertile ground for deep conversations, reflections, and connections. Further to these stakeholders' engagement in the project was the research team who contributed an additional layer of AfL learning to the project, adding layers of commitment to a professional learning project substantiates its value and increases energy for participation. In this case, there was a purposeful move from schoolwide learning to systemic learning, which motivated learning for those involved in the project.

Finally, success for many teachers and administrator came when they realized that AfL applied equally to their learning as to their students' learning. To this end, the program was predicated on high levels of non-evaluative feedback, a hallmark of the instructional rounds approach ([City et al., 2009](#); [Marzano, 2011](#)). When teachers began utilizing the AfL principles intuitively to guide their own professional learning and to provide meaningful feedback to one another, they could see its application more clearly to their classroom teaching context. Their own professional learning became yet another context to practice AfL and deepened their AfL skills, knowledge, and questions. Fundamentally, as the project progressed two components of AfL began to take hold: (1) teachers and school principals became better able (and more comfortable) at providing constructive and valuable feedback; and (2) their questions about AfL became more nuanced, critical, and generative. Both of these components led to greater gains in learning.

Learning from the cases

Although both cases describe unique projects across different international contexts, there was a substantial degree in commonality of focus and outcomes. Hence, a cross-case analysis provides insights into common findings to show what worked to facilitate implementation of AfL.

Findings from both cases highlight the importance of considering the temporal dimension of AfL implementation. Both projects were conducted across a three-year period and showed that participants gradually deepened their understandings and shaped their practices over this period of time. This finding has important implications for anyone willing to implement AfL. It shows that the implementation process is likely to be gradual, for teachers as well as for students. At times, implementation may also require taking steps back before moving forward. For example, the Australian case highlighted the need for teachers to take a step back to more closely examine curriculum expectations, before clarifying what success looks like with their students. Although all of this takes time, it vitally underpins successful AfL implementation.

Further, both cases highlight the interplay between actors within different layers of the education system. Although AfL ultimately concerns students and happens in day-to-day classroom practice, these cases show that implementation requires a focus beyond the classroom. We conclude that AfL needs to be viewed as embedded within a multilayered community of practice. Several key findings pertain to this multilayered contextual dimension.

Across both cases, implementation of AfL reflected the very nature of AfL, with school leaders, teachers and students learning alongside one another. Although school leaders and teachers were at different qualitative stages of implementation, they continuously reviewed and refined their practices, taking it one step at a time. This was exactly the approach taken with students. As illustrated by the Australian case; rather than showing students a model of high quality, teachers worked with students to work out where they were currently at, and aimed to progress their learning by helping them identify next steps on the road to improvement. School leaders, teachers and students all gradually became more proficient and confident in enacting AfL.

Interactions between actors in different layers of the education system played important catalyzing roles in AfL implementation. Establishing a learning culture and common language to talk about learning was one important driver of implementation. Further, professional accountability functioned as an important lever for AfL implementation by enhancing commitment of actors at different layers of the system. Teachers and school leaders were accountable to each other, external partners, as well as students in their schools. In these cases, external experts played an important role in initiating conversations using a shared language, challenging previous conceptions and holding teachers and school leaders accountable for implementing learnings from the professional learning intervention. Yet, we believe that even outside a funded professional learning intervention, the same can be achieved when there is a clearly articulated vision and articulation of intention across the school.

As previously identified, teachers needed time to figure out what AfL might look like in their classrooms, which took time. The cases showed that a deep understanding of why AfL results in improved student learning fostered teacher agency and depth of implementation. Teachers particularly valued being able to see what a particular strategy might look like in action, either through expert demonstration or peer observation and discussion. Middle leaders and coaches played a vital role in connecting theory and practice. These findings further highlight the importance of school-wide and/or systemic support structures. For example, principals in both cases set aside time for teachers to learn about AfL, reflect on practice and plan for implementation. This signals to teachers that engaging in in-depth AfL practice is valued and meaningful. The cases thus provide compelling evidence of the importance of school leader commitment and the provision of organizational strategic support. Importantly, collaboration does not need to be limited to the bounds of the school; both cases highlight the potential of cross-school collaboration.

Across both cases, teacher learning and implementation was propelled by seeing the impact of changed classroom practices on student motivation, agency and learning outcomes. As argued by Hattie (2012), seeing the impact of one's practice on students is one of the most important drivers of teacher motivation to continuously improve their practice. Coupling professional learning with active research and evaluation allowed for teachers and school leaders in these two cases to see the gains their learning was having on students. In many ways, this partnership between research and professional learning leveraged a key AfL strategy; to make visible the learning for the learner.

The cases also highlighted the contextual nature of AfL within subjects. These findings attest to the notion that generic principles are insufficient to meaningfully guide AfL practice (Hodgen and Marshall, 2005; James and McCormick, 2009; Van der Kleij et al., 2018); teachers need more specific guidance which shows what this might look like in their or across different curriculum areas. Within the Canadian case we saw that teachers had latitude to narrow their focus of learning to their own context of teaching. As such, a primary task of their professional learning was to translate, with support, general AfL principles into context and disciplinary specific actions. Going forward it may be useful for researchers and the policy community to provide teachers with greater granularity in the application of AfL principles across teaching and disciplinary contexts.

Future implementation

In writing this article, we set out to consider what was required for systemic adoption and implementation of AfL. From the two cases presented, and corroborated by the literature, it is clear that several structures influence AfL uptake across schools. At the heart of these factors is a drive to change the culture of assessment in schools—largely from summative motivators toward feedback-driven teaching and learning (Laveault and Allal, 2018). Changing assessment cultures takes time and dedication by stakeholders

across system levels - teachers, school leaders, and policy makers. Involvement of students and parents is also critical, and will often follow once systemic changes have been achieved. Centrally, the structures known to shape systemic AfL implementation are:

- (a) Facilitating co-learning and co-commitment to learning by classroom teachers, school leaders, and where possible, policy makers;
- (b) Facilitating cycles of application, observation, and reflection (i.e., providing opportunities to implement theory in practice with feedback and to reflect on implementation practices);
- (c) Enabling teacher autonomy to narrow and specify learning goals to their unique teaching contexts while maintaining broad alignment with systemic goals of school districts;
- (d) Leveraging systemic and district policies that support professional learning aims and that build coherence across schools' and teachers' professional learning goals;
- (e) Coupling professional learning initiatives with research and evaluation activities to make visible for teachers and school leaders the effects of teacher change on students' and school assessment culture; and
- (f) Investing in longitudinal professional learning programs that are resourced for proactive and sustained teacher development.

Through such structures, there is evidence that communities of professional learning emerge, and new assessment practices and cultures flourish within schools.

What remains less well established in the literature, and which requires greater action by both educational researchers and policy makers, is (a) the narrowing of broad AfL principles to diverse teaching contexts and disciplines, which would enable teachers to gain greater clarity on the application of AfL across teaching contexts; (b) the charting of school leader and school system progressions of AfL adoption and growth in order to make visible how entire systems ebb and flow as they strive toward greater implementation; and (c) the development of robust policies aimed at teacher professional learning that make progressive provisions for teachers' intensive training in key competency areas, such as assessment. To this final point, building more consistent policy narratives that establish congruence between various assessment mandates across the system (i.e., formative, summative and large-scale) and that couples such mandates with provisions for teacher professional learning would enable a progressive policy response that functions as a productive foundation for systemic AfL implementation.

In order to move forward with AfL adoption at scale, it is clear that a multi-layer, strategically resourced approach is required that involves key stakeholders in progressive and sustained learning. Additionally, policy and research must continue to advance models that make visible what systemic AfL looks like in terms of resources, policies, and professional learning commitments.

Acknowledgment

The "Improving student outcomes: Coaching teachers in the power of feedback" project (2017–2019) was supported by the Australian Government through the Australian Research Council [LP160101604].

References

- Andersson, C., Palm, T., 2017. The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learn. Instr.* 49, 92–102. <https://doi.org/10.1016/j.learninstruc.2016.12.006>.
- Assessment Reform Group, 1999. *Assessment for Learning: Beyond the Black Box*. http://assessmentreformgroup.files.wordpress.com/2012/01/beyond_black_box.pdf.
- Assessment Reform Group, 2002. *Assessment is for Learning: 10 Principles. Research- Based Principles to Guide Classroom Practice*. Assessment Reform Group. http://assessmentreform-group.files.wordpress.com/2012/01/10principles_english.pdf.
- Baird, J.-A., Hopfenbeck, T.N., Newton, P., Stobart, G., Steen-Utheim, A.T., 2014. *Assessment and Learning: State of the Field Review*. Knowledge Centre for Education. https://www.forskningssradet.no/servelet/Satellite?c=Rapport&cid=1253996755700&lang=en&pagename=kunnskapssenter_%2FHovedsidemal.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., 2015. International trends in the implementation of assessment for learning: implications for policy and practice. *Policy Futures Educ.* 13, 117–140. <https://doi.org/10.1177/1478210314566733>.
- Black, P., William, D., 1998a. Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5, 7–74.
- Black, P., William, D., 1998b. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan* 80, 139–148.
- Brookhart, S.M., 2019. Summative and formative feedback. In: Lipnevich, A.A., Smith, J.K. (Eds.), *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press, Cambridge, pp. 52–78.
- Brooks, C., Burton, R., Van der Kleij, F., Ablaza, C., Carroll, A., Hattie, J., Neill, S., 2021a. Teachers activating learners: the effects of a student-centred feedback approach on writing achievement. *Teaching and Teacher Education* 105, 103387. <https://doi.org/10.1016/j.tate.2021.103387>.
- Brooks, C., Burton, R., van der Kleij, F., Carroll, A., Hattie, J., 2021b. Towards student-centred feedback practices: evaluating the impact of a professional learning intervention in primary schools. *Assessment in Education. Principles, Policy and Practice* 28 (5–6), 633–656. <https://doi.org/10.1080/0969594X.2021.1976108>.
- Brooks, C., Burton, R., Van der Kleij, F., Carroll, A., Olave, K., Hattie, J., 2021c. From fixing the work to improving the learner: an initial evaluation of a professional learning intervention using a new student-centred feedback model. *Stud. Educ. Eval.* 68, 100943.
- City, E.A., Elmore, R.F., Fiarman, S.F., Teitel, L., 2009. *Instructional Rounds in Education: A Network Approach to Improving Teaching and Learning*. Harvard Education Press, Cambridge, MA.
- Cumming, J.J., Van Der Kleij, F.M., Adie, L., 2019. Contesting educational assessment policies in Australia. *J. Educ. Pol.* 34, 836–857.
- DeLuca, C., Chapman-Chin, A., Klinger, D.A., 2019. Toward a teacher professional learning continuum in Assessment for Learning. *Educ. Assess.* 24, 267–285.
- DeLuca, C., Chapman-Chin, A.E., LaPointe-McEwan, D., Klinger, D., 2018. Students' perspectives on assessment for learning. *Curric. J.* 29, 77–94.
- DeLuca, C., Klinger, D., Pyper, J., Woods, J., 2015. Building school district capacity in assessment for learning: a study on the effectiveness of an instructional rounds professional learning approach. *Assess. Educ. Princ. Pol. Pract.* 22, 122–139.

- Hattie, J., 2012. *Visible Learning for Teachers*. Routledge.
- Hattie, J., Timperley, H., 2007. The power of feedback. *Rev. Educ. Res.* 77, 81–112.
- Hawe, E.M., Dixon, H.R., 2014. Building students' evaluative and productive expertise in the writing classroom. *Assess. Writ.* 19, 66–79.
- Hawe, E., Parr, J., 2014. Assessment for Learning in the writing classroom: an incomplete realisation. *Curric. J.* 25, 210–237.
- Heitink, M.C., Van der Kleij, F.M., Veldkamp, B.P., Schildkamp, K., Kippers, W.B., 2016. A systematic review of prerequisites for implementing Assessment for Learning in classroom practice. *Educ. Res. Rev.* 17, 50–62.
- Hodgen, J., Marshall, B., 2005. Assessment for learning in English and mathematics: a comparison. *Curric. J.* 16, 153–176.
- Hopfenbeck, T.N., Flórez Petour, M.T., Tolo, A., 2015. Balancing tensions in educational policy reforms: large-scale implementation of Assessment for Learning in Norway. *Assess. Educ. Princ. Pol. Pract.* 22, 44–60.
- Hopkins, D., 2013. Exploding the myths of school reform. *Sch. Leader. Manag.* 33, 304–321.
- James, M., McCormick, R., 2009. Teachers learning how to learn. *Teach. Teach. Educ.* 25, 973–982.
- Klenowski, V., 2009. Assessment for learning revisited: an Asia-Pacific perspective. *Assess. Educ. Princ. Pol. Pract.* 16, 263–268.
- LaPointe-McEwan, D., 2018. *Exploring How Video-Informed Feedback Supports Embedded Educator Professional Learning Between Middle Leader Facilitators and Teachers*. Unpublished Doctoral Dissertation. Queen's University, Kingston, ON.
- Lee, I., Mak, P., Yuan, R.E., 2019. Assessment as learning in primary writing classrooms: an exploratory study. *Stud. Educ. Eval.* 62, 72–81.
- Louis, K.S., Lee, M., 2016. Teachers' capacity for organizational learning: the effects of school culture and context. *Sch. Effect. Sch. Improv.* 27, 534–556.
- Marshall, B., Drummond, M.J., 2006. How teachers engage with assessment for learning: lessons from the classroom. *Res. Pap. Educ.* 21, 133–149.
- Marzano, R.J., 2011. Making the most of instructional rounds. *Educ. Leadersh* 68, 80–82.
- Moss, C., Brookhart, M., 2009. *Advancing Formative Assessment in Every Classroom*. ASCD, Alexandria, VA.
- Ontario Ministry of Education, 2010. *Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools*. Queen's Printer for Ontario, Toronto.
- Swaffield, S., 2011. Getting to the heart of authentic assessment for learning. *Assess. Educ. Princ. Pol. Pract.* 18, 433–449.
- Sadler, R., 1989. Formative assessment and the design of instructional systems. *Instr. Sci.* 18, 119–144.
- Shepard, L., 2000. The role of assessment in a learning culture. *Educ. Res.* 29, 4–14.
- Van der Kleij, F.M., Cumming, J.J., Looney, A., 2018. Policy expectations and support for teacher formative assessment in Australian education reform. *Assess. Educ. Princ. Pol. Pract.* 25, 620–637.
- Van der Kleij, F.M., Vermeulen, J.A., Schildkamp, K., Eggen, T.J.H.M., 2015. Integrating data-based decision making, Assessment for Learning and diagnostic testing in formative assessment. *Assess. Educ. Princ. Pol. Pract.* 22, 324–343.
- William, D., 2011. What is assessment for learning? *Stud. Educ. Eval.* 37, 3–14.
- William, D., 2017. Assessment and learning: some reflections. *Assess. Educ. Princ. Pol. Pract.* 24, 394–403.
- Willis, J., 2010. Assessment for learning as a participative pedagogy. *Assess. Matters* 2, 65–84.
- William, D., Thompson, M., 2008. Integrating assessment with learning: what will it take to make it work? In: Dwyer, C.A. (Ed.), *The Future of Assessment. Shaping Teaching and Learning*. Lawrence Erlbaum Associates, New York, pp. 53–84.

Further Reading

- Laveault, D., Allal, L., 2018. *Assessment for Learning: Meeting the Challenge of Implementation*. Springer.

A more intelligent accountability—future directions

Sølvi Lillejord^{a,b}, ^a SLATE, The Center for the Science of Learning and Technology, University of Bergen, Bergen, Norway; and
^b Department of Education, Oxford University, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Aim of the chapter	155
The “Coleman report”	155
The accountability reform movement	156
Holding teachers accountable	157
Expert group on schools’ contribution to students’ learning	158
The quest for simplicity—human and social capital	159
What can be expected of teachers, given their working conditions?	160
A more intelligent accountability	160
Conclusion	161
References	161

For every complicated problem, there is a solution that is short, simple, and wrong.

H. L. Mencken (1880–1956).

Aim of the chapter

This chapter addresses political concerns raised about education worldwide the last decades. Many countries, including Norway, have developed *school contribution indicators*, statistical measures of how much “the school” contributes to students’ learning, or the schools’ “value added” (Adnett et al., 2002). As these indicators reveal that schools’ contributions differ, the Norwegian Government in 2019 appointed an Expert group on schools’ contribution. The group’s mandate stated that, compared to other nations, social differences are small in Norway, but also here, a systematic connection is found between students’ home background and their learning outcomes. The Expert group published two reports (Lillejord et al., 2020, 2021). The first gave the background for the interest in the school’s contribution to students’ learning and policy initiatives taken in relation to this. The second drew on data from school visits, where the Expert group had dialogs with school leaders, local administrators, teachers, and students, about how they engage in quality development. The chapter starts where the Expert group started, with the so-called Coleman-report, often labeled a “watershed” report (Barr and Dreeben, 1977; Kantor and Lowe, 2017) because it reformulated what was meant by equal education, or at least how it should be measured.

Economists used the Coleman-report to juxtapose schools to firms and introduced the concept *education production function*, equaling students’ learning outcomes to businesses’ profits. The quest for simplicity in this black-box-thinking was accompanied by traditional, somewhat outdated policy assumptions about teaching and how people learn. Based on the idea that investment in human capital will boost the knowledge economy, the OECD launched PISA in 2000, to prepare member states for increased global competition. The focus on human capital and cognition favored individual learning over the school’s social mandate. By interpreting human capital as individual cognitive skills, only half of educations’ *raison d’être* was emphasized, and this policy favored those students who needed it the least. In conclusion, it is argued that to hold teachers accountable for results of work processes that they do not control is a short-sighted, simple, and wrong solution. A better idea is to strengthen the teaching profession and educational knowledge at all levels.

The “Coleman report”

In 1966, sociologist James S. Coleman and colleagues published *Equality of Education Opportunity* (Coleman et al., 1966). The report was massive (Hill, 2017), with data from 4000 schools, 66,000 teachers, and almost 600,000 1st-, 3rd-, 6th-, 9th-, and 12th graders. It revealed a segregated US education system that favored white children in school facilities, class size, and the availability of advanced coursework; and race-based inequality on tests of academic achievement. What mattered more than the conventional measures of school quality were the socioeconomic backgrounds and educational aspirations of classmates, students’ sense of control over their own destinies, and, most importantly, the socioeconomic background of the individual child. In other words, the research team found small effects of school resources and facilities on students’ achievement, relative to those of family background.

These findings were misinterpreted as having shown that schools did not matter or had no effect on student achievement (Barr and Dreeben, 1977; Sørensen and Morgan, 2000). Fifty years later, Downey and Condrón (2016) found still no consensus among scholars regarding the relationship between schools and inequality. They asked if sociologists tended to perceive schools as engines of inequality and suggested a more balanced perspective because schools might reproduce or exacerbate some inequalities, while

compensating for others. Barr and Dreeben (1977) raised concern about research that examined school effects using what they called a quasi-production function formulation to examine the connections between school resources and achievement. In this tradition, from economics, schools were analyzed as firms which transform inputs (resources such as books, teachers' time, activities, and equipment) into outputs (such as increments of knowledge, change in attitude, or gains in achievement).

A driving force in the production function research was Eric Hanushek, senior fellow at Stanford University's think-tank Hoover Institution. In the wake of the Coleman-report, Hanushek (1971) argued that more knowledge was needed about the relationship between inputs and outputs of the educational process. His main goal was to estimate the relationship between educational output and variables which can be controlled by public policy, by addressing three fundamental educational policy questions: (1) do teachers count? (2) are schools operated efficiently? (3) what characteristics of teachers and classrooms are important? Based on equations such as this:

$A_{it} = f(A_{it*}, B_i^{(t-t*)}, P_i^{(t-t*)}, I_i, S_i^{(t-t*)})$ Hanushek (1971, p. 280) concluded that teaching experience and graduate education do not contribute to gains in student achievement scores. Later, Hanushek (1979, p. 355) described education as a "service" which transforms fixed quantities of inputs (i.e., individuals) into individuals with different quality attributes. Since the Coleman-report, he argued, the policy relevance of input-output studies has led to a rapid growth in number of studies with contradictory results. He therefore suggested, as a next step of statistical inquiry into the education process, the education production function (EPF), based on the following argument: All students of economics are introduced to the concept production function, that describes the firm's maximum output feasible with different sets of inputs. Firms are attempting to maximize profits, and a production function represents the maximum achievable output for given inputs. Hanushek described the production function as a powerful pedagogical tool since it provided a basis for describing efficient production in all industries, steel industry as well as education industry.

Hanushek found it unfortunate that EPF-studies were statistical analyses, with few measures of the quality of education, since the most pressing policy questions concern how to improve the quality of education. He regarded cognitive skills the chief measure of educational quality, probably the most important outcome of schooling in determining individuals' future success (Hanushek, 1979, 357) and formulated this definition of teacher quality: good teachers get large gains in student achievement for their classes; bad teachers are just the opposite (Hanushek, 2002, p. 3).

A recurring theme in his research was cognitive skills. Hanushek and Woessmann (2015) saw a direct link between a country's prosperity and the skills of its population and argued that nothing is more important to the future knowledge capital of nations than the cognitive knowledge, and skills, of those being educated today. They assumed a causal relation between the level of human capital acquired by individual students and a nation's economic growth. The greater investment in human capital, the larger gross domestic product (GDP). While this assumption was never tested, policymakers appear to have found the idea compelling.

Vanderberghe (1999) described the education production functions debate as problematic, because researchers used different proxies of educational achievement. Some referred to test results, others to wages. He also problematized human capital theory's "black box" approach to education because it assumes that schools and teachers have no objectives of their own, regards educational institutions as passive, mechanically responding to their "clients," and as not submitted to any constraints.

With increased globalization and the emerging knowledge society (Stehr, 1994), politicians' understanding of the purposes of education gradually shifted. Ydesen and Bomholt (2020) showed how the Organization for Economic Cooperation and Development (OECD) since its inception in 1961, promoted a global vision of education as a source of human capital needed to deal with social challenges and improve the economies of member-states. Education was a prerequisite for nations' "global survival" and global actors, such as the World Bank and the OECD, took the stage as educational experts, focusing on efficiency, and social equity. To oversee that educationalist did their utmost, *accountability* was a favored policy tool.

Economists understood knowledge as the cognitive capacity of the individual student to respond to tests and education as a tool for maximizing the outcomes of the available human resources. In this logic, higher test scores indicated more knowledge stored in individual students' brains. Instead, Stehr (2003), proposed to define knowledge as a *capacity to act*. This is more in line with how modern educational institutions prepare their students, better aligned with cognitive psychology (Sloman and Fernbach, 2017) and in congruence with how Coleman (1988) described the potential of social capital. As the next section shows, a central tenet of the accountability movement was that it *reduced* complexity—by ignoring it, by equaling schools to businesses and by reducing learning to individual, cognitive activities.

The accountability reform movement

Half a century ago, Dalton (1972) warned that any method of educational accountability was likely to be fraught with difficulties because time is spent on measurement and evaluation instead of on the quality of teaching and learning. While questions such as accountable for what, how much and to whom appear simple, they take on a complexity hitherto unmatched in education, he argued. The US report *A Nation at Risk* (National Commission on Excellence in Education 1983), initiated a period of mistrust in schools and introduced policies that supported traditional, teacher-led instruction and ignored the dual function of education—as an individual and social endeavor—to both prepare students for society and improve their learning outcomes. In an early account of US accountability reform, Kister (1993) described how states were required to adopt systems of outcomes-based standards and measures of performance. A decade later, Darling-Hammond (2004) reported that virtually all states were creating new curriculum frameworks and standards for student learning. Outcomes-based education was assumed to help students develop

essential occupational 21st century skills, such as higher-order thinking skills, and be prepared for the demands of the new economy.

In a review of research on accountability reform, Wang et al. (2006) identified a polarized debate around four controversies: Assessment-driven reform, standard-based assessment, assessment-centered accountability, and high-stakes consequences. In the US, standardized assessment, defined as a large-scale, externally developed, and mandated, uniformly administered, and scored evaluation of student learning, had characterized education reform from the late 1800s Committee of Ten, to the National Defense Education Act in the 1950s, the Elementary and Secondary Education Act of 1965, A Nation at Risk in 1983, the Goals 2000 in 1994, up to No Child Left Behind (NCLB) Act of 2001. Each emphasized standardized assessment as a reform catalyst and quality control mechanism. Wang et al. (2006) argued that it is a formidable task to set reasonably attainable standards that can be measured reliably and validly and can challenge all students. Few researchers have publicly defended standardized testing (e.g., Cizek, 2001; Phelps, 2005) and measurement professionals who know and make standardized tests have largely remained silent (Cizek, 2001, p. 20). Therefore, Wang et al. (2006) suggested that future test designs should involve teachers and cognitive-developmental and social psychologists.

With respect to who is holding whom accountable for what, differences are noted among legal, professional, or market accountability systems and administrative/bureaucratic accountability. Based on an assumption that information, its existence, and usage, is key to improving schools, O'Day (2002) studied how information travels through complex systems. She found that the combination of administrative and professional accountability presents a more promising approach for lasting and meaningful school reform than the outcomes-based bureaucratic accountability approach where the emphasis is on student outcomes, and consequences attached to performance levels.

The accountability movement, argued Ingersoll and Collins (2017), assumed that low-quality teachers and teaching explain inadequate school performance. Because teachers too often are not held accountable, they simply do what they want behind closed classroom doors. Tröhler (2017) labeled such unfounded assumptions about education, teachers and school leaders, *organized mistrust*. Onora O'Neill (2013) called for a more *intelligent* accountability, characterized by *intelligible* information that can help stakeholders make informed judgments. She argued that systems of accountability are *second order* ways of using evidence of the standard to which *first order* tasks are carried out for a great variety of purposes. The distinction between first- and second order constructs was initially developed by Schütz (1962). He labeled words and concepts used by actors *first-order constructs*, while the researchers' interpretation of the informants' explanations was second-order construct. Researchers try to ascribe meaning to actors' descriptions, meanings that reach beyond contexts, to make them more general so that they can be part of a theory (Aspers, 2009).

Holding teachers accountable

For decades, teachers have been held accountable by highly abstract information generated by standardized tests. Berliner (2020) explained how he suddenly realized that the OECD-developed Program for International Student Assessment (PISA), is merely another Standardized Achievement Test, because like other Standardized Achievement Tests (SATs) PISA adheres to certain principles of design, have similar correlates, and limits on the interpretations of the results obtained. A problem with standardized tests is that teachers are held accountable for data at an aggregated level, generated through equations that very few people understand. This abstract, second-order information has a summative, restrictive function that is not necessarily helpful for teachers who deal with people in complex, contextual, first-order situations.

In the USA, noted Ingersoll and Collins (2017), teachers have been blamed for the erosion of economic competitiveness, the decline in student academic achievement, teenage pregnancy, juvenile delinquency, a decline in morals, gender and racial stereotyping and discrimination, and on and on. Lillejord (2020) traced the mistrust in teachers, teacher education and the education system to the 1983 report *A Nation at Risk*, commissioned by the US president Ronald Reagan, and the UK election of Margaret Thatcher as Prime Minister in 1979. *A Nation at Risk* claimed that because too many did the minimum work necessary, the ideal of academic excellence as the primary goal of schooling was fading across US education. Both Reagan and Thatcher were proponents of *neoliberal* policies that applied the logic and rules of market competition on the public sector. The goal was to dismantle economic and social structures that obstructed the market and let the "cleansing waters" of marketplace competition eliminate "inferior" schools (Hess, 2001). An individualized, consumer-driven, free market choice approach that superseded the democratizing function of public schools, led to a commodification of educational knowledge (Ball, 2010).

The key discursive strategy in the UK, argued Beck (2008), was to delegitimize teachers through "discourses of derision" (Ball, 1990) and marginalize organized groups representing teachers. Exley and Ball (2011) described UK conservative education policy as founded on the neo-liberal belief in markets and a minimal state. The Conservatives believed in traditional, "old" teaching methods, "real subjects," detailed instruction over what should be taught in schools and surveillance from above. Parents were consumers of educational services and schools subject to market accountability. Pupils succeed thanks to ability, hard work and traditional teaching methods. Where they fail, it is because they, their families or their teachers have not tried hard enough.

Not just teachers and schools were under attack. In England, educational research was criticized for being fragmented, noncumulative, methodologically flawed; often tendentious and politically motivated (Biesta, 2007). It did not provide the answers policymakers needed or equip educational professionals with clear guidance for their work. In the US it was decided that large-scale effect studies were the only form of research eligible for government subsidies (Berliner, 2002). While researchers agreed that

more rigorous research was needed, they argued that there is still a need for correlational, descriptive, and other disciplined inquiry in education (Slavin, 2002). Lareau (2009) noted that politicians' emphasis on scientific *rigor* and references to a *gold* standard left little room to discuss the relevance of randomized controlled trials (RCTs) for education. This approach, she argued, cannot answer the most important questions in education, as advocates of RCTs have misunderstood the research model and misapplied it to the field of education. The idea was simple: when schools are being told what is proven by research to yield the best results or effects, they implement this practice. This simplistic idea, however, does not work, argued Biesta (2010). The introduction of standards and best practice influenced the training and daily work of principals, often practiced as a controlling mechanism that encouraged a narrow "tick box" approach (Cowie and Crawford, 2007).

To sum up, policymakers have tried to solve problems in education by applying a *deficit* perspective and an individualized understanding of education. They have treated schools as private businesses and assumed that schools lack the necessary knowledge and skills and are in desperate need of external assistance. The individualized deficit perspective, combined with organized mistrust, has separated people from each other and eroded the educational system's social fabric. An alternative approach for policymakers who see weaknesses in schools, is to *strengthen* education. Strengthening educational research and the education sector supports knowledge development and contributes to restoring trust among actors and between levels. This is elaborated in the next section, where findings from the Norwegian Expert group on schools' contribution is presented.

Expert group on schools' contribution to students' learning

The case used in this chapter is from a policy experiment (Lillejord et al., 2021). The groups' mandate was anchored in politicians' concern that in Norway, albeit to a lesser degree than in many other countries, family background is the most significant predictor for students' learning outcomes. Along with several other countries (Grek and Ydesen, 2021), Norway has developed *indicators* that measure the school's contribution to students' learning. These indicators show that schools contribute in varying degree to students' learning. On this background, the Government appointed an expert group with a mandate to generate knowledge about how local administrators and schools can enhance their systematic quality work.

A group with varied background and experience was nominated, to identify how schools and local administrators can work systematically to increase students' learning outcomes. The expert group was assigned two tasks: First, to establish a knowledge base on quality work in schools and publish a report on the findings. Second, to visit schools and have dialogs with teachers, students, school leaders and local administrators, to learn how they organize and engage in quality work. Along with Denmark, Finland, Iceland, and Sweden, Norway belongs to the Nordic region, where education is influenced by the German *Bildung* tradition, described by Herdt (2019) as rooted in Enlightenment ideals of freedom and liberation; that sees education as the forming of humanity. A person of *Bildung* can make sound judgment, demonstrate civil courage, and take responsibility for her own life rather than leaving this in the hands of paternalistic authorities. The National Norwegian Curriculum emphasizes the schools' *double mandate*: *Bildung* and students' learning outcomes.

As the mandate was anchored in Norwegian politicians' concern about the persistent "effect" of socio-economic background (SES) on students' learning, the group used the Coleman report (Coleman et al., 1966) as starting point. The mandate expected the group to describe activities in schools and municipalities that enhance students' learning and advice on how schools with low contribution to students' learning can improve their work. Surprisingly, the school contribution indicators were not sufficiently robust to select schools. They were therefore identified by students' grades, national test results, indicators, and other available measures such as student and parent satisfaction and level of bullying in school. Eight schools were selected and paired in four categories, based on their results the last 5 years:

Two schools with substantial contribution to students' learning the last 5 years
 Two schools that contributed to students' learning the last 5 years, with many students from homes with low SES
 Two schools with progression to students' learning the last 5 years
 Two schools with low contribution to students' learning the last 5 years

Fall 2020 and spring 2021, during the Covid 19 pandemic, members from the expert group visited eight primary schools all over Norway, from the north to the south, east, and west, and interviewed local administrators, school leaders, teachers, and students. They all got the same questions. Members of the expert group facilitated the dialogs, and notes were taken both by individual group members and the groups' secretariat. The day after each school visit was used to compare notes and write up a document for each school.

The expert group's main finding was that the schools' contribution to students' learning is primarily a question of leadership and organization. Schools with a substantial contribution to students' learning (the last 5 years) worked *with* complexity instead of reducing it. They were extremely loyal to the schools' double mandate, showed a deep respect for the school's knowledge and practiced a high degree of student participation. These schools regularly evaluated and developed practice as one intertwined process. Leaders assumed that the knowledge needed to develop quality was in the school, and teachers were empowered to lead the professional learning communities, in close collaboration with the school leadership. Information flow was systemic. In one school, an

assistant school leader surveilled all available data and met regularly with the teachers to discuss pedagogical and organizational implications of the data. Joint decisions on which actions to take, were swiftly implemented, and students kept in the loop by their teachers.

In these schools, students, teachers, and leaders worked inquiry based. They continuously developed knowledge about practice and used it to improve practice. Leaders and teachers took for granted that students need a safe learning environment if they are to concentrate on learning. They therefore started with establishing a social foundation for teaching and learning. Everyone was perceived as a resource—for each other and to the school. Each student had a learning partner. If a student struggled with a task, he or she would ask the learning partner. If the partner could not answer, they both approached the teacher to get the help they needed. This is but one example on how the organizational structures aimed at minimizing vulnerability. In these schools, the focus was on learning for all, and no one was exposed or left alone. There was no “my” students and “your” students. All students were the responsibility of all employed, from janitor to principal, librarians, cleaning staff and teachers.

The two schools with many students from homes with low SES focused on information flow and communication between all actors involved. The complexity in these schools was overwhelming as many students needed individual attention and follow-up. As students’ needs varied from day to day, the work was continuously organized and reorganized. The two schools had established specialized teams “around” students *and* teachers, to care for students’ need and to allow the teachers time to teach. Specialized teams assisted students with additional training in reading, writing and math. Other teams were responsible for the learning environment or special education. Ideally, in Norway, all students are taught in class and therefore special educators come to the classroom instead of taking students out.

District level representatives and the school leaders at the six schools in the three profiles where schools contributed to students’ learning, collaborated regularly, in a climate of trust and shared responsibility. At the two schools with low contribution to students’ learning the last 5 years, dialogs between district level representatives and principals were more formal, focused on intentions and what ought to be done. Rhetoric of district level representatives indicated hopes that better alignment between district level documents and the schools’ documents could improve teachers’ practice and students’ learning outcomes.

The expert group concluded that local administrators must strengthen their educational knowledge, improve their understanding of schools, and use this to support struggling schools.

The quest for simplicity—human and social capital

While educational institutions are overwhelmingly complex, too often, policies aim at reducing the complexity by treating education as a series of simple cause-effect-activities. Policies often assume that schools need advice on what works, summed up by Berends (2004) as: the school effects research (1960–1970) effective schools research (1970–1980), and standards-based reform (1980–1990). Underlying such policy-initiatives, argued Ingersoll and Collins (2017), is a “teacher-deficit” assumption where the primary source of low-quality teaching in schools lies in various deficits in teachers, their ability, commitment, or effort. The best way to fix schools, is therefore to fix these deficits through increased external regulations, incentives, and sanctions.

This reasoning was central to neoliberal policies that emerged in the US under the presidency of Ronald Reagan and in the UK under Prime Minister Margaret Thatcher. Etzioni (1988) described the neoliberal paradigm as utilitarian, rationalist, and individualist. It assumes that individuals seek to maximize their utility, rationally choosing the best means to serve their goals. He suggested an alternative paradigm of *codetermination*, encompassing factors that form society and personality as well as rational decision-making. As individuals act within social contexts, social groups are not aggregates of individuals but have structures of their own, Etzioni argued. To explain neoliberal policy problems, Michael Walzer (2007) used the term *neoliberal fallacy*. While liberalism had contributed to allowing people to live within autonomous institutions with institutional integrity, neoliberal policies separated people: “people are not liberated when separated from their fellow beings. Lines encircle churches, schools, markets and families—not you and me” (Walzer, 2007, p. 63). According to Dewey (1915, 40), the primary waste in education is that of human life—all waste is due to isolation, which is a result of how schools are organized.

The OECD understands human capital as individual knowledge, skills, competencies, and attributes embodied in individuals that facilitate the creation of personal, social, and economic well-being (Keeley, 2007). They see human capital as critical to sustainability, productivity, and the current and future economic health of a country. Having linked human capital and productivity, Fraumeni (2012) concluded that while labor productivity can be measured, human capital productivity is neither easily nor reliably measured.

Coleman (1988) argued that while sociologists assume that actors act in social contexts of norms and rules that affects their actions, economists assume that people act freely and independently, have individual goals, and seek to maximize their return on investment. He therefore suggested to see social and human capital as entangled. Individual and goal rational action can be studied in a social context which makes it possible to explain the development of social structures that schools must establish and maintain. Coleman (1988) distinguished between three forms of social capital, (1) obligations and expectations, (2) information channels and (3) social norms. Having analyzed large datasets on school dropout, he concluded that the lack of social capital explains dropout. Social capital is embodied in the relations among people, a resource that showed evidence of considerable value in reducing the probability of dropping out of high school.

Powers et al. (2016) argued that policymakers had neglected the Coleman report’s fundamental insights about the relationship between poverty and student achievement by emphasizing measurement; human capital theory; and the use of the White, middle-

class male student as the implicit norm against which all other students were assessed and ranked. They labeled the systematic denial of class-based inequalities in shaping individuals' and groups' life chances *willful ignorance*.

Researchers have studied social capital as a resource. For instance, Spengler et al. (2015) assumed that social-emotional skills students learn in school, might help them succeed later in life. They conducted a longitudinal study spanning 40 years (1968 and 2008) to examine how student characteristics and behaviors in late childhood (age 12) predicted career success in adulthood. Having controlled for home background and intelligence, they found that student personality-related characteristics and behaviors played significant roles in important life outcomes, over and above socioeconomic factors and cognitive abilities. Being successful, they argued, is more than "just" having good cognitive resources and coming from a socially advantaged family. Therefore, educational systems should pay more attention to social-emotional skills. Merritt and Buboltz (2015) found that socioeconomic status (SES) was significantly related to students' self-efficacy, and parental influence was a significant predictor of academic self-efficacy.

If we perceive education as preparation for life instead of comparing schools to firms, education should compensate for what groups of students do *not* get from their family or neighborhood, instead of stressing their cognitive skills to increase the schools' test results.

What can be expected of teachers, given their working conditions?

Ingersoll and Collins (2017) argued that the teacher accountability perspective involves a flawed problem diagnosis. The simple logic is that when teachers are not held accountable, they simply do what they want behind closed classroom doors, with low-quality performance as predictable result. This logic, however, overlooks that how schools are managed and organized, influence teacher quality. Poorly run schools can make otherwise excellent teachers not so excellent.

Based on data from PISA (2012), Ingersoll and Collins (2017) showed that while the US education system is relatively decentralized, schools are not. While 26% of principals reported to have high power over curriculum in the 1990s, the number was 49% in 2012. Teachers are delegated much responsibility but have little real power. For most decisions, they lie at the bottom of the hierarchy. Principals decide teachers' course assignments, courses, and grade levels they teach. Teachers can be assigned to teach subjects which do not match their fields of preparation, they have no control over decisions regarding school curricula, limited control over the school's discipline policy, little say over what kind of student ability grouping the school has and which students are placed into which tracks. Teachers have little influence over the assignment of students to their courses and rarely the power to remove disruptive students from their classrooms. Ingersoll & Collins refer to Lortie (1975), who described the relationship between teacher and student as one of "dual captivity"; teachers are public servants who cannot choose to *not* serve their clients. Their clients are recipients of a public service and cannot choose to *not* be served. This stands in sharp contrast to traditional professions, such as lawyers, psychologists, or physicians, who often have a substantial degree of choice over whom they serve and may have the option to not work with certain clients.

While teachers are delegated the responsibility for implementing decisions within their classrooms, they are not typically part of the school management, nor are they the workers, but responsible for the workers— their students.

To sum up, the article started with politicians' concern about the strong relationship between students' family background and their learning outcomes. Since the Coleman-report, questions of the school's contribution to students' learning have been high on the agenda. Accountability reforms, based on a deficit perspective, aimed at pressuring teachers and schools to better achievement. Economists launched the statistical measure *education production function*, that supposedly, would show policymakers where their investments would yield the largest profits. Teacher quality was equated with largest possible increase in students' results. The understanding of schools as factories of human capital, i.e., that the main product produced by the school is individual cognitive skills that can be tested, measured, and compared at an aggregate level to indicate nations' relative competitive advantage, has narrowed the school's broader mission and frustrated teachers. Few have interrogated into the relationship between what is expected of teachers and their room for maneuver in implementing decisions. Ingersoll and Collins (2017) showed that teachers' work is constrained by larger national, district- and school-wide decisions, over which they have little influence. Old-fashioned neoliberal ideas about learning—back to basics, rote learning, drilling, and testing of skills contributed to de-professionalizing teachers, as they were not only told *what* to teach, but also *how*.

A more intelligent accountability

If policymakers cannot trust schools, teachers, and educational researchers, then who should they trust? The chapter has shown that over the last decades, many policymakers have trusted economists as educational experts. They have invested their trust, not only in the economists' equations, but also in their understanding of education, what matters in education and what is quality teaching. The economists' black-box approach and the idea that students' learning outcomes can be compared to firms' return on investments has expected schools to invest in human capital to "produce" gains in students' learning outcome. Old-fashioned policy-ideas about how we learn has favored educational practices that John Dewey deemed outdated more than a century ago (Dewey, 1915). Cognitive psychologists (Sloman and Fernbach, 2017), argued that individual intelligence is overrated, as human intelligence resides not in individual brains but in the collective mind. What goes on between our ears ultimately depends on what goes on elsewhere, and we therefore learn best when we collaborate. Some of the best learning techniques at every level of education is to have students

learning as a team. While this is not new to educational researchers, the insight is not implemented in classrooms as widely as it could be. Education is a complex field that suffers under single-factor solutions, and policymakers should therefore consult research from various fields.

O'Neill (2018) describe trust as valuable only when directed to agents and activities that are trustworthy. Having studied the work of 43 teachers and school- and district-leaders, Butler et al. (2015) concluded that most participants experienced increased efficacy when they engaged in collaborative inquiry. They ask if approaches that assess students' performance using large-scale, standardized tests and then mandate practices and monitor improvement on that basis fail to recognize the knowledge and expertise of educators, underestimate the importance of adapting practices to meet local needs, and undermine teachers' morale and investment in innovation. Based on findings from the Norwegian expert group, it can be added that these policies also neglect the importance of knowledgeable school leaders who organize for learning.

Ingersoll and Collins (2017) found substantial variation among schools in the degree of teacher control. Schools where teachers have more control over key school-wide and classroom decisions had fewer problems with student misbehavior, more collegiality and cooperation among teachers and administrators and higher teacher retainment. They also found a strong relationship between the average amount of teacher influence over discipline decisions and the likelihood that teachers stay in, or voluntarily depart from, their schools. Too much organizational control may deny teachers the very control and flexibility they need to do the job effectively and undermine their motivation. The variable that stood out in their study was how much autonomy teachers were allowed in their own classrooms over key issues such as selecting textbooks and other instructional materials, selecting content, topics, and skills to be taught, selecting teaching techniques, evaluating, and grading students, determining the amount of homework to be assigned and disciplining students. They conclude that at the crux of successful teaching is the level of control and power teachers have on issues for which they are responsible. This aligns with findings from the Norwegian expert group (Lillejord et al., 2021). In schools with a substantial contribution to students' learning, the principals knew how to utilize and strengthen teachers' educational knowledge. Teachers were empowered to lead the schools' professional learning communities and continuously reported suggestions for improvement to the school leaders. Swift information flow, organization for the purpose of learning, the combination of professional knowledge and administrative capabilities (O'Day, 2002) enabled these schools to organize robust social learning units where learning was both individual *and* social. These school leaders adhered to the three forms of social capital identified by Coleman (1988): obligations and expectations were systemic; information channels were effectively used, and social norms were clearly communicated, adhered to and followed up. Shain (2016) argued that we cannot expect that schools alone can address the impact of wider social inequalities, certainly not without significant investment.

Conclusion

A question raised in this chapter is if accountability reforms created more problems than they solved because they ignored the complexity of schools. When Onora O'Neill (2013) called for a more intelligent accountability, she used the word *intelligible*. This chapter contributes to the debate waging over how to establish frameworks that empower all stakeholders to exercise responsibility in an intelligent accountability system. If teachers are to improve their work, they must understand what they are being held accountable for, they must have the necessary authority, be able to control their professional working conditions and have leadership support. Elliott (1994) argued that the rationale for involving teachers as researchers of their own practice is connected to an aspiration to give them control over what is to count as knowledge about practice. Instead of applying knowledge generated by outsiders, teachers should generate their own professional knowledge. Aggregated data is an abstract form of knowledge with low explanatory power as the data is not the evidence, but the evidence is *in* the data (Ströing, 2018) and needs interpretation to be relevant for teachers' context. In trusting educational systems, all stakeholders are reciprocally responsible to one another, students, and the public to fulfill their joint responsibilities for educational outcomes. The case used in this chapter revealed that schools that contribute to students' learning have leaders—at school and district level—who understand that abstract, second order data must be translated to the first-order work of teachers. As accountability reforms have lost some of their momentum, the search for new and better policies continues. This chapter has suggested a strengthening of the teaching profession and their leaders as an obvious next step. Combined with strengthening the sectors and politicians' educational knowledge (Brezinka, 2012) the education sector can be better prepared for future attempts on a hostile take-over.

References

- Adnett, N., Bougheas, S., Davies, P., 2002. Market-based reforms of public schooling: some unpleasant dynamics. *Econ. Educ. Rev.* 21 (4), 323–330.
- Aspers, P., 2009. Empirical phenomenology: a qualitative research approach (The Cologne Seminars). *Indo-Pacific J. Phenomenol.* 9 (2), 1–12.
- Ball, S.J., 1990. *Politics and Policy Making in Education: Explorations in Policy Sociology*. Routledge, London.
- Ball, S.J., 2010. New voices, new knowledges and the new politics of education research: the gathering of a perfect storm? *Eur. Educ. Res. J.* 9 (2), 124–137.
- Barr, R., Dreeben, R., 1977. Instruction in classrooms. *Rev. Res. Educ.* 5, 89–162.
- Beck, J., 2008. Governmental professionalism: re-professionalizing or de-professionalizing teachers in England? *Br. J. Educ. Stud.* 56 (2), 119–143.
- Berends, M., 2004. In the wake of A Nation at Risk: new American Schools' private sector school reform initiative. *Peabody J. Educ.* 79 (1), 130–163.
- Berliner, D.C., 2002. Educational research: the hardest science of all. *Educ. Res.* 31 (8), 18–20.

- Berliner, D.C., 2020. The implications of understanding that PISA is simply another standardized achievement test. In: *Handbook of Education Policy Studies*. Springer, Singapore, pp. 239–258.
- Biesta, G.J., 2007. Why “what works” won’t work: evidence-based practice and the democratic deficit in educational research. *Educ. Theor.* 57 (1), 1–22.
- Biesta, G.J., 2010. Why “what works” still won’t work: from evidence-based education to value-based education. *Stud. Philos. Educ.* 29 (5), 491–503.
- Brezinka, W., 2012. *Philosophy of Educational Knowledge: An Introduction to the Foundations of Science of Education, Philosophy of Education and Practical Pedagogics*, vol. 3. Springer Science & Business Media.
- Butler, D.L., Schnellert, L., MacNeil, K., 2015. Collaborative inquiry and distributed agency in educational change: a case study of a multi-level community of inquiry. *J. Educ. Change* 16 (1), 1–26.
- Cizek, G.J., 2001. More unintended consequences of high-stakes testing. *Educ. Meas.* 20 (4), 19–27.
- Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F.D., York, R.L., 1966. *Equality of Educational Opportunity*. U.S. Government Printing Office, Washington, DC.
- Coleman, J.S., 1988. Social capital in the creation of human capital. *Am. J. Sociol.* 94, S95–S120.
- Cowie, M., Crawford, M., 2007. Principal preparation—still an act of faith? *Sch. Leader. Manag.* 27 (2), 129–146.
- Dalton, J.L., 1972. Accountability in Education! Why? ERIC—ED075481—Accountability in Education! Why? 1972.
- Darling-Hammond, L., 2004. Standards, accountability, and school reform. *Teach. Coll. Rec.* 106 (6), 1047–1085.
- Dewey, J., 1915. *The School and Society & The Child and the Curriculum*. Dover Publications Inc., Mineola, NY.
- Downey, D.B., Condon, D.J., 2016. Fifty years since the Coleman Report: rethinking the relationship between schools and inequality. *Sociol. Educ.* 89 (3), 207–220.
- Elliott, J., 1994. Research on teachers’ knowledge and action research. *Educ. Action Res.* 2 (1), 133–137.
- Etzioni, A., 1988. The moral dimension. In: *Toward a New Economics*. The Free Press, New York.
- Exley, S., Ball, S.J., 2011. Something old, something new... understanding conservative education policy. In: Bochel, H. (Ed.), *The Conservative Party and Social Policy*. Policy Press, Bristol.
- Fraumeni, B.M., 2012. Human capital productivity: a new concept for productivity analysis. *Int. Prod. Mon.* (24), 20.
- Grek, S., Ydesen, C., 2021. Where science met policy: governing by indicators and the OECD’s INES program. *Glob. Soc. Educ.* 19 (2), 122–137.
- Hanushek, E., Woessmann, L., 2015. *The Knowledge Capital of Nations: Education and the Economics of Growth*. MIT Press, Cambridge, MA.
- Hanushek, E., 1971. Teacher characteristics and gains in student achievement: estimation using micro data. *Am. Econ. Rev.* 61 (2), 280–288.
- Hanushek, E., 1979. Conceptual and empirical issues in the estimation of educational production functions. *J. Hum. Resour.* 14 (3), 351–388.
- Hanushek, E., 2002. Teacher quality. In: Izumi, L., Evers, W. (Eds.), *Teacher Quality*. Hoover Institution Teacher Quality I Eric A. Hanushek (stanford.edu), Palo Alto, CA, pp. 1–12.
- Herdt, J.A., 2019. *Forming Humanity: Redeeming the German Bildung Tradition*. University of Chicago Press, Chicago.
- Hess, F., 2001. *Tear Down This Wall: The Case for a Radical Overhaul of Teacher Certification*. Progressive Policy Institute, Washington, DC from: <https://fordhaminstitute.org/national/commentary/tear-down-wall-case-radical-overhaul-teacher-certification>. (Accessed 31 October 2019).
- Hill, H.C., 2017. The Coleman Report, 50 years on: what do we know about the role of schools in academic inequality? *Ann. Am. Acad. Polit. Soc. Sci.* 674 (1), 9–26.
- Ingersoll, R.M., Collins, G.J., 2017. Accountability and control in American schools. *J. Curric. Stud.* 49 (1), 75–95.
- Kantor, H., Lowe, R., 2017. Introduction: what difference did the Coleman Report make? *Hist. Educ. Q.* 57 (4), 570–578.
- Keeley, B., 2007. *Human Capital. How What You Know Shapes Your Life*. OECD Insights. OECD, Paris.
- Kister, J., 1993. Two Perspectives: Perkins II Accountability. OBE/Education Reform Movement. Paper Presented at the Annual Conference of the American Vocational Association, Nashville, December 3–7, 1993.
- Lareau, A., 2009. Narrow questions, narrow answers: the limited value of randomized controlled trials for education research. In: *Education Research on Trial*. Routledge, pp. 155–172.
- Lillejord, S., Bolstad, A.K., Fjeld, S.-E., Lund, T., Myhr, L.A., Ohm, H., 2020. Kunnskapsgrunnlag (Knowledge Base) Kunnskapsgrunnlag—Ekspertgruppe for skolebidrag (regjeringen.no).
- Lillejord, S., Bolstad, A.K., Fjeld, S.-E., Lund, T., Myhr, L.A., Ohm, H., 2021. En skole for vår tid. Sluttrapport fra ekspertgruppe for skolebidrag [A School for our Time. Final report from Expertgroup for School Contribution] (regjeringen.no) Sluttrapport – Ekspertgruppe for skolebidrag (regjeringen.no).
- Lillejord, S., 2020. From “unintelligent” to intelligent accountability. *J. Educ. Change* 21 (1), 1–18.
- Lortie, D., 1975. *School Teacher*. University of Chicago Press, Chicago, IL.
- Merritt, D.L., Buboltz, W., 2015. Academic success in college: socioeconomic status and parental influence as predictors of outcome. *Open J. Soc. Sci.* 3 (05), 127.
- O’Day, J., 2002. Complexity, accountability, and school improvement. *Harv. Educ. Rev.* 72 (3), 293–329.
- O’Neill, O., 2013. Intelligent accountability in education. *Oxf. Rev. Educ.* 39 (1), 4–16.
- O’Neill, O., 2018. Linking trust to trustworthiness. *Int. J. Phil. Stud.* 26 (2), 293–300. <https://doi.org/10.1080/09672559.2018.145463>.
- Phelps, R.P. (Ed.), 2005. *Defending Standardized Testing*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.
- OECD, 2013. *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices*, vol. IV. PISA, OECD Publishing. <https://doi.org/10.1787/9789264201156-en>.
- Powers, J.M., Fischman, G.E., Berliner, D.C., 2016. Making the visible invisible: willful ignorance of poverty and social inequalities in the research-policy nexus. *Rev. Res. Educ.* 40 (1), 744–776.
- Schütz, A., 1962. *Collected Papers I: The Problem of Social Reality*. Springer, Dorderech.
- Shain, F., 2016. Succeeding against the odds: can schools “compensate for society”? *Educ.* 3-13 44 (1), 8–18.
- Slavin, R.E., 2002. Evidence-based education policies: transforming educational practice and research. *Educ. Res.* 31 (7), 15–21.
- Slooman, S., Fernbach, P., 2017. *The Knowledge Illusion. Why We Never Think Alone*. Macmillan, UK.
- Sørensen, A.B., Morgan, S.L., 2000. School effects. In: *Handbook of the Sociology of Education*. Springer, Boston, MA, pp. 137–160.
- Spengler, M., Brunner, M., Damian, R.I., Lüdtke, O., Martin, R., Roberts, B.W., 2015. Student characteristics and behaviors at age 12 predict occupational success 40 years later over and above childhood IQ and parental socioeconomic status. *Dev. Psychol.* 51 (9), 1329.
- Stehr, N., 1994. *Knowledge Societies*. Sage publications, London.
- Stehr, N., 2003. Modern societies as knowledge societies. In: Ritzer, G., Smart, B. (Eds.), *Handbook of Social Theory*. Sage Publications, London, pp. 494–508.
- Strömg, P., 2018. Data, evidence, and explanatory power. *Philos. Sci.* 85 (3), 422–441.
- Tröhler, D., 2017. The paradox of being a teacher: institutionalized relevance and organized mistrust. In: *Schools in Transition*. Brill Sense, pp. 79–95.
- Vandenbergh, V., 1999. Economics of education. The need to go beyond human capital theory and production-function analysis. *Educ. Stud.* 25 (2), 129–143.
- Walzer, M., 2007. In: Miller, D. (Ed.), *Thinking Politically: Essays in Political Theory*. Yale University Press, New Haven, CT.
- Wang, L., Beckett, G.H., Brown, L., 2006. Controversies of standardized assessment in school accountability reform: a critical synthesis of multidisciplinary research evidence. *Appl. Meas. Educ.* 19 (4), 305–328.
- Ydesen, C., Bomholt, A., 2020. Accountability implications of the OECD’s economic approach to education: a historical case analysis. *J. Educ. Change* 21 (1), 37–57.

Educational accountability: the use and abuse of assessment data

Christopher M. Saldaña^a and Lorrie A. Shepard^b, ^a School of Education, University of Wisconsin, Madison, WI, United States; and ^b School of Education, University of Colorado, Boulder, CO, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	164
The use and misuse of assessment in the United States	164
Assessment in the 1840s–1930s—program evaluation and ability tracking	164
Assessment in the 1960s–1970s—the passage of ESEA and the emergence of NAEP	165
The passage of ESEA and requirement of assessment in title 1 schools	165
The creation of NAEP	165
Assessment in the 1980s–2015s—standards-based reform (SBR), incentive theory, & NCLB	165
A shift to “excellence,” SBR, and incentive theory	165
The passage of No Child Left Behind (NCLB)	166
The use of assessment across and within international contexts	166
The use of assessment across international contexts	166
Assessment within international contexts	167
Chile	167
Norway	168
Singapore	168
Conclusion	169
References	169

Glossary¹

Assessment The process of collecting data to document or measure student knowledge, skills, or dispositions. Testing is the most common formal means of assessment

Educational Accountability The concept of holding schools, administrators, teachers, and/or students responsible for students’ academic performance. Accountability policies vary across time and across jurisdictions in the types of school and student data that are used and whether rewards or sanctions are attached to performance. Most forms of educational accountability include student standardized achievement test scores as a key element and report the information publicly

High-Stakes Tests Assessments that carry serious consequences for students or for educators. Their outcomes determine such important things as promotion to the next grade or graduation or may be used to administer rewards or sanctions to schools or teachers. Even when there are no official rewards or punishments attached, reporting results publicly can have high stakes for teachers

Large-Scale Assessment Standardized tests and other forms of assessment designed to be administered to large numbers of students and to provide information on a standardized scale

Matrix Sampling An assessment administration method by which a large set of test items is organized into a number of relatively short test forms, each of which is randomly assigned to a subsample of test-takers, thereby avoiding the need to administer all items to all examinees

Program Evaluation Involves the systematic collection and analysis of data to determine the effectiveness of educational projects, programs, or policies. The use of assessment data to evaluate programs is distinguished from the use of assessment data to make instructional decisions about individual students. Program level analysis may entail disaggregating total school performance to check on program effectiveness for important subgroups or examining curricular strengths and weaknesses, as when students excel in basic skills but perform relatively less well on tasks requiring critical thinking

Standardized Test Tests that are administered and scored under uniform conditions for all students. Standardization ensures comparability of scores across individuals

Standards-Based Reform An approach to education reform that sets standards of performance in academic subject areas as a means to improve the substance of school curricula and increase the motivation and effort of students, teachers, and school systems and thereby improve student achievement. The reform assumes high standards for all students and has two

¹Definitions for terms associated with Standards Based Reform are adapted from McDonnell, L. M., McLaughlin, M. J., and Morison, P. (eds.). (1997). *Educating one and all: students with disabilities and standards-based reform*. Washington, D. C.: National Academy Press. And from McLaughlin, M. W., and Shepard, L. A. (1995). *Improving education through standards-based reform*. Stanford, CA: The National Academy of Education.

components challenging standards that set out what students should know and be able to do, and an accompanying agenda for educational equity

Introduction

When used well, “large-scale standardized assessments” provide educational stakeholders with valuable information about schooling while having a limited impact on the everyday operation of schools. Policymakers, for example, can use large-scale assessment to track trends in achievement and identify effective educational programs. When used in this way, assessment can be a tool for improving equity within and across schools. When misused or used poorly—for example, as a tool for punishing under-resourced schools—assessments can harm students by misdirecting the focus of school administrators, distorting how practitioners teach, and narrowing the purpose of education to performance on standardized tests. When used in this way, large-scale assessment can widen “opportunity gaps” by encouraging repetitive drill and practice for low-performing students and further disadvantaging schools serving students with the greatest needs. This entry examines how large-scale assessment has been used in educational policy over time, focusing on both the uses and misuses of assessment data.

The first section summarizes the history of assessment and accountability in the United States. The earliest uses of achievement and IQ tests in the mid-19th and early 20th centuries were for local program evaluation and to track students by ability. The role of achievement testing then shifted in the 1960s and 1970s to the relatively low-key beginnings of accountability, with Title 1 of the Elementary and Secondary Education Act (ESEA), and informational data collection for policymakers with the creation of the National Assessment of Educational Progress (NAEP) in 1969. The role of assessment in the United States evolved further during the 1980s because of complaints of educational stagnation and fear of “falling behind” economic competitors. During this time, assessment changed from a tool for program evaluation and public information to a lever for applying pressure to “failing” schools and school districts. Within this latter period, the role assessment played in policy is described, with particular attention to the “theories of action” informing its use, and the effect of this approach on the delivery of instruction and student academic outcomes on standardized tests.

The second section describes the use of assessment and accountability across and within international contexts. The use of large-scale international assessments such as The Program for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS) is presented first, considering the motivation for administering these assessments and the potential pitfalls when interpreting results. Then, as examples of large-scale national educational assessment systems across three continents, brief descriptions are given of the assessment systems and data use in Chile, Norway, and Singapore.

The use and misuse of assessment in the United States

Although standardized measures of academic achievement are used throughout the world, a large portion of the research on the use (and misuse) of large-scale assessment for accountability purposes has focused on the United States. This section examines the role of assessment in US K-12 educational policy as it shifted from 1840 to the present.

Assessment in the 1840s–1930s—program evaluation and ability tracking

The use of achievement tests in the United States began in the mid-19th century. Horace Mann, considered one of the founders of public education and the common school movement, sought to use evidence from student examinations to evaluate school programs and identify effective practices. Known for his vision of a universal common school, Mann saw value in creating common schools that could be scaled across contexts. Assessment offered Mann a tool for realizing that vision. To this end, Mann encouraged schools to adopt writing exams that could be used to inform state policymakers and inform student placement. Beginning in 1908, roughly 60 years after Mann’s initial push for using assessments in public schools, Edward Thorndike and other early experts in educational measurement began to design hundreds of achievement tests that could be used for educational program evaluation (Shepard, 2008).

In the same year, Henry Goddard published his version of the “The Binet and Simon Tests of Intellectual Capacity.” By 1913, the IQ test, as it would come to be known, was in use in public schools. At the time, the common belief about the intelligence quotient was that it could accurately measure a student’s inherited intellectual ability and capacity to learn. The promises of scientific measurement seemingly arrived just in time to address problems arising from a rapid expansion in both total enrollment in public schools and the diversity of students enrolling. These enrollment and diversity pressures, resulting from greater accessibility to high schools and an influx of immigrants from Europe in the years between 1880 and 1920, posed a challenge for policymakers—how to expand the common school while maintaining the vision outlined by Horace Mann for the “one best system.”

Policymakers would eventually adopt a system known as a social efficiency approach – a public education system where schools offer differentiated instruction (i.e., tracks) within schools ostensibly to account for variation in student capabilities (Tyack, 1974).

While there were some critics of IQ testing from the beginning, full-throated objections and scientific reexamination did not occur until the civil rights movement of the 1960s and 70s. Belatedly, it was recognized, at least among researchers, that assessment results are an indication of the learning opportunities an individual has received to that point, not a measure of an individual's innate ability (National Research Council, 1982). Thus, a system relying on assessment results for student placement cemented within schools the economic inequality and cultural, racial, and language discrimination that existed outside of schools.

This entry focuses primarily on the use of large-scale assessment data for policy and program evaluation purposes. Nonetheless, the legacy of IQ testing for individual student placement is worth noting because the implication that some groups of students are less capable than others continues to influence public discourse about achievement test results and the actions of some educators. For example, a recent consensus report on learning research from the National Academies of Sciences, Engineering, and Medicine (2018) found it necessary still to amass research evidence arguing against a “deficit” model of cultural difference.

Assessment in the 1960s–1970s—the passage of ESEA and the emergence of NAEP

The passage of ESEA and requirement of assessment in title 1 schools

The passage of the Elementary and Secondary Education (ESEA) of 1965 shifted the landscape of testing policy in the United States. Title 1 of ESEA provided federal dollars to schools serving students in poverty. In return, policymakers mandated that schools that accepted the federal monies provide evidence that the money was being used effectively to help economically disadvantaged students. The mandate marked a shift in the management approach of policymakers and the role of assessment within that approach (Resnick, 1980). The assessments tied to the first ESEA remained low stakes, however. Moreover, the desire to prevent the encroachment of the federal government into local practices meant that assessment remained “chaotically diverse,” leaving little policymakers could use the assessments for beyond local evaluation (Cronbach et al., 1980, p. 33).

The creation of NAEP

Because the various standardized achievement tests used for Title I evaluation offered little information that could be used to report on the nation as a whole or to monitor progress over time, policymakers sought a very different kind of achievement test (Lehmann, 2004). In 1969, efforts of Ralph Tyler and others culminated in the creation of the National Assessment of Educational Progress as a data collection tool analogous to the US census. NAEP would provide information to policymakers on the academic standing of students, nationally and by region, in different subject areas at key age levels. The program was not designed to apply pressure—specifically, there were not to be rewards and punishments attached to performance on the test.

Further, NAEP was designed to be minimally intrusive in schools. The assessment was administered using matrix sampling, a technique designed so that multiple students take different parts of the test. Sampling of students avoids the problem of requiring every student in a state, district, or school to take the total test, thus, lowering the cost of testing. Sampling of test content across sampled students also allows much greater coverage of content domains without undue time demands on participating students, thus ensuring a much higher quality test.

Over time, policymakers have sought to shift the relevance and usefulness of NAEP. For example, policymakers sought changes to NAEP that made it possible to compare results across states. At the same time, the National Assessment Governing Board (NAGB), instituted by Congress in 1988 to oversee NAEP, decided to change the format for reporting NAEP results to make them more digestible for public consumption. This included setting subjectively determined “achievement levels” indicating basic, proficient, and advanced performance. While varying across grades and subject areas, the proficiency threshold for “what students should know” was generally set quite high, at roughly the 75th percentile, which led to more worrisome results than when district administered tests norm-referenced tests had claimed the majority of students to be above average. These changes to NAEP marked a shift that coincided with increased political interest in the performance of schools and a desire to focus public attention on inadequacies in the quality of American schools. While the move did not involve direct accountability sanctions, it again signaled a shift in general beliefs about how assessment could and should be used in educational policy.

Assessment in the 1980s–2015s—standards-based reform (SBR), incentive theory, & NCLB

A shift to “excellence,” SBR, and incentive theory

Early in the 1980s, politicians called for improvements to the educational system, fearful that an apparent test score decline (on the SAT, not NAEP) threatened to diminish the US's economic and political hegemony in the world. This sentiment—which was known as the “excellence movement”—is perhaps best captured by the publication of the 1983 U.S. Department of Education's report entitled *A Nation at Risk*. The report, commissioned by the Reagan administration, called for a reinvestment in public education and a greater role for assessment and accountability in K-12 public schools. Unfortunately, the movement focused primarily on tests of basic skills, resulting in a watered-down curriculum (Resnick and Resnick, 1992).

It was not until the early 1990s that policymakers sought new ideas about what large-scale assessments should measure and how they should be used as part of accountability systems. Educational researchers and policymakers eventually adopted the theory of action known as systemic reform or standards-based reform (SBR) (Smith and O'Day, 1991; Resnick and Resnick, 1992). SBR holds that clear and visible standards—i.e., those identifying what students should know and be able to do—when accompanied with high-quality teacher training, instructional materials, and improved assessment practices, could enable students to reach higher levels of performance by prompting schools to organize themselves around best practices and high expectations. As part of SBR,

assessments would be used to track student performance toward those high expectations. The theory stopped short of mandating a national curriculum or attaching sanctions when schools failed to meet expectations.

Simultaneously, however, a different group of policy leaders and researchers, working more from the perspective of economic theory than learning research, were advocating for the use of incentives to motivate educational change in US K-12 public schools (Elmore and Rothman, 1999). When applied to education, incentive theory holds that policymakers should tie school performance on academic standards to rewards and punishments. The threat of sanctions, they argued, would motivate schools to identify best practices that more effectively served students. The incentives approach was in contrast to the capacity building approach encouraged by advocates of standards-based reform. According to the incentives view, capacity building—e.g., teacher training, improved curricula, and better assessments—would be accomplished indirectly by encouraging schools to pursue those resources through their own initiative. Incentive theory, still named SBR, gained national attention because of the perceived success of such reforms in places like Texas—otherwise known as the “Texas Miracle.” Years later, studies comparing Texas test scores on state standardized tests to Texas scores on NAEP would conclude that much of the “success” was the result of test-score inflation, not real academic progress (Klein et al., 2000).

The passage of No Child Left Behind (NCLB)

No Child Left Behind, the 2002 reauthorization of ESEA, intensified the strictures of incentive theory in federal law. NCLB continued the idea of assessments aligned to challenging content and performance standards, but instead of requiring tests once at each level of school (elementary, middle, and high school), reading and math tests were mandated every year from grades 3–8. In addition, state policymakers were required to establish targets for adequate yearly progress and to report disaggregated assessment data by race and socioeconomic status. NCLB, thereby, provided state officials with a tool to punish “underperforming” schools.

Specifically, the use of assessment to hold schools accountable was intended to impact multiple parts of the K-12 educational system in the United States—i.e., it would improve system alignment, increase test performance, close achievement gaps, enable school turnaround for the worst performing schools, and improve educational outcomes for students with disabilities. While there is evidence that high-stakes accountability made a difference in some of these areas, the results were not as substantial as some would have hoped. A review of studies published by the National Academies of Sciences, Engineering, and Medicine, for example, found that the effect of incentives and test-based accountability on student achievement was small (National Research Council, 2011), about 0.08 of a standard deviation, on average, with most of the improvements being concentrated in elementary mathematics.

Despite these small positive effects, however, many, perhaps a majority of, educational researchers consider the use of assessment in high-stakes accountability policy to be an overall net failure because of the negative effects high-stakes assessment has had on teaching and learning. Prior to the implementation of NCLB, experts in the field of measurement had cautioned that a narrow use of assessment to punish schools would result in the corruption of measures and would be a distraction for policymakers, turning attention away from the work required to build local capacity for schools to successfully serve students. They were right—NCLB encouraged school administrators and teachers to focus most of their attention on standardized tests, resulting in test-score inflation and curriculum narrowing, stress and de-professionalization of teaching, and deficit views of students and dead-end student placements. A meta-synthesis of 49 qualitative studies, for example, found only rare exceptions to the familiar pattern of curricula narrowed to tested subjects, fragmentation of subject matter knowledge into test-related pieces, and increased use of teacher-centered pedagogies (Au, 2007). The fact of test-score inflation is such a well-recognized problem that all of the best studies, of NCLB and accountability effects reported above, rely on independent NAEP results rather than gains on likely-inflated state test results.

While the blunt use of large-scale assessment to sanction schools was seemingly reduced in the 2015 reauthorization of ESEA—the Every Student Succeeds Act signed by President Obama—there was no reduction in the amount of testing required. Thus, the federal approach to educational policymaking continues to rely on large-scale assessment as a tool for educational accountability.

The use of assessment across and within international contexts

This section reviews the use of large-scale standardized assessment across and within international contexts. The first part of the section focuses on the use of large-scale international assessments, considering specifically the motivation for using those assessments and the potential pitfalls in interpreting testing outcomes. The second part examines the use of large-scale assessment in educational systems in Chile, Norway, and Singapore.

The use of assessment across international contexts

Around the time assessment began to grow in the United States in response to the passage of the first ESEA, the United States, along with other high- and upper middle-income countries, began exploring the possibility of conducting large-scale standardized assessments across international contexts (Lockheed et al., 2015). In 1960, the first meeting for the International Association for the Evaluation of Educational Achievement Pilot Study (IEA) was held. The convening included representatives from Belgium, England,

Germany, France, Finland, Israel, Poland, Scotland, Sweden, Switzerland, Yugoslavia, and the United States. The conclusion from initial pilot studies was that large-scale international assessments could be conducted despite concerns about “translation and administration” across contexts. Over time, participation in international assessments grew, as did the number of international assessments in which countries could participate. Among these are the Trends in International Mathematics and Science Study (TIMSS), The Program for International Student Assessment (PISA), and the Progress in International Reading Literacy Study (PIRLS) (see [Scherer and Nilsen, 2022](#), for more discussion of international assessments).

Arising at the same time as increasing economic globalization and economic development, international assessments such as the TIMSS and PISA have served as indicators of the effectiveness of international educational systems. Much like the use of measures of economic development—such as, the growth rate of a country’s Gross Domestic Product (GDP)—international assessments offer policymakers across the globe a tool to make international comparisons and track improvements in educational systems.

However, making comparisons across international contexts can be challenging. For example, a report published by the Economic Policy Institute argues, among other things, that it can be challenging for test-makers and administrators to create truly representative national samples and select which topics will be included in tests ([Carnoy and Rothstein, 2018](#)). The authors point out that in 2009, PISA oversampled low-income US students attending schools with similarly economically disadvantaged students. They argue this artificially lowered US results. With respect to subjects tested, the authors note that on the TIMSS, the United States and Finland performed about the same on mathematics. The underlying results, however, showed that the United States performed better in Algebra but worse in number properties (e.g., fractions). Algebra is weighted less on the TIMSS, however. Caution should therefore be practiced when interpreting the results of international assessments to make international comparisons.

Assessment within international contexts

Much like the globalization of economic policies experienced throughout the 1990s and 2000s, educational “policy borrowing” has resulted in the spread of policies and practices between the United States and other regions of the world ([Lingard and Lewis, 2016](#)). Large-scale assessment is an example of a tool policymakers have adopted across international contexts. This section examines how assessment is used in Chile, Norway, and Singapore.

Chile

In the United States, high-stakes accountability policy developed over time in a confluence with school choice policy. That is, school choice advocates found assessment to be a useful tool to rank and grade schools as part of a marketized system where parents could use test data to choose schools. In Chile, high-stakes accountability policy came after the establishment of school choice. In the 1970s, Chile underwent an economic revolution during the dictatorship of Augusto Pinochet. Under Pinochet, leaders in Chile created a school voucher system to facilitate school choice ([Weinstein et al., 2016](#)).

Chile began using national large-scale assessments in 1988. The system—called the Sistema Nacional de Evaluación de Calidad de la Educación (SIMCE—The National System for Evaluating the Quality of Education)—requires that tests be administered to students in fourth, eighth, and tenth grade. Initially, the system was low-to moderate-stakes. SIMCE results were not made public until 1995, when Chile’s Ministry of Education changed its policy to create a school ranking system that would, in theory, help families make decisions about school choice. Between 2003 through 2013, Chile used the system to rate schools as “advanced,” “intermediate,” or “inadequate.” The logic behind the ratings was that “pressure” on schools would be applied by threats to student enrollment—i.e., parents could use the information offered by assessments and the power of choice offered through a school voucher to move their children to a different school.

Chile’s leaders quickly realized that its system of public education had resulted in large gaps in school performance with the gaps largely determined by socioeconomic status. These gaps were caused in part by the socioeconomic segregation that had developed as a result of parental choice ([Hsieh and Urquiola, 2006](#)). To address the problem, in 2008 Chile offered families from economically disadvantaged backgrounds a “preferential student voucher” (subvención escolar preferencial (SEP)). In Chile, all children can receive a voucher, but the SEP program offers vulnerable students a voucher worth more money. In 2015, the law reached the lowest 60% of the country’s income distribution ([Elacqua et al., 2019](#)).

Along with the new monies, however, Chile required schools that accepted the SEP voucher to accept a harsher accountability system. Schools that agreed to receive the funds had to sign a four-year education improvement plan (Plan de Mejoramiento Educativo) that included a commitment to raise the school’s SIMCE scores, innovate, reduce dropout and grade repetition, improve working conditions, and increase parental involvement. Much of the assessment of schools (70%) was determined by SIMCE scores. If the school failed, it could be labeled “in recovery,” have its budget cut, be taken over, and/or be closed.

Since the implementation of the policy, student scores on national assessments have improved, although it remains difficult to estimate how much of the increase was caused by the implementation of the SEP. Recent studies have found that the policy did have some benefit by (1) providing more funding to schools already enrolling a large number of economically disadvantaged students and (2) increasing the number of private schools participating in the program. Qualitative studies of changes in school behavior, however, have raised concerns over narrowing instruction to focus on test performance and arrangements for low-performing students to be absent for the test. Such findings have raised questions about inflation in test performance or whether additional funds tied to school sanctions determined by assessments are resulting in meaningful academic improvement for Chile’s economically disadvantaged students ([Elacqua et al., 2019](#)).

Norway

While the use of large-scale assessment in Norway has changed in recent decades with policymakers placing a greater emphasis on the collection of student academic outcome data through assessment, Norway's approach remains light relative to the use of assessment data in educational systems in other countries (Skedsmo and Mausesthaugen, 2016). This is due in part to the emphasis on trust and professionalism Norway places in its teachers (Lillejord, 2022).

Norway began using assessment as part of a "school-based evaluation" system in the 1990s. Much like the United States, however, the general public in Norway was motivated to increase the role of assessment for information about student outcomes after the country's educational system was perceived to have underperformed relative to international economic competitors. In 2001, Norway faced "PISA shock," as reported in national media, when its system ranked below average in math and science and underperformed in literacy as well. In response, the country sought to increase pressure on schools by publishing the results of assessments to the public. The country also implemented significant curricular reforms (Hopfenbeck et al., 2013; Lillejord, 2022).

In 2002 and 2004, the Government of Norway published two green papers encouraging more attention to individual learning outcomes and improved instructional practices. Those two papers provided the framework for what would become Norway's National Quality Assessment System (NQAS) and national curriculum (the Knowledge Promotion Reform), which emphasizes basic skills in reading, writing, oral communication, digital literacy and numeracy. Underlying the implementation of NQAS were calls for more data and information about school and individual student performance.

NQAS uses a combination of outcomes on national tests, international assessments (i.e., PISA, TIMSS, and PIRLS), the "school-leaving" examination (exit exam taken at the end of secondary school, the "Craft Certificate" (the certificate for qualifying for vocational education), and individual student grades. In Norway, municipalities use performance on NQAS to publish school results and rank schools. Not at all municipalities, however, publish school rankings. Moreover, Norway does not attach any government-regulated sanctions to school performance on NQAS. Again, this is due in part to the trust placed in school leaders and teachers as capable professionals pursuing the best interest of students (Lillejord, 2022).

Internationally, Norway regularly places above the average in student performance on assessments across grade levels and subject areas. The country is considered to have one of the more efficient and effective educational systems in the world. Despite this performance when viewed externally, educational stakeholders within Norway continue to raise concerns about how the use of national assessments to publicly pressure schools might distort behavior and motivation in classrooms.

Singapore

The Singaporean system of education makes use of assessment in similar ways to each of the countries reviewed above. As in other countries, the shift over time in educational policy and in the use of assessment has been spurred by the perception of educational failure and a greater emphasis on connecting educational outcomes to economic outcomes. In the 1980s, for example, Singapore sought to incorporate into its educational system the practices of successful businesses and non-profit organizations. The shift to greater educational accountability was motivated in part by the 1985 global economic crisis. This movement was further spurred by the economic collapse in Singapore in 1995; and recent changes in school design and goals have reflected shifts in the skills Singapore believes students will need to remain competitive on the global stage. Currently, the state uses assessment as part of a three-pronged system of accountability, which includes Singapore's school ranking system, the School Excellence Model (SEM), and the School Awards system (Tee Ng, 2010).

Singapore introduced its school ranking system in 1992. The logic behind the system was that competition would improve school performance. This presentation of school performance was not tied to sanctions or rewards like those used in the United States. Instead, Singapore's ranking system ranks schools based on three criteria: student performance on national exams, the "value-added-ness" of a school (i.e., the gain in test scores that can be attributed to a school), student performance on the National Physical Fitness Test, and the percentage of students in a school who are overweight. While school rankings were initially published to the public, Singapore eventually discontinued the practice to avoid an overemphasis on academic results (Loh and Hu, 2019). The state now provides school principals with a list of comparable schools, where comparability is determined by school-level student characteristics, such as student socioeconomic status, student ethnicity, and parental educational attainment.

The SEM was introduced in 2000. It requires that schools conduct a self-evaluation once a year, with a state evaluation occurring every five years. The SEM model rests on the assumption that schools are operated by "enablers" who produce "results." In the SEM model, "enablers" are judged on leadership, strategic planning, staff management, resources, and student-focused processes. Alternatively, "results" are judged by measuring administrative and operational outcomes, staff outcomes, partnership and society outcomes, and key performance outcomes, where "key performance" accounts for 30% of a school's total score and includes data from a school's performance on national assessments and learning outcomes.

The School Awards system, or the Masterplan of Awards (MoA), is part of Singapore's SEM system. It recognizes schools for their performance in several areas. For example, achievement awards are given to schools for the year's achievement; development awards are awarded for achievements in character development and national education; best practice awards are given to schools based on "enabler" outcome scores; sustained achievement Awards are given to schools for sustained performance in "results" outcomes. Singapore also gives out outstanding development awards for outstanding and sustained achievement in national education and character education and school distinction awards and school excellence awards that recognize exemplary school processes and practices and overall performance, respectively. Singapore reports the number and type of awards won by each school. It does not, however, report "failing" schools—or schools that perform poorly on the SEM measured outcomes. As an example of the criteria considered by the Singaporean Ministry of Education in determining an award, consider the characteristics used to describe three schools that

received “school excellence awards” and “best practice awards.” The schools St. Hilda’s Primary, Chung Cheng School (Main), and Dunman Secondary received these awards because they prepared students holistically—i.e., emotionally, physically, and cognitively, emphasized leadership in their curriculum, and tied learning to constant development aimed at fostering a global outlook and 21st century skills (Ministry of Education – Singapore, 2010).

Unlike the United States, Singapore does not attach sanctions to poor school performance. The system is designed to apply pressure through public scrutiny and centralized forms of government control. That is, the Singaporean system of education does not rely on local control as an aspect of governance. As a city and small country, there are not intermediate district or state levels of governance. All schools in the country are accountable to the Ministry of Education. Most schools in Singapore are public and most students attend public schools. Like Norway, Singapore also holds much greater control over the curriculum used in schools, giving it greater control over the alignment of curriculum, assessment, and expectations.

Singapore continues to be one of the top performers on international assessments. It regularly scores in the top five countries on PISA and TIMSS in various subject areas and across all grade levels. Singapore serves just over half a million students. For comparison, the United States serves close to 50 million students.

Conclusion

When used well, assessment offers valuable information that can support student learning. For example, state and national-level standardized assessments can offer information on the performance of individual schools and school districts and the effectiveness of educational programs and resources. However, when misused, assessment can diminish the educational opportunity of students by narrowing the curriculum and distorting the practices of teachers.

In the United States, large-scale assessments have shifted from tools for program evaluation and public information to levers for state policymakers to “pressure” schools into educational change. Although policymakers had hoped this approach would result in better academic outcomes for economically disadvantaged students and students from racial minority groups, the results have been underwhelming. The costs of the misuse of assessment, however, have been substantial. Since the passage of NCLB, for example, the focus on large scale standardized testing outcomes has crowded out policy discussions of more systemic changes to schools—such as the capacity building approaches that had been recommended by SBR, e.g., high-quality teacher professional development, instructional materials, and assessment.

Similarly, the role of international assessment has grown both within countries and across countries, albeit not at the pace and scope as in the United States. International large-scale assessments, like the PISA, PIRLS, and TIMSS, for example, have become popular metrics for assessing trends in the quality of educational systems around the world. And within countries, like Norway, Chile, and Singapore, large scale assessments have become a powerful lever for policymakers to track and pressure schools, although the variation in accountability policy approach remains significant.

References

- Au, W., 2007. High-stakes testing and curricular control: a qualitative metasynthesis. *Educ. Res.* 36 (5), 258–267. <https://doi.org/10.3102/0013189X07306523>.
- Carnoy, M., Rothstein, R., January 28, 2018. What Do International Tests Really Show about U.S. Student Performance? Economic Policy Institute. <https://www.epi.org/publication/us-student-performance-testing/>. (Accessed 7 July 2021).
- Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., et al., 1980. *Toward Reform of Program Evaluation: Aims, Methods, and Institutional Arrangements*. Jossey Bass, San Francisco.
- Elacqua, G., Hincapié, D., Martínez, M., 2019. The Effects of the Chilean School Accountability System on Teacher Turnover. No. IDB-WP-1075. IDB Working Paper Series.
- Elmore, R.F., Rothman, R., 1999. *Testing, Teaching and Learning: A Guide for States and School Districts*. National Academy Press, Washington, DC.
- Hopfenbeck, T., Tolo, A., Florez, T., El Masri, Y., 2013. *Balancing Trust and Accountability? the Assessment for Learning Programme in Norway*. Organization for Co-operation and Development (OECD).
- Hsieh, C.T., Urquiola, M., 2006. The effects of generalized school choice on achievement and stratification: evidence from Chile’s voucher program. *J. Publ. Econ.* 90 (8–9), 1477–1503.
- Klein, S.P., Hamilton, L.S., McCaffrey, D.F., Stetcher, B.M., 2000. *What Do Test Scores in Texas Tell Us?* The RAND Corporation, Santa Monica.
- Lehmann, I.J., 2004. The genesis of NAEP. In: Jones, L.V., Olkin, I. (Eds.), *The Nation’s Report Card: Evolution and Perspectives*. Phi Delta Kappa Education Foundation, Bloomington, IN, pp. 25–92.
- Lillejord, S., 2022. A more intelligent accountability: future directions. In: Davies, P., Rajeswari, R., Lincoln, A. (Eds.), *International Encyclopedia of Education*, fourth ed.
- Lingard, B., Lewis, S., 2016. Globalization of the Anglo-American approach to top-down, test-based educational accountability. In: *Handbook of Human and Social Conditions in Assessment*. Routledge, pp. 403–419.
- Lockheed, M., Prokic-Breuer, T., Shadrova, A., 2015. *The Experience of Middle-Income Countries Participating in PISA 2000–2015*. OECD Publishing, Paris.
- Loh, J., Hu, G., 2019. The Impact of Educational Neoliberalism on Teachers in Singapore. *Oxford Research Encyclopedia of Education*. <https://oxfordre.com/education/view/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-294>. (Accessed 6 July 2021).
- Ministry of Education – Singapore, 2010. *Recognising School Achievements in 2010*. https://www.nas.gov.sg/archivesonline/data/pdfdoc/20100926001/press_release_moa_2010.pdf. (Accessed 7 July 2021).
- National Academies of Sciences, Engineering, and Medicine, 2018. *How People Learn II: Learners, Contexts, and Cultures*. National Academy Press, Washington, DC. <https://doi.org/10.17226/24783>.
- National Research Council, 2011. *Incentives and Test-Based Accountability in Education*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/12521>.
- National Research Council, 1982. *Placing Children in Special Education: A Strategy for Equity*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/9440>.
- Ng, P.T., 2010. The evolution and nature of school accountability in the Singapore education system. *Educ. Assess. Eval. Account.* 22 (4), 275–292.

- Resnick, D.P., 1980. Chapter 1: minimum competency testing historically considered. *Rev. Res. Educ.* 8 (1), 3–29. <https://doi.org/10.3102/0091732X008001003>.
- Resnick, L.B., Resnick, D.P., 1992. Assessing the thinking curriculum: new tools for educational reform. In: Gifford, B.G., O'Conner, M.C. (Eds.), *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*. Kluwer Academic Publishers, Boston, pp. 37–75.
- Scherer, R., Nilsen, T., 2022. Accountability for the future? International testing and future directions. In: Davies, P., Rajeswari, R., Lincoln, A. (Eds.), *International Encyclopedia of Education*, fourth ed.
- Skedsmo, G., Mausesthaugen, S., 2016. Emerging accountability policies and practices in education: the case of Norway. In: Easley, J., Tulowitzki, P. (Eds.), *Educational Accountability: International Perspectives on Challenges and Possibilities for School Leadership*, first ed. Routledge, pp. 205–223.
- Shepard, L.S., 2008. A brief history of accountability testing, 1965–2007. In: Ryan, K.E., Shepard, L.S. (Eds.), *The Future of Test-Based Educational Accountability*. Routledge, New York, NY, pp. 25–46.
- Smith, M.S., O'Day, J., 1991. Systemic school reform. In: Fuhrman, S.H., Malen, B. (Eds.), *The Politics of Curriculum and Testing: The 1990 Yearbook of the Politics of Education Association*. Falmer Press, Bristol, PA, pp. 233–267.
- Tyack, D.B., 1974. *The One Best System: A History of American Urban Education*. Harvard University Press.
- Weinstein, J., Marfán, J., Horn, A., 2016. Chile: school leadership challenged by double accountability towards schools. In: Easley, J., Tulowitzki, P. (Eds.), *Educational Accountability: International Perspectives on Challenges and Possibilities for School Leadership*, first ed. Routledge, pp. 54–72.

Social, educational and legal perspectives on assessment and accountability measures in inclusive education for students with disability

J. Joy Cumming^a and Elizabeth A. Dickson^b, ^a Institute for Learning Sciences and Teacher Education, Australian Catholic University, Brisbane, QLD, Australia; and ^b School of Law, Faculty of Business and Law, QUT, Brisbane, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	171
Examining social, educational and legal considerations of inclusive assessment: Australia as a case study	172
Frameworks for inclusive assessment: social, education policy and practice, and legal	173
Social perspective	173
Education policy perspective	174
Legal perspective	176
Accountability at national and international level and inclusion of students with disability	176
Conclusion	178
References	178

Glossary

Adjustment A change made to an assessment item, form or condition to facilitate the needs of a student with disability. A change to the content demand of assessment may be termed accommodation.

Disability An impairment that affects a person's cognitive, physical or affective abilities that may be short term or long term.

Assessment The process of obtaining evidence of a person's learning or achievement.

Accountability The use of assessment evidence to make judgements, especially comparative judgements, about the effectiveness of an educational system, program or teaching.

Introduction

In this chapter we explore and extend our work on equity of assessment for students with disability in Australia (Cumming, 2008, 2012; Cumming and Dickson, 2007, 2013; Cumming et al., 2013; Dickson, 2003, 2012; Dickson and Cumming, 2018; Dickson et al., 2021; Rasooli et al., 2021). Over this time, we find there has been some crystallization of issues in the educational delivery of equitable assessments for students with disability but not necessarily clarity of procedures and processes for their implementation.

Three universally-accepted principles inform the education of children with disability. First, all children with disability have the right to education (UN Convention on the Rights of the Child, 1989; UN Convention on the Rights of Persons with Disabilities [CRPD], 2008; UNESCO, 1994). Second, such education should be "effective" (Salamanca Statement) to maximize opportunities and achievements for these students, with the expectation that policy and policy strategies will engage with a framework of "inclusive education" for as many students as possible, within inclusive education settings, and further creating more inclusive society (UNESCO, 1994). Third, students with disability have the same right as their peers without disability to have their achievement monitored to inform future instruction as well as to report progress (Cumming, 2012). A fourth principle, not so universal in implementation, is that students with disability have the right to participate in national and international monitoring of student achievements.

These principles have a strong basis in underlying philosophies of social justice and equity. In translation into conventions and policies, focus moves to equitable opportunities and removal of barriers and oppression (Rasooli et al., 2021; Sabbagh and Schmitt, 2016). Taking a human rights perspective, Rioux (2014) identified the responsibility of society to educate all students, including students with disability.

The responsibility of the school system is to develop and sustain a place of learning that enables every child to exercise her or his fundamental right to education and learning. It is about the right of children to an equal place in society (p. 135).

Increasingly, in policies to enact inclusive education, the categorization of students with disability has changed from reliance on a medical model that identifies a disability as a deficit in the person that can be diagnosed (medically) and labeled, to a social model of disability, whereby barriers to participation are created by society, and context, not the individual themselves. A more complex

model of disability that bridges the medical and social, the biopsychosocial model of disability, perceives barriers to participation as an interaction between the characteristics of the individual and the social context (World Health Organization, 2002).

From a social perspective, the idealized vision of education for students with disability is inclusive education without restrictions. Inclusive education developed philosophically as the opposite of exclusive education. Historically, since the introduction of public education for the masses, students with disability were educated separately from other students, in “special” education settings; many were not educated at all. In inclusive education, as many students with disability are educated within mainstream classrooms as possible. In the 21st century, this is interpreted as more than the mere presence of students with disability in a classroom; it is the provision of education that supports each student and seeks to remove restrictions and barriers to make the educational experience of these students as similar as possible to the experiences of their peers without disability.

Inclusive education requires an educational framework of policy and practice, including inclusive pedagogy (Florian and Black-Hawkins, 2011), and, as an integral component of education, inclusive assessment, identified in part as “a range of methods and strategies that aim to gather clear evidence about learners learning in non-academic areas as well as academic subjects” (European Agency for Special Needs and Inclusive Education [EASNIE], 2007). A common theme in international statements, policies and guidelines, as noted earlier in the Salamanca Statement, is to focus on optimal development and outcomes for students with disability. Therefore, by corollary, inclusive assessment practices that will enable students with disability to demonstrate their achievement optimally are essential. In keeping with changing models of disability, inclusive assessment is seen as a shift “from approaches that are used for most learners with something ‘additional’ or ‘different’” to address “deficit” needs of a student with disability, “to an approach that provides rich learning opportunities for everyone, so that they can all participate in classroom life” (Florian and Black-Hawkins, 2011, p. 216).

Inclusive education is often paired or discussed in terms of Universal Design of Learning (UDL), which originated from the development of inclusive building design to remove barriers that prevented accessibility for individuals with disability. Principles of UDL are intended to optimize access and provision for all through considering diverse and variable needs of individuals and providing, expressed most simply, “multiple means” of “engagement,” “representation” and “action and expression,” that is, for the last, to “differentiate the ways that students can express what they know” (CAST, 2021). Therefore UDL, and, more generally, inclusive assessment approaches, address “the design and use of fair and effective assessment methods and practices that enable all students to demonstrate to their full potential what they know, understand and can do” (Hockings, 2010, p. 34).

Internationally, in many countries, policies and frameworks for inclusive education for students with disability, and in some countries for inclusive assessment, are supported by legislation such as the introduction in the US of the *No Child Left Behind Act* (NCLB) (2002). Without elaborating on subsequent issues with this act for students with disability, its introduction was seen by educators in the US as significant at that time in recognizing that students with disability should be included in, not excluded from, national assessments to monitor student learning (McDonnell et al., 1997). Hence, the progress of these students was identified in law as important.

Examining social, educational and legal considerations of inclusive assessment: Australia as a case study

The Australian context¹ is similar to those of many other countries (Cumming, 2012). As a signatory to the many international conventions and policy statements about the rights of persons with disability and of children in general and in education, Australia has established education policy that endorses principles of inclusive education for students with disability (Australian Curriculum Assessment and Reporting Authority [ACARA], 2012; Cumming et al., 2013; Education Council, 2019). Statements of national priorities in Australian education, most recently the Alice Springs (Mparntwe) Education Declaration (Education Council, 2019), establish the goal that all students “at risk of educational disadvantage” should be able to reach their potential, including students with disability. “Equity” of opportunity is a priority to support their achievement and “achieve their highest educational outcomes” (p. 17). For students with disability, this can be seen as recognizing the individuality of each student and their learning needs, rather than categorizing students according to a medical definition of disability (Nisbet and Shaw, 2020).

Data and evidence of student learning and improvement are highlighted in the Mparntwe Declaration, both through national and international accountability assessment and school-level assessment that is “world-class,” including assessment “for,” “as” and “of” learning and assessment (Education Council, 2019, p. 15). Focus on assessment for learning has grown in the last two decades, seen as encompassing principles that make learning a key factor of assessment (Assessment Reform Group) (Broadfoot et al., 2002), and consequently making the student the center of learning, and assessment processes and goals. Such a central role is consistent with a social perspective of democracy that informs social justice views of inclusion (van Swet et al., 2011) as well as focusing on the experiences of the individual (Rasooli et al., 2021). Implicitly, such policies are intended to be enacted in practice at the system, school and individual classroom level.

¹A note: The Australian legal system is similar to many other systems in Western countries, emerging from the English common law system and involving both criminal and civil law. Unlike England but similar to the US, Australia does have a written Constitution which can also be the basis of legal challenges. However, unlike the US, Australia in general is not a litigious society, especially in education, with many challenges resolved through mediation or tribunals, rather than formal law courts. When Australian challenges in education reach the courts, they are frequently settled out of court, often to ensure no precedent is established (Cumming and Dickson, 2007). There is evidence, however, that legal challenges in education, and in education assessment, in Australia are increasing, especially in areas such as higher education which may have consequences for employment and where ambiguity exists about the role of students as consumers (Cumming, 2008; Pullin and Cumming, 2022).

Inclusive education policies are endorsed at state and territory level in Australia, with the intention to maximize success for all students with disability (see, e.g., [Education Queensland, 2021](#)), although we have found that practical guidance for teachers in how to provide such world-class assessment for learning, and especially for students with disability, is limited ([Cumming and van der Kleij, 2016](#); [van der Kleij et al., 2018](#)). These limits are reflected in the key issues in inclusive assessment that are universal that are discussed later.

As in other countries, Australia's inclusive education policies are reflected in and strengthened through legal frameworks that address potential discrimination against students with disability in education ([Dickson and Cumming, 2018](#); *Disability Discrimination Act 1992* (Cth)² [DDA]; *Disability Standards for Education 2005* ["Standards"]). Discrimination may be direct, treating a student less favorably because of their disability. Examples would be denying enrollment based on a disability, or denying a right to participate in an assessment because of a disability ([Dickson and Cumming, 2018](#)). It may also be indirect, requiring a student to comply with an unreasonable condition that they cannot meet because of a disability, and which causes them disadvantage. An example would be requiring a student with vision impairment to read the same printed format assessment as students without such an impairment ([Dickson and Cumming, 2018](#)). In Australian law, disability is defined broadly (*Disability Discrimination Act 1992* (Cth)), encompassing physical, cognitive, psychiatric, behavioral and sensory disabilities, including disability that is shortterm, ongoing and/or potential. The *Standards*, which draw on the DDA definition of disability, identify the requirements for education providers for equitable education for students with disability.

Two core principles are identified in the *Standards*. Students should be able to engage in their education "on the same basis" (§ 2.2) as students without disability and all aspects of their education should be "accessible" (§ 6.3). To provide such equal opportunity and accessibility, providers—system and school level—are to provide "adjustments," "a measure or action" that assists a student with disability (§ 3.3) with the proviso that such adjustments must be "reasonable" (§ 3.4). Reasonableness is to be determined by consideration of a student's needs, in consultation with the student and their parent or carer, about the student's disability and its intersection with their education (§ 3.5). The effect of adjustments on a student's ability to participate in courses or programs, the effect on others, costs and benefits of making an adjustment, and the "balance of interests of all parties affected" are also relevant to reasonableness. Provision of an adjustment should not impose unjustifiable "hardship" on the provider (§ 3.4 Note).

Assessment and certification requirements are identified as an area where adjustments to enable full participation should be made. Assessment requirements should be accessible for the student, with "assessment procedures ... adapted to enable the student to demonstrate the knowledge, skills or competencies being assessed" (§ 6.3). Again, consultation with the student or their parent or carer should identify whether an adjustment is necessary and whether an adjustment can be identified to achieve this aim (§ 6.2). Assessment adjustments must also be reasonable, and importantly,

a provider may continue to ensure the integrity of its courses or programs and assessment requirements and processes, so that those on whom it confers an award can present themselves as having the appropriate knowledge, experience and expertise implicit in the holding of that particular award.
(§ 3.4 Note).

Under law, the principles of inclusive education are to be met by adjustments to suit the needs of the student with disability. As such, delivering education "on the same basis" to students with disability may mean treating them differently, rather than the same, potentially a deficit model of disability. Moreover, in practice, adjustments are often determined through considerations from the medical model of disability, relying on diagnosis, often medical, and certification as evidence of disability. The extent to which adjustments should be provided are limited by perceptions of reasonableness and, for assessment, integrity. However, the underlying philosophy is to ensure that students with disability are consulted about and able to access full education. Students with disability are accorded rights to quality education that are not generalized to other students in Australia law.

Frameworks for inclusive assessment: social, education policy and practice, and legal

It is clear that different philosophies, policies and legislation create different frameworks for inclusive assessment for students with disability.

Social perspective

From the social perspective, inclusion is "aspirational in nature" ([Finkelstein et al., 2021](#), p. 737). The fullest social perspective would provide every student with disability with all educational opportunities they need to learn, and to demonstrate their learning, that

²Australian discrimination legislation exists at both federal ("Commonwealth") and state or territory levels. The state and territory acts (*Discrimination Act 1991* (ACT); *Anti-Discrimination Act 1977* (NSW); *Anti-Discrimination Act* (NT); *Anti-Discrimination Act 1991* (Qld); *Equal Opportunity Act 1984* (SA); *Anti-Discrimination Act 1988* (Tas); *Equal Opportunity Act 2010* (Vic); *Equal Opportunity Act 1984* (WA)) address a range of attributes on which discrimination may occur and specifically include disability or impairment. The federal act is specific to disability. The federal, state and territory acts are similar and do not cause conflict in application in courts and tribunals. The *Disability Standards for Education* are authorized by and subordinate to the *Disability Discrimination Act 1992* (Cth).

is, there would be no restrictive parameters. However, social perspectives of equity and fairness in assessment, in themselves, are not simple. The most recent edition of the *American Standards for educational and psychological testing* (American Educational Research Association et al., 2014) provides a chapter on fairness in testing, without defining the concept but noting that “fairness ... is an overriding, foundational concern” (p. 49). Social justice and fairness derive from social context (Nisbet and Shaw, 2020; Rasooli et al., 2021). Education systems in all countries rely on funding allocations, mostly public funding derived from taxation, for provision. Education funding is in competition with health funding, resource funding and social welfare provision, to name a few. Resources, both financial and human, are finite. When resources are finite, decisions have to be made about how to distribute these resources. Not only may education be perceived as a different priority in many countries, different decisions within education may prioritize different student groups or different outcomes (Deutsch, 1975). In terms of equity and fairness in social justice, one area where alternative considerations arise in assessment is whether the focus is on fair process or fair outcome (McArthur, 2016). These are complex issues for which there is no right or wrong answer, despite the sense that most countries are committed to inclusive education for students with disability, and, implicitly, inclusive assessment practices.

Education policy perspective

From an education policy perspective, policy expressions of social inclusion goals may also be aspirational, such as the first Australian national goal to “provide all young Australians with access to high-quality education that is inclusive and free from any form of discrimination” (Education Council, 2019, p. 5). However, given, as noted, that resource allocations may restrict inclusive education in practice, the goal of inclusive education may be more appropriately considered as ambitious. In most countries it is also seen as work-in-progress. For example, a policy document noted earlier marries social ambition with practical resource constraints, referring to the “next phase of reform as we continue our journey toward a more inclusive education system” while also noting “[transition] to a new resource allocation model for students with disability” (Education Queensland, 2021). Professional development of education personnel is seen as critical to developing inclusive education, although inclusive assessment is rarely mentioned in policies (see, e.g., Education Queensland, 2021), despite the recognized difficulty of changing teacher assessment practices to be more inclusive (Deneed and Boud, 2014; Offerdahl and Tomanek, 2011). Teacher preparation programs are expected to develop graduate teachers who are able to engage with students with disability (EASNIE, 2021), however, again, guidelines linked to such expectations are usually generic rather than explicitly focused on inclusive assessment (see, e.g., Australian Institute for Teaching and School Leadership, 2020).

Policy frameworks do not intentionally impose barriers on inclusive assessment. However, *from an educational practice perspective*, three parameters emerge that affect implementation. The first parameter, which is linked closely to the third, relates to teachers’ knowledge of how to practice inclusive assessment and the limitations on their willingness to engage with such assessment practice. As noted, practical guidance is still limited. The second parameter, which is not to be dismissed lightly, is the perspectives of students with disability themselves. These students may not wish to identify as having a disability, setting themselves apart from their peers (O’Rourke and Houghton, 2008; Woodgate et al., 2020). They may not perceive themselves as having a “legitimate” disability that would enable them to have considerations of assessment approaches in place, and, they themselves may also perceive that inclusive assessment practices may be “unfair” and create an advantage for them (Tomlinson and Moon, 2013).

The third parameter is the approach most frequently taken to inclusive assessment in practice, as identified in the Australian law, the provision of adjustments to mass assessments designed for the majority of students to make them suitable for students with disability. In the US such adjustments are referred to as accommodations or adaptations (AERA et al., 2014). The principle of an adjustment is that it makes changes to an assessment design to be accessible to a student with disability, based on their needs. Ideally, under UDL, assessments would be designed to suit all students so that adjustments or accommodations are not necessary. However, from the perspective of “critical realism” (Pawson, 2013), this is unlikely, given the need for due consideration of a multiplicity of factors: the disability of a student, the context and background of the student, the area being assessed, and the setting. In practice, this is difficult, especially for larger-scale tests. Differentiation of assessment design and modes becomes the critical focus in practice. Principles of UDL provide guidelines for diverse assessment development to reduce “noise” in assessment tasks that affects the ability of students with disability to demonstrate their knowledge (AERA et al., 2014). From a validity perspective, the most important theoretical construct in quality assessment, that is, that an assessment is measuring the learning it is intended to assess, and an appropriate interpretation can therefore be made about the learning that has occurred. “Noise” may mean that an assessment outcome and interpretations made from the outcome may be invalid.

The key point in assessment adjustments in accessibility. Although assessors may have demonstrated reluctance to change assessment practices, many alternative modes of determining what a student with disability knows and can do are available. One approach to identifying what learning is to be demonstrated, drawing on alignment with the curriculum that has been taught (or was intended to be taught), is to identify the “target” knowledge and skills that are to be assessed, and “access” skills, the skills required to complete the assessment (Dembitzer and Kettler, 2018; Kettler and Elliot, 2010). Target skills are the essential elements of an assessment. “Access” skills may impede the outcomes for target knowledge and skills that a student with disability can demonstrate, for example, reading print material may not be a suitable target for students with vision impairment, but a different form of

“reading” may be able to be substituted. Where a student’s disability is such that the intended target skills are not appropriate for their age/peer level, then target skills may also be addressed.³

Core assessment issues of validity, interpretation and reporting arise when target skills, and on occasion, access skills, are adjusted. As noted, Australian law indicates that adjustments are not required if the “integrity,” hence, validity, of an assessment may be compromised. What is reported should be what a student can do. It is frequently commented that even students with disability have the right to fail if they do not know what is being assessed, if the assessment itself is a fair task. For many assessment adjustments, an issue is whether there has been reduction of cognitive load, with consequent impact on the validity of the assessment (Andersen et al., 2018). Conversely, cognitive load may be increased for students with disabilities in comparison with their peers without disability, if a task requires excessive and inappropriate access skills.

When different forms of adjustments (and modifications) occur, grading and reporting are an issue that authorities, teachers, students and parents must address. One consideration is whether grading and progress should be peer-related, standards-related or ipsative, depending on the nature of a student’s disability and their education program (Quenemoen et al., 2003). Again, there is no easy answer to this. If the import of an assessment adjustment is to enable the student with disability to be assessed on the same basis as other students, while preserving the integrity of an assessment so that the outcome does not purport that a person can do something that they are not independently capable of doing, then the question is whether the reported outcome should indicate that the assessment has been modified. Again, this is contextual. A noted legal case in Ireland concerned a challenge against an amendment to certification that noted the nature of an adjustment (for discussion of the case and issue see Nisbet and Shaw, 2020, p. 82). Many drivers, like the authors of this chapter, wear glasses to compensate for a degree of vision impairment. Driving assessment is undertaken with the candidate wearing glasses. A resulting driving license has a condition that glasses should be worn when driving. This does not cause sensitivity for the driver but does alert others to conditions under which the driver can operate independently and safely.

In principle, schools and classrooms have a range of options available to assess the learning of students with disability. In principle also, such options could be made available for all students. A simple example is a critical “reading” of a novel for more mature students. The target skill would be understanding the author’s intentions, character development, the nature of a plot and so on. There may be a choice of novels. If some students with disability are not able to work at the level of their peers, they could be set a more simple novel. The act of reading as decoding is not necessarily a criterion for successful completion of the task. Students with a visual impairment may have an audiotape of the novel, some students may view a fair visual representation of the novel through a video or film. A degree of scaffolding to structure responses might be provided for all students to assist in the task. Students could then be offered a range of presentation formats from an essay, handwritten or with an option to use technology, including voice technology—(really more appropriate in the 21st century for all students), to an oral presentation with or without technology, or pre-recorded, to the class of students or in a singular setting with a teacher for students whose disability characteristics include anxiety. These are just some of the options that could be available for the task, with the focus on whether the student can demonstrate the understandings of the novel that are the essential target skills. These options have been recorded in research on classroom assessment with students with disability, and adjustments in education (Dickson and Cumming, 2018). They do require teachers who are confident in equating the demands of assessment tasks and understanding the needs of their students with disability.

However, in practice, most classroom assessment adjustments mirror the range of assessment options available in national assessment programs and/or standardized test settings. It is in this area that most research on adjustments/accommodations/adaptations for students with disability has been undertaken. In these settings, adjustments commonly include different formats such as font size or technology assistance, availability of scribes and readers in some circumstances, students placed in separate venues to minimize distraction to the student with disability and to other students, and, most frequently, extra time to complete an examination or test. There is limited research still on the effectiveness of these adjustments and optimal adjustments, such as the sufficiency of extent of extra time. A number of theoretical frameworks have been used to examine the impact of accommodations. For example, one approach is that any accommodation, even if provided to all students, should provide most benefit to students with disability, a differential boost (Sireci et al., 2005), alternatively considered as a maximum potential thesis [MPT] (Zuriff, 2000).

A concern noted with most empirical research on adjustments in these contexts is that they address students as group members, whereas most students with disability have idiosyncratic characteristics affecting their learning and assessment. A further issue with most of the available empirical research is that while studies which address adjustments need to address access skills, the focus is on student achievement outcomes, rather than considering the appropriateness, that is, fairness, of the process (Nisbet and Shaw, 2020), and accessibility of assessment tasks. Little empirical evidence still exists on the effectiveness of different adjustments: how much extra time is desirable to enable optimal demonstration of a student’s knowledge and skills without leading to boredom or lack of concentration; under what circumstances should tasks be read aloud or does this increase the cognitive load for the student with disability who has to remember as well as respond, as noted by students without disability (Kim, 2012); how much time is needed for a student who is reading braille (Ragosta and Wendler, 1992). By contrast, both educators and students express concern about the fairness and potential advantage offered to students with disability when alternative forms of assessment are available (O’Neill, 2017; Rasooli et al., 2021; Tomlinson and Moon, 2013).

³In the US, such adjustments are recognized as modifications to the assessment goal (Kettler et al., 2018).

Legal perspective

The appropriateness of adjustments, and student outcomes, are the focus of most considerations of assessments for students with disability from the *perspective of law*. In Australian law,⁴ most challenges are based on the discrimination laws, and the DSE, including the concept of “reasonable adjustments” versus “hardship.” As [Dickson and Cumming \(2018, p. 321\)](#) summarized, such challenges have addressed denial of adjustments by the education provider, and hence indirect discrimination by requiring students to comply with unreasonable conditions which disadvantage them in completing assessments. Assessment adjustments considered in the context of legal challenges include

- Extra time to finish an exam or assignment
- Completing a course over a longer than usual time
- Supervised rest, food and medication breaks in examinations
- Adjustment to the format of an exam paper—font, paper color and size, paper “masks”
- Separate venues to minimize distraction or to accommodate assistance animals
- Alternatives to writing—viva voce, examination, scribe, assistive technology
- Alternatives to reading—brailled and/or taped materials, assistive technology
- Alternatives to hearing—written stimulus materials, translators, assistive technology
- Alternatives to speaking—written rather than spoken responses allowed
- Adjusted scheduling of assessment.

Three challenges noted by Dickson and Cumming are of a different nature. Two relate more specifically to how assessments are used to obtain a grade and report on the grade.

- Adjusted level of achievement thresholds where performance on a particular item has been compromised by disability (“special consideration”)
- Adjusted weighting of assessment for students who have missed an assessment item because of disability

One challenge relates to direct discrimination, less favorable treatment alleged by a student who wished to undertake an assessment but was excluded by the authority on the basis of their disability. This enforces the sense that students with disability do not wish to be seen as different from their peers. A further note is that in some of the cases, students were doing well or had potential to do well in their studies, that is, better than average, if assessment adjustments they identified as appropriate were provided. In these cases, concern was expressed in the court judgments that adjustments should not advantage students with a disability. Students are likely to be successful in challenges under the DSE if they are able to establish that an adjustment would have enabled them to achieve a better result, where an inappropriate or no adjustment has been provided. They are generally not likely to be successful if the education provider can show that consideration had been made of the nature of the student disability, adjustments were provided, and these were deemed reasonable ([Dickson and Cumming, 2018](#)).

From a legal perspective in Australia, therefore, assessment adjustments should be provided to address the consequences of a disability of a student, unless this compromises the integrity of the assessment. Integrity or validity are not issues which are argued frequently in Australian court cases, although these have occurred in cases in higher education ([Dickson, 2012](#)), but are more frequently raised as validity issues in cases elsewhere, especially in the US ([Pullin and Cumming, 2022](#)). However, clear evidence of a failure to make reasonable adjustment and related indirect or direct discrimination may lead to a successful challenge by an individual student with disability. From the student’s perspective, this should ensure a fairer process and lead to a better outcome. From an assessment perspective, this would increase the validity of the assessment and most importantly interpretations made from the assessment. This brings us to the core issue of national (Australia) and international test programs and the assessment of students with disability.

Accountability at national and international level and inclusion of students with disability

Many countries have national assessment programs intended to monitor student progress. Different reasons are proffered for these programs from monitoring the educational quality of a state or district, the quality of a school, the quality of a teacher, and the learning and progress of an individual student ([Cumming et al., 2019](#)). Test formats vary from traditional paper-and-pencil to more recent use of technology and ICT ([Eurydice, 2009](#)). Where digital assessments are being used, countries are also making use of adaptive testing, where the achievement level of students is gauged from initial questions, and the student is then scaffolded as appropriate into more difficult or more easy questions ([ACARA, 2016c; Eurydice, 2009](#)). Adaptive tests are seen as more suitable for assessment with students with disability as they enable greater differentiation without increasing the length of a test, reduce

⁴As noted, Australia does not have a history as a litigious nation in education, with most challenges settled through mediation or out of court, including cases where students were found to have been taught the wrong curriculum for a certification examination, similar to cases regarding assessment and opportunity to learn in the US. The cases noted in [Dickson and Cumming \(2018\)](#) represent those that have made it to the courts. For discussion on US assessment law cases, see, for example.

anxiety, increase student engagement, and reduce student discouragement due to repeated failure, while at the same time enabling high achieving students to have more challenging tasks (ACARA, 2016c).

The Australian National Assessment Program (NAP) monitors the achievement of all students in literacy and numeracy in specific year levels in Australian schools (National Assessment Program—Literacy and Numeracy [NAPLAN]) and conducts sample assessments in other areas. In addition to national assessments, the NAP incorporates participation in international sample tests such as PISA, TIMSS and PIRLS to enable international comparisons.

The Australian NAPLAN and other national NAP tests make provision for adjustments for students with disability. With the recent introduction of online testing, these include

- Assistive technology
- Alternative format tests
- Scribes and support persons
- Reading to students
- Extra time and rest breaks
- Separate location (ACARA, 2016a).

Online tests also enable “alternative questions” that can be responded to by students with disability engaging with, to use prior terminology, the same target skills as the original question formats. Alternative questions may be through audio, visual and colored presentations (ACARA, 2016a). Some adjustments will assist in reducing the cognitive demands on a student with disability. Clear guidelines are provided to ensure that no advantage is provided to a student with disability and that the resources used are those that the student would normally access in the classroom. By contrast, not all resources a student might use in the classroom will be available for NAPLAN (ACARA, 2016b; Davies et al., 2016). The expectation is that all students will complete NAPLAN and NAP assessments, however, “some students with significant intellectual disability and/or students with significant coexisting conditions, which severely limit their capacity to participate in the tests, may be exempted from sitting the tests.” (ACARA, 2016b). In 2019, between 1.6% and 1.8% of students were exempted from assessments across the year levels of testing (ACARA, 2019). According to national data, in 2020, some 20% of Australian students are identified as having a disability that requires some form of school support, with 1.6% having an “extensive” disability, and 3.4% having a “substantial disability” (ACARA, n.d.). While these figures may be consistent with those who were exempted from NAPLAN testing, the learning progress of a percentage of students is not being captured. Students who are excluded are directly discriminated against in the NAPLAN assessments on the basis of their disability. In terms of adjustments, NAPLAN is more inclusive than previously (Cumming and Dickson, 2013; Cumming et al., 2013). However, a key issue remains. A large number of students with disability may be receiving adjustments in completing the assessments. Information on the performance of these students and the nature and quantity of adjustments are not reported to the public. The implicit assumption is that the testing format and adjustments provided are valid and appropriate to allow students with disability to demonstrate their learning in these key areas at an optimal level. Conversely, a further implicit assumption is that interpretation of the achievement of these students in comparison with other students is also valid.⁵ These are major assumptions regarding the appropriateness of these assessments to judge the learning of students with disability, and of their cohort and school. For students who do participate in the tests, where the adjustments that are made are not sufficient to address their needs due to the consequences of their disability, indirect discrimination has occurred.

Like a large number of countries internationally, the NAP includes participation in international tests such as TIMSS and PIRLS (<https://timssandpirls.bc.edu>) and PISA (<https://www.oecd.org/pisa/>). The purposes of these tests are mixed. They provide global comparisons of achievement for participating countries, although they have also been defined as showing “how education systems provide educational opportunities to their students, as well as factors that influence how students use these opportunities” (International Association for the Evaluation of Educational Achievement [IEA], 2021). This could be interpreted as consideration of how countries address education of the diversity of students, including students with disability.

Tests are conducted by sample testing. However, TIMSS and PIRLS still exclude students in schools that are only for “special needs students” including students with “functional disabilities” who cannot complete the TIMSS tests and students with “intellectual disabilities” including students “who are emotionally or mentally unable to follow even the general instructions of the test.” Students with learning disabilities are to be “accommodated,” rather than excluded (LaRoche et al., 2020, Chapter 3.6). The Australian TIMSS report indicates that “non-mainstream schools (such as schools for students with special needs)” were excluded with further within-school exclusions of students with intellectual and functional disabilities (Thomson et al., 2020, p. 153). In Australia, for the 2018 PISA round, students were excluded on the basis of “moderate to severe permanent physical disability,” or cognitive delay due to “cognitive, behavioral or emotional disability” (Thomson et al., 2019, p. 247). OECD statistics for the same round indicate 623 Australian students were excluded due to special education needs (OECD, 2019, Table I. A2.4). Standard adjustments for standardized tests are now available for school administration (cf Cumming and Dickson, 2013; Cumming et al., 2013) and PISA is noted as aiming “to be as inclusive as possible ... [with] strict guidelines enforced with regard to the percentage of schools and of students that could be excluded” (Thomson et al., 2019, p. 6). For PISA, the intention was that special education settings would be included in sampling, however, in Australia, these schools are only attended by students with disability who would

⁵It should be noted, however, that interpretation of NAPLAN scores for individual students with or without disability, and their learning progress, is problematic due to measurement error (McGaw et al., 2020; Wu, n.d.). For students at the extremes of the distribution of scores for their Year level, as might occur for many (but not all) students with disability, the problem increases (McGaw et al., 2020).

meet the exclusion criteria. Alternative assessment booklets are available for students with disability, but were not used in Australia. Many countries have been reported as providing participating students with disability with the standard range of test adjustments to ensure that they are not practicing indirect discrimination according to their legislation (LeRoy et al., 2019).

National and international test regimes will undertake preparatory, and ex post facto, analyses to establish the construct validity, that is, the match of the test focus, to desired curriculum learning. In administering such tests, given the focus on comparability for interpretation, the emphasis is on reliability, rather than validity, with limited involvement of students with disability a consequence (LeRoy et al., 2019). Consistency in administration and scoring become important. Standardization of procedures ensures such comparability. There have been developments to make these tests more inclusive, enabling participation of students with disability. However, students may not only be excluded but also little information is provided about the nature and extent of any adjustments provided for students with disability—"their presence is invisible" (LeRoy et al., 2019, p. 3). If one purpose of such assessments is to inform countries on the quality of their education provision for all students, including students with disability who must be included in the sampling, there is an "informational vacuum" with respect to the education of students with disability (p. 3). The overall validity of these international accountability assessments in terms of student outcomes and interpretation of overall performance may be questioned.

Conclusion

Inclusive education of students with disability is an idealized international and national goal. Inclusive assessment practices with students with disability are an essential component of education overall. Models of disability increasingly recognize the situational nature of disability. Disability is not a medical diagnosis but arises from the social and practical consequences of the characteristics of a disability in interaction with society. From a social perspective, inclusive education is sited within human rights, social justice and equity. Through such perspectives concern is with fairness, a concept which has defied precise definition. At the ideal, inclusive education would result from principles of Universal Design, so that all are included, all are different, and education goals, engagement and outcomes are clear for all students. Inclusive education as a social endeavor has been translated into national and international policies. Providers and schools endeavor to implement inclusive education through the full inclusion of students with disability, and hence to ensure that students with disability are able to access and engage in educational assessments. At a realistic level, and within societal context in the provision of education, Universal Design is seen to have limitations. Differentiation in provision and assessment become the goal, keeping expectations for the achievements of all students, including students with disability, high. In educational practice in schools, administrators, teachers and students work through the daily challenge of addressing the needs of students with disability to ensure that the education provided is as fulsome as possible, and that educational assessments are accessible and allow students to demonstrate their achievement at as high a level as possible. In most countries, as exemplified by Australia, these ambitions are supported by legislation that prevents discrimination, either direct or indirect, against students with disability.

Inclusive education has been noted to be aspirational, ambitious, and hopefully moving forward. Some clarity of processes and procedures is emerging through policies, practices, and to a lesser extent, the law. Differentiation addresses the consequences of the characteristics of students with disability to enable them to be on a level playing field with their peers without disability. It cannot remove the disability, it is not a deficit model, it is a practical approach. However, enacting differentiation in assessment is complex and demanding. It requires considerable experience and confidence of educators, that is yet to be fully realized. It requires further empirical research on what is the best approach to assessment for students with disability to enable optimal but valid achievement. The challenges that remain are the beliefs and experiences that shape teacher, student and community experiences: perceptions of fairness, equity, advantage and disadvantage. From a social perspective these are broad issues, from the educational assessment perspective, the options are more restricted, from a legal perspective, within educational settings, interpretation of legislation and provision are further restricted. Unfortunately, from national and international accountability assessment perspectives, these are unknowns.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Australian Curriculum, Assessment, Reporting Authority, 2012. Curriculum, Assessment and Reporting in Special Educational Needs and Disability: A Thematic Overview of Recent Literature. https://docs.acara.edu.au/resources/ACARA_Research_for_Publication_Final.pdf.
- Andersen, L., Nash, B.L., Bechard, S., 2018. Articulating the validity evidence for a science alternate assessment. *J. Res. Sci. Teach.* 55 (6), 826–848. <https://doi.org/10.1002/tea.21441>.
- Australian Curriculum, Assessment and Reporting Authority, 2016a. NAPLAN. Adjustments for Students With Disability. <https://www.nap.edu.au/naplan/school-support/adjustments-for-students-with-disability>.
- Australian Curriculum, Assessment and Reporting Authority, 2016b. NAPLAN. Parent/Carer Support. <https://www.nap.edu.au/naplan/parent-carer-support/>.
- Australian Curriculum, Assessment and Reporting Authority, 2016c. Tailored Tests. <https://www.nap.edu.au/online-assessment/research-and-development/tailored-tests>.
- Australian Curriculum, Assessment and Reporting Authority, 2019. National Assessment Program. Literacy and Numeracy. Achievement in Reading, Writing, Language Conventions and Numeracy. National Report for 2019. <https://nap.edu.au/docs/default-source/default-document-library/2019-naplan-national-report.pdf?Status=Temp&sfvrsn=2>.
- Australian Curriculum, Assessment and Reporting Authority, (n.d.). School Students With Disability. <https://www.acara.edu.au/reporting/national-report-on-schooling-in-australia/national-report-on-schooling-in-australia-data-portal/school-students-with-disability#view1>.

- Australian Institute for Teaching and School Leadership, 2020. Spotlight. Inclusive Education: Teaching Students With Disability. <https://www.aitsl.edu.au/research/spotlight/inclusive-education-teaching-students-with-disability>.
- Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M., Stobart, G., 2002. *Assessment for Learning: 10 Principles. Research-Based Principles to Guide Classroom Practice*. Assessment for Learning, England.
- CAST Inc, 2021. About Universal Design for Learning. <https://www.cast.org/impact/universal-design-for-learning-udl>.
- Cumming, J., 2008. Legal and educational perspectives of equity in assessment. *Assess. Educ. Princ. Pol. Pract.* 15 (2), 123–135. <https://doi.org/10.1080/09695940802164168>.
- Cumming, J., 2012. *Valuing Students With Impairment*. Springer.
- Cumming, J.J., Dickson, E.A., 2007. Equity in assessment: discrimination and disability issues from an Australian legal perspective. *Educ. Law* 19 (3–4), 201–220.
- Cumming, J., Dickson, E., 2013. Educational accountability tests, social and legal inclusion approaches to discrimination for students with disability: a national case study from Australia. *Assess. Educ. Princ. Pol. Pract.* 20 (2), 221–239. <https://doi.org/10.1080/0969594X.2012.730499>.
- Cumming, J., van der Kleij, F., 2016. Effective enactment of assessment for learning with students with diverse backgrounds in Australia. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer, Dordrecht, The Netherlands, pp. 55–73.
- Cumming, J., Dickson, E., Webster, A., 2013. Reasonable adjustments in assessment: putting law and policy into practice. *Int. J. Disabil. Dev. Educ.* 60 (4), 295–311. <https://doi.org/10.1080/1034912X.2013.846467>.
- Cumming, J.J., van der Kleij, F.M., Adie, ., 2019. Contesting educational assessment policies in Australia. *J. Educ. Pol.* 34 (6), 836–857. <https://doi.org/10.1080/02680939.2019.1608375>.
- Davies, M., Elliot, S., Cumming, J.J., 2016. Documenting support needs and adjustment gaps for students with disabilities: teacher practices in Australian classrooms and on national tests. *Int. J. Incl. Educ.* 20 (4), 1252–1269. <https://doi.org/10.1080/13603116.2016.1159256>.
- Dembitzer, L., Kettler, R.J., 2018. Testing adaptations: research to guide practice. In: Elliott, S., Kettler, R., Beddow, P., Kurz, A. (Eds.), *Handbook of Accessible Instruction and Testing Practices*. Springer, Dordrecht, The Netherlands, pp. 213–230.
- Deneen, C., Boud, D., 2014. Patterns of resistance in managing assessment change. *Assess. Eval. High. Educ.* 39 (5), 577–591. <https://doi.org/10.1080/02602938.2013.859654>.
- Deutsch, M., 1975. Equity, equality, and need: what determines which value will be used as the basis of distributive justice? *J. Soc. Issues* 31 (3), 137–149.
- Dickson, E., 2003. Understanding disability: an analysis of the influence of the social model of disability in the drafting of the Anti-Discrimination Act 1991 (Qld) and in its interpretation and application. *Aust. N. Z. J. Law Educ.* 8 (1), 47–70.
- Dickson, E., 2012. The assessment of students with disabilities: the Australian law as to reasonable adjustment and academic integrity. *Int. J. Law Educ.* 17 (2), 49–62.
- Dickson, E., Cumming, J., 2018. Reasonable adjustment in assessment: the Australian experience. In: Trimmer, K., Dixon, R.M., Findlay, Y. (Eds.), *The Palgrave Handbook of Education Law for Schools*. Palgrave Macmillan, pp. 315–333.
- Dickson, E., Cumming, J., Wyatt-Smith, C., Webster, A., 2021. Reasonable adjustment to assessment: what do Australian teachers need to know and consider? *Int. J. Law & Educ.* 24, 62–78.
- Education Council, 2019. Alice Springs (Mparntwe) Education Declaration. <https://www.dese.gov.au/alice-springs-mparntwe-education-declaration>.
- Education Queensland, 2021. Every Student With Disability Succeeding Plan 2021–2025. <https://education.qld.gov.au/student/Documents/every-student-with-disability-succeeding-plan.pdf>.
- Eurydice, 2009. *National Testing of Pupils in Europe: Objectives, Organisation and Use of Results. A Report for the Education, Audiovisual and Culture Executive Agency*. Eurydice, Brussels, Belgium.
- European Agency for Special Needs, Inclusive Education, 2007. *Assessment in Inclusive Settings—Key Issues for Policy and Practice Report*. <https://www.european-agency.org/resources/publications/assessment-inclusive-settings-key-policy-messages>.
- European Agency for Special Needs, Inclusive Education, 2021. *Teacher Education for Inclusion. Country Report*. <https://www.european-agency.org/projects/te4i>.
- Finkelstein, S., Sharma, U., Furlonger, B., 2021. The inclusive practices of classroom teachers: a scoping review and thematic analysis. *Int. J. Incl. Educ.* 25, 735–762. <https://doi.org/10.1080/13603116.2019.1572232>.
- Florian, L., Black-Hawkins, K., 2011. Exploring inclusive pedagogy. *Br. Educ. Res. J.* 37 (5), 813–828.
- Hockings, C., 2010. *Inclusive Learning and Teaching in Higher Education: A Synthesis of Research*. EvidenceNet, Higher Education Academy. <https://www.advance-he.ac.uk/knowledge-hub/inclusive-learning-and-teaching-higher-education-synthesis-research>.
- International Association for the Evaluation of Educational Achievemen, 2021. *PIRLS 2021*. <https://www.iea.nl/studies/iea/pirls/2021>.
- Kettler, R., Elliott, S., 2010. Assessment accommodations for children with special needs. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*. Elsevier, pp. 530–536.
- Kettler, R.J., Elliott, S., Beddow, P.A., Kurz, A., 2018. Accessible instruction and testing today. In: Elliott, S., Kettler, R., Beddow, P., Kurz, A. (Eds.), *Handbook of Accessible Instruction and Testing Practices*. Springer, Dordrecht, The Netherlands, pp. 1–16.
- Kim, J.S., 2012. The effect of “read-aloud” as a test accommodation for students with visual impairments in South Korea. *J. Vis. Impair. Blind.* 106 (6), 356–361.
- LaRoche, S., Joncas, M., Foy, P., 2020. Sample design in TIMSS 2019. In: Martin, M.O., von Davier, M., Mullis, I.V.S. (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*. IEA. Chapter 3.
- LeRoy, B.W., Samuel, P., Deluca, M., Evans, P., 2019. Students with special educational needs within PISA. *Assess. Educ. Princ. Pol. Pract.* 26 (4), 386–396. <https://doi.org/10.1080/0969594X.2017.1421523>.
- McArthur, J., 2016. Assessment for social justice: the role of assessment in achieving social justice. *Assess. Eval. High. Educ.* 41 (7), 967–981. <https://doi.org/10.1080/02602938.2015.1053429>.
- McDonnell, L.M., McLaughlin, M.J., Morison, P. (Eds.), 1997. *Educating One & All. Students With Disabilities and Standards-Based Reform*. National Academy Press, Washington, D.C.
- McGaw, B., Loudon, W., Wyatt-Smith, C., 2020. *NAPLAN Review. Final Report*. https://naplanreview.com.au/pdfs/2020_NAPLAN_review_final_report.pdf.
- Nisbet, I., Shaw, S., 2020. *Is Assessment Fair?* SAGE, London, England.
- No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319, 2002.
- OECD, 2019. *PISA 2018 results. In: What Students Know and Can Do, vol. I*. <https://www.oecd-ilibrary.org/docserver/5f07c754-en.pdf?expires=1635916546&id=id&accname=guest&checksum=4A9AD4C65C6F533F0E6E97352A3F6376>.
- Offerdahl, E.G., Tomanek, D., 2011. Changes in instructors’ assessment thinking related to experimentation with new strategies. *Assess. Eval. High. Educ.* 36 (7), 781–795. <https://doi.org/10.1080/02602938.2010.488794>.
- O’Neill, G., 2017. It’s not fair! Students and staff views on the equity of the procedures and outcomes of students’ choice of assessment methods. *Ir. Educ. Stud.* 36 (2), 221–236. <https://doi.org/10.1080/03323315.2017.1324805>.
- O’Rourke, J., Houghton, S., 2008. Perceptions of secondary school students with mild disabilities to the academic and social support mechanisms implemented in regular classrooms. *Int. J. Disabil. Dev. Educ.* 55 (3), 227–237.
- Pawson, R., 2013. *The Science of Evaluation: A Realist Manifesto*. SAGE, London, England.
- Pullin, D., Cumming, J., 2022. Teacher performance assessment and considerations of potential legal challenges. Chapter 10. In: Wyatt-Smith, C., Adie, L., Haynes, M., Day, C. (Eds.), *Professionalizing Teacher Education: Performance Assessment, Standards, Moderation, and Evidence*. Routledge.
- Quenemoen, R., Thurlow, M., Moen, R., Thompson, S., Morse, A.B., 2003. *Progress Monitoring in an Inclusive Standards-Based Assessment and Accountability System (Synthesis Report 53)*. University of Minnesota, National Center on Educational Outcomes, Minneapolis, Minnesota. <http://education.umn.edu/NCEO/OnlinePubs/Synthesis53.html>.

- Ragosta, M., Wendler, C., 1992. Eligibility Issues and Comparable Time Limits for Disabled and Nondisabled SAT Examinees. College Entrance Examination Board Report No. 92-95. (ED 349 337), New York, NY.
- Rasooli, A., Razmjoei, M., Cumming, J., Dickson, E., Webster, A., 2021. Conceptualising a fairness framework for assessment adjusted practices for students with disability: an empirical study. *Assess. Educ. Princ. Pol. Pract.* 28 (3), 301–321. <https://doi.org/10.1080/0969594X.2021.1932736>.
- Rioux, M., 2014. Disability rights in education. In: Florian, L. (Ed.), *SAGE Handbook of Special Education*. SAGE, London, England, pp. 131–147.
- Sabbagh, C., Schmitt, M., 2016. Past, present, and future of social justice theory and research. In: Sabbagh, C., Schmitt, M. (Eds.), *Handbook of Social Justice Theory and Research*. Springer, New York, NY, pp. 1–11.
- Sireci, S.G., Scarpato, S.E., Li, S., 2005. Test accommodations for students with disabilities: an analysis of the interaction hypothesis. *Rev. Educ. Res.* 75 (4), 457–490. <https://doi.org/10.3102/00346543075004457>.
- Thomson, S., De Bortoli, L., Underwood, C., Schmid, M., 2019. PISA 2018. Reporting Australia's Results. Volume I. Student Performance. ACER, Melbourne, Victoria.
- Thomson, S., Wernert, N., Rodrigues, S., O'Grady, E., 2020. TIMSS 2019 Australia. Volume 1. Student Performance. ACER, Melbourne, Victoria.
- Tomlinson, C.A., Moon, T.R., 2013. *Assessment and Student Success in a Differentiated Classroom*. ASCD.
- United Nations Convention on Rights of Persons with Disabilities, 2008. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>.
- United Nations Convention on the Rights of the Child, 1989. <https://www.ohchr.org/en/professionalinterest/pages/crc.aspx>.
- UNESCO, 1994. The Salamanca Statement and Framework for Action on Special Needs Education. <https://www.european-agency.org/sites/default/files/salamanca-statement-and-framework.pdf>.
- van der Kleij, F.M., Cumming, J.J., Looney, A., 2018. Policy expectations and support for teacher formative assessment in Australian education reform. *Assess. Educ. Princ. Pol. Pract.* 25 (6), 620–637. <https://doi.org/10.1080/0969594X.2017.1374924>.
- van Swet, J., Wichers-Bots, J., Brown, K., 2011. Solution-focused assessment: rethinking labels to support inclusive education. *Int. J. Incl. Educ.* 15 (9), 909–923. <https://doi.org/10.1080/13603110903456615>.
- Woodgate, R.L., Gonzalez, M., Demczuk, L., Snow, W.M., Barriage, S., Kirk, S., 2020. How do peers promote social inclusion of children with disabilities? A mixed-methods systematic review. *Disabil. Rehabil.* 42 (18), 2553–2579. <https://doi.org/10.1080/09638288.2018.1561955>.
- World Health Organization, 2002. International Classification of Functioning, Disability and Health: ICF. WHO. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>.
- Wu, M., (n.d.). Interpreting NAPLAN Results for the Layperson. https://numeracy4schools.files.wordpress.com/2015/03/naplan_for_lay_person.pdf.
- Zuriff, G., 2000. Extra examination time for students with learning disabilities: an examination of the maximum potential thesis. *Appl. Meas. Educ.* 13 (1), 99–117. https://doi.org/10.1207/s15324818ame1301_5.

Educational accountability, international large scale assessments and PISA: powerful influences based on weak foundations

Sue Grey^a and Paul Morris^b, ^aGraduate School of Education, University of Exeter, Exeter, United Kingdom; and ^bProfessor of Comparative Education IOE, Department of Education Practice and Society, University College London, London, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

International large scale student assessments	181
Logic and assumptions	185
Where next?	189
References	190

“Accountability” is one of many polysemous concepts which have become increasingly employed in academia and policymaking in education. Its core meaning refers to someone being held to account for their actions or outcomes by a third party and implies that the latter possesses a form of legitimate authority over the former. School principals have the authority to hold teachers accountable for their actions and outcomes; professional bodies are authorised to hold teachers accountable for their actions; and inspection agencies are charged with the task of holding schools accountable for their pupils’ performance and the quality of schooling. The core purpose of accountability based on outcomes, as opposed to process, is the assumption that improvements can be elicited by measuring. Högberg and Lindgren (2021) describe the three steps by which this is realized: the production of standardized performance data which serves as a proxy measure of the quality of education; its evaluation by external parties; and incentives and or sanctions resulting from the evaluation.

An extensive literature has described the increased role of such systems of outcomes-based accountability and portrayed them as central to the “New public management” (NPM) reforms associated with the rise of neoliberalism which promoted the practices of the private sector into the education sector (Hogan et al., 2021). This encouraged a climate of “performativity” which resulted in individual practitioners organizing themselves to constantly respond to targets, indicators, and evaluators. This shift has been manifested in policies designed to increase the autonomy of schools and from states governing schools directly to forms of more indirect governance based on “steering at a distance”. Modes of outcomes-based accountability exercised through national agencies like the schools inspectorate OFSTED in England, have become the primary means by which states operationalize accountability, especially within contexts where governance reforms have sought to decrease direct state control, increase school autonomy and parental choice (Asadolahi et al., 2022).

The rise of such systems of accountability has developed in tandem with the shift from envisioning schools as providing pupils with forms of knowledge based on different school subjects to expecting that they equip them with useful skills or competencies (currently described as “21st Century Skills”). Originating in the US economic policy reports “A Nation at Risk” (National Commission on Excellence in Education, 1983, online) and “What Work Requires of Schools” (Secretary’s Commission, 1991), the discourse around the necessity of skills like collaboration and critical thinking has increasingly dominated global education initiatives, and is now reflected in local education policies across a range of countries. Anderson-Levitt and Gardinier (2021) note that UNESCO’s original interest in such skills was to enable individuals to be better citizens in increasingly multicultural democratic societies (Delors, 1996); and the OECD extended that rationale by stressing the need to prepare citizens who can address environmental emergencies facing the globe (OECD, 2019). However, those humanistic aims have been largely replaced by an economic rationale which claims that, in line with human capital theory, in the new “global knowledge economy”, citizens must be prepared in schools to exercise interdisciplinary skills, or “competencies” – that are increasingly demanded of workers and professionals for the sake of their own individual success as well as the prosperity of national economies (Robeyns, 2006).

Our focus is on Outcomes-Based accountability that derives from cross national studies of pupil performance, or international large-scale assessments (ILSAs), which have emerged in the 21st Century as a powerful influence on education policies in many nations.

The chapter begins with a brief analysis of the history of ILSAs and their key features, focusing on their sources of legitimacy and the conflicts inherent in their modus operandi. It then identifies the logic and central assumptions which underlie ILSAs as a form of accountability. Finally, the chapter provides an overview of the major critiques of this form of accountability, focusing on PISA.

International large scale student assessments

The first ILSA was undertaken by UNESCO’s Institute for Education in 1962. The Pilot Twelve Country study involved 12 000 13-year-olds and its results were published in 1962 (Forshay et al., 1962). From this the International Association for the Evaluation of Educational Achievement (IEA) emerged which then conducted a series of international evaluations. The IEA now conducts the 4-yearly Trends in Mathematics and Science Study (TIMSS) and the 5-yearly Progress in Reading Literacy Study (PIRLS). The IEA

studies are not wholly standardized as they test pupils on the curriculum they have studied, which clearly varies across nations; and advice arising from them is tailored to individual countries in separate reports.

The latest generation of cross-national tests has been developed by the OECD as part of the Program for International Student Assessment (PISA). PISA assessments have been conducted every 3 years since 2000, in reading, mathematics, and science. Each cycle also includes an “innovative domain” – these have included problem solving, global competences and critical thinking. PISA is a standardized test designed to measure the skills required by 15-year-old pupils entering the workforce. In 2000, 43 countries participated and by 2018 that had increased to 79 participants, including what the OECD calls “jurisdictions” such as Shanghai. The triennial release of the PISA League Table rankings of the performance of pupils across nations attracts a great deal of media scrutiny and has generated reactions varying from national pride, (for example, in Singapore), shock and shame (Germany, Denmark and England) to ambivalence (USA, France) (Waldow et al., 2014; Engel and Frizzell, 2015)

The OECD has worked in conjunction with the media (Grey and Morris, 2018) and with a powerful global network of actors to make PISA the most prominent and influential ILSA and contributed to the emergence of a global system of educational accountability and governance. That network comprises other cross-national organisations, such as the World Bank and UNESCO, and companies such as McKinsey and Pearson which have positioned “competitive comparisons” based on ILSAs as the basis for measuring, evaluating and reforming national educational systems. ILSAs, by virtue of comparing pupils across nations and cultures, are far easier to operationalize if they are assessing and comparing based on a universal and standardized measure (defined as competencies) rather than one which, like those of the IEA, reflects the curricula in different nations. As will be demonstrated, while their legitimacy and mode of operation differ from national systems of outcomes accountability, their influence on many of those systems has been powerful; although it has been far greater in some nations than others (Pultar, 2021).

AS PISA has grown, so the OECD has also extended its testing empire to include more populations and concepts. Fig. 1 shows the range of tests it has developed.

These now include tests for pre-school children (IELS); university students (AHELO); and adults (PIAAC). PISA has primarily involved more affluent nations, having originally been offered only to OECD member states. Few low- and middle-income nations have joined and some which did take part, performed poorly and subsequently withdrew. India, for example, withdrew, claiming that the test was not sufficiently suited to the Indian context after it was outperformed by all other nations except for Kazakhstan (Addey, 2016). In an attempt to attract low and middle income nations, the OECD ran a pilot project, termed PISA for Development

Programme/Start date	Who is assessed?	Subject Areas	Sources of context information	Frequency	Coverage
OECD PISA (2000) Programme of International Student Assessment	15 year olds in FT education.	- Reading - Mathematics - Science - Collaborative problem-solving (2015) - Global Competency(2018)	- Students - Parents (optional) - Teachers (optional) - School principals	Every three years since 2000	OECD countries: 34 Non-OECD participants: 38 (PISA 2015)
OECD PIAAC (2013) Programme for the International assessment of Adult Competencies.	16-65 year olds	- Literacy - Numeracy - Reading components - Problem-solving in technology-rich environments	- The individuals who are assessed	Every 3 years	OECD countries:24 Non-OECD participants: 2 (PIAAC 2011)
OECD TALIS (2008) Teaching and Learning International Survey	Teachers of lower secondary education	(Focuses on the learning environment and working conditions of teachers)	- Teachers - School principals	Five years between first two cycles Since 2008	OECD countries: 16 Non-OECD participants: 7 (TALIS 2008)
OECD AHELO (2012) Assessment of Higher education learning Outcomes	University students at the end of B.A. programme	- Generic skills common to all university students (such as critical thinking) - Skills specific to economics and engineering	- Students - Faculties - Institutions	A feasibility study was carried out in 2012	Institutions from 17 countries participated in the feasibility study in 2012
PISA for Schools (2013)	Cohort of 15 year olds in a school – at least 42 students and 80% participation	Same assessment framework as PISA	As for PISA	Rolling programme, ‘on demand’, maximum of once per school year	Currently available in 11 countries and 10 languages
OECD PISA-D (2018) PISA for Development	15 year olds in school and some ‘Out of school’ 14-16 year olds.	- Mathematics - Reading - Science (in school only)	- Students - Teachers - School principals	Results of a pilot round in 9 countries to be released in Dec 2018, merging with PISA in 2021	Low- and middle- income countries (9 pilot countries)
OECD IELS (2018) The International Early Learning and Child Well-being Study	4/5 year olds.	- Emergent literacy - Emergent numeracy - Empathy and trust - Self-regulation	- Parents - Staff	2017 field trial/Pilot 2018 Main Study	Pilot Study in USA, England & Estonia

Fig. 1 OECD international comparative education tests (Taken from Grey, 2020).

(PISA-D) with modified questions and metrics tailored toward the educational realities in the global south; which was completed and declared a success by the OECD in 2018, notwithstanding some critical analyses (Auld et al., 2020; Li and Auld, 2020).

In contrast to national systems of inspection and evaluation designed to hold individual teachers or schools to account and/or to provide comparative data to promote school competition and school choice, the OECD, as an organization concerned with economic coordination and the promotion of market economies, has no official mandate in the field of education. The focus of PISA is on educational systems, and it has no power to provide rewards or impose sanctions. Thus, in contrast to national systems of outcome-based accountability, the legitimacy of the external agent, the entity being held to account and the potential incentives/sanctions of cross-national systems are opaque and loosely coupled. As Meyer comments:

the din about changes in rankings is masking the even more profound problem that in the wake of 15 years of PISA assessments, democratically elected national authorities are increasingly ceding control over their schools to an unelected, unaccountable body of OECD experts.

Meyer (2013).

These “experts” are, significantly, not “experts” in education, nor even in policy, but are largely economists and statisticians – the Director of Education himself has a background in theoretical physics. Notwithstanding, PISA has emerged as the largest and most influential cross-national test that has been instrumental in initiating education reforms across nations and specifically in promoting systems of outcomes-based accountability within many nations (Verger et al., 2021).

The OECD Director of Education, Andreas Schleicher, who pioneered PISA after a brief career at the IEA, describes PISA as “the world’s premier yardstick of educational quality” and explains its value thus:

Comparative international evaluations can extend and enrich the national picture by providing a larger context within which to interpret national performance. They have gained prominence – over recent years – since the benchmarks for public policy in education are no longer solely national goals or standards, but, increasingly, the performance of the most successful education systems internationally. International evaluations can provide countries with information that allows them to identify areas of relative strengths and weaknesses, and monitor the pace of progress of their educational system. They can also stimulate countries to raise aspirations by showing what is possible in education, in terms of the quality, equity, and efficiency of educational services provided in other countries or regions, and they can foster better understanding of how different education systems address similar policy challenges

Schleicher (2010).

ILSAs are thus portrayed as an accountability tool that can assess policies and policymakers against “benchmarks” which compare the performance of nations with the most successful and provide them the opportunity to learn from “world class” policies and “best practices”. Unlike more established and local forms of accountability, which in themselves are contested and ambiguous (Christensen and Lægriid, 2015) this new form of accountability to globalized norms based on numbers and datasets lacks a direct hierarchy in which an “agent” (a manager, or an organization) offers feedback to an actor (a teacher, for example), who is then expected to demonstrate “improvement”. Personalized or tailored targets which previously would have served to focus performance on a set of outcomes are now replaced by fuzzy exhortations to match the world’s best, and to strive for continued improvement against a set of shifting criteria defined by scores and rankings on PISA.

Thus, schools are no longer only accountable to their students or to local education policymakers and evaluations. Since around 2012 the OECD has made explicit the underlying rationale for the instrument: that improving PISA scores contributes directly to economic growth, and that the future prosperity of a nation is also at stake. The logic is that PISA will not only help nations to improve the intrinsic quality of their schooling but that those improvements will increase economic growth and their capacity to compete in the global knowledge economy. As the OECD explains:

For the USA, an increase in GDP of 25 PISA score points will increase GDP by \$41 trillion (OECD, 2012, p. 2)

And:

If Ghana could educate all of its students to at least the basic level of reading and mathematics skills [on PISA], it would see a gain over the lifetime of children born today that, in present value terms, is 38 times its current GDP (Schleicher, 2018, p. 139).

The OECD, in contrast to the IEA, has become heavily involved in policy advocacy and thus acts as not only the collector of performance data but also interprets it to promote the “best practices” inferred from features of high performing nations. This dual role allows the OECD to design, collate and disseminate data which it then also translates into carefully framed policy messages, aimed at entire school systems but largely unamenable to school-level intervention. A form of “Soundbite governance”, resting on slogans such as “the quality of an education system is only as good as its teachers”; “Money makes a difference ... but only up to a point”, and the all-important mantra, “Your school system today is your economy tomorrow” (OECD presentations, various

sources), is backed up with literally hundreds of documents, videos and webpages, which are aimed at an ill-defined audience, but are generally consumed by media actors and transferred to media audiences in edited and diluted format.

The scientific expertise necessary for test design, data collection and reporting has thus become entangled with the construction and narration of policy stories (Stone, 1989) which are passed into the public arena via the media. This narrative revolves around identifying what is wrong (usually framed as a crisis), the ideal promised future and the proposed solution. The solutions are derived from the myriad of possible policy stories that can be constructed from the data and the problems are framed in a way which leads inevitably toward the solution identified (Auld and Morris, 2021).

Two fundamental problems arise from the OECD performing both those roles: firstly, a tendency for the organization to collect, interpret and highlight data in ways which confirm the policies and “best practices” which they advocate should be adopted by nations; and secondly, an unwillingness to acknowledge any weaknesses in the test design, data analysis or interpretation given that they have been directly used to promote and initiate policy actions. The OECD has addressed these conflicts of interest by a strategy of either not engaging with critics, or by resorting to *ad hominem* responses. “Who is opposing this?” Schleicher asked in 2015. He answered his own question; “people from pedagogy, not really anyone else” (Schleicher, 2015). Criticisms are skilfully deflected by a man who is highly experienced in dealing with both the media and academia, as the example below demonstrates. In his drive to ensure that the world’s school systems are accountable to data, Schleicher himself remains essentially unaccountable:

I think the methodological robustness is beyond question. I don’t worry about it, when I get those questions, I provide the answers and I think very soon this dies down ... the bigger question is, are our metrics complete ... ?

Schleicher (2015).

PISA’s legitimacy as a powerful source of influence on educational policy making derives from a multitude of interconnected factors. As a global metric presented as a numerical league table of comparative performance it bears the hallmarks of an objective test undertaken by scientific experts unsullied by ideological considerations. It provides a ranking of performance which is shaped by and tailored for media stories (Grey and Morris, 2018) of national pride or shame; crisis, failure and blame. It provides national politicians with a wide range of correlations from which to select and develop an “evidence based” policy story, which supports their own ideological preferences and related policy stories (Auld and Morris, 2021), in which the reasons for poor performance and the path to salvation are identified. This path requires emulating the “best practices” and policies of high performing nations or “world class” systems.

In domestic political arenas the global nature and source of the data can also help overcome political resistance and encourage coalition building. PISA is also constantly validated and legitimized by the Global Education network and industry (Mohammed and Morris, 2019), which operates a business model reliant on interpreting PISA data to identify “what works” and sell solutions. The OECD also provides a convincing strategic narrative which includes portraying those who do not follow its advice as reactionaries or ideologues who are not committed to improving the quality of education; claiming that PISA offers a path to economic growth; and portraying a vision of an idealized future for those who follow its advice.

The impact of PISA on national education policies emerges from the ways in which evaluation by both the OECD and external parties of a nation’s performance on PISA is undertaken in a public arena by a combination of stakeholders, almost all of whom lack a democratic mandate. The media, domestic politicians, commercial consultancies, global agencies like the World Bank, as well as the OECD itself, all act on and shape the perceptions of PISA in local political spaces, and all benefit, directly or indirectly, from its ongoing and dynamic presence. PISA has thus contributed to a continuing shift to what Gleckman (2018) describes as “multi-stakeholder global governance”; the key feature of which is that control, in this case of the system of accountability, shifts from nation states to those who provide or benefit from the activity. Agency is often obscured, and the rhetoric of competition and the imperative of economic growth become “common sense” objectives, against a backdrop of the need to skill future workers to thrive in a globalized arena. This narrative presents some difficulties for the OECD; for example, with the passage of time it has become clear that there is a gap between the promissory vision and reality. PISA is marketed on the basis that its comparisons provide the basis for improving pupils’ learning in participating nations (and that this will result in economic growth). However, despite many nations apparently following the best practices promoted by the OECD, that improvement has not generally materialized. As the OECD explained in 2019: “it is disappointing that most OECD countries saw virtually no improvement in their students’ performance since PISA was first conducted in 2000”. (OECD, 2019, p. 3).

The same outcome is evident outside the OECD. The nations of the Arabian Gulf have relied heavily on the policy advice derived from ILSAs and have, in conjunction with their network of consultants and private outside companies identified improving their scores on ILSA’s as the goal of their education reforms (Mohamed and Morris, 2019). However, a review of the reforms implemented on the basis of the recommendations arising from the involvement of these agencies concluded:

Over the past decade and a half, key countries of the Gulf... have invested considerable resources in education. Driven by a desire to better prepare their economies and societies for an increasingly globalised and competitive world.... And yet despite the availability of ample financial resources and expert policy and management advice, most independent reviewers have concluded that the actual results have fallen short of initial expectations.

Alfadala (2015, p. 4).

The reason given for this was:

Most, if not all, of these outcomes were found to stem from a single, underlying response to the reforms being introduced: resistance to the change. Alfadala (2015, p. 7).

While allocating the blame for failure to local teachers and decision makers is a standard strategy employed in the “gray literature”, most particularly by the media (Grey, 2020), it conveniently avoids any consideration of the policy making process and its reliance on ILSAs.

Much of the academic literature on PISA has involved secondary analyses of the data designed to identify trends, patterns and correlations (Hopfenbeck et al., 2018). Here we instead focus on how PISA is used to influence policy by undertaking a critical interrogation of the logic and assumptions on which it is constructed. We argue that a comprehensive understanding of why ILSAs have failed to deliver requires the recognition that the PISA enterprise is built on weak foundations. However, by locating blame for failure within local conditions and actors, agency is removed from the OECD or its recommendations, which, ironically, are made without reference to those same local conditions. The quest for world-classness is culture-blind; not so the post-mortems when policy interventions are considered to have failed to deliver the hoped-for results. “They didn’t change the culture, they changed what they did in education”, said Andreas Schleicher in 2015 of Poland’s improved scores on PISA (Schleicher, 2015, online), illustrating that while failures can be blamed on local conditions, successes are seen to validate the mission and support the strategies promoted by PISA.

Logic and assumptions

The logic underpinning the policy advocacy which is central to PISA’s influence derives from asserting a strong connection between economic performance, competition/survival in the global/knowledge economy and the quality of schooling/education, as measured by cross national tests. The basic logic can be stated thus (Morris, 2015):

- a. Country A is growing rapidly and is competing successfully in the global knowledge economy, which is largely the result of the educational system producing workers with the required skills as evidenced by international tests.
- b. Country B has the opposite features to Country A.
- c. If Country B adopts the critical features of the educational system of Country A, it will improve pupil performance, the state of the economy and its capacity to compete in the global knowledge economy.

This logic, while presented in a far more subtle and convincing manner which constantly recognizes and then ignores its reductive nature and limitations (Auld and Morris, 2014), provides policy-makers with a clear logic for reform. It is premised on the following assumptions:

- (1) economic success in a nation state is primarily a function of schooling within that nation and the development of human capital. Tests of pupil achievement, such as PISA, provide a reliable measure of a nation’s stock of human capital and therefore of future economic growth and competitiveness.
- (2) The tests are valid and reliable.
- (3) The causes of pupil performance lie primarily within the educational system.
- (4) The cause(s) of high performance can be isolated and policies based on those causes can be transplanted into a different context.
- (5) The evidence is faithfully and consistently considered and the policies advocated are coherent insofar as they are validly derived from that evidence.
- (6) The policy actions taken to improve test scores also serve to improve the quality of education.

These assumptions provide the foundations upon which this form of accountability is constructed and how that is translated into policy advocacy. Below we highlight the key issues associated with each of them.

The focus of the *first assumption*, that economic success is primarily the function of a nation’s quality of schooling is oft repeated and central to the discourse of global governance. It is not supported by the evidence other than that provided by the work of Erick Hanushek, on whom the OECD rely (with Kimko [Hanushek and Kimko, 2000]; then later with Ludger Woessman [Hanushek and Woessmann, 2010, 2015]). A study by Kamens (2015) who traced the historical relationship between test scores and economic growth from 1963, concluded that “the linkage between national test scores and economic growth is weak and inconsistent” (Kamens, 2015, p. 436) and that “countries with higher test scores ... are not more likely to experience later higher economic growth. (Kamens, 2015, p. 441). Komatsu and Rapple (2017) scrutinized the claims of Hanushek and Woessman, addressing in particular the ease with which their methodology was misunderstood and therefore remained largely unchallenged. Their conclusion – that the work upon which PISA is based is both inaccurate and misleading – has important implications for policymaking and governance. In a later article (Komatsu and Rapple, 2017a), the authors go further, suggesting that explicit linkage of economic objectives with student performance actually serves to reduce student performance, when promised economic gains fail to materialize, thus “undermin[ing] the core premise of the OECD’s whole approach to PISA itself” (Komatsu and Rapple, 2017a, p. 269).

Feniger and Atia (2018) also examined the relationship between PISA scores and economic growth and concluded that success in the former increased after the latter. This is not surprising as Hanushek and Woessman's analysis does not build in a time lag between educational outcomes and economic performance; and it ignores the influence of other factors (such as economic policies, capital investment and natural resources). Further, as Lauder (2015) argues, it does not recognize the global nature of modern labor markets. The relative influence of a range of factors is underlined by the evidence on labor productivity. If increasing PISA scores resulted in faster economic growth, this would be manifested in higher levels of labor productivity. Many nations with high levels of productivity, such as the USA, Italy and France, do not perform highly on PISA; and many high performers on PISA, such as Hong Kong and Singapore, have low levels of labor productivity (OECD, 2021). A similar pattern is seen in measures of innovation and inventiveness (Global Innovation Index, 2021).

The *second assumption* relates to the validity of the tests and the key question of whether what they measure is important.

It is not clear that the cognitive competencies measured by PISA are those that employers seek in their employees; PISA measures skills the OECD believes to be essential in the workforce and assumes that it is primarily success in STEM subjects that will generate economic growth. When surveyed, employers routinely identify teamwork, collaboration and oral communication skills as among the most valuable yet hard-to-find qualities in workers (e.g. Casner-Lotto and Barrington, 2006; Jerald, 2009). Deming (2015) also argues that social skills are a more powerful predictor of workers' productivity than cognitive skills, suggesting that advances in technology have primarily reduced the number of jobs that require cognitive skills rather than those which require social skills. Consequently, it is social skills that are increasingly valued and rewarded in the labor market.

PISA compares pupil performance in science, mathematics and reading, so inevitably, the tests focus on cognitive skills which are amenable to measurement. Competencies which are less easily quantified, such as critical thinking, cooperation and creativity are now assessed through proxies constructed to make them amenable to measurement; while competencies associated with other domains such as Art, Music, PE, Morality and Ethics are excluded. In essence, ILSAs measure that which is most amenable to measurement; and Biesta (2009) argues we have ended up valuing what we measure rather than measuring what we value. The potential reductive effect on the school curriculum is clear, and while education systems continue to be governed at a distance by the metrics of PISA, this narrowing remains one of the most significant dangers.

Acknowledging weaknesses around PISA, in 2018, Schleicher stated:

The current PISA doesn't include many elements that are important. We need to do better on social skills, on creative skills, so the new framework – the 2030 framework is really about developing new ideas for what PISA should assess. Education 2030 is about what should PISA assess, and what are the skills that are really important for success.

Cited in Li and Auld (2020).

We return to discuss OECD's vision of the future in the final section.

The *third assumption* relates to the longstanding debate as to whether student outcomes are primarily a result of the impact of schooling or of factors outside school (e.g. culture, parental expectations, private tutoring). If the levels of pupil achievement are significantly influenced by factors outside the school then the inferences that can be derived for educational policies are severely limited. While those advocating educational policies often acknowledge the influence of non-school factors, they tend to subsequently ignore their significance as they promote their own preferred education policy solutions.

The performance of populations of pupils from one culture schooled in another provides a basis for considering these claims. In all cross-national studies of academic achievement, Chinese pupils (from Hong Kong, Singapore, Taiwan and China) have been ranked among the top performers and these societies have been portrayed in England as education utopias. Studies by Feniger and Lefstein (2014) and Jerim (2015) examined the performance of Chinese students who have been born and schooled in Australia and New Zealand, respectively. They found that Chinese immigrant students performed at a significantly higher level than pupils who were native to those countries. This would suggest that influences outside the school have a powerful impact on pupil performance.

The *fourth assumption* is that the cause(s) of high performance can be isolated, and policies based on those causes can be transplanted into a different context. Unless a causal relationship can be established between the factor identified as promoting a high level of pupil performance, and improvement in that performance, any claim to have identified "what works" lacks a strong foundation. Yet attempts to promote policy interventions are often based on loose correlations, or confirmatory and selective observations. The claimed sources of high performance are often contradictory or pre-selected and confirmed or cited as a feature of a high performing system; but, importantly, there is no corresponding consideration as to whether poorly performing systems lack this factor (Auld and Morris, 2014). Thus, high performance on PISA has variously been associated with both traditional and progressive teaching methods; centralized and decentralized systems of control and with teachers and leaders. Claims regarding the features of selected high performing systems are used to deliver prescriptions for improving matters elsewhere. Further, many of the assertions, for example, that good teachers lead to high levels of pupil performance, are self-evident analytical claims (i.e. they are necessarily connected – good teaching is defined by its impact on pupil outcomes) and not empirical findings. Similarly, claims that good educational systems are characterized by "good/strong leadership" are axiomatic and banal. The network that uses PISA data to identify and advocate reforms has developed a sophisticated range of strategies to mask the often-weak basis of such claims about "what works" (Auld and Morris, 2016).

The proposition that "best practices" can be transplanted from one context to another assumes that the goals of education systems, the problems they face, and the desired solutions are universal. It ignores the desires and interests of communities and what and how they decide what their children should learn. What is "best" is, therefore, as Samoff (2021) notes:

.... necessarily local. Of course we can learn by observing what others do. But there are not and cannot be universal best practices, detached from their context.

Samoff (2021).

Further, a substantial body of evidence demonstrates that educational reforms and practices do not remain constant when transplanted to a new context. Transfer is followed by a process of local translation and as Cowen (2009) comments: “when it moves it morphs”. Accordingly, “best practices” are modified, recontextualized and often fail to achieve their stated goals. This was evident in the various attempts by the UK Government to introduce Mathematics curricula and pedagogy from Singapore and Shanghai into English schools; a move premised on the belief that this would drive English pupils to emulate the high levels of performance on PISA of pupils in those areas. Chinese curricula and maths teachers were brought to the UK; English teachers modified the curriculum to suit their own context and an independent evaluation concluded that the imported “best practice” resulted in no improvement in pupils learning outcomes (DfE, 2019).

The fifth assumption relates to the faithful use of PISA data by governments and is a variant of the second assumption on the faithful use of data by the OECD. This issue is especially pertinent in England, where policymakers have used pupil performance data extensively to justify reform in ways that are often highly selective or which distort the evidence. For example, in terms of selectivity, there has been a marked tendency for policymakers, the media and the network of policy advocates to focus on reports that are negative and serve to portray schooling in England as an educational dystopia in need of urgent and radical reform (Morris, 2012).

Even when England’s results show a slight improvement in its ranking or no significant change in the level of pupil performance, this has been characterized in the media and policy discourse as a decline in standards or, at best, stagnation. In contrast, there are reports that indicate relatively good performance in England, such as the 2007 TIMSS Report and the OECD Report (2013a).

Problem Solving in PISA 2012, receive comparatively little attention. The response in England to the 2013 PIACC Report (OECD, 2013b, online) was also illustrative of a tendency to interpret data selectively. England scored poorly on the tests of the literacy and numeracy skills of adults, and the OECD associated this with the impact of poverty and the low-skilled nature of much work in England. However, in the subsequent public discussion of the results by politicians and the media in England these factors were ignored and a range of other factors were proposed as the causes of the low performance. These included low curriculum standards, low standards of teaching, a dumbed-down curriculum and the absence of performance-related pay for teachers. The linkage in England between the specific policies promoted and the comparative evidence from which they are derived has been, at best, tenuous. For example, the reforms promoted in the 2010 White Paper (Morris, 2012) were primarily justified on the grounds that they were derived from the best practices of high-performing (i.e. primarily East Asian) systems. However, many of the policy initiatives – the establishment of Free Schools, multiple and shortened routes into teacher education, the recruitment of untrained teachers, and a national curriculum that was optional for academies and Free Schools – are not characteristics of those school systems in East Asia that topped the PISA league tables. Similarly, neither the PISA data, as interpreted by the OECD, nor the prevailing systems of school governance prevailing in East Asia provide a rationale to justify the form of school autonomy promoted in England since 2010 (You and Morris, 2015). Further, the direction of educational reform in many East Asian states is toward reducing the prevalence of the specific features (direct instruction, exam orientation and a focus on academic achievement) that English politicians admire and seek to emulate (Forestier and Crossley, 2014).

The final assumption, that the policy actions designed to improve PISA scores result in improvements to the quality of education, is fundamental, as the ultimate purpose of PISA is to improve the quality of schooling pupils receive and subsequently generate economic growth. We suggested above that by the OECD’s own analysis there is no evidence to indicate an improvement in educational outcomes, as measured by PISA scores. Instead, policy actions are justified by retrospective reference to what successful countries have done: these are then used to advocate the adoption of such interventions by less successful nations. Much of this “evidence” is often little more than axiomatic commentary, as is demonstrated by the following extracts from the.

PISA 2018 results report:

In 60 countries and economies (or four in five countries/economies), students who reported that the longest piece of text they had to read for school was 101 pages or more scored significantly higher on reading performance than those who reported 10 pages or less after accounting for students and schools’ socio-economic profile and students’ gender.

OECD (2021, online, p. 127).

Education systems with a higher proportion of students who were taught how to detect biased information in school were more likely to distinguish fact from opinion in the PISA reading assessment, even after accounting for country per capita GDP (OECD, 2021, online, p. 139).

In much the same way as the absence of particular factors is not noted in unsuccessful countries (see above), so by the time the following cycle is reported, three years have elapsed, and a robust follow-up to determine whether policy recommendations have resulted in the promised improvements is not undertaken. Instead, a new edition of glossy brochures, colourful charts and narrative commentaries is issued, glorifying the new successes and noting what it is that they have done to demonstrate improvement.

Emler et al. (2019) have identified a number of side effects or unintended consequences of PISA on school systems which serve to undermine their quality. These primarily arise from the introduction of policies designed to improve scores which result in a narrowing

of the curriculum and a focus on preparing for the test. Examples include: modify the curriculum content to better align with the test; increase the time allocated to areas tested in PISA by reducing the time allocated to other areas; train teachers to coach pupils how to pass the test; align existing systems of inspection and accountability to the test; and, introduce performance pay which links teachers' pay to improving pupils results in PISA subjects.

These actions illustrate [Goodhart's Law \(1975\)](#) which states: "when a measure becomes a target it ceases to be an effective measure" as the actions taken to achieve the target are designed to game the system and have no clear connection to achieving the ultimate goal, in this case improving the quality of education. Moreover, the "ultimate goal" is not evaluated in terms of the actions taken in response to earlier recommendations, so that the target is constantly moving.

The unintended negative consequences of such actions in education are evident in England, where successive governments have sought to use PISA data to further an already frenzied environment of serial educational assessment and reform. This "policy hysteria", characterized by short reform cycles and frequent policy switches ([Stronach and Morris, 1994](#)) took a "data turn" in 1997 when the then (Labour) education secretary David Blunkett declared,

The publication of performance data benefits parents and acts as a spur to improve performance. We will publish more such data than ever before.
[DfEE \(1997, p. 25\).](#)

Subsequent Coalition and Conservative administrations have enthusiastically followed in Blunkett's footsteps, and PISA has provided a rich seam of performance data which has been mined by Education ministers since-with one, Michael Gove, evangelizing on behalf of Andreas Schleicher during a visit by the latter to the UK:

[he is] a man I recently have described as the most important man in the British education system - but he could equally be the most important man in world education. ... in truth Andreas is the father of more revolutions than any German since Karl Marx.
[Gove \(2011, online\).](#)

While his junior minister, Liz Truss, faithfully repeated the economic messages underlying PISA:

The link between student scores in international tests and real GDP growth rates per head is growing stronger; doubling between 1960 to 1980 and 1980 to 2000.
[Truss \(2012, online\).](#)

Thus, there has been a neat alignment between the assessment ambitions of UK governments and the focus of PISA on measurable skills, especially in reading, mathematics and science, which has been reflected in a narrowing of the English curriculum to sideline Arts and creative subjects in favor of these same subject areas ([Johnes, 2017](#)).

Such actions may help to improve PISA scores but they undermine the requirement that schools provide a "broad and balanced curriculum" and it is unclear how they improve the quality of schooling. Whether or not it improves a country's ranking in the PISA League Tables is also less certain if the fallacy of composition operates and other nations adopt the same strategies.

More importantly, the focus on PISA scores results more generally in education policies being narrowly framed as solely designed to find solutions which will help improve those scores. The result is a form of policy tinkering designed to prioritize and resource activities directly related to influencing the target outcome. The focus on performativity ascribed to practitioners is mirrored by policy makers as they organize themselves around responding to targets, indicators, and evaluators. This has resulted in a focus on schools, curricula, teachers and leadership, while more complex and systemic issues of poverty, health, social deprivation which affect pupils' education, as well as more generic threats to humanity such as climate change ([Rapplee and Komatsu, 2020](#)) are ignored or marginalised.

A pattern has emerged: the OECD promotes policies based on claims which are "loosely coupled" to their own data; governments adopt those OECD recommendations and promote them to domestic audiences as "evidence based policy" and global "best practices"; and national governments develop their own policy stories which are tenuously linked to the evidence provided by the OECD. The outcome is often not the faithful use of evidence but rather that PISA emerges as an unlimited source of legitimizing data that can be used for asserting a multitude of potential policy stories. The OECD produces a vast quantity of data and an ongoing steady flow of documents, videos, social media posts and web-pages based on PISA. The sheer volume of material, and the regularity of updates - for example, the monthly "PISA in Focus" bulletins, billed as "a series of concise monthly education policy-oriented notes designed to describe a PISA topic" (OECD, online) - make it easy for policymakers to find global sources to support almost any policy initiative, which can then be described as "evidence-based" or be justified by reference to "world class" systems. Notwithstanding that PISA scores have not improved after two decades of adopting "best practices" the triennial cycle repeats itself as new crises and priorities are identified.

Where next?

We have analyzed the ways in which PISA operates as a new form of global accountability, and highlighted the weak foundations on which it is constructed, through an examination of its core assumptions. We identified four systemic issues, namely: a failure by the OECD to deliver on its promise to improve educational quality (as they define it); a conflict of interest arising from its mixing the roles of data collection and policy advocacy; the development by the OECD of a model of policy advocacy which claims agency for improvements but assigns failure to others; and, the use of PISA data by policymakers to deflect attention from deeper, systemic sources of disadvantage such as poverty and wider existential issues like climate change. We now briefly look to the future, and to where the PISA enterprise might next seek to cast its reach.

There are signs that the OECD is shifting its organisational discourse (Li and Auld, 2020). From its inception the OECD's core claim could be summarized as: *PISA accurately measures educational quality; high performance in Maths, Science and Literacy generates economic growth; high performers provide the source of "best practice" and transferrable policies.* More recent pronouncements suggest the OECD is now aligning itself with the UN's Sustainable Development Goals and adopting a more humanitarian discourse which can be summarized as: *we need to plan for an uncertain future (the fourth Industrial Revolution) which requires 21C/transversal/soft skills such as Creativity, Global Competencies & Well-being (not traditional/Legacy skills) and technology. The private sector and AI will play a central role in shaping these skills.*

Acknowledging this "future turn", Schleicher made the following statements:

the industrial approach to schooling, where children were educated in batches of age and all go through the same standardised and prefabricated curriculum and instruction", "hasn't really been conducive to nurturing the kind of creativity that we need in this age of accelerations. Whichever way you construct your list of 21st century skills, creativity will be at the heart of it.

Schleicher (2017).

And:

The kinds of things that are easy to teach are nowadays also easy to digitise and automate. In the age of artificial intelligence (AI) we need to think harder about how to develop *first-class humans*, and how we can pair the AI of computers with the cognitive, social and emotional skills, and values of people.

OECD (2018).

This is indicative of a move from "diagnostic expertise", whereby PISA has identified factors linked to success in high-performing systems, to a type of "prognostic expertise" in global governance, which "render[s] knowable what has not yet happened – the future" (Berten and Kranke, 2021). In this way, accountability is not only entirely decoupled from agency, as in current iterations of PISA, it is also predicated on an uncertain and unimaginable set of hierarchies and imaginaries which may exist in future.

A further set of challenges lies in the move to measuring non-cognitive competencies and including low- and middle-income nations in the PISA testing regime. Extending PISA to include poorer nations exacerbates the longstanding problems arising from using a universal metric across diverse cultures/contexts. As the OECD explained with reference to measuring Global Competencies:

The most salient challenge for the PISA assessment is that — through a single international instrument — it needs to account for the large variety of geographic and cultural contexts represented in participating countries ...

OECD (2018, online)

While developing the original, universal PISA tests of Mathematics, Reading and Science presented significant technical issues, these are now dwarfed by the more fundamental problems which arise from assuming that different cultures and societies have a common understanding of concepts such as "well-being" "creativity" and "Global Competencies" and that these can be measured through a standardized test. Studies have pointed to the ways in which the OECD's conception of a globally competent citizen appears to be a proxy for a globally mobile middle class knowledge worker (Gardinier, 2021). Goren (2020) demonstrates the deeply embedded nature of the challenge arising from the multiple understandings of the concept across cultures and its measurement by the OECD. She found that certain items, deemed sensitive by particular nations, for example on immigration, were required to be deleted from the OECD's Global Competence test; while one nation decided not to administer the questionnaire to a specific religious group (as it was contrary to their beliefs). Notwithstanding, the results and the associated comparative league tables were presented as if all participants had taken the same test, raising questions about its validity.

We have presented the PISA paradox; it is on the one hand the most successful and influential ILSA that is at the core of global educational governance; on the other hand it is constructed on weak foundations and has failed to deliver on its promissory claims

to improve the quality of schooling. The challenges it faces in maintaining its centrality on the global education stage require it not only to be seen to “respond” to a world which it tells us “is changing”; it also needs to be instrumental in helping to engineer that change, in order to capture it with metrics.

References

- Addy, C., 2016. PISA for development and the sacrifice of policy-relevant data. *Educ. Sociedade* 37, 685–706.
- Alfadala, A., 2015. K-12 Reform in the Gulf Cooperation Council (GCC) Countries: Challenges and Policy Recommendations. Qatar Foundation, Doha.
- Anderson-Levitt, K., Gardinier, M.P., 2021. Introduction contextualising global flows of competency-based education: polysemy, hybridity and silences. *Comp. Educ.* 57, 1. <https://doi.org/10.1080/03050068.2020.1852719>.
- Asadolahi, S., Farney, J., Triadafilopoulos, T., White, L.A., 2022. School choice, policy feedback effects, and policy outcomes: understanding the relationship between government policy design and parent decisions to “stay” or “defect” from public education. *Comp. Educ.* 58 (2), 242–259. <https://doi.org/10.1080/03050068.2021.2021129>.
- Auld, E., Li, X., Morris, P., 2020. Piloting PISA for development to success: an analysis of its findings, framework and recommendations. *Compare* 0 (0), 1–25. <https://doi.org/10.1080/03057925.2020.1852914>.
- Auld, E., Morris, P., 2021. A NeverEnding story: tracing the OECD’s evolving narratives within a global development complex. *Glob. Soc. Educ.* 19 (2), 183–197. <https://doi.org/10.1080/14767724.2021.1882959>.
- Auld, E., Morris, P., 2014. Comparative education, the “New Paradigm” and policy borrowing: constructing knowledge for educational reform. *Comp. Educ.* 50 (2), 129–155. <https://doi.org/10.1080/03050068.2013.826497>.
- Auld, E., Morris, P., 2016. PISA, policy and persuasion: translating complex conditions into education “best practice”. *Comp. Educ.* 52 (2), 202–229. <https://doi.org/10.1080/03050068.2016.1143278>.
- Berten, J., Kranke, M., 22-23 September 2021. Knowing the future: prognostic expertise in global governance. In: Conference Presentation at the Metro Conference: Quantifying the World. www.metro-project.eu.
- Biesta, G., 2009. Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educ. Assess. Eval. Account.* 21 (1), 33–46.
- Casner-Lotto, J., Barrington, J., 2006. Are They Really Ready to Work? Employers’ Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century US Workforce. Partnership for 21st Century Skills. 1 Massachusetts Avenue NW Suite 700, Washington, DC 20001.
- Christensen, T., Lægred, P., 2015. Performance and accountability—a theoretical discussion and an empirical assessment. *Publ. Organ. Rev.* 15 (2), 207–225.
- Cowen, R., 2009. The transfer, translation and transformation of educational processes: and their shape-shifting? *Comp. Educ.* 45 (3), 315–327.
- Delors, J., 1996. *Learning: The Treasure within*. UNESCO, Paris.
- Deming, D.J., 2015. The Growing Importance of Social Skills in the Labor Market. NBER Working Paper (21473) (issued August 2015).
- DfE, 2019. Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England - Final Report. Sheffield Hallam University.
- DfEE, 1997. *The Learning Age: A Renaissance for A New Britain*. Green Paper. <https://www.leeds.ac.uk/educol/documents/00000654.htm>. (Accessed 4 September 2014).
- Emler, T.E., Zhao, Y., Deng, J., Yin, D., 2019. Side effects of large-scale assessments in education. *ECNU Rev. Educ.* 2 (3), 279–296.
- Engel, L.C., Frizzell, M.O., 2015. Competitive comparison and PISA bragging rights: sub-national uses of the OECD’s PISA in Canada and the USA. *Discourse: Stud. Cult. Polit. Educ.* 36 (5), 665–682.
- Feniger, Y., Atia, M., 2018. Rethinking cause and effect: analyzing economic growth and PISA scores over a period of 15 years. In: *World Yearbook of Education 2019*. Routledge, pp. 96–110.
- Feniger, Y., Lefstein, A., 2014. How not to reason with PISA data: an ironic investigation. *J. Educ. Policy* 29 (6), 845–855.
- Forestier, K., Crossley, M., 2014. International education policy transfer borrowing both ways: the Hong Kong and England experience. *Compare*. <https://doi.org/10.1080/03057925.2014.928508>.
- Forshay, A.W., Thorndike, R.L., Hotyat, F., Pidgeon, D.A., Walker, D.A., 1962. Educational Achievements of Thirteen-Year-olds in Twelve Countries: Results of an International Research Project, 1959–1961. UNESCO Institute for Education, Hamburg. Retrieved from. http://iea.nl/pilot_twelve-country_study.html.
- Gardinier, M.P., 2021. Imagining globally competent learners: experts and education policy-making beyond the nation-state. *Comp. Educ.* 57 (1), 130–146. <https://doi.org/10.1080/03050068.2020.1845064>.
- Gleckman, H., 2018. *Multistakeholder Governance and Democracy: A Global Challenge*, first ed. Routledge. <https://doi.org/10.4324/9781315144740>.
- Global Innovation Index 2021, <https://www.globalinnovationindex.org/Home>, accessed 28 September 2021.
- Goodhart, C., 1975. Problems of monetary management: the U.K. experience. In: Courakis, A.S. (Ed.), *Inflation, Depression, and Economic Policy in the West*. Barnes and Noble Books, Totowa, New Jersey, p. 116.
- Gove, M., 2011. “Education for Economic Success”, Speech to the Education World Forum available online from: 26 April 2012. (Accessed November 2014).
- Goren, H., 2020. Measuring Global Citizenship: Opting Out and Obscure Comparisons in OECD’s Latest Assessment. NORRAG Blog.
- Grey, S.C., 2020. PISA, the Media and the Governance of Education Policy in England and Sweden. Doctoral dissertation. UCL (University College London).
- Grey, S., Morris, P., 2018. PISA: multiple “truths” and mediated global governance. *Comp. Educ.* 54 (2), 109–131. <https://doi.org/10.1080/03050068.2018.1425243>.
- Hanushek, E.A., Kimko, D.D., 2000. Schooling, labor-force quality, and the growth of nations. *Am. Econ. Rev.* 90 (5), 1184–1208.
- Hanushek, E., Woessmann, L., 2010. The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes. OECD, Paris. Online. www.oecd.org/edu/school/programme-for-international-student-assessment-pisa/the-high-cost-of-low-educational-performance.htm.
- Hanushek, E., Woessmann, L., 2015. Universal Basic Skills: What Countries Stand to Gain. OECD, Paris. Online. www.oecd.org/edu/universal-basic-skills-9789264234833-en.htm.
- Hogan, A., Thompson, G., Mockler, N., 2021. Romancing the public school: attachment, publicness and privatisation. *Comp. Educ.* <https://doi.org/10.1080/03050068.2021.1958995>.
- Högberg, B., Lindgren, J., 2021. Outcome-based accountability regimes in OECD countries: a global policy model? *Comp. Educ.* 57 (3), 301–321. <https://doi.org/10.1080/03050068.2020.1849614>.
- Hopfenbeck, T.N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., Baird, J.-A., 2018. Lessons learned from PISA: a systematic review of peer-reviewed articles on the programme for international student assessment. *Scand. J. Educ. Res.* 62 (3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>.
- Jerald, C.D., 2009. Defining a 21st century education. *Cent. Public Educ.* 16, 1–10.
- Jerrim, J., 2015. Why do East Asian children perform so well in PISA? An investigation of Western-born children of East Asian descent. *Oxf. Rev. Educ.* 41 (3), 310–333.
- Johnes, R., 2017. Entries to Arts Subjects at Key Stage Four. Education Policy Institute, London.
- Kamens, D.H., 2015. A maturing global testing regime meets the world economy: test scores and economic growth, 1960–2012. *Comp. Educ. Rev.* 59 (3), 420–446.
- Komatsu, H., Rappleye, J., 2017. A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth. *Comp. Educ.* 53 (2), 166–191.
- Komatsu, H., Rappleye, J., 2017a. A PISA paradox? An alternative theory of learning as a possible solution for variations in PISA scores. *Comp. Educ. Rev.* 61 (2), 269–297.
- Lauder, H., 2015. Human capital theory, the power of transnational companies and a political response in relation to education and economic development. *Compare* 45 (3), 490–493.

- Li, X., Auld, E., 2020. A historical perspective on the OECD's "humanitarian turn": PISA for development and the learning framework 2030. *Comp. Educ.* 56 (4), 503–521. <https://doi.org/10.1080/03050068.2020.1781397>.
- Meyer, H.D., 2013. OECD's PISA: A Tale of Flaws and Hubris. ID Number: 17371. Teachers College Record. <http://www.tcrecord.org>. (Accessed 28 September 2021).
- Mohamed, M., Morris, P., 2019. "Buying, selling and outsourcing educational reform: the Global Education Industry and "policy borrowing" in the Gulf", *Compare. J. Comp. Int. Educ.* <https://doi.org/10.1080/03057925.2019.1607255>.
- Morris, P., 2012. Pick'n'mix, select and project; policy borrowing and the quest for 'world class' schooling: an analysis of the 2010 schools White Paper. *J. Educ. Policy* 27 (1), 89–107.
- Morris, P., 2015. Comparative education, PISA, politics and educational reform: a cautionary note, *Compare. J. Comp. Int. Educ.* 45 (3), 470–474. <https://doi.org/10.1080/03057925.2015.1027510>.
- National Commission on Excellence in Education, 1983. *A Nation at Risk*, 1983. <https://www2.ed.gov/pubs/NatAtRisk/risk.html>. (Accessed September 2018).
- OECD, 2013a. PISA 2012 Results: Creative Problem Solving Students' Skills in Tackling Real-Life Problems Volume V. <https://www.oecd.org/pisa/keyfindings/PISA-2012-results-volume-V.pdf>. (Accessed 1 October 2021).
- OECD, 2013b. OECD Skills Surveys: About PIAAC. <https://www.oecd.org/skills/piaac/>. (Accessed 1 October 2021).
- OECD, 2018. Preparing Our Youth for an Inclusive and Sustainable World. The OECD PISA global competence framework, Paris. Retrieved from OECD website. www.oecd.org/pisa/Handbook-PISA-2018-Global-Competence.pdf.
- OECD, 2019. PISA 2018 Results (Volume III): What School Life Means for Students' Lives. PISA, OECD Publishing, Paris. <https://doi.org/10.1787/acd78851-en>.
- OECD, 2021. PISA 2018 Results. <https://www.oecd.org/pisa/publications/pisa-2018-results.htm>. (Accessed 1 October 2021).
- Pultar, A., 2021. The elusiveness of the common school in Austria. PISA, politics and the survival of selection in a conservative regime. *Comp. Educ.* 57 (3), 341–359. <https://doi.org/10.1080/03050068.2021.1885131>.
- Rapplee, J., Komatsu, H., 2020. Towards (comparative) educational research for a finite future. *Comp. Educ.* 56 (2), 190–217. <https://doi.org/10.1080/03050068.2020.1741197>.
- Robeyns, I., 2006. Three models of education: rights, capabilities and human capital. *Theor. Res. Educ.* 4 (1), 69–84. <https://doi.org/10.1177/1477878506060683>.
- Samoff, J., 2021. Great Buys, Good Buys, and Bad Buys. *Norrag blog*. <https://www.norrag.org/great-buys-good-buys-and-bad-buys-by-joel-samoff/>. (Accessed 12 September 2021).
- Schleicher, A., 2010. International comparisons of student learning outcomes. In: Hargreaves, A., Lieberman, A., Fullan, M., Hopkins, D. (Eds.), *Second International Handbook of Educational Change*, Springer International Handbooks of Education, vol. 23. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2660-6_29.
- Schleicher, A., 2015. "Meet the Mentor" Interview with Denise Galluci, CEO of the Education Partners, USA, Global Education and Skills Forum 2015. <https://www.youtube.com/watch?v=7jvtsNclMGA>. (Accessed April 2016).
- Schleicher, A., 2018. World Class: How to Build a 21st Century School System. <https://doi.org/10.1787/9789264300002-en>.
- Schleicher, A., 2017. The Importance of Creativity and Gradual Learning for the 21st Century Worker [online video] available at: <https://www.youtube.com/watch?v=RGtWfP14-B4>. (Accessed 15 September 2021).
- Secretary's Commission, 1991. What Work Requires of Schools (SCANS Report). <https://eric.ed.gov/?id=ED332054>. (Accessed 28 July 2018).
- Stone, D.A., 1989. Causal stories and the formation of policy agendas. *Polit. Sci. Q.* 104 (2), 281–300.
- Stronach, I., Morris, B., 1994. Polemical notes on educational evaluation in the age of "policy hysteria". *Eval. Res. Educ.* 8 (1–2), 5–19.
- Truss, E., 2012. Speech on International Evidence. Delivered 11 December 2012, available at: <https://www.gov.uk/government/speeches/education-minister-elizabeth-truss-gives-speech-on-international-evidence>. (Accessed 9 January 2015).
- Verger, A., Moschetti, M.C., Fontdevila, C., 2021. How and why policy design matters: understanding the diverging effects of public-private partnerships in education. In: *Realizing the Abidjan Principles on the Right to Education*. Edward Elgar Publishing.
- Waldow, F., Takayama, K., Sung, Y.K., 2014. Rethinking the pattern of external policy referencing: media discourses over the "Asian Tigers" PISA success in Australia, Germany and South Korea. *Comp. Educ.* 50 (3), 302–321.
- You, Y., Morris, P., 2015. Imagining school autonomy in high-performing education systems: East Asia as a source of policy referencing in England. *Compare*. <https://doi.org/10.1080/03057925.2015.1080115>.
- 2010 The High Cost of Low Educational Performance THE LONG-RUN ECONOMIC IMPACT OF IMPROVING PISA OUTCOMES. (Accessed 5 July 2022).

Assessment, evaluation, and accountability: a brief introduction

Anastasiya A. Lipnevich^a, Lale Khorramdel^b, and Jeffery K. Smith^c, ^a Queens College and the Graduate Center, The City University of New York, New York, NY, United States; ^b TIMSS & PIRLS International Study Center at Boston College, Chestnut Hill, MA, United States; and ^c University of Otago, Dunedin, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Definitions of terms	192
Assessment	192
Evaluation	193
Measurement	193
Testing	193
Accountability	194
Definitions of modifiers	194
Formative/summative	194
Criterion-referenced/norm-referenced	194
High stakes/low stakes	194
Large-scale	195
Standards-based	195
Minimum competency (basic skills)	195
Key characteristics of assessments	195
Validity	195
Reliability	196
Fairness	196
Purposes of educational assessment and types of educational decisions	196
Formative and summative assessment (and assessment for learning)	197
Evaluation and accountability through large-scale assessments	197
Purpose and policy relevance	198
Examples of large-scale assessments	198
Domains and constructs	198
Measurement characteristics and design considerations	199
Trends and innovations	199
Reporting of results	199
Conclusion	199
References	200

The purpose of this chapter is to provide a brief introduction to a very broad set of concepts: assessment, evaluation, and accountability. In doing so, it will be necessary to include the terms “measurement” and “testing” as part of the list; omitting them would leave a gap in one’s understanding of this area of education. It should be noted at the outset that these terms are overlapping, and that their definitions have evolved and changed somewhat over the past 50 years or so. Thus, one might find different definitions in older works on these topics. Finally, there really is not 100% agreement on what all these terms mean. We will try to provide the broadest consensus on their definitions and how they relate to one another. We begin with a series of definitions of terms, then a list of modifiers (adjectives) that have been used in relation to these terms, and then turn to an examination of the categories and sub-categories of these concepts, and how they are related. Following these definitions, we take a step back and look at the process of educational measurement in general—how it is accomplished, key characteristics, types of decisions made from measurements, and trends and innovations in testing. We focus in particular on formative and summative assessment at the student level, and large-scale assessments at the state, national, and international levels, as these issues are critical in the current educational climate. In order to simplify the text, we will often refer to instances of the entire enterprise as “assessment.” as it is the most popular term in usage today. But in looking at the field in general, “educational measurement” is probably a better umbrella term, and we use that occasionally as well.

Definitions of terms

Assessment

Of all the terms to be defined here, assessment is perhaps the most nebulous. It can mean a standardized test used to determine course achievement or how well a school is performing, or it can refer to a conversation between a teacher and a child who is

working on a personal reflection. Bloom (1969) provided an excellent definition of assessment, and although it is not by any means a standard definition, to us it captures the fundamental idea of what assessment is about. He argued that assessment concerns finding the fit between individuals and their environment. Such a definition would require taking into account the individuals and what they are doing/working on/learning at any given point in time. It would relate their efforts and abilities to where they were in their learning, what the learning environment was, and where the learners were headed. But over the past half century, the term “assessment” has been used in countless different ways and today can range from referring to national programs looking at school achievement across the years of schooling, or a learning story between a teacher and a preschooler (Carr and Lee, 2019). It has also become the term of choice for referring to any of a number of *approaches* to determining how well a student is learning. It appears to be a more palatable term than “testing” or “measurement.”

Evaluation

Evaluation has had a number of definitions over the years. House (1980) provided an excellent and simple description of evaluation: “Evaluation leads to a judgment about the worth of something.” (p. 18) House goes on to point out that this does not have to lead to action, although it often does. He notes that one can say that Caesar was a great general without doing anything about that. In education today, evaluation seems to be used primarily for looking at how well instructional programs (or schools, or teachers, etc.) are performing. It is typically assessed by looking at change in the intended outcomes (usually achievement) over time in comparison to a reasonable contrast. For example, how well did this new program do with these students in terms of growth over the year compared to students from the year before who didn’t get this program? Up until the mid 1980s, the term evaluation was often used to refer to evaluating student progress in learning. It was considered a broader term than testing; thus evaluation of student learning might include looking at essays, projects, presentations, etc. In 1971, Bloom, Hastings, and Madaus published the classic *Handbook on the Formative and Summative Evaluation of Student Learning*. “Evaluation” was the term of choice in referring to the determination of student progress and needs. They (Bloom et al., 1981) later published *Evaluation to Improve Learning*, which for us indicates that the term was still in popular use to refer to student assessment then. In 1988, Crooks produced a landmark review of work on classroom assessment activities entitled, “The impact of classroom evaluation practices on students.” Thus, the term evaluation was still in use to refer to this type of activity. Crooks, however, regretted the use of evaluation in the title. He would have preferred assessment instead of evaluation, as he thought that would generate a wider readership because people were using assessment more than evaluation (Crooks, personal communication, 2005). Thus, the mid to late 1980s appears to be when the shift in usage was realized.

Measurement

Measurement is a very broad term and relates to activities both within and outside of education. If we say that someone is six feet tall, that is a measurement. If we say that someone reads at the sixth-grade level, that, too is a measurement. One is based on physiology, the other on education. The fundamental idea behind measurement is that one defines a characteristic of an individual (typically) that is of interest and assigns a number to it that quantifies how much of that characteristic the individual possesses. It is inherently a method of quantification. Educational measurement is the process of obtaining information about the characteristics of students, schools, or educational systems. These characteristics may include knowledge, skills, abilities, as well as range of contextual descriptors, such as socio-economic status (SES), school climate, and others. The measurement process includes the development of protocols and tools for gathering information, procedures for collecting and analyzing the data, as well as evaluating the quality of obtained information. Furthermore, it concerns strategies for communicating scores to audiences, which may include policymakers, teachers, parents, and students, to name a few. Although the parties participating in the measurement process may vary, there are common requirements that are to be achieved in the educational measurement process, such as arriving at valid conclusions about students’ knowledge, skills, abilities, and other characteristics and, consequently, contributing to the improvement of teaching and learning.

Testing

Testing (in education as opposed to testing things such as tensile strength) is a process that typically yields a measurement. It is a specific form of measurement that uses a series of items or an essay (or essays) to ascertain where a student is with regard to some characteristic of interest. Thus, a test typically consists of a set of items (which is how testing experts refer to what is commonly known as “questions”) that are given to a testee. The testee attempts to get the right answers to the questions. Usually, a total score of the number correct is tallied and then either used directly or modified in some fashion to come up with a score. Sometimes the answers provided by the student are given scores other than simply right/wrong (such as a mark on an essay test). Modern testing, especially testing using computers, takes a slightly different approach to this process. A computer-based test usually estimates the testee’s ability after each item is completed, and new items are provided to the testee based on that estimate of ability (if a student gets an item wrong, an easier one is typically given next, and vice versa, until a stable estimate of ability has been obtained). As testing typically generates a scores (sometimes multiple scores on sub-scales—literal comprehension, inferential comprehension, vocabulary on a reading test for example), testing is generally considered to be a subset of the concept of measurement.

Accountability

If you were to ask a group of assessment and measurement specialists to list all the terms pertinent to what they do, many would leave “accountability” off that list. Accountability is more the stuff of political debates than it is of student learning and its assessment. Accountability is the process of determining whether a system is functioning as it should be and who should be held responsible if it isn’t. In education, this translates into the notion that children should be making certain levels of progress, and that testing/assessment can be used to determine whether those levels are being achieved. If not, then actions can be taken to determine how to make improvements. [Haertel and Herman \(2005\)](#) point out that accountability debates in education have been going on for at least a century in the United States. A movement toward accountability began at the state level in the 1970s and crystallized with the publication of “A Nation at Risk,” (National Commission on Excellence in Education, 1983). This report argued that student performance in the US had dropped precipitously over the previous several decades and that immediate steps needed to be taken to rectify the situation. Many observers saw this report and subsequent actions as a political effort to hold teachers accountable, with arguments for incentive—or performance-based pay regimens being called for.

Definitions of modifiers

Most of the terms defined above are used with modifiers to specify just what type of activity is being undertaken. There are many such modifiers that might be used; we pick what we believe are the most widely used here.

Formative/summative

Probably the most widely used modifiers for assessments are formative and summative. They come from the work of [Scriven \(1967\)](#) who used them to describe different purposes for evaluating educational programs. [Bloom \(1969\)](#) took the idea of formative and summative and applied it to use with evaluating (assessing) individual student learning. Formative refers to assessment that is primarily used for helping students learn. Formative assessment provides information to the student and the teacher that will help guide subsequent instruction. Summative assessment provides information that summarizes where students stand at a particular point in time with regard to their achievement. It is not concerned with improving that learning, but rather simply documenting the current state of affairs. Closely related to these terms is the popular nomenclature “assessment for learning” and “assessment of learning” ([Assessment Reform Group, 1999](#)). These terms have come to mean roughly the equivalent of formative assessment and summative assessment (respectively), but it should be pointed out that initially they were defined within a particular, student-centered philosophical framework.

Criterion-referenced/norm-referenced

The results that one derives from a test (often a quantification of student performance) are meaningless unless they are given a context for interpretation. For example, a score of 64, or 720, has no meaning unless a context is provided. Glaser and Nitko argued that we make meaning out of scores by referring them either to some standard of performance, or to a reference group (see, e.g., [Glaser, 1963](#); [Nitko, 1980](#)). A norm-referenced test (or, more appropriately, a norm-referenced interpretation of a test score) compares the score of an individual student to some reference group. This might be as simple as saying, “You got the best score in the class!”, or as sophisticated as saying, “This score is in the 83rd percentile of all children in the national norming group.” It allows the person looking at the score some basis for determining how good it is. A criterion-referenced score makes mean by referring the score to a preset standard. “You have passed your driver’s license examination” is an example of a criterion-referenced interpretation. There are no other people being compared with this score. This score passes. That is all one really needs to know. In education, criterion-referenced tests are often used to determine if a student is ready for some next level of instruction.

Combining the first two sets of modifiers, it is possible for one to have a criterion-referenced formative assessment, or a norm-referenced summative assessment (or other combinations as well).

High stakes/low stakes

The “stakes” of a test, evaluation, or assessment have to do with the consequences for those involved in the activity. The term is taken from the game of poker, where the stakes are the amounts that can be gambled. A high stakes test is one where there is substantial consequence, and low stakes the opposite. The consequences of the test or assessment may be different for different stakeholders in the assessment. For example, a statewide assessment may be very important for the school district in terms of rankings of quality or the possibility that the state will intervene in the school system if results are poor. Thus, at the district level, the stakes are high. But the assessment may have been given at the end of the school year and the results not provided until the next school year, and the results might not be used at all at the pupil level. Thus, the stakes would be very low for individual school children. This is important as it is well-known that the consequences, or stakes, of an assessment can greatly influence student performance ([Wolf et al., 1995](#)).

Large-scale

There is an entire category of tests and assessments that are referred to as “large-scale.” This means that the test is given to all the students (or a sample thereof) at some societal level (a nation, a state, internationally, etc.), and the results are used to talk about where the students are at a given point in time, or how much progress has been made since the previous assessment. Such assessments are usually professionally designed and administered, and carry high stakes at whatever level they are given. The National Assessment of Education Progress (NAEP) in the United States, or the Progress in International Literacy Study (PIRLS), given internationally. Obviously, assessments can be small scale as well, but that term is usually not applied.

Standards-based

Standards-based assessments are those which are based on a clear set of progressions/objectives/goals that are deemed necessary for the students who are being given the assessments. The argument here is that there is a knowable set of progressions that students move through on their way to literacy, numeracy, and other skills, and that these progressions are assessable. In theory, standards-based assessments allow for determining not simply whether students are progressing well, but whether larger entities, such as classes, schools, school districts, or even states and nations, are progressing satisfactorily. Many standards-based assessment programs have led to what is called “teaching to the test.” That is, since statistically equating test forms from 1 year to the next requires that the items on the test remain fairly similar, teachers become familiar with the nature of the items, and work with students on practice items of that type. This can lead to an increase in scores on that test, but not necessarily to other tests that measure the same content, but in a different form.

Minimum competency (basic skills)

Minimum competency testing, sometimes known as basic skills testing, is an idea that tends to go in and out of favor. In the 1980s there was a “back to basics” movement in the United States that saw a focus on what the minimum skills were that a student needed to succeed (at some level) in society—hence, minimum competency. Minimum competency testing is typically large-scale, standards-based, and high stakes, either for students or schools, or both.

Key characteristics of assessments

We turn now to the qualities that assessments need to possess. There are three key characteristics of educational assessments that must be taken into account at all times: validity, reliability, and fairness. Each one of these could be turned into a separate chapter, if not volume, but we will give the reader a flavor of what these concepts represent and offer *Standards for Educational and Psychological Testing* (AERA et al., 2014) as one of the most cited sources to glean additional information.

Validity

Validity may be defined as the soundness of interpretations and uses of assessment results. It is important to remember that the assessment results have different degrees of validity for different purposes and for different occasions. The *Standards for Educational and Psychological Testing* (2014) defines validity as the “degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (p. 11). This definition further emphasizes that validity is not only an attribute of the test; it describes what we intend to do with the results. We would take just a bit of an issue with the definition here as the validity of a measure is not actually the degree to which evidence supports that validity. A test could be valid without any evidence testifying to that status (One can be guilty of a crime even if there is not evidence of that guilt). But, that is a quibble.

The concept of validity went through a series of rather dramatic transformations over the past three-quarters of a century. So, in early editions of the *Testing Standards* (1954, 1974), validity was considered to be three-fold in nature. That is, *content validity* was thought to demonstrate how well the test covered the content area of interest; *criterion validity* compared scores of a test to external variables that measured the behavior or characteristic of interest; and *construct validity* was thought to describe the qualities that the test measures. Cronbach and Meehl (1955) considered construct validity as essential. In 1985, the testing standards revised their description of the three types of validity. Instead, they discussed them as three types of validity evidence, thus emphasizing that a unified concept of validity. Messick (1989) further emphasized a unified conception of validity as well as the importance of considering consequences of scores as part of the validation efforts. He defined validity as: “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment” (p. 14). He discussed different categories in the validity inquiry—from convergent (i.e., whether test scores expectedly relate to other, similar measures) to discriminant (i.e., whether the test is unnecessarily highly related to measures to which it should not relate). Messick argued that construct validation should include testing plausible alternate explanations or meanings of the test score (Again, we would argue that validity is not a judgment about a measure, but rather an inherent characteristic of the measure itself for any given use).

A more recent conception of validity was introduced by Kane (2006) who offered a more pragmatic approach to defining and establishing validity. He proposed that validation comprises two kinds of arguments. The first one is an interpretive argument that “lays out the network of inferences and assumptions leading from the observed performances to the conclusions and decisions based on performances” (p. 23). The second one, called validity argument, provides an evaluation of the overall coherence of the interpretive argument and the plausibility of assumptions. He defined validation as the process of evaluating the plausibility of interpretations and uses, and validity as the extent to which accrued evidence supports suggested uses and interpretations. We would like to refer our reader to Kane (1992, 2006) and to an excellent article of Moss (2013), both of which cover the topic of validity in great detail.

Reliability

Reliability is defined as the consistency of assessment results across replications. That is, would we expect to get similar results if an assessment (or a highly similar form of it) were given to the same students a few days after the initial assessment? Are the results consistent over time? One might think of reliability as the degree to which a measure correlates with itself on two different administrations. A complementary concept, measurement error, focuses on how much a score may vary on a subsequent administration. In education contexts, the sources of measurement error may include items, raters, or occasions. There are numerous indices for reporting reliability, including reliability coefficients, generalizability coefficients, standard errors of measurement, and information functions, among others. The indices are used differently depending on a test theory that is used as a basis.

Reliability may be considered to be a subset of the concept of validity. If a measure doesn’t provide consistent results from one administration to the next, how can it provide valid results? The level of reliability needed in a measure depends on its uses and stakes. For example, reliability is critically important when the consequences of test use are high stakes. In making decisions about classifying a student with a particular type of special need, one would want the reliability to be very high. For looking at the efficacy of an instructional program, reliability at the individual student level would be less critical.

Fairness

The third key feature of educational assessment is fairness. The essence of fairness is that if two students have the same ability with regard to what is being measured, then they should get the same score irrespective of race, ethnicity, gender, or other characteristics. Tierney (2016) breaks this down into three features of fairness. The first is that educational assessments must be democratic, inclusive, and supportive of student diversity. This requires assessment designers to consider ensuring that some students are not in any way favored over others. Specifically, when student characteristics, background, and abilities (e.g., language spoken at home, gender) are not relevant to the construct being gauged, then they logically should not influence the scores resulting from the assessment. The second characteristic relates to stakeholders’ *perceptions* of fairness. When students and stakeholders perceive an assessment as fair, the buy-in and engagement is more likely, as opposed to superficial completion of the assessment. Finally, the pedagogical imperative is of key importance to fairness. Educational assessment has been seen as a social process that may increase opportunities to learn (Moss et al., 2008). The fairness of classroom assessment, regardless of purpose, is affected by classroom relationships and interactions, which in turn influences the learning environment. This loop can be scaled up from individual classrooms to larger systems, based on the assumption that all educational assessments should ultimately inform teaching and learning (Tierney, 2016).

Purposes of educational assessment and types of educational decisions

Assessment activities may focus on educational functions and outcomes at various levels. As noted above, assessments may be used to gather information at the country, state, institutional, program, and individual levels, to name a few. At higher levels, assessments typically incorporate aggregated individual data, whereas at the individual level, the collected data reflect student performance and learning or teachers’ performance indicators (Darling-Hammond et al., 2013). Assessment data from the various levels can be reported in both aggregate and individual level. For example, student graduation data can be reported and analyzed at individual, district, as well as state and national levels (Brookhart, 2011).

Assessments may also be differentiated depending on different types of decisions that rely on the outcomes of educational assessment. Depending on the type of decision, different tools and assessment procedures will be used. For example, assessment can be used to inform *selection decisions*, where school and university admissions consider student scores on tests as a condition for selection. One of the important considerations of this type of decisions has to do with its predictive value. In other words, candidates’ performance on these types of assessments should relate to their success in the program for which they were selected (Dahlke et al., 2019).

A related type of decision that is based on results of educational assessment is *placement decisions*. For example, tests that screen students into advanced, gifted, and talented programs represent placement decisions. The difference between selection and placement is that in the former case, students may be rejected from the institution, whereas in the latter case their elimination is not possible. Students are placed into appropriate educational levels that are likely to result in their improved achievement (Lakin, 2018).

Whereas a single test may be enough to complete selection and placement process, *counseling and guidance* relies on a suite of educational assessments (Chandler et al., 2018). The types of different tools that can be used for this purpose may include interviews, interest inventories, aptitude tests, and achievement batteries. The outcomes of these assessments are discussed with the students during a series of counseling sessions.

Certification decisions represent yet another category, with teacher certification being one of the most illustrative examples. Some states in the US employ paper-and-pencil tests to assess preservice teachers' knowledge. These tests may gauge content knowledge, pedagogy, and professional knowledge. Other states use portfolio assessments or classroom observations to certify teachers (Parfitt and Rose, 2018; Gitomer et al., 2021).

Instructional decisions are probably the decisions most frequently taken in any educational context. These are ongoing assessments that a teacher makes to evaluate students' current level of performance and share information on how they can achieve their desired level of performance. These assessments are also used to provide information to the teacher about how well the students have mastered a task in order to adjust instructional activities (Black and Wiliam, 1998; Brookhart, 2011). These types of assessments are typically referred to as "formative" and are juxtaposed with "summative" ones. We devote the next section to this dichotomy.

Formative and summative assessment (and assessment for learning)

For teachers, the most important aspects of assessment to understand are the ideas of formative and summative assessment, and the closely related ideas of assessment *of* learning and assessment *for* learning. These concepts were introduced above, but they are so important that we elaborate on them here. When Bloom first developed his notion of mastery learning theory (Bloom et al., 1971; Guskey, 2007), he incorporated formative assessment into the theory. He argued that after a relatively brief period of instruction, a short test should be administered to students to see how well they were doing. This test would only be used to guide specific additional instruction for those who needed it, and not be included as part of a grade. In Bloom's conceptualization, the information was primarily for the use of the teacher. Over time, the focus of formative assessment has turned to providing information and guidance on improvement directly to the student. In the mastery learning approach, students would proceed through the course of instruction, and at the end, they would take a summative assessment based on the entire course, and that assessment would be used for a final grade. But summative assessment could also include statewide, national, or international assessments that are used to assess the progress of large groups of students.

The distinction of formative and summative assessment has become widespread throughout education, and formative assessment, along with instructional feedback, is an essential component of teaching almost worldwide (Lipnevich and Smith, 2009, 2018; Kluger and DeNisi, 1996; Ramaprasad, 1983; Sadler, 1989; Wiliam, 2011, 2013). The notion of "assessment for learning" (as opposed to assessment of learning) has become as popular, if not more popular than formative assessment. This notion was developed by the Assessment Reform Group in the United Kingdom in 1989. They proposed the following set of characteristics for assessment for learning that they argued would lead to better instructional practice:

- it is embedded in a view of teaching and learning of which it is an essential part;
- it involves sharing learning goals with pupils;
- it aims to help pupils to know and to recognize the standards they are aiming for;
- it involves pupils in self-assessment;
- it provides feedback which leads to pupils recognizing their next steps and how to take them;
- it is underpinned by confidence that every student can improve;
- it involves both teacher and pupils reviewing and reflecting on assessment data (Assessment Reform Group, 1999, p. 7).

The difference between assessment for learning and formative assessment was in the emphasis placed upon a philosophical commitment to child-centeredness that is found in the assessment for learning conceptualization. It must be said, however, that assessment for learning and formative assessment are used pretty much interchangeably today.

Evaluation and accountability through large-scale assessments

At national and systemic levels many countries have well-designed accountability systems in place, such as large-scale assessments. That is, assessments at the population level are used to monitor and evaluate different educational systems and how well student populations emerging from these systems do at the national level and in cross-country comparisons. However, there is frequently a disconnect between the national agenda and classroom practice. Generally, large-scale assessments do not feed into individual schools, let alone individual classrooms, because they are not designed for this purpose. However, they can potentially be used to compare different school types within a country if the drawn school samples are large enough. Let us explicate the nature of large-scale assessments (LSAs).

Different LSAs show some variation in their aims and frameworks. NAEP, TIMSS and PIRLS are more driven by countries' school curricula (Mullis and Martin, 2015, 2017); NAEP Large-scale assessments (LSAs) are comparative educational surveys that are widely used to inform policy makers and provide large data bases for educational researchers. They compare skills, knowledge, attitudes, and behaviors across different populations and subpopulations of interest. The focus of these assessments is on group-level scores in contrast to large-scale testing programs that aim to assess individuals. This is reflected in the test and sampling design of

LSAs as well as in the reporting of results. LSAs have grown and expanded over the past decades (since the late 1950s) to include a broader range of skill domains and populations, and to better meet the questions posed by policymakers and researchers worldwide (Kirsch et al., 2013). We next provide a general overview of the most important aspects of LSAs; more detailed information on each aspect is provided in Khorramdel et al. (2022).

Purpose and policy relevance

The overall purpose of LSA is to provide policymakers and researchers with valid data and meaningful group-level information about populations and subpopulations of interest with regard to skills, knowledge and behavior and how these relate to educational, economic and social outcomes. The focus is especially on skills that “contribute to the prosperity of nations and economies and to better lives for individuals in those nations” (Kirsch et al., 2013). LSA can be utilized for monitoring educational effectiveness and can help develop ways to improve the quality and equality of education worldwide. With their rich data bases, LSA provide comparative data about countries’ student achievement in relation to key home, school, and classroom variables (Mullis and Martin, 2017). Thus, they enable researchers to examine the relation of students’ proficiencies to their contexts for learning. They also allow for monitoring the effectiveness of a country’s educational system compared to other countries and own expectations, and for measuring progress over time (trends).

Examples of large-scale assessments

LSAs can be roughly categorized into national and international assessments. They assess either student or adult populations of different age groups or school grades within or across countries. There are also regional assessments which compare populations within a particular geographical region. The National Assessment of Educational Progress (NAEP), conducted by the National Center for Education Statistics (NCES), was one of the first of its kind and has assessed elementary and secondary grade students across the US since 1969. A cross-country feasibility study that was conducted for NAEP at the Educational Testing Service (ETS) in the 1980s (Beaton and Barone, 2017; Kirsch et al., 2013), paved the way for other LSA that followed. Some of the most prominent LSAs are the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS). They are both conducted by the International Association for the Evaluation of Educational Achievement (IEA) and the TIMSS & PIRLS International Study Center. These assess students at grades 4 and 8 in over 60 countries. The Program for the International Student Assessment (PISA) and the Program for the International Assessment of Adult Competencies (PIAAC) are both conducted by the Organization for Economic Co-operation and Development (OECD). Here the focus is on 15-year-old students in over 90 countries and 16–65-year-old examinees in over 40 countries, respectively.

Some assessment frameworks¹ focus on school-related skills while others are interested in skills needed to meet real-life challenges at home, work and in the wider community (PISA 2022 Assessment and Analytic Framework²; OECD, 2021). However, the designs and methods utilized in LSAs, as described below, are very similar.

Domains and constructs

The core achievement domains in most LSAs are reading and literacy, mathematics and numeracy, and science. Each domain consists of a content dimension, specifying the subject matter to be assessed, and a cognitive dimension, specifying the thinking processes to be assessed. This leads to different subdomains for each domain. Whereas international reporting and comparisons usually focus on results in the overall domain (e.g., reading), information about subdomains (e.g., locate, understand, and evaluate information) is available in the database as well (note that subdomains are generally highly correlated).

In addition to the listed core domains, NAEP also provides measures for arts, civics, economics, geography, US history, technology and engineering, and writing. TIMSS and PIAAC provide additional measures in different types of problem solving (problem solving and inquiry as part of mathematics and science, and problem solving in technology rich environments and adaptive problem solving, respectively). PIRLS, PISA and PIAAC also measure digital reading, and PISA offers the measurement of financial literacy as well as the option of a new innovative domain in each cycle, which is only assessed once, such as collaborative problem solving, global competence, creative thinking, or learning in the digital world.

LSAs not only collect achievement data but a large set of policy relevant background or context variables through student, parent and school questionnaires as well. These include home and school contexts for teaching and learning, students’ home background, school and classroom resources, teacher and teaching quality, and equity. Context variables can be used to examine and understand group differences in cognitive proficiencies and achievement, and can be related to other external variables like social economic status or school and teaching systems. They also have potential to be leveraged to maximize educational outcomes (Lee and Chen, 2019; Lee and Stankov, 2018).

¹<https://nces.ed.gov/nationsreportcard/assessments/frameworks.aspx>.

²<https://www.oecd.org/pisa/publications/pisa-2021-assessment-and-analytical-framework.htm>.

Measurement characteristics and design considerations

Establishing comparable achievement and contextual scales across populations is an important goal of international large-scale assessments and the basis for group comparisons and the reporting of trends (cf. Khorramdel et al., 2020a,b). This is achieved through high standards in instrument development, specific sampling and test design considerations, standardized data collection, the use of item response theory (IRT; Lord and Novick, 1968), and population modeling (Mislevy et al., 1992; von Davier et al., 2006, 2009; von Davier and Sinharay, 2013). To enable the reporting of trends, LSA are administered in cycles. Depending on the subject domain, NAEP is administered every 2 or 4 years; TIMSS has a 4-year cycle; PIRLS has a 5-year cycle; PISA is administered every 3 years; and PIAAC follows approximately a 10-year cycle. Each LSA cycle is preceded by a field test for evaluating the quality of the instruments, procedures and building capabilities (e.g., preparing countries for the main assessment).

LSAs are based on two important sampling considerations: population and domain sampling (Beaton and Barone, 2017). Population sampling refers to the selection of representative samples with similar distribution of characteristics like in the defined target population. The aim is to provide reliable and accurate measures at the population and subpopulation level. The sample has to be large enough to provide estimates with sufficiently small standard errors, yet small enough to be cost effective. Hence, the goal is to select the smallest sample size that will meet the defined measurement standards (Beaton and Barone, 2017). Domain sampling refers to the makeup of the test itself. The achievement of a valid representation of frameworks and constructs would require a large number of items for each assessed domain, resulting in more items than can be possibly taken by one test-taker. Therefore, LSAs make use of sampling designs (Messick et al., 1983; von Davier et al., 2006; von Davier and Sinharay, 2013) that reduce the burden for individual students and the time and costs for participating countries. Each student is assessed with only a subset of the available items while the overall design ensures that the complete item pool is administered at the national level.

Thus, LSAs do not provide reliable achievement estimates at the individual student level, but allow for reliable estimates at the group and country level (von Davier et al., 2006, 2009; von Davier and Sinharay, 2013). Since these are group-level estimates, they cannot be used or interpreted like individual test scores. More information on this work can be found in Yamamoto et al. (2013, 2019) and the PISA 2018 technical report.

Trends and innovations

Providing trends over time in an everchanging world is one of the biggest challenges of LSAs. As the world around us changes, requiring our skills to evolve with it, LSAs strive to evolve and improve with each assessment cycle. However, the reporting of trends is based on comparable scales over time and significant changes from one assessment cycle to the next can break the trend. Therefore, changes in content or methods have to be made gradually to ensure the trend measure while, at the same time, moving into the future.

One of the most important recent changes in LSAs was the transitioning from a paper-based to a largely digital-based administration. This introduced the possibility of assessing new skills through innovative and interactive tasks and item types, as well as the advancement of new methods for analyzing LSA data (von Davier et al., 2019). Adaptive test designs can accommodate the participation of countries with a larger variation in average student proficiency (Yamamoto et al., 2018a,b, 2019) and the use of automated scoring for open-ended response items can improve the scoring accuracy (Sukkarieh et al., 2012; von Davier et al., 2022; Yamamoto et al., 2017; Zehner et al., 2016).

The expansion of LSA to include a wide range of participating countries and populations within those countries, as well as to assess constructs beyond typical academic domains like reading, mathematics, and science reflect the “growing recognition of the need for lifelong learning as a tool to succeed in rapidly changing economies” (Kirsch et al., 2013).

Reporting of results

Results of LSAs are reported at the international and at the national level. At the international level, the data bases, technical documentation, and policy related reports (including cross-country comparisons) are available on the websites of the organizations that conduct these LSAs (e.g., NCES, IEA, TIMSS & PIRLS International Study Center, OECD). At the national level, each country produces detailed reports on their own country-level results (including the comparison of subgroups and samples within countries and descriptions of educational systems).

Results are reported with regard to different benchmarks (e.g., in TIMSS and PIRLS) or proficiency levels (PISA and PIAAC) with descriptions of each benchmark and proficiency level. International databases, international reports, and national reports are usually released to the public at the same time and covered by the media in each country. The full transparency of the data and results are instrumental in the impact of LSAs on educational policies and their use in educational research.

Conclusion

Our goal with this chapter was to offer a broad overview of a number of concepts that are discussed under the large umbrella of educational assessment such as measurement, testing, evaluation and accountability. We discussed these and related concepts in the context of high versus low stakes assessments as well as individual-based versus group-based assessments. We also described

criterion and norm referenced assessments, standards-based assessments, formative and summative assessment, and presented an introduction into large-scale assessments. Different types of assessments provide a vast amount of information for policy makers and stakeholders and, when used properly, can create opportunities for everyone involved in the educational process. The field of assessment, evaluation, and accountability in education is evolving, and we will continue to harness the power of these assessments and processes to inform decisions that help learners of all levels and educational contexts to achieve their highest life potential.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1954; 1974; 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Assessment Reform Group, 1999. Assessment for Learning: Beyond the Black Box. Qualifications and Curriculum Authority, Cambridge Institute of Education, Cambridge, London.
- Beaton, A.E., Barone, J.L., 2017. Large-scale group-score assessment. In: Bennett, R.E., von Davier, M. (Eds.), *Advancing Human Assessment, Methodology of Educational Measurement and Assessment*. https://doi.org/10.1007/978-3-319-58689-2_8.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess. Educ.: Princ. Policy Pract.* 5 (1), 7–74.
- Bloom, B.S., 1969. Some theoretical issues relating to educational evaluation. In: Tyler, R.W. (Ed.), *Educational Evaluation: New Roles, New Means* (National Society for the Study of Education Yearbook, vol. 68. University of Chicago Press, Chicago, IL, pp. 26–50, 2.
- Bloom, B.S., Hastings, J.T., Madaus, G.F. (Eds.), 1971. *Handbook on the Formative and Summative Evaluation of Student Learning*. McGraw-Hill, New York, NY.
- Bloom, B.S., Madaus, G.F., Hastings, G.F., 1981. *Evaluation to Improve Learning*. McGraw-Hill, New York.
- Brookhart, S.M., 2011. Educational assessment knowledge and skills for teachers. *Educ. Meas.* 30 (1), 3–12.
- Carr, M., Lee, W., 2019. *Learning Stories in Practice*. SAGE Publications, London.
- Chandler, J.W., Burnham, J.J., Riechel, M.E.K., Dahir, C.A., Stone, C.B., Oliver, D.F., et al., 2018. Assessing the counseling and non-counseling roles of school counselors. *J. Sch. Couns.* 16 (7), 7.
- Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. *Psychol. Bull.* 52 (4), 281.
- Crooks, T.J., 1988. The impact of classroom evaluation practices on students. *Rev. Educ. Res.* 58 (4), 438–481.
- Dahlke, J.A., Sackett, P.R., Kuncel, N.R., 2019. Effects of range restriction and criterion contamination on differential validity of the SAT by race/ethnicity and sex. *J. Appl. Psychol.* 104 (6), 814.
- Darling-Hammond, L., Herman, J., Pellegrino, J., Abedi, J., Aber, J.L., Baker, E., et al., 2013. Criteria for high-quality assessment. *Stanf. Center Oppor. Pol. Educ.* 2, 171–192.
- Gitomer, D.H., Martinez, J.F., Battey, D., Hyland, N.E., 2021. Assessing the assessment: evidence of reliability and validity in the edTPA. *Am. Educ. Res. J.* 58 (1), 3–31.
- Glaser, R., 1963. Instructional technology and the measurement of learning outcomes. *Am. Psychol.* 18, 519–521.
- Guskey, T.R., 2007. Closing achievement gaps: revisiting Benjamin S. Bloom's "Learning for Mastery". *J. Adv. Acad.* 19 (1), 8–31.
- Haertel, E.H., Herman, J.L., 2005. A historical perspective on validity arguments for accountability testing. *Teach. Coll. Rec.* 107 (14), 1–34.
- House, E.R., 1980. *Evaluating With Validity*. SAGE Publications, Beverly Hills.
- Kane, M.T., 1992. An argument-based approach to validity. *Psychol. Bull.* 112 (3), 527.
- Kane, M.T., 2006. Validation. *Educ. Meas.* 4, 17–64.
- Khorramdel, L., Pokropek, A., van Rijn, P., 2020a. Editorial—special topic, Part I: establishing comparability and measurement invariance in large-scale assessments, Part I. *Psychol. Test Assess. Model.* 62, 3–10.
- Khorramdel, L., Pokropek, A., van Rijn, P., 2020b. Editorial—special topic Part II: establishing comparability and measurement invariance in large-scale assessments, Part II—old questions, new challenges and possible solutions. *Psychol. Test Assess. Model.* 62, 139–145.
- Khorramdel, L., von Davier, M., Yamamoto, K., Kirsch, I., 2022. Educational surveys: conceptual overview. In: McCaffrey, D., Rupp, A. (Eds.), *Quantitative Research and Educational Measurement, International Encyclopedia of Education, fourth ed., vol. 10*.
- Kirsch, I., Lennon, M., von Davier, M., Gonzalez, E., Yamamoto, K., 2013. On the growing importance of international large-scale assessments. In: von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K. (Eds.), *The Role of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-4629-9_1.
- Kluger, A.N., DeNisi, A., 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention Theory. *Psychol. Bull.* 119 (2), 254–284.
- Lakin, J.M., 2018. Making the cut in gifted selection: score combination rules and their impact on program diversity. *Gift. Child. Q.* 62 (2), 210–219.
- Lee, J., Chen, M., 2019. Cross-country predictive validities of non-cognitive variables for mathematics achievement: evidence based on TIMSS 2015. *Eurasia J. Math. Sci. Technol. Educ.* 15 (8), em1725. <https://doi.org/10.29333/ejmste/106230>.
- Lee, J., Stankov, L., 2018. Non-cognitive predictors of academic achievement: evidence from TIMSS and PISA. *Learn. Individ. Differ.* 65, 50–64. <https://doi.org/10.1016/j.lindif.2018.05.009>.
- Lipnevich, A.A., Smith, J.K., 2009. Effects of differential feedback on students' examination performance. *J. Exp. Psychol. Appl.* 15 (4), 319–333. <https://doi.org/10.1037/a0017841>.
- Lipnevich, A.A., Smith, J.K. (Eds.), 2018. *The Cambridge Handbook of Instructional Feedback*. Cambridge University Press.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA. <https://psycnet.apa.org/record/1968-35040-000>.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education/Macmillan, New York, pp. 13–103.
- Messick, S.J., Beaton, A.E., Lord, F.M., 1983. NAEP Reconsidered: A New Design for a New Era (NAEP Report No. 83-1). Educational Testing Service, Princeton, NJ. <https://files.eric.ed.gov/fulltext/ED236156.pdf>.
- Mislevy, R.J., Beaton, A.E., Kaplan, B., Sheehan, K.M., 1992. Estimating population characteristics from sparse matrix samples of item responses. *J. Educ. Meas.* 29, 133–162. <https://doi.org/10.1111/j.1745-3984.1992.tb00371.x>.
- Moss, P.A., 2013. Validity in action: lessons from studies of data use. *J. Educ. Meas.* 50 (1), 91–98. <https://doi.org/10.1111/jedm.12003>.
- Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J., 2008. *Assessment, Equity and Opportunity to Learn*. Cambridge University Press, Cambridge.
- Mullis, I.V.S., Martin, M.O. (Eds.), 2015. *PIRLS 2016 Assessment Framework*, second ed. TIMSS & PIRLS International Study Center Website <http://timssandpirls.bc.edu/pirls2016/framework.html>.
- Mullis, I.V.S., Martin, M.O. (Eds.), 2017. *TIMSS 2019 Assessment Frameworks*. TIMSS & PIRLS International Study Center Website. <http://timssandpirls.bc.edu/timss2019/frameworks/>.
- Nitko, A.J., 1980. Distinguishing the many varieties of criterion-referenced tests. *Rev. Educ. Res.* 50 (3), 461–485.
- OECD, 2021. *The Assessment Frameworks of Cycle 2 of the Programme for the International Assessment of Adult Competencies*, OECD Skills Studies. OECD Publishing, Paris. <https://doi.org/10.1787/4bc2342d-en>.
- Parfitt, C.M., Rose, A.L., 2018. Collaborating to meet the needs of alternative certification teachers using formative design. *J. Form. Des. Learn.* 2 (1), 49–55.

- Ramaprasad, A., 1983. On the definition of feedback. *Behav. Sci.* 28 (1), 4–13.
- Sadler, R.D., 1989. Formative assessment and the design of instructional systems. *Instr. Sci.* 18 (2), 119–144.
- Scriven, M., 1967. *The Methodology of Evaluation*. American Educational Research Association, Washington, DC.
- Sukkarieh, J.Z., von Davier, M., Yamamoto, K., 2012. From Biology to Education: Scoring and Clustering Multilingual Text Sequences and Other Sequential Tasks. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/j.2333-8504.2012.tb02307.x>.
- Tierney, R.D., 2016. Fairness in educational assessment. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Science+Business Media, Singapore.
- von Davier, M., Sinharay, S., 2013. Analytics in international large-scale assessments: item response theory and population models. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, Boca Raton, FL, pp. 155–174. <https://doi.org/10.1201/b16061-12>.
- von Davier, M., Sinharay, S., Oranje, A., Beaton, A., 2006. Statistical procedures used in the national assessment of educational progress (NAEP): recent developments and future directions. In: Rao, C.R., Sinharay, S. (Eds.), *Handbook of Statistics*, vol. 26. Elsevier, Amsterdam, Netherlands. [https://doi.org/10.1016/S0169-7161\(06\)26032-2](https://doi.org/10.1016/S0169-7161(06)26032-2).
- von Davier, M., Gonzalez, E., Mislevy, R., 2009. What are plausible values and why are they useful?. In: von Davier, M., Hastedt, D. (Eds.), *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments*, vol. 2, pp. 9–36. Retrieved from. https://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02_Chapter_01.pdf.
- von Davier, M., Khorramdel, L., He, Q., Shin, H., Chen, H., 2019. Developments in psychometric population models for technology-based large-scale assessments—an overview of challenges and opportunities. *J. Educ. Behav. Stat.* 44, 671–705. <https://doi.org/10.3102/1076998619881789>.
- von Davier, M., Tyack, L., Khorramdel, L., 2022. Automated Scoring of Graphical Open-Ended Responses Using Artificial Neural Networks. <https://arxiv.org/abs/2201.01783>.
- William, D., 2011. What is assessment for learning? *Stud. Educ. Eval.* 37 (1), 3–14.
- William, D., 2013. Feedback and instructional correctives. In: McMillan, J.H. (Ed.), *SAGE Handbook of Research on Classroom Assessment*. SAGE Publications, Inc., pp. 196–214. <https://www.doi.org/10.4135/9781452218649.n12>
- Wolf, L.F., Smith, J.K., Birnbaum, M.E., 1995. Consequence of performance, test motivation and mentally taxing items. *Appl. Meas. Educ.* 8, 341–352.
- Yamamoto, K., He, Q., Shin, H.J., von Davier, M., 2017. Developing a Machine-Supported Coding System for Constructed-Response Items in PISA. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12169>.
- Yamamoto, K., Khorramdel, L., von Davier, M., Ali, U.S., 2013. Scaling outcomes. In: OECD, *Technical Report of the Survey of Adult Skills (PIAAC)*, third ed. PIAAC, OECD Publishing. https://www.oecd.org/skills/piaac/publications/PIAAC_Technical_Report_2019.pdf
- Yamamoto, K., Khorramdel, L., Shin, H.J., 2018a. Introducing multistage adaptive testing into international large-scale assessments designs using the example of PIAAC. *Psychol. Test Assess. Model.* 60, 347–368. https://www.psychologie-aktuell.com/fileadmin/Redaktion/Journale/ptam_3-2018_347-368.pdf.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2018b. Multistage adaptive testing design in international large-scale Assessments. *Educ. Meas.* 37, 16–27. <https://doi.org/10.1111/emip.12226>.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2019. Introduction of Multistage Adaptive Testing Design in PISA 2018. OECD Working Paper. <https://doi.org/10.1787/b9435d4b-en>.
- Zehner, F., Sälzer, C., Goldhammer, F., 2016. Automatic coding of short text responses via clustering in educational assessment. *Educ. Psychol. Meas.* 76 (2), 280–303. <https://doi.org/10.1177/0013164415590022>.

Building teachers' capacity in assessment *for* learning in rural East Africa

Nicholas Wachira^a and Don A. Klinger^b, ^a Aga Khan Education Services, Dar-es-Salaam, Tanzania; and ^b Te Wānanga Toi Tangata, Division of Education, University of Waikato, Hamilton, New Zealand

© 2023 Elsevier Ltd. All rights reserved.

Introduction	202
The research and learning context	203
Methodology	203
Results	204
Building an AfL culture through questions	205
Building assessment competence	205
The assessment guidelines	206
Discussion	206
Appendix	208
References	211

Introduction

Consider a typical primary grade classroom. To an untrained observer, such a classroom may seem quite chaotic as the teacher shifts from one activity to another and the students shift from quiet attention to boisterous action. Nevertheless, underlying this apparent chaos is a complex learning environment in which the teacher is continually observing, monitoring and assessing students' learning to direct their teaching and support students' learning. Researchers and educators have long understood the value of these actions within the context of teaching; however, over the past 25 years, there has been an even greater emphasis on the potential of such assessment methods to enhance both teaching and learning (Andrade and Brookhart, 2016; Assessment Reform Group, 2002; Birenbaum et al., 2015; Black and Wiliam, 1998, 2006; Chappuis et al., 2012; DeLuca et al., 2018; Stiggins, 2005).

The concept of formative assessment or "Assessment for Learning" has come to represent the critical shifts in the role of assessment as an integral part of classroom learning (e.g., Assessment Reform Group, 2002; Birenbaum et al., 2015; Brookhart, 2011; Laveault and Allal, 2016; Stiggins, 2005). Within the framework, both the teacher and the students are active users of ongoing assessments that are integrated, formally and informally within the interplay of teaching and learning. There is strong evidence that these practices have the potential to positively impact students' educational outcomes (Flóres and Sammons, 2013; Wiliam et al., 2004). In response, teachers across international jurisdictions have been encouraged to implement comprehensive formative assessment practices to better support students' learning. These calls are explicit in national and board policies (e.g., Ryan and Feller, 2009) and classroom assessment research (e.g., Klinger et al., 2016; Laveault and Allal, 2016; Stiggins, 2005; Wylie and Lyon, 2015), and standards of practice (e.g., Brookhart, 2011; Klinger et al., 2015; Ndalichako, 2013). In support, researchers and policymakers have developed guidelines, principles and models to help teachers understand and implement these current assessment conceptions. Despite these supports, research continues to show that teachers struggle to interpret assessment policies and implement practices aligned with contemporary formative assessment notions (Bennett, 2011; DeLuca and Klinger, 2010; DeLuca et al., 2012; Laveault and Allal, 2016).

While there is little debate regarding the potential of formative assessment practices to be a powerful classroom "tool," there has been much debate with respect to developing teachers' capacity to effectively implement these practices (e.g., Bennett, 2011; DeLuca et al., 2012; Laveault and Allal, 2016; Kingston and Nash, 2011; Klinger et al., 2012). While some have focused their critiques on the lack of empirical evidence regarding the effectiveness of formative assessment to impact educational outcomes, (e.g., Dunn and Mulvenon, 2009; Kingston and Nash, 2011), the greatest concerns tend to focus on the challenges of developing teachers' assessment competence. This focus on teachers' competence has highlighted that principles and models alone are insufficient to create the necessary changes in classroom assessment (Wylie and Lyon, 2015). Rather, there is a need for integrated and focused professional learning initiatives that enable teachers to collaboratively explore new assessment practices (City et al., 2009; DeLuca et al., 2015; Flóres and Sammons, 2013). In summarizing their systematic review of assessment for learning and efforts to support its implementation, Flóres and Sammons (2013) identified a series of strategies and practices that would support relevant and effective professional learning. These included "an atmosphere in which teachers are expected to watch others in action—to actively support peer observation" (p. 2), and professional learning opportunities that enable teachers to develop a "flexible and deep understanding" (p.

2) of formative assessment approaches rather than the accumulation of techniques (see also DeLuca et al., 2015, 2019; Guskey, 2000; Klinger et al., 2016).

Of critical importance for our work is the recognition that the professional context in which teachers are working may have large impacts on the success of any efforts to develop teachers' assessment competence and skills (Birenbaum et al., 2015; Flóres and Sammons, 2013). And as Flóres and Sammons (2013) note, these challenges are even further exacerbated "in those contexts in which the ideal conditions observed in research are not given" (p. 4). It is within these "far less than ideal conditions" for professional learning and classroom practice that our work resides. With support from the Aga Khan Foundation and Global Affairs Canada, our work and research in the East Africa region was an attempt to enhance the formative assessment practices of rural primary teachers in this region. Our research not only addressed the paucity of information about classroom assessment in rural East Africa (Ndlichako, 2004, 2013; Rea-Dickins, 2006; Wachira and Klinger, 2017) but also provided an empirical evidence-base on how effective formative assessment practices could facilitate students' learning and promote classroom inclusion and equity within this context.

The research and learning context

As developing countries, the nations of East Africa are typically characterized by one or two large metropolitan centers with substantial portions of their populations living and working in rural communities. As a result, there are substantial differences in terms of access to educational resources, training and opportunities. When asked to consider a typical primary classroom, it is unlikely that the generated image was one in which there are between 140 and 200 children in a room with a single teacher, and that the classroom lacks electricity, has a single blackboard and there are seats for less than half of the children. Nevertheless, such an image could easily be observed of a teacher working in rural eastern Africa. Using Tanzania as an example, as this is where the majority of our work was completed, concerns have long been raised about the quality of primary education in the region, and several policies and reforms have occurred over the years to address these concerns (e.g., Magidanga, 2021; Ministry of Education and Vocational Training, 2013, 2015; Tanzania Institute of Education, 2005; Twaweza-ni sisi, 2015; USAID, 2013).

Similar calls for classroom assessment reforms can be traced back to 1995. The 1995 Tanzanian Education and Training Policy (ETP) acknowledged that teachers in East Africa needed to strengthen their assessment skills to provide quality education (URT, 1995). The ETP further highlighted the need to introduce courses in assessment and evaluation. Yet the primary focus of these subsequent teacher education courses was on the administration of classroom tests and examination and measurement theory and research (Meena, 2009; MOEC, 1996a,b), rather than on the development of skills and competencies of using assessment as part and parcel in classroom teaching. As Ndlichako (2004) argued, this likely contributed to teachers' tendency to largely rely on the use of tests and examinations to evaluate students' learning (see also Halai et al., 2022). Since 2013 the Primary School pre-service teacher training in Tanzania has been managed and coordinated by the National Council for Technical Education (NACTE). Primary school teacher education programs include modules on assessment, measurement and evaluation. Given that the majority of teachers in the rural regions received their teacher education before this time, and the lack of professional development available to teachers in rural communities, it is unlikely these teachers have developed sufficient assessment competence to effectively enact AfL practices (Halai et al., 2022). As those familiar with formative assessment attest, the lack of assessment education poses challenges, especially when considering the crucial role of teachers' assessment practices for teaching and students' learning (Brookhart, 2011; DeLuca et al., 2012; Kingston and Nash, 2011; Klinger et al., 2012).

Methodology

The foundations for our work are based on three pillars for supporting teachers "professional learning and influencing their subsequent instructional practices. First, as described above, advances in our knowledge about formative assessment reflect a valuable pedagogical practice to support teachers' efforts. Second, professional learning interventions for teachers are more successful when they include an on-site, sustained effort, supported by peers and knowledgeable experts. Lastly, professional learning initiatives must be responsive and aligned with the cultural and environmental context in which teachers and students work.

Our collaborative research was conducted in Western Nile, Uganda and Southern Tanzania. The study was structured as follows: first investigated teachers' classroom assessment conceptions and practices, and subsequently developed and implemented a contextually relevant, professional learning "intervention" to strengthen teachers' assessment capacities. Instructional Rounds (IR) has been identified as a powerful professional learning approach that combines long-term, professional learning guided by experts with teacher practice and peer observation (e.g., City et al., 2009; DeLuca et al., 2015; Marzano, 2011). IR was ideal for strengthening teachers' assessment capacity as it builds on teachers' current skills in a manner that teachers find highly supportive and non-

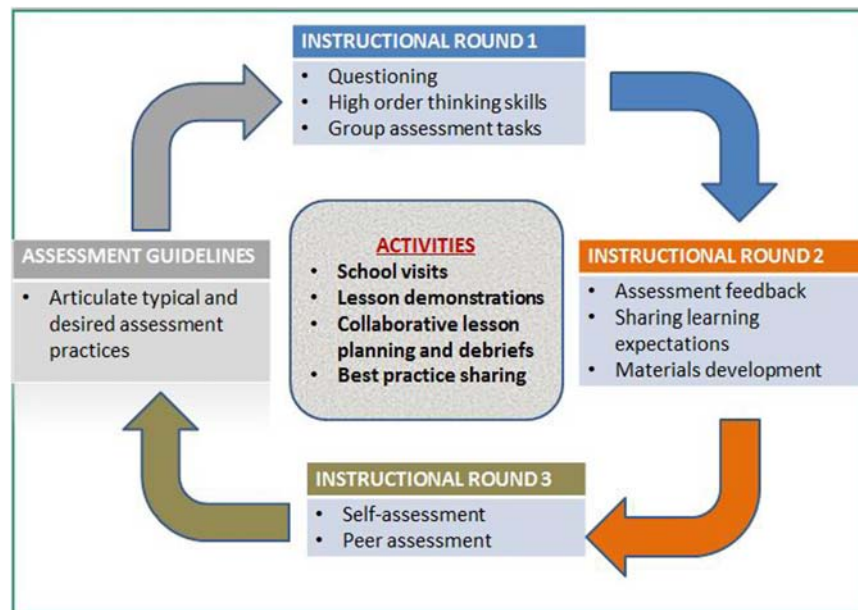


Fig. 1 The stage 2 instructional rounds process.

threatening (see Fig. 1). Importantly, the inclusive nature of IR and the supportive professional learning environment it creates led to the development of exemplary assessment practices, Assessment Guidelines for early literacy and numeracy development and persuasive policy briefs based on our findings (Wachira et al., 2017).

We used a multi-method approach to provide deeper understanding and greater explanatory power, using a progressively focused design with three inter-related stages. In the first stage, we used a mixed-methods survey ($n = 366$) and an ethnographic study (40 teachers and head teachers participating in Stage 2) to ascertain early years teachers' current assessment knowledge and practices. Our primary focus was to determine teachers' current conceptions of assessment and the assessment methods and strategies teachers used in their classrooms. Stage 2 comprised of a 9-month longitudinal professional learning, IR "intervention" to strengthen teachers' assessment capacities through a collaborative method to support students' and teachers' learning (Appendix 1).

We used a two-phased approach, combining three-day teacher workshops with centralized days in which teachers observed each other within their schools and with other schools in the project implementing a recently discussed assessment strategy, followed by an expert-led debriefing session. Teachers subsequently practiced these strategies in their classrooms. The process was repeated on a bimonthly basis. Teachers selected specific AFL learning goals and supported one another in planning for and implementing these goals to support teaching and learning. Through peer observations and assessment expert debriefing, teachers collaboratively engaged in learning about AFL within their local teaching contexts. Through structured activities during the workshops, teachers documented their assessment experiences and identified empirical "assessment episodes" and samples collected during Stage 2. These contextually relevant materials were then used to develop Teachers' Assessment Guidelines with high curriculum validity. The assessment guidelines were developed both in English and Kiswahili so that they could be easily accessible to teachers in both the Tanzanian and Uganda contexts.

Results

The first stage of our research provided critical information regarding teachers' assessment practices, and enabled us to better focus our efforts in subsequent stages. Over 80% of the teachers associated assessment with tests. They considered assessment to be individual paper and pencil written tasks given to students at the end of instruction and subsequently assigned a grade. For these teachers, assessment was a formal written activity done with the intention of grading students. A much smaller number of teachers viewed assessment as being both formal and summative, although it was associated with measuring students' increasing proficiency as instruction progressed. This seems to suggest that teachers' conception of assessment was more focused on performance rather

than on the substance of learning. Assessment was also understood as a diagnostic tool used either at entry to different classes, a new topic or at the beginning of the lesson to find out the students' knowledge and skills before engaging them in learning. While this can be considered one aspect of *AfL*, the vast majority of the teachers only used this information to measure students' attainment rather than as a tool to influence teaching and learning.

Our findings further indicated that more than 87% of the participating teachers had never received any professional development on assessment. Coupled with our review of teacher education materials which found that less than 5% of their education addressed matters of assessment, these teachers would not have been familiar with the growing body of research related to formative assessment and would not have yet obtained the necessary assessment competencies to effectively use *AfL*.

Building an *AfL* culture through questions

The results obtained in Stage 1 demonstrated that teachers considered assessment as being separate from learning. They believed that learning came first, to be followed by a formal, graded assessment. In contrast to these expressed views of assessment, our classroom observations revealed that all of the teachers consistently assessed their students through oral questions. When asked if this was an example of assessment many were hesitant to call it such, and as one teacher insisted *"this is too informal, and has no record."* Regardless, these observations suggested that questioning practices could serve as a doorway into *AfL* as it would build on practices already in place. Admittedly these questioning practices, as with the other assessment activities we observed were very teacher-centric. Teachers' questions tended to solicit low-order thinking responses, offered little feedback to probe or enable students to reflect on their learning, and students rarely directed any questions toward the teacher or their peers.

In contrast to the formality and summative nature of teachers' assessment practices, they did recognize that assessment could be used to improve teaching and learning, along with the more formal evaluative purposes which would keep students and teachers accountable. While we rarely observed this in practice, it once again highlighted an opportunity to enhance these teachers' assessment capacities.

As identified in the first stage, teachers' questioning practices provided an opportunity to introduce the foundations of *AfL* while also enhancing their assessment practices. While questioning may not be a principle of *AfL*, effective questioning practices can provide opportunities for rich feedback and student engagement, both of which are fundamental to *AfL*. The intervention began with a workshop that oriented teachers to three major questioning practices (e.g., Poehner and Rea-Dickins, 2013). Subsequently, IR sessions provided opportunities for teachers to visit and observe each other's developing questioning practices, and to debrief and analyze these experiences. The interventions focused on three major aspects of questioning: (1) high-order thinking questions; (2) wait time; and (3) Student questions, self assessment and peer assessment. Teachers were introduced to Bloom's Taxonomy and its use to design questions at higher levels of thinking. As a next step, teachers were encouraged to design questions before class to ensure the questions would focus on the learning objective and to make sure the questions solicit high-order thinking. Teachers were also taught to give much more wait time after asking questions, from 6 to 10 s before asking for a response. Lastly, teachers learned to provide opportunities for students to ask questions to their teacher and peers. To accomplish this, teachers not only had to build students' confidence to ask questions but also equip students with questioning and reflection skills.

Building assessment competence

The IR sessions provided opportunities to shift practice, and these were supported by workshops that helped to advance teachers' knowledge, skills and competence in using various *AfL* techniques such as learning intentions, feedback, and assessment development. Our focus was to demonstrate the important connections between learning and assessment. Teachers were first encouraged to shift from "learning objectives" to "learning intentions" in order to place a greater emphasis on the process of learning rather than the end product. From here, the value of sharing these learning intentions with students was introduced, along with the use of success criteria as a strategy to monitor progress toward the learning intentions.

If students are to efficiently progress toward the achievement of learning intentions, they must receive learning focused feedback and assess their learning. The types of feedback we promoted were in contrast to the typical feedback we observed in Stage 1 in which the feedback teachers gave to students was short, unhelpful or unclear, and sometimes even irrelevant and demoralizing. The workshops provided teachers with the knowledge and skills to make the shift toward using feedback to support learning. Those familiar with *AfL* know that a key component of *AfL* is in empowering students to take a role in understanding and communicating their progress. Learning is enhanced when students are able to use feedback and also self assess their own progress to assume responsibility and independence. During the workshops, we explored the use of having students ask questions of themselves such as: "Do I understand?" "Do I need help?" "Have I met the learning intention?"

Lastly, workshop sessions focused on the development of teachers' skills in writing assessment items. The workshops sought to create an awareness of different types of assessment tasks under the general categories of selected and constructed response items.

The advantages and challenges of each were discussed in an effort to make teachers aware of the implication of using different assessment items. Teachers were first provided with an opportunity to develop items which were then collaboratively analyzed for item structure, suitability and validity.

The assessment guidelines

Professional learning projects last for only a given time period and have a direct impact on the limited number of those able to participate. Hence a critical component of our work was to document our collective learning and use this to develop guidelines that both teachers who were in the project and others in the region could use to enact contextually relevant AfL practices. These guidelines are intended for use by teachers in a variety of curricular areas and across grade levels, along with head teachers and school principals who have responsibility for supporting the professional practices of their teachers. Further, the guidelines provide a strong foundation for professional learning in classroom assessment, both for pre-service teachers and for teachers who are currently working in schools. The *Classroom Assessment Guidelines* document the experiences of teachers to highlight, with examples, AfL practices used by those who participated in the Instructional Rounds and workshops. As a result, the development of the *Classroom Assessment Guidelines* was a result of collaborative efforts between researchers and teachers. This collaborative process enabled us to translate current knowledge about effective classroom assessment into a narrative that was based on contextually applicable classroom examples, and reflected the developing, classroom assessment practices of early years' teachers.

Our intention here is not to reproduce the *Classroom Assessment Guidelines* published by Wachira et al. (2017) but rather to highlight the format and structure that resulted from our collaborative efforts. The Guidelines were organized into four broad components: (1) Foundations of Effective Assessment; (2) Power of Questions; (3) Creating Assessment Items; and (4) Student Led Assessment. Within each component, a series of narratives described a specific formative assessment practice that the project teachers implemented. The narratives were based on our observations and the descriptions provided by participating teachers. Along with each narrative, key aspects of the practice were identified, extensions to the practice provided, and points to remember listed.

Each component begins with a brief definition and purpose of the component to introduce the assessment concept under discussion and highlight its value and potential impact on students' learning. The section is then followed by various types of assessment under the specific component. For example, the section on student led assessment is then followed by a specific focus on self assessment, peer assessment and group assessment. These specific areas of focus are represented by a narrative describing an assessment episode adopted from one of the actual classroom lessons observed. To capture its substance teachers were interviewed after the class so as to capture their thoughts about the assessment mode they used: purpose, intended outcome and its impact on learning. The purpose of the episode was to raise the contextual relevance of the assessment mode by showing that it can be used in a class that is similar to many teachers' classrooms in East Africa. To do this, the episodes indicate the number of students, the teacher's name and the competency being developed. The episode further focuses on the effective use of a particular assessment mode by indicating the teacher's thoughts and reflections when implementing it. The episode then serves as a point of reference for discussions about the assessment mode.

The episode is then followed by three summaries: (1) Key Highlights; (2) Extending the Formative Assessment Process; and (3) Points to Remember. The "Key Highlights" describe the main points a teacher needs to consider to effectively use the assessment practice under discussion. These highlights capture important procedural points from the assessment episode that make it a more effective practice. "Extending the Formative Assessment Process" provides teachers with strategies to further enhance the assessment practice when they are ready to do so. Through our ongoing work with these teachers, we observed they were at different places in terms of their readiness to use these assessment practices. Hence we wanted to provide guidelines that would provide opportunities for teachers to continue their formative assessment "journey." Lastly, "Points to Remember" identifies the major technical aspects of the assessment strategy and breaks down the assessment strategy into various pedagogical activities that ensure the assessment strategy is successfully implemented. The episode finishes with a "Link to the Competency Based Curriculum" to link assessment to learning as provided in the curriculum. This serves to link the assessment strategy to a specific topic and competency while highlighting the benefits of using that assessment (see Appendix 1 for an example).

Discussion

In an ongoing need to improve the learning outcomes of children, education systems in East Africa have adopted a Competency Based Curriculum (MOEVT, 2013, 2015). The current curriculum emphasizes the application of knowledge, skills and abilities to be learned (learning intentions), along with what students need to demonstrate the learning (competencies). As our research

has highlighted, and in alignment with research in other contexts, the barriers to implementation continue to be related to teachers' exposure to current conceptions about teaching, learning and assessment. As agents of educational change, teachers are critical for effective implementation but they require the supports and resources to meet these expectations (Deluca et al., 2012; Wylie and Lyon, 2015), a real challenge in the rural regions of East Africa (Halai et al., 2022; Magidanga, 2021; Ndalichako, 2013). As Halai et al. noted, and as our work also highlighted, these large primary classrooms were characterized by teacher directed feedback that was overly general and focused on the correct answer. These observations were key to our efforts to build teachers' assessment capacities. We realize the explicit work we did with a group of teachers and the Assessment Guidelines we produced will be insufficient to meet this challenge. Nevertheless, our work and research provide examples and evidence of effective implementation, thus providing a pathway for future success.

Overall, it was evident that teachers' knowledge and AfL related practices improved as teachers became more attuned to students' learning gaps and adjusted their teaching appropriately. Consequently, teachers, albeit at a very initial level, were attempting to meet diverse students' needs and attempted to differentiate and adapt their teaching to raise levels of student achievement. While an attempt was made by teachers to better engage students, we did not observe substantial pedagogical changes in this aspect. Throughout the IR process, teachers expressed concerns about student attention amidst concerns a large class would become chaotic, even though participants' IR observations demonstrated that such engagement was very productive rather than chaotic. Unlike questioning, in which teachers were building on already existing practices, self and peer assessment were a complete paradigm shift for both students and teachers from their former assessment practices. For some teachers, this was a jump too far in the 9 months we worked with them.

Not surprisingly, the level of AfL implementation varied amongst teachers. For some teachers, AfL practices were becoming more naturally integrated into instruction (see also Halai et al., 2022). They used feedback from students to adjust their instruction, routinely shared and used learning intentions to assess students learning and used questions effectively. For others, we observed evidence of improved questioning but a lack of continued use of learning intentions and peer and self-assessment. Nevertheless, Our classroom observations and workshops contained numerous examples of shifting assessment practices. In these classrooms, students were becoming more active in class and teachers were using a variety of pedagogical strategies to engage them (Wachira et al., 2017). We also observed teachers as they began to:

- make connections on how assessment supported teaching and learning and they demonstrated an emerging understanding of the different components and practices of AfL.
- provide task-specific feedback rather than grades that helped create a sense of self-efficacy in their students, leading to higher student participation and engagement in class activities.
- use success criteria that shifted the way teachers viewed their role to include the facilitation of learning, and involve students in determining what successful learning looked like.
- Develop skills in the design of different assessment item formats, resulting in improved validity across assessment items
- accept that when teachers share with students what they will be learning (the learning intentions) and what they are looking for in their students' work (success criteria), student engagement and learning improves.

Those of us who are very familiar with the concepts of formative assessment and AfL might conclude that the practices our research highlighted as a result of the instructional rounds, workshops and *Classroom Assessment Guidelines* do not reflect the "leading edge" of work related to formative classroom assessment. Yet for these teachers, the introduction to Bloom's taxonomy and the value of "wait time" represented substantial shifts in teachers' knowledge and classroom assessment practices, further illustrating the importance of the context in which the professional learning is occurring (e.g., Flóres and Sammons, 2013). Perhaps most importantly, our work with teachers has provided further evidence of the need for sustained resources and support if teachers are to develop the capacity to enact the shifts required. While we agree that a focus on assessment techniques is not enough (DeLuca et al., 2015, 2019; Guskey, 2000; Klinger et al., 2016), our research has highlighted that teachers require an entry point in which to build their assessment capacity, and this entry point must build on teachers' current practices and understandings. Those conducting research and working with teachers need to develop a strong understanding of the educational context and work collaboratively with teachers to identify and meet their professional learning needs.

Appendix

The Power of Questions in Assessment

As an early years' teacher teaching children in Class 1 to Class 3, it is very important to continually assess students' understanding and learning during each class. While there are many formal ways to assess students' learning, one of the most important assessment skills a teacher can develop is that of oral questioning. Some teachers believe there is no special skill to asking questions. This is not at all true. The most effective teachers have worked very hard to develop their questioning skills. These teachers know how to ask simple questions to help students begin to think about their learning, and then, using increasingly more difficult questions, these teachers guide students' further thinking and learning. Teachers who understand the Power of Questions, have learned to give students time to think of answers to the question asked. They have developed ways to seek out the responses of many of the students before identifying the correct or desired response. Effective teachers provide opportunities for all of their students to engage in a questioning dialogue, expanding on the answers of their peers, or asking their own questions of the teacher and their peers to further their own learning. Throughout the questioning process, the teacher carefully listens and observes the students in the classroom, using the students' answers to guide subsequent teaching.



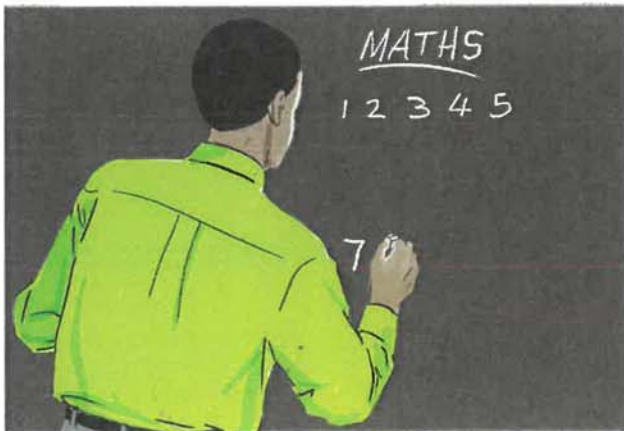
The Assessment Guidelines in this Section provide methods to help you develop more effective questioning process to use in your classroom. The use of these guide-lines and your continued practice will enable you to better use questioning methods to support your teaching and to better engage your students. Help your students use fewer assessments and have a better understanding of your students' learning and achievement. It will take practice to develop these skills but your efforts will be rewarded, saving you time and energy while also improving your students' achievement.



The Power of Questioning

A Teacher's Use of Questions

Juma is a Class 1 teacher. She has 135 students in her classroom, many of whom have only started to attend school over the past 2 months. Juma has worked very hard with the class to ensure the children are able to write the numerical versions for numbers between 1 and 1000. The learning intention Juma is trying to meet is "recognition of the concept of numbers" (see Curriculum for Basic Education, Years 1 and 2). Through her years of experience, Juma knows that students often struggle with numbers over 100, often adding 0's into the numbers by mistake.



The previous class, the students began to learn how to write numbers such as 100, 200, 300. Today Juma wants to introduce numbers between 100 and 200. As a review, Juma orally asks all of the students to work in groups at their desk and write down the numbers 20, 35, and 73. Given the size of her class, she asks each table to hold up the numbers when they have finished. Juma does a quick observation of the numbers and notes that each table has the correct numbers.

Juma does a further review at the board, writing the numbers 99, 100, and 200. As per the routine she has created in her classroom, she asks the class to call out in unison each number as she points to it.

To introduce the next part, Juma asks her students to picture in their head, the number 199. She waits five seconds and then asks the students to put up their hands if they think they know the answer.

Juma asks one of the students to come to the board to write out the number 199. Sofya comes to the board and writes the following. And this is where effective formative assessment practices are very important.

Commonly, a teacher would state that the answer was incorrect and would ask another student to come to the board. Such a strategy would not effectively support the learning of the class. Rather, Juma asks students who have the same answer to put their hand up, noticing that approximately 30 students responded this way. She then asks the class to put up their hand if they have a different answer. After 5 seconds, Juma asks Alexander to tell the class how his answer is different than Sofya's answer. Juma then asks Malai if her answer is the same as Alexander's, to which Malai states yes.



The keys to good questioning

The example of Juma highlights several features of effective teacher questioning strategies.

- Juma has the students work in small groups to complete the review and to check their own understanding with each other.
- Juma uses a variety of methods for students to respond to her questions (writing at the desk, oral response, writing at the blackboard).
- Juma gives the students time to think about their answers to her questions (3 -5 seconds depending on the complexity of the question).
- Juma does not immediately tell the students if they are correct or incorrect. Rather she first seeks to determine what her students are thinking.
- Critical to the effectiveness of these teacher questioning strategies, Juma would follow this questioning structure even if the first response was correct.

Extending the Formative Assessment Process

- Rather than having the students raise their hands if they have the same answer, have all the students with the same answer move to one part of the room and students with a different answer go to another part of the room. Ask the students in these groups another question, for example, write down 201.
- Give students a series of numbers to place in order, include the numbers 199 and 1909.

Points to Remember

1. A teacher's questions during learning can provide important information about students' learning, and their levels of understanding and misunderstanding.
2. Always give students sufficient time to think about their answers to the questions you ask.
3. Thoughtful questions can push students' learning beyond recall of information, helping them to explain their own thinking or understanding.
4. Good questions take time to develop and time to ask and answer in class. Take the time and your students will be more successful.

Competency	Topic	Assessment Tool	Benefit
Arithmetic	Identifying numbers	Questioning	Quickly assessing students learning • Increased feedback • Increased student participation • Identify misconceptions quickly • Probing and scaffolding

References

- Andrade, H., Brookhart, S.M., 2016. The role of classroom assessment in supporting self-regulated learning. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer International Publishing, pp. 293–309.
- Assessment Reform Group, Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M., Stobart, G., 2002. *Assessment for Learning: 10 Principles*. University of Cambridge School of Education, Cambridge, UK.
- Bennett, R.E., 2011. Formative assessment: a critical review. *Assess Educ. Princ. Pol. Pract.* 18 (1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., Wyatt-Smith, C., 2015. International trends in the implementation of assessment for learning: implications for policy and practice. *Pol. Futures Educ. Internet* 13 (1), 117–140.
- Black, P.J., Wiliam, D., 1998. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5 (1), 7–73.
- Black, P., Wiliam, D., 2006. Developing a theory of formative assessment. In: Gardner, J. (Ed.), *Assessment and Learning*. Sage Publications, London, pp. 81–100.
- Brookhart, S.M., 2011. Educational assessment knowledge and skills for teachers. *Educ. Meas.* 30 (1), 3–12.
- Chappuis, J., Stiggins, R., Chappuis, S., Arter, J., 2012. *Classroom Assessment for Student Learning*. Pearson Education, Upper Saddle River, NJ.
- City, E.A., Elmore, R.F., Fiarman, S.F., Teitel, L., 2009. *Instructional Rounds in Education: A Network Approach to Improving Teaching and Learning*. Harvard Education Press, Cambridge, MA.
- DeLuca, C., Klinger, D.A., 2010. Assessment literacy development: identifying gaps in teacher candidates' learning. *Assess. Educ.: Princ. Policy Pract.* 17 (4), 419–438.
- DeLuca, C., Luu, K., Sun, Y., Klinger, D.A., 2012. Assessment for learning in the classroom: barriers to implementation and possibilities for teacher professional learning. *Assess. Matters* 4, 5–29.
- DeLuca, C., Klinger, D.A., Pyper, J., Woods, J., 2015. Instructional rounds as a professional learning model for systemic implementation of assessment for learning. *Assess Educ. Princ. Pol. Pract.* 22 (1), 122–139.
- DeLuca, C., Klinger, D.A., Chapman, A., LaPointe-McEwan, D., 2018. Student perspectives on assessment for learning. *Curric. J.* 29, 77–94 epub 2017.
- DeLuca, C., Klinger, D.A., Chapman-Chin, A., 2019. Towards a teacher professional learning continuum in assessment for learning. *Educ. Assess.* 24 (4), 267–285.
- Dunn, K.E., Mulvenon, S.W., 2009. A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education. *Practical Assess. Res. Eval.* 14 (7), 1–11.
- Flores, M.T., Sammons, P., 2013. *Assessment for Learning: Effects and Impact*. CfBT Education Trust, Reading, UK. <https://files.eric.ed.gov/fulltext/ED546817.pdf>.
- Guskey, T.R., 2000. *Evaluating Professional Development*. Corwin Press, Thousand Oaks, CA.
- Halai, A., Sarungi, V., Hopfebeck, T., 2022. Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania. In: Tierney, R., Rizvi, F., Ercikan, K., Smith, G. (Eds.), *Elsevier International Encyclopaedia of Education*, vol. IV. Elsevier.
- Kingston, N., Nash, B., 2011. Formative assessment: a meta-analysis and a call for research. *Educ. Meas.* 30 (4), 28–37.
- Klinger, D.A., Volante, L., DeLuca, C., 2012. Building teacher capacity within the evolving assessment culture in Canadian education. *Policy Futures Educ.* 10, 447–460. Special Edition: Developing Sustainable Assessment Cultures in School Learning Organisations.
- Klinger, D.A., McDivitt, P., Howard, B.A., Rogers, W.T., Wylie, C., Munoz, M., 2015. *The Classroom Assessment Standards*. Joint Committee on Standards for Educational Evaluation (Amazon).
- Klinger, D.A., DeLuca, C., Chapman, A.E.A., 2016. Implementing Assessment for Learning: A Continuum for Teacher Professional Learning. Paper presented at The Annual Conference of the American Educational Research Association, April 8–12, Washington, DC.
- Laveault, D., Allal, L. (Eds.), 2016. *Assessment for Learning: Meeting the Challenge of Implementation*. Springer International Publishing.
- Magidanga, F.S., 2021. Issues and challenges of primary education in Tanzania; and experiences from other countries. *Asia-Pac. Collab. Educ. J.* 17 (2), 1–20. <https://doi.org/10.14580/apcj.2021.17.2.1>.
- Marzano, R.J., 2011. Making the most of instructional rounds. *Educ. Leader* 68, 80–82.
- Meena, W.E., 2009. Curriculum innovation in teacher education: exploring conceptions among Tanzanian teacher educators. Åbo: Åbo Akad. Förlag.
- Ministry of Education and Culture (MoEC), 1996a. Syllabus for Diploma in Education. Curriculum and Teaching Part A: Theory Guidance and Counselling. Ministry of Education and Culture, Dar es Salaam.
- Ministry of Education and Culture (MoEC), 1996b. Syllabus for Diploma in Education, Education Research, Measurement and Evaluation (Dar es Salaam).
- Ministry of Education and Vocational Training (MOEVT), 2013. *School Improvement Toolkit: Practical Guide for Head Teachers and Heads of School*. MOEVT, Dar es Salaam, Tanzania.
- Ministry of Education and Vocational Training (MOEVT), 2015. Curriculum for Primary School Standard One and Two. MOEVT, Dar es Salaam, Tanzania.
- Ndalichako, J.L., 2004. Towards an understanding of assessment practices of primary school teachers in Tanzania. *Zimbabwe J. Educ. Res.* 16 (3), 168–177. <https://doi.org/10.4314/zjer.v16i3.26046>. Available on.
- Ndalichako, J.L., 2013. Examining Classroom Assessment Practices of Secondary School Teachers in Tanzania. http://www.iaea.info/documents/paper_226dc3fd1.pdf.
- Poehner, M.E., Rea-Dickins, P. (Eds.), 2013. *Addressing Issues of Access and Fairness in Education Through Dynamic Assessment*. Routledge.
- Rea-Dickins, P., 2006. Currents and eddies in the discourse of assessment: a learning-focused interpretation. *Int. J. Appl. Ling.* 16 (2), 163–188.
- Ryan, K.E., Feller, I., 2009. Evaluation, accountability and performance measurement in national education systems: trends, methods, and issues. In: Ryan, K.E., Cousins, J.B. (Eds.), *Sage International Handbook of Educational Evaluation*. SAGE, Thousand Oaks, CA, pp. 171–190.
- Stiggins, R., 2005. From formative assessment to assessment for learning: a path to success in standards-based schools. *Phi Delta Kappan* 87 (4), 324–328.
- Tanzania Institute of Education, 2005. Curriculum for Primary School in Tanzania. Dar es Salaam, Tanzania.
- Twaweza-ni sisi, 2015. Are Our Children Learning? Literacy and Numeracy Across East Africa 2014. https://twaweza.org/wp-content/uploads/2021/01/Uwezo_EA_Report-EN-FINAL.pdf.
- URT, 1995. Tanzania Education and Training Policy. Ministry of Education and Vocational Training.
- USAID, 2013. Proposing Benchmarks for Early Grade Reading and Mathematics in Tanzania. https://ierc-publicfiles.s3.amazonaws.com/public/resources/Report_Reading%20and%20Math%20Benchmarks_Zambia_11_2014.pdf.
- Wachira, N., Klinger, D.A., 2017. Teachers as Agents of Classroom Change: Focus on Early Years' Classroom Assessment. Symposium Presented at The International Congress for School Effectiveness and Improvement, January 7–10. Ottawa, ON.
- Wachira, N., Klinger, D.A., Mrutu, N., Matenga, C., 2017. *Assessment Guidelines: Early Years and Class 1-3*. Aga Khan University and Global Affairs Canada. Institute for Educational Development, East Africa, Dar es Salaam, Tanzania.
- William, D., Lee, C., Harrison, C., Black, P., 2004. Teachers developing assessment for learning: impact on student achievement. *Assess Educ. Princ. Pol. Pract.* 11 (1), 49–65.
- Wylie, E.C., Lyon, C.J., 2015. The fidelity of formative assessment implementation: issues of breadth and quality. *Assess Educ. Princ. Pol. Pract.* 22 (1), 140–160.

Validity for through-year assessments

Scott F. Marion, National Center for the Improvement of Educational Assessment

© 2023 Elsevier Ltd. All rights reserved.

Current examples of through-year assessment systems	212
Additional goals of through-year assessment	212
What's the problem?	213
Additional goals	213
Claims and evidence	213
Claims of instructional and accountability utility	213
Argument-based approach to validity and validation	214
The interpretation-and-use argument	214
Applying an argument-based approach to the validation of through-year assessments	215
Scoring inferences	215
Scoring considerations for accountability uses of through-year assessments	215
Scoring considerations for instructional uses of through-year assessments	215
Generalization inferences	216
Generalizability considerations for accountability uses of through-year assessments	216
Generalizability considerations for instructional uses of through-year assessments	216
Extrapolation inferences	216
Extrapolation considerations for accountability uses of through-year assessments	217
Extrapolation considerations for instructional uses of through-year assessments	217
Decision/use inferences	217
Decision considerations for accountability uses of through-year assessments	217
Decision considerations for instructional uses of through-year assessments	218
The validity argument: synthesizing evidence and logic	218
A call for evidence	219
References	219

"Through-year" (or "through-course") assessment systems are proliferating in the United States. What do such systems entail? Assessments in a through-year system are given in multiple, distinct administrations (also referred to as events and components) across a school year and, further, are intended to support (a) creating a summative determination of student proficiency on state content standards and (b) one additional goal. There are at least ten states and associated assessment companies that are in various stages of exploration, design, and development of such systems. Florida is the most recent case, using what its governor calls a "progress monitoring" system.

The current interest in through-year assessment systems appears to be based on strong assumptions about the multiple purposes that such assessments may serve. These assumptions, in turn, lead to claims about the types of inferences and uses the assessment system is intended to support. But strong assumptions require strong evidence. Is there evidence to support the ambitious assumptions associated with through-year assessment? A validity evaluation is an important way to answer this question. This chapter outlines a path for conducting such an evaluation.

Current examples of through-year assessment systems

Many people regard the administration of three or four independent assessments as the typical through-year design, consistent with how districts administer commercial interim assessments. There are many design variations, however, including (a) weekly, standards-aligned assessments based on a diagnostic classification modeling system, (b) delayed-stage adaptive systems, where through-year components serve as "priors" for the end-of-year assessment, and (c) unit-based assessments tied to a common curriculum. And there doubtless will be additional designs by the time this chapter is in print.

Additional goals of through-year assessment

Crafting a successful through-year assessment system design involves the prioritization of goals, acknowledging that not all goals can be met within a single system. The first step in specifying the additional goal(s) is to ascertain the problem users are trying to address that led to their interest in a through-year design.

What's the problem?

In my work, I have found five general problems with the typical end-of-year test that foster interest in the adoption of through-year assessment designs. First, many users, including district and school leaders, teachers, and even parents, have expressed concerns that their current assessments do not produce instructionally useful information. This concern often is directed at the end-of-year state summative assessment, but it also may reflect the lack of useful information from low-quality local assessments. Most assessment experts acknowledge that end-of-year tests are not designed to provide fine-grained instructional information, but that does not prevent users from wanting more.

Second, most school districts already administer interim assessments multiple times a year, with the last interim assessment usually occurring in the spring. This often leads to concerns about double testing: administering an interim and state summative test in closeness, proximity. It is the district leader's choice to double test, of course, but that does not stop them from pushing for the elimination of the state test because they argue they are getting the information they already need.

Third, some state leaders would like to design and implement an end-of-year test that includes rich performance tasks, both to better measure depth of learning and to signal the types of activities teachers should structure in their classrooms. Unfortunately, administering such tasks takes considerable time. State leaders are reluctant to face stakeholder pushback about excessive testing time, and they also are concerned about the associated validity threats due to student fatigue. Thus, they may turn to a through-year design to spread out the end-of-year testing.

Fourth, education leaders and teachers may be concerned that assessing students just once each year inaccurately captures the student's true level of knowledge because of test anxiety, personal factors, and other idiosyncrasies. Therefore, these stakeholders claim that offering students multiple opportunities throughout the school year should provide for more accurate measurement of students' knowledge and skills.

Finally, the cancellation of state testing in spring 2020, because of the COVID-19 pandemic, revealed the single point of failure of state testing. Specifically, stakeholders had put so much emphasis on end-of-year test results that, when the test disappeared, they were left with no data (in their view). Having a set of assessments, aligned with state content standards, administered throughout the year would have avoided this situation.

Additional goals

As stated at the outset, a through-year assessment system must address at least one additional goal beyond the goal of accountability. At least rhetorically, all assessments included in the through-year system are stated to support accountability purposes as well as at least one other use. These goals derive from the problems described above. The following are some common goals associated with the current set of through-year systems.

- State leaders hope to provide instructionally useful information to educators and students throughout the year to enhance the value proposition of state tests beyond accountability. The design could include a flexible schedule for the administration of through-year assessments by districts to better align with the local scope and sequence.
- State leaders hope to reduce the burden of a single-test administration. The end-of-year summative assessment could be spread out so that the end-of-year test administration requires less time. This was the motivation behind the original through-year design of PARCC, one of the US assessment consortia funded through Race to the Top funds in 2010 (although this design was never operationalized because of concerns regarding the impact on curricular sequencing in participating states and the time requirements for multiple administrations).
- State leaders hope to build an assessment system more coherent than the current constellation of an end-of-year state summative assessment paired with a variety of interim assessments. Through-year designs may help states and districts create more balanced systems of assessments than is currently the case.

The additional goals inform the claims the assessment system is designed to support, and these claims inform the validity argument necessary to evaluate the inferences derived from the scores produced by assessment system. For example, if I assert that changing to a through-year assessment system with multiple assessments administered throughout the school year will address needs for providing test score information that can be used to address instructional needs of students, then I am claiming that the scores from through-year assessments yield information at the appropriate grain size and timing to productively support instructional decisions.

Claims and evidence

Gong (2021a, 2021b) outlined some challenges associated with the claims for through-year systems, such as the comparability of (a) our inferences about student achievement from assessments given throughout the school year and (b) those we would make from a single, end-of-year test. I unpack these issues in the sections that follow.

Claims of instructional and accountability utility

Those claiming that components of the through-year assessment system will support improved instruction should specify their claims by drawing on a theory of action or other heuristic. For example, theory of action might specify that teachers should be

able to gain insights at a small enough grain size to target skills and knowledge students had not yet grasped or identify what students need to learn next to maximize their progress. The theory of action is often a first step in validity evaluation.

The through-year assessment system must produce what the U.S. Department of Education refers to as comparable annual determinations (USED, 2016) in order to service accountability purposes. There is no particular time of year at which these determinations are to be made, but they must be comparable across the state's students and schools. In other words, test scores should support inferences regarding what students know and are able to do that is reflective of the student's learning for each school year/grade level or course. Again, a validity evaluation is how measurement professionals examine the warrants for these sorts of claims. I adopt Kane's (2006, 2013) approach to test validity. Thus, before I outline the specific argument, I provide an overview of Kane's argument-based approach to validity evaluations.

Argument-based approach to validity and validation

Kane (2006) proposed two types of arguments: "An interpretive argument specifies the proposed interpretations and uses of test results by laying out the network of inferences and assumptions leading from the observed performances to the conclusions and decisions based on the performances, [whereas] the validity argument provides an evaluation of the interpretive argument" (p. 17). Kane (2013, 2016) later renamed his interpretive argument as the interpretation-and-use argument (IUA). This was more than a simple name change: It signifies the inseparability of interpretations and uses. Kane (2016) summarized this succinctly: "I think of validity as the extent to which the proposed interpretations and uses of test scores are justified" (p. 198).

The interpretation-and-use argument outlines what the user thinks should occur (and why it should occur) as a result of testing and related systemic endeavors, whereas the validity argument involves weighing the available evidence and logic to decide whether the specific interpretations and uses are justified.

The interpretation-and-use argument

The interpretation-and-use argument can be viewed as a mini-theory for understanding the inferences from test scores in the context of a specific use. Like theory, the IUA guides data collection and methods for conducting validity analyses. Validation is not an exercise in confirmation: Just as theories are falsifiable, so too are IUAs. In this light, one can think of the IUA as a series of if-then statements, such as, "If a student scores at the proficient level on the various through-year components, then they have a solid understanding of the required grade-level content and skills." Of course, there are many more if-then statements for any given test and test score across and within the different categories of inference—scoring, generalization, extrapolation, and decision.

An effective challenge to an IUA is to propose and substantiate an alternative hypothesis that is at least equally plausible as the original assertion. Using the present example above, an alternative inference could be as simple as, "If the student scores at the proficient level on the through-year components, but not on the end-of-year test, then the student possesses only partial knowledge of grade-level standards." When an alternate hypothesis is not supported by the evidence, the user can have greater confidence in the original assertion.

Test validation is the process of making assertions about a test, or testing program, and then collecting data and posing logical arguments to challenge those assertions. Using the original assertion and alternative inference in the previous example, the validity evaluator would design studies to determine the degree to which test scores are sensitive to differences in the instructional experience of students who differ in background knowledge and socio-economic status. The evaluator would then analyze these data, considering the original assertion and alternative hypothesis. For example, if the studies indicate that certain student groups, particularly historically marginalized groups, scored well on the through-year assessments, but not on the end-of-year test, in a pattern that differed from students with more privileged backgrounds, we might conclude there are different expectations on the through-year compared to the end-of-year test or these students were not afforded the opportunities to move beyond proximal and fragile understanding of the required knowledge and skills. In short, validity evaluators are continually attempting to challenge the justifiability of a testing program's claims.

Kane (2006) provides examples of four types of inferences—scoring, generalization, extrapolation, and decision—that are useful for structuring an IUA for a testing program. For each intended inference, a set of arguments is established and thoroughly investigated until the claims can withstand counterargument. Kane provides two examples for scoring: "the scoring rule is appropriate," and "the scoring rule is applied accurately and consistently" (p. 24). A person would then identify the most plausible counterclaims/alternative hypotheses, evaluate the most likely of these plausible counterclaims, and, finally, determine whether the arguments for validating the scoring process still hold after investigating different avenues to debunk each stated argument.

Some might question why an evaluator should adopt Kane's framework when we have *The Standards for Educational and Psychological Tests* (AERA et al., 2014) for collecting and organizing validity evidence with respect to test content, response processes, internal structure, relationships with external variables, and the consequences of testing.

To be sure, *The Standards* address critical sources of validity evidence, but they provide little guidance for organizing, implementing, and synthesizing validity evaluations. Therefore, an IUA should be used to frame validity evaluations while integrating the five sources of evidence from *The Standards* as appropriate.

Insofar as through-year assessment systems are intended to have positive instructional consequences, any validity evaluation of a through-year assessment system must search for evidence of both positive and negative consequences. Because such evidence is

a key aspect of decision inferences, naming consequences as a required source of evidence, as is done in *The Standards*, is absolutely necessary.

Applying an argument-based approach to the validation of through-year assessments

In the sections that follow, I describe Kane's four types of inferences (scoring, generalizability, extrapolation, and decision), offer suggestions for evaluating such inferences for through-year designs, and suggest potential sources of evidence where relevant. I limit this discussion to the accountability and instructional uses of through-year systems. I emphasize that the claims and types of evidence I describe in the next section are just examples. Users and evaluators might find these examples useful, depending on the particular context and through-year design, but I have little doubt that the interpretative argument I outline below will need to be tailored to each case.

Scoring inferences

"The validation of a score interpretation involves an investigation of whether the scores mean what they are supposed to mean, and the interpretation is said to be valid if the claims inherent in the interpretation are supported by appropriate evidence" (Kane, 2016, p. 199). Scoring inferences arguably lie at the heart of any validity evaluation, for if the scores do not "mean what they are supposed to mean," then the validity of any subsequent inference—generalizability, extrapolation, decision—is moot.

Scoring considerations for accountability uses of through-year assessments

The US federal assessment and accountability laws require that students are assessed on the full breadth and depth of the state's content standards. This often is demonstrated through alignment studies and cognitive laboratories for end-of-year assessments. In both cases, students are assumed to have had an adequate opportunity to learn the required content standards by the time of testing. Through-year assessments carry the added assumption that the through-year components include items and tasks appropriate to the student's stage of development at the time of testing. This is especially important where students are expected to progress through the year on the domain in question.

The following examples capture how evaluators, examining the plausibility of scoring inferences for accountability uses, likely would collect validity evidence.

- The scoring rules are consistently, equitably, and accurately applied to all examinees. Supporting evidence would come from differential item-function studies and quality-control documentation.
- There are adequate procedures for ensuring consistency and accuracy of scoring for all students and all identifiable student groups. Supporting evidence would be drawn from scoring consistency and accuracy studies as well as reliability/generalizability analyses.
- Where open-ended items are used, the quality of the selection, training, and monitoring of scorers is adequately described. Supporting evidence would come from documentation of training and monitoring protocols, interrater analyses, and generalizability studies.

Scoring considerations for instructional uses of through-year assessments

Scores from a through-year assessment system must provide information for meaningfully shaping instructional decisions. In order to inform teaching and learning, a research-based model of learning must be specified explicitly in the design. A few examples of claims and potential evidence regarding such inferences include:

- Scoring rules for through-year components maximize the utility of the scores for instructional purposes. Supporting evidence could be drawn from studies employing different scoring rules (particularly for subscores) and working with teachers to evaluate which rules yield better instructional insights.
- The use of approved accommodations supports the intended scoring inferences for different types of test items and examinees. Evidence would be derived from studies indicating the relative comparability of accommodated and non-accommodated scores or perhaps differential boost studies that demonstrate the effectiveness of the accommodation for the students who need them relative to students not needing that accommodation.
- The scoring rules for creating a total score from the multiple assessments are defensible and fair; further, these rules accommodate student development through the year, both for all students and for identifiable student groups. Supporting evidence would come from simulation and empirical studies evaluating various approaches against a priori criteria for fairness and defensibility.
- The scoring rules are designed to yield scores, to the extent practicable, that support instructional decision-making by educators and supporting students' insight into their own learning. Evidence supporting this claim could be derived cognitive laboratories using different scoring models to document the instructional insights from the various models.

Generalization inferences

Generalizability theory treats an assessment's items as representing a sample drawn from the universe of all possible items related to the specific learning targets to which the scores are intended to generalize (Brennan, 1992; Cronbach et al., 1972). Generalizability theory addresses the accuracy with which we can generalize from a person's score on a particular set of items to the person's universe score over the population of items, raters, occasions, and other specified facets. Measurement specialists have developed extensive tools and processes to evaluate error (measurement and sampling) and reliability, both of which are subsumed by generalizability theory. Mapping to the *Standards*, generalizability requires a careful specification of the universe of generalization, which can be part of the evidence related to test content, but evidence related to internal structure is also important.

Generalizability considerations for accountability uses of through-year assessments

Through-year assessments must be able to support the same sorts of generalization inferences as end-of-year tests. However, the former have the added burden of considering how occasions (an important facet in generalizability studies, e.g., Shavelson et al., 1993) affect generalizations regarding students' knowledge and skills. While there are many designs that separate the through-year components from the end-of-year proficiency determination, I assume here that the through-year components count in the annual determination. Potential generalizability claims include:

- Scores from the multiple assessment components support generalizability inferences comparable to the end-of-year test in terms of student performance relative to the test blueprint.
- Scores have the appropriate level of generalizability to support accountability-related uses of these scores, for all students and identifiable student groups.
- The combination of the various assessment components support the intended internal structure of the overall test: how items fit together. Supporting evidence would be a confirmatory item-level factor analysis.
- The internal structure based on the combinations of through-year components is similar across identifiable groups (e.g., cultural, racial, language).
- The use of approved accommodations supports the intended generalizability claims.
- The various errors associated with test scores are evaluated appropriately (e.g., through analyses addressing inter-rater consistency or person-by-item interactions) and are found not to hinder generalizability inferences.
- The various sources of error associated with items and tasks are minimized to the extent practicable.

Generalizability considerations for instructional uses of through-year assessments

The most important generalizability considerations for through year assessments when intended to support instructional uses involves the degree to which the items and tasks included on the through-year assessments support instructional decisions concerning the learning targets of interest. Knowing which items a student got right or wrong is less important for supporting instruction than knowing how well a student understands a particular content domain or subdomain. While many of the generalizability inferences noted above also apply to instructional use, the following inferences are specific to instructional uses.

- Scores from each through-year assessment support inferences regarding students' knowledge and skills in specific subdomains, which are meaningful targets of curriculum and instruction. Supporting evidence would include confirmatory factor analyses demonstrating the unique information produced by each subscore (subdomain). Additionally, evidence should be collected demonstrating that these subscores are at a grain size at which teachers can productively act to support student learning.
- The scores from each assessment event have the appropriate level of generalizability to support the intended instructional use for all students and identifiable student groups. Support for this claim includes evidence that the assessments administered throughout the year produce scores that generalize to instructionally relevant aspects of the domain for that time of year.
- Considered together, the scores from all through-year assessments support inferences about students' knowledge and skills relative to the target domain, for all students and for each identifiable student group. Evidence to support this claim could include generalizability studies demonstrating reasonable coefficients as well as studies indicating the scores provide an indication of students' progress through the target domain.

Extrapolation inferences

Many validity evaluations ignore or downplay extrapolation, in part, because it is challenging to gather appropriate evidence (e.g., Haertel, 2018). Generalization is geared to evaluate inferences related to the specified domain (e.g., the eligible knowledge and skills as specified in the test blueprint or related document). Extrapolation evaluates the degree to which inferences about what students know and can do related to the larger domain (e.g., "biology" beyond what is defined by the test blueprint). In some sense, the extrapolation inference is intended to evaluate the degree to which the domain specified by the test represents the larger domain from which the test specifications were drawn.

The relationship of the test scores to other variables, such as test scores thought to measure the same general domain, can help evaluate inferences regarding whether the test scores represent the larger domain beyond the universe of generalization. Supporting evidence includes results from concurrent and predictive validity analyses involving a criterion representing the domain of interest.

For example, college entrance examination providers should provide content-related validity evidence that the test items and blueprints represent knowledge and skills deemed important for postsecondary education. While such content-related evidence is important, the test's predictive relationship with the criterion—postsecondary success as measured by 1st year GPA, for example—is a critical component of the validity argument for using college entrance exams.

Extrapolation considerations for accountability uses of through-year assessments

The extrapolation inference tends to be less important than the scoring, generalizability, and decision inferences when it comes to accountability use cases. Accountability tests must support inferences related to the defined set of knowledge and skills as well as certain transformations of test scores such as for student growth measures (Haertel, 2018). The examples below highlight some types of evidence an evaluator would collect to support inferences related to extrapolation specific to through-year assessments.

- The full set of test scores relates positively to measures of similar constructs (convergent validity) and negatively to measures of dissimilar constructs (discriminant validity). Through-year assessments used for accountability should have evidence as strong as that produced for an end-of-year test measuring the same domain.
- The relationships of test scores and results from other measures are similar across identifiable groups (e.g., cultural, racial, language).
- The knowledge, skills, and other attributes, as measured by the various through-year components are required for success in the target domain, such as in postsecondary education or employment. The judgment of experts, for example, would be helpful in this regard.

Extrapolation considerations for instructional uses of through-year assessments

Extrapolation inferences related to instructional uses address the extent to which students can apply their knowledge and skills to the larger domain—even beyond that defined by the test blueprint. This pertains to the concept of transfer—particularly distal transfer—which is facilitated by deep understanding (NASEM, 2018). Thus, if through-year assessment systems are intended to support instructional uses, system designers must specify the degree to which assessments support deeper learning and transfer. I am not clear if this is a goal of many (all) of the current through-year systems—it does not appear to be the case based on the designs I have seen (e.g., extensive/exclusive use of multiple-choice items), but for those users with a goal of supporting deeper learning, the following are examples of claims related to the extrapolation inference.

- Through-year assessments call for cognitively complex thinking. Supporting evidence could be drawn from alignment studies that include a focus on depth-of-knowledge and cognitive laboratories.
- The assessments' items and tasks are aligned with the target domain, but also are representative of the larger domain. Supporting evidence could come from the judgments of experts.

Decision/use inferences

Kane (2006) supported the investigation of consequences as part of validity evaluation, even though he did not single out consequences as its own category as did *The Standards* (AERA et al., 2014). Nonetheless, he was unequivocal about the relationship between decision inferences and consequences: "A score use involves a sequence of claims that culminates in a decision, and decisions are justified in terms of their expected consequences" (Kane, 2016, p. 198). As such, evidence related to test consequences play a critical role in evaluating the validity of decision/use inferences. Note, because of the connection between decisions and intended uses, I term this inference the decision/use inference following Kane's interpretative and use argument.

Decision considerations for accountability uses of through-year assessments

Through-year assessment systems used in support of accountability decisions must meet rigorous assessment quality criteria, such as validity (except for consequences), inclusion, and reliability. Critically, these systems must support the classification of students into various achievement levels—the resulting proportions being one of the most important accountability indicators. Evaluators examining the plausibility of decision/use inferences for accountability purposes likely would collect evidence along the lines of these examples:

- There is evidence that high-stakes classification decisions, based on the combination of through-year test scores, are supported through multiple types of studies. These studies should address such considerations as decision consistency and accuracy, content alignment, and subsequent performances (e.g., performance in the next grade).
- Classification decisions (e.g., proficient), based on the combination of through-year test scores or based on individual through-year assessments, do not lead to negative unintended consequences, such as restricted (or denied) opportunities and negative labeling.
- Decisions, such as advanced course opportunities, promotion, or graduation, based on the combination of through-year test scores do not carry negative consequences for identifiable student groups.
- The standard-setting process, based on the multiple through-year assessments, results in cutscores that validly and consistently reflect important differences in student performance along the achievement continuum based on documentation of the process

and materials used, surveys of panelists about how they considered the various through-year components, and other relevant factors. Further, there should be evidence that the score aggregation method used results in more defensible cutscores and achievement levels than another potential approach for aggregating through-year scores.

Decision considerations for instructional uses of through-year assessments

Decision inferences to support instructional uses must attend to the types of decisions educators make in response to the test results to guide and enhance student learning. Further, this aspect of the argument must also, based on through-year assessments, must address the potential unintended negative consequences associated with mining instructional information from accountability tests. More specifically, there is the risk that using such tests for accountability purposes corrupts their potential for instructional benefit (e.g., [Campbell, 1979](#)). What follows are examples of the claims and evidence an evaluator would collect to support inferences related to the decision information for instructional uses of through-year assessments.

- Through-year test components yield results at a grain size sufficient to inform instructional decisions. Supporting evidence could come from teacher surveys, classroom observations, and think-aloud protocols with teachers while they review score reports.
- Through-year assessment components measure important subdomains aligned with local curricula. Supporting evidence could be derived from surveying teachers and reviewing local curricula.
- Teachers can interpret and use the results from the through-year components to appropriately adjust instruction for individual students as well for identifiable student groups. Supporting evidence could be derived from classroom observations, video simulations, think-aloud studies, and other means to gather direct observations (rather than simply a survey of teachers' impressions).
- Teachers do not engage in test preparation or other activities that distract from the instructional aims of the assessment system. Supporting evidence could be drawn from interviews with principals and surveys of educators and students.

The validity argument: synthesizing evidence and logic

Cronbach once proclaimed that “construct validation is better seen as an ever-extending inquiry into the processes that produce a high- or low-test score and into the other effects of those processes” ([Cronbach, 1971](#), p. 452)—which may have frightened people from engaging in validation for fear of getting mired in never-ending studies! In fact, the second part of Kane's approach to validation—the validity argument—involves the synthesis of the various evidence collected according to the IUA as well as the theory and logic underlying the assessment system. Unfortunately, this part of the framework has not been described as well as the first, the IUA. Nevertheless, one of the advantages of Kane's approach is that it provides considerably more structure to validity evaluations than the construct model.

Similarly, [Haertel \(1999\)](#) emphasized that individual pieces of evidence do not determine an assessment system's validity. Evidence and logic must be synthesized to evaluate the IUA. Such activities should allow key stakeholders to make overall judgments about the efficacy of the testing program. Yet, there are few examples of validity argument syntheses in either the literature or in practice.

The Standards ([AERA et al., 2014](#)) include a brief note at the end of the validity chapter that provides general guidance about integrating the various sources of validity evidence into a summary judgment:

A test interpretation for a given use rests on evidence for a set of propositions making up the validity argument, and at some point validation evidence allows for a summary judgment of the intended interpretation that is well supported and defensible. At some point the effort to provide sufficient validity evidence to support a given test interpretation for a specific use does end (at least provisionally, pending the emergence of a strong basis for questioning that judgment) (p. 22).

Unfortunately, *The Standards* do not provide guidance for how evaluators should weigh the various sources of evidence when making a judgment about test use. For example, does strong evidence regarding response processes outweigh weak evidence regarding the relationship to other variables, assuming both are relevant to a particular test use?

Kane's argument-based approach appears fairly conjunctive in that the evaluator first weighs evidence related to scoring before moving on to generalization, extrapolation, and ultimately decision/use inferences. This makes sense in many regards. For example, it would be hard to substantiate a generalization claim without confidence in the scoring inference. But it still seems like a high bar for many reasons, including the challenges associated with satisfying extrapolation claims.

The process of synthesizing the empirical evidence and logical arguments into an evaluative judgment can be a challenging activity even when all evidence is available. But in operational assessment programs, new evidence comes in over an extended time frame, especially the consequence evidence, which amplifies this challenge. [Kane \(2006\)](#) anticipated this problem, suggesting that the initial evaluation of an IUA therefore could adopt a confirmationist bias, whereby the developer and key users look for evidence to support the system design and implementation plans in order to get the program off the ground. However, this does not absolve evaluators of planning longer term studies, such as those related to consequences and fairness; but synthesis

and evaluation in the early years of an assessment program should at least confirm that assessment development generally has proceeded as planned, and the design can be supported with early empirical evidence as well as sound logic.

I recognize the notion of synthesizing the various pieces of evidence into an overall judgment is somewhat abstract. Further, as Ryan (2002) pointed out, different stakeholders and users would weigh the evidence differently. An overall judgment might obscure these different needs. Therefore, instead of an overall judgment by a single evaluator, it makes more sense to array the evidence in a profile or similar format. For example, the assessment company, Cognia and MSAA Consortium (2019), has been working with its clients to produce a “validity scorecard,” whereby the various sources of evidence, according to the IUA, are summarized and rated (e.g., strong evidence, weak evidence, no evidence) such that the user is able to determine which components and strength of evidence are most important to them. My only quibble with this particular approach is that there is a conflict of interest with the test company rating the strength of evidence. The evidence could be produced and an independent entity, such as a state’s technical advisory committee (TAC), could offer recommendations regarding the strength of each aspect of evidence. The TAC because of its unique position could also offer recommendations for the state, for example, regarding which categories of evidence should be prioritized.

A call for evidence

I started this essay by pointing out that the current interest in through-year assessment systems appears to be based on the multiple purposes that such assessments are assumed to serve. These assumptions, in turn, lead to claims regarding the types of inferences and uses that the assessment system is intended to support. But strong assumptions require strong evidence. I have not yet seen the evidence, nor the logical arguments, needed to support many of the assumptions associated with through-year designs. Those advocating through-year systems have a responsibility to produce evidence to support their major assumptions and to evaluate this evidence fairly. Perhaps that is because through-year assessment systems are too new to generate the necessary evidence. Nevertheless, there is significant potential for negative consequences as a result of mixing instructional and accountability uses in the same system. Therefore, there is an urgency to begin constructing an interpretative argument and collecting evidence, especially in terms of both positive and negative consequences.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Brennan, R.L., 1992. Elements of Generalizability Theory. American College Testing Program, Iowa City, IA.
- Campbell, D.T., 1979. Assessing the impact of planned social change. *Eval. Progr. Plann.* 2 (1), 67–90. [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X).
- Cognia & MSAA Consortium, 2019. Multi-state Alternate Assessment 2019 Technical Report. Available from. <https://www.azed.gov/assessment/technical-and-legal-resources>.
- Cronbach, L.J., 1971. Test validation. In: Thorndike, R.L. (Ed.), *Educational Measurement*, second ed. American Council on Education, Washington, DC, pp. 443–507.
- Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N., 1972. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. Wiley, New York, NY.
- Gong, B., 2021a. Why Has It Been so Difficult to Develop a Viable Through Year Assessment? National Center for the Improvement of Educational Assessment. Retrieved from. <https://www.nciea.org/blog/state-testing/why-has-it-been-so-difficult-develop-viable-through-year-assessment>.
- Gong, B., 2021b. Could Two Through Year Assessment Designs Provide Both Summative and Instructional Information? National Center for the Improvement of Educational Assessment. Retrieved from. <https://www.nciea.org/blog/state-testing/could-two-through-year-assessment-designs-provide-both-summative-and>.
- Haertel, E.H., 1999. Validity arguments for high-stakes testing: in search of the evidence. *Educ. Meas.* 18 (4), 5–9.
- Haertel, E.H., 2018. Tests, test scores, and constructs. *Educ. Psychol.* 53 (3), 203–216. <https://doi.org/10.1080/00461520.2018.1476868>.
- Kane, M.T., 2006. Validation. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger, Washington, DC, pp. 17–64.
- Kane, M.T., 2013. Validating the interpretations and uses of test scores. *J. Educ. Meas.* 50 (1), 1–73.
- Kane, M.T., 2016. Explicating validity. *Assess. Educ. Princ. Pol. Pract.* 23 (2), 198–211.
- National Academies of Sciences, Engineering, and Medicine, 2018. *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24783>.
- Ryan, K., 2002. Assessment validation in the context of high stakes assessments. *Educ. Meas.* 21 (1), 7–15.
- Shavelson, R.J., Baxter, G.P., Gao, X., 1993. Sampling variability of performance assessments. *J. Educ. Meas.* 30 (3), 215–232.
- U.S. Department of Education, December 8, 2016. 34 CFR Part 200 Title I—improving the academic achievement of the disadvantaged—academic assessments. *Fed. Regist.* 81 (236).

Assessment and mathematical literacy: a brief introduction

Hugh Burkhardt^a and Alan H. Schoenfeld^b, ^a School of Education, University of Nottingham, Nottingham, United Kingdom; and ^b School of Education, University of California Berkeley, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	220
Mathematical literacy—what is it and why is it important?	222
The defining role of tasks	222
Implications for teaching and learning	224
Assessment tools and processes	225
Desirable properties of an assessment system	225
The importance of the tasks	226
The assessment process	226
WYTIWYG: what you test is what you get	227
The illusion of accuracy	227
A spectrum of assessment models	227
Assessment as part of instruction	227
Coursework	228
Timed examinations	229
Short item multiple choice tests	230
Routes to better assessment	232
The centrality of teacher assessment	233
Reporting—the value of multi-dimensional profiles	233
Acknowledgment	233
References	233

Overview

This section identifies the key themes of this entry, which are elaborated in the sections that follow.

Mathematical literacy (ML) is the ability to use one's mathematical capabilities to better understand the world, its structures and events, and with that understanding to make better-informed decisions. It has become increasingly important as the world changes rapidly; people are faced with more choices, many with uncertain outcomes.

Mathematical literacy is part of literacy in its broader meaning: those aspects of understanding and capability that are of value to most people in enriching their adult lives. These include the ability to understand arguments and to create persuasive ones. What is that politician (or other salesman) saying? What might they actually mean? Given their selection of data, what else might be significant, affecting the outcome? What meaning can I glean from that graph? How can I find good reasons for the action I think desirable, collect information, and build it all into a reasoned persuasive argument? These abilities are important across all school subjects.

The ability to think logically and mathematically about problems as they arise provides key advantages in facing the opportunities and challenges of life and work. The Covid-19 pandemic made this emphatically clear. It led people who are normally busy “just getting on with life” to think about controlling the risks they face, adapting their finances, and wondering about the implications for them of the changing data and model predictions that filled the news.

Learning formal mathematics does not necessarily prepare one for its meaningful use. In a classic study, 70% of the upper elementary students who were asked how many two-foot boards could be cut out of two five-foot board lengths answered “five,” completely ignoring the context in which the problem was set (making a bookcase). The same is the case for what would appear to be directly relevant mathematical tools and techniques learned in secondary school. Although “real world” contexts are increasingly used in presenting tasks, the focus of instruction is usually on the mathematical procedures involved—not the practical implications.

Assessing mathematical literacy. This neglect of mathematical literacy in curricula is mirrored in assessment—with the major exception of PISA, the widely influential OECD *Programme for International Student Assessment*. Since the first tests in 2000, the PISA Mathematics Expert Groups, working with the creative design teams at the Australian Council for Educational Research (ACER) and elsewhere, have developed a sophisticated analysis of the issues involved (see [Stacey and Turner, 2015](#)). The PISA 2021/2022 draft Mathematics Framework, reflecting earlier versions, defines mathematical literacy as

... an individual's capacity to reason mathematically and to formulate, employ, and interpret mathematics to solve problems in a variety of real-world contexts. It includes concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to know the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective 21st century citizens.

Source: <https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf>.

Two tasks are shown in Fig. 1. Note that The OECD website for PISA refers to “units” of connected “questions”, also called “items”—the statistical term for a score point that is often used for any task. The tasks in Fig. 1 are items.

Stacey and Turner (2015) review ML and its assessment in depth in the context of PISA. We refer to their analysis throughout, while expanding beyond the tight constraints on assessment allowed by an international comparison study.

The abilities students need to employ, even within one subject like mathematics, are complex. They include not only knowledge and skills, but the ability use them strategically in tackling diverse types of tasks. A valid, full and fair assessment is correspondingly *complex*—as is a valid, full and fair report on the results. This calls for alignment between curricular goals, instruction, and assessment. ML expands the range of learning goals for Mathematics, increasing the dimensions of complexity in designing assessment.

Assessment and accountability: an unhappy marriage. Assessment and accountability arise from very different needs and reflect fundamentally different philosophies although they are often tightly bound in practice. Government at every level needs to document effectiveness at reasonable cost, presenting results in a form that is simple enough for those concerned to understand and interpret. Complexity in thinking and writing as students work through tasks, and in the judgment process, are condensed into simple results, often just one number. Does a score of 70% on an exam, for example, reflect a student's lack of familiarity with some central concepts, or lack of computational facility? Absent more fine-grained feedback, a report of scores may suffice for accountability purposes but may not at all be useful for identifying areas of needed improvement.

The tension between complex curriculum/performance goals and the need for simple summary feedback is resolved in very different ways across different countries, levels of education, and subject disciplines. We explore and evaluate a range of mechanisms for resolving these differences. Our analysis also applies more widely, wherever complex thinking is assessed.

A mathematical literacy curriculum. The strategic purpose of assessment is to improve teaching and learning. The challenge is, what changes are needed to help *all* students develop the mathematical literacy that they need? Necessary changes involve rebalancing the

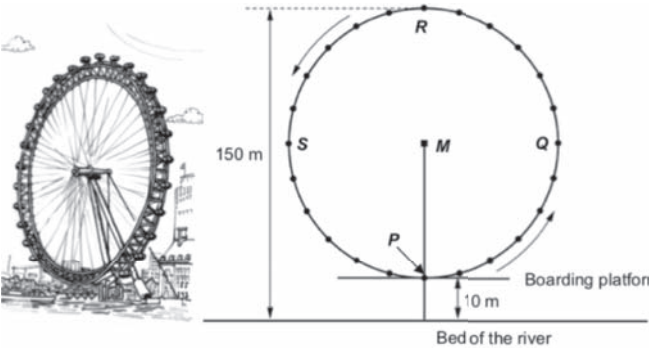
SAUCE		FERRIS WHEEL						
You are making your own dressing for a salad. Here is a recipe for 100 millilitres (mL) of dressing. <table border="1"> <tbody> <tr> <td>Salad oil:</td> <td>60 mL</td> </tr> <tr> <td>Vinegar:</td> <td>30 mL</td> </tr> <tr> <td>Soy sauce:</td> <td>10 mL</td> </tr> </tbody> </table> Q1: How many millilitres (mL) of salad oil do you need to make 150 mL of this dressing?		Salad oil:	60 mL	Vinegar:	30 mL	Soy sauce:	10 mL	A giant Ferris wheel is on the bank of a river. See the picture and diagram below.  The Ferris wheel rotates at a constant speed. The wheel makes one full rotation in exactly 40 minutes. Q2: John starts his ride on the Ferris wheel at the boarding point, <i>P</i>. Where will John be after half an hour?
Salad oil:	60 mL							
Vinegar:	30 mL							
Soy sauce:	10 mL							

Fig. 1 Mathematics in the real world—PISA tasks.

mathematics curriculum with meaningful real world tasks and modifying instruction in ways that give students more responsibility and agency in their learning (see Schoenfeld and The Teaching for Robust Understanding Project, 2016 and Teaching for Robust Understanding, 2015, TRUframework.org). Numerous small-scale curriculum projects (typically at the school district level, affecting at most a few thousand students), demonstrate that these issues can be dealt with productively. However, there is still much to learn about the design factors that enable such projects to succeed, and how to implement such projects at scale. Assessment will play a central role in any transition because of the way it both reflects and influences the curriculum. Mathematical Literacy is in most governments' stated goals but it is rarely supported in practice (e.g., in curricula or assessment). There remain unsolved issues of systemic change (Burkhardt, 2019).

What you test is what you get (WYTIWYG). Assessment providers traditionally claim that examinations merely sample performance; "Of course, all good teachers teach much more than that." However, in systems where schools and teachers are judged on student test scores, most teachers focus their teaching on tasks similar to task in the tests on which their students will be judged. This effect has gradually become recognized, as is the need for broader, more balanced assessments. However, any comparison of the official learning goals with the types of performance actually assessed in most official examinations shows that a profound difference remains. Since most teachers will, understandably, teach to the test, developing "tests worth teaching to" is a key responsibility of those in authority.

Teaching for Robust Understanding (TRU). High quality instruction—instruction from which students emerge as knowledgeable and resourceful thinkers and problem solvers—has been shown to depend on the degree to which the following five dimensions are present: (1) rich disciplinary content and practices; (2) opportunities for "productive struggle" and sensemaking; (3) equitable opportunities for *all* students to engage with core disciplinary ideas; (4) opportunities to contribute to classroom activities in ways that contribute to the development of a sense of agency and positive disciplinary identity; and (5) the use of formative assessment in the service of dimensions 1–4 (Schoenfeld, 2020; see also <https://truframework.org/>).

The sections that follow explore these issues. The section **Mathematical literacy—what is it and why is it important?** sets out the essentials of mathematical literacy, with illustrative examples. The section **Assessment tools and processes** addresses key factors in assessment. The section **A spectrum of assessment models** describes four very different models of assessment in wide use, and their strengths and limitations for assessing ML. The final discussion considers ways in which assessment of ML could change for the better.

Mathematical literacy—what is it and why is it important?

This section provides tasks that exemplify mathematical literacy and assessments that capture it.

The defining role of tasks

Currently the teaching and learning of mathematics is focused on the acquisition of procedural fluency—arithmetic in the early grades, symbolic (primarily algebraic) in later grades. The assessment tasks in Fig. 2 are typical both in their nature and brevity. It is assumed that students should *remember* how to do the task from practicing closely similar tasks in class.

In contrast, the task-types at the heart of mathematical literacy involve much more substantial reasoning. Typically, such tasks ask students to:

- Interpret and translate—*Extract meaning from data, move between representations*
- Explore and discover—*Find relationships, make predictions*
- Evaluate and improve—*Review and improve an argument, a plan, or an artifact.*
- Model and explain—*Invent, interpret, and explain models of real world phenomena*

or, more proactively, to

- Plan and organize—*Find an optimum solution subject to constraints*
- Design and make—*Design an artifact or procedure, and test it.*

Arithmetic	Algebra
• $735 - 546 = \frac{[1]}{[SEP]}$ • Find the sum of the integers from 11 to 19 • In a grocery store, 4 Kg of tomatoes cost \$5. How much will 7 Kg cost?	• Solve $3x + 5 = 21 - 5x$ $\frac{[1]}{[SEP]}$ • Factor $x^2 + 3x - 4$ $\frac{[1]}{[SEP]}$ • Write $\sin(A + B)$ in terms of $\sin$ and $\cos$ of A and B. $\frac{[1]}{[SEP]}$

Fig. 2 Some typical tasks from school mathematics.

Fig. 3 illustrates four of these task-types. While students are expected to have encountered tasks of these general types in instruction, and to be familiar with the real world contexts from which the tasks are drawn, the specific situations to be analyzed are meant to be new. The mathematical techniques needed for these tasks are standard, but each task also requires strategic thinking including thinking through appropriate representations, which aspects of the context need to be captured, which approaches might be profitable, and how to judge the reasonableness of the answers that emerge.

In examinations successful students would typically take 10–15 min to work tasks like each of these. If there is more time available in the curriculum or in assessment, each problem can be expanded in ways that reward deeper exploration.

There is the question of how “real” a task needs to be in order to reflect ML. Burkhardt (1981) classified tasks as: Action, Believable, Curious and Dubious (just to practice the mathematics, as in Fig. 2). PISA 2012 describes problem contexts as *personal*, *societal*, *occupational* or *scientific*, using the classification in balancing the variety of tasks chosen for the survey sample (Stacey and Turner, 2015).

Task difficulty increases with *technical demand*, *complexity*, *unfamiliarity*, and *the length of the chain of reasoning* that the student needs to construct. In traditional mathematics tasks such as those in Fig. 2, difficulty is modified by adjusting technical demand. In contrast, the challenges of mathematical literacy tend to lie in the other three, less technical dimensions. For assessment purposes, where comparable overall difficulty across a range of tasks is called for, technical demand must be lowered to allow for challenges of interpretation, representation, and reasoning. Indeed, Steen (2002) described mathematical literacy as the sophisticated use of relatively simple mathematics, contrasting it with the reverse priorities found in school mathematics. Task difficulty is notoriously easy to misjudge. Stacey and Turner (2015) describe how the creative team for PISA 2012 first estimated the likely difficulty of each task. They then established the difficulty by trialing the task with a cohort of appropriately prepared students in each country, analyzing the responses. This is the only reliable way to establish the difficulty of a complex task.

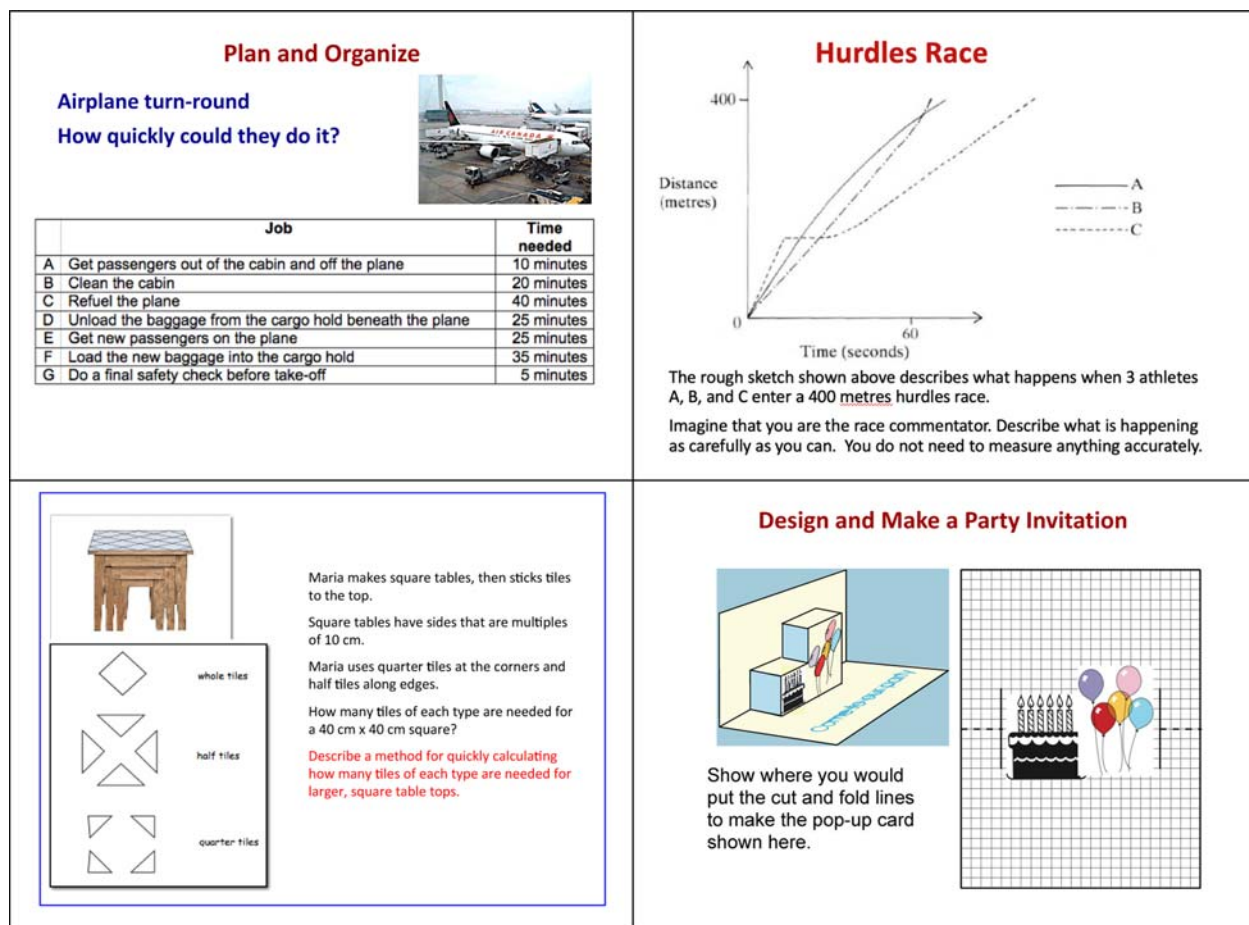


Fig. 3 Mathematics in the real world. Tasks reproduced, with permission, from the Mathematics Assessment Project (2015).

Implications for teaching and learning

The tasks in Fig. 2 reflect a common teaching style, based on “demonstrate and practice” of skills. However, practice alone sustains skills; long-term maintenance depends on having an understanding of the underlying mathematical concepts (see, e.g., Black and Wiliam, 1998).

Due to the fact that most instruction aims at the kind of narrowly focused content-based learning exemplified in Fig. 2, the strategies and extended reasoning one needs to use mathematics in a literate way do not figure much in current school mathematics—or in the understandings of the adults that emerge. Very few classrooms focus on real world problems as substantial as those in Fig. 3, let alone more extended investigations of issues and data of concern.

Typically the purpose of “real world” problems is to provide concrete examples of a concept or skill—perhaps through the counting of stones, or their pairing to illustrate odd and even numbers. A more advanced example is the exemplification of exponential growth by the famous “rice grains on a chessboard” puzzle: “If the number of grains doubles from square to square: 1, 2, 4, 8, ... until all 64 squares are covered, how many grains are there in total?” The initially surprising answer is $2^{64} - 1$, with half of the grains on the 64th square. In fact, this number is around 20,000,000,000,000,000,000. The size of this number vividly makes the point that the phrase “growing exponentially” means much more than its common (mis)use to mean “growing very fast.” Equally clear is that the “chessboard” in the problem statement is conceptual and not possibly real; clearly that many grains of rice don’t fit! “Transfer” to the real world is dubious, as evidenced by the difficulty people had in understanding the truly exponential growth of Covid-19. In reality, when the “doubling time” is a few days, a week or two’s delay in making a decision can make an epidemic a massively different challenge for the health system.

Addressing ML seriously calls for significant changes in instruction:

- Longer timescales—Getting to understand any significant real world situation takes time. Even the classroom problems in Fig. 3 require 10–15 min of thought. Many richer problems are more suited to extended project work. This is common in other subjects but rare in mathematics.
- Reasoning, not just answers—The analytical thinking that is at the core of ML involves building extended chains of reasoning, setting out assumptions and developing them through to a set of conclusions. Such chains of reasoning are essential.

For example, in the Airplane Turnround task in Fig. 3, “the answer” (65 min) depends on the assumptions as to which of the processes listed are necessarily sequential (e.g., unloading before loading; but can you refuel while there are still passengers on board?).

- Communication—Explaining your chain of reasoning in a form that others can understand is an essential part of mathematical literacy—and literacy more widely.

It is clear that all these elements are aspects of a strong response to the assessment tasks in Fig. 3. The modeling and data analysis processes involved in understanding a situation of interest like this, asking and answering questions about it, are summarized in the PISA elaboration of the now-standard modeling diagram, Fig. 4.

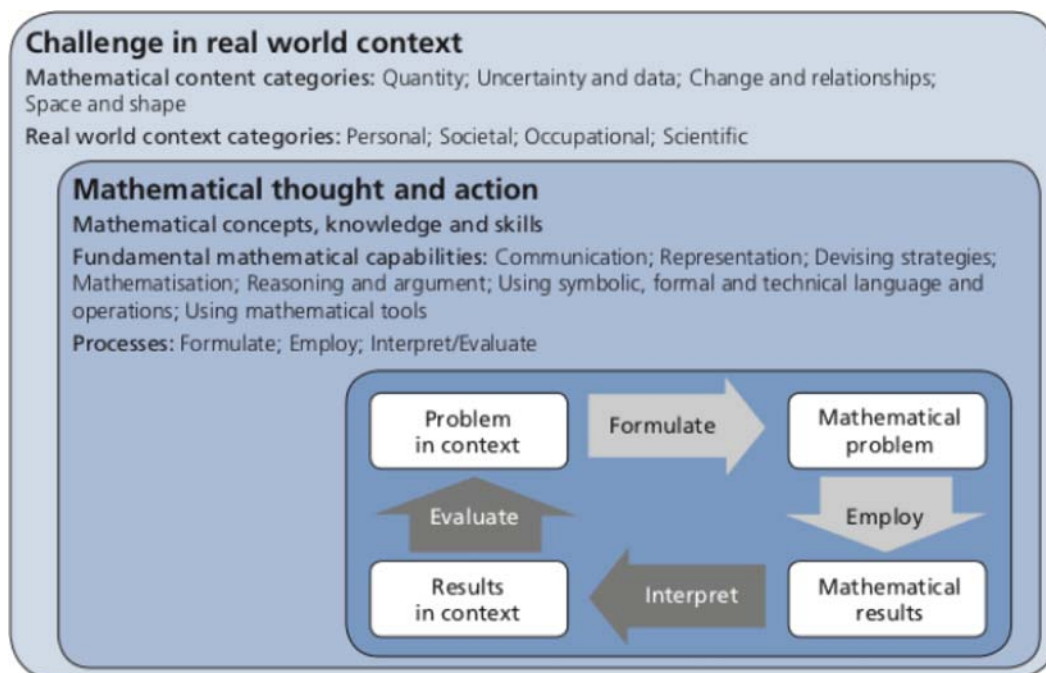


Fig. 4 The PISA 2012 model of mathematical literacy.

The skill set implicated in Fig. 4 is quite complex but, with appropriate teaching, has proved to be accessible to all students. For students to employ the skills in Fig. 4, the mathematical techniques needed must well-integrated into their network of understanding. Developing increasing sophistication involves working through a wide variety of tasks and task types, along with some explicit teaching of problem solving strategies and tactics.

The essentials of classrooms that produce mathematically literate thinkers are summarized in the five dimensions of the TRU framework (Schoenfeld and The Teaching for Robust Understanding Project, 2016; TRUframework.org). Teaching for Robust Understanding requires: (1) high quality content; (2) appropriate cognitive demand, engaging students in productive struggle; (3) equity, ensuring that all students are productively engaged; (4) opportunities for agency and identity, so that students feel the reasoning is their own; (5) formative feedback, from fellow students and the teacher. This takes many teachers of mathematics beyond their current practice and experience. For a discussion see Burkhardt and Schoenfeld (2019).

Assessment tools and processes

The range of task types in a *balanced assessment* should mirror those used in learning activities in classrooms, and be aligned with the performance goals of the curriculum. Such alignment is rarely realized in practice for reasons that are explored here.

Assessment designs intended to provide useful feedback at the individual student or school or district level normally specify the following elements:

- a desired range of knowledge and performance
- a set of task types sampling this range
- the circumstances in which the students will work on the tasks
- a scoring process, using rubrics for each task in evaluating student responses
- an aggregation process for assembling the scoring information into reports
- a way of communicating the results to users.

To explore how these are realized in practice we first examine properties of assessment systems, and then compare specific modes of assessment that exemplify different approaches and priorities.

Desirable properties of an assessment system

What features do reasonable people expect from an official assessment of their children? Using non-technical language the desirable properties of assessment include:

Fairness—widely agreed to be essential. Two similar performances under similar circumstances should be judged as comparable, independent of the individuals involved in the assessment. Fairness demands *equity*: that the assessment should not be in a form that favors one social group over another. Fig. 5 illustrates a quote from Albert Einstein: “Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing it is stupid.”

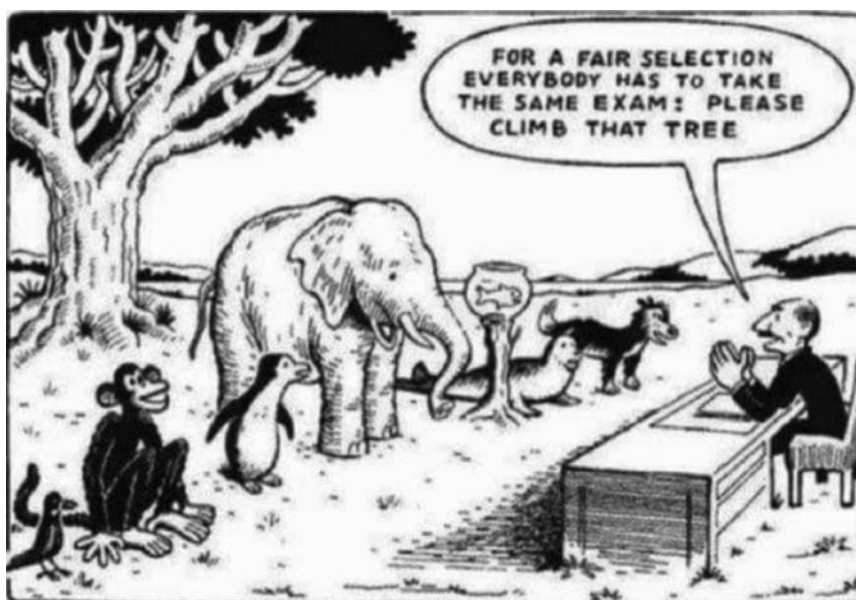


Fig. 5 One illustration of equity as an issue.

Validity of coverage and balance—what is actually assessed. Assessment should match what students have been taught; this, in turn, should fully address the learning and performance goals of the curriculum. In track and field athletics the 10 events in the decathlon are seen as a valid measure of overall “all around” ability. Any one component, such as the 100 m race, is not—it is a valid assessment only of speed over 100 m. Similarly, a spelling test is not a valid assessment of English proficiency, and a set of arithmetic exercises is not a valid assessment of mathematics literacy.

Assessment that has a narrow range, e.g., assessment focused primarily on learned procedures in mathematics, cannot validly assess mathematical literacy when curricular goals include substantial chains of reasoning, communication, and problem solving (Similarly in English, tests of spelling and grammar cannot be said to address issues of literacy).

Interpretability of test scores. What judgments can reasonably be inferred from the scores? Are differences in scores meaningful enough to justify the consequences for students and for teachers and schools that result? This is often called *consequential validity*.

Objectivity of scoring. Are all responses scored on the same basis, avoiding subjective judgments that may vary? This is a key aspect of fairness.

Reproducibility of results (often misleadingly called *Reliability*). If a student takes a different but supposedly equivalent assessment with a different scorer, what is the likely difference in the scores? What are the inter-scorer and test-retest contributions to this variation?

Cost. How much money, and time that may otherwise be used for teaching and learning, does the assessment require?

The ordering above puts educational priorities first. Where accountability is the priority, the order is different, possibly even the reverse. The science of *psychometrics* handles these issues in sophisticated statistical ways that we will not address, other than to note that in nations that employ high-stakes assessments, a key factor in assessment design is the legal defensibility of scoring systems, which is typically insured by task simplicity. Tendencies in that direction are in tension with the desired complexity of meaningful assessments of mathematical literacy.

The importance of the tasks

The tasks and rubrics used in an assessment provide an upper bar on its validity: no meaningful statement regarding any aspect of domain performance is possible if tasks from that domain are not included in the examination. For example, tests that exclude meaningful opportunities to engage in extended chains of reasoning or in problem solving cannot plausibly claim to assess mathematical thinking. As the examples that follow show, many assessments fail this test.

The mismatch between ostensible coverage and exam content is more egregious in Mathematics than, say in English, where examination tasks and their rubrics often cover a broad spectrum of performance that ranges from technical aspects like spelling and grammar to extended pieces of writing varying from commentary and argument to poems and drama. In Mathematics examinations, tasks that require the student to construct extended chains of reasoning on the timescale of an essay are rare at school level; most exam tasks take only a minute or two, assessing recall of learned procedures. Even “problem solving” tasks rarely take more than a few minutes.

The assessment process

Converting a set of tasks that cover specific learning and performance goals into a meaningful and useable assessment involves taking into account the elements noted above: circumstances of performance, scoring, and aggregating scores into a report for users of the results.

Circumstances of performance. Validity depends on students working on the assessment in a mode that enables them to show what they know, understand and can do—not, for example, having to respond on a computer when they are used to working on paper, or vice versa. A fair and valid assessment gives students enough time to show what they can do; if not, students are being assessed on their speed or ability to work under pressure, not on what they know and can do.

Scoring. The evaluation of each student’s response to a task is typically guided by a more-or-less detailed “scoring rubric”, which specifies how to distribute credit across the elements of performance. The design of the scoring rubric is thus integral to the design of the task. It implies value judgments on which people may and do differ. In mathematics, for example, how far the real world context of the task is taken seriously? This is absolutely central for assessing mathematical literacy but still rare in Mathematics assessments. Machine scoring may be reliable for short items, but there are costs in task complexity.

Aggregation and reporting. Different users of the results of an assessment have very different needs and priorities. Teachers can profit from access to the rich information contained in each student’s work—full scripts reveal the student’s thinking. For teachers there is no substitute for seeing the actual scored responses. Although this happens naturally in classroom assessment, it is unfortunately rare in external assessment for practical reasons including cost—and to avoid disputation over scores. The simple results on external accountability measures provide no useful formative guidance for improvement. Computer technology offers the opportunity for sophisticated ways of integrating the data from a student’s responses, allowing profile scores with several dimensions, but this use of technology is rare.

WYTIWYG: what you test is what you get

In practice, where assessment has significant consequences for teachers and students, it is not surprising that teachers in most classrooms focus on the types of tasks that are in the mandated assessments. With high-stakes assessment, what you test is what you get; the tested curriculum *becomes* the implemented curriculum. As a result, the true validity of testing depends on how close the tested curriculum is to the intended curriculum.

The illusion of accuracy

People have come to view examination results as a true and just measure of a student's ability in the subject of the test. If this were true, a response scored by a different scorer would get the same score; but this *inter-scorer consistency* is far from perfect—particularly for extended response tasks such as essays but not only for these. Then there is *test-retest variation*—if the same student tackled a supposedly equivalent examination, would they get the same score or grade? The answer is “sometimes”. For example, in the important school leaving “A-level” examination in England, experts recognize that about half the students would get a higher or lower grade (on the scale A-F) on an equivalent examination—the previous year's, for example. This change could affect whether students get into the university of their choice. Even more serious are tests where the outcome is more life-changing—for example, the “11+” test that determines whether students go to a more, or a less academic high school. One study had the same students take tests from two different years; only half the students that “passed” in the two tests were the same. Even in multiple choice tests that are designed to optimize consistency, random variability of about a quarter of the points assigned is normal.

In summary, high-stakes examinations have not only limited scope but, more importantly, limited accuracy—and thus an element of lottery. We do *not* advocate efforts to improve accuracy, because this is often done at the expense of validity, often by narrowing the task set to exclude richer tasks that involve extended reasoning. Quite the reverse. We believe that society would benefit from openly recognizing the random element in examination scores, then lowering the stakes by giving weight to other information, such as the student's record in their school. This shift in emphasis would open up the possibility using more open tasks that better reflect mathematical literacy.

We note that in the UK, studies show that students from disadvantaged backgrounds do better at university than more advantaged students with the same examination grades. Studies in the US have shown that school grade point average (GPA) is a better predictor of undergraduate performance than the psychometrics-driven school aptitude test (SAT). For this reason among others (e.g., serious concerns about equity), the University of California and other prestigious institutions in the US no longer consider SAT or ACT test scores as a factor in admissions or fellowship decisions. Rich data can give a more useful picture than that from narrower formal tests.

A spectrum of assessment models

There is a wide variety of ways that assessment is realized in practice. The four examples below span the range of styles. The sequence in 4A to 4D may be characterized as a ladder of decreasing richness of information and complexity. These reflect the different needs of their primary users. We shall look at each of these approaches from various perspectives, noting the different balance among the desirable properties we described above.

Assessment as part of instruction

This takes various forms including regularly scheduled tests and, less formally, teacher observations of student work while teaching. Fig. 6 presents an ML task from a professional development module from bowlandmaths.org.uk—“believable” in a societal context. The module focus is on *Handling unstructured problems* in the classroom, where the teachers compare this task with a “school math” version that guides the student through the process step by step.

The primary purpose of classroom assessment is, or should be, *formative*—to guide the teacher's subsequent instruction to better fit the needs of each student. Realizing formative assessment in their classroom is challenging for teachers: while all teachers diagnose student difficulties at some level, making effective modifications to one's next move with a class of 25 individual students is not easy. Doing so requires adaptive expertise (Hatano and Inagaki, 1986) and a supportive curriculum that utilizes student-student discussion, both to reveal misconceptions and to help students resolve them. Black and Wiliam (1998) showed that formative assessment can substantially improve student learning. Burkhardt and Schoenfeld (2019) documented the support that specifically designed teaching materials, exemplified by the Formative Assessment Lessons at map.mathshell.org, can provide teachers in meeting the challenges.

Many of these lessons focus on mathematical literacy. For example, *Matchsticks* asks students to estimate how many matchsticks could be made from a pine tree, given the dimensions of both. The lesson structure provides feedback for the teacher and for peer assessment by students on: the students' assumptions; their concepts of length-area-volume; changes of units; and the proportional reasoning that is involved at multiple points in the problem. *Medical testing* addresses some counter-intuitive but essential understandings regarding proportions of false positives.

Organising a table tennis tournament

You have the job of organising a table tennis league.

- 7 players will take part
- All matches are singles.
- Every player has to play each of the other players once.
- There are four tables at the club.
- Games will take up to half an hour.
- The first match will start at 1:00 PM.



Plan how to organize the games, so that the tournament will take the shortest time.
Put all the information on a poster so that the players can easily understand what to do.

Fig. 6 A planning and organizing task.

Summative assessment is a standard part of teaching practice, with class tests using tasks of the same type as in the curriculum. The tasks in [Fig. 3](#) were designed for assessing ML in this way. Test scores often become part of a student's record. How this record is used depends on the school and the system. In the US, the continuing Grade Point Average is important in guiding decisions about a student's future, within and beyond the school. In the UK it plays a less important role, mainly guiding within-school decisions about future curriculum options.

Analysis. It is straightforward to design classroom assessments of mathematical literacy with high validity—provided that ML is part of the curriculum. The rich tasks and chains of reasoning that ML involves are suitable for both the normal pattern of lessons and for assessing student performance. Test results are directly interpretable in terms of the task types included; interpretation in terms of modeling skills and conceptual understanding may require the deeper probing that is characteristic of formative assessment. Objectivity is limited, particularly when teachers work alone. The same applies to reproducibility. There is no additional cost to this kind of assessment. Teacher time spent evaluating student work may provide diagnostic information that informs better-focused instruction or summative testing for accountability through the student record.

Coursework

Sometimes called *portfolio assessment*, coursework assessment involves selecting and evaluating of pieces of work produced by each student during the school term. If a portfolio includes substantial pieces of work on rich problems from real world contexts, this approach fits the assessment of mathematical literacy well.

To address the possibilities of limited objectivity and reproducibility, portfolio assessment imposes constraints that improve consistency. Detailed guidance may be given regarding the range and balance of task types to be included, and detailed guidance on the scoring. In some cases actual tasks are specified although it is common to offer a choice of tasks, particularly for the extended pieces of work. In some cases, as in the practical assessment of Fine Art, the task-set is left entirely to each student and their teacher. Scoring such tasks employs generic quality criteria.

Given the extended timescale and the resources available from the classroom and beyond, a wide variety of tasks can be included. Two ML examples illustrate this. A UK examination for age 16 included among the coursework options the task *Design a Leisure Center*, the response being an entry for the town's design competition in the form of a scale plan of the site and buildings, with explanations of the design choices made. The extended problems in the school leaving examination for the Victorian Certificate of Education (VCE) were more closely specified, as with *Oil Pipelines* in [Fig. 7](#) (There is more about this VCE examination, which achieved remarkably high validity through the variety of task types included, in the section [A spectrum of assessment models](#) in [Burkhardt, 2009](#)).

There is similar diversity in approaches to scoring coursework and in how such scoring is monitored. Scoring rubrics can be task specific and detailed, or more generic. The open task *Design a Leisure Center* asks for wide ranging thinking and for data collection regarding plausible assumptions, such as the size of areas needed for different recreation activities and how they may fit together efficiently and pleasingly. The design of scoring rubrics that are objective and reproducible for such tasks is less straightforward than for well-specified tasks like "Oil Pipelines"—but consistency can be at least comparable to that in scoring English essays. Student choice among a group of tasks offered also presents comparability challenges. Such challenges can be tackled by trialing the tasks and adjusting the rubrics. Generic rubrics that give credit for higher-level aspects of the response are sometimes used for coursework.

When high stakes are involved monitoring systems are typically employed to improve consistency of scoring. Sophisticated systems involve groups of teacher-scorers comparing samples of scored student work—first within a school and then across a groups of schools, with a chair from the exam board. This activity is also a mechanism for powerful professional development, spreading

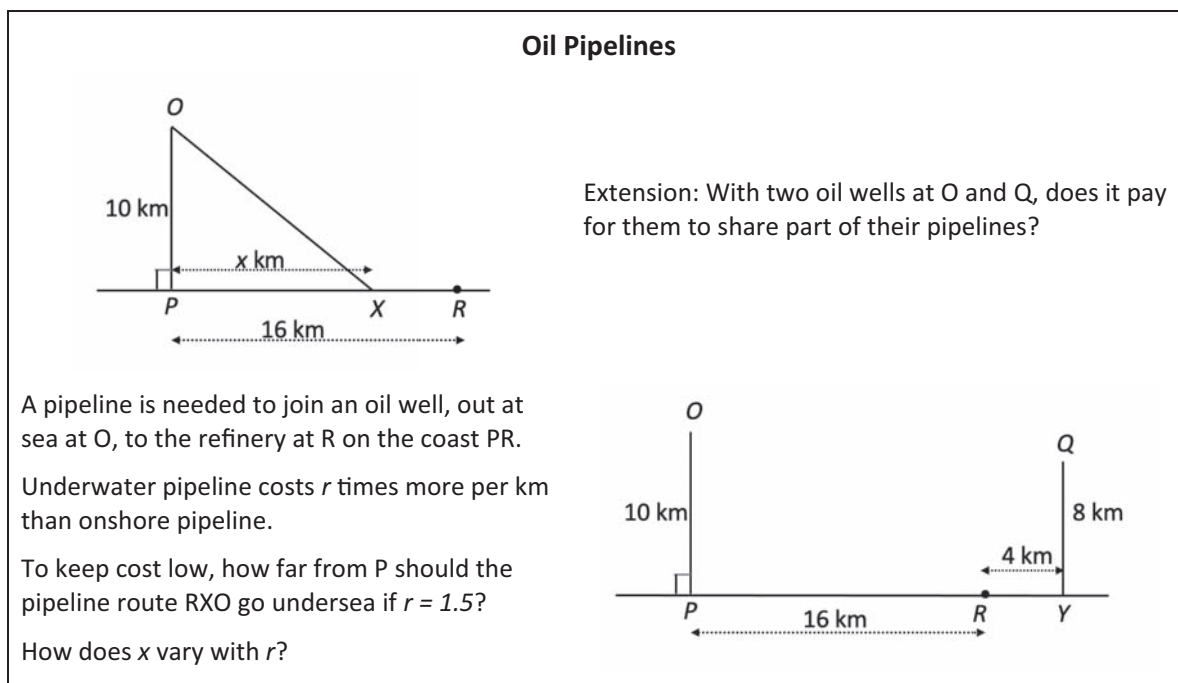


Fig. 7 A VCE problem solving investigation (modified).

good practice through discussions of examples of student work. In a simpler model, teachers score work from other classes than their own, then a lead scorer monitors samples of each scorer's work. In both cases scores for a whole group are adjusted on the basis of the sample.

Analysis The central characteristics of coursework are task timescale and circumstances of performance—not only can extended pieces of work, including projects, be included in the package assessed but students have time to check and refine their responses. This is much closer to the way adults are assessed in their life and work. Provided there is some control over the range of tasks, validity can be exceptionally high. The body of information contained in the responses of each student provides a rich, multidimensional picture of the student's capabilities, in contrast to the single number that results when student performance is boiled down to one numerical indicator for accountability purposes. Objectivity with a strong monitoring system is high, as is reproducibility because of the large body of evidence included. The cost of this approach is high both in the teacher time needed for careful scoring and in the monitoring process. However, any cost-benefit analysis should take other aspects into account: careful attention to and feedback on student work can yield significant student learning gains.

British work from the 1960s and 70s on has shown that carefully designed assessment based entirely on coursework could combine their intrinsic validity with a level of consistency comparable to timed examinations, when there is a process of consensus monitoring of teachers' scoring overseen by an exam board (Gibbons and Marshall, 2010). Some US states also developed comparable systems of "collaborative standardized assessment". However, trust in coursework remains a significant political problem. Policy makers have used the specter of cheating as a reason for reducing the role of coursework or eliminating it altogether from many high-stakes assessments. This was the fate of the expert group's recommendation for carefully structured and monitored coursework for the assessment of the National Curriculum in England—the politicians chose tests.

Timed examinations

Examinations are the typical mode of assessment in many countries, with tasks varying from short correct/incorrect items to open or "essay" questions involving more extended chains of reasoning. A few-hour timescale allows a wide variety of task types to be included, each taking the successful student anywhere from a few minutes to an hour. This approach can cover most of the kinds of performance implied by the learning goals in most subjects, including mathematical literacy. Students responses are scored centrally by teachers or others with subject expertise following a more-or-less detailed rubric that distributes credit among different aspects of a successful response. The rubric is normally specific to each task. Mathematical literacy tasks like those in Fig. 3 fit well into such examinations.

Opportunities for high-validity timed examinations are rarely realized in practice, particularly in Mathematics where expected task time tends to be much shorter than for other subjects. There have been enough successful examples of Mathematics examinations including rich tasks that require 10–20 min of solution time (see Burkhardt, 2009) to suggest that their typical absence is a systemic issue: changing high-stakes examinations causes controversy, which politicians tend to avoid. Equally, tasks that require

extended reasoning necessitate the kinds of changes in the ways that mathematics is taught and learned that were described in the section **Mathematical literacy—what is it and why is it important?**. Thus what one sees in mathematics assessments of this type is a trade-off between complexity (of tasks, supportive pedagogy, assessment mechanisms, and interpretation thereof) and tests that are easy to score and reliable—even if the reliability comes at the cost of representing true proficiency.

The variation of scores between different scorers is likely to be much greater for extended chains of reasoning, as in ML or in English, than in Mathematics. In Mathematics high inter-scorer consistency, often a source of pride, is largely the result of limiting the task types to short items with correct/incorrect answers. It is notable that the wide disparities between subjects, in both validity and accuracy, is socially accepted as part of the general practice of assessment, even though a difference of one grade in a school leaving examination can have life-affecting consequences. Hence tasks that assess mathematical literacy should not be excluded on those grounds. The task in Fig. 8 and its associated rubric provide an example of how complex reasoning can be assessed in a reasonably consistent way.

The original “Mathematics version” of this task called for identifying the student with the best average score, rendering the context just a pretext for computation. The version in Fig. 8 is much more realistic. It calls for an understanding of the context and the use of more sophisticated mathematical reasoning regarding what each measure does and doesn’t tell you.

Analysis. Specific rubrics of this kind tend to have high objectivity and reproducibility. Regarding validity, the limited time of an examination and the need to assess a reasonable sample of the curriculum limits task length compared with coursework—but a range of task types listed in the section **Mathematical literacy—what is it and why is it important?** can still be sampled. Such assessment for ML is much broader than in Mathematics, where actual examinations rarely cover even their own stated performance goals. Interpretability is limited by the common practice of reporting just a single score; “profile reporting” with several scores (e.g., for skills, reasoning, problem solving) is an improvement, but remains rare. The cost of the examination is relatively low, as is the class time required—a few hours per subject. Provided that the balance of task types reflects the learning goals for the course in a balanced way, time spent preparing for the examination is valuable learning time.

Short item multiple choice tests

This kind of test of Mathematics has long been favored in the United States for large-scale implementation, e.g., for many state-wide assessments and for pre-college “achievement” tests. Machine-scoring means the objectivity is high. Professional input comes in the design and development of the test, which is often a complex and expensive procedure ruled by psychometric considerations. The direct cost per student is low and little classroom time is needed to administer the test. However, the real cost of such exams can be

Long Jump

Our school needs to select a girl for the long jump at the regional championship.

At the school jump-off three girls have 6 jumps each. Their results, in meters, are given in the table.

Elsa	Ilse	Olga
3.25	3.55	3.67
3.84	3.99	3.78
4.1	3.61	3.92
2.95	3.97	3.62
3.66	3.69	3.85
3.86	3.59	3.73

Hans says:
“Olga has the longest average. She should go.”

Do you think Hans is right? Explain your reasoning.

Suggested rubric
7 minute task >> total 7 points,
partitioned to reflect importance

Points	Criteria
1	Recognize that this is about choice of criteria for likely success in long jump.
1+1 each up to max of 4 points	List criteria (other than average) + some explanation, e.g. • Longest jump; ultimate criterion • Average of two longest jumps; gives some weight to consistency • Average of longest four jumps; allows slips such as “no jumps”.
2	Correctly compute numbers required (-1 for each error)
7	Total

Fig. 8 Long jump—a mathematical literacy task with scoring rubric.

high because teachers typically spend several weeks of class time on otherwise unproductive test preparation. Validity is an issue because the design is constrained by the format: tasks must be short, usually taking a successful student less than a minute—which is very different from the extended chains of reasoning that Mathematical Literacy (or indeed, any substantial problem solving) typically involves.

Despite this PISA, the outstanding assessment of ML and arguably the most influential test in mathematics worldwide, largely in this format. The choice of format was driven by the constraints on any international comparison assessment—notably:

- the tasks must present an equivalent level of challenge when translated into many languages and cultures;
- consistency of scoring is essential if the politically important comparisons for which PISA is used are to be trusted. It is only possible to achieve this across multiple cultures through the use of very simple rubrics.

The quality of the PISA tests is widely respected. That raises questions of how this has been achieved and what the limitations of this model for assessing ML might be.

The design and development of these tests (see [Stacey and Turner, 2015](#)) involved combining principles developed by the PISA Mathematics Expert Group with creativity in task design and academic rigor from the item writing teams at the Australian Council for Educational Research (ACER) and in the contributing organizations around the world. These groups developed coherent units that involved a sequence of questions/items of increasing difficulty, each item with a short response. A rich collection of these units can be found as “released items” on the OECD PISA website. These include the tasks in [Figs. 9 and 10](#), which address the two key elements of mathematical literacy—using data and modeling.

Which Car? involves reasoning to find meaning in quantitative data—in this case using the three criteria in sequence to eliminate alternative cars. *Revolving Door* is a very challenging modeling task: How big can the openings be such that air can never flow freely between the entrance and the exit.

PISA contributes significantly to the assessment of ML worldwide, not least because its name, *PISA Mathematics*, opens the door to such tasks being used in regular mathematics courses and assessments. However, the limitations of PISA’s assessment format are clear. Tasks are longer than is typical in multiple-choice mathematics tests but still short. They require somewhat longer chains of reasoning but still exclude the coherent explanations of extended reasoning, characteristic of ML that other assessment modes allow. Longer tasks within the test time mean a smaller sample of each student’s performance range but coverage is achieved in the national samples by spreading a large number of tasks across different students.

WHICH CAR?

Chris has just received her car driving license and wants to buy her first car.

This table below shows the details of four cars she finds at a local car dealer.

Model:	Alpha	Bolte	Castel	Dezal
Year	2003	2000	2001	1999
Advertised price (zeds)	4800	4450	4250	3990
Distance travelled (kilometres)	105 000	115 000	128 000	109 000
Engine capacity (litres)	1.79	1.796	1.82	1.783

Chris wants a car that meets **all** of these conditions:

- The distance travelled is **not** higher than 120 000 kilometers. [1] [SEP]
- It was made in the year 2000 or a later year. [1] [SEP]
- The advertised price is **not** higher than 4500 zeds.

Q1: Which car meets Chris’s conditions? [1] [SEP]

Fig. 9 Data analysis—a PISA task.

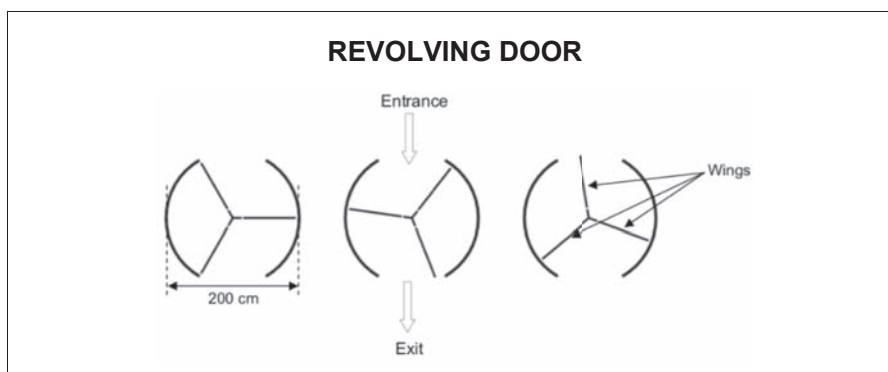


Fig. 10 Modeling with mathematics—based on a PISA task.

Given the international nature of PISA there are cultural limits on tasks that require students to bring in information that is not explicitly provided, an essential part of ML. Nonetheless, increased use of such tasks would be beneficial. That they remain rare is, in view of PISA's political significance in many countries, further evidence of the system-level difficulties of changing established high-stakes assessments.

Analysis. A student's responses to well-designed multiple-choice items, with distractors reflecting common misconceptions, can form valuable diagnostic information for teachers on conceptual understanding, but only when viewed at the item level. As assessment of a complex subject like ML (or Mathematics or English) multiple choice tests alone have low validity. These are not the performances that teachers should focus on developing in their students; the psychometric justifications for validity depend on correlations to meaningful tasks, not the tasks themselves. Objectivity is indeed high, reproducibility is moderate—and machine scoring can keep the cost low if class time spent on "test prep" is excluded. Interpretability is usually limited by the small range of task types that are included, as well as by single number reporting. PISA, however, has shown how far these limitations can be minimized.

Finally, the skill of the PISA design team and the costs of task development should not be underestimated. Finding the talent and the funding to modify PISA for purposes other than large-scale assessment would be a challenge.

Routes to better assessment

Policy makers need assessments that provide them with data to monitor the education systems for which they are responsible. We have shown that there is a great variety of ways this is done, each having a different balance of desirable features. But an assessment cannot be more reliable than it is valid; in any normal understanding of those terms, you have to assess what you want teachers to teach and students to learn. Those who choose accountability that relies primarily on multiple-choice tests are thus accountable for the damage done to the teaching and learning in their systems. In contrast, there do exist systems that provide the necessary feedback for accountability at reasonable cost and provide pressure to enhance the quality of teaching and learning. This implies giving priority to validity.

We have set out the many reasons why an accountability system that strategically aims to improve learning should be based on the use of assessment tasks that together represent a broad and balanced sample of the knowledge and performances that represent high-quality learning in each subject. We have pointed to examples of large-scale implementation of assessment systems that have met these standards. The central issue in the design of such a system is the balance between validity and objectivity, so different across the models in 4A to 4D above, taking into account the likely effect on what is taught. A combination of coursework and timed examinations ranks high on both these desirables.

The best overall approach to meeting the multiple goals of assessment for any domain as complex as Mathematical Literacy is a combination of coursework and timed examinations, as described in 4B and 4C above, each with sufficient weight to ensure it is not neglected. This approach has been successfully designed, developed and implemented in many systems over the years and in a variety of ways. The VCE model mentioned in 4C contained all these features—and research showed it led to positive changes in what was taught throughout secondary schools in Victoria.

We have noted that this approach is more expensive than multiple choice tests only if the cost in classroom time spent on otherwise unproductive "test prep" is not included. A system that includes "tests worth teaching to" makes effective use of classroom testing time.

We conclude with comments on two features that have proved important.

The centrality of teacher assessment

An element of classroom assessment by the teacher is crucial. The direct involvement of teachers in assessment enables the inclusion of important, rich kinds of performance that cannot fit within a timed examination—hence coursework. Externally provided coursework tasks and rubrics help teachers become more aware of the nuances of the assessment and the performances assessed.

This complements the element of timed examinations. Most kinds of important performances can be validly assessed through tasks that take less than half an hour, many much less than that. The more tightly controlled and familiar format of a timed examination builds society's confidence in the overall results. But even here there is advantage if teachers take some part as scorers, building knowledge of the exam and a sense of agency—it becomes “our assessment” not just one imposed from on high, in addition to serving as a form of professional development.

In summary, the detailed examination of extended pieces of student work on rich tasks, and the comparative discussions in the monitoring process, are an invaluable contribution to teacher development. The controlled environment of the timed examination builds confidence and allows many aspects of a subject to be covered.

Reporting—the value of multi-dimensional profiles

A fair and valid assessment of any complex subject, including mathematical literacy, must collect a balanced sample of the broad range of knowledge and performance skills that define the domain. Different uses of the results of the assessment will require that information to be condensed to a greater or lesser extent—in the most extreme (and also most common) case to a single number for each student.

The danger in this condensation is that a 1-number score gives a false impression, obliterating the multi-dimensional complexity of a student's abilities in different aspects of the subject. Further, it gives an impression of accuracy that is unwarranted. There is a middle ground where the profile of a student's performance is represented with a handful of scores; in Mathematics, four dimensions are widely accepted as reasonable: separate concepts and skills; extended mathematical reasoning; tackling non-routine problems within mathematics; understanding situations from the real world through modeling and data analysis—the essence of mathematical literacy. Such systems have been demonstrated to be both feasible and psychometrically defensible. Accountability systems based on assessment of this kind can enhance teaching and learning rather than, as too often now, the reverse.

Acknowledgment

The authors gratefully acknowledge the Kaye Stacey's assistance in thinking through the broad range of issues dealt with in this manuscript.

References

- Black, P.J., William, D., 1998. Assessment and classroom learning. *Assess. Educ.* 5, 7–74.
- Burkhardt, H., 1981. *The Real World and Mathematics*. Blackie, Reissued Nottingham. Shell Centre Publications, Glasgow, 2000.
- Burkhardt, H., 2009. On strategic design. *Educ. Des.* 1 (3). Retrieved from: <http://www.educationaldesigner.org/ed/volume1/issue3/article9>.
- Burkhardt, H., 2019. Improving policy and practice. *Educ. Des.* 3 (12). Retrieved from: <http://www.educationaldesigner.org/ed/volume3/issue12/article46/>.
- Burkhardt, H., Schoenfeld, A.H., 2019. Formative assessment in mathematics. In: Bennett, R., Cizek, G., Andrade, H. (Eds.), *Handbook of Formative Assessment in the Disciplines*. Routledge, New York.
- Gibbons, S., Marshall, B., 2010. Assessing English: a trial collaborative standardised marking project. *Engl. Teach. Pract. Critiq.* 9, 26–39.
- Hatano, G., Inagaki, K., 1986. Two courses of expertise. In: Stevenson, H.W., Azuma, H., Hakuta, K. (Eds.), *Child Development and Education in Japan*. W.H. Freeman/Times Books/Henry Holt & Co., New York, pp. 262–272.
- Mathematics Assessment Project, 2015. <https://www.map.mathshell.org/>.
- Schoenfeld, A.H., The Teaching for Robust Understanding Project, 2016. *The Teaching for Robust Understanding (TRU) Observation Guide for Mathematics: A Tool for Teachers, Coaches, Administrators, and Professional Learning Communities*. Graduate School of Education, University of California, Berkeley, Berkeley, CA. Retrieved December 3, 2017 from: <http://TRUframework.org>.
- Schoenfeld, A.H., 2020. Addressing horizontal and vertical gaps in educational systems. *Eur. Rev.* <https://doi.org/10.1017/S1062798720000940>.
- Stacey, K., Turner, R. (Eds.), 2015. *Assessing Mathematical Literacy: The PISA Experience*. Springer, Berlin. <https://doi.org/10.1007/978-3-319-10121-7>.
- Steen, L.A. (Ed.), 2002. *Mathematics and Democracy: The Case for Quantitative Literacy*. National Council on Education and the Disciplines, Washington, DC. <http://www.maa.org/q/mathanddemocracy.html>.
- Teaching for Robust Understanding (TRU) Framework, 2015. <https://truframework.org/>.

Assessment of collaborative problem-solving: past achievements and current challenges

Areum Han^a, Florian Krieger^b, and Samuel Greiff^a, ^a Department of Behavioral and Cognitive Sciences, University of Luxembourg, Luxembourg; and ^b Department of Rehabilitation Sciences, TU Dortmund University, Dortmund, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	234
Definitions of collaborative problem solving	235
A classification framework for collaborative problem-solving skill assessments	235
The past landscape	237
The current landscape	239
Future directions and agendas	241
References	242

Introduction

Collaborative problem-solving (ColPS)¹ skills are defined as “the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution” (OECD, 2017a, p. 6). ColPS skills are prominently featured on lists of 21st-century skills people should possess in today’s fast-changing societies (National Research Council, 2011; Swiecki et al., 2020), as the problems that we face in today’s world have become increasingly non-routine and intractable. Consequently, ColPS skills are considered critical in various dimensions of life, including the workplace, educational and “civic contexts” (Graesser et al., 2018; OECD, 2017a, p. 4; O’Neil et al., 2003; Polyak et al., 2017; Rosen and Tager, 2013; Stoeffler et al., 2020). The significance of ColPS skills even extends to our private lives, such as interactions with our families or among teammates in the sports context (Stadler et al., 2020).

Due to its importance in both everyday life and the workplace, it comes as no surprise that a great deal of attention is placed on assessing ColPS. For instance, it has been included recently in large-scale assessment formats of significance to a broad community of both researchers and policymakers, such as the Program for International Student Assessment (PISA) convened by the Organization for Economic Co-operation and Development (OECD) and the Assessment and Teaching 21st Century Skills (ATC21S).

However, the assessment of ColPS skills does not have a short pedigree. It has evolved at different levels such as measurements at the group or individual level and across various disciplines such as cognitive psychology, business, computer science, education, and other (sub-)disciplines (Andrews-Todd and Forsyth, 2020; Care et al., 2016; Graesser et al., 2018; Oliveri et al., 2017). Since ColPS is related to different disciplines, each with different perspectives and methodological approaches, it is no surprise that the history of ColPS assessment is quite diverse and linked to numerous research traditions.

In a nutshell, the history of ColPS assessment can be described on different levels. That is, there are multiple theoretical perspectives underlying assessments. In addition, various methods have been utilized to measure ColPS, including scoring strategies for processes and outcomes (e.g., Hao et al., 2015; National Research Council, 2011; OECD, 2017a). More specifically, ColPS assessments can be divided into several categories, such as measurements at the group or individual level and at a large or small scale.

Consequently, this book chapter’s overarching aim is to (1) introduce some significant milestones in the development of ColPS assessments, (2) introduce a taxonomy of different assessment approaches, and (3) provide an outlook on the future development (or development needs) of ColPS assessments. Specifically, examining and articulating the levels mentioned above can be beneficial for the further development of assessments, as lessons from the past and the present can serve as a basis for future assessments.

To achieve these aims, in the following sections, we will first discuss various definitions of ColPS by illustrating the similarities and differences between different definitions by means of examples (with a focus on the educational context). We then propose a generic classification framework, which serves as a foundation for classifying different assessment approaches based on methodological aspects and implications. Next, we will illustrate several assessment approaches via examples from the past and the current landscape and connect them with the main aspects of the proposed framework. Finally, we will propose several agenda points to be considered in the future from the perspective of educational psychology and education.

¹Please note that collaborative problem-solving is also abbreviated with acronyms such as CPS in some literature (e.g., Care et al., 2016; Sun et al., 2020).

Definitions of collaborative problem solving

As previously mentioned, there are considerable differences between definitions and perspectives on ColPS. There is still no widely agreed-upon definition or perspective on ColPS (Greiff et al., 2013; Stadler et al., 2020). Many studies intermingle the relevant terms, such as group decision making, teamwork, collaborative learning, and collaborative problem-solving (e.g., Graesser et al., 2018; Kaendler et al., 2015; Oliveri et al., 2017).

On the most basic level, however, existing perspectives and definitions of ColPS share some key similarities. One such similarity concerns the composition of ColPS as a multitude of interrelated (sub-)constructs (Andrews-Todd and Forsyth, 2020; Care et al., 2016; Ercikan and Oliveri, 2016; Hao et al., 2019; von Davier et al., 2017). Here, a construct refers to “the concept or characteristic that a test is designed to measure” (American Educational Research Association et al., 1999, p. 173, as cited in Birk, 2013, p. 4; Care and Griffin, 2014). Another distinct commonality reflected in all definitions of ColPS is that both cognitive aspects (i.e., task work) and social aspects (i.e., teamwork) are required (Avry et al., 2020; Graesser et al., 2018).

However, when taking a more detailed view, differences emerge, particularly with respect to the different facets. That is, the specific sub-skills identified and the frameworks into which they are organized vary considerably (Graesser et al., 2018). For example, the OECD constructed a matrix of collaboration skills and individual problem-solving skills for PISA 2015 (Care and Griffin, 2014; OECD, 2017a). By contrast, the framework for the ATC21S, another large-scale ColPS assessment project, divided each skill into two components, cognitive and social aspects (Care and Griffin, 2014). Meanwhile, von Davier and Halpin (2013) interpret collaboration from a *statistical* viewpoint as “a function of the departure from statistical independence among the activities of two or more individuals” (von Davier et al., 2017, p. 633). Thus, while the PISA and ATC21S approaches highlight “the *types* of interdependence,” this latter approach highlights “the *degree* of interdependence” (von Davier et al., 2017, p. 634). These examples illustrate that various ColPS assessments are built on divergent perspectives.

This diversity of definitions is natural because ColPS is rooted in multiple disciplines and requires cross-functional measurement approaches. Accordingly, it is impossible to settle upon a single correct definition; we should acknowledge the construct’s heterogeneity. In other words, for optimal assessment, we should not distinguish between correct and incorrect definitions, but rather articulate ColPS in a fine-grained way that enjoys broad consent. This is because assessment results stemmed from constructs are in turn likely to derive valid inferences and consequences (see Messick, 1994; Mislevy et al., 2003; Schmidgall et al., 2019).

However, the lack of such a fine-grained and widely agreed-upon definition has not hindered the development of numerous assessment instruments for ColPS. Table 1 provides an overview of examples of CPS assessment, mainly in educational settings. The instruments differ from one another not only in terms of their perspectives but also their methods. Beginning in the next section, we discuss such methodological aspects, another diverse element of ColPS assessments, by means of a classification framework.

A classification framework for collaborative problem-solving skill assessments

To discuss different methodological approaches to ColPS assessment and present an overarching framework for discussion, this chapter draws on classifications by Bunderson et al. (1988), Cope and Kalantzis (2016), and DiCerbo and Behrens (2012, as cited in Shute et al., 2016). Table 2 summarizes these classifications, based mainly on Cope and Kalantzis (2016) and Dicerbo and Behrens (2012) and extended with relevant characteristics from Bunderson et al. (1988). Although these classifications are broad in nature, they can be adapted to ColPS assessments. As seen in Table 2, the classifications focus on methods, particularly with respect to the gradual integration of computer technology. Thus, there is a need to create a custom classification describing the consolidation of computer technology in *ColPS assessments* based on existing generic frameworks. These frameworks can help to categorize progress in the development of ColPS assessments and facilitate improvements to the assessments’ methods or modes.

The framework in Table 2 consists of five *developmental* levels: Level 0 to Level 4. The first level (i.e., Level 0) encompasses traditional assessments with non-computerized tasks and modes. The second level (i.e., Level 1) is characterized by computerized paper-and-pencil modes. The third level (i.e., Level 2) provides test-takers with simulation-based tasks. The fourth level (i.e., Level 3) encompasses natural digital activities. The last level (i.e., Level 4) is characterized by collecting information from different natural digital activities. This hierarchy is utilized in this chapter to categorize and evaluate different assessments. It also highlights a potential point for development, even though not all assessments need to be at Level 4.

Elaborating each level further, Level 0 represents traditional formats of assessments such as paper-and-pencil questionnaires, traditional situational judgment tests, evaluators’ observations, or narratives (e.g., letters of recommendation). That is, assessments at Level 0 are implemented as non-computerized tasks, which mostly consist of discrete items that test only one element at a time so induce limited responses by test-takers. Moreover, assessments are not integrated with other learning activities but are carried out separately.

Next, Level 1 (i.e., computerized paper-and-pencil tests) refers to the simplest transfer of assessments into the “technologically rich environment” (Shute et al., 2016, p. 34). At this level, it is most common for traditional item types (e.g., “discrete, independent items”) to be incorporated into computer environments, although they can have more than one of the following elements: “innovative item types, automated scoring, adaptivity, and/or advanced feedback” (DiCerbo and Behrens, 2012, p. 6, as cited in

Table 1 Example assessments for ColPS and its sub-skills

Example	Field	Inferential level	Method	Description
Behavioral Observation Scales (BOS) (Taggar and Brown, 2001)	Education	Individual	Observation and ratings (likert-type scales)	Anonymous peer assessment of team members' behaviors (e.g., in terms of synthesizing ideas within the team)
Team Problem Solving Assessment Tool (TPSAT) (Rotondi, 1999)	Health and medical workforce	Group	Observation and Ratings	While test-takers participate in simulated performance tasks, an evaluator (potentially their manager) assigns scores between 0 and 100 for each variable
ETS Personal Potential Index (as cited in Oliveri et al., 2017)	Education	Individual	Observation and ratings	- An evaluator assesses students' readiness for graduate school - Parts of constructs overlap with ColPS skills (e.g., teamwork, critical thinking, and problem-solving)
Predictor Situational Judgment Test (PSJT) (Waugh and Russell, 2005, as cited in Oliveri et al., 2017)	Military workforce	Individual	Situational judgment test (paper-and-pencil)	- Written scenarios consisting of 20 items - Utilized for improving the selection and classification of personnel in the United States military - Parts of constructs overlap with ColPS skills (e.g., teamwork, relating to and supporting peers)
Vanderbilt University's Jasper Woodbury Problem Solving Series (Cognition and Technology Group at Vanderbilt, 1992)	Education	Group	Video-based assessment	- Embedded assessment in various content areas (i.e., all the information necessary to solve the problems is included in the video.) - Students need to work together collaboratively (in pairs)
ATC21S (e.g., Griffin et al., 2012)	Education	Individual	Computer-based assessment	- Domain-general - Simulated tasks with human-to-human approach
PISA 2015 (OECD, 2017a)	Education	Individual	Computer-based assessment	- Domain-general - Simulated tasks with human-to-agent approach
ETS Tetralogue (Hao et al., 2015; Scoular and Care, 2019)	Education	Group and individual	Computer-based assessment	- Constructed from a natural sciences perspective (i.e., online simulation in the science domain is provided) - Used in schools for research purposes - Hybrid approach (human-to-human and human-to-agent)

Note. Source: Kyllonen, 2012; OECD, 2017a; Oliveri et al., 2017; Particularly adapted from "A literature review on collaborative problem solving for college and workforce readiness," by M. E. Oliveri, R. Lawless, and H. Molloy, 2017, ETS Research Report Series, 1, pp. 12–16 (<https://doi.org/10.1002/ets2.12133>). Copyright 2017 by Educational Testing Service.

Shute et al., 2016, p. 43). In this regard, Level 1 can be seen as part of the "computerized testing (CT)" or "computerized adaptive testing (CAT)" generations identified by Bunderson et al. (1988, p. 4). According to Bunderson et al. (1988), the CT generation began when traditional tests were converted to computer-based but *non-adaptive* tests. In contrast, the CAT generation led to the emergence of *adaptive* tests that considered students' performance/responses to prior tasks when presenting new ones. In other words, assessments on Level 1 can belong to either the CT or the CAT generation depending on their degree of adaptivity.

Level 2 goes beyond Level 1 by providing more simulation-based tasks (Shute et al., 2016). Test-takers' actions within tasks can be observed via technology, but instruction and assessment are still separated and remain distinguishable from each other (DiCerbo and Behrens, 2012, as cited in Shute et al., 2016). For this reason, assessments at this level still belong to the CAT generation as defined by Bunderson et al. (1988), in which the students' previous responses and the measurement model determine item selection (Quellmalz and Pellegrino, 2009). Moreover, in assessments at this level, item presentation times and test length can change depending on test-takers' previous response times or performance; thus, for assessments at this level, a huge calibrated item bank should be constructed to estimate test-takers' skills (Bunderson et al., 1988).

At Level 3, the line between instruction and assessment is blurred (Shute et al., 2016). In other words, instruction is embedded in the assessment, with both becoming technology-rich activities (Shute et al., 2016). Consequently, the "continuous measurement (CM)" generation begins at this level (see also Bunderson et al., 1988, p. 4). This means that changes or trajectories in students'

Table 2 Classification framework of assessments

	Level 0	Level 1	Level 2	Level 3	Level 4
General description	Non-computerized tasks and modes	Computerized paper-and-pencil CT → CAT generation	Integrated context and tasks CAT generation (continuous)	Lifelike activities with evaluative features in digital environments CM generation	Accrual of information from various lifelike activities in digital environments IM generation
Tasks	Discrete items (i.e., test only one element at a time)	Discrete items	Complex tasks/problems	Complex tasks/problems	Variety of tasks/problems
Integration with learning activities	Low	Low	Low	Moderate	High
Variety of activities/responses	Low	Low	Moderate	Moderate	High
Data type	Traditional assessments (no machine involvement)	Computerized assessments	Structured, embedded data (e.g., semantic tagging and annotation) Unstructured, incidental data (e.g., keystroke patterns, social interaction patterns)		
Mode of data collection	Paper-and-pencil questionnaire, traditional situational judgment test, observation, narratives (e.g., recommendation letters) etc.	Computerized adaptive testing	Natural language processing Procedure-defined processes Argument-defined processes Machine-learning processes Incidental “data exhaust” Dedicated devices for collecting unstructured data		

Note. CT: Computerized Testing; CAT: Computerized Adaptive Testing; CM: Continuous Measurement; IM: Intelligent Measurement.

Source: Bunderson et al., 1988; Cope & Kalantzis, 2016; DiCerbo & Behrens, 2012; Shute et al., 2016; Particularly adapted from “Big data comes to school: Implications for learning, assessment, and research,” by B. Cope, and M. Kalantzis, 2016, AERA Open, 2(2), p. 3 (<https://doi.org/10.1177/2332858416641907>). CC BY 3.0.; and “Advances in the science of assessment,” by V. J. Shute, J. P. Leighton, E. E. Jang, and M-W. Chu, 2016, Educational Assessment, 21(1), p. 42 (<https://doi.org/10.1080/10627197.2015.1127752>). Copyright 2016 by Taylor & Francis.

proficiency can be observed and estimated in a continuous and unobtrusive way. For example, students’ strategies (e.g., the methods they use to approach or avoid tasks) can be automatically and instantaneously tested while they complete learning tasks. Moreover, individual learner profiles (e.g., abilities or preferences) can be constructed. Thus, assessments at this level (1) can be more dynamic than the previous levels, which might necessitate more complex psychometric procedures, and (2) focus on “individual purposes, not institutional purposes” (Bunderson et al., 1988, p. 79).

Lastly, at Level 4, assessment data can be accumulated via numerous activities in technology-rich environments that are not limited to only a few settings (Shute et al., 2016). Thus, more knowledge-based computing (i.e., artificial intelligence) can be applied to deal with the mass of information (Bunderson et al., 1988). That is to say, the fourth generation, “intelligent measurement (IM)” would begin at this level with the application of more advanced technologies (Bunderson et al., 1988, p. 4). By using artificial intelligence, more sophisticated individual response processes and trajectories can be automatically measured and interpreted. Furthermore, individualized recommendations can be provided to students and teachers “during continuous measurement,” although scrutiny by professionals should be involved in the case of recommendations with far-reaching consequences (Bunderson et al., 1988, p. 125). Professional guidance may also be crucial because the generated advice can be interpreted or prescribed in several ways.

As will be described in the following sections, an overall trend of ColPS assessments moving from Level 1 to the more advanced levels of this framework can arguably be observed. However, assessments at the lowest level (i.e., Level 0; traditional formats) are still used today as well.

The past landscape

In this section, we discuss ColPS assessments that have been employed in the past based on the framework in Table 2. To evaluate and categorize these assessments, we apply the levels and definitions described in Table 2. In addition, we explore the advantages and disadvantages of several assessment formats, which could be cornerstones for developing optimal assessments in the future (i.e., by helping to identify what to strengthen and what to keep).

As explained in Level 0 of Table 2, most past ColPS assessments were implemented in the non-computerized mode (Dowell et al., 2020; O’Neil et al., 2003; National Research Council, 2011). They were used for selection (e.g., to measure readiness for the workforce/educational institutions), education, and research-related purposes. Specifically for the field of education, several assessment tools were designed and utilized for purposes such as monitoring the development of students’ skills (Kyllonen, 2012). Some examples of assessment formats at Level 0 are mentioned in the *mode of data collection* section of Table 2 and outlined in greater detail below.

- *Self-ratings* (i.e., questionnaire or self-assessment): self-ratings are a popular way to measure inter- and intrapersonal skills, which are components of ColPS skills (Kyllonen, 2012). This assessment format often applies Likert scales or forced choices and asks test-takers in a questionnaire to what extent (or whether or not) they agree with statements about preferences, beliefs, tendencies, and behaviors (Kyllonen, 2012; Oliveri et al., 2017). Some example questions for self-ratings would be “I like to work with people” (Kyllonen, 2012, p. 10), “I helped develop a plan to solve the problem” (Forsyth et al., 2020, p. 404), or “I share ideas with others to accomplish a task” (Stadler et al., 2020, p. 5).
- *Situational Judgment Test*: This type of test confronts test-takers with particular ColPS scenarios (via text or video) and asks them to decide how to react out of a set of plausible options or to rate how appropriate each option is (Oliveri et al., 2017). An example task would be:

Suppose that you find yourself in an argument with several co-workers about who should do a very disagreeable but routine task. Which of the following would likely be the most effective way to resolve this situation?

A. Have your supervisor decide, because this would avoid any personal bias.

B. Arrange for a rotating schedule so everyone shares the chore.

C. Let the workers who show up earliest choose on a first-come, first-served basis.

D. Randomly assign a person to do the task and don't change it.

Stevens and Campion (1999, pp. 225–226).

- *Observation* (ratings by peers/evaluators): Ratings by others are another commonly used assessment format (Kyllonen, 2012). In this format, test-takers' ColPS subskills during real or simulated tasks are assessed based on behavior checklists (Oliveri et al., 2017). That is, external observers note (1) the emergence of (un)desirable behaviors during the task, (2) the quality of behaviors, and (3) their frequencies (Oliveri et al., 2017). For example, evaluators can select a score between 0 and 100 regarding “the process a team uses to explore the problem and to develop a solution” on a simple rubric (Rotondi, 1999, p. 209; see Table 3).
- *Narratives based on biodata*: Narratives such as personal statements or recommendation letters can be utilized to measure a collection of previous life experiences (i.e., biodata), for example, social interaction experiences (Kyllonen, 2012; National Research Council, 2011). Alongside other data, including but not limited to standardized test scores, such narratives can be a tool to assess students' proficiency with respect to particular skills (Kyllonen, 2012; National Research Council, 2011). For example, applicants can be asked to describe a group problem-solving or teamwork experience from their life in personal statements or employment interviews. In such cases, applicants present biodata in order to demonstrate their skills, such as how they solved problems with group members, what difficulties arose, or what the final result was.

The four formats described above have both advantages and disadvantages. As previously mentioned, familiarizing ourselves with these advantages and disadvantages is important, as it can lead us to consider points to be kept or strengthened in future assessments.

Firstly, one advantage of self-assessments is that they are cost-effective and convenient to administrate (Oliveri et al., 2017; Zhuang et al., 2008). Nevertheless, response style bias or reference group bias may influence self-reports (Herde et al., 2019; Oliveri et al., 2017). Also, it is impossible to objectively measure *interactions* between problem-solvers or between tasks and problem-solvers during ColPS with such methods.

Conversely, third-party observation (by peers or evaluators) frees the individual from evaluating him/herself. However, other limitations exist, such as raters' biases and low reliability between different raters. Raters, even when well-trained, are unable to observe all of test-takers' behaviors during assessment. For these reasons, questions remain concerning the objectivity of ColPS evaluations using this method (see Oliveri et al., 2017).

Situational judgment tests can make manifest judgment processes that are normally too subtle to be detected (Oliveri et al., 2017). However, only “procedural knowledge (knowing what to do and how to do)” about effective behavior in the situation can be captured with close-ended response formats, once again leaving questions about the objectivity of the evaluation unanswered (Herde et al., 2019, p. 66; Lievens and Motowidlo, 2016).

Finally, narratives based on biodata can provide additional evidence regarding applicants' skills in the personnel selection context, for example (Kyllonen, 2012). On the other hand, distinguishing skills, knowledge, and cognitive processes is not easy. Also, such biodata may lead to inaccurate predictions about how the applicant will react in future situations, leaving much room for speculation about the real impact of such assessments in terms of criterion validity (National Research Council, 2011).

Table 3 Selected elements of an example simple rubric for “the process a team uses to explore the problem and to develop a solution” (Rotondi, 1999, p. 209).

Scale (from 0 to 100)	
100	The team will break the problem into manageable parts and thoroughly explore each part, in an effort to identify a customer’s particular needs, and the “root causes” of problems. Where appropriate, data will be used to pinpoint problem causes. The team will assess potential problems resulting from interaction of the parts with the whole. A solution will then be developed to address the identified needs and causes of each part of the problem.
40	The team will have a general discussion about the problem. The discussion will jump with no clear plan from one part of the problem to another, and no part will be explored in great depth. Some ideas to solve the problem will be briefly discussed. Finally there will be some debate about how a “standard” approach might be modified to solve this problem.
0	Members will throw out solutions and the team will react to them. The team will not seriously attempt to identify and discuss the possible components of the problem in order to understand it better. The team will basically jump from the problem to a solution. The problem will be solved by using a standard and obvious approach.

Note. Adapted from “Assessing a team’s problem solving ability: evaluation of the Team Problem Solving Assessment Tool (TPSAT),” by A. J. Rotondi, 1999, *Health Care Management Science*, 2, pp. 209–210 (<https://doi.org/10.1023/A:1019004332123>). Copyright 1969 by Kluwer Academic Publishers.

In sum, these traditional methods were (and sometimes still are) the first suitable tools for assessing ColPS for certain purposes. They tend to give a rough estimation of skills *associated* with ColPS, which might have some advantages for both researchers and practitioners, such as generating hypotheses or explanatory analyses. However, these tools tend to neglect the cognitive *problem-solving* part of ColPS. Also, their ability to validly assess real interactions during ColPS is limited due to their methodological pitfalls, such as subjective understandings of scales, as described previously. Hence, in order to appropriately describe and analyze ColPS in its full complexity, other tools need to be applied (Andrews-Todd and Forsyth, 2020; Care and Griffin, 2014; Herde et al., 2019; Quellmalz and Pellegrino, 2009; Yuan et al., 2019). The merely descriptive data on ColPS offered by traditional measures—such as self-ratings, situational judgment, observation, and narratives based on biodata—might have created some reluctance and failure to include 21st-century skills (including ColPS) as an educational objective before the development of more creative and innovative solutions (Kyllonen, 2012).

The current landscape

Leaving the issues associated with assessing ColPS skills with traditional methods behind, technological advances opened new doors for measuring various learner characteristics, including ColPS (Shute et al., 2016). As part of an innovative wave of new technologies, a series of computer-based assessments separated from instruction were developed, which tend to fall into Level 2 or the middle of Level 1 in Table 2 (i.e., computer-based simulations or computerized paper-and-pencil modes).

By implementing simulations with realistic problem scenarios and other computer-based environments, ColPS assessments can capture more complexities of the *interactions* between different individuals as well as between individuals and the computer interface in the form of various types of data. The data used for this purpose include “access-related, response-related and process-related” log data (Kroehne and Goldhammer, 2018, p. 528; O’Neil et al., 2003). Accordingly, current assessments offer measurement opportunities that incorporate social and cognitive aspects (O’Neil et al., 2003).

Broadly speaking, interactions, a core target of such measures, can be divided into two categories in terms of group composition: human-to-agent and human-to-human. Current approaches usually employ one of these two approaches in that they either initiate interactions between a group of human individuals or use computers to simulate some group members. Specifically, the distinction between these two approaches hinges on whether test-takers complete the ColPS assessment in collaboration with pre-programmed computer agents with various characteristics or with real humans such as peers (Rosen, 2015).

Several tools have been developed to measure ColPS using either or both of these approaches. The most prominent implementations were probably those used in ATC21S and PISA 2015 (Hao et al., 2019). Both assessments were interested in domain-general individual ColPS skills and measured them in computer-based simulations (De Boeck and Scalise, 2019; Graesser et al., 2018; Hao et al., 2019; Scoular and Care, 2019).

However, there were also some differences between ATC21S and PISA 2015. One distinctive difference concerns the assessment approach, as stated above (Scoular and Care, 2019). In PISA 2015, a human-to-agent approach was employed, while ATC21S adopted a human-to-human approach (Scoular and Care, 2019). For instance, in the PISA 2015 task *Xandar*, presented in Fig. 1, each student is required to interact with two computer agents in the collaborative setting, such as by proposing a common strategy for solving problems (e.g., “We can answer more questions if we divide them among us,” OECD, 2016, p. 5). As seen in the figure, participants can select from a limited set of pre-defined responses. In contrast, in the example



Fig. 1 User interface for an example unit in the PISA 2015 ColPS assessment, Xandar (OECD, 2016). Note. From Description of the released unit from 2015 PISA collaborative problem-solving assessment, collaborative problem-solving skills, and proficiency levels (p. 4), by OECD, 2016 (<https://www.oecd.org/pisa/test/CPS-Xandar-scoring-guide.pdf>). Copyright 2016 by OECD.

ATC21S task *olive oil*, two students should collaboratively identify and decide on problem-solving tools and processes to achieve the goal of “fill[ing] student B’s jar with 4 L of olive oil,” with each student having different information, resources, and instructions (Care et al., 2016, p. 257). In this task, students’ chats and actions are neither limited nor pre-defined, unlike in the PISA 2015 tasks.

Due to their different approaches (i.e., human-to-agent vs. human-to-human), each assessment had advantages and disadvantages. For example, because it analyzed *real human interactions*, ATC21S has more face and ecological validity, but is more laborious, complex, and time-consuming than PISA 2015 (Andrews-Todd and Forsyth, 2020; Graesser et al., 2018; Krkovic et al., 2018). In contrast, by employing the *human-to-agent* approach, PISA 2015 has higher internal validity, as it was able to construct more standardized problem spaces, which curtail several biases and confounding variables (Graesser et al., 2018). However, due to the limited conversation possibilities, interactions in PISA 2015 are somewhat removed from those in real-life collaboration (Herborn et al., 2020).

After ATC21S and PISA 2015, numerous ColPS assessments inspired by the two assessments, have been conducted by various institutions and researchers. Some trends can be observed in the current landscape of ColPS assessments. Firstly, although such assessments are often implemented at a large scale, only subsets of skills have been focused on due to the multidimensionality and complexity of ColPS. That is, some researchers admit to these limitations and narrow the scope of assessment. For example, the Educational Testing Service (ETS) developed its own assessment tool, ECSAP, and tried to highlight subsets of constructs by studying communication via text chats during collaboration (Hao et al., 2019).

In contrast, other assessments are implemented on a small scale but with more sophisticated analytics, such as machine learning techniques or state-of-the-art devices (Hao et al., 2019). The types of data on collaborative problem-solvers have become increasingly multifaceted, which provides meaningful “evidence of processes such as problem-solving sequences and strategy use” (e.g., time allocation or number of attempts; Quellmalz and Pellegrino, 2009, p. 77). The collected data is analyzed with algorithms to identify characteristics such as different patterns according to expertise (Quellmalz and Pellegrino, 2009). For instance, Polyak et al. (2017) measured individual ColPS ($n = 159$) by applying computational psychometrics—“a blend of data-driven computer science methods, stochastic theory, and theory-driven psychometrics in order to measure latent abilities in real-time” (p. 3)—to data such as log files of “conversation flow” with a virtual agent (p. 5) collected from the tool, *Circuit Runner* of ACT, Inc., an American assessment organization. Recently, Swiecki et al. (2020) proposed the novel method of epistemic network analysis (i.e., “a set of techniques that identifies and measures connections among elements in coded data and represents them in dynamic network models”; Shaffer et al., 2016, pp. 9–10) to measure how an individual’s contribution is related to those of their teammates within a given timeframe and how different it is compared to those of others.

Interestingly, these sophisticated analytics have also been applied to group assessments. Although slightly beyond the scope of the *computerized* generations described in Table 2, it is worth noting here that Fiore and Kapalo (2017) identify three levels of collaborative team assessments (i.e., including individuals and groups) using cutting-edge technologies. At the first level, the micro-level, individual collaboration can be measured through modern neurobiological and physiological technologies such as looking for “neural synchrony” (p. 54) in electroencephalograms. At the second level, the meso-level, interactions between bodies—including “mediating artifacts,” “movements and non-verbal behaviors” (p. 53)—can be employed to study complex constructs (e.g., shared awareness) utilizing technologies such as sensors or “sociometric badges” (p. 4). At the third level, the macro-level, inter-team or inter-network collaboration can be investigated over time using network science and social network analysis. Although the three levels of assessments suggested by Fiore and Kapalo (2017) are not the main focus of this book chapter and beyond the scope of computerized assessments, they can provide us with additional food for thought concerning the wider scope of technologies (i.e., not only computers but also various evolving devices and techniques) for future research.

Alongside these (at the moment still rather) philosophical considerations, the endeavor to refine assessments has continued. For example, educational researchers are illuminating complex constructs empirically or devising scoring methods for group/individual assessments beyond traditional test theories. For example, Andrews-Todd and Forsyth (2020) outline an ontology of ColPS constructs composed of multiple layers of concepts, behaviors, or strategies that identifies a hierarchy of skills (i.e., high- or low-level) while specifying how evidence for them can be collected from log data. In doing so, the ontology identifies explicit evidence that can be used to infer students’ skills (Graesser et al., 2018). In terms of scoring, von Davier and Halpin (2013) investigate several statistical models to measure individual ColPS with process data, such as Bayesian knowledge tracing or hidden Markov models, and introduce an alternative model for measuring individual cognitive skills during ColPS. They claim that their novel model is distinct in that individual latent ability, individual actions, and entire process trajectories during ColPS are all conditions of the model.

It might be too early to claim to have provided a full account of these current streams in ColPS assessment, as most approaches are still in their infancy. However, some consequences of the assessments employed in ATC21S and PISA 2015 can already be identified. PISA 2015 led to unexpected consequences due to innate characteristics such as the use of a standardized test across nations and the publication of results based on a ranking system. For instance, in contrast to the original intention, few educational reforms to cultivate students’ ColPS skills took place following PISA 2015, particularly in high-ranking countries (i.e., off-label use of the assessment, Zumbo, 2015). Although the ATC21S project was launched to offer feedback to stakeholders (i.e., not only students and teachers but also schools and systems; Griffin et al., 2012), the results may also have somewhat deviated from the original intention in practice. Poon et al. (2015, p. 204) report the following vignettes from a teacher and a student participating in the project in Singapore:

I do not see any formative feedback given to students in the task. Students are not aware of whether they got the answer right or wrong, and how do they improve based on the mistakes they have made in solving the questions? (Teacher).

I received no feedback (Student).

Nevertheless, more empirical evidence is needed in order to evaluate the impact of current ColPS assessments on educational policies and curricula. What can be highlighted is that despite several limitations and unintended repercussions, ATC21S and PISA 2015 made laudable, ground-breaking progress and set the foundation for subsequent research on such assessments (Hao et al., 2019; Poon et al., 2015).

Future directions and agendas

The importance of ColPS in the 21st century is undisputed. Despite the introduction of new technologies and a growing body of research on ColPS using a variety of tools, there remain open questions to be tackled in order to further develop ColPS assessments toward or beyond Levels 3 and 4 (i.e., amassing information from various lifelike activities in digital environments; see Table 2). In other words, this will allow evidence on ColPS to be collected in manifold settings simultaneously and instantaneously integrated with learning activities. Moreover, the collected data will be analyzed more elaborately with the help of innovative technologies.

One way to approach this goal is cooperation on an interdisciplinary level. The field of ColPS assessment can benefit from research from multiple perspectives in the form of great insights on multilevel analysis and a deep understanding of such assessments (Fiore and Kapalo, 2017). For instance, although current artificial intelligence technologies have limitations for studying

human-to-human collaboration, it is recommended that psychologists work with the developers of such technologies (Graesser et al., 2018; Hao et al., 2019). Such collaboration will assist in the automatic, multilevel analysis of larger samples while ensuring measurement validity (Graesser et al., 2018). It can also result in empirical studies in the human-to-human as well as human-to-agent modes.²

Another way of advancing the field is to continue revamping and optimizing existing assessments in sophisticated ways (Krkvic et al., 2018). First, appropriate methods such as Bayesian logic should be explored more deeply to solidify “evidentiary arguments” such as those based on the evidence-centered design (ECD) framework (Mislevy et al., 2003, p. i; Shute et al., 2016). The ECD framework is a well-known logical approach to constructing a test with valid reasoning while eliciting the target assessment constructs, obtaining evidence from students’ behaviors, and creating tasks to elicit such evidence (Mislevy et al., 2003; Shute et al., 2016; Zieky, 2014). In addition to studying ColPS constructs, the effectiveness of assessment task design can also be examined (Lin et al., 2015). Specifically, tasks should be constructed such that they optimally elicit the target constructs and cover the levels of proficiency defined in the ECD framework. A study Child and Shaw (2016) fleshes out this point by identifying five task design criteria for ColPS assessments: (1) sufficient complexity, (2) ill-structured tasks, (3) use of technologies to facilitate the ColPS process, (4) dynamics (including negotiation) among team members, and (5) team members’ motivation for collaboration. An additional suggestion is (6) use of technology to *explore* the ColPS process. For instance, strategic behaviors (e.g., how to share information) can be highlighted through the pattern of turn-taking among problem-solvers (Krkvic et al., 2018). In addition, tasks are authentic when they involve “the challenges and standards of [real-world] performances” (Wiggins, 1989, p. 703) and “the ways in which knowledge and skills are used in real-world contexts” (Frederiksen and Collins, 1989, p. 30) despite the underrepresentation of complex “functioning” in the real-world ColPS situation (Messick, 1994, p. 17). What and how to score such assessments should also be studied more extensively, particularly given the quantity and variability of mass data collected by technologies, which includes final outputs as well as processes (Scoular et al., 2020).

Furthermore, further empirical studies should be conducted to disentangle social, cognitive, and metacognitive aspects of the ColPS framework or investigate influencing factors. This idea is associated with building a common understanding of ColPS and consequently obtaining more valid results from data collected in multiple settings. In addition, such research can influence policy-makers to identify ways to facilitate students’ education based on the results of assessments embedded within learning activities. For instance, the PISA 2015 ColPS assessment results indicate that girls tend to outperform boys, which is influenced by disparities in personality traits (i.e., “more agreeable and more conscientious”; OECD, 2017b, p. 93). As this case demonstrates, discerning factors that influence ColPS assessment scores can not only increase flexibility in enhancing students’ ColPS in education contexts but can also lead to the construction of more valid assessments.

Finally, there is a need to increase the level of acceptance and understanding of ColPS assessments in educational settings. If teachers, students, and other stakeholders do not have a good understanding of what such assessments measure and how they can be interpreted, even good assessment tools cannot be implemented sustainably and integrated into learning. According to Poon et al. (2015), Singaporean teachers who participated in ATC21S had trouble with “concepts in the new teaching and assessment paradigm—including the introduction of ambiguity into assessment tasks, tracking dynamic behaviors in collaborative settings, and the debate over content-rich and content-free assessment of 21st century competencies” (p. 199). To alleviate these difficulties, professional development opportunities for teachers and relevant stakeholders on ColPS assessments should be promoted more strongly, and students should be given sufficient opportunities to familiarize themselves with the new assessments.

The journey to more advanced levels of assessment might be quite arduous. However, concrete plans to cultivate students’ ColPS can only be studied once a thorough understanding of ColPS skills, processes, performance levels, and measurement has been achieved (Krkvic et al., 2018). Despite the already observable shift away from non-computerized assessment, considerable work still needs to be done in order to reach this goal. Once a more mature level of understanding based on enhanced assessments has been achieved, not only national policies but also personalized learning strategies or interventions to foster ColPS can be crafted based on a stronger evidence base.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999. Standards for Educational and Psychological Testing. AERA, Washington, DC.
- Andrews-Todd, J., Forsyth, C.M., 2020. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Comput. Hum. Behav.* 104, 1–12. <https://doi.org/10.1016/j.chb.2018.10.025>.
- Avry, S., Molinari, G., Bétrancourt, M., Chanel, G., 2020. Sharing emotions contributes to regulating collaborative intentions in group problem-solving. *Front. Psychol.* 11, 1–14. <https://doi.org/10.3389/fpsyg.2020.01160>.
- Birk, L.E., 2013. Construct Relevant and Irrelevant Variables in Math Problem Solving Assessment. Doctoral Dissertation. University of Oregon. <https://core.ac.uk/download/pdf/36687798.pdf>.
- Bunderson, C.V., Inouye, D.K., Olsen, J.B., 1988. The four generations of computerized educational measurement. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. Macmillan, New York.
- Care, E., Griffin, P., 2014. An approach to assessment of collaborative problem solving. *Res. Pract. Technol. Enhanc. Learn.* 9 (3), 367–388.

²The replacement of the human-to-human approach with the human-to-agent approach is still controversial. According to Herborn et al. (2020), there are no significant differences between the two approaches. However, given the constraints on the human-to-human approach applied in their study, the generalizability of their results should be explored in greater detail (Scoular et al., 2020).

- Care, E., Scoular, C., Griffin, P., 2016. Assessment of collaborative problem-solving in education environments. *Appl. Meas. Educ.* 29 (4), 250–264. <https://doi.org/10.1080/08957347.2016.1209204>.
- Child, S.F.J., Shaw, S., 2016. Collaboration in the 21st century: implications for assessment. *Res. Matters* 22, 17–22.
- Cognition and Technology Group at Vanderbilt, 1992. The Jasper experiment: an exploration of issues in learning and instructional design. *Educ. Technol. Res. Dev.* 40, 65–80. <https://doi.org/10.1007/BF02296707>.
- Cope, B., Kalantzis, M., 2016. Big data comes to school: implications for learning, assessment, and research. *AERA Open* 2 (2), 1–19. <https://doi.org/10.1177/2332858416641907>.
- De Boeck, P., Scalise, K., 2019. Collaborative problem solving: processing actions, time, and performance. *Front. Psychol.* 10, 1–9. <https://doi.org/10.3389/fpsyg.2019.01280>.
- DiCerbo, K.E., Behrens, J.T., 2012. From Technology-Enhanced Assessments to Assessment-Enhanced Technology. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, Vancouver, British Columbia, Canada.
- Dowell, N., Lin, Y., Godfrey, A., Brooks, C., 2020. Exploring the relationship between emergent sociocognitive roles, collaborative problem-solving skills and outcomes: a group communication analysis. *J. Learn. Anal.* 7 (1), 38–57. <https://doi.org/10.18608/jla.2020.71.4>.
- Ercikan, K., Oliveri, M.E., 2016. In search of validity evidence in support of the interpretation and use of assessments of complex constructs: discussion of research on assessing 21st century skills. *Appl. Meas. Educ.* 29 (4), 310–318. <https://doi.org/10.1080/08957347.2016.1209210>.
- Fiore, S.M., Kapalo, K.A., 2017. Innovation in team interaction: new methods for assessing collaboration between brains and bodies using a multi-level framework. In: von Davier, A.A., Zhu, M., Kyllonen, P.C. (Eds.), *Innovative Assessment of Collaboration*. Springer International Publishing, Cham, pp. 51–64.
- Forsyth, C., Andrews-Todd, J., Steinberg, J., 2020. Are you really a team player? Profiling of collaborative problem-solvers in an online environment. In: *Proceedings of the 13th International Conference on Educational Data Mining (EDM'20)*, pp. 403–408.
- Frederiksen, J.R., Collins, A., 1989. A systems approach to educational testing. *Educ. Res.* 18 (9), 27–32. <https://doi.org/10.2307/1176716>.
- Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W., Hesse, F.W., 2018. Advancing the science of collaborative problem solving. *Psychol. Sci. Publ. Interest* 19 (2), 59–92. <https://doi.org/10.1177/1529100618808244>.
- Greiff, S., Holt, D.V., Funke, J., 2013. Perspectives on problem solving in educational assessment: analytical, interactive, and collaborative problem solving. *J. Probl. Solving* 5 (2), 71–91. <https://doi.org/10.7771/1932-6246.1153>.
- Griffin, P., Care, E., McGaw, B., 2012. The changing role of education and schools. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer International Publishing, Cham, pp. 1–16.
- Hao, J., Liu, L., von Davier, A., Kyllonen, P., 2015. Assessing collaborative problem solving with simulation based tasks. In: Lindwall, O., Häkkinen, P., Koschman, T., Tchounikine, P., Ludvigsen, S. (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015*. The International Society of the Learning Sciences, Gothenburg, Sweden, pp. 544–547.
- Hao, J., Liu, L., Kyllonen, P., Flor, M., von Davier, A.A., 2019. Psychometric considerations and a general scoring strategy for assessments of collaborative problem solving. *ETS Res. Rep. Ser.* 1, 1–17. <https://doi.org/10.1002/ets2.12276>.
- Herborn, K., Stadler, M., Mustafić, M., Greiff, S., 2020. The assessment of collaborative problem solving in PISA 2015: can computer agents replace humans? *Comput. Hum. Behav.* 104, 1–9. <https://doi.org/10.1016/j.chb.2018.07.035>.
- Herde, C.N., Lievens, F., Solberg, E.G., Harbaugh, J.L., Strong, M.H., Burkholder, G.J., 2019. Situational judgement tests as measures of 21st century skills: evidence across Europe and Latin America. *J. Work Organ. Psychol.* 35 (2), 65–74.
- Kaendler, C., Wiedmann, M., Rummel, N., Spada, H., 2015. Teacher competencies for the implementation of collaborative learning in the classroom: a framework and research review. *Educ. Psychol. Rev.* 27, 505–536. <https://doi.org/10.1007/s10648-014-9288-9>.
- Krkovic, K., Mustafić, M., Wustenberg, S., Greiff, S., 2018. Shifts in the assessment of problem solving. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer International Publishing, Cham, pp. 55–74.
- Kroehne, U., Goldhammer, F., 2018. How to conceptualize, represent, and analyze log data from technology-based assessment? A generic framework and an application to questionnaire items. *Behaviormetrika* 45, 527–563. <https://doi.org/10.1007/s41237-018-0063-y>.
- Kyllonen, P.C., 2012. Measurement of 21st Century Skills Within the Common Core State Standards. *Invitational Research Symposium on Technology Enhanced Assessments [Symposium]*. <http://www.ets.org/Media/Research/pdf/session5-kyllonen-paper-tea2012.pdf>.
- Lievens, F., Motowidlo, S.J., 2016. Situational judgement tests: from measures of situational judgment to measures of general domain knowledge. *Ind. Organ. Psychol.* 9 (1), 3–22. <https://doi.org/10.1017/lop.2015.71>.
- Lin, K.-Y., Yu, K.-C., Hsiao, H.-S., Chu, Y.-H., Chang, Y.-S., Chien, Y.-H., 2015. Design of an assessment system for collaborative problem solving in STEM education. *J. Comput. Educ.* 2, 301–322. <https://doi.org/10.1007/s40692-015-0038-x>.
- Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educ. Res.* 23 (2), 13–23. <https://doi.org/10.3102/0013189X023002013>.
- Mislevy, R.J., Almond, R.G., Lukas, L.F., 2003. *A Brief Introduction to Evidence-Centered Design*. Educational Testing Service, Princeton, NJ. <https://www.ets.org/Media/Research/pdf/RR-03-16.pdf>.
- National Research Council, 2011. *Assessing 21st Century Skills: Summary of a Workshop*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/13215>.
- OECD, 2016. Description of the Released Unit From 2015 PISA Collaborative Problem-Solving Assessment, Collaborative Problem-Solving Skills, and Proficiency Levels. <https://www.oecd.org/pisa/test/CPS-Xandar-scoring-guide.pdf>.
- OECD, 2017a. PISA 2015 Collaborative Problem-Solving Framework. <https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>.
- OECD, 2017b. PISA 2015 Results (Volume V): Collaborative Problem Solving. <https://doi.org/10.1787/9789264285521-en>.
- Oliveri, M.E., Lawless, R., Molloy, H., 2017. A literature review on collaborative problem solving for college and workforce readiness. *ETS Res. Rep. Ser.* 1, 1–27. <https://doi.org/10.1002/ets2.12133>.
- O'Neil, H.F., Chuang, S.-H., Chung, G.K.W.K., 2003. Issues in the computer-based assessment of collaborative problem solving. *Assess. Educ. Princ. Pol. Pract.* 10 (3), 361–373. <https://doi.org/10.1080/0969594032000148190>.
- Polyak, S.T., von Davier, A.A., Peterschmidt, K., 2017. Computational psychometrics for the measurement of collaborative problem solving skills. *Front. Psychol.* 8, 1–16. <https://doi.org/10.3389/fpsyg.2017.02029>.
- Poon, C.L., Tan, S., Cheah, H.M., Lim, P.Y., Ng, H.L., 2015. Student and teacher responses to collaborative problem solving and learning through digital networks in Singapore. In: Griffin, P., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer International Publishing, Cham, pp. 199–212.
- Quellmalz, E.S., Pellegrino, J.W., 2009. Technology and testing. *Science* 323, 75–79. <https://doi.org/10.1126/science.1168046>.
- Rosen, Y., 2015. Computer-based assessment of collaborative problem solving: exploring the feasibility of human-to-agent approach. *Int. J. Artif. Intell. Educ.* 25, 380–406. <https://doi.org/10.1007/s40593-015-0042-3>.
- Rosen, Y., Tager, M., 2013. Computer-Based Assessment of Collaborative Problem-Solving Skills: Human-to-Agent Versus Human-to-Human Approach, pp. 1–35. <https://pdfs.semanticscholar.org/caa8/d4f05d57ea2be94e707d71c6a15cb94cfc79.pdf>.
- Rotondi, A.J., 1999. Assessing a team's problem solving ability: evaluation of the Team Problem Solving Assessment Tool (TPSAT). *Health Care Manag. Sci.* 2, 205–214. <https://doi.org/10.1023/A:1019004332123>.
- Schmidgall, J., Oliveri, M.E., Duke, T., Grissom, E.C., 2019. Justifying the construct definition for language proficiency assessment: the redesigned TOEIC Bridge tests—framework paper. *ETS Res. Rep. Ser.* 1, 1–20. <https://doi.org/10.1002/ets2.12267>.

- Scoular, C., Care, E., 2019. A generalized scoring process to measure collaborative problem solving in online environments. *Educ. Assess.* 24 (3), 213–234. <https://doi.org/10.1080/10627197.2019.1615372>.
- Scoular, C., Eleftheriadou, S., Ramalingam, D., Cloney, D., 2020. Comparative analysis of student performance in collaborative problem solving: what does it tell us? *Aust. J. Educ.* 64 (3), 282–303. <https://doi.org/10.1177/0004944120957390>.
- Shaffer, D.W., Collier, W., Ruis, A.R., 2016. A tutorial on epistemic network analysis: analyzing the structure of connections in cognitive, social, and interaction data. *J. Learn. Anal.* 3 (3), 9–45. <https://doi.org/10.18608/jla.2016.33.3>.
- Shute, V.J., Leighton, J.P., Jang, E.E., Chu, M.-W., 2016. Advances in the science of assessment. *Educ. Assess.* 21 (1), 34–59. <https://doi.org/10.1080/10627197.2015.1127752>.
- Stadler, M., Herborn, K., Mustafić, M., Greiff, S., 2020. The assessment of collaborative problem solving in PISA 2015: an investigation of the validity of the PISA 2015 CPS tasks. *Comput. Educ.* 157, 1–11. <https://doi.org/10.1016/j.compedu.2020.103964>.
- Stevens, M.J., Campion, M.A., 1999. Staffing work teams: development and validation of a selection test for teamwork settings. *J. Manag.* 25 (2), 207–228. <https://doi.org/10.1177/014920639902500205>.
- Stoeffler, K., Rosen, Y., Bolsinova, M., von Davier, A.A., 2020. Gamified performance assessment of collaborative problem solving skills. *Comput. Hum. Behav.* 104, 1–9. <https://doi.org/10.1016/j.chb.2019.05.033>.
- Sun, C., Shute, V.J., Stewart, A., Yonehiro, J., Duran, N., D'Mello, S., 2020. Towards a generalized competency model of collaborative problem solving. *Comput. Educ.* 143, 1–17. <https://doi.org/10.1016/j.compedu.2019.103672>.
- Swiecki, Z., Ruis, A.R., Farrell, C., Shaffer, D.W., 2020. Assessing individual contributions to collaborative problem solving: a network analysis approach. *Comput. Hum. Behav.* 104, 1–15. <https://doi.org/10.1016/j.chb.2019.01.009>.
- Taggar, S., Brown, T.C., 2001. Problem-solving team behaviors: development and validation of BOS and a hierarchical factor structure. *Small Group Res.* 32 (6), 698–726. <https://doi.org/10.1177/104649640103200602>.
- Von Davier, A.A., Halpin, P.F., 2013. Collaborative problem solving and the assessment of cognitive skills: psychometric considerations. *ETS Res. Rep. Ser.* 2, i–36. <https://doi.org/10.1002/j.2333-8504.2013.tb02348.x>.
- Von Davier, A.A., Hao, J., Liu, L., Kyllonen, P., 2017. Interdisciplinary research agenda in support of assessment of collaborative problem solving: lessons learned from developing a collaborative science assessment prototype. *Comput. Hum. Behav.* 76, 631–640. <https://doi.org/10.1016/j.chb.2017.04.059>.
- Waugh, G.W., Russell, T.L., 2005. Predictor situational judgment test (PSJT). In: Knapp, D.J., Sager, C.E., Tremble, T.R. (Eds.), *Development of Experimental Army Enlisted Personnel Selection and Classification Tests and Job Performance Criteria*. US. Army Research Institute for the Behavioural and Social Sciences, Arlington, VA, pp. 59–66.
- Wiggins, G., 1989. A true test: toward more authentic and equitable assessment. *Phi Delta Kappan* 70 (9), 703–713. <https://www.jstor.org/stable/20404004>.
- Yuan, J., Xian, Y., Liu, H., 2019. Assessment of collaborative problem solving based on process stream data: a new paradigm for extracting indicators and modeling dyad data. *Front. Psychol.* 10, 1–14. <https://doi.org/10.3389/fpsyg.2019.00369>.
- Zhuang, X., MacCann, C., Wang, L., Liu, L., Roberts, R.D., 2008. Development and validity evidence supporting a teamwork and collaboration assessment for high school students. *ETS Res. Rep. Ser.* 2, i–51. <https://doi.org/10.1002/j.2333-8504.2008.tb02136.x>.
- Zieky, M.J., 2014. An introduction to the use of evidence-centered design in test development. *Psycologia Educativa* 20, 79–87. <https://doi.org/10.1016/j.pse.2014.11.003>.
- Zumbo, B.D., 2015. Consequences, Side Effects and the Ecology of Testing: Keys to Considering Assessment In Vivo. YouTube. <https://www.youtube.com/watch?v=0L6Lr2BzuSQ&feature=youtu.be>.

Educational assessment of 21st century skills—novel initiatives, yet a lack of systemic transformation

Ola Erstad^a and Fazilat Siddiq^b, ^a Department of Education, University of Oslo, Oslo, Norway; and ^b Unit for Digitalization and Educational Quality, University of South-Eastern Norway, Drammen, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	245
Defining 21st century skills	245
Evolving frameworks of 21st century skills	246
From policy to practice	247
Assessment in the 21st century—digitalization	248
International assessment initiatives	250
Case 1: Assessment of 21st century skills (ATC21S)	250
Case 2: Assessment of computer and information literacy	251
National 21st century skills assessment—projects	251
Example 1: “Teaching and transfer effects of 21st century skills—collaborative problem solving in digital environments” (TEACH21st)	252
Example 2: “Educational assessments of the 21st century: measuring and understanding students’ adaptability in complex problem solving situations” (ADAPT21)	252
Example 3: “Modeling collaborative problem-solving data: new approaches for individual and team level inferences”	252
The transformative aspects of assessment in the 21st century	253
References	254

There is a belief that ease of measurement often dictates which subjects are assessed, and “hard to measure” subjects are ignored. Assessment and measurement are in turn seen as reducing learning and curriculum to what is easy to measure. In fact, nothing is too hard to measure. As [Thurstone \(1959\)](#) said, “If it exists it can be measured, and if it can’t be measured it doesn’t exist.” It all depends on how measurement is defined and how we organize the evidence of more difficult learning concepts.

[Griffin et al. \(2012, p. 10\).](#)

Introduction

The transition from the 20th to the 21st century has diverted increased attention toward the societal changes and transformations that have taken place over the last two decades. Change and transformation have become buzzwords indicating how our societies are adjusting to the contemporary and future challenges of labor markets, economic models, education and the environment ([Hölscher et al., 2018](#)). However, there is no agreement concerning what change and transformation really mean in policy and practice, which leaves an open discourse about how to define key parameters for such development in education. Further, the most important driving force for these perspectives has been the technological developments that have taken place during this period ([Mahlow and Hediger, 2019](#)). Technological developments and innovations—along with increased access to them—and media culture have created entirely new conditions for communicative interaction, information flow, creation and so forth ([Mahlow and Hediger, 2019](#)).

In this chapter, we focus on the educational assessment of 21st century skills, aiming to put forward critical issues related to the assessment of such competence areas and frameworks in contemporary societies. 21st century skills frameworks have become important international policy directives, transforming curricula in many countries over the last decade. However, there is uncertainty surrounding how to understand and measure such skills. To move forward, there is a great need to systematize our knowledge and experiences in this area.

We start by introducing some definitions of 21st century skills, addressing theoretical and empirical perspectives and issues. We also provide a short overview of key transformations in educational assessment and an outline of international and national initiatives, using examples from two countries, Norway and Finland. In the next section, we present some examples of projects and assessment instruments relevant to our argument that assessment, education and technological development has to be defined as interrelated parts of how to understand and assess 21st century skills. In the last section, we detail some perspectives on assessments and transformations in education that point toward opportunities and pitfalls for the future. By taking a meta-perspective, we identify evolving themes within this field and provide a critical perspective on further developments.

Defining 21st century skills

Education has been defined as a key mechanism for providing future citizens with the qualifications and competences needed for life in the 21st century ([Ananiadou and Claro, 2009](#)). Instructing 21st century skills has thus become a key initiative, especially over

the last decade, and part of the educational policy agenda (Tahirsylaj and Sundberg, 2020). Different frameworks have been presented, and such skills have now entered curricula in many countries (Erstad et al., 2021). There is not a clear consensus on what defines such frameworks, but some key concepts and terminology are common across most of them (Joynt et al., 2019). While some have criticized the focus on 21st century skills as reliant on the metanarrative of “salvation through technology,” primarily informed by human capital theory or similar schools of thought that reduce student identity by praising homogenizing practices (Hilt et al., 2019), others have welcomed it as a vital part of a knowledge-based society (Voogt and Roblin, 2012).

A number of definitions have been proposed to pin down what exactly 21st century skills are. Nearly all the definitions emphasize their importance for individuals, society and future work life (Siddiq, 2016). In the following, we provide three definitions of 21st century skills. These examples were chosen because they are the most cited sources in related frameworks and definitions (Tahirsylaj and Sundberg, 2020). The first is provided by the Organization for Economic Co-operation and Development (OECD) and defines 21st century skills as “those skills and competencies young people will be required to have in order to be effective workers and citizens in the knowledge society of the 21st century” (Ananiadou and Claro, 2009, p. 8). The second is proposed by the Partnership for 21st Century Skills (P21), stating that they are “the skills, knowledge and expertise students must master to succeed in work and life—a blend of content, knowledge, specific skills, expertise and literacies” (2010, p. 8). The third is provided by the Assessment and Teaching of 21st Century Skills (ATC21S) project, which asserts that 21st century skills are underpinned by knowledge, skills, attitudes, values and ethics—referred to as the KSAVE model (see the next section for more on this initiative) (Binkley et al., 2012).

In a recent systematic review, Tahirsylaj and Sundberg conclude that “there is still unfinished business for educational researchers in critically engaging with framing and defining competences for the 21st century” (2020, p. 1). Acknowledging this need, we also agree with scholars in the field that 21st century skills are broad, complex and generic (Geisinger, 2016). They refer to a wide range of competences, attitudes, values and ways of working that are cross-curricular and changing alongside the developments and transformations brought by technology. A number of initiatives further unpack and detail 21st century skills, identifying the most common ones as collaboration, communication, ICT literacy (also labeled digital competence or digital literacy; Siddiq, 2016), citizenship, problem solving (reference to Greiff in the Encyclopedia), critical thinking, creativity and productivity (Geisinger, 2016; Voogt and Roblin, 2012). Note that some frameworks also mention learning to learn, self-direction, adaptability, mathematics, science, history and art (Voogt and Roblin, 2012). We note that even though the study by Voogt and Roblin was conducted almost a decade ago, it is still relevant and often referred to. Most importantly, to our knowledge, no update to its comparative analysis of 21st century skills frameworks has been published. In the following section, we provide an overview of some of the evolving 21st century skills frameworks that have laid much of the foundational focus and emphasis on these skills in education and we discuss how assessment has become a key aspect of such frameworks.

Evolving frameworks of 21st century skills

A key initiative that has been influential for the way policy makers and researchers address challenges and transformations for education in the 21st century is the Definition and Selection of Key Competencies (DeSeCo) project, which the OECD began at the end of the 1990s, about the same time as the Program for International Student Assessment (PISA). A group of researchers and experts were brought together to discuss the concept of “competence” and how it relates to social processes of transformation, as well as “cross-curricular competencies” as the “knowledge and skills related to outcomes of education in a broad sense” (Salganik et al., 1999, p. 13). The authors of the DeSeCo reports mention that concepts like “competence” and “skills” are quite vague. Their approach focuses on the complex challenges that young people today face when entering work and life, indicating that the notion of competencies encompasses cognitive but also motivational, ethical, social and behavioral components, combining stable traits, learning outcomes (e.g., knowledge and skills), belief–value systems, habits and other psychological features. Basic reading, writing and calculating are thus skills that make up the critical foundation of numerous competencies. That is, 21st century skills are wider than pure subject knowledge; they are multifunctional, multi-faceted and inter-related, consisting of several competencies and sub-competencies (Griffin et al., 2012; Tahirsylaj and Sundberg, 2020; Voogt and Roblin, 2012). Rychen and Salganik (2001, p. 8) emphasize that they are “transversal across social fields and refer to a higher order of mental complexity.” In particular, understanding what these skills are is utterly important from an assessment perspective, which has traditionally focused on one single subject at a time (Siaorova et al., 2017). We do not hold the view that assessment within a single subject is easy or effortless. Rather, our viewpoint is that 21st century skills frameworks and issues have added several dimensions to the assessment field.

In 2006, the European Commission published the Key Competences for Lifelong Learning: a European Reference Framework (European Parliament, 2006). This framework of key competences was revised 10 years later and implemented in 2018 with an increased focus on 21st century skills. Eight key competences are outlined therein as necessary for personal fulfillment, active citizenship, social cohesion and employability. These competences are: (1) literacy; (2) multilingualism; (3) numerical, scientific and engineering skills; (4) digital and technology-based competences; (5) interpersonal skills and the ability to adopt new competences; (6) active citizenship; (7) entrepreneurship; and (8) cultural awareness and expression.

The Key Competences Network (KeyCoNet) project, organized by the European Schoolnet (Looney and Michel, 2014), brings together definitions and initiatives surrounding these key competences, especially digital competence. One report, based on surveys conducted across Europe, shows that many countries struggle in their efforts to implement such competences into school curricula (Arjomand et al., 2013). Curricula developers and schools have a tendency to focus on competences that are familiar, like STEM and mother tongue and foreign language learning, while more unfamiliar and complex competences, like creativity, collaborative

problem solving and digital competence, are not implemented in the compulsory curriculum. Moreover, the assessment of such competences is oftentimes not addressed (Care et al., 2018a,b). The reason for this oversight is a key area to address in order to understand future possibilities and tensions in educational assessment of key competences and 21st century skills.

An important initiative in this field is the Assessment and Teaching of 21st Century Skills project (ATC21S) (Griffin et al., 2012). This is a unique project in the sense that it not only identifies and defines 21st century skills but also focuses on assessment as a core part of its mission. One of the working groups, of which Ola Erstad was a core member, published a white paper, which later became a book chapter, called “Identification and definition of 21st century skills.” This group identified ten skillsets defining 21st century skills, grouping them into four areas: ways of thinking (1. creativity and innovation; 2. critical thinking, problem solving, decision making; 3. learning to learn, metacognition); ways of working (4. communication; 5. collaboration, teamwork); tools for working (6. information literacy; 7. ICT literacy); and living in the world (8. citizenship—local and global; 9. life and career; 10. personal and social responsibility—cultural awareness and competence). What is unique to this project compared to other national and international frameworks from this time is that each skillset is operationalized according to its components: knowledge, skills, attitudes, values and ethics, thus called the KSAVE model (Binkley et al., 2012, p. 37).

These frameworks have set the agenda for instructing 21st century skills and the future orientation of education systems around the world. In the next sections, we will focus on more concrete examples of how these frameworks have been addressed before returning to overall perspectives of transformation and change as part of the educational assessment of these skills. One fundamental difficulty is that of gathering information on assessment of 21st century skills in non-European contexts. For example, the International Computer and Information Literacy Study (ICILS), across three cycles (2013, 2018, 2023) has no representation from African countries and an overrepresentation of European countries.

From policy to practice

Central to how the assessment of 21st century skills is implemented and explored as part of learning practices is the way curricula specify how such skills should be integrated into teaching and instruction—in which subjects and on which levels. Curricula are the tools through which policy makers influence practices. Over the last 5–6 years, new curriculum developments have emerged in several countries, referring to what they call 21st century skills and reflecting on the ways that teachers conceive of assessment as part of the process of developing such skills and as being linked to the digitalization of schools. However, surprisingly, not much has been written about how digital technologies transform the development and content of curricula for schools (Williamson, 2013; Erstad et al., 2021).

We will here introduce some recent developments in the countries we know best and focus on assessment in the next part. Within the Nordic region, Norway and Finland are the countries that have most explicitly referred to 21st century skills in recently updated curricula inspired by the contemporary frameworks mentioned above. Interestingly, the ways in which these two countries have dealt with such skills when developing new national curricula are slightly different.

In Finland, a re-orientation of the school system has been implemented through the new National Core Curriculum of 2016 (Finnish National Agency for Education, 2016). One of its main goals has been to enhance competence building and the development of 21st century learning skills among students (Erstad et al., 2021). The rationale for this re-orientation is that the world is changing rapidly; the school system needs to evolve to meet the needs of the future while still enabling equal opportunities. Even though Finland has scored high on international tests like PISA (OECD), policy makers have identified several challenges emanating from major societal, demographic, cultural and economic changes. This recognition has pushed forward a sense of urgency in building a school system oriented toward the future. One study clarifies that the new educational curriculum in Finland holds the view that “the purpose of education is to promote life-long and life-wide learning, holistic development and well-being of all learners, as well as to improve their skills for living in a sustainable way” (Halinen, 2018, p. 75).

One of the main intentions of this curriculum reform was to promote an integrative approach to teaching and learning. The purpose of this approach was to enable students to see the relationships and interdependencies between the subjects studied at school and the phenomena of life outside school (Halinen, 2018, p. 84). The most important tools for this purpose, defined in the core curricula, are transversal competencies, multi-disciplinary learning modules and pupil assessment. These “transversal competences” are developed by all subjects and are based on the Act on Early Childhood Education and Care and the National Core Curriculum for ECEC (2018). These competences are:

- (1) Thinking and learning to learn
- (2) Cultural competence, interaction and expression
- (3) Taking care of oneself and managing daily life
- (4) Multiliteracy
- (5) Digital (ICT) competence
- (6) Working life competence, entrepreneurship
- (7) Participation, involvement and building a sustainable future

These seven transversal competences have been taken into account when defining the objectives and main content areas of compulsory school subjects. Moreover, the content load has been reduced in all subjects to allow for more depth in understanding of domain knowledge and spend time on transversal competences. The subject descriptions also point out the links between the objectives of the subject and the objectives of the transversal competences.

Interestingly, these transversal competences are not defined as 21st century skills specifically, but rather as broader principles of learning and literacy. Still, they are framed within conceptions of 21st century skills and the changes necessary within schools to meet future challenges.

Assessment has also been included in this strategy to transform the methods through which schools support a more integrative approach to teaching and learning. As stated by Halinen (2018): “In Finland, especially in basic education, the focus is on formative assessment and giving supportive and encouraging feedback. In addition to the assessment of learning, more and more assessment for learning and assessment as learning takes place. The development of skills for self-assessment and for peer-assessment is promoted during the school years. This means that students learn step by step, and are supported by their teachers, to set goals for their own learning, to discuss the assessment criteria, to plan and reflect upon their working processes, and to assess the results of their own work” (2018, p. 86). Halinen (2018) also makes the point that “in Finland, we seldom use the phrase ‘implementation of the curriculum’ but instead talk about teachers who construct their own professional guidelines based on the local curriculum. They are the ones who bring the curriculum into reality” (p. 87). However, a system and tools of assessment that fulfill these objectives for schools are still in development.

In Norway, the process of revising the national curriculum started with a broader process of exploring a framework for future schools in Norway, led by a governmental committee consisting of researchers, stakeholders and representatives from the field of practice but also involving teachers and school leaders around the country through a blog. This process produced a lot of engagement from teachers with suggestions about a revision of school subjects. As part of preparing for the new curriculum, the governmental committee commissioned a systematic review of research and frameworks on 21st century skills, led by Prof. Erstad. This work summarized common concepts and frameworks across the literature by defining ten key concepts therein. The committee further used this review in writing up a green paper for the government (NOU, 2015). In this green paper, they define four key areas, described as future competences, that should be included in all subjects. These areas are:

- Subject-specific knowledge
- Competence in learning
- Competence in exploring and creating
- Competence in communicating, cooperation and participation

However, in writing up the final curriculum, the Renewal of the Subjects (“Fagfornyelsen,” White Paper, 2016), these competence areas are not defined; the focus is rather on how the subjects can be renewed. Moreover, these areas are integrated into principles for learning, development and *Bildung*. The term *Bildung* refers to the German concept of self-cultivation, to become literate/educated, and the process of formation as a lifelong process of becoming human (Siljander and Sutinen, 2012). *Bildung* corresponds to the Humboldtian (1767–1835) model of higher education of spiritual and cultural growth, and refers to both a *process* and *outcome* in developing a person through education (Schneuwly and Vollmer, 2018). This tradition has been criticized for being too elitistic, and during the last two decades attempts have been made to re-write conceptions of *Bildung* better suited for the digital age. Hence, in lack of an equally comprehensive concept, Digital *bildung* has been introduced. Digital *bildung*, among others, focuses on how students’ participation, multi-membership of different communities and identity development in the digital era are influenced by the digitization of society. This implies ethical and moral reflections on the technology’s role in human development (Krumsvik, 2008).

In the Norwegian curriculum *Bildung* traditionally is important for the overall aims of education. Recent curriculum developments privilege a connection between basic skills to prepare students for the future and certain aspects of schooling, such as three key trans-disciplinary themes (democracy and citizenship; sustainable development; public health and coping with life) and the specific principle of deep learning. This focus has garnered a lot of interest from teachers, but it is not clear what these themes actually mean in practice nor the conceptual understanding of deep learning. Another important new aspect of the national curriculum is its emphasis on computational thinking and programming, which also brings new aspects to the framework of 21st century skills that need consideration for the purpose of assessment.

Curriculum developments in Finland, Norway and other countries point to the existence of important challenges in implementing 21st century skills frameworks from policy to practice. In both national curricula, assessment is defined as very important in supporting students’ learning. However, they are vague in explaining how assessment is embedded in instruction of 21st century skills, and they do not give any indication about a renewal of the assessment system tuned to future challenges and new orientations toward skills and competences.

Assessment in the 21st century—digitalization

Assessments serve many functions within education systems, motivating and helping students to learn and assisting teachers in refining their practice and supporting their students’ learning. Some additional functions include certifying student accomplishments, evaluating the output of educational programs, measuring the progress of educational systems, and facilitating cross-national comparisons (OECD, 2013). Linked to students’ learning, assessment is often divided into two main categories: formative assessment, understood as “assessment for learning,” and summative assessment, understood as “assessment of learning” (Siarova et al., 2017; Boitshwarelo et al., 2017, James in Encyclopeida). While formative assessment is undertaken throughout the school year and functions as feedback for the student to increase learning, summative assessment is often conducted at the end of the

learning cycle (e.g., at the end of the school year or semester). The latter is often high stakes, constituting the students' knowledge in the form of a grade, which may impact their future plans. Some researchers emphasize a third approach toward assessment, "assessment as learning," which extends the use of formative assessment for learning and is considered a process through which students are largely involved in the assessment, playing a key role in their learning (Dann, 2002).

In the field of digitalization and the transformation of education and assessment, increased attention has been paid to developing new and improved instruments to assess new competence areas. Nevertheless, this field is still very much in development, and there is a basic challenge in adjusting both formative and summative assessment strategies to be more in tune with the needs of education systems concerning new areas of skills and competences among students. An important dimension to this shift is the role of digital technologies and the assessment of 21st century skills. Technological developments can facilitate innovations and alternative approaches to assessments (Siarova et al., 2017). Even in a key international journal like *Assessment in Education: Principles, Policy & Practice*, there are not many articles that target assessment and digital technologies specifically. There is an interesting special issue from 2003 in the sense that it addressed important issues of technology and assessment early in the 2000s when technologies were starting to enter schools on a large scale and that this is an important journal in the field setting the agenda for such issues in this field of research. During the early 2000s there are not many other articles or journal issues with the digital dimension as a core focus. However, after this special issue an important development was how ways of using digital technologies became associated with key skills development of the 21st century, both as part of the above-mentioned frameworks, and as opportunities to transform assessment methods in fundamental ways (Griffin et al., 2012; van Laar et al., 2017).

Revisiting an article written by Erstad (2008), which summarizes issues surrounding assessment practices and the role of new technologies therein, it is clear that a major concern in much of the research at that time was the transition from paper-and-pencil to computer-based assessment. Several studies trying to compare these testing methods describe the latter as highly problematic, especially concerning issues of test validity (Russell et al., 2003). However, since 2008, most assessments have shifted to using computers and including digital interfaces, while paper-and-pencil tests are less common in most countries. Furthermore, the assessments—in particular summative assessments—have largely moved from being driven by individual teachers and schools to being derived from national tests and international studies, like PISA and the Trends in International Mathematics and Science Study (TIMSS).

However, as described then (Erstad, 2008), it is still possible to see three combinations of assessment and digital technologies. First, such technologies can be used to make traditional tests and assessments more efficient, since data can be processed using digital software. Second, digital technologies are used to assess competences and areas that traditionally have been difficult to assess and to develop new instruments and approaches. Technological developments have created a new interest in what some describe as "assessing the inaccessible" (Nunes et al., 2003), like metacognition, learning strategies, attitudes and lifelong learning skills (Anderson, 2008; Deakin Crick et al., 2004; Siddiq et al., 2017). These innovations have promoted the use of formative ways of assessment in order to better monitor and assess student progress through different modes of expression (multimodality). An important point is also the way digital tools can support collaborative work. Students can send documents and files to each other and, in this way, work on tasks together. Within the field of computer-supported collaborative learning (CSCL), there are many examples of computer-based learning environments designed for collaboration to stimulate student learning and the process of inquiry (Wasson et al., 2003). Such collaborative problem-solving skills are considered necessary for success in today's world of work and school. Online collaborative problem-solving tasks offer new measurement opportunities when information on what individuals and teams are doing is synthesized along the cognitive dimension. This development raises issues both about interface design features that can support online measurement and how to evaluate collaborative problem-solving processes in an online context (O'Neil et al., 2003). There are also examples of web-based peer assessment strategies in ways that students themselves are put in the position of evaluators as well as of learners (Lin et al., 2001; Lee et al., 2006). Third, digital technologies can be used to test digital skills specifically. Since digital skills and competences have become key elements of national curricula over the last decade, as we will return to below, the challenge has been to develop items and instruments for assessment of such new competence areas (see **Case 2: Assessment of computer and information literacy**).

Another dimension that has entered the field of assessment and digital technologies in the last decade is that of "learning analytics." The development of online environments and participation, also referred to as "platformization" (Sefton-Green, 2021), means that digital data has become a new resource for the assessment of skills such as online reading, writing and collaboration, self-regulated learning over time, creativity and lifelong learning. Digital data also provides information about how students navigate and click within online environments as part of solving tasks, thereby visualizing how students solve tasks, individually or in collaboration, and where they may have problems. Enormous amounts of data are collected as part of educational online platforms linked to basic and higher education, as well as in Massive Open Online Courses (MOOCs) and throughout our everyday lives. However, how we can systematically collect, analyze and visualize data generated by digital tools as part of assessment activities is still being investigated (Cui et al., 2020). After years of over-optimistic arguments about the possibilities provided by data tracing and gathering for creating affordances for students' learning processes, this field seems to become more nuanced in what can actually be done in approaches toward "big data," along with digital tools that need to be developed to make sense of the vastness of the available data and how it might help us in developing new tools and approaches to assess 21st century skills. Moreover, as highlighted by Murchan and Siddiq (2021) there is a critical need to address ethical and privacy issues related to use of process data in educational assessment.

International assessment initiatives

Acknowledging that issues of educational transformations, and specifically those related to 21st century skills frameworks and assessment practices, are still in development, it is all the more important to point out the main challenges and learn from previous experiences. In this section, we will present some international initiatives that have developed instruments and practices for assessment. We identify two competence areas within these frameworks, namely digital competence and collaborative problem solving, and focus on them as cases through which to illustrate developments in assessment instruments, approaches and research linked to educational practices. After these case presentations, we will also briefly describe three ongoing projects to demonstrate how former international initiatives are now being followed up with new national projects in this field. There are of course other relevant examples internationally, but we concentrate on those that we know well as illustrations of further developments.

Case 1: Assessment of 21st century skills (ATC21S)

The team behind the above-mentioned ATC21S project (Griffin et al., 2012) wanted to follow up on the initial framework and perspectives to develop more concrete tools for the assessment of 21st century skills. Two institutions took part—the Assessment Research Center at the University of Melbourne, Australia and the Berkeley Evaluation and Assessment Research (BEAR) Center at UC Berkeley, USA. They identified two competence areas, combining several of the skillsets in the initial framework, to be assessed on a broader scale internationally. They aimed at measuring collaborative problem solving (including a combination of critical thinking, problem solving, decision-making and collaboration) and digital literacy through an assessment they called the Learning through Digital Networks—ICT Literacy Test, which evaluates the competence areas of information literacy, ICT literacy and personal and social responsibility within the 21st century skills framework. The team, together with international scholars, initialized and developed two main assessment instruments that were piloted in six countries (Australia, Costa Rica, Finland, Netherlands, Singapore and USA), and they reported their experiences in a book series (Care et al., 2018a; Griffin and Care, 2015).

Given our focus in this chapter, it is interesting that these 21st century skills were conceptualized as an outline for developing assessment instruments and approaches. Collaborative problem solving was understood as consisting of five broad strands (see Fig. 1) and was defined as the capacity of an individual to recognize the perspectives of other persons in a group, participate as a member of the group by contributing their knowledge, experience and expertise in a constructive way, recognize the need for contributions and how to manage them, identify structure and the procedure involved in resolving a problem and, as a member of the collaborative group, to build and develop knowledge and understanding. In the process of creating and field-testing collaborative problem solving tools, broad types of scenarios and tasks were developed and trialed (Fig. 1).

The learning through digital networks—ICT literacy competence was conceptualized as consisting of the following strands: learning as a consumer of information; learning as a producer of information; learning in the development of social capital; and learning in the development of intellectual capital (Griffin and Care, 2015). Several broad scenarios were developed that engaged up to four students at a time in identifying procedures and collaborative tools that enabled them to learn and develop (see Fig. 2).

For these two skill areas, the assessment tasks were checked with teachers to ensure that they were realistic, that students would be able to work with them, that they could differentiate between high and low levels of ability, and that the skills underpinning the task resolutions were teachable. Think-aloud protocols were used in small-scale studies (cognitive laboratories) with students representing the target population in order to generate bases for the automatic coding and scoring of student performances on the tasks. A series of small-scale pilot studies were undertaken in a small number of intact classes to determine the technological and administrative requirements for assessment implementation. These studies represented a rehearsal for the large-scale trials that were carried out in the above-mentioned six countries. Trial data was collected using a matrix sampling approach to identify a uniform cross-national sample of students to maximize calibration accuracy (Care et al., 2018a,b; Griffin and Care, 2015).

Since the ATC21ST project ended its funding period in about 2013, several smaller national and international groups have continued working with the two assessments. The Melbourne group has kept their focus on the assessment of collaborative problem solving by further developing tasks within this assessment framework alongside methods and practices for analyzing such data,

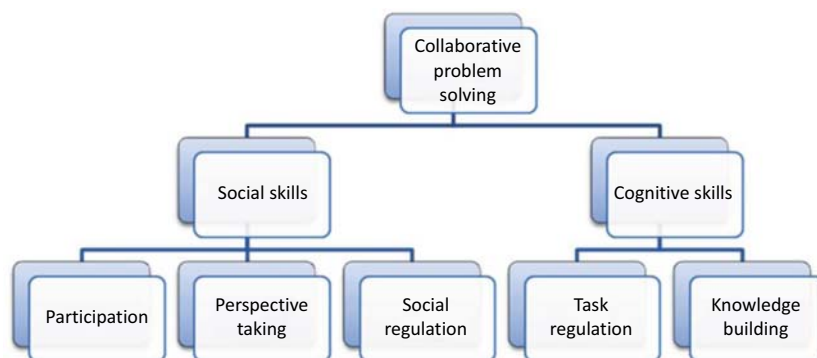


Fig. 1 Conceptual framework for collaborative problem solving (Griffin et al., 2012).

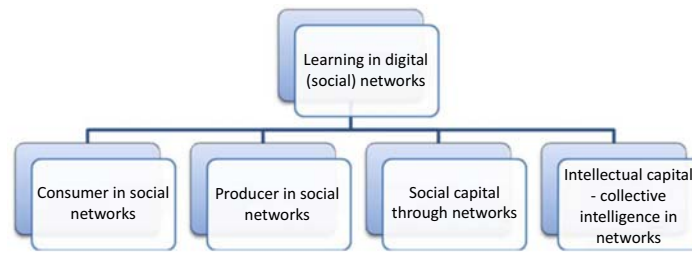


Fig. 2 Conceptual framework for learning in digital networks (Griffin et al., 2012).

while still disseminating knowledge gained from the continued project (e.g., Pöysä-Tarhonen et al., 2018; Scoular and Care, 2019, 2020).

The assessment of digital literacy also garnered continued attention by Fazilat Siddiq and others (Siddiq, 2016). Siddiq translated, adapted and revised the learning through digital networks—ICT literacy test, piloting it in the Norwegian educational context to investigate its validity and reliability (Siddiq et al., 2017). She also published a second study adding a novel task aimed at assessing students' collaborative problem solving (Siddiq and Scherer, 2017). In both these studies, it was apparent that it was not sufficient to score the students' responses to the tasks as correct or incorrect. To gain a further understanding of students' competence within the underlying domains and to assess these constructs authentically, it was also necessary to investigate and score the processes by which the answer is reached (e.g., following which paths they took through click streams). In sum, technology offers new opportunities for assessing the 21st century skills that could not easily be detected in paper-pencil assessments. At the same time, it also creates new challenges that needs to be dealt with, such as determining which patterns of movement within a test environment should be scored and what it would mean to leave them out.

Case 2: Assessment of computer and information literacy

The only international, large-scale assessment of digital literacy and 21st century curricula is The International Computer and Information Literacy Study (ICILS) developed and conducted by IEA (International Association for the Evaluation of Educational Achievement). ICILS measures and monitors developments and changes across time in students' computer and information literacy achievement and their home and school environments with regard to access to technology, use and support of digital literacy in classrooms (Fraillon et al., 2014). The assessment consists of both questionnaires about technology use among different stakeholders (e.g., teachers, principals) and a performance assessment. For the participating Nordic countries (Denmark, Norway and Finland), the study reveals a high rate of access to technologies both at school and at home compared to other countries, as well as higher performance on most of the items on the ICILS scale (Fraillon et al., 2014). The study also reports similarities in the impact of students' socio-economic background and gender—girls performed better than boys on most items—on the results. Nordic teachers are more positive about using digital technologies compared to many other participating countries; however, a majority of students report that such technologies are not used much in learning activities in subject domains, even though a majority of students use them at large at home (Fraillon et al., 2014).

ICILS was initiated in 2013 and runs every 5 years. Even though this large-scale assessment provides education systems and policymakers with important knowledge on the contexts and outcomes of CIL-related education programs, the participating countries vary. More importantly, some countries, for instance in Africa and in parts of Asia did not participate in ICILS. This entails that our knowledge regarding the assessment of digital literacy more globally is limited, and the unavailability of robust data in these contexts restricts our understanding surrounding these topics. Consequently, such issues might increase the digital divide with regard to access to technology and the development of students' and teachers' digital competence.

There are examples of tests and instruments that are being developed at national levels (e.g., in Australia, Norway and Hong Kong) in order to measure the digital skills and competences of students and teachers, but many of these are still in development (Erstad, 2008; Scherer et al., 2017; Tondeur et al., 2016). Reports from the ICILS 2013 and 2018 document the ways digital technologies are used in education in the participating countries, the competences of students and the self-reported attitudes and self-beliefs of the teachers (Fraillon et al., 2014, 2020). Moreover, these assessments have slightly changed across testing cycles. For instance, the 2018 ICILS included tasks for measuring computational thinking in addition to digital literacy, and both will also be assessed in the 2023 cycle. This ongoing transformation, along with developments in the processes of both collecting and processing assessment data, demonstrates that the assessment of 21st century skills is complex and changing as our knowledge regarding the constructs at play evolves and new technology and procedures for dealing with such assessment and data are developed.

National 21st century skills assessment—projects

To illustrate that some of the international cases mentioned above are also being emphasized and investigated through new initiatives and projects nationally, we will briefly present three examples of ongoing projects within our own research communities that target 21st century skills and new assessment initiatives.

Example 1: “Teaching and transfer effects of 21st century skills—collaborative problem solving in digital environments” (TEACH21st)

The TEACH21st project¹ is aimed at developing teaching material and practices for increasing upper secondary students’ “collaborative problem-solving competencies” (ColPS) and investigating whether the training of ColPS skills affects competences in both the domains under instruction as well as in other domains through transfer effects. This is a 4-year project (2019–2023) funded by the Norwegian Research Council. The project is designed to narrow the gap between research and practice by involving key stakeholders, focusing on the relevant curricular goals and including a classroom and student perspective (Siddiq et al., 2020). The initial phase (see Fig. 3) of the project focuses on the development and piloting of materials and practices to teach ColPS. To achieve this goal, a teacher design team (TDT) approach (Tondeur et al., 2016) has been employed. The TDTs consist of teacher educators, in-service teachers and student teachers who design educational material integrating ColPS with subject teaching. By involving different stakeholders in the design process, the implementation of new curricula in educational practice and its gradual application to classrooms is more likely to succeed (Wang and Hannafin, 2005). The TDTs are led by teacher educators from three different teacher-training institutions and are supported by the international project team. Throughout the developmental phase of the project, indicators of ecological and content validity are being developed to ensure that the teaching program is feasible, relevant for upper secondary curricula and classrooms, easy to implement and inclusive of the core elements of ColPS. In the second phase of the project (Fig. 3), one of the ColPS teaching programs developed by a TDT will be used in a randomized controlled trial (RCT). The project has just entered the second phase (fall 2021), and its focus is on preparing the intervention (i.e., one of the teaching materials developed in this project) and identifying, selecting, revising and developing the relevant assessments for the pre- and post-tests.

Example 2: “Educational assessments of the 21st century: measuring and understanding students’ adaptability in complex problem solving situations” (ADAPT21)

The ADAPT21² project aims at assessing students’ adaptability in complex problem-solving situations, building on the need for more knowledge about this as a critical 21st century competence. Adaptability is understood as the ability to adjust one’s thinking, behavior, and affect to novelty and change. Given the lack of performance-based assessments of such skills (Scherer, 2015), the project aims at developing such assessments by building on so-called “minimal complex systems”—computer-based assessments that capture the core cognitive processes involved in complex problem solving. These systems will present students with complex problems that are subject to novelty and change. Data from around 1000 students has been collected, including the logfiles of each subject, which do not only record the students’ responses for each task but also how they went about finding the solutions. The project is entering its last phase (fall 2021), focusing on data analysis and dissemination.

Example 3: “Modeling collaborative problem-solving data: new approaches for individual and team level inferences”

The project titled “Modeling collaborative problem-solving data: new approaches for individual and team level inferences”³ aims at developing a set of psychometric and statistical modeling approaches to the hierarchical nature of ColPS tasks. This is a collaborative project between two assessment units—the Centre for Educational Measurement at the University of Oslo (CEMO), Norway and the BEAR Center at UC Berkeley, USA—and is led by assistant professor Perman Gochyyev. This project builds on a background of increased focus on students’ collaborative problem-solving skills, and it gathers both individual and group-level datasets. The complexity of ColPS, along with the shortage of methods for processing and analyzing such data, underpins the focus of this project. For instance, in order to handle data from ColPS tasks properly, one needs to account for (1) the clustering of individuals within

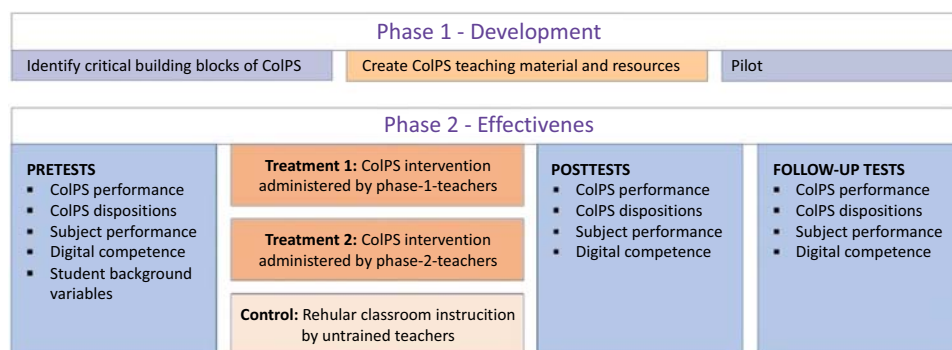


Fig. 3 Design of the TEACH21st project.

¹<https://teach21st.usn.no/>.

²<https://www.uv.uio.no/ceмо/english/research/projects/adapt21/>.

³Research Project (<https://www.researchgate.net/project/Modeling-collaborative-problem-solving-data-new-approaches-for-individual-and-team-level-inferences>) | Modeling Collaborative Problem-solving data: New Approaches.

teams, where individuals (i.e., students) are level-1 units and teams are level-2 units of analysis and (2) scoring items at different levels. One challenge lies in disaggregating individuals' cognitive (e.g., science) or noncognitive (e.g., perspective taking) skills under circumstances wherein their performance partially depends on their group's performance. Also, challenging is the extraction of cognitive and noncognitive skills at the group level after teasing out individual performances. Given that most ColPS designs require an elaborate approach to modeling the measurement data arising from their hierarchical structures, there are no readily available methods for researchers in this field. In particular, there is a need for testing cross-level differential item functioning (DIF) in the item parameters—that is, to what extent item difficulties and discriminations vary between teams. This project was initialized in 2020 but delayed because of the Covid-19 pandemic. The project is thus still running, and its results will be disseminated over time.

The transformative aspects of assessment in the 21st century

We end this chapter by raising some general reflections on the opportunities and possible pitfalls in understanding how we can and cannot assess 21st century skills, along with some future directions for this field of research.

Knowing that assessment is one of the most powerful determinants of practice in the classroom (Griffin et al., 2012, p. vi) and that it must be developed in order to “improve learning and support systemic education reform” (Kozma, 2009, p. 13), the extent to which our present educational systems manage to facilitate the development of 21st century skills curricula and assessment remains in question. However, as described earlier in this chapter, the attention given to the assessment of 21st century skills is rather superficial in both the leading international policy initiatives and national curricula. Along these lines, Scheuerman and Björnsson (2009) point out in their report focusing on new approaches to skills assessment that “existing models of assessment typically fail to measure the skills, knowledge, attitudes and characteristics of self-directed and collaborative learning that are increasingly important for our global economy and fast changing world. New assessments are required that measure these” (2009, p. 13).

The transformative potential of digital technologies for assessment was also raised by the Assessment and Teaching of 21st Century Skills project. These potentials are illustrated in Fig. 4 (Binkley et al., 2012, p. 27).

Fig. 4 provides a representation of the contrast between two key aspects of change and transformation as defined by this project—the business efficiency gains versus the educational transformation gains. The lower-left quadrant represents traditional assessments, which 10 years ago were typically paper-based but are now more often computer-based. Many school- and higher education-based assessments are of this type. Moving from the lower-left to the lower-right quadrant represents a migratory strategy by which paper-based assessments are shifted to a screen-based environment, which we see more examples of now. Delivery is more efficient, but the assessments are qualitatively unchanged. In contrast, moving to the upper-right quadrant is a transformational strategy in which technology is used to support innovative assessment designed to influence innovation in curriculum design and learning. The aim for the projects presented in this chapter is to reach this transformative stage combining innovative and technology-rich assessment. However, this has not reached a systemic transformation in any countries so far.

Most research on the use of digital technologies for assessment has dealt with its improvement of assessing traditional skills—improvement in the sense that such technologies have the potential for the large-scale delivery of tests and scoring procedures, easily providing the learner accessible feedback on their performance (Scheuerman and Björnsson, 2009; Siarova et al., 2017).

What concluding remarks can we draw based on the presentation in this chapter? We have argued that we need to understand how 21st century skills are defined and the key frameworks that contextualize these developments internationally, as well as how

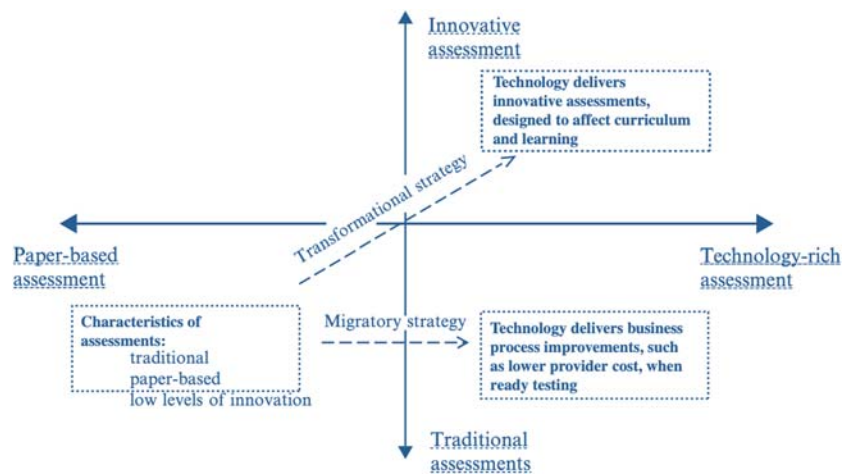


Fig. 4 The dimensions of computer-based assessment innovations (Binkley et al., 2012, p. 27).

they are implemented at the national level. We believe there are several pitfalls that can move us in the wrong direction, especially if we create unfounded expectations for the ways digital technologies will change assessment systems by themselves or else fail to develop substantial research insights into how we can operationalize and measure not only “easy-to-measure” areas but also diverse 21st century skills. There is also a challenge of representation from historically underrepresented countries in large-scale assessment initiatives (Anjum et al. this Encyclopeida).

However, as we have emphasized in this text, there are several interesting initiatives in the making, internationally and nationally, that document the possibilities that exist in assessing 21st century skills. Over the last decade, and especially in the last couple of years, several examples of projects and initiatives have emerged that take advantage of developments in digital technologies and produce better understandings of complex skill-sets, using them to create items and instruments better suited to assessments and to building a knowledge base for future developments. A key possibility is also provided by recent curriculum developments in several countries around the world, highlighting the instruction of 21st century skills and key competences for the future as important dimensions of how education systems must now proceed in order to tackle future challenges in our societies, forwarding the skills and competences needed by young people growing up today.

As stated above, issues surrounding assessing and defining 21st century skills are part of a much broader discussion about systemic change in our education systems and models for that change. Assessment is a key element in any discussion about educational change, as well as a premise for the integration of teaching and learning 21st century skills in the classroom. If we do not further develop our assessment systems, it will be hard to accomplish any change in education at a systemic level. That is why the growing interest in and orientation toward 21st century skills have issued some directions for fundamental changes in which assessment is a key component.

In order to move in the direction argued for above, we believe there is a great need for more interdisciplinary approaches and projects. In order to develop good and sustainable tests and measurement instruments, we need to combine insights from fields such as psychology, technology, education, assessment/psychometrics and entrepreneurship/innovation. There is also a need to develop international large-scale projects targeting the assessment of 21st century skills in order to further collect experiences and insights in this field. There is a danger that many projects may stay as national initiatives or rely on a concentration of just a few skills that are defined as easier to measure. We need to produce approaches that better combine teaching and assessment that can support the learning processes of all students, a change already indicated by new curriculum developments. There are also opportunities for teachers and researchers to better collect and analyze the data that is now created by the digital footprints of students, and these areas will probably develop in a more robust and substantial way than we have seen until now in the years to come.

References

- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries. OECD Education Working Papers, No. 41. OECD Publishing. <https://doi.org/10.1787/218525261154>.
- Anderson, R., 2008. Implications of the information and knowledge society for education. In: Voogt, J., Knezek, G. (Eds.), *International Handbook of Information Technology in Primary and Secondary Education*. Springer, Berlin Heidelberg New York.
- Arjomand, G., Erstad, O., Gilje, O., Gordon, J., Kallunki, V., Kearney, C., Rey, O., Siewiorek, A., Vivitsou, M., von Reis Saari, J., 2013. Literature Review: Key Competence Development in School Education in Europe. Available at: <http://keyconet.eun.org/literature-review>. (Accessed 26 June 2014).
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., Rumble, M., 2012. Defining twenty-first century skills. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht, pp. 17–66.
- Boitshwarelo, B., Reedy, A.K., Billany, T., 2017. Envisioning the use of online tests in assessing twenty-first century learning: a literature review. *Res. Pract. Technol. Enhanc. Learn.* 12, 16. <https://doi.org/10.1186/s41039-017-0055-7>.
- Care, E., Griffin, P., Wilson, M., 2018a. *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-65368-6>.
- Care, E., Kim, H., Vista, A., Anderson, K., 2018b. Education System Alignment for 21st Century Skills: Focus on Assessment. Available at: <https://www.brookings.edu/wp-content/uploads/2018/11/Education-system-alignment-for-21st-century-skills-012819.pdf>.
- Cui, Y., Guo, Q., Leighton, J.P., Chu, M., 2020. Log data analysis with ANFIS: a fuzzy neural network approach. *Int. J. Test.* 20 (1), 78–96.
- Dann, R., 2002. *Promoting Assessment as Learning: Improving the Learning Process*. Routledge Falmer, London.
- Deakin Crick, R.D., Broadfoot, P., Claxton, G., 2004. Developing an effective lifelong learning inventory: the ELLI project. *Assess. Educ.* 11, 247–318.
- Erstad, O., 2008. Changing assessment practices and the role of IT. In: Voogt, J., Knezek, G. (Eds.), *International Handbook of Information Technology in Primary and Secondary Education*. Springer Publishing Company, pp. 181–194.
- Erstad, O., Kjällander, S., Järvelä, S., 2021. Facing the challenges of “digital competence”: a Nordic agenda for curriculum development for the 21st century. *Nord. J. Digit. Lit.* 16 (2), 77–87.
- European Parliament, 2006. Recommendation of the European Parliament and the Council of 18 December 2006 on key competences for lifelong learning (2006/962/EC).
- Finnish National Agency for Education, 2016. National Core Curriculum. Ministry of Education, Helsinki.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt, E., 2014. *Preparing for Life in a Digital Age: The IEA International Computer and Information Literacy Study International Report*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-14222-7>.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Duckworth, T.D., 2020. *Preparing for Life in a Digital Age: IEA International Computer and Information Literacy Study 2018*. IEA Springer, Cham. <https://doi.org/10.1007/978-3-030-38781-5>.
- Geisinger, K.D., 2016. 21st century skills: what are they and how do we assess them? *Appl. Meas. Educ.* 29 (4), 245–249. <https://doi.org/10.1080/08957347.2016.1209207>.
- Griffin, P., Care, E. (Eds.), 2015. *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer, Dordrecht. <https://doi.org/10.1007/978-94-017-9395-7>.
- Griffin, P., McGaw, B., Care, E. (Eds.), 2012. *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht. <https://doi.org/10.1007/978-94-007-2324-5>.
- Halinen, I., 2018. The new educational curriculum in Finland. In: Matthes, M., Pulkkinen, L., Clouder, C., Heys, B. (Eds.), *Improving the Quality of Childhood in Europe*, vol. 7. Alliance for Childhood European Network Foundation, Brussels, pp. 75–89.

- Hilt, L.T., Riese, H., Sørdeide, G.E., 2019. Narrow identity resources for future students: the 21st century skills movement encounters the Norwegian education policy context. *J. Curric. Stud.* 51 (3), 384–402. <https://doi.org/10.1080/00220272.2018.1502356>.
- Hölscher, K., Wittmayer, J.M., Lorbach, D., 2018. Transition versus transformation: what's the difference? *Environ. Innov. Soc. Transit.* 27, 1–3. <https://doi.org/10.1016/j.eist.2017.10.007>.
- Joynes, C., Rossignoli, S., Fenyiwa Amonoo-Kuofi, E., 2019. 21st Century Skills: Evidence of Issues in Definition, Demand and Delivery for Development Contexts. K4D Helpdesk Report. Institute of Development Studies, Brighton, UK. Available at: https://assets.publishing.service.gov.uk/media/5d71187ce5274a097c07b985/21st_century.pdf.
- Kozma, R., 2009. Transforming education: assessing and teaching 21st century skills assessment, call to action. In: Scheuerman, F., Björnsson, B. (Eds.), *The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing*. Office for Official Publications of the European Communities, Luxembourg. <https://doi.org/10.2788/60083>.
- Krumsvik, R.J., 2008. Situated learning and teachers' digital competence. *Educ. Inf. Technol.* 13, 279–290. <https://doi.org/10.1007/s10639-008-9069-5>.
- Lee, E.Y.C., Chan, C.K., van Aalst, J., 2006. Students assessing their own collaborative knowledge building. *Int. J. Comput. Support. Collab. Learn.* 1, 57–87.
- Lin, S.S.J., Liu, E.Z.F., Yuan, S.M., 2001. Web-based peer assessment: feedback for students with various thinking styles. *J. Comput. Assist. Learn.* 17, 420–432.
- Looney, J., Michel, A., 2014. KeyCoNet's Conclusions and Recommendations for Strengthening Key Competence Development in Policy and Practice: Final Report. European Schoolnet, Brussels. Available at: http://keyconet.eun.org/c/document_library/get_file?uid=78469b98-b49c-4e9a-a1ce-501199f7e8b3&groupld=11028. (Accessed 12 October 2021).
- Mahlow, C., Hediger, A., 2019. Digital transformation in higher education—buzzword or opportunity? *eLearn* 5, 13. <https://doi.org/10.1145/3329488/3331171>.
- Murchan, D., Siddiq, F., 2021. A call to action: a systematic review of ethical and regulatory issues in using process data in educational assessment. Accepted for publication in *Large-Scale Assess. Educ.* <https://doi.org/10.1186/s40536-021-00115-3>.
- NOU, 2015. The school of the future—renewal of subjects and competences. Norges Offentlige Utredninger 8. Available at: <https://www.regjeringen.no/contentassets/da148fec8c4a4ab88daa6b77a700292/en-gb/pdfs/nou201520150008000engpdfs.pdf>. (Accessed 31 October 2017).
- Nunes, C.A.A., Nunes, M.M.R., Davis, C., 2003. Assessing the inaccessible: metacognition and attitudes. *Assess. Educ.* 10, 375–388.
- OECD, 2013. Student assessment: putting the learner at the centre. In: *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264190658-7-en>.
- O'Neil, H.F., Chuang, S., Chung, G.K.W.K., 2003. Issues in the computer-based assessment of collaborative problem solving. *Assess. Educ.* 10, 361–374.
- Pöysä-Tarhonen, J., Care, E., Awwal, N., Häkkinen, P., 2018. Pair interactions in online assessments of collaborative problem solving: case-based portraits. *Res. Pract. Technol. Enhanc. Learn.* 13, 1. <https://doi.org/10.1186/s41039-018-0079-7>.
- Russell, M., Goldberg, A., O'Connor, K., 2003. Computer-based testing and validity: a look into the future. *Assess. Educ.* 10, 279–294.
- Rychen, D.S., Salganik, L.H. (Eds.), 2001. *Defining and Selecting Key Competencies*. Hogrefe & Huber, Seattle.
- Salganik, L.H., Rychen, D.S., Moser, U., Konstant, J.W., 1999. Projects on competencies in OECD contexts: analysis of theoretical and conceptual foundations. Neuchâtel 1–50.
- Scherer, R., 2015. Is it time for a new measurement approach? A closer look at the assessment of cognitive adaptability in complex problem solving. *Front. Psychol.* 6, 1664.
- Scherer, R., Tondeur, J., Siddiq, F., 2017. On the quest for validity: testing the factor structure and measurement invariance of the technology-dimensions in the Technological, Pedagogical, and Content Knowledge (TPACK) model. *Comput. Educ.* 112, 1–17. <https://doi.org/10.1016/j.compedu.2017.04.012>.
- Scheuerman, F., Björnsson, J., 2009. *The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing*. European Commission Joint Research Centre. <https://doi.org/10.2788/60083>.
- Schneuwly, B., Vollmer, H.J., 2018. Bildung and subject didactics: exploring a classical concept for building new insights. *Eur. Educ. Res. J.* 17 (1), 37–50. <https://doi.org/10.1177/1474904117696096>.
- Scoular, C., Care, E., 2019. A generalized scoring process to measure collaborative problem solving in online environments. *Educ. Assess.* 24 (3), 213–234. <https://doi.org/10.1080/10627197.2019.1615372>.
- Scoular, C., Care, E., 2020. Monitoring patterns of social and cognitive student behaviors in online collaborative problem solving assessments. *Comput. Hum. Behav.* 104. <https://doi.org/10.1016/j.chb.2019.01.007>.
- Sefton-Green, J., 2021. Towards platform pedagogies: why thinking about digital platforms as pedagogic devices might be useful. *Discourse*. <https://doi.org/10.1080/01596306.2021.1919999>.
- Siarova, H., Sternadel, D., Mašidlauskaitė, R., 2017. *Assessment Practices for 21st Century Learning: Review of Evidence (NESET II Report)*. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2766/71491>.
- Siddiq, F., 2016. *Assessment of ICT Literacy. A Comprehensive Inquiry of the Educational Readiness for the Digital Era*. Doctoral Dissertation, University of Oslo. <http://urn.nb.no/URN:NBN:no-56585>.
- Siddiq, F., Scherer, R., 2017. Revealing the processes of students' interaction with a novel collaborative problem solving task: an in-depth analysis of think-aloud protocols. *Comput. Hum. Behav.* 76, 509–525. <https://doi.org/10.1016/j.chb.2017.08.007>.
- Siddiq, F., Gochyyev, P., Wilson, M., 2017. Learning in digital networks—ICT literacy: a novel assessment of students' 21st century skills. *Comput. Educ.* 109, 11–37. <https://doi.org/10.1016/j.compedu.2017.01.014>.
- Siddiq, F., Scherer, R., Tondeur, J., 2020. Developing Students' Collaborative Problem-Solving Competence—(TEACH21st): Towards a Linking Science. *The European Association for Research on Learning and Instruction*. <https://doi.org/10.13140/RG.2.2.24949.50409>.
- Sijlander, Kivela, A., Sutinen, A. (Eds.), 2012. *Theories of Bildung and Growth: Connections and Controversies Between Continental Educational Thinking and American Pragmatism*. Sense Publishers, Rotterdam.
- Tahirsylaj, A., Sundberg, D., 2020. The unfinished business of defining competences for 21st century curricula—a systematic research review. *Curric. Perspect.* 40, 131–145. <https://doi.org/10.1007/s41297-020-00112-6>.
- Thurstone, L.L., 1959. *The measurement of values*. The University of Chicago Press, Chicago.
- Tondeur, J., Aesaert, K., Pynoo, B., van Braak, J., Fraeyman, N., Erstad, O., 2016. Developing a validated instrument to measure preservice teachers' ICT competencies: meeting the demands of the 21st century. *Br. J. Educ. Technol.* 48. <https://doi.org/10.1111/bjet.12380>.
- van Laar, E.A., van Deursen, A.J.A.M., van Dijk, J.A.G.M., de Haan, J., 2017. The relation between 21st-century skills and digital skills: a systematic literature review. *Comput. Hum. Behav.* 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>.
- Voogt, J., Roblin, N.P., 2012. A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J. Curric. Stud.* 44 (3), 299–321. <https://doi.org/10.1080/00220272.2012.668938>.
- Wang, F., Hannafin, M.J., 2005. Design-based research and technology-enhanced learning environments. *Educ. Technol. Res. Dev.* 53 (4), 5–23.
- Designing for change in networked learning environments. In: Wasson, B., Ludvigsen, S., Hoppe, U. (Eds.), 2003. *Proceedings of the International Conference on Computer Support for Collaborative Learning*. Computer-Supported Collaborative Learning Series, vol. 2. Kluwer, Dordrecht, the Netherlands.
- White Paper, 2016. *Subjects—Specialization—Understanding*. Meld.St. 28 (2015–2016). Ministry of Education.
- Williamson, B., 2013. *The Future of the Curriculum: School Knowledge in the Digital Age*. The MIT Press.

Educational assessment and the rise of immigrant students

Nina Jude, Ruprecht-Karls-Universität Heidelberg, Institut für Bildungswissenschaft, Heidelberg, Germany

© 2023 Elsevier Ltd. All rights reserved.

Immigrant students in ILSA—indicators for policy and practice	256
Definition of immigrant status	257
Educational outcomes of immigrant students	258
Education policy and equality of opportunity	259
New challenges and outlook	260
References	261

In our globalized world, immigration and temporarily movement between countries is an ever-growing phenomenon and an essential part of our society. Movement between countries happens for different reasons and with different prospects (IOM, 2019). While education is the driving factor usually for students in the tertiary sector, it can be assumed that immigrants also consider future prospects of their children including educational opportunities as being of major importance. In the OECD countries, the percentage of students with an immigrant background, or whose parents emigrated from another country, is continuously rising (OECD, 2019a).

Education systems are facilitators in the process of integration for all students. As societies are becoming more heterogeneous, education needs to adapt when catering to students from all over the world. Moreover, increasing global mobility has been here for several decades, which leads to a growing proportion of immigrant's children and their grandchildren in all nations (Sweetman and Ours, 2015). While immigration is an obvious part of all societies, inclusion of students with an immigration background still seems to be a challenge for many education systems. Indicators of system monitoring show significant differences in educational outcomes when comparing native and immigrant students. These differences also impact the prospects of the adult population (OECD, 2019a).

Consequently, education systems must face the fact that dealing with a heterogeneous student body might not come by itself, or as a byproduct of a heterogeneous society.

This chapter focusses on latest findings regarding immigrant students' educational attainment and the growing importance of immigration as a topic in education policy, as emerged from the data and results of International Large Scale Assessment (ILSA). It informs about different definitions of immigrant students, summarizes major findings regarding education outcomes, and highlights possible mechanisms on the system and school level for promoting inclusive education systems and school settings.

Immigrant students in ILSA—indicators for policy and practice

Indicators addressing the educational attainment of different groups in a society are relevant when it comes to evaluating the quality aspects of an education system. Policy makers need to be informed about progress and developments toward specific goals, for example the UN sustainable development goals (United Nations, 2015). These goals state—among others—equality for all and quality in education. Countries acknowledging these goals are asked to state their progress, using comparable indicators to track their success. Indicators included in ILSA can be used to monitor progress toward these goals (IEA and UNESCO, 2020; Kaplan and Lee, 2018). Moreover, ILSA inform about the equality of educational opportunities and allow estimating the effects of the family background, including the immigrant status, on educational outcomes (Hanushek and Wössmann, 2011).

Education policy needs to be aware of the development of these indicators as schooling in the host country is obviously the key factor for educational success, which in turn is a prerequisite for labor market prospects. When it comes to evidence informed policy focusing on immigrant students, ILSA are a state-of-the-art source for data on educational outcomes of different student populations. They offer indicators on change over time and provide an insight into teaching and learning contexts which can be related to specific countries policies, including immigration policies (for an overview on potential policies see de la Rica et al., 2015).

While workforce migration and student population in tertiary education is often in the focus of national monitoring, key indicators and policy regulations, little is known about immigrant students on the primary or secondary level (K-12) on a global scale. The diversity of datasets available—and in some cases the lack thereof—might pose a challenge for policy makers if they want to determine the needs for adapting primary and secondary education to a growing and more heterogeneous society. With the OECDs migration outlook, the EUs approach on monitoring, and the Global Education Monitor, different sources of information have been merged and provide a comprehensive overview (EU, 2017; Unesco, 2018; OECD, 2020a).

Many parts of these reports are based on data from ILSA. Large scale assessments as valid instruments for monitoring students' achievement in different competence areas and in education systems around the world regularly include different indicators of the student's heritage. The relevance of assessing indicators of immigration background arises from empirical findings which highlight

significant gaps between immigrant and non-immigrant students when it comes to educational outcomes and performance (see below).

The most prominent ILSA which continuously provide indicators of achievement, context factors and immigration status of students are the Trends in International Mathematics and Science Study (TIMSS), the Program for International Student Assessment (PISA), and the Program for International Reading Literacy (PIRLS). TIMSS and PIRLS use grade-based samples of fourth and eighth grade students while PISA targets 15-year-old students. PISA assesses competencies in reading, mathematics, and science as well as in innovative fields like problem-solving or global competence. PIRLS measures reading competencies of primary school children and TIMSS focusses on skills in mathematics and science in the fourth and eighth grade. Both PIRLS and TIMSS use a curriculum-based approach in their competence frameworks, while PISA focusses on a broader concept of literacy (for an overview of differences and similarities see [Klieme, 2016](#)). In addition to the tests assessing cognitive skills, all studies use questionnaires for students to capture attitudes and context factors of learning which are related to the educational outcomes ([Jude, 2016](#); [Jude and Kuger, 2018](#)).

The results of ILSA regularly raise awareness for the situation of immigrant students and the need for respective education policies. At present, ILSA offer a rich system of indicators which can be used to monitor the progress countries make in integrating immigrant students. Moreover, these studies allow to analyze context factors on different levels of the education system which support or hinder the success of immigrant students. These context factors include the input into the education system regarding resources as well as the processes on school level. Regular monitoring of respective indicators can point education policy to challenges countries will need to face in the future ([OECD, 2020d](#)). For many countries, studies like TIMSS or PISA were the first ever datasets to exist that included indicators of students heritage, showing in many cases a significant gap between native and immigrant students regarding mean values of educational outcomes. The results differ between countries and groups of immigrants within countries, still the overall discrepancies are of major importance. Interestingly though, gaps are not necessarily present in non-cognitive indicators and attitudes.

It needs to be stated that immigrant students are by no means a homogenous group. Consequently, different indicators need to be considered when describing challenges and prospects of immigrant students and defining an immigrant status.

Definition of immigrant status

The immigration status of students participating in ILSA is typically assessed with a questionnaire given to all the students taking part in the assessment. These questionnaires ask the students where they themselves and their parents were born. The indicators register if the students were born in the country where they currently attend school (host country), or in another country. In some studies, questionnaires for parents of the participating students assessing their heritage are added to the database.

Using these indicators, the migration status of students can be classified into different categories: students born in the country where they currently attend schooling and students born abroad in relation to the country where they currently attend schooling. Students born abroad are also called first-generation immigrants. In addition, students born in the host country can still be attributed an immigration background if one or both of their parents were born abroad. They are called second-generation immigrants. Some ILSA add questions asking about the specific country of origin which can be used to track migration movements as well as differences in educational achievement which might be related to specific heritage groups ([OECD, 2019b](#)).

As immigration movements change over time, so does the share of first and second generation students within a country. In countries with a long history of immigration, more students of a second, or even third generation with immigration background will attend schools and universities, whereas in countries where immigration is a relatively new phenomenon, the students themselves will have migrated with their parents. In recent years, new waves of immigration have started, bringing with them once again first-generation students and their families ([OECD, 2020a](#)). As it can be assumed that the educational needs of these different groups vary, tracking the development of these indicators is crucial when education systems want to adapt their resources.

Another indicator of immigration background assessed with questionnaires in ILSA is the language spoken in the home of the students. Questions in the questionnaire ask the students if the language they speak in their family most of the time if the language of the instruction in the country or another language. This indicator can be used for in-depth analysis in addition to the country of birth, as many immigrants speak their heritage language in the families and are thus often multi-lingual, whereas education in most countries is mono-lingual. As language competence is the key factor when it comes to educational success, and most subject areas even require their own technical language, education systems need to enable support structures for students whose native language is not the language of instruction in the country they are currently living in ([Li, 2013](#); [Hoff, 2018](#)).

Some ILSA assess the age in which the students moved to the country where they currently attend schooling. This information can help identify areas of action for policy which can differ depending on the respective institutions on different levels of the education systems, for example kindergarten and pre-school, primary school, and secondary education. Moreover, tracking the success of student groups which migrated at different time points in their educational career can give insights where integration in the host country works well ([OECD, 2012a](#)). When looking at factors influencing immigrant students' performance in ILSA compared to those of non-immigrants, the age of arrival in the host country indeed needs to be considered. On average, performance is correlated with the age of arrival: The later students arrive in the host country, the lower they perform in assessments. However, the effect of arrival age differs between countries, as covariates like the language barrier, any differences between the school-systems like tracking or immersion practices, as well as differences in educational standards need to be taken into account ([OECD, 2012a](#)).

ILSA data, though it is currently the most prominent indicator system in this area, only provides limited information on the sample of immigrant students. Even though these studies comprise a representative sample of students on the country level for each participating nation, the prerequisite for being tested is that students have at least received one year of instruction in the language of assessment in the host country (OECD, 2019b). Accordingly, students who just arrived in the country and do not have sufficient language competencies are not represented in the samples. It can thus be assumed that effects of immigration might be underestimated when relying solely on ILSA data.

Overall, the immigrant population is becoming more and more diverse, i.e., the number of different heritage countries is rising with the increasing phenomenon of refugee children attending school in various countries (OECD, 2015a; Dryden-Peterson, 2015). Thus it could be argued that immigrant status should not necessarily be a defining attribute for those students which were born in the country, i.e., second-generation immigrants. This would mean that they inherit an attribute which does not really belong to them but has been carried over from parents, sometimes even grandparents. Still the phenomenon exists that we can see differences when it comes to educational outcomes of immigrant students even in the second generation, and in many cases also when only one parent was born abroad.

Several studies have investigated and compared findings from ILSA regarding the performance gaps. As different surveys vary in their definition and assessment of competencies and focus on different age groups, they can be used to analyze the overall phenomenon from different perspectives.

Educational outcomes of immigrant students

Ideally, education systems aim for fair opportunities for all students. Fair opportunities imply that educational outcome and success only depends on one's abilities and is not influenced by unjust distribution of resources or unequal possibilities. Thus, the challenge of system monitoring is to identify potential sources of these gaps and highlight the role of education policy and schools to overcome any systematic barriers which would lead to discrimination of specific groups of students. Students with an immigration history can be seen as a specific group of students despite their heterogeneity. Accordingly, system monitoring needs to analyze if differences in the educational success of native and immigrant students exist on a broad scale, and if this is due to systematic mechanisms in the education system.

The current discussion on the rise of immigrant students reflects contradictory findings when it comes to their educational outcome. While most studies strongly suggest that achievements of students with an immigration background on average are lower than those of the native students in the respective country, latest research points to considerable variation in these findings. The variations are reflected between the different host countries, between specific heritage countries, and also between school subjects and educational grade. Accordingly, even though performance gaps can be seen when looking at the average, they are not persistent across all conditions.

In most countries, it is the native students which on average perform best. First-generation immigrant students show lower performance and second-generation immigrant students are still outperformed by native students but show higher performance than their first-generation peers. It is important to note that these results only refer to the average across all students with an immigration background on a global scale. The gap is however not homogenous across countries, differs between specific subject areas and has been found to be related to different covariates (Andon et al., 2014). In some countries, the gap closes over time and in a few exceptions, immigrant students outperform the native ones (Alieva et al., 2018; OECD, 2016).

While significant differences in mean performance between native and immigrant students can be found in all ILSA, across countries and subjects, there is also a substantial variation to be noted. This leads to the conclusion that country specific policies could mediate the correlation (Andon et al., 2014). While immigrant students on average show lower outcomes compared to native students in mathematics and science (Hastedt, 2016), the variance of achievement is higher for students with an immigration background than among non-immigrant students (Schnepf, 2008; OECD, 2012a).

Performance can also be analyzed regarding percentiles, e.g., the share of immigrant students among the 25th and 75th percentile marking the group of low and high performers. For most but not all countries, immigrant students are over-represented among low achievers and under-represented among high achievers (OECD, 2012). On average in the OECD countries, the proportion of low performers among immigrant students is twice as high as among native students. The opposite is true for the highest competence levels: Here, the share of top-performing native students is almost twice as high as the proportion of their immigrant peers. This means that on average of all students with an immigration background, half of them do not reach basic levels of competence at the age of 15 (Schnepf, 2004; OECD, 2018).

An important aspect of educational prospects is the language immigrant students speak at home with their families. On average in the OECD countries, more than 60% of first generation immigrants and more than 1/3 of the second generation speak their heritage language at home which is different from the language of instruction in the host country (OECD, 2016). Regarding educational performance, the gap is significantly smaller between migrant and native students if the host-country language is spoken at home, still language proficiency seems to have a stronger impact on achievement in, for example, reading assessments than in the subject of mathematics (Schnepf, 2007; Cordero and Mateos, 2021).

The educational and economical level of the country of migration also reflects in the educational achievement in the host country—immigrant students arriving from a country with high educational achievements on average show a smaller achievement

gap in the host country (Gianellie and Rappalini, 2016). Thus the differences between the heritage and the host country need to be taken into account, as the heritage education system reflects in the students' performance also in the host country (Arikan et al., 2020).

While a substantial influence of the family background can be seen in educational performance for all students independent of their immigration background, this influence varies between countries (for an overview see Hanushek and Wössmann, 2011). Still the question remains if differential effects can be found when it comes to the impact of the family (Sirin, 2005; Kim et al., 2019). In their meta-analysis, Kim et al. (2020) show that parents proximal engagement, i.e., supporting students psychologically and getting involved in their learning, had stronger effects than the overall family background in immigrant families. Parents' knowledge about the host countries education system, their involvement and educational aspirations are also important when it comes to a successful integration (Hamish et al., 2016). It has been shown that the socio-economic status of students plays a crucial role when it comes to inclusion and can influence social cohesion in school and society. Consequently, current studies controlling for variation in the immigrant students background, find that performance gaps are reduced or even vanish when controlling the socio-economic background (Arikan et al., 2020).

When describing student populations with an immigration background, especially when comparing their educational achievements between countries, we need to be aware of determinative factors on the system level: Immigration policies between countries vary widely, along with policies in admitting refugees, which can reflect in the performance of immigrant students (OECD, 2015b). Consequently, cultures and socio-economic backgrounds of the immigrant population vary widely between and within countries, as do the education policies and school-level practices (Buchmann and Parrado, 2006). When looking at trend in performance gaps over time, while the differences between native and immigrant students still are significant, a decline in the overall performance gap can be seen in several countries (Riederer and Verwiebe, 2015). Thus, a closer look at policies on system and school level can reveal factors which support immigrant students' educational success.

Education policy and equality of opportunity

To guide education policy, it is importance to identify factors that support immigrant students. A topic that has widely been discussed in ILSA is the question of resilience and students well-being. Resilient students are those non-native students that are among the top-performers in the country of immigration, i.e., performing well against the odds (OECD, 2019a). When looking at the average achievement in reading across OCED countries, 17% of immigrant students can be found in the top-performing group, while in some countries, even 30% of immigrant student showed excellent reading scores (OECD, 2019a).

Factors supporting resilience can be found on the individual and the system level: Individual factors like self-efficacy, a growth-mindset and family support help when it comes to building resilience. Links between a growth-mindset and educational outcome are associated with a larger gain for immigrant students compared to their native peers (OECD, 2021).

On the school level, perceived teacher enthusiasm, co-operation in school and a more positive disciplinary climate in school help students to build resilience. When it comes to well-being, results show significant differences: High well-being at school is reported by fewer immigrant students than native students, and on the other hand, immigrant students are more likely to report feeling like an outsider at school (OECD, 2019a). When schools want to support students with heterogeneous backgrounds, their specific circumstances need to be recognized by offering support structures (Cordero and Mateos, 2021).

This leads to the interpretation that resilience and well-being are not effects of migration itself, but can be supported by schools and communities (OECD, 2018). There is no single factor which could be focused on when it comes to supporting educational participation and success of immigrant students. Rather, complex relationships between factors on the individual level (students and their families), the school level (teaching and learning), and the host countries society contribute to the overall educational outcomes. (Volante et al., 2019).

On the *system level*, most studies find effects of tracking and segregation policies on the achievement of immigrant students: Differentiation usually goes along with a larger gap depending on socio-economic background, which is in turn most often related to the immigration background (Bodovski et al., 2020). Tracking within the school system usually correlates negatively with equality of educational opportunities. It also might impact educational pathways of students: while on average, many studies show that immigrant students have higher educational expectation than their native peers, this difference is smaller in tracked systems. Bigger inequalities can be found in systems which implement early tracking, increasing inequalities in achievement between students with different family backgrounds (Contini and Cugnata, 2020; Dronkers et al., 2014). Moreover, the country of origin of the students seems to be of importance when selection mechanisms in highly tracked systems are in place (Alieva and Hildebrand, 2019).

However, latest studies point to the need for differentiation of effects when distinguishing between groups of immigrant students, for example taking into account the age of arrival in the host country and students language competence (Borgonovi and Ferrara, 2020), but also the age in which tracking occurs and the language spoken at home, still showing that late tracking enables equity and opportunities especially when students are less integrated. (Ruhose and Schwerdt, 2016).

On the *school level*, segregation is a strongly debated topic. In-school ability grouping seems to be related negatively to achievement with stronger effects for immigrant than non-immigrant student (He and Fischer, 2020). Cross-national studies show that school segregation and early tracking are usually related to lower equality of educational opportunities in the respective systems (Schütz et al., 2008). Stratification between schools is usually associated with a lower performance of immigrant students, especially when this is paired with segregation on the school level. Still, immigrant students can benefit from ability grouping within schools

(Teltemann and Schunk, 2016). Regarding the support of immigrant students, a high concentration of these students in disadvantaged schools is correlated with lower educational attainment (OECD, 2012b), one factor being that highly educated teachers usually can be found in high-SES schools (Rubinstein-Avila, 2016).

Schooling in all levels of education usually requires fundamental knowledge of the language of instruction. Hence, the success of immigrant students is related to their possibilities of acquiring *sufficient language proficiency* to follow instruction as well as their peers. Systematic policies to provide language learning opportunities for all immigrant student, independent of the age of arrival could support educational attainment, though language support especially in early childhood education is essential for later educational success (Entorf and Minoiu, 2005; Rubinstein-Avila, 2016). Policies which seems to positively influence inclusion can be found in countries which enforce education standards with centralized curricula and corresponding programs which reinforce language learning at all levels in the educational biography of students. This requires a skilled teacher workforce and adequate training for teachers in supporting a multi-lingual and heterogeneous student population (Christensen and Stanat, 2007).

Language competence can be seen as prerequisite for *participation and integration* in school and community, which reflects in a so-called sense of belonging. ILSA regularly show a correlation between sense of belonging and educational success. Data from the International Civic and Citizenship Education Study (ICCS) show that immigrant students report a lower sense of belonging on average (Ham et al., 2017), the PISA study finds that more than 40% of immigrant students report a weak sense of belonging, compared to 30% of native students (OECD, 2018). Still positive effects can be found when integration policies and anti-discrimination measures are in place which support social inclusion both for sense of belonging on the school level, but also on a national level when it comes to reducing the disparity between immigrant and non-immigrant students (Ham and Yang, 2020). Sense of belonging can be supported by integrating immigrant students in schools and fostering a positive relationship with teachers. Greater inclusion can be promoted by encouraging civic participation in school and the community (Rutkowski et al., 2014).

The background of students with refugee experience also needs to be considered, as requirements of these students differ significantly from other migration experiences. Disrupted or sporadic schooling in often traumatizing conditions influences not only performance, but also attitudes and educational expectations (Dryden-Peterson, 2015). Students' attitudes are related to the values shared in the society they are living in. The more inclusive and multicultural a society is, the higher the positive attitudes of students toward ethnic minorities (Kim and Byun, 2019).

Building an inclusive and welcoming setting for all students can be seen as a major challenge. *School climate and positive relationships* between students, their peers and teachers as well as civic engagement in schools and communities contribute to generalized trust, a comparable effect can be seen for ethnic diversity in the schools community (Choi and Lee, 2021). Openness and open classroom climate are positively associated with students civic and citizenship knowledge (Knowles and MacCafferty-Wright, 2015). Inclusive views regarding immigrants are correlated with the immigrant status of students; both the first- and the second-generation migrants are significantly more inclusive in their views on immigrants than native students. Mixed classrooms contribute positively to inclusive views of students as well as to positive attitudes regarding immigrants' rights (Janmaat, 2014; Isac et al., 2012).

For education policy as the target audience of ILSA results, it is important to learn not only about potential factors influencing differences between immigrant students and native students. Consequently, when discussing the rise of immigrant students in our education systems, the focus should not be only on cultural heritage, but rather on building society and schools that are able to support a growing heterogeneity amongst their student population in various aspects.

New challenges and outlook

Existing policies on system and school level were disrupted in the course of the pandemic that hit the education systems around the globe in early 2020. As a result, schools remained closed or switched to distance teaching for more than a school year (for an overview on national policies see OECD, 2020c). While it is yet to early to estimate the overall impact of this disruption on educational outcomes, current studies point to the fact that vulnerable students will face additional challenges: Students with an immigration background miss out on opportunities solely offered by schools, such as secure learning settings with the opportunity for immersion in the host countries language as well as participating in the society. Reducing these learning opportunities might impact not only educational outcome, but also the students' sense of belonging and educational aspiration. Moreover, learning support at home including technical resources is comparably limited and shows huge differences depending on the socio-economic and immigrant background (OECD, 2020b; Murat and Bonacici, 2020).

Education systems will need to face a challenge also in the years to come, when not only learning gaps, but also strains in regarding health and well-being in the student population will show their impact. Several ILSA have adapted their indicators to measure challenges and coping approaches of education systems and schools in dealing with the pandemic. The IES Responses to Educational Disruption Survey (REDS) will deliver insights on the country and school level policies which were implemented in dealing with disruption and distance learning. The PISA assessment in 2022 will be addressing differential effects of the pandemic on student learning and well-being across different education systems (Bertling et al., 2020).

Society and education systems will need to continue adapting to a changing and more heterogeneous population, and schools have a central function and responsibility not only to support students with an immigration background, but also to help develop intercultural skills in schools and society at large.

Immigration and worldwide movement of labor force are constant phenomenon in a globalized world. While some countries might be more affected than others, most countries will continue to be faced with challenges when it comes to adapting their education system to a more and more heterogeneous body of students in all levels of education. All stakeholders in education need to react to enable successful education for all students independent of their heritage.

As can be seen when analyzing ILSA data, the variation between students with an immigration background is high—variation refers to countries of origin, heritage language but also educational achievement. Differences in educational outcomes depend on heritage, countries of origin, language spoken at home but also on the education policies in the host countries. While in most countries immigrant students are outperformed by their native peers in standardized tests of achievement, they prosper in a limited number of countries. Moreover, immigrant students show higher competences when it comes to civic attitudes and intercultural aspects.

Meanwhile, when looking at the trend of ILSA data over the last 20 years, changes in the performance of immigrant student over time can be seen. In many countries, this can be attributed to education and social policies fostering integration. For some countries their overall average score point in ILSA improved solely because the educational outcomes of immigrant students increased, while those of the native students remained stable. There is a need for policies to regard immigrant students especially when it comes to teacher training and inclusion policies in schools. To overcome the specificity in the discussion of equity in education, monitoring should not only focus on minorities, but needs to broach the topic of inclusion for all independent of their heritage, without neglecting needs of groups of students which might face specific challenges. Education system which are truly global value and welcome heterogeneity in all aspects by adapting to the needs of all students to help them succeed.

References

- Alieva, A., Hildebrand, V., 2019. Does Prior Achievement Matter? Early Tracking and Immigrant Children in Europe. Institute of Socio-Economic Research, Luxembourg.
- Alieva, A., Hildebrand, V.A., Van Kerm, P., 2018. How Does the Achievement Gap between Immigrant and Native-Born Pupils Progress from Primary to Secondary Education? Institute of Socio-Economic Research, Luxembourg.
- Andon, A., Thompson, C., Becker, B., 2014. A quantitative synthesis of the immigrant achievement gap across OECD countries. *Large Scale Assess. Educ.* 2 (7). <https://doi.org/10.1186/s40536-014-0007-2>.
- Arikan, S., van de Vijver, F.J.R., Yagmur, K., 2020. Mainstream and immigrant students primary school mathematics achievement differences in European countries. *Eur. J. Psychol. Educ.* 35, 819–837.
- Bertling, J., Rojas, N., Alegre, J., Faherty, K., 2020. A Tool to Capture Learning Experiences during Covid-19: The PISA Global Crises Questionnaire Module. OECD Publishing, Paris.
- Bodovski, K., Munoz, I.G., Byun, S., Chykina, V., 2020. Do education system characteristics moderate the socioeconomic, gender and immigrant gaps in math and science achievement? *Int. J. Sociol. Educ.* 9 (2), 122–154.
- Borgonovi, F., Ferrara, A., 2020. Academic achievement and sense of belonging among non-native-speaking immigrant students: the role of linguistic distance. *Learn. Individ. Differ.* 81. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S1041608020300911>.
- Buchmann, C., Parrado, E.A., 2006. Educational achievement of immigrant-origin and native students: a comparative analysis informed by institutional theory. In: Baker, D.P., Wiseman, A.W. (Eds.), *The Impact of Comparative Education Research on Institutional Theory*. Emerald Group Publishing Limited, Bingley, pp. 335–366.
- Choi, S., Lee, M., 2021. Diversity as an opportunity or a challenge? A cross-national study of ethnic diversity and students generalized trust. *Int. J. Educ. Dev.* 80. <https://doi.org/10.1016/j.ijedudev.2020.102301>. Available at:
- Christensen, G., Stanat, P., 2007. *Language Policies and Practices for Helping Immigrants and Second-Generation Students Succeed*. Migration Policy Institute, Washington, DC.
- Contini, D., Cugnata, F., 2020. Does early tracking affect learning inequalities? Revisiting difference-in-differences modeling strategies with international assessments. *Large Scale Assess. Educ.* 8 (14). <https://doi.org/10.1186/s40536-020-00094-x>.
- Cordero, J.M., Mateos-Romero, L., 2021. Exploring factors related with resilience in primary education: evidence from European countries. *Stud. Educ. Eval.* 70. <https://doi.org/10.1016/j.stueduc.2021.101045>.
- de la Rica, S., Glitz, A., Ortega, F., 2015. Immigration in Europe: trends, policies, and empirical evidence. In: Chiswick, B.R., Miller, P.W. (Eds.), *Handbook of the Economics of International Migration*. Elsevier, Amsterdam, pp. 1303–1362.
- Dronkers, J., Levels, M., de Heus, M., 2014. Migrant pupils scientific performance: the influence of educational system features of origin and destination countries. *Large Scale Assess. Educ.* 2 (3). Available at: <https://largescaleassessmentsineducation.springeropen.com/track/pdf/10.1186/2196-0739-2-3.pdf>.
- Dryden-Peterson, S., 2015. *The Educational Experiences of Refugee Children in Countries of First Asylum*. Migration Policy Institute, Washington, DC.
- Entorf, H., Minioiu, N., 2005. What a difference immigration policy makes: a comparison of PISA scores in Europe and traditional countries of immigration. *German Econ. Rev.* 6 (3), 355–376.
- European Union, 2017. *Research for Cult Committee – Migrant Education: Monitoring and Assessment* [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2017/585903/IPOL_STU\(2017\)585903_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2017/585903/IPOL_STU(2017)585903_EN.pdf). (Accessed 13 July 2021).
- Giannelli, G.C., Rapallini, C., 2016. Immigrant student performance in Math: does it matter where you come from? *Econ. Educ. Rev.* 52, 291–304.
- Ham, S.-H., Yang, K.-E., 2020. The impact of anti-discrimination policies on immigrant youths sense of national belonging: an institutional account. *Pol. Stud.* <https://doi.org/10.1080/01442872.2020.1782869> [online]. (Accessed 13 July 2021).
- Ham, S.-H., Yang, K.-E., Cha, Y.-K., 2017. Immigrant integration policy for future generations? A cross-national multilevel analysis of immigrant-background adolescents' sense of belonging at school. *Int. J. Intercult. Relat.* 60, 40–50.
- Hanushek, E.A., Woessmann, L., 2011. The economics of international differences in educational achievement. In: Hanushek, E.A., Machin, S., Woessmann, L. (Eds.), *Handbook of the Economics of Education*. Elsevier, Amsterdam, pp. 189–200.
- Hastedt, D., 2016. *Mathematics Achievement of Immigrant Students*. Springer, Berlin.
- He, J., Fischer, H., 2020. Differential associations of school practices with achievement and sense of belonging of immigrant and non-immigrant students. *J. Appl. Dev. Psychol.* 66. <https://doi.org/10.1016/j.appdev.2019.101089> [online]. Available at: (Accessed 13 July 2021).
- Hoff, E., 2018. Bilingual development in children of immigrant families. *Child Dev. Perspect.* 12 (2), 80–86.
- IEA, UNESCO, 2020. *Measuring Global Education Goals: How TIMSS Helps. Monitoring Progress Towards Sustainable Development Goal 4 Using TIMSS. Trends in International Mathematics and Science Study*. Available at: <https://online.flippingbook.com/view/763810/>. (Accessed 13 July 2021).
- IOM, 2019. *World Migration Report 2020*. International Organization for Migration, Geneva.
- Isac, M.M., Maslowski, R., van der Werf, G., 2012. Native students attitudes towards equal rights for immigrants. A study in 18 European countries. *J. Soc. Sci. Educ.* <https://doi.org/10.4119/jsse-586> [online]. Available at: (Accessed 13 July 2021).

- Janmaat, J.G., 2014. Do ethnically mixed classrooms promote inclusive attitudes towards immigrants everywhere? A study among native adolescents in 14 countries. *Eur. Socio Rev.* 30 (6), 810–822.
- Jude, N., 2016. The assessment of learning contexts in PISA. In: Kuger, S., Klieme, E., Jude, N., Kaplan, D. (Eds.), *Assessing Contexts of Learning. An International Perspective*. Springer International Publishing, New York, pp. 39–51.
- Jude, N., Kuger, S., 2018. Questionnaire development and design for International Large-Scale Assessments (ILSAs): current practice, challenges, and recommendations. *National Academy of Education Workshop Series on Methods and Policy Uses of International Large-Scale Assessments (ILSAs)* [online]. Available at: <https://naeducation.org/wp-content/uploads/2018/04/Jude-and-Kuger-2018-FINAL.pdf>. (Accessed 13 July 2021).
- Kaplan, D., Lee, C., 2018. Optimizing prediction using Bayesian model averaging: examples using large-scale educational assessments. *Eval. Rev.* 42 (4), 423–457.
- Kim, H., Byun, S., 2019. Immigrant integration policy and native adolescents attitudes towards ethnic minorities: a comparative study of European Countries. *Multicult. Educ. Rev.* 11 (3), 172–188.
- Kim, S., Cho, H., Kim, L.Y., 2019. Socioeconomic status and academic outcomes in developing countries: a meta-analysis. *Rev. Educ. Res.* 89 (6), 875–916.
- Kim, Y., Mok, S.Y., Seidel, T., 2020. Parental influences on immigrant students achievement-related motivation and achievement: a meta-analysis. *Educ. Res. Rev.* 30. <https://doi.org/10.1016/j.edurev.2020.100327> [online]. Available at: (Accessed 13 July 2021).
- Klieme, E., 2016. TIMSS 2015 and PISA 2015. How Are They Related on the Country Level? DIPF Working Paper [online]. Available at: https://pisa.dipf.de/de/pdf-ordner/Klieme_TIMSS2015andPISA2015.pdf. (Accessed 13 July 2021).
- Knöwles, R., McCafferty-Wright, J., 2015. Connecting an open classroom climate to social movement citizenship: a study of 8th graders in Europe using IEA ICCS data. *J. Soc. Stud. Res.* 39 [online]. Available at: [10.1016/j.jssr.2015.03.002](https://doi.org/10.1016/j.jssr.2015.03.002). (Accessed 13 July 2021).
- Li, G., 2013. Immigrant language acquisition: an international review. In: Gold, S.J., Nawyn, S.J. (Eds.), *Routledge International Handbook of Migration Studies*. Routledge, London, pp. 271–282.
- Murat, M., Bonacini, L., 2020. Coronavirus Pandemic, Remote Learning and Education Inequalities. GLO Discussion Paper No. 679 [online]. Available at: <http://hdl.handle.net/10419/224765>. (Accessed 13 July 2021).
- OECD, 2012a. *Untapped Skills: Releasing the Potential of Immigrant Students*. OECD Publishing, Paris.
- OECD, 2012b. *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. OECD Publishing, Paris.
- OECD, 2015a. *Immigrant Students at School: Easing the Journey Towards Integration*. OECD Publishing, Paris.
- OECD, 2015b. *Helping Immigrant Students to Succeed at School – and Beyond*. OECD Publishing, Paris.
- OECD, 2016. *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD Publishing, Paris.
- OECD, 2018. *The Resilience of Students with an Immigrant Background: Factors that Shape Well-Being*. OECD Publishing, Paris.
- OECD, 2019a. *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OECD Publishing, Paris.
- OECD, 2019b. *PISA 2018 Technical Report*. OECD Publishing, Paris.
- OECD, 2020a. *International Migration Outlook 2020*. OECD Publishing, Paris.
- OECD, 2020b. *What Is the Impact of the Covid-19 Pandemic on Immigrants and Their Children?* OECD Publishing, Paris.
- OECD, 2020c. *A Framework to Guide an Education Response to the COVID-19 Pandemic of 2020*. OECD Publishing, Paris.
- OECD, 2020d. *Education at a Glance 2020: OECD Indicators*. OECD Publishing, Paris.
- OECD, 2021. *Skys the Limit: Growth Mindset, Students, and Schools in PISA*. OECD Publishing, Paris.
- Riederer, B., Verwiebe, R., 2015. Changes in the educational achievement of immigrant youth in western societies: the contextual effects of national (educational) policies. *Eur. Socio Rev.* 31 (5), 628–642.
- Rubinstein-Avila, E., 2016. Immigrant and refugee students across “receiving” nations: to what extent can educators rely on PISA for answers? *Clear. House* 89 (3), 79–84.
- Ruhose, J., Schwerdt, G., 2016. Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Econ. Educ. Rev.* 52, 134–154.
- Rutkowski, D., Rutkowski, L., Engel, L.C., 2014. Inclusive schooling: fostering citizenship among immigrant students in Europe. *Intercult. Educ.* 25 (4), 269–282.
- Schnepf, S.V., 2004. How Different Are Immigrants? A Cross-Country and Cross-Survey Analysis of Educational Achievement [online]. Available at: SSRN: <https://ssrn.com/abstract=621062>. (Accessed 13 July 2021).
- Schnepf, S.V., 2007. Immigrants educational disadvantage: an examination across ten countries and three surveys. *J. Popul. Econ.* 20 (3), 527–545.
- Schnepf, S., 2008. *Inequality of Learning Amongst Immigrant Children in Industrialised Countries*. Institute for the Study of Labor (IZA). IZA Discussion Papers.
- Schütz, G., Ursprung, H.W., Wößmann, L., 2008. Education policy and equality of opportunity. *Int. Rev. Soc. Sci.* 61 (2), 279–308.
- Sirin, S.R., 2005. Socioeconomic status and academic achievement: a meta-analytic review of research. *Rev. Educ. Res.* 75 (3), 417–453.
- Sweetman, A., van Ours, J.C., 2015. Immigration: what about the children and grandchildren? In: Chiswick, B.R., Miller, P.W. (Eds.), *Handbook of the Economics of International Migration*. Elsevier, Amsterdam, pp. 1141–1193.
- Teltemann, J., Schunck, R., 2016. Education systems, school segregation, and second-generation immigrants' educational success. *Int. J. Comp. Sociol.* 57 (6), 401–424.
- UNESCO, 2018. *Global Education Monitoring Report 2019: Migration, Displacement and Education – Building Bridges, Not Walls*. UNESCO, Paris.
- United Nations, 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development* [online]. Available at: https://www.unfpa.org/sites/default/files/resource-pdf/Resolution_A_RES_70_1_EN.pdf. (Accessed 13 July 2021).
- Volante, L., Klinger, D., Siegel, M., Yahia, L., 2019. Raising the achievement of immigrant students: towards a multi-layered framework for enhanced student outcomes. *Pol. Futures Educ. Internet* 17, 1037–1056.

Metaevaluation: key concepts explained[☆]

Lori A. Wingate, Zachary D. Tilton, and Michael A. Harnar, Western Michigan University, Kalamazoo, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of L.A. Wingate, Meta-evaluation: Purpose, Prescription, and Practice, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 765–774, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01614-6>.

Metaevaluation purposes	264
Formative metaevaluation	264
Summative metaevaluation	264
Ascriptive metaevaluation	265
Distinguishing metaevaluation from synthesis studies	265
Metaevaluation questions and criteria	265
Questions	265
Criteria	266
Program evaluation standards—joint committee on standards for educational evaluation	266
African evaluation guidelines—African Evaluation Association (AfrEA)	267
Evaluation standards—Aotearoa New Zealand Evaluation Association (ANZEA)	267
Quality standards for development evaluation—Organization for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC)	267
Norms and standards for evaluation in the United Nations systems—United Nations Evaluation Group (UNEG)	268
Idiosyncratic criteria and connoisseurship in metaevaluation	268
Metaevaluation methods	268
Metaevaluation methods recommended by stufflebeam	269
Metaevaluation methods recommended by Cook and Gruder	269
Metaevaluation methods recommended by Scriven	270
Metaevaluation practice	270
New York integrated learning system metaevaluation	271
Metaevaluation of a technology literacy program in underserved communities	271
Responsive metaevaluation of the Appalachia regional comprehensive center's evaluation of a capacity-building program	271
Metaevaluation of HIV/AIDS program evaluations using the African evaluation guidelines	272
Summary	272
Relevant websites	273
References	273

Evaluation is the determination of something's quality, value, or importance. *Metaevaluation* is the evaluation of evaluation.

The cradle of metaevaluation was educational evaluation. Orata's (1940) article "Evaluating Evaluation" in the *Journal of Educational Research* introduced the evaluation field to the general function of formal metaevaluation. In that article, Orata assessed the state of educational and student evaluation and made evaluative claims about the merits of well-known studies, including Ralph Tyler's landmark Eight-Year Study (Smith and Tyler, 1942). Nearly 30 years later, in 1969, Michael Scriven, a preeminent evaluation theorist, coined the term *metaevaluation* in an *Educational Products Report* article about evaluating educational products. In defining metaevaluation, Scriven described both a theoretical function and a practical one. The theoretical function of metaevaluation "involves the methodological assessment of the role of evaluation." The practical function "is concerned with the evaluation of specific evaluative performances" (Scriven, 1969, p. 36).

As defined by Scriven, the theoretical and methodological roles of metaevaluation parallel the two meanings of *meta* as summarized by Thomas (1984, p. 18): (1) "knowledge about" or "analysis of" and (2) "encompassing multiple aspects of instances of the phenomenon under consideration." Stufflebeam (2001, p. 183) argued that metaevaluation (i.e., engaging in evaluation at these meta-levels) "is a professional obligation of evaluators." Scriven (2009, 2011) extended this professional imperative to claim metaevaluation is an ethical imperative. He argued that metaevaluation is the "conscience of evaluation" and that the survival of evaluation as a discipline depends on it.

Metaevaluation's theoretical function is realized through two streams of activity: evaluation theorizing and research on evaluation.

[☆]Change History: May 2021. L. Wingate, Z. Tilton, and M. Harnar updated the text of the entire article.

Evaluation theorizing is broadly viewed as any writing on evaluation (Shadish, 1998). More specifically, it involves generating and refining “explanations of underlying fundamental issues such as the nature of evaluation purpose, valuing, evidence, use and so on” (Smith, 2010, p. 384). Various frameworks for conceptualizing the use and influence of evaluation and the numerous distinct approaches to evaluation fall under the umbrella of evaluation theorizing.

Research on evaluation refers to “any purposeful, systematic, empirical inquiry intended to test existing knowledge, contribute to existing knowledge, or generate new knowledge related to [evaluation]” (Coryn et al., 2016, p. 161). Such studies serve the theoretical function of metaevaluation by producing insights about the use and influence of evaluation within organizations and within society.

Scriven’s second dimension of metaevaluation—the evaluation of evaluation performances—is currently the dominant connotation of metaevaluation. This practice involves systemically examining specific evaluations to determine their quality.

This article focuses on practical metaevaluation as applied to program evaluations, specifically educational program evaluations. It introduces and explains metaevaluation’s purposes, metaevaluation questions and criteria, and methods for conducting metaevaluations. A few practice examples illustrate how these key concepts manifest in real-world metaevaluations of educational programs.

Metaevaluation purposes

Metaevaluations are conducted for the same reasons that evaluations are conducted. The chief purposes of evaluations are (1) to identify ways to improve an effort while it is underway or (2) to determine an effort’s overall quality and effectiveness to inform decisions about whether it should be continued, expanded, reduced, or discontinued. Additionally, evaluations may be conducted (3) to generate knowledge about the type of effort being examined, for example, to gain insights about program implementation. Respectively, these purposes are *formative*, *summative*, and *ascriptive*.

In primary evaluation, the effort in question is typically a coordinated set of activities intended to improve some aspect of the human condition or organizational functioning (i.e., a program). In metaevaluation, an evaluation itself is the focus. However, the primary purposes are the same, whether the focus is a program or its evaluation.

Formative metaevaluation

Formative metaevaluation occurs while an evaluation is underway. Its purpose is to identify opportunities to improve the primary evaluation. If no improvement is needed, then the formative metaevaluation serves as a confirmation of the evaluation’s quality. Such confirmations may enhance stakeholders’ confidence in the evaluation process and subsequent results.

In a formative metaevaluation, draft evaluation materials, such as evaluation plans, instruments, protocols, and reports, may be subject to review and feedback by a metaevaluator prior to finalization, implementation, or dissemination. Less tangible aspects of a primary evaluation may also be assessed in formative metaevaluation; for example, a metaevaluation may assess data collection and analysis procedures; evaluators’ competence; and the evaluation’s responsiveness to stakeholders’ information needs.

The most significant potential benefit of formative metaevaluation is that it can reveal deficiencies in the primary evaluation at a point in time when they can be remedied. In this way, a formative metaevaluation could prevent the primary evaluator from making serious errors such as overlooking critical stakeholder groups, reaching faulty conclusions due to poor data quality, or failing to demonstrate cultural competence. Some scholars have argued that the mere presence of a metaevaluator while an evaluation is underway reduces bias and increases rigor (Fitzpatrick et al., 2004; Scriven, 1975). That is, even evaluators are susceptible to the Hawthorne effect (Landsberger, 1958), a form of reactivity that takes place when those being evaluated perform better simply because they are aware they are being observed.

By surfacing potential problems early, a formative evaluation can increase a primary evaluation’s quality. It can also enhance an evaluation’s cost-effectiveness by ensuring resources are used effectively (Yarborough et al., 2011; Stufflebeam and Coryn, 2014).

Summative metaevaluation

A summative metaevaluation assesses the quality of a completed evaluation. It involves judging the appropriateness of an evaluation’s processes, the validity of its conclusions, or the usefulness and influence of the evaluation as a whole. For example, a summative metaevaluation may examine the extent to which evaluators delivered what was promised, used resources efficiently, and acted ethically. In addition to informing such accountability issues, a summative metaevaluation may serve a verification purpose. That is, a summative evaluation might confirm or refute an evaluation’s conclusions (and recommendations, if present). Summative metaevaluation that checks the veracity of findings and appropriateness of methods is especially important when the stakes are high and important decisions hinge on the evaluation’s results (Patton, 2008; Stufflebeam and Coryn, 2014), as was the case with the summative metaevaluation of the Paris Declaration on Aid Effectiveness (OECD, 2005). That metaevaluation was conducted concurrently with the evaluation of the Paris Declaration, in which the primary evaluation spanned 4 years and involved 50 studies across 22 partner countries with reviews by 18 donor agencies (Patton and Gornick, 2011).

When a summative metaevaluation validates a primary evaluation, it adds credibility to the evaluation. It may enhance the primary evaluation stakeholders’ confidence in using the evaluation findings to inform decisions to expand, modify, or cancel

programs. When a summative metaevaluation finds severe flaws in a primary evaluation, it can help prevent decision-makers from taking actions based on faulty information. Evaluation consumers, clients, and practitioners can use summative metaevaluations in a formative manner by applying what they learned in the summative metaevaluation to improve future evaluations.

Ascriptive metaevaluation

Metaevaluation is ascriptive when it is done for the sake of generating knowledge about evaluation (Scriven, 2011). This type of metaevaluation is not done to provide instrumental decision support as are formative and summative metaevaluations. Instead, ascriptive metaevaluation is done to gain insights about the state of evaluation, the impact or effectiveness of certain evaluation practices, and other issues whose study contributes to advancing evaluation as a professional practice and form of inquiry. Ascriptive metaevaluations are often conducted by graduate students for their theses as a means of conducting research on evaluation (Henry and Mark, 2003; Mark, 2008). Here, Scriven's conceptualization of the theoretical and practical functions of metaevaluations converge. That is, ascriptive metaevaluation serves the practical purpose of determining the quality of a specific evaluative performance, and it does so in the service of advancing knowledge about evaluation itself as a professional practice or form of inquiry—i.e., it is also an example of the theoretical function of metaevaluation.

Distinguishing metaevaluation from synthesis studies

Metaevaluation is often confused with meta-analysis, which is a process for synthesizing quantitative results from multiple studies to determine an overall effect for the type of intervention being studied. Qualitative analogs to meta-analysis include systematic reviews and qualitative reviews. Whether qualitative or quantitative, such synthesis studies involve finding, screening, combining, and extracting data from primary research and evaluation studies. The screening process involves assessing the primary studies to determine whether they meet established thresholds for quality or rigor for inclusion in the synthesis. Thus, there is a meta-evaluative step in syntheses of evaluation studies. However, meta-analysis and other types of synthesis studies are not the same as metaevaluation (although some noted organizations and scholarly publications conflate the concepts, adding to the confusion).

We join other evaluation scholars who distinguish metaevaluation from synthesis studies such as meta-analyses and systematic reviews (Scriven, 2009; Schwandt, 2015; Stufflebeam and Zhang, 2017; Mertens and Wilson, 2019). In short, if a study's purpose is to judge the quality of evaluations, it is a metaevaluation. If its purpose is to reach conclusions about the programs or other initiatives that were evaluated, it is not a metaevaluation.

Metaevaluation questions and criteria

Metaevaluation questions define the scope of the inquiry. Criteria provide the basis for reaching judgments in relation to those questions. Both questions and criteria stem from evaluators' and associated stakeholders' values and assumptions about how the object of an evaluation should perform and what it should achieve. Therefore, questions and criteria are inextricably linked and must be considered in tandem.

Questions

As in a primary evaluation, the focal questions in a metaevaluation are determined based on the study's purpose and intended uses. However, various evaluation scholars have recommended sets of general metaevaluative questions. These general questions can be used to determine more specific questions for a given metaevaluation.

Regardless of the specific questions, metaevaluations by definition focus on questions about an evaluation's quality, value, or importance (Evaluation scholars tend to use the terms *merit*, *worth*, and *significance* as synonyms for those more commonplace terms).

Stufflebeam and Coryn (2014, pp. 653–654) said the fundamental question about an evaluation's merit is "how well it meets the requirements of a sound evaluation." They said the critical issue regarding an evaluation's worth is "its sufficiency in meeting the audience's evaluative information needs."

Patton (2008, p. 570) suggested these basic metaevaluation questions that focus on evaluation processes: "To what extent was the [evaluation] implemented as designed? What adaptations occurred during implementation and what are the implications of those adaptations for intended and unintended outcomes?"

The Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2011, p. 241) identified the following questions among those that a metaevaluation might address:

- Were the findings for the program evaluation clear?
- Were the design and implementation of the design appropriate for the study?
- Were the conclusions justified by the data?
- Were the results adequately disseminated? Were they used?

Stakeholders involved in a metaevaluation can use these general questions as prompts to generate more specific questions that reflect their values, concerns, and information needs.

Criteria

Commenting on the salience of criteria in any evaluation context, [Patton \(2020, p. 581\)](#) explained,

It's all about criteria. Criteria are the basis for evaluative judgment. Determining that something is good or bad, successful or unsuccessful, works or doesn't work, and so forth and so on requires criteria.

Since the point of a metaevaluation is to reach conclusions about evaluation quality, it is essential to identify what quality means in a given evaluation situation. The selected criteria should make sense within the geographic and cultural context in which the primary evaluation was conducted, as well as with regard to its philosophical or theoretical underpinnings.

Various voluntary organizations for professional evaluation (VOPEs) and multilateral organizations have developed criteria that define evaluation quality for their contexts. [Table 1](#) lists 11 prominent evaluation quality frameworks. (Websites for these frameworks are listed under Relevant Websites at the end of this article.)

These frameworks define criteria for high-quality evaluation in their respective contexts. Evaluators can use them to guide their practice. Familiarity with them helps evaluation commissioners and other stakeholders to be more informed evaluation consumers. In the context of metaevaluation, these frameworks provide the foundational criteria for judging evaluation quality.

Many organizations that have authored evaluation quality frameworks have also put forth companion documents that define professional competencies for evaluators working in their contexts, as well as separate guidelines for ethical practice. Strictly speaking, these are not criteria for judging the quality of evaluations. Instead, they define how professional evaluators should conduct themselves. While the competence of evaluators figures into evaluation quality, a competent evaluator can certainly conduct a poor evaluation.

Grounding metaevaluation in codified, widely accepted criteria for evaluation quality promotes consistency and professionalism in evaluation practice. Evaluators do not need special certification or licensure processes to practice evaluation or call themselves evaluators. As rare exceptions, Japan and Canada have voluntary credentialing systems ([Kuji-Shikatani and Imata, 2018](#); [Canadian Evaluation Society, 2014](#)). In its lack of regulation, the evaluation profession contrasts sharply with the fields of accounting, medicine, law, and psychology, and the construction trades, in which regulatory practices are well-established. In the absence of regulation, it is incumbent on evaluation practitioners, clients, managers, and consumers to educate themselves about what constitutes evaluation quality. The use of established evaluation quality frameworks to both guide and assess evaluation practice helps advance a shared expectation for professional evaluation practice.

Next, we review a few examples of evaluation quality frameworks to highlight criteria that have been developed for use in different regions and sectors.

Program evaluation standards—joint committee on standards for educational evaluation

The Program Evaluation Standards were developed by the Joint Committee on Standards for Educational Evaluation ([Yarbrough et al., 2011](#)). The Joint Committee formed in 1975 as a coalition of several US and Canadian organizations with that each had a stake

Table 1 Region- and organization-specific evaluation quality frameworks.

<i>Authoring organization</i>	<i>Evaluation quality framework</i>
African Evaluation Association	African Evaluation Guidelines
Aotearoa New Zealand Evaluation Association	Evaluation Standards
Department of Performance Monitoring and Evaluation, Presidency of the Republic of South Africa	Standards for Evaluation in Government
German Evaluation Society	Standards for Evaluation
Joint Committee on Standards for Educational Evaluation	Program Evaluation Standards
Latin American and Caribbean Network of Monitoring, Evaluation and Systematization	Evaluation standards for Latin american and the Caribbean
Organization for Economic Co-operation and Development, Development Assistance Committee	Quality Standards for Development Evaluation
Swiss Evaluation Society	Evaluation Standards
Uganda Evaluation Association	Uganda Evaluation Standards
United Nations Evaluation Group	Norms and Standards for Evaluation
United States Government Accountability Office	Government Auditing Standards (Specifically for Performance Audits)

in educational evaluation. The Committee published its first set of standards for program evaluation in 1981 and subsequent editions in 1994 and 2011.

Several distinguished evaluation scholars and practitioners have identified the Program Evaluation Standards as one of the most important sources of criteria for assessing evaluation quality in the US context (Fitzpatrick et al., 2022; Patton, 2008; Schwandt, 2015; Scriven, 2013; Stufflebeam and Coryn, 2014). They include 30 distinct standards that define evaluation quality in five domains: Utility, Feasibility, Propriety, Accuracy, and Accountability.

The *Utility* standards explain that evaluation should be useful and valuable to stakeholders. The *Feasibility* standards call for evaluations to be conducted efficiently and be responsive to their cultural and political contexts. The *Propriety* standards address issues such as consent, confidentiality, ethics, and social justice. The *Accuracy* standards call for evaluations to be methodologically sound, transparent, and culturally appropriate. Finally, the *Evaluation Accountability* standards explain that evaluators should fully document and evaluate (i.e., metaevaluate) their work.

The individual Program Evaluation Standards are broad statements; for example, the accuracy standard on communication and reporting states, “Evaluation communications should have adequate scope and guard against misconceptions, biases, distortions, and errors.” Each standard is elaborated with a rationale, guidance for meeting the standards, a description of hazards evaluators may encounter related to the standard, and a scenario to help readers understand how the standard manifests in practice.

The Joint Committee’s Program Evaluation Standards, developed for use in the US and Canada, were the basis for evaluation standards defined by the following VOPEs:

- African Evaluation Association
- German Evaluation Society
- Latin American and Caribbean Network for Evaluation
- South African Department of Performance Monitoring and Evaluation
- Ugandan Evaluation Association

African evaluation guidelines—African Evaluation Association (AfrEA)

AfrEA’s framework is made up of 22 principles that define evaluation quality in the African context (AfrEA, 2007, 2020; Nairobi M&E Network et al., 2002). The 2002 and 2007 versions of AfrEA’s standards were closely aligned with the Joint Committee’s Program Evaluation Standards. In contrast, the latest iteration of these principles, published in draft form in 2020, reflects AfrEA’s Made in Africa initiative to develop and promote a “uniquely African approach to evaluation.”

The *Powerful for Africans* principles emphasize empowerment, learning, and capacity development. The *Technically Robust* principles address rigor, transparency, efficiency, cultural responsiveness, and other factors that affect an evaluation’s technical quality and credibility. The *Ethically Sound* principles call for protecting and respecting the people involved in an evaluation and attending to equity, diversity, and inclusion. Principles under the heading *Africa Centric, Yet Open* are about prioritizing African values and methods while engaging with indigenous communities and, more broadly, the global south. The *Connected With The World* principles are about acknowledging and fostering interconnectedness across communities and with nature, calling for evaluations to contribute to transformative change.

Evaluation standards—Aotearoa New Zealand Evaluation Association (ANZEA)

ANZEA formed in 2006 to represent the needs and interests of evaluators working in New Zealand (Aotearoa is the Māori name for New Zealand). The ANZEA standards are grounded in the Treaty of Waitangi, New Zealand’s founding document, and reflect the values of New Zealand’s indigenous populations (ANZEA, 2015). There are 20 ANZEA evaluation standards that define four overarching evaluation principles.

The standards for *Respectful, Meaningful Relationships* call for nurturing positive relationships among everyone involved in an evaluation in mutually beneficial and culturally appropriate ways. The *Ethic of Care* standards convey the importance of kindness, respect, humanity, reciprocity, cultural sensitivity, safety, inclusion, privacy, and confidentiality in evaluation. The standards for *Responsive Methodologies and Trustworthy Results* call for evaluators to use methods that are culturally and technically appropriate and ensure findings and conclusions are meaningful and credible. The *Competence and Usefulness* standards explain that both evaluators and evaluation commissioners must be professionally, contextually, and culturally competent and that evaluations should generate useful information.

Quality standards for development evaluation—Organization for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC)

OECD includes 31 member countries that make development investments worldwide. The OECD’s *Quality Standards for Development Evaluation* (2010) include 27 standards in four categories. The standards apply to evaluations of international aid initiatives and development assistance, including humanitarianism, sustainable development, conflict response, and peacebuilding. They are intended for evaluation commissioners and practitioners within the OECD network, including those from bilateral donors and multilateral development agencies and especially those who use the OECD-DAC criteria in primary evaluations.

The OECD’s evaluation quality standards under the heading of *Overarching Considerations* include high-level statements about the nature of development evaluation, ethics, capacity development, and other general topics. The *Purpose, Planning, and Design* standards address a diverse set of topics related to defining an evaluation’s purpose and focus, selecting criteria and methods, and

management considerations. The standards related to *Implementation and Reporting* cover a range of issues related to people, resources, contexts, execution of the evaluation's design, and the evaluation's products. Finally, the *Follow-Up, Use, and Learning* standards underscore the importance of taking action after an evaluation to enhance its use and influence.

Norms and standards for evaluation in the United Nations systems—United Nations Evaluation Group (UNEG)

The *Norms and Standards for Evaluations in the United Nations Systems* (UNEG, 2016) include 14 norms and 24 standards. This quality framework is intended to strengthen and harmonize evaluation practice within the U.N. system, especially with regard to developing national-level capacities for evaluating progress toward the Sustainable Development Goals (United Nations, 2015). Despite the focus on the U.N. system, these standards are intended to be used beyond the U.N. and benefit the global evaluation community. The norms represent overarching principles and values. The standards are more specific statements about how organizations, evaluation managers and commissioners, and evaluators should uphold the norms in practice.

The first ten norms are described as “general norms for evaluation.” These include *Internationally Agreed Principles, Goals and Targets; Utility; Credibility; Independence; Impartiality; Ethics; Transparency; Human Rights and Gender Equality; National Evaluation Capacities; and Professionalism*. The last four norms are “institutional norms for evaluation in the United Nations System.” These include *Enabling Environment; Evaluation Policy; Responsibility for the Evaluation Function; and Evaluation Use and Follow-Up*.

The 24 standards are organized into five categories: *Institutional Framework, Management of the Evaluation Function, Evaluation Competencies, Conduct of Evaluations, and Quality*. The latter two are the most relevant for judging the quality of specific evaluations. The eleven standards under *Conduct of Evaluations* explain that evaluations should:

- be timely and reflect specific intentionality
- be informed by evaluability assessments
- be guided by clear terms of reference
- have clearly defined and feasible scope and objectives
- employ rigorous methods
- engage stakeholders in evaluation planning and implementation
- address human rights and gender equity by both (1) assessing these aspects of programs being evaluated and (2) ensuring the evaluation itself is supportive of human rights and gender equity
- be conducted by competent and diverse evaluation teams selected through transparent processes
- generate reports and other products that are responsive to users' needs and include evidence-based findings, conclusions, and recommendations
- include realistic, evidence-based recommendations
- be communicated and disseminated effectively to enhance use

The final three UNEG standards are under the heading of *Quality*. These standards call for quality assurance systems for evaluation, especially at the design and final stages of evaluation.

Idiosyncratic criteria and connoisseurship in metaevaluation

Metaevaluators and/or stakeholders may opt, based on a metaevaluation's overarching purpose or what they want to learn from it, to use criteria different from or additional to those represented in widely accepted evaluation quality frameworks. Such criteria may be developed specifically for the metaevaluation or drawn from other sources. For example, suppose a metaevaluation involved assessing the degree to which an evaluation was carried out with cultural competence. In that case, a metaevaluator could use the resource that the Centers for Disease Control and Prevention developed that links the Program Evaluation Standards to specific strategies to increase and demonstrate cultural competence (CDC, 2014). In some cases, an issue of interest may be the degree to which an evaluation adheres to a specific theoretical approach to evaluation, such as utilization-focused evaluation (Patton, 2008), empowerment evaluation (Fetterman and Wandersman, 2018), or program-theory-based evaluation (Chen, 2015; Donaldson, 2007). In these cases, the tenets of those approaches could be treated as metaevaluation criteria (Scriven, 2009). However, it should be noted that these idiosyncratic criteria are sometimes contested among evaluators and stakeholders, depending on their professional values, methodological persuasions, and philosophical leanings.

Regardless of the criteria one uses in metaevaluation, it is almost unavoidable that there be a connoisseurial (Eisner, 1983) dimension to it. That is, to some extent, metaevaluators must bring to bear their own expert knowledge about evaluation, in combination with what they learn about a particular evaluation through additional means. Datta (1999, p. 346) called this the “I know it when I see it” approach to metaevaluation. Extensive data collection and explicit guidelines for interpreting metaevaluative data are not always practical for metaevaluations, especially ones that examine a single evaluation.

Metaevaluation methods

The methods used in a metaevaluation—including the types of data selected and how they are collected, analyzed, and interpreted—depend on the metaevaluation's purpose and scope. For example, a formative evaluation intended to improve the primary evaluation must involve examining evaluation plans and materials before they are finalized. In contrast, a summative evaluation would necessarily examine final evaluation products. A metaevaluation of multiple evaluations, whether a few or hundreds—must

include systematic procedures for extracting and coalescing consistent information about each evaluation, including steps to calibrate multiple raters, if applicable.

Evaluation reports serve as an important data source in metaevaluation because they are typically the most prominent deliverables related to primary evaluations. However, evaluation reports are rarely adequate for judging the overall quality of evaluations. There is an inherent difference between a report designed to meet the needs to a primary evaluation's stakeholders and one designed to convey all the information needed for a thorough metaevaluation. Long, highly technical reports tend to be counterproductive to the aim of making evaluations readily useful to stakeholders. Yet, it's often the technical details that are needed in order the technical details in order to judge an evaluation's technical adequacy, among other important aspects of its quality. Further, not all relevant aspects of an evaluation can be conveyed well in a written report. For instance, appropriate treatment and engagement of stakeholders is a common criterion across many of the evaluation quality frameworks discussed above, but this aspect of evaluation is difficult to describe in a written report; it is more easily understood through interviewing stakeholders about their experience with the evaluation. Reports are rarely adequate for comprehensively judging the overall quality of evaluations.

Next, we review methods and procedures recommended by prominent evaluation scholars. The relevance of some of these methods depends on whether the metaevaluation is examining one or several evaluations, the stakes involved and decisions in the balance, and the pertinent metaevaluation questions and criteria.

Metaevaluation methods recommended by stufflebeam

Stufflebeam (Stufflebeam, 2001; Stufflebeam and Coryn, 2014) recommended a metaevaluation approach that involves eleven tasks which parallel the tasks in primary evaluation. Stufflebeam and Coryn suggested these generic tasks (listed in Table 2) be used as a heuristic when planning metaevaluations. They pointed out that some metaevaluations may omit some tasks or include other tasks not on this list.

Metaevaluations, whether formative or summative, should almost always examine evaluation documents, such as plans, meeting minutes, instruments, protocols, and reports—i.e., the “pertinent available information” mentioned in Task 6, as defined by Stufflebeam and Coryn. However, documentation cannot tell the full story about an evaluation. That is why Task 7, which calls for the acquisition of additional information, is often warranted. Additional information may be collected through various means, such as interviews, focus groups, surveys, and observations.

Metaevaluation methods recommended by Cook and Gruder

Cook and Gruder's (1978) typology of metaevaluation approaches offers a different way to conceptualize metaevaluation methods. They argued that primary considerations are whether the metaevaluation occurs after a primary evaluation or concurrently, whether data collected for the primary evaluation are manipulated, and whether there are one or multiple data sets. They organized the possible options into seven basic approaches for metaevaluation:

1. Essay review of an evaluation report;
2. Review of the literature about a specific program;
3. Empirical reevaluation of an evaluation or program;
4. Empirical reevaluation of multiple data sets about the same program;
5. Consultant metaevaluation;
6. Simultaneous secondary analysis of raw data; and
7. Multiple independent replications (p. 17).

Table 2 Generic metaevaluation tasks.

1. Staffing	Staff the metaevaluation team with one or more qualified evaluators.
2. Stakeholder engagement	Identify and arrange to interact with the metaevaluation's stakeholders.
3. Standards	Agree on standards, principles, or criteria to judge the evaluation system or particular evaluation.
4. Questions	Define the metaevaluation questions.
5. Formal agreements	Issue a memo of understanding or negotiate a formal metaevaluation contract.
6. Existing information	Collect and determine the adequacy of pertinent available information.
7. New information	Collect new information as needed.
8. Analysis and synthesis	Analyze and synthesize the obtained information.
9. Reaching conclusions	Judge the evaluation's adherence to appropriate standards, principles, and/or criteria.
10. Reporting	Convey the metaevaluation findings through reports, correspondence, oral presentations, workshops, and other means.
11. Follow-up	As appropriate, help the client and other stakeholders interpret and apply the findings.

From Stufflebeam and Coryn (2014, p. 644).

This typology influenced early conceptualizations of metaevaluation. However, it does not reflect important developments in evaluation over the past 40+ years. For example, Cook and Gruder developed their typology before the first edition of the Program Evaluation Standards was published in 1981, so it lacks references to these or other standards for judging evaluations. Nonetheless, the range of options suggested by Cook and Gruder is broader than what is offered by more contemporary writings on metaevaluation, and persons interested or engaged in metaevaluation would be well served by revisiting this seminal work.

Metaevaluation methods recommended by Scriven

Scriven (2013) suggested three options that substantially overlap Cook and Gruder's (1978) typology:

1. Replicate the evaluation using the same methods and compare results.
2. Evaluate the same program using different methods and compare results.
3. Apply the Key Evaluation Checklist (Scriven, 2013), the Program Evaluation Standards (Yarbrough et al., 2011) or the United States Government Accountability Office (2011) standards to the evaluation.

The first two of these recommended methods and their correlates in Cook and Gruder's typology are applicable to summative metaevaluations for assessing a primary evaluation's accuracy. Scriven (1975) recommended a procedure for reaching convergence when independent evaluations of the same program are conducted. He said the involved evaluators should critique each other's reports, then modify their own based on the feedback. This process is critical for achieving metaevaluative objectives. Otherwise the client just has two evaluations of the same program, which may or may not agree. However, from a feasibility standpoint, it seems unlikely that many evaluation clients would be willing to pay for two full evaluations of the same program at the same time.

Metaevaluation practice

What does real-world metaevaluation practice look like? There have been few systematic investigations into the extent and nature of metaevaluation practice, but even these limited offerings provide useful insights.

Jacob and Affodegon (2015) described 28 metaevaluations published between 1980 and 2012. Across this period, they identified four generations of metaevaluation, which they defined as follows:

1. The *Development Period* (early 1970s to early '80s) was marked by attempts to define and clarify the concept of metaevaluation.
2. The *Propagation Period* (early 1980s to early '90s) was characterized by efforts to understand how metaevaluation should be practiced.
3. The *Enrollment Period* (mid-1990s to mid-2000s) was a time of generating quality frameworks, metaevaluation instruments, and actual metaevaluations
4. The *Maturity Period* (mid-2000s to 2012) was characterized by further codification, adoptions and adaptations of the Joint Committee's Program Evaluation Standards across regional and national VOPEs, and two peaks of activity in terms of the number of published metaevaluation reports annually.

Cooksy and Caracelli (2009) reviewed 18 metaevaluations of single evaluations that they located via evaluation listservs, journals, school websites, and professional contacts. They used these examples to describe and reflect on issues in selecting and applying metaevaluative criteria in metaevaluation practice. Among these 18 studies,

- 17 were summative
- most used multiple data sources, with evaluation reports being the most common source of data about the primary evaluation
- more than two-thirds included reviews of primary evaluation data
- two were conducted internally
- half used narrative reviews and five used checklists
- most used emergent criteria instead of predefined criteria; only four used the Joint Committee's Program Evaluation Standards
- all but one came from three university-based evaluation centers and graduate programs.

From their review, Cooksy and Caracelli concluded it was not clear that metaevaluation had become standard practice.

Motivated by intuition that metaevaluation practice was more common than suggested by these reviews of metaevaluation, this article's third author commenced an ambitious study to examine the extent and nature of metaevaluation practice over half a century (1969 to 2019). (Harnar et al., 2022). A robust search of scholarly and gray literature resulted in 2660 unique materials with the word *metaevaluation* in their title, abstract, or executive summary. Of these, 194 were true positive empirical examples of metaevaluation, indicating that metaevaluation is far more commonplace than suggested in previous reviews.

A review of a subset (nonrandom) of 27 of these metaevaluations published between 2000 and 2019 provided a preliminary view of metaevaluation practice in the education sector. The majority of published educational metaevaluations were external, summative metaevaluations of program evaluations. The metaevaluations assessed multiple evaluations and multiple aspects of those evaluations, including design, process, and results. Most used the Joint Committee's Program Evaluation Standards, either indirectly or explicitly, as their criteria for metaevaluative reasoning and made individual value-claims for each criterion. Data collection methods across these metaevaluations were diverse. Document review was the most common method, followed by

interviews, checklist application, surveys, and observations. These metaevaluations were typically conducted after the primary evaluation concluded, but some took place concurrently. Most made recommendations about the primary evaluations, and some included recommendations about conducting metaevaluation.

Next, we describe four examples of metaevaluations to illustrate the range of ways metaevaluations are carried out.

New York integrated learning system metaevaluation

Finn et al.'s (1997) metaevaluation was a hybrid formative and summative metaevaluation that took place both during and after the evaluation of the New York City Public Schools Integrated Learning System Project (Miller, 1997). It was conducted by an external metaevaluation team and written for diverse audiences, including schools, district offices, the state board of education, parents, community leaders, and the broader evaluation and educational research community. The metaevaluation was intended to "help these audiences assess the evaluation report and draw warranted conclusions about the merit and worth of the Integrated Learning Systems Project" (Finn et al., 1997, p. 159). Specifically, its purpose was to determine "the extent to which the overall study and each sub-study fulfilled the requirements of each of the 30 Joint Commission [sic] Standards" (p. 161). Accordingly, the metaevaluation used as metaevaluation criteria the second edition of the Program Evaluation Standards, which at the time included the domains of *Utility*, *Feasibility*, *Propriety*, and *Accuracy* (the *Accountability* domain was added in 2011). The metaevaluators used a robust assortment of procedures for data collection, analysis, and reporting:

- periodic reviews of evaluation plans, processes, and data
- site visits to observe the project's operation
- discussions with the primary evaluation team and other stakeholders about interim data, stakeholders' information needs, and findings from the formative metaevaluation
- systematic assessment of the final evaluation report and associated materials against the Program Evaluation Standards
- discussions by the metaevaluation team in order to reach consensus about findings and the contents of the metaevaluation report

The metaevaluation team concluded the primary evaluation met requirements for the *Feasibility*, *Propriety*, and *Accuracy* standards, but that the performance of the evaluation for the *Utility* criteria was "marginal" (p. 159).

Metaevaluation of a technology literacy program in underserved communities

This summative metaevaluation assessed an evaluation of an out-of-school-time technology literacy program (Rodríguez-Campos and Ashford, 2015). Specifically, the metaevaluation sought to "assess the evaluation's purpose and questions, in addition to its methods, strengths, and weaknesses" (p. 3). The metaevaluation addressed these five questions:

1. What is the nature of the report?
2. To what extent is the report technically sound?
3. To what extent are the reports [sic] useful?
4. To what extent did the report employ ethical procedures?
5. To what extent was the evaluation method practical? (p. 6)

The external metaevaluation team used Stufflebeam's (1999) Program Evaluation Metaevaluation Checklist to judge the quality of the primary evaluation based on their review of its final report. Stufflebeam's metaevaluation checklist is based on the 1994 version of Program Evaluation Standards by the Joint Committee on Standards for Educational Evaluation (Note that Questions 2 and 3 above parallel the Program Evaluation Standards domains of *Accuracy*, *Utility*, *Propriety*, and *Feasibility*). The checklist includes multiple checkpoints for each standard. Users are instructed to indicate whether the primary evaluation met each checkpoint and use the checklist's scoring guide to reach conclusions about the evaluation quality within each domain.

The metaevaluators adapted the checklist's scoring guidelines by using both a "conservative" and "moderate" approach. Their conservative scoring approach assumed that all information needed to assess the evaluation based on the checklist would be included in the evaluation report and that all checkpoints were relevant. The moderate approach allowed for a more liberal interpretation, acknowledging that some checkpoints may not be relevant or justifiably not be addressed in the evaluation report. The metaevaluators reported their conclusions based on both approaches. According to their conservative analysis, they rated the evaluation's *Utility* as "fair" and its *Feasibility*, *Propriety*, and *Accuracy* as "good." Using the moderate approach, they judged the evaluation's *Utility* and *Accuracy* as "very good" and its *Feasibility* and *Propriety* as "excellent." Ultimately, they synthesized these results and concluded that the evaluation was credible and high quality.

Responsive metaevaluation of the Appalachia regional comprehensive center's evaluation of a capacity-building program

Sturges and Howley's (2017) formative metaevaluation of the evaluation of a "five-year program to build a state department of education capacity to implement federal education law" (p. 126) took place concurrently with the primary evaluation. The metaevaluation merged standards-based metaevaluation, similar to the first two examples in this section, with a primary evaluation

approach of responsive evaluation (Stake, 2004). The metaevaluation's purpose was to "build evaluator capacity, enhance the evaluation's quality, optimize evaluation use, and minimize conflict" (Sturges and Howley, 2017, p. 126).

The metaevaluators described eight steps that closely resembled those recommended by Stufflebeam (Section "Metaevaluation methods recommended by Scriven" of this article; Stufflebeam, 2001; Stufflebeam and Coryn, 2014). However, their final two steps reflect the responsive and iterative nature of their metaevaluation approach.

Their Step 7 was to "facilitate dialogue, deliberation, and agreement" (p. 131). They explained that this involved two types of "facilitated events." They engaged the evaluation teams in periodic reflection sessions to "examine challenges and create plans to remedy them" (p. 131). They also held annual planning retreats to review the year's completed evaluation work, discuss its strengths and weaknesses, and set objectives for the next year.

Their Step 8 was to "review and refine the metaevaluation" (p. 131). In essence, this step was meta-metaevaluative in nature. The metaevaluators sought feedback about the strengths and weaknesses of their work (i.e., the metaevaluation) and made adjustments to be responsive to their stakeholders' needs.

The metaevaluation criteria were derived from the Program Evaluation Standards (Yarbrough et al., 2011), Stufflebeam's (1999) Metaevaluation Checklist, and the American Evaluation Association's (2018) Guiding Principles for Evaluators. After engaging stakeholders to review and prioritize criteria from these sources, the metaevaluators created a matrix that represented the "collectively agreed-upon set of criteria, expectations, and critical success factors for the evaluation" (p. 129). It also listed the sources of information for the metaevaluation, which included document reviews, interviews, surveys, and observations. Like Stufflebeam's metaevaluation checklist, it included a five-point rating scale for each criterion.

Due to the formative, iterative, and concurrent nature of the multi-year metaevaluation, there was no single final metaevaluative synthesis. Instead, the metaevaluators delivered annual reports with their quality ratings and recommendations to the primary evaluators and the project's leadership, along with periodic "targeted feedback briefs" (p. 134).

Metaevaluation of HIV/AIDS program evaluations using the African evaluation guidelines

This metaevaluation included 14 evaluations of HIV/AIDS-related programs funded by UNICEF and UNAIDS across nine African countries (Patel, 2002). Its purpose was two-fold: (1) identify opportunities to improve evaluations of HIV/AIDS-related programs and (2) pilot use of the then-new African Evaluation Guidelines (AEG) (Nairobi M&E Network et al., 2002) on a large scale. The metaevaluation applied 20 of the 30 guidelines that made up the 2002 AEGs. The other ten were not included because, Patel explained,

Some of the Guidelines are best used *ex ante* since some aspects of quality cannot easily be assessed after the fact. Some guidelines required information that was not available in the papers submitted for publication. For example, it was often difficult for the reviewers to assess the local 'Political viability' of an evaluation (p. 330).

Using the prompt, "How well did each evaluation meet the guidelines?" three reviewers independently rated each of the 14 evaluations on the 20 relevant guidelines using a six-point (0–5) rating scale. Compiling the evaluations' scores across the 20 guidelines in a table enabled the team to compare their relative strength and weakness.

Across all the guidelines, the median scores for the set of 14 evaluations were either 3 or 4, except for the guideline of *Values Identification* under the *Utility* domain, which scored a 2. Patel observed that the low scores in this area were "especially troubling" and a "serious weakness" given that "variations in cultural and social values play a significant role in the [HIV/AIDS] epidemic" (p. 330).

Summary

Metaevaluation serves several important functions for a wide range of evaluation stakeholders. Formative metaevaluation helps evaluators ensure their work is relevant and technically sound. Summative metaevaluation adds credibility to primary evaluations and helps evaluation audiences gauge how much confidence they should place in the findings. Ascriptive metaevaluation, as a vehicle for research on evaluation and theory-building, aids in enhancing the evaluation discipline's self-knowledge and developing more sophisticated models and tools for evaluation. Independently, each type of metaevaluation represents what Scriven (1969, p. 36) referred to as the practical role of metaevaluation: to evaluate "specific evaluative performances." Collectively, these metaevaluation efforts contribute to our knowledge about what constitutes sound evaluation practice, where evaluation discipline falls short of what it strives to accomplish, evaluation's role in contributing to the betterment of society, and how to advance the discipline. This is what Scriven described as the theoretical function of metaevaluation, which goes beyond looking at particular evaluations to "the evaluation of the very function and practice of evaluation itself" (Fitzpatrick et al., 2004, p. 443).

Relevant websites

<i>Evaluation quality framework</i>	<i>Website</i>
African evaluation association - african evaluation guidelines	afrea.org
Aotearoa New Zealand evaluation association - evaluation standards	anzea.org.nz/evaluation-standards
German evaluation society - standards for evaluation	degeval.org
Joint committee on standards for educational evaluation - program evaluation standards	evaluationstandards.org
Latin american and caribbean network of monitoring, evaluation and systematization - evaluation standards for Latin America and the Caribbean	relac.net/biblioteca/evaluation-standards
Organization for economic Co-operation and development - quality standards for development evaluation	oecd.org/dac/evaluation/publicationsanddocuments.htm
Republic of South Africa - standards for evaluation in government	dpme.gov.za
Swiss evaluation society - evaluation standards	seval.ch/en
Uganda evaluation association - Uganda evaluation standards	ugandaevaluationassociation.org
United nations evaluation group - norms and standards for evaluation	unevaluation.org/document/detail/1914
United States government accountability office - government auditing standards	www.gao.gov/products/gao-12-331g

References

- African Evaluation Association (AfREA), 2007. African Evaluation Guidelines: Norms and Standards. Available at: https://member.samea.org.za/img/AEG_2007_EN.pdf. (Accessed 24 April 2021).
- African Evaluation Association, 2020. African Evaluation Guidelines: 2020 Version. Available at: <https://afrea.org/>. (Accessed 24 April 2021).
- American Evaluation Association, 2018. AEA Guiding Principles (2018 Update). Available at: <https://www.eval.org/cm/ld/fid=51>. (Accessed 24 April 2021).
- Aotearoa New Zealand Evaluation Association (ANZEA), 2015. Evaluation Standards for Aotearoa New Zealand. Available at: <https://www.anzea.org.nz/about-us/resources/>. (Accessed 24 April 2021).
- Canadian Evaluation Society, 2014. About the C.E. Designation. Available at: <https://evaluationcanada.ca/ce>. (Accessed 24 April 2021).
- Centers for Disease Control and Prevention (CDC), 2014. Practical Strategies for Culturally Competent Evaluation. Available at: https://www.cdc.gov/asthma/program_eval/cultural_competence_guide.pdf. (Accessed 1 May 2021).
- Chen, H.T., 2015. Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective, second ed. Sage, Los Angeles.
- Cook, T.D., Gruder, C.L., 1978. Metaevaluation research. *Eval. Rev.* 2 (1), 5–51.
- Cooksy, L.J., Caracelli, V.J., 2009. Metaevaluation in practice. *J. Multidiscip. Eval.* 6 (11), 1–15.
- Coryn, C.L., Ozeki, S., Wilson, L.N., Greenman, G.D., Schroeder, D.C., Hobson, K.A., et al., 2016. Does research on evaluation matter? Findings from a survey of American Evaluation Association members and prominent evaluation theorists and scholars. *Am. J. Eval.* 37 (2), 159–173.
- Datta, L., 1999. CIRCE's demonstration of a close-to-ideal evaluation in a less-than-ideal world. *Am. J. Eval.* 20 (2), 345–354.
- Donaldson, S.I., 2007. Program Theory-Driven Evaluation Science: Strategies and Applications. Erlbaum, Mahwah, NJ.
- Eisner, E.W., 1983. Educational connoisseurship and criticism: their form and functions in educational evaluation. In: Madaus, G.F., Scriven, M., Stufflebeam, D.L. (Eds.), *Evaluation Models*. Kluwer-Nijhoff, Boston, pp. 87–102.
- Fetterman, D.M., Wandersman, A., 2018. Essentials of empowerment evaluation. In: Fetterman, D.M., Rodríguez-Campos, L., Zukoski, A.P. (Eds.), *Collaborative, Participatory, and Empowerment Evaluation: Stakeholder Involvement Approaches*. Guilford, New York, pp. 74–89.
- Finn Jr., C.E., Stevens, F.I., Stufflebeam, D.L., 1997. A meta-evaluation. *Int. J. Educ. Res.* 27 (2), 159–174.
- Fitzpatrick, J.L., Sanders, J.R., Worthen, B.R., 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines*, third ed. Pearson, Boston.
- Fitzpatrick, J.L., Sanders, J.R., Worthen, B.R., Wingate, L.A., 2022. *Program Evaluation: Alternative Approaches and Practical Guidelines*, fifth ed. Pearson, Boston.
- Harnar, M. A., Jersild, A. C., Marshall, V., Rasheed, S. I., Tsuzaki, T., and Tilton, Z. D., 2022. Half a century of transdisciplinary metaevaluation practice: a systematic review of metaevaluations from 1969–2019. Manuscript in preparation.
- Henry, G.T., Mark, M.M., 2003. Toward an agenda for research on evaluation. *N. Dir. Eval.* 97, 69–80.
- Jacob, S., Affodegon, W.S., 2015. Conducting quality evaluations: four generations of meta-evaluation. *SpazioFilosofico* 13, 165–175.
- Kuji-Shikatan, K., Imata, K., 2018. AEA's Evaluator Competencies: What Can We Learn From Other Evaluation Societies About Next Steps? The Japan Evaluation Society's Case. Available at: <https://comm.eval.org/viewdocument/aea-evaluation-2018-session1518aea>. (Accessed 24 April 2021).
- Landsberger, H.A., 1958. Hawthorne Revisited. Cornell University, Ithaca.
- Mark, M.M., 2008. Building a better evidence base for evaluation theory. In: Smith, N.L., Brandon, P.R. (Eds.), *Fundamental Issues in Evaluation*. Guilford, New York, pp. 111–134.
- Mertens, D.M., Wilson, A.T., 2019. *Program Evaluation Theory and Practice*, second ed. Guilford, New York, NY.
- Miller, H.L., 1997. An overview of the evaluation. *Int. J. Educ. Res.* 27 (2), 93–97.
- Nairobi M&E Network, African Evaluation Association Secretariat, Réseau Nigérien de Suivi et Evaluation, Cape Verde Evaluation Network, Réseau Malagache de Suivi et Evaluation, Comoros Evaluation Network, Eritrean Evaluation Network, Malawi M&E Network, Réseau National de Chercheurs et Evaluateurs de Burundi, Rwanda Evaluation Network and UNICEF Eastern and Southern Africa Region M&E Network, 2002. The African Evaluation Guidelines: 2002. A checklist to assist in planning evaluations, negotiating clear contracts, reviewing progress and ensuring adequate completion of an evaluation. *Eval. Progr. Plann.* 25 (4), 481–492.

- Orata, P., 1940. Evaluating evaluation. *J. Educ. Res.* 33 (9), 641–666.
- Organisation for Economic Development and Cooperation, 2005. Paris Declaration on Aid Effectiveness. Available at: https://www.oecd-ilibrary.org/development/paris-declaration-on-aid-effectiveness_9789264098084-en. (Accessed 30 April 2021).
- Patel, M., 2002. A meta-evaluation, or quality assessment, of the evaluations in this issue, based on the African Evaluation Guidelines: 2002. *Eval. Progr. Plann.* 25 (4), 329–332.
- Patton, M.Q., 2008. *Utilization-focused Evaluation*, fourth ed. Sage, Los Angeles.
- Patton, M.Q., Gornick, J., 2011. Evaluation of the Phase 2 Evaluation of the Paris Declaration: An Independent Review of Strengths, Weaknesses, and Lessons. *Utilization-Focused Evaluation*, Saint Paul. Available at: <https://www.oecd.org/dac/evaluation/dcdndep/48620425.pdf>. (Accessed 1 May 2021).
- Patton, M.Q., 2020. Evaluation use theory, practice, and future research: reflections on the Alkin and King AJE series. *Am. J. Eval.* 41 (4), 581–602.
- Rodríguez-Campos, L., Ashford, S., 2015. Metaevaluation of a technology literacy program for under-served communities. *Int. J. Interdiscip. Stud. Commun.* 10 (4), 1–8.
- Schwandt, T.A., 2015. *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford University Press, Stanford, CA.
- Scriven, M., 1969. An introduction to meta-evaluation. *Educ. Prod. Rep.* 2, 36–38.
- Scriven, M., 1975. Evaluation bias and its control. *J. Multidiscip. Eval.* 7 (15), 79–98.
- Scriven, M., 2013. Key evaluation checklist. Available at: http://www.michaelscriven.info/images/KEC_3.22.2013.pdf. (Accessed 24 April 2021).
- Scriven, M., 2009. Meta-evaluation revisited. *J. Multidiscip. Eval.* 6 (11), iii–viii.
- Scriven, M., 2011. Evaluating Evaluations: A Meta-evaluation Checklist. Available at: http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.16.11.pdf. (Accessed 24 April 2021).
- Smith, N.L., 2010. Characterizing the evaluand in evaluating theory. *Am. J. Eval.* 31 (3), 383–389.
- Smith, E.R., Tyler, R.W., 1942. *Appraising and Recording Student Progress*. Harper & Row.
- Shadish, W.R., 1998. Evaluation theory is who we are. *Am. J. Eval.* 19 (1), 1–19.
- Stake, R.E., 2004. *Standards-Based and Responsive Evaluation*. Sage, Thousand Oaks, CA.
- Stufflebeam, D.L., 1999. Program evaluations metaevaluation checklist. Available at: <https://wmich.edu/evaluation/legacychecklists>. (Accessed 1 May 2021).
- Stufflebeam, D.L., 2001. The metaevaluation imperative. *Am. J. Eval.* 22 (2), 183–209.
- Stufflebeam, D.L., Coryn, C.L.S., 2014. *Evaluation Theory, Models, and Applications*. Jossey-Bass, San Francisco.
- Stufflebeam, D.L., Zhang, G., 2017. *The CIPP Evaluation Model: How to Evaluate for Improvement and Accountability*. Guilford, New York, NY.
- Sturges, K.M., Howley, C., 2017. Responsive meta-evaluation: a participatory approach to enhancing evaluation quality. *Am. J. Eval.* 38 (1), 126–137.
- Thomas, R.M., 1984. Mapping meta-territory. *Educ. Res.* 13, 161–18.
- United Nations, 2015. The 17 Goals. Available at: <https://sdgs.un.org/goals>. (Accessed 30 April 2021).
- United Nations Evaluation Group (UNEG), 2016. Norms and Standards for Evaluations. Available at: <http://www.unevaluation.org/document/detail/1914>. (Accessed 24 April 2021).
- United States Government Accountability Office, 2011. *The Government Auditing Standards*. Available at: <https://www.gao.gov/products/gao-12-331g>. (Accessed 24 April 2021).
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K., Caruthers, F.A., 2011. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*, third ed. Sage, Los Angeles, CA.

Evaluating schools as professional learning communities[☆]

Sally M. Thomas, School of Education, University of Bristol, Bristol, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

This is an update of S.M. Thomas, Evaluating Schools as Learning Communities, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 539-547, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01638-9>.

Introduction	275
Three approaches to evaluate schools as professional learning communities	276
Characteristics and processes of schools as PLCs	277
Stages of PLC development	278
Summary indicators of PLCs	280
Contrasting the extent of PLCs in primary and secondary schools	281
Relationship between pupil outcomes and PLC processes and contexts	281
Calculating VA and raw school performance	282
Correlations between school performance and PLC indicators	282
Correlations between school performance and PLC factual characteristics	283
Conclusion	283
References	285
Further reading	286
Relevant website	286

Glossary

CPD Continuing Professional Development
DFE UK's Department for Education
DCSF UK's Department for Children, Schools and Families (replaced by DFE in 2010)
DFES UK's Department for Education and Science (replaced by DCSF in 2007)
GCSE General Certificate of Secondary Education
GNVQ General National Vocational Qualifications
GTC UK's General Teaching Council
KS2 National Curriculum Key Stage Two (England)
KS4 National Curriculum Key Stage Four (England)
NCSL UK's National College for School Leadership.
PLC Professional Learning Community
PLCs Professional Learning Communities
VA Value-Added measure

Introduction

Continuing emphasis on the fundamental—but not particularly new—idea that teachers working and learning together can enhance the quality of learning and teaching in schools has led to widespread international interest, focusing on the potential of professional learning and professional learning communities (PLCs) to advance school improvement. For over 15 years this approach has been a key thread in England's Department for Education (DfE) National Standards for Head teachers and teachers' professional development (DfE, 2020, 2016). Research studies in a range of international contexts, particularly in the USA, Netherlands and more recently China, have similarly emphasized the value and role of professional learning communities to enhance professional practice (Stoll and Louis, 2007; Hipp et al., 2008; Slegers et al., 2013; OECD, 2016; Hairon et al., 2017; Thomas et al., 2017; Zhang and Sun, 2018; Zahedi et al., 2021). Effective PLCs are seen as more likely to generate and support sustainable improvements in student and other outcomes because they build the necessary professional skill, capacity, and shared good practice. However, can the effectiveness and quality of a PLC be measured, and does it have a quantifiable positive impact on pupil

[☆] Change History: March 2022. S.M. Thomas updated the text and references throughout this entire article, updated the article title and added new Tables 2, 4 and 7.

outcomes? Although relatively few research studies address these specific questions, this article draws on the available evidence and reports in detail selected findings from two PLC studies conducted in UK and China. The first is a large-scale research project funded by the UK's Department for Education and Science (DFES), General Teaching Council (GTC), and National College for School Leadership (NCSL) 'Creating and Sustaining Effective Professional Learning Communities (EPLCs)' (Bolam et al., 2005). The second is a similar research project funded by UK Economic and Social Science Research Council (ESRC)¹ and UKAID 'Improving Teacher Development and Educational Quality in China (ITDEQC)' (Thomas et al., 2015). In particular, both the EPLC and ITDEQC projects illustrate the quantitative research tools used to define and measure specific aspects of PLCs and examines the relationship between these measures and pupils' academic outcomes. At this point, it is important to point out that the term professional learning community, used in this article, overlaps with—but is, nevertheless, more focused on—professional staff learning, than the term school learning community which is, generally, seen as broader and refers to all learners in the school community.

Three approaches to evaluate schools as professional learning communities

One common approach to evaluating schools as professional learning communities is to examine the school characteristics and processes that have been identified as key features of a PLC. International research suggests that PLCs are characterized by five aspects: shared values and vision; collective responsibility for pupils' learning; reflective professional inquiry; collaboration; and the promotion of group, as well as individual learning (Stoll et al., 2003, 2006a; Hord, 1997; Louis et al., 1996). However, further research identified three new characteristics and four processes relevant to the creation and effective functioning of PLCs (Bolam et al., 2005). Therefore, although PLCs do not look exactly the same in all schools, research has found that they seem to display eight characteristics which—irrespective of the educational phase or context—they exhibit to a greater or lesser degree:

1. Shared values and vision;
2. Collective responsibility for pupils' learning;
3. Collaboration focused on learning;
4. Individual, group, and collective professional learning;
5. Reflective professional enquiry;
6. Openness, networks, and partnerships;
7. Inclusive membership; and
8. Mutual trust, respect, and support.

In addition, four processes are required to create and sustain a PLC:

1. Optimizing resources and structures to promote the PLC;
2. Promoting individual and collective professional learning;
3. Evaluating and sustaining the PLC over time; and
4. Leading and managing to promote the PLC.

A range of professional development activities have been developed drawing on these 8 characteristics and 4 processes to assist school leaders and staff in reviewing, evaluating and sustaining their own school as a PLC (Stoll et al., 2006b). Other more recent research has explored the benefits of PLC interventions to support professional development in Dutch schools (Admiraal et al., 2021) as well as sought to further refine the PLC as a multidimensional construct (Hairon et al., 2017).

A second approach to evaluating schools as professional learning communities is to use survey and statistical techniques to identify and examine a finite set of key indicators or factors related to the processes of continuing professional development (CDP) and, more broadly, to the development of PLCs. This approach has been used in a study in the Netherlands by Visscher and Witziers (2004) who identified six PLC concepts (policy and evaluation; consultation and cooperation; consensus; decision-making, school leadership; and departmental leadership) and measured these via a teacher questionnaire administered to mathematics teachers in 39 schools. Pedder et al. (2005) have also explored the nature of professional learning via a factor analysis of teachers' responses to a questionnaire and identified four factors related to teacher learning. The factors were:

1. *Inquiry*. Using and responding to different sources of evidence; carrying out joint research and evaluation with colleagues.
2. *Building social capital*. Learning, working, supporting, and talking with each other.
3. *Critical and responsive learning*. Through reflection, self-evaluation, experimentation, and by responding to feedback.
4. *Valuing learning*.

One purpose of this kind of approach is to establish the existence of one or more factors that could be useful as summary indicators to measure and evaluate particular aspects of schools' capacity and success in developing as a PLC. More recent research has similarly sought to examine and refine PLC measures in different country contexts in Europe and East Asia (Slegers et al., 2013; Pang et al., 2016; Chen et al., 2016; Lomos, 2017; Zhang and Sun, 2018; Antinluoma et al., 2018). Importantly, indicators of this type could provide helpful feedback to inform school self-evaluation and improvement processes (see, e.g., Scheerens et al., 2003). This is

¹This research is supported by ESRC/UKAID funded projects Improving Educational Evaluation and Quality in China (RES-167-25-0353) and Improving Teacher Development and Educational Quality (RES-167-25-0428). See <http://ieeqc.bristol.ac.uk>

illustrated in greater detail below by presenting selected results from the EPLC questionnaire survey used to test out the relevance of different characteristics and processes of PLCs in different kinds of schools in England (Bolam et al., 2005) and equivalent findings from the ITDEQC study conducted in China (Thomas et al., 2015).

A third and, arguably, the most revealing approach to evaluating schools as professional learning communities is to examine the relationship between the indicators of PLCs and a range of pupil outcome data to establish whether or not the extent of school characteristics and processes that are associated with PLCs are also linked to improved student outcomes. There is some evidence with regard to the impact of continuing professional development—and, more generally, of PLCs—on student outcomes and progress (see, e.g., Lee and Smith, 1996; Louis and Marks, 1998; Wiley, 2001; Vescio et al., 2008; Lomos et al., 2011; Hofman et al., 2015; Doğan and Adams, 2018). However, findings in this area can be illustrated by both the EPLC study (Bolam et al., 2005), and the more recent comparable ITDEQC study in China (Thomas et al., 2015; Thomas, 2020) which provide rare opportunity—via the collection of school-questionnaire—survey data alongside access to the national pupil-assessment databases and school value added measures—to produce some relatively hard data. The nature of this data is, of course, correlational—not causal—but, nevertheless, this evidence is arguably more robust than many studies and is employed in this context to illustrate, in detail, the findings and methods used to evaluate the effectiveness of PLCs in terms of promoting student outcomes.

Characteristics and processes of schools as PLCs

As noted above, previous research in England has identified eight key characteristics and four processes relevant to the creation and sustainability of schools as PLCs. The study conducted by Bolam et al. (2005) employed both case-study and questionnaire-survey methods to draw its final conclusions. However, the detailed findings from the questionnaire administered to almost 400 English schools provide a useful example of the most and least common PLC features reported by schools.

The questionnaire was developed from a review of the literature on PLCs and was designed to be completed by the headteacher or CDP coordinator of each school participating in the study—in consultation with other staff. The questionnaire comprised three parts:

- *Part 1.* Items designed to gather opinions regarding aspects of PLC in the school; these questions are framed in terms of the proportion of staff in the school engaged in particular activities or having particular views. These items reflect different aspects of CDP and teachers' professional life: teacher support for pupil learning; professional learning; support for professional learning; collaboration and culture; research and professional enquiry; external links and networking; and participation of teaching assistants and nonteaching support staff in PLC.
- *Part 2.* Items exploring the perceptions of a definition of a PLC and the factors that the respondent felt facilitated, or were barriers, to the school becoming a PLC.
- *Part 3.* Items related to factual information with regard to the range and extent of professional development and school self-evaluation activities in the school.

The questionnaire findings indicated which PLC characteristics and processes the respondents were most and least likely to report, on average—in terms of the proportion of staff engaged in particular PLC activities, as well as which activities were the most and least variable across schools—thereby suggesting agreement among schools on the importance of some key PLC features (see Table 1).

The findings show that two question items are crucial in terms of primary and secondary schools reporting the highest mean level of teacher involvement. These are: collective responsibility for pupil learning and creating conditions for pupils to feel the confidence to learn. Responses to these two items also demonstrate a high level of consensus among schools, given the low item-score variability in comparison to other items. These findings suggest, not surprisingly, that pupil learning is the foremost concern of teachers, and that this PLC characteristic is very common in all schools.

In contrast, the results for two different items (experience job rotation and have opportunities for work shadowing) indicate that both primary and secondary schools, typically, reported a low level of teacher involvement in these activities. The finding suggests that these particular aspects of CDP related to learning and developing staff roles, were not a common characteristic in most schools. Finally, the results for two further items (have dedicated time for classroom observation and have some protected time for joint planning and development) suggest that, although a medium level of staff involvement is reported by primary and secondary schools, there is a notable lack of consensus (i.e., high item-score variability) concerning these two activities across schools. This finding suggests that some aspects of CDP—concerned with improving classroom practice—are very variable across English schools, with some schools reporting a high level of staff involvement and others reporting a low level of staff involvement. The issues of teachers' time management and professional culture may have an impact here.

Interestingly by comparing similar part 1 teacher survey results indicating the extent of staff engagement in PLC activities from England (EPLC project) and China (ITDEQC project), out of 34 equivalent questionnaire items, head teacher/CPD coordinators' responses were mostly found to rank similarly (see Table 2). However, three items (set learning targets for individual pupils; routinely share information with parents; and the community; and Non-Teaching support staff actively contribute to the school as a PLC) were ranked much higher by English than Chinese respondents (difference in ranking >16) in terms of the greater extent these activities were observed by head teachers/CPD coordinators in school. In contrast, four items (carry out classroom-based research; have dedicated time for classroom observation; engage in team teaching; and Non-Teaching Support Staff (NTSS) share

Table 1 Perceived aspects of PLC by primary schools and secondary schools: Extent of staff engagement for the three items (indicated with the original questionnaire numbers) with the highest/lowest means and standard deviations (England; EPLC project).

	<i>Mean</i>	<i>Standard deviation</i>	<i>n</i>
Highest mean—primary schools			
Q1: Collective responsibility for pupil learning	3.77	0.56	224
Q3: Create conditions for pupils to feel the confidence to learn	3.74	0.52	225
Q4: Learn together with colleagues	3.73	0.51	223
Highest mean—secondary schools			
Q1: Collective responsibility for pupil learning	3.64	0.54	164
Q3: Create conditions for pupils to feel the confidence to learn	3.40	0.61	164
Q24: are members of at least one professional team	3.80	0.55	160
Lowest mean—primary schools			
Q9: Carry out classroom-based research	1.81	1.00	209
Q36: Experience job rotation	1.83	1.07	214
Q38: Have opportunities for work shadowing	1.75	1.03	215
Lowest mean—secondary schools			
Q36: Experience job rotation	1.13	0.34	160
Q38: Have opportunities for work shadowing	1.36	0.71	157
Q44: Receive financial support from the school for award-bearing courses	1.42	0.85	153
Highest standard deviation—primary schools			
Q13: Have dedicated time for classroom observation	2.48	1.25	222
Q44: Receive financial support from the school for award-bearing courses	1.95	1.26	207
Q45: Have some protected time for joint planning and development	2.57	1.37	218
Highest standard deviation—secondary schools			
Q13: Have dedicated time for classroom observation	2.28	1.13	162
Q43: Use professional development profiles/portfolios	2.41	1.14	158
Q45: Have some protected time for joint planning and development	2.36	1.18	163
Lowest standard deviation—primary schools			
Q1: Collective responsibility for pupil learning	3.77	0.56	224
Q4: Learn together with colleagues	3.73	0.51	223
Q18: Share a common core of educational values	3.71	0.55	225
Lowest standard deviation—secondary schools			
Q1: Collective responsibility for pupil learning	3.64	0.54	164
Q24: are members of at least one professional team	3.80	0.55	160
Q36: Experience job rotation	1.13	0.34	160

Only includes items from the PLC survey part 1. The four-item response categories indicating the extent of staff engagement are treated as a numerical scale from one to four as follows: nearly all staff (>80%) = 4, most staff (50–80%) = 3, some staff (20–49%) = 2, a few staff (<20%) = 1. Adapted from Bolam et al. (2005).

responsibility for pupil learning) were ranked much higher by Chinese than English respondents (difference in ranking >16) in terms of the greater extent these activities were observed by head teachers/CPD coordinators in school. The findings suggest that PLCs in England and China may operate similarly overall, but some key elements such as PLC involvement of parents and non-teaching staff, and the opportunities for team teaching, classroom observation and research may differ (see also Qiao et al., 2018; Chen, 2020). Importantly this suggests that context specificity may be a critical issue when considering the nature and contribution of PLCs in supporting teacher development in different country contexts, and points to the need for new research exploring the usefulness of PLC concept in different education systems.

Stages of PLC development

Schools in the English EPLC sample were found to be adequately representative of the national picture in terms of context factors, such as the percentage of pupils in the school entitled to free school meals. However, it is important to note that schools were asked specifically, in part 2 of the survey, to assess their current stage of development as a PLC. The provisional working definition of a PLC quoted in the survey was as follows: “An effective professional learning community has the capacity to promote and sustain the learning of all professionals and other staff in the school community with the collective purpose of enhancing pupil learning.”

Table 2 Comparison of findings from England (EPLC project) and China (ITDEQC project): high (1–12) and low (23–34) ranked items based on 34 similar PLC survey questions regarding extent of teachers/ staff engagement with PLC activities.

	<i>Rank: China 2010</i>	<i>Rank: England 2003</i>
1. Shared values and vision		
Share a common core of educational values	1	6
2. Collective responsibility for pupils' learning		
Collective responsibility for pupil learning	2	2
Non-teaching support staff (NTSS) share responsibility for pupillearning	5	23
Pupil learning		
Set learning targets for individual pupils	27	5
Ensure pupils receive constructive feedback about their work	3	11
3. Collaboration focused on learning		
Engage in team teaching	8	31
Have dedicated time for classroom observation	9.5	26
Regularly discuss teaching methods	7	8
Have opportunities for work shadowing	25	33
Learn from each other	11.5	9
4. Individual, group and collective professional learning		
Receive training in how to work and learn in teams	28	25.5
5. Reflective professional practice		
Carry out classroom-based research	11.5	32
6. Openness, networks and partnerships		
Use university staff for professional learning	34	34
Use professional/subject associations for professional learning	30	28
7. Inclusive membership		
Routinely share information with parents and the community	31	7
Non-teaching support staff (NTSS) are valued by teachers	6	1
NTSS actively contribute to the school as a professional learning community	29	12

Note: Different but comparable item scales were used: China (never = 1; Always = 5); England (1 = <20% staff; 4 = >80% staff). For both EPLC and ITDEQC project the survey was designed to be completed by one respondent—CPD coordinator/head teacher. EPLC sample: secondary schools $n = 164$; ITDEQC sample: 98 senior secondary schools from 3 west/east LEAs.

Table 3 Percentage of English EPLC schools reporting different stages of PLC development.

<i>PLC stage</i>	<i>Nursery and primary schools (n = 227) %</i>	<i>Secondary schools (n = 166) %</i>
Mature/established	25	16
Developer	57	67
Starter	14	15
Working to reestablish PLC	2	1
No response	2	1

Note: The primary and the secondary school samples included special schools, respectively. Adapted from Bolam et al. (2005)

Most respondents—57% from nursery and primary schools and 67% from secondary schools—reported that their school was a developing PLC (see Table 3). Therefore, the EPLC survey findings reported above reflect the features of this self-reported group of schools more than any other. Moreover, it is important to note that the survey findings are a summary of data collected at one point in time from one respondent in each school—either the head or CDP coordinator—responding to the survey. Of course, the situation may have changed since the data were collected and, in addition, other members of the school may have different opinions concerning the questions. Clearly, these issues should be considered when interpreting the survey findings, although interviews in case-study schools—also conducted as part of the study—generally confirmed survey respondents' perceptions of their school's current stage of PLC development.

To contrast with the EPLC findings from English schools, where almost all questionnaire participants recognized their school at least to some extent as a PLC, the findings from China indicate that not all teachers agree that a 'Whole' school PLC exists in this school (see Table 4). Regardless of the three locations (LEAs in two central and one west region of China), 80% of teachers or more

Table 4 ITDEQC Teacher questionnaire 2012: Existence of PLCs in Chinese senior high schools.

Percentage response to item: "Whole school PLC exists in this school"	Strongly disagree	Disagree	Uncertain	Agree	Strongly agree
LEA1	1.3	3.2	16.5	61.1	18.0
LEA2	4.6	2.2	12.2	50.6	30.4
LEA3	0.9	4.4	22.5	54.7	17.5

Note: (i) calculated using individual teacher level data; (ii) sample size = LEA1 (8093), LEA2 (7608), LEA3 (1814) with item missing values less than 2% for total sample (17,515).

either agree or strongly agree that PLC is a meaningful concept in Chinese schools. However, the results are up to ten percent lower in terms of whether a PLC actually exists in their own school. Maybe not surprisingly LEA3 in the poorer west of China reported the lowest percentages on these issues, suggesting the greatest uncertainty regarding the PLC concept and the strength of traditional CPD approaches. This again highlights the need for new research exploring the usefulness of PLC concept in different education systems, and ways in which it might need to be adapted for different country contexts such as China.

Summary indicators of PLCs

In addition to the questionnaire-item results outlined above, the EPLC study of English schools also provides a useful example of findings that employ quantitative factor-analysis techniques to establish the existence of one or more summary indicators or factors that could be used to measure and evaluate schools as professional learning communities. Four independent factors were identified from the EPLC questionnaire results which relate to:

- Factor 1: professional and pupil learning ethos;
- Factor 2: within-school policy, management, and support for professional learning;
- Factor 3: enquiry orientation (external and internal); and
- Factor 4: participation of nonteaching support staff in PLC.

As an illustration of the results, the questionnaire items loading on Factor 1 (professional and pupil learning ethos) are shown in [Table 5](#) (for further details, see [Bolam et al., 2005](#)).

Correlations between the EPLC factor scores showed that, although there is some overlap between factors (i.e., all correlations are positive), there are, nevertheless, observable differences between these PLC summary indicators. Therefore, the evidence indicates that at least four dimensions reflecting different aspects of PLCs exist and individual schools can have a very different profile of results across the four factors. The factor findings related to professional and pupil learning, and enquiry orientation are also

Table 5 EPLC questionnaire items included on the factor of "professional and pupil learning ethos".

<i>Factor 1 professional and pupil learning ethos</i>
1 Collective responsibility for pupil learning
2 Base their approach to change on the use of good evidence
3 Create conditions for pupils to feel the confidence to learn
4 Learn together with colleagues
5 Ensure pupils receive constructive feedback about their work
6 Set learning targets for individual pupils
7 Routinely collect, analyze, and use data and evidence to inform their practice
8 Have low expectations of children
9 Actively seek and use feedback from pupils
10 Regularly monitor the learning and progress of individual pupils
11 Share a common core of educational values
12 Use the staff room at break times for professional links
13 Are involved in seeking solutions to problems facing the school
14 Regularly discuss teaching methods
15 Share their professional experiences and successes
16 Experiment and innovate in their work
17 See the school as stimulating and professionally challenging
18 Routinely share information with parents and the community
19 Learn from each other
20 Take responsibility for their own professional learning
21 Give priority to learning more about pupils learning

reflected in a comparable questionnaire study which focused specifically on professional learning and was conducted on the basis of a teacher survey—over 1000 individual teacher responses from 32 schools (Pedder et al., 2005). As previously noted, in the study carried out by Pedder and his colleagues, four factors were identified using standard factor-analysis techniques which comprised aspects of inquiry and using different sources of evidence, building social capital, critical and responsive learning, and valuing learning. Broad similarities could also be found with the PLC concept factors identified by Visscher and Witziers (2004).

Contrasting the extent of PLCs in primary and secondary schools

The validity of the PLC concept at the primary and secondary phases of education was also examined by Bolam and colleagues in terms of whether English schools from different educational phases report similar levels of staff involvement as measured by the PLC factors. Table 6 depicts the mean and standard deviation of the standardised factor scores from primary and secondary schools. Overall, the findings indicate that primary schools, generally, report higher levels of staff involvement in PLC processes than secondary schools. Interestingly, the highest level of primary-phase staff involvement suggested by the factor scores is in terms of pupil and professional learning ethos (Factor 1). In contrast, for the secondary phase, the highest level of staff involvement is in terms of enquiry orientation (external and internal; Factor 3).

As discussed above, in the EPLC survey, respondents were also asked to assess their current stage of development as a PLC in three developmental stages: mature/established, developing, and starting the journey to become a PLC. The validity of conceptualizing PLCs in three stages was examined by looking at the relationship between schools' four-factor scores and their self-reported stage as a mature, developer, or starter. In some cases, the findings suggest an overlap between mature and developer PLCs—with starter PLCs falling behind on some factors. In addition—not surprisingly—in all cases, the primary school mean factor score for a particular PLC stage is higher than the equivalent secondary school mean factor score, again reflecting the overall tendency for primary schools to have reported higher teacher PLC involvement than secondary schools.

Given these findings, it is not clear whether it should necessarily be expected that primary and secondary phases would report a similar level of staff involvement in the PLC. This is because secondary schools are, usually, much larger and it is possible that smaller PLCs exist within the schools surveyed but were not identified because the focus was at the whole-school level. For example, one or more secondary school subject departments may operate as a PLC and display the characteristics and features of a mature PLC, but the proportion of total school staff involved may be low.

Nevertheless, overall—from these findings—it appears that whole-school PLCs within English secondary schools may be less developed than those within primary and nursery schools, because the results indicate that a smaller percentage of staff are perceived to be engaging in the measured PLC behaviors. Of course, such a response pattern is not uncommon in questionnaires of this type (e.g. McCall et al., 2001), and achieving secondary school improvement is notoriously more difficult than that in primary schools (Louis and Miles, 1990).

Overall, the results from the four PLC factors do appear to be broadly in line with schools' self-reported stage of development—in all cases, schools reported as mature PLCs have higher average factor scores than starter PLCs. Despite this, it is also important to point out that by combining the results from different questionnaire items into summary indicators (i.e., factor scores), a greater degree of differentiation can be employed in measuring each factor than the three crude PLC developmental stages of mature, developers, and starter. This evidence does seem to suggest that the conceptualization of a PLC needs to shift from the idea of a single dimension in three developmental stages to multiple dimensions—each being on a continuum—although the two approaches could also be seen as broadly complementary.

Relationship between pupil outcomes and PLC processes and contexts

The EPLC research conducted by Bolam and colleagues also provides evidence of the relationship between the characteristics of PLCs and pupil outcomes and these findings address the crucial question of whether there is a statistically significant relationship

Table 6 Means and standard deviations of the four PLC factor scores by educational phase.

	Educational phase			
	Nursery and primary schools n 227		Secondary schools n 165	
	Mean	Standard deviation	Mean	Standard deviation
Learning ethos (factor 1)	0.29	0.94	−0.41	0.95
Support for professional development (factor 2)	0.28	1.03	−0.38	0.82
Enquiry orientation (factor 3)	0.22	1.06	−0.32	0.81
Participation of nonteaching staff (factor 4)	0.26	0.82	−0.36	1.12

Factor scores were standardized across the whole sample with mean 0 and standard deviation 1. The primary and the secondary school samples included special schools, respectively. Adapted from Bolam et al. (2005).

between school PLC indicators (e.g., factor scores) and school academic performance measures. As outlined in greater detail below, only tentative evidence of a positive relationship has been found. Nevertheless, the findings support previous research by several researchers. Lee and Smith (1996, p. 103) found that “Achievement gains are significantly higher in schools where teachers take collective responsibility for students’ academic success or failure rather than blaming students for their own failure. Achievement gains were also higher in schools with more cooperation among staff.” Louis and Marks (1998) found that the promotion of a professional community for teachers has a positive relationship with the organization of classrooms for learning and the academic performance of students. Similarly, Wiley (2001) found a positive effect for professional community on student achievement—but only in situations where teachers’ experience of (transformational) leadership was also positive. In line with meta-analyses conducted by Vescio et al. (2008) and Lomos et al. (2011), a more recent study by Doğan and Adams (2018) comparing 13 international studies also reported evidence to support the conclusion that participation in PLCs results in improved teacher practice and increased student achievement. However, they highlight that there are methodological challenges in some PLC studies that limit interpretation of results which still need to be addressed.

Arguably, the most rigorous approach to evaluate the effectiveness of PLCs in promoting student outcomes involves sophisticated statistical analyses (multilevel modeling) conducted to establish whether or not there is a statistically significant relationship between schools’ value added in terms of pupils’ relative progress and their capacity as a PLC. Crucially, despite the limitations that apply to all numerical measures, longitudinal value-added (VA) measures of relative pupil progress provide a more valid measure of educational and school effectiveness than the raw results (Thomas and Mortimore, 1996). However, for comparison purposes, raw (unadjusted) assessment and examination results are typically also employed—as published in the annual league tables in the UK.

Calculating VA and raw school performance

In the EPLC project a variety of different multilevel models were tested prior to deciding on the final ones to use for calculating VA school performance (e.g. see Thomas, 2001). This approach was seen as preferable to assess the long-term VA impact of a school on a pupil’s progress in learning and attainment as the models utilize outcome and baseline data that span, by and large, the period spent in that school phase by pupils. However, both the VA and raw (unadjusted) models have been used to analyze the end of phase schooling outcomes. The Outcomes for primary schools comprise the mean National Curriculum Key Stage 2 (KS2) assessment score across the core subjects of English, mathematics, and science for 11-year-olds. Outcomes for secondary schools comprise the Total National Curriculum Key Stage 4 (KS4) assessment score, including grade scores for all subjects taken by 16-year-olds in the General Certificate for Secondary Education and in General National Vocational Qualifications (GCSE/GNVQ). The final multilevel VA models take account of the variability in pupil outcomes due to the characteristics of the pupil intake for a school. Thus, pupils’ prior attainment, gender, socioeconomic background, and other variables are controlled for in the statistical analysis (see Bolam et al., 2005). The VA score (or residual) for each school is determined from the model results and used to make more robust (reliable) comparisons about how pupils in a school perform compared with pupils in other schools. The raw model is used to calculate school performance in terms of raw pupil outcomes when no adjustment for intake and other factors has been made.

Correlations between school performance and PLC indicators

The four EPLC factors were correlated with primary and secondary schools’ VA and raw performance. However, only a minority of these comparisons were found to be statistically significant at the 0.05 level and the findings are summarized below (see, Bolam et al., 2005, for further details of statistical results). At the primary level, positive and statistically significant correlations were found between schools’ Factor 1 score professional and pupil learning ethos and their 2002 KS2 performance—both raw and VA. At the secondary level, positive and statistically significant correlations were also found for both Factor 1 professional and pupil learning ethos and Factor 2 support for professional learning—but only in terms of schools’ VA GCSE/GNVQ performance. Of course, it should be noted that the relationships are fairly weak (no correlations reported are greater than 0.3). However, it could be argued that this would be expected, given that any statistical relationship is likely to be tenuous between process measures collected as a snapshot at one point in time and pupil performance and progress over a relatively long time period (4–5 years).

The relationship between pupil outcomes and the individual items from part 1 of the survey, were also examined. As expected, the findings support those from the EPLC factor results reported above. Several positive and statistically significant correlations (at 0.05 level) were found, but only two were fairly strong ($r > 0.3$) and these relate only to the primary phase (teachers have low expectations of pupils and teachers share a common core of educational values—note that both of these items are included on Factor 1). There are only three items that are statistically significant (at 0.05 level) across both the primary and secondary phases and the evidence tentatively indicates that these aspects of a PLC are key in terms of enhancing pupil progress. Positive relationships were found between schools’ VA scores and their responses to the following three items: teachers create the conditions for pupils to feel the confidence to learn, teachers share a common core of educational values, and teachers see the school as stimulating and professionally challenging. A positive relationship (at the 0.05 level) was also found between primary VA scores and one additional item: teachers’ experiment and innovate in their work.

Correlations between school performance and PLC factual characteristics

The relationship between pupil outcomes and the PLC factual characteristics, from part 3 of the survey, were also examined in the EPLC project involving English schools. Both positive and negative statistically significant correlations (at 0.05 level) were found—six indicating a fairly strong relationship ($r > 0.3$). These six strongest correlations were found between pupil outcomes and items related to the status of the school—that is, being in an Education Action Zone or part of an Excellence in Cities initiative (negative relationships) or being a Beacon School, having other working links or Specialist status (positive relationships). Thus, a school's involvement in these kinds of initiatives seems to be related to school performance and effectiveness. A school being in an Education Action Zone or Excellence in Cities initiative indicates that it receives additional funding and support due to being located in a particularly economically disadvantaged area. Beacon School or specialist status indicates that a school has met the criteria for excellence and/or additional resources in particular aspects of the curriculum.

In addition, interestingly at the primary level, weak but nevertheless statistically significant positive associations were found between school VA outcomes and the degree to which staff were monitoring pupil progress. This finding supports previous research that underlines the need for systematic monitoring of pupil progress in schools to improve school effectiveness (Thomas et al., 2007). At the primary level, the total number of supply teaching days for 2001/02 was also significantly negatively correlated with raw pupil outcomes suggesting that frequent teacher absences have a detrimental effect on pupil outcomes.

Finally, in addition to the correlational analyses described above, further multilevel analyses were also carried out to determine whether the previously identified PLC factors and other school characteristics have a combined (as well as individual) impact on pupil progress in the VA models (i.e., given pupil intake factors have already been accounted for in the analysis) and whether they further reduced the remaining unexplained school and total variance. In line with the correlational analyses, it was found that, in the primary school analyses, the addition of Factor 1 professional and pupil learning ethos to the VA model was statistically significant (at the 0.05 level); however, this resulted in a very small reduction in school and total variability in KS2 mean score outcomes. Similarly, in the secondary school analyses, the addition of both Factor 1 professional and pupil learning ethos and Factor 2 internal support for professional learning into the VA model, was statistically significant. However, overall, the findings indicated that any positive link between school PLC factors and pupil outcomes is fairly weak and may be difficult to distinguish when other (possibly confounding) school factors and context—such as percentage of pupils entitled to free school meals—is taken into account. In summary, these findings are interesting as they do suggest that a weak, but positive, link exists between individual PLC factors and pupil outcomes—particularly, VA performance, which measures the relative progress of pupils in a school in comparison to pupils in other schools, arguably one of the most valid measures of school effectiveness.

As a further illustration of the relationship between pupil VA measures and PLC/CPD factors, equivalent ITDEQC evidence from one region in an affluent central area of China (LEA1) is reported in Table 7. The findings are tentative and identify association not causality but provide additional support regarding the type of CPD variables that seem to be most significant in explaining secondary student examination outcomes in Chinese, Mathematics and English, once student prior attainment, background and school context have been controlled for in analyses. There is a mixed pattern of positive and negative statistically significant results for CPD factors, which makes the interpretation somewhat difficult given negative results may suggest that more CPD activities are being undertaken because student outcomes are relatively low. Nevertheless, there appear to be more positive than negative impacts associated with CDP factors, including: number of formal CPD hours, participation in CPD related to research projects and new teaching methods, and CPD factors related to collaboration and resources or support. Overall, these results from China ITDEQC project are quite encouraging given, alongside the EPLC findings reported above, they suggest that some CPD activities do appear to have a statistically significant association with students VA progress at school. Moreover, the significant factors seem to be those typically associated with the concept of professional learning communities—such as teacher collaboration.

Providing a more rigorous analysis, correlational and multilevel modeling findings of the kind reported above are important because, arguably, particular PLC and CPD factors can only be validated as effective if a positive impact on student progress and outcomes can be demonstrated. On the other hand, it may be that quantitative measures used in the research conducted so far are not sufficiently sensitive to fully reflect the complex relationship among PLCs and pupil performance and progress. Therefore, further research is necessary to extend the knowledge base in a broader range of contexts and contribute stronger evidence on this issue which has only rarely been addressed using sophisticated quantitative techniques.

Conclusion

In conclusion, a limited number of relevant studies have been reviewed and the findings indicate that some key PLC characteristics in England are more (or less) commonly reported in schools (e.g., opportunities for work shadowing and experience job rotation were, generally, less common). In addition, some key PLC characteristics are more-or-less variable among some schools than others (e.g., dedicated time for classroom observation and protected time for joint planning and development were more variable). These findings reveal an important message for policymakers—that there are some key areas of teachers' professional development that either happen very infrequently in schools or are highly variable across schools. Moreover, evidence comparing PLC characteristics in England and China indicates that context specificity is an issue, pointing to the need for the PLC concept to be refined and adapted in different country contexts.

Table 7 Statistically Significant teacher questionnaire CPD variables related to value added performance (2012).

<i>CPD variables</i>	<i>LEA1</i>		
	<i>Chinese</i>	<i>English</i>	<i>Maths</i>
Total formal CPD hours		+	
Total informal CPD hours			
Participation in qualification program			
Participation in a network of teachers formed			
Participation in individual or collaborative research			
Participation in mentoring and/or peer observation and coaching	-		-
Participation in reading professional literature			-
Participation/Engaging in informal dialogue with your colleagues			+
Participation in research project, writing papers		+	+
Participation/Engaging in activities and planning			
Participation in how to teach in a multicultural setting			
Participation in instructional strategies and teaching methods	+	+	
Participation in teaching content and performance standards			
Participation in class management	-		
Participation in student discipline and behavior problems			
Participation in how to counsel students			
% CPD activities at school level	+	+	
Factor of perceived future CPD need related to student development and class management			
Factor of perceived future CPD need related to curriculum and teaching methods			
Factor of frequency appraisal received			
Factor of CPD through collaboration and feedback	+	+	
Factor of CPD through educational resources and support			+

Note: CPD variables (– negative coefficient; + positive coefficient) statistically significant (0.05) in explaining student outcomes (University entrance examinations), once student prior attainment, background and school context have been controlled for in the multilevel model (VA Model Xc). LEA1 sample: Schools 40 teachers 326 Pupils 28,197. See [Thomas \(2020\)](#) for further details of analyses.

Regarding PLC measurement, the [Bolam et al. \(2005\)](#) and other similar studies found that the summary indicators of different aspects of a PLC can be identified using quantitative factor-analysis techniques. However, it is important to recognize that the different factors identified are not, necessarily, comprehensive for the purpose of evaluating the nature and quality of PLCs. The qualitative case-study data, also reported by [Bolam et al. \(2005\)](#), tapped into considerable detail and depth concerning the nature of PLCs and, subsequently, revealed important new aspects of PLC processes and characteristics such as trust and distributed leadership, which were not directly included on their questionnaire survey. Moreover, in Bolam and colleague's final conclusions, it was suggested that PLCs should be evaluated not only in terms of pupil outcomes—indeed, overall the quantitative evidence linking student outcomes and PLC functioning, although positive seems to be fairly weak—but also via impact on staff morale and practice, and potential for developing leadership capacity, as well as by establishing that PLC characteristics are in place and processes are operating smoothly. Further research is, of course, needed to explore these issues in more detail, given that future theory building will require stronger evidence of this kind to clearly differentiate between the effects of the wide varieties of learning, development, and improvement activities that teachers, schools, and PLCs engage in.

This article indicates fairly strongly that a school, as a PLC, is multidimensional and complex in nature. Evidence also, tentatively, suggests that some PLC characteristics and processes may be enhanced or hindered by the advantages or disadvantages of particular contexts (e.g., smaller size may allow greater opportunities for collaboration, higher percentage of student disadvantage indicated by their eligibility for free school meals may indicate higher funding levels for professional development activities). The idea of three developmental stages (mature, developer, and starter) may be helpful as a starting point in understanding the PLC concept but, in the end, may be too crude to measure the complexity of PLC characteristics and processes. This article demonstrates that a more detailed approach is needed to help teachers and other professionals working in schools to understand the conditions

and contexts that underpin systematic and effective professional learning as well as to enhance PLC development in schools. Further evidence is also needed to better understand the links between professional learning and the learning of other members of the school community—such as students and parents—to explore whether the processes and outcomes of a PLC mirror those of the school learning community as a whole. Interestingly, fruitful areas of new research have also developed the idea of PLCs to look more at external links via Networked Learning Communities (Katz and Earl, 2010; Poortman et al., 2021; Prenger et al., 2019) and Research Learning Communities (Dimmock, 2016; Rose et al., 2017).

Overall, the review findings indicate that quantitative and qualitative methods can be used to evaluate different aspects of the effectiveness and quality of a PLC and demonstrates the potential for creating valid measures of professional learning within schools in England, China and elsewhere. These kinds of measures could provide a useful method of enhancing information available to schools and, if used alongside other types of evidence, could meaningfully inform school improvement and self-evaluation activities (see Stoll et al. (2006b) and ITDEQC (2018) for examples of PLC evaluation materials produced for use in English and Chinese schools).

References

- Admiraal a, W., Schenkeb, W., De Jonga, L., Emmelotb, Y., Sligteb, H., 2021. Schools as professional learning communities: what can schools do to support professional development of their teachers? *Prof. Dev. Educ.* 47 (4), 684–698.
- Antinluoma, M., Ilomäki, L., Lahti-Nuuttila, P., Toom, A., 2018. Schools as professional learning communities. *J. Educ. Learn.* 7 (5), 76–91.
- Bolam, R., McMahon, A., Stoll, L., et al., 2005. Creating and Sustaining Effective Professional Learning Communities. DfES Research Report RR637. University of Bristol. <http://www.dfes.gov.uk/research/data/uploads/files/RR637.pdf>. (Accessed September 2009).
- Chen, P., Che-Di Lee, Lin, H., Zhang, C.-X., 2016. Factors that develop effective professional learning communities in Taiwan. *Asia Pac. J. Educ.* 36 (2), 248–265.
- Chen, L., 2020. A historical review of professional learning communities in China (1949–2019): some implications for collaborative teacher professional development. *Asia Pac. J. Educ.* 40 (3), 373–385.
- DfE, 2016. Standard for Teachers' Professional Development. Department for Education.
- DfE, 2020. Headteachers' Standards 2020. Department for Education. <https://www.gov.uk/government/publications/national-standards-of-excellence-for-headteachers/headteachers-standards-2020#section-2-headte%E2%80%A6>.
- Dimmock, C., 2016. Conceptualising the research–practice–professional development nexus: mobilising schools as 'research-engaged' professional learning communities. *Prof. Dev. Educ.* 42 (1), 36–53.
- Doğan, S., Adams, A., 2018. Effect of professional learning communities on teachers and students: reporting updated results and raising questions about research design. *Sch. Effect. Sch. Improv.* 29 (4), 634–659.
- Hairon, S., Goh, J.W.P., Chua, C.S.K., Wang, L., 2017. A research agenda for professional learning communities: moving forward. *Prof. Dev. Educ.* 43 (1), 72–86.
- Hipp, K., Huffman, J., Pankake, A., Olivier, D., 2008. Sustaining professional learning communities. *J. Educ. Change* 9, 173–195.
- Hofman, R.H., Hofman, W.H.A., Gray, J.M., 2015. Three conjectures about school effectiveness: an exploratory study. *Cogent Education* 2, 1.
- Hord, S.M., 1997. Professional Learning Communities: Communities of Continuous Inquiry and Improvement. Southwest Educational Development Laboratory.
- IEEQC, 2018. IEEQC and ITDEQC Research Impact. <http://www.bristol.ac.uk/education/research/sites/ieeqc/research-impact/>.
- Katz, S., Earl, L., 2010. Learning about networked learning communities. *Sch. Effect. Sch. Improv.* 21 (1), 27–51.
- Lee, V.E., Smith, J.B., 1996. Collective responsibility for learning and its effects on gains in achievement for early secondary students. *Am. J. Educ.* 104, 103–147.
- Lomos, C., Hofman, R.H., Bosker, R.J., 2011. Professional communities and student achievement: a meta-analysis. *Sch. Effect. Sch. Improv.* 22 (2), 121–148.
- Lomos, C., 2017. To what extent do teachers in European countries differ in their professional community practices? *Sch. Effect. Sch. Improv.* 28 (2), 276–291.
- Louis, K.S., Marks, H., 1998. Does professional community affect the classroom? Teachers' work and student experience in restructured schools. *Am. J. Educ.* 106 (4), 532–575.
- Louis, K.S., Miles, M.B., 1990. Improving the Urban High School: What Works and Why. Teachers College Press, New York.
- Louis, K.S., Marks, H.M., Kruse, S., 1996. Teachers' professional community in restructuring schools. *Am. Educ. Res. J.* 33 (4), 757–798.
- McCall, J., Smith, I., Stoll, L., et al., 2001. Views of pupils, parents and teachers: vital indicators of effectiveness and for improvement. In: MacBeath, J., Mortimore, P. (Eds.), *Improving School Effectiveness*. Open University Press, Buckingham.
- OECD, 2016. School Leadership for Learning: Insights from TALIS 2013. OECD Publishing, Paris.
- Pang, N.S., Wang, T., Leung, Z.L.-M., 2016. Educational reforms and the practices of professional learning community in Hong Kong primary schools. *Asia Pac. J. Educ.* 36 (2), 231–247.
- Pedder, D., James, M., MacBeath, J., 2005. How teachers value and practise professional learning. *Res. Pap. Educ.* 20 (3), 209–243.
- Poortman, C.L., Brown, C., Kim, S., 2021. Professional learning networks: a conceptual model and research opportunities. *Educ. Res.* <https://doi.org/10.1080/00131881.2021.1985398>.
- Prenger, R., Poortman, C.L., Handzelalts, A., 2019. The effects of networked professional learning communities. *J. Teach. Educ.* 70 (5), 441–452.
- Qiao, X., Yu, S., Zhang, L., 2018. A review of research on professional learning communities in mainland China (2006–2015) Key findings and emerging themes. *Educ. Manag. Adm. Leader* 46 (5), 713–728.
- Rose, J., Thomas, S., Zhang, L., Edwards, A., Anwandter, A., Roney, P., 2017. Evaluation of Research Learning Communities. Educational Endowment Foundation.
- Scheerens, J., Glas, C., Thomas, S., 2003. Educational Evaluation, Assessment and Monitoring: A Systemic Approach. Swets & Zeitlinger, Tokyo.
- Sleegers, P., den Brok, P., Verbiest, E., Moolenaar, N.M., Daly, A.J., 2013. Toward conceptual clarity: a multidimensional, multilevel model of professional learning communities in Dutch elementary schools. *Elem. Sch. J.* 114 (1), 118–137.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., Thomas, S., 2006a. Professional learning communities: a review of the literature. *J. Educ. Change* 7 (4), 221–258.
- Stoll, L., Bolam, R., McMahon, A., et al., 2006b. Professional Learning Communities: Source Materials for School Leaders and Other Leaders of Professional Learning. Innovation Unit, DfES, NCSL and GTC, London.
- Stoll, L., Louis, K.S. (Eds.), 2007. Professional learning communities: elaborating new approaches. In: *Professional Learning Communities*. Open University Press, Maidenhead, pp. 1–14.
- Stoll, L., McMahon, A., Bolam, R., et al., 2003. Creating and Sustaining Effective Professional Learning Communities – Questions Arising from the Literature and How They Are Being Addressed in the Project. DfES research Brief no. RBX 12-03.
- Thomas, S.M., 2001. Dimensions of secondary school effectiveness: comparative analyses across regions. *Sch. Effect. Sch. Improv.* 12 (4), 285–322.
- Thomas, S.M., 2020. School and teacher value added performance and the relationship with teacher professional development in mainland China. In: Hall, Sammons, Lindorff (Eds.), *International Perspectives in Educational Effectiveness Research*. Springer.
- Thomas, S., Mortimore, P., 1996. Comparison of value added models for secondary school effectiveness. *Res. Pap. Educ.* 11 (1), 5–33.

- Thomas, S.M., Peng, W.-J., Gray, J., 2007. Value added trends in English secondary school performance over ten years. *Oxf. Rev. Educ.* 33 (3), 261–295.
- Thomas, S.M., Peng, W.-J., Li, J., 2015. 中国学校增值表现趋势研究—兼评增值与教师专业发展的关联性 (time trends in school value added performance and the relationship between value added and teachers professional development in China). *Educ. Res.* 35 (7), 64–72 (National Institute for Education Sciences: Beijing).
- Thomas, S.M., Peng, W.-J., Triggs, P., 2017. Professional learning communities in Chinese senior secondary schools. In: Harris, A., Jones, M., Huffman, J. (Eds.), *Teachers Leading Educational Reform: The Power of Professional Learning Communities*. Routledge, Abingdon UK.
- Vescio, V.A., Ross, D., Adams, A., 2008. A review of research on the impact of professional learning communities on teaching practice and student learning. *Teach. Teach. Educ.* 24 (1), 80–91.
- Visscher, A.J., Witziers, B., 2004. Subject departments as professional communities? *Br. Educ. Res. J.* 30 (6), 785–800.
- Wiley, S., 2001. Contextual effects of student achievement: school leadership and professional community. *J. Educ. Change* 2 (1), 1–33.
- Zahedi, S., Bryant, C.L., Iyer, A., Jaffer, R., 2021. Professional learning communities at a primary and secondary school network in India. *Asia Pac. Educ. Rev.* 22, 291–303.
- Zhang, J., Sun, Y., 2018. Development of a conceptual model for understanding professional learning communities in China: a mixed-method study. *Asia Pac. Educ. Rev.* 19, 445–457, 2018.

Further reading

- Stoll, L., McMahon, A., Thomas, S., 2006. Identifying and leading effective professional learning communities. Special issue: leadership and student learning in professional learning communities. *J. Sch. Leader.* 16 (5), 611–623.

Relevant website

<http://www.bristol.ac.uk/education/eplc/>—Creating and Sustaining Effective Professional Learning Communities.

International trends on teacher evaluation

Leonidas Kyriakides, Evi Charalambous, and Tryfonas Papadamou, Department of Education, University of Cyprus, Nicosia, Cyprus

© 2023 Elsevier Ltd. All rights reserved.

Introduction	287
Developing a comprehensive teacher evaluation system	288
Evaluation purposes	288
Evaluation criteria	288
Sources for collecting relevant data	290
Establishing meta-evaluation mechanisms to support teacher evaluation systems	290
Meta-analysis of studies on teacher and school evaluation: an overview	291
Methods	291
Selection of studies	291
Criteria for including studies	292
Main findings: descriptive analysis	292
Conclusions	294
References	295

Introduction

One of the most crucial problems that almost all educational systems around the world are facing concerns the need for developing an effective teacher evaluation system based on a valid theoretical framework. Existing teacher evaluation systems fail to either identify effective teachers (Flores and Derrington, 2018; Rockoff, 2004; Sanders et al., 1997) or contribute to teacher professional development (Creemers et al., 2013). This is mainly because links between teacher evaluation policies and research on educational effectiveness and improvement have not been established. The use of empirical evidence in introducing or implementing new mechanisms of teacher evaluation can be questioned (Huber and Skedsmo, 2016). Hallinger et al. (2014) argue that the policies and politics driving teacher evaluation processes remain considerably stronger than the empirical evidence. Consequently, existing teacher evaluation systems are not able to promote educational quality since they are unable to support teachers become more effective and through that promote student learning outcomes (Peterson, 2000; Weisberg et al., 2009). Thus, by pursuing policymakers to establish an effective teacher evaluation system in their countries, will eventually enhance each country's educational outcomes since student achievement may be promoted.

Teacher quality is seen as the highest priority factor for affecting student achievement (Hattie, 2009; Looney, 2011; Nye et al., 2004; Muijs et al., 2014) and studies in educational effectiveness research (Coleman et al., 1966; Gordon et al., 2006; Rivkin et al., 2005; Rockoff, 2004; Muijs and Reynolds, 2011) describe the importance of maintaining high quality of teaching so that students can improve their learning outcomes. Therefore, the field of educational effectiveness turned its focus on elaborating on the features of teacher evaluation (Mortimore, 2001). However, the field has become rather ethnocentric, with groups of scholars working in different countries pursuing isolated agendas, largely uninfluenced by the agendas in other countries (Reynolds et al., 2002; Chapman et al., 2016). This situation can partly be attributed to the lack of specific criteria for measuring teacher effectiveness (Kyriakides and Campbell, 2003). Unless the criteria for evaluating teachers emerge from tested theories on teacher and school effectiveness, evaluators cannot be held accountable for how their judgements about teacher performance have been arrived at. The main theoretical models of teacher effectiveness research (TER) can be seen as a source for generating a set of criteria for teacher evaluation that captures the multiple teacher roles in changing the educational environment.

Teacher evaluation can be described as "the systematic assessment of a [teacher]'s performance and/or qualifications in relation to a professional role and some specified and defensible institutional purpose" (Gullickson and Howard, 2009, p. 27). Teacher evaluation is seen as a systematic process, which means that it is purposeful, it requires preparation, and specific products/outcomes are expected to occur (Liu, 2011). Consequently, data on teachers' performance need to be collected and analyzed through precise methods even if the tools for collection and the type of data collected might differ. Moreover, teacher evaluation should aim to improve teachers' professional development by providing them with an understanding of both how they are performing based on the expectations and what needs to be done to perform better (Tarhan et al., 2019). Nowadays, educational policies are introduced to evaluate teachers and their performance to achieve public accountability of schools and teachers (Huber and Skedsmo, 2016). Nevertheless, a second aim of teacher evaluation is the development of teachers and their practice. Teacher evaluation processes may assess not only individual teacher performance but also the achievement of school and/or educational system goals (Stronge, 2006b; Hallinger et al., 2014). These different approaches to teacher evaluation are formed by the relative emphasis education systems place on the use of results for either accountability (e.g., high-stakes rewards and/or sanctions) and/or for improvement purposes (e.g., professional development targeted to identified needs and support for innovation) (Looney, 2011; Robertson-Kraft and Zhang, 2018). Thus, it is important to set priorities on features of evaluation that are more important than

others and to investigate which combinations of policies and actions situated on different levels enhance educational effectiveness, so as to establish successful reforms on evaluation.

In this context, the main elements of a comprehensive teacher evaluation system are described in the next section of this chapter. In the third section of the chapter, we argue for establishing meta-evaluation mechanisms when designing a teacher evaluation system. In the fourth section, an overview of the findings of a recent meta-analysis of 156 international studies on teacher evaluation is provided pointing out the main aspects that should be taken into account when conducting research to design a valid and reliable evaluation system. In the last section of the chapter, implications for research, policy and practice on teacher evaluation are drawn.

Developing a comprehensive teacher evaluation system

There are three main aspects that need to be taken into account when developing a comprehensive teacher evaluation system: (1) *the evaluation purposes*, (2) *the evaluation criteria* and (3) *the sources for collecting relevant data* (Ellett et al., 1996). It is important to note that the establishment of criteria depends on the rationale for which the evaluation is conducted and on the availability of appropriate sources of data, thus all three main aspects are interrelated (Kyriakides et al., 2006).

Evaluation purposes

A *twofold purpose of teacher evaluation* is recognized in the literature (see Barile et al., 2012; Kimball et al., 2004; Rigby, 2015; Steinberg and Donaldson, 2016; Santelices and Taut, 2011; Vekeman et al., 2015; Zhang and Ng, 2011) since teacher evaluation is intended to hold teachers accountable (i.e., summative purpose) and/or to help them in their professional development (i.e., formative purpose). These two general functions of any teacher evaluation system may refer to more specific aims that might include licensing/credentialing, tenure, self-assessment, and professional development (Beerens, 2000; Harlen and James, 1997; Stronge, 2006a). Sometimes combinations of purposes may apply to a teacher evaluation system, contributing to either completing multiple tasks or creating unnecessary deficiency to the initial task. This is because, on the one hand, the accountability purpose reflects the need to verify the ability of teachers to ensure that services delivered are effective, and this has typically been considered to be summative in nature, and, on the other hand, the improvement purpose reflects the need for the professional growth and development of the individual teacher, and typically has been considered to be formative in nature. Further to the summative purpose, teachers are held accountable (by rewarding and/or “punishing” them) based on their students’ performance, and it is implied that this will improve their effectiveness (Amrein-Beardsley, 2014; Papay, 2012; Paufler, 2018). Accountability-focused evaluations regularly tend to rate and rank teachers and these practices are, thus, used to develop decisions related to compensation and staff issues (Goldstein, 2014). Summative assessment practices were found, though, to have a negative impact on teachers, especially on psychological aspects such as increased pressure (Collins, 2014) and reduced morale (Collins, 2014; Paufler, 2018).

When teacher evaluation is used for formative purposes, evaluation is related mainly to the provision of constructive feedback that contributes to teachers’ professional growth and improvement (Glickman et al., 2014). This form of teacher evaluation can have invaluable positive effects on teacher instructional practices (Delvaux et al., 2013; Taylor and Tyler, 2012). Moreover, teachers have indicated that a combination of practices such as classroom observations accompanied with formative feedback was crucial for improving teacher performance and their morale for teaching (Anast-May et al., 2011). Although there are studies reporting on the frequent use of feedback and its positive outcomes (Prado Tuma, et al., 2018), other researchers emphasize the lack of such an approach during the evaluation process (Donaldson, 2012, 2016; Weisberg et al., 2009). It can, therefore, be claimed that teacher evaluation without formative feedback is unlikely to improve the quality of teaching (Goe et al., 2012; Papay, 2012).

However, it is practically impossible that the dual character of evaluation can be achieved within a single teacher evaluation mechanism, since determining the evaluation purpose inevitably influences the design of the evaluation methods and instruments, as well as the interpretation of the results (Black and Wiliam, 1998; Kyriakides et al., 2021). Given that, a comprehensive system of teacher evaluation should be built serving both purposes of evaluation, as it is important to give emphasis on both outcomes and improvement (McGreal, 1988), policy makers should be able to establish different mechanisms of formative and summative teacher evaluation. Since the summative and formative purposes of evaluation are based on different value assumptions, different criteria might be used for conducting each type of teacher evaluation. The relevance of the teacher evaluation criteria cannot be judged without the purpose for which evaluation is conducted being made explicit. However, in order for the teacher evaluation systems to serve both purposes, there must be a rational link between them (Stronge, 1995); this should not allow the summative function of evaluation to dominate the formative function, as the experiences of many countries seem to reveal (e.g., Ellett et al., 1996; Kyriakides et al., 2018). Additionally, this link cannot be established, unless the criteria for both the formative and summative evaluation mechanisms are based on the same theoretical framework concerning what constitutes an effective teacher. Consequently, in the next part of this section, we argue on generating criteria for teacher evaluation.

Evaluation criteria

Several sources could guide the development of *evaluation criteria*, including teachers’ job descriptions and their professional code (Johnson, 1997), the theoretical models of TER (Cheng and Tsui, 1999), as well as teachers’ opinions (Patterson et al., 1986; Firestone, 1989). However, in many countries there is a lack of professional code and the teacher’s role is vaguely described. Also, each

country according to its teaching tradition may contain different value assumptions on teacher evaluation, thus creating a universal set of criteria is not an easy task. Therefore, we argue that TER could be the basis upon which measurable criteria of teacher evaluation could be established since the characteristics of effective teachers and their multiple features in the educational environments are captured through the different theoretical models of teacher effectiveness (Fessler and Ungaretti, 1994; Kyriakides et al., 2002; Murphy, 1995). These theoretical models are presented below in an effort to outline both the advantages and disadvantages of using each model to generate criteria for teacher evaluation. Identifying the strengths and weaknesses of employing the various TER models in developing criteria for teacher evaluation, helps policy makers to take advantage of their best features and determine well-suited evaluation mechanisms for each context.

- (a) *Goal and Task Model*: The goal and task model conception of teacher effectiveness is based on the formation of aims and tasks that are justified through the achievement level of students. A distinct advantage of this model is the fact that goals and tasks are defined accurately. This model has also the advantage that the outcomes of teachers' performance are easily observed since policymakers could evaluate teachers directly based on their students' learning outcomes. In this sense, performance standards should exist. Indicators of teacher effectiveness that may arise through such an approach include achievement of teaching objectives, fulfilment of job specifications, performance standards, student learning outcomes and many more (Cheng and Tsui, 1999). Furthermore, the model relies on the genuine effort of teachers to achieve their goals and teacher effectiveness is set to be perceived to the extent that educators are able or not to reach the goals initially set. This model could express the demands of summative evaluation as it enables evaluators to examine specific dimensions of the teachers' work and thus offers the possibility of comparable results. Despite the obvious advantages of the model, a series of concerns might arise when applying this evaluation model mainly concerning fairness issues for setting complicated or numerous tasks in small time frames as well as employing methods that are not pedagogically appropriate but help students to achieve the expected goals. Moreover, questions are raised on the types of goals that are meaningful both for the school and teachers and on who is responsible for setting these goals (Kyriakides et al., 2020).
- (b) *Working-Process Model*: The working process model emphasizes on the smooth teaching and working processes. This model is basically supporting that a number of educational actions by teachers such as effective teaching and the learning process occurring to students are fulfilling the predetermined objectives measured through the various types of students' learning outcomes. Hence, effective teachers are described as those that can ensure the quality of their teaching (Creemers and Kyriakides, 2008). This model refers to the "process-product" model, which is responsible for the classification of a range of teacher behaviors, which are linked to the improvement of student achievement gains (Brophy and Good, 1986; Doyle, 1986; Everston et al., 1980; Galton, 1987). Factors that seem to carry evidence of significance are quantity of instruction and likewise, the pace of instruction (Brophy and Good, 1986). In this model, effective teachers are those that are able to create a learning environment in their classroom that promotes their students' learning outcomes (Creemers and Reezigt, 1996; Doyle, 1986). Under this condition, student achievement is linked to variables that refer to the form and quality of teachers' educational planning and instruction, such as structuring of the lesson, asking questions (soliciting), providing feedback (reacting) and providing opportunities to engage students in the learning process (Kyriakides and Demetriou, 2007). Therefore, this model could be appropriate to assess teachers according to their teaching behaviors, the classroom environment that they create, their willingness to participate in decision making and the working relationship with other key actors in the school. However, it is implied that by using this model as the unique source of teacher evaluation may not provide valid results on examining the appropriateness of the educational aims initially set. This is because there are questions regarding to what standards are teachers who operate in different contexts evaluated and how are these standards related to students' progress, given that students have diverse instructional needs, due to their different learning styles and personality types.
- (c) *Resource-Utilization Model*: The resource utilization model refers to a more bureaucratic approach and is based on the use of allocated resources and acquisition of inputs to produce a desirable output. In this model, it is made clear that the relationship of resources to teaching and its outputs is established, so as to explain an effective teacher evaluation process. Teachers are expected to complete numerous tasks in their school, usually under time pressure. Hence, the provision of the appropriate resources is considered of crucial importance for maximizing teacher effectiveness. According to this model teachers need to utilize existing resources but also, locate new resources, especially when there is evidence of linking those resources with achieving desirable outcomes. However, there are questions related to the kind of resources to be used, the availability of them, if they are sufficient for the whole staff of each school and most importantly if these resources facilitate teacher performance and education outcomes.
- (d) *The School-Constituencies Satisfaction Model*: This model links teacher quality with "client (e.g. students, parents) satisfaction", which is basically demonstrated through the clients' expectations and demands. Teachers are considered effective when their performance satisfies the students and parents (Peterson et al., 2000). This is a model that could provide solutions on the weaknesses of the other theoretical models, due to the various criteria for evaluating teachers that it contains, based on the different needs of the various education stakeholders. However, when employing this model, there are issues that need to be considered such as who are the key stakeholders that strongly influence school functioning and teachers' work and how their different needs and expectations can be related to the predefined objectives and duties of the teachers.
- (e) *The Accountability Model*: The focus of this model is to demonstrate competence and responsibility of teachers in their instruction and other educational practices in school in conjunction with related professional decisions. This model is based on the assumption that the overall quality of the educational system is ensured by demonstrating evidence of accountability

(Scheerens, 2013). Therefore, this model is in various aspects similar to the school-constituencies model, since effective teachers are identified by the satisfaction that they provide to other stakeholders involved in the educational process. The concept of this model requires that teachers are reporting to the various school constituencies about their work, the decisions they make and their standards. To that end, the usefulness of the model relies on the case that teachers are requested to provide evidence of accountability, or even further, in the case that a process of monitoring or evaluation occurs, accountability measures are the focus for many educational systems. The present model includes new factors to teachers' evaluation such as keeping social and school stakeholders informed on the progress of schools' performance. These kinds of factors may raise negative reactions from teachers since there is a high possibility of perceiving this form of monitoring as an intervention in their professional role.

- (f) *The Absence of Problems Model*: This model constitutes that the absence of problems in the duties of teachers creates a framework of effectiveness. Accordingly, effective teachers are defined as those who satisfy the minimal requirements, for example the personnel that do not cause any problems or conflict by their actions in their professional environment and furthermore are able to cope with conflicts that may arise in the school environment. The usefulness of this model relies on the specification of certain criteria for teacher evaluation referring to certain problematic areas of schooling. The model is mainly used by evaluators to identify weaknesses and problematic aspects of school life and make suggestions for improvement.
- (g) *The Continuous Learning Model*: This model assumes that teacher professional development is linked to school improvement (Fullan and Hargreaves, 1992) and that professional development programs can help teachers become more effective. The model proves to be effective especially in cases where teachers change educational settings and are in need of adapting to constant new challenges. Evaluators, though, should have in mind that teachers' adjustment to new school environments may not be the single reason for improvement, and before setting up any type of professional development, they need to be sure that these changes derive from existing program objectives.

Sources for collecting relevant data

In terms of the *sources for collecting evaluation data*, high performance educational systems nowadays tend to prefer multiple sources of data for evaluating their personnel. This multidimensional approach for collecting different types of data (see Danielson and McGreal, 2000) may increase the internal validity of an evaluation system (Cronbach, 1990). Studies in the field of TER stress the need for implementing multiple methods for assessing the impact of teachers' instruction upon student achievement as it aims at gaining a greater variation of teacher effectiveness. Such methods include techniques like classroom observations, value-added measures of the contribution of individual teachers, student ratings and self-evaluation processes (see Creemers and Kyriakides, 2012; Creemers et al., 2010). Classroom observation is widely recognized as a mean to teacher evaluation and teacher development. Observational systems in which quality observational instruments, well-trained raters, and vigorous scoring designs are combined, can produce reliable teacher scores (Hill et al., 2012). To achieve this, generalizability theory (Brennan, 2001; Cronbach et al., 1972; Marcoulides, 1989; Shavelson and Webb, 2006) can assist in the design of observation systems that produce reliable scores. This is because, generalizability theory can provide a comprehensive framework for making decisions about the multiple elements of observational systems and provide empirical evidence regarding the optimal number of raters and lessons required to produce desired consistencies. In addition, student ratings of teacher performance can provide information about teachers' behavior in the classroom. As students are direct recipients of the teaching-learning process, they are in a key position to provide information regarding the development of motivation in the classroom, opportunities for learning, degree of understanding and communication developed between them and the teacher (Carle, 2009; Kyriakides et al., 2014; Marsh, 1987). However, in spite of the several strong reasons for using student ratings to evaluate teachers (Lasagabaster and Sierra, 2011), ratings of teacher performance remain suspect as a means of evaluating instructional effectiveness and very little effort has gone into the development of principles and practices in relation to this source of data in primary and secondary education. Nevertheless, there are several effectiveness studies that have shown that primary students are capable of providing valid data on the classroom behavior of their teachers (Bacher-Hicks et al., 2019; Elstad et al., 2017; Kyriakides et al., 2014; Peterson et al., 2000). Of course, it is argued that at the very first stage of analysis of student ratings, researchers need to investigate the generalisability of such ratings, especially when single scores per teacher for each teaching skill are generated.

To conclude, it is important to stress that adopting a multiple-source approach for collecting evaluation data may cause problems of practicality and validity. For instance, evaluators may encounter difficulties in developing psychometric tools based on data gathered from different sources that may incorporate different sources of measurement error associated with each source of data (Kyriakides et al., 2006). Hence, it is a challenge for evaluators to locate the appropriate data sources for an effective teacher evaluation system by selecting the data through a criterion-based approach suggested by theory that supports TER. Thus, it is important that there is transparency on stakeholders' perceptions on the reliability and validity of data collection, before deciding what evaluation scheme should be deployed.

Establishing meta-evaluation mechanisms to support teacher evaluation systems

When designing a teacher evaluation system, monitoring is the key factor of the process for assuring quality of the desired outcomes. Either in the case of setting up a whole teacher evaluation system or redesigning parts of an existing evaluation system, establishing

a quality assurance policy is imperative. Thus, it is important to examine whether evaluation systems set up such meta-evaluation mechanisms so as to identify their strengths and weaknesses and take actions in order to further improve the mechanisms as well as the evaluators' skills to achieve their evaluation aims.

It is not a given that every evaluation system is properly designed, just as it is not a given that the evaluation results are always successful, and the suggestions derived from these results are always valid and reliable. To ensure that an evaluation system is working properly, meta-evaluation procedures should be used (Cooksy and Caracelli, 2009; Stufflebeam, 2011). Scriven (1969) introduced the term "meta-evaluation" underlining the importance of evaluating the evaluation processes, but since then only few educational systems around the world developed meta-evaluation procedures. One can claim that evaluating an educational system or a part of it is a challenging process depending on the policy makers' and various stakeholders' consent. Often teachers and other bodies show a negative reaction to new evaluation reforms hindering their implementation. However, data has shown that the evaluators tend to be more responsible when they know that their performance will be evaluated in the end (Scriven, 2009; Stufflebeam, 2011).

The meta-evaluation procedure has generally a two-fold purpose: (a) to help the evaluators to design, implement, and improve their evaluation system and to properly interpret the evaluation results (formative purpose), and (b) to identify the strengths and weaknesses of an evaluation system (summative purpose). Also, there are two main forms of meta-evaluation that the educational systems seem to apply. First, there is an integrated measure of meta-evaluation, where the process co-exists with the teacher evaluation process but still it is distinctly separated. This form of meta-evaluation is designed alongside the teacher evaluation mechanism; thus, the elements of the evaluation process direct the nature of the meta-evaluation. Evaluators in this concept are mainly external evaluators, but there are also opportunities for enhancing other approaches too. Second, there is an independent form of meta-evaluation where it is composed separately from the main evaluation procedure, initiated by the constituency that ordered the evaluation process in the first place. Choosing the form of meta-evaluation also determines its time of occurrence. Elaborating on the time of evaluation occurrence there are two possible scenarios: a) in the case of integrated meta-evaluation it runs simultaneously with the main evaluation process and b) in the case of independent meta-evaluation there are scenarios where it runs simultaneously alongside to the main evaluation process, but mostly after the completion of it (Fitzpatrick et al., 2004).

Meta-evaluation processes are reported only rarely in the related literature and there is a number of reasons for this. The most important reason is oriented toward the resistance of stakeholders participating in the design of a teacher evaluation process and providing accountability transparency. In this perspective it should also be mentioned that an evaluation process incorporates a wide range of stakeholders. Moreover, the actions of evaluators have either direct or indirect impact on the participating counterparts. Hence, developing meta-evaluation systems causes the political dimension of evaluation to become the main point of discussion (Pfeffer, 1992; Stronge and Tucker, 2000; Fitzpatrick et al., 2004).

To sum up, an effective meta-evaluation system needs to be characterized by a wide consensus among all the interested parties. Decisions should be taken not only by policymakers, but also by other stakeholders that take part in the evaluation process. Creating consensus could be achieved by simulating the impact of meta-evaluation outcomes on the related parties and speculating the probable reactions that may be caused during and/or after its occurrence.

In the next section of this chapter, a brief overview of the findings of a recent meta-analysis of international studies on teacher evaluation is presented. This meta-analysis investigates the extent to which research on teacher evaluation covers the main aspects of the proposed comprehensive framework on teacher evaluation. It also helps readers to identify the international dimension of research on teacher evaluation. Suggestions for further research are also provided in the last section of this chapter.

Meta-analysis of studies on teacher and school evaluation: an overview

Methods

In the first part of this section, the main methods of the meta-analysis such as the selection of studies and the criteria for including specific studies in the analysis, are briefly presented. The way that data from each study were coded and entered in the analysis as well as the estimation of effect sizes of the results reported in those studies investigating the impact of teacher evaluation and the specific information on the multilevel model used to conduct the meta-analysis are not mentioned since this information is beyond the scope of this chapter.

Selection of studies

For the selection of studies included in this meta-analysis we searched documentary databases containing abstracts of empirical studies. The following databases were considered: Educational Resources Information Center, Educational Research Abstracts Online, Social Sciences Citation Index, Educational Administration Abstracts, SCOPUS, ProQuest 5000, and PsycArticles. We also paged through volumes of education-focused peer-reviewed journals with interest in teacher evaluation, teacher professional development, and effective teaching, including the journals "School Effectiveness and School Improvement", "Educational Assessment, Evaluation and Accountability", "American Educational Research Journal", "Educational Management Administration and Leadership", "Educational Policy", "Studies in Educational Evaluation" and "Assessment in Education: Principles, Policy and Practice". The key term "teacher evaluation" was used and a total of 1766 articles appeared as related to the topic. From these articles, 156 were eventually selected as they were relevant to the topic, coming from 29 different countries.

Criteria for including studies

The following two criteria for selecting and including studies were used. First, we only selected studies conducted during the last 20 years (i.e., articles published between 2000 and 2019). At this point it should be noted that more than half of the studies included in this meta-analysis (i.e., 58%) were conducted between 2014 and 2019. This shows an increase of the interest of researchers in conducting studies on teacher evaluation during the last 5 years. Secondly, only empirical studies were selected (e.g., validation, experimental, and longitudinal studies). For example, studies testing the validity of an evaluation system (e.g., Borman and Kimball, 2005; Bradshaw, 2002; Derrington and Campbell, 2018) or instruments used in this system (e.g., Amrein-Beardsley et al., 2016; Koedel et al., 2017; Namaghi, 2010), and/or studies investigating the effectiveness of a teacher evaluation reform (e.g., Taut and Sun, 2014; Tuytens and Devos, 2018; Yoo, 2018) were included in this meta-analysis. The studies that met both criteria were, therefore, included in our analysis ($n = 156$). These studies can be found in the web page www.uci.ac.cy/eeerucy/en/.

Main findings: descriptive analysis

In this part, the main results of the descriptive analysis of the data are briefly presented. These results may help readers understand the international dimension of teacher evaluation and come up with conclusions for the main trends of research on teacher evaluation. The nine findings emerged from this meta-analysis are presented below.

First, teacher evaluation seems to be an international concern, considering the fact that the search for studies concluded to sources from 29 different countries. Although 60% of these studies are reported in North America and the United States, we were able to identify a significant percentage of studies (i.e., 10%) which were conducted in developing countries. It should however be acknowledged that papers coming from developing countries seem to criticize the “borrowing” of teacher evaluation policies established in developed countries. This is clearly stated in the papers included in this meta-analysis. Chile is one of the countries where teacher evaluation policy was based on “borrowing” teacher evaluation policies developed in other countries and especially the United States. It is, however, obvious that the research community has doubts on the effectiveness of such cross-national “policy borrowing”, as regards to their impact on improving teacher evaluation practices in the respective countries.

Second, there are several studies ($n = 26$) conducted in various countries ($n = 8$) that put efforts into changing their existing teacher evaluation systems. These studies aim to evaluate the effectiveness of a teacher evaluation reform and are concerned with the purposes of the teacher evaluation reform ($n = 25$), the criteria of the proposed evaluation system ($n = 26$), and the sources of data that the reform has introduced ($n = 20$). In detail, 96% (i.e., 25 out of 26) of the studies report on the purposes of the newly developed teacher evaluation reforms, whereas 35% (9 out of 26) of the studies refer to summative evaluation purposes and 62% of studies refer to both evaluation purposes (i.e., summative and formative). Consequently, none of the studies was found to refer to systems that were developed on the premise of promoting only the formative evaluation purposes. Moreover, all studies that were found to refer to newly developed systems of teacher evaluation are related to specific criteria of evaluation which are frequently in line with more than one model of evaluation. For example, in this set of studies, evaluation criteria were found to be mostly related with the accountability model (96%), the goal and task model (69%) and the working process model (27%). Also, 77% of studies report on the sources of data that are used in each evaluation system. It is important to note that only few ($n = 7$) of the proposed evaluation reforms address meta-evaluation and this seems to be in line with our argument that meta-evaluation mechanisms should be established in most countries (see third section of this chapter).

Third, almost all the studies reviewed here collected data from either the pre-primary (29%, $n = 46$), primary (80%, $n = 125$) and/or the secondary school level (71%, $n = 111$). No more than 10% of the studies ($n = 16$) are concerned with teacher evaluation in higher education. In each phase of schooling, one can notice that the relevant studies were conducted in various countries and not only in developed countries. For example, studies carried out in pre-primary education did not take place only in the United States, but also in countries such as Chile, China, Finland and the United Kingdom. Similar observations can be drawn for studies carried out in primary education. Most of these studies were conducted in the United States (64%), whereas other studies took place in countries such as Chile, Portugal and the Netherlands. Studies conducted in secondary education also appear in several countries including Chile, China and Belgium. These results reveal that teacher evaluation is a topic of interest for researchers in various countries and interest was given to the role of teacher evaluation in each phase of compulsory education (i.e., pre-primary, primary and secondary education). Despite the fact that most studies were conducted in the United States (i.e., pre-primary: 34 out of 46; primary education: 80 out of 125; secondary education 70 out of 111), researchers in other countries show also an interest on conducting studies about the role of teacher evaluation in each phase of schooling. It should, however, be acknowledged, that only recently studies on teacher evaluation in pre-primary education were conducted. More specifically, more than 70% of studies on teacher evaluation in pre-primary education took place during the last 5 years and 13 out of 46 were conducted in 2019. On the other hand, the number of studies on teacher evaluation in higher education, remains low throughout the year range set in this meta-analysis ($n = 16$). To conclude, the fact that almost 30% of the studies in this meta-analysis report findings concerning pre-primary education, shows the importance of establishing valid teacher evaluation mechanisms from the very early stages of schooling.

Fourth, in regard to the research strategy that each study has used one can identify an uneven number of empirical studies across the eleven types that were detected. More specifically, the most frequent research strategies were (a) cross-sectional studies (26%), (b) studies investigating stakeholder perceptions using the interview technique (16%), (c) mixed-research methods (i.e., combination of qualitative and quantitative measures) (15%) and secondary analysis of large scale studies (13%). Notably, only a limited number of studies (6%) were experimental (either randomization studies or quasi-experimental). Moreover, a very small percentage

of studies (8%) were found to be longitudinal. Countries that concentrate the largest number of studies in the meta-analysis, such as the USA, Chile, Belgium and China tend to use a variation of research strategies, but in each of these countries a very small number of studies were found to use experimental and/or longitudinal research strategies. For example, in the case of the United States, only 10 out of 88 studies (11%) were either experimental or longitudinal.

Fifth, almost all studies included in this meta-analysis (91%) were concerned with the criteria of teacher evaluation. It was, therefore, possible to examine the emphasis that was given to each model of TER in generating evaluation criteria. It was found that most studies, and especially those conducted in the United States, paid attention to the *Accountability model* ($n = 111$) and to the *Goal and Task model* ($n = 100$). An interest to criteria deriving from the *Working Process model* can also be observed ($n = 48$). It is, however, important to note that those studies which were concerned with the *Working Process model*, took place mainly in European countries where teachers are usually not held accountable for their students' learning outcomes. In regard to the other models of TER, very few studies were concerned with criteria deriving from the *Continuous Learning model* ($n = 23$) and almost no studies with the *Absence of Problems model* ($n = 2$) and the *Resource Utilization model* ($n = 1$). It is, however, important to note that 64% of studies (i.e., 98 out of 156) acknowledged the importance of criteria deriving from more than one model of TER. The most common combination of criteria observed in these studies had to do with criteria deriving from the *Goal and Task model* and the *Accountability model* and this combination was mainly observed in studies conducted in the United States, Chile, and China where strong teacher accountability systems seem to exist. Moreover, studies using criteria deriving from the *Continuous Learning model* were also concerned with the importance of criteria associated with either the *Working Process* or the *Goal and Task model*. These results reveal that the research community supports the importance of using more than a single model of TER to generate criteria of teacher evaluation, especially since teachers are expected to take on multiple roles and additional responsibilities, such as curriculum developers, action researchers, team leaders and staff development facilitators. Thus, a broader and multidimensional conception of teacher effectiveness that incorporates measures of effectiveness across these different roles, is necessary.

Sixth, in terms of the sources for collecting teacher evaluation data, these are allocated in two broad categories according to: (a) the authority/stakeholder responsible for providing teacher evaluation data and (b) the methods used for collecting the data. As regards to the first category, headteachers (65%), ministries or other educational agencies (38%) and external experts on evaluation (21%) are mostly reported in the studies included in this meta-analysis. It is acknowledged that 63% of the studies included in this meta-analysis, have used more than one stakeholder (i.e., source) to provide teacher evaluation data. In regard to the techniques used to generate data, questionnaires were often used to collect data, either from students (50%) or from headteachers (33%). It is interesting to note that only 13 out of 156 (8%) studies searched for the extent to which other teachers/colleagues can provide teacher evaluation data. On the other hand, the use of teacher self-evaluation as a technique for conducting formative evaluation was examined by more than 20% of studies. Furthermore, a significant percentage of studies (32%) were concerned with the use of value-added methods to evaluate teacher effectiveness. This technique was observed in countries with a strong teacher accountability system. Finally, more than 3 out of 4 studies (78%), were concerned with the use of more than one source of data. This implies that the research community recognizes the importance of using multiple methods for collecting teacher evaluation data and this trend can be observed in studies conducted in various countries around the world.

Seventh, the studies considered in this meta-analysis were classified into four broad categories depending on the purpose that they served. The four categories are: (1) Evaluation of recent reforms related to teacher evaluation, (2) Validation studies, (3) Investigation of new approaches to teacher evaluation and (4) Investigations of the impact of teacher evaluation on student learning outcomes. It is acknowledged here that there are studies that belong to more than one category. More specifically, 26 studies were found to be situated in more than one category. The most frequent combination ($n = 21$) was found to be studies examining the validity of a teacher evaluation system and at the same time investigating the impact of the evaluation on teacher outcomes and student achievement. Studies that were found to be situated in the first category ($n = 26$) aim at evaluating a reform policy in teacher evaluation and most of these studies were concerned with the perceptions of different stakeholders toward the reform policy ($n = 21$) whereas only 4 attempted to identify the extent to which the reform managed to achieve its aims which implies that the objective evaluation model was rarely used by this group of researchers. On the other hand, studies that are placed in the second category ($n = 121$), seek to investigate the validity of different evaluation tools and/or of existing mechanisms and/or frameworks of evaluation. It is important to note that this area of studies forms the largest set of studies. The third area of studies ($n = 16$) includes the studies that aim to examine new evaluation features and approaches either because of their recent appearance or due to an experimental set up that may provide new insights to all related stakeholders of the educational process. The fourth category of studies ($n = 30$) aims to link teacher evaluation with the ultimate aim of the educational process; the improvement of student learning outcomes. This category includes a very limited set of studies and represents the efforts to measure the impact of teacher evaluation on improving student learning outcomes. Researchers in different countries gave more emphasis to the second category of studies rather than others. For instance, in the United States, more than 75% of studies (66 out of 88) belong to the second category which aims to validate specific mechanisms of teacher evaluation systems. It is finally important to note that in each country, only few studies belonged either to the third or fourth category. However, an interest on reform evaluation studies is only observed in those countries where changes in teacher evaluation mechanisms take place. These results reveal some variation in the interest of researchers in different countries on issues related to teacher evaluation which can partly be attributed to contextual factors.

Eighth, the great majority of studies identified in this meta-analysis are not theory driven. Specifically, only 18% of the studies reported the use of any form of theoretical structure as a medium for designing the teacher evaluation system that they are focusing on. The presence of only 18% of studies relying on a specific theory reveals that policy makers need more evidence in order to turn their attention to appropriate directions for improvement of teacher evaluation systems. This argument is also supported, by the fact

that in the validation of evaluation systems category, the use of experimental studies (7%, $n = 8/121$) and longitudinal studies (7%, $n = 9/121$) were also very limited.

Lastly, it is important to note the lack of studies examining the aspect of meta-evaluation in an evaluation system. The absence of a meta-evaluation system is also emphasized in the previous section of this chapter. More specifically, in this meta-analysis, only 14 studies in the general set of studies (9%) were found to report on a meta-evaluation system. This finding can be attributed to the difficulty of the educational systems to change their evaluation mechanisms over time (Fitzpatrick et al., 2004; Scriven, 2009; Stufflebeam, 2011). In addition, this finding shows that the research community has not yet developed enough initiatives to explore the effectiveness of the existing teacher evaluation systems. Taking this into consideration, in the last section of this chapter below, we stress the significance of taking all aspects of a comprehensive teacher evaluation system described in this chapter, into account and we propose actions that could be initiated by the researchers in this field in order to establish links between research on teacher effectiveness and reform policies on teacher evaluation.

Conclusions

Teacher evaluation is considered a vital element for assuring teacher quality and for promoting not only the professional development of teachers but also students' learning outcomes (Seidel and Shavelson, 2007; Stronge, 2006b, 2007). In the first three sections of this chapter a literature review on teacher evaluation has given emphasis to four main aspects that need to be taken into account when developing a comprehensive and effective teacher evaluation system: (1) the evaluation purposes, (2) the evaluation criteria, (3) the sources for collecting relevant data and (4) the meta-evaluation mechanisms. We argued that a teacher evaluation system should be built on both purposes of evaluation (summative and formative), as it is important to give emphasis on both outcome and improvement. However, policy makers should be able to establish different mechanisms of formative and summative teacher evaluation, since the two purposes of evaluation are based on different value assumptions. Consequently, different criteria can be used for conducting each type of teacher evaluation. Thus, in the second section of the chapter we explicitly refer to seven TER models in developing criteria for teacher evaluation. There is a need for using TER to generate separate criteria for the two purposes of evaluation, since a certain group of criteria may be more appropriate for formative evaluation and a different (or at least different in some aspects) set of criteria may be more suitable for summative evaluation. Hence, we suggest the incorporation of multiple models of teacher evaluation, as different perspectives may generate a detailed outline of the teaching and learning situation/process. For example, focusing solely on the goal and task model, we may have evidence of effective and ineffective teachers, but there is no guidance of how teachers could improve their practices. On the other hand, focusing only on the working process model evaluators may distinguish what features of teaching could be improved by consequently setting up appropriate teacher professional development programs, but questions regarding to what standards are teachers evaluated and how are those standards related to students' progress, should also be raised. In addition, for each criterion that is judged important for developing a teacher evaluation system, appropriate sources of data upon which the measurement of teacher effectiveness could be based should be identified. Accordingly, combining multiple sources for measuring the quality of teaching, may reveal unseen aspects of teacher effectiveness and lead to specific actions for improvement. Moreover, specific instruments need to be developed to collect data from each source and their construct validity should be systematically investigated.

In the third section of the chapter the fourth aspect of a comprehensive teacher evaluation system is emphasized, since when designing such an evaluation system, monitoring is the key factor of the process for assuring quality of the desired outcomes. Meta-evaluation processes are reported only rarely in the related literature either due to the resistance of stakeholders participating in the design of a teacher evaluation process and providing accountability transparency or because the research community has not yet developed enough initiatives to explore the effectiveness of the existing teacher evaluation systems. We, thus, emphasize through the results of the descriptive analysis of the data of the meta-analysis reported in the fourth section of this chapter, that it is important for countries to design their own teacher evaluation system based on empirical data as well as on a valid theoretical framework. The existence of a theoretical framework is valuable when designing a teacher evaluation system since this framework will guide the decisions of policymakers as regards to the criteria that can be used for evaluating teachers as well as the sources of data that will be used.

The results of the meta-analysis reported in this chapter, showed an increasing interest of the research community on teacher evaluation, during the last 5 years. This is an important finding especially since an interest on teacher evaluation research was also observed in various developing countries. More specifically, some countries were found to introduce policies and practices that are currently used in other developed countries. This "borrowing policy" is found to be ineffective by the research community. There are matters concerning the relevance of the "borrowed policies" with each country's context and framework relating to teacher evaluation. It is also important to note that policymakers should not simply introduce reforms that are in line with the perceptions of various stakeholders, but only promote those reforms which were found to promote quality of teaching and student learning outcomes. Finally, most studies conducted in the field of teacher evaluation concern the investigation of the effectiveness of existing evaluation systems, rather than suggesting new approaches to teacher evaluation and testing their impact on improving teacher effectiveness and student learning outcomes. The research community is encouraged to search for effective approaches to teacher evaluation that are based on theories. Findings of these studies are more likely to be used by policymakers who are expected to develop comprehensive teacher evaluation reforms, which will not only be valid and reliable but should also have an effect on promoting quality and equity in education. We, therefore, argue for the use of a theory-driven and evidence-based approach to

generate teacher evaluation systems in each country which will cover the purposes of the evaluation, the evaluation criteria, the sources for collecting relevant data, as well as the meta-evaluation mechanisms for supporting the continuous improvement of teacher evaluation systems.

References

- Amrein-Beardsley, A., 2014. *Rethinking Value-Added Models in Education: Critical Perspectives on Tests and Assessment-Based Accountability*. Routledge, London and New York.
- Amrein-Beardsley, A., Polasky, S., Holloway-Libell, J., 2016. Validating “value added” in the primary grades: one district’s attempts to increase fairness and inclusivity in its teacher evaluation system. *Educ. Assess. Eval. Account.* 28 (2), 139–159.
- Anast-May, L., Penick, D., Schroyer, R., Howell, A., 2011. Teacher conferencing and feedback: necessary but missing! *Int. J. Educ. Leadership Preparation* 6 (2).
- Bacher-Hicks, A., Chin, M.J., Kane, T.J., Staiger, D.O., 2019. An experimental evaluation of three teacher quality measures: value-added, classroom observations, and student surveys. *Econ. Educ. Rev.* 73, 101919.
- Barile, J.P., Donohue, D.K., Anthony, E.R., Baker, A.M., Weaver, S.R., Henrich, C.C., 2012. Teacher–student relationship climate and school outcomes: implications for educational policy initiatives. *J. Youth Adolesc.* 41 (3), 256–267.
- Beerens, D.R., 2000. *Evaluating Teachers for Professional Growth: Creating a Culture of Motivation and Learning*. Corwin Press, Thousand Oaks, CA.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assess Educ. Princ. Pol. Pract.* 5 (1), 7–74.
- Borman, G.D., Kimball, S.M., 2005. Teacher quality and educational equality: do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elem. Sch. J.* 106 (1), 3–20.
- Bradshaw, L.K., 2002. Local district implementation of state mandated teacher Evaluation policies and procedures: the North Carolina case. *J. Person. Eval. Educ.* 16 (2), 113–127.
- Brennan, N., 2001. Reporting intellectual capital in annual reports: evidence from Ireland. *Account Audit. Account. J.* 14 (4), 423–436.
- Brophy, J.Y., Good, T.L., 1986. Teacher behaviour and student achievement. In: Wittrock, M.C. (Ed.), *Handbook of Research on Teaching*, third ed. Macmillan, New York, pp. 328–375.
- Carle, A.C., 2009. Fitting multilevel models in complex survey data with design weights: recommendations. *BMC Med. Res. Methodol.* 9 (1), 1–13.
- Chapman, C., Muijs, D., Reynolds, D., Sammons, P., Teddlie, C. (Eds.), 2016. *The Routledge International Handbook of Educational Effectiveness and Improvement*. Routledge, London and New York.
- Cheng, Y.C., Tsui, K.T., 1999. Multimodels of teacher effectiveness: implications for research. *J. Educ. Res.* 92 (3), 141–150.
- Coleman, J.S., Campbell, E.Q., Hobson, C.F., Mcpartland, J., Mood, A.M., Weinfeld, F.D., York, R.L., 1996. *Equality of Educational Opportunity*. US Government Printing Office, Washington, DC.
- Collins, C., 2014. Houston, we have a problem: teachers find no value in the SAS education value-added assessment system (EVAAS®). *Educ. Pol. Anal. Arch.* 22 (98).
- Cooksy, L.J., Caracelli, V.J., 2009. Meta-evaluation in practice. *J. Multidiscip. Eval.* 6 (11), 1–15.
- Creemers, B.P.M., Kyriakides, L., 2008. *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. Routledge, London and New York.
- Creemers, B.P.M., Kyriakides, L., 2012. *Improving Quality in Education: Dynamic Approaches to School Improvement*. Routledge, London and New York.
- Creemers, B.P.M., Kyriakides, L., Antoniou, P., 2013. *Teacher Professional Development for Improving Quality of Teaching*. Springer, Dordrecht, the Netherlands.
- Creemers, B.P.M., Kyriakides, L., Sammons, P., 2010. *Methodological Advances in Educational Effectiveness Research*. Routledge, London and New York.
- Creemers, B.P.M., Reezigt, G.J., 1996. School level conditions affecting the effectiveness of instruction. *Sch. Effect. Sch. Improv.* 7 (3), 197–228.
- Cronbach, L.J., 1990. *Essentials of Psychological Testing*, third ed. Harper & Row, New York.
- Cronbach, L.J., Gleser, J.C., Nanda, H., Rajaratnam, N., 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability of Scores and Profiles*. John Wiley, New York.
- Danielson, C., McGreal, T.L., 2000. *Teacher Evaluation to Enhance Professional Practice*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., Van Petegem, P., 2013. How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teach. Teach. Educ.* 36, 1–11.
- Derrington, M.L., Campbell, J.W., 2018. High-stakes teacher evaluation policy: US principals’ perspectives and variations in practice. *Teach. Teach.* 24 (3), 246–262.
- Donaldson, M.L., 2012. *Teachers’ Perspectives on Evaluation Reform*. Center for American Progress.
- Donaldson, M.L., 2016. Teacher evaluation reform: focus, feedback, and fear. *Educ. Leader* 73 (8), 72–76.
- Doyle, W., 1986. Classroom organization and management. In: Wittrock, M.C. (Ed.), *Handbook of Research on Teaching*, third ed. Macmillan, New York, pp. 392–431.
- Elett, C.D., Wren, C.Y., Callender, K.E., Loup, K.S., Liu, X., 1996. Looking backwards with the personnel evaluation standards: an analysis of the development and implementation of a statewide teacher assessment program. *Stud. Educ. Eval.* 22 (1), 79–113.
- Elstad, E., Lejonberg, E., Christophersen, K.A., 2017. Student evaluation of high-school teaching: which factors are associated with teachers’ perception of the usefulness of being evaluated? *J. Educ. Res.* 9 (1), 99–117.
- Evertson, C.M., Anderson, C.W., Anderson, L.M., Brophy, J.E., 1980. Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *Am. Educ. Res. J.* 17 (1), 43–60.
- Fessler, R., Ungaretti, A., 1994. Expanding Opportunities for Teacher Leadership. In: *Teachers as Leaders: Perspectives on the Professional Development of Teachers*, pp. 211–222.
- Firestone, W.A., 1989. Using reform: conceptualizing district initiative. *Educ. Eval. Pol. Anal.* 11 (2), 151–164.
- Fitzpatrick, J.L., Sanders, J.R., Worthen, B.R., 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines*. Pearson Education, London.
- Flores, M.A., Derrington, M.L., 2018. Improving teacher evaluation: key issues for appraisers in a globalised era. *Teach. Teach.* 24 (3), 203–208.
- Galton, M., 1987. An ORACLE chronicle: a decade of classroom research. *Teach. Teach. Educ.* 3 (4), 299–313.
- Glickman, C.D., Gordon, S.P., Ross-Gordon, J.M., 2014. *Supervision and Instructional Leadership: A Developmental Approach*. Pearson, New York.
- Goe, L., Biggers, K., Croft, A., 2012. *Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning*. National Comprehensive Center for Teacher Quality, Washington, DC.
- Goldstein, D., 2014. *The Teacher Wars: A History of America’s Most Embattled Profession*. Doubleday, New York.
- Gordon, R.J., Kane, T.J., Staiger, D., 2006. *Identifying Effective Teachers Using Performance on the Job*. Brookings Institution, Washington.
- The Joint Committee on Standards for Educational Evaluation. In: Howard, B., Gullickson, A. (Eds.), 2009. *Personnel Evaluation Standards: How to Assess Systems for Evaluating Educators*. Corwin Press, Thousand Oaks, CA.
- Hallinger, P., Heck, R.H., Murphy, J., 2014. Teacher evaluation and school improvement: an analysis of the evidence. *Educ. Assess. Eval. Account.* 26 (1), 5–28.
- Hargreaves, A., Fullan, M.G., 1992. *Understanding Teacher Development*. Teachers College Press, New York.
- Harlen, W., James, M., 1997. Assessment and learning: differences and relationships between formative and summative assessment. *Assess Educ. Princ. Pol. Pract.* 4 (3), 365–379.
- Hattie, J., 2009. The black box of tertiary assessment: an impending revolution. In: *Tertiary Assessment and Higher Education Student Outcomes: Policy, Practice and Research*, 259, p. 275.

- Hill, H.C., Charalambous, C.Y., Blazar, D., McGinn, D., Kraft, M.A., Beisiegel, M., Humez, A., Litke, E., Lynch, K., 2012. Validating arguments for observational instruments: attending to multiple sources of variation. *Educ. Assess.* 17 (2–3), 88–106.
- Huber, S.G., Skedsmo, G., 2016. Teacher evaluation—accountability and improving teaching practices. *Educ. Assess. Eval. Account.* 28 (2), 105–109.
- Johnson, B.L., 1997. An organizational analysis of multiple perspectives of effective teaching: implications for teacher evaluation. *J. Person. Eval. Educ.* 11 (1), 69–87.
- Kimball, S.M., White, B., Milanowski, A.T., Borman, G., 2004. Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody J. Educ.* 79 (4), 54–78.
- Koedel, C., Li, J., Springer, M.G., Tan, L., 2017. The impact of performance ratings on job satisfaction for public school teachers. *Am. Educ. Res. J.* 54 (2), 241–278.
- Kyriakides, L., Anthimou, M., Panayiotou, A., 2020. Searching for the impact of teacher behavior on promoting students' cognitive and metacognitive skills. *Stud. Educ. Eval.* 64, 100810.
- Kyriakides, L., Campbell, R.J., 2003. Teacher evaluation in Cyprus: some conceptual and methodological issues arising from teacher and school effectiveness research. *J. Person. Eval. Educ.* 17 (1), 21–40.
- Kyriakides, L., Campbell, R.J., Christofidou, E., 2002. Generating criteria for measuring teacher effectiveness through a self-evaluation approach: a complementary way of measuring teacher effectiveness. *Sch. Effect. Sch. Improv.* 13 (3), 291–325.
- Kyriakides, L., Creemers, B.P.M., Charalambous, E., 2018. *Equity and Quality Dimensions in Educational Effectiveness*. Springer, Dordrecht, the Netherlands.
- Kyriakides, L., Demetriou, D., 2007. Introducing a teacher evaluation system based on teacher effectiveness research: an investigation of stakeholders' perceptions. *J. Person. Eval. Educ.* 20 (1–2), 43–64.
- Kyriakides, L., Demetriou, D., Charalambous, C., 2006. Generating criteria for evaluating teachers through teacher effectiveness research. *Educ. Res.* 48 (1), 1–20.
- Kyriakides, L., Creemers, B.P.M., Panayiotou, A., Charalambous, E., 2021. *Quality and Equity in Education: Revisiting Theory and Research on Educational Effectiveness and Improvement*. Routledge, London and New York.
- Kyriakides, L., Creemers, B.P.M., Panayiotou, A., Vanlaar, G., Pfeifer, M., Cankar, G., McMahon, L., 2014. Using student ratings to measure quality of teaching in six European countries. *Eur. J. Teach. Educ.* 37 (2), 125–143.
- Lasagabaster, D., Sierra, J.M., 2011. Classroom observation: desirable conditions established by teachers. *Eur. J. Teach. Educ.* 34 (4), 449–463.
- Liu, C.W., 2011. The Implementation of Teacher Evaluation for Professional Development in Primary Education in Taiwan. Universal-Publishers, Boca Raton, Florida.
- Looney, J., 2011. Developing high-quality teachers: teacher evaluation for improvement. *Eur. J. Educ.* 46 (4), 440–455.
- Marcoulides, G.A., 1989. The application of generalizability analysis to observational studies. *Qual. Quantity* 23 (2), 115–127.
- Marsh, H.W., 1987. The big-fish-little-pond effect on academic self-concept. *J. Educ. Psychol.* 79 (3), 280.
- McGreal, C.E., 1988. In great-grandma's nursery. *Infant Ment. Health J.* 9 (4), 305–318.
- Mortimore, P., 2001. Globalisation, effectiveness and improvement. *Sch. Eff. Sch. Improv.* 12 (2), 229–249.
- Muijs, D., Reynolds, D., 2011. *Effective Teaching: Evidence and Practice*. Paul Chapman, London.
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., Earl, L., 2014. State of the art—teacher effectiveness and professional learning. *Sch. Effect. Sch. Improv.* 25 (2), 231–256.
- Murphy, J., 1995. Changing role of the teacher. *Educating Teachers for Leadership and Change* 311–323.
- Namaghi, S.A.O., 2010. A data-driven conceptualization of teacher evaluation. *Qual. Rep.* 15 (6), 1504–1522.
- Nye, B., Konstantopoulos, S., Hedges, L.V., 2004. How large are teacher effects? *Educ. Eval. Pol. Anal.* 26 (3), 237–257.
- Papay, J., 2012. Refocusing the debate: assessing the purposes and tools of teacher evaluation. *Harv. Educ. Rev.* 82 (1), 123–141.
- Pauffer, N.A., 2018. Declining morale, diminishing autonomy, and decreasing value: principal reflections on a high-stakes teacher evaluation system. *Int. J. Educ. Policy Leadersh.* 13 (8).
- Patterson, J.L., Purkey, S.C., Parker, J.V., Association for Supervision and Curriculum Development, 1986. *Productive School Systems for a Nonrational World*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Peterson, K.D., 2000. *Teacher Evaluation: A Comprehensive Guide to New Directions and Practices*. Corwin Press, Thousand Oaks, CA.
- Peterson, K.D., Wahlquist, C., Bone, K., 2000. Student surveys for school teacher evaluation. *J. Person. Eval. Educ.* 14 (2), 135–153.
- Pfeffer, J., 1992. *Managing With Power: Politics and Influence in Organizations*. Harvard Business Press, Boston.
- Reynolds, D., Creemers, B.P.M., Stringfield, S., Teddlie, C., Schaffer, G., 2002. *World Class Schools. International Perspectives on School Effectiveness*. Routledge Falmer, New York and London.
- Rigby, J., 2015. Principals' sensemaking and enactment of teacher evaluation. *J. Educ. Adm.* 53 (3), 374–392.
- Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools, and academic achievement. *Econometrica* 73 (2), 417–458.
- Robertson-Kraft, C., Zhang, R.S., 2018. Keeping great teachers: a case study on the impact and implementation of a pilot teacher evaluation system. *Educ. Pol.* 32 (3), 363–394.
- Rockoff, J.E., 2004. The impact of individual teachers on student achievement: evidence from panel data. *Am. Econ. Rev.* 94 (2), 247–252.
- Sanders, W.L., Wright, S.P., Horn, S.P., 1997. Teacher and classroom context effects on student achievement: implications for teacher evaluation. *J. Person. Eval. Educ.* 11 (1), 57–67.
- Santelices, M.V., Taut, S., 2011. Convergent validity evidence regarding the validity of the Chilean standards-based teacher evaluation system. *Assess. Educ. Princ. Pol. Pract.* 18 (1), 73–93.
- Scheerens, J., 2013. The use of theory in school effectiveness research revisited. *Sch. Effect. Sch. Improv.* 24 (1), 1–38.
- Scriven, M., 1969. *An Introduction to Meta-Evaluation*. Computer Science.
- Scriven, M., 2009. Meta-evaluation revisited. *Journal of Multidisciplinary Evaluation* 6 (11), 3–8.
- Seidel, T., Shavelson, R.J., 2007. Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Rev. Educ. Res.* 77 (4), 454–499.
- Shavelson, R.J., Webb, N.M., 2006. Generalizability theory. In: Green, J.L., Camilli, G., Elmore, P.B. (Eds.), *Handbook of Complementary Methods in Education Research*. Erlbaum, Mahwah, NJ, pp. 309–322.
- Steinberg, M.P., Donaldson, M.L., 2016. The new educational accountability: understanding the landscape of teacher evaluation in the post-NCLB era. *Educ. Finance Policy* 11 (3), 340–359.
- Stronge, J.H., 1995. Balancing individual and institutional goals in educational personnel evaluation: a conceptual framework. *Stud. Educ. Eval.* 21 (2), 131–151.
- Stronge, J.H., 2006a. *Evaluating Teaching: A Guide to Current Thinking and Best Practice*. Corwin Press, Thousand Oaks, CA.
- Stronge, J.H., 2006b. Teacher evaluation and school improvement: improving the educational landscape. In: *Evaluating Teaching: A Guide to Current Thinking and Best Practice*, vol. 2, pp. 1–23.
- Stronge, J.H., Tucker, P.D., 2000. *Teacher Evaluation and Student Achievement*. Student Assessment Series. National Education Association, Washington, DC.
- Stufflebeam, D.L., 2011. Meta-evaluation. *J. Multidiscip. Eval.* 7 (15), 99–158.
- Tarhan, H., Karaman, A., Lauri, K., Aerial, J.A., 2019. Understanding teacher evaluation in Finland: a professional development framework. *Aust. J. Teach. Educ.* 44 (4), 33–50.
- Taut, S., Sun, Y., 2014. The development and implementation of a national, standards-based, multi-method teacher performance assessment system in Chile. *Educ. Policy Anal. Archiv.* 22, 1–30.
- Taylor, E.S., Tyler, J.H., 2012. The effect of evaluation on teacher performance. *Am. Econ. Rev.* 102 (7), 3628–3651.
- Tuma, A.P., Hamilton, L.S., Tsai, T., 2018. *A Nationwide Look at Teacher Perceptions of Feedback and Evaluation Systems*. RAND Corporation, Santa Monica, CA.
- Tuytens, M., Devos, G., 2018. Teacher evaluation policy as perceived by school principals: the case of Flanders (Belgium). *Teach. Teach.* 24 (3), 209–222.

- Vekeman, E., Devos, G., Tuytens, M., 2015. The influence of teachers' expectations on principals' implementation of a new teacher evaluation policy in Flemish secondary education. *Educ. Assess. Eval. Account.* 27 (2), 129–151.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., Morgan, K., 2009. The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness. New Teacher Project.
- Yoo, J., 2018. Evaluating the new teacher evaluation system in South Korea: case studies of successful implementation, adaptation, and transformation of mandated policy. *Pol. Futures Educ. Internet* 16 (3), 277–290.
- Zhang, X.F., Ng, H.M., 2011. A case study of teacher appraisal in Shanghai, China: in relation to teacher professional development. *Asia Pac. Educ. Rev.* 12 (4), 569–580.

Leading school-based assessment for educational improvement: rethinking accountability

Gavin T.L. Brown^{a,b}, ^aFaculty of Education & Social Work, The University of Auckland, Auckland, New Zealand; and ^bInstitutionen för Tillämpad Utbildningsvetenskap [Department of Applied Educational Science], Umeå Universitet, Umeå, Sweden

© 2023 Elsevier Ltd. All rights reserved.

Constraints on accountability assessment	299
Consequences for assessment	299
Assessment validity	300
Leading for learning	301
New Zealand research	302
Conclusion	303
Acknowledgment	303
References	303

Globally, there is increasing pressure from politicians, public policy, and parents to know (1) that students in schools are learning what is expected and (2) that schools (i.e., leaders and teachers) are adding value to student learning (Buckendahl, 2016). The pressure comes from a valid concern about the importance of academic outcomes for student life chances (i.e., educationally successful students get better jobs, earn more, and have more fulfilling lives) and for societal development (i.e., our country and its economy will be better if more and more students are well educated). While some experts may reject the notion that schools should be held to account for their use of funding or practices (Biesta, 2010), that is not how most members of society see good educational practice. A simple view simply expects that teachers cause students to do better on valued and presumably accurate measures of learning (Harris and Brown, 2016).

Accountability pressure also arises from the unpalatable realization that not every teacher is as good as another (McNaughton and Lai, 2009) and that it is difficult to improve teachers (Richardson and Placier, 2001). For example, studies in New Zealand have shown that teachers make curriculum decisions independent of assessment information that would show what the student really needs (Parr and Timperley, 2008). Teachers' overall judgments of student capability in New Zealand have been shown to be biased against student standardized test performance scores for all students except white girls (Meissel et al., 2017). Hence, the quality and effectiveness of teachers is not universal.

Educational assessments are frequently used to demonstrate that individual students have fulfilled our expectations (e.g., tests and examinations for entry or qualifications) and test scores from surveys of school performance are used to identify schools and teachers who are associated with students whose performance meets societal expectations (Linn, 2000). The trend to use student tests (most commonly of reading comprehension and mathematics) to evaluate the quality of teachers and schools has become especially prevalent in the United States of America and England (Lingard and Lewis, 2016). In other jurisdictions where public examinations are commonplace (e.g., Asia, Africa, the Middle East; Brown, 2017; Gebril, 2016), the prestige of schools and teachers often depends on the success rate of their students on such high-stakes assessments. Hence, judging teachers and schools by the test performance of their students is both widespread, easy to achieve, and very influential on teachers and school leaders. After all, no school leader wants to be associated with what may be seen as "failure" in the face of its constituent communities (Ngan et al., 2010).

The problem with this pressure is that, teachers feel compelled to deliver exactly what those in authority want; that is, higher test scores. This is a natural human response to accountability pressure; subordinates deliver to those in authority whatever they consider it is that superiors want—the first consequence of accountability (Lerner and Tetlock, 1999). Putting aside whether or not teachers are doing a substandard job, we need to consider whether it is possible to raise test scores without any real learning taking place. Are there ways for teachers to meet the expectation that student test scores will go up, despite students not achieving the learning goals expected?

Unfortunately, the answer is yes. Scores can be increased by a number of means (Cannell, 1989):

- teaching students directly what will be on the test,
- using old tests that have invalid comparison norms,
- getting an early copy of the test and ensuring items are taught prior to the test,
- encouraging low performing students to NOT come to school on testing days,
- giving students hints during testing,
- correcting student test responses before submission to a central agency, and

- a plethora of such dubious acts.

In 2015, teachers and school leaders were jailed in Atlanta, USA for falsifying student performances on school accountability tests.¹ My sense is that in light of the unfair negative consequences (e.g., my school has many children whose parents are poor, uneducated; my school does not have resources; my school is in a location where it is difficult to attract highly skilled teachers; labeling kids as failures is wrong; or using standardized tests to determine a child's worth is wrong) for low scores, "cheating" to raise test scores may be a highly rational, albeit unprofessional, response. It seems highly likely under a regime of accountability, which ignores the many complicating factors that limit student test performance, resistance and subversion of the regime is likely and possibly defensible.

Constraints on accountability assessment

Leading with assessment data poses substantial challenges for teachers. Unfortunately, in many societies school leaders face a context in which they are driven to respond positively to accountability demands, that seem disconnected from circumstances outside the control of schools, leaders, and teachers. As demonstrated powerfully over and over, factors in the home resources of any child account for a large proportion (usually estimated at 50%) of school achievement (Coleman, 1966; Hattie, 2009). Hence, the imposition of negative and strong consequences for "poor" performance that does not take into account input factors outside the control of educators is inherently unfair (Nichols and Harris, 2016). Nonetheless, school leaders have to face those accountability demands and lead teachers to instructional practices that fulfill adequately those demands (Daniëls et al., 2019). In my view, two important conditions must be present if school leaders or teachers are to use address accountability demands transformationally.

Consequences for assessment

It is relatively clear that school systems have the choice to control the level of consequence they attach to school-based assessments. This decision will be in part determined by how policy makers conceive of the role education professionals play. If they view teachers as the problem, they will probably attempt to put into place teacher-proof curricula or impose a regime of consequences for low test scores (e.g., firing, redeployment, salary impact, etc.; Nichols and Harris, 2016). If they view teachers as professionals through whom policy is implemented and manifested, then a substantially more formative approach might be taken to accountability tests (Hattie et al., 2003). In a formative approach, consequences for low scores focus on support and resources for changing the factors that generated the results. Whatever the mindset for accountability testing, policy makers need to take into account the psychology of the agents through whom the policy is deployed.

The willing compliance and cooperation of teachers and school leaders is needed for any policy, not just assessment. This means that the beliefs and attitudes of school teachers matter to how school-based assessment reforms are implemented. There is strong evidence that teacher beliefs contribute to how teachers understand their work contexts, what they are supposed to do, and how they should do it (Fives and Buehl, 2012). Clearly, the meaning of assessment to teachers matters to how school-based assessment can be affected. Brown (2008) has argued that there are four major purposes of assessment (i.e., improved teaching & learning; evaluation of students; evaluation of schools & teachers; negative purposes). Evidence has accumulated that teacher beliefs about assessment depend in part upon the policy and cultural constraints in which teachers work (Barnes et al., 2015; Bonner, 2016; Fulmer et al., 2015).

Research in many countries has found that the relationship of improvement (generally endorsed by all) has an interesting tension with evaluation or accountability purposes (especially school evaluation). For example, school evaluation was weakly correlated with improvement in New Zealand (Brown, 2004, 2011), Queensland (Brown et al., 2011b), and Cyprus (Brown and Michaelides, 2011) all of which had low-stakes assessment environments in which schools and teachers were not punished for low test scores. It's interesting to note that in a New Zealand study (Brown, 2009a), primary school teachers associated student accountability with surface (memorization) learning, whereas school accountability was associated with deep (transformational) learning. This was understood to mean that the teachers thought good schools created deep learning which was not easily captured by standardized tests.

In contrast, improvement and accountability beliefs are strongly correlated in Hong Kong (Brown et al., 2009), China (Brown et al., 2011a), Egypt (Gebril and Brown, 2014), and India (Brown et al., 2015), all of which are high-stakes public examination environments, in which school and teacher quality are inferred from student test scores. Hence, understanding what accountability assessment does to schools and teachers depends on knowing what consequences or purposes are attached to the tests and their scores.

Contextual effects exist within societies where the age and stage of schooling changes. High school teachers tend to generally agree that student evaluation is an actual purpose for assessment by statistically significant margins over primary school teachers. This probably happens because qualifications assessments are a natural, normal, and expected part of schooling for adolescent children (Brown, 2011; Brown et al., 2011b). This indicates teachers are not unrealistic about the importance of ensuring students meet expectations for young people prior to entering work or higher education.

¹https://en.wikipedia.org/wiki/Atlanta_Public_Schools_cheating_scandal.

In societies where examinations are used to select students for limited resources (e.g., scholarship to university), such assessments are seen as both a way to make students accountable and simultaneously as a way to improve. In addition to improved academic performance through assessment, in China, it seems teachers associate greater academic achievement with greater acquisition of scholarly virtues as a better person, giving even greater impetus to the belief that assessment leads to improvement (Brown et al., 2009, 2011; Brown et al., 2011a). The downside to this perspective is simple; the outcomes of teaching depend not on the school leadership or teacher quality, but rather on the student. The practices of instruction do not need to change if student effort or virtue are responsible for outcomes. Hence, this perspective can lend itself to defending teachers from criticism and/or improvement.

Commitment to improvement depends upon both policy and leadership. National surveys carried out among New Zealand primary and secondary teachers under conditions of low-stakes school accountability mechanisms (i.e., tests were voluntary and inspections were triennial), teachers generally endorsed the notion that assessment was for improvement and not much for school evaluation (Brown, 2004, 2011). Subsequently, the Ministry of Education supported schools in adopting a schooling improvement initiative in which school leaders were to actively use school-wide achievement data to guide changes in curriculum and teaching. However, a survey of Auckland teachers in those schools indicated that assessment was much more irrelevant and much more for school accountability, than for improvement, than the previous national surveys (Brown and Harris, 2009). It seemed, when talking with teachers, that some were persuaded that educational assessment is about evaluating schools, rather than improving teaching and learning. If assessment was something for school leadership then it became legitimate to ignore or reject the assessment process.

Thus, making it safe to discover “bad” news (i.e., low scores) has to be real, otherwise teachers will treat assessments as a game in which the goal is to maximize scores. The environment in which assessment is deployed impacts teacher understanding of assessment practices; Saldaña and Shepard (2023) provide multiple examples of the impact of local context upon school use of test scores. Teacher beliefs about what assessment is and what it is for contribute to deep transformational improvement or surface compliance. Reducing external consequences in favor of internally-monitored professionalism is a powerful strategy to take advantage of teachers’ commitment to educational improvement. Nonetheless, as Bonner (2016) demonstrates the tension between meeting accountability demands and improvement goals places a substantial burden on educators. Lillejord (2023) provides further insights into policy frameworks that constructively support teachers in meeting this goal of improvement.

Assessment validity

In order to contribute to either or both formative and summative decision-making, assessments are expected to be both valid samples of valued domains and provide reliable, consistent, and truthful scores about performance (Messick, 1989). Accurate scoring of performance on a sample of items about competencies or capabilities is a feature of standardized, well-prepared tests. This contrasts considerably to the inconsistency in teacher judgements on performance assessments (e.g., essays; Brown, 2009b, 2010). Notwithstanding concerns about the validity of tests, they are relatively cheap, quick, and easy to develop and tend to provide the kinds of scores which easily satisfy politicians, media, and parents (Linn, 2000).

Nonetheless, a considerable problem with tests is the limited information provided through total or rank order scores (Brown and Hattie, 2012; Hattie and Brown, 2010). Standardized tests or examinations tend to produce information about the total score (i.e., percent correct score or letter grade) and information about how well students have done relative to normative samples (i.e., stanine, percentile, z-score, etc.). Total scores are not overly useful in schools because they do not identify which aspects of the domain a student is weak or strong in; neither do they identify which students have which learning needs in each important area. Rank order scores likewise have limited usefulness, since most teachers already know which students are at the top or bottom of their own classes (Elley and Livingstone, 1972). Teachers tend not to know how good their class is relative to the rest of the country population; this means a weak student in my class is seen as a weak student in the jurisdiction, even if the class is in a top performing school. While it is useful to know if a score differs a little or a lot compared to the norm, this does not tell teachers how to use that score to make a difference. Thus, tests that only provide total and rank scores fail to generate the information school leaders and teachers need to make systematic changes to instruction so that students learn more. Another trap with these kinds of statistical data lies in the relative inability of teachers to make proper sense of test scores, the illusion that such test scores are objective and correct, and the loss of details about individual learners in aggregation of scores into means (Fjørtoft and Lai, 2020).

To address this problem, rather than dispense with standardized tests altogether, in New Zealand, a standardized, school-based test system was developed to provide high-quality diagnostic and formative information to both teachers and school leaders (Brown, 2013; Hattie and Brown, 2008; Hattie et al., 2003, 2004). The Assessment Tools for Teaching and Learning (asTTle) computer-assisted software system (now a fully online web-based system) was developed to assist the teacher with two onerous tasks; that is, creating a test and analyzing student performance. The asTTle software uses a sophisticated linear programming heuristic to create both paper and computer-adaptive online test options using a mixture of open- and closed-response items to fit the teacher curriculum and difficulty requirements. Teachers determine the validity of the test, thus ensuring it fits their students and their teaching. The system software calculates student achievement, strengths, and weaknesses using an item response theory formula so that regardless of the items on a test, student scores are on a common scale. This ensures interpretation against nationally representative norms, curriculum levels, and curriculum achievement objectives.

Like other tests, the system provides nationally comparative rank-order scores and standardized total scores. Unlike other systems, it also reports the strengths and weaknesses of individuals and groups right down to learning objectives within the various strands of the curriculum for reading comprehension, writing, and mathematics (Brown et al., 2018). Additionally, it points

teachers to high-quality curriculum resources to support better teaching of all students regardless of whether they are behind or ahead of expectations. Further, since the system is designed to be formative, the stakes or consequences are generally low because schools see and respond to the information before external authorities. This allows improvement to take place before external accountability inspections take place and helps to show the competence of educational professionals. This model of assessment has the potential to enable school leadership for improved outcomes as the priority for educational assessment.

Leading for learning

It is thus necessary for school leaders to resist the temptation to inflate test scores through devious means. It also means that the environment of consequence and test resources has to be in place for leaders to turn assessment into an educational practice. Real, transformational learning is not a simple input-output factory-like process; rather it is a complex set of interactions between students and insightful, enthusiastic, skilled educators over a long period of time. In essence, instructional leadership around the challenges of assessment for accountability require that:

Leaders who take into account and test the theories of those who are responsible for improvements are more likely to better understand the problem that they are trying to solve, better able to develop solutions that are accepted, and more likely to effect changes in the practices that they wish to impact. [Wilson and Lai \(2019, p. 458\)](#).

The school leader is responsible for leading instruction not just ensuring administrative compliance with accountability pressures. The real goal of school leadership is improved learning and instruction. To achieve this, school leaders need to adopt instructional leadership practices that research has shown lead to greater learning outcomes for students. [Daniëls et al. \(2019\)](#) demonstrate that instructional leadership frameworks now sit comfortably within the more recent Leadership for Learning framework, which focuses on learning, teaching, curricula, and instruction as well as ensuring all the other dimensions of schooling (e.g. administration, organization, finance) contribute to improved student learning. Nonetheless, the impact of school leadership on student learning outcomes is largely indirect as it depends on the subordinate middle managers and teachers implementing the learning vision of the leader.

In her synthesis of leadership practices that impact student learning outcomes, [Robinson \(2007\)](#) has shown that the most powerful effects on student learning are associated with leaders who: (1) promote and participate in teacher learning; (2) plan, coordinate, and evaluate teaching; (3) establish goals and expectations; (4) deal with strategic resourcing; and (5) provide an orderly and supportive environment. This list makes it clear that administrative competence is necessary but not sufficient. The meta-analysis showed that the largest gains occurred when school leaders were actively involved with their teachers in collecting and evaluating assessment data and in designing and monitoring pedagogical responses to the assessment results. This result is consistent with Hattie's meta-synthesis (2009) of effective practices that lead to greater student learning; teachers who use assessment data to inform them about their impact on student learning generate large educational gains ($d = 0.90$).

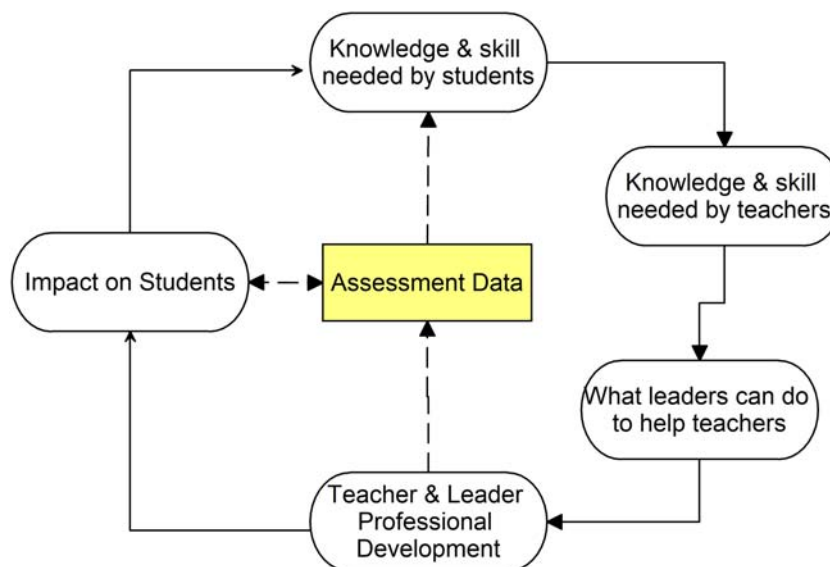


Fig. 1 Basic professional learning inquiry and knowledge building cycles. Adapted from [Timperley \(2011c\)](#).

Based on these findings a comprehensive, general model (Timperley, 2011b) of five sequential steps has been developed (Fig. 1). This has been modified for this chapter to highlight the role assessment data play in the process. Assessment data, alongside curricular objectives, provide insights into the knowledge and skills students need so as to be successful in learning in any domain. To exploit these data leaders and teachers need to jointly review the data to identify strengths and deficiencies in the student performance (Parr and Timperley, 2008; Timperley, 2005). One of the side effects of this approach is that educators may be surprised by what students can actually do, despite their perceptions.

Then, an analysis of the professional learning needs of the teachers is conducted; it may well be that teachers know the students need “x,” but teachers may not know how to deliver “x,” without replicating their previous practices. This is critical since we cannot assume that teachers’ competencies include the skills needed to meet the gaps identified; after all, if they already had the skills to teach “x,” why do the students still need to learn “x”? This approach is based on evidence that teacher expectations and pedagogical skills can limit children’s educational achievement (Timperley and Phillips, 2003; Parr and Timperley, 2008).

Leaders must then take action to understand the professional development needs of teachers and take action to ensure that those resources are available. A key component of this stage is the active involvement of the various leaders; teachers will not take seriously this reform or change. If teachers see that the leader is not actively committed to the change and supporting them in adopting the change, then it is unlikely teachers will lead. Leaders who retreat to their offices to carry on administration are not acting as instructional leaders.

As and after the planned changes are implemented, assessment of impact on students is needed. Using an action research outlook (Robinson and Lai, 2006), this approach takes a tentative view of the efficacy of the changes on the basis that the theory underlying the analysis may be mistaken. It is easy to misunderstand the problem, to misdiagnose the causes, to mismeasure the domain, or ineffectively implement the change. Assessment data here may reveal new insights into the problem, suggesting further modifications to the solution. Having seen (or not) impact of the professional development on student learning, the assessment data are used to identify next steps in what students need next. It is important to monitor teacher practices to ensure that they do not slip back into old ineffective practices; continued assessment data can help identify reduced impact on student learning.

New Zealand research

Effective leadership in times of accountability will depend on the context in which schools and leaders are situated (Daniëls et al., 2019). So how are school leaders using educational assessment in New Zealand and does it really lead to improved student outcomes? It is worth noting first that the “approach to assessment, for example, allows for an emphasis on collaborative analysis of achievement data and a relentless focus on achievement using a range of assessments without the ‘baggage’ of high stakes assessment” (Lai et al., 2020, p. 35).

New Zealand has a somewhat unique assessment environment, that helps school leaders meet accountability demands in accordance with Leading for Learning principles. The *New Zealand Government* (2011, p. 9) stated that “In New Zealand, there is a deliberate focus on the use of professional teacher judgment underpinned by *assessment for learning* principles rather than a narrow testing regime.” The national assessment policy prior to Year 11 (students nominally 15 years old) emphasizes voluntary, school-based assessment for the purposes of raising achievement and improving instruction relative to the outcomes and objectives specified in the national curriculum (Crooks, 2010). New Zealand has standardized diagnostic tests that are aimed at informing improved teaching, a reliance on teachers’ judgment about learning, and a culture of collaboration within and between schools (Lai et al., 2020).

These affordances mean that leaders can treat assessment data as one part of the data profile of a school setting, allowing them to collaboratively understand why results are as they are, and experiment with solutions that are likely to redress learning outcome challenges. All of this can take place with little fear of being seen as a “bad” school because school accountability testing is not a practice in the environment. Furthermore, given that external school evaluations are carried out by the Education Review Office (ERO) through triennial site visits and school self-evaluations, the need for school accountability testing has been eschewed. The ERO does not require that schools demonstrate effectiveness with any one assessment method, but rather allows them to select from a range of assessment methods to demonstrate schooling effectiveness.

Research into improving outcomes in schools has relied on a number of support factors. Partnerships between schools and university researchers receive substantial funding from government. Interventions focus on improving teacher capabilities in response to robust diagnostic, formative assessment data. School and subject leaders were actively involved in analysis, interpretation, design, and monitoring of professional development activities. Substantial time for teacher changes to become visible on student learning performance were given. Under these conditions, evidence that leadership makes a difference is found.

Timperley (2005) found small but statistically significant effects ($d \approx 0.20$) in an 18-month action research study involving teachers and a school leader in learning to use assessment data to guide reading instruction. Government funded professional development projects in reading and writing have involved more than 300 schools and nearly 4000 teachers in 2–3 year-long school development studies (Parr et al., 2007; Sīlata et al., 2011; Timperley et al., 2010). These studies consistently show that after 2–3 years of professional development student learning in reading and writing gained more than national norms (observed changes in student learning were above national norm gains by between 0.50 and 2.50 times) and greater gains were seen among the lowest achieving students. However, not all school data analysis projects produce such large learning gains.

The more important changes lie in how data are used by leaders and teachers together. The Learning Schools Model (LSM; Lai and McNaughton, 2016; Lai and Schildkamp, 2016; McNaughton et al., 2012) involves school leaders and teachers working with professional researchers. The model involves creating baseline profiles and collaborative analysis of data, a series of workshops focused on analysis of data, and design and monitoring of interventions that should redress problems identified at the start. LSM has been implemented in schools with high proportions of low achieving students and contributed to increased student learning of about 3–4 months additional progress per year. While the gains are sustained and large, even after 3 years, they did not result in these below average students reaching the national norms. Large effects in student literacy were found in a 3-year longitudinal LSM study (Lai et al., 2009). In this study, learning gains were initially small, but showed evidence of acceleration over time suggesting persistence in the intervention process is required to bring about sustained change in teaching and learning. Furthermore, the changes were sustained because they became part of normal school routines for conducting organizational inquiry focused on solving teaching and learning problems through professional learning communities (Lai et al., 2009). Part of what makes the changes sustainable is that they are for the most part small modifications to existing practices (McNaughton and Lai, 2009); schools already (1) assess students, (2) plan teaching, (3) modify teaching, and (4) evaluate performance.

What is different throughout these interventions is the joint participation of the school leader with the teachers and a focus on student learning. The quality of the principal's involvement in these changes is critical; a study of five schools (Timperley, 2011a) in which there was exceptionally high learning gains revealed that the teachers considered their leader had scored 5 or 6 out of 6 for each of the five instructional leadership strategies identified by Robinson (2007), with the principals themselves emphasizing the function of promoting and participating in teacher learning and development. This really meant that the leaders not only had high levels of content and pedagogical knowledge, but they also led teachers into learning and implementing both kinds of knowledge (Timperley, 2011a). Hence, if the school leader lacks confidence or competence to act as a learning leader, changes around assessment practices or data use will be insufficient.

Conclusion

The challenge for school leaders is substantial. Leaders must not just manage and administer schools. If learning is to happen they must be competent in analyzing data and leading their teachers into changed classroom practices which are based on and evaluated with achievement data. Further, leaders must have deep expert knowledge in the content areas being considered; this of course requires distributed leadership through teams of leaders with expertise in relevant and complementary domains. Additionally, leaders must understand how teachers think about accountability pressures and be adept at raising professional thinking about the use of assessment data, while avoiding score inflation for the sake of conforming to accountability expectations. Many school leaders may be reluctant to initiate such action for fear of punitive responses from the political system. Hence, the challenge is also for government policy makers and inspection or evaluation officials to ensure that it is safe for schools to discover bad news about student learning. Under fear of consequences, the honesty to know and grapple with problems in teaching or leadership practices is unlikely. The New Zealand research especially shows that university researchers need to participate with schools to introduce new theories, new instructional techniques, and robust data analysis methods. Without combined efforts and resources, it is unlikely accountability assessment will achieve more than compliance—higher scores without deeper or richer learning outcomes. A significant challenge exists, but there are promising results from New Zealand, under contexts that are highly conducive to success, that school leaders, who use assessment for learning with their teaching staff in the ways described here, contribute to real learning gains.

Acknowledgment

I wish to thank my colleagues at the University of Auckland who helped me understand the methods and progress they had made in helping school leaders make a difference to staff and student learning. Specifically, thanks go to Dr. Brian Annan, Assoc. Prof. Mei Kuin Lai, Prof. Stuart McNaughton, Prof. Judy Parr, and Prof. Helen Timperley.

References

- Barnes, N., Fives, H., Dacey, C.M., 2015. Teachers' beliefs about assessment. In: Fives, H., Gregoire Gill, M. (Eds.), *International Handbook of Research on Teacher Beliefs*. Routledge, New York.
- Biesta, G.J.J., 2010. *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Paradigm Publishers, Boulder, CO.
- Bonner, S.M., 2016. Teachers' perceptions about assessment: competing narratives. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Brown, G.T.L., Harris, L.R., 2009. Unintended consequences of using tests to improve learning: how improvement-oriented resources heighten conceptions of assessment as school accountability. *J. MultiDiscip. Eval.* 6, 68–91.
- Brown, G.T.L., Hattie, J.A., 2012. The benefits of regular standardized assessment in childhood education: guiding improved instruction and learning. In: Suggate, S., Reese, E. (Eds.), *Contemporary Debates in Childhood Education and Development*. Routledge, London.
- Brown, G.T.L., Michaelides, M.P., 2011. Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *Eur. J. Psychol. Educ.* 26, 319–337.
- Brown, G.T.L., 2004. Teachers' conceptions of assessment: implications for policy and professional development. *Assess Educ. Princ. Pol. Pract.* 11, 301–318.

- Brown, G.T.L., 2008. *Conceptions of Assessment : Understanding What Assessment Means to Teachers and Students*. Nova Science Publishers, New York.
- Brown, G.T.L., 2009a. Teachers' self-reported assessment practices and conceptions: using structural equation modelling to examine measurement and structural models. In: Teo, T., Khine, M.S. (Eds.), *Structural Equation Modelling in Educational Research: Concepts and Applications*. SensePublishers, Rotterdam, NL.
- Brown, G.T.L., 2009b. The reliability of essay scores: the necessity of rubrics and moderation. In: Meyer, L.H., Davidson, S., Anderson, H., Fletcher, R., Johnston, P.M., Rees, M. (Eds.), *Tertiary Assessment and Higher Education Student Outcomes: Policy, Practice and Research*. Ako Aotearoa, Wellington, NZ.
- Brown, G.T.L., 2010. The validity of examination essays in higher education: issues and responses. *High Educ. Q.* 64, 276–291.
- Brown, G.T.L., 2011. Teachers' conceptions of assessment: comparing primary and secondary teachers in New Zealand. *Assess. Matters* 3, 45–70.
- Brown, G.T.L., 2013. asTTle—a national testing system for formative assessment: how the national testing policy ended up helping schools and teachers. In: Lai, M.K., Kushner, S. (Eds.), *A National Developmental and Negotiated Approach to School and Curriculum Evaluation*. Emerald Group Publishing, London.
- Brown, G.T.L., 2017. What we know we don't know about teacher education. In: Clandinin, D.J., Husu, J. (Eds.), *The SAGE International Handbook of Research on Teacher Education*. Sage, Thousand Oaks, CA.
- Brown, G.T.L., Chaudhry, H., Dhamija, R., 2015. The impact of an assessment policy upon teachers' self-reported assessment beliefs and practices: a quasi-experimental study of Indian teachers in private schools. *Int. J. Educ. Res.* 71, 50–64.
- Brown, G.T.L., Hui, S.K.F., Yu, F.W.M., Kennedy, K.J., 2011a. Teachers' conceptions of assessment in Chinese contexts: a tripartite model of accountability, improvement, and irrelevance. *Int. J. Educ. Res.* 50, 307–320.
- Brown, G.T.L., Lake, R., Matters, G., 2011b. Queensland teachers' conceptions of assessment: the impact of policy priorities on teacher attitudes. *Teach. Teach. Educ.* 27, 210–220.
- Brown, G.T.L., O'leary, T.M., Hattie, J.A.C., 2018. Effective reporting for formative assessment: the asTTle case example. In: Zapata-Rivera, D. (Ed.), *Score Reporting: Research and Applications*. Routledge, New York.
- Brown, G., Kennedy, K., Fok, P., Sang, J., 2009. Assessment for student improvement: understanding Hong Kong teachers' conceptions and practices of assessment. *Assess Educ. Princ. Pol. Pract.* 16, 347–363.
- Buckendahl, C.W., 2016. Public perceptions about assessment in education. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Cannell, J.J., 1989. How Public Educators Cheat on Standardized Achievement Tests. Friends for Education, Albuquerque, NM.
- Coleman, J., 1966. *Equality of Educational Opportunity*. National Center for Educational Statistics, Washington, DC.
- Crooks, T.J., 2010. Classroom assessment in policy context (New Zealand). In: Mcgraw, B., Peterson, P., Baker, E.L. (Eds.), *The International Encyclopedia of Education*, third ed. Elsevier, Oxford, UK.
- Daniëls, E., Hondeghem, A., Dochy, F., 2019. A review on leadership and leadership development in educational settings. *Educ. Res. Rev.* 27, 110–125.
- Elley, W.B., Livingstone, I.D., 1972. *External Examinations and Internal Assessments: Alternative Plans for Reform*. New Zealand Council for Educational Research, Wellington, NZ.
- Fives, H., Buehl, M.M., 2012. Spring cleaning for the "messy" construct of teachers' beliefs: what are they? Which have been examined? What can they tell us? In: Harris, K.R., Graham, S., Urdan, T. (Eds.), *APA Educational Psychology Handbook: Individual Differences and Cultural and Contextual Factors*. APA, Washington, DC.
- Fjortoft, H., Lai, M.K., 2020. Affordances of narrative and numerical data: a social-semiotic approach to data use. *Stud. Educ. Eval.* 100846.
- Fulmer, G.W., Lee, I.C.H., Tan, K.H.K., 2015. Multi-level model of contextual factors and teachers' assessment practices: an integrative review of research. *Assess Educ. Princ. Pol. Pract.* 22, 475–494.
- Gebril, A., Brown, G.T.L., 2014. The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assess Educ. Princ. Pol. Pract.* 21, 16–33.
- Gebril, A., 2016. Educational assessment in Muslim countries: values, policies, and practices. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Harris, L.R., Brown, G.T.L., 2016. Assessment and parents. In: Peters, M.A. (Ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Singapore, Singapore.
- Hattie, J., 2009. *Visible Learning: A Synthesis of Meta-analyses in Education*. Routledge, London.
- Hattie, J.A., Brown, G.T.L., 2008. Technology for school-based assessment and assessment for learning: development principles from New Zealand. *J. Educ. Technol. Syst.* 36, 189–201.
- Hattie, J.A., Brown, G.T.L., 2010. Assessment and evaluation. In: Rubie-Davies, C. (Ed.), *Educational Psychology: Concepts, Research and Challenges*. UK Routledge, Abingdon.
- Hattie, J.A.C., Brown, G.T.L., Keegan, P.J., 2003. A national teacher-managed, curriculum-based assessment system: assessment Tools for Teaching & Learning (asTTle). *Int. J. Learn.* 10, 771–778.
- Hattie, J.A.C., Brown, G.T.L., Keegan, P.J., Mackay, A.J., Irving, S.E., Cutforth, S., Campbell, A., Patel, P., Sussex, K., Sutherland, T., McCall, S., Mooyman, D., Yu, J., December 2004. *Assessment Tools for Teaching and Learning (asTTle) Manual (Version 4, 2005)*. University of Auckland/Ministry of Education/Learning Media, Wellington, NZ.
- Lai, M.K., McNaughton, S., 2016. The impact of data use professional development on student achievement. *Teach. Teach. Educ.* 60, 434–443.
- Lai, M.K., Schildkamp, K., 2016. In-service teacher professional learning: use of assessment in data-based decision-making. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Lai, M.K., McNaughton, S., Amituanai-Tolosa, M., Turner, R., Hsiao, S., 2009a. Sustained acceleration of achievement in reading comprehension: the New Zealand experience. *Read. Res. Q.* 44, 30–56.
- Lai, M.K., McNaughton, S., Jesson, R., Wilson, A., 2020. *Research-Practice Partnerships for School Improvement the Learning Schools Model*. Emerald Publishing Limited, Bingley, UK.
- Lai, M.K., McNaughton, S., Timperley, H., Hsiao, S., 2009b. Sustaining continued acceleration in reading comprehension achievement following an intervention. *Educ. Assess. Eval. Account.* 21, 81–100.
- Lerner, J.S., Tetlock, P.E., 1999. Accounting for the effects of accountability. *Psychol. Bull.* 125, 255–275.
- Lillejord, S., 2023. A more intelligent accountability – future directions. In: *International Encyclopedia of Education*, fourth ed. Elsevier.
- Lingard, B., Lewis, S., 2016. Globalization of the Anglo-American approach to top-down, test-based educational accountability. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Linn, R.L., 2000. Assessments and accountability. *Educ. Res.* 29, 4–16.
- McNaughton, S., Lai, M.K., 2009. A model of school change for culturally and linguistically diverse students in New Zealand: a summary and evidence from systematic replication. *Teach. Educ.* 20, 55–75.
- McNaughton, S., Lai, M.K., Hsiao, S., 2012. Testing the effectiveness of an intervention model based on data use: a replication series across clusters of schools. *Sch. Effect. Sch. Improv.* 23, 203–228.
- Meissel, K., Meyer, F., Yao, E.S., Rubie-Davies, C.M., 2017. Subjectivity of teacher judgments: exploring student characteristics that influence teacher judgments of student ability. *Teach. Teach. Educ.* 65, 48–60.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. MacMillan, Old Tappan, NJ.
- Ngan, M.Y., Lee, J.C.K., Brown, G.T.L., 2010. Hong Kong principals' perceptions on changes in evaluation and assessment policies: they're not for learning. *Asian J. Educ. Res. Synergy* 1.
- New Zealand Government, 2011. *Ministry of Education Position Paper: Assessment [Schooling Sector]*. Learning Media, Wellington, NZ.
- Nichols, S.L., Harris, L.R., 2016. Accountability assessment's effects on teachers and schools. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.

- Parr, J.M., Timperley, H., 2008. Teachers, schools and using evidence: considerations of preparedness. *Assess Educ. Princ. Pol. Pract.* 15, 57–71.
- Parr, J.M., Timperley, H., Reddish, P., Jesson, R., Adams, R., 2007. Literacy Professional Development Project: Identifying Effective Teaching and Professional Development Practices for Enhanced Student Learning. Ministry of Education, Research Division, Wellington, NZ.
- Richardson, V., Placier, P., 2001. Teacher change. In: Richardson, V. (Ed.), *Handbook of Research on Teaching*, fourth ed. AERA, Washington, DC.
- Robinson, V., Lai, M.K., 2006. *Practitioner Research for Educators: A Guide to Improving Classrooms and Schools*. Corwin Press, Thousand Oaks, CA.
- Robinson, V.M.J., October 2007. School leadership and student outcomes: identifying what works and why (William Walker Oration). In: Gurr, D. (Ed.), *National Conference of the Australian Council for Educational Leaders*. Australian Council for Educational Leaders, Melbourne, Aus.
- Saldaña, C.M., Shepard, L.A., 2023. Educational accountability: the use and abuse of assessment data. In: *International Encyclopedia of Education*, fourth ed. Elsevier.
- Si'ilata, R., Dreaver, K., Parr, J., Timperley, H., Meissel, K., 2011. Tula'i Mai! Making a Difference to Pasifika Student Achievement in Literacy Literacy Professional Development Project 2009–2010. Auckland UniServices Ltd, Auckland, NZ.
- Timperley, H., 2011a. Knowledge and the leadership of learning. *Leader. Pol. Sch.* 10, 145–170.
- Timperley, H., 2011b. Leading teachers' professional learning. In: Robertson, J., Timperley, H. (Eds.), *Leadership and Learning*. Sage, London.
- Timperley, H., 2011c. Using Student Assessment for Professional Learning: Focusing on Students' Outcomes to Identify Teachers' Needs (Paper No. 21). Department of Education and Early Childhood Development, Education Policy and Research Division, Office for Policy, Research and Innovation, Melbourne, Aus.
- Timperley, H.S., Phillips, G., 2003. Changing and sustaining teachers' expectations through professional development in literacy. *Teach. Teach. Educ.* 19, 627–641.
- Timperley, H.S., 2005. Instructional leadership challenges: the case of using student achievement information for instructional improvement. *Leader. Pol. Sch.* 4, 3–22.
- Timperley, H., Parr, J., Meissel, K., 2010. Making a Difference to Student Achievement in Literacy: Final Research Project (Milestone 3) on the Literacy Professional Development Project 2009–2010. Auckland Uniservices Limited, Auckland, NZ.
- Wilson, A., Lai, M.K., 2019. Leading literacy change through collaborative analysis of practice. *J. Adolesc. Adult Literacy* 62, 455–458.

Accountability in an age of international testing

Jenny Ozga, University of Oxford, Department of Education, Oxford, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

Introduction	306
Drivers of TBA: globalization, NPM, data	306
International organizations	308
The OECD's PISA and beyond	308
References	311

Introduction

Many discussions of accountability start by attempting its definition, and most conclude that there are many types—for example, professional accountability, bureaucratic accountability and market-based or technical-managerial accountability (Barbana et al., 2020; Grek et al., 2021). These different definitions illuminate the need to approach accountability from a perspective that is attentive to the social and political conditions within which the term is being employed, the actors employing it, and the contexts within which they operate. Such a framework necessarily connects changes in accountability to changes in governance, in education and beyond. This discussion draws on research on education primarily, but not exclusively, focusing on Europe, which explores knowledge-governing relations and highlights the knowledge-based resources available to transnational governing elites, including regulatory instruments and performance data (see for example, Grek, 2009; Ozga et al., 2011; Lawn and Grek, 2012; Fenwick et al., 2014; Verger et al., 2019). It shows how those instruments make possible the shift toward governing through networks of new actors, often located in International Organizations (IOs) who define education in global terms, prioritize the redefinition of education as learning, and mobilize the application of policy “solutions” derived from benchmarking, indicators and performance measurement across nation-states. (Gorur, 2017; Waldow and Steiner-Khamsi, 2019).

Accountability, then, is placed here within the context of neo-liberal globalization, that drives the economizing of education, and the promotion of competition as the lever that produces improved performance. In tandem, new public management (NPM) installs monitoring and performance management systems, while the growth of technical capacity enables the collection, analysis and distribution of unprecedented amounts of performance data, and supports the production of benchmarks, indicators and rankings, which together constitute a dominant way of seeing and thus governing, education. The more recent growth in complexity of data creates a space for expert “translators” and mediators, especially those based in or connected with IOs, as well as those in private companies, to take active roles in governing, and to promote as “natural” (Novoa and Yariv-Mashal, 2003) evidence-based solutions to educational problems. Raising attainment levels and connecting this drive to increased economic performance become the dominant drivers of education policy and create pressures for standardization and comparison, including through the promotion of test-based accountabilities. These developments are highly interdependent and are discussed below.

Drivers of TBA: globalization, NPM, data

Neo-liberal globalization drives the integration of markets worldwide (Bishop and Payne, 2021) and erodes the distinction between states and markets. It brings about changes in the architecture of power globally and within nation-states, privileging the market and enhancing the power of economic institutions, such as multi-nationals, and other IOs. It can thus be understood as the overarching context in relation to changing governing, while the associated agenda of knowledge economy-driven change installs competition and suppresses historically embedded shared values and practices that have informed national education provision, including in relation to accountability. Neo-liberal globalization has shaped and is shaping a global structured agenda for education (Robertson and Dale, 2017), as broadly similar education “reform” programs are implemented in widely varying settings (Sahlberg, 2016).

As the measurement and comparison of learning outcomes is understood to be central to establishing and improving levels of economic competitiveness globally (Verger et al., 2019), educational processes are collapsed into outcomes, especially demonstrable and quantifiable learning outcomes. All forms of education, from early childhood provision, through schools, vocational education and training, and higher education are interconnected through the attitudes and dispositions encouraged by new learning agendas and practices, including international testing, which claim consistency across populations, institutions and national contexts. Schooling and post-school education is focused on developing learners’ capacity to learn, while learners also learn to constantly evaluate their learning through self-evaluation, and through the use of international performance testing that produces comparative data from which systems supposedly learn how to improve (Addey et al., 2017; Falabella, 2021; Grek, 2013).

Globalization thus erodes state responsibility and state capacity for service provision. The dispersion of state responsibilities produced by the combined effects of neo-liberal ideology and NPM principles weakens the responsibility of national politicians for key areas of public policy and disrupts traditional lines of accountability. These lines of accountability were designed—in

principle, if not always in practice—to enable the effective communication of decisions, to respond to the expression of public feeling, and to sustain policy-makers' claims to legitimate authority. The privatization and decentralization that accompanied the implementation of New Public Management (NPM), combined with simultaneous decentralization and re-centralization of government functions, obscures and complicates accountability, making it hard to see clearly who is accountable for what. Direct lines of political accountability from governments to the electorate and through elected local representatives with responsibility for provision are reduced or obscured. Instead, technical-managerial accountability operates to ensure that, through the publication of data across a range of locales—from individual learners, teachers, institutions, municipalities and national governments, citizens, redefined as stakeholders, can see and judge performance, and make choices informed by this evidence. In this form, as Newman and Clarke suggest, accountability has been adopted internationally as part of the conception of information or evidence-based “good governance” in which transparency, participation and performance are apparently combined (see Newman and Clarke, 2009: 100–102).

In Europe, it is possible to trace a shift, to a greater or lesser degree, from professional accountability to what Ranson (2003) calls neo-liberal corporate accountability, and then-again, in differing degrees-towards test-based accountability. That transition expressed movement from the hierarchical practices of public sector bureaucracies in which accountability “upwards” involved teachers being accountable to parents, to the local authority/municipality, to governing boards, while local and national politicians were “downwards” accountable to their electorates. These forms of accountability involved exchanges of “accounts,” that is, narratives that conveyed meaning about values in contexts of multiple and reciprocal answer abilities, and which could differ, that is they could involve conflicts of purpose in different contexts (Ranson, 2003: 461). It should be stressed that these narratives are not without problems, and it is recognized that the outcomes of conflicts of purpose within that system tended to favor the more powerful. However, the point is that spaces for debate and conflict about priorities, purposes and values in the provision of education existed, at least acknowledging the contested politics and power relations that are an inherent feature of education policy (Ozga, 2020).

From the 1980s onwards, especially in the Anglosphere, policy makers adopted the practices and assumptions of NPM and so redefined accountability as the management and scrutiny of performance at all levels—from the international competition of league table positioning though to the constant review of performance against hard and soft targets that is installed in teachers and pupils through forms of self-evaluation. More recently, test based accountability is acknowledged as spreading internationally (Addey et al., 2017; Barbana et al., 2020) throughout continental Europe, including the Nordic countries (Camphuijsen et al., 2020; Mau-Sethagen, 2013) and globally, in the USA, in Australia, Canada and Latin America. As well as the ideological dominance of neo-liberalism it is important, in accounting for the international take up of this version of accountability, albeit at different speeds and to different degrees, to note the increased capacity and use of data in education as in other fields, and the contribution that data growth and development make to TBA. Data enter the frame because the transmission of information is necessary to hold together the diverse and networked governing forms that accompany globalization. Data enable this complex education arena to be calculated, for example through the Bologna Accord, that harmonizes different European higher education systems by creating a single degree system and the Lisbon Process or Open Method of Coordination (OMC) in which benchmarking and comparison are core governing processes across a “learning society” shaped by economic reform, citizenship obligations, employability and international comparison and performance testing in education (Grek et al., 2021).

Local government and schools are now rendered visible and calculable (Ozga et al., 2011). Data expressed as public rankings, league tables and international test results enable and consolidate control over a wide network of actors and institutions—national governments, local authorities, schools and teachers included. The construction of indicators and the collection and processing of performance data install comparison as a basic principle in building policy consensus and also, importantly, in embedding self-regulation (Ozga et al., 2011; Grek, 2014) as data systems not only monitor the activity of teachers and learners, but also shape individual conduct while apparently enabling autonomous, choice-making activity supported by information (Piattoeva, 2014). Throughout the countries in the various research projects on which this paper draws, there is a growth in knowledge/information expressed in quantitative forms, through comparative statistics, charts, diagrams and comparative tables, competency-based curricula and pedagogy that is partly driven by measurements of competence. There is a stronger emphasis on the production of knowledge about performance in quantitative forms that can be shared, analyzed and translated into benchmarks and indicators. There is a growth of evaluation, accompanied by “scientisation.” Accountability is clearly linked to the production and circulation of data-based knowledge and is linked to introducing self-regulation (of teachers and pupils). There is increasing emphasis on technical, “objective” knowledge as the basis for evaluation of performance by external agencies and internal audiences alike.

These different developments together change the way in which education is understood and visualized: data enable schools and education systems to be seen as “computational” projects (Selwyn, 2018) and the “modeling” of education through digital data creates algorithmically driven “systems thinking” where complex social problems are converted into complex but apparently solvable statistical problems. Thus, digital data encourage “solutionism” and suppress debate and discussion about the purposes and content of education, or about different assessment forms, while the space available for political conflict or contention about priorities, purposes and process is effectively shut down. In this way data use produces education settings (Selwyn, 2018). Put differently, data and data practices are not neutral representations or measurements of educational activities, but ways of establishing these practices as defining features of education.

Digital data work within education/schooling is now normalized—it is understood almost universally as the basis of improvement, and it is increasing in scope, including in Higher education, where multi-nationals see market opportunities. From this perspective, the very significant development of data resources in education from the early 2000s is not only a story of increased

technical capacity but connects directly to the prioritization and spread of test-based assessments. Furthermore, data enable the growth of networks engaged in governing work, often through IOs or transnational organizations, whose influence is discussed below.

International organizations

The combined effects of neo-liberal globalization, new public management and the growth and increased sophistication of data systems together enable international organizations (IOs) such as the OECD, the World Bank or the European Commission, or private, corporate actors such as McKinsey or Pearson, to occupy a key space in the governing of education. The global governance of education is increasingly organized through multi-layered international networks of collaboration linking national and local actors to worldwide networks and structures, with the capacity for global political action.

The term “governance” rather than government is used by many contemporary scholars to draw attention to the growing diversity of actors and organizational forms engaged in governance work, including in education. The term indicates a changed landscape: one in which the traditional models of government, understood as restricted to specific, legally-recognized functions and levels within the nation-state, no longer apply. Governance implies acceptance of a shift away from fixed and bordered institutions, from stable bureaucratic forms and hierarchies toward the flexible and adaptive modes of co-ordination of markets and networks (Ozga, 2019) and may also imply a decline in the influence and capacity of the state. Indeed the shift to governance follows logically from the increased complexity of service provision, as the traditional state form became more and more differentiated to meet more complex demands. Networks then emerge as the natural form for the coordination and management of activity across different sectors (see, e.g., Ball et al., 2016; Shiroma, 2014), and the state becomes one actor among many.

These networks are engaged in what has been called “epistemic governance” that reflects the contemporary interdependency of governing and comparative, data-based knowledge (Alasuutari, 2016; Normand, 2016). The growth of this form of expertise is recognized as a transnational phenomenon, with experts increasingly working between national and transnational arenas, and identified as a “new governing elite” (Stone, 2013: 41, 2019) and a new “European technocracy” (Normand, 2016: 129), located in IOs. These IOs draw on the authority of science and expertise to promote their agendas and that politicization of science weakens its claims to independence. The knowledge distilled from test-based assessments supposedly enables IOs to engage in policy “teaching,” extracting from data about comparative performance lessons which enable systems to improve the quality of education provision and move their systems in productive, evidence-based policy directions. As indicated earlier, these agendas are highly dependent on growth of data use and capacity-in education and across public sector service provision. Data and data systems construct policy problems and frame policy solutions beyond and across the national scale. Epistemic governance moves more freely, uses more artifacts, and involves new actors in its production. The knowledge that is valued in governing is “actionable” (Grundmann and Stehr, 2012): that is, it contains advice about the direction to be taken. The massive expansion of information requires particular skills in simplifying, synthesizing and brokering to reduce possibilities and make knowledge actionable:

The rapid growth of experts, advisers and consultants in education arises from the rapid expansion of knowledge/information, this provides opportunities for simplification of the problem of endless competing interpretation in order to provide a basis for action.

Grundmann and Stehr (2012: 20–21).

This network of experts operates to produce and disseminate a “cognitive consensus” (Grundmann and Stehr, 2012) that is, a system of standardized policy agendas and their repertoire of benchmarks, indicators and competitive testing regimes. The homogeneity of programs of whole system structural education reform (Sahlberg, 2016) is striking, and this consensus is promoted through appeals to its “evidence based” nature, or its “common sense” about “what works” which conceals the political nature of translation, because, as Shiroma (2014) points out, the label “expert” confers scientific status and authority and seems to be above politics. The work that policy actors do in networks, whether as “chosen” experts, mediators and translators for OECD, or in private consultancy, strengthens the trend toward comparison and the search for trends and patterns in comparative data, and also increases the influence of analysts who can interpret data and turn it into policy advice (Grundmann and Stehr, 2012: 20–21).

The OECD's PISA and beyond

The Organization for Economic Co-operation and development (OECD) provides a paradigmatic case study of the increasing transnational influence of IOs in promoting standardized policies in education globally, and in using test-based accountability to sustain this governing activity. The OECD is able to present itself as an expert knowledge-based organization, “above” national politics, and sufficiently distanced from national contexts to be dispassionate and objective (Rautalin et al., 2021). The most visible manifestation of OECD's influence is the Program for International Student Assessment (PISA) first implemented in 2000. PISA measures students' knowledge and skills in three main domains: reading, scientific, and mathematical literacy as well as problem-solving capacity and financial literacy. PISA was redeveloped in 2012 into PISA for Development (PISA-D) with the aim of making the

instruments more relevant to low- and middle-income countries, a shift that also illustrated an expansion from simple data provision to stronger steering that seeks to change local cultures of assessment (Hopfenbeck et al., 2017). These assessments are decontextualized: that is, they are not concerned with national curricula and their success in promoting learning, but only with cross-national comparisons of levels of skill.

More recently, PISA has been further developed into the PISA Global Competence (PGC) program of assessment, which claims to measure:

'students' capacity to examine local, global and intercultural issues, to engage in open, appropriate and effective interactions with people from different cultures, and to act for collective well-being and sustainable development.

OECD (2018).

OECD's move into global competences, suggests, as Sellar (2015: 425, quoted in Engel et al., 2019) argues, that the organization is seeking to capture a broader range of skills and capacities that are presented as measurable globally and as "free" of cultural shaping or influences. By making competence a core concern of education systems, OECD ensures its continuing dominance in assessing progress in education, and by changing the basis of measurement, enables its even wider spread.

Global Competence is defined as "a multi-dimensional construct that requires a combination of knowledge, skills, attitudes and values successfully applied to global issues or intercultural situations" (OECD, 2018). This "multi-dimensional construct" was developed by experts who designed it to capture capacities that are understood to be measurable and assessable globally. As Gardinier (2020: 8–9) demonstrates, the measurement of global competence promotes an idea of global learners as "autonomous individuals who can separate themselves from their social identities." The imagined ideal learner is "future oriented, socially-mature, goal oriented, and extremely self-aware" (*op. cit.*: 9–10).

OECD describes global competence as follows:

"Global competence can help young people:

- develop cultural awareness and respectful interactions in increasingly diverse societies;
- recognize and challenge cultural biases and stereotypes, and facilitate harmonious living in multicultural communities;
- prepare for the world of work, which increasingly demands individuals who are effective communicators, are open to people from different cultural backgrounds, can build trust in diverse teams and can demonstrate respect for others, especially as technology continues to make it easier to connect on a global scale;
- capitalize on inherently interconnected digital spaces, question biased media representations, and express their voice responsibly online;
- care about global issues and engage in tackling social, political, economic and environmental challenges.

Global competence thus supports the Sustainable Development Goals (SDGs) both by providing the vision of education the SDGs advocate for, and by encouraging young people to act in the general interest of collective-wellbeing and sustainable development that the SDGs embody" (OECD, 2018).

This form of assessment, that seamlessly blends quite divergent and contested issues reflecting substantial social, political and economic problems, very clearly reflects concerns to manage problematic issues in global governance, problems that have arisen because of the exclusive focus on the economy, for example, increasing racism, unemployment and environmental damage. However, PGC places the responsibility for responding to these issues on individual learners, who must demonstrate their competence in "harmonious living," labor market readiness and environmental awareness.

OECD's description of PGC implies that there is "global consensus" on global competence and that experts have simply expressed this "cognitive consensus" and produced an assessment that countries can use to measure this concept. Yet it is apparent that PGC contains conflicting rationales and as Engel et al. (2019) point out, there is no consensus about who is imagined to be and gets the opportunity to become a globally competent citizen." Indeed, the promotion of the idea of consensus enables PGC to be shaped in ways that "may exclude orientations that do not fit the profile and values of the Northern/Western national citizen" (Engel et al., 2019: 125).

Concerns about the global applicability and comparability of the constructs that shape PISA in all its forms have long been expressed, and there is also an important literature on technical and conceptual issues with the design of PISA for schools (see, e.g., Baird et al., 2017). Furthermore, researchers have questioned whether there is, in fact, persuasive evidence that PISA and test-based accountability policies in general do improve the quality and efficiency of public education, while also highlighting the evidence that increased high-stakes testing is restricting students' conceptual learning, and damaging their understanding of creative action and innovation (Sahlberg, 2016).

PISA as a governing instrument is best understood as the direct product of the combination of the economization of education, the spread of NPM principles, and the influence of data systems and data-based expertise. The comparative basis of PISA enables participants to demonstrate globally and to their national publics that they embrace the modernizing logic of competition and reliance on evidence, it locates countries in the worldwide competitive space, identifies their relative positioning in the knowledge economy and holds out the promise of learning from the example of more successful competitors. Indeed, PISA has been shown to be a very powerful policy instrument that is able to penetrate different kinds of policy-making regimes and very diverse political

circumstances, because of its promotion as the “gold standard” of performance assessment, and because of its apparent adaptability to many disparate political ends. Carvahlo (2012) identified three modes of PISA knowledge use (i) as a basis for knowledge-based analyses of specific issues on the policy agenda; (ii) as a source for studies of national education “problems”; and (iii) as a resource for improving domestic regulation tools—in particular accountability mechanisms and quality assurance policies. PISA knowledge was mobilized to support the legitimacy of specific education policies within a broad policy framework, promoted by the European Commission, of enhanced economic competitiveness, as other commentators have noted (Lawn and Grek, 2012:104–105).

These are highly performative forms of accountability: policy makers are acutely conscious of the need to be seen to succeed both within the setting of the transnational construction of PISA and by their national media in relation to PISA outcomes. The OECD’s diagnosis of underperformance is accepted as the key problem that needs to be tackled through competition driven by the monitoring of the relative positioning of nations, schools, teachers and pupils. Policy makers are not rendered accountable to their electorates for poor performance as judged by PISA, rather they use PISA to justify policy directions that they had already sought to implement in their own contexts, especially where these data enable stronger measures of deregulation and privatization. Accountability is subsumed by visibility: where poor performance is acknowledged by a national policy maker it is attributed to poor teacher performance and not to national policy.

The nation-states participating in PISA and studied in research on its early implementation and effects were found to engage actively with this regulatory instrument despite the threat that PISA may pose to the nation-state’s capacity to act independently in the governing of its education system. This engagement requires the building of consensus about how problems are translated into national contexts through the creation of consensus and the incorporation and dissemination of ways of thinking and acting that draw policy makers into monitoring and mutual surveillance of education systems—practices that define their new identities and roles as system managers, operating within a technical-managerial form of accountability. PISA texts enable policy-makers to underline the rationality and scientific basis of their actions; to embrace modernization agendas that break away from national traditions, and to see their system’s performance in comparative terms within a European and global competitive space (Ozga et al., 2011). More recent scholarship, focusing on newer participants, suggests that participation may contribute to nation-building to access to political, economic or trade entities (Addey et al., 2017; Hopfenbeck et al., 2017); or is due to pressure to participate as signatories of global commitments. At the same time, PISA creates a space of discussion and brokering in which national and transnational experts and policy makers to interact in ways that allow nation states to be seen and to see themselves as independent and voluntary participants (Waldow and Steiner-Khamsi, 2019) while encouraging “best practice” and policy learning that enables national adaptation of the OECD “script.”

Indeed that “continuous crafting of its relationship with key education actors in other IOs and within national contexts” is a key element in explaining the continuous and growing influence of OECD (Grek, 2017: 298), and the interaction of local, national and international actors in “iterative processes of collective learning” is key to OECD’s achievement of a paradigm shift in the thinking and framing of education. The influence of global networks of experts is further apparent in the growing use and impact of Country Reviews (see, for example, Grek, 2013, 2017; Browes and Verger, 2020). PISA data first identifies “problems” that need to be addressed, and those problems may then be further investigated and remedies proposed through OECD’s process of Country Review, in which international teams of experts are invited by national governments to work with “local” experts to gather and analyze data to highlight issues at the national level, and then to provide advice and tools judged necessary to put things right. For example, Scotland has engaged substantially with OECD, following poor performance in PISA for schools 2015, where scores for maths, reading and science all declined. The Scottish Government commissioned OECD expert reviews (OECD, 2015) as well as appointing an International Council of Education Advisers, which together called for intensive reform processes that signaled disruptions and departures from established relationships between expertise, knowledge and policy. These required substantial system restructuring, increased concentration on attainment measures and stronger assessment and testing regimes. Media reporting of the PISA “crisis” fueled criticism of the teaching profession and encouraged an attempted policy shift toward the closer alignment of Scottish education with OECD’s norms (Ozga, 2021).

Research has identified different kinds of responses to the PISA results and very different effects on national policies (Waldow and Steiner-Khamsi, 2019; Grek et al., 2021). Research in Europe shows that in some spaces PISA results are suppressed while in others they become part of a triennial ephemeral spectacle (Grek, 2009). However, PISA’s significance in relation to the widespread equation of accountability with international test-based assessments is unquestioned; the literature illustrates its significance as a policy instrument that builds consensus through comparison, ordering and meaning-making within the framework of accountability understood as performance-based reporting. PISA draws its power from the combined effects of globalization, NPM and data, enabling policy to use comparative data to legitimizing preferred national policies on the basis of “international measures.” However what counts is not just the prevalence of comparison, but its use as the basis for:

the circulation of languages that tend to impose as ‘evident’ and ‘natural’ specific solutions for educational problems.

Novoa and Yariv-Mashal (2003).

Here Novoa and Yariv-Mashal highlight one of the key consequences of international comparison, and especially international performance testing: the centrality of competition to this mode of governance, and the push toward standardization through the elimination of alternative conceptualizations of what the key issues, problems or solutions in education might be. The example

of the assessment of global competence illustrates this well. The necessity of improving performance to meet economic need, and the assumption of improvement through competitive testing and ranking become “natural” orthodoxies that cannot be challenged. They further suggest that the continuous collection, production and publication of surveys of attainment create an “international spectacle” that replaces political debate; they identify a “logic of comparison” working through policy exchange, transfer and agreements to construct “a common understanding of ‘best practices’, leading to standardization of policy in Europe and beyond. They point out that the protected status of education in Europe as a sphere of national policy making through the principle of subsidiarity does not impede this process, as the European Commission’s construction of indicators serves as a “reference point” that eventually leads the various national institutions to adopt “freely” very similar actions and perspectives within the educational field. Novoa and Yariv-Mashal conclude that:

This new approach to European affairs reveals, clearly, a strategy to move the discussion away from matters of government (habited by citizens, elections, representation, etc.) and place it in the more diffused level of governance (habited by networks, peer review, agreements, etc.).
Novoa and Yariv-Mashal (2003).

Here Novoa and Yariv-Mashal capture key elements in the argument made in this piece about understanding international assessments as instruments for the governing of education, expressing as they do the combined force of change driven by neo-liberal globalization, by new public management practices, the growth of data and technical capacity, and the increasing influence of international organizations and the experts who work with them. These together function to exclude or conceal politics, and to erode and suppress spaces of debate and contestation around alternatives to the cognitive consensus.

References

- Addey, C., Sellar, S., Steiner-Khamisi, G., Lingard, B., Verger, A., 2017. The rise of international large-scale assessments and rationales for participation. *Compare* 47 (3), 434–452.
- Alasuutari, P., 2016. *The Synchronization of National Policies*. Routledge, London.
- Baird, J.A., Andrich, D., Hopfenbeck, T., Stobart, G., 2017. Assessment and learning: fields apart. *Assess. Educ.* 24 (3), 317–350.
- Ball, S.J., Junemann, C., Santori, D., 2016. Joined -up policy: network connectivity and global education. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley Blackwell, Chichester, pp. 535–553.
- Bishop, M.L., Payne, A., 2021. The political economies of different globalizations; theorizing reglobalization. *Globalizations* 18 (1), 1–21.
- Barbana, S., Dumay, X., Dupriez, V., 2020. Accountability policy forms in European education systems, and introduction. *Eur. Educ. Res. J.* 19 (2), 87–93.
- Browes, N., Verger, A., 2020. Global governance through peer review: the Dutch experience of OECD reviews of National Policies for Education. *Crit. Pol. Stud.* <https://doi.org/10.1080/19460171.2020.1865177>.
- Camphuijsen, M., Moeller, J., Skedsmo, G., 2020. Test based accountability in the Norwegian context. *J. Educ. Pol.* 36 (5), 624–642.
- Carvalho, L.M., 2012. The fabrications and travels of a knowledge-policy instrument. *Eur. Educ. Res. J.* 11 (2), 172–188. <https://doi.org/10.2304/eej.2012.11.2.172>.
- Engel, L.C., Rutkowski, D., Thompson, G., 2019. Toward an international measure of global competence? A critical look at the PISA 2018 framework. *Glob. Soc. Educ.* 17 (2), 117–131. <https://doi.org/10.1080/14767724.2019.1642183>.
- Falabella, A., 2021. The seduction of hyper-surveillance: standards, testing and accountability. *Educ. Adm. Q.* 57 (1), 113–142.
- Fenwick, T., Mangez, E., Ozga, J. (Eds.), 2014. *Governing Knowledge: Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education*. World Yearbook of Education 2014. Routledge, London.
- Gardinier, M.P., 2020. Imagining globally competent learners: experts and education policy making beyond the nation state. *Comp. Educ.* 57 (1), 130–146.
- Gorur, R., 2017. Statistics and statecraft: exploring the potentials politics and practices of international educational assessment. *Crit. Stud. Educ.* 58 (3), 261–265.
- Grek, S., 2009. Governing by numbers: the PISA ‘effect’ in Europe. *J. Educ. Pol.* 24:1, 23–37. <https://doi.org/10.1080/02680930802412669>.
- Grek, S., 2013. Expert moves: international comparative testing and the rise of expertocracy. *J. Educ. Pol.* 28 (5), 695–709. <https://doi.org/10.1080/02680939.2012.758825>.
- Grek, S., 2014. OECD as a site of coproduction: European education governance and the new politics of “policy mobilization”. *Crit. Pol. Stud.* 8 (3), 266–281.
- Grek, S., 2017. Socialisation, learning and the OECD’s reviews of national policies for education: the case of Sweden. *Crit. Stud. Educ.* 58 (3), 295–310.
- Grek, S., Maroy, C., Verger, A., 2021. In: *Accountability and Datafication in the Governance of Education*. World Yearbook of Education 2021. Routledge, London.
- Grundmann, R., Stehr, N., 2012. *The Power of Scientific Knowledge*. Cambridge University Press.
- Hopfenbeck, T.N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J.M., Baird, J., 2017. Lessons learned from PISA: a systematic review of peer-reviewed articles on the Programme for International Student Assessment. *Scand. J. Educ. Res.* <https://doi.org/10.1080/00313831.2016.1258726>.
- Lawn, M., Grek, S., 2012. *Europeanizing Education: Governing an Emerging Policy Space*. Symposium Books, Oxford.
- Mausethagen, S., 2013. Governance through concepts: OECD and the construction of “competence” in Norwegian education policy. *Berkley Rev. Educ.* 4 (1), 161–181. <https://doi.org/10.5070/B84110058>.
- Newman, J., Clarke, J., 2009. *Publics, Politics and Power: Remaking the Public in Public Services*. SAGE, London.
- Normand, R., 2016. *The Changing Epistemic Governance of European Education*. Springer.
- Novoa, A., Yariv-Mashal, T., 2003. Comparative research in education: a mode of governance or a historical journey? *Comp. Educ.* 39 (4), 423–438. <https://doi.org/10.1080/030506032000162002>.
- OECD, 2015. *Improving Schools in Scotland*. OECD, Paris.
- OECD, 2018. *Preparing Our Youth for an Inclusive and Sustainable World: OECD PISA Global Competence Framework*. OECD, France. Retrieved from. <https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf>.
- Ozga, J., Dahler-Larsen, P., Segerholm, C., Simola, H. (Eds.), 2011. *Fabricating Quality in Education: Data and Governance in Europe*. Routledge, London.
- Ozga, J., 2019. Governing and knowledge: theorising the relationship. In: Langer, R., Brusemeister, T. (Eds.), *Handbuch Educational Governance Theorien*. Springer.
- Ozga, J., 2020. The politics of accountability. *J. Educ. Change* 21, 19–35. <https://doi.org/10.1007/s10833-019-09354-2>.
- Ozga, J., 2021. Who governs? Political leadership in transnational times. *Sch. Leader. Manag.* 41 (1–2), 6–21. <https://doi.org/10.1080/13632434.2020.1789857>.
- Piattoeva, N., 2014. Elastic numbers: national examinations data as a technology of government. *J. Educ. Pol.* 30 (3), 316–334.
- Rautalin, M., Syväterä, J., Vento, E., 2021. International organisations establishing their scientific authority: periodizing the legitimization of policy advice by the OECD. *Int. Sociol.* 36 (1), 3–24.

- Robertson, S., Dale, R., 2017. Comparing policies in a globalizing world. *Educacao e Realidade* 42 (3), 859–875.
- Ranson, S., 2003. Public accountability in an age of neo-liberal governance. *J. Educ. Pol.* 18 (5), 459–480.
- Sahlberg, P., 2016. The global educational reform movement and its impact on schooling. In: Mundy, K., Green, A., Lingard, B., Verger, A. (Eds.), *The Handbook of Global Education Policy*. Wiley-Blackwell, West Sussex, pp. 128–144.
- Sellar, S., 2015. A strange craving to be motivated. *Deleuze Stud.* 9 (3), 424–436.
- Selwyn, N., 2018. Technology as a focus of education policy. In: Papa, R., Armfield, S.W.J. (Eds.), *The Wiley Handbook of Educational Policy*. John Wiley & Sons, pp. 459–477.
- Shiroma, E.O., 2014. Expert consultants and knowledge production. In: Fenwick, T., Mangez, E., Ozga, J. (Eds.), *Governing Knowledge: Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education*, *World Yearbook of Education* 2014. Routledge, London, pp. 101–113.
- Stone, D., 2013. *Knowledge Actors and Transnational Governance: The Private-Public Policy Nexus in the Global Agora*. Palgrave Macmillan.
- Stone, D., 2019. *Making Global Policy. Elements in Public Policy*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781108661690>.
- Verger, A., Fontdevila, C., Parcerisa, L., 2019. Reforming governance through policy instruments: how and to what extent standards, tests and accountability in education spread worldwide. *Discourse* 40 (2), 248–270. <https://doi.org/10.1080/01596306.2019.1569882>.
- Waldow, F., Steiner-Khamsi, G. (Eds.), 2019. *Understanding PISAs Attractiveness: Critical Analyses in Comparative Policy Studies*. Bloomsbury, London.

Assessment and accountability global trends and future directions in 21st century competencies

Graham S. Maxwell, Institute for Learning Sciences and Teacher Education, Australian Catholic University, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

Introduction	313
Specifying 21st century competencies	313
Domains of knowledge (domains of learning)	314
Information and communication technologies (ICT)-related competencies	314
Conceptual competencies	314
Cognitive competencies	314
Metacognitive competencies	315
Personal (or intrapersonal) competencies	315
Interpersonal competencies	315
International commitments to 21st century competencies	315
Progress in assessing 21st century competencies	315
Assessing ICT-related competencies	316
Assessing conceptual knowledge competencies	317
Cognitive competencies	317
Metacognitive competencies	317
Personal (or intrapersonal) competencies	318
Interpersonal competencies	318
Implications for future assessment practice	319
References	320

Introduction

It is generally accepted that the world is experiencing a transition from an industrial age based on mechanization to an information age based on computerization (Autor et al., 2003; Autor and Price, 2013; Levy and Murnane, 2004, 2013; Reich, 1992; United Nations Conference on Trade and Development, 2017). The reach of computerization is extensive, affecting all aspects of life and work through the development of information and communication technologies (ICTs), replacing many existing human activities, and setting a premium on new competencies (Ananiadou and Claro, 2009; Deming, 2017; Gordon et al., 2009; Joyes et al., 2019; Lake, 2019; Levy and Murnane, 2004; Suto, 2013). This transition requires a reframing of educational aims for the information age, and these are typically represented as 21st century, key or core competencies, covering a range of knowledge, understandings, capabilities, skills and attributes.

Here, “competencies” is preferred to “competences”, though both are in common use and appear to have the same meaning (though “competency” and “competence” would appear to differ). There is also widespread reference to “21st century skills”, though skill would seem to be a narrower concept than competency, and not always applicable

Assessment needs to accommodate these reframed educational aims. Assessment, both in the classroom and for accountability, is recognized as a driver of what is learned and as a signifier of what is valued in learning. Haertel (2013) refers to these as incentive effects and messaging effects. That is, assessment focusses effort on the targeted learning outcomes and it also shapes perceptions of what learning is important and worth attention. The narrowness and shallowness of much educational testing is therefore problematic. Attention to 21st century competencies indicates an emerging awareness that the range of both classroom assessments and accountability assessments needs to be extended (Hill and Barber, 2014).

Specifying 21st century competencies

There is no universally agreed definition of 21st century competencies. A large number of relevant frameworks exists (for example, Ananiadou and Claro, 2009; Bereiter and Scardamalia, 2012; Binkley et al., 2012; Birenbaum, 1996; Chalkiadaki, 2018; Conley, 2003, 2005, 2010, 2011; Cook and Weaving, 2012; de Corte, 2010; Dumont and Istance, 2010; European Union, 2019; Gordon et al., 2009; Partnership for 21st Century Skills, 2019; Pellegrino and Hilton, 2012; Voogt and Roblin, 2012; Wagner, 2008; Wang

et al., 2018; World Education Forum, 2015). While these frameworks reveal differences of perspective, terminology and emphasis, there are substantial conceptual similarities, and a collated framework can be constructed, consisting of the following domains (Maxwell, 2021, pp. 80–85):

- Domains of knowledge (or domains of learning, or substantive learning)
- Information and communication technologies (ICT)-related competencies
- Conceptual competencies
- Cognitive competencies
- Metacognitive competencies
- Personal (or Intrapersonal) competencies
- Interpersonal competencies.

Each of these domains is made up of a variety of components or dimensions indicating different characteristics of the domain. These components can be represented as separate learning aims and assessable competencies.

Domains of knowledge (domains of learning)

Knowledge domains or disciplines remain the main organizers of curriculum in most countries and are seen as foundational for all 21st century competencies (Australian Curriculum, Assessment and Reporting Authority, 2019; Binkley et al., 2012; European Union, 2019). Knowledge domains typically include the arts and humanities and science-technology-engineering-mathematics (STEM), as well as other specialized areas of knowledge (Partnership for 21st Century Learning, 2019; UNESCO, 2013), and sometimes also literacy and numeracy. While domains of knowledge are not themselves usually conceived as competencies, they are the sites where most competencies are realized, and need to be mentioned for this reason. Some 21st century competencies assume specific characteristics within each knowledge domain; others may be cross-curriculum (“transversal”), that is, perhaps developed within several knowledge domains but transferable to other situations with a different knowledge base.

Information and communication technologies (ICT)-related competencies

These are sometimes seen as defining a knowledge domain in itself, or otherwise as cross-curriculum competencies. Three sub-domains can be identified:

- ICT literacy (technical skills for accessing and using digital technologies and communication tools)
- Information literacy (capabilities for accessing, managing, evaluating and using information)
- Technological literacy (understanding the ways in which technologies can be used to solve complex problems)

Conceptual competencies

Different facets of knowledge can be categorized as factual, conceptual and procedural (Krathwohl, 2002). Factual knowledge encompasses the terms, definitions and elements that define a field of knowledge. Conceptual knowledge encompasses understanding and includes theories, principles, ideas and ways in which facts can be organized and related. Both of these are concerned with knowing “what”. Procedural knowledge encompasses techniques and procedures, that is, knowing “how”. In the classification of competencies, factual knowledge is not included, perhaps taken for granted as foundational within domains of knowledge. In the classification of competencies here, conceptual knowledge competencies and cognitive competencies offer a different division and elaboration of conceptual and procedural knowledge. The main issue is that, while factual knowledge provides the “component parts”, their integration and application is essential for building a meaningful and useful knowledge “engine”.

Conceptual competencies represent desirable qualities for conceptual knowledge, and typically include:

- conceptual depth and complexity
- structuring and interconnection of key concepts
- working with abstractions (connecting theory and practice through abstract modeling)
- transferability (capability of applying knowledge learned in one context to another context).

Cognitive competencies

These may also infuse the knowledge domains or else be treated as cross-curriculum competencies. They typically include:

- higher-order thinking skills (understanding, analyzing, interpreting, evaluating, generalizing, organizing, structuring)
- problem solving and problem formulation (hypothesizing, strategizing, experimenting, investigating)
- critical thinking and reasoning (searching for relevant information, selecting relevant information, being systematic and logical, weighing evidence, identifying errors, constructing defensible arguments, making informed decisions)
- constructive and systematic inquiry and knowledge creation (posing questions, seeking answers, planning, investigating, producing, applying and extending)

- information management and retrieval (searching, selecting, evaluating, organizing, drawing conclusions and communicating findings)
- communication skills (oral, written and multimedia).

Metacognitive competencies

Metacognition is thinking about thinking, especially managing one's own learning, sometimes identified as a fourth facet of knowledge (Krathwohl, 2002). Metacognitive competencies typically include:

- self-reflection, self-evaluation and self-adjustment
- self-regulation, self-management and self-direction
- self-awareness and self-efficacy
- capabilities for life-long-learning (learning-to-learn).

Personal (or intrapersonal) competencies

These are usually considered as attributes, dispositions or propensities, including:

- adaptability and flexibility
- creative endeavor, enterprise, initiative and entrepreneurship
- curiosity and imagination resilience and persistence
- civic mindedness.

Whether, or to what extent, these can be taught is unclear (Bereiter and Scardamalia, 2012), and this has implications for their assessment.

Interpersonal competencies

These are based on a recognized need in the modern world for empathy and respect in accommodating different opinions, lifestyles and cultures, which can also be beneficial in solving problems. They include:

- collaboration and teamwork
- social, cultural and intercultural awareness
- leadership, interpersonal support and conflict management.

International commitments to 21st century competencies

There is widespread "in principle" commitment around the world for students to develop these kinds of competencies (Australian Curriculum, Assessment and Reporting Authority, n.d., 2019; Ananiadou and Claro, 2009; Care et al., 2016a, 2018; Cheng, 2017; Cook and Weaving, 2012; European Commission, 2020; Gordon et al., 2009; Russell, 2016; Taylor et al., 2020; and UNESCO, 2019). For example, in a scan of 102 countries, Care et al. (2016) concluded that most countries "acknowledge the importance of breadth of skills somewhere in their public documents, and many list a number of specific skills or competencies that the education system intends to provide" (p. 4). Even so, "few countries show evidence that these skills are consistently identified in national documents, including mission and vision statements, curriculum frameworks, and descriptions of how skills progress for learners" (p. 4). Commitment has been largely verbal rather than practical.

Declarations of intent are clearly insufficient to create system-wide change in educational practice. Important requirements for change include system-wide commitment, development of resources and professional development. Even more critical is producing operational definitions of the competencies and developing sound principles and procedures for their assessment. These are matters on which a start has been made but progress has been slow (Care, 2018; Care et al., 2019; Russell, 2016; Saavedra and Opfer, 2012).

Progress in assessing 21st century competencies

Commitment to 21st century competencies around the world can be summarized as an emphasis on

substantive knowledge in (at least) a curriculum core; knowledge having structure, complexity and transferability; cognitive strategies covering a range of thinking, enquiry and communication skills; metacognitive strategies relating to self-direction, monitoring and regulation; personal attributes such as adaptability, creativity, curiosity and persistence; and interpersonal skills such as collaboration and cultural awareness.

Maxwell (2021, p. 97).

While these competencies are not newly discovered (Kereluik et al., 2013), what is different is their importance and applicability from an educated elite to everyone (Kay, 2010; Pellegrino and Hilton, 2012). They have also acquired new contextual characteristics and changes in their relative importance. For example, “problem solving” now requires accessing, managing, integrating and evaluating large amounts of “disorderly” information, while “interpersonal skills” such as collaboration now become more salient than previously (Dede, 2010).

It is sometimes suggested that specific competencies may be assessable using current assessment approaches; for example, the Program for International Student Assessment [PISA] includes complex assessment of critical thinking (OECD, 2018). However, this may misrepresent the transformative implications of the competencies (Care, 2018; Care and Luo, 2016; Cook and Weaving, 2012; Pepper, 2011; Silva, 2008). Voogt and Roblin (2012) suggest that: “current assessment models, which mostly focus on the measurement of discrete “bits” of knowledge, fail to assess 21st century competences and call for assessment grounded in authentic and complex tasks” (p. 312). The need for “real world” or “authentic” assessment contexts is a common concern (Care and Kim, 2018; Care et al., 2018; Lai and Viering, 2012).

The use of complex assessment tasks also recognizes the practical impossibility of assessing all of the competency categories and sub-categories separately. Rather, assessment needs to focus on constellations of competencies operating synergistically. Examples include literacy tasks in combination with critical thinking and self-management skills, problem-solving tasks that involve communication and collaboration, portfolios (as collections of student work), presentations and constructive activities, and problems involving computer simulations of real-world scenarios (Care et al., 2016; Koenig, 2011; Microsoft Partners in Learning, 2012; Soland et al., 2013).

A further complication is whether competencies are embedded within domains of learning or can be cross-curriculum (transversal). The latter assumes generalization of the competency beyond the confines of particular subject matter, that is, that the competency is generic, and can transfer to contexts beyond those in which it is initially acquired. This remains a contentious issue (Shepard, 2019).

While progress has been slow and piecemeal, developments in assessing 21st century competencies can be identified for each of the domains. The following discussion identifies several key developments around the world.

Assessing ICT-related competencies

Siddiq et al. (2016) conducted a substantial international review of existing ICT literacy assessments. They found substantial convergence in emphasis but considerable ambiguity in definitions of ICT literacy (otherwise referred to as digital literacy). They chose a well-researched and comprehensive framework, DIGCOMP (Ferrari, 2013), as the basis for their review, with modifications. The revised framework consisted of six competence areas (here paraphrased): browsing, searching, filtering, evaluating, storing and retrieving information; communication involving interacting, sharing, engaging and collaborating through digital technologies; digital content creation; problem solving through innovation and collaboration; technical and operational skills; and safety management and ethics. A total of 38 tests from 17 countries was analyzed. They note that Australia was the only country to systematically embed ICT literacy in the curriculum and assess it at a national level, in this case, through sampling testing every third year at two grade levels (6 and 10); the Australian national curriculum includes ICT Literacy as one of several general capabilities (ACARA, n.d) with relevant content descriptions and elaborations in each learning area of the curriculum, and rubrics detailing several elements and sub-elements across a learning continuum of six levels [<https://www.australiancurriculum.edu.au/media/1074/general-capabilities-information-and-communication-ict-capability-learning-continuum.pdf>]

In their findings, Siddiq et al. (2016) conclude that the complexity of the ICT literacy concept is not yet matched by commensurate complexity in its assessment, indicating “a gap between how ICT literacy is theorized in research and dominant frameworks, and how the concept is operationalized in curriculum and assessment” (p. 77). They report a majority focus on handling digital information, with some attention to digital content creation, communication and technical/operational skills, but little attention to other aspects of ICT Literacy, such as “problem solving with ICT, digital communication, and online collaboration” (p. 78). They also found a preponderance of traditional testing, with limited use of authentic tasks, such as would be relevant, for example, for assessing collaborative problem solving.

One framework, taking up the authentic assessment challenge, charted ICT literacy across four strands for Computer and Information Literacy (CIL)—understanding computer use; gathering information; producing information; and digital information—and two strands for Computational Thinking (CT)—conceptualizing problems; and operationalizing solutions. Other developments are reported by Ainley et al. (2016), Fraillon et al. (2019), Kaarakainen et al. (2018), Siddiq et al. (2017), van Deursen et al. (2016) and Wilson et al. (2015, 2018a,b).

It can be noted that ICT-related competencies include other competencies, such as cognitive (e.g., problem solving) and interpersonal (e.g., collaboration) competencies, in the context of ICT. This highlights the interrelationships among various 21st century competencies, and the possibilities for combining two or more different competencies (from the different categories listed earlier) in a single assessment. ICT-related competencies are typically represented as cross-curriculum (generic or transversal) competencies, but they may be learned through application within specific domains of knowledge.

Assessing conceptual knowledge competencies

Conceptual knowledge is located within domains of knowledge because it represents desirable cognitive dimensions of the subject matter. To become expert in a domain of knowledge is to develop conceptual depth and complexity, going beyond factual knowledge to include knowledge structures and their application in solving authentic problems. Factual knowledge is “surface learning”, which can be usefully assessed as foundational for deeper learning, but is inadequate as the ultimate aim of domain learning (Hattie, 2009).

Assessment of domain knowledge therefore needs to go beyond assessment of routine memorization. Binkley et al. (2012) say:

The crux of twenty-first century skills is the need to integrate, synthesize, and creatively apply content knowledge in novel situations. Consequently, twenty-first century assessments must systematically ask students to apply content knowledge to critical thinking, problem solving, and analytical tasks throughout their education, so that we can help them hone this ability and come to understand that successful learning is as much about the process as it is about facts and figures (p. 25).

Consequently, assessment needs to extend beyond factual recall and simple understanding to reveal deeper conceptual understanding and knowledge structures, something that typical standardized tests cannot do (Sawyer, 2006). Assessments that can assess such competencies are well known but require further development. They include performance assessments, projects, presentations and portfolios, focusing on the deep structure of the knowledge domain and its application in authentic situations (Maxwell, 2021). The recommendations of Lai and Viering (2012) still hold true, especially the need for: open-ended and ill-structured tasks; meaningful, authentic or real-world tasks; making student thinking and reasoning visible; and innovative use of technology. Interesting examples can be found in curriculum materials around the world, for example, *Australian Curriculum Sample Assessments and Senior Assessment Hub* [Queensland Curriculum and Assessment Authority: <https://www.qcaa.qld.edu.au>], *New Zealand Assessment Resource Banks* [New Zealand Council for Educational research: <http://arbs.nzcer.org.nz>].

Cognitive competencies

Cognitive competencies are about thinking, reasoning, inquiry and problem solving. The extent to which these competencies can be learned and assessed generically, that is, applicable in any context, is contentious. Typically, in education systems, these competencies are embedded as learning aims and assessable components of subject domains (Leutner et al., 2017). Clearly, problem solving in different domains, for example, history versus science, has different characteristics, and capability in one domain may not transfer to another domain. A knowledge base is essential for the thinking. For assessment of generic competencies, the relevant knowledge base can be provided as part of the task specifications, but competence may not then transfer to more realistic situations (Ercikan and Oliveri, 2016).

The “complex problem solving” component of cognitive competencies has been given substantial attention and exploration in recent years (Krkvic et al., 2018). Yet, Funke et al. (2018) conclude that its definition and assessment remain inadequate and require more research. One approach that has shown some promise for large-scale testing is reported by Herde et al. (2016). Their approach involves computer-based tasks that set a problem stemming from provided information. They illustrate their explanation with a task, *Climate Control*, from PISA 2012 (OECD, 2013a). The approach allows assessment of both knowledge acquisition and knowledge application as two fundamental stages of problem solving. A potential advantage of computer-based tasks is that the information available from the computer log-files can provide insights into the strategies that students used in solving the problem, something about which little is currently known.

A different approach was taken by Griffin and Care (2015), focusing on collaborative problem solving (combining interpersonal and cognitive components) and using ICT as a tool for group collaboration rather than as a tool for solving the problem (see also Care et al., 2016b). Examples of such tasks were developed, and scored in terms of various cognitive and collaborative skills. Some success was experienced in implementing the assessments across the six countries involved in the research.

Metacognitive competencies

The long history of research on metacognition (Azevedo, 2020; Lai, 2011) has not led to its systematic assessment in educational practice. Assessment of metacognitive competency typically relies on self-report to reveal the thinking processes, thinking strategies, self-awareness, self-monitoring and self-perceptions of the student. This can involve questionnaires, observation schedules and interview protocols (Karlen, 2016; Kuzle, 2018; Marulis et al., 2016; Roebers et al., 2020; Ozturk, 2017; Rovers et al., 2019; Tuck and Moxley, 2017). One useful strategy is asking students to “think aloud” during or after completing an assessment task to reveal the strategies used (Bannert and Mengelkamp, 2007; Desoete, 2008). Such approaches are more useable in classroom contexts than large-scale testing. Their veridicality is often doubtful. It could be argued that successful metacognitive strategies should be implicit in the successful completion of tasks that require the use of such strategies. Assessment of metacognitive competence is therefore more important as formative assessment than as summative assessment.

A substantial analysis reported by Caena (2019) combines metacognitive with personal and interpersonal competencies. This combination is expressed as the Personal, Social and Learning-to-Learn Competency (LifeComp), one of eight competencies

adopted by the European Union (European Commission, 2018; European Union, 2019). A framework was proposed consisting of “cross-cutting core elements (self-regulation, growth mindset, empathy), which support compound competences for personal development (adaptability and wellbeing), social development (communication and collaboration), and learning to learn (critical thinking and managing learning)” (Caena, 2019, p. 32). This framework provides a useful foundation for further development of assessment processes.

Personal (or intrapersonal) competencies

Calls for the development of personal competencies (Herman and Hilton, 2017; Pelligrino and Hilton, 2012) have not seen much progress. Duckworth and Yeager (2015) analyze the inadequacies of typical assessments of personal competencies, such as self-report surveys, teacher/parent-report surveys, and teacher observations for program evaluation, school accountability and individual diagnosis. Surveys suffer from reference bias, social desirability bias, acquiescence bias, and faking; observations reference exhibited behaviors rather than the targeted internal emotions and thoughts, and these behaviours may not generalize to other tasks and other contexts. Nevertheless, development of a suite of assessment approaches for use in the classroom would assist teachers in focusing on the development of these competencies.

One such approach, the Mission Skills Assessment, was developed for the Independent Schools Exchange (INDEX) by Educational Testing Service (ETS) (Petway et al., 2016). This was a web-based assessment of six constructs, covering personal and interpersonal competencies: creativity; curiosity; ethics; resilience; teamwork; and time management. Despite the admitted inadequacies, six Likert-item scales were constructed for use in self-report and teacher-report surveys. Soland et al. (2013) report the usefulness of this approach for drawing attention to these competencies and creating conversations in classrooms about them, but there does not seem to have been further development of these scales. An alternative that has received some attention is Situational Judgment Tests (Herde et al., 2019).

A distinction can be made between a skill and a quality, for example, critical thinking versus critical thinker (Bereiter and Scardamalia, 2012). Critical thinking involves logical evaluation of arguments; becoming a critical thinker requires a deeper level of commitment and engagement. In other words, critical thinking is a cognitive competency while being a critical thinker is a personal quality. Critical thinking capability might be assessed by a test consisting of multiple-choice items, but whether a person is developing into a critical thinker requires judgment by an observer over time.

Creativity is a related personal quality. Lucas (2016) reports a promising model of creativity in terms of five habits of mind: inquisitive; imaginative; persistent; collaborative; and disciplined. Assessment trials of the inquisitive dimension were reported, assessing students on strength, breadth and depth along a four-step progression from “awakening” to “adept”. This is a promising set of ideas, further developed in Lucas and Venckutė (2020) and Venckutė et al. (2020), who review international approaches to teaching and assessing creativity.

Other developments are difficult to find. Bolden et al. (2020) surveyed developments in assessing creativity, focusing on its formative development, and reveal some of the complexities and confusions about creativity and its assessment, pointing to the need for greater construct consistency and clarity. They conclude that “assessment strategies that include (a) explicitly establishing and defining criteria for creativity and creative processes, behaviors, and products; and (b) engaging students in supported and particularly teacher-supported self-assessment emerged as positive supports for creativity in classrooms” (p. 369).

Interpersonal competencies

The Rand Education Assessment Finder (<https://www.rand.org/education-and-labor/projects/assessments.html>) is a web-based repository of assessments of interpersonal, (intra)personal and cognitive competencies. Its construction is described in Schweig et al. (2018). This shows that a large number of such assessments exist, evidencing a variety of different constructs of relative quality and value.

There is a long history of interest in social and emotional learning, as a combination of interpersonal and (intra)personal competencies (Jones et al., 2019). Berg et al. (2017) found substantial diversity in how these competencies are represented and defined. In order to impose some coherence, their study used a classification of competencies organized into six domains: cognitive regulation (e.g., attention control, critical thinking); emotional processes; interpersonal processes; values (e.g., civic, ethical); personal perspectives (e.g., optimism, openness); and self-identify (e.g., efficacy, esteem) with explanations of each (p. 43). Their review classified a range of existing measures in terms of their characteristics. However, the appropriateness and usability of different measures for different contexts needs to be considered.

At an international level, OECD has committed to conducting a regular survey of social and emotional skills (defined as how people manage their emotions, perceive themselves and engage with others) targeting 10 and 15 year olds and using 5-point Likert items. The assessment covers 17 skills across five domains: task performance; emotional regulation, collaboration; open-mindedness; and engaging with others (<https://www.oecd.org/education/ceri/social-emotional-skills-study/>). The first cycle of assessment (2018–20) has been reported (OECD, n.d), including a discussion of the development of self-efficacy, metacognition and critical thinking, and their interrelationships.

Another dimension of interpersonal competency receiving substantial attention is collaboration. Child and Shaw (2016) provide an overview of assessment of collaboration, defined in terms of six facets: social interdependence; introduction of new ideas; cooperation; conflict resolution; sharing of resources; and communication. They make an important distinction between

the collaborative process (the quality of the team activity) and the collaborative outcome (the quality of the knowledge and understanding produced for solution of the problem).

Collaboration is typically, almost inevitably, linked with problem solving. [Care et al. \(2016\)](#) elaborate both strands in collaborative problem solving, with collaboration consisting of three social components—participation, perspective taking and social regulation—and problem solving consisting of two cognitive processes—task regulation and knowledge building. Contributing sub-skills for each of these five components enable design of assessment tasks and mapping the progression from lower to higher competency. Construction and implementation of an online task requiring collaborative problem solving between two students is described. The authors emphasize the inadequacy of traditional methods of assessment for such competencies and the need for “innovative and advanced” approaches of assessment if they are to become part of the expected learning outcomes in education systems.

Implications for future assessment practice

The implications of a commitment to 21st century competencies extend not just to developing new assessments for those competencies but to reframing the way in which assessment is designed and implemented in schools. For example, assessment needs to incorporate application of knowledge and processes in different contexts, access to and use of information resources, and collaborative as well as individual problem solving ([Wilson et al., 2015](#)). This requires assessments that are situated in authentic real-life contexts, assessments requiring multiple steps and a chain of reasoning, assessments that incorporate several competencies, and assessments with open response formats ([Care and Kim, 2018](#); [Cook and Weaving, 2012](#); [Voogt and Roblin, 2012](#)). Progress in the development of such assessments has been slow and uncoordinated. Most assessments in use around the world remain traditional and narrow, limited to simple recall and routine tasks, with limited attention to most of the typically endorsed 21st century competencies ([Darling-Hammond, 2017](#)). While the complexities are substantial, examples of innovative approaches are beginning to emerge (for example, [Care et al., 2018](#)).

There remains a problem of definition of various competencies—giving a clear and precise account of the relevant construct and the learning goals indicative of each competency ([Care and Luo, 2016](#); [Care et al., 2019](#); [Lucas, 2021](#)). Preferably, too, levels of progression or stages of development are needed to represent increasing complexity in the construct, indicating greater depth of learning rather than greater quantity ([Bereiter and Scardamalia, 2012](#); [Forster and Masters, 2010](#); [Lucas, 2016](#); [Jin et al., 2019](#); [Shepard, 2018](#)).

Once the constituent components have been determined, assessment tasks can be constructed. Given the synergistic nature of competencies, such assessment tasks need to cover constellations of competencies ([Care et al., 2016](#)). However, task development may not be the greatest challenge. [Care et al. \(2020\)](#), working with teams from Cambodia, Mongolia and Nepal on assessment of problem solving, critical thinking and collaboration, found that the bigger problem was the adaptation needed in teaching practice and classroom culture. Both philosophical and logistic issues were involved. Commitment to change required teachers being convinced it was worthwhile. It was concluded that an evolutionary approach, based on building teacher capacity and confidence, has most chance of success.

The prospect of this happening quickly appears slim. For political and cultural reasons, national education systems are firmly wedded to existing practices and are difficult to change. As [Issacs and Gorgen \(2018\)](#) note:

Educational cultures differ across jurisdictions, permeating their assessment structures and processes in idiosyncratic ways. Different cultures and contexts give rise to a variety of accepted practices in national assessment and examination systems across the world and often act as impediments to change (p. 308).

Emancipation from existing practices therefore requires different strategies in different countries, for example breaking the monopoly of standardized testing in the US versus breaking the monopoly of public examinations in many other countries. Comprehensive calls for change in the US, such as the “next generation assessment” proposals of [Darling-Hammond \(2014\)](#) and [Conley \(2018\)](#), have seen some take-up and some backsliding. In other parts of the world, the need for change in public examination systems is recognized but requires careful management of potential political and practical repercussions ([Baird et al., 2018](#); [Kellaghan and Greaney, 2020](#)). A standout case is Queensland Australia where all high-school graduation assessment was school-based for fifty years, allowing for the development of broader and deeper assessment ([Cumming, 2020](#); [Maxwell and Cumming, 2011](#)), but which has succumbed recently to pressures for national conformity and the reintroduction of public examinations ([Campbell, 2018](#)). As well, Australian national testing of literacy and numeracy in Years 3, 5, 7 and 9 has had typical narrowing effects on the curriculum through earlier years ([Cumming et al., 2019](#)).

Establishing a link between national aspirations for broader learning, as expressed in 21st century competencies, and assessments at all levels of an educational system needs an holistic approach in which “various components of assessment and evaluation should form a coherent whole ... generate synergies between components, avoid duplication and prevent inconsistency of objectives” ([OECD, 2013b](#), p. 2). Ensuring that all relevant learning aims are given attention at appropriate levels within an educational system requires a systems approach. In a systems approach, assessment is planned as a whole, ensuring appropriate coverage and

balance across all components of the education system (state/district, school and classroom; year levels; subjects or areas of study). This allows the overall system to be coherent (so that different components are part of an overall plan), comprehensive (so that the components contribute toward coverage of all relevant learning outcomes), and continuous (so that there is systematic developmental progression from earlier assessments to later assessments) (Maxwell, 2021). This also establishes a link and complementarity between accountability assessment and classroom assessment.

Judging progress in learning on only some of the learning aims that are considered important, whether the unit of analysis is the nation/state, district, school, classroom or student, provides incomplete and distorted information. What is needed is a profile approach to monitoring and reporting progress, where the different learning outcomes are considered as separate dimensions, and all aspects of learning are recognized and valued. Further discussion of these issues is found in chapters 3 and 5 of Maxwell (2021).

References

- Ainley, J., Fraillon, J., Schulz, W., Gebhardt, E., 2016. Conceptualizing and measuring computer and information literacy in cross-national contexts. *Appl. Meas. Educ.* 2, 291–309. <https://doi.org/10.1080/08957347.2016.1209205>.
- Ananiadou, K., Claro, M., 2009. 21st Century Skills and Competences for New Millennium Learners in OECD Countries. OECD Working Papers 41. OECD Publishing, Paris. <https://doi.org/10.1787/218525261154>.
- Australian Curriculum, Assessment and Reporting Authority, (n.d.). About the Australian Curriculum. Melbourne: ACARA. <https://www.australiancurriculum.edu.au/about-the-australian-curriculum/>.
- Australian Curriculum, Assessment and Reporting Authority, 2019. Program of Research: Key Findings From Four International Comparative Studies. ACARA, Melbourne. <https://www.australiancurriculum.edu.au/media/4394/ac-bcc-fncc-sc-nzc-key-findings.pdf>.
- Autor, D.H., Levy, F., Murnane, R.J., 2003. The skill content of recent technological change: an empirical exploration. *Q. J. Econ.* 118, 1279–1333. <https://doi.org/10.1162/003355303322552801>.
- Autor, D.H., Price, B., 2013. The Changing Task Composition of the US Labor Market. MIT Economics, Cambridge, MA. <http://economics.mit.edu/files/9758>.
- Azevedo, R., 2020. Reflections on the field of metacognition: issues, challenges, and opportunities. *Metacogn. Learn.* 1, 91–98. <https://doi.org/10.1007/s11409-020-09231-x>.
- Baird, J.-A., Isaacs, T., Opposs, D., Gray, L. (Eds.), 2018. Examination Standards: How Measures and Meanings Differ Around the World. University College London, Institute of Education Press, London.
- Bannert, M., Mengelkamp, C., 2007. Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted: does the verbalisation method affect learning? *Metacogn. Learn.* 3, 39–58. <https://doi.org/10.1007/s11409-007-9009-6>.
- Bereiter, C., Scardamalia, M., 2012. What Will It Mean to Be an Educated Person in the 21st Century? The Gordon Commission on the Future of Assessment in Education, Washington, DC. https://www.ets.org/research/policy_research_reports/gordon_commission?WT.ac=research_policyresearchreports_gordoncommission_180223.
- Berg, J., Osher, D., Same, M.R., Nolan, E., Benson, D., Jacobs, N., 2017. Identifying, Defining, and Measuring Social and Emotional Competencies: Final Report. American Institutes for Research, Washington, DC. <https://www.air.org/sites/default/files/downloads/report/Identifying-Defining-and-Measuring-Social-and-Emotional-Competencies-December-2017-rev.pdf>.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., Rumble, M., 2012. Defining twenty-first century skill. In: Griffin, P., McGaw, B., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht, pp. 17–66. https://doi.org/10.1007/978-94-007-2324-5_2.
- Birenbaum, M., 1996. Assessment 2000: towards a pluralistic approach to assessment. In: Birenbaum, M., Dochy, F.J.R.C. (Eds.), *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge*. Springer, New York, pp. 3–29. https://doi.org/10.1007/978-94-011-0657-3_1.
- Bolden, B., DeLuca, C., Kukkonen, T., Roy, S., Wearing, J., 2020. Assessment of creativity in K–12 education: a scoping review. *Rev. Educ.* 8 (2), 343–376. <https://doi.org/10.1002/rev3.3188>.
- Caena, F., 2019. Developing a European Framework for the Personal, Social and Learning to Learn Competence (LifeComp): Literature Review and Analysis of Frameworks. Publications Office of the European Union, Luxembourg. <https://publications.jrc.ec.europa.eu/repository/handle/JRC117987>.
- Campbell, M., 2018. Standard setting in Queensland: the Queensland certificate of education. In: Baird, J.-A., Isaacs, T., Opposs, D., Gray, L. (Eds.), *Examination Standards: How Measures and Meanings Differ Around the World*. University College London, Institute of Education Press, London.
- Care, E., 2018. Twenty-first century skills: from theory to action. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer, Cham, Switzerland, pp. 3–17. https://doi.org/10.1007/978-3-319-65368-6_1.
- Care, E., Anderson, K., Ki, H., 2016a. Visualizing the Breadth of Skills Movement Across Education Systems: Skills for a Changing World. The Brookings Institution, Washington, DC. <https://www.brookings.edu/research/visualizing-the-breadth-of-skills-movement-across-education-systems/>.
- Care, E., Kim, H., 2018. Assessment of 21st century skills: the issue of authenticity. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer, Cham, Switzerland, pp. 21–39. <https://doi.org/10.1007/978-3-319-65368-6>.
- Care, E., Kim, H., Vista, A., Anderson, K., 2018. Education System Alignment for 21st Century Skills: Focus on Assessment. The Brookings Institution, Washington, DC. <https://www.brookings.edu/wp-content/uploads/2018/11/Education-system-alignment-for-21st-century-skills-012819.pdf>.
- Care, E., Luo, R., 2016. Assessment of Transversal Competencies: Policy and Practice in the Asia-Pacific Region. UNESCO, Bangkok. <https://bangkok.unesco.org/content/assessment-transversal-competencies-policy-and-practice-asia-pacific-region>.
- Care, E., Scoular, C., Griffin, P., 2016b. Assessment of collaborative problem solving in education environments. *Appl. Meas. Educ.* 29 (4), 250–264. <https://doi.org/10.1080/08957347.2016.1209204>.
- Care, E., Vista, A., Kim, H., 2019. Assessment of Transversal Competencies: Current Tools in the Asian Region. UNESCO, Bangkok. <https://bangkok.unesco.org/content/assessment-transversal-competencies-current-tools-asian-region>.
- Care, E., Vista, A., Kim, H., 2020. Optimizing Assessment for All: Focus on Asia. The Brookings Institution, Washington, DC.
- Cheng, K.-M., 2017. Advancing 21st Century Competencies in East Asian Education Systems. Asia Society, Center for Global Education, New York. <https://asiasociety.org/files/21st-century-competencies-east-asian-education-systems.pdf>.
- Chalkiadaki, A., 2018. A systematic literature review of 21st century skills and competencies in primary education. *Int. J. Instruct.* 3, 1–16. <https://doi.org/10.12973/iji.2018.1131a>.
- Child, S., Shaw, S., 2016. Collaboration in the 21st century: implications for assessment. *Camb. Assess. Res. Matter.* 22. <http://www.cambridgeassessment.org.uk/research-matters/>.
- Conley, D.T., 2003. Understanding University Success. University of Oregon, Center for Educational Policy Research, Eugene, OR. <https://www.math.mnr.net/standards/SandS%20Introduction%20and%20Mathematics.pdf>.
- Conley, D.T., 2005. College Knowledge: What It Really Takes for Students to Succeed and What We Can Do to Get Them Ready. Jossey Bass, San Francisco, CA.
- Conley, D.T., 2010. College and Career Readiness: Helping All Students Succeed Beyond High School. Jossey Bass, San Francisco, CA. <https://doi.org/10.1002/9781118269411>.
- Conley, D.T., 2011. Redefining College Readiness, vol. 5. Educational Policy Improvement Center, Eugene, OR. <https://www.ct.edu/files/pdfs/p20/RedefiningCollegeReadiness.pdf>.

- Conley, D.T., 2018. *The Promise and Practice of Next Generation Assessment*. Harvard Education Press, Cambridge, MA.
- Cook, R., Weaving, H., 2012. Key Competence Development in School Education Europe: KeyCoNet's Review of the Literature: A Summary. European Schoolnet, Brussels. http://keyconet.eun.org/c/document_library/get_file?uuid=3a7a093c-4c8f-473c-8702-f38ed86bb730&groupId=11028.
- Cumming, J.J., 2020. Senior secondary school assessment and standards-setting in Queensland, Australia: social context and paradigmatic change. *Assess Educ. Princ. Pol. Pract.* 27 (2), 160–177. <https://doi.org/10.1080/0969594X.2019.1684877>.
- Cumming, C., Jackson, C., Day, C., Maxwell, G., Adie, L., Lingard, B., Haynes, M., Heck, E., 2019. *The 2018 Queensland NAPLAN Review: School and System Perceptions Report and Literature Review*. Queensland Department of Education, Brisbane. <https://qed.qld.gov.au/programs/initiatives/education/Documents/naplan-2018-school-perceptions-report.pdf>.
- Darling-Hammond, L., 2014. *Next Generation Assessment: Moving Beyond the Bubble Test to Support 21st Century Learning*. Jossey-Bass, San Francisco, CA.
- Darling-Hammond, L., 2017. *Developing and Measuring Higher-Order Skills: Models for State Performance Systems*. Council of Chief State School Officers, Washington, DC. <https://learningpolicyinstitute.org/product/models-state-performance-assessment-systems-report>.
- de Corte, E., 2010. Historical developments in the understanding of learning. In: Dumont, H., Istancé, D., Beavides, F. (Eds.), *The Nature of Learning: Using Research to Inspire Practice*. OECD, Paris, pp. 35–67. <https://doi.org/10.1787/9789264086487-en>.
- Dede, C., 2010. Comparing frameworks for 21st century skills. In: Bellanca, J.A., Brandt, R. (Eds.), *21st Century Skills: Rethinking How Students Learn*. Solution Tree Press, Bloomington, IN, pp. 51–76.
- Deming, D.J., 2017. The growing importance of social skills in the labor market. *Q. J. Econ.* 132 (4), 1593–1640. <https://doi.org/10.1093/qje/qjx022>.
- Desoete, A., 2008. Multi-method assessment of cognitive skills in elementary school children: how you test is what you get. *Metacogn. Learn.* 3, 189–206. <https://doi.org/10.1007/s11409-008-9026-0>.
- Duckworth, A., Yeager, D.S., 2015. Measurement matters: assessing personal qualities other than cognitive ability for educational purposes. *Educ. Res.* 44 (4), 237–251. <https://doi.org/10.3102/0013189X15584327>.
- Dumont, H., Istancé, D., 2010. Analysing and designing learning environments for the 21st century. In: Dumont, H., Istancé, D., Beavides, F. (Eds.), *The Nature of Learning: Using Research to Inspire Practice*. OECD, Paris, pp. 19–34. <https://doi.org/10.1787/9789264086487-en>.
- Ercikan, K., Oliveri, M.E., 2016. In search of validity evidence in support of the interpretation and use of assessments of complex constructs: discussion of research on assessing 21st century skills. *Appl. Meas. Educ.* 29 (4), 310–318. <https://doi.org/10.1080/08957347.2016.1209210>.
- European Commission, 2018. *Commission Staff Working Document Accompanying Proposal for a Council Recommendation on Key Competences for Lifelong Learning*. European Commission, Brussels. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018SC0014>.
- European Commission, 2020. *Supporting Key Competence Development: Learning Approaches and Environments in School Education: Conference Report*. European Commission, Brussels. <https://op.europa.eu/en/publication-detail/-/publication/37cc844f-4248-11ea-9099-01aa75ed71a1/language-en/format-PDF/source-search>.
- European Union, 2019. *Key Competencies and Lifelong Learning*. European Union, Luxembourg. <https://op.europa.eu/en/publication-detail/-/publication/297a33c8-a1f3-11e9-9d01-01aa75ed71a1/language-en>.
- Ferrari, A., 2013. *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*. European Union, Luxembourg. <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC83167/lb-na-26035-enn.pdf>.
- Forster, M., Masters, G., 2010. Progression and assessment: developmental assessment. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*. Elsevier, Oxford, pp. 369–377. <https://doi.org/10.1016/B978-0-08-044894-7.00354-7>.
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., Friedman, T., 2019. *IEA International Computer and Information Literacy Study 2018 Assessment Framework*. Springer, Cham, Switzerland. <https://doi.org/10.1007/978-3-030-19389-8>.
- Funke, J., Fischer, A., Holt, D.V., 2018. Competencies for complexity: problem solving in the 21st century. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer, Cham, Switzerland, pp. 41–53. https://doi.org/10.1007/978-3-319-65368-6_3.
- Gordon, J., Halasz, G., Krawczyk, M., Leney, T., Michel, A., Pepper, D., Putkiewicz, E., Wiśniewski, J., 2009. *Key Competences in Europe: Opening Doors for Lifelong Learners Across the School Curriculum and Teacher Education*. Center for Social and Economic Research, Warsaw. <https://doi.org/10.2139/ssrn.1517804>.
- Griffin, P., Care, E. (Eds.), 2015. *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer, Dordrecht, pp. 85–104. https://doi.org/10.1007/978-94-017-9395-7_2.
- Haertel, E.H., 2013. How is testing supposed to improve schooling? *Measurement* 11, 1–2. <https://doi.org/10.1080/15366367.2013.783752>, 1–18.
- Hattie, J.A.C., 2009. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>.
- Herde, C.N., Wustenberg, S., Greiff, S., 2016. Assessment of complex problem solving: what we know and what we don't know. *Appl. Meas. Educ.* 29 (4), 265–277. <https://doi.org/10.1080/08957347.2016.1209208>.
- Herde, C.N., Lievens, F., Solberg, E.G., Harbaugh, J.L., Strong, M.H., Burkholder, G.J., 2019. Situational judgment tests as measures of 21st century skills: evidence across Europe and Latin America. *J. Work. Organ. Psychol.* 35, 65–74. <https://doi.org/10.5093/jwop2019a8>.
- Herman, J., Hilton, M. (Eds.), 2017. *Supporting Students' College Success: Assessment of Intrapersonal and Interpersonal Competencies*. The National Academies Press, Washington DC. <http://www.nap.edu/24687>.
- Hill, P., Barber, M., 2014. *Preparing for a Renaissance in Assessment*. Pearson, London. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/uk/documents/educator/primary/preparing_for_a_renaissance_in_assessment_and_summary_text_december_2014.pdf.
- Isaacs, T., Gorgen, C., 2018. Culture, context and controversy in setting national examination standards. In: Baird, J.-A., Isaacs, T., Opposs, D., Gray, L. (Eds.), *Examination Standards: How Measures and Meanings Differ Around the World*. University College London, Institute of Education Press, London, pp. 307–330.
- Jin, H., Shin, H.J., Hokayem, H., Qureshi, F., Jenkins, T., 2019. Secondary students' understanding of ecosystems: a learning progression approach. *Int. J. Sci. Math. Educ.* 17 (2), 217–235. <https://doi.org/10.1007/s10763-017-9864-9>.
- Jones, S.M., McGarrah, M.W., Kahn, J., 2019. Social and emotional learning: a principled science of human development in context. *Educ. Psychol.* 54 (3), 129–143. <https://doi.org/10.1080/00461520.2019.1625776>.
- Joynes, C., Rossignoli, S., Fenyiwa Amonoo-Kuofi, E., 2019. *21st Century Skills: Evidence of Issues in Definition, Demand and Delivery for Development Contexts*. K4D Helpdesk Report. Institute of Development Studies, Brighton, UK. <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/14674>.
- Kaarakainen, M.-T., Kivinen, O., Vainio, T., 2018. Performance-based testing for ICT skills assessing: A case study of students' and teachers' ICT skills in Finnish schools. *Univ. Access Inf. Soc.* 17 (2), 349–360. <https://doi.org/10.1007/s10209-017-0553-9>.
- Karlen, Y., 2016. Perceived learning environments and metacognitive strategy knowledge at the upper secondary school level. *J. Educ. Res. Online* 8 (2), 212–232. https://www.pedocs.de/volltexte/2016/12442/pdf/JERO_2016_2_Karlen_Perceived_learning_environments.pdf.
- Kay, K., 2010. 21st century skills: why they matter, what they are, and how we get there. In: Bellanca, J.A., Brandt, R. (Eds.), *21st Century Skills: Rethinking How Students Learn*. Solution Tree Press, Bloomington, pp. xiii–xxxi.
- Kelleghan, T., Greaney, V., 2020. *Public Examinations Examined*. The World Bank, Washington, DC. <https://openknowledge.worldbank.org/bitstream/handle/10986/32352/9781464814181.pdf?sequence=2>.
- Kereluik, K., Mishra, P., Fahnoe, C., Terry, L., 2013. What knowledge is of most worth? *J. Digit. Learn. Teach. Educ.* 29 (4), 127–140. <https://doi.org/10.1080/21532974.2013.10784716>.
- Koenig, J.A., 2011. *Assessing 21st Century Skill: Summary of a Workshop*. The National Academies Press, Washington, DC.
- Krathwohl, D.R., 2002. A revision of Bloom's taxonomy: an overview. *Theory Into Pract.* 41 (4), 212–218.

- Krkovic, K., Mustafic, M., Wüstenberg, S., Greiff, S., 2018. Shifts in the assessment of problem solving. In: Griffin, P., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer, Dordrecht, p. 73. https://doi.org/10.1007/978-94-017-9395-7_2, 56.
- Kuzle, A., 2018. Assessing metacognition of grade 2 and grade 4 students using an adaptation of multi-method interview approach during mathematics problem-solving. *Math. Educ. Res. J.* 30, 185–207. <https://doi.org/10.1007/s13394-017-0227-1>.
- Lai, E.R., 2011. Metacognition: A Literature Review. Pearson. http://images.pearsonassessments.com/images/tmrs/metacognition_literature_review_final.pdf.
- Lai, E.R., Viering, M., 2012. Assessing 21st Century Skills: Integrating Research Findings. National Council on Measurement in Education, Washington, DC. http://images.pearsonassessments.com/images/tmrs/Assessing_21st_Century_Skills_NCME.pdf.
- Lake, R., 2019. Preparing Students for the Uncertain Future: Why America's Educators Are Ready to Innovate—But Their Education Systems Are Not. The 74, New York. <https://www.the74million.org/article/preparing-students-for-the-uncertain-future->.
- Leutner, D., Fleischer, J., Grünkorn, J., Klieme, E., 2017. *Competence Assessment in Education: Research, Models and Instruments*. Springer, Cham, Switzerland. <https://doi.org/10.1007/978-3-319-50030-0>.
- Levy, F., Murnane, R., 2004. The New Division of Labor: How Computers Are Creating the Next Job Market. Princeton University Press, Princeton, NJ. <https://doi.org/10.1515/9781400845927>.
- Levy, F., Murnane, R., 2013. *Dancing With Robots: Human Skills for Computerized Work*. Third Way, Washington, DC. <https://www.thirdway.org/report/dancing-with-robots-human-skills-for-computerized-work>.
- Lucas, B., 2016. A five-dimensional model of creativity and its assessment in schools. *Appl. Meas. Educ.* 29 (4), 278–290. <https://doi.org/10.1080/08957347.2016.1209206>.
- Lucas, B., 2021. Rethinking Assessment in Education: The Case for Change. Centre for Strategic Education, Melbourne. <https://drive.google.com/file/d/1tQJAHdj7kEjcnt0Ztah19QffsliQrP/view>.
- Lucas, B., Venckutė, M., 2020. Creativity: A Transversal Skill for Lifelong Learning. An Overview of Existing Concepts and Practices. Literature Review Report. European Union Joint Research Centre, Publications Office of the European Union, Luxembourg.
- Marulis, L.M., Palincsar, A.S., Berhenke, A.L., Whitebread, D., 2016. Assessing metacognitive knowledge in 3–5 year olds: the development of a metacognitive knowledge interview. *Metacogn. Learn.* 11 (3), 339–368. <https://doi.org/10.1007/s11409-016-9157-7>.
- Maxwell, G.S., 2021. *Using Data to Improve Student Learning: Theory, Research and Practice*. Springer, Cham, Switzerland. <https://doi.org/10.1007/978-3-030-63539-8>.
- Maxwell, G.S., Cumming, J.J., 2011. Managing without public examinations: successful and sustained curriculum and assessment reform in Queensland. In: Yates, L., Collins, C., O'Connor, K. (Eds.), *Australia's Curriculum Dilemmas: State Cultures and Big Issues*. Melbourne University Press, Melbourne, pp. 185–206. https://pivotal.acer.org/files/11_YATES.pdf.
- Microsoft Partners in Learning, 2012. 21st Century Learning Design: Student Work Rubrics. Microsoft Partners in Learning. https://www.creatinggrounds.com/uploads/9/6/2/4/96240662/21cld_student_work_rubrics_2012.pdf.
- Organisation for Economic Co-operation and Development [OECD], 2013a. PISA 2015: Draft Collaborative Problem Solving Framework. OECD, Paris. <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>.
- Organisation for Economic Co-operation and Development [OECD], 2013b. Synergies for Better Learning: An International Perspective on Evaluation and Assessment. OECD, Paris. <https://www.oecd.org/education/school/synergies-for-better-learning.htm>.
- Organisation for Economic and Cultural Development [OECD], 2018. PISA 2018 Assessment and Analytical Framework. OECD, Paris. <https://doi.org/10.1787/b25efab8-en>.
- Organisation for Economic Co-operation and Development [OECD], (n.d.). Social and Emotional Skills: Well-Being, Connectedness and Success. Paris: OECD. [https://www.oecd.org/education/school/UPDATED%20Social%20and%20Emotional%20Skills%20-%20Well-being,%20connectedness%20and%20success.pdf%20\(website\).pdf](https://www.oecd.org/education/school/UPDATED%20Social%20and%20Emotional%20Skills%20-%20Well-being,%20connectedness%20and%20success.pdf%20(website).pdf).
- Ozturk, N., 2017. Assessing metacognition: theory and practices. *Int. J. Assess. Tool. Educ.* 4 (2), 134–148. <https://doi.org/10.21449/ijate.298299>.
- Partnership for 21st Century Learning, 2019. Framework for 21st Century Learning Definitions. Battelle for Kids, Columbus OH. http://static.battelleforkids.org/documents/p21/P21_Framework_DefinitionsBfK.pdf.
- Pelligrino, J.W., Hilton, M.L. (Eds.), 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. The National Academies Press, Washington, DC.
- Pepper, D., 2011. Assessing key competences across the curriculum—and Europe. *Eur. J. Educ.* 46 (3), 335–353. <https://doi.org/10.1111/j.1465-3435.2011.01484.x>.
- Petway, K.T., Rikoon, S.H., Brennenman, M.W., Burrus, J., Roberts, R.D., 2016. Development of the Mission Skills Assessment and Evidence of Its Reliability and Internal Structure. ETS Research Report Series, ETS RR-16-19, pp. 1–15. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ets2.12107>.
- Reich, R., 1992. *The Work of Nations: Preparing Ourselves for 21st Century Capitalism*. Vintage Books, New York.
- Roebbers, C.M., Kälin, S., Aeschlimann, E.A., 2020. A comparison of non-verbal and verbal indicators of young children's metacognition. *Metacogn. Learn.* 15 (1), 31–49. <https://doi.org/10.1007/s11409-019-09217-4>.
- Rovers, S.F.E., Clarebout, G., Savelberg, H.H.C.M., de Bruin, A.B.H., Merriënboer, J.J.G., 2019. Granularity matters: comparing different ways of measuring self-regulated learning. *Metacogn. Learn.* 14 (1), 1–19. <https://doi.org/10.1007/s11409-019-09188-6>.
- Russell, C.A., 2016. System Supports for 21st Century Competencies. Asia Society, Center for Global Education, New York: Santa Monica, CA. https://asiasociety.org/files/system-supports-for-21st-century-competencies-2016_0.pdf.
- Saavedra, A.R., Opfer, V.D., 2012. Teaching and Learning 21st Century Skills: Lessons From the Learning Sciences. Asia Society and RAND Corporation, Santa Monica, CA. <https://asiasociety.org/sites/default/files/Tteaching-and-learning-21st-century-skills-rand.pdf>.
- Sawyer, R.K., 2006. The schools of the future. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, Cambridge, pp. 567–580. <https://doi.org/10.1017/CBO9780511816833.035m>.
- Schweig, J., Baker, G., Hamilton, L.S., Stecher, B.M., 2018. Building a Repository of Assessments of Interpersonal, Intrapersonal, and Higher-Order Cognitive Competencies. RAND Corporation, Santa Monica, CA. <https://doi.org/10.7249/RR2508>.
- Shepard, L.A., 2018. Learning progressions as tools for assessment and learning. *Appl. Meas. Educ.* 31 (2), 165–174. <https://doi.org/10.1080/08957347.2017.1408628>.
- Shepard, L.A., 2019. NEPC Review: "Education System Alignment for 21st Century Skills: Focus on Assessment". National Education Policy Center, Boulder, CO. <https://nepc.colorado.edu/thinktank/assessment>.
- Siddiq, F., Gochyyev, P., Wilson, M., 2017. Learning in digital networks: ICT literacy: a novel assessment of students' 21st century skills. *Comput. Educ.* 109, 11–37. <https://doi.org/10.1016/j.compedu.2017.01.014>.
- Siddiq, F., Hatlevik, O.E., Olsen, R.V., Throndsen, I., Scherer, R., 2016. Taking a future perspective by learning from the past: a systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educ. Res. Rev.* 19, 58–84. <https://doi.org/10.1016/j.edurev.2016.05.002>.
- Silva, E., 2008. Measuring Skills for the 21st Century. Education Sector, Washington, DC. <https://eric.ed.gov/?id=ED503236>.
- Soland, J., Hamilton, L.S., Stecher, B.M., 2013. Measuring 21st Century Competencies: Guidance for Educators. Asia Society and RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/external_publications/EP50463.html.
- Suto, I., 2013. 21st Century Skills: Ancient, Ubiquitous, Enigmatic?, vol. 15. Research Matters: A Cambridge Assessment Publication, pp. 2–8. <https://www.cambridgeassessment.org.uk/Images/467762-research-matters-15-editorial-.pdf>.
- Taylor, R., Fadel, C., Kim, H., Care, E., 2020. Competencies for the 21st Century: Jurisdictional Progress. Center for Universal Education at the Brookings Institution, Washington, DC. <https://www.brookings.edu/research/visualizingthe-breadth-of-skills-movement-across-education-systems/>.
- Tuck, J.L., Moxley, J.H., 2017. A comprehensive reanalysis of the metacognitive self-regulation scale from MSLQ. *Metacogn. Learn.* 12 (1), 79–111. <https://doi.org/10.1007/s11409-016-9161-y>.

- United Nations Conference on Trade and Development [UNCTAD], 2017. Information Economy Report 2017: Digitalization, Trade and Development. UNCTAD, Geneva, Switzerland. https://unctad.org/system/files/official-document/ier2017_en.pdf.
- United Nations Educational, Scientific and Cultural Organisation [UNESCO], 2013. Toward Universal Learning: What Every Child Should Learn. The Brookings Institution, Washington, DC. <https://www.brookings.edu/wp-content/uploads/2016/06/LMTRpt1TowardUnivrsLearng.pdf>.
- United Nations Educational, Scientific and Cultural Organization [UNESCO], 2019. Education Content Up Close: Examining the Learning Dimensions of Education for Sustainable Development and Global Citizenship Education. UNESCO, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000372327/PDF/372327eng.pdf.multi>.
- van Deursen, A.J.A.M., Helsper, E.J., Eyon, R., 2016. Development and validation of the internet skills scale (ISS). *Inf. Commun. Soc.* 19 (4), 804–823. <https://doi.org/10.1080/1369118X.2015.1078834>.
- Venckutė, M., Mulvik, I.B., Lucas, B., 2020. Creativity: A Transversal Skill for Lifelong Learning. An Overview of Existing Concepts and Practices. Final Report. European Union Joint Research Centre, Publications Office of the European Union, Luxembourg. <https://publications.jrc.ec.europa.eu/repository/handle/JRC122016>.
- Voogt, J., Roblin, N.P., 2012. A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J. Curric. Stud.* 44 (3), 299–321. <https://doi.org/10.1080/00220272.2012.668938>.
- Wagner, T., 2008. The Global Achievement Gap: Why Even Our Best Schools Don't Teach the New Survival Skills Our Children Need—And What We Can Do About It. Basic Books, New York.
- Wang, Y., Lavonen, J., Tirri, K., 2018. Aims for learning 21st century competencies in national primary science curricula in China and Finland. *Eurasia J. Math. Sci. Technol. Educ.* 14 (6), 2081–2095. <https://doi.org/10.29333/ejmste/86363>.
- Wilson, M., Scalise, K., Gochyyev, P., 2015. Rethinking ICT literacy: from computer skills to social network. *Think. Skills Creativ.* 18, 65–80. <https://doi.org/10.1016/j.tsc.2015.05.001>.
- Wilson, M., Scalise, K., Gochyyev, P., 2018a. Learning in digital networks as a modern approach to ICT literacy. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Research and Applications*. Springer, Cham, Switzerland, pp. 181–210. https://doi.org/10.1007/978-3-319-65368-6_11.
- Wilson, M., Scalise, K., Gochyyev, P., 2018b. ICT literacy as a 21st century skill: learning in digital networks with agile development approaches. In: Jiao, H., Lissitz, R.W. (Eds.), *Technology Enhanced Innovative Assessment: Development, Modeling, and Scoring from an Interdisciplinary Perspective*. Information Age Publishing, Charlotte, NC, pp. 249–264.
- World Education Forum, 2015. New Vision for Education: Unlocking the Potential of Technology. WEF, Geneva, Switzerland. http://www3.weforum.org/docs/WEFUSA_NewVisionforEducation_Report2015.pdf.

Accountability for the future? International testing and future directions

Ronny Scherer^a and Trude Nilsen^{b, a} Center for Educational Measurement at the University of Oslo (CEMO), Faculty of Educational Sciences, University of Oslo, Oslo, Norway; and ^b University of Oslo, Faculty of Educational Sciences, Department of Teacher Education and School Research, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	324
Context, history, and purposes of ILSAs	325
Definition and characteristics of ILSAs	325
Historical perspectives	325
Main purposes of ILSAs	326
Future directions of ILSAs	327
Domains and populations	327
Current status	327
Future directions and needs	327
Assessment of contextual information	328
Current status	328
Future directions and needs	329
Striving for good indicators of teaching practice: student ratings, teacher ratings, and observational classroom data	329
Types of data and data sources	330
Linking multiple ILSAs	330
Longitudinal designs and extensions of ILSAs	331
Linking ILSA data to non-ILSA data	332
Methodological advancements	333
Novel task and test designs	334
Utilizing process data from technology-based assessments	334
Comparability of measures	335
Concluding remarks	337
References	337

Glossary

CFA Confirmatory factor analysis
GDP Gross domestic product
GPA Grade point average
ILSA International large-scale assessment (e.g., PIRLS, PISA, TALIS, TIMSS)
IRT Item response theory
MST Multi-stage adaptive testing
NEPS German National Educational Panel Study
PIRLS Progress in International Reading Literacy Study (IEA)
PISA Programme for International Student Assessment (OECD)
TALIS Teaching and Learning International Survey (OECD)
TIMSS Trends in International Mathematics and Science Study (IEA)
VOTAT Vary-one-thing-at-a-time—a strategy of controlling variables systematically

Introduction

International large-scale assessments (ILSAs) have become key sources of information for educational research and policymaking. As they provide information about relevant educational indicators, such as students' achievement in core domains, teachers' job satisfaction, or employees' skills to solve problems with technology, both nationally and internationally, ILSAs are pieces in the "puzzle" of accountability. Accountability in education has many facets, focusing on the processes of performance reporting as well as the technical, political, and institutional processes (Levin, 1974). It broadly refers to "a process which involves the duty both of individuals and the organizations of which they are part to render periodically accounts for tasks performed"

(Neave, 1987, p. 70). The following sections are based on this broad definition and focuses on the performance reporting and the information ILSAs can provide about the possible processes leading up to it.

This chapter discusses future directions of ILSAs in the context of accountability and describes their current status. To achieve this, we specifically aim to:

- Provide **background knowledge** of ILSAs through describing and discussing their characteristics, developments, and purposes;
- Provide brief overviews of the **current ILSA domains, populations, and contextual information** and to discuss the **future directions and needs** within these areas;
- Describe and illustrate the **potential and innovative approaches of ILSAs** with regards to the types of data, data sources, and methodological advances.

Context, history, and purposes of ILSAs

As we review the current status and future directions of international large-scale assessments in the context of accountability, we refer to their key characteristics and outline their innovations in light of their historical developments and contexts. To provide readers with some background knowledge of these issues, we (a) define and characterize ILSAs; (b) sketch four main eras of their development; and (c) describe their main purposes.

Definition and characteristics of ILSAs

The common definition of ILSAs is much broader today than it had been when ILSAs started (see **Definition and characteristics of ILSAs** section). The early ILSAs targeted representative samples of student populations across countries and mainly assessed their mathematics and science competence.

The definition of ILSAs lies in the very name which refers to assessment studies that are **international** and include **large-scale samples**. Yet, today many would include both assessment studies as well as surveys when referring to ILSAs. For instance, TIMSS is an **assessment study** capturing student competence in science and mathematics (Mullis and Martin, 2017), while TALIS is a **survey** of teachers capturing aspects of the teaching profession and teacher education (OECD, 2019b). Moreover, international and regional studies are often referred to as ILSAs (IEA, 2017). Furthermore, ILSAs may target students in school or samples outside of school, such as adults or children in pre-primary education. ILSAs may assess core subject domains, such as mathematics and reading, or cross-curricular domains, such as problem solving, digital literacy, and global competence. Nowadays, ILSAs would refer to studies that are meant for accountability informing educational policy at the national level rather than studies designed solely for research purposes (Addey et al., 2017).

ILSAs have developed extensively since the first IEA studies. One driver of this development has been the implicit alignment of ILSAs with the comprehensive model of educational outcomes—a model emphasizing all aspects of a dynamically changing school system at the system, school, class, and student level and their interactions (Stancel-Piątak and Schwippert, 2022; Blömeke et al., 2022; Wagemaker, 2020; Creemers and Kyriakides, 2008). Another driver was to accommodate the needs of societies, educational research, policymaking, and accountability systems (Addey et al., 2017). This development motivated the inclusion of new types of educational outcomes, new populations, and countries. Regional studies, such as Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) are examples of new ILSAs that were established to better adjust for the local context. PIAAC is an example of an ILSA targeting an adult population rather than students, and OECD International Early Learning and Child Well-being Study (IELS) is an example of an ILSA that samples children in pre-primary education. Examples of ILSAs that include new types of outcomes other than achievement in mathematics, reading, and science, are ICILS that assesses students' digital skills, and PISA that started to include financial literacy in 2012 (Fraillon et al., 2020; OECD, 2020b). **Table 1** lists examples of ongoing ILSAs (for a complete list, please see <https://ilsa-gateway.org/studies/factsheets>).

Historical perspectives

There are a number of publications describing in great detail the history and development of ILSAs (see e.g., Wagemaker, 2020; Rutkowski et al., 2013; von Davier et al., 2012; Nilsen et al., 2022). The present chapter, however, provides only a brief overview.

The history of ILSAs can be considered in three main eras. In the **first era**, from the 1950s to 1995, ILSAs consisted of the early IEA studies, such as the *First International Mathematics Study* (FIMS) and the *Second International Mathematics and Science Studies* SIMS and SISS (Wagemaker, 2014). Research, rather than accountability, was the main purpose of these studies, and these studies focused on the opportunities to learn (OTL) and classroom characteristics, contributing to the research fields of curricular studies, OTL and teachers' competence and instruction (Scheerens, 2016). These ILSAs included representative samples of students from a narrow selection of developed, Western countries and assessed students' competence in mathematics and science in different cohorts (Medrich and Griffith, 1992).

The **second era** consisted of the ILSAs as we know them today and started with TIMSS in 1995 (Beaton, 1996). This era introduced ILSAs as steered by policy, and the results from PISA 2000 that triggered a shock for many countries, the so-called

Table 1 Examples of ILSAs, their main focus, and governing organizations.

<i>ILSA</i>	<i>Main focus</i>	<i>Organization</i>
ICCS (International Civic and Citizenship Education Study)	Students' knowledge, understanding, beliefs, attitudes, and behaviors with respect to civics and citizenship	IEA
ICILS (International Computer and Information Literacy Study)	Students' knowledge, skills, and understanding for successfully dealing with (dynamic) information and computer technology	IEA
PIAAC (Programme for the International Assessment of Adult Competencies)	Adults' proficiency in crucial information-processing skills, including numeracy, literacy, and problem-solving in technology-rich environments, and skills at home, at work, and in their communities	OECD
PIRLS (Progress in International Reading Literacy Study)	Students' reading performance at grade 4 in relation to their reading education, monitoring transitions from "learning to read" to "reading to learn"	IEA
PISA (Programme for International Student Assessment)	Students' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges	OECD
TALIS (Teaching and Learning International Survey)	Teachers' working life in school and principals' experiences with school policy, resources, professional development and training	OECD
TIMSS (Trends in International Mathematics and Science Study)	Students' performance in mathematics and science at grades 4 and 8, monitoring achievement trends and possible changes in education	IEA

"PISA-shock." This era reflects a shift for the ILSAs, from a focus on research to emphasis on accountability, from being dominated by researchers to being influenced by policy, and from receiving little attention by others than researchers to massive media attention (Addey et al., 2017).

The **third and present era** is characterized by recent developments of old ILSAs and the addition of new ILSAs. The number of ILSAs has increased, and new populations and types of outcomes were included (von Davier et al., 2012). The community working with the ILSAs and using the data increased immensely, and methodological competence capacity was built along with the increasing complexity of the data (Blömeke et al., 2022). This era is also characterized by a rapid development within information and communication technology and ILSAs' adopting of new assessment modes, such as adaptive testing. The massive attention from the media that started in the second era continued, but ILSAs received more critique from researchers and practitioners about the heavy influence of policy on education and the emphasis on testing, among other things (Addey et al., 2017).

Main purposes of ILSAs

Today, the main purpose of ILSAs is embedded deep in accountability, meaning that ILSAs provide knowledge to countries' authorities, informing evidence-based decisions. However, as Braun and Singer (2019) have noted, the purposes of ILSAs are diverse and vary in their appropriateness. Specifically, while ILSAs can deliver on comparative descriptions of educational outcomes, conditions, and inputs, drawing causal inferences on the basis of ILSA data is the least appropriate purpose. Indeed, the limitations of the sampling and study designs mainly allow for correlational inferences (Rutkowski and Delandshere, 2016). These correlations are, however, too often interpreted as causal effects (Gustafsson and Nilsen, 2022). The main threats to correlational inferences are omitted variables, confounding variables, reversed causality and, in the case of ILSAs, self-reports (Gustafsson, 2013). Specifically, reversed causality entails that one would not know whether the predictor (e.g., student motivation) affects the outcome (e.g., student achievement) or the other way around. Furthermore, omitted and confounding variables entail that one would not know whether it is the predictor (e.g., class size) that affects the outcome (e.g., student achievement) or whether there is an omitted variable (e.g., socio-economic status, SES) that actually predicts the outcome. Omitted variables can affect both the outcome and the predictor (Gustafsson, 2013).

To what extent ILSAs fulfill the purpose of providing knowledge for evidence-based decisions is a complex question and depends on the type of knowledge in question, on the type of ILSAs one refers to, and what characterizes the purposes of these studies. In the following, we will provide a brief overview of the purposes and characteristics of two of the most influential ILSAs, namely TIMSS and PISA.

PISA aims to measure the reading, mathematical, and scientific literacy students need in the future, while TIMSS aims to measure students' competence in mathematics and science corresponding to the countries' curricula (OECD, 2019a; Mullis and Martin, 2017). To this end, it means that PISA has a future perspective, focusing on what student should learn to manage their lives in future societies, while TIMSS has a retrospective perspective, measuring what students should have learned following their curricula (Blömeke et al., 2022). PISA includes cross-disciplinary domains (e.g., global competence; see also the chapter by Ozga, 2022; in this encyclopedia) and provides advice to countries' policymakers (Addey et al., 2017). TIMSS, on the other hand, provides

descriptive statistics rather than correlational analyses, and does not provide advice to countries' policymakers directly (Wagemaker, 2014).

PISA takes a top-down approach, where decisions are made by expert groups and policy makers (Blömeke et al., 2022). The OECD countries are in the main focus, although more developing countries are gradually joining the study. TIMSS takes a more bottom-up approach based on the collaborative enterprise by the participating countries (Mullis et al., 2020). While TIMSS is, to a large extent, motivated by research rather than policy, there is an increasing influence from policymakers (Blömeke et al., 2022). Compared to PISA, TIMSS has included more developing countries, although both studies have an increasing, heterogeneous mix of countries (Rutkowski and Rutkowski, 2018; see also the chapter by Rutkowski and Rutkowski, 2022, in this encyclopedia).

TIMSS measures "educational achievement in mathematics and science at the fourth and eighth grades, and sometimes the final year of secondary school" (IEA, 2021). The stated goal of PISA is to assess "the extent to which 15-year-old students near the end of their compulsory education have acquired the knowledge and skills that are essential for full participation in modern societies" (OECD, 2019a). While the goals of TIMSS and PISA are quite different, the goal perceived by policymakers is one and the same—accountability.

Future directions of ILSAs

In this section, we first provide an overview of *ILSA's domains and populations* (3.1) and their *assessments of contextual information* (3.2) and, second, illustrate their future directions. The following sections focus on the potential and innovative approaches of ILSAs with respect to the *types of data and data sources* (3.3) and the *methodological advancements* (3.4).

Domains and populations

Current status

The current status of **domains included in ILSAs** is rapidly changing. Currently, the main and oldest ILSAs (i.e., TIMSS, PISA, and PIRLS) focus on traditional cognitive domains, such as mathematics, science, and reading. However, the inclusion of new cross-disciplinary areas, such as problem solving, global competence, and financial literacy, is increasing. Especially PISA has lately included one new domain per cycle (see, for example, OECD, 2020b). In addition, there are new ILSAs that measure cross-disciplinary outcomes, such as ICILS measuring digital skills and computational thinking (Fraillon et al., 2020). Very few ILSAs measure other types of cognitive outcomes in domains beyond mathematics, science, and reading. For instance, DICE¹ measures the effects of educational theater and drama on five of the eight Lisbon Key Competences in a longitudinal EU-funded project in 12 countries (Gjærum et al., 2022; Cziboly, 2010).

The current status of **populations included in ILSAs** is also rapidly changing. Traditionally, the main ILSAs have targeted students in school. The target grades of TIMSS are grades 4 and 8, PIRLS includes only fourth graders, and PISA samples 15-year-old students. However, newer ILSAs include populations outside of school, such as PIAAC that includes adults, IELS that includes children in kindergarten, and TALIS that includes teachers. In addition, the inclusion of low- and middle-income countries in PISA and TIMSS has increased over the years (Rutkowski and Rutkowski, 2019; see also the chapter by Rutkowski and Rutkowski, 2022, in this encyclopedia). New regional studies and a new PISA study called PISA-D targets developing countries and strive to take into account the local context (Kamens and Benavot, 2011).

Future directions and needs

As education and society change, the **need for measuring new domains**, at least from the perspective of accountability, is increasing. There is a common understanding that students need other types of skills in the 21st century, and that these skills are more cross-disciplinary in nature (Griffin and Care, 2014). While some ILSAs already include some of these cross-disciplinary domains, there is still a need to extend these to include other key domains, for instance, creative thinking, self-regulation, or socio-emotional skills. Lately, democracy has been endangered in some countries (Afsahi et al., 2020); hence, the International Civic and Citizenship Education Study (ICCS) is becoming increasingly important (Schulz et al., 2018).

There is further a **need for better assessments**, both for those that have already been included as well as for possible future domains. There are many ways to improve the measurement of these domains, for instance, through computer-based assessment that would allow for interactive items, the collection and analyses of process data, and adaptive testing (see sections **Novel task and test designs** and **Utilizing process data from technology-based assessments**). At the same time, many cross-disciplinary skills might be hard to measure, such as emotional intelligence like empathy or real collaboration.

There is also a **need for including subject domains other than mathematics, science, and reading**, for instance, art subjects or history. Including other subject domains would provide a more holistic picture of students' competence and a more complete picture in terms of accountability. Moreover, operationalizing the outcomes in new subject domains entails research on curricula and theories underlying the subjects, which again could contribute to enhance the quality of the subjects. However, creating ILSAs that measure "everything" could compromise the validity of inferences. To use an analogy, there is a reason why there are different

¹Drama Improves Lisbon Key Competences in Education, <http://www.dramanetwork.eu/acknowledgements.html>.

types of airplanes, for instance one type for cargo, another to transport passengers, and a third for military purposes. Having one type of plane to cover all these needs, would not be efficient and would reduce the quality of the transport.

In terms of **populations**, ILSAs already cover pre-primary, school, to adult samples. Among other, a key challenge in terms of populations is the balance between high- and low-income countries, between developed and developing countries, or between Western and non-Western countries. ILSAs were primarily built on the theories and contexts of Western countries (Blömeke et al., 2022), and even though non-Western countries' voices are increasingly impacting the processes and decision-making around the development and implementation of the ILSAs, there is still a need to include these voices to a larger degree (Blömeke et al., 2022). Including theories and contexts of low- and middle-income countries is important. The regional studies may be a consequence and response of countries striving for an ILSA that takes into consideration the country contexts and cultures (Kamens and Benavot, 2011). Previous research (Rutkowski and Rutkowski, 2018, 2019; Rutkowski et al., 2018) demonstrated that the increased heterogeneous mix of countries participating in ILSAs is problematic for establishing achievement scales, as test items are not equally difficult across countries. This means that the validity of the achievement scale is threatened, and this is especially problematic at the ends of the scales. Hence, for the low-performing countries, the ILSA tests may not provide an accurate picture, and are of less value. This issue could be partly addressed in PISA-D and via adaptive testing (see **Novel task and test designs** section).

Assessment of contextual information

Current status

A **mutual relationship between research and ILSAs** exists, in that ILSAs inform research and vice versa (Blömeke et al., 2022). For instance, school climate is a broad construct that is defined differently across many studies (Wang and Degol, 2015). However, through consensus and feedback, the ILSAs have continuously improved their indicators of school climate (e.g., school safety, bullying, and students' sense of school belonging; Mullis and Fishbein, 2020; Mullis and Martin, 2017). Perhaps as consequence of a more clearly defined framework for school climate and of high-quality measures and data, many secondary analyses of school climate ILSA data have been published (Teig and Nilsen, 2022). In this way, ILSAs have generated research evidence and opportunities. At the same time, researchers influence the shaping of the ILSA constructs as part of expert groups (Mullis and Fishbein, 2020; OECD, 2019a). In addition, the frameworks are built on previous research (OECD, 2019a; Mullis and Martin, 2017). In other words, research contributes back to ILSAs and influences their development.

The **types of constructs and contextual information** included in the contextual questionnaires in ILSAs, is primarily guided by educational effectiveness on the one hand (Klieme, 2013), and policy and accountability on the other hand (Addey et al., 2017). Educational effectiveness is a field concerned with identifying factors that contribute to make a school "good" (e.g., student cognitive and affective outcome; Reynolds et al., 2014), and takes into account the dynamic nature and hierarchical design of educational systems (Creemers and Kyriakides, 2008). Other than including contextual information, ILSAs also include constructs and variables that may not be directly related to student outcomes, such as the availability of a laboratory at school (Mullis et al., 2020) and the content coverage in classrooms (Scheerens, 2016), but that would still inform the countries' accountability systems and policymaking.

Including new contextual information entails extending and revising the contextual frameworks. This is an ongoing process, and during this process the quality of the overall framework is often improved (Blömeke et al., 2022). Some constructs included in ILSAs have long had a solid foundation in theory, such as student motivation that is based on the Expectancy-Value Theory (Hooper et al., 2013; Eccles and Wigfield, 2002). Other constructs, such as teachers' instructional quality has been recently included, extended, and improved, utilizing the framework of the basic dimensions of teaching (Nilsen and Gustafsson, 2016; Kuger et al., 2016; Mullis and Martin, 2017; OECD, 2013). In section **Striving for good indicators of teaching practice: student ratings, teacher ratings, and observational classroom data**, we provide an example of how practices pointing to instructional quality are assessed.

Even though different ILSAs include different contextual information, there are some **key constructs and variables** that are similar across the studies and that align with educational effectiveness (Creemers and Kyriakides, 2008). To use TIMSS and PISA as examples, they both administer student and school leader questionnaires, and TIMSS, in addition, includes teacher questionnaires in both grades and parent questionnaires in grade 4. In the student questionnaires in these two studies, some of the key constructs and variables include student home background and characteristics (e.g., SES, minority, and gender), learning environment (e.g., school belonging, and bullying), perceptions of teachers' instructional quality, and student motivation. Key constructs and variables in the school questionnaires include principals' background and education, school climate (e.g., school safety and climate), and information about the school (e.g., characteristics of the neighborhood, school composition). Some of the key constructs and variables in the TIMSS' teacher questionnaire include teachers' formal qualifications and education, their experience, instruction, self-efficacy in teaching mathematics and science, job satisfaction, and the implemented curriculum (the degree to which the content in TIMSS framework has been taught). Key constructs and variables in the parent questionnaires include early literacy and numeracy of their child before attending school, background and characteristics of child and parent (e.g., SES and minority), and parents' attitudes toward school and to mathematics and science. Just like the test of students' cognitive outcomes, the contextual constructs and variables are subject to strict quality assurance procedures (Wagemaker, 2020). These procedures include feedback from the countries and expert groups, piloting, and psychometric evaluation to ensure high reliability and validity.

Future directions and needs

Although the underlying theory of constructs and variables in the contextual questionnaires has improved, there is still room for improvement (Blömeke et al., 2022). Measures for SES is a typical example of this and has been subject to much criticism (e.g., Rutkowski and Rutkowski, 2018). One of the major problems with SES and many other constructs in ILSAs, is that their conceptualizations and operationalizations may vary across cultures (van de Vijver, 2018). In PISA, for instance, one of the indicators of SES is whether students have their own room, and while this may help differentiate between low- and high-SES students in developing countries, it may not be the case for high-income countries. While much effort has been taken in psychometrics to ensure comparability of constructs across different cultures (see *Comparability of measures* section), there is still a need for the same type of effort within the theoretical foundation underlying these constructs.

Striving for good indicators of teaching practice: student ratings, teacher ratings, and observational classroom data

As ILSAs strive for describing contextual information about teaching and learning, assessing teaching practices has gained considerable importance. However, measuring teaching practices is not an easy task, given the heterogeneous mix of countries in ILSAs (see *Future directions and needs* section). Despite this challenge, it has become the hope to identify “effective” practices that have the potential to foster student learning and compare the occurrence of these practices across countries (Klieme, 2013). However, researchers are struggling with defining the practices that promote learning and motivation (Praetorius et al., 2018), and many different theoretical frameworks exist in different parts of the world (e.g., Praetorius et al., 2018; Kane and Cantrell, 2010; Ferguson and Danielson, 2015; Kyriakides et al., 2000). Indeed, most research within this area has been developed in Germany and the USA, and several studies have found that these frameworks may not create a good foundation for comparisons across different cultures (Blömeke and Olsen, 2019).

ILSAs have progressed toward including several indicators of teaching practices and examine their relations to educational outcomes (as indicators of educational effectiveness). As most ILSAs do not follow a strict experimental design, inferences on educational “effectiveness” are, nevertheless, limited to the description of relations among classroom, school, and student variables (Klieme, 2020).

Student and teacher ratings of teaching practice

Depending on the source of information (e.g., student or teacher ratings), measures of teaching practices may have different interpretations. For instance, while student ratings primarily represent individual perceptions, classroom-average student ratings represent shared perceptions and could be considered a classroom rather than an individual characteristic (Scherer et al., 2016). Marsh et al. (2012) argued that researchers targeting the effectiveness or benefits of some teaching practice for student achievement should consider the classroom level and aggregated student ratings rather than using the student and individual ratings at the individual level in their analyses. In this sense, variation between student ratings represents a form of measurement error. As such, student ratings of teaching practices reference the classroom and/or the teacher and operate at these levels.

Klieme (2020) cautioned that not all ILSAs qualify for addressing research questions on educational effectiveness and practice. For instance, although students were asked to indicate the frequency of core teaching practices, such as classroom management, teacher support, and cognitive activation in mathematics in PISA 2012, the sampling design of this ILSA does not allow for inferences on educational effectiveness per se. As PISA samples students and schools randomly in most participating countries (with some exceptions, such as Germany), the level at which student ratings of teaching practice operate (i.e., the classroom/teacher level) is not contained (Eriksson et al., 2019). In contrast, the sampling design of TIMSS 2019 (and, in fact, all TIMSS cycles) contains the classroom and teacher level, so that student ratings are meaningful and have the potential to represent teaching practices meaningfully. Similarly, TALIS 2018 assessed teacher ratings of teaching practices and conceptualized them as teacher-level variables. Already in the national extensions of TIMSS 2015 in Belgium (Flanders), Germany, and Norway, student ratings of teaching practices were included. Bellens et al. (2019) took a multilevel structural equation modeling approach to describe the structure of these ratings and made the classroom-average ratings the core of their investigation. The authors have found empirical support for the hypothesized three dimensions of teaching practices at both the student and the classroom level, thus crafting a validity argument (see Fig. 1A).

This example highlights the importance of establishing appropriate representations of teaching practice constructs and clarifying at which level of analysis they may operate. Aligning the levels of measurement and construct representations is a key direction in ILSAs’ striving for information about the link between teaching and learning (Stapleton et al., 2016).

Although student and teacher ratings of teaching practices are accessible in ILSA via background questionnaires, these ratings still represent only perceptions of the classroom activities or environment or self-perceptions of the implemented practices. In fact, some evidence suggests that student and teacher ratings can differ considerably, even in clear-cut practices, such as classroom management (Wagner et al., 2016). Moreover, ratings may be specific to the teaching domains, contents, and situations (Praetorius et al., 2014; Scherer and Gustafsson, 2015). These observations have led to some criticisms of the student and teacher ratings of teaching practices in ILSAs and called into question the validity of the inferences drawn from them (Klieme, 2020).

TIMSS 1995 and 1999 video studies

To address these issues and move beyond using student or teacher ratings as indicators of teaching practices, some ILSAs have initiated observational classroom studies. For instance, TIMSS 1995 and 1999 have been supplemented by video studies (Neubrand, 2006). The TIMSS 1995 Video Study examined the mathematics instruction of 231 German, Japanese, and US-American classrooms from nationally representative teacher samples. Besides its focus on developing and testing the methodology and feasibility of large-

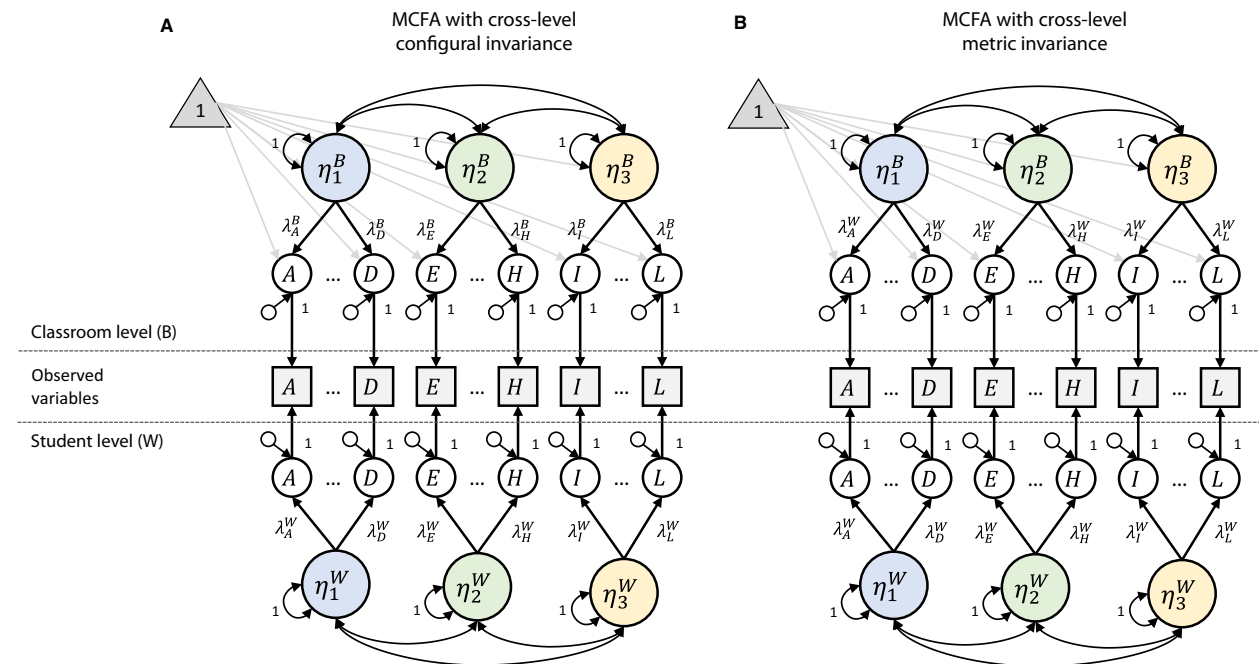


Fig. 1 Multilevel Confirmatory Factor Analysis (MCFA) models with (A) cross-level configural and (B) cross-level metric invariance.

Note. B = Between (school) level, W = Within (teacher) level, A–L = Items measuring the constructs. Residual variances are omitted in the figure to improve the clarity of presentation.

scale classroom observations, this study also resulted in video vignettes to inform teachers, educators, and policymakers around the world. Extending the range of participating countries, the TIMSS 1999 Video Study captured classroom practices not only in mathematics but also in science classrooms. Among others, this study has shed light on the typical practices of teaching both subjects, how students are provided with opportunities to learn, the extent to which teachers support student learning and scientific practices. These two video studies have initiated a plethora of follow-up video studies at small and at large scale (e.g., Klette, 2016), and have shaped the development of methodological standards for assessing and operationalizing teaching practices in mathematics and science (Jacobs et al., 2007). Identifying the similarities and differences in classroom instruction has become a key interest (Stigler et al., 2000).

TALIS 2018 video study and Global Teaching InSights

As an extension of the TALIS 2018 study, the OECD has launched an international large-scale video study to shed light on (a) the types of teaching practices; (b) their interrelations; and (c) their relations to students' cognitive and non-cognitive outcomes (OECD, 2018). This study utilized an advanced video observation methodology, focusing on a single curricular unit and including a large and diverse teacher sample. This study implemented a longitudinal design with pre- and post-test measures of student achievement, student and teacher ratings of instruction, videotaped classroom lessons, and an evaluation of the teaching material. Besides the insights this study provides into the effectiveness and implementation of specific teaching practices, it provides rich material to evaluate and code classroom instruction. Selected video sequences are being made available for teachers, teacher educators, and researchers via a digital platform for supporting professional development, teacher training, and research projects (OECD, 2020a).

Types of data and data sources

International large-scale assessment data provide researchers and policymakers with many opportunities to evaluate educational systems nationally and internationally. Over the last two decades, ILSA data sources have been extended in several ways: Some countries utilized the data from multiple ILSA cycles to examine changes in educationally relevant constructs and indicators at the country or system level over time, and/or added longitudinal follow-up studies to examine such changes at the level of students, classrooms, or schools. Apart from these extensions, under certain conditions, the data from one ILSA could also be linked the data from another ILSA (e.g., TIMSS and PIRLS, PISA and TALIS) and to national data sources (Scherer and Gustafsson, 2015; Barth et al., 2021; Kaplan and McCarty, 2013). In the following, we describe these opportunities and trends and showcase their potential.

Linking multiple ILSAs

ILSAs vary in their sampling designs, assessment domains, and the levels at which information is extracted. Linking different ILSAs can therefore offer opportunities to examine the interrelations between unique and common assessment domains and to link data at multiple analytic levels.

Linking multiple ILSAs at the individual, classroom, or school level

There have been several occasions at which multiple ILSAs were administered in the same year. Some countries have utilized these occasions and linked different ILSAs, so that one data set can inform the other. Ideally, the sampling designs are aligned, and direct links between the different data sets can be established. However, if such direct links cannot be established, matching procedures may approximate them (Kaplan and McCarty, 2013). For instance, TIMSS and PIRLS were linked for fourth graders in 2011 (Martin and Mullis, 2013). This allowed researchers to examine the relations among student achievement in reading, mathematics, and science and contextual variables. Scherer and Gustafsson (2015) utilized this design and studied student ratings of teaching practices across the three domains at both the individual and the classroom level. Their study supported the domain specificity of student ratings, even for generic teaching practices—an insight that may not have been gained without this link. In other instances, researchers have utilized the possibilities for data fusion between PISA and TALIS (Gil-Izquierdo and Cordero, 2018), PISA and TIMSS (Carnoy et al., 2016), PIAAC and PISA (Meroni et al., 2015), and PIAAC and TALIS (Hämäläinen et al., 2021).

Linking multiple ILSAs at the country level

At the level of participating countries, many creative links between ILSAs can be established (Gustafsson, 2018). To ensure their validity, however, researchers need to pay close attention to the extent to which the corresponding data, measures, or effect sizes are comparable across ILSAs. If sufficient comparability can be established, integrative data analysis and meta-analysis provide a rich methodological repertoire to quantify the variation of empirical results across ILSAs, countries, and over time, and the evidence for the generalizability and replicability of phenomena (Cheung and Jak, 2016; Hedges and Schauer, 2019).

For instance, Keller et al. (2021) examined the nonlinear relations between students' academic self-concept and achievement and analyzed the data from 13 countries participating in TIMSS, PIRLS, and PISA for each study cycle separately. Once the relevant effect sizes had been extracted, the authors combined them across the different ILSAs via random-effects meta-analysis. These analyses resulted in overall effect sizes for different achievement domains and age groups and their respective variance components. In contrast to ordinary random-effects meta-analysis, in which two variance components are estimated, the linking of multiple ILSAs resulted in a more complex data structure with more variance components. Fig. 2 exemplifies such a structure for the domain of mathematics in Keller et al.'s study. Given that the 13 countries participated in multiple ILSAs and ILSA cycles, the meta-analytic data structure is *cross-classified*, that is, next to the level of effect sizes, the levels of countries and ILSAs co-exist independently (Fernández-Castilla et al., 2019). Generally, the combination of analyzing the ILSA data to obtain the effect sizes of interest and their meta-analysis across countries bear great potential for synthesizing findings and examining their generalizability.

Longitudinal designs and extensions of ILSAs

Longitudinal studies in education have the potential to shed light on learning gains, developmental processes over time, and their mitigating factors. Under certain conditions, such designs can aid the crafting of a causality argument (Rutkowski and Delandshere, 2016). In the context of ILSAs, trends in educational success and changes in performance in core educational domains may serve as direct accountability measures (Rutkowski et al., 2010). Two main approaches to generating longitudinal data have mainly been taken and developed further: (a) Longitudinal designs of ILSAs across cycles at the level of countries; and (b) Longitudinal extensions of ILSAs within countries.

Longitudinal ILSA designs at the country level

As many ILSAs repeat their cycles, countries are provided with information about the possible changes in educationally relevant constructs over time. These changes, however, may only be interpreted at the (aggregated) country level, if no direct link between the different ILSA cycles is established at the student, classroom, teacher, or school level. Interpretations along these lines should also account for possible linking error between ILSA cycles (Robitzsch and Lüdtke, 2019).

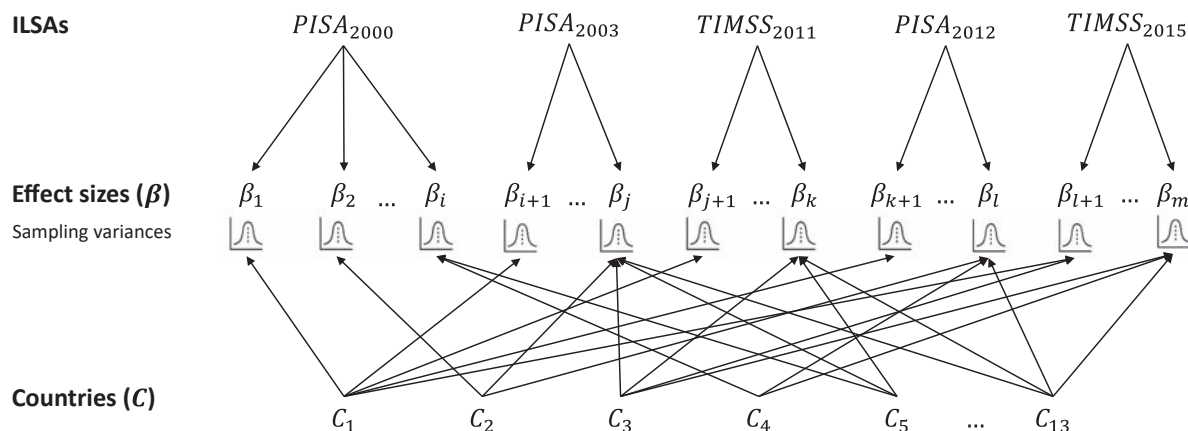


Fig. 2 Meta-analytic data structure quantifying heterogeneity between the ILSAs and 13 selected countries in mathematics.

Drawing from the evidence of more than twenty years of ILSAs, including PIRLS, PISA, and TIMSS, [Strello et al. \(2021\)](#) meta-analyzed the effects of early tracking and educational inequalities on student achievement. This study provided evidence that tracking may have increased achievement gaps in core educational domains over time. [Nilsen and Gustafsson \(2014\)](#) utilized the longitudinal design of TIMSS and studied the extent to which changes in students' science achievement were related to changes in schools' emphasis on academic success in Norway between 2007 and 2011. The authors found support for the mediating role of the schools' emphasis on academic success for the relation between the year of the TIMSS cycle and science performance, thus partly explaining performance changes. Extending this research to 38 countries participating in TIMSS 2007 and 2011, [Gustafsson and Nilsen \(2016\)](#) showed that differences in mathematics achievement and differences in aspects of school climate were related over time. These examples highlight the potential that lies within the country-level longitudinal designs of ILSAs.

Longitudinal extensions of ILSAs at the individual level within countries

Some countries have strived for extending ILSAs longitudinally at the individual participants' level. Such extensions provide rich information, as they make visible not only country-level but also individual-level changes over time ([OECD, 2013](#)). For instance, in Germany, a second measurement occasion has been added to PISA 2012, transforming this ILSA into a pretest-posttest study to examine achievement gains ([Nagy et al., 2016](#)). Similarly, the PIAAC 2012 cycle had been extended longitudinally ([Rammstedt et al., 2017](#)). Such extensions can provide evidence on the direct link between changes within variables over time and could inform value-added models in education ([Braun and von Davier, 2017](#)). Ultimately, these extensions are key future directions of ILSAs to inform educational practice, policymaking, and accountability ([OECD, 2013](#)).

Linking ILSA data to non-ILSA data

Exploiting the potential of ILSAs for accountability, many countries have aimed for linking them to national data, such as students' performance on national tests covering specific curricular content areas, register data on participants' pathways to their current profession, or official statistics on countries' economic development and innovation potential. As [Gustafsson \(2018\)](#) noted, the increasing availability of additional data sources provides researchers with "virtually unlimited possibilities for creative combinations of data which allow investigation of phenomena in several different disciplinary areas" (p. 332). In this way, it becomes possible to anchor the results of ILSAs in national contexts and to draw inferences that may aid national policymaking in education. Reviewing these attempts, we observed that ILSA data have mainly been linked to non-ILSA data at the individual, classroom, school, or national (country) level (see also [Strietholt and Scherer, 2018](#)). The following studies illustrate these links and their potential.

Bringing together ILSAs and national assessments

By and large, national and international assessments in education follow similar goals, such as monitoring students' proficiency levels and progress in core domains. While national assessments are oftentimes focusing on knowledge, skills, and other constructs in specific curricular domains or content areas, ILSAs cover knowledge, skills, and other constructs in cross-curricular domains or content areas relevant for a broad range of countries. In this sense, national assessments bear the potential to assess relevant educational outcomes more specifically tied to the national contexts—hence, linking ILSA and national assessment data can provide more contextual information about educational systems. For instance, [Jerrim et al. \(2019\)](#) linked the PISA 2015 data with the data from national assessments in England. This link provided evidence against the common belief that inquiry-based teaching practices in science are effective for fostering students' scientific literacy and science achievement. There was no support for the relation between the frequency of teaching scientific inquiry and performance on science exams, although this relation existed for the PISA science tests.

However, the designs, sampling procedures, and assessment characteristics may differ substantially between ILSAs and national assessments. Such differences necessitate studying the extent to which the two types of assessments may capture educational outcomes similarly. Linking the PISA 2012 data to the data of the German National Educational Panel Study (NEPS) at the student level, [Ehmke et al. \(2020\)](#) evaluated the comparability of the PISA and NEPS mathematics scores. Utilizing item response theory modeling, the authors found a high correlation between these scores ($r = 00.90$) and obtained further evidence for the correspondence between the two assessments. On the contrary, [Kampa et al. \(2021\)](#) provided some evidence that such a correspondence may not be a given. Specifically, they showed that a PISA scientific literacy test and a curricular-oriented science test from NEPS differed in the extent to which they assessed general and science-specific abilities. Overall, the linking of ILSA and national assessment data can aid the crafting of a validity argument for the national or international assessments.

Linking ILSA data to national registers

National register data can contain rich information about participants' educational pathways, academic success, employments, health, or wages. Linking these data with educational indicators derived from ILSAs could ultimately aid the study of lifelong learning and predictors of academic success at later stages in life. For instance, [Barth et al. \(2021\)](#) combined the Norwegian PIAAC 2011 data of young adults on numeracy and literacy with the registered GPA at age 16 and participants' employment and education two years after the PIAAC study. After controlling for indicators of home background, health status, and other variables, the authors found that GPA explained about 37 % of the variation in numeracy. Furthermore, GPA and the PIAAC scores were significantly related to later educational and economic success. This example highlights the value added by linking ILSA and register data.

Making sense of contextual information about countries

Studying the extent to which teachers' team innovativeness related to relevant teacher constructs, such as job satisfaction, teaching practices, and collaboration, Blömeke et al. (2021) analyzed the TALIS 2018 data from 48 participating countries and educational systems. In their study, the authors identified meta-analytically a substantial variation in the relationships among constructs across the countries. For instance, the contextual effect of teachers' shared perceptions of innovativeness in their schools on their individual satisfaction with the school environment ranged between $B = -0.10$ (small negative effect) and $B = 0.53$ (large positive effect). Similarly, the relations between perceived innovativeness and the extent to which teachers engaged in innovative teaching practices, for instance including technology and focusing on the teaching of 21st century skills, exhibited substantial between-country heterogeneity. In an attempt to explain this heterogeneity, Blömeke et al. (2021) extracted contextual information about the participating countries from additional data sources and examined whether the GDP per capita and the degree of individualism in a country moderated the innovativeness–teaching practices relation (see Fig. 3). While there was no evidence suggesting that GDP explained the variation between countries, there was evidence on the moderating role of individualism. Countries with higher scores on Hofstede's individualism scale were more likely to show a weaker innovativeness–teaching practices relation. Overall, about 25.4% of the variation could be explained. This example shows how additional context information can shed light on the possible sources or correlates of the variation between countries.

Once again, correlation is not causation

While the linking of ILSA data with external data sources can shed light on the “creative” relations between ILSA and non-ILSA measures in education (Gustafsson, 2018), the inferences drawn from such relations need to be treated with caution. At the country level, such inferences only represent correlations, yet not causation (Aldrich, 1995). For instance, a recent report on the status of school education after one year into the COVID-19 pandemic displayed the relationship between countries' average performance score on the PISA 2018 reading assessment and the number of days with full closures of upper secondary schools in 2020 (OECD, 2021). For the selected sample of countries, the relationship was negative, statistically significant, and corresponded to 54.1% of variance explanation. This observation, however, only showed that countries with lower average reading performance in 2018 *were likely* to experience school closures in 2020. Similarly, countries experiencing many school closures in 2020 were likely to be among the low-performing countries in PISA 2018 reading. Given the design of the PISA 2018 study and the nature of the school-closure data (e.g., different measurement occasions, different student and school samples as compared to PISA, 2018), inferences on poor learning outcomes causing disruptions of face-to-face learning or limited face-to-face learning in schools causing poor reading performance cannot be drawn. This example illustrates the need for carefully interpreting the links between ILSA and non-ILSA data and for communicating clearly which types of inferences can be drawn. Hence, researchers, educators, and policymakers should always consider the design and contexts of the respective studies providing these data.

Methodological advancements

Over the last two decades, international large-scale assessments have undergone major developments in their methodology. Especially with the advancements of technology, novel test and task designs have made their way into the ILSAs and have innovated the measurement of skills in the core and supplemental domains (Greiff et al., 2017). Besides these innovations, key issues of both the reliability of scores and the validity of the inferences drawn from these scores have been and are currently being addressed. These issues include but are not limited to the comparability of scales and their respective scores across countries and educational systems and the inclusion of information about test-taking behavior (e.g., Gustafsson, 2018; Schulz and Carstens, 2020; Beller, 2013).

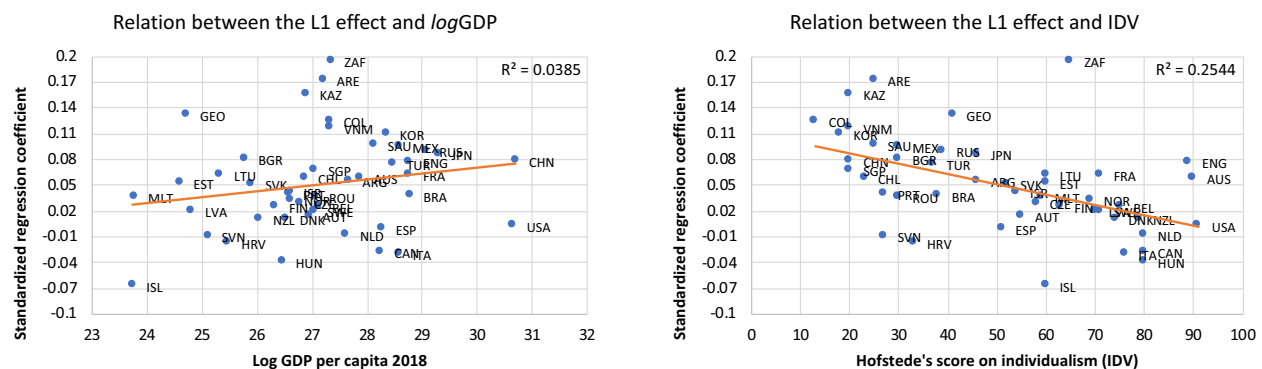


Fig. 3 Relations between *logGDP*, individualism, and the teacher-level effect of perceived team innovativeness on innovative teaching practices. Note. Countries' three-digit ISO codes are shown in the figures. Based on the TALIS (2018) data.

Novel task and test designs

Novel task and test designs in ILSA are aimed at improving the assessments of core educational outcomes, which are ultimately used to inform accountability. At the same time, such designs make accessible “hard-to-measure” skills, such as collaboration, complex problem solving, or computational thinking (Greiff et al., 2017).

Innovative item and task formats

With the shift toward computerized testing, new items and task types have been made available. For instance, in the PISA 2012 problem-solving assessment, interactive items were included to assess students’ abilities to generate information about an unknown system and acquire knowledge about its functioning (OECD, 2014). These items required a direct interaction and could be used to trace students’ task-taking behavior in computer log files. Adding a component of collaboration, through which students could collaborate and communicate with an avatar, PISA 2015 assessed collaborative problem-solving skills—skills that would not have been accessible with paper-and-pencil tests. Similarly, PISA 2015 contained scientific inquiry tasks, in which students could explore the effects of varying one variable at a time in an experiment (Teig et al., 2020). These examples illustrate how the assessments of key educational constructs could be extended by computerized tasks.

Multi-stage adaptive testing

Utilizing the potential of adaptive testing and striving for improving the precision of proficiency estimates, some ILSAs, such as PISA 2018 and PIAAC 2012, have introduced multi-stage adaptive testing (MST). This test designs allows test-takers to take different testing routes, depending on their responses to a set of pre-defined items or tasks (Yamamoto et al., 2018). For instance, in the first PIAAC cycle, participants were first assessed in their digital skills and then routed to the assessments of literacy, numeracy, and problem solving. This way, participants with little or no computer experience could be directed to a paper-and-pencil test and/or interview to adequately assess their skills (Kirsch and Lennon, 2017). In PISA 2018, the digital reading assessment contained a set of core tasks which were then used to administer follow-up tasks. This MST design has the potential to improve measurement precision and keep the time students spend on the assessment reasonable. Alongside the development of MST designs, the psychometric models describing performance and directing the selection of items and tasks have also been advanced (Magis et al., 2017). Moreover, current research evaluates the optimal routing strategies to be implemented in ILSAs (Svetina et al., 2019).

Utilizing process data from technology-based assessments

As noted earlier, ILSAs are shifting toward technology-based assessments (Beller, 2013; Cresswell et al., 2015). This shift offers several opportunities to measure skills and to gain information about the task-taking processes via so-called log files—files that store the interactions and activities while participants work on tasks. The resultant process data can (a) inform test designers about the match between the intended and observed response processes to craft a validity argument (Pellegrino et al., 2016; Ulitzsch et al., 2021); (b) uncover systematic or unsystematic strategies or behaviors needed to solve problems (Greiff et al., 2015); (c) indicate the effort participants have spent while taking an assessment (Scherer et al., 2015); (d) explain why some test-takers succeed and other do not (OECD, 2015); or (e) improve the estimation of proficiency scores in test domains (Reis Costa et al., 2021).

Uncovering problem solving strategies and response behaviors

In PISA 2012, students’ complex problem-solving skills were assessed by interactive tasks, in which the information needed to solve the task had not been provided at the outset of the task. A typical example is the so-called “Climate Control” task, which contained three input variables and two outcomes (OECD, 2014). Hence, students had to interact with the task by manipulating variables systematically, observe the outcomes, and interpret them. This process of knowledge acquisition ultimately aids the development of a mental model about how input and outcome variables are related (Greiff et al., 2013). In their secondary analyses of the PISA 2012 data, Greiff et al. (2015) explored students’ problem-solving strategies, focusing on the Vary-One-Thing-At-a-Time (VOTAT) strategy. The authors found that students who performed this strategy were more likely to succeed with the overall problem-solving task, including the application of the knowledge acquired in an exploration stage. This way, Greiff et al. (2015) were able to classify students’ performance not only into successful or unsuccessful but also into systematic (i.e., following the VOTAT strategy) or unsystematic. The actions stored in log files uncovered the performance of this particular problem-solving strategy.

While the VOTAT strategy may be straightforward to identify, other strategies or sequences of actions may be more challenging to extract. To address this challenge, data-mining approaches are currently explored to make visible more complex test-taking behavior and their profiles (e.g., Qiao and Jiao, 2018; Han et al., 2019; Teig et al., 2020). While many of these approaches are exploratory, Goldhammer and Zehner (2017) cautioned that any process indicator extracted from log file data must undergo a validation process to facilitate meaningful inferences and interpretations.

Indicating test-taking effort

Process data can further provide information about test-taking effort and indicate motivation or achievement goals (Scherer et al., 2015). In recent years, the number of interactions with a task and the time spent on a task have served as indicators of these constructs (Goldhammer et al., 2020). In this respect, process data could provide ways of measuring test motivation beyond self-report scales. For instance, Pools and Monseur (2021) studied the relationship between test-taking effort and science achievement for the PISA 2015 data of six English-speaking countries. Utilizing students’ response times as indicators of effort, their study uncovered moderate to strong correlations, with increasing values toward the end of the test. Moreover, typical measures of

expectancy and value did not explain substantial variation in response times, thus questioning their correspondence. Hence, response times—key process data extracted from log files—can inform researchers and evaluators about the effort spent on the tasks and thus facilitate a nuanced interpretation of performance (Kroehne and Goldhammer, 2018).

Improving the estimation of performance scores

The information about test-taking behavior and effort can further inform the estimation of proficiency. Specifically, several indicators extracted from log files, such as response times, the number of actions, or the occurrence of specific strategies, may be used as variables informing the estimation of proficiency scores in the ILSA scaling models (Reis Costa et al., 2021) and modeling the processes underlying missing data (Pohl et al., 2019). For instance, in their study of the PISA 2012 interactive mathematics tasks, Reis Costa et al. (2021) showed that including information about the time students have spent on the computer-based tasks can improve the precision of the ability estimates and make visible country differences in response processes. Integrating process data into the ability estimates is a future direction of research on ILSAs that bears the potential to improve the current scaling models which form the basis for the inferences drawn in the context of accountability (von Davier et al., 2019a).

Comparability of measures

With the increasing number of countries and populations included in ILSAs, establishing the comparability of measures is key to crafting a validity argument (Rutkowski and Rutkowski, 2016). As a consequence, ILSAs have included analytic procedures to test for the invariance of the measurement models of important scales and indicators. These procedures are aimed at ruling out that observed country differences may only be due to the differential functioning of items or scales (Avvisati et al., 2019; Scherer, 2020). While perfect invariance is oftentimes not met, for instance due to culture-specific item formulations, several analytic options have been developed to retain some degree of invariance by including country-specific and common item parameters in the scaling (von Davier et al., 2019b).

Measurement invariance

Measurement invariance refers to the psychometric equivalence of construct representations (i.e., measurement models) and establishes the degree to which constructs have the same meaning across groups and/or over time (Putnick and Bornstein, 2016). Typically, measurement invariance is tested by specifying and estimating a sequence of measurement models with increasing equality constraints on the model parameters. These models are then compared to a baseline model, the *configural invariance model*, which establishes the same setup of the factor structure but freely estimates all model parameters across groups and/or over time (Millsap, 2011). Fig. 4 displays such a baseline model with three latent variables and two groups.

The subsequent invariance models constrain the factor loadings (*metric invariance*), factor loadings and item intercepts (*scalar invariance*), and factor loadings, item intercepts, and item residual variances (*strict invariance*). However, other combinations of constraints are possible and may shed light on additional degrees of invariance, such as the equality of factor means or (co-)variances (Marsh et al., 2009). These models are then compared to the configural invariance model or to other models with less constraints on the basis of goodness-of-fit indices (Chen, 2007; Rutkowski and Svetina, 2013). The respective guidelines and their specific application to ILSAs were discussed by Rutkowski and Svetina (2017). For instance, Scherer et al. (2016) and Bellens et al. (2019) tested the invariance of student ratings of teaching practice measures across countries, differentiating between three dimensions and utilizing the PISA 2012 and TIMSS 2015 data. Evaluating the above-mentioned invariance models, the authors found support for metric invariance and thus argued that structural relations among teaching practices, yet not the average levels (i.e., factor means) could be compared.

Approaches to invariance testing

Given the complexities of ILSA data, typical measurement invariance testing procedures, such as multi-group confirmatory factor analysis (MG-CFA) or multi-group item response theory (MG-IRT) models, reach their limits. As a consequence, psychometricians

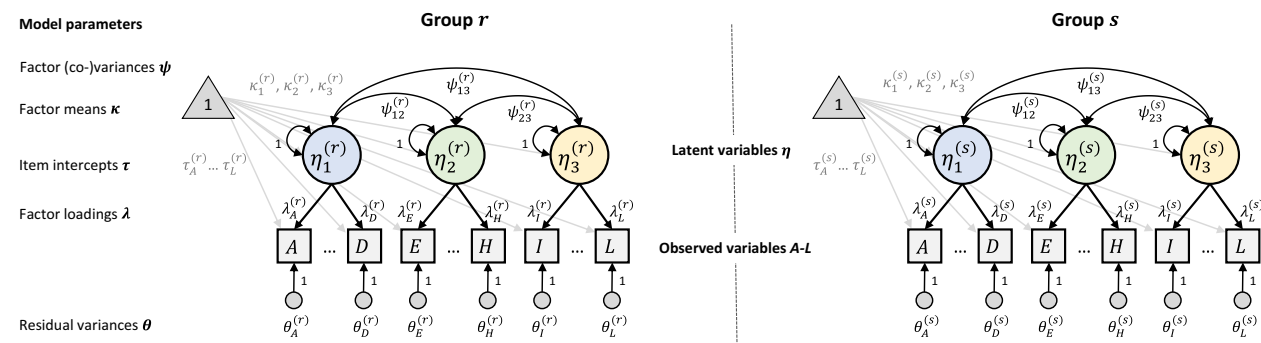


Fig. 4 Multiple-groups measurement model with continuous observed variables.

have explored and developed alternative testing procedures, which may address these complexities. **Table 2** provides an overview of these procedures.

For instance, with many countries participating in ILSAs, MG-CFA and MG-IRT can hardly identify invariance levels beyond metric invariance. Hence, measurement invariance testing with many groups can be performed via multilevel CFA/IRT or alignment methods. [Marsh et al. \(2018\)](#) found that alignment methods can retain the comparability of factor means if scalar invariance fails. Next to these methods, other procedures provide information about item-specific differential item functioning ([Buchholz and Hartig, 2017](#)), the invariance along continuous variables ([Bauer et al., 2020](#)), or the invariance across initially unobserved groups ([Kim et al., 2017](#)). [Scherer \(2020\)](#) reviewed these procedures in greater detail and illustrated their implementation in the context of ILSAs. As the research on measurement invariance testing is continuing, new procedures may be developed which could examine and establish sufficient degrees of construct equivalence across a broad range of countries efficiently in the context of ILSAs. Such procedures are key to interpreting the results of ILSAs for accountability purposes.

Invariance of multilevel constructs

As noted earlier, some constructs have a meaning at multiple levels of analysis (e.g., classroom teaching practices assessed by student ratings). Hence, multilevel measurement models may exist, with model parameters that can vary between countries at different levels of analysis. As a consequence, the procedures of invariance testing have been extended to multilevel situations ([Ryu, 2014](#)). [Scherer and Gustafsson \(2015\)](#) evaluated the invariance of student ratings of teaching practices in reading, mathematics, and science for the TIMSS and PIRLS 2011 data across several countries and developed a testing framework with specific invariance testing steps. These steps include the testing of invariance across countries at the student level, the classroom level, and both levels at the same time.

Besides invariance across countries, another form of invariance is key to establishing the same meaning of constructs across levels of analysis, that is, cross-level measurement invariance ([Jak, 2019](#)). **Fig. 1** shows two models of multilevel confirmatory factor analysis (MCFA) that may describe the structure of multilevel constructs, such as teaching practices in classrooms. The configural cross-level invariance model (a) assumes the same setup of the factor structure across levels, and the metric cross-level measurement invariance model (b) assumes the same factor loadings across levels in addition. In a situation where multilevel constructs are compared across countries, both cross-country and cross-level measurement invariance are important to test and establish (e.g., [Blömeke et al., 2021](#)).

Despite these psychometric perspectives, the main perspective behind measurement invariance testing is to ensure valid inferences drawn from possible differences between countries or levels of analysis. Further research will develop the methodological approaches and assessment designs to improve the invariance testing procedures but also deal with situations in which full invariance cannot be established.

Table 2 Overview of selected invariance testing procedures applied to ILSA data.

Procedure	Description	ILSA examples
Multi-group CFA/IRT	Measurement models with a sequence of equality constraints on model parameters across groups	Rutkowski and Rutkowski (2013)
Multi-group ESEM	Exploratory structural equation models with a sequence of equality constraints on model parameters across groups	Scherer et al. (2016)
Multi-group Bayesian CFA/IRT	Measurement models with a sequence of equality constraints on model parameter distributions across groups and sensitivity analyses of model priors	Braeken and Blömeke (2016)
Alignment optimization	Approximation of scalar invariance by estimating group-specific factor means and variances	Glassow et al. (2021)
Alignment-within-CFA/IRT	Two-step procedure combining alignment optimization and multi-group CFA/IRT	Marsh et al. (2018)
Multilevel CFA/IRT	Multilevel measurement models utilizing the grouping as a level of analysis and testing random effects of model parameters	Fox and Glas (2016)
Moderated nonlinear factor analysis	Differential item functioning tests for (non-) linear moderation effects on model parameters of grouping variables and/or covariates (e.g., MIMIC-DIF models)	Chen and Hwu (2017)
Local SEM	Differential item functioning tests for (non-) linear moderation effects on model parameters at given points of grouping variables and/or covariates	Gnambs and Schroeders (2017)
Factor mixture modeling	Invariance testing for unobserved (latent) groups	Kim et al. (2016)

Note. MIMIC-DIF = Multiple Indicators Multiple Causes Model for Differential Item Functioning testing, SEM = Structural equation modeling.

Concluding remarks

To summarize, ILSAs include various *domains*, *populations*, and *contextual information*, yet could be further extended to improve the crafting of an accountability argument and ultimately provide countries with a more holistic picture of educational outcomes. Concerning the populations, longitudinal extensions and a balanced selection of countries could enhance the information gained from ILSAs substantially, informing educational effectiveness and cultural comparisons. With regards to contextual information, there is a need to further develop the theoretical foundation underlying constructs to be able to make meaningful comparisons across cultures. It is hence important that researchers from different cultures, from both Western and non-Western countries, contribute to developing the contextual questionnaires.

ILSAs provide rich *data sources* have benefitted from the *innovations in assessments* and *methodological advancements*. The novel task and test designs (e.g., innovative items, task formats, and multi-stage adaptive testing) are all aimed at assessing better constructs in the traditional domains and at making assessable new constructs that require interactions between test-takers and tasks, and among test-takers. The shift to computer-based assessments has facilitated the implementation of these innovations and have made available process data that can uncover problem-solving strategies, response behaviors, or test-taking effort. These data inform the crafting of a validity argument and have the potential to improve the estimation of achievement scores.

With these insights, we return to our question in the section: To what extent can ILSAs inform evidence-based decision-making and accountability? What is needed? Which developments can support these processes?

Certain characteristics and strengths of ILSAs enable them to fulfill their purpose on accountability. One such strength is their trend design which allows countries to monitor cognitive and affective outcomes as well as contextual information over time. In addition, comparisons across countries could, to a certain extent, be useful as they may provide some indication of success factors in educational systems. Internationalization may render countries that lag behind on ILSAs less fit for competition in the future, global job market. At the same time, comparisons of both outcomes and contextual information across countries need to be treated with caution. As has been discussed, there are major issues with the non-invariance of measures and indicators due to the heterogeneous sets of countries.

The comparative descriptions, for instance, of outcomes and contextual information over time, reflect the very strength of the ILSAs as the data are of high quality and based on representative samples. However, correlations between predictors and outcomes may not provide useful information on accountability, as drawing causal inferences on the basis of ILSA data is the least appropriate purpose with regards to accountability. That being said, we showcased how causal methods (e.g., differences in differences), longitudinal extensions of ILSAs, meta-analyses of ILSAs, and linking ILSAs with other data may enhance the validity and robustness of inferences. To this end, even correlations may be useful, especially with methods that allow for stronger and more robust inferences (e.g., multi-level SEM, differences in differences, or propensity score matching), but need to be interpreted with care. Here, ILSAs have a responsibility to inform stakeholders on how to interpret results, and researchers using ILSA data are responsible for valid and correct reports and interpretations of their findings. However, ILSAs should continue to play an important role for informing the accountability of educational systems, but not alone.

References

- Addey, C., Sellar, S., Steiner-Khamsi, G., Lingard, B., Verger, A., 2017. The rise of international large-scale assessments and rationales for participation. *Compare* 47, 434–452.
- Afsahi, A., Beausoleil, E., Dean, R., Ercan, S.A., Gagnon, J.-P., 2020. Democracy in a global emergency: five lessons learned from the COVID-19 pandemic. *Democr. Theory* 7, v–xix.
- Aldrich, J., 1995. Correlations genuine and spurious in Pearson and Yule. *Stat. Sci.* 10, 364–376, 13.
- Avisati, F., Le Donne, N., Paccagnella, M., 2019. A meeting report: cross-cultural comparability of questionnaire measures in large-scale international surveys. *Meas. Instr. Soc. Sci.* 1, 8.
- Barth, E., Keute, A.L., Schöne, P., Von Simson, K., Steffensen, K., 2021. NEET status and early versus later skills among young adults: evidence from linked register-PIAAC data. *Scand. J. Educ. Res.* 65, 140–152.
- Bauer, D.J., Belzak, W.C.M., Cole, V.T., 2020. Simplifying the assessment of measurement invariance over multiple background variables: using regularized moderated nonlinear factor analysis to detect differential item functioning. *Struct. Equ. Model.* 27, 43–55.
- Beaton, A.E., 1996. Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study (TIMSS). ERIC.
- Bellens, K., Van Damme, J., Van Den Noortgate, W., Wendt, H., Nilsen, T., 2019. Instructional quality: catalyst or pitfall in educational systems' aim for high achievement and equity? An answer based on multilevel SEM analyses of TIMSS 2015 data in Flanders (Belgium), Germany, and Norway. *Large Scale Assess. Educ.* 7.
- Beller, M., 2013. Technologies in large-scale assessments: new directions, challenges, and opportunities. In: Von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K. (Eds.), *The Role of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Springer Netherlands, Dordrecht.
- Blömeke, S., Nilsen, T., Olsen, R.V., Gustafsson, J.E., 2022. Conceptual and methodological accomplishments of ILSAs, remaining criticism and limitations. In: Nilsen, T., Stancul-Piatak, A., Gustafsson, J.E. (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer International Handbooks of Education.
- Blömeke, S., Nilsen, T., Scherer, R., 2021. School innovativeness is associated with enhanced teacher collaboration, innovative classroom practices, and job satisfaction. *J. Educ. Psychol.* 113, 1645–1667.
- Blömeke, S., Olsen, R.V., 2019. Consistency of results regarding teacher effects across subjects, school levels, outcomes and countries. *Teach. Teach. Educ.* 77, 170–182.
- Braeken, J., Blömeke, S., 2016. Comparing future teachers' beliefs across countries: approximate measurement invariance with Bayesian elastic constraints for local item dependence and differential item functioning. *Assess. Eval. High Educ.* 41, 733–749.
- Braun, H., Von Davier, M., 2017. The use of test scores from large-scale assessment surveys: psychometric and statistical considerations. *Large Scale Assess. Educ.* 5.
- Braun, H.I., Singer, J.D., 2019. Assessment for monitoring of education systems: international comparisons. *Ann. Am. Acad. Polit. Soc. Sci.* 683, 75–92.
- Buchholz, J., Hartig, J., 2017. Comparing attitudes across groups: an IRT-based item-fit statistic for the analysis of measurement invariance. *Appl. Psychol. Meas.* 43, 241–250.

- Carnoy, M., Khavenson, T., Loyalka, P., Schmidt, W.H., Zakharov, A., 2016. Revisiting the relationship between international assessment outcomes and educational production: evidence from a longitudinal PISA-TIMSS sample. *Am. Educ. Res. J.* 53, 1054–1085.
- Chen, C.-T., Hwu, B.-S., 2017. Improving the assessment of differential item functioning in large-scale programs with dual-scale purification of Rasch models: the PISA example. *Appl. Psychol. Meas.* 42, 206–220.
- Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* 14, 464–504.
- Cheung, M.W.-L., Jak, S., 2016. Analyzing big data in psychology: a split/analyze/meta-analyze approach. *Front. Psychol.* 7.
- Creemers, B., Kyriakides, L., 2008. *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. Routledge, Abingdon.
- Cresswell, J., Schwanter, U., Waters, C., 2015. *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data*. OECD Publishing, Paris.
- Cziboly, A. (Ed.), 2010. *The DICE Has Been Cast. A DICE Resource - Research Findings and Recommendations on Educational Theatre and Drama*. DICE Consortium, Budapest.
- Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132.
- Ehmke, T., Van Den Ham, A.-K., Sälzer, C., Heine, J., Prenzel, M., 2020. Measuring mathematics competence in international and national large scale assessments: linking PISA and the national educational panel study in Germany. *Stud. Educ. Eval.* 65, 100847.
- Eriksson, K., Helenius, O., Ryve, A., 2019. Using TIMSS items to evaluate the effectiveness of different instructional practices. *Instr. Sci.* 47, 1–18.
- Ferguson, R.F., Danielson, C., 2015. How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In: Kane, T., Kerr, K.A., Pianta, R.C. (Eds.), *Designing Teacher Evaluation Systems: New Guidance From the Measures of Effective Teaching Project*.
- Fernández-Castilla, B., Maes, M., Declercq, L., Jamshidi, L., Beretvas, S.N., Onghena, P., Van Den Noortgate, W., 2019. A demonstration and evaluation of the use of cross-classified random-effects models for meta-analysis. *Behav. Res. Methods* 51, 1286–1304.
- Fox, J.-P., Glas, C.A., 2016. Multilevel response models with covariates and multiple groups. In: Linden, W.J.V.D. (Ed.), *Handbook of Item Response Theory*. Chapman and Hall/CRC, New York, NY.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Duckworth, D., 2020. *Preparing for Life in a Digital World: IEA International Computer and Information Literacy Study 2018 International Report*. Springer Nature.
- Gil-Izquierdo, M., Cordero, J.M., 2018. Guidelines for data fusion with international large scale assessments: insights from the TALIS-PISA link database. *Stud. Educ. Eval.* 59, 10–18.
- Gjærum, R., Cziboly, A., Eriksson, S., 2022. ILSA in arts education: the effect of drama on competences. In: Nilsen, T., Stancel-Piatak, A., Gustafsson, J.E. (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer International Handbooks of Education, Cham.
- Glassow, L.N., Rolfe, V., Hansen, K.Y., 2021. Assessing the comparability of teacher-related constructs in TIMSS 2015 across 46 education systems: an alignment optimization approach. *Educ. Assess. Eval. Account.* 33, 105–137.
- Gnamb, T., Schroeders, U., 2017. Cognitive abilities explain wording effects in the Rosenberg self-esteem scale. *Assessment* 27, 404–418.
- Goldhammer, F., Hahnel, C., Kroehne, U., 2020. Analysing log file data from PIAAC. In: Maehler, D.B., Rammstedt, B. (Eds.), *Large-Scale Cognitive Assessment: Analyzing PIAAC Data*. Springer International Publishing, Cham.
- Goldhammer, F., Zehner, F., 2017. What to make of and how to interpret process data. *Measurement* 15, 128–132.
- Greiff, S., Scherer, R., Kirschner, P.A., 2017. Some critical reflections on the special issue: current innovations in computer-based assessments. *Comput. Hum. Behav.* 76, 715–718.
- Greiff, S., Wüstenberg, S., Avisati, F., 2015. Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Comput. Educ.* 91, 92–105.
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., Csapó, B., 2013. Complex problem solving in educational contexts—something beyond g: concept, assessment, measurement invariance, and construct validity. *J. Educ. Psychol.* 105, 364–379.
- Griffin, P., Care, E., 2014. *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer.
- Gustafsson, J.-E., 2018. International large scale assessments: current status and ways forward. *Scand. J. Educ. Res.* 62, 328–332.
- Gustafsson, J.E., 2013. Causal inference in educational effectiveness research: a comparison of three methods to investigate effects of homework on student achievement 1. *Sch. Effect. Sch. Improv.* 24, 275–295.
- Gustafsson, J.E., Nilsen, T., 2016. The impact of school climate and teacher quality on mathematics achievement: a difference-in-differences approach. In: Nilsen, T., Gustafsson, J.-E. (Eds.), *Teacher Quality, Instructional Quality and Student Outcomes: Relationships Across Countries, Cohorts and Time*. Springer International Publishing, Cham.
- Gustafsson, J.E., Nilsen, T., 2022. Methods of causal analysis with ILSA data. In: Nilsen, T., Stancel-Piatak, A., Gustafsson, J.E. (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer International Handbooks of Education, Cham.
- Hämäläinen, R., Nissinen, K., Mannonen, J., Lämsä, J., Leino, K., Taajamo, M., 2021. Understanding teaching professionals' digital competence: what do PIAAC and TALIS reveal about technology-related skills, attitudes, and knowledge? *Comput. Hum. Behav.* 117, 106672.
- Han, Z., He, Q., Von Davier, M., 2019. Predictive feature generation and selection using process data from PISA interactive problem-solving items: an application of random forests. *Front. Psychol.* 10.
- Hedges, L.V., Schauer, J.M., 2019. Statistical analyses for studying replication: meta-analytic perspectives. *Psychol. Methods* 24, 557–570.
- Hooper, M., Mullis, I.V.S., Martin, M.O., 2013. TIMSS 2015 context questionnaire framework. In: Mullis, I.V.S., Martin, M.O. (Eds.), *TIMSS 2015 Assessment Frameworks*. TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, MA.
- IEA, 2017. *ILSA Gateway* [Online]. Available: <https://ilsa-gateway.org/>. (Accessed 28 May 2021).
- IEA, 2021. *TIMSS: Trends in International Mathematics and Science Study* [Online]. Available: <https://www.iea.nl/studies/iea/timss>. (Accessed 28 May 2021).
- Jacobs, J.K., Hollingsworth, H., Givvin, K.B., 2007. Video-based research made “easy”: methodological lessons learned from the TIMSS video studies. *Field Methods* 19, 284–299.
- Jak, S., 2019. Cross-level invariance in multilevel factor models. *Struct. Equ. Model.* 26, 607–622.
- Jerrim, J., Oliver, M., Sims, S., 2019. The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learn. Instruct.* 35–44.
- Kamens, D.H., Benavot, A., 2011. National, regional and international learning assessments: trends among developing countries, 1960–2009. *Glob. Soc. Educ.* 9, 285–300.
- Kampa, N., Scherer, R., Sass, S., Schipolowski, S., 2021. The relation between science achievement and general cognitive abilities in large-scale assessments. *Intelligence* 86, 101529.
- Kane, T., Cantrell, S., 2010. *Learning About Teaching: Initial Findings From the Measures of Effective Teaching Project*. MET Project Research Paper, vol. 9. Bill & Melinda Gates Foundation.
- Kaplan, D., Mccarty, A.T., 2013. Data fusion with international large scale assessments: a case study using the OECD PISA and TALIS surveys. *Large Scale Assess. Educ.* 1.
- Keller, L., Preckel, F., Brunner, M., 2021. Nonlinear relations between achievement and academic self-concepts in elementary and secondary school: an integrative data analysis across 13 countries. *J. Educ. Psychol.* 113, 585–604.
- Kim, E.S., Cao, C., Wang, Y., Nguyen, D.T., 2017. Measurement invariance testing with many groups: a comparison of five approaches. *Struct. Equ. Model.* 24, 524–544.
- Kim, E.S., Joo, S.-H., Lee, P., Wang, Y., Stark, S., 2016. Measurement invariance testing across between-level latent classes using multilevel factor mixture modeling. *Struct. Equ. Model.* 23, 870–887.
- Kirsch, I., Lennon, M.L., 2017. PIAAC: a new design for a new era. *Large Scale Assess. Educ.* 5.
- Klette, K., 2016. Introduction: studying interaction and instructional patterns in classrooms. In: Klette, K., Bergem, O.K., Roe, A. (Eds.), *Teaching and Learning in Lower Secondary Schools in the Era of PISA and TIMSS*. Springer International Publishing, Cham.

- Klieme, E., 2013. The role of large-scale Assessments in research on educational effectiveness and school development. In: Von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K. (Eds.), *The Role of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Springer Netherlands, Dordrecht.
- Klieme, E., 2020. Policies and practices of assessment: a showcase for the use (and misuse) of international large scale assessments in educational effectiveness research. In: Hall, J., Lindorff, A., Sammons, P. (Eds.), *International Perspectives in Educational Effectiveness Research*. Springer, Cham.
- Kroehne, U., Goldhammer, F., 2018. How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika* 45, 527–563.
- Kuger, S., Klieme, E., Jude, N., Kaplan, D., 2016. *Assessing Contexts of Learning*. Springer International Publishing, Cham.
- Kyriakides, L., Campbell, R., Gagatsis, A., 2000. The significance of the classroom effect in primary schools: an application of Creemers' comprehensive model of educational effectiveness. *Sch. Effect. Sch. Improv.* 11, 501–529.
- Levin, H.M., 1974. A conceptual framework for accountability in education. *Sch. Rev.* 82, 363–391.
- Magis, D., Yan, D., Von Davier, A.A., 2017. *Computerized Adaptive and Multistage Testing With R: Using Packages CatR and MstR*. Springer International Publishing, Cham.
- Marsh, H.W., Guo, J., Parker, P.D., Nagengast, B., Asparouhov, T., Muthén, B., Dicke, T., 2018. What to do when scalar invariance fails: the extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychol. Methods* 23, 524–545.
- Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A.J.S., Abduljabbar, A.S., Köller, O., 2012. Classroom climate and contextual effects: conceptual and methodological issues in the evaluation of group-level effects. *Educ. Psychol.* 47, 106–124.
- Marsh, H.W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A.J.S., Trautwein, U., 2009. Exploratory structural equation modeling, integrating CFA and EFA: application to students' evaluations of university teaching. *Struct. Equ. Model.* 16, 439–476.
- Martin, M.O., Mullis, I.V.S., 2013. *TIMSS and PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade—Implications for Early Learning*. Timss & PIRLS International Study Center, Lynch School of Education, Boston College, and International Association for the Evaluation of Educational Achievement (IEA), Chestnut Hill, MA.
- Medrich, E.A., Griffith, J.E., 1992. *International Mathematics and Science Assessments: What Have We Learned?* US Department of Education, Office of Educational Research and Improvement.
- Meroni, E.C., Vera-Toscano, E., Costa, P., 2015. Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. *J. Pol. Model.* 37, 308–323.
- Millap, R.E., 2011. *Statistical Approaches to Measurement Invariance*. Routledge/Taylor & Francis Group, New York, NY, US.
- Mullis, I.V.S., Fishbein, B., 2020. Updating the TIMSS 2019 instruments for describing the contexts for student learning. In: Martin, M.O., Von Davier, M., Mullis, I.V.S. (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*. Boston College, TIMSS & PIRLS International Study Center, Boston.
- Mullis, I.V.S., Martin, M.O., 2017. *TIMSS 2019 Assessment Frameworks* [Online]. Boston. Available: <http://timssandpirls.bc.edu/timss2019/frameworks/>. (Accessed 28 May 2021).
- Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D., Fishbein, B., 2020. *TIMSS 2019 International Results in Mathematics and Science* [Online]. TIMSS & PIRLS International Study Center, Boston, USA. Retrieved from: <https://timssandpirls.bc.edu/timss2019/international-results/>. (Accessed 28 May 2021).
- Nagy, G., Lüdtke, O., Köller, O., 2016. Modeling test context effects in longitudinal achievement data: examining position effects in the longitudinal German PISA 2012 assessment. *Psychol. Test Assess. Model.* 58, 641–670.
- Neave, G., 1987. *Accountability in education*. In: Psacharopoulos, G. (Ed.), *Economics of Education*. Pergamon.
- Neubrand, J., 2006. The TIMSS 1995 and 1999 video studies. In: Leung, F.K.S., Graf, K.-D., Lopez-Real, F.J. (Eds.), *Mathematics Education in Different Cultural Traditions—A Comparative Study of East Asia and the West: The 13th ICMI Study*. Springer US, Boston, MA.
- Nilsen, T., Gustafsson, J.-E., 2014. School emphasis on academic success: exploring changes in science performance in Norway between 2007 and 2011 employing two-level SEM. *Educ. Res. Eval.* 20, 308–327.
- Nilsen, T., Gustafsson, J.-E., 2016. *Teacher Quality, Instructional Quality and Student Outcomes: Relationships across Countries, Cohorts and Time*. Springer, Cham.
- Nilsen, T., Stancel-Piatak, A., Gustafsson, J.E. (Eds.), 2022. *International Handbook of Comparative Large-Scale Studies in Education*. Springer, Cham.
- OECD, 2013. *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing, Paris.
- OECD, 2014. *PISA 2012 Results: Creative Problem Solving*, vol. V. OECD Publishing, Paris.
- OECD, 2015. *Students, Computers and Learning*. OECD Publishing, Paris.
- OECD, 2018. *What does teaching look like?* Teach. Focus 20.
- OECD, 2019a. *PISA 2018 Assessment and Analytical Framework*.
- OECD, 2019b. *TALIS 2018 Results. In: Teachers and School Leaders as Lifelong Learners*, vol. I. OECD Publishing, Paris.
- OECD, 2020a. *Global Teaching Insights*. OECD Publishing, Paris.
- OECD, 2020b. *PISA 2018 Results. In: Are Students Smart About Money?*, vol. IV.
- OECD, 2021. *The State of School Education*. OECD Publishing, Paris.
- Pellegrino, J.W., Dibello, L.V., Goldman, S.R., 2016. A framework for conceptualizing and evaluating the validity of instructionally relevant assessments. *Educ. Psychol.* 51, 59–81.
- Pohl, S., Ulitzsch, E., Von Davier, M., 2019. Using response times to model not-reached items due to time limits. *Psychometrika* 84, 892–920.
- Pools, E., Monseur, C., 2021. Student test-taking effort in low-stakes assessments: evidence from the English version of the PISA 2015 science test. *Large Scale Assess. Educ.* 9, 10.
- Praetorius, A.-K., Klieme, E., Herbert, B., Pinger, P., 2018. Generic dimensions of teaching quality: the German framework of three basic dimensions. *ZDM* 50, 407–426.
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., Klieme, E., 2014. One lesson is all you need? Stability of instructional quality across lessons. *Learn. Instr.* 31, 2–12.
- Putnick, D.L., Bornstein, M.H., 2016. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev. Rev.* 41, 71–90.
- Qiao, X., Jiao, H., 2018. Data mining techniques in analyzing process data: a didactic. *Front. Psychol.* 9.
- Rammstedt, B., Martin, S., Zabal, A., Carstensen, C., Schupp, J., 2017. The PIAAC longitudinal study in Germany: rationale and design. *Large Scale Assess. Educ.* 5.
- Reis Costa, D., Bolsinova, M., Tijmstra, J., Andersson, B., 2021. Improving the precision of ability estimates using time-on-task variables: insights from the PISA 2012 computer-based assessment of mathematics. *Front. Psychol.* 12.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., Stringfield, S., 2014. *Educational effectiveness research (EER): a state-of-the-art review*. Sch. Eff. Sch. Improv. 25, 197–230.
- Robitzsch, A., Lüdtke, O., 2019. Linking errors in international large-scale assessments: calculation of standard errors for trend estimation. *Assess Educ. Princ. Pol. Pract.* 26, 444–465.
- Rutkowski, D., Delandshere, G., 2016. Causal inferences with large scale assessment data: using a validity framework. *Large Scale Assess. Educ.* 4, 1–18.
- Rutkowski, D., Rutkowski, L., 2013. Measuring socioeconomic background in PISA: one size might not fit all. *Res. Comp. Int. Educ.* 8, 259–278.
- Rutkowski, D., Rutkowski, L., Liaw, Y.-L., 2018. Measuring widening proficiency differences in international assessments: are current approaches enough? *Educ. Meas.* 37, 40–48.
- Rutkowski, L., Gonzalez, E., Joncas, M., Von Davier, M., 2010. International large-scale assessment data: issues in secondary analysis and reporting. *Educ. Res.* 39, 142–151.
- Rutkowski, L., Rutkowski, D., 2016. A call for a more measured approach to reporting and interpreting PISA results. *Educ. Res.* 45, 252–257.
- Rutkowski, L., Rutkowski, D., 2018. Improving the comparability and local usefulness of international assessments: a look back and a way forward. *Scand. J. Educ. Res.* 62, 354–367.
- Rutkowski, L., Rutkowski, D., 2019. Methodological challenges to measuring heterogeneous populations internationally. In: Suter, L., Smith, E., Denman, B. (Eds.), *The SAGE Handbook of Comparative Studies in Education*. SAGE, London, pp. 126–140.
- Rutkowski, L., Svetina, D., 2013. Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educ. Psychol. Meas.* 74, 31–57.
- Rutkowski, L., Svetina, D., 2017. Measurement invariance in international surveys: categorical indicators and fit measure performance. *Appl. Meas. Educ.* 30, 39–51.

- Rutkowski, L., Von Davier, M., Rutkowski, D., 2013. *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, Boca Raton, FL.
- Ryu, E., 2014. Factorial invariance in multilevel confirmatory factor analysis. *Br. J. Math. Stat. Psychol.* 67, 172–194.
- Scheerens, J., 2016. *Opportunity to Learn, Curriculum Alignment and Test Preparation: A Research Review*. Springer.
- Scherer, R., 2020. Analysing PIAAC data with structural equation modelling in Mplus. In: Maehler, D.B., Rammstedt, B. (Eds.), *Large-Scale Cognitive Assessment: Analyzing PIAAC Data*. Springer International Publishing, Cham.
- Scherer, R., Greiff, S., Hautamäki, J., 2015. Exploring the relation between time on task and ability in complex problem solving. *Intelligence* 48, 37–50.
- Scherer, R., Gustafsson, J.-E., 2015. Student assessment of teaching as a source of information about aspects of teaching quality in multiple subject domains: an application of multilevel bifactor structural equation modeling. *Front. Psychol.* 6.
- Scherer, R., Nilsen, T., Jansen, M., 2016. Evaluating individual students' perceptions of instructional quality: an investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Front. Psychol.* 7.
- Schulz, W., Ainley, J., Cox, C., Friedman, T., 2018. *Young People's Views of Government, Peaceful Coexistence, and Diversity in Five Latin American Countries: IEA International Civic and Citizenship Education Study 2016 Latin American Report*. Springer Nature, Cham.
- Schulz, W., Carstens, R., 2020. Questionnaire development in international large-scale assessment studies. In: Wagemaker, H. (Ed.), *Reliability and Validity of International Large-Scale Assessment: Understanding IEA's Comparative Studies of Student Achievement*. Springer International Publishing, Cham.
- Stancel-Piatak, A., Schwippert, K., 2022. Comprehensive frameworks of school learning in ILSAs. In: Nilsen, T., Stancel-Piatak, A., Gustafsson, J.E. (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer International Handbooks of Education, Cham.
- Stapleton, L.M., Yang, J.S., Hancock, G.R., 2016. Construct meaning in multilevel settings. *J. Educ. Behav. Stat.* 41, 481–520.
- Stigler, J.W., Gallimore, R., Hiebert, J., 2000. Using video surveys to compare classrooms and teaching across cultures: examples and lessons from the TIMSS video studies. *Educ. Psychol.* 35, 87–100.
- Strello, A., Strietholt, R., Steinmann, I., Siepmann, C., 2021. Early tracking and different types of inequalities in achievement: difference-in-differences evidence from 20 years of large-scale assessments. *Educ. Assess. Eval. Account.* 33, 139–167.
- Strietholt, R., Scherer, R., 2018. The contribution of international large-scale assessments to educational research: combining individual and institutional data sources. *Scand. J. Educ. Res.* 62, 368–385.
- Svetina, D., Liaw, Y.-L., Rutkowski, L., Rutkowski, D., 2019. Routing strategies and optimizing design for multistage testing in international large-scale assessments. *J. Educ. Meas.* 56, 192–213.
- Teig, N., Nilsen, T., 2022. A systematic review of studies investigating the relationships between school climate and student outcomes in Timss, Pisa, and PIRLS. In: Nilsen, T., Stancel-Piatak, A., Gustafsson, J.E. (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Cham: Springer International Handbooks of Education.
- Teig, N., Scherer, R., Kjærnsli, M., 2020. Identifying patterns of students' performance on simulated inquiry tasks using PISA 2015 log-file data. *J. Res. Sci. Teach.* 57, 1400–1429.
- Ulitzsch, E., He, Q., Ulitzsch, V., Molter, H., Nichterlein, A., Niedermeier, R., Pohl, S., 2021. Combining clickstream analyses and graph-modeled data clustering for identifying common response processes. *Psychometrika* 86, 190–214.
- Van De Vijver, F.J., 2018. Towards an integrated framework of bias in noncognitive assessment in international large-scale studies: challenges and prospects. *Educ. Meas.* 37, 49–56.
- Von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K., 2012. *The Role of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Springer, Cham.
- Von Davier, M., Khorramdel, L., He, Q., Shin, H.J., Chen, H., 2019a. Developments in psychometric population models for technology-based large-scale assessments: an overview of challenges and opportunities. *J. Educ. Behav. Stat.* 44, 671–705.
- Von Davier, M., Yamamoto, K., Shin, H.J., Chen, H., Khorramdel, L., Weeks, J., Davis, S., Kong, N., Kandathil, M., 2019b. Evaluating item response theory linking and model fit for data from PISA 2000–2012. *Assess. Educ. Princ. Pol. Pract.* 26, 466–488.
- Wagemaker, H., 2014. International large-scale assessments: from research to policy. In: Rutkowski, L., Von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, Boca Raton, FL, pp. 11–36.
- Wagemaker, H., 2020. *Reliability and Validity of International Large-Scale Assessment: Understanding IEA's Comparative Studies of Student Achievement*. Springer Nature, Cham.
- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., Trautwein, U., 2016. Student and teacher ratings of instructional quality: consistency of ratings over time, agreement, and predictive power. *J. Educ. Psychol.* 108, 705–721.
- Wang, M.-T., Degol, J.L., 2015. School climate: a review of the construct, measurement, and impact on student outcomes. *Educ. Psychol. Rev.* 1–38.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2018. Multistage adaptive testing design in international large-scale assessments. *Educ. Meas.* 37, 16–27.

FOURTH EDITION

International Encyclopedia of Education

Quantitative Research
and Educational Measurement

VOLUME
14

Editors-In-Chief

Robert J. Tierney

Fazal Rizvi

Kadriye Ercikan

Volume Editor

Daniel McCaffrey



INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

This page intentionally left blank

INTERNATIONAL ENCYCLOPEDIA OF EDUCATION

FOURTH EDITION

EDITORS IN CHIEF

Robert J Tierney

*The University of British Columbia, Vancouver, Canada,
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China*

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

*Vice President for Research and Measurement Sciences
Educational Testing Service, Princeton, NJ, United States*

VOLUME 14

Quantitative Research and Educational Measurement

VOLUME 14 EDITOR

Daniel McCaffrey



Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge MA 02139, United States

Copyright © 2023 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-818630-5

For information on all publications visit our website at
<http://store.elsevier.com>.



Publisher: Oliver Walter
Acquisition Editors: Fiona Geraghty
Content Project Manager: Paula Davies
Associate Content Project Manager: Greetal Carolyn, Gayathri S, Ebin Clinton, Abraham Lincoln and Rajeswari R
Designer: Mark Rogers

EDITORS IN CHIEF

Robert J Tierney

The University of British Columbia, Vancouver, Canada
University of Sydney, Sydney, Australia and
Beijing Normal University, Beijing, China

Fazal Rizvi

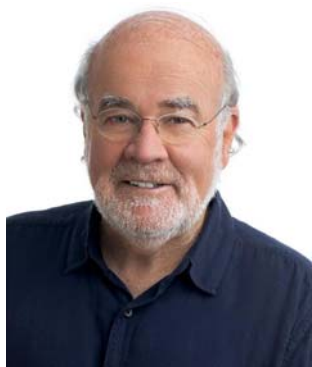
The University of Melbourne, Melbourne, Australia

Kadriye Ercikan

Vice President for Research and Measurement Sciences, Educational Testing Service, Princeton, NJ, United States

This page intentionally left blank

EDITORS IN CHIEF – BIOGRAPHIES



Robert J Tierney

Rob Tierney is an international educator whose passion is for developing research partnerships to address local literacy needs with educators in different countries. Rob began his career as a classroom teacher in Australia, then proceeded to work in the United States, Canada, Australia and China. He is Dean Emeritus and Professor Emeritus of Language and Literacy Education at the University of British Columbia, the former Dean and Honorary Professor of the Faculty of Education and Social Work at University of Sydney, and a Distinguished University Professor at Beijing Normal University. In the United States, he has served on the faculties of the University of Illinois, University of Arizona, The Ohio State University, Harvard University, and the University of California. Rob has written a number of books and numerous articles on literacy especially related to global issues, diversity, and literacy development. His most recent book was entitled *A History of Literacy Education: Waves of Research and Practice* published by Teachers College Press in 2021. During his tenure as an educator, he has served as the President of the International Literacy Association, President of the Literacy Research Association, an Editor for the *Reading Research Quarterly*, the President of the Association of Canadian Deans of Education/ L'Association Canadienne des Doyens et Doyennes d'éducation, as well on a number of committees and boards for professional groups and government and nongovernment

agencies. In addition, he has been involved on a range of groups on projects globally including in Africa and Asia including for UNESCO, Children's Television Workshop, and Apple Computer.



Fazal Rizvi

Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely, and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge 2010) is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge 2022) has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.

**Kadriye Ercikan**

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and the NAEP contract for ETS. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation, and use of scores from artificial intelligence-based automated scoring.

Ercikan is a Fellow of the International Academy of Education. Her research has resulted in 6 books, 4 special issues of refereed journals, and over 100 publications. She was awarded the AERA Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another

coedited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative*

Polarization, and received an Early Career Award from the University of British Columbia. Ercikan has been selected to serve as the NCME Book Series Editor (2021–2026).

EDITORIAL BOARD

Jason Beech

School of Education Culture and Society, Faculty of Education, Monash University, Melbourne, Australia

Qing Gu

UCL Centre for Educational Leadership, UCL Institute of Education, University College London (UCL), London, United Kingdom

Graham Hingangaroa Smith

Distinguished Professorial Chair, Te Toi Ihorei ki Purehuroa Massey University, Palmerston North. New Zealand

Therese Hopfenbeck

University of Melbourne, Melbourne, Australia

Nataliya Ivankova

The University of Alabama at Birmingham, Alabama, USA

Huia Jahnke

Professor of Māori and Indigenous Education in Te Pūtahi a Toi School of Māori Studies at Massey University, Palmerston North New Zealand

Berit Karseth

Professor at the University of Oslo, Department of Education. Oslo, Norway

Kenneth Leithwood

University at Buffalo, State University of New York, USA

Xiufeng Liu

Emeritus Professor, university of Toronto, Toronto, Canada

Margaret Maaka

Professor, College of Education, University of Hawai'i at Mānoa, Hawai'i, USA

Diane Mayer

Professor Emeritus, University of Oxford, UK

Dan McCaffrey

Educational Testing Service, Princeton, NJ, United States

Judy Parr

University of Auckland, Auckland, New Zealand

Jeong-eun Rhee

Long Island University, New York, USA

Fazal Rizvi

The University of Melbourne, Melbourne, Australia

Theresa Rogers

Professor, Faculty of Education, University of British Columbia, Canada

Chika Sehoole

University of Pretoria, Pretoria, South Africa

Roland Sintos Coloma

Wayne State University, Michigan, USA

Roger Slee

Diamond Jubilee Chair of Disability & Inclusion, School of Sociology & Social Policy, University of Leeds, UK

Ninni Wahlström

Linnaeus University, Växjö, Sweden

Lei Wang

Beijing Normal University, Beijing, China

Everard Weber

Research Associate, South African Chair in Teaching and Learning—UJ Academic Advisor, Future Professors Programme—UJ University of Johannesburg (UJ), Johannesburg, South Africa

Ian Wilkinson

The Ohio State University, University of Auckland, New Zealand

David Yaden

Dept. of Teaching, Learning and Sociocultural Studies, College of Education, University of Arizona, USA

SECTION EDITORS



Jason Beech is Associate Professor (Global Policy) at the University of Melbourne and visiting professor at Universidad de San Andrés in Buenos Aires where he holds a UNESCO Chair in Education for Sustainability and Global Citizenship. He is senior researcher of the National Council for Scientific and Technical Research of Argentina (CONICET), and Associate Editor of *Education Policy Analysis Archives*. He has been a member of the Board of Directors of the Comparative and International Education Society (2014-2016). Jason has published extensively on the globalization of knowledge and policies related to education. He has also written and is passionate about the challenges of educating for global citizenship and a sustainable future. He has taught in several universities in the Americas, Europe and Australia.



Professor Qing Gu is Director of the UCL Centre for Educational Leadership and Professor of Leadership in Education. She is the Past Chair of the *British Association of Comparative and International Education* (BAICE), Co-Editor of *Teachers and Teaching: Theory and Practice*, Associate Editor of *International Journal of Educational Development*, a member of the Research Standing Committee of the World Council of Comparative Education Societies (WCCES), and a member of the Research Evidence and Impact Panel for the Leadership College for UK Government. Professor Gu has directed and codirected many government and research council funded projects in the areas of teacher professional development, school improvement, and systemic reform and change. She has led a £1.9 m UK Economic and Social Research Council (ESRC) funded project *Schools as Enabling Spaces to Improve Learning and Health-Related Quality of Life for Primary School Children in Rural Communities in South Africa*. Some of her books have been translated into Chinese, Japanese, and Spanish.



Professor Smith is a New Zealand Maori of Te Aitanga a Hauiti, Ngati Porou, Kati MamoeNgati Apa, Ngati Kahungunu descent. He is a highly regarded and internationally recognized, Maori educationalist and scholar. He has been at the forefront of transforming Maori and Indigenous social, economic and cultural conditions through re-developing education and schooling outcomes. His work links theoretical thinking and practical applications within an ongoing cycle of transforming praxis.

Professor Smith trained as a teacher at Auckland Teachers College and taught in Auckland schools. He also lectured in Education at the Auckland College of Education. He was one of the first teachers and developers of a Kura Kaupapa Maori, an alternative, Maori philosophy and principles-based School. These Maori immersion schools have grown from a single entity in 1988 to over seventy-five publicly funded schools in 2015. His theoretical leadership in the Education Department at the University of Auckland helped inform the emergence of Maori Education as a distinct field of study across the New Zealand Tertiary Sector. This work has encouraged a wide range of academic studies focused on overthrowing persisting inequities within and as a result of education and schooling in New Zealand. His significant contribution to New Zealand education is to be seen in his work related to Kaupapa Maori as a theory of Transformative Praxis. All of these elements derive from his PhD research entitled, "The Development of Kaupapa Maori: Theory and Praxis (1997)".

Professor Smith's later academic work centered on developing theoretically informed strategies related to transforming high and disproportionate levels of Maori cultural, political, social, educational, and economic marginalization. He has been a significant scholar involved in the emergence of Tribal "university" institutions (Wananga) and is the former foundation Chair of the Council for Te Whare Wananga O Awanuiarangi. He led the successful Treaty of Waitangi Claim and settlement for Awanuiarangi. In his former position as Pro Vice Chancellor (Maori) at the University of Auckland, he was responsible for developing significant structural change within the institution while working within the Office of the then Vice Chancellor, Dr. John Hood. Notable achievements included implementing a Maori development strategy within the University which required all Deans to be responsive in their performance results for Maori and Pacific advancement; the oversight of the Proposal and winning of a National Centre of Research Excellence (Nga Pae o te Maramatanga); the initiation of the Maori and Indigenous doctoral program aimed at creating 500 Maori PhD graduates across the country in five years; the recruitment of top Maori academic scholars into the University of Auckland; the development of Kaupapa Maori as a theory of transformative praxis and which has had national and international influence. Professor Smith also served on the University of Auckland Council for three years as the Professorial representative for Academic staff.

In 2007 Professor Smith accepted a role as the Universitas 21 Distinguished Visiting Professor in Indigenous Education hosted by the University of British Columbia (within the Faculty of Education). Although the position was initially intended for two years, the role was extended for a further three years. While situated at the University of British Columbia, Professor Smith also chaired the Educational Policy Studies Department for three years. These positions allowed engagement with many Canadian universities as well as other universities across the world.

Professor Smith has made significant contributions to the political, social, economic, and cultural advancement of Maori and indigenous communities in New Zealand and around the Pacific Rim. He has worked extensively with other Native/Indigenous/First Nation's peoples across the world, including in Canada, India, US mainland (including extensive work in Hawaii and Alaska), Norway, Taiwan, Chile, Australia, as well as with selected Micronesian and Pacific nations. He is a regular contributor to national and international forums and conferences. Professor Smith has been an authoritative voice within the critical debates on *social, cultural and economic inequality* both in New Zealand and abroad, and he maintains a strong influence in the Maori/Indigenous languages revitalization movements.

Professor Smith stepped down from the administrative position of CEO/Vice Chancellor at Te Whare Wananga o Awanuiarangi in 2015 after eight years in the role. He resumed his academic research and writing career. In 2019 he was invited to take up a special Chair – Te Toi Ihorei ki Purehuroa – A "distinguished Professorial position at large" established for him at Massey University.

Professor Smith has been recognized and honored for his international contributions, receiving an Honorary doctorate (D.Litt.) from the Okanagan University College in 2005, and an Honorary doctorate (LLD) from the University of Northern British Columbia in 2013; he was elected as a Fellow of the American Education Research Association in 2014, and was awarded a Queens Honour – a CNZM in 2014. He was given the Prime Minister's Life-time Achievement Award (2017); and received the Te Ururangi National Award for Education (Matariki Awards, 2019).

Professor Smith is married to Professor Linda Tuhiwai Smith and they have one daughter and two mokopuna (grandchildren).



Therese N. Hopfenbeck is Professor of Measurement Analytics in the Melbourne Graduate School of Education, The University of Melbourne. Before starting her position in Melbourne, she was Professor in Educational Assessment and Director of the Oxford University Centre for Educational Assessment. She is a Visiting Fellow at Kellogg College, University of Oxford, elected Vice-President of *The Association for Educational Assessment-Europe* and Lead Editor of the journal *Assessment in Education, Principle, Policy and Practice*. Dr. Hopfenbeck's research agenda focuses upon bridging research on self-regulation and classroom-based assessment and making sense of international large-scale studies in education. In collaboration with Professor Nancy Perry, University of British Columbia, she is currently leading an international network of researchers disseminating classroom-based research, funded by Social Sciences and Humanities Research Council of Canada (2020–2022). She is also currently Principal Investigator for two research projects funded by IB, on critical thinking in PYP schools internationally and evaluation of education reforms in Kent, UK (2020–2022). In 2020, she led the research on Critical Thinking in the Diploma Program in Australia, England, and Norway (<https://ibo.org/research/outcomes-research/diploma-studies/critical-thinking-skills-of-dp-students/>). Dr. Hopfenbeck was Principal Investigator for the PISA (2022) study in England, Northern Ireland, and Wales, in collaboration with Pearson UK (2018–2023). She was the Research Manager of PIRLS (2016), funded by The Department of Education, UK.gov, and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.

tion, [UK.gov](https://www.uk.gov/), and Principal Investigator of a major ESRC-DFID research study, Assessment for Learning in Africa (ES/N010515/1) (2016–2019). Since coming to Oxford in 2012, she has been the recipient of funding from ESRC-DFID, OECD, The Norwegian Research Council, Education Endowment Foundation, State Examinations Commissions Ireland, Jacob Foundation, and the International Baccalaureate totaling more than £2 mill in addition to a single grant of £4 mill in collaboration with SLATE: Centre for the Science of Learning & Technology at the University of Bergen, Norway. Prior to her appointment at Oxford, she worked as a postdoctoral researcher at the University of Oslo's research group for Measurement and Evaluation of Student Achievement at the Unit for Quantitative Analysis of Education (2010–2011). She is Adjunct Professor of the Norwegian University of Science and Technology (NTNU), member of the Visiting Panel for Research at the Educational Testing Service (ETS) in Princeton, expert member of the PISA (2022) Questionnaire Framework group, appointed by ETS and OECD (2014–2023). She has advised on the implementation of formative assessment programs in India, South Africa, Norway, and the Emirates and carried out policy work for UNESCO/OECD and the Norwegian Ministry of Education Norway. Therese has a presence on LinkedIn, ResearchGate, [Academia.edu](https://www.academia.edu/), and Twitter: @TNHopfenbeck.



Nataliya Ivankova, PhD, MPH, is Professor in the School of Health Professions at the University of Alabama at Birmingham (UAB). As an applied research methodologist working at the intersection of mixed methods, qualitative, community-based, action, and translational research, she is internationally recognized for her empirical and methodological publications. She shares her research expertise through teaching mixed methods and qualitative research courses, conducting workshops and guest lectures, mentoring doctoral students and junior faculty, and serving as key personnel on funded research projects. At UAB, she directs an online graduate certificate program in *Applications of Mixed Methods Research* and oversees the design and analytical aspects of qualitative and mixed methods studies through the Center for Health Informatics for Patient Safety and Quality. She is a founding Coeditor of the *Mixed Methods Research Series* (Sage) and serves as an Associate Editor for the *Journal of Mixed Methods Research*.



Professor Jahnke is Director of Toi Kura Centre for Maori & Indigenous Education. She was the 2019 inaugural Toi o Te Wananga Fellow. Professor Jahnke holds professional qualifications in education and is cofounder and immediate past Chair of the Indigenous Peoples of the Pacific Special Interest Group of the American Education Research Association (AERA). She currently serves on the AERA Executive Committee. She was a member of the Social & Human Sciences Sub Commission of the NZ National Commission for UNESCO, the NZ Bioethics Council, the NZ Ethics Committee on Assisted Reproduction Technology, and Chair of NZ Health Research Council Ethics Committee. She is currently Cochair of the Aotearoa Research Ethics Board and a Fulbright Scholar.



Berit Karseth, PhD, is Professor at the University of Oslo, Department of Education. Karseth's main research and publications lie in the field of education policy, reforms, and curriculum studies. She has edited four books and written several journal articles and book chapters in both Norwegian and English. Her most recent publications in English include articles published in the *Journal of Education Policy*, *Scandinavian Journal of Educational Research*, *Journal of Curriculum Studies*, and *European Educational Research Journal*. Karseth was the elected dean of the Faculty of Educational Sciences at the University of Oslo 2013–2016 and is currently in the leadership team of the research group Curriculum studies, educational leadership, and governance (CLEG) at the same institution.



Dr. Kenneth Leithwood is Emeritus Professor at OISE/University of Toronto and an honorary professor at the University College London and the University of Nottingham. His research and writing is about school leadership, educational policy, and organizational change. He has published extensively on these topics. For example, his most recent books include *Leadership Development on a Large Scale: Lessons for Long-term Success* (2018, Corwin), *How School Leaders Contribute to Student Success* (2017, Springer), and *Linking Leadership to Student Learning* (2012, Jossey Bass). Professor Leithwood is the inaugural recipient of the University of Toronto's *Impact on Public Policy* award, AERA (2011) *Outstanding Leadership Researcher Award*, the 2012 *Roald F. Campbell Lifetime Achievement Award* from the University Council for Educational Administration, and the Ontario Principals' Council *Outstanding Educator of the Year* award for 2017. He is a *Fellow of the Royal Society of Canada*. A recently published bibliometric analysis found that Dr. Leithwood was the most frequently cited, highest impact educational leadership researcher in the world.¹ Research.com's 2022 Ranking of the Top 1000 Scientists in the area of Social Sciences and Humanities ranked Dr. Leithwood 70th in the world and 7th in Canada.



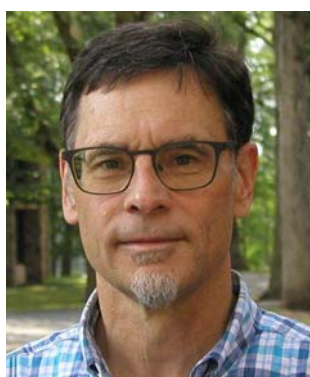
Xiufeng Liu is a Professor of Science Education at State University of New York, Buffalo, and a Fellow of the American Association for the Advancement of Science (AAAS). He obtained his PhD from the University of British Columbia in 1993. Before his current position, he taught high school chemistry in China, and was a tenured faculty member at St. Francis Xavier University and University of Prince Edward Island, both in Canada. From September 2020 to August 2022, he served as a program director in the Division of Research on Learning of the National Science Foundation. Dr. Liu conducts research on measurement and evaluation of STEM education and student learning of cross-cutting concepts of matter and energy. Dr. Liu is currently a Coeditor-in-chief for the journal *Disciplinary and International Science Education Research* (Springer) and an Associate Editor for the *Journal of Research in Science Teaching*, a premier journal in science education.



Margie Maaka is a Professor of Education at the University of Hawai'i at Mānoa. Her PhD is in educational psychology. She also holds professional qualifications in education from the University of Waikato and Hamilton Teachers College. She is cofounder and current Program Chair of the Indigenous Peoples of the Pacific Special Interest Group (SIG) of the American Educational Research Association (AERA). From 2017 to 2020, she served as a member of the AERA Executive Council and Chair of the SIG Executive Committee. Before that (2014–2017), Professor Maaka served as a member of the SIG Executive Committee. She has also served as a member of the Ngā Pae o te Māramatanga Center of Research Excellence International Research Advisory Panel. Her research on community-based educational renewal saw her recognized as a national scholar with the Goodlad Institute for Educational Renewal. As Cochair of the Sovereign Councils of the Hawaiian Homelands Assembly's Committee on Education, she prepared educators and educational leaders for public schools in the Hawaiian Home Lands Trust communities. Currently, Professor Maaka is committed to the preparation of indigenous educational leaders in higher education. Her research interests include community partnerships, educational psychology/indigenous educational psychology, indigenous development and advancement, educational policy, multiliteracies, and language and cognitive development.



Diane Mayer is Professor Emerita at the University of Oxford. Her research and scholarship focus on teacher education and early career teaching, examining issues associated with the policy and practice of teachers' work and teacher education.



Daniel McCaffrey is the Associate Vice President of Psychometric Analysis and Research in the Research and Measurement Science unit of the Research and Development division at ETS. Dan's research interests include the evaluation of artificial intelligence scoring of constructed responses, developing tools and methods for causal modeling, evaluating the psychometric properties of student achievement growth measures derived from annual test scores, and methods to account for measurement error in secondary data analysis. Dan received his PhD in Statistics from North Carolina State University in 1991.



Judy Parr is Professor Emerita at the University of Auckland, Faculty of Education, where she was the founding Head of the School of Curriculum and Pedagogy. A psychologist, Dr. Parr's research program is located broadly within improvement science. Her expertise is in literacy, specifically writing and its development, encompassing instructional issues like teacher knowledge and practice and the assessment of writing. Much research has been collaborative, with colleagues and practitioners within major national projects for school improvement aimed at raising achievement in literacy and, internationally, as part of funded large-scale research projects, for example in Canada and Norway. Dr. Parr has published widely in the areas of literacy, assessment, and professional learning.



Jeong-eun Rhee is a Professor of Education, Long Island University, Post, New York. Born and raised in Seoul, Korea, she earned her bachelor's degree in Educational Psychology from Ewha Women's University. In 1992, she came to the US as an international student and received an MA in Educational Psychology from West Virginia University and a PhD in Educational Policy and Leadership from the Ohio State University. As an interdisciplinary cultural researcher and educator, her recent work focuses on developing and supporting transnational, intergenerational, and decolonial feminist knowledge projects in/through antioppressive education. She recently published *Decolonial Feminist Research: Haunting, Rememory, and Mothers*, which received an honorable mention for the 2022 Qualitative Book Award of International Congress of Qualitative Inquiry.



Fazal Rizvi is an Emeritus Professor of Global Studies in Education at the University of Melbourne, as well as at the University of Illinois at Urbana-Champaign. He is also a Visiting Professor at the Universities of Turku in Finland and Pretoria in South Africa. He has written extensively on issues of identity and culture in transnational contexts, globalization and education policy, and Australia-Asia relations. He has published widely and his book, coauthored with Bob Lingard, *Globalizing Education Policy* (Routledge, 2010), is used widely in courses around the world. A sequel to this book, *Reimagining Globalization and Education* (Routledge, 2022), has been published recently. Fazal is currently researching educational reform in Bhutan. He is a Fellow of the Australian Academy of the Social Sciences, a past Editor of the journal, *Discourse: Studies in Cultural Politics of Education*, and a past President of the Australian Association of Research in Education.



Theresa Rogers is Professor of Language and Literacy Education at the University of British Columbia. Her research interests include youth arts, media and critical literacies, adolescent literature, and teacher education in literacy. Most recently she published "Youth Activism through Critical Arts, Transmedia, and Multiliteracies" in the *Oxford Research Encyclopedia*. She is author of the book, *Youth, Critical Literacies and Civic Engagement: Arts, Media and Literacy in the lives of adolescents* (Routledge) and numerous articles in scholarly journals such as *Literacy Research: Theory, Method and Practice*, *Reading Research Quarterly*, *Pedagogies*, *Language Arts*, *Children's Literature in Education* and *Journal of Literacy Research*. See: <https://ubc.academia.edu/TheresaRogers>.



Professor Chika Sehoole is a Professor of Higher Education and the Dean of the Faculty of Education at the University of Pretoria. He obtained his PhD at the University of the Witwatersrand, Johannesburg, in December 2002. In 2003/4 he was a visiting Rockefeller postdoctoral fellow at the Centre for African Studies at the University of Illinois in Urbana Champaign. In 2005/06 he was a New Century Fulbright Fellow, again, at the same university. His research interests are in the areas of higher education policy, internationalization of higher education, higher education in Africa, and teacher education. He has served in several international research projects focusing on higher education and providing expertise from the South African and African perspectives. In 2013–2017 he served as the Chairperson of the Board of the African Network for Internationalisation of Education (ANIE). He currently serves as the chairperson of the Education Deans' Forum of South Africa. He has published a singled authored book titled "Democratising higher education policy, constraints of reform in post-apartheid South Africa" by (Routledge); three edited scholarly books on higher education and more than 40 articles in journal articles, book chapters, and popular media.



Roland Sintos Coloma is a tenured Full Professor of Teacher Education at Wayne State University in Detroit, Michigan. A scholar of history, cultural studies, and education, his research addresses critical questions of race, gender, and sexuality from transnational and intersectional perspectives. His publications include *Asian Canadian Studies Reader* (2017), *Filipinos in Canada: Disturbing Invisibility* (2012), and *Postcolonial Challenges in Education* (2009). His articles have been published in top-ranked peer-reviewed journals, including *Curriculum Inquiry*, *Discourse: Studies in the Cultural Politics of Education*, *Educational Studies*, *Educational Theory*, *History of Education Quarterly*, *International Journal of Qualitative Studies in Education*, *Paedagogica Historica*, and *Race Ethnicity and Education*. He has successfully garnered over \$2 million of external funding from federal, education, and nonprofit agencies, including the US Department of Education, National Education Association, Spencer Foundation, and Schultz Family Foundation as part of the Bill & Melinda Gates Foundation's Grand Economic Challenges. Roland has served in various university leadership roles as an assistant dean, department chair, graduate coordinator, and center codirector. He was president of the American Educational Studies Association (2018–19), editor of the *Educational Studies* journal (2014–17), and program cochair of Division G (Social Context of Education) in the American Educational Research Association (AERA). He received the 2017 Distinguished Scholar Award from the AERA Research on the Education of Asian Pacific Americans SIG and the 2015 Article of the Year Award from the AERA Queer Studies and Education SIG. In 2020, he was appointed to the Governor of Michigan's statewide Asian Pacific American Affairs Commission. Born in the Philippines and raised in California, he completed his PhD in Cultural Studies in Education and Minor in African American and African Studies from The Ohio State University in 2004.



Roger Slee is the Founding Editor of *The International Journal of Inclusive Education* and *The Journal of Disability Studies in Education*. A former Deputy Director-General of Education in Queensland, Australia, he has advised governments on inclusive education policy and practice. His published books include: *The Irregular School*; *Inclusive Education Isn't Dead, It Just Smells Funny*; and *Ethics and Inclusive Education: Disability, Schooling & Justice* (with Gordon Tait). Roger was commissioned by the UNESCO to write the scoping paper for the 2020 Global Education Monitoring Report, *Inclusion and Education*. Roger also held the Chair of Inclusive Education at the Institute of Education, London, and is now the Diamond Jubilee Professor of Disability & Inclusion in the School of Sociology & Social Policy at the University of Leeds.



Ninni Wahlström is a Professor of Education at Linnæus University, Sweden. Her current research focuses on transnational and national policy discourses and their implications for national curriculum and classroom teaching from a perspective of curriculum theory and educational philosophy. Most recent publications are "School and democratic hope: the school as a space for civic literacy" in *European Educational Research Journal* (2022) and Wahlström (Ed.) *Equity, Teaching Practice and the Curriculum*, Routledge (2022).



Lei Wang is Professor and PHD advisor of the Institute of Chemical Education, the College of Chemistry in Beijing Normal University (BNU). She is also the member of Teaching Steering Committee on Basic Education of Ministry of Education of China and the director of its branch Committee on Secondary Chemistry Education, the director assistant of Faculty of Education of BNU, vice director of the Institute of Science Education of BNU, and Chairman of the Chemistry Education Committee of the Chinese Chemical Society. She is one of the chief experts in developing the *Secondary School Chemistry Curriculum Standards* of MOE. She works as the chief editor of high school chemistry textbooks (Shandong Science and Technology Press, 2004 and 2019 editions) and junior high school chemistry project-based learning textbooks (Shanxi Education Press, 2018 edition). She achieved the first prize of the National Basic Education Teaching Achievement Award, the second prize of the National Higher Education Teaching Achievement Award, and the BAOGANG Education Outstanding Teacher Award.



Everard Weber is a retired professor. He has worked in the field of education for more than 40 years. Currently he is a research associate at the University of Johannesburg (UJ) and an academic advisor in the Future Professors Programme (FPP) at UJ. He studied African Languages and History at the University of Cape Town and Comparative and International Education at Harvard University. In South Africa he has worked at Crestway High School in Cape Town, the University of the Witwatersrand, and the University of Pretoria. A recent book is *A home away from home: a community of international and South African university students*. The FPP is a government-sponsored initiative that seeks to develop the next generation of professors and academic leaders in all disciplines, across all universities in South Africa. He is in addition presently conducting fieldwork on curriculum change in South African higher education with Professor Shireen Motala and Dr. Venise Joubert at UJ. This project is grounded in the history of southern Africa and hopes to contribute to the scholarship of postcolonialism.



Ian A.G. Wilkinson is Professor in the Department of Teaching and Learning at The Ohio State University and Honorary Professor in the School of Curriculum and Pedagogy at the University of Auckland, Aotearoa/New Zealand. He has a background in educational psychology with special interests in cognition, instruction, and research methodology as they relate to the study of literacy. He served as coauthor of the influential report, *Becoming a Nation of Readers* (National Academy of Education, 1985). He was Co-Editor of *Reading Research Quarterly* from 2006 to 2012 and was inducted into the Reading Hall of Fame in 2021. His research examines the impact of classroom discussion and dialogue on students' reading comprehension and argument literacy and the implications for professional development of teachers. Recently he coauthored books include *The Most Reasonable Answer: Helping Students Build Better Arguments Together* and *Quality Talk About Text: Discussion Practices for Talking and Thinking About Text*.



David B. Yaden, Jr., PhD, is the *Jewell M. Lewis Endowed Distinguished Professor of Literacy* in the Department of Teaching, Learning and Sociocultural Studies at the University of Arizona. Prior to his present position at the University of Arizona, he held appointments at Emory University, the University of Houston, and the University of Southern California. He served as Co-Editor of the *Journal of Literacy Research*, one of the premier research journals in the field of literacy studies, and was a Principal Investigator in the federally funded Center for the Improvement of Early Reading Achievement (CIERA) where he implemented an early literacy curriculum for linguistically diverse preschoolers in inner-city Los Angeles. He also directed a statewide consortium of researchers from Arizona's major universities in the evaluation of the state's early childhood initiative to evaluate health, developmental, and educational factors contributing to school readiness. His research interests and specializations include developmental issues in early childhood education, the acquisition of literacy and biliteracy in young emergent bilingual children, family literacy, theories of reading disability, microgenetic research design, and the application of complex adaptive systems theory to growth in reading and writing.

Volume 14: Quantitative Research and Educational Measurement

Daniel F McCaffrey

© 2023 Elsevier Ltd. All rights reserved.

Introduction

The chapters that follow in Volume 14, *Quantitative Research and Educational Measurement*, present an overview and introduction to these two foundational areas of education research and practice. Quantitative research consists of “explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics)” (Creswell, 1994). This short description captures the two fundamental elements of quantitative research: data and analysis using mathematical or statistical methods. The data and analysis are meant to support *empirical statements*, i.e., statements expressed in numerical terms that are about what “is” the case in the “real world” rather than what “ought” to be the case (Cohen and Manion, 1980). The unique feature of quantitative research is its systematic use of data to derive conclusions. Quantitative research studies are designed so that the data and analysis offer an accurate description of the real world and with samples that are large enough to provide conclusive evidence concerning hypotheses about the world. Typically, the statistical analysis methods used in quantitative research allow the researchers to assess the level of uncertainty or error in the empirical results and provide a measure of confidence in any hypotheses tested with the data.

Quantitative research is important because it provides insights not offered by other types of research. Quantitative research in education is used to test theories against the real world, whereas qualitative research typically is seen as generating hypotheses and theories about “how” the phenomena develop, progress, and exist. Theories are based on human interpretations and understandings of how the world works. Theories must be tested against the world and updated and revised. The empirical methods of quantitative research are needed for such tests. Quantitative research draws conclusions from the empirical analyses and studies are designed to test hypotheses with confidence not to generate them (Queirós et al., 2017).

Quantitative research is a massive topic area that encompasses all educational research involving quantitative data and methods. To provide a manageable introduction to this topic, we narrowed the focus of Volume 14 to statistical methods and educational measurement. These two areas cover the two fundamental elements of quantitative research: data and analysis.

“Modern problems and needs are forcing statistical methods and statistical ideas more and more to the fore” (Garrett, 1926, p. v). The sentiment captured by Henry Garrett almost 100 years ago is probably even truer today. With the ever-growing availability of data and empirical results, an understanding of the foundations of statistical methods has been and remains essential for any educational researcher or practitioner. The discipline of statistics methods is itself a large topic area with many subdisciplines and lines of research. Because the focus of the Encyclopedia is education, we restricted the entries on statistical methods and models that are pertinent to educational research and applications.

The other fundamental component of quantitative research is data. As Sukamolson (2007) notes, a challenge to quantitative education research is the fact that many of the phenomena of interest in education such as students’ knowledge, attitudes, or beliefs do not occur in the form of “naturally” quantitative data. The field of educational measurement arose to overcome this need by providing statistical methods, theories, and practices to collect data on educational outcomes of interest. Like statistical methods, the practice of measuring student knowledge has a long history dating back to the mid-19th century in the England (Clauser and Margolis, 2022).

Educational measurement and statistical methods are overlapping fields. Educational measurement is an application of the general statistical approach:

- a. Collect data related to question of interest or to support a decision;
- b. Analyze and summarize the data;
- c. Using the laws of probability and assumptions about the relationship between the data and the question of interest, draw conclusions from the data including measures of uncertainty about inferences.

However, the discipline of statistics, even restricted to methods used in educational research and practice, entails far more than the models and methods used in educational measurement. Additionally, educational measurement fundamentally is more than simply a branch of statistics. There are unique foundational issues underlying measurement, particularly for the unique attributes measured in educational applications. Educational measurement not only entails the analysis of the data but also includes the theory and practice of collecting measurements through educational testing or assessments. Furthermore, educational measurement is not an abstract field of research. It involves drawing inferences about humans engaged in educational endeavors and the educational systems that support them. Educational measurements can have profound impacts on individuals and educational policies and practices. Consequently, the research on educational measurement goes beyond the mathematics of psychometrics and abstract theories on the foundations of measurement and includes considerations of the responsibilities associated with the use of educational assessments and the rights on people being assessed. These issues are fundamental to educational measurement and woven into the research in ways that are distinct from statistical methods.

The goal of this section is to provide readers with a useful survey of the topics in educational measurement and the statistical methods used in educational research and practice. The topics included are broad but not fully comprehensive. Readers coming to this area for the first time should be able to find an introduction to the field including topics that are the basic foundations

of the field (e.g., reliability or linear regression) and emerging topics (e.g., advanced machine learning or intelligent tutoring). For individual topics, the goal is to give readers a brief introduction that highlights some of the key aspects of the topic. For basic statistical methods, chapters typically include an example to highlight how the methods work in practice. Our goal is that after reading a chapter, users will be familiar with the topic and have enough information to decide if a statistical method might be of use for an application or if a topic is something they should explore further to help them pursue their research interests or solve a problem.

Key Themes and Topics Covered

One of the defining and long-term persistent features of educational practice and research is the interest in the measurement of students' educational knowledge and skills. The goal of education is for students to learn and acquire new knowledge and skills; so measuring the knowledge and skills is clearly important for studying, evaluating, and directing education. Thus, educational measurement is an essential topic to be covered as part of the section on quantitative methods in educational research.

Beyond educational measurement, quantitative methods for education research and practice encompass many topics in multiple fields. We chose to focus on statistical methods because they comprise an essential component of educational research and because there exists a set of statistical methods that are commonly associated with educational applications and truly key for any researcher, student, or practitioner interested in quantitative methods for education. Hence, the topics covered in this section pertain to either educational measurement or statistical methods.

Educational Measurement

The field of educational measurement encompasses theoretical concepts such as the philosophy of measurement and abstract mathematical psychometrics as well as topics related to the practice of developing, administering, and scoring assessments of people's knowledge and skills. Someone coming to the field should have a familiarity with the theoretical foundation of educational measurement and the practice and application of educational assessments. In recent years, the field of assessment has undergone significant changes in response to the widespread introduction of digital assessment and the data that it produces. In this section we attempted to find topics across the spectrum from foundational theory to the most recent applications of digital assessment. Education is a core component of all human societies and educational measurement must support the societal role of education and be responsive to it. Hence, we also include topics to give some sense of how educational measurement fits into the broader social context. We group the chapters into seven broad topic areas roughly ordered from more theoretical topics to more applied topics. The section only covers a portion of all the topics we could cover in educational measurement. They give readers a broad introduction to educational measurement and include topics everyone interested in the areas should be familiar with.

Foundational Concepts

We begin the section with an overall history of the discipline of educational measurement in the chapter titled "Past, Present, and Future of Educational Measurement." This chapter sets the material on educational measurement in its historical context helping readers to understand where practices and the conceptual foundations of educational measurement arose and came to be the standards of the field. It also describes the fields troubling past as part of the eugenics movement and sets the stage for the future looking at coming trends including issues of equity.

The historical perspective is followed with a discussion of the role of standards in the practice of educational measurement in "Educational Testing: A Review of Standards, Guidelines, and Practices." This chapter acknowledges the dual nature of the field of educational measurement as both an academic discipline and a set of methods and practices used in assessment of individuals for consequential decisions for the individual or the systems that educate them.

The first two chapters are meant to ensure readers understand how fundamentally important the ethical responsibility of protecting test takers is to the field of educational measurement. Measurement of cognitive skills at the core of educational measurement is distinct from measuring physical objects which is the heart of physical science. Measurement of cognitive skills is indirect and based on inference. This allows for imprecision and biases. Moreover, educational measurement makes inferences about the abilities and skills of individual people. These inferences can be used for high stakes decisions such as employment or acceptance to higher education. Hence, it is critically important that educational measurement takes steps to protect the individuals being measured and the field has established ethical codes and practices to meet that requirement. The code is most explicitly laid out in the standards discussed in the chapter "Educational Testing: A Review of Standards, Guidelines, and Practices." However, it is also clearly part of the history of educational measurement and woven into the fabric of all the aspects of educational measurement and all the topics and chapters in this section of the Encyclopedia. These first two chapters are an important read for anyone coming to the field to understand the ethical underpinnings of the field.

The remaining three chapters in Foundational Concepts describe the three properties that are the basis for all measurement. Educational measurements must be reliable, support valid inferences, and be fair to the people being measured. Given these are such foundational properties for educational measurement, the chapters "Reliability," "Validity," and "Fairness" provide an introduction to the core concepts, but their authors also acknowledge that even core properties evolve and the chapters provide perspectives on how the concepts are changing and might continue to change in the future.

Modeling Frameworks

The chapters in the section Modeling Frameworks provide an introduction to the basic psychometric models of educational measurement. Psychometric modeling has generally followed two broad paradigms: classical test theory and latent variable/item response theory. Classical test theory is the more traditional approach and focuses on identifying and assessing the errors in test scores. The first two chapters, “Classical Test Theory” and “Generalizability Theory and Applications,” present models from the classical test theory paradigm.

Item response theory (IRT) posits a latent variable or ability and models test takers’ responses as probabilistic functions of their latent variables. The remaining chapters in this section are all from the IRT paradigm. Although models of the form of the IRT models have been around for over 80 years, widespread use of these models for measurement practice is more recent, occurring in the last 30 or so years, substantially after the use of classical test theory. There has been and continues to be very active research into developing new models in the latent variable paradigm.

The chapters in this section survey the diverse models from this paradigm. They start with an introduction to the basic IRT model and then consider adaptations and extensions such as modeling multiple latent variables for each test taker and the adaptation for classifying test takers skills via diagnostic models. The chapters also consider both classical (frequentist) and Bayesian methods for fitting the models and making inferences about test takers. The chapters not only give readers an introduction to models used in practice and research, but they also give readers a sense of how the models have evolved and might be further adapted for additional applications.

Test Development

Whereas Model Frameworks section presents an introduction to the mathematical theory underlying educational measurement, the chapters in the Test Development section turn to the practical side of educational measurement focusing on the design of assessment systems and tests. The chapters in this section are forward looking often focusing on the use of technology in testing and test development and other topics related to recent changes in testing and society such as increased globalization.

A theme to this section is the importance of the design for ensuring that inferences drawn from test scores are valid and supported with strong evidence. This theme runs through the first three chapters, “Principled Assessment Design,” “Evidence-Centered Design,” and “Cognition-Centered Design Principles for Digital Assessment Tasks and Items.”

A second theme is the role of automation and technology in assessments and their development. This theme is related to the broader theme of the impact of the transition to digital assessment and the increasing role of technology in testing. Automation of the development of tests through automated item generation and test assembly is increasing in practice and likely to be more transformational in the future as methods and technology evolve. Hence, these are key chapters for readers who want to see the direction of educational measurement. Adaptive testing is an example of the way digital assessments are changing testing and will continue to do so.

The section ends with another key component for ensuring the validity of inferences: score reporting. Again, digitalization is affecting reporting and offers great opportunities for making scores more informative and maintaining the validity of inferences drawn from the scores. Growing globalization makes translation and adaptation of testing increasingly important. This chapter gives readers a brief introduction to this topic and helps prepare them for educational measurement practices in the future.

Establishing the Quality of Scores

The section on quality control also focuses on the practice of assessment. Chapters in this section cover many practices and methods that are used to ensure test scores are reliable, fair, and support the inferences and decisions made about test takers. The chapters start with the very traditional and universal practices of equating, linking, and scaling of scores and testing for differential item functioning. These are foundational concepts for which anyone interested in educational measurement should be familiar.

All the models presented in Modeling Frameworks make various assumptions about the nature of the data generated by the assessment and its relationship to the attributes of the test takers that are of interest. For those models to provide useful information, the models must fit the data. Consequently, an active field in educational measurement is development of methods for evaluating model fit. Multiple chapters in the section provide readers with an introduction to this important practice and its growing research base.

Respondents’ actions and tendencies when responding to assessments can invalidate inference made from their responses. Methods to detect such behaviors and tendencies are a standard part of practice, and the development of new detection methods is an active area of research. Multiple chapters in this section provide readers with broad introduction to the traditional and emerging research and practices in detecting problem behaviors of test takers. Topics covered include the detection of cheating and aberrant responses, modeling response times, and response styles. Cheating and other forms of aberrant responses pose significant risks to the validity of inferences drawn from test scores because the scores might not reflect the test taker’s ability, knowledge, or skills.

Methods to detect possible problem behaviors or scores also are an active area of research and checks for cheating are commonly part of the practice of assessment. Response times can also provide valuable information about a test taker’s interaction with the assessment. This information can have many purposes, and with the rapid expansion of digital assessments, use of timing data to evaluate test takers or their behaviors has been growing. The chapter “Modeling Response Times” introduces this field of research and practice.

Test takers may have tendencies in their responses to assessment items such as not using extreme responses when responding to Likert items. We provide readers with an introduction to the emerging field of modeling such behaviors in the chapter “Response Style Modeling.”

A final source of potential error considered in the section is the errors introduced by human raters of constructed response items. Readers can find information on the burgeoning field of modeling the tendencies of raters to make systematic errors in the chapter “Rater effects modeling.”

Educational Policy

In addition to assessing individuals, educational measurements are used to evaluate education systems such as schools or school districts or entire countrywide systems. In certain countries, e.g., the United States, the assessments are used to hold educators accountable for student outcomes. More broadly, around the world a goal is to use assessment scores to improve education systems. A major component of educational assessment’s contribution to educational policies is through surveys of the knowledge and skills of student or adult populations in international programs such as the Programme for International Student Assessment (PISA; [OECD, n.d.a](#)), the Programme for the International Assessment of Adult Competencies (PIAAC; [OECD n.d.b](#)), the Trends in International Mathematics and Science Study (TIMSS; [IEA, n.d.](#)), the Progress in International Reading Literacy Study (PIRLS; [IEA, n.d.](#)), or the Programme for the Analysis of Education Systems (PASEC; [The World Bank, n.d.](#)) or programs in individual countries such as the National Assessment of Educational Progress (NAEP; [National Center for Education Statistics, n.d.](#)) in the United States. The programs are described in two chapters. One, “Educational surveys: Conceptual overview,” gives an overview of the goals and objectives of the programs, and the other, “Educational surveys: Methodological foundations,” presents the common test design and statistical modeling approaches used in these surveys. The methods are distinct from most other assessments because the intended inferences for the surveys are about population groups and relationship among variables for populations. They are not intended for inferences about particular individuals, which are the target of other educational measurements.

The remaining chapters in this section relate to practices and models used to support educational policy. Standard settings are used in many contexts where test scores are used to classify test takers by their performance, including licensure. In the policy context in the United States, they are used to set cut scores for classifying students into proficiency categories and this classification is widely used in educational policy. The “Standard Setting” chapter reviews procedures for setting cut scores and considerations in making classifications.

Growth and value-added modeling are used for educational accountability in the United States and other countries. The chapters “Achievement Growth Modeling” and “Value-Added Modeling” provide readers with a general overview of both modeling approaches and discussion of their applications and the concerns about their use for policy decisions. These chapters again echo the importance of issues around the ethical use of educational measurement through their discussion of the potential unintended consequences of using test scores in growth or value-added modeling for educator accountability.

Interdisciplinary Frontiers

The chapters in the section Interdisciplinary Frontiers provide readers with further examples of how digital assessments and computer technology have changed the landscape of educational measurement. These chapters are likely to be important to readers in the coming years as the use of the practices described in the chapters expands and basic understanding of the applications becomes increasingly valuable. The topics covered in the chapters “Automated Scoring of Constructed-Response Items in Educational Assessment,” “The Current Trends and Opportunities for Machine Learning in Learning Analytics,” and “Computer-Based Process Data Analysis” share a common theme of the use of computer algorithms to extract and process data from test takers interactions with digital assessments to yield scores or other measures for individual test takers. Automated scoring uses algorithms to extract information from constructed (e.g., written or spoken) responses. Process data use computer algorithms to extract information from data about students’ interactions with the assessment other than their actual responses such as timing data, changes to responses, or the use of hints or aids that might be provided. Both methods tend to use machine learning ([Hastie et al., 2009](#)) as part of the extraction or synthesis of the data. Analytics uses various computer algorithms, potentially including algorithms used in automated scoring or extracting process data, to provide information to support learning.

The final chapter in this section, “Modeling of Network Structures,” focuses on another area of rapid growth in methodological development and applications: network analysis. Network analysis models connections in data such as modeling social connections, but it is also commonly used in analytics research.

Application Areas

The final section under Educational Measurement offers readers additional insights into educational measurement through a series of measurement applications. The applications include more traditional ones such as the use of educational measurements in certification and licensure, higher education admissions, and classroom assessment. They also include application areas that have evolved and continue to evolve with changes in technology and digital assessments. These include intelligent tutoring, simulation and game-based assessment, conversation-based assessment, and the assessment of collaborative problem-solving. Educational measurement is often associated with the assessment of quantitative and literacy skills or other traditional disciplinary subject matter such as science or history. However, research continues to demonstrate the importance of interpersonal and intrapersonal skills for success in school and work and the assessment of these skills is of increasing interest and seems likely to continue to be so in the future. Hence, we include an introduction to these skills and their assessment in the chapter “Assessment of Interpersonal and

Intrapersonal Skills.” Like the other sections, the applications included in this section provide readers with a grounding in the traditional and foundational aspects of educational measurement as well as an introduction to topics that are likely to be of importance in the future.

Statistical Methods

The chapters in the section Statistical Methods provide readers with an introduction to statistical methods that have wide application for educational research or practice. The chapters are grouped into two sections titled Modeling Frameworks and Foundational Methods. The chapters in Modeling Frameworks present statistical modeling approaches that can be used in data analysis. The chapters in Foundational Methods focus more on foundational concepts such as families of distributions, paradigms for inference (frequentist, Bayesian, or resampling), and the statistical approaches to study design. The chapters describe the models or methods including guidance on how they can be used to support quantitative research. To demonstrate their use, most chapters include an application from the field of education. The treatment of these topics focuses on their application rather than the mathematical foundations. The goal is for readers to have a basic understanding of what the method and models are and some insights into how they might apply them to their own research and applications.

Modeling Frameworks

The models described in the section Modeling Frameworks include classical models that make up the foundation of any applied statistics toolkit. They include analysis of variance, analysis of covariance, linear and nonlinear models, categorical data models, and generalized linear models. In addition to these standard models, there is a chapter on hierarchical linear models which are of great importance to educational research given that data on students are typically hierarchical with students grouped into classrooms and schools and sometimes into groups of schools. There are also chapters on standard multivariate modeling techniques including principal component analysis, exploratory and confirmatory factor analysis, clustering methods, and latent class analysis. Beyond these classic models and approaches there are chapters on structural equation models and growth mixture models. Recent advances in the field are also included in chapters on causal modeling and advanced machine learning.

Foundational Methods

The chapters in Foundational Methods review general statistical principles that are the foundation for the statistical models used in quantitative educational research. These principles include approaches to statistical inference and general theories of estimation. There are introductions to the primary paradigms of inference Bayesian and frequentist approaches. In addition, we include an introduction to resampling methods for inference. Resampling is an approach to conducting inference from the frequentist perspective that relies on fast computing power to provide inferences when traditional methods are difficult to apply. The chapter on estimation covers three methods for parameter estimation: least-squares, maximum likelihood, and Bayesian. These methods are all increasingly widely used and nearly all applications use one of these methods for estimation.

In addition to chapters on the foundations of parameter estimation and inference, Foundational Methods includes chapters covering continuous and discrete distributions and measures of correlation and association. Statistics are key to designing studies, and this section includes three chapters on study design. The first, “Research Study Design” is a general overview of study design including randomized experiments and observational studies. The second chapter on design, “Sampling Designs for Educational Research,” focuses more narrowly on the sampling of study participants and the concepts of survey sampling. The final chapter on design is “Power Analysis.” A critical component of every study or sample design is the determination of the sample size needed to conduct the study to produce precise estimates and meaningful inferences. Power analysis provides information to support the decision on sample size and hence is a critical component of study design. The chapter provides the basic concepts in power analysis but focuses on study designs for hierarchical or nested data (i.e., students in classrooms) because such designs are extremely common for educational studies.

The final chapters in the Foundational Methods section cover the challenge of the analysis of incomplete data, i.e., data with missing values, the graphical displays of data, and meta-analysis. Missing data are a challenge faced by nearly all quantitative studies. The chapter “Missing Data Analysis” provides theoretical basis for conducting analysis with incomplete data and describes common approaches including some more recent advances. The chapter “Displaying Information: From Woolly Mammoths to the Great Migration” describes some general principles and advantages of displaying data graphically and uses multiple examples to demonstrate the power of graphical data displays. The final chapter presents an introduction to meta-analysis which combines data from multiple other studies. Meta-analysis has and continues to play an important role in educational research. This chapter offers a general introduction to the method of meta-analysis and notes some recent advances.

Emerging Issues

New issues emerge constantly in all fields of study including education. The tools of statistical methods and basic educational measurement continue to be applicable to whatever new issues arise. For this reason, this section of the Encyclopedia offers readers introductions to many of the foundational concepts, models, and methods used in quantitative research generally and specifically for educational applications. Although the foundational principles and standard methods remain mostly unchanged and

applicable, statistical methods and models and practices in educational measurement continue to evolve. New models become a standard part of the research landscape as happened with the application of hierarchical linear models in the 1980s and 1990s. Increasing computational power makes the estimation of some models feasible as in the case of the estimation of Bayesian models and multidimensional latent variable models (e.g., multidimensional IRT). There are also innovations and new perspectives such as the use of multiple imputation for incomplete data. However, all these changes build on the fundamental concepts presented in this section.

The foundations of educational measurement also remain stable, but the practices change more rapidly. For example, the conceptualizations of validity evolved substantially from the 1950s to today. Also, there continue to be adaptations and variations to IRT models that work their way into common practice.

One of the main drivers of change to general statistical methods and educational measurement is the continuing growth in computational resources and technology. As noted earlier, this has allowed for new approaches to fitting statistical models, use of more complex models, and expanded the uses of resampling and Bayesian methods for inference. Artificial intelligence and machine learning also continue to evolve with new algorithms, computational methods, and models. These changes are affecting many areas of our lives and they will very likely have notable impact on quantitative educational research. The continuing growth in the use of digital testing and the large volumes of data this can produce provides even more opportunities for AI to affect educational measurement and quantitative research.

Education is deeply embedded into society; so societal changes outside of education per se are likely to affect quantitative methods in education. One example is the growing concerns about equity and justice in the United States and other countries. General concerns and focus on inequity and structural racism in society ignited concerns about the equity and the validity of educational assessments for all test takers, in particular, people of color and other historically disadvantaged groups (Sireci, 2021). These concerns have recently led researchers to suggest the need for participation by individuals from diverse groups in the design and development of assessments so that their perspectives can be considered throughout the process and the results will give all test takers equal opportunity to demonstrate their knowledge, skills, and abilities. There are also calls for changing test items and test designs to be socioculturally responsive and actively acknowledge the inherent biases in nearly all aspects of current educational assessment practices (Randall, 2021). Practices and methods are only catching up to the challenges being identified. It is highly likely that issues around assessment and equity and developing tests that are responsive to sociocultural context will have a substantial impact on educational measurement in the coming years both in practice and theory. In terms of practice, there is a need for more specific guidance on how to create assessments that are socioculturally responsive. There also is likely to be a need for new psychometric models for analyzing data created by such assessments.

References

- Clauser, B.E., Margolis, M.J., 2022. Past, present, and future of educational measurement. In: Tierney, R., Rizvi, F., Ercikan, K. (Eds.), *International Encyclopedia of Education*, fourth ed. Elsevier. pp. XX-XX.
- Creswell, J.W., 1994. *Research Design: Qualitative and Quantitative Approaches*. SAGE Publications.
- Cohen, L., Manion, L., 1980. *Research Methods in Education*. Groom Helm Ltd.
- Garrett, H.E., 1926. *Statistics in Psychology and Education*. Longmans, Green and Company.
- Hastie, T., Tibshirani, R., Friedman, J.H., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- IEA. (n.d.). TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/>.
- National Center for Education Statistics (n.d.). NAEP: National Assessment of Educational Progress. <https://nces.ed.gov/nationsreportcard/>.
- OECD. (n.d.a). PISA: Programme for International Student Assessment. <https://www.oecd.org/pisa/>.
- OECD. (n.d.b). OECD Skills Surveys. <https://www.oecd.org/skills/piaac/>.
- Queirós, A., Faria, D., Almeida, F., 2017. Strengths and limitations of qualitative and quantitative research methods. *Eur. J. Educ. Stud.* 3 (9), 369–386. <https://doi.org/10.5281/zenodo.887089>.
- Randall, J., 2021. “Color-Neutral” is not a thing: redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educ. Meas.* 40 (4), 82–90. <https://doi.org/10.1111/emip.12429>.
- Sireci, S.G., 2021. NCME presidential address 2020: valuing educational measurement. *Educ. Meas.* 40 (1), 7–16. <https://doi.org/10.1111/emip.12415>.
- Sukamolson, S., 2007. Fundamentals of quantitative research. *Lang. Ins. Chulalongkorn Univ.* 1 (3), 1–20.
- The World Bank. (n.d.). The Programme for the Analysis of Education Systems 2014. <https://www.worldbank.org/en/topic/education/publication/the-programme-for-the-analysis-of-education-systems-2014>.

CONTRIBUTORS TO VOLUME 14

Terry Ackerman
University of Iowa, Iowa City, Iowa, United States

Samrachana Adhikari
Division of Biostatistics, Department of Population Health, New York University Grossman School of Medicine, New York, NY, United States

Craig A Albers
University of Wisconsin, Madison, WI, United States

Allison J Ames
University of Arkansas, Fayetteville, AR, United States

Carolyn J Anderson
Department of Educational Psychology, University of Illinois, Urbana-Champaign, Champaign, IL, United States

Jessica Andrews-Todd
Educational Testing Service, Princeton, NJ, United States

Burcu Arslan
Educational Testing Service, Princeton, NJ, United States; and ETS Global B.V., Amsterdam, the Netherlands

Doug Baldwin
Senior Strategic Advisor, Educational Testing Service, Princeton, NJ, United States

Betsy Jane Becker
Department of Educational Psychology and Learning Systems, College of Education, Florida State University, Tallahassee, FL, United States

Randy E Bennett
Educational Testing Service, Princeton, NJ, United States

Paulina Biernacki
Stanford University, Palo Alto, CA, United States

Dennis D Boos
Department of Statistics, North Carolina State University, Raleigh, NC, United States

Angela D Bowzer
WestEd, San Francisco, CA, United States

Allison Ames Boykin
University of Arkansas, Fayetteville, AR, United States

Laine Bradshaw
Mary Frances Early College of Education, University of Georgia, Athens, GA, United States

Jennifer E Broatch
Arizona State University, Glendale, AZ, United States

Matthew A Brunetti
WestEd, San Francisco, CA, United States

Li Cai
University of California, Los Angeles, CA, United States

Jodi M Casabianca
Educational Testing Service, Princeton, NJ, United States

Michael Chajewski
Principal Research Scientist, Psychometric & Research Services, Pearson, New York, NY, Brooklyn, NY, United States

Dandan Chen
Department of Educational Psychology, University of Illinois, Urbana-Champaign, Champaign, IL, United States

Dakota W Cintron
Evidence for Action, San Francisco, CA, United States

Brian E Clauser
National Board of Medical Examiners, Philadelphia, PA, United States

James E Corter
Department of Human Development, Teachers College, Columbia University, New York, NY, United States

Tim Davey
Educational Testing Service, Princeton, NJ, United States

Paul De Boeck
The Ohio State University, Columbus, OH, United States

Kristen DiCerbo
Khan Academy, Avondale, AZ, United States

Chris Domaleski
National Center for the Improvement of Educational Assessment, Dover, NH, United States

Nianbo Dong
University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Neil J Dorans
Consultant, Lambertville, NJ, United States

Peter K Dunn
School of Science, Technology and Engineering, University of the Sunshine Coast, Sippy Downs, QLD, Australia

Norbert Elliot
New Jersey Institute of Technology, Newark, NJ, United States

Wilco Emons
Cito, National Institute for Educational Measurement, Arnhem, the Netherlands

Nnamdi C Ezike
University of Arkansas, Fayetteville, AR, United States

Raul Ferraz
University of Barcelona, Barcelona, Spain

Tahereh Firoozi
Centre for Research in Applied Measurement and Evaluation, University of Alberta, Edmonton, AB, Canada

Carol M Forsyth
Educational Testing Service, Princeton, NJ, United States; and Educational Testing Service, Princeton, NJ, United States

Andreas Frey
Goethe University, Frankfurt, Germany; and Centre for Educational Measurement (CEMO), University of Oslo, Oslo, Norway

Michael Friendly
York University, Toronto, ON, Canada

Matthew N Gaertner
WestEd, San Francisco, CA, United States

Michele Gallo
Department of Human and Social Sciences, University of Naples "L'Orientale", Naples, Italy

Mark J Gierl
Centre for Research in Applied Measurement and Evaluation, University of Alberta, Edmonton, AB, Canada

Janice D Gobert
Rutgers Graduate School of Education, New Brunswick, NJ, United States; and Apprendis, LLC, Massachusetts, MA, United States

Sakine Gocer Sahin
WIDA, Wisconsin Center for Education Research, University of Wisconsin-Madison, Madison, WI, United States

Brenna Gomer
University of Notre Dame, Notre Dame, IN, United States

Jennifer L Green
Michigan State University, East Lansing, MI, United States

Samuel Greiff
University of Luxembourg, School of Humanities, Education and Social Sciences, Department of Behavioral and Cognitive Sciences, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, Esch-sur-Alzette, Luxembourg

Debbie L Hahs-Vaughn
University of Central Florida, Orlando, FL, United States

Garret J Hall
Florida State University, Tallahassee, FL, United States

Zhuangzhuang Han
Educational Testing Service, Princeton, NJ, United States

Timothy A Hansen
Michigan State University, East Lansing, MI, United States

Emma K Harring
University of Maryland, College Park, MD, United States

Jeffrey R Harring
University of Maryland, College Park, MD, United States

Qing He

Department of Statistics and Data Science, University of Central Florida, Orlando, FL, United States

Mirka Henninger

University of Zurich, Zurich, Switzerland

Kei Hirose

Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan

Hsin-Hsiung Huang

Department of Statistics and Data Science, University of Central Florida, Orlando, FL, United States

Sijia Huang

Indiana University Bloomington, Bloomington, IN, United States; and University of California, Los Angeles, CA, United States

Paul A Jewsbury

Educational Testing Service, Princeton, NJ, United States

Matthew S Johnson

Educational Testing Service, Princeton, NJ, United States

Priya Kannan

Educational Testing Service, Princeton, NJ, United States

Madeleine Keehner

Educational Testing Service, Princeton, NJ, United States

Benjamin Kelcey

University of Cincinnati, Cincinnati, OH, United States

Bryan Keller

Teachers College, Columbia University, New York, NY, United States

Lale Khorramdel

Boston College, Chestnut Hill, MA, United States

Jee-Seon Kim

Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States

Stella Y Kim

University of North Carolina at Charlotte, Charlotte, NC, United States

Bruce M King

Clemson University, Clemson, SC, United States

Irwin Kirsch

Educational Testing Service, Princeton, NJ, United States

Florian Krieger

Technical University Dortmund, Department of Rehabilitation Sciences, Dortmund, Germany

Patrick C Kyllonen

Educational Testing Service, Princeton, NJ, United States

Suzanne Lane

University of Pittsburgh, Pittsburgh, PA, United States

Mina Lee

University of Massachusetts, Amherst, MA, United States

Tae Kyoung Lee

Department of Social Innovation Convergence, Child Psychology and Education, Sungkyunkwan University, Seoul, South Korea

Won-Chan Lee

University of Iowa, Iowa City, IA, United States

Blair Lehman

Educational Testing Service, Princeton, NJ, United States

Pui-Wa Lei

Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA, United States

Mariann Lemke

WestEd, San Francisco, CA, United States

Brian C Leventhal

James Madison University, Harrisonburg, VA, United States

Haiying Li

Iowa College Aid, IA, United States

Marlit Annalena Lindner

IPN—Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Xiang Liu

Educational Testing Service, Princeton, NJ, United States

Yang Liu

Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, United States

Oren E Livne

Educational Testing Service, Princeton, NJ, United States

Wen Wei Loh

*Department of Quantitative Theory and Methods,
Emory University, Atlanta, GA, United States*

Sharon L Lohr

Arizona State University, Tempe, AZ, United States

Christine Lott

*Rutgers Graduate School of Education, New Brunswick,
NJ, United States*

Richard M Luecht

*University of North Carolina at Greensboro, Greensboro,
NC, United States*

Juliette Lyons-Thomas

Educational Testing Service, Princeton NJ, United States

Mingjia Ma

University of Iowa, Iowa City, Iowa, United States

Ye Ma

Amazon Web Services, Seattle, WA, United States

Jaime L Malatesta

Graduate Management Admission Council (GMAC)

Sarah E Manski

*Michigan State University, East Lansing, MI,
United States*

Dobrin Marchev

*Teachers College, Columbia University, New York, NY,
United States*

Melissa J Margolis

*National Board of Medical Examiners, Philadelphia,
PA, United States*

Alberto Maydeu-Olivares

*University of South Carolina, Columbia, SC, United
States; and University of Barcelona, Barcelona, Spain*

James H McMillan

*Virginia Commonwealth University, Richmond, VA,
United States*

Thorsten Meiser

University of Mannheim, Mannheim, Germany

Robert J Mislevy

*University of Maryland, College Park, MD,
United States*

Scott Monroe

*University of Massachusetts, Amherst, MA,
United States*

Tim Moses

College Board, New York, NY, United States

Aaron J Myers

University of Arkansas, Fayetteville, AR, United States

Björn Nicolay

*University of Luxembourg, School of Humanities,
Education and Social Sciences, Department of
Behavioral and Cognitive Sciences, Campus Belval,
Maison des Sciences Humaines, 11, Porte des Sciences,
Esch-sur-Alzette, Luxembourg*

Maria Elena Oliveri

*Buros Center for Testing, University of Nebraska,
Lincoln, NE, United States*

Catherine Walker O'Neal

*Department of Human Development and Family
Science, University of Georgia, Athens, GA, United
States*

Juliana Cerentini Pacico

University of Iowa, Iowa City, Iowa, United States

Goran Pavlov

University of Barcelona, Barcelona, Spain

Alexandra Pérez

University of Connecticut, Storrs, CT, United States

Julius M Pfadt

*University of Ulm, Department of Psychological
Research Methods, Ulm, Germany*

Mya Poe

Northeastern University, Harvard, MA, United States

Sonya Powers

WestEd, San Francisco, CA, United States

Larry R Price

*Department of Methodology, Measurement and
Statistical Analysis, Psychometrics and Statistics, Texas
State University, San Marcos, TX, United States*

Gautam Puhan

*Educational Testing Service, Princeton, NJ,
United States*

D Rindskopf

CUNY Graduate Center, New York, NY, United States

Patrick J Rosopa

Clemson University, Clemson, SC, United States

Michael A Sao Pedro

Apprendis, LLC, Massachusetts, MA, United States

Madeline Schellman

*Mary Frances Early College of Education, University of
Georgia, Athens, GA, United States*

Dexin Shi

University of Barcelona, Barcelona, Spain

Jinnie Shin

University of Florida, Gainesville, FL, United States

Klaas Sijtsma
*Tilburg University, Department of Methodology and
 Statistics TSB, Tilburg, the Netherlands*

Sandip Sinharay
*Educational Testing Service, Princeton, NJ,
 United States*

Jesse R Sparks
*Educational Testing Service, Princeton, NJ,
 United States*

Jessaca Spybrook
*Western Michigan University, Kalamazoo, MI, United
 States*

Tracy M Sweet
*Measurement, Statistics and Evaluation Program,
 Department of Human Development and Quantitative
 Methodology, University of Maryland, College Park,
 MD, United States*

Sean Tan
*University of California, Graduate School of Education,
 UC Berkeley, Berkeley, CA, United States*

Jorge N Tendeiro
*Office of Research and Academia-Government-
 Community Collaboration, Education and Research
 Center for Artificial Intelligence and Data Innovation,
 Hiroshima University, Higashihiroshima, Japan*

Caitlin Tenison
*Educational Testing Service, Princeton, NJ, United
 States*

Christopher G Thompson
*Department of Educational Psychology, College of
 Education, Texas A & M University, College Station, TX,
 United States*

Kathryn N Thompson
*James Madison University, Harrisonburg, VA,
 United States*

Nickolay Trendafilov
*Department of Human and Social Sciences, University
 of Naples "L'Orientale", Naples, Italy*

Peter W van Rijn
ETS Global, Amsterdam, the Netherlands

Jeroen K Vermunt
*Department of Methodology and Statistics, Tilburg
 University, Tilburg, the Netherlands*

Matthias von Davier
*Boston College, Chestnut Hill, MA,
 United States*

Howard Wainer
Pennington, NJ, United States

Cindy M Walker
*Research Analytics Consulting, LLC, Pittsburgh, PA,
 United States*

Yibo Wang
University of Iowa, Iowa City, Iowa, United States

Kandauda AS Wickrama
*Department of Human Development and Family
 Science, University of Georgia, Athens, GA, United
 States*

Mark Wilson
*University of California, Graduate School of Education,
 UC Berkeley, Berkeley, CA, United States*

Stefanie A Wind
*The University of Alabama, Tuscaloosa, AL,
 United States*

Guanlan Xu
University of Iowa, Iowa City, Iowa, United States

Jiajun Xu
*Mary Frances Early College of Education, University of
 Georgia, Athens, GA, United States*

Kentaro Yamamoto
*Retired, former Educational Testing Service (Princeton,
 NJ, USA)*

Tianpeng Ye
University of Iowa, Iowa City, Iowa, United States

Chengbin Yin
*University of Maryland, College Park, MD,
 United States*

Ke-Hai Yuan
*University of Notre Dame, Notre Dame, IN,
 United States*

Diego Zapata-Rivera
*Educational Testing Service, Princeton, NJ,
 United States*

Jihong Zhang
University of Iowa, Iowa City, Iowa, United States

Mo Zhang
*Educational Testing Service, Princeton, NJ,
 United States*

Qi Zhang
*American Institutes for Research, Washington, VA,
United States*

Hui Zhao
*Department of Educational Psychology, Counseling, and
Special Education, The Pennsylvania State University,
University Park, PA, United States*

Mingying Zheng
University of Iowa, Iowa City, Iowa, United States

Rebecca Zwick
*Educational Testing Service, Princeton, NJ,
United States*

CONTENTS OF VOLUME 14

<i>Editors in Chief</i>	<i>v</i>
<i>Editors in Chief – Biographies</i>	<i>vii</i>
<i>Editorial Board</i>	<i>ix</i>
<i>Section Editors</i>	<i>xi</i>
<i>Volume 14: Quantitative Research and Educational Measurement</i>	<i>xix</i>
<i>Contributors to Volume 14</i>	<i>xxv</i>

EDUCATIONAL MEASUREMENT

Foundational Concepts

Past, present, and future of educational measurement <i>Brian E Clauser and Melissa J Margolis</i>	1
Educational testing: a review of standards, guidelines, and practices <i>Chris Domaleski</i>	15
Reliability <i>Klaas Sijtsma and Julius M Pfadt</i>	21
Validity <i>Suzanne Lane and Alexandra Pérez</i>	35
Fairness <i>Maria Elena Oliveri, Mya Poe, and Norbert Elliot</i>	45

Modeling Frameworks

Classical test theory <i>Michael Chajewski</i>	51
Generalizability theory and applications <i>Stella Y Kim, Jaime L Malatesta, and Won-Chan Lee</i>	59
Item Response Theory <i>Terry Ackerman, Ye Ma, Mingjia Ma, Juliana Cerentini Pacico, Yibo Wang, Guanlan Xu, Tianpeng Ye, Jihong Zhang, and Mingying Zheng</i>	72
Multidimensional item response theory <i>Sijia Huang and Li Cai</i>	86

Explanatory measurement	99
<i>Paul De Boeck and Mark Wilson</i>	
Diagnostic classification modeling	109
<i>Laine Bradshaw, Madeline Schellman, and Jiajun Xu</i>	
Bayesian networks	119
<i>Dakota W Cintron</i>	
Bayesian psychometrics	130
<i>Xiang Liu, Matthew S Johnson, and Zhuangzhuang Han</i>	
Nonparametric methods for modeling item responses: Mokken Scale Analysis	138
<i>Stefanie A Wind</i>	
Test Development	
Principled assessment design	146
<i>Mark Wilson and Sean Tan</i>	
Evidence-centered design	163
<i>Chengbin Yin and Robert J Mislevy</i>	
Cognition-centered design principles for digital assessment tasks and items	171
<i>Madeleine Keehner, Burcu Arslan, and Marlit Annalena Lindner</i>	
Common item formats and evaluation practices	185
<i>Doug Baldwin</i>	
Automatic item generation	193
<i>Mark J Gierl, Jinnie Shin, and Tahereh Firoozi</i>	
Automated test assembly	201
<i>Tim Davey</i>	
Computerized adaptive testing and multistage testing	209
<i>Andreas Frey</i>	
Score reporting: design and evaluation methods informed by research	217
<i>Priya Kannan</i>	
Test translation and adaptation	230
<i>Juliette Lyons-Thomas and Paulina Biernacki</i>	
Establishing the Quality of Scores	
Score equating: an aspirational form of score linking	236
<i>Neil J Dorans and Tim Moses</i>	
Differential functioning	249
<i>Cindy M Walker and Sakine Gocer Sahin</i>	
Model-data fit evaluation: item fit and model selection	260
<i>Alberto Maydeu-Olivares, Dexin Shi, Goran Pavlov, and Raul Ferraz</i>	
Model-data fit evaluation: aberrant response detection	273
<i>Jorge N Tendeiro</i>	
Model-data fit evaluation: posterior checks and Bayesian model selection	279
<i>Allison Ames Boykin, Nnamdi C Ezike, and Aaron J Myers</i>	

Subscores; what they are, why they are in high demand, and how to evaluate their quality <i>Sandip Sinharay and Gautam Puhan</i>	290
Statistical methods for detection of test fraud on educational assessments <i>Sandip Sinharay</i>	298
Rater effects modeling <i>Jodi M Casabianca</i>	308
Modeling item response times <i>Peter W van Rijn and Sandip Sinharay</i>	321
Quality control: response style modeling <i>Mirka Henninger and Thorsten Meiser</i>	331
Simulation studies for psychometrics <i>Brian C Leventhal, Allison J Ames, and Kathryn N Thompson</i>	341
Educational Policy	
Educational surveys: conceptual overview <i>Lale Khorramdel, Matthias von Davier, Irwin Kirsch, and Kentaro Yamamoto</i>	347
Educational surveys: methodological foundations <i>Paul A Jewsbury</i>	359
A brief introduction to contemporary standard setting methods <i>Sonya Powers, Angela D Bowzer, Mariann Lemke, Matthew A Brunetti, and Matthew N Gaertner</i>	369
Achievement growth modeling <i>Scott Monroe and Mina Lee</i>	381
Statistics: Value-Added Models <i>Jennifer E Broatch, Sarah E Manski, and Jennifer L Green</i>	390
Interdisciplinary Frontiers	
Automated scoring of constructed-response items in educational assessment <i>Mo Zhang and Randy E Bennett</i>	397
The current trends and opportunities for machine learning in learning analytics <i>Carol M Forsyth, Caitlin Tenison, and Burcu Arslan</i>	404
Interdisciplinary frontiers: computer-based process data analysis in educational measurement <i>Björn Nicolay, Florian Krieger, and Samuel Greiff</i>	417
Modeling of network structures <i>Tracy M Sweet and Samrachana Adhikari</i>	430
Application Areas	
Certification and licensure testing <i>Richard M Luecht</i>	441
Admissions testing in higher education <i>Rebecca Zwick</i>	451
Intelligent tutoring systems: a history and an example of an ITS for science <i>Janice D Gobert, Michael A Sao Pedro, Haiying Li, and Christine Lott</i>	460

Simulation- and game-based assessment <i>Kristen DiCerbo</i>	471
Assessment of interpersonal and intrapersonal skills <i>Patrick C Kyllonen</i>	477
Assessment of collaborative problem solving skills <i>Jessica Andrews-Todd and Carol M Forsyth</i>	494
Conversation-based assessment: current findings and future work <i>Diego Zapata-Rivera, Jesse R Sparks, Carol M Forsyth, and Blair Lehman</i>	504
Classroom assessment <i>James H McMillan</i>	519

STATISTICAL METHODS

Modeling Frameworks

Analysis of variance: univariate and multivariate approaches <i>Patrick J Rosopa and Bruce M King</i>	529
Analysis of covariance: univariate and multivariate approaches <i>Bryan Keller and Dobrin Marchev</i>	536
Linear regression analysis <i>Sijia Huang</i>	548
Nonlinear regression analysis <i>Hsin-Hsiung Huang and Qing He</i>	558
Hierarchical linear models <i>D Rindskopf</i>	568
Categorical data analysis <i>Dandan Chen and Carolyn J Anderson</i>	575
Generalized linear models <i>Peter K Dunn</i>	583
PCA and other dimensionality-reduction techniques <i>Nickolay Trendafilov and Michele Gallo</i>	590
Exploratory factor analysis <i>Nickolay Trendafilov and Kei Hirose</i>	600
Confirmatory factor analysis: foundations and extensions <i>Larry R Price</i>	607
A tutorial on structural equation modeling <i>Pui-Wa Lei and Hui Zhao</i>	619
Clustering approaches <i>James E Corter</i>	627
Latent class analysis <i>Jeroen K Vermunt</i>	637
An introduction to growth mixture models (GMM) <i>Tae Kyoung Lee, Kandauda AS Wickrama, and Catherine Walker O'Neal</i>	646

Nonparametric statistical methods <i>Wilco Emons and Klaas Sijtsma</i>	657
Causal models <i>Wen Wei Loh and Jee-Seon Kim</i>	670
Advanced machine learning <i>Oren E Livne</i>	684
Foundational Methods	
Research study design <i>Garret J Hall and Craig A Albers</i>	695
Sampling designs for educational research <i>Sharon L Lohr</i>	703
Continuous probability distributions <i>Sandip Sinharay</i>	710
Discrete probability distributions <i>Sandip Sinharay</i>	718
Descriptive statistics <i>Jennifer L Green, Sarah E Manski, Timothy A Hansen, and Jennifer E Broatch</i>	723
Foundational methods: descriptive statistics: bivariate and multivariate data (correlations, associations) <i>Debbie L Hahs-Vaughn</i>	734
Statistical inference: frequentist approaches <i>Jeffrey R Harring</i>	751
Resampling methods <i>Dennis D Boos</i>	762
Statistical inference: Bayesian approaches <i>Yang Liu and Tracy Sweet</i>	773
Foundational methods: power analysis <i>Qi Zhang, Jessaca Spybrook, Benjamin Kelcey, and Nianbo Dong</i>	784
Estimation approaches: least-squares, maximum likelihood and Bayesian <i>Jeffrey R Harring and Emma K Harring</i>	792
Missing data analysis <i>Brenna Gomer and Ke-Hai Yuan</i>	805
Displaying information: from woolly mammoths to the Great Migration <i>Howard Wainer and Michael Friendly</i>	819
Meta-analysis <i>Betsy Jane Becker and Christopher G Thompson</i>	842
<i>Index</i>	861

This page intentionally left blank

Past, present, and future of educational measurement

Brian E. Clauser and Melissa J. Margolis, National Board of Medical Examiners, Philadelphia, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

The past: the foundations of educational measurement	1
The London school and the development of classical test theory: Galton, Pearson, Spearman and beyond	2
Francis Galton	2
Pearson and Spearman	3
Contributions after Spearman	4
Intelligence testing and the eugenics movement	5
Alfred Binet	5
The Army Alpha test	6
The eugenics movement	6
The Stanford-Binet and the rise of the anti-testing movement	6
The present: development of current practice in educational measurement	7
Modern test theory	7
Item response theory	7
Generalizability theory	9
Validity theory	9
Testing practice	10
The future of educational measurement	11
Increased use of computational power	11
The science of test development	11
Process data	11
Fairness and consequences	12
Conclusion	12
References	12

As with most scientific fields, the discipline of educational measurement has been—and continues to be—defined by both theory and practice. In what follows, we provide an overview of the past, present, and future of the field. We begin by examining the historical events in the 1800s that led to the development of a coherent mathematical theory of test scores in the first half of the 20th century. We also describe the rise of large-scale testing beginning with the Army Alpha test in 1917 and the administration of IQ tests to millions of school children in the decade that followed. We continue by discussing the current state of educational measurement theory and practice and the ways that both continue to be shaped by cultural context. Finally, we consider directions and developments that are likely to define the future of the field.

The past: the foundations of educational measurement

Systematic attempts to measure educational achievement date back more than 1000 years to the advent of Chinese civil service testing (DuBois, 1970). By the end of the first millennium, this process had developed into something that we would now recognize as standardized testing: tests based on pre-established content administered under controlled conditions with set time limits. In some cases, written responses were even transcribed so that a test taker's calligraphy could not be recognized by the person scoring the response (DuBois, 1970).

Within the western tradition, university testing for ranking and graduation dates back hundreds of years; by the middle of the 18th century, written tests were being used to evaluate and rank Cambridge students seeking to graduate with honors in mathematics (Neale, 1907). The same testing approach soon spread to other subjects and other universities, and by the late 19th century standardized tests were being used in England to select individuals for the Indian civil service. These types of tests were the model that Francis Edgeworth had in mind when he wrote what may be the first paper to develop a framework for understanding the precision of test scores (Edgeworth, 1888).

Given this history, it is impossible to identify a single point that represents the beginning of the field of educational measurement. It is clear, however, that much of what we consider the theoretical underpinnings of educational measurement and a good bit of the practice began in the intellectual climate that existed in late 19th century England (See Table 1 for an overview of the early history of educational measurement.).

Table 1 Major events in educational measurement: 1850 to 1950.

1859	Darwin	Publication of <i>On the Origin of Species</i>
1869	Galton	Publication of <i>Hereditary Genius</i>
1888	Edgeworth	Publication of <i>The Statistics of Examinations</i> (first paper on the precision of test scores)
1889	Galton	Publication of <i>Natural Inheritance</i> and introduction of the correlation coefficient (in a separate paper)
1896	Pearson	Publication of <i>Mathematical Contributions to the Theory of Evolution, III: Regression, Heredity and Panmixia</i> (introduced the product moment correlation coefficient)
1904	Spearman	Publication of <i>The Proof and Measurement of Association between Two Things</i> (introduced formula for true-score correlation)
1905	Binet and Simon	Begin publication of a series of papers on the development of intelligence in children (1905–11)
1909	Galton	Publication of <i>Essays in Eugenics</i> (introduced the science of eugenics)
1910	Spearman and Brown	Publication of <i>Correlation Calculated from Faulty Data and Some Experimental Results in the Correlation of Mental Abilities</i> (introduced the Spearman-Brown formula)
1916	Lippman and Terman	Lippman-Terman Debate published in <i>The New Republic</i>
1924	Kelley	Publication of <i>Note on the Reliability of a Test</i> (Began Kelley's program of research that formalized classical test theory)
1925	Thurstone	Publication of <i>A Method of Scaling Psychological and Educational Tests</i> (presented model for placing items and examinees on a common scale)
1926	College Board	First administration of multiple-choice based SAT
1937	Kuder and Richardson	Publication of <i>The Theory of Estimation of Reliability</i> (introduced the KR-20 and KR-21 measures of reliability)

The London school and the development of classical test theory: Galton, Pearson, Spearman and beyond

In describing the foundation of classical test theory, Gulliksen (1950) stated that, “Nearly all the basic formulas that are particularly useful in test theory are found in Spearman’s early papers”. The theoretical framework that grew out of those formulas remains an essential part of educational measurement. We begin by examining the genesis of Spearman’s work.

The chain of events that led to these formulas can be traced back to November 24th, 1859: the date of the first publication of Charles Darwin’s *On the Origin of Species by Means of Natural Selection*. The direct impact of this text on the biological sciences is obvious, but its indirect impact on the social sciences also was profound; that impact was mediated through Darwin’s first cousin, Sir Francis Galton.

Francis Galton

Galton was the quintessential Victorian polymath. Like Darwin, Galton began his intellectual career studying medicine; also like Darwin, he never completed his medical studies. Galton contributed to numerous fields during his long life: he explored a part of Africa previously unknown to Europeans (Galton, 1853), created some of the first weather maps and discovered the phenomenon known as the anticyclone (Gillham, 2001), and wrote three books on fingerprints which significantly influenced the adoption of the methodology for individual identification (Galton, 1892, 1893, 1895; Gillham, 2001).¹

Of all of Galton’s efforts, those with the most lasting impact were motivated by his reading of Darwin’s *Origin of Species*. As Stigler (1986) commented, “Darwin’s theories opened an intellectual continent” (p. 267) for Galton to explore, and he spent much of the rest of his life mapping that continent. Though heredity played an important role in Darwin’s theory, the mechanisms of genetics and laws of heredity were not known.² Galton’s study of heredity had an important direct impact on what would become the field of genetics: it produced insights into how characteristics are passed from parents to children and launched the biometrics movement (Gillham, 2001). His work also had a critical—if somewhat less direct—impact on the development of educational measurement. Galton’s study of heredity led to the creation of the correlational procedures that Spearman would use as the basis for the formulas referenced by Gulliksen. Galton’s interest in heredity also led him to launch the science of eugenics (Galton, 1909), which gained widespread acceptance and shaped the cultural context in which early mental tests were developed and applied. (The role of eugenics in shaping educational measurement is discussed in more detail in a subsequent section.)

¹He also was responsible for a number of activities that make him seem like a bit of a crank: he created a beauty map of England (Galton, 1908) and produced a data-based argument to evaluate the efficacy of prayer (Galton, 1872).

²Although Mendel’s work was published when Galton began his experiments with pea plants (Mendel, 1866), it remained essentially unknown to the scientific community until late in the century.

Galton's work to understand heredity was both observational and experimental. His first major work on the topic, *Hereditary Genius*, cataloged instances in which individuals who were outstanding in a field (e.g., art, science, jurisprudence) tended to have relatives who were outstanding in the same field (Galton, 1869). Galton viewed this as evidence that psychological characteristics are inherited in the same way that physical characteristics are passed from parents to children.³

Galton also collected data to explore the laws of heredity. Like Mendel, Galton experimented with peas. Unlike Galton's work on hereditary genius, the results from his experiments with peas and the data he began collecting on human inheritance of physical characteristics provided numerical measures that called for mathematical analysis. His efforts to understand the relationships inherent in these data led to two important contributions to the field of statistics: The first was the concept of regression to the mean, and the second was the correlation coefficient (Stigler, 1986). These statistical contributions led to the widespread use of statistical methods in the social sciences and created a theoretical foundation that would be critical for the development of educational measurement.

Galton's work with peas was motivated by a simple question about how it was that if tall parents tended to have tall children and short parents tended to have short children, over time the characteristics of the population remained stable. Lacking data on humans, at Darwin's suggestion he decided to conduct experiments on peas (Galton, 1908). The results of his experiments showed that the characteristics of the next generation of peas regressed to the mean in a manner completely consistent with maintaining stable characteristics in the population (Galton, 1877). This phenomenon subsequently has been observed and described across a wide range of contexts in which probabilistic events play a role.

Twelve years later, Galton's quantitative approach to the study of heredity led to the introduction of the correlation coefficient. In his autobiography, Galton described how his examination of data relating the characteristics of parents and children led to a breakthrough: "At length, one morning, while waiting at a roadside station near Ramsgate for a train, and poring over the diagram in my notebook, it struck me that the lines of equal frequency ran in concentric ellipses" (Galton, 1908). Although Galton had graduated from Cambridge with a degree in mathematics, his mathematical skills were not outstanding. He collaborated with J. Hamilton Dickson to understand the quantitative relationship represented by the concentric ellipses. This collaboration yielded an index of correlation. Galton described his general approach in detail in *Natural Inheritance* (Galton, 1889). Although the focus of his work and that volume was on heredity, he suggested to the reader that understanding the methodology would be of broad value because "It familiarizes us with the measurement of variability, and with the curious laws of chance that apply to a vast diversity of social subjects" (p. 3).

Galton's work directly influenced numerous scientists that advanced the study of educational measurement. The two most critical links in that intellectual chain were Karl Pearson (who extended Galton's work on the correlation coefficient) and Charles Spearman (who applied correlation to the study of intelligence and in the process developed the tools that would become classical test theory).

Pearson and Spearman

Pearson studied mathematics at Cambridge. Unlike Galton, he excelled in those studies, becoming Third Wrangler in his class (indicating third place in the competitive examination for honors in mathematics). He became professor at University College London and later held the Galton Chair of Eugenics. Pearson was an incredibly productive researcher: The annotated bibliography of his works contains 648 entries (Morant and Welch, 1939) including the paper that introduced the chi squared statistic (Pearson, 1900). Although he wrote at least one paper on the statistics of testing (Pearson and Moul, 1927), most of his work focused on applying statistics to heredity and biology more generally.

For the purposes of this discussion, Pearson's most important work related to refining Galton's approach to correlation (Pearson, 1896). Pearson's name remains attached to the product moment correlation coefficient; his 1896 paper introduced that formulation, replacing Galton's method based on median and interquartile range with the more mathematically tractable version—based on means and variances—that we use today. Pearson's contributions created the foundation upon which Spearman would build to create the tools he needed to study intelligence.

Spearman served the British military for 15 years before resigning his commission and moving to Leipzig to study psychology with Wundt (Spearman, 1930). His doctoral work focused on the psychophysics-based measures that were the essence of the Leipzig school of psychology; interestingly, his dissertation *Die Normaltäuschungen in der Lagewahrnehmung*, which focused on spatial localization, seems to have had no impact on science (Tinker, 1932). His first important papers—published before he completed his doctoral work—were based on his work on the nature of intelligence, which was independent of the research going on at Wundt's laboratory. Those papers were *The Proof and Measurement of Association between Two Things* and "General Intelligence" Objectively Determined and Measured, both published in 1904.

In order to understand the nature of intelligence, Spearman examined how different measures of intelligence are related. He came to believe that intelligence was heavily influenced by a general factor. This interest made correlational analysis essential. Spearman was not the first to use correlation to explore this relationship. Clark Wissler's Ph.D. dissertation from Columbia University (Wissler, 1901) preceded Spearman's work and may have been the first published study in psychology to use Pearson's product moment correlation coefficient (Lovie and Lovie, 2010). Spearman's work, however, went beyond previous studies in an important

³There obviously are alternative explanations for Galton's observations. See Clauser (2021) for a discussion.

way: he recognized that to interpret the correlations between measures he needed to account for the impact of measurement error. To do this, he introduced the formula for estimating true score correlation (the correction for attenuation):

$$R_{\infty} = \frac{r_{11}}{\sqrt{r_{11}r_{11}}}$$

In this formula, the true correlation is equal to the observed correlation between the two measures divided by the square root of the product of the reliability of the measures. The formula appeared in *The Proof and Measurement of Association between Two Things* (Spearman, 1904). In that same paper, Spearman criticized Pearson for failing to account for the reliability of the measures in comparing the magnitude of correlations between physical and psychological characteristics.

Spearman's work was immediately attacked by Pearson (1904). Spearman then responded to the critique with *Demonstration of Formulae for True Measurement of Correlation* (Spearman, 1907). Spearman's last important contribution to classical test theory was *Correlation Calculated with Faulty Data* (Spearman, 1910); the paper was yet another defense against Pearson's attacks. More importantly, however, this paper presented what has become known as the Spearman-Brown⁴ formula,

$$R_{KK} = \frac{Kr_{11}}{1+(K-1)r_{11}}.$$

The formula shows the reliability of a test that has been lengthened as a function of the proportion by which it has been lengthened (K) and the reliability of the original test. The Spearman-Brown formula was of general use, but the most important application was for estimating the reliability of a single test based on the correlation between two half tests (such as those comprising only odd and only even items from the original test). The result is not only practical; it did much to define the way reliability would be conceptualized for the next 100 years.

Contributions after Spearman

Spearman created a foundation for test theory. In the years immediately following his contributions, his work was both applied and evaluated. Several authors examined the empirical accuracy of the Spearman-Brown formula by comparing predictions to empirical results (Holzinger and Clayton, 1925; Remmers et al., 1927; Ruch et al., 1926; Wood, 1926). Each of these studies confirmed the predictions based on the Spearman-Brown formula.

As Pearson's attacks on Spearman began to fade from view, the importance of Spearman's approach became more widely recognized and other researchers extended his work to create *classical test theory* as it was presented by Gulliksen (1950) and Lord and Novick (1968). No one played a more instrumental role in this process than Truman Kelley. Where Spearman had developed tools to advance his program of correlational psychology, Kelley was interested in advancing assessment; in this regard he may be the first true psychometrician. In his 1924 and 1925 papers, Kelley set out the building blocks of test theory, showing that if we have two forms of a test such that $x_1 = a + e_1$ and $x_2 = a + e_2$, where x is the observed score, a is the true score, and e is a "chance factor",⁵

$$r_{12} = \frac{\sigma_a^2}{\sigma_x^2}.$$

Using this starting point, he shows the relationship between the reliability of a measure and the correlation between the measure and the related true score:

$$r_{1\infty} = \sqrt{r_{11}}$$

or equivalently,

$$r_{11} = r_{1\infty}^2.$$

The next critical contribution to the development of test theory came in the form of a paper by Kuder and Richardson (1937). At a time when computing was done by hand, with a slide rule, or with a cumbersome mechanical calculator, the paper provided simplified computational procedures for calculating reliability. This was valuable to practitioners, but perhaps more importantly it established the view that from the perspective of test theory, reliability was to be conceptualized as split-half rather than test-retest reliability. Kuder and Richardson begin with the assumption that $a, b, \dots, n$ and $A, B, \dots, N$ are corresponding items in two forms of a test. They state that the tests are equivalent if the items in each pair (a and A , b and B , etc.) are interchangeable (i.e., the items in one form have the same difficulties and inter-item correlations as those in the other form). Based on these assumptions, they present the KR-3:

$$r_{tt} = \frac{\sigma_t^2 - \sum_1^n p_i q_i + \sum_1^n r_{ii} p_i q_i}{\sigma_t^2}.$$

⁴Brown refers to William Brown. He independently developed a proof of the same formula in a paper published immediately following Spearman's in the same journal (Brown, 1910). Interestingly, Brown was a follower of Pearson and much of his dissertation—on which this article was based—recapitulated Pearson's previous attacks on Spearman.

⁵The original equation given by Kelley did not include a subscript for the variance in the denominator. The subscript has been added because this form is likely to be more familiar to contemporary readers.

In this equation, r_{tt} is the reliability of interest, σ_t^2 is the variance of the test, p is the difficulty for a specific item, and q is $1-p$. Seventeen equations later they present the KR-20, derived from KR-3 with the assumption that all inter-correlations are equal:

$$r_{tt} = \frac{n}{n-1} \cdot \frac{\sigma_t^2 - npq}{\sigma_t^2}.$$

Fourteen years later, Cronbach's paper on coefficient alpha extended this work and represents a kind of capstone on the development of classical test theory (Cronbach, 1951). In that paper, Cronbach showed that alpha represents the mean of all possible split-half coefficients and that it is also "the value expected when two *random* samples of items from a pool like those in the given test are correlated" (p. 331). As Cronbach pointed out, KR-20 is a special case of alpha that applies to tests comprising dichotomously scored items.

From what we have described in the previous paragraphs it should be clear that Spearman's focus on correlational analysis and the introduction of the Spearman-Brown formula provide the foundation on which the theoretical basis for educational measurement was built. The introduction of the Spearman-Brown formula represented the first important step toward conceptualizing reliability as a statistical property of the test (although it is not clear that Spearman recognized that fact). Developments over the next half century were dependent on the way of thinking about test scores that Spearman introduced. During the 50 years over which classical test theory developed, the practice of testing also evolved. Though Galton, Pearson, and Spearman influenced that development as well, their influence was less direct. We discuss the evolution of testing practice in the next section.

Intelligence testing and the eugenics movement

Alfred Binet

During the same time period in which Spearman was writing the papers that would spur the development of classical test theory, Alfred Binet was in France doing research that would shape intelligence testing across the globe for much of the 20th century. Binet's career was eclectic if not totally disjointed. In 1882 he met Jean Martin Charcot and Charles Fere and began spending much of his time in Charcot's laboratory observing and experimenting with the effects of hypnosis on hysterics (Binet and Fere, 1887). Binet's doctoral dissertation was entitled, *A Contribution to the Study of the Subintestinal Nervous System of Insects* (Wolf, 1973). By 1905, Binet had turned his attention to education—specifically to the identification of school children who would benefit from "special" educational programs. As Binet and Simon (1916) described the motivation for their work on intelligence tests:

In October 1904 the Minister of Public Instruction named a commission which was charged with the study of measures to be taken for insuring the benefits of instruction to defective children. ... They decided that no child suspected of retardation should be eliminated from the ordinary school and admitted to a special class, without being subjected to a pedagogical and medical examination from which it could be certified that because of the state of his intelligence, he was unable to profit, in an average measure, from the instruction given in the ordinary schools (p. 9).

They go on to say that such decisions about placement,

... have at all times too much the nature of the arbitrary, of caprice, of indifference. Such a condition is quite unfortunate because the interests of the child demand a more careful method. To be a member of a special class can never be a mark of distinction, and such as do not merit it, must be spared the record (pp. 9–10).

This work resulted in a series of papers that were published between 1905 and 1911; the papers were subsequently translated and published in the United States in Binet and Simon (1916). In the first paper, Binet and Simon discuss the limitations of the then-current approaches to diagnosis, arguing that (a) approaches based on the physical cause of the disability provide no insight into the extent of the intellectual deficit (e.g., *epileptic idiocy*) and (b) approaches that provide diagnostic criteria based on observable manifestations of intelligence have been too vague to be useful. For example, one diagnostic framework stated that "in profound idiocy *the attention is fugitive*," while in imbecility, *the attention is fleeting*" (p. 23).

Binet and Simon (1916) then present procedures for carrying out 30 tests under standardized conditions. The tests were designed to discriminate between different levels of development. At the simple end of the continuum, the experimenter is instructed to move a lit match across the subject's field of vision and record whether he or she follows it visually. The most sophisticated test requires the subject to define abstract terms: e.g. "What difference is there between weariness and sadness?"

With respect to contributions to educational measurement, Binet's work may be viewed as contributing more to practice than to theory, but in developing the tests he made important methodological and conceptual breakthroughs. The first of these breakthroughs relates to the empirical evaluation of tests (items) for inclusion in the scale. Binet and Simon tried out numerous tests and retained those that discriminated between children of different ages. There does not appear to be evidence that they actually plotted what we might now call *item characteristic curves* to evaluate tests for inclusion, but several of the tables in their paper on

application of these new methods present scores as a function of age and could not be more informative if they were converted to graphic form (Binet and Simon, 1916). The second conceptual contribution that Binet and Simon made was the idea of a norm-referenced interpretation of scores.

The Army Alpha test

As noted previously, Binet and Simon created the prototype for what we know as intelligence tests. Like Spearman's work, which ultimately was advanced by Truman Kelley and numerous other psychologists and statisticians, Binet's work became influential because it was advanced by two Americans psychologists: Henry Goddard and Lewis Terman. Goddard had been invited to open a research center at the New Jersey School for Feeble-minded Boys and Girls, in Vineland. Goddard's efforts to evaluate the children at the school left him in search of improved means of assessing developmental status. During a tour of Europe in 1908 he encountered Binet's work and immediately recognized that it would provide a framework for the improved assessment he was seeking (Zenderland, 1998). Terman was researching intelligence in California at the same time; he also recognized the importance of Binet's approach.

When the United States entered the First World War, both Goddard and Terman were invited to be part of a group convened by Robert Yerkes—the president of the American Psychological Association—to explore how psychology could contribute to the war effort. The group met at Vineland and created the basis for the Army Alpha test. The intention was to create an intelligence test that could be used to make placement decisions for the thousands of recruits who would support the war effort (Sokal, 1990). The focus of their effort began with a variation on the model created by Binet. The group quickly realized that for large-scale testing of the sort that they were proposing, an individually administered test like Binet's would not be feasible. Terman had brought a copy of an intelligence test appropriate for group administration to the Vineland meeting.⁶ This test, which relied heavily on a format we would come to know as multiple-choice questions, became the prototype for the Army Alpha. The format allowed for efficient scoring of the test by clerical staff, and by the time the war ended in November 1918, more than 1,700,000 men had been examined (Yoakum and Yerkes, 1920). Although there is some question as to whether the scores were widely used to make decisions about the recruits, there is no question that the test was instrumental in launching large-scale intelligence testing in the United States; it also was a critical step in the rise of the multiple-choice item format.

The eugenics movement

As we mentioned, in addition to introducing the analytic procedures that would be critical to the development of test theory, Galton was a founder of the eugenics movement. The movement became widely popular in England and the United States and was closely tied to intelligence testing. In Galton's writing, the focus often was on approaches for encouraging eugenically superior individuals to have more children (Galton, 1909; Gillham, 2001). Later eugenicists promoted more extreme measures, including preventing those viewed as inferior from having children and preventing members of specific ethnic groups from entering the country. In 1912 Spearman cowrote, "One can conceive the establishment of a minimum index (of general intelligence) to qualify for parliamentary vote, and above all, for the right to have offspring" (Hart and Spearman, 1912, p. 79). After Galton's death, Pearson held the Galton Chair in Eugenics at University College London, and in a speech given at his retirement in 1934 he praised Reichskanzler Hitler's "vast experiment" "to regenerate the German people" (Pearson, 1934, p. 23).⁷

Goddard also explicitly promoted a eugenics agenda. His book *The Kallikak Family* (Goddard, 1912) was designed to show the harm to society that resulted from allowing eugenically inferior individuals to breed. (Kallikak was a pseudonym for a family that Goddard studied; Deborah Kallikak lived at Vineland). The book was a kind of mirror image of Galton's *Hereditary Genius*; instead of tracing recurrent instances of genius within the same family, Goddard traced instances of mental deficiency within a single family.

Before World War II, there can be little doubt that educational measurement and eugenics were closely tied. Intelligence testing served this connection well. At the same time, it should be remembered that during this period the eugenics movement was mainstream and respected. This changed dramatically after the atrocities of the German eugenics project became widely known.

The Stanford-Binet and the rise of the anti-testing movement

In 1916, Terman published his revision of Binet's scale. In 1918, this was followed by the publication of group tests; within 30 months, approximately four million children were tested (Cronbach, 1975). Although intelligence testing and eugenics were mainstream in the 1920s, they were far from universally accepted. In 1922, Walter Lippmann—author and cofounder of the *New Republic*—wrote a six-part attack on intelligence testing and in so doing launched the anti-testing movement that has continued in one form or another for a century. Much of the article was specifically targeted at Terman. Although Lippmann conceded that standardized testing may be fairer and more accurate than the subjective system of grading prevailing in public schools at the time, he raised concerns about the potential of the tests to stigmatize children. He also expressed concern that rather than motivating remediation, low scores may provide justification for failing to provide the resources needed for education. He then moved on to argue against the conclusion that intelligence—as measured by these tests—is inherited and largely unchanged over time. Lippmann's presentation and Terman's response are heavily dosed with sarcasm, but the main points are not so far removed

⁶The test had been developed by his graduate student, Arthur Otis.

⁷This statement clearly reflects Pearson's enthusiasm for eugenics. It should be noted, however, that it was made years before Kristallnacht and a decade before the atrocities of the Holocaust were widely known.

from the essence of the debate that followed the 1969 *Harvard Educational Review* publication of Jensen's "How Much Can We Boost IQ and Achievement?" and Herrnstein and Murray's 1994 publication of *The Bell Curve*. Jensen (1969) and Herrnstein and Murray (1994) both entertain the idea that group differences in scores on intelligence tests are the result of genetic differences. These views led to backlash and accusations—from both within and outside of the academic community—that the authors were racist; the publications also fanned the flames of the anti-testing movement.

The previous pages describe how test theory and testing programs developed in Europe and the United States over the course of the 19th and 20th centuries. We now turn our focus to the measurement theories and applications that define educational measurement in the beginning of the 21st century.

The present: development of current practice in educational measurement

In this section, we examine the issues in theory and practice that define contemporary educational measurement. From the theoretical perspective we consider item response theory (IRT), generalizability theory (g-theory), and the development of a coherent theory of test validity. Practical considerations include a broadening of testing applications and widespread adoption of computer technology for test administration (See Table 2 for an overview of the major events of this period.).

Modern test theory

By the 1950s, classical test theory provided a well-developed framework for evaluating test scores. The value of the theoretical framework was broadly understood; the limitations of the model also were apparent. The framework Spearman had introduced was well suited to the study of correlational psychology; it was less well suited to the understanding, development, administration, and improvement of tests. Classical test statistics are a function of the set of items administered as well as the sample of test takers responding to those items. The classical model for reliability also assumes a true score and an undifferentiated source of error. This model is less useful for understanding the performance of assessments with more complex designs (e.g., tests where error is introduced by the selection of items and the raters who score those items). Similarly, early conceptualizations of validity tended to focus on correlation with criterion measures (Kane and Bridgeman, 2021; Thurstone, 1931); this narrow focus neglected a range of issues that now are recognized as important in contemporary validity theory.

Item response theory

IRT has become all but ubiquitous within large-scale educational assessment. As noted above, classical test theory is limited in that the resulting statistics are a function of the set of items that were administered and the sample of test takers responding to those items. For example, the difficulty of an item is defined as the proportion of test takers in the sample responding correctly to that item; if the item were administered to a different (i.e., more or less proficient) sample of test takers, the item difficulty would change.

Table 2 Major events in educational measurement: 1950 to the present.

1950	Gulliksen	Publication of <i>Theory of Mental Tests</i> (the first comprehensive monograph on test theory)
1951	Cronbach	Publication of <i>Coefficient Alpha and the Internal Structure of Tests</i>
1952	Lord	Publication of <i>A Theory of Test Scores</i> (introduces framework for item response theory)
1960	Rasch	Publication of <i>Probabilistic Models for Some Intelligence and Attainment Tests</i> (introduces Rasch model)
1968	Lord and Novick	Publication of <i>Statistical Theories of Mental Test Scores</i> (makes Birnbaum's work on logistic IRT models widely available)
1969	Jensen	Publication of <i>How Much Can We Boost IQ and Scholastic Achievement?</i> (raises questions about the connection between intelligence and race)
1969	NAEP	First administration of the National Assessment of Educational Progress
1971	Cronbach	Publication of <i>Test Validation</i> in 2nd edition of <i>Educational Measurement</i> (incorporates construct validity into the structure of validity theory)
1972	Cronbach et al.	Publication of <i>The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles</i> (introduces generalizability theory)
1989	Messick	Publication of <i>Validity</i> in 3rd edition of <i>Educational Measurement</i> (highlights construct validity and the role of consequences in validity)
1995	TIMSS	First administration of the Trends in International Mathematics and Science Study
2000	PISA	First administration of the Program for International Student Assessment
2001	PIRLS	First administration of the Progress in International Reading Literacy Study
2006	Kane	Publication of <i>Validation</i> in 4th edition of <i>Educational Measurement</i> (presents validation as a structured argument)

IRT represents an explicit effort to produce estimates of item and test-taker characteristics (parameters) that are invariant across samples. The common way of formulating the resulting relationships is through an equation such as:

$$P(\theta) = \frac{1}{1 + e^{-a(\theta-b)}},$$

which shows that the probability of a test taker with a specific proficiency level θ responding correctly to a(n) (dichotomously scored) item is a function of the difference between test taker proficiency, item difficulty b , and item discrimination a .⁸ Creating estimates of item and test-taker parameters that are invariant across samples clearly has theoretical appeal; doing so also has numerous practical advantages. These models created new and more flexible ways to equate test forms. They also created a framework for targeting test items to the proficiency level of a specific test taker—opening the way for computer-adaptive testing—and targeting items to a cut score to minimize classification error. Use of IRT for optimizing test development, equating, and administration is now commonplace, but development of practical applications took time, as estimation procedures and computer technology needed to catch up with the underlying mathematical theories.

Development of item response theory in the United States

In science, it is occasionally the case that researchers working independently arrive at similar breakthroughs at nearly the same time.⁹ The development of IRT—the origins of which occurred primarily in the United States and Denmark with contributions from researchers in other countries (in particular, the Netherlands)—is such a situation. Although the birth of the theory tends to be associated with the publication of Frederic Lord's (1952) *A Theory of Test Scores*, several researchers had contributed important ideas during previous decades. For example, Thurstone (1925) demonstrated how items and people could be scaled together; during the Second World War, Lazarsfeld (1950) did important work on scaling latent traits such as the proficiency that would be integral to IRT. Despite these other contributions, it is clear that Lord's monograph was the first to lay out the basics of the theory as we know it today. Publication of the monograph was soon followed by a paper published in *Psychometrika* in which Lord (1953) presented the two-parameter normal-ogive model:

$$P(\theta) = \int_{-\infty}^{a(\theta-b)} \frac{1}{\sqrt{2\pi}} e^{-.5z^2} dz.$$

This equation defined the theoretical basis for IRT, but the mathematical intractability of the formulation made actual use of the model all but impossible. It took nearly 15 years and the publication of Birnbaum's work in Lord and Novick's (1968) *Statistical Theories of Mental Test Scores* for the normal-ogive model to be replaced with the logistic form that we use today.

Rasch Measurement

This two-decade-long effort by Lord to create a practically useful IRT model was paralleled by Georg Rasch in Denmark. Rasch was a trained statistician who had studied with R. A. Fisher; he began working on educational data in the 1950s. As he later described his work, "When I got the data I made the guess that it might be a good idea to try the multiplicative Poisson model. It turned out to fit the data quite well." (Rasch, 1980, p. xiii). As this quotation suggests, Rasch started by fitting a model to the data. Over time, however, Rasch came to stress the importance of the data fitting the model rather than the model fitting the data.

What has become known as *Rasch measurement* has focused on the importance of two characteristics of measurement procedures. The first is that the raw score is a *sufficient statistic* for the modeled score. Given that sufficiency is a concept central to Fisher's work, it is not surprising that Rasch recognized its importance. The other characteristic is what Rasch referred to as *specific objectivity*. This refers to the fact that measurement is invariant across samples. Because of the emphasis Rasch placed on the data fitting the model and the need for sufficient statistics, the basic Rasch model is equivalent to Birnbaum's logistic model with a single parameter representing item difficulty—it explicitly excludes a discrimination parameter.

Extensions and Applications

The models developed in the US and Denmark were rapidly extended by other researchers. Within the Rasch framework, Andrich (1978) and Masters (1982) developed models for items that are scored polytomously. Samejima (1969, 1972) and Bock (1972) provided related models that can be viewed as extensions of Lord and Birnbaum's work. van der Linden and Hambleton's (1997) *Handbook of Modern Item Response Theory* presented twenty-seven models - including polytomous and multidimensional extensions of the logistic models as well as models incorporating response time.

Although the development of these models has been important, the direct application of IRT to testing problems has been even more critical to the practice of educational measurement. IRT created a practical approach to constructing large item banks that can be used to build equivalent test forms (through an approach referred to as pre-equating) and to support

⁸When applied to multiple-choice questions, the model is also commonly modified to reflect the fact that even test takers with very low proficiency will have a non-zero chance of selecting the correct answer.

⁹Well-known examples of this phenomenon include the development of the theory of evolution through natural selection by both Darwin and Wallace, the creation of calculus by Newton and Leibniz, and development of the Spearman-Brown formula.

computer-adaptive testing (sometimes referred to as CAT). The mathematical theories and conceptual applications largely were defined by the time Lord published *Applications of Item Response Theory to Practical Testing Problems* (Lord, 1980), but practice lagged because estimation procedures and robust software for implementing these models were unavailable. This would change in the 1990s.

Generalizability theory

The development of IRT was motivated by the desire to produce test-taker and item parameter estimates that are invariant across samples. As the theory developed, it also provided a convenient model-based approach to estimating conditional standard errors for proficiency estimates. IRT has, however, remained linked to the basic classical test theory model in that it conceptualizes error as being undifferentiated. In contexts in which tests are comprised of a random sample of items that can be mechanically scored as correct or incorrect, this mode is adequate. For conditions in which tasks are scored by human judges or items are sampled from predefined content categories, however, this model likely will be inadequate (and in some cases, grossly inadequate). Generalizability theory is an effort to model the different facets of the testing process so that their contributions to measurement error can be appropriately represented in determining the precision of test scores and in planning administration procedures.

The concepts formalized as generalizability theory by Cronbach et al. (1972) were described first by Francis Edgeworth (1888). The basic approach is based on analysis of variance procedures which allow for isolating the different components (or facets) of the measurement procedure that contribute to score variance—particularly those facets that contribute to measurement error.¹⁰ Edgeworth conceptualized tests as being constructed through random sampling (the central assumption of generalizability theory). He conceptualized the error impacting a test score as comprising multiple facets including what we now call task effects, rater effects, rater-by-task interactions, and person-by-task interactions. He noted both that the total error would equal the square root of the sum of the squared effects and that the impact of any effect will be reduced by the square root of the number of observations.

Although Edgeworth outlined much of what we would currently describe as generalizability theory, his papers appear to have been forgotten by later researchers. Some researchers (e.g., Burt, 1936, 1955) returned to this problem earlier in the 20th century, but the major contribution came when Cronbach et al. (1972) published their monograph. This provided what the authors described as a *liberalization of classical test theory*, and their approach has defined the state of the art for evaluating the precision of test scores for the last half century. Despite the obvious theoretical improvement that generalizability theory offered, classical test theory approaches—such as coefficient alpha—have remained in common use; in his final published paper, Cronbach (2004) continued to argue that correlational methods of estimating reliability should be replaced with analysis-of-variance-based procedures. He believed that the earlier approaches had served correlational psychology well, but that they were inadequate for evaluating the kind of testing that had become common by the end of the 20th century.

Validity theory

Many of the theoretical developments that define educational measurement are based explicitly on statistical models. Interestingly, perhaps the most important theory—*validity theory*—is more philosophical, or conceptual. As with the other theoretical frameworks that support educational measurement, this one has developed over the course of the last century. There is a sense in which the concept of validity has been stable over time, consistently referring to the extent that a test score is useful for its intended purpose.¹¹ What has changed is the range of intended purposes that have been considered and the extent to which validity has gone from an *ad hoc* evaluation to a systematic and principled framework.

In the period following World War II, considerable attention was being given to tests that would predict future performance. This included selection and placement tests in industrial and educational settings. Because the purpose of these tests was to predict a criterion, validity was frequently expressed as a correlation between the test score and the criterion. Thurstone (1931) defined validity exclusively within this framework, and Cureton's article on validity (1951) in the first edition of *Educational Measurement* (Lindquist, 1951) focused substantially on this aspect of validity. By the time the second edition of *Educational Measurement* was published (Thorndike, 1971), the conceptualization of validity had expanded substantially. It was recognized that different approaches to validity may be necessary for different types of tests: Correlations may be appropriate for tests that are used to make predictions, and approaches that evaluate the content of the test may be needed for achievement tests. In the case of tests for which there is no readily available criterion and no definitive content (such as psychological tests), Cronbach (1971) recommended a construct validity framework. Unlike predictive validity and content validity, which focus on a single outcome, construct validity requires a program of research to evaluate whether the test scores behave in a manner that is consistent with the research and theory related to the construct. Cronbach's (1971) presentation of validity was perhaps most influential in establishing that test validity would require an ongoing program of research regardless of the purpose of the test.

By the publication of the third edition of *Educational Measurement* (Linn, 1989), construct validity had become an organizing principle for validity theory. In addition to elevating the importance of construct validity, Messick (1989) gave considerable

¹⁰Edgeworth (1885) published a paper on analysis of variance decades before Fisher's work (Fisher, 1925) made the methodology common practice.

¹¹Kelley (1927) defined validity as the extent to which a test "really measures what it purports to measure".

attention to the role of consequences in validity. Incorporating consequences into the conceptualization of test validity consciously extends the scope beyond an empirical verification of the usefulness of test scores and explicitly introduces value judgments into the process.

The 2006 edition of *Educational Measurement* (Brennan, 2006) again expanded the view of validity. Kane's (2006) presentation acknowledges the importance of the types of validity evidence discussed in the three earlier editions of *Educational Measurement* and reconceptualizes that evidence as part of a structured argument that supports the interpretations and uses of test scores. This argument-based approach emphasizes that the overall validity argument is only as strong as its weakest part. This approach is consistent with the views of Cronbach and Messick that the process of test validation requires a program of research. It also serves to focus that research on the aspects of the argument which might be most dubious.

Testing practice

Educational tests have become ubiquitous in the modern world, and much of this testing can be seen as a natural evolution from the intelligence and achievement tests that were in wide use in the United States by the middle of the 20th century. Although standardized tests are used in most, if not all, countries, the use of these tests varies across countries.

High-school graduation tests and tests for admission to colleges and universities are common, although the specifics vary by country. In China, the National College Entrance Examination is taken by approximately 10,000,000 students each year. This test determines eligibility for post-secondary education, and because of the impact of the results on later career options, many students spend years actively preparing for it (Larmer, 2014). In Finland, the Matriculation Examination is administered at the end of high school; it acts both as a graduation examination and a requirement for candidates who want to continue their studies at university. In England, students typically take numerous standardized tests (as many as 20) during the final years in high school to earn their general certificate of secondary education and to complete their A levels (which are required for university admissions). In the United States, the two major college entrance examinations—SAT and ACT—are taken by roughly 4,000,000 students each year. The SAT and ACT have been particularly targeted by the anti-testing movement during recent years, and some colleges and universities are no longer requiring test scores as part of the application process.

In addition to high school graduation and college entrance examinations, many countries administer achievement tests. Both the use and purpose of these tests vary across countries and across jurisdictions within countries. Finnish children are unlikely to take this sort of standardized test except in the form of the Matriculation Examination mentioned in the previous paragraph. Germany administers standardized tests in grades 3 and 8 to support improvements in educational practice. The United States makes substantially greater use of standardized achievement tests (although the specifics vary from state to state). Among the first large-scale achievement tests in the United States was *The Iowa Test of Basic Skills* (ITBS). It was introduced in Iowa in 1935 and was quickly adopted by states across the country (Briggs, 2021). Though the initial focus of this and other large-scale achievement tests was on evaluating individual students, that focus has shifted over time. School-level achievement testing continues in widespread use; for example, some states require passing a standardized achievement test for high school graduation. Beyond this, in a controversial and much debated practice, achievement tests are being used to evaluate the performance of teachers. These tests also are used to evaluate schools and school systems.¹²

The most dramatic aspect of the evolution of achievement tests is reflected in national examinations such as the *National Assessment of Educational Progress* (NAEP) in the US and international examinations such as the *Trends in International Mathematics and Science Study* (TIMSS), *Progress in International Reading Literacy Study* (PIRLS), and *Program for International Student Assessment* (PISA). These tests often are based on complex statistical and psychometric models that use sophisticated sampling procedures to support aggregate proficiency estimates without supporting estimates at the level of individual test takers. The results allow for tracking of how states/provinces or countries are performing.

In addition to the ever-broadening applications of testing, another defining feature of current practice is the use of computers in assessment and the potential offered by use of this technology. From a content perspective, paper-and-pencil administration is perfectly acceptable if test material is limited to text and simple graphics; for more complex content, computerization has opened the way for high-quality and low-cost use of color pictures, has greatly simplified the use of audio and video as part of the stimulus material, and makes interactive tasks and simulations possible. In addition to broadening the range of materials that can be presented, computerization also has made a range of approaches to test administration possible: Computer-adaptive testing—in which test items are targeted at the estimated proficiency level of the test taker—has become common, and linear-on-the-fly testing (LOFT)—in which item exposure is controlled by constructing a potentially unique test form for each test taker—is well established. Computerization also has facilitated automated scoring of constructed responses (including essay and short answer questions) and complex performance tasks such as design problems for architects (Braun et al., 2006) and patient management simulations for physicians (Margolis and Clauser, 2020).

In addition to the benefits associated with administration and scoring, computerization has created previously unavailable sources of data about how test takers respond to test items. These data have the potential to substantially improve testing.

¹²Interestingly, evaluation of public schools in Boston was the motivation behind the introduction of some of the first standardized written examinations used in American schools: those introduced by Horace Mann in 1845.

Information on how test takers change answers is now routinely collected as is information about how much time test takers use and how they distribute available testing time across test items. Data of this sort can be useful in evaluating the validity of the intended interpretations and uses of test scores as well as for improving test materials and administration procedures (Margolis and Feinberg, 2020).

The future of educational measurement

Although it is impossible to know the future of educational measurement, a number of current trends are almost certain to shape the future of the field. In the following paragraphs, we describe four of these trends.

Increased use of computational power

Computational power has defined practice in educational measurement for generations. Spearman and Pearson introduced formulas for approximating the correlation coefficient with less computational effort, and the culmination of Kuder and Richardson's landmark paper was the KR-21, which allowed practitioners to calculate reliability based on statistics that were likely to be readily available and so save computational effort. Similarly, widespread use of many IRT applications had to wait for computational technology to catch up with theory (see Luecht and Hambleton, 2021). As practice historically was limited by computational power, more recently it has been enabled by that power. Artificial intelligence, machine learning, and computational linguistics (the application of computer science to the analysis of written and spoken language) already have made an impact on educational measurement, and automated scoring procedures that use these technologies for essays and other constructed responses have been studied for more than a half century (Page, 1967) and used operationally for decades (Margolis and Clauser, 2020). These technologies are likely to become more prevalent in the future, and they have the potential to support automated assessment of both written and spoken responses. Systems driven by artificial intelligence will not only tutor students to improve their writing and math skills; they will allow medical students to practice interviewing patients to improve their clinical decision making and allow for both teaching and assessment of a wide range of similarly sophisticated skills in other disciplines.

The science of test development

For as long as test questions have existed, test developers have engaged in the art of item writing. In recent years, test development efforts have slowly shifted away from the artistic toward the scientific, adopting principled approaches to test development (e.g., Ferrara et al., 2017). *Principled design* refers to constructing assessments in a way that ensures that the interpretations made based on the test scores are aligned with and supported by the evidence from the assessment; it includes approaches such as evidence-centered design (Mislevy et al., 2006), assessment engineering (understanding the characteristics of items that impact item performance; Luecht, 2012), and efforts to generate items using computer technology (e.g., automated item generation, or AIG). Although evidence-centered design has been well delineated, assessment engineering and computer-generation of test items represent technologies that are still in the early stages of development. It is likely that this work will help to define the next generation of educational assessments and that computational linguistics and machine learning will support both the development and operationalization of these efforts. These areas of research are critical because the field of educational measurement can never advance beyond the quality of the test material used in that measurement.

Process data

Closely associated with a more scientific approach to test development is the use of a wide range of data to evaluate how test takers respond to test items. These data often are referred to as *process data* or *meta data*. In some cases, process data can be collected seamlessly and effortlessly when the test is delivered via computer. For example, commonly captured process data include the amount of time test takers spend on each item—both before and after recording an answer—and how this timing information relates to the probability that an item is answered correctly. This type of timing data has been useful for many purposes, such as evaluating whether an assessment is speeded, determining a test taker's level of engagement, and even identifying test takers who may have had prior access to the test material (Margolis and Feinberg, 2020). Another type of process data can be collected with the use of eye-tracking technology, which provides even more specific information about how test takers read, process, and respond to items (Oranje et al., 2017). It is a near certainty that these types of data will be used more widely in the future to provide test developers with insight into the cognitive processes involved with responding to test items. As more use is made of process data, test developers will better understand *how* examinees respond to test items, and educational measurement is likely to become more closely linked to cognitive psychology.

Fairness and consequences

The role of *consequences* and the issue of fairness within the framework of test validity has received considerable attention during the last three decades. Issues that may limit opportunities for already disadvantaged minorities—such as score differences across racial or socioeconomic groups—have been a particular focus of the anti-testing movement. The historical association of intelligence testing with the eugenics movement makes the issue of fairness even more critical. Messick (1989) often is viewed as the major driver for the inclusion of consequences in the understanding of validity, but a careful reading suggests that Cronbach (1971, 1988) and Kane (2013) are equally strong advocates for the importance of consequential validity. As the centrality of consequential validity becomes more widely recognized, issues of fairness in testing will have a greater role in the future of educational measurement (Dorans and Cook, 2016).

Conclusion

We conclude with two points. First, though we acknowledge that predictions are risky, it nevertheless seems a certainty that the basic methodologies that have shaped the past and present of educational measurement—classical test theory, IRT, generalizability theory, and validity theory—will remain an essential part of the future of the field. The second point is that we have provided a perspective on educational measurement that is far from exhaustive. We have emphasized the theoretical advances that make educational measurement a scientific discipline, and in so doing we have neglected large domains—such as day-to-day classroom assessment—and many essential aspects of test administration—such as equating and standard setting—that continue to be the focus of careful study. Although the picture we have painted is unavoidably incomplete, we hope that what is clear are the ways in which the past, present, and future of educational measurement are connected through the theories and practice that have created and sustained this field.

References

- Andrich, D., 1978. A rating formulation for ordered response categories. *Psychometrika* 43, 561–573.
- Binet, A., Fere, C., 1887. *Le magnetisme animal*. Alcan, Paris, FR.
- Binet, A., Simon, T., 1916. *The Development of Intelligence in Children*. Translated by E. S. Kite. Williams and Williams Company, Baltimore, MD.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37, 29–51.
- Braun, H., Bejar, I.I., Williamson, D.M., 2006. Rule-based methods for automated scoring: application in licensing assessment. In: Williamson, D.M., Mislevy, R.J., Bejar, I.I. (Eds.), *Automated Scoring for Complex Tasks in Computer-Based Testing*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 123–167.
- Brennan, R. (Ed.), 2006. *Educational Measurement*, fourth ed. American Council on Education and Praeger, Westport, CT.
- Briggs, D.C., 2021. A history of scaling and its relationship to measurement. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement*. Routledge, New York, NY, pp. 263–291.
- Brown, W., 1910. Some experimental results in the correlation of mental abilities. *Br. J. Psychol.* 3, 296–322.
- Burt, C., 1936. The analysis of examination marks. In: Hartog, P., Rhodes, E.C. (Eds.), *The Marks of Examiners*. The Macmillan Co, London, UK, pp. 245–314.
- Burt, C., 1955. Test reliability estimated by analysis of variance. *Br. J. Stat. Psychol.* 8, 103–118.
- Clauser, B.E., 2021. A history of classical test theory. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement: Key Advancements in Theory, Policy, and Practice*. Routledge, New York, NY, pp. 157–180.
- Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.
- Cronbach, L.J., 1971. Test validation. In: Thorndike, R.L. (Ed.), *Educational Measurement*, second ed. American Council on Education, Washington, DC, pp. 443–507.
- Cronbach, L.J., 1975. Five decades of public controversy over mental testing. *Am. Psychol.* 30, 1–14.
- Cronbach, L.J., 1988. Five perspectives on the validity argument. In: Wainer, H., Braun, H. (Eds.), *Test Validity*. Lawrence Erlbaum, Hillsdale, NJ, pp. 3–17.
- Cronbach, L.J., 2004. My current thoughts on coefficient alpha and successor procedures. *Educ. Psychol. Meas.* 64, 391–418.
- Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N., 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. John Wiley, New York, NY.
- Cureton, E.E., 1951. Validity. In: Lindquist, E.F. (Ed.), *Educational Measurement*. American Council on Education, Washington, DC, pp. 621–694.
- Darwin, C., 1859. *On the Origin of Species*. John Murray, London.
- Dorans, N.J., Cook, L.L. (Eds.), 2016. *Fairness in Educational Assessment and Measurement*. Routledge, New York, NY.
- DuBois, P., 1970. *A History of Psychological Testing*. Allyn and Bacon Inc, Boston, MA.
- Edgeworth, F.Y., 1888. The statistics of examinations. *J. Roy. Stat. Soc.* 51, 346–368.
- Edgeworth, F.Y., 1885. On methods of ascertain variations in the rate of births, deaths and marriages. *J. Roy. Stat. Soc.* 48, 628–649.
- Ferrara, S., Lai, E., Reilly, A., Nichols, P.D., 2017. Principled approaches to assessment design, development, and implementation. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley Blackwell, New York, pp. 41–74.
- Fisher, R.A., 1925. *Statistical Methods for Research Workers*. Oliver & Boyd, Edinburgh, UK.
- Galton, F., 1853. *Tropical South Africa*. John Murray, London.
- Galton, F., 1869. *Hereditary Genius*. Macmillan, London.
- Galton, F., 1872. Statistical inquiries into the efficacy of prayer. *Fortn. Rev.* 12, 125–135.
- Galton, F., 1877. Typical laws of heredity. *Proc. R. Soc.* 8, 282–301.
- Galton, F., 1889. *Natural Inheritance*. Macmillan, London.
- Galton, F., 1892. *Finger Prints*. Macmillan, London.
- Galton, F., 1893. *Decipherment of Blurred Finger Prints*. Macmillan, London.
- Galton, F., 1895. *Finger Print Directories*. Macmillan, London.
- Galton, F., 1908. *Memories of My Life*. Methuen, London.

- Galton, F., 1909. *Essays in Eugenics*. Eugenics Education Society, London.
- Gillham, N.W., 2001. *A Life of Sir Francis Galton: From African Exploration to the Birth of Eugenics*. Oxford University Press, New York, NY.
- Goddard, H.H., 1912. *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness*. The MacMillan Company, New York, NY.
- Gulliksen, H., 1950. *Theory of Mental Tests*. Wiley, New York.
- Hart, B., Spearman, C., 1912. General ability, its existence and nature. *Br. J. Psychol.* 5, 51–84.
- Herrnstein, R.J., Murray, C., 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Simon and Schuster, New York, NY.
- Holzinger, K.J., Clayton, B., 1925. Further experiments in the application of Spearman's prophecy formula. *J. Educ. Psychol.* 16, 289–299.
- Jensen, A.R., 1969. How much can we boost IQ and scholastic achievement? *Harv. Educ. Rev.* 39 (1), 1–123.
- Kane, M., 2006. Validation. In: Brennan, R. (Ed.), *Educational Measurement*, fourth ed. American Council on Education and Praeger, Westport, CT, pp. 17–64.
- Kane, M., 2013. Validating the interpretations and uses of assessment scores. *J. Educ. Meas.* 50, 1–73.
- Kane, M., Bridgeman, B., 2021. The evolution of the concepts of validity. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement*. Routledge, New York, NY, pp. 181–205.
- Kelley, T.L., 1924. Note on the reliability of a test: a reply to Dr. Crum's criticism. *J. Educ. Psychol.* 15, 193–204.
- Kelley, T.L., 1925. The applicability of the Spearman-Brown formula for the measurement of reliability. *J. Educ. Psychol.* 16, 300–303.
- Kelley, T.L., 1927. *Interpretation of Educational Measurements*. World Book, Yonkers-on-Hudson, N.Y.
- Kuder, G.F., Richardson, M.W., 1937. The theory of estimation of reliability. *Psychometrika* 2, 151–160.
- Larmer, B., 2014. Inside a Chinese test prep factory. *N. Y. Times*. <https://www.nytimes.com/2015/01/04/magazine/inside-a-chinese-test-prep-factory.html>.
- Lazarsfeld, P.F., 1950. The logical and mathematical foundation of latent structure analysis. In: Stouffer, S.A., et al. (Eds.), *Measurement and Prediction*, Vol. 4 of *Studies in Social Psychology in World War II*. Princeton University Press, Princeton, NJ, pp. 362–412.
- Lindquist, E.F. (Ed.), 1951. *Educational Measurement*. American Council on Education, Washington, DC.
- Linn, R.L. (Ed.), 1989. *Educational Measurement*, third ed. American Council on Education and Macmillan, New York.
- Lord, F.M., 1952. *A Theory of Test Scores* (Psychometric Monograph No. 7). Psychometric Corporation, Richmond: Psychometric Corporation.
- Lord, F.M., 1953. An application of confidence intervals and maximum likelihood to the estimation of an examinee's ability. *Psychometrika* 18, 57–75.
- Lord, F.M., 1980. Applications of Item Response Theory to Practical Testing Problems. Erlbaum, Mahwah, NJ.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley Publishing Company, Reading, MA.
- Lovie, S., Lovie, P., 2010. Commentary: Charles Spearman and correlation: a commentary on "the proof and measurement of association between two things". *Int. J. Epidemiol.* 39, 1151–1153.
- Luecht, R.M., 2012. An introduction to assessment engineering for automatic item generation. In: Gierl, M., Haladyna, T. (Eds.), *Automatic Item Generation*. Taylor-Francis/Routledge, New York, pp. 59–101.
- Luecht, R.M., Hambleton, R.K., 2021. Item response theory: a historical perspective and brief introduction to applications. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement: Key Advancements in Theory, Policy, and Practice*. Routledge, New York, NY, pp. 232–262.
- Margolis, M.J., Clauser, B.E., 2020. Automated scoring in medical licensing. In: Yan, D., Rupp, A.A., Foltz, P. (Eds.), *Handbook of Automated Scoring: Theory into Practice*. Taylor and Francis, New York, NY, pp. 445–467.
- Margolis, M.J., Feinberg, R.A. (Eds.), 2020. *Integrating Timing Considerations to Improve Standardized Testing Practices*. Routledge, Abingdon.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.
- Mendel, J.G., 1866. Versuche über Pflanzeng-hybriden. Verhandlungen des naturforschenden Ver-eines in Brünn, Bd. IV für das Jahr 1865. Abhand-lungen 3–47.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education and Macmillan, New York, pp. 13–103.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., Lucas, J.F., 2006. Concepts, terminology, and basic models of evidence centered design. In: Williamson, D.M., Mislevy, R.J., Bejar, I.I. (Eds.), *Automated Scoring for Complex Tasks in Computer-Based Testing*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 123–167.
- Morant, G.M., Welch, B.L., 1939. *A Bibliography of the Statistical and Other Writings of Karl Pearson*. Issued by the Biometrika Office, University College, London. Cambridge University Press, Cambridge, UK.
- Neale, C.M., 1907. *The Senior Wranglers of Cambridge University From 1748 to 1907*. F. T. Groom & Son, Bury St Edmunds.
- Oranje, A., Gorin, J., Jia, Y., Kerr, D., Ercikan, K., Pellegrino, J.W., 2017. Collecting, analysing, and interpreting response time, eye tracking and log data. In: Ercikan, K., Pellegrino (Eds.), *Validation of Score Meaning for the Next Generation of Assessments*. Routledge, New York, NY, pp. 39–51.
- Page, E.B., 1967. The imminence of grading essays by computer. *Phi Delta Kappan* 47, 238–243.
- Pearson, K., 1896. Mathematical contributions to the theory of evolution, III: regression, heredity and panmixia. *Phil. Trans. Roy. Soc. Lond.* 187, 253–318.
- Pearson, K., 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Magazine* 50, 157–175.
- Pearson, K., 1904. On the laws of inheritance in man. II. On the inheritance of mental and moral characters in man, and its comparison with the inheritance of physical characters. *Biometrika* 3, 131–190.
- Pearson, K., 1934. Professor Karl Pearson's reply. In: *Speeches Delivered at a Dinner Held in University College, London in Honor of Professor Karl Pearson 23 April 1934*. Privately Printed at the University Press, Cambridge, pp. 19–24.
- Pearson, K., Moul, M., 1927. The mathematics of intelligence. I. The sampling errors in the theory of a generalized factor. *Biometrika* 19, 246–291.
- Rasch, G., 1960, 1980. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research, Copenhagen. Republished in 1980 with Foreword and Afterword by B. Wright, Chicago, IL: MESA Press/The University of Chicago.
- Remmers, H.H., Shock, N.W., Kelly, E.L., 1927. An empirical study of the validity of the Spearman-Brown formula as applied to the Purdue rating scale. *J. Educ. Psychol.* 18, 187–195.
- Ruch, G.M., Ackerson, L., Jackson, J.D., 1926. An empirical study of the Spearman-Brown formula as applied to educational test material. *J. Educ. Psychol.* 17, 309–313.
- Samejima, F., 1969. Estimation of Latent Trait Ability Using a Response Pattern of Graded Scores. *Psychometrika Monograph Supplement*. No. 17.
- Samejima, F., 1972. A General Model for Free-Response Data. *Psychometrika. Monograph Supplement*, No. 18.
- Sokal, M.M. (Ed.), 1990. *Psychological Testing and American Society, 1890–1930*. Rutgers University Press, New Brunswick, NJ.
- Spearman, C.E., 1904a. The proofs and measurement of association between two things. *Am. J. Psychol.* 15, 72–101.
- Spearman, C.E., 1904b. "General intelligence" objectively determined and measured. *Am. J. Psychol.* 15, 201–293.
- Spearman, C., 1907. Demonstration of formulae for true measurement of correlation. *Am. J. Psychol.* 18, 161–169.
- Spearman, C., 1910. Correlation calculated from faulty data. *Br. J. Psychol.* 3, 271–295.
- Spearman, C., 1930. *Autobiography*. In: Murchison, C. (Ed.), *A History of Psychology in Autobiography*, vol. 1. Clark University Press, Worcester, MA, pp. 299–333.
- Stigler, S.M., 1986. *The History of Statistics: The Measurement of Uncertainty before 1900*. The Belknap Press of Harvard University Press, Cambridge, MA.
- Thorndike, R.L. (Ed.), 1971. *Educational Measurement*, second ed. American Council on Education, Washington, DC.
- Thurstone, L.L., 1925. A method of scaling psychological and educational tests. *J. Educ. Psychol.* 16, 433–451.
- Thurstone, L.L., 1931. *The Reliability and Validity of Tests*. Edwards Brothers, Inc, Ann Arbor, MI.
- Tinker, M.A., 1932. Wundt's doctorate students and their theses 1875–1920. *Am. J. Psychol.* 44, 630–637.

- van der Linden, W.J., Hambleton, R.K. (Eds.), 1997. Handbook of Modern Item Response Theory. Springer, New York.
- Wissler, C., 1901. The correlation of mental and physical tests. *Psychol. Rev.*: Monograph Supplements 3 (6), i–62.
- Wolf, T.A., 1973. Alfred Binet. University of Chicago Press, Chicago, IL.
- Wood, B.D., 1926. Studies in achievement tests. Part III. Spearman-Brown reliability predictions. *J. Educ. Psychol.* 17, 263–269.
- Yoakum, C.S., Yerkes, R.M. (Eds.), 1920. *Mental Tests in the American Army*. Sidgwick & Jackson Limited, London, UK.
- Zenderland, L., 1998. *Measuring Minds: Henry Herbert Goddard and the Origins of American Intelligence Testing*. Cambridge University Press, Cambridge, UK.

Educational testing: a review of standards, guidelines, and practices

Chris Domaleski, National Center for the Improvement of Educational Assessment, Dover, NH, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	15
Sources	15
Early attempts	15
Contemporary sources	16
Academic and professional organizations	16
Testing organizations and sponsors	16
Government agencies	17
Principles	18
Foundations	18
Operations	18
Applications	19
Issues and future directions	19
References	20

Introduction

Given the central role that testing plays in education, it is important to understand the applicable professional standards that guide responsible and ethical practice. Testing is a broad term and is used to describe many formal and informal practices in service to various purposes and uses, such as informing classroom instruction or as a factor in college admissions. The focus of this work is primarily restricted to large-scale, often standardized achievement tests, especially those used for consequential purposes, such as tests used to inform school accountability. This limited focus is not meant to diminish the full depth and breadth of testing in particular or assessment more broadly. Rather, it is a recognition that the literature prominently emphasizes these applications, which is understandable given that more consequential use demands more attention to standards and ethics.

This article begins with a review of the various sources for professional guidelines and standards, starting with early attempts before exploring more recent initiatives. This section attempts to address both the origins of professional standards and some of the mechanisms that influence practice. Next, the principles or core areas addressed by contemporary standards are explored, including foundations, operations, and applications. The final section addresses issues and future directions, focusing on both critiques of current practice and ideas to promote improvement.

Sources

Early attempts

Initial attempts to establish and document standards and practices for large scale testing in education can be traced to the early 20th century. As described in [Madaus \(2001\)](#) early efforts are associated with Giles Ruch followed by Oscar Buros. In 1925 Giles Ruch lamented the lack of information to guide selection and use of tests and he called for a set of “working criteria for test construction” to be established as well as the adoption of “some fairly uniform practices in reporting on tests” (p. 349). In the 1920s and 1930s Ruch produced numerous publications proposing various technical standards for testing and advocated for a “consumers’ research bureau” on tests ([Madaus, 2001](#)).

A more widely-known and lasting attempt was initiated by Oscar K. Buros which ultimately led to the *Mental Measurement Yearbook* (MMY) starting in 1938. This enterprise grew into what is now the Buros Center for Testing, which to date has produced twenty-one editions of the MMY in addition to the Tests in Print series ([Buros Center for Testing, 2021](#)). From the start, Buros’ objective was to promote improved standards for practice through transparency. He writes in the first edition, “Up to the present time, test authors and publishers have been able to place on the market inadequately validated tests without fear of criticism. Test users have every right to demand that test authors and publishers present full particulars concerning the methods used in constructing and validating the tests” ([Buros, 1938](#), pp. xiii–xiv).

In subsequent years, the MMY served as a form of de facto peer review for test publishers. The process, as described in [Carlson and Geisinger \(2012\)](#), relies on the selection of independent and expert reviewers who examine documentation focused on the development process and evidence in support of reliability and validity with respect to the instrument’s stated uses.

The first initiative to develop standards that had the backing of an influential professional association occurred in 1954. In this year the American Psychological Association (APA) published *Technical Recommendations for Psychological Tests and Diagnostic Techniques*. This was followed quickly by a joint effort of the American Educational Research Association (AERA) and the National

Council on Measurement Used in Education (NCMUE), who produced *Technical Recommendations for Achievement Tests*, published by the National Education Association (NEA) in 1955. Similar to the initiatives that preceded it, these early statements continued to focus on the development process and emphasized the type of information that should be made available for inspection by test users (Plake and Wise, 2014).

Contemporary sources

Academic and professional organizations

Academic and professional organizations have historically made the most prominent contributions to standards and practices, which continues today. Standards refer to explicit statements, guidelines, or codes published by organizations generally regarded to have influence or oversight on professional practices related to testing.

Perhaps the most widely recognized statement of professional standards is termed the *Standards for Educational and Psychological Testing*. These standards were originally produced by a joint committee representing AERA, APA, and the National Council on Measurement in Education (NCME) in 1966. Four subsequent versions were released since then, the most recent of these in 2014. It would be difficult to overstate the scope and reach of this document, which is commonly (and hereafter) referenced as simply: *the Standards*.

The Standards represent an important step forward in establishing guidelines for professional practice not solely because of the influence of the joint sponsors. Perhaps more importantly, contemporary editions of *the Standards* represented the first substantial initiative to go beyond calls for transparency in technical documentation and address the ethics of testing practices. As stated in the introduction of the 2014 edition, “the purpose of the standards is to provide criteria for the development and evaluation of tests and testing practices and to provide guidelines for assessing the validity of interpretations of test score for the intended test uses (American Educational Research Association et al., 2014, p. 1).

Given the influence of *the Standards* and recognizing that the sponsoring organizations are based in the United States, it is worth asking to what extent the standards have implications more globally. Zumbo (2014) explores this question and concludes that both the surface features and content position the document as useful across borders. He observes, “except for the translation into a language other than English, one would have to do very little to make the Standards international” (p. 32). Further, Zumbo notes that although the standards are not widely cited in academic literature overall, the frequency of citations in the US versus international journals is similar (p. 32). Additionally, the *Standards*, according to Zumbo, “play an important pedagogical role” in other countries because they cover foundational topics of common interest across borders (p. 32). Finally, their global spread may be indicative of the fact that the sponsoring organizations have many influential international members (Zumbo, 2014).

Notwithstanding the broad influence of *the Standards*, many other organizations have sponsored testing standards and guidelines. For example, the Association of Educational Assessment—Europe developed a set of standards for the European environment culminating in the publication of the *European Framework of Standards for Educational Assessment* (AEA, 2012). The framework is described as, “a practical instrument that test providers, users of assessment results and (educational) authorities can use to compare, contrast and evaluate practices in the development, administration, grading and reporting of a wide variety of assessments and assessment programs, and to review and present evidence that the methods used meet certain quality criteria” (2012, p. 11). The AEA framework emphasizes ethics, fairness, and rights of the individual and highlights the importance of validity, reliability, practicality, and impact on stakeholders (AEA, 2012).

Other initiatives fall into the category of providing more specific guidelines and recommendations for practice, often addressing a specific topic. For example, the International Test Commission (ITC) produces a number of professional guidelines addressing various testing practices such as: *The ITC Guidelines for Translating and Adapting Tests* (2017) or *The ITC Guidelines on Computer Based Testing* (2005). These and other documents produced by the ITC provide guidance and suggestions for practice, many of which build on and extend the work of *the Standards* among other sources.

In some cases industry standards and practices are more focused on implementation than development. For example, the Council of Chief State School Officers (CCSSO) and the Association of Test Publishers (ATP) collaborated in 2013 to produce the *Operational Best Practices for Large-Scale Assessment Programs*. As the name implies, these standards provide guidelines to “strengthen implementation” of assessment programs, in particular those developed to meet federal testing requirements in the United States (2013). Accordingly, the Operational Best Practices go into more detail than *the Standards* on issues such as test security, administration, scoring, reporting, and data management.

Testing organizations and sponsors

Another source for standards and guidelines related to educational measurement comes from organizations that develop or sponsor the tests. Some of these organizations are government agencies, others may be non-profit or commercial testing companies. Regardless, it is considered best practice for organizations to develop or adopt professional practices that are consistent with industry standards. Moreover, organizations should make technical information publicly available and work with experts to monitor, review and advise on technical and operational practices.

The monitoring and review functions are often fulfilled by expert technical advisors. In the best case these groups, which go by different naming conventions such as Technical Advisory Committees (TAC), comprise experts from various disciplines (e.g., psychometrics, accessibility, content development) who are external to and independent of the test developer or sponsor. Domaleski (2017) suggests TAC members should fulfill at least three roles: auditor, problem solver, and strategic partner. The auditor role

refers to the task of reviewing information or program activities, such as plans for a standard setting workshop or results from equating multiple forms, to ensure they are in keeping with established professional practices. The problem solver role refers to providing guidance with respect to a specific challenge, such as helping the client appropriately respond to the impact of a computer failure during testing. The strategic partner role places the TAC in a position to help identify new directions and innovations aligned to the organization's goals and priorities. For example, an organization may wish to transition a testing program that is based on fixed-forms to a computer adaptive format. The TAC may be asked to help develop a process including acceptance criteria to guide this initiative.

For example, the Organization for Economic Co-operation and Development (OECD), an inter-governmental organization comprising 38 member countries, sponsors the Program for International Student Assessment (PISA), a well-known international assessment of 15-year-old students in math, science and reading. OECD regularly and publicly disseminates technical reports and other documents to address standards, practices, and outcomes for the PISA program, such as the *Technical Standards and Guidelines* to, "contribute to creating an internal dataset of quality that allows for valid cross-national inferences to be made" (2018). Additionally, OECD convenes a Technical Advisory Group (TAG), "comprising independent, world-renowned experts in the field that underpin the PISA methodology, such as sampling, survey design, scaling, and analysis." The PISA TAG helps determine technical methods and procedures and provides guidance to ensure published results are useful for their intended purpose.

Similar examples could be presented to further illustrate how testing organizations or sponsors such as public agencies, companies, or collaboratives work to uphold standards through developing guidelines, disseminating public information, and enlisting the support of independent, expert advisors.

Government agencies

Although lines of distinction among groups such as professional associations, testing sponsors, and government can be unclear, government agencies are singled out in this section to describe some distinctive contributions to professional standards and guidelines, particularly related to regulation.

Heretofore, the sources and initiatives described are all voluntary and non-regulatory. There is not necessarily a formal consequence if a testing company decides to ignore widely-accepted professional standards, forgoes disseminating public information, or fails to seek counsel from independent advisors. Presumably, such behaviors would result in a loss of public confidence in the program and diminish support. However, there is no shortage of critiques about widely used assessments that raise questions about the scant evidence to support lofty claims, properties, purposes, or uses.

Enforcement, then, is largely the province of government agencies. In some cases, government agencies can fulfill that role by publishing their own laws, policies, guidelines, or requirements. For example, the United States Department of Education (ED) has produced a very influential document for state education agencies to meet the assessment requirements set forth in federal law—the Elementary and Secondary Education Act (ESEA). The document, informally referred to as simply, *Peer Review Guidance*, describes thirty elements for which states must assemble evidence to meet federal requirements (United States Department of Education, 2015). As the name implies, this evidence is evaluated by a panel of peer experts to judge whether it is sufficient. Failure to receive a favorable outcome could ultimately result in a loss of funding. While this outcome is rare, in practice, the *Peer Review Guidance* are the de facto principal set of guiding standards for American testing programs developed to meet federal law. Given the high stakes of failing to meet the standards, state agencies use this document to establish criteria for contractors that compete for lucrative testing contracts.

In other instances, government entities may be directly responsible for development and administration of assessments. Australia's National Assessment Program—Literacy and Numeracy (NAPLAN), which is overseen by the Australian Curriculum, Assessment and Reporting Authority (ACARA) provides an example of this practice. In this case, ACARA operates under a charter issued by the Australian Government's Education Council which outlines requirements for the development, implementation and monitoring of the NAPLAN. In turn, ACARA produces numerous reports to establish and communicate standards for development and implementation, such as the *Guidelines for the Development of Accessible NAPLAN Online Items* (2021). ACARA also produces annual technical reports and works with an independent advisory group, as addressed earlier in this article.

Another way that government agencies can both influence and enforce professional standards is through litigation. Accordingly, the corpus of case law from such litigation becomes yet another source of professional standards and practices. This is particularly germane when assessment is used for consequential or high-stakes purposes. O'Neill (2003) observes, "suits challenging high stakes testing programs have focused on a variety of factors, but the common denominator is fairness" (p. 634). For example, fairness may pertain to whether a test can be used as a diploma eligibility criterion or whether the outcomes are biased in favor of racial or ethnic groups. While a full review of legal precedent as pertains to testing is beyond the scope of this work, it is instructive to highlight the influence of some prominent cases.

While numerous cases in United States case law reveal that courts generally affirm the role of educational institutions to use assessments for consequential purposes, when property rights are at stake due process must be established (O'Neill, 2003). *Debra P. v. Turlington* (1981) is an oft cited case in which a Florida court determined that evidence of opportunity to learn and adequate notification must be established to warrant the use of a minimum test score as a diploma eligibility requirement. Another property right was adjudicated in *Bartlett v. New York State Board of Law Examiners* (2001); in this case the plaintiff sought extended time accommodations on a professional licensure test, the New York bar exam. The court ruled in favor of the plaintiff, which adds to the growing body of laws and public policy internationally that affirm the obligation to provide accommodations to examinees

with disabilities. For example, under the Equality Act of 2010 in the United Kingdom, education providers are required to make reasonable adjustments for students with disabilities, which includes educational assessment.

Principles

This section is intended to serve as a broad review of the core areas or principles that are often addressed from the body of contemporary professional standards and practices for testing. It is no small challenge to succinctly categorize or organize any description of the prominent areas addressed in professional standards, not least because most topics are cross-cutting. For example, an attempt to separate topics related to test validity from those related to test administration is confounded by the fact that administration is one factor that contributes to validity.

Ultimately, this section is organized using the same major sections in *the Standards*: foundations, operations, and applications. However imperfect, this may promote clarity and consistency with prevailing practice.

Foundations

Foundations refer to the canonical concepts of validity, reliability, and fairness. Validity is often regarded as the most important of the foundational concepts, in part because it is thought to subsume reliability and fairness (i.e., a test use that is not reliable or fair cannot be valid). However, as will be discussed in a later section of this article, the meaning and scope of validity continues to be a source of debate in the measurement field. In most cases, validity is typically described as the extent to which evidence supports the claim that a test measures what it is designed to measure and can be interpreted and used as intended (see e.g., McDonald, 1999; Kane, 2006). Professional standards that address validity typically emphasize the need for validity evidence to be (1) integrated across multiple areas (2) customized to the intended use(s) or inference(s) and (3) regarded as ongoing, insofar as validation is more of a process than an event. There are numerous frameworks for understanding and addressing validity that are beyond the scope of this article. However, most any credible attempt to address validity will draw on evidence connecting the stated purposes and claims of the test to evidence related to the development process, the content and structure of the test, the technical properties of the items and forms, the cognitive processes elicited, the relationship to external variables, and the manner in which results are communicated and used.

Reliability is often addressed in professional standards and generally refers to the extent to which results are stable or consistent for examinees across administrations under similar conditions. Reliability includes the concepts of precision or accuracy, often addressed by sources of evidence that may include: reliability or generalizability studies (e.g., internal consistency, test-retest, parallel forms, inter-rater consistency), reports of measurement error, and decision consistency studies. Standards typically call for these or similar sources of evidence to be produced for all relevant groups and conditions according to the intended nature and use of the test. For example, a test given on paper and computer must be equally reliable across both modes.

Like validity, the concept of fairness is both broad and often eludes a consensus definition with respect to its application to educational measurement. *The Standards* (2014) note of fairness, “fairness is a fundamental validity issue and requires attention throughout all stages of test development and use,” (p. 49). Commonly, professional standards related to fairness will address issues of inclusion, such as requirements to ensure that tests are developed, designed, administered, and scored in a manner to maximize accessibility. Moreover, the process should work to minimize barriers for all examinees such as by applying the principles of universal design in item and form development and ensuring appropriate accessibility features and accommodations are available (e.g., for students with disabilities or language learners). The concept of fairness also includes procedures to screen items and forms for differential performance by examinee group using both empirical and judgment-based procedures. Differential Item Functioning (DIF) studies are a common source of evidence to evaluate items for disparate impact among examinee groups and item bias and sensitivity reviews are typically employed to detect potential sources of bias.

More broadly, the issue of fairness also includes responsibilities of the test developer or sponsor to safeguard against uses that are inappropriate or potentially harmful. For example, test descriptions, documentation, and score reports should be accompanied by clear statements that indicate intended purposes, uses, and interpretations and, to the extent practicable, anticipate and caution against those that are not supported.

Operations

Operations refers generally to the wide-range of activities or steps involved in developing, administering, scoring, and reporting tests. A review of common sources for standards and practices consistently uncovers two dominant principles that emerge in this area. First, test developers have a responsibility to follow established, research-based practices. Second, developers have a responsibility to make supporting evidence available for review and scrutiny by experts and stakeholders.

The first principle speaks to the need to develop a clear rationale for the selected procedures and ensure that the methods in support of those procedures are grounded in a research-base such as from academic literature or established professional practices. For example, standards that pertain to scoring open response items address steps widely-supported in research and deployed in practice such as developing clear scoring rubrics and protocols and providing scorer training to include qualifying and ongoing monitoring. As another example, when judgment-based cut-scores are established for assessments, relevant standards address

the importance of selecting qualified participants, providing feedback throughout the process, and quantifying variation in judgments. Particularly when tests are used for high-stakes or consequential purposes, it is ill-advised to draw on new, unverified methods and procedures.

The second principle consistently found in a review of operational standards is the need for clear documentation and review throughout the cycle of test development and implementation. In fact, nearly every standard statement in *the Standards* is framed as a call to provide description, information, evidence, etc. in support of a particular decision, attribute, or claim. This includes nearly every operational task from developing the test blueprint to determining specifications for field testing to defending methods for score equating. In fact, numerous standards across multiple sources call for cross-validation, review, and transparency. As noted previously, strong practices for test developers or sponsors call for not only widely disseminating information publicly, but also engaging independent experts to advise on program decisions and conduct ongoing monitoring.

Applications

Heretofore, most of the standards and practices discussed have centered responsibly on test developers or sponsors. However, it is also important to recognize the ethical principles and practices primarily, but not exclusively, under the purview of test users. In education, test users may refer to such personnel as district or school administrators, counselors, or classroom teachers. The principles related to applications cut across many areas previously described focusing on responsibilities related to being appropriately credentialed and trained to fulfill expectations, complying with relevant standards and policies (e.g., related to security and confidentiality), and interpreting and using results in a responsible manner consistent with the intended purpose of the test. For example, responsible test use requires being informed about the error associated with test scores and factoring this information into decisions, such as by corroborating scores with other sources to inform consequential decisions.

Issues and future directions

There is no shortage of opinions about the current state of professional standards for educational measurement. This section acknowledges some of the prominent critiques, related to both content and process, and identifies some emerging focus areas in the short and longer term.

Starting with the content, one criticism of the standards is that the manner in which construct representation is addressed impedes progress toward more culturally sustaining assessment practices. Jennifer Randall argues for redefining construct definition and representation from an antiracist perspective contending that current methods, “provide the illusion of fairness, but in reality, these practices serve only to support the current white supremacist hegemony” (2021, p. 4). Randall urges the field to rethink so-called “color-neutral” approaches that are reified in the standards.

This critique relates to broader and growing area of concern about the extent which prevailing standards do, or should, address the wider range of ethical or social consequences associated with test use. Newton and Shaw (2015) describe this tension as it manifests in widely different perspectives in the field about the meaning of validity. The authors contrast more conservative positions that are primarily focused on validity from a narrower technical perspective to more liberal views that may address “overall defensibility, or acceptability, of the policy which sanctions the use of test scores for a particular purpose” (2015, p. 4). The latter view has given rise to the term “consequential validity,” the meaning of which is widely debated, but it is often used to describe the necessity of evaluating the extent to which any benefits of testing outweigh the negative consequences.

This lack of clarity, illustrated by the competing views about validity, is not simply academic. It gives rise to very real and practical questions about the scope of professional standards and ethics. As society increasingly grapples with the role of tests in educational policy and practice, questions about how professional standards can or should address these uses and who bears responsibility require an answer. For example, results from achievement tests are increasingly used as prominent components of high-stakes school accountability systems. Should professional standards address these uses in part or whole? If such a use falls outside that which is supported by the test developer, the assumption is that the responsibility falls to the organization that adopts or implements the policy. Importantly, *the Standards* caution, “If a test score is interpreted for a given use in a way that has not been validated, it is incumbent on the user to justify the new interpretation for that use, providing rationale and collecting new evidence, if necessary” (p. 24). Additionally, NCME released a position statement in 2020 calling for program sponsors, designers, or users to take responsibility for developing theories of actions for consequential test uses designed to effect change. Such theories of action should “document and disseminate arguments and evidence in support of the expected (causal) relationships among the program’s constituent parts, the implementation actions, and the intended outcomes, and also plan for ongoing evaluation” (2020, p. 1).

However, it is reasonable to ask if professional associations or others should do more to support such endeavors. As Sinharay et al. (2014) observe about *the Standards*, “In our opinion, the Standards are useful to a certain extent in guiding researchers in our field with respect to general principles, but lack the type of specific guidance that practitioners often need. The Standards are thus more academic than pragmatic” (p. 36). This critique suggests that priorities for future contributions to professional standards should provide more specific guidance about how to assemble a coherent and convincing collection of evidence to support prominent uses.

Consider, too, that contemporary professional standards rely almost exclusively on professional judgment to judge the adequacy of evidence. Questions such as what constitutes adequate alignment or what level of rater agreement is acceptable are largely left to

the province of either convention or judgment. While it is unreasonable to expect sufficient clarity from professional standards and guidelines to direct all decisions related to development, implementation, and use, the field will likely benefit from ongoing efforts to create and disseminate more guidance about both cautions and promising practices. This is likely to be especially beneficial when directed to both test developers and a range of users, such as educators, leaders and policy makers.

References

- American Educational Research Association, Committee on Test Standards, National Council on Measurements Used in Education, 1955. Technical Recommendations for Achievement Tests. American Educational Research Association, Washington, DC.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- American Psychological Association, 1954. Technical recommendations for psychological tests and diagnostic techniques. *Psychol. Bull.* 51, 201–238.
- Association for Educational Assessment—Europe, 2012. European Framework of Standards for Educational Assessment 1.0. Nuova Cultura, Italy.
- Australian Curriculum, Assessment and Reporting Authority, 2021. Guidelines for the Development of Accessible NAPLAN Online Items. Available at: <https://www.nap.edu.au/about/test-development>.
- Bartlett v. New York State Board of Law Examiners, 93 Civ. 4986 (SS) (S.D.N.Y. 2001).
- Buros Center for Testing, 2021. History of the Buros Center for Testing. Available at: <https://buros.org/history>.
- Buros, O.K., 1938. The Nineteen Thirty-Eight Mental Measurements Yearbook. Rutgers University Press, New Brunswick, New Jersey.
- Carlson, J.F., Geisinger, K.F., 2012. Test reviewing at the Buros center for testing. *Int. J. Test.* 12, 122–135.
- Council of Chief State School Officers and the Association of Test Publishers, 2013. Operational Best Practices for Statewide Large-Scale Assessment Programs. Washington, DC.
- Debra P. v. Turlington, 474 F. Supp. 244 (M.D. Fla. 1979), 644 F. 2d 397 (5th Cir. 1981).
- Domaleski, C., 2017. Perspective from a TAC member. In: *Walking a Tightrope: Navigating the Balance of Policy and Psychometrics*. Presentation at the Annual Meeting of the National Council on Measurement in Education, San Antonio, TX.
- International Test Commission, 2005. International Guidelines on Computer-Based and Internet Delivered Testing. Available at: www.intestcom.org.
- International Test Commission, 2017. The ITC Guidelines for Translating and Adapting Tests, second ed. Available at: www.InTestCom.org.
- Kane, M.T., 2006. Validation. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education, Westport, CT, pp. 17–64.
- Madaus, G., 2001. A Brief History of Attempts to Monitor Testing. National Board on Educational Testing and Public Policy, vol. 2. NBETPP Statements. Number 2. Available at: <https://www.bc.edu/research/nbetpp/publications/v2n2.html>.
- McDonald, R.P., 1999. *Test Theory: A Unified Treatment*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Newton, P.E., Shaw, D.S., 2015. Disagreement over the best way to use the word “validity” and options for reaching consensus. *Assess Educ. Princ. Pol. Pract.* <https://doi.org/10.1080/0969594X.2015.1037241>.
- O'Neill, P.T., 2003. High stakes testing law and litigation. *BYU Educ. L. J.* 2003, 623.
- PISA, 2018. PISA 2018 Technical Standards. Available at: <https://www.oecd.org/pisa/pisaproducts/PISA-2018-Technical-Standards.pdf>.
- Plake, B.S., Wise, L.L., 2014. What is the role and importance of the revised AERA, APA, NCME Standards for Educational and Psychological Testing? *Educ. Meas.* 33, 4–12.
- Randall, J., 2021. “Color-Neutral” is not a thing: redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educ. Meas.* <https://doi.org/10.1111/emip.12429>.
- Ruch, G.M., 1925. Minimum essentials in reporting data on standard tests. *J. Educ. Res.* 12, 349–358.
- Sinharay, S., Rijmen, F., Choi, S., 2014. The revised standards and its role in research on educational measurement. *Educ. Meas.* 33, 36–38.
- United States Department of Education, 2015. Non-regulatory Guidance for States for Meeting Requirements of the Elementary and Secondary Education Act of 1965. Washington, DC.
- Zumbo, B.D., 2014. What role does and should the test standards play outside the United States of America? *Educ. Meas.* 33, 31–33.

Reliability

Klaas Sijtsma^a and Julius M. Pfadt^b, ^a Tilburg University, Department of Methodology and Statistics TSB, Tilburg, the Netherlands; and ^b University of Ulm, Department of Psychological Research Methods, Ulm, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	21
Classical test theory	22
Test design	22
Model	22
Methods	23
Real-data example	24
Measurement precision	25
Comments	25
Generalizability theory	26
Test design	26
Model	26
Methods	27
Real-data example	28
Measurement precision	28
Comments	28
Factor analysis	28
Test design	28
Model	28
Methods	29
Real-data example, continued	31
Measurement precision	31
Comments	31
Item response theory	31
Test design	31
Model	31
Methods	32
Real-data example, continued	32
Measurement precision	32
Comments	33
Conclusion	33
References	33

Introduction

Reliability is a key concept in the measurement of psychological attributes, such as the central intelligence ability of analytical reasoning and the personality trait of neuroticism, educational attributes, such as arithmetic ability and knowledge of American history, and health attributes, such as quality of life and pain experience. Reliability refers to the degree to which measurement values obtained from repetitions of the measurement procedure using the same test or questionnaire in the same population are equal. Different approaches to reliability exist that depart from slightly different definitions and develop different formal estimation procedures. Classical test theory (CTT) provided the traditional approach that focuses on single tests administered once in one population. Generalizability theory (GT) is a broader approach that identifies variance sources other than the attribute of interest that affect the relative or absolute measurement of persons and enables adapting the test procedure to reduce these sources while increasing reliability. Factor analysis (FA) takes a factor model as a point of departure, often a common-factor model, and defines reliability based on person variance with respect to the common or other factors. Item response theory (IRT) has also proposed reliability approaches, but in IRT the focus is on the information function conditional on the scale of measurement.

Our examples discussed next refer to the repeatability of an individual diagnosis and an individual test score, mainly because this optimally allows explaining the idea behind reliability. However, when referring to an individual result, the correct terminology is measurement precision, whereas reliability refers to a group characteristic (Mellenbergh, 1996). Measurement precision expresses the statistical precision of estimating an individual's measurement value and varies across the scale of measurement. Using a particular arithmetic test, John's test score may be estimated with less precision than Mary's. The reason for this apparent discrepancy is that the mean arithmetic level the test represents aligns better with Mary's level than with John's. For example, if the test is difficult

and Mary has a high arithmetic ability, the estimation of ability is relatively precise, whereas the estimation of John's ability, which is well below the test's level is less precise than Mary's. Reliability is a group characteristic summarizing individuals' differences in measurement precision at a higher aggregation level. In each approach, reliability expresses the ratio of the variance in a group relevant to person measurement and this variance plus sources considered irrelevant for person measurement. For example, in the FA approach, reliability can be the proportion of variance the common factor explains in the variance the total factor model explains. Because reliability is an aggregated quantity, a reliable test or questionnaire can still measure some individuals' performance with little precision and an unreliable instrument can be precise for some other individuals.

In this entry, for the CTT, GT, FA and IRT approaches to reliability, we discuss the test design, the model for reliability, methods for estimating reliability, and measurement precision aimed at the measurement of individuals. We point out the approaches' similarities and differences. We end with a typology of the four approaches emphasizing their differences but note that each has its specific merits and applications.

Classical test theory

Test design

Reliability defined by CTT expresses the degree to which a measurement value is repeatable under the same measurement conditions in a population of people. Classical reliability aligns very well with the classical question in statistics: What would happen when I do it over again? While in statistics the question refers to drawing a new random sample and assessing the degree to which statistical results are different from initial results, it is also well applicable in daily life (e.g., Lord and Novick, 1968, pp. 29–30). For example, consulting your family doctor about a resistant cough, you expect her to come up with the same diagnosis irrespective of the occasion of your visit. If she would diagnose your cough as a mild cold on Tuesday and a serious bronchitis on Wednesday, simply having forgotten that she saw you a day earlier—admittedly, not a likely assumption but needed to mimic statistical independence—you would probably feel uneasy about her trustworthiness or reliability. Similarly, an intelligence test that would diagnose you one day a mediocre person and the next day as highly intelligent would raise some eyebrows. In both cases, reliability of measurement seems to be low. Statistically, in these cases, the predictability of the measurement upon repetition is poor, a result one wishes to avoid.

The measurement value can be anything, ranging from the score on individual items to total scores on subsets of items assessing different subattributes and on the whole item set, but also response times and assessments raters provide of educational tasks. Repeatability under the same conditions refers to controlling the administration circumstances for the test or the questionnaire, such that only random events influence performance on the test or the questionnaire. Controlling the administration circumstances is known as standardization and is key to reliable measurement. Reliability is defined on a well-defined population of people and expresses the degree of measurement repeatability in this population. Reliability of measurement for a particular test or questionnaire varies across different populations and across subgroups from one population. It is important to note that a test or questionnaire does not have one reliability but as many reliabilities as there are populations in which the test is used. A test or questionnaire that is reliable in the group of all high school grades may be unreliable in one high-school grade. A test or questionnaire administered may have a different reliability under different administration procedures. For example, it may make a difference whether the test was administered in high-stakes or low-stakes conditions. In both situations, motivation may be different, affecting test performance.

Model

Let X_j be the score on item j , and let the test or questionnaire consist of J items, so that items are indexed $j = 1, \dots, J$. Let the total score on the J items be

$$X = \sum_{j=1}^J X_j.$$

CTT splits the test score into the true score, denoted T , and a random measurement error, E , so that

$$X = T + E.$$

Assume that it is possible to administer the test repeatedly to the same group of people under precisely the same circumstances assuming that everybody's attribute level stays fixed between administrations. Then, between administrations, test scores vary only due to random measurement error. For individual i , in administration r , the model would be

$$X_{ir} = T_i + E_{ir},$$

where her true score is defined as the expected value across test scores obtained in different test administrations,

$$T_i = \mathcal{E}_r(X_{ir}).$$

In a population of people, the reliability of test score X is defined as the proportion of test-score variance, denoted σ_X^2 , that is true-score variance, denoted σ_T^2 . Reliability is denoted $\rho_{XX'}$, so that

$$\rho_{XX'} = \frac{\sigma_T^2}{\sigma_X^2}.$$

Given that reliability is a ratio of two positive numbers, it cannot have negative values. If $\sigma_T^2 = 0$, that is, if all persons have the same true score, reliability equals zero; that is, $\rho_{XX'} = 0$. If all test-score variance equals true-score variance, $\sigma_X^2 = \sigma_T^2$, then reliability equals 1; that is, $\rho_{XX'} = 1$. Because CTT assumes random measurement error covaries zero with all other variables including true scores, that is, $\sigma_{ET} = 0$, one can derive that

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2,$$

and reliability can also be written as

$$\rho_{XX'} = 1 - \frac{\sigma_E^2}{\sigma_X^2}.$$

From this equation one can derive that $\rho_{XX'} = 1$ also if measurement-error variance equals zero; that is, $\sigma_E^2 = 0$. Because from the definition of the true score, we can also derive that for individual i ,

$$\mathcal{E}_r(E_{ir}) = \mathcal{E}_r(X_{ir} - T_i) = 0,$$

across individuals we find that the mean measurement error equals

$$\mathcal{E}_i(E_i) = 0.$$

Thus, variance $\sigma_E^2 = 0$ and expectation $\mathcal{E}_i(E_i) = 0$ together imply that all measurement errors equal zero, and $X_i = T_i$, for all individuals. Therefore, reliability equal to 1 implies that the test scores are without random measurement error.

To summarize, we have that

$$0 \leq \rho_{XX'} \leq 1.$$

The definition of reliability is an equation with two unknowns. On the right-hand side, either σ_T^2 or σ_E^2 are unknown, so that we cannot compute the left-hand side, $\rho_{XX'}$. Many methods have been proposed to estimate reliability by approximating σ_T^2 or σ_E^2 using only parameters that can be estimated directly from the data. Using these methods, for real tests and questionnaires, it is common to find values of at least 0.7, but for well-constructed tests and questionnaires values greater than 0.8 or 0.9 are attainable. Reliability almost equal to 1 is possible in principle when one lets the number of items grow without limit, clearly unattainable in practical applications due to person fatigue, time constraints, and item shortage.

A result we will encounter later in this entry discussing IRT reliability, is the equation

$$\rho_{XX'} = \rho_{TX}^2,$$

where the right-hand-side is the squared correlation between test score and true score. Correlation $\rho_{TX} = \sqrt{\rho_{XX'}}$ is known as the reliability index. It provides an upper bound to the correlation test score X can have with any other variable $Y \neq T$. It also plays a theoretically important role in regression approaches to CTT.

Methods

The best-known method for estimating reliability based on only observable variables, is coefficient α (Cronbach, 1951; first proposal by Kuder and Richardson, 1937). We focus on the test score based on the individual items in the test or questionnaire. Let the covariance between the item scores X_j and X_k be denoted σ_{jk} . Based only on assumptions of CTT applied to individual item scores, it can be proven that the next inequality holds (Novick and Lewis, 1967; Ten Berge and Sočan, 2004),

$$\alpha = \frac{J}{J-1} \frac{\sum_{j \neq k} \sigma_{jk}}{\sigma_X^2} \leq \frac{\sigma_T^2}{\sigma_X^2} = \rho_{XX'}.$$

The left-hand side of the inequality is coefficient α . The inequality shows that theoretically coefficient α is always smaller than or equal to reliability $\rho_{XX'}$. Unlike the right-hand side, which contains unobservable true-score variance σ_T^2 , coefficient α depends on the inter-item covariances σ_{jk} and test-score variances σ_X^2 , parameters that can be estimated from the data.

Equality $\alpha = \rho_{XX'}$ is obtained when the items are highly similar, which is mathematically defined as essential τ -equivalence (Lord and Novick, 1968) and too strict for items in practical applications. Hence, for real tests and questionnaires, at the population level, coefficient α is a lower bound to the reliability, $\alpha < \rho_{XX'}$. For item sets that are approximately unidimensional sharing a common factor, coefficient α is nearly equal to the reliability, $\rho_{XX'}$. Estimates of coefficient α are nearly equal to the true reliability. For clearly multidimensional data, coefficient α seriously underestimates reliability. The test constructor may consider whether having a test

score that expresses several attributes simultaneously is desirable. Estimates of α based on unidimensional item subsets would be close to reliability.

Many lower bounds for reliability exist that depend only on observable statistics (Guttman, 1945; Ten Berge and Zegers, 1978). For example, coefficient λ_2 is marginally higher than coefficient α and may be used alternatively. Other coefficients are greater lower bounds than coefficients α and λ_2 , but are based on optimization of sample characteristics including sampling error, known as chance capitalization, and as a result tend to overestimate true reliability seriously. One of these lower bounds readily producing overestimates is method λ_4 (e.g., Guttman, 1945; Jackson and Agunwamba, 1977), which is based on the historically well-known split-half method, and Hunt and Bentler (2015) demonstrate how to adapt the estimation procedure such that estimates are nearly equal to the reliability. The greatest lower bound (GLB; Bentler and Woodward, 1980; Woodward and Bentler, 1978) is theoretically the lower bound closest to the true reliability, but like method λ_4 , the GLB is known to overestimate reliability especially when sample sizes are smaller than 1000 and numbers of items are at least 10. Oosterwijk et al. (2017) provide more information about lower bound methods overestimating reliability in real data.

Other methods for estimating reliability having a longer tradition in test construction tend to be less efficient than the lower-bound methods but are still in use. They all mimic the idea that reliability should inform about the degree to which the test score is repeatable in a group when administered more often to the same group under the same administration procedure. We discuss the best-known methods next.

The retest method is based on two test administrations with a time interval in between that suits the researcher's purposes. The interest is with the correlation between the two sets of test scores collected, expressing the degree to which the test performance is repeatable or stable across the chosen time interval. Obviously, in this setup, all changes that take place in the individuals between the two administrations relevant to the measured attribute exert their influence on the correlation. Examples are developmental effects, learning effects, but also memory of the items answered at the first administration when presented at the second administration. Moreover, it has been suggested that across time attributes may become factorially more complex (Embretson, 1991), thus changing the meaning of measurement. At least, the retest method provides an impression of the measurement's stability across a given time interval. The retest method is expected to produce values smaller than the reliability, $\rho_{XX'}$.

The parallel-test method circumvents many of the problems discussed with the retest method by administering in the same session two different tests for the same attribute that are interchangeable with respect to content, such that they are also interchangeable with respect to their psychometric properties. We refrain from defining this mathematical convertibility known as parallelism but note that the theoretical correlation between the two sets of test scores obtained from two parallel tests, denoted X and X' , is precisely equal to the reliability of either test version. This correlation is denoted $\rho_{XX'}$, which explains the notation we introduced previously without further ado. Because two parallel tests can be shown to have the same reliabilities, we write

$$\rho_{XX'} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_{T'}^2}{\sigma_{X'}^2}.$$

The practical problem with the method is the construction of two test versions that are parallel without the items looking like copies of one another. One of the complexities of the measurement of human attributes are the badly understood intricacies of cognitive and affective systems reacting to complex stimuli like the items from a test or questionnaire. Correlations between test scores obtained from tests intended to be parallel are expected to be smaller than reliability, $\rho_{XX'}$.

The split-half method is an efficient version of the parallel-test method, where the test constructor divides the set of items into two halves that she assumes function as parallel half-tests. The correlation between the total scores on the two halves is an estimate of the reliability of a test half the length of the intended test. The Spearman-Brown prophecy formula, assuming that the halves are parallel, can then be used to estimate the reliability of a test twice the length of half a test; that is, a test as intended having length J . Although more efficient than the parallel-test method because it uses only half the number of items, the split-half method suffers from the same problems. Most important, test halves are not precisely parallel. In addition, there are many ways to construct test halves or near test-halves, each producing different correlations. Guttman's coefficient λ_4 provides a way out of these problems and Hunt and Bentler (2015) propose an estimation procedure taking it close to the true reliability.

Revelle and Condon (2019) and Sijtsma and Van der Ark (2021, Chap. 2) provide up-to-date overviews of CTT reliability estimation methods.

Real-data example

We used the correct/incorrect (1/0) data from 24 four-choice items comprising an educational test in psychological test theory. The test was administered to 247 second-year undergraduate students in psychology at Tilburg University in 2013. The data were used previously by Van der Palm et al. (2014). We deleted one invalid case that was accidentally included in the original file ($N = 248$) and one item that was answered correctly by all students, so that $J = 23$. For 247 students answering 23 items, we found $\hat{\alpha} = 0.546$ and $\hat{\lambda}_2 = 0.568$, the latter result marginally greater, consistent with expectation. In addition, $\hat{\lambda}_4 = 0.731$, and following Hunt and Bentler (2015), the λ_4 estimates based on the 5th, the 50th, and the 95th percentile of the estimated sampling distribution equaled 0.624, 0.669, and 0.707. Hunt and Bentler (2015) suggested using the 5th percentile estimate, which they claim protects best against overestimation while being close to the true reliability. The GLB estimate was 0.747.

Coefficients α and λ_2 provide almost unbiased estimates of reliability $\rho_{XX'}$ when the item set has a one-factor structure, but the λ_4 variations and the GLB capitalize on chance and tend to overestimate the reliability. Educational tests administered at universities are constructed to cover highly heterogeneous knowledge domains rendering a multi-factor structure likely. An exploratory minimum-residuals two-factor analysis produced five items loading higher than 0.3 on the first Varimax rotated factor and six items loading at least 0.3 on the second factor, with a maximum loading of 0.49. For a three-factor solution the numbers of items were four, four, and two, with a maximum loading of 0.62. For multi-factor structures, coefficients α and λ_2 provide reliability underestimates, but the coefficient λ_4 variations and the GLB tend to overestimate reliability seriously. Given the messy factor structures, we did not attempt to estimate reliability for item subsets or based on methods taking multiple factors into account.

Measurement precision

A reliability equal to, say, 0.85, may seem quite satisfactory but represents a group characteristic, leaving the issue open whether the test or questionnaire measures individuals with enough precision. A simple but crude approach is using the standard error of measurement (SEM),

$$\sigma_E = \sigma_X \sqrt{(1 - \rho_{XX'})},$$

derived directly from the definition of reliability. Inserting sample standard deviation S_X and an estimate for $\rho_{XX'}$ provides an estimate of the SEM, denoted S_E . The equation of the SEM contains only group characteristics. Hence, similar to reliability, the SEM is a group characteristic providing the same measure of imprecision for each tested individual. In the simplest case, assume that we take the test score as an estimate of the true score, and we assume that across independent repetitions the distribution of person i 's test scores is normal and except for mean T_i varies across individuals with the same standard deviation, estimated by means of S_E . A 95% confidence interval equals

$$[\hat{T}_i - z_{(\alpha/2)} S_E; \hat{T}_i + z_{(\alpha/2)} S_E],$$

where α is the 5% significance level used to test the null hypothesis that $T_i = X_i$. Another approach known as Kelley's formula uses a regression model for estimating the true score. The model uses group characteristics to transform the test score to a true score estimate and like the SEM, the regression approach provides a residual standard error that is the same for everyone.

The weakness of the SEM approach to measurement precision is that it assigns the same measure of imprecision, σ_E (and the regression residual standard error), to everyone irrespective of whether the test or questionnaire matches the individual's attribute level. For example, an easy arithmetic test is more informative of a student having low arithmetic ability than a student with a higher level. Because the latter student gives a positive response to most items, we only know that her level is higher than the items' levels, but we need more difficult (ability) or less popular (trait, attitude) items to pinpoint the individual's scale position. [Sijtsma and Van der Ark \(2021, pp. 80–82\)](#) discussed methods for estimating scale-dependent SEMs; also see [Feldt et al. \(1985\)](#). [Lek and Van de Schoot \(2018\)](#) studied SEMs for subgroups and individuals and recommended the following. An individual-dependent SEM presumes many items enabling several "parallel" subtests with approximate continuous subtest scores, conditions rarely realistic. In other cases, a SEM conditional on subgroups may be appropriate when a relationship exists between the true score and the conditional error variance. Otherwise, the overall SEM is in order.

Comments

A CTT feature often overlooked is that CTT is an error theory but not a latent variable model that seeks to explain the relationships among a set of items. At the level of items, for items j and k , CTT equals

$$X_j = T_j + E_j \text{ and } X_k = T_k + E_k.$$

For individuals, the item true-scores are expectations across independent repetitions of the same items, but there are no model features such as latent variables that dictate the items must have something in common. To modern-day test constructors and researchers using tests and questionnaires this may seem odd, but they should realize that they are the actors who select a set of items that purport to measure the same attribute, collect data, and then use principal component analysis or factor analysis to assess whether the items measure enough in common to constitute one scale. Principally, it is only then that CTT is brought in to estimate the reliability of the test score. Of course, CTT has produced many methods to help construct a useful test or questionnaire, but as an error theory, it basically is concerned with reliability but not with what items or other test parts measure in common. FA and IRT do assume items or test parts to have one or more latent variables in common, and in this respect are different from CTT. Still, the fact that a set of items have a latent variable in common does not logically imply they measure the attribute of interest. As with CTT, it is up to the researcher to establish this as part of validity research.

The fact that modern-day psychometrics has embraced latent variables as explanatory tools underlying sets of observable variables has convinced several psychometricians they should view CTT and reliability from the perspective of latent-variable models, especially FA. Hence, the true score is considered a latent variable that the items have in common. The condition of essentially τ -equivalent items is interpreted as a 1-factor model, and because the 1-factor model is too restrictive to fit a data set, coefficient

α is criticized for not equaling true reliability. This result is the main reason why adherents of other approaches reject coefficient α and other CTT-based reliability methods in favor of other approaches mainly based on FA. We have argued that coefficient α simply does what it is supposed to do, but that it tends to underestimate reliability when data are multidimensional, a situation confounding different attributes test constructors often wish to avoid. One may consult recent discussions on reliability; see [Sijtsma \(2009\)](#), [Cho and Kim \(2015\)](#), [McNeish \(2018\)](#), [Raykov and Marcoulides \(2019\)](#), [Revelle and Condon \(2019\)](#), and [Sijtsma and Pfadt \(2021\)](#). [Mellenbergh \(1998\)](#) developed the idea of one latent variable in the context of CTT and studied test-score reliability for the estimated latent variable.

Generalizability theory

Test design

Chronologically, GT was the next approach to reliability, dating back to the 1950s. Its first full account was given by [Cronbach et al. \(1972\)](#). CTT defines and estimates reliability of test scores based on one test administered on one occasion and scored by one rater in case of constructed-response items. One may think of a set of items assessing reading ability at the beginning of the school year that is the same for each student tested, and one rater who scores the responses the students provide. One could argue that the test score is only interesting to the degree it generalizes across item sets, raters, and occasions, but other facets—sources of test-score variation—may be of interest depending on the purpose of testing. When different item sets, raters, and occasions, and possibly combinations of these facets produce different test scores, the exact conditions under which a test is administered are important for the result, but one may wish the assessment of test performance were independent of these rather arbitrary conditions and adapt the test to minimize their influences.

GT defines the sources that influence test performance in addition to people's attribute as error sources. The test constructor wishes to minimize the influence of these error sources on the test's reliability. Based on a design including the facets and the facet conditions of interest for collecting data, the generalizability (G) study estimates the variance components each of the error sources contribute to the total-score variance, and for large variance components the decision (D) study assesses how the test should be adapted to optimize reliability. Not the single test score that is of interest to CTT is of interest to the purpose of testing as adopted by GT, but the mean score taken across the levels or conditions of the different facets. This mean score is known as the universe score, which parallels CTT's true score. The question addressed in GT is how to assemble the test and the conditions for its administration so that an estimate of people's universe scores is obtained.

In the G study, GT identifies different sources of variance, in the example attributable to item sets, raters, and occasions, their two-way and three-way interactions, and a combined variance component for four-way interactions, unexplained sources of test-score variance and random measurement error. The latter component reflects the confounded influence of various sources on test performance that the designs do not allow to be disentangled. The variance components are used to estimate two kinds of reliability, one assessing the relative interpretation of test performance, by which GT means the ordering of persons relative to one another, and the other assessing the absolute interpretation of individuals' test performance independent of other people's test performance. CTT only considers random measurement error as an error source and all other variance as true-score variance and defines and estimates reliability for relative measurement. Depending on the relative or absolute interpretation of test performance, GT reliability is based on a variance decomposition different from that of CTT.

GT's contribution is to identify the various variance components based on a factorial design for collecting data and estimate a coefficient that corrects for these components depending on the kind of person assessment. The smaller the error variances, the higher the reliability and the more generalizable the test performance across the identified facets to universe scores.

Model

We limit attention to two facets considered relevant, which are item sets and raters. By relevant we mean that we suspect one fixed item set and one arbitrarily chosen rater to be insufficient for reliable measurement. Therefore, we prefer to study the degree to which different item sets and raters affect reliability and based on this information how to improve the test, for example, by using more items and raters. We discuss the GT model for two random facets, item sets and raters, in a crossed design. Let X_{ijr} be the score of person i on item j ($j = 1, \dots, J$) scored by rater r ($r = 1, \dots, R$). Persons are the objects of measurement and item sets and raters constitute potential sources of error we wish to estimate and use for improving the test.

Score X_{ijr} is decomposed into effects of facets represented in the design according to analysis-of-variance logic. This yields a person effect, an item effect, a rater effect, three two-way interaction effects, and a residual effect. The residual represents both the three-way interaction and all the effects not represented thus far; that is, effects stemming from facets relevant to the generalization not included in the design and all other systematic and random influences. Let μ be the grand population mean, defined as $\mu = \mathcal{E}_i \mathcal{E}_j \mathcal{E}_r (X_{ijr})$, let $\mu_i = \mathcal{E}_j \mathcal{E}_r (X_{ijr})$ be the mean for person i or her universe score, and so on; then

$$\text{Main effects: } X_{ijr} = \mu + (\mu_i - \mu) + (\mu_j - \mu) + (\mu_r - \mu) +$$

$$\text{Two-way interactions: } + (\mu_{ij} - \mu_i - \mu_j + \mu) + (\mu_{ir} - \mu_i - \mu_r + \mu) + (\mu_{jr} - \mu_j - \mu_r + \mu) +$$

$$\text{Residual effect : } + (X_{ijr} - \mu_{ij} - \mu_{ir} - \mu_{jr} + \mu_i + \mu_j + \mu_r - \mu).$$

We assume that each effect has a distribution. For example, the universe scores μ_i have mean equal to 0 and positive variance, reflecting person variance with respect to, for example, reading ability. The total variance of all X_{ijr} scores in the design can be decomposed in

$$\sigma_{X_{ijr}}^2 = \sigma_i^2 + \sigma_j^2 + \sigma_r^2 + \sigma_{ij}^2 + \sigma_{ir}^2 + \sigma_{jr}^2 + \sigma_{ijr,e}^2.$$

The residual variance, $\sigma_{ijr,e}^2$, represents all variation due to the three-way interaction and all unaccounted influences, systematic and random, impossible to disentangle because design cell (i, j, r) contains just one observation X_{ijr} . Variance σ_i^2 is the variance of the universe scores and shows the degree to which persons vary with respect to reading ability. Variance σ_{ij}^2 for person-by-item interaction represents the extent to which different item sets order persons differently, and so on. These variance components are the basis for estimating reliability based on a relative or an absolute interpretation of test performance. How this is done precisely depends on the design, which may be fully crossed (or incomplete), nested, (un)balanced, random/fixed or based on any combination of these possibilities, comparable to the large variety of experimental designs known from analysis-of-variance methods (e.g., [Winer, 1962](#)). Given the space admitted in an encyclopedia, we limit ourselves to the completely-crossed random two-facet design we discussed thus far.

Methods

Variance components are estimated using methods well known from analysis of variance. See [Shavelson and Webb \(1991\)](#) for a primer, and [Brennan \(2001\)](#) for an advanced treatment of GT using analysis-of-variance methods. The computations are beyond the scope of this entry, so we confine the discussion to noticing that one starts the computations with the sums-of-squares, then divide by the appropriate number of degrees of freedom to obtain the mean squares, and then derive the expected mean squares, which are functions of the unknown variance components desired. The result is a set of K equations from which K variance components can be solved. These variance estimates are the input for reliability estimation. In what follows, we assume we have the variance components available and focus on reliability. Two possibilities exist, one for relative person measurement and one for absolute person measurement, the first producing a *generalizability coefficient* and the second an *index of dependability*. We focus on the general idea and refer the interested reader to the literature. The difference between the two coefficients resides in identifying variance sources that confound the relative measurement of the ordering of persons, the absolute measurement of persons, or both.

For relative measurement, obviously σ_i^2 is relevant, but the interaction sources including persons are not and, in this context, qualify as error. That is, a positive interaction variance between persons and item sets means that different item sets order persons differently, and a positive interaction variance between persons and raters means that different raters order persons differently. A positive three-way interaction variance, included in the residual variance, $\sigma_{ijr,e}^2$, means that different combinations of item set and rater order persons differently. However, a positive interaction variance between item set and rater means that different raters order item sets differently by mean item scores, but does not necessarily influence person ordering, hence does qualify as an error source relative to person ordering. The same conclusion holds for positive item set variance and positive rater variance; these variances show only that mean item scores computed across persons and raters vary ($\sigma_j^2 > 0$) and that mean rater scores computed across persons and items vary ($\sigma_r^2 > 0$). For other designs, different variance sources constitute error and lead to different generalizability coefficients and dependability indices.

For our completely-crossed random two-facet design example, the error variance, denoted as σ_{Rel}^2 (Rel for relative), equals

$$\sigma_{\text{Rel}}^2 = \frac{\sigma_{ij}^2}{J} + \frac{\sigma_{ir}^2}{R} + \frac{\sigma_{jr}^2}{JR}.$$

The denominators reflect that larger numbers of item sets and raters reduce error sources. This knowledge may be used in the D study to make optimal decisions about the design maximizing reliability. The other ingredient for determining generalizability is person variance, σ_i^2 . Following notational conventions in GT (e.g., [Brennan, 2001](#), p. 13), generalizability for relative measurement is defined as

$$\mathcal{E}(\rho^2) = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\text{Rel}}^2} = \frac{\sigma_i^2}{\sigma_i^2 + \frac{\sigma_{ij}^2}{J} + \frac{\sigma_{ir}^2}{R} + \frac{\sigma_{jr}^2}{JR}}.$$

It is interesting to notice that the denominator is different from $\sigma_{X_{ijr}}^2$, the variance of all observed mean item scores collected in the design. [Shavelson and Webb \(1991\)](#) noticed that the denominator equals $\mathcal{E}_{ijm}(\sigma_{X_{ijm}}^2)$.

For absolute measurement, the error variance σ_{Abs}^2 (Abs for absolute) equals

$$\sigma_{\text{Abs}}^2 = \frac{\sigma_j^2}{J} + \frac{\sigma_r^2}{R} + \frac{\sigma_{ij}^2}{J} + \frac{\sigma_{ir}^2}{R} + \frac{\sigma_{jr}^2}{JR} + \frac{\sigma_{ijr,e}^2}{JR}.$$

Reliability for absolute measurement is now defined using the index of dependability and equals

$$\Phi = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_{\text{Abs}}^2} = \frac{\sigma_i^2}{\sigma_i^2 + \frac{\sigma_j^2}{J} + \frac{\sigma_r^2}{R} + \frac{\sigma_{ij}^2}{J} + \frac{\sigma_{ir}^2}{R} + \frac{\sigma_{jr}^2}{JR} + \frac{\sigma_{jre}^2}{JR}}.$$

Given that the error term in the index of dependability Φ contains more variance components than in the generalizability coefficient $\mathcal{E}\rho^2$, one will often find smaller values for Φ .

Real-data example

We did not provide GT reliability estimates for real data because we would need data collected based on a factorial design, different from the data used in the CTT section. We do not have such more complex data available, but for easily accessible data examples refer the reader to [Shavelson and Webb \(1991\)](#).

Measurement precision

[Brennan \(2001\)](#) discussed standard errors of measurement for relative and absolute measurement in several designs. He also discussed GT approaches to conditional measurement precision. These discussions are important but beyond the scope of this entry.

Comments

Conceptually, GT is more complex than CTT. Whereas CTT focuses on one measurement instrument that is administered to a particular population under well-defined conditions, GT asks what happens to measurement when each of these characteristics and others deemed relevant are varied. Thus, GT assesses reliability—generalizability and dependability in GT parlance—of estimates of universe scores obtained across the various conditions, which replace true scores. Variation in measurement caused by variation in facet conditions is considered error that confounds measurement in the sense that it reduces reliability. A G study is used to estimate the variance components in the total-score variance due to the choice of items and the choice of raters. Depending on the application of the planned measurement instrument and based on the results, new estimates are made of the variance components when the numbers of items and raters are increased. Next, a D study is used to estimate reliability for relative and absolute measurement. The outcomes determine test length and deployment of raters. Thus, the D study determines the final measurement instrument.

Facets discussed in treatments of GT are usually quantitative, such as numbers of item sets and raters, but an interesting question is how GT handles qualitative facets, such as different presentation modes for the items. Suppose a test constructor wonders whether the measurement of reading ability depends to some degree on item presentation in multiple-choice mode or constructed-response mode. When there are only two conditions for the facet of presentation mode, if it is included in the design, it is a fixed facet, which by itself does not pose a problem to analysis of variance. The problem here is that conceptually it is difficult to imagine how the test procedure should be adapted if the variance component for presentation mode is large; surely not by increasing the number of different presentation modes. [Shavelson et al. \(1989\)](#) and [Briesch et al. \(2014\)](#) suggest approaches that do justice to qualitatively different item-presentation modes in the context of GT.

Factor analysis

Test design

One administers a test or questionnaire once to a group of persons, comparable to the test design discussed for CTT. The resulting data are the input for a reliability estimation using FA methods.

Model

Unlike CTT, FA models the relationship between pairs of items j and k in a test or a questionnaire by means of one or more latent variables, in the context of FA called factors, as ([Bollen, 1989](#))

$$X_j = b_j + \sum_{m=1}^M a_{jm}\xi_m + \delta_j \text{ and } X_k = b_k + \sum_{m=1}^M a_{km}\xi_m + \delta_k.$$

The latent variables are denoted by ξ_m , with $m = 1, \dots, M$. Each item has a loading on each latent variable, denoted a_{jm} , indicating the strength of the relationship between the item and the latent variable, and an intercept b_j that represents the item's difficulty or (lack of) popularity. Residual δ_j represent unexplained systematic score components and random measurement error. In a multidimensional model, some latent variables may be unimportant for some items, in which case loading $a_{jm} \approx 0$. For example, for some arithmetic items, language skills may play a role while for other items language is immaterial for solving the problem. In a unidimensional model, only one common latent variable is of interest and $M = 1$.

Whereas researchers using CTT reliability use principal component analysis and FA to select items from a preliminary item set into the final test or questionnaire and then estimate reliability, researchers using the FA approach to reliability estimate a confirmatory factor model and a reliability measure tailored to the instrument's factor structure. The item set may be adapted based on the goodness-of-fit of the FA model to the data. Items having small loadings on each factor or cross loadings across factors are often candidates for removal, a strategy also used in CTT-based instrument construction. In FA reliability estimation, the true-score variance is modeled by latent variables that replace the true score. In this section, we discuss two factor models resulting in several reliability methods for test score X .

Methods

One-Factor Model. For one item and one latent variable, [Bollen \(1989, pp. 218–221\)](#) discussed the next model, dropping latent-variable index m ,

$$X_j = b_j + a_j\xi + \delta_j.$$

This model is also known as the model for congeneric measures ([Jöreskog, 1971](#)). Residual δ_j consists of two components, which are a unique error U_j typical of item j , and random measurement error consistent with CTT, resulting in $\delta_j = U_j + E_j$. The unique error U_j has unknown variance, $\sigma_{U_j}^2$, and thus cannot be separated from $\sigma_{E_j}^2$. Without going into further detail and consistent with the practice of FA reliability estimation, we assume $\delta_j = E_j$, implying that residual variance is random measurement error variance. The one-factor model thus equals

$$X_j = b_j + a_j\xi + E_j,$$

with CTT true score modeled as

$$T_j = b_j + a_j\xi.$$

Thus, in the FA reliability approach, the true score depends on latent variable ξ , which is a quantity the different items in the test have in common, and the item parameters b_j (intercept, difficulty) and a_j (loading on the latent variable). CTT defines the true score of an item as the expected item score across independent repetitions. The item true score depends on the item only, but different items may measure unrelated attributes or unrelated sets of attributes. This is what the model represents mathematically, but not what a test constructor will pursue. Instead, the test constructor will select items to assess the attribute of interest as well as possible, based both on theoretical and statistical analysis, but this does not affect the fact that CTT does not model the true score and estimates test-score reliability irrespective of the dimensional composition of the item set. This is the fundamental difference with the FA approach to reliability.

Based on the classical definition of reliability, we define ([Bollen, 1989, p. 220](#)) the reliability of test score X as in the one-factor model (OFM) as

$$\rho_{XX'}^{\text{OFM}} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma^2(b_j + a_j\xi)}{\sigma^2(b_j + a_j\xi) + \sigma_E^2}.$$

Under particular restrictions not further discussed, the variance $\sigma^2(b_j + a_j\xi)$ can be estimated using the squared sum of the factor loadings $(\sum_{j=1}^J a_j)^2$ and variance σ_E^2 can be estimated using the sum of the items' residual variances. The result is coefficient ω ([McDonald, 1999; Revelle and Zinbarg, 2009](#)). The right-hand side of the definition represents the one-factor model reliability, which we have equated to the CTT reliability on the left-hand side. However, equality will hold only if the one-factor model perfectly fits the data for a real test or questionnaire, a situation that will not hold in the practice of test and questionnaire construction. Notation $\rho_{XX'}^{\text{OFM}}$ represents the former situation.

We already pointed out the opinion that coefficient α can only be used when the items satisfy the mathematical requirement of essential τ -equivalence, which is then interpreted as a one-factor model. This one-factor model is considered the FA interpretation of CTT. Rather than defining reliability for test score X , [Mellenbergh \(1994, 1998\)](#) studied reliability of the estimated latent-variable score rather than the test score as an estimate of the true score.

The model for congeneric measures is the point of departure in Mellenbergh's proposal. The model allows items to vary with respect to item parameters b_j and a_j , and population error variance, thus allowing $\sigma_{E_j}^2 \neq \sigma_{E_k}^2$ ([Bollen, 1989, p. 208](#)). Mellenbergh noticed that one can only estimate the item parameters b_j and a_j when the origin and unit of latent variable ξ are defined. This is done by means of standardization, such that in the population $\mathcal{E}(\xi) = 0$ and $\sigma_\xi^2 = 1$. We skip the detailed development of the model and follow [Mellenbergh \(1994\)](#) to define a reliability coefficient in which the variance of the latent variable, σ_ξ^2 , replaces the variance of the true score, σ_T^2 , and the variance of the estimated latent-variable score, $\hat{\xi}$, conditional upon latent-variable score

ξ , $\sigma_{\xi|\xi}^2$, also known as the residual variance, replaces error variance, σ_E^2 . The conditional variance of the estimated latent-variable score equals

$$\sigma_{\xi|\xi}^2 = \left(\sum_{j=1}^J \frac{a_j^2}{\sigma_{\varepsilon_j}^2} \right)^{-1}.$$

Following the definition of CTT reliability for test score X , the reliability of the estimated latent-variable score is defined as

$$\rho_{\xi}^{\text{OFM}} = \frac{\sigma_{\xi}^2}{\sigma_{\xi}^2 + \sigma_{\xi|\xi}^2}.$$

Because $\sigma_{\xi}^2 = 1$ by definition, reliability reduces to

$$\rho_{\xi}^{\text{OFM}} = \left[1 + \sigma_{\xi|\xi}^2 \right]^{-1}.$$

These are theoretical definitions, and simple methods exist for estimating latent-variable scores ξ and variance $\sigma_{\xi|\xi}^2$, thus allowing the estimation of reliability ρ_{ξ}^{OFM} (Mellenbergh, 1994, 1998). Raykov (1997, 2001) proposed the adaptation of coefficient α to congeneric measures. What matters here, is that adopting a factor model to replace CTT produces a new model different from CTT and has consequences for the reliability definition adopted. This consequence is important but it often is not recognized.

Multi-Factor Model. As we saw already, the multi-factor model (MFM) equals

$$X_j = b_j + \sum_{m=1}^M a_{jm} \xi_m + \delta_j.$$

Clearly, the MFM opens a wide range of possible models for the item scores, and to avoid trial and error it is advisable to depart from a well-founded attribute theory, if available. This limits candidate models to only a few the theory predicts. This approach implies fixing loadings for some items on some factors to 0, $a_{jm} = 0$, when latent variable ξ_m is irrelevant for the response to item j . In the ideal case, one tests whether a set of hypothesized loadings produces a fitting model, but we do not know examples of this scenario.

A simple MFM assumes that each item loads on at least one latent variable and that at least two items load on the same latent variable, so that the model does not allow unique components. In FA reliability, one usually assumes intercept $b_j = 0$, so that the true score relates to the factor model by

$$T_j = \sum_{m=1}^M a_{jm} \xi_m.$$

The test score is written as

$$X = \sum_{j=1}^J X_j = \sum_{j=1}^J \sum_{m=1}^M a_{jm} \xi_m + \sum_{j=1}^J E_j.$$

Reliability of test score X for the MFM is defined as (Raykov and Shrout, 2002)

$$\rho_{XX'}^{\text{MFM}} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma^2 \left(\sum_j \sum_m a_{jm} \xi_m \right)}{\sigma^2 \left(\sum_j \sum_m a_{jm} \xi_m \right) + \sigma_E^2}.$$

Only if the factor model fits the data well does $\rho_{XX'}^{\text{MFM}}$ equal CTT reliability. In practice, ignoring relevant latent variables or including irrelevant latent variables, incorrectly specifying the loading structure, or a combination of these errors, leads to a misfitting factor model and a biased estimate of reliability $\rho_{XX'}$.

Revelle and Zinbarg (2009) rephrased $\rho_{XX'}^{\text{MFM}}$ as follows. Let the item communality be denoted h_j^2 . This is the proportion of variance of item score, X_j , explained by the common latent variable and the group variables, which are factors on which subsets of items load but not all items. Let M relevant latent variables be orthogonal or uncorrelated; then,

$$h_j^2 = \sum_{m=1}^M a_{jm}^2.$$

The complement of item communality h_j^2 , denoted d_j^2 , is the proportion of unexplained or residual variance, so that $d_j^2 = 1 - h_j^2$. The complement approximates the item's proportion of error variance. For standardized item scores, coefficient ω_t is defined as (McDonald, 1999; Revelle and Zinbarg, 2009)

$$\omega_t = 1 - \frac{\sum_j (1 - h_j^2)}{\sigma_X^2} = 1 - \frac{\sum_j d_j^2}{\sigma_X^2}.$$

In CTT reliability, σ_E^2 replaces numerator $\sum_j d_j^2$, but because the models underlying the two equations are different, equating the coefficient ω_t to the classical reliability definition might be misleading.

The bi-factor model is a popular multi-factor model for which we briefly discuss the FA approach to reliability. The model defines a group factor for subsets of items and a common factor for all items. Thus, residual d_j in coefficient ω_t includes the variance explained by the common factor and the group factor for item j . Another method known as coefficient ω_h quantifies the degree to which the item set measures the common factor in a bi-factor model (e.g., Flora, 2020; Zinbarg et al., 2005). Raykov and Shrout (2002) estimated coefficient ω by means of a structural equation modeling approach. Pfadt et al. (2021) discussed a Bayesian approach to estimating the ω coefficient for the one-factor model, and Pfadt and Sijtsma (in press) studied the statistical properties of several ω coefficients.

Real-data example, continued

For the same data as analyzed in the section on CTT, we found estimate $\hat{\omega}$ based on one factor equal to 0.547, almost equal to CTT estimates $\hat{\alpha}$ and $\hat{\lambda}_2$. Given the messy factor structure we identified, we refrained from estimating variations of coefficient ω for the multifactor case.

Measurement precision

FA approaches to measurement precision are rare, a circumstance that could be explained from the focus of FA approaches on research rather than the diagnosis of individual behavior. Recently, Martin and Rast (2022) proposed an FA-based approach to modeling measurement precision which they called case-specific reliability, applicable to subgroups and covariates in addition to individual scores. Their approach estimates a non-linear, nonmonotonic function for reliability conditional on the latent variable, thus providing a function analogous to the test information functions in IRT we discuss shortly.

Comments

Whereas CTT and GT focus on error sources in test performance, estimate the magnitude of their influence on reliability, and aim at reducing this influence thus increasing reliability, the FA approach to reliability focuses on the factorial composition of the test data and defines reliability as the proportion of variance the common factor explains in the data. This proportion equals CTT reliability when the FA model fits the data, and then the common-factor variance equals the true-score variance. Given that FA models fit the data by approximation at best, an FA reliability method does not estimate reliability without bias, an observation that is true for all approaches to reliability.

The FA approach to reliability models the true score, and this is often seen as an advantage of the approach. However, the finding that a factor model fits the data by approximation does not imply that the item set measures the attribute of interest, unless a well-founded theory of the attribute laid at the foundation of the test construction. CTT and GT do not model the true score. This places the establishment of validity separate from estimating reliability. Reliability approaches modeling test performance using latent variables are faced with the same task.

Item response theory

Test design

Tests and questionnaires can be fixed as in the CTT, GT, and FA context, but a special feature of IRT is that its models allow different examinees to respond to different sets of items taken from large item collections known as item banks. These examinee-tailored tests allow the optimally precise measurement of the attribute of interest. This feature makes IRT reliability and measurement precision stand out from the other three approaches.

Model

FA is a linear latent-variable model for continuous dependent variables, and IRT is a non-linear model for discrete dependent or response variables. These response variables are dichotomous (e.g., correct/incorrect) or polytomous (e.g., ordered rating-scale

scores), and are transformed to response probabilities as functions of the latent variable (the factor in FA) and latent item characteristics. Like FA, one or more latent variables representing attributes model the relationship among the items. The number of IRT models is large (e.g., [Sijtsma and Van der Ark, 2021](#); [Van der Linden, 2016](#)). Therefore, we use general notation for the functional relation, denoted f , between a response probability for an item, P , and latent variables, θ , and item properties such as difficulty (comparable to the intercept in FA), and item discrimination (comparable to the item loading), quantified by parameters in η . For two items j and k , the response functions, which are usually assumed to be increasing, are

$$P(X_j \geq x|\theta) = f(\theta; \eta_j) \text{ and } P(X_k \geq x|\theta) = f(\theta; \eta_k).$$

For scores 0/1, the response probabilities are $P(X_j = 1|\theta)$ and $P(X_k = 1|\theta)$. Given that tests are mostly constructed to measure one attribute, there is usually just one latent variable, so that $\theta = \theta$. Function f is often a normal ogive (cumulative normal) or a logistic function. Like FA, IRT provides a model for the relationship between items allegedly measuring the same attribute or attributes.

Methods

[Kim \(2012\)](#) defined two items as parallel if their response functions are equal. Several methods are available for estimating latent variable θ , resulting in estimate $\hat{\theta}_i$ for everyone taking the test. Expectation $\mathcal{E}(\hat{\theta}|\theta)$ is the expected value of estimate $\hat{\theta}$ for a particular scale value θ , comparable to the true score in CTT; hence, $\sigma^2_{\mathcal{E}(\hat{\theta}|\theta)}$ can be compared to CTT true-score variance, σ_T^2 . Variance $\sigma^2_{\hat{\theta}|\theta}$ is conditional estimation-error variance, and expectation $\mathcal{E}(\sigma^2_{\hat{\theta}|\theta})$ is comparable to error variance σ_E^2 assessed for the whole group.

Because estimate $\hat{\theta}$ is biased ([Lord, 1983](#)), one cannot simply assume all results from CTT to be valid in the IRT context. Recalling CTT equality $\rho_{XX'} = \rho_{TX}^2$, biasedness of $\hat{\theta}$ implies that a similar equality in IRT does not hold ([Kim, 2012](#)). Two parallel tests result in estimates $\hat{\theta}$ and $\hat{\theta}'$. Item parameters η_j have been estimated previously with enough precision and are assumed known. Based on [Lord \(1980, p. 52\)](#), [Kim \(2012\)](#) defines reliability of parallel measures as

$$\rho_{\hat{\theta}\hat{\theta}'} = \frac{\sigma^2_{\mathcal{E}(\hat{\theta}|\theta)}}{\sigma_{\hat{\theta}}^2} = 1 - \frac{\mathcal{E}(\sigma^2_{\hat{\theta}|\theta})}{\sigma_{\hat{\theta}}^2}.$$

In general, because of biasedness in $\hat{\theta}$, we have that $\rho_{\hat{\theta}\hat{\theta}'} \neq \rho_{\hat{\theta}}^2$. See [Holland and Hoskens \(2003\)](#) for further relationships between CTT and IRT.

Real-data example, continued

Estimates of reliability in an IRT analysis are rarely encountered in research reports because the focus in IRT is on the Fisher information function ([Sijtsma and Van der Ark, 2021](#), chap. 4). As an illustration for this entry, we obtained a CTT-like reliability for the data used in the CTT section by replacing the true-score variance with the variance of the estimated latent variable $\hat{\theta}$ from the unidimensional two-parameters logistic model and the error-score variance with the sample mean of the squared standard errors of the estimates $\hat{\theta}_i$ across the various $\hat{\theta}_i$ values ([Chalmers, 2012](#)). To maintain comparability with the CTT and FA estimates, we ignored the multidimensional data structure and refrained from a goodness-of-fit analysis common to an IRT data analysis. Using the R package *mirt* ([Chalmers, 2012](#)), we obtained estimate $\rho_{\hat{\theta}\hat{\theta}'} = 0.620$. This value is almost equal to $\hat{\lambda}_4$ for the 5th percentile [Hunt and Bentler \(2015\)](#) proposed, but this may be coincidental.

Measurement precision

Reliability never played a major role in IRT because the availability of the Fisher information function, which equals the inverse of the standard error of estimate $\hat{\theta}$ conditional on θ , denoted $\sigma_{\hat{\theta}|\theta}$, made it superfluous. The information function and its inverse can be estimated for the J items and each of the individual items. Conditional standard error $\sigma_{\hat{\theta}|\theta}$ is the IRT analogy of the standard error of measurement in CTT and has the advantage that it depends on the scale of θ . The standard error thus shows for item j where it measures most precise on the θ scale, and items can be selected such that the resulting test has measurement precision where it is most useful. An example is a cut score θ_c for selecting students for a follow-up program, where precision at the cut score is mandatory and items are selected that optimize precision at θ_c . In intelligence testing, one may require equal precision along the whole scale and items that provide measurement precision at levels varying from low to high covering the whole scale range.

A feature of IRT is the possibility to tailor a test to an individual's unknown θ value in a search process, which starts with a few "average" items, estimates θ_i , then based on $\hat{\theta}_i$ selects the next item, updates the estimate of θ_i , selects the next items, and so on until $\sigma_{\hat{\theta}|\theta}$ is considered small enough for the measurement's purpose. This procedure is known as tailored or adaptive testing. It can only

work well when for the attribute of interest, the researcher has access to a large item bank of which the items have been pretested in large samples, so that their parameter estimates are known with enough precision to consider them the true parameter values.

Comments

From the advent of IRT in the 1950s, reliability never played an important role and attention soon was directed at the information function and the conditional standard error. The availability of the information function undeniably has great advantages for constructing tests. Like the FA approach, the IRT approach to reliability is based on latent variables, whereas CTT and GT approaches deal with observable variables. A wide variety of methods allows the estimation of reliability of observable test scores and latent variables, but the question whether the test measures the attribute of interest requires an independent study.

Conclusion

We discussed the four main approaches to reliability discussed in the psychometric literature. CTT is a random measurement-error theory that does not model the true score. In this sense, CTT makes abundantly clear that the question whether the test measures the attribute of interest well must be answered separately from a reliability analysis by means of a validity study. Reliability concerns the issue whether the measurement is repeatable, and validity addresses the question whether the test measures the attribute of interest. Whereas CTT focuses on a single test administered to a well-defined population under standardized conditions, GT separates person measurement from situational influences the test constructor wishes to control when they have a confounding effect on the person measurement. Getting these sources, such as test length and performance raters, under control means estimating the variance components in the test scores they represent and assessing the numbers of items and raters needed to reduce their confounding influence on person measurement. GT does not criticize CTT and is not at odds with CTT but generalizes it to obtain better control over testing conditions considered relevant.

The FA approach to reliability models the true score by means of a factor model, introducing a distinction between wanted and unwanted factors, the latter sometimes referred to as systematic error sources. Reliability is limited to wanted factors, often a factor common to the items. Whereas FA is a linear model that assumes continuous item scores modeled by one or more factors, IRT is a non-linear model for discrete item scores dependent upon one or more latent variables playing the same role in IRT as factors do in FA. Compared to CTT and GT, FA and IRT posit a particular structure on the items and the relationship between the items. Whether this is useful depends on the goodness-of-fit of the hypothesized FA and IRT models, and whether the fit is good depends on the strength of the theory about the attribute underlying the construction of the measurement instrument. As with CTT and GT, validity research is key to useful measurement giving it its meaning, and reliability expresses the precision of person measurement given the model on which it is based.

References

- Bentler, P.M., Woodward, J.A., 1980. Inequalities among lower bounds to reliability: With applications to test construction and factor analysis. *Psychometrika* 45, 249–267.
- Bollen, K.A., 1989. *Structural Equations With Latent Variables*. Wiley, New York, NY.
- Brennan, R.L., 2001. *Generalizability Theory*. Springer, New York, NY.
- Briesch, A.M., Swaminathan, H., Welsh, M., Chafouleas, S.M., 2014. Generalizability theory: A practical guide to study design, implementation, and interpretation. *J. Sch. Psychol.* 52, 13–35. <https://doi.org/10.1016/j.jsp.2013.11.008>.
- Chalmers, R.P., 2012. Mirt: A multidimensional item response theory package for the R environment. *J. Stat. Software* 48 (6), 1–29. <https://doi.org/10.18637/jss.v048.i06>.
- Cho, E., Kim, S., 2015. Cronbach's coefficient alpha: Well known but poorly understood. *Organ. Res. Methods* 18, 207–230.
- Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N., 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. Wiley, New York.
- Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.
- Embretson, S.E., 1991. A multidimensional latent trait model for measuring learning and change. *Psychometrika* 56, 495–515.
- Feldt, L.S., Steffen, M., Gupta, N.C., 1985. A comparison of five methods for estimating the standard error of measurement at specific score levels. *Appl. Psychol. Meas.* 9, 351–361.
- Flora, D.B., 2020. Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Pract. Psychol. Sci.* 3, 484–501. <https://doi.org/10.1177/2515245920951747>.
- Guttman, L., 1945. A basis for analyzing test-retest reliability. *Psychometrika* 10, 255–282.
- Holland, P.H., Hoskens, M., 2003. Classical test theory as a first-order item response theory: Application to true-score prediction from a possibly nonparallel test. *Psychometrika* 68, 123–149.
- Hunt, T.D., Bentler, P.M., 2015. Quantile lower bounds to reliability based on locally optimal splits. *Psychometrika* 80, 182–195.
- Jackson, P.H., Agunwamba, C.C., 1977. Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: algebraic lower bounds. *Psychometrika* 42, 567–578.
- Jöreskog, K.G., 1971. Statistical analysis of sets of congeneric tests. *Psychometrika* 36, 109–133.
- Kim, S., 2012. A note on the reliability coefficients for item response model-based ability estimates. *Psychometrika* 77, 153–162. <https://doi.org/10.1007/S11336-011-9238-0>.
- Kuder, G.F., Richardson, M.W., 1937. The theory of estimation of test reliability. *Psychometrika* 2, 151–160.
- Lek, K.M., Van de Schoot, R., 2018. A comparison of the single, conditional and person-specific standard error of measurement: What do they measure and when to use them? *Front. Appl. Math. Stat.* 4, 40. <https://doi.org/10.3389/fams.2018.00040>.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Erlbaum, Hillsdale, NJ.

- Lord, F.M., 1983. Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. *Psychometrika* 48, 233–245.
- Martin, S.R., Rast, P.S., 2022. The reliability factor: modeling individual reliability with multiple items from a single assessment. *Psychometrika*.
- McDonald, R.P., 1999. *Test Theory: A Unified Treatment*. Erlbaum, Mahwah, NJ.
- McNeish, D., 2018. Thanks coefficient alpha, we'll take it from here. *Psychol. Methods* 23 (3), 412–433. <https://doi.org/10.1037/met0000144>.
- Mellenbergh, G.J., 1994. A unidimensional latent trait model for continuous item responses. *Multivariate Behav. Res.* 29, 223–236.
- Mellenbergh, G.J., 1996. Measurement precision in test score and item response models. *Psychol. Methods* 1, 293–299.
- Mellenbergh, G.J., 1998. Het één-factor model voor continue en metrische responsen (The one-factor model for continuous and metric responses). In: Van den Brink, W.P., Mellenbergh, G.J. (Eds.), *Testleer en testconstructie*. Boom, Amsterdam, pp. 155–186.
- Novick, M.R., Lewis, C., 1967. Coefficient alpha and the reliability of composite measurements. *Psychometrika* 32, 1–13.
- Oosterwijk, P.R., Van der Ark, L.A., Sijtsma, K., 2017. Overestimation of reliability by Guttman's λ_4 , λ_5 , and λ_6 and the greatest lower bound. In: Van der Ark, L.A., Cuipepper, S., Douglas, J.A., Wang, W.-C., Wiberg, M. (Eds.), *Quantitative psychology research: The 81th Annual Meeting of the Psychometric Society 2016*, Asheville NC, USA. Springer, New York, pp. 159–172.
- Pradt, J.M., Sijtsma, K., (in press). Statistical properties of reliability methods from classical test theory and factor analysis. In: Wiberg, M., Molenaar, D., González, J., Kim, J.S. (Eds.), *Quantitative Psychology: The 86th Annual Meeting of the Psychometric Society, Virtual 2021*. Springer, New York, NY, pp. xxx–yyy.
- Pradt, J.M., Van den Bergh, D., Sijtsma, K., Moshagen, M., Wagenmakers, E.J., 2021. Bayesian estimation of single-test reliability coefficients. *Multivariate Behav. Res.* <https://doi.org/10.1080/00273171.2021.1891855>.
- Raykov, T., Marcoulides, G.A., 2019. Thanks coefficient alpha, we still need you! *Educ. Psychol. Meas.* 79 (1), 200–210. <https://doi.org/10.1177/0013164417725127>.
- Raykov, T., Shrout, P.E., 2002. Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Struct. Equ. Model.* 9, 195–212.
- Raykov, T., 1997. Estimation of composite reliability for congeneric measures. *Appl. Psychol. Meas.* 21, 173–184.
- Raykov, T., 2001. Bias of coefficient α for fixed congeneric measures with correlated errors. *Appl. Psychol. Meas.* 25, 69–76.
- Revelle, W., Condon, D.M., 2019. Reliability from α to ω : a tutorial. *Psychol. Assess.* 31 (12), 1395–1411. <https://doi.org/10.1037/pas0000754>.
- Revelle, W., Zinbarg, R.E., 2009. Coefficients alpha, beta, omega, and the GLB: comments on Sijtsma. *Psychometrika* 74, 145–154.
- Shavelson, R.J., Webb, N.M., 1991. *Generalizability Theory. A Primer*. Sage, Thousand Oaks, CA.
- Shavelson, R.J., Webb, N.M., Rowley, G.L., 1989. Generalizability theory. *Am. Psychol.* 44, 922–932.
- Sijtsma, K., Pradt, J.M., 2021. Part II: On the use, the misuse, and the very limited usefulness of Cronbach's alpha: Discussing lower bounds and correlated errors. *Psychometrika* 86, 843–860. <https://doi.org/10.1007/s11336-021-09789-8>.
- Sijtsma, K., Van der Ark, L.A., 2021. *Measurement Models for Psychological Attributes*. Chapman & Hall/CRC, Boca Raton, FL.
- Sijtsma, K., 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, 107–120.
- Ten Berge, J.M.F., Sočan, G., 2004. The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika* 69, 613–625.
- Ten Berge, J.M.F., Zegers, F.E., 1978. A series of lower bounds to the reliability of a test. *Psychometrika* 43, 575–579.
- Van der Linden, W.J. (Ed.), 2016. *Handbook of Item Response Theory: Volume One. Models*. Chapman & Hall/CRC, Boca Raton, FL.
- Van der Palm, D.W., Van der Ark, L.A., Sijtsma, K., 2014. A flexible latent class approach to estimating test-score reliability. *J. Educ. Meas.* 51, 339–357.
- Winer, B.J., 1962. *Statistical Principles in Experimental Design*. McGraw-Hill, Tokyo, Japan.
- Woodward, J.A., Bentler, P.M., 1978. A statistical lower-bound to population reliability. *Psychol. Bull.* 85, 1323–1326.
- Zinbarg, R.E., Revelle, W., Yovel, I., Li, W., 2005. Cronbach's α , Revelle's β , and McDonald's ω_H^2 : their relations with each other and two alternative conceptualizations of reliability. *Psychometrika* 70, 123–133.

Validity

Suzanne Lane^a and Alexandra Pérez^b, ^a University of Pittsburgh, Pittsburgh, PA, United States; and ^b University of Connecticut, Storrs, CT, United States

© 2023 Elsevier Ltd. All rights reserved.

Fairness as validity	35
Linking values to validity and fairness	36
Threats to validity and fairness	36
Argument-based approach to validity	36
Inferences in the IUA	37
Enacted interpretations and uses	37
Principled approaches in test design to support the IUA	38
Sources of validity evidence	38
IUA and validity argument use case	38
Performance inference	39
Scoring inference	39
Generalization inference	39
Extrapolation inference	39
Decision/use inference	40
Theories of action to support the IUA	40
Alternate perspectives on the meaning of validity	40
Consequences: an essential aspect of IUA and validity arguments	40
Consequences and tests as levers of educational change	41
Evaluating the consequences of tests used as levers of educational change	42
Concluding remarks	42
References	43

Validity is at the core of all aspects of testing from enumerating the intended score interpretations and uses of the test to evaluating the consequences of intended and actual test use. Validity refers to the extent to which evidence supports or refutes intended and actual test score interpretations and uses as well as resulting consequences. Validation is the systematic process of obtaining and evaluating this evidence. We argue that validity must be addressed throughout the testing process, beginning with defining the construct to be assessed. A construct refers to the knowledge, cognitive skills and strategies, and other attributes to be assessed.

We advocate the argument-based approach to validity because it provides a framework for articulating intended score interpretations and test uses as well as their consequences, both intended positive consequences and potentially unintended negative consequences, and for identifying the body of evidence needed to support score interpretations and uses of tests. Such a framework also allows for the articulation and evaluation of potential alternative explanations. The validity of score interpretations and test use rests on intentional test design that is responsive to different groups in the targeted population. Such groups typically are defined by ethnicity, race, culture, language dominance, disability, economic status, and other relevant experiential, social, and personal characteristics. To address fairness and equity issues in testing an essential aspect of the argument-based approach to validity are logical and empirical evaluations of the validity of score meaning and test use for different groups in the targeted population.

In this article, we begin with a brief discussion on the inherent relationship between validity and fairness. We agree with others that fairness investigations are validity investigations. Following this, we discuss the argument-based approach to validity, including how principled approaches to test design and theories of actions can support the interpretation and use argument and the validity argument. As an illustration, we lay out an interpretation and use argument and a validity argument for a particular use case. We then discuss the central role of the consequences of test use in the argument-based approach.

Fairness as validity

We adopt the prevailing view that fairness in testing is essential to the validity of score interpretation and test use, and concerns with fairness must be addressed throughout the testing process, from defining and labeling the construct, to designing, developing, administering, and scoring the test, to interpreting scores and evaluating test use and their consequences. Validity theorists have forwarded the view that fairness investigations are validity investigations. Messick (1998) advocated that beliefs and values bear on fairness evaluations of testing programs and he considered validity and fairness appraisals akin to value judgments. Kane (2012) furthered this view by indicating that fairness and validity investigations are similar ways of examining the same question: “Are the proposed interpretations and uses of the test scores appropriate for a population over some range

of contexts" (p. 177). Fairness investigations have emphasized the comparability of score interpretations, appropriateness of rules used to make decisions, and comparability of consequences across groups (Kane, 2010) however we contend that fairness investigations must begin with an examination of how the construct of interest is defined and by whom. These fundamental fairness investigations fall within the realm of validity. The *Standards for Educational and Psychological Testing* (American Educational Research Association (AERA) et al., 2014) also views fairness as a fundamental validity issue that requires attention throughout the design and development of the test as well as when interpreting and using scores. We agree with this perspective that fairness investigations *are* validity investigations.

For assessments to support the validity and fairness of score meaning and test use, individuals, regardless of group affiliation, should have equitable opportunity to access the construct measured. This requires that validity and fairness investigations begin with an appraisal of the appropriateness of the conceptualization of the construct for different groups in the targeted population. It is not necessarily the targeted construct that is assessed, instead it is the individual's construction of knowledge and ways of knowing that arise from one's social, cultural, and personal experiences (Gordon and Rajagopalan, 2016; Hood, 1998; Messick, 1998; Nortvedt et al., 2020; Randall, 2021). The knowledge, skills, and other attributes evoked by examinees when responding to assessment tasks depend on the assessment context as well as the individual's cultural, linguistic, and experiential histories. Assessments, in particular educational assessments, that account for the different ways individuals know and express their understandings in their design and development have the potential to enhance the validity and fairness of score interpretations and test use (Trumbull and Nelson-Barber, 2019). Being responsive to different cultures in assessment design recognizes testing as a situated culture practice (Gordon and Rajagopalan, 2016; Hood, 1998). We and others argue that this calls for a critical appraisal of the values, beliefs, and assumptions held by those who mandate, design, develop, and use tests.

Linking values to validity and fairness

Decisions made in test design, development, and use rest on value judgments (Cronbach, 1976; Messick, 1989; Kane, 2006), and such judgments can affect the validity and fairness of score interpretations and test use for individuals with different cultural and linguistic backgrounds. Each step of the testing process rests on value judgements including the articulation of the intended purposes and uses of a test, the conceptualization of the construct, the design and development of tasks and scoring procedures, the development of performance level descriptors, the development of performance standards for determining competency and certification, and the design of score reports and accompanying interpretive guides. Value judgments also underlie the identification of the intended positive consequences and the potential unintended, negative consequences.

Attention to differing values can uncover fairness and equity issues in the design, development, and use of tests for individuals from different backgrounds. As an example, consider the case of using a test for decisions regarding high school graduation. Although its intended consequence of raising standards and improving learning for all students is worthy, such use has had a differential effect on students who traditionally have been underserved. Judgments regarding whether the benefit of using tests in graduation decisions to raise standards outweigh their undesirable effect on different groups of students depends on the beliefs, values, and assumptions held by different stakeholders (Lane, 2020).

Threats to validity and fairness

Two sources of potential threats to the validity of score interpretations and uses are construct underrepresentation and construct-irrelevant variance (Messick, 1989). Construct underrepresentation occurs when the assessment does not capture important facets of the targeted construct. The extent to which the assessment is too narrow or underrepresents the construct will impact the generalizability of score inferences to the broader construct domain. As an example, if a science inquiry assessment does not full represent inquiry skills as laid out in the construct definition for the test, generalizing performance on the assessment to the full range of those inquiry skills is jeopardized.

Construct-irrelevant variance occurs when the assessment measures one or more constructs that are irrelevant to the targeted construct in addition to measuring the targeted construct. Construct-irrelevant variance may differentially impact the performance of different groups of individuals. As an example, a mathematics assessment which has a substantial reading load may depress the performance of learners whose primary language is not the tested language. These two threats to validity play an essential role in the appraisal of the validity and fairness of score interpretations and test use.

Argument-based approach to validity

The argument-based approach to validity has been widely recognized for evaluating score interpretations and test use (Cronbach, 1971, 1988; Kane, 1992, 2006, 2013; Shepard, 1993). The concept of an argument-based approach to validity can be traced to the seminal article by Cronbach and Meehl (1955) where they discuss how validation involves defending a proposed interpretation of test scores. Later, Cronbach (1988) extended his thinking regarding the argument-based approach to validity by acknowledging the importance of linking social and personal consequences and values in the argument.

Kane (2006) expanded on the argument-based approach to validity by specifying an interpretive argument and a validity argument. To emphasize a focus on the validity of test use as well as score interpretation, Kane (2013) shifted from using the term

“interpretive argument” to “interpretation and use argument”, although it should be noted that he emphasized test use in the interpretive argument. The interpretation and use argument (IUA) explicitly links the inferences from test performance to decisions and uses, including the actions resulting from the decisions and uses. Consequences of test decisions and uses therefore are integral to the argument-based approach to validity. The validity argument provides a comprehensive framework, including the rationales, assumptions, and backing or evidence, to help support or refute the inferences and uses specified in the IUA. The IUA guides all aspects of the testing process, including test design and use, provides the foundation for validation, and is critically evaluated against plausible alternative interpretations and uses (Kane, 2004).

The validity argument entails obtaining both logical and empirical evidence to evaluate the soundness of the claims and underlying assumptions in the IUA. As the backing for the IUA, the validity argument involves an overall evaluation of the appropriateness of the proposed score interpretations and test uses by providing a coherent analysis of the evidence for and against proposed score interpretations and uses (AERA et al., 2014; Cronbach, 1988; Kane, 1992, 2006, 2013). The validity of a score interpretation and use depends on the strength of the chain of evidence leading from making inferences from the task to performance to the inferences and uses of the scores and their resulting consequences. Shepard (1993) succinctly stated that the validity argument frames the questions:

What does the testing practice claim to do? What are the arguments and the intended aims of the test? What does the test do in the system other than what it claims, for good or bad? (p. 429).

Cronbach (1998) reminds us that when articulating a validity argument, one needs to play the “devil’s advocate” (p. 5) by identifying competing or rival hypotheses about score interpretations and uses. The argument-based approach provides a practical approach to validation by specifying the inferences and assumptions for appraisal and allows for prioritizing the forms of evidence needed in terms of the most pertinent inferences of the argument that require support (Kane, 1992, 2004).

Inferences in the IUA

The IUA consists of inferences connecting the observed test performance to the conclusions of interest. Inferences relevant to most IUAs are the scoring, generalization, extrapolation, and decision/use inferences proposed by Kane (2006, 2013). However, every IUA does not need to address every inference, but every IUA should include the most critical inferences. As a presumptive argument, Kane (2006, 2013) illustrated the IUA and its inferences using Toulmin’s (1958) model for presumptive inferences. As Kane (2006, 2013) describes, each inference goes from a data point to a claim. The inference employs a warrant which is a general rule for making claims based on data and this warrant is based on assumptions. The backing, or evidence, provides support for the warrant and its assumptions or it might refute the warrant and its assumptions.

The scoring inference links the observed performance of an examinee to a score by a scoring procedure developed during test design. The warrant is that a scoring rule assigns a score to the test performance accurately and consistently. The generalization inference refers to generalizing the examinee’s score over various measurement conditions such as items, occasions, and raters. The observed score is an estimate of the universe score and score interpretations are “extended from a claim about observed performances to a claim about expected performances over the universe of generalization” (Kane, 2016, p. 25). The extrapolation inference extends the claims of examinees’ observed scores to their performance in “real-world” scenarios. The warrant is that examinees who can apply their knowledge and skills to the assessment tasks can apply those knowledge and skills in the real world. The decision/use inference involves making decisions, such as selection, placement, certification, and proficiency decisions, based on the examinee’s observed score. The warrant is a rule for how the decision will be made. Evaluating decisions about individual and aggregate scores involves appraisal of both intended and unintended consequences.

In addition to these inferences in the interpretive chain, we propose the “performance inference” that would occur prior to the scoring inference, providing a link from the task to the performance. The claim would be that the task elicits the intended knowledge and skills, and the assessment is representative of the defined construct. The performance inference positions the conceptualization of the construct and the design and development of the assessment at the forefront of the IUA and validity argument. In their applications of the IUA, Chapelle et al. (2007, 2010) include domain description prior to the scoring inference and similarly, Shaw et al. (2012) include construct representation. The evaluation of the construct representation of tasks in terms of knowledge, skills, and other relevant attributes as well as principled approaches to test design provide backing, or evidence, for the performance inference. The performance inference explicitly links the IUA and validity argument to assessment design and development.

Enacted interpretations and uses

The interpretation and use of scores in practice may be quite different from the intended interpretations and uses. Consequently, it is insufficient to evaluate only the intended interpretation and uses of test scores, the actual or enacted interpretations and uses must be part of the IUA and validity argument (Coburn and Turner, 2011; Moss, 2013, 2016; O’Leary et al., 2017). The enacted interpretations and uses are more diverse and complex than the intended interpretations and uses (Coburn and

Turner, 2011; Moss, 2013, 2016), and may result in undesirable consequences for some individuals or groups. Test use varies depending on the context and needs of those who are using the tests, and these uses and needs, including local uses and needs, as well as their consequences require an evaluation in terms of their appropriateness and potential negative impact on individuals and groups.

Principled approaches in test design to support the IUA

Principled approaches to test design, such as Evidence Centered Design (ECD) (Mislevy et al., 2003), provide a framework for articulating the knowledge, skills and other attributes needed for successful task performance. As an example, within the ECD framework, evidence observed in examinee performance on tasks is used to support claims about individual's knowledge and skills. An explicit delineation of the claims and evidence to support these claims can help minimize construct-irrelevant variance and help ensure appropriate representation of the targeted knowledge and skills. Such principled approaches to assessment design provide validity evidence for evaluating score meaning and use.

The use of principled approaches to test design in conjunction with universal design principles provides a mechanism to make assessments responsive to different cultural, linguistic, and educational groups. Universal design principles for learning include the need for multiple means of representation, multiple means of action and expression, and multiple means of engagement (Thompson et al., 2002). The application of these universal design principles to test design, development, and administration recognizes that individuals from different cultural, linguistic, and educational backgrounds may have different ways of knowing and understanding and different ways of demonstrating their knowing and understanding. The application of universal design principles throughout every phase of test design, development, and implementation affords accessible assessments by minimizing construct-irrelevant barriers that may impede test performance.

Sources of validity evidence

The 2014 *Standards* provide five sources of evidence, evidence based on test content, response processes, internal structure, relations to other variables, and consequences of testing.¹ These five forms of validity evidence can be used to inform our thinking when designing validity investigations. Some researchers (Sireci, 2013; Zumbo and Chan, 2014) have suggested to use these sources of validity evidence as a framework for validation research, however others believe they are insufficient and lead to inadequate validation programs (Chapelle, 2011; Newton, 2016). As an example, Chapelle (2011) argued for examining the inferences in the argument instead of consulting a list of sources of validity evidence. We and others contend that the IUA provides a more comprehensive framework for validation by linking intended inferences and uses to evidence as well as linking plausible alternative inferences and uses to evidence (Lane and Marion, in press). The five sources of evidence proposed by the *Standards* in addition to other sources of evidence however can be encompassed within IUA and validity arguments.

IUA and validity argument use case

To illustrate the inferences in an IUA consider an assessment system used to examine the proficiency of students in disciplines such as mathematics, science, civics, history, and language arts. The assessment system includes performance tasks that are designed locally as well as performance tasks that are designed at the state or country level, and these performance tasks are administered at instructionally appropriate times throughout the year.² Such a performance assessment system provides an opportunity for teachers, culture experts, measurement professionals, and other relevant experts to collaboratively develop local performance tasks and scoring rubrics that reflect how students come to know and understand content within their own culture while adhering to rigorous content standards. A goal of this type of assessment system is that it leads to more valid score interpretations and uses.

For this use case, we provide a description of the inferences, some underlying assumptions, and evidence to support these assumptions. Of course, there may be other inferences and assumptions dependent on the specific use case. We also address the need to provide evidence in support of the inferences for different groups in the targeted population. For example, the response processes evoked by the assessment should be similar across different groups, the scoring of the performances should not bias one group over another, the generalizability of the inferences across groups should be comparable, and the consequences of the uses of the assessment should not inappropriately favor one group over another.

¹The 2014 *Standards* used the phrase "evidence for validity and consequences of testing" as if to separate validity from consequential evidence. The *Standards* also links validity "to interpretation of test scores for proposed uses of tests" (p. 11). As expressed by Lane and her colleagues (Lane and Niepokoj, 2018, Lane and Marion, in press), although the *Standards* offers much needed advice on validation efforts, it provides a narrow view of validity, stating that validity refers to score interpretations "for proposed uses" instead of explicitly stating "and their uses". The *Standards* also provides an ambiguous view on whether the use of tests and their consequences are part of the IUA and validity argument.

²This proposed performance assessment case draws on features of the Performance Assessment of Competency-Based Education (PACE) pilot initiative (New Hampshire Department of Education, 2016) that consists of a combination of locally developed performance tasks and common performance tasks shared by participating districts with a smaller statewide assessment component.

Performance inference

The performance inference links the tasks to the student performance. For this use case, the underlying assumption is that the tasks elicit the intended knowledge and skills and the assessment represents the intended construct for all students. Evidence to support this assumption includes, but is not limited to, the following:

1. Documentation specifying how the construct was conceptualized and defined and what groups were involved in defining the construct. (Content standards may be used to define the construct.)
2. Documentation specifying how principled approaches were used in designing the assessment tasks and rubrics, including the use of cognitive, socio-cognitive, and sociocultural theories of learning.
3. Examinee response evidence such as think alouds for different cultural and linguistic groups to evaluate whether the intended knowledge and skills are evoked.
4. Content and cultural expert review of the tasks and rubrics.
5. Alignment studies that evaluate the link among the defined construct (e.g., content standards), assessment specifications, and assessment tasks.

As an example, asking students to think aloud as they perform the task as well as evaluating any written responses to evaluate the knowledge, skills, and strategies evoked by the students provides backing or evidence to support or refute the performance inference.

Scoring inference

The scoring inference links the performance to the score. Pertinent assumptions underlying the scoring inference for this use case is that (1) rubrics capture the intended knowledge and skills and account for different ways of expressing those knowledge and skills, (2) raters appropriately internalize the rubric criteria and apply them accurately, and (3) tasks are administered appropriately. Evidence to support these assumptions includes the following:

1. Documentation of the design of the scoring rubrics and weighting of the criteria to produce scores.
2. Evaluation of the accuracy and consistency of implementation of the rubrics by the teachers, including an evaluation for responses from different groups of individuals.
3. Evidence of the appropriate administration of the performance tasks.
4. Audit of the scores assigned by the teachers.
5. Evaluation of the comparability of scores from locally developed tasks and state/country developed tasks.
6. Evaluation of the validity/comparability of score meaning across different groups of individuals.

An evaluation of the accuracy and consistency of the implementation of the rubrics may include rater think alouds to examine whether raters are interpreting and applying the rubric criteria to the responses appropriately. It may also involve having more than one teacher in a locale score a subset of the responses to the task. An audit of the teacher ratings may involve a random sample of performance tasks from a random sample of teachers to be scored by trained raters designated by the state. A systematic evaluation of the comparability of the scores from locally developed tasks and state developed tasks may also shed light on the accuracy of local scoring and the rigor of locally developed tasks. The comparability of scores in this assessment scenario however may be reduced to some extent in favor of using tasks that are more culturally and instructionally responsive.

Generalization inference

The generalization inference links the score to the broader construct or domain. Generalizability studies (Cronbach et al., 1972) that examine the generalizability of the scores over tasks, raters, and other relevant measurement conditions provide evidence to evaluate the generalization inference. An assumption is that students should be able to generalize their knowledge and skills to other tasks within the domain. Evidence to support the assumption underlying the generalization inference includes the following.

1. Generalizability studies for tasks, raters, and other relevant measurement conditions.
2. Evidence to support the generalizability of scores across groups.
3. Studies of the internal structure of the assessment across groups to lend support for reporting one or more scores as intended.

As an example, a multi-group factor analysis provides evidence of whether the structure holds across relevant groups. The generalization inference may be threatened if the performance assessment consists of a relatively small number of tasks, but in this use case the performance tasks may better represent the construct domain because they are implemented throughout the instructional year, leading to a broader range of tasks.

Extrapolation inference

The extrapolation inference links the test performance to the expected performance in real-world contexts. Because performance assessments intend to assess students at a deeper, more critical-thinking level, extrapolating to the real-world context may be less of an issue than with other forms of assessments. Evidence to support the assumption that student performance on the assessment is similar to their performance in the real world includes the following.

1. Expert review of the alignment of knowledge and skills assessed to the knowledge and skills used in diverse real-world contexts and communities.
2. Expected relationships between assessment scores and scores from other criterion measures hold and are similar across different groups.

Because a goal of the locally developed tasks is to design them to be more relevant to cultural groups, an examination of whether they do reflect the way students construct, organize, and display knowledge and skills in their own communities would support this goal.

Decision/use inference

The decision/use inference links the score to the decisions, uses, and resulting consequences. The assumptions underlying the decision/use inference for this use case includes (1) assignment to performance levels is appropriate, (2) uses of test results are appropriate, and (3) positive consequences are maximized while negative consequences are minimized. Evidence to support the assumptions underlying the decision/use inference includes the following.

1. Standard setting method and procedures are appropriate.
2. Decision consistency and accuracy is appropriate and similar across groups.
3. Intended positive consequences are maximized while unintended negative consequences are minimized for all groups of individuals.

Standard setting evidence includes an evaluation of whether the standard setting procedures result in reasonable performance standards and whether potential adverse impact was considered. Performance assessments have been used as levers of educational change, so an evaluation of the consequences is essential, such as whether instruction and learning for all students has benefited because of the use of performance assessments (Lane, 2014, 2020).

This use case which specifies the inferences, assumptions, and the evidence to support them requires tailoring to specific assessment purposes and contexts. It highlights the most pertinent assumptions and evidence, but it is not exhaustive. A goal of ours was also to address fairness and equity issues in laying out the IUA and validity argument. We acknowledge however that additional validity and fairness evidence to support score meaning and use for different groups would strengthen the IUA and validity argument.

Other IUAs and validity arguments have been developed for different types of assessments. As an example, Kane and his colleagues (Kane, 2004; Kane et al., 2017) illustrated a general IUA for credentialing testing programs that delineates associated warrants, assumptions, and evidence for the assumptions. Lane and Marion (in press) propose a general IUA and validity argument for the design of more culturally responsive and sustainable educational assessments.

Theories of action to support the IUA

Theories of action provide a framework for the IUA and validity argument by identifying critical testing components and their logical points of impact. A theory of action for an assessment system should include the rationale for the use of the assessment, the mechanisms that will lead to the intended outcomes, and the intended and potentially unintended outcomes or consequences. Bennett (2010) indicated that the theory of action for an assessment system provides a framework for bringing together score meaning and impact claims, with an emphasis on claims about the intended consequences of the assessment system for both individuals and institutions. In describing the theory of action for the Cognitively Based Assessment of, for and as Learning (CBAL) initiative, Bennet (2010) illustrated how the assessment processes and mechanisms connect the design of the assessment with the intended outcomes and potential unintended outcomes, emphasizing the effects of the assessment system on individuals and institutions as well as the mechanisms behind the effects.

Alternate perspectives on the meaning of validity

Although the conceptualization of validity as pertaining to the appropriateness of test score interpretations and uses, as well as resulting consequences has been widely received some still adhere to more restricted conceptualizations of validity (Borsboom et al., 2004; Cizek, 2012). Borsboom et al. (2004) indicated that “a test is valid for measuring an attribute if and only if: (a) the attribute exists; and (b) variations in the attribute causally produce variations in the outcomes of the measurement procedure” (1061). This view links validity to a causal explanation, implying that changes in the tested attribute cause changes in the item responses, and that validation does not apply to score use. We argue that such restrictive views are limited because they do not consider test use and their consequences as integral aspects of the IUA and validity argument.

Consequences: an essential aspect of IUA and validity arguments

We adopt the view that the appropriateness of test-based decisions and uses is central to validity and therefore validity investigations must attend to consequences of score interpretations and test use, both the intended positive consequences and the potential

unintended negative consequences (Cronbach, 1971; Kane, 1992, 2006, 2013; Lane et al., 1998; Linn, 1997; Messick, 1989; Shepard, 1997). Five decades ago, Cronbach (1971) considered the appraisal of decisions and actions based on tests as part of validity evaluation and Messick (1983) made the role of consequences an inherent aspect of validity theory. In response to an ongoing debate regarding the inclusion of consequences of test use in validity investigations, Linn (1997) pushed for consequences to be integral to validity:

the best way of encouraging adequate consideration of major intended positive effects and plausible unintended negative effects of test use is to recognize the evaluation of such effects as a central aspect of test validation (p. 16).

In Messick's (1989) conceptualization of validity, he discussed the relationship of consequences with test use

Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy and appropriateness of inferences and actions* based on test score and other modes of assessment (p. 13).

Messick's (1989, 1994) perspective on adverse social consequences as an aspect of validity however, limited them to being traced back to construct-irrelevant score variance or to construct underrepresentation. Unfortunately, the validity chapter in the *Standards* (2014) adopted this restricted view of the relationship between consequences and validity. We agree with Kane (2006) when he pressed that Messick's perspective gave consequences a secondary role in validation. Instead, like Kane and Cronbach (1971, 1988), we support the position that consequences have an essential role in the validity argument in that negative consequences could invalidate test use even if they cannot be traced to construct-irrelevant variance or construct underrepresentation. Cronbach (1988) pushed further indicating that those who mandate tests are obligated to evaluate whether a test use has "appropriate consequences for individuals and institutions, and especially to argue against adverse consequences" (p. 6). Consequences that may have adverse impact and systemic consequences are integral to the IUA and validity argument (Kane, 2013). Adverse social and personal consequences of testing bear on both validity and fairness issues.

In the context of language assessments, Bachman and his colleague (Bachman, 2005; Bachman and Parker, 2010) proposed an assessment use argument (AUA) that provides a framework for linking performance on assessments to decisions and use, laying out the intended consequences and the potentially unintended negative consequences. The AUA links score interpretations to decisions and consequences, specifying claims about outcomes, intended and actual interpretations, decisions, and consequences. Like Kane's IUA, the AUA positions the articulation of the intended positive consequences and potentially unintended consequences of test use at the beginning of the testing process. The warrant associated with consequences in the AUA states that "the *consequences* of using the assessment and making intended decisions will be beneficial to individuals, the program, company, institution, or system, or to society at large" (Bachman, 2005, p. 19). The AUA emphasizes not only intended but also *actual* interpretations, decisions, and consequences, broadening the need for appraisal to actual test use and their resulting consequences. As indicated by Bachman (2005), including consequences in the AUA positioned investigations of test consequences on instruction and learning as part of the AUA instead of them being external to the validity framework.

Consequences and tests as levers of educational change

The goal of test-based accountability systems in the US is to improve the educational opportunities provided to all students, narrow the achievement gaps, and prepare students for college and careers by holding schools, educators, and sometimes, students accountable. This section discusses validity issues with respect to the use of tests as levers of educational change in the US. It should be noted however that these validity issues are relevant to test-based accountability systems and the use of tests as policy levers in other countries. An early example was England's initiation of the 1862 "payment-by-results" policy which was implemented because of an increased demand for student enrollment in elementary schools (Stobart and Eggen, 2012). This policy required that teacher salaries reflect their student performance in reading, writing and mathematics which resulted in an instructional focus on drilling students in these subjects that lasted 30 years. The use of tests to serve as tools for educational reforms by policy makers requires evidence to support this intended use and resulting intended and actual consequences (Cronbach, 1982). There is an inherent conflict however in using tests for both instruction and accountability purposes as test design and development decisions differentially affect the validity of using the same test for these two purposes (Lane and DePascale, 2016).

The use of tests as policy levers bears on fairness and equity issues. Consequences of tests used in accountability systems typically have different impacts on different groups of students and in different schools and economic areas, and these differential impacts require evaluation as part of the validity argument (Lane and Stone, 2002). Differential test outcomes for students living in economically marginalized communities in the US, have led to undesirable consequences, including a narrowing of the curriculum and instruction and an increase in high school dropout rates. Evaluations of intended and actual test use in accountability systems need to establish whether the intended positive consequences are realized and whether potential unintended negative consequences are minimized. A systematic evaluation includes an appraisal of whether the test achieves its goals and the costs in attempting to achieve its goals.

Different stakeholders, including policymakers, advocacy groups, administrators, educators, parents, students, business leaders, and the community, have different, and sometimes conflicting, perspectives and values on the purposes and uses of tests as well as their consequences. While the stakeholders may change depending on the purpose of the test, the context for use, and the locale, differing values and perspectives need attention throughout the testing process. [Bachman \(2002\)](#) proposed the need for different IUAs that account for the conflicting values of different stakeholders which is likely impractical in many contexts. At the very least, the articulation of competing values from stakeholders will allow for a better understanding of the potential consequences, both positive and negative, of test use ([Lane, 2014, 2020](#)).

Evaluating the consequences of tests used as levers of educational change

Evaluating the effectiveness of tests used for accountability purposes requires attention to how well the tests are achieving their goals in terms of the motivation and practices of school administrators and educators and in terms of student achievement and learning ([Lane and Stone, 2002](#)). Those who mandate and use tests must make a case for the appropriateness of the decision rules in the use context, and consideration of whether any potential unintended negative consequences outweigh intended positive consequences prior to the use of a test as an instrument of policy. The articulation of intended and potential unintended consequences should occur in conjunction with the specification of the claims in the IUA. The validity argument should entail a systematic evaluation of how the testing achieves its goals and the cost of attempting to obtain its goals.

In previous research, a conceptual framework for evaluating consequences of assessment and accountability systems, both intended and unintended, was proposed ([Lane et al., 1998, 2002; Lane and Stone, 2002; Lane, 2020](#)). The framework provides a principled approach for evaluating consequences in the IUA and validity argument and can be modified for different purposes of educational assessments. We developed and implemented the framework to evaluate the consequences of the Maryland State Performance Assessment Program (MSPAP) which was comprised of performance tasks integrated across content disciplines. The goal of MSPAP was to promote performance-based instruction by providing opportunities for students to engage in problem solving and reasoning. The IUA addressed the intended consequences at multiple levels, individual, school, and state, by articulating a set of claims that required logical and empirical validity evidence at each level. It involved the specification of the forms of evidence to accumulate for both positive and potentially negative consequences for different participants, including students, teachers, and administrators, and at different levels, including state, district, school, class, and individual. To provide evidence of the assessment's effect on instruction, forms of evidence include surveys, focus groups, instructional logs, and a systematic evaluation of instructional tasks and assessments.

Using this framework to evaluate the consequences of MSPAP, *we found that teacher reported use of performance-based instruction accounted for differences in school performance on MSPAP in reading, writing, mathematics, and science. Schools that focused on performance-based instruction, such as the engagement of students in critical thinking and reasoning, had higher MSPAP scores than schools in which their instruction was less performance-based* ([Lane et al., 2002](#)). Another example of a performance-based assessment system and an examination of its impact is PACE (Performance Assessment of Competency-Based Education) which was an innovative assessment program for some districts in a state in the US. One assumption underlying the use of this performance-based assessment system, which was implemented throughout the school year, was that the knowledge and skills acquired by students in the PACE program would transfer to the statewide test that primarily consisted of selected-response items. Results from two studies provided evidence of this assumption in that for those students who participated in PACE there was not a negative impact on their statewide test scores ([Evans, 2019; Stone and Evans, 2021](#)).

When using test scores to make inferences regarding the quality of instruction, contextual information informs the inferences and resulting actions ([Haertel, 1999](#)). As an example, in the MSPAP study, we found that the percent of students receiving free or reduced lunch, which served as a proxy for socio-economic status, albeit an imperfect proxy, was not significantly related to school performance gains on MSPAP in mathematics, writing, science, and social studies ([Lane et al., 2002](#)). Although this can be conceived as a positive finding, some would argue that schools in economically disadvantaged areas should have shown greater gains because of their lower initial performance. The framework my colleagues and I developed stipulates the need for contextual information, such as socioeconomic status, funding per student, access to quality curriculum and materials, and stability of administrators, teachers, and students, to inform the evaluation of consequences. An understanding of contextual variables is critical for identifying barriers in achieving the intended outcomes.

Concluding remarks

In this article, we discussed how validity is central to every aspect of the testing process, from conceptualizing the targeted construct to evaluating the consequences of test use. Furthermore, fairness and equity issues are fundamental to the validity of score meaning and test use for individuals from different population groups. We and others contend that assessment design, development, and use should begin with an interpretation and use argument that explicitly links the inferences from the tasks to the decisions and uses, and to the resulting consequences. A validity argument then provides an evaluation of the claims and assumptions that are laid out in the IUA.

We maintain that values and beliefs affect the choices made when delineating the IUA and validity argument as well as when designing, developing, and using assessments. These values and beliefs must be understood and attended to because they bear

on the validity and fairness of score meaning and use for different groups of individuals. Educational testing programs that explicitly address how cultural, linguistic, and social experiences and practices shape how individuals learn, know, and demonstrate their understanding have the potential to lead to more equitable and valid test score interpretations, uses, and consequences.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Bachman, L.F., 2002. Alternative interpretations of alternative assessments: some validity issues in education performance assessments. *Educ. Meas.: Issues Pract.* 21 (3), 5–18. <https://doi.org/10.1111/j.1745-3992.2002.tb00095.x>.
- Bachman, L.F., 2005. Building and supporting a case for test use. *Lang. Assess. Q.: Int. J.* 2 (1), 1–34. <https://doi.org/10.1207/s15434311laq0201.1>.
- Bachman, L.F., Palmer, A.S., 2010. *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World*. Oxford University, Oxford, UK.
- Bennett, R., 2010. Cognitively based assessment of, for, and as learning: a preliminary theory of action for summative and formative assessment. *Measurement* 8 (2–3), 70–91. <https://doi.org/10.1080/15366367.2010.508686>.
- Borsboom, D., Mellenbergh, G., van Heerden, J., 2004. The concept of validity. *Psychol. Rev.* 111, 1061–1071.
- Chapelle, C.A., 2011. Validity argument for language assessment: the framework is simple.... *Lang. Test.* 29 (1), 19–27.
- Chapelle, C.A., Enright, M.K., Jamieson, J.M., 2007. *Building a Validity Argument for the Test of English as a Foreign Language*. Routledge, Abingdon.
- Chapelle, C.A., Enright, M.K., Jamieson, J., 2010. Does and argument-based approach to validity make a difference? *Educ. Meas.* 29 (1), 3–13.
- Coburn, C.E., Turner, E.O., 2011. Research on data use: a framework and analysis. *Measurement* 9 (4), 173–206.
- Cizek, G., 2012. Defining and distinguishing validity: interpretations of score meaning and justification of test use. *Psychol. Methods* 17 (1), 31–43.
- Cronbach, L.J., 1971. Test validation. In: Thorndike, R.L. (Ed.), *Educational Measurement*, second ed. American Council on Education, Washington, DC, pp. 443–507.
- Cronbach, L.L., 1976. Equity in selection – where psychometrics and political philosophy meet. *J. Educ. Meas.* 13 (1), 31–41.
- Cronbach, L.J., 1982. *Designing Evaluations of Educational and Social Programs*. Jossey-Bass, San Francisco, CA.
- Cronbach, L.J., 1988. Five perspectives on validity argument. In: Wainer, H., Braun, H. (Eds.), *Test Validity*. Erlbaum, Hillsdale, NJ, pp. 3–17.
- Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N., 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. Wiley, New York, NY.
- Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. *Psychol. Bull.* 52 (4), 281–302.
- Evans, C., 2019. Effects of New Hampshire's innovative assessment and accountability system on student achievement outcomes after three years. *Educ. Pol. Anal. Arch.* 27 (10). <https://doi.org/10.14507/epaa.27.4014>.
- Gordon, E.W., Rajagopalan, K., 2016. *The Testing and Learning Revolution: The Future of Assessment in Education*. Palgrave Macmillan, New York, NY.
- Haertel, E.H., 1999. Validity arguments for high-stakes testing: in search of the evidence. *Educ. Meas.* 18 (4), 5–9.
- Hood, S., 1998. Introduction and overview: assessment in the context of culture and pedagogy: a collaborative effort, a meaningful goal. *J. Negro Educ.* 67 (3), 184–186.
- Kane, M.T., 1992. An argument-based approach to validity. *Psychol. Bull.* 112 (3), 527–535.
- Kane, M.T., 2004. Certification testing as an illustration of argument-based validation. *Measurement* 2 (3), 135–170.
- Kane, M.T., 2006. Validation. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger, Washington, DC, pp. 17–64.
- Kane, M.T., 2010. Validity and fairness. *Lang. Test.* 27 (2), 177–182.
- Kane, M.T., 2012. Validating score interpretations and uses. *Lang. Test.* 29 (1), 3–17.
- Kane, M.T., 2013. Validating the interpretations and uses of test scores. *J. Educ. Meas.* 50 (1), 1–73.
- Kane, M.T., 2016. Validation strategies: delineating and validating proposed interpretations. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 64–80.
- Kane, M.T., Clauser, B.E., Kane, J., 2017. A validation framework for credentialing tests. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, NY, pp. 21–40.
- Lane, S., 2014. Validity evidence based on testing consequences. *Psicothema* 26 (1), 127–135.
- Lane, S., 2020. Test-based accountability systems: the importance of paying attention to consequences. In: ETS Research Report Series: 17th William H. Angoff Memorial Lecture. ETS, Princeton, NJ.
- Lane, S., DePascale, C., 2016. Psychometric considerations for performance-based assessments and student learning objectives. In: Braun, H. (Ed.), *Meeting the Challenges to Measurement in an era of Accountability*. Routledge, New York, pp. 77–106.
- Lane, S., Marion, S., in press. Validity in educational measurement. In: Cook, L., Pitoniak, M. (Eds.), *Educational Measurement*, fifth ed.
- Lane, S., Niepokoj, D., 2018. Validity Evidence Based on Consequences of Assessment and Accountability Programs (Paper presented at a coordinated session at the annual meeting of the National Council on Measurement in Education, New York, NY).
- Lane, S., Parke, C.S., Stone, C.A., 1998. A framework for evaluating the consequences of assessment programs. *Educ. Meas.* 17 (2), 24–28.
- Lane, S., Parke, C.S., Stone, C.A., 2002. The impact of a state performance-based assessment and accountability program on mathematics instruction and student learning: evidence from survey data and school performance. *Educ. Assess.* 8 (4), 279–315.
- Lane, S., Stone, C.A., 2002. Strategies for examining the consequences of assessment and accountability programs. *Educ. Meas.* 21 (1), 23–30.
- Linn, R.L., 1997. Evaluating the validity of assessments: the consequences of use. *Educ. Meas.* 21 (1), 14–16.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education and Macmillan, New York, pp. 13–104.
- Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educ. Res.* 23 (2), 13–23.
- Messick, S., 1998. *Consequences of Test Interpretation and Use: The Fusion of Validity and Values in Psychological Assessment*. ETS, Princeton, NJ.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., 2003. On the structure of educational assessments. *Measurement* 1 (1), 3–62.
- Moss, P.A., 2016. Shifting the focus of validity for test use. *Assess Educ. Princ. Pol. Pract.* 23 (2), 236–251. <https://doi.org/10.1080/0969594X.2015.1072085>.
- Moss, P.A., 2013. Validity in action: lessons from studies of data use. *J. Educ. Meas.* 50 (1), 91–98.
- New Hampshire Department of Education, 2016. *Moving from Good to Great in New Hampshire: Performance Assessment of Competency Education (PACE)*. Author, Concord, NH.
- Newton, P.E., 2016. Macro- and micro-validation: beyond the “five sources” framework for classifying validation evidence and analysis. *Practical Assess. Res. Eval.* 21 (12).
- Nortvedt, G.A., Wiese, E., Brown, M., Burns, D., McNamara, G., O'Hara, J., et al., 2020. Aiding culturally responsive assessment in schools in a globalising world. *Educ. Assess. Eval. Account.* 32 (1), 5–27.
- O'Leary, T.M., Hattie, J.A.C., Griffin, P., 2017. Actual interpretations and uses of scores as aspects of validity. *Educ. Meas.* 36 (2), 16–23.
- Randall, J., 2021. “Color-Neutral” is not a thing: redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educ. Meas.* 40, 2. <https://doi.org/10.1111/emip.12429>.
- Shaw, S., Crisp, V., Johnson, N., 2012. A framework for evidencing assessment validity in large-scale, high-stakes international examinations. *Assess Educ. Princ. Pol. Pract.* 19 (2), 159–176.

- Shepard, L.A., 1993. Evaluating test validity. *Rev. Res. Educ.* 19 (1), 405–450.
- Shepard, L., 1997. The centrality of test use and consequences for test validity. *Educ. Meas.* 21 (1), 5–8, 13.
- Sireci, S.G., 2013. Agreeing on validity arguments. *J. Educ. Meas.* 50, 99–104.
- Stobart, G., Eggen, T., 2012. High-stakes testing-value, fairness, and consequences. *Assess Educ. Princ. Pol. Pract.* 19 (1), 1–6.
- Stone, A.L., Evans, C., 2021. NH PACE Student Outcome Evaluation After Five Years (Coordinated Paper Sessions at annual conference of the National Council on Measurement in Education).
- Thompson, S.J., Johnstone, C.J., Thurlow, M.L., 2002. Universal Design Applied to Large Scale Assessments. NCEO Synthesis Report 44. University of Minnesota, National Center on Educational Outcomes, Minneapolis, MN.
- Toulmin, S., 1958. *The Uses of Argument*. Cambridge University press, Cambridge, UK.
- Trumbull, E., Nelson-Barber, S., 2019. The ongoing quest for culturally responsive assessment for indigenous students in the US. *Front. Educ.* <https://doi.org/10.3389/educ.2019.00040>.
- Zumbo, B.D., Chan, E.K.H. (Eds.), 2014. *Validity and Validation in Social, Behavioral, and Health Sciences*. Springer, New York.

Fairness

Maria Elena Oliveri^a, Mya Poe^b, and Norbert Elliot^c, ^a Buros Center for Testing, University of Nebraska, Lincoln, NE, United States; ^b Northeastern University, Harvard, MA, United States; and ^c New Jersey Institute of Technology, Newark, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	45
Setting the context: international examples of (un)fairness	45
Definitions: key terms related to fairness	46
Considerations: The future of fairness in measurement	48
Conclusion	49
References	49

Introduction

In 2016, a discussion about fairness might have begun with a general acknowledgment of the varied ways that fairness is understood in social, legal, and measurement contexts and then focus on issues specific to test design: “The goal of the fairness argument is to present evidence that the test is fair for various groups within the test-taking population” (Zieky, 2016, p. 96). Today, in the wake of the COVID-19 global pandemic beginning in 2019 and international Black Lives Matter demonstrations in 2020, conversations about fairness must take these global issues into account. Inattention to social context and justice is not an option anymore. Consequently, in this chapter, we acknowledge that fairness cannot be achieved without jointly considering inequity, anti-racism, and justice. The conventional definition of fairness offered in test design guides and professional standards seem woefully insufficient for a post-pandemic, anti-racist worldview. Fairness cannot be circumscribed solely to discussions of test design but should also include discussions of the processes used to make evaluative judgments of the meaning of test scores, the stakeholders involved in making them, the cultural and linguistic backgrounds of test takers, and the many ways test-based decisions have disadvantaged and continue to disadvantage subgroups. To this end, like Gipps and Stobart (2010), we take an inclusive view of educational assessment, including classroom and program assessment as well as high-stakes testing. We also see fairness as an ongoing project of justice (i.e., using assessment toward addressing injustice; Poe et al., 2018; Randall, 2021).

This entry has four sections. In the first section, we begin our discussion of fairness by presenting examples of (un)fairness in assessment systems used around the world and the types of challenges stakeholders and the public face in relation to addressing fairness considerations within those systems. Those examples inform the key terms we define in the second section. In our definitions, we include both terms used traditionally in measurement and newer terms. In the third section, we describe considerations on the future of fairness in measurement. The fourth section is the conclusion in which we provide a brief reflection on issues discussed in the article.

Setting the context: international examples of (un)fairness

Real-world examples, drawn from international contexts, such as Chile, the Middle East and North Africa, and the United States, demonstrate the difficulties of achieving fairness when resources are limited and tests are seen as barriers to accessing prestigious universities:

- On January 6 and 7, 2020, the Prueba de Selección Universitaria (PSU)—a battery of multiple-choice achievement tests taken by high school students applying for university admissions—was scheduled to be taken by approximately 300,000 applicants across Chile. Led by student organizations, the Coordinating Assembly of High-School Students, and the National Coordinator of High-School Students, protests erupted against the PSU. Protestors argued that the exams privilege wealthier students and increase economic inequality. Access to test centers were blocked and examination papers were burned. More than 27 people died, with thousands injured and arrested (Nugent, 2020). In October of 2021, large-scale demonstrations marked the anniversary of the protests (Cambero, 2021).
- On June 19, 2021, the Ministry of Education in Sudan ordered the nation’s internet service providers to shut down all mobile internet connections in advance of national secondary school examinations. In Algeria, the practice of shutting down the nation’s internet has occurred each year since 2016 during periods of national examinations. In Syria, the government stopped internet service across the country between May 31 and June 22 for approximately four and a half hours each day while students took their high-school examinations. And in Jordan, the government blocked the internet between June 24 and July 15, 2021, as students took their high-school examinations. Journalists noted that these practices violated United Nations Resolution 44/12 regarding internet shutdowns (Fatafta et al., 2021).
- On January 20, 2020, the Centers for Disease control confirmed the first case of COVID-19 in the US. By June 2022, more than one million people in the US had died of COVID 19 (Centers for Disease Control and Prevention, 2022a,b). The disease has had

a disproportionate effect on historically marginalized groups, resulting in huge disparities in hospitalizations and deaths (Centers for Disease Control and Prevention, 2022a,b; Lopez et al., 2021). The pandemic upended classroom teaching, forcing students from preschool to college age into online learning contexts. In US K-12 schooling, test scores in math and reading declined (Kuhfeld et al., 2022). In US higher education, the pandemic brought about many issues, including declines in student enrollment, declines in students' post-college labor outcomes, especially among students of color (Department of Education, 2021, p. 39, 41), and changes to admissions testing. By March of 2020, approximately 1000 of the 2300 private nonprofit and public bachelor's-granting colleges and universities offered students the option to apply without submitting standardized test scores. By October of 2020, nearly two-thirds of colleges and universities had adopted test-optional, or test score masked, policies in place (Conley and Masa, 2020). A February 2021 ACT survey of 4-year colleges found that 80% ended up being test-optional for the previous year and between 60 and 70% said they would likely remain test-optional or test blind post-COVID (ACT, 2021).

In terms of fairness, the PSU in Chile has been criticized for equity deficits in terms of economic inequality and opportunity to learn; in 2018, approximately 80,000 students did not have the opportunity to learn the material covered in the technical and professional section of the test—the educational track taken by the poorest in the nation. Politically, the PSU has come to symbolize a country divided between those desiring social change and those devoted to conservative policies; an October PSU anniversary demonstration, held in advance of November 2021 presidential and legislative elections, reveals the polarized nature of the country. Internet shutdowns in Middle East and North Africa have been framed as violating the United National General Assembly resolution “strongly condemning the use of Internet shutdowns to intentionally and arbitrarily prevent or disrupt access to or dissemination of information online” (United National General Assembly, 2020, p. 3)—an example of how educational testing and human rights are conflated when technology becomes a feature of the testing environment. In the US, the decline of standardized college admissions testing during the global COVID-19 pandemic strengthened the long-standing concerns that standardized admissions tests may not provide information on students that is sufficient to compensate for inequitable barriers to access created by the tests themselves. One early study of 100 private institutions has shown that test-optional admissions increase Pell Grant applicants, women, and first-time students from underrepresented racial/ethnic backgrounds (Bennett, 2021).

The examples above illustrate the complexities of fairness when we consider the vary unfair conditions of assessment globally. If we are to determine complexities related to fairness, including the ways assessment is interwoven with larger social issues such as public health crises, further understanding can also be gained by identifying key terms related to fairness. To this end, in the next section, we provide definitions useful in conceptualizing fairness.

Definitions: key terms related to fairness

Inherently, fundamental considerations of fairness are a part of education. When we narrow fairness applications to educational measurement alone, discussions of unfairness (fairness violations) arise across all settings, from elementary to secondary school group scoring programs to higher education admissions testing—as well as across purposes of testing, from classroom assessments to testing for licensure and certification. In this broad range of testing applications, it seems nearly impossible to begin a discussion of fairness without considering the larger context in which tests are used. This issue is ever more important when we use situated language frameworks in which local environments have a substantial influence on stakeholders impacted by tests and its scores: advisory board members, administrators, faculty, parents, professional organizations, students, tribes and tribal members, and the public (Gee, 2020; Mislevy, 2018). Within this context, we next describe key terms related to fairness. Our list of key terms includes both terms from the educational assessment community used traditionally in fairness related discussions and newer terms that reflect global calls for fairness and justice. These terms are organized in alphabetical order.

- *Anti-racist assessment.* These assessment practices disrupt racist beliefs around learning by actively confronting the economic, structural, and historical roots of inequality (Kendi, 2016). With anti-racist approaches to assessment, teachers and test designers “acknowledge both the current and historical role of race and racism in their own pedagogical and assessment practices” (Randall et al., 2021, p. 596).
- *Bias:* Sources of construct underrepresentation or construct-irrelevance variance may differentially disadvantage test-taker groups. In statistics, bias is identified through systematic error in a test score (American Educational Research Association et al., 2014).
- *Consequence:* Impact identifies the intended or unintended outcomes of using tests in particular ways, in certain contexts, with certain populations. Some consequences are directly related to the score interpretation and use, while others are related to testing contexts or unintended negative consequences (American Educational Research Association et al., 2014).
- *Culturally responsive assessment:* As explained by Montenegro and Jankowski (2017), culturally responsive assessment is “mindful of the student populations the institution serves, using language that is appropriate for all students when developing learning outcomes, acknowledging students’ differences in the planning phases of an assessment effort, developing and/or using assessment tools that are appropriate for different students, and being intentional in using assessment results to improve learning for all students” (p. 10) (See also Hood et al., 2015).
- *Culturally responsive indigenous evaluation:* Also known as an Indigenous Self-Determination Evaluation Model, CRIE “respects, recognizes, and values the inherent worth of Indian culture; is responsive to the communities’ needs as voiced by all members of

the tribal community; builds evaluation designs and processes around Indian assets and resources; and literally and figuratively employs Indians in every part of the process (program, policy, implementation, evaluation) to heal, strengthen, and preserve Indigenous societies for the next 7 generations" (Bowman, 2005, p. 8).

- *Culturally sustaining assessment*: As explained by Randall et al. (2021), "culturally sustaining approaches to assessment (a) draw on Black, Indigenous, People of Color (BIPOC) students' funds of knowledge (Vélez-Ibáñez and Greenberg, 1992), (b) are connected to the lives of these students, (c) allow them to demonstrate their competence in a variety of ways through community cultural wealth (Yosso, 2005), and (d) are embedded within a culturally sustaining curriculum. Such an approach would 'perpetuate and foster ... linguistic, literate, and cultural pluralism as part of the democratic project of schooling and as a needed response to demographic and social change' (Paris and Alim, 2014, p. 88)" (p. 596) (See also Ladson-Billings, 1995).
- *Disparate impact*: Differences in educational outcomes may result from facially neutral policies or practices. When used as both a statistical and conceptual tool, disparate impact analysis allows detailed analysis educational consequences, as well as legal strategies that may be used to redress inequity (Poe and Cogan, 2016).
- *Equity*: Gee (2008) calls for assurances that every individual has the opportunity to develop an identity as a competent learner, participate in communities of learning, and sustain a learning trajectory. Equity-centered design acknowledges that students have rich backgrounds of experience that, when acknowledged, provide opportunities for knowledge demonstration (Oliveri et al., 2020; Pullin, 2008).
- *Fairness*: "Validity of test score interpretations for intended use(s) for individuals from all relevant subgroups. A test that is fair minimizes the construct-irrelevant variance associated with individual characteristics and testing contexts that otherwise would compromise the validity of scores for some individuals" (American Educational Research Association et al., 2014, p. 219).
- *Opportunity to Learn (OTL)*: Exposure to particular knowledge, skills, and attitudes needed for curricular success. Consideration of OTL is important for the study of alignment between tests and curriculum, as well as allocation of resources needed to assure learning for all students (Moss et al., 2008).
- *Validity*: Extent to which evidence and theory support a specific interpretation and use argument. As is the case with fairness and reliability—the foundations of educational measurement—validity is considered as a process of evidence gathering (American Educational Research Association et al., 2014).
- *Reliability/Precision*: The degree to which test scores are consistent over repeated test applications. A test is said to provide evidence of reliability/precision when a measurement process is dependable, consistent, and free from random error (American Educational Research Association et al., 2014).
- *Social Justice*: Theories holding that fairness is understood through analysis of social contracts, distributive principles, and communal interaction. While theories of social justice vary, each maintains a commitment to principled strategies aimed at understanding and reducing inequality (Rawls, 1971/1999; Young, 2011).
- *Universal Design*: Approaches to educational measurement that maximize test accessibility for all test takers. Practices of universal design ensure that a test is fair for all test-takers, regardless of gender, age, language background, socioeconomic status, or disability (American Educational Research Association et al., 2014).

In general, some concepts—such as bias and reliability/precision—fall into technical work associated with educational measurement. Other concepts—such as culturally responsive assessment—are often associated with pedagogical practice. And yet other concepts, such as disparate impact, are associated with legal precedent. When studying fairness, it is important to emphasize commonalities among communities, from obligations to stakeholders to consideration of multidisciplinary collaboration. However, it is equally important to recognize that these definitions themselves are not without critics. For example, Randall et al. (2021) argue that OTL "has been used to justify construct definitions that exclude/attempt to erase the ways of knowing and understanding of BIPOC students. Simply put, constructs are defined in such a way that privilege whiteness and then BIPOC communities are informed that they simply have not had the opportunity to perfect/improve their whiteness" (p. 85). When defined with recognized qualifications, the terms defined above are one way of opening portals to common frames of reference (Meyer and Land, 2006).

Various fairness definitions have been offered in the *Standards for Educational and Psychological Testing* as well as educational measurement guidelines (American Educational Research Association et al., 2014; ETS, 2014; ITC, 2018). For a historical review of the *Standards*, see Sireci and Randall (2022). As Zwick (2019) has noted, increasing awareness of standards allow stakeholders to reach informed selection decisions, use multiple measures in making those decisions, and maintain decision transparency.

The best-known articulation of standards is that issued, in its present form, by three leading disciplinary organizations: American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME]. With origins in 1952, the present version of the *Standards* established fairness as a center of evidence equal to validity and reliability as a foundational concept (American Educational Research Association et al., 2014; American Psychological Association and Committee on Test Standards, 1952). As noted in the definition provided above, fairness is understood as the validity of test score interpretations for intended uses applied to individuals from all relevant subgroups. Further, a test is said to be fair if it minimizes the construct-irrelevant variance associated with individual characteristics and testing contexts that otherwise would compromise the validity of scores for some individuals. In associating fairness with validity, the authors demonstrate the integrative value of fairness as a fundamental category of evidence. Beyond the definition, the *Standards* for fairness focus on minimizing barriers to valid score interpretations, providing detailed comment on accommodations and safeguards, and emphasizing the general conditions needed allow diverse student groups opportunities to demonstrate the construct under examination. As such,

the authors adopt a distributed perspective in which fairness is both the goal for all testing applications and the responsibility of all parties involved in testing.

Of equal interest are the *ITC Guidelines for the Large-Scale Assessment of Linguistically Diverse Populations* (International Test Commission, 2018) and the *ETS International Principles for Fairness Review of Assessments* (Educational Testing Service, 2016). Both are focused on testing within or across countries or regions that are linguistically and culturally diverse. For example, the *International Test Commission's (2018) Guidelines* "are designed to inform test developers, psychometricians, and test users of the considerations that should be made to help ensure test fairness and score comparability to support meaningful inferences in culturally and linguistically diverse contexts" (p. 3). Emphasizing diversity, the *ITC Guidelines* define fairness as evidence that an adapted assessment—changes made in content, format, or test administration to increase access for cultural or linguistic groups—will result in valid demonstration of test-takers construct knowledge. In its definition of fairness, ETS cites the *Standards* (2014). These guidelines and principles are important to understanding the factors that impact fair assessment in diverse populations of test-takers, including legal language status, language of instruction, and resources for adapting tests. Adopting the sociocognitive and sociocultural perspectives identified above are a great advantage in understating diversity.

Considerations: The future of fairness in measurement

While Zwick (2019) has noted the importance of using existing standards to deepen stakeholder perspectives, she has also advocated the importance of candid and clear communication. As she observes, frank and honest communication allows identification of common ground among stakeholders, including discussions of test fallibility and avoidance of defensiveness. She is cautious, however, in recommending a single solution model to the solving the challenges of fairness in educational measurement. In her discussion of bias, for example, she quotes Cronbach (1976, p. 31): "Make no mistake. The issues will not be settled by mathematical specialists." Standards can't stop student protests, make governments restore the internet, or erase decades of racist, classist, sexist gatekeeping.

To try to imagine a discussion of fairness today without considering these larger social implications is shortsighted, but there are real questions as to what future of fairness is to be in educational assessment. To those ends, we offer four considerations as a way to reimagine the Gipps and Stobart discussion of relationships among equity, access, curriculum, and assessment in the previous edition of this Encyclopedia (2010, Table 1, p. 58).

- *Consider fairness as a form of purpose pluralism.* Newton (2017) has called attention to the significance of purpose pluralism by noting that "assessment design should be driven by a multiplicity of assessment purposes simultaneously" (p. 5). In advancing purpose pluralism, he describes the tensions between purism (a fixed belief that an assessment should be used for a single purpose) and pluralism (a fluid belief acknowledging the many aims and perspectives involved in coordinating the activities of test designers and users). In proposing purpose pluralism, Newton identifies important perspectives related to information quality (targeting use of assessment results to make decisions, with emphasis on a trustworthy and transparent basis for making important decisions), expertise (emphasizing fidelity to a specific and prespecified cluster of learning outcomes, with attention to providing assessment justification), and engagement (focusing on the consequences arising from anticipating assessment impact for various participants, with variations regarding how positive consequences may be achieved). In applications of purpose pluralism to fairness, it may be useful to think about a central definition that can be applied to a specific assessment, with tailored versions of that definition shaped by various stakeholders. Evidence of fairness can then be gathered in accordance with these definitions, with claims made in cases where the definitions are congruent, and qualifications made in cases where the definitions are incongruent.
- *Consider fairness as an ongoing process.* While statisticians are likely correct in noting that answers related to bias may not be settled by mathematical specialists, it is equally true that much of our understanding of fairness has benefitted enormously by psychometric analysis in which statistical methods are used determine disparate impact. In seminal work, Cleary (1968) used general linear modeling techniques to target evidence of test bias based on whether or not the criterion score was consistently too high or too low for members of the subgroup. This work combined emphasis on both social goals of equity and statistical goals of production. Perhaps, in the future, it will be equally true that questions regarding fairness will be answered by more fully explicating interpretation and use of scores (Kane, 2013) as well as their social consequences. As Camilli (2006) has presently observed, test developers and teachers alike are embedded in social contexts. The future may therefore well be determined by reconciliation of that situatedness—in terms of transparent test construction and clear community communication on the part of test developers, careful test selection and anticipated consequences on the part of teachers, and sustained research and action on the social uptake of test results. Sustained research and action on the social uptake of test results would invite the linking of historical injustice to current action and ongoing work toward achieving justice for all test takers (Hood and Hobson, 2008).
- *Consider fairness as advanced through anticipatory design practices.* The 21st century has witnessed a variety of consequence-centered, anticipatory approaches to designing a variety of assessment genres. These approaches find their origin in work by Mislevy et al. (2003) in Western theory. Using a conceptual assessment framework, evidence-centered design uses model-based reasoning to identify conceptual, delivery, scoring, and reporting models for any given measurement episode. This assessment framework has been followed by related design approaches, including Integrated Design and Appraisal Framework (IDAF, Slomp, 2016),

expanded evidence-centered design (Arieli-Attali et al., 2019), and equity-centered design (Oliveri et al., 2020). Similarly, theory of action techniques have been used to conceptualize the educational effectiveness of an innovation in terms of governing variables (curricular models), action strategies (curricular delivery strategies) and impact (consequences of those strategies) (Argyris, 1997). When applied to educational tests, classroom assessment, and community education evaluation, this approach holds the potential to help identify patterns of interactions among stakeholders and various forms of consequences, both intermediate and long term (Hazelton et al., 2021). In addition, a decolonial approach to anticipatory design is culturally responsive indigenous evaluation in which evaluators do not “impose a model, evaluation design, instruments, or tools upon the members of that community” (Waapalaneexkweew, 2018, p. 546). Instead, through “recognition of the legal implications of Indigenous sovereignty,” designers move beyond the “beads and feathers’ linguistic and cultural aspects of evaluation to work with tribal governments and communities as sovereign partners,” in the design, adaptation, problem-solving, and reporting of evidence-based evaluations (Waapalaneexkweew, 2018, p. 550). Taken together, evidence-centered design approaches, theory of action techniques, and culturally responsive indigenous evaluation provide powerful ways for assessment specialists and classroom teachers to consider factors related to evidence of fairness, validity, and reliability in a given measurement episode.

- *Consider fairness as understood through multidisciplinary and community collaboration.* While the twentieth century may be said to have proposed solutions to social challenges, the 21st century may be said to have viewed those very solutions as problems. From the present vantage point, addressing these problems will require multidisciplinary collaboration. As Oliveri et al. (2021) have shown in the case of a complex digital assessment, coordination of multidisciplinary action involving a wide variety of community stakeholders and experts (from computer science to policy analysis) is needed to maximize the desired, intended consequences from the use of an assessments—and to minimize undesirable, unintended effects.

Conclusion

Whether framed in terms of broad societal issues such as resource distribution or specific applications such as admissions testing, fairness is a critical subject to discuss in educational assessment. While the concepts involving evidence of fairness may be complex, they are nevertheless accessible. Even the most complex legal, statistical, and community-based discussions of fairness are possible, especially when these discussions are framed—at the outset—as accessible through multidisciplinary collaboration. Often, fairness is envisioned as an unattainable goal, but present international circumstances described in the introduction to this Encyclopedia entry reveal that position to be untenable. While it may be argued that it was impossible to prevent the adverse consequences identified in the case studies drawn from Chile, the Middle East and North Africa, and the United States, it can be equally argued that intentional use of the perspectives provided here could have lessened their impact. And in the wake of the COVID 19 pandemic and Black Lives Matter movement, discussions of fairness can never again be extracted from the social fabric in which fairness is woven.

References

- ACT, 2021. Summary findings: Survey of higher education enrollment and admissions officers. ACT. <https://chronicle.brightspotcdn.com/a1/52/df530674ab59217fc5f025d6a76/210212-hedsurveysummaryfindings-externaluse.pdf>.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association.
- American Psychological Association, Committee on Test Standards, 1952. Technical recommendations for psychological tests and diagnostic techniques: preliminary proposal. *Am. Psychol.* 7, 461–475.
- Argyris, C., 1997. Learning and teaching: a theory of action perspective. *J. Manag. Educ.* 21, 9–26. <https://doi.org/10.1177/105256299702100102>.
- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., von Davier, A.A., 2019. The expanded evidence-centered design (e-ECD) for learning and assessment systems: a framework for incorporating learning goals and processes within assessment design. *Front. Psychol.* 10, 1–17. <https://doi.org/10.3389/fpsyg.2019.00853>.
- Bennett, C.T., 2021. Untested admissions: examining changes in application behaviors and student demographics under test-optional policies. *Am. Educ. Res. J.* <https://doi.org/10.3102/00028312211003526>.
- Bowman, N.R., 2005. Government to government evaluation: Issues and strategies for conducting evaluation with tribal governments. Paper Presented at the Annual Conference of the American Evaluation Association, Toronto, Ontario, Canada.
- Camero, F., 2021. Polarized Chile marks anniversary of 2019 protests as election nears. Reuters. <https://www.reuters.com/world/americas/chile-braces-protests-crossroads-election-nears-2021-10-18/>.
- Camilli, G., 2006. Test fairness. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger, pp. 221–256.
- Centers for Disease Control and Prevention, 2022a. Health equity considerations and racial and ethnic minority groups. <https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html>.
- Centers for Disease Control and Prevention, 2022b. COVID Data Tracker. <https://covid.cdc.gov/covid-data-tracker/#datatracker-home>.
- Cleary, T.A., 1968. Test bias: prediction of Negro and white students in integrated colleges. *J. Educ. Meas.* 5, 115–124.
- Conley, B., Masa, B., 2020. The decline of testing affects more than testing. Inside High. Ed. <https://www.insidehighered.com/admissions/views/2020/12/14/decline-standardized-testing-affects-more-testing-opinion>.
- Cronbach, L.J., 1976. Equity in selection: where psychometrics and political philosophy meet. *J. Educ. Meas.* 13, 31–42.
- Department of Education, 2021. Education in a Pandemic: The Disparate Impacts of COVID-19 on America’s Students. Office for Civil Rights. <https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf>.
- Educational Testing Service, 2014. ETS Standards for Quality and Fairness. Educational Testing Service.
- Educational Testing Service, 2016. ETS International Principles for Fairness Review of Assessments. Educational Testing Service.

- Fatafta, M., Mnejja, K., Anthonio, F., 2021. Internet shutdowns during exams: When MENA governments fail the test. *accessnow*. <https://www.accessnow.org/mena-internet-shutdowns-during-exams/>.
- Gee, J.P., 2008. A sociocultural perspective on opportunity to learn. In: Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J. (Eds.), *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press, pp. 76–108.
- Gee, J.P., 2020. *What is a Human? Language, Mind, and Culture*. Palgrave Macmillan.
- Gipps, C., Stobart, G., 2010. Fairness. In: McGraw, B., Baker, E., Peterson, P. (Eds.), *International Encyclopedia of Education*, third ed. Elsevier, pp. 56–60.
- Hazelton, L., Nastal, J., Elliot, N., Burstein, J., McCaffrey, D.F., 2021. Formative automated writing evaluation: a standpoint theory of action. *J. Resp. Wri.* 7, 37–91. <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1041&context=journalrw>.
- Hood, S., Hopson, R.K., 2008. Evaluation roots reconsidered: Asa Hilliard, a fallen hero in the “Nobody Knows My Name” Project, and African educational excellence. *Rev. of Educ. Res.* 78 (3), 410–426. <https://doi.org/10.3102/0034654308321211>.
- Hood, S., Hopson, R.K., Kirkhart, K.E., 2015. Culturally responsive evaluation: Theory, practice, and future implications. In: Newcomer, K.E., Hatry, H.P., Wholey, J.S. (Eds.), *Handbook of Practical Program Evaluation*, fourth ed. Jossey-Bass, pp. 281–317.
- International Test Commission, 2018. ITC Guidelines for the Large-Scale Assessment of Linguistically Diverse Populations. Version 4.2. Buros Center for Testing.
- Kane, M.T., 2013. Validating the interpretation and uses of test scores. *J. Educ. Meas.* 50, 1–73. <https://doi.org/10.1111/jedm.12000>.
- Kendi, I.X., 2016. *Stamped from the Beginning: The Definitive History of Racist Ideas in America*. Nation Books.
- Kuhfeld, M., Soland, J., Lewis, K., 2022. Test score patterns across three COVID-19-impacted school years. In: Annenberg Institute at Brown University EdWorkingPaper 22–521. <https://doi.org/10.26300/ga82-6v47>.
- Ladson-Billings, G., 1995. Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32 (3), 465–491.
- Lopez, L., Hart, L.H., Katz, M.H., 2021. Racial and ethnic health disparities related to COVID-19. *JAMA* 325 (8), 719–720. <https://doi.org/10.1001/jama.2020.26443>.
- Meyer, J.H.F., Land, R. (Eds.), 2006. *Overcoming Barriers to Social Understanding: Threshold Concepts and Troublesome Knowledge*. Routledge.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., 2003. On the structure of educational assessment. *Measurement* 1, 3–62. https://doi.org/10.1207/S15366359MEA0101_02.
- Mislevy, R.J., 2018. *Sociocognitive Foundations of Educational Measurement*. Routledge.
- Montenegro, E., Jankowski, N.A., 2017. Equity and assessment: moving towards culturally responsive assessment (Occasional Paper No. 29). National Institute for Learning Outcomes Assessment. <https://files.eric.ed.gov/fulltext/ED574461.pdf>.
- Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J. (Eds.), 2008. *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press.
- Newton, P.E., 2017. There is more to educational measurement than measuring: the importance of embracing purpose pluralism. *Educ. Meas.* 36, 5–15. <https://doi.org/10.1111/emip.12146>.
- Nugent, C., 2020. Why Chile’s SATs have become the new frontline of inequality protests. *Time*. <https://time.com/5770308/chilestudent-protests/>.
- Oliveri, M.E., Slomp, D., Elliot, N., Rupp, A., Mislevy, R., Vezzu, M., Tackitt, A., Nastal, J., Phelps, J., Osborn, M., 2021. Introduction: meeting the challenges of workplace English communication in the 21st Century. *J. Wri. Anal.* 5, 1–33. <https://wac.colostate.edu/docs/jwa/vol5/intro.pdf>.
- Oliveri, M.E., Nastal, J., Slomp, D., 2020. Reflections on equity-centered design. In: ETS Research Report 20–22. Educational Testing Service.
- Paris, D., Alim, H.S., 2014. What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harv. Educ. Rev.* 84 (1), 85–100.
- Poe, M., Cogan, J.A., 2016. Civil rights and writing assessment: using the disparate impact approach as a fairness methodology to evaluate social impact. *J. Writ. Assess.* 9. <http://journalofwritingassessment.org/article.php?article=97>.
- Poe, M., Inoue, Asao, A.B., Elliot, N., 2018. *Writing Assessment, Social Justice, and the Advancement of Opportunity*. The WAC Clearinghouse; University Press of Colorado. <https://doi.org/10.37514/PER-B.2018.0155>.
- Pullin, D.C., 2008. Assessment, equity, and opportunity to learn. In: Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., Young, L.J. (Eds.), *Assessment, Equity, and Opportunity to Learn*. Cambridge University Press, pp. 333–351.
- Randall, J., 2021. ‘Color-neutral’ is not a thing: redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educ. Meas.* 40, 82–90. <https://doi.org/10.1111/emip.12429>.
- Randall, J., Poe, M., Slomp, D., 2021. Ain’t Oughta Be in the dictionary: getting to justice by dismantling anti-black literacy assessment practices. *J. Adolesc. Adult Literacy* 64 (5), 594–599.
- Rawls, J., 1999. *A Theory of Justice*. Cambridge University Press (Original work published 1971).
- Sireci, S., Randall, J., 2022. Evolving notions of fairness in testing in the United States. In: Clauser, B.E., Bunch, M.B. (Eds.), *The History of Educational Measurement*. Routledge, pp. 111–135.
- Slomp, D., 2016. An integrated design and appraisal framework for ethical writing assessment. *J. Writ. Assess.* 9 (1). <http://journalofwritingassessment.org/article.php?article=91>.
- United National General Assembly, 2020. Promotion and Protection of All Human Rights, Civil, Political, Economic, Social and Cultural Rights, including the Right to Development. Resolution 44.12. <https://undocs.org/en/A/HRC/RES/44/12>.
- Vélez-Ibáñez, C.G., Greenberg, J.B., 1992. Formation and transformation of funds of knowledge among U.S. Mexican households. *Anthropol. Educ. Q.* 23 (4), 313–335.
- Waapalaneekweew (Nicole R. Bowman-Farrell), 2018. Looking backward but moving forward: honoring the sacred and asserting the sovereign in indigenous evaluation. *Am. J. Eval.* 39 (4), 543–568.
- Yosso, T.J., 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethn. Educ.* 8, 69–91.
- Young, I.M., 2011. *Responsibility for Justice*. Oxford University Press.
- Zieky, M., 2016. Developing fair tests. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Routledge, pp. 81–99.
- Zwick, R., 2019. Fairness in measurement and selection: statistical, philosophical, and public perspectives. *Educ. Meas.* 38, 34–41. <https://doi.org/10.1111/emip.12299>.

Classical test theory

Michael Chajewski, Principal Research Scientist, Psychometric & Research Services, Pearson, New York, NY, Brooklyn, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

History	51
True score	51
Models	52
Parallel	53
Tau-equivalent	53
Essentially tau-equivalent	53
Congeneric	54
CTT as a special case	54
Interpretational contexts	54
Reliability	55
Validity	55
Generalizability	55
Test development	56
Discussion	57
References	58
Relevant website	58

History

The term Classical Test Theory (CTT) started appearing with the publication of Lord and Novick's book *Statistical Theories of Mental Test Scores* (1968), primarily to differentiate Item Response Theory (IRT) and the assessment topics that preceded it (see "Item Response Theory (Unidimensional)" by Terry Ackerman this volume). Since CTT encompasses a wide range of psychometric measurement issues, which continued to expand into the formulation of contemporary test models, the term is somewhat misleading in that CTT does not imply a singular test measurement theory divergent from modern assessment development practices.

The historic context that led to the development of what is presently referred to as CTT, has long cultural antecedence in assessment practices going back to the early 7th century CE. The imperial Chinese civil service examination (Keju) system, believed to be in continuous operation from approximately 606 CE to 1905, is possibly the oldest systematic testing program in the world (Suen and Yu, 2006). What differentiates the Keju system from its contemporaries and earlier assessments, and thus warrants its inclusion in a discussion about CTT, is its appreciable focus on scoring consistency and the reduction of measurement error. Because of its use of oral examinations and constructed responses, the Keju system imposed specific question and response formats with the aim to assure consistency of performance evaluations. Albeit not statistically defined in psychology until the late 19th century, the desire for consistency in evaluating skills and knowledge would give rise to the initial interpretational context of classical test theory: reliability.

In a simplified overview, the development of classical test theory can be viewed through the lens of three measurement aspects: (1) Reliability, (2) validity, and (3) generalizability (McDonald, 1999). While these attributes are not independent of one another in contemporary general educational measurement/psychometric practice, their sequential historic formalization provides an evident progression of the CTT interpretational context in which observable test scores were defined.

True score

The scientific revolution of the 17th century provided the conceptual antecedents to many areas of modern science, including psychometrics. Scientific measurements and the concern with their precision morphed into what Traub (1997) discussed as the "zeitgeist at the turn of the 20th century, [which] was the notion of errors in scientific observations" (p. 8). The historic developments leading up to that point in time introduced the notion that an observable score, which we shall refer to as X , was not necessarily the true value of the quantified construct. Instead, the observable score was an adjusted true state or value of the measurement, which we shall refer to as the true score, or τ , affected by an exogenous source(s) of error, which we shall refer to as ϵ . In its simplest form, as reaffirmed by Gulliksen (1950) in his seminal book on CTT, *Theory of Mental Tests*, the relationship between the three terms can be expressed as:

$$X_{ij} = \tau_i + \varepsilon_i \quad (1)$$

where X is an observable score for person i on a test or item j , τ is the person's true score on that test or item, and ε is the measurement error for person i .

Eq. (1) represents the very core of classical test theory and is the equation most associated with it. It provides the general relationship between a person's observable quantified performance and the underlying true score of the criterion of interest. Deceptively simple, nearly the next 100 years were spent in providing the statistical links and relational definitions among these three terms in the framework of cognitive and non-cognitive assessments.

The ongoing statistical developments of the early 20th century and their adoption in psychology would express observed scores as realizations of a random variables, where a random variable was defined as a variable that takes on values according to a probabilistic function. The relationship between the observed score and the unobservable true score would be formalized as:

$$X|\tau_X \sim N(\tau_X, \sigma_\varepsilon^2) \quad (2)$$

where the observed scores in a population, given the same true score, are normally distributed with a mean equal to that true score and a variance equal to the variance across all observed scores in the population with the same true score. By extension the observed score and the true score have a bivariate Normal relationship as the true score itself is believed to vary across the population.

$$\tau \sim N(\mu_\tau, \sigma_\tau^2) \quad (3)$$

Here the true score, τ , is normally distributed with its specific population mean, μ_τ , and variance, σ_τ^2 .

It should be noted that the average of measurement errors, such as that depicted for a single measurement in Eq. (1), is believed to be equal to 0,

$$\mu_\varepsilon = 0 \quad (4)$$

as the errors in the measurements are believed to be random and independent of the true scores. Because the errors are independent of the true scores, errors should not covary systematically with the true score, meaning that their covariance (and thus correlation) should also be 0:

$$\sigma_{\tau,\varepsilon} = \rho_{\tau,\varepsilon} = 0 \quad (5)$$

It follows from the assumption of statistical independence between the true scores and the errors, and the assumption in Eq. (4) as it relates to Eq. (1), that the expectations (or means) of the three terms should relate to one another just as the terms in Eq. (1):

$$\mu_X = \mu_\tau + \mu_\varepsilon \quad (6)$$

However, given that the average error is believed to be zero, Eq. (6) can be simplified to:

$$\mu_X = \mu_\tau \quad (7)$$

where the average observed score is believed to be equal to the average true score. This notion further expanded from multiple measurements in the population to parallel assessments of the same examinee postulating that an examinee's average observed score across parallel assessments should equal their average true score. By the same convention it can be derived that the variance of the observed scores is equal to the sum of the true score and error variances:

$$\sigma_X^2 = \sigma_\tau^2 + \sigma_\varepsilon^2 \quad (8)$$

It follows that in Eq. (1) the observable quantity of the score X is associated with its known descriptive statistics and distributional characteristics, while the terms on the right side of the equation require modeling. While the relational aspects laid out in Eqs. (2)–(8) form the principal assumptions of CTT, much attention has been given to the estimation of the true score and error variances, as these are not directly observable. Due to the nature of Eq. (1), and the established relationships between terms, if one of the two, true score or error quantities, could be approximated, Eq. (1) would become a simple equation with one unknown for which we can solve. Arriving at a universally accepted approximation, or practical analog, of either the true score or error, has proven to be challenging as measurement instrument attributes (viz. test characteristics) have made finding efficacious solutions elusive. Alternatively, approximating distributional estimates of these quantities has enabled further inferences about multiple measurements from either the same student or for a sample of examinees from the population.

Models

The CTT model in Eq. (1) is typically introduced, as we have done, in the context of an observable measurement akin to a performance evaluation. In educational measurement we could think of this performance evaluation, the observed score X , as the score a student may get on either a single test item or an overall assessment instrument. Because of this, the terms "item" and "test" will largely be used interchangeably. In fact, the assumptions of the CTT framework remain the same, regardless of the granularity at which they are applied, providing that the tenets of the theory that the observable score is comprised of a true score and error are applicable. Subsequently, CTT extends into the treatment of multiple observed scores measuring a common construct. The

approaches by which the relationship of multiple observed scores and related true scores can vary are commonly referred to as the classical theory true score measurement models.

Parallel

The parallel, or classically parallel, CTT true score model reflects the original sentiments of CTT in that each observable item score, or test score, measuring the same construct, is parallel to one another in the sense that for person i , their observed score j ($X_{i,j}$), is comprised of a person specific true score, τ_i , and person random error realization, ε_i ; basically the CTT model in Eq. (1). What becomes more explicit in the parallel measurement model is that this relationship is expected to hold across all parallel construct measurements for observed scores $j, j+1$, etc. Table 1 provides the general assumptions of this assumption for the expected observed score, the definition of the person specific true score, and the standard deviations of true scores and errors across multiple measurements.

If the variability of true scores and errors across measurements is constant then the average observed scores should also be equivalent (model 1 assumptions a, c and d in Table 1). This restrictive assumption of the parallel model, while computationally desirable as any construct relevant measurement would depend on only two parameters, is practically unrealistic.

Tau-equivalent

The assumptions of the parallel model (1.a and 1.d in Table 1) suggest that any measurement instance of the construct is affected by the random error, which is independent of the item. Whether considering a repeat measure on a single item (j), or a measurement of the same construct on a subsequent item ($j+1$), assuming that there are no changes in error variability across these instances is unlikely given that a variety of exogenous factors typically affect measurement.

The tau-equivalent measurement model relaxes the assumption of equivalent errors with the resulting assumption that the hypothetical measurement error across instances (items, test, etc.) will be different (Eq. 9 and assumption 2.d in Table 1).

$$\sigma_{\varepsilon j} \neq \sigma_{\varepsilon j+1} \quad (9)$$

Given the nature of observed test/item scores, the tau-equivalent model (and all subsequent models) is more closely aligned to practical observations of psychological measurements than the classically parallel true score model. Since the underlying true score that generates an observed score on item j and subsequent measurement instances ($j+1$) is believed to be the same, we refer to the models as tau-equivalent.

Essentially tau-equivalent

All CTT measurement models postulate that a universal person specific true score is the generative source of the observed score. However, given that the true score is believed to follow a distribution in the population, Eq. (3), a person's true score on a specific item or test (j), $\tau_{i,j}$, can be thought of as a realization of the person's specific true score, τ_i . When discussing distributional

Table 1 Classical test theory measurement models relating observable test scores with unobservable true scores.

	Model	Observed score	Assumptions
(1)	Parallel	$X_{i,j} = \tau_i + \varepsilon_i$	(a) $\mu_{Xj} = \mu_{Xj+1}$ (b) $\tau_i = \tau_{i,j} = \tau_{i,j+1}$ (c) $\sigma_{\tau j} = \sigma_{\tau j+1}$ (d) $\sigma_{\varepsilon j} = \sigma_{\varepsilon j+1}$
(2)	Tau-equivalent	$X_{i,j} = \tau_i + \varepsilon_{i,j}$	(a) $\mu_{Xj} = \mu_{Xj+1}$ (b) $\tau_i = \tau_{i,j} = \tau_{i,j+1}$ (c) $\sigma_{\tau j} = \sigma_{\tau j+1}$ (d) $\sigma_{\varepsilon j} \neq \sigma_{\varepsilon j+1}$
(3)	Essentially tau-equivalent	$X_{i,j} = (\tau_i + \mu_j) + \varepsilon_{i,j}$	(a) $\mu_{Xj} \neq \mu_{Xj+1}$ (b) $\tau_i = \tau_{i,j} - \mu_j = \tau_{i,j+1} - \mu_{j+1}$ (c) $\sigma_{\tau j} = \sigma_{\tau j+1}$ (d) $\sigma_{\varepsilon j} \neq \sigma_{\varepsilon j+1}$
(4)	Congeneric	$X_{i,j} = (\lambda_j \tau_i + \mu_j) + \varepsilon_{i,j}$	(a) $\mu_{Xj} \neq \mu_{Xj+1}$ (b) $\tau_i = (\tau_{i,j} - \mu_j) / \lambda_j = (\tau_{i,j+1} - \mu_{j+1}) / \lambda_{j+1}$ (c) $\sigma_{\tau j} \neq \sigma_{\tau j+1}$ (d) $\sigma_{\varepsilon j} \neq \sigma_{\varepsilon j+1}$

Note. In the table above X represents the observable score, τ the unobservable true score and ε the measurement error for person i on item j .

characteristics, such as mean and variance, we tend to think of these parameters in the context of a latent trait distribution as location (or precision) and scale respectively (Graham, 2006). Subsequently, the essentially tau-equivalent measurement model holds that while essentially all measurement instances for a person share the specific true score, they vary in the items' precision to measure the true score. Table 1 shows the corresponding CTT model (3), and how the essentially tau-equivalent model varies from the tau-equivalent predecessor by an additional parameter μ_i .

The assumption of a person item true score (Eq. 10 and assumption 3.b in Table 1) is that a person's specific true score, τ_i , is shifted by an item location or precision parameter, μ_i .

$$\tau_i = \tau_{i,j} - \mu_j = \tau_{i,j+1} - \mu_{j+1} \quad (10)$$

It thus follows that if items' precision parameters (for any item j , $j+1$, etc.) could be subtracted from the corresponding person specific true scores on those items ($\tau_{i,j}$, $\tau_{i,j+1}$, etc.), we should recover the person specific (overall) true score.

Congeneric

Comparing the CTT measurement models side-by-side in Table 1, we can deduce that as assumptions about an examinee's true score and measurement error become less restrictive, additional parameters are needed to accurately capture the relationship (Jöreskog, 1971). Thus, the congeneric model expands further on the essentially tau-equivalent model in that in addition to its precision parameter the congeneric measurement model introduces an item true score specific scale parameter, λ_i .

$$X_{i,j} = (\lambda_j \tau_i + \mu_j) + \varepsilon_{i,j} \quad (11)$$

The scale parameter of the congeneric model provides a mechanism by which the dispersion of the centered item true score ($\tau_{i,j} - \mu_j$), which we encountered in the essentially tau-equivalent model, can be adjusted so as to reproduce the person's true score scale. In essence assumption 4.b in Table 1 shows that an item's true score requires correction in terms of both precision and scale in order to ascertain the person's true score on the measured construct.

With the specification of the congeneric model, CTT acquired a more flexible form, and one that would lend itself naturally to modern psychometric practices. For example, Raykov (1997) demonstrated how Structural Equation Modeling (SEM) could be used to fit the congeneric measurement model to observed data (and therefore all other discussed CTT models by imposing constraints) to determine both the fit of the model, given the data, as well as to derive latent trait variance estimates for purposes of computing CTT derived reliability estimates (see Interpretational Contexts below).

CTT as a special case

Based on the focus of the afore discussed models, CTT's principal concern is one of attempting to understand the nature of the true score and how concrete assumptions about its realization and its relationship to error relate to observable data. Given the continued development of models that attempted to ascertain the representative link between observable data and the true score, it should come as no surprise that a general form, one to encompass a multitude of varying levels of measurement and distributional attributes, was derived. Much as it is the case with the generalized linear model, a generalization of the link between the latent trait true score and observable scores lead to the development of the generalized latent variable model by Skrondal and Rabe-Hesketh (2004). While their work provides conceptual extensions to the classical test theory framework, the relationship between the components of the generalized latent trait model to CTT may not be immediately clear. Arguing for the similarities between IRT and CTT, Lewis (2007) extends backwards from the generalized latent trait model and provides a discussion of how components of CTT are represented in the generalized form. In his treatment, CTT is viewed as a constrained special case of the generalized latent trait model, "with the qualification that a suitable substitute for $[\tau]$ must be found to approximate the biserial correlation", $\rho_{\tau,j}$ between the random latent true score and the item score (p. 42) in order to establish the link.

Interpretational contexts

In isolation, Eq. (1) and the true score models provide a rather disenfranchized view of CTT and their interpretation within the educational measurement domain. As noted earlier, historically CTT has been contextualized by a series of frameworks within which aspects of the theory were used to provide solutions to measurement concerns. These interpretational contexts have ranged from the very fundamental, such as reliability, to the very applied, such as maintaining assessment score interpretation across multiple test forms. Most of these contexts have merited the writing of separate books, and countless articles, dedicated to the intricacies of those respective domains. It is therefore far beyond the scope of this brief introduction to CTT to enumerate or elaborate on these applications adequately. Nevertheless, it would be remiss not to acknowledge the progressive historic applications in which CTT has been, and continues to be, a prominent keystone.

Reliability

Growing out of the works by [Pearson \(1901\)](#), [Spearman \(1904\)](#) and [Thorndike \(1907\)](#), a new direction formed in the social sciences of the early 20th century in which the variance of one variable, attribute or characteristic was systematically explained by the variance of another in what is today familiar to most as the Pearson product-moment correlation. The resulting application of correlation to multiple performance scores, believed to assess the same or common parts of a single construct, provided further statistical connectivity between observed measurement scores, as realizations from distributions, and the hypothetical person latent (unobservable) construct true score.

Today, measures of consistency are inseparable from the development and evaluation of assessment instruments. However, as psychological measurement was developing and the edicts of CTT were gaining traction among the practitioners and thought-leaders of the times, the context growing around CTT became one of reliability. This drive for CTT's application in this context is evident from [Kelley's \(1921\)](#) verbalization of "the need of a standardized procedure in measuring the reliability of the tests which are in use, and with the hope that it will encourage such study and discussion as will lead to a unified practice" (p. 377).

As [Kelley \(1916\)](#) himself would show, and [Thurstone \(1927\)](#), [Gulliksen \(1950\)](#), and [Lord and Novick \(1968\)](#) would reintroduce, the reliability of the observed scores (ρ_X) could be expressed as the squared correlation between the observed scores and latent true scores ($\rho_{X,\tau}$), a concept fundamental to CTT. This squared correlation could be expressed as the ratio of the respective variances, or the proportion of systematic variance:

$$\rho_X = \rho_{X,\tau}^2 = \sigma_\tau^2 \sigma_X^{-2} = 1 - (\sigma_\epsilon^2 \sigma_X^{-2}) \quad (12)$$

where if the true score variance, σ_τ^2 , numerator is larger than the variance of the observed score, σ_X^2 , the expected theoretical reliability will always be positive, suggesting a higher degree of consistency as the quotient nears to 1.0.

As was conveyed through the previous sections, ascertaining the standard deviation of the unobservable true score would require additional assumptions. It is herein that the initial CTT formulation intersects with the first true score model, the classically parallel measurement model, in that in the absence of knowledge of the true scores, variance of the true score can be derived as

$$\sigma_\tau^2 = \sigma_X^2 \rho_{XW} \quad (13)$$

where the estimate ρ_{XW} in [Eq. \(13\)](#), is the correlation between observed and classically parallel scores (or observed scores X and W) and is assumed to be equivalent to the reliability in [Eq. \(12\)](#) ([Lord and Novick, 1968](#), pp. 60–61).

Substituting [Eq. \(13\)](#) into [Eq. \(12\)](#), [Gulliksen \(1950\)](#) showed that the estimate of reliability becomes just the square of the correlation between observed and a classically parallel score, and is equivalent to the proportion of variance in the observed score accounted for by the variance in the classically parallel score (p. 23). For further reading see "[Reliability](#)" by Klaas Sijtsma and Julius Pfadt this volume.

Validity

CTT has been critiqued and exalted, deemed in- and out of fashion, and experienced more than one revival in which its models have been reevaluated and cast in a new perspective. Regardless of the context in which CTT is used to address psychological measurement features, at the heart of the theory resides the belief in a person specific true score, τ_i , that gives rise to the observed score. Consequently, if that belief is incorrect, or is inappropriately specified, then CTT's utility is brought into question (see "[Validity](#)" by Suzanne Lane this volume).

In the 1950s and 1960s, an ongoing deliberation about what an unobservable true score actually represents, CTT experienced a conceptual impasse. [Rust and Golombok \(2009\)](#) provided an overview of some of the principal criticisms of the theory of true scores, and discussed how subsequent conceptualizations enabled the support for CTT (pp. 35–42). Notably, the treatment of the true score as a strong statistical theoretical construct ([Novick, 1966](#)) by which computational solutions are derived seemed to prevail until interpretational requirements moved the debate to arguments for defining the true score in terms of trait or function. [Angoff \(1988\)](#) further elaborates on the practical aspects in which the true score is described via a collection of evidence to substantiate the construct with which it is associated where the "responsibility for validation falls to [...] the person willing to claim that certain inferences may be validly drawn from the test scores" (p. 29).

Many of the discussions regarding the true score definition and its relationship to CTT concentrated on substantive meaning and practical applications ([Messick, 1994](#)). However, the statistical methods associated with the diverse discussions of validity, had been surprisingly limited, frequently simplifying to a correlation between the observed score and a criterion score. [Zimmerman \(1998\)](#), in a reversal of this process, focuses on the problematic CTT assumption of uncorrelated errors and parallelism, and explains validity in the context of scores on a criterion and predictive validity of a test with respect to itself. "It should be emphasized, however, that the modified definition does not overcome all difficulties arising from correlated errors. Rather, the complications are transferred to the task of estimating validity by practical methods" (p. 248).

Generalizability

With two out of the three terms in [Eq. \(1\)](#) being unobservable, it is perhaps to be expected that as assumptions and constraints were imposed on the true score and error that questions regarding their nature became focal points of CTT. Just as was the case with

validity, in which a growing concern around the definition of the unobservable true scores became a central context for understanding CTT, so the nature of the errors became an imminent topic of concern.

While the treatment of the person specific true score in the presented measurement models in Table 1 evolves, the only discernible change regarding the error occurs between the parallel model and all others in that the random error is either equally variant across items or varies. Cronbach et al. (1963) recognized that this was problematic as it assumes a single homogenous source of error (regardless of the model), which is not only improbable but also impractical given the desire to disentangle the nature of the error affecting the true score. The concern about the inability to model the measurement error, particularly as it related to the estimations of reliability, evolved into generalizability theory (or G-theory), which “liberalizes classical theory by employing ANOVA methods that allow an investigator to untangle multiple sources of error that contribute to the undifferentiated [e] in classical theory” (Brennan, 2001, p. 3). Ultimately, G-theory provides a flexibility in evaluating measurement consistency and provides an estimated proportion in the total variance attributable to examinee facets, something not possible in CTT (see “Generalizability Theory and Applications” by Stella Kim, this volume, for an extended discussion of the framework).

Test development

All three of the previously identified CTT interpretational frameworks make important implicit connections to assessment design and development. However, one of the most effective applications of CTT has been in the realm of test development, which deserves a more direct treatment and elaboration.

From the previous delineation of the CTT models it follows that the parallel nature of tests is an important feature in CTT. Typical operationalized parallel test form development follows strict content blueprint guidelines, scoring and score reporting objectives, consistent testing experience directives and psychometric measurement model and test form statistical targets. In order to target parallel test development an additional set of assumptions in the form of domain sampling theory needs to be made. Domain sampling, in which all item samples considered for a specific blueprint attribute or student expectation are believed to be randomly equivalent, and thus item sets created from such item samples are assumed to be randomly parallel. Consequently, CTT test development implemented operational procedures aimed at preserving a CTT true score model, utilize information about item candidates (available items to be included in a test form) to deduce future item set (and total test form) statistics in order to approximate form parallelism.

The two fundamental classical item statistics used to describe the functioning of an item are item difficulty and item effectiveness. Item difficulty (p_j), for a dichotomously scored item (j), is given as the proportion of examinees who answered the item correctly out of all examinees.

$$p_j = \sum x_{ij} N^{-1} \quad (14)$$

where x_{ij} denotes the dichotomously (0—incorrect, 1—correct) scored responses for person i on item j , and N , total number examinees who received credit on the item (viz. provided a correct response for which they received 1 point).

Item effectiveness, also referred to as item discrimination, is given by the point-biserial correlation coefficient:

$$r_{xj} = ((\mu_+ - \mu_x) s_x^{-1}) \sqrt{(p_j (1 - p_j)^{-1})} \quad (15)$$

where μ_+ is the test score mean for those examinees who answered the item correctly, μ_x is the overall test score mean, s_x is the test score standard deviation, and p_j is the item difficulty in Eq. (14) (Crocker and Aligna, 2006, p. 317). Armed with the two estimates from Eqs. (14) and (15) several important CTT parallel form targets can be determined.

Consider the hypothetical item statistics in Table 2 for 10 items. Of the provided estimates, only item difficulty (first row) and item point-biserial correlation with the total score (fourth row) are computed based on the performance of an examinee population representative sample. All other estimates are derived.

Using the properties of the binomial distribution, the proportion correct can be transformed into the corresponding item variance as follows:

Table 2 Sample classical test theory item (j) analysis statistics.

Estimate		$j1$	$j2$	$j3$	$j4$	$j5$	$j6$	$j7$	$j8$	$j9$	$j10$	Sum
Proportion correct	p_j	0.73	0.70	0.63	0.56	0.54	0.50	0.43	0.36	0.32	0.30	5.07
Item variance	s_j^2	0.20	0.21	0.23	0.25	0.25	0.25	0.25	0.23	0.22	0.21	2.29
Item standard deviation	s_j	0.44	0.46	0.48	0.50	0.50	0.50	0.50	0.48	0.47	0.46	4.78
Point-biserial correlation	r_{xj}	0.25	0.26	0.30	0.39	0.44	0.43	0.38	0.37	0.35	0.31	3.48
Item “reliability index”	$s_j r_{xj}$	0.11	0.12	0.14	0.19	0.22	0.22	0.19	0.18	0.16	0.14	1.67

$$s_j^2 = p_j(1 - p_j) \quad (16)$$

where the item standard deviation (third row in [Table 2](#)) is simply the square root of the estimate in [Eq. \(16\)](#) above. It follows that the item “reliability index” ([Gulliksen, 1950](#), p. 377, [Eq. 22](#)), is the product of the item standard deviation and its corresponding point-biserial correlation.

Using these estimates the expected mean test score for a test form consisting of *any subset* of items can be taken as,

$$\mu_x = \sum p_j \quad (17)$$

with a corresponding test score standard deviation equal to the sum of all item reliability indices (products of item standard deviations and their point-biserial correlations):

$$s_x = \sum s_j r_{xj} \quad (18)$$

Thus, the 10-item set in [Table 2](#) yields an expected test score mean of 5.07,

$$\mu_x = p_1 + p_2 + \dots + p_9 + p_{10} = 0.73 + 0.70 + \dots + 0.32 + 0.30 = 5.07 \quad (19)$$

and a test score standard deviation of 1.67,

$$s_x = s_1 r_{x1} + s_2 r_{x2} + \dots + s_9 r_{x9} + s_{10} r_{x10} = 0.11 + 0.12 + \dots + 0.16 + 0.14 = 1.67 \quad (20)$$

The item estimates, or item parameters, can be further employed to determine test reliability (test item internal consistency) as demonstrated by [Kuder and Richardson \(1937; KR-20\)](#):

$$\rho_X = \left(K(K-1)^{-1} \left(1 - \sum s_j^2 s_{x-1}^2 \right) \right) \quad (21)$$

where K is the number of items on the test, $\sum s_j^2$ is the sum of all item variances and s_x^2 is the total test score variance. All of these estimates can be derived from the two CTT item analysis statistics provided in [Table 2](#) discussed earlier. In the provided example in [Table 2](#) the resulting test reliability can be computed as:

$$\rho_X = (10 \times (10-1)^{-1}) \left(1 - (2.29 \times (1.67^2)^{-1}) \right) = 1.11 \times (1 - 0.82) = 0.20 \quad (22)$$

While not a strong indication of measurement consistency, the low test reliability is to be expected given the small number of items on the test (see [Gulliksen, 1950](#), pp. 74–86). Following [Eq. \(12\)](#) we can further express the square root of the test reliability to denote the approximation of the correlation between the test score and the unobservable true score,

$$\rho_{X_t} = \sqrt{\rho_X} = \sqrt{0.20} = 0.45. \quad (23)$$

In the provided example, drawing subsequent 10-item sample sets, which meet blueprint and test assembly requirements, can be evaluated in the context of their psychometric properties. Computing the aforementioned estimates enables the practitioner to evaluate competing sets’ statistical parallelism. Using the CTT framework, we demonstrated how selecting items based on their CTT item parameters can target parallel test form score means, standard deviations and test reliability properties. These relationships between CTT item analysis statistics (from such sources as pilot studies, field testing, etc.) and test development objectives have found applications in a variety of uses, far beyond their original conceptualizations. For example, using the above demonstrated relationships, psychometricians have imbedded these estimate relationships as test development targets in automated test assembly ([Parshall et al., 2002](#), pp. 113–123), and computer-based testing ([Gibson and Weiner, 1998](#)) to name a few.

Discussion

As is the case with many scientific fields, foundational concepts echo through the ages of innovation and development, often lending themselves to new solutions via original precepts. Arguably, CTT has provided a lasting framework through which psychometrics has grown into an indispensable expertise for many industries and utilizations in assessment development, educational measurement and cognitive- and non-cognitive evaluations. While perhaps most prominently immortalized in its application to reliability estimation and test development, CTT remains relevant as the original principle of the psychological measurement framework.

With the cyclical resurgence of the CTT specific reliability, validity and generalizability contexts, CTT has undergone several modernizations. Analogous to the validity and fairness discussions of the 1950s to 1990s, as issues of inclusivity and antiracism are proactively voiced, CTT will require a transparent and honest substantiation within a just assessment approach ([Randall, 2021](#)). However, whether moving the test theory discussion from parallel test form reliability to SEM estimation of true score and error variances, or expanding the congeneric true score model to a generalized latent trait design, CTT, despite its name, continues to motivate the evolution of test theory.

References

- Angoff, W.H., 1988. Validity: an evolving concept. In: Wainer, W., Braun, H.I. (Eds.), *Test Validity*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, pp. 19–32. Ch. 2.
- Brennan, R.L., 2001. *Generalizability Theory*. Springer, New York, NY.
- Crocker, L., Algina, J., 2006. *Introduction to Classical and Modern Test Theory*. Wadsworth Thomson Learning, Mason, OH.
- Cronbach, L.J., Rajaratnam, N., Gleser, G., 1963. Theory of generalizability: a liberalization of reliability theory. *Br. J. Statistical Psychol.* 16, 137–163.
- Gibson, W.M., Weiner, J.A., 1998. Generating random parallel test forms using CTT in a computer-based environment. *J. Educ. Meas.* 35, 297–310.
- Graham, J.M., 2006. Congeneric and (essentially) tau-equivalent estimates of score reliability: what they are and how to use them. *Educ. Psychol. Meas.* 66, 930–944.
- Gulliksen, H., 1950. *Theory of Mental Tests*. Wiley & Sons, New York, NY.
- Jöreskog, K., 1971. Statistical analysis of sets of congeneric tests. *Psychometrika* 6, 109–333.
- Kelley, T.L., 1916. A simplified method for using scaled data for purposes of testing. *Sch. Soc.* 4, 34–71.
- Kelley, T.L., 1921. The reliability of test scores. *J. Educ. Res.* 3, 370–379.
- Kuder, G.F., Richardson, M.W., 1937. The theory of the estimation of test reliability. *Psychometrika* 2, 151–160.
- Lewis, C., 2007. Selected topics in classical test theory. In: Rao, C.R., Sinharay, S. (Eds.), *Handbook of Statistics*, vol. 26. Elsevier B.V, New York, NY, pp. 29–43.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.
- McDonald, R.P., 1999. *Test Theory: A Unified Treatment*. Lawrence Erlbaum Associates, Mahaw, NJ.
- Messick, S., 1994. Validity of Psychological Assessment: Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning (RR-94-45). Educational Testing Service, Princeton, NJ.
- Novick, M.R., 1966. The axioms and principal results of classical test theory. *J. Math. Psychol.* 3, 1–18.
- Parshall, C.G., Spray, J.A., Kalohn, J.C., Davy, T., 2002. *Practical Considerations in Computer-Based Testing*. Spring, New York, NY.
- Pearson, K., 1901. Mathematical contributions to the theory of evolution. VII. On the inheritance of characters not capable of exact quantitative measurement. *Philos. Trans. R. Soc. A* 195, 79–150.
- Randall, J., 2021. "Color-neutral" is not a thing: redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educ. Meas.* 40, 1–9.
- Raykov, T., 1997. Estimation of composite reliability for congeneric measures. *Appl. Psychol. Meas.* 21, 173–184.
- Rust, J., Golombok, S., 2009. *Modern Psychometrics: The Science of Psychological Assessment*, third ed. Routledge, New York, NY.
- Skrondal, A., Rabe-Hesketh, S., 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC Press, Boca Raton, FL.
- Spearman, C., 1904. The proof and measurement of association between two things. *Am. J. Psychol.* 15, 72–101.
- Suen, H.K., Yu, L., 2006. Chronic consequences of high-stakes testing? Lessons from the Chinese civil service exam. *Comp. Educ. Rev.* 50, 46–65.
- Thorndike, E.L., 1907. Empirical studies in the theory of measurement. *Arch. Psychol.* 1, 1–45.
- Thurstone, L.L., 1927. The unit of measurement in educational scales. *J. Educ. Psychol.* 18, 505–524.
- Traub, R.E., 1997. Classical test theory in historical perspective. *Educ. Meas.* 14, 8–14.
- Zimmerman, D.W., 1998. How should classical test theory have defined validity. *Soc. Indic. Res.* 45, 233–251.

Relevant website

National Council on Measurement in Education (NCME). Instructional Topics in Educational Measurement Series (ITEMS). <https://www.ncme.org/resources/items>.

Generalizability theory and applications

Stella Y. Kim^a, Jaime L. Malatesta^b, and Won-Chan Lee^c, ^a University of North Carolina at Charlotte, Charlotte, NC, United States; ^b Graduate Management Admission Council (GMAC); and ^c University of Iowa, Iowa City, IA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to generalizability theory	59
Universe of admissible observations and G studies	60
G study variance components (VCs)	60
Infinite universe of generalizability and D studies	61
Universe scores	61
D study variance components (VCs)	61
Error variances	62
Reliability-like coefficients	62
Multivariate generalizability theory	63
Practical and conceptual considerations in application of generalizability theory	64
Random versus fixed facets	64
Model specification	64
Real data illustration	64
Univariate $p \times i$ design	65
Univariate $p \times (i : h)$ design	66
Multivariate $p' \times i'$ design	68
Closing remarks	70
References	70

Introduction to generalizability theory

Measuring student achievement is an essential part of education, allowing for tracking of student progress and evaluation of teachers, schools, or educational systems. Thus, it is critical to educators to ensure quality measurement. The notion that measurement is subject to various sources of error spawned the development of psychometric theories that each takes a unique approach to quantifying measurement error in observed data. In the context of educational measurement, the word *error* does not mean a mistake that is correctable or avoidable, but it is rather an innate property of measurement. Unlike the natural sciences, most educational and psychological research centers around a trait or construct that is not directly observable. Thus, a measurement procedure designed to capture the amount of that trait a person possesses, inevitably results in fluctuating observed scores. The term “reliability” is often used to quantify the extent to which a measurement procedure yields consistent results, (i.e., observed scores) over replications of the measurement procedure.

There are two broad perspectives on true score: expected-value and latent-trait perspectives. Our focus here is on the former. From the expected-value perspectives, *true score* for a person is defined as an expected value of the person’s observed scores obtained from a hypothetically infinite number of replications of a measurement procedure, each being implemented under similar conditions. In this context, one’s observed score presumably fluctuates from one instance of the measurement procedure to the next. This fluctuation is a result of error. An error score is defined as a deviation of an observed score from a true score. In practice, the variance of error scores is quantified and used to determine the extent to which measurement results are consistent across hypothetical replications of a measurement procedure. The larger the variance of error scores (meaning one’s observed scores vary greatly from one administration to another), the less reliable measurement results are.

Of course, it is almost impossible in reality to repeat the measurement procedure over and over on a sample of examinees from a population and thus, the conceptualization of true scores, observed scores, and error requires a conceptual and hypothetical proposition. Among well-established psychometric theories, generalizability theory has been recognized as a powerful and comprehensive framework for identifying and quantifying multiple sources of error in observed scores over investigator-identified conditions of measurement. Generalizability theory breaks measurement errors into the contribution of multiple sources by decomposing the variance of the overall error into variance component associated with each source of error. In doing so, the analysis of variance (ANOVA) procedure typically is used. Often, classical test theory (CTT; Lord and Novick, 1968) is conceived as the predecessor to generalizability theory with the analytical support of ANOVA. However, Brennan (2001a) pointed out that generalizability theory is not a just simple association of the two parental components, and its power mainly comes from its conceptual framework.

Roughly speaking, generalizability theory consists of two distinct conceptual and analytical phases: (1) G (generalizability) studies and the universe of admissible observations (UAO) and (2) D (decision) studies and the universe of generalization (UG). This chapter provides some fundamental concepts of generalizability theory and its application to a set of example data. Note that the terminologies and notation used in this chapter are similar to Brennan (2001a). Equations presented in this chapter

are mostly expressed in terms of parameters, and can be easily replaced by estimators which are denoted with a hat ($\hat{\cdot}$). For a more thorough treatment of generalizability theory, readers might refer to Brennan (2001a).

Universe of admissible observations and G studies

The first phase starts with defining the UAO for a given measurement. The UAO is determined by a set of conditions of measurement called “facets”, as well as the way that these facets are associated with one another (crossed vs. nested). A facet refers to a set of conditions of measurement that constitutes a measurement procedure. Some examples of facets that may characterize a typical achievement test include items, occasions, testing modes, raters, item formats, and content domains. For example, consider a situation where a test is being newly developed to measure spelling knowledge for young children. The test developer might identify some characteristics of items that are considered appropriate for the intended purpose. Any item that possesses such presumed characteristics constitutes an admissible condition of measurement for the item facet. Then, any specific items used in a particular measurement procedure can be viewed as a sample randomly selected from the very large, or possibly infinite, universe of such items. The terms “facet” and “universe” are used for conditions of measurement (i.e., sources of error), while the terms “object of measurement” and “population” are reserved for those being assessed (i.e., persons in most educational settings).

Now suppose the investigator assumes that any person (p) in the population can take any item (i) in the universe. This is a verbal description of a “crossed” UAO structure, which is denoted $p \times i$. Once the UAO is defined, a study is conducted to collect data for estimating variance components. Such a study is called a G study. The design of a G study is identified by the investigator such that the variance components of all important facets can be estimated. When the UAO has a crossed structure, there are two possible data collection designs the investigator can use for a G study: the $p \times i$ or $i : p$ design. The $p \times i$ design involves each person taking the same set of n_i items. For the $i : p$ design, by contrast, each person responds to a different set of n_i items. The primary purpose of a G study is to estimate variance components (VCs) based on the data collected using either the crossed or nested design. As seen in this example, the G study design may or may not be the same as the structure of the UAO.

The UAO can have as many facets as the investigator believes are important to the particular measurement procedure. For example, the investigator might consider testing occasions (o) as another substantive source of measurement error, in addition to items. Then, the UAO structure will likely be $p \times i \times o$, and the G study data can be collected using the fully crossed design, $p \times i \times o$, or one of the nested designs such as $p \times (i : o)$ or $(i \times o) : p$. Generalizability theory indeed is most useful when multiple sources of error are identified and hence disentangled. Defining the UAO also depends on the size of the universe of each facet. When the population for the object of measurement and the universe of every facet are infinite, at least in theory, the model is called a random effects model. Strictly speaking, this assumption of an infinite universe is less likely to be perfectly true, but often a necessary conceptualization for most studies. When some facets have a finite universe while other facets are infinite in size, it is a mixed model. For example, suppose that each item in the universe is associated with one of a few content categories and the content-category facet (k) has a finite number of conditions in the universe. In this case, the UAO structure will be $p \times (i : k)$ and it is likely that a G study involves using all conditions in the universe for k .

G study variance components (VCs)

In generalizability theory, observed scores are represented with respect to effects of several components. Often, defining the components involved in a certain G study design can be supported by use of a Venn diagram. Venn diagrams for the two designs ($p \times i$ and $i : p$) described earlier can be found in Fig. 1. Each circle in the Venn diagram is associated with a facet, and the total number of components in any design is equal to the number of distinct areas in the Venn diagram. A distinction between the crossed and nested designs is noted by the intersection of circles. For the crossed design, for instance, two main components, p and i , form the intersection, pi . By contrast, the nested design is represented by the i circle being completely enclosed by the p circle.

Any observed score for a single data point can be presented as:

$$X_{\omega} = \mu + \sum_{\omega} v_{\alpha} \quad (1)$$

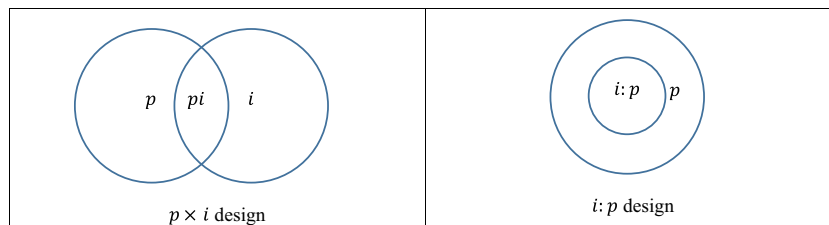


Fig. 1 Venn diagrams for $p \times i$ and $i : p$ designs.

where ω represents a set of all components in the design, μ is the grand mean in the population, and ν_α pertains to an effect associated with component α . Note that the summation is taken over all components in the design. Provided that all score effects in the model are uncorrelated and random, the variance of observed scores defined in Eq. (1) can be decomposed into independent VCs as

$$\sigma^2(X_\omega) = \sum_{\omega} \sigma^2(\alpha) \quad (2)$$

where $\sigma^2(X_\omega)$ indicates the variance of observed score X_ω . Applying these equations to the example of the crossed design ($p \times i$), any observable score for a person to a single item can be presented as

$$X_{pi} = \mu + \nu_p + \nu_i + \nu_{pi} \quad (3)$$

and consequently,

$$\sigma^2(X_{pi}) = \sigma^2(p) + \sigma^2(i) + \sigma^2(pi) \quad (4)$$

Although not denoted explicitly, ν_{pi} includes the interaction effect between p and i as well as all other errors in the observed score that are not specified by the design (i.e., residual effects). The number of VCs in a model depends on the number of facets and the G study design. The VCs can be estimated using the mean square procedure or a matrix procedure (see appendices B & C in Brennan, 2001a). Both procedures typically yield very similar results unless negative estimates of variance components are encountered. These estimated variance components are the building blocks of generalizability theory analysis and are used in the subsequent D studies.

Infinite universe of generalizability and D studies

As its name implies, generalizability theory enables investigators to generalize a decision about the objects of measurement made from a limited set of responses to a larger (hypothetically infinite) set of measurement conditions of interest. The universe to which the investigators hope to generalize results is referred to as “the universe of generalization” (UG). In general, decisions are made based on an average score over all measurement conditions (e.g., a set of items). Up to this point, however, the VCs estimated in a G study are for single scores (i.e., a response to a single item), not for average scores. To underscore a distinction between the UAO and the UG, an uppercase index (e.g., I for the item facet) is used to represent a facet for D study designs. Unlike CTT where most of the basic results are expressed in terms of the total score metric, generalizability theory employs the mean score metric. Consequently, equations in D study designs are almost always expressed in terms of observed mean scores taken over sets of conditions. For example, a person’s observed mean score over a certain number of items can be decomposed into four effects for the crossed design described previously:

$$X_{pI} = \mu + \nu_p + \nu_I + \nu_{pI} \quad (5)$$

Note that ν_I represents the effect of mean scores over a set of items, not a single item. Note also that the uppercase subscripts are used to denote the mean score metric for D studies.

Universe scores

Under CTT, defining measurement error as the discrepancy between observed scores and true scores requires conceptualization of true scores. The true score in CTT is conceived as an average or expected observed score taken over replications of parallel forms of a measurement procedure. Most basic results in CTT are obtained under the assumptions of classically parallel forms, while some reliability estimates can be derived under essentially tau equivalent or congeneric forms assumptions (Feldt and Brennan, 1989). As discussed later, generalizability theory utilizes the notion of randomly parallel forms. In generalizability theory, a close counterpart to the true score is the universe score. The universe score for a person is defined as the expected value of observed mean scores taken over all random facets in the UG, and is expressed as

$$\mu_p \equiv E_R X_{pR} \quad (6)$$

where R is a set of random facets in the UG. For example, for the $p \times i$ design, R includes the item facet and the expectation is taken over observed scores on different random samples of n'_i items. These random samples of n'_i items are referred to as randomly parallel forms. Note that the D study sample size is denoted using the prime symbol and it needs not be the same as the G study sample size (i.e., n_i). The variance of universe scores over all persons in the population is called universe score variance, which is analogous to true score variance in CTT.

D study variance components (VCs)

Because D studies are concerned with mean scores and not individual scores, the variance components estimated from the G study need adjustments to be used for decision-making for various D study considerations. The D study VCs (i.e., the variance of mean

scores) are found by dividing G study VCs (i.e., the variance of individual scores) by their corresponding sample size. Letting α be any component consisting of $\bar{\alpha}$, then the VC for $\bar{\alpha}$ is

$$\sigma^2(\bar{\alpha}) = \frac{\sigma^2(\alpha)}{\prod_{\alpha \neq p} n'_\alpha} \quad (7)$$

where $\sigma^2(\alpha)$ is the estimated VC from a G study, n'_α denotes the D study sample size associated with α , and the product in the denominator is taken over all components in $\bar{\alpha}$ except for the object of measurement (p). Note that the bar above α indicates that it is a D study effect in the mean-score metric (e.g., for the item facet, $\bar{\alpha} = I$). Application of Eq. (7) to the $p \times I$ design gives:

$$\sigma^2(I) = \frac{\sigma^2(i)}{n'_i} \text{ and } \sigma^2(pI) = \frac{\sigma^2(pi)}{n'_i} \quad (8)$$

It is worth noting that the division is not performed with respect to the object of measurement, p . Consequently, the estimated VC for universe scores $\sigma^2(p)$ remains unchanged for both G and D studies.

Error variances

Two types of error are frequently discussed in generalizability theory. One is called absolute error and is denoted as Δ . Absolute error treats a person's observed mean score as an estimated universe score, and is defined as a discrepancy between a person's observed mean score and universe score. For the $p \times I$ design, for example, absolute error for person p is expressed as,

$$\Delta_p \equiv X_{pi} - \mu_p \quad (9)$$

and its variance for the population of persons, $\sigma^2(\Delta)$, is the sum of all VCs in the D study design except the variance of universe scores:

$$\sigma^2(\Delta) = \sum_{\bar{\alpha} \neq p} \sigma^2(\bar{\alpha}) \quad (10)$$

Consequently, for the $p \times I$ design:

$$\sigma^2(\Delta) = \sigma^2(I) + \sigma^2(pI) \quad (11)$$

Taking a square root of the estimated absolute error variance results in an estimate of the absolute standard error of measurement (SEM).

The second type of error is referred to as relative error, and is denoted δ . Relative error for person p under the $p \times I$ design can be computed as

$$\delta_p \equiv (X_{pi} - \mu_I) - (\mu_p - \mu) \quad (12)$$

As seen in Eq. (12), relative error captures the difference between a deviation of a person's observed mean score X_{pi} from the item mean score, μ_I , for the population and a deviation of a person's universe score μ_p from the grand mean score μ . Due to its reflection of relative status, relative error variance is found as the sum of all VCs that involve p and at least one other index, excluding the components that do not include p in their index:

$$\sigma^2(\delta) = \sum_{p\bar{\alpha}} \sigma^2(\bar{\alpha}) \quad (13)$$

Application of Eq. (13) to the $p \times I$ design yields relative error variance of

$$\sigma^2(\delta) = \sigma^2(pI) \quad (14)$$

An estimate of relative SEM can be obtained by taking the square root of $\sigma^2(\delta)$ in Eq. (14). Relative error is more appropriate to consider when comparing the relative standing or rank ordering of persons' scores. Whereas, absolute error is more appropriately used whenever examinee scores are compared to predetermined criteria such as a cut score.

Reliability-like coefficients

Generalizability theory offers two reliability-like coefficients: a generalizability coefficient, $E\rho^2$, and an index of dependability, Φ . These two coefficients differ with respect to the type of error they incorporate. More specifically, the generalizability coefficient involves relative error variance and is computed as:

$$E\rho^2 = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\delta)} \quad (15)$$

This equation becomes identical to Cronbach's coefficient alpha under the $p \times I$ design. By contrast, the index of dependability, or a phi coefficient, incorporates absolute error variance and is computed as:

$$\Phi = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\Delta)} \quad (16)$$

The generalizability coefficient is most appropriate for norm-referenced interpretations of scores whereas the index of dependability is best suited for criterion-referenced or domain-referenced interpretations.

Multivariate generalizability theory

The development of multivariate generalizability theory (MGT) has further advanced the contribution of generalizability theory to the field of measurement (Brennan, 2001a). Imagine a measurement procedure where multiple universe scores exist for an object of measurement and a composite of the universe scores is of primary interest. One such example is a test measuring more than one skills or content areas. Typically, content categories are considered as a fixed facet while items are a random facet. In this case, a person has a universe score associated with each content category, and within each content category, there is a random effects model. A multivariate design is identified as a linkage of these multiple random-effects designs through covariance components.

In MGT, there can be more than one fixed facets. A facet is fixed in the sense that a set of measurement conditions defined by that facet remains the same over replications of a measurement procedure. Specifying the fixed facets is often guided by a table of specifications (Brennan, 2000). Due to this nature, multivariate generalizability theory is sometimes called the table of specifications model (Jarjoura and Brennan, 1982, 1983). In practice, the content areas or cognitive skills outlined in the table specifications are often described as a fixed facet, because every form of the test is expected to include items assessing the same set of content categories or cognitive skills. Then, for each level of different content categories or cognitive skills, any observable score can be represented in a similar fashion to Eq. (1) as a random effects model. Statistical computation and estimation under multivariate designs involving variance and covariance components can be facilitated by the use of matrices. The general estimation approach is similar to the univariate analysis, except now parameters are presented in matrices.

Suppose that the investigator identified two distinct types of items considered to be necessary to assess spelling knowledge: (1) writing down a word after hearing it and (2) completing partially spelled words with the appropriate missing letter(s). Hence, there are two distinct item types, and each corresponds to a distinct level of the fixed facet, item format. If the investigator were to replicate the measurement procedure, items would be randomly sampled from the two aforementioned levels of the fixed facet, item type, and each level would have its own universe of items. The item types are fixed in a sense that every form would include the same types of items over replications. This is a verbal description of a $p' \times i'$ multivariate design. A superscript filled circle associated with the person facet emphasizes that the facet is crossed with the multivariate variable (i.e., item type) whereas a superscript empty circle associated with the item facet indicates that the item facet is nested within the multivariate variable. Under the $p' \times i'$ multivariate design, covariances are only present for the object of measurement (i.e., persons) because examinees respond to both item types.

Because items are nested within item type, all remaining matrices are diagonal as follows:

$$\Sigma_p = \begin{bmatrix} \sigma_v^2(p) & \sigma_{vv'}^2(p) \\ \sigma_{vv'}^2(p) & \sigma_{v'}^2(p) \end{bmatrix} \quad (17)$$

$$\Sigma_i = \begin{bmatrix} \sigma_v^2(i) & \\ & \sigma_{v'}^2(i) \end{bmatrix} \quad (18)$$

$$\Sigma_{pi} = \begin{bmatrix} \sigma_v^2(pi) & \\ & \sigma_{v'}^2(pi) \end{bmatrix} \quad (19)$$

where the notation v and v' is used to refer to different levels of the fixed variable, item type. The elements in the D study matrices Σ_I and Σ_{pI} are obtained by dividing their corresponding element in Σ_i and Σ_{pi} , respectively, in a similar fashion to the univariate analysis. For this particular example, the matrices for universe score, relative error, and absolute error are obtained using the following formulas:

$$\Sigma_\tau = \Sigma_p \quad (20)$$

$$\Sigma_\delta = \Sigma_{pI} \quad (21)$$

$$\Sigma_\Delta = \Sigma_I + \Sigma_{pI} \quad (22)$$

Most often, researchers are interested in results for composite scores. The composite scores are defined as a weighted sum of scores over all levels of a fixed variable. In most applications, weights are determined such that composite scores reflect the relative contribution of each level to the total score in terms of the number of score points. Let w_v be a weight of level v , then the composite universe score is defined as

$$\mu_{pC} = \sum_v w_v \mu_{pv} \quad (23)$$

Composite universe score variance, relative error variance, and absolute error variance for the composite are found using the following formulas, respectively:

$$\sigma_C^2(p) = \sum_v w_v^2 \sigma_v^2(p) + \sum_{v \neq v'} w_v w_{v'} \sigma_{vv'}(p) \quad (24)$$

$$\sigma_C^2(\delta) = \sum_v w_v^2 \sigma_v^2(\delta) \quad (25)$$

$$\sigma_C^2(\Delta) = \sum_v w_v^2 \sigma_v^2(\Delta) \quad (26)$$

Finally, the multivariate generalizability coefficient and phi coefficient for composite scores, are computed by substituting $\sigma_C^2(p)$, $\sigma_C^2(\delta)$, and $\sigma_C^2(\Delta)$ into Eqs. (15) and (16), respectively.

Practical and conceptual considerations in application of generalizability theory

There are several practical and conceptual considerations that an investigator should take into account when using generalizability theory as an analysis framework. This section describes some of those considerations that many researchers might encounter in practice. Then, the following section provides several illustrative examples of how generalizability theory can be used to analyze educational data.

Random versus fixed facets

The essence of generalizability theory lies in the fact that it permits an investigator to identify and isolate facets contributing to measurement error that occurs over replications of a measurement procedure. Inclusion of a certain facet in a design relies on the investigator's identification of potential facets and the decision should be evidenced by the significance of each facet to the precision of the measurement procedure. The investigator not only determines the set of facets, but also defines the UG associated with each of the facets.

Suppose the same conditions appear in every repeated measurement procedure for a certain facet, and the investigator has no interest in generalizing results over the other conditions. Then, we can say that the UG involves a fixed facet. In other words, if there is a fixed number of conditions for the facet in the UG and the D study sample size is equal to the fixed number of conditions in the UG, the facet is fixed. This restricted UG typically leads to increased consistency among observed scores for a given individual but prohibits generalizing results to other conditions of the fixed facet. Some examples of a design that involves a fixed facet are presented later.

Model specification

There are several key questions that must be addressed by the investigator prior to application of generalizability theory to real-world data. For example, it is the investigator's task to answer questions like: What constitutes replications of a measurement procedure? Which facets are random and which are fixed? Are the facets crossed or nested? Answering these questions probably is the most challenging aspect of putting generalizability theory into practice. It is crucial that the investigator identify all potential sources of error, incorporate those major facets in the data collection design, and provide evidence about the importance/contribution of the considered facets to the precision of the measurement procedure. Sometimes, however, practical constraints (e.g., estimation complexity, limited data collection designs) may limit the number of facets that can be considered in a design, or the form of the UG. If this happens, the investigator should be cautious while interpreting results due to the increased bias in analysis invoked by the model specification.

Real data illustration

The next section provides several illustrations of how univariate generalizability theory (UGT) and MGT can be used to analyze educational data. The data used in this section was originally collected in 2015 as part of the Organization for Economic

Cooperation and Development's Program for International Student Assessment (PISA). PISA exams are administered internationally to 15 and 16 year-olds every 3 years and are aimed primarily at assessing students' ability to apply reading, math, and science skills to solve real-world problems (OECD, 2017).

In this section, data from PISA's reading assessment are used first to illustrate univariate single-facet and multi-facet designs. Later, PISA science assessment data is used to demonstrate the use of a multivariate single-facet design. Analyses in this section are based on modified PISA data and therefore, generalizations cannot be made to the operational PISA program. For example, to reduce design complexity, 5 of the 29 reading assessment items were dropped to obtain a balanced design in which each level of a facet has an equal sample size. Furthermore, PISA uses a matrix sampling and administers over 30 forms per content area. However, in this illustration, analyses are carried out using a single form and hence are not representative of the PISA population. Furthermore, PISA assessment results are rarely, if ever, presented for individual students, and instead, focus on country-level reporting. However, each of the illustrations presented next is for designs that specify students as the objects of measurement. Thus, results in this section are purely for illustration purposes.

Univariate $p \times i$ design

The $p \times i$ design is perhaps one of the most straightforward univariate single-facet designs available. In this design, persons or students (p) are the objects of measurement and respond to all items (i). The data used to illustrate this design were collected from students in three countries (i.e., Australia, Brazil, and Spain) who all took the same form of the 2015 PISA reading assessment. A total of 848 examinees were included in the analyses – 331 from Australia, 336 from Brazil, and 181 from Spain. Therefore, we are treating students from these countries as if they belong to the same general population, but an alternative approach would be to carry out analyses within each country. The reading assessment originally contained 29 items, but 5 items were removed, leaving 24 items (i.e., $n_i = 24$). Three D studies were carried out with a varying number of items – the first for a shortened test (i.e., $n'_i = 20$ items), the second for a test of the same length (i.e., $n'_i = 24$ items), and the third for a longer test (i.e., $n'_i = 28$ items). It should be noted that the term “default D study” is often used to refer to a D study that has the same sample sizes as the G study (i.e., $n_i = n'_i$). The software program, GENOVA (Version 3.1; Crick and Brennan, 2001) was used to carry out G and D study analyses. Results for the $p \times i$ design can be seen in Table 1.

VCs from the G study analysis can be seen in the first column of Table 1. These VCs represent the breakdown of the estimated total variance of single examinees on single items, as noted in Eq. (4). There is one VC for each random effect, $\hat{\sigma}^2(\alpha)$, in the design. Note that the VCs and other statistics in Table 1 are denoted with a hat to highlight the fact that they are estimates, not parameters. As seen in Table 1, the VC's for persons and items are roughly equal to each other and roughly 1/3 the magnitude of the VC for the pi interaction. Since PISA reading items are scored dichotomously, the sum of all G study VCs is always smaller than 0.25.

The same design employed in the G study was also carried over to three D studies. VCs for these D studies are representative of the population of PISA examinees and the UG, which, in this example, is made up of randomly parallel forms of the PISA reading assessment. In a D study, decisions about persons are based on their average or mean performance to all n'_i items. $\hat{\sigma}^2(I)$ can be interpreted as the variance of item mean scores, where each mean score comes from a different random sample of n'_i items from the UG. For D studies, capitalized letters are commonly used to indicate results pertaining to mean scores (Brennan, 2001a) whereas in traditional statistics a bar would be used to indicate a mean (i.e., $\bar{X}$).

To obtain $\hat{\sigma}^2(I)$ and $\hat{\sigma}^2(pI)$ for any given D study, one simply divides each G study VC by n'_i , as presented in Eq. (7). For example, $\hat{\sigma}^2(I)$ for the D study with $n_i = 20$ equals $\hat{\sigma}^2(i)/20 = 0.0512/20 = 0.0026$. Universe score variance $\hat{\sigma}^2(\tau)$ for the D studies is the same as its estimated value from the G study, $\hat{\sigma}^2(p)$. Absolute error variance, $\hat{\sigma}^2(\Delta)$ represents the error associated with using a person's actual mean score as an estimate of their universe score. In the $p \times I$ design, absolute error variance for the UG is the

Table 1 G Study and D Study Results for the $p \times I$ Design.

G study		D studies		
$\hat{\sigma}^2(\alpha)$		$n'_i = 20$	$n'_i = 24^a$	$n'_i = 28$
$\hat{\sigma}^2(p) = 0.0471$	$\hat{\sigma}^2(p)$	0.0471	0.0471	0.0471
$\hat{\sigma}^2(i) = 0.0512$	$\hat{\sigma}^2(I)$	0.0026	0.0021	0.0018
$\hat{\sigma}^2(pi) = 0.1539$	$\hat{\sigma}^2(pI)$	0.0077	0.0064	0.0055
	$\hat{\sigma}^2(\tau)$	0.0471	0.0471	0.0471
	$\hat{\sigma}^2(\delta)$	0.0077	0.0064	0.0055
	$\hat{\sigma}^2(\Delta)$	0.0103	0.0086	0.0073
	$E\hat{\rho}^2$	0.8595	0.8801	0.8954
	$\hat{\Phi}$	0.8212	0.8464	0.8653

^aIndicates the Default D study.

sum of $\hat{\sigma}^2(I)$ and $\hat{\sigma}^2(pI)$. Relative error variance in generalizability theory is often viewed as an analog to the error variance in classical test theory. Relative error variance, $\hat{\sigma}^2(\delta)$, is concerned with making decisions about a particular examinee, relative to the group of examinees to which that individual belongs. In most cases, individuals are compared to a group mean and that group mean is constant for all individuals. Hence, in the $p \times I$ design, $\hat{\sigma}^2(I)$ does not contribute to $\hat{\sigma}^2(\delta)$ and therefore $\hat{\sigma}^2(\delta)$ is equal to $\hat{\sigma}^2(pI)$. This implies that, unless, $\hat{\sigma}^2(I)$ is zero, absolute error variance is always larger than relative error variance in the $p \times I$ design.

The size of $\hat{\sigma}^2(\Delta)$ and $\hat{\sigma}^2(\delta)$ directly impacts the size of the phi coefficient ($\hat{\Phi}$) and generalizability coefficient ($E\hat{\rho}^2$), respectively. For the $p \times I$ design, the generalizability coefficient and Cronbach's alpha give the same result, both involving relative error variance. Meanwhile, the phi coefficient involves absolute error variance and is more appropriate when examinee performance is being compared to a performance criterion such as a cut score. In the $p \times I$ design, the two reliability-like coefficients are computed using Eqs. (15) and (16). The D Study trends seen in Table 1 are largely predictable based on the equations used to compute D study results. Namely, as n'_i increases, VCs for the random effects (i.e., $\hat{\sigma}^2(I)$ and $\hat{\sigma}^2(pI)$) decrease, error variances decrease, and coefficients increase. These trends can be seen more easily in Fig. 2. Also noticeable from Fig. 2 is the inverse relationship between the size of the error variance and the size of its associated coefficient.

The question of how large coefficients ought to be, is at the discretion of the practitioner or researcher and should be based on the purpose and use of the testing program. As seen in Figs. 2 and 3, the measurement precision gained by going from a 20 item test to a 24 item test is greater than that gained by going from a 24 item test to a 28 item test. Hence, as more and more items are added, measurement precision will tend to improve but at a diminishing rate. It is up to the investigator to weigh the costs and benefits associated with various test lengths.

Univariate $p \times (i : h)$ design

The univariate $p \times (i : h)$ design is similar to the $p \times i$ design except now, items are nested within random passages (h). The same PISA reading data used in the $p \times i$ illustration is also used to illustrate the $p \times (i : h)$ design. However, the PISA reading data used in the $p \times i$ illustration were not modeled faithfully because in reality, those items were nested within reading passages. The PISA data used in this illustration contains 8 passages and each has 3 items associated with it, thus resulting in a total of 24 items. The $p \times (i : h)$ design more faithfully models the table of specifications for the PISA data by incorporating the passage facet into the design. More specifically, $p \times (i : h)$ design consists of three main effects (i.e., $p, h, i : h$) and two interaction effects (i.e., $ph, pi : h$). These effects are easily identified through the use of the Venn diagram as shown in Fig. 4. Also, Fig. 4 provides a comparison between the two designs used for the same set of data.

G study VCs for the $p \times (i : h)$ design can be seen in the first column of Table 2. From the size of the VCs, it can be seen that the VC associated with passages, $\hat{\sigma}^2(h)$, is smaller than the VC associated with items within passages, $\hat{\sigma}^2(i : h)$. This implies that passages are more homogenous than the cluster of items within them.

Three D studies were carried out and each contained a total of 24 items but varied with respect to the number of passages (i.e., 12, 8, or 6) and the number of items nested within each passage (i.e., 2, 3, or 4). Across the three D studies, the total number of items was set equal to the original test length of 24 items. Results for these D studies can be seen in the columns 3 to 5 in Table 2. The size of the D study VCs is directly related to n'_i and n'_h as seen in the following equations:

$$\hat{\sigma}^2(H) = \frac{\hat{\sigma}^2(h)}{n'_h} \quad (27)$$

$$\hat{\sigma}^2(I : H) = \frac{\hat{\sigma}^2(i : h)}{n'_i n'_h} \quad (28)$$

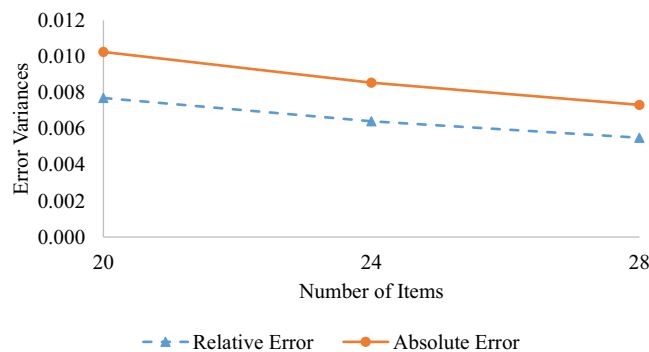


Fig. 2 Error variances for three D studies based on the $p \times I$ design.

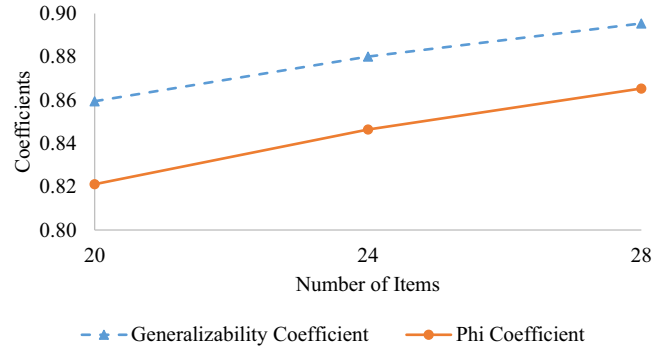


Fig. 3 Coefficients for three D studies based on the $p \times I$ design.

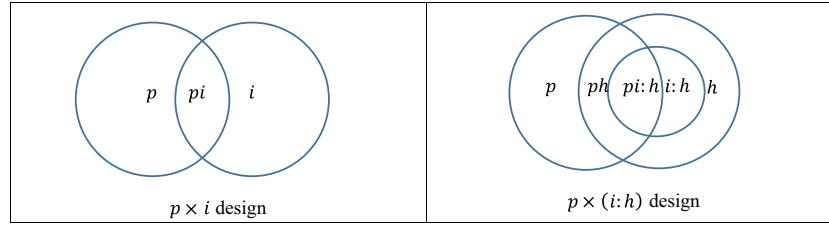


Fig. 4 Venn diagrams for $p \times i$ and $p \times (i : h)$ designs.

Table 2 D study results for the $p \times (I : H)$ design.

<i>G study</i>	<i>D studies</i>			
$\hat{\sigma}^2(\alpha)$	n'_h	12	8	6
	n'_i	2	3	4
	$n'_{i:h}$	24	24 ^a	24
$\hat{\sigma}^2(p) = 0.0464$	$\hat{\sigma}^2(p)$	0.0464	0.0464	0.0464
$\hat{\sigma}^2(h) = 0.0015$	$\hat{\sigma}^2(H)$	0.0001	0.0002	0.0003
$\hat{\sigma}^2(i : h) = 0.0498$	$\hat{\sigma}^2(I : H)$	0.0021	0.0021	0.0021
$\hat{\sigma}^2(ph) = 0.0078$	$\hat{\sigma}^2(PH)$	0.0007	0.0010	0.0013
$\hat{\sigma}^2(pi : h) = 0.1469$	$\hat{\sigma}^2(pI : H)$	0.0061	0.0061	0.0061
	$\hat{\sigma}^2(\tau)$	0.0464	0.0464	0.0464
	$\hat{\sigma}^2(\delta)$	0.0068	0.0071	0.0074
	$\hat{\sigma}^2(\Delta)$	0.0090	0.0094	0.0097
	$E\hat{\rho}^2$	0.8728	0.8675	0.8623
	$\hat{\Phi}$	0.8381	0.8323	0.8266

^aIndicates the Default D study.

$$\hat{\sigma}^2(pH) = \frac{\hat{\sigma}^2(ph)}{n'_h} \quad (29)$$

$$\hat{\sigma}^2(pI : H) = \frac{\hat{\sigma}^2(pi : h)}{n'_i n'_h} \quad (30)$$

Universe scores and their variance are defined for the UG and are not impacted by D study sample sizes. Hence, universe score variance for the G study (i.e., $\hat{\sigma}^2(p)$) carries over to any given D study.

For the $p \times (I : H)$ design, $\hat{\sigma}^2(pI : H)$ and $\hat{\sigma}^2(pH)$ contribute to $\hat{\sigma}^2(\delta)$ and $\hat{\sigma}^2(\Delta)$ consists of all the VCs contained in $\hat{\sigma}^2(\delta)$ plus $\hat{\sigma}^2(H)$ and $\hat{\sigma}^2(I : H)$. The generalizability and phi coefficients are computed using Eqs. (15) and (16), respectively. As seen in Table 2, increasing the number of passages had a greater impact on measurement precision than increasing the number of items within passages. This is because we constrained the total number of items to be 24, which thereby resulted in $\hat{\sigma}^2(pI : H)$ and $\hat{\sigma}^2(I : H)$ being equal for all D studies. As seen in Figs. 5 and 6, the scenario with the greatest number of passages (i.e., $n'_h = 12$) led to the smallest error variances and largest coefficients. These findings wouldn't necessarily be expected to hold among D studies with varying test lengths.

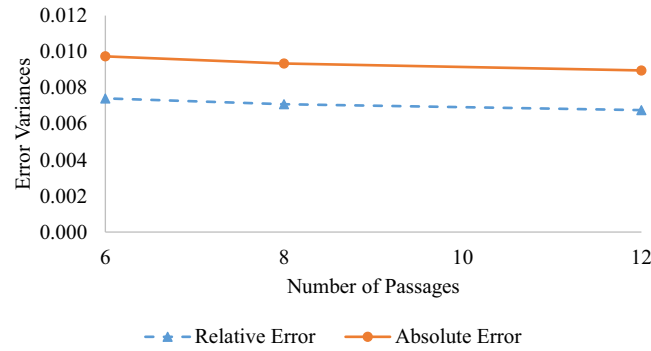


Fig. 5 Error variances for three D studies based on the $p \times (I : H)$ design.

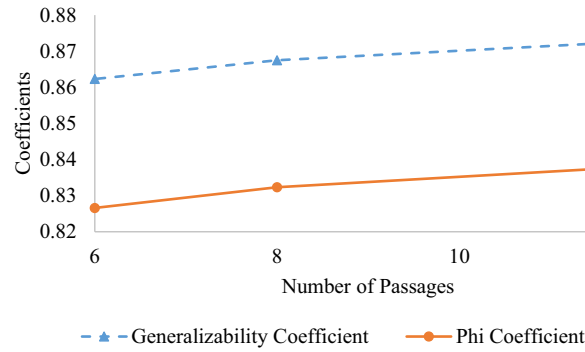


Fig. 6 Coefficients for three D studies based on the $p \times (I : H)$ design.

A comparison between D study results from the univariate $p \times i$ design and the current $p \times (i : h)$ design is useful for illustrating the potential consequences of ignoring a facet. First it is important to recall that the G study and D study designs were always identical, meaning that, for example, the $p \times i$ design was used in the G study and the $p \times I$ design was also used in the D study. If one were to compare the facets contained in the $p \times i$ design and $p \times (i : h)$ design, it becomes apparent that passages (h) are ignored in the $p \times i$ design but are taken into account for the $p \times (i : h)$ design. Since the PISA reading exam contains passages, the estimated VCs in the $p \times i$ design are likely biased and this bias will carry over to the D study results such as coefficients and error variances. The inclusion of the passage facet more accurately depicts the exam and results in slightly lower, albeit more defensible, estimates of its measurement precision. This can be seen by comparing error variances and coefficients based on the default D studies from both designs.

Multivariate $p' \times i''$ design

The multivariate $p' \times i''$ design, which is also referred to as the table of specifications model, is similar to a mixed-model univariate $p \times (i : h)$ design with h being fixed (Brennan, 2001a). However, in the $p' \times i''$ design, items are nested within a fixed multivariate variable that has several exclusive levels. Having fixed levels implies that each replication of the measurement procedure will entail the same exact set of levels. In educational contexts, the fixed multivariate variable often represents content domains or item format. Compared to its univariate counterpart, the multivariate design is computationally easier for unbalanced data in which different levels of the fixed facet can have unequal sample sizes. For the $p' \times i''$ design, each level of the fixed facet is represented by a random effects univariate $p \times i$ design. The multivariate design also has an advantage of producing covariance estimates for the linked facet (i.e., object of measurement) as well as estimates of VCs and covariance components for the distinct levels of the fixed multivariate variable.

The PISA science data is used to illustrate this multivariate design. Data in this example were collected from five countries – Australia, Brazil, Spain, Thailand, and Uruguay. Again, examinees were treated as though they belong to the same population but alternatively, analyses could have been carried out within each country. Fifty examinees from each country were retained at random, resulting in a sample size of 250. It should be noted that data manipulation was made for illustrative purposes only, because a balanced design often requires relatively reduced computational complexity and allows for more straightforward interpretations of results over an unbalanced design. The arbitrary manipulation of the data, however, does not necessarily imply lack of capacity of generalizability theory to deal with unbalanced designs.

Table 3 D study results for the multivariate $p' \times i'$ design.

<i>G study</i>	<i>D studies</i>		
	$n'_{MC} = 24$ & $n'_{FR} = 7$	$n'_{MC} = 28$ & $n'_{FR} = 5^a$	$n'_{MC} = 32$ & $n'_{FR} = 3$
$\hat{\Sigma}_p = \begin{bmatrix} 0.0559 & 0.0652 \\ 0.0652 & 0.0831 \end{bmatrix}$	$\hat{\Sigma}_p = \begin{bmatrix} 0.0559 & 0.0652 \\ 0.0652 & 0.0831 \end{bmatrix}$	$\hat{\Sigma}_p = \begin{bmatrix} 0.0559 & 0.0652 \\ 0.0652 & 0.0831 \end{bmatrix}$	$\hat{\Sigma}_p = \begin{bmatrix} 0.0559 & 0.0652 \\ 0.0652 & 0.0831 \end{bmatrix}$
$\hat{\Sigma}_i = \begin{bmatrix} 0.0531 & & \\ & 0.0784 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_i = \begin{bmatrix} 0.0022 & & \\ & 0.0112 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_i = \begin{bmatrix} 0.0019 & & \\ & 0.0157 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_i = \begin{bmatrix} 0.0017 & & \\ & 0.0261 & \\ & & \end{bmatrix}$
$\hat{\Sigma}_{pi} = \begin{bmatrix} 0.2116 & & \\ & 0.2285 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_{pi} = \begin{bmatrix} 0.0088 & & \\ & 0.0326 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_{pi} = \begin{bmatrix} 0.0076 & & \\ & 0.0457 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_{pi} = \begin{bmatrix} 0.0066 & & \\ & 0.0762 & \\ & & \end{bmatrix}$
	$\hat{\Sigma}_\delta = \begin{bmatrix} 0.0088 & & \\ & 0.0326 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_\delta = \begin{bmatrix} 0.0076 & & \\ & 0.0457 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_\delta = \begin{bmatrix} 0.0066 & & \\ & 0.0762 & \\ & & \end{bmatrix}$
	$\hat{\Sigma}_\Delta = \begin{bmatrix} 0.0110 & & \\ & 0.0438 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_\Delta = \begin{bmatrix} 0.0095 & & \\ & 0.0614 & \\ & & \end{bmatrix}$	$\hat{\Sigma}_\Delta = \begin{bmatrix} 0.0083 & & \\ & 0.1023 & \\ & & \end{bmatrix}$
	$\hat{\sigma}_C^2(\tau) = 0.0606$	$\hat{\sigma}_C^2(\tau) = 0.0589$	$\hat{\sigma}_C^2(\tau) = 0.0576$
	$\hat{\sigma}_C^2(\delta) = 0.0070$	$\hat{\sigma}_C^2(\delta) = 0.0065$	$\hat{\sigma}_C^2(\delta) = 0.0061$
	$\hat{\sigma}_C^2(\Delta) = 0.0088$	$\hat{\sigma}_C^2(\Delta) = 0.0082$	$\hat{\sigma}_C^2(\Delta) = 0.0077$
	$E\hat{p}^2 = 0.8971$	$E\hat{p}^2 = 0.9008$	$E\hat{p}^2 = 0.9044$
	$\hat{\Phi} = 0.8726$	$\hat{\Phi} = 0.8777$	$\hat{\Phi} = 0.8825$

^aIndicates the default D study.

In the $p' \times i'$ design, item format is modeled as a fixed multivariate variable with two levels: multiple choice (MC) and free response (FR). Persons are considered crossed with respect to item format (and denoted with a closed circle, p') because they respond to both MC and FR items. Items are considered nested within item format, which is denoted using an open circle, i' . Even though both levels of the fixed facet are comprised of a separate $p \times i$ design, these univariate designs are connected through persons because all persons respond to both item formats.

Using the program, *mGENOVA* (version 2.1; Brennan, 2001b), a G study was carried out using the 2015 PISA science assessment data. In total, the data contained item responses from 250 students, each of whom were administered the same form. Analyses were conducted based on 33 items, 28 of which were MC and 5 being FR items (scored 0, 1, and 2). Results from the G study can be seen in the first column in Table 3.

The diagonal matrix elements in Table 3 represent VCs and off-diagonal elements represent covariance components for v and v' . Three D studies were carried out and results for them are presented in columns 2–4 in Table 3. The maximum number-correct score for each D study was 38 points, but the proportion of MC and FR items varied. The v th diagonal element in the D study matrices $\hat{\Sigma}_i$ and $\hat{\Sigma}_{pi}$ was obtained by dividing the v th diagonal element in $\hat{\Sigma}_i$ and $\hat{\Sigma}_{pi}$, respectively, by n_{iv} . Matrices for universe score, relative error, and absolute error were obtained using Eqs. (20)–(22), respectively.

For the PISA science data, composite mean scores were computed as a simple weighted average with weights being proportional to the number of score points associated with each item format. $\hat{\sigma}_C^2(p)$, $\hat{\sigma}_C^2(\delta)$, and $\hat{\sigma}_C^2(\Delta)$ were found using Eqs. (24)–(26), respectively. The multivariate generalizability coefficient and phi coefficient for composite scores, were computed by substituting $\hat{\sigma}_C^2(\tau)$, $\hat{\sigma}_C^2(\delta)$, and $\hat{\sigma}_C^2(\Delta)$ into Eqs. (15) and (16), respectively. When the weights are proportional to item sample size as in this example, the generalizability coefficient for this design is equivalent to the stratified alpha reliability coefficient in CTT (Brennan, 2001a).

For the PISA science data, there was a clear trend in that composite score error variances decreased and composite score coefficients increased as the ratio of MC to FR items increased (Figs. 7 and 8). This finding is not surprising as measurement error for MC items tends to be smaller than that for FR items because FR items generally contain additional measurement error from raters. Relative and absolute error variances for MC and FR item formats can be seen in the diagonal elements in $\hat{\Sigma}_\delta$ and $\hat{\Sigma}_\Delta$, respectively. As long as all VCs > 0 , relative error is consistently smaller than absolute error. In addition, error variances are larger for FR item types compared to MC item types. This can be seen in rows four and five of Table 3 in the diagonal elements of $\hat{\Sigma}_\delta$ and $\hat{\Sigma}_\Delta$, respectively. Univariate coefficients specific to MC and FR items (considered separately) can be computed by substituting the diagonal elements of $\hat{\Sigma}_p$, $\hat{\Sigma}_\delta$, and $\hat{\Sigma}_\Delta$ into Eqs. (15) and (16).

Because the table of specifications plays a central role in all testing programs, it is often desirable to have estimates of measurement precision broken down according to the categories defined within that table. At the same time, there is likely to be equal importance placed on examining psychometric properties for composite scores. The multivariate $p' \times i'$ design is able to evaluate measurement precision at both grain sizes, and for unbalanced data, without much added computational complexity.

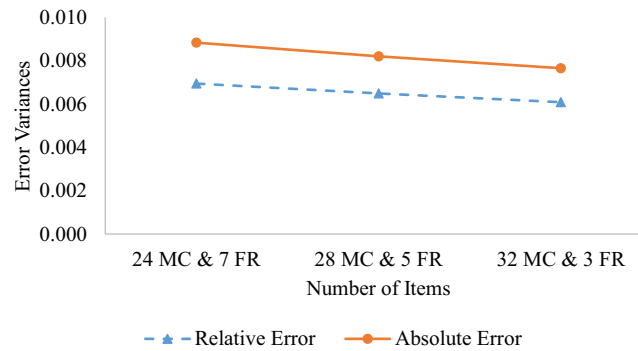


Fig. 7 Error variances for three D studies based on the $p \times i$ design.

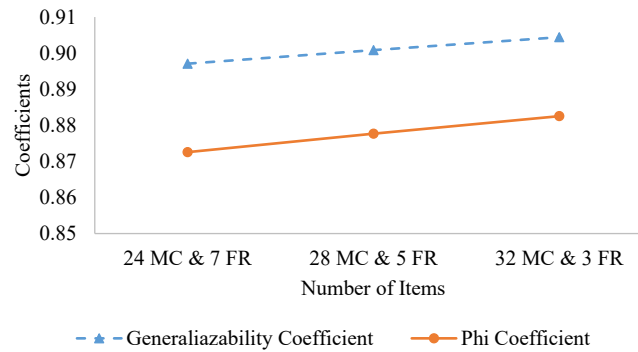


Fig. 8 Coefficients for three D studies based on the $p \times i$ design.

Closing remarks

Generalizability theory is a powerful tool to address numerous measurement issues and has been extensively applied to various real-world measurement settings. However, its powerful features come at the cost of added technical complexities, which in turn, has generally limited its use among educational researchers. This article is intended to provide a brief summary of generalizability theory and demonstrate its applicability using real world data. In addition, several important considerations such as what constitutes a random or fixed facet and the important role of the data collection design, are provided. In general, whenever generalizability theory is used, it is imperative for the researcher to model the measurement procedure in its completeness if defensible and accurate results are to be obtained.

This chapter mainly focuses on balanced univariate designs and multivariate designs. Readers might consider the use of an unbalanced univariate design if it better captures the measurement procedure under investigation. However, unbalanced designs often bring complexities for estimating random VCs. A detailed description of unbalanced designs can be found in Brennan (2001a, Chapters 7 and 8).

For readers interested in carrying out their own analyses using generalizability theory, several free computer programs exist. As presented earlier, GENOVA (Crick and Brennan, 2001) can be used for univariate balanced designs, urGENOVA (Brennan, 2001c) for univariate unbalanced designs, and mGENOVA (Brennan, 2001b) for multivariate designs. Also, an R package, gtheory (Moore, 2016), is available for R users. Alternatively, random effects G-study VCs can be estimated using computer packages designed for general statistical analyses such as Varcomp (SAS Institute Inc, 1996). Specific details on how to use of SPSS, SAS, and MATLAB programs for generalizability analyses can be found in Mushquash and O'Connor (2006).

References

- Brennan, R.L., 2000. (Mis)conceptions about generalizability theory. *Educ. Meas.* 19 (1), 5–10.
- Brennan, R.L., 2001a. *Generalizability Theory*. Springer, New York.
- Brennan, R.L., 2001b. *Manual for mGENOVA*. CASMA, University of Iowa, Iowa City, IA.
- Brennan, R.L., 2001c. *Manual for urGENOVA*. CASMA, University of Iowa, Iowa City, IA.
- Crick, J.E., Brennan, R.L., 2001. *Manual for GENOVA*. CASMA, University of Iowa, Iowa City, IA.
- Feldt, L.S., Brennan, R.L., 1989. Reliability. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education and Macmillan, New York, pp. 105–146.
- Jarjoura, D., Brennan, R.L., 1982. A variance components model for measurement procedures associated with a table of specifications. *Appl. Psychol. Meas.* 6, 161–171.

- Jarjoura, D., Brennan, R.L., 1983. Multivariate generalizability models for tests developed according to a table of specifications. In: Fyans, L.J. (Ed.), *New Directions for Testing and Measurement: Generalizability Theory*, vol. 18. Jossey-Bass, San Francisco, pp. 83–101.
- Lord, F.M., Novick, R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading MA.
- Moore, C.T., 2016. gtheory: apply generalizability theory with R. R package version 0.1.2. <https://CRAN.R-project.org/package=gtheory>.
- Mushquash, C., O'Connor, B.P., 2006. SPSS and SAS programs for generalizability theory analyses. *Behav. Res. Methods* 38, 542–547.
- OECD, 2017. PISA 2015 Technical Report. OECD, Paris. Retrieved from. https://www.oecd.org/pisa/data/2015-technical-report/PISA2015_TechRep_Final.pdf.
- SAS Institute Inc, 1996. *The SAS System for Windows Release 6.12*. SAS Institute, Cary, NC.

Item Response Theory

Terry Ackerman^a, Ye Ma^b, Mingjia Ma^a, Juliana Cerentini Pacico^a, Yibo Wang^a, Guanlan Xu^a, Tianpeng Ye^a, Jihong Zhang^a, and Mingying Zheng^a, ^a University of Iowa, Iowa City, Iowa, United States; and ^b Amazon Web Services, Seattle, WA, United States

© 2023 Elsevier Ltd. All rights reserved.

Item response theory	72
Review of IRT models	74
Dichotomous response models	74
Polytomous response models	75
Partial credit model and generalized partial credit model	75
Graded response model	76
Nominal model	77
Generalized graded unfolding model	78
Multidimensional models	79
Parameter estimation of IRT models	80
Joint maximum likelihood estimation	80
Marginal maximum likelihood estimation	81
Markov chain Monte Carlo	82
Goodness-of-fit	83
References	84

Item response theory

Item response theory (IRT) attempts to mathematically model the probability of an identified response (e.g., correctness, Likert category) based upon the interaction of a person's latent ability and characteristics (e.g., difficulty and discrimination) of an item. IRT has its origins in the work of several key researchers. They include Binet and Simon (1905) and their work on creating an intelligence scale for children. Binet developed the concept of a standardized metric of chronological age and mapped test items on to this scale using pretesting. Thurstone (1925) extended the work of Binet by developing a latent variable scale. Unlike Binet's intelligence scale which could be directly measured, Thurstone proposed an absolute measure of scaling that assumed intelligence at each age was normally distributed but differed in mean and variability. He used response data to map items onto a latent scale using the normal cumulative distribution function. A key contribution of Thurstone was that intelligence could not be directly measured but was latent and had to be inferred from the responses provided by examinees. Several researchers contributed to the concept of a response function of a latent variable (Richardson, 1936; Ferguson, 1943) and estimation of item and ability parameters (Lawley, 1943).

The seminal work can be attributed to Lord (1952) and Rasch (1960). IRT developers tried to overcome the shortfalls of Classical Test Theory (CTT) simplistic linear model, that an observed test score (X) (i.e., the sum of the item scores) is the sum of a latent true score (T) and measurement error (E), where T and E are unobservable. CTT statistics are always dependent on the group of examinees being sampled. Two other shortfalls were that there was no basis for predicting how an examinee would perform on an item and the error score variance is equal for all examinees (Lord et al., 1968; Allen and Yen, 1979).

IRT modeling makes specific assumptions about the test data being studied. The first is that the latent trait(s) being modeled define the complete latent ability space and explain examinee performance. For many tests it is assumed that only one latent ability is required (i.e., unidimensional models). For other tests, the IRT model may require several latent abilities (i.e., multidimensional models). A second assumption is that items are locally independent of each other when conditioning on the latent ability. A third assumption is that an examinees response to an item can be modeled by the item response function. This IRF which depends on the form of the IRT model expresses the probability of getting a score of X (for dichotomous item, IRF can represent the probability of $X = 1$ or $X = 0$) given each level of the latent ability, θ . This probabilistic relationship is independent of the distribution of θ in the population. Most high stakes standardized achievement and licensure tests use IRT to set the scale for reported test results. When fitting IRT models to data it is important to examine the degree to which the assumptions are met, and that the response data provide a reasonable fit the model.

Lord (1952) developed one of the first IRT models in which the interaction between an item and an examinee was expressed in terms of the probability, $p_i(\theta)$, of correctly responding to a dichotomously scored (correct/incorrect) item i , as

$$p_i(\theta) = \text{pr}\{U_i = 1 | \theta\} = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad (1)$$

where the latent ability $\theta \in (-\infty, +\infty)$, the discrimination parameter $a_i \in (0, +\infty)$, and the difficulty parameter $b_i \in (-\infty, +\infty)$. Conceptually this formula can best be understood by examining Fig. 1.

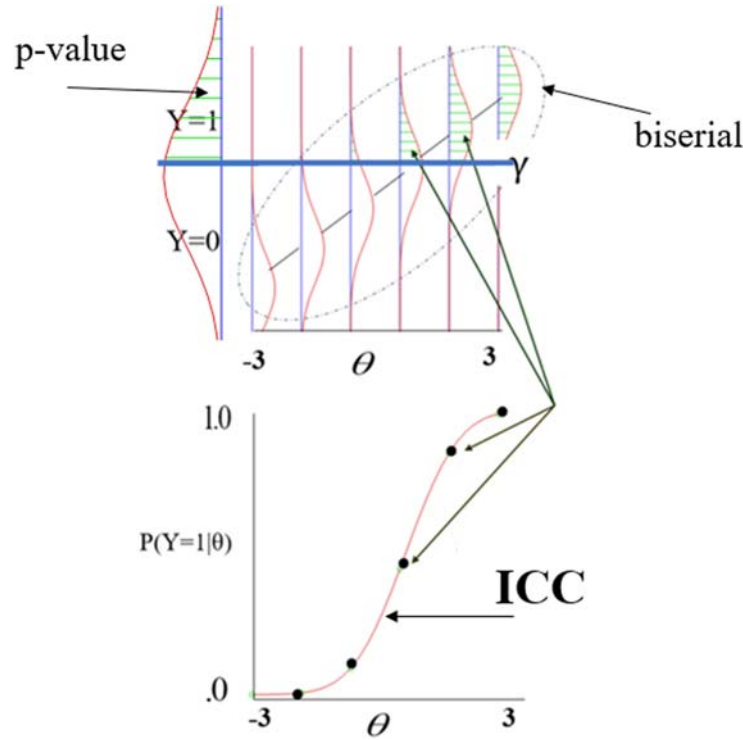


Fig. 1 An illustration of the IRT normal ogive model and the resulting ICC.

In this diagram, the top figure illustrates a dashed scatterplot of the theoretical biserial correlation between a latent ability θ , and the artificially dichotomization (where 0 = incorrect vs. 1 = correct response) of proficiency needed to respond correctly to an item. Examinees' whose ability are greater than a threshold γ , will get the item correct. Note, the CTT p -value is the proportion of the marginal distribution above γ . Following linear regression theory, it is hypothesized that at each θ level there is a normal shaped conditional distribution of proficiency, and the proportion of this distribution greater than γ , represents the probability of a correct response. The integration in the formula is determining this proportion of area. When calculated across a range of θ , one would obtain a logistic shaped curve, known as Item Characteristic Curve (ICC), illustrated in the bottom figure (Lord, 1980, p. 30–32). Richardson (1936) first developed the relationship between IRT parameters and CTT item parameters. The steepness of the ICC represents the degree of discrimination, or a -parameter in IRT parameters. This formula was modified to a more tractable logistic form for examinee j on item i , by Birnbaum (1968) and included a third item parameter, $c_i \in [0, 1]$, to account for examinee guessing.

One important contribution of IRT models is the expansion of the concept of measurement accuracy or reliability. In IRT, accuracy of θ estimation is a function of the item parameters and is determined by the Fisher information function described as,

$$I(\theta) = \frac{P_i'^2}{P_i Q_i} \quad (2)$$

where $P_i'^2$ is the squared first derivative of the IRT model with respect θ . For the 3-PL model shown in Eq. (2) the information function is given as

$$I(\theta) = \frac{a_i^2 Q_i}{P_i} \left[\frac{(P_i - c_i)^2}{(1 - c_i)^2} \right] \quad (3)$$

The standard error of estimation of θ is equal to $1/\sqrt{I(\theta)}$. A plot of an information curve and corresponding standard error is displayed in Fig. 2. Unlike the standard error of measurement calculated in CTT which is described as a constant across the ability scale, the IRT $SE(\theta)$ is a function of the item parameters and its value changes across the range of ability. Ideally a licensure or certification test would be composed of items that optimized the test information function (i.e., minimize the measurement error) at the cut score (van der Linden, 2005).

The information function for a test is equal to the sum of the individual item information functions values conditional on θ . The information function is a key element that greatly enhances the utility of IRT. Such applications directly involving the information function include test construction, computer adaptive testing, equating, and differential item functioning (DIF).

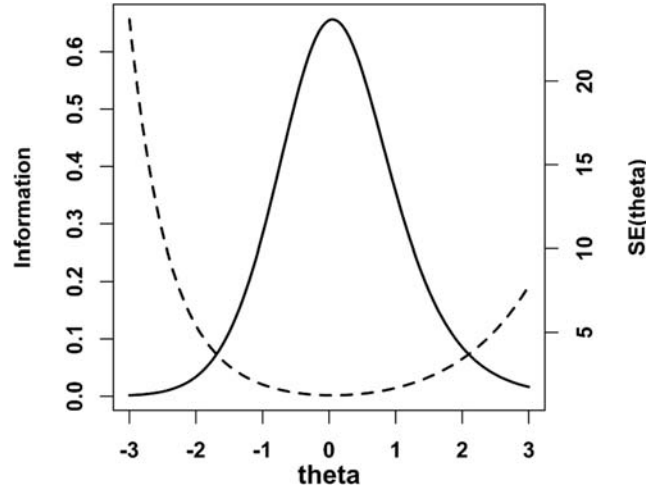


Fig. 2 Information (solid) and SE (dashed) curves for an item with $a = 1.0$, $b = 0.0$, and $c = 0.05$.

Review of IRT models

Dichotomous response models

One of the earliest IRT models was developed by [Rasch \(1960\)](#) in which he developed his model to analyze categorical data (e.g., a reading test). The model denotes a response to item I as U_i , where $U_i = 1$ for a correct response and $U_i = 0$ for an incorrect response. The probability of a correct response for an examinee having ability θ can be expressed as

$$p_i(\theta) = \Pr\{U_i = 1|\theta\} = \frac{\theta}{\theta + \delta_i} \quad (4)$$

where $\delta_i \in (-\infty, +\infty)$ is an item difficulty parameter. When transformed to a logarithmic scale the model takes the form,

$$p_i(\theta) = \Pr\{U_i = 1|\theta\} = \frac{1}{1 + \exp(-1.702(\theta - \delta_i))} \quad (5)$$

It should be noted that the 1.702 is a scaling factor to make the function approximate the normal ogive model described in [Eq. \(1\)](#). The greater a person's ability is relative to an item's difficulty the greater the probability of correct response. There is a unique distinction that sets the Rasch apart from other IRT models.

Quite often when parameters are estimated for individuals and items based on response data, their goodness-of-fit can easily be estimated. However, in the strict Rasch perspective ([Andrich, 1988](#); [Wright, 1984](#)) the goal is to find data that fit the model. Thus, items and examinees would be dropped, if they did not fit the model. Another unique feature of this model is that every examinee who has the same number correct score, will have the same ability estimate, regardless of which items they answered correctly. That is an examinee who has the five easiest items correct will get the same ability estimate as an examinee who answered the five most difficult items correctly.

The leading program for Rasch model and accompanying statistics is the program Winsteps ([Linacre, 2019](#)). The RWinsteps ([Albano and Babcock, 2019](#)) is a developed R package, which makes the use of Winsteps more convenient.

ICCs for the Rasch model never intersect. In [Fig. 3](#) the ICC's labeled A and B represent Rasch ICCs with parameters ($a = 1$, $b = -1.1$) and ($a = 1$, $b = -0.8$), respectively.

In the two-parameter logistic (2PL) model a discrimination parameter, a_i , is added to the model. This parameter allows the slopes of the ICCs to vary from item to item. The model is written as

$$p_i(\theta) = \Pr\{U_i = 1|\theta\} = \frac{1}{1 + \exp(-1.702a_i(\theta - b_i))} \quad (6)$$

The a -parameter is proportional to the slope of the ICC at the point of inflexion (i.e., $a = \text{slope}/0.425$). In the Rasch model all a 's are set to 1.0, whereas in the one-parameter logistic (1PL) model in which the a -parameter for all items is fixed to the mean a across all items. For both the 1PL/Rasch model and the 2PL model, the b value equals the θ -value at which $p_i(\theta) = 0.5$. In [Fig. 3](#), curves C and D are 2PL ICC's that have parameters ($a = 0.8$, $b = 0.0$) and ($a = 1.3$, $b = 0.3$), respectively. If the response model is logistic and the n item parameters are known, then the weighted item score s , where

$$s = \sum_i^n a_i u_i \quad (7)$$

is a sufficient statistic for estimating an examinee latent ability θ .

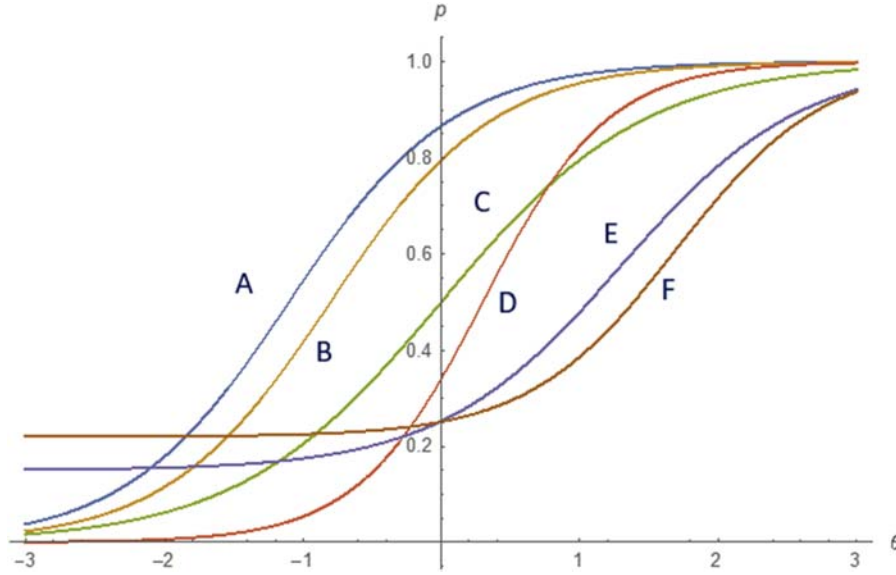


Fig. 3 ICCs for the 1PL (A,B), 2PL (C,D) and 3PL (E,F) models.

In some testing scenarios guessing is assumed to occur. In this scenario the appropriate model to describe the data may be the 3PL model, given as

$$p_i(\theta) = \Pr\{U_i = 1|\theta\} = c_i + \frac{(1 - c_i)}{1 + \exp(-1.7a_i(\theta - b_i))} \quad (8)$$

where c_i is the guessing parameter. The guessing parameter is the lower asymptote of the 3PL ICC and represents the probability of an examinee randomly guessing and getting an item correct. In this model the slope of the ICC at the point of inflexion is $0.425a(1-c)$. Note that when $c = 0$, the 3PL is equivalent to the 2PL model. The two 3PL ICC in Fig. 3 are E and F with parameters $(a = 0.9, b = 1.3, c = 0.15)$ and $(a = 1.1, b = 1.7, c = 0.22)$, respectively.

Because there is an indeterminacy for the latent θ scale, it is usually assumed that the latent abilities are distributed as standard normal distribution or $N(0,1)$ and thus the latent abilities are rescaled to $N(0,1)$ after initial estimates are obtained. Estimation programs that are used to estimate the parameters for the dichotomous models include BILOG (Mislevy and Bock, 1982), PARSCALE (Muraki and Bock, 1993) and flexMIRT (Cai, 2022).

Polytomous response models

The dichotomous response models above are used under situations when the responses are scored dichotomously. In other words, when one answers one item, the response can only be scored correctly (1) or incorrectly (0). However, depending on the items and how the items are scored, there might be more score categories other than 0 and 1. For instance, depending on the rubric used, the response can be credited partially, which leads to more than two score categories. Therefore, in contrast to dichotomous response models, to model data with more than two score categories, the polytomous response models were developed. Please note that polytomous response models with modeling data containing only two score categories is equivalent to dichotomous response models. The following section introduces the most popular polytomous response models.

Partial credit model and generalized partial credit model

The generalized partial credit model (GPCM) is an extension of the partial credit model. The partial credit model (PCM) (Masters, 1982) was designed for items in which there was increasing credit given for responses $b = 1, \dots, m_{ih}$. This model was a variation of the Rasch model requiring items to be equally discriminating. It was formulated as

$$p_{ib}(\theta) = \Pr\{U_i = b|\theta\} = \frac{\exp\left(\sum_{h=1}^b (\theta - b_{ih})\right)}{\sum_{b=1}^{m_i} \exp\left(\sum_{h=1}^b (\theta - b_{ih})\right)} \quad (9)$$

Muraki (1992) created the generalized partial credit model (GPCM) in which item discrimination parameters, α_i , were free to vary. In GPCM the probability of getting response b ($b = 1, \dots, m_{ih}$) in the h th score category for item i is:

$$p_{ib}(\theta) = \Pr\{U_i = b|\theta\} = \frac{\exp\left(\sum_{h=1}^b \alpha_i(\theta - \delta_{ih})\right)}{\sum_{b=1}^{m_i} \exp\left(\sum_{h=1}^b \alpha_i(\theta - \delta_{ih})\right)} \quad (10)$$

where $\delta_{ih} = \delta_i - d_h$, and $\alpha_i \in (0, +\infty)$, is a slope parameter for item i , δ_{ih} represents the transition location parameter between the h th category and the $h-1$ category and m_i denotes number of response categories for item i . The item-category parameter δ_{ih} can be further decomposed into two parts: the item-location parameter δ_i , and the threshold parameter d_h for each category d_h . Note, that item i containing m_i response categories has m_i-1 item-category parameters, because the first item-category parameter is arbitrarily set to 0 ($\delta_{i1} \equiv 0$). The location parameter δ_i represents the overall difficulty of item i . The threshold parameter d_h represents the relative difficulty of category h compared to other categories within an item. The summation of d_h of a same item is constrained to equal to 0 to overcome the issue of indeterminacy:

$$\sum_{h=2}^{m_i} d_h = 0 \quad (11)$$

From the equation, one can readily see that when $\theta = \delta_{ih}$, $P_{ih}(\theta) = P_{i,h-1}(\theta)$, thus, the ICC for h th category on item i will intersect with the ICC for $h-1$ st category. When $\theta > \delta_{ih}$, $P_{ih}(\theta) > P_{i,h-1}(\theta)$, meaning subjects with ability θ is more likely to fall into h th category compared to $h-1$ st category. When $\theta < \delta_{ih}$, $P_{ih}(\theta) < P_{i,h-1}(\theta)$, subjects with ability θ is more likely to fall into $h-1$ st category compared to h th category. An example of the category characteristic curves for a given set of parameters is shown in Fig. 4.

The marginal maximum likelihood estimation (MMLE) method is commonly used for GPCM estimation. Other estimation methods have also been developed, such as weighted maximum likelihood estimation (Penfield and Bergeron, 2005) and Bayesian estimation (Li and Baser, 2012).

Previous research has proposed 500 as the minimal required sample size for GPCM. Test length has little influence with respect to model-data fit (Naumenko, 2014). GPCM item parameter estimation can be conducted in the following developed software including PARSCALE (Muraki and Bock, 1993) or MULTILOG (Thissen et al., 2003) computer programs. There are also developed R packages which can conduct parameter estimation for GPCM, such as *mirt* (Chalmers, 2012) and *ltm* (Rizopoulos, 2006).

Graded response model

In contrast to GPMC/PCM, which focuses on the probability of obtaining adjacent score categories, the GR model specifies the probability of obtaining category score x_i or higher versus responding in lower category scores. Samejima (1969, 1972) developed the Graded Response Model (GRM) which includes a family of mathematical models for *ordered polytomous categories*. This family of models includes the general case (Samejima, 1969), the homogeneous (Samejima, 1973), and the heterogeneous case (Samejima, 1995). Similar to the GPCM mode, the GR model also contains a discrimination parameter for each item. The examinee responses to item i are categorized into graded $m_i + 1$ categories (m_i is the number of categories boundaries). The model can be expressed as

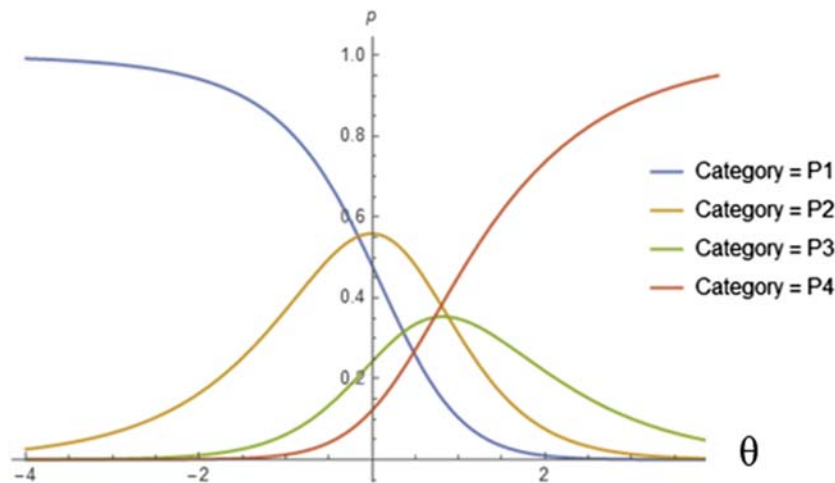


Fig. 4 GPCM category characteristic curves for an item with parameters $\alpha = 0.95$, $\delta = 0.48$, $d_1 = 0.0$, $d_2 = 0.64$, $d_3 = -0.40$ and $d_4 = -0.24$.

$$p_{ib}(\theta) = \begin{cases} 1 & \text{for } b = 1 \\ \Pr\{U_i \geq b|\theta\} & \text{for } b = 2, \dots, m_i \\ 0 & \text{for } b > m_i. \end{cases} \quad (12)$$

$P_{ib}(\theta)$ for $b = 2, \dots, m_i$ is the 2PL model expressed in Eq. (4). The response functions for these categories can be expressed as

$$p_{ib}(\theta) = p_{ib}(\theta) - p_{i(b+1)}(\theta) \quad (13)$$

When there are only two possible response categories (i.e., correct and incorrect), the GRM reduces to the 2PL model.

GRM has been used in various assessment and test contexts, including physician clinical competence assessment (de Ayala et al., 1992), questions for psychological/non-cognitive assessment (Lubbe and Schuster, 2019), educational writing tests, classroom assessment, personality inventories and computerized adaptive tests. GRM has always attracted researchers' attention. For example, recently Falk (2020) proposed an extension of GRM called Monotonic Polynomial Graded Response Model, which results in more flexible category response functions.

The GRM can be estimated using various developed standard software such as MULTILOG (Thissen et al., 2003) and PARSCALE (Muraki and Bock, 1993). DeMars (2002) and Yavuz and Hambleton (2017) compared different kinds of software to estimate the GRM.

Nominal model

The nominal response model was developed by Bock (1972) for polytomous items which have unordered response categories. It has been used for situations as: (1) item option analysis for multiple-choice items, (2) response option analysis for a survey or attitude instrument, and (3) rater judgments, etc. (De Ayala, 2013). The model can be written as

$$p_{ib}(\theta) = \Pr(U_i = b|\theta) = \frac{\exp(\alpha_{ib}(\theta - \gamma_{ib}))}{\sum_{b=1}^{m_i} \exp(\alpha_{ib}\theta - \gamma_{ib})} \quad (14)$$

where $P_{ib}(\theta)$ is the probability of choosing the response category b_{ib} , $b = 1, \dots, m_i$, and the slope parameter, α_{ib} , and the intercept parameter γ_{ib} have the same interpretation as the discrimination and difficulty parameters in the 2PL model. The model is not identified in a sense that the addition of any constant to either all of α_{ib} 's or all of γ_{ib} 's yield different parameter sets but the same value of $p_{ib}(\theta)$. To rectify this problem, identification constraints are imposed such that $\alpha_{i0} = \gamma_{i0} = 0$. Despite an item having a nominal format, the values for the intercept parameter do imply and order. An example of the category response curves for a given set of parameters is illustrated in Fig. 5.

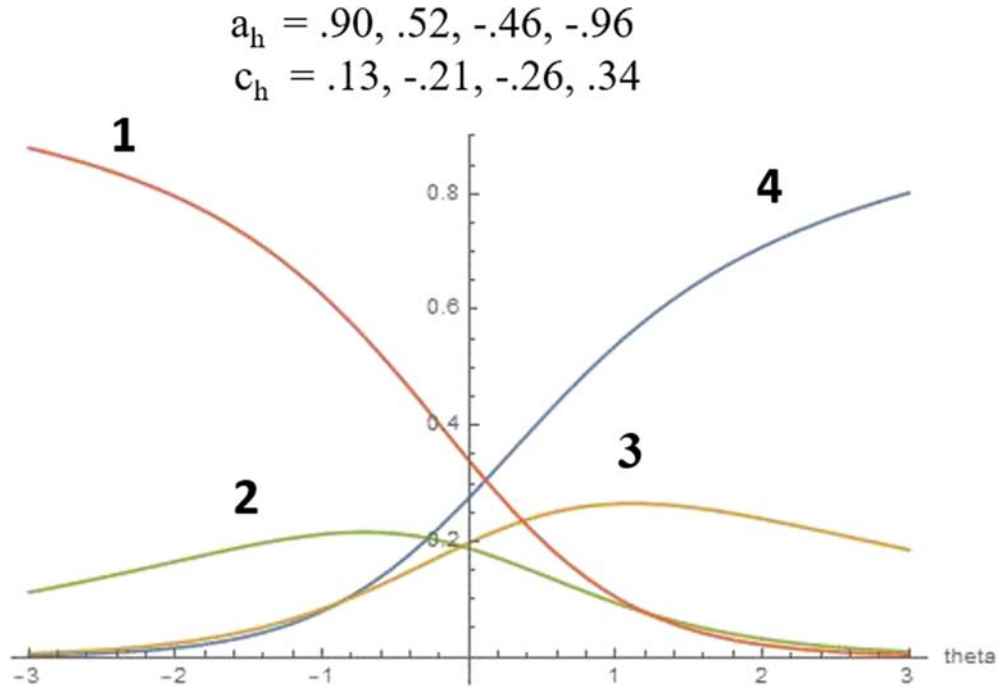


Fig. 5 Category characteristic curves for a nominal model.

Marginal maximum likelihood estimation of both item parameters and latent traits are often used for Nominal model's parameter estimation, which is described in detail in Baker and Kim (2004). The parameter estimation can also be conducted using computer software such as PARSCALE (Muraki and Bock, 2003), MULTILOG (Thissen et al., 2003), FlexMIRT (Cai, 2022), or R packages including "mirt" (Chalmers, 2012). De Ayala and Sava-Bolesta (1999) suggested that a ratio of the sample size to number of parameters (SSR) as 10:1 can yield reasonable accurate item parameter estimation when the latent trait is normally distributed. However, DeMars (2002) found that a key in estimation is the ratio of sample size to number of categories, not SSR.

Generalized graded unfolding model

Like the nominal IRT model, generalized graded unfolding model (GGUM) (Roberts et al., 1998), was also developed for self-report attitude questionnaires that measures subjects' disagreement-agreement to attitude statement. What makes attitude questionnaire special to educational data is that subjects with opposite attitudes can give same responses. For a same statement, one subject may think it is too conservative and disagree with it; another subject may think it is too liberal and disagree with it. GGUM model is designed to handle this special condition. In GGUM, data can be either binary or polytomous (e.g., a Likert scale).

Let a_i and δ_i denote the discrimination and the location on attitude continuum for statement i . θ_j denotes the location on the attitude continuum for person j . Let τ_{ik} denotes threshold parameters of category k for statement i , and Z_i denotes the observed response to statement i . The probability of person j falls in z th observed category on statement i is given by the following equation:

$$p_{iz}(\theta_j) = \Pr(Z_i = z | \theta_j) = \frac{\exp\left\{a_i \left[z(\theta_j - \delta_i) - \sum_{k=0}^z \tau_{ik}\right]\right\} + \exp\left\{a_i \left[(M - z)(\theta_j - \delta_i) - \sum_{k=0}^z \tau_{ik}\right]\right\}}{\sum_{w=0}^C \left\{ \exp\left\{a_i \left[w(\theta_j - \delta_i) - \sum_{k=0}^w \tau_{ik}\right]\right\} + \exp\left\{a_i \left[(M - w)(\theta_j - \delta_i) - \sum_{k=0}^w \tau_{ik}\right]\right\} \right\}} \quad (15)$$

where: C is the number of observable response categories minus 1, and M is calculated by $2C + 1$. The response z , can take any integer values range from $[0, C]$ and $z = 0$ corresponds to strongest level of disagreement and $z = C$ refers to strongest level of agreement.

The latent attitude continuum of the GGUM model is on the $\theta - \delta$ scale, instead of the traditional θ scale, to contain both the direction and magnitude of the distances between subject's attitude location and statement's location. Since subjects on opposite location have same probability to choose same category, the shape of ICC curve of GGUM model is symmetrical for each category.

GGUM is an extension of GPCM. Parameters of GGUM also has similar constrains as in the GPCM. Statement discrimination parameters are constrained to be all positive ($a_i > 0$); threshold τ_{i0} is constrained to be 0 ($\tau_{i0} \equiv 0$), and

$$\sum_{k=0}^C \tau_{ik} = 0 \quad (16)$$

The GGUM equation is a reparameterization of two GPCM models, one for the lower part of the attitude continuum, the other for its symmetric opposite counterpart representing the upper part of the attitude continuum. Specifically, the separate subjective response curves, SRC, are combined to form the Objective Response Curves ORC curves for one category are illustrated in Fig. 6 and the ORC for a complete set of categories is shown in Fig. 7.

Like the GPCM, the parameters in GGUM are estimated using the marginal maximum likelihood estimation. Statement parameters a_i , δ_i , and τ_{ik} are first solved via the iterative EM algorithm. Subject attitude parameter θ_j is then estimated based on a_i , δ_i , and τ_{ik} .

The minimal required sample size for GGUM is 500. Simulation studies have shown that a sample size of 750 is able to obtain reasonably well estimates for all parameters. Test length and its interaction with sample size have relatively less influence on

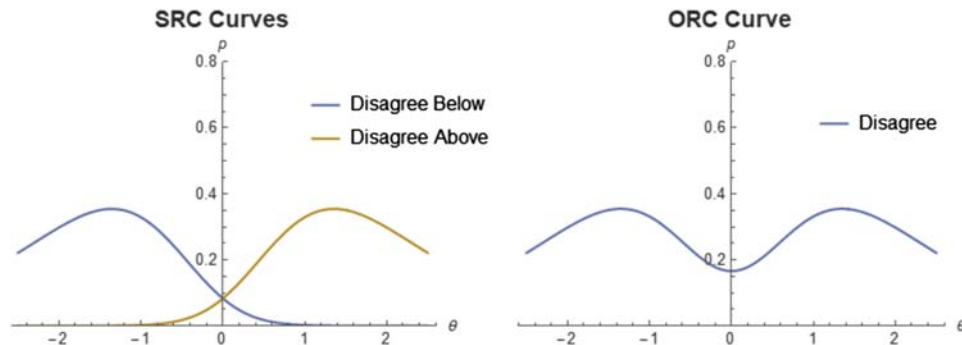


Fig. 6 GGUM Subject Response Category, SRC, curves (left) and Objective Response Category, ORC, curves (right) for a disagree category.

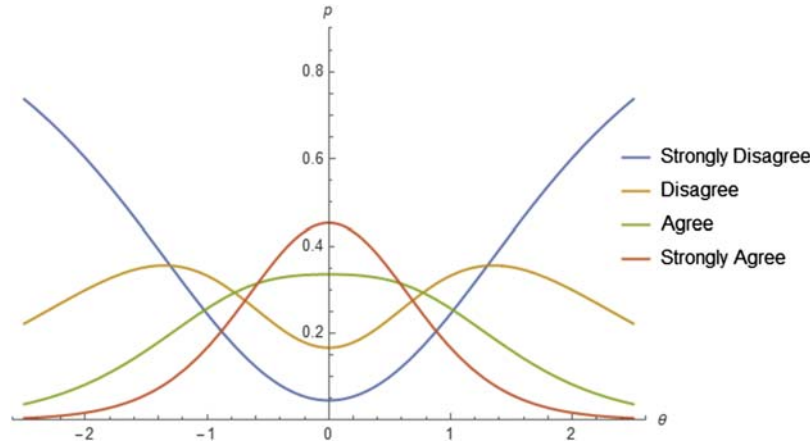


Fig. 7 GGUM category characteristic curves for an item with parameters $\tau_0 = 0$; $\tau_1 = -1.3$; $\tau_2 = -0.7$; $\tau_3 = -0.3$; $\tau_4 = 0.0$; $\tau_5 = 0.3$; $\tau_6 = 0.7$; $\tau_7 = 1.3$; $a = 0.95$, $b = 0.48$, $d_1 = 0.0$, $d_2 = 0.64$, $d_3 = -0.40$ and $d_4 = -0.24$.

parameter estimates (Roberts et al., 1998). The GGUM parameters can be estimated using the computer program GGUM2004 (Roberts et al., 2006). Wang et al. (2015) developed a MCMC GGUM program that uses Markov chain Monte Carlo (MCMC) approach. Recently, an R package named “GGUM” is also available (Tendeiro and Castro-Alvarez, 2019).

Multidimensional models

A two-dimensional extension of the 2-PL unidimensional model for dichotomously scored items multidimensional compensatory (M2PL) model (Reckase, 2009), which can be expressed as

$$p_i(\theta_1, \theta_2) = \Pr(u_i = 1 | \theta_{1j}, \theta_{2j}) = \frac{1.0}{1.0 + e^{-(a_{1i}\theta_{1j} + a_{2i}\theta_{2j} + d_i)}} \quad (17)$$

where a_{1i} and a_{2i} represent the discrimination parameters for item i on dimension one and two, respectively; and, θ_{1j} and θ_{2j} denote the latent abilities for subject j . In this model, item difficulty for each dimension is indeterminate, thus there is only one overall difficulty parameter for item i , d_i . Similarly, a guessing parameter, c_i , could be added to this model (see Eq. 6), but affects only the overall probability and would not be broken down for each dimension. Thus, if this model were extended to n -dimensions, there would be an a_{ni} and θ_{nj} parameter for each dimension, but only one d_i .

This model is described as “compensatory” because the abilities weighted by their an item’s respective discrimination parameters are additive in the logit. Thus, being “low” on one ability can be compensated by being “high” on the other ability. The degree of compensation depends on the a -parameters and is maximal when $a_{1i} = a_{2i}$.

Another way to express the interaction between the two traits is in a noncompensatory fashion as shown in the noncompensatory MIRT (Reckase, 2009) given as

$$p_i(\theta_1, \theta_2) = \Pr(u_i = 1 | \theta_{1j}, \theta_{2j}) = \left[\frac{1.0}{1.0 + e^{-1.7(a_{1i}\theta_{1j} - b_{1i})}} \right] \left[\frac{1.0}{1.0 + e^{-1.7(a_{2i}\theta_{2j} - b_{2i})}} \right] \quad (18)$$

In this model, for a given item i each dimension has discrimination parameters, a_{1i} and a_{2i} , as well as difficulty parameters, b_{1i} and b_{2i} . Also θ_{1j} and θ_{2j} denote the latent abilities on the two dimensions for subject j . This model is essentially the product two 2PL unidimensional IRT models, one for each dimension. The multiplicative nature of this model also implies that being “low” on one dimension cannot be compensated by being “high” on the other dimension. That is, the overall probability of correct response is never greater than the largest probability of the two dimensions. Specifically, if the first dimension (θ_1) p -component is 0.20 and the second dimensional p -component (θ_2) was 1.0, the over all probability of correct response would only be 0.20.

Researchers have developed numerous ways to illustrate and represent items, item and test surfaces and information in a multidimensional context (Ackerman, 1994, 1996; Reckase, 2009; Ackerman, Ma & Ip, 2019). Examples of compensatory and non-compensatory surfaces and contour plots are illustrated in Fig. 8. Many of the previously reviewed models do have multidimensional extensions. However, MIRT models are not currently being used in high stakes testing or licensure and certification testing. MIRT model parameters can be estimated using such programs as NOHARM (Fraser and McDonald, 1988), TESTFACT (Wilson et al., 1984), and flexMIRT (Cai, 2022).

One variation of MIRT models is the diagnostic classification model (DCM). These models share some of the same scaling intent as MIRT, but with more fine-grained, dichotomous attributes rather than continuous latent scales, dichotomous “loadings” (i.e., a Q-matrix), and no explicit covariance structures that relate the latent attributes to one another. The Q-matrix is the quintessential component in any DCMs because it represents an empirical hypothesis about the attribute structure that can be tested. DCMs are

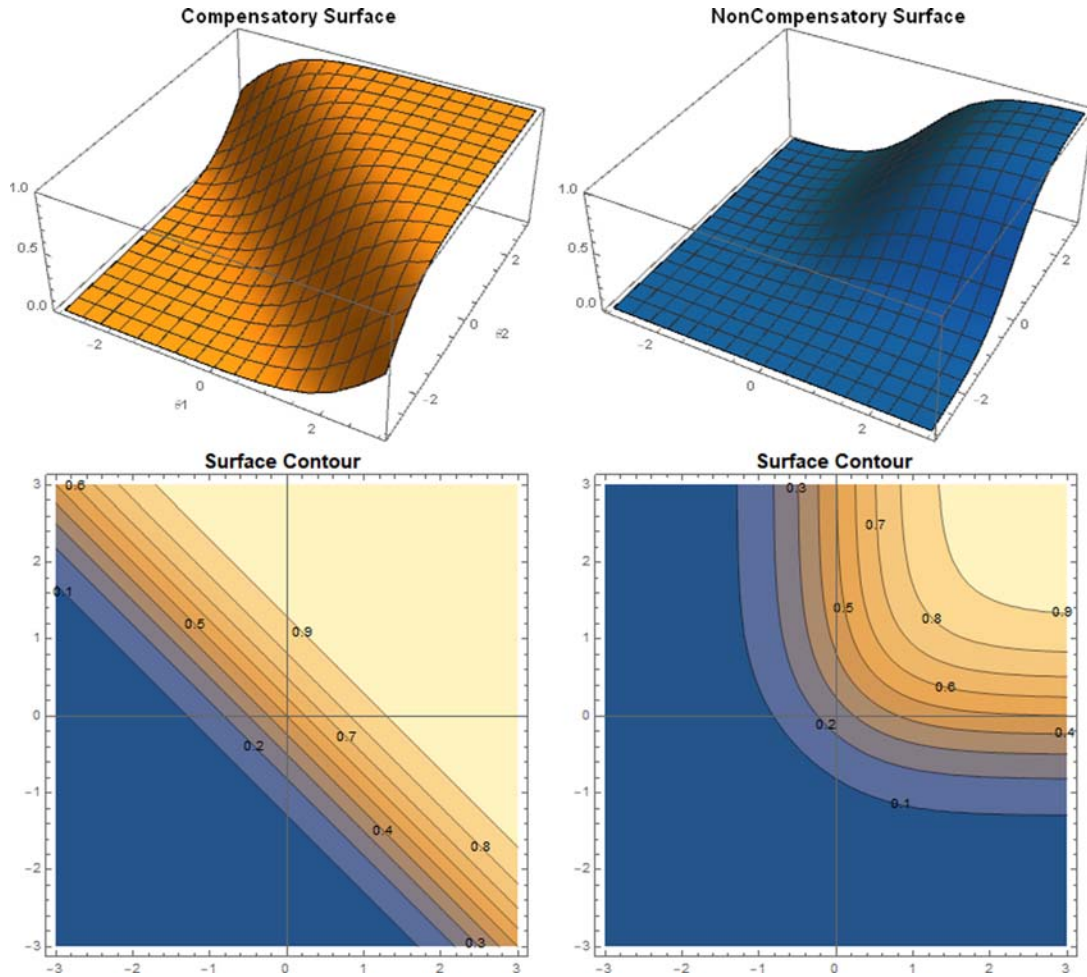


Fig. 8 Compensatory ($a_1 = 1$, $a_2 = 1$, $d = 0$) and noncompensatory ($a_1 = 1.0$, $a_2 = 1.0$, $b_1 = 0.0$, $b_2 = 0.0$) surfaces and corresponding contour plots.

also considered special case of constrained latent class models (LCMs), which are developed to provide more accurate and informative mastery profile for respondents at a fine diagnostic grain size. Two seminal sources detailing DCM models, their use and their estimation are [Rupp et al. \(2010\)](#), and [von Davier and Lee \(2019\)](#).

Parameter estimation of IRT models

The item response model does not allow for a closed form expression to estimate either the item parameters or the latent ability. There are, however, several estimation procedures to get around this limitation involving the likelihood function $L(u|\theta, a, b, c)$. These include joint maximum likelihood estimation, marginal maximum likelihood estimation, and Markov chain Monte Carlo estimation.

Joint maximum likelihood estimation

One of the major problems in estimating IRT item and ability parameters is that there is no closed form expression to estimate each parameter. One proposed approach proposed by [Birnbaum \(1968\)](#) was a two-stage joint maximum likelihood estimation (JMLE) procedure. Fitting a 3PL model to a test having n items and N examinees requires the estimation of $3n + N$ parameters. In each stage the estimation is centered around the likelihood function given as

$$\ln L(u|\theta, a, b, c) = \sum_{j=1}^N \sum_{i=1}^n [u_{ij} \ln p_{ij} + (1 - u_{ij}) \ln q_{ij}] \quad (19)$$

where u_{ij} for item i by examinee j is the scored 0/1 response and p_{ij} is the 3PL model shown in Eq. (6). In the first stage, the item parameters are estimated assuming the examinee abilities are known. Usually, a standardized raw test score would serve as an initial ability estimate for each examinee. In the second stage, the examinee abilities are estimated assuming the item parameters are known. The item parameter estimates used in the second stage are the estimates obtained from the first stage. In each stage Newton-Raphson iteration is used to maximize the log of the likelihood function. The two-stage process would iteratively run and terminate when the change in each parameter estimate from successive stages is less than a given convergence criterion (e.g., 0.01).

A graphical example of a second stage estimation for an examinee has a item response vector of [10110] for five items whose Items parameters were $\hat{a} = [0.94, 1.4, 0.88, 1.0, 0.9]$, $\hat{b} = [0.94, 1.4, 0.88, 1.0, 0.9]$, $\hat{c} = [0.0, 0.0, 0.0, 0.0, 0.0]$. The graph of $\ln L$ is shown in Fig. 9 along with the four Newton Raphson iteration approximations. The initial estimate was $\hat{\theta} = 0$. It can be seen that convergence to the maximum likelihood θ -estimate takes only a few iterations before the difference between successive iterations becomes negligible, with the final estimate being $\hat{\theta} = 0.74534$.

Note this process can be carried out for multidimensional models below. In Figs. 10 and 12 is a graphical example of three Newton-Raphson iterations for a two-dimensional model for an examinee with a response vector of [00100] and the estimated two-dimensional discrimination parameters are $\hat{a} = [\{0.94, 0.0\}, \{0.0, 1.4\}, \{0.88, 1.0\}, \{1.0, 0.9\}, \{0.08, 1.5\}]$ and the vector of difficulty parameters are $\hat{d} = [1.2, 0.8, -0.4, 0.2, 0.9]$. The final MLE (θ_1, θ_2) ability estimate $(-0.63, -1.28)$ is cross validated using the contour plot of the log likelihood surface.

Several researchers including Hambleton and Swaminathan (1985), Baker and Kim (2004), Finch and French (2019) have detailed joint maximum likelihood estimation.

Marginal maximum likelihood estimation

Marginal maximum likelihood estimation (MMLE) which is a technique used for item parameter estimation. In the MMLE technique it is assumed that examinees represent a sample from the population in which the latent ability is distributed according to a known density function (e.g., $N(0,1)$). Bock and Lieberman (1970) provided a solution in which they integrated over the ability distribution which removed the ability parameters from the likelihood function allowing for the item parameters to be estimated in the marginal distribution. A further improvement on the MMLE was the EM algorithm developed by Bock and Aitkin (1981). The EM algorithm is an iterative procedure that enables the maximum likelihood estimation of the item parameters by maximizing the formed marginal likelihood function (Thissen, 1982; De Ayala, 2013; van der Linden, 2018). One advantage of MMLE is that because the persons are integrated out from the estimation process, increasing sample size does not affect the estimation process as it would in using JMLE. Person parameters, θ 's, are usually calibrated using Newton-Raphson iteration after the item parameter estimation results achieve the satisfactory of the model fit.

Bayesian priors are frequently employed in the estimation of IRT item parameters to constrain the range of the parameter estimates as well. Common priors include a normal distribution for both ability and b parameter estimates, lognormal distribution for a -parameter estimates, and beta distribution for c -parameters estimates. Examples of these priors are illustrated in Fig. 11.

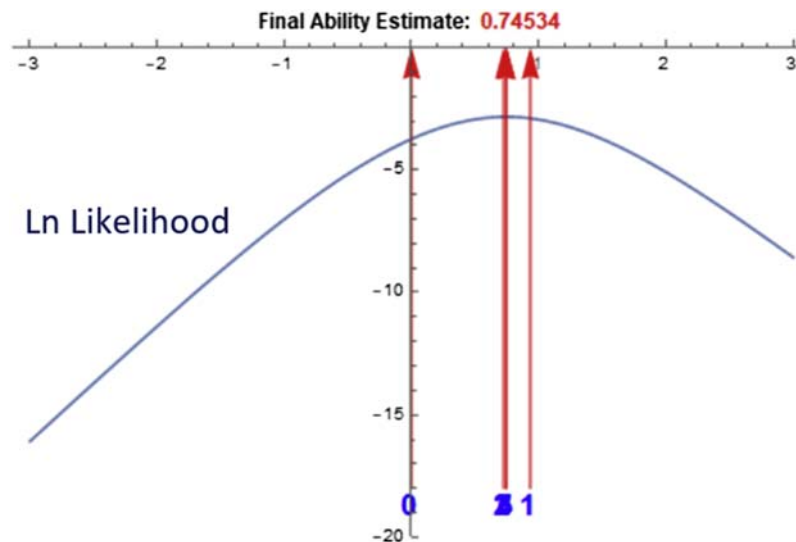


Fig. 9 An illustration of Newton-Raphson iteration example for a five-item test.

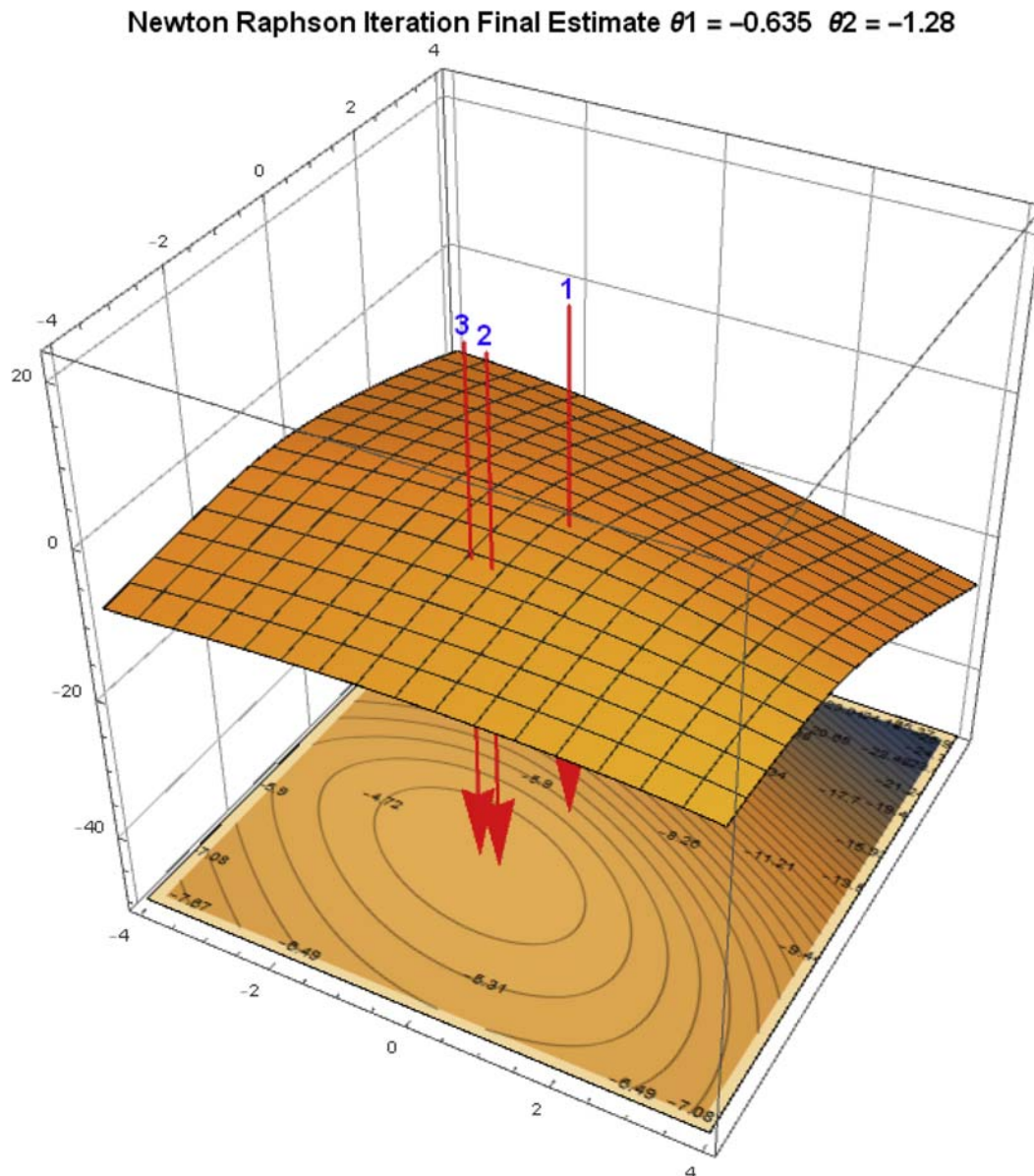


Fig. 10 An illustration of two-dimensional compensatory model Newton-Raphson MLE estimation.

An example of an application of using a normal prior is illustrated in Fig. 12. In this figure, on the left the likelihood MLE would be $+$. Using a $N(0,1)$ prior we can create a posterior distribution, which is the likelihood $\times$ prior, and then estimate the MLE from the posterior (figure on the left).

Markov chain Monte Carlo

Another approach to IRT model parameter estimation is Markov chain Monte Carlo (MCMC) sampling. MCMC has been successfully applied in various IRT and IRT-related settings. For instance, [Fox \(2010\)](#) extended the data-augmentation-based MCMC for normal ogive models to hierarchical IRT models. Other IRT related models using MCMC includes testlet models ([Bradlow et al., 1999](#)), models for test compromise ([Segall, 2002](#); [Shu, 2010](#)). There are numerous existing standard software which can conduct MCMC algorithm including WinBUGS ([Lunn et al., 2009](#)), OpenBUGS ([Spiegelhalter et al., 2012](#)), [Stan Development Team \(2019\)](#) and JAGS ([Plummer, 2012a,b](#)).

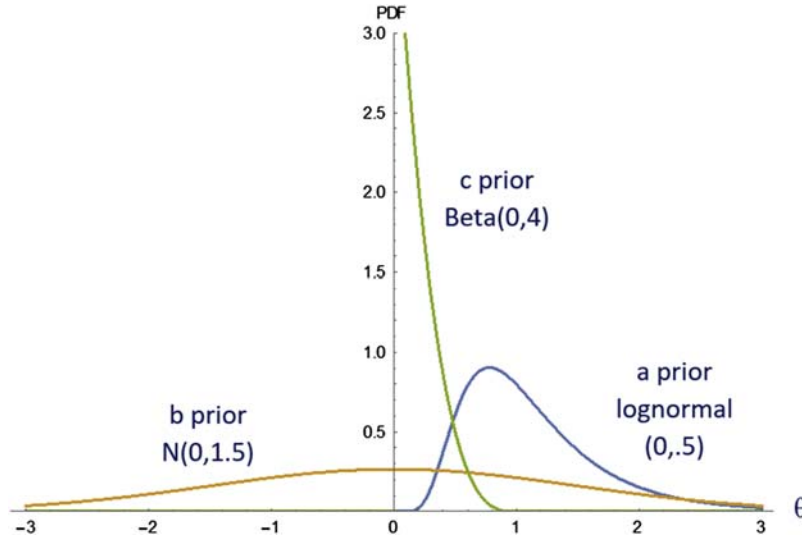


Fig. 11 Example prior distributions used for dichotomous IRT models.

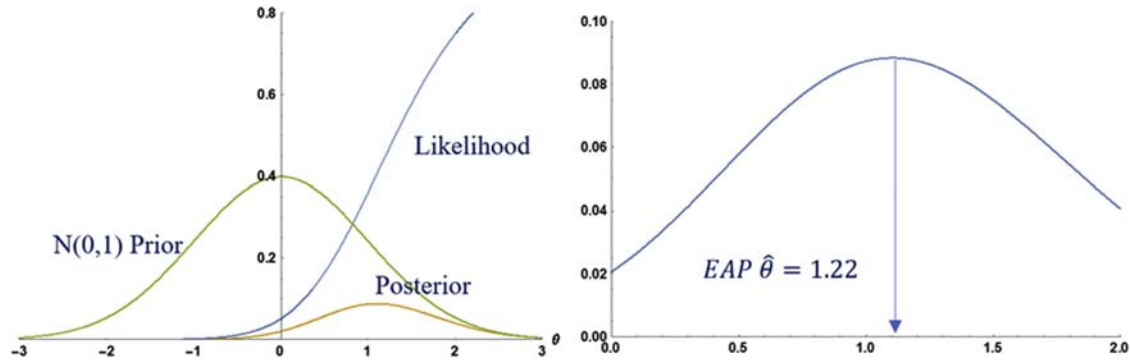


Fig. 12 Graphical illustration of Bayesian ability estimation.

Goodness-of-fit

It is important to note that no model will fit a given set of response data perfectly. To assess how closely estimated parameters fit both at the item and test level there are standard procedures that can be performed. For example, fit statistics designed specifically for the Rasch model (Wright and Masters, 1983; Linacre, 2003) are infit and outfit. Infit is inlier oriented and is sensitive to the pattern of responses to item targeted on the examinee. Outfit is outlier-sensitive and more sensitive to responses to items₁ with difficulty far from the person.

Another approach is to examine the accuracy of model predictions versus what we observed in the observed data. At the item level we can examine residuals for item j . That is, dividing the θ -scale into k -intervals (usually where the number of subjects in the interval is greater than 5) and determining the difference between observed proportion correct $p_i^*(\bar{\theta}_i)$, and the predicted (expected) probability, $p_i(\bar{\theta}_i)$, based on estimated parameters. The concept of item model misfit is illustrated in Fig. 13.

One commonly used approach to assess the fit of an item is Yen's Q_1 statistic, Yen (1984). This statistic is computed for an item i by dividing the θ -scale from $[-4,4]$ into K cells, where the number of examinees in each cell >5 and then comparing the observed proportion correct with the expected proportion correct. Specifically,

$$Q_{1i} = \sum_{j=1}^K \frac{N_j (O_{ij} - E_{ij})^2}{E_{ij} (1 - E_{ij})} \quad (20)$$

where: N_j is the number of examinees in cell j , O_{ij} is the observed proportion of examinees in cell j that respond correctly to item i , and E_{ij} is $p_i(\bar{\theta})$, the probability of correct response (e.g., Eq. 6) given the estimated item parameters and $\bar{\theta}$ is the mean or median $\hat{\theta}$ of the N_j examinees. Q_1 is distributed as a χ^2 statistic with K minus the number of estimated item parameters.

Fit statistics and graphical measures of fit are provided by such programs as BILOGMG (Mislevy and Bock, 1990). Approaches for conducting goodness-of-fit investigations for checking model assumptions are outlined in Hambleton and Swaminathan (1985).

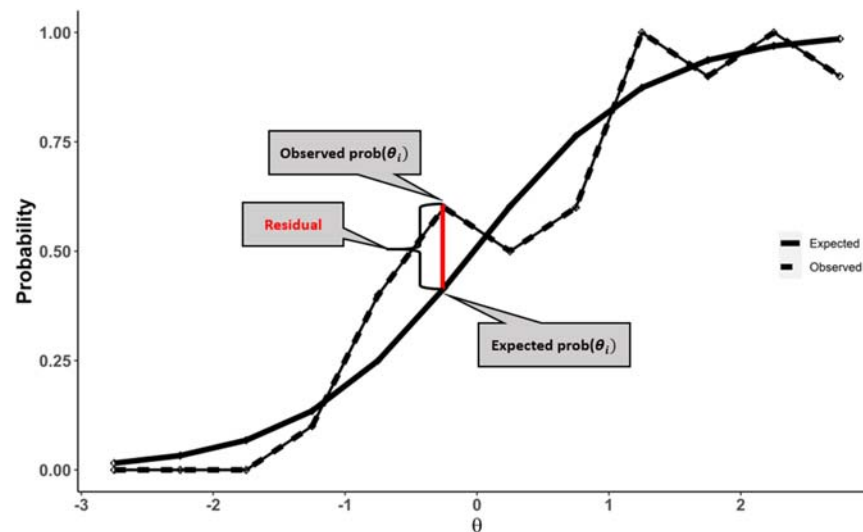


Fig. 13 Example of misfitting 2PL model and accompanying residual.

References

- Ackerman, T., 1994. Creating a test information profile for a two-dimensional latent space. *Appl. Psychol. Meas.* 18 (3), 257–275.
- Ackerman, T., 1996. Graphical representation of multidimensional item response theory analyses. *Appl. Psychol. Meas.* 20 (4), 311–330.
- Ackerman, Ma, Y., Ip, E., 2019. Comparison of Three Unidimensional Approaches to Represent a Two-Dimensional Latent Ability Space. In: *Quantitative Psychology*. Springer International Publishing, pp. 195–204. https://doi.org/10.1007/978-3-030-01310-3_18.
- Albano, A., Babcock, B., 2019. RWinsteps: Running Winsteps in R. <https://CRAN.R-project.org/package=RWinsteps>, version 1.0-1.1.
- Allen, M.J., Yen, W.M., 1979. *Introduction to Measurement Theory*. Brooks/Cole Pub. Co, Monterey, CA.
- Andrich, D., 1988. *Rasch Models for Measurement*. Sage Publications, Beverly Hills.
- Baker, F., Kim, S.H., 2004. *Item Response Theory: Parameter Estimation Techniques*. Marcel Dekker, New York.
- Binet, A., Simon, T.A., 1905. Méthodes Nouvelles pour le diagnostic du niveau intellectuel des anormaux. *Psychol. Year* 11, 191–336.
- Birnbaum, A., 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord, F.M., Novick, M.R. (Eds.), *Statistical Theories of Mental Test Scores*. Addison Wesley, Reading, MA, pp. 397–479.
- Bock, R.D., Aitkin, M., 1981. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika* 46 (4), 443–459.
- Bock, R.D., Lieberman, M., 1970. Fitting a response model for n dichotomously scored items. *Psychometrika* 35 (2), 179–197.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37 (1), 29–51.
- Bradlow, E., Wainer, H., Wang, X., 1999. A Bayesian random effect model for testlets. *Psychometrika* 64 (2), 153–168.
- Cai, L., 2022. flexMIRT®: Flexible Multilevel Multidimensional Item Analysis and Test Scoring (Version 3.6.5) [Computer Software]. Vector Psychometric Group, Chapel Hill, NC.
- Chalmers, R.P., 2012. Mirt: a multidimensional item response theory package for the environment. *J. Stat. Software* 48 (6), 1–29.
- De Ayala, R.J., Sava-Bolesta, M., 1999. Item parameter recovery for the nominal response model. *Appl. Psychol. Meas.* 23 (1), 3–19.
- De Ayala, R.J., Dodd, B.G., Koch, W.R., 1992. A comparison of the partial credit and graded response models in computerized adaptive testing. *Appl. Meas. Educ.* 5 (1), 17–34.
- De Ayala, R.J., 2013. *The Theory and Practice of Item Response Theory*. Guilford Publications.
- DeMars, C.E., 2002. Recovery of Graded Response and Partial Credit Parameters in MULTILOG and PARSCALE.
- Falk, C.F., 2020. The monotonic polynomial graded response model: implementation and a comparative study. *Appl. Psychol. Meas.* 44 (6), 465–481.
- Ferguson, G.A., 1943. Items selection by the constant process. *Psychometrika* 7 (1), 19–29.
- Finch, H., French, B.F., 2019. A comparison of estimation techniques for IRT models with small samples. *Appl. Meas. Educ.* 32 (2), 77–96.
- Fox, J., 2010. *Bayesian Item Response Modeling: Theory and Applications*. Springer, New York.
- Fraser, C., McDonald, R.P., 1988. NOHARM: least squares item factor analysis. *Multivariate Behav. Res.* 23, 267–269.
- Hambleton, R., Swaminathan, H., 1985. *Item Response Theory: Principles and Applications*. Kluwer-Nijhoff, Boston.
- Lawley, D.N., 1943. On problems connected with item selection and test construction. *Proc. R. Soc. Edinb.* 62-A (Part 1), 74–82.
- Li, Y., Baser, R., 2012. Using R and WinBUGS to fit a generalized partial credit model for developing and evaluating patient-reported outcomes assessments. *Stat. Med.* 31 (18), 2010–2026.
- Linacre, J.M., 2003. Size vs. significance: infit and outfit mean-square and standardized Chi-Square fit statistic. *Rasch Meas. Trans.* 17, 918.
- Linacre, J.M., 2019. Winsteps® Rasch Measurement Computer Program. Winsteps.com, Beaverton, Oregon.
- Lord, F.M., Novick, M.R., Birnbaum, A., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley.
- Lord, F.M., 1952. *A Theory of Test Scores*. Psychometrika Monographs, No. 7.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Lawrence Erlbaum, Hillsdale, NJ.
- Lubbe, D., Schuster, C., 2019. A graded response model framework for questionnaires with uniform response formats. *Appl. Psychol. Meas.* 43 (4), 290–302.
- Lunn, D., Spiegelhalter, D., Thomas, A., Best, N., 2009. The BUGS project: evolution, critique and future directions. *Stat. Med.* 28, 3049–3067.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.
- Mislevy, R.J., Bock, R.D., 1982. BILOG: Item Analysis and Test Scoring with Binary Logistic Models [Computer Software]. Scientific Software, Chicago.
- Mislevy, R.J., Bock, R.D., 1990. BILOGMG: Item analysis and test scoring with binary logistic models [Computer software]. Scientific Software, Chicago.
- Muraki, E., Bock, R.D., 1993. PARSCALE: IRT Based Test Scoring and Item Analysis for Graded Open-Ended Exercises and Performance Tasks. Scientific Software International.
- Muraki, E., Bock, R.D., 2003. PARSCALE 4: IRT Based Test Scoring and Item Analysis for Graded Items and Rating Scales. Scientific Software International, Chicago (IL).
- Muraki, E., 1992. A generalized partial credit model: application of an EM algorithm. *Appl. Psychol. Meas.* 16 (2), 159–176.

- Naumenko, O., 2014. Comparison of Various Polytomous Item Response Theory Modeling Approaches for Task-Based Simulation CPA Exam Data. AICPA 2014 Summer Internship Project.
- Penfield, R.D., Bergeron, J.M., 2005. Applying a weighted maximum likelihood latent trait estimator to the generalized partial credit model. *Appl. Psychol. Meas.* 29 (3), 218–233.
- Plummer, M., 2012a. JAGS Version 3.2.0 User Manual. Tech. Rep. <http://mcmc-jags.sourceforge.net>.
- Plummer, M., 2012b. Package “rjags”. Tech. Rep. <http://cran.r-project.org/web/packages/rjags/index.html>.
- Rasch, G., 1960. Probabilistic Models for Some Intelligence and Attainment Tests. Danish Institution for Educational Research, Copenhagen.
- Reckase, M.D., 2009. Multidimensional Item Response Theory. Springer, New York, NY.
- Richardson, M.W., 1936. The relation between the difficulty and the differential validity of a test. *Psychometrika* 1 (2), 33–49.
- Rizopoulos, D., 2006. ltm: an R package for latent variable modeling and item response theory analyses. *J. Stat. Software* 17 (5), 1–25.
- Roberts, J.S., Donoghue, J.R., Laughlin, J.E., 1998. The generalized graded unfolding model: a general parametric item response model for unfolding graded responses. *ETS Res.* 1998 (2), 1–53. <https://doi.org/10.1002/j.2333-8504.1998.tb01781.x>.
- Roberts, J.S., Fang, H.R., Cui, W., Wang, Y., 2006. GGUM2004: a windows-based program to estimate parameters in the generalized graded unfolding model. *Appl. Psychol. Meas.* 30 (1), 64–65.
- Rupp, A.A., Templin, J., Henson, R.A., 2010. Diagnostic Measurement: Theory, Methods, and Applications. Guilford Press, New York.
- Samejima, F., 1969. Estimation of latent ability using a response pattern of graded scores. *Psychometr. Monogr.* 34 (4, Pt. 2), 100.
- Samejima, F., 1972. A general model for free-response data. *Psychometr. Monogr.* 37 (1, Pt. 2), 68.
- Samejima, F., 1973. Homogeneous case of the continuous response model. *Psychometrika* 38 (2), 203–219.
- Samejima, F., 1995. Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika* 60 (4), 549–572.
- Segall, D.O., 2002. An item response model for characterizing test compromise. *J. Educ. Behav. Stat.* 27 (2), 163–179.
- Shu, Z., 2010. Detecting test cheating using a deterministic, gated item response theory. *Appl. Psychol. Meas.* 7, 201–210.
- Spiegelhalter, D.J., Thomas, A., Best, N., Lunn, D., 2012. The OpenBUGS User Manual Version 3.2. 2. MRC Biostatistics Unit, Cambridge, UK.
- Stan Development Team, 2019. Stan Modeling Language Users Guide and Reference Manual, Version 2.26. <https://mc-stan.org>.
- Tendeiro, J.N., Castro-Alvarez, S., 2019. GGUM: an R package for fitting the generalized graded unfolding model. *Appl. Psychol. Meas.* 43 (2), 172–173.
- Thissen, D., Chen, W.H., Bock, D., 2003. MULTILOG 7.03. Scientific Software International, Lincolnwood, IL.
- Thissen, D., 1982. Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika* 47 (2), 175–186.
- Thurstone, L.L., 1925. A method of scaling educational and psychological tests. *J. Educ. Psychol.* 16 (7), 433–451.
- van der Linden, W.J., 2005. Linear Models for Optimal Test Assembly. Springer, New York.
- van der Linden, W.J., 2018. Handbook of Item Response Theory. CRC Press, Boca Raton, FL.
- von Davier, M., Lee, Y.-S., 2019. In: von Davier, M., Lee, Y.-S. (Eds.), Handbook of Diagnostic Classification Models. Springer, Switzerland.
- Wang, W., de la Torre, J., Drasgow, F., 2015. MCMC GGUM: a new computer program for estimating unfolding IRT models. *Appl. Psychol. Meas.* 39 (2), 160–161.
- Wilson, D., Wood, R., Gibbons, R., 1984. TESTFACT: Test Scoring, Item Statistics, and Item Factor Analysis [Computer Program]. Scientific Software, Inc., Mooresville, IN.
- Wright, B., Masters, G., 1983. Rating Scale Analysis. Rasch Measurement. Mesa Press, Chicago.
- Wright, B.D., 1984. Despair and hope for educational measurement. *Contemp. Educ. Psychol.* 3 (1), 281–288.
- Yavuz, G., Hambleton, R.K., 2017. Comparative analyses of MIRT models and software (BMIRT and flexMIRT). *Educ. Psychol. Meas.* 77 (2), 263–274.
- Yen, W., 1984. Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Appl. Psychol. Meas.* 8 (2), 125–145.

Multidimensional item response theory

Sijia Huang and Li Cai, University of California, Los Angeles, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	86
From unidimensional IRT models to MIRT models	87
MIRT models	89
Correlated-factors MIRT models	89
2PL model correlated factors extension	89
3PL model correlated factors extension	90
GRM correlated factors extension	90
The bifactor/testlet models	91
2PL model bifactor extension	92
GRM bifactor extension	92
The two-tier model	92
An aside: multidimensional NRM	93
Estimation of MIRT models and scoring individuals	94
Bock-Aitkin EM algorithm	95
Metropolis-Hastings Robbins-Monro (MH-RM) algorithm	95
Markov chain Monte Carlo (MCMC) approach	95
Scaled score based on response patterns	95
Summed score to IRT scaled scores	96
Example applications and future directions	96
Educational measurement	96
Clinical outcomes assessment	96
Future directions	96
References	96

Introduction

Item response theory (IRT) can be viewed minimally as a set of psychometric models for categorical item-level response data (Cai et al., 2016). It can also be viewed from the statistical angle as a set of interrelated discrete multivariate statistical models and methods that help address the needs arising out of item analysis, test development and validation, linking and equating, test assembly and adaptive testing, and test scoring and reporting (Thissen and Steinberg, 2009). Cai et al. (2016) also demonstrated IRT's usefulness in data analysis for large-scale educational assessments and in evaluation studies using experimental or quasi-experimental designs. IRT has a long history that can be traced at least to Thurstone's (1925) effort on locating Binet test questions on the chronological age scale (see Bock, 1997). Since the publication of Lord and Novick's (1968) *Statistical Theories of Mental Test Scores*, IRT gradually became an active research area. IRT is now one of the most important modern psychometric methodologies, and is broadly applied to data from educational assessments and other areas such as clinical outcome assessments.

IRT models describe the relationship among the items' psychometric properties, item responses, and individual latent variable(s) theorized to influence the item responses. The formulation of IRT models relies on careful descriptions of the individual response process, assumptions about the latent variable distributions, and the critical *conditional independence* assumption to turn the psychometric models into multivariate statistical models. The conditional independence assumption states that conditional on the latent variable(s), responses to items are independent. Stated in another way, the items are correlated because of the influence of the latent variables.

In the context of traditional unidimensional IRT models, the conditional independence assumption implies that the dependence or covariation among observed item responses is solely attributable to the presence of one single continuous latent variable. This unidimensionality assumption is often violated in practice because many constructs in social and behavioral sciences are inherently multifaceted (e.g., quality of life), and observed item responses may also be influenced by more than one latent variables because of features of test design (e.g., passage-based reading tests or language tests).

Multidimensional item response theory (MIRT) models naturally extend traditional unidimensional IRT models by including more than a single latent variable in model formulation. MIRT models can also be viewed as factor analysis models for categorical item level data (Wirth and Edwards, 2007). In this article, we present recent developments in MIRT and major applications of MIRT models.

From unidimensional IRT models to MIRT models

Let us consider the two-parameter logistic (2PL) model, a classic unidimensional IRT model for items scored in 2 categories (e.g., correct/incorrect, symptom present/not present, endorse/not endorse). For a generic item i , the item response is a Bernoulli (0–1) random variable U_i . As its name suggests, the 2PL model (Eq. 1) uses two item parameters to describe how the conditional probability of the correct response for item i changes as a function of the latent variable θ , which characterizes individual differences on the dimension of interest. The latent variable is often, though not always required, assumed to follow a standard normal distribution, $\theta \sim \mathcal{N}(0, 1)$.

$$P(U_i = 1|\theta) = \frac{1}{1 + \exp[-(c_i + a_i\theta)]} \quad (1)$$

The slope/discrimination parameter a_i is the increase in the logit of the correct response probability that corresponds to one unit increase in θ . The intercept parameter c_i is the logit when $\theta = 0$. Fig. 1A shows the item characteristic curve (ICC) for an item with $a_i = 1.5$ and $c_i = 0.5$, whereas Fig. 1B presents the y -axis in logits. A higher a_i value suggests that the item can better differentiate individuals with similar θ values. The classical item difficulty parameter b_i , i.e., the θ point at which $P(U_i = 1|\theta)$ is 0.5, can be computed as $b_i = -c_i/a_i$.

The item i can best differentiate individuals whose θ values are close to its 50–50 point. That is to say, it provides the most psychometric information in the range near b_i along the θ continuum. The amount of information an item can provide depends on its discriminating power. Under the 2PL model, the item information function, denoted by $\mathcal{I}(\theta)$, is given by the formula,

$$\mathcal{I}(\theta) = a_i^2 P(U_i = 1|\theta)[1 - P(U_i = 1|\theta)] \quad (2)$$

Fig. 2 shows the item information functions of three hypothetical items with $a_1 = 0.75$, $c_1 = 0.5$, $b_1 = -0.67$; $a_2 = 1.5$, $c_2 = 1$, $b_2 = -0.67$; and $a_3 = 0.75$, $c_3 = 1$, $b_3 = -1.33$. Having the same slope parameter, the relatively easy item (item 3) provides more

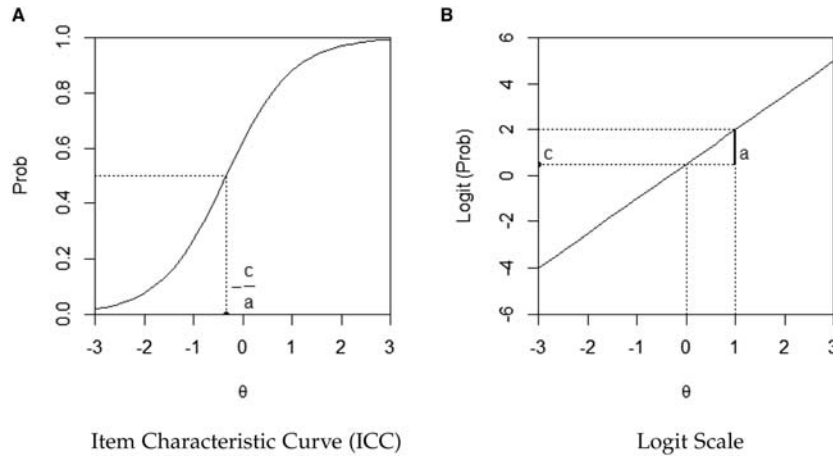


Fig. 1 Unidimensional two-parameter logistic (2PL) model. Item characteristic curve (ICC) for a 2PL model when $a = 1.5$, and $c = 0.5$.

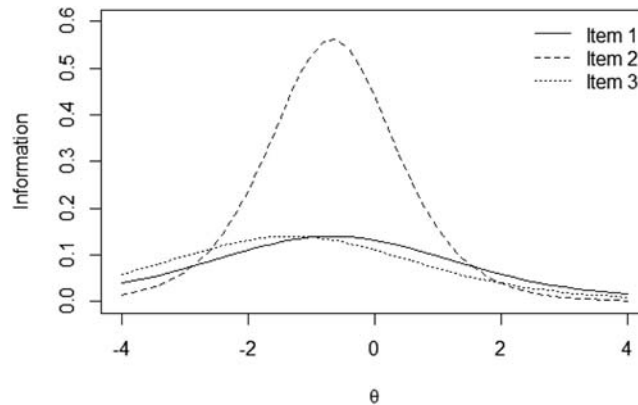


Fig. 2 Item information functions of three hypothetical items. Item parameters are respectively $a_1 = 0.75$, $c_1 = 0.5$, $b_1 = -0.67$; $a_2 = 1.5$, $c_2 = 1$, $b_2 = -0.67$; and $a_3 = 0.75$, $c_3 = 1$, $b_3 = -1.33$.

information in the lower- θ range than the difficult item (item 1). The more-discriminating item (item 2) provides more information than the less-discriminating item (item 1). Statistically, the item information function also indicates the precision of θ estimates, and its reciprocal $1/\mathcal{I}(\theta)$ equals the asymptotic variance of the maximum likelihood estimates (MLEs) of the latent variable θ , the square root of which is also referred to as the *Standard Error of Measurement*.

In MIRT models, a D -dimensional latent variable vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d, \dots, \theta_D)$ is used instead of the scalar θ . Two classes of MIRT models, *compensatory* and *noncompensatory* models, can be distinguished by the nature of how the latent variables are combined to influence the item response probability.

In the more widely used *compensatory* models, the response probability depends on a linear combination of the latent variables inside the logit (Eq. 3).

$$P(U_i = 1|\boldsymbol{\theta}) = \frac{1}{1 + \exp[-(c_i + a_{i1}\theta_1 + \dots + a_{iD}\theta_D)]} \quad (3)$$

The slope parameter a_{id} indicates the rate of change in the logit-probability with respect to θ_d . The relative item difficulty that corresponds to dimension d is $-c_i/a_{id}$, which is interpretable as the 50–50 point for the response probability on the dimension corresponding to θ_d , while holding $\boldsymbol{\theta}$ at 0 on all other dimensions. The multidimensional generalization of the a_i in the unidimensional model is the so-called multidimensional discrimination index (MDISC; Reckase and McKinley, 1991), which is often denoted by A_i . The MDISC is the overall measure of an item's ability to distinguish individuals in the D -dimensional latent space, and is defined as the following,

$$A_i = \sqrt{\sum_{d=1}^D a_{id}^2} \quad (4)$$

The important point here is that with compensatory models, a low value on one or more of the latent variables can be offset by high values of other latent variables. Fig. 3 shows the item response surface (instead of a curve as in a unidimensional model) of an item when $D = 2$ with $a_{i1} = 1.5$, $a_{i2} = 0.75$ and $c_i = 0.5$. When θ_1 approaches the lower extreme of the latent continuum (i.e., $\theta_1 = -3$), there is still a substantial correct/positive response probability when θ_2 is high.

Analogous to the unidimensional case, where the item information varies over levels of θ , the item information in MIRT models is defined as a function of $\boldsymbol{\theta}$'s position in the D -dimensional latent space,

$$\mathcal{I}(\boldsymbol{\theta}) = \left(\sum_{d=1}^D a_{id} \cos \alpha_d \right)^2 P(U_i = 1|\boldsymbol{\theta}) [1 - P(U_i = 1|\boldsymbol{\theta})] \quad (5)$$

where α_d indicates the angle between the d th axis and a line from the origin to the $\boldsymbol{\theta}$ point, and is introduced to accommodate the difference in slopes of a point on different dimensions.

In a *noncompensatory* model for dichotomous data, the probability of a correct response is the product of probabilities of D unidimensional models (Eq. 6).

$$P(U_i = 1|\boldsymbol{\theta}) = \prod_{d=1}^D \frac{1}{1 + \exp[-(c_{id} + a_{id}\theta_d)]} \quad (6)$$

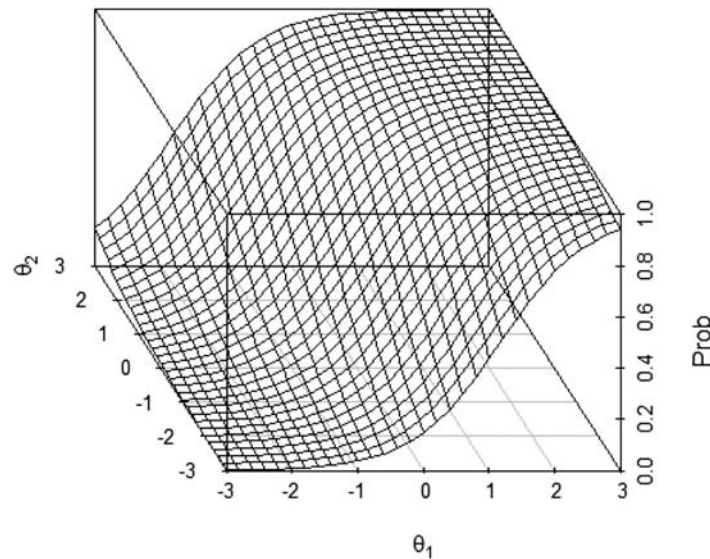


Fig. 3 Compensatory multidimensional item response theory model. Item response curve for a compensatory MIRT model when $a_1 = 1.5$, $a_2 = 0.5$, and $c = 0.5$.

Each of the unidimensional models includes a slope parameter a_{id} and an intercept parameter c_{id} . This can lead to estimation challenges. Noncompensatory MIRT models are appropriate when an item can be hypothesized to involve several distinct parts that depend on different latent variables in the response process. The reader is referred to sources such as Reckase (2009) for more in-depth discussions about compensatory and non-compensatory MIRT models. Again, though not required, θ is often assumed to follow a multivariate normal distribution $\theta \sim \mathcal{N}(\mathbf{0}, \Sigma)$. Elements of Σ show how the latent variables covary.

MIRT models

We focus on the compensatory MIRT models (i.e., Thurstonian item factor analysis models; Bock et al., 1988; Bock and Aitkin, 1981) and show how different MIRT models can be formulated for items with two response categories before moving on to polytomous data. Due to space constraints, we present in detail only the multidimensional extension of Samejima's (1969) graded response model (GRM) for ordered polytomous responses. Discussions about multidimensional extensions of the generalized partial credit model and partial credit model can be found in Yao and Schwarz (2006), Kelderman and Rijkes (1994), and Reckase (2009). At the end of this section, we introduce the nominal response model (NRM; Bock, 1972), which does not assume ordered categories, and review a new parameterization (Thissen et al., 2010; Thissen and Cai, 2016) that allows the generalization of the NRM to the case of MIRT.

To provide a context of discussions, let us consider an assessment with four sections that respectively measure four domains of students' language proficiency (reading, listening, writing, and speaking). The English Language Proficiency Assessment for the 21st Century (ELPA21; Cai and Hansen, 2018) is a concrete example. Let us assume that the test includes both multiple choice (MC) items, which have two scored categories (i.e., correct and incorrect), and constructed response (CR) items, which have more than two ordered score categories to reflect partial credit scoring of test taker performance using a rubric. The four sections in the assessment can be treated as four separate tests and modeled with four sets of unidimensional IRT models, as is often done in practice. However, by doing so, the correlations among the four domains of language proficiency are ignored, to the detriment of psychometric evaluations and validation studies, as well as limiting the range of possibilities in score reporting. Depending on the assumption regarding how latent variables influence observed item responses, different MIRT models can be specified.

Correlated-factors MIRT models

One of the many sensible ways to formulate a MIRT model in our context is shown in Fig. 4, in which items (squares) in the four sections measure respectively four correlated latent variables (circles) that represent students' (correlated) proficiencies in reading, listening, writing, and speaking. With this correlated-factors MIRT model, one can obtain estimated proficiency levels on each of the four latent variables (i.e., four scaled scores) simultaneously. Below we present how the 2PL model, the three-parameter logistic (3PL) model, and the GRM (Samejima, 1969) can be extended to their MIRT counterparts.

2PL model correlated factors extension

Let $\theta = (\theta_1, \dots, \theta_4)'$ denote a 4-dimensional vector whose elements represent students' proficiencies in reading, listening, writing, and speaking. Items may load on a single domain, or cross domains, reflecting the interactive language modality that requires more than one domain-based skills. The probability of answering a dichotomous item correctly can be specified as a function of θ and the item parameters,

$$P(U_i = 1 | \theta) = \frac{1}{1 + \exp[-(c_i + a_{i1}\theta_1 + a_{i2}\theta_2 + a_{i3}\theta_3 + a_{i4}\theta_4)]} = \frac{1}{1 + \exp[-(c_i + \mathbf{a}_i'\theta)]} \quad (7)$$

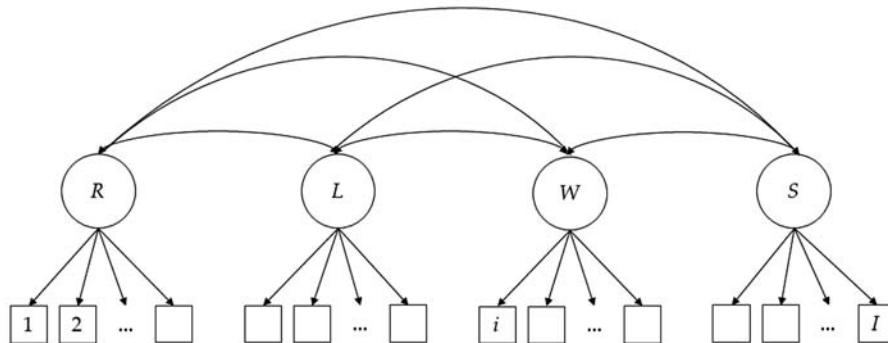


Fig. 4 Correlated-trait MIRT model. The squares in the figure represent items, while the circles represent latent variables measured with the assessment, which are reading (*R*), listening (*L*), writing (*W*) and speaking (*S*). The single-headed arrows from the latent variables to the items indicates that levels of latent variables impact responses to items. The double-headed arrows show that the latent variables are correlated with each other.

in which $\mathbf{a}_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4})'$ is a vector of slope/discrimination parameters, and c_i is the intercept parameter.

As mentioned earlier, the correlated factors extension of the 2PL model is closely related to the categorical confirmatory factor analysis (CCFA) model. Parameter estimates of $\mathbf{a}_i$ and c_i can be easily converted to parameters of the CCFA model, and vice versa (Takane and De Leeuw, 1987; Wirth and Edwards, 2007; Cai, 2010a). The CCFA model assumes that each individual possesses a latent item score x_i^* , reflecting his/her position on a normally-distributed latent response scale. It is also assumed that for the item i (which has 2 categories), there exists a threshold parameter τ_i , so that

$$U_i = \begin{cases} 1, & x_i^* \geq \tau_i, \\ 0, & x_i^* < \tau_i. \end{cases} \quad (8)$$

The covariance structure of the latent scores can then be modeled as

$$\Sigma^* = \Lambda^* \Phi^* \Lambda^{*'} + \Psi^* \quad (9)$$

where Σ^* represents the covariance between latent item scores, Λ^* denotes the factor loading matrix, Φ^* is the covariance matrix of latent common factors, and Ψ^* is the unique variance matrix. The asterisk indicates that both sides of Eq. (9) are latent.

Suppose each of the four domains in the example (reading, listening, writing, and speaking) is measured by 2 unique items. The factor loading matrix Λ^* , whose rows correspond to items and columns correspond to common latent factors, is a 8×4 matrix

$$\Lambda^* = \begin{bmatrix} \lambda_1^* & 0 & 0 & 0 \\ \lambda_2^* & 0 & 0 & 0 \\ 0 & \lambda_3^* & 0 & 0 \\ 0 & \lambda_4^* & 0 & 0 \\ 0 & 0 & \lambda_5^* & 0 \\ 0 & 0 & \lambda_6^* & 0 \\ 0 & 0 & 0 & \lambda_7^* \\ 0 & 0 & 0 & \lambda_8^* \end{bmatrix}$$

If estimates of $\mathbf{a}_i$ and c_i were obtained, elements of Λ^* can be computed as

$$\lambda_i^* = \frac{a_i/1.702}{\sqrt{1 + (a_i/1.702)^2}} \quad (10)$$

and the threshold parameter τ_i is

$$\tau_i = \frac{-c_i/1.702}{\sqrt{1 + (a_i/1.702)^2}} \quad (11)$$

in which 1.702 is a traditionally-used scaling constant that puts logistic parameters on a normal ogive metric.

3PL model correlated factors extension

For an MC item, it is possible that a student with lower proficiencies could still answer it correctly by guessing. The 3PL model is an extension of the 2PL model and uses a non-zero lower asymptote to account for the impact of guessing on item characteristics. Its multidimensional extension may be expressed as

$$P(U_i = 1|\boldsymbol{\theta}) = g_i + (1 - g_i) \frac{1}{1 + \exp[-(c_i + \mathbf{a}_i' \boldsymbol{\theta})]} \quad (12)$$

in which $\mathbf{a}_i$ and c_i continue to be the slope vector and intercept parameters, respectively, and g_i is the lower asymptote (pseudo-guessing) parameter (Bock et al., 1988).

GRM correlated factors extension

Muraki and Carlson (1995) proposed the multidimensional extension of Samejima's (1969) GRM. For an item with K ordered categories, e.g., an item in the CR section, a set of cumulative probabilities are specified,

$$P(U_i \geq 0|\boldsymbol{\theta}) = 1$$

$$P(U_i \geq 1|\boldsymbol{\theta}) = \frac{1}{1 + \exp[-(c_{i,1} + \mathbf{a}_i' \boldsymbol{\theta})]}; \quad (13)$$

$$P(U_i \geq K-1|\theta) = \frac{1}{1 + \exp[-(c_{i,K-1} + \mathbf{a}_i'\theta)]}$$

$$P(U_i \geq K|\theta) = 0$$

such that the probability that the item response U_i is in category k , $k = 0, \dots, K-1$ can be computed as the difference between two adjacent cumulative probabilities,

$$P(U_i = k|\theta) = P(U_i \geq k|\theta) - P(U_i \geq k+1|\theta) \quad (14)$$

The bifactor/testlet models

In addition to assessing the proficiencies in the language domains, the goal of the assessment may be to provide an overall evaluation of students' language proficiency. That is to say, instead of reporting four domain scores with the correlated-trait model, one overall score may be desired. In this new context, the bifactor model (Holzinger and Swineford, 1937; Cai et al., 2011) can be useful.

The bifactor model captures the factorial structure of many educational and psychological tests that consist of items selected from domains (or multifaceted subdomains) and aim to measure a more general construct when assembled together. Specifically, the bifactor model assumes that there is one general latent variable (the general factor) that contribute to the covariance among all items and several latent variables that are group-specific that explain additional dependence among items in the same domain/sub-domain (Fig. 5). In addition to assuming that each item loads on the general factor and at most one group-specific factor, it further assumes that the group factors are conditionally independent given the general factor.

Restricted versions of the bifactor model exist and the most prominent example is the testlet response model (Bradlow et al., 1999). Specifically, to obtain a testlet response model from the bifactor model, one can constrain an item's loading on the group-specific factor to be proportional to its loading on the general factor (Rijmen, 2010). Another well-known restricted version of the bifactor model is the second-order model (Fig. 6), which assumes conditionally-independent group-specific factors and

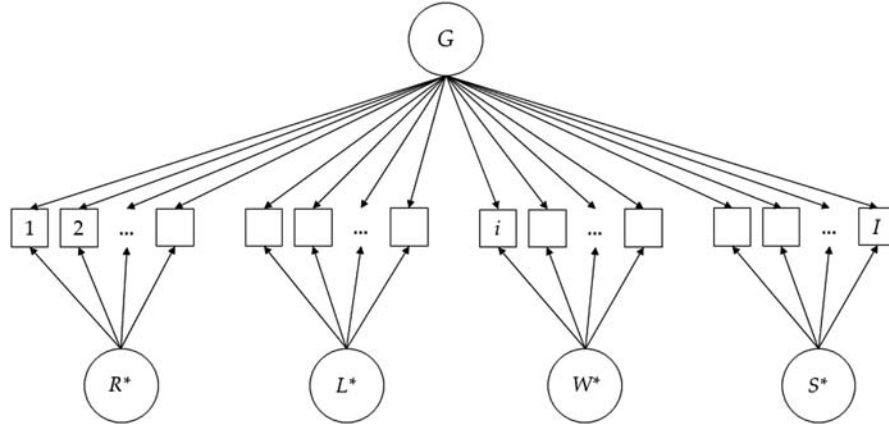


Fig. 5 Bi-factor model. The latent variable G represents students' overall language proficiency, and impacts responses to all items. The latent variables R^* , L^* , W^* , and S^* impact only items that load on them. Note, the asterisk indicates that R^* , L^* , W^* , and S^* are different from R , L , W , and S in Fig. 4.

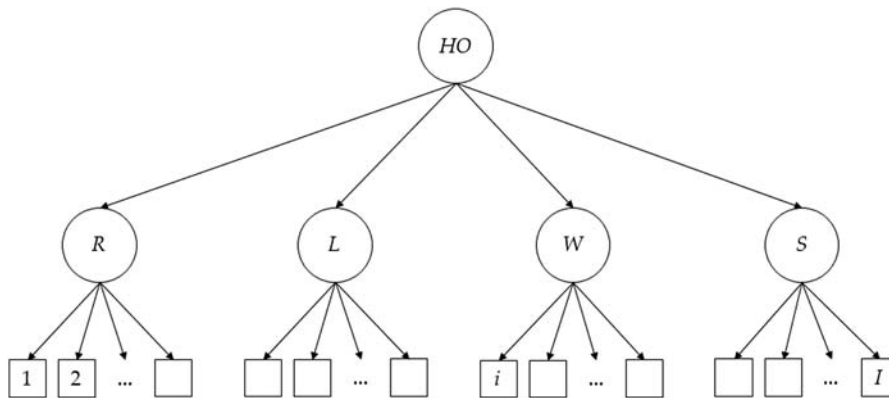


Fig. 6 Second-order factor model. In the second-order factor model, a second-order factor HO is introduced to explain the covariance between group-specific factors.

introduces a second-order factor to explain the covariance between the group-specific factors. The second-order factor model is equivalent to the testlet response model and can be obtained from the bifactor model by imposing proportionality constraints as well (Rijmen, 2010; Yung et al., 1999).

2PL model bifactor extension

Gibbons and Hedeker (1992) developed the bifactor item response model for binary data with the normal-ogive (probit) link function. The logistic counterpart, which is a direct extension of the 2PL model is

$$P(U_i = 1|\theta) = \frac{1}{1 + \exp[-(c_i + a_i\eta + \beta_{id}\xi_d)]} \quad (15)$$

for an item that belong to domain $d = 1, \dots, 4$. The latent variable $\theta = (\eta, \xi_1, \xi_2, \xi_3, \xi_4)'$ consists of one general latent variable η and four domain-specific latent variables ξ_d . For example, if there were 8 items and each domain is measured by 2 unique items, the pattern of item slopes can be represented as a 8×5 matrix as follows

$$\begin{bmatrix} a_1 & \beta_{11} & 0 & 0 & 0 \\ a_2 & \beta_{21} & 0 & 0 & 0 \\ a_3 & 0 & \beta_{32} & 0 & 0 \\ a_4 & 0 & \beta_{42} & 0 & 0 \\ a_5 & 0 & 0 & \beta_{53} & 0 \\ a_6 & 0 & 0 & \beta_{63} & 0 \\ a_7 & 0 & 0 & 0 & \beta_{74} \\ a_8 & 0 & 0 & 0 & \beta_{84} \end{bmatrix}$$

As shown in Eq. (15) and the matrix above, for any item, only two factors (i.e., the general factor and the group factor it loads on) are involved, even though the total latent dimensionality of the model is five. The key advantage of the bifactor pattern is in computational savings for item parameter estimation, to be discussed in **Estimation of MIRT models and scoring individuals** section.

GRM bifactor extension

Gibbons et al. (2007) developed a bifactor item response model for graded response data, which may be useful for the CR items. We can express the bifactor extension of the GRM as follows:

$$\begin{aligned} P(U_i \geq 0|\theta) &= 1 \\ P(U_i \geq 1|\theta) &= \frac{1}{1 + \exp[-(c_{i,1} + a_i\eta + \beta_{id}\xi_d)]}; \\ P(U_i \geq K-1|\theta) &= \frac{1}{1 + \exp[-(c_{i,K-1} + a_i\eta + \beta_{id}\xi_d)]} \\ P(U_i \geq K|\theta) &= 0, \end{aligned} \quad (16)$$

and

$$P(U_i = k|\theta) = P(U_i \geq k|\theta) - P(U_i \geq k+1|\theta).$$

The two-tier model

In the assessment context described above, the reading and listening domains focus on the test taker's receptive language modality, while the writing and speaking sections focus on the productive language modality. If proficiencies in these language modalities are of primary interest, we may utilize the two-tier model (Cai, 2010c) for data analysis and scaled score estimation. The two-tier model includes two types of latent variables. The primary latent variables (e.g., students' receptive and productive proficiencies) are in the first tier, and group-specific latent variables in the second tier (Fig. 7). The primary latent variables are analogous to the general factor in the item bifactor model, and they can be correlated, while the specific latent variables are conditionally independent given the primary latent variables, similar to the group-specific factors in a bifactor model, accounting for residual covariance among non-overlapping subsets of items above and beyond the primary dimensions. The standard correlated-factors MIRT model and the various bifactor models discussed in the previous subsections can all be viewed as special cases of the two-tier model.

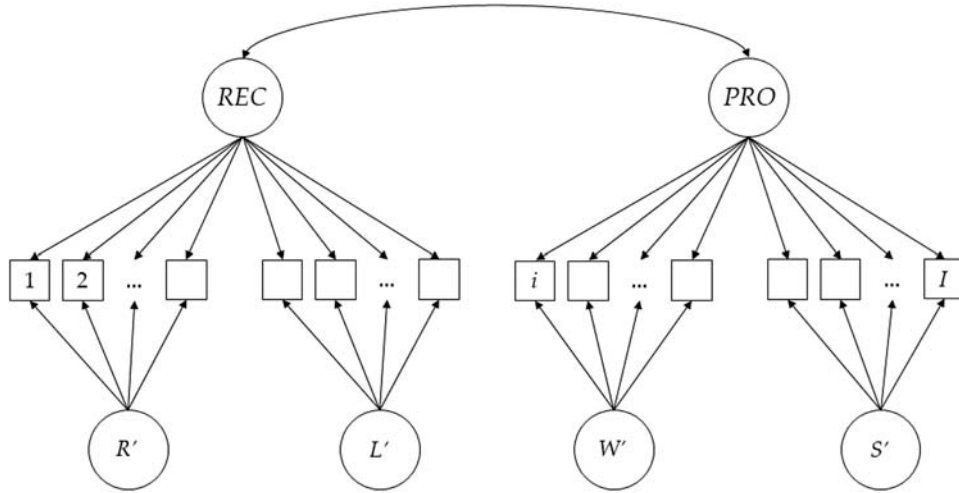


Fig. 7 Two-tier model. The *REC* and *PRO* are the two primary latent variables and represent students' receptive and productive language proficiencies. The *R'*, *L'*, *W'*, and *S'* are the specific latent variables. Note when there is no specific latent variables, the two-tier model becomes a correlated-trait MIRT model. When there is only one primary latent variable, it becomes a bi-factor model.

In the example depicted in Fig. 7, we have a total of 6 latent variables: $\theta = (\eta_1, \eta_2, \xi_1, \xi_2, \xi_3, \xi_4)'$. The extension of the 2PL model in the two-tier modeling framework is

$$P(U_i = 1|\theta) = \frac{1}{1 + \exp[-(c_i + a_{i1}\eta_1 + a_{i2}\eta_2 + \beta_{id}\xi_d)]} \quad (17)$$

for an item that belongs to domain indexed by latent variable d . Eq. (17) states that the probability of endorsing an item depends on both primary latent variables and the specific latent variable on which the item loads. For example, if there were 8 items and each domain is measured by 2 items, the pattern of item slopes can be represented as a 8×6 matrix as follows

$$\begin{bmatrix} a_{11} & 0 & \beta_{11} & 0 & 0 & 0 \\ a_{21} & 0 & \beta_{21} & 0 & 0 & 0 \\ a_{31} & 0 & 0 & \beta_{32} & 0 & 0 \\ a_{41} & 0 & 0 & \beta_{42} & 0 & 0 \\ 0 & a_{52} & 0 & 0 & \beta_{53} & 0 \\ 0 & a_{62} & 0 & 0 & \beta_{63} & 0 \\ 0 & a_{72} & 0 & 0 & 0 & \beta_{74} \\ 0 & a_{82} & 0 & 0 & 0 & \beta_{84} \end{bmatrix}$$

Cai (2010c) illustrated the use of the two-tier model in the context of longitudinal IRT modeling, complex educational assessments, and response style analysis, with demonstrations of how other IRT models can be extended to the two-tier case.

An aside: multidimensional NRM

In contrast to the GRM, Bock's (1972) nominal response model (NRM) does not require ordered categorical data. The original formulation of the NRM is based on the multinomial logit link function:

$$P(U_i = k|\theta) = \frac{\exp(a_{ik}\theta + c_{ik})}{\sum_{j=0}^{K-1} \exp(a_{ij}\theta + c_{ij})} \quad (18)$$

which states that the conditional probability that the response U_i falls in category $k = 0, \dots, K-1$ depends on the latent variable θ and item parameters $\mathbf{a}_i$ and $\mathbf{c}_i$. Due to the translational invariance of the multinomial logit, not all elements of $\mathbf{a}_i$ and $\mathbf{c}_i$ are simultaneously estimable. For identification purposes, constraints on item slopes and intercepts are necessary, e.g., $\sum_{k=0}^{K-1} a_{ik} = \sum_{k=0}^{K-1} c_{ik} = 0$. As Bock (1972) showed, these constraints can be implemented by reparameterization using various contrast matrices akin to those used in analysis of variance, i.e.,

$$\mathbf{a}_i = \mathbf{T}_a \boldsymbol{\alpha}_i, \quad \mathbf{c}_i = \mathbf{T}_c \boldsymbol{\gamma}_i,$$

where the T_a and T_c are contrast matrices of dimension $K \times (K-1)$ and α_i and γ_i contain the estimable contrasts for slopes and intercepts. There are altogether $2(K-1)$ free parameters in the unidimensional NRM.

Thissen et al. (2010) presented a new parameterization of the NRM to facilitate its extension to the case of MIRT. Specifically, the new parameterization separates a_{ik} (in Eq. 18), the $(k+1)$ -th element of the slope parameter, into two components: an overall slope a_i and scoring function parameters for category k , s_{ik}

$$P(U_i = k | \theta) = \frac{\exp[s_{ik}(a_i\theta) + c_{ik}]}{\sum_{j=0}^{K-1} \exp[s_{ij}(a_i\theta) + c_{ij}]} \quad (19)$$

Dropping the i subscript for a moment, the following identification constraints,

$$s_0 = 0, \quad s_{K-1} = 0 = K - 1, \quad \text{and} \quad c_0 = 0 \quad (20)$$

are implemented by reparameterizing $\mathbf{s}$ and $\mathbf{c}$ using contrast matrices

$$\mathbf{s} = T_s \alpha, \quad \mathbf{c} = T_c \gamma, \quad (21)$$

where α takes on the role of scoring function contrasts. Thissen et al. (2010) demonstrated the model with a Fourier contrast matrix augmented with a linear column. In this new parameterization, $K-2$ elements of α are freely estimated, $K-1$ intercept contrasts are free, and the overall slope a is freely estimated. In other words, the number of free parameters is identical to the original model in Eq. (18). Thissen et al. (2010) also showed that popular IRT models for polytomous items such as the rating scale (RS; Andrich, 1978) model, the partial credit model (Masters, 1982), and the generalized partial credit (GPC; Muraki, 1992) model can be seen as constrained cases of the NRM, with this new parameterization.

The key benefit of this new parameterization lies in the fact that it offers a smooth path to extend the NRM to a MIRT model by expanding both the scalar θ and the overall slope a of item i into vectors,

$$P(U_i = k | \theta) = \frac{\exp[s_{ik}(\mathbf{a}'_i \theta) + c_{ik}]}{\sum_{j=0}^{K-1} \exp[s_{ij}(\mathbf{a}'_i \theta) + c_{ij}]} \quad (22)$$

where θ is the vector of latent variables, $\mathbf{a}_i$ is a vector of item slopes of the same dimension as θ . This model is nominal in the sense that for the linear combination of latent variables ($\mathbf{a}'_i \theta$), the category may not be ordered. A more general form of the model (Thissen and Cai, 2016) is,

$$P(U_i = k | \theta) = \frac{\exp\{[\mathbf{a}_i \circ \mathbf{s}_{ik}]' \theta + c_{ik}\}}{\sum_{j=0}^{K-1} \exp\{[\mathbf{a}_i \circ \mathbf{s}_{ij}]' \theta + c_{ij}\}} \quad (23)$$

in which $\circ$ denotes the Schur product, $\mathbf{s}_{ik}$ is the $(k+1)$ -th column of a $D \times K$ scoring function matrix $\mathbf{S}_i$, whose rows correspond to latent dimensions and columns correspond to item categories. By using the $\mathbf{S}_i$ (instead of the K -dimensional vector $\mathbf{s}_i$ as in Eq. 22), scoring functions are allowed to vary across latent dimensions.

Estimation of MIRT models and scoring individuals

The two major uses of IRT are item analysis and test scoring (Thissen and Steinberg, 2009). The former task requires statistical tools to obtain item parameters, and the latter involves the computation of individual scores (and standard errors) along the latent continua. From a statistical point of view, the two steps corresponds to the standard procedure in empirical Bayes inference. Item analysis produces structural parameter estimates, typically with maximum marginal likelihood, and the test scoring step produces scale scores, which are empirical Bayes estimates of the random effects/latent variables.

Using notation established above, the conditional likelihood of response pattern $\mathbf{u} = (u_1, \dots, u_i, \dots, u_I)$ can be written as

$$L_{\omega}(\mathbf{u} | \theta) = P\left(\bigcap_{i=1}^I U_i = u_i | \theta\right) = \prod_{i=1}^I P(U_i = u_i | \theta) = \prod_{i=1}^I \prod_{k=0}^{K-1} [P(U_i = k | \theta)]^{\chi_k(u_i)} \quad (24)$$

where we have utilized the conditional independence assumption such that the joint response pattern probability factors into a product, and $\chi_k(u_i)$ is an indicator function that equals to 1 if and only if u_i is equal to k , 0 otherwise. The likelihood is a function of ω , which contains all the item parameters and any other structural parameters (e.g., means and covariances of the population distribution). The marginal likelihood is an integral of the conditional likelihood over the population/prior distribution $h(\theta)$ of the latent variables θ :

$$L_{\omega}(\mathbf{u}) = \int \prod_{i=1}^I P(U_i = u_i | \theta) h(\theta) d\theta \quad (25)$$

The marginal likelihood, when maximized over ω , leads to the marginal maximum likelihood estimates of the item parameters.

In this section, we outline approaches of estimating MIRT model parameters and methods of scoring individuals. The test scoring section is more abbreviated, because the methods are parallel to those for estimating θ in unidimensional IRT models. We refer interested readers to [Thissen and Wainer \(2001\)](#) for details.

Bock-Aitkin EM algorithm

[Bock and Aitkin \(1981\)](#) proposed a procedure to obtain maximum likelihood estimates (MLEs) of IRT model parameters based on [Dempster et al.'s \(1977\)](#) expectation-maximization (EM) principle. The idea is to treat observed item responses as incomplete data and latent variables as missing data so that they together form complete data. The Bock-Aitken EM algorithm alternates between the *Expectation* step and the *Maximization* step. In the *E*-step, the integrals in the conditional expected complete data log-likelihood is approximated with numerical quadrature. In the *M*-step, the conditional expected complete data log-likelihood is maximized to yield updated item parameters. The two steps are repeated until the sequence of parameter estimates stabilize (convergence).

The item parameter estimates based on the Bock-Aitkin EM algorithm is often considered the gold standard in IRT. However, a major drawback of the EM algorithm is the “curse of dimensionality.” As the number of latent variables increases linearly, the total number of quadrature points grows exponentially. Thus, the Bock-Aitkin EM algorithm can be cumbersome for MIRT models.

Metropolis-Hastings Robbins-Monro (MH-RM) algorithm

[Cai \(2008, 2010a,b\)](#) proposed the Metropolis-Hastings Robbins-Monro (MH-RM) algorithm for maximum likelihood estimation in IRT models. The dominating insight of the MH-RM algorithm is that were the values of the latent variables (i.e., missing data) known, the complete likelihoods maximization would be easy. The MH-RM algorithm combines the Metropolis-Hastings (MH; [Metropolis et al., 1953](#); [Hastings, 1970](#)) sampler with the Robbins-Monro (RM) stochastic approximation algorithm ([Robbins and Monro, 1951](#)). Each iteration of the MH-RM algorithm involves three steps. In the *Stochastic Imputation* step, the MH sampler is used to simulate draws of the latent variable scores from their posterior predictive distribution to form complete data. In the *Stochastic Approximation* step, the gradient ascent directions of the observed data log-likelihood is computed indirectly from the Monte Carlo average of the gradient of the complete data log-likelihood, invoking the Fisher's Identity ([Fisher, 1925](#)). In the *Robbins-Monro Update* step, parameters are updated through the RM algorithm, which appropriately handles the Monte Carlo noise introduced in earlier step and guarantees pointwise convergence of the sequence of parameter estimates.

The MH-RM algorithm has been successfully applied in many IRT contexts (e.g., [Monroe and Cai, 2014](#); [Chalmers and Flora, 2014](#); [Yang and Cai, 2014](#)). It outperforms the Bock-Aitkin EM algorithm for high-dimensional problems in efficiency because it does not involve numerical quadrature.

Markov chain Monte Carlo (MCMC) approach

Markov chain Monte Carlo (MCMC; [Tierney, 1994](#)) involves drawing random samples with the help of a Markov chain from target distributions that are otherwise difficult to sample from. When applied in the fully Bayesian context, the target distribution is the joint posterior distribution of parameters. The samples from the Markov chain offer simple MCMC-based procedures have been proposed to estimate MIRT models (e.g., [Béguin and Glas, 2001](#); [Kim and Bolt, 2007](#); [Edwards, 2010](#); [Babcock, 2011](#)), building upon previous work on MCMC estimation of unidimensional IRT models (e.g., [Albert, 1992](#); [Patz and Junker, 1999a,b](#)). While applying MCMC algorithms, prior distributions of model parameters have to be specified. However, this has to be done with caution since the specification of priors will impact parameter estimates, even *non-informative* prior.

Scaled score based on response patterns

Once the items parameters are estimated, we can plug them into the MIRT model and make predictions of the latent variable θ based on an individual's responses to all I items, $\mathbf{u}$. Suppose the maximum marginal likelihood estimate is $\hat{\omega}$. From the Bayes rule, the posterior distribution of the latent variables is

$$P_{\hat{\omega}}(\theta|\mathbf{u}) = \frac{L_{\hat{\omega}}(\mathbf{u}|\theta)h(\theta)}{\int L_{\hat{\omega}}(\mathbf{u}|\theta)h(\theta)d\theta} = \frac{L_{\hat{\omega}}(\mathbf{u}|\theta)h(\theta)}{L_{\hat{\omega}}(\mathbf{u})} \quad (26)$$

where structural (item) parameters have been replaced with their maximum marginal likelihood estimates. Based on the posterior, we can compute the *Expected A Posteriori* (EAP) and *Maximum a Posteriori* (MAP) estimates by computing the mean and mode of the posterior. If a stochastic algorithm such as the MH-RM algorithm is used, multiple imputation or plausible values of θ may also be obtained ([Mislevy et al., 1992](#)).

Summed score to IRT scaled scores

An estimate of θ can be also obtained from the posteriors conditional on observed summed scores (Thissen and Wainer, 2001). The computation typically relies on the Lord-Wingersky algorithm (Lord and Wingersky, 1984), which is a recursive algorithm for computing summed score based likelihoods. A summed score to IRT scale score conversion table can be produced at the end, facilitating the use of IRT scoring without the more complex computation involved in IRT scoring based on response patterns. Cai (2015) extended the Lord-Wingersky algorithm to the case of bifactor and two-tier MIRT models.

Example applications and future directions

Educational measurement

Educational assessments routinely measure multi-faceted constructs, and subscores are of particular interest. Several approaches can be used to obtain subscores, including approaches based on classical test theory (CTT), unidimensional IRT and MIRT. Subscores derived through the MIRT approach are more accurate than those derived by CTT (Haberman and Sinharay, 2010). It is also shown that the MIRT-based approach provides more reliable subscores, overall scores (Yao, 2010), and more accurate classification information (Yao and Boughton, 2007).

MIRT models are also applied to longitudinal data to measure students' change over time. For example, Lockwood and Castellano (2015) proposed to apply the MIRT framework as an alternative to quantile regression to estimate the popular student growth percentile (SGP; Betebenner, 2009). The SGP is used in many states to help understand a student's academic progress by comparing his/her growth in scale score to the growth of students with similar prior year performance on standardized assessments (e.g., a student with SGP 75 indicates this student outperforms 75% of students who scored similarly in prior years). Monroe and Cai (2015) presented a straightforward method for estimating SGP reliability in the MIRT framework and evaluated how it is heavily impacted by the correlation between prior and current year score distributions.

te Marvelde et al. (2006) applied multidimensional generalized partial credit models to longitudinal survey data to measure students' change in three aspects as related to school effectiveness ("well-being at school", "academic self concept", and "attentiveness in the classroom"). Paek et al. (2016) pointed out the equivalence between MIRT models and longitudinal IRT models and elaborated how to analyze longitudinal categorical item response data with MIRT models. An MIRT-based approach is also proposed to monitor ability changes in computerized practice environments (Park et al., 2019).

Clinical outcomes assessment

A clinical outcome assessment (COA) typically evaluates health-related endpoints for clinical care, research, and trials. MIRT models, especially the bifactor model, play an important role in assessment development, short form construction, and refinement (e.g., Gibbons et al., 2009; DeWitt et al., 2011; Hansen et al., 2014; Shadel et al., 2014).

In COA, self-report questionnaires involving Likert-type items (e.g., strongly disagree, disagree, neutral, agree, strongly agree) are often used. Participants' response styles (e.g., tendencies to use endpoints, or the middle point), along with their genuine symptom severity levels, are both reflected in the item responses. One area in which MIRT can contribute further information is in explicit response style modeling. Falk and Cai (2016) proposed an approach based on the multidimensional NRM. This approach is further extended to multilevel data (Ju and Falk, 2019).

Future directions

In this article, we have reviewed recent developments in the MIRT literature. We conclude this article with discussions of future directions of research on MIRT. MIRT models discussed in this article assume that the individuals are uncorrelated. However, in educational and clinical settings, data usually exhibit multilevel structures with nesting or cross-classification. Thus, MIRT models for hierarchical and cross-classified data need to be studied. Current MIRT models also assume fixed item effects (i.e., fixed item parameters) and random person effects (i.e., θ follow a multivariate normal distribution). A new class of MIRT models with random item effects can be developed and would lead to new applications such as item bank development. In addition, we have assumed that all latent variables are continuous in this article. If the latent variables in MIRT models become discrete—we obtain the so-called diagnostic classification models (DCMs; Rupp et al., 2010), which have important research and practical implications in educational measurement.

References

- Albert, J.H., 1992. Bayesian estimation of normal ogive item response curves using Gibbs sampling. *J. Educ. Stat.* 17, 251–269.
- Andrich, D., 1978. A rating formulation for ordered response categories. *Psychometrika* 43, 561–573.
- Babcock, B., 2011. Estimating a noncompensatory IRT model using Metropolis within Gibbs sampling. *Appl. Psychol. Meas.* 35, 317–329.
- Béguin, A.A., Glas, C.A., 2001. MCMC estimation and some model-fit analysis of multidimensional IRT models. *Psychometrika* 66, 541–561.
- Betebenner, D., 2009. Norm- and criterion-referenced student growth. *Educ. Meas.* 28, 42–51.

- Bock, R.D., Aitkin, M., 1981. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika* 46, 443–459.
- Bock, R.D., Gibbons, R., Muraki, E., 1988. Full-information item factor analysis. *Appl. Psychol. Meas.* 12, 261–280.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37, 29–51.
- Bock, R.D., 1997. A brief history of item response theory. *Educ. Meas.* 16, 21–33.
- Bradlow, E.T., Wainer, H., Wang, X., 1999. A Bayesian random effects model for testlets. *Psychometrika* 64, 153–168.
- Cai, L., Hansen, M., 2018. Improving educational assessment: multivariate statistical methods. *Pol. Insights Behav. Brain Sci.* 5, 19–24.
- Cai, L., Yang, J.S., Hansen, M., 2011. Generalized full-information item bifactor analysis. *Psychol. Methods* 16, 221.
- Cai, L., Choi, K., Hansen, M., Harrell, L., 2016. Item response theory. *Annu. Rev. Stat. Appl.* 3, 297–321.
- Cai, L., 2008. A Metropolis-Hastings Robbins-Monro Algorithm for Maximum Likelihood Nonlinear Latent Structure Analysis with a Comprehensive Measurement Model (Unpublished Doctoral Dissertation). The University of North Carolina at Chapel Hill.
- Cai, L., 2010a. High-dimensional exploratory item factor analysis by a Metropolis-Hastings Robbins-Monro algorithm. *Psychometrika* 75, 33–57.
- Cai, L., 2010b. Metropolis-Hastings Robbins-Monro algorithm for confirmatory item factor analysis. *J. Educ. Behav. Stat.* 35, 307–335.
- Cai, L., 2010c. A two-tier full-information item factor analysis model with applications. *Psychometrika* 75, 581–612.
- Cai, L., 2015. Lord-Wingersky algorithm version 2.0 for hierarchical item factor models with applications in test scoring, scale alignment, and model fit testing. *Psychometrika* 80, 535–559.
- Chalmers, R.P., Flora, D.B., 2014. Maximum-likelihood estimation of noncompensatory IRT models with the MH-RM algorithm. *Appl. Psychol. Meas.* 38, 339–358.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B* 39, 1–22.
- DeWitt, E.M., Stucky, B.D., Thissen, D., Irwin, D.E., Langer, M., Varni, J.W., et al., 2011. Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory. *J. Clin. Epidemiol.* 64, 794–804.
- Edwards, M.C., 2010. A Markov chain Monte Carlo approach to confirmatory item factor analysis. *Psychometrika* 75, 474–497.
- Falk, C.F., Cai, L., 2016. A flexible full-information approach to the modeling of response styles. *Psychol. Methods* 21, 328.
- Fisher, R.A., 1925. Theory of statistical estimation. *Math. Proc. Camb. Phil. Soc.* 22, 700–725.
- Gibbons, R.D., Hedeker, D.R., 1992. Full-information item bi-factor analysis. *Psychometrika* 57, 423–436.
- Gibbons, R.D., Bock, R.D., Hedeker, D., Weiss, D.J., Segawa, E., Bhaumik, D.K., et al., 2007. Full-information item bifactor analysis of graded response data. *Appl. Psychol. Meas.* 31, 4–19.
- Gibbons, R.D., Rush, A.J., Immekus, J.C., 2009. On the psychometric validity of the domains of the PDSQ: an illustration of the bi-factor item response theory model. *J. Psychiatr. Res.* 43, 401–410.
- Haberman, S.J., Sinharay, S., 2010. Reporting of subscores using multidimensional item response theory. *Psychometrika* 75, 209–227.
- Hansen, M., Cai, L., Stucky, B.D., Tucker, J.S., Shadel, W.G., Edelen, M.O., 2014. Methodology for developing and evaluating the PROMIS[®] smoking item banks. *Nicotine Tob. Res.* 16, S175–S189.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 97–109.
- Holzinger, K.J., Swineford, F., 1937. The bi-factor method. *Psychometrika* 2, 41–54.
- Ju, U., Falk, C.F., 2019. Modeling response styles in cross-country self-reports: an application of a multilevel multidimensional Nominal Response Model. *J. Educ. Meas.* 56, 169–191.
- Kelderman, H., Rijkes, C.P., 1994. Loglinear multidimensional IRT models for polytomously scored items. *Psychometrika* 59, 149–176.
- Kim, J.-S., Bolt, D.M., 2007. Estimating item response theory models using Markov chain Monte Carlo methods. *Educ. Meas.* 26, 38–51.
- Lockwood, J., Castellano, K.E., 2015. Alternative statistical frameworks for student growth percentile estimation. *Stat. Public Policy* 2, 1–9.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.
- Lord, F.M., Wingersky, M.S., 1984. Comparison of IRT true-score and equipercentile observed-score “equatings”. *Appl. Psychol. Meas.* 8, 453–461.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092.
- Mislevy, R.J., Johnson, E.G., Muraki, E., 1992. Chapter 3: scaling procedures in NAEP. *J. Educ. Stat.* 17, 131–154.
- Monroe, S., Cai, L., 2014. Estimation of a Ramsay-curve item response theory model by the Metropolis-Hastings Robbins-Monro algorithm. *Educ. Psychol. Meas.* 74, 343–369.
- Monroe, S., Cai, L., 2015. Examining the reliability of student growth percentiles using multidimensional IRT. *Educ. Meas.* 34, 21–30.
- Muraki, E., Carlson, J.E., 1995. Full-information factor analysis for polytomous item responses. *Appl. Psychol. Meas.* 19, 73–90.
- Muraki, E., 1992. A generalized partial credit model: application of an EM algorithm. *Appl. Psychol. Meas.* 16, 159–176.
- Paek, I., Li, Z., Park, H.-J., 2016. Specifying ability growth models using a multidimensional item response model for repeated measures categorical ordinal item response data. *Multivariate Behav. Res.* 51, 569–580.
- Park, J.Y., Cornille, F., van der Maas, H.L., Van Den Noortgate, W., 2019. A multidimensional IRT approach for dynamically monitoring ability growth in computerized practice environments. *Front. Psychol.* 10, 620.
- Patz, R.J., Junker, B.W., 1999a. Applications and extensions of MCMC in IRT: multiple item types, missing data, and rated responses. *J. Educ. Behav. Stat.* 24, 342–366.
- Patz, R.J., Junker, B.W., 1999b. A straightforward approach to Markov chain Monte Carlo methods for item response models. *J. Educ. Behav. Stat.* 24, 146–178.
- Reckase, M.D., McKinley, R.L., 1991. The discriminating power of items that measure more than one dimension. *Appl. Psychol. Meas.* 15, 361–373.
- Reckase, M.D., 2009. *Multidimensional Item Response Theory*. Springer-Verlag, New York.
- Rijmen, F., 2010. Formal relations and an empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *J. Educ. Meas.* 47, 361–372.
- Robbins, H., Monro, S., 1951. A stochastic approximation method. *Ann. Math. Stat.* 400–407.
- Rupp, A.A., Templin, J.L., Henson, R.A., 2010. *Diagnostic Measurement: Theory, Methods, and Applications*. Guilford Press, New York.
- Samejima, F., 1969. Estimation of latent ability using a response pattern of graded scores. *Psychometr. Monogr.* 17.
- Shadel, W.G., Edelen, M.O., Tucker, J.S., Stucky, B.D., Hansen, M., Cai, L., 2014. Development of the PROMIS[®] coping expectancies of smoking item banks. *Nicotine Tob. Res.* 16, S202–S211.
- Takane, Y., De Leeuw, J., 1987. On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika* 52, 393–408.
- te Marvelde, J.M., Glas, C.A., Van Landeghem, G., Van Damme, J., 2006. Application of multidimensional item response theory models to longitudinal data. *Educ. Psychol. Meas.* 66, 5–34.
- Thissen, D., Cai, L., 2016. Nominal categories models. In: van der Linden, W.J. (Ed.), *Handbook of Item Response Theory*, vol. 1. Chapman and Hall/CRC, Boca Raton, FL, pp. 51–74.
- Thissen, D., Steinberg, L., 2009. Item response theory. In: Millsap, R., Maydeu-Olivares, A. (Eds.), *The Sage Handbook of Quantitative Methods in Psychology*. Sage, London, pp. 148–177.
- Thissen, D., Wainer, H.E., 2001. *Test Scoring*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Thissen, D., Cai, L., Bock, R.D., 2010. The nominal categories item response model. In: Nering, M., Ostini, R. (Eds.), *Handbook of Polytomous Item Response Theory Models: Developments and Applications*. Taylor & Francis, New York, NY, pp. 43–75.
- Thurstone, L.L., 1925. A method of scaling psychological and educational tests. *J. Educ. Psychol.* 16, 433.
- Tierney, L., 1994. Markov chains for exploring posterior distributions. *Ann. Stat.* 22, 1701–1728.
- Wirth, R., Edwards, M.C., 2007. Item factor analysis: current approaches and future directions. *Psychol. Methods* 12, 58.

- Yang, J.S., Cai, L., 2014. Estimation of contextual effects through nonlinear multilevel latent variable modeling with a Metropolis-Hastings Robbins-Monro algorithm. *J. Educ. Behav. Stat.* 39, 550–582.
- Yao, L., Boughton, K.A., 2007. A multidimensional item response modeling approach for improving subscale proficiency estimation and classification. *Appl. Psychol. Meas.* 31, 83–105.
- Yao, L., Schwarz, R.D., 2006. A multidimensional partial credit model with associated item and test statistics: an application to mixed-format tests. *Appl. Psychol. Meas.* 30, 469–492.
- Yao, L., 2010. Reporting valid and reliable overall scores and domain scores. *J. Educ. Meas.* 47, 339–360.
- Yung, Y.-F., Thissen, D., McLeod, L.D., 1999. On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika* 64, 113–128.

Explanatory measurement

Paul De Boeck^a and Mark Wilson^b, ^aThe Ohio State University, Columbus, OH, United States; and ^bUniversity of California Berkeley, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Explanatory measurement as an alternative	99
Examples	101
An example of modeling item difficulties	101
An example of modeling item loadings	101
Random respondent coefficients in the explanative models (random slopes)	101
Differences between the common approach and the explanatory approach	102
Sets of items as indicators	102
The effects of item covariates can be dimensions	102
Measure and explain at the same time	102
Descriptive and explanatory	103
Generality of explanatory measurement	103
Looking beyond data analysis: the construct map	103
Research issues	106
Software	106
References	107

Typically, the focus of measurement is on obtaining an estimate of a person's location on a construct. In contrast, explanatory measurement seeks to provide an explanation of the measurement. Thus, compared with the common measurement approaches, explanatory measurement offers extra opportunities for the explanation and validation of the measurement result. Explanation is a core activity of science (Hand, 2016) and measurement is considered an important preliminary condition for explanatory research. Explanatory measurement brings this perspective of explanation into the conceptualization of the statistical models used in measurement, so that the measurement can help lead to scientific advancements by shaping further explanatory research.

In psychology and the educational sciences, and in science in general, constructs play a crucial role. The measurement of constructs is often based on sum scores across a set of items that are formulated to reflect the targeted construct. The constructs can be cognitive abilities or non-cognitive variables, whether stable or not, and they may represent typical performance or maximum performance. In a statistical model used in measurement, the item responses are modeled to be a probabilistic function of an underlying latent variable that is conceptualized as the representation of a construct. Hence the estimates of an individual's latent variable value on a chosen scale function as measurements; that is, as estimates of the construct (e.g., cognitive abilities, etc.).

The common measurement practice can be characterized as follows:

1. A set of items is thought of as generating a set of indicators of the construct. This is the basic transduction in the measurement situation (Mari et al., 2021).
2. The measurements are averages of, or more complex functions of, the item responses. The scores obtained from the item responses indicate the average level of the responses (as in a sum score or a weighted average), or the scores reflect the most likely latent variable value given the responses depending on the model. Scores are summaries of the overall level of the responses (higher or lower), whether model based or not.
3. First measure, then validate. The conceptual link between items and construct is the basis to develop or select items, but the conceptual link needs to be empirically validated. In this separate step, one should test whether the item responses are in line with the conceptualized link using internal validity criteria such as internal consistency based reliability, internal structural consistency (e.g., difficulty order, factor loadings, etc.), goodness of fit of the model and/or the items, and using external validity criteria such as the relationship with external variables, the validity of which will be dependent on the assumed validity of the external structure (nomothetic span).
4. Measurement is descriptive. The scores indicate the level of the construct (e.g., a cognitive ability) without explaining why it is as high or low as it is found to be. To answer the deeper question of "why?," one needs to rely on external variables.

Explanatory measurement as an alternative

Explanatory measurement offers an alternative to this common approach—it provides the same outcomes, the scores or estimates of the construct, but goes well beyond that. The approach was laid out for item response modeling by Rijmen et al. (2003), and expanded on in a volume edited by De Boeck and Wilson (2004) on explanatory item response models (EIRM). Various models

are illustrated with verbal aggression data (Vansteelandt, 2000) using the NLMIXED module from SAS (SAS Institute, 1999). While this software was convenient at the time, the approach can be more readily implemented using newer software, as will be briefly discussed later. The explanatory approach is inspired by generalized linear (and nonlinear) mixed models (McCulloch and Searle, 2001) and related multilevel modeling (Goldstein, 2003; Raudenbush and Bryk, 2002), by Robert Sternberg's work on processes underlying intelligence using individual regression models to link item covariates to item responses (response times and accuracy) (Sternberg, 1977), and by Susan Embretson's covariate-approach for cognitive processes (Whitely, 1976, 1977). The approach is also in line with Ben Wright's insistence (Wright and Masters, 1981) on the importance of the interpretation of both item difficulties and person locations using item maps—commonly referred to as Wright maps (Wilson, 2017).

In terms of ability measurement based on binary data, the rationale underlying EIRM is that if one wants to know what makes, say, a cognitive ability item difficult (as reflected in the item difficulty parameter) and what makes an item more sensitive to the ability level (as reflected in the discrimination parameter), one can begin to answer questions such as why an individual's ability level is higher or lower and what is being measured beyond just referring to the specific set of items used in the calibration. For other types of constructs and for polytomous items, the rationale is analogous, with difficulty replaced with one or more thresholds.

All explanation is based on covariation, whether empirical or conceptual as in Wright maps. EIRM focuses on the covariation between properties of items and parameters of the same items, such as their intercept (difficulty, threshold, or thresholds) and slope (discrimination). In principle, the covariation can be tested in a second step after model estimation, but in EIRM the covariation is part of the model, which is why measurement implies explanation. In a similar way, person-level covariates can be included in the model to explain the person parameters (the latent variable values), but those are not discussed here; see Van den Noortgate and Paek (2004) for more information.

Let us start from the conditional logit for binary data assuming a two-parameter formulation:

$$\eta_{pi} = \alpha_i \theta_p - \beta_i \quad (1)$$

where η_{pi} is the conditional logit (conditional on θ_p but omitted in the notation) for respondent p ($p = 1, \dots, P$) and item i ($i = 1, \dots, I$),

θ_p is the latent variable value for respondent p (η_{pi} is conditional on θ_p) with $\theta_p \sim N(0, 1)$, which can be relaxed under certain conditions, for example, with a histogram distribution (Adams et al., 1997).

α_i is the item discrimination of item i , and

β_i is the item difficulty,

so that the conditional probability of a correct response (π_{pi}) can be written as:

$$\pi_{pi} = \frac{\exp(\eta_{pi})}{1 + \exp(\eta_{pi})} \quad (2)$$

In an EIRM focussing on explanation in terms of item difficulties, the item difficulty is modeled as:

$$\beta_i = \sum_{k=0}^K \beta'_k X_{ik} + \varepsilon_{\beta i} \quad (3)$$

with $X_{i0} = 1$ for all values of i , and β'_0 as the intercept, and where.

k is an index for item properties, $k = 0, 1, \dots, K$,

X_{ik} is the value of item i on item covariate k , and

$\varepsilon_{\beta i}$ is the residual term for the difficulty, $\varepsilon_{\beta i} \sim N(0, \sigma_{\beta}^2)$.

Without the residual term, the model is a fixed-effects model and implicitly assumes that the item difficulties can be fully explained by the item covariates (De Boeck, 2008; Janssen et al., 2004). But this is a strong assumption, seldom realized in empirical applications (see Wilson and Carstensen (2007) for one such example), hence the model with a residual term is more realistic. This leads to a crossed-random effects model because both respondents and the items are random.

In an EIRM, the item discrimination or the log of the item discrimination can also be modeled as a function of item properties. Eq. (4) is written without the log of the discrimination parameter:

$$\alpha_i = \sum_{k=0}^K \alpha'_k X_{ik} + \varepsilon_{\alpha i} \quad (4)$$

with the β notation replaced with an α notation. A similar distinction is made for the discrimination as for the difficulty between a fixed effects model and a random effects model. The model with fixed effects on the discrimination was first described by Embretson (1999). One complication, however, is that the item response model with a discrimination parameter can also be formulated as $\eta_{pi} = \alpha_i(\theta_p - \beta_i)$, which implies a different interpretation, not of α_i , but of β_i , with possible consequences for the explanation of the difficulty parameter in terms of item covariates (Smits et al., 2003). Using Eq. (1), $\exp(-\beta_i)/[1 + \exp(-\beta_i)]$ is the probability of a 1-observation if $\theta_p = 0$ or if $\alpha_i = 0$. No such interpretation is possible for the $\alpha_i(\theta_p - \beta_i)$ parameterization, and when $\alpha_i = 0$, the probability of a 1-observation is 0.50, which does not always make sense.

Examples

An example of modeling item difficulties

As an example of modeling the item difficulty, we briefly summarize the model and results for a one-dimensional mental addition test for first-grade students (Janssen et al., 1999). The data used for the model are speeded responses (with a time limit of 4 s) to 10 addition tasks with integers (from 1 to 4) and a maximum sum of 5: $1 + 1$, $2 + 1$, $1 + 2$, $2 + 2$, $3 + 1$, $1 + 3$, $4 + 1$, $1 + 4$, $3 + 2$, $2 + 3$ presented to (a) a normal ability group and (b) a disability group. The model chosen by the authors did not employ a discrimination parameter and the difficulty was modeled as a linear function of four item covariates based on the following item features:

- (1) whether one of the numbers is a 1,
- (2) whether the two numbers to be added are equal (a tie problem),
- (3) when the second number is larger than the first, the size of the second number, and
- (4) a covariate specific for the $1 + 2$ addition, because one can simply count 1, 2, 3 to find the correct response (i.e., it is very familiar to students).

An additional complexity of the model is an item-related group difference parameter to estimate a hypothesized effect that the five easiest items are easier for the normal ability group than for the disability group. This group difference parameter, specific to a subset of the items, is called a saltus parameter (Wilson, 1989). The model also contains an overall group effect. All four item covariates were found to have statistically significant effects. The effects of item covariates 1, 2, and 4 make the items easier, and the effect of item covariate 3 makes items more difficult. Together the covariates predict almost all of the differences between item difficulties. The correlation between the item difficulty estimates from the Rasch model (with a saltus parameter) and the predicted difficulties is 0.995, indicating that the model is doing a good job of encompassing the Rasch model item difficulty estimates; the likelihood ratio test comparing the explanatory model with the Rasch model (both with a saltus parameter) was not significant. Because the explanation of the item difficulties was nearly perfect, no residual term was needed.

An example of modeling item loadings

As a second example, we briefly summarize the model and results for a two-dimensional depression and anxiety questionnaire with 15 items (Xu et al., 2018). Here the model is not an item response model but a linear factor model with two factors, which illustrates that explanatory measurement can be generalized to multidimensional models. The item loadings (functioning as item discrimination parameters in factor analysis) are modeled as a linear function (not using the log as in Eq. 4) of two item covariates: distress and arousal (the relevance of each of these two aspects of affect, as rated for each of the 15 items). A random item residual term is added to accommodate for an imperfect explanation of the loadings. The item covariates were not used to explain the item intercepts because the item covariates do not contain level information but relevance information instead.

It was found that the loadings on one factor depend on the distress property of the items and the loadings on the other factor depend on the arousal property of the items. The model fit is comparable with a correlated confirmatory two-factor model with depression and anxiety as factors and better than a bifactor model with a general factor and group factors for depression and anxiety. The approach shows that measurement can be brought closer to theory than with confirmatory factor models (CFA). In a CFA, the confirmatory aspect refers to the exclusive association of groups of items with only one factor. The loadings on other factors are constrained to zero, so that there is only one item group per factor. With item covariates that directly refer to important affective features of depression and anxiety as specified in the core affect theory (Russell, 2003) the confirmatory model is at the same time more fine-grained and more flexible (because of the residual term), while preserving the measured constructs.

While the possible uses of such explanations of item parameters will vary due to the context of the application, two examples can be readily generalized to other situations as follows. For a classroom teacher knowing that some features tend to make items easier, she might start the class with examples that have these features (initially singly, but also in combination where that makes sense), and then move onto other examples with features that make the problem more difficult. For a test developer, the features related to item discrimination and item difficulty can be used to generate a new set of items intended to measure the same construct as with the original set, a parallel form, for example, for a pretest/posttest situation.

Random respondent coefficients in the explanative models (random slopes)

Explanatory measurement models, as presented in Eq. (3) and in principle also as presented in Eq. (4), can be extended with random respondent coefficients (i.e., random in p) reflecting the random effects of item covariates. This formulation is referred to as “random slopes” in the multilevel modeling literature; we will refer to it this way in what follows. This means that the β'_k differ depending on the respondent, which can be denoted with an extra subscript p : β'_{pk} . For example, for the mental addition case referred to above (Example 1), the effect of the size of the second number if it is larger than the first could be considered as a source of individual differences rather than as a uniform effect. To explain the corresponding cognitive process, we use a metaphor of weightlifting. To solve the mental addition task (such as $1 + 4$, where the second number is greater than the first) in an efficient

way, the second and larger number is “lifted” to be in front of the smaller first number (i.e., $1 + 4$ is transformed into $4 + 1$). The assumption is that the larger the “weight” to be lifted, the more power one needs to succeed (e.g., $1 + 4$ is more difficult than $1 + 3$, because a 4 is larger (“heavier”) than a 3). This random slope extension of an explanatory measurement model is a way to capture meaningful specific respondent abilities that can help to identify strengths and weaknesses in subcomponents of tasks.

Dimensions based on the coefficients of explanative models (i.e., “slopes-based”) are latent variables and have some interesting features:

- (1) Similar to cognitive diagnostic models (Rupp et al., 2010), they can fulfill a diagnostic role without the limitations of (i) the mixture modeling required for cognitive diagnostic models and (ii) the binary conceptualization of mastery inherent in those models.
- (2) The item covariate values fulfill the role of fixed discriminations for items on the latent variables.
- (3) The latent variables are effects of item covariates as independent variables. In item response models and in factor models the latent variables are commonly conceptualized as independent variables in a regression model with the item responses as the dependent variables, such that the discriminations (i.e., the loadings) are seen as the effects on the item responses. This conceptualization can be problematic because the latent variables are in fact derived from the item responses. Slope-based latent variables can solve this problem; in this new conceptualization, the independent variables are the item covariates while the latent variables are seen as their *effects* on the item responses, with the added flexibility that these effects can differ across respondents. In this way, the independent variables are no longer hypothesized variables. Instead, they are item covariates.
- (4) An interpretation of slope-based latent variables is more objective in that the interpretation does not depend just on a conceptual link in the mind of the test developer or item writer nor a link between a construct and a set of items. Instead, the item covariates themselves are seen as defining the latent variable. This can be empirically grounded and tested.

Differences between the common approach and the explanatory approach

Compared with the common measurement approach, EIRM can be characterized as follows.

Sets of items as indicators

Explanatory measurement shares the usual operationalization of a construct through a set of items, but it also goes beyond the set of items in that item covariates are specified. The additional requirement leads to a better elaboration of the construct. If one cannot specify for an ability construct the kind of difficulty that needs to be overcome, then the construct is not fully specified and the validity of the measure must be called into question.

This kind of difficulty is an item covariate, a property of the items. If the difficulties vary across items (which they almost always do), the hypothesis that this item property makes items more difficult can be tested. This requires that items be designed based on hypothesized difficulty covariates. The result is that the test can be thought as a repeated measurements “experiment” with a covariate for the repeated measures.

The effects of item covariates can be dimensions

Explanatory measurement offers the opportunity to go beyond the average of a respondent’s responses and incorporates a more sophisticated model-based indication of a construct than just the average response or respondent estimate in that the effect of item covariates on the item responses can also be measured. For example, for mental addition, some students may have more issues than others when the second number is larger.

This random slope implies a separate dimension, different from the one defined by a sum score or respondent estimate. Covariates with a random effect across persons define dimensions of the task that can be measured as random slopes.

Measure and explain at the same time

The validity question is formulated in a partially different way for explanatory measurement. Traditionally, a distinction is made between internal validity and external validity, with the former referring to the relationships between variables within a test (items, factors) and relationships with external variables, respectively. For example, internal consistency concerns the relationship between items, item fit indicates that an item fits with the latent variable as assumed by the model, and a finding that a confirmatory factor model for a scale fits well provides evidence for how the scale is conceptualized. In all or almost cases, these indications refer to relationships across respondents, of which correlations are the prototypical examples. As an example of external validity, convergent and discriminant validity is based on high correlations with alternate measurements of the same construct and low correlations with measures of different constructs. Similarly, SEM models can also be used to investigate external validity.

Our discussion of the role that explanatory measurement can play in validation is based on Embretson’s concepts of construct representation and nomothetic span (Whitely, 1983). Explanatory measurement adds something important to the investigation of

internal and external validity based on a different kind of relationship, one between item responses and item covariates. It provides evidence for internal validity in the sense that we can now specify the specific item properties that make an item more difficult and we can investigate whether the expected relationship (this time across items) is empirically supported. This is the construct representation aspect of internal validity in Embretson's terms. When there also are individual differences in the expected relationship, these can be related to external variables. This is the nomothetic span aspect of external validity in Embretson's terms. Thus construct representation and nomothetic span can be simultaneously investigated with an explanatory measurement approach.

An additional, interesting potential use of explanatory measurement is the investigation of possibly confounding aspects of items. If an item covariate that should not affect the item difficulty does in fact make items more difficult—especially when the effect is different depending on the respondent—this is a source of invalidity, and the items should be adjusted accordingly. See De Boeck et al. (2016) for an example.

Descriptive and explanatory

The common approach to the investigation of validity in measurement is to first measure and then explain. Often, the reason to measure is an explanatory research question. Once a construct is measured, one can think of testing a hypothesis in which the construct has an explanatory role, or another variable has an explanatory role vis-a-vis the construct. Explanatory measurement brings an explanatory component into the measurement itself, without an additional external step.

Generality of explanatory measurement

Although explanatory measurement has been discussed primarily for the Rasch family of models, without discrimination parameters (De Boeck and Wilson, 2004; De Boeck et al., 2011), the explanatory measurement idea can be implemented for any parameters of a measurement model, whether it is an item response model (IRT) or a factor model. In theory, any modeled parameter can be made a function of item covariates and person covariates: difficulties and intercepts, discriminations (slopes) and loadings, guessing parameters, person parameters (the latent variables), mixture probabilities, etc. The covariates can be item covariates, person covariates, and covariates of person-by-item pairs. As an example of the latter, De Boeck et al. (2016) have modeled word spelling responses with an item covariate that indicates whether the respondent knows the word in question or not, which is specific to person-by-item pairs. It also is possible to use response category covariates, such as the extremity of response categories on a Likert scale, whether the response category is a midpoint category, and the social desirability of the response category, etc. For multiple-choice response options covariates can also be considered. This then leads to response tree models (e.g., Böckenholt, 2012; De Boeck and Partchev, 2012; Jeon and De Boeck, 2016). Item tree models have not yet been framed as explanatory models in that way in the literature, but it is possible for sequential or continuation ratio models (e.g., Tutz, 1990) as well as for divide-by-total models (Masters, 1982).

Although, in principle, models can be extended to include explanatory components for their parameters, the model extensions can add too much complexity such that the practical feasibility is an important consideration. The most straightforward extensions are those with item covariates for difficulty parameters and person covariates for the latent variable, as noted above.

The explanatory function for a parameter either does or does not include a residual term. In practice it is rare that a full explanation or nearly full explanation can be obtained for a real data set. The earlier example for mental addition tasks is an exception where the explanation is nearly perfect, as is the situation described in the Wilson and Carstensen (2007) paper. A discussion of the issue is provided in De Boeck et al. (2016). The inclusion of a random residual tends to increase the standard error of covariate effects, so that a test of their effects is more stringent. Another consequence of introducing a residual term is that the model contains a random item effect (i.e., the residual item term) (De Boeck, 2008). As a result, the common maximum likelihood estimation methods are no longer feasible, and one needs to rely on approximate methods such as alternating imputation posterior estimation (Cho and Rabe-Hesketh, 2011), variational approaches (Rijmen and Jeon, 2013), or Bayesian estimation methods (Janssen et al., 2004).

Looking beyond data analysis: the construct map

The discussion about EIRM above may convey the impression that the main place in measurement where explanation is important is during the analysis of the data from the respondents interacting with items. But, indeed, the explanative expressions in Eqs. (3) and (4) above are useful only when the possibility of explanative features has previously been built into the thinking and planning about the item design. There are several critical steps in the cycle of measurement instrument development where this design strategy can be employed: (a) during the development of the underlying construct, (b) during the development of the item design, (c) during the coding and scoring of the responses from the respondents, and as extensively explained in the previous sections of this chapter, (d) during the analysis of the resulting data (These form the four building blocks of the BEAR Assessment System—Wilson, 2005). The impact of taking this explanative design strategy will take on a mainly qualitative rather than quantitative nature during the first three of these steps; hence, these design activities will generally not be expressed using equations as in (3) and (4) above, but nevertheless will take on a particular nature due to the aims of the strategy.

Although other foci may be possible, here we will focus on the expression of the explanatory strategy via the difficulty of the items involved, bearing in mind that, due to the nature of the modeling we are using, this will also inherently involve the ability of the respondents. To this end, a somewhat generalized strategy has been introduced, called a Construct Map, which embodies a developmental perspective on the latent variable in question. From this perspective the aim of the measurement is conceptualized not simply as a matter of quantitatively expressing an amount of knowledge, skills or attitude. Focussing on the measurement of knowledge or skills, one can see the underlying cognitive process being the addition of progressively higher levels of competence as new knowledge/skills are linked to existing knowledge/skills and as deeper understandings are developed from and take the place of earlier understandings. The underlying variable here needs to be derived in part from research into the hypothesized cognitive structure of the domain and in part from professional knowledge about what constitutes higher and lower levels of performance or competence but is also informed by empirical research into how respondents respond to training/instruction or perform in practice (NRC, 2001).

Consider the example of a construct map illustrated in Table 1. This Functions construct map is based on an important topic within the mathematics curriculum—making connections across multiple representations of mathematical functions (e.g., National Council of Teachers of Mathematics, 2000). The Functions construct map is based on the SOLO taxonomy (Biggs and Collis, 1982; Biggs, 1999), a general assessment-development framework that is used to assess student ability to make connections within and across topic areas. This general approach was readily applied to the area of mathematical functions, as the aim was to assess student ability to do just that for the topic of multiple representations of mathematical functions—see the successively complex levels represented by the rows in Table 1. The standard SOLO format of levels displays five levels of students' increasingly complex understanding. Here, this is modified by adding a lower and initial level called "Prealgebraic." Above that, going from bottom to top, the qualitatively differentiable levels represent increasing sophistication of *both* student understanding of the connections among mathematical functions *and* of relative difficulty of items designed to generate responses at these levels. For more information about the levels and to see examples of common student responses, see Wilmot et al. (2011, pp. 265–6).

Even though the distinctions among qualitatively different levels of sophistication is seen as crucial to this approach, so too is the expectation that, as a student is making progress toward those higher levels, their progress in between those levels can be discerned as being consistent with a continuous conception of development. Thus, although one might be able to design two items that are

Table 1 The functions construct map.

<i>Functions learning progression: Making connections across multiple representations of mathematical functions</i>		
<i>Levels of sophistication</i>	<i>Student is able to:</i>	<i>Response to items</i>
Extended Abstract (EA)	Make connections not only within a given subject area, but also beyond it, able to generalize and transfer the principles and ideas underlying the specific instance.	• Predict, explain and synthesize their understanding to a real-world context. • Solve non-routine problems including non-algorithmic functions.
Relational (R)	Demonstrate understanding of the significance of the parts in relation to the whole.	• Compare/contrast information given in multiple representations of functions to demonstrate understanding of the content. • Recognize which representations to choose for the context. • Select representations and move fluently among them to achieve a solution.
Multistructural (MS)	Make a number of connections, but the meta-connections between them are missed, as is their significance for the whole.	• Make connections across more than two representations (e.g., tabular, symbolic, graphic and verbal representations) • Recognize more than one feature of a functional relationship.
Unistructural (US)	Make simple and obvious connections, but their significance is not grasped.	• Make pairwise connections between representations.
Prestructural (PS)	Acquire bits of unconnected information, which have no organization and make no sense.	• Interpret graphs where both variables have to be interpreted, or where time is independent variable. • Demonstrate understanding of function in one representation.
Prealgebraic (PA)	Acquire prerequisite skills	Demonstrate understanding of: • Functional dependence, where a change in one variable affects another variable. • Continuous variables (time, distance) and dichotomous variables (hot or cold) • Relative order and measurement to define variables

Note: Read from bottom to top.

distinctively aimed at assessing two qualitatively different consecutive levels in a construct map, one can also design items that will fall at many points in between those distinct levels. Equivalently, two specific students may well generate responses that are typical of two consecutive levels, but other students can be found whose responses reflect many intervening positions. To embody this expectation of latent structure, and because the relationship of the item difficulty parameters to respondent ability is particularly straightforward for the Rasch-type models, the family of Rasch models is typically used to calibrate the items designed to match a construct map (e.g., see Wilson, 2005).

Apart from being a design principle used to create new items, and to code and score the responses from those items, a construct map can also serve as a principal way to interpret the respondent and item parameters that are the output from a test calibration. An example of how these are displayed in an item map (aka Wright map) is shown in Fig. 1. The first column of this Figure (from the far left) shows empirical locations for levels on the construct map—these are the combined levels of the Construct Map that were eventually found to be empirically supported (using an interpretative procedure known as “construct mapping” (Draney and Wilson, 2011)) as being distinguishable after the data have been calibrated and the results appraised. The second column indicates the estimated ability of the students, and difficulties for item thresholds, in logit units. The third column on this map shows the distribution of the students (as an “on-its-side histogram”) who took the Functions items in an administration of the test, and where each “X” in the histogram corresponds to 4 students. Then there are 8 more columns, one for each of the 8 items (named at the top of the 8 columns). The calibrated difficulty of the thresholds between the levels of each of the tasks is indicated by the entry within each of those columns. The Wright map allows one to estimate the probability of a student (or group of students) for each of the items, and, by extension, for the set (or “band”) of students in each of the construct map levels.

Considering now the interpretations of the resulting empirical construct levels, one can see for example, that the combined top construct map levels, EA and R, include only a few students and that for these they would be likely to give a Relational response to several of the tasks, for example, the “24 h crude oil” and the “100-years Crude Oil” tasks (Note: these item thresholds are denoted by “R” in the columns on the right hand side of the Figure). Below that, one can see that most of the students are located within the combined MS and US band, while the lowest levels, comprising the combined PS and PA band, are represented by fewer students.

This map then, allows for an explanation of student results on this set of items by showing how different locations on the map can be interpreted as relating to the likelihood of certain types of student response indicating levels of student understanding about mathematical functions. The outcomes are not so clear as one would wish however, as the detail in the hypothesized construct map has not been fully realized, but rather a grainier interpretation has been attained, distinguishing only 3 consecutive levels of interpretation. This can be ascribed in this case to (a) inconsistent difficulties of location of the MS and US thresholds in the middle bands of the construct map (which may be due to the distinction between these being not so important in this case as is assumed in the SOLO structure), and (b) fewer students being found at the two extremes (that is, the calibration sample might be more

College Readiness Assessment – Wright Map			Distribution of item difficulty (n=8)							
Level on the Connections Construct Map	Estimate	Distribution of student proficiency (n=688)	HEXAGON PATTERN	EQUIVALENT FUNCTIONS	EDUCATIONAL ACCOMPLISHMENTS	GENDER GAP	COST OF POSTAGE	STAIRCASE TOOTHPICK	24 HOUR CRUDE OIL	100 YEARS CRUDE OIL
Extended Abstract (EA) Relational (R)	4						R		EA	EA
	3					MS		MS		
	2	X XXX		R	R			US	R	R
Multistructural (MS) Unistructural (US)	1	XXXXXX XXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX -XXXXXXXXXXXXXXXXXXXX	EA R		MS	US	MS		MS	MS
	0	XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX	MS US	MS	US	PS	US PS	PS PA		US
	-1	XXXXXX XX XXX		PA PS	PS		PA		PS	
Pre-structural (PS) Pre-algebraic (PA)		X	PS PA						PA	
	-2	X			PA	PA				PS PA
EACH X REPRESENTS 4 STUDENTS, EACH ROW REPRESENTS .255 LOGITS										

Fig. 1 Wright map for the functions construct map.

informative if students were sampled from a wider grade level range). Note that this example has been chosen to reflect both strengths and weaknesses of the approach. Other examples of this approach may be found in [Duckor et al. \(2017\)](#), [Schwartz et al. \(2017\)](#), and [Wilson and Lehrer \(2021\)](#).

Research issues

There are several important research issues that deserve more attention to further develop explanatory measurement. Some of these issues are discussed here: (1) extensions to other parameters than item difficulty and to factor models, (2) goodness of fit measures, (3) diagnostic potential and relationships with cognitive diagnostic modeling, (4) test design.

1. Explanatory measurement has been widely used, primarily with item response models (IRT) and for modeling the difficulty parameter and the latent variable. For extensions to two-parameter models, a choice needs to be made between two formulations: $\alpha_i(\theta_p - \beta_i)$, and $\alpha_i\theta_p - \beta_i$, with consequences for the meaning of the β_i parameter and possibly for the precision of the β_i estimate as well. A greater challenge exists with respect to multidimensional models when items are associated with more than one dimension, in the so-called “within-dimension” models ([Adams et al., 1997](#)), or when none of the discriminations or loadings (i.e., in a factor model) are fixed to zero and the discriminations (or loadings) are modeled as a linear function of item covariates. This is especially a challenge when a residual term is used, as in the model of [Xu et al. \(2018\)](#). For this promising development to be more broadly used, an intense study of model identification is needed. Such a development may allow us to overcome the too heavy restrictions of confirmatory models which impose discrimination and loading values to be zero. While on the one hand, explanatory models of this type would be more flexible than confirmatory models, on the other hand, they would be more theory-based and confirmatory because all loadings are considered at least tentatively subject to linear effects of the item covariates.
2. Goodness of fit is an important issue for all models. It is rather straightforward to evaluate relative goodness of fit with AIC and BIC for maximum likelihood estimation, and with DIC and some more recently developed indices such as the WAIC and LOO-CV for Bayesian methods ([Vehtari et al., 2017](#)). If the models are nested, likelihood ratio methods can also be used for maximum likelihood results. It is less straightforward to evaluate absolute goodness of fit. One may consider bootstrap procedures and data recovery methods based on residuals combined with out of sample prediction, but it would also be helpful to develop limited information goodness of fit indices, such as RMSEA, as have been developed for item response models (e.g., [Maydeu-Olivarez, 2011](#)).
3. The random slopes of explanatory measurement models have the potential to capture components of item responses so that more differentiated information can be extracted from the item responses, which would be useful for diagnostic purposes. Thus far, explanatory measurement models have been used primarily for research purposes and for the evaluation of internal validity. With this formulation, we believe that explanatory measurement can fulfill the same role that cognitive diagnostic models aspire to. It would help if research and practice-based comparisons would be made between the two approaches. The comparison can be facilitated with the Q-matrix, which is, in essence, an item covariate matrix and thus can be used directly in explanatory measurement, although in a different way than in cognitive diagnostic models.
4. Related to the previous point and similarly for the cognitive diagnostic modeling approach, it will be beneficial to develop tests based on a test design. Test design has also been promoted by [Embretson \(1985, 1998\)](#), and also by authors associated with Principled Assessment Design PAD ([Wilson and Tan](#), this volume) based on the same ideas that underlay explanatory measurement and the use of covariates. For example, in Embretson’s writings and research, the focus is on cognitive processes underlying item responses, which is a highly valuable approach. Process oriented evidence is an important basis for the validation of tests, as also stressed in the Standards for Psychological and Educational Testing ([AERA/APA/NCME, 2014](#)). Moving to design-based test development would make explanatory measurement a more standard and feasible approach, but it also implies challenges and a need for further research and development, including regarding the practicality of test design and whether it can be combined with criteria such as content validity and representativeness of a domain.

Software

For the estimations of EIRMs, general purpose software is available as well as more specific IRT software. The following are general purpose software tools: The SAS NLMIXED procedure ([SAS Institute, 1999](#)) is used for all chapters of [De Boeck and Wilson, 2004](#)). The R package lme4 ([Bates et al., 2015](#)) can be used for models without a discrimination parameter, and is illustrated by [De Boeck et al. \(2011\)](#). Gllamm ([Rabe-Hesketh and Skrondal, 2016](#)) can also be considered, for example, see <http://www.gllamm.org/aggression.html>.

The following are more specific IRT software programs: Conquest ([Adams et al., 2020](#)) can estimate item covariate models for threshold parameters, as well as person covariates, using design matrices. EIRM ([Bulut, 2021](#)) estimates item covariate models for binary and polytomous item responses. FLIRT ([Jeon and Rijmen, 2016](#)) estimates item covariate models for both the difficulty and the discrimination parameters. Mirt ([Chalmers, 2012, 2015](#)) estimates item covariate models for item difficulty and discrimination with random item residuals, using a Metropolis-Hastings Robbins-Monro algorithm for model estimation, which is an advantage

for high dimensionality. TAM (Robitzsch et al., 2021) estimates item covariate models for item difficulty and item slopes. Finally, brms (Bürkner, 2017, 2021) is more general than IRT: It is a general Bayesian multilevel program and is the most general software package thus far for explanatory item response models.

References

- Adams, R.J., Wilson, M., Wang, W., 1997. The multidimensional random coefficients multinomial logit model. *Appl. Psychol. Meas.* 21, 1–23.
- Adams, R.J., Wu, M.L., Cloney, D., Wilson, M.R., 2020. ACER ConQuest: Generalised Item Response Modelling Software [Computer software]. Version 5. Australian Council for Educational Research, Camberwell, Victoria.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA, APA & NCME), 2014. Standards for Educational and Psychological Testing. AERA, APA & NCME, Washington, DC.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Software* 67 (1).
- Biggs, J.B., Collis, K.F., 1982. Evaluating the Quality of Learning: The SOLO Taxonomy. Academic Press, New York, NY.
- Biggs, J.B., 1999. Teaching for Quality Learning at University. SRHE/Open University Press, Buckingham, UK.
- Böckenholt, U., 2012. Modeling multiple response processes in judgment and choice. *Psychol. Methods* 17, 665–678.
- Bulut, O., 2021. Eirm: Explanatory Item Response Modeling for Dichotomous and Polytomous Item Responses. R package version 0.3.0. <https://CRAN.R-project.org/package=eirm>.
- Bürkner, P., 2017. Brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Software* 80 (1), 1–28.
- Bürkner, P.C., 2021. Bayesian item response modeling in R with brms and Stan. *J. Stat. Software* 100 (5).
- Chalmers, R.P., 2012. Mirt: a multidimensional item response theory package for the R environment. *J. Stat. Software* 48 (6), 1–29.
- Chalmers, R.P., 2015. Extended mixed-effects item response models with the MH-RM algorithm. *J. Educ. Meas.* 52, 200–222.
- Cho, S.-J., Rabe-Hesketh, S., 2011. Alternating imputation posterior estimation of models with crossed random effects. *Comput. Stat. Data Anal.* 55, 12–25.
- De Boeck, P., Partchev, I., 2012. IRTrees: tree based item response models of the GLMM family. *J. Stat. Software* 48 (1), 1–28.
- De Boeck, P., Wilson, M. (Eds.), 2004. Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach. Springer-Verlag, New York, NY.
- De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., Partchev, I., 2011. The estimation of item response models with the lmer function from the lme4 package in R. *J. Stat. Software* 39 (12), 1–28.
- De Boeck, P., Cho, S.-J., Wilson, M., 2016. Explanatory item response models: an approach to cognitive assessment. In: Rupp, A., Leighton, J. (Eds.), *Handbook of Cognition and Assessment*. Wiley Blackwell, Harvard, MA, pp. 249–266.
- De Boeck, P., 2008. Random item IRT models. *Psychometrika* 73, 533–559.
- Draney, K., Wilson, M., 2011. Understanding Rasch measurement: selecting cut scores with a composite of item types: the construct mapping procedure. *J. Appl. Meas.* 12, 298–309.
- Duckor, B., Draney, K., Wilson, M., 2017. Assessing assessment literacy: an item response modeling approach for teacher educators. *Pensamiento Educativo* 54, 1–26.
- Embretson, S.E. (Ed.), 1985. *Test Design: Developments in Psychology and Psychometrics*. Academic Press, New York, NY.
- Embretson, S.E., 1998. A cognitive design system approach to generating valid tests: application to abstract reasoning. *Psychol. Methods* 3, 380–396.
- Embretson, S.E., 1999. Generating items during testing: psychometric issues and models. *Psychometrika* 64, 407–433.
- Goldstein, H., 2003. *Multilevel Statistical Models*, third ed. Arnold, London, UK.
- Hand, D.J., 2016. *Measurement. A Very Short Introduction*. Oxford University Press, Oxford, UK.
- Janssen, R., De Boeck, P., Vaena, M., Vallaes, L., 1999. Simple mental addition in children with and without mild mental retardation. *J. Exp. Child Psychol.* 74, 261–281.
- Janssen, R., Schepers, J., Peres, D., 2004. Models with item and item group predictors. In: De Boeck, P., Wilson, M. (Eds.), *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. Springer-Verlag, New York, NY, pp. 189–212.
- Jeon, M., De Boeck, P., 2016. A generalized item response tree model for psychological assessments. *Behav. Res. Methods* 48, 1070–1085.
- Jeon, M., Rijmen, F., 2016. A modular approach for item response theory modeling with the R package FLIRT. *Behav. Res. Methods* 48, 742–755.
- Mari, L., Wilson, M., Maul, A., 2021. *Measurement Across the Sciences. Developing a Shared Concept System for Measurement*. Springer, New York, NY.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.
- Maydeu-Olivares, A., Cai, L., Hernandez, A., 2011. Comparing the fit of IRT and factor analysis models. *Struct. Equ. Model.* 18, 333–356.
- McCulloch, C.E., Searle, S.R., 2001. *Generalized, Linear, and Mixed Models*. Wiley, New York, NY.
- National Council of Teachers of Mathematics, 2000. *Principles and Standards for School Mathematics*. National Council of Teachers of Mathematics, Reston, VA.
- National Research Council (NRC), 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. In: Committee on the Foundations of Assessment, Pellegrino, J., Chudowsky, N., Glaser, R. (Eds.), Division on Behavioral and Social Sciences and Education. National Academy Press, Washington, DC.
- Rabe-Hesketh, S., Skrondal, A., 2016. Gilmm. In: van der Linden, W. (Ed.), *Handbook of Item Response Theory, Statistical Tools, and Applications*, vol. 3. Chapman & Hall, New York, NY.
- Raudenbush, S.W., Bryk, A.S., 2002. *Hierarchical Linear Models. Applications and Data Analysis Methods*. Sage, Newbury Park, CA.
- Rijmen, F., Jeon, M., 2013. Fitting an item response theory model with random item effects across groups by a variational approximation method. *Ann. Oper. Res.* 206, 647–662.
- Rijmen, F., Tuerlinckx, F., De Boeck, P., Kuppens, P., 2003. A nonlinear mixed model framework for IRT models. *Psychol. Methods* 8, 185–205.
- Robitzsch, A., Kiefer, T., Wu, M., 2021. TAM: Test Analysis Modules, pp. 7–16. R package version 3. <https://CRAN.R-project.org/package=TAM>.
- Rupp, A.A., Templin, J., Henson, R.A., 2010. *Diagnostic Measurement. Theory, Methods, and Applications*. Guilford, New York, NY.
- Russell, J.A., 2003. Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172.
- SAS Institute, 1999. *SAS Online Doc (Version 8) Software Manual on CD-Rom*. SAS Institute Inc, Cary, NC.
- Schwartz, R., Ayers, E., Wilson, M., 2017. Mapping a learning progression using unidimensional and multidimensional item response models. *J. Appl. Meas.* 18, 268–298.
- Smits, D., De Boeck, P., Hoskens, M., 2003. Examining the structure of concepts: using interactions between items. *Appl. Psychol. Meas.* 27, 415–439.
- Sternberg, R.J., 1977. Component processes in analogical reasoning. *Psychol. Rev.* 84, 353–378.
- Tutz, G., 1990. Sequential item response models with an ordered response. *Br. J. Math. Stat. Psychol.* 43, 39–55.
- Van den Noortgate, W., Paek, I., 2004. Person regression models. In: De Boeck, P., Wilson, M. (Eds.), *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. Springer-Verlag, New York, NY, pp. 167–187.
- Vansteelandt, K., 2000. *Formal Models for Contextualized Personality Psychology*. Unpublished doctoral dissertation. KU, Leuven, Belgium.
- Vehari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27 (5), 1413–1432.
- Whitley, S.E., 1976. Solving verbal analogies: some cognitive components of intelligence test items. *J. Educ. Psychol.* 68, 234–242.
- Whitley, S.E., 1977. Information processing on intelligence test items: some response components. *Appl. Psychol. Meas.* 1, 465–476.
- Whitley, S.E., 1983. Construct validity: construct representation versus nomothetic span. *Psychol. Bull.* 93, 179–197.
- Wilmot, D.B., Schoenfeld, A., Wilson, M., Champney, D., Zahner, W., 2011. Validating a learning progression in mathematical functions for college readiness. *Math. Think. Learn.* 13, 259–291.

- Wilson, M., Carstensen, C., 2007. Assessment to improve learning in mathematics: the BEAR Assessment System. In: Schoenfeld, A. (Ed.), *Assessing Mathematical Proficiency*. Cambridge University Press, London, UK.
- Wilson, M., Lehrer, R., 2021. Improving learning: using a learning progression to coordinate instruction and assessment. *Front. Educ.* 6, 654212.
- Wilson, M., 1989. Saltus: a psychometric model of discontinuity in cognitive development. *Psychol. Bull.* 105, 276–289.
- Wilson, M., 2005. *Constructing Measures: An Item Response Modeling Approach*. Erlbaum, Mahwah, NJ.
- Wilson, M., 2017. Things I learned from Ben. In: Wilson, M., Fisher Jr., W.P. (Eds.), *Psychological and Social Measurement: The Career and Contributions of Benjamin D. Wright*. Springer, New York, NY, pp. 75–82.
- Wilson, M., Tan, S., (to be published). Principled Assessment Design. In: *International Encyclopedia of Education, Quantitative Research / Educational Measurement*. Elsevier. In preparation.
- Wright, B.D., Masters, G.N., 1981. *Rating Scale Analysis*. MESA Press, Chicago, IL.
- Xu, M., De Boeck, P., Strunk, D., 2018. An affective space model for depression and anxiety. *Int. J. Methods Psychiatr. Res.* e1747.

Diagnostic classification modeling

Laine Bradshaw, Madeline Schellman, and Jiajun Xu, Mary Frances Early College of Education, University of Georgia, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

Determining the measurement domain	109
Determining the measurement model and developing items	110
The general latent class model	111
Specifying the structural model	111
Specifying the measurement model	111
DCM measurement model parameterization	112
The log-linear cognitive diagnosis model	113
Constraining the LCDM to yield common DCMs	114
Noncompensatory models	114
Compensatory models	115
Extended DCM measurement models	116
Polytomous DCMs	116
Hierarchical DCMs	116
Longitudinal DCMs	116
Measuring misconceptions	117
Conclusion	117
References	117

Diagnostic classification models (DCMs; Rupp et al., 2010) are classification-based psychometric models that efficiently group examinees by their levels of competency for the latent variables of interest. The DCM framework assumes that the latent variables, referred to as *attributes*, have categorical rather than continuous distributions. The discrete nature of the latent variables utilized in DCMs distinguishes them from other psychometric models such as those in the Item Response Theory (IRT) family of models. IRT models utilize continuous latent variables to characterize and estimate latent traits. The choice of whether to assume that the distribution of a latent construct is continuous or categorical depends on the assessment purpose and developers' and users' intended interpretations of the assessment results. An IRT model produces continuous scores along one or more latent continua allowing for rank ordering and comparing examinees' estimated positions on a given continuum. A DCM, on the other hand, produces statistically-driven groupings of examinees according to attribute states allowing for classifying examinees by sets of characteristics.

Diagnostic assessments—assessments that yield classifications of attributes via a DCM—are designed to provide more fine-grained detail about examinee characteristics with reasonable test lengths. Reliable classifications require shorter assessments than reliable continuous scales do; thus, utilizing DCMs in coordination with a careful diagnostic assessment design process enables a multidimensional assessment to remain reasonable in length. The intended use of diagnostic assessment results may be to better understand the multifaceted nature of a construct or to better inform next steps for tailoring an intervention or treatment. In education, a student's diagnostic result may be used to tailor instructional supports for areas identified as weaknesses, providing a more personalized and efficient path to learning.

To support intended diagnostic interpretations of examinee classifications, diagnostic assessments require an evidence-based assessment design process. This process begins with determining the measurement domain, or defining the construct of interest, and then includes developing assessment items to adequately represent the domain. Then, the DCM can be specified to model these attributes and item-attribute relationships, as well as estimate examinee classifications. DCM specifications include choices for modeling how attributes are related to each other, as well as how attributes behave to produce an item response. Once sample data is collected to calibrate the DCM, results provide empirical evidence to evaluate hypotheses of the specified attributes. Results may be used to adjust model specifications or may provide sufficient evidence to support the interpretations of the examinee classifications as meaningful groups according to the attribute states.

In this entry, we briefly discuss the process of specifying attributes and writing items to specify item-attribute relationships, and then we focus on specifying the DCM.

Determining the measurement domain

One of the first steps in constructing a diagnostic assessment requires developers to determine the construct(s) that they want to measure. Then, prior to item development, knowledge domain or content experts operationalize or define a set of attributes as the target construct(s). To accomplish this task, experts must identify a set of attributes that span the essential competencies that

an examinee typically needs in the target knowledge domain (e.g., Bradshaw et al., 2014). This set of attributes is the set of latent traits measured in the diagnostic assessment. For each examinee-attribute combination, the DCM will provide a diagnosis.

The diagnoses or groupings that result from DCMs are typically assigned labels that have meaning and guide decision making. The labels for the groupings depend on the nature of the attributes and the intended interpretation(s) of the classifications. Attributes can represent specific abilities, misconceptions, knowledge components, affective characteristics (e.g., motivation, attitudes, values), skills, or other characteristics that are defined *a priori* by content experts and, ideally, supported by cognitive theories (Leighton and Gierl, 2007). Although polytomous attributes (attributes with more than two levels) that each follow a categorical distribution may be utilized within a DCM framework (Bao, 2019), this entry assumes that attributes are binary and that each attribute follows a Bernoulli distribution. Binary attributes typically have contrasting labels such as “mastery” versus “non-mastery.” Alternate labels include (a) “present” versus “absent,” (b) “proficient” versus “emerging proficiency,” (c) “on-track” versus “needs attention,” (d) “competent” versus “needs improvement,” or (e) attribute *possession* versus no attribute possession (Bradshaw and Levy, 2019). This entry uses the labels “competency” and “non-competency” to depict the two attribute levels, or the two examinee groups for each attribute.

Classifications such as competent or non-competent refer to the examinee’s status on individual attributes. When considering the set of all attributes measured by the diagnostic assessment, the pattern of competencies across the set of attributes describes the examinee’s overall group classification. This collection of attribute-level classifications is referred to as an *attribute profile*, *attribute pattern*, or *latent class*.

To make these concepts more concrete, consider a diagnostic assessment that measures four attributes. This hypothetical assessment and its items will be used for illustration throughout this entry. The DCM will determine whether each examinee is competent or non-competent for each attribute. Then, each examinee will have four attribute-level classifications. For example, suppose the DCM determines that Examinee J is competent on Attribute 1, competent on Attribute 2, non-competent on Attribute 3, and competent on Attribute 4. Examinee J’s set of attribute-level classifications (competent, competent, non-competent, competent) represents a specific attribute profile.

More generally, when using only binary attributes, the total number of all possible unique attribute profiles is 2^A where A is the number of attributes measured by a diagnostic assessment. A DCM-based latent class, α_c , can be represented by an A -length vector: $\alpha_c = [\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{cA}]$, where $c \in \{1, 2, \dots, 2^A\}$ refers to the class number and $a \in \{1, 2, \dots, A\}$ refers to the attribute number. Each term α_{ca} is binary and has a value of 1 when class c is competent for attribute a and 0 when class c is non-competent for attribute a . For the example diagnostic assessment that measures four attributes, $A = 4$ and the DCM will assign each examinee to one of the 2^4 (i.e., 16) possible attribute profiles. These 16 profiles are represented by $\alpha_1 = [0000]$, $\alpha_2 = [0001]$, $\alpha_3 = [0010]$, $\alpha_4 = [0011]$, $\alpha_5 = [0100]$, $\alpha_6 = [0101]$, $\alpha_7 = [0110]$, $\alpha_8 = [0111]$, $\alpha_9 = [1000]$, $\alpha_{10} = [1001]$, $\alpha_{11} = [1010]$, $\alpha_{12} = [1011]$, $\alpha_{13} = [1100]$, $\alpha_{14} = [1101]$, $\alpha_{15} = [1110]$, and $\alpha_{16} = [1111]$. Therefore, Examinee J has the attribute profile $\alpha_{14} = [1101]$ because $\alpha_{14,1} = 1$, $\alpha_{14,2} = 1$, $\alpha_{14,3} = 0$, and $\alpha_{14,4} = 1$.

Because the number and nature of the attributes is determined *a priori*, the latent classes are hypothesized and specified prior to item development, item administration, and data analysis. Thus, DCMs are confirmatory latent class models because after fitting a DCM to collected response data, the results will provide evidence that supports or refutes the hypothesized specification of attribute profiles. Additionally, content experts can update or refine the theories from which the attributes were drawn to accommodate the empirical evidence obtained from DCM analyses (e.g., Bradshaw et al., 2014, 2017–2021; Jacobson et al., 2016). The cycle of creating theories, using theories to create assessments, and using assessment results to advance and improve theories results in stronger knowledge bases on which to build stronger assessments.

Specifying and confirming the latent dimensionality structure for multidimensional assessments is a complex endeavor. Initial empirical evidence may be utilized to refine attribute hypotheses. Updated hypotheses then may be evaluated using subsequent samples. Informing DCM specifications with empirical evidence in this way may be repeated until strong evidence for the model specifications is amassed through the iterative process.

Determining the measurement model and developing items

After specifying the target attributes and attribute profiles, developers and content experts create items to measure the attributes. Substantial knowledge of content and the target examinee population is required to write items that successfully elicit the intended attributes.

Item development highlights another way in which DCMs are confirmatory models: Like the attributes and latent classes, the item-attribute relationships are hypothesized and specified *a priori*, and the subsequent DCM-based data analysis provides evidence that supports or refutes the hypothesized item-attribute relationships. These hypothesized relationships are typically displayed in an array called the *Q-matrix* (Tatsuoka, 1983). The rows of the Q-matrix represent the n items included in the assessment, and the columns represent the A attributes measured by the assessment. Thus, the Q-matrix has the dimensions $n \times A$. If item i is hypothesized to measure attribute a , then the corresponding entry (i, a) in the Q-matrix has a value of 1. If item i is not hypothesized to measure attribute a , then the corresponding entry (i, a) in the Q-matrix has a value of 0. A q -vector for an item is the row of the Q-matrix that corresponds to the item. More specifically, q -vectors are A -length vectors of 0s and 1s, with each entry indicating whether the item measures the corresponding attribute.

The accuracy of the hypothesized item-attribute relationships is important as Q-matrix misspecification decreases the estimation accuracy of the model which prevents valid interpretation of assessment results (e.g., [Rupp and Templin, 2008](#)). Evidence to evaluate the hypothesized item-attribute relationships may be collected via content expert review and/or examinee interviews prior to specifying and estimating the DCM to strengthen the accuracy of the Q-matrix and, as a result, the accuracy of the DCM classifications.

The general latent class model

Classical Test Theory (CTT) is a sum score-based approach to psychological measurement that hinges on the theory that, for a given latent construct, each examinee has a true score with respect to a given construct. However, the construct is latent, so we cannot directly measure the true score in the same way that we can use a ruler to measure the length of a table. Therefore, we must rely on observed scores from assessments to estimate examinees' true scores, and we must accept that such an approach includes unavoidable uncertainty or error and do our best to minimize measurement error.

The goal with CTT is to estimate an examinee's true score. The goal with DCMs is to estimate, or diagnose, an examinee's level of competency with respect to each measured attribute. In other words, the goal with DCMs is to estimate an examinee's true latent class. Like a true score, we cannot directly observe an examinee's true latent class, so we rely on their scored item responses and use the responses to estimate the examinee's true latent class. This estimate, from a measurement perspective, is the class to which the examinee *most likely* belongs. DCMs compute the probability that an examinee belongs to each latent class, given the examinee's responses and the model's estimated parameters which include estimates of how prevalent the attribute profiles are and how items are related to attributes in the population. These probabilities are known as *posterior probabilities*. The examinee's most likely class is the class whose posterior probability is greatest. Based on each examinee's most likely class, one can compute the proportion of examinees who are in each latent class. Using the class proportions, one can compute the attribute *base rates*—the marginal proportion of examinees who are competent for each measured attribute.

DCMs determine posterior probabilities using the general latent class model, which is parameterized as

$$P(\mathbf{X}_e = \mathbf{x}_e) = \sum_{c=1}^{2^A} v_c \prod_{i=1}^n \pi_{ic}^{x_{ie}} (1 - \pi_{ic})^{1-x_{ie}} \quad (1)$$

In the equation, $\mathbf{x}_e$ is examinee e 's *item response vector*—the n -length vector that includes the observed, scored item responses for examinee e where $e \in \{1, 2, \dots, N\}$ and N is the total number of examinees in a sample. [Eq. \(1\)](#) computes the probability ($P(\cdot)$) of observing $\mathbf{x}_e$ given that the examinee belongs to latent class c . The probability that an examinee belongs to latent class c is denoted by the structural parameter v_c . The probability that an examinee in class c provides a correct response to item i is denoted by the measurement parameter π_{ic} . Examinee e 's response to item i is denoted as x_{ie} . This entry focuses on dichotomous (e.g., correct versus incorrect for cognitive domains; endorsement versus non-endorsement for affective domains) rather than polytomous (e.g., partial credit; nominal) item responses. Therefore, $x_{ie} = 1$ if examinee e 's response to item i is correct, and $x_{ie} = 0$ if examinee e 's response to item i is incorrect.

The general latent class model includes then two components: the *structural component* of the model with structural parameter estimates that describe the latent attribute distributions and relationships, or correlations, among the latent attributes, and the *measurement component* of the model with measurement parameter estimates that describe the item-attribute relationships. Each of these components requires its own model specification.

Specifying the structural model

The structural model in the DCM framework is analogous to models in the Structural Equation Modeling (SEM; e.g., [Kline, 2011](#)) framework where relationships among latent variables are modeled.

The structural parameter v_c represents the weight of latent class c among the other classes. Because each examinee is assigned to exactly one latent class, the latent classes are mutually exclusive. As a result, the class proportions sum to one ($v_1 + \dots + v_c + \dots + v_{2^A} = 1$).

[Rupp et al. \(2010\)](#) describe multiple ways to specify the structural model for a DCM. Options include using a log-linear model which parameterizes the latent attributes and their interactions for predicting the probability of latent classes, higher-order models which parameterize the latent attributes for predicting latent class membership through higher-order continuous variables, or an unstructured or a structured tetrachoric model.

Specifying the measurement model

The measurement model in the DCM framework links the observed response data and the unobserved latent attributes. A number of different options for a DCM measurement model have been developed. The choice of which model to use depends on the nature of the attributes and items and the desired interpretations of the diagnostic assessment results. The next section will describe some of the most commonly used DCM measurement models as well as extensions of these models that can be used in more complex contexts.

The measurement model parameterizes the *item response probability* (IRP) conditional upon class, represented by π_{ic} in [Eq. \(1\)](#). The IRP is the probability of answering an item correctly. Examinees with the same attribute profiles share the same IRPs; however,

depending on the number of attributes that an item measures, examinees who do not have identical attribute profiles may also have the same IRPs, as long as their profile is equivalent with respect to the measured attributes.

Consider again the example diagnostic assessment that measures four attributes. Suppose that Item 1 measures all four attributes. The A -length q -vector for Item 1 is [1111]. For a general DCM, because Item 1 measures all attributes included in the assessment, it has $2^4 = 16$ unique IRPs. Now suppose that Item 2 has the q -vector [0101]. Item 2 has only four unique IRPs because it only measures Attributes 2 and 4. Therefore, all attribute profiles that have the same competency states for Attributes 2 and 4 will have the same IRP for Item 2. In other words, it does not matter whether the attribute profile shows 0 or 1 for Attributes 1 and 3 because the item only measures Attributes 2 and 4. Table 1 demonstrates the conditional IRPs for Item 2. Notice that some attribute profiles are partially equivalent because they have the same competency states for Attributes 2 and 4 and, therefore, have the same IRPs for Item 2 regardless of their competency states for Attributes 1 and 3.

For non-general DCMs, an item's attribute behavior may be constrained to form other equivalence classes for the IRPs based on the constraint assumptions. For example, the chosen DCM for the example assessment might set a rule for Item 2 indicating that only examinees who are competent for both attributes measured by the item are likely to answer correctly while examinees who are competent for only one or neither of the two measured attributes have the same chance of answering correctly, and they are unlikely to answer correctly. Such a rule would reduce to a total of two IRPs for Item 2. With different model constraints applied to the item, different types of measurement models for DCMs are defined.

Local independence of items is a key assumption in item response models. In the DCM framework, local independence assumes that item responses are independent given the latent class. This assumption ultimately requires that the attribute(s) measured by an item completely explain the variation in examinee responses. In other words, local independence assumes that the Q -matrix captures all latent traits within the construct domain that influence an examinee's ability to respond to the item. The product part of Eq. (1) is only meaningful under the local independence assumption because of the multiplicative rule for independent events: the probability that independent Events X and Y (e.g., Event X is the selection of a correct response for Item 1 and Event Y is the selection of an incorrect response for Item 2) both occur is the product of their individual probabilities, $P(X) \cdot P(Y)$. As a result, correct specification of the Q -matrix becomes vitally important because the Q -matrix determines how class-specific probabilities are exhibited in the model. If the Q -matrix is misspecified, some variance in examinees' item responses is left unexplained, resulting in a violation of the local independence assumption (e.g., Rupp and Templin, 2008).

DCM measurement model parameterization

Different DCMs have been developed to have different approaches for modeling the item-attribute relationships and most have been shown to be nested within several general DCMs, contingent on model reparameterization. These flexible, general models can specify a wide variety of DCMs (de la Torre, 2011; Henson et al., 2009; von Davier, 2005). This entry focuses on one such model: the log-linear cognitive diagnosis model (LCDM; Henson et al., 2009). The first subsection shows the saturated form of the LCDM

Table 1 Conditional item response probabilities and corresponding attribute profiles for Item 2 with the q -vector [0101].

Item response probability	Abbreviated attribute profile	Attribute profiles included
1	[*0*0]	$c_1 = [0000]$ $c_3 = [0010]$ $c_9 = [1000]$ $c_{11} = [1010]$
2	[*0*1]	$c_2 = [0001]$ $c_4 = [0011]$ $c_{10} = [1001]$ $c_{12} = [1011]$
3	[*1*0]	$c_5 = [0100]$ $c_7 = [0110]$ $c_{13} = [1100]$ $c_{15} = [1110]$
4	[*1*1]	$c_6 = [0101]$ $c_8 = [0111]$ $c_{14} = [1101]$ $c_{16} = [1111]$

Note. The asterisk indicates that it does not matter whether the attribute profile shows 0 or 1 for Attributes 1 and 3 because the item only measures Attributes 2 and 4.

which can fully investigate how examinee responses are connected to latent attributes and their interactions. The second subsection demonstrates how the LCDM reparametrizes some previously-developed DCMs that, in isolation and with their original parameterizations, may appear unrelated.

The log-linear cognitive diagnosis model

In educational and psychological assessment, the LCDM models the log-odds of a correct response to item i for an examinee with the attribute profile for class c , α_c . Eq. (2) shows that a log-linear model expressed in its logit form is equivalent to the linear predictor $\lambda_{i,0} + \lambda_i^T h(\alpha_c, q_i)$.

$$\log\left(\frac{\pi_{ic}}{1 - \pi_{ic}}\right) = \log\left(\frac{P(x_{ie} = 1 | \alpha_c)}{P(x_{ie} = 0 | \alpha_c)}\right) = \lambda_{i,0} + \lambda_i^T h(\alpha_c, q_i) \quad (2)$$

Applying the exponential function to both sides of Eq. (2) and solving for the IRP, π_{ic} , yields the logistic regression form of Eq. (2):

$$\pi_{ic} = P(x_{ie} = 1 | \alpha_c) = \frac{\exp(\lambda_{i,0} + \lambda_i^T h(\alpha_c, q_i))}{1 + \exp(\lambda_{i,0} + \lambda_i^T h(\alpha_c, q_i))} \quad (3)$$

The linear predictor, called a *kernel*, is a combination of linear regression effects including intercepts, attributes' main effects, and higher-order attribute interaction effects. The kernel is written in vector notation to show its most general form. The vector λ_i^T contains all attribute main effects and higher order attribute interaction effects, which are similar to those in analysis of variance models where predictors are also categorical. These effects are item parameters estimated through the LCDM. Each effect has a set of corresponding attributes. For example, each main effect corresponds to one attribute, and each two-way interaction effect corresponds to two of the measured attributes.

The A -dimensional attribute profile for an examinee in class c is denoted as $\alpha_c = [\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{ca}, \dots, \alpha_{cA}]$. The A -dimensional q -vector for item i is denoted as $q_i = [q_{i1}, q_{i2}, \dots, q_{ia}, \dots, q_{iA}]$. The function $h(\alpha_c, q_i)$ reduces to a vector of the same length as λ_i^T that consists of 0s and 1s which are determined by whether item i measures the corresponding attributes (via q_i) and whether an examinee in class c is competent for the corresponding attributes (via α_c). The product of vector λ_i^T and vector $h(\alpha_c, q_i)$ results in a sum of the main and interaction effects, where the effects are retained in the equation only if item i measures all of the corresponding attributes and if examinees in class c are competent for all of the corresponding attributes. Eq. (4) gives the expanded form of the kernel.

$$\lambda_{i,0} + \lambda_i^T h(\alpha_c, q_i) = \lambda_{i,0} + \sum_{a=1}^A \lambda_{i,1(a)} (\alpha_{ca} q_{ia}) + \sum_{a=1}^{A-1} \sum_{a' > 1}^A \lambda_{i,2(a,a')} (\alpha_{ca} \alpha_{ca'} q_{ia} q_{ia'}) + \dots \quad (4)$$

As mentioned above, the kernel contains three types of parameters: an intercept, main effects, and interaction effects. The first term in Eq. (4) is the intercept parameter, $\lambda_{i,0}$, which gives the log-odds that an examinee who is not competent for any attribute measured by item i will provide a correct response to item i .

The second term in Eq. (4) includes the sum of the products of the main effect parameters, $\lambda_{i,1(a)}$, and their corresponding α_{ca} and q_{ia} statuses. Each main effect parameter indicates the increase in the log-odds of a correct response that comes with achieving competency for attribute a conditioned on being non-competent for the other measured attribute(s). Main effect parameters are reflective of an item's discrimination, or the item's ability to discriminate between different competency statuses for attribute a . Additionally, main effect parameters must be positive to adhere to the monotonicity constraint that examinees who are competent for attribute a have a greater IRP than do examinees who are non-competent for attribute a .

The third term in Eq. (4) includes the double summation of the products of the two-way interaction parameters, $\lambda_{i,2(a,a')}$, and their corresponding α_{ca} , $\alpha_{ca'}$, q_{ia} , and $q_{ia'}$ statuses. Each two-way interaction effect parameter indicates the change in the log-odds of a correct response that comes with achieving competency for both attributes a and a' . Notice that for the two-way (and higher-order) interaction effect parameters we cannot say that the change is an increase as it is for the main effect parameters because this is not always the case. Interaction effects may be positive, negative, or zero depending on how the attributes interact and impact examinees' IRPs. However, like main effect parameters, interaction parameters employ constraints that support the assumption that acquiring additional competencies with respect to the measured attributes cannot decrease an examinee's IRP.

The ellipses in Eq. (4) represent the three-way and other higher-order (up to A -way) interactions that can be incorporated into the model if appropriate.

The subscripts of linear lambda coefficients follow particular naming rules. The first subscript represents the item number. The second subscript represents the level of effect where "0" represents the intercept, "1" represents the main effects, "2" represents the two-way interaction effects, and so on. The third subscript indicates the attribute(s) that contribute to the effect. For example, $\lambda_{i,1,(a)}$ is the main effect of attribute a for item i and $\lambda_{i,2,(a,a')}$ is the two-way interaction of attributes a and a' for item i .

Consider the example diagnostic assessment that measures four attributes. The kernels for the items in this assessment could have up to four main effects, six two-way interaction effects, four three-way interaction effects, and one four-way interaction effect, as shown in Eq. (5):

$$\begin{aligned}
\text{kernel}_i = & \lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{c1}q_{i1} + \lambda_{i,1,(2)}\alpha_{c2}q_{i2} + \lambda_{i,1,(3)}\alpha_{c3}q_{i3} + \lambda_{i,1,(4)}\alpha_{c4}q_{i4} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}q_{i1}q_{i2} \\
& + \lambda_{i,2,(1,3)}\alpha_{c1}\alpha_{c3}q_{i1}q_{i3} + \lambda_{i,2,(1,4)}\alpha_{c1}\alpha_{c4}q_{i1}q_{i4} + \lambda_{i,2,(2,3)}\alpha_{c2}\alpha_{c3}q_{i2}q_{i3} + \lambda_{i,2,(2,4)}\alpha_{c2}\alpha_{c4}q_{i2}q_{i4} \\
& + \lambda_{i,2,(3,4)}\alpha_{c3}\alpha_{c4}q_{i3}q_{i4} + \lambda_{i,3,(1,2,3)}\alpha_{c1}\alpha_{c2}\alpha_{c3}q_{i1}q_{i2}q_{i3} + \lambda_{i,3,(1,2,4)}\alpha_{c1}\alpha_{c2}\alpha_{c4}q_{i1}q_{i2}q_{i4} + \lambda_{i,3,(1,3,4)}\alpha_{c1}\alpha_{c3}\alpha_{c4}q_{i1}q_{i3}q_{i4} \\
& + \lambda_{i,3,(2,3,4)}\alpha_{c2}\alpha_{c3}\alpha_{c4}q_{i2}q_{i3}q_{i4} + \lambda_{i,4,(1,2,3,4)}\alpha_{c1}\alpha_{c2}\alpha_{c3}\alpha_{c4}q_{i1}q_{i2}q_{i3}q_{i4}
\end{aligned} \tag{5}$$

Recall example Items 1 and 2 with the respective q -vectors [1111] and [0101]. Because Item 1 measures all of the attributes included in the assessment, $q_{11} = 1$, $q_{12} = 1$, $q_{13} = 1$, and $q_{14} = 1$. The Item 1 kernel retains all possible lambda parameters, yielding an equation identical to Eq. (5) where all of the q 's equal one. Because Item 2 measures only Attributes 2 and 4, $q_{21} = 0$, $q_{22} = 1$, $q_{23} = 0$, and $q_{24} = 1$. Therefore, every term in Eq. (5) that contains q_{i1} or q_{i3} equals zero. Eq. (6) gives the LCDM kernel for Item 2.

$$\text{kernel}_2 = \lambda_{2,0} + \lambda_{2,1,(2)}\alpha_{c2} + \lambda_{2,1,(4)}\alpha_{c4} + \lambda_{2,2,(2,4)}\alpha_{c2}\alpha_{c4} \tag{6}$$

As discussed above, Item 2 has four distinct IRPs. The first IRP is for examinees who are non-competent for Attributes 2 and 4 ($\alpha_{c2} = 0$ and $\alpha_{c4} = 0$), denoted as [*0*0]. The kernel used to compute the first IRP equals $\lambda_{2,0}$. The second IRP is for examinees who are non-competent for Attribute 2 but competent for Attribute 4 ($\alpha_{c2} = 0$ and $\alpha_{c4} = 1$), denoted as [*0*1]. The kernel used to compute the second IRP equals $\lambda_{2,0} + \lambda_{2,1,(4)}$. The third IRP is for examinees who are competent for Attribute 2 but non-competent for Attribute 4 ($\alpha_{c2} = 1$ and $\alpha_{c4} = 0$), denoted as [*1*0]. The kernel used to compute the third IRP equals $\lambda_{2,0} + \lambda_{2,1,(2)}$. Finally, the fourth IRP is for examinees who are competent for Attributes 2 and 4 ($\alpha_{c2} = 1$ and $\alpha_{c4} = 1$), denoted as [*1*1]. The kernel used to compute the fourth IRP equals $\lambda_{2,0} + \lambda_{2,1,(2)} + \lambda_{2,1,(4)} + \lambda_{2,2,(2,4)}$.

Suppose we fix $\lambda_{2,0} = -1.5$, $\lambda_{2,1,(2)} = 1$, $\lambda_{2,1,(4)} = 4$, and $\lambda_{2,2,(2,4)} = 1$ for Item 2. Table 2 shows the kernels and IRPs for the four different types of attribute profiles. As we can see, examinees who are non-competent for Attributes 2 and 4 have the lowest IRP while examinees who are competent for both Attributes 2 and 4 have the greatest IRP. The large main effect for Attribute 4 results in an increase in the IRP for examinees who have achieved competency for only Attribute 4 that is greater than the increase in the IRP for examinees who have achieved competency for only Attribute 2.

Constraining the LCDM to yield common DCMs

As described above, the LCDM is a general DCM that can yield other common DCMs when various parameters in the saturated LCDM are constrained. To demonstrate how the LCDM can be constrained differently for different items, this entry provides parameterization for three sub-models of the LCDM. These three models and others can be categorized into two groups: non-compensatory models and compensatory models. The critical difference between the two models is the characteristic, or the assumption, of attribute behavior. The attribute behavior assumption directs how the attribute competency contributes to the IRPs (i.e., π_{ic}).

Noncompensatory models

Noncompensatory models assume that non-competency in one attribute cannot be compensated for by competency in another attribute. Thus, the lack of one measured attribute is statistically equivalent to a lack of all measured attribute(s). Three well-known noncompensatory DCMs include the deterministic inputs, noisy "and" gate (DINA; e.g., Junker and Sijtsma, 2001) model, the noisy-inputs, deterministic "and" gate (NIDA; e.g., Junker and Sijtsma, 2001) model, and the noncompensatory reparametrized unified model (NC-RUM; e.g., Hartz, 2002).

The DINA model has exactly two IRPs and assumes monotonicity: the IRP for examinees who are competent on all measured attributes (*completely competent examinees*) is greater than the IRP for examinees who are not competent on all measured attributes (*non-competent examinees*). Therefore, achieving competency on all measured attributes in a noncompensatory model is the only way for an examinee to increase their IRP. To illustrate an item that follows DINA rules, consider the math problem: $\frac{1}{2} \left(\frac{3}{4} + \frac{7}{8} \right)$, which

Table 2 Kernels and item response probabilities for Item 2.

α_c	LCDM kernel	π_{ic}
[*0*0]	$\lambda_{2,0} = -1.5 = -1.5$	$\frac{\exp(-1.5)}{1 + \exp(-1.5)} = .18$
[*0*1]	$\lambda_{2,0} + \lambda_{2,1,(4)} = -1.5 + 4 = 2.5$	$\frac{\exp(2.5)}{1 + \exp(2.5)} = .92$
[*1*0]	$\lambda_{2,0} + \lambda_{2,1,(2)} = -1.5 + 1 = -0.5$	$\frac{\exp(-0.5)}{1 + \exp(-0.5)} = .38$
[*1*1]	$\lambda_{2,0} + \lambda_{2,1,(4)} + \lambda_{2,1,(2)} + \lambda_{2,2,(2,4)} = -1.5 + 1 + 4 + 1 = 4.5$	$\frac{\exp(4.5)}{1 + \exp(4.5)} = .99$

measures two attributes: fraction addition and fraction multiplication. If an examinee is competent on fraction addition but not fraction multiplication, their addition competency will not be enough for them to answer correctly. Likewise, if an examinee is competent on fraction multiplication but not on fraction addition, their multiplication competency will not compensate for addition non-competency, and the examinee is unlikely to answer correctly. A correct response to this item requires dual competency on fraction addition and fraction multiplication.

For the DINA model, only completely competent examinees are expected to answer an item correctly. To adhere to the noncompensatory nature of the DINA model, one can obtain the DINA model by constraining the LCDM main and lower-order interaction effect parameters to be zero, resulting in a kernel with only two parameters: the intercept and the highest order interaction effect parameter. Eqs. (7.1) and (7.2) give the LCDM parameterizations of the DINA model for example Items 1 and 2, respectively.

$$\text{kernel}_{1,\text{DINA}} = \lambda_{1,0} + \lambda_{1,4,(1,2,3,4)}\alpha_{c1}\alpha_{c2}\alpha_{c3}\alpha_{c4} \quad (7.1)$$

$$\text{kernel}_{2,\text{DINA}} = \lambda_{2,0} + \lambda_{2,2,(2,4)}\alpha_{c2}\alpha_{c4} \quad (7.2)$$

In the DINA framework, each item has exactly two kernels, which correspond to the two IRPs. The first kernel for Item 2, $\lambda_{2,0}$, is used to compute the IRPs for examinees with any of the 12 attribute profiles that do not have the partial competency $[*1*1]$. The second kernel for Item 2, $\lambda_{2,0} + \lambda_{2,2,(2,4)}$, is used to compute the IRPs for examinees with any of the four attribute profiles that have the partial competency $[*1*1]$. For Item 1, the second kernel, $\lambda_{1,0} + \lambda_{1,4,(1,2,3,4)}$, is used to compute the IRPs only for examinees with the attribute profile $[1111]$. The first kernel for Item 1, $\lambda_{1,0}$, is used to compute the IRPs for examinees with any of the other 15 attribute profiles.

Compensatory models

Unlike noncompensatory models, compensatory models assume that non-competency in one attribute can be compensated by competency in another attribute. Examinees lacking one attribute can still have a high IRP if they are competent for other attributes.

For example, an extreme case of a compensatory DCM, the deterministic inputs, noisy “or” gate (DINO; Templin and Henson, 2006) model, assumes that competency for one of the measured attributes is statistically equivalent to competency for multiple or all attributes. The DINO model is therefore also referred to as a fully *disjunctive* model, the “opposite” of DINA, a fully *conjunctive* model. The DINO model has exactly two IRPs and assumes monotonicity: the IRP for examinees who are competent for at least one measured attribute (*competent examinees*) is greater than the IRP for examinees who are not competent for any measured attributes (*completely non-competent examinees*). Therefore, examinees can increase their log odds of a correct response by achieving competency for any number of the measured attributes.

The LCDM parameterization of the DINO model imposes two constraints for each item: (1) an item’s main effect and interaction effect parameters are constrained to all have the same value, referred to as a *common main effect* and denoted as $\lambda_{i,1}$, and (2) the main effect parameters are constrained to be positive, the two-way interaction parameters are constrained to be negative, the three-way interaction parameters are constrained to be positive, and so on, alternating positive and negative signs for each parameter level. Eqs. (8.1) and (8.2) give the LCDM parameterizations of the DINO model for example Items 1 and 2, respectively.

$$\begin{aligned} \text{kernel}_{1,\text{DINO}} &= \lambda_{1,0} + \lambda_{1,1}\alpha_{c1} + \lambda_{1,1}\alpha_{c2} + \lambda_{1,1}\alpha_{c3} + \lambda_{1,1}\alpha_{c4} - \lambda_{1,2}\alpha_{c1}\alpha_{c2} - \lambda_{1,2}\alpha_{c1}\alpha_{c3} - \lambda_{1,2}\alpha_{c1}\alpha_{c4} - \lambda_{1,2}\alpha_{c2}\alpha_{c3} - \lambda_{1,2}\alpha_{c2}\alpha_{c4} \\ &\quad - \lambda_{1,2}\alpha_{c3}\alpha_{c4} + \lambda_{1,3}\alpha_{c1}\alpha_{c2}\alpha_{c3} + \lambda_{1,3}\alpha_{c1}\alpha_{c2}\alpha_{c4} + \lambda_{1,3}\alpha_{c1}\alpha_{c3}\alpha_{c4} + \lambda_{1,3}\alpha_{c2}\alpha_{c3}\alpha_{c4} - \lambda_{1,4}\alpha_{c1}\alpha_{c2}\alpha_{c3}\alpha_{c4} \\ &= \lambda_{1,0} + \lambda_{1,1}(\alpha_{c1} + \alpha_{c2} + \alpha_{c3} + \alpha_{c4} - \alpha_{c1}\alpha_{c2} - \alpha_{c1}\alpha_{c3} - \alpha_{c1}\alpha_{c4} - \alpha_{c2}\alpha_{c3} - \alpha_{c2}\alpha_{c4} - \alpha_{c3}\alpha_{c4} \\ &\quad + \alpha_{c1}\alpha_{c2}\alpha_{c3} + \alpha_{c1}\alpha_{c2}\alpha_{c4} + \alpha_{c1}\alpha_{c3}\alpha_{c4} + \alpha_{c2}\alpha_{c3}\alpha_{c4} - \alpha_{c1}\alpha_{c2}\alpha_{c3}\alpha_{c4}) \end{aligned} \quad (8.1)$$

$$\text{kernel}_{2,\text{DINO}} = \lambda_{2,0} + \lambda_{2,1}\alpha_{c2} + \lambda_{2,1}\alpha_{c4} - \lambda_{2,1}\alpha_{c2}\alpha_{c4} = \lambda_{2,0} + \lambda_{2,1}(\alpha_{c2} + \alpha_{c4} - \alpha_{c2}\alpha_{c4}) \quad (8.2)$$

Notice that these parameterizations can be written in the form $\lambda_{i,0} + \lambda_{i,1}(\cdot)$, where the alternating signs in $(\cdot)$ are designed so that the values for $(\cdot)$ are given by the indicator function in Eq. (9).

$$(\cdot) = \begin{cases} 1 & \text{for competent examinees} \\ 0 & \text{completely non-competent examinees} \end{cases} \quad (9)$$

Another commonly-used compensatory model is the compensatory reparametrized unified model (C-RUM; e.g., Hartz, 2002). In C-RUM items, achieving competency for one or more of the measured attributes would yield a greater IRP than would being competent for no attributes. The C-RUM uses the main effects to represent increases in the IRPs for being competent in additional attributes, but sets all interactions effects to zero. Eqs. (10.1) and (10.2) show the LCDM parameterizations of the C-RUM for example Items 1 and 2, respectively.

$$\text{kernel}_{1,\text{C-RUM}} = \lambda_{1,0} + \lambda_{1,1(1)}\alpha_{c1} + \lambda_{1,1(2)}\alpha_{c2} + \lambda_{1,1(3)}\alpha_{c3} + \lambda_{1,1(4)}\alpha_{c4} \quad (10.1)$$

$$\text{kernel}_{2,\text{C-RUM}} = \lambda_{2,0} + \lambda_{2,1(2)}\alpha_{c2} + \lambda_{2,1(4)}\alpha_{c4} \quad (10.2)$$

To illustrate the item response functions for the submodels, [Tables 3](#) and [4](#) show the specific DINA, DINO, and C-RUM kernels for Items 1 and 2, respectively, for each of the 16 attribute profiles in the example diagnostic assessment that measures four attributes.

Extended DCM measurement models

Thus far, this entry has discussed only dichotomous item responses and binary attributes with competent and non-competent diagnoses—labels that are most commonly associated with achievement testing. This section briefly introduces readers to extensions of the base DCMs that accommodate different types of observed variables and modeling frameworks.

Polytomous DCMs

DCMs have been developed to handle polytomously-scored responses ([Chen and de la Torre, 2018](#)), similar to other psychometric models such as graded or nominal item response models ([Kim et al., 2020](#)). The structural and measurement components of the DCM would need to change to accommodate polytomous responses. The structural model adopts a polychoric instead of a tetrachoric correlation. Depending on the contexts, the response variable in the measurement model may be ordinal or nominal. The distribution of the response variable, the link function connecting the response variable, and the linear predictor (the kernel) are then modeled accordingly.

Hierarchical DCMs

The models discussed thus far assume that attributes are not perfectly correlated. In fact, they assume that attributes are correlated but not so much that they cannot be distinguished from one another. However, DCMs have been developed that can manage complete dependency between attributes which is known as an *attribute hierarchy*. When achieving competency for an attribute completely relies on one's competency status with respect to another attribute, the two attributes are said to be part of an attribute hierarchy ([Templin and Bradshaw, 2014](#)). For example, consider a diagnostic assessment that measures two attributes: (1) addition of single-digit numbers and (2) addition of two-digit numbers. One cannot achieve competency for Attribute 2 without achieving competency for Attribute 1. Therefore, Attribute 2 completely depends on the competency of Attribute 1. Because this example assessment measures two attributes, it has four attribute profiles: [00], [01], [10], and [11]. However, the profile [01] is not possible because an examinee cannot be competent on Attribute 2 without being competent on Attribute 1. This profile can be removed, and the total number of possible profiles would reduce to three: [00], [01], and [11], which would also reduce the number of parameters that need to be estimated in the structural model, as well as have implications for the measurement model.

Longitudinal DCMs

Recent research has developed DCMs for longitudinal settings. The longitudinal scenario refers to evaluating examinees' knowledge components over time throughout an extended lesson or intervention, a semester, a school year, or other time frames. The longitudinal DCM can assess examinee growth in the form of transitions between attribute states, while providing diagnostic feedback about examinee understanding at each time point (e.g., [Madison and Bradshaw, 2018](#); [Wang et al., 2018](#); [Xu and Bradshaw, 2020](#)).

Table 3 DINA, DINO, and C-RUM kernels for Item 1 [1111] in the example diagnostic assessment.

Class	Profile	DINA	DINO	C-RUM
1	[0000]	$\lambda_{1,0}$	$\lambda_{1,0}$	$\lambda_{1,0}$
2	[0001]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(4)}$
3	[0010]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(3)}$
4	[0011]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(3)} + \lambda_{1,1(4)}$
5	[0100]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(2)}$
6	[0101]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(2)} + \lambda_{1,1(4)}$
7	[0110]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(2)} + \lambda_{1,1(3)}$
8	[0111]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(2)} + \lambda_{1,1(3)} + \lambda_{1,1(4)}$
9	[1000]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)}$
10	[1001]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(4)}$
11	[1010]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(3)}$
12	[1011]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(3)} + \lambda_{1,1(4)}$
13	[1100]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(2)}$
14	[1101]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(2)} + \lambda_{1,1(4)}$
15	[1110]	$\lambda_{1,0}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(2)} + \lambda_{1,1(3)}$
16	[1111]	$\lambda_{1,0} + \lambda_{1,4,(1,2,3,4)}$	$\lambda_{1,0} + \lambda_{1,1}$	$\lambda_{1,0} + \lambda_{1,1(1)} + \lambda_{1,1(2)} + \lambda_{1,1(3)} + \lambda_{1,1(4)}$

Note. Item 1 measures all 4 attributes on the example assessment.

Table 4 DINA, DINO, and C-RUM kernels for Item 2 [0101] in the example diagnostic assessment.

Class	Profile	DINA	DINO	C-RUM
1	[*0*0]	$\lambda_{2,0}$	$\lambda_{2,0}$	$\lambda_{2,0}$
2	[*0*1]	$\lambda_{2,0}$	$\lambda_{2,0} + \lambda_{2,1}$	$\lambda_{2,0} + \lambda_{2,1(4)}$
3	[*1*0]	$\lambda_{2,0}$	$\lambda_{2,0} + \lambda_{2,1}$	$\lambda_{2,0} + \lambda_{2,1(2)}$
4	[*1*1]	$\lambda_{2,0} + \lambda_{2,2,(2,4)}$	$\lambda_{2,0} + \lambda_{2,1}$	$\lambda_{2,0} + \lambda_{2,1(2)} + \lambda_{2,1(4)}$

Note. Item 2 measures Attribute 2 and Attribute 4 on the example assessment.

Longitudinal DCMs provide a framework to model theoretical learning trajectories with empirical evidence. These learning trajectories reflect when attribute competencies are acquired relative to each other and provide a discrete description about how examinees evolve from novice to sophisticated learners. Quantitatively vetting theories around how examinees learn, and possibly forget, can put forth new perspectives to curriculum-based learning trajectory research (e.g., Confrey et al., 2014).

Researchers and practitioners can also utilize longitudinal DCMs to evaluate the added value of instruction or learning materials over time and help investigate the effectiveness of an intervention using a pre-/post-test design.

Measuring misconceptions

Traditional assessments score responses as correct or incorrect without considering the specific selected response options. This approach discards potentially valuable information about examinee reasoning that can be found in their specific incorrect responses. A family of assessments called *concept inventories* (e.g., Hestenes et al., 1992) are specifically designed to include incorrect response options that target known plausible, yet incorrect reasoning in the population (*misconceptions*). DCMs have been developed to model misconceptions as the underlying categorical latent traits (de la Torre, 2009; DiBello et al., 2015; Kuo et al., 2016). Some of these models also estimate ability on an underlying continuum (Bradshaw and Templin, 2014) or diagnose skills (Kuo et al., 2018).

In education, identifying and correcting misconceptions as early as possible is critical because misconception can worsen over time (e.g., Fischbein and Schnarch, 1997) and impede learning of new concepts. Diagnostic concept inventories (Bradshaw, 2011; Bradshaw and Templin, 2014; Schellman, 2021) can help teachers identify which examinees possess misconceptions and use the model-based diagnoses to provide differentiated instruction.

Conclusion

This entry presented details of the DCM framework specifically through the lens of the LCDM—a flexible tool that can model multiple, categorical attributes, making it an important step in the process of compiling evidence to support or refute the validity of the intended assessment-based interpretations. Diagnostic assessments provide criterion-referenced information about examinees' individual strengths and weaknesses, which can be used to tailor remediation or interventions in ways that unidimensional information from scaling- or IRT-based assessments cannot. Assessment developers and users, however, must employ the psychometric framework that best supports the purpose of the assessment and the intended assessment-based interpretations, while aligning with practical constraints.

References

- Bao, Y., 2019. A Diagnostic Classification Model for Polytomous Attributes. Unpublished Doctoral Dissertation. University of Georgia, Athens, GA.
- Bradshaw, L.P., 2011. Combining Scaling and Classification: A Psychometric Model for Scaling Ability and Diagnosing Misconceptions. Unpublished Doctoral Dissertation. University of Georgia, Athens, GA.
- Bradshaw, L., Levy, R., 2019. Interpreting probabilistic classifications from diagnostic psychometric models. *Educ. Meas.* 38, 79–88.
- Bradshaw, L., Templin, J., 2014. Combining item response theory and diagnostic classification models: a psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika* 79 (3), 403–425.
- Bradshaw, L., Izsák, A., Templin, J., Jacobson, E., 2014. Diagnosing teachers' understandings of rational numbers: building a multidimensional test within the diagnostic classification framework. *Educ. Meas.* 33, 2–14.
- Bradshaw, L., Masters, J., Famularo, L., Lee, H., Azevedo, R., 2017–2021. Diagnostic Inventories of Cognition in Education (DICE) (Project No. R305A170441) [Grant]. U.S. Department of Education. Institute of Education Sciences.
- Chen, J., de la Torre, J., 2018. Introducing the general polytomous diagnosis modeling framework. *Front. Psychol.* 9.
- Confrey, J., Maloney, A.P., Corley, A.K., 2014. Learning trajectories: a framework for connecting standards with curriculum. *ZDM* 46 (5), 719–733.
- de la Torre, J., 2009. A cognitive diagnosis model for cognitively-based multiple choice options. *Appl. Psychol. Meas.* 33 (3), 163–183.
- de la Torre, J., 2011. The generalized DINA model framework. *Psychometrika* 76 (2), 179–199.
- DiBello, L.V., Henson, R.A., Stout, W.F., 2015. A family of generalized diagnostic classification models for multiple choice option-based scoring. *Appl. Psychol. Meas.* 39, 62–79.
- Fischbein, E., Schnarch, D., 1997. The evolution with age of probabilistic, intuitively based misconceptions. *J. Res. Math. Educ.* 28 (1), 96–105.
- Hartz, S.M., 2002. A Bayesian Framework for the Unified Model for Assessing Cognitive Abilities: Blending Theory With Practicality. ProQuest LLC, Ann Arbor, MI.

- Henson, R.A., Templin, J.L., Willse, J.T., 2009. Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika* 72 (4), 191–210.
- Hestenes, D., Wells, M., Swackhamer, G., 1992. Force concept inventory. *Phys. Teach.* 30, 141–158.
- Jacobson, E., Remillard, J.T., Hoover, M., Aaron, W., 2016. The interaction between measure design and construct development: building validity arguments. *J. Res. Math. Educ. Monogr.* 15, 155–174.
- Junker, B.W., Sijtsma, K., 2001. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Appl. Psychol. Meas.* 25 (3), 258–272.
- Kim, S.-H., Kwak, M., Bian, M., Feldberg, Z., Henry, T., Lee, J., Ölmez, İ.B., Shen, Y., Tan, Y., Tanaka, V., Wang, J., Xu, J., Cohen, A.S., 2020. Item response models in *Psychometrika* and psychometric textbooks. *Front. Educ.* 5.
- Kline, R.B., 2011. *Principles and Practice of Structural Equation Modeling*, fourth ed. Guilford, New York, NY.
- Kuo, B.-C., Chen, C.-H., Yang, C.-W., Mok, M.M.C., 2016. Cognitive diagnostic models for tests with multiple-choice and constructed-response items. *Educ. Psychol.* 36 (6), 1115–1133.
- Kuo, B.-C., Chen, C.-H., de la Torre, J., 2018. A cognitive diagnosis model for identifying coexisting skills and misconceptions. *Appl. Psychol. Meas.* 42 (3), 179–191.
- Leighton, J.P., Gierl, M.J., 2007. Why cognitive diagnostic assessment? In: Leighton, J.P., Gierl, M.J. (Eds.), *Cognitive Diagnostic Assessment for Education*. Cambridge University Press, New York, NY, pp. 3–18.
- Madison, M.J., Bradshaw, L.P., 2018. Assessing growth in a diagnostic classification model framework. *Psychometrika* 83 (4), 963–990.
- Rupp, A.A., Templin, J., 2008. The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educ. Psychol. Meas.* 68 (1), 78–96.
- Rupp, A.A., Templin, J., Henson, R.A., 2010. *Diagnostic Measurement: Theory, Methods, and Applications*. The Guilford Press, New York, NY.
- Schellman, M.A., 2021. *Diagnostic Concept Inventories for Misconception Classification Accuracy and Reliability*. Unpublished Master's Thesis. University of Georgia, Athens, GA.
- Tatsuoka, K.K., 1983. Rule space: an approach for dealing with misconceptions based on item response theory. *J. Educ. Meas.* 20, 345–354.
- Templin, J., Bradshaw, L., 2014. Hierarchical diagnostic classification models: a family of models for estimating and testing attribute hierarchies. *Psychometrika* 79 (2), 317–339.
- Templin, J.L., Henson, R.A., 2006. Measurement of psychological disorders using cognitive diagnosis models. *Psychol. Methods* 11 (3), 287–305.
- von Davier, M., 2005. *A General Diagnostic Model Applied to Language Testing Data*. Research Report. ETS RR-05-16. ETS Research Report Series.
- Wang, S., Yang, Y., Culpepper, S.A., Douglas, J.A., 2018. Tracking skill acquisition with cognitive diagnosis models: a higher-order, hidden Markov model with covariates. *J. Educ. Behav. Stat.* 43 (1), 57–87.
- Xu, J., Bradshaw, L., 2020. The Impact of Different Missing Data Patterns on Longitudinal Diagnostic Classifications. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, San Diego, CA.

Bayesian networks

Dakota W. Cintron, Evidence for Action, San Francisco, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	119
Graphical models	119
Bayesian networks	120
Maps	120
D-separation	121
Probability factorization	121
Message-passing, belief propagation, and probability updating	122
A discrete item response theory model	123
Additional topics in Bayesian networks	126
Evidence-centered design	126
From Bayesian networks to causal networks	126
Structure learning	126
Parameter learning	126
Summary	127
Appendix A	127
References	129

Introduction

Graphical models

The notion of *graphical models* appeared in the work of Speed (1978), Lauritzen et al. (1984), and Darroch et al. (1980). The group of authors developed a class of models—graphical models (Lauritzen, 1996)—with a general formalism, referred to as the Markov property, that enables explicit factorization (decomposition) of graphical models in terms of conditional independencies. According to Lauritzen (1996), graphical models have a fundamental and universal applicability due to three factors. First, graphs provide an easy way to visualize the scientific content of a given model as well as facilitate communication. Second, as a graph, scientific models are naturally modular, and as such, facilitate the description of complex problems, often as a particular combination of simpler elements. Third, graphs offer a useful data structure for communicating with digital computers.

Models as graphs have a long history. Early examples include Wright's path analysis (1934) and the representation of the total energy of a system (Gibbs, 1902)—gas or solid particles occupying a space as neighbors with interaction. The development of graphical models by Lauritzen and colleagues initially focused on undirected graphical models from contingency table data because of their direct formalism with Markov random fields. Kiiveri and Speed (1982) linked this general formalism to linear causal models that were popular in sociology (see Duncan, 1966; Heise, 1975). In the 1980s, Judea Pearl detailed several properties of graphical models (Pearl and Verma, 1987) and developed an algorithmic approach known as *d-separation* (Pearl, 1988) for determining whether any directed graphical model (without loops or cycles) implies any given conditional independence relation (Glymour, 2004).

This work by Pearl ultimately led to the development Bayesian Networks. Pearl initially referred to Bayesian networks as "belief networks, Bayesian networks, or influence networks, interchangeably, the former two to emphasize the judgmental origin and the probabilistic nature of the quantifiers, the latter to reflect the directionality of the links" (p. 246, Pearl, 1986). Bayesian networks have found use in several domains of research including artificial intelligence, computer science, economics, educational assessment, and causal inference (Almond et al., 2015; Hünemund and Bareinboim, 2019; Pearl, 1995; Quintana, 2020). The Markov property of graphical models has also contributed significantly to the development of computational and estimation techniques on graphs (e.g., see Lauritzen and Jensen, 2001; Cowell, 2005; Rijmen, 2008,2009; Rijmen et al., 2008).

In this chapter, we provide a basic overview of the notion and properties of Bayesian networks. We then discuss belief propagation, message-passing, and the recently developed R package, BayesNetBP, that facilitates belief propagation on Bayesian networks (Yu et al., 2020). BayesNetBP is the "first open source package to facilitate exact probabilistic reasoning in the three most common types of Bayesian networks (BNs) (discrete multinomial, continuous Gaussian and conditional Gaussian Bayesian networks)." (p. 2, Yu et al., 2020). We provide a motivating example using BayesNetBP for probabilistic reasoning in the form of a familiar multinomial Bayesian network known as the discrete item response theory (IRT) model (Almond et al., 2015). Finally, we conclude with comments on important topics in the field of Bayesian networks.

Before preceding to our discussion on Bayesian networks, it is important to note that several alternative types of graphical models are not discussed in this chapter. Graphical models generally fall into one of three domains: directed, undirected, and chain graphical models. Directed graphs include graphical models with directed edges and include the focus of this chapter, Bayesian

networks. Undirected graphs include graphical models with undirected edges and are known also as Markov random fields and include, for example, the Ising model or Boltzmann machine. Lastly, chain graphs include graphical models with both directed and undirected edges and are also known as partially directed acyclic graphs. While not discussed in this chapter, these types of graphical models are increasingly finding use in educational assessment and statistics. For example, the use of undirected graphs in educational assessment, statistics, and psychometrics has seen substantial growth in the past decade (e.g., see discussion on “network psychometrics” in [Cramer et al., 2010](#), [Epskamp et al., 2017](#); [Marsman et al., 2018](#); [Savi et al., 2019](#)).

Bayesian networks

Bayesian networks are a subclass of graphical models known as directed acyclic graphs (DAGs). DAGs are graphical models with only directed edges and no cycles. Acyclic refers to the structure of a directed graph defining a partial topological ordering of the nodes ([Scutari and Denis, 2014](#)). Bayesian networks as DAGs allow a concise decomposition of the probabilistic dependencies between a set of random variables $X = \{X_1, X_2, \dots, X_p\}$ that enables the joint distribution of the variables to be represented in a compact factored form. Let $G = (V, E)$ represent a DAG where each node $v_i \in V$ corresponds to a random variable X_i and E corresponds to the directed edges (also known as arcs) in the graph. If there is a direct edge from v_i to v_j , v_i precedes v_j in the topological ordering, and v_i is called an ancestor of v_j and v_j is called a descendant of v_i . If a path between two nodes is composed of only one arc, v_i is a parent of v_j and v_j is a child of v_i (see [Fig. 1](#)).

Maps

The correspondence between the graphical structure ($\perp_G$) and probabilistic structure ($\perp_P$) in Bayesian networks can be formalized in terms of three different mappings ([Nagarajan et al., 2013](#); [Scutari and Denis, 2014](#)). The mappings help categorize the relationship between a probabilistic model and a graph. A graph G is said to be an independence map (I-map) of the probabilistic structure P of X “if there is a one-to-one correspondence between the random variables in X and the nodes V of G , such that for all disjoint subsets A, B, C ” (p. 14, [Nagarajan et al., 2013](#))

$$A \perp_G B | C \Rightarrow A \perp_P B | C \quad (1)$$

That is, G is an I-map of P if B separates A and C in G implies A is independent of C given B in P . If a graph G is an I-map of the probability structure P , then P is said to be Markov with respect to G (i.e., variables separated in G are conditionally independent given the variables that separate them). A graph is said to be a dependency map (D-map) of P of X if

$$A \perp_G B | C \Leftarrow A \perp_P B | C \quad (2)$$

A D-map indicates that for any edge in a graph there is a dependence relationship between variables in the probability model ([Almond et al., 2015](#)). In contrast, in an I-map, there may be edges in a graph whether there are not actually dependencies. Lastly, a graph is said to be a perfect map of P if it is both an I-map and D-map. That is,

$$A \perp_G B | C \Leftrightarrow A \perp_P B | C \quad (3)$$

and we say that P is isomorphic or faithful to G . [Almond et al. \(2015\)](#) note that perfect maps are rare whereas I-maps and D-maps always exist. For instance, connecting all nodes in a graph (i.e., a complete graph) is a trivial I-map and a disconnected graph (no edges) is a trivial D-map ([Almond et al., 2015](#)). When we care about eliciting distributions it is typical to work with D-maps in the

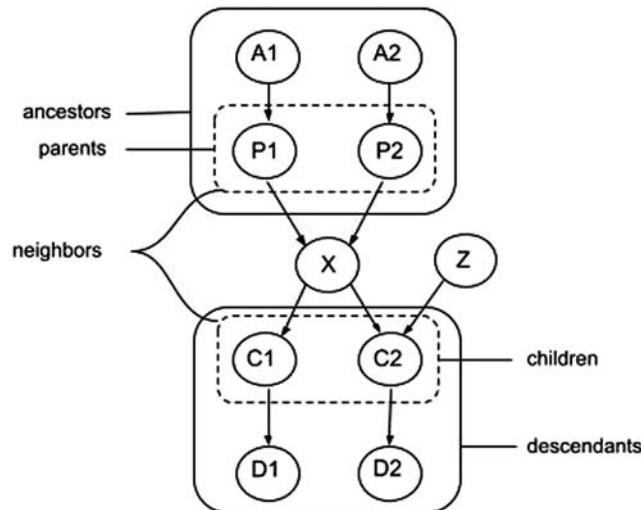


Fig. 1 An example of parents, children, ancestors, descendants, and neighbors of node X in a directed graph.

form of directed graphs (e.g., see the [Probability factorization](#) section where we factorize a graph according to its structure to represent the joint distribution). However, when constructing message-passing algorithms on Bayesian networks it is common to work with I-maps that transform directed graphs into an undirected representation for probability computation (e.g., an elimination tree or junction tree)—we return to the idea of probability computations on trees in the [Message-passing, belief propagation, and probability updating](#) section.

D-separation

The correspondence between the structure of a DAG G and the conditional independence relationships represented in the graph is formally codified by the directed separation (d-separation) algorithm (Pearl, 1988). The d-separation algorithm (here, d stands for directional) allows a computer to easily study independence properties for subgraphs. As such, the d-separation algorithm is a useful process for predicting the shared dependencies in a graph. A formal definition of d-separation is as follows, if A , B , C are three disjoint subsets of nodes in a DAG G , C is said to d-separate A from B , denoted $A \perp\!\!\!\perp_C B|C$, if along every sequence of edges (i.e., a path) between a node in A and a node in B , there is a node v meeting one of two conditions:

1. v is in C and does not have converging arc where C could be a chain ($A \rightarrow C \rightarrow B$) or fork ($A \leftarrow C \rightarrow B$); note that by being in C the node v is conditioned on, or
2. v has converging edges (i.e., there are two directed edges pointing to v) and none of v or its descendants are in C —known as a collider ($A \rightarrow C \leftarrow B$).

The d-separation criterion enables us to assess, for any pair of nodes, whether the nodes are d-connected—that is, there exists a connecting path between the nodes—or d-separated—there is no such path between a pair of nodes. Pearl et al. (2016) note that “When we say that a pair of nodes are d-separated, we mean that the variables they represent are definitely independent; when we say that a pair of nodes are d-connected, we mean that they are possibly, or likely, dependent” (p. 46). With the d-separation criterion, we can now evaluate complex Bayesian networks and determine whether variables are in the graphical model are independent or dependent, both marginally and conditional on other variables.

For example, in Fig. 1, consider the relationship between $C1$ and Z . Given we do not condition on any information (i.e., we have an empty conditioning set), we know that $C1$ and Z are d-separated and that $C1$ and Z are independent. This is because there is no unblocked path between $C1$ and Z . There is only one path between $C1$ and Z and it is blocked by a collider ($Z \rightarrow C2 \leftarrow X$). Conversely, if we now condition on $C2$, the d-separation criterion implies that $C1$ and Z are now d-connected—this is because $C2$ is in the conditioning set. In this case, since the path between $C1$ and Z contains a fork (X) that is not in the conditioning set, and the only collider $C2$ on the path is conditioned on, the path between $C1$ and Z is not blocked. Conditioning on colliders “unblocks” the path between variables. Note the same would be true if we conditioned on $D2$, because $D2$ is a descendant of a collider on the path between $C1$ and Z .

As we will see in the motivating example below, when we condition on a variable to probabilistically reason about the remaining variables in the graph, the d-separation criterion enables us to predict how new observed information flows (propagates) throughout the graph—the probabilities of the remaining d-connected variables in the graph are updated. Specifically, using a graph of the discrete IRT factor model (Fig. 2), we demonstrate (1) how we can observe evidence about item responses to probabilistically reason (or update our beliefs) about an individual’s ability and (2) how we can observe evidence about an individual’s ability to probabilistically reason (or update our beliefs) about that individual’s likelihood of answering an item correctly.

Probability factorization

The Markov property of Bayesian networks follows directly from the definition of d-separation (Nagarajan et al., 2013). The Markov property of Bayesian networks formalizes the representation of a joint probability distribution of the random variables in X (the global distribution) as a product of conditional probability distributions (the local distributions associated with each variable X_i). The Markov property is an application of the chain rule or total law of probability (Almond et al., 2015; Pearl, 2009). In

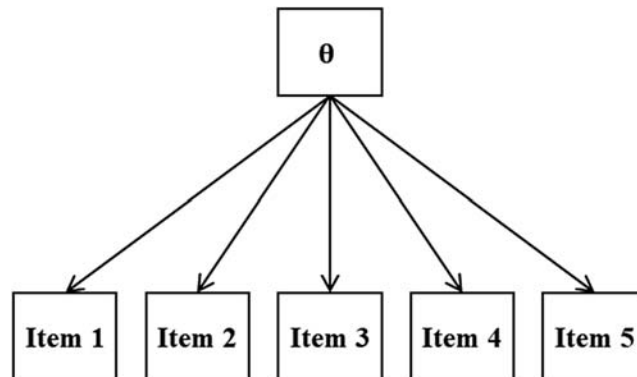


Fig. 2 Five-item discrete IRT Bayesian network. Squares are used to represent discrete nodes.

the case of discrete variables, the joint probability distribution can be decomposed or factorized as a product of conditional probabilities

$$P_X(X) = \sum_{i=1}^p P(X_i | pa(X_i)), \quad (4)$$

where $pa(X_i)$ is the set of the parents of X_i . In the case of continuous random variables, the joint density function f_X is given by

$$f_X(X) = \sum_{i=1}^p f_{X_i}(X_i | pa(X_i)), \quad (5)$$

Similar results can be derived for the case whether there are both discrete and continuous variables and thus mixed probability distributions. For further details on the properties of Bayesian networks we recommend (Scutari and Denis, 2014; Pearl et al., 2016).

Message-passing, belief propagation, and probability updating

Graphical models embed probability models with graphical structure. Graphical models, as a union of probability and graph theory, provide a visual and modular representation of probability models that facilitates flexible computation of probabilities. Message-passing algorithms (e.g., belief propagation or junction-tree algorithms) on Bayesian networks are basic examples of embedding probability models with graphical structure to compute probabilities. Given that we know the joint distribution from the graphical structure, in principle, we can always compute the effect of new information (observed evidence) for any variable, or subset of variables, on “our beliefs about the remaining variables (i.e., to propagate the evidence)” (p. 105, Almond et al., 2015).

As a simple example, consider the basic DAG ($X \rightarrow Y$). Let X represent a student’s proficiency in mathematics where X can take two values ($x_1 = \text{proficient}$ and $x_2 = \text{not – proficient}$), and Y represent whether the student provides a correct answer to a specific mathematics problem ($y_1 = \text{yes}$ and $y_2 = \text{no}$). Assuming the student is equally likely to be proficient or not-proficient means $p(\text{proficient}) = p(\text{not – proficient}) = 0.5$. Furthermore, assume that a proficient student’s probability of providing a correct answer is 0.6 and a not-proficient student’s probability is 0.3—that is, $p(y_1 | x_1) = 0.6$ and $p(y_2 | x_1) = 0.4$, and $p(y_1 | x_2) = 0.3$ and $p(y_2 | x_2) = 0.7$. In an educational assessment, we would like to observe a student’s performance on a mathematics problem, evaluate its correctness, and update our belief about the student’s proficiency status.

For instance, say we observe that Lynn correctly answers the mathematics problem (i.e., $Y = y_1 = \text{yes}$). With this simple DAG, we can use Bayes theorem to update our current beliefs regarding Lynn’s proficiency status (i.e., 0.5/0.5). In particular, now that we have the news that $Y = y_1$, our updated beliefs about Lynn’s proficiency status (proficient or not-proficient) are as follows

$$p(x_1 | y_1) = \frac{p(y_1 | x_1)p(x_1)}{p(y_1 | x_1)p(x_1) + p(y_1 | x_2)p(x_2)} = \frac{0.6 \times 0.5}{0.6 \times 0.5 + 0.3 \times 0.5} = 0.67 \quad (6)$$

and

$$p(x_2 | y_1) = 1 - p(x_1 | y_1) = 0.33. \quad (7)$$

Almond et al. (2015) demonstrate how these operations for this simple DAG also work on a simple set of conditional probability tables. We briefly note that these operations involve first starting with a table for the initial joint and marginal probabilities and a table for the observed evidence that $Y = y_1$. These initial tables are combined and normalized by $p(y_1)$ to obtain a conditional probability table for X given that $Y = y_1$ —the resulting table allows us to interpret the probabilities $p(x_1 | y_1)$ and $p(x_2 | y_1)$ (See pp. 107–109 of Almond et al., 2015, for further details on these operations).

With trivial examples like the DAG $X \rightarrow Y$, we can update our beliefs or probabilistically reason by hand using Bayes theorem. However, in more complicated DAGs, computer-assisted message-passing algorithms enable more rapid belief updating. The messages that are passed in these algorithms are updated probabilities based on observed evidence. However, the computations involved with updating probabilities on graphs becomes prohibitive as the number of variables in a Bayesian network increases. Almond et al. (2015) note, “A model with 15 variables with four values each ... means working with a joint probability table with over a trillion entries” (p. 105). Message-passing algorithms attempt to modularly represent the topology of probability models on graph structure such that calculations are more computationally efficient. Mathematical topologies that are amenable for computational efficiency are those with small cliques (or low tree width; Almond et al., 2015). Belief propagation is a message passing algorithm for inference on graphical models such as Bayesian networks. After a Bayesian network is constructed, the belief propagation algorithm may be applied to the network.

Almond et al. (2015) note that most Bayesian network software that implement message-passing algorithms operate in two modes: model construction/editing and model manipulation/inference. In the model construction mode, we (1) specify a node for each variable in the DAG and the names and the states of the variables, (2) draw edges between the nodes to depict the conditional dependence relationships in a model, and (3) supply a conditional probability table for each node in the model, given the parents of the node. If a node has no parents, an unconditional probability table is used. After the model is constructed, in the inference model, the model is compiled—that is our initial beliefs are propagated throughout the network—in several examples in Almond et al. (2015) a model is compiled using the Lauritzen and Jensen (2001) approach with junction trees in the commercial software Netica (Norsys). After the model is compiled, most Bayesian network software packages allow for immediate inspection of

the marginal distributions for all nodes. Furthermore, after evidence is propagated throughout the network (known as compiling), we are now ready to probabilistic reason through evidence absorption (i.e., we can make inquiries about the effect of observing or absorbing new evidence in the Bayesian network).

The BayesNetBP R package (Yu et al., 2020) is no different, and has both a model construction step and an inference or probabilistic reasoning step. BayesNetBP is a recently developed R package for implementing the belief propagation algorithm on Bayesian networks that was previously only available in the commercial software such as Hugin (Hugin Expert A/S, 2017) or Netica (Norsys). Each software performs belief propagation, but they differ in their approach or algorithms for implementing belief propagation. For example, Hugin, Netica, and BayesNetBP are based on Lauritzen and Spiegelhalter (1988), Lauritzen and Jensen (2001), and Cowell (2005), respectively. A key difference between the Lauritzen and Cowell approaches is that “message passing takes place on an elimination tree structure rather than the more compact (and usual) junction tree of cliques” (p. 1517). Yu et al. note that the approach of Cowell is “simpler to implement, and can be generalized to accommodate purely discrete or continuous BNs [Bayesian networks]” (p. 10). We do not provide further details on the differences between elimination trees and junction trees (see, Cowell, 2005). Furthermore, we do not cover important ideas such as cliques, elimination order, moralization, or triangulation (see, Yu et al., 2020). We note that the results provided by BayesNetBP in the motivating example below are roughly identical to Almond et al. (2015), albeit negligible differences are noted due to the fact that the local distributions were computed using simulated data. We encourage the reader to consult the references provided herein with particular focus on Chapters 4–6 in Almond et al. (2015).

A discrete item response theory model

Our motivating example for Bayesian networks uses a classical example of a discrete IRT model from Almond et al. (2015). We simulated data to conform with conditional probability table of the discrete IRT model in Almond et al. (see Table 6.1; Appendix A for code). The conditional probability table for the simulated data in this example is (Table 1):

All variables are treated as categorical. In this example, we have five items scored to yield observable values of right (1) or wrong (0). We also have a discrete measure $\theta = \{-2, -1, 0, 1, 2\}$ that represents a student’s proficiency in the mathematics knowledge and skills that this set of items requires. The items range from easy (item 1) to difficult (item 5) with the $\beta = \{-1.5 - 0.75, 0, 0.75, 1.5\}$.

A teacher may be interested in drawing inferences about the mathematics proficiency of each student based on the observed score patterns from this five item test. Given the assumption of local item independence, that is $X_j \perp\!\!\!\perp X'_j \mid \theta$, we can write the joint probability distribution over both the proficiency variable (θ) and evidence variables (items) as

$$P(X_1, \dots, X_J, \theta) = P(\theta) \sum_{j=1}^J P_j(X_j \mid \theta). \quad (8)$$

The factorization structure of the joint probability distribution is represented by the Bayesian network in Fig. 2. Note that a classic IRT model treats the θ variable as continuous. As such, there would be directed edges from the continuous θ to the discrete items which indicates that the standard IRT model does not fall into the computationally convenient class of conditional Gaussian Bayesian networks (see Lauritzen, 1992, and Lauritzen, 1996, for more details). In Fig. 2, since all the variables are discrete, there are no integrals that need to be solved numerically. Almond et al. (2015) note that this discretization of θ may not be that problematic citing Haberman’s (2005) demonstration of how an ordered latent class model (Bayesian network in Fig. 2) and a unidimensional IRT model both fit multiple data sets equally well.

Bayesian networks and belief propagation allow (1) the visual representation of domain information (in our example θ , or proficiency, influences item response behavior), (2) encoding of domain information in terms of conditional independencies on a joint distribution of variables (here, the model encodes the typical IRT assumption of local independence—that is the item responses are independent conditional on θ), (3) simple specification of the joint distribution (see Eq. 8), and (4) rapid computation and updating of given probabilities of interest after observing evidence—Almond et al. (2015) describes this as the facility to “enter data and

Table 1 Conditional probabilities of a correct response for the five-item IRT model in Almond et al. (2015).

θ	Item 1	Item 2	Item 3	Item 4	Item 5
−2	0.3770	0.2150	0.1100	0.0650	0.0290
−1	0.6230	0.4440	0.2480	0.1445	0.0820
0	0.8050	0.6788	0.5005	0.3195	0.1850
1	0.9255	0.8540	0.7460	0.5670	0.3790
2	0.9660	0.9530	0.8180	0.7820	0.6170

play ‘what-if’ games for various hypotheses” (p. 163). In the next few sections, we demonstrate how to use the BayesNetBP package to answer identical questions formulated in Almond et al. (2015). R code is provided in supplemental material (see Appendix A).

In the motivating example below, we first inspect initial marginal probabilities that represent our “best predictions of how a student who we know nothing about, other than that the student is representative of the population for which the model was built, would do on a quiz” (p. 161, Almond et al., 2015). These initial marginal probabilities are in Fig. 3. These marginal probabilities are calculated from the conditional probabilities of a correct response given θ and the initial marginal distribution for θ using the law of total probability. The marginal probabilities are simply the expected proportion correct on each item in a population for whom the quiz was designed (Almond et al., 2015). The marginal probabilities for the items indicate that Items 1 and 2 are relatively easy items with expected proportion correct greater than 60%. On the other hand, Item 4 and 5 are relatively difficult items with expected proportion correct less than 35%.

An advantage of Bayesian network software, versus doing the calculations by hand, is that we can quickly observe evidence on the graph (i.e., absorb or instantiate evidence on the graph). This is the facility to enter data and play “what-if” games for various hypotheses. The first example we demonstrate this facility on is absorbing evidence from the simple five item mathematics quiz. For instance, suppose that Merve, a student in a class taking this mathematics quiz, answers Item 2 and Item 3 correctly (i.e., Item 2 = Right and Item 3 = Right). Note that we have shaded these items gray in Fig. 4 to indicate that they are now observed. We can absorb this new evidence into the graph using BayesNetBP that immediately performs belief propagation. The result is in Fig. 4. The result produces a posterior distribution for Merve’s θ . In particular, we can see that Merve answering Item 2 and 3 correctly has shifted our beliefs in her θ such that there are now higher probabilities for higher values of θ —for instance, we can see that the probability that Merve’s θ is -2 is nearly zero. Furthermore, we can see that the probability that Merve would answer the remaining items correct has also increased—Merve’s probability of answering Item 1 correctly has shifted to nearly 90%. If instead we had observed that Merve answered Items 3 and 4 correctly, we would expect an even greater shift toward higher values of θ since Items 3 and 4 are more difficult than Items 2 and 3 collectively. Note, in this example, we can immediately see that all the variables in Fig. 2 are d-connected because after observing new evidence, all probabilities in the graph were updated. If there were d-separated variables (nodes) in the graph, the probabilities for those nodes would not be updated.

Likewise, we can evaluate the effect of observing evidence about θ —note in practice θ is typically unobservable but Bayesian networks enable us to specify θ values in order to test out various scenarios. For instance, say we know that there are groups of students in the class for which this quiz was designed that had different θ values (e.g., Group A’s θ values are 2 and Group B’s θ values are 1, etc.). We can use the Bayesian network to answer the question, what is the expected performance on each of the items for a member from Group B? In BayesNetBP, we can set $\theta = 1$ and propagate this information through out the Bayesian network to update our beliefs about the probability of a member from Group B answering a given item correctly. Note that we have shaded θ gray in Fig. 5 to indicate that it is now observed. Fig. 5 reports the results of absorbing $\theta = 1$ in the BayesNetBP package (i.e., the updated item probabilities). We can see that a member from Group B is very likely to answer Items 1–3 correctly (probabilities all greater than 75%). However, while the probability of answering Item 4 correctly has increased from the expected proportion correct in the population this test was designed for, Group B members still only have a roughly 50–50 chance of answering the item correctly. Item 5 (the most difficult item) still has quite a low probability of being answered correctly by members of Group B.

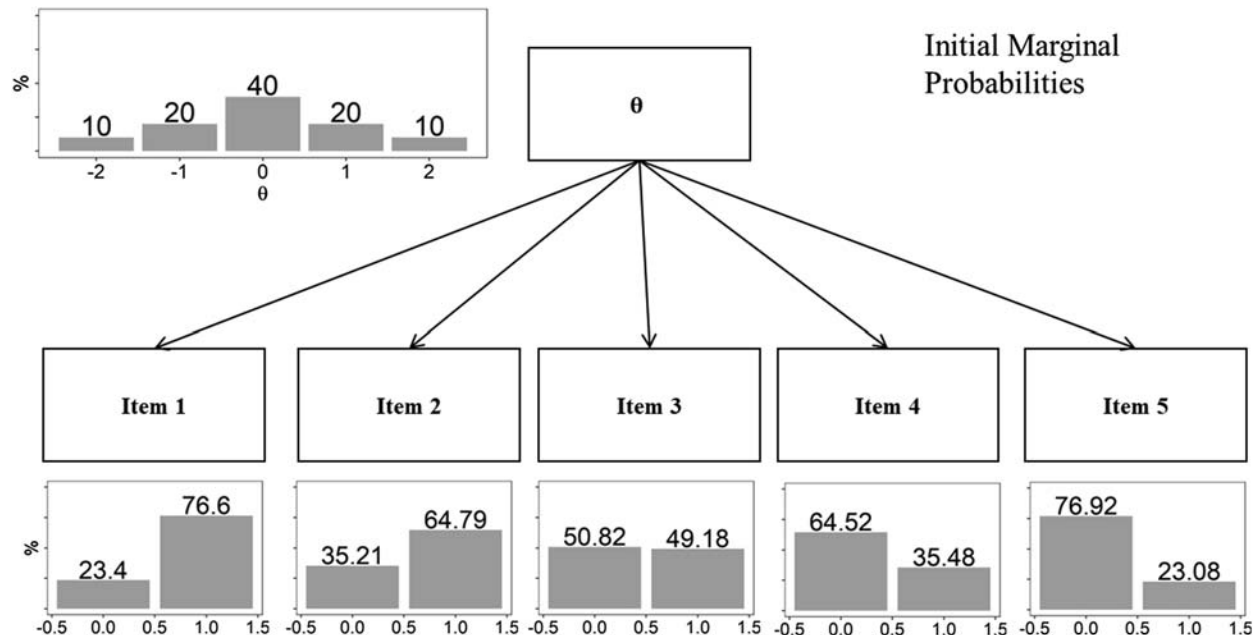


Fig. 3 Initial marginal probabilities and propagating prior evidence throughout the Bayesian network.

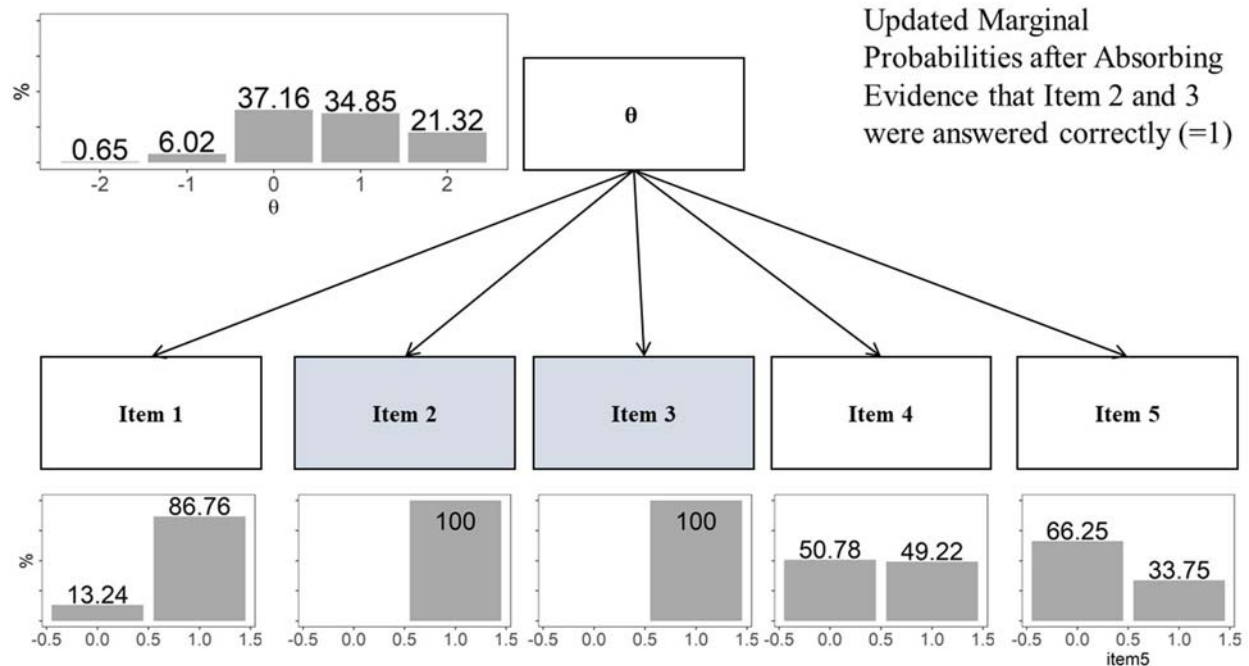


Fig. 4 Marginal probabilities after absorbing evidence that item 2 and 3 were correct. The shaded grade node indicates that evidence for the node is absorbed.

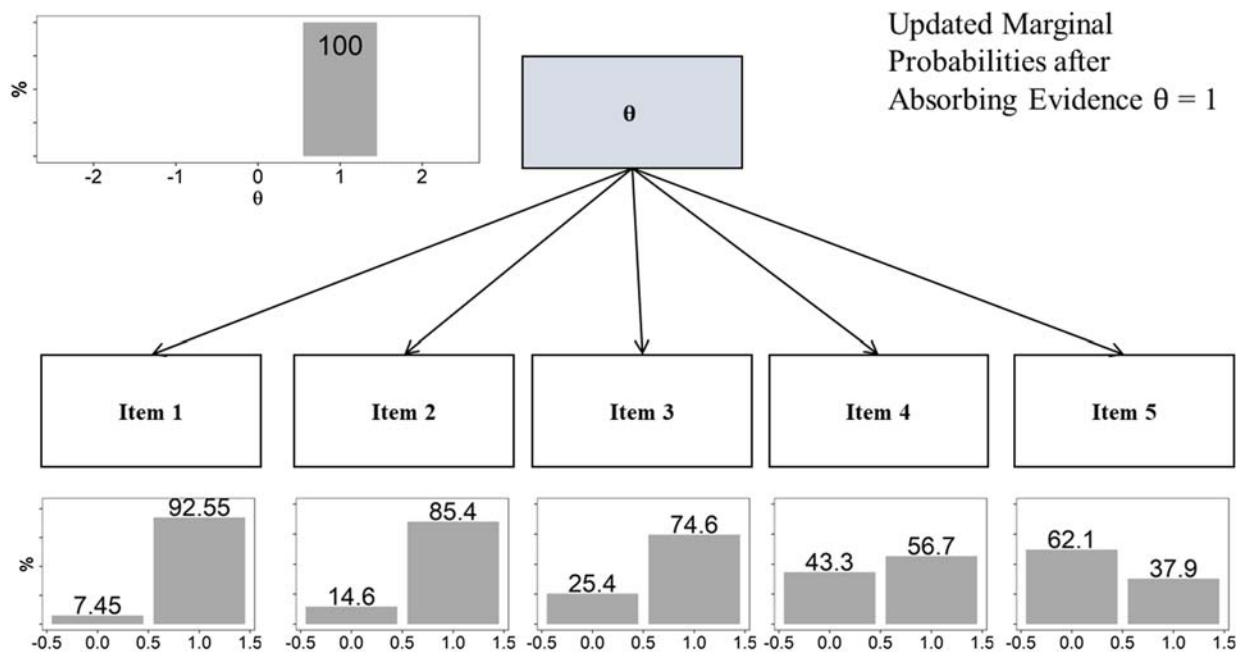


Fig. 5 Marginal probabilities after absorbing θ equal to 1. The shaded grade node indicates that evidence for the node is absorbed.

These examples demonstrate how we can perform both deductive and inductive reasoning on Bayesian networks. The example in Fig. 4 is an example of inductive reasoning (Almond et al., 2015). From the observed values of the items, the Bayesian network induces a posterior distribution for the unknown proficiency. Likewise, the example in Fig. 5 is an example of deductive reasoning. From the observed value of the proficiency variable, we can deduce the likely consequences (i.e., probability of answering an item correctly). Furthermore, another possible application that is readily available from the Bayesian network in Fig. 5 is that we can calculate the expected weight of evidence (or information) for a particular unobserved item. In the current example, using a result from IRT theory that states that under the Rasch model an item for which a student has 50–50 chance of getting the item right

provides the most information, we can see how the Bayesian network approach provides a simple approach to adaptive testing or item selection.

Additional topics in Bayesian networks

Evidence-centered design

The Evidence-Centered Design (ECD) project created at the Educational Testing Service (ETS) is an approach to constructing educational assessments in terms of evidentiary arguments (Mislevy et al., 2003). The ECD process involves a series of procedures that focus on questions such as: What can we observe about a test takers performance that will provide evidence that the test taker has the knowledge or ability we wish to make inferences about? How can we structure situations to be able to make these inferences? The ECD process results in a formal design for educational assessment that is known as the conceptual assessment framework (CAF). The CAF uses Bayesian networks to represent what we measure (student proficiency model), how we measure (evidence model), where we measure (task model), how much we measure (assembly model), and customization (presentation and delivery models). As with the examples here, ECD aims to use Bayesian networks to probabilistically reason about accumulating evidence across task performance to characterize the knowledge, skills, and/or abilities of students (Mislevy et al., 2003). For information, the reader is referred to Mislevy et al. (2003).

From Bayesian networks to causal networks

The idea of observing evidence on a Bayesian network is also important to Judea Pearl's work on causal inference (Pearl, 2009). This idea has paved the way for the notion of an intervention or do-calculus in Pearl's work on causal inference. For instance, Pearl, using the notion of intervening and do-calculus (i.e., set a variable to a specific observed value like we did in the example above), provides a way to probabilistically reason about causal inferences based on the structure of a graph (or causal model; Pearl, 1995). In fact, a recent book, embodying utility of interventions on Bayesian networks is entitled "Causal Inference: What If" (Hernán and Robins, 2020) where notions of interventions (and planned interventions) are prominently utilized to make causal inferences—recall Almond's quote above about "what-if" games on Bayesian networks. We simply provide a note here to make a connection between Bayesian networks and causal networks. Further information can be found in Pearl (1995) where a discussion is provided about how we can move from Bayesian networks to causal networks as well as in Pearl (2009).

Structure learning

Structure learning is another important idea worth mentioning that has emerged from the Bayesian network literature. In all of our examples herein, the structure of the Bayesian network (i.e., the direction of the edges between the nodes in the graph) was assumed to be known. Structure learning algorithms alternatively attempt to learn the underlying (causal) structure of a Bayesian network or particular features of a causal graphical model (see, for example, Heinze-Deml et al., 2018). Common causal structure learning algorithms come in the form of: (1) constraint-based algorithms, (2) score-based algorithms, and (3) hybrid-based algorithms (i.e., constraint- and score-based). Further, causal structure learning algorithms have been developed for structural equation models with non-Gaussian noise, for exploring invariance properties, for time series data, for mixed data (i.e., continuous and discrete), and Bayesian methods (see Heinze-Deml et al., 2018 for more details). Causal structure learning algorithms have not found widespread use in education. Notable examples include Fancsali (2014), Quintana (2020), and Rau and Scheines (2012).

Parameter learning

In the example above, we build on a common presentation of Bayesian networks in educational assessment that utilizes message-passing, belief propagation, and probability updating. The example is useful because it demonstrates how we can update our beliefs and probabilistically reason about individual students' proficiency based on information in their responses to items. However, it is important to note that this example assumes that the proficiency and conditional probability distributions for the observed variables are known. In other scenarios where we use Bayesian networks, we may be interested in estimating these distributions or parameters. Note, that the practice of "parameter estimation" in the statistics literature is commonly referred to as "parameter learning" in the literature on Bayesian networks and graphical models (e.g., Koller and Friedman, 2009). The term learning stems from the use of Bayesian networks in the machine learning and artificial intelligence communities.

Parameter learning in Bayesian networks is distinct from structure learning. Parameter learning generally assumes that the structure of the Bayesian network is already known or already learned from the data. Once the structure of the Bayesian network is known, estimating and updating the parameters is often simplified by decomposing the global distribution into a set of local distributions. From these local distributions, two approaches to parameter estimation are common: maximum likelihood estimation (MLE) and Bayesian estimation. For example, Koller and Friedman demonstrate how, using the properties of Bayesian networks, the likelihood function of MLE decomposes as a product of independent terms, one for each of conditional probability distribution in the network. They refer to this as the *global decomposition* of the likelihood function. We do not cover any further details on the maximum likelihood and Bayesian estimation of Bayesian networks in this chapter. However, detailed discussion on both

approaches to parameter estimation or learning, for both binary and continuous nodes, is provided in [Almond et al. \(2015\)](#) and [Koller and Friedman \(2009\)](#).

Summary

In this chapter, we provided a basic overview of the notion and properties of Bayesian networks. We discussed belief propagation, message-passing, and the recently developed R package, BayesNetBP for performing belief propagation on Bayesian networks. We also provided a motivating example using BayesNetBP for probabilistic reasoning in the form of a familiar multinomial Bayesian network known as the discrete IRT model ([Almond et al., 2015](#)). We provided two examples that demonstrate how we can observe evidence to perform inductive and deductive reasoning on a Bayesian network (i.e., either performance on an item or a student's proficiency level).

There are important topics not discussed here including belief propagation on conditional Gaussian Bayesian networks ([Lauritzen, 1996](#)), which is also available in the BayesNetBP package. Conditional Gaussian Bayesian networks facilitate message-passing techniques when there are both continuous and discrete nodes in a Bayesian Network, albeit without directed edges from continuous nodes to discrete nodes. Directed edges from continuous nodes to discrete nodes do not fall into the computationally convenient class of conditional Gaussian Bayesian networks ([Lauritzen, 1992, 1996](#))—that is, all integrals involved in a Bayesian network can be solved in closed form ([Almond et al., 2015](#)). We also do not discuss alternative Bayesian network formulations such as factor graphs or influence diagrams (see [Sucar, 2015](#)) nor extensions to Bayesian networks such as dynamic Bayesian networks (see [Levy, 2019](#)).

In summary, Bayesian networks are a type of graphical model that represents a set of variables and their conditional dependencies in a DAG. Bayesian networks provide a useful tool for visualizing the knowledge about some domain and embedding a probabilistic model for this domain. As the examples in this chapter demonstrate, Bayesian networks are useful for taking knowledge about an event or events that have happened and predicting the likelihood of some variable of interest. That is, Bayesian networks facilitate probabilistic reasoning by allowing us to make inquiries about the state of some variable given some observed evidence. Bayesian networks have been extended to develop educational assessments, make causal inferences, perform parameter and structure learning, and design efficient computational and estimation techniques.

Appendix A

```
#####
# Discrete IRT Data #
#####
library(tidyverse)
library(purrr)
set.seed(13)
n <- 10000
# t is theta and 1,2,3,4,5 are labels for -2,-1,0,1,2
t <- c(rep(1, n*.1), rep(2, n*.2),
      rep(3, n*.4), rep(4, n*.2), rep(5, n*.1))
i1 <- ifelse(t==1, rbernoulli(n, p = 0.3775),
            ifelse(t==2, rbernoulli(n, p = 0.6225),
            ifelse(t==3, rbernoulli(n, p = 0.8176),
            ifelse(t==4, rbernoulli(n, p = 0.9241),
            rbernoulli(n, p = 0.9707))))))
i2 <- ifelse(t==1, rbernoulli(n, p = 0.2227),
            ifelse(t==2, rbernoulli(n, p = 0.4378),
            ifelse(t==3, rbernoulli(n, p = 0.6792),
            ifelse(t==4, rbernoulli(n, p = 0.8520),
            rbernoulli(n, p = 0.9399))))))
i3 <- ifelse(t==1, rbernoulli(n, p = 0.1192),
            ifelse(t==2, rbernoulli(n, p = 0.2689),
            ifelse(t==3, rbernoulli(n, p = 0.5000),
            ifelse(t==4, rbernoulli(n, p = 0.7311),
            rbernoulli(n, p = 0.8088))))))
i4 <- ifelse(t==1, rbernoulli(n, p = 0.0601),
            ifelse(t==2, rbernoulli(n, p = 0.1480),
            ifelse(t==3, rbernoulli(n, p = 0.3208),
```

```
ifelse (t==4, rbernoulli (n, p = 0.5622),
rbernoulli (n, p = 0.7773))))
i5 <- ifelse (t==1, rbernoulli (n, p = 0.0293),
ifelse (t==2, rbernoulli (n, p = 0.0759),
ifelse (t==3, rbernoulli (n, p = 0.1824),
ifelse (t==4, rbernoulli (n, p = 0.3775),
rbernoulli (n, p = 0.6225))))))
dat <- as . data . frame (cbind (t, i1, i2, i3, i4, i5))
cols <- sapply (dat, is . logical)
dat[, cols] <- lapply (dat[, cols], as . numeric)
#####
# BayesNetBP Example Code #
#####
# you may need to install the graph package from bioconductor
library (bnlearn)
library (igraph)
library (BayesNetBP)
library (qtl)
library (qtlnet)
# defining graph with nodes theta and items 1-5
ug = empty.graph(c("t", "i1", "i2", "i3", "i4", "i5"))
arcs(ug) = matrix(
c("t", "i1",
"t", "i2",
"t", "i3",
"t", "i4",
"t", "i5"),
ncol = 2, byrow = TRUE,
dimnames = list(c(), c("from", "to")))
plot(ug)
student_gnel <- bn_to_graphNEL(ug) # converting to graphNEL
plot(student_gnel)
# defining node class TRUE = discrete and FALSE = continuous
node.class <- rep(c(TRUE), c(6))
names(node.class) <- c("t", "i1", "i2", "i3", "i4", "i5")
# initialize elimination tree and propagate evidence
tree.init <- Initializer(dag = student_gnel,
data = dat, node.class = node.class,
propagate = TRUE)
# view conditional probability tables cpt
FactorQuery(tree.init, vars = c("i1"), mode = "conditional")
FactorQuery(tree.init, vars = c("i2"), mode = "conditional")
FactorQuery(tree.init, vars = c("i3"), mode = "conditional")
FactorQuery(tree.init, vars = c("i4"), mode = "conditional")
FactorQuery(tree.init, vars = c("i5"), mode = "conditional")
# absorb evidence (probabilistic inquiries)
# item2 and item 3 correct
tree.i2i3 <- AbsorbEvidence(tree.init, vars = c("i2", "i3"),
values = list("1", "1"))
# theta = -1 (note 4 is a nominal label for -1)
tree.t1 <- AbsorbEvidence(tree.init, vars = c("t"),
values = list("4"))
# initial probabilities
marg.init <- Marginals(tree.init, c("t", "i1", "i2", "i3", "i4", "i5"))
# probabilities after absorbing item 2 and item 3 correct
marg.i2i3 <- Marginals(tree.i2i3, c("t", "i1", "i4", "i5"))
# probabilities after absorbing theta == 1
marg.t1 <- Marginals(tree.t1, c("i1", "i2", "i3", "i4", "i5"))
```


References

- Almond, R.G., Mislevy, R.J., Steinberg, L.S., Yan, D., Williamson, D.M., 2015. *Bayesian Networks in Educational Assessment*. Springer.
- Cowell, R.G., 2005. Local propagation in conditional Gaussian Bayesian networks. *J. Mach. Learn. Res.* 6, 1517–1550.
- Cramer, A.O., Waldorp, L.J., Van Der Maas, H.L., Borsboom, D., 2010. Comorbidity: a network perspective. *Behav. Brain Sci.* 33 (2–3), 137–150.
- Darroch, J.N., Lauritzen, S.L., Speed, T.P., et al., 1980. Markov fields and log-linear interaction models for contingency tables. *Ann. Stat.* 8 (3), 522–539.
- Duncan, O.D., 1966. Path analysis: sociological examples. *Am. J. Sociol.* 72 (1), 1–16.
- Epskamp, S., Rhemtulla, M., Borsboom, D., 2017. Generalized network psychometrics: combining network and latent variable models. *Psychometrika* 82 (4), 904–927.
- Fancsal, S., 2014. Causal discovery with models: behavior, affect, and learning in cognitive tutor algebra. In: *Proceedings of the 7th International Conference on Educational Data Mining*.
- Gibbs, W., 1902. *Elementary Principles of Statistical Mechanics*. Yale University Press.
- Glymour, C., 2004. James Woodward, making things happen: a theory of causal explanation. *Br. J. Philos. Sci.* 55 (4), 779–790.
- Haberman, S.J., 2005. Latent-class item response models. *ETS Res. Rep. Ser.* 2005 (2), i–7.
- Heinze-Deml, C., Maathuis, M.H., Meinshausen, N., 2018. Causal structure learning. *Annu. Rev.* 5, 371–391.
- Heise, D.R., 1975. *Causal Analysis*. John Wiley & Sons.
- Hernán, M.A., Robins, J.M., 2020. *Causal Inference: What If*. Chapman & Hall/CRC, Boca Raton.
- Hugin Expert A/S, 2017. Hugin Software. URL: <http://www.hugin.com/>.
- Hünemann, P., Bareinboim, E., 2019. Causal Inference and Data-Fusion in Econometrics arXiv preprint arXiv:1912.09104.
- Kivieri, H., Speed, T.P., 1982. Structural analysis of multivariate data: a review. *Sociol. Methodol.* 13, 209–289.
- Koller, D., Friedman, N., 2009. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.
- Lauritzen, S.L., Jensen, F., 2001. Stable local computation with conditional Gaussian distributions. *Stat. Comput.* 11 (2), 191–203.
- Lauritzen, S.L., Spiegelhalter, D.J., 1988. Local computations with probabilities on graphical structures and their application to expert systems. *J. Roy. Stat. Soc. B* 50 (2), 157–194.
- Lauritzen, S.L., Speed, T.P., Vijayan, K., 1984. Decomposable graphs and hypergraphs. *J. Aust. Math. Soc.* 36 (1), 12–29.
- Lauritzen, S.L., 1992. Propagation of probabilities, means, and variances in mixed graphical association models. *JASA* 87 (420).
- Lauritzen, S.L., 1996. *Graphical Models*. Oxford University Press.
- Levy, R., 2019. Dynamic Bayesian network modeling of game-based diagnostic assessments. *Multivariate Behav. Res.* 54 (6), 771–794.
- Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., van Bork, R.V., Waldorp, L.J., et al., 2018. An introduction to network psychometrics: relating Ising network models to item response theory models. *Multivariate Behav. Res.* 53 (1), 15–35.
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A brief introduction to evidence-centered design. *ETS Res. Rep. Ser.* 2003 (1), i–29.
- Nagarajan, R., Scutari, M., Lèbre, S., 2013. *Bayesian Networks in R*, vol. 122. Springer, pp. 125–127.
- Pearl, J., Verma, T., 1987. *The Logic of Representing Dependencies by Directed Graphs*. University of California (Los Angeles) Computer Science Department.
- Pearl, J., Glymour, M., Jewell, N.P., 2016. *Causal Inference in Statistics: A Primer*. John Wiley & Sons.
- Pearl, J., 1986. Fusion, propagation, and structuring in belief networks. *Artif. Intell.* 29 (3), 241–288.
- Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Pearl, J., 1995. From Bayesian networks to causal networks. In: *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*. Plenum Press, pp. 157–181.
- Pearl, J., 2009. *Causality*. Cambridge University Press.
- Quintana, R., 2020. The Structure of Academic Achievement: Searching for Proximal Mechanisms Using Causal Discovery Algorithms. *Sociological Methods & Research*, 0049124120926208.
- Rau, M.A., Scheines, R., 2012. Searching for Variables and Models to Investigate Mediators of Learning From Multiple Representations. *International Educational Data Mining Society*.
- Rijmen, F., 2008. Bayesian networks with a logistic regression model for the conditional probabilities. *Int. J. Approx. Reason.* 48 (2), 659–666.
- Rijmen, F., 2009. An efficient EM algorithm for multidimensional IRT models: full information maximum likelihood estimation in limited time. *ETS Res. Rep. Ser.*
- Rijmen, F., Vansteelandt, K., De Boeck, P., et al., 2008. Latent class models for diary method data: parameter estimation by local computations. *Psychometrika* 73 (2).
- Savi, A.O., Marsman, M., van der Maas, H.L., Maris, G.K., 2019. The wiring of intelligence. *Perspect. Psychol. Sci.* 14 (6), 1034–1061.
- Scutari, M., Denis, J.B., 2014. *Bayesian Networks: With Examples in R*. CRC Press.
- Speed, T.P., 1978. Relations between models for spatial data, contingency tables and Markov fields on graphs. *Adv. Appl. Probab.* 10, 111–122.
- Sugar, L.E., 2015. *Probabilistic Graphical Models: Principles and Applications*. Springer.
- Yu, H., Moharil, J., Blair, R.H., 2020. BayesNetBP: an R package for probabilistic reasoning in Bayesian networks. *J. Stat. Software* 94 (3), 1–31.

Bayesian psychometrics

Xiang Liu, Matthew S. Johnson, and Zhuangzhuang Han, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	130
Bayesian data analysis	130
Model specification	131
Posterior inference	132
A real data example	134
Posterior predictive model checking	134
Observed sum scores	134
Polychoric correlation coefficient	134
Odds ratio	136
Conclusion	136
References	137

Introduction

Bayesian statistical methods have become popular in the field of psychometrics. Thanks to the development of general purpose Bayesian inference software (e.g., *Stan*; Gelman et al., 2015), researchers are now able to quickly develop and fit sophisticated and complex psychometric models that once required significant computational expertise to estimate. In addition, Bayesian methods also provide a class of approaches alternative to classical or frequentist statistics (e.g., maximum likelihood, confidence intervals, p-values). In certain cases, Bayesian approaches can alleviate some difficulties classical approaches may have. There is abundant literature on applying Bayesian methods to psychometric modeling. Levy (2009) gives a nice overview and survey of the applications of Bayesian estimation for psychometric models. Some specific examples include Bayesian estimation of a broad class of cognitive diagnostic models (Culpepper, 2015; Liu and Johnson, 2019), the extended marginal Rasch model (Maris et al., 2015), a multilevel item response theory (IRT) model (Fox and Glas, 2001), and an unfolding response model (Johnson and Junker, 2003). Other examples include estimating uncertain elements of the Q-matrix in a CDM model (DeCarlo, 2012), nonparametric ordered latent class modeling (Liu, 2019), robust IRT outlier detection (Öztürk and Karabatsos, 2017), and evaluating model fit (Sinharay, 2005, 2015; Sinharay and Almond, 2007). These mentioned examples of Bayesian methods in psychometrics are far from exhaustive. The main purpose of this entry is to give readers a brief introduction to some of the most essential concepts and steps involved in a Bayesian analysis of psychometric models.

Bayesian data analysis

Let X denote the observed random variables (i.e., the data), and θ denote the vector of unobserved random variables (i.e., the parameters). In Bayesian statistics, the parameters are treated as random and are given prior distributions. This is in contrast to the classical likelihood methods where parameters are treated as fixed but unknown. We are typically interested in using the data to estimate and make inferences about the vector of unknown parameters. A Bayesian analysis involves three essential steps:

1. Model specification. In this step, we set up a full probability model that describes the joint probability distribution for all observed and unobserved random variables, i.e., $p(X, \theta)$. According to the definition of conditional probability, this joint probability distribution can be factorized as $p(X, \theta) = p(X|\theta)p(\theta)$. The marginal density, $p(\theta)$, is often referred to as the *prior distribution* which expresses any prior belief we may have on the parameters with uncertainty, and the conditional density, $p(X|\theta)$, is called the *sampling distribution* (or the generating distribution) which describes the data generation process. Setting up these models generally requires some knowledge of the underlying problems with, sometimes, considerations for mathematical convenience.
2. Posterior inference. In this step, we calculate and interpret the *posterior distribution* from which we make inferences about the parameters. The posterior distribution is the conditional distribution of the unobserved random variables (i.e., parameters) which we are typically interested in, given the observed data, i.e., $p(\theta|X) = p(\theta, X)/p(X) = p(X|\theta)p(\theta)/p(X)$. This is known as Bayes' rule. As a function of parameters θ , $p(X|\theta)$ is also referred to as the likelihood function through which the data X influences the posterior distribution of the parameters. To obtain the density of the marginal distribution of the data, we need to evaluate $p(X) = \int p(X|\theta)p(\theta)d\theta$. Except for certain extremely simple models, this (potentially high dimensional) integral cannot be evaluated analytically. Consequently, the posterior is generally not available in closed forms. Instead, the posterior distribution $p(\theta|X)$ is usually approximated using sampling methods such as the Markov chain Monte Carlo (MCMC; Gelman et al., 2013; Neal, 1998).

3. Model fit evaluation. Once a model has been set up and the posterior distribution has been computed, it is important to assess the fit of the model. Inferences drawn from a model that does not fit well could be very misleading. In a Bayesian framework, the assessment of model fit is typically performed through the posterior predictive model checking (PPMC; Gelman et al., 2013; Rubin, 1984). If the model fits reasonably well, the observed data or the derived statistics of our practical interest should look plausible under the posterior predictive distribution. The identified aspects of model misfit could help improve or expand the model. Sometimes, identifying misfit itself could be of practical interest, for example, to identify aberrant response patterns (e.g., Sinharay, 2015).

These steps are true to all Bayesian modeling applications. Bayesian psychometrics is an example of such applications to psychometric modeling. In the following sections, we demonstrate these essential steps of the Bayesian data analysis with a little more details by using a simple 1-parameter normal ogive IRT model as an example.

Model specification

Let $X_{ij} = x_{ij}$ be a dichotomous random variable of the item response from the i^{th} student to the j^{th} item, $i = 1, 2, \dots, N$ and $j = 1, 2, \dots, J$. And $x_{ij} = 1$ if the response is correct, $x_{ij} = 0$ otherwise. Under the 1-parameter normal ogive (1PNO) IRT model, the probability of a correct response is given by

$$P(X_{ij} = 1 | \theta_i, a, b_j) = \Phi(a(\theta_i - b_j)), \quad (1)$$

where $\Phi(\cdot)$ denotes the cumulative probability density function (CDF) of the standard normal distribution, θ_i is the person parameter, b_j is the item difficulty parameter, and a is the common discrimination parameter shared among J items. The person parameter is assumed to follow a standard normal distribution, i.e., $\theta_i \sim N(0, 1)$. Because a is a constant across items, the item characteristic curves (ICC) never cross each other and only differ by locations (see Fig. 1).

The 1PNO as in Eq. (1) can be expressed as a Bayesian hierarchical model. For computational efficiency, we reparameterize the model so that the model is in its slope-intercept form,

$$P(X_{ij} = 1 | \theta_i, \psi_j) = \Phi(\theta_i + \psi_j), \quad (2)$$

and $\theta_i \sim N(0, a^2)$, where a^2 is the variance of the normal distribution.

A prior distribution describes any prior belief we may have (or the lack of it) on the parameter with uncertainty. For instance, in the current example, there may be reasonable prior information on the item parameter ψ_j from similar items previously administered to a comparable population of students. In this case, we may choose a normal prior with mean μ_{ψ_j} informed by prior information and variance $\sigma_{\psi_j}^2$ expressing the level of uncertainty, i.e., $\psi_j \sim N(\mu_{\psi_j}, \sigma_{\psi_j}^2)$. A prior with larger variance (or a flatter prior, see Fig. 2) is less informative and appropriate when there is greater uncertainty about a parameter, thus having less impact on the posterior inference of the parameter. However, prior information may not always be available. In this case, instead of choosing a fixed prior with a large variance, an alternative approach is to use a hierarchical prior. For each item parameter ψ_j , we assign a same Gaussian prior, $\psi_j \sim N(\mu_{\psi}, \sigma_{\psi}^2)$. Hyper-priors are assigned to the parameters of this prior distribution, i.e., $\mu_{\psi} \sim N(0, 10^2)$ and

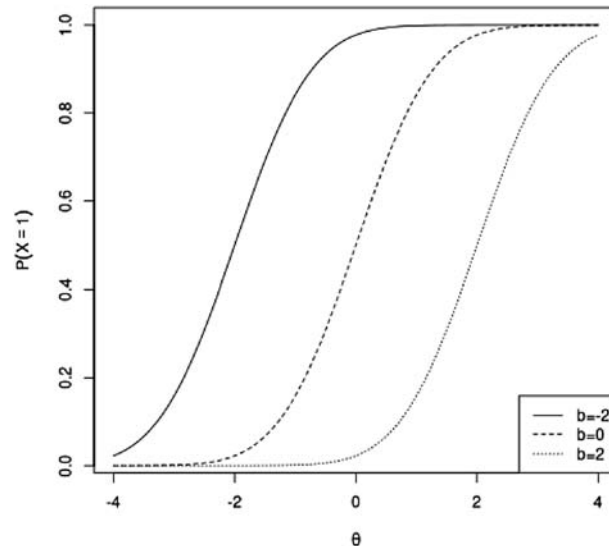


Fig. 1 ICC of the 1PNO IRT model.

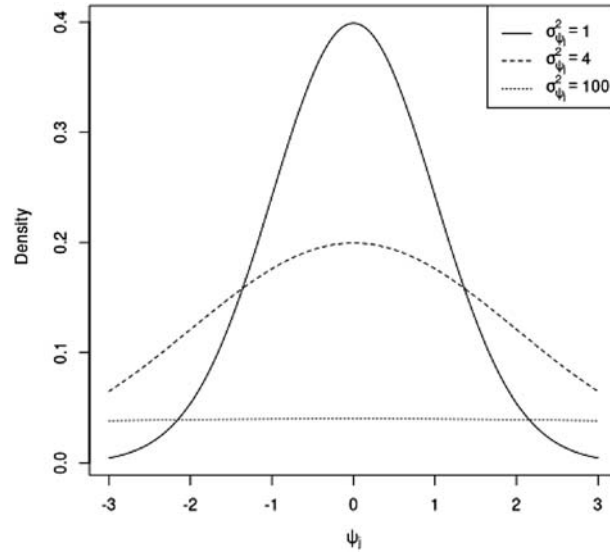


Fig. 2 Density of prior distribution of ψ_j .

$\sigma_\psi^2 \sim Ga^{-1}(10^{-4}, 10^{-4})$, where $Ga^{-1}(\cdot)$ denotes the inverse-gamma distribution (see Gelman et al., 2013, p. 43). The chosen inverse-gamma distribution is noninformative if most of the posterior density is away from 0 as the prior density is flat in the tail. The choice of this set of priors is also mathematically convenient. The normal prior and inverse-gamma prior are conditionally conjugate for the model. A prior is said to be conjugate for the likelihood function if the resulting posterior distribution is in the same probability distribution family. In this case, the conditional posteriors for the mean parameter and the variance parameter normal and inverse-gamma. We choose a similar noninformative inverse-gamma distribution for a^2 , i.e., $a^2 \sim Ga^{-1}(10^{-4}, 10^{-4})$. We can then write the 1PNO IRT model has a Bayesian hierarchical model,

$$\begin{aligned}
 x_{ij}|\theta_i, \psi_j &\sim \text{Bernoulli}\left(\Phi\left(\theta_i + \psi_j\right)\right) \\
 \psi_j|\mu_\psi, \sigma_\psi^2 &\sim N\left(\mu_\psi, \sigma_\psi^2\right) \\
 \mu_\psi &\sim N(0, 10^2) \\
 \sigma_\psi^2 &\sim Ga^{-1}\left(10^{-4}, 10^{-4}\right) \\
 \theta_i|a^2 &\sim N(0, a^2) \\
 a^2 &\sim Ga^{-1}\left(10^{-4}, 10^{-4}\right).
 \end{aligned} \tag{3}$$

Posterior inference

For any practical applications, we are interested in inferring model parameters given the data, in other words, computing the posterior distribution. The *joint* posterior distribution of the parameters, $p(\theta, \psi, \mu_\psi, \sigma_\psi^2, a^2|\mathbf{x})$, cannot be analytically computed. As a result, we use MCMC to sample from this unknown distribution. The simplest MCMC algorithm is perhaps the Gibbs sampling. A Gibbs sampler (Geman and Geman, 1984) works by iteratively sampling from the full conditional distributions of the parameters. Generally, this is feasible if the full conditional distributions can be computed in closed form and sampled from easily. However, this may not be true for some models. In these cases, other MCMC algorithms are required, e.g., the Metropolis-Hastings algorithm (Hastings, 1970). The literature on MCMC algorithms is vast, and the development of new MCMC methods is still an extremely active research area. Readers may refer to Gelman et al. (2013) and Neal (1998) for some detailed introduction and review on this topic.

For our example of the 1PNO model, a direct Gibbs sampling is possible. However, we would need to augment latent variables. For every item response x_{ij} , there is a latent variable y_{ij} associated. The relationship between them is deterministic such that

$$x_{ij} = \begin{cases} 1 & \text{if } y_{ij} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $y_{ij} \sim N(\theta_i + \psi_j, 1)$. It is straightforward to verify that, under this specification, the probability of a correct response remains $\int_{-\infty}^{\theta_i + \psi_j} \phi(y) dy = \Phi(\theta_i + \psi_j)$.

To derive the densities full conditional distributions, we use the definition of conditional probabilities and the chain rule. For example,

$$p(y_{ij}|x_{ij}, \theta_i, \psi_j) = \frac{p(x_{ij}|y_{ij})p(y_{ij}|\theta_i, \psi_j)}{\sum_{y_{ij}} p(x_{ij}|y_{ij})p(y_{ij}|\theta_i, \psi_j)}. \quad (4)$$

Notice that the denominator of Eq. (4) is a constant (with respect to y_{ij}) that normalizes the full conditional density. Therefore, we can conveniently write

$$\begin{aligned} p(y_{ij}|x_{ij}, \theta_i, \psi_j) &= Cp(x_{ij}|y_{ij})p(y_{ij}|\theta_i, \psi_j) \\ &\propto p(x_{ij}|y_{ij})p(y_{ij}|\theta_i, \psi_j). \end{aligned}$$

Given that $p(x_{ij}|y_{ij})$ is a piecewise point density with support on either $x_{ij} = 1$ or $x_{ij} = 0$ depending on y_{ij} , the full conditional distribution of y_{ij} is a truncated normal distribution, i.e.,

$$y_{ij}|x_{ij}, \theta_i, \psi_j \sim \begin{cases} N(\theta_i + \psi_j, 1)T(0, \infty) & \text{if } x_{ij} = 1, \\ N(\theta_i + \psi_j, 1)T(\infty, 0) & \text{if } x_{ij} = 0. \end{cases}$$

$T(0, \infty)$ denotes that the distribution is truncated with a lower bound of 0; while $T(\infty, 0)$ denotes a truncation with an upper bound of 0.

The full conditional distributions of other parameters can be derived similarly. Since the priors are chosen so that they are *conditionally* conjugate, the full conditional distributions of the parameters are within the same distribution families as their priors. The standard conjugacy results apply (for a reference, see Gelman et al., 2013). Here, we state the results without detailed derivations.

Now we can summarize the Gibbs sampler. After assigning initial values to parameters, at the M^{th} iteration,

1. for each i and j , update y_{ij} by drawing from $N(\theta_i + \psi_j, 1)T(0, \infty)$ if $x_{ij} = 1$; else draw from $N(\theta_i + \psi_j, 1)T(\infty, 0)$;
2. for each i , update θ_i by drawing from $N\left(\frac{1}{1/a^2 + J} \left(\sum_j (y_{ij} - \psi_j) \right), (1/a^2 + J)^{-1}\right)$;
3. for each j , update ψ_j by drawing from $N\left(\frac{1}{1/\sigma_\psi^2 + N} \left(\mu_\psi / \sigma_\psi^2 + \sum_i (y_{ij} - \theta_i) \right), \frac{1}{1/\sigma_\psi^2 + N}\right)$;
4. update a^2 by drawing from $Ga^{-1}\left(10^{-4} + \frac{N}{2}, 10^{-4} + \frac{\sum_i \theta_i^2}{2}\right)$;
5. update μ_ψ by drawing from $N\left(\frac{1}{1/10^2 + J/\sigma_\psi^2} \left(\frac{\sum_j \psi_j}{\sigma_\psi^2} \right), (1/10^2 + J/\sigma_\psi^2)^{-1}\right)$;
6. update σ_ψ^2 by drawing from $Ga^{-1}\left(10^{-4} + J/2, 10^{-4} + \frac{\sum_i (\psi_i - \mu_\psi)^2}{2}\right)$.

Keep updating long enough, the draws are eventually random samples from the joint posterior distribution. Even though the theory of Markov chains guarantees the eventual convergence to the desired target distribution which is the joint posterior (see Neal, 1998, for a review), in practice, we can only draw finite number of samples. It is important to assess the convergence so that we are reasonably confident using the draws to form valid posterior inference. There are many methods for monitoring the convergence of MCMC (Cowles and Carlin, 1996). For example, the Gelman-Rubin statistic (Gelman and Rubin, 1992) is one of the most popular methods and implemented in many Bayesian computation software.

Once the samples are drawn, to make inferences about parameters, we need to summarize the posterior distribution. Typically, we use mean, mode, and median to describe the central tendencies of the posterior distribution and serve as a point estimate. Among them, posterior mean is probably the most common. In addition to the point estimate, the variance or standard deviation can be used to describe the dispersion of the posterior distribution. We may also be interested in forming posterior intervals which is referred to as the credible intervals. One way to construct a $100(1 - \alpha)\%$ credible interval is to form the equal-tailed interval, i.e., $I_\alpha = [\theta_{\alpha/2}, \theta_{1-\alpha/2}]$, where θ_z is the z -quantile of the posterior. However, this interval could be unnecessarily long and not representative of the posterior if the posterior distribution is asymmetric or even multi-modal. A better approach is to construct the Highest Posterior Density (HPD) intervals/sets. A $100(1 - \alpha)\%$ HPD set is defined as $I_\alpha = \{\theta : p(\theta|\mathbf{x}) \geq k\}$ where

$$k = \max \left\{ k : \int_{\theta: p(\theta|\mathbf{x}) \geq k} p(\theta|\mathbf{x}) d\theta = 1 - \alpha \right\}.$$

If the posterior is approximately uni-modal, then the HPD set is guaranteed to be the HPD interval.

Table 1 Posterior summary.

	<i>Mean</i>	<i>SD</i>	<i>Lower</i>	<i>Upper</i>
b_1	-3.626	0.337	-4.299	-3.049
b_2	-1.405	0.157	-1.716	-1.128
b_3	-0.349	0.107	-0.569	-0.146
b_4	-1.823	0.188	-2.202	-1.500
b_5	-2.857	0.270	-3.405	-2.368
a	0.432	0.040	0.349	0.504

A real data example

We use the LSAT dataset provided by the *ltm* R package (Rizopoulos, 2015) for demonstration. The dataset was first described in Bock and Lieberman (1970). It contains dichotomous item responses from 1000 students to 5 items that are designed to measure a single latent trait. We fit the Bayesian 1PNO model with the Gibbs sampler described in the previous section. The posterior samples are then transformed back to the more common parameterization as in Eq. (1). The posterior summary are in Table 1. The first two columns are posterior means and posterior standard deviations. The last two columns are the lower limit and the upper limit of the 95% HPD intervals.

Posterior predictive model checking

A statistical model is essentially a set of assumptions. Before any inferences are drawn from the model, it is important to assess the fit of the model or, in other words, to check the degree to which the model fails to explain the data that are of our practical interest. This is particularly important to psychometrics where we are interested in inferring or measuring some theoretical attributes (e.g., intelligence) from observable quantities (e.g., responses to test items). A test item may not function as intended, and the subject may not respond to items as assumed by the model. Before making any conclusions about the attribute of interest, it is critical to evaluate the assumptions. Model checking procedures can help identify unreasonable assumptions which may be potentially mitigated in some way later.

The basic intuition of the PPMC method is that we can check the observed data against the replicated data predicted by the model using some statistics. The distribution of the replicated data is also referred to as the posterior predictive distribution, i.e.,

$$p(\mathbf{x}^{rep}|\mathbf{x}) = \int p(\mathbf{x}^{rep}|\boldsymbol{\zeta})p(\boldsymbol{\zeta}|\mathbf{x})d\boldsymbol{\zeta}, \quad (5)$$

where $\boldsymbol{\zeta}$ is a vector of all model parameters and $p(\boldsymbol{\zeta}|\mathbf{x})$ is the posterior distribution. Except for some simple models, finding the posterior predictive distribution requires sampling in general.

A statistic is a function of the data. Different statistics may describe different aspects of the data. Thus, it is important to choose appropriate statistics that reflect aspects of the model that are of interest to us. In this chapter, we follow the suggestions given by Sinharay (2005) and demonstrate the PPMC method using some most common and important statistics.

Observed sum scores

In many tests, ranking students is one of the most important objective. A model that does not predict the number of correct responses given by students well is very unlikely could rank students fairly. For the 5 dichotomous items in the example, there are 6 possible score categories. We observe a number of students within each score category. So the observed distribution of sum scores are $NC = (NC_0, NC_1, \dots, NC_5)$. Using the posterior samples produced by the Gibbs sampler, we can generate replicated datasets and compute the posterior predictive distribution of the sum scores. The result is in Fig. 3.

The observed sum scores all fall within the 80% HPD intervals of the posterior predictive distributions. For a given set of observed data and a chosen model, the choice of the level of HPD intervals has an effect on their lengths. The smaller percentages are associated with shorter intervals, therefore representing a more strict criteria for evaluating model fit. Compared to higher levels such as 90% or 95%, using a 80% HPD intervals leads to higher chances of detecting model misfits. In our case, the model does seem to be able to explain the distribution of number of correct responses well.

Polychoric correlation coefficient

An important assumption of the 1PNO IRT model is that discrimination parameters across all items are equal. To check this assumption we use the polychoric correlation coefficient. A polychoric correlation coefficient describes the correlation between

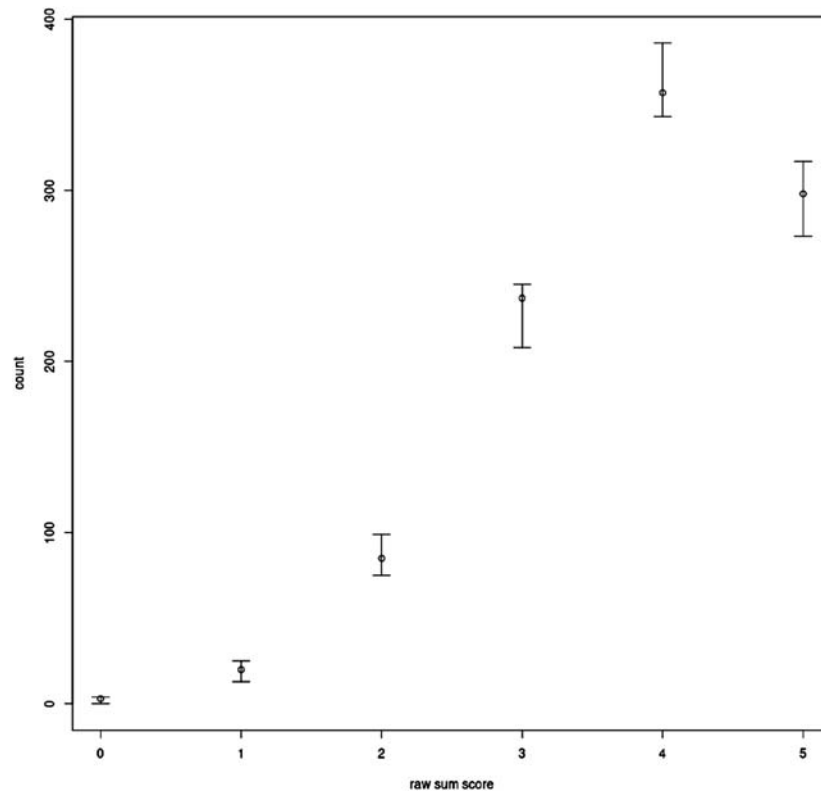


Fig. 3 The vertical bars represent the 80% HPD intervals of the predicted frequency distributions at each sum score, and the circles are the observed frequency of each sum score.

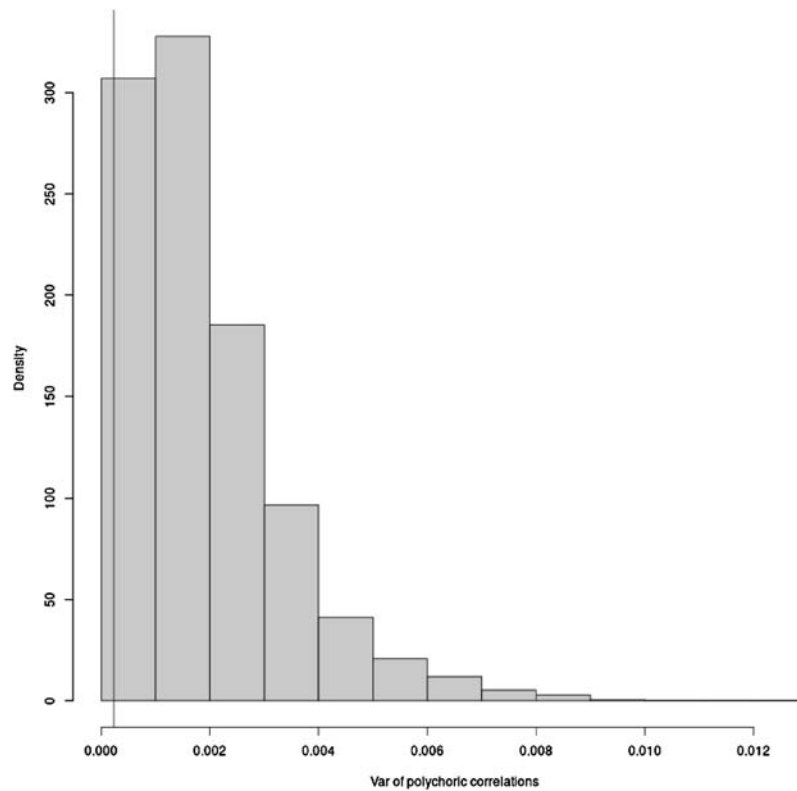


Fig. 4 The histogram describes the posterior predictive distribution of the variance of the polychoric correlations of the 5 items. The vertical line is the observed variance of the polychoric correlations. The observed is in a high posterior density region.

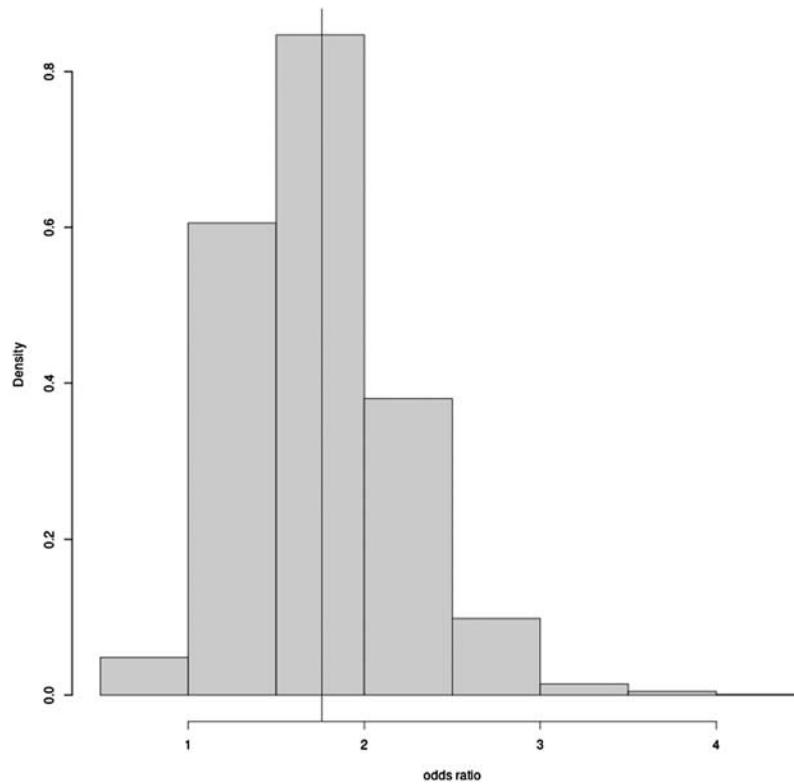


Fig. 5 The histogram describes the posterior predictive distribution of the odds ratio of an item pair. The vertical line marks the observed odds ratio of that item pair.

the normal variables derived from two observed ordinal variables - in our case, the total number of correct responses of a student and that student's item response to the j^{th} item. If the model is reasonable, then we would expect to see the polychoric correlations for all J items are similar or the variance of the polychoric correlations is close to 0. The results in Fig. 4 shows the variance of the observed polychoric correlations are very close to 0 and within a high density area of the posterior predictive distribution of the same statistic. Consequently, there is no evidence of this aspect of misfit of the model.

Odds ratio

The 1PNO IRT model assumes that item responses are independent of each other given latent abilities θ . This is also commonly referred to as the local independence (LI) assumption. If LI holds, the model should predict the odds ratio of any item pairs,

$$OR = \frac{n_{00}n_{11}}{n_{01}n_{10}}, \quad (6)$$

where $n_{kk'}$ denotes the number of students has a score k on the first item and a score k' on the second item. For 5 items, there are $\binom{5}{2} = 10$ pairs. For demonstration purpose, we plot the results of one item pair in Fig. 5. The observed OR is within a high density region of the posterior predictive distribution. The examined statistic does not show evidence of violation of the LI assumption for this pair of items.

Conclusion

Thanks to the development of general purpose Bayesian inference software such as the Stan, applied researchers now have access to a set of powerful tools to develop and estimate complex and sophisticated Bayesian models. While the software can help automate the Bayesian computation, it should be emphasized that all three steps outlined earlier in the entry are important and integral parts of the Bayesian modeling. Especially, the importance of evaluating model fit should not be overlooked. A model is a set of assumptions. Drawing inferences based on assumptions without checking them can be dangerous. Furthermore, the Bayesian modeling is really a continuous process where researchers iterate through the three steps. After the aspects of model-data misfits are identified in model fit evaluations, the model can be expanded and improved. In this entry, we offered a high level introduction to the basics of

Bayesian psychometrics. For a more detailed treatment of these introductory topics, readers may refer to Gelman et al. (2013) and Levy and Mislevy (2016).

References

- Bock, D.R., Lieberman, M., 1970. Fitting a response model for n dichotomously scored items. *Psychometrika* 35 (2), 179–197. <https://doi.org/10.1007/BF02291262>.
- Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. *J. Am. Stat. Assoc.* <https://doi.org/10.1080/01621459.1996.10476956>.
- Culpepper, S.A., 2015. Bayesian estimation of the DINA model with Gibbs sampling. *J. Educ. Behav. Stat.* 40 (5), 454–476. <https://doi.org/10.3102/1076998615595403>.
- DeCarlo, L.T., 2012. Recognizing uncertainty in the Q-matrix via a Bayesian extension of the DINA model. *Appl. Psychol. Meas.* 36 (6), 447–468. <https://doi.org/10.1177/0146621612449069>.
- Fox, J.P., Glas, C.A., 2001. Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika* 66 (2). <https://doi.org/10.1007/BF02294839>.
- Gelman, A., Lee, D., Guo, J., 2015. Stan: a probabilistic programming language for Bayesian inference and optimization. *J. Educ. Behav. Stat.* 40 (5), 530–543.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*. Chapman; Hall/CRC.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* <https://doi.org/10.1214/ss/1177011136>.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6 (6), 721–741. <https://doi.org/10.1109/TPAMI.1984.4767596>.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57 (1), 97–109. <http://www.jstor.org/stable/2334940>.
- Johnson, M.S., Junker, B.W., 2003. Using data augmentation and Markov chain Monte Carlo for the estimation of unfolding response models. *J. Educ. Behav. Stat.* 28 (3), 195–230. <https://doi.org/10.3102/10769986028003195>.
- Levy, R., 2009. The rise of Markov chain Monte Carlo estimation for psychometric modeling. *J. Probab. Stat.* 1–18. <https://doi.org/10.1155/2009/537139>.
- Levy, R., Mislevy, R.J., 2016. *Bayesian Psychometric Modeling*, first ed. Hall/CRC, Chapman.
- Liu, X., 2019. *Three Contributions to Latent Variable Modeling*. Doctoral Dissertation. Columbia University.
- Liu, X., Johnson, M.S., 2019. Estimating CDMs using MCMC. In: von Davier, M., Lee, Y.-S. (Eds.), *Handbook of Diagnostic Classification Models*. Springer, pp. 629–646.
- Maris, G., Bechger, T., San Martín, E., 2015. A Gibbs sampler for the (extended) marginal Rasch model. *Psychometrika* 80 (4), 859–879. <https://doi.org/10.1007/s11336-015-9479-4>.
- Neal, R.M., 1998. Probabilistic inference using Markov chain Monte Carlo methods. *Tech. Rep.* 1, 1–144. <https://doi.org/10.1021/np100920q>.
- Öztürk, N.K., Karabatsos, G., 2017. A Bayesian robust IRT outlier-detection model. *Appl. Psychol. Meas.* 41 (3), 195–208. <https://doi.org/10.1177/0146621616679394>.
- Rizopoulos, D., 2015. ltm: an R package for latent variable modeling and item response theory analyses. *J. Stat. Software* 17 (5), 1–25. <https://doi.org/10.18637/jss.v017.i05>.
- Rubin, D.B., 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Stat.* 12 (4), 1151–1172. <https://doi.org/10.1214/aos/1176346785>.
- Sinharay, S., 2015. Assessment of person fit for mixed-format tests. *J. Educ. Behav. Stat.* 40 (4), 343–365. <https://doi.org/10.3102/1076998615589128>.
- Sinharay, S., 2005. Assessing fit of unidimensional item response theory models using a Bayesian approach. *J. Educ. Meas.* 42 (4), 375–394. <https://doi.org/10.1111/j.1745-3984.2005.00021.x>.
- Sinharay, S., Almond, R.G., 2007. Assessing fit of cognitive diagnostic models a case study. *Educ. Psychol. Meas.* 67 (2), 239–257. <https://doi.org/10.1177/0013164406292025>.

Nonparametric methods for modeling item responses: Mokken Scale Analysis

Stefanie A. Wind, The University of Alabama, Tuscaloosa, AL, United States

© 2023 Elsevier Ltd. All rights reserved.

Nonparametric methods for modeling item responses: Mokken Scale Analysis	138
Nonparametric modeling frameworks for educational research	139
Mokken Scale Analysis framework	140
Mokken Scale Analysis models	140
Variations on MSA	143
Other nonparametric approaches	143
Summary and conclusions	144
References	144

Glossary

Invariance Evidence that a characteristic of interest holds over some set of conditions

Invariant item ordering Evidence that relative item difficulty ordering is consistent for all examinees, regardless of their latent variable location

Latent variable A continuous psychological variable that represents a construct of interest that cannot be directly measured. In item response modeling, items, persons, and other elements of an assessment procedure are hypothesized to hold locations on a latent variable

Nonparametric methods Descriptive and inferential statistical techniques that are not directly tied to a particular probability distribution

Modeling framework An approach to modeling quantitative data that provides researchers with a structure for analyzing the data and interpreting the results based on guiding principles or requirements

Mokken scale analysis A nonparametric item response modeling framework based on principles of invariant ordering

Monotonicity For dichotomous items, evidence that the probability for a correct or positive response is non-decreasing as examinee locations on the latent variable increase; For polytomous items, evidence that average ratings are non-decreasing as examinee locations on the latent variable increase

Nonparametric modeling framework Modeling framework for item response data in which the statistical techniques do not impose a specific mathematical structure on the shape of the relationship between the latent variable and the probability for a response (i.e., the item response function [IRF]) and instead rely on ordering requirements

Parametric modeling framework Modeling framework for item response data in which the statistical techniques impose a specific mathematical structure on the shape of the relationship between the latent variable and the probability for a response (i.e., the item response function [IRF])

Restscore Examinee sum scores minus scores from an item or a set of items of interest

Restscore group Groups of examinees with the same or similar restscores. Restscore groups are used to provide more statistical power for evaluating item properties such as monotonicity or invariant item ordering

Scalability Evidence that item response patterns correspond to what would be expected given relative item difficulty ordering

Nonparametric methods for modeling item responses: Mokken Scale Analysis

Nonparametric methods are descriptive and inferential statistical techniques that are not directly tied to a particular probability distribution. Nonparametric methods can be usefully applied to data that are used in educational research in order to address a variety of research questions. Nonparametric methods include standalone (a-theoretical) statistical tests (e.g., kappa coefficients), data exploration techniques (e.g., functional data analysis), as well as methods situated within modeling frameworks (e.g., Mokken Scale Analysis; discussed later in the article). The aim of this article is to provide an overview of nonparametric modeling frameworks that can be used to evaluate measurement instruments in educational research, such as achievement tests and attitude scales.

The structure of this article is as follows: The article begins with a description of nonparametric modeling frameworks in general. Then, a detailed discussion of Mokken Scale Analysis (MSA; Mokken, 1971), which is a theory-driven nonparametric modeling framework for social science measurement procedures, is provided with brief illustrations. The article continues with a brief overview of other nonparametric modeling approaches. The article concludes with a summary and discussion of the value of nonparametric modeling frameworks for educational research.

Nonparametric modeling frameworks for educational research

Modeling frameworks provide researchers with a structure for analyzing quantitative data and interpreting the results based on guiding principles or requirements. Particularly in the area of educational and psychological measurement, the general goal of modeling frameworks is to help analysts examine responses with the purpose of improving measurement procedures and informing the interpretation and use of the results from the measurement procedure. These approaches stand in contrast to data analysis techniques in which the ultimate goal is to describe relationships among variables or to replicate the structure of a dataset.

A variety of modeling frameworks for educational and psychological measurement are available that include both parametric and nonparametric techniques. Modeling frameworks based on *parametric* techniques impose a specific mathematical structure on the shape of the relationship between the latent variable and the probability for a response (i.e., the item response function [IRF]). For example, Rasch measurement theory (Rasch, 1960) is a parametric modeling framework in which the measurement models transform raw responses from an ordinal scale to an interval-level scale and require that IRFs conform to a logistic ogive (see Fig. 1). On the other hand, nonparametric modeling frameworks include, at most, ordinal requirements for the characteristics of the IRF. Nonparametric models do not involve mathematical transformations to achieve interval-level units. For example, models within the Mokken Scale Analysis (MSA (Mokken, 1971); framework allow IRFs to take on any shape as long as basic ordering requirements for individual items and groups of items are satisfied.

Parametric modeling frameworks are useful because their models result in convenient mathematical properties, including interval-level measures for items and persons, which are required for many statistical analyses. For example, interval-level estimates of item and person locations on a latent variable calculated using parametric IRT models can be applied to statistical techniques including Computer-Adaptive Testing and test equating. Although these applications are critical for many practical purposes, parametric modeling is not always necessary or appropriate for the measurement procedures that are used in educational research (Wind, 2017).

Parametric IRT models are characterized by several strong assumptions that are not always appropriate for educational and psychological measurement procedures (Santor and Ramsay, 1998; Junker and Sijtsma, 2001; Molenaar, 2001; Meijer and Baneke, 2004). For example, the appropriate interpretation of parameters such as examinee locations and item locations on a latent variable requires approximate adherence to the assumption that a logistic ogive (see Fig. 1A) accurately represents the relationship between examinee locations and item locations, and that this relationship holds for all examinees and all items. In addition, these models assume that the sample of examinees and items is sufficient to support precise parameter estimates for examinees and items that hold for the entire sample. Violations of these assumptions imply that the results from parametric IRT models may not have a meaningful interpretation, may be imprecise, and may not hold over replications of the measurement procedure. In addition, the

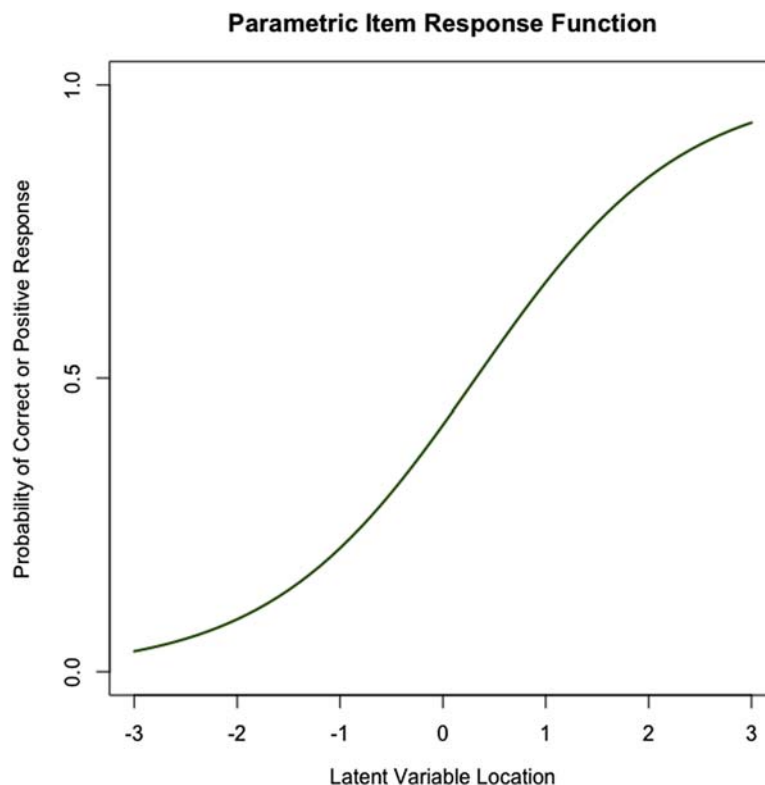


Fig. 1 Parametric item response function.

relatively stringent requirements of parametric IRT models may lead researchers to discard items for which model violations are observed at some locations on the latent variable, but not others (Santor and Ramsay, 1998; Meijer and Baneke, 2004). Reflecting on these challenges with parametric IRT models, Molenaar (2001) observed “there remains a gap ... between the organized world of a mathematical measurement model and the messy world of real people reacting to a real set of items” (p. 295). Although nonparametric IRT models are also imperfect representations of the “messy” nature of item responses, they provide researchers with tools for modeling these responses that provide additional flexibility that can be more appropriate and realistic for many contexts in which parametric IRT models may not be appropriate.

MSA is a practically useful approach to nonparametric IRT because it relies on requirements related to *invariance* in measurement procedures. The current article will focus on the nonparametric MSA modeling framework because it is a comprehensive nonparametric framework that can be usefully applied to evaluate measurement procedures in educational research. Following the discussion of MSA, some other nonparametric techniques that are relevant for educational research are briefly considered.

Mokken Scale Analysis framework

MSA is a probabilistic nonparametric framework for modeling responses from measurement procedures in the social and behavioral sciences. In his original presentation of MSA, Mokken (1971) discussed the usefulness of parametric IRT models in general, and Rasch measurement theory models (Rasch, 1960) in particular. He noted that these models have properties that are ideal for conducting mathematical analyses with social science data, but that they may not be appropriate for many practical applications. Mokken (1971) summarized the rationale for his approach to item response modeling as follows:

In vast areas of social research the application of parametric models may often be too far fetched. Their application presupposes a relatively deep insight into the structure of the variable to be measured and the properties of the items by which it can be measured. For these reasons it seems legitimate to try to find starting points for scaling models which do not rely too heavily upon specific parametric assumptions, as these lead to procedures of inference and estimation that are too pretentious and intricate for the level of information and the precision that can be claimed for the data used in actual measurement (p. 173).

In his presentation of MSA, Mokken noted that his scaling models are useful starting points for analyzing item response data before the application of parametric IRT models. However, Mokken (1971) and other researchers (e.g., Cliff and Keats, 2003; van Schuur, 2011) have also noted that nonparametric IRT models are also useful in their own right in contexts where an ordinal level of measurement is sufficient to inform the decisions that will be made based on a measurement procedure.

Working within the context of political science, Mokken proposed two nonparametric scaling models that reflect the emphasis on invariance that characterizes Rasch measurement theory, but that do not involve parametric transformations. Because of this shared emphasis on invariance, several researchers have discussed the theoretical and empirical similarities between MSA and Rasch measurement theory (e.g., Meijer et al., 1990; Engelhard, 2008). After their original presentation in the context of political science, MSA models earned popularity in contexts such as quality of life measurement (Vaughan et al., 2020), intelligence testing (Abdelhamid et al., 2020), clinical psychology (Tabares et al., 2020), physical therapy and rehabilitation (Rinkel et al., 2019), among others. In addition, researchers have recently applied MSA to educational assessment contexts, including multiple-choice assessments (e.g., Wind, 2016b) and performance assessments that involve rater judgments (Wind, 2016a, 2019).

Mokken Scale Analysis models

MSA includes two nonparametric IRT models: (1) the Monotone Homogeneity model and (2) the Double Monotonicity model. Mokken (1971) proposed the dichotomous formulations of these models in his original presentation of MSA. Extending these models, Molenaar (1982, 1997) presented polytomous formulations that have been applied across a wide range of social science contexts. The next sections introduce these models, followed by a discussion of recent variations on MSA models.

Monotone Homogeneity model

The first MSA model is the Monotone Homogeneity (MH) model. This model is characterized by three key assumptions: (1) *Unidimensionality*: One latent variable is sufficient to explain the variation in examinee responses; (2) *Local independence*: After controlling for the primary latent variable, there is no dependence between examinee responses to unique components of the measurement procedure (e.g., items), and (3) *Monotonicity*: As examinee locations on the latent variable increase, so does the probability for a correct or positive response. Both the dichotomous and the polytomous formulations of the MH model are characterized by these assumptions.

Because MSA models are nonparametric, the assumptions are evaluated using examinee locations calculated with *restscores*. Restscores are examinee total scores minus their score on the item of interest. Restscores avoid the problem that the item of interest is included the total score, such that researchers can evaluate each item independently. The MH model assumptions can be evaluated at the item level using restscores calculated as:

$$R_{ij} = X_{+i} - X_{ij},$$

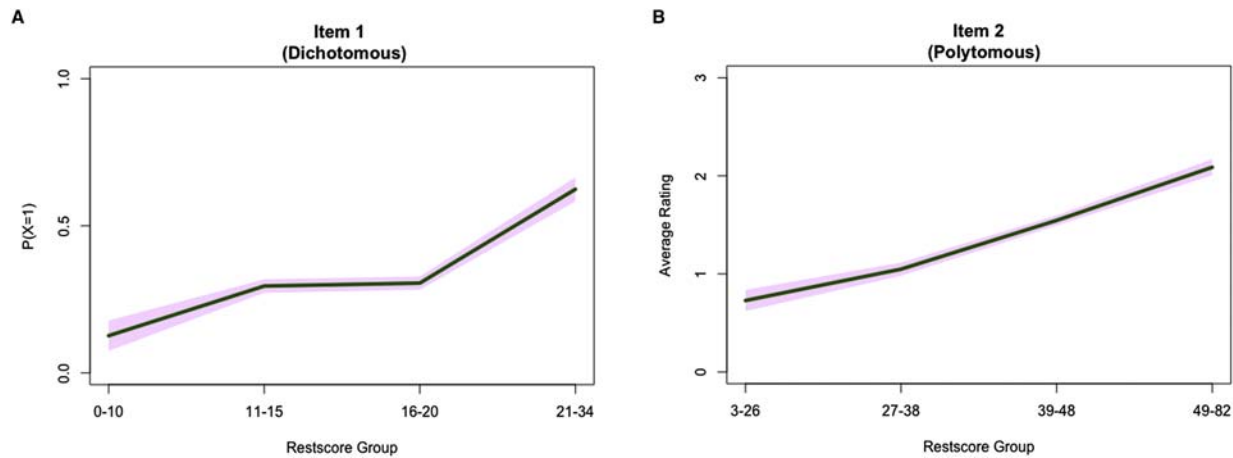


Fig. 2 Item monotonicity plots. A: Two dichotomous items B: Two polytomous items.

where R_{ij} is the restsore for Examinee i specific to Item j , X_{+i} is the total score for Examinee i on all items, and X_{ij} is Examinee i 's score on Item j . In practice, researchers use restscores to evaluate adherence to the MH model assumptions with two primary indices: (1) numeric and graphical tests for monotonicity, and (2) scalability coefficients.

When it is applied to dichotomous data (scored such that $x = 0$ or $x = 1$) the monotonicity assumption states that as examinee locations on the latent variable increase, the probability for a correct or positive response ($x = 1$) does not decrease. **Fig. 2A** is a nonparametric IRF that illustrates a graphical approach to evaluating monotonicity for dichotomous Item 1. The x -axis shows *restscore groups*, which are subgroups of examinees with similar restscores who have been combined into groups to improve the statistical power of the test for monotonicity (see [Molenaar and Sijtsma \(2000\)](#) for a description of typical procedures for creating restsore groups in MSA). The y -axis shows the conditional probability for a correct or positive response specific to each restsore group. The solid line is the nonparametric IRF that shows the probability for a correct or positive response within each restsore group, and the light shading around the line shows a 95% confidence interval for these probabilities. Item 1 adheres to the monotonicity assumption because the probability for a correct or positive response is non-decreasing over increasing restsore groups.

When it is applied to polytomous data scored in categories $x = 0, \dots, k, \dots, m$, the monotonicity assumption states that as examinee locations on the latent variable increase, the probability for a rating in category k or higher ($x \geq k$) does not decrease. Although it is possible to evaluate item monotonicity within rating scale categories for polytomous items, most applications of MSA focus on overall item-level analyses. To evaluate monotonicity at the overall item level, analysts evaluate the extent to which average ratings on an item of interest are non-decreasing as examinee locations on the latent variable increase. Using restsore groups along the x -axis, **Fig. 2B** illustrates a graphical approach to evaluating monotonicity for polytomous Item 2, which was scored in four ordered categories ($x = 0, 1, 2, 3$). In this figure, the y -axis is the average rating observed within each restsore group. Item 2 adheres to the monotonicity assumption at the overall item level because the average rating observed within restsore groups does not decrease as restscores increase. In addition to graphical analyses, researchers can also evaluate item adherence to monotonicity using statistical hypothesis tests. For additional details about these tests, please see [Sijtsma and Molenaar \(2002\)](#).

Several properties are important to note regarding the nonparametric IRFs for the MH model, as illustrated in **Fig. 2**. First, the MH model IRFs can take on any shape as long as they adhere to the ordering requirement specified by the model; they do not have to conform to the logistic ogive shape that characterizes parametric IRT model IRFs. Second, when items adhere to the MH model assumptions, some important inferences can be made related to invariance. For the dichotomous model, adherence to the MH model implies *invariant person ordering over items*, which means that inferences about person ordering on the latent variable are consistent for all of the items in the scale. Specifically, the order of student total scores on the raw score scale (X_{+}) reflects their ordering on the latent variable (θ ; i.e., stochastic ordering on the latent variable [SOL]; [Hemker et al., 1997](#)). For the polytomous MH model, adherence to the MH model implies a weaker form of SOL. Specifically, *weak SOL* ([van der Ark and Bergsma, 2010](#)) means that examinee sum scores (X_{+}) can be used to divide the sample into two groups (low and high), and examinee classification within those groups can be used to identify students with low and high locations on the latent variable. Invariant person ordering is important because it means that the interpretation of examinee locations on the latent variable do not depend on the specific items or components of the scale used to measure them. As a result of these properties, researchers can check for adherence to the MH model assumptions to provide support for the interpretation and use of examinee sum scores with respect to the latent variable.

A second major indicator of adherence to the MH model is *scalability analyses*. Within the context of MSA, scalability analysis involves evaluating item responses for evidence that response patterns match what would be expected given item ordering. Specifically, scalability is evaluated by identifying *Guttman errors*. For dichotomous items, Guttman errors occur for a pair of items ordered such that Item C is easier than Item D, when a correct response to Item D is observed in combination with an *incorrect* response to Item C. For polytomous items, Guttman errors are identified at the rating scale category level, where the difficulty of each category is

calculated using cumulative category probabilities. MSA scalability coefficients were originally adapted from the scalability coefficients that [Loevinger \(1948\)](#) proposed for evaluating item pairs for evidence of Guttman errors. [Mokken \(1971\)](#) expanded Loevinger's coefficients to include scalability coefficients for individual items and pairs of items.

Scalability coefficients can be calculated using the ratio of observed-to-expected Guttman errors for individual items (H_i), pairs of items (H_{ij}), and a set of 3 or more items (H). The general form of these coefficients can be illustrated using the formula for the item-pair scalability coefficient:

$$H_{ij} = 1 - \frac{F_{ij}}{E_{ij}},$$

Where F_{ij} is the observed frequency of Guttman errors for the item pair made up of Item i and Item j , and E_{ij} is the expected frequency of Guttman errors given the marginal distributions of the responses for Item i and Item j . For individual items, scalability is calculated as:

$$H_i = 1 - \frac{\sum_{j \neq i} F_{ij}}{\sum_{j \neq i} E_{ij}},$$

For a set of three or more items, scalability is calculated as:

$$H = 1 - \frac{\sum_i F_{i,R(i)}}{\sum_i E_{i,R(i)}},$$

where $R(i)$ is the Restscore for Item i . It is also possible to calculate scalability coefficients using the covariances between pairs of items. For details about the covariance formulations of the scalability coefficients, please see [Sijtsma and Molenaar \(2002\)](#).

In applications of MSA, researchers frequently interpret these coefficients using critical values that [Mokken \(1971\)](#) proposed: $H \geq 0.5$: strong scalability, $0.3 \leq H < 0.5$: moderate scalability, and $H < 0.30$: weak scale. However, these critical values have not been thoroughly examined in methodological research, particularly for polytomous item responses. Instead of relying on critical values, researchers have recommended using standard errors (SE) for scalability coefficients ([Kuijpers et al., 2013](#)), interpreting scalability coefficients as continuous variables, and considering the implications of their values in the context of the specific application.

Double Monotonicity model

The second MSA model is the Double Monotonicity (DM) model. The dichotomous DM model ([Mokken, 1971](#)) and the polytomous DM model ([Molenaar, 1982, 1997](#)) share the same assumptions as the MH model, but include one additional assumption: (4) *Invariant item ordering* (IIO): Items have a consistent relative ordering for all examinees, regardless of their location on the latent variable. When data adhere to the DM model, a second important invariance property is observed in addition to invariant person ordering from the MH model: item ordering is invariant across examinees. This means that conclusions about the relative difficulty of a set of items are consistent for all examinees in a sample of interest. As a result, the DM model provides information about invariance related to person ordering and item ordering. These properties are useful for scale construction and evaluation when a consistent interpretation of item and person order on the latent variable is important for theoretical and practical purposes ([Ligtvoet, 2010](#)).

When it is applied to dichotomous item responses, the DM model assumption of IIO can be evaluated by examining the probability for a correct or positive response for *pairs* of items for evidence that the item ordering is consistent across the range of the latent variable. As in the MH model, an ordinal formulation of the latent variable is needed to evaluate IIO. Because the procedure for checking the DM model assumption of IIO is based on item pairs, restscores are calculated as follows:

$$R_{ijk} = X_{+i} - X_{ij} - X_{ik},$$

where R_{ijk} is the restscore for Examinee i specific to the item pair made up of Item j and Item k , X_{+i} is the total score for Examinee j on all items, X_{ij} is Examinee i 's score on Item j , and X_{ik} is Examinee i 's score on Item k . As with the MH model, examinee restscores are combined into restscore groups for the purpose of evaluating the DM model assumption of IIO.

In practice, researchers evaluate IIO for dichotomous items using graphical displays of nonparametric IRFs and statistical hypothesis tests. [Fig. 3A](#) illustrates nonparametric IRFs that were constructed for the purpose of evaluating IIO for the item pair made up of dichotomous Item 3 and dichotomous Item 4. In the figure, the x -axis shows restscore groups, and the y -axis shows the probability for a correct or positive response. For this item pair, IIO is observed because Item 4 is always easier (higher probabilities for a correct or positive rating) compared to Item 3 for all restscore groups. Importantly, the nonparametric IRFs can take on any shape as long as they satisfy the ordering requirements for the MH and DM models.

For polytomous items, IIO can be evaluated at the overall item level or within rating scale categories. Importantly, some researchers have observed that invariant ordering at the rating scale category level does not always correspond to invariant ordering for overall item scores ([Sijtsma and Hemker, 1998](#); [Ligtvoet et al., 2010, 2011](#); [Sijtsma et al., 2011](#)). As a result, researchers typically focus on invariance at the overall item level because this indicator has more practical utility compared to investigations of invariance for specific rating scale categories ([Ligtvoet et al., 2010](#)). Specifically, researchers have encouraged analysts to use the manifest IIO

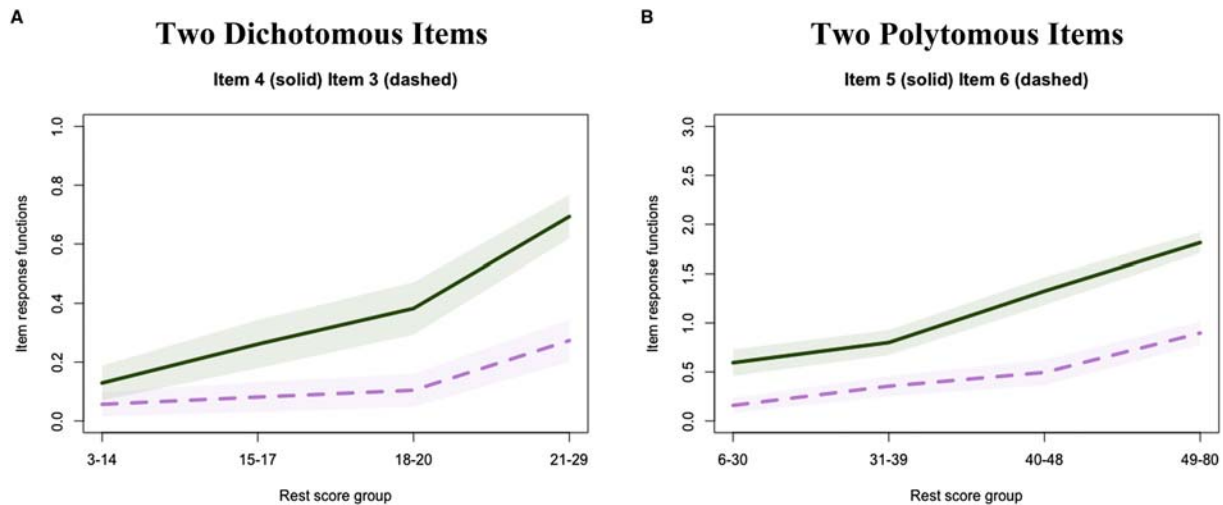


Fig. 3 Invariant item ordering plots.

procedure (MIIO) that [Ligtvoet et al. \(2010, 2011\)](#) proposed. The MIIO procedure involves evaluating IIO for overall polytomous items, rather than evaluating individual rating scale categories for invariant ordering. [Fig. 3B](#) shows a MIIO plot for polytomous Item 5 and Item 6, both of which were scored in four categories ($x = 0, 1, 2, 3$). The x -axis shows restsore groups, and the y -axis shows the raw score scale. The IRFs show the average rating on each item observed within each restsore group. Item 5 and Item 6 adhere to IIO because Item 5 is easier (higher average ratings) compared to Item 6 for all restsore groups.

In practice, researchers typically evaluate IIO using plots such as those shown in [Fig. 3](#). In addition, researchers use statistical hypothesis tests that take into account the magnitude of violations of IIO to identify statistically significant violations that warrant additional investigation. For more details about statistical hypothesis tests for evaluating IIO, please see [Sijtsma and Molenaar \(2002\)](#).

Software for MSA

Several software programs are available within which researchers can apply MSA models. Specifically, [Molenaar and Sijtsma \(2000\)](#) presented *MSP5*, which is a standalone software program that researchers can use to analyze data with MSA models. However, *MSP5* is no longer being updated to reflect recent developments in MSA. In addition, a Mokken scaling module has been developed for use with the Stata software program ([Hardouin et al., 2011](#)). Finally, researchers can conduct MSA using the *mokken* package for the R software program ([van der Ark, 2007, 2012](#)). This package is frequently updated to reflect methodological developments in MSA.

Variations on MSA

Recently, researchers have proposed several variations on MSA that facilitate analyses that are better suited to specific data collection designs or research questions than the original dichotomous and polytomous MSA models. For example, [Koopman et al., 2016; Koopman et al., 2020; Koopman et al., 2020](#)) and colleagues have presented adaptations of MSA statistics that are designed for use with multi-level (i.e., nested) data structures. In addition, [Wind \(2016a, 2019; Wind and Schumacker, 2017\)](#) presented an adjacent-categories adaptation of MSA indices that can be used to identify rating scale category disordering in contexts where this information is relevant, such as in rater-mediated educational performance assessments.

Other nonparametric approaches

This article focused on MSA because it is a nonparametric approach that is situated within a modeling framework that provides guidance for interpreting analyses based on a measurement theory. However, other nonparametric techniques are available that researchers can use to analyze educational research data that are not suited for analysis with MSA, in contexts where a measurement framework is not needed for modeling item response data, or in contexts where an exploratory approach is needed before measurement-focused analyses are conducted.

For example, a number of nonparametric statistical tests are available that are not based on any particular modeling framework but provide insight into a variety of research questions that may be of interest in educational research. For example, kappa statistics ([Cohen, 1968; Berry and Mielke, 1988; Schuster, 2004](#)) provide information about rater agreement, but these statistics are not based on a measurement framework.

Another type of nonparametric analysis that is useful in educational research is functional data analysis ([Ramsay and Silverman, 2005](#)). Functional data analysis provides valuable exploratory tools for examining data in a variety of contexts, including a number of variables that are relevant in educational research. Among functional data analysis techniques, kernel smoothing ([Ramsay, 1991](#);

Mazza et al., 2014) can be used to estimate and visualize response functions for item response data in educational research. Although functional data analysis is not directly tied to a measurement framework, it offers insight into response patterns and other characteristics of data that can help inform the interpretation and use of measurement instruments and can also inform future analytic decisions.

Summary and conclusions

The purpose of this article was to provide an overview of nonparametric modeling frameworks that are useful for examining the properties of measurement instruments in educational research. First, a general consideration of modeling frameworks was presented and a distinction was made between parametric and nonparametric modeling frameworks for analyzing responses to educational and psychological measurement instruments. The major focus of the article was on Mokken Scale Analysis (MSA; Mokken, 1971) because MSA is a theory-driven nonparametric modeling framework that offers several useful indices for evaluating measurement instruments that are particularly useful in educational research. An overview of MSA was provided, including a description of the two major MSA models and procedures for checking their assumptions. The article concluded with a brief discussion of other nonparametric analysis techniques that may be relevant for educational research.

The major conclusion from this article is that nonparametric modeling frameworks such as MSA are available that can provide educational researchers with theory-based analytic techniques for evaluating their measurement instruments but that do not rely on potentially inappropriate parametric techniques. These methods can be used as a standalone procedure and they can also be used to supplement parametric analyses either as a preliminary analytic step or to provide additional information beyond what is captured by parametric models. These nonparametric frameworks include numeric and graphical techniques for evaluating the properties of measurement instruments in educational research that can provide insight into the degree to which instruments exhibit important measurement properties, such as invariant item and person ordering.

References

- Abdelhamid, G.S.M., et al., 2020. A demonstration of mokken scale Analysis methods applied to cognitive test validation using the Egyptian WAIS-IV. *J. Psychoeduc. Assess.* 38 (4), 493–506. <https://doi.org/10.1177/0734282919862144>.
- van der Ark, L.A., 2007. Mokken scale analysis in R. *J. Stat. Software* 20 (11), 1–19.
- van der Ark, L.A., 2012. New developments in Mokken scale analysis in R. *J. Stat. Software* 48 (5), 1–27.
- van der Ark, L.A., Bergsma, W.P., 2010. A note on stochastic ordering of the latent trait using the sum of polytomous item scores. *Psychometrika* 75 (2), 272–279. <https://doi.org/10.1007/s11336-010-9147-7>.
- Berry, K.J., Mielke, P.W., 1988. 'A generalization of cohen's kappa agreement measure to interval measurement and multiple raters'. *Educ. Psychol. Meas.* 48 (4), 921–933. <https://doi.org/10.1177/0013164488484007>.
- Cliff, N., Keats, J.A., 2003. *Ordinal Measurement in the Behavioral Sciences*. Lawrence Erlbaum Associates Publishers. Available at: <http://psycnet.apa.org/psycinfo/2002-06630-000>. (Accessed 20 October 2016).
- Cohen, J., 1968. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol. Bull.* 70 (4), 213–220. <https://doi.org/10.1037/h0026256>.
- Engelhard, G., 2008. Historical perspectives on invariant measurement: guttman, Rasch, and mokken. *Measurement* 6 (3), 155–189. <https://doi.org/10.1080/15366360802197792>.
- Hardouin, J.-B., Bonnaud-Antignac, A., Sebille, V., 2011. Nonparametric item response theory using Stata. *STATA J.* 11 (1), 30–51.
- Hemker, B.T., et al., 1997. Stochastic ordering using the latent trait and the sum score in polytomous IRT models. *Psychometrika* 62 (3), 331–347. <https://doi.org/10.1007/BF02294555>.
- Junker, B.W., Sijtsma, K., 2001. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Appl. Psychol. Meas.* 25 (3), 258–272. <https://doi.org/10.1177/01466210122032064>.
- Koopman, L., et al., 2020a. Bias of two-level scalability coefficients and their standard errors. *Appl. Psychol. Meas.* 44 (3), 197–214. <https://doi.org/10.1177/0146621619843821>.
- Koopman, L., Zijlstra, B.J.H., van der Ark, L.A., 2016. Weighted guttman errors: handling ties and two-level data. In: *Quantitative Psychology. The Annual Meeting of the Psychometric Society*. Springer, Cham, pp. 183–190. https://doi.org/10.1007/978-3-319-56294-0_17 (Springer Proceedings in Mathematics & Statistics).
- Koopman, L., Zijlstra, B.J.H., van der Ark, L.A., 2020b. New advances in Mokken scale analysis. In: *Annual Meeting of the Psychometric Society*.
- Kuijpers, R.E., van der Ark, L.A., Croon, M.A., 2013. Standard errors and confidence intervals for scalability coefficients in mokken scale analysis using marginal models. *Socio. Methodol.* 43 (1), 42–69. <https://doi.org/10.1177/0081175013481958>.
- Ligtvoet, R., 2010. *Essays on Invariant Item Ordering*. Tilburg University. Available at: [https://pure.uvt.nl/portal/en/publications/essays-on-invariant-item-ordering\(872d8f6b-778e-49f7-8fdb-c74237a70ffc\).html](https://pure.uvt.nl/portal/en/publications/essays-on-invariant-item-ordering(872d8f6b-778e-49f7-8fdb-c74237a70ffc).html). (Accessed 22 February 2016).
- Ligtvoet, R., et al., 2010. Investigating an invariant item ordering for polytomously scored items. *Educ. Psychol. Meas.* 70 (4), 578–595. <https://doi.org/10.1177/0013164409355697>.
- Ligtvoet, R., et al., 2011. Polytomous latent scales for the investigation of the ordering of items. *Psychometrika* 76, 200–216.
- Loevinger, J., 1948. The technic of homogeneous tests compared with some aspects of scale analysis and factor analysis. *Psychol. Bull.* 45 (6), 507–529.
- Mazza, A., Punzo, A., McGuire, B., 2014. KernSmoothIRT: an R package for kernel smoothing in item response theory. *J. Stat. Software* 58 (6). <https://doi.org/10.18637/jss.v058.i06>.
- Meijer, R.R., Baneke, J.J., 2004. Analyzing psychopathology items: a case for nonparametric item response theory modeling. *Psychol. Methods* 9 (3), 354–368. <https://doi.org/10.1037/1082-989X.9.3.354>.
- Meijer, R.R., Sijtsma, K., Smid, N.G., 1990. Theoretical and empirical comparison of the mokken and the Rasch approach to IRT. *Appl. Psychol. Meas.* 14 (3), 283–298. <https://doi.org/10.1177/014662169001400306>.
- Mokken, R.J., 1971. *A Theory and Procedure of Scale Analysis*. De Gruyter, The Hague: Mouton/Berlin.
- Molenaar, I.W., 1982. Mokken scaling revisited. *Kwantitative Methoden* 3 (8), 145–164.
- Molenaar, I.W., 1997. Nonparametric models for polytomous responses. In: *Handbook of Modern Item Response Theory*. Springer, New York, NY, pp. 369–380.
- Molenaar, I.W., 2001. Thirty years of nonparametric item response theory. *Appl. Psychol. Meas.* 25 (3), 295–299.

- Molenaar, I.W., Sijtsma, K., 2000. MPS5 for Windows: A Program for Mokken Scale Analysis for Polytomous Items. ProGAMMA, Groningen, The Netherlands.
- Ramsay, J.O., 1991. Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika* 56 (4), 611–630. <https://doi.org/10.1007/BF02294494>.
- Ramsay, J.O., Silverman, B.W., 2005. *Functional Data Analysis*, second ed. Springer, New York.
- Rasch, G., 1960. *Probabilistic Models for Some Intelligence and Achievement Tests*. Expanded Edition, 1980. University of Chicago Press, Chicago.
- Rinkel, W.D., et al., 2019. Development of grading scales of pedal sensory loss using Mokken scale analysis on the Rotterdam Diabetic Foot Study Test Battery data. *Muscle Nerve* 60 (5), 520–527. <https://doi.org/10.1002/mus.26628>.
- Santor, D.A., Ramsay, J.O., 1998. Progress in the technology of measurement: applications of item response models. *Psychol. Assess.* 10 (4), 345–359.
- Schuster, C., 2004. A note on the interpretation of weighted kappa and its relations to other rater agreement statistics for metric scales. *Educ. Psychol. Meas.* 64 (2), 243–253. <https://doi.org/10.1177/0013164403260197>.
- van Schuur, W.H., 2011. *Ordinal Item Response Theory: Mokken Scale Analysis*. Sage, Los Angeles, CA (Quantitative Applications in the Social Sciences, 09–169).
- Sijtsma, K., Hemker, B.T., 1998. Nonparametric polytomous IRT models for invariant item ordering, with results for parametric models. *Psychometrika* 63 (2), 183–200. <https://doi.org/10.1007/BF02294774>.
- Sijtsma, K., Meijer, R.R., van der Ark, L.A., 2011. Mokken scale analysis as time goes by: an update for scaling practitioners. *Pers. Individ. Differ.* 50 (1), 31–37. <https://doi.org/10.1016/j.paid.2010.08.016>.
- Sijtsma, K., Molenaar, I.W., 2002. *Introduction to Nonparametric Item Response Theory*. Sage, Thousand Oaks.
- Tabares, J.V., et al., 2020. Mokken scale Analysis of lifetime responses on the columbia suicide severity rating scale's severity of ideation subscale. *Assessment*. <https://doi.org/10.1177/1073191120913626>, 1073191120913626.
- Vaughan, B., Mulcahy, J., Fitzgerald, K., 2020. PROMIS[®] General Life Satisfaction scale: construct validity in musculoskeletal pain patients. *Chiropr. Man. Ther.* 28 (1), 27. <https://doi.org/10.1186/s12998-020-00320-x>.
- Wind, S.A., 2016a. Adjacent-categories Mokken models for rater-mediated assessments. *Educ. Psychol. Meas.* <https://doi.org/10.1177/0013164416643826>.
- Wind, S.A., 2016b. Examining the psychometric quality of multiple-choice assessment items using mokken scale Analysis. *J. Appl. Meas.* 17 (2), 142–165.
- Wind, S.A., 2017. An instructional module on Mokken scale analysis. *Educ. Meas.* 36 (2), 50–66.
- Wind, S.A., 2019. Nonparametric evidence of validity, reliability, and fairness for rater-mediated assessments: an illustration using mokken scale Analysis. *J. Educ. Meas.* 56 (3), 478–504.
- Wind, S.A., Schumacker, R.E., 2017. Exploring within-rater category ordering: a simulation study using adjacent-categories mokken scale Analysis. *Educ. Psychol. Meas.* <https://doi.org/10.1177/0013164417724841>, 001316441772484.

Principled assessment design

Mark Wilson and Sean Tan, University of California, Graduate School of Education, UC Berkeley, Berkeley, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	146
The BEAR assessment system	147
The construct map (building block 1)	148
The items design (building block 2)	149
The outcome space (building block 3)	149
The measurement model (building block 4)	151
Selected principled assessment design approaches	153
Evidence-centered design	153
Cognitive design system	155
Assessment engineering	155
Principled design for efficacy	157
Going “outside the triangle”	159
Conclusion	160
References	161

Introduction

What is *principled assessment design*, or PAD? And, where do those principles come from? One part of the answer to these questions is that the current usage of the term arose in the impact of cognitive psychology on testing, principally in education, but also in psychology. Under the impetus of writing by figures such as [Glaser \(1981\)](#), and [Snow and Lohman \(1989\)](#), new demands were made of tests to respond to criticisms like the following from a very influential National Research Council (NRC) Committee Report:

The central problem ... is that most widely used assessments of academic achievement are based on highly restrictive beliefs about learning and competence not fully in keeping with current knowledge about human cognition and learning. Likewise, the observation and interpretation elements underlying most current assessments were created to fit prior conceptions of learning and need enhancement to support the kinds of inferences people now want to draw about student achievement.

NRC (2001, pp. 2–3).

The suggested solution was to base the logic of the measurement that is embodied in the test in the theories and practices of cognitive psychology. In one sense, this was just a re-statement of an older adage about the question of test validity—that a test is valid if it does what it is supposed to do—but the suggestion is that cognitive psychology should be the source of the property that is being measured by the test. However, it was also a serious challenge to the then-dominant role of “content” in defining the goal of testing, a goal that is so centrally embodied in the mastery-testing formulation of pedagogic design ([Bloom, 1981](#)).

With the help of some forward thinking among psychometricians (as argued, for example, by [Mislevy \(1996\)](#) in his aptly-named paper “Test theory reconceived”), the NRC established a Committee to acknowledge and advance this perspective ([NRC, 2001](#)). The Committee established a common-ground delineation of what aspects should be incorporated into this perspective, which they referred to as foundations. In the words of the Committee:

First, every assessment is grounded in a conception or theory about how people learn, what people know, and how knowledge and understanding progress over time. Second, each assessment embodies certain assumptions about which kinds of observations, or tasks, are most likely to elicit demonstrations of important knowledge and skills from students. Third, every assessment is premised on certain assumptions about how best to interpret the evidence from the observations in order to make meaningful inferences about what students know and can do (p. 16).

This is illustrated in the NRC Assessment Triangle shown in [Fig. 1](#), with the three foundations succinctly labeled as Cognition, Observation, and Interpretation. The foundations then, are seen as constituting a guide for “the process of collecting evidence to support the types of inferences one wants to draw ... referred to as *reasoning from evidence* [emphasis in original]” ([Mislevy, 1996](#), p. 38). These ideas then constitute a founding attempt at laying out the principles in PAD. The fact that the Committee labeled them as “foundations” reflects the implicit view that they also constitute a philosophical foundation for the measurement involved in the assessments (see [Mari et al. \(2021\)](#) for a perspective on this). Generally, the collection of such evidence is referred to in the

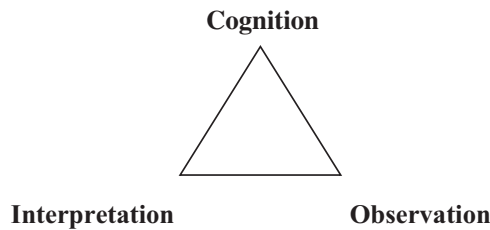


Fig. 1 The National Research Council (NRC) assessment triangle.

testing literature as test validity, and the appropriate standards for current professional practices regarding such evidence is delineated in the “Testing Standards” (American Educational Research Association et al., 2014). Although these Standards are designed to be industry-wide, and hence apply to many forms and varieties of assessment that would not be seen as conforming to PAD ideas (or even be vaguely associated with such), in the main, the types of validity evidence listed in the Testing Standards are very consistent with the perspectives described in this article, which are much influenced by the newer perspective described above. For example, to read an account of how the strands of validity evidence given in the Testing Standards can be related to one of the PAD approaches—the BEAR Assessment System, described in the next section—see chapter 8 of Wilson (2004).

Some reflection of the influence of the NRC Assessment Triangle conceptualization can be seen by considering two subsequent reviews of the topic of principled assessment designs (Ferrara et al., 2016; Nichols et al., 2016). Both are useful and insightful accounts, and we acknowledge a debt to their work. We have summarized their delineations of the topic in Table 1—we will not go into depth on their systems but save this level of detail for the specific PAD approaches discussed below. As can be seen, Ferrara et al. (2016) postulate five “foundation” elements of PAD, and these can be readily assigned to the three vertices of the NRC Assessment Triangle, with the first three aligned to the Cognition vertex. In contrast, the Nichols et al. (2016)’s system is composed of three elements, and each of these is quite consistent with one of the three NRC vertices. In addition, both systems argue that the elements should be seen as part of a larger whole—each contributing to the crucially important validity argument, as is the case for the NRC Assessment Triangle.

In the sections that follow, we will first give a detailed example of one particular PAD approach, pointing out how the NRC foundations play a central role in the structure of that approach. This is followed by descriptions of several other selected PAD approaches to show the diversity of ways in which the idea of a principled design can be exploited, while still maintaining a similarity with the NRC formulation. The article finishes with a section containing a discussion of links beyond assessment, and some final concluding remarks.

The BEAR assessment system

The BEAR Assessment System (BAS; Wilson and Sloane, 2000) uses four “building blocks” to address the PAD challenges for good assessment: (a) construct map, (b) items design, (c) outcome space, and (d) measurement model. As expressed in Fig. 2, these building blocks are seen as being used in a specific cyclical sequence during assessment development, but the cycle may also iterate during development. These four building blocks embody the three foundations from the NRC Assessment Triangle. In other words, the foundations are also seen as being principles for assessment development. The match, in sequence, is (a) construct map to Cognition, (b) items design to Observation, and (c) outcome space and measurement model (together) to Interpretation. This matching is enlarged upon below, in the respective subsections of this section. For a detailed account of an instrument development

Table 1 A summary of three different, but related, systems for Principled Assessment Design (PAD).

NRC (2001): <i>Assessment triangle</i>	Nichols et al. (2016): <i>Three common characteristics of PAD approaches</i>	Ferrara et al. (2016): <i>Five foundation elements and one organizing element of PAD approaches</i>
Cognition	Construct-centered approach	Clearly defined assessment targets Statement of intended score interpretations and uses Model of cognition, learning, or performance
Observation	Explicit design decisions and rationales	Manipulation of assessment activities to align with assessment
Interpretation	Engineering toward intended interpretations and uses	Aligned measurement models and reporting scales
Overall evidentiary reasoning goal	Evidentiary coherence to enhance validity	Ongoing accumulation of evidence to support validity arguments

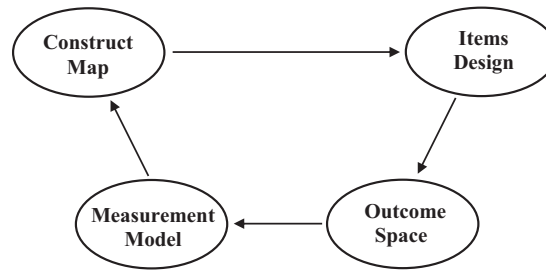


Fig. 2 Four building blocks in the BEAR Assessment System (BAS).

process that works through these building blocks, see [Wilson \(2004\)](#). Below, each of these building blocks is considered in turn, and related to the PAD challenges.

The construct map (building block 1)

The construct map (see [Fig. 3](#) for an example) represents a cognitive and developmental perspective on student learning (corresponding to Cognition—the first, top vertex of the NRC Assessment Triangle). Rather than conceptualizing learning as simple as a matter of acquiring quantitatively more knowledge, skills, and understandings, learning is conceptualized in BAS as progress along the road to higher levels of cognition where new knowledge is linked to existing knowledge, and as less sophisticated

<i>Functions Learning Progression: Making Connections Across Multiple Representations of Mathematical Functions</i>		
Level of sophistication:	Student is able to:	Response to items:
(5) Extended Abstract	Make connections not only within the given subject area, but also beyond it, able to generalize and transfer the principles and ideas underlying the specific instance.	- Predict, explain, and synthesize their understanding to a real-world context. - Solve non-routine problems including non-algebraic functions.
(4) Relational	Demonstrate understanding of the significance of the parts in relation to the whole.	- Compare/contrast information given in multiple representations of functions to demonstrate understanding of the content. - Recognize which representations to choose for the context. - Select representations and move fluently among them to achieve a solution.
(3) Multistructural	Make a number of connections, but the meta-connections between them are missed, as is their significance for the whole.	- Make connections across more than two representations (e.g., symbolic, tabular, graphic, and verbal representations). - Recognize more than one relevant feature of a functional relationship
(2) Unistructural	Make simple and obvious connections, but their significance is not grasped.	Make pair-wise connections between representations.
(1) Prestructural	Acquire bits of unconnected information, which have no organization and make no sense.	- Interpret graphs where both variables have to be interpreted, or where time as independent variable. - Demonstrate understanding of function in one representation.
(0) Prealgebraic	Acquire prerequisite skills.	Demonstrate understanding of: - functional dependence, where a change in one variable affects another variable. - continuous variables (time, distance) and dichotomous variables (hot or cold). - relative order and measurement to define variables.

Fig. 3 Functions construct map (Note: read from bottom to top).

understandings are substituted by deeper understandings. It is assumed that learning can be described and mapped as a progression toward qualitatively richer knowledge, higher-order skills, and deeper understandings. This progression is not necessarily along a single path, but may be represented as progression along multiple paths, in say, a multi-dimensional representation.

These progression paths are thought of as developing variables of student ability. They are drawn, in part, from research on the foundational cognitive structure of the assessment domain, and also from educators' professional views about the nature of higher and lower levels of performance, and how it is manifested in student work. An important component of that is the empirical research on how students respond to instruction or perform in practice (NRC, 2001), including in typical testing situations, as well as in classroom-embedded contexts. The construct map can be seen as a "boundary object" mediating between the assessment and the instruction that a teacher might devise based on the assessment outcomes, but also a mediator between the assessment developer and the curriculum structure that defines the context in which the assessment will be used (Fisher and Wilson, 2019; Lehrer et al., 2014). An example using the "Functions" construct from an assessment of college-readiness in mathematics (Wilmot et al., 2011) is described below to illustrate this.

The Functions construct was conceived using the Structure Of the Learning Outcome (SOLO) taxonomy (Biggs and Collis, 1982), a conceptual framework that is used to assess student ability to make relational connections within and across subject areas. The aim was to model multiple representations of mathematical functions, and then design an assessment that comports with that structure. The typical application of the SOLO taxonomy has five levels of student understanding. This assessment extended the taxonomy by adding a sixth level at the bottom called "Prealgebraic." The Functions Learning Progression consists of a single construct map shown in Fig. 3—note that this is somewhat unusual, as most learning progressions are conceived as being more complex (Wilson, 2009)—it has been chosen as an example because it is relatively easier to compress it into a short article (see in Schwartz et al. (2017) for an example of a more typical, complex learning progression consisting of several inter-related construct maps). Student responses are classified from lowest to highest, representing increasing sophistication in student understanding of the connections among mathematical functions. For details about these levels and the expected student responses, see Wilmot et al. (2011, pp. 265–266).

Creating such a construct map is not a trivial task—see Wilmot et al. (2011) for an account of the developmental steps involved in developing this one. However, having succeeded in formulating a set of construct maps within a given curriculum, the task of dealing with the challenges raised in the first part of this article can be readily assayed, though the steps are also not at all trivial. This approach assumes a match between instruction and assessment, which will be addressed next.

The items design (building block 2)

The items design, which is a plan for systematic observations (the second vertex of the NRC Assessment Triangle) using assessment tasks or activities, along with the construct maps, to establish the match between the instruction and assessment in BAS. Essentially, the core framework for the curriculum and instruction must be the same as the set of construct maps for the assessments. This is not to imply that the curriculum description should completely encompass the assessment, nor that the curriculum should be driven by the assessment, but instead that the two (assessment and instruction) need to be aligned. More specifically, the range and contents of the instructional practices used in the curriculum should be reflected in the assessment tasks, and the assessment tasks must be placed in appropriate spots to adhere to the "rhythm" of the instruction—they must occur when it makes instructional sense to include them, typically be at where instructors want to know how much progress their students have made on a specific topic. This brings the content-depth and topic-relevancy of curriculum development into assessment, and at the same time, instills the discipline and hard-headedness of assessment data into the design of instruction.

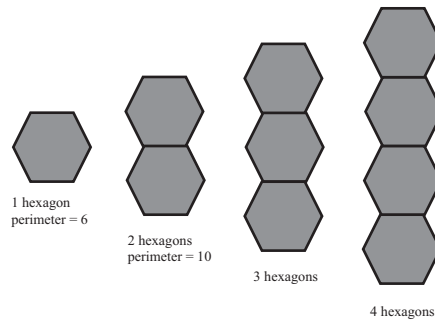
BAS can utilize a wide range of styles of assessment tasks, including tasks that involve student readings, laboratory exercises, interactive exercises, individual and group challenges, and data interpretation questions. The test based on the Functions construct map was composed of 12 constructed response items (for a full text of the test [in 6 different administration forms], see Wilmot (2008, pp. 173–302)). Students were asked to write their responses with words, algebraic formulas, and/or numbers. An example item, the Hexagon Pattern task, is shown in Fig. 4. If student work is generated or students can be observed at work, and this performance can be scored by matching it to the construct maps, then it is amenable for BAS regardless of the form of instruction. Developing both the instructional materials and the assessment tasks at (roughly) the same time is a good way to make this happen—by adapting and/or cloning good instructional sequences to produce assessable responses, and/or by developing successful assessments into full-blown instructional activities.

The outcome space (building block 3)

For instructors and students to use information from the assessment tasks and the BAS reports, it must be represented in terms of the levels in the construct maps. This is the first part of interpretation (the third vertex of the NRC Assessment Triangle). For selected response items, the outcome space is usually directly related to subsets of the selections, but for open-ended items, the situation is more complicated—open-ended tasks (with written, performance, or product outputs) must be speedily and reliably scorable. Thus, for open-ended tasks, one must ensure that the outcome space (or the "rubric") is either efficiently scorable by relevant professionals, such as teachers, students themselves, readers, teaching assistants and/or instructors, or scorable by machine, say, using web-based interfaces with real-time delivery of instructional material and feedback.

The application of the outcome space related to a specific construct map for a specific item is called a scoring guide. To make scoring guides more readily useable, instructors and students need concrete examples (called "exemplars") that show the rating

1. For the following geometric pattern, there is a chain of regular hexagons (meaning all 6 sides are equal):



- 1a.) Complete the table showing the number of hexagons in a chain and the perimeter (number of outside edges).

NUMBER OF HEXAGONS	PERIMETER
1	6
2	10
3	
4	
5	

- 1b.) Describe the process for determining the perimeter for 100 hexagons, without knowing the perimeter for 99 hexagons.

- 1c.) Write a formula to describe the perimeter for any number of hexagons in the chain (it does not need to be simplified).

- 1d.) Explain why you think your formula in (1c) is correct.

- 1e.) Suppose instead of regular hexagons, you form a chain with regular n -gons (where n is any number of sides). Describe the process for determining the perimeter for any number of regular n -gons in a chain.

- 1f.) Write a formula to describe the perimeter for any number of regular n -gons in a chain (it does not need to be simplified).

- 1g.) Explain why you think your formula in (1f) is correct.

Fig. 4 Hexagon Pattern task.

Connection Construct	GP-1 Exemplars
EA - Extended Abstract: Applies knowledge of functions for hexagons to a generalizable formula for n-gons.	Symbolic function for n-gon (1g) is correct and related to the symbolic function for hexagon (1c).
R - Relational: Relates information from hexagons to determine functions for n-gons.	Verbal function for n-gon (1e/1g) is correct and related to verbal function for hexagon (1b).
MS - Multistructural: More than two functions (tabular, verbal, symbolic) are clearly connected.	Table (1a), verbal description (1b), and equation (1c) are correct and connected.
US - Unistructural: Two functions are connected.	Table (1a) and equation (1c) is connected OR Table (1a) and verbal description of 100 hexagons is connected (1b) (Note: Verbal description must demonstrate knowledge beyond “adding four”).
PS - Prestructural: One function is correct.	Table (1a) is correct.
PA - Prealgebraic: One function is partially correct.	Table (1a) is partially correct (Note: Student may mention “adding four” as the only substance to the problem).
OT - Off Target	Nothing is correct. Makes no sense.
DK - Don’t Know	Writes “I don’t know”.
NR - No Response	Blank.

Fig. 5 Scoring guide for the Hexagon Pattern task (Note: read from bottom to top).

of specific samples of student work. These examples should illustrate what an instructor might expect from students at varying levels of development along each variable (i.e. at the different levels of the construct map). It turns out that this is also a critical element of how raters learn to understand the rationale of the construct maps and scoring guides. Fig. 5 shows a summary of the “scoring guide” for the Hexagon Pattern task (from Fig. 4), which shows how student responses are coded into the six construct map levels from Fig. 3, and hence would be scored as shown in Fig. 3. Note that in Fig. 5 there are a few more codes for responses that are only partial, and these would generally be scored as zero, alongside Prealgebraic. Beyond this, genuine samples of student work, scored and moderated by project members, are helpful as a resource for each activity—they illustrate typical responses for each score level, as well as atypical responses which can require careful consideration by a rater.

The measurement model (building block 4)

This is the technical end of the measurement development, and also the second part of interpretation (the third vertex of the NRC Assessment Triangle). The approach adopted here is to use the special version of item response modeling (IRT), known as the Rasch model (Rasch, 1980). In particular, the techniques pioneered by Adams et al. (1997a,b), which are broadly used in the Program for International Student Assessment (PISA) testing, are applied. These measurement models are well-developed enough for use in classroom-based assessments in a fairly routine and feasible way, with accompanying software that is available to be used by district- and school-based professionals and classroom teachers (Wilson et al., 2019). Quality control information is available as an output from these models. The output can be used to estimate where individual students are located on the Functions variable using a map, called a Wright Map, as shown in Fig. 6. The analyses for these Wright Maps were performed using the ConQuest software (Adams et al., 2020), which implements an expectation-maximization (EM) algorithm for the estimation of multi-dimensional Rasch-type models.

In Fig. 6, the Functions construct map levels are shown in the first column on the left-hand side. The second column from the left shows the units of the Functions scale (in logits). The third column from the left shows the estimated distribution of a set of students who responded to the Functions items in a data set gathered as part of the field trial of the test. The right-hand side of Wright Map is devoted to the items. There are eight items, named at the top of the first eight columns from the right.¹ Within each of those columns, the calibrated difficulty thresholds between the levels of the respective item are shown.

¹Since there is not enough room on the Wright map to show all 12 items, eight of the 12 items are represented on the right side, with their respective titles written across the top of the map. The Equivalent Functions item is similar to the missing four items on the map.

College Readiness Assessment – Wright Map			Distribution of item difficulty (n=8)							
Level on the Connections Construct Map	Estimate	Distribution of student proficiency (n=688)	HEXAGON PATTERN	EQUIVALENT FUNCTIONS	EDUCATIONAL ACCOMPLISHMENTS	GENDER GAP	COST OF POSTAGE	STAIRCASE TOOTHPICK	24 HOUR CRUDE OIL	100 YEARS CRUDE OIL
Extended Abstract (EA) Relational (R)	4						R		EA	EA
	3				EA	MS		MS US		
	2	X XXX		R	R				R	R
Multistructural (MS) Unistructural (US)	1	XXXXXX XXXXXXXXXX XXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXX -XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX	EA R		MS	US	MS		MS	MS
	0		MS US	MS	US	PS	US PS	PS PA		US
				US						
Prestructural (PS) Prealgebraic (PA)	-1	XXXXX XX XXX X X X	PS PA	PA PS	PS		PA		PS PA	
	-2				PA	PA				PS PA
EACH X REPRESENTS 4 STUDENTS, EACH ROW REPRESENTS .255 LOGITS										

Fig. 6 Wight map for the functions construct.

The Wright Map can be used for multiple purposes. The main idea is to use it to show how the items relate to the students. In this Figure, for example, one can note that the top levels (denoted by “EA”—see Figs. 3 and 5 for the meaning of these abbreviations) for two of the items, “24 Hour Crude Oil” and “100 Years Crude Oil,” are located above the highest student location. This means that it is very unusual for students to respond at those levels. In contrast, one can consider the lower end of the map, and see that it is very likely that some students responded at level PA or below. Most students have responded at about the US or MS levels.

The Wright Map is just one aspect of the functioning of the Partial Credit Model (Masters, 1982). It can also be used to control instructor rating consistency, and to maintain a meaningful measurement scale throughout the school year. This regularity in measurement makes the information from the items more useful for instructors in the classroom. Based on the Wright Map, advice on “next steps” in instruction can be generated by and for teachers. Many types of maps are available, derived from the analysis of student data collected in coursework. For example, Wright Maps can be used to track and assess student progress, and to exemplify the skills that a student has not yet mastered or is working on. Teachers can illustrate students’ progress with respect to the goals and expectations of the course by locating students’ performance on the scale in the map. The map can also be used to inform the planning of instruction. Wilmot (2008) interviewed a small sample of teachers to examine the usages that they made of the Wright Maps and associated reports of results for the Functions test. In her summary of her interviews, she noted that the maps helped teachers with the following activities: (a) checking student progress, (b) making instructional decisions, (c) seeing learning as a trajectory, and (d) planning curriculum across years. Further details and examples can be found in chapter 5 of her dissertation (Wilmot, 2008, pp. 126–131).

The BAS approach was originally envisaged in the context of classroom assessments (Kennedy, 2005; Wilson, 2018) and formative assessments embedded in curricula (Wilson and Sloane, 2000). Recent examples illustrating this can be found in the subject areas of reading (e.g. Dray et al., 2019), science (e.g. Junpeng et al., 2018; Metz et al., 2019; and for a Next Generation Science Standards [NGSS] application, Morell et al., 2017), mathematics (e.g. Arneson et al., 2019; Lehrer et al., 2014), and 21st century skills (e.g. Siddiq et al., 2017). BAS has also been applied in large-scale assessments in the topic areas of collaborative problem solving (Zhang et al., 2018), and information and communications technology (Wilson et al., 2018). All of these are in typical testing situations (both online and pencil-and-paper), but its largest application to date has been to an observational system for tracking child development in five early childhood and kindergarten learning domains (the Desired Results Developmental Profile [DRDP], see <https://www.desiredresults.us>), which is used in several US states, most conspicuously in California. Although initially proposed for cognitive assessments, BAS has also been found to be useful in the affective domain (e.g. Koo et al., 2021; Hermisson et al., 2019; Rocca et al., 2018), and in assessing teacher skills and behaviors (Duckor and Holmberg, 2017; Duckor et al., 2017).

Selected principled assessment design approaches

In this section, we give an account of selected PAD approaches, and briefly point out how they are consistent with and aligned to the NRC formulation in Fig. 1. This account is based, in part, on the earlier work of Ferrara et al. (2016).

Evidence-centered design

Evidence-Centered Design (ECD)² uses a logical and systematic approach based on evidentiary arguments to design, implement, and deliver educational assessments (Almond et al., 2015; Mislevy and Haertel, 2006; Mislevy and Riconscente, 2006; Mislevy et al., 2003b). Stemming from Messick (1989)’s views on validity—defined as “the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores” (p. 13)—ECD focuses on the process of reasoning using evidence from observations of what students say, do, or make in a particular testing situation, to make broader inferences about their knowledge, abilities, or other attributes. What this implies is that an assessment needs to begin around (a) the inferences to be made, (b) the observations that are needed to support those inferences, and (c) the chain of reasoning that links the observations to the inferences. This shows the clear alignment of ECD to the three-part NRC Assessment Triangle. In this way, the ECD framework “helps one understand the connections among an assessment’s purpose, a conception of proficiency in the domain, the evidentiary argument, the design of the assessment elements, and operational processes,” and “organizes thinking about domains, purposes, and delivery capabilities in a way that leads to coherent assessments” (Mislevy et al., 2003b).

Drawing on ideas from architecture and software engineering, ECD enhances the NRC Assessment Triangle model to encompass the entire assessment design and development process by viewing it in terms of five “layers” or clusters of activities: (a) domain analysis, (b) domain modeling, (c) conceptual assessment framework, (d) assessment implementation, and (e) assessment delivery. Generally, across the layers, ECD follows a chain of reasoning starting with an analysis of the domain of knowledge, skills, or attributes of interest. This is followed by the specification of the claims to be made based on the test, the decision on the evidence to support the claims, the development of tasks to provide the required evidence, the assembly of an appropriate set of tasks to support the claims with sufficient evidence, the establishment of rules for scoring and score aggregation to extract the necessary evidence, and

²This gives just a brief overview of ECD—for greater depth, the reader is directed to the article on “Evidence Centered Design” by Chengbin Yin and Bob Mislevy in this volume for more details.

a description of the explicit logic links across the entire process. While the layers may suggest a sequential design process, each layer may be executed and tweaked independently without affecting the rest of the layers, and yet, the different layers are closely inter-related. Consequently, as for BAS, the ECD process is iterative—problems that surface in a later layer may require us to return to and refine an earlier layer, or to make changes across layers.

As the first stage in the assessment design, the domain analysis layer collects background and fundamental information about the domain of interest, usually with the help of subject-matter experts. The analysis provides understanding of the valued knowledge and skills in the domain, and the situations that elicit the use of such knowledge and skills. The second domain modeling layer organizes the information gathered from the investigation in the domain analysis to focus on a selection of aspects of the domain, including claims, data, and warrants, in order to construct an assessment argument. This transitions to more precise specifications for the assessment system (e.g. item types, delivery system, scoring scheme, measurement model) in the third conceptual assessment framework layer. The conceptual assessment framework, as depicted in Fig. 7, is akin to the “blueprint” for the assessment. It comprises several models, such as the student (or proficiency) model, task model, evidence model, and assembly model, which are connected through key elements called student model variables, task model variables, observable variables, and work products. The first three models show the link to the NRC Assessment Triangle—the student model defines the variables related to the knowledge, skills, or attributes that the assessment is trying to measure; the task model provides a description of the task characteristics that would elicit the desired observable behavior or product; and the evidence model provides a bridge across the student and task models by describing the responses to the task that would provide evidence on the knowledge or skill of interest. The assembly model then describes what the test would look like as a whole, and makes the linkages among claims, evidence, and tasks explicit. The next layer is assessment implementation, which concerns the operational elements specified in the conceptual assessment framework layer, including item writing, scoring mechanisms, and measurement models. Finally, the assessment delivery layer provides the interface for task interaction, performance evaluation, and scoring. It comprises four principal processes: activity selection, task presentation, response processing, and summary scoring.

The use of evidentiary reasoning in ECD can be applied to a wide variety of assessments, from traditional paper-and-pencil assessments based on classical test theory, to more complex computer-based assessments that use advanced measurement models. Importantly, ECD does not replace traditional test development; instead, ECD augments traditional test development. ECD has been widely implemented in various assessment initiatives by SRI International, Cisco Systems, Educational Testing Service (ETS), and the College Board, as well as in statewide assessment programs (e.g. National Center and State Collaborative [NCSC], Partnership for Assessment of Readiness for College and Careers [PARCC], Smarter Balanced) under the Race to the Top initiative (Ferrara et al., 2016). The ECD approach has also been formulated as Evidence-Centered Game Design (ECgD; Mislevy et al., 2014) to focus on building principled assessments within gaming situations. More recently, the Next Generation Science Assessment (NGSA, see <http://nextgenscienceassessment.org/>) project has adopted ECD as the design framework to develop formative assessments aligned with the NGSS. An open-source version of the full ECD package has also been established—the Principled Assessment Designs for Inquiry (PADI, see <https://padi.sri.com/>).

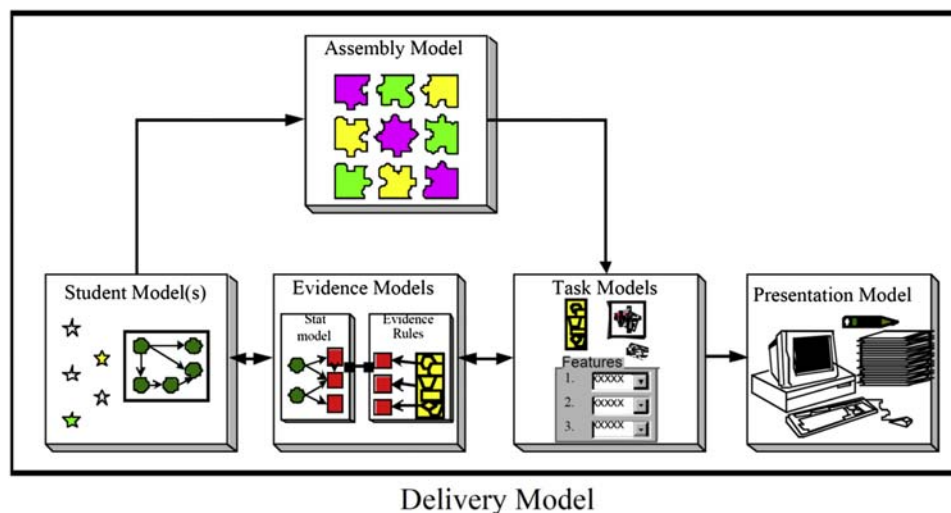


Fig. 7 Principal design objects of the Evidence-Centered Design (ECD) conceptual assessment framework. Note. These models are a bridge between the assessment argument and the operational activities of an assessment system. Looking to the assessment argument, they provide a formal framework for specifying the knowledge and skills to be measured, the conditions under which observations will be made, and the nature of the evidence that will be gathered to support the intended inference. Looking to the operational assessment, they describe the requirements for the processes in the assessment delivery system. Reproduced from Mislevy et al. (2003a) with permission from John Wiley and Sons.

Cognitive design system

The Cognitive Design System (CDS) is a framework that “centralize(s) the role of cognitive theory in [the development of] ability and achievement tests” (Embretson, 1998; Embretson and Gorin, 2001). According to Embretson (1994), there are several advantages in applying cognitive design to testing: (a) cognitive theory elaborates the constructs and processes that are involved in item solving, and thus supports construct validity of a test; (b) the correlations of test scores with other variables (or nomothetic span) can be explained by cognitive theory; (c) item design moves from an artistic process to a scientific process that reflects the cognitive theory influencing the processing requirements in the task; and (d) cognitive theory allows items with difficulties that are more precisely targeted to be developed, thereby increasing the efficiency of test development. Consequently, CDS has the following properties: “(1) test content is prescribed by explicit principles; (2) score meaning is linked to underlying cognitive processes, rather than just to correlations with other measures; (3) item parameters represent the sources of cognitive complexity in the item; and (4) abilities are linked to the processes that underlie task difficulty.”

CDS comprises a conceptual framework and a procedural framework. The conceptual framework distinguishes two empirically separable aspects of construct validity: (a) construct representation, and (b) nomothetic span. Construct representation “concerns the meaning of test scores,” and “is elaborated by understanding the processes, strategies, and knowledge [structures] that persons use to solve items” (Embretson, 1998). It is supported by research from cognitive psychology. For example, empirical data from modeling item difficulty in terms of accuracy and response time can be used to identify elements of the stimuli that can be manipulated to vary the cognitive demands in the assessment task. On the other hand, nomothetic span “concerns the utility [or significance] of the test for measuring individual differences,” and “is elaborated by the correlations of test scores with other measures” (Embretson, 1998). Importantly, the correlations are a consequence of construct representation. In other words, test meaning precedes test significance. This is in contrast with the traditional conceptualization of construct validity where test meaning is elaborated by the external correlations (Cronbach and Meehl, 1955). Since meaning is established after the test is developed, the correlations provide little information about how to improve the test design.

The CDS procedural framework, as outlined in Table 2, specifies a sequence of developmental stages that integrates cognitive theory into the test design. While the stages are presented sequentially, the actual test development process is iterative, that is, if problems are surfaced in a particular stage, one may need to re-visit and refine the earlier stages. Embretson and Gorin (2001) described seven key procedural steps for CDS: (a) specify the goals of measurement in terms of construct representation and nomothetic span; (b) develop a cognitive model that consolidates the relevant processes, strategies, and knowledge structures; (c) identify the task features that can be manipulated to vary construct representation by influencing the cognitive processes, strategies, and knowledge structures required for item-solving, using research from cognitive psychology based on research and theory for the particular item type, operationalization of the item features that affect item solving processes, and empirical evidence of the impact of those features on psychometric properties (item difficulty and discrimination); (d) generate items by manipulating the task features to vary the cognitive complexity of the item-solving processes; (e) evaluate both the cognitive model (by using task features to predict performance in terms of response time and item difficulty) and psychometric model (by examining fit to the item response data) of the item generating system to establish construct representation and validity; (f) bank items by their sources of cognitive complexity; and (g) validate that the cognitive model provides a basis for specific hypotheses about the nomothetic span of the test. Referring back to the NRC Assessment Triangle, we can readily see the emphasis on the Cognition vertex in this approach; in fact, one can draw analogies between steps (a) and (b) above and Cognition, steps (c) and (d) and Observation, and steps (e) through (g) and Interpretation.

The CDS approach has been used for the design and development of mental rotation items and object assembly items (measuring spatial ability), progressive matrix items and non-verbal reasoning items (measuring abstract reasoning), and mathematical word problems (Embretson and Gorin, 2001, as cited in Ferrara et al., 2016). Examples include the Spatial Learning Ability Test (SLAT) developed by Embretson (1994), the Abstract Reasoning Test (ART) of the Armed Services Vocational Aptitude Battery (Embretson, 1998), and more recently, a new form for a test of fluid intelligence (ART-E1) used for the selection of law enforcement personnel (Embretson, 2016). The latest application of CDS is in the development of a computer-adaptive test for computational thinking that was administered to over 90,000 primary school students in the Netherlands (Hogenboom et al., 2021), which involved the creation of close to 4500 test items in an online game environment, and the validation of the cognitive model for the assessment.

Assessment engineering

Assessment Engineering (AE) is a “highly structured and formalized manufacturing-engineering process” to efficiently design and implement coherent cognitively based assessments (Luecht, 2013). According to Leucht (2012), the AE framework has several operational goals: “(1) to design and consistently measure the same intended construct(s) over time and conditions of measurement—for formative or summative purposes; (2) to fully or partially automate tedious, costly or time-consuming, human-dominant procedures; (3) to provide low-cost, scalable and efficient methods of producing a large number of items that may not have to be individually pretested, calibrated and then discarded after operational use (or re-purposed); and (4) to relax some of the pressure on data-hungry psychometric calibration and scaling models to do that “heavy-lifting” in establishing and maintaining one or more scale scores that measure the intended construct(s).”

Table 2 Procedural framework of the Cognitive Design System (CDS).

<i>Stage</i>
Specify general goals of measurement
Construct representation (meaning)
Nomothetic span (significance)
Identify design features in task domain
Task-general features (mode, format, conditions)
Task-specific features
Develop a cognitive model
Review theories
Select or develop model for psychometric domain
Revise model
Test model
Evaluate cognitive model for psychometric potential
Evaluate cognitive model plausibility on current test
Evaluate impact of complexity factors on psychometric properties
Anticipate properties of new test
Specify item distributions on cognitive complexity
Distribution of item complexity parameters
Distribution of item features
Generate items to fit specifications
Artificial intelligence?
Evaluate cognitive and psychometric properties for revised test domain
Estimate component latent trait model parameters
Evaluate plausibility of complexity factors on psychometric properties
Evaluate plausibility of the psychometric model
Calibrate final item parameters and ability distributions
Psychometric evaluation
Measuring processing abilities
Bank items by cognitive processing demands
Assemble test forms to represent specifications
Fixed content test
Adaptive test
Validation: strong program of hypothesis testing

Reproduced from Embretson (1998) with permission from the American Psychological Association.

To achieve those goals, AE makes “three fundamental assertions” (Leucht, 2012): (a) task content and complexity should increase across a proficiency or score scale to reflect the increasing level of knowledge and skill requirements along the scale; (b) a class or family of items can be designed from a model of task complexity that includes specification of the declarative knowledge (static representations of knowledge in memory) and procedural knowledge (cognitive skills), available tools, and overall complexity of the context or task setting required for items in the family; and (c) a large number of items with the same task complexity and statistical or psychometric properties can be generated within the same family using task templates and other quality control mechanisms.

There are four key stages in the AE process (Leucht, 2012, 2013): (a) construct mapping and evidence modeling; (b) task modeling; (c) template design and item writing; and (d) item calibration and quality control. These four stages can be seen as being analogous to the four building blocks in BAS (though each with its particular emphases), and hence share the same relationships with the NRC Assessment Triangle. Like the other principled approaches, AE begins with the fundamental design process of developing a construct map, which articulates the ordered proficiency-related claims on respondents’ knowledge and skills (akin to achievement- or performance-level descriptors) along an intended score scale. An evidence model is also developed alongside the construct map to describe the expected performance at each level of the scale, and this could include exemplars of work products or other performance-based evidence that support the knowledge and skills required by each proficiency claim.

In the next stage, a task model is created to provide detailed specifications for a class or family of items that share the same cognitive demands as well as other task features such as content, context, and auxiliary resources or tools. These specifications can be represented in a task model using a task model grammar, which allows items generated for the same family to be isomorphic in terms of cognitive complexity. With the specifications, each task model can be located on a specific region of the measurement scale by relating the task model’s cognitive complexity to the proficiency claims and associated evidence models. Each task model also provides the basis and efficiency for the building of assessment task templates, and the subsequent generation of items from each template. A task model map, as illustrated in Fig. 8, can then be created by arranging the task models along the scale to provide a visualization of the distribution of task models at different locations. The task model map is a formal design specification for the intended scale, and we would expect locations that require greater precision or richer interpretations to have a higher density of task models.

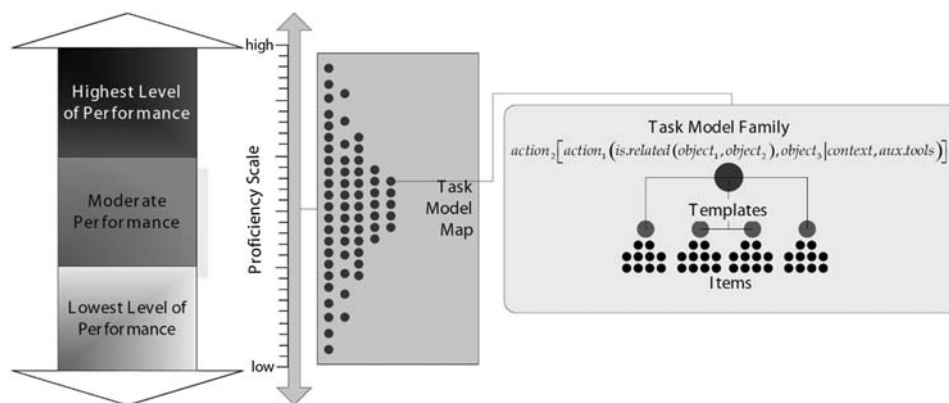


Fig. 8 Linking templates and items to the task model map and construct map in the Assessment Engineering (AE). Reproduced from [Luecht \(2012\)](#) with permission from Taylor & Francis Group LLC.

During the next stage of designing templates and item writing, the task model is implemented through template and item development. In particular, templates are created to maintain the interpretation of the knowledge and skills specified by a particular task model, and items are created as part of a family based on a particular template. Given that each task model has an implicit or ordered location, the corresponding templates and items would also need to adhere to that location or ordering. To facilitate that, each template includes three components that specify the presentation format, manipulable content, and scoring rules, namely, (a) the rendering model (also called item shell or item model) which sets the item type format and look-and-feel of the item; (b) the scoring evaluator which determines which response or other data is collected and used; and (c) the data model which contains all relevant data for presenting and scoring the item. The templates also form part of the quality control system, where quality assurance would require that items in the same template family behave similarly in terms of their psychometric and other characteristics. This leads to the final stage where the assessment tasks are field-tested and calibrated to validate the hypothesized complexity-proficiency ordering of items, templates, and task models.

AE is particularly helpful to the design and development of cognitively based assessments for both formative and summative uses. It also provides an efficient and cost-effective way to generate large numbers of isomorphic items for an item bank. A study by [Gierl et al. \(2008\)](#), for instance, has demonstrated that over 330,000 unique items could be generated from 10 mathematics item templates using an automatic item generator software.³ [Luecht \(2012\)](#) has further suggested that AE may be most beneficial for formative and diagnostic assessments, given that educators need to have a test which can produce profiles of student performance on a regular basis as part of the learning process. One related application is the College Board's Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT/NMSQ), where the AE approach is used to generate and evaluate over 10,300 items based on 15 item templates ([Lai et al., 2010](#), as cited in [Ferrara et al., 2016](#)).

Principled design for efficacy

The Principled Design for Efficacy (PDE) is a "practical way to implement principled assessment design and development" by "helping test designers to make design decisions thoughtfully, explicitly, and systematically" ([Nichols et al., 2015](#)). Notably, PDE shares some similarities with other principled approaches, particularly ECD, such as being construct-centered and the documentation of validity arguments, and hence, it also shares ECD's relationships with the NRC Assessment Triangle. However, some features are unique to PDE. For example, PDE aims to make the assessment development process more efficient and effective by: (a) using knowledge, skills, and abilities (KSA) research from the human resources domain to inform test design and development; (b) emphasizing communication with various stakeholders; (c) working within internal and external constraints; and (d) considering the intended interpretations and uses of the test scores. Note that there has been some controversy about using KSAs in this way—see [Kay \(2011\)](#).

The PDE approach is driven by a conceptual framework with six inter-related design concepts (as illustrated in [Fig. 9](#)): (a) construct, (b) consequences, (c) content, (d) evidence, (e) communication, and (f) implementation. Not only does the conceptual framework guide the test development process, but it also provides a common language to facilitate reasoning and communication among different stakeholders of the assessment program. First, the concept of construct is associated with the question "What are you assessing?", which is answered by a description of the KSAs that inferences will be made about test-takers, and a validity framework documenting the chain of reasoning that links the intended interpretation and use of the test results to the test scores. Second, the concept of consequences is associated with the questions "What do you expect to happen when you assess?" and "What are the general processes through which the identified consequences are expected to occur?" and can be addressed by

³The reader is directed to the "Automatic Item Generation" article by Mark Gierl in this volume.

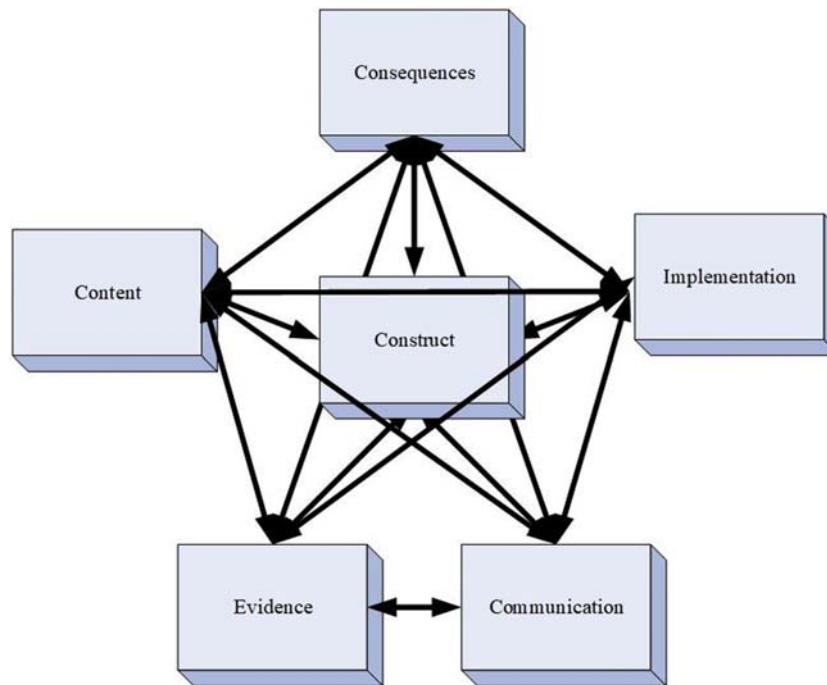


Fig. 9 Conceptual framework for the Principled Design for Efficacy (PDE). Reproduced from [Nichols \(2021\)](#) with permission from Planful Learning and Assessment LLC.

a theory of action that describes the process of change or mechanisms assumed to lead to the intended outcomes of the assessment. Third, the concept of content is associated with the question “How do you assess it?” and is concerned with how task content features can be manipulated to elicit the target KSAs, and to vary the task demand and complexity. Fourth, the concept of evidence is associated with the question “How do you know what you are assessing when you see it?”, which corresponds to specific instructions for interpreting test-takers’ responses, comprising work product specifications (what performances will be captured), evidence rules (how performances will be coded), and the statistical model (how coded responses will be aggregated to make inferences). Fifth, the concept of communication is associated with the question “How do you talk to people about what you are assessing?” and focuses on effective communication with different stakeholders (sponsors, content developers, test-takers, information users), who may bring different values and perspectives to the assessment program. Finally, the concept of implementation is associated with the question “How do you want to carry out the assessment program?”—this is about the thoughtful consideration of the assessment goals and purpose, needs of stakeholders and use of test results, test requirements and constraints, and context of administration in the design, development, and delivery of the assessment.

PDE is operationalized through four key stages of design and development activities: (a) construct definition, (b) content creation, (c) feature generalization, and (d) content re-creation. During construct definition, the assessment designer uses the research literature to define the content standards in terms of cognitive processes, knowledge structures, strategies, and mental models, and to articulate the general content features of stimuli and items that will elicit the explicated KSAs. This is followed by content creation where stimuli, items, and rubrics are developed from those general content features to capture the full range of performance in the target KSAs. Next, in feature generalization, more detailed guidelines and specifications are written for those stimuli and items created. Lastly, content developers use those guidelines and specifications to efficiently generate more stimuli and items during content re-creation. It should also be noted that the design concepts cut across the various stages of design and development. For example, the concept of communication is addressed throughout all stages of activities to achieve a shared understanding of the assessment (e.g. assessment goals, KSAs, content features) among stakeholders.

As seen above, PDE focuses specifically on the development of the test content instead of the entire assessment development process, and thus benefits from the application of learning sciences research. Further, the element of implementing the design and development process within practical constraints means that PDE can be used for existing assessments or new assessments that already have certain design decisions established. These two features support the intent of PDE as an enhancement rather than a replacement for current assessment practices ([Ferrara et al., 2016](#)). Thus, the PDE approach has mainly been used for educational achievement tests, especially in state-wide assessment programs such as the elementary and middle school formative assessment program for the Baltimore County Public Schools, and the NGSS assessment program in Maryland ([Ferrara et al., 2016](#)). Notwithstanding that, PDE has also been applied to performance tests, particularly embedded assessments in simulation-based and learning games—for example, the Insight Learning System, which is an integrated learning system for the teaching and

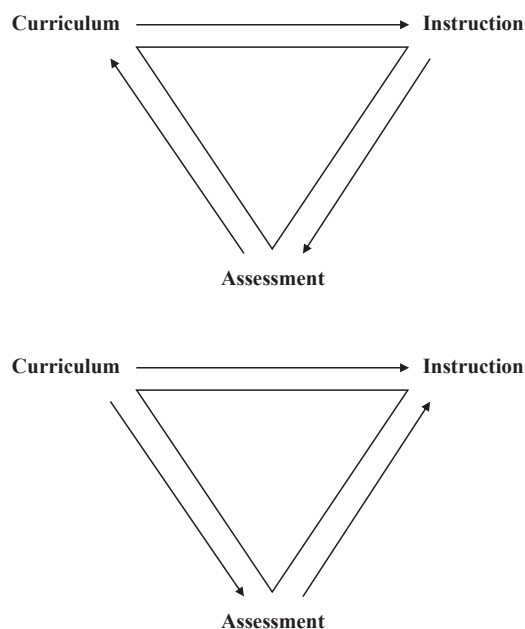


Fig. 10 The CIA triangle (top) and Black’s “vicious triangle” (bottom).

assessment of geometry (length and area) for Grades K-4 students; and the Common Core Speaking and Listening (CCSL) formative assessment system, which incorporates learning and feedback through simulation of real presentational speaking and two-way communication (DiCerbo et al., 2017).

Going “outside the triangle”

There is no more hackneyed phrase these days than “Going outside the box,” but indeed that is what we need to consider in surveying the relations of PAD approaches to the rest of the educational enterprise (though we have a triangle rather than a box!). We can see assessment as playing a key role in education, as is represented in the traditional “CIA Triangle”—Curriculum, Instruction, and Assessment—as shown in the top part of Fig. 10. The conceptualization in this diagram is that the aims of the educational activity are specified in the curriculum, then the method of delivery is embodied in the instruction (or, for those in the UK, the pedagogy), and then the evaluation of the activity is carried out via the assessments. However, Paul Black, from King’s College, London, has suggested a modification of this figure to illustrate what he sees as the deleterious effects of assessment on education. This, which he labels as the “vicious triangle” (Black et al., 2011), is illustrated in the bottom part of Fig. 10.

In this situation, the teachers’ instructional plans and efforts are “squeezed” between the overarching goals of the curriculum, and the concentrating and winnowing effects of external (often large-scale) tests which have been designed to hold teachers and schools to certain accountability targets. According to Wilson (2018):

The central issue is that the assessments (both classroom and large-scale) need to be working coherently with the curriculum. When that coherence breaks down, or as is sometimes the case, it was never there in the first place, the sorts of negative outcomes mentioned in the previous section can be expected to occur (pp. 7–8).

Black and his colleagues have suggested that what we require to avoid this dilemma is to build curriculum, instruction, and assessment on unifying theories of learning and teaching that assemble these three elements into a coherent educational process (Black et al., 2011). This is illustrated in the “Revised CIA Triangle” shown in Fig. 11, which symbolizes both the presence of additional forms of feedback among the three parts of the CIA Triangle, as well as the need for principles of coherence in the shape of theories of learning.

The reason that this account is relevant to the PAD approaches is that the part of this diagram related to assessment is very much what is the focus of PAD itself. That is, we must create structures for understanding and expressing the architecture of *both* curriculum and assessments together so that (a) the curriculum can be used to define the goals of the assessment (i.e. constructs to be assessed), and (b) the results of the assessments can be interpreted directly in terms of those curriculum constructs. What the Revised CIA Triangle adds in here is the element of instruction, which, of course, is the motor of the educational enterprise.

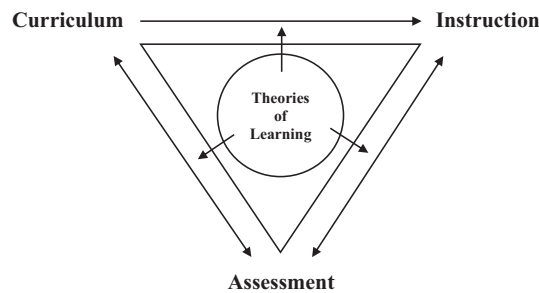


Fig. 11 The revised CIA triangle.

Thinking along these lines leads one to consider a theory of action concerning the implementation of the PAD approaches. This is ground that has already been trodden by another researcher and his colleagues (Bennett, 2010)—in this ground-breaking effort, researchers at ETS have endorsed a PAD-like approach but then gone beyond it (i.e. “outside the Triangle”), and considered not just the usual usages of assessment *of* learning, but also assessment *for*, and *as* learning—the Cognitively Based Assessment of, for, and as Learning or CBAL (see <https://www.ets.org/cbal>). More specifically, they are working to develop assessments in which (a) the activities of learning integrate assessments and their products (student outcomes and ways to interpret them), and (b) the activity of taking the assessments can itself act as a learning activity. The latter, of course, breaks with one of the usual assumptions about measurement, that is, measurement is based on indicators of student abilities, but the process of measurement does not, in itself, bring about changes in those same student abilities—modeling this phenomenon is a considerable challenge for modern psychometrics. See, for example, Mari et al. (2021), where this is discussed under the heading of “measurement that involves feedback” (pp. 227–228). The CBAL effort is an important initiative to take the advantages of the PAD approach, as discussed above, into the real world of testing.

Another related formulation for going outside the Triangle is designed to link into the current expansion of web-based learning technologies: *Expanded Evidence-Centered Design* (e-ECD; Arieli-Attali et al., 2019) extends the ECD framework to the design and development of technology-enhanced learning and assessment systems by adding specifications for describing aspects of learning associated with each of the three core models in the original ECD formulation. This enables a close detailing of the linkage and coordination between learning and assessment, through parallel layers or models, in the design framework. Importantly, this framework is written to a generic version of a learning theory, so that it can be used with a variety of learning systems.

Conclusion

One of the common complaints among the society of those who make their professional life in and around the creation and usage of assessments in education is that many lay-persons, and even professionals in education, such as teachers, do not see anything in an assessment instrument beyond the explicit set of items that constitutes the “test.” This is in marked contrast to the common thread among all of the PAD approaches described above—a determination that (a) the target variable underlying the assessment should be well-specified, and (b) there is a strong argument for the validity of inference from the assessment outcomes to that target variable. The items, though necessary of course, are seen in this light as ephemera—what was needed to provoke the interesting information.

Considering this, one of us has argued in the past (perhaps somewhat tersely) that “The instrument [i.e. the test], then, is seen as a logical argument that the results can be interpreted to help make a decision as the measurer intended them to be” (Wilson, 2004). This embodiment of the PAD argument in the form of the instrument acknowledges the crucial importance of the linkage between (a) and (b) in the previous paragraph that is supplied by the instrument itself. In this view, the test must be seen as being like an iceberg, with hidden (to the examinee) underpinnings in the shape of the relevant PAD components—for example, for BAS, the hidden parts are the construct map, the items design, the outcome space, the measurement model, the validity and reliability evidence, and the interpretative materials. One way to see the efforts we have described above to come up with PAD approaches is as a struggle to both understand what the basis is for using an assessment, but also to find a way to communicate this basis to the many types of assessment users, from students and teachers to policy-makers. For example, although professional standards (such as the “Testing Standards”) have an important role in maintaining quality assessment, it is only through the broad appreciation of what makes for a good assessment among the stakeholders (i.e. from students and teachers to policymakers) that there will be a continuing stream of support for the costly business of developing and utilizing sound assessments.

References

- Adams, R.J., Wilson, M., Wang, W., 1997a. The multidimensional random coefficients multinomial logit model. *Appl. Psychol. Meas.* 21 (1), 1–23. <https://doi.org/10.1177/0146621697211001>.
- Adams, R.J., Wilson, M., Wu, M., 1997b. Multilevel item response models: an approach to errors in variables regression. *J. Educ. Behav. Stat.* 22 (1), 47–76. <https://doi.org/10.3102/10769986022001047>.
- Adams, R.J., Wu, M.L., Cloney, D., Wilson, M.R., 2020. ACER ConQuest: Generalised Item Response Modeling Software [Computer software]. Version 5. Australian Council for Educational Research (ACER), Camberwell, Australia.
- Almond, R.G., Mislevy, R.J., Steinberg, L.S., Yan, D., Williamson, D.M., 2015. *Statistics for Social and Behavioral Sciences: Bayesian Networks in Educational Assessment*. Springer, New York, NY. <https://doi.org/10.1007/978-1-4939-2125-6>, pp. 19–40, 411–505.
- American Educational Research Association, American Psychological Association, National Council for Measurement in Education, 2014. *Standards for Educational and Psychological Testing*. American Educational Research Association, Washington, DC.
- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., von Davier, A.A., 2019. The expanded evidence-centered design (e-ECD) for learning and assessment systems: a framework for incorporating learning goals and processes within assessment design. *Front. Psychol.* 10. <https://doi.org/10.3389/fpsyg.2019.00853>.
- Arneson, A., Wihardini, D., Wilson, M., 2019. Assessing college-ready data-based reasoning. In: Bostic, J., Krupa, E., Shih, J. (Eds.), *Quantitative Measures of Mathematical Knowledge: Researching Instruments and Perspectives*. Routledge, New York, NY. <https://doi.org/10.4324/9780429486197>.
- Bennett, R., 2010. Cognitively based assessment of, for, and as learning (CBAL): a preliminary theory of action for summative and formative assessment. *Measurement* 8 (2–3), 70–91. <https://doi.org/10.1080/15366367.2010.508686>.
- Biggs, J.B., Collis, K.F., 1982. *Evaluating the Quality of Learning: The SOLO Taxonomy*. Academic Press, New York, NY. <https://doi.org/10.1016/C2013-0-10375-3>.
- Black, P., Wilson, M., Yao, S., 2011. Road maps for learning: a guide to the navigation of learning progressions. *Measurement* 9, 71–123.
- Bloom, B.S., 1981. *All our Children Learning: A Primer for Parents, Teachers, and Other Educators*. McGraw-Hill, New York, NY.
- Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. *Psychol. Bull.* 52, 281–302. <https://doi.org/10.1037/h0040957>.
- DiCerno, K., Ferrara, S., Lai, E., 2017. Principled design and development for embedding assessment for learning in games and simulations. In: Jiao, H., Lissitz, R.W. (Eds.), *Technology Enhanced Innovative Assessment: Development, Modeling, and Scoring from an Interdisciplinary Perspective*. Information Age Publishing, Charlotte, NC, pp. 163–196.
- Dray, A.J., Brown, N.J.S., Diakow, R., Lee, Y., Wilson, M., 2019. A construct modeling approach to the assessment of reading comprehension for adolescent readers. *Read. Psychol.* 40 (2), 191–241. <https://doi.org/10.1080/02702711.2019.1614125>.
- Duckor, B., Holmberg, C., 2017. *Mastering Formative Assessment Moves: 7 High-Leverage Practices to Advance Student Learning*. ASCD, Washington, DC.
- Duckor, B., Draney, K., Wilson, M., 2017. Assessing assessment literacy: an item response modeling approach for teacher educators. *Pensamiento Educ.* 54 (2), 1–26.
- Embretson, S.E., Gorin, J., 2001. Improving construct validity with cognitive psychology principles. *J. Educ. Meas.* 38 (4), 343–368. <https://doi.org/10.1111/j.1745-3984.2001.tb01131.x>.
- Embretson, S., 1994. Applications of cognitive design systems to test development. In: Reynolds, C.R. (Ed.), *Perspectives on Individual Differences: Cognitive Assessment. A Multidisciplinary Perspective*. Plenum Press, New York, NY, pp. 107–135. https://doi.org/10.1007/978-1-4757-9730-5_6.
- Embretson, S., 1998. A cognitive design system approach to generating valid tests: application to abstract reasoning. *Psychol. Methods* 3, 300–396. <https://doi.org/10.1037/1082-989X.3.3.380>.
- Embretson, S., 2016. An integrative framework for construct validity. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, New York, NY, pp. 102–123. <https://doi.org/10.1002/9781118956588.ch5>.
- Ferrara, S., Lai, E., Reilly, A., Nichols, P.D., 2016. Principled approaches to assessment design, development, and implementation. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, New York, NY, pp. 41–74. <https://doi.org/10.1002/9781118956588.ch3>.
- Fisher, W.P., Wilson, M., 2019. An online platform for sociocognitive metrology: the BEAR assessment system software. *Meas. Sci. Technol.* 31 (3). <https://doi.org/10.1088/1361-6501/ab5397>.
- Gierl, M.J., Zhou, J., Alves, C., 2008. Developing a taxonomy of item model types to promote assessment engineering. *J. Technol. Learn. Assess.* 7 (2). Retrieved from: <https://ejournals.bc.edu/index.php/jtla/article/view/1629>.
- Glaser, R., 1981. The future of testing: a research agenda for cognitive psychology and psychometrics. *Am. Psychol.* 36, 923–936. <https://doi.org/10.1037/0003-066X.36.9.923>.
- Hermisson, S., Gochyyev, P., Wilson, M., 2019. Assessing pupils' attitudes towards religious and worldview diversity—development and validation of a nuanced measurement instrument. *Br. J. Relig. Educ.* 41, 371–387. <https://doi.org/10.1080/01416200.2018.1556604>.
- Hogenboom, S.A.M., Hermans, F.F.J., Van der Maas, H.L.J., 2021. Computerized adaptive assessment of understanding of programming concepts in primary school children. *Comput. Sci. Educ.* <https://doi.org/10.1080/08993408.2021.1914461>.
- Junpeng, P., Inprasitha, P., Wilson, M., 2018. Modeling of the open-ended items for assessing multiple proficiencies in mathematical problem solving. *Turkish Online J. Educ. Technol.* 2, 142–149. Retrieved from: http://www.tojet.net/special/2018_12_3.pdf.
- Kay, J., 2011. No more KSAs? How the federal job application process has changed. FedSmith.com. Retrieved from: <https://www.fedsmith.com/2011/03/21/no-more-ksas-how-federal-job>.
- Kennedy, K.A., 2005. *The BEAR Assessment System: A Brief Summary for the Classroom Context*. Technical Report No. 2005-03-01. Retrieved from: <https://bearcenter.berkeley.edu/bibliography/bear-assessment-system-brief-summary-classroom-context>.
- Koo, B.W., Bathia, S., Morell, L., Gochyyev, P., Phillips, M., Wilson, M., Smith, R., 2021. Examining the effects of a peer-learning research community on the development of students' researcher identity, confidence, and STEM interest & engagement. *J. STEM Outreach* 4 (1).
- Lai, H., Gierl, M., Alves, C., 2010. Generating Items Under the Assessment Engineering Framework [Invited Symposium: An Application of Assessment Engineering to Multidimensional Diagnostic Testing in an Education Setting], 2010 Annual Meeting of the National Council on Measurement in Education (NCME), Denver, CO, United States.
- Lehrer, R., Kim, M.-J., Ayers, E., Wilson, M., 2014. Toward establishing a learning progression to support the development of statistical reasoning. In: Maloney, A., Confrey, J., Nguyen, K. (Eds.), *Learning Over Time: Learning Trajectories in Mathematics Education*. Information Age Publishing, Charlotte, NC, pp. 31–60.
- Luecht, R.M., 2012. An introduction to assessment engineering for automatic item generation. In: Gierl, M., Haladyna, T. (Eds.), *Automatic Item Generation*. Routledge, Taylor-Francis, New York, NY, pp. 59–101. <https://doi.org/10.4324/9780203803912>.
- Luecht, R.M., 2013. Assessment engineering task model maps, task models and templates as a new way to develop and implement test specifications. *J. Appl. Test. Technol.* 14 (Special Issue 1). Retrieved from: <http://www.jattjournal.net/index.php/atp/article/view/45254>.
- Mari, L., Wilson, M., Maul, A., 2021. *Measurement Across the Sciences: Developing a Shared Concept System for Measurement*. Springer, New York, NY. <https://doi.org/10.1007/978-3-030-65558-7>.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174. <https://doi.org/10.1007/BF02296272>.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. Macmillan Publishing; American Council on Education, New York, NY, pp. 13–103.
- Metz, K.E., Cardace, A., Berson, E., Ly, U., Wong, N., Sisk-Hilton, S., Metz, S.E., Wilson, M., 2019. Primary grade children's capacity to understand microevolution: the power of leveraging their fruitful intuitions and engagement in scientific practices. *J. Learn. Sci.* 28 (4–5), 556–615. <https://doi.org/10.1080/10508406.2019.1667806>.

- Mislevy, R.J., Haertel, G.D., 2006. Implications of evidence-centered design for educational testing. *Educ. Meas.* 25 (4), 6–20. <https://doi.org/10.1111/j.1745-3992.2006.00075.x>.
- Mislevy, R.J., Riconscente, M.M., 2006. Evidence-centered assessment design: layers, concepts, and terminology. In: Downing, S., Haladyna, T. (Eds.), *Handbook of Test Development*. Erlbaum, Mahwah, NJ, pp. 61–90.
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003a. A Brief Introduction to Evidence-Centered Design (CSE Technical Report No. 632). The National Center for Research on Evaluation, Standards, Student Testing (CREST), Los Angeles, CA. <https://doi.org/10.1002/j.2333-8504.2003.tb01908.x>.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., 2003b. On the structure of educational assessments. *Measurement* 1, 3–67. https://doi.org/10.1207/S15366359MEA0101_02.
- Mislevy, R.J., Corrigan, S., Oranje, A., Dicerbo, K., John, M., Bauer, M.I., Hoffman, E., von Davier, A.A., Hao, J., 2014. Psychometric Considerations in Game-Based Assessment. GlassLab, Redwood City, CA. Retrieved from: https://www.ets.org/research/policy_research_reports/publications/white_paper/2014/jrx.
- Mislevy, R., 1996. Test theory reconceived. *J. Educ. Meas.* 33 (4), 379–416. <https://doi.org/10.1111/j.1745-3984.1996.tb00498.x>.
- Morell, L., Collier, T., Black, P., Wilson, M., 2017. A construct-modeling approach to develop a learning progression of how students understand the structure of matter. *J. Res. Sci. Teach.* 54 (8), 1024–1048. <https://doi.org/10.1002/tea.21397>.
- National Research Council (NRC), 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/10019>.
- Nichols, P.D., Ferrara, S., Lai, E., 2015. Principled design for efficacy: design and development for the next generation of assessments. In: Jiao, H., Lissitz, R. (Eds.), *The Next Generation of Testing: Common Core Standards, Smarter Balanced, PARCC, and the Nationwide Testing Movement*. Information Age Publishing, Charlotte, NC, pp. 49–81.
- Nichols, P.D., Kobrin, J.L., Lai, E., Koepfler, J.D., 2016. The role of theories of learning and cognition in assessment design and development. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, New York, NY, pp. 15–40. <https://doi.org/10.1002/9781118956588.ch2>.
- Nichols, P., 2021. Why is Principled Assessment Design not Sufficient to Build Successful Assessment Products? [Blog] Planful Learning and Assessment LLC. Available at: <https://planfullearning.com/2021/11/01/why-is-principled-assessment-design-not-sufficient-to-build-successful-assessment-products>. (Accessed 1 November 2021).
- Rasch, G., 1980. *Probabilistic Models for Some Intelligence and Attainment Tests*. University of Chicago Press, Chicago, IL (Original work published in 1960).
- Rocca, C.H., Ralph, L.J., Wilson, M., Gould, H., Foster, D.G., 2018. Psychometric evaluation of an instrument to measure prospective pregnancy preferences: the desire to avoid pregnancy scale. *Med. Care* 57 (2), 152–158. <https://doi.org/10.1097/MLR.0000000000001048>.
- Schwartz, R., Ayers, E., Wilson, M., 2017. Mapping a learning progression using unidimensional and multidimensional item response models. *J. Appl. Meas.* 18 (3), 268–298.
- Siddiq, F., Gochyyev, P., Wilson, M., 2017. Learning in digital networks—ICT literacy: a novel assessment of students' 21st century skills. *Comput. Educ.* 109, 11–37. <https://doi.org/10.1016/j.compedu.2017.01.014>.
- Snow, R.E., Lohman, D.F., 1989. Implications of cognitive psychology for educational measurement. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. Macmillan Publishing; American Council on Education, New York, NY, pp. 263–331.
- Wilmot, D.B., Schoenfeld, A., Wilson, M., Champney, D., Zahner, W., 2011. Validating a learning progression in mathematical functions for college readiness. *Math. Think. Learn.* 13 (4), 259–291. <https://doi.org/10.1080/10986065.2011.608344>.
- Wilmot, D.B., 2008. *Assessing Progress Toward College Readiness with Psychometric and Cognitive Models of Student Learning in Mathematics*. Doctoral dissertation. University of California, Berkeley.
- Wilson, M., Sloane, K., 2000. From principles to practice: an embedded assessment system. *Appl. Meas. Educ.* 13 (2), 181–208. https://doi.org/10.1207/S15324818AME1302_4.
- Wilson, M., Scalise, K., Gochyyev, P., 2018. Learning in digital networks as a modern approach to ICT literacy. In: Griffin, P., Care, E., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills: Volume 3—Research and Applications*. Springer, Dordrecht, Germany, pp. 181–210. https://doi.org/10.1007/978-3-319-65368-6_11.
- Wilson, M., Scalise, K., Gochyyev, P., 2019. Domain modeling for advanced learning environments: the BEAR assessment system software. *Educ. Psychol.* 39 (10), 1199–1217. <https://doi.org/10.1080/01443410.2018.1481934>.
- Wilson, M., 2004. *Constructing Measures: An Item Response Modeling Approach*. Erlbaum, Mahwah, NJ. <https://doi.org/10.4324/9781410611697>.
- Wilson, M., 2009. Measuring progressions: assessment structures underlying a learning progression. *J. Res. Sci. Teach.* 46 (6), 716–730. <https://doi.org/10.1002/tea.20318>.
- Wilson, M., 2018. Making measurement important for education: the crucial role of classroom assessment. *Educ. Meas.* 37 (1), 5–20. <https://doi.org/10.1111/emip.12188>.
- Zhang, Z., Wilson, M., Alom, M., Awwal, N., Griffin, P., 2018. Adopting a Process Perspective of Collaborative Problem Solving [Paper presentation]. 2018 Annual Meeting of the National Council on Measurement in Education (NCME), New York, NY, United States.

Evidence-centered design

Chengbin Yin and Robert J. Mislevy, University of Maryland, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	163
Assessment as evidentiary argument	163
Test specifications	164
Assessment engineering	164
Layers in ECD	164
Domain analysis	165
Domain modeling	165
Conceptual assessment framework	167
The implementation and delivery layers	167
Applications and extensions	167
Future directions	168
References	169
Further reading	169

Introduction

Grounded in the concept of assessment as evidentiary argument, Evidence-centered Design (ECD) is a comprehensive framework for articulating the conceptual, inferential, and computational elements of educational assessment and explicating the interrelationships among substantive arguments, assessment designs, and operational processes. Within and across domains and contexts, assessments are intended to support inferences regarding learners' knowledge, skills, and abilities (KSAs), based upon observations of what they say, do, or make in a handful of situations (Mislevy et al., 2003). In ECD assessment is conceptualized as five coordinated layers of components, processes, and representations. Relating directly to test development are the Domain Modeling layer, which delineates the assessment design and use argument, and the Conceptual Assessment Framework (CAF), which instantiates the argument with specifics such as scoring methods, measurement models, and schemas for tasks. This article describes the basic ideas of ECD and illustrates and discusses its applications and extensions across domains. For further discussion of ECD, see the "Further Reading" list at the end of the article.

Assessment as evidentiary argument

The most fundamental consideration in test development and evaluation is validity, which refers to "the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests" (AERA/APA/NCME, 2014, p. 11). Validity research thus concerns ways in which we can construct and use assessments to support interpretation and uses of test scores (Bachman and Palmer, 2010; Kane, 2006, 2013; Messick, 1994). Building an assessment argument means specifying elements that concord with Toulmin's (1958) argument schema. As is shown in Fig. 1 (Mislevy, 2018), the schema consists of a *claim* (C), which is a proposition we wish to support with *data* (D), and a *warrant* (W), which is a generalization that justifies the inference from the particular *data*

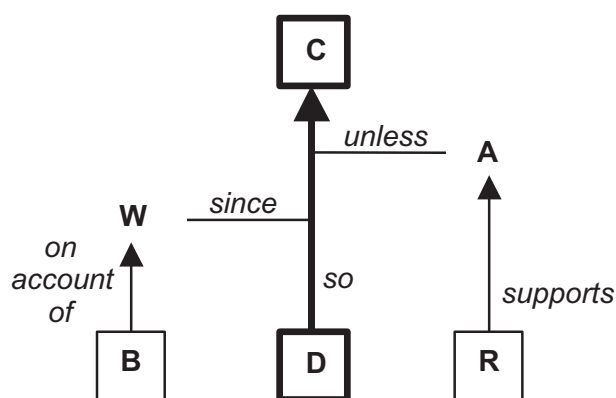


Fig. 1 Toulmin's general structure for arguments. The letters are the abbreviations for the following terms: B for *Backing*, D for *Data*, R for *Rebuttal*, W for *Warrant*, A for *Alternative Explanations*, and C for *Claims*.

(D) to the particular *claim* (C). When using the test, we reason back up through the *warrant* (W) from the observations to particular *claims* (C). Usually we must qualify our inferences because of *alternative explanations* (A), which *rebuttal data* (R) might support or undercut.

Assessment design means defining the kinds of claims we wish to make, and then crafting the operational elements of an assessment that allow us to reason in this manner. Messick (1994) outlines the essential structure:

A construct-centered approach would begin by asking what complex of knowledge, skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviors or performances should reveal those constructs, and what tasks or situation should elicit those behaviors? Thus, the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics (p. 17).

Test specifications

A traditional and useful representational form in the development of a test, regardless of its nature and intended uses, is the test specifications. The original focus was content and item types. Considerations of cognition were incorporated as in Bloom's (1956) taxonomy. Continued extensions provide additional support for the many specific considerations a test developer must make. Broadly-defined, test specifications document the test's purpose and intended uses while also including detailed decisions regarding its content, format, length, psychometric properties, delivery mode, administration, scoring, and score reporting (AERA/APA/NCME, 2014). From an ECD perspective, test specifications express selected information developed in the Domain Analysis, Domain Modeling, and CAF layers described in **Layers in ECD** section, in a form that supports the crafting of individual tasks and the composition of the test as a whole.

Assessment engineering

Another line of contributing research is assessment engineering and developing tools for improving the efficiency and quality of task design. Early theory-based approaches to task development were Hively et al. (1968) item shells, Haladyna and Shindoll's (1989) item forms, and Bormuth's (1970) transformational grammar item generation algorithms. Embretson (1985) provided a path toward integrating test theory, task design, and cognitive psychology. Irvine and Kyllonen (2002) reviewed articles and illustrative examples of theory-based item generation through 2000. More recent work varies physical features of tasks considered important cognitively (e.g., Gierl et al., 2008). Embretson's (1998) cognitive assessment design system and Luecht's (2013) integrated test design, development and delivery approach are two examples of technology supports for principled task design. From an ECD perspective, each of these efforts inform the development and the principled implementation of assessment tasks from a guiding construct, with results expressed in Domain Analysis and/or CAF objects.

Layers in ECD

ECD is best understood through its representation of layers in the design and implementation of educational assessment (Mislevy et al., 2002, 2003). Layering facilitates intrafield investigations as well as providing structures that support communication across realms of expertise (Riconscente et al., 2015). Table 1 summarizes the roles and key entities of layers, which are discussed below using examples and applications of ECD.

Table 1 Summary of ECD layers.

Layer	Role	Key entities	Examples of knowledge representations
Domain analysis	Gather substantive information about what is to be assessed.	Concepts, terminology, tools and representation forms; analysis of information use.	Representation forms and symbol systems employed in a given domain; common core state standards; state curriculum standards; syllabi.
Domain modeling	Express assessment argument in narrative form.	Knowledge, skills, and capabilities; task features; potential observations.	Toulmin diagrams; design patterns; claims, evidence, and task worksheets.
Conceptual assessment framework	Express assessment argument in specifications for tasks, measurement models, scoring procedures.	Student, evidence, and task model; task model variables; rubrics; measurement models; test specifications.	Task shells; task templates; graphical representation of measurement models.
Assessment implementation	Implement assessment, including creating tasks and fitting measurement models.	Task materials; pretest data for item and task scoring.	Items and tasks; design of test materials and score reports; item parameter data files.
Assessment delivery	Coordinate interaction of students and tasks; task and test-level scoring and reporting.	Task as presented; responses as collected; scores as evaluated.	Rendering of test materials; individual and group-level reports; data files of responses or logs of students' actions.

Adapted from Riconscente et al. (2015).

Domain analysis

Domain analysis marshals substantive information that has implications for assessment in the target domain. It is a preliminary synthesis of concepts, terminology, tools and representational forms that people engage with. While not constructed or organized in terms of the structures of an assessment, most of this information lays the foundation for developing it (Mislevy et al., 2002, 2003; Riconscente et al., 2015). Examples of this are beliefs, theory and research, subject-matter expertise, and exemplars from other assessments. Outcomes of this analysis include identification of KSAs that are essential for success in the domain, task features, and competency-task relationships. The design team for the Test of English for International Communication (TOEIC) speaking and writing test, for instance, initiated task design analysis by reviewing the framework papers for a related test, Test of English as a Foreign Language (TOEFL), in addition to the other materials about design and research regarding workplace and general English tests, which provided them with ideas about potential test tasks for the subsequent steps (Hines, 2010).

The inferences we wish to draw and the psychological perspective we take on the nature of KSAs serve as the starting point for domain analysis: our conception of the nature of human capabilities shapes the design and interpretation of elements, relationships, and rationale in an assessment argument. Four perspectives on construing instruction and assessment are trait, behavioral, information-processing, and sociocognitive, which holds implications for designing and using assessment (Mislevy, 2018). A trait is defined as a relatively stable characteristic of a person, consistently manifested to some degree, albeit with considerable variations across settings and circumstances (Messick, 1989). A prototypical example of tests cast in this perspective are aptitude tests, locating people along continua of mental characteristics, and supporting inferences regarding the degree to which individuals possess the capabilities. The domain analysis for these tests is thus an analysis of assemblies of abilities that capture individual between-person differences in the aptitude of interest. A behavioral perspective foregrounds targeted behavior in relevant situations, while placing cognitive and social considerations in the background. The domain analysis for tests cast in this perspective concerns features of what people do and features of situations in which the relevant behaviors are elicited. An information-processing perspective examines the procedures and cognitive processes individuals engage with for problem-solving. Assessment from this perspective focuses on the production system, i.e., knowledge, procedures, and strategies for problem-solving, in the domain of interest. An example of this is mixed-number subtraction, which provides diagnostic information regarding the strategies students use to solve the problems. From a sociocognitive perspective, people's thinking, acting, and learning in the physical world are mediated by "socially constructed practices, built around linguistic, cultural, and substantive (LCS) patterns and representations" (Mislevy, 2018, p. 47). From this perspective, individuals' capabilities are viewed as resource deployment and activation of cognitive resources attuned to LCS patterns that characterize regularities and variations in people and situations.

Domain modeling

The Domain Modeling layer articulates the assessment argument. It does not address technical details such as measurement models or data formats, but rather ways to fill in an argument scheme using information identified in the domain analysis layer. Schematic tools for this purpose are proficiency-, evidence-, and task paradigms, Toulmin diagrams, and design patterns. The paradigms are non-technical sketches of the student-, evidence-, and task models discussed in the ensuing section: proficiency paradigm organizes potential claims about students; evidence paradigm are the kinds of things students say or do that constitute evidence; and task paradigm are the kinds of situations that can be configured to obtain the evidence (Mislevy et al., 2003).

Fig. 2 is an extended Toulmin Diagram in the context of portfolio-based assessment of student interns pursuing K-12 teaching certification. The claim is an aspect of a student intern's KSAs for teaching a particular content area in K-12 classrooms. The evidence is an assembly of artifacts that reflect the student intern's teaching preparation and teaching practices in the assessment situations. In this assessment, student intern A is placed in a months-long internship program and has completed a number of teaching assignments for inclusion in the teaching portfolio, all requiring an integration of knowledge of a particular content area, of national and state teaching standards, and of unit- and content-based formative and summative assessment of student learning with teaching and assessment practices, which the student intern then demonstrates through completing tasks that include and are not limited to lesson planning, lesson delivery, and implementing unit-based formative and summative assessment. Theory and research regarding teaching and learning in that particular content area and national and state teaching standards provide backing in the assessment argument. A sociocognitive perspective sees the student acting in a given situation and frames claims, tasks, and evidence in terms of the student's regularities and variations for acting across situations, all informed by and contingent upon across-person regularities and variation at many levels of LCS patterns associated with the situations, and other information concerning the interaction between the student and the assessment situations (Mislevy, 2018).

Design patterns organize information from *domain analysis* to define a design space and an assessment structure for a task designer who needs to write tasks addressing the capabilities at issue (Mislevy et al., 2003; Riconscente et al., 2015). Table 2 shows the attributes of a design pattern. Focal KSAs are the capabilities which the tasks will be constructed to assess (i.e., the claims). Characteristic features of tasks must be included in some form in order to evoke evidence of the targeted capabilities, and variable features are ways a designer can shift focus or difficulty. Potential Work Products are the kinds of things student can say, do, or make to produce evidence, and Potential Observations are the qualities of the work that constitute evidence (e.g., accuracy of the inferences in a reading assessment).

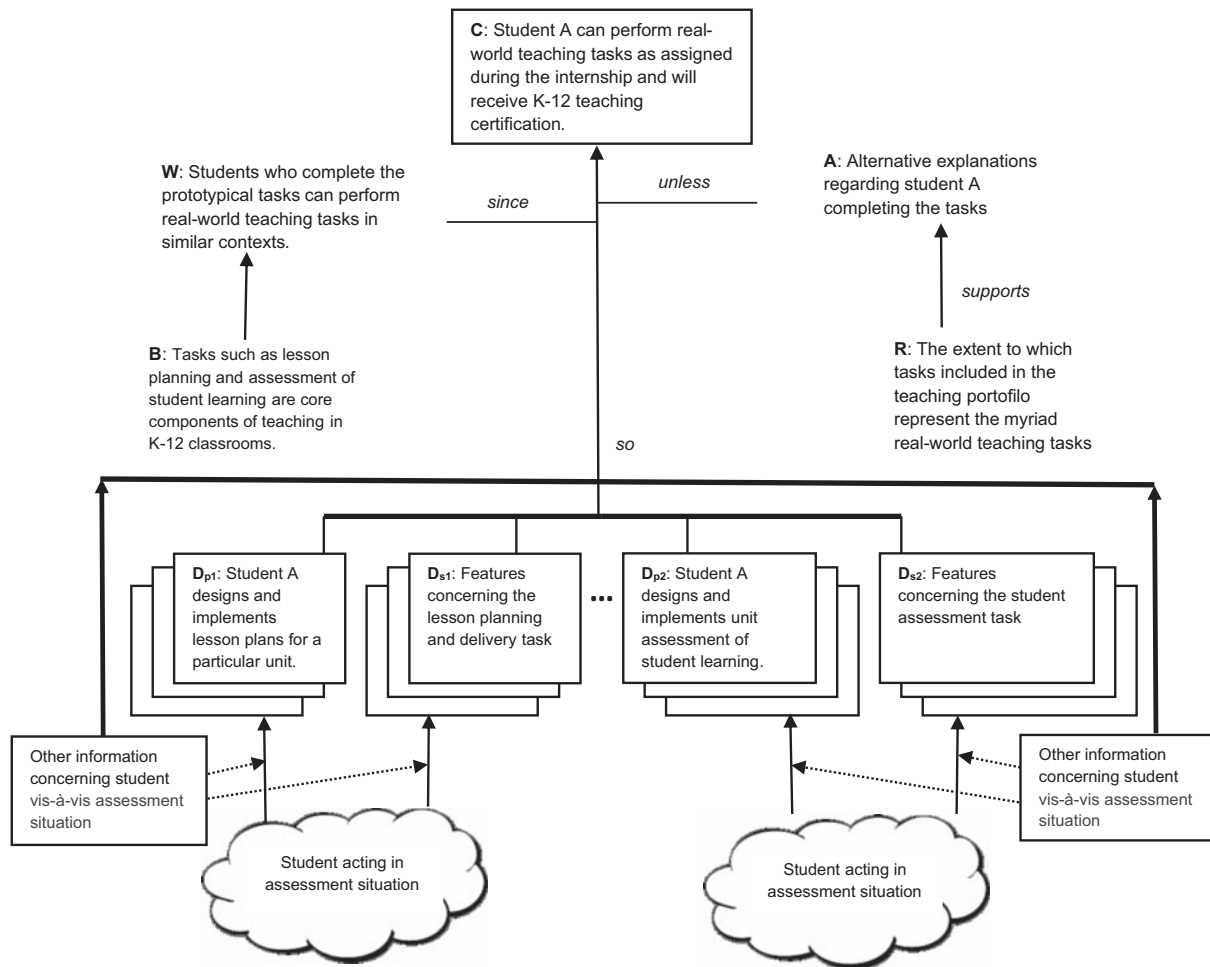


Fig. 2 Extended Toulmin diagram in the context of portfolio-based assessment of student interns.

Table 2 Design pattern attributes and relationships to assessment argument.

Attribute	Description	Assessment argument component
Rationale	Provide a rationale for linkage between potential observations and focal knowledge, skills and abilities (KSAs).	Warrant
Focal knowledge, skills and abilities (focal KSAs)	The primary knowledge/skill/abilities targeted by this design pattern.	Student model
Additional knowledge, skills and abilities	Other knowledge/skills/abilities that may be required by tasks generated under this design pattern.	Student model, if construct relevant; alternative explanations, if construct irrelevant.
Characteristic features of tasks	Aspects of assessment situations that must be present in the task situation in order to evoke the desired evidence.	Task model
Variable features of tasks	Aspects of assessment situations that can be varied in order to shift difficulty or focus.	Task model
Potential observations	Characteristics of performance that constitute evidence about KSAs (e.g., accuracy, efficiency, correct inference).	Evidence model
Potential work products	Performances by which students can produce evidence about focal KSAs (e.g., written text, response choice, oral speech).	Task model

Conceptual assessment framework

The *conceptual assessment framework* (CAF) provides a structure for coordinating the substantive, statistical, and operational aspects of an assessment. The technical detail required for implementing the assessment is laid out in this layer. This layer specifies the operational elements of assessment machinery in the form of three models that instantiate the assessment argument. The Student Model operationalizes claims in terms of student variables, which can be as simple as an overall score, or as complex as the person variables in an item response theory or cognitive diagnostic model. Values assigned to student model variables, based on observing performances, correspond to claims about students' capabilities.

The Evidence Model contains two components, the *evaluation component* and the *measurement model*. The *evaluation component* contains procedures for identifying and characterizing salient features of examinee performances—that is, how to score task performances. Through evaluation procedures using answer keys, rating rubrics with examples, and so on, we determine the values of *Observable Variables* from students' Work Products. What aspects of student work are important and how to evaluate them are governed by the psychological perspective in which the assessment is framed. Tests cast in behaviorist terms, with its focus on stimulus-response bonds, may focus on qualities such as accuracy of the responses. Tests cast in sociocognitive terms highlight the appropriateness and effectiveness of the choices an examinee makes in the assessment situation. In interactive tasks, evaluation may draw on natural language processing methods, learning analytics, or data mining techniques, and may comprise multiple stages (as in the automated scoring of spoken responses; Higgins et al., 2011). The *measurement model* provides a mechanism for synthesizing such data across tasks or interactions and for quantifying the strength of evidence for claims, such as reliability coefficients and standard errors of measurement. Psychometric models such as classical test theory, item response theory (Wilson, 2005), and cognitive diagnostic models (Rupp et al., 2010) are used in this step.

The Task Model consists of blueprints for assessment tasks and work products. In applications using design patterns, task models reflect choices from the suggested range of Potential Work Products, and Characteristic and Variable Features and Potential Observations of the Work Products. This layer specifies work products and narrows down to central and optional task features that evoke evidence about the targeted claims. For example, performance-based assessment of student teachers will necessarily involve observations of a student teaching a content class, and evaluations of qualities of lesson planning and lesson delivery. In contrast, if the claim concerns mastery of content or pedagogical knowledge, a computer-delivered knowledge assessment in the form of multiple-choice questions or short-essay questions may serve the purpose. In assessments consisting of multiple discrete tasks, constellations of tasks are brought together as an assessment in the *assembly model* to determine the breadth and depth of the domain being assessed. Test specifications and blueprints are generated in this model on the basis of CAF models. In interactive tasks such as in simulations and games, specifications for how interactions between test-takers and the task environment are effected, data are captured, and evaluations are produced are specified in more complex forms of these basic ECD models (Conrad et al., 2014; Ke et al., 2019).

The implementation and delivery layers

The *assessment implementation* layer includes activities such as developing interfaces, authoring tasks, fine-tuning rubrics and examples, programming automated scoring routines, and fitting measurement models. It is in this layer that assessment-engineering developments in technology for creating elements such as adaptive testing, interaction in simulation environments, and capture and evaluation of complex performances are effected, based upon specifications laid out in the CAF. Behrens et al. (2012), for instance, provide worked-through examples of how Cisco Networking Academies designed its next generation automated performance-based assessment, the Packet Tracer system, using the ECD framework.

The *assessment delivery* layer includes processes for selecting and administering test items, interacting with examinees, reporting scores, and providing feedback to appropriate parties. The information, structure, and rationale of these processes embody the assessment argument laid out in domain analysis, are structured by the blueprint laid out in the CAF, and utilize the operational elements created in *Assessment Implementation*. Details of delivery system architecture are provided in Almond et al. (2002) and, as extended to interactive digit assessments, in Conrad et al. (2014).

Applications and extensions

Across domains ECD applications involve its principles in action for developing large-scale assessment, for revising existing tests to more coherently articulate the assessment argument and align learning and assessment, and for generating simulation- and game-based assessment (GBA). The procedures Educational Testing Service used to develop test blueprints for the TOEFL iBT illustrates ECD principles as applied to a large-scale assessment of English learners (ELs) (Chapelle et al., 2008). Defining the construct to be measured and claims and subclaims about examinee abilities constitutes the proficiency paradigm in the domain modeling layer. For example, a claim for the speaking test is defined as "[examinees] can communicate effectively by speaking in English-speaking environments" (Pearlman, 2008: 232). Subsequent steps characterize observations for grounding them and identifying real-world tasks to provide evidence, resulting in evidence paradigms and task paradigms in the same layer. Based upon these results, templates for generating items or tasks, tasks shells, are constructed to incorporate essential elements of tasks models in the conceptual assessment framework (e.g., Pearlman, 2008: 237–238).

The redesign of Advanced Placement (AP) examinations is an illustration of the use of ECD for revising the existing suite of AP exams designed to measure skills and knowledge essential for success in introductory-level college courses (College Board, n.d.). Parallel to the redesign of the AP curriculum which begins with clearly defined learning outcomes essential for success in subsequent college courses, revision of the AP exams using the ECD approach begins by asking what kinds of claims we want to make about the students for the purpose of predicting their success in subsequent college courses, what kinds of evidence can support these claims, and what assessment tasks can provide evidence of the claims. The Domain Modeling layer focuses on developing exemplar claims and supporting evidence that characterize student performance at each of the achievement levels, i.e., Achievement Level Descriptors (ALDs), which are then used to construct the assessment framework (Brennan, 2010). Task models are embedded within each achievement model, which becomes the basis for generating items and test specifications. The use of ECD in redesigning AP exams explicitly articulates an evidentiary argument to support score interpretations as well as provides direct linkage among claims, evidence, task models, items, achievement levels, and test specifications (Hendrickson et al., 2010).

In the emerging field of GBA, an integration of ECD into the game development process leverages game design principles and mechanics to generate engagement, and assessment design principles and methods to generate useful evidence pertaining student learners' understandings and capabilities (Mislevy et al., 2014). An example of this is the GlassLab's game SimCityEDU: Pollution Challenge! which addresses the concept of "Systems and Systems Modeling" in the Next Generation Science Standards. The domain modeling layer draws on theory and research regarding the developmental trajectory of students' scientific reasoning and system thinking to gather substantive information about the assessment and assessment situations. Built upon this, an assessment design team consisting of game designers, psychometricians, subject-matter experts, teachers, psychologists, and experts in other relevant areas, each using their own language and tools, sketched out the essential elements of an assessment argument using the Toulmin Diagram and design patterns for each of the developmental levels in system thinking, while simultaneously addressing competing constraints. These elements translate into specifications of the student-, task-, and evidence models in the CAF, which define the operational elements subsequently constructed, prepared, and delivered in the assessment implementation and delivery layers.

Advances in computer technology and availability of rich data have brought forth myriad opportunities for assessment designers to apply and expand on ECD to develop innovative complex learning and assessment systems (e.g., Arieli-Attali et al., 2019; Behrens et al., 2012; Mislevy et al., 2012). One way assessment designers expand on ECD is to utilize rich data generated from simulation and game-based assessment to inform scoring rules and game design (Conrad et al., 2014; Grover et al., 2017; Shute et al., 2010). Adaptations also include retrofitting extensions to learning systems, or "ECD for Learning" (ECDL); Arieli-Attali et al. (2019) and Feng et al. (2009), for example, extend ECD with pedagogical models to utilize data within existing learning products. More generally, the expanded ECD framework (e-ECD; Arieli-Attali et al., 2019) expands all three basic models—for students, evidence, and tasks—by adding an additional learning layer in each model to specify targeted learning objectives, to provide task-based learning support, and to capture evidence for changes in learning. These point to important ways the ECD framework can coherently blend learning and assessment and simultaneously inform and guide learning in an era when the use of computer technology and digital devices has become ubiquitous.

Future directions

Scientific understanding of how students learn has advanced radically, and new technologies are coming online continually to provide interactive task environments and capture the resulting evidence. The challenge is harnessing this new knowledge for practical testing purposes—effectively, validly, affordably. Evidence-centered design provides a coherent conceptual framework that makes explicit the principles and interacting structures of argumentation, analysis, and operation that underlie assessment. It provides language and concepts to help developers design and coordinate the elements of familiar assessments as well as new ones. It motivates definitions of data structures and operational processes that can be used and re-used flexibly, across projects and organizations.

The work described in this article represents current practices in developing and using ECD. Future work will fill the framework out more fully, provide additional worked examples, and develop software tools and knowledge representations that make it easy for test developers to put the ideas to use. In the domain modeling layer, design patterns are needed to provide and support performance assessments and simulation-based assessments that focus on interactive capabilities. In the CAF layer, we need task design frameworks, adaptable modules for measurement models, and automated or rubric-based evaluation procedures, all explicated in terms of their relationship to assessment arguments. Recent work mentioned earlier shows how researchers are extending the ECD framework to incorporate learning with assessment and to detail the more complex design and analysis interplay in game- and simulation-based assessments. In the implementation and delivery layers, we need further representations and data structures that exploit computer capabilities to build, administer, and evaluate tests. These are ambitious and broad undertakings, spanning years or decades. At the same time, however, the most fundamental ideas of ECD—carefully thinking through what capabilities one wants to assess and the kinds of things we need to see learners do in what kinds of situations—can be put into action immediately by every teacher and every test designer in the field.

References

- Almond, R., Steinberg, L., Mislevy, R., 2002. Enhancing the design and delivery of assessment systems: a four-process architecture. *J. Technol. Learn. Assess.* 1. Retrieved from. <https://ejournals.bc.edu/index.php/jtla/article/view/1671>.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing, 2014. *Standards for Educational & Psychological Testing*. American Educational Research Association, Washington, DC.
- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., von Davier, A.A., 2019. The expanded evidence-centered design (e-ECD) for learning and assessment systems: a framework for incorporating learning goals and processes within assessment design. *Front. Psychol.* 10. <https://doi.org/10.3389/fpsyg.2019.00853>.
- Bachman, L.F., Palmer, A.S., 2010. *Language Assessment Practice: Developing Language Assessments and Justifying Their Use in the Real World*. Oxford University Press, Oxford.
- Behrens, J.T., Mislevy, R.J., DiCerbo, K.E., Levy, R., 2012. Evidence centered design for learning and assessment in the digital world. In: Michael, C.M., Clarke-Midura, J., Robinson, D.H. (Eds.), *Technology-Based Assessments for 21st Century Skills*. Information Age Publishing, Charlotte, NC.
- Bloom, B.S. (Ed.), 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I, Cognitive Domain*. Longman, London.
- Bornmuth, J.R., 1970. *On the Theory of Achievement Test Items*. University of Chicago Press, Chicago.
- Brennan, R.L., 2010. Evidence-centered assessment design and the Advanced Placement Program®: a psychometrician's perspective. *Appl. Meas. Educ.* 23 (4), 392–400. <https://doi.org/10.1080/08957347.2010.510973>.
- Chapelle, C., Enright, M.K., Jamieson, J.M. (Eds.), 2008. *Building a Validity Argument for the Test of English as a Foreign Language™*. Routledge, London.
- College Board, (n.d.). AP Course and Exam Redesign. Retrieved from <https://aphighered.collegeboard.org/courses-exams/course-exam-redesign>.
- Conrad, S., Clarke-Midura, J., Klopfer, E., 2014. A framework for structuring learning assessment in an educational massively multiplayer online educational game—experiment centered design. *Int. J. Game Base. Learn.* 4, 37–59. <https://doi.org/10.4018/IJGBL.2014010103>.
- Embretson, S.E. (Ed.), 1985. *Test Design: Developments in Psychology and Psychometrics*. Academic Press, New York.
- Embretson, S.E., 1998. A cognitive design system approach to generating valid tests: application to abstract reasoning. *Psychol. Methods* 3, 380–396. <https://doi.org/10.1037/1082-989X.3.3.380>.
- Feng, M., Hansen, E.G., Zapata-Rivera, D., 2009. Using Evidence Centered Design for Learning (ECDL) to Examine the Assistments System. Paper Presentation. American Educational Research Association Annual Meeting, San Diego, CA.
- Gierl, M.J., Zhou, J., Alves, C., 2008. Developing a taxonomy of item model types to promote assessment engineering. *J. Technol. Learn. Assess.* 7. Retrieved from. <https://files.eric.ed.gov/fulltext/EJ838625.pdf>.
- Grover, S., Bienkowski, M., Basu, S., Eagle, M., Diana, N., Stamper, J., 2017. A Framework for Hypothesis-Driven Approaches to Support Data-Driven Learning Analytics in Measuring Computational Thinking in Block-Based Programming. Seventh International Learning Analytics & Knowledge Conference (ACM), New York, NY.
- Haladyna, T.M., Shindoll, R.R., 1989. Shells: a method for writing effective multiple-choice test items. *Eval. Health Prof.* 12, 97–104. <https://doi.org/10.1177/016327878901200106>.
- Hendrickson, A., Huff, K., Luecht, R., 2010. Claims, evidence, and achievement-level descriptors as a foundation for item design and test specifications. *Appl. Meas. Educ.* 23 (4), 358–377. <https://doi.org/10.1080/08957347.2010.510966>.
- Higgins, D., Xi, X., Zechner, K., Williamson, D., 2011. A three-stage approach to the automated scoring of spontaneous spoken responses. *Comput. Speech Lang* 25, 282–306. <https://doi.org/10.1016/j.csl.2010.06.001>.
- Hines, S., 2010. Evidence-Centered Design: The TOEIC® Speaking and Writing Tests. Educational Testing Service. www.ets.org/research/policy_research_reports/publications/report/2010/itjx.
- Hively, W., Patterson, H.L., Page, S.H., 1968. A “universe-defined” system of arithmetic achievement tests. *J. Educ. Meas.* 5, 275–290. Retrieved from. www.jstor.org/stable/1433778.
- Irvine, S.H., Kyllonen, P.C. (Eds.), 2002. *Item Generation for Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Kane, M., 2013. Validating the interpretation and use of test scores. *J. Educ. Meas.* 50, 1–73. <https://doi.org/10.1111/jedm.12000>.
- Kane, M., 2006. Validation. In: Brennan, R.J. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger, Westport, CT, pp. 17–64.
- Ke, F., Shute, V., Clark, K.M., Erlebach, G., 2019. *Interdisciplinary Design of Game-Based Learning Platforms*. Springer Nature, Heidelberg, Germany.
- Luecht, R.M., 2013. Assessment engineering task model maps, task models and templates as a new way to develop and implement test specifications. *J. Appl. Test. Technol.* 14, 1–38. <http://jattjournal.com/index.php/atp/article/view/45254/36645>.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education/Macmillan, New York, pp. 13–103.
- Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educ. Res.* 32, 13–23. <https://doi.org/10.3102/0013189X023002013>.
- Mislevy, R.J., Steinberg, L.S., Almond, R.A., 2002. Design and analysis in task-based language assessment. *Lang. Test.* 19, 477–496. <https://doi.org/10.1191/0265532202lt2410a>.
- Mislevy, R.J., Hamel, L., Fried, R.G., Gaffney, T., Haertel, G., Hafter, A., et al., 2003. *Design Patterns for Assessing Science Inquiry (PADI Technical Report 1)*. SRI International. Retrieved from. http://padi.sri.com/downloads/TR1_Design_Patterns.pdf.
- Mislevy, R.J., Steinberg, L.S., Almond, R.A., 2003. On the structure of educational assessments. *Measurement* 1, 3–67. https://doi.org/10.1207/S15366359MEA0101_02.
- Mislevy, R.J., Behrens, J.T., DiCerbo, K.E., Levy, R., 2012. Design and discovery in educational assessment: evidence-centered design, psychometrics, and educational data mining. *J. Educ. Data Min.* 4, 11–48. <https://doi.org/10.5281/zenodo.3554641>.
- Mislevy, R.J., Corrigan, S., Oranje, A., DiCerbo, K., John, M., Bauer, M.I., et al., 2014. *Psychometric Considerations in Game-Based Assessment*. Institute of Play, New York, NY.
- Mislevy, R.J., 2018. *Sociocognitive Foundations of Educational Measurement*. Routledge, New York.
- Pearlman, M., 2008. Finalizing the test blueprint. In: Chapelle, C., Enright, M.K., Jamieson, J.M. (Eds.), *Building a Validity Argument for the Test of English as a Foreign Language™*. Routledge, London.
- Riconscente, M.M., Mislevy, R.J., Corrigan, S., 2015. Evidence-centered design. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, London.
- Rupp, A., Templin, J., Henson, R., 2010. *Diagnostic Measurement: Theory, Methods, and Applications*. Guilford Press, New York, NY.
- Shute, V.J., Masduki, I., Donmez, O., 2010. Conceptual framework for modeling, assessing and supporting competencies within game environments. *Nonlinear Opt. Quant. Opt. Concepts Mod. Opt.* 8, 137–161.
- Toulmin, S.E., 1958. *The Uses of Argument*. Cambridge University Press, Cambridge.
- Wilson, M., 2005. *Constructing Measures*. Lawrence Erlbaum, Mahwah, NJ.

Further reading

- Arieli-Attali, M., Ward, S., Thomas, J., Deonovic, B., von Davier, A.A., 2019. The expanded evidence-centered-design (e-ECD) for learning and assessment systems: a framework to incorporating learning goals and processes within assessment design. *Front. Psychol.* 10, 853.

- Chapelle, C., Enright, M.K., Jamieson, J.M. (Eds.), 2008. *Building a Validity Argument for the Test of English as a Foreign Language™*. Routledge, London.
- Luecht, R.M., 2013. Assessment engineering task model maps, task models and templates as a new way to develop and implement test specifications. *J. Appl. Test. Technol.* 14, 1–38.
- Mislevy, R.J., Yin, C., 2009. If language is a complex system, what is language assessment? *Lang. Learn.* 59 (Suppl. 1), 249–267.
- Mislevy, R.J., Almond, R.G., Lukas, J., 2004. A Brief Introduction to Evidence-Centered Design. CSE Technical Report 632. The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. <https://files.eric.ed.gov/fulltext/ED483399.pdf>.
- Mislevy, R.J., 2018. *Sociocognitive Foundations of Educational Measurement*. Routledge, New York.
- Riconscente, M.M., Mislevy, R.J., Corrigan, S., 2015. Evidence-centered design. In: Lane, S., Haladyna, T.M., Raymond, M. (Eds.), *Handbook of Test Development*. Routledge, London, pp. 40–63.
- Wilson, M., 2005. *Constructing Measures*. Lawrence Erlbaum, Mahwah, NJ.

Cognition-centered design principles for digital assessment tasks and items

Makeleine Keehner^a, Burcu Arslan^b, and Marlit Annalena Lindner^c, ^a Educational Testing Service, Princeton, NJ, United States; ^b ETS Global B.V., Amsterdam, the Netherlands; and ^c IPN—Leibniz Institute for Science and Mathematics Education, Kiel, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	171
Goals and scope of this chapter	171
Characteristics of digital assessment tasks and technology-enhanced items	172
Principles for cognition-centered task and item design	173
Working memory and the constraints of human information processing	173
Cognitive load	174
Gestalt principles of visual perception	174
Affordances	175
Interactivity	175
Multimedia	176
Visual media	176
Use of visual media in education and assessment	176
Representational and decorative pictures	176
Graphs and charts	178
Dynamic visual media: animations and videos	179
Designing tasks for process data modeling, analysis, and reporting	179
Conclusions and future directions for researchers and practitioners	181
References	181

Introduction

Digital assessment tasks and technology-enhanced items provide many design palettes and choices for assessment developers that go far beyond the options available in paper-and-pencil assessments. The interactive affordances of digitally-based assessment have led to some dramatic changes to the front-end design and the test-taker experience, as well as the types of data collected. Test takers' interactive behaviors, together with system events, can be captured at the back-end of digital assessment environments. With continuous logging of events in this "process data", and with timings captured to millisecond precision, test creators and administrators can capture detailed information about the observable external processes that precede, co-occur with, and broadly correlate with test-taker responses.

As a consequence, digital tasks and items have the potential to gather more evidence than was previously possible for inferences about test-taker knowledge, skills, and abilities, and offer a wider range of design possibilities for targeting so-called "hard to measure" constructs (Bennett et al., 1990; Parshall and Brunner, 2017; Parshall and Cadle, 2014; Parshall et al., 2011; Sireci and Zenisky, 2011). In some cases, they can also support inferences about contextual information such as test-taker response or solution strategies, (dis)engagement, and metacognitive processes (e.g., Arslan et al., 2020; Eichmann et al., 2020; He et al., 2021; Wise, 2017).

Nonetheless, the use of digital tasks and items in educational assessment is not on its own guaranteed to improve the quality of the assessment results or the validity of the inferences made from the data gathered. Purposeful, cognition-centered design of both the front-end and the back-end of digital items and tasks is crucial for achieving these goals.

Goals and scope of this chapter

Although computerized assessments offer many exciting design options, a lack of evidentiary basis for design decisions can threaten the validity of the interpretation of test scores (Parshall and Brunner, 2017; Parshall and Cadle, 2014; Parshall et al., 2011). In order to capitalize effectively on the tremendous potential of digital assessments, we argue that it is important to apply an evidence-based, theoretically-grounded *cognitive* framework to support digital design decisions from a principled basis. Starting from a cognition-centered foundation can help to ensure that interactive digital tasks and items are: (a) designed in ways that are consistent with human cognition; and (b) designed to optimize the interactive affordances available for gathering process data as evidence of the non-visible cognitive processes occurring in the mind of the test taker. In addition, applying cognitive theory provides an evidence-based theoretical grounding for evaluating the potential strengths and drawbacks of using different task types and design features for particular goals. This in turn should help to support the overarching assessment goal of making valid inferences about test takers (Katz et al., 2017).

Applying a cognition-centered approach means drawing on insights and principles from the decades-long cognitive and learning sciences research literature to proactively inform task and item design. The approaches and principles we present in this chapter, drawn from this body of knowledge, are well-aligned with the goal of designing interactive items and tasks to elicit construct-relevant and valid evidence for target constructs in the form of both product (i.e., response) and process data, and should support the effective design of interactive digital assessment tasks of all kinds. Note that many of these principles were originally established in instructional design research with a focus on learning materials and have not yet been subject to systematic research in assessment contexts; therefore, further research is certainly required to better understand their effects in testing (see also [Kirschner et al., 2017](#); [Lindner, 2021](#)). Nonetheless, there is evolving evidence that these principles can be helpful for better design of assessment tasks and items.

Digitally-based assessment tasks range from technology-enhanced test items with various response formats, such as drag-and-drop, inline-choice, or grid (e.g., [Arslan et al., 2020](#); [Moon et al., 2019](#)), to interactive performance tasks, including simulations and games (see [DiCerbo, 2022](#), this volume, for a detailed review), and extended scenario-based tasks and microworlds (e.g., [Krieger et al., 2021](#); [Sheehan and O'Reilly, 2012](#); [Zhang et al., 2019](#)). There is no single agreed-upon classification system or naming protocol, nor are there absolute boundaries between different kinds of tasks and items (e.g., a simulation may be designed to be more or less game-like; extended “scenario-based tasks” are often created from sequences of related short items with multimedia content and interactive features). Digital assessments also typically include additional materials such as tutorials and instructions, universal design elements, and accessibility tools, as well as the assessment interface. Given this complex background, our goal is to describe digital task design principles that are broadly domain-agnostic and not restricted to a certain task type or only relevant to assessment in a particular discipline.

What we discuss below are cognitively-grounded, domain-agnostic design considerations, which we believe should be considered throughout the course of digital task design and development. These must be balanced and integrated with content-specific domain-expert knowledge and assessment development expertise in order to meet the specific needs for construct-validity in a given assessment. The primary use-case we consider is summative assessment (i.e., assessments *of* learning), but much of what we cover is also relevant to formative applications (i.e., assessments *for* learning, [Bennett, 2011](#); [Black and Wiliam, 2009](#)).

In the remainder of this chapter, we first review some common features of digital tasks and items (**Characteristics of digital assessment tasks and technology-enhanced items**). We then discuss cognitively-based principles to support the design of digital tasks to be consistent with human cognition and information processing constraints (**Principles for cognition-centered task and item design**). Following this, we discuss how to design interactive digital tasks to collect useful process data for making valid inferences about test takers (**Designing tasks for process data modeling, analysis, and reporting**). Finally, we provide take-home conclusions and future directions (**Conclusions and future directions for researchers and practitioners**).

Characteristics of digital assessment tasks and technology-enhanced items

Although there is no single universally-agreed set of characteristics for categorizing and defining digital assessment tasks, various authors have attempted to describe the range of resources that exist in digital assessments. One high-level distinction is whether the problem has a single definitive correct response. [Bull and McKenna \(2003\)](#) describe *objective testing* as tasks in which the correct answer is predetermined. These tasks tend to include recognizable assessment components such as a stem and a response field or a set of response options. They might also make use of some interactive digital design affordances to present information or to allow different kinds of response modes. Since the correct answers are predetermined, they have the potential to support automated feedback to test takers. Objective tests are well suited to testing factual knowledge, but they can also be creatively designed to assess more complex constructs at the higher end of Bloom’s taxonomy (e.g., analysis, synthesis, evaluation; [Bloom, 1956](#); see also [Anderson et al., 2001](#)), especially if interactive capabilities or multimedia elements are incorporated in a construct-relevant manner.

[Bull and McKenna \(2003\)](#) defined a second category comprising activities in which there is not a single predetermined “correct” way of responding. These open-ended activities tend to make full use of interactive digital capabilities such as simulations, interactive multimedia, computer-mediated communication, virtual environments, and complex technology-enhanced items for the evaluation (or acquisition) of complex skills and knowledge. They are also potentially harder to score. The distinction between objective testing and open-ended testing is similar to the long-established distinction in cognitive science between well-defined problems and ill-defined problems ([Lynch et al., 2009](#)). Although not a strict dichotomy ([Le et al., 2013](#)), there is nonetheless substantial evidence that different cognitive processes occur when tackling these two kinds of problems (see, e.g., [Dörner and Funke, 2017](#); [Schraw et al., 1995](#)), and therefore this is an important high-level task characteristic for assessment designers to consider.

Several authors have focused primarily on the interactive affordances of computerized tasks. [Parshall et al. \(2002\)](#) used the term *innovative item types* to refer to items that have features or functions available as a result of being administered on a digital device. They identify five dimensions of innovation in computer-based assessment: item format, response action, media inclusion, level of interactivity, and scoring method. They claim that innovative items can improve measurement by: (a) reducing guessing (e.g., increasing the degrees of freedom for options that can be chosen in selected responses, or having test takers produce new types of constructed responses); (b) measuring some knowledge or skill more directly than traditional item types (i.e., having response formats that better map onto the cognitive process or mental representation of the test taker, such as clicking directly on a section of a graphic, or dragging and dropping sources into targets to represent a sequence of events); (c) expanding the content coverage (e.g.,

by presenting information via audio or visual media); and (d) expanding the cognitive skills measured (i.e., by having test takers participate in activities that are only feasible through interactive digital technologies).

Scalise and Gifford (2006) also focused on the interactive affordances of technology-enhanced items. Their taxonomy organizes innovative item types into seven categories, arranged according to the degree of complexity and the constraints placed on test takers in terms of the available ways of interacting and responding, ranging from fully constrained responses—traditional multiple-choice items—to fully constructed responses—essay items (see also Bennett et al., 1990, for an earlier classification of constructed response items). Their taxonomy is applicable to all innovative item types, including formats that use simulations, multimedia, animations, audio, and/or video resources.

Scalise and Gifford (2006) also emphasize that more research is needed to ensure that the potential advantages of innovative item types do not exceed the risks of incorporating them into assessment or learning environments (see also Russell, 2016, for a framework regarding the utility of technology-enhanced items; Sireci and Zenisky, 2016, for a review of the implications of the use of innovative item formats on score validity; and Bryant, 2017, for a review of the potential advantages and limitations of technology-enhanced items). Areas of potential risk include the possibility of cognitive overload occurring in test takers, as well as the risk of test takers engaging in unintended and construct-irrelevant processes, resulting from the rich and potentially complex design options offered by digital tasks. In the next section, we highlight some domain-general principles for designing digital assessment tasks and items that take into account the limitations of human cognition and can help to reduce the risk of cognitive overload.

Principles for cognition-centered task and item design

The cognitive and learning sciences have well-established knowledge bases and have developed broadly applicable principles relevant to design, generated over many decades from extensive empirical research conducted in fields such as cognitive psychology, educational psychology, human factors, human-computer interaction, attention and perception, usability design, multimedia design, embodied cognition, and interactivity research. The constraints of human cognition apply regardless of whether a task is designed for paper-and-pencil or computerized administration. But digital assessments offer many more design possibilities and therefore understanding how their varied information types and formats will be processed, and how test takers will think and behave with their interactive affordances, is especially important. As in all human information processing, the processing of test material involves perceptual, attentional, cognitive, and action systems that are tightly interconnected and have known properties and constraints. Designing digital materials in ways that are consistent with human cognition can help test takers:

- attend to key information rather than task-irrelevant information,
- be less likely to experience overall sensory and cognitive overload,
- more easily integrate text, images, and other types of information, and
- expend less cognitive capacity on extraneous (irrelevant) processes.

In the next sections we highlight some key properties of human cognition and related design principles that we believe are especially relevant to digital assessment design.

Working memory and the constraints of human information processing

When digital assessment tasks or learning materials are presented, the information is perceived via sensory channels, filtered and selectively attended to, and a mental representation is formed. The mental representation is synthesized with preexisting long-term knowledge from prior experience, mentally manipulated, and continually refreshed and updated with newly-encountered external information. Working memory (WM) is the mind's online workspace where this moment-to-moment representation and processing of information occurs (Baddeley, 1986, 2003; Baddeley and Hitch, 1974).

WM has dedicated buffers for handling and translating between different kinds of information—a phonological loop for generating and processing verbal-auditory information (sometimes called the “inner voice” and “inner ear”), and a visuo-spatial buffer for generating and processing visual and spatial images (sometimes called the “inner scribe” and “inner eye”; Baddeley, 1986). For example, when we read, we initially perceive and encode the text in its visual form, but the words are converted into a phonological form (the inner voice) in WM, to allow sense-making and manipulation of the information. WM also supports complex cognition involving higher-level “executive functions,” such as synthesizing different types and sources of information, generating new mental representations, and allocating cognitive resources to different mental activities (Baddeley, 2003).

Although WM is highly flexible and invaluable for our ability to process and respond to incoming information, it is also a relatively fragile and limited-capacity system (Cowan, 2010). That is, WM can only hold a relatively small amount of information at any given time; the information in WM is easily disrupted, such as by new incoming information; and it takes conscious mental effort to maintain and manipulate information in WM—for example, think of the mental effort it takes to hold complex information in your mind and how that information can be instantly lost if someone unexpectedly asks you an unrelated question. Understanding the ways in which WM and human information processing can create a bottleneck, and how information overload can be ameliorated, is important for designers of digital materials.

First, even before information reaches WM, competition can occur between the incoming signals if they involve the same sensory channel—for example, if an animation is presented simultaneously with captions underneath, trying to attend to both is likely to cause split attention (Ayres and Sweller, 2005). By contrast, auditory narration uses a different sensory channel and therefore does not compete with the animation for the viewer’s visual attention (Mayer, 2008).

Second, the buffers that hold information in WM are capacity-limited and easily disrupted by new incoming information, such that interference may occur even when information is received through different sensory channels. For example, reading text (visual) while hearing unrelated speech (auditory) causes interference. Even though these signals enter WM through separate external sensory channels, both contain linguistic information that must be processed in the phonological loop, and therefore they compete for the limited capacity of this WM buffer (Baddeley, 1986; Baddeley and Hitch, 2019).

Further, the higher-level executive capacities of WM that allow us to engage in complex cognitive activities are also limited, so that multipart or very demanding complex tasks cause difficulty and detriment in task performance (e.g., see Borst et al., 2010). Thus, designers of digital assessments should think about both the *external* representations they are presenting and the *internal* or *mental* representations and processes that are expected to derive from and be required by the external information, when considering test-taker capacity-limits in relation to these different aspects of information processing.

Cognitive load

The term “cognitive load” is often used as informal shorthand for overall mental effort during an activity that calls for WM resources. In the literature, theories of cognitive load have been developed and many empirical studies have been conducted to examine the effects on learning and performance from different types and sources of cognitive load (e.g., Sweller et al., 2011; Wirzberger et al., 2020). Given the role of WM in consciously-controlled, cognitively effortful information processing, and given its known capacity limitations (Baddeley, 1986; Engle, 2002), cognitive load in complex tasks can also be thought of as primarily WM load, that is, the amount of mental work conducted in WM (e.g., see Sweller, 2020).

Researchers have characterized different types of cognitive load. The so-called “intrinsic load” (Sweller et al., 2011) or “essential processing” (Mayer and Fiorella, 2014) of a given task has been characterized as the inherent mental work or the necessary cognitive steps required to complete the task. However, it is also recognized that the amount of task-relevant load is actually a product of the interaction between the task characteristics and the knowledge, skills, and aptitudes of the individual working on the task (Kalyuga et al., 2003). For example, a mathematical equation may be solved rapidly and easily by a proficient student, for whom the problem triggers previously learned, automated solution procedures (i.e., schemas), thus reducing the amount of work that has to be done in WM (e.g., Kalyuga, 2009). By contrast, for a less proficient student who does not have these automated schemas, the same task may pose more demands on WM.

So-called “extraneous load” refers to the mental effort expended on activities that are irrelevant to or unnecessary for the task at hand. For example, unproductive effort may be required to process information from displays that are poorly designed (Paas et al., 2004; Saß et al., 2017). Thus, extraneous load is not inherent in the task itself but arises from the way information is presented. For example, the way a task is designed may cause additional effort to locate, comprehend, or integrate task-relevant information, such as having to switch between tabs or having to mentally integrate text and diagrams that are spatially or temporally separated in the display (e.g., see Moon et al., 2022). Consequently, extraneous load can be altered by adjusting the way a task is designed and how the information is presented (see also the related distinction between informational equivalence and computational equivalence; Simon, 1978).

Understanding these ideas should be helpful for designers of assessment materials when thinking about both the task itself and the way in which it is represented. A central goal of assessment is to achieve desired item difficulty while avoiding construct-irrelevant variance (i.e., “excess reliable variance that is irrelevant to the interpreted construct”; Messick, 1989). It is important for assessment developers to consider potential sources of extraneous load or construct-irrelevant difficulty arising from the way information is presented. It should also be remembered that, as stated above, test takers with relevant prior knowledge may be able to offload otherwise cognitively effortful WM processes onto automated schemas, and thus total WM load for those individuals should be lower regardless (Koedinger and Anderson, 1990).

Using cognitive theory to predict and explain when and why different kinds of cognitive load might arise in a given task can provide a theoretically-grounded framework to inform design choices and reduce construct-irrelevant variance. Below we outline some principles that have been developed to reduce undesirable and task-irrelevant sources of WM load.

Gestalt principles of visual perception

The Gestalt school of psychology, which emerged in the early 20th Century, gave rise to a set of universal laws describing the ways in which humans perceive the world (e.g., see Wagemans et al., 2012 for original and more recently added Gestalt principles of perception). Of particular relevance to task design are the Gestalt principles of grouping, which refer to visual characteristics that tend to make viewers perceive elements as belonging together. These phenomena can be co-opted in the design of visual displays to support the perception of relatedness among elements and help learners attend to and perceive, or make connections among, aspects of the information presented.

The principle of grouping (see Wagemans et al., 2012) states that elements in the display will be perceived as related or belonging together if they are visually similar, which can include similarity of color, size, shape, orientation, motion, timing, font, and

synchronized changes. For example, color-coding a key term in the text and using the same color to highlight part of an image signals to the viewer that these two aspects are related (i.e., that the color-coded term is the name of the depicted object; [Richter et al., 2016](#)). Elements will also be perceptually grouped by the viewer if they are located close to each other; if they are symmetrical or parallel to each other; if they appear to form a continuous entity or a complete closed figure; and if they are contained within the same region or area (e.g., objects that are located inside the same boundary or the same shaded region of a screen will be perceived as belonging together).

These principles can be used to “pre-organize” information on the screen to reduce extraneous effort spent searching for connections, or to guide test-takers to attend to information in the intended order. Grouping principles of various types can also be used to proactively signal meaningful connections between elements to help in building a mental representation of the information or the problem to be solved (e.g., see [Kirschner et al., 2017](#); [Moon et al., 2022](#); [Saß and Schütte, 2016](#); [Saß et al., 2017](#)).

Affordances

When someone sees an object for the first time, its appearance can provide clues about how to interact with it. For instance, the shape and construction of a chair affords the action of sitting, the handle on a cup is designed to be grasped. These design aspects of objects, which provide opportunities for acting on them ([Gibson, 1979](#)) and may even “invite” certain behaviors ([Norman, 1999](#)), are known as affordances. Affordances are therefore easily perceivable action possibilities in external objects ([Norman, 1988, 1999](#)).

Controlled experimental studies have shown that different affordances in the design or layout of technology-enhanced items that are otherwise content-equivalent can affect test taker behaviors, time on task, and/or performance. The effects observed in these studies include differential behaviors elicited when response options were formatted in a grid layout versus a multiple-choice list ([Moon et al., 2019, 2020](#)); performance and response strategy differences generated by drag-and-drop items designed with different layouts and interactive features ([Arslan et al., 2019, 2020](#)); impacts on problem-solving efficiency and time on task in items incorporating an interactive signaling tool and/or spatially integrated information ([Moon et al., 2022](#)).

Such findings indicate that the affordances of different item formats and designs, including technology-enhanced interactive components, are critical to consider because they affect test-taker behavior and cognition even when the content is identical. [Kirsh \(2005\)](#) argues that good design helps users “discover” the affordances of external representations in such a way that they can make effective use of them. Item designers should therefore carefully consider the affordances that result from different design decisions. It is also important that researchers continue to explore these effects in carefully controlled experimental studies grounded in cognitive theory and communicate the implications of their findings to practitioners working in digitally-based assessment design.

Interactivity

Closely related to the notion of affordances is the issue of interactivity. Interactive resources, where the user can control what is on the screen to varying degrees, have become commonplace in digital assessments. The types of interactions available can range from relatively constrained interactions involved in selecting and providing a response (e.g., drag-and-drop; see [Arslan et al., 2020](#)) to resources that provide almost limitless degrees of freedom, such as simulations, games, and virtual worlds (e.g., [Ke and Shute, 2015](#)). In the cognitive science literature, researchers have identified different types and degrees of interactivity, from low-level control over pace and direction (play, pause, rewind, etc.), to altering parameters or exploring information from different perspectives. Interactions with materials can range from sequential, hierarchical, or nested structures to fully conditional interactivity that is responsive to all user manipulations (for some discussions of different types, levels, and definitions of interactivity, see [Bétrancourt, 2005](#); [Domagk et al., 2010](#); [Kioussis, 2002](#); [McMillan, 2002](#); [Renkl and Atkinson, 2007](#)).

In terms of whether having interactive control is helpful for cognition, the results of previous studies are mixed. With different tasks, users, and types of interactivity, different studies have found advantages, no differences, and even disadvantages for participants with interactive control of resources, relative to those not given control (e.g., see [Keehner et al., 2008](#), for an overview). Drawing conclusions about the benefits of interactivity is complicated by the fact that studies differ in the comparability of interactive and non-interactive conditions and in the levels and types of control available. However, the literature to date suggests that providing an interactive resource for communicating key information does not guarantee that it will improve comprehension or performance.

The ability of a user to grasp how to use an interactive resource effectively probably depends on internal factors, such as previous experience and expertise, and on the affordances of the external information itself, such as what is represented (and how familiar it is to the user), the method of control, the range of interactive behaviors that are possible, and the complexity of the information. For example, in a review of the literature on interactive animations, [Bétrancourt \(2005\)](#) hypothesized that only certain types of individual (e.g., those with more experience) will benefit from certain types of interactivity. Designers of interactive tools should therefore consider how the features of their design might interact with the characteristics of the intended user. For example, if a tool is designed primarily for domain experts, different principles would apply compared to a representation designed for domain novices. Assessment designers should also aim to make sure (ideally confirmed with empirical evidence gathered from user studies) that the intended interactive affordances are easily discoverable and useable in the design of the task, interface, or tool ([Kirsh, 2005](#); [Norman, 1999](#)).

Multimedia

Multimedia representations are typically defined as resources that present information using both words and pictures (Mayer and Fiorella, 2021). Multimedia design principles (e.g., see Mayer, 2014; Mayer and Fiorella, 2021) are research-based guidelines for designing these kinds of instructional materials to be consistent with the capabilities and constraints of WM and its associated sensory and attentional processes. Designing educational materials in line with multimedia principles has been shown to have significant positive effects on information processing and learning outcomes (e.g., see Mayer and Moreno, 2003).

It should be noted that the multimedia design principles have been established and empirically investigated mainly in relation to learning goals. Given the differences between learning and testing in terms of both educational goals and educational settings, it is unclear to what extent multimedia principles can simply be transferred from learning to testing materials (see Lindner, 2021 for a discussion). Below we provide descriptive suggestions with potential implementation examples in the assessment context, but more research in the context of assessment design is needed to validate the principles and investigate their boundary conditions for this specific use case.

Nonetheless, assessment developers should still be aware of multimedia design principles and consider them in their design decisions, since these principles respect and account for the limits of human cognition, and their implementation will therefore most likely reduce unnecessary WM load in assessment materials as well. In Table 1, we present a summary of some key design principles (synthesized from Mayer, 2014; Mayer and Fiorella, 2014) that could be particularly relevant to item and task design in assessment materials.

Visual media

Visual representations are ubiquitous in education and also play an important role in assessment. Given their proliferation and the many different types of visual displays that are used, we devote an extended section to discussing different types of visual media, their uses in assessment, and what has been learned about their cognitive properties and effects.

Use of visual media in education and assessment

With digitally-based assessments, the choice of available visual formats spans both static resources (e.g., pictures, diagrams, graphs) and dynamic formats with varying degrees of user control and interactivity (e.g., playable animations, interactive simulations, user-controlled dynamic data visualizations). Educational content, especially in STEM subjects, is often focused on phenomena involving objects, processes, relationships, systems, quantities, and theories or models. Including representational pictures along with text (i.e., multiple representations) can help to make this kind of information clearer than text only, less potentially ambiguous, and easier to understand, by directly displaying aspects such as physical characteristics or relationships, rather than making students attempt to reconstruct these properties mentally from text-based descriptions alone (e.g., see Ainsworth, 2006).

There are several ways in which well-designed visual displays can augment cognition and offload some of the mental effort associated with WM processes (e.g., see Ainsworth, 2006; Hegarty, 2011; Mayer and Fiorella, 2021; Padilla et al., 2018; Schnotz and Bannert, 2003). These include offloading the effort of mentally maintaining information in WM by making it available for reference in the external display; organizing the information in ways that support processes such as filtering, attending, visual search and integration; and offloading the process of perceiving patterns or key information (e.g., plotting data in a line graph instead of a bar graph to reveal a trend). Nonetheless, interpretations of displays are impacted by prior knowledge of the task content, and success in understanding and using a display depends on attending to it, interacting with it effectively, and having knowledge of relevant design conventions (e.g., see Hegarty, 2011; Schnotz, 2005).

For different types of visual displays in a task, assessment designers should support and guide test-taker attention by adhering to the multimedia and Gestalt principles referenced above (Mayer, 2014; Wagemans et al., 2012), such as signaling meaningful connections between sections and organizing information to make synthesis and integration of elements easier (e.g., Moon et al., 2022; Richter et al., 2016). However, test designers always need to keep in mind that support for test takers' mental model construction must be aligned with the measurement of the target construct, which is an individual decision tied to a specific assessment goal (e.g., see Lindner et al., 2017a,b).

To support informed decisions for selecting, modifying, or designing visual displays for assessments, it is important to evaluate the type of representation in terms of its general cognitive properties and how it might impact test-taker cognition and behavior (e.g., see Lindner, 2021). Images can be distinguished in various ways in the context of assessments, mainly based on their informational content (i.e., decorative, representational, organizational, informational, see Hu et al., 2021; Lindner, 2021, for classifications) and their relation to the explanation that is given in accompanying text. In the following subsections, we discuss in more detail some commonly used types of visual representations and their effects in assessment.

Representational and decorative pictures

Static pictures are frequently added to assessment tasks, but it is important to distinguish between different types of pictures and their functions. In assessment research, *representational pictures* can be defined as images that represent information from a corresponding text item stem in a pictorial way, without adding novel task-critical information (Lindner, 2021). Representational pictures are frequently used to display task- or construct-relevant information that is integral to the cognitive processes targeted

Table 1 Summary of instructional design principles that may be relevant for item and task design in assessment contexts, assuming the supported cognitive process is **not** part of the target construct of the measurement.

Principle or finding	Core idea	Potential implications for test item design	Example references for further information
Multimedia principle	The presentation of text with relevant pictures is more effective for learning than presenting text alone (i.e., better exploitation of WM capacity of the visual and verbal channel) and facilitates a proper mental model construction (e.g., by constraining interpretations).	Consider presenting task-relevant text information together with a corresponding representational or organizational picture to facilitate test-takers' mental model construction and understanding of a situation, if this is a prerequisite for a proper measurement of the target construct.	see, e.g., Ainsworth (2006), Butcher (2014)*, Lindner et al. (2017a*, 2017b*, 2018*), Lindner (2021)*, Mayer (1997), Saß et al. (2012)*
Coherence principle	Adding “seductive details” that are conceptually irrelevant can result in less deep processing of the core information, possibly due to priming of inappropriate schemas and limited WM resources.	Avoid unnecessary information such as “seductive details” in test materials (e.g., decorative elements) that add WM load but are not relevant to the task or the measurement of the core construct, especially if the seductive information is distracting in nature.	see, e.g., Lindner (2020)*, Mayer et al. (2001), Strobel et al. (2019)*
Signaling principle	Executive functions of attending, selecting, organizing, and integrating given information with existing knowledge can exceed WM capacity. Semantic and/or visual signals can support cognitive processing of complex information by guiding attention to the most important aspects of a material.	Use <i>semantic signals</i> such as preview summaries, section headings, labels and pointer words (e.g., “because”) for complex written explanations. Use <i>visual cues</i> (color, emphasis, pointers, etc.) to make relevant information more visually salient and support connections between corresponding text and picture information in multimedia displays.	see, e.g., Hegarty et al. (2010), Mautone and Mayer (2001), Moon et al. (2022)*, Richter et al. (2016), Saß et al. (2017)*, Saß and Schütte (2016)*
Spatial contiguity principle	Mental integration of spatially disconnected information is demanding of WM capacity. Proximity of corresponding text and images helps build connections between the information and reduce extraneous WM load during information integration.	Add chunks of text in closeness, proximity to semantically related parts (e.g., a picture) to avoid spatial split attention, enhance efficient processing, and to support a meaningful integration of corresponding information.	see, e.g., Jarodzka et al. (2015)*, Moon et al. (2022)*, Moreno and Mayer (1999), Saß et al. (2017)*
Temporal contiguity principle	Mentally integrating sources of information that are separated by time places a burden on WM. Presenting information simultaneously or close together in time (provided they do not compete for the same sensory channel) reduces WM load.	Present related verbal and pictorial information simultaneously and in non-competing ways (e.g., as an animation with a simultaneous narration), if supporting test-takers' mental model construction reduces task-irrelevant cognitive load.	see, e.g., Ginns (2006), Mayer and Anderson (1992)
Static media hypothesis	Static illustrations tend to reduce extraneous processing and promote active processing of information (e.g., imagery or reasoning) compared to animations, which may facilitate mental processes but can also introduce passive viewing (underwhelming effect) or lead to WM overload (overwhelming effect).	Consider using static images or sequential pictures over animations (even when presenting dynamic processes) and encourage the viewer to “mentally animate” sequential pictures, if understanding a process is a relevant prerequisite for the measurement of a target construct.	see, e.g., Hegarty (2004), Lowe (2003), Mayer et al. (2005)
Segmenting principle	Viewer-controlled pace of a presentation or animation allows to properly integrate presented information to build a mental representation while avoiding WM overload.	Break up complex information or animations into segments/chunks and give test takers control over the pace at which they view the segments. Allow user control of speed, or play/pause in animations to manage WM load in complex tasks.	see, e.g., Mayer and Chandler (2001), Mayer et al. (2019)
Pre-training principle	Presenting basic information (e.g., terminology, parts of complex processes) in a pre-training can help to reduce WM capacity that is needed during more complex mental model construction of the main material.	Consider offloading mental demands by introducing general or complex information in a familiarization phase in which test-takers are, for example, informed about item formats in a test, task requirements, technical terms and functions (e.g., in interactive tasks) or provided with required prior knowledge before they solve the target items.	see, e.g., Mayer et al. (2002), Mayer and Pilegard (2005)

Note. WM = working memory.

*Indicates research studies explicitly focused on assessment contexts and materials (unmarked references are focused on learning contexts and materials).

by the task (e.g., see Lindner et al., 2018; Saß et al., 2012). By contrast, *decorative pictures* can be defined as task-irrelevant visual additions that are not required to solve an item and are usually added to tests with the goal of enhancing test-taking pleasure. Experimental findings indicate that adding decorative elements (such as attractive photographs) to test items does not enhance motivation (Lindner, 2020), yet in contrast to research on seductive details in learning materials (e.g., see Wiley, 2019), decorative elements do not seem to significantly harm test-taker performance by distracting them from the main task (Berends and van Lieshout, 2009; Cooper et al., 2018; Lindner, 2020).

Several studies in the assessment field have examined the effects of including representational pictures that visualize task-relevant information from the test item stem (e.g., as compared to text-only items). Research from controlled experimental studies (e.g., Lindner et al., 2017a,b, 2018, 2020; Lindner, 2020, 2021; Ott et al., 2018; Saß et al., 2012), has shown a cognitive benefit from representational images in terms of improved task performance. An eye-tracking study by Lindner et al. (2017a), further revealed the potential of representational pictures to improve effective information processing by supporting students' mental model construction as well as their decision-making processes. Eye movement data showed that the picture attracted attention early in the onset of the item on-screen (i.e., mental scaffolding) and in the later phase of the decision-making process (i.e., mental model updating) and that students needed less time to read the accompanying text when a representational picture was present. Studies also suggest that adding representational pictures to text-based test items has the potential to improve self-reported item-solving satisfaction (e.g., Dewolf et al., 2015; Lindner et al., 2018, 2020; Lindner, 2020) and lowers the incidence of disengaged rapid-guessing behavior (Lindner et al., 2017b, 2019; Wise et al., 2009).

Taken together, these findings suggest that representational pictures help to provide cognitive scaffolds to support students' understanding of the task they are asked to solve. However, the effects of representational pictures also seem to depend at least to some extent on the specifics of the layout, the task and domain (e.g., see Hu et al., 2021; Lindner, 2021, for reviews). For example, even the specific layout of a picture (e.g., a linear format vs. a circular arrangement of the same 24-hour clock) can affect its impact on different tasks (see Schnotz and Kürschner, 2008). Thus, it is important not only to consider the display *content* but also its *format* and how well it supports the cognitive task to be completed.

Graphs and charts

Many different kinds of graphs and charts are used in education to convey quantitative information. In assessment, these types of pictures can generally be classified as *informative* and/or *organizational* pictures, since they display additional task-critical information that is complementary to information provided in a verbal format (see Lindner, 2021). Graphs are used to assess skills of data interpretation and graph generation, but they can also be used as a stimulus for unrelated measures such as speaking or writing skills.

Multiple studies in cognitive science have shown that graph comprehension is influenced by prior knowledge and long-term schema, such as knowledge of how to read a specific type of graph, as well as a person's schemas and assumptions about the referents represented in the graph. Most importantly for the assessment context, graph comprehension is also affected by bottom-up factors such as graph design features, as well as the complexity of the graph and the data represented (e.g., see Freedman and Shah, 2002; Kozhevnikov et al., 2007; Ratwani and Trafton, 2008; Shah, 2002; Shah and Freedman, 2011; Shah et al., 2005; Shah and Hoeffner, 2002).

Since particular conventions govern the formatting and design of different graph types, viewers must know these conventions in order to interpret a graph correctly (Hegarty, 2011; Ratwani and Trafton, 2008; Shah et al., 2005). Whether to include graphs and complex tables can therefore be a sensitive decision in assessment materials. In cases where differences in graph reading skills can introduce construct-irrelevant variance (i.e., when graph reading is not part of the targeted construct), supporting accurate interpretation of the data represented, by including design features to scaffold and guide graph comprehension, is especially critical for valid interpretation of test-taker performance. For example, in assessment tasks that used graphs as stimuli for speaking tasks in tests of language skills, the number of visual chunks in a graph (i.e., the amount of information to be integrated, which contributes to general mental effort when processing graphs) predicted performance (Katz et al., 2004; Xi, 2005). Moreover, reducing graph complexity and giving participants additional time to process the graph information in advance of the speaking task reduced construct-irrelevant variance arising from performance differences between students with differing degrees of graph familiarity (Xi, 2005, 2010).

Eye-tracking studies are especially well-suited to understanding how individuals read graphs and how different layout decisions impact cognitive information processing and the steps that occur as students answer questions about the information conveyed by a graph. Carpenter and Shah (1998) found that the majority of graph readers' time was spent looking at the axes and labels. Moreover, participants reexamined the same areas over and over, suggesting that mentally maintaining all of the displayed information in WM was challenging, especially for more complex, higher-dimensional graphs. Applying graphical features in these areas to capture viewers' attention and guide interpretation (e.g., salient visual signals to direct attention or Gestalt grouping principles to indicate connections) could help viewers to process the graph and free-up more WM capacity for interpreting the data (Shah and Hoeffner, 2002; Shah et al., 1999).

These methods have also been used to empirically study the effects of including graphs in assessment items. An eye-tracking study by Strobel et al. (2018) showed that adding task-irrelevant data to a graph can significantly increase task-processing time, error rates, and reported cognitive load. In particular, the addition of task-irrelevant data series (i.e., second factor) was detrimental, whereas the addition of task-irrelevant data points was less harmful, indicating that test designers should reduce graph information to the relevant core, if graph reading skills are not the targeted construct. Task-irrelevant additions to graphs should generally be avoided, even if they are not all equally harmful. For example, seductive text and seductive decorative pictures did not affect

students' performance, but eye-tracking data showed that the processing time (especially on the seductive content) was significantly increased (Strobel et al., 2019). Finally, to reduce task-irrelevant cognitive load, test developers should make sure that they display the graph type that is best suited for the task goal (e.g., line graphs for reading trends; e.g., Strobel et al., 2016).

Dynamic visual media: animations and videos

When the information to be communicated is dynamic in nature, such as a process that occurs over time or a system in action, it may seem logical to use dynamic playable media such as video or animations to convey that information. Well-designed dynamic visualizations can more directly represent dynamic processes than static images (e.g., see Lowe and Boucheix, 2016), but cognitive benefits of animated graphics over static representations have not been consistently found for learning (Bétrancourt and Tversky, 2000). Decades of research in cognitive science indicate that animations can be difficult for viewers to process and comprehend (e.g., see Höffler and Leutner, 2007 for an overview) and therefore should probably be used somewhat sparingly in assessment items.

The question of when to use dynamic rather than static representations in assessment is complicated by inconsistent findings on the advantages and disadvantages of different kinds of dynamic playable representations in educational settings, making it hard to conclusively establish when to use video or animation rather than static images (Berney and Bétrancourt, 2016; Höffler and Leutner, 2007). Nonetheless, Bétrancourt (2005) provides some helpful suggestions in the context of learning materials to answer the question "Do I really need to use an animation?" She suggests that animations may be beneficial when communicating changes over time if learners may not be able to infer the steps between static images; or when users are novices, or are under a very high cognitive load, and may not be able to form a mental model of dynamic phenomena without actually seeing them. She also suggests using animations for particular educational goals—for example, to visualize phenomena that are not easily directly perceptible (e.g., too fast, too small, or too dangerous to observe directly). However, it is known that prior domain knowledge, as well as visuospatial cognitive skills, impact how readily and effectively students can benefit from dynamic educational resources (e.g., see Berney and Bétrancourt, 2016), and this should also be taken into consideration when making design decisions about dynamic visualizations.

If a dynamic visualization becomes part of an item, test designers should consider established instructional design guidelines such as matching the type of visualization to what is being represented (Ainsworth and VanLabeke, 2004), giving the user control over the speed of the animation (as well as play, pause, rewind, etc.) so they can match the playing speed with their speed of comprehension (Schwan and Riempp, 2004), and providing complementary verbal-auditory information to scaffold the information without overloading visual working memory (Huff and Schwan, 2012). It can also be a helpful design strategy to break apart complex processes into serialized static pictures or shorter animations that show different parts of causal chains, with corresponding narrations and explanations of how they are related to each other (Narayanan and Hegarty, 1998, 2002).

Designing tasks for process data modeling, analysis, and reporting

As mentioned above, digital assessment platforms have made it possible to include interactive response formats, such as drag-and-drop, and interactive components and multimedia stimuli, such as text-to-speech, key-term translations, online tutorials, and calculators (Bartram and Hambleton, 2005; Ercikan and Pellegrino, 2017). With purposeful back-end design, process data can capture test-taker interactions along with their respective timing for later modeling, analysis, or reporting. This includes interactions with on-screen elements, in addition to interactions related to entering the response, such as drag-and-drop actions, keystrokes, and mouse clicks. In this regard, a key benefit of digital assessments and technology-enhanced items is the potential to collect information about the interactions of test takers with the item functions while completing a task (Ercikan and Pellegrino, 2017).

There are different use cases for process data in educational assessment. Process data can be used: (a) to investigate potential user interface and experience issues that are irrelevant to the construct in pilot studies or experimental studies to make sure that the interface design or the response format choices are functioning in the way they are intended to function (e.g., Arslan et al., 2020; Moon et al., 2020); (b) to make fine-grained inferences about test-takers' cognition, such as information processing, decision making (e.g., Katz et al., 2017; Oranje et al., 2017; Lindner et al., 2017a,b), or higher-level cognition such as knowledge, skills, or strategies (e.g., Greiff et al., 2018); (c) to detect rapid guessing and aberrant response behaviors (e.g., Wise, 2017); (d) to validate score meaning (e.g., Ercikan and Pellegrino, 2017) and support score reporting (e.g., The Nation's Report Card, <https://nces.ed.gov/nationsreportcard/>); and e) to provide feedback to test takers (e.g., Wise et al., 2006) or proctors (e.g., Wise et al., 2019). See also Nikolay et al. (2022, this volume), for a thorough review of computer-based process data analysis in education and further use cases of process data in formative assessment (i.e., assessment for learning).

However, these low-level logged interactions do not by themselves represent test-taker knowledge, skills, and abilities. The context of the interaction (i.e., the preceding and subsequent test-taker actions or system events) needs to be considered and aggregated (e.g., see Goldhammer et al., 2021). Moreover, there is always uncertainty associated with inferences made from process data, since captured behavior patterns can also be related to off-task behavior or cognitively disengaged test-taking behavior, such as mind wandering or interacting with a task without being cognitively engaged.

In order to make valid and relevant inferences about cognitive processes from process data, it is important to have theoretically-grounded approaches for process data modeling (e.g., Forsyth et al., 2014; Zapata-Rivera and Arslan, 2021). It is also important that interactive task design is grounded in a *cognitive model*—that is, a theoretical account of the processes and steps assumed to occur

during that cognitive activity. A cognitive model is not the same as a target construct—a target construct, such as physics knowledge or computer programming skills, is what we hope to measure through an assessment, defined in a way to allow assessment developers to devise a task to elicit evidence from which they can infer a student’s competency in that construct (e.g., see Evidence Centered Design, [Mislevy et al., 2003](#); and In-Task Assessment framework; [Kerr et al., 2016](#)). This can in principle be done without any formal model of the cognitive processes involved in engaging in those activities ([Leighton and Gierl, 2007](#)).

By contrast, one can use a cognitive model to describe and/or explain the mental steps or processes that occur when test takers engage in an activity. This provides a cognitively grounded account of the mental processes associated with the construct targeted by a task. For problems in a given domain, we may attempt to specify the kinds of cognitive processes that are recruited and identify steps or strategies that are commonly applied by test takers who are engaging in solving specific problems (e.g., see [Goldhammer et al., 2021](#)). This can also include individual differences in the relevant processes and solution steps. For example, depending on test-taker proficiency, due to different starting points of prior knowledge levels, we may expect differing quality of mental representations of problems or situations, and different approaches to reasoning and problem solving (e.g., [Chi et al., 1981](#); [Larkin et al., 1980](#); [Newell and Simon, 1972](#); [Newell, 1990](#); see also enacted constructs: [Gorin, 2006](#); [Keehner et al., 2016](#)).

Thus, when designing an assessment task, if we want to make evidence-based inferences from process data about cognition, a cognitive model is needed in which the mental and behavioral processes that are expected to occur during a certain activity are specified. This theoretical grounding can then be used to design interactive elements that not only support, frame, or constrain the activity to the specified processes, but that will also make it possible to find evidence of the expected cognitive steps or processes in the log data of actions as recorded when the test taker interacts with the task elements. If we wish to apply quantitative scoring rules or qualitative criteria for evaluating the processes inferred from the log data, we also need to specify the ways in which the relevant cognitive processes might differ in individuals with more or less proficiency, and how the process data arising from the task would reflect those differences (e.g., [Katz et al., 1998](#)).

From this perspective, once a task or item is envisioned, there are some iterative steps that can help to ensure that it will produce process data that is able to support fine-grained and targeted inferences about test takers’ cognitive processes. These iterative design steps are illustrated in [Fig. 1](#).

As [Fig. 1](#) shows, the first foundational step is to model the mental processes that are associated with the target construct in the context of the proposed task or item (i.e., determine the required solution process or expected reasoning steps that test takers will engage in to provide their response). The next step is to conceive and design a set of activities that would elicit these mental steps and capture evidence of them in the form of external actions, such as test-taker interactions with tools and technology-enhanced response formats that reflect the steps in the reasoning process (e.g., interaction with a calculator, drag-and-drop response actions). By designing the task or item with interactive affordances that require or encourage construct-related external actions, which capture the sequences of test takers’ reasoning processes as they unfold in time, the task or item has the capability of “embodying” students’ mental steps in its process data outputs, to provide observable evidence of these internal processes.

Once these two key design steps are accomplished, it is helpful to generate research questions and hypotheses about construct-relevant observable behaviors (e.g., action sequences that indicate misconceptions; potential range of tool use patterns such as calculator use) and ensure that the logged actions include the dimensions and the details needed to distinguish construct-relevant individual differences among test takers in the anticipated range of key behaviors. If, during this phase, it becomes evident

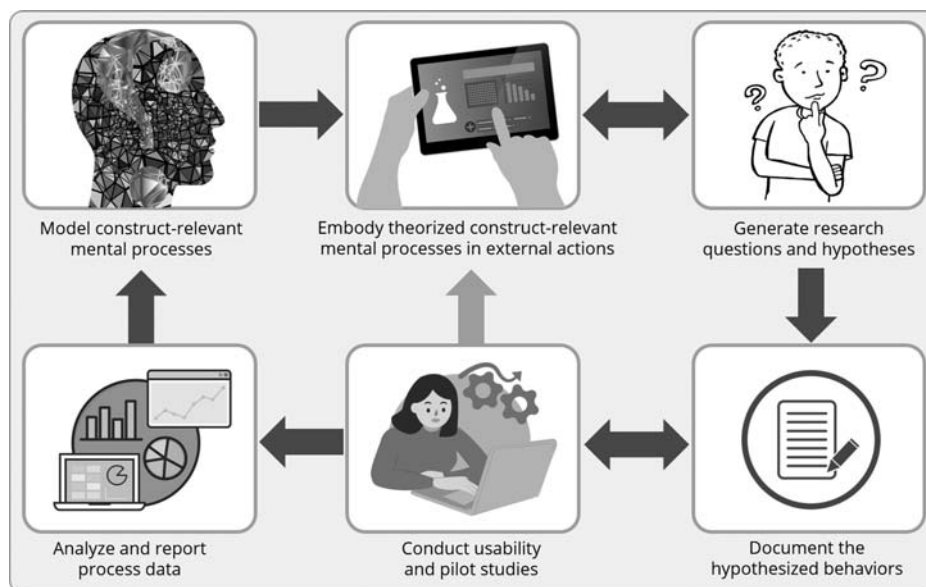


Fig. 1 Iterative steps in designing tasks for process data modeling, analysis, and reporting.

that the current design of the task is not able to provide evidence for the expected range of hypothesized construct-relevant observable behaviors, the design should be revised accordingly. A further important related step is to document the hypothesized behaviors so that there is a theoretical basis for making inferences about student knowledge, skills, and abilities in any subsequent process data modeling and analyses. Usability and/or pilot studies should also be conducted, with further revisions to the item or task after these, if necessary, to ensure the task design and the process data it produces include the properties and features needed to make valid inferences about the mental processes associated with the target construct.

The resulting process data from the task or item should be analyzed from both a top-down theory-driven perspective (i.e., based on the previously documented research questions) and a data-driven exploratory approach (i.e., exploring the process data to discover new insights). If the exploratory analyses reveal additional constructs or hypotheses that were not anticipated or have not been encountered before, the task designer should revisit the first stages of the design process (modeling the mental processes and documenting hypothesized behaviors) and the new observations and inferences should be incorporated.

Ideally, tasks resulting from this design approach should be evaluated to establish whether they are eliciting the theorized cognitive processes. Cognitive scientists have developed a range of research methodologies to support such inferences, which can be used for task evaluation. These include asking individuals to verbalize their thoughts as they work through a task (think aloud methodology; e.g., see [Somerén et al., 1994](#) for an introduction) or examining how individuals attend to and shift their gaze between different parts of a stimulus (eye tracking methodology; e.g., see [Holmqvist et al., 2011](#) for an introduction). These methods can produce external traces of the presumed mental processes, known as *protocol data* (and, in fact, process data is another form of protocol data, in that the log of test-taker interactions reflects the invisible mental processes occurring during task completion as they unfold in time). Validation studies using such methods can help researchers to evaluate whether the processes and steps predicted by the cognitive model are indeed being observed during the activity ([Keehner et al., 2016](#)).

Conclusions and future directions for researchers and practitioners

Advancements in technology have drastically increased the design options for interactive digital assessments and technology-enhanced items. With these increased options, it is more critical than ever to consider the constraints of human cognition when designing items or tasks, so as not to put the validity of the interpretation of the test outcomes at risk (see [Principles for cognition-centered task and item design](#)). To this end, we have briefly reviewed and summarized some key insights from the decades of relevant research findings and principles from the cognitive and learning sciences, to encourage assessment developers, testing companies, and educational institutions to apply a cognition-centered design approach instead of a technology-centered approach.

Although there are principles for cognition-centered design (see [Table 1](#)) that may provide some guidance, it is important to further investigate the effects of specific design choices on test-taker cognition and performance. Therefore, as the field of assessment design principles is still evolving (and it cannot be taken for granted that we can simply transfer all of the principles from learning to testing materials), it is critical to conduct more research in which the effects of design choices are investigated by conducting think-aloud studies, eye-tracking studies, and randomized, controlled experiments, so that researchers can develop cognition-centered theories about the design of digital assessment tasks and technology-enhanced items. Also, it is important that researchers disseminate their findings and communicate practical implications for assessment professionals.

Moreover, to be able to make valid, fine-grained inferences about test-taker cognition from digital assessments, we have argued that it is important to design tasks and items with the explicit goal of capturing meaningful and interpretable process data (see [Designing tasks for process data modeling, analysis, and reporting](#) and [Fig. 1](#) for detailed suggestions on how to design digital tasks and items for process data). Designing for valid inferences about test-taker cognition from process data requires cognitively-grounded design considerations at both the front-end and the back-end of tasks and items to ensure that externalized construct-relevant mental processes are captured in the log files.

To put all of the above-mentioned suggestions, principles, and methods into action, content experts, item writers, and assessment developers need to work with an interdisciplinary team that involves cognitive scientists, user experience and user interface designers, psychometricians, and software developers, from the onset of digitally-based task creation and design, and throughout the development process. Designing tasks and items with cognitive processes in mind helps to support the overarching assessment goal of making valid inferences about test-taker cognition.

References

- Ainsworth, S., VanLabeke, N., 2004. Multiple forms of dynamic representation. *Learn. InStruct.* 14 (3), 241–255.
- Ainsworth, S., 2006. DeFT: a conceptual framework for considering learning with multiple representations. *Learn. InStruct.* 16, 183–198.
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Rath, J., Wittrock, M.C., 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, abridged ed. Longman, White Plains, NY.
- Arslan, B., Jiang, Y., Gong, T., Keehner, M., Katz, I.R., 2019. A Computational Cognitive Modeling Approach to Understand Test-Takers' Strategy Use in Drag-and-Drop Math Questions. *CogSci*.
- Arslan, B., Jiang, Y., Keehner, M., Gong, T., Katz, I.R., Yan, F., 2020. The effect of drag-and-drop item features on test-taker performance and response strategies. *Educ. Meas.* 39 (2), 96–106.

- Ayres, P., Sweller, J., 2005. The split-attention principle in multimedia learning. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*, second ed. Cambridge University Press, Cambridge, pp. 135–146.
- Baddeley, A.D., Hitch, G., 1974. Working memory. *Psychol. Learn. Motiv.* 8, 47–89.
- Baddeley, A.D., Hitch, G., 2019. The phonological loop as a buffer store: an update. *Cortex* 112, 91–106.
- Baddeley, A.D., 1986. *Working Memory*. Oxford University Press, Oxford.
- Baddeley, A.D., 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4 (10), 829–839.
- Bartram, D., Hambleton, R., 2005. *Computer-Based Testing and the Internet: Issues and Advances*. John Wiley & Sons, Chichester.
- Bennett, R.E., Ward, W.C., Rock, D.A., LaHart, C., 1990. Toward a Framework for Constructed-Response Items (RR-90-7). Educational Testing Service, Princeton, NJ.
- Bennett, R.E., 2011. Formative assessment: a critical review. *Assess. Educ. Princ. Pol. Pract.* 18 (1), 5–25.
- Berends, I.E., van Lieshout, E.C.D.M., 2009. The effect of illustrations in arithmetic problem-solving: effects of increased cognitive load. *Learn. Instr.* 19, 345–353.
- Berney, S., Bétrancourt, M., 2016. Does animation enhance learning? A meta-analysis. *Comput. Educ.* 101, 150–167.
- Bétrancourt, M., Tversky, B., 2000. Effect of computer animation on users' performance: a review (Effet de l'animation sur les performances des utilisateurs: une synthèse). *Trav. Hum. Le.* 63 (4), 311–329.
- Bétrancourt, M., 2005. The animation and interactivity principles in multimedia learning. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, Cambridge, pp. 287–296.
- Black, P., William, D., 2009. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* 21 (1), 5–31.
- Bloom, B.S., 1956. Taxonomy of Educational Objectives. In: *Cognitive Domain*, vol. 1, p. 1. New York: McKay, 20(24).
- Borst, J.P., Taatgen, N.A., Van Rijn, H., 2010. The problem state: a cognitive bottleneck in multitasking. *J. Exp. Psychol. Learn. Mem. Cogn.* 36 (2), 363–382.
- Bryant, W., 2017. Developing a strategy for using technology-enhanced items in large-scale standardized tests. *Practical Assess. Res. Eval.* 22 (1), 1.
- Bull, J., McKenna, C., 2003. *A Blueprint for Computer-Assisted Assessment*. Routledge, New York, NY.
- Butcher, K.R., 2014. The multimedia principle. In: Mayer, Richard E. (Ed.), *The Cambridge handbook of multimedia learning*, 2nd. Cambridge University Press, pp. 174–205.
- Carpenter, P.A., Shah, P., 1998. A model of the perceptual and conceptual processes in graph comprehension. *J. Exp. Psychol. Appl.* 4 (2), 75–100.
- Chi, M.T., Feltovich, P.J., Glaser, R., 1981. Categorization and representation of physics problems by experts and novices. *Cognit. Sci.* 5 (2), 121–152.
- Cooper, J.L., Sidney, P.G., Alibali, M.W., 2018. Who benefits from diagrams and illustrations in math problems? Ability and attitudes matter. *Appl. Cognit. Psychol.* 32 (1), 24–38.
- Cowan, N., 2010. The magical mystery four: how is working memory capacity limited, and why? *Curr. Dir. Psychol. Sci.* 19 (1), 51–57.
- Dewolf, T., Van Dooren, W., Hermens, F., Verschaffel, L., 2015. Do students attend to representational illustrations of non-standard mathematical word problems, and, if so, how helpful are they? *Instr. Sci.* 43 (1), 147–171.
- DiCerbo, K., 2022. Simulation- and game-based assessment. In: Rivzi, F., Ericikan, K., Smith, G. (Eds.), *International Encyclopedia of Education*, fourth ed. Elsevier, Oxford, UK.
- Domagk, S., Schwartz, R.N., Plass, J.L., 2010. Interactivity in multimedia learning: an integrated model. *Comput. Hum. Behav.* 26 (5), 1024–1033.
- Dörner, D., Funke, J., 2017. Complex problem solving: what it is and what it is not. *Front. Psychol.* 8, 1153.
- Eichmann, B., Goldhammer, F., Greiff, S., Brandhuber, L., Naumann, J., 2020. Using process data to explain group differences in complex problem solving. *J. Educ. Psychol.* 112 (8), 1546.
- Engle, R.W., 2002. Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11 (1), 19–23.
- Ericikan, K., Pellegrino, J.W., 2017. *Validation of Score Meaning for the Next Generation of Assessments*. Routledge, New York, NY.
- Forsyth, C., Graesser, A., Pavlik Jr., P.I., Millis, K., Samei, B., 2014. Discovering theoretically grounded predictors of shallow vs. deep-level learning. In: *Proceedings of the 7th International Conference on Educational Data Mining*, pp. 229–232.
- Freedman, E.G., Shah, P., 2002. Toward a Model of Knowledge-Based Graph Comprehension. Paper Presented at the International Conference on Theory and Application of Diagrams.
- Gibson, J.J., 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.
- Ginns, P., 2006. Integrating information: a meta-analysis of the spatial contiguity and temporal contiguity effects. *Learn. Instr.* 16 (6), 511–525. <https://doi.org/10.1016/j.learninstruc.2006.10.00>.
- Goldhammer, F., Hahnel, C., Kroehne, U., Zehner, F., 2021. From byproduct to design factor: on validating the interpretation of process indicators based on log data. In: *Large-Scale Assessments in Education*, vol. 9, pp. 1–25, 1.
- Gorin, J.S., 2006. Test design with cognition in mind. *Educ. Meas.* 25 (4), 21–35.
- Greiff, S., Molnár, G., Martin, R., Zimmermann, J., Csapó, B., 2018. Students' exploration strategies in computer-simulated complex problem environments: a latent class approach. *Comput. Educ.* 126, 248–263.
- He, Q., Borgonovi, F., Paccagnella, M., 2021. Leveraging process data to assess adults' problem-solving skills: using sequence mining to identify behavioral patterns across digital tasks. *Comput. Educ.* 166, 104170.
- Hegarty, M., 2004. Dynamic visualizations and learning: getting to the difficult questions. *Learn. Instr.* 14 (3), 343–351. <https://doi.org/10.1016/j.learninstruc.2004.06.007>.
- Hegarty, M., 2011. The cognitive science of visual-spatial displays: implications for design. *Top. Cogn. Sci.* 33 (3), 446–474.
- Hegarty, M., Canham, M.S., Fabrikant, S.I., 2010. Thinking about the weather: how display salience and knowledge affect performance in a graphic inference task. *J. Exp. Psychol.: Learn. Mem. Cogn.* 36 (1), 37–53. <https://doi.org/10.1037/a0017683>.
- Höfler, T.N., Leutner, D., 2007. Instructional animation versus static pictures: a meta-analysis. *Learn. Instr.* 17 (6), 722–738.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., De Weijer, J., 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, Oxford, UK.
- Hu, L., Chen, G., Li, P., Huang, J., 2021. Multimedia effect in problem-solving: a meta-analysis. *Educ. Psychol. Rev.* 33, 1717–1747.
- Huff, M., Schwan, S., 2012. The verbal facilitation effect in learning to tie nautical knots. *Learn. Instr.* 22 (5), 376–385.
- Jarodzka, H., Janssen, N., Kirschner, P.A., Erkins, G., 2015. Avoiding split attention in computer-based testing: is neglecting additional information facilitative? *Br. J. Educ. Technol.* 46, 803–817.
- Kalyuga, S., Ayres, P., Chandler, P., Sweller, J., 2003. The expertise reversal effect. *Educ. Psychol.* 38 (1), 23–31.
- Kalyuga, S., 2009. Knowledge elaboration: a cognitive load perspective. *Learn. Instr.* 19 (5), 402–410.
- Katz, I.R., Martinez, M.E., Sheehan, K.M., Tatsuoka, K.K., 1998. Extending the rule space methodology to a semantically-rich domain: diagnostic assessment in architecture. *J. Educ. Behav. Stat.* 23 (3), 254–278.
- Katz, I.R., Xi, X., Kim, H.J., Cheng, P.C.H., 2004. Elicited speech from graph items on the test of Spoken English™. *ETS Res. Rep. Ser.* 2004 (1), i–31.
- Katz, I.R., LaMar, M.M., Spain, R., Zapata-Rivera, J.D., Baird, J.-A., Greiff, S., 2017. Validity issues and concerns for technology-based performance assessments. In: Sottilare, R.A., Graesser, A.C., Hu, X., Goodwin, G.A. (Eds.), *Design Recommendations for Intelligent Tutoring System-Volume 5: Assessment Methods*. US Army Research Laboratory (ARL), Orlando, Florida, pp. 209–224.
- Ke, F., Shute, V., 2015. Design of game-based stealth assessment and learning support. In: Loh, C., Sheng, Y., Ifenthaler, D. (Eds.), *Serious Games Analytics*. Springer, New York, NY, pp. 301–318.
- Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., Montello, D.R., 2008. Spatial reasoning with external visualizations: what matters is what you see, not whether you interact. *Cognit. Sci.* 32 (7), 1099–1132.
- Keehner, M., Gorin, J.S., Feng, G., Katz, I.R., 2016. Developing and validating cognitive models in assessment. In: Rupp, A.P., Leighton, J.P. (Eds.), *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, New York, pp. 75–101.

- Kerr, D., Andrews, J.J., Mislevy, R.J., 2016. The in-task assessment framework for behavioral data. In: Rupp, A.P., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, New York, pp. 472–507.
- Kiousis, S., 2002. Interactivity: a concept explication. *New Media Soc.* 4 (3), 355–383.
- Kirschner, P., Park, B., Malone, S., Jarodzka, H., 2017. Towards a cognitive theory of multimedia assessment (CTMMA). In: Spector, J.M., Lockee, B.B., Childress, M.D. (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*. Springer, Cham, pp. 1–23.
- Kirsh, D., 2005. Metacognition, distributed cognition and visual design. In: Gardenfors, P., Johansson, P. (Eds.), *Cognition, Education, and Communication Technology*. L. Erlbaum Associates, Mahwah, N.J., pp. 147–180.
- Koedinger, K.R., Anderson, J.R., 1990. Abstract planning and perceptual chunks: elements of expertise in geometry. *Cognit. Sci.* 14 (4), 511–550.
- Kozhevnikov, M., Motes, M.A., Hegarty, M., 2007. Spatial visualization in physics problem solving. *Cognit. Sci.* 31 (4), 549–579.
- Krieger, F., Stadler, M., Bühner, M., Fischer, F., Greiff, S., 2021. Assessing complex problem-solving skills in under 20 minutes. In: *Psychological Test Adaptation and Development*.
- Larkin, J., McDermott, J., Simon, D.P., Simon, H.A., 1980. Expert and novice performance in solving physics problems. *Science* 208 (4450), 1335–1342.
- Le, N.-T., Loll, F., Pinkwart, N., 2013. Operationalizing the continuum between well-defined and ill-defined problems for educational technology. *IEEE Trans. Learn. Technol.* 6 (3), 258–270.
- Leighton, J.P., Gierl, M.J., 2007. Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educ. Meas.* 26 (2), 3–16.
- Lindner, M.A., Eitel, A., Strobel, B., Köller, O., 2017a. Identifying processes underlying the multimedia effect in testing: an eye-movement analysis. *Learn. Instruct.* 47, 91–102.
- Lindner, M.A., Lüdtke, O., Grund, S., Köller, O., 2017b. The merits of representational pictures in educational assessment: evidence for cognitive and motivational effects in a time-on-task analysis. *Contemp. Educ. Psychol.* 51, 482–492.
- Lindner, M.A., Ihme, J.M., Saß, S., Köller, O., 2018. How representational pictures enhance students' performance and test-taking pleasure in low-stakes assessment. *Eur. J. Psychol. Assess.* 34, 376–385.
- Lindner, M.A., Lüdtke, O., Nagy, G., 2019. The onset of rapid-guessing behavior over the course of testing time: a matter of motivation and cognitive resources. *Front. Psychol.* 10 (1533), 1–15.
- Lindner, M.A., Schult, J., Mayer, R.M., 2020. A multimedia effect for multiple-choice and constructed-response test items. *J. Educ. Psychol.* 1–17.
- Lindner, M.A., Eitel, A., Barenthien, J., Köller, O., 2021. An integrative study on learning and testing with multimedia: effects on students' performance and metacognition. *Learn. Instruct.* 71, 101100.
- Lindner, M.A., 2020. Representational and decorative pictures in science and mathematics tests: do they make a difference? *Learn. Instruct.* 68, 101345.
- Lindner, M.A., 2021. Principles for educational assessment with multimedia. In: Mayer, R., Fiorella, F. (Eds.), *Cambridge Handbook of Multimedia Learning*, third ed. Cambridge University Press, Cambridge.
- Lowe, R.K., Boucheix, J.-M., 2016. Principled animation design improves comprehension of complex dynamics. *Learn. Instruct.* 45, 72–84.
- Lowe, R.K., 2003. Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction* 13 (2), 157–176. [https://doi.org/10.1016/S0959-4752\(02\)00018-X](https://doi.org/10.1016/S0959-4752(02)00018-X).
- Lynch, C., Ashley, K.D., Pinkwart, N., Alevy, V., 2009. Concepts, structures, and goals: redefining ill-definedness. *Int. J. Artif. Intell. Educ.* 19 (3), 253–266.
- Mayer, R.E., Anderson, R.B., 1992. The instructive animation: helping students build connections between words and pictures in multimedia learning. *J. Educ. Psychol.* 84 (4), 444–452.
- Mayer, R.E., Chandler, P., 2001. When learning is just a click away: does simple user interaction foster deeper understanding of multimedia messages? *J. Educ. Psychol.* 93 (2), 390–397. <https://doi.org/10.1037/0022-0663.93.2.390>.
- Mayer, R.E., Fiorella, L., 2014. Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, Cambridge, pp. 279–315.
- Mayer, R.E., Fiorella, L. (Eds.), 2021. *The Cambridge Handbook of Multimedia Learning*, third ed. Cambridge University Press.
- Mayer, R.E., Hegarty, M., Mayer, S., Campbell, J., 2005. When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *J. Exp. Psychol. Appl.* 11 (4), 256–265. <https://doi.org/10.1037/1076-898X.11.4.256>.
- Mayer, R.E., Heiser, J., Lonn, S., 2001. Cognitive constraints on multimedia learning: when presenting more material results in less understanding. *J. Educ. Psychol.* 93 (1), 187–198. <https://doi.org/10.1037/0022-0663.93.1.187>.
- Mayer, R.E., Matthias, A., Wetzell, K., 2002. Fostering understanding of multimedia messages through pre-training: evidence for a two-stage theory of mental model construction. *J. Exp. Psychol. Appl.* 8 (3), 147–154. <https://doi.org/10.1037/1076-898X.8.3.147>.
- Mayer, R.E., Moreno, R., 2003. Nine ways to reduce cognitive load in multimedia learning. *Educ. Psychol.* 38 (1), 43–52.
- Mautone, P.D., Mayer, R.E., 2001. Signaling as a cognitive guide in multimedia learning. *J. Educ. Psychol.* 93 (2), 377–389. <https://doi.org/10.1037/0022-0663.93.2.377>.
- Mayer, R.E., 1997. Multimedia learning: are we asking the right questions? *Educ. Psychol.* 32, 1–19.
- Mayer, R.E., 2008. Research-based principles for learning with animation. In: *Learning With Animation: Research Implications for Design*. Cambridge University Press, New York, NY, pp. 30–46.
- Mayer, R.E., 2014. Research-based principles for designing multimedia instruction. In: Benassi, V.A., Overson, C., Hakala, C.M. (Eds.), *Applying Science of Learning in Education: Infusing Psychological Science Into the Curriculum*.
- Mayer, R.E., Pilegard, C., 2005. Principles for managing essential processing in multimedia learning: segmenting, pretraining, and modality principles. In: Mayer, R.E. (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge University Press, pp. 169–182.
- Mayer, R.E., Wells, A., Parong, J., Howarth, J.T., 2019. Learner control of the pacing of an online slideshow lesson: does segmenting help? *Appl. Cogn. Psychol.* 33 (5), 930–935. <https://doi.org/10.1002/acp.3560>.
- McMillan, S.J., 2002. Exploring models of interactivity from multiple research traditions: users, documents, and systems. In: Lievrouw, L.A., Livingstone, S. (Eds.), *Handbook of New Media: Social Shaping and Consequences*. SAGE Publications, Ltd, pp. 205–229.
- Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. Macmillan, Washington, pp. 13–104.
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A brief introduction to evidence-centered design. *ETS Res. Rep. Ser.* 2003 (1), i–29.
- Moon, J.A., Keehner, M., Katz, I.R., 2019. Affordances of item formats and their effects on test-taker cognition under uncertainty. *Educ. Meas.* 38 (1), 54–62.
- Moon, J.A., Keehner, M., Katz, I.R., 2020. Test takers' response tendencies in alternative item formats: a cognitive science approach. *Educ. Assess.* 25 (3), 236–250.
- Moon, J.A., Lindner, M.A., Arslan, B., Keehner, M., 2022. Investigating the Split-Attention Effect in Computer-Based Assessment: Spatial Integration and Interactive Signaling Approaches. *Educ. Meas.: Issues Pract.* <https://doi.org/10.1111/emip.12485>.
- Moreno, R., Mayer, R.E., 1999. Cognitive principles of multimedia learning: the role of modality and contiguity. *J. Educ. Psychol.* 91 (2), 358–368.
- Narayanan, N.H., Hegarty, M., 1998. On designing comprehensible interactive hypermedia manuals. *Int. J. Hum. Comput. Stud.* 48 (2), 267–301.
- Narayanan, N.H., Hegarty, M., 2002. Multimedia design for communication of dynamic information. *Int. J. Hum. Comput. Stud.* 57 (4), 279–315.
- Newell, A., Simon, H., 1972. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.
- Newell, A., 1990. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA.
- Nikolay, B., Krieger, F., Greiff, S., 2022. Simulation- and game-based assessment. In: Rivzi, F., Ercikan, K., Smith, G. (Eds.), *International Encyclopedia of Education*, fourth ed. Elsevier.
- Norman, D.A., 1988. *The Design of Everyday Things*. Basic Books, New York.
- Norman, D.A., 1999. Affordances, conventions, and design. *Interactions* 6, 38–42.

- Oranje, A., Gorin, J., Jia, Y., Kerr, D., Ercikan, K., Pellegrino, J., 2017. Collecting, analyzing, and interpreting response time, eye tracking and log data. In: Ercikan, K., Pellegrino, J.W. (Eds.), *Validation of Score Meaning for the Next Generation of Assessments*. Routledge, New York, pp. 39–51.
- Ott, N., Brünken, R., Vogel, M., Malone, S., 2018. Multiple symbolic representations: the combination of formula and text supports problem solving in the mathematical field of propositional logic. *Learn. InStruct.* 58, 88–105.
- Paas, F., Renkl, A., Sweller, J., 2004. Cognitive load theory: instructional implications of the interaction between information structures and cognitive architecture. *Instr. Sci.* 32 (1/2), 1–8.
- Padilla, L., Creem-Regehr, S., Hegarty, M., Stefanucci, J., 2018. Decision making with visualizations: a cognitive framework across disciplines. *Cogn. Res.* 3, 3–29.
- Parshall, C.G., Brunner, B., 2017. Content development and review. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practice*. Routledge, Abingdon, pp. 85–104.
- Parshall, C.G., Cadle, A.W., 2014. How to identify, develop and implement innovative items. In: *The Annual Conference of the Association of Test Publishers*.
- Parshall, C.G., Spray, J.A., Kalohn, J., Davey, T., 2002. *Practical Considerations in Computer-Based Testing*. Springer Science & Business Media, New York.
- Parshall, C.G., Depascale, C., Skinner, L., 2011. Adapting Parshall and Harmes' Six-Stepmodel to the Integration of Alternative Item Types Into a Multiple-Choice Testing Program. Presented at the Annual Meeting of ATP, Phoenix, AZ, February 27–March 2.
- Ratwani, R.M., Trafton, J.G., 2008. Shedding light on the graph schema: perceptual features versus invariant structure. *Psychon. Bull. Rev.* 15 (4), 757–762.
- Renkl, A., Atkinson, R.K., 2007. Interactive learning environments: contemporary issues and trends. An introduction to the special issue. *Educ. Psychol. Rev.* 19 (3), 235–238.
- Richter, J., Scheiter, K., Eitel, A., 2016. Signaling text-picture relations in multimedia learning: a comprehensive meta-analysis. *Educ. Res. Rev.* 17, 19–36.
- Russell, M., 2016. A framework for examining the utility of technology-enhanced items. *J. Appl. Test. Technol.* 17 (1), 20–32.
- Saß, S., Schütte, K., 2016. Helping poor readers demonstrate their science competence: item characteristics supporting text–picture integration. *J. Psychoeduc. Assess.* 34 (1), 91–96.
- Saß, S., Wittwer, J., Senkbeil, M., Köller, O., 2012. Pictures in test items: effects on response time and response correctness. *Appl. Cognit. Psychol.* 26, 70–81.
- Saß, S., Schütte, K., Lindner, M.A., 2017. Test-takers' eye movements: effects of integration aids and types of graphical representations. *Comput. Educ.* 109, 85–97.
- Scalise, K., Gifford, B., 2006. Computer-based assessment in e-learning: a framework for constructing "intermediate constraint" questions and tasks for technology platforms. *J. Technol. Learn. Assess.* 4 (6).
- Schnotz, W., Bannert, M., 2003. Construction and interference in learning from multiple representation. *Learn. InStruct.* 13, 141–156.
- Schnotz, W., Kürschner, C., 2008. External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instr. Sci.* 36 (3), 175–190.
- Schnotz, W., 2005. An integrated model of text and picture comprehension. In: Mayer, R.E. (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, Cambridge, pp. 49–69.
- Schraw, G., Dunkle, M.E., Bendixen, L.D., 1995. Cognitive processes in well-defined and ill-defined problem solving. *Appl. Cognit. Psychol.* 9 (6), 523–538.
- Schwan, S., Riempp, R., 2004. The cognitive benefits of interactive videos: learning to tie nautical knots. *Learn. InStruct.* 14 (3), 293–305.
- Shah, P., Freedman, E.G., 2011. Bar and line graph comprehension: an interaction of top-down and bottom-up processes. *Top. Cogn. Sci.* 3 (3), 560–578.
- Shah, P., Hoeffner, J., 2002. Review of graph comprehension research: implications for instruction. *Educ. Psychol. Rev.* 14 (1), 47–69.
- Shah, P., Mayer, R.E., Hegarty, M., 1999. Graphs as aids to knowledge construction: signaling techniques for guiding the process of graph comprehension. *J. Educ. Psychol.* 91 (4), 690–702.
- Shah, P., Freedman, E.G., Vekiri, I., 2005. The comprehension of quantitative information in graphical displays. In: Shah, P., Miyake, A. (Eds.), *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press, Cambridge, pp. 426–476.
- Shah, P., 2002. Graph comprehension: the role of format, content and individual differences. In: Anderson, M., Meyer, B., Olivier, P. (Eds.), *Diagrammatic Representation and Reasoning*. Springer, London, pp. 173–185.
- Sheehan, K., O'Reilly, T., 2012. The case for scenario-based assessments of reading competency. In: *Reaching an Understanding: Innovations in How We View Reading Assessment*, pp. 19–33.
- Simon, H.A., 1978. On the forms of mental representation. In: Savage, C.W. (Ed.), *Minnesota Studies in the Philosophy of Science, Perception and Cognition: Issues in the Foundations of Psychology*, vol. ix. University of Minnesota Press, Minneapolis.
- Sireci, S.G., Zenisky, A.L., 2011. Innovative item formats in computer-based testing: in pursuit of improved construct representation. In: Haladyna, T.M., Downing, S.M. (Eds.), *Handbook of Test Development*. Routledge, New York, pp. 343–362.
- Sireci, S.G., Zenisky, A.L., 2016. Computerized innovative item formats: achievement and credentialing. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Routledge, New York, pp. 313–334.
- Someren, M.W., Barnard, Y.F., Sandberg, J.A.C., 1994. *The Think Aloud Method. A Practical Guide to Modeling Cognitive Processes*. Academic Press, London.
- Strobel, B., Saß, S., Lindner, M.A., Köller, O., 2016. Do graph readers prefer the graph type most suited to a given task? Insights from eye tracking. *J. Eye Mov. Res.* 9, 1–15.
- Strobel, B., Lindner, M.A., Saß, S., Köller, O., 2018. Task-irrelevant data impair processing of graph reading tasks: an eye tracking study. *Learn. InStruct.* 55, 139–147.
- Strobel, B., Grund, S., Lindner, M.A., 2019. Do seductive details do their damage in the context of graph comprehension? Insights from eye movements. *Appl. Cognit. Psychol.* 33, 95–108.
- Sweller, J., Ayres, P., Kalyuga, S., 2011. *Cognitive load theory in perspective*. In: *Cognitive Load Theory*. Springer, New York, pp. 237–242.
- Sweller, J., 2020. Cognitive load theory and educational technology. *Educ. Technol. Res. Dev.* 68, 1–16.
- Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., von der Heydt, R., 2012. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychol. Bull.* 138 (6), 1172.
- Wiley, J., 2019. Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Appl. Cognit. Psychol.* 33 (1), 9–19.
- Wirzberger, M., Borst, J.P., Krems, J.F., Rey, G.D., 2020. Memory-related cognitive load effects in an interrupted learning task: a model-based explanation. *Trends Neurosci. Educ.* 20, 100139.
- Wise, S.L., Bhola, D.S., Yang, S.T., 2006. Taking the time to improve the validity of low-stakes tests: the effort-monitoring CBT. *Educ. Meas.* 25 (2), 21–30.
- Wise, S.L., Pastor, D.A., Kong, X.J., 2009. Correlates of rapid-guessing behavior in low-stakes testing: implications for test development and measurement practice. *Appl. Meas. Educ.* 22, 185–205.
- Wise, S.L., Kuhfeld, M.R., Soland, J., 2019. The effects of effort monitoring with proctor notification on test-taking engagement, test performance, and validity. *Appl. Meas. Educ.* 32 (2), 183–192.
- Wise, S.L., 2017. Rapid-guessing behavior: its identification, interpretation, and implications. *Educ. Meas.* 36 (4), 52–61.
- Xi, X., 2005. Do visual chunks and planning impact performance on the graph description task in the SPEAK exam? *Lang. Test.* 22 (4), 463–508.
- Xi, X., 2010. Aspects of performance on line graph description tasks: influenced by graph familiarity and different task features. *Lang. Test.* 27 (1), 73–100.
- Zapata-Rivera, D., Arslan, B., 2021. July. Enhancing personalization by integrating top-down and bottom-up approaches to learner modeling. In: Sottillare, R.A., Schwarz, J. (Eds.), *Adaptive Instructional Systems. Adaptation Strategies and Methods*. Springer, pp. 234–246.
- Zhang, M., van Rijn, P.W., Deane, P., Bennett, R.E., 2019. Scenario-based assessments in writing: an experimental study. *Educ. Assess.* 24 (2), 73–90.

Common item formats and evaluation practices

Doug Baldwin, Senior Strategic Advisor, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	185
Item types in the context of test design	185
General principles in selecting item types	186
Selected-response items—some advantages and disadvantages	187
Constructed-response items—some advantages and disadvantages	187
Principles for developing items	188
Developing common selected-response item types	189
Developing and scoring constructed-response item types	190
Other item types	191
References	191

After more than 10 years as a teacher and 25 years as a test developer, one lesson I have learned again and again: It's extremely easy to write a bad test question.

—Overheard in the hallway at a testing company.

Introduction

This article focuses on one facet of overall test development. It presents an overview on the nature of traditional item types typically used in educational measurement, and it describes the advantages and disadvantages each item type contributes to the test design, and, therefore, to test use and to test validity, reliability, and fairness.

For the purposes of this discussion, 'item' is used to describe all versions of what are typically known as test questions. Item as a general term here includes everything from individual discrete true-false questions, to multiple-select multiple-choice questions, to sets of items related to a single stimulus (e.g., a reading passage or a video clip), to prompts (e.g., essay questions), and to larger tasks (e.g., more complex sets of interrelated questions, often including test takers doing some kind of performance). This article also looks briefly at non-traditional item types, including technology-enhanced items and, as often used in noncognitive assessments, Likert-scale item types.

The larger context and guiding principle for this discussion is overall test development: "Test development or test construction refers to the science and art of planning, preparing, administering, scoring, statistically analyzing, and reporting results of tests" (Downing, 2010, p 159). (For especially useful overviews of test design and development, see Schmeiser and Welch, 2006; Downing, 2010; Odendahl, 2011; Lane et al., 2016.) That is, the writing of appropriate types of high-quality items directly contributes to the three guiding goals for all assessments: validity, fairness, and reliability, about which more will be said as we look at how these goals are reached during the process of item development. Although there are references to basic statistics (i.e., at the level of how test data informs item writing and reviewing), the scope of this article does not include any detailed discussion of psychometrics, information covered elsewhere in the encyclopedia. The focus of this discussion is on specific item types used in tests that are primarily designed to assess general cognitive skills and/or specific subject content in either academic settings (e.g., summative or formative classroom-based; admissions; student learning outcomes; large-scale assessments of state or national learning progression) or in career settings (e.g., screening for hiring; promotion within a corporation).

One major overall aspect of developing and administering an assessment is that the process is iterative, not linear, and, reflecting that process, this discussion is also iterative. That is, for instance, the text describes item types, then returns to discuss the major steps in item development, and then returns to discussing those steps as they function in different test design contexts, and then returns to discussing how they are evaluated in operational settings, and so on. The ongoing focus, however, is on the different kinds of item types most often seen in assessments.

Item types in the context of test design

The initial decisions around determining which item types are most appropriate for a given assessment evolve from the early stages of test design, which, in turn, are typically informed by test designers following procedures of Evidence-Centered Design (ECD) (For detailed discussions of ECD, see Mislevy et al., 2003; Mislevy and Riconscente, 2006; Zieky, 2014; Riconscente et al., 2016).

In brief, the process of ECD helps test developers—often working with external committees of subject-matter experts (SMEs)—systematically determine, and document rationales for, such test design decisions as precisely defining, among many other steps in the process, at least all of the following:

- the assessment's construct (or domain: the specific knowledge, skills, or other attributes [KSAs] to be measured for the clearly specified test-taking population),
- the timing,
- the number of items in a given test form,
- the item formats (e.g., the number of options for multiple-choice items; the descriptors in scoring guides, or rubrics, for constructed-response prompts),
- the test blueprint or test specs,
- the specific intended use of the test,
- whether the test will be norm-referenced or criterion-referenced (and, if the latter, possible cut scores),
- the initial high-level design of the likely test report,
- how the test will be made accessible,
- in what mode(s) the test will be administered (e.g., paper-and-pencil and/or via computer),
- what psychometric data will be collected,
- if/how different forms of the test will be statistically equated, and
- the intended test use.

At each stage, participants in the ECD process collect clear and definable rationales for every decision, linking specific test taker performance as evidence (specific claims defining that evidence) of their KSAs—the foundation to evaluating the assessment's validity—knowing, however, that the overall test development process is inevitably recursive and that many of these decisions will be revisited and emended during the entire process of developing and field testing early tryout versions of the assessment. (For useful overviews on general developments in assessment practices, see [Odendahl, 2011](#), and [Bennett and von Davier, 2017](#).)

General principles in selecting item types

As previously suggested, the ECD process, in brief, includes defining the following: the domain model (what KSAs are being measured); the claims (what we can say about test takers); the task models (precisely what we ask test takers to do to demonstrate their knowledge of the construct's KSAs); the evidence models (what we observe to substantiate those claims); and the overall scoring models (how items will be scored and how scores will be aggregated across items and what scores will be reported). From this foundational design work, item writers proceed to compose specific items that the test developers and the external SMEs think will best measure the KSAs of the assessment's intended construct.

Throughout the twentieth century, in large-scale standardized academic assessments, the most common item type has been the multiple-choice item. But as [Odendahl \(2011, pp 1–17\)](#) points out, for centuries back to early screening assessments for civil service employment in China during the Han Dynasty (202 BCE – 220 CE), the main item type was the essay, variations of which continued to be delivered well into the twentieth century. However, in the mid-twentieth century, with technical improvements in capturing and scanning answer sheets, multiple-choice was seen as being more objective and, given efficiencies, more cost-effective, while essay responses were viewed as a less reliable (and more subjective) way to assess skills or content. Many academic assessments include a large set of multiple-choice questions to sample students' knowledge in a broad domain (e.g., biology) and a few constructed-response questions to assess the students' ability to apply that knowledge (e.g., design a controlled experiment or analyze data and draw a conclusion).

'Item type' is defined by the specific response method used by a test taker. Response methods can include, for instance, selecting one or more options (e.g., multiple-choice and multiple-response multiple-choice, matching, re-ordering); selecting a cell in a grid; completing a graph; writing a short answer to a specific question; filling in a one- or two-word or more answer; composing a brief response; composing an extended response; performing one or more tasks; and so forth. Each item type can assess a wide variety of KSAs, and each item type can have a wide variety of components, as will be discussed.

Although multiple-choice is the most common (and well-known) item type of the last hundred years, it is but one of many possible item types, and even the multiple-choice item type itself comes in variety of flavors. But the first, high-level, decision about selecting which item type(s) to use in an assessment is, broadly speaking, choosing between the two main types of items: selected-response (SR) items (in which the test taker selects from a list of possible answers) and constructed-response (CR) items (in which the test taker constructs, or produces, an answer or performs one or more tasks). In brief, whereas selected response items allow reliable and economical measurement of a wide range of content in a short testing time, constructed-response items allow students to provide evidence of their ability to develop an argument, explain their reasoning with examples or data, draw graphs, or otherwise produce rather than merely recognize an answer.

Before looking more closely at each of these two main item types, we need to look at the specific primary advantages and disadvantages of each. Consider the standardized test to qualify for a state driver's license permit: there is usually a multiple-choice section consisting of some number of discrete items designed to assess whether the test taker is familiar with the basic rules of driving and traffic laws as outlined in a state department of motor vehicles manual (e.g., "Which of the following is indicated by a broken white line?"). However, even a perfect score on such a multiple-choice test section does not begin to indicate whether the potential driver has even rudimentary actual driving skills, so there is a second test section, a road test, in which the test taker has to demonstrate (by performing) an actual ability to drive on local roads, merge into moving traffic, parallel park, and so forth.

This combination of the two primary item types is analogous to traditional writing assessments: many tests combine one section of (SR) questions (an 'indirect' measure of writing ability) along with one or two (CR) prompts, each requiring the test taker to write an essay (a 'direct' measure of writing ability, which is often described as 'authentic' assessment). Why do so many writing assessments require both, instead of relying on just the multiple-choice or relying on just the essay(s)? The answer, as will be seen, helps delineate the advantages and disadvantages of each of these two major item types. (For a useful discussion of some of the reasons test designs might use CR along with, or instead of, SR item types, see [Livingston, 2009](#).)

Note that many innovative assessments (e.g., situational judgment tests) can use either SR or CR item types. That is, the item types often reflect, but they do not necessarily determine, the developments in new kinds of innovative assessments. The same affordances for each hold true: for instance, on one recently developed assessment of confirmation bias (designed to assess how willing test takers are to look at alternatives other than the ones they had previously chosen at random), the SR items measure the degree to which test takers demonstrate confirmation bias, but the CR responses reveal the test takers' rationales for their choices. Another recent assessment presents video performances dramatizing various kinds of potential conflicts in the work environment; some of these stimuli are followed by multiple-choice options, which quickly measure the test takers' understanding of, for instance, teamwork, while other stimuli are followed by questions asking the test takers to describe what they would do in that situation, with these latter item types helping measure the test takers' ability to think on their own of good ways to demonstrate teamwork while also capturing their rationales for given actions.

Selected-response items—some advantages and disadvantages

In brief, SR items are simple, versatile, effective, efficient, and accessible. They are relatively inexpensive to write, and they are relatively inexpensive to score. They can be easily delivered in both paper-and-paper and computer-based modes. They take test takers relatively little time to answer. For instance, some recent college admissions assessments had approximately 150 Reading, Writing, and Math multiple-choice items to be answered in 3 h, or an average of a little over a minute to answer each question, with per-item time on different sections ranging from 48 to 87 s; meanwhile, test takers had 50 min to respond to the single optional essay prompt. Given the large number of SR items that can be included in a relatively short time frame, SR items can efficiently assess a large portion of the intended construct, and, for the same reason, the SR items can provide higher reliability, including allowing for reliable subscores. That is, the power of SR assessments is not in any given item, but, rather, in the aggregate of a collection of items that cover a large portion of the construct.

However, SR items measure recognition, not actual test taker production. They do not reveal test takers' thinking processes. They allow for random guessing. They are often decried in anti-testing commentaries: educators in general do not trust them as accurate measurements of student abilities, and clients often see them as untrustworthy measures of educational attainment.

Constructed-response items—some advantages and disadvantages

In contrast, CR prompts require test takers to actively demonstrate their knowledge, rather than merely recognize and select correct answers. At the same time, including CR items that ask for written responses, for example, on an assessment means that test takers prepare by doing more actual writing and that their teachers include more writing in their instruction and classroom assignments. This, in turn, leads to greater face value in the educational community, and the 'washback effect' of encouraging test takers to do more writing has significant pedagogical value, as does the fact that CR items can provide real-world validity, can reveal the test takers' thought process and reasoning, and can assess test takers' ability to organize and clearly present complex information. The same general advantages apply to other, non-written, CR item types (e.g., those requiring an oral response or some other performative demonstration of the test taker's KSA).

Note, however, that some CR item types can in fact require test takers to write very little and in that sense can approximate what can be assessed more efficiently and inexpensively by SR items: such item types as fill-ins (requiring just one or just a very few words) or focused questions that require very short answers (one or just a few sentences) have many of the advantages of SR items, except that in general SR items are much more efficient to score and are still able to roughly assess the same KSAs, so the CR versions are not commonly used. (As AI scoring engines become more sophisticated, the difference in efficiency will inevitably begin to disappear.) Note that, for instance, some assessments in England have traditionally used CRs exclusively, including fill-in items, and these item types do not assess the KSAs any more accurately than do multiple-choice items, and they also require hand 'marking' by raters, which scoring itself can be unreliable. In contrast, many assessments in the United States require test takers to generate answers, for instance specific numbers on a math test, but the entries are still marked on scannable answer sheets so that the scoring can be done by machines, and not by human raters, thus combining the CR advantage of asking test takers to produce, not just recognize, correct answers with the SR advantage of fast and accurate scoring. Some writing assessments include 'sentence construction shift' SR items—here test takers must mentally rewrite the sentence according to specific criteria in order to choose the correct option that maintains the original sentence's meaning, which is another kind of blend of the properties of both CR and SR item types.

Because many CRs typically require more time to answer (especially essay prompts requiring anywhere from one paragraph to a multi-page response), fewer CR items can be included on a typical assessment, which thus leads to less coverage of the construct, and with fewer items on a test form, the result is lower reliability (without even taking into account the exacerbating challenge of maintaining high rater consistency during the scoring, another contributor to lower test reliability). It is also harder to ensure

comparability of different CR prompts, and therefore it is harder to equate different forms of the assessment. Although CR prompts can be relatively easy to develop, they require significantly more time and expense to score than do SR items. (Note that assessments with CR items that require written or oral responses are more and more relying on computer software for the scoring of the word-processed or spoken and recorded responses; use of AI scoring engines can significantly reduce the time and expense associated with all-human-scored assessments.) Finally, CR prompts are easier to memorize, and thus they can make assessments less secure. (For extended discussions of the methods, complexities, and challenges in developing and scoring CR and other performance assessments, see Baldwin et al., 2005; McClellan, 2010; Baldwin, 2012.)

It is important to note that it is misleading to claim that SRs are more objective than CRs. While it is definitely true that scoring SRs is generally objective and accurate in comparison to the potentially subjective human scoring (or even computer software trained on human scoring) of a response to a CR item type, an SR item can in fact have an objectively scored but in fact subjective answer (e.g., “Which of the following cities is the most fun for a traveler to visit?”), while a CR that asks a very specific question can be purely objective (e.g., “State the three primary examples the author presents to support his argument.”). Similarly, it is also misleading to claim that by their nature CR item types require higher-level thinking skills; CR items often do so, but so do well-crafted SR items.

Given the advantages and disadvantages of SR and CR item types, many assessments compromise by including some combination of both, taking into account the purposes, desired time allotment, budgeted costs, test development resources, and other aspects of the planned assessment. (For a history of writing assessment in the twentieth century and the decisions to use or not use essay prompts, see Elliot, 2005; for a survey of more recent practices in writing assessment, see Baldwin, 2020.) One general approach taken in some elementary and secondary school systems in the United States is to use constructed-response items in the classroom as formative assessment while relying on selected-response items for large-scale standardized summative assessments, especially when reporting at the institutional level.

Principles for developing items

Before looking in more detail at the two main item types, it is important to note some guiding principles that inform all item development.

As previously noted, the types of items are determined during the ECD process. For instance, typical academic summative assessments for elementary and secondary school students in the United States are often directly aligned to specific standards. Each state develops its own standards in order to define the specific construct for each subject at each grade level. That is, the standards define the KSAs that the students are expected to exhibit at the end of the academic year. Therefore, the SMEs and test developers who are designing and developing the assessments will develop test specifications that directly reflect those standards. Often the test design will specify that each item be directly linked to a specific and identified standard; at the same time, the test specifications will align the difficulty of each item to the appropriate level. The intent is to represent the range of item difficulty across the entire test. Item difficulty represents the item’s complexity, which is often based on the level of abstraction on Bloom’s taxonomy (1956) and/or at the correct level of the four levels of Webb’s ‘depth of knowledge’ (1997). These decisions about the test design are also closely related to whether the assessment is norm-referenced or whether it is criterion-referenced; the latter is more common for assessments at the elementary and secondary levels. Assessments may be aligned to specific curricula, and some curricula are defined simultaneously with the intended assessments. The same decisions are made in terms of specifying the cognitive load (and ‘working memory’) of the item types and the specific items. Often during the ECD process, item templates are created for each of the test’s item types, templates that specify the kinds of stimuli (e.g., length; lexical complexity; whether based on actual texts, or written specifically for the assessment), the kinds of content, and the kinds of options that the test designers feel will best assess the test takers’ KSAs for a given construct. These templates will also indicate, for instance, if the SR items will be primarily ‘discrete’ items (i.e., not part of a set of items) or if, instead, there will mostly be sets of items based on a single stimulus, or whether the assessment will include a combination of both kinds of items. Similarly, the overall test design will specify, for instance, whether there are composite items made up of different parts that together lead to a single overall score.

Meanwhile, for many tests (licensing, admissions, job screening), test takers are typically motivated and will do their best on any items; in contrast, test takers are often less motivated for tests where the stakes are low (e.g., large-scale assessments such as PISA, TIMSS, or NAEP, which do not report individual test taker scores but, rather, are used to assess progress across large populations of students). Low test taker motivation means that item writers must try to generate items that are more intrinsically of interest to the test taker population, and the item templates will reflect that goal. Dealing with low test-taker motivation can be especially challenging when using CR item types, although here at least test developers do not need to determine the motivation effect, since test takers cannot just randomly guess.

All item writers and reviewers must have deep understanding and knowledge of the requisite content or skill that is being assessed, and they must be carefully and extensively trained in all aspects of item writing, including, for instance, fairness, accessibility, and appropriate editorial style. The item writers might be any combination of professional test developers, trained outside freelance item writers, and external committee members (e.g., college professors in a specified discipline, often termed ‘subject matter experts,’ or SMEs). More recently, some testing companies have tried various ways to clone items, and some have even begun using automated item generation (Gierl, Lai, 2016).

Items for most assessments go through an extensive and rigorous set of review steps, including, for instance: initial item writing; two or three separate content reviews; separate formal fairness and separate formal edit reviews; and a final ‘fresh eyes’ review. (For outlines of fairness as an essential standard for test development, see [American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014](#) and [Educational Testing Service, 2014](#). For a discussion of achieving fairness in assessments, and for a discussion of the kinds of sensitivity guidelines reviewers follow when reviewing items to ensure that there are no construct-irrelevant factors that could impact test takers, see [Camilli, 2006](#), and [Zieky, 2016](#).) At each step, all review comments are then resolved with the ‘owner’ of the previous step. Typically, completed items are collected in some kind of item pool or item banking system. When the different items are assembled into a test, following psychometric test specifications, the entire form often receives an overall final review, which includes insuring that the overall group of items fits the test content specifications, that there’s no unintended duplication of item content, and that no items ‘cue,’ or provide hints to, other items.

Items are typically tried out in different contexts before they are administered in an operational form. Test developers will often administer new item types in ‘cog labs,’ where a representative subpopulation of the intended test taking population will answer the try-out items. These test takers will then provide written and/or oral feedback: how they interpreted the items, what if any possible difficulties they had understanding what was being asked, and what cognitive processes they followed as they tried to answer the items. These cog labs help test developers ensure that the items are functioning as intended, i.e., that the items are measuring the intended construct.

For new assessments, test forms made of new items are often administered to small groups of test takers representative of the intended test-taking population, often as separate ‘field tests,’ so that psychometricians can generate statistical data before the items are assembled into final forms. The data in turn can identify items that are not performing as intended; can generate accurate statistics in order to be able to equate different forms; and can individually flag possibly problematic items that appear to be unfair to specific subgroups, termed differential item functioning (DIF). (For a discussion of DIF analyses, see [Camilli, 2006](#); [Zwick, 2012](#); [Zieky, 2016](#), pp 93–94.). Many well-established programs will embed pretest items into operational forms to determine item performance, although this practice is limited almost exclusively to SR items—CR items are often field tested separately in order to evaluate how potential test takers respond to a given prompt, but they are rarely if ever embedded into operational forms.

Developing common selected-response item types

The guiding principles described in [Principles for Developing Items](#) section inform both SR and CR item development, but there are processes specific to SR that are discussed in this section, with the focus being on multiple-choice items, the most common SR item type.

Whether a discrete item or an item in a set of items based on a single stimulus, the most common multiple-choice item type has a structure that typically includes the following:

- A stem, or the question, which can be open-ended (the answer completes a sentence, along the lines of, “Mixing sodium with water produces _____”) or closed (the stem ends typically with a question mark, along the lines of, “Which of the following is an example of a metaphor?”).
- Three or four or five options, with four being most common; however, note that [Rodriguez \(2005\)](#) and [Rodriguez and Haladyna \(2013\)](#) argue that three options are optimal.
- One key (the intended answer), with the other options being distractors. Note: Distractors that are intentionally written to represent specific types of common misunderstandings of the construct can be used to provide diagnostic feedback to the test taker and/or to the test taker’s teacher. Note that there are also multiple-select multiple-choice items (e.g., a test taker is required to indicate two answers from among six options—the test design would determine whether there is partial-credit scoring or whether credit is given only for test takers who mark both intended keys).
- SR items, discretized parts of sets, are often associated with a stimulus (a passage, chart, graph, map, audio or video clip, or other piece of work that provides the basis for the test questions).

When developing SR items, the goal is to have a single best answer, one that is defensible and that is directly linked to the KSAs defined as the construct being assessed and that is at the appropriate level of difficulty. There should be no errors in content or in grammar and usage. Wording throughout should be precise, concise, and accurate. The stem and key should reflect a clearly defined problem or task, and the stem should be focused specifically on a specific KSA from the construct. The options should be parallel, and they should be roughly the same length. The ‘distractors’ (options that are not the intended key) should be plausible, but they cannot be arguably correct responses. That is, there cannot be more correct answers than the test taker is asked to identify (some SR items, known as Multiple-Select Multiple-Choice, ask test takers to identify two or more correct answers). Distractors should also be mutually exclusive, or test takers can eliminate them as possible keys. Distractors should be wrong for non-trivial reasons. In general, more attractive distractors lead to more difficult items. “All the above” and “None of the above” are in general not acceptable options. To the degree possible, item writers should avoid using negative stems, but, if it’s deemed necessary to use them, the item writers should emphasize the negative word (“Which of the following is NOT an example of hyperbole...” “All of the following are examples of inert gases EXCEPT...”) so that test takers aren’t easily misled by construct-irrelevant phrasing of the stem. In general, item writers try to avoid anything possibly tricky or misleading in the language used in the item: if

test takers get a question wrong, it should be because they lack the required skill or knowledge, not because they have difficulty understanding the question or the options. Any accompanying graphs or images should be clear and appropriate. (For more detailed discussions of best practices in SR item writing, see [Roid and Haladyna, 1982](#); [Haladyna, 2004](#); [Schmeiser and Welch, 2006](#); [Downing, 2006a, 2006b, 2010](#); [Rodriguez, 2016](#).)

Developing and scoring constructed-response item types

Best practices in writing and reviewing CR prompts parallel those used in developing SR items.

However, the process of developing and scoring CR items includes the need to address challenges specific to CR item types.

CR (and other performance) item types range from entering numbers on a scannable answer sheet to fill-in-the-blank sentence completions to short paragraph responses to multi-page essays to in-class projects presented over two or more days. Assessments can require test takers to submit portfolios, which can include not just essays but also, for instance, video recorded performances. One alternative college admissions assessment required test takers to participate in team-based efforts to replicate a Lego structure; to participate in a group discussion; and to be interviewed. An American Sign Language assessment consisted of interactive interviewing of each test taker, with the ASL responses scored by two independent raters. A student learning outcomes assessment for college freshmen was intended to measure test takers' collaborative problem solving skills: the test required test takers to coordinate online with two other anonymous test takers in order to reach a group consensus, with raters evaluating the teamwork skills demonstrated by the test takers' online communications. Another college outcomes assessment asked test takers to present two directly opposing perspectives on a real-life topic. All of these use CR item types. (For a useful overview of performance assessments, see [Lane and Stone, 2006](#).)

But, in general, CR item types most typically require test takers to write essays of varying lengths. Note that although test takers increasingly use word processing software to generate their responses, many CR assessments are still given in a paper-and-pencil format. The items are typically used to assess test takers' writing and reasoning skills or their knowledge of specific content. For the former, prompts may ask test takers to develop a position on an issue or analyze a piece of writing; for the latter, test takers may, for instance, analyze historical documents, explain a scientific principle, or synthesize information from several passages. Stimuli for CR essay prompts can range from very brief one-sentence questions to two or more developed passages which test takers are asked to synthesize and reference in their responses. (For a useful overview of developing constructed-response and other performance assessments, see [Welch, 2006](#), and [Lane and Iwatani, 2016](#).)

SR item types can generally be scored by machine; CR items require raters to read, view, or listen to test takers' responses and to assign a score to each response. Typically, essays are scored on a 1–4 or 1–6 scale, using the sets of descriptors from a scoring guide (or rubric) that defines the typical qualities of a response at each score point along with exemplary responses at each of the score points, though many CR item types used in K12 content assessments are scored on a 0–2 or 0–3 scale. Some assessments have had much larger scales (e.g., 1–19). In general, the principle is that there should be as many score points as raters are able to consistently and accurately distinguish.

The two most common approaches to scoring CR prompts are with either holistic or with analytic rubrics. In holistic scoring, raters assign a single, overall score that reflects the overall quality of the response (it was originally called 'wholistically'). Holistic scoring has the advantage of being efficient; it provides a useful overview of student achievement. Analytic scoring (often called 'trait' or 'dimension' scoring) is increasingly common; this has been especially true for assessments of elementary and secondary school students in the United States. Here, raters assign separate scores for each of several traits, or dimensions (e.g., "Analysis & Argumentation," "Organization & Logic," and "Grammar & Mechanics" each scored on a separate scale, sometimes with traits being reported as subscores and also having different weights when combined for an overall reported score). Trait scoring can provide more detailed feedback, or diagnostic information, to test takers and their teachers (especially useful for assessments of content), but the scoring effort is significantly more time consuming. With either approach, large-scale assessments with CR item types have organized and standardized scoring procedures in order to ensure that all raters are using the same criteria when assigning scores.

When developing CR items, test developers will often themselves write out responses to prompts in order to evaluate how difficult it might be to generate a reasonable response in a reasonable time frame. Newly developed CR items are typically field tested in order for the prompts to be evaluated based on criteria reflecting, for instance, such questions as how well test takers seem to be able to engage the topic, whether responses are unfairly rewarded for outside knowledge that is not part of the intended construct, and how accurately and consistently raters are able to apply the scoring guide. During the scoring of the field test responses, test developers usually also select benchmarks, papers that are clear examples of what is representative of the qualities at each score point. In the early stages of the development of a new CR assessment, before it is administered operationally, the scoring guides are often changed to reflect what raters encounter while reading a range of test takers' responses. Occasionally changes are made to the time allotted for each task during the iterative process of test design and field testing of prompts, items, and entire test forms. Some programs will develop generic scoring guides for a large pool of prompts but will then also develop prompt-specific guides or notes for each individual prompt in order to help raters reach a high interrater reliability (i.e., the degree to which different raters independently give the same scores as do the other raters). Typically, during the scoring of operational CR prompts, essay responses receive two independent scores from two different raters, and if the two scores vary by more than a predetermined amount, the response typically goes to a third (often more experienced) rater to determine the final, reported score. During scoring, trained raters

are monitored by senior and more experienced scoring leaders. All efforts are made to ensure that the scoring is valid, reliable, and fair.

Although there are few assessments that rely exclusively on automated scoring engines, more and more scoring of CR item types is done with at least one ‘reading’ performed by computer software. Computational linguists have developed scoring engines for assessing writing, specific content, and math. While the initial development costs can be high, these scoring programs, once ready for operational use, are significantly less expensive and of course significantly more efficient than human scoring. (For further discussions of automated scoring of CR assessments, see [Ericsson and Haswell, 2006](#), and [Shermis et al., 2016](#). For further discussions of both human and AI scoring of CR item types, see [Educational Testing Service, 2021](#).)

Other item types

Technology, specifically the increased capabilities of computers (along with the creativity of task designers and item writers), has not only changed how CR item types are scored, but it has also led to a greatly increased use of technology-enhanced items (TEIs). TEIs rely on test takers being able to use computers to produce their responses (for either SR or CR item types, although the latter is far more common). Early and less technologically demanding TEIs included, for instance, item types such as asking test takers to use the computer’s mouse to highlight the appropriate sentence in a reading passage. More recent TEIs have included such item types as inline choice, grids, fill-ins, graphs, ordering events, and drag-and-drop. TEIs can be more engaging for test takers, can use realistic simulations, and can capture more aspects of a response than traditional items. For instance, drag-and-drop items allow test takers to construct an appropriate sequence rather than just recognize one ([Arslan et al., 2020](#)). In that sense, the drag-and-drop item type shares some of the advantages of CR item types, but it can be machine scored like SR items. The new item types and their many variations have given task designers more tools to use, but they have also increased the amount of information that task designers need to make appropriate decisions about which item types and features to use. TEIs have disadvantages as well as advantages and should not be used merely to add variety or novelty to a test without regard to their impact on the quality of the construct-relevant evidence provided by the test; accessibility can be a special challenge with TEIs. (For discussions for and against the argument that TEIs have the potential to better represent knowledge, skills, and abilities of students in a more authentic way compared to traditional multiple-choice items, see [Bryant, 2017](#), and [Scalise, 2012](#)).

Also worth mentioning is the Likert-type item. This item type is widely used in ‘noncognitive’ assessments of, for instance, test takers’ physical and mental health, attitudes, opinions, and other traits. Although common in anonymous surveys, this item type is also found in assessments designed, for instance, to measure an individual test taker’s personality. A Likert-type item asks test takers to rate themselves on some attribute, using a scale (e.g., numeric, as in 1–5 [from “Strongly Disagree” to “Strongly Agree”], or by descriptors, as in “Never,” “Sometimes,” “Always”). (For a brief but useful overview of this item type, see [McLeod, 2019](#).) However, such self-ratings can be misleading in that test takers may well intentionally select not what they actually think is an accurate self-description but, instead, what they think the employer or academic program wants to see (i.e., putting themselves in a positive light, or answering based not on their actual self-evaluation but instead on social desirability). As [Kyllonen \(2016\)](#) notes, “The forced-choice response format was designed to address the faking problem” in self-reporting (p. 201), since test takers have to select which of two similarly socially positive or similarly socially negative Likert-type statements best describes them (e.g., having to select between the two statements: “In mathematics I always try to do better than the other students in my class” and “In mathematics, I enjoy helping others to work well in a group”). Over the last two decades there has been a greatly increased use of these kinds of assessments used to measure personality traits ([Kyllonen, 2016](#)), primarily in the context of careers, but more and more also in the academic arena. Employers and academic programs are interested in candidates’ strengths and weaknesses in such psychological areas as motivation, ‘grit’ (or resiliency), creativity, leadership, and teamwork.

References

- American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Arslan, B., Jiang, Y., Keehner, M., et al., 2020. The effect of drag-and-drop item features on test-taker performance and response strategies. *Educ. Meas.* 39 (2), 96–106.
- Baldwin, D., 2012. Fundamental challenges in developing and scoring constructed-response assessments. In: Elliot, N., Perelman, L. (Eds.), *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Hampton, New York, pp. 327–341.
- Baldwin, D., 2020. “The times, they are a-changin’”: reflections on the evolution of research and policy in large-scale writing assessments. In: Elliot, N., Horning, A.S. (Eds.), *Talking Back: Senior Scholars and Their Colleagues Deliberate the Past, Present, and Future of Writing Studies*. Utah State UP, Louisville, CO, pp. 56–72.
- Baldwin, D., Fowles, M., Livingston, S., 2005. Guidelines for Constructed-Response and Other Performance Assessments. Educational Testing Service, Princeton, NJ.
- Bennett, R.E., von Davier, M., 2017. *Advancing Human Assessment: The Methodological, Psychological and Policy Contributions of ETS*. Springer Nature, Cham.
- Bloom, B.S., 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. David McKay Company, Inc, New York, NY.
- Bryant, W., 2017. Developing a strategy for using technology-enhanced items in large-scale standardized tests. *Practical Assess. Res. Eval.* 22 (1), 1–10.
- Camilli, G., 2006. Test fairness. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. Praeger, Westport, CT, pp. 221–256.
- Downing, S.M., 2006a. Selected-response item formats in test development. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 287–301.
- Downing, S.M., 2006b. Twelve steps for effective test development. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 3–25.

- Downing, S.M., 2010. Test development. In: International Encyclopedia of Education, third ed. <https://doi.org/10.1016/B978-0-08-044894-7.00270-0>
- Educational Testing Service, 2014. ETS Standards for Quality and Fairness. Educational Testing Service, Princeton, NJ.
- Educational Testing Service, 2021. Best Practices for Constructed-Response Scoring. Educational Testing Service, Princeton, NJ.
- Elliot, N., 2005. *On a Scale: A Social History of Writing Assessment in America*. Peter Lang, New York, NY.
- Ericsson, P.F., Haswell, R. (Eds.), 2006. Machine Scoring of Student Essays: Truth and Consequences. Utah State UP, Logan, UT.
- Gierl, M.J., Lai, H., 2016. Automatic item generation. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 410–429.
- Haladyna, T.M., 2004. *Developing and Validating Multiple-Choice Test Items*, third ed. Lawrence Erlbaum Associates, Mahwah, NJ.
- Kyllonen, P.S., 2016. Designing tests to measure personal attributes and noncognitive skills. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 190–211.
- Lane, S., Stone, C.A., 2006. Performance assessment. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. Praeger, Westport, CT, pp. 387–432.
- Lane, S., Iwatani, E., 2016. Design of performance assessments in education. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 274–293.
- Lane, S., Raymond, M.R., Haladyna, T.M., Downing, S.M., 2016. Test development process. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 3–18.
- Livingston, S.A., 2009. Constructed-response Test Questions: Why We Use Them; How We Score Them. ETS R&D Connections, Princeton, NJ.
- McClellan, C., 2010. Constructed-response Scoring – Doing it Right. R&D Connections, Princeton, NJ.
- McLeod, S.A., 2019. Likert scale. In: Simply Psychology. www.simplypsychology.org/likert-scale.html.
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A Brief Introduction to Evidence-Centered Design. Educational Testing Service, Princeton, NJ (Research Report 03-16).
- Mislevy, R.J., Riconscente, M.M., 2006. Evidence-centered design. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 61–90.
- Odendahl, N.V., 2011. *Testwise: Understanding Educational Assessment*. Rowman & Littlefield Education, Lanham, MD.
- Riconscente, M.M., Mislevy, R.J., Corrigan, S., 2016. Evidence-centered design. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 40–63.
- Rodriguez, M.C., 2005. Three options are optimal for multiple-choice items: a meta-analysis of 80 years of research. *Educ. Meas.* 24 (2), 3–13.
- Rodriguez, M.C., 2016. Selected-response item development. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 259–273.
- Rodriguez, M.C., Haladyna, T.M., 2013. Selected-response classroom assessment. In: McMillan, J.H. (Ed.), *SAGE Handbook of Research on Classroom Assessment*. SAGE, Thousand Oaks, CA, pp. 293–311.
- Roid, G.H., Haladyna, T.M., 1982. *A Technology for Test-Item Writing*. Academic Press, New York.
- Scalise, K., 2012. Using technology to assess hard-to-measure constructs in the common core state standards and to expand accessibility. In: ETS K-12 Center: International Research Symposium on Technology-Enhanced Assessments.
- Schmeiser, C.B., Welch, C.J., 2006. Test development. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Greenwood, New York, pp. 307–353.
- Shermis, M.D., Burstein, J., Brew, C., Higgins, D., Zechner, K., 2016. Recent innovations in machine scoring of student- and test taker-written and -spoken responses. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 335–354.
- Webb, N., 1997. Research Monograph Number 6: Criteria for Alignment of Expectations and Assessments on Mathematics and Science Education. CCSSO, Washington, DC.
- Welch, C., 2006. Item and prompt development in performance testing. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 303–328.
- Zieky, M., 2014. An introduction to the use of evidence-centered design in test development. *Psicol. Educ.* 20, 79–87.
- Zieky, M., 2016. Developing fair tests. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 81–99.
- Zwick, R., 2012. A Review of ETS Differential Item Functioning Assessment Procedures (ETS Research Rept. No. 12-08). ETS, Princeton, NJ.

Automatic item generation

Mark J. Gierl^a, Jinnie Shin^b, and Tahereh Firoozi^a, ^aCentre for Research in Applied Measurement and Evaluation, University of Alberta, Edmonton, AB, Canada; and ^bUniversity of Florida, Gainesville, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	193
Automatic item generation for test development	194
Template-based AIG	194
Example of template-based AIG	195
Evaluation of template-based AIG	196
Non-template-based AIG and example items	197
Evaluation of non-template-based AIG	198
Conclusion	199
References	199

Glossary

Automatic Item Generation A process of generating test items with the aid of computer technology.

Cognitive Model for AIG A representation that highlights the knowledge, skills, problem-solving processes and/or content a student requires to answer test items.

Item Model A template that highlights the features in an item that can be manipulated to generate new items.

Key Features Cognitive Model A model used to guide item generation based on the relationships among the key features specified in the cognitive model, which include the attributes or features of a task that are systematically combined to produce meaningful outcomes across the item feature set.

Logical Structures Cognitive Model A model used to guide item generation based on the students' ability to use a concept across a variety of different content representations. The concept is often used to implement a formula, algorithm, and/or logical outcome.

Semantic-Based AIG The semantic-based approaches use information acquired from a semantic analysis of the content in a corpus or knowledge base. Semantics focuses on the meaning of a sentence which is often expressed using a structured combination of words.

Sequence-Based AIG The sequential-based neural network approaches are used to generate test items automatically. These approaches focus on generating test items by disambiguating and learning the primary characteristics from the content in a corpus or knowledge base to create test items.

Syntax-Based AIG The syntax-based approaches are designed to generate items by uncovering the underlying syntactic structures of the content in a corpus or knowledge base. Syntax is a description of how words are combined in sentences, where the syntactic structure of a sentence conveys meaning.

Systematic Distractor Generation A method for generating distractors where specific information related to errors and misconceptions is used to create plausible but incorrect options.

Introduction

Changes in educational technology are profoundly affecting educational testing. Developers of international, national, and local educational tests are now implementing computer-based tests at an extraordinary rate. The world's most popular and visible achievement test—the Program for International Student Assessment (PISA) which was developed, administered, and analyzed by the Organization for Economic Cooperation and Development (OECD)—is computerized. While the first five cycles of the test administered were paper-based, 58 of the 72 participating countries in PISA 2015 took the computer-based version. The number of countries which opted for CBT increased to 89% (70 of the 79 participating countries) for PISA 2018 in keeping with the OECD view that computer-based testing has become part of the educational experience for most students (OECD, 2019). To cite another recent milestone: 2016 marked the first academic year that US states administered more computers than paper-based tests (He and Lao, 2018). Provincial agencies in Canada are also moving to computer-based testing. Alberta Education, for example, introduced computer-based tests in 2010, but limited its use to the Grades 6 and 9 achievement testing programs. Computer-based testing has now expanded to include all provincial tests including the achievement tests as well as the Grade 12 diploma examinations. These examples demonstrate that computer-based testing has quickly become the hallmark of 21st-century assessment and that testing

agencies are either administering their tests using a computer or migrating to this platform. Computer-based testing offers important benefits. For example, computers permit testing on-demand which means exams can be administered using a schedule that is flexible, frequent, and convenient. Computer-based testing allows students to take the test during instruction. Items on computer-based tests are scored immediately thereby providing students with feedback.

But the advent and growing popularity of computer-based testing has also raised daunting new challenges, particularly in the area of item development. Educators must have access to large numbers of diverse, high-quality test items to effectively implement computer-based testing because items are continuously administered to students. Hence, thousands of new items are needed to develop the test item banks necessary for computer-based testing. A bank is a repository of test items, which includes both the individual items and item characteristics. These banks must be initially created then constantly replenished to ensure that students receive a continuous supply of items while, at the same time, limiting item exposure. Unfortunately, educational test items, as they are currently created, are time consuming and expensive to develop because each individual item is written, initially, by a subject matter expert (SME) and, then, reviewed and revised by groups of SMEs to ensure the items are reliable and valid (Haladyna and Rodriguez, 2013; Lane et al., 2016). Hence, test item development is one of the most important problems that must be solved before educators can fully migrate to computer-based testing and realize the benefits of formative assessment.

Automatic item generation for test development

Automatic item generation (AIG) is a method for creating test content that can be used to address this item development challenge (Gierl and Haladyna, 2013; Irvine and Kyllonen, 2002). AIG is a rapidly evolving research area where cognitive and psychometric modeling practices guide the production of tests that include items generated with the aid of computer technology. AIG researchers and practitioners have adopted two different methods for generating test items—the first method relies on the use of templates to structure the testing content while the second is template free, relying, instead, on the use of advanced statistical techniques such as natural language processing and machine learning to structure the testing content. In this chapter we describe, illustrate, and evaluate these two methods.

Template-based AIG

Templates are commonly referred to as item models. An item model is a template that specifies the features in an item that can be manipulated to generate new items (LaDuca et al., 1986). Item modeling can be embedded in a three-step approach for generating test items (Gierl and Lai, 2013, 2016). In Step 1, the content for item generation is identified by the SME. This content is identified using design principles and guidelines that highlight the knowledge, skills, and abilities required to solve problems and perform tasks in a specific domain. A cognitive model for AIG is a summary that organizes the cognitive- and content-specific information into a structured representation of how the SME expects students will think about and solve problems. Two types of cognitive models are often used to generate educational test items (Gierl and Lai, 2017). The first is called the logical structures cognitive model. It is most suitable for measuring students' ability to use a concept across a variety of different content presentations. The concept is often used to implement a formula, algorithm, and/or logical outcome. Examples of logical structures cognitive models for AIG can be found in the content areas of science (Gierl and Lai, 2017; Gierl et al., 2016) and mathematics (Gierl and Lai, 2016; Gierl et al., 2015). The second is called the key features cognitive model. It is used when the attributes or features of a task are systematically combined to produce meaningful outcomes across the item feature set. This model is most suitable for measuring students' ability to assemble and apply key features within a domain as well as to solve problems using these key features. Examples of key features cognitive models for AIG can be found in the medical content areas of abdominal injury (Gierl et al., 2018), infection and pregnancy (Gierl and Lai, 2016), chest trauma (Gierl and Lai, 2018), hernia (Gierl and Lai, 2013), post-operative fever (Gierl et al., 2012), and jaundice (Gierl and Lai, 2017).

Cognitive models contain constraints. Constraints are rules specified by the SME that, in turn, are applied to the content in the cognitive model. Without the use of constraints, all the content in each part of the model would be combined to create new items. However, many of these combinations would not produce items that are sensible. Constraints serve as descriptions of conditions that must be applied during the assembly step (i.e., Step 3) so that meaningful items are generated. When the goal of item generation is to produce large banks containing many similar items, a relatively small number of constraints are used in the cognitive model. Alternatively, when the goal is to produce comparatively smaller banks of diverse items, a relatively large number of constraints are needed in the cognitive model. Hence, constraint coding is the method by which AIG captures the specific decisions, judgments, and rules used by SMEs to produce diverse, high-quality, content-specific test items.

In Step 2, an item model is developed to specify where the content from the cognitive model must be placed to generate new items. Item models provide the foundation for template-based AIG. An item model is a template created by the SME that specifies the features in an item that can be manipulated to generate new items (Bejar, 1996, 2002; Bejar et al., 2003; LaDuca et al., 1986). Item models also specify which parts of the assessment task can be manipulated for item generation. These parts include the stem, the options, and the auxiliary information. The stem is the part of an item which formulates context and/or content the student is required to answer. The options contain the alternative answers with one correct option and one or more incorrect options or distracters. Auxiliary information includes text, images, tables, diagrams, sound, and/or video, which are required to convey additional information required for examinees to answer the item. The stem and options can be divided further into elements. These elements

are denoted as non-numeric values (i.e., strings) and numeric values (i.e., integers). By systematically manipulating the elements, many items can be generated from one item model. For selected-response items like multiple choice questions, the stem and the complete list of options are required, including the correct option and distractors. For constructed-response items like short answer questions, only the stem and correct option are required to be specified in an item model.

One of the important differences between a constructed- and selected-response item is the outcome space for incorrect options or distractors. The incorrect options for a selected-response item are typically designed from a list of plausible but incorrect alternatives linked to misconceptions or errors in thinking, reasoning, and problem solving. Using a traditional item development approach, distractors are created by a SME that are specific to each item. But because AIG produces hundreds or thousands of items, an alternative approach is used to produce a correspondingly large number of plausible but erroneous distractors. This method is called systematic distractor generation (Lai et al., 2016). It is closely linked to the cognitive modeling step. Recall that to generate a stem and correct option, specific information must be collected in the form of a cognitive model (i.e., Step 1 in the three-step AIG process). Systematic distractor generation also involves specifying information related to errors and misconceptions in the form of a cognitive model in order to generate plausible but incorrect options. The process of modeling incorrect options is more complex compared to modeling the correct option because the item stem is constantly changing in a generative item development system. As a result, the small number of constraints required to ensure the information presented in the stem yields a correct response must be counterbalanced with a much larger number of constraints required to ensure the information presented in the incorrect options is erroneous but plausible. Also, more than one incorrect option is required for a selected-response item. Typically, three, sometimes four, incorrect options must be produced for each generated item.

In Step 3, a computer algorithm is used to place the cognitive model content specified in Step 1 into the item model developed in Step 2. Different types of software have been written to generate test items (e.g., Gierl et al., 2008; Gütl et al., 2011; Higgins, 2007; Higgins et al., 2005; Singley and Bennett, 2002). The most common approach for template-based AIG is to create an algorithm based on logical constraints. A logical constraints approach provides a straight-forward way to generate items using iterations where the generated content is specified as an element, as described earlier, and where each element contains all possible values to be displayed and substituted in each generated item. The presentation of the elements is organized in a cloze test format, where all values of an element can be displayed. Then, all combinations of the element values are assembled. With logical constraints, not all combinations of the values will produce meaningful test items. To prevent implausible combinations, the constraints defined by the SME in the cognitive model are used to limit the generated outcomes to those combinations that are deemed to be meaningful. The generation process can be described, conceptually, as an iterator that permutes through all combination of elements and, in the process, eliminates combinations (i.e., meaningless generated items) that do not meet the constraint requirements. In other words, once the content is specified in Step 1 and the template is created in Step 2, a logical constraints algorithm can be used to systematically place the content into the template in Step 3 to produce new items subject to the constraints in the cognitive model.

Example of template-based AIG

Because template-based AIG using item modeling is currently the preferred method for generating items, we present an elaborated example using content from the natural sciences. This example is drawn from the content area of botany where the student is required to identify a coniferous tree based on the characteristics of the tree. This example uses a key features cognitive model with systematic distractor generation.

In Step 1, a cognitive model for AIG is created. The model in this example was created inductively by a SME who was asked to identify and describe the information that could be used to differentiate coniferous trees. Three types of information are specified in the cognitive model for AIG. The first type of information is related to the main concept or problem and its associated scenarios. In our example, the problem is coniferous tree identification. Three associated scenarios include three different types of coniferous trees: Rocky Mountain Juniper, Limber Pine, and Tamarack. The second type of information specifies the relevant sources of information presented with the problem. In our example, four sources of information are presented: needle shape, needle group, needle length, and cones. The third type of information highlights the salient features, including the elements and constraints, for each source of information. The first component for a feature is the element. Elements contain content specific to each feature that can be manipulated for item generation. The second component for a feature is the constraint. Constraints serve as restrictions that must be applied during the generation process to ensure that content in the elements are combined in a meaningful way so useful items can be generated. For example, the needle shape feature includes three elements and three constraints where the constraints in this task are the correct option as shown in brackets: scaly, needle shaped leaves (Rocky Mountain Juniper), triangular needles (Limber Pine), and needles that are flattened above and keeled below (Tamarack).

In Step 2, the item model is created (see Table 1). The item model contains the template that will be used to generate test items. It also specifies the elements for the features. Four sources of information are included: needle shape, needle group, needle length, and cones. Each source of information contains unique elements: needle shaped leaves, triangular needles, and needles that are flattened above and keeled below (needle shape feature); opposite pairs along the branch, five-needle tufts, 15–60 need tufts (needle groups feature); 12 mm, 8–20 cm, 2–5 cm (needle length feature); berry-like seed cones, cylindrical seed cones, red and pink broadly ovoid seed cones (cone feature). The same logic is also used to produce the incorrect options using the systematic distractor generation method. Distractor 1 is Yellow Cedar. This coniferous tree shares some, but not all, of the elements with Rocky Mountain Pine. For instance, a Rocky Mountain Pine and a Yellow Cedar both contain scale-like needles. As a result, Yellow Cedar is a plausible but

Table 1 An item model for generating botany test items.

Stem	A first-year forestry student is hiking through Jasper National Park. He comes across a coniferous tree with <NEEDLE SHAPE> that are in <NEEDLE GROUPS> and measure <NEEDLE LENGTH> in length. Along the branches are <CONE>. These characteristics best describe which coniferous tree?
Elements	<NEEDLE SHAPE> 1. scaly, needle shaped leaves (RJ) 2. triangular needles (LP) 3. needles that are flattened above and keeled below (T) <NEEDLE GROUPS> 1. opposite pairs along the branch (RJ) 2. five-needle tufts (LP) 3. 15–60 needle tufts (T) <NEEDLE LENGTH> 1. 12 mm (RJ) 2. 8–20 cm (LP) 3. 2–5 cm (T) <CONE> 1. berry-like seed cones (RJ) 2. cylindrical seed cones (LP) 3. red and pink broadly ovoid seed cones (T)
Options	Rocky Mountain Juniper (RJ) Limber Pine (LP) Tamarack (T) Yellow Cedar (scale-like needles like the Rocky Mountain Pine) (YC) Jack Pine (triangular needles like the Limber Pine) (JP) Subalpine Larch (ovoid red seed cones like the Tamarack) (SL)
Key	Rocky Mountain Juniper (RJ) Limber Pine (LP) Tamarack (T)

erroneous option because it shares some but not all of the elements related to Rocky Mountain Pine. An incorrect option can be selected from this distractor model for inclusion in the generated item whenever it shares some but not all of the elements with the correct option.

In Step 3, the logical constraint algorithm assembles the content specified in an item model, subject to elements and constraints described in the cognitive model by the SME for both the correct and incorrect options. Iterations are conducted to assemble all possible combinations of elements and options, subject to the constraints. The outcome of the Step 3 assembly process is shown in Fig. 1. The model used in this example is small and simple because it is used to illustrate key concepts in template-based AIG. Item generation in operational testing programs is much more complex and often includes large cognitive models that address one problem with 4–7 different scenarios, 5–7 different sources of information, and 8–12 different features.

Evaluation of template-based AIG

Template-based AIG is based on item modeling. An item model specifies the parts and content in the task that will be manipulated to create new test items. Item modeling is flexible because it allows SMEs to create content for tests that are administered at different levels in the education system, for a wide range of purposes, and for different subject areas using a range of item formats and item types. Item modeling also allows SMEs to create diverse, complex, and high-quality items using an approach that is both efficient and cost effective. For these reasons, template-based AIG using item modeling is currently the preferred method for generating items. But template-based AIG has one important limitation. It relies on the SME's ability to create clearly structured cognitive and item models to guide the generation process. However, some content areas on educational tests do not permit an explicit disambiguation and extraction of features. For example, reading comprehension test items can be used to evaluate students' inferential knowledge. Reading comprehension items are designed to assess students' ability to evaluate the main idea and the sentiment of the reading passages of varying genre, length, and types. Generating test items using template-based AIG is challenging in a content area like reading comprehension because of the complexity required to identify and capture the features related to the correct option and to manipulate these features to create new items. For example, to create cognitive models to produce inference test items in reading comprehension, the SME must identify all text features (e.g., words, phrases, and sentences) used as evidence to support the key answer (e.g., overall sentiment, main topic, or theme). Without any systematic methods to analyze the deep semantic connections between the text features, and the overall themes, main ideas, and the sentiment presented in the text, this can be a challenging and laborious task for SMEs.

1. A first-year forestry student is hiking through Jasper National Park. She comes across a coniferous tree with scaly, needle shaped leaves that are in opposite pairs along the branch and measure 12 millimeters in length. Along the branches are berry-like seed cones.

These characteristics best describe which coniferous tree?

- A. Jack Pine
- B. Yellow Cedar
- C. Subalpine Larch
- D. Rocky Mountain Juniper *

*correct option

2. A first-year forestry student is hiking through Jasper National Park. She comes across a coniferous tree with triangular needles that are in five-needle tufts and measure 8-20 centimeters in length. Along the branches are cylindrical seed cones.

These characteristics best describe which coniferous tree?

- A. Jack Pine
- B. Limber Pine *
- C. Yellow Cedar
- D. Subalpine Larch

3. A first-year forestry student is hiking through Jasper National Park. She comes across a coniferous tree with needles that are flattened above and keeled below that are in 15-60 needle tufts and measure 2-5 centimeters in length. Along the branches are red and pink broadly ovoid seed cones.

These characteristics best describe which coniferous tree?

- A. Jack Pine
- B. Tamarack *
- C. Yellow Cedar
- D. Subalpine Larch

Fig. 1 Sample generated items in botany.

Non-template-based AIG and example items

To address the challenge of producing large numbers of items in a content area such as reading comprehension, non-template-based AIG can be used. Unlike the template-based method, the purpose of non-template-based AIG is to generate items directly with reduced human intervention. In other words, this method produces the stem, options, and/or auxiliary information without the use of an item model created by an SME. To accomplish this goal, the method relies on different types of natural language processing techniques and neural network systems which, in turn, draw on information from corpora and knowledge bases to generate the test items.

Three different approaches for non-template-based AIG have been described and illustrated. The first approach is syntax based. The syntax-based approach is designed to generate items by uncovering the underlying syntactic structures of the content in a corpus or knowledge base. Syntax is a description of how words are combined in sentences, where the syntactic structure of a sentence conveys meaning. Two commonly used syntactic techniques that can be used to generate items are syntax parsing and part-of-speech tagging (Brown et al., 2005; Chen et al., 2006; Danon and Last, 2017). Chen et al. (2006), for example, described a semi-automated grammar item generation system. Their system focused on identifying parts of the text or the sentences that could be used for a grammar test. Their non-template-based AIG method matched the regular expressions of the grammatical structure with predefined patterns to identify test items. Their system could also identify a set of plausible but incorrect options based on pattern matching. Brown et al. (2005) created a non-template-based AIG method that could generate items for assessing students' vocabulary knowledge. Their approach used a large lexical resource, WordNet, to generate six types of vocabulary knowledge items that measure students' ability to use definitions, synonyms, antonyms, hypernyms, hyponyms, and cloze questions. They focused on identifying the correct meaning of the word using the word and its part-of-speech tag information. Once the correct sense meaning was identified for the target word, their approach could also be used to identify a list of words that served as the correct and incorrect options. Examples of generated items using syntactic approaches are shown in Table 2.

The second approach is semantic based. The semantic-based approaches use information acquired from a semantic analysis of the content in a corpus or knowledge base. Semantics focus on the meaning of a sentence, which is often expressed using a structured combination of words. Semantic approaches for non-template-based AIG use natural language processing techniques such as topic

Table 2 Sample generated items using three non-template-based AIG methods.

Approach	Source	Sample test item
Syntactic	Chen et al. (2006)	I intend ____ you that we cannot approve your application. (A) to inform (B) to informing (C) informing (D) inform
	Brown et al. (2005)	Choose the word that best completes the phrase below: The child's misery would move even the most ____ heart. (A) torpid (B) invidious (C) stolid (D) obdurate
Semantic	Flor and Riordan (2018)	Text: Peter called on Monday./The particles from the Sun also carry an electric charge. Question: When did Peter call?/Do the particles from the Sun carry an electric charge?
Sequential	Du et al. (2017)	Text: Inflammation is one of the first responses of the immune system to infection. Stem: What is one of the first objections of the immune system to infection?
	Narayan et al. (2020)	Text: Former Beatle Sir Paul McCartney has topped the Sunday Times rich list of musicians with his £730 m fortune. Stem: Who is the richest musician in the world? Answer: Sir Paul McCartney
	Gao et al. (2018)	Text: The electric guitar is often emphasized, used with distortion and other effects, both as a rhythm instrument using repetitive riffs with a varying degree of complexity, and as a solo lead instrument. Stem: What is a solo lead instrument? Answer: The electric guitar

modeling and keyword extraction to explicitly model text that can be used as test items (Flor and Riordan, 2018; Gütl et al., 2011; Mazidi, 2017; Susanti et al., 2015). For instance, Flor and Riordan (2018) described a semantic-based AIG approach for generating factual test items using role labeling. Semantic role labeling is the process of assigning labels to different parts of a sentence in order to represent their semantic roles. Their system used the information gathered from role labeling with rule-based models to create *wh*-questions (e.g., what, who, where) and yes/no questions. Flor and Riordan (2018) used open language processing tools to analyze the grammatical structure of sentence parts required for assigning particular semantic roles. *SENNA*, for instance, is a popular tool for identifying generalizable core arguments in a sentence to provide specific labels, such as agent (or subject), patient (or object), location, and time of the event. Then, the system directly generated test items from the labeled sentences focusing on a focal label (e.g., agent) to select the most appropriate *wh*-question. Generating yes/no questions followed a similar logic. Examples of generated items using a semantic approach are shown in Table 2.

The third approach is sequence based. Sequential-based approaches are used to generate test items automatically using deep neural networks. Unlike the syntactic and semantic approaches, neural systems focus on generating test items by disambiguating and then learning the primary characteristics from the content in a corpus or knowledge base in order to generate test items without any manual feature engineering. This non-template-based AIG method does not identify the syntactic and semantic characteristics in the text. Instead, it attempts to learn from the test items that are used during training. Du, Shao, and Cardie (2017) described a sequential-based approach for generating reading comprehension items using a neural framework. Their system focused on mapping the problem statement and the reading passages to generate alternative questions to the problem statement using a recurrent neural network. Their system could generate various *wh*-questions (e.g., what, who, where). The generated items were then evaluated by SMEs to identify those items that were judged to be high in naturalness (e.g., grammatical and fluent) and difficulty (e.g., difficulty in syntactic divergence and reasoning). Narayan, Simoes, Ma, Craighead, and McDonald (2020) introduced a model focused on generating items that could be directly answered using the information from the reading passage. Their AIG approach used a transformer-based sequence-to-sequence model (Devlin et al., 2018). The system used the reading passages as input to generate the most suitable stem for the correct option. The model was based on publicly available question-answering databases (e.g., Rajpurkar et al., 2016). Gao, Bing, Chen, Lyu, and King (2018) described an approach for generating factual items from reading passages. Unlike the previous systems, Gao et al. controlled item difficulty in the generation process by using a sequence-to-sequence model that could use a sentence in the text that contained a corresponding answer with a specified difficulty level (e.g., "easy", "hard") to generate items. Examples of generated items using sequential approaches are shown in Table 2.

Evaluation of non-template-based AIG

The rapid advancement in non-template-based AIG offers a glimpse into the future of item development. This method not only overcomes a key limitation with template-based AIG but it also dramatically reduces the role of the SME in creating test items. Sequential neural network methods in particular, require no human input for engineering the features or identifying the set of

patterns or rules required to generate test items. Despite these potential benefits, the current applications of non-template-based AIG remains limited because they require a significant amount of training, instruction, and prior information that is very challenging to acquire. Most notably, all the approaches we describe for non-template-based AIG require a corpus or knowledge base. Moreover, the data in the corpus or knowledge base must be directly related to the purpose of the test because this information will be used to generate test items. When the content in the corpus is available, it can be structured using NLP methods defined at the syntactic, semantic, or sequential level to generate items. However, the data in corpora and knowledge bases that are currently available can only be used to produce very limited results that are characterized by items that are short (e.g., one or two-sentence) and simple. The quality of these simple items also varies across the approaches we described when evaluated by SMEs. The generated items are either overly simplistic or require heavy editing, filtering, and revisions to produce content that meets operational testing standard of quality. Conversely, when we evaluate the quality of the generated items using template-based AIG a markedly different result is reported. Gierl, Latifi et al. (2016) found that generated science test items were judged to be of equal or higher quality than items created by SMEs. In another study where a template-based AIG method was used to generate science items in English and Spanish, Gierl, Lai, Houston, Rich, and Boughton (2015) reported that 90% of the items were considered acceptable by SMEs while the remaining 10% required only minor revisions in order to meet an operational testing standard of quality.

Conclusion

AIG is a research area where cognitive and psychometric modeling practices guide the production of tests that include items generated with the aid of computer technology. AIG can be used to produce large numbers of high-quality test items required to support the current transition from paper to computer-based educational testing. Computer-based testing is one important form of educational technology that can be used to foster the development of a skilled workforce because computers allow educators to administer tests more frequently so students can receive continuous feedback as they develop their knowledge and skills. But when tests are administered more frequently using a computer, a large and continuous supply of new test items is needed. It is challenging—perhaps even impossible—to produce this large and continuous supply of new content using a traditional item development where SMEs author individual items. Hence, alternative item development methods are needed. Two general methods are currently available for producing large numbers of items using automated processes. The first method is template based. Template-based AIG can be characterized as a three-step approach where item modeling serves as the foundation. In the first step, content for the generated items is described by the SME. In the second step, the template for the content is created by the SME. In the third step, an algorithm places the Step 1 content into the Step 2 template using rules and constraints. Template-based AIG is a flexible approach that can be used to generate items for different levels in the education system, for a wide range of purposes, and different subject areas using a range of item formats and item types. The generated items are typically produced at a high standard of quality and therefore can be used directly in operational computer-based testing administrations. Because of the flexibility offered by this method and because the generated items meet high standards of quality, template-based AIG is currently the preferred method for generating items. But template-based AIG also relies on the SME's ability to create clearly structured cognitive and item models to guide the generation process. Not all content areas on educational tests permit explicit disambiguation and extraction of features. To address this limitation, a non-template-based AIG method can be used. Non-template-based AIG can be guided by the syntactic, semantic, or sequential structure of a text. These approaches all rely on a combination of natural language processing techniques and knowledge bases to generate statements, questions, and options from inputs such as texts, databases, and corpora of existing information without the need for templates. But to produce new items, a non-template-based AIG method requires large amounts of data specific to the educational testing area of interest. The data in the corpus must also be directly related to the purpose of the test because this information will be used to generate test items. Currently, corpora and knowledge bases are challenging to find and, even when they are available, produce limited results because the approaches used with non-template-based AIG generate short and simple test items that do not meet the standards of quality required for operational computer-based testing. Non-template AIG clearly holds tremendous promise for the future. But to meet our current item development challenges, template-based AIG is the method of choice for generating large numbers of items because it offers SMEs a great deal of control and flexibility for producing items in a wide range of content needed for computer-based testing.

References

- Bejar, I.I., 1996. Generative Response Modeling: Leveraging the Computer as a Test Delivery Medium (ETS Research Report No. 96–13). Educational Testing Service, Princeton, NJ.
- Bejar, I.I., 2002. Generative testing: from conception to implementation. In: Irvine, S.H., Kyllonen, P.C. (Eds.), *Item Generation for Test Development*. Lawrence Erlbaum, Hillsdale, NJ, pp. 199–217.
- Bejar, I.I., Lawless, R., Morley, M.E., Wagner, M.E., Bennett, R.E., Revuelta, J., 2003. A feasibility study of on-the-fly item generation in adaptive testing. *J. Technol. Learn. Assess.* 2 (3). Retrieved from: <http://www.jtla.org>.
- Brown, J., Frishkoff, G., Eskenazi, M., 2005, October. Automatic question generation for vocabulary assessment. In: *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pp. 819–826.
- Chen, C.Y., Liou, H.C., Chang, J.S., 2006, July. FAST—An automatic generation system for grammar tests. In: *Proceedings of the COLING/ACL 2006 interactive presentation sessions*, pp. 1–4.
- Danon, G., Last, M., 2017. A Syntactic Approach to Domain-Specific Automatic Question Generation arXiv:1712.09827 [cs.CL].

- Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding arXiv preprint arXiv:1810.04805.
- Du, X., Shao, J., Cardie, C., 2017. Learning to Ask: Neural Question Generation for Reading Comprehension arXiv preprint arXiv:1705.00106.
- Flor, M., Riordan, B., 2018, June. A semantic role-based approach to open-domain automatic question generation. In: *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, pp. 254–263.
- Gao, Y., Bing, L., Chen, W., Lyu, M.R., King, I., 2018. Difficulty Controllable Generation of Reading Comprehension Questions arXiv preprint arXiv:1807.03586.
- Gierl, M.J., Haladyna, T., 2013. *Automatic Item Generation: Theory and Practice*. Routledge, New York, NY.
- Gierl, M.J., Lai, H., 2013. Using automated processes to generate test items. *Educ. Meas.* 32, 36–50.
- Gierl, M.J., Lai, H., 2016. Automatic item generation. In: Lane, S., Raymond, M., Haladyna, T. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 410–429.
- Gierl, M.J., Lai, H., 2017. The role of cognitive models in automatic item generation. In: Rupp, A., Leighton, J. (Eds.), *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley, New York, NY, pp. 124–145.
- Gierl, M.J., Lai, H., 2018. Using automatic item generation to create solutions and rationales for computerized formative testing. *Appl. Psychol. Meas.* 42, 42–57.
- Gierl, M.J., Lai, H., Hogan, J., Matovinovic, D., 2015. A method for generating test items that are aligned to the Common Core State Standards. *J Appl. Test. Technol.* 16, 1–18.
- Gierl, M.J., Lai, H., Turner, S., 2012. Using automatic item generation to create multiple-choice items for assessments in medical education. *Med. Educ.* 46, 757–765.
- Gierl, M.J., Lai, H., Zhang, X., 2018. Automatic item generation. In: Khosrow-Pour, M. (Ed.), *Encyclopedia of Information Science and Technology*, fourth ed. IGI Global, Hershey, PA, pp. 2369–2379.
- Gierl, M.J., Latifi, F., Lai, H., Matovinovic, D., Boughton, K., 2016. Using automated processes to generate items to measure K-12 science skills. In: Rosen, Y., Ferrara, S., Mosharraf, M. (Eds.), *Handbook of Research on Computational Tools for Real-World Skill Development*. IGI Global, Hershey, PA, pp. 590–610.
- Gierl, M.J., Zhou, J., Alves, C., 2008. Developing a taxonomy of item model types to promote assessment engineering. *J. Technol. Learn. Assess.* 7 (2). Retrieved January 20, 2011 from: <http://www.jtla.org>.
- Gütl, C., Lankmayr, K., Weinhofer, J., Höfler, M., 2011. Enhanced Automatic Question Creator – EAQC: concept, development and evaluation of an automatic test item creation tool to foster modern e-education. *Electron. J. Elearn.* 9, 23–38.
- Haladyna, T.M., Rodriguez, M.C., 2013. *Developing and Validating Test Items*. Routledge, New York, NY.
- He, D., Lao, H., 2018. Paper-and-pencil assessment. In: *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Sage, Thousand Oaks, CA, pp. 1199–1200.
- Higgins, D., 2007. Item Distiller: text retrieval for computer-assisted test item creation. In: *Educational Testing Service Research Memorandum*. Educational Testing Service, Princeton, NJ (RM-07-05).
- Higgins, D., Futagi, Y., Deane, P., 2005. Multilingual Generalization of the Model Creator Software for Math Item Generation. Educational Testing Service Research Report (RR-05-02). Educational Testing Service, Princeton, NJ.
- Irvine, S.H., Kyllonen, P.C., 2002. *Item Generation for Test Development*. Erlbaum, Hillsdale, NJ.
- LaDuca, A., Staples, W.I., Templeton, B., Holzman, G.B., 1986. Item modeling procedures for constructing content-equivalent multiple-choice questions. *Med. Educ.* 20, 53–56.
- Lai, H., Gierl, M.J., Touchie, C., Pugh, D., Boulais, A., De Champlain, A., 2016. Using automatic item generation to improve the quality of MCQ distractors. *Teach. Learn. Med.* 28, 166–173.
- Lane, S., Raymond, M., Haladyna, R., 2016. Test development process. In: Lane, S., Raymond, M., Haladyna, T. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, NY, pp. 3–18.
- Mazidi, K., 2017, April. Automatic question generation from passages. In: *International Conference on Computational Linguistics and Intelligent Text Processing*. Springer, Cham, pp. 655–665.
- Narayan, S., Simoes, G., Ma, J., Craighead, H., McDonald, R., 2020. QURIOUS: Question Generation Pretraining for Text Generation arXiv preprint arXiv:2004.11026.
- OECD, 2019. *PISA 2018 Technical Report*. PISA, OECD Publishing, Paris.
- Rajpurkar, P., Zhang, J., Lopyrev, K., Liang, P., 2016. Squad: 100,000+ Questions for Machine Comprehension of Text arXiv preprint arXiv:1606.05250.
- Singley, M.K., Bennett, R.E., 2002. Item generation and beyond: applications of schema theory to mathematics assessment. In: Irvine, S.H., Kyllonen, P.C. (Eds.), *Item Generation for Test Development*. Erlbaum, Mahwah, NJ, pp. 361–384.
- Susanti, Y., Iida, R., Tokunaga, T., 2015, May. Automatic generation of English vocabulary tests. In: *CSEDU*, vol. 1, pp. 77–87.

Automated test assembly

Tim Davey, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	201
Test blueprints	201
Test assembly as a combinatorial optimization problem	202
Constraints	202
Objective functions	202
Solution algorithms	205
Test assembly use cases	206
Simultaneous assembly of multiple forms or item blocks	206
Adaptive test assembly	206
As an adjunct to manual assembly	207
Limitations	207
Conclusion	207
References	208

Introduction

Much of what constitutes human progress is marked by the development of machines that perform what were formerly laborious manual tasks. Until recently the tasks best suited to mechanization were those that lacked a creative element, a domain once believed the exclusive province of human thought. On its surface, assembling test forms from pools of available items may seem more prosaic than artistic. Indeed, [Campion and Miller \(2006\)](#) characterize test forms as utilitarian objects whose form should follow their function. However, many would argue that test assembly, at least done properly, has always benefitted from a creative flair. And those who did it or saw it done years ago, when index cards representing each item were strewn about large tables and gradually, after much deliberation, organized into test forms, would certainly agree that considerable expert judgment was being exercised. However, the ongoing *software revolution* is now blurring the line between aesthetics and algorithms, with computers generating music, literature, and visual art ([McCormack and d’Inverno, 2012](#)). This chapter describes past and current efforts to apply algorithms that mechanize and so routinize the test assembly process. These algorithms are generally referred to as *automated test assembly* (ATA).

Test blueprints

Proper construction of almost anything begins with a blueprint, or an organized set of rules and instructions that govern the characteristics of the assembled object. Test forms are no exception. Blueprints that guide test assembly dictate the substantive, and ideally the statistical characteristics of the items that comprise a proper test form. A blueprint should be specific and detailed enough to ensure that all tests built to it measure the same traits in the same or equivalent ways ([Lane et al., 2016](#)).

Test specifications are usually based on a taxonomy that positions each item within the substantive domain that the test measures. For example, the taxonomy for a high-school level math test may include branches for arithmetic, algebra, geometry, and trigonometry items. Beneath each of these major classes, increasingly specific distinctions among items may be made. For example, algebra items might be classified as elementary or intermediate. Going further, an intermediate algebra item might be cataloged as dealing with exponents.

Specifications usually dictate the number of items that each test form should contain from each branch of the item taxonomy. The math test in our example might specify 15 items from the algebra branch, 10 from geometry, and 5 from trigonometry. These general requirements might be supplemented by rules that dictate the composition of the test at greater levels of detail. For example, the fifteen algebra items may need to include nine intermediate and six elementary items, with the nine intermediate items perhaps including two with exponents and seven without. The numbers of items required at a given level of specification may be fuzzy rather than absolute, representing indecision or indifference on the part of test developers. For example, a specification may call for between eight and ten intermediate algebra items rather than exactly nine.

Specifications at the lowest levels may be fuzzier still, dealt with by what can be viewed as akin to “randomization” in sampling or experimental design. Rather than demand that each form contain exactly two exponent items among the eight to ten intermediate algebra items, the specifications might restrict the number of such items to no more than two. Such a requirement does not insist that an exponent item be present on every form (since none is less than two). However, if exponent items are present they will not be so numerous as to dominate or bias what the form measures.

Test specifications can also dictate the statistical characteristics of a test form. Rules here can be expressed either in terms of classical item statistics (item difficulty and discrimination) or item response theory (IRT) information functions ([Davey and](#)

Hendrickson, 2010). Statistical goals might seek to fix the average difficulty or even the distribution of item difficulties across a form. IRT information functions characterize the measurement properties of a test more thoroughly, indicating where along the score scale the tests measures well or poorly.

Test assembly as a combinatorial optimization problem

The goal of test assembly is to select items from an available pool such that the blueprint specifications are uniformly satisfied. Although simple test construction problems can be and long have been solved by hand, more complex problems are good candidates for *combinatorial optimization* algorithms (Gill et al., 1982). These algorithms work to find combinations of items that maximize some desirable test property (the *objective function*) while at the same time satisfying the most important blueprint specifications (called *constraints*).

Theunissen (1985) developed the general approach to test construction as a combinatorial optimization problem and introduced linear and integer programming as a means of its solution. Stocking and Swanson (1993) were among the first to apply automated assembly in an operational setting. To do so they also introduced a heuristic called the *weighted deviations model* (WDM) as a means of finding approximate solutions to test assembly problems (Swanson and Stocking, 1993). Others have expanded the range of problems to which automated assembly can be applied, and the sophistication and practicality of the solution algorithms (van der Linden and Boekkooi-Timminga, 1989; van der Linden and Adema, 1997; van der Linden, 2005).

Blueprint specifications are generally divided between the objective and constraint sides of the optimization problem, the basic difference being that those assigned to the objective function are test characteristics that we would *like* to have, while those designated as constraints are *must* haves. Software packages differ in the way constraints and objective functions are defined, with a common arrangement having constraints dictate the substantive or content characteristics of the items that comprise an assembled test form while the objective function handles the statistical properties of the form. Typical choices of constraints and objective functions are described in the following sections.

Constraints

The substantive characteristics of each test item must be encoded into a series of classification variables or data fields. The content taxonomy for the test generally provides several of these fields, as for the sample item below:

Field	Value
Construct:	Math
Primary content:	Algebra
Secondary content:	Intermediate
Specific content:	Operations with exponents

Other substantive characteristics may also be encoded and considered during test assembly, including item type or format, whether the item stands alone or must appear in conjunction with other items (for example, a geometric figure on which several distinct items are jointly based), and cognitive level or depth of knowledge. Essentially, any item characteristic that is dictated or controlled by the test blueprint must be encoded as a data field.

The test blueprint is translated into constraints of two basic sorts: equality and inequality. Examples of both were briefly given above. An equality constraint requires a test form to contain a specific number of items of a given sort, e.g., the number of algebra items = 15. An inequality constraint requires that the number of items of a given sort be either less than or equal to, or greater than or equal to a given value, e.g., the number of algebra items ≤ 16 . Inequality constraints can be used singly to impose an upper bound on the number of items of a given sort allowed on a test, e.g., number of exponent items ≤ 2 . Pairing inequality constraints allows the number of algebra items to take on any value at or between given boundaries, e.g., the number of algebra items is ≥ 13 and the number of algebra items is ≤ 16 .

Objective functions

The objective function is commonly used to control the statistical or performance characteristics of test forms. It is defined as a real-valued function of the item performance statistics or values derived from those statistics. The goal of optimization is to minimize the objective function in some sense, subject to the constraints. Several examples are provided below to illustrate in a practical way how objective functions are used. Each example assumes that either classical or IRT performance statistics are available for all items in the assembly pool.

- (1) *Specifying average item difficulty.* This function sets the average item difficulty (e.g., proportion correct) equal to some specified value across the n items in a form. Let d_i be the difficulty for each item i in the pool, $i = 1, \dots, N$, where N is the number of items in the pool. Define a *selection variable* x_i such that $x_i = 1$ if item i is selected into the test form and 0 if not. Finally, let $\bar{d}$ be the

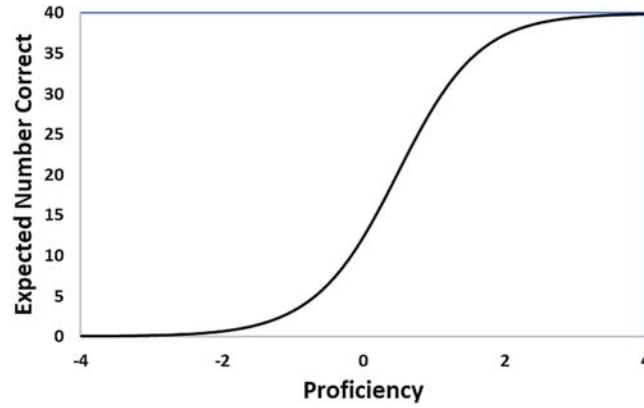


Fig. 1 Test characteristic curve.

average difficulty we would like a form to have. The objective function to be minimized is then the difference between the observed and targeted average item difficulty:

$$F = \bar{d} - \frac{1}{n} \sum_{i=1}^N x_i d_i \quad (1)$$

- (2) *Setting the test characteristic curve (TCC) of a form near to some target function.* The TCC is a function of IRT item parameter estimates. More specifically, the TCC is the sum of the *item characteristic curves* (ICCs) across the items in a test form. As shown in Fig. 1, the TCC is defined as a function relating examinee proficiency with expected test scores. As such it serves an important role in test equating, or the statistical adjustment of scores achieved on test forms that differ in difficulty in order that these scores be made directly comparable. Because equating works best when it's needed least, a common goal is to produce new test forms that are similar in difficulty to all preceding forms. Setting the TCC of a newly assembled test form as near as possible to some target function can accomplish this.

Let $p_i(t_j)$ be the value of the ICC for item i at proficiency level j . Under assumptions common to most IRT models, the TCC function at each proficiency level t_j is then the sum across the n items on the test form of the ICC values at that proficiency level. Suppose we have some target value for the TCC function, denoted as T_j across some grid of proficiency values $j = 1, \dots, J$. Then an objective function for selecting items that provide a close fit to this function can be written as:

$$F = \sum_{j=1}^J \sum_{i=1}^N x_i (p_i(t_j) - T_j)^2 \quad (2)$$

Although this function defines “close fit” in terms of squared errors, other definitions are possible (e.g., the sum of absolute distances between the observed TCC and the target). The function in Eq. (2) also weighs differences equally across all points of the proficiency scale. It is possible, and often sensible, to weight differences proportionally to the density of proficiency at each value j .

- (3) *Setting the test information function of a form close to some target.* The IRT information function characterizes the measurement properties of items and test forms (Lord, 1980). Information functions trace across the proficiency scale the measurement power items and forms contribute toward the estimation of test-taker proficiency. Fig. 2 shows example information functions for items that differ both in the amount of measurement power they possess and in where along the proficiency scale that power is most concentrated. Easier items, like Item 1, provide most of their measurement power at lower proficiency levels while more difficult items (like Item 2) measure most strongly at higher proficiency levels. Some items, like Item 3, provide little measurement power at any proficiency level.

Test information functions, like test characteristic curves, are additive across the items in a test form, again accepting assumptions common to most IRT models. Like item information functions, test information functions indicate the measurement power of a test form across the proficiency scale. Fig. 3 shows examples of easier and more difficult forms as well as forms that provide broader or more focused measurement.

Because they so concisely characterize the measurement properties of a test form, information functions are a common focus of the test assembly process. Given a test's measurement needs, target test information functions can be specified either judgmentally or empirically (e.g., by mimicking the information function of an effective, previously administered form). For example, the information functions in Fig. 3 might be set as targets for forms intended to prioritize measurement of either lower-performing test takers (*Easy*), or higher-performers (*Hard*). Similarly, the form labeled *Broad* would serve as a good target for a test intended to measure well across a broad swath of the proficiency range while the form labeled *Focused* is an effective target for a test intended to support a classification decision (e.g., pass or fail) at proficiency levels where measurement power is focused.

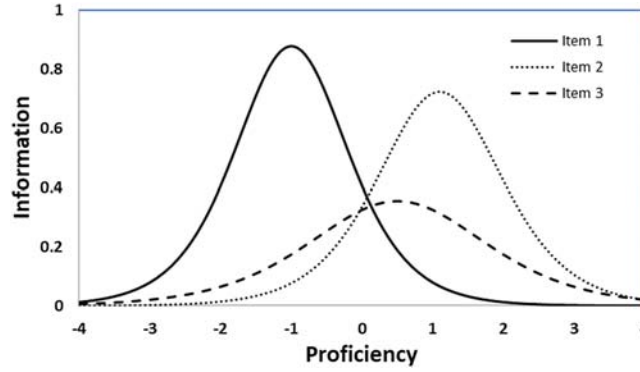


Fig. 2 Item information functions.

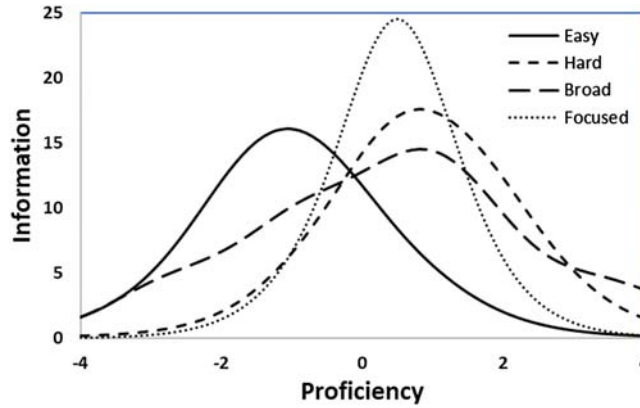


Fig. 3 Test information functions.

Test information can be incorporated into an assembly objective function exactly as a TCC in case (2) above. Let inf_{ij} be the item information function for item i evaluated at each of a grid of proficiency values $j = 1, \dots, J$. If the target information function across this same grid of proficiency values is again denoted as T_j , an objective function intended to meet this target as closely as possible is then:

$$F = \sum_{j=1}^J \sum_{i=1}^N x_i (inf_{ij} - T_j)^2 \quad (3)$$

The same considerations for use of this objective apply as above, namely that “close” can be defined by any distance metric and weights can be introduced to prioritize close fit at certain proficiency levels.

- (4) *Incorporating content considerations into the objective function.* It was noted above that the content properties of a test are often ensured through the constraint side of the optimization problem. However, doing so has both advantages and drawbacks. The advantage is that under most solution algorithms, satisfying constraints has a higher priority than optimizing the objective function, thereby reflecting the primacy of content in most assembly applications. The potential drawback is that some solution algorithms require that all constraints be met in order to return a solution. Otherwise, the problem is declared as *infeasible* (van der Linden, 2005).

Even if a feasible solution is found that meets all content constraints, the result may still be one that a human test assembler finds unappealing. An example is the case where a minor violation of the content requirements produces a significant improvement in the statistical properties of the form. A human test assembler might well prefer this tradeoff.

A better, or at least more transparent compromise between content and statistical properties might be struck by incorporating all or most of both into the objective function. This can be done by first defining some distance metric that compares the observed numbers of items in each content domain with the numbers required by the assembly blueprint.¹ Let o_k and r_k be the number of items observed and required in content domain $k = 1, 2, \dots, K$. Then let $f(o_k, r_k)$ be some function measuring

¹Considerable care must be taken both in the definition of distance metrics for content discrepancies and in the incorporation of both content and other objectives (e.g., test information) into a single objective function. The concern is that the things combined—either discrepancies across item types or content with other objectives—are not naturally on the same scale or even of the same order of magnitude. As such, the relative weighting of the various components of the objective function may not be explicit. A far more elegant solution than that presented here is to make use of solution algorithms that simultaneously optimize multiple objective functions (see Veldkamp, 1999).

the distance between observed and required values. Common choices for f are $(o_k - r_k)^2$ and $|o_k - r_k|$. It is also possible to have the difference function operate similarly to the inequality constraints discussed above, restricting the observed item counts to a permitted range rather than fixed to a single value. A disjunctive function like that below will do this:

$$f(o_k, r_k) = \begin{cases} (o_k - r_k^-)^2 & \text{if } o_k < r_k^- \\ 0 & \text{if } r_k^- \leq o_k \leq r_k^+ \\ (r_k^+ - o_k)^2 & \text{if } o_k > r_k^+ \end{cases} \quad (4)$$

where r_k^- and r_k^+ are the lower and upper boundaries of the range that should contain the observed number of items in domain k . The content contribution to the objective function is then:

$$\sum_{k=1}^K f(o_k, r_k), \quad (5)$$

to which can be added one (or several) of the statistical objectives in Eqs. (1)–(3) to complete the objective function.

Test developers may also judge that it is more important to meet some content requirements than others. These priorities can be reflected by assigning weight values, w_k , to each content domain, k in Eq. (5), resulting in the content contribution:

$$\sum_{k=1}^K w_k f(o_k, r_k). \quad (6)$$

Because both content and statistical requirements are represented in the objective function, it is further possible to assign a priority weight to each, making the compromise between content and statistical properties explicit.

- (5) *Practical objectives.* Practical considerations may impose requirements on test assembly beyond the content and statistical properties of test forms or item blocks. For example, when test security is a concern, requirements may stipulate that all items in the assembly pool be administered at roughly equal rates across a series of administration occasions. In this case, item usage rates need to be tracked and made available to the assembly routine, which would prioritize use in blocks currently under assembly of items that were more rarely used in prior assemblies.

Solution algorithms

Anyone who has assembled a test form by hand can attest to the difficulty of the problem. The task of identifying, accounting for, and choosing among a host of often competing priorities requires sharp judgment, endless patience, and steely determination. Translating the problem into a computerized algorithm has been equally difficult.

Test assembly belongs to a large class of combinatoric optimization problems referred to as *knapsack problems* (Fletcher, 2000). In its simplest form, the knapsack problem considers a set of objects, each of a certain size and carrying a given value. The problem is to select a subset of these objects that fit within a knapsack of given size such that their collective value is maximized. Test assembly generalizes this paradigm by typically imposing more than a single constraint; items are characterized by multiple content properties rather than a single weight. Such problems can always be solved by *brute force*, or by trying all possible combinations of objects (items, in our case). However, the number of possible item combinations (and so the computer processing time) grows at an alarming rate as the length of the test and, especially, the size of the assembly pool increase. For example, selecting a 40-item test from a modest 200-item pool results in a number of possible combinations that extends to 42 digits, rendering brute force an unworkable strategy for real-world problems. That leaves a series of algorithms of increasing sophistication that promise results that approximate the best (optimal) solution to increasing degree.

The first algorithms applied to operational test assembly problems were so-called *greedy heuristics*. Examples of these algorithms were developed by Swanson and Stocking (1993) and Luecht (1998). A greedy heuristic works by breaking the combinatoric problem down into bite-sized chunks, here the selection of a single item from the pool. The algorithm works to build the test from an item at a time, selecting at each point the “best” item from the remaining pool, given those items already selected. “Best” is usually defined in terms of the objective functions described under scenario (4), above, those that incorporate both content and statistical properties.

As a simple example, consider the item pool and test assembly requirements set out in the table below.

Classification	Items available in the pool	Items required on each form
Arithmetic	25	10
Algebra	50	5
Geometry	25	5
Total	100	20

The distribution of item content is seen to differ between what is available in the pool and what is required on the test form. Notably, Arithmetic items must comprise half of each test form but constitute only one-quarter of the pool. An effective heuristic algorithm will recognize this discrepancy and so prioritize Arithmetic items as early selections. Algebra items, in contrast, are over-represented in the pool relative to what is needed on each form. Selection of Algebra items can therefore wait until more immediate needs are met. Because these items are plentiful in the pool, they will likely be instrumental in meeting the statistical needs of the form.

Because heuristic algorithms proceed sequentially, they are subject to be trapped in *local* optima and so fail to minimize the objective function completely (Fletcher, 2000). A good heuristic will then attempt to adjust the completed form to further minimize the objective function (Swanson and Stocking, 1993). For example, a common strategy is to consider swapping each item currently selected within the test with each item remaining in the assembly pool. Trades that result in improvement are accepted and the process continues until no further improvement is possible.

Closer approximations to the optimal solution can be obtained through algorithms developed to solve general *integer programming* problems, of which our knapsack problem is a special case (Fletcher, 2000). These algorithms work to define and then divide the *solution space*, which in our case consists of all test forms that could possibly be assembled from the item pool and their associated objective function values. Strategies for doing so include the simplex method, branch-and-bound, and cutting-plane algorithms (Wolsey, 1998).

The *simplex method* was developed in the more general context of *linear programming*, of which integer programming is a special case (Katta, 1983). However, the method is applicable only in cases where the selection variable (x_i) is real-rather than integer-valued, meaning that items may contribute fractionally to a test form, clearly not a sensible result. Although such fractional values can be rounded to the nearest integer, doing so does not guarantee an optimal (or even feasible) solution. A better solution is found by applying a *branch-and-bound* or *cutting-plane* strategy to the rounded selection variables. These methods work to successively divide and constrain the problem space, ultimately converging on an optimal and integer-valued solution (Wolsey, 1998).

Branch-and-bound and cutting-plane algorithms are found in most modern integer programming software packages. Dozens of such packages are available, ranging in capabilities from limited to expansive, in complexity from simple to daunting, and in price from free to exorbitant. Performance of some of these packages is compared in Cor et al. (2008) and Luo (2020). Given the widespread availability of these packages and the speed of current computers, there is little reason to settle for a heuristic solution in most operational settings today.

Test assembly use cases

Although the discussion above is set in the context of assembling a single, conventional test form, ATA capabilities extend beyond this scenario. Most importantly, the assembled item collections need not constitute complete, intact test forms. Rather, they can be sets of items (often called *blocks* or *modules*) that serve as components of tests. Common use cases involving full test forms or item blocks are described below.

Simultaneous assembly of multiple forms or item blocks

Under this scenario, multiple test forms or item blocks are simultaneously assembled from a common item pool. Forms or blocks may be assembled to identical specifications or each may be unique. Sharing of items across forms or blocks may be prohibited or restricted to certain limits (e.g., no more than 10% of the items in a block may be shared with other blocks).² The trick under this scenario is to ensure that all forms or blocks meet specifications to a nearly equal degree. Heuristic algorithms often seek to ensure equality by selecting items for each form in turn rather than selecting an entire form and then moving on to the next, i.e., forms are selected in parallel rather than in series.

Adaptive test assembly

Under adaptive testing, different test takers receive different test forms, the difficulty of which is tailored by computer to the performance level of the test taker. Adaptive tests make a series of decisions as the test proceeds, selecting items or item blocks for administration based on the test-taker's demonstrated level of performance. Adaptive testing follows either of two basic strategies: *item-adaptive* and *multi-stage adaptive*, the difference being in the number of decision points and the size of the item block selected at each point. ATA has an important role to play under both strategies.

As conceived, item-adaptive tests select items individually, deciding what item to administer next based on the test-taker's performance on those items administered so far. A unique or nearly unique test form is then sequentially assembled for each test taker as the test proceeds. This assembly is entirely under control of the computer administering the test, with no human intervention possible. As such, item-adaptive testing was the object of much of the early research into automated test assembly. Stocking and Swanson (1993) developed their weighted deviations model heuristic, for use with adaptive testing systems. More recently, van

²Restricting sharing of items across forms can greatly increase the number of constraints imposed, and so the complexity of the assembly problem.

der Linden and others have applied the more formal integer programming algorithms described above to the same purpose (van der Linden and Reese, 1998).

Multi-stage testing is more constrained than item-adaptive, routing each test taker through a series of pre-assembled item blocks, with each block selected to be of a difficulty appropriate for the test-taker's performance level. Under this design, ATA is employed prior to rather than during test administration, assembling the collection of item blocks through which test takers are routed.

A particular sort of practical objective arises when data collected from administration of multiple test forms or adaptively selected blocks are expected to support IRT item calibration. In this case, assembly objectives may be added that ensure sufficient administration sample sizes for each item (Jewsbury and van Rijn, 2020).

As an adjunct to manual assembly

Apart from item-adaptive testing, where human inspection of test forms prior to administration is impossible, most applications of ATA involve human oversight of some sort. For example, test developers typically review and often revise test forms initially drafted by ATA. The advisability of this practice will be made clear in the section below detailing some of ATA's limitations. However, closer interaction between human and machine is possible, with both working cooperatively to produce a form. For example, a human test developer could manually assemble the core of a test form and then ask ATA to best complete it. ATA can also be used to suggest replacement items for those removed from a form under review. Modern item banking systems that include ATA capabilities typically allow for some level of interaction.

Limitations

The author has long recounted an anecdote wherein he used an early ATA system to assemble what he naively believed to be a "perfect" test form measuring high-school level science. He proudly delivered this form to the test development team expecting fulsome praise. However, the test developers instead noted that the form contained an item pertaining to lake-effect snow, another concerning stream ecology and a third dealing with ocean currents. "There is", they declared, "an awful lot of water on this form". There was, of course, no formal content requirement that dealt with water level. Rather, an excess of watery content is simply one of those things that a good, experienced test developer recognizes when one sees it.

Liquid is in fact one of thousands of such things that don't rise to the level of requiring a formal requirement being encoded and enforced. These things instead fall under the realm of learned experience and expert judgment. Therein lies the primary limitation of ATA: the system can attend to only those test characteristics that are explicitly identified and called out, and these characteristics will always be a fraction of those that test developers might deem important. Of course, it is ever a matter of debate as to the extent that these characteristics impact test-taker performance and so warrant attention.

Human review of computer-generated forms is the principal way of dealing with the limitations of ATA systems. However, there are situations where review is not possible and ATA, and the inputs it's provided, need to be trusted more completely. Chief among these situations is item-adaptive testing, where test forms are assembled uniquely for each test taker in real time as the test proceeds. Where review is not possible, it seems prudent to constrain test assembly more strictly and at greater levels of detail than would be necessary for human assembly (Davey et al., 2016).

One of the primary benefits of human review is the identification of item "enemies". These are pairs or larger combinations of items that for any of a variety of reasons should not appear together on a form. They may cue one another, contradict one another, or be too similar in content or appearance. The three water-based items discussed above fall into this last category. Although item enemies can be identified and coded as such prior to assembly, this becomes a daunting and ultimately impossible task as the size of an item pool (or the number of factors that render items as enemies) increases.

A final limitation pertains to test assembly generally but seems more salient under ATA. This is that item characteristics are often poorly determined meaning that efforts to constrain them are necessarily murky. This is most obvious in the case of IRT item parameter estimates which often contain substantial error, meaning that test characteristic curves or information functions used as objectives may be poorly estimated. The impact of this is most pronounced in the case where the objective function is to maximize test information. This will likely result in the selection of items whose information functions have been over-estimated, meaning the resulting test information function will be grossly over-estimated.

Even item content classifications are sometimes imprecise, with test developers often disagreeing as to how an item should be classified. What it means to "optimize" test characteristics given uncertainty in both item content and statistical properties is a question to ponder. At the very least, it may mean that there may be little practical difference between solutions returned by simpler heuristic algorithms and the "more optimal" solutions returned by more complex algorithms.

Conclusion

Most technologies move over time from being complex, expensive, and specialized to being simple, cheap, and widely accessible. This has certainly been true of automated test assembly, which is now available to all testing programs through a variety of convenient and inexpensive software systems.

ATA can make test assembly much easier and far more efficient than manual procedures. However, it requires time and attention up front to define the assembly problem and create the system inputs needed to solve it. As such, the capability is best applied by testing programs that assemble test forms routinely, to the same or similar requirements.

Finally, it is best in nearly all cases that human test developers be as closely involved in the assembly process as possible, either in a review and oversight capacity or interacting with the ATA system throughout the assembly process in complete cooperation. “Trust, but verify” ever has been and ever will be a useful motto to live by.

References

- Campion, D., Miller, S., 2006. Test production effects on validity. In: Haladyna, T., Downing, S. (Eds.), *Handbook of Test Development*, first ed. Routledge, New York.
- Cor, M., Alves, C., Gierl, M., 2008. Computer software review: conducting automated test assembly using the premium solver platform version 7.0 with Microsoft Excel and the large-scale LP/QP solver engine add-in. *Appl. Psychol. Meas.* 32, 652–663.
- Davey, T., Hendrickson, A., 2010. Classical versus IRT statistical specifications for creating test forms. Presented at the annual meeting of the National Council for Measurement in Education. Denver.
- Davey, T., Slater, S., Pitoniak, M.J., 2016. Designing computerized adaptive tests. In: Lane, S., Raymond, M., Haladyna, T. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York.
- Fletcher, R., 2000. *Practical Methods of Optimization*, second ed. Wiley, New York.
- Gill, P.E., Murray, W., Wright, M.H., 1982. *Practical Optimization*. Academic Press, London.
- Jewsbury, P.A., van Rijn, P.W., 2020. IRT and MIRT models for item parameter estimation with multidimensional multistage tests. *J. Educ. Behav. Stat.* 45, 383–402.
- Katta, M.G., 1983. *Linear Programming*. John Wiley, New York.
- Lane, S., Raymond, M., Haladyna, T., 2016. *Handbook of Test Development*, second ed. Routledge, New York.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Erlbaum, Hillsdale, NJ.
- Luecht, R.M., 1998. Computer-assisted test assembly using optimization heuristics. *Appl. Psychol. Meas.* 22, 224–236.
- Luo, X., 2020. Automated test assembly with mixed-integer programming: the effects of modeling approaches and solvers. *J. Educ. Meas.* 57, 547–565.
- McCormack, J., d’Inverno, M., 2012. *Computers and Creativity*. Springer, Heidelberg.
- Stocking, M.L., Swanson, L., 1993. A method for severely constrained item selection in adaptive testing. *Appl. Psychol. Meas.* 17, 277–292.
- Swanson, L., Stocking, M.L., 1993. A model and heuristic for solving very large item selection problems. *Appl. Psychol. Meas.* 17, 151–166.
- Theunissen, T.J., 1985. Binary programming and test design. *Psychometrika* 50 (4), 411–420.
- van der Linden, W.J., Adema, J.J., 1997. Simultaneous assembly of multiple test forms. *J. Educ. Meas.* 35, 185–198.
- van der Linden, W.J., Boekkool-Timminga, E., 1989. A maximin model for test design with practical constraints. *Psychometrika* 54, 237–247.
- van der Linden, W.J., Reese, L.M., 1998. A model for optimal constrained adaptive testing. *Appl. Psychol. Meas.* 22, 259–270.
- van der Linden, W.J., 2005. *Linear Models for Optimal Test Design*. Springer, NYC, NY.
- Veldkamp, B.P., 1999. Multiple objective test assembly problems. *J. Educ. Meas.* 36 (3), 253–266.
- Wolsey, L.A., 1998. *Integer Programming*. John Wiley, New York.

Computerized adaptive testing and multistage testing

Andreas Frey^{a,b}, ^aGoethe University, Frankfurt, Germany; and ^bCentre for Educational Measurement (CEMO), University of Oslo, Oslo, Norway

© 2023 Elsevier Ltd. All rights reserved.

Introduction	209
How computerized adaptive testing works	209
Building blocks	210
Item pool	210
Start of the test	211
Person parameter estimation	211
Item selection	211
Fully adaptive item selection	211
Item selection with limited adaptivity	212
Constraint management	212
End of the test	213
Effects of adaptive testing	213
Measurement efficiency	214
Validity	214
Test-taking motivation	215
Conclusion	215
References	215

Introduction

Tests are an important type of quantitative measurement instruments. They are used to measure one or more constructs of interest. Most tests use a fixed set of test items and the test items are presented in a fixed order. Computerized adaptive tests differ from this in that the set of items presented to the examinee is based on the examinee's responses to previous items in the same test and can make use of existing preknowledge about the standing of the individual regarding the construct of interest. Computerized adaptive testing (CAT) is typically based on psychometric item response theory models (IRT; van der Linden, 2016a) and comprises six essential building blocks.

The history of IRT-based CAT dates back to the 1970s. Although individual attempts had already been made to develop response-dependent test procedures before then (see Chang, 2015 and Weiss, 2004 for overviews), it was only from this time on that viable CAT concepts were formulated and tested. With their groundbreaking work, Lord (1980) and Weiss (1983) were the first to present the essential aspects of adaptive testing in an integrated form. Starting in the 1980s, numerous fundamental questions relevant for an operational application of CAT were investigated in the development of the Computerized Adaptive Testing version of the Armed Services Vocational Aptitude Battery (CAT-ASVAB; Sands et al., 1997). Since then, the US military has been using that test battery for personnel selection with approximately 400,000 test takers per year. In addition, CAT applications can also be found in a wide range of areas such as health measurement, competence measurement (e.g., in large-scale assessments of student achievement), intelligence testing, aptitude testing, and testing in higher education.

In this article I will first describe the basic idea of CAT. In order to configure a modern CAT-system, six building blocks need to be specified. I will describe the basic methods used in each of these building blocks. I will thereby clarify that multistage testing (MST) is a special kind of adaptive testing. Next, I will summarize the current state of knowledge about the effects of CAT on (a) measurement efficiency, (b) the validity of test score interpretations, and (c) test-taking motivation. I will end the article with a conclusion, an overview of current research directions in the area of CAT, and a brief outlook.

How computerized adaptive testing works

The flowchart shown in Fig. 1 shows the general principle behind CAT. The test starts with the selection and presentation of an initial item, which is drawn from an existing item pool (Building blocks 1 and 2; see **Item pool** and **Start of the test** sections). The response given to this item is recorded and scored. With this empirical information a little bit more is known than before about the level of the examinee regarding the construct to be measured. This information is used to derive a provisional estimate $\hat{\theta}$ of the person parameter, that is, the examinee's score on the construct of interest (Building block 3; see **Person parameter estimation** section). After this, it is checked whether one or more previously defined termination criteria are fulfilled. After the first

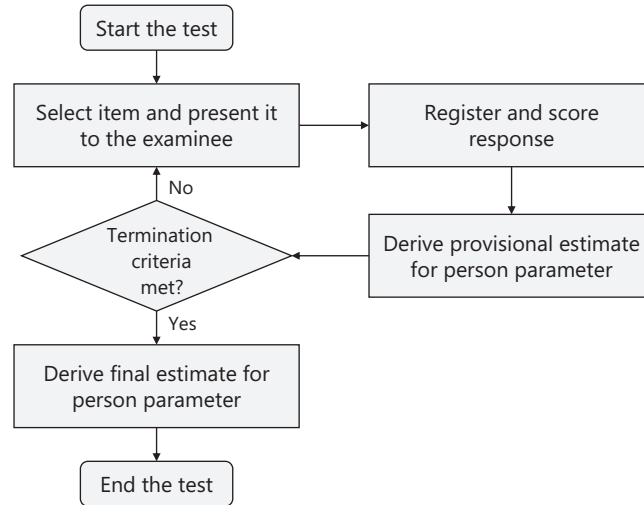


Fig. 1 Flowchart of the core principles of computerized adaptive testing.

item, this will normally not be the case. Therefore, Steps 1 to 3 are iterated until all termination criteria are met. The selection of one (fully adaptive CAT) or more items (MST, testlet-based CAT) in Step 1 uses the responses that the examinee has given in previous iterations. Note that, to achieve good readability, the rest of the text refers to the selection of individual items. This includes the possibility of selecting groups of items (testlets or modules in MST), even if this is not explicitly added each time. The previous responses are used in Step 1 to be able to present only those items that fit well to the examinee with regard to an optimality criterion (Building block 4; see [Item selection](#) section). The item with the maximum value of this optimality criterion is presented next. For most optimality criteria used in CAT, examinees with high levels regarding the construct of interest answer more difficult items than examinees with lower levels; this fosters the precision of the measurement (see [Measurement efficiency](#) section). In addition to this, adaptive item selection also avoids the situation that examinees with a high level regarding the construct being measured (e.g., high ability level) have to answer items that are much too easy for them and that examinees with a low level have to answer items that are much too difficult for them. This tailored item presentation was traditionally expected to foster the test-taking motivation of the examinees. Empirical results on whether this expectation is met in operational testing will be discussed in [Test-taking motivation](#) section. In addition to the statistical optimality criterion, other nonstatistical constraints (Building block 5; see [Constraint management](#) section) are usually taken into account during item selection in Step 1. After all of the termination criteria have been met, the final person parameter estimate $\hat{\theta}$ is determined and the test is terminated (Building block 6; see [End of the test](#) section).

Building blocks

Item pool

The item pool is the complete set of items that can be used for testing. Over the course of an adaptive test, the item that meets the optimality criterion best is selected from the remaining eligible items in the item pool. To calculate the optimality criteria typically employed in CAT, statistical properties are required for all items in the item pool. These statistical properties are estimated using IRT models from responses gathered in a calibration study. Many adaptive tests use logistic IRT models with one (1PL), two (2PL), or three (3PL) item parameters, where the item parameters represent these item properties. The 3PL is the most general of these three models. This model defines the probability of a correct response $u_{ij} = 1$ of person $j = 1, \dots, N$ with latent ability θ_j to an item $i = 1, \dots, I$ as:

$$P(u_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{\exp[a_i(\theta_j - b_i)]}{1 + \exp[a_i(\theta_j - b_i)]} \quad (1)$$

where a_i is the discrimination parameter, b_i is the difficulty parameter, and c_i is the pseudo-guessing parameter of item i . [Eq. \(1\)](#) reduces to the 2PL when c_i is set to zero for all items, which in turn reduces to the 1PL when a_i is set to a constant ([van der Linden, 2016a](#)). The use of IRT models is beneficial in CAT because the item difficulties and the persons are located on the same common scale. Thereby, examinees can be compared with each other even if they have answered different items. Based on their responses, they can be placed between the items and thus on the common scale that represents the levels of the construct of interest.

For typical adaptive tests, the item parameters are estimated in the calibration phase and then considered as known in the operational CAT phase in which the test is used to test examinees. Therefore, the calibration study is of utmost importance to ensure that an adaptive test functions properly. In order to obtain precise item parameter estimates, relatively large calibration samples are needed. As a rough rule of thumb, [de Ayala \(2009\)](#) recommends a minimum of a few hundred responses per item for the 1PL

model, 500 responses for the 2PL model, and 1000 responses for the 3PL model under good conditions (good fit between item difficulties and ability distribution, many items per examinee). Online calibration methods provide possibilities to reduce these relatively large demands regarding the sample size in the CAT construction phase. The Continuous Calibration Strategy (CCS) by Fink et al. (2018), for example, allows the complete calibration to be carried out while the test is already in use.

Start of the test

At the beginning of an adaptive test, no responses are available yet, which means that the selection of the first item cannot be informed by previously given responses. Frequently, the first item is therefore drawn randomly from a set of “icebreaker” items, which have a low item difficulty. Alternative procedures are to make completely random draws from the item pool or to use information from previous tests (if available) to select the first item. However, both simulations and empirical studies have shown that, in CAT, the selection of the first item has a comparatively small influence on the precision of the final test results. Accordingly, most authors assign only minor importance to the choice of the first item in CAT (e.g., Hambleton et al., 1991).

Person parameter estimation

In CAT, provisional estimates for the person parameters are derived in Step 3 and final estimates in Step 4 (see Fig. 1). Common IRT estimation methods (van der Linden, 2016b) can be used for the estimation of the person parameters. Both maximum likelihood (ML) estimators, which include the WLE (Weighted Likelihood Estimate; Warm, 1989), as well as Bayesian estimators such as the EAP (Expected a Posteriori; Bock and Mislevy, 1982) or the BME (Bayes Modal Estimate; Mislevy, 1986) are frequently used in CAT. For unidimensional CAT, comparisons of different person parameter estimation methods can be found in van der Linden and Pashley (2010) and Cheng and Liou (2000). Because ML estimators are undefined for invariant response patterns (i.e., if only correct or only incorrect answers have been given), in practice, the provisional person parameter estimate $\hat{\theta}$ is often derived with Bayesian estimators until correct and incorrect answers are recorded and then an ML-based estimator is used.

Item selection

In prototypical CAT, item selection is accomplished on the level of single items; this is also referred to as fully adaptive item selection (Weiss, 2011). However, there are also approaches with limited adaptivity in which groups of items are selected jointly. This section describes both approaches successively.

Fully adaptive item selection

Adaptation at the item level offers the most flexible way of adapting the presented items to the examinee’s response behavior. The adaptation is based on an optimality criterion that is calculated after each response for all eligible items in the item pool. The item with the most desirable value of this optimality criterion is then selected for presentation. A frequently used optimality criterion is maximum item information given $\hat{\theta}$. The calculation of the item information varies with respect to the IRT model used. For the 3PL model, the item information function is given by:

$$I_i(\theta) = a_i^2 \left(\frac{p_i(\theta) - c_i}{1 - c_i} \right)^2 \frac{q_i(\theta)}{p_i(\theta)} \quad (2)$$

where $p_i(\theta)$ is the probability that an examinee with ability θ can solve item i (see Eq. 1) and $q_i(\theta)$ is the probability that item i cannot be solved by this examinee, so that $q_i(\theta) = 1 - p_i(\theta)$ applies. The equation shows that the discrimination parameter a_i as well as the pseudo-guessing parameter c_i have a direct influence on the item information $I_i(\theta)$. While high a -parameters are associated with high item information, high c -parameters are associated with low item information. With regard to the c -parameter, the highest item information results for $c_i = 0$.

The 2PL model and the 1PL model have fewer item parameters than the 3PL model. Therefore, the equation for the item information for the 2PL is reduced to:

$$I_i(\theta) = a_i^2 p_i(\theta) q_i(\theta) \quad (3)$$

and for the 1PL to:

$$I_i(\theta) = p_i(\theta) q_i(\theta) \quad (4)$$

Thus, in item selection that is conducted according to the maximum item information criterion, parameters a and c are relevant in the 3PL and a is relevant in the 2PL. In the 1PL, neither a nor c appears in the item information function. In this IRT model, the maximum value for I_i is achieved if $p_i(\theta) = q_i(\theta) = 0.5$, which is the case if the item difficulty is equal to the provisional person parameter estimate $\hat{\theta}$. Item selection that uses the maximum item information criterion therefore selects the item whose item difficulty is closest to $\hat{\theta}$.

Note that, because the standard error of the person parameter estimate can be represented by $SE(\theta) = 1/\sqrt{I(\theta)}$ for the 1PL and the 2PL, applying the maximum item information criterion means reducing the SE in the best possible way each time an item is selected if these models are used.

An overview of alternative optimality criteria for CAT item selection can be found, for example, in [van der Linden and Pashley \(2010\)](#). In some testing situations, these alternative optimality criteria can be used to increase the performance of CAT compared to the performance achieved with item selection conducted according to the maximum item information criterion. However, the maximum item information criterion has been proved to be only slightly worse in these cases and is applicable across a wide range of situations. Therefore, this robust criterion is considered the standard criterion for item selection in CAT.

Item selection with limited adaptivity

In some studies, such as large-scale assessments of student achievement (e.g., PISA, PIRLS, or TIMSS), items are grouped into testlets. Testlets consist of several items that refer to a common stimulus and should be presented as a whole. Solutions for the adaptive selection of complete testlets instead of single items while considering the local dependence of the items which share the same stimulus have been proposed by [Keng \(2011\)](#) for unidimensional CAT and by [Frey et al. \(2016\)](#) for multidimensional CAT.

Another type of tests from the early days of CAT that use observed response behavior for item selection purposes have recently experienced a renaissance. These type of tests are tests with two to five stages (multistage tests; e.g., [Yan et al., 2014](#); initially referred to as flexilevel tests, [Lord, 1971](#)). Multistage testing (MST) has probably come back into focus because—at first glance—organizational requirements of LSAs can be solved well with it. Each stage of a multistage test consists of several items and the stages are assembled in advance. In MST, examinees initially answer the items included in a “routing test” ([Lord, 1971, 1980](#)). Examinees who do well in this routing test (i.e., give many correct answers) get a harder subtest (called module) and examinees who do not do so well in the routing test (i.e., give many incorrect answers) get an easier subtest in the next stage. This principle is followed for all stages of the test. [Fig. 2](#) shows a sample MST design with a routing test in stage 1, three modules with different average difficulty in stages 2 and 3 each.

MST also incorporates an adaptation even though it is relatively unrefined compared to fully adaptive item selection on the level of single items. Therefore, MST represents an intermediate stage between nonadaptive and fully adaptive computerized testing that does not make use of the full potential of CAT. The fact that all possible paths are known a-priori in MST is sometimes mentioned as an advantage compared to fully adaptive item selection. With all paths known, it is straightforward to plan the exposure rates of items and to ensure that individual test takers are confronted with certain proportions of items per facet of the measured construct (e.g., subdimensions). However, this is not fully convincing, as such test properties can also be achieved (and often even more precise) through constraint management methods (see [Constraint management](#) section) in fully adaptive item selection. Detailed information on MST can be found in [Magis et al. \(2017\)](#) and [Yan et al. \(2014\)](#).

Constraint management

Adaptive item selection is based on a statistical optimality criterion (see [Item selection](#) section). However, in practical CAT applications, additional nonstatistical requirements usually have to be met. These typically mean selecting items that are not optimal in a statistical sense. In such a case, a compromise between meeting statistical optimality and nonstatistical requirements is necessary. The methods used to consider nonstatistical requirements in CAT are subsumed under the term constraint management methods. These methods can be divided into two groups. The first group of constraint management methods controls how often individual items are presented across examinees and is called exposure control. The second group of methods regulates the handling of content requirements for each examinee and is referred to as content management.

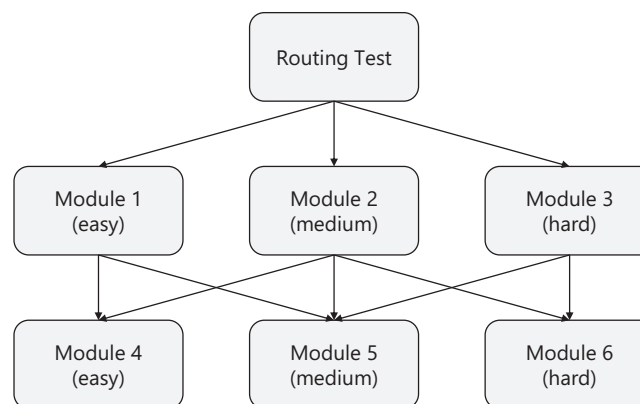


Fig. 2 1-3-Multistage testing design with one routing test, three modules with differing average difficulty at stage 2 and three modules with differing average difficulty at stage 3.

Exposure control methods are often necessary in CAT because item selection that aims to optimize only the statistical optimality can easily lead to situations in which some items are presented to many or even all examinees while other items are rarely or never presented. This can lead to test security issues because frequently presented items have an increased probability of becoming publicly known in the population of future test takers. In this case, the item parameter estimates from the calibration study (with examinees without item preknowledge) can become invalid. Various exposure control strategies have been developed in order to control the frequency with which items are presented. The simple Randomesque Method suggested by [Kingsbury and Zara \(1989\)](#) consists of adding a stochastic component to the statistical optimality criterion. Instead of selecting the single best item according to the optimality criterion, a random draw from the m optimal items is carried out; with m typically being a number between 4 and 10. The Randomesque Method does not guarantee that the maximum item exposure rate is limited to a target rate, but it can prevent the same item from being used repeatedly for examinees with similar levels of the construct to be measured. A more thorough control of item exposure rates can be achieved with more elaborate algorithms such as the Sympson-Hetter Method ([Sympson and Hetter, 1985](#)), the Progressive Method ([Revuelta and Ponsoda, 1998](#)), and the Alpha Stratification Method ([Cheng et al., 2009](#)). A comparison of the performance achieved with the difference exposure control methods can be found in [Leroux et al. \(2013\)](#).

In addition to controlling item exposure rates across examinees, it is often desirable to control the type of items presented to the individual examinees. Such methods are referred to as content management methods. They can be used to control the proportions of items presented per facet of the construct of interest, per response format, per position of the correct answer, and others. Modern content management methods make it possible to consider several of these requirements simultaneously. Probably, the most powerful content management method is the Shadow Testing Method proposed by [van der Linden and Reese \(1998\)](#). With this method, very good results are achieved even for high numbers of nonstatistical constraints. Heuristic methods such as the Weighted Deviation Model ([Stocking and Swanson, 1993](#)), the Weighted Penalty Model ([Shin et al., 2009](#)), and the Maximum Priority Index Method ([Cheng and Chang, 2009](#)) represent alternatives to the Shadow Testing Method for many situations. They are easy to implement and perform well, even though they might not be the best choice for highly constrained testing situations. Comparisons of the heuristic content balancing methods can be found in [Born and Frey \(2017\)](#) and [He et al. \(2014\)](#).

End of the test

An adaptive test is continued until one or more predefined test termination criteria are met. Four frequently used test termination criteria are listed below ([Linacre, 2000](#)).

An adaptive test is terminated when

1. a specified number of items has been answered and/or
2. the standard error of the person parameter estimate is sufficiently small and/or
3. the maximum testing time is reached and/or
4. the complete item pool has been presented.

[Segall \(2005\)](#) points out that the choice of the termination criterion strongly depends on the context, the item pool, and the testing conditions. If, on the one hand, the test scores are to be used for decisions that have a high personal relevance (e.g., university admission tests), person parameter estimates that have comparable precision across examinees are desirable, as they result from the second test termination criterion. If, on the other hand, complete groups of examinees are tested at the same time (e.g., complete school classes), tests with varying test lengths are impractical or even not feasible. In such conditions, adaptive tests are usually terminated for all members of the tested group after a fixed time has elapsed (test termination Criterion 3). Because data stemming from such studies is typically analyzed with group-level statistics, varying precision between examinees plays a minor role here.

In practice, a combination of test termination criteria is often used. A simulation study can be found in [Babcock and Weiss \(2012\)](#), where different test termination criteria as well as combinations of them were examined while taking item-pool characteristics into account.

Effects of adaptive testing

Because CAT is linked to the use of a computer, it has the same advantages and disadvantages as computer-based testing (CBT). This is solely due to the use of a computer to administer the test and is not connected to the use of adaptive test algorithms. Because these advantages and disadvantages have been intensively discussed before (e.g., [Parshall et al., 2002](#)), they are only briefly summarized here. The advantages that can be achieved with CBT compared to paper-and-pencil testing include high test security, highly standardized testing procedures, quick and accurate determination of test results, application without psychometric expertise, fast reporting of test results, and the possibility of using innovative item formats (e.g., [Sireci and Zenisky, 2016](#)). The potential disadvantages of CBT compared to paper-and-pencil testing are less of a conceptual and more of an organizational nature. They include higher test development effort, expenses incurred by providing computers at the testing location, and potential fairness issues caused by individual differences in computer-related personality variables (e.g., familiarity with computers, computer anxiety) that may be correlated with the construct of interest.

Measurement efficiency

The main advantage of CAT compared to nonadaptive testing is its superiority in terms of measurement efficiency. The measurement efficiency of a test can be conceptualized as the ratio of measurement precision to test length (Segall, 2005). The test length is usually quantified by the number of items presented. Measurement precision is defined as the accuracy of the test scores. It can vary with respect to θ in IRT modeling. One way to represent it is with the test information function. If a single value for a test's measurement efficiency is of interest, the average test information can be calculated for a defined theta range and divided by the average number of items presented.

Item selection in CAT is often based on the criterion of maximum item information (see **Fully adaptive item selection** section). This maximizes the measurement precision and thus the measurement efficiency. This is connected to two advantages of CAT. Firstly, the number of items needed in CAT to reach the same measurement precision as in nonadaptive testing is much lower than the number needed in nonadaptive testing (typically only 40%–60% of the items). Secondly, CAT can align the standard errors of the person parameter estimates across examinees, especially if a corresponding test termination criterion such as Criterion 2 (see **End of the test** section) is used. The degree of the alignment of the standard errors across examinees that can be reached also depends on the characteristics of the item pool. To achieve perfectly equal standard errors for all examinees, the item pool needs to include enough items for the θ -value of each examinee. This means that the test information function of the item pool must exceed a certain level across the entire θ -range for which examinees are expected. If this information function falls below this level for a certain θ -range, suboptimal items will be selected for examinees in this range and this will result in higher standard errors of the final person parameter estimates.

Validity

According to the Standards for Educational and Psychological Testing (AERA et al., 2014, p. 11), validity “refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests”. In empirical validation studies, it is common practice to clarify, among other aspects, whether test results (a) are correlated as expected with other measures representing the same construct (convergent evidence) and (b) are uncorrelated with measures representing theoretically unrelated constructs (discriminant evidence).

With regard to convergent validity evidence (a), CAT results in higher correlations with convergent variables than nonadaptive testing if its superior measurement efficiency is utilized for an increase in measurement precision. This is the case because the increased measurement precision means an increase in the proportion of systematic variance of the person parameter estimates that is due to the construct that the test aims to measure in relation to unsystematic error variance (Frey, 2006). The additional systematic variance is likely to correlate with convergent variables, while the unsystematic error variance has a zero correlation as expectancy with all variables by definition. The result is a higher correlation with convergent variables. If, however, CAT's superior measurement efficiency is used to reduce the number of items presented while maintaining the same measurement precision as in a nonadaptive test, it is reasonable to expect no effects on the convergent validity evidence.

While the effects of CAT on convergent validity evidence can be derived directly from the optimality criterion used for adaptive item selection, the effects of CAT on discriminant validity evidence are somewhat more difficult to predict. By increasing the proportion of systematic variance to the total variance of the estimated person parameters, no systematic changes in the correlation of the test score with discriminant variables are to be expected. This is the case because both the systematic variance and the unsystematic error variance are expected to be uncorrelated with discriminant variables. Increasing the proportion of systematic variance will thus not affect the correlation with discriminant variables. Nevertheless, the special type of testing that is realized in CAT obviously can have an effect on psychological variables, which may in turn interfere with the test results. For example, the fact that only about half of the items are answered correctly may lead to negative emotional reactions such as anxiety or anger, especially in highly proficient examinees who are not used to failing in such a large proportion of test items. These negative emotions can then have a (typically negative) impact on the test scores. However, the goal of most achievement tests is to measure maximum individual performance. If the test results are lowered by confounding variables such as anxiety or anger, they do not represent maximum performance. This will limit or may even jeopardize the validity of test score interpretations. In order to derive valid test score interpretations, near-zero correlations between test scores and confounding variables are thus aimed for and considered to provide discriminant validity evidence. Empirical results presented in Frey (2006) in fact showed near-zero correlations between confounding variables (such as trait test anxiety and state anger) measured before the test and negative affect measured during the test for the adaptive version of a test and significant correlations for a nonadaptive version of the same test. However, this does not mean that CAT generally increases discriminant validity. Ortner and Caspers (2011) reported opposite effects for their study with $N = 110$ university students: Significant correlations between the confounding variable trait test anxiety and intelligence as the construct of interest were found in CAT but not in nonadaptive testing. However, the problematic correlations observed for CAT disappeared in this study when the test instruction included information about the adaptivity and the fact that the proportion of items that can be solved was around 50% for all examinees. Thus, transparent test instruction is important in CAT with regard to the discriminant validity of the test score interpretations.

Test-taking motivation

Previous research has repeatedly claimed that CAT has a positive effect on test-taking motivation. This claim can be traced back to early works on CAT by [Betz \(1975\)](#) and [Betz and Weiss \(1976a,b\)](#). On the basis of the results of several experiments, [Betz and Weiss \(1976b\)](#) concluded that CAT increases test-taking motivation, especially for examinees with low proficiency. This finding was explained by the fact that examinees get test items whose item difficulty is tailored to their proficiency level. CAT prevents examinees from being presented both with items that are much too easy for them and hence may cause boredom and with items that are much too difficult for them and hence may cause frustration. Beyond these early findings, however, empirical evidence that supports the positive effects of CAT on test-taking motivation is sparse ([Wise, 2014](#)).

Several studies from the 1990s and 2000s even point in the opposite direction, indicating a negative effect of CAT on test-taking motivation. Those studies frequently argue that presenting items with an average probability of a correct response of .50 in CAT does not lead to high test-taking motivation (e.g., [Bergstrom et al., 1992](#); [Eggen, 2004](#); [Ponsoda et al., 1999](#)). Such a test is considered to be demotivating, especially for high-performing examinees who are used to being able to solve a much larger proportion of test items. This argument is supported by the results of [Ortner et al. \(2014\)](#). They compared an adaptive and a nonadaptive version of a nonverbal matrices test with regard to the two motivational aspects of situational fear of failure and perceived probability of success in the test. The examinees in the adaptive condition reported higher fear of failure and lower perceived probability of success compared to the examinees in the nonadaptive condition.

The expectancy-value model of test-taking motivation proposed by [Asseburg \(2011\)](#) unites these seemingly contradicting results regarding the effects of CAT on test-taking motivation by considering the expectation and value components separately. The model was examined in a comprehensive experiment with $N = 703$ students in Grade 9. The findings of the Asseburg study are manifold and lead to differentiated practical recommendations for the configuration of computerized adaptive tests for different test settings and groups of test takers. One conclusion of the study concerns the test instruction. If the test instructions explain the adaptivity to the examinees, as well as the fact that a probability of .50 for correct answers is normal in CAT, unwanted effects on test-taking motivation can be avoided.

Another important point with regard to the motivational effects of CAT is connected with the average difficulty of the items presented. If an optimality criterion is used for item selection that leads to an average proportion of correct answers of about 70%, the resulting test-taking motivation is substantially higher compared to a typical CAT with a proportion of around 50% correct answers ([Ling et al., 2017](#)) even though this is connected to a somewhat lower measurement efficiency. In conclusion, there is no straightforward answer to the question of whether CAT fosters test-taking motivation or not. Nonetheless, the way in which CAT needs to be specified in order for a given test situation to have positive motivational effects is clearly documented.

Conclusion

CAT is a special way of testing that uses the examinees' responses to present only statistically optimal items while taking nonstatistical constraints into account. CAT is made up of six building blocks: item pool, start of the test, person parameter estimation, item selection, constraint management, and end of the test. A collection of methods is used to specify these building blocks. If these methods are configured in such a way that they fit well to a given assessment situation, the use of CAT can be expected to have positive effects on measurement efficiency and on the validity of test score interpretations while also avoiding negative effects on test-taking motivation. In addition to the basic aspects of CAT that are summarized in this article, CAT-related research covers many additional areas. Several of these are concerned with expanding the application range of CAT to novel testing situations and to complex requirements regarding the type of measures to be obtained. Such requirements involve, for example, incorporating more complex IRT models, such as compensatory multidimensional models ([Frey and Seitz, 2009](#)), noncompensatory multidimensional models ([Hsu and Wang, 2019](#)), or higher-order models ([Wang and Kingston, 2019](#)). Another line of research focuses on the possibility of using CAT in order to overcome general measurement and testing problems, for example, to use the available testing time optimally by using the maximum information per time unit as the item selection criterion ([Cheng et al., 2017](#)) or to control for item position effects ([Frey and Fink, in press](#)).

As the transition from paper-based to computer-based testing will certainly not stop, it is most likely that CAT will play an even greater role in the future than it does now. CAT is not only able to realize highly efficient measurements; it can also consider an enormous number of constraints automatically and can thereby help to make tests better and to reduce costs. These features make CAT especially attractive for large and complex assessments such as large-scale assessments of student achievement ([Frey and Seitz, 2011](#)). In the future, the trend might go to CAT as novel, self-optimizing testing systems in which procedures that are now carried out sequentially, such as item calibration, fit analysis, drift and cheating detection, and test score reporting, are represented by continuous processes.

References

- AERA, APA, NCME, 2014. Standards for Educational and Psychological Testing. AERA, Washington.
- Asseburg, R., 2011. Leistungsbereitschaft in Testsituationen. Motivation zur Bearbeitung adaptiver und nicht-adaptiver Leistungstests [Preparedness to Perform in Test Situations. Test Taking Motivation of Adaptive and Non-adaptive Achievement Tests]. Tectum, Marburg.
- Babcock, B., Weiss, D.J., 2012. Termination criteria in computerized adaptive tests: do variable-length CATs provide efficient and effective measurement? *J. Comput. Adapt. Test.* 1, 1–18.

- Bergstrom, B.A., Lunz, M.E., Gershon, R.C., 1992. Altering the level of difficulty in computer adaptive testing. *Appl. Meas. Educ.* 5, 137–149.
- Betz, N.E., Weiss, D.J., 1976a. Effects of Immediate Knowledge of Results and Adaptive Testing on Ability Test Performance. University of Minnesota, Minneapolis.
- Betz, N.E., Weiss, D.J., 1976b. Psychological Effect of Immediate Knowledge of Results and Adaptive Ability Testing. University of Minnesota, Minneapolis.
- Betz, N.E., 1975. New types of information and psychological implications. In: Weiss, D.J. (Ed.), *Computerized Adaptive Trait Measurement: Problems and Prospects*. University of Minnesota, Minneapolis, pp. 32–43.
- Bock, R.D., Mislevy, R.J., 1982. Adaptive EAP estimation of ability in a microcomputer environment. *Appl. Psychol. Meas.* 6, 431–444.
- Born, S., Frey, A., 2017. Heuristic constraint management methods in multidimensional adaptive testing. *Educ. Psychol. Meas.* 77, 241–262.
- Chang, H.H., 2015. Psychometrics behind computerized adaptive testing. *Psychometrika* 80, 1–20.
- Cheng, Y., Chang, H.H., 2009. The maximum priority index method for severely constrained item selection in computerized adaptive testing. *Br. J. Math. Stat. Psychol.* 62, 369–383.
- Cheng, P.E., Liou, M., 2000. Estimation of trait level in computerized adaptive testing. *Appl. Psychol. Meas.* 24, 257–265.
- Cheng, Y., Chang, H.H., Douglas, J., Guo, F., 2009. Constraint-weighted α -stratification for computerized adaptive testing with nonstatistical constraints balancing measurement efficiency and exposure control. *Educ. Psychol. Meas.* 69, 35–49.
- Cheng, Y., Diao, Q., Behrens, J.T., 2017. A simplified version of the maximum information per time unit method in computerized adaptive testing. *Behav. Res. Methods* 49, 502–512.
- de Ayala, R.J., 2009. *The Theory and Practice of Item Response Theory*. Guilford, New York.
- Engen, T.J.H.M., 2004. Contributions to the Theory and Practice of Computerized Adaptive Testing. Print Partners Ipskamp, Enschede.
- Fink, A., Born, S., Spoden, C., Frey, A., 2018. A continuous calibration strategy for computerized adaptive testing. *Psychol. Test Assess. Model.* 60, 327–346.
- Frey, A., Fink, A., (in press). Controlling for item position effects when adaptive testing is used in large-scale assessments. In: Khorramdel, L., von Davier, M., Yamamoto, K. (Eds.), *Innovative Computer-Based International Large-Scale Assessments—Foundations, Methodologies, and Quality Assurance Procedures*, Springer.
- Frey, A., Seitz, N.N., 2009. Multidimensional adaptive testing in educational and psychological measurement: current state and future challenges. *Stud. Educ. Eval.* 35, 89–94.
- Frey, A., Seitz, N.N., 2011. Hypothetical use of multidimensional adaptive testing for the assessment of student achievement in PISA. *Educ. Psychol. Meas.* 71, 503–522.
- Frey, A., Seitz, N.N., Brandt, S., 2016. Testlet-based multidimensional adaptive testing. *Front. Psychol.* 7, 1–14.
- Frey, A., 2006. Validitätssteigerungen durch adaptives Testen [Increasing Validity by Adaptive Testing]. Lang, Frankfurt.
- Hambleton, R.K., Zaal, J.N., Pieters, J.P.M., 1991. Computerized adaptive testing: theory, applications, and standards. In: Hambleton, R.K., Zaal, J.N. (Eds.), *Advances in Educational and Psychological Testing: Theory and Applications*. Kluwer Academic/Plenum Publishers, New York, pp. 341–366.
- He, W., Diao, Q., Hauser, C., 2014. A comparison of four item-selection methods for severely constrained CATs. *Educ. Psychol. Meas.* 74, 677–696.
- Hsu, C.L., Wang, W.C., 2019. Multidimensional computerized adaptive testing using non-compensatory item response theory models. *Appl. Psychol. Meas.* 43, 464–480.
- Keng, L., 2011. A Comparison of the Performance of Testlet-Based Computer Adaptive Tests and Multistage Tests. Proquest, Ann Arbor.
- Kingsbury, G.G., Zara, A.R., 1989. Procedures for selecting items for computerized adaptive tests. *Appl. Meas. Educ.* 4, 359–375.
- Leroux, A.J., Lopez, M., Hembry, I., Dodd, B.G., 2013. A comparison of exposure control procedures in CATs using the 3PL model. *Educ. Psychol. Meas.* 73, 857–874.
- Linacre, J.M., 2000. Computer-Adaptive Testing: a methodology whose time has come. In: Sunhee, C., Unson, K., Eunhwa, J., Linacre, J.M. (Eds.), *Development of Computerized Middle School Achievement Test*. Komesa Press, Seoul.
- Ling, G., Attali, Y., Finn, B., Stone, E.A., 2017. Is a computerized adaptive test more motivating than a fixed-item test? *Appl. Psychol. Meas.* 41, 495–511.
- Lord, F.M., 1971. The self-scoring flexilevel test. *Educ. Psychol. Meas.* 8, 147–151.
- Lord, F.M., 1980. Applications of Item Response Theory to Practical Testing Problems. Lawrence Erlbaum Associates, Hillsdale.
- Magis, D., Yan, D., von Davier, A.A., 2017. *Computerized Adaptive and Multistage Testing With R*. Springer International Publishing, Cham, Switzerland.
- Mislevy, R.J., 1986. Bayes modal estimation in item response models. *Psychometrika* 51, 177–195.
- Ortner, T.M., Caspers, J., 2011. Consequences of test anxiety on adaptive versus fixed item testing. *Eur. J. Psychol. Assess.* 27, 157–163.
- Ortner, T.M., Weißkopf, E., Koch, T., 2014. I will probably fail: higher ability students' motivational experiences during adaptive achievement testing. *Eur. J. Psychol. Assess.* 30, 48–56.
- Parshall, C.G., Spray, J.A., Kalohn, J.C., Davey, T. (Eds.), 2002. *Practical Considerations in Computer-Based Testing*. Springer, New York.
- Ponsoda, V., Olea, J., Rodriguez, M.S., Revuelta, J., 1999. The effects of test difficulty manipulation in computerized adaptive testing and self-adapted testing. *Appl. Meas. Educ.* 12, 167–184.
- Revuelta, J., Ponsoda, V., 1998. A comparison of item exposure control methods in computerized adaptive testing. *J. Educ. Meas.* 35, 311–327.
- Sands, W.A., Waters, B.K., McBride, J.R. (Eds.), 1997. *Computerized Adaptive Testing: From Inquiry to Operation*. American Psychological Association, Washington.
- Segall, D.O., 2005. Computerized adaptive testing. In: Kempf-Leonard, K. (Ed.), *Encyclopedia of Social Measurement*. Elsevier, Amsterdam, pp. 429–438.
- Shin, C.D., Chien, Y., Way, W.D., Swanson, L., 2009. Weighted Penalty Model for Content Balancing in CATs. Pearson, San Antonio.
- Sireci, S.G., Zenisky, A.L., 2016. Computerized innovative item formats: achievement and credentialing. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Routledge/Taylor & Francis Group, London, pp. 313–334.
- Stocking, M.L., Swanson, L., 1993. A method for severely constrained item selection in adaptive testing. *Appl. Psychol. Meas.* 17, 277–292.
- Symposium, J.B., Hetter, R.D., 1985. Controlling item-exposure rates in computerized adaptive testing. In: *Proceedings of the 27th Annual Meeting of the Military Testing Association*. Navy Personnel Research and Development Center, San Diego, pp. 973–977.
- van der Linden, W.J., Pashley, P.J., 2010. Item selection and ability estimation in adaptive testing. In: van der Linden, W.J., Glas, C.A.W. (Eds.), *Elements of Adaptive Testing*. Springer, New York, pp. 3–30.
- van der Linden, W.J., Reese, L.M., 1998. A model for optimal constrained adaptive testing. *Appl. Psychol. Meas.* 22, 259–270.
- van der Linden, W.J. (Ed.), 2016a. *Handbook of Item Response Theory. Volume One: Models*. Chapman & Hall/CRC, Boca Raton.
- van der Linden, W.J. (Ed.), 2016b. *Handbook of Item Response Theory. Volume Two: Statistical Tools*. Chapman & Hall/CRC, Boca Raton.
- Wang, W., Kingston, N., 2019. Adaptive testing with a hierarchical item response theory model. *Appl. Psychol. Meas.* 43, 51–67.
- Warm, T.A., 1989. Weighted likelihood estimation of ability in item response models. *Psychometrika* 54, 427–450.
- Weiss, D.J., 1983. *New Horizons in Testing: Latent Trait Test Theory and Computerized Adaptive Testing*. Academic Press, New York.
- Weiss, D.J., 2004. Computerized adaptive testing for effective and efficient measurement in counseling and education. *Meas. Eval. Counsel. Dev.* 37, 70–84.
- Weiss, D.J., 2011. Better data from better measurements using computerized adaptive testing. *J. Methods Meas. Soc. Sci.* 2, 1–27.
- Wise, S.L., 2014. The utility of adaptive testing in addressing the problem of unmotivated examinees. *J. Comput. Adapt. Test.* 2, 1–17.
- Yan, D., von Davier, A.A., Lewis, C. (Eds.), 2014. *Computerized Multistage Testing: Theory and Applications*. Chapman & Hall/CRC, Boca Raton.

Score reporting: design and evaluation methods informed by research

Priya Kannan, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Score reporting research	217
Audience specificity	218
Focus on stakeholder interpretation and use	219
Iterative multistep framework for designing and evaluating score reports	219
Gather assessment information needs	220
Design preliminary mockups	220
Data visualization principles	221
Development of score report mockups and prototypes	222
Conduct user-centered needs assessments	223
Using focus groups and surveys to conduct user-centered needs assessments	223
Develop prototypes and iterate with systematic feedback	226
Cognitive interviews and surveys to evaluate stakeholder comprehension and usability	226
Surveys and interviews to evaluate stakeholder preference	226
Operational deployment	227
Large-scale surveys to evaluate intended uses and decisions	227
Future of score reporting research	227
Interactive reporting systems	228
References	228

Tests or assessments play a major role in education. Their main purpose is to gain insights into the test taker's knowledge, skills, and abilities (KSAs) on the focal content measured by the test. These insights have to be communicated to various stakeholders or users (e.g., teachers, principals, administrators, admissions officers, parents, and often the test-takers themselves) using some form of a score report. Score reports, which include both static (paper/online) reports and interactive dashboards, are the vehicles used to communicate results of the assessment to score users. These results are often communicated through graphical representations and data tables for individual students or groups of test takers (Hambleton and Zenisky, 2013). To be useful, however, score reports should convey the results from the assessment in a manner that makes it easy for score users to make clear interpretations about the test taker's KSAs as measured on the test and use those test results to take appropriate actions that are warranted based on the test's purpose. In fact, a test's validity is dependent on the extent to which inferences drawn from the test results are appropriate and meaningful for the intended stakeholders (Hubley and Zumbo, 2011; Kane, 2006, 2013; Tannenbaum, 2019; Zapata-Rivera and Katz, 2014). Some authors (e.g., O'Leary et al., 2017) have even suggested that the evidence of interpretability of the information contained in score reports should be included as one of the sources of a test's validity evidence. Therefore, it is critical that score reports are appropriately designed for specific audiences and presented in simple and clear language to enable score users to interpret the information accurately and use the results appropriately.

In this chapter, I first broadly discuss the development of score reporting research with a particular focus on two broad areas: (i) audience specificity and (ii) stakeholder interpretation and use of information presented in score reports. Then, I introduce the iterative multistep framework for the design and evaluation of score reports. Within this framework, I focus on recommended evaluation methods that would provide practitioners with the necessary tools to help with the implementation of this iterative multistep framework. I conclude the chapter by suggesting areas for future research, particularly surrounding the design and development of interactive reporting systems.

Score reporting research

Score reporting, as a field, goes beyond just communicating the scores obtained on a test. Zapata-Rivera and Katz (2014) rightly point out that score reports are "where the rubber hits the road" (p. 442) in the validity argument for a test. For some stakeholders, such as parents, score reports are the first, and perhaps only, point of interaction with the assessment and its results—therefore, score reports should not only communicate scores from the assessment, but also serve as the vehicle for communicating information about the test's purpose, convey the meaning of various scores, sub-domain scores, and any performance levels, and serve as a resource for actionable next steps for the intended stakeholder. In other words, an ideal score report would be designed to serve as a comprehensive one-stop shop to communicate the test's purpose, the intended interpretations from the scores, and the suggested next steps to the score user. As Hambleton and Zenisky (2013) point out, "good reporting" (p. 481) should be contextualized

and provide intended stakeholders with results that are meaningful and practical by linking them to actions. And, as previously mentioned, the validity of the assessment is dependent upon the appropriate and accurate inferences drawn from score reports by stakeholders.

However, until about 10 years ago, score reporting was by far the most under-researched area within educational measurement (Hambleton and Zenisky, 2013). This was because score reporting was historically not conducted as a planned activity during test development, but more as an “ad hoc activity” (p. 480) at the end of the test development process (Hambleton and Zenisky, 2013). Over the last couple of decades this has been changing. Following some ground-breaking work by Howard Wainer around the inaccessibility of data presentation formats (e.g., Wainer, 1997; Wainer et al., 1999), one of the first papers to call attention to the dire need for research around score reporting was published by Goodman and Hambleton in 2004. This seminal paper reviewed current practices on score reporting and highlighted some of the main concerns around the ways in which score reports were designed and developed at the time. Primary among the issues highlighted in the Goodman and Hambleton paper was the inaccessibility of the communicated information. They highlighted the profusion of the use of technical language and inaccessible data presentation formats that are not comprehensible to the intended user. Following this paper there has been an increase in awareness of the critical importance of score reporting when it comes to stakeholder interpretation and use and its role in establishing validity arguments.

Since then, and particularly, in the last decade, score reporting has become a critical requirement within the request for K-12 state contract proposals in the United States, with the necessary push for some focus groups with the intended stakeholders to ensure appropriate interpretations and use of reported test scores. Moreover, there is a strong emphasis on score reporting across several chapters of the latest *Standards* (American Educational Research Association [AERA] et al., 2014), and particular emphasis on score reporting in Chapter 6 (Test scoring, reporting, and interpretation) where best practices around score reporting are set in standards 6.10, 6.11, 6.12, 6.13 and 6.16. Finally, research on issues surrounding score reporting has also substantially increased in the past decade with the necessary focus on audience specificity (Zapata-Rivera and Katz, 2014) in the design and development and on stakeholder interpretation and use (Kane, 2013) in the evaluation of score reports. Recently, NCME sponsored an edited volume on score reporting research and applications (Zapata-Rivera, 2019a), which includes a comprehensive set of edited chapters on the latest developments in score reporting research and applications. In this section, I will provide a broad overview of score reporting research that has been conducted in the last decade or more with particular focus on audience specificity and stakeholder interpretation and use of the results.

Audience specificity

Educational tests are developed and administered in various contexts with various purposes, each using a unique score scale, and typically have more than one intended audience or score user. Researchers (e.g., Marshall and Drummond, 2006; Underwood et al., 2007) have found that even teachers and administrators who are more in-grained within the larger context of education and assessments find a lot of the information reported in score reports to be inaccessible or impractical to their needs. Therefore, the results reported from each of these types of assessments should consider the needs and attitudes of the diversity of stakeholders (e.g., teachers, administrators, parents) who are likely to receive reports from and take actions based on the results.

In their influential paper on applying audience analysis to the design of score reports, Zapata-Rivera and Katz (2014) recommend that the needs, pre-existing knowledge, and attitudes (NKAs) of each stakeholder group be taken into consideration in developing audience-specific score reports. They point out that each stakeholder group such as parents, teachers, and administrators are likely to have different needs for information from the score report describing student performance on an assessment. In addition, they are each likely to have different levels of pre-existing knowledge about the assessment and its context, and have different attitudes, feelings, or biases that might color their interpretations of the information shown in the reports (Zapata-Rivera and Katz, 2014).

The majority of score reporting research, particularly in the last decade, has focused on identifying the NKAs of specific stakeholder groups such as parents (e.g., Kannan et al., 2018a, 2021b), teachers (e.g., Kannan et al., 2017, 2021a; Marshall and Drummond, 2006; Underwood et al., 2007; Zapata-Rivera et al., 2010, 2016), students (e.g., Corrin, 2019; Vezzu et al., 2012), and administrators (e.g., Hambleton and Slater, 1997; Kannan et al., 2017; Underwood et al., 2010).

Results from these studies highlight the diverse NKAs for each stakeholder group. For example, parents mainly want to know how their child has performed in an assessment, what these scores mean, and how they can help their child improve (Kannan et al., 2018a), but they may not be intimately aware of the assessment context and their attitudes toward assessments may sometimes be impacted by the media (Bennett, 2016). In contrast, teachers are particularly interested in information that can directly guide their instruction whether it is in the formative context (Brown et al., 2019) or for “receiving teachers” (p. 72) in the summative context (Kannan et al., 2017). And, administrators are most interested in average performance of the school or district population, resource allocation based on school/district performance, and program evaluation based on comparisons with other similar schools and districts (Kannan et al., 2017). However, teachers and administrators are often busy and have limited time to analyze volumes of data (Underwood et al., 2010). For example, teachers have often reported that they feel inundated with large volumes of raw student data which they find overwhelming to sift through, and due to this, researchers (Zapata-Rivera et al., 2010) have found that teachers may sometimes draw unwarranted conclusions from the scores.

Therefore, for the intended stakeholders to interpret the information accurately and use it appropriately, test developers should conduct an in-depth audience analysis before beginning any score report design process. Moreover, as Zapata-Rivera and Katz (2014) point out, the audience analysis should not just be considered a “mental exercise” (p. 13) at the beginning of the process. Rather, the intended stakeholder should be at the center of all decisions made at every step of the score report design and evaluation process.

Focus on stakeholder interpretation and use

As critical as it is to take the intended stakeholder's needs into consideration, it is equally, if not more important to evaluate the extent to which these intended stakeholders understand the information presented in the score reports and use these results appropriately. If the intended score user is either unable to understand the presented information or unable to glean how to use these results appropriately, the validity of the entire assessment may be called into question. Therefore, a lot of the focus within score reporting research (e.g., Kannan et al., 2017, 2018a; Zapata-Rivera et al., 2010, 2019; Zwick et al., 2014) has fittingly centered around stakeholder interpretation and use of the information presented in score reports. However, there is also a clear audience specificity in this line of research since stakeholder comprehension is understandably unique for each stakeholder group.

Prior to the recent momentum in score reporting research, score reporting was conducted as an ad hoc activity without much deliberation on the way in which information was presented in the score reports. This led to the inclusion of a lot of technical jargon and unclear data presentation formats that left score users confused and befuddled. In addition, cognitive aspects (Hegarty, 2019) such as perception, attention, working memory, and prior knowledge, which varies largely across individuals but also particularly between various stakeholder groups, also tends to play a critical role in the extent to which stakeholders can comprehend the information presented in score reports. Therefore, evaluating the appropriateness of the information for each stakeholder group is essential.

Parents were generally the most disadvantaged since they do not typically have the same type of background or orientation to the testing context as other stakeholders. Particularly with the unprecedented increase in large scale accountability testing since 2001, numerous score reports with metrics such as student growth percentiles (SGPs; Betebenner, 2009) and a variety of subscore indicators were being sent home to parents. A few researchers (e.g., Barber et al., 1992; Kannan et al., 2018a; Sireci et al., 2013; Zapata-Rivera et al., 2014) have focused on parents' interpretation and use of information presented in score reports. Research (Kannan et al., 2018a) has found that parents, who are a diverse and heterogeneous group, particularly, in terms of their education levels and English language proficiency, in general struggle to understand technical terms such as standard error of measurement (SEM) presented in score reports, and this is particularly the case for parents from underserved groups (Kannan, 2020).

However, rather than dismiss the need for the inclusion of such information in parent score reports, research indicates that parents desire the inclusion of technical information such as measurement error as long as it is clearly presented and explained (Kannan et al., 2021b). Moreover, the majority of the parents in previous research (Kannan et al., 2021b) have indicated that they are keen to understand the technical details of their child's performance especially when high-stakes decisions (e.g., retention or eligibility for participation in special programs) are to be made from their child's results. Overall, these results point to the importance of including clear supplementary information, possibly in multiple languages (Kannan, 2020), that helps aid parent comprehension of their child's results, and the need for clear guidance for both the interpretation and appropriate use of scores in parent score reports.

Even though teachers and administrators are likely to be more in-grained within the larger context of educational assessments, researchers (e.g., Marshall and Drummond, 2006; Underwood et al., 2007, 2010) have found that they also find a lot of the information reported in score reports to be inaccessible and/or tend to use the information for unwarranted purposes (Zapata-Rivera et al., 2010). Teachers have often been found to struggle to parse out some of the technical information presented in score reports (Impara et al., 1991; Underwood et al., 2007; Zapata-Rivera et al., 2012). In particular, teachers have been shown to have several misconceptions of complex technical concepts such as measurement error (Zwick et al., 2014), and the inclusion of brief tutorials (Zapata-Rivera et al., 2016) have been shown to be beneficial to aid teacher comprehension of such information.

Similarly, administrators and policy makers who are strapped for time have also been shown to be overwhelmed with large volumes of data (Kannan et al., 2017; Underwood et al., 2010), have trouble understanding the way in which data are presented (Hambleton and Slater, 1997) and draw unwarranted conclusions from the information presented in score reports. As a solution, Wainer et al. (1999) first proposed the evaluation of alternative visual displays with stakeholders—they found that simplified visual displays led to more accurate and faster responses by the policymakers. But, more importantly, Underwood et al. (2010) proposed an evidence-based framework for designing administrator and policy maker reports that links student data to focal questions that are directly linked to the types of decisions made by these stakeholders. Such reports designed to answer focal questions have been shown to result in better comprehension and foster appropriate use among administrators and policymakers (VanWinkle et al., 2011).

Iterative multistep framework for designing and evaluating score reports

Two related, but separately proposed, models for designing and evaluating score reports for large-scale educational assessments were offered by Zapata-Rivera et al. in 2012 and Hambleton and Zenisky in 2013. These models were largely informed by best practices in various aspects of test development (Hambleton and Zenisky, 2013) and by best practices in human-computer interaction (Zapata-Rivera et al., 2012), and offer a series of iterative steps to guide the score report development process. These models have been subsequently adapted and implemented in a number of research studies (e.g., Kannan et al., 2018b, 2019, 2021a) as an iterative multistep framework (see Fig. 1 below) that specifically focuses on stakeholder needs in informing the design and stakeholder interpretation and use in iteratively evaluating the score reports. In this section of the chapter, I will discuss the five major steps within the iterative multistep framework as presented in Fig. 1 with a major emphasis on the tools and evaluation methods that help with the implementation of this framework in practice.

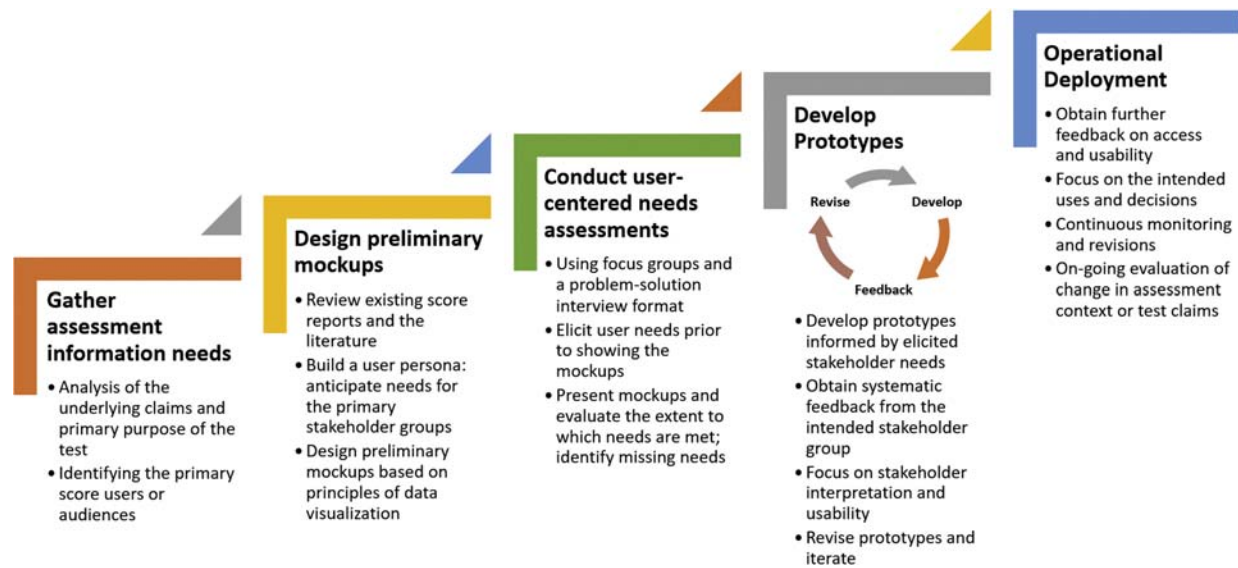


Fig. 1 Iterative multistep framework for the design and development of score reports and reporting systems. Based on models proposed by Zapata-Rivera et al. (2012) and Hambleton and Zenisky (2013).

The iterative multistep framework starts with an analysis of the underlying claims of the assessment to ensure that the information communicated within the score reports are aligned to the test's purpose and an identification of the primary score users to ensure audience specificity in the design and development of the score reports. Next, preliminary mockups are designed based on an in-depth review of the literature and anticipating stakeholder needs. These preliminary mockups are used to elicit stakeholder needs by having them react to what's useful, confusing, and missing in the mockups. Once stakeholder needs are identified, the multistep framework continues through ongoing and iterative score report development and evaluation cycles. Finally, once score reports are operationally deployed, it is ideal to have continuous monitoring of the accessibility and usability of the reports within the changing context of the assessment and stakeholder needs to cycle back to the iterative process as needed. Each of the five major steps in this framework (see Fig. 1) are discussed in detail below.

Gather assessment information needs

A test's validity is dependent upon the extent to which the inferences made from the assessment results are appropriate and warranted based on the purpose of the test (Hubley and Zumbo, 2011; Tannenbaum, 2019). It is critical therefore to ensure that the score report, as the primary vehicle through which such inferences are made, is designed to reflect the test's underlying purpose and directly presents scores and results that are warranted based on the claims of the assessment. As such, the score report design and development process should always start with an evaluation of the assessment claims by conducting a client interview with the testing program (Slater et al., 2019) and evaluating the test's underlying claims and purpose.

Slater et al. (2019) provide a comprehensive list of questions that could help guide this conversation with the client, and include questions that are intended to identify the target audience or score user(s), the basic types of knowledge and skills measured by the test, the types of decisions that would be made based on this information, and if there are any complex metrics such as sub-scores or growth scores that need to be reported. These questions are intended to provide the necessary starting point to design some preliminary mockups.

Zapata-Rivera et al. (2012) introduced the concept of creating prospective score reports (or PSRs) while gathering assessment information needs. The PSR is a document that is used to gather client requirements and can include various pieces of information such as representations for scores and subscores from other reports developed in the past, definitions of skills, possible performance levels, etc. The PSR is then shared with content and measurement experts who can identify possible discrepancies and provide appropriate suggestions for including information relevant and appropriate to the assessment. The PSR can then serve as a starting point for designing the preliminary mockups that are used in the needs assessments with potential stakeholders. Moreover, designing PSRs early in the test design process also serves to guide test development process—in particular, this helps test designers to develop the test in a manner that helps capture the necessary data to support desired claims in an evidence-centered design approach (ECD; Mislevy et al., 2003; Underwood et al., 2010).

Design preliminary mockups

The design of preliminary mockups may start with some "napkin sketches" (p. 96) that are drafted during the initial client interview (Slater et al., 2019); this can also take the form of the PSR mentioned above. In developing these mockups, it is also useful to review

other existing score reports for similar testing contexts and stakeholders, which can provide both good and bad examples in terms of layout and graphical design. For ongoing maintenance and updates to existing score reports, the current score report can serve as the preliminary mockup. However, if substantial changes have occurred to the test's context and claims, then the current score report can be used to inform designers of any program-specific constraints that must be modified. Moreover, these preliminary mockups may also be informed by any audience-centric research studies in the past that may shed some light on the NKAs for the stakeholder group of interest. This review of the literature to identify audience-specific needs should not automatically be considered a replacement for conducting a needs assessment—these prior studies may provide guidance, but it is critical that the iterative score report design and development process also include stakeholder engagement to elicit specific needs for the context and purpose of the current assessment. In addition, for any static or interactive reports that will be housed in an online portal it is important to consider the latest Web Content Accessibility Guidelines (WCAG; currently WCAG 2.1). Finally, and most importantly, these preliminary mockups and subsequent score report prototypes should be informed by data visualization principles (Hegarty, 2019; Slater et al., 2019; Tufte, 2001).

Data visualization principles

Data visualization refers to the use of graphical representations to visually communicate data be it for communicating the pre-election poll results (see <https://projects.fivethirtyeight.com/2020-election-forecast/>), or the rate of disease spread such as with COVID 19 (see <https://covid19.healthdata.org/united-states-of-america?view=total-deaths&tab=compare>), or for communicating results from scientific experiments and large-scale educational assessments (Hambleton and Zenisky, 2013). These data can range from qualitative data (e.g., obtained from brainstorming activities or focus groups) to quantitative data (e.g., obtained from large-scale surveys and represented as a stacked bar chart as shown in Fig. 2).

When used as a tool for communicating information, data visualizations should be able to communicate the key message in a nutshell and be presented in a manner that aids the user's understanding of the data. For example, in Fig. 2, the main pattern to be observed is the high proportion of aging population in these first-world countries suggesting a possible decrease in birthrate. All other information, such as the actual proportions represented by each bar of the stacked bar chart is irrelevant to the main message, and would result in visual crowding, and potentially confuse the recipient of the data.

A lot of the work by Howard Wainer in the late 1990s (e.g., Wainer, 1997; Wainer et al., 1999) was focused on alternative visual representations of quantitative data that are intended to simplify the visual load, and aid in the comprehension of the illustrated relationships for the intended stakeholder. These explorations of alternative visual displays of quantitative information were responsible for sparking the increased interest in data visualization issues within score reporting. In addition, the influential work by Edward Tufte (2001) on the "Visual Display of Quantitative Information" serves as a foundational source of reference to score report designers and researchers alike in informing the design of quantitative data visualizations for score reporting.

Data visualizations should be informed by principles derived from graphic design and user-interface design (Evergreen and Metzner, 2013; Slater et al., 2019) and principles derived from cognitive science (Hegarty, 2019) that help delineate how people understand and reason about data. Slater et al. (2019) provide a list of principles that are specifically intended to guide the score report design process that include highlighting the key information that can be communicated at first glance and creating a visual hierarchy that guides the user as they review the report. Hegarty (2019) identifies a set of principles, informed by cognitive science and human information processing, that could be considered as heuristics for the design of data visualizations. She emphasizes the importance of presenting "no more or no less information" (p. 27) than is needed for the effective communication of the relevant

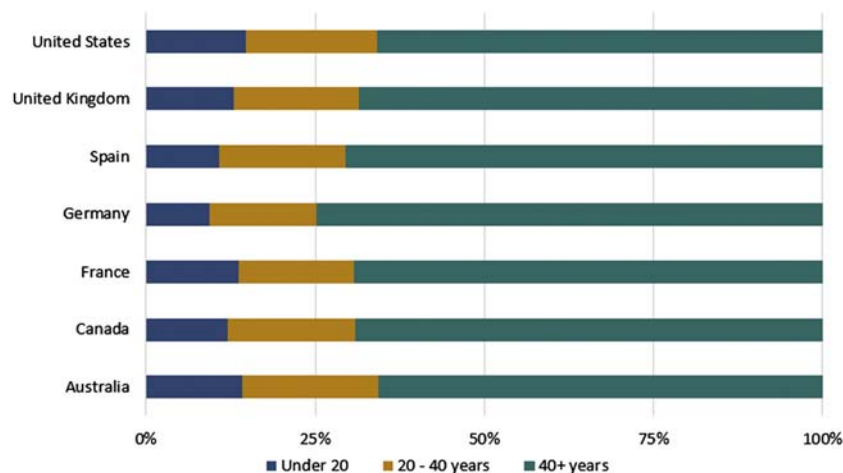


Fig. 2 An example of a quantitative data visualization (i.e., a stacked bar chart) which shows the proportion of population by age bands across select OECD countries in 2012. Data extracted from <https://stats.oecd.org/>.

message. In addition, [Hegarty \(2019\)](#) emphasizes the importance of ensuring that these representations are accurately interpreted. This highlights the importance of evaluating the interpretation and use of these visual displays which is highlighted in Step 4 of the iterative multistep framework.

Development of score report mockups and prototypes

In the context of reporting from large-scale assessment data, we are often interested in data visualizations that help to communicate results from student performance on standardized assessments to a variety of stakeholders. The users of a score report are typically neither experts of statistics nor experts in graphic design conventions and thus are not able to digest complex visual representations of technical and statistical information. In many cases, the users that receive the score reports do not have knowledge of the assessment context or content domain being assessed. Therefore, it is important to keep the pre-existing knowledge and background of the intended audience in mind as the score report mockups and prototypes are developed. A couple of hypothetical examples of such audience-centric score report mockups are presented in [Figs. 3 and 4](#).

A snapshot of a hypothetical student's score on a K-12 end-of-year summative assessment as would be typically presented in an individual student score report (or ISR) is shown in [Fig. 3](#). The intended stakeholder or recipient of the ISR is typically the student's parent or, in higher grade levels, the test-taker or student themselves. In this context and for these stakeholders, the most important pieces of information that should be highlighted in the score report are the student's score on the test, their performance level, any available normative comparisons (such as comparisons to average performance or percentile ranks), and any associated interpretive information that helps these users understand and appropriately use the test results (see [Kannan et al., 2018a](#)). As such, in this hypothetical example, the student's overall test score on the end of year mathematics test—as the most important piece of information—is highlighted prominently in this report. It is also easy to glean at a glance that there are three performance levels into which students can be classified and that this particular student has scored in the highest performance level (i.e., exceeds expectations). Additional information that is also likely to interest parents such as the student's percentile rank (or comparative performance) is highlighted. Finally, required interpretive information about the score reliability (i.e., the bar around the student's score) is also provided to ensure appropriate interpretation and use of the score results.

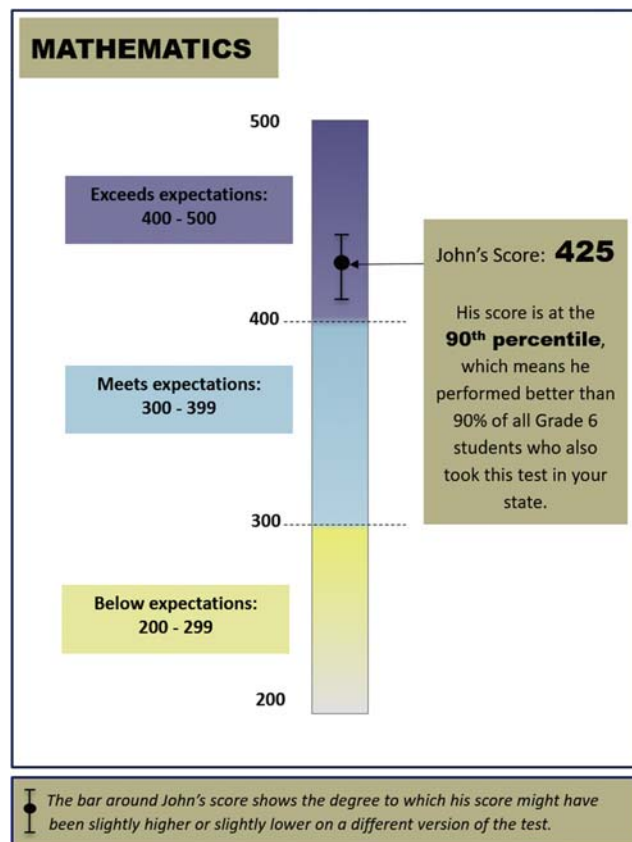


Fig. 3 Hypothetical example of the visual representation of an individual student's performance results on a K-12 end-of-year standardized assessment.

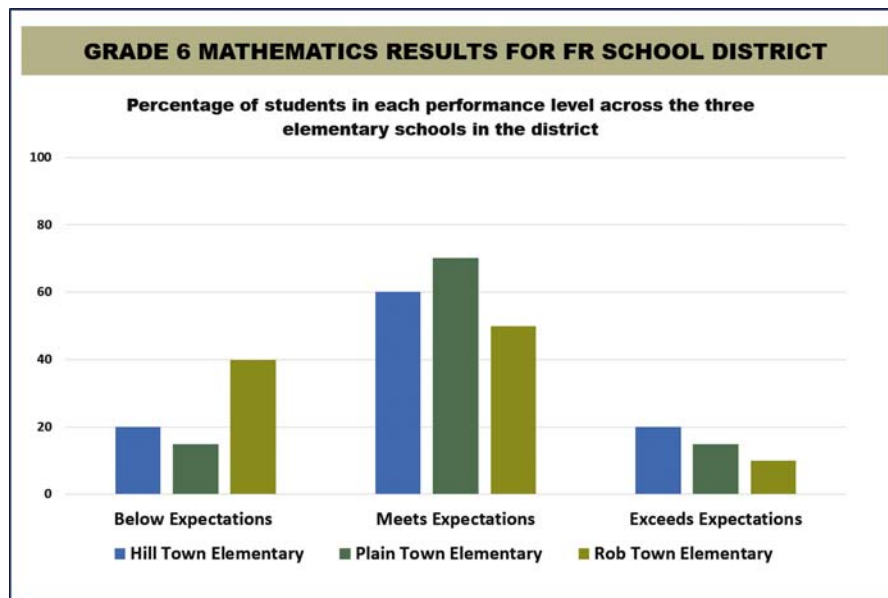


Fig. 4 Hypothetical example of the visual representation of a district report on a standardized assessment that shows the percentages of students in each performance level across the three elementary schools in the district.

In contrast, one slice of a hypothetical district-level performance report on a K-12 end-of-year summative assessment is shown in Fig. 4. The intended stakeholder or user of this information is likely a district administrator or data analyst in the district superintendent's office. District administrators are likely to have various questions as they explore district-level performance on a summative assessment (see Kannan et al., 2017; Underwood et al., 2010). The data visualization presented in Fig. 4 is but one of many slices of the district-level data. One critical need for a district administrator is to identify patterns in the data that inform their resource allocation decisions (see Kannan et al., 2017). The data presented in Fig. 4 are intended to help a district administrator identify patterns of performance across three elementary schools in a hypothetical district. The administrator may review these data to make appropriate resource allocation decisions and perhaps move some experienced teachers to Rob Town Elementary and focus on some targeted professional development for teachers from this school where the students may be struggling to "meet expectations."

In summary, it might seem counter-intuitive to design mockups before conducting an audience-focused needs assessment. However, in order to be able to orient the study participants to the purpose of the assessment, and provide some possible solutions to the needs that are elicited to see what works and what does not work, it is recommended that some preliminary mockups are designed ahead of the stakeholder needs assessments. This methodology is borrowed from user experience research (Kinzie et al., 2002; Kujala, 2008; Silver, 2005) wherein preliminary solutions are presented after eliciting needs in a user-focused needs assessment, and users are asked to evaluate the preliminary solution to determine what works and what does not. In the audience-centric focus that has been followed in score report design and development (e.g., Kannan et al., 2018b, 2019, 2021a), a similar methodology is followed where preliminary mockups are presented after stakeholder-specific needs are elicited, and participants are asked to evaluate these preliminary mockups for what they like, dislike, and find confusing.

Conduct user-centered needs assessments

A "needs assessment" is defined as a systematic process for addressing the gaps between "what is" and "what should be" (p. 3)—in other words, identifying the discrepancies between what is currently available and what is desired (Altschuld and Kumar, 2010). Since score reports play such a critical role within the validity argument for a test, reports should be designed to address the gaps in terms of stakeholders needs, so that they are able to interpret and use the assessment results accurately and appropriately. Moreover, it is clear from score reporting research, that each stakeholder group has unique needs, necessitating user-specific needs assessments for each stakeholder group.

Using focus groups and surveys to conduct user-centered needs assessments

Focus groups can serve as an economical and efficient way of getting in-depth feedback from multiple participants in a setting where they do not feel threatened (Onwuegbuzie et al., 2009). In the context of score report design, focus groups serve as an efficient method to obtain a deep understanding of stakeholder needs by delving into questions about their current experiences and pain points, before eliciting their context-specific needs, and subsequently having them react to some preliminary mockups. For each stakeholder group, it is advisable to schedule and conduct at least 3–5 separate focus group meetings with about 5–8 participants each in order to achieve data saturation—that is, until no additional new categories are identified in subsequent focus groups.

(Onwuegbuzie et al., 2009). In some previous studies (e.g., Kannan et al., 2021a), up to six separate focus groups have been conducted with the intended stakeholder group (i.e., teachers) in order to achieve data saturation.

A sample schedule of activities (as implemented in Kannan et al., 2018b, 2019, 2021a) is presented in Fig. 5. Five broad types of activities, starting from a pre-meeting survey and ending with a post-meeting survey, are recommended in conducting a user-centered needs assessment. In the current environment (i.e., during and post COVID-19), it is also important to note that all the activities detailed within this section can be conducted remotely via a combination of pre-meeting and post-meeting surveys conducted online, and a virtual focus group meeting conducted via a platform such as Zoom (<https://zoom.us/>). Virtual meetings can help with increasing the diversity of study participants by expanding access to participants in different locations; however, it may also serve to exclude those without access to reliable internet—though, this concern has become less ubiquitous with increasing access to high-speed internet globally. The pre- and post-meeting surveys help supplement the focus groups in identifying gaps ahead of the meeting and prioritizing needs at the end of the meeting. In this section, I will provide a broad overview of the activities listed in Fig. 5.

Pre-meeting survey

As a first step, previous studies (Kannan et al., 2018b, 2019, 2021a) have recommended the use of online pre-meeting surveys to elicit participant feedback. The purpose of this pre-meeting survey is to get a general idea of the types of reports that stakeholders currently receive and what they may like or dislike about these reports in order to identify some pain points and gaps. It is a good idea to send participants a link to the survey about 2–3 weeks prior to the scheduled focus group and request responses at least 3–5 days before the meeting to generate some conversation starters for the first whole-group discussion.

Whole-group discussion

As indicated before, participants may be convened for either an in-person or virtual focus group. Once the participants have been provided with a basic introduction to the assessment, its claims, purpose, and context, this whole group discussion can be used to set the stage for the overall needs assessment. A structured protocol may be used for this whole group discussion to identify the types of reports that stakeholders currently receive, what they like and dislike about the current reports, and the types of feedback they would desire and consider most useful in this context. The responses provided to the pre-meeting survey should be analyzed ahead of time for themes and used as conversation starters during this discussion.

In addition, participants may be called upon to comment about specific types of reports or report slices that have been used in the past and identified based on either a review of the literature or a review of other existing reports. For example, in a focus group with teachers, they may be specifically asked to react to the need and importance of various types of feedback typically found in class-level reports, such as rosters with tabular information, or distributions of student scores based on strengths and weaknesses.

Overall, the goal of this discussion should be to elicit a list of needs based on some of the gaps identified in terms of what they currently have and what they would ideally desire in score reports. The focus of the structured prompts should be to identify gaps based on intended uses and decisions for the stakeholder group, and thereby identifying the types of data or supporting information

Sample Schedule of Activities for a Needs Assessment Focus Group



I. Pre-meeting survey

Used to get a **general idea of current experiences** and generate some conversation starters during the focus group.



II. Whole-group discussion

Used to **build on the responses provided in the pre-meeting survey**, and facilitate a conversation around what types of reports they currently receive and what types of feedback they would value.



III. Individual work

Used as a **proxy for a problem/solution interview** and elicit individual needs – where needs are elicited, mockups are presented as one potential solution, and then mockups are evaluated



IV. Small group discussion

Used to explicitly **identify needs not addressed** in the mockups.



V. Post-Meeting Survey

Used to **prioritize and rank-order needs**

Fig. 5 Sample schedule of activities in a user-centered needs assessment focus group study [as implemented in Kannan et al., 2018b, 2019, 2021a].

that can be provided in the score reports to specifically support those uses and decisions. During the whole group discussion, a structured call-out format is recommended, where each participant is called upon to respond in a systematic pre-determined order so that no one person goes first or goes last always, and everyone gets a chance to weigh-in on the discussion.

Individual work

The overall goal with the individual work is to be able to evaluate the preliminary mockups and to identify specific rank-ordered needs for each individual participant and see if some needs are seemingly more important than others. This part of the focus-group is mainly meant to serve as a “problem-solution interview” (a terminology borrowed from UX research) where individual participant needs are identified and the extent to which the preliminary mockups serve as a solution to the identified problems (i.e., needs) is evaluated. To identify individual rank-ordered needs, in previous studies (e.g., Kannan et al., 2018b, 2019, 2021a), participants have been asked to reflect upon the preceding whole-group discussion and list their individual top 5 needs for information in the score report. Once the rank-ordered needs are elicited, the score report mockups are then presented to the entire group, and participants are asked to evaluate each report page or screen. In order to evaluate the mockups, participants are typically provided with printed screenshots of the mockups with multiple colored pens, and asked to mark up what they like, what they dislike, and what they find confusing.

In one previous study, Kannan et al. (2021a) used virtual focus groups so as to include a representative sample of teachers from across the country. In this study, an online heatmapping tool, as presented in Fig. 6, was used. Participants used the heatmapping tool to click on various areas of each report screen to indicate whether this was something they particularly liked, or disliked, or found confusing. Then, participants elaborated by providing detailed comments associated with each click. These clicks and associated comments were later analyzed and summarized to understand the extent to which the preliminary mockups served participant needs and what if anything could be changed or further clarified. In addition to having participants evaluate what they like and dislike about the current mockups, they should identify if any of their most important needs are not being addressed in the current mockups. The subsequent small-group discussions are used for this purpose.

Small-group discussion

The main purpose of the small group discussion is to consolidate participants’ individual evaluations of the mockups and identify what is missing from the mockups vis-à-vis needs that were identified both at the individual and group levels. In previous needs assessment studies (e.g., Kannan et al., 2018b, 2019, 2021a), participants have been asked to reflect upon the rank-ordered needs and consolidate these needs within their groups. Then participants use these group-level needs to evaluate the mockups and identify needs from their group-level list that have not been addressed in the preliminary mockups. Participants can be provided with flip-charts or notepads to discuss and summarize this information that can then be collected at the end of the meeting. In an online/virtual setting, breakout rooms with individual facilitators assigned to each breakout room may be used to achieve the same results.

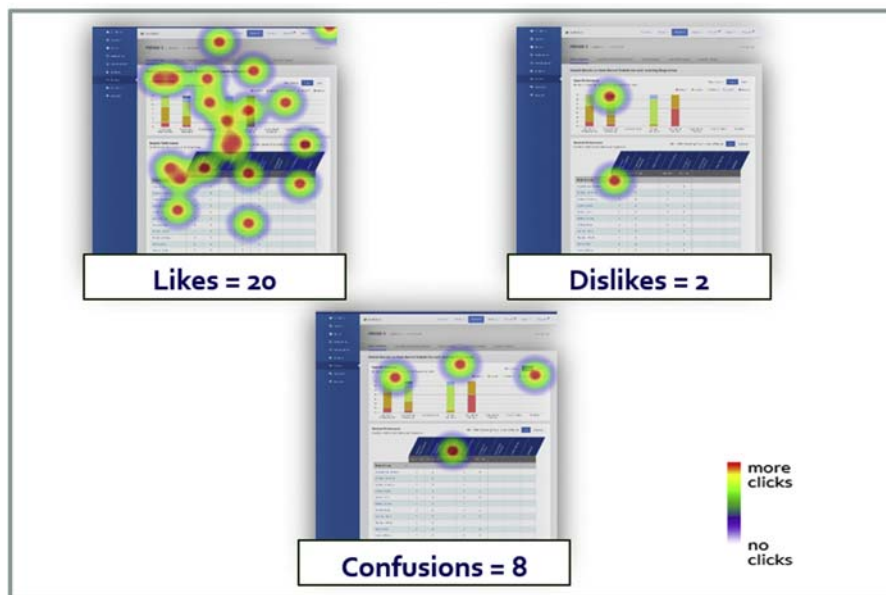


Fig. 6 Heatmap results for one of the 6 screenshots for mathematics teachers identified in Kannan et al. (2021a). Each thumbnail in this figure shows the density of clicks for likes, dislikes, and confusions indicated by participants as they evaluated one mockup screen; each of these clicks was also linked with associated comments that helped identify the report elements that were particularly liked, disliked, and confusing to the participants and informed subsequent prototype revisions.

The room facilitators can use online tools such as MS-OneNote to help capture the discussions and to identify the critical stakeholder needs that have not been addressed in the preliminary mockups.

Post-meeting survey

A variety of needs may be identified during the needs assessment focus groups, but for various practical considerations, it may not be feasible for the testing program to prioritize all of these needs in their score report prototype development. Moreover, regardless of the total number of focus groups conducted in the attempts to achieve data saturation, it should be recognized that data from focus groups are based on qualitative evaluations of small samples and may need to be validated. Therefore, the needs identified in the focus groups should be prioritized based on their importance by conducting a post-meeting survey with a larger group of stakeholders. In one previous study, [Kannan et al. \(2019\)](#) listed all the needs generated during the focus groups in an online survey using a Likert-type scale where participants rated the importance of each need for themselves and for other teachers in a similar context (e.g., ESL teachers, special education or remedial teachers, regular classroom teachers). It is important to try to include a reasonably large and diverse sample for the post-meeting survey to be able to generalize the results from the needs assessment to the larger stakeholder group. The results from the survey may then be used to rank order and prioritize needs for subsequent prototype development. For example, all needs with average importance scores of 4 or above (on a 5-point Likert) may be identified as top needs that are then summarized and provided to the designers for implementation into the next iteration of prototypes.

Develop prototypes and iterate with systematic feedback

Once the most critical needs are identified for each intended stakeholder group, the iterative prototype design process can then begin. The evaluation of the preliminary mockups and additional needs, if any, that are identified during the needs assessment are shared with the UI/UX graphic designers. The data visualization principles and WCAG compliance noted in Step 2 are again critical to consider in the design of functional prototypes. However, the most important aspect of the iterative prototype development cycle is the adoption of an agile methodology and repeated engagement with the stakeholders to evaluate that the scores, tables, or visual representations, and any additional supplementary information presented in the report prototypes are accurately interpreted and appropriately used. Therefore, during each iterative design and revision cycle, cognitive interviews and usability studies should be scheduled with the intended stakeholders to evaluate their comprehension of the presented information, their preference for alternate visual representations which are substantiated by the interpretability of each alternate representation, and the usability of any interactive interface.

Cognitive interviews and surveys to evaluate stakeholder comprehension and usability

Since a test's validity is dependent on the extent to which stakeholders can understand and interpret the information presented in score reports, it is critical to iteratively evaluate how stakeholders comprehend this information and make appropriate revisions where additional clarity may be needed. Study participants should be ensured that the comprehension questions are not intended as a test of the participant's knowledge, but rather to evaluate how well the reports have been designed to convey necessary information. Moreover, for online and interactive reporting systems, participants' ability to engage with the interface, access all relevant information, and use the interactive features designed for data manipulation, also have an influence on stakeholder comprehension of the information and should also be systematically evaluated. Evaluation of participant comprehension and usability may be done either through embedding report snapshots with associated questions in an online survey or through engaging with participants in cognitive interviews, where participants are asked to verbally respond to focused questions either during engagement in an activity or after they have completed the activity ([Keehner et al., 2016](#)).

Some previous studies (e.g., [Kannan et al., 2018a](#); [Kannan, 2020](#)) have used a hybrid cognitive interview style which combines retrospective verbal probing where participants respond to directed questions with concurrent think-aloud methods where participants verbalize their thoughts as they are interacting with the report or reporting system. These cognitive interviews are intended to identify the elements in the score report that are most salient to the stakeholders and if they are able to access and use all the information as intended in addition to evaluating if the information presented in these reports is interpreted accurately. In other studies (e.g., [Kannan et al., 2017, 2021b](#)), specific comprehension questions pertinent to the range of information provided in the reports were embedded in online surveys. These survey studies have included several questions that are quick to answer (such as multiple choice, true/false), and focused these questions on aspects of the representations that may likely be confusing given the specific knowledge and constraints of the intended stakeholder group. Participant responses to these comprehension questions then enable test developers to evaluate the extent to which the data visualizations and other information presented in the reports are being understood correctly and identify areas where additional clarity may be needed.

Surveys and interviews to evaluate stakeholder preference

In addition, it may be important to evaluate if different alternative visual representations lead to more accurate interpretation of the key information presented in the score reports. It is possible that certain representations are not as easily understood by certain stakeholder groups. For example, for parents, who may not have the necessary statistical background, simple bar and line charts may be more readily understood than more complicated representations such as box and whisker plots and bubble charts. In addition, information on stakeholder preference for one visual representation over another can be collected; however, feedback on

preference should always be substantiated by improved comprehension. Just preference of one representation does not justify its use if it does not also lead to better comprehension of the presented information.

In one previous study (Kannan et al., 2017), researchers anticipated that a preliminary design that represented sub-area performance in terms of a spider-web diagram (or radar chart) would possibly be less familiar to some stakeholder groups (i.e., teachers, principals, and district administrators). Therefore, two alternative representations for sub-area performance were designed and included for evaluation, specifically, stacked bar charts with scale scores and multiple column bar charts with percentages. Results from this study indicated that principals and district administrators had a strong preference for the multiple column bar charts with percentages, whereas teachers had a slight preference for the stacked bar charts with scale scores. Both these alternative representations resulted in significantly higher levels of comprehension than the original spider-web diagrams, which were unanimously not preferred by any of the stakeholder groups (Kannan et al., 2017).

Operational deployment

Score report prototypes should be iteratively and systematically evaluated with both internal clients and the larger stakeholder group until the interviews and surveys indicate that there are no undue misinterpretations of the reported results. There is no ground rule on the required number of iterations before score reports are finalized. This has to be guided by the data and may also be based on program directions and other operational constraints such as pre-determined test launch-date. Slater et al. (2019) provide a broad overview of the operational score report design process and identify several operational constraints that may be encountered during the iterative design cycles. Once score reports are operationally deployed, a program for on-going maintenance and continuous monitoring of the score report interpretation and use should be set in place. Large-scale surveys may be used to evaluate how stakeholders intend to use the data presented in these reports and the decisions they will make based on the reports.

Large-scale surveys to evaluate intended uses and decisions

A systematic evaluation of the uses and decisions to be made from the reported data is also important to ensure a test's validity. In other words, the extent to which stakeholders understand how the results may be used (i.e., appropriate next steps), and the types of decisions that they intend to make from the reported results should be evaluated for each stakeholder group. The uses and decisions made from the score reports should not lead to negative unintended consequences that undermine the validity of the test. One previous study (Kannan et al., 2017) specifically evaluated stakeholder use and decisions made from the data presented in reports by embedding directed questions about intended use and decisions in an online survey with large samples of teachers, principals, and district administrators. Results from this survey helped the testing program ensure that appropriate and warranted decisions were being made based on the reported results. Similarly, all testing programs should continue to evaluate if the operational reports are functioning as intended and that the results presented in the reports are being used in the manner intended. If there are indications of inappropriate use, actions should be taken to revise and update the reports to clearly present information about intended uses so that the program can mitigate any unintended consequences from inappropriate use of test results.

In addition, state departments and testing programs should continuously monitor that reports continue to remain useful over time as changes are made to the test or the testing context. There may be changes initiated to the ways in which the data from the test are analyzed and scored, which results in the need to communicate results from new models. There may be new federal or state mandates that require the reporting of new metrics, for example to report student growth across the years. Each of these types of changes may result in the need to report either new metrics or include additional visual representations or both. Such additions/revisions to the existing score report should be evaluated for interpretability and appropriateness for each stakeholder group. Moreover, tests are often revised over time, and if these revisions are substantial wherein they impact the underlying claims based on the test, then the entire iterative design and development process starting with a needs assessment should again be undertaken. For on-going program maintenance, the iterative process can directly start in Step 3, and the current score reports can serve as preliminary mockups to identify gaps and conduct a needs assessment with the intended stakeholder group(s). Regardless, and for a multitude of reasons only some of which are listed here, it is important to establish a program of on-going maintenance and monitor if the score reports continue to communicate test results to the intended stakeholders in a manner that fosters appropriate interpretation and use.

Future of score reporting research

Score reports are the “bridge” (p. 9, Tannenbaum, 2019) that help link what is measured in a test to how stakeholders use that information. Therefore, the information communicated through the score reports should be iteratively evaluated to ensure that stakeholders are able to interpret and use the information in appropriate ways. In this chapter, we reviewed research that has been conducted on audience-centric score reports and how specific audiences interpret and use score reports. We also reviewed an iterative multistep framework with suggested research-based methods for the design and evaluation of score reports. This framework provides a clear, research-based, set of steps with a guidance on specific methods to follow in developing and evaluating score reports for new and existing testing programs. Though a number of applied research studies have been conducted in this context, additional research should be conducted that focuses on supporting stakeholder interpretation and use, particularly around sub-

score reporting and growth metrics (Zenisky and Hambleton, 2012). But overall, interest around interactive reporting systems has been burgeoning and this focus for future research would be very apropos given the current shift toward a predominantly virtual world.

Interactive reporting systems

Research in score reporting has largely focused on static score reports, whether these are printed score reports or accessed through an online portal. National and International assessment programs such as the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) have pioneered in the area of dashboard reporting and interactive reporting systems. In recent years, there has been an increasing momentum in the development of interactive reporting systems in the context of various large-scale K-12 assessment programs, especially for some stakeholders such as teachers and administrators, who have to review group-based data and make instructional or administrative decisions based on data available in score reports. Moreover, there has also been some interest in developing and housing interactive score reports with various other supporting and supplementary information even for stakeholders such as parents so that they may access all assessment related information in a single location such as a parent portal.

These interactive reporting systems have a “high degree of user choice” (p. 25) in the ways in which stakeholders access and obtain the information (Zenisky and Hambleton, 2012). These reporting systems are typically housed in websites where users arrive at a landing page, and then make a series of decisions about the type of data they want to explore and the order in which they want to explore these data. Therefore, these interactive reporting systems should offer a high degree of flexibility where each user can take their own unique path to explore the system. However, it is important to ensure that each slice of the reporting system is developed with a keen focus on audience-specificity and ensuring that these slices do not lead to unwarranted interpretations and uses for the stakeholders.

There are a variety of issues to consider in designing and developing such reporting systems. In essence, every slice or view of the reporting system may be considered a pre-canned report that is developed using the iterative framework. The important slices that are relevant to each stakeholder group can be identified through a needs assessment, and question-based reporting (VanWinkle et al., 2011) may help target users to specific slices that clearly answer a user-based need within the score report. In addition, higher-level decisions about what the user would want to see first and the ability for the user to navigate back and forth from each screen or reporting slice to any other location in the system should be carefully considered. Research in the area of software and web content development and user interface design can largely inform the research on developing reporting systems. These reporting systems, dashboards, and portals should be carefully designed and evaluated by applying and extending the iterative multistep framework to the design and evaluation of reporting systems. Some of the early work around the evaluation of NAEP dashboards (Zenisky et al., 2009) is a great start in this area, but additional work is needed.

In addition to the need for research on the design, development, and evaluation of interactive reporting systems, there are several other emerging areas of research within score reporting. These include: score reporting for formative purposes (Kannan et al., 2018b; Zapata-Rivera et al., 2018), score reporting of process data (e.g. Zapata-Rivera et al., 2018), research around data analytics and dashboards (e.g., Bodily et al., 2018; Corrin, 2019), and work on visualizing information from adaptive instructional systems (e.g., Zapata-Rivera, 2019b). Additional research in each of these areas is warranted, but continuing to understand the extent to which stakeholders are able to interpret and use the score reports in each unique assessment context continues to remain important. Overall, though score reporting research has come a long way from being one of the most under-researched areas in measurement about 10 years ago (Hambleton and Zenisky, 2013), much still remains to be understood.

References

- Altschuld, J.W., Kumar, D.D., 2010. Needs Assessment: An Overview. SAGE Publications Inc, Thousand Oaks, CA.
- American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME), 2014. Standards for Educational and Psychological Testing. AERA, Washington, DC.
- Barber, B.L., Paris, S.G., Evans, M., Gadsden, V.L., 1992. Policies for reporting test results to parents. *Educ. Meas.: Issues Pract.* 11 (1), 15–20.
- Bennett, R.E., 2016. Opt Out: An Examination of Issues (Research Report No. RR-16-13). Educational Testing Service, Princeton, NJ.
- Betebenner, D.W., 2009. Norm- and criterion-referenced student growth. *Educ. Meas.* 28 (4), 42–51.
- Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., Verbert, K., 2018. Open learner models and learning analytics dashboards: a systematic review. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. Association for Computing Machinery, pp. 41–50.
- Brown, G.T.L., O’Leary, T.M., Hattie, J.A.C., 2019. Effective reporting for formative assessment: the asTTle case example. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, NY, pp. 107–125.
- Corrin, L., 2019. Evaluating students’ interpretation of feedback in interactive dashboards. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, pp. 145–159.
- Evergreen, S., Metzner, C., 2013. Design principles for data visualization in evaluation. In: Azzam, T., Evergreen, S. (Eds.), *Data Visualization*, vol. 140. New Directions for Evaluation, pp. 5–20 part 2.
- Goodman, D.P., Hambleton, R.K., 2004. Student test score reports and interpretive guides: review of current practices and suggestions for future research. *Appl. Meas. Educ.* 17 (2), 145–220.
- Hambleton, R.K., Slater, S., 1997. Are NAEP Executive Summary Reports Understandable to Policy Makers and Educators? (CSE Technical Report 430) National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.

- Hambleton, R., Zenisky, A., 2013. Reporting test scores in more meaningful ways: a research-based approach to score report design. In: *APA Handbook of Testing and Assessment in Psychology*. APA, Washington, DC.
- Hegarty, M., 2019. Advances in cognitive science and information visualization. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, NY, pp. 19–34.
- Hubley, A.M., Zumbo, B.D., 2011. Validity and the consequences of test interpretation and use. *Soc. Indic. Res.* 103, 219–230.
- Impara, J.C., Divine, K.P., Bruce, F.A., Liverman, M.R., Gay, A., 1991. Does interpretive test score information help teachers? *Educ. Meas.* 10 (4), 16–18.
- Kane, M.T., 2006. Validation. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Praeger, Westport, CT, pp. 17–64.
- Kane, M., 2013. Validating the interpretation and use of test scores. *J. Educ. Meas.* 50, 1–73.
- Kannan, P., Bryant, A.D., Shao, S., 2017. Evaluating needs, preferences, intended uses and decisions from data presented in an online reporting system for teachers, principals, and district administrators. In: *Final Report Submitted to the Testing Program to Inform the Development of a Reporting System for a Summative Assessment Suite*. Princeton, NJ.
- Kannan, P., Zapata-Rivera, D., Leibowitz, E.A., 2018a. The interpretation of score reports by diverse subgroups of parents. *Educ. Assess.* 23 (3), 173–194. <https://doi.org/10.1080/10627197.2018.1477584>.
- Kannan, P., Zapata-Rivera, D., Mikeska, J., Bryant, A.D., Long, R., Howell, H., 2018b. Providing formative feedback to pre-service teachers as they practice facilitation of high-quality discussions in simulated mathematics and science methods classrooms. In: Langran, E., Borup, J. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference*, pp. 1570–1575.
- Kannan, P., Beigman-Klebanov, B., Shao, S., Long, R., 2019. Evaluating teachers' needs for on-going feedback from a technology-based book reading intervention. In: *Paper Presented at the Annual Meeting of the National Council on Measurement in Education*, Toronto, ON.
- Kannan, P., Bryant, A.D., Shao, S., Wylie, E.C., 2021a. Identifying Teachers' Needs for Results from Interim Unit Assessments (Research Report No. RR-21-08). ETS. <https://doi.org/10.1002/ets2.12320>.
- Kannan, P., Zapata-Rivera, D., Bryant, A.D., 2021b. Evaluating parent comprehension of measurement error information presented in score reports. *Pract. Assess. Res. Eval.* 26 (12). <https://scholarworks.umass.edu/pare/vol26/iss1/12>.
- Kannan, P., 2020. Supporting the interpretation of summative assessment results for parents from underserved groups. In: *Final Report Submitted to the Testing Program to Inform the Development of a Parent Portal with Supporting Resources in Two Languages for a Summative Assessment Suite*. Princeton, NJ.
- Keenhner, M., Gorin, J.S., Feng, G., Katz, I.R., 2016. Developing and validating cognitive models in assessment. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. John Wiley & Sons, West Sussex, UK, pp. 75–101.
- Kinzie, M.B., Cohn, W.F., Julian, M.F., Knaus, W.A., 2002. A user-centered model for website design: needs assessment, user interface design, and rapid prototyping. *J. Am. Med. Inf. Assoc.* 9 (4), 320–330.
- Kujala, S., 2008. Effective user involvement in product development by improving the analysis of user needs. *Behav. Inf. Technol.* 27 (6), 457–473.
- Marshall, B., Drummond, M.J., 2006. How teachers engage with assessment for learning: lessons from the classroom. *Res. Pap. Educ.* 21, 133–149.
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A Brief Introduction to Evidence-Centered Design (Research Report No. RR-03-16). Educational Testing Service, Princeton, NJ.
- O'Leary, T.M., Hattie, J.C., Griffin, P., 2017. Actual interpretations and use of scores as aspects of validity. *Educ. Meas.* 36 (2), 16–23.
- Onwuegbuzie, A.J., Dickinson, W.B., Leech, N.L., Zoran, A.G., 2009. A qualitative framework for collecting and analyzing data in focus group research. *Int. J. Qual. Methods* 8 (3), 1–21. <https://doi.org/10.1177/160940690900800301>.
- Silver, M.A., 2005. *Exploring Interface Design*. Thompson/Cengage Learning, Clifton Park, NY.
- Sireci, S.G., Wells, C.S., Bahry, L., 2013. Student Growth Percentiles: More Noise Than Signal? Paper presented at the annual meeting of the National Council on Measurement in Education (NCME). San Francisco, CA.
- Slater, S., Livingston, S.A., Silver, M., 2019. Score reports for large-scale testing programs: managing the design process. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, NY, pp. 91–106.
- Tannenbaum, R.J., 2019. Validity aspects of score reporting. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, NY, pp. 9–18.
- Tufte, E.T., 2001. *The Visual Display of Quantitative Information*, second ed. Graphics Press, Cheshire, CT.
- Underwood, J.S., Zapata-Rivera, D., VanWinkle, W., 2007. Growing Pains: Teachers Using and Learning to Use IDMS (Research Memorandum No. RM-08-07). Educational Testing Service, Princeton, NJ.
- Underwood, J.S., Zapata-Rivera, D., VanWinkle, W., 2010. An Evidence-Centered Approach to Using Assessment Data for Policymakers (Research Report No. RR-10-03). Educational Testing Service, Princeton, NJ.
- VanWinkle, W., Vezzu, M., Zapata-Rivera, D., 2011. Question-Based Reports for Policymakers (Research Memorandum No. RM-11-16). Educational Testing Service, Princeton, NJ.
- Vezzu, M., VanWinkle, W., Zapata-Rivera, D., 2012. Designing and Evaluating an Interactive Score Report for Students (Research Memorandum No. RM-12-10). Educational Testing Service, Princeton, NJ.
- Wainer, H., Hambleton, R.K., Meara, K., 1999. Alternative displays for communicating NAEP results: a redesign and validity study. *J. Educ. Meas.* 36 (4), 301–335.
- Wainer, H., 1997. Improving tabular displays: with NAEP tables as examples and inspirations. *J. Educ. Behav. Stat.* 22 (1), 1–30.
- Zapata-Rivera, D., Katz, I., 2014. Keeping your audience in mind: applying audience analysis to the design of score reports. *Assess. Educ. Princ. Pol. Pract.* 21 (4), 442–463.
- Zapata-Rivera, D., VanWinkle, W., Zwick, R., 2010. Exploring effective communication and appropriate use of assessment results through teacher score reports. In: *Paper Presented at the Annual Meeting of the National Council on Measurement in Education*, Denver, CO.
- Zapata-Rivera, D., VanWinkle, W., Zwick, R., 2012. Applying Score Design Principles in the Design of Score Reports for CBAL™ Teachers (Research Memorandum No. RM-12-20). Educational Testing Service, Princeton, NJ.
- Zapata-Rivera, D., Vezzu, M., Nabors Olah, L., Leusner, D., Biggers, K., Bertling, M., 2014. Designing and evaluating score reports for parents who are English Language Learners. In: *Paper Presented at the Annual Meeting of the American Educational Research Association (AERA)*, Philadelphia.
- Zapata-Rivera, D., Zwick, R., Vezzu, M., 2016. Exploring the effectiveness of a measurement error tutorial in helping teachers understand score report results. *Educ. Assess.* 21 (3), 215–229. <https://doi.org/10.1080/10627197.2016.1202110>.
- Zapata-Rivera, D., Kannan, P., Forsyth, C., Peters, S., Bryant, A.D., Guo, E., Long, R., 2018. Designing and evaluating reporting systems in the context of new assessments. In: Schmorow, D., Fidopiastis, C. (Eds.), *Augmented Cognition: Users and Contexts*. AC 2018, Lecture Notes in Computer Science, vol. 10916. Springer, Cham, pp. 143–153.
- Zapata-Rivera, D., Kannan, P., Zwick, R., 2019. Communicating measurement error information to teachers and parents. In: Zapata-Rivera, D. (Ed.), *Score Reporting Research and Applications*. Routledge, New York, NY, pp. 63–74.
- Zapata-Rivera, D. (Ed.), 2019a. *Score Reporting Research and Applications*. Routledge, New York, NY.
- Zapata-Rivera, D., 2019b. Supporting human inspection of adaptive instructional systems. In: Sottile, R., Schwarz, J. (Eds.), *Adaptive Instructional Systems*. HCII 2019, Lecture Notes in Computer Science, vol. 11597. Springer, Cham, pp. 482–490.
- Zenisky, A.L., Hambleton, R.K., 2012. Developing test score reports that work: the process and best practices for effective communication. *Educ. Meas.* 31 (2), 21–26.
- Zenisky, A.L., Hambleton, R.K., Sireci, S.G., 2009. Getting the message out: an evaluation of NAEP score reporting practices with implications for disseminating test results. *Appl. Meas. Educ.* 22 (4), 359–375.
- Zwick, R., Zapata-Rivera, D., Hegarty, M., 2014. Comparing graphical and verbal representations of measurement error in test score reports. *Educ. Assess.* 19 (2), 116–138.

Test translation and adaptation

Juliette Lyons-Thomas^a and Paulina Biernacki^b, ^a Educational Testing Service, Princeton NJ, United States; and ^b Stanford University, Palo Alto, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	230
The importance of translation and adaptation	230
Translation and adaptation defined	231
Validity and construct irrelevance	231
The process of translation	231
Poor translation and adaptation	232
Inappropriate uses of translation and adaptation	232
ITC guidelines	233
Statistical measures of comparability	234
Conclusions and future implications	234
References	235

Introduction

In the last fifty years, international migration of people living in a country other than their birth country has grown steadily. In 2019, the number of international migrants grew globally to 272 million, or 3.5% of the world's population, up from 84 million or 2.3% in the year 1970 ([International Organization for Migration, 2019](#)). A recent Pew Research Center analysis found that individuals in the United States who were born after the year 2000 are the most racially and culturally diverse yet ([Fry and Parker, 2018](#)). These changes in diversity greatly impact student populations and place unique demands on systems of education. Students categorized as English Language Learners (ELLs) – those who are learning content in English while also gaining proficiency in this language ([Valdés, 2004](#)) are the fastest-growing segment of the school-aged population ([Grantmakers for Education, 2013](#); [National Education Association, 2015](#)).

For decades, educators have been concerned that differences in English language proficiency skills may contribute to measurement error on assessments, with a wide range of potential consequences for ELLs' educational programming and opportunities to learn. Language proficiency has been shown to greatly affect student performance on academic assessments, even when language is not the focal construct of the assessment ([Abella et al., 2005](#)). ELLs tend to perform lower than native English speakers on standardized tests that are offered in English, especially on tests with higher language demands ([Abedi, 2002, 2009](#)). Lower test scores may be misinterpreted as evidence of academic deficits, resulting in placement into remedial courses ([Callahan and Shiffrer, 2016](#); [Kanno and Kangas, 2014](#)) and potentially incorrect special education diagnoses. In fact, educators and researchers have been alarmed at the disproportionate overrepresentation of ELLs into special education programs and underrepresentation in gifted and talented programs, relative to ELLs' representation in the overall student population ([Abedi, 2009](#); [Donovan and Cross, 2002](#); [Ford et al., 2008](#); [Sullivan, 2011](#)). It is thus vital that test developers and users attend to linguistic factors in assessment, which include translation and adaptation processes.

In situations where the focal construct is unrelated to language proficiency, a translation of the assessment may be an appropriate step in ensuring the validity of the assessment results. For example, in New York State, a number of statewide Regents exams are translated from English into the top five home languages of the state (i.e., Chinese, Haitian Creole, Korean, Russian, and Spanish). These exams are eligible for translation because the constructs being measured, such as Algebra or Global History, are considered unrelated to English language proficiency. In the international context, large scale assessments of student proficiency (e.g., the Programme for International Student Assessment) are often translated into various languages so that students' academic skills are assessed in their home language, and subsequently their performance on a construct can be compared across regions. Beyond the translation of an assessment from one language to another, assessments may also need to be adapted for cultural purposes so that test questions are appropriate for the students' context. This article describes these types of situations and processes, and in doing so, contributes to an understanding of the complexity of translation and adaptation of assessments for ELL students.

The importance of translation and adaptation

When test items include high linguistic demands, language factors may prevent students from demonstrating their domain-relevant knowledge, thus introducing sources of systematic error into students' test scores ([Abedi, 2014](#); [Messick, 1994](#)). When students from a wide variety of linguistic backgrounds are assessed using the same assessment instrument, the test questions may be adapted so that students can experience the questions in a comparable way. The true comparability of test questions is a central concern because test results are often used for making important decisions based on performance. For instance, using the examples described

above, New York State uses their Regents exams for the high stakes purpose of graduation. Students must successfully pass five Regents exams in order to graduate with a New York State Regents diploma. However, given the diversity of home languages of students from across New York, translated assessments are one effort toward promoting equitable graduation criteria. International assessments are another example, wherein students from across the globe are assessed so that the educational practices of various countries and regions can be compared. Policy decisions are frequently made based on outcomes from these international assessments (Carnoy, 2015). Clearly, it is imperative that linguistic and cultural factors are considered and addressed so that assessment results can be interpreted in an appropriate manner, both locally and internationally.

Translation and adaptation defined

It is important to distinguish between test translation and adaptation for the purposes of cross-cultural and linguistic assessment before moving on to the other sections of this article. Test translation refers to converting test items and associated text, such as directions or tutorials, from one language into a second target language. Adaptation refers to the broader practice of ensuring the appropriateness of various test forms, with the goal of achieving true comparability. In the context of culturally and linguistically diverse test takers, adaptation encompasses traditional translation activities; however, it may also include additional actions such as the modification of test layout to fit a new language or culture (e.g., formatting text so that reads from left to right) and determining whether two tests measure the same constructs equivalently after adjustments are made for a new language or culture.

Adaptation is sometimes needed when translation is not necessary, such as when two populations use the same language, but have different meanings associated with some words or phrases. There are many examples from languages that are used in more than one region. Consider, for instance, that Quebec French spoken in Canada varies from Metropolitan French which is spoken in France. In Metropolitan French, the word for bicycle is *vélo*, however the Quebec French equivalent is *bicyclette*. Another example is how lemon and lime are translated into Spanish, or whether the two types of fruit are distinguished at all. Lime can be referred to as *limon verde*, *limón*, or *lima*, depending on the locale of the speaker. However, *limón* or *lima* is also used to refer to lemons in some Spanish-speaking countries. For instance, requesting a *limón* in Chile will result in the large yellow citrus fruit, whereas asking for a *limón* in the Dominican Republic will produce the smaller green fruit. In each country, the alternative citrus fruit (that is, a lime in Chile or a lemon in the Dominican Republic) is rarely found, which can further complicate the issue of differentiating between the two. These examples of French and Spanish lexicons demonstrate that test adaptation may be a necessary step beyond full text translation because the meaning of words may differ across regions that share the same language.

Validity and construct irrelevance

Validity refers to the degree to which empirical evidence supports the appropriateness of test score interpretations (American Educational Research Association et al., 2014). Test validation involves searching for sources of both confirming and disconfirming evidence related to interpretive arguments, as well as identifying underlying assumptions that may or may not be credible in relation to the evidence collected (Haertel, 1999; Kane, 1992). Test developers collect various types of evidence to build a validity argument, including but not limited to evidence from expert judgments, test taker response processes, cognitive interviews, and statistical analyses. For instance, if the purpose of a test is to make a judgment about a student's mathematical skills, a test developer would collect validity evidence to support the conclusion that the test score is in fact a reflection of students' mathematical knowledge.

Construct irrelevant variance is "the degree to which test scores are affected by processes that are extraneous to the test's intended purpose" (American Educational Research Association et al., 2014, p. 12). Language ability is a potential source of construct irrelevant variance that may artificially deflate ELLs' scores on assessments. For example, a test would be affected by construct irrelevant variance if the test were designed to measure mathematics ability, despite inadvertently measuring language proficiency due to high linguistic demands in the wording of test items. The test validation process may reveal that test takers' language abilities may affect the test score interpretation. In this scenario, test developers may find that a translation or adaptation of a test is a useful way to reduce construct irrelevant variance. It should be noted that interpretations of a translated or adapted test also must undergo the same validation processes.

The process of translation

The actual process of translating or adapting a test may be time consuming and laborious. However, it is necessary in order to ensure the comparability of test forms. During the translation of a test, a common first step is to perform what is known as a forward translation. This procedure involves translating the test from its original language into a target language. Great care should be given to how the translators are chosen in that they should be fully bilingual in both languages, but also aware of variations in the languages that they are working with. Beyond extensive knowledge of language and context, the translators should also have knowledge about the construct that is being assessed.

A back translation may then be subsequently completed, wherein the test is translated from the target language back to the original language in which the test was developed. This step should be conducted by independent translators, that is, individuals

who did not conduct the initial forward translation. This stage would ultimately yield two versions of the same test in the original language. It can be very helpful for test developers to compare the two versions to determine if the original meaning of the test has been preserved after a back translation.

Finally, following the back translation, a harmonization step takes place. During harmonization, a bilingual panel reviews both the original and the translated test items. This panel aims to reach a consensus on the quality of the translation that took place, and whether meaning of words or concepts have remained constant during the translation process. The panel also explores how to remedy any possibly discrepancies that have emerged due to the translation process (Abubakar et al., 2013).

Poor translation and adaptation

When a test is translated from one source language into another, differences between the test forms may be unfair for a variety of reasons. These differences can include text length, complexity, or changes in actual meaning. For instance, Goodrich and Ercikan (2019) describe the example of French and English versions of the 2011 Progress in International Reading Literacy Study (PIRLS), which is an international assessment of reading for students in the fourth grade. The authors used expert reviews to examine the construct equivalence of French and English passages used in the assessment. In some passages, reviewers found differences such as longer and more confusing French text when compared to the English version, and also noted that for one passage, font size was smaller in the French version than the English version. It was noted that some words in the French version of the assessment were not commonly used by French language groups in Canada, or were more difficult than the English term that was translated. For instance, in one passage, the English term “looked like” was translated to “*apparence extérieure*” – a term that translates more closely to “external appearance” in English. This example highlights the care that must be taken to ensure the resulting translated items are truly measuring equivalent constructs.

At times, poor translations provide one group of students with more information about an item, or a hint at the correct answer, compared to what the other group of students will see. An example of this was presented in Ercikan and McCreith (2002) who also examined French and English versions of another international assessment, the Trends in International Mathematics and Science Study (TIMSS). In one translation of an item, the term “oxygen equipment” is translated to the more informative “*bouteilles d’oxygène pour respirer*”. The addition of “... *pour respirer*” inserts the additional term “for breathing” into the phrase. Another example is found in research conducted by Ercikan et al. (2010) in which an item from a Canadian nationwide assessment presented a diagram of clocks with two different times displayed. Students were then asked to measure the degrees of movement of the minute hand. The English version of the question asked students, “How many degrees does the *minute-hand* move from 3:25 pm to 3:45pm?” whereas the French version asked students “De combien de degrés *la grande aiguille* s’est-elle déplacée de 15 h 25 à 15 h 45?”. In this situation, the term “minute-hand” was translated to “*la grande aiguille*”, which back-translates into “long hand”. In order to correctly answer the question, students taking the English-version of the assessment would need to first know that the minute hand was the longer hand of the two hands of a clock, and then calculate the degrees of movement of that hand. The French students are required to take one less step to answer the question, given that the stem of the question indicates to the students the clock hand that they need to examine.

While comparability of two language versions of a test is a primary goal during test translation, one area of research has highlighted how traditional translation procedures may focus too narrowly on the equivalence of the translated versions. With the potential to be heavily influenced by confirmation bias, scholars argue that it is critical to focus on actively searching for translation issues throughout the entire test development process. The Theory of Test Translation Error (TTTE) was proposed to address these important issues (Solano-Flores et al., 2009). According to TTTE, translation errors are observed when the original and translated test versions differ in terms of design, language and content. Such errors are inevitable, given that each language encodes meaning differently. By actively anticipating such issues, the goal is to minimize translation errors to the greatest extent possible (Solano-Flores et al., 2009).

Inappropriate uses of translation and adaptation

As noted above, translations may be an appropriate step to take if the assessment content can be translated into a target language or context without altering the construct. However, there are scenarios where translations may not be appropriate. For example, the translation of an assessment measuring English language proficiency into any home language would not produce results that can be used in an appropriate manner, because the construct of measurement would be altered. English language proficiency assessments are often administered to students to determine the appropriate level of English language services to which a student is entitled. Therefore, translating any section of this type of assessment into a student’s home language would be inappropriate. Another example would be the translation of an English Language Arts (ELA) assessment into a students’ first language, given that language arts are unique to each language (e.g., poetry rhymes cannot be replicated exactly) and the structure of language differs across languages (Solano-Flores and Hakuta, 2017). Thus, translating an ELA assessment would distort the construct being measured.

When it has been determined that a language arts assessment should be made available to students in another language, some assessment programs have created a new assessment in the target language. For example, the Colorado Department of Education has

created a Spanish Language Arts (SLA) assessment as an accommodated form of its ELA assessment. This tool is aligned to its English counterpart, but it is not a translation of the ELA assessment. Rather, it is a high-quality assessment that was created to measure the language skills of those students who are designated as English Learners and are also native Spanish speakers. The benefit of this type of tool is that students' knowledge and skills can be measured in their home language, as the construct being measured is now independent of English language ability.

ITC guidelines

Given the frequent need for test translation and adaptation, and the common mishandlings of this process, the International Test Commission (ITC) has created guidelines for translating and adapting tests (ITC, 2017). The first edition of these guidelines was released in 2005 and the second edition, revised to reflect changes in both technology and practices, was released in 2017. The ITC identified six unique categories for translation and adaptation, into which 18 guidelines are categorized. These six categories include (1) Pre-condition, which refers to the conditions that take place prior to translation and adaptation; (2) Test development, or the point at which adaptation takes place; (3) Confirmation, which are the practices needed in order to verify construct equivalence, validity, and reliability after adaptation has taken place; (4) Administration; (5) Score scales and interpretation; and (6) Documentation.

In the pre-condition category, proprietary issues as well as general context are taken into consideration. The guidelines for the pre-condition category are:

- (1) Obtain the necessary permission from the holder of the intellectual property rights relating to the test before carrying out any adaptation.
- (2) Evaluate that the amount of overlap in the definition and content of the construct measured by the test and the item content in the populations of interest is sufficient for the intended use (or uses) of the scores.
- (3) Minimize the influence of any cultural and linguistic differences that are irrelevant to the intended uses of the test in the populations of interest.

The test development category takes into consideration the issues that may arise during the process of translating and adapting the test:

- (4) Ensure that the translation and adaptation processes consider linguistic, psychological, and cultural differences in the intended populations through the choice of experts with relevant expertise.
- (5) Use appropriate translation designs and procedures to maximize the suitability of the test adaptation in the intended populations.
- (6) Provide evidence that the test instructions and item content have similar meaning for all intended populations.
- (7) Provide evidence that the item formats, rating scales, scoring categories, test conventions, modes of administration, and other procedures are suitable for all intended populations.
- (8) Collect pilot data on the adapted test to enable item analysis, reliability assessment and small-scale validity studies so that any necessary revisions to the adapted test can be made.

Confirmation guidelines provide direction on best practices for ensuring that the translation and adaptation are suitable for the target population, after the process has taken place. These guidelines include:

- (9) Select sample with characteristics that are relevant for the intended use of the test and of sufficient size and relevance for the empirical analyses.
- (10) Provide relevant statistical evidence about the construct equivalence, method equivalence, and item equivalence for all intended populations.
- (11) Provide evidence supporting the norms, reliability and validity of the adapted version of the test in the intended populations
- (12) Use an appropriate equating design and data analysis procedures when linking score scales from different language versions of a test.

Administration guidelines are given to ensure that translation and adaptation are appropriate when the assessment is administered to examinees. These two guidelines are:

- (13) Prepare administration materials and instructions to minimize any culture- and language-related problems that are caused by administration procedures and response modes that can affect the validity of the inferences drawn from the scores.
- (14) Specify testing conditions that should be followed closely in all populations of interest.

The score scales and interpretation guidelines are provided to support inferences about examinees' scores:

- (15) Interpret any group score differences with reference to all relevant available information.
- (16) Only compare scores across populations when the level of invariance has been established on the scale on which scores are reported.

Finally, documentation guidelines are provided. The authors of the ITC guidelines note that documentation is often neglected in test adaptation. These guidelines are particularly important because they support the test user, and may also be used as a blueprint in future test adaptations. Guidelines pertaining to documentation include:

- (17) Provide technical documentation of any changes, including an account of the evidence obtained to support equivalence, when a test is adapted for use in another population.
- (18) Provide documentation for test users that will support good practice in the use of an adapted test with people in the context of the new population.

Statistical measures of comparability

The ITC establishes the need for statistical evidence to support various types of translation equivalence (ITC, 2017). Statistical evidence can be extremely useful to researchers who aim to establish construct and score equivalence among translated versions of a test.

At the item level, a common analysis employed to understand score equivalence is called differential item functioning (DIF). This statistical method compares groups of test takers and determines whether one set of test takers (e.g. English Language Learners or native English speakers) is more likely to correctly answer an item, despite comparable ability levels of the two groups. The key element here is that one group has a lower probability of answering an item correctly, despite the two groups being matched on their ability levels using an indicator such as the total test score.

While DIF can be used to understand group differences on any assessment, it is an especially useful tool in translation and adaptation because it can identify items that may have non-equivalence. If an item is identified as exhibiting DIF, an investigation of the item is warranted. For instance, an expert review of the DIF items may need to take place. The expert participants of a review should be familiar with the original and target languages, or both cultural contexts if applicable. These participants would examine both original and adapted DIF items to determine if the meaning of any words or phrases changed in the adaptation, or the equivalence of the terms used in both versions. Another type of investigation that can take place with DIF items is analysis of examinee cognitive processes through think aloud protocols (TAPs), in which students verbalize their thought processes either concurrently while responding to a test item or retrospectively after they have responded to an item. Examining cognitive processes can reveal important differences in translation or cultural contexts that impact students' abilities to demonstrate their knowledge (Ercikan et al., 2012). For instance, it may be determined that the translated item included words that differ from the target language in terms of difficulty, frequency of use, or overall meaning.

At the test level, construct comparability between forms may be examined through analyses such as exploratory or confirmatory factor analysis. This statistical method uses inter-item correlations to determine which items group together into "factors" or dimensions. Using this approach, it is possible to compare the dimensions of the original and translated test versions. Similarities across the two versions imply construct comparability, while dissimilarities may suggest a difference in the construct being measured by the two test versions. Exploratory factor analysis is performed separately for both groups and does not require that the factor structure be predicted prior to conducting the analysis, whereas confirmatory factor analysis does require a predicted factor structure, but may be used for multiple groups simultaneously.

Conclusions and future implications

Educational practices continue to evolve alongside technological innovations. Such advances have made translation more accessible than ever. For example, the proliferation of computer-based testing offers considerable opportunities to provide various forms of translated and adapted versions, including pop-up glossaries (Solano-Flores et al., 2019). Additionally, software programs that allow for immediate translation are numerous and often free to users. However, such accessibility also increases the danger of inaccurate or imprecise adaptations. Considering the many high stakes uses of test scores (e.g. graduation, special education placement), test developers should prioritize the implementation of best practices for translation and adaptation to promote the valid use of translated test scores.

This article opened with a description of the ever-increasing changes in migration that have been observed in the last 50 years, and its implications for more diverse schools and classrooms, as well as the general globalization of education. There is no indication that this trend will slow or pause; if anything, diversity in language and culture will only continue to intensify in schools. While exposure to diverse students based on language and culture offers many positive benefits to students, it is the responsibility of educators and researchers to ensure that assessment results are interpreted in a fair and valid manner.

The issues described in this article demonstrate how fairness comes into play when assessments are translated or adapted for diverse populations. Test developers and users may use this article to better understand the complex issues surrounding translation and adaptation in educational assessments. By being aware of these concerns, and following the practices described in the article, test developers and users can minimize construct irrelevant variance in assessment while promoting test fairness for ELL students.

References

- Abedi, J., 2002. Standardized achievement tests and English language learners: psychometric issues. *Educ. Assess.* 8 (3), 231–257.
- Abedi, J., 2009. English language learners with disabilities: classification, assessment, and accommodation issues. *J. Appl. Test. Technol.* 10 (2), 1–31.
- Abedi, J., 2014. English language learners with disabilities: classification, assessment, and accommodation issues. *Appl. Meas. Educ.* 27 (4), 261–272.
- Abella, R., Urrutia, J., Shneyderman, A., 2005. An examination of the validity of English-language achievement test scores in an English language learner population. *Biling. Res. J.* 29 (1), 127–144. <https://doi.org/10.1080/15235882.2005.10162827>.
- Abubakar, A., Dimitrova, R., Adams, B., Jordanov, V., Stefenel, D., 2013. Procedures for translating and evaluating equivalence of questionnaires for use in cross-cultural studies. *Bull. Transilv. Univ. Braşov* 6, 55.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. *Standards for Educational and Psychological Testing*. AERA, Washington, DC.
- Callahan, R.M., Shiffrer, D., 2016. Equitable Access for secondary English learner students: course taking as evidence of EL program effectiveness. *Educ. Adm. Q.* 52 (3), 463–496. <https://doi.org/10.1177/0013161X16648190>.
- Carnoy, M., 2015. *International Test Score Comparisons and Educational Policy: A Review of the Critiques*. NEPC.
- Donovan, M.S., Cross, C.T. (Eds.), 2002. *Minority Students in Special and Gifted Education*. National Academies Press, Washington, DC.
- Ercikan, K., McCreith, T., 2002. Effects of adaptations on comparability of test items and test scores. In: *Secondary Analysis of the TIMSS Data*. Springer, Dordrecht, pp. 391–405.
- Ercikan, K., Arim, R., Law, D., Domene, J., Gagnon, F., Lacroix, S., 2010. Application of think aloud protocols for examining and confirming sources of differential item functioning identified by expert reviews. *Educ. Meas.* 29 (2), 24–35.
- Ercikan, K., Simon, M., Oliveri, M.E., 2012. Score comparability of multiple language versions of assessments within jurisdictions. In: *Improving Large-Scale Assessment in Education: Theory, Issues and Practice*, pp. 110–124.
- Ford, D.Y., Grantham, T.C., Whiting, G.W., 2008. Culturally and linguistically diverse students in gifted education: recruitment and retention issues. *Except. Child.* 74 (3), 289–306.
- Fry, R., Parker, K., 2018. Early Benchmarks Show 'Post-Millennials' on Track to Be Most Diverse, Best-Educated Generation Yet. Pew Research Center. <https://www.pewsocialtrends.org/2018/11/15/early-benchmarks-show-post-millennials-on-track-to-be-most-diverse-best-educated-generation-yet/>.
- Goodrich, S., Ercikan, K., 2019. Measurement comparability of reading in English and French Canadian population: special case of the 2011 progress in international reading literacy Study. In: *Frontiers in Education*, vol. 4. Frontiers, p. 120.
- Grantmakers for Education, 2013. *Educating English Language Learners: Grantmaking Strategies for Closing America's Other Achievement Gap*. Retrieved from. https://edfunders.org/sites/default/files/Educating%20English%20Language%20Learners_April%202013.pdf.
- Haertel, E.H., 1999. Validity arguments for high-stakes testing: in search of the evidence. *Educ. Meas.* 18 (4), 5–9.
- International Organization for Migration, 2019. *World Migration Report 2020*. https://www.un.org/sites/un2.un.org/files/wmr_2020.pdf.
- International Test Commission, 2017. *The ITC Guidelines for Translating and Adapting Tests*, second ed. www.InTestCom.org.
- Kane, M.T., 1992. An argument-based approach to validity. *Psychol. Bull.* 112 (3), 527–535.
- Kanno, Y., Kangas, S.E.N., 2014. "I'm not going to Be, like, for the AP": English language learners' limited access to advanced college-preparatory courses in high school. *Am. Educ. Res. J.* 51 (5), 848–878. <https://doi.org/10.3102/0002831214544716>.
- Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educ. Res.* 23 (2), 13–23.
- National Education Association, 2015. *All in! How Educators Can Advocate for English Language Learners*. NEA Professional Library, Washington, DC.
- Solano-Flores, G., Hakuta, K., 2017. *Assessing Students in Their Home Language*. Retrieved from Stanford University. Understanding Language website. <https://stanford.app.box.com/s/uvwlgjbmeuokts6c2wnibucms4up9c2>.
- Solano-Flores, G., Backhoff, E., Contreras-Niño, L.A., 2009. Theory of test translation error. *Int. J. Test.* 9, 78–91.
- Solano-Flores, G., Chia, M., Kachchaf, K., 2019. Design and use of pop-up illustration glossaries as accessibility resources for second language learners in computer-administered tests in a large-scale assessment system. *Int. Multidiscip. Res. J.* 13 (4), 277–293.
- Sullivan, A., 2011. Disproportionality in special education identification and placement of English language learners. *Except. Child.* 77 (3), 317–334.
- Valdés, G., 2004. Between support and marginalization: the development of academic language in linguistic minority children. *Int. J. Biling. Educ. Biling.* 7 (2–3), 102–123. <https://doi.org/10.1080/13670050408667804>.

Score equating: an aspirational form of score linking

Neil J. Dorans^a and Tim Moses^b, ^a Consultant, Lambertville, NJ, United States; and ^b College Board, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Score equating: a special case of scale alignment	237
Data collection designs	238
The single group (SG) and counterbalanced (CB) designs	238
The equivalent groups design (EG)	238
Anchor test designs	239
Procedures for equating scores	239
Data pre-processing practices	239
Sample selection	240
Checking that anchor items act like common items	240
The need to continuize the discrete distributions of scores	240
Presmoothing score distributions	240
Observed score equating methods	241
Observed score procedures for equating scores in a common population	241
Procedures for equating scores on complete tests when using common items	242
Best practices for common item equating	243
Best practices for the anchor test design	243
Lack of invariance	244
Score linkings that can be misconstrued as score equatings	244
Linkages across tests from different testing programs	244
Recent examples of concordances within testing programs	245
Best practices for reporting score linkings that are not equatings	246
Summary	246
References	246

People like to make comparisons. Sometimes, this urge to compare is sensible. For example, we can order objects by length or by weight. We may order students by perceived ability to read according to how reading is measured by a certain test. Scores from different assessments are sometimes compared sensibly; sometimes they are not. Rarely does a person consider using weight, one measure of size, as a surrogate for height, another measure of size. Yet that same person may be misled to assume that scores from any one educational assessment can be used as substitutes for scores from another assessment simply because they are said to measure the same thing.

The process that converts scores from different assessments to a common metric is called score linking. Simply because scores are on a common metric, however, does not mean they can be compared as if they measure the same attribute or construct. [Holland and Dorans \(2006\)](#) provided a framework for score linking that considered key aspects of the measurement process, including what is being measured by the tests to be linked, the trustworthiness of the scores produced by the tests, and the reason for linking the scores. Score linking can be divided into two fundamental categories: prediction and scale alignment, of which there are many subcategories, including score equating. The goal of prediction is to minimize errors for predicted scores on the dependent/criterion variable based on information from other predictor variables.

The goal of scale alignment is to transform the scores from two different tests onto a common scale. Scaling procedures are about 100 years old ([Hull, 1922](#); [Kelley, 1914, 1923](#)), yet scale alignment and score equating are often confused because the statistical procedures used for scale alignment can also be used to equate tests. As a consequence of this procedural identity, some view the distinction between equating and scale alignment (as well as distinctions among different types of scale alignment) as unnecessary because the numerical procedures are the same ([Livingston, 2014b](#); [van der Linden, 2013](#)). [Dorans \(2013\)](#) maintains that the distinctions made in the [Holland and Dorans \(2006\)](#) framework, and other frameworks, refer to the meanings that can be legitimately attached to the outcomes of the numerical operations, not to the procedures themselves. For example, it is easy to align distributions of height and weight in a specified population, and assign them to a common scale, e.g., largeness, such that a largeness unit of height is as far above the average height as a largeness unit of weight is above the average weight. That numerical alignment, however, does not mean that a largeness unit of height is *equal in meaning* to a largeness unit of weight.

It is important to distinguish between the subcategories of scale alignment because they are often seen as similar or identical when, in fact, they are not. The distinctions between different types of linkages are related to the similarities and differences between the tests as well as to the ways in which the linked scores are intended to be used and interpreted. The terms used for the various methods of score linking should have restricted meanings so that they may be used with precision. Understanding the distinctions

among these categories can promote appropriate score interpretations and discourage misinterpretations, misuse, and violations of professional practice.

The previous version of this chapter (von Davier, 2010) focused on different types of linking including item parameter linking in the context of item response theory (IRT). The emphasis of this chapter is on score equating where the goal is to produce interchangeable scores from different editions of the same test that are assembled from a common blueprint and administered under a standard set of measurement conditions. Equating adjusts for minor differences in the difficulties of alternate test forms so that interchangeable scores can be produced across the forms. Large representative samples of examinees, sound data collection practices, and appropriate methods are needed to produce equated scores on these test editions. Ideally, the tests should produce highly reliable scores as well.

The section **Score equating: a special case of scale alignment** addresses what distinguishes score equating from other forms of scale alignment. All methods for linking test scores include scores on two tests obtained by one or more samples of test takers, obtained via data collection methods as described in **Data collection designs** section, and one or more procedures for computing the transformation function, as described in **Procedures for equating scores** section. All methods of observed score equating specify a population of test takers to which the test linking is to be applied. Our focus is on observed score equating because true scores are unobserved and consequently of primarily theoretical interest only. The reality is that the observed score is all we can observe. Even sophisticated model-based scoring procedures that have theoretical true scores as their target only produce manifest observed scores, not latent unobservable scores, a point that is often not stated explicitly.

After describing score equating and what is needed to produce equated scores, in the **Score linkings that can be misconstrued as score equatings** section we address other forms of score linking that fail to achieve interchangeable scores even though they use the same procedures as those used for score equating. Unfortunately, equating-like interpretations are erroneously attached to these ersatz equating functions. In the last section we close with a discussion of best practices for score equating.

Score equating: a special case of scale alignment

Score equating is the strongest form of linking between the scores on two tests. Score equating may be viewed as a form of scale alignment in which very strong requirements are placed on the tests being linked. The goal of equating is to produce a linkage between scores on two test forms such that the scores from each test form can be used as if they had come from the same test. Among other things, the two tests built from the same blueprint or test specifications measure the same construct at the same or nearly the same level of difficulty and are administered under similar measurement conditions to produce scores with a similar degree of reliability.

Score equating involves scale aligning computations that link raw scores on a new form or edition of test X (the x 's) to those of an old form of test Y (the y 's). Then these equated raw scores from X , $e(x)$, are converted to a reporting scale. In practice, there is a scaling function that maps the raw scores of Y to the reporting scale, call it $S = s(y)$. The composition function, $S = s[e(x)]$, is the scaling function for X .

In practice, equating is necessary because of the inability to construct multiple forms of a test that are strictly parallel. Equating is used to fine-tune the test construction process. While impossible to achieve in practice, ideal equating can be approximated well with the type of sound equating practices promoted by Dorans et al. (2011).

There are four requirements that are widely viewed as necessary for a linking to be an equating. Those requirements are:

- a. **The Equal Construct Requirement:** The two tests should both be measures of the same construct.
- b. **The Equity Requirement:** It should be a matter of indifference to a test taker to be tested by either of the tests that have been equated.
- c. **Common Measurement Condition Requirement:** The test should be built to the same set of specifications and taken under the same measurement conditions.
- d. **The Symmetry Requirement:** The equating transformation for mapping the scores of Y to those of X should be the *inverse* of the equating transformation for mapping the scores of X to those of Y .

If requirements a, b and c hold, then the following consequences should hold:

- I. **Subpopulation Invariance:** The function used to link the scores of X and Y should be the same regardless of the choice of (sub) population from which it is derived.
- II. **Equal and High Reliability:** The test will produce scores that have equal and high reliability in any subpopulation taking the tests

Statements related to these appear in a variety of earlier and subsequent sources including Angoff (1971), Lord (1980), Petersen et al. (1989), Kolen and Brennan (2014), Holland and Dorans (2006), and Moses (2022). In practice, requirements (a) and (b) mean that the tests need to be built to the same content and statistical specifications. Requirement (c) acknowledges the important role that context plays in testing practice, as noted by Kolen and Brennan (2014). Consequences (I) and (II) may be used to explain why Lord (1980) noted that both (a) and (b) are necessary, and why we believe that (c) is also needed. If the two tests

measure different constructs, test takers will prefer the one on which they believe they will score higher, and some may prefer a less reliable measure than others. Dorans and Holland (2000) use consequence (I) to develop quantitative measures of equatability that indicate the degree to which equating functions depend on the subpopulations used to estimate them. Liu and Dorans (2013) explain how these measures and related measures of equatability can be used to assess the degree to which test scores are equatable. Requirement (b) is hard to evaluate empirically, and its use is primarily theoretical (Lord, 1980; Hanson, 1991). If consequence (II) does not hold then requirement (b) cannot be met. Brennan (2010) demonstrated analytically that equity can only be achieved with tests that produce highly reliable scores, confirming a hypothesis posited by Dorans and Walker (2007) that high reliability, rather than simply equal reliability, characterizes equated scores.

Data collection designs

To obtain the clearest estimates of test form difficulty differences, all score equating methods must control for differential ability of the test taker groups. Data collection strategies used for score equating strive to obtain equivalent groups, either directly or indirectly. Often, two different, non-strictly parallel tests are taken by different groups of test takers of unequal ability. Assuming that the samples are large enough to ignore sampling error, differences in the distributions of the resulting scores can be caused by either differences in the relative difficulty of the two tests or differences in the relative ability of the two groups taking these tests, or both. Difference in ability between the groups is a confounding factor that needs to be eliminated before the equating process can adjust for differences in test difficulty.

There are two different approaches for addressing the separation of test difficulty and group ability differences. The first approach is to use a common population of test takers. The other approach is to use an anchor measure of the construct being assessed by both X and Y.

When the same population of test takers takes both tests, we achieve direct control over differential test taker ability. In practice, it is more common to use two equivalent samples from a common population instead of identical test takers.

The second approach assumes that performance on a set of common items or an anchor test can quantify the ability differences between two distinct samples of test takers. The use of an anchor test allows for flexible data collection designs, but it comes with a cost. These designs employ procedures that make various assumptions that are not needed when the test takers are from the same population. Different methods make different assumptions and the various statistical adjustments for these group differences in ability can produce different results.

All procedures that follow, including those for samples from a common population, make a set of common assumptions needed to ensure that the linking can produce interchangeable scores. All samples are assumed to be random samples from their respective populations. In practice this may be only an approximation. In addition, all test takers are assumed to have had the opportunity to answer all the questions on one or both tests.

Because it is desirable for the statistical uncertainty associated with test equating to be much smaller than the other sources of variation in test results, it is important that the results of test equating be based on samples that are large enough to insure this. This fact should always be kept in mind when selecting a data collection design.

The single group (SG) and counterbalanced (CB) designs

The single group (SG) design is the simplest data collection design. In the single group design, all test takers in a single sample from population P take both tests. This design can provide accurate equating results with relatively small sample sizes. The SG design controls for differential group proficiency by having the same test takers take both tests. In using this design, however, it is necessary to assume that a test taker's score on the second test is unaffected by the fact that she or he has previously taken the first form.

To study whether order affects the results obtained with the SG design, the sample is sometimes randomly divided in half, and in each half sample the two tests are taken in different orders—X first and then Y or Y first and then X. This is the counterbalanced (CB) data collection design. The CB design contains both the SG and EG (described in [The equivalent groups design \(EG\)](#) section) designs within it.

The equivalent groups design (EG)

Typically, there is not enough testing time for every test taker to take more than one test. In the equivalent groups (EG) design, two equivalent samples are taken from a common population P, one is tested with X and the other with Y. The EG design avoids the issue of possible order effects that can arise in the SG design where each test taker takes both tests.

Sometimes test booklets are assigned randomly to groups, which is why this design is sometimes called the random groups design (Kolen and Brennan, 2014). While it does not require that the two tests have any items in common, this design can be used with items in common, which makes it an anchor test design (see next subsection). When sample sizes are large and tests forms are administered under secure conditions, the EG is the preferred data collection design.

Anchor test designs

In anchor test designs there are two populations, P and Q, with a sample of test takers from P taking test X, and a sample from Q taking test Y. In addition, both samples take an anchor test, A. von Davier et al. (2004) and Holland and Dorans (2006) call this the “Non-Equivalent groups with Anchor Test” (or NEAT) design. Kolen and Brennan (2014) and others refer to this as the “common-item nonequivalent groups design”.

The role of the anchor test is to quantify the differences in ability between samples from P and Q that affect their performance on the two tests to be equated, X and Y. The anchor A is usually, but not always, a shorter and less reliable test than the tests to be equated, but still measures the same construct as X and Y.

The relationship of A to X and Y can be of two different types. If A is a subset of items contained in both X and Y, then it is called an internal anchor test and scores on A are used in computing scores on the total tests. If A is a separately-timed test that each test taker takes in addition to one or the other of the tests being equated, then it is called an external anchor test.

The anchor test design contains two single-group designs within it. This structure underlies both anchor scaling (Holland and Dorans, 2006), and the chain equating methods to be discussed later. The anchor test design improves upon the flexibility of the EG design by allowing the two samples taking X and Y to be non-equivalent. Unlike the SG design, it does not require test takers to take both X and Y.

While the use of anchor tests may appear to be a minor variation of the previous data collection designs, the use of common items rather than common test takers involves new assumptions that are not necessary in the use of SG, EG, and CB designs. Some type of assumption is required to make up for the fact that X is never observed for test takers in Q and Y is never observed for test takers in P. For this reason, there are several distinct methods of scaling and equating tests using the NEAT design. Each of these methods corresponds to making different untestable assumptions about the missing data.

The difference between the anchor test design and the SG, EG, and CB designs can be thought of as the difference between observational studies versus experimental designs (Rosenbaum, 1995). The SG design is a repeated measures design with a single group and two treatments, the EG design is a randomized comparison with two treatment groups, while the CB design is a repeated measures design with a single group and counterbalanced order of treatments. In contrast, the common item design is an observational study where there are two non-randomized study groups that are possibly subject to varying amounts of self-selection. Liou et al. (2001) point out that the properly executed anchor test design produces missing data that are *missing at random* in the technical sense discussed in Little and Rubin (2002).

Unrepresentative or unmotivated samples undermine score equating; hence, special care should be taken to ensure that only members of the population of interest are included in the samples. If possible, the sample should be representative of the population as well. In addition, equating cannot be done effectively in small samples. The smaller the sample size, the more restricted is the class of stable equating methods. The single group (SG) design requires the smallest sample sizes, while the equivalent groups (EG) design requires the largest sample sizes to achieve the same level of accuracy, as measured by the standard error of equating (see Lord, 1950; Holland and Dorans, 2006). The anchor test designs require sample sizes somewhere in between those of the SG and EG designs; these requirements depend on how strongly correlated the anchor test is with the two tests to be equated and how similar the two populations are.

More extended discussions of data collection designs are given in Angoff (1971), Petersen et al. (1989), von Davier et al. (2004), Kolen and Brennan (2014), and Holland and Dorans (2006).

Procedures for equating scores

In this section, we will limit our discussion to observed score equating methods that use the data collection designs just described. Much of this section is an abridged version of the treatment provided by Holland and Dorans (2006). Consult Holland and Dorans (2006) for more complete treatments of observed score and true score procedures.

First, pre-processing steps are described. Next, observed score procedures are described, first for the common population case, and then for the case where anchor tests are needed to adjust for population differences before the test scores can be properly equated.

Some notation is used in this section. The *cumulative distribution function*, CDF, of the scores of test takers in the target population, T, on test X is denoted by $F_T(x)$, and it is defined as the proportion of test takers in T who score at or below x on test X. More formally, $F_T(x) = P\{X \leq x \mid T\}$, where $P\{\cdot \mid T\}$ denotes the population proportion or probability in T. Similarly, $G_T(y) = P\{Y \leq y \mid T\}$, is the CDF of Y over T. Cumulative distribution functions increase from 0 up to 1 as x (or y) moves from left to right along the horizontal axis in a two way plot of test score by proportion of test takers. In this notation, x and y may be any real values, not necessarily just the possible scores on the two tests. For distributions of observed scores such as number right or rounded formula scores, the CDFs are step functions that have points of increase only at each possible score (Kolen and Brennan, 2014).

Data pre-processing practices

Having collected the data, several steps should be taken to improve the quality of the data prior to score equating. These data processing steps deal with sample selection, item screening, and continuizing and smoothing score distributions.

Sample selection

Tests are designed with a target population T in mind. For example, admissions tests are used to gather standardized information about candidates who plan to enter a college or university. Test takers who are not part of this group can be removed from the equating analyses.

Statistical outlier analysis can be used to identify those test takers whose anchor test performance is substantially different from their performance on the operational test, i.e., the scores are so different that both scores cannot be plausible indicators of the test taker's ability. Removing these test takers from the equating sample prevents their unlikely performance from having an undue effect on the resulting equating function.

Checking that anchor items act like common items

The statistical properties of the common items should be evaluated to make sure they have not differentially changed from one test administration to the other. Internal common items are susceptible to context effects because they may be embedded within different sets of items in the two tests. Changes in widely held knowledge may also lead to changes in performance on anchor test items.

Simple plots of item difficulty values and other statistics may also be used to detect changes in items. It is important to check whether the difficulties of the common items are ranked in the same way across the old and new form samples. Item difficulty estimates from old and new form samples should correlate in the high 90s. That said, even correlations as high as the low 90s can still result in problematic score equatings. Differential item functioning methods may be used to compare the performance of the common items with the two test administrations treated as the reference and focal groups, and the total score on the common items as the matching criterion.

The need to continuize the discrete distributions of scores

Test scores are typically integers, such as number-right scores or rounded formula-scores. The equipercentile function defined below can depend on how the cumulative frequency distribution of X on T , $F_T(x)$, and Y on T , $G_T(y)$, are made continuous or "continuized". There are two methods of "continuization" of $F_T(x)$ and $G_T(y)$ that are in current use. [Holland and Dorans \(2006\)](#) treat two of these methods, the linear interpolation and kernel smoothing methods, in detail.

There are two primary differences between these two approaches to continuization. First, the use of linear interpolation results in an equipercentile function that is piecewise linear and continuous, resulting in functions that have kinks that some prefer to post-smooth ([Kolen and Brennan, 2014](#)). In contrast, kernel smoothing results in equipercentile functions that are very smooth (i.e., differentiable everywhere) and that do not need further post-smoothing. Second, the equipercentile functions obtained by linear interpolation always map the highest score on X into the highest score on Y and the same for the lowest scores (unlike kernel smoothing and the linear equating function). For more discussion on how the highest and lowest scores of two tests should be mapped, see [Petersen et al. \(1989\)](#), [Kolen and Brennan \(2014\)](#) and [von Davier et al. \(2004\)](#).

Presmoothing score distributions

Irregularities in the score distributions can produce irregularities in the equipercentile equating function that do not generalize to other groups of test-takers. Consequently, it is generally considered advisable to *smooth* the raw-score frequencies, the CDFs, or the equipercentile equating function itself ([Holland and Thayer, 1987, 2000](#); [Kolen and Jarjoura, 1987](#); [Kolen and Brennan, 2014](#); [von Davier et al., 2004](#); [Moses, 2011](#)). The purpose of this step is to eliminate some of the sampling variability present in the raw-score frequencies, in order to produce smoother CDFs for computation of the equipercentile function.

When presmoothing data, it is important to achieve a balance between a good representation of the original data and smoothness. Smoothness reduces sampling variability while a good representation of the data reduces the possibility of bias. The more parameters employed in the smoothing, the better the model will represent the original data, but the less smooth the fitted model becomes.

One way to presmooth data, proposed by [Lord \(1980\)](#), involves the use of an IRT model to pre-smooth a number-right score distribution prior to equipercentile equating. Let X and Y denote number-right scores for two tests with dichotomous test items. Using IRT, the conditional distributions of X and Y given the common latent ability, q , may be estimated. This computation exploits the assumption of local independence to compute $P\{X = x | q\}$ and $P\{Y = y | q\}$ as the probabilities derived from independent but not identically distributed conditional probabilities of 0/1 variables. See [Lord and Wingersky \(1984\)](#) and [Kolen and Brennan \(2014\)](#) for a detailed example and additional references. The estimated density function for the ability distribution over T , $f_T(q)$, may be approximated in various ways, and these results can be combined via numerical integration. The resulting estimated probabilities, $P\{X = x | T\}$ and $P\{Y = y | T\}$, are IRT model-based estimates of the original raw-score sample proportions, i.e., the proportion in the sample getting a score of x on X , or the proportion getting a score of y on Y . Thus, the estimates, $P\{X = x | T\}$ and $P\{Y = y | T\}$, may be regarded as a form of presmoothing of the two, unsmoothed sample score distributions. This is a use of an IRT model to presmooth the data, prior to continuization during equipercentile equating. Once the presmoothed estimated score probabilities are in hand, they may be continuized by either equipercentile-based linear interpolation or Gaussian kernel smoothing and used to produce a version of the equipercentile function, $Equi_{IRT}(x)$.

IRT observed score equating is seldom used in practice (see [Kolen and Brennan, 2014](#)). Compared to other presmoothing methods, it requires strong assumptions at the item level and is computationally more complex. Furthermore, there is no guarantee that these presmoothed estimates of the score distributions will reproduce the types of score frequencies found in real data. For this

reason, it is unlikely to be used except in research studies or very special situations with incomplete test data. For situations with data from complete tests, it is neither as efficient nor as flexible as the direct presmoothing of score frequencies using loglinear models, as advocated by [Holland and Thayer \(2000\)](#), [von Davier et al. \(2004\)](#) and others.

Smoothing score distributions works in moderately-sized samples, but does not help much with very small samples, especially when it is not clear how representative the sample is of the intended population. In these situations, one option may be to make strong assumptions about the equating function. The best practices solution to the small sample size problem may be to report raw scores and state that they cannot be compared across test forms. If the sample size suggested by consideration of standard errors is not achieved, raw scores could be reported with the caveat that they are not comparable to other scores, but that they could be made comparable when adequate data become available.

Observed score equating methods

Many procedures have been developed over the years for equating tests. [Holland and Dorans \(2006\)](#) considered three factors when attempting to develop a taxonomy of equating methods:

- (1) common population versus common item data collection designs,
- (2) observed-score versus true-score procedures, and
- (3) linear versus non-linear methods.

Because equating is an empirical procedure, it requires a data collection design and a procedure for transforming scores on one test form to scores on another. Linear methods produce a linear function for mapping the scores from X to Y , while non-linear methods allow the transformation to be curved. Observed-score procedures directly transform (or equate) the observed scores on X to those on Y . True-score methods are designed to transform the *true scores* on X to the true scores of Y . True score methods employ a statistical model with a test taker's true score defined as their expected observed test score based on the chosen statistical model. The psychometric models used to date are those of classical test theory and item response theory. [Holland and Hoskens \(2003\)](#) show how these two psychometric models may be viewed as aspects of the same model.

Observed score procedures for equating scores in a common population

As noted previously, there are three data collection designs that make use of a common population of test takers: the Single Group (SG), the Equivalent Groups (EG) and the Counterbalanced (CB) designs. They all involve a single population, P , which is also the target population, T .

We will use a definition of observed-score equating that applies to either linear or non-linear procedures depending on whether additional assumptions are satisfied. This allows us to consider both linear and non-linear observed-score equating methods from a single point of view.

The equipercentile equating function

The equipercentile definition of *comparable* scores is that x (an X -score) and y (a Y -score) are *comparable* in T if $F_T(x) = G_T(y)$. This means that x and y have the same percentile in the target population, T . When the two CDFs are continuous and strictly increasing, the equation $F_T(x) = G_T(y)$ can always be satisfied and can be solved for y in terms of x . Solving for y leads to the *equipercentile function*, $Equi_{YT}(x)$, that links x to y on T , defined by:

$$y = Equi_{YT}(x) = G_T^{-1}(F_T(x)) \quad (1)$$

In [Eq. \(1\)](#), $y = G_T^{-1}(p)$ denotes the inverse function of $p = G_T(y)$. Note that with discrete data, this relationship does not hold because for most x scores there is no y score for which the two cumulative distributions, one for x and one for y are exactly equal. Hence, with most applications, continuization steps are taken to make the data appear continuous (see [The need to continuize the discrete distributions of scores](#) section); different steps can yield different answers.

If [Eq. \(1\)](#) is satisfied, then $Equi_{YT}(x)$ will transform the distribution of X -scores on T so that it is the same as the distribution of Y -scores on T . Note that the target population T is explicit in the definition of $Equi_{YT}(x)$ ([Dorans and Holland, 2000](#); [von Davier et al., 2004](#); [Holland and Dorans, 2006](#)). In general, there is nothing to prevent $Equi_{YT}(x)$ from varying with the choice of T . For equating, we expect the influence of T to be small or negligible and we call the scores *equivalent*. In other kinds of linking, T can have a substantial effect and we call the scales *comparable in T*.

The linear equating function

It is sometimes appropriate to assume that the two CDFs, $F_T(x)$ and $G_T(y)$, have the same shape and only differ in their means and standard deviations. To formalize the idea of a common shape, suppose that $F_T(x)$ and $G_T(y)$ both have the form,

$$F_T(x) = K[(x - m_{XT}) / s_{XT}] \text{ and } G_T(y) = K[(y - m_{YT}) / s_{YT}] \quad (2)$$

where K is a CDF with mean 0 and standard deviation 1, m_{XT} and m_{YT} are the means of X and Y , and s_{XT} and s_{YT} are the standard deviations of X and Y .

When Eq. (2) holds, $F_T(x)$ and $G_T(y)$ both have the shape determined by K . In this case, it can be shown that the equipercentile function is the linear function, $Lin_{YT}(x)$, defined as

$$Lin_{YT}(x) = m_{YT} + (s_{YT} / s_{XT})(x - m_{XT}) \quad (3)$$

The linear function may also be derived as the transformation that gives the X-scores the same mean and standard deviation as the Y-scores on T. Both the linear and equipercentile functions satisfy the symmetry requirement (d) but reflect different definitions of comparable score scales (Kelley, 1923). This means that $Lin_{XT}(y) = Lin_{YT}^{-1}(x)$, and $Equi_{XT}(y) = Equi_{YT}^{-1}(x)$, i.e., equating Y to X is the inverse of the function for equating X to Y. In general, the function, $Equi_{YT}(x)$, curves around the function, $Lin_{YT}(x)$, such that the differences in these functions attributable to differences in the shapes of the CDFs fluctuate around 0 as the linear function approximates the first term of the expansion of the curvilinear function (von Davier et al., 2004, pp. 11–13).

The linear function requires estimates of the means and standard deviations of X- and Y-scores over the target population, T. It is easy to obtain these estimates for the SG and EG designs described in the Data collection designs section (see Angoff, 1971, or Kolen and Brennan, 2014). It is less straightforward to obtain estimates for the CB design, as noted by Holland and Dorans (2006).

Procedures for equating scores on complete tests when using common items

The anchor test design is widely used for equating scores because its use of common items to control for differential test taker ability gives it greater operational flexibility than the approaches using common test takers. Test takers need only take one test, and the samples need not be from a common population. However, this flexibility comes with a price. First, the target population is less clear-cut for the anchor test design—there are two populations, P and Q, and either one could serve as the target population. In addition, the use of the NEAT design requires additional assumptions to allow for the missing data—X is never observed in Q and Y is never observed in P. We use the term complete test to indicate that everyone in P sees all items on X and that everyone in Q see all items on Y. Our use of the term “missing data” is restricted to data that are missing by design. The assumptions needed to make allowances for the missing data are not easily tested with the observed data, and they are often unstated. As a result of this design, the equating of X and Y must be produced indirectly through the use of A (Holland and Dorans, 2006; Holland, 2007). We will discuss two distinct sets of assumptions that may be used to justify the *observed score* procedures that are commonly used with the anchor test design.

Braun and Holland (1982) proposed that the target population for the anchor test design, or what they called the *synthetic population*, be created by weighting P and Q. They denoted the synthetic population by $T = wP + (1-w)Q$, by which they meant that distributions (or moments) of X or Y over T are obtained by first computing them over P and Q, separately, and then averaging them with w and $(1-w)$ to get the distribution over T. There is considerable evidence that the choice of w has a relatively minor influence on equating results, for example, see von Davier et al. (2004). This insensitivity to w is an example of the population invariance consequence. The definition of the synthetic population forces the user to confront the need to create distributions (or moments) for X on Q and Y in P, where there are no data. In order to do this, assumptions must be made about the missing data.

Equating methods used with the anchor test design can be classified into two major types: *post stratification equating* (PSE) and *chain equating* (CE). Each of these types employs assumptions that assert that an important distributional property that connects scores on X or Y to scores on the anchor test A is the same for any $T = wP + (1-w)Q$, i.e., is population invariant.

Post-stratification equating methods

The PSE types of assumptions all have the form that the conditional distribution of X given A (or of Y given A) is the same for any synthetic population, $T = wP + (1-w)Q$. In this approach, we estimate, for each score on the anchor test, the distribution of scores on the new form and on the old form in T. We then use these estimates for equating purposes as if they had actually been observed in T. The PSE type of equating assumes that the relationship that generalizes from each equating sample to the target population is a conditional relationship. In terms of the missing data in the NEAT design, this means that conditional on the anchor test score, A, the distribution of X in Q (where it is missing) is the same as in P (where it is not missing). In the special case of an EG design with anchor test, $P = Q$ and the PSE assumptions hold exactly. When P and Q are different, the PSE assumptions are not necessarily valid, but there are no data to contradict them.

Chained equating methods

The CE assumptions all have the form that a linking function from X to A (or from Y to A) is the same for any synthetic population, $T = wP + (1-w)Q$. In this approach, we link the scores on the new form to scores on the anchor and then link the scores on the anchor to the scores on the old form. The “chain” formed by these two linkages links the scores on the new form to those on the old form. The CE type of equating approach assumes that the linking relationship that generalizes from each equating sample to the target population is an equating relationship. It is less clear for the CE assumptions than for the PSE assumptions what is implied about the missing data in the NEAT design (Kolen and Brennan, 2014).

In the special case of an EG design with anchor test, $P = Q$ and the CE assumptions hold exactly. In this special situation, the corresponding methods based on either the PSE or the CE assumptions will produce identical results. When P and Q are different, the PSE assumptions and CE assumptions can result in equating functions that are different and there are no data to allow us to contradict or help us choose between either set of assumptions.

Levine observed score equating method

In addition to the PSE and CE types of procedures, classical test theory may be used to derive an additional *linear observed-score* procedure for the NEAT design—the Levine observed-score equating function, $Lev_{YT}(x)$ (Kolen and Brennan, 2014). $Lev_{YT}(x)$ may be derived from two population invariance assumptions that are different from those that we have considered so far and that are based on classical test theory. Holland and Dorans (2006) use the form of classical test theory discussed in Holland and Hoskens (2003) to derive the Levine observed score equating function from two assumptions. To formalize the intuition that X, Y and A all measure the same construct, they assume their true scores are linearly related in a way that holds for all T, i.e., the three measures are congeneric.

Local linking

van der Linden and Wiberg (2010) propose an ability- or anchor-specific procedure for observed score equating in which the equipercentile linking function between test X and Y is defined at each level of an ability measure, such as the scores of anchor test A. This procedure has both theoretical and practical limitations as noted by Dorans (2013), and Holland (2013).

Best practices for common item equating

The following sections address issues that can make score equating challenging, especially when an anchor test design is employed. The quality and similarity of tests to be equated, choice of data collection design, characteristics of the anchor test in relation to the total tests, sample sizes and test taker characteristics, screening items and tests for outliers, and the choice of analyses all involve best practices that contribute to a successful equating.

The ideal design, in theory and in terms of best practice, is a large sample equivalent groups (EG) design with an external anchor test. If the anchor test is administered last, only the anchor test can be affected by possible order effects. A comparison of the distributions of the anchor test in the two (equivalent) samples then allows differential order effects to be identified; if they are substantial, the anchor test can be ignored, leaving a simple EG design, where no order effects are possible. If the anchor test is internal to the two tests, then context or order (e.g., item location) effects may arise and need to be dealt with.

In practice, however, large representative motivated samples that result from a random assignment of test forms to test takers are not always attainable. Reliability is not always as high as desired. Anchor tests may not be very reliable, especially internal anchors with few items. Anchors, especially external anchors, are not always highly related to the tests being equated. Lastly, tests are not always appropriate for the group that takes them. These issues often arise when best practices in design and data collection are not followed.

Best practices for the anchor test design

The anchor test design is often used because of the greater flexibility it provides. The role of the anchor test is to provide a common score that can be used to adjust for group ability differences before adjusting for test difficulty differences via equating. Scores from short anchor tests tend to have inadequate reliabilities, and consequently less than desirable correlations with the test scores. Low correlations may also result when the content of the anchor test differs from the test. Context effects can affect the comparability of anchor items. Anchors that are too hard or too easy for the target population produce skewed score distributions that are not helpful for equating.

Statistical procedures are needed to adjust for ability differences between groups when the anchor test design is used. Assumptions need to be made in order to make these adjustments; however, the assumptions may be flawed. In the case of the anchor test design, it is particularly important to employ multiple models as each model rests on different sets of assumptions. It is valuable to equate with several different models, including both linear and equipercentile models.

The equipercentile version of a particular method (say PSE) can be compared to the linear method version using the standard error of equating which describes sampling error and the difference that matters (DTM), an effect size that can be used to assess whether differences in equating functions have practical significance or are an artifact of rounding. Holland and Dorans (2006) describe the DTM, the standard error of equating and the standard error of the difference in equating or SEED. If the departures from linearity are less than the DTM and less than what would be expected due to sampling error, the linear model is often chosen on the grounds of parsimony because it is not sufficiently falsified by the data. Otherwise, the more general, less falsifiable, equipercentile model is selected.

With the anchor test design, the old and new form sample may perform very differently on the anchor test. Large ability differences on the anchor test tend to yield situations where equating is unsatisfactory unless the anchor is highly related to both tests to be equated. In this setting, different equating methods tend to give different answers unless the anchor test is strongly related to both the old and new tests. This divergence of results is indicative of a poor data collection design. The search for a single best model that could be employed universally would be unwise data analysis (Tukey, 1963).

Whether an equating test is an external anchor or an internal anchor also has an impact, as does the number of anchor tests and the type of score linking plan employed. For the observed score equating methods, described above, where the score on the anchor test plays an important role, it is often advised that the anchor test be a “mini-version” of the two tests being equated (Angoff, 1971).

Best practices for external anchors

The phrase external anchor usually refers to items that are administered in a separately-timed section and that do not count toward the test taker's score. For best practices, it is important to disguise the external anchor test so that it appears to be just another section of the test. This cannot always be achieved for tests that are composed of multiple distinct sections in which no one section contains all the item types. In these cases the section with the external anchor test is made to look like one of the test sections, preferably the one with the most item types represented.

Best practices for internal anchors

Items in an internal anchor test are part of the assessment and count toward each test taker's score. Internal anchor items are usually spread throughout the test. Because the items in an internal anchor test count toward the score, test takers are unlikely to skip them. On the other hand, once anchor test items have been used in the test administration of the old form, the items may become susceptible to security breaches and become known by test takers taking the new form to be equated. Security breaches are an unfortunate reality for many testing programs, and due diligence is required to prevent them or to recognize them when they occur.

Changes in the statistical properties related to item difficulty can occur with internal anchor tests due to context effects. Context effects can occur when common items are administered in different locations (e.g., common item 10 in one form is item 20 in the other form), or under different testing conditions (i.e., paper and pencil vs. computer delivered), or when they are adjacent to different kinds of items in the two tests. These effects are well documented (Leary and Dorans, 1985; Brennan, 1992; Harris and Gao, 2003). Mode of administration can also have an effect on the properties of an item, as noted by Gu et al. (2006), Keng et al. (2008), Kim and Huynh (2010) and Pommerich (2004, 2007).

Lack of invariance

An important property and consequence of score equating is that equating functions, to the extent possible, should be subpopulation invariant. These subpopulations should not be defined on the basis of the tests to be equated or the anchor test because the assumptions made by equating methods are sensitive to direct selection on the test or anchor.

The "same construct" and "equity" requirements are prerequisites for subpopulation invariance. One way to demonstrate that two tests are not equatable is to show that the equating functions used to link their scores are not invariant across different subpopulations of test takers. Lack of invariance in a linking function indicates that the differential difficulty of the two tests is not consistent across different groups. Note that subpopulation invariance is a matter of degree. In the situations where equating is usually performed, subpopulation invariance implies that the dependence of the equating function on the subpopulation used to compute it is small enough to be ignored.

Dorans and Liu (2009) and Liu and Dorans (2013) describe procedures that assess whether test scores on different forms that are expected to be used interchangeably are in fact interchangeable across different subpopulations. It uses the subpopulation invariance of linking functions across important subgroups (e.g., gender groups) to assess the degree of score exchangeability. These procedures focus on invariance at the reported score level. It is a basic quality control tool that can be used to assess whether a test construction process is under control, as can checks on the consistency of raw-to-scale conversions across test forms built to the same specifications (Haberman et al., 2008).

Score linkings that can be misconstrued as score equatings

Only tests that measure the same construct can be equated. Even though both height and weight are measures of size, height cannot be equated to weight. Likewise, a math test can be linked to a reading test, but it cannot be equated to a reading test. In this section, we discuss examples of score linkings that are not score equatings, but which might be misconstrued as score equating.

Dorans et al. (2007) use the linking taxonomy suggested by Holland and Dorans (2006) to examine score linkings that ranged from the relatively easy task of producing interchangeable scores on alternative versions of a test to the daunting challenge of aligning state standards with the National Assessment of Educational Progress (NAEP) scales. Since 2007, many more examples of attempts to link scores from different assessments have occurred, leading Moses (2022) to consider updating the framework provided by Holland and Dorans (2006) in an attempt to accommodate these various linkages.

Linkages across tests from different testing programs

The desire to compare scores from different assessments administered under different measurement conditions has always existed. Dorans et al. (2007) considered different types of linking that might be mistaken for score equating because they use data collection designs and statistical procedures associated with score equating. The interchangeability of scores associated with equating is not achieved, however, simply because certain numerical operations have been performed. Inferences that can be made from each of these types of linking are constrained by the features of the testing situation. Some prominent examples are discussed in this section.

Score linking is challenging when it involves two different tests, such as the SAT and ACT, that are produced by different assessment companies. Without a concordance, which provides a data-based path between the score scales of the two tests, the temptation might be to presume that the percentile ranks found in the norms tables for the two tests are comparable. That is rarely the case. For example, as reported in the summary of linkages in a joint cohort of ACT and SAT test takers between the ACT and the 2005 version of the SAT provided in Dorans (2020), a particular test taker in 2006 was more likely to achieve a score at the 75th percentile or above on the ACT than on the SAT because the SAT norms group was more able than the ACT norms group. Concordance tables attempt to address this problem.

The concordances between the ACT and the 2005 version of the SAT in a joint cohort of ACT and SAT test takers, summarized in Dorans (2020), provided linkages between ACT and SAT scores. Unlike equating, which produces the same link between scores across subgroups, concordances between scores from different tests such as the SAT and ACT can be subgroup dependent; for example, the linking between math score differed for males and females. Rather than replicate the design used in the earlier concordance based on a joint cohort of ACT and SAT test-takers, the College Board decided to update the ACT/SAT concordances through chains of concordances, some of which are based on self-selected examinees participating in administrations given prior to the official launch of the 2015/2016 version of the SAT (College Board, 2017, pp. 114–119). These concordances were criticized (Anderson, 2016a,b) and had to be replaced with direct ACT/SAT concordances that employed a superior collection design (College Board, 2018) similar to the design used jointly by ACT and the SAT programs to link the 2005 version of the SAT to the ACT (ACT, 2009).

Even when concordances are based on a single population of test takers, it is important to recognize that concordances are based on atypical subpopulations of their respective cohorts of test takers, namely those who take both tests. They may not work well where needed with test takers who take only one of the two tests. Lee and Brennan (2021) critique an attempt to use scale alignments of predicted scores to link the GRE and LSAT tests on a sample of matriculated law school students who took both tests. See Pommerich (2016) for an extensive treatment of the fairness issues associated with concordances.

There is often an interest in comparing performance across tests of different levels of difficulty for a given construct. Test scores are often compared across grades even though test content and test-taker groups differ. Vertical scaling is the term used to describe linkages that attempt to make these comparisons meaningful despite the changes in content and examinees that occur between grade levels, and despite the potential sensitivity of scaling results to differences in populations, grades and ages (Holland, 2007). Yen (2007), Briggs and Weeks (2009) and Kolen and Brennan (2014) address issues in vertical linking in greater detail.

Late last century the National Research Council published *Uncommon Measures* (Feuer et al., 1999). The motivation for that report was the debate in the late 1990s between those who favored voluntary national tests as a means of assessing the educational progress of students across the nation, and those who believed that statistical linkages among existing tests could be used to achieve that purpose. That volume examined the feasibility of linking the results of commercial and state tests to compare one student's achievement with national and international benchmarks as well as with those of students in other places. Feuer et al. demonstrated that it was not possible to link state assessments for reasons which were summarized by Thissen (2007, 2016). Dorans (2020) noted that the issues raised in *Uncommon Measures* (Feuer et al., 1999) remained relevant twenty years later.

Reardon et al. (2021) reported linkings of school district means on state tests to the NAEP scale, using statistical procedures that infer the district means on the state tests from published distributions of passing rates. Commentaries on Reardon et al. by Bolt (2021), Davison (2021), Moses and Dorans (2021) and von Davier (2021) re-emphasized longstanding cautions that the large variation in blueprints used by different state tests, as well as differences with NAEP with respect to content, administration conditions, stakes and examinee motivation are likely to make the interpretability of linking results restricted to comparing districts within states. Moses and Dorans (2021) provided empirical demonstrations that state-based linkings of district means from one test to another are not invariant across states even when the correlation of means within each state is very high.

Recent examples of concordances within testing programs

Recent events like the Covid-19 pandemic, repurposing of exams to meet multiple needs, competition for a dwindling supply of test takers, and innovative assessments have all contributed to a combination of increased demand for comparisons across different versions of the same test across disparate measurement conditions within testing programs. Responses to this increased demand for comparisons can produce linkings that may not be equatings.

Some testing programs have decided to shorten their tests. In order to be equated, tests need to produce equally reliable scores. A short test that produces erratic scores cannot be equated to a long test that produces very stable scores. Holland and Dorans (2006) used the term calibration to refer this type of linkage that does not produce interchangeable scores.

In addition to—or as an alternative to—shortening their tests, many large-scale testing programs transitioned from paper-pencil to online administration models in an effort to continue administering their tests with COVID-based closures to schools and testing centers. Camara (2020) described the risks of integrity, accuracy, reliability, validity and fairness of the scores produced from admissions and placement testing programs that administered online forms that were very different from their usual paper-pencil forms and standardized administrations. Linking across different modes of assessment depends on the similarity of the administration modes. Whereas the linking of linear digital administration forms to paper-pencil administration forms may be closer to equatings, the linking of adaptive digital administration forms to paper-pencil forms that involve not only different modes of administration but also different blueprints are less likely to be equatings (Eignor, 2007; Pommerich, 2016).

Sometimes, the administration of different tests in different modes or other different conditions of measurement involve groups that are not equivalent with respect to the constructs being measured and for which there are no common items. Baldwin and Clauser (2022) and Kim and Walker (2021) mention a variety of approaches to dealing with collateral quantitative information in place of common items. Some approaches to this problem involve using examinee background characteristics that are common to the administration groups to statistically weight or match the groups (Haberman, 2015; Livingston, 2014a; Wiberg and Branberg, 2015). These background characteristics measure something different than the tests being linked. When the background variables are weakly related to the scores to be linked, these approaches will not adequately adjust for group ability differences, as demonstrated in studies cited by Puhan and Kim (2022).

One option exercised in the absence of anchor test data may be to present a case for presuming that the scores from different modes or conditions are already aligned, and that no linking is needed. When users begin to compare the unlinked scores to the equated scores or when they assess whether different versions of the unlinked scores are comparable to each other; however, they may question the correctness of the properly equated scores as well as the unlinked scores. See Baldwin and Clauser (2022) for a discussion of interpretational difficulties that arose due to differences in unlinked and weakly linked scores for translated forms of the Psychometric Entrance Test (Beller, 1994).

Best practices for reporting score linkings that are not equatings

Best practices associated with unlinked or weakly-linked scores are to either avoid releasing scores based on questionable linkings and assumptions, or to release the scores by communicating both appropriate and inappropriate interpretations of those scores. For example, in the **Presmoothing score distributions** section, we recommended that the reporting of unequated raw scores for situations where the test administration sample sizes were too small for equating be accompanied with the caveat that the raw scores are not comparable to other scores, and that they would be made comparable when adequate data become available. Recommended communications might include cautions about limitations and potential misinterpretations of concordances within and across testing programs (Brennan, 2007; Pommerich, 2016). Other examples urging prudent interpretation of linkings between tests that do not measure the same thing and are designed for different purposes include the recommendation that linked GRE and LSAT scores not be used interchangeably for law school admissions without additional study (Lee and Brennan, 2021), and that linked district-means on NAEP and state tests have uses that should be restricted to within state comparisons (Moses and Dorans, 2021).

Summary

The comparability of measurements made in differing circumstances by different methods and investigators is a fundamental precondition for all of science. Psychological and educational measurement is no exception to this rule. Test equating techniques are those statistical and psychometric methods used to adjust scores obtained on different tests measuring the same construct so that they are comparable (Dorans and Holland, 2000, p. 281).

In high stakes settings, all test takers should be given an opportunity to maximize their performance under fair measurement conditions. Scores across different versions of the test should be as equivalent as possible. Some aspects of fairness lie outside the locus of control of test developers, such as opportunity to learn, quality of teaching, and test-taker motivation. Some components, however, can and should be controlled by the test developer and test administrator. Multiple versions of a test should be as parallel as possible. Score reliability should be large enough to justify the trust that users place in the scores. Test instructions should be clear and widely accessible to prospective test takers. Measurement conditions should be standardized and secure. Score equating procedures, when properly applied to reliable scores obtained under sound data collection conditions, can yield scores that may be treated as equivalent. The aspirations of score equating are worth achieving. When the same procedures are applied to unreliable scores from tests built to different specifications and administered under insecure, unstandardized conditions they fail to achieve the goal of score equating. Checking for equal and high reliability across versions of the test and the subpopulation invariance of the linking function are empirical checks on the quality of the linking results. These checks are essential for assessing the extent to which interchangeability and comparability are achieved, for communications about the quality and appropriate interpretations of linking results, and also for discouraging misinterpretations and misuse.

References

- ACT, 2009. ACT-SAT Concordance Tables. Retrieved from. <https://www.act.org/content/dam/act/unsecured/documents/ACTCollegeBoardJointStatement.pdf>.
- Anderson, N., 2016a. Why Your New SAT Score is Not as Strong as You Think It is. Washington Post. Retrieved from. <https://www.washingtonpost.com/news/grade-point/wp/2016/05/11/why-your-new-sat-score-is-not-as-strong-as-you-think-it-is/>.
- Anderson, N., 2016b. What's a College Test Score Worth? An ACT-vs.-SAT Dispute. Washington Post. Retrieved from. <https://www.washingtonpost.com/news/grade-point/wp/2016/05/12/whats-a-college-test-score-worth-an-sat-vs-act-dispute/>.
- Angoff, W.H., 1971. Scales, norms and equivalent scores. In: Thorndike, R.L. (Ed.), *Educational Measurement*, second ed. American Council on Education, Washington DC, pp. 508–600.
- Baldwin, P., Clauser, B., 2022. Historical perspectives on score comparability issues raised by innovations in testing. *J. Educ. Meas.* 59 (2), 231–250.

- Beller, M., 1994. Psychometric and social issues in admissions to Israeli universities. *Educ. Meas.* 13 (2), 12–20.
- Bolt, D., 2021. Commentary on Reardon, Kalogrides, and Ho's "Validation methods for aggregate-level test scale linking: a case study mapping school district test score distributions to a common scale". *J. Educ. Behav. Stat.* 46 (2), 168–172.
- Braun, H.I., Holland, P.W., 1982. Observed-score test equating: a mathematical analysis of some ETS equating procedures. In: Holland, P.W., Rubin, D.B. (Eds.), *Test Equating*. Academic, New York, pp. 9–49.
- Brennan, R.L., 1992. The context of context effects. *Appl. Meas. Educ.* 5 (3), 225–264.
- Brennan, R.L., 2007. Tests in transition: discussion and synthesis. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, NY, pp. 161–175.
- Brennan, R.L., 2010. First-Order and Second-Order Equity in Equating. CASMA Research Report Number 30. CASMA, Iowa City: IA.
- Briggs, D.C., Weeks, J.P., 2009. The impact of vertical scaling decisions on growth interpretations. *Educ. Meas.* 28, 3–14.
- Camara, W., 2020. Never let a crisis go to waste: large-scale assessment and the response to COVID-19. *Educ. Meas.* 39, 10–18.
- College Board, 2017. SAT Suite of Assessments Technical Manual: Characteristics of the SAT. Retrieved from. <https://collegereadiness.collegeboard.org/pdf/sat-suite-assessments-technical-manual.pdf>.
- College Board, 2018. Guide to the 2018 ACT/SAT Concordance. Retrieved from. <https://satsuite.collegeboard.org/media/pdf/guide-2018-act-sat-concordance.pdf>.
- Davison, M.L., 2021. Commentary on "Validation methods for aggregate-level test scale linking: a case study mapping school district test score distributions to a common scale". *J. Educ. Behav. Stat.* 46 (2), 173–186.
- Dorans, N.J., Holland, P.J., 2000. Population invariance and the equitability of tests: basic theory and the linear case. *J. Educ. Meas.* 37 (4), 281–306.
- Dorans, N.J., Liu, J., 2009. Score Equity Assessment: Development of a Prototype Analysis Using SAT Mathematics Test Data Across Several Administrations. ETS Research Report No. RR-09-08. Educational Testing Service, Princeton, NJ.
- Dorans, N.J., Walker, M.E., 2007. Sizing up linkages. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, pp. 179–198.
- Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), 2007. *Linking and Aligning Scores and Scales*. Springer-Verlag, New York.
- Dorans, N.J., Moses, T., Eignor, D.E., 2011. Equating test scores: towards best practices. In: von Davier, A.A. (Ed.), *Statistical Models for Scaling, Equating and Linking*. Springer-Verlag, New York, pp. 21–42.
- Dorans, N.J., 2013. On attempting to do what Lord said was impossible: commentary on van der Linden's conceptual issues in observed-score equating. *J. Educ. Meas.* 50 (3), 304–314.
- Dorans, N.J., 2020. Uncommon Measures Revisited. ETS Research Report No. RR-20-04. Educational Testing Service, Princeton, NJ.
- Eignor, D.R., 2007. Linking scores derived under different modes of test administration. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, NY, pp. 135–159.
- Feuer, M.J., Holland, P.W., Green, B.F., Bertenthal, M.W., Hemphill, F.C. (Eds.), 1999. *Uncommon measures: Equivalence and linkage among educational tests* (Report of the Committee on Equivalency and Linkage of Educational Tests, National Research Council). Washington DC: National Academy Press.
- Gu, L., Drake, S., Wolfe, E.W., 2006. Differential item functioning of GRE Mathematics items across computerized and paper-and-pencil testing media. *JTLA* 5 (4), 1–30.
- Haberman, S., Guo, H., Liu, J., Dorans, N.J., 2008. Consistency of SAT[®] Reasoning Score Conversions. ETS Research Report No. RR-08-67. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/j.2333-8504.2008.tb02153.x>.
- Haberman, S.J., 2015. Pseudo-equivalent groups and linking. *J. Educ. Behav. Stat.* 40 (3), 254–273.
- Hanson, B.A., 1991. A note on Levine's formula for equating unequally reliable tests using data from the common item nonequivalent groups design. *J. Educ. Stat.* 16, 93–100.
- Harris, D.J., Gao, X., 2003. A conceptual synthesis of context effect. In: *Context Effects: Implications for Pretesting and CBT*. Symposium Conducted at the 2003 Annual Meeting of the American Educational Research Association, Chicago, IL.
- Holland, P.W., Dorans, N.J., 2006. Linking and equating. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education/Prager, Westport, CT, pp. 187–220.
- Holland, P.W., Hoskens, M., 2003. Classical test theory as a first-order item response theory: application to true-score prediction from a possibly nonparallel test. *Psychometrika* 68, 123–149.
- Holland, P.W., Thayer, D.T., 1987. Notes on the use of log-linear models for fitting discrete probability distributions. Educational Testing Service (Technical Report 87-79), Princeton, NJ.
- Holland, P.W., Thayer, D.T., 2000. Univariate and bivariate loglinear models for discrete test score distributions. *J. Educ. Behav. Stat.* 25, 133–183.
- Holland, P.W., 2007. A framework and history for score linking. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, NY, pp. 7–30.
- Holland, P.W., 2013. Comments on van der Linden's critique and proposal for equating. *J. Educ. Meas.* 50, 286–294.
- Hull, C.L., 1922. The conversion of test scores into series which shall have any assigned mean and degree of dispersion. *J. Appl. Psychol.* 6 (4), 298–300.
- Kelley, T.L., 1914. Comparable measures. *J. Educ. Psychol.* 5 (10), 589–595.
- Kelley, T.L., 1923. *Statistical Method*. Macmillan, New York.
- Keng, L., McClarty, K.L., Davis, L.L., 2008. Item-level comparative analysis of online and paper administrations of the Texas assessment of knowledge and skills. *Appl. Meas. Educ.* 21, 207–226.
- Kim, D., Huynh, H., 2010. Equivalence of paper-and-pencil and online administration modes of the statewide English test for students with and without disabilities. *Educ. Assess.* 15, 107–121.
- Kim, S., Walker, M., 2021. Assessing Mode Effects of at Home Testing Without a Randomized Trial. Research Report No. RR-21-10. ETS. <https://doi.org/10.1002/ets2.12323>.
- Kolen, M.J., Brennan, R.L., 2014. *Test Equating, Scaling, and Linking: Methods and Practices*, third ed. Springer, New York.
- Kolen, M.J., Jarjoura, D., 1987. Analytic smoothing for equipercentile equating under the common item nonequivalent populations design. *Psychometrika* 52, 43–59.
- Leary, L.F., Dorans, N.J., 1985. Implications for altering the context in which test items appear: an historical perspective on an immediate concern. *Rev. Educ. Res.* 55, 387–413.
- Lee, W., Brennan, R.L., 2021. An Independent Review of Educational Testing Service's Study on the Predictive Validity of the Graduate Record Exams for Making Law School Admissions Decisions. Retrieved from. https://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar/2021/20210907-casma-report-final.pdf.
- Liou, M., Cheng, P.E., Li, M.Y., 2001. Estimating comparable scores using surrogate variables. *Appl. Psychol. Meas.* 25, 197–207.
- Little, R.J.A., Rubin, D.B., 2002. *Statistical Analysis With Missing Data*, second ed. Wiley, Hoboken, NJ.
- Liu, J., Dorans, N.J., 2013. Assessing a critical aspect of construct continuity when test specifications change or test forms deviate from specifications. *Educ. Meas.* 32 (1), 15–22.
- Livingston, S.A., 2014a. Demographically Adjusted Groups for Equating Test Scores. Research Report No. RR-14-30. ETS. <https://doi.org/10.1002/ets2.12030>.
- Livingston, S.A., 2014b. *Equating Test Scores (Without IRT)*, second ed. Educational Testing Service, Princeton, NJ.
- Lord, F.M., Wingersky, M.S., 1984. Comparison of IRT true-score and equipercentile observed-score equatings. *Appl. Psychol. Meas.* 8, 453–461.
- Lord, F.M., 1950. Notes on Comparable Scales for Test Scores. ETS RB-50-48. Educational Testing Service, Princeton, NJ.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Erlbaum, Hillsdale, NJ.
- Moses, T., Dorans, N.J., 2021. Aggregate-level test-scale linking: a new solution for an old problem? *J. Educ. Behav. Stat.* 46 (2), 187–202.
- Moses, T.P., 2011. Log-linear models as smooth operators: Holland's statistical applications and their practical uses. In: Dorans, N., Sinharay, S. (Eds.), *Looking Back: Proceedings of a Conference in Honor of Paul W. Holland*. Springer, New York, pp. 185–202. https://doi.org/10.1007/978-1-4419-9389-2_10.

- Moses, T.P., 2022. Linking by definition and presumed invariance across conditions of measurement. *J. Educ. Meas.* 59 (2), 231–250.
- Petersen, N.S., Kolen, M.J., Hoover, H.D., 1989. Scaling, norming, and equating. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. Macmillan, New York, pp. 221–262.
- Pommerich, M., 2004. Developing computerized versions of paper-and-pencil tests: mode effects for passage-based tests. *J. Technol. Learn. Assess.* 2 (6), 1–44.
- Pommerich, M., 2007. The effect of using item parameters calibrated from paper administrations in computer adaptive test administrations. *J. Technol. Learn. Assess.* 5 (7), 1–28.
- Pommerich, M., 2016. The fairness of comparing test scores across different tests or modes of administration. In: Dorans, N.J., Cook, L.L. (Eds.), *Fairness in Educational Assessment and Measurement*. Routledge, New York, pp. 111–134.
- Puhan, G., Kim, S., 2022. Score comparability issues with at-home testing and how to address them. *J. Educ. Meas.*
- Reardon, S.F., Kalogrides, D., Ho, A.D., 2021. Validation methods for aggregate-level test scale linking: a case study mapping school district distributions to a common scale. *J. Educ. Behav. Stat.* 46 (2), 138–167.
- Rosenbaum, P.R., 1995. *Observational Studies*. Springer-Verlag, New York.
- Thissen, D., 2007. Linking assessments based on aggregate reporting: background and issues. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, pp. 287–310.
- Thissen, D., 2016. Commentary on the assessment of the fairness of comparisons under divergent measurement conditions. In: Dorans, N.J., Cook, L.L. (Eds.), *Fairness in Educational Assessment and Measurement*. Routledge, New York, pp. 203–214.
- Tukey, J.W., 1963. Mathematics 596—an introduction to the frequency analysis of time series. In: Brillinger, D.R. (Ed.), *The Collected Works of John W. Tukey, Volume I: Time Series, 1949–1964*. Chapman & Hall, London.
- van der Linden, W.J., Wiberg, M., 2010. Local observed-score equating with anchor test designs. *Appl. Psychol. Meas.* 34, 620–640.
- van der Linden, W.J., 2013. Some conceptual issues in observed-score equating. *J. Educ. Meas.* 50, 249–285.
- von Davier, A.A., Holland, P.W., Thayer, D.T., 2004. *The Kernel Method of Test Equating*. Springer, New York.
- von Davier, A.A., 2010. Equating and scaling. In: Peterson, P., Baker, E., McGaw, B. (Eds.), *International Encyclopedia of Education*, third ed. Educational Assessment, pp. 50–55.
- von Davier, A.A., 2021. Commentary on “Validation methods for aggregate-level test scale linking: a case study mapping school district test score distributions to a common scale”. *J. Educ. Behav. Stat.* 46 (2), 203–208.
- Wiberg, M., Bränberg, K., 2015. Kernel equating under the non-equivalent groups with covariates design. *Appl. Psychol. Meas.* 39 (5), 349–361.
- Yen, W.M., 2007. Vertical scaling and no child left behind. In: Dorans, N.J., Pommerich, M., Holland, P.W. (Eds.), *Linking and Aligning Scores and Scales*. Springer-Verlag, New York, pp. 273–283.

Differential functioning

Cindy M. Walker^a and Sakine Gocer Sahin^b, ^a Research Analytics Consulting, LLC, Pittsburgh, PA, United States; and ^b WIDA, Wisconsin Center for Education Research, University of Wisconsin-Madison, Madison, WI, United States

© 2023 Elsevier Ltd. All rights reserved.

What is differential functioning?	249
Differential item functioning	249
Uniform and non-uniform DIF	250
Differential bundle functioning	251
Differential test functioning	252
Differential person functioning	252
Methods used to detect DIF	252
Mantel Haenszel (MH)	253
SIBTEST	254
Poly-SIBTEST	254
Logistic regression	255
Methods based on item response theory (IRT)	255
MIMIC model	256
Latent DIF analysis	256
How are DIF analyses used operationally?	256
Why are DF analyses an important part of test development and validation?	257
Practical considerations	257
References	257

What is differential functioning?

Differential Functioning (DF) is a term derived from the concept of item bias which resulted in the development of Differential Item Functioning (DIF). Studies using this methodological approach date back to the 1960s, in attempt to determine why there were statistically significant differences, on average, between Minority and Caucasian students on achievement test performance, as measured by large-scale standardized tests (Angoff, 1993). The word bias can be defined as a divergence from expression or indication of truth. Statistically, it is a systematic distortion of a statistic, in this case the mean, for a factor that was not considered in its derivation. The concept underlying item bias is that the statistically significant differences found between Minority and Caucasian students on overall test performance was due to the additive impact of multiple items that were biased (Angoff, 1993).

Differential item functioning

Although the methodology of DIF was developed primarily for use on large-scale standardized tests, this approach can be used on any data collection tool for which an individual responds to multiple items that, taken as a whole, measure a construct. For example, DIF analyses could be conducted on surveys that measure self-esteem, self-efficacy, attitudes, beliefs, or personality. DIF analyses are conceptually based on the idea that some items are measuring one or more constructs in addition to the primary construct being measured; and that there are two or more groups that differ in their underlying distribution on these secondary dimension(s). If groups do not differ in their underlying distribution on the additional secondary measured by some items then DIF will not occur because, theoretically, the reason that items function differentially is based on the interaction between item and examinee characteristics.

Many educational and psychological tests are believed to multidimensional, meaning they measure at least two constructs. For example, research has demonstrated that mathematics items that ask examinees to explain something in writing, or to read complex passages, function differentially against less proficient writers and readers, respectively (Walker and Beretvas, 2001; Walker et al., 2008). In the first example the DIF analyses were conducting using writing proficiency as the grouping variable, because it was hypothesized that the secondary dimension measured by test items was writing ability. In the second example the DIF analyses were conducting using reading proficiency as the matching variable, because it was hypothesized that the secondary dimension measured by test items was reading ability. Then examinees were matched using total score on the mathematics items that did not require examinees to write, in the first case, or read, in the second case. In both cases the existence of DIF does not imply that these items are biased, since the addition of these item types increases the construct validity of the test, with respect to the National Council of Teachers of Mathematics (NCTM) Standards (NCTM, 2000). Therefore, the additional dimensions are considered *auxiliary* and the DIF is considered *benign* (Roussos and Stout, 1996a). It should be noted, however, that if only one score is reported, as opposed to reporting scores on mathematics ability and mathematical writing or mathematical reading ability, some

proficiency classification will be biased if a unidimensional *item response theory* model was used to create scale scores (Walker and Beretvas, 2001). This is particularly true for students whose overall test scores are near the cut points that determine proficiency classifications.

If the additional constructs measured by some items are unintended then item bias is said to exist (Ackerman, 1992; Zumbo, 1999). This type of DIF is sometimes referred to as *adverse* DIF while the additional constructs measured by some items are considered *nuisance* dimensions (Roussos and Stout, 1996a). For example, if a mathematics problem-solving test item also required examinees to understand how points are scored in the game of soccer then this item would function differentially against examinees who were not familiar with how points are scored in soccer. If it were also true that, on average, males were more familiar with these rules than females then this item would also function differentially against females. In both cases, examinees with an understanding of how points are scored in soccer have an unfair advantage in obtaining a correct answer to this test item, that has nothing to do with the construct of mathematical problem solving.

In DIF analyses, the group that is suspected to have an advantage due to the multidimensionality of some items is referred to as the reference group. Conversely, the group that is suspected to have a disadvantage is referred to as the focal group. There are many ways to test for DIF. However, in all cases the sign of the statistic obtained from the analysis indicates which group the item favors. Therefore, while researchers suggest that DIF analyses be conducted based on *a priori* hypotheses about the additional dimensions that are thought to influence performance (Gierl, 2005); it is also possible to conduct exploratory DIF analyses. The examples presented earlier were both conducted based on *a priori* hypotheses about the secondary dimensions. In the first example students that were classified as minimal or basic in writing composed the focal group, while those that were classified as proficient or advanced composed the reference group. A similar approach was taken for the second example presented earlier. Exploratory DIF analyses are routinely conducted during the test development process, when every test item is evaluated for the presence of DIF, using performance on all, or some, of the remaining test items as the matching variable. If exploratory DIF analyses are conducted, however; one should adjust the alpha level of the statistical test, to control for Type I errors, just as one would do when conducting many hypotheses tests for any other statistic.

Uniform and non-uniform DIF

DIF analyses can be conducted on both dichotomous and polytomous items. However, for the sake of simplicity we will focus only on dichotomous items in this section, knowing that all of the concepts presented can be generalized to polytomous items. There are two types of DIF that can occur, uniform and non-uniform. When *uniform DIF* occurs, the difference between the probability of responding favorably, or correctly, to an item is constant for all respondents in the reference group, regardless of their level of the underlying trait being measured. Fig. 1 illustrates an Item Characteristic Curve (ICC) for an item that is not functioning differentially. The y-axis reflects the conditional probability of endorsing, or correctly answering, this item. The x-axis reflects the latent trait scale. As Fig. 1 illustrates, when DIF does not occur there is no difference between respondents in the reference and focal group, in the conditional probability of endorsing, or correctly answering, an item. Fig. 2 illustrates uniform DIF. As the figure illustrates, the conditional probability of correctly responding to, or endorsing, an item is always higher for respondents in the reference group. In addition, the difference in the probability of correctly responding to this item is approximately constant across all ability levels for

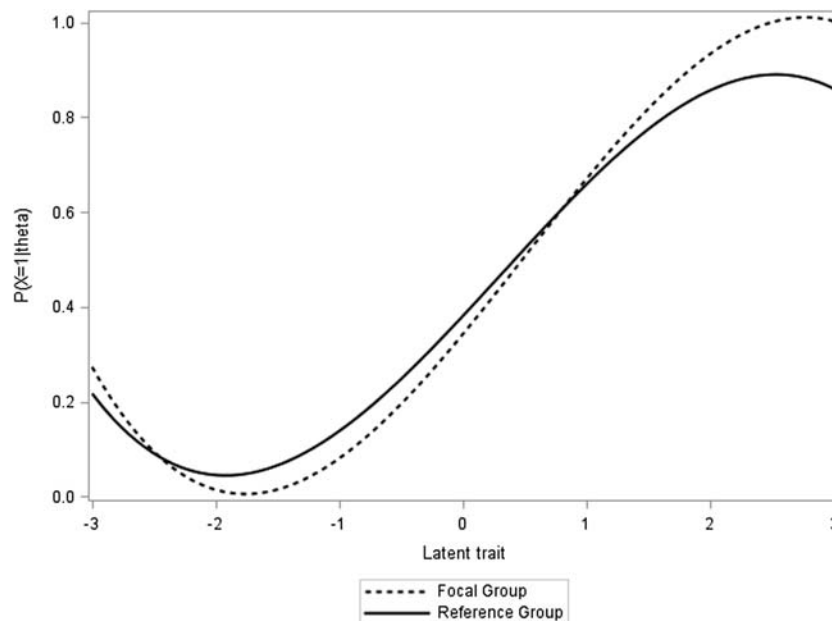


Fig. 1 An ICC of an item when there is no DIF.

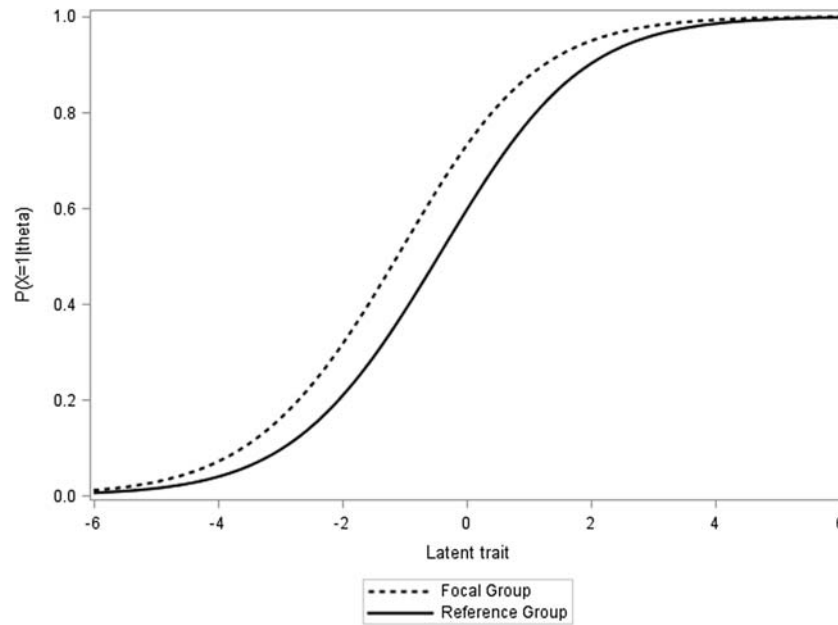


Fig. 2 An ICC of an item when there is uniform DIF.

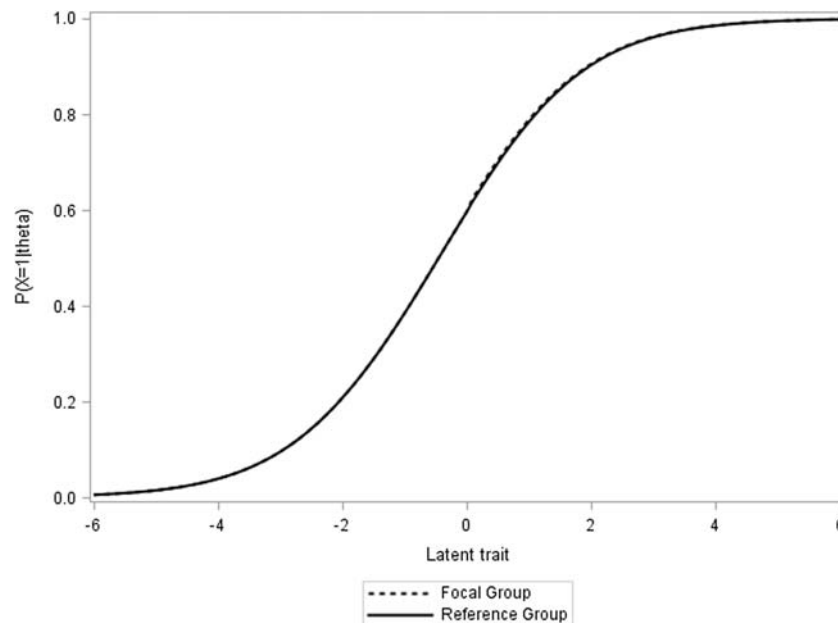


Fig. 3 An ICC of an item when there is non-uniform DIF.

reference and focal group examinees. Conversely, when the difference in the probability of correctly responding to an item is not constant across all ability levels for the two groups, but rather depends on whether a respondent is high or low on the latent trait then *non-uniform DIF* is said to occur. In extreme cases, non-uniform DIF can result in *crossing DIF* as illustrated in Fig. 3. As the figure illustrates the conditional probability of endorsing or answering an item correctly is lower for respondents in the reference group with lower estimates of the latent trait and higher for respondents that are higher on the latent trait.

Differential bundle functioning

When multiple items on a measure are found to function differentially then Differential Bundling Functioning (DBF) is said to occur. DBF analyses should be conducted when one hypothesizes that the small DIF amount on single items are caused by the same secondary dimension because taken as a group they can be amplified to produce a significant amount of DIF across all the

items in a bundle (Nandakumar, 1993). In fact, many researchers suggest that DBF analyses based on substantive hypotheses be conducted whenever possible because they help us to understand *why* DIF occurs. This approach has great potential for helping researchers understand the cultural and cognitive reasons for the occurrence of DIF. It has been used to study the impact of language spoken, culture, accommodation status, reading ability, writing ability, and gender on test performance (Abbott, 2006, 2007; Banks, 2006; Finch et al., 2009; Gierl et al., 2003, 2001; Mendes-Barnett and Ercikan, 2006; Roussos and Stout, 1996b; Ryan and Chiu, 2001; Ryan and Fan, 1996; Walker and Beretvas, 2001; Walker et al., 2008).

An item bundle can be a group of items that have something in common that researchers believe constitute a secondary dimension. For example, one might hypothesize that all science items that require an examinee to mentally manipulate two- or three-dimensional figures will function differentially against females, since research has demonstrated that males, on average, are better at spatial visualization than females, on average. However, if a set of test items all share a common reading passage or graph then they should always be tested as a bundle because they are based on a common stimulus (Rosenbaum, 1988). Bundle can be created in a confirmatory manner, based on hypotheses and expert opinion, or in an exploratory manner, based on a statistical procedure such as factor or cluster analysis.

Differential test functioning

When DBF occurs, if all the items in the bundle are found to function against the focal group then Differential Test Functioning (DTF) is said to occur. In this case, depending on the number of items found to function differentially, one might hypothesize that statistically significant differences would be found on the mean of overall test scores between the two groups. This is not to be confused with impact, which refers to group differences on the primary construct measured by test items. Rather, DTF implies that there are group differences based on the secondary construct(s) measured by test items. On the other hand, when DBF occurs for both groups, meaning some items are found to function in favor of the focal while other items are found to function in favor of the referenced group, then cancellation can occur (Nandakumar, 1993). In this case, depending on the number of items that are found to cancel each other out, one would hypothesize that statistically significant differences would not be found on the mean of overall scores between the two groups.

In fact, research has demonstrated that there is a strong correlation between finding statistically significant differences between focal and reference group examinees on overall test score and the number of DIF items, as well as the magnitude of the DIF (Walker et al., 2012). However, this will only occur if the magnitude of the DIF is moderate and all items function differentially against the same group. Therefore, items that are found to function differentially should not automatically be removed from the test or measure. When the magnitude of DIF is negligible the aggregate of DIF items may not impact overall performance. Likewise, when cancellation occurs the impact of DTF on overall performance may be minimal. In these cases, automatically removing items that are found to function differentially may jeopardize the reliability and content validity of the test or measure. Conversely, when the magnitude of DIF amount is high and all items favor the reference group, the presence of DTF is problematic and items should be eliminated or rewritten.

Taken as a group, the concepts of DIF, DBF, and DTF are all different forms of DF. In all cases, the presence of any type of DF, be it adverse or benign, implies that there is not measurement invariance in the measure being used. This implies that the same overall score on the measure means different things for examinees, depending on what group they belong to which, in turn, implies that the validity of the test may be threatened.

Differential person functioning

Differential Person Functioning (DPF) is another conceptualization of differential functioning in which persons, as opposed to items, are investigated individually to determine if they perform similarly on two subsets of test items after controlling for item difficulty (Johanson and Alsmadi, 2002). In this procedure, item and person parameters are held constant, except for the person for whom DPF is being calculated. This procedure can be used to identify examinees who have cheated on a high-stakes exam, such as a certification exam, if a set of test items are administered to examinees that are known to be secure (O'Leary and Smith, 2017). Specifically, the performance of each individual examinee on the set of operational items that are suspected of being compromised is compared to their performance on the set of items that are known to be secure. If an examinee performs similarly on the operational items and the secure items, then an examinee would not be flagged for suspect behavior. If, on the other hand, an examinee performs well on the operational items but does poorly on the items that are known to be secure then they would be flagged for suspect behavior (Smith and Davis-Becker, 2011). This procedure can also be used to determine if an examinee performs equally well on different item formats or items in different content domains (Johanson and Alsmadi, 2002). It has also been referred to in the literature as person-fit statistics (Sinharay, 2016) and is the conceptualization behind the science of data forensics (Wollack, 1997).

Methods used to detect DIF

There are more than 20 different methods to detect DIF, all of which are conceptually based on the same general idea. Specifically, DIF detection methodologies statistically match two groups on total score or ability, the reference and focal group, that are created

based on a criterion variable (such as race or gender) and compare the performances of both groups on a particular item, after controlling for impact. Impact can be defined as true differences that exist between the two groups on the primary construct measured by non-DIF items. These methods differ in how they match respondents in the two groups—raw score or latent trait, how the groups are statistically equated or matched to control for impact; whether they can be used to test dichotomous or polytomous items; whether they are parametric or non-parametric tests; and, the statistical framework they are based on—categorical data analysis, item response theory, structural equation modeling, or latent class analysis. It should be noted that very few procedures allow one to test for DBF or DTF. Specifically, at the writing of this article, only SIBTEST and a specific application of Structural Equation Modeling (SEM) can be used to test for DBF or DTF.

Next, we explain, in greater detail, a representative subset of these methods.

Mantel Haenszel (MH)

Mantel–Haenszel (MH) is a non-parametric chi-squared test introduced by Mantel and Haenszel (1959) and later developed by Holland and Thayer (1988). In this method the total overall score on the test or measure is used as a matching variable.

The first step in calculating the MH test statistic is to form k subgroups of respondents for each possible total score. For example, on a ten-item multiple choice test there would be 11 groups of respondents, assuming that all possible scores from zero to ten were present in the data. Then a 2×2 contingency table is created, as shown in Table 1, for each of the k total scores to compare the performance of reference and focal group examinees on each individual item. The null hypothesis for this procedure is that the odds of responding correctly to item i are equivalent for respondents in the focal and reference groups for all k possible total scores on the measure.

$$\alpha_{MH} = \frac{\sum_j A_j D_j / T_j}{\sum_j B_j C_j / T_j} \quad (1)$$

Eq. (1) depicts the formula used to calculate the MH test statistic, which is used to estimate the odds ratio for reference and focal group respondents with a total score of j on item i . The value of the test statistic can range between 0 and infinity. A value of 1.0 indicates that one cannot reject the null hypothesis and thus there is no DIF. A value greater than 1.0 is indicative of DIF in favor of reference group while a value less than 1.0 is indicative of DIF in favor of the focal group. To simplify interpretation of the test statistic, Holland and Thayer (1985) proposed standardizing it by using a log transformation, thereby centering the value around 0. This transformed test statistic, $MH\Delta$ –DIF is depicted in Eq. (2).

$$MH\Delta - DIF = -2.35 \ln(\alpha_{MH}) \quad (2)$$

For the $MH\Delta$ –DIF statistic, a value of 0 indicates no DIF; negative values indicate DIF in favor of the reference group; and, positive values indicate DIF in favor of the focal group. Educational Test Service (ETS) developed guidelines for evaluating the effect size of this test statistic (Doran and Holland, 1993; Zwick, 2012). Specifically,

- Type A items—negligible DIF: items with $|MH\Delta - DIF| < |1|$ or $MH\Delta - DIF$ is not statistically significantly different from zero.
- Type B items—moderate DIF: items with $1 \leq |MH\Delta - DIF| \leq 1.5$, and statistically significant.
- Type C items—large DIF: items with $|MH\Delta - DIF| > 1.5$, and statistically significantly different from 1.

Type A items are considered to be DIF free items while Type B items should only be used if there are no equivalent items to replace them and Type C items should only be used if they are essential to meet test specifications (Zwick and Ercikan, 1989).

Utilizing MH to test for DIF is relatively easy and the results obtained from this DIF detection procedure are consistent with the results obtained using an IRT approach for DIF detection, if the 1-PL model is assumed (Holland and Thayer, 1988; Guo and Dorans, 2020); therefore, it is commonly used. However, one disadvantage of using this method is that it can only detect non-uniform DIF under some conditions (Masor et al., 1994). Another disadvantage of using this methodology is that the reference and focal groups are matched based on the number of correct responses on a test which is not an efficient matching variable when DIF is present in some items. Finally, this procedure cannot be used to test for DTF or DBF.

Table 1 Contingency table used to calculate the MH test statistic for item i .

	<i>Correct</i>	<i>Incorrect</i>	<i>Total</i>
Reference	A_j	B_j	$A_j + B_j$
Focal	C_j	D_j	$C_j + D_j$
Total	$A_j + C_j$	$B_j + D_j$	$T_j = (A_j + B_j + C_j + D_j)$

Note: A, B, C, and D corresponds to frequencies of each cell, j corresponds to the level of the total score.

SIBTEST

The SIBTEST procedure is also a nonparametric procedure for detecting DIF. SIBTEST was developed by [Shealy and Stout \(1993\)](#) in order to detect uniform DIF in dichotomous data and was further developed to test for non-uniform DIF and to detect DIF in polytomous items. In the SIBTEST approach, items are divided into two subsets, a matching subtest and the suspect subtest that is hypothesized to be functioning differentially. The suspect subtest can consist of a single item, when testing for DIF, or a bundle of items, when testing for DBF or DTF. The matching subtest consists of the rest of the items or a subset thereof. It is important that this matching subtest not include any items that are functioning differentially. Therefore, researchers have suggested using an iterative approach to purify the matching subtest by removing any items that are flagged for DIF in an exploratory analysis ([Gierl, 2005](#)). The null hypothesis tested using this procedure is expressed as:

$$H_0 : \beta = 0 \text{ vs. } H_1 : \beta \neq 0 \quad (3)$$

Using total test score on the matching subtest as a substitute for θ , an estimate of β is given by

$$\hat{\beta} = \sum_{k=0}^K P_k d_k \quad (4)$$

where P_k represents the proportion of examinees in that obtain a score of k on the conditioning valid subtest and $d_k = \bar{Y}_{Rk}^* - \bar{Y}_{Fk}^*$. $\bar{Y}_{Gk}^*$ is defined to be the average performance on the studied item for examinees in group G (i.e., the reference or focal group) who obtained a score of k on the conditioning subtest, corrected for bias by controlling for differences in items thought to be measuring the secondary dimension. SIBTEST also yields an overall statistical test for β . The test statistic, SIB, for evaluating the null hypothesis is

$$\text{SIB} = \frac{\hat{\beta}}{\hat{\sigma}(\hat{\beta})} \quad (5)$$

where $\hat{\sigma}(\hat{\beta})$ is the estimated standard error of $\hat{\beta}$. [Shealy and Stout \(1993\)](#) indicated that this test statistics asymptotically follows a normal distribution with mean of 0 and variance of 1 under the null hypothesis of no DIF.

[Roussos and Stout \(1996a,b\)](#) developed guidelines for interpreting the results obtained SIBTEST. Specifically:

- Type A items—negligible DIF: items with $|\hat{\beta}| < 0.059$
- Type B items—moderate DIF: items with $0.059 \leq |\hat{\beta}| < 0.088$
- Type C items—large DIF: items with $|\hat{\beta}| \geq 0.088$.

Poly-SIBTEST

Poly-SIBTEST is an extension of SIBTEST for polytomous items. In this procedure DIF-free subtest scores are used to represent the ability level of the latent trait of interest. The expected item scores can be estimated as:

$$ES_R(t) = \sum_{k=1}^m k P_{Rk}(t) \quad (6)$$

and

$$ES_F(t) = \sum_{k=1}^m k P_{Fk}(t) \quad (7)$$

where P_{Rk} and P_{Fk} are equal to the proportion of examinees in the reference group and focal group, respectively, that obtained a score of k on the studied item and have a subtest score of t .

A DIF index is estimated using an average difference of these adjusted scores, weighted by the proportion of examinees obtaining DIF-free subtest score equal to t . The estimated DIF index is represented by $\hat{\beta}$ where,

$$\hat{\beta} = \sum_{t=1}^T \left([ES_R^*(t) - ES_F^*(t)] \frac{N_R(t) + N_F(t)}{N} \right) \quad (8)$$

where T equals the maximum possible score on the DIF-free subtest, $N_R(t)$ and $N_F(t)$ are equal to the numbers of examinees obtaining DIF-free subtest score t from the reference group and the focal group respectively, and N is the total number of examinees. For large samples this DIF index, $\hat{\beta}$, can be assumed to be approximately normally distributed when the null hypothesis of no DIF is assumed ([Bolt, 2002](#)). Therefore, $\hat{\sigma}_{\beta}$ can be found using the following formula:

$$\hat{\sigma}_{\beta} = \sqrt{\sum_{t=0}^T \left[\frac{N_R(t) + N_F(t)}{N} \right]^2 \left[\frac{\hat{\sigma}_{Rt}^2}{N_{Rt}} + \frac{\hat{\sigma}_{Ft}^2}{N_{Ft}} \right]} \quad (9)$$

Then a test statistic, SIB, can be obtained by dividing the DIF effect measure by its standard error (Bolt, 2002):

$$\text{SIB} = \frac{\hat{\beta}}{\hat{\sigma}_{\beta}} \quad (10)$$

This test statistic can be evaluated against a standard normal distribution.

Logistic regression

Swaminathan and Rogers (1990) proposed the use of logistic regression for DIF detection with dichotomous items and it has been shown to be very effective (French and Miller, 1996; Swaminathan and Rogers, 1990; Zumbo, 1999). Like the MH procedure, the matching variable used is the overall total score on the test or measure. However, whereas MH considers the total score to be a categorical variable it is considered a continuous variable in logistic regression. When using logistic regression to test for DIF three different logistic regression models are fit to the data, a full model and two models that are nested within the full model. These models are depicted in Eqs. (11)–(13). In the full model the logit of the probability of a correct response to a particular item is predicted using total score, group (reference or focal), and the interaction of total score and group as depicted in Eq. (11). If this model is the best fit to the data then non-uniform DIF is present in the item being tested. This fit of this model can be compared to the fit of a nested model that eliminates the interaction term as depicted in Eq. (12). Specifically, if L_1 = the log Likelihood obtained from fitting the model depicted in Eq. (11) to the data and L_0 = the log Likelihood obtained from fitting the model depicted Eq. (12) to the data then $G^2 = -2[\ln(L_0) - \ln(L_1)]$, which reflects the difference between the deviance of the two models, can be used to determine if the interaction term is needed in the model. G^2 follows a χ^2 distribution with one degree of freedom, since one less parameter is estimated in this nested model. If G^2 is statistically significant then the interaction term is needed in the model, which implies that the item being tested has non-uniform DIF. If G^2 is not statistically significant then the interaction term is not needed. The next step is to determine if uniform DIF is present in the item being tested. To do so the difference between the Likelihood Ratio Tests obtained from fitting the model depicted in Eq. (12) (L_1) and Eq. (13) (L_0) is calculated to obtain a new G^2 statistic with one degree of freedom. If this test is statistically significant then the group variable is needed in the model, which implies that the item being tests has uniform DIF. If this test is not statistically significant, then the item being tested is not functioning differentially.

$$\text{Logit}(P(Y = 1)) = \beta_0 + \beta_1 \text{total} + \beta_2 \text{group} + \beta_3 (\text{total} * \text{group}) \quad (11)$$

$$\text{Logit}(P(Y = 1)) = \beta_0 + \beta_1 \text{total} + \beta_2 \text{group} \quad (12)$$

$$\text{Logit}(P(Y = 1)) = \beta_0 + \beta_1 \text{total} \quad (13)$$

Since power increases as sample size increases, very large samples can result in statistically significant G^2 statistics that are not practically significant. To determine the practical implication of including the interaction term and/or the group term in the model, changes in R^2 can be utilized (i.e., ΔR^2). In fact, Jodoin and Gierl (2001) developed guidelines for interpreting the results obtained when using logistic regression to detect DIF. Specifically,

- Type A items—negligible DIF: items with $\Delta R^2 < 0.035$
- Type B items—Moderate DIF: items with $0.035 \leq \Delta R^2 < 0.070$
- Type C items—High DIF: items with $\Delta R^2 \geq 0.070$

One of the advantages of using logistic regression to detect DIF, when compared to MH, is that it is not necessary to categorize the continuous matching variable and it is possible to model both uniform and/or non-uniform DIF (Zumbo, 1999). Another advantage of the logistic regression method is that it can detect both uniform and non-uniform DIF. However, one of the disadvantages, when compared to SIBTEST or Poly-SIBTEST, is that only single items can be tested. In addition, when the focal and reference groups have unequal ability distributions, and the sample size of at least one of the groups is small, there is a loss of power (Narayanan and Swaminathan, 1996).

It should be noted that several researchers have extended this methodology for use with polytomous items (French and Miller, 1996; Miller and Spray, 1993; Zumbo, 1999).

Methods based on item response theory (IRT)

There are several DIF detection methodologies that are based on the framework of item response theory including the Likelihood Ratio Test (Thissen et al., 1986, 1993, 1998), Raju's Area Measures (Raju, 1988, 1990), and Lord's χ^2 (Lord, 1980). These methods differ in terms of what is used to compare the performance of reference and focal group examinees on the studied item. The Likelihood Ratio Test compares the fit of two models—one in which all item parameters are assumed to be equivalent for reference and focal group examinees (L_0) and one in which item parameters for the items being studied are allowed to vary. The calculation of the test statistic, G^2 , is identical to what was done when logistic regression was used to test for DIF. The degrees of freedom for this test are equal to the number of parameters allowed to vary. Raju's area measure evaluates the magnitude of the area between two ICCs,

one for the reference group and one for the focal group. Specifically, a z-statistic is used to test the null hypothesis that the area between the ICCs is zero. Lord's χ^2 is used to compare the item parameters obtained from scaling the reference and focal groups separately.

These methods are similar in that they all must somehow control for impact, or true differences between the two groups on the primary construct measured by test items. Raju's area measures and Lord's χ^2 both require separate calibrations of the item parameters for the reference and focal group. Therefore, an equating procedure must be used to ensure that the item parameter estimates for the two groups are placed on the same scale prior to testing for DIF. The Likelihood Ratio Test utilizes a set of DIF-free anchor items, that is akin to the matching subtest used to test for DIF when using SIBTEST. They are also similar in that they can all be used to test for either uniform or non-uniform DIF and none of them can be used to test for DBF or DTF.

MIMIC model

The Multiple Indicators Multiple Causes (MIMIC) model (Joreskog and Goldberger, 1975) is rooted in the methodology of structural equation modeling (SEM) and can be used to test for DIF. In fact, it can be shown that the parameters obtained from fitting the MIMIC model can be easily transformed to the item parameters that would be obtained if IRT were used to scale the data or test for DIF (MacIntosh and Hashim, 2003). In general, SEM is a multivariate statistical approach that combines factor analysis and multiple regression to model the relationships between latent constructs and observed variables. When using a MIMIC model to test for DIF, the latent variable in the model is the underlying construct measured by the instrument. All items on the measure are modeled as indicators of the latent trait. This is the measurement part of the model and, in fact, is a confirmatory factor analysis. The structural part of the model includes regressing the grouping variable on the latent trait. This models impact, or true group mean difference on the latent trait. Items are tested for DIF by regressing the responses to the suspect item on the grouping variable. If group membership is a statistically significant predictor of responses to that item, after controlling for group mean differences on the latent trait, then it is functioning differentially. Similar to the use of the Likelihood Ratio Tests that have already been discussed, full and reduced models can be compared to determine which model is the best fit to the data.

MIMIC models can be used to test for uniform and non-uniform DIF with both dichotomous and polytomous items. They also are the only other procedure that can be used to test for DBF or DTF (Finch, 2012). Some of the other advantages of using the method to test for DIF are that the groups are matched based on a latent variable which is more accurate than using total score which include measurement error. It is also possible to model multidimensional items and test for multiple grouping variables at once, as opposed to only using two groups, using this methodology.

Latent DIF analysis

All of the DIF detection procedures that have been discussed utilize a manifest grouping variable, such as race or gender. However, mixture IRT models can be used to uncover latent grouping variables inherent in the data that are not measured (Cho et al., 2016). These models integrate IRT with Latent Class Modeling (LCM), which is a statistical procedure used to identify underlying groups, or clusters, of respondents that respond similarly to items on a measure. The models assume that respondents are from two or more latent populations that impact how they respond to the items. These different populations could be based on different problem-solving strategies on an educational test or different personality types on a psychological measure. They are simply variables that have not been directly measured, although it is possible that the latent classes that are identified really represent a manifest grouping variable. This will occur when there is a large overlap between the latent class identified by the model with an existing manifest grouping variable. There are four reasons why a researcher might want to consider a latent, as opposed to manifest, grouping variable: group membership is missing or one has no idea what group is of interest; group membership is unobservable or unmeasured, such as the examples provided earlier; the manifest variable is not reliable because there is heterogeneity in one or more groups, and the manifest variable may not be a valid indicator of true group membership (De Boeck et al., 2011).

How are DIF analyses used operationally?

Most educational testing companies periodically update the content of their tests. This is done by field testing items, which means they are administered but not used in the calculation of a respondent's latent trait or reported score. Field test items are then analyzed, with respect to their difficulty, discrimination, and whether they are functioning differentially to determine which items should be operationalized. Items should be tested for both uniform and non-uniform DIF. If an item is found to have C level DIF then a panel of experts, in the content area being assessed by the test, is convened. For instance, if the item is related to language proficiency and differentially functioning against to males, then the panel must be diverse, with respect to gender, country of origin, and language spoken, and must not be familiar with the item, as would be the case if they were involved in the writing of the item. Panelists discuss potential reasons for the occurrence of DIF and are asked to come to a consensus decision about *why* the item might be functioning differentially. If they can come up with a reason as to why they all believe that the item is functioning differentially then the item is rewritten and field tested again, to determine if the changes have resulted in a DIF-free item. If they cannot come up with a reason as to why the item is functioning differentially then the item should not be used operationally unless there is a very

strong rationale. This decision should be based on the direction and magnitude of the DIF; whether or not there is cancellation, which occurs when some items function in favor of the reference group and other function in favor of the focal group, and the test specifications, to ensure that content validity will not be compromised.

Why are DF analyses an important part of test development and validation?

DF analyses are an important part of test development and validation because they help to ensure that scores obtained from educational tests and psychological measures are unbiased and reflect the same construct for all examinees or respondents. As previously described, DF analyses are routinely used by test developers and they are equally as important in the development of psychological assessments. It is equally vital that equivalent scores obtained from a measure of anxiety or an interest inventory reflect the same construct for all respondents, regardless of gender, race, or ethnicity, as it is for equivalent scores from an educational assessment to reflect the appropriate construct. This verification represents an important aspect of the test validation process, in terms of defining the construct or constructs that are being assessed. Thinking carefully about what test items may be measuring, in terms of the primary construct an assessment is designed to measure as well as any additional constructs that may be measured by test items, is critical. Moreover, how groups of examinees may differ in their underlying distribution on any supplemental dimensions measured by items can allow one to come up with testable hypotheses that can be assessed in a confirmatory manner to better understand the underlying structure of a psychological measure.

Practical considerations

With so many methods available for evaluating items for DIF, the reader may be wondering which one to use. As is often the case, there is no simple answer to this question. Specialized software is needed for almost all procedures discussed, which may be a reason for not selecting a particular method. Exceptions to this are the MH procedure and logistic regression, which can be conducted using any standard statistical software program. Some of the procedures can only be used to test for uniform DIF, which may be a reason for not using them. While all methods work better with larger sample sizes which provide better estimates, some, such as those based on IRT, require large samples for estimation. Some of the procedures can test for DBF and DTF which may be a reason for selecting them for use. In general, one should choose the method that is most aligned with their hypothesis, the nature of data (dichotomous vs. polytomous), and research questions posed. In addition, given that each method differs, with respect to how they ensure that the results are not biased by impact, it is advisable to use more than one method to cross-validate findings.

As stated previously, all methods work better with larger sample sizes. This naturally leads to the question of how many respondents are needed to conduct DIF analyses. Once again, the answer is complicated. One thing that is known to affect type I error is having comparable sample sizes in both the reference and focal group (Miller and Oshima, 1992). Balanced samples are typically better than unbalanced samples when it comes to type I error and power. With unbalanced designs, it has been suggested that multiple random samples be drawn from the larger group and multiple analyses be conducted with the smaller group to cross-validate results, with unbalanced sample sizes (Sireci and Rios, 2013). However, even for small focal group sample sizes, as may be the case in practice, adequate power is achieved if the sample size of the reference group is large, when Mantel-Haenszel is used to test for DIF (Paek and Guo, 2011). In general, it has been suggested that the focal group consist of at least 200 during the test assembly phase and at least 300 during the item analysis phase (Zwick, 2012). Non-parametric procedures require smaller sample sizes than do parametric procedures. Typically, when using IRT based methods it is recommended that each group have at least 500 respondents (Narayan and Swaminathan, 1996). For non-IRT based methods, it is recommended that each group have at least 200 respondents, if the measure has at least 20 questions, and at least 300 in each group if the measure has less than 20 items (Scott et al., 2009).

One thing that has already been touched on, is the importance of having a DIF-free subtest score, when using this as the matching variable. This implies that one should not simply utilize a respondent's total score on all remaining test items, that are not being tested for DIF, as the matching variable. It is possible that some of these items are actually functioning differentially, even if they are not suspected of doing so. Therefore, it is suggested that all items that are found to be functioning differentially, when conducting single-item analyses, be removed from the matching subtest score and the analyses be re-run.

References

- Abbott, M.L., 2006. ESL reading strategies: differences in Arabic and Mandarin speakers test performance. *Lang. Learn.* 56, 633–670.
- Abbott, M.L., 2007. A confirmatory approach to differential item functioning on an ESL reading assessment. *Lang. Test.* 24, 7–36.
- Ackerman, T., 1992. A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *J. Educ. Meas.* 29, 67–91.
- Angoff, W.H., 1993. Perspectives on differential item functioning methodology. In: Holland, P.W., Wainer, H. (Eds.), *Differential Item Functioning*. Lawrence Erlbaum Associates, Inc, pp. 3–23.
- Banks, K., 2006. A comprehensive framework for evaluating hypotheses about cultural bias in educational testing. *Appl. Meas. Educ.* 19, 115–132.

- Bolt, D., 2002. A Monte Carlo comparison of parametric and nonparametric polytomous DIF detection methods. *Appl. Meas. Educ.* 15, 113–141.
- Cho, S., Suh, Y., Lee, W., 2016. An NCME instructional module on latent DIF using mixture item response modules. *Educ. Meas.* 35, 48–61.
- De Boeck, P., Cho, S., Wilson, M., 2011. Explanatory secondary dimension modeling of latent differential item functioning: a mixture distribution conceptualization. *Appl. Psychol. Meas.* 35, 583–603.
- Doran, N.J., Holland, P.W., 1993. DIF detection and description: Mantel–Haenszel and Standardization. In: Holland, P.W., Wainer, H. (Eds.), *Differential Item Functioning*. Lawrence Erlbaum & Associates, Inc, pp. 35–66.
- Finch, H., Barton, K., Meyer, P., 2009. Differential item functioning analyses for accommodated versus non-accommodated students. *Educ. Assess.* 14, 38–56.
- Finch, W.H., 2012. The MIMIC model as a tool for differential bundle functioning detection. *Appl. Meas. Educ.* 36, 40–59.
- French, A.W., Miller, T.R., 1996. Logistic regression and its use in detecting differential item functioning in polytomous items. *J. Educ. Meas.* 33 (3), 315–332. <https://doi.org/10.1111/j.1745-3984.1996.tb00495.x>.
- Gierl, M.J., Bisanz, J., Bisanz, G.L., Boughton, K.A., Khaliq, S.N., 2001. Illustrating the utility of differential bundle functioning analyses to identify and interpret group differences on achievement tests. *Educ. Meas.* 20 (2), 26–36.
- Gierl, M.J., Bisanz, J., Bisanz, G.L., Boughton, K.A., 2003. Identifying content and cognitive skills that produce gender differences in mathematics: a demonstration of the multidimensionality-based DIF analysis paradigm. *J. Educ. Meas.* 40, 281–307.
- Gierl, M.J., 2005. Using dimensionality-based DIF analyses to identify and interpret constructs that elicit group differences. *Educ. Meas.* 24 (1), 3–14.
- Guo, H., Dorans, N., 2020. Using weighted sum scores to close the gap between DIF practice and theory. *J. Educ. Meas.* 57 (4), 484–510.
- Holland, P.W., Thayer, D.T., 1985. An Alternative Definition of the ETS Delta Scale of Item Difficulty. ETS Research Report No. RR 85-43. Educational Testing Service, Princeton, NJ.
- Holland, P.W., Thayer, D.T., 1988. Differential item functioning and the Mantel–Haenszel procedure. In: Wainer, H., Braun, H.I. (Eds.), *Test Validity*. Erlbaum, Hillsdale, NJ, pp. 129–145.
- Jodoin, M.G., Gierl, M.J., 2001. Evaluating Type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Appl. Meas. Educ.* 14, 329–349.
- Johanson, F., Alsmadi, A., 2002. Differential person functioning. *Educ. Psychol. Meas.* 62, 435–443.
- Joreskog, K., Goldberger, A., 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *J. Am. Stat. Assoc.* 70 (351), 631–639. <https://doi.org/10.2307/2285946>.
- Lord, F.M., 1980. Applications of Item Response Theory to Practical Testing Problems. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ.
- MacIntosh, R., Hashim, S., 2003. Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. *Appl. Psychol. Meas.* 27 (5), 372–379. <https://doi.org/10.1177/0146621603256021>.
- Mantel, N., Haenszel, W., 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* 22, 719–748.
- Masor, K.M., Clauser, B.E., Hambleton, R.K., 1994. Identification of nonuniform differential item functioning using a variation of the Mantel–Haenszel procedure. *Educ. Psychol. Meas.* 54, 284–291.
- Mendes-Barnett, S., Ericson, K., 2006. Examining sources of gender DIF in mathematics using a confirmatory multidimensional model approach. *Appl. Meas. Educ.* 19, 289–304.
- Miller, M.D., Oshima, T.C., 1992. Effect of sample size, number of biased items, and magnitude of bias on a two-stage item bias estimation method. *Appl. Psychol. Meas.* 16, 381–388.
- Miller, T.R., Spray, J.A., 1993. Logistic discriminant function analysis for DIF identification of polytomously scored items. *J. Educ. Meas.* 30, 107–122.
- Nandakumar, R., 1993. Simultaneous DIF amplification and cancellation: Shealy–Stout’s measure test for DIF. *J. Educ. Meas.* 30, 293–311.
- Narayanan, P., Swaminathan, H., 1996. Identification of items that show nonuniform DIF. *Appl. Psychol. Meas.* 20, 257–274.
- National Council of Teachers of Mathematics, 2000. Principles and Standards for School Mathematics. National Council of Teachers of Mathematics, Reston, VA.
- O’Leary, L.S., Smith, R.W., 2017. Detecting candidate pre-knowledge and comprised content using differential person and item functioning. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Quantitative Methods for Detecting Cheating on Tests*. Routledge, New York, pp. 200–216.
- Pae, I., Guo, H., 2011. Accuracy of DIF estimates and power in unbalanced designs using the Mantel–Haenszel DIF detection procedure. *Appl. Psychol. Meas.* 35 (7), 518–535. <https://doi.org/10.1177/0146621611420559>.
- Raju, N.S., 1988. The area between two item characteristic curves. *Psychometrika* 53, 495–502.
- Raju, N.S., 1990. Determining the significance of estimated signed and unsigned areas between two item response functions. *Appl. Psychol. Meas.* 14, 197–207.
- Rosenbaum, P.R., 1988. Item bundles. *Psychometrika* 53, 349–359.
- Roussos, L.A., Stout, W.F., 1996a. A multidimensionality-based DIF analysis paradigm. *Appl. Psychol. Meas.* 20, 353–371.
- Roussos, L.A., Stout, W.F., 1996b. Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel–Haenszel Type I error performance. *J. Educ. Meas.* 33 (2), 215–230. <https://doi.org/10.1111/j.1745-3984.1996.tb00490.x>.
- Ryan, K.E., Chiu, S., 2001. An examination of item context effects, DIF, and gender DIF. *Appl. Meas. Educ.* 14, 73–90.
- Ryan, K.E., Fan, M., 1996. Examining gender DIF on a multiple-choice test of mathematics: a confirmatory approach. *Educ. Meas.* 15 (4), 15–20.
- Scott, N.W., Fayers, P.M., Aaronson, N.K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M.A., Sprangers, M.A., EORTC Quality of Life Group; Quality of Life Cross-Cultural Meta-analysis Group, 2009. A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J. Clin. Epidemiol.* 62 (3), 288–295. <https://doi.org/10.1016/j.jclinepi.2008.06.003>.
- Shealy, R., Stout, W., 1993. A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika* 58, 159–194.
- Sinharay, S., 2016. Asymptotically correct standardization of person-fit statistics beyond dichotomous items. *Psychometrika* 81, 992–1013. <https://doi.org/10.1007/s11336-015-9465-x>.
- Sireci, S.G., Rios, J.A., 2013. Decisions that make a difference in detecting differential item functioning. *Educ. Res. Eval.* 19, 170187. <https://doi.org/10.1080/13803611.2013.767621>.
- Smith, R.W., Davis-Becker, S., 2011. Detecting Suspect Candidates: An Application of Differential Person Functioning Analysis. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.
- Swaminathan, H., Rogers, H.J., 1990. Detecting differential item functioning using logistic regression procedures. *J. Educ. Meas.* 27, 361–370.
- Thissen, D., Steinberg, L., Gerrard, M., 1986. Beyond group mean differences: the concept of item bias. *Psychol. Bull.* 99, 118–128.
- Thissen, D., Steinberg, L., Wainer, H., 1988. Use of item response theory in the study of group differences in trace lines. In: Wainer, H., Braun, H.I. (Eds.), *Test Validity*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, pp. 147–169.
- Thissen, D., Steinberg, L., Wainer, H., 1993. Detection of differential item functioning using the parameters of item response models. In: Holland, P.W., Wainer, H. (Eds.), *Differential Item Functioning*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, pp. 67–113.
- Walker, C.M., Beretvas, S.N., 2001. An empirical investigation demonstrating the multidimensional DIF paradigm: a cognitive explanation for DIF. *J. Educ. Meas.* 38, 147–163.
- Walker, C.M., Zhang, B., Surber, J., 2008. Using a multidimensional differential item functioning framework to determine if reading ability affects student performance in mathematics. *Appl. Meas. Educ.* 21 (2), 162–181. <https://doi.org/10.1080/08957340801926201>.

- Walker, C.M., Zhang, B., Banks, K., Cappaert, K., 2012. Establishing effect size guidelines for interpreting the results of differential bundle functioning analyses using SIBTEST. *Educ. Psychol. Meas.* 72, 415–434.
- Wollack, J.A., 1997. A nominal response model approach for detecting answer copying. *Appl. Psychol. Meas.* 21 (4), 307–320. <https://doi.org/10.1177/01466216970214002>.
- Zumbo, B.D., 1999. A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores. Directorate of Human Resources Research and Evaluation, Department of National Defense, Ottawa, ON.
- Zwick, R., Ercikan, K., 1989. Analysis of differential item functioning in the NAEP history assessment. *J. Educ. Meas.* 26 (1), 55–66. <https://doi.org/10.1111/j.1745-3984.1989.tb00318.x>.
- Zwick, R., 2012. A Review of DIF. <https://www.ets.org/Media/Research/pdf/RR-12-08.pdf>.

Model-data fit evaluation: item fit and model selection

Alberto Maydeu-Olivares^{a, b}, Dexin Shi^b, Goran Pavlov^b, and Raul Ferraz^b, ^a University of South Carolina, Columbia, SC, United States; and ^b University of Barcelona, Barcelona, Spain

© 2023 Elsevier Ltd. All rights reserved.

Introduction	260
Goodness of fit vs. model selection	260
Model fit in factor analysis and structural equation modeling	261
Overall goodness of fit statistics	262
Identifying the source of misfit and model modification	263
Goodness of fit indices	263
Why should we care about model fit? Fit predicts parameter bias	264
Model fit with discrete observed data	266
Obtaining accurate <i>p</i> -values by pooling cells a priori	266
Overall goodness of fit statistics and goodness of fit indices: using univariate and bivariate residuals, item means, cross-products and correlations under the model	267
Piecewise assessment of fit and model modification	268
Model selection indices	268
Numerical examples	268
Factor analysis	268
Item response theory	269
Concluding remarks	270
References	271

Glossary

Goodness of fit index A numerical summary of the discrepancy between the observed data values and the values expected under a statistical model

Goodness of fit statistics A goodness of fit index with known sampling distribution that may be used in statistical hypothesis testing

Absolute goodness of fit The discrepancy between a statistical model and the data generating mechanism

Relative goodness of fit The discrepancy between two statistical models

Introduction

The **goodness of fit** (GOF) of a statistical model refers to the extent to which a statistical model captures the data generating mechanism. When more than one model is under consideration **model selection** is necessary. Thus, goodness of fit refers to assessing the discrepancy between the model and the data, whereas model selection refers to assessing the discrepancy between two models. GOF is often referred to as model-data fit, or absolute fit. Model selection is often referred to as relative model fit.

Two different perspectives exist regarding goodness of fit and model selection. One perspective views modeling as describing the data (Cudeck and Henly, 1991; MacCallum, 2003). From this perspective, all models are simply approximations to real life phenomena (“all models are wrong”) and the researcher’s aim is to choose the model that provides the “best” approximation to the data generating process. From this perspective, it makes sense to focus on model selection. An alternative perspective views modeling as estimating structural or “causal” relations even when data is observational (Pearl, 2009; Bollen and Pearl, 2013). From this perspective, inferences are to be as precise and valid as possible, and the focus is on model fit.

Goodness of fit vs. model selection

The goodness of fit vs. model selection distinction, and how they are intertwined, can be illustrated by considering a simple regression model $y = \beta_0 + \beta_1 x + \varepsilon$ with homoscedastic errors. The model’s R^2 is a model selection statistic, which can be used for instance to choose between (a) the linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$, and (b) a mean-only model $y = \beta_0 + \varepsilon$. We illustrate this in Fig. 1, in which R^2 can be used to successfully select between both models in all instances. R^2 is **not** a goodness of fit statistic. We also illustrate this in Fig. 1, where interpreting R^2 as a goodness of fit statistic would lead us to conclude that the model fits well in (b) but

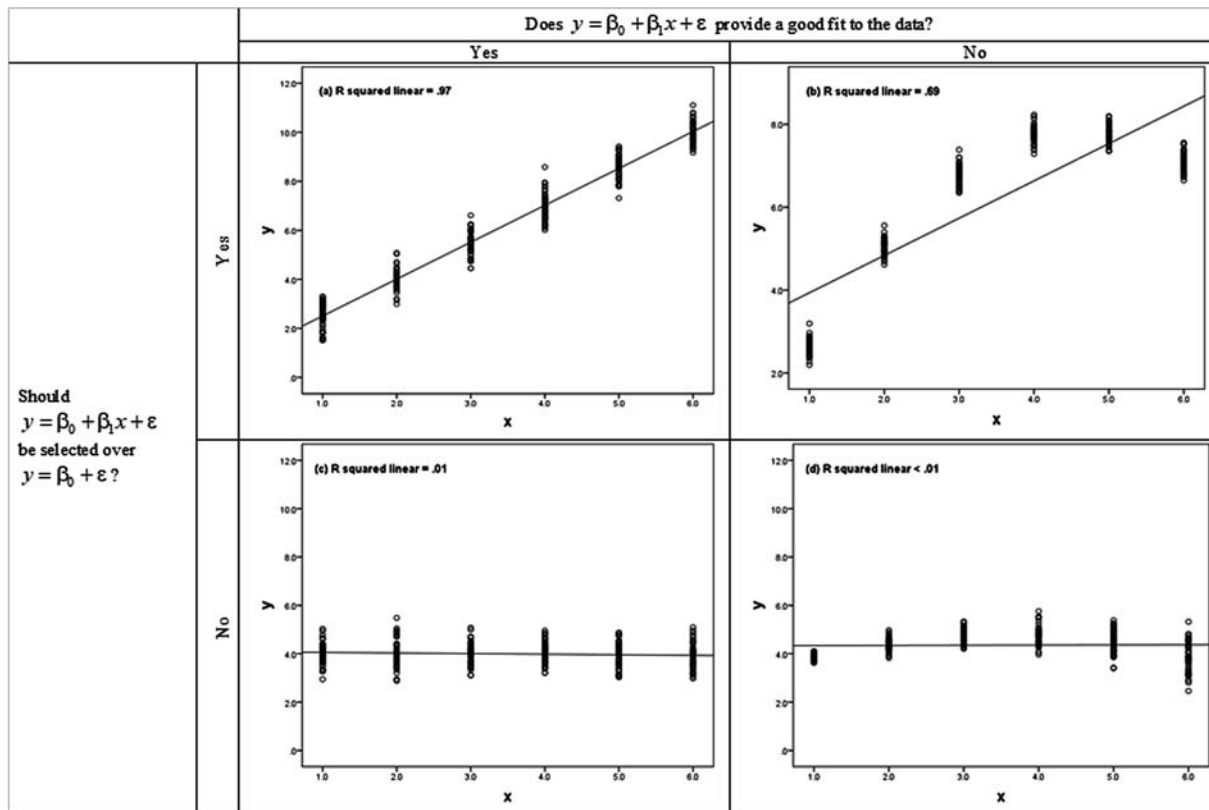


Fig. 1 Goodness of fit and model selection for a simple linear regression model.

not in (c), while it fits excellently in (c) but not in (b). Rather, goodness of fit in linear regression is usually assessed using plots, although if repeated observations are available for each value of x , a goodness of fit test may be performed. This test statistic is usually referred to in the regression literature as lack of fit test (Weisberg, 2005). Only after the fit of the regression model has been established, R^2 , a model selection statistic, is applied and interpreted.

In applications, we recommend that both goodness of fit assessment and model selection be performed. When fitting several models to a dataset, after determining that more than one model fits well, model selection must be performed. Alternatively, after selecting the best fitting model, some evidence of the fit of the model must be obtained. When model selection is performed, but goodness of fit assessment is not, the resulting model may fit the data rather poorly, and lead to misleading substantive conclusions (Maydeu-Olivares, 2017b). When goodness of fit assessment is performed but no alternative model is considered (i.e., no model selection), the resulting model may overfit the data and not cross-validate well.

In some applications different conclusions may be reached if a goodness of fit or a model selection perspective is adopted. This is illustrated in Fig. 1D. If a pure model selection perspective is adopted, correctly concluding that the mean-only model is to be selected over the linear model is all that matters. This is consistent with considering models as approximations. From a goodness of fit perspective, because we observe some non-linearity and heteroscedasticity, it may be concluded that neither the linear regression model nor the mean only model capture the data generating mechanism, and that a better model must be sought.

Model fit in factor analysis and structural equation modeling

Moving from assessing the fit of a single linear equation to assessing the fit of a set of linear equations we should apply the same principles. Plots are to be used to assess model fit for each model equation. When latent variables (i.e., common factors) are involved, a complication arises as latent variable values are not observed, and they must be estimated, i.e., factor scores must be used (Bollen and Arminger, 1991).

When a system of equations is fit, residual means and covariances (or correlations) provide an additional procedure to assess the fit of the model. For instance, a test of the suitability of the restrictions imposed on the covariance among two variables can be obtained using a z -statistic for the residual covariance (Maydeu-Olivares and Shi, 2017),

$$z_{ij} = \frac{s_{ij} - \hat{\sigma}_{ij}}{SE(s_{ij} - \hat{\sigma}_{ij})}, \quad (1)$$

where s_{ij} denote the sample covariance between variables i and j , $\hat{\sigma}_{ij}$ denotes the covariance predicted by the model, and $SE(\bullet)$ denotes standard error.

Often, some residual means and covariances are zero by construction and they cannot be tested. For instance, in a multivariate regression model all residual means and residual covariances are zero by construction and, as a result, they do not provide information about the fit of the model. As another example, researchers often use structural equation models (SEM) that do not impose restrictions on the intercepts of the model equations (intercepts are freely estimated). In this case, residual means are zero by construction, they are known to be zero before any data is gathered, and therefore z statistics cannot be computed.

To control for multiple testing across all z statistics, overall goodness of fit tests are used.

Overall goodness of fit statistics

Historically, the likelihood ratio test statistic has been the most widely used overall goodness of fit test statistic. When no restrictions are imposed on the means of the observed variables (aka covariance structure models), the likelihood ratio test can be written as

$$T_{LR} = N \{ \ln |\hat{\Sigma}| - \ln |S| + \text{tr}(\hat{S}\hat{\Sigma}^{-1}) - n \}, \quad (2)$$

where S denotes the sample covariance matrix among the n observed variables, $\hat{\Sigma}$ denotes the covariance matrix expected under the model evaluated at the values of the estimated model parameters, and N denotes sample size. If the observed data follows a multivariate normal distribution, the distribution of the likelihood ratio test statistic can be approximated by a chi-square distribution in large samples. Because of its popularity, T_{LR} is commonly referred to as *the* chi-square test in the factor analysis and structural equations modeling literature.

However, T_{LR} should not be used in small samples, nor when data departs from normality. Recent research has also shown that p -values obtained using this test are too small when all assumptions are satisfied but the number of variables exceeds 20 or 30, even in very large samples (Moshagen, 2012). As a result, well-fitting models will be rejected when the number of variables is large.

T_{LR} is not the only overall test statistic for structural equation modeling that follows a chi-square distribution. For instance, Browne (1974, 1982, 1984) proposed a test statistic that for covariance structure models can be written as

$$T_B = N \hat{\mathbf{e}}' \hat{\mathbf{C}} \hat{\mathbf{e}}, \quad \mathbf{C} = \mathbf{\Gamma}^{-1} - \mathbf{\Gamma}^{-1} \Delta (\Delta' \mathbf{\Gamma}^{-1} \Delta)^{-1} \Delta' \mathbf{\Gamma}^{-1}. \quad (3)$$

In (3), $\hat{\mathbf{e}} = \mathbf{s} - \hat{\boldsymbol{\sigma}}$ denote the residual covariances transformed from a matrix layout to a single vector, $\mathbf{\Gamma}$ denotes the covariance matrix of the sample covariances in large samples, and Δ is a matrix of derivatives of the population covariances with respect to the model parameters. There are two variants of this statistic, depending on how $\mathbf{\Gamma}$ is computed: under normality assumptions, $T_B^{(NT)}$, or under the very general asymptotically distribution free (ADF) assumptions, $T_B^{(ADF)}$ (Browne, 1982). In both cases, p -values can be obtained in large samples using a chi-square distribution. We note that under normality assumptions and when maximum likelihood is used to estimate the model parameters, (3) simplifies to $T_B^{(NT)} = N \hat{\mathbf{e}}' \hat{\mathbf{\Gamma}}_{NT}^{-1} \hat{\mathbf{e}}$, where $\mathbf{\Gamma}_{NT}$ denotes $\mathbf{\Gamma}$ computed under normality assumptions.

$T_B^{(NT)}$ is a direct competitor to T_{LR} (both assume normality), whereas $T_B^{(ADF)}$ provides a goodness of fit test robust to non-normality. Unfortunately, very large sample sizes are needed for the p -values of $T_B^{(ADF)}$ to be accurate, and sample size requirements increase as the number of observed variables increase (Curran et al., 1996). As a result, $T_B^{(ADF)}$ should only be used in small models and very large samples.

An alternative way to obtain p -values when data is non-normal was proposed by Satorra and Bentler (1994): adjust a test statistic for normally distributed data by its asymptotic mean (or by its asymptotic mean and variance) so that it agrees in mean (or in mean and variance) with a chi-square distribution. Mean and variance corrected statistic provide more accurate p -values than mean corrected statistics (Maydeu-Olivares, 2017a), and even outperform uncorrected statistics when their assumptions are met (i.e., under normality).

The current standard to assess the overall goodness of fit of a structural equation model is the likelihood ratio test statistic T_{LR} , suitable for normally distributed data, and its mean (or mean and variance) adjustment suitable for non-normal data, $\bar{T}_{LR}$ and $\bar{\bar{T}}_{LR}$, respectively. However, Browne's test statistic computed under normality, $T_B^{(NT)}$ can also be adjusted by its mean (or mean and variance), leading to $\bar{T}_B^{(NT)}$ and $\bar{\bar{T}}_B^{(NT)}$. Hayakawa (2019) has shown that $T_B^{(NT)}$ provides more accurate p -values than the likelihood ratio test statistic, T_{LR} . In particular, $T_B^{(NT)}$ appears to overcome the overrejection problem of T_{LR} when the number of observed variables is large relative to the sample size T_{LR} . Hayakawa's results strongly suggest that Browne's statistic (3), not the likelihood ratio statistic (2), possibly with a Satorra-Bentler adjustment to account for non-normality, should be the statistic of choice to assess the overall goodness of fit test in for SEM and other multivariate latent-variable models.

Identifying the source of misfit and model modification

Overall goodness of fit statistics are summary measures of the residual means and covariances. For instance, Browne's statistic (3) is a weighted average of residual covariances. As a result, before interpreting an overall test of model fit it is necessary to examine whether there is any outlying residual means/covariances that unduly inflates the value of the statistic. Also, residuals means and covariances should be inspected to identify parts of the data that are not well fit by the model. In so doing, it is necessary to consider that multiple tests are performed. The simplest approach to deal with multiple testing is to apply a Bonferroni correction to the critical value used in the z -statistics (1), but more accurate procedures are preferable (Benjamini and Hochberg, 1995; Thissen et al., 2002). In poorly fitting models, inspection of residual means and covariances may suggest the source of model misfit. However, often modification indices are more useful to this aim.

Models impose restrictions, implicitly or explicitly, on their parameter values. Examples in factor analysis are setting a factor loading equal to zero, or the correlation between the errors of two equations equal to zero. Modification indices may be used to test those restrictions one at a time. These are *score-test* statistics (Buse, 1982) with one degree of freedom that contrast the fitted model (where this parameter is set to a fixed value) against a model in which this parameter is estimated. Together, the modification index (i.e., the value of the test statistic), the expected parameter change, and the power of the test enable researchers to investigate whether there is a small misspecification for which the test statistic is very sensitive, or a large misspecification that the test statistic is not sensitive enough to identify (Saris et al., 2009).

Goodness of fit indices

As the number of variables being modeled increases, so do degrees of freedom (intuitively, ways that the model can be wrong), and it becomes increasingly difficult to find a well-fitting model (Maydeu-Olivares, 2017b). In these situations, researchers resort to the use of goodness of indices, coupled with cutoff criteria that have been offered in the literature (Hu and Bentler, 1999) and use them as test statistics: models are retained if they meet the cutoff criteria, and rejected otherwise (Barrett, 2007). The most widely used goodness of fit indices are probably the Root Mean Squared Residual (RMSEA: Browne and Cudeck, 1993), the Standardized Root Mean Residual (SRMR: Jöreskog and Sörbom, 1988), the Tucker-Lewis index (TLI: Tucker and Lewis, 1973), and the Comparative Fit Index (CFI: Bentler, 1990).

Letting t be the number of residual covariances in a model, the sample SRMR is

$$\widehat{SRMR} = \sqrt{\frac{1}{t} \hat{\mathbf{e}}' \hat{\mathbf{G}}^{-1} \hat{\mathbf{e}}} = \sqrt{\frac{\sum_{i \leq j} (s_{ij} - \hat{\sigma}_{ij})^2}{s_{ii} s_{jj}}}, \quad (4)$$

where $\hat{\mathbf{G}}$ is a diagonal matrix with elements $s_{ii} s_{jj}$. The SRMR can be approximately interpreted as the average residual correlation. In turn, the sample RMSEA is

$$\widehat{RMSEA} = \sqrt{\max\left(\frac{T - df}{N \times df}, 0\right)}. \quad (5)$$

where df refers to the degrees of freedom of the model, and T denotes the chi-square statistic used to assess whether the model fits exactly in the population (generally the likelihood ratio test statistic). When Satorra-Bentler corrected statistic is used to account for non-normality, (5) needs to be adjusted accordingly (Savalei, 2018).

The SRMR and RMSEA assess the degree of absolute fit of a model. The closer the values of the SRMR and RMSEA to zero, the better the fit of the model. The RMSEA includes a degrees of freedom adjustment, to penalize models with too many parameters, whereas the SRMR does not. In turn, the sample TLI and CFI are relative model fit indices. They involve two models, M_0 and M_1 , where the former is more restrictive. The sample TLI and CFI are

$$\widehat{TLI} = \frac{\frac{T_0}{df_0} - \frac{T_1}{df_1}}{\frac{T_0}{df_0} - 1}, \quad (6)$$

$$\widehat{CFI} = \frac{(T_0 - df_0) - (T_1 - df_1)}{T_0 - df_0}. \quad (7)$$

where T_0 and T_1 are the chi-square statistics, of each model; and df_0 and df_1 are their degrees of freedom. Both the CFI and TLI are normed to lie between 0 and 1. Larger values indicate better fit. Almost invariably the CFI and TLI are used to compare the fitted model against a model that assumes that variables are uncorrelated. When used in this fashion, TLI and CFI values are very large.

Consider using goodness of fit indices to determine whether to retain a one factor model when the data generating model involves two factors. A one factor model is close to the two factor model when the inter-factor correlation is very high (say 0.9), but it provides a poor approximation when the correlation is low (say 0.2). As a result, we would like the goodness of fit indices to suggest retaining a one factor model in the first case, but not in the second case, regardless of sample size, model size (number of variables), and factor loading used to generate the data. We illustrate in Fig. 2, taken from Ximénez et al. (2022), that this is not the case. In this figure we have plotted the population value and average sample behavior of the RMSEA and CFI across a number of

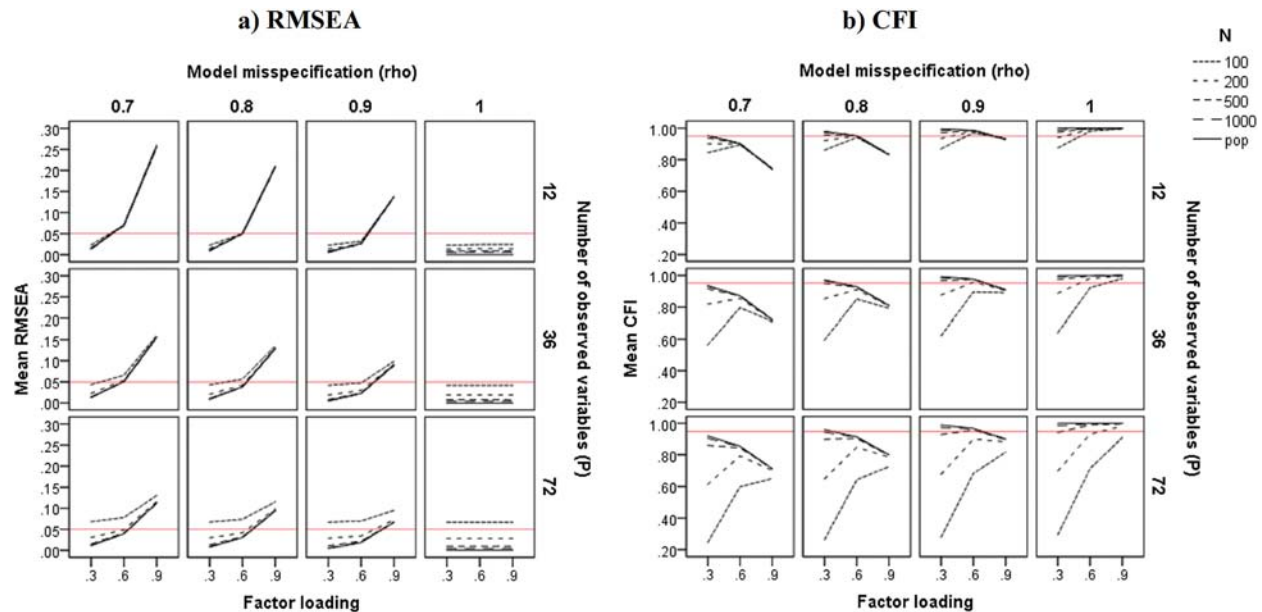


Fig. 2 Population and average sample CFI and RMSEA when fitting a one factor model to two factor model data

Note: ρ denotes the correlation between the factors, N is sample size. Data were generated so that all items had a common standardized factor loading (0.30, 0.60, and 0.90), 1000 replications were obtained per condition, pop is the population values for each index, and the red solid line corresponds to [Hu and Bentler's \(1999\)](#) cutoff for each index. Models are to be retained if estimated CFI is above the suggested cutoff, and if estimated RMSEA is below the suggested cutoff.

conditions. We have also included in this figure the cut-off values suggested by [Hu and Bentler \(1999\)](#). As we see in this figure, these cutoffs would lead to reject that a one factor model is a close approximation to a two-factor model with correlation 0.9 when sample size is small, the number of observed variables is large, and factor loadings are strong. Two findings are particularly striking in [Fig. 2](#): (a) the effect of sample size on the average behavior of the CFI is considerably larger than on the average behavior of the RMSEA, and (b) with large number of variables (e.g., $n = 72$) and small samples ($N \leq 200$), the CFI would lead us to conclude, on average, that a one factor model does not fit closely even when it fits exactly (i.e., $\rho = 1$). These findings are due to the fact that the RMSEA estimate (5) was formulated to be an unbiased estimator of the RMSEA population parameter, whereas the CFI estimate (7) is not an unbiased estimate of the CFI parameter; it is simply the sample counterpart of the population parameter. When an unbiased estimator of a goodness of fit parameter is used, cutoff values can be determined using the parameter, as sample size plays no role in the average behavior of the index. This can be seen in [Fig. 3](#), where we display the behavior of the sample SRMR (4), along with its unbiased counterpart ([Maydeu-Olivares, 2017b](#)).

Cutoff criteria for goodness of fit indices (SRMR, RMSEA, CFI, etc.) to distinguish between close and non-close fitting models are (a) to be put forth at the population level, (b) an unbiased estimator is to be used so that the average behavior of the estimate matches the behavior of the parameter, and (c) a confidence interval for the population value is needed to take into account sampling variability. Also, population cutoff criteria probably need to be adjusted to take into account extraneous factors, so that the population index just reflects misfit. For instance, letting $\bar{R}^2$ be the average R^2 of the observed variables, a cutoff of $\text{SRMR} / \bar{R}^2$ perfectly separates between close and non-close fitting models in [Figs. 2 and 3](#), regardless of model size, factor loading, or sample size ([Shi et al., 2018](#)).

Why should we care about model fit? Fit predicts parameter bias

A great deal of research has been devoted to model fit. Yet, very little research has been devoted to the implications of poor model fit, i.e., why model fit matters (but see [Bandalos, 2002](#)). [Steenkamp and Maydeu-Olivares \(2022\)](#) recently showed how the degree of misfit (i.e., goodness of fit indices values) predicts parameter bias, and hence the accuracy of substantive inferences. More specifically, they showed how cross-loadings incorrectly set to zero in confirmatory factor analysis models are expressed through the correlations between the factors, leading to biased factor correlations and to biased structural (regression) parameter estimates. Most importantly, they showed how the magnitude of bias can be predicted from goodness of fit indices. We show population results graphically in [Fig. 4](#) for two goodness of fit indices: the RMSEA (an absolute fit index), and the CFI (a relative fit index). For the RMSEA, we see in [Fig. 4](#) that the number of indicators per factor, the magnitude of the substantive loading, and the RMSEA, along with their interactions predict the natural log of the difference between the true and estimated inter-factor correlation. Across the

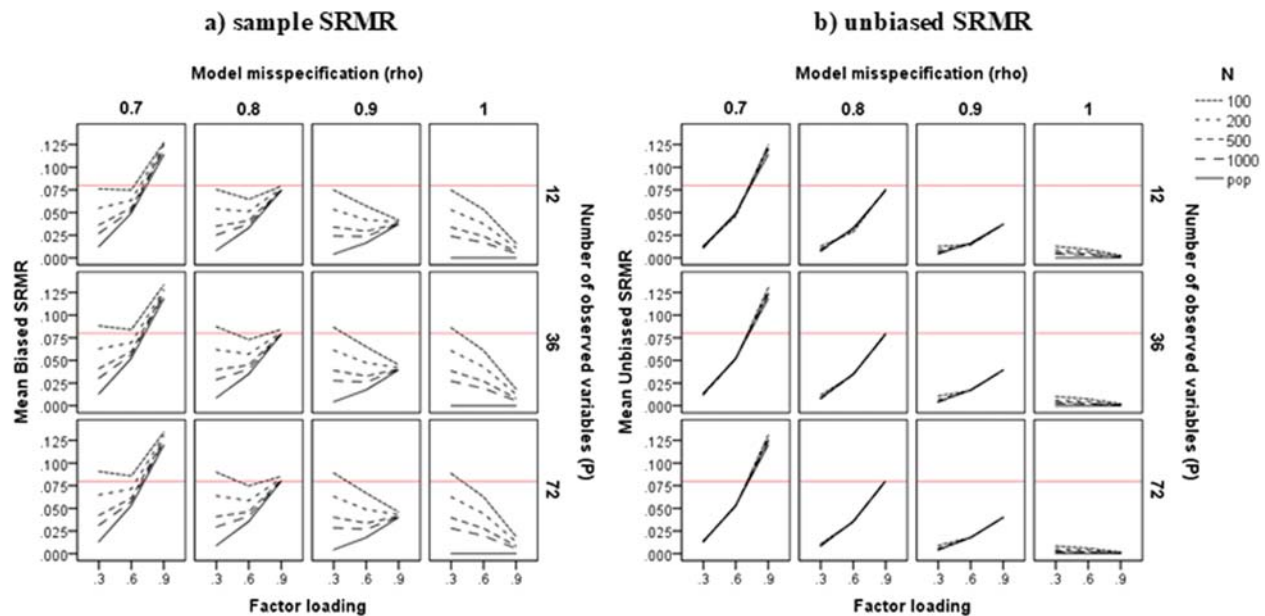


Fig. 3 Population and average sample SRMR when fitting a one factor model to two factor model data

Note: rho denotes the correlation between the factors, N is sample size. Data were generated so that all items had a common standardized factor loading (0.30, 0.60, and 0.90), 1000 replications were obtained per condition, pop is the population value -common for (A) and B), and the red solid line corresponds to Hu and Bentler's (1999) cutoff. Models are to be retained if estimated SRMR is below the suggested cutoff.

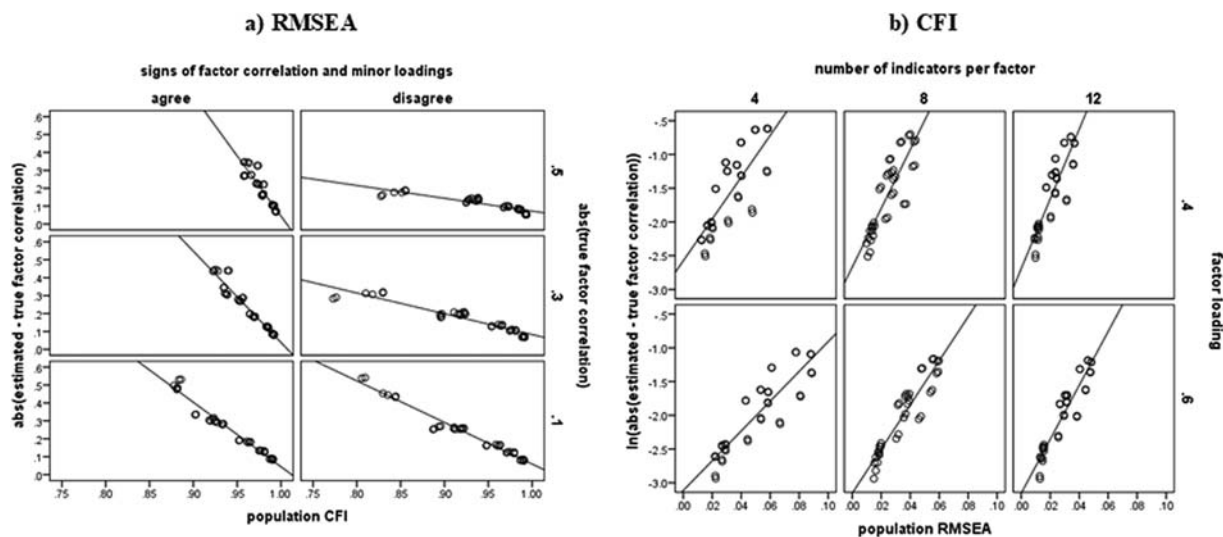


Fig. 4 Relationship between population model fit and parameter bias in confirmatory factor analysis models when the data generating model involves cross-loadings for 25% of the items and the fitted model has no cross-loadings

Note: The data generating models is a two-factor model. 252 conditions were obtained by crossing true factor correlation (−0.5, −0.3, −0.1, 0.1, 0.3, 0.5), number of indicators per factor (4, 8, 12), magnitude of the substantive factor loadings (0.4, 0.6), and magnitude of the cross-loadings (−0.3, −0.2, −0.1, 0, 0.1, 0.2, 0.3). Error variances were set so that the resulting covariance matrix has a unit diagonal. The fitted model was a two-factor analysis model without cross-loadings. For each index, bias is plotted against the main drivers of the behavior of the index.

panels of Fig. 4A, R^2 ranges between 70% and 88%. For the CFI, we see in Fig. 4 that whether the sign of the cross-loadings agrees with the sign of the interfactor correlation, the absolute value of the interfactor correlation, and the CFI, along with their interactions predict the difference between the true and estimated inter-factor correlation. Across the panels of Fig. 4B, the percentage of variance of parameter bias explained ranges between 81% and 97%

These results provide convincing evidence that the size of model misfit matters.

Model fit with discrete observed data

When the observed variables are discrete (as in responses to test items) and they are modeled as discrete, the observed responses can be gathered in a contingency table with C cells. When all n items consist of the same number of categories, K , $C = K^n$. The two standard goodness of fit statistics for discrete data are Pearson's statistic, $X^2 = N \sum_{c=1}^C (p_c - \hat{\pi}_c)^2 / \hat{\pi}_c$, and the likelihood ratio statistic, $G^2 = 2N \sum_{c=1}^C p_c \ln(p_c / \hat{\pi}_c)$, where $\hat{\pi}_c$ denotes the probability of cell c under the model, and p_c denotes the sample proportion. When maximum likelihood estimation is used, asymptotic p -values for both statistics can be obtained using a chi-square distribution. However, these asymptotic p -values are only correct when all expected frequencies $N\hat{\pi}_c$ are large (>5 is the usual rule of thumb). A practical way to evaluate whether the asymptotic p -values for X^2 and G^2 are valid is to compare them. If the p -values are similar, then both are likely to be correct. If they are very different, it is most likely that both p -values are incorrect. Unfortunately, as the number of cells in the table (i.e., possible response patterns) increases, the expected frequencies must become small (as the sum of all C probabilities must be equal to 1). As a result, when the number of item categories is large (say >4) the asymptotic p -values almost invariably become inaccurate as soon as $n > 5$.

Obtaining accurate p -values by pooling cells a priori

Accurate p -values for test statistics can be obtained by pooling cells of the contingency table so that larger expected frequencies are obtained. Pooling must be performed a priori, before the data is observed, so that the distribution of test statistics can be approximated by a reference distribution. The margins of the contingency table provide a straightforward way to pool cells a priori. We illustrate the relationship between cell probabilities and marginal probabilities using the following 2×3 contingency table:

	$X_2 = 0$	$X_2 = 1$	$X_2 = 2$
$X_1 = 0$	π_{00}	π_{01}	π_{02}
$X_1 = 1$	π_{11}	π_{11}	π_{12}

This table can be characterized using the 6 cell probabilities $\pi = (\pi_{00}, \dots, \pi_{12})'$ that add up to one (so that there are only 5 mathematically independent cell probabilities). Alternatively, this table can be characterized using the 3 univariate probabilities, $\pi_1 = (\pi_1^{(1)}, \pi_2^{(1)}, \pi_2^{(2)})$,

$$\pi_i^{(k)} = \Pr(X_i = k), \quad (8)$$

and 2 bivariate probabilities, $\pi_2 = (\pi_{12}^{(1)}, \pi_{12}^{(2)})$,

$$\pi_{ij}^{(k,l)} = \Pr(X_i = k, X_j = l), \quad (9)$$

Both characterizations (5 cell probabilities, or 5 marginal probabilities) are equivalent: Given $\pi = (\pi_{00}, \dots, \pi_{12})'$ we can determine $\pi = (\pi_1^{(1)}, \pi_2^{(1)}, \pi_2^{(2)}, \pi_{12}^{(1)}, \pi_{12}^{(2)})$, and vice versa. As a result, we can write Pearson's X^2 as a quadratic form of residuals using one characterization or the other (Maydeu-Olivares and Joe, 2005, 2006).

The equivalence extends to contingency tables of any dimension (i.e., regardless of K and n). For instance, the responses to n binary items can be characterized by the $2^n - 1$ cell probabilities or by the $\binom{n}{1} = n$ univariate probabilities $\pi_i^{(k)}$, plus $\binom{n}{2} = n(n-1)/2$ bivariate probabilities $\pi_{ij}^{(k,l)}$, plus $\binom{n}{3}$ trivariate probabilities, ..., up to $\binom{n}{n}$ n -way probabilities, and Pearson's X^2 can be written as a quadratic form in residual proportions up to n -way order.

Since data sparseness leads to a poor approximation of the sampling distribution of test statistics by their reference asymptotic distribution, Maydeu-Olivares and Joe (2005, 2006) proposed for goodness of fit testing to use only residuals in proportions up to the highest level that the model is identified. This amounts to pooling cells a priori so that the resulting statistic can be well approximated by its reference asymptotic distribution in small samples.

Since most models are identified (they can be estimated) using only univariate and bivariate probabilities (as described in the 2×3 table above), they suggested to use only univariate and bivariate residual proportions (the difference between the observed and expected probabilities). This yields their test statistic M_2 which has become the gold standard for goodness of fit testing when the observed variables are discrete. Statistics such as M_2 are referred to as limited information goodness of fit statistics because they disregard information contained in the higher order marginals of the table.

Overall goodness of fit statistics and goodness of fit indices: using univariate and bivariate residuals, item means, cross-products and correlations under the model

M_2 is a member of a larger family of statistics of the form (Joe and Maydeu-Olivares, 2010)

$$M_K = N \hat{\mathbf{e}}_K' \hat{\mathbf{C}}_K \hat{\mathbf{e}}_K, \quad \mathbf{C}_K = \mathbf{\Gamma}_K^{-1} - \mathbf{\Gamma}_K^{-1} \Delta_K (\Delta_K' \mathbf{\Gamma}_K^{-1} \Delta_K)^{-1} \Delta_K' \mathbf{\Gamma}_K^{-1}, \quad (10)$$

In (10), $\hat{\mathbf{e}}_K = \mathbf{k} - \hat{\mathbf{k}}$ denote the residuals: $\mathbf{k}$ denote some sample summary statistics with asymptotic covariance matrix $\mathbf{\Gamma}_K$, $\hat{\mathbf{k}}$ denotes the expected summary statistics (which depend on a smaller set of parameters $\boldsymbol{\theta}$) evaluated at the values of the estimated parameters, and Δ_K denotes the matrix of derivatives of κ with respect to $\boldsymbol{\theta}$. Joe and Maydeu-Olivares (2010) showed that the more highly summarized the data are, the more accurate the empirical Type I errors, and the higher the power of the test.

More specifically, M_2 is of the form (10) where the summary statistics κ contain all the univariate and bivariate probabilities (8) and (9). When all items consist of the same number of categories, K , there are $n(K-1)$ univariate probabilities $\pi_i^{(k)}$, $\frac{n(n-1)}{2}(K-1)^2$ bivariate probabilities $\pi_{ij}^{(k)(l)}$, and M_2 is asymptotically distributed as a chi-square with degrees of freedom equal to $n(K-1) + \frac{n(n-1)}{2}(K-1)^2$ minus the number of estimated parameters.

When n and K are large, very large matrices need to be stored in memory. Also, for large K , bivariate tables may be sparse, leading to slightly inaccurate p -values of M_2 (Cai and Hansen, 2013). When data is ordinal, $K \geq 3$, we can summarize the data further, which would lead to more accurate p -values and higher power. Returning to the 2×3 table of the previous section, if X_2 , which consists of 3 categories, is ordinal, we can compute the item means and cross-products expected under the fitted model as follows

$$\mu_1 = \pi_1^{(1)} \quad \mu_2 = 1\pi_2^{(1)} + 2\pi_2^{(2)} \quad (11)$$

$$\mu_{12} = 1 \times 1 \times \pi_1^{(1)(1)} + 1 \times 2 \times \pi_1^{(1)(2)} \quad (12)$$

Then, a quadratic form (10) can be constructed using the n item means μ_i and $n(n-1)/2$ cross-products μ_{ij} . We refer to this statistic for ordinal data as M_{ord} (Cai and Hansen, 2013; Maydeu-Olivares, 2013). If the data are nominal is not meaningful to compute this statistic (but see Starr et al., 2021). When all items are binary $M_{ord} = M_2$ and no further summarization of the univariate and bivariate tables is obtained.

M_{ord} cannot be used if the number of categories is large and the number of items is small due to lack of degrees of freedom for testing. For instance, for a unidimensional logistic graded-response model (e.g., Samejima, 1969), the number of items must be larger than the number of categories plus two (i.e., $n \geq K + 2$) for the degrees of freedom M_{ord} to be positive. Note that the test statistics (10) are completely general and can be used with any model, for instance with cognitive diagnostic models (Hansen et al., 2016; Liu et al., 2015).

For ordinal data, a statistic that combines features of M_2 and M_{ord} , M_{mixed} may be the best choice (Cai and Monroe, 2014). More specifically, this test uses the $n(K-1)$ univariate residual proportions (8) (as in M_2) and the $n(n-1)/2$ bivariate cross-products (12) (as in M_{ord}).

Note that the limited information test statistics (10) and Browne's test statistic have the same form. This is not accidental, as there is an increasing integration of methods between the categorical and continuous outcomes literatures. For instance, tests of close fit are now routinely employed with discrete data and goodness of fit indices have recently been proposed that mirror those used with continuous outcomes such as RMSEAs (Maydeu-Olivares and Joe, 2014), and TLI (Cai et al., 2021) using M statistics, or the Correlation Root Mean Squared Residual (CRMR: Maydeu-Olivares and Joe, 2014) suitable for binary or ordinal data.

The CRMR can be estimated as

$$\widehat{CRMR} = \sqrt{\sum_{i < j} \frac{(r_{ij} - \hat{\rho}_{ij})^2}{n(n-1)/2}}, \quad (13)$$

where r_{ij} and $\hat{\rho}_{ij}$ denote the sample and expected Pearson correlations. Using (11) and (12), the expected variances, covariances, and expected Pearson correlation among the categorical variables in our example are

$$\sigma_{11} = 1^2 \pi_1^{(1)} - \mu_1^2, \quad \sigma_{22} = 1^2 \pi_2^{(1)} + 2^2 \pi_2^{(2)} - \mu_2^2, \quad (14)$$

$$\sigma_{12} = \mu_{12} - \mu_1 \mu_2, \quad \rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11} \sigma_{22}}}. \quad (15)$$

Maydeu-Olivares and Joe (2014) reported that the population RMSEA₂ computed using the M_2 statistics depend on the number of categories K and suggested that a cutoff of RMSEA₂ = 0.089, 0.05, and 0.05/($K-1$) indicate adequate, close, and excellent fit, respectively. For the CRMR, they suggested cutoffs of 0.05, 0.027, and 0.027/($K-1$) indicate adequate, close, and excellent fit, respectively.

Piecewise assessment of fit and model modification

One of the main tasks performed in educational research is test construction. As a result, after the parameters have been estimated using all the items, there is particular interest in identifying misfitting items. Pearson's X^2 cannot be applied to assess the fit of single items as there are no degrees of freedom available for testing. X^2 can be applied to pairs of items but it does not follow a chi-square distribution; parameters would need to be estimated using only the items involved in the pair for X^2 to follow a chi-square distribution (Maydeu-Olivares and Liu, 2015). Liu and Maydeu-Olivares (2014) proposed using a mean and variance adjustment to X^2 to overcome this problem. They also compared the performance of a variety of statistics with known asymptotic distribution in identifying misfitting items. They concluded that z statistics for means and bivariate cross products performed best in terms of Type I and II errors. These statistics are

$$z_i = \frac{\bar{x}_i - \hat{\mu}_i}{SE(\bar{x}_i - \hat{\mu}_i)}, \quad z_{ij} = \frac{\bar{x}_{ij} - \hat{\mu}_{ij}}{SE(\bar{x}_{ij} - \hat{\mu}_{ij})}, \quad (16)$$

where $\bar{x}_i$ and $\bar{x}_{ij}$ denotes the sample means and cross products, respectively. Alternatively, a mean and variance corrected X^2 could be used. In closing, as in structural equation modeling, modification indices can also be employed in these models (Glas, 1999; Liu and Thissen, 2012, 2014).

Model selection indices

For some models with discrete or continuous indicators, it may not be possible to assess model-data fit. In those cases, researchers often simply use information criteria indices to select the best fitting model among a set of competing alternatives. The two most popular information criteria GOF indices are Akaike's (1974) information criterion (AIC), and Schwarz (1978) Bayesian information criterion (BIC),

$$AIC = -2L + 2q, \quad BIC = -2L + q \ln(N). \quad (17)$$

where L denotes the loglikelihood of the data, q the number of estimated parameters, and N sample size. We note that the AIC and BIC are not goodness of fit indices, as they do not provide interpretable information about the magnitude of model-data fit. Rather, they are model selection indices. Given a data set, a researcher chooses either the AIC or BIC, and computes it for all models under consideration. Then the model with the lowest index is selected. Notice that both the AIC and BIC combine absolute fit with model parsimony. That is, they penalize adding parameters to the model, but they do so differently. Of the two, the BIC penalizes adding parameters to the model more strongly than the AIC.

Numerical examples

Factor analysis

The most widely used instrument to assess satisfaction with life is the SWLS (Diener et al., 1985). It consists of 5 items which are to be rated on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). We used a US nationally representative sample. Sample size is $N = 543$. We treated the data as continuous and fitted a one factor model. In this example, the likelihood ratio test statistic and Browne's test provide nearly identical results and we use the more popular likelihood ratio test, adjusted by its the mean and variance to account for possible non-normality of the data, to assess model-data fit. Fit results are provided in Table 1.

We see in this table that the model does not fit the data exactly: $\bar{T}_{LR} = 24.88$, 5 df, $p < 0.001$. Hu and Bentler (1999) suggested that "for ML estimation, a cutoff value close to 0.95 for TLI and CFI, a cutoff value close to 0.08 for SRMR, and a cutoff value close to 0.06 for RMSEA are needed before we can conclude that there is a relatively good fit between the hypothesized model and the observed data". The results provided in Table 1 suggest that the model provides a relatively good fit according to the CFI, TLI, and SRMR, but not according to the RMSEA. The use of the RMSEA in models with few degrees of freedom is not recommended (Kenny et al., 2015; Shi et al., 2022). With few degrees of freedom, the sampling variability of the RMSEA increases, and sample estimates tend to reject well-fitting models.

Since the model provides a good enough fit using commonly used criteria, we could stop here, but inferences based on this model could be somewhat biased, as the model is rejected. To identify the source of misfit, we used z statistics (1) and modification indices.

The z statistics are presented in Table 3 and enable us to identify parts of the population covariance matrix that are not well reproduced by the fitted one factor model. In this table we also provide residual correlations which may be used to assess the magnitude of the misfit. The only bivariate z statistic that is statistically significant at the 5% level involves items 4 and 5. These two items involve satisfaction with past life, whereas the first three items involve satisfaction with present life. This suggests that an alternative model for these data is a two factor model: satisfaction with present life (measured by items 1 to 3) and satisfaction with past life (measured by items 4 and 5).

Table 1 Modeling satisfaction with life using factor analysis: overall goodness of fit results.

Criterion	Model		
	One factor model	Two factor model	One factor model dropping item 5
χ^2	24.88	0.689	0.489
Df	5	4	5
p	<0.001	0.95	0.78
RMSEA	0.086	0	0
$p(\text{RMSEA} \leq 0.05)$	0.033	0.996	0.935
SRMR	0.040	0.005	0.005
CFI	0.955	1	1
TLI	0.911	1	1
AIC	9959.843	9934.273	7831.607
BIC	10024.300	10003.027	7883.172

Notes: The two factor model involves correlated factors with indicators {1,2,3}, and {4,5}, respectively; χ^2 = mean and variance corrected likelihood ratio statistic, df = degrees of freedom of the null hypothesis that the fitted model reproduces exactly the population means and covariances; p = p -value associated with the null hypothesis of exact fit; RMSEA = Root Mean Squared Error of Approximation; $p(\text{RMSEA} \leq 0.05)$ = p -value associated with a null hypothesis that the population RMSEA is less or equal to 5%; SRMR = Standardized Root Mean Squared Residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike's information criterion; BIC = Bayesian information criterion. Results were obtained using Mplus (Muthén and Muthén, 2017). RMSEA, CFI, TLI estimates are not adjusted for non-normality. SRMR denotes the sample (biased) estimate.

On the other hand, the largest modification index obtained when fitting a one factor model involves the covariance between the errors of items 4 and 5, MI = 24.489, leading to an expected correlation of the errors of these items of 0.25. Thus, one could fit a one factor model with a correlation between the errors of these items. In this case, a one factor model with correlated errors and a confirmatory two factor model are equivalent (they yield the same fit) and we selected the two factor model. Goodness of fit results for this model are also provided in Table 2. We see that this model fits extraordinarily well with just one more parameter: $\bar{T}_{LR} = 0.69$, 4 df, $p = .95$. Yet, an alternative model can be obtained by dropping the worst fitting item, which can be computed as the average of the absolute value of the z statistics. This is item 5, and upon removing this item we obtain $\bar{T}_{LR} = 0.49$, 2 df, $p = 0.78$. The goodness of fit indices for this model (see Table 2) are very similar to those of the two factor model. A higher precision of measurement is obtained with the models for 5 items, than with this four item model; therefore, we still choose the two factor model.

We also provide in Table 2 the AIC and BIC for the models considered. These indices suggest choosing the two factor model over the one factor model. We note that the AIC and BIC are considerably lower for the four item model. However, the AIC and BIC should not be used to compare models with a different number of variables.

Item response theory

In the second application, we consider fitting a one-dimensional item response theory (IRT) model to the Positive Problem Orientation (PPO) scale of the Social Problem Inventory-Revised (SPSI-R D'Zurilla et al., 2002). The female Spanish normative data (Maydeu-Olivares et al., 2000) are analyzed. The sample size is 692. The scale consists of five five-category items. A number of IRT models can be fitted to these data, such as Samejima's (1969) graded response model, or Bock's (1972), Thissen et al. (2015) nominal model. There is some evidence (Maydeu-Olivares, 2005) that Samejima's model is the best fitting model in empirical applications involving ordinal data, and it is the model that we fitted. Goodness of fit results are provided in Table 4. There are $5^5 = 3125$ possible response patterns. These data are sparse and Pearson's χ^2 and the likelihood ratio G^2 statistic yield conflicting conclusions, both on 3099 degrees of freedom: χ^2 yields a p -value of 0 and G^2 yields a p -value of 1. Neither p -value is accurate. M_2

Table 2 Modeling satisfaction with life using factor analysis: residual correlations (below the diagonal) and z statistics (above the diagonal).

Item	1	2	3	4	5
1	0	0.378	0.226	-0.204	-0.717
2	0.006	0	1.665	-1.371	-1.717
3	0.003	0.020	0	-1.601	-1.344
4	-0.005	-0.028	-0.024	0	4.249
5	-0.020	-0.039	-0.026	0.166	0

Notes: z statistics larger than 2.93 are statistically significant at the 5% level with a Bonferroni correction.

Table 3 Modeling positive problem orientation using a unidimensional IRT graded model: Overall goodness of fit results.

Statistic	Statistic	df	P
χ^2	66,054.18	3099	0
G^2	1098.26	3099	1
M_2	463.11	155	<0.001
M_{mixed}	8.96	5	0.11
RMSEA ₂	0.054		
RMSEA _{mixed}	0.034		
CRMR	0.030		

Notes: Pearson's χ^2 and the likelihood ratio test statistic G^2 are full information test statistics. M_2 and M_{mixed} are limited information test statistics. M_2 assesses fit to all bivariate tables, M_{mixed} assesses fit to all univariate table and the ordinal cross-products. RMSEA₂ and RMSEA_{mixed} are RMSEAs computed based on M_2 and M_{mixed} , respectively. CRMR = Correlation Root Mean Squared Residual.

Table 4 Modeling positive problem orientation using a unidimensional IRT graded model: univariate and bivariate ordinal z statistics.

Item	1	2	3	4	5	Average
1	-3.31	-3.29	-2.10	-3.83	-2.48	3.00
2	-3.29	1.29	-1.62	-0.10	-2.00	1.66
3	-2.10	-1.62	-2.00	-4.41	-1.76	2.38
4	-3.83	-0.10	-4.41	-2.10	-2.68	2.62
5	-2.48	-2.00	-1.76	-2.68	-1.35	2.06

Notes: Univariate z statistics in the diagonal. We have boldfaced the z statistics significant at an overall 5% level using the Benjamini-Hochberg procedure. The average column is the row average of absolute values of z statistics.

yields 463.11 on 155 degrees of freedom, and the 90% confidence interval for RMSEA₂ is (0.05; 0.06). Since $p(\text{RMSEA}_2 < 0.05) = 0.14$, we conclude that the hypothesized model provides a close fit to the data using the cutoff of [Maydeu-Olivares and Joe \(2014\)](#). The ordinal M statistic cannot be computed for these data due to lack of degrees of freedom for testing. We can compute the M_{mixed} statistic which assesses fit using the univariate margins and the ordinal cross-products. We obtain $M_{\text{mixed}} = 8.96$ on 5 df, $p = 0.11$, so we cannot reject that the model fits exactly using this statistic, but we feel this result needs to be interpreted cautiously in this application, as there are no degrees of freedom to assess the ordinal associations. We also computed the CRMR (13), obtaining 0.030. Hence, on average, the residual Pearson's correlations after fitting the graded model are around this magnitude, which suggests that the degree of misfit of the model is quite small.

To investigate how the fit could be improved, we computed all univariate ordinal z_i statistics and ordinal bivariate z_{ij} statistics (16). They are displayed in [Table 4](#). The statistics that are significant at the 5% level using the Benjamini-Hochberg method are boldfaced. There is only one statistically significant univariate statistic in this table, that of item 1. There are five statistically significant bivariate statistics, three of which involve item 1, and this item shows the largest average z statistic. Item 1, "When my first attempt to solve a problem fails, I believe if I don't give up, I will eventually succeed" is considerably longer than the remaining items in this scale and it includes a conditional clause, which we believe explains why this item fits slightly worse than the rest.

Concluding remarks

A model may fit poorly and yet yield useful predictions. The fact that a model fits poorly simply means that in principle a model could be found to capture more closely the data generating model, whose predictions could be very different. Because inferences are to be as precise and valid as possible, it is important that models reproduce the data generating process as close as possible.

Researchers should always assess the goodness of fit of their models using an overall statistic. If the selected model fits well, researchers should: (a) Assess the power of the statistic against meaningful deviations from the selected model, as it may well be that the statistic has no power to distinguish between the selected model and substantively meaningful alternative models. (b) Perform a piecewise assessment of the selected model to examine if some parts of the data are not well captured by the model. On the other hand, if the model does not fit well, researchers should: (a) Perform a piecewise assessment of the model attempting to determine the source of the misfit. (b) Assess the magnitude of the discrepancy between the fitted model and the data generating model to gauge the extent to which substantive inferences drawn on the fitted model may be biased.

In closing, we encourage researchers to use goodness of fit statistics and indices for which a trustworthy confidence interval (or *p*-value) can be obtained, as doing so captures the effect of sample size. When the observed variables are discrete, this leads to additional computing time, but it protects researchers from discarding items or tests based on faulty statistics.

References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19 (6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>.
- Bandalos, D.L., 2002. The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Struct. Equ. Model.* 9 (1), 78–102. https://doi.org/10.1207/S15328007SEM0901_5.
- Barrett, P., 2007. Structural equation modelling: adjudging model fit. *Pers. Individ. Differ.* 42 (5), 815–824. <https://doi.org/10.1016/j.paid.2006.09.018>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.* 57 (1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Bentler, P.M., 1990. Comparative fit indexes in structural models. *Psychol. Bull.* 107 (2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37 (1), 29–51. <https://doi.org/10.1007/BF02291411>.
- Bollen, K.A., Arminger, G., 1991. Observational residuals in factor analysis and structural equation models. *Socio. Methodol.* 21, 235. <https://doi.org/10.2307/270937>.
- Bollen, K.A., Pearl, J., 2013. Eight myths about causality and structural equation models. In: Morgan, S.L. (Ed.), *Handbook of Causal Analysis for Social Research*. Springer Netherlands, pp. 301–328. <https://doi.org/10.1007/978-94-007-6094-3>.
- Browne, M.W., 1974. Generalized least squares estimators in the analysis of covariance structures. *S. Afr. Stat. J.* 8 (1), 1–24. https://doi.org/10.10520/AJA0038271X_175.
- Browne, M.W., 1982. Covariance structures. In: Hawkins, D.M. (Ed.), *Topics in Applied Multivariate Analysis*. Cambridge University Press, pp. 72–141.
- Browne, M.W., 1984. Asymptotically distribution-free methods for the analysis of covariance structures. *Br. J. Math. Stat. Psychol.* 37 (1), 62–83. <https://doi.org/10.1111/j.2044-8317.1984.tb00789.x>.
- Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, K.A., Long, J.S. (Eds.), *Testing Structural Equation Models*. Sage, pp. 136–162.
- Buse, A., 1982. The likelihood ratio, wald, and lagrange multiplier tests: an expository note. *Am. Statistician* 36 (3), 153–157. <https://doi.org/10.2307/2683166>.
- Cai, L., Hansen, M., 2013. Limited-information goodness-of-fit testing of hierarchical item factor models. *Br. J. Math. Stat. Psychol.* 66 (2), 245–276. <https://doi.org/10.1111/j.2044-8317.2012.02050.x>.
- Cai, L., Monroe, S., 2014. A New Statistic for Evaluating Item Response Theory Models for Ordinal Data (No. 839).
- Cai, L., Chung, S.W., Lee, T., 2021. Incremental model fit assessment in the case of categorical data: Tucker–Lewis index for item response theory modeling. *Prev. Sci.* <https://doi.org/10.1007/s1121-021-01253-4>.
- Cudeck, R., Henly, S.J., 1991. Model selection in covariance structures analysis and the “problem” of sample size: a clarification. *Psychol. Bull.* 109 (3), 512–519. <https://doi.org/10.1037/0033-2909.109.3.512>.
- Curran, P., West, S., Finch, J., 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol. Methods* 1 (1), 16–29. <https://doi.org/10.1037/1082-989X.1.1.16>.
- Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S., 1985. The satisfaction with life scale. *J. Pers. Assess.* 49 (1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13.
- D’Zurilla, T.J., Nezu, A.M., Maydeu-Olivares, A., 2002. *Manual of the Social Problem-Solving Inventory-Revised*. Multi-Health Systems, Inc.
- Glas, C.A.W., 1999. Modification indices for the 2-PL and the nominal response model. *Psychometrika* 64 (3), 273–294. <https://doi.org/10.1007/BF02294296>.
- Hansen, M., Cai, L., Monroe, S., Li, Z., 2016. Limited-information goodness-of-fit testing of diagnostic classification item response models. *Br. J. Math. Stat. Psychol.* 69 (3), 225–252. <https://doi.org/10.1111/bmsp.12074>.
- Hayakawa, K., 2019. Corrected goodness-of-fit test in covariance structure analysis. *Psychol. Methods* 24 (3), 371–389. <https://doi.org/10.1037/met0000180>.
- Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6 (1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Joe, H., Maydeu-Olivares, A., 2010. A general family of limited information goodness-of-fit statistics for multinomial data. *Psychometrika* 75 (3), 393–419. <https://doi.org/10.1007/s11336-010-9165-5>.
- Jöreskog, K.G., Sörbom, D., 1988. *LISREL 7. A Guide to the Program and Applications*, second ed. International Education Services.
- Kenny, D.A., Kaniskan, B., McCoach, D.B., 2015. The performance of RMSEA in models with small degrees of freedom. *Socio. Methods Res.* 44 (3), 486–507. <https://doi.org/10.1177/0049124114543236>.
- Liu, Y., Maydeu-Olivares, A., 2014. Identifying the source of misfit in item response theory models. *Multivariate Behav. Res.* 49 (4), 354–371. <https://doi.org/10.1080/00273171.2014.910744>.
- Liu, Y., Thissen, D., 2012. Identifying local dependence with a score test statistic based on the bifactor logistic model. *Appl. Psychol. Meas.* 36, 670–688.
- Liu, Y., Thissen, D., 2014. Comparing score tests and other local dependence diagnostics for the graded response model. *Br. J. Math. Stat. Psychol.* 67 (3), 496–513. <https://doi.org/10.1111/bmsp.12030>.
- Liu, Y., Tian, W., Xin, T., 2015. An application of M2 statistic to evaluate the fit of cognitive diagnostic models. *J. Educ. Behav. Stat.* XX (X), 1–24. <https://doi.org/10.3102/1076998615621293>.
- MacCallum, R.C., 2003. Working with imperfect models. *Multivariate Behav. Res.* 38 (1), 113–139. https://doi.org/10.1207/S15327906MBR3801_5.
- Maydeu-Olivares, A., 2005. Further empirical results on parametric versus non-parametric IRT modeling of Likert-type personality data. *Multivariate Behav. Res.* 40 (2), 261–279. https://doi.org/10.1207/s15327906mbr4002_5.
- Maydeu-Olivares, A., 2013. Goodness-of-fit assessment of item response theory models. *Measurement* 11 (3), 71–101. <https://doi.org/10.1080/15366367.2013.831680>.
- Maydeu-Olivares, A., 2017a. Maximum likelihood estimation of structural equation models for continuous data: standard errors and goodness of fit. *Struct. Equ. Model.* 24 (3), 383–394. <https://doi.org/10.1080/10705511.2016.1269606>.
- Maydeu-Olivares, A., 2017b. Assessing the size of model misfit in structural equation models. *Psychometrika* 82 (3), 533–558. <https://doi.org/10.1007/s11336-016-9552-7>.
- Maydeu-Olivares, A., Joe, H., 2005. Limited- and full-information estimation and goodness-of-fit testing in 2 n contingency tables. *J. Am. Stat. Assoc.* 100 (471), 1009–1020. <https://doi.org/10.1198/016214504000002069>.
- Maydeu-Olivares, A., Joe, H., 2006. Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika* 71 (4), 713–732. <https://doi.org/10.1007/s11336-005-1295-9>.
- Maydeu-Olivares, A., Joe, H., 2014. Assessing approximate fit in categorical data analysis. *Multivariate Behav. Res.* 49 (4), 305–328. <https://doi.org/10.1080/00273171.2014.911075>.
- Maydeu-Olivares, A., Liu, Y., 2015. Item diagnostics in multivariate discrete data. *Psychol. Methods* 20 (2), 276–292. <https://doi.org/10.1037/a0039015>.
- Maydeu-Olivares, A., Shi, D., 2017. Effect sizes of model misfit in structural equation models: standardized residual covariances and residual correlations. *Methodology* 13 (Suppl. 1), 23–30. <https://doi.org/10.1027/1614-2241/a000129>.

- Maydeu-Olivares, A., Rodríguez-Fornells, A., Gómez-Benito, J., D'Zurilla, T.J., 2000. Psychometric properties of the Spanish adaptation of the Social problem-Solving Inventory-Revised (SPSI-R). *Pers. Individ. Differ.* 29 (4), 699–708. [https://doi.org/10.1016/S0191-8869\(99\)00226-3](https://doi.org/10.1016/S0191-8869(99)00226-3).
- Moshagen, M., 2012. The model size effect in SEM: inflated goodness-of-fit statistics are due to the size of the covariance matrix. *Struct. Equ. Model.* 19 (1), 86–98. <https://doi.org/10.1080/10705511.2012.634724>.
- Muthén, L.K., Muthén, B.O., 2017. MPLUS 8. Muthén & Muthén.
- Pearl, J., 2009. *Causality: Models, Reasoning, and Inference*, second ed. Cambridge University Press. <https://doi.org/10.1017/S0266466603004109>.
- Samejima, F., 1969. Estimation of latent ability using a response pattern of graded scores. *Psychometric* 34. Monograph No. 17.
- Saris, W.E., Satorra, A., van der Veld, W.M., 2009. Testing structural equation models or detection of misspecifications? *Struct. Equ. Model.* 16 (4), 561–582. <https://doi.org/10.1080/10705510903203433>.
- Satorra, A., Bentler, P.M., 1994. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye, A., Clogg, C.C. (Eds.), *Latent Variable Analysis. Applications for Developmental Research*. Sage, pp. 399–419.
- Savalei, V., 2018. On the computation of the RMSEA and CFI from the mean-and-variance corrected test statistic with nonnormal data in SEM. *Multivariate Behav. Res.* 0 (0), 1–11. <https://doi.org/10.1080/00273171.2018.1455142>.
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Stat.* 6 (2), 461–464. <https://doi.org/10.1214/aos/1176344136>.
- Shi, D., Maydeu-Olivares, A., DiStefano, C., 2018. The relationship between the standardized root mean square residual and model misspecification in factor analysis models. *Multivariate Behav. Res.* 53 (5), 676–694. <https://doi.org/10.1080/00273171.2018.1476221>.
- Shi, D., DiStefano, C., Maydeu-Olivares, A., Lee, T., 2022. Evaluating SEM model fit with small degrees of freedom. *Multivariate Behav. Res.* 0 (0), 1–36. <https://doi.org/10.1080/00273171.2020.1868965>.
- Starr, J., Falk, C.F., Monroe, S.L., Vachon, D.D., 2021. A comparison of limited-information test statistics for a response style MIRT model. *Multivariate Behav. Res.* 56 (4), 687–702. <https://doi.org/10.1080/00273171.2020.1828024>.
- Steenkamp, J.E.M., Maydeu-Olivares, A., 2022. Unrestricted factor analysis: a powerful alternative to confirmatory factor analysis. *J. Acad. Market. Sci.* [10.1007/s11747-022-00888-1](https://doi.org/10.1007/s11747-022-00888-1).
- Thissen, D., Steinberg, L., Kuang, D., 2002. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *J. Educ. Behav. Stat.* 27 (1), 77–83. <https://doi.org/10.3102/10769986027001077>.
- Thissen, D., Cai, L., Bock, R.D., 2015. The nominal categories item response model. In: Nering, M.L., Ostini, R. (Eds.), *Handbook of Polytomous Item Response Theory Models*. Routledge, pp. 43–75. <https://doi.org/10.4324/9780203861264.ch3>.
- Tucker, L.R., Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38 (1), 1–10. <https://doi.org/10.1007/BF02291170>.
- Weisberg, S., 2005. *Applied Linear Regression*, third ed. Wiley.
- Ximénez, C., Maydeu-Olivares, A., Shi, D., Revuelta, J., 2022. Assessing cutoff values of SEM fit indices: advantages of the unbiased SRMR index and its cutoff criterion based on communality. *Struct. Equ. Model.* 1–13. <https://doi.org/10.1080/10705511.2021.1992596>.

Model-data fit evaluation: aberrant response detection

Jorge N. Tendeiro, Office of Research and Academia-Government-Community Collaboration, Education and Research Center for Artificial Intelligence and Data Innovation, Hiroshima University, Higashihiroshima, Japan

© 2023 Elsevier Ltd. All rights reserved.

Introduction	273
Item response theory	273
Person-fit analysis	274
The I_2 statistic family	274
Guttman errors	275
The extended caution indices family	275
Person response function	275
Computerized adaptive testing	276
Comparative performance of person-fit statistics	276
Applications	276
Aspects to consider	276
Summary	276
References	277

Introduction

Consider a large-scale test consisting of I items being administered to a group of N students. The test's main goal is typically to estimate each student's proficiency level on the course content under evaluation. A suitable statistical model must be considered to analyze such data. One popular statistical framework commonly used is Item Response Theory (IRT; e.g., Embretson and Reise, 2000; van der Linden, 2018). As it is always the case with statistical modeling, the question of model (mis)fit is of crucial interest. The most common IRT model fit strategy consists of focusing on the item level to determine whether there are large discrepancies between the observed and expected item scores. However, it is arguably the case that the test manager is more interested in collecting model fit information from the student's perspective. In other words, one is interested in assessing whether the response pattern provided by each student is in accordance with the statistical model being used. The question is therefore one of the validity of each student's response pattern, in light of the statistical model at hand. Lack of validity of a student's response pattern will cast doubt on his or her test result. Worse, depending on the proportion of misfitting response patterns in the data set, potentially more test results may be invalid (St-Onge et al., 2011). This is a real concern that should not be ignored (Walker, 2017).

Researchers have spent time studying how to detect misfitting response patterns, in what is nowadays best known as appropriateness measurement or person-fit analysis (PFA). In what follows, we first briefly introduce the main ideas behind IRT. Next, we explain what misfitting response patterns consist of, and what statistical techniques currently exist to aid their detection. Finally, we mention some recent applications of PFA and discuss important aspects to take into account when applying PFA.

Item response theory

IRT models are a mathematical means of relating a person's proficiency to an item's particular characteristics. Fig. 1 display the curves of three such models; these are known as item characteristic functions (IRFs). These simpler models assume that a single latent person trait accounts for the effect on the item score. The more general of the three models is the three-parameter logistic model (3PLM; Bimbaum, 1968), given by

$$P(X_i = 1 | \theta_n) = c_i + (1 - c_i) \frac{1}{\exp[-a_i(\theta_n - b_i)]} \quad (1)$$

In Eq. (1), a_i , b_i , and c_i denote respectively the discrimination, difficulty, and pseudo-guessing parameters of item i ($i = 1, \dots, I$), θ_n is the trait level of person n ($n = 1, \dots, N$), and X_i is the dichotomous score of item i (correct = 1; incorrect = 0). The two-parameter logistic model (2PLM) is a particular case of the 3PLM when $c_i = 0$. The one-parameter logistic model (1PLM), also known as the Rasch model (Rasch, 1960), further constraints the 2PLM by setting the discrimination constant (typically to 1).

There is a large plethora of IRT models beyond those mentioned above, including models that accommodate polytomous item scores, non-parametric models, multidimensional models, models for response times, and models for nonmonotone items (Embretson and Reise, 2000; Sijsma and Molenaar, 2002). Here we focus on unidimensional dichotomous IRT models.

There are various algorithms available for parameter estimation, comparing models, and assessing the fit of an IRT model. Below we focus on one particular model fit strategy.

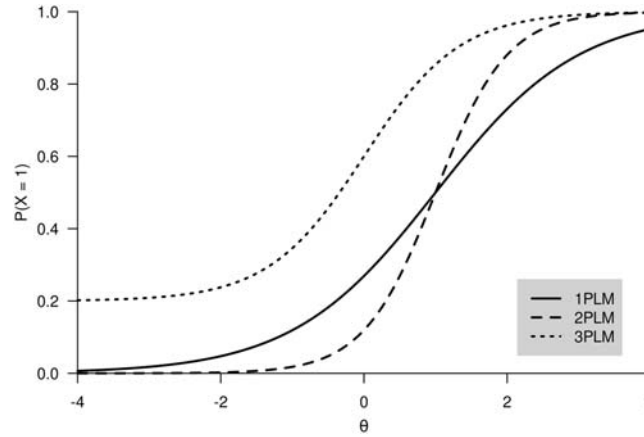


Fig. 1 Three IRFs. Item parameters: $b = 1$ (1PLM); $a = 2, b = 1$ (2PLM); $a = 1.5, b = 0, c = 0.2$ (3PLM). θ stands for the person's trait parameter.

Person-fit analysis

From a model fit perspective, PFA aims at determining whether person response patterns are as expected based on the model. For example, consider the following response patterns, with scores to ten items in increasing order of item difficulty:

Person A: 1101110100.

Person B: 0100011011.

Person A's response pattern is not very strange and could pertain to an average trait person: Most easy items were correctly answered whereas most difficult items were incorrectly answered. Person B, on the other hand, displays a suspicious set of answers: Most of the correct answers concern the most difficult items in the test. Such response behavior can be the result of one or more idiosyncratic response behavior: Warm-up issues, unfortunate lack of familiarity with the content covered by the easier items, cheating, or something else. Whatever the reason, this response pattern warrants the analyst to exert caution when analyzing and reporting the test results.

PFA allows conducting a systematic analysis of all individual response patterns in a data set. There is a large variety of statistical strategies that can be used in person-fit analysis (e.g., Karabatsos, 2003; Meijer and Sijtsma, 2001; Tendeiro et al., 2016). We will focus on some of the most popular statistics currently in use.

The l_z statistic family

Dragow et al. (1985) introduced the most popular person-fit statistic for dichotomous data in use. The l_z is the standardized log-likelihood of a person's response pattern:

$$l_z = \frac{l_0 - E(l_0)}{\sqrt{V(l_0)}} \quad (2)$$

where $l_0 = \sum_{i=1}^I X_i \ln[P_i(\theta_n)] + (1 - X_i) \ln[1 - P_i(\theta_n)]$, and $E(l_0)$ and $V(l_0)$ are the mean and variance of l_0 , respectively:

$$E(l_0) = \sum_{i=1}^I P_i(\theta_n) \ln[P_i(\theta_n)] + [1 - P_i(\theta_n)] \ln[1 - P_i(\theta_n)]$$

$$V(l_0) = \sum_{i=1}^I P_i(\theta_n) [1 - P_i(\theta_n)] \left[\ln \left(\frac{P_i(\theta_n)}{1 - P_i(\theta_n)} \right) \right]^2.$$

The l_z statistic is asymptotically standard normally distributed under the assumption of no aberrant response behavior, provided the true ability parameters θ_n ($n = 1, \dots, N$) are used. Low scores of l_z (e.g., $l_z < -2$) are indicative of misfit.

One problem with l_z is that practitioners never know the true ability values. These must be estimated and plugged in Eq.(2) in order to compute the statistic. Replacing true with estimated person parameters alters the reference distribution, and that affects computing the statistic threshold needed to determine whether a response pattern should be flagged as aberrant or not. Snijders (2001) developed an improvement now known as the l_z^* statistic which fixes the problem (see also Magis et al., 2012 for an accessible read). Sinharay (2016) further extended Snijders' statistic to mixed format tests (i.e., including both dichotomous and polyt-omous items), and Albers et al. (2016) developed a related statistic suitable for a multiple subtests setting.

Guttman errors

Guttman (1950) proposed a statistic that counts the number of “illogical” pairs of scores within a response pattern. This can easily be understood by considering again the response patterns from persons A and B above. Recall that the scores are ordered in increasing item difficulty. Consequently, one considers an error whenever a 0 precedes a 1, that is, whenever the person answered an easier item incorrectly but a harder item correctly. Such a pair of item scores is known as a Guttman error. Counting all the 01 pairs for each person, we conclude that Person A has $G = 5$ Guttman errors whereas Person B has $G = 19$ Guttman errors. Thus, we tentatively conclude that Person B displays a stranger response pattern than Person A.

However, as can be readily seen, the number of Guttman errors depends on the total number of items for a given sum score. In particular, for a sum S , G has a conditional maximum equal to $S(I-S)$. This property makes the comparison of G between subjects with different sum scores impractical. Luckily, we can norm G by dividing it by this conditional maximum, leading to a statistic known as “ G normed” (Tendeiro et al., 2016), which does allow comparisons between subjects with different sum scores. Extensions of both G and its normed version to polytomous data were developed by Molenaar (1991) and Emons (2008), respectively.

The extended caution indices family

Sato (1975) proposed the caution index (see also Harnisch and Linn, 1981). In what follows, consider that the items are ordered in increasing order of difficulty. The caution index is given by

$$C_n = 1 - \frac{\text{Cov}(\mathbf{x}_n, \mathbf{i})}{\text{Cov}(\mathbf{x}_n^*, \mathbf{i})} \quad (3)$$

where $\mathbf{x}_n$ is the response pattern of respondent n , $\mathbf{x}_n^*$ is the associated Guttman vector, and $\mathbf{i}$ is the vector with item proportions of correct answers, across examinees. To understand what the Guttman vector is, consider the response pattern of Person B introduced before: $\mathbf{x}_n = 0100011011$. Recall that the item scores are ordered in increasing item difficulty. For this response pattern the sum-score equals 5, which implies a perfect Guttman response pattern $\mathbf{x}_n^* = 1111100000$. Thus, ideally, the 5 correct items should be the 5 easier items. C_n is zero for Guttman vectors and its value tends to increase for misfitting response patterns, hence cautioning the researcher about interpreting such item scores.

Tatsuoka and Tatsuoka (1982) and Tatsuoka (1984) generalized the caution index in what is now known as the extended caution indices (ECIs). There are six such ECIs, all of which consist of variations to Sato’s caution indexed shown above (see Meijer and Sijtsma, 2001, Eqs. 42–47). Also some standardized versions of these coefficients are available (Tatsuoka, 1984). Sinharay (2016) offered asymptotic corrections to the standardized versions of the ECIs, following a similar procedure developed by Snijders (2001) for the I_z .

Person response function

Fig. 1 previously displayed examples of IRFs. A similar plot can be conceived such that the person and item swap roles; the result is the person response function (PRF; Lumsden, 1978; Weiss, 1973). One example is shown in Fig. 2. Thus, for a fixed person, one can conceive of a functional relation between a continuous item difficulty parameter and the probability of answering an item of that difficulty correctly. This function in theory is non-increasing. Thus, observed deviations from this pattern may give away deficiencies in a person’s response pattern. For example, an observed increasing PRF should warrant concern. This idea was first explored by Trabin and Weiss (1983), and later expanded by Klauer and Rettig (1990), Nering and Meijer (1998), Sijtsma and Meijer (2001), and Walker et al. (2018).

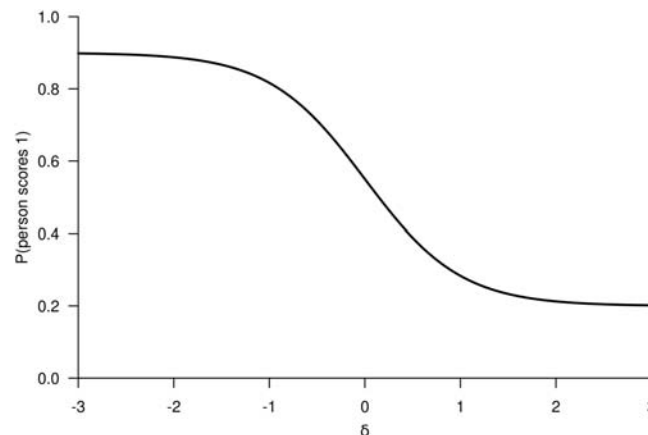


Fig. 2 Example of a PRF. δ stands for a generic item difficulty parameter.

Computerized adaptive testing

One important form of assessment is the computerized adaptive test (CAT). A CAT is a dynamic assessment which unfolds as the person takes the test. For a given test taker, each item is selected depending on his or her test performance up to that moment. In this way, different items may be assigned to different persons. Because items are optimally selected based on the up-to-date information for a given person, CATs have the potential of being shorter and of assessing each person more efficiently than traditional tests.

Some person-fit strategies adapted to CATs have been developed (Armstrong and Shi, 2009; Tendeiro et al., 2013; van Krimpen-Stoop and Meijer, 2001). These techniques are based on the theory of statistical process control (Page, 1954). In particular, cumulative sum control charts have been used in PFA. These charts are sensitive to sequential types of aberrant responding, both locally and globally. A response pattern is flagged once the chart line surpasses either an upper or lower control limit.

Comparative performance of person-fit statistics

Because the offer is so wide, practitioners may feel at a loss when they decide to choose one person-fit approach. Many simulation studies have compared the performance of various person-fit statistics; some examples include Karabatsos (2003) and references therein, Tendeiro and Meijer (2014), Ferrando (2015), and Sinharay (2017). Performance is typically judged by considering the power and false positive rates under varying item, test, and person characteristics (e.g., dichotomous vs. polytomous items, test length, aberrant response behavior). The conclusion is that there is no overall winner across the board, as results differ according to the many testing covariates. Practitioners instead tend to resort to popularity, and that also helps explaining why for example the l_z statistic and its derivatives have met a great deal of endorsement.

Applications

Recent years have witnessed an increase of interest of person-fit analysis in various fields, educational and otherwise. For example, de Vroege et al. (2018) used PF analysis to detect aberrant response patterns before fitting an exploratory factor analysis to data from a clinical study of alexithymia (a personality trait). Wanders et al. (2015) used PF to identify atypical profiles of depressive symptoms. Conrad et al. (2010) used PF to screen for suicide risk among people undergoing alcohol and drug treatment. PF is also useful to detect careless responding, for instance in internet-based quality of life assessments (Felt et al., 2017; Schneider et al., 2018). PF has also been applied to detect test collusion in large-scale testing (Belov, 2013). Cui and Li (2015) applied PF in the framework of cognitive diagnostic assessment. Response time analysis has also received attention, as response times for aberrant responders tend to be differently distributed (e.g., Lee and Chen, 2011). Finally, we can mention the work of Egberink et al. (2010) as one empirical application of CUSUM PFA.

Aspects to consider

In setting up a person-fit analytical strategy, several aspects must be considered. For example, how to choose a threshold value is important for decision making. The value -1.96 is commonly used for l_z^* , so persons with scores below this threshold are flagged as potentially aberrant. However, the specification of the threshold is far from obvious, especially when the person-fit statistic's sampling distribution is unknown. Mousavi et al. (2019) have recently explored and compared various strategies for determining suitable threshold values.

Another important consideration is IRT model misspecification. Poor fit of an IRT model may have detrimental consequences at various levels (Crişan et al., 2017), including PF analyses (Hong et al., 2020). It is therefore important to evaluate the quality of model fit before capitalizing on the results of a PFA.

Implementing PFA may be difficult for practitioners with limited programming expertise. Luckily, the R package PerFit (Tendeiro et al., 2016; see also Meijer et al., 2016) offers an accessible implementation of the most popular person-fit statistics currently in use.

Summary

Ascertaining whether response patterns are valid is a basic requirement in test scoring. PFA offers a useful set of tools that can assist practitioners. We offered a broad overview over PFA, highlighting both its features and potential. Research in PFA is still ongoing and further developments can be expected in the near future. It is our hope that PFA is more consistently used, thus increasing the robustness and fairness of test results.

References

- Albers, C.J., Meijer, R.R., Tendeiro, J.N., 2016. Derivation and applicability of asymptotic results for multiple subtests person-fit statistics. *Appl. Psychol. Meas.* 40, 274–288. <https://doi.org/10.1177/0146621615622832>.
- Armstrong, R.D., Shi, M., 2009. A parametric cumulative sum statistic for person fit. *Appl. Psychol. Meas.* 33, 391–410. <https://doi.org/10.1177/0146621609331961>.
- Belov, D.I., 2013. Detection of test collusion via Kullback-Leibler divergence: detection of test collusion. *J. Educ. Meas.* 50, 141–163. <https://doi.org/10.1111/jedm.12008>.
- Birnbaum, A., 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord, F.M., Novick, M.R. (Eds.), *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, pp. 397–479.
- Conrad, K.J., Bezruczko, N., Chan, Y.F., Riley, B., Diamond, G., Dennis, M.L., 2010. Screening for atypical suicide risk with person fit statistics among people presenting to alcohol and other drug treatment. *Drug Alcohol Depend.* 106, 92–100. <https://doi.org/10.1016/j.drugalcdep.2009.07.023>.
- Crisan, D.R., Tendeiro, J.N., Meijer, R.R., 2017. Investigating the practical consequences of model misfit in unidimensional IRT models. *Appl. Psychol. Meas.* 41, 439–455. <https://doi.org/10.1177/0146621617695522>.
- Cui, Y., Li, J., 2015. Evaluating person fit for cognitive diagnostic assessment. *Appl. Psychol. Meas.* 39, 223–238. <https://doi.org/10.1177/0146621614557272>.
- de Vroege, L., Emons, W.H.M., Sijtsma, K., van der Feltz-Cornelis, C.M., 2018. Psychometric properties of the Bermond-Vorst Alexithymia Questionnaire (BVAQ) in the general population and a clinical population. *Front. Psychiatr.* 9. <https://doi.org/10.3389/fpsy.2018.00111>.
- Drasgow, F., Levine, M.V., Williams, E.A., 1985. Appropriateness measurement with polychotomous item response models and standardized indices. *Br. J. Math. Stat. Psychol.* 38, 67–86. <https://doi.org/10.1111/j.2044-8317.1985.tb00817.x>.
- Egberink, I.J., Meijer, R.R., Veldkamp, B.P., Schakel, L., Smid, N.G., 2010. Detection of aberrant item score patterns in computerized adaptive testing: an empirical example using the CUSUM. *Pers. Individ. Differ.* 48, 921–925. <https://doi.org/10.1016/j.paid.2010.02.023>.
- Embretson, S.E., Reise, S.P., 2000. Item Response Theory for Psychologists. In: *Multivariate Applications Book Series*. L. Erlbaum Associates, Mahwah, N.J.
- Emons, W.H.M., 2008. Nonparametric person-fit analysis of polytomous item scores. *Appl. Psychol. Meas.* 32, 224–247. <https://doi.org/10.1177/0146621607302479>.
- Felt, J.M., Castaneda, R., Tiemensma, J., Depaoli, S., 2017. Using person fit statistics to detect outliers in survey research. *Front. Psychol.* 8. <https://doi.org/10.3389/fpsyg.2017.00863>.
- Ferrando, P.J., 2015. Assessing person-fit in typical-response measures. In: Reise, S.P., Revicki, D. (Eds.), *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment*. Routledge, New York, NY, pp. 128–156.
- Guttman, L., 1950. The basis for scalogram analysis. In: Stouffer, S.A., Guttman, L., Suchman, E.A. (Eds.), *Measurement and Prediction*. Princeton University Press, Princeton, NJ, pp. 60–90.
- Harnisch, D.L., Linn, R.L., 1981. Analysis of item response patterns: questionable test data and dissimilar curriculum practices. *J. Educ. Meas.* 18, 133–146. <https://doi.org/10.1111/j.1745-3984.1981.tb00848.x>.
- Hong, S.E., Monroe, S., Falk, C.F., 2020. Performance of person-fit statistics under model misspecification. *J. Educ. Meas.* 57, 423–442. <https://doi.org/10.1111/jedm.12207>.
- Karabatsos, G., 2003. Comparing the aberrant response detection performance of thirty-six person-fit statistics. *Appl. Meas. Educ.* 16, 277–298. https://doi.org/10.1207/S15324818AME1604_2.
- Klauser, K.C., Rettig, K., 1990. An approximately standardized person test for assessing consistency with a latent trait model. *Br. J. Math. Stat. Psychol.* 43, 193–206. <https://doi.org/10.1111/j.2044-8317.1990.tb00935.x>.
- Lee, Y., Chen, H., 2011. A review of recent response-time analyses in educational testing. *Psychol. Test Assess. Model.* 53, 359–379.
- Lumsden, J., 1978. Tests are perfectly reliable. *Br. J. Math. Stat. Psychol.* 31, 19–26. <https://doi.org/10.1111/j.2044-8317.1978.tb00568.x>.
- Magis, D., Raiche, G., Béland, S., 2012. A didactic presentation of Snijders's I_z^* index of person fit with emphasis on response model selection and ability estimation. *J. Educ. Behav. Stat.* 37, 57–81. <https://doi.org/10.3102/1076998610396894>.
- Meijer, R.R., Sijtsma, K., 2001. Methodology review: evaluating person fit. *Appl. Psychol. Meas.* 25, 107–135. <https://doi.org/10.1177/01466210122031957>.
- Meijer, R.R., Niessen, A.S.M., Tendeiro, J.N., 2016. A practical guide to check the consistency of item response patterns in clinical research through person-fit statistics: examples and a computer program. *Assessment* 23, 52–62. <https://doi.org/10.1177/1073191115577800>.
- Molenaar, I., 1991. A weighted Loevinger H-coefficient extending Mokken scaling to multicategory items. *Kwantitatieve Methoden* 12, 97–117.
- Mousavi, A., Cui, Y., Rogers, T., 2019. An examination of different methods of setting cutoff values in person fit research. *Int. J. Test.* 19, 1–22. <https://doi.org/10.1080/15305058.2018.1464010>.
- Nering, M.L., Meijer, R.R., 1998. A comparison of the person response function and the I_z person-fit statistic. *Appl. Psychol. Meas.* 22, 53–69. <https://doi.org/10.1177/01466216980221004>.
- Page, E.S., 1954. Continuous inspection schemes. *Biometrika* 41, 100. <https://doi.org/10.2307/2333009>.
- Rasch, G., 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research, Copenhagen.
- Sato, T., 1975. *The Construction and Interpretation of S-P Tables*. Meiji Toshio, Tokyo.
- Schneider, S., May, M., Stone, A.A., 2018. Careless responding in internet-based quality of life assessments. *Qual. Life Res.* 27, 1077–1088. <https://doi.org/10.1007/s11136-017-1767-2>.
- Sijtsma, K., Meijer, R.R., 2001. The person response function as a tool in person-fit research. *Psychometrika* 66, 191–207. <https://doi.org/10.1007/BF02294835>.
- Sijtsma, K., Molenaar, I.W., 2002. Introduction to Nonparametric Item Response Theory. In: *Number v. 5 in Measurement Methods for the Social Sciences Series*. SAGE Publications, Thousand Oaks, Calif.
- Sinharay, S., 2016. Asymptotic corrections of standardized extended caution indices. *Appl. Psychol. Meas.* 40, 418–433. <https://doi.org/10.1177/0146621616649963>.
- Sinharay, S., 2017. How to compare parametric and nonparametric person-fit statistics using real data. *J. Educ. Meas.* 54, 420–439. <https://doi.org/10.1111/jedm.12155>.
- Snijders, T.A.B., 2001. Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika* 66, 331–342. <https://doi.org/10.1007/BF02294437>.
- St-Onge, C., Valois, P., Abdous, B., Germain, S., 2011. Accuracy of person-fit statistics: a Monte Carlo study of the influence of aberrance rates. *Appl. Psychol. Meas.* 35, 419–432. <https://doi.org/10.1177/0146621610391777>.
- Tatsuoka, K.K., Tatsuoka, M.M., 1982. Standardized Extended Caution Indices and Comparisons of Their Rule Detection Rates. Research Report 82-4-ONR. Computer-Based Education Research Laboratory, University of Illinois, Urbana.
- Tatsuoka, K.K., 1984. Caution indices based on item response theory. *Psychometrika* 49, 95–110. <https://doi.org/10.1007/BF02294208>.
- Tendeiro, J.N., Meijer, R.R., 2014. Detection of invalid test scores: the usefulness of simple nonparametric statistics: detection of invalid test scores. *J. Educ. Meas.* 51, 239–259. <https://doi.org/10.1111/jedm.12046>.
- Tendeiro, J.N., Meijer, R.R., Schakel, L., Maji-de Meij, A.M., 2013. Using cumulative sum statistics to detect inconsistencies in unproctored internet testing. *Educ. Psychol. Meas.* 73, 143–161. <https://doi.org/10.1177/0013164412444787>.
- Tendeiro, J.N., Meijer, R.R., Niessen, A.S.M., 2016. PerFit: an R package for person-fit analysis in IRT. *J. Stat. Software* 74. <https://doi.org/10.18637/jss.v074.i05>.
- Trabin, T., Weiss, D., 1983. The person response curve: fit of individuals to item response theory models. In: Weiss, D. (Ed.), *New Horizons in Testing*. Academic Press, New York, pp. 83–108.
- van der Linden, W.J., 2018. *Handbook of Item Response Theory*, Three Volume Set, first ed. Chapman and Hall/CRC Press, Boca Raton, FL. <https://doi.org/10.1201/9781315119144>.

- van Krimpen-Stoop, E.M., Meijer, R.R., 2001. CUSUM-based person-fit statistics for adaptive testing. *J. Educ. Behav. Stat.* 26, 199–217. <https://doi.org/10.3102/10769986026002199>.
- Walker, A.A., Jennings, J.K., Engelhard, G., 2018. Using person response functions to investigate areas of person misfit related to item characteristics. *Educ. Assess.* 23, 47–68. <https://doi.org/10.1080/10627197.2017.1415143>.
- Walker, A.A., 2017. Why education practitioners and stakeholders should care about person fit in educational assessments. *Harv. Educ. Rev.* 87, 426–444. <https://doi.org/10.17763/1943-5045-87.3.426>.
- Wanders, R.B., Wardenaar, K.J., Penninx, B.W., Meijer, R.R., de Jonge, P., 2015. Data-driven atypical profiles of depressive symptoms: identification and validation in a large cohort. *J. Affect. Disord.* 180, 36–43. <https://doi.org/10.1016/j.jad.2015.03.043>.
- Weiss, D., 1973. The Stratified Adaptive Computerized Ability Test. Research Report 73-3. In: Department of Psychology, Psychometric Methods Program. University of Minnesota, Minneapolis.

Model-data fit evaluation: posterior checks and Bayesian model selection

Allison Ames Boykin, Nnamdi C. Ezike, and Aaron J. Myers, University of Arkansas, Fayetteville, AR, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	279
Absolute fit: posterior predictive model checking	280
Educational measurement example	280
Discrepancy measures for PPMC	282
Posterior predictive p -value	283
Relative fit	284
Bayes Factors	284
Information criterion	285
Bayesian information criterion	285
Akaike information criterion	286
Deviance information criterion	286
Watanabe-Akaike information criterion	287
Leave-one-out cross validation	287
Educational measurement example	288
References	288

Glossary

Posterior predictive distribution Posterior predictive distribution is a distribution of future observable data conditional on observed data

Posterior predictive p -value Posterior predictive p -value is the probability that the replicated data could be more extreme than the observed data as measured by the discrepancy measure

Posterior predictive model checking Posterior predictive model checking is a methodology used for assessing Bayesian model fit using replicated data simulated from the posterior predictive distribution

Relative model-data fit Critical evaluation and comparison of two or more models under consideration for the purpose selecting a model based on the balance between best fit and predictive accuracy

Multilevel model Also known as hierarchical models, random-effects models, and mixed-effects models. Models in which parameters are considered random variables with some specified distribution at one or more levels and are fixed at other levels

Introduction

The goal of statistical modeling, whether Bayesian or frequentist, is for researchers and practitioners to construct a model or series of models to help describe and make inferences about complex processes. Before the researcher can make inferences from any model, the model must be evaluated for its adequacy, regardless of estimation framework. This process of model evaluation is often referred to as model-data fit or model checking. It is the intended inferences and interpretation of model results that drive model checking procedures, with the researcher determining if the inferences made from the results are reasonable. Researchers seek to learn, through model checking procedures, specifics of the model's limitations and failures. For example, one model may provide appropriate inferences for individuals at the lower end of an observed score distribution, but inappropriate conclusions for individuals at the upper end of the distribution.

Gelman et al. (1996) provide three approaches of evaluating a model's appropriateness for the data and desired inferences for Bayesian methods. These approaches are: (1) examining the sensitivity of inferences to reasonable changes in the prior distribution and the likelihood; (2) checking that the posterior inferences are reasonable, given the substantive context of the model; and (3) checking that the model fits the data, which is the focus of this entry. Bayesian model checking encompasses both absolute and relative evaluation of model-data fit, mirroring the purposes of model checking in the frequentist framework. Absolute checking of model-data fit involves checking the appropriateness of the model for the observed data to model assumptions and desired inferences. Continuing the observed score example, absolute fit would evaluate, for example, one model fit to the observed score distribution. Relative checking of model-data fit is the process of evaluating one model against other, competing models. Relative fit might compare two competing models to each other (e.g., one symmetric and one allowing for skew), considering both parsimony and inferential errors of the models for the empirical data.

Models estimated in a Bayesian framework rely on posterior predictive model checking (PPMC) for evaluation of absolute fit. While the specifics differ for each class of model (e.g., item response theory, confirmatory factor models), the general PPMC procedure remains consistent. Details of PPMC are provided in the next section of this entry. Relative fit encompasses a broad set of procedures, including Bayes Factors and information criteria. In the final section of the entry, details of model comparison procedures are presented.

Regardless of the type fit evaluation, absolute or relative, Bayesian data analysis proceeds according to five general steps: (1) identify and collect data required to answer the research question, (2) choose a statistical model for the relevant data, (3) specify a prior distribution on the parameters, (4) obtain the posterior distributions for the model parameters (e.g., through Markov chain Monte Carlo [MCMC] estimation), and (5) evaluate model adequacy (Kruschke, 2014). Steps (2) through (5) may repeat until model adequacy is obtained. To ensure that credible model parameters are obtained in the fourth step, it is crucial to ensure convergence of the parameter(s) (Patz and Junker, 1999). For example, the Geweke statistic (Geweke, 1992), Raftery-Lewis diagnostic (Raftery and Lewis, 1992), potential scale reduction factor (PSRF) statistic (Gelman and Rubin, 1992), and a generalization of the PSRF (Brooks and Gelman, 1998) are all popular measures available to evaluate convergence in Bayesian statistics. It is after the convergence checking step that the researcher can begin to evaluate model-data fit.

Absolute fit: posterior predictive model checking

Posterior predictive model checking is a popular Bayesian approach for assessing fit; it was developed out of the posterior predictive methodology (Rubin, 1984). PPMC is based on the simple, intuitive principle that the data generated from the posterior distribution should resemble the observed data if the model provides adequate fit for the data. The posterior predictive distribution (PPD) for replicated data is a distribution of future observable data conditioned on the observed values (Rubin, 1984). That is, PPD is the distribution of simulated, or “replicated,” data that would arise from the model and parameters estimated from observed data. Specifically, the PPD is defined via:

$$p(x_{\text{new}}|\mathbf{x}) = \int p(x_{\text{new}}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}$$

where $\boldsymbol{\theta}$ is used as generic notation for all of the model parameters, x_{new} is the replicated data, and $\mathbf{x}$ is the observed data. In practice, the PPD is computed via simulation (Stern and Cressie, 2000). Computing the replicated samples is accomplished by randomly drawing values of the parameter(s) from the posterior distribution, which are then used to simulate the data using the model of interest. This process is repeated many times – often 1000 or more. The end result is a large number of data sets, each the same number of observations as the observed data.

In a simple example, assume a researcher is using a normal distribution to model an outcome variable, X , for n individuals. The model parameters are the mean, μ , and the standard deviation, σ . Although a simple problem such as this may be solved analytically via use of conjugate priors, this and future examples assume the use of MCMC estimation algorithms. In the first replicated data set ($\text{rep} = 1$), parameter values are randomly drawn from the joint posterior distribution (i.e., $\mu_{\text{rep}}, \sigma_{\text{rep}}$) and a data set of sample size n is simulated from a $\text{Normal}(\mu_{\text{rep}}, \sigma_{\text{rep}})$ distribution. In the second replicated data set ($\text{rep} = 2$), a new set of values are randomly drawn from the joint posterior distribution and a new data set of sample size n is simulated from a distribution. This process continues until a large number of data sets are replicated, resulting in the PPD, which can then be compared to the observed data. Specifically, PPMC methods assess which aspects of the data, if any, are or are not captured by the model. Any systematic differences between the observed data set and those of the replicated data sets indicate a failure of the model to explain those aspects of the data (Sinharay et al., 2006).

Systematic differences can be seen graphically, by comparing plots of observed and replicated data, or numerically, through comparison of statistical features of the observed data to the same statistical features of the replicated data. Graphical approaches include commonly used graphs such as the histogram, boxplot, and scatterplot. Numerical approaches make use of discrepancy measures, which are statistical features computed for the observed data and for each of the replicated data sets. Continuing the simple normal distribution example, skewness could be used as the discrepancy measure. Because the model is a normal distribution, which is symmetric with an excess skew of zero, the PPD will reflect this symmetry and lack of skewness. If the observed data does not have significant skewness, then the skewness value from the observed data will be close to the skewness values from the replicated data. But if the observed data does have significant skewness, either positive or negative, then the skewness value from the observed data will not be similar to the skewness values from the PPD. In this way, the specific discrepancy measure can diagnose important aspects of model-data misfit.

Educational measurement example

To demonstrate the utility of the PPMC in educational measurement, PPMC procedures are applied to the Law School Admission Test Section 6 (LSAT6; Bock and Lieberman, 1970). The LSAT6 data has been extensively used in IRT (see Bock and Lieberman, 1970; Thissen, 1982; McDonald, 2013; Kim, 2001) and is publicly available. The data set contains dichotomously-scored item responses from 1000 examinees to 5 items. Fig. 1 presents the distribution of the number of examinees across the total scores obtained on the LSAT6.

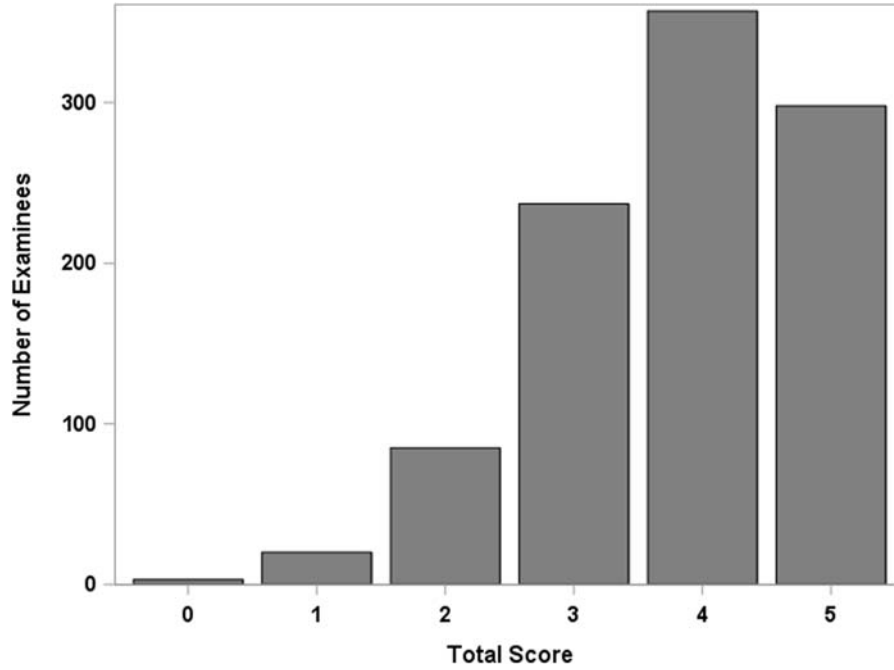


Fig. 1 Distribution of the number of examinees obtaining each score on the LSAT6.

The LSAT6 has shown adequate fit to the one-parameter logistic model (1PL; Rasch, 1960). The 1PL IRT model is the simplest IRT model for items that are scored dichotomously and models the probability of a correct response for person i to item j . The mathematical representation is given by:

$$P_{ij}(U_{ij} = 1 | b_j, \theta_i) = \frac{1}{1 + e^{-(\theta_i - b_j)}} \quad (1)$$

where θ_i is the latent ability of examinee i , b_j is difficulty parameter of item j , and $P_{ij}(U_{ij} = 1 | b_j, \theta_i)$ is the probability of a correct response by examinee i on item j given the examinee's latent ability, θ_i , and item difficulty, b_j . The 1PL model is based on a restrictive hypothesis that all items have equal discriminating power and no guessing was present. For the purpose of this illustration, we will demonstrate absolute model-data fit using the 1PL and the two-parameter logistic (2PL; Birnbaum, 1957) models. The 2PL model relaxes one of the restrictive assumptions of the 1PL model by allowing items to vary in discrimination. The 2PL model is given as:

$$P_{ij}(U_{ij} = 1 | a_j, b_j, \theta_i) = \frac{1}{1 + e^{-a_j(\theta_i - b_j)}} \quad (2)$$

where $P_{ij}(U_{ij} = 1 | a_j, b_j, \theta_i)$ represents the probability of a correct response by examinee i on item j given the examinee's latent ability, θ_i , the item difficulty, b_j , and discrimination, a_j , parameters. The Bayesian formulation of the IRT model requires the specification of priors on the item and person parameters. The prior distributions specified in fitting these models are, in this example:

$$a_j \sim \text{lognormal}(0, 4)$$

$$b_j \sim \text{normal}(0, 25)$$

$$\theta_i \sim \text{normal}(0, 1)$$

After fitting the model, we generated replicated data for the PPD. **Fig. 2** displays the distribution of the total scores for the replicated data simulated under the 1PL (top panel) and 2PL (bottom panel) models for a single replicated data set. The two distributions in **Fig. 2** closely resemble the distribution of the observed data shown in **Fig. 1**, indicating that the 1PL and 2PL models both provide reasonable approximations of the observed total score distribution.

This example illustrated one aspect of model-data fit, but it does not illustrate how well other key aspects of the observed data are represented by the models. That is, the resemblance of the distributions in **Fig. 2** shows that graphical checks of the total score distribution alone may not be conclusive. Moreover, interpretation of graphical checks may be subjective. For instance, the PPD may approximate the observed data with respect to total score distribution, but fail to describe other aspects of the posited model, such as the appropriateness of the assumption that all items have the same discriminating power.

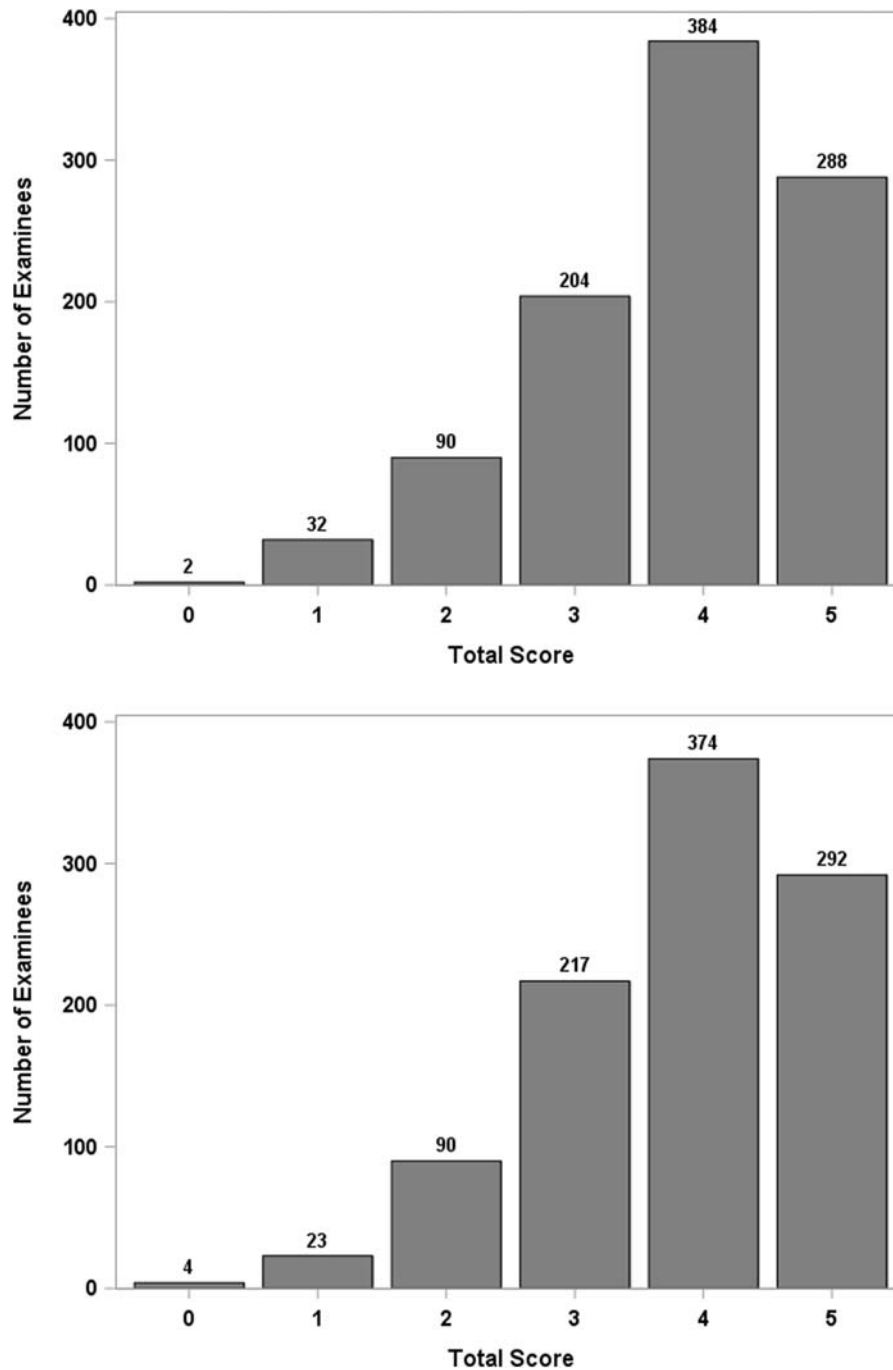


Fig. 2 The top panel shows the distribution the total score of the replicated data for the Bayesian 1PL IRT model. The bottom panel shows the distribution of the total score of the replicated data for the Bayesian 2PL IRT model.

Discrepancy measures for PPMC

Meng (1994) and Gelman et al. (1996) have further extended the methodology proposed by Rubin (1984) to quantify discrepancies between the posited model and observed data. Numerical quantification is accomplished by using discrepancy measures and posterior predictive p -values. Given there is no true perfect model, it is important to assess how the model fits the features and inferences of importance to the researcher. These specific aspects of the data can be assessed with the aid of discrepancy measures. Discrepancy measures are chosen to allow researchers to evaluate compatibility of the model with the observed data (Bayarri and Berger, 2000). In principle, any basic descriptive quantities (e.g., mean, minimum, maximum, standard deviation) or inferential statistics (e.g., correlation coefficient, fit indices) can serve as discrepancy measures. Define $T(D)$ to be a discrepancy measure of the observed values, also called the *test statistic*, and $T(D_{\text{rep}})$ is the test statistic for each replicated data set. There is an

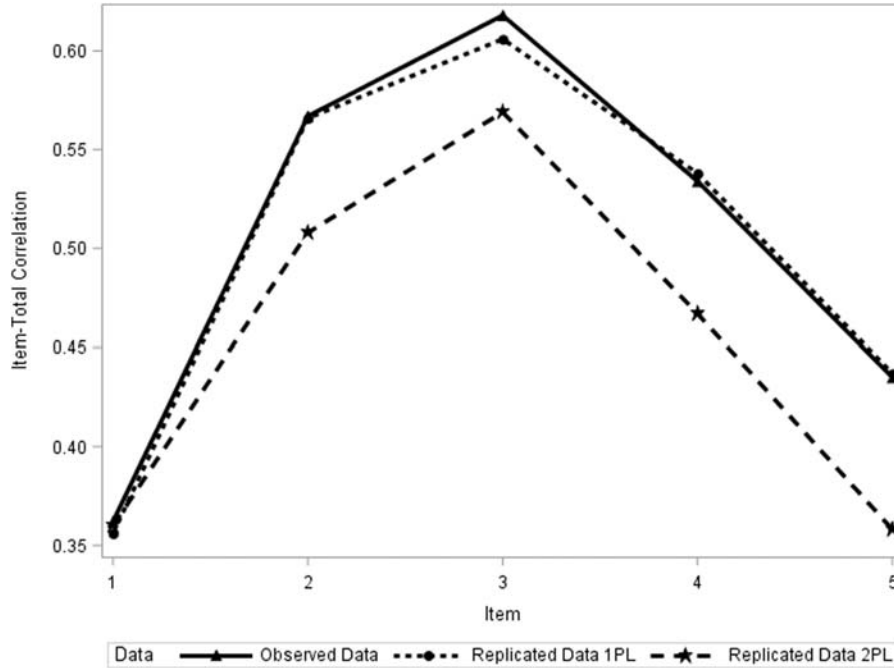


Fig. 3 Fit of the replicated 1PL and 2PL data sets to the item-total correlation.

entire distribution of discrepancy measures for the replicated data because there are multiple replicated data sets. Any significant deviation between these discrepancy measures, $T(D)$ and $T(D_{rep})$, is an indication of model misfit.

Zhu and Stone (2012) caution that discrepancy measures should be chosen to reflect sources of potential misfit relevant to a testing application. In essence, the choice of the discrepancy measure is crucial to accurately detect misfit in the model, which could lead to questionable inferences. For example, Sinharay et al. (2006) noted that the correlation of examinee scores with their binary outcomes for a particular item is effective in assessing misfit of a 1PL model associated with the assumption of equal discriminating power across all items. This correlation is the item-total score correlation, which is related to the discrimination, or slope, parameter. If the true model is the 2PL, with varied discrimination parameters, then PPD data generated from the 1PL analysis model, with a single discrimination parameter, may not have similar item-total score correlations.

Continuing the LSAT example, the item-total score correlation was computed for the observed data and for each of the 1000 data sets in the PPD. Fig. 3 illustrates the observed and median item-total correlations across 1000 data sets replicated from the 1PL and 2PL models. The item-total correlations for the 1PL replicated data closely resembles item-total correlations computed from the observed data for all five items. Thus, the 1PL appears to provide an adequate fit of the model to the data for this feature of the data. In contrast, the 2PL appears to underestimate the item-total correlation for items 2 through 5, as indicated by the 50th percentile of the item-total correlation discrepancy measure falling below the observed discrepancy measure's value for those items.

Posterior predictive p -value

In the frequentist framework, a p -value is defined as the probability of observing the test statistic, or one more extreme, under a true null hypothesis (H_0). In the Bayesian framework, the posterior predictive p -value serves as a measure of deviance between the model and the data (Berkhof et al., 2000). The Bayesian posterior predictive p -value is defined as the probability that the replicated data is more extreme than the observed data as measured by the discrepancy measure (Gelman et al., 2013). This probability is given by:

$$p_B = \Pr\{T(D_{rep}) \geq T(D) | x, H_0\} \quad (3)$$

Posterior predictive p -values near 0.5 indicate adequate model-data fit, whereas, values greater than 0.95 or less than 0.05 are often used to indicate poor model-data fit (Sinharay, 2006). The posterior predictive p -value is computed as follows:

1. Compute the discrepancy measure, $T(D)$, from the observed data set
2. Simulate many posterior predictive data sets (typically 1000 data sets)
3. Calculate $T(D_{rep})$ for each posterior predictive data set
4. Count the proportion of replications for which $T(D_{rep})$ exceeds $T(D)$.

We computed the posterior predictive p -value of the item-total score correlations as the discrepancy measure, for the 1000 replicated data sets. Fig. 4 presents a histogram of the discrepancy measure $T(D_{rep})$, when the 1PL model is used, based on 1000 replications

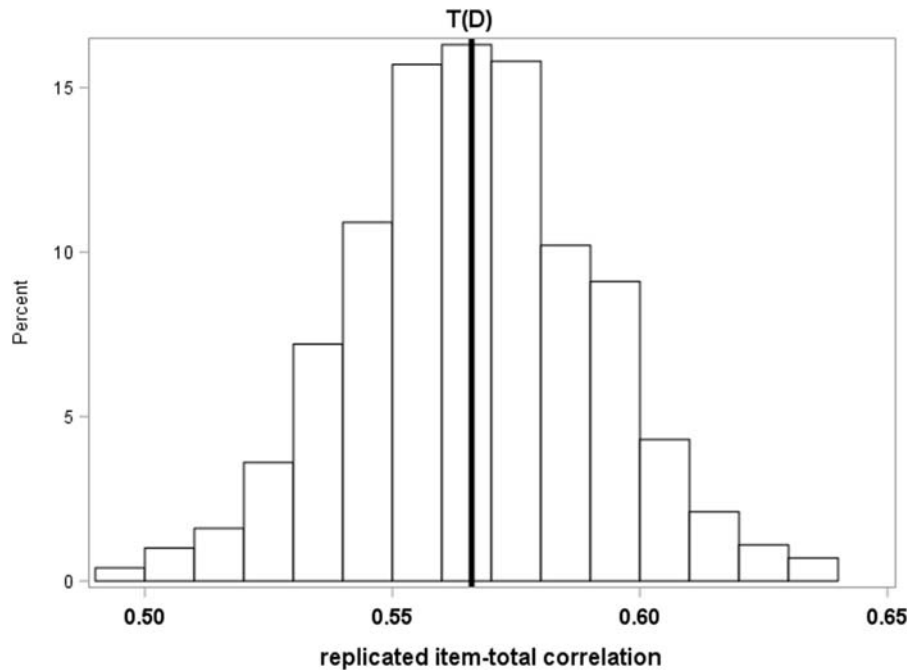


Fig. 4 Histogram of 1000 1PL model replications for the item-total discrepancy measure, $T(D_{rep})$, for item 2. The vertical line represents the observed discrepancy measure, $T(D)$.

for item 2. The vertical line in the plot indicates the observed discrepancy measure, $T(D)$. Of the 1000 $T(D_{rep})$ values, 503 values were greater than $T(D) = 0.566$, resulting in a posterior predictive p -value of $503/1000 = 0.503$. This posterior predictive p -value suggests the 1PL model is a good fit for item 2 in terms of discrimination, according to the item-total correlation discrepancy measure.

Relative fit

In contrast to evaluation of absolute model-data fit, when evaluating relative model-data fit, a researcher is not generally concerned with checking a particular model, but rather comparison among models with the goal of improving upon and finding or selecting the most useful model. Although the previously discussed PPMC framework is generally referenced when evaluating absolute fit, it may also be used for evaluation of relative model-data fit when particular characteristics of model-data fit are of interest. For example, the item-total score correlation discrepancy measure example illustrated in [Discrepancy measures for PPMC](#) provided evidence of superior fit of the 1PL model relative to the 2PL model. [Li et al. \(2006\)](#) successfully used odds ratios among item pairs as a discrepancy measure to assess the fit of an array of potential IRT testlet models fit to testlet data violating the assumption of conditional independence. Nonetheless, PPMC is generally thought of as a method of assessing absolute fit and is considered concurrently with assessment of relative fit.

Bayes Factors

Given the attention to the frequentist perspective of hypothesis testing, education practitioners and researchers often focus on research designs and analyses centered around null hypothesis significance testing (NHST). For example, pedagogical researchers might question if addition of an inquiry-based-learning component results in higher scientific literacy than simply an experiential learning approach. Such a research question, with competing theoretical models, is framed in such a way that it lends itself to the traditional method of NHST. The model comparison method of Bayes factors is the standard Bayesian hypothesis testing alternative to NHST ([Lewis and Raftery, 1997](#)). In contrast to interpretation of NHST p -values, as illustrated above, Bayes factors quantify the evidence of one hypothesis compared to another on a continuous scale ([Jeffreys, 1961](#)). More specifically, Bayes factors provide a formal method of evaluating the relative probabilities of competing models, translated to hypotheses, given the data. Thus, Bayesian hypothesis testing allows researchers to assess the strength of evidence in favor of one hypothesis over another rather than the NHST practice of simply rejecting, or failing to reject, a null hypothesis.

Consider two competing theoretical models represented by hypotheses H_0 and H_1 . The hypotheses could represent null and alternative hypotheses, regression models with different subsets of predictors, or nested 1PL and 2PL IRT models. The researcher assigns prior probabilities to each hypothesis such that their sum equals 1, $p(H_0) + p(H_1) = 1$. The data x are theorized to come

from H_0 or H_1 , where the sum of the posterior probabilities of each hypothesis equals 1, $p(H_0|x) + p(H_1|x) = 1$. Weighing the relative evidence of H_0 is facilitated by expressing the probabilities of each hypothesis in terms of their odds. Bayes theorem thus gives:

$$\frac{p(H_0|x)}{p(H_1|x)} = \frac{p(x|H_0)}{p(x|H_1)} \times \frac{p(H_0)}{p(H_1)} \quad (4)$$

where the term on the left is the posterior odds of H_0 , the first term on the right is the Bayes factor, which has the form of a likelihood ratio, and the second term on the right is the prior odds of H_0 . Thus, the Bayes factor, denoted as BF_{01} , is the ratio of the marginal densities of the data given the hypotheses (e.g., $p(x|H_0)$). The marginal densities are obtained by integrating out the model parameters θ_j ($j \in \{0, 1, \dots\}$). Note, θ_j may represent a single model parameter or vector of model parameters for model j . The marginal density for hypothesis 0, for example, is obtained by:

$$p(x|H_0) = \int p(x|H_0, \theta_0) p(\theta_0|H_0) d\theta_0 \quad (5)$$

where $p(x|H_0, \theta_0)$ is the likelihood and $p(\theta_0|H_0)$ is the prior for the parameters. Rearranging terms of Eq. (4), we can see the Bayes factor is a ratio of the posterior odds of H_0 to its prior odds:

$$BF_{01} = \frac{p(x|H_0)}{p(x|H_1)} = \frac{p(H_0|x)}{p(H_1|x)} \frac{p(H_0)}{p(H_1)} \quad (6)$$

Accordingly, BF_{01} can be interpreted as the strength of evidence in favor of H_0 or against H_1 . Consider the pedagogical researchers' hypothetical results in Table 1.

Assuming the prior probabilities of both hypotheses are equal ($p(H_0) = p(H_1) = 0.50$), the obtained $BF_{01} = \frac{p(x|H_0)}{p(x|H_1)} = \frac{.56}{.44} = 1.27$ indicates that after conducting the experiment, H_0 is 1.27 times more likely than H_1 . Kass and Raftery (1995, p. 777) provided guidelines for interpreting the strength of H_0 relative to H_1 in terms of the raw metric of BF_{01} and in the metric of twice the natural log of BF_{01} (see Table 2). Note that in practice it is more common to express Bayes factors in terms of twice their natural log:

$$2 \log BF_{01} = 2 \log \frac{p(x|H_0)}{p(x|H_1)} = 2 [\log p(x|H_0) - \log p(x|H_1)] \quad (7)$$

In practice, calculations of exact marginal densities, and therefore exact Bayes factors, are computationally difficult for more complex models (e.g., when non-conjugate priors are used). Accordingly, several approximations have been developed to facilitate use of Bayes factors (for different approximations, see Bollen et al., 2012; Han and Carlin, 2001; Kass and Raftery, 1995; Newton and Raftery, 1994; Sinharay, 2001). For example, Entink et al. (2009) demonstrate the use of Newton and Raftery's (1994) importance sampling method for comparing nested multilevel IRT models. One of the most common approximations, however, is Schwarz's (1978) Bayesian information criterion (BIC; Bollen et al., 2012).

Information criterion

Bayesian information criterion

BIC is an information criteria-based relative-fit index that was developed as an approximation of marginal densities (e.g., $p(x|H_1)$) to simplify computation of Bayes factors (Kass and Raftery, 1995; Gelman et al., 2013). As with other information criteria-based indices, the simplicity of BIC lies in the two basic components required for its computation: deviance and a penalty term. In general,

Table 1

	Hypothesis	Probability
Inquiry-based-learning component	$H_{10}H_1$	0.56
Experiential learning approach only	$H_{21}H_2$	0.44

Table 2

Guidelines for Interpreting Bayes Factors		
Evidence against H_{21}	$BF_{1201}BF_{12}$	$2\log BF_{1201}$
Not worth more than a bare mention	1–3	0–2
Positive	3–20	2–6
Strong	20–150	6–10
Very strong	> 150	> 10

Source: Kass and Raftery (1995, p. 777).

the deviance is equal to minus 2 times the log-likelihood given a point estimate, $-2 \log p(x|\hat{\theta})$. The penalty term represents the effective number of parameters in the model and is a measure of model complexity. With respect to BIC, the maximum likelihood estimate, $\hat{\theta}_{MLE}$, is used as the point estimate and the penalty term consists of the product of the number of free parameters, k , and the natural log of the sample size, N . The formula for BIC is thus:

$$BIC = -2 \log p(x|\hat{\theta}_{MLE}) + k \log N \quad (8)$$

When used for relative model comparison, models with smaller values of BIC are preferable. Notice as sample size increases, BIC provides an increasingly larger penalty per parameter and thus tends to select more parsimonious models. In fact, BIC tends to overly penalize complex models (Bollen et al., 2012).

To approximate Bayes factors, Raftery (1995) showed that BIC values for each model can be substituted for the respective marginal densities as follows:

$$2 \log BF_{01} = 2 \log \frac{p(x|H_0)}{p(x|H_1)} \approx \frac{BIC_0}{BIC_1} = BIC_0 - BIC_1 \quad (9)$$

The difference between BIC values can then be evaluated using the guidelines in Table 2. Note, the original Bayes factor metric can be obtained via $BF_{01} \approx \exp\left(\frac{1}{2}BIC_0 - BIC_1\right)$. A common criticism of Bayes factors is their sensitivity to the prior distributions used on the parameters (Kass and Raftery, 1995). Accordingly, Bayes factors should be interpreted with caution, and prior sensitivity analyses should be performed when possible. Graphical prior sensitivity analyses have been shown to be useful approaches for visualizing prior sensitivity across ranges of hyperparameter specifications (Franck and Gramacy, 2019; Sinharay and Stern, 2002).

Akaike information criterion

Akaike information criterion (AIC) is an information criteria-based relative fit index that was developed as an approximation of out-of-sample predictive accuracy of a model given the available data (Akaike, 1974). Like BIC, AIC's deviance term is based on the log-likelihood (also known as the log predictive density; Gelman et al., 2014) given the maximum likelihood point estimate $\hat{\theta}_{MLE}$. The formula for AIC is:

$$AIC = -2 \log p(x|\hat{\theta}_{MLE}) + 2k \quad (10)$$

In general, increasing the number of estimated parameters k increases the predictive accuracy of a model. Thus, the penalty term $2k$ is added as a correction for the increased predictive accuracy associated with additional estimated parameters. The correction for the number of estimated parameters, or effective number of parameters, however, presupposes linear models with normal posterior distributions. Moreover, the effective number of parameters may diverge from the number of estimated parameters when random effects (e.g., multilevel models) or informative prior information is specified, as the effective number of parameters is not clearly defined in these situations (Spiegelhalter et al., 2002). For a detailed illustration of this problem, see Gelman et al. (2014).

Interpretation of AIC is similar to BIC—and other information criteria-based indices—in that models with smaller values are preferred. The difference in penalty terms, however, results in AIC tending to prefer more complex models than BIC. Moreover, AIC tends to perform better with smaller sample sizes (Wasserman, 2000; c.f., Zhang, 1993). Despite the relative ease of interpretations, AIC and BIC are not generally recommended among Bayesian researchers (e.g., Gelman et al., 2013). A common criticism among Bayesian researchers is AIC and BIC are not generally considered fully-Bayesian approaches, as the uncertainty of the posterior distribution is not accounted for in their computations. Bayesian researchers have thus pursued relative model-data fit indices that account for the uncertainty associated with Bayesian methodology.

Deviance information criterion

Deviance information criterion (DIC) was developed as a partially Bayesian generalization of AIC that accommodates multilevel models and non-flat (i.e., informative) prior specifications (Spiegelhalter et al., 2002). DIC is analogous to other information criteria-based indices in that it is expressed as a sum of a deviance and a penalty term, and smaller values are preferable when comparing models. However, the deviance is function of the posterior mean $\hat{\theta}_{EAP}$ rather than $\hat{\theta}_{MLE}$ and thus DIC does not perform well when the posterior is not well summarized by the mean. That is, for posteriors that are not symmetric and unimodal, DIC may not be an effective criteria for evaluating relative fit.

The effective number of parameters is based on posterior estimates of model parameters. Specifically, it is the difference between the posterior mean of the deviance and the deviance of the posterior means of the parameters. An important consideration often relevant to educational researchers using models with latent variables, such as multilevel, IRT, and structural equation models (SEM), is the conditional versus marginal likelihood DIC definition (Zhang et al., 2019). The two definitions of the likelihood are concerned with the parameters of focus to the researcher. That is, the conditional likelihood DIC definition (DIC_C) is relevant when the latent variable parameters are of focus, and the marginal likelihood DIC definition (DIC_M) is relevant when the latent variables are not of focus (Merkle et al., 2019; Spiegelhalter et al., 2002). Merkle et al. (2019) provide a detailed comparison of these and other indices and illustrate the differential results obtained when the two definitions of DIC are used for model comparison.

DIC is among the most frequently used and widely implemented relative fit indices in Bayesian psychometric software. Given the two definitions of DIC can give disparate results, researchers should be careful to note which definition is implemented in their software of choice. DIC_C is currently implemented in BUGS (Lunn et al., 2012), JAGS (Plummer, 2003), the SAS MCMC procedure (SAS Institute Inc., 2018), and Stan (Stan Development Team, 2019), and DIC_M is currently implemented in blavaan (Merkle and Rosseel, 2018) and Mplus, (Muthén and Muthén, 1998–2017).

For brevity, we will limit our discussion to the DIC_C definition. The deviance term associated with DIC_C is minus 2 times the natural log of the likelihood conditional on the model and latent variables:

$$Deviance = -2 \log p_C(x|\hat{\theta}_{EAP}, \hat{\omega}_{EAP}) \quad (11)$$

where $\hat{\omega}_{EAP}$ represents the latent variables. With respect to DIC_C , the effective number of parameters p_{DIC} is expressed as:

$$p_{DIC} = -2 \left(\log p_C(x|\hat{\theta}_{EAP}, \hat{\omega}_{EAP}) + E_{x|\theta, \omega}(\log p_C(x|\theta, \omega)) \right) \quad (12)$$

where the first term is the log predictive density given the posterior parameter means and the second term is the expected value of the log predictive density. In practice, this expectation is computed as follows:

$$E_{x|\theta, \omega}(\log p_C(x|\theta, \omega)) = \frac{1}{S} \sum_{s=1}^S \log p_C(x|\theta^{(s)}, \omega^{(s)}) \quad (13)$$

where S represents the number of simulations, $s = 1, 2, \dots, S$. This use of the full posterior is why some consider DIC a partially Bayesian fit index. Of note, the computation of DIC_M follows the same general formula except the latent variables ω have been integrated out. The definition of DIC is further complicated when used in the context of multilevel latent variable models and multilevel models with greater than two levels, as the level in which the marginalization occurs must be carefully considered. Further discussion regarding the conditional versus marginal likelihood choice can be found in Millar (2009) and specifically applied to psychometric models by Merkle et al. (2019) and Zhang et al. (2019).

Watanabe-Akaike information criterion

Watanabe-Akaike information criterion (WAIC), or widely applicable information criterion, is a more fully-Bayesian approach to estimating out-of-sample predictive accuracy (Watanabe, 2010). Given the applicability to multilevel and latent variable models, conditional versus marginal likelihood definitions must be considered. Although WAIC is considered a more novel relative fit index for Bayesian inference, its use in educational measurement is somewhat limited. When it is used, the conditional versus marginal WAIC definition is often not detailed (Revuelta and Ximénez, 2017). Luo and Al-Harbi (2017); Luo (2017); da Silva et al. (2019); and Qu et al. (2020) have demonstrated the use of $WAIC_C$ in IRT, and Lu et al. (2017) demonstrated the use of $WAIC_M$ in factor analysis. Merkle et al. (2019) illustrate both likelihood definitions, along with its DIC analogs, and discuss implications for choosing among the indices.

WAIC is considered a more fully-Bayesian fit index because both the deviance and penalty terms account for the full posterior distribution. The deviance term for the conditional WAIC definition ($WAIC_C$) is computed by using the log pointwise predictive density:

$$Deviance = -2 \sum_{j=1}^J \sum_{i=1}^N \log \left(\frac{1}{S} \sum_{s=1}^S p(x_{ji}|\theta_j^{(s)}, \omega_i^{(s)}) \right) \quad (14)$$

for items $j = 1, 2, \dots, J$ and persons $i = 1, 2, \dots, N$. The penalty term p_{WAIC} is then computed as the sum of the posterior variances of the log pointwise predictive density for each data point:

$$p_{WAIC} = \sum_{j=1}^J \sum_{i=1}^N Var(\log p(x_{ji}|\theta_j, \omega_i)) \quad (15)$$

This definition makes clear why the WAIC is often considered a more fully Bayesian method of model comparison as it accounts for the uncertainty in the posterior distribution rather than strictly a point estimate of the posterior.

Leave-one-out cross validation

A more recent model comparison index introduced to educational research is the Bayesian leave-one-out cross validation approximation (LOOCV; Watanabe, 2010). LOOCV is a fully Bayesian method that accounts for the uncertainty of the posterior distribution and provides an estimate of out-of-sample predictive accuracy. WAIC and LOOCV are asymptotically equivalent (Watanabe, 2010); however, little is known about their performance in educational research (for examples, see Lu et al., 2017; Revuelta and Ximénez, 2017). Like DIC and WAIC, conditional versus marginal likelihood definitions of LOOCV must be considered. Merkle et al. (2019) discuss the connection among conditional and marginal definitions of WAIC and LOOCV and their implications. Interested readers are further referred to Gelman et al. (2013) and Vehtari et al. (2017).

Given the relatively recent appearance of WAIC and LOOCV in educational and psychometric research, there is little known regarding the relative performance of DIC, WAIC, and LOOCV and even less known with respect to their conditional and marginal likelihood definitions. Luo and Al-Harbi (2017) found $WAIC_C$ performed better than AIC, BIC, DIC_C , and similar to $LOOIC_C$ when selecting the best dichotomous IRT model. However, findings were mixed when AIC, BIC, DIC_C , and $LOOCV_C$ were used to select the best polytomous IRT model (Luo, 2017). Moreover, there have been mixed findings comparing relative performance of DIC_C and $WAIC_C$ applied to growth IRT models (Qu et al., 2020), and when comparing DIC, WAIC, and LOOCV applied to multidimensional polytomous IRT models (Revuelta and Ximénez, 2017).

Educational measurement example

Returning to the 1PL and 2PL models fit to the LSAT6 data, $WAIC_C$ and $LOOCV_C$ estimates were obtained from the R package loo (Vehtari et al., 2017). The smaller obtained $WAIC_C$ value indicated the 1PL model ($WAIC_C = 4903.20$) fit the data better than the 2PL model ($WAIC_C = 4910.60$). Note, however, some posterior variances of the log pointwise predictive density used in the calculation of p_{WAIC} were greater than 0.40 and thus the approximation of the effective number of parameters may be unreliable (Vehtari et al., 2017). Nonetheless, the smaller value of $LOOCV_C$ further supported the better fit of the 1PL model ($LOOCV_C = 4907.10$) compared to the 2PL model ($LOOCV_C = 4922.00$).

References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19 (6), 716–723.
- Bayarri, M.J., Berger, J.O., 2000. P values for composite null models. *J. Am. Stat. Assoc.* 95 (452), 1127–1142.
- Berkhof, J., Van Mechelen, I., Hoijtink, H., 2000. Posterior predictive checks: principles and discussion. *Comput. Stat.* 15 (3), 337–354.
- Birnbaum, A., 1957. Efficient Design and Use of Tests of a Mental Ability for Various Decision-Making Problems (Series Report No. 58-16). USAF School of Aviation Medicine, Randolph Air Force Base, TX.
- Bock, R.D., Lieberman, M., 1970. Fitting a response model for dichotomously scored items. *Psychometrika* 35 (2), 179–197.
- Bollen, K.A., Ray, S., Zavisca, J., Harden, J.J., 2012. A comparison of Bayes factor approximation methods including two new methods. *Socio. Methods Res.* 41 (2), 294–324.
- Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph Stat.* 7 (4), 434–455.
- daSilva, M.A., Bazán, J.L., Huggins-Manley, A.C., 2019. Sensitivity analysis and choosing between alternative polytomous IRT models using Bayesian model comparison criteria. *Commun. Stat. Simulat. Comput.* 48 (2), 601–620.
- Entink, R.K., Fox, J.P., van der Linden, W.J., 2009. A multivariate multilevel approach to the modeling of accuracy and speed of test takers. *Psychometrika* 74 (1), 21–48.
- Franck, C.T., Gramacy, R.B., 2020. Assessing Bayes factor surfaces using interactive visualization and computer surrogate modeling. *Am. Statistician* 74 (4), 359–369. <https://doi.org/10.1080/00031305.2019.1671219>.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7 (4), 457–472.
- Gelman, A., Meng, X.L., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* (6), 733–760.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*, third ed. Chapman and Hall/CRC, New York.
- Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* 24 (6), 997–1016.
- Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. In: *Bayesian Statistics*, vol. 4, pp. 641–649.
- Han, C., Carlin, B., 2001. Markov chain Monte Carlo methods for computing Bayes factors: a comparative review. *J. Am. Stat. Assoc.* 96 (455), 1122–1132.
- Jeffreys, H., 1961. *Theory of Probability*, third ed. Oxford University, Oxford, U.K.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90 (430), 773–795.
- Kim, S.H., 2001. An evaluation of a Markov chain Monte Carlo method for the Rasch model. *Appl. Psychol. Meas.* 25 (2), 163–176.
- Kruschke, J., 2014. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.
- Lewis, S.M., Raftery, A.E., 1997. Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. *J. Am. Stat. Assoc.* 92 (438), 648–655.
- Li, Y., Bolt, D.M., Fu, J., 2006. A comparison of alternative models for testlets. *Appl. Psychol. Meas.* 30 (1), 3–21.
- Lu, Z.H., Chow, S.M., Loken, E., 2017. A comparison of Bayesian and frequentist model selection methods for factor analysis models. *Psychol. Methods* 22 (2), 361–381.
- Lunn, D., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D., 2012. *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Chapman and Hall/CRC, New York, NY.
- Luo, Y., Al-Harbi, K., 2017. Performances of LOO and WAIC as IRT model selection methods. *Psychol. Test Assess. Model.* 59 (2), 183–205.
- Luo, Y., 2017. LOO and WAIC as model selection methods for polytomous items. *Psychol. Test Assess. Model.* 61 (2), 161–185.
- McDonald, R.P., 2013. *Test Theory: A Unified Treatment*. Psychology Press.
- Meng, X.L., 1994. Posterior predictive p-values. *Ann. Stat.* 22 (3), 1142–1160.
- Merkle, E.C., Rosseel, Y., 2018. blavaan: Bayesian structural equation models via parameter expansion. *J. Stat. Software* 85 (4). <https://doi.org/10.18637/jss.v085.i04>.
- Merkle, E.C., Furr, D., Rabe-Hesketh, S., 2019. Bayesian comparison of latent variable models: conditional versus marginal likelihoods. *Psychometrika* 84 (3), 802–829.
- Millar, R.B., 2009. Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors. *Biometrics* 65 (3), 962–969.
- Muthén, L.K., Muthén, B.O., 1998–2017. *Mplus User's Guide*, eighth ed. Muthén and Muthén, Los Angeles, CA. Available at: www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf. (Accessed 30 June 2020).
- Newton, M.A., Raftery, A.E., 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Stat. Soc. B* 56 (1), 3–58.
- Patz, R.J., Junker, B.W., 1999. A straightforward approach to Markov chain Monte Carlo methods for item response models. *J. Educ. Behav. Stat.* 24 (2), 146–178.
- Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Hornik, K., Leisch, F., Zeileis, A. (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. Vienna, Austria, 20–22 March 2003. Achim Zeileis.
- Qu, S., Zhang, J., Tao, J., 2020. Bayesian longitudinal multilevel item response modeling approach for studying individual growth differences. *Stat. Interface* 13 (1), 1–16.
- Raftery, A.E., Lewis, S., 1992. How many iterations in the gibbs sampler?. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), *Bayesian Statistics*, vol. 4.
- Raftery, A.E., 1995. Bayesian model selection in social research. *Socio. Methodol.* 25, 111–163.
- Rasch, G., 1960. *Studies in Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment Tests*. Nielsen & Lydiche.
- Revuelta, J., Ximénez, C., 2017. Bayesian dimensionality assessment for the multidimensional nominal response model. *Front. Psychol.* 8. <https://doi.org/10.3389/fpsyg.2017.00961> (Article 961).
- Rubin, D.B., 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Stat.* 1151–1172.

- SAS Institute Inc, 2018. SAS/STAT® 9.4 User's Guide, second ed. SAS Institute Inc, Cary, NC. Available at: <https://documentation.sas.com/api/docsets/shrref/9.4/content/shrref.pdf?locale=en>. (Accessed 30 June 2020).
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Stat.* 6 (2), 461–464.
- Sinharay, S., Stern, H.S., 2002. On the sensitivity of Bayes factors to the prior distributions. *Am. Statistician* 56 (3), 196–201.
- Sinharay, S., Johnson, M.S., Stern, H.S., 2006. Posterior predictive assessment of item response theory models. *Appl. Psychol. Meas.* 30 (4), 298–321.
- Sinharay, S., 2001. Bayes Factors for Variance Component Testing in Generalized Linear Mixed Models. Ph. D. Iowa State University. Available at: <https://lib.dr.iastate.edu/rtd/1079/>. (Accessed 30 July 2020).
- Sinharay, S., 2006. Bayesian item fit analysis for unidimensional item response theory models. *Br. J. Math. Stat. Psychol.* 59 (2), 429–449.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. B* 64 (4), 583–639.
- Stan Development Team, 2019. Stan Modeling Language Users Guide and Reference Manual (Version 2.22.0). Available at: <http://mc-stan.org/documentation/>. (Accessed 30 June 2020).
- Stern, H., Cressie, N., 2000. Posterior predictive model checks for disease mapping models. *Stat. Med.* 19, 2377–2397.
- Thissen, D., 1982. Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika* 47 (2), 175–186.
- Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27 (5), 1413–1432.
- Wasserman, L., 2000. Bayesian model selection and model averaging. *J. Math. Psychol.* 44 (1), 92–107.
- Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* 11 (116), 3571–3594.
- Zhang, X., Tao, J., Wang, C., Shi, N.Z., 2019. Bayesian model selection methods for multilevel IRT models: a comparison of five DIC-based indices. *J. Educ. Meas.* 56 (1), 3–27.
- Zhang, P., 1993. On the convergence rate of model selection criteria. *Commun. Stat.* 22 (10), 2765–2775.
- Zhu, X., Stone, C.A., 2012. Bayesian comparison of alternative graded response models for performance assessment applications. *Educ. Psychol. Meas.* 72 (5), 774–799.

Subscores; what they are, why they are in high demand, and how to evaluate their quality

Sandip Sinharay and Gautam Puhan, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	290
Subscore reporting: current practice and recommendations	290
Psychometric quality of subscores: a descriptive discussion	292
Methods to evaluate when subscores are worth reporting	292
Use of correlation or reliability of subscores	292
Application of principal component analysis and factor analysis	293
Application of the beta-binomial model	293
Fitting of multidimensional item response theory models	293
Using dimensionality assessment software such as DETECT and DIMTEST	293
Application of Haberman's classical test theory based method	294
How often do subscores have adequate psychometric quality: a brief review of existing research	294
A real data example	294
Conclusions and recommendations	295
References	296

Introduction

Subscores are scores based on subsections of a test. For example, scores on the algebra and geometry sections on a mathematics test or physical and human geography sections on a geography test could be reported as subscores. Brennan (2012) stated that users of test scores often want (indeed demand) that subscores be reported, along with total test scores, due to their potential diagnostic, remedial and instructional benefits. According to the National Research Council report "Knowing What Students Know" (2001), the purpose of assessments is to provide information about an examinee's knowledge, skills, and abilities and subscores may have the potential to provide such information. Furthermore, the US Government's No Child Left Behind (NCLB) Act of 2001 and Every Student Succeeds Act (ESSA; Every Student Succeeds Act, 2015–2016) require that state assessments provide more detailed and formative information. In particular, the latter act requires that state assessments "produce individual student interpretive, descriptive, and diagnostic reports" (p. 26); subscores might be used in such a diagnostic report. Thus, it seems quite obvious that there is substantial pressure on testing programs to report subscores, both at the individual (e.g., test taker) and aggregate (e.g., institution) levels. It is therefore not surprising that subscores are reported by several large-scale testing programs, such as SAT[®], ACT[®], Praxis, and LSAT.

In the next section, we provide a review of the literature on how subscores are typically reported and on recommendations regarding how they ought to be reported. This review is followed by a section on the quality of subscores, which includes a discussion of how existing subscores often do not have satisfactory psychometric properties. Then we include a section on techniques to evaluate the quality of subscores and a review of the existing analyses of their quality. A real data example is provided in the penultimate section. The final section includes several conclusions and recommendations.

Subscore reporting: current practice and recommendations

Fig. 1 shows an example of one section of a sample Praxis[®] test taker score report (the report can also be found at https://www.ets.org/s/praxis/pdf/sample_score_report.pdf). A total of five subscores with maximum possible values ranging from 15 to 37 are reported for this test that has the title *Praxis Elementary Education: Curriculum, Instruction, and Assessment*. The average range of subscores earned by the middle 50% of test takers is reported and can be used to compare how well an examinee did versus the other test takers who took this test. Fig. 2 shows an example of a portion of a sample institutional subscore report for an assessment designed to measure critical thinking. The subscore report is intended to enable institutions to compare their mean scaled scores with the mean scaled scores of the institutions in the comparison group. Reports like those shown in Figs. 1 and 2 might be useful to assess examinees' or institutions' specific strengths and weaknesses in different sub-content areas. However, operational score reports often provide the word of caution that these subscores are based on a small number of items and are less reliable than the official scaled scores. See Fig. 1 for an example.

Goodman and Hambleton (2004) provided a comprehensive review and critique of score reporting practices from 16 large-scale assessments offered by US states and Canadian provinces and commercial testing companies. Figures 17–21 of Goodman and Hambleton (2004) provide examples of several operational score reports that include subscores. The subscores in the score reports examined by them were reported in one of three forms—number-correct or raw scores, percent-correct scores, or percentile raw scores.

Goodman and Hambleton (2004) identified the following problems in the score reports that they examined: (a) reports that assume a high level of statistical knowledge on part of the users, (b) use of statistical jargon such as "statistical significance" or

Test / Test Category *	Your Raw Points Earned	Average Performance Range **
ELEMENTARY EDUCATION: CURRICULUM, INSTRUCTION, AND ASSESSMENT (5017)		
I. READING AND LANGUAGE ARTS	33 out of 37	23–29
II. MATHEMATICS	26 out of 31	19–25
III. SCIENCE	15 out of 20	11–15
IV. SOCIAL STUDIES	14 out of 17	9–13
V. ART, MUSIC, AND PHYSICAL EDUCATION	13 out of 15	8–12

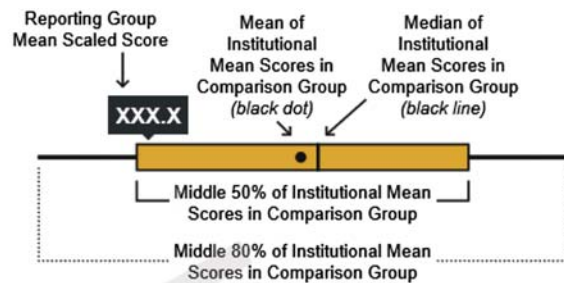
* Category-level information indicates the number of test questions answered correctly for relatively small subsets of the questions. Because they are based on small numbers of questions, category scores are less reliable than the official scaled scores, which are based on the full sets of questions. Furthermore, the questions in a category may vary in difficulty from one test to another. Therefore, the category scores of individuals who have taken different forms of the test are not necessarily comparable. For these reasons, category scores should not be considered a precise reflection of a candidate's level of knowledge in that category and ETS recommends that category information not be used to inform any decisions affecting candidates without careful consideration of such inherent lack of precision.

** The range of scores earned by the middle 50% of a group of test takers who took this form of the test at the most recent national administration or other comparable time period. N/C means that this range was not computed because fewer than 30 test takers took this form of the test or because there were fewer than 8 questions in the category or, for a constructed-response module, fewer than 8 points to be awarded by the raters. N/A indicates that this test section was not taken and, therefore, the information is not applicable.

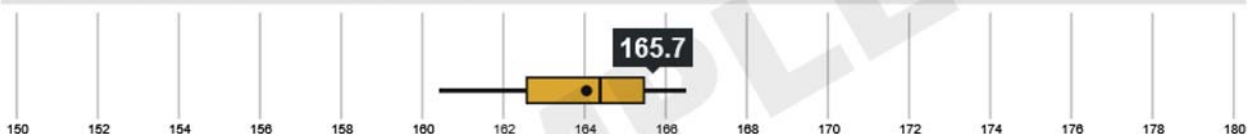
Fig. 1 A sample *Praxis*® test taker subscore report. Reproduced with permission from https://www.ets.org/s/praxis/pdf/sample_score_report.pdf.

The chart below enables you to compare the mean scaled scores for your Reporting Group with the mean scaled scores of the institutions in the Comparison Group.

The number in the dark rectangle is the mean scaled score of your Reporting Group. The figure below is a "box-and-whisker" graph of the mean scores of the institutions in the Comparison Group. The yellow bar (the "box") shows the range of the middle 50% of the institutions. The black horizontal lines (the "whiskers") extend to the range of the middle 80%. The vertical line through the box indicates the median — the point that separates the upper half of the institutions from the lower half. The black dot indicates the mean of the institutions' mean scores.



OVERALL SCALED SCORES (Scale of 150–180)



SUBSCORES (Scale of 1–10)

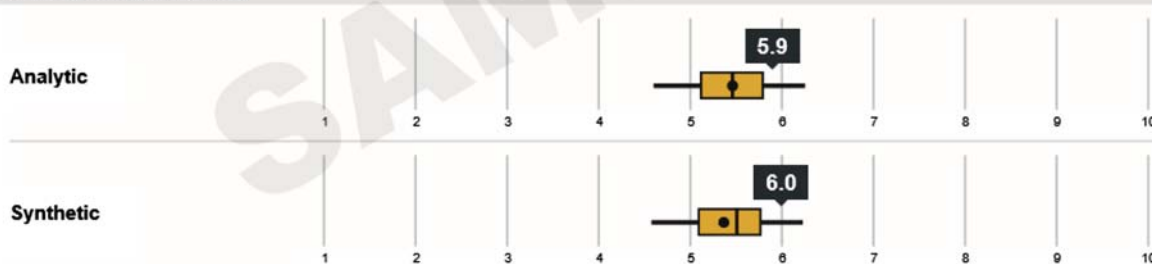


Fig. 2 Sample institutional subscore report for an assessment measuring critical thinking. Reproduced with permission from <https://www.ets.org/s/heighten/pdf/critical-thinking-sample-institutions-score-report.pdf>.

"standard error" that confused and often intimidated the users, (c) misunderstanding or ignoring of technical symbols, concepts and footnotes by the users of score reports, (d) reports that provided too much information that made it difficult for users to extract what they needed most, (e) including excessively dense graphics and displays that were daunting for readers, and (f) lack of descriptive information such as definitions and examples to aid interpretation of the results. Even in the examples provided in Figs. 1 and 2, audience without basic statistical knowledge might fail to understand some of the text provided such as "scaled score" and "box and whisker plots." Further details on more recent score-reporting practices and recommendations on how to improve upon them can

be found in [Rick and Park \(2017\)](#) who extended the findings of [Goodman and Hambleton \(2004\)](#) through an investigation using recent score reports from 23 US States and one US territory, and also in [Tannenbaum \(2018\)](#), [Slater et al. \(2018\)](#), [Zapata-Rivera et al. \(2012\)](#), and [Zenisky and Hambleton \(2012\)](#). The practices and recommendations in the aforementioned reports primarily relate to score reports for the overall test scores but several of them relate to subscores as well.

[Roberts and Gierl \(2010\)](#) noted that integration and application of interdisciplinary techniques from education, information design (e.g., [Pettersson, 2002](#)), and technology are required for effective score reporting. They provided guidelines for reporting and presenting diagnostic scores originating from cognitive diagnostic assessments (see the article on cognitive diagnostic models by Bradshaw in this encyclopedia) based on a review of educational score reporting practices and literature from the area of information design and presented a sample diagnostic report to illustrate their approach. Because subscores constitute a type of diagnostic score, several recommendations of Roberts and Gierl apply to subscore reporting as well. For example, their sample diagnostic report included three sections: (a) a top section contained an overview of the contents of the report, (b) the middle section contained diagnostic information along with item-level performance, and (c) the bottom section contained a narrative summary of the examinee's performance across all the subareas. A score report that is supposed to include information on subscores could consist of similar sections as in Roberts and Gierl; for example, the middle section could include diagnostic information (obtained from the subscores) along with item-level performance.

Psychometric quality of subscores: a descriptive discussion

Despite the demand for and apparent usefulness of subscores, they need to satisfy certain professional standards in order for them to be reported. According to [Haberman \(2008\)](#), a subscore may be reported if it has high reliability and it is distinct from the other subscores. Similarly, [Tate \(2004\)](#) has emphasized the importance of ensuring reasonable subscore performance in terms of high reliability and validity to minimize incorrect instructional and remediation decisions. These concerns agree with Standards 1.14 and 2.3 of *Standards for Educational and Psychological Testing* ([American Educational Research Association et al., 2014](#)) that require proof of adequate reliability (see the article on the topic by Chajewski in this encyclopedia), validity (see the article on the topic by Lane in this encyclopedia), and distinctness of subscores. [Monaghan \(2006\)](#) pointed out that "While they want to be responsive to the desires of the educational marketplace, testing organizations are also very concerned about the appropriate use and interpretation of subscores." Just as inaccurate information at the total test score level may lead to decisions with damaging consequences (e.g., wrongly certifying someone to be a health care professional), inaccurate information at the subscore level can also lead to incorrect remedial decisions resulting in large and needless expenses for examinees, states, or institutions ([Sinharay et al., 2011](#)). Moreover, if subscores are used in combination with total scores for high-stakes purposes such as hiring or certification, then the consequences of using subscores of inadequate psychometric quality (i.e., low reliability and validity) can be especially damaging.

Low reliabilities of subscores on educational tests are common because these subscores are often based on only a few items. For example, the score report shown in Figure 20 of [Goodman and Hambleton \(2004\)](#) includes seven subscores that are based on only five items each. Such subscores are most often outcomes of retrofitting, where reporting subscores was not a primary goal but subscores were later provided to comply with clients' requests for more diagnostic information on examinees. Furthermore, as [Sinharay and Haberman \(2008\)](#) pointed out, real data may not provide information as fine-grained as suggested or hoped for by the assessment specialist. A theory of response processes based on cognitive psychology may suggest several skills, but a test includes a limited number of items and the test may not have enough items to provide adequate information about all of these skills. For example, for the iSkills™ test (e.g., [Katz et al., 2008](#)), an expert committee identified seven performance Information and Communications Technology literacy skills. However, an exploratory factor analysis (e.g., [Tukey, 1977](#); also see the article by Trendafilov on this topic in this encyclopedia) of the data revealed only one factor and confirmatory factor models (e.g., [Gorsuch, 1983](#); also see the article by Price on confirmatory factor analysis in this encyclopedia) in which the factors corresponded to performance areas or a combination thereof did not fit the data any better than a 1-factor solution ([Katz et al., 2008](#)). As a result, only an overall Information and Communications Technology literacy score was reported. Clearly, a testing practitioner attempting to report subscores has to make an informed judgment on how much evidence the data can reliably provide and report only as much information as is reliably supported.

If subscores are used as diagnostic tools, then the consequences of using subscores are less severe compared to those of using scores for high stakes use such as certification. Nevertheless, subscores should at least be moderately reliable (e.g., 0.80 or higher, a cutoff that seems reasonable from [Nunnally, 1978](#), p. 245) to be of any remedial use. Otherwise, users would mostly be chasing noise and making incorrect remedial decisions. Even when the subscore reliabilities are high, reporting subscores is not automatically justified if the subscores do not provide unique and additional information than what is already provided by the total score or the other subscores.

There are several statistical techniques, starting from fairly basic statistical operations such as computing correlations and reliabilities to more sophisticated analyses such as factor analysis and dimensionality analysis using specialized software such as DIMTEST ([Stout, 1987](#)) and DETECT ([Kim, 1994](#)), which have been used to assess whether the subscores are worth reporting. Some of these techniques are briefly reviewed in the next section.

Methods to evaluate when subscores are worth reporting

Use of correlation or reliability of subscores

Researchers and practitioners often use simple rules to determine if subscores are worth reporting. Several researchers have used the correlations corrected for attenuation due to unreliability (see, for example, [Crocker and Algina, 1986](#), p. 237) between the different

subscores to decide whether it is reasonable to report subscores. Fairly large disattenuated correlations among subscores essentially indicate that the subscores are not distinct from each other and therefore are not worth reporting. For example, [McPeck et al. \(1976\)](#) and [Haladyna and Kramer \(2004\)](#) used the criterion that reporting of subscores is not warranted if the correlations corrected for attenuation among them are larger than 0.90. Similarly, researchers often judge the subscores to be useful only if their reliabilities are sufficiently large. For example, [Wainer et al. \(2001, p. 363\)](#) commented that the subscores for three of the four sections of the item tryout administration of the North Carolina Test of Computer Skills are insufficiently reliable to allow individual student reporting.

Application of principal component analysis and factor analysis

A simple approach to evaluate whether the subscores are distinct enough involves the computation of the eigenvalues from the correlation matrix among the subscores (or from the correlation matrix among the items) in a principal component analysis (e.g., [Hotelling, 1933](#); also see the article by [Trendafilov](#) on this topic in this encyclopedia). If, for example, for a data set that involves 5 hypothesized subscores, four of the five eigenvalues computed from the 5×5 correlation matrix of the subscores are smaller than 1 or if a plot of the eigenvalues in a decreasing order shows that the eigenvalues abruptly level out after one or two subscores, then the claim of five distinct subscores is not justified. On the contrary, the presence of five large eigenvalues would support the reporting of the five subscores. [Sinharay et al. \(2007\)](#) computed the eigenvalues from the 6×6 correlation matrix of 6 reported subscores from two forms of a test for paraprofessionals. They found that the largest eigenvalue was 4.3 for both forms while the remaining 5 eigenvalues were smaller than 0.5, suggesting that the test was unidimensional. Although the test was developed with the intention of measuring knowledge and skills in 6 distinct subareas, analysis of the data justified the reporting of a single composite score and not 6 distinct subscores. Similarly, [Stone et al. \(2010\)](#) reported, using an exploratory factor analysis of the inter-item correlation matrix, the presence of only one factor in the Eighth Grade Mathematics portion of the Spring 2006 assessment of the Delaware State Testing Program; hence, subscores were not worth reporting for the assessment.

Application of the beta-binomial model

[Harris and Hanson \(1991\)](#) suggested the approach of fitting the beta-binomial model ([Lord, 1965](#)) to the observed subscore distributions to determine if the subscores have added value over and beyond the total score. Consider a test with two purported subscores. If the subscores are not distinct, then their corresponding true subscores are functionally related. In this approach, one computes the hypothetical bivariate distribution of the two subscores under the assumptions that (a) their corresponding true subscores are functionally related, and that (b) the distributions of the subscores can be modeled by the four parameter beta binomial model ([Lord, 1965](#)). Rejection of the null hypothesis that this hypothetical bivariate distribution fits the observed bivariate distribution of the two subscores would provide evidence of the distinctness of the subscores. Harris and Hanson used a χ^2 -type statistic to determine the goodness of fit of the hypothetical bivariate distribution of the two subscores to the observed bivariate distribution of subscores. However, the method of Harris and Hanson involves significance testing with a chi-square statistic whose null distribution is not well established.

Fitting of multidimensional item response theory models

One can examine if subscores have added value by fitting a multidimensional item response theory (MIRT) model (e.g., [Reckase, 1997](#); [Ackerman et al., 2003](#); see also the article by [Cai](#) on these models in this encyclopedia) to the data. MIRT is a tool to model examinee responses to a test that measures more than one ability, for example, a test that measures both mathematical and verbal ability. In MIRT, the probability of a correct item response is a function of several abilities, rather than a single measure of ability. If subscores have added value, then a statistical test of model fit should indicate a better fit when a MIRT model instead of an unidimensional IRT model (see the article by [Ackerman et al.](#) on unidimensional IRT models in this encyclopedia) is applied to the data as long as the sample size is not too small. See, for example, [Sinharay \(2019\)](#) for a demonstration of this sort of use of MIRT.

Using dimensionality assessment software such as DETECT and DIMTEST

The DETECT software ([Zhang and Stout, 1999](#)) uses an algorithm that searches through all of the possible partitions of items into clusters to find the one that maximizes the DETECT statistic that was suggested by [Kim \(1994\)](#). The data can be considered to be unidimensional if the DETECT statistic is less than 0.10. Values between 0.10 and 0.50 would indicate a weak amount of multidimensionality, values between 0.51 and 1.00 would indicate a moderate amount of multidimensionality, and values higher than 1.00 would indicate high level of multidimensionality ([Kim, 1994](#)). The use of the DIMTEST software ([Stout, 1987](#)) involves the assessment of the statistical significance of the possible dimensional distinctiveness between two specified subtests (the Assessment Subtest or AT and the Partitioning Subtest or PT, which are supposed to measure different composite abilities if the test is multidimensional). The test statistic T calculated by DIMTEST represents the degree of dimensional distinctiveness between these two specified subsets. For example, if the testing practitioner wants to know if an Economics test has distinct dimensions such as macro or micro economics that are dimensionally distinct from the rest of the items in the test, then the macro (or micro) items can form the assessment subtest and the remaining items can form the partitioning subtest; then, a significant value of the DIMTEST index T would indicate that the two subsets are dimensionally distinct.

Application of Haberman's classical test theory based method

Haberman (2008) suggested a method based on classical test theory (see the article by Chajewski in this encyclopedia) to determine if a subscore has added value. According to the method, a subscore has added value only if it can be predicted better by the corresponding subscore on a parallel form than by the total score on the parallel form (Sinharay, 2013). To apply the method, one examines whether the reliability of a subscore is larger than another reliability-like measure referred to as the proportional reduction in mean squared error (PRMSE) of the total score. Like the reliability coefficient, the PRMSE ranges between 0 and 1 with 0 and 1 indicating the lowest and highest degrees of trustworthiness, respectively. Using the Haberman method, a subscore has added value and may be considered worth reporting only when the subscore reliability is larger than the PRMSE of the total score, which happens if and only if the observed subscore predicts the true subscore more accurately than does the observed total score.

Sinharay et al. (2007) discussed why the strategy suggested by Haberman is reasonable and how it ensures that a subscore satisfies professional standards. Conceptually, if the subscore is highly correlated with the total score (i.e., the subscore and the total score measure the same basic underlying skill), then the subscore does not provide any added value over what is already provided by the total score. A subscore is more likely to have added value according to the method of Haberman if it has high reliability and if it is distinct from the other subscores. Applications of the Haberman method can be found in Lyren (2009), Puhan et al. (2010), Meijer et al. (2017), Sinharay (2010, 2019), Sinharay et al. (2010), and Sinharay et al. (2011). For example, Sinharay et al. (2011) analyzed data from a teacher certification test with four subscores. They found the values of the reliabilities of the subscores to be 0.71, 0.85, 0.67, and 0.72, for the reading, math, social studies and science subscores, respectively. The corresponding values of the PRMSE for the total score are 0.72, 0.76, 0.74, and 0.79, respectively. Thus, only the mathematics subscore of the test had added value. Similar results were also reported by Puhan et al. (2010) for several other teacher certification tests.

How often do subscores have adequate psychometric quality: a brief review of existing research

In the previous subsection, we presented methods that can be used to examine the psychometric quality of subscores to determine if they are worth reporting. In this subsection, we will present research that shows how often subscores have adequate psychometric quality so that they can be considered worth reporting. Both operational and simulated data were used in these studies.

Sinharay (2010) performed an extensive survey regarding whether subscores have added value over the total score using the Haberman (2008) method. He used data from 25 large-scale operational tests (e.g., P-ACT+ English, P-ACT+ Math, SAT Verbal, SAT Math, SAT, SweSAT, Major Field Test in Business, etc), several of which reported subscores operationally. Of the 25 tests, sixteen had no subscores with added value according to the method of Haberman. Among the remaining nine tests, all of which had at least an average of 24 items contributing to each subscore, only some of the subscores had added value. Sinharay also performed a detailed simulation study to find out when subscores can be expected to have added value. The simulation study showed that in order to have added value, subscores have to be based on a sufficient number of (roughly 20) items, and be sufficiently distinct from one another—the disattenuated correlation between subscores has to be less than about 0.85.

Similarly, from various types of analyses, subscores were found to have little added value by Stone et al. (2010) for the 2006 eighth-grade Delaware Math Assessment, Harris and Hanson (1991) for the P-ACT+ examination, Wainer et al. (2001) for the Just-in-Time Examination conducted by the American Production and Inventory Control Society certification, Wainer et al. (2000) for the Education in the Elementary School Assessment, and Ackerman and Shu (2009) for a fifth grade end-of-grade assessment.

These studies collectively indicate that subscores on operational tests have most often been found to not satisfy professional standards.

A real data example

Some of the abovementioned approaches are demonstrated using a data set that was considered in Dai et al. (2017) and is from the 2011 administration of the Trends in International Mathematics and Science Study (TIMSS; Kastberg et al., 2013) eighth-grade mathematics assessments. TIMSS is used to assess achievement in countries around the world every four years; the assessment is used to collect a rich array of information about the educational contexts for learning mathematics and science (Mullis et al., 2004). The assessment is one of the most widely used large-scale assessments, and its public-use data are available for free. As described under the TIMSS 2011 Mathematics Framework, the following four content subscales were defined to specify the mathematics subject matter at the eighth grade: (1) number, (2) algebra, (3) geometry, and (4) data and chance. Due to the booklet design in TIMSS, students participating in the assessment were only administered a fraction (i.e., a booklet) of the total item pool. In this article, data from Booklet 11 that includes two blocks (M11 and M12) were utilized. The data contained responses from 765 students to 32 items with 6–9 items on each of the four subscales. Omitted responses were treated as incorrect. Note that although individual-level subscores are considered in the analysis here, TIMSS is not designed to report individual-level scores or subscores.

Table 1 shows the correlation coefficients (above the diagonal), deattenuated correlations (below the diagonal), and the reliabilities (Cronbach's alpha; diagonals) of the four subscores. The table shows that the subscores have low to moderate reliabilities that range between 0.57 and 0.68. The correlation coefficients range between 0.51 and 0.65 while the disattenuated correlations range between 0.75 and 0.96—so several disattenuated correlations are smaller than 0.90. Table 2 shows the reliability and the PRMSE for the total score for each subscore. The latter quantity is larger than the former for each subscore, which indicates that none of the four

Table 1 Correlation coefficients (above the diagonal), deattenuated correlations (below the diagonal, in bold font), and the reliabilities (diagonals,) of the four subscores for the TIMSS data.

	<i>Number</i>	<i>Algebra</i>	<i>Geometry</i>	<i>Data and chance</i>
Number	0.68	0.57	0.56	0.65
Algebra	0.84	0.67	0.52	0.51
Geometry	0.89	0.83	0.57	0.54
Data & chance	0.96	0.75	0.86	0.68

Table 2 Reliabilities and PRMSEs for the TIMSS data.

	<i>Number</i>	<i>Algebra</i>	<i>Geometry</i>	<i>Data and chance</i>
Reliability	0.68	0.67	0.57	0.68
PRMSE for the total score	0.84	0.72	0.77	0.78

```

> library(subscore) #Load the R package 'subscore'
> data("TIMSS11G8M.data") #Load the TIMSS data
> TIMSSdata = data.prep(TIMSS11G8M.data, c(4,9,9,6,8),
  c("Number","Algebra","Geometry","Data&Chance","Math")) #Prepare data for analysis
> Output=CTTsub(TIMSSdata,method="Haberman") #Do the analysis suggested by Haberman (2008)
> names(Output) #Get the quantities produced from the analysis
> Output$subscore.original #The subscores for each examinee
> Output$Correlation #Correlation coefficients between the subscores and the total score
> Output$Disattenuated.correlation #Disattenuated correlations between the subscores
> Output$PRMSE #Reliabilities and the PRMSEs required to determine if the subscores have added value
> library(mirt) #Load the R package 'mirt'
> m1=mirt(TIMSS11G8M.data,1) #Fit a unidimensional IRT Model to the data
> Q=matrix(0,32,4) # Write the Q-matrix for fitting a multidimensional IRT model to the data
> Q[1:9,1]=1
> Q[10:18,2]=1
> Q[19:24,3]=1
> Q[25:32,4]=1
> COV =matrix(c(FALSE,TRUE,TRUE,TRUE,TRUE,FALSE,TRUE,TRUE,TRUE,TRUE,FALSE,TRUE,TRUE,TRUE,
  TRUE,FALSE), 4)
> model=mirt.model(Q,COV=COV) #Define the multidimensional IRT model to be fitted
> m2=mirt(TIMSS11G8M.data,model) #Fit a multidimensional IRT Model to the data
> anova(m1,m2) #Compare the fit of the unidimensional and multidimensional IRT models

```

Fig. 3 Annotated R code to perform the analysis of Haberman and compare the fit of unidimensional and multidimensional models for the TIMSS data.

subscores has added value. Annotated R code that can be used to perform the computations that produced the values in [Tables 1](#) and [2](#) are provided in [Fig. 3](#). The R package “subscore” ([Dai et al., 2019](#)) was used to perform the computations. [Fig. 3](#) also includes the R code for using the R package *mirt* ([Chalmers, 2020](#)) to determine if the subscores have added value; in this analysis, a 4-dimensional IRT model did not fit the data better than a unidimensional IRT model, which supports the results from the Haberman method that the subscores do not have added value.

Conclusions and recommendations

While subscores are highly sought after by test users and are reported operationally for several large-scale assessments, operationally reported subscores often lack adequate psychometric quality. Researchers such as [Goodman and Hambleton \(2004\)](#) and [Sinharay \(2010\)](#) provided examples of several operational tests for which reported subscores are based on too few items and/or do not have added value.

While the use of content blueprints as a basis for defining subscore structure does not necessarily guarantee that the different subscores will produce highly distinct subscores, the use of psychometric approaches as a basis of constructing subscore structure

might be more useful. For example, Reckase and Xu (2014) suggested an approach based on MIRT to identify the subscore structure for a data set. However, psychometric-based approaches may lead to subscores that would be difficult to interpret. In addition, such subscores may not be reproducible when applied to other data.

Subscores should be reported only when there is sufficient evidence of adequate reliability, validity and distinctiveness of the subscores. Any reported subscore, in order to be reliable, should be based on a sufficient number of carefully constructed items. Although there are no clear guidelines in the psychometric literature on how many items are needed to provide a reliable subscore, some research has shown that at least 20 multiple-choice items are typically needed in each subtest for the corresponding subscore to yield a moderately high reliability (e.g., Puhan et al., 2010; Sinharay, 2010). Combining some subscores can increase subdomain test length and therefore might result in subscores that have higher reliability and hence added value. For example, subscores for “Physics: theory” and “Physics: applications” may be combined to yield one subscore for “Physics.” Combining subscores to meet reliability considerations, however, can create interpretation problems. For given lengths of a test and its component subtests, it is possible to increase subscore reliability by having some items measure two subdomains. But common items may make the two corresponding subscores more related and therefore less useful for diagnostic purposes. It is also important to ensure that the skills of interest are as distinct as possible from each other, though this is quite a difficult task to accomplish before seeing the data.

Although we primarily discussed subscores for individual examinees, subscores can be reported at an aggregate level as well (e.g., institution or district levels). Longford (1990) and Haberman et al. (2009) have suggested methods to examine whether aggregate-level subscores are of added value, and presented examples of situations when aggregate-level subscores do not have added value. They also presented examples where it is not justified to report a subscore for individual examinees, but it is justified to report averages of that subscore for groups of examinees.

References

- Ackerman, T., Shu, Z., 2009. Using confirmatory MIRT modeling to provide diagnostic information in large scale assessment. In: Paper Presented at the Annual Meeting of the National Council of Measurement in Education, San Diego, CA, April 2009.
- Ackerman, T.A., Gierl, M.J., Walker, C.M., 2003. Using multidimensional item response theory to evaluate educational and psychological tests. *Educ. Meas.* 22 (3), 37–51.
- American Educational Research Association, American Psychological Association, National Council for Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Brennan, R.L., 2012. Utility Indexes for Decisions About Subscores. CASMA Research Report No. 33. Center for Advanced Studies in Measurement and Assessment, Iowa City, IA.
- Chalmers, R.P., 2020. Mirt: Multidimensional Item Response Theory (R Package Version 1.32.1) [Computer Software]. Retrieved from: <http://CRAN.R-project.org/package=mirt>.
- Crocker, L., Algina, J., 1986. Introduction to Classical and Modern Test Theory. Harcourt, New York, NY.
- Dai, S., Svetina, D., Wang, X., 2017. Reporting subscores using R: a software review. *J. Educ. Behav. Stat.* 42, 617–638. <https://doi.org/10.3102/1076998617716462>.
- Dai, S., Wang, X., Svetina, D., 2019. Subscore: Sub-score Computing Functions in Classical Test Theory (R Package Version 3.1) [Computer Software]. Retrieved from: <http://CRAN.R-project.org/package=subscore>.
- Every Student Succeeds Act, Pub. L. No. 114-95 § 114 Stat. 1177, 2015–2016.
- Goodman, D.P., Hambleton, R.K., 2004. Student test score reports and interpretive guides: review of current practices and suggestions for future research. *Appl. Meas. Educ.* 17, 145–220.
- Gorsuch, R.L., 1983. Factor Analysis, second ed. Lawrence Erlbaum, Hillsdale, NJ.
- Haberman, S.J., 2008. When can subscores have value? *J. Educ. Behav. Stat.* 33, 204–229.
- Haberman, S.J., Sinharay, S., Puhan, G., 2009. Reporting subscores for institutions. *Br. J. Math. Stat. Psychol.* 62, 79–95.
- Haladyna, T.M., Kramer, G.A., 2004. The validity of subscores for a credentialing test. *Eval. Health Prof.* 24, 349–368.
- Harris, D.J., Hanson, B.A., 1991. Methods of examining the usefulness of subscores. In: Paper Presented at the Meeting of the National Council on Measurement in Education, Chicago, IL.
- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441.
- Kastberg, D., Roey, S., Ferraro, D., Lemanski, N., Erberber, E., 2013. U.S. TIMSS and PIRLS 2011 technical report and user's guide, (NCES 2013-046), U.S. Department of Education. National Center for Education Statistics, Washington, DC.
- Katz, I.R., Attali, Y., Rijmen, F., Williamson, D.M., 2008. ETS's iSkills™ assessment: measurement of information and communication technology literacy. In: Paper Presented at the Conference of the Society for Industrial and Organizational Psychology, San Francisco, CA.
- Kim, H.R., 1994. New Techniques for the Dimensionality Assessment of Standardized Test Data. Unpublished doctoral dissertation. University of Illinois, Urbana-Champaign.
- Longford, N.T., 1990. Multivariate variance component analysis: an application in test development. *J. Educ. Stat.* 15, 91–112.
- Lord, F.M., 1965. A strong true-score theory, with applications. *Psychometrika* 30, 239–270.
- Lyren, P., 2009. Reporting subscores from college admission tests. *Practical Assess. Res. Eval.* 14, 1–10.
- McPeck, M., Altman, R., Wallmark, M., Wingersky, B.C., 1976. An Investigation of the Feasibility of Obtaining Additional Subscores on the GRE Advanced Psychology Test (GRE Board Professional Report No. 74-4P). Educational Testing Service, Princeton, NJ (ERIC Document No. ED163090).
- Meijer, R.R., Boev, A.J., Tendeiro, J.N., Bosker, R.J., Albers, C.J., 2017. The use of subscores in higher education: when is this useful? *Front. Psychol.* 8. <https://doi.org/10.3389/fpsyg.2017.00305>.
- Monaghan, W., 2006. The Facts About Subscores (ETS RDC-04). Educational Testing Service, Princeton, NJ.
- Mullis, I.V., Martin, M.O., Gonzalez, E.J., Chrostowski, S.J., 2004. TIMSS 2003 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades. TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, MA.
- Nunnally, J.C., 1978. Psychometric Theory, second ed. McGraw-Hill, New York.
- Pettersson, R., 2002. Information Design: An Introduction. John Benjamins Publishing, Philadelphia, PA.
- Puhan, G., Sinharay, S., Haberman, S.J., Larkin, K., 2010. Comparison of subscores based on classical test theory. *Appl. Meas. Educ.* 23, 1–20.
- Reckase, M.D., Xu, J.R., 2014. The evidence for a subscore structure in a test of English language competency for English language learners. *Educ. Psychol. Meas.* 75, 805–825. <https://doi.org/10.1177/0013164414554416>.
- Reckase, M.D., 1997. The past and future of multidimensional item response theory. *Appl. Psychol. Meas.* 21, 25–36.
- Rick, F., Park, Y., 2017. How far have we come? A review and evaluation of changes in score reporting. In: Roberts, M.R. (Ed.), Looking Back and Moving Forward on Score Reporting Research and Practice. Symposium Held at the Annual Meeting of the National Council on Measurement in Education, San Antonio, TX.
- Roberts, M.R., Gierl, M.J., 2010. Developing score reports for cognitive diagnostic assessments. *Educ. Meas.* 29 (3), 25–38.

- Sinharay, S., Haberman, S.J., 2008. How much can we reliably know about what examinees know? *Measurement* 6, 46–49.
- Sinharay, S., Haberman, S.J., Puhane, G., 2007. Subscores based on classical test theory: to report or not to report. *Educ. Meas.* 26 (4), 21–28.
- Sinharay, S., Puhane, G., Haberman, S.J., 2010. Reporting diagnostic subscores in educational testing: temptations, pitfalls, and some solutions. *Multivariate Behav. Res.* 45, 553–573.
- Sinharay, S., Puhane, G., Haberman, S.J., 2011. An NCME instructional module on subscores. *Educ. Meas.* 30 (3), 29–40.
- Sinharay, S., 2010. How often do subscores have added value? Results from operational and simulated data. *J. Educ. Meas.* 47, 150–174.
- Sinharay, S., 2013. A note on added value of subscores. *Educ. Meas.* 32 (4), 38–42.
- Sinharay, S., 2019. Digital module 07: subscores: evaluation and reporting. *Educ. Meas.* 38, 114–115. <https://doi.org/10.1111/emip.12288>. <https://ncme.elevate.commpartners.com>.
- Slater, S., Livingston, S.A., Silver, M., 2018. Designing score reports for large-scale assessment programs: managing the design process. In: Zapata, D. (Ed.), *Score Reporting: Research and Applications*. Routledge, New York, NY, pp. 91–106.
- Stone, C.A., Ye, F., Zhu, X., Lane, S., 2010. Providing subscale scores for diagnostic information: a case study when the test is essentially unidimensional. *Appl. Meas. Educ.* 23, 266–285.
- Stout, W., 1987. A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika* 52, 589–617.
- Tannenbaum, R.J., 2018. Validity aspects of score reporting. In: Zapata, D. (Ed.), *Score Reporting: Research and Applications*. Routledge, New York, NY, pp. 9–18.
- Tate, R.L., 2004. Implications of multidimensionality for total score and subscore performance. *Appl. Meas. Educ.* 17, 89–112.
- Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison-Wesley Pub. Co, Reading, Massachusetts.
- Wainer, H., Sheehan, K.M., Wang, X., 2000. Some paths toward making Praxis scores more useful. *J. Educ. Meas.* 37, 113–140.
- Wainer, H., Vevea, J.L., Camacho, F., Reeve, B.B., Rosa, K., Nelson, L., et al., 2001. Augmented scores—“borrowing strength” to compute scores based on small numbers of items. In: Thissen, D., Wainer, H. (Eds.), *Test Scoring*. Erlbaum Associates, Mahwah, NJ, pp. 343–387.
- Zapata-Rivera, D., VanWinkle, W., Zwick, R., 2012. *Applying Score Design Principles in the Design of Score Reports for CBAL™ Teachers* (ETS Research Memorandum RM-12-20). ETS, Princeton, NJ.
- Zenisky, A.L., Hambleton, R.K., 2012. Developing test score reports that work: the process and best practices for effective communication. *Educ. Meas.* 31 (2), 21–26.
- Zhang, J., Stout, W., 1999. The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika* 64, 213–249.

Statistical methods for detection of test fraud on educational assessments

Sandip Sinharay, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

The types of test fraud and the statistical approaches to detect them	298
Score differencing	299
Analysis of erasures and answer changes	299
Copying-detection and similarity analyses	300
Person-fit analysis	300
Use of response-time models	300
Other analyses	301
Software for detection of test fraud	301
Real data examples	301
Example 1: similarity analysis	302
Example 2: analysis of answer changes	302
Example 3: person-fit analysis	303
Example 4: use of score differencing to detect item preknowledge	303
Conclusions and recommendations	305
References	306

Testing programs employ a range of tools including innovative test design, use of identification procedures and biometric data to authenticate examinees, and proctoring to prevent, deter, and impede cheating on educational tests. Despite the testing programs' best efforts, examinees manage to successfully cheat on these tests (Wollack and Schoenig, 2018). In fact, numerous recent instances of fraud on educational tests (e.g., Jacob and Levitt, 2003; Maynes, 2013) severely undermine the integrity of the test results. While various types of fraudulent behavior have been observed, Wollack and Schoenig (2018) divided the types of fraudulent behavior into three main types: (a) test tampering, (b) answer-copying and collusion and (c) item preknowledge.

Standard 6.6 of the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014) includes the recommendation among others that testing programs with high-stakes consequences should have defined procedures for detecting potential testing irregularities. Some guidelines on methods for detecting fraudulent behavior on tests can be found in a special report on testing and data integrity published by NCME (2012).

Naturally, there is a growing interest in statistical/psychometric methods for detecting fraudulent behavior on tests, which is evident from the publication of several edited volumes (Cizek and Wollack, 2017; Kingston and Clark, 2014; Wollack and Fremer, 2013) and numerous journal articles, book articles, and conference presentations (e.g., Allen and Ghattas, 2016; McLeod et al., 2003; Sinharay, 2017; 2018a; 2018b; Sinharay et al., 2017; van der Linden, 2009; van der Linden and Guo, 2008; van der Linden and Lewis, 2015; van der Linden and Sotaridona, 2006; Wollack et al., 2015; Wang et al., 2017; Wollack and Schoenig, 2018) on the topic. Note that these recent publications add to several often-cited publications on the topic from the previous millennium and early in this millennium (e.g., Buss and Novick, 1980; Karabatsos, 2003; Lewis and Thayer, 1998; van der Linden and van Krimpen-Stoop, 2003; Wollack, 1997). Summaries of some of the statistical/psychometric methods presented in these publications can be found in, for example, Ferrara (2017) and He et al. (2020). Most of the statistical approaches for detection of test fraud employ item response theory (IRT) models (e.g., de Ayala, 2009; Hambleton and Swaminathan, 1985, and the Item Response Theory (Unidimensional) article in this volume) to model the item scores and some of the approaches employ response-time models (RTMs; e.g., De Boeck and Jeon, 2019; van der Linden, 2016, and the Response Time Modeling article in this volume).

The goal of this paper is to provide a brief review of the statistical/psychometric methods for detecting fraud on educational assessments that are embodied in the aforementioned publications. Brief descriptions are provided of score differencing, analysis of erasures and answer changes, similarity analysis, person-fit analysis, and use of response-time models, which are the five main types of statistical/psychometric methods for detecting test fraud. Only the underlying concepts and principles behind the various statistical methods are provided. The mathematical descriptions of the methods are not described and can be found in the original references that are provided. Several real data examples demonstrate how to apply some of these methods.

The types of test fraud and the statistical approaches to detect them

Wollack and Schoenig (2018) divided the types of test fraud that are common in educational tests into the three following main types:

- *Test tampering*: In its most common form, classroom teachers or administrators change student answers after the tests are complete (e.g., Wollack et al., 2015). Although less common, test tampering can take place on computerized tests as well.

Table 1 The three types of test fraud and five types of statistical approaches to detect them.

Approaches	Test tampering	Answer-copying and collusion	Item preknowledge
Score differencing	✓		✓
Analysis of erasures and answer changes	✓		
Copying-detection and similarity analyses		✓	
Person-fit analysis	✓	✓	✓
Use of response-time models		✓	✓

- *Answer-copying and collusion*: Examinees often copy answers from other examinees, most often those sitting close to them. Collusion involves two or more examinees working together (when they are not allowed to) to answer the test items.
- *Item preknowledge*: Examinees benefit from item preknowledge when a “source” (such as a teacher, a test preparation company, or a website) shares test questions and/or their answers and then several beneficiaries memorize the assessment questions and/or answers. The items that are shared are usually referred to as “compromised” items.

Wollack and Schoenig (2018) also noted the existence of five types of statistical/psychometric approaches to detect the three aforementioned types of fraud.

Table 1 shows the three types of test fraud (the column names) and the five types of statistical/psychometric approaches (row names) identified by Wollack and Schoenig (2018). A check mark for a cell indicates that the approach in the heading of the corresponding row can be applied to detect the type of fraud corresponding to that column. For example, a check mark for score differencing and test tampering indicates that score differencing can be used to detect test tampering.

The five main types of statistical approaches for detecting test fraud that were mentioned by Wollack and Schoenig (2018) are described below. A couple of the most popular approaches under each type are briefly mentioned as well. Various types of test fraud lead to various types of aberrance or irregularity in the examinee data. For example, multiple types of test fraud lead to a better performance on a subset of the test compared to the rest of the test (e.g., Sinharay and Jensen, 2019), answer-copying often leads to unusual similarity between the answer options chosen by pairs of examinees (e.g., Wollack, 1997), and most types of test fraud leads to a general misfit of an IRT model to the examinee data. Each statistical approach is designed to detect one type of irregularity caused by test fraud. For example, score differencing is designed to detect better performance on a subset of the test while person-fit analysis is designed to detect the misfit of an IRT model to the examinee data.

Score differencing

Score differencing involves a test of the null hypothesis of equal performance of an examinee over two sets of items, S_1 and S_2 . The alternative hypothesis is that the performance is better, presumably due to test fraud, on one of the item sets. To perform score differencing, one adopts an IRT framework and assumes that the true ability of the examinee on S_1 and S_2 respectively are θ_1 and θ_2 , respectively, and performs the null hypothesis $H_0: \theta_1 = \theta_2$. Score differencing can be used to detect several types of test fraud including item preknowledge (e.g., Sinharay, 2017) where the item sets S_1 and S_2 are the set of compromised items and the remaining items, and test tampering (e.g., Sinharay et al., 2017), where the item sets S_1 and S_2 are the set of items with erasures and the remaining items. Existing approaches for score differencing include (a) a variation of the optimal index (Drasgow et al., 1996), which is essentially the ratio of the likelihoods of an examinee under test fraud and that under no fraud, (b) the L_s statistic (Sinharay, 2017) that is a variation of the likelihood ratio test statistic for testing the null hypothesis $H_0: \theta_1 = \theta_2$, and (c) the predictive checking method (Wang et al., 2017) that involves the computation of a Bayesian p-value indicating the extremeness of the performance on the compromised items.

Analysis of erasures and answer changes

Erasure analysis gained prominence after the alleged educator cheating in 2009 in the public schools of the US city of Atlanta (e.g., Kingston, 2013). Erasures, or, in general, answer changes (that is a more appropriate term for computerized tests that allow answer changes), are found roughly for one out of every 50 answers irrespective of test fraud (e.g., Primoli et al., 2011). While most of these answer changes are benign (that is, not fraudulent), fraudulent answer changes arise when teachers or school administrators help students by correcting students' wrong answers on their answer sheets after the test, often to hide underachievement by their class or school (van der Linden and Lewis, 2015). In analysis of answer changes, an investigator attempts to find out if the answer changes are fraudulent. In 2013, as many as 33 US states reported using analysis of answer changes to detect possible fraudulent test-taking behavior in their K-12 assessments (Government Accountability Office, 2013). However, analysis of answer changes in several states involves only the use of the average wrong-to-right answer-change count for a school or school-district (e.g., Bishop and Egan, 2017, pp. 204–205), a statistic whose power has been shown to be rather low, that is, the statistic would fail to flag many cheaters (e.g., Sinharay, 2018b). Additional indices that have been suggested for performing analysis of answer changes include the erasure detection index (EDI; Wollack et al., 2015), EDI for groups of examinees (Wollack and Eckerly, 2017), the L-index (Sinharay et al., 2017), the posterior predictive p-value corresponding to the EDI (Sinharay and

Johnson, 2017), and extensions of the EDI for groups of examinees (Peng and Sinharay, 2020; Sinharay, 2018b). These indices mostly involve an examination of whether the examinees' performance on the items with answer changes is much better given the performance on the items with no answer changes—much better performance on the former items typically indicates that the answer changes are fraudulent. The EDI is the standardized value, computed under an IRT model, of the wrong-to-right answer-change count of an examinee and the L-index is a variation of the likelihood ratio test statistic (and is conceptually similar to the L_s statistic; Sinharay, 2017) for the null hypothesis that the examinee performance on the items with answer changes is the same as that on the items with no answer changes.

Copying-detection and similarity analyses

The goal of *copying-detection and similarity analyses* is to detect potential collusion among a pair or group of examinees by investigating the similarity of their answers and flagging the examinees who produce an observed number of answer matches that significantly exceeds the number of matches predicted under the model. A “match” in this context may refer to matching incorrect answers or to matching correct and incorrect answers. Copying-detection and similarity analyses are performed using *answer-copying indices* and *similarity indices*. While both of these two types of indices quantify the similarity of the answers of pairs of examinees, they differ in the way they are computed. In computing the answer-copying indices for a pair of examinees, one first specifies a source and a copier in the pair and computes the expected number of answer matches conditional on the estimated ability of the copier and the answer string of the source. In computing the similarity indices for a pair of examinees, one does not specify a source or a copier and computes the expected number of answer matches conditional on the estimated abilities of the pair of examinees.

The answer-copying indices that are most popular include the K-index (e.g., Holland, 1996), which, for a pair of examinees, is the chance of observing at least the observed number of matching incorrect responses if the pair were working independently, or the closely related probability of matching incorrect responses (PMIR; e.g., Lewis and Thayer, 1998) the ω index (Wollack, 1997) that is the standardized value of the number of matching responses; and the probability of observing at least the observed number of matching responses under the generalized binomial model if the examinee pair were working independently (van der Linden and Sotaridona, 2006). van der Linden and Lewis (2015) suggested an approach to compute the posterior odds of answer-copying between a pair of examinees, where

$$\text{posterior odds of answer-copying} = \frac{\text{posterior probability that answer-copying occurred}}{\text{posterior probability that answer-copying did not occur}}$$

The approach does not employ any existing answer-copying index and instead employs combinatorial expressions and a reformulation of the simple recursive algorithm of Lord and Wingersky (1984) for the calculation of the distribution of the number-correct score. The approach is computation-intensive, but they suggested an efficient recursive algorithm for the calculation of the posterior odds of answer-copying.

Maynes (2014) suggested the similarity index M4 that consists of the number of identical correct and the number of identical incorrect answers. The probability distribution of the statistic is assumed to follow a generalized trinomial distribution, where probabilities of matching responses are estimated using item response probabilities from the nominal response model (Bock, 1972). Wollack and Maynes (2017) suggested a similarity index that can be used to detect groups of test-takers engaged in collusion. The approach employs cluster-analysis methods (e.g., Everitt et al., 2011) in conjunction with the M4 index (Maynes, 2014) to find clusters of examinees whose responses are similar.

Person-fit analysis

In person-fit analysis, an investigator tries to determine whether the item scores of an examinee follows an IRT model. In the presence of cheating, the item scores usually do not follow an IRT model because, for example, an examinee who cheats may answer more difficult items correctly than what an IRT model predicts. There exist several indices for performing person-fit analysis—these are referred to as person-fit statistics. Meijer and Sijtsma (2001) provided a review of several of these indices. Popular person-fit statistics include the caution indices (Tatsuoka, 1984), the H^T statistic (Smith, 1986), the l_z statistic (Drasgow et al., 1985), and the l_z^* statistic (Snijders, 2001; Sinharay, 2016c).

Glas and Meijer (2003), Sinharay (2015) and Sinharay (2016a) suggested performing person-fit analysis using the Bayesian p -values or *posterior predictive p -values* (e.g., Gelman et al., 1996) corresponding to several person-fit indices. While their Bayesian approach is computation-intensive, the approach does not depend on asymptotic results, and, may lead to more satisfactory false positive rate and/or power compared to frequentist approaches.

Use of response-time models

While there exist several RTMs, the most popular of them, arguably, is the lognormal model for response times (LNMRT; van der Linden, 2006). Let t_i denote the response time of a randomly chosen examinee on item i , where $i = 1, 2, \dots, I$. Let us define

$$y_i = \log(t_i)$$

Under the lognormal model for response times (LNMRT; van der Linden, 2006), y_i , $i = 1, 2, \dots, I$, are independent given τ and

$$y_i | \tau \sim \mathcal{N}\left(\beta_i - \tau, \frac{1}{\alpha_i^2}\right) \quad (1)$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 . That is, the probability density function of y_i is given by

$$f(y_i | \tau, \alpha_i, \beta_i) = \frac{\alpha_i}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha_i^2(y_i - \beta_i + \tau)^2} \quad (2)$$

The parameter τ is the examinee's speed parameter; a larger value of the parameter results in smaller expected response times on all the items for the examinee. The parameter β_i is the time-intensity parameter for item i ; a larger value of the parameter results in larger expected response times for all examinees on the item. The parameter α_i is the discrimination parameter for item i ; a larger value of the parameter leads to more information on and hence smaller standard error of the examinee speed parameters.

The hierarchical/joint model of van der Linden (2007) involves the application of a model for response times in combination with an IRT model for the item scores. In applying the joint model, one typically assumes that

- the response time follows the lognormal model (LNMRT; van der Linden, 2006) defined above.
- the item scores follow an IRT model; the flexibility of the approach allows the use of any IRT model such as the 2-parameter logistic model (2PLM).

van der Linden and Guo (2008) suggested a Bayesian approach that employs the hierarchical model of van der Linden (2007). The examinee-item combinations whose observed response times are too extreme compared to the corresponding posterior predictive density according to the joint model are flagged. An examinee with too many flagged examinee-item combinations is flagged for possible fraudulent behavior. van der Linden (2009) suggested a bivariate lognormal model for the distribution of the response times on a test for a pair of examinees. The correlation parameter of the model indicates the degree to which the response-times of the pair are similar. Therefore, the estimated value of the correlation and a test of hypothesis about the correlation provides information on whether the pair of examinees were involved in collusion, which is one type of test fraud. Mariani et al. (2014) and Sinharay (2018c) respectively suggested a Bayesian and frequentist variations of an approach in which the LNMRT is employed and an aberrant behavior is flagged when the sum of squared standardized residuals is large. Wang et al. (2016) suggested a two-stage approach to detect fraudulent behavior; in the first stage, a mixture hierarchical model is fitted to both the item scores and response times to distinguish solution behavior from aberrant behavior; in the second stage, a Bayesian residual index is constructed to differentiate rapid guessing from fraudulent behavior. Fox and Mariani (2017) suggested a Bayesian approach to person-fit analysis that adopts the hierarchical modeling approach of van der Linden (2007); large posterior probabilities of aberrant item-scores, aberrant response times, or both indicate possible fraudulent behavior. Sinharay and Johnson (2020) suggested an approach to detect item preknowledge that also adopts the hierarchical modeling approach of van der Linden (2007); much better and faster performance on the compromised items, found from a statistical hypothesis test, indicates potential preknowledge.

Other analyses

There are several other types of analyses that have been suggested for detecting fraudulent behavior on tests. Such analyses include the matching of response patterns (e.g., Haberman and Lee, 2017) to detect key sharing among examinees, modified IRT models for detecting various types of test fraud (e.g., McLeod et al., 2003; Segall, 2002), and application of data mining methods to detect various types of cheating (Man et al., 2019; Toton and Maynes, 2018; Zopluoglu, 2019).

Software for detection of test fraud

While companies specializing on test security have their proprietary software packages for the detection of test fraud, there is a lack of publicly available software packages for the purpose, with the exception of the SIFT (Assessment Systems Corporation, ASC) software, the R (R Core Team, 2021) packages CopyDetect (Zopluoglu, 2018) and PerFit (Tendeiro et al., 2016), and the R function lzstarMix (Sinharay, 2016b).

Real data examples

In the following discussion, four data examples demonstrate the use of various statistical tools to detect various types of test fraud. Each tool/approach that is used in an analysis has been found to perform satisfactorily compared to competing approaches in comparison studies. For example, Example 2 involves the frequentist L-index that was found to perform satisfactorily compared to other frequentist statistics for erasure analysis by Sinharay et al. (2017).

Example 1: similarity analysis

A data set including the item scores of 1644 examinees on one form of an international credentialing examination was available. While about three-fourths of the examinees are from the USA, the other one-fourth are from 27 different countries, with India ($n = 159$) and Philippines ($n = 156$) being the most popular countries among them. The data set was analyzed in several articles of Cizek and Wollack (2017) including in Zopluoglu (2017) who computed the values of several answer-copying indices for all examinee pairs within each test center. The examinees took the test in a total of 326 test centers. The test form includes 170 operational items that are dichotomously scored. The credentialing organization that administers the examination identified 61 items on the form as compromised. The organization also flagged 48 individuals on the form as possible cheaters from a variety of statistical analyses and an investigative process that brought in other information. These 48 individuals can be treated as truly aberrant for all practical purposes because of the rigor of the investigative process. Six of these flagged individuals belonged to a specific test center (that was the test center for 27 examinees in the data set) referred to as TC1.

The nominal response model (NRM; Bock, 1972) was fit to the data using the R package mirt (Chalmers, 2012). The ω index (Wollack, 1997) was computed for all within-center examinee pairs across all the test centers (5640 of them). The ω index was statistically significant at the 1% level for 367 pairs of examinees among the 5640 pairs, that is, for 6.5% of all pairs—this percentage indicates an unusually large extent of answer-copying present for the data. The percentage of statistically significant values of ω was 27.5% at the 1% significance level in test center TC1. Among the 48 examinees flagged by the credentialing organization, 15 examinees feature in at least one of the 367 pairs with statistically significant values of ω ; interestingly, this set of 15 examinees includes all six individuals belonging to test center TC1 who were flagged by the credentialing organization. For a pair of examinees from test center #2305 both of whom were flagged by the credentialing organization, the number of matching incorrect answer options is 17 (that is the third quartile of the corresponding variable over the within-center examinee pairs) and the p-value for the ω index is 0.00002. Thus, the application of the ω index seems to produce useful information in detecting test fraud in the form of similarly answered items for examinees at the same test center.

Example 2: analysis of answer changes

The data set consisted of the scores of a random sample of 2710 high-school students on 46 dichotomous items on a computerized end-of-course assessment for a particular subject area from a US state. The test is high-stakes for both the students and teachers. The students in the sample were from about 30 schools. The item parameters for the three-parameter logistic model (3PLM) were estimated for the data using marginal maximum likelihood estimation—the 3PLM showed adequate item fit for the data. The median number of answer changes among the students in the sample was 1 (that represents about 2% of all items), which is in agreement with the prevalence of erasures reported by Primoli et al. (2011).

Multiple strong sources of evidence pointed to test fraud at one particular high school, henceforth denoted as school HS_1 . The data set included 205 students from School HS_1 among whom 159 made at least one answer change. The median number of answer changes among the 205 students of School HS_1 was 4, which is four times that in the full sample. The (frequentist) L-index (Sinharay et al., 2017), which quantifies the difference between the estimated ability on the items with answer changes and items with no answer changes and approximately follows a standard normal distribution under no test fraud, was computed for each examinee. The EDI for each examinee was computed using a continuity correction of 0.5 as in Wollack et al. (2015). The weighted maximum likelihood estimate (WLE; Warm, 1989) was used as the estimate of the examinee ability.

Table 2 shows the percentages of p-values that were statistically significant at the 1% level for the two approaches among examinees who made at least one answer change on the assessment. For the L-index and the EDI, a value larger than 2.33 (that is the 99th percentile of the standard normal distribution) of the statistic was considered statistically significant. The first row of numbers in the table provide the percentages among the examinees who were not in School HS_1 (that is, in any school other than HS_1) and made at least one answer change. The second row of numbers provide the percentages among the 159 examinees in School HS_1 who made at least one answer change; thus, for example, the percentage of 12.6 for the L-index in the second row indicates that among the 159 examinees in School HS_1 who made at least one answer change, the index was statistically significant at the 1% level, that is, larger than 2.33, for 20 examinees (note that $\frac{100 \times 20}{159} \approx 12.6$).

The percentage of significant values for either examinee group is larger for the L-index than for the EDI, which is consistent with the finding of Sinharay et al. (2017) that the latter statistic is conservative. One important finding in Table 2 is that the percent significant for each approach is considerably larger among students from school HS_1 than among other students. For example, while the L-index was significant for only 1.4% examinees among other schools, the index was significant for about 12.6% of the students in School HS_1 . This result is favorable to the approaches in the sense that they seem to work as intended in practice by flagging many students from a school (HS_1) that appears suspicious from several sources of evidence.

Table 2 The percentages of examinees with at least one answer change and statistically significant p-values for the L-index and EDI.

Examinees	L-index	EDI
Not in HS_1	1.4	0.7
In HS_1	12.6	7.5

Example 3: person-fit analysis

The data set from the international credentialing examination, which was considered in Example 1, was used to perform person-fit analysis. Two parametric person-fit statistics— ζ_2 (Tatsuoka, 1984) and I_z^* (Snijders, 2001)—were used in the analysis.

The Rasch model is operationally used in the credentialing test; however, the two-parameter logistic model (2PLM) was found to fit the data better than the Rasch model—therefore the 2PLM was used in the analysis. The item parameters under the 2PLM were estimated from the data set and were used to compute the person-fit statistics. The weighted maximum likelihood estimate (WLE; Warm, 1989) was used as the estimate of the examinee ability. For ζ_2 and I_z^* , the critical values at the 1% significance level were 2.33 and -2.33, respectively, which are the 99th and first percentiles of the standard normal distribution.

Table 3 shows the percentages of examinees for which the p-values were statistically significant at the 1% level for the two person-fit statistics. The first row of numbers provides the percentages among the examinees that were not flagged by the credentialing organization. The second row of numbers provides the percentages among the 48 examinees who were flagged by the credentialing organization. Thus, for example, the percentage of 8.3 for I_z^* in the second row indicates that among the 48 examinees flagged by the credentialing organization, I_z^* was statistically significant at the 1% level, that is, smaller than -2.33, for four examinees (note that $4/48 = 0.083$).

Table 3 shows that the ζ_2 statistic led to fewer significant values among non-flagged examinees and more significant values among flagged examinees compared to I_z^* . The table also shows that the percent significant for each person-fit statistic is much larger among the examinees flagged by the credentialing organization compared to among non-flagged examinees—this result provides some evidence that the person-fit statistics provide useful information by being significant at a larger rate among the examinees who are truly aberrant for all practical purposes.

To obtain further insight regarding person fit, the 170 items in the data set were divided into 10 groups using their proportion correct values. Thus, the 17 items with 10 smallest proportion correct values formed item group 1, the 17 items with next 10 smallest proportion correct values formed item group 2, ..., the 17 items with the 10 largest proportion correct values formed item group 10. For each examinee, the proportion correct on the 10 item groups were computed. One would expect the proportion corrects on item groups 1 to 10 to be roughly in an increasing order for each examinee. The four panels in Fig. 1 show, for each of four examinees, the proportion corrects on the 10 groups of items. A loess-smoothed curve (drawn using the scatter smooth function in R; R Core Team, 2021) is also shown in each panel. The range of the vertical axis is the same for all four examinees. The values of ζ_2 and I_z^* were significant for Examinees 1–3 while they were not significant for Examinee 4. While the proportion corrects mostly are in increasing order for Examinee 4, they are not so for Examinees 1–3 (for Examinee 3, the proportion corrects are mostly increasing except a sharp drop for item group 6)—it is then quite natural that ζ_2 and I_z^* indicate misfit for the latter examinees.

Example 4: use of score differencing to detect item preknowledge

The credentialing data set described in Examples 1 and 3 can be used to detect item preknowledge because 61 items were identified as compromised by the credentialing organization. A frequentist and a Bayesian approach were used to detect item preknowledge. The 2PLM was used in all the analyses.

The frequentist approach comprised the use of the signed likelihood ratio statistic L_s (Sinharay, 2017) that quantifies the difference between the estimated ability on the compromised items and non-compromised items—substantial item preknowledge would cause the statistic to be a large positive number. The large-sample distribution of L_s is the standard normal distribution under the null hypothesis of no item preknowledge (Sinharay, 2017).

The Bayesian approach comprised the use of the predictive p-value (Wang et al., 2017) of the raw score on the compromised items. A Markov chain Monte Carlo algorithm was used to simulate draws from the posterior distribution of the examinee ability given the scores on the non-compromised items, posterior predictive data sets were generated using the draws from the algorithm, and then predictive p-values were computed for each examinee using the raw score on the compromised items as the test statistic. A standard normal prior distribution was used on the examinee ability.

Table 4 shows the percentages of examinees for which the p-values were statistically significant at the 1% level for the frequentist and Bayesian approaches. The two rows of numbers provide the percentages among the examinees who were not flagged and flagged, respectively, by the credentialing organization. While the two percentages are similar in each row of

Table 3 The percentages of examinees with statistically significant p-values for the person-fit statistics for the credentialing data.

Examinees	I_z^*	ζ_2
Not flagged	2.2	1.6
Flagged	8.3	10.4

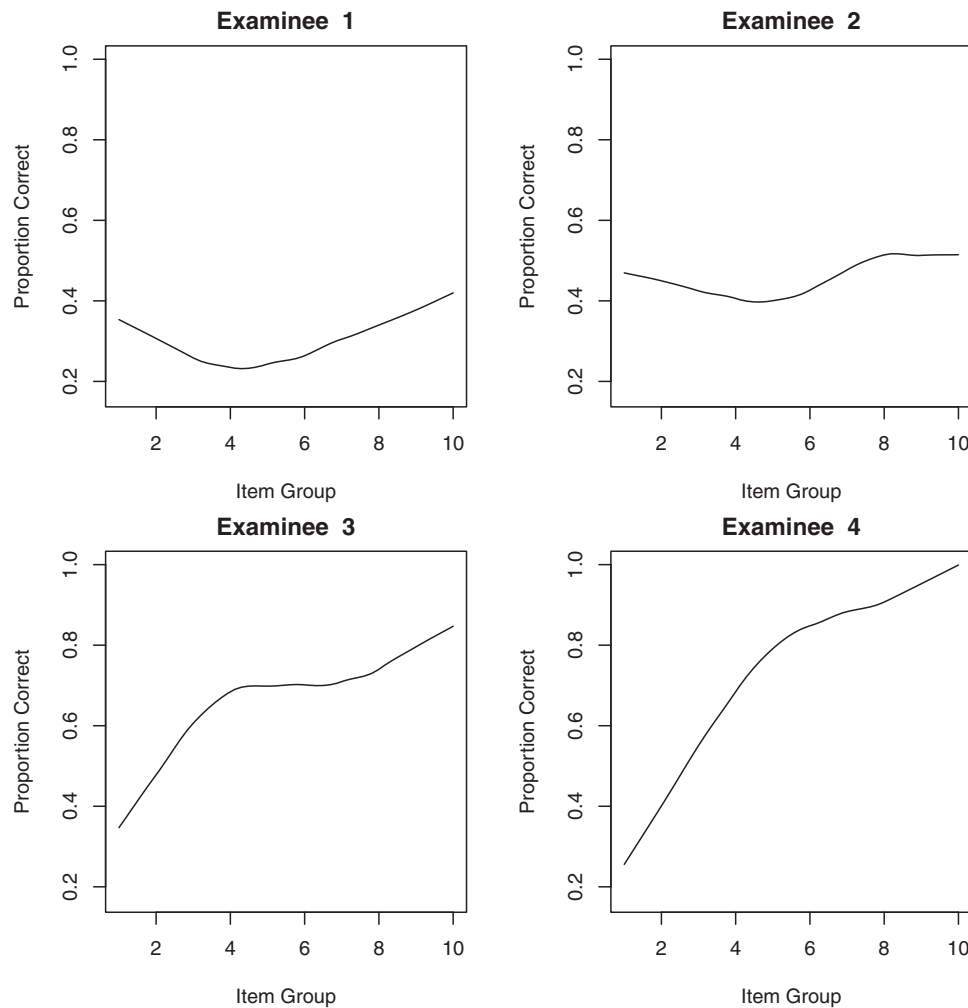


Fig. 1 Average proportion corrects on sorted item groups (and smooth curves) for four examinees.

Table 4 The percentages of examinees with significant p-values in the analysis to detect item preknowledge for the credentialing data.

Examinees	Frequentist	Bayesian
Not flagged	3.1	3.8
Flagged	18.8	21.2

the table, the Bayesian approach leads to slightly larger numbers of significant p-values for both group of examinees. The table also shows that the percent significant for each approach is much larger among the examinees flagged by the credentialing organization compared to among non-flagged examinees—this result provides some evidence that the approaches provide useful information by being significant at a larger rate among the examinees who truly were involved in test fraud for all practical purposes.

Fig. 2 shows the kernel-density estimates (computed and plotted using the function “density” in the R software [R Core Team \(2021\)](#)) of the distributions of the proportion corrects on the compromised and non-compromised items of the examinees who had statistically significant p-values at the 1% level for both the frequentist and Bayesian approaches and the remaining examinees. These two groups of examinees are referred to as “detected” and “non-detected,” respectively, in the legend of the figure. The figure indicates that the detected examinees performed much better on the compromised items than on the non-compromised items while the non-detected examinees performed slightly better on the non-compromised items than on the compromised items. Also, the performance of the detected examinees is much better than the non-detected examinees on the compromised items.

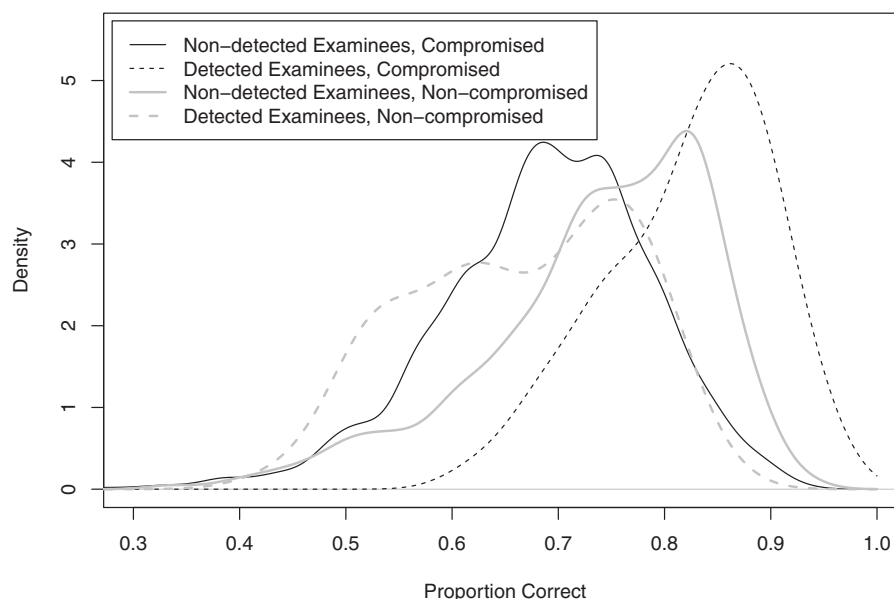


Fig. 2 Proportion corrects for the detected and non-detected examinees on the compromised and non-compromised items.

Conclusions and recommendations

A brief review of the five main statistical methods for detecting test fraud identified by [Wollack and Schoenig \(2018\)](#) is provided. For each of the five areas, some popular approaches are briefly described. Four real data examples are used to demonstrate the application of some of these approaches.

One feature of the real data examples is that the statistics (for detecting test fraud) are significant for only a fraction of the known cheaters. For example, for the data set from the international credentialing examination, 48 examinees were known to have been involved in test fraud and yet, only 15 (that is, about 31%) of them were associated with a statistically significant value of the ω index in the first example, about 8% of them had a significant value of the I_z^* statistic in the third example, and about 19% of them had a significant value of the L-index in the fourth example. There are two primary reasons of this apparent lack of effectiveness of the statistics. First, while it is known that 48 examinees were involved in test fraud, information on which examinee was involved in what type of fraud was unavailable; if such information were available, the statistics would most likely have been more effective. Second, almost all statistics for detecting test fraud suffer from modest power (e.g., [Ferrara, 2017](#)). Together with a small percentage of cheaters in a typical sample of examinees, the low power often leads to the number of false positives larger than that of true positives in an examinee sample (e.g., [Skorupski and Wainer, 2017](#)). The lack of high power often arises from the lack of data. For example, in the second example, erasure analysis for each examinee was based on 46 dichotomous item responses—larger number of items on the test would have led to considerably larger power. However, the fact that these statistics are significant at a substantially larger rate among the known cheaters than among the rest of the examinees provides evidence that these statistics are somewhat useful.

As the above review shows, there exist several approaches for detecting test fraud. However, there remain several unanswered questions on how these approaches should be used in practice. For example, how many approaches should be used and which ones for a specific data set? [Wollack and Cizek \(2017b, p. 397\)](#) warned against the use of multiple conceptually similar methods and commented that evidence of cheating from distinctly different indices can be quite compelling. They also recommended controlling for false positives in the context of test fraud where several statistical tests are typically performed on many examinees; one way to limit the number of false positives is to choose a critical value that adjusts for multiple comparisons by controlling the family-wise error rate (using, for example, a Bonferroni correction) or controlling the false discovery rate (using the procedure of [Benjamini and Hochberg, 1995](#)). [Buss and Novick \(1980\)](#) asserted that “statistical methods of detection should generally not be the sole basis for a judgment that an examinee cheated (or that an examinees scores are sufficiently questionable to justify non-reporting) in the absence of corroboration from other types of evidence.” Other researchers such as [Holland \(1996\)](#), [Hanson et al. \(1987\)](#) and [Tendeiro and Meijer \(2014\)](#) made similar recommendations. However, test-security experts such as [Wollack and Cizek \(2017a, p. 200\)](#) have recently presented the viewpoint that statistical evidence based on even a single statistic may constitute conclusive proof of cheating provided the statistic has been properly vetted and accepted by the research community and the degree of aberrance is clearly extreme.

References

- Allen, J., Ghattas, A., 2016. Estimating the probability of traditional copying, conditional on answer-copying statistics. *Appl. Psychol. Meas.* 40, 258–273.
- American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B* 57, 289–300.
- Bishop, S., Egan, K., 2017. Detecting erasures and unusual gain scores: understanding the status quo. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 193–213.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37, 29–51.
- Buss, W.G., Novick, M.R., 1980. The detection of cheating on standardized tests: statistical and legal analysis. *J. Law Educ.* 9, 1–64.
- Chalmers, R.P., 2012. mirt: a multidimensional item response theory package for the R environment. *J. Stat. Software* 48 (6), 1–29.
- Cizek, G.J., Wollack, J.A., 2017. *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC.
- de Ayala, R.J., 2009. *The Theory and Practice of Item Response Theory*. Guilford Press, New York, NY.
- De Boeck, P., Jeon, M., 2019. An overview of models for response times and processes in cognitive tests. *Front. Psychol.* 10.
- Drasgow, F., Levine, M.V., Williams, E.A., 1985. Appropriateness measurement with polychotomous item response models and standardized indices. *Br. J. Math. Stat. Psychol.* 38, 67–86.
- Drasgow, F., Levine, M.V., Zickar, M.J., 1996. Optimal identification of mismeasured individuals. *Appl. Meas. Educ.* 9, 47–64.
- Everitt, B.S., Landau, S., Leese, M., Stahl, D., 2011. *Cluster Analysis*. John Wiley & Sons, Ltd.
- Ferrara, S., 2017. A framework for policies and practices to improve test security programs: prevention, detection, investigation, and resolution (PDIR). *Educ. Meas.* 36 (3), 5–24.
- Fox, J.-P., Mariani, S., 2017. Person-fit statistics for joint models for accuracy and speed. *J. Educ. Meas.* 54, 243–262.
- Gelman, A., Meng, X., Stern, H.S., 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* 6, 733–807.
- Glas, C.A.W., Meijer, R.R., 2003. A Bayesian approach to person fit analysis in item response theory models. *Appl. Psychol. Meas.* 27, 217–233.
- Government Accountability Office, 2013. K-12 Education: States' Test Security Policies and Procedures Varied (GAO-13-495R) (Tech. Rep.). Author, Washington, DC.
- Haberman, S.J., Lee, Y.-H., 2017. A Statistical Procedure for Testing Unusually Frequent Exactly Matching Responses and Nearly Matching Responses (ETS Research Report No. RR-17-23). ETS, Princeton, NJ.
- Hambleton, R.K., Swaminathan, H., 1985. *Item Response Theory: Principles and Applications*. Kluwer Academic Publishers, Boston.
- Hanson, B.A., Harris, D.J., Brennan, R.L., 1987. A Comparison of Several Statistical Methods for Examining Allegations of Copying (ACT Research Report Series no. 87-15). American College Testing, Iowa City, IA.
- He, Q., Meadows, M., Black, B., 2020. An introduction to statistical techniques used for detecting anomaly in test results. *Res. Pap. Educ.* <https://doi.org/10.1080/02671522.2020.1812108>.
- Holland, P.W., 1996. Assessing Unusual Agreement Between the Incorrect Answers of Two Examinees Using the K-Index: Statistical Theory and Empirical Support (ETS Research Report No. RR-94-4). ETS, Princeton, NJ.
- Jacob, B., Levitt, S., 2003. Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Q. J. Econ.* 118, 843–877.
- Karabatsos, G., 2003. Comparing the aberrant response detection performance of thirty-six person-fit statistics. *Appl. Meas. Educ.* 16, 277–298.
- Kingston, N., Clark, A., 2014. *Test Fraud: Statistical Detection and Methodology*. Routledge, New York, NY.
- Kingston, N., 2013. Educator testing case studies. In: Wollack, J.A., Fremer, J.J. (Eds.), *Handbook of Test Security*. Routledge, New York, NY, pp. 299–311.
- Lewis, C., Thayer, D.T., 1998. The Power of the K-Index (or PMIR) to Detect Copying (ETS Research Report No. RR-98-49). ETS, Princeton, NJ.
- Lord, F.M., Wingersky, K., 2001. Comparison of IRT true-score and equipercentile observed-score "equatings". *Appl. Psychol. Meas.* 8, 453–461.
- Man, K., Harring, J.R., Sinharay, S., 2019. Use of data mining methods to detect test fraud. *J. Educ. Meas.* 56, 251–279.
- Mariani, S., Fox, J.-P., Aветisyan, M., Veldkamp, B.P., Tijmstra, J., 2014. Testing for aberrant behavior in response time modeling. *J. Educ. Behav. Stat.* 39, 426–451.
- Maynes, D., 2013. Educator cheating and the statistical detection of group-based test security threats. In: Wollack, J.A., Fremer, J.J. (Eds.), *Handbook of Test Security*. Routledge, New York, NY, pp. 173–199.
- Maynes, D., 2014. Detection of non-independent test-taking by similarity analysis. In: Kingston, N.M., Clark, A.K. (Eds.), *Test Fraud: Statistical Detection and Methodology*. Routledge, New York, NY, pp. 53–82.
- McLeod, L.D., Lewis, C., Thissen, D., 2003. A Bayesian method for the detection of item preknowledge in computerized adaptive testing. *Appl. Psychol. Meas.* 27, 121–137.
- Meijer, R.R., Sijtsma, K., 2001. Methodology review: evaluating person fit. *Appl. Psychol. Meas.* 25, 107–135.
- NCME, 2012. *Testing and Data Integrity in the Administration of Statewide Student Assessment Programs* (Tech. Rep.). Author, Madison, WI.
- Peng, L., Sinharay, S., 2020. The use of theory of linear mixed-effects models to detect fraudulent erasures at an aggregate level. *Educ. Psychol. Meas.* <https://doi.org/10.1177/0013164421994893>.
- Primoli, V., Liassou, D., Bishop, N.S., Nhuyvanisvong, A., 2011. Erasure descriptive statistics and covariates. In: Paper Presented at the Annual Meeting of the National Council of Measurement in Education, New Orleans, LA.
- R Core Team, 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.
- Segall, D.O., 2002. An item response model for characterizing test compromise. *J. Educ. Behav. Stat.* 27, 163–179.
- Sinharay, S., Jensen, J.L., 2019. Higher-order asymptotics and its application to testing the equality of the examinee ability over two sets of items. *Psychometrika* 84, 484–510.
- Sinharay, S., Johnson, M.S., 2017. Three new methods for analysis of answer changes. *Educ. Psychol. Meas.* 77, 54–81.
- Sinharay, S., Johnson, M.S., 2020. The use of item scores and response times to detect examinees who may have benefited from item preknowledge. *Br. J. Math. Stat. Psychol.* 73, 397–419.
- Sinharay, S., Duong, M.Q., Wood, S.W., 2017. A new statistic for detection of aberrant answer changes. *J. Educ. Meas.* 54, 200–217.
- Sinharay, S., 2015. Assessing person fit using I_2^* and the posterior predictive model checking method for dichotomous item response theory models. *Int. J. Quant. Res. Educ.* 2, 265–284.
- Sinharay, S., 2016a. Assessment of person fit using resampling-based approaches. *J. Educ. Meas.* 53, 63–85.
- Sinharay, S., 2016b. Iztarmix: assessment of person fit for mixed-format tests. *Appl. Psychol. Meas.* 40, 76–77.
- Sinharay, S., 2016c. Asymptotically correct standardization of person-fit statistics beyond dichotomous items. *Psychometrika* 81, 992–1013.
- Sinharay, S., 2017. Detection of item preknowledge using likelihood ratio test and score test. *J. Educ. Behav. Stat.* 42, 46–68.
- Sinharay, S., 2018a. Application of Bayesian methods for detecting fraudulent behavior on tests. *Measurement* 16, 100–113.
- Sinharay, S., 2018b. Detecting fraudulent erasures at an aggregate level. *J. Educ. Behav. Stat.* 43, 286–315.
- Sinharay, S., 2018c. A new person-fit statistic for the lognormal model for response times. *J. Educ. Meas.* 55, 457–476.
- Skorupski, W.P., Wainer, H., 2017. The case for Bayesian methods when investigating test fraud. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 214–231.
- Smith, R.M., 1986. Person fit in the Rasch model. *Educ. Psychol. Meas.* 46, 359–372.
- Snijders, T., 2001. Asymptotic distribution of person-fit statistics with estimated person parameter. *Psychometrika* 66, 331–342.
- Tatsuoka, K.K., 1984. Caution indices based on item response theory. *Psychometrika* 49, 95–110.

- Tendeiro, J.N., Meijer, R.R., 2014. Detection of invalid test scores: the usefulness of simple nonparametric statistics. *J. Educ. Meas.* 51, 239–259.
- Tendeiro, J.N., Meijer, R.R., Niessen, A.S.M., 2016. PerFit: an R package for person-fit analysis in IRT. *J. Stat. Software* 74 (5).
- Toton, S.L., Maynes, D.D., 2018. The use of data mining techniques to detect cheating. In: Jiao, H., Lissitz, R.W., Van Wie, A. (Eds.), *Data Analytics and Psychometrics: Informing Assessment Practices*. Information Age Publishing, New York, NY.
- van der Linden, W.J., Guo, F., 2008. Bayesian procedures for identifying aberrant response-time patterns in adaptive testing. *Psychometrika* 73, 365–384.
- van der Linden, W.J., Lewis, C., 2015. Bayesian checks on cheating on tests. *Psychometrika* 80, 689–706.
- van der Linden, W.J., Sotaridona, L., 2006. Detecting answer copying when the regular response process follows a known response model. *J. Educ. Behav. Stat.* 31, 283–304.
- van der Linden, W.J., van Krimpen-Stoop, E.M.L.A., 2003. Using response times to detect aberrant responses in computerized adaptive testing. *Psychometrika* 68, 251–265.
- van der Linden, W.J., 2006. A lognormal model for response times on test items. *J. Educ. Behav. Stat.* 31, 181–204.
- van der Linden, W.J., 2007. A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika* 72, 287–308.
- van der Linden, W.J., 2009. Conceptual issues in response-time modeling. *J. Educ. Meas.* 46, 247–272.
- van der Linden, W.J., 2016. Lognormal response-time model. In: van der Linden, W. (Ed.), *Handbook of Item Response Theory, Models*, vol. 1. Chapman and Hall/CRC, Boca Raton, FL.
- Wang, C., Xu, G., Shang, Z., 2016. A two-stage approach to differentiating normal and aberrant behavior in computer based testing. *Psychometrika*. <https://doi.org/10.1007/s11336-016-9525-x>.
- Wang, X., Liu, Y., Hambleton, R.K., 2017. Detecting item preknowledge using a predictive checking method. *Appl. Psychol. Meas.* 41, 243–263.
- Warm, T.A., 1989. Weighted likelihood estimation of ability in item response theory. *Psychometrika* 54, 427–450.
- Wollack, J.A., Cizek, G.J., 2017a. Test security for licensure and certification examination programs. In: Davis-Becker, S., Buckendahl, C. (Eds.), *Testing in the Professions*. Routledge/Taylor & Francis, New York, NY, pp. 178–209.
- Wollack, J.A., Cizek, G.J., 2017b. The future of quantitative methods for detecting cheating. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 390–399.
- Wollack, J.A., Eckerly, C., 2017. Detecting test tampering at the group level. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 214–231.
- Wollack, J.A., Fremer, J.J., 2013. *Handbook of Test Security*. Routledge, New York, NY.
- Wollack, J.A., Maynes, D., 2017. Detection of test collusion using cluster analysis. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 124–150.
- Wollack, J.A., Schoenig, R.W., 2018. Cheating. In: Frey, B.B. (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Sage, Thousand Oaks, CA, pp. 260–265.
- Wollack, J.A., Cohen, A.S., Eckerly, C.A., 2015. Detecting test tampering using item response theory. *Educ. Psychol. Meas.* 75, 931–953.
- Wollack, J.A., 1997. A nominal response model approach for detecting answer copying. *Appl. Psychol. Meas.* 21, 307–320.
- Zopluoglu, C., 2017. Similarity, answer copying, and aberrance: understanding the status quo. In: Cizek, G.J., Wollack, J.A. (Eds.), *Handbook of Detecting Cheating on Tests*. Routledge, Washington, DC, pp. 25–46.
- Zopluoglu, C., 2018. Copydetect: Computing Response-Similarity Indices for Multiple-Choice Tests (R package version 1.3).
- Zopluoglu, C., 2019. Detecting examinees with item preknowledge in large-scale testing using extreme gradient boosting (XGBoost). *Educ. Psychol. Meas.* 79, 931–961.

Rater effects modeling

Jodi M. Casabianca, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Background on rater effects in constructed response scores	308
Common rater metrics and rater effects	308
Rater effects modeling	309
Observed score rater effects indices	309
IRT model-based rater effects indices	310
Models with rater parameters	310
Models without rater parameters	312
Empirical example: monitoring raters in a teacher licensure performance assessment	312
Summary of the FACT pilot data	312
IRT rater model results compared to observed score metrics	312
Identifying aberrant raters	318
Factors in model selection	318
Summary	319
Acknowledgments	319
References	319

Background on rater effects in constructed response scores

Tasks that measure complex skills and constructs such as critical thinking, collaboration, or technological literacy, typically cannot be scored by a simple machine like selected-response items—they require human raters to score them. Human raters are trained and instructed to follow a scoring rubric or scoring guide, however many factors may lead to raters' deviation from the rubrics, thereby yielding scores on tasks that do not solely reflect the test takers' knowledge, skills, and abilities. Instead, scores and their corresponding consequences such as decisions about individuals and groups, are impacted by rater biases and errors. These errors contribute to the scores' construct irrelevant variance which in turn degrades the test takers' scores and decisions made about them. Rater effects are patterns or systematic variance in scores resulting from common rater errors. Part of upholding professional standards for testing is explicitly addressing the issue of rater errors in order to minimize them (e.g., see Section 2.3.3 of [International Test Commission, 2012](#); [American Educational Research Association, the American Psychological Association and the National Council on Measurement in Education, 2014](#)).

In order to reduce these errors from occurring in the first place, the best practice is to take measures during the test and item development process to generate items with clearly defined rubrics that will minimize construct irrelevant variance ([ETS, 2021](#)). In addition, rater selection and training should also be carefully performed, with consideration of the types of raters and the most effective training for accurate scoring. Unfortunately, despite these countermeasures, raters still make errors, as they are human.

The literature on rater effects and the impacts of them is vast and well-developed (see [Saal et al., 1980](#); [Wolfe and McVay, 2012](#)). In general, estimates of rater effects can be used in many ways. First, estimates can be used to statistically adjust scores by using modeling techniques that will account for rater effects. There is evidence to show that ignoring rater effects, especially in certain assessment situations, can have a detrimental impact on scores thereby reducing their reliability and validity of their interpretation and uses ([Casabianca and Wolfe, 2017](#); [Guo, 2014](#); [Hombo et al., 2001](#)). It is also possible that rater effects do not have a detrimental impact as they may “cancel each other out” in certain rating designs ([Hombo et al., 2001](#)). Second, rater effects estimates can be used to monitor raters and/or give them feedback to improve their use of the score scale. This chapter describes the common rater effects, metrics and modeling techniques that can be used to detect rater effects, and some factors in model selection.

Common rater metrics and rater effects

A major focus of quality control monitoring in most CR scoring systems is the estimation of interrater reliability or consistency. The ability for multiple raters to use the same scoring criteria and yield consistent scores on the same responses shows the extent to which training procedures, the rubric, and other scoring materials provide the necessary foundation for raters to derive the same evaluation of test taker responses. Low levels of rater consistency reveal weaknesses in this foundation. However, just because interrater reliability is high, does not necessarily mean that the scores they yield are technically accurate. Ratets can be in agreement on incorrect scoring decisions. They may have similarly misunderstood the rubric, for example.

For this reason, another traditional focus of quality control monitoring is on rater accuracy (or inaccuracy). The notion of rater accuracy is complex, or “multi-componential” as per [Cronbach \(1955\)](#). He decomposed accuracy into four parts: (i) elevation, or overall accuracy, (ii) differential elevation, or discrimination among test takers, (iii) stereotype accuracy, or discrimination among traits, and (4) differential accuracy, or discrimination among test takers within traits. Different measures of rater accuracy that are linked to specific components may not be correlated. The more common, or traditional, measures of rater inaccuracy capture the extent to which a rater assigns scores inconsistent with the true score (or approximations of the true score), which is the test taker’s score if there was no measurement error. Rater accuracy measures may be criterion-referenced, when the rater is compared to a score assigned by an expert, or norm-referenced, when the rater is compared to a score based on a consensus from many raters scoring the response. It is very common in practice to use criterion-referenced measures of rater accuracy to monitor raters. To do this, a set of responses each are scored by an expert to produce a set of “validity responses.” The validity responses are then embedded into the queue of responses scored by each rater and the raters’ scores on the validity responses are compared to the experts’ scores to evaluate the raters’ accuracy. An example of a norm-referenced approach is the correlation between a rater’s scores and the average score across raters, denoted by $(r_{x, \bar{x}})$.

Rater effects manifest as systematic variance in scores, however, rater inaccuracy is often discussed as a rater effect even though it is not necessarily systematic. As Wolfe puts it, rater inaccuracy is the extent to which “raters assign scores that deviate from true scores in a random, unpredictable manner” (2014, p. 8). Rater accuracy can be captured with different metrics depending on the data available and these different metrics may be capturing different components of rater accuracy. Despite established complexities in measuring components of rater (in)accuracy, it is mostly basic agreement statistics that are used to measure it. These measures include the percent exact agreement, correlations, kappa, or quadratically weighted kappa (QWK) (see [Burry-Stock, 1996](#)) for an accessible description of how these measures are computed and interpreted).

Taken together, these traditional measures of interrater reliability and rater accuracy can provide a picture of how consistently raters use the scoring rubric and score scale and the accuracy of individual raters and the rater pool overall. However, other measures may be able to better provide more detailed feedback to raters to address specific issues with their use of the rubric. Rater effects are patterns in a rater’s scoring behavior and are not isolated when using these traditional metrics. Some common rater effects include:

- **Severity/Leniency:** One of the most commonly discussed rater effects is the error of rater severity. Severity occurs when a rater tends to score harshly, or lower than they should. Leniency occurs when a rater assigns a higher score than they should. This is sometimes referred to as rater bias, as the error can be in either direction. Differential severity/leniency errors occur when a rater is severe/lenient in their scoring behavior with a particular subgroup, but on average, does not exhibit that behavior with other subgroups.
- **Centrality/Extremity:** While not the most commonly discussed, centrality is a rater error that may be the most common error made in practice. For any number of reasons, a rater may rely heavily on the middle of the score scale. This scoring behavior is akin to “playing it safe” whereas the opposite use of the score scale leads to extremity. Extreme raters essentially use the score scale as a dichotomy, giving either the highest or lowest possible score.
- **Restriction of Range:** Restriction of range occurs when a rater clusters his or her ratings around an area of the score scale, not necessarily the middle. It has also been defined as the “extent to which obtained ratings discriminate among different ratees in terms of their respective performance levels” ([Saal et al., 1980](#), p. 417). There has been some confusion in the literature about how to distinguish centrality from restriction of range. Centrality can be thought of as a special case of restriction of range in which the cluster of ratings is in the middle of the score scale. Restriction of range may occur if a rater is severe, lenient, and/or central in their scoring.
- **Halo Effects:** Halo effects manifest as correlated errors in the ratings which occur when a rater assigns multiple ratings to a test taker’s response(s), whether they are ratings for responses to multiple prompts, or multiple ratings of the same response using different domains. The errors are correlated because information from one response affects the rater’s rating of other responses. For example, if a rater is assigned to rate three responses from the same test taker and the rater is more lenient on rating the second and third responses as a result of a positive feature they observed on the first response.

Importantly, when raters are scoring over a duration of time, they may drift in the scoring behavior ([Casabianca et al., 2015](#); [Condon and McQueen, 2000](#); [Leckie and Baird, 2011](#); [Wilson and Case, 2000](#)). This is not a constant rater effect as the other rater effects described above, but drift is a contextual concern and rater monitoring is an important step to maintain the quality of scores.

For a thorough treatment of rater effects and how they manifest in ratings, refer to [Wolfe and McVay \(2012\)](#), [Myford and Wolfe \(2003\)](#), and/or [Wolfe \(2014\)](#).

Rater effects modeling

Observed score rater effects indices

The term “observed score” rater effects indices refers to metrics that quantify rater effects based on the raw scores or ratings that were observed or collected. This is converse to model-based indices, such as the item response theory (IRT) metrics later discussed in the section on [IRT model-based rater effects indices](#).

Performing an analysis of rater score distributions can reveal much about how a rater uses the score scale. For example, a rater who assigns 75% 5s on a 1–5 scale may be labeled as a lenient rater. This is an analysis that is easy to understand and explain, but only feasible for a small rater pool and relies on assumptions such as a random sample of test takers and also a random assignment of raters to responses. When there is a large number of raters, it is more convenient to rely on basic summary statistics that have been shown to correlate with specific rater behaviors (Wolfe, 2014). This also assumes random assignment of responses and large numbers of responses used in the calculation. Rater severity is captured by the mean of individual rater's ratings ($\bar{X}_r$)—raters with means higher than the overall mean tend to be lenient, and raters with means lower than the overall average tend to be severe. Rater centrality can be captured by the standard deviation (SD) of individual rater's ratings (SD_r). Raters with relatively large SDs might be extreme in their scoring and use the ends of the score scale. Raters with relatively small SDs might be central in their scoring and use the middle categories. Rater SDs can also capture range restriction in general, and thus it is also important to consider the distance of the rater's mean to the middle of the score scale. As mentioned above, observed score metrics for rater inaccuracy can include agreement measures based on a criterion- or norm-referenced approach.

IRT model-based rater effects indices

There are limitations to the information observed score metrics can provide because of the differences in the other facets of the measurement process that go unadjusted for in basic summary statistics. A statistical model can make adjustments for item effects, person effects, and more, if necessary. There are many modeling frameworks within which to examine rater errors. Generalizability theory, in particular, is an approach to decomposing the variance in scores attributable to various facets (Brennan, 1992). However, another natural modeling framework to address this in the constructed-response scoring scenario is IRT (Lord, 1980).

IRT models are probabilistic models used to scale item response data. Readers new to IRT are encouraged to learn more in the chapter on IRT in this Encyclopedia and in other introductory sources (see the ITEMS module: Penfield, 2014). In a common set of IRT models called the 2-parameter logistic IRT models, items are characterized by a set of parameters that describe item difficulty and item discrimination, or an item's ability to distinguish test takers (Lord, 1980). An even simpler model is the Rasch model or the 1-parameter logistic IRT model, which only includes the item difficulty parameters (Rasch, 1960). In all IRT model formulations, test takers are characterized by a latent ability parameter θ_i , which reflects their standing on the construct being measured by the test on a scale ranging from $-\infty$ to $+\infty$ (commonly specified as a standard normal distribution, centered at 0 with standard deviation of 1.0). For polytomous items, scored on a discrete scale with more than 2 points, there are specific IRT models used to appropriately capture response patterns and determine the locations of the boundaries for the score levels along the scale of the underlying latent ability. For example, the generalized partial credit model (GPCM; Muraki, 1992) given by

$$\pi_{ijx} = \frac{\exp \left[\sum_{k=1}^x \alpha_j (\theta_i - \beta_j - \gamma_{jk}) \right]}{\sum_{c=0}^{m_j} \exp \left[\sum_{k=1}^c \alpha_j (\theta_i - \beta_j - \gamma_{jk}) \right]}, x = 0, 1, \dots, m_j \quad (1)$$

expresses the probability π_{ijx} that a test taker i ($i = 1, \dots, N$) will achieve a score x on their response as a function of four types of parameters: latent ability (θ_i), item discrimination (α_j), item location/difficulty (β_j), and the item category threshold (γ_{jk}), where k refers to the threshold between score category x and the adjacent lower category $x-1$, and m_j refers to the number of possible categories for item j . Note that Model 1 implies that:

$$\ln \left[\frac{\pi_{ijk}}{\pi_{ij(k-1)}} \right] = \theta_i - \beta_j - \gamma_{jk}. \text{ The partial credit model (PCM; Masters, 1982) is a polytomous Rasch model and is a special case of the}$$

GPCM in which the item discrimination is equal to 1.0 for all items. IRT rater models are based on these same principles, but they are used to measure and describe rater behaviors.

Models with rater parameters

Models incorporating rater parameters are perhaps the most widely used when the analysis goal is to account for rater effects when scoring test takers. The multi-faceted Rasch model (MFRM; Linacre, 1989) is probably the most popular because of its relative accessibility and ease of interpretation. The traditional Rasch model as discussed earlier, includes a difficulty parameter for the item. The MFRM is an expanded version that also includes parameters for other facets, including raters. Note that the MFRM is recognized as a special case of Fischer's linear logistic trait model (1973). The MFRM based on the PCM (Masters, 1982) is given by

$$\ln \left[\frac{\pi_{ijrk}}{\pi_{ijr(k-1)}} \right] = \theta_i - \beta_j - \gamma_{jk} - \phi_r \quad (2)$$

where θ_i is the latent trait for test taker i , β_j is the location or difficulty for item j , γ_{jk} is the k th threshold for item j , and ϕ_r is the severity for rater r . Both β_j and ϕ_r are quantities on the logit scale (centered at 0) that affect the probability of observing a score

category. Specifically, when $\phi_r > 0$ the probability will be lower as the quantity is subtracted. This means that the rater tends to score low relative to the average rater ($\bar{\phi} = 0$) and so the rater is (relatively) severe. The opposite is true when $\phi_r < 0$; in this case, the rater is lenient relative to the average rater. Much research on rater effects and IRT modeling is based on the MFRM and how it can be used to detect rater effects (Linacre, 1989; Myford and Wolfe, 2003, 2004). Expanding the MFRM is relatively straightforward and therefore it offers a lot of flexibility to those with complex designs and needs—there are numerous other variants of the MFRM that capture more than just rater severity (Eckes, 2009; Myford and Wolfe, 2003).

The model in (2) is sometimes referred to as an item-related three-facet PCM since the thresholds term γ_{jk} is allowed to vary by item j (Eckes, 2009). A simpler version models one set of thresholds γ_k for all items. Yet another variant is the rater-related three-facet PCM which models γ_{rk} , or rater-specific thresholds representing the difficulty of score category k relative to rater r , removing the item-specific thresholds. In this case the rating scale is allowed to vary by rater, but applies to all items. The SD of these rater-specific thresholds is a measure of rater centrality in that when most of the ratings are in the center of the score scale, the rater will have a wider dispersion of thresholds yielding a large SD. In the item- and rater-related three-facet PCM, γ_{jrk} measures the rating scale for each rater and item. All of these variants are based on facets as main effects, but there are variants that utilize interaction effects to model other aspects of scoring, including interactions between raters and test takers (Eckes, 2009; Myford and Wolfe, 2003). Further, though the MFRM does not explicitly capture rater accuracy with a parameter, other measures are used to approximate it (Linacre, 2004; Wind and Engelhard, 2012; Wind et al., 2016; Wright and Masters, 1982).

Importantly, the MFRM and its variations treat multiple ratings of the same item response as “pseudoitems” or “virtual items” (Wang et al., 2014) in which rater-item combinations are treated as separate items. IRT models assume that item responses are independent conditional on the latent trait, or they assume *local independence* given the latent trait. However, it is unreasonable to make that assumption in these models since it is unlikely that ratings assigned to the same response will be independent. Treating ratings in this way ignores the dependencies that exist in the ratings of the same responses given by different raters. There are ramifications to this, which are discussed in **Factors in model selection** section.

The class of hierarchical rater models incorporate the hierarchical structure of the data that exists when there are multiple ratings of the same response thereby accounting for the local dependency that the MFRM ignores (DeCarlo et al., 2011; Patz et al., 2002). Depending on the model specification, these may incorporate parameters for rater severity, centrality, and/or intra-rater reliability or consistency, among other rater effects. In the original hierarchical rater model (HRM; Patz et al., 2002), the first level of the hierarchy uses a signal detection approach to model the distribution of ratings given the “ideal” rating which is the rating free of bias or unreliability, the second level models the distribution of a test taker’s ideal ratings given their latent trait θ_i , and the third level models the distribution of the latent trait θ_i .¹ The hierarchical representation of the HRM is given by

$$\begin{aligned} X_{ijr} | \xi_{ij}, \tau_r^2, \phi_r &\sim \text{polytomous signal detection model, } r = 1, \dots, R, \text{ for each } i, j. \\ \xi_{ij} | \theta_i, \beta_j, \gamma_{jk} &\sim \text{GPCM as in (1), } j = 1, \dots, J, \text{ for each } i \\ \theta_i &\sim N(\mu_0, \sigma_0^2), i = 1, \dots, N \end{aligned} \quad (3)$$

Here, θ_i , the latent trait for test taker i ($i = 1, \dots, N$) is normally distributed with mean μ_0 and σ_0^2 . This is a typical IRT specification for the latent trait, and depending on the model constraints, often $\mu_0 = 0$ and $\sigma_0^2 = 1.0$. The ξ_{ij} is the discrete ideal rating for test taker i on item j ($j = 1, \dots, J$) taking on values 1 to m_j , modeled using the GPCM as given in (1) but with the observed score replaced by the ideal ratings, which are latent variables. Finally, X_{ijr} is the observed rating given by rater r for test taker i ’s response to item j . In the original HRM, the signal detection model for rater bias (severity or leniency) ϕ_r and rater variability τ_r^2 is:

$$P[X_{ijr} = k | \xi_{ij} = \xi] \propto \exp \left\{ -\frac{1}{2\tau_r^2} [k - (\xi + \phi_r)]^2 \right\}, \quad (4)$$

where for each set of rating probabilities conditional on $\xi_{ij} = \xi$, the probability that the observed rating equals k is approximately equal to normal density with mean $\xi + \phi_r$ and variance τ_r^2 (Patz et al., 2002). The bias parameter is a continuous measure of bias on the scale of the raw score scale. When $\phi_r > 0$, rater r has positive bias or is lenient, and when $\phi_r < 0$, rater r has negative bias, or is severe. The variability parameter is often analyzed in its reciprocal form as a rater precision parameter, capturing the consistency with which a rater scores around the ideal rating offset by their individual bias or ϕ_r (Casabianca and Wolfe, 2017). The model is typically estimated as a Bayesian model with Markov chain Monte Carlo (MCMC) estimation. The model in (3) assumes prior distributions for unknown parameters including the precision of the latent traits $\omega = 1/\sigma_0^2$, the difficulty parameters β_j and step parameters of the IRT model γ_{jk} , and the rater parameters, $1/\tau^2$ and ϕ . Due to the model’s modularity, the basic HRM as discussed in Patz et al. (2002) has been expanded into a framework of HRMs with various special cases. For example, the signal detection layer of the model can be specified in different ways to incorporate more rater effects and idiosyncratic scoring behaviors than just bias and consistency (DeCarlo et al., 2011). In addition, other levels of the model can be adapted for longitudinal data designs (Casabianca et al., 2017) as well as multidimensional assessments (Nieto and Casabianca, 2019). For more information on full HRM framework, see Casabianca (2021).

Other IRT rater models such as the rater bundle model (RBM; Wilson and Hoskens, 2001) and the model for multiple ratings (MMR; Verhelst and Verstralen, 2001) were introduced around the same time as the HRM to address the same problem,

¹See DeCarlo (2005) for a primer on signal detection theory applied to the ratings context.

and more recently the MFRM was extended to address local dependence among multiple ratings for the same response (Wang et al., 2014).

Models without rater parameters

IRT models *without* rater parameters may be fit to ratings data with different assumptions. First, there may be the case that rater effects are completely ignored and a polytomous IRT model is used to scale the data, ignoring rater differences. This becomes complicated when there are multiple ratings for responses, in which case the multiple ratings may be aggregated somehow (e.g., averaged and then rounded), or the multiple ratings may be included as multiple item responses (or multiple items). Alternatively, IRT models without rater parameters may be used purposefully to analyze rater effects in the single-item context. In this case, in a rater response model, there are N test takers responding to one item, and the model treats raters as items. In this approach the model estimates the traditional item parameters (discrimination, difficulty, and thresholds) but in this context they take on a different interpretation as measures of rater effects. For example, in the rater response GPCM (RR-GPCM), which is an extension of the basic GPCM, the difficulty parameter is a measure of rater severity, the discrimination parameter is a measure of rater accuracy, and the SD of the thresholds is a measure of rater centrality (Stafford et al., 2018; Wolfe and McVay, 2012; Wolfe, 2014).

The RR-GPCM given by

$$\pi_{irx} = \frac{\exp\left[\sum_{k=1}^x \alpha_r(\theta_i - \phi_r - \gamma_{rk})\right]}{\sum_{c=0}^m \exp\left[\sum_{k=1}^c \alpha_r(\theta_i - \phi_r - \gamma_{rk})\right]}, x = 0, 1, \dots, m \quad (5)$$

is the same formulation as in a traditional GPCM in (1) except for the item parameters are now indexed by r instead of j because they are associated with raters. That is, π_{irx} is the probability that a particular rater r ($r = 1, \dots, R$) will assign rating x to test taker i ($i = 1 \dots N$) as a function of test taker location (θ_i), rater accuracy (α_r), rater location (ϕ_r), and the rating scale category threshold (γ_{rk}) of rater r .² Rater accuracy in this model is a norm-referenced indicator describing the rater's ability to discriminate test takers. Fitting this model to one item at a time may not be optimal if there is more than one CR item on the test, but it can be a valuable model to explore if the goal is to analyze and flag aberrant raters.³

Other polytomous IRT models can also be used in a similar fashion (Andersen, 1997; Andrich, 1978; Rasch, 1961; Samejima, 1969). Even dichotomous IRT models may be useful for ratings data (Engelhard, 1996). For example, a Rasch model fit to validity agreement data (agree = 1, disagree = 0) can be useful for capturing rater accuracy, or scaling raters. See Chin et al. (2014) for an example on how this can be applied in practice.

Empirical example: monitoring raters in a teacher licensure performance assessment

Understanding how best to measure teachers and teaching effectiveness with standardized tools has been a focus in recent educational research (Hill et al., 2008; Kane et al., 2013). Assessments for teacher licensure have mostly been focused on content knowledge, however, the *Foundational Assessment of Competencies for Teaching* (FACT) is an online performance assessment for teacher licensure that assesses not only teachers' content knowledge, but teaching skills (Phelps et al., 2020). In this section, FACT pilot study data demonstrate and compare rater effects indices from different modeling approaches.

Summary of the FACT pilot data

The FACT was pilot tested using 57 teacher candidates, each of whom were rated on 14 tasks (7 Math, 7 Reading & Language Arts) by 14 raters. All tasks were scored on a 3-point scale. In total there were 617 scorable performances scored by 3 randomly selected raters resulting in 1851 ratings. In addition to scoring these responses, each rater scored 74 validity responses for purposes of rater evaluation. Scoring accuracy was measured with validity agreement rates (see Table 1). Percent exact agreement ranged across raters from 0.51 to 0.68. Considering the 3-point scale, this level of agreement is suboptimal, however, these data were collected as part of an experimental formative training exercise and not meant to contribute to scores for teachers.

IRT rater model results compared to observed score metrics

An analysis using the MFRM, HRM, and RR-GPCM (as specified in Eqs. (2), (3) and (5)) demonstrates the IRT modeling approach, comparing it to the observed score metrics of rater effects and the accuracy estimates based on validity responses.^{4,5} As specified, all three IRT models provide an estimate of rater severity, the RR-GPCM and HRM provide measures of rater accuracy, and the RR-

²Note the Greek letter used for the item location has changed from (1) to be consistent with the severity parameters of other models.

³Though the GPCM is discussed here, see Masters and Wright (1997) for an example demonstrating the PCM to calibrate raters.

⁴The RR-GPCM was fit to the FACT pilot data using maximum likelihood estimation. The HRM and MFRM were fit to the data using MCMC estimation.

⁵For convenience the RR-GPCM analysis uses the full dataset (ignoring any item differences) since the score characteristics for the 14 items were very similar (Phelps et al., 2020). In practice, it may be beneficial to fit the RR-GPCM for each item to obtain a series of estimates for each rater.

Table 1 Rater agreement rates on validity responses for the FACT pilot data.

Rater	Validity agreement rates		
	Exact	Adjacent	Discrepant
03	0.68	0.32	0.00
11	0.68	0.31	0.01
06	0.66	0.32	0.01
09	0.66	0.32	0.01
12	0.66	0.32	0.01
05	0.62	0.35	0.03
07	0.61	0.39	0.00
13	0.61	0.36	0.03
04	0.59	0.38	0.03
02	0.58	0.41	0.01
08	0.58	0.39	0.03
10	0.57	0.41	0.03
01	0.51	0.43	0.05
14	0.51	0.43	0.05

Note. Raters are sorted descending by exact agreement rate.

GPCM also provides a measure of rater centrality.⁶ These measures may be compared to their observed score counterparts (mean, SD, and the correlation between raters' scores and the average of the triple scores). Table 2 provides the rater effects indices for all models, grouped by rater effect to facilitate comparisons across models.

With respect to severity, low rater mean scores reflect severity while high mean scores reflect leniency. Both the RR-GPCM and MFRM estimates are centered at 0 in this analysis, thus estimates for severe raters will be relatively more positive than other raters, and estimates for lenient raters will fall in the other direction—they will be negative. The HRM's signal detection model puts rater bias on the raw score scale, so the lenient raters will receive positive bias estimates in the HRM and severe raters will have negative bias estimates.

Table 3 provides the correlation matrix for all rater effects indices and Fig. 1 is a scatterplot matrix of the severity estimates, which shows direct linear relationships between all measures. The strongest relationships were between the severity estimates from the HRM and the GPCM ($r = -0.976$) and the HRM and the MFRM ($r = -0.973$). The high correlation between the severity estimates

Table 2 Rater effects indices and flagged raters for the FACT pilot data.

Rater ID	Severity				Centrality		Accuracy			
	RR-GPCM			SD	RR-GPCM		% Exact agree	Corr with mean	RR-GPCM discrim.	HRM rater SD
	Rater mean location	MFRM severity	HRM bias		SD	SD(thres)				
01	2.17	-0.22	-0.26	0.26	0.74	0.92	0.51	0.85	2.34	0.58
02	2.03	0.12	0.15	-0.01	0.69	1.24	0.58	0.83	2.01	0.47
03	2.01	0.11	0.25	-0.07	0.64	1.35	0.68	0.86	3.62	0.50
04	2.02	0.20	0.07	-0.05	0.72	1.04	0.59	0.87	3.04	0.56
05	2.08	0.02	-0.01	0.10	0.74	0.94	0.62	0.87	2.81	0.55
06	2.12	0.01	0.07	0.05	0.71	1.16	0.66	0.83	2.12	0.60
07	1.96	0.16	0.26	-0.04	0.73	0.92	0.61	0.84	3.12	0.58
08	2.01	0.21	0.17	-0.06	0.76	0.86	0.58	0.88	3.72	0.53
09	2.08	0.02	-0.04	0.10	0.73	0.99	0.66	0.87	3.54	0.64
10	2.35	-0.42	-0.58	0.40	0.74	0.73	0.57	0.84	2.09	0.64
11	2.05	-0.01	0.08	0.06	0.70	1.19	0.68	0.82	2.23	0.50
12	1.93	0.14	0.27	-0.02	0.70	1.04	0.66	0.88	4.64	0.48
13	2.07	0.04	-0.04	0.09	0.70	1.03	0.61	0.89	5.60	0.81
14	2.20	-0.39	-0.40	0.31	0.72	1.25	0.51	0.76	1.01	0.58

Note. **Bolded** cells indicate rater effects indices that might be flagged in practice.

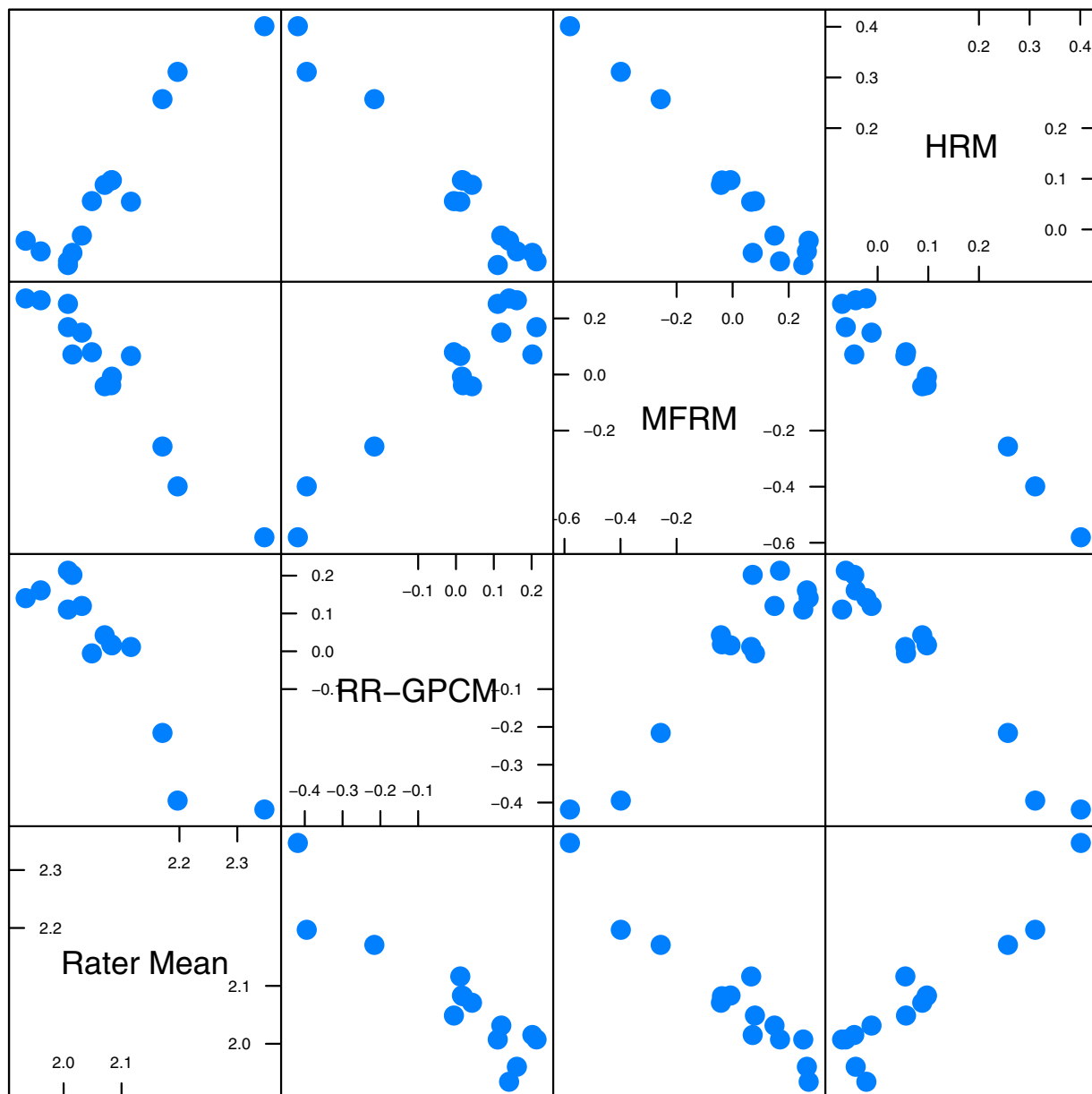
⁶Fit indices and model-based expected values from the MFRM have been used as measures of rater accuracy and centrality (Wolfe, 2004) but they were not used in this analysis.

Table 3 Correlation matrix of rater effects indices.

Rater effect	Index	Severity				Centrality		Accuracy			
		Rater mean	RR-GPCM location	MFRM severity	HRM bias	Rater SD	RR-GPCM SD(thres)	Percent exact agree	Corr with mean	RR-GPCM discrim	HRM rater var
Rater severity	Rater mean	1.000									
	RR-GPCM—location	−0.918	1.000								
	MFRM severity	−0.960	0.938	1.000							
	HRM bias	0.939	−0.976	−0.973	1.000						
Rater centrality	Rater SD	0.358	−0.259	−0.427	0.393	1.000					
	RR-GPCM—SD (thresholds)	−0.312	0.158	0.346	−0.322	−0.815	1.000				
Rater accuracy	Percent exact agreement	−0.502	0.536	0.623	−0.572	−0.529	0.303	1.000			
	Corr with mean	−0.423	0.593	0.424	−0.464	0.004	−0.379	0.392	1.000		
	RR-GPCM—discrimination	−0.530	0.568	0.474	−0.482	−0.180	−0.172	0.417	0.864	1.000	
	HRM—rater variability	0.500	−0.662	−0.557	0.579	0.329	0.053	−0.517	−0.884	−0.760	1.000

Note: Grayed cells of the correlation matrix highlight correlations between measures of the same rater effect. No significance tests were conducted due to small sample size ($n = 14$).

Measures of Rater Severity – FACT Pilot Data



Scatter Plot Matrix

Fig. 1 Scatterplot matrix of rater severity based on observed rater means (Rater Mean), and the rater effects estimates from the RR-GPCM, the MFRM, and the HRM.

from the MFRM and the HRM indicates that the MFRM's failure to properly model the correlation from multiple ratings of the same response does not drastically impact this index—instead, the impact would be more prominent in the standard errors of the latent trait estimates which would be larger in the HRM (Patz et al., 2002). The rater mean had the strongest correlation with the rater effect from the MFRM ($r = -0.96$) and that is likely because this version of the MFRM accounts for the least rater behavior compared to the other models and therefore it sorts raters more consistently.

Taking a closer look at Rater 12 reveals a low mean score (1.93) and the largest positive MFRM severity estimate (0.27), indicating that Rater 12 rates relatively severely. However, the RR-GPCM and HRM severity estimates are much smaller and do not indicate a tendency to be severe. In terms of leniency, Rater 10 has the highest mean rating (2.35) and the largest negative MFRM severity estimate (-0.58) as well as the most extreme RR-GPCM and HRM estimates. The RR-GPCM and the HRM estimates are

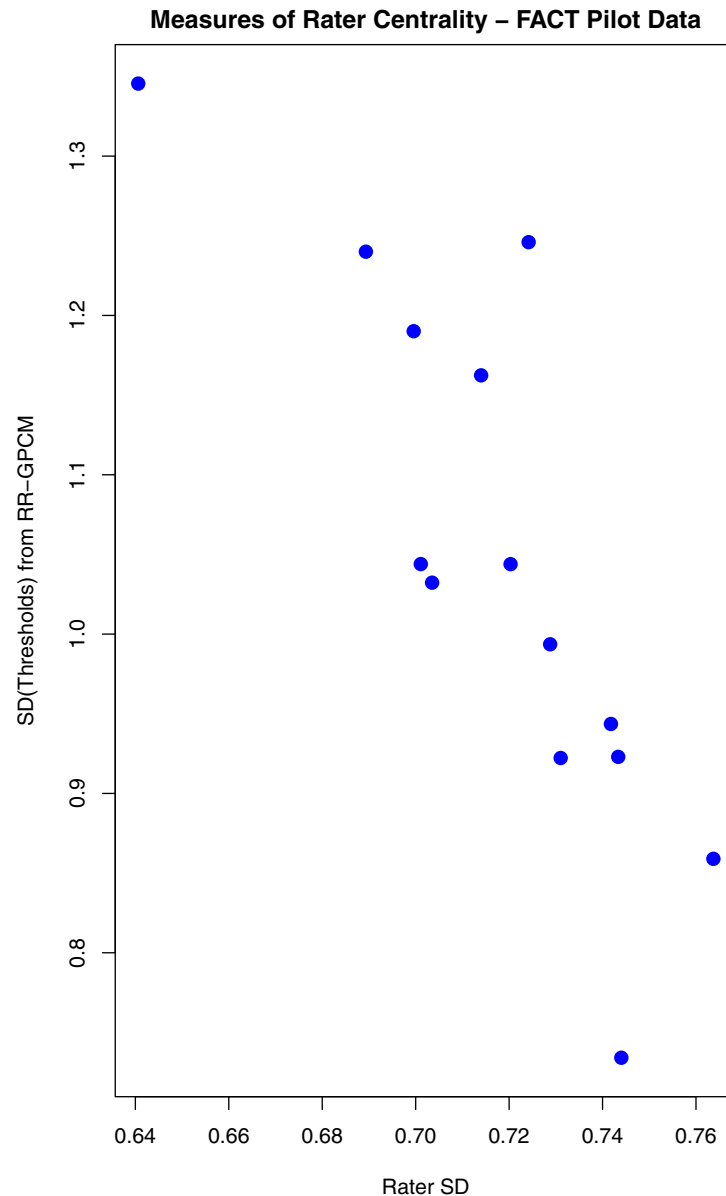


Fig. 2 Scatterplot of rater centrality based on observed rater SDs (x-axis) and the SD of the rater thresholds from the RR-GPCM (y-axis).

not large enough to indicate substantial leniency and this could be because they account for additional factors. For example, the RR-GPCM accounts for rater behavior along the score scale by including thresholds for raters. The HRM accounts for item effects via the GPCM and in addition to severity it also accounts for rater variability/inconsistency. Also note that all of these IRT models have latent traits and thus they account for test taker effects or the quality of the responses, which the observed mean does not. Accounting for differences in the quality of the response via the IRT modeling approach might change the way raters are ranked with respect to rater effects especially if a rater happens to score a disproportionate number of responses from test takers of low (or high) ability.

Fig. 2 plots the rater SDs against the SD of the rater thresholds from the RR-GPCM. These measures were highly negatively correlated ($r = -0.815$). A central rater will have a relatively small observed score SD compared to other raters as they are scoring in the middle of the score scale and not using the extremes. Conversely, a central rater will have a larger SD of thresholds from the RR-GPCM because their outer thresholds are pushed out into the extremes as a result of scoring so heavily in the middle. This in turn yields the negative correlation. While the correlation is high, it is not perfect and there are raters who would be ranked differently with respect to centrality when using these different indices. Much like when measuring severity with different models, the correlation between centrality indices will not be perfect because of the differences in what these models account for. The SD of thresholds from the RR-GPCM are based on model estimates that account for several factors (quality of the response, rater

Measures of Rater Accuracy – FACT Pilot Data

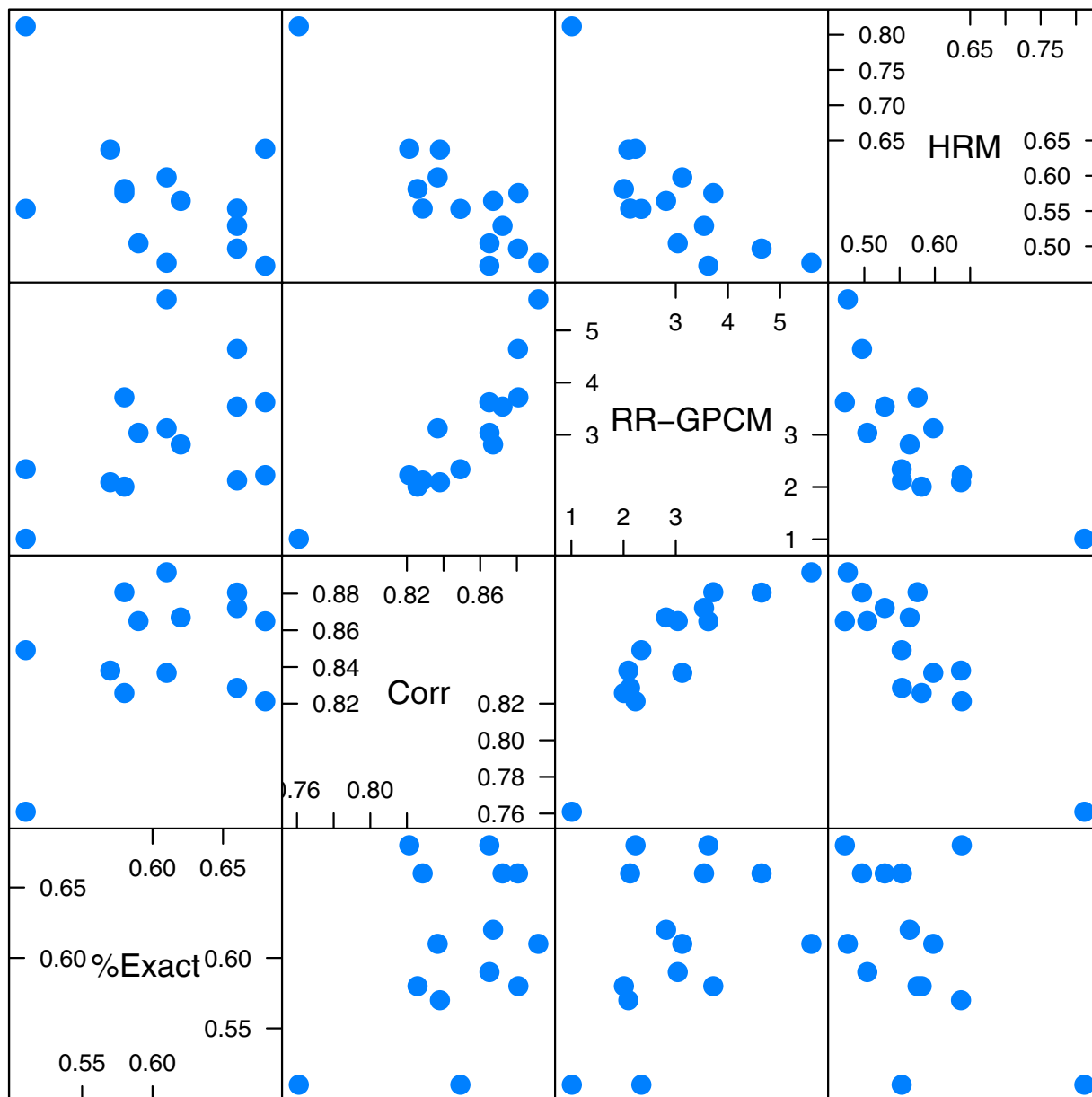


Fig. 3 Scatterplot matrix of rater accuracy based on the percent exact agreement (%Exact) on validity responses, the correlation with a rater's rating and the mean of all raters' ratings (Corr), the discrimination estimate from the RR-GPCM (RR-GPCM), and the rater SD estimate from the HRM (HRM).

severity, discrimination/accuracy, and the scoring behavior along the scale via thresholds). However, both indices are in clear agreement on the one rater at the top left of Fig. 2 (Rater 03) who is exhibiting a persistent tendency to score in the middle of score scale. Almost 60% of the ratings Rater 03 assigned were at score point 2 (the middle) and the rest of their ratings were split evenly between score levels 1 and 3. Restriction of range would also manifest in these indices as not using the full score scale will impact the SD even if the rater is using different points in between the extremes (not necessarily in the middle). However, since the FACT scale only has three score points there is not much room along the scale to have different clusters of ratings (aside from the one that exists for Rater 03).

Fig. 3 plots a series of rater accuracy measures: the percentage of exact agreement on validity responses, the correlation between the rater's rating and the mean rating, the discrimination estimate from the RR-GPCM, and the rater variability (SD) parameter estimate from the HRM which is a measure of intra-rater consistency. The strongest relationships are between the correlation with the mean rating and the HRM estimate, and the correlation and the GPCM discrimination estimate. The HRM estimate is also strongly correlated with the GPCM discrimination estimate ($r = -0.760$)—it is negatively correlated because larger discrimination estimates reflect higher levels of accuracy, but the HRM variability parameter will be smaller when a rater is more consistent or reliable. These strong relationships are expected since they are norm-referenced measures of accuracy. These measures were only moderately correlated with the percent exact agreement on the validity responses.

Identifying aberrant raters

These metrics provide much information about the raters' behavior which might be used to flag raters for remediation and retraining. Some metrics do not have established thresholds for flagging and thus an empirical approach may be applied, such as traditional outlier detection methods. For example, Rater 10 has a mean score that is extreme relative to the other raters, corresponding to a z -score > 2.5 . This is a large z score and depending on the outlier detection rules used, would probably be flagged as an outlier. Another example of an empirical approach is using the 95th percentile of the rater location estimates and using that value as the threshold for flagging severe raters (Stafford et al., 2018). Some metrics have established lower or upper bounds or rule sets used to flag raters. The MFRM's rater location term ϕ_r represents the relative locations of the raters' mean scores. A rater with $|\hat{\phi}_r - \bar{\phi}| > 0.5$ (more than 0.5 logits from the mean of the ϕ distribution, denoted by $\bar{\phi}$) might be flagged as being severe or lenient compared to other raters (Wolfe and McVay, 2012). In the FACT example, $\bar{\phi} = 0$ and therefore only Rater 10 would be flagged since they are the only rater with $|\hat{\phi}_r| > 0.5$. Other approaches are also used including computing Wald statistics to capture statistically significant deviations from the group mean (Wolfe, 2004).

Thresholds have been suggested for the HRM's rater effects indices. For the rater bias/severity estimate ϕ_r , an absolute value of 0.5 or greater is typically considered nonnegligible bias because it reflects the rater's tendency to score consistently 1 point higher or lower (since differences of 0.5 are rounded to a score point) (Patz et al., 2002). A rater SD of around 0.75 or higher has been noted as an indication of significant rater unreliability because it means that the rater is scoring inconsistently with a fairly large SD (Casanova and Wolfe, 2017).

Rater 03 very clearly has a SD lower than the other raters (albeit not that large in the grand scheme of things), and the RR-GPCM based estimate of centrality supports the notion that this rater is exhibiting centrality. A large(r) SD of thresholds is related to a rater's relatively heavier use of the middle categories since this pushes the thresholds farther apart. While this is only a 3-point scale, Rater 03 did use the middle category 60% of the time compared to the overall raters' behavior in which a score of 2 was only used for 47% of the responses. This difference in score distributions is reflected in the indices.

With respect to rater accuracy, both the correlation and the RR-GPCM accuracy estimate might flag Rater 14 as they had lowest values. Applying the outlier detection method mentioned earlier would flag this rater based on the correlation with the mean ($z < -2.5$) but not on the RR-GPCM accuracy estimate. Certainly, the exact agreement rates would lead to flagging Rater 14 and also Rater 01 however, that measure is based on a different set of responses so it may be considered an external criterion to the other indices and not directly comparable.

In summary, from this analysis the following raters may be flagged:

- Rater 10 would be flagged for scoring relatively higher (or leniently) than the other raters, and this would be supported by the MFRM. The RR-GPCM location estimate and the HRM bias estimate did not reach a "flaggable" quantities, which is typically a magnitude of 0.5.
- The RR-GPCM index would flag Rater 14 for a low level of accuracy. The observed score accuracy indices support this finding.
- Rater 03 would be flagged for centrality by both indices.
- Rater 01 might be flagged using the validity response data (51% exact agreement).

It is important to note that flagging raters on validity agreement statistics alone would have missed possible pertinent retraining and remediation for Raters 03 and 10. Feedback to raters based on their relative standing on rater effects indices can be targeted to address those specific issues. Myford and Wolfe (2003) provide strategies to minimize various rater effects via rater feedback and training.

Factors in model selection

All IRT rater models have specific affordances and limitations and may be appropriate in different use cases and contexts. As such, there are several considerations when selecting a model. Perhaps the first is the rating design, which dictates how raters are assigned to responses to be scored, and how many ratings each response receives. There are spiral designs in which raters score responses to a particular subset of items. There are nested designs in which raters score all responses per a test taker. In a random assignment, often used in online scoring systems for large-scale assessments, raters may score all or any subset of items and not necessarily the same test taker. The rating design is directly related to how the data (test taker responses and raters) are linked to each other,

or the extent to which there is “connectedness” (Engelhard, 1997; Schumacker, 1999). When there is insufficient connectedness of the data, the ability to estimate model parameters may be limited because information about raters and test takers cannot be separated or identified (Stafford et al., 2018; Wind and Jones, 2018).

Independent of the rating design, responses may be scored by multiple raters, or a small subset may be double-scored for the purposes of reliability estimation. How these multiple ratings are treated in the IRT model is naturally a defining factor for model selection. In the IRT context, ignoring the dependencies from multiple raters scoring the same response leads to an accumulation of information which in turn leads to falsely low latent trait standard errors (Mariano, 2002). Users should take caution about the models they choose in this scoring context. In some cases, the user may prefer to consider each rater’s evaluation of the response to be separate information (Linacre, 2003). However, typically when a rater is trained to score according to a rubric, the expectation is that raters should be similar in their evaluation as they are serving as “scoring machines” and their individual views on the response are not valued or considered as separate information. The importance of the rater model in this case is to account for ratings that are not consistent with the rubric or not assigned by raters who are equally good at using the rubric. Ratings should cluster around what is an appropriate value for the response.

Importantly, these models capture different rater effects. Most, if not all, IRT rater models will capture rater severity, but some do not capture centrality and/or accuracy, at least not directly with model parameters. It may be obvious to select the model that captures the most rater behavior or the most rater effects, however, there are several factors to consider which may lead to selecting a simpler model. The relatively more complex models offering more rater effects estimates may not be necessary. Complex models can be difficult to estimate and/or interpret, so it is important to apply them when appropriate and necessary. A preliminary check of observed score rater effects indices may suggest that a specific subset of rater effects must be captured by the IRT model and maybe others are less important. In addition, depending on the rater design, it may be especially important to capture specific rater effects such as halo effects. These types of considerations and determinations should be made a priori.

Another important factor is to understand the goals of the analysis. Is the goal to score test takers? Is the goal to monitor raters and provide feedback? Is the goal to do both? All the IRT rater models discussed here will provide estimates of the latent trait, however, the RR-(G)PCM will do this based on only one item (and multiple raters). This may not be the best way to scale test takers, and so when there is another scoring approach in place for computing scores, there may be a preference to apply this model in the context of rater monitoring only. If scoring is the main goal, using the HRM may be preferable to a model such as the MFRM, which will ignore the dependencies as discussed above.

Finally, a persistent limiting factor in the application of IRT models is sample size. IRT model estimation tends to require somewhat large samples to yield trustworthy estimates. In the event that the dataset is too small, it is advised to use the simpler models such as the MFRM and observed score statistics. Other more complex models such as the HRM may not be feasible due to the number of estimated parameters.

Summary

With the reliance on open-ended tasks and performance assessments and human raters to score these assessments, it is imperative to minimize construct irrelevant variance wherever possible. This chapter addressed different methods for modeling rater effects and flagging raters for aberrant behavior. While most rudimentary scoring systems may use basic agreement measures to compare raters’ scores to other raters’ scores or consensus scores, rater effects indices, both observed score metrics and IRT-based metrics, can provide more diagnostic information with which to provide feedback to raters. Augmenting existing scoring systems with this type of information can increase opportunities for raters to reduce their errors with targeted training. Certainly, creating a scoring system for a new assessment permits the use of the IRT framework from the very start. However, if there are logistical barriers to implementing some of these methods, it is clear that even examining simple observed score statistics can be very insightful.

Rater effects indices can be used in various ways including tracking raters and providing feedback to raters. Ideally future work will be able to connect established thresholds and subsequent decisions to improvements in the validity, reliability, and fairness in test scores and uses. However, overall there is a consensus that the quality control of constructed-response scores and the monitoring of raters is vital to maintaining efficient and effective scoring systems, and some well-defended plan should be in place to ensure the optimal conditions are in place. The interested reader is encouraged to further study the psychometric literature which contains an abundance of works on rater-mediated assessments and rater effects modeling.

Acknowledgments

Special thanks to Drs. Geoffrey Phelps and Brent Bridgeman for permission to use the FACT data.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association, Washington, DC.

- Andersen, E.B., 1997. The rating scale model. In: *Handbook of Modern Item Response Theory*. Springer, New York, NY, pp. 67–84.
- Andrich, D., 1978. A rating formulation for ordered response categories. *Psychometrika* 43, 561–573.
- Brennan, R.L., 1992. Generalizability theory. *Educ. Meas.* 11 (4), 27–34.
- Burry-Stock, J.A., Shaw, D.G., Laurie, C., Chissom, B.S., 1996. Rater agreement indexes for performance assessment. *Educ. Psychol. Meas.* 56 (2), 251–262.
- Casabianca, J.M., Wolfe, E.W., 2017. The impact of design decisions on measurement accuracy demonstrated using the hierarchical rater model. *Psychol. Test Assess. Model.* 59 (4), 471–492.
- Casabianca, J.M., Lockwood, J.R., McCaffrey, D.F., 2015. Trends in classroom observation scores. *Educ. Psychol. Meas.* 75 (2), 311–337.
- Casabianca, J.M., Junker, B.W., Nieto, R., Bond, M.A., 2017. A hierarchical rater model for longitudinal data. *Multivar. Behav. Res.* 52 (5), 576–592.
- Casabianca, J.M., 2021. Hierarchical rater models (Digital ITEMS Module 27). *Educ. Meas.* 40 (4).
- Chin, M., Pollard, C., Beisiegel, M., Hill, H., 2014. Exploring Rater Ability in Identifying Effective Mathematics Instruction. Paper Presented at the Annual Meeting of the National Council for Measurement in Education. San Francisco, California.
- Congdon, P.J., McQueen, J., 2000. The stability of rater severity in large-scale assessment programs. *J. Educ. Meas.* 37 (2), 163–178.
- Cronbach, L.J., 1955. Processes affecting scores on “understanding of others” and “assumed similarity”. *Psychol. Bull.* 52, 177–193.
- DeCarlo, L.T., Kim, Y., Johnson, M.S., 2011. A hierarchical rater model for constructed responses, with a signal detection rater model. *J. Educ. Meas.* 48 (3), 333–356.
- DeCarlo, L.T., 2005. A model of rater behavior in essay grading based on signal detection theory. *J. Educ. Meas.* 42 (1), 53–76.
- Eckes, T., 2009. Many-facet Rasch measurement. In: Takala, S. (Ed.), *Reference Supplement to the Manual for Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (Section H)*. Strasbourg, France: Council of Europe/Language Policy Division. Retrieved from: http://www.coe.int/t/dg4/linguistic/manual1_EN.asp?P19_2121.
- Educational Testing Service, 2021. Best Practices for Constructed-Response Scoring. Educational Testing Service, Princeton, NJ. https://www.ets.org/s/about/pdf/cr_best_practices.pdf.
- Engelhard, G., 1996. Evaluating rater accuracy in performance assessments. *J. Educ. Meas.* 33 (1), 56–70.
- Engelhard, G., 1997. Constructing rater and task banks for performance assessments. *J. Outcome Meas.* 1, 19–33.
- Fischer, G.H., 1973. The linear logistic test model as an instrument in educational research. *Acta Psychol.* 37 (6), 359–374.
- Guo, S., 2014. Correction of rater effects in longitudinal research with a cross-classified random effects model. *Appl. Psychol. Meas.* 38 (1), 37–60.
- Hill, H.C., Blunk, M.L., Charalambous, C.Y., Lewis, J.M., Phelps, G.C., Sleep, L., Ball, D.L., 2008. Mathematical knowledge for teaching and the mathematical quality of instruction: an exploratory study. *Cognit. Instr.* 26 (4), 430–511.
- Hombo, C.M., Donoghue, J.R., Thayer, D.T., 2001. A Simulation Study of the Effect of Rater Designs on Ability Estimation. ETS Research Report Series, pp. i–41.
- International Test Commission, 2012. International Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores. www.intestcom.org.
- Kane, T.J., McCaffrey, D.F., Miller, T., Staiger, D.O., 2013. Have we identified effective teachers? Validating measures of effective teaching using random assignment. In: *Research Paper. MET Project. Bill & Melinda Gates Foundation*.
- Leckie, G., Baird, J.A., 2011. Rater effects on essay scoring: a multilevel analysis of severity drift, central tendency, and rater experience. *J. Educ. Meas.* 48 (4), 399–418.
- Linacre, J.M., 1989. *Many-Facet Rasch Measurement*. Mesa Press, Chicago, IL.
- Linacre, J.M., 2003. The hierarchical rater model from a Rasch perspective. *Rasch Meas. Trans.* 17, 928–928.
- Linacre, J.M., 2004. Guessing and carelessness asymptotes: estimating IRT parameters with Rasch. *Rasch Meas. Trans.* 18, 959–960.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Routledge.
- Mariano, L.T., 2002. *Information Accumulation, Model Selection and Rater Behavior in Constructed Response Student Assessments*. Doctoral dissertation. Carnegie Mellon University.
- Masters, G.N., Wright, B.D., 1997. The partial credit model. In: van der Linden, W.J., Hambleton, R.K. (Eds.), *Handbook of Modern Item Response Theory*. Springer, New York, NY, pp. 101–121.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.
- Muraki, E., 1992. A generalized partial credit model: application of an EM algorithm. *ETS Res. Rep. Ser.* 1992 (1), i–30.
- Muraki, E., 1993. Variations of polytomous item response models: raters’ effect model, DIF model, and trend model. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Myford, C.M., Wolfe, E.W., 2003. Detecting and measuring rater effects using many-facet Rasch measurement: Part I. *J. Appl. Meas.* 4, 386–422.
- Myford, C.M., Wolfe, E.W., 2004. Detecting and measuring rater effects using many-facet Rasch measurement: Part II. *J. Appl. Meas.* 5, 189–227.
- Nieto, R., Casabianca, J.M., 2019. Accounting for rater effects with the hierarchical rater model framework when scoring simple structured constructed response tests. *J. Educ. Meas.* 56 (3), 547–581.
- Patz, R.J., Junker, B.W., Johnson, M.S., Mariano, L.T., 2002. The hierarchical rater model for rated test items and its application to large-scale educational assessment data. *J. Educ. Behav. Stat.* 27, 341–384. <https://doi.org/10.3102/10769986027004341>.
- Penfield, R.D., 2014. An NCME instructional module on polytomous item response theory models. *Educ. Meas.* 33 (1), 36–48.
- Phelps, G., Bridgeman, B., Yan, F., Steinberg, J., Weren, B., Zhou, J., 2020. Preliminary evidence on measurement characteristics for the foundational assessment of competencies for teaching performance tasks. *ETS Res. Rep. Ser.* 2020 (1), 1–50.
- Rasch, G., 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danmarks Paedagogiske Institute, Copenhagen, Denmark.
- Rasch, G., 1961. On general laws and the meaning of measurement in psychology. *Proceedings of the IV. Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA, pp. 321–333.
- Saal, F.E., Downey, R.G., Lahey, M.A., 1980. Rating the ratings: assessing the psychometric quality of rating data. *Psychol. Bull.* 88 (2), 413–428.
- Samejima, F., 1969. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monogr. Suppl.* 34 (4, Pt. 2), 100.
- Schumacker, R.E., 1999. Many-facet Rasch analysis with crossed, nested, and mixed designs. *J. Outcome Meas.* 3, 323–338.
- Stafford, R.E., Wolfe, E.W., Casabianca, J.M., Song, T., 2018. Detecting rater effects under rating designs with varying levels of missingness. *J. Appl. Meas.* 19 (3), 243–257.
- Verhelst, N.D., Verstralen, H.H., 2001. An IRT model for multiple raters. In: *Essays on Item Response Theory*. Springer, New York, NY, pp. 89–108.
- Wang, W.C., Su, C.M., Qiu, X.L., 2014. Item response models for local dependence among multiple ratings. *J. Educ. Meas.* 51 (3), 260–280.
- Wilson, M., Case, H., 2000. An examination of variation in rater severity over time: a study in rater drift. *Objective Meas.* 5, 113–134.
- Wilson, M., Hoskens, M., 2001. The rater bundle model. *J. Educ. Behav. Stat.* 26 (3), 283–306.
- Wind, S.A., Engelhard, G., 2012. Examining rating quality in writing assessment: rater agreement, error, and accuracy. *J. Appl. Meas.* 13 (4), 321–335.
- Wind, S.A., Jones, E., 2018. The stabilizing influences of linking set size and model–data fit in sparse rater-mediated assessment networks. *Educ. Psychol. Meas.* 78 (4), 679–707.
- Wind, S.A., Engelhard Jr., G., Wesolowski, B., 2016. Exploring the effects of rater linking designs and rater fit on achievement estimates within the context of music performance assessments. *Educ. Assess.* 21 (4), 278–299.
- Wolfe, E.W., McVay, A., 2012. Application of latent trait models to identifying substantively interesting raters. *Educ. Meas.* 31 (3), 31–37.
- Wolfe, E.W., 2004. Identifying rater effects using latent trait models. *Psychol. Sci.* 46, 35–51.
- Wolfe, E.W., 2014. *Methods for Monitoring Rating Quality: Current Practices and Suggested Changes*. Pearson, Iowa City.
- Wright, B.D., Masters, G.N., 1982. *Rating Scale Analysis*. MESA press.

Modeling item response times

Peter W. van Rijn^a and Sandip Sinharay^b, ^a ETS Global, Amsterdam, the Netherlands; and ^b Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Measuring item response times	322
Modeling item response times	323
Separate models for response times	324
Joint models for item responses and response times	325
Estimation	326
Model fit	326
Illustration	327
Discussion	329
References	329

Although many process variables recorded from computer-based educational assessments can nowadays be studied (e.g., Provasnik, 2021), response time, generally defined as the time it takes to respond to a test item, remains popular because it can be used for several purposes, is relatively easy to collect, and is typically available for all items. Consequently, the modeling of response times, often alongside item responses, has become an important research topic in educational assessment. For example, response times are studied in a variety of contexts, ranging from group-score assessments such as PISA (Kerzabi et al., 2020) to high-stakes college admission tests (van Rijn et al., 2021). In this entry, we aim to provide a brief review of modeling aspects of item responses and response times in the context of educational assessments.

Response times are often modeled (often under slightly different names such as “reaction time” and “duration”) in scientific fields such as psychology (Luce, 1986) and survival analysis (Cox and Oakes, 1984). However, it is not always easy to transfer concepts across fields in a meaningful way. For example, the speed-accuracy tradeoff often studied in psychological reaction-time experiments is usually observed within persons across time-limit conditions (Goldhammer, 2015), a situation not often encountered in educational assessment. Alternatively, the concept of hazard, the risk of an item not “surviving” the student’s solution process per time unit, may not be that useful in our context. Furthermore, even within the field of educational assessment, the analysis and modeling of item response times may require different approaches. That is, the response type and response time of multiple-choice items and that of extended constructed-response items can be qualitatively and quantitatively different.

Analyses of response times in educational assessment can serve a variety of purposes that can be roughly be divided into two types: (1) enhancing data quality and security, and (2) improving reliability and validity of the test scores. Response-time analyses of the first type can involve basic data-quality checks such as computation of the time test takers spend on the assessment relative to the allotted testing time (Schnipke, 1995) and sophisticated analyses such as detection of cheating behaviors using response times (van der Linden, 2009a; Sinharay, 2020; Sinharay and Johnson, 2019). Response-time analyses of the second type can, for instance, pertain to enhancing measurement precision beyond response accuracy (Bolsinova and Tijmstra, 2018) and evaluating construct-irrelevant factors such as disengagement (Wise, 2019).

Some caution in modeling response times is needed too. For example, Pohl et al. (2021) argue that response-time differences related to proficiency can be confounded with response-time differences related to test-taking behaviors (e.g., working speed). Therefore, there is a need to disentangle these two types in order to enhance test score interpretations (Kane, 2013) and fairness (Dorans and Cook, 2016). Furthermore, response times, and process data in general, may not always be comparable between test takers, for example, if testing accommodations were made available for test takers with disabilities or health-related needs.

Before discussing statistical modeling of response times, it is important to determine what the role of response times is with respect to the outcomes of the educational assessment. By outcomes, we mean results such as the final scores given to individual test takers in an admissions test and the group statistics reported in educational surveys such as PISA. A distinction can be made between levels of importance of response times with respect to the outcomes of an educational assessment (Bergner and von Davier, 2019). Response times can be considered *collateral* to the outcome when they are either irrelevant to the outcome, auxiliary to the outcome, or only important in understanding the outcome. Response times can be considered *essential* to the outcome when they are used in a scoring rubric or in the statistical model that produces the outcome. Consider, for example, the historical distinction between speed and power tests (Gulliksen, 1950). On a speed test, several easy items are to be answered as quickly as possible so that response times are essential to the outcome. In contrast, on a power test, items with a wide range of difficulty are to be answered correctly so that response times are typically collateral to the outcome. Given that most educational assessments are intended to be power tests and time limits are generally imposed for practical reasons, it is important to ask the question what the role of response times should be. This should not only be clear to the test developer, but also to the test taker.

Three additional, conceptual issues pertaining to response times are relevant in the context of educational assessment (van der Linden, 2009b). The first issue concerns the representation of response times as random variables. That is, response times are treated

as realizations of random variables so that methods from mathematical statistics can be used for making inferences (Luce, 1986). The second issue pertains to distinguishing notions of labor intensity to describe items and speed to describe test takers (van Breukelen, 1997). In certain response-time models (e.g., the log-normal model, see the section on separate models for response times), this is done in a similar fashion as the distinction between item difficulty and test-taker ability in case of item responses. Third, conditional independence assumptions such as the independence of response times conditional on latent speed play an important role in response-time modeling. The exemplified assumption is akin to the assumption of conditional (or “local”) independence in item response theory (IRT) models where item responses are independent conditional on latent ability. However, conditional independence assumptions cannot be automatically translated from IRT models to response-time models. For example, between-item dependencies can arise for response times when there is a total test time limit. That is, test takers can spend their time only once, and sequential dependencies between response times can occur if, for instance, test takers spend more time in the first half than in the second half of the assessment. If the time limit is too strict to allow all test takers to see all items, the test is considered speeded, which can lead to methodological and validity issues in the scoring of examinees (Lu and Sireci, 2007). Furthermore, in joint models for item responses and response times, modeling assumptions need to be specified to deal with within-item dependencies between item response and response time. For example, when the difference in mean response time for incorrect and correct item responses varies substantially across items, and this difference is both positive and negative, a flexible model would be needed.

We begin with a section on defining and measuring response times. This is followed by an overview of response-time modeling approaches, supplemented by estimation and model fit. An example of response-time modeling is discussed in some detail. The entry ends with a brief discussion and some suggestions for further reading.

Measuring item response times

Although it is sometimes thought that response times have only become available with the start of computer-based testing (CBT), item-level response times have been collected with paper-based testing as well. For example, Blommers and Lindquist (1944) measured response times by displaying consecutive numbers at 10 s intervals during test administration and asking test takers to record the number before and after each item. Obviously, CBT has simplified the recording of response time, but, in practice, many design features of the educational assessment can impact response time. Before response time can be defined, it needs to be clear on exactly what response it is to be captured, and which design features can impact the recording of response time. We discuss the issue of defining response time by referring to the IMS question and test interoperability (QTI) specification. This is a standard for the digital representation of test content and results developed by the IMS Global Learning Consortium.¹

In CBT, response time can typically be defined as the time difference between two events in the process data log of a test administration. In the QTI specification, “A duration value specifies a distance (in time) between two time points. In other words, a time period as defined by [ISO 8601], but represented as a single float that records time in seconds. Durations may have a fractional part.”² The key issue in defining response time lies in determining which two events are to be used. Furthermore, in the QTI specification, an item session is defined as “the accumulation of all the attempts at a particular instance of an item made by a candidate.” Although a test taker visiting an item multiple times is more or less captured in this definition, it is not explicit. That is, the extent to which test takers are allowed to navigate forward and backward in the test can substantially complicate the definition of response time and the correct extraction of response time from the log files (Kroehne and Goldhammer, 2018). Furthermore, time limits at the item and/or test level, multiple attempts, and feedback can all have an impact on the definition, extraction, and comparability of response time.

Other issues to be taken into account when measuring response time are related to the time it took for the item to load on the screen, which may be hardware-dependent, or, if the test is administered online, dependent on the speed of the internet connection. An interesting aspect of response-time data is that even if test takers did not provide an item response, the log file may reveal how much time was spent on the item. Such information can be helpful in the treatment of missing data.

Fig. 1 shows a simplified four-stage process that represents what happens when an examinee visits an item. In the first stage, the item is initiated. Different events here pertain to loading the item on the screen. For example, images, text, and interactive elements (e.g., buttons, text boxes, spreadsheets, navigation) are loaded on the screen. In the second stage, the test taker can interact with the item (e.g., enter text response, mouse clicks, click next item). In stage three, scoring events (to the extent possible given the type of item administered) take place (e.g., scoring multiple-choice or numeric-entry response), including display of feedback. In stage four, the item session ends. Depending on what is allowed by the design, item sessions can be restarted (e.g., if items can be revisited). Now, for defining response time, the first time stamp can be the beginning of the item session, but can also be the beginning of the interaction. The second time stamp can be the end of the item session, but can also be the beginning of the scoring events. In any case, it is relevant to be aware of the exact definition of response time in order to be able to detect any issues in the response time data. For example, if the beginning of the interaction is used as the first time stamp, but navigation buttons can become active earlier, negative response times could show up in the data if one is not careful in the implementation of the exact definition of response time.

¹See www.imsglobal.org/question/index.html and www.wikipedia.org/wiki/QTI.

²See www.imsglobal.org/question/qti2p1/imsqti_infov2p1.html and www.wikipedia.org/wiki/ISO_8601.

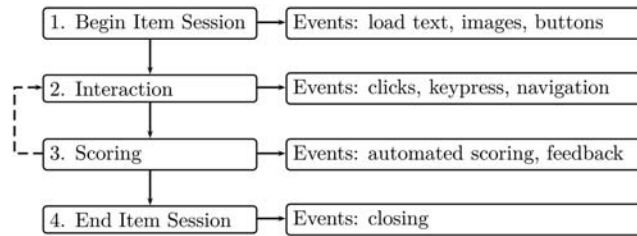


Fig. 1 Simplified lifecycle of item session.

Fig. 2 shows an example of multiple item visits by a test taker. In this example, the test taker navigates from item 1 to 2, item 2 to 3, item 3 to 2, item 2 to 3, and item 3 to 4. Response times are recorded for four items in six item sessions. There are multiple ways to treat response time in this case. The most basic way is to sum the response times of item sessions that are associated with the same item. However, it may be relevant to see how this test taker split response time across the multiple visits, and perhaps to link this information with the response data (e.g., to see if the original response was changed in the second visit).

In sum, measuring response time in educational assessment is not always straightforward. One has to know the exact definition of response time and be aware of the potential impact of other design features on response time in order to be able to make the correct inferences from response-time data.

Modeling item response times

In this entry, we do not intend to provide the reader with a comprehensive overview of models for response time in educational assessment, but rather to give a list of modeling options to start out with. Extensive reviews of response-time models are provided by Lee and Chen (2011), van Rijn and Ali (2017), De Boeck and Jeon (2019), and Kyllonen and Zu (2016). Although the distinction between speed and power tests mentioned before is often made, most educational assessments have elements of both types with generally more focus on power. We distinguish two approaches to modeling item response times. In the first approach, a *separate* model is specified for response times. In the second approach, a *joint* model is specified for item responses and response times.

In addition to the distinction between separate and joint modeling approaches, three conceptual approaches can be distinguished (van Rijn and Ali, 2017). In the *statistical* approach, the focus is on adequate modeling of the observed distributions of item responses and response times (van der Linden, 2006, 2007). In the *measurement* approach, the goal is to formulate a scoring rule based on accuracy and speed that leads to a sufficient statistic of some latent proficiency (Maris and van der Maas, 2012; van Rijn and Ali, 2018a). In the *process* approach, the goal is to describe the cognitive process underlying the item responses (van der Maas et al., 2011). Although the starting points in these conceptual approaches are different, the resulting models do not necessarily have to be different.

Before discussing models, it is useful to establish some notational conventions. We denote the continuous random variable representing the response time for test taker i on item j by Y_{ij} and its realization by y_{ij} . The density of response time is denoted as $f(y_{ij})$ and the corresponding cumulative distribution is given by $F(y_{ij}) = \int_0^{y_{ij}} f(u)du$. Although response time can be discretized, the focus in this entry is on response time as a continuous variable. In addition, item responses are assumed to be discrete with the associated random variable denoted by X_{ij} . Finally, the number of test takers is denoted by n and the number of items by m .

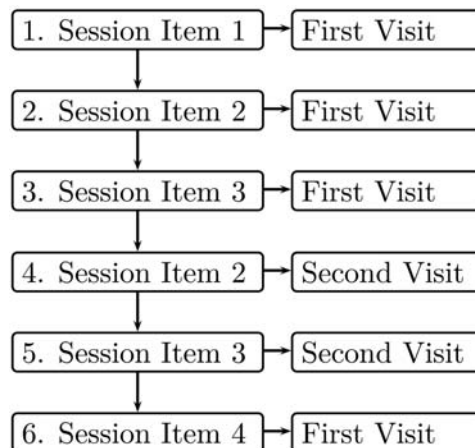


Fig. 2 Example of test-taker path with multiple item visits.

Separate models for response times

We discuss three models for response times that are based on the log-normal, Weibull, and gamma distributions, respectively. Since observed response-time distributions are often positively skewed (due to the inevitable presence of some large values of response time), distributions that can accommodate skewness are typically chosen for modeling. Probably the most popular model for response times in educational assessments is the log-normal model. Although this model was popularized by [van der Linden \(2006\)](#), it was also used by [Thissen \(1983\)](#) and the log transformation for response times was suggested even much earlier ([Furieux, 1948](#); [Tate, 1950](#)). Other models for response time include the inverse Gaussian, ex-Gaussian, ex-Wald, and Gumbel distributions ([van Zandt, 2000](#); [De Boeck and Jeon, 2019](#)), but we focus on the aforementioned three to simplify our exposition.

Regardless of the chosen distribution, experts have recommended the use of parameters that can account for shift, scale, and shape to accurately capture observed response-time distributions ([Rouder et al., 2003](#)). In applications of response-time models in educational assessments, shift parameters relate to the minimum amount of processing time that is needed before any response can be produced (e.g., reading the item), scale relates to the variability in response times, and shape to skewness and kurtosis.

The log-normal model is given by

$$f(y_{ij} | \tau_i) = \frac{1}{y_{ij} \sigma_j \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{[\log y_{ij} - (a_j + b_j \tau_i)]^2}{\sigma_j^2}\right), \quad (1)$$

where τ_i is a person parameter, and a_j , b_j , and σ_j^2 are item parameters. Equivalently, we can write $\log(y_{ij}) = a_j + b_j \tau_i + e_{ij}$ with $e_{ij} \sim N(0, \sigma_j^2)$. This model is essentially a normal linear factor model for $\log y_{ij}$ in which a_j is the intercept, b_j the factor loading, and σ_j^2 is the error variance ([Finger and Chee, 2009](#); [Molenaar et al., 2015a](#)). In the log-normal model described by [van der Linden \(2006\)](#), the factor loading b_j is fixed at -1, which allows for interpreting τ_i as a latent speed variable and a_j as the time intensity of item j . We shall refer to the van der Linden model as the one-parameter log-normal (1PLN) model and the model given by [Eq. \(1\)](#) as the two-parameter log-normal (2PLN) model. Further skewness can be accommodated by applying a one-parameter Box-Cox transformation to the response times ([Klein Entink et al., 2009](#)).

Different log-normal distributions for different types of item responses (such as correct and incorrect responses) can be specified without assuming a model for the item responses. In the following model, for example, the response times are allowed to be different for incorrect and correct responses ([Glas and van der Linden, 2010](#))

$$\log y_{ij} = a_j + b_j \tau_i + d_j x_{ij} + e_{ij}, \quad e_{ij} \sim N(0, \sigma_j), \quad (2)$$

where d_j is the additional item parameter and x_{ij} is the dichotomous item response with one for a correct response, and zero for an incorrect response. Other extensions of the log-normal model are possible, for example, by combining the extensions discussed above (e.g., a Box-Cox normal model with item response as predictor). However, the number of item parameters in such extensions can become quite large and large samples would be needed to obtain stable estimates.

The response-time model based on the three-parameter Weibull distribution is given by

$$f(y_{ij}) = \frac{k_j}{\lambda_j} \left(\frac{y_{ij} - c_j}{\lambda_j}\right)^{k_j-1} \exp\left[-\left(\frac{y_{ij} - c_j}{\lambda_j}\right)^{k_j}\right], \quad (3)$$

with parameters k_j for shape, λ_j for scale, and c_j for shift. In their model, [Tatsuoka and Tatsuoka \(1979\)](#) center the response times by $y_{ij}^* = y_{ij} - \bar{y}_i$, let $\tau_i + y_{ij}^*$, where τ_i is a person parameter, and fix $c_j = 0$ to simplify estimation. If the shape and scale parameters in [Eq. \(3\)](#) are fixed at 1, the model is similar to the exponential model by [Scheiblechner \(1979\)](#), which is given by $f(y_{ij} | \tau_i) = (\tau_i + a_j) \exp[-(\tau_i + a_j)y_{ij}]$.

The response-time model based on the gamma distribution is given by

$$f(y_{ij}) = \frac{y_{ij}^{k_j-1} \exp\left(-\frac{y_{ij}}{\lambda_j}\right)}{\lambda_j^{k_j} \Gamma(k_j)}, \quad (4)$$

with parameters k_j for shape and λ_j for scale, and $\Gamma(k_j)$ is the gamma function $\Gamma(k_j) = \int_0^\infty u^{k_j-1} \exp(-u) du$ for $k > 0$. [Maris \(1993\)](#) describes several models for response time based on the gamma distribution, assuming both additive and multiplicative relationships between item and person parameters.

We note that many other approaches can be used. Since response times can be considered to be durations, connections can be made to survival analysis (e.g., [Cox and Oakes, 1984](#)). We mention here the proportional-hazard modeling approach ([Ranger and Kuhn, 2012](#)) and the log-normal race model ([Rouder et al., 2015](#)). Also, if the interest is in modeling response times of correct responses only, special models can be constructed (see, e.g., [Muraki and Hagiwara, 2003](#); [Gaviria, 2005](#)). In addition, many different types of models can be constructed for special applications such as the Rasch Poisson model for reading speed ([Rasch, 1960](#), Chapter 3).

Alternatively, if the interest is in just fitting the response-time distribution and the sample is large, nonparametric methods such as Gaussian kernel estimators can be used ([van Zandt, 2000](#)). Such nonparametric methods can be useful for basic response-time

analysis (e.g., to evaluate timing issues or speededness), possibly combined with basic item-response analysis (e.g., to evaluate distractors and scoring). Particularly, such smoothing techniques allow for convenient graphical display.

Joint models for item responses and response times

Before discussing joint models, we briefly mention models for item responses. For dichotomously scored item responses, the following two-parameter setup can be employed

$$p(X_{ij} = 1 | \theta_i) = F[\alpha_j(\theta_i - \beta_j)], \quad (5)$$

where θ_i is a latent ability parameter, F is a (cumulative) distribution function, α_j is an item discrimination parameter, and β_j is an item difficulty parameter. Typical choices of F are the normal distribution and the more popular logistic distribution, referred to as the two-parameter logistic (2PL) model (Birnbaum, 1968). If the α_j are constant across items in the logistic model, the one-parameter logistic or Rasch model is obtained (Rasch, 1960). The three-volume handbook of item response theory (IRT) by van der Linden (2016) provides numerous other IRT models, including models with (pseudo-)guessing parameters, models for polytomous items, and multidimensional models.

In joint models, the focus is on modeling item responses and response times simultaneously. The speed-accuracy tradeoff (e.g., responses are less accurate if one is forced to work faster) is often taken into account in joint modeling. As noted in the introduction, however, this tradeoff originates from research on reaction times in experimental psychology and describes a within-person phenomenon resulting from performing tasks under different time limits. In educational assessments, it translates to a *speed-ability tradeoff* that can have both within- and between-person elements, but it can also be related to items and their positions in the test (van der Linden, 2009b; Goldhammer, 2015).

We discuss four different joint models for item responses and response times. The first model is the *hierarchical model* in which a model for item responses is combined with a model for response times (van der Linden, 2007). For example, the 2PL model can be used for item responses and the 2PLN model for response times. A bivariate normal distribution is assumed for latent ability θ and latent speed τ . This approach has been extended in several directions. For example, instead of IRT models, cognitive diagnostic models have also been used for item responses in combination with models for response time (e.g., Zhan et al., 2018). Furthermore, multidimensional latent variables can be specified for both ability and speed (Man et al., 2019).

The *Rasch response-time model* proposed by Roskam (1987, 1997) specifies the probability of a correct response as follows

$$p(X_{ij} = 1 | \theta_i, \gamma_{ij}) = \frac{\exp(\theta_i + \log \gamma_{ij} - \beta_j)}{1 + \exp(\theta_i + \log \gamma_{ij} - \beta_j)}, \quad (6)$$

where the log response time is used as a predictor of the item response. For response times, the following Weibull density is assumed:

$$f(\gamma_{ij} | \zeta_i, \theta_i) = \lambda_{ij} \gamma_{ij} \exp\left(-\frac{\lambda_{ij}}{2} \gamma_{ij}^2\right), \quad (7)$$

where $\lambda_{ij} = \frac{\exp(\theta_i)}{\exp(\beta_j) \zeta_i}$ and ζ_i is a person persistence parameter (van Breukelen and Roskam, 1991). A limitation of this model is the assumption that the probability of a correct response increases as response time increases, which makes the model suitable only for speed tests. A similar model was developed by Verhelst et al. (1997), making use of a gamma distribution for response time.

A different type of joint model can be specified if an explicit scoring rule is employed for the accuracy and speed of item responses. Maris and van der Maas (2012) used the signed-residual-time scoring rule that can be expressed as $s_{ij} = (2x_{ij} - 1)(h_j - \gamma_{ij})$, where h_j is the time limit for item j . In this scoring rule, fast correct responses are rewarded, while fast incorrect responses are penalized. A model in which this scoring rule is a sufficient statistic can be set up as

$$f(x_{ij}, \gamma_{ij} | \theta_i) = f(s_{ij} | \theta_i) = \frac{\exp(s_{ij} \eta_{ij})}{C(\eta_{ij})}, \quad (8)$$

where $\eta_{ij} = \alpha_j(\theta_i - \beta_j)$ and $C(\eta_{ij}) = \int_{-h_j}^{h_j} \exp(u \eta_{ij}) du$ is a normalizing constant (van Rijn and Ali, 2018a). This model, referred to as the *speed-accuracy response model*, is related to models for continuous item responses (Samejima, 1973; Müller, 1987; Verhelst, 2019).

Another joint model is the *diffusion model* (Ratcliff, 1978), which was mostly developed for two-choice memory retrieval tasks (Vandekerckhove et al., 2011), but has been applied in the context of ability testing (Tuerlinckx and de Boeck, 2005; van der Maas et al., 2011). In this model, the joint density of response accuracy and response time is given by

$$f(x_{ij}, \gamma_{ij} | \theta_i, \tau_i) = \frac{\pi}{B^2} \exp\left[BD\left(x_{ij} - \frac{1}{2}\right) - \frac{D^2}{2}(\gamma_{ij} - T_{nd})\right] \\ \times \sum_{k=1}^{\infty} k \sin\left(\frac{1}{2} \pi k\right) \exp\left[-\frac{\pi^2 k^2}{2B^2}(\gamma_{ij} - T_{nd})\right], \quad (9)$$

where B is a parameter referred to as boundary separation, D is the drift rate, and T_{nd} is the non-decision time. Here, we only mention the drift-diffusion model in which the boundary separation and drift rate are split into item and person parameters as follows: $D_{ij} = \theta_i - \beta_j$ and $B_{ij} = \tau_i/\alpha_j$. A simplified version of the diffusion model that eliminates the infinite sum in Eq. (9) is discussed by Boehm et al. (2021).

For the hierarchical model using the 2PL, the speed-accuracy response model, and the drift-diffusion model as discussed here, the marginal model for item responses is equivalent to a 2PL model (van Rijn and Ali, 2017). When joint models are used, it can be relevant to study the properties of the conditional item-response function $P(X_{ij} = 1|\theta_i, y_{ij})$ (also referred to as the conditional accuracy function) and the conditional response-time function $f(y_{ij}|\theta_i, x_{ij})$ for the purpose of model and parameter interpretation. However, we refrain from discussing such conditional functions here and refer to the original papers and comparisons described by Lee and Chen (2011), van Rijn and Ali (2017), and De Boeck and Jeon (2019). Finally, several other joint models and extensions exist, but the approaches discussed in this section should provide a good starting point. It is useful to keep in mind that several of the models can be framed as a multivariate generalized linear mixed model.

Estimation

Both frequentist and Bayesian estimation methods are generally available to estimate model parameters of the abovementioned models. The relationship of the log-normal models for response time with normal linear factor analysis (Lawley and Maxwell, 1971) greatly simplifies inference. However, this is not the case for some extensions of the log-normal model and other models for response times.

In the context of joint models, a number of assumptions are commonly made. For the hierarchical model, for example, the first assumption specifies that the response times for various items are independent of each other conditional on the latent speed variable, so that the joint conditional density can be written as a simple product. The second assumption specifies that item scores are independent conditional on the latent ability variable. Third, the item responses and response times are independent conditional on both latent variables (see van der Linden, 2007, p. 292). These three assumptions lead to the following simple product

$$f(\mathbf{x}_i, \mathbf{y}_i|\theta_i, \tau_i) = p(\mathbf{x}_i|\theta_i)f(\mathbf{y}_i|\tau_i). \quad (10)$$

In addition, a bivariate normal distribution is commonly assumed for latent variables θ and τ . Then, if we denote the model parameters by ξ , the marginal log-likelihood function is found by integrating out the latent variables as

$$\log L(\xi) = \sum_{i=1}^n \log \left[\int p(\mathbf{x}_i|\theta)f(\mathbf{y}_i|\tau)f(\theta, \tau)d\theta d\tau \right], \quad (11)$$

where the integral can be approximated using a Gauss-Hermite quadrature procedure. Maximum likelihood estimation can be carried out by finding the solution of the first-order derivatives set equal to zero. Details of marginal likelihood inference for the hierarchical model can be found in Glas and van der Linden (2010) and Kang et al. (2020).

The parameters of the speed-accuracy response model can also be estimated by maximizing its marginal log-likelihood function, which can be found under similar assumptions of conditional independence and a normally distributed latent variable. A computer program to estimate this model is described by van Rijn and Ali (2018b). Analogously, the parameters of the drift-diffusion model can be estimated by maximizing the marginal log-likelihood function. This method is implemented in the R package diffIRT (Moleenaar et al., 2015b).

Parameters of the hierarchical model can also be estimated with the method of moments by using an estimate of the covariance matrix. For example, the R package lavaan (Rosseel, 2012) can estimate the hierarchical model using least-squares methods if a normal ogive model is assumed for the item responses. This method is very flexible as it allows for easy generalizations to multi-dimensional models, inclusion of predictors, multi-group models, and correlated residuals.

Bayesian estimation can be performed after assuming prior densities for the model parameters, by applying Markov chain Monte Carlo (MCMC) methods. For the 1PLN, van der Linden (2006) used a Gibbs sampler to approximate the posterior densities. This approach is implemented for both 1PLN and 2PLN models in the R package LNIRT (<https://cran.r-project.org/web/packages/LNIRT/index.html>). Details of Bayesian inference for the hierarchical model can be found in van der Linden (2007) and this model can also be estimated in the R package LNIRT. However, the more general R package blavaan (<https://cran.r-project.org/web/packages/blavaan>) may be used as well.

Model fit

Depending on the purpose of modeling response times, different approaches to model fit may be relevant. For example, response time modeling for test security purposes (e.g., detecting item preknowledge) may require examining and evaluating different aspects of model fit than response time modeling for measurement purposes (e.g., improving measurement precision). Although many specific model features can be evaluated, we discuss three key elements with more focus on the fit to response times. The first element is that of the distributional assumption at the item level (e.g., see Eq. 1). Model fit can be evaluated by inspecting plots of the observed and predicted distributions of response times using estimated parameters, but P-P and Q-Q plots can be used as well. In addition, this type of fit can be summarized by using the root mean squared difference (RMSD) between observed and

predicted distributions at each ventile (i.e., 5th, 10th, ..., 95th percentile), as described by Schnipke and Scrams (1999). This approach is most useful when comparing different distributions, because the null distribution of RMSD is unknown.

Sinharay and van Rijn (2020) evaluated several statistics for testing the absolute fit of the log-normal model for response times. The Shapiro-Wilks test, which tests if log response time is normally distributed, was found to be the most powerful statistic for detecting item misfit. The statistic is based on order statistics and is easily computed using the R function `shapiro.test` (see <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/shapiro.test.html>).

Residuals from estimated response-time functions similar to those for item-response functions developed by Haberman et al. (2013) can also be used to evaluate item fit. Under the log-normal model of Eq. (1), for example, the expected log response time as a function of τ for item j is $\hat{F}_j(\tau) = E(\log y_j | \tau) = \hat{b}_j - \hat{a}_j \tau$. An alternative estimate of this function can be defined as the ratio

$$\bar{F}_j(\tau) = \frac{\sum_{i=1}^n \log(y_{ij}) \hat{g}(\tau | y_{ij})}{\sum_{i=1}^n \hat{g}(\tau | y_{ij})}, \quad (12)$$

where the posterior $\hat{g}(\tau | y_{ij})$ is normal (Bartholomew, 1981, Eq. 5). A fixed number of values τ can be chosen based on its prior and posterior density. After defining the residual $r_j(\tau) = \bar{F}_j(\tau) - \hat{F}_j(\tau)$, a standardized residual can be computed as $z_j(\tau) = r_j(\tau) / \sigma[r_j(\tau)]$, where an estimate of the $\sigma[r_j(\tau)]$ can be found following Haberman et al. (2013, Eq. 46). Under the log-normal model, this standardized residual would converge in distribution to a standard normal variable, and plotting these as a function of τ should look like random scatter. A similar approach is available for the speed-accuracy response model (van Rijn and Ali, 2018a).

The second element of model fit relates to conditional-independence assumptions. In joint models, these assumptions can be tested both within items (e.g., item response and response time) and between items (e.g., item responses for item pairs, response times for item pairs). Glas and van der Linden (2010) describe an approach to evaluate the three conditional-independence assumptions for the hierarchical model mentioned in the section on estimation. They use Lagrange multipliers to develop tests against a specific alternative model for each of the three assumptions. For example, for testing conditional independence of item responses and response times conditional on both latent variables, the alternative model they used is that of Eq. (2). Alternatively, Sinharay and van Rijn (2020) found that a test on the difference between the observed and expected correlation between log response times for item pairs was quite powerful for testing the between-item conditional-independence assumption for response times. A benefit is that this test is easily obtained using the function `lavResiduals` in the aforementioned R package `lavaan`. For a critical discussion of conditional-independence assumptions in joint models, see Bolsinova et al. (2017).

The third element in evaluating model fit adheres to the assumed distributions for the latent variables in the model. Glas and Verhelst (1989) developed a chi-squared test statistic for the fit of the ability distribution in the Rasch model and its polytomous extension, the partial credit model. Monroe (2021) describes a method based on posterior residuals in the context of IRT models. To our knowledge, however, there is little research in a frequentist context on the fit of the latent speed distribution in both separate and joint models.

If Bayesian methods are used, model fit can be evaluated using, for example, Bayesian model fit and model selection tools such as Bayes factors and posterior predictive model checking (e.g., Sinharay et al., 2006).

Illustration

The purpose of this illustration is to provide a simple demonstration of how different distributions and models for response times can be compared. We analyzed data obtained from 9355 test takers to 17 four-choice items of one out of two sections of a listening test, which is part of a test of English language ability used for international students (van Rijn and Ali, 2018b). Test takers had 10 min to answer all 17 items in this section of the test.

To illustrate the different distributions for response time described earlier, we fit log-normal, Weibull, and gamma distributions to the response-time data of each item. Following Schnipke and Scrams (1999), we summarized the fit using the root mean squared difference (RMSD) between observed and predicted distributions at each ventile (i.e., 5th, 10th, ..., 95th percentile). The mean and standard deviation (SD) of this statistic across all 17 items are shown for the three distributions in Table 1. The log-normal distribution seems to fit best in the sense that it has smaller mean RMSD values than the other two distributions both overall, and for incorrect and correct responses separately. To further illustrate the differences between the distributions, Fig. 3 shows histograms of response time and fitted log-normal, Weibull, and gamma densities for correct responses of four items. Items 2 and 10 were selected because they had relatively high RMSD values (0.043 and 0.051, respectively, for the overall log-normal), while items 3 and 14 had relatively low values (0.015 and 0.010, respectively, for the overall log-normal). Gaussian kernels were also fitted using the R function `density` and added to the plots (RMSDs were not computed). Not surprisingly, the Gaussian kernels appear to fit best, but these cannot be used in further modeling. Generally, the graphical results are in line with those in the table in that the log-normal distribution outperforms the Weibull and gamma distributions, which supports the use of the log-normal model in further modeling.

We used the R package `lavaan` to fit the log-normal model of Eq. (1) to the response times and evaluated model fit. The CFI is 0.802 and the RMSEA is 0.056, indicating poor model fit (Rigdon, 1996). We used the R function `shapiro.test` on a random subsample of 1000 test takers instead of the full sample to avoid detection of very small effects. Nevertheless, the test on the log

Table 1 Mean RMSD (SD) for log-normal, Weibull, and gamma distributions across 17 items of listening test.

Distribution	Mean RMSD (SD)		
	Overall	Incorrect	Correct
Log-normal	0.021 (0.012)	0.012 (0.003)	0.026 (0.015)
Weibull	0.053 (0.013)	0.038 (0.010)	0.061 (0.015)
Gamma	0.034 (0.014)	0.015 (0.006)	0.042 (0.016)

response time for each item returns 17 out of 17 significant results at the 0.05 level. So, even though the log normal distribution produced lower RMSD values and better fitting densities than the other two distributions, there still seem to be issues with model fit at the item level (e.g., due to differences in response-time distributions for incorrect and correct responses). In addition, the z-test for conditional independence on the full sample results in 98 significant results at the 0.05 level out of 136 pairs of response times. Of course, the sample size is rather large here and the power to detect misfit is substantial.

Notwithstanding these model fit results, we continued with fitting the hierarchical model for which we also used the lavaan package. This resulted in a CFI of 0.817 and an RMSEA of 0.062, again indicating poor model fit. The estimated correlation between latent ability and latent speed was 0.40, so higher ability is associated with higher speed in this case. In order to improve model fit, we allowed different response-time distributions for incorrect and correct answers. To this end, we fit the hierarchical model with the dichotomous item response as a predictor of response time as in Eq. (2). This resulted in a CFI of 0.930 and an RMSEA of 0.039, a drastic improvement. The latent correlation between ability and speed was estimated at 0.04. Noteworthy is that all 17 additional

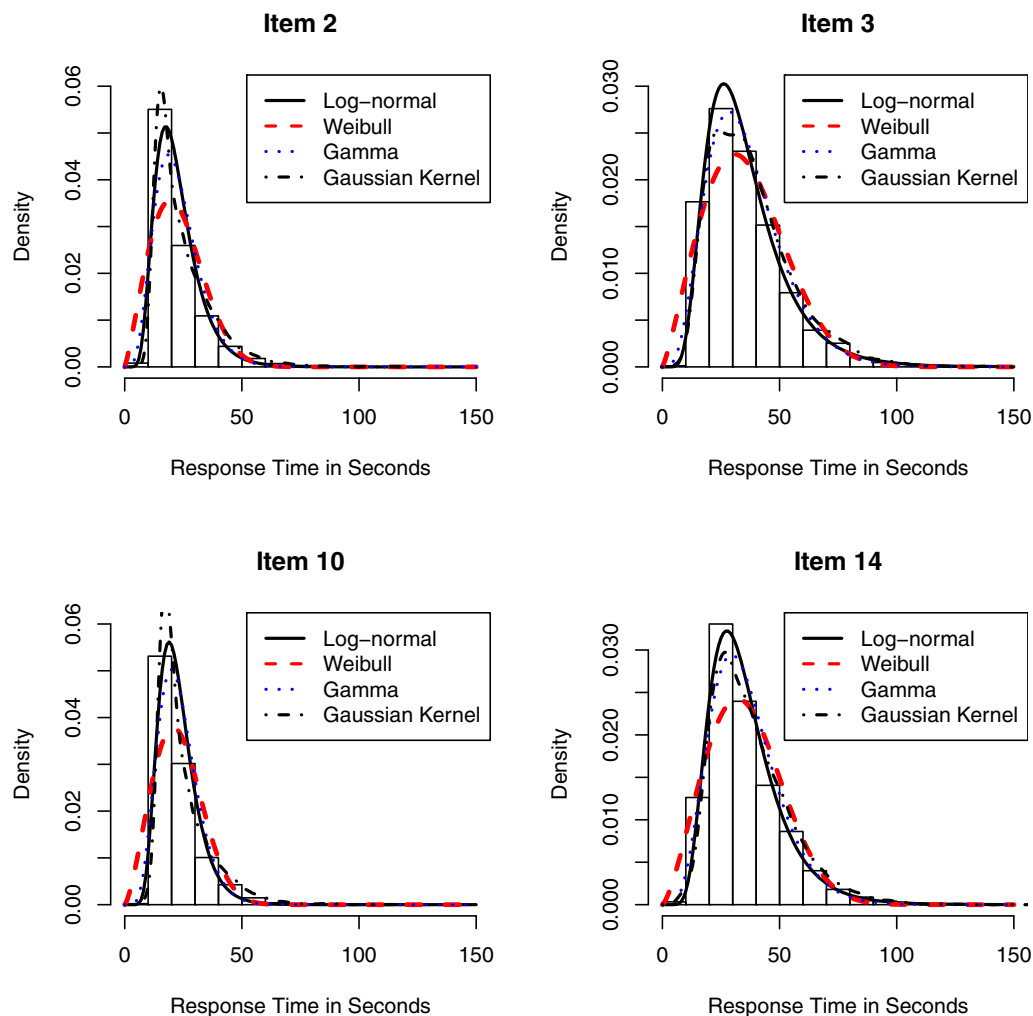


Fig. 3 Histogram of response times of correct responses with fitted log-normal, Weibull, gamma, and Gaussian kernel densities for four listening items.

regression parameters were significantly different from zero, which may explain the misfit discussed above. Alternatively, the original hierarchical model can also be relaxed by allowing residual correlations between item responses and response times within items, in which case the CFI becomes 0.870, RMSEA 0.053, and latent correlation 0.33. Although this is a substantial improvement compared to the original hierarchical model, it is not as good as the hierarchical model with item response as predictor.

Discussion

In this entry, we discussed both practical and theoretical aspects of measuring and modeling item response times in educational assessment. We discussed both separate models for item response times and joint models for item responses and response times. We briefly discussed methods for estimation and evaluating model fit. Response times can provide valuable information in addition to item responses, but poor model fit can be an issue, as shown in the illustration. Notwithstanding fit issues, models ultimately need to be judged based on their utility as well (Box, 1976) and the practical implications of misfit should be carefully evaluated before discarding them. In addition, the extent to which misfit has a practically significant impact depends on how response time is used (i.e., collateral vs. essential to the assessment outcome). The approach described by Sinharay and Haberman (2014) for IRT models in which practically significant misfit is determined by whether a better-fitting model leads to different decisions based on test scores can be tailored to models for response times.

In closing this entry, we would like to point out some topics that may guide readers in further study and research on modeling item response times. The generalized linear factor modeling framework described by Molenaar et al. (2015c) could be developed further. It can especially include more distributions for response time, as most distributions we discussed fit into the exponential family. Furthermore, the evaluation and relaxation of conditional-independence assumptions has been an active area of research (Bolsinova et al., 2017); although this research has focused mostly on within-item conditional independence of item responses and response times, there are examples of between-item studies (e.g., Li et al., 2012). Finally, mixture modeling approaches can be useful if one is interested in distinguishing different unobserved classes of response behaviors and processes (Molenaar and de Boeck, 2018).

References

- Bartholomew, D.J., 1981. Posterior analysis of the factor model. *Br. J. Math. Stat. Psychol.* 34, 93–99. <https://doi.org/10.1111/j.2044-8317.1981.tb00620.x>.
- Bergner, Y., von Davier, A.A., 2019. Process data in NAEP: past, present, and future. *J. Educ. Behav. Stat.* 44 (6), 706–732.
- Birnbaum, A., 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord, F., Novick, M. (Eds.), *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA, pp. 397–472.
- Blommers, P., Lindquist, E.F., 1944. Rate of comprehension of reading; its measurement and its relation to comprehension. *J. Educ. Psychol.* 35 (8), 449–473.
- Boehm, U., Marsman, M., van der Maas, H.L.J., Maris, G., 2021. An attention-based diffusion model for psychometric analyses. *Psychometrika* 86, 938–972.
- Bolsinova, M., Tijmstra, J., 2018. Improving precision of ability estimation: getting more from response times. *Br. J. Math. Stat. Psychol.* 71 (1), 13–38.
- Bolsinova, M., Tijmstra, J., Molenaar, D., De Boeck, P., 2017. Conditional independence between response time and accuracy: an overview of its possible sources and directions for distinguishing between them. *Front. Psychol.* 8. <https://doi.org/10.3389/fpsyg.2017.00202>.
- Box, G.E.P., 1976. Science and statistics. *J. Am. Stat. Assoc.* 71 (356), 791–799.
- Cox, D.R., Oakes, D., 1984. *Analysis of Survival Data*. Chapman and Hall.
- De Boeck, P., Jeon, M., 2019. An overview of models for response times and processes in cognitive tests. *Front. Psychol.* 10 (102). <https://doi.org/10.3389/fpsyg.2019.00102>.
- Dorans, N.J., Cook, L.L. (Eds.), 2016. *Fairness in Educational Assessment and Measurement*. Routledge, New York.
- Finger, M.S., Chee, C.S., 2009. Response-Time Model Estimation Via Confirmatory Factor Analysis. Paper presented at NCME, San Diego, CA.
- Furneaux, W.D., 1948. The Structure of "G", With Particular Reference to Speed and Power. *Proceedings of the 12th International Congress of Psychology*, Edinburgh.
- Gaviria, J.-L., 2005. Increase in precision when estimating parameters in computer assisted testing using response time. *Qual. Quant.* 39 (1), 45–69.
- Glas, C.A.W., van der Linden, W.J., 2010. Marginal likelihood inference for a model for item responses and response times. *Br. J. Math. Stat. Psychol.* 63, 603–626. <https://doi.org/10.1348/000711009x481360>.
- Glas, C.A.W., Verhelst, N.D., 1989. Extensions of the partial credit model. *Psychometrika* 54 (4), 635–659.
- Goldhammer, F., 2015. Measuring, ability, speed, or both? Challenges, psychometric solutions, and what can be gained from experimental control. *Measurement* 13, 133–164. <https://doi.org/10.1080/15366367.2015.1100020>.
- Gulliksen, H., 1950. *Theory of Mental Tests*. Wiley, New York.
- Haberman, S.J., Sinharay, S., Chon, K.H., 2013. Assessing item fit for unidimensional item response theory models using residuals from estimated item response functions. *Psychometrika* 78, 417–440. <https://doi.org/10.1007/s11336-012-9305-1>.
- Kane, M.T., 2013. Validating the interpretations and uses of test scores. *J. Educ. Meas.* 50, 1–77.
- Kang, H.-A., Zheng, Y., Chang, H.-H., 2020. Online calibration of a joint model of item responses and response times in computerized adaptive testing. *J. Educ. Behav. Stat.* 45 (2), 175–208.
- Kerzabi, E., Joo, S.-H., Robin, F., Yamamoto, K., 2020. Comparability of response time scales in pisa. *Psychol. Test Assess. Model.* 62 (1), 107–135.
- Klein Entink, R.H., van der Linden, W.J., Fox, J.-P., 2009. A Box-Cox normal model for response times. *Br. J. Math. Stat. Psychol.* 62, 621–640.
- Kroehne, U., Goldhammer, F., 2018. How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika* 45 (2), 527–563.
- Kyllonen, P.C., Zu, J., 2016. Use of response time for measuring cognitive ability. *J. Intell.* 4 (4), 14.
- Lawley, D., Maxwell, A., 1971. *Factor Analysis as a Statistical Method*. Butterworth, London.
- Lee, Y.-H., Chen, H., 2011. A review of recent response-time analyses in educational testing. *Psychol. Test Assess. Model.* 3, 359–379.
- Li, F., Cohen, A., Shen, L., 2012. Investigating the effect of item position in computer-based tests. *J. Educ. Meas.* 49 (4), 362–379.
- Lu, Y., Sireci, S.G., 2007. Validity issues in test speededness. *Educ. Meas.* 4, 29–37.
- Luce, R.D., 1986. *Response Times*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195070019.001.0001>.

- Man, K., Harring, J.R., Jiao, H., Zhan, P., 2019. Joint modeling of compensatory multidimensional item responses and response times. *Appl. Psychol. Meas.* 43 (8), 639–654.
- Maris, E., 1993. Additive and multiplicative models for gamma distributed random variables, and their application as psychometric models for response times. *Psychometrika* 58 (3), 445–469.
- Maris, G., van der Maas, H.L.J., 2012. Speed-accuracy response models: scoring rules based on response time and accuracy. *Psychometrika* 77, 615–633. <https://doi.org/10.1007/s11336-012-9288-y>.
- Molenaar, D., de Boeck, P., 2018. Response mixture modeling: accounting for heterogeneity in item characteristics across response times. *Psychometrika* 83 (2), 279–297.
- Molenaar, D., Tuerlinckx, F., van der Maas, H.L.J., 2015a. A bivariate generalized linear item response theory modeling framework to the analysis of responses and response times. *Multivariate Behav. Res.* 50, 56–74. <https://doi.org/10.1080/00273171.2014.962684>.
- Molenaar, D., Tuerlinckx, F., van der Maas, H.L.J., 2015b. Fitting diffusion item response theory models for responses and response times using the R-package diffIRT. *J. Stat. Software* 25 (2). <https://doi.org/10.18637/jss.v066.i04>.
- Molenaar, D., Tuerlinckx, F., van der Maas, H.L.J., 2015c. A generalized linear factor model approach to the hierarchical framework for responses and response times. *Br. J. Math. Stat. Psychol.* 68, 197–219. <https://doi.org/10.1111/bmsp.12042>.
- Monroe, S., 2021. Testing latent variable distribution fit in IRT using posterior residuals. *J. Educ. Behav. Stat.* 46 (3), 374–398.
- Müller, H., 1987. A Rasch model for continuous responses. *Psychometrika* 52, 165–181. <https://doi.org/10.1007/BF02294232>.
- Muraki, E., Hagiwara, Y., 2003. Samejima's continuous response model for item response time. In: International Meeting of Psychometric Society (IMPS). Cagliari, Italy.
- Pohl, S., Ulitzsch, E., von Davier, M., 2021. Reframing rankings in educational assessments. *Science* 372 (6540), 338–340.
- Provasnik, S., 2021. Process data, the new frontier for assessment development: rich new soil or a quixotic quest? *Large-Scale Assess. Educ.* 9 (1), 1–17.
- Ranger, J., Kuhn, J.-T., 2012. A flexible latent trait model for response times in tests. *Psychometrika* 77, 31–47. <https://doi.org/10.1007/s11336-011-9231-7>.
- Rasch, G., 1960. Probabilistic Models for Some Intelligence and Attainment Tests. Paedagogike Institute, Copenhagen.
- Ratcliff, R., 1978. A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. <https://doi.org/10.1037//0033-295x.85.2.59>.
- Rigdon, E.E., 1996. CFI versus RMSEA: a comparison of two fit indexes for structural equation modeling. *Struct. Equ. Model.* 3 (4), 369–379.
- Roskam, E.E., 1987. Toward a psychometric theory of intelligence. In: Roskam, E.E., Suck, R. (Eds.), *Progress in Mathematical Psychology*. Elsevier Science, pp. 151–174.
- Roskam, E.E., 1997. Models for speed and time-limit tests. In: Hambleton, R.K., van der Linden, W.J. (Eds.), *Handbook of Modern Item Response Theory*. Springer, New York, pp. 187–208.
- Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. *J. Stat. Software* 48 (2).
- Rouder, J.N., Sun, D., Speckman, P.L., Lu, J., Zhou, D., 2003. A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika* 68, 589–606.
- Rouder, J.N., Province, J.M., Morey, R.D., Gomez, P., Heathcote, A., 2015. The lognormal race: a cognitive-process model of choice and latency with desirable psychometric properties. *Psychometrika* 80 (2), 491–513.
- Samejima, F., 1973. Homogeneous case of the continuous response model. *Psychometrika* 38 (2), 203–219.
- Scheibelechner, H., 1979. Specifically objective stochastic latency mechanisms. *J. Math. Psychol.* 19 (1), 18–38.
- Schnipke, D.L., 1995. Assessing speededness in computer-based tests using item response times. Paper presented at the annual meeting of NCME, San Francisco, CA.
- Schnipke, D.L., Scrams, D.J., 1999. Representing Response-Time Information in Item Banks. Law School Admission Council Computerized Testing Report. LSAC Research Report Series. Tech. Rep. No. LSAC-R-97-09. Princeton, NJ.
- Sinharay, S., 2020. Detection of item preknowledge using response times. *Appl. Psychol. Meas.* 44 (5), 376–392.
- Sinharay, S., Haberman, S.J., 2014. How often is the misfit of item response theory models practically significant? *Educ. Meas.* 33, 23–35.
- Sinharay, S., Johnson, M.S., 2019. The use of item scores and response times to detect examinees who may have benefited from item preknowledge. *Br. J. Math. Stat. Psychol.* 73, 397–419. <https://doi.org/10.1111/bmsp.12187>.
- Sinharay, S., van Rijn, P.W., 2020. Assessing fit of the lognormal model for response times. *J. Educ. Behav. Stat.* 45 (5), 534–568.
- Sinharay, S., Johnson, M.S., Stern, H.S., 2006. Posterior predictive assessment of item response theory models. *Appl. Psychol. Meas.* 30, 298–321.
- Tate, M.W., 1950. Notes on the measurement of mental speed. *J. Educ. Psychol.* 41 (4), 219.
- Tatsuoka, K., Tatsuoka, M., 1979. A Model for Incorporating Response-Time Data in Scoring Achievement Tests Tech. Rep. No. CERL Report E-7.
- Thissen, D., 1983. Timed testing: an approach using item response theory. In: Weiss, D.J. (Ed.), *New Horizons in Testing: Latent Trait Test Theory and Computerized Adaptive Testing*. Academic Press, New York. <https://doi.org/10.1016/b978-0-12-742780-5.50019-6>.
- Tuerlinckx, F., de Boeck, P., 2005. Two interpretations of the discrimination parameter. *Psychometrika* 70, 629–650. <https://doi.org/10.1007/s11336-000-0810-3>.
- van Breukelen, G.J.P., 1997. Separability of item and person parameters in response time models. *Psychometrika* 62, 525–544. <https://doi.org/10.1007/bf02294641>.
- van Breukelen, G.J.P., Roskam, E.E., 1991. A Rasch model for the speed-accuracy tradeoff in time limited tests. In: Falmagne, J.-C., Doignon, J.-P. (Eds.), *Mathematical Psychology, Current Developments*. Springer, New York, pp. 251–272.
- Vandekerckhove, J., Tuerlinckx, F., Lee, M.D., 2011. Hierarchical diffusion models for two-choice response times. *Psychol. Methods* 16, 44–62. <https://doi.org/10.1037/a0021765>.
- van der Linden, W.J., 2006. A lognormal model for response times on test items. *J. Educ. Behav. Stat.* 31, 181–204. <https://doi.org/10.3102/10769986031002181>.
- van der Linden, W.J., 2007. A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika* 72, 287–308. <https://doi.org/10.1007/s11336-006-1478-z>.
- van der Linden, W.J., 2009a. A bivariate lognormal response-time model for the detection of collusion between test takers. *J. Educ. Behav. Stat.* 34, 378–394.
- van der Linden, W.J., 2009b. Conceptual issues in response-time modeling. *J. Educ. Meas.* 46, 247–272.
- van der Linden, W.J., 2016. *Handbook of Item Response Theory*. Chapman and Hall/CRC, Boca Raton, FL.
- van der Maas, H.L.J., Molenaar, D., Maris, G., Kievit, R., Borsboom, D., 2011. Cognitive psychology meets psychometric theory: on the relation between process models for decision making and latent variable models for individual differences. *Psychol. Rev.* 118, 339–356. <https://doi.org/10.1037/a0022749>.
- van Rijn, P.W., Ali, U.S., 2017. A comparison of item response models for accuracy and speed of item responses with applications to adaptive testing. *Br. J. Math. Stat. Psychol.* 70, 317–345. <https://doi.org/10.1111/bmsp.12101>.
- van Rijn, P.W., Ali, U.S., 2018a. A generalized speed-accuracy response model for dichotomous items. *Psychometrika* 83, 109–131. <https://doi.org/10.1007/s11336-017-9590-9>.
- van Rijn, P.W., Ali, U.S., 2018b. SARME: A Computer Program for Estimating Speed-Accuracy Response Models ETS Research Report ETS RR-18-15. Princeton, NJ.
- van Rijn, P.W., Attali, Y., Ali, U.S., 2021. Impact of scoring instructions, timing, and feedback on measurement: an experimental study. *J. Exp. Educ.* 1–23. <https://doi.org/10.1080/00220973.2021.1969532>.
- van Zandt, T., 2000. How to fit a response time distribution. *Psychon. Bull. Rev.* 7 (3), 424–465.
- Verhelst, N.D., 2019. Exponential family models for continuous responses. In: Veldkamp, B.P., Sluiter, C. (Eds.), *Theoretical and Practical Advances in Computer-Based Educational Measurement*. Springer, pp. 135–160.
- Verhelst, N.D., Verstralen, H.H.F.M., Jansen, M.G.H., 1997. A logistic model for time-limit tests. In: *Handbook of Modern Item Response Theory*. Springer, pp. 169–185.
- Wise, S.L., 2019. Controlling construct-irrelevant factors through computer-based testing: disengagement, anxiety, & cheating. *Educ. Inq.* 10 (1), 21–33.
- Zhan, P., Jiao, H., Liao, D., 2018. Cognitive diagnosis modelling incorporating item response times. *Br. J. Math. Stat. Psychol.* 71 (2), 262–286.

Quality control: response style modeling

Mirka Henninger^a and Thorsten Meiser^b, ^a University of Zurich, Zurich, Switzerland; and ^b University of Mannheim, Mannheim, Germany

© 2023 Elsevier Ltd. All rights reserved.

Introduction	331
Multidimensional extensions of ordinal IRT models	332
IRTree models	333
Empirical illustration	335
Multidimensional extension of the PCM	336
IRTree models	336
Comparison of extended PCMs and IRTrees	338
Discussion	338
References	339

Introduction

In psychological and educational measurement, many researchers use rating scales to assess abilities, personality dispositions, attitudes, beliefs, or opinions. However, observed rating responses reflect not only the latent construct that is to be measured, but the responses are also influenced by the respondents' interpretation and use of the rating categories (Baumgartner and Steenkamp, 2001). Examples of such response tendencies or *response styles* are the preference for extreme categories (i.e., extreme response style, ERS) or for the middle category (i.e., midscale response style, MRS), or the tendency to agree with the item independent of item content (i.e., acquiescent response style, ARS; e.g., Couch and Keniston, 1960; Cronbach, 1942; Hamilton, 1968; Hui and Triandis, 1985; Paulhus, 1991; Rost et al., 1999; Van Vaerenbergh and Thomas, 2013).

The awareness of response styles, and the interest in controlling their effects to safeguard valid measurement, have increased in recent years. Because it has been shown that response styles exert substantial effects on rating responses (e.g., Adams et al., 2019; Böckenholt and Meiser, 2017; Bolt and Johnson, 2009) and that individual differences in response styles are stable over content domains and over time (Van Vaerenbergh and Thomas, 2013; Weijters et al., 2010a, 2010b; Wetzel et al., 2013b), response styles may put data quality at risk. Hence, if unaccounted for, response styles can bias the measurement of latent constructs in educational research. For instance, they can influence diagnostic decisions through biasing individual estimates (Bolt et al., 2014; Plieninger, 2017), inflate relations between trait estimates or total scores (Böckenholt and Meiser, 2017), and affect group comparisons in cross-cultural research (De Jong et al., 2008; Moors, 2004; Morren et al., 2012; van Herk et al., 2004).

Early attempts to control for response styles used descriptive statistics, such as the number of extreme categories to measure ERS (Bachman and O'Malley, 1984; Cronbach, 1942; Greenleaf, 1992). Later, latent variable models, such as Structural Equation Models (SEM) or Item Response Theory (IRT) models, were extended to incorporate, and therewith account for, response styles. For example, mixture distribution IRT models allowed item parameters to differ between latent subpopulations (Rost, 1991) and hence, they can be used to identify subpopulations that use the rating scale in different ways. Indeed, applications of mixture models have identified two major subpopulations: one subpopulation with moderate response behavior preferring non-extreme categories, and one subpopulation with a preference for extreme response categories (e.g., Meiser and Machunsky, 2008; Moors, 2003; Wetzel et al., 2013a). However, although the identification of unknown latent subpopulations of respondents is an interesting characteristic of mixture distribution models, the assumption of categorical response style types (such as a type of moderate responders vs. a type of extreme responders) appears too strict and implausible given that most psychological and educational constructs are continuous in nature.

Consequently, several approaches have been proposed to incorporate response styles as continuous dimensions in IRT models, and these approaches will be the focus of the present chapter. First, IRT models for ordinal responses have been extended by additional response style dimensions accompanying the constructs of interest (e.g., Bolt and Newton, 2011; Falk and Cai, 2016; Henninger and Meiser, 2020a; Wetzel and Carstensen, 2017). Second, by decomposing rating responses into multiple discrete judgments with appropriate IRT representations, so-called IRTree models reflect response processes that are related to the traits of interest and to response styles as continuous dimensions (e.g., Böckenholt, 2012; De Boeck and Partchev, 2012; Khorramdel and von Davier, 2014; Meiser et al., 2019).

Besides modeling response styles as continuous dimensions, extended IRT models for ordinal responses and IRTree models have several advantages in common. First, both approaches are flexible in accounting for numerous variants of response styles. Second, they allow researchers to specify, model, and test specific theoretical assumptions about response styles in a confirmatory way. Third, each approach enables researchers to estimate latent correlations between the traits of interest and response styles. Fourth, both modeling approaches are easy to implement in existing software packages. And finally, even though we focus on the modeling

of response styles in this chapter, both modeling approaches can easily be adapted and extended to model other response tendencies or processes, such as socially desirable responding, missing responses, or test quitting behavior (see e.g., Falk and Cai, 2016; Jeon and De Boeck, 2016). Despite their common advantages, multidimensional IRT models and IRTree models for response styles differ in decisive ways, as will be discussed in the following sections.

The chapter is structured as follows. In the next sections, we provide introductions into multidimensional extensions of ordinal IRT models (Section **Multidimensional extensions of ordinal IRT models**) and into IRTree models (Section **IRTree models**) that accommodate response styles. In the remainder, we illustrate the similarities and differences of ordinal IRT and IRTree models for response styles using background questionnaire data from the PISA study 2018 (Section **Empirical illustration**). In the final discussion (Section **Discussion**), we summarize further modeling extensions, critical aspects for model application, and alternative response formats as a remedy of response style effects.

Multidimensional extensions of ordinal IRT models

Responses to rating scales (e.g., with $K + 1$ response categories $k \in \{0, \dots, K\}$) can be described through IRT models for ordinal data, such as the Graded Response Model (GRM; Samejima, 1969) and models from the Divide-by-Total model family (Thissen and Steinberg, 1986), such as the Nominal Response Model (NRM; Bock, 1972; Takane and de Leeuw, 1987), the Partial Credit Model (PCM; Masters, 1982), or the Rating Scale Model (RSM; Andrich, 1978).

These modeling approaches describe the probability that respondent n chooses category k in item i as a function of a latent person parameter θ_n describing the respondent's level on the construct of interest, an item location parameter β_i , and an item-specific threshold parameter τ_{ik} . Here we present the PCM as one typical model that can be used for ordinal rating data in further detail: In the PCM (adapted from Masters, 1982), the probability of choosing a certain category k is given by

$$p(X_{in} = k) = \frac{\exp\left(s_k \theta_n - \sum_{k'=0}^k (\beta_i + \tau_{ik'})\right)}{\sum_{j=0}^K \exp\left(s_j \theta_n - \sum_{k'=0}^j (\beta_i + \tau_{ik'})\right)} \quad (1)$$

with $\sum_{k=1}^K \tau_{ik} = 0$. The scoring weights s_k describe the relation between the latent person parameter and the observed category, and in the PCM they are usually fixed to $s = (0, \dots, K)$ for the $K + 1$ categories to reflect the ordinal nature of rating data. For example, for an item with 4 response categories, scoring weights can be fixed to $s = (0, 1, 2, 3)$. For identification, the sum of the item and threshold parameter for the first response category $k = 0$ is set to zero, that is, $(\beta_i + \tau_{i0}) \equiv 0$. The denominator of Eq. (1) accumulates the exponential of the linear parameter combination for each category, which ensures that the sum of all category probabilities is equal to 1 for all values of θ .

Fig. 1 shows the *category characteristic curves* (CCCs) for an item with 4 response categories (e.g., 0: “strongly disagree”, 1: “disagree”, 2: “agree”, 3: “strongly agree”), where each CCC depicts the probability of a response category for a given item location and set of thresholds as a function of the person parameter θ . The concept of category probabilities is best illustrated in the middle panel of Fig. 1 that shows the CCCs in the case of a neutral response style. For instance, respondents with a person parameter of $\theta_n = 1$ have the highest probability for the “agree” category as compared to the other categories, while for respondents with person parameter $\theta_n = 2$, it is most likely that they choose the “strongly agree” category. At the same time, the model is probabilistic, meaning that it makes a prediction for the most probable category, but also allows that other categories are chosen with a smaller probability.

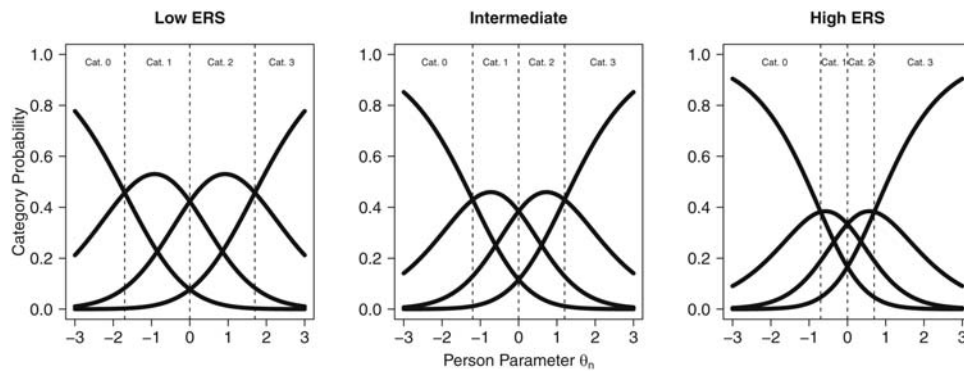


Fig. 1 Category characteristic curves for an item with four response categories (Cat. 0 to Cat. 3) and low (left), intermediate (middle), and high (right) levels of extreme response style.

Eq. (1) can be used to model responses to rating scales of ordinal nature. However, to accommodate response styles, the model equation must be extended to incorporate additional dimensions that reflect the way respondents use the rating categories. Hence, in multidimensional PCMs (and also in multidimensional GRMs, NRMs, or RSMs), additional latent dimensions can be specified that capture systematic influences on item responses other than individual differences in θ . For example, we can assume that not only the construct we are interested in, but also the ERS level influences the selection among the “strongly disagree” versus “disagree” categories, and among the “agree” versus “strongly agree” categories, respectively. Similarly, for rating scales with uneven numbers of response categories, a response in the middle category may not only depend on the construct level θ , but also on the respondent’s MRS level (see e.g., Bolt and Newton, 2011; Falk and Cai, 2016; Henninger and Meiser, 2020a, 2020b; Wetzel and Carstensen, 2017).

In an extended model that accommodates extreme responding with an additional ERS dimension, the category probabilities can be specified as

$$p(X_{in} = k) = \frac{\exp\left(s_k \theta_n - \sum_{k'=0}^k (\beta_i + \tau_{ik'}) + s_k^{\text{ERS}} \theta_n^{\text{ERS}}\right)}{\sum_{j=0}^K \exp\left(s_j \theta_n - \sum_{k'=0}^j (\beta_i + \tau_{ik'}) + s_j^{\text{ERS}} \theta_n^{\text{ERS}}\right)} \quad (2)$$

The scoring weights of the response style dimension (here s_k^{ERS}) define which categories are influenced by the response style level. For example for an item with four response categories, the scoring weights for ERS can be set to $s_k^{\text{ERS}} = (1, 0, 0, 1)$ indicating that positive levels of θ_n^{ERS} increase the probability of choosing one of the extreme categories, while negative levels of θ_n^{ERS} decrease the probability for an extreme response.

The left and right panels of Fig. 1 illustrate how the response style dimension θ^{ERS} affects the category probabilities, where CCCs are displayed for low and high levels of ERS, respectively. When ERS is low, the probability of passing the first threshold increases, while the probability of passing the last threshold decreases (first and last vertical line), making the non-extreme categories the modal categories over a larger range of θ as compared to a neutral response tendency (middle panel). The opposite applies in the case of high ERS: The extreme categories become more likely for smaller and larger values of θ as it is harder to pass the first, and more difficult to pass the last threshold (see also Henninger and Meiser, 2020a, 2020b, for an extensive discussion).

Hence, multidimensional IRT models allow for different levels of response styles, so that the probability of choosing a certain category does not only depend on the person parameter θ , but also on a respondent’s value on the response style dimension. For instance, in Fig. 1 a respondent with $\theta_n = 1$ and *low* ERS level (left panel) has the highest probability for choosing the “agree” category (Cat. 2), while a respondent with the same person parameter $\theta_n = 1$ and *high* ERS level (right panel) has the highest probability for choosing the “strongly agree” category (Cat. 3).

In Fig. 1 and in the empirical example below, we use responses to 4-point rating scales to illustrate IRT models that incorporate response styles, because 4-point rating scales are commonly used in the background questionnaires of the PISA studies and in other educational assessments. To model response styles in 4-point rating scales, researchers usually add only an ERS dimension (see Eq. (2)), although preferences for other categories can be considered in an analogous way. When the data is based on rating scales with five or more response categories (as is typical in the measurement of personality constructs, such as the Big Five), there is more flexibility in the modeling of various types of response styles. For instance, Table 1 shows exemplary scoring weights for the construct of interest θ as well as for ERS and MRS in items with a five-point rating scale (see also Bolt and Newton, 2011; Falk and Cai, 2016; Henninger, 2020; Wetzel and Carstensen, 2017, for further examples of scoring weights). Besides the typical response styles ERS and MRS, scoring weights can be used to specify other types of response tendencies, such as ARS through $s_k^{\text{ARS}} = (0, 0, 0, 1, 1)$, socially desirable responding (see e.g., Falk and Cai, 2016), individually varying thresholds (Henninger, 2020; Wang et al., 2006), or preferences for specific categories (Adams et al., 2019; Bolt et al., 2014). For overviews on how scoring weights can be defined to reflect different types of response tendencies, see Falk and Cai (2016) and Henninger and Meiser (2020a, b).

IRTree models

Following the rationale of decision trees, IRTree models decompose rating responses into a series of underlying judgments that terminate in the choice of an observed category (Böckenholt, 2012; De Boeck and Partchev, 2012). The judgments can be specified

Table 1 Scoring weights for different constructs and response styles in items with five response categories $k \in (0, \dots, 4)$.

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$s_{\text{Construct}}$	0	1	2	3	4
$s_{\text{Construct(Reversed Items)}}$	4	3	2	1	0
s_{ERS}	1	0	0	0	1
s_{MRS}	0	0	1	0	0

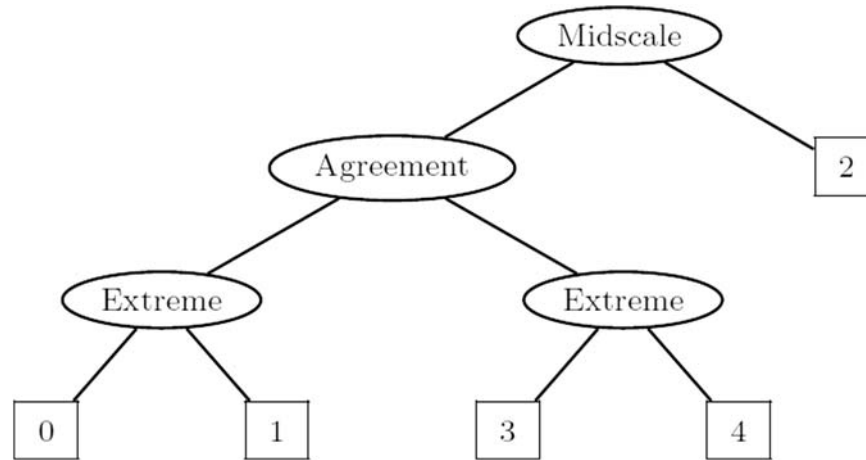


Fig. 2 IRTree structure for 5-point rating items with midscale, agreement and extreme response judgments.

as decision nodes which reflect trait-based response processes of endorsing the item content and general response tendencies affecting the selection of categories. For a 5-point rating item as in [Table 1](#), [Fig. 2](#) shows an IRTree structure with three judgment processes (for applications, see [Böckenholt, 2012](#); [Khorramdel and von Davier, 2014](#); [Zettler et al., 2016](#)). The first node refers to the decision to choose the midscale response 2 as opposed to a rating category of either disagreement (i.e., 0 or 1) or agreement (i.e., 3 or 4). If a respondent does not choose the midscale category, they decide between disagreement or agreement with the item at the second node. The third decision node relates to the selection of an extreme versus non-extreme response category given a disagreement or agreement judgment.

The assumed latent judgment processes of an IRTree model are specified by means of virtual pseudo-items together with an IRT parameterization for each decision node, as illustrated in [Table 2](#). Concerning response X_i to rating item i , the first node in [Fig. 2](#) can be represented via a binary pseudo-item Y_{1i} which takes on value 1 if the midscale category is chosen and 0 otherwise. Likewise, the second node can be captured by a pseudo-item Y_{2i} which is scored 1 for a non-extreme or extreme agreement response, 0 for the two disagreement responses, and missing for midscale responses. The third judgment, finally, can be represented through a pseudo-item Y_{3i} with value 1 for extreme responses and value 0 for non-extreme responses of agreement or disagreement. The pseudo-items of an IRTree model are parameterized with an appropriate IRT model, like the Rasch, 2 PL or probit model, with separate person parameters for the assumed judgment processes. The probability of observing category k in rating item i for respondent n , $p(X_{in} = k)$, follows from multiplying over the model probabilities of the pseudo-items that are involved in the category response (for details, see [Böckenholt and Meiser, 2017](#); [Jeon and De Boeck, 2016](#)). [Table 2](#) illustrates the parameterization of pseudo-items with the Rasch model, including person parameters for midscale responding θ^{MRS} , agreement with the item content θ , and extreme responding θ^{ERS} . The table also shows the model equation of the category probabilities for the IRTree structure in [Fig. 2](#).

For rating items with four categories that are often used in educational studies, the midscale judgment can be omitted, yielding an IRTree model with two latent dimensions of agreement θ and extreme responding θ^{ERS} ([Kim and Bolt, 2021](#); [Plieninger and Meiser, 2014](#)). The IRTree structure follows from the diagram in [Fig. 2](#) if the first node is dropped and the remaining terminal nodes are relabeled as categories $k = 0, \dots, 3$. The corresponding definition of binary pseudo-items and their parameterization with the Rasch model are displayed in the upper panel of [Table 3](#).

Like in [Table 2](#), the pseudo-items for 4-point rating scales in [Table 3](#) maintain the assumption that the extremity judgment is based on a unidimensional person dimension θ^{ERS} . Inasmuch as the selection of an extreme response category of disagreement or agreement is driven by (a) the tendency to prefer extreme response options and (b) the strength of the trait that is measured by the item, however, θ^{ERS} reflects a confound of the general response style and the specific trait of interest. We therefore used multidimensional parameterizations of the IRTree node probabilities (see also [Böckenholt, 2019](#); [Jeon and De Boeck, 2016](#); [von Davier and Khorramdel, 2013](#)) to account for the joint influence of the response style and the target trait in the choice of (non-)extreme response categories ([Meiser et al., 2019](#)). For this purpose, the pseudo-item of extremity has to be split into separate pseudo-items for the disagreement and agreement categories to allow for different signs of the target trait θ , as displayed in the lower panel of [Table 3](#): According to the ordinal nature of the rating scale, the probability of the extreme category 0 should decrease with increasing values of θ , whereas the probability of extreme category 3 should increase for larger values of θ . Hence, θ has a negative sign for judgments of extreme disagreement but a positive sign for judgments of extreme agreement. In addition, θ is weighted by the parameter $\alpha_{\text{dis}} > 0$ in the pseudo-item of extreme disagreement and by $\alpha_{\text{ag}} > 0$ in the pseudo-item of extreme agreement to allow for a differential impact of the target trait on the overall agreement judgment in the first node and on the gradual judgments of disagreement and agreement in the subsequent nodes. To summarize, the IRTree model with multidimensional extremity nodes in the lower panel of [Table 3](#) combines an ordinal trait-based response process over the full range of the rating scale with an individual tendency of response extremity.

Table 2 Pseudo-items of midscale, agreement and extreme responding with resulting model probabilities for rating categories k .

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	
Midscale (Y_{1i})	0	0	1	0	0	$\frac{\exp(\theta_n^{\text{MRS}} - \beta_{1i})}{1 + \exp(\theta_n^{\text{MRS}} - \beta_{1i})}$
Agreement (Y_{2i})	0	0	–	1	1	$\frac{\exp(\theta_n - \beta_{2i})}{1 + \exp(\theta_n - \beta_{2i})}$
Extremity (Y_{3i})	1	0	–	0	1	$\frac{\exp(\theta_n^{\text{ERS}} - \beta_{3i})}{1 + \exp(\theta_n^{\text{ERS}} - \beta_{3i})}$
$p(X_{in} = k) = \frac{\exp(y_{1in}(\theta_n^{\text{MRS}} - \beta_{1i}))}{1 + \exp(\theta_n^{\text{MRS}} - \beta_{1i})} \left(\frac{\exp(y_{2in}(\theta_n - \beta_{2i}))}{1 + \exp(\theta_n - \beta_{2i})} \right)^{(1-y_{1in})} \left(\frac{\exp(y_{3in}(\theta_n^{\text{ERS}} - \beta_{3i}))}{1 + \exp(\theta_n^{\text{ERS}} - \beta_{3i})} \right)^{(1-y_{1in})}$						

Table 3 Pseudo-items of agreement and extreme responding for 4-point rating items.

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	
Unidimensional nodes					
Agreement	0	0	1	1	$\frac{\exp(\theta_n - \beta_{1i})}{1 + \exp(\theta_n - \beta_{1i})}$
Extremity	1	0	0	1	$\frac{\exp(\theta_n^{\text{ERS}} - \beta_{2i})}{1 + \exp(\theta_n^{\text{ERS}} - \beta_{2i})}$
Multidimensional nodes					
Agreement	0	0	1	1	$\frac{\exp(\theta_n - \beta_{1i})}{1 + \exp(\theta_n - \beta_{1i})}$
Extremity	1	0	–	–	$\frac{\exp(\theta_n^{\text{ERS}} - \alpha_{\text{dis}} \theta_n - \beta_{2i})}{1 + \exp(\theta_n^{\text{ERS}} - \alpha_{\text{dis}} \theta_n - \beta_{2i})}$
	–	–	0	1	$\frac{\exp(\theta_n^{\text{ERS}} + \alpha_{\text{ag}} \theta_n - \beta_{2i})}{1 + \exp(\theta_n^{\text{ERS}} + \alpha_{\text{ag}} \theta_n - \beta_{2i})}$

Empirical illustration

In this section, we demonstrate that response styles are a substantial variance component of rating data in the field of education. The following analyses use data from the Program for International Student Assessment (PISA) and illustrate how psychometric models can help researchers to account for ERS in the estimation of latent constructs. In addition, we compare the multidimensional extension of the PCM with the IRTree model and demonstrate to what extent the approaches differ with respect to the estimation of person parameters and estimated correlations between latent constructs.

For the empirical illustration, we selected two subscales about reading from the background questionnaire¹ of the PISA 2018 study.² The “reading enjoyment” scale (ST160) contains five items including three inverted items, while the “reading self-evaluation” scale (ST161) contains six items with three inverted items. The selection of these two scales allows us, first, to model different constructs at the same time, and second, to use scales with reverse-coded items. Both aspects facilitate the separation of construct and response style dimensions when estimating response style models and are hence recommended modeling strategies for multidimensional PCMs and IRTree models (see e.g., Henninger and Meiser, 2020b). All items had a 4-point response format with categories from “strongly disagree” to “strongly agree”. For the present analyses, we selected data from native English speaking US participants. We paid careful attention to retain only valid data and therefore removed participants with incomplete data and those who did not select at least three different response categories over the selected items. The final sample comprised data from $N = 4,411$ participants from the US that gave responses to the 11 items.

All analyses reported below were conducted with the R package `mirt` (Chalmers, 2012) using the EM algorithm with 25 quadrature points. Full latent variance-covariance matrices were estimated to allow for correlated reading constructs and response style dimensions. Inverted items were recoded prior to the analyses.³ Exemplary R code, that allows users to apply the models presented here to their own data, is provided on OSF.⁴

¹https://nces.ed.gov/surveys/pisa/pisa2018/questionnaires/Student_Q_Booklet_English.html.

²<http://www.oecd.org/pisa/data/2018database/>.

³As an alternative to recoding inverted items prior to model specification, original responses can be analyzed with scoring weights $s^{\text{Construct Reversed}} = (3, 2, 1, 0)$ for reversed-coded items, as shown in Table 1.

⁴<https://osf.io/bu5qt/>.

Multidimensional extension of the PCM

To illustrate the estimation and interpretation of a multidimensional extension of the PCM, we fit two models, one ignoring ERS and one accounting for ERS according to Eqs. (1) and (2). Model 1 is a two-dimensional PCM with the items from the two scales “reading enjoyment” and “reading self-evaluation” each defining one latent dimension. Model 2 is an extension of Model 1 with an additional dimension that accounts for ERS. The scoring weights were fixed to $s^{\text{Construct}} = (0, 1, 2, 3)$ for the two reading dimensions and to $s^{\text{ERS}} = (1, 0, 0, 1)$ for the ERS dimension. Table 4 shows the log-likelihood, number of estimated parameters, AIC and BIC for the two models.

Model comparisons with the likelihood ratio test ($\chi^2 = 2819.33, df = 3$) and the information criteria AIC and BIC indicated that Model 2 fits the data better than Model 1, corroborating the role of ERS. Moreover, the latent correlations show that the two reading dimensions are substantially related, $\hat{\rho} = .44$, while the ERS dimension forms a distinct factor with weaker relations to the two reading dimensions, $\hat{\rho} = .23$ and $\hat{\rho} = .16$. In addition, Model 2 shows substantial variance of the ERS dimension ($\hat{\sigma}^2 = 1.96$), compared to the two reading dimensions ($\hat{\sigma}^2 = 4.00$ and $\hat{\sigma}^2 = 2.58$).

Fig. 3 illustrates the impact of ERS on rating responses. Panel A shows the observed relative frequencies of each category for three subgroups of respondents with different ERS levels based on ERS estimates in Model 2. From left to right, the columns show the relative frequencies of category choices for the 20% of respondents with lowest ERS levels, for respondents with intermediate ERS levels, and for respondents with the highest ERS levels (80% to 100% quantiles). We can see that the distribution of category choices differs substantially between the three subgroups, where respondents with low ERS levels tend to avoid the extreme categories, while respondents with high ERS levels tend to overselect the extreme categories. Panel B shows the construct estimates of the reading enjoyment scale based on a PCM ignoring ERS (x-axis) and the PCM accounting for ERS (y-axis). The color shades indicate the ERS level based on ERS estimates in Model 2. We can see that by accounting for ERS through the multidimensional model, low reading enjoyment estimates are corrected downwards for low ERS levels (light blue circles), but they are corrected upwards for high ERS levels (dark blue squares). For high estimates, in contrast, the correction works in the opposite direction, such that larger estimates result for low ERS and smaller estimates for high ERS. Thereby, the multidimensional PCM accounts for an over- or underselection of the extreme categories through an extreme response tendency.

IRTTree models

From the class of IRTree models, the model with unidimensional nodes of agreement and extremity in the upper panel of Table 3 was applied to the reading data as Model 3, and the model with two-dimensional nodes of extremity in the lower panel of Table 3 was applied as Model 4. As in the case of the PCMs above, separate dimensions were specified for the items of the subscales “reading enjoyment” and “reading self-evaluation”. The analysis allowed us to compare the two versions of the IRTree model, and it afforded comparisons of parameter estimates between the IRTree that accommodates ERS and the standard PCM that does not capture response styles.⁵ Fit indices of the Models 3 and 4 are displayed in Table 4.

The traditional IRTree Model 3 with unidimensional node parameterizations showed a similar estimated variance-covariance matrix as the extended PCM, including substantial variance of the ERS dimension ($\hat{\sigma}^2 = 1.84$), compared to the two reading dimensions ($\hat{\sigma}^2 = 5.57$ and $\hat{\sigma}^2 = 3.70$).

The latent correlation of the two reading dimensions amounts to $\hat{\rho} = .48$, and their estimated correlations with the latent ERS dimensions are $\hat{\rho} = .15$ and $\hat{\rho} = .31$. The notable correlations of the reading constructs with the response style dimension hint at the potential confound of the reading dimensions and the response style in the factor θ^{ERS} of Model 3. Model 4, in contrast, separates effects of the target traits and of ERS in the pseudo-item for extreme disagreement and agreement. As can be seen in Table 4, Model 4 provides a better fit to the reading data as compared with the traditional IRTree Model 3. Moreover, due to accounting for

Table 4 Model fit of the ordinal IRT models and IRTree models for response styles.

	Log-likelihood	Parameters	AIC	BIC
Model 1: two-dimensional PCM ignoring ERS	− 49,160.73	36	98,393.47	98,623.58
Model 2: three-dimensional PCM with ERS dimension	− 47,751.07	39	95,580.14	95,829.43
Model 3: IRTree model with unidimensional nodes	− 50,872.82	28	101,801.60	101,980.60
Model 4: IRTree model with multidimensional nodes	− 48,422.44	32	96,908.88	97,113.42

⁵Please note that the two-dimensional PCM ignoring ERS is a special case of the three-dimensional PCM accounting for ERS within the same model class. However, there is no equivalent restriction of the IRTree model one could directly compare the IRTree model to. Therefore, we use the construct estimates of the PCM ignoring ERS for comparison (see e.g., Böckenholt, 2017, for a similar approach).

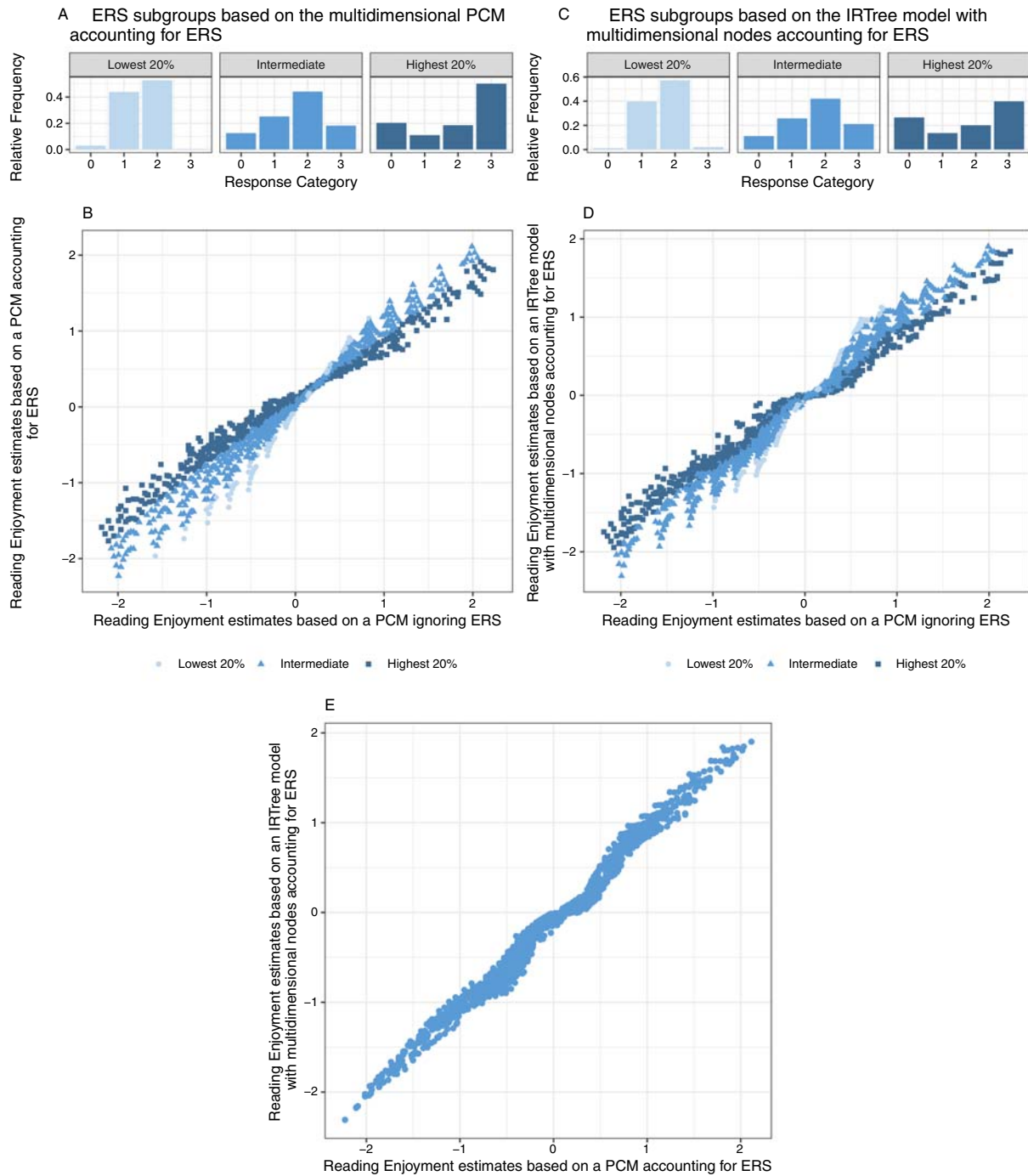


Fig. 3 Panel A: relative frequencies of response category choices for lower and upper 20% quantiles and intermediate levels of Extreme Response Style (ERS) based on ERS estimates in Model 2; Panel B: correction of “Reading Enjoyment” estimates when ERS is accounted for by Model 2 (Multidimensional PCM with ERS dimension). Panel C and D: as Panels A and B but based on the ERS estimates in Model 4; Panels E: “Reading Enjoyment” estimates when ERS is accounted for by Model 4 (IRTTree model with multidimensional nodes) compared to a PCM accounting for ERS.

both the substantive traits and ERS in the extremity node, Model 4 accomplishes a clear distinction between the trait and ERS factors: $\hat{\rho} = .46$ for the correlation among the two reading dimensions, and $\hat{\rho} = .002$ and $\hat{\rho} = -.006$ for the correlations of the two reading dimensions with the latent ERS dimension.

Panel C in Fig. 3 shows how different subgroups of respondents with varying ERS levels in the IRTree Model 4 with multidimensional nodes differ in their category choices. Panel D shows the de-biasing role of Model 4 with respect to the “Reading Enjoyment” estimates in comparison to the PCM ignoring ERS. As both model correct the estimates for the biasing effect of ERS, the frequencies of category choices (Panels A and C) as well as the correction of trait estimates (Panels B and D) are highly similar between the two

modeling approaches. The similarity can also be seen in Panel E, where the “Reading Enjoyment” estimates from the multidimensional PCM accounting for ERS and the IRTree model with multidimensional nodes are directly compared to each other.

Comparison of extended PCMs and IRTrees

The extended PCM and the IRTree models have shown substantial variance of the latent ERS dimension, attesting to the role of the response tendency across the items of the two reading subscales. In addition, Fig. 3 has illustrated how ERS influences response behavior, and that extended PCMs and IRTree models allow researchers to correct construct estimates from biasing ERS effects.

Although the two approaches to modeling response styles have several commonalities and differences, a direct comparison of multidimensional PCMs and IRTrees for response styles is precluded by the different structures of the likelihood functions and by different kinds of restrictions. Looking at the multidimensional PCM Model 2 and the IRTree Model 4 with multidimensional nodes, Model 2 is more restrictive in terms of equal discrimination parameters across categories and thresholds, whereas Model 4 allows for different weights of θ for extreme disagreement and extreme agreement. In contrast, Model 4 is more restrictive than Model 2 in terms of equal threshold parameters for extreme responses at either side of the rating scale, whereas the multidimensional PCM does not include restrictions on the threshold parameters.

Beyond the specific model definitions and restrictions used in the present illustration with 4-point rating data, both multidimensional PCMs and IRTree models form flexible modeling frameworks that allow researchers to specify a range of theoretical assumptions for a variety of rating response formats.

Discussion

We have presented and illustrated two different modeling classes that incorporate and control for response styles: multidimensional extensions of ordinal IRT models and IRTree modeling approaches. Both modeling classes have advantages and disadvantages with respect to the modeling of response styles that we would like to highlight and discuss.

An advantage of multidimensional extensions of ordinal models, such as PCMs, is that they retain the ordinal scoring for content traits. Hence, they allow that, for instance, a response in the “strongly agree” category is influenced by both, a positive level of the construct to be measured and a positive level of ERS. Another advantage is that multidimensional PCMs do not require to restructure the data, as scoring weights can be used to specify how each of the categories is related to the latent dimensions. Besides, scoring weights allow the applicant of this model family to flexibly specify additional response style dimensions to be accounted for. In this chapter, we presented scoring weights for the most commonly investigated response styles ERS and MRS, but other types such as ARS (see Falk and Cai, 2016; Henninger and Meiser, 2020b), varying thresholds (Henninger, 2020; Wang et al., 2006) or socially desirable responding (Falk and Cai, 2016) can also be incorporated into the IRT model with specific definitions of scoring weights. Please note that when scoring weights are fixed in a way that they are not redundant to the construct of interest, the covariance matrix between constructs and response style dimensions can be estimated and interpreted to evaluate whether the constructs of interest are related to certain response tendencies.

Similar to multidimensional PCMs, IRTree models provide the opportunity to flexibly specify different subprocesses through recoding of the original responses into specific pseudo-items. Therewith, IRTrees cannot only account for ERS or MRS, but also for other response processes such as not-reached or omitted responses (De Boeck and Partchev, 2012). Like in multidimensional extensions of PCMs, the latent covariance matrix between the trait and response style dimensions can be estimated and interpreted. An advantage of IRTrees over multidimensional PCM is the direct interpretation of person and item parameters for each subprocess, allowing for example to determine which items specifically foster the use of the extreme or middle categories. A potential disadvantage of IRTrees may be seen in the dichotomization of responses in the definition of pseudo-items pertaining to agreement judgments or other subprocesses, which may reduce the precision of estimating the underlying construct (but see Plieninger and Meiser, 2014). Recent development of IRTrees, however, allow for ordinal pseudo-items and for assessing the same construct over several binary pseudo-items (e.g. Meiser et al., 2019), which remedies the potential disadvantage as demonstrated in our empirical application with a multidimensional node parameterization.

A particular challenge of IRTrees compared to multidimensional PCMs concerns the modeling of response styles that go along with the target trait, such as ARS. Modeling ARS has been implemented in IRTrees with the aid of a Bayesian hierarchical estimation approach to multinomial processing trees (Pleninger and Heck, 2018), but it has not yet been implemented in a version that is easily accessible to applied researchers. Lastly, we would like to caution against a strict interpretation of the order of processes in IRTrees, which may be suggested by tree diagrams: The order of processes in an IRTree diagram does not imply a temporal sequence, but it entails logical dependencies as reflected in the conditional definition of pseudo-items. Besides other caveats and potential pitfalls, Pleninger (2020a) provides an extensive discussion of when and why the order of processes should not be interpreted and which logical dependencies may be important.

While in this chapter, we have used one-parameter variants of PCMs and IRTrees with discrimination parameters fixed to 1 or constrained to be equal across rating items, the response style models can be extended to two-parameter variants with discrimination parameters varying over rating items. For the dimensions reflecting the construct of interest, the discrimination parameters can be interpreted as the strength of the influence of the construct on the item response. Analogously, discrimination parameters of response style dimensions reflect the extent to which single items are impacted by response style dimensions. Therefore, item-specific discrimination parameters of response style dimensions might be helpful to identify items that are influenced by response styles to a larger or smaller extent (see e.g., Henninger and Meiser, 2020b, for an application, discussion, and explanatory model).

Response style models can also be extended to account for different strengths or types of response styles across groups or subpopulations, such as in cross-cultural research (see e.g., Bolt et al., 2014; Henninger, 2020; Morren et al., 2012; Plieninger, 2020a, for examples) or mixture distribution analyses (Kim and Bolt, 2021; Tijmstra et al., 2018). Those extensions of response style models, allow us to compare latent means and variances, not only of the constructs to be measured, but also of response style dimensions. In consequence, they might provide interesting insights into how response processes differ between countries, socio-demographic groups or latent subpopulations of respondents.

In our empirical example, we used the free, open-source programming environment R (R Core Team, 2020) together with the package *mirt* (Chalmers, 2012, OSF⁴ repository for example R code), but there are multiple software alternatives that can be used for estimating response style models. For instance, the R package *TAM* (Kiefer et al., 2017) provides flexible options to specify diverse psychometric models using design matrices (see Henninger, 2020; Henninger and Meiser, 2020b, for example code). To fit IRTree models to polytomous data, the R packages *IRTrees* (De Boeck and Partchev, 2012) or *ItemResponseTrees* (Plieninger, 2020b) provide user-friendly functions that facilitate the creation of pseudo-items. Alternatively, the commercial software program *Mplus* (Muthén and Muthén, 2012) can be used to specify multidimensional extensions of ordinal IRT models as well as IRTree approaches (for examples of *Mplus* syntax, see Böckenholt and Meiser, 2017; Meiser et al., 2019). As an alternative, multidimensional IRT and IRTree models can be estimated in a Bayesian estimation framework, as discussed, for example, by Leventhal and Stone (2018), Bolt et al. (2014) or Plieninger and Heck (2018). Independent of the choice of software, modeling response styles across different constructs, or using additional heterogeneous items or anchoring vignettes (Bolt et al., 2014), facilitates estimation of these complex psychometric models and enhances the interpretation of latent factors.

Aside from using statistical models to control for response styles in the measurement of latent propensities, recent developments in psychometrics have considered alternative response formats as means to reduce biasing effects, such as response formats that distinguish categorical agreement judgments from ratings of response intensity (Böckenholt, 2017) or rankings of items pertaining to different latent constructs (so-called multidimensional forced-choice format; Wetzel and Frick, 2020). The combination of statistical modeling in data analysis and innovative response formats in test construction allows researchers to enhance the validity of educational measurement and to better specify and test theoretical assumptions on latent response processes.

References

- Adams, D.J., Bolt, D.M., Deng, S., Smith, S.S., Baker, T.B., 2019. Using multidimensional item response theory to evaluate how response styles impact measurement. *Br. J. Math. Stat. Psychol.* 72, 466–485. <https://doi.org/10.1111/bmsp.12169>.
- Andrich, D., 1978. A rating formulation for ordered response categories. *Psychometrika* 43 (4), 561–573. <https://doi.org/10.1007/BF02293814>.
- Bachman, J.G., O'Malley, P.M., 1984. Yea-saying, nay-saying, and going to extremes: black-white differences in response styles. *Publ. Opin. Q.* 48, 491–509. <https://doi.org/10.1086/268845>.
- Baumgartner, H., Steenkamp, J.-B.E., 2001. Response styles in marketing research: a cross-national investigation. *J. Market. Res.* 38 (2), 143–156. <https://doi.org/10.1509/jmkr.38.2.143.18840>.
- Bock, R.D., 1972. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37, 29–51. <https://doi.org/10.1007/BF02291411>.
- Böckenholt, U., Meiser, T., 2017. Response style analysis with threshold and multi-process IRT models: a review and tutorial. *Br. J. Math. Stat. Psychol.* 70, 159–181. <https://doi.org/10.1111/bmsp.12086>.
- Böckenholt, U., 2012. Modeling multiple response processes in judgment and choice. *Psychol. Methods* 17, 665–678. <https://doi.org/10.1037/a0028111>.
- Böckenholt, U., 2017. Measuring response styles in Likert items. *Psychol. Methods* 22, 69–83. <https://doi.org/10.1037/met0000106>.
- Böckenholt, U., 2019. Assessing item-feature effects with item response tree models. *Br. J. Math. Stat. Psychol.* 72 (3), 486–500. <https://doi.org/10.1111/bmsp.12163>.
- Bolt, D.M., Johnson, T.R., 2009. Addressing score bias and differential item functioning due to individual differences in response style. *Appl. Psychol. Meas.* 33, 335–352. <https://doi.org/10.1177/0146621608329891>.
- Bolt, D.M., Newton, J.R., 2011. Multiscale measurement of extreme response style. *Educ. Psychol. Meas.* 71 (5), 814–833. <https://doi.org/10.1177/0013164410388411>.
- Bolt, D.M., Lu, Y., Kim, J.-S., 2014. Measurement and control of response styles using anchoring vignettes: a model-based approach. *Psychol. Methods* 19, 528–541. <https://doi.org/10.1037/met0000016>.
- Chalmers, R.P., 2012. *mirt*: a multidimensional item response theory package for the R environment. *J. Stat. Software* 48 (6), 1–29. <https://doi.org/10.18637/jss.v048.i06>.
- Couch, A., Keniston, K., 1960. Yeasayers and naysayers: agreeing response set as a personality variable. *J. Abnorm. Soc. Psychol.* 60, 151–174. <https://doi.org/10.1037/h0040372>.
- Cronbach, L.J., 1942. Studies of acquiescence as a factor in the truefalse test. *J. Educ. Psychol.* 33 (6), 401–415. <https://doi.org/10.1037/h0054677>.
- De Boeck, P., Partchev, I., 2012. IRTrees: tree-based item response models of the GLMM family. *J. Stat. Software* 48, 1–28. <https://doi.org/10.18637/jss.v048.c01>.
- De Jong, M.G., Steenkamp, J.-B.E., Fox, J.-P., Baumgartner, H., 2008. Using item response theory to measure extreme response style in marketing research: a global investigation. *J. Market. Res.* 45, 104–115. <https://doi.org/10.1509/jmkr.45.1.104>.
- Falk, C.F., Cai, L., 2016. A flexible full-information approach to the modeling of response styles. *Psychol. Methods* 21, 328–347. <https://doi.org/10.1037/met0000059>.
- Greenleaf, E.A., 1992. Measuring extreme response style. *Publ. Opin. Q.* 56, 328–351. <https://doi.org/10.1086/269326>.
- Hamilton, D.L., 1968. Personality attributes associated with extreme response style. *Psychol. Bull.* 69, 192–203. <https://doi.org/10.1037/h0025606>.
- Henninger, M., Meiser, T., 2020a. Different approaches to modeling response styles in Divide-by-Total IRT models (Part I): a model integration. *Psychol. Methods* 25, 560–576. <https://doi.org/10.1037/met0000249>.
- Henninger, M., Meiser, T., 2020b. Different approaches to modeling response styles in Divide-by-Total IRT models (Part II): applications and novel extensions. *Psychol. Methods* 25, 577–595. <https://doi.org/10.1037/met0000268>.
- Henninger, M., 2020. A novel Partial Credit extension using varying thresholds to account for response styles. *J. Educ. Meas.* <https://doi.org/10.1111/jedm.12268>.
- Hui, C.H., Triandis, H.C., 1985. The instability of response sets. *Publ. Opin. Q.* 49, 253–260. <https://doi.org/10.1086/268918>. *Ausgedruckt*.
- Jeon, M., De Boeck, P., 2016. A generalized item response tree model for psychological assessments. *Behav. Res. Methods* 48, 1070–1085. <https://doi.org/10.3758/s13428-015-0631-y>.

- Khorramdel, L., von Davier, M., 2014. Measuring response styles across the Big Five: a multiscale extension of an approach using multinomial processing trees. *Multivariate Behav. Res.* 49, 161–177. <https://doi.org/10.1080/00273171.2013.866536>.
- Kiefer, T., Robitzsch, A., Wu, M., 2017. TAM: test analysis modules. Version 2.8-21. <http://cran.r-project.org/package=TAM>.
- Kim, N., Bolt, D.M., 2021. A mixture IRTree model for extreme response style: accounting for response process uncertainty. *Educ. Psychol. Meas.* 81, 131–154. <https://doi.org/10.1177/0013164420913915>.
- Leventhal, B.C., Stone, C.A., 2018. Bayesian analysis of multidimensional item response theory models : a discussion and illustration of three response style models. *Measurement: Interdisciplinary Research and Perspectives* 16 (2), 114–128. <https://doi.org/10.1080/15366367.2018.1437306>.
- Masters, G.N., 1982. A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174. <https://doi.org/10.1007/BF02296272>.
- Meiser, T., Machunsky, M., 2008. The personal structure of personal need for structure: a mixture-distribution Rasch analysis. *Eur. J. Psychol. Assess.* 24, 27–34. <https://doi.org/10.1027/1015-5759.24.1.27>.
- Meiser, T., Plieninger, H., Henninger, M., 2019. IRTree models with ordinal and multidimensional decision nodes for response styles and trait-based rating responses. *Br. J. Math. Stat. Psychol.* 72 (3), 501–516. <https://doi.org/10.1111/bmsp.12158>.
- Moors, G., 2003. Diagnosing response style behavior by means of a latent class factor approach. Socio-demographic correlates of gender role attitudes and perceptions of ethnic discrimination reexamined. *Qual. Quant.* 37, 277–302. <https://doi.org/10.1023/A:1024472110002>.
- Moors, G., 2004. Facts and artefacts in the comparison of attitudes among ethnic minorities. A multigroup latent class structure model with adjustment for response style behavior. *Eur. Socio Rev.* 20, 303–320. <https://doi.org/10.1093/esr/jch026>. NULL.
- Morren, M., Gelissen, J.P., Vermunt, J.K., 2012. The impact of controlling for extreme responding on measurement equivalence in crosscultural research. *Methodology* 8, 159–170. <https://doi.org/10.1027/1614-2241/a000048>. NULL.
- Muthén, L.K., Muthén, B.O., 2012. Mplus user's guide (7th). Muthén & Muthén, Los Angeles, CA.
- Paulhus, D.L., 1991. Measurement and control of response bias. In: Robinson, J.P., Shaver, P.R., Wrightsman, L.S. (Eds.), *Measures of Personality and Social Psychological Attitudes*. Academic Press, San Diego, CA, pp. 17–59. <https://doi.org/10.1016/B978-0-12-590241-0.50006-X>.
- Plieninger, H., Heck, D.W., 2018. A new model for acquiescence at the interface of psychometrics and cognitive psychology. *Multivariate Behav. Res.* 53, 633–654. <https://doi.org/10.1080/00273171.2018.1469966>.
- Plieninger, H., Meiser, T., 2014. Validity of multiprocess IRT models for separating content and response styles. *Educ. Psychol. Meas.* 74, 875–899. <https://doi.org/10.1177/0013164413514998>.
- Plieninger, H., 2017. Mountain or molehill? A simulation study on the impact of response styles. *Educ. Psychol. Meas.* 77, 32–53. <https://doi.org/10.1177/0013164416636655>.
- Plieninger, H., 2020a. Developing and applying IR-Tree models : guidelines, caveats, and an extension to multiple groups. *Organ. Res. Methods* 1–17. <https://doi.org/10.1177/109428120911096>.
- Plieninger, H., 2020b. ItemResponseTrees. R package version 0.2.5. <https://cran.r-project.org/web/packages/ItemResponseTrees/index.html>.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org>.
- Rost, J., Carstensen, C.H., von Davier, M., 1999. Sind die Big Five Rasch-skalerbar? Eine Reanalyse der NEO-FFI-Normierungsdaten./Are the Big Five Rasch scaleable? A reanalysis of the NEO-FFI norm data. *Diagnostica* 45 (3), 119–127. <https://doi.org/10.1026/0012-1924.45.3.119>.
- Rost, J., 1991. A logistic mixture distribution model for polychotomous item responses. *Br. J. Math. Stat. Psychol.* 44, 75–92. <https://doi.org/10.1111/j.2044-8317.1991.tb00951.x>.
- Samejima, F., 1969. Estimation of latent ability using a response pattern of graded scores [Monograph]. *Psychometrika* 34 (Suppl. 17), 1–100. https://www.psychometricsociety.org/sites/default/_les/pdf/MN17.pdf.
- Takane, Y., de Leeuw, J., 1987. On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika* 52, 393–408. <https://doi.org/10.1007/BF02294363>.
- Thissen, D., Steinberg, L., 1986. A taxonomy of item response models. *Psychometrika* 51, 567–577. <https://doi.org/10.1007/BF02295596>. NULL.
- Tijmstra, J., Bolsinova, M., Jeon, M., 2018. General mixture item response models with different item response structures: exposition with an application to Likert scales. *Behav. Res. Methods* 50, 2325–2344. <https://doi.org/10.3758/s13428-017-0997-0>.
- van Herk, H., Poortinga, Y.H., Verhallen, T.M.M., 2004. Response styles in rating scales: evidence of method bias in data from six EU countries. *J. Cross Cult. Psychol.* 35, 346–360. <https://doi.org/10.1177/0022022104264126>.
- Van Vaerenbergh, Y., Thomas, T.D., 2013. Response styles in survey research: a literature review of antecedents, consequences, and remedies. *Int. J. Publ. Opin. Res.* 25, 195–217. <https://doi.org/10.1093/ijpor/eds021>.
- von Davier, M., Khorramdel, L., 2013. Differentiating response styles and construct-related responses: a new irt approach using bifactor and second-order models. In: Millsap, R.E., van der Ark, L.A., Bolt, D.M., Woods, C.M. (Eds.), *New Developments in Quantitative Psychology*. Springer, New York, pp. 463–487.
- Wang, W.-C., Wilson, M., Shih, C.-L., 2006. Modeling randomness in judging rating scales with a random-effects rating scale model. *J. Educ. Meas.* 43, 335–353. <https://doi.org/10.1111/j.1745-3984.2006.00020.x>.
- Weijters, B., Geuens, M., Schillewaert, N., 2010a. The individual consistency of acquiescence and extreme response style in self-report questionnaires. *Appl. Psychol. Meas.* 34, 105–121. <https://doi.org/10.1177/0146621609338593>.
- Weijters, B., Geuens, M., Schillewaert, N., 2010b. The stability of individual response styles. *Psychol. Methods* 15, 96–110. <https://doi.org/10.1037/a0018721>.
- Wetzel, E., Carstensen, C.H., 2017. Multidimensional modeling of traits and response styles. *Eur. J. Psychol. Assess.* 33, 352–364. <https://doi.org/10.1027/1015-5759/a000291>.
- Wetzel, E., Frick, S., 2020. Comparing the validity of trait estimates from the multidimensional forced-choice format and the rating scale format. *Psychol. Assess.* 32, 239–253. <https://doi.org/10.1037/pas0000781>.
- Wetzel, E., Böhne, J.R., Carstensen, C.H., Ziegler, M., Ostendorf, F., 2013a. Do individual response styles matter? Assessing differential item functioning for men and women in the NEO-PI-R. *J. Indiv. Differ.* 34, 69–81. <https://doi.org/10.1027/1614-0001/a000102>.
- Wetzel, E., Carstensen, C.H., Böhne, J.R., 2013b. Consistency of extreme response style and non-extreme response style across traits. *J. Res. Pers.* 47, 178–189. <https://doi.org/10.1016/j.jrp.2012.10.010>.
- Zettler, I., Lang, J.W.B., Hülshager, U.R., Hilbig, B.E., 2016. Dissociating indifferent, directional, and extreme responding in personality data: applying the three-process model to self- and observer reports. *J. Pers.* 84, 461–472. <https://doi.org/10.1111/jopy.12172>.

Simulation studies for psychometrics

Brian C. Leventhal^a, Allison J. Ames^b, and Kathryn N. Thompson^a, ^a James Madison University, Harrisonburg, VA, United States; and ^b University of Arkansas, Fayetteville, AR, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to psychometric simulation	341
Step 1: Specifying the research Question(s)	342
Step 2: Defining and justifying conditions	342
Step 3: Specifying the experimental design and outcomes of interest	343
Step 4: Simulating data under the specified conditions	343
Step 5: Estimating parameters	344
MCMC estimation considerations	344
Maximum likelihood estimation considerations	344
Step 6: Comparing true and estimated parameters	344
Step 7: Replicating the procedure a specified number of times	345
Step 8: Analyzing results based on the design and research question	345
Example	345
Conclusion	346
References	346

Introduction to psychometric simulation

Psychometric research questions may be answered by theoretical derivations, empirical data collection, or simulation-based methods. Theoretical studies typically derive new methods or evaluate current methods mathematically. Empirical studies require the collection of data or analysis of secondary data. Although in a majority of studies using psychometric methods data is collected through the use of achievement tests or attitudinal measures to help answer important questions, these empirical studies are unable to answer all questions of interest to psychometric researchers. For example, questions in a parameter recovery study investigating the accuracy of estimating an item discrimination parameter using a new estimation routine, or questions in a comparative study that focus on Type I error rates of a model-fit statistic are not answerable with empirical methods. This is because in empirical studies, the researcher does not know “truth”, such as the true population parameter values, or whether the mechanism giving rise to the data is the same as the model being fit to the data, each of which are necessary for answering these types of methodological research questions. As an alternative to empirical approaches, simulation studies address a unique class of research questions that require knowing the truth, such as those that focus on parameter recovery or a comparison of methods. In answering these questions, Monte Carlo simulation studies, often simply called “simulation studies”, use data generated by the researcher in a carefully planned experiment.

Simulation studies are prevalent in recent psychometric literature. In five major educational and psychological measurement journals from the years 2015 through 2019, at least one-quarter of published articles utilized simulation methods, as shown in **Table 1**. Furthermore, of the five journals reviewed, the majority of articles were simulation based in four, with *Applied Psychological Measurement* publishing the largest percentage of simulation studies (86.24%), followed by *Journal of Educational and Behavioral Statistics* (74.29%), *Journal of Educational Measurement* (55.56%), and *Educational and Psychological Measurement* (51.97%). These simulation entries come from a wide variety of research areas, such as structural equation modeling, hierarchical linear modeling/multilevel modeling, differential item functioning, computer adaptive testing, cognitive diagnostic models, response styles, with item response theory being the most common across all. It is also quite common for publications to include studies that contained both empirical and simulation components, such as those in the *Journal of Educational and Behavior Statistics* where a majority of published studies contained both simulation and empirical parts.

One reason for the popularity and prevalence of Monte Carlo simulations is that a simulation study facilitates answering questions not possible using empirical methods as data is generated under the control of the psychometric researcher. Specification of data generation models under a variety of conditions allows the researcher to understand psychometric model behavior in varying circumstances, helping provide context for situations where the model is applied empirically. For example, survey responses may be generated using a structural equation model with varying conditions of sample size or patterns in factor loadings. Specifying the levels of these conditions is critical as the generalizability of the results are a direct consequence of the conditions selected. Simulation researchers attest to the need to balance the number of factors in a study, the levels of each of those factors, and the generalizability of the results.

In addition to studies answering simulation-based research questions, multiple publications provide guidance to psychometric researchers performing simulation studies (e.g., [Harwell et al., 1996](#); [Lee, 2015](#); [Feinberg and Rubright, 2016](#)). Other publications are tutorial-focused, providing application based guidance to performing simulation (e.g., [Ames et al., 2020](#); [Beaujean, 2018](#);

Table 1 Percentage of simulation entries in educational measurement journals.

<i>Journal</i>	<i>Total publications</i>	<i>Simulation study only %</i>	<i>Simulation & empirical studies %</i>
Applied Measurement in Education	119	11	9
Applied Psychological Measurement	189	54	32
Educational and Psychological Measurement	254	38	14
Journal of Educational and Behavioral Statistics	105	8	67
Journal of Educational Measurement	135	26	30

Leventhal and Ames, 2020) as well as details to performing simulation studies for different classes of models (e.g., Luecht and Ackerman, 2018). As with empirical studies, most simulation studies have the same key components, even though specifics will vary among studies. Harwell et al.'s (1996) eight steps for simulation studies provide a viable framework for simulation studies in psychometrics. This eight-step framework includes: (1) specifying the research question(s), (2) defining and justifying conditions, (3) choosing an experimental design, (4) generating data, (5) estimating parameters, (6) comparing true and estimated parameters, (7) replicating the procedure a specified number of times, and (8) analyzing results.

Step 1: Specifying the research Question(s)

Questions that can be answered using simulation techniques are classified as one of three types: (1) parameter recovery studies, (2) comparative studies, and (3) tool-based where the simulation techniques are used to enhance empirical-based methods. Parameter recovery studies typically focus on the robustness of statistical procedures, estimators, or test statistics. Common outcomes of parameter recovery studies are parameter estimate bias, efficiency, and consistency. Comparative studies are used to evaluate error rates (e.g., Type I, Type II), statistical power, and model comparison procedures. Even though the term “simulation studies” is commonly used for either parameter recovery or comparative studies, simulation techniques are useful in other areas of psychometric research. For example, simulation techniques assist in generating empirical sampling distributions under a null hypothesis in a frequentist framework, posterior predictive model checking in a Bayesian framework, qualities of interest reporting, model cross validation studies, and resampling methods such as bootstrapping, jackknifing, and permutation tests. Additionally, simulation techniques can be particularly helpful in answering research questions that would alternatively stretch ethical boundaries when using human subjects. For example, researchers may want to investigate the impact of a technology disruption on a computer adaptive test in a high-stakes setting. Because it would be unethical to disrupt an exam session in real life, researchers may turn to simulation techniques to gauge the impact of the disruption on ability estimation.

Step 2: Defining and justifying conditions

Simulation studies require delicately balancing the comprehensive nature of the design, feasibility of the study, and generalizability of the results. To help, the factors selected should meaningfully impact practice, necessitating an examination of their appropriateness in each study on a case-by-case basis. Factors of the simulation are analogous to independent variables of an experiment: they are manipulated by the researcher in order to examine their effect on the outcome (i.e., dependent variable). For example, the model selected to generate and analyze the data, typically dictated by the research questions, must be worthwhile of study with broad enough application for interest beyond the scope of simulation. Results from the simulation study apply only to the specific models used and parameter values selected to generate data. Therefore, in addition to justifying the model, parameter values used for generation (i.e. true parameter values) must be grounded in literature or practice.

In order to generate data, model parameters are considered known, or “true”. Three ways of selecting these true generating parameters exist. First, generating parameters may be fixed at specified values within a range. For example, item difficulties in an IRT framework are set to equally spaced values between -2 and 2 . Second, model parameters may be randomly selected from a range of values such as item difficulties being randomly drawn from a uniform distribution of values between -2 and 2 . Third, generating parameter are based on empirical results such as item difficulties fixed to empirical estimates from a prior or paired study. In most simulations, it is common that multiple methods are employed. Continuing the IRT example, item difficulties may be randomly drawn from a range of empirical estimates. Whichever technique is used to select these true parameters, the method used with a sound justification grounded in literature or a paired study must be explicit in the manuscript. This ensures reproducibility and helps contextualize the generalizability of the results. In the case where multiple models are used to generate and analyze data, careful consideration must be made to the selection of model parameter values to assure comparability. In these examples, the generating model or analysis model become independent variables of the study.

Step 3: Specifying the experimental design and outcomes of interest

Simulation studies commonly use factorial designs, where independent variables, known as simulation factors; dependent variables, known as outcomes of interest; and simulation constants, known as fixed factors, are specified and justified. A combination of levels from all factors is called a simulation condition. Simulation studies are defined by the size of the factorial designs or the number of simulation conditions. For example, a $3 \times 2 \times 5$ simulation study has three independent variables with 3, 2, and 5 levels, respectively, resulting in 30 simulation conditions. Interpretability of the factorial design must be balanced with generalizability of the results. Partly due to this balance, in Monte Carlo simulations it is equally important to specify and justify fixed factors. Fixed factors limit the generalizability of the results so their use must add to usefulness of the study in some other way. To justify fixed conditions, previous research and theory can be beneficial. For example, if investigating the effects of estimating item difficulties, it may be worth fixing test length to make the study feasible, but include sample size as a simulation factor with multiple levels to improve generalizability. Or, if a specific empirical scenario is of interest to the researcher, conditions from the empirical study can be used as fixed conditions in the simulation.

Factor selection is critical, and is influenced by previous literature in the domain, research questions of the study, a paired empirical study, or application interests. For example, SEM studies may focus on manipulating estimation methods, number of latent factors, or factor loading patterns while IRT studies may focus on varying the number of items, sample size, or true ability distributions. In classical test theory simulations, many of these same factors might be considered important by researchers (e.g., sample size), but other factors like variability of the observed score distribution may also be manipulated. Once factors are selected, each level must be meaningful and justified. As an illustration, unless a specific case can be made, there likely is no practical difference between a test with 10 items and a test with 12 items. There may, however, be a meaningful difference between a test with 10 items and one with 20 items. More general interpretations, such as the effects on the outcome when test length is doubled, can be drawn if the factor levels are sufficiently distinct.

Factor level selections typically relate to properties and assumptions of estimators. If trying to estimate latent ability in IRT, for example, test length should be included as a factor rather than sample size, but sample size must be fixed to be large enough to assure no confounding effects of low sample size. Some experts consider including levels where good results are to be expected such as a large number of items when estimating latent ability. This can serve as a baseline condition although inclusion needs to be balanced with practicality of the study.

Simulation factors also depend on the estimation framework. The two primary estimation frameworks used in psychometrics are Bayesian Markov chain Monte Carlo (MCMC) and frequentist Maximum Likelihood estimation (MLE). Even though both methods require specification of the analytic and generating models, which may differ from each other, with the stochastic model often referred to as the likelihood, Bayesian MCMC methods differ from MLE in many ways.

One primary difference is in the use of prior distributions, which are hypotheses about the shape of the distribution of parameter values. For each parameter, the prior distribution and the likelihood combine to form the posterior distribution. In simplified terms, the MCMC process updates the prior distribution and yields a new probability distribution, reflecting plausible values of the parameter. When the prior is combined with the likelihood, each value in the likelihood is weighted by its prior probability. With little to no prior information contained in the prior distribution, each value in the likelihood is equally weighted, and the posterior distribution is nearly identical to the likelihood function. However, with more prior information (i.e., the researcher uses informative priors), values in the likelihood receive more weight where the prior has more density and less weight where the prior has a lower density. The result is a posterior that is shrunk toward the prior. Further, as sample size increases, the prior becomes less and less influential. Because many Monte Carlo simulation studies vary sample size as a condition, priors may be influential in some levels of the condition, but not influential in others. Because of this interaction, the choice of prior may be included as an additional factor (e.g., [Natesan et al., 2016](#)) when sample size is manipulated.

Regardless of estimation framework, the outcomes of interest are dictated by the type of simulation study. For parameter recovery studies, these dependent variables commonly reflect parameter estimates, empirical error rates, or sampling distributions of parameter estimates. Alternatively, comparative studies examine test statistics and model-fit statistics to calculate error rates or statistical power. For other studies that use simulation techniques, outcomes of interest widely vary. In one example, the study of an estimation algorithm may focus on the number of iterations it takes the method to converge under varying factors or the resulting number of improper solutions (e.g., negative variance estimates in structural equation modeling). Specific details of calculating these outcomes does depend on the estimation framework used as highlighted in steps 5 and 6.

Step 4: Simulating data under the specified conditions

Methods and software used to generate data should be adequately described in a manuscript to allow replication of the study. If programs are developed, or common routines are modified, they must be verified for accuracy. All properties of generated data must be validated to confirm simulation conditions, both manipulated and fixed, are accurately reflected in the data. This is known as data generation validation and is analogous to mitigating and documenting validity concerns during data collection in an empirical study. Validation may be performed by using previously well-documented estimation routines or through graphical means. For example, in IRT, a researcher may simulate a large number of individuals (whether or not it is a condition), estimate item parameters with a known estimation algorithm or software package (e.g., marginal maximum likelihood in FlexMIRT), and compare the true and

estimated item parameters to make sure the data generation process is accurate. Alternatively, validation can be performed graphically. In some cases, a simple histogram can be used, such as visually confirming that a latent trait follows a normal or skewed distribution as intended. Data validation is not commonly reported in full detail, but all key aspects of the data generation must be validated just as any data collection procedure would be validated in an empirical study.

Step 5: Estimating parameters

After the data is generated and validated, the next step is to fit the analysis model(s) to the data. There may be a single analysis model if performing a parameter recovery study or multiple analysis models if performing comparative studies. Fit of analysis models in parameter recovery studies results in estimated parameters of interest, whereas estimation of models in comparative studies result in, for example, the calculation of model-fit statistics. Each simulation should document the exact estimation methods, providing details of the estimation algorithm and software choice. Simulation studies using either MCMC methods or MLE must assess convergence. MCMC methods may rely on visual tools, such as trace plots, or statistical measures, such as Geweke (1992) and the Gelman-Rubin diagnostic (Gelman and Rubin, 1992). MLE algorithms may use tolerance. Both methods should specify the number of iterations necessary to achieve convergence, which helps the simulation audience understand the applicability and generalizability of the results.

MCMC estimation considerations

While there are few guidelines for Bayesian simulation studies, several authors have provided reporting guidelines on the details of MCMC (e.g., van Doorn et al., 2019; Sung et al., 2005). Many of these same guidelines apply to reporting Bayesian simulation studies and they reflect the decisions that must be made by the researcher in setting up and conducting the Monte Carlo simulation. For example, Sung et al. (2005) suggest reporting the exact prior distribution specifications, as well as providing a justification for the choice of priors. In terms of computational issues, authors should report the specific software choice, MCMC sampling algorithm, and a range of other technical details (Sung et al., 2005). Specific MCMC algorithm examples include: Metropolis Hastings (Metropolis et al., 1953; Hastings, 1970), implemented in the SAS MCMC procedure (SAS Institute Inc., 2015); Gibbs (Geman and Geman, 1984), implemented in Mplus (Muthén and Muthén, 1998–2017); and Hamiltonian Monte Carlo (Duane et al., 1987), implemented in Stan (Stan Development Team, 2017). Because Bayesian estimation results in a full parameter posterior distribution, many researchers choose to summarize the posterior using a point estimate, such as the posterior mode (i.e., Maximum a Posteriori, MAP) or mean (i.e., Expected a Posteriori, EAP). The posterior is a probability distribution; as such, a single value may not be adequate to reflect the estimated parameter and an indicator of posterior precision is critical.

Maximum likelihood estimation considerations

MLE-specific algorithms used in psychometrics include Marginal MLE (Bock and Aitkin, 1981) and joint MLE (Andersen, 1972; Fischer, 1981; Haberman, 1977). Despite the differences in algorithms and underlying philosophies between Bayesian and MLE methods, for many of the important details in Bayesian MCMC estimation, there is a MLE equivalent. For example, the precision of the point estimate in MCMC estimation is summarized using quantities such as posterior highest density interval or the posterior standard deviation. In MLE, the parameter standard error is used to summarize the precision of the parameter.

Step 6: Comparing true and estimated parameters

Comparison depends on the type of simulation study and specific research questions. Feinberg and Rubright (2016) state, “It is also important to consider how simulations were evaluated in previous research, such that similar dependent variables can be applied to facilitate comparisons and allow results from your simulation to build upon previous research” (p. 40). If the study is a parameter recovery study, the research question is asking if the estimation routine is accurate across multiple levels of each condition. In frequentist studies relying on MLE, the parameter point estimate is compared to the generating value. These types of simulations use metrics like bias, standard error, root mean square error (RMSE), and absolute bias to determine accuracy and precision. The deviation between the estimated parameter and true generating parameter is the fundamental unit for many parameter recovery studies. Frequentist studies may also look at interval estimation, such as the number of confidence intervals that capture the generating value. When using MCMC methods, researchers commonly use a point estimate like the EAP or MAP for this comparison to compute the deviation. Coverage, a measure of whether a specified posterior interval encompasses the generating value, may be used in an effort to avoid discarding information present in the full posterior that is missed by the using of only a point estimate.

If the purpose of the simulation is to evaluate a new procedure or to compare models, additional measures need to be computed beyond the parameter point estimates. For example, McKinley and Mills (1985) evaluated four goodness-of-fit procedures for item response models on their ability to discriminate cases of fit from cases of misfit. The comparison in this type of study involves three conditions: (1) the erroneous conclusion that the analysis model does not fit the data when, in fact, the data were generated with that model; (2) the erroneous conclusion of fit when the analysis model is not a restricted form of the generating model; and (3) the

correct conclusion that the analysis model does not fit the data when the data were generated with a more complex version. Simulation studies often calculate the prevalence of each condition. Because MCMC studies do not rely upon null hypothesis statistical testing procedures, equivalent quantities might be “hit rates” for correct conclusions and “false positives” for incorrect conclusions.

Step 7: Replicating the procedure a specified number of times

Replicating within conditions allows for more generalizable simulation results through introduction of variability in generating parameter values and their estimates. [Feinberg and Rubright \(2016\)](#) provide recommendations on the number of replications for a simulation study in IRT, pointing out that “More replications are always better in terms of producing a more accurate and reliable estimate of the parameters of interest” (p. 39). However, more replications require more computational time which will often constrain the number of replications and there becomes a point in which diminishing returns are reached. [Bandalos \(2006\)](#) suggested that 500 replications were large enough for structural equation modeling simulations in the frequentist framework, arguing this number of replications would provide stable standard error estimates. In IRT, some researchers have used as few as 30 replications (e.g., [Wang et al., 2018](#)), but other have used 1000 or more (e.g., [Miyazaki et al., 2019](#)). There are few guidelines on the number of replications for Bayesian simulation studies, but a smaller number of replications tend to be used in comparison to frequentist simulations, as Bayesian methods tend to be more computationally time consuming. For a multidimensional item response application, [Leventhal \(2019\)](#) used 25 replications per condition where within-replication posterior distributions were built using 20,000 sampled values of the MCMC chain.

Determining the number of replications requires balancing feasibility of the study and generalizability of the results. Methods to increase the efficiency of the simulation can alleviate simulation resource cost, such as the time to perform the study, allowing more conditions. For example, if stability of results across replications is likely, the number of replications can be reduced, however if stability is unlikely, the number of replications should be increased limiting the number of feasible simulation conditions. Noise or variability can be reduced by randomly sampling from the same generated dataset across multiple conditions increasing the efficiency of the design. Finally, it is important to understand that different outcomes require a different number of replications. Parameter recovery studies focused on bias generally require fewer replications than parameter recovery studies investigating the standard error of parameter estimates. Prior studies or theoretical literature can be used to determine and justify the number of replications necessary for simulation conditions.

Step 8: Analyzing results based on the design and research question

The crossing of conditions, combined with a large number of replications, and the potential for multiple outcomes produces abundant data and output. As such, the logical, coherent summarization and presentation of results is critical to drawing conclusions from the simulation. The interpretable presentation of results typically takes on one or more of the following forms: (1) graphical displays, (2) tabular summarization by condition, and (3) statistical analysis. The statistical analysis appropriate when conditions have discrete levels, for example, is an ANOVA.

The outcome measures of the analysis, like bias or Type I error, are summarized across all replications within a single condition. For example, bias for parameter θ in condition levels i, j , and k is the average deviation between the true and generating value:

$$Bias_{i,j,k} = \sum_R \frac{(\hat{\theta}_{i,j,k,R} - \theta_{True,i,j,k})}{R} \quad (1)$$

where R is the number of replications, $\hat{\theta}_{i,j,k,R}$ is the estimated parameter value and $\theta_{True,i,j,k}$ is the generating parameter value, which is commonly fixed across all replications for studies of frequentist methods. The researchers could determine, using ANOVA, which conditions have a significant effect on the outcome. Analysis of variance techniques facilitate succinctly analyzing each factor and their interactions. Factors are commonly treated as fixed effects and may be considered a within- or between-subjects independent variable. All interactions among factors crossed in the simulation design are analyzed. Effect sizes, such as partial eta-squared (η_p^2), which is a measure of variance accounted for, are used to determine meaningful interactions rather than statistical significance due to the influence of the large replication numbers. Interactions and main effects determined to be of significant interest are interpreted further through follow-up analysis and graphical displays.

Another approach is for values of the outcome to be compared to some heuristic of an acceptable level of the outcome and illustrated graphically. For example, a Type I error rate is often set at 0.05. Any condition combinations for which the Type I error exceeds 0.05 could then be used to guide practitioners in their applied research. There is no single best analysis approach, as the analysis and display of results depend upon the research question and research audience.

Example

To illustrate the concepts, we provide a complete example of a simulation study in the IRT framework. Consider an example scenario in which a researcher is investigating the Type I error rates of a hypothesis test for model-data fit for a two-parameter logistic

model using a chi-square statistic (Step 1). They vary latent trait distribution (positive skew, negative skew, normal) using small, moderate, and large sample sizes (Step 2). They generate responses to a 20-item test (test length is a fixed factor) using item difficulties randomly selected from a standard normal distribution and item discrimination parameters from a log-normal distribution. To evaluate their research questions, sample size levels are crossed with trait distributions resulting in a 3×3 or 9 condition simulation design. For each replication, the outcome considered is the chi-square fit statistic (Step 3). The generating model simulates data under the null hypothesis of model data fit in each condition. Latent trait distributions are validated using a histogram of generated latent abilities and item difficulties are confirmed via visual inspection of generated item characteristics curves (Step 4). Using marginal maximum likelihood, the analysis model is fit to the data (Step 5) and the hypothesis test is conducted calculating the chi-square statistic. Because Type I error is the research focus, the analysis model and generating model are the same. The researcher specifically records whether the null hypothesis of model-data fit is rejected within each replication (Step 6). Steps 4, 5, and 6 are replicated 1000 times (Step 7). After replication, Type I error rate is calculated as the proportion of times the null hypothesis is rejected out of 1000 replications (Step 8). Using a tabular display, the effects on Type I error rates of sample size, latent trait distribution, and their interaction are investigated for the Chi-Square test statistic.

Conclusion

Simulation studies are a subset of psychometric studies used to answer questions where knowing true parameters or true model fit is necessary. An understanding of simulations is critical for both consumers and producers of psychometric research. The eight steps outlined provide a framework to most applications of simulation techniques whether it is for parameter recovery studies or comparison studies. As with empirical studies, simulation studies need to be evaluated on a case-by-case basis but have many common elements. Justification of each decision, validation of all syntax, and sufficient detail must be described in order to assure the reproducibility of the study. Simulation studies can quickly become far too complex to be feasible or interpretable, so these concerns must be balanced with the generalizability of the results.

References

- Ames, A.J., Leventhal, B.C., Ezike, N.C., 2020. Monte Carlo simulation in item response theory applications using SAS. *Measurement* 18, 55–74.
- Andersen, E.B., 1972. The numerical solution of a set of conditional estimation equations. *J. Roy. Stat. Soc. B* 34, 42–54.
- Bandalos, D.L., 2006. The use of Monte Carlo studies in structural equation modeling research. In: Hancock, G.R., Mueller, R.O. (Eds.), *Structural Equation Modeling: A Second Course*. Information Age, Greenwich, CT, pp. 385–426.
- Beaujean, A.A., 2018. Simulating data for clinical research: a tutorial. *J. Psychoeduc. Assess.* 36, 7–20.
- Bock, R.D., Aitkin, M., 1981. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika* 46, 443–459.
- Duane, S., Kennedy, A.D., Pendleton, B.J., Roweth, D., 1987. Hybrid Monte Carlo. *Phys. Lett. B* 195, 216–222.
- Feinberg, R.A., Rubright, J.D., 2016. Conducting simulation studies in psychometrics. *Educ. Meas.* 35, 36–49.
- Fischer, G.H., 1981. On the existence and uniqueness of maximum likelihood estimates in the Rasch model. *Psychometrika* 46, 59–77.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* 6, 721–741.
- Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), *Bayesian Statistics*, fourth ed. Oxford University Press, Oxford, p. 169–193.
- Haberman, S.J., 1977. Maximum likelihood estimates in exponential response models. *Ann. Stat.* 5, 815–841.
- Harwell, M., Stone, C.A., Hsu, T.C., Kirisci, L., 1996. Monte Carlo studies in item response theory. *Appl. Psychol. Meas.* 20, 101–125.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–109.
- Lee, S., 2015. Implementing a simulation study using multiple software packages for structural equation modeling. *SAGE Open* 5, 1–16.
- Leventhal, B.C., 2019. Extreme response style: a simulation study comparison of three multidimensional item response models. *Appl. Psychol. Meas.* 43, 322–335.
- Leventhal, B.C., Ames, A.J., 2020. Monte Carlo simulation studies in IRT [Digital ITEMS Module 13]. *Educ. Meas.* 39, 109–110.
- Luecht, R., Ackerman, T.A., 2018. A Technical note on IRT simulation studies: dealing with truth, estimates, observed data, and residuals. *Educ. Meas.* 37, 65–76.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092.
- McKinley, R.L., Mills, C.N., 1985. A comparison of several goodness-of-fit statistics. *Appl. Psychol. Meas.* 9, 49–57.
- Miyazaki, Y., Chungbaek, Y., Shropshire, K.O., Hedeker, D., 2019. Consequences of ignoring nested data structure on item parameters in Rasch/1P-IRT model. *Behaviormetrika* 46, 401–434.
- Muthén, L.K., Muthén, B.O., 1998–2017. *Mplus User's Guide*, sixth ed. Muthén & Muthén, Los Angeles, CA.
- Natesan, P., Nandakumar, R., Minka, T., Rubright, J.D., 2016. Bayesian prior choice in IRT estimation using MCMC and variational Bayes. *Front. Psychol.* 7, 1–11.
- SAS Institute Inc, 2015. *SAS/STAT® 14.1 User's Guide*. SAS Institute Inc, Cary, NC.
- Stan Development Team, 2017. *Stan: A C++ Library for Probability and Sampling*, Version 2.14.0. <http://mc-stan.org/>.
- Sung, L., Hayden, J., Greenberg, M.L., Koren, G., Feldman, B.R., Tomlinson, G.A., 2005. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J. Clin. Epidemiol.* 58, 261–268.
- van Doorn, J., van den Bergh, D., Boehm, U., Dablander, F., Derks, K., Draws, T., Evans, N., Gronau, Q., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Raj, A., Sarafoglou, A., Stefan, A., Voelkel, J., Wagenmakers, E.-J., 2019. The JASP guidelines for conducting and reporting a bayesian analysis. *Psychon. Bull. Rev.* 28 (3), 813–826.
- Wang, C., Su, S., Weiss, D.J., 2018. Robustness of parameter estimation to assumptions of normality in the multidimensional graded response model. *Multivariate Behav. Res.* 53, 403–418.

Educational surveys: conceptual overview

Lale Khorramdel^a, Matthias von Davier^a, Irwin Kirsch^b, and Kentaro Yamamoto^c, ^a Boston College, Chestnut Hill, MA, United States; ^b Educational Testing Service, Princeton, NJ, United States; and ^c Retired, former Educational Testing Service (Princeton, NJ, USA)

© 2023 Elsevier Ltd. All rights reserved.

Introduction	347
Purpose and policy relevance of large-scale assessments	348
National and international large-scale assessments	348
NAEP	348
TIMSS	348
PIRLS	348
CIVED	349
ICILS	349
PISA	349
PIAAC	349
PISA for schools	349
Regional large-scale assessments	350
PASEC	350
ERCE	350
SACMEQ	350
SEA-PLM	351
PILNA	351
Comparing regional and international large-scale assessments for building capabilities in low- and middle-income countries	351
PISA-D	351
IEA's Rosetta Stone	351
IEA's LaNA	352
Content, cognition and context information	352
Core domains	352
Example, PIRLS 2021 reading	352
Example, PISA 2018 reading	353
Additional and innovative domains	353
Context variables	353
Measurement characteristics	354
General design considerations	354
Population sampling	354
Domain sampling	354
Matrix sampling	354
Adaptive testing	355
Group adaptive designs	355
Multistage adaptive designs	355
Establishing comparable scales and plausible values	355
Trends and innovations	356
Reporting of results	356
Conclusion	357
References	357

Introduction

Large-scale assessments (LSAs) are comparative educational surveys that are widely used to inform policy and provide large datasets for educational researchers. They compare skills, knowledge, attitudes and behaviors across different populations and subpopulations of interest. The focus of these assessments is on group-level scores in contrast to large-scale testing programs that aim to report scores for individual test takers. These differences are reflected in the test and sampling design of LSAs as well as in the reporting of results. LSAs have grown and expanded over the past decades (since the late 1950s) to include a broader range of skill domains and populations, and to better meet the questions posed by policymakers and researchers worldwide (Kirsch et al., 2013). They can be

used for cross-country comparisons, school monitoring, the evaluation of growth and changes in learning progress over time (reporting of trends), and monitoring progress toward educational development goals.

Purpose and policy relevance of large-scale assessments

The overall purpose of LSA is to provide researchers and policymakers with valid data and meaningful group-level information about populations and subpopulations of interest with regard to skills, knowledge and behavior and how these relate to educational, economic and social outcomes. The focus is especially on skills that “contribute to the prosperity of nations and economies and to better lives for individuals in those nations” (Kirsch et al., 2013, p. 7). LSAs can be utilized for monitoring educational effectiveness and can help develop ways to improve the quality and equality of education worldwide.

With their rich databases, LSAs provide comparative data about countries’ student achievement in relation to key home, school, and classroom variables (Mullis and Martin, 2017). Therewith, they enable researchers to examine the relation of students’ proficiencies to their contexts for learning. They also allow for monitoring the effectiveness of a country’s educational system compared to other countries and own expectations, and for measuring progress over time (trends).

National and international large-scale assessments

LSAs can be roughly separated into national and international assessments that assess either student or adult populations of different age groups or school grades within or across countries. While student populations are assessed in classroom settings, adult populations are usually assessed in household surveys. There are also regional assessments which compare populations within a particular geographical region. In the following, examples of national and international large-scale assessments are provided.

NAEP

The National Assessment of Educational Progress (NAEP¹) conducted by the National Center for Education Statistics (NCES) was one of the first of its kind and assesses elementary and secondary grade students across the US since 1969. While a cross-country feasibility study had been conducted prior between 1959 and 1962 (see Kirsch et al., 2013), NAEP and the methods developed for it, especially at the Educational Testing Service (ETS) in the 1980s (Beaton and Barone, 2017; Kirsch et al., 2013), paved the way for other LSA that followed. NAEP assesses the core domains of reading, mathematics and science next to a number of other domains (described in the section “Content, cognition and context information” below); it is administered every 2 or 4 years depending on the subject domain, and transitioned to a digital-based administration (tablets) in 2017.

TIMSS

The Trends in International Mathematics and Science Study (TIMSS²) is conducted by the International Association for the Evaluation of Educational Achievement (IEA) and the TIMSS & PIRLS International Study Center at Boston College since 1995 and assesses students’ mathematics and science achievement. It was first implemented at grades 3 and 4, 7 and 8 and the final year of secondary school in 40 countries and since then is being implemented at grades 4 and 8 in over 60 countries. TIMSS assesses the domains mathematics and science (including problem solving related to both domains) and is administered every 4 years. It started transitioning to a digital-based assessment administration in 2019. In 1995, 2008 and 2015, the *TIMSS Advanced* study was conducted in addition to TIMSS for assessing advanced mathematics and physics achievement for students in their final year of secondary school.

PIRLS

The Progress in International Reading Literacy Study (PIRLS³) is conducted by the IEA and the TIMSS & PIRLS International Study Center since 2001; it assesses students reading achievement at grade 4 in over 60 countries. As indicated by its name, PIRLS assesses reading and is administered every 5 years. It began transitioning to a digital-based assessment administration in 2021 and implemented the additional *ePIRLS* assessment, which monitors how well students read online information.

¹<https://nces.ed.gov/nationsreportcard/>.

²<https://timssandpirls.bc.edu/timss-landing.html>.

³<https://timssandpirls.bc.edu/pirls-landing.html>.

CIVED

The Civic Education Study (CIVED⁴) was conducted by the IEA in collaboration with the Humboldt University of Berlin in two phases between 1996 and 2000. It represented the largest and most rigorous study of civic education implemented internationally up to that time. CIVED assessed 14-year-old students in 30 countries in the domains of democracy and citizenship, national identity, social cohesion and diversity, as well as the engagement in civil society. Some of the participating countries additionally conducted an optional survey of upper-secondary school students (ages 17–19).

ICILS

The International Computer and Information Literacy Study (ICILS⁵) has been conducted by the IEA since 2013 (in collaboration with the ACER in the first two cycles) and is administered in a 5-year cycle. It was designed to assess how well students in grades 8 and 9 (i.e., at the average age of 13.5 years or above) are prepared for study, work, and life in a digital world; it also measures international differences in students' computer and information literacy as well as the optional domain computational thinking. In addition to helping countries to monitor their own national targets regarding students' digital competences, ICILS provides researchers, educators, and policymakers with a data base for monitoring progress toward the UN's Sustainable Development Goals (SDGs). Furthermore, the resulting data are recognized as an official European Union (EU) target by the EU Council and Member States to support strategic priorities toward the European Education Area and beyond (2021–2030).

PISA

The Programme for the International Student Assessment (PISA⁶) is conducted by the Organization for Economic Co-operation and Development (OECD) in collaboration with the Australian Council for Educational Research (ACER) and ETS since 2000. PISA assesses 15-year-old students in over 80 countries in the core domains reading, mathematics and science, and offers an optional assessment of financial literacy since 2012. It is administered every 3 years with one of the core domains being the main domain in a particular assessment cycle. That is, new items are being developed for the main domain based on an updated framework while the other two domains are assessed with items from previous assessment cycles only (called trend items). In 2015, science was the main domain, in 2018 reading was the main domain, and in 2021 (administered in 2022 due to the covid-19 pandemic) mathematics was the main domain. Moreover, PISA offers the optional assessment of a new innovative domain in each cycle (as described in the section **“Content, cognition and context information”** below), transitioned to a fully digital-based assessment in 2015 after offering a partially digital-based administration in 2012, and started implementing multistage adaptive test designs in 2018.

PIAAC

The Programme for the International Assessment of Adult Competencies (PIAAC⁷) is a household survey conducted by the OECD in collaboration with ETS since 2012 and was preceded by the International Adult Literacy Survey (IALS) and the Adult Literacy and Life skills Survey (ALL),⁸ administered between 1994 and 1998 as well as between 2003 and 2007, respectively. PIAAC assesses 16–65-year-old examinees in some 40 countries in the domains of literacy, numeracy and problem solving in technology-rich environments. In its second cycle, PIAAC is assessing the domain adaptive problem solving. PIAAC follows approximately a 10-year cycle and was the first international large-scale assessment to offer a digital-based assessment administration for all participating countries as well as a multistage adaptive testing design (see the section **“Measurement characteristics”** below for more information on the design).

PISA for schools

The PISA-based Test for Schools (PBTS⁹) has been offered by the OECD since 2017. PBTS is based on the same framework as PISA and is assessing 15-year-old students' abilities to think critically and apply their knowledge creatively in novel contexts in the domains reading, mathematics and science. It is administered as a digital-based assessment on demand in interested schools and, to this date, available in 15 countries. PBTS aims to improving student learning opportunities and well-being by empowering teachers and school leaders through global connections and international benchmarking based on a common scale provided by PISA. By comparing their performance to international benchmarks, schools can set goals against the best school systems worldwide, better understand the challenges faced by low-performing students, and create peer-learning communities and networks with other schools and teachers.

⁴<https://www.iea.nl/studies/iea/cived>.

⁵<https://www.iea.nl/studies/iea/icils>.

⁶<https://www.oecd.org/pisa/>.

⁷<https://www.oecd.org/skills/piaac/>.

⁸<https://www.oecd.org/education/innovation-education/adultliteracy.htm>.

⁹<https://www.oecd.org/pisa/pisa-for-schools/>.

These different LSAs show some variation with regard to their aims and frameworks. NAEP, TIMSS and PIRLS are more driven by countries' school curricula (Mullis and Martin, 2015, 2017; NAEP assessment frameworks¹⁰) while PISA and PIAAC are rather interested in skills needed to meet real-life challenges at home, work and in the wider community (PISA 2022 Assessment and Analytic Framework¹¹; OECD, 2021). However, the designs and methods utilized in LSAs are very similar. All LSAs are based on matrix sampling, item response theory (IRT; Lord and Novick, 1968) and conditioning or population modeling (Mislevy et al., 1992; von Davier et al., 2006; von Davier et al., 2009), and are administered in cycles to provide trend measures for evaluating changes in the performance of populations over time.

Regional large-scale assessments

This section provides examples of regional large-scale assessments which provide comparisons across different countries belonging to specific regions that share the same language and a similar culture. Like LSAs in general, regional assessments aim to provide relevant information on the status and progress of learning, monitor, evaluate, and compare the general conditions of schooling and the quality of basic education, generate information that can be used by decision-makers to formulate plans for improving the quality of education, and provide a basis for policy and practice. Because these assessments are largely implemented for low- and middle-income countries, they are also an important tool for providing data that help measure progress toward the United Nations (UN) Sustainable Development Goal 4 of the 2030 Agenda (SDG 4, ¹² Target 4.1) which proposes to “ensure inclusive and equitable quality education and promote lifelong learning opportunities for all” especially with regard to target or indicator 4.1.1 which aims to ensure that by 2030 “all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.” Note that the regional LSAs presented below are all paper-based assessments.

PASEC

The Programme for the Analysis of Education Systems (PASEC¹³; Programme d'Analyse des Systemes Educatifs) has been conducted by the Conference of Ministers of Education of French-Speaking Countries (CONFEMEN) since 1993. It started as a diagnostic tool to assess the progress of students from the beginning to the end of the same school year in grades 2 and 5 in countries of Sub-Saharan Africa, the Indian Ocean, the Middle East and South-East Asia, but has evolved since 2012. PASEC's new focus is on the comparison of national assessments in francophone Sub-Saharan countries for enabling a better understanding of the effectiveness and equity of education systems, in line with other international programs such as PISA, PIRLS or TIMSS. Since then, PASEC has been administered in 13 countries in Francophone West Africa to assess students' achievement in mathematics and reading French (including listening comprehension) at grades 2 and 6 (i.e., at the beginning and end of primary school). Results are available for the 2014 and 2019 assessment cycles and more cycles will follow.

ERCE

The Regional Comparative and Explanatory Study (ERCE¹⁴) is conducted by the Latin American Laboratory for Evaluation of the Quality of Education (LLECE) of the United Nations Educational, Scientific and Cultural Organization (UNESCO) since 1997. It assesses students at grades 3 and 6 in Latin America and the Caribbean in the domains of reading, writing, mathematics, and science, and is an important tool for monitoring the SDG 4. ERCE was last implemented in 18 countries in 2019.

SACMEQ

The Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ¹⁵) has conducted cross-national assessments in 15 anglophone countries in the Southern and Eastern Africa region since 1995. To date, three cross-national educational research projects have been implemented to assess achievements in reading English and mathematics at grade 6: SACMEQ I (1995–1999), SACMEQ II (1998–2004) and SACMEQ III (2005–2010). While SACMEQ I assessed only students' achievement (reading), SACMEQ II and III assess students' achievement as well as the performance level of teachers (reading and mathematics). Moreover, SACMEQ II and III included linking items from TIMSS and PIRLS with the goal of enabling comparisons between these regional and international assessments (Cresswell et al., 2015).

¹⁰<https://nces.ed.gov/nationsreportcard/assessments/frameworks.aspx>.

¹¹<https://www.oecd.org/pisa/publications/pisa-2021-assessment-and-analytical-framework.htm>.

¹²<https://sdgs.un.org/goals/goal4>.

¹³<http://www.pasec.confemen.org>.

¹⁴<https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2019/05/GAML6-Session5-LLECE.pdf>.

¹⁵<http://www.sacmeq.org>.

SEA-PLM

The Southeast Asia Primary Learning Metrics (SEA-PLM¹⁶) was initiated by the Southeast Asian Ministers of Education Association (SEAMEO) and UNICEF in 2012 with the aim to realize the SDG 4 and to further improve the quality of primary education in Southeast Asia. SEA-PLM assess and monitors students' achievement in mathematics, reading, writing, and global citizenship at primary grade 5. It is based on the SEAMEO political framework and is embedded within national education systems in order to be considerate of the regional context and values. The assessment was first implemented in 6 countries in 2019, in collaboration with ACER.

PILNA

The Pacific Island Literacy and Numeracy Assessment (PILNA¹⁷) is administered by the Educational Quality and Assessment Division (EQAP) of the Pacific Community in collaboration with the Pacific Ministers for Education and ACER, and has received financial support from the Australian Department of foreign Affairs and Trade as well as from the New Zealand Aid Programme. It has been implemented in 3-year cycles since 2012, and was administered to 15 Pacific countries in 2018. PILNA was developed from the component skills in literacy and numeracy common to all Pacific countries, and assess students' basic numeracy and literacy skills after completing four and 6 years of primary education.

Comparing regional and international large-scale assessments for building capabilities in low- and middle-income countries

This section describes studies that aim to build capabilities in low- and middle-income countries by either providing assessments as a preliminary step toward the participation in international large-scale assessments (PISA for Development and IEA's LaNA), by establishing a comparison between regional and international large-scale assessments for countries that are not yet ready to participate (IEA's Rosetta Stone), or both (IEA's LaNA). The main goal is to provide countries with data and international comparisons that help improve their education systems (students' learning, teachers' education, school systems' efficacy) so that students can meet minimum proficiency standards in reading and mathematics, as defined in the UN's SDG 4. For more than half of the world's children this has not been the case, even before the COVID pandemic, and with school closures during the pandemic impacting about 90% of students worldwide, it is more important than ever to concentrate on improving global education. Moreover, participation in such studies provides educational officials, researchers and analysts with training in implementing standardized LSA procedures, assessment administration, and analyzing the data and reporting the outcomes in preparation for participation in international large-scale assessments such as TIMSS, PIRLS or PISA in the future.

PISA-D

The PISA for Development programme (PISA-D¹⁸) was designed as a 6-year project and conducted by the OECD between 2014 and 2019 through a partnership involving eight low-income countries, concerned development partners (DAC members plus the World Bank, UNESCO, UNICEF and other UN bodies and regional organizations), ETS and The Learning Bar. It aimed to encourage and facilitate the participation of interested low- and middle-income countries in PISA by providing instruments that are more relevant for the contexts found in such countries but which produce scores that are on the same scales as the main PISA assessment. PISA-D assessed students in the domains of reading, mathematics, and science, and was administered as a paper-based assessment to in school student populations. It was also administered to out-of-school children as a digital-based household survey (tablets) following the PIAAC model. The goal was to provide policy makers with insights on how to improve their educational systems, help build the capacity of participating countries to conduct LSA and make use of resulting data, and contribute to monitoring progress toward nationally-set targets for improvement as well as toward the SDG 4. Since 2021, PISA-D was integrated into the main PISA test, making the assessment more accessible and relevant to middle- and low-income countries.

IEA's Rosetta Stone

Named after the famous archeological discovery that enabled translation between different written languages, the Rosetta Stone¹⁹ study, conducted by IEA and the TIMSS & PIRLS International Study Center, enables relating children's achievement on countries' regional and national assessments to achievement at the TIMSS and PIRLS international benchmarks. Therewith, IEA's Rosetta Stone study allows countries to compare their students' achievement internationally and provides data that help countries measure global progress toward the SDG 4²⁰. Administering the paper-based Rosetta Stone linking booklets containing TIMSS and PIRLS items

¹⁶<https://www.seaplm.org>.

¹⁷<https://eqap.spc.int/PILNA>.

¹⁸<https://www.oecd.org/pisa/pisa-for-development/>.

¹⁹<https://www.iea.nl/studies/additionalstudies/rosetta>.

²⁰Please refer to the following document: <http://unesdoc.unesco.org/images/0026/002606/260607E.pdf>.

together with the national or regional assessments provides data for the TIMSS & PIRLS International Study Center to establish concordance tables that relate countries' achievement on their national or regional assessment to the TIMSS and PIRLS International Benchmarks. Countries who participate in the Rosetta Stone study are provided with information about the proportions of primary school students reaching the TIMSS and PIRLS achievement benchmarks. In 2022, the first concordance tables for reading and mathematics were constructed for the regional assessments PASEC and ERCE.

IEA's LaNA

The Literacy and Numeracy Assessment (LaNA²¹) was developed by the TIMSS & PIRLS International Study Center upon request by the IEA, with the final instrument being established in 2021 after a multi-country pilot, including student samples from Africa, Asia, the Caribbean, and Europe. LaNA is aimed at measuring basic literacy and numeracy skills at the end of primary school, and targets student populations for which TIMSS and PIRLS are too challenging. LaNA is a shorter, less demanding paper-based assessment and can be used to identify strengths and weaknesses in education systems and as a precursor to implementing TIMSS or PIRLS. LaNA also aims to establish a linkage between its literacy and numeracy achievement scales and the TIMSS and PIRLS reading and mathematics scales. To accomplish this, LaNA uses additional paper-based linking booklets containing material sourced from applicable (easier) parts of the original TIMSS and PIRLS assessments. Countries administering LaNA will be able to use their students' performance on LaNA to estimate how they would perform against the lower TIMSS and PIRLS achievement benchmarks which can serve as measures for SDG 4. In contrast to IEA's Rosetta Stone, where the constructs measured by regional assessments differ from the constructs measured in TIMSS and PIRLS (yet, are similar enough to establish a concordance), the constructs measured with LaNA are actually based on the comprehensive TIMSS and PIRLS frameworks.

Content, cognition and context information

The constructs measured by LSAs are based on comprehensive conceptual assessment frameworks which provide definitions and examples for each measure, and the background for why these are important for everyday life, work, school or economy. These frameworks are established by content experts from different fields such as education, psychology, sociology and measurement. In addition to what an assessment intends to measure, conceptual assessment frameworks provide information on how the assessment should be structured. According to [Mislevy and Riconscente \(2005\)](#) the conceptual assessment framework includes the student or proficiency model, the evidence model, the task model and the assembly model. The student model describes what is being measured and defines one or more variables that are related to the construct(s) of interest (knowledge, skills, abilities). The evidence model defines how the responses of examinees to tasks should be scored (evaluation component) and how these scores are related to the measurement (measurement model). The task model describes how items need to be developed and structured to obtain the behavior or responses needed for the evidence model (e.g., remembering versus recognizing or problem solving), including item forms (e.g., multiple-choice or constructed response items), and the materials used (including directives, manipulatives, and features). The assembly model relates the student, evidence and task model to one another, and represents the accuracy, breadth and diversity of the domain being assessed (e.g., selecting a given items from particular subdomains, and whether to present them in a linear or adaptive test administration).

The purpose of assessment frameworks is not only to guide the development of the assessment, but also to guide the reporting of results and provide a common language and accountability for stakeholders. To ensure the relevance of the measured constructs in an ever-changing world, assessment frameworks are constantly being updated for each assessment cycle. To ensure construct validity, the item development based on these frameworks passes through multiple development, review and revision stages overseen by different (often international) experts, including translation and verification processes, and actively involves participating countries to ensure that the content and difficulty levels are suitable for most students in each country.

Core domains

The core achievement domains in most LSA are reading and literacy, mathematics and numeracy, and science. Each domain consists of a content dimension, specifying the subject matter to be assessed, and a cognitive dimension, specifying the thinking processes to be assessed. Hence, each achievement domain consists of different subdomains.

Example, PIRLS 2021 reading

The domain reading in PIRLS is defined as "... the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment." ([Mullis and Martin, 2019](#)). This definition comprises online reading skills as well, and focusses on the following subdomains:

²¹<https://www.iea.nl/studies/iea/LaNA>.

- Purposes for reading
 - Literacy experience
 - Acquire and use information
- Processes of comprehension
 - Focus on and retrieve explicitly stated information
 - Make straightforward inferences
 - Interpret and integrate ideas and information
 - Evaluate and critique content and textual elements

Example, PISA 2018 reading

The domain reading in PISA 2018, where reading was the main domain, is defined as “understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society” (OECD, 2019) and comprises the following subscales:

- Text formats
 - Continuous texts (newspaper reports, essays, novels, short stories, reviews and letters)
 - Non-continuous texts (lists, tables, graphs, diagrams, advertisements, schedules, catalogs, indexes and forms)
 - Mixed texts (a combination of continuous and non-continuous texts)
- Text sources (units of texts)
 - Single source texts
 - Multiple source texts
- Cognitive processes
 - Locate information
 - Understand
 - Evaluate and reflect

The definitions and subdomains of reading in PIRLS and PISA show some similarities and overlap, but also a number of differences. Moreover, the goals of these LSAs have subtle but important distinctions. Similar to NAEP, which focuses on the US school system, as opposed to cross-country comparisons, PIRLS aims to measure school-based curricular attainment while PISA is a literacy-based assessment that focuses on the “yield” of education systems and the application of skills in real-world contexts (Stephens and Coleman, 2007). Hence, constructs across different LSAs cannot be compared directly.

International reporting and comparisons usually focus on results in the overall domain (e.g., reading), but information at the subdomain level (e.g., locate, understand, and evaluate information) is available in the database as well for further analysis at the national level (note that subdomains are generally highly correlated).

Additional and innovative domains

Next to these core achievement domains, different LSAs typically assess additional domains of interest. NAEP, for example, provides measures for arts, civics, economics, geography, US history, technology and engineering, and writing. The transition of LSAs to a computer-based administration also offers the assessment of new and innovative domains through interactive item types and new materials and environments. TIMSS, PISA and PIAAC provide different types of problem-solving measures, for example. TIMSS assesses problem solving and inquiry as part of mathematics and science; PIAAC assessed problem solving in technology rich environments in cycle 1 and will assess adaptive problem solving in cycle 2. PISA assessed collaborative problem solving in 2015. PIRLS, PISA and PIAAC also measure digital reading, and PISA offers the measurement of financial literacy (which is not specific to computer-based administration) as well as the option of a new innovative domain in each cycle, which is only assessed once, such as collaborative problem solving (2015), global competence (2018), creative thinking (2021); (due to the COVID pandemic administered in 2022), or learning in the digital world (2025). ICILS offers the option of assessing computational thinking (including programming as well as structuring and manipulating data), in addition to its core assessment of computer and information literacy skills.

Context variables

In addition to collecting achievement data, LSAs collect a large set of policy relevant background or context variables through student, parent, and school questionnaires as well, which are based on self-ratings. These include student attributes, home and school contexts for teaching and learning, students’ home background, school and classroom resources, teacher and teaching quality, and equity. In addition to demographic information, background questionnaires provide valuable information on the learning climate within a school, students socioeconomic background and motivation for learning, and they can help to measure a wider range of 21st century skills. Similar to achievement domains, the development and structure of background questionnaires is based on comprehensive assessment frameworks, and their content is updated with every assessment cycle to reflect new topics of

emerging relevance in addition to topics of ongoing importance across cycles. Context variables can be used to examine and understand group differences in cognitive proficiencies, can be related to other external variables like socioeconomic status, school, teaching or economic systems, and have the potential to be leveraged to maximize educational outcomes (Lee and Chen, 2019; Lee and Stankov, 2018).

Measurement characteristics

Abilities, skills, or proficiencies measured by LSAs are not directly observable but constitute hidden or latent traits. They can be modeled based on the evidence collected through examinees' responses to assessment items (i.e., observable behavior) using measurement models such as IRT and conditioning or population models. These models are used to construct unidimensional (if possible) and comparable scales for each construct of interest by estimating common item parameters (cf. Khorramdel et al., 2020a,b,c; Yamamoto et al., 2013a, updated 2019). Establishing comparable achievement and contextual scales across populations within countries, across countries, and across assessment cycles is an important goal of international large-scale assessments and the basis for group comparisons and the reporting of trend (cf. Khorramdel et al., 2020a,b). Consequently, measurement invariance of item parameters and scales has to be achieved to ensure that results can be interpreted similarly across different groups of interest. Conceptually, measurement invariance is concerned with minimizing construct irrelevant sources of variance in item responses so that the relationship between items and the latent trait measured by those items (represented by test scores) is not affected by group membership (Mellenbergh, 1989). Measurement invariance in LSA is achieved through high standards in instrument development, translation and adaptation, specific sampling and test design considerations, standardized data collection, the use of IRT and conditioning models in data analysis, and rigorous quality evaluation on each step of the way. Each LSA cycle is preceded by a field test for evaluating the quality of the instruments, procedures and building capabilities (i.e., preparing countries for the main assessment).

General design considerations

The design of a LSA has to fulfill multiple goals to establish reliable, accurate, comparable, meaningful, and policy relevant results. It has to ensure comprehensive construct coverage for all groups of interest, comprehensive coverage of the target population(s), and comprehensive reporting coverage at the population and subpopulation level. The design has to account for a wide range of student achievement, allow for the measurement of trends and, at the same time, incorporate innovations (regarding constructs and test development, test administration, methods and analysis) (cf. Kirsch et al., 2020).

LSA designs are based on two important sampling considerations, taking advantage of the fact that group and subgroup comparisons are of interest as opposed to comparing individual test-takers: population sampling and domain sampling (Beaton and Barone, 2017).

Population sampling

Population sampling refers to the selection of samples with the aim to provide reliable and accurate measures at the population and subpopulation level. The sample has to be large enough to provide estimates with sufficiently small standard errors, yet small enough to be cost effective. Hence, the goal is to select the smallest sample size that will meet the defined measurement standards (Beaton and Barone, 2017). Moreover, the drawn samples need to be equivalent and representative, and have a similar distribution of characteristics like in the defined target population. To achieve this, sampling designs make use of methods like randomization, stratification and multistage-sampling (e.g., LaRoche et al., 2020). The way in which results have to be reported (e.g., for which subpopulations results are required) is another important variable that influences the sample size and the sampling design.

Domain sampling

Domain sampling determines the breadth of the measurement within a subject area and refers to the achievement of a valid representation of frameworks and constructs (domains and subdomains) and reliable construct measures. For this purpose, a large number of items is required for each assessed domain, more items than can be possibly taken by one test-taker. Moreover, assessing multiple domains will increase the number of items further. Therefore, LSAs make use of matrix sampling and balanced incomplete block (BIB) designs, which includes multiple test forms with a rotated block design (Messick et al., 1983; von Davier et al., 2006; von Davier and Sinharay, 2013).

Matrix sampling

Matrix sampling and BIB designs allow the administration of a large number of items to achieve broad construct coverage and ensure that the complete item pool is administered to the population of interest while keeping the individual student burden (i.e., time spend on the assessment) as low as possible. More precisely, each student usually responds to one test booklet that comprises only a subset of the available items. These test booklets are linked through common item sets (blocks) ensuring that a common scale can be constructed even though students respond to different sets of items. Moreover, the position of items across test booklets is varied (rotated) to balance out possible position or order effects. Because matrix sampling allows for the incorporation of a large item pool in the design, trend items from prior assessment cycles can be administered together with new items to account for construct updates and innovations. Based on this linking across different cycles, matrix sampling allows for the reporting

of trends and changes in achievement over time. It should be noted that matrix sampling does not provide reliable achievement estimates at the individual student level, but allows for reliable estimates at the group and country level. Matrix sampling reduces the time and costs for participating countries, immensely, compared to census testing where all individuals of a population of interest are sampled.

Adaptive testing

One important challenge faced by international LSAs, in particular, is the variability of proficiencies across and even within countries, as examinees with a broad range of proficiency levels, school curricula and sometimes age (e.g., in PIAAC) are tested in multiple languages often associated with diverse cultural backgrounds. For example, in lower performing countries students might struggle to answer items while for students in higher performing countries these items might not be challenging enough. With more and more countries participating in international LSAs, this is an increasing concern. For example, 32 countries participated in PISA in 2000 while 90 countries participated in 2021/2022. Hence, LSA test designs benefit from some form of adaptiveness to better match item difficulty with examinee ability (cf. [Martin et al., 2019](#); [Kirsch et al., 2020](#); [Yamamoto et al., 2018b](#)). Adaptive test designs can increase test efficiency and the accuracy of LSAs by providing better coverage for a diverse range of ability distributions in addition to increasing student motivation and decreasing item-level nonresponse (omitted and not reached items) ([Yamamoto et al., 2018a,b](#)).

Group adaptive designs

This type of design applies adaptive testing at the country level without the need for individual routing of students. Item difficulty and average country proficiency are estimated based on prior assessment cycles and/or the field test. Based on this information, items can be labeled as difficult, medium and easy, for example, and combined into booklets or forms with different difficulty levels. PISA used such a design in its 2015 cycle and similar ones in prior cycles. Countries could choose between standard forms or easier forms, where two of the standard item sets were replaced by easier item sets in approximately one-third of the forms ([OECD, 2017](#)). Until 2015, PISA also offered a special 1 h assessment instrument (Une Heure or UH) for students with special educational needs containing about half as many items as the other PISA instruments. PIRLS 2021 and TIMSS 2023 use a more advanced design where all booklets are differentiated into either less difficult or more difficult booklets ([Martin et al., 2019](#); [Yin and Foy, 2021](#)). All booklets are administered in all countries, but the distribution of booklets depends on a countries' average proficiency.

Multistage adaptive designs

A computer-based test administration of LSAs allows for the implementation of more advanced adaptive test designs, such as multistage testing (MST) with a higher level of adaptiveness. That is, the administration of item sets with different difficulty levels depends on the student's proficiencies in prior item sets and follows a routing system. PIAAC was the first international LSA implementation of an MST design, in 2012. The PIAAC MST design in the first cycle was based on information from both the background questionnaire and the cognitive assessment, and had two levels of adaptation ([Kirsch et al., 2020](#); [Yamamoto et al., 2018a](#)): (1) based on their computer skills and experience, examinees were routed to either a paper- or a computer-based assessment and (2) within the computer-based assessment, examinees' proficiency levels were determined based on responses to prior cognitive items as well as information about their educational level and native language, and then used to assign examinees to the different adaptive stages. A probability-based multistage adaptive algorithm was used to control the item exposure rate for enabling broad construct coverage and for minimizing differential item functioning (DIF). The second cycle of PIAAC will be digital-based only (using tablets) and will use an adjusted MST design. PISA was the second international LSA to implement an MST design for its reading domain in 2018. The PISA MST design can be viewed as a three-stage (core, stage 1, and stage 2), two-level (difficult and easy item blocks²²) design ([OECD, 2020](#); [Yamamoto et al., 2019](#)) where students are routed to either more difficult or easier item blocks, first in stage 1 and then in stage 2, based on their performance in the previous core stage. The assignment in stage 2 is based on the combined performance of students in the core stage and stage 1. The PISA design also accounts for possible item position effects by reversing stage 1 and 2. Both, PIAAC and PISA MST designs use an additional probability layer matrix to control for item exposure to ensure a more accurate item parameter estimation and sufficient construct coverage at the population level.

Establishing comparable scales and plausible values

Given the complexities of the data collection and the need to describe student achievement on a scale that is comparable across samples and that represents the entirety of the assessment frameworks, LSAs utilize IRT models ([Lord and Novick, 1968](#)) to provide accurate measures of student proficiency distributions. IRT models have important advantages compared to classical methods in that they directly support the creation of comparable scales across multiple forms of a test, regardless of which student took which items. IRT has become one of the most important tools of educational measurement as it provides a flexible framework for estimating proficiency scores that allows establishing scales with a common meaning over different sets of items measuring the

²²In the PISA MST design, each item block comprises multiple item sets and each item set comprises multiple items.

same construct, different populations, and different time periods., which is especially useful in international large-scale assessments (von Davier et al., 2019b).

Based on the estimates from the IRT model and the available context variables from background questionnaires (usually the student questionnaire), a conditioning or population model is estimated. This approach can be viewed as an imputation model for the unobserved proficiency distribution that aims at obtaining unbiased group-level proficiency distributions by utilizing information about the extent to which background or context variables are related to the proficiency variable (von Davier et al., 2006, 2009; von Davier and Sinharay, 2013). The goal of population modeling is to produce posterior distributions of proficiencies from which plausible values (PVs) can be drawn. PVs are the basis for all reporting of proficiency data in LSAs, allowing reliable group level comparisons and the reporting of trend. Since PVs are group-level estimates, they cannot be used or interpreted like individual test scores.

The use of IRT and population models in the context of LSAs such as PIAAC, PISA, and TIMSS, and the proper use and interpretation of PVs, are illustrated by Khorramdel et al. (2020c), von Davier (2020), Yamamoto et al. (2013a,b) or in the PISA 2018 technical report (OECD, 2020). See also Jewsbury (2023), for more information on this subject.

Trends and innovations

Providing trends over time in an everchanging world is one of the biggest challenges of LSAs. In the same way that the world and skills around us change—what and how we learn (e.g., on computers, tablets, or smartphones), and the skills we need to participate successfully in school, in the workplace and in our everyday lives (e.g., technology related skills or digital reading)—LSAs strive to evolve and improve with each assessment cycle. However, the reporting of trends is based on comparable scales over time; significant changes from one assessment cycle to the next can potentially break the trend. Therefore, changes in content or methods have to be made gradually to ensure that trends can be measured while, at the same time, moving into the future.

One of the most important recent changes in LSAs was the transition from a paper-based to a digital-based administration (DBA). PIAAC has used a DBA since 2012, PISA moved to a DBA in 2015 for the majority of countries, TIMSS did so in 2019, followed by PIRLS in 2021, and NAEP in 2022. DBA enabled the implementation of a number of innovations that enhance LSA in different ways. DBA can make the assessment potentially more engaging and provide enhanced accommodations for students with special educational needs. Furthermore, it introduces higher standardization for test administration and test security, the possibility to measure existing domains in new ways and new domains through innovative and interactive tasks (e.g., experiments and simulations), new item types and tools (e.g., highlighting text, drop down, calculator, drawing), and the collection log and process data (such as response time, types of actions and action sequences, mouse movements, different nonresponse types).

These new data types can be used to evaluate the quality and validity of LSA data (Yamamoto and Lennon, 2018), and to advance methods for analyzing LSA data and, therewith, the accuracy and fairness of test scores (von Davier et al., 2019a). Process data can also potentially deepen our understanding of what students know and how they approach problem solving tasks (Mullis et al., 2021), as well as how they engage with cognitive and noncognitive assessment content, by examining behaviors such as guessing behavior and test-taking effort (Pokropek, 2016; Pokropek et al., 2022; Ulitzsch et al., 2020a, 2021), revisit behavior (Bezirhan et al., 2021), nonresponse behavior (Pohl et al., 2019; Ulitzsch et al., 2020b) and navigation patterns (He and von Davier, 2015; He et al., 2021).

Another innovation introduced by DBA is the use of automated scoring of text and image responses to reduce the workload of human raters and related costs, and to improve scoring reliability (Sukkarieh et al., 2012; von Davier et al., 2022; Yamamoto et al., 2017; Zehner et al., 2016). Furthermore, DBA allows for the implementation of MST designs for more accurate measurement at lower and higher proficiency levels, and to accommodate the participation of countries with a larger variation in average student proficiency (Yamamoto et al., 2018a,b, 2019).

The expansion of LSA to include a wide range of participating countries and populations within those countries, as well as to assess constructs beyond typical academic domains of reading, mathematics and science reflect the “growing recognition of the need for lifelong learning as a tool to succeed in rapidly changing economies” (Kirsch et al., 2013, p. 14). However, every change needs to be supported by extensive research and scientific studies before they are operationally implemented in LSA, especially with regard to possible impacts on the measurement of trends. The transition to DBA, for example, required the study of potential mode effects, and the IRT scaling methods needed to be extended to treat such effects in cases where they occur (Fishbein et al., 2018; von Davier et al., 2019a). Similarly, the implementation of MST in PISA was preceded by a study on item position effects (Yamamoto et al., 2018b) and simulation studies that compared the efficacy of different MST designs (Yamamoto et al., 2019).

Reporting of results

Results of LSAs are reported at the international and at the national level. At the international level, the international databases, technical documentation (including information on instrument development, sampling and test design, as well as methods and procedures), and policy related reports (including cross-country comparisons) are available on the websites of the organizations that conduct these LSAs (e.g., NCES, IEA, TIMSS & PIRLS International Study Center, OECD). At the national level, each country

produces detailed reports based on their own country-level results, including the comparison of subgroups and samples within countries and descriptions of educational systems.

To provide meaningful interpretation, results are usually reported with regard to different benchmarks or proficiency levels that describe what students or examinees at each level typically know and can do. TIMSS and PIRLS, for example, provide regular reports of the percentages of students reaching advanced, high, intermediate, and low international benchmarks of achievement in reading and mathematics, which can also be used as minimal proficiency standards according to SDG 4.²³ PIAAC and PISA report results at different proficiency levels, usually 5 to 6 levels respectively. Each benchmark or proficiency level is defined by certain score ranges and cut-points. Examinees whose proficiency estimate places them at a certain point on a proficiency scale would be more likely to successfully complete items at or below that point.

Full transparency of the data and results are instrumental in the impact of LSAs on educational policies and their use in educational research. Following the publication of these reports and databases, researchers engage in secondary analysis that can provide additional information and insight, for example, by explaining reasons for group differences or education inequality. It is important to stress again, that LSA results are not meant to be interpreted or used at the individual student or examinee level, nor as school rankings; they are designed to provide group-level results only and aim to support improvement efforts for whole educational systems or defined subgroups such as school types or certain student subgroups (e.g., gender) given that the sample size allows for it based on the sampling design. von Davier et al., 2019b

Conclusion

There is an increasing need to evaluate the effectiveness of school systems and to improve education given that we live in a more and more technology driven world with growing complexity, global challenges and educational inequity. This is not only true for young students but also for adult populations, especially given the recent focus on lifelong learning and the need to upskill and reskill young and prime working age adults. This educational challenge is reflected in the growing importance of LSAs at the national, regional and international level. To balance the need for both the reporting of trends (look to the past) and the implementation of innovations (look to the future), LSA are constantly striving to advance. They do so by (1) continuously extending the comprehensive frameworks to measure cognitive proficiencies and noncognitive variables and to include 21st century skills that are relevant for success in school, work and everyday life, and (2) by advancing the methods and procedures for establishing higher measurement accuracy and validity, and more inclusiveness for a larger variety of countries and economies. LSAs, now more than ever, can influence the development of skills that are needed to successfully navigate through our fast-changing world and adapt to it. To provide relevant information to policy makers means to identify mechanisms of change and the related personal characteristics and skills, to develop appropriate measures of these skills and to deliver learning opportunities to those who need them.

References

- Beaton, A.E., Barone, J.L., 2017. Large-scale group-score assessment. In: Bennett, R.E., von Davier, M. (Eds.), *Advancing Human Assessment, Methodology of Educational Measurement and Assessment*. https://doi.org/10.1007/978-3-319-58689-2_8.
- Bezirhan, U., von Davier, M., Grabovsky, I., 2021. Modeling item revisit behavior: the hierarchical speed–accuracy–revisits model. *Educ. Psychol. Meas.* 81 (2), 363–387. <https://doi.org/10.1177/0013164420950556>.
- Cresswell, J., Schwantner, U., Waters, C., 2015. A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data, PISA, the World Bank. OECD Publishing, Washington, D.C. <https://doi.org/10.1787/9789264248373-en>. Paris.
- Fishbein, B., Martin, M.O., Mullis, I.V.S., Foy, P., 2018. The TIMSS 2019 Item Equivalence Study: examining mode effects for computer-based assessment and implications for measuring trends. *Large-Scale Assess. Educ.* 6, 11. <https://largescaleassessmentsineducation.springeropen.com/articles/10.1186/s40536-018-0064-z>.
- He, Q., von Davier, M., 2015. Identifying feature sequences from process data in problem-solving items with n-grams. In: van der Ark, A., Bolt, D., Chow, S., Douglas, J., Wang, W. (Eds.), *Quantitative Psychology Research: Proceedings of the 79th Annual Meeting of the Psychometric Society*. Springer, New York, NY, pp. 173–190.
- He, Q., Borgonovi, F., Paccagnella, M., 2021. Leveraging process data to assess adults' problem-solving skills: using sequence mining to identify behavioral patterns across digital tasks. *Comput. Educ.* 166 (2021), 104170. <https://doi.org/10.1016/j.compedu.2021.104170>. ISSN 0360-1315.
- Jewsbury, P., 2023. Educational surveys: methodological foundations. In: Tierney, R.J., Rizvi, F., Ercikan, K. (Eds.), *International Encyclopedia of Education*, 4th edition. Elsevier, In press.
- Khorramdel, L., Pokropek, A., van Rijn, P., 2020a. Editorial—special topic, part I: establishing comparability and measurement invariance in large-scale assessments, part I. *Psychol. Test Assess. Model.* 62, 3–10. https://www.psychologie-aktuell.com/fileadmin/Redaktion/Journale/ptam-2020-1/01_Khorramdel.pdf.
- Khorramdel, L., Pokropek, A., van Rijn, P., 2020b. Editorial—special topic part II: establishing comparability and measurement invariance in large-scale assessments, part II—old questions, new challenges and possible solutions. *Psychol. Test Assess. Model.* 62, 139–145. https://www.psychologie-aktuell.com/fileadmin/Redaktion/Journale/ptam-2020-2/00_Editorial_Khorramdel.pdf.
- Khorramdel, L., von Davier, M., Gonzalez, E., Yamamoto, K., 2020c. Plausible values: principles of item response theory and multiple imputations. In: Maehler, D.B., Rammstedt, B. (Eds.), *Large-Scale Cognitive Assessment: Analysing PIAAC Data*. Springer, Cham, New York, pp. 27–47. https://doi.org/10.1007/978-3-030-47515-4_3.
- Kirsch, I., Lennon, M., von Davier, M., Gonzalez, E., Yamamoto, K., 2013. On the growing importance of international large-scale assessments. In: von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K. (Eds.), *The Role of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Springer, Dordrecht, pp. 5–22. https://doi.org/10.1007/978-94-007-4629-9_1.
- Kirsch, I., Yamamoto, K., Khorramdel, L., 2020. Design and key features of the PIAAC survey of adults. In: Maehler, D.B., Rammstedt, B. (Eds.), *Large-Scale Cognitive Assessment: Analysing PIAAC Data*. Springer, Cham, New York, pp. 7–26. https://doi.org/10.1007/978-3-030-47515-4_2.

²³Please refer to the following document: <http://unesdoc.unesco.org/images/0026/002606/260607E.pdf>.

- LaRoche, S., Joncas, M., Foy, P., 2020. Sample design in TIMSS 2019. In: Martin, M.O., von Davier, M., Mullis, I.V.S. (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*, pp. 3.1–3.33. <https://timssandpirls.bc.edu/timss2019/methods/chapter-3.html>.
- Lee, J., Chen, M., 2019. Cross-country predictive validities of non-cognitive variables for mathematics achievement: evidence based on TIMSS 2015. *Eurasia J. Math. Sci. Technol. Educ.* 15 (8), em1725. <https://doi.org/10.29333/ejmste/106230>.
- Lee, J., Stankov, L., 2018. Non-cognitive predictors of academic achievement: evidence from TIMSS and PISA. *Learn. Indiv. Differ.* 65, 50–64. <https://doi.org/10.1016/j.lindif.2018.05.009>.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA. <https://psycnet.apa.org/record/1968-35040-000>.
- Martin, M.O., von Davier, M., Foy, P., Mullis, I.V.S., 2019. PIRLS 2021 assessment design. In: Mullis, I.V.S., Martin, M.O. (Eds.), *PIRLS 2021 Assessment Frameworks*. <https://timssandpirls.bc.edu/pirls2021/frameworks/>.
- Mislevy, R.J., Riconscente, M.M., 2005. *Evidence-Centered Assessment Design: Layers, Structures, and Terminology*. PADI Technical Report. http://padi.sri.com/downloads/TR9_ECD.pdf.
- Mislevy, R.J., Beaton, A.E., Kaplan, B., Sheehan, K.M., 1992. Estimating population characteristics from sparse matrix samples of item responses. *J. Educ. Meas.* 29, 133–162. <https://doi.org/10.1111/j.1745-3984.1992.tb00371.x>.
- Mellenbergh, G.J., 1989. Item bias and item response theory. *Int. J. Educ. Res.* 13, 127–143. [https://doi.org/10.1016/0883-0355\(89\)90002-5](https://doi.org/10.1016/0883-0355(89)90002-5).
- Messick, S.J., Beaton, A.E., Lord, F.M., 1983. *NAEP Reconsidered: A New Design for a New Era*. NAEP Report No. 83-1. Educational Testing Service, Princeton, NJ. <https://files.eric.ed.gov/fulltext/ED236156.pdf>.
- Mullis, I.V.S., Martin, M.O. (Eds.), 2015. *PIRLS 2016 Assessment Framework*, second ed. <http://timssandpirls.bc.edu/pirls2016/framework.html>.
- Mullis, I.V.S., Martin, M.O. (Eds.), 2017. *TIMSS 2019 Assessment Frameworks*. <http://timssandpirls.bc.edu/timss2019/frameworks/>.
- Mullis, I.V.S., Martin, M.O. (Eds.), 2019. *PIRLS 2021 Assessment Frameworks*. <https://timssandpirls.bc.edu/pirls2021/frameworks/>.
- Mullis, I.V.S., Martin, M.O., Fishbein, B., Foy, P., Moncaleano, S., 2021. In: *Findings From the TIMSS 2019 Problem Solving and Inquiry Tasks*. <https://timssandpirls.bc.edu/timss2019/psi/>.
- OECD, 2017. *PISA 2015 Technical Report*. OECD Publishing, Paris. https://www.oecd.org/pisa/data/2015-technical-report/PISA2015_TechRep_Final.pdf.
- OECD, 2019. *PISA 2018 reading framework*. In: *PISA 2018 Assessment and Analytical Framework*. OECD Publishing, Paris. <https://doi.org/10.1787/5c07e4f1-en>.
- OECD, 2020. *PISA 2018 Technical Report*. OECD Publishing, Paris. <https://www.oecd.org/pisa/data/pisa2018technicalreport/>.
- OECD, 2021. *The Assessment Frameworks of Cycle 2 of the Programme for the International Assessment of Adult Competencies*, OECD Skills Studies. OECD Publishing, Paris. <https://doi.org/10.1787/4bc2342d-en>.
- Pohl, S., Ulitzsch, E., von Davier, M., 2019. Using response times to model not-reached items due to time limits. *Psychometrika* 84 (3), 892–920.
- Pokropek, A., 2016. Grade of membership response time model for detecting guessing behaviors. *J. Educ. Behav. Stat.* 41 (3). <https://doi.org/10.3102/1076998616636618>.
- Pokropek, A., Żółtak, T., Muszyński, M., 2022. Mouse Chase: Detecting Careless and Unmotivated Responders Using Cursor Movements in Web-Based Surveys. <https://doi.org/10.31234/osf.io/9kpmn>.
- Stephens, M., Coleman, M., 2007. *Comparing PIRLS and PISA With NAEP in Reading, Mathematics, and Science* (Working Paper). U.S. Department of Education. National Center for Education Statistics, Washington, DC. Available at: <http://nces.ed.gov/Surveys/PISA/pdf/comppaper12082004.pdf>.
- Sukkarieh, J.Z., von Davier, M., Yamamoto, K., 2012. *From Biology to Education: Scoring and Clustering Multilingual Text Sequences and Other Sequential Tasks*. Research Report No. RR-12-25. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/j.2333-8504.2012.tb02307.x>.
- Ulitzsch, E., von Davier, M., Pohl, S., 2020a. A hierarchical latent response model for inferences about examinee engagement in terms of guessing and item level non-response. *Br. J. Math. Stat. Psychol.* 73, 83–112. <https://doi.org/10.1111/bmsp.12188>.
- Ulitzsch, E., von Davier, M., Pohl, S., 2020b. Using response times for joint modeling of response and omission behavior. *Multivariate Behav. Res.* 55 (3), 425–453. <https://doi.org/10.1080/00273171.2019.1643699>.
- Ulitzsch, E., Pohl, S., Khorramdel, L., Kroehne, U., von Davier, M., 2021. A Response-Time-Based Latent Response Mixture Model for Identifying and Modeling Careless and Insufficient Effort Responding in Survey Data. <https://doi.org/10.1007/s11336-021-09817-7>.
- Von Davier, M., 2020. TIMSS 2019 scaling methodology: item response theory, population models, and linking across modes (chapter 11). In: Martin, M.O., von Davier, M., Mullis, I.V.S. (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*. <https://timssandpirls.bc.edu/timss2019/methods>.
- von Davier, M., Sinharay, S., 2013. Analytics in international large-scale assessments: item response theory and population models. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, Boca Raton, FL, pp. 155–174. <https://doi.org/10.1201/b16061-12>.
- von Davier, M., Sinharay, S., Oranje, A., Beaton, A., 2006. Statistical procedures used in the national assessment of educational progress (NAEP): recent developments and future directions. In: Rao, C.R., Sinharay, S. (Eds.), *Handbook of Statistics, Psychometrics*, vol. 26. Elsevier, Amsterdam, Netherlands. [https://doi.org/10.1016/S0169-7161\(06\)26032-2](https://doi.org/10.1016/S0169-7161(06)26032-2).
- von Davier, M., Gonzalez, E., Mislevy, R., 2009. What are plausible values and why are they useful?. In: von Davier, M., Hastedt, D. (Eds.), *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments*, vol. 2, pp. 9–36. https://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02_Chapter_01.pdf.
- von Davier, M., Khorramdel, L., He, Q., Shin, H., Chen, H., 2019a. Developments in psychometric population models for technology-based large-scale assessments—an overview of challenges and opportunities. *J. Educ. Behav. Stat.* 44, 671–705. <https://doi.org/10.3102/1076998619881789>.
- von Davier, M., Yamamoto, K., Shin, H., Chen, H., Khorramdel, L., Weeks, J., Davis, S., Kong, N., Kandathil, M., 2019b. Evaluating item response theory linking and model fit for data from PISA 2000–2012. *Assess. Educ. Princ. Pol. Pract.* 26, 466–488. <https://doi.org/10.1080/0969594X.2019.1586642>.
- von Davier, M., Tyack, L., Khorramdel, L., 2022. Automated Scoring of Graphical Open-Ended Responses Using Artificial Neural Networks. <https://arxiv.org/abs/2201.01783>.
- Yamamoto, K., Lennon, M.L., 2018. Understanding and detecting data fabrication in large-scale assessments. *Qual. Assur. Educ.* 26 (2), 196–212. <https://doi.org/10.1108/QAE-07-2017-0038>.
- Yamamoto, K., Khorramdel, L., von Davier, M., Ali, U.S., Robin, F., 2013a. Scaling PIAAC cognitive data. In: OECD, *Technical Report of the Survey of Adult Skills (PIAAC)*, third ed. PIAAC, OECD Publishing https://www.oecd.org/skills/piaac/publications/PIAAC_Technical_Report_2019.pdf.
- Yamamoto, K., Khorramdel, L., von Davier, M., Ali, U.S., 2013b. Scaling outcomes. In: OECD, *Technical Report of the Survey of Adult Skills (PIAAC)*, third ed. PIAAC, OECD Publishing https://www.oecd.org/skills/piaac/publications/PIAAC_Technical_Report_2019.pdf.
- Yamamoto, K., He, Q., Shin, H.J., von Davier, M., 2017. Developing a Machine-Supported Coding System for Constructed-Response Items in PISA. Research Report No. RR-17-47. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12169>.
- Yamamoto, K., Khorramdel, L., Shin, H.J., 2018a. Introducing multistage adaptive testing into international large-scale assessments designs using the example of PIAAC. *Psychol. Test Assess. Model.* 60, 347–368. https://www.psychologie-aktuell.com/fileadmin/Redaktion/Journale/ptam_3-2018_347-368.pdf.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2018b. Multistage adaptive testing design in international large-scale assessments. *Educ. Meas.* 37, 16–27. <https://doi.org/10.1111/emip.12226>.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2019. *Introduction of Multistage Adaptive Testing Design in PISA 2018*. OECD Working Paper. <https://doi.org/10.1787/b9435d4b-en>.
- Yin, L., Foy, P., 2021. TIMSS 2023 assessment design. In: Mullis, I.V.S., Martin, M.O., von Davier, M. (Eds.), *TIMSS 2023 Assessment Frameworks*. <https://timssandpirls.bc.edu/timss2023>.
- Zehner, F., Sälzer, C., Goldhammer, F., 2016. Automatic coding of short text responses via clustering in educational assessment. *Educ. Psychol. Meas.* 76 (2), 280–303. <https://doi.org/10.1177/0013164415590022>.

Educational surveys: methodological foundations

Paul A. Jewsbury, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction and design	359
Focus on group-level inference	359
Matrix sampling design	359
Multistage sampling design	360
Emphasis on trend reporting	360
Collection of contextual information	360
Core analysis methodology	360
Item response theory	360
IRT-latent regression	361
Multiple imputations	361
Generation of plausible values	362
Interpretation of plausible values	362
Analysis of plausible values	363
Extensions and advances in analysis methodology	364
Bridge study designs	364
Adaptive testing designs	364
Process data	365
Questionnaire matrix sampling	365
Alternatives to plausible values	366
Concluding remarks	366
References	366
List of relevant websites	368

Introduction and design

Large-scale educational surveys (LSES) are designed to measure what individuals know and can do within a variety of subject domains. Examples include the National Assessment of Educational Progress (NAEP) and the National Adult Literacy Survey (NALS) in the United States; the National Educational Panel Study (NEPS) in Germany; and the Program for International Student Assessment (PISA), the International Assessment of Adult Competencies (PIAAC), the Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS) internationally. LSES typically have at least five major features in common that motivate commonalities in their methodological foundations:

Focus on group-level inference

Perhaps the most defining feature of LSES is the focus on group-level inference. Unlike other education assessments that may provide scores to individual test takers to inform high-stakes decisions such as student placement, LSES do not provide scores to individual test takers and in fact may be prevented by legislation from reporting on individual test results. Instead, LSES report on summaries of the ability distribution within groups of interest, such as the mean of ability for male students, the difference in means of ability for students with low and high socioeconomic status, and the proportion of students above a threshold of ability known as an achievement level.

Matrix sampling design

The desire to have a large pool of items in the assessment to satisfy content and validity requirements in LSES conflicts with the desire to minimize test taker burden from lengthy tests. Longer testing times may be associated with fatigue, reduced motivation, reduced participation, and disruption, particularly because LSES do not provide individual scores. To relieve this conflict, test takers receive only a sample of items from a larger item pool, known as matrix sampling. Matrix sampling in LSES is typically implemented using a balanced incomplete block (BIB) design (Messick et al., 1983). In a BIB design, items in the item pool are allocated to mutually exclusive item blocks. Test takers receive two or more item blocks in a design that ensures all blocks are observed an equal number of times in each position and all possible block pairings are observed.

Multistage sampling design

For logistical and cost reasons, LSES typically use a multistage probability sampling design. For example, NAEP samples students within schools, schools within primary sampling units, and primary sampling units within the population (Rust and Johnson, 1992; Rust, 2013).

Emphasis on trend reporting

Typically, LSES are regularly re-assessed, with a duration between surveys as short as two years (in the case of NAEP reading and mathematics at grades 4 and 8) and as long as ten years (in the case of PIAAC). While the regular assessments are rarely identical in terms of assessment design, item pool, and administration, changes to the surveys are typically minimized to reduce potential confounds on comparisons across results from surveys administered at different times (Zwick, 1992).

Collection of contextual information

In addition to the cognitive assessment of what the test taker knows and can do, LSES collect contextual information about the test taker. The contextual information may be collected from existing databases or may be collected with questionnaires provided to the test taker during the administration of the LSES. The contextual information is typically used to contextualize differences in performance and to define the groups for group-level inference but may also be used as an independent source of data separate from the cognitive assessment (Bertling et al., 2016b).

In this chapter, we first detail the core analysis methodology common to most LSES that was motivated to handle the major features detailed above. Next, we review recent research on extensions and advancements to the core analysis methodology.

Core analysis methodology

Item response theory

The design of LSES to efficiently collect information most useful for policymakers and other stakeholders introduces several technical challenges in the analysis. As test takers only receive a subset of the entire assessment under the matrix sampling design, the assessment form each test taker receives may differ in terms of content, subscale or subject domain distribution, as well as assessment difficulty. These differences in the psychometric properties of the assessments preclude the applicability of some conventional methods of aggregating the data.

In education assessment and research, best practice to estimate scores while accounting for differences in the assessment design is considered to be latent trait models (Embretson and Reise, 2000; van der Linden, 2016). LSES typically use latent trait models within the framework of item response theory (IRT), a model that is suited for the item types used in LSES (see Ackerman, this volume, and Cai, this volume, for unidimensional and multidimensional item response theory, respectively).

With IRT, different item types may be modeled with different item response functions, which describe the dependency of item responses on the latent ability of the test taker. For example, the two-parameter logistic model (2PL; Birnbaum, 1968) is commonly used for dichotomously-scored items where the correct answer cannot be selected randomly,

$$P(Y_j = 1 | \theta_i, \beta_j) = \text{logistic}(a_j(\theta_i - b_j))$$

where Y_j is the response variable for item j scored as 1 (correct) or 0 (incorrect), θ_i is the latent ability for test taker i , $\beta_j = \{a_j, b_j\}$ is the item parameters for item j , a_j is the discrimination for item j , b_j is the difficulty of item j , $P(\cdot | \cdot)$ is a conditional probability and $\text{logistic}()$ is the standard logistic function.

For other item types, other models may be used. For example, NAEP uses the three-parameter logistic model (3PL; Birnbaum, 1968) for multiple choice items and the generalized partial credit model (GPCM; Muraki, 1992) for polytomous items. Also, latent ability may be multidimensional, as in the case of a survey with multiple subscales. In this chapter, latent ability is treated as unidimensional for simplicity, but the equations may be readily generalized for the multidimensional case.

To model the pattern of item responses on the assessment, the item responses are assumed to be independent to each other conditional on latent ability (locally independent), that is,

$$P(Y = y_i | \theta_i, \beta) = \prod_{j=1}^J P(Y_j = y_{ij} | \theta_i, \beta_j)$$

where $Y = \{Y_j\}_{j=1}^J$, $y_i = \{y_{ij}\}_{j=1}^J$, y_{ij} is the response for test taker i on item j , $\beta = \{\beta_j\}_{j=1}^J$, and J is the total number of items.

When IRT is used to analyze LSES data, the matrix sampling design may be conceptualized as a missing data problem: most item responses for each test taker are missing, as most items are not presented to each test taker. The assignment of assessment forms to each test taker is carefully randomized to ensure this missing data can be classified as Missing Completely at Random (MCAR; Rubin, 1976; see Johnson, 1992).

LSES typically use marginal maximum likelihood (MML) estimation via an expectation-maximization (EM) algorithm (Bock and Aitkin, 1981; Dempster et al., 1977). In MML, items that were not presented to a test taker are ignored in that test taker's likelihood calculation (Rubin, 1976). Specifically, the following likelihood is maximized to estimate the parameters of the IRT model:

$$L(\beta, \mu, \sigma^2 | y) = \prod_{i=1}^N w_i \int \prod_{j \in s_i} P(Y_j = y_{ij} | \theta_i, \beta_j) f(\theta_i | \mu, \sigma^2) d\theta_i$$

where $L(\cdot | \cdot)$ is a likelihood, μ is the mean of the latent variable, σ^2 is the variance of the latent variable, $y = \{y_i\}_{i=1}^N$, N is the total number of test takers, w_i is the sampling weight for test taker i , and s_i is the set of items that test taker i received.

MML estimation is consistent provided that any ignored missing data is Missing at Random (MAR; Rubin, 1976). MCAR missing data, such as that resulting from matrix sampling, is MAR. Furthermore, provided that the IRT model is specified appropriately, missing data due to an adaptive or targeted testing design is also MAR (e.g., Eggen and Verhelst, 2011; Jewsbury and van Rijn, 2020, 2021), as well as missing data due to time limits (Mislevy, 2016).

IRT-latent regression

While IRT enables the analysis of LSES data on a common ability metric, the core technical challenge is the estimation of the distribution of ability for each group of interest. With this aim, LSES use IRT-latent regression models (IRT-LRM; Mislevy, 1984, 1985). IRT-latent regression models can be estimated by maximizing the following marginal likelihood,

$$L(\Gamma, \sigma^2 | x, y, \beta) = \prod_{i=1}^N w_i \int \prod_{j \in s_i} P(Y_j = y_{ij} | \theta_i, \beta_j) f(\theta_i | \Gamma, \sigma^2, x_i) d\theta_i$$

where $x = \{x_i\}_{i=1}^N$, x_i is the vector of contextual information for test taker i , Γ is the vector of regression coefficients, σ^2 is the residual variance, and $f(\cdot | \cdot)$ is a density. Usually, θ_i conditional on x_i is assumed to be normally distributed, with a mean of $\Gamma'x_i$ and a variance of σ^2 .

The parameters are typically estimated via an EM algorithm (Dempster et al., 1977; Mislevy, 1984, 1985). The E-step is often performed with either numerical integration or Laplace approximation (Thomas, 1993), depending on the dimensionality of ability. The item parameters are estimated prior to the IRT-latent regression, and the estimates of the item parameters are treated as the known values of β when estimating the latent regression parameters $\{\Gamma, \sigma^2\}$ (see Thomas and Gan, 1997).

Based on the IRT-latent regression, estimators for group-level parameters of interest may be defined for any group g , where a dummy coded variable indicating membership in group g is included in x . For example, the estimator for the mean of ability in group g is

$$\hat{\mu}_g = \hat{\Gamma}' \bar{x}_g$$

where $\hat{\Gamma}$ is the maximum likelihood estimator for Γ and $\bar{x}_g$ is the expectation of x conditional on group g .

The estimator for the variance of ability in group g is,

$$\hat{\sigma}_g^2 = \hat{\Gamma}' S_{xg} \hat{\Gamma} + \hat{\sigma}^2$$

where S_{xg} is the variance-covariance matrix of x conditional on group g .

Multiple imputations

While the institutional machinery of IRT-latent regression models estimated with marginal maximum likelihood could be used to yield estimators for group-level parameters of interest as described above, the widespread interest in analyzing LSES data prompts a more accessible approach to analysis.

To aid in secondary user analysis and allow for official results to be easily replicated by secondary users, LSES use imputation methods to produce imputed values for latent proficiency. The intention underlying the use of imputed values is to allow secondary users to calculate estimates for group-level parameters of interest that are consistent with the IRT-latent regression estimators described above, but with more familiar statistical methods.

While methods to produce a single imputed value for latent proficiency for each respondent are available, methods involving multiply imputed values are generally recommended when there is substantial uncertainty in the imputation (Sinharay et al., 2001). Specifically, the method of multiple imputation (Rubin, 1987) is commonly recommended as a statistically principled approach for dealing with missing data (e.g., Schafer and Graham, 2002), enabling consistent estimation of parameters in data with missingness under the assumption the missing data is missing at random (MAR; Rubin, 1976).

Multiple imputation has two benefits over many other imputation methods. First, multiple imputation is justified by statistical theory, in contrast to some common methods that are ad hoc and generally produce invalid inferences such as mean substitution (Rässler et al., 2013). Relatedly, these common methods produce estimates that are inconsistent with the IRT-latent regression estimators. Second, the use of multiple imputations instead of a single imputation both reduces the uncertainty in estimates associated

with the imputation process as well as enables the uncertainty due to imputation to be accounted for in the standard errors (Rubin, 1976; Sinharay et al., 2001). As a result, the institution sponsoring the large-scale assessment creates the imputed data for both official and secondary user analysis and reporting. The institution packages the imputations in data products for secondary users for them to replicate published reports and create new ones of their own. This serves as a measure that provides transparency to the officially reported results and allows for external quality control and verification of the reports.

A key concept in multiple imputation is congeniality, where the analysis model applied to the imputed data must correspond to the imputation model used to generate the imputed values (Meng, 1994). For example, consider a regression analysis where some values on the outcome variable are MAR. Multiple imputation may be used to impute values for these missing responses, allowing an institution to provide a completed dataset to secondary users. While the imputed values are generally not the responses that would have been observed, the secondary user may still conduct the regression analysis with the imputed values and obtain the correct regression parameter estimates. In this context, congeniality would be guaranteed if the predictors used by the secondary user were also used as predictors in a regression-based imputation model (Meng, 1994).

Generation of plausible values

Multiple imputation is usually applied to impute values for missing responses on observed variables (Rubin, 1987). Mislevy's (1991) insight was that the latent variable in IRT representing ability may be considered a variable missing for all test takers. As missingness on the latent variable in IRT is guaranteed to be MAR, the assumption underlying multiple imputation, values may be imputed on the latent variable for all test takers. With values imputed for the latent variable, secondary users may apply standard statistical methods that require complete data, such as least-squares regressions and *t*-tests. With plausible values, the challenges associated with the latency of ability and the massive amount of missingness in the item responses are eliminated for the secondary user.

Specifically, plausible values are generated as random draws from an approximation to each test taker's conditional ability distribution,

$$f(\theta_i | x_i, \gamma_i, \beta, \Gamma, \sigma^2) \propto \prod_{j \in S_i} P(Y_j = y_{ij} | \theta_i, \beta_j) f(\theta_i | \Gamma, \sigma^2, x_i)$$

Multiple plausible values are drawn for each test taker to account for uncertainty in the test taker's ability. Specific details and approximations made in this procedure are provided in the technical documentation for each program (e.g., National Center for Education Statistics, 2018; Organization for Economic Co-operation and Development, 2016, 2018), and may differ between programs.

In the context of LSES, the requirement of congeniality implies that in order to support a potential secondary user's analysis involving estimating the mean ability for a given group, a dummy coded variable indicating membership in the group should be accounted for in the IRT-latent regression model (Mislevy, 1991). For example, a categorical ethnicity variable with seven categories may be converted into six dummy coded variables, with one category treated as the reference category and one dummy coded variable indicating membership in each of the other six categories. To account for the dummy coded variable, the variable may be included as a predictor in the regression (within x). But if the variable is independent of ability conditional on the predictors included in the regression, the variable does not need to be included (Meng, 1994; Mislevy, 1991). Note that the amount of bias that may be introduced by a non-congenial analysis is not only dependent on the magnitude of the relationship between ability and that variable unaccounted for by the variables included in the regression model, but also the reliability of the cognitive assessment (Marsman et al., 2016).

To account for many potential secondary user analyses, many predictors are typically included in the IRT-latent regression model (Mislevy, 1991). This follows generally recommended practices in the multiple imputation literature that support being as inclusive as possible in constructing the imputation model (Collins et al., 2001). To reduce overfitting, principal component analysis is typically applied to the full set of contrast-coded contextual variables as a data reduction technique. In practice, the predictors in the regression are the subset of the principal components that explain a predetermined amount of variance in the contextual variables (von Davier et al., 2006).

While LSES may have been the earliest use of multiple imputation for latent variables (see Rubin, 1996), the method has been applied more widely in other latent variable contexts more recently (e.g., Boeschoten et al., 2018; Carlin, 1992; Dicke et al., 2015; Gortler et al., 2015; Joinson et al., 2012; Rhee et al., 2013). The popular structural equation modeling software Mplus, for example, enables researchers to generate plausible values for a large class of latent trait models (see Asparouhov and Muthén, 2010). Many of these applications employ plausible values for the purpose of missing data treatment even in primary user analysis, as recommended by some authors (Bray et al., 2015; Graham, 2003), similar to how multiple imputation is often used as a general purpose missing data treatment with missingness in observed variables (Schafer and Graham, 2002; Sinharay et al., 2001).

Interpretation of plausible values

While this chapter describes the history of plausible values as a generalization of Rubin's multiple imputation methodology for missing data, plausible values may also be understood within a Bayesian framework without reference to missing data. Here,

plausible values are random draws from each test taker's ability distribution conditional on their item responses, contextual information, and the IRT-latent regression model parameter estimates.

Within this framework, the distinction between plausible values in large-scale educational assessments and conventional latent scores used in individual-score assessments is clarified:

- (1) Plausible values are random draws from each test taker's conditional ability distribution, which may be associated with greater bias and error for individual-level inference than some other statistics such as the mean or mode (expected a posteriori, EAP, or maximum a posteriori, MAP, respectively; [Mislevy et al., 1992](#)).
- (2) Plausible values in LSES are drawn from ability distributions conditional on both item responses and contextual information. As the contextual information is not part of the assessment, individual-level inference with these plausible values may be inappropriate and unfair.

Notably, with either interpretation of plausible values—as either multiply imputed values for the missing ability variable or as random draws from each test taker's conditional ability distribution—the same key conclusions are reached. Within the multiple imputation interpretation, plausible values are intermediate values that may be used in place of ability scores so that simple statistical methods such as *t*-tests may be applied in order to reproduce the results from the IRT-latent regression imputation model. As imputations of missing values, plausible values for any given test taker may not best represent the ability of the test taker. Within the Bayesian interpretation, plausible values are not statistically optimal for individual-level inference and may not facilitate fair individual-level inference due to the incorporation of contextual information in the plausible value generation. With both interpretations, plausible values are used to calculate estimates of summaries of latent ability distributions, conditional on contextual variables of interest, but are not suited for individual-level inference.

Analysis of plausible values

Plausible values, or multiple imputations more generally, result in M complete datasets: one for each set of imputations. A general principle in the analysis of plausible values is that each of the M complete datasets may be analyzed with standard complete-data analysis methods, as if latent ability for every test taker had been directly observed and recorded. Once the standard complete-data analysis method has been applied to each of M complete datasets, the results are averaged across the datasets to produce the final point estimates.

For example, an estimate of the mean of latent ability for a group g is calculated within each set of plausible values,

$$\hat{\mu}_g^{(m)} = \left(\sum_{i \in s_g} w_i \right)^{-1} \sum_{i \in s_g} w_i \hat{\theta}_i^{(m)}$$

where $\hat{\mu}_g^{(m)}$ is the estimated ability mean for group g using the m th set of plausible values, s_g is the set of all test takers in group g , $\hat{\theta}_i^{(m)}$ is the m th plausible value for test taker i .

Once M estimates are obtained, these estimates may be averaged to obtain the final estimate,

$$\hat{\mu}_g = \frac{1}{M} \sum_{m=1}^M \hat{\mu}_g^{(m)}$$

This procedure can be applied more generally to other types of statistics,

$$\hat{G}_g(\hat{\theta}) = \frac{1}{M} \sum_{m=1}^M \hat{G}_g^{(m)}(\hat{\theta})$$

For example, $\hat{G}_g(\hat{\theta})$ could be a standard deviation, so

$$\hat{G}_g^{(m)}(\hat{\theta}) = \sqrt{\left(\sum_{i \in s_g} w_i - 1 \right)^{-1} \sum_{i \in s_g} w_i \left(\hat{\theta}_i^{(m)} - \hat{\mu}_g^{(m)} \right)^2}$$

And $\hat{G}_g^{(m)}$ could be the proportion of test takers within an achievement level, defined as ability within the range of $T_L \leq \theta < T_U$,

$$\hat{G}_g^{(m)}(\hat{\theta}) = \left(\sum_{i \in s_g} w_i \right)^{-1} \sum_{i \in s_g} w_i I(T_L \leq \hat{\theta}_i^{(m)} < T_U)$$

where $I(\cdot)$ is an indicator function that is 1 if the condition is satisfied or 0 otherwise.

The variance of the estimator is estimated as the sum of two components according to Rubin's rules ([Rubin, 1987](#)),

$$\widehat{\text{var}}(\hat{G}_g(\hat{\theta})) = \widehat{\text{var}}_{\text{WI}}(\hat{G}_g(\hat{\theta})) + (1 + M^{-1}) \widehat{\text{var}}_{\text{BI}}(\hat{G}_g(\hat{\theta}))$$

The first component, $\widehat{\text{var}}_{\text{WI}}(\widehat{G}_g(\widehat{\theta}))$, denotes the variance component within imputations (plausible values), and is calculated as the mean variance estimate with each set of imputations,

$$\widehat{\text{var}}_{\text{WI}}(\widehat{G}_g(\widehat{\theta})) = \frac{1}{M} \widehat{\text{var}}_{\text{WI}}^{(m)}(\widehat{G}_g(\widehat{\theta}))$$

where $\widehat{\text{var}}_{\text{WI}}^{(m)}(\widehat{G}_g(\widehat{\theta}))$ is the estimated variance of $\widehat{G}_g(\widehat{\theta})$ when estimated from the m th set of plausible values.

To account for the complex sampling design in large-scale assessments, replicate weights from a resampling procedure such as the jackknife (Rust and Johnson, 1992) are provided by the institution for use in the calculation of the within-imputation variance component,

$$\widehat{\text{var}}_{\text{WI}}^{(m)}(\widehat{G}_g(\widehat{\theta})) = \sum_{p=1}^P \left(\widehat{G}_g^{(m,p)}(\widehat{\theta}) - \widehat{G}_g^{(m)}(\widehat{\theta}) \right)^2$$

where $\widehat{G}_g^{(m,p)}(\widehat{\theta})$ is the estimate calculated using the m th plausible value and the p th replicate weight.

In practice, the within-imputation variance component may be approximated using the estimated variance component based on the first plausible value only (as done in NAEP; von Davier et al., 2006) as

$$\widehat{\text{var}}_{\text{WI}}(\widehat{G}_g(\widehat{\theta})) \approx \widehat{\text{var}}_{\text{WI}}^{(1)}(\widehat{G}_g(\widehat{\theta}))$$

The second component $\widehat{\text{var}}_{\text{BI}}(\widehat{G}_g(\widehat{\theta}))$ denotes the variance component between imputations (plausible values), and is calculated as

$$\widehat{\text{var}}_{\text{BI}}(\widehat{G}_g(\widehat{\theta})) = \frac{1}{M-1} \sum_{p=1}^P \left(\widehat{G}_g^{(m)}(\widehat{\theta}) - \sum_{m=1}^M \widehat{G}_g^{(m)}(\widehat{\theta}) \right)^2$$

Accessible introductions including practical examples on analyzing plausible values include Mislevy et al. (1992), von Davier et al. (2009), and Wu (2005).

Extensions and advances in analysis methodology

Bridge study designs

Given the emphasis on the comparability of results from assessments administered in different years, LSES balance two competing pressures (Beaton, 1990a). First, the pressure to maintain equivalence in the measurement instrument across years to minimize confounding by instrument differences (Beaton, 1990b). Second, the pressure to adapt to changes in the interaction of test takers and the measurement instrument, changes in the skills and knowledge most relevant to stake holders, and methodological advancements (Zwick, 1992).

When changes to the assessment or administration are required, a bridge study is typically employed to study the impact of these changes and to link results from the new administration to results from previous administrations. For example, NAEP has employed bridge studies to link results after transitioning to assess each student in only one subject area (Johnson and Zwick, 1990), standardizing and redesigning the NAEP long-term trend assessments (Perie et al., 2005), and more recently, transitioning from a paper-based assessment to a digitally based assessment (Jewsbury et al., 2020, 2021a).

While bridge studies can differ in methods to evaluate the impact of the changes and to link the results, the most common approach is to administer the new and previous administrations to two randomly equivalent samples. With randomly equivalent samples, the impact of the change may be evaluated without confounds from population differences. Furthermore, a common population linking method may be employed to link the results (Mazzeo and von Davier, 2013; Yamamoto and Mazzeo, 1992). Common population linking assumes that the new administration still measures the same construct as the previous administration and that both samples have been drawn from the same population. The theoretical properties of error introduced by such linking, as well as methods to extend the conventional approach to LSES variance estimation to account for linking in bridge studies, can be found in Jewsbury (2019).

Adaptive testing designs

With the transition of several LSES to digitally based assessments, interactive assessments where item selection adapts to the performance of the test taker in real time become possible. Such designs are referred to as computerized adaptive tests (CATs) and serve to increase measurement precision by selecting items during the assessment that are more informative about ability (Wainer et al., 1990).

To balance content and quality control in LSES, instead of adapting after each item, adapting after blocks of items may be preferred (Oranje et al., 2014; Yamamoto et al., 2018). Block-level adaptive tests are known as multistage tests (MSTs), and have been implemented in PISA and PIAAC (Yamamoto et al., 2018). Introducing MST to the 2018 PISA assessment improved measurement precision of individual test takers by 4–5% (Shin et al., 2021).

Latent trait models, such as IRT used in LSES, enable the design and analysis of adaptive tests (Lord, 1980; Wainer et al., 1990). As described earlier in this chapter, LSES methodology involving marginal maximum likelihood estimation allows for missing data in the item responses provided that the missing data is missing at random (MAR). Missing data resulting from adaptation in an adaptive test is MAR provided that the model used to estimate the item parameters is appropriate with respect to the features of the adaptive test design (e.g., Eggen and Verhelst, 2011; Jewsbury and van Rijn, 2020, 2021).

Process data

Process data are data that describe cognitive and behavioral processes and are a new type of contextual information that may be collected by digitally-based LSES (Bergner and von Davier, 2019; Ercikan and Pellegrino, 2017; Nicolay et al. this volume). Examples of process data are response time, and the sequence of actions a test taker took in completing an item. Process data has substantial utility to improve the reliability and validity of assessments. For example, to evaluate motivation and effort (Michaelides et al., 2020); to screen for data quality (Yamamoto and Lennon, 2018); and to inform missing data treatment and identification of guessing behaviors (Ulitzsch et al., 2020a,b). In addition, process data may help contextualize ability differences between groups and years, and inclusion of process data in the imputation model for the plausible values may reduce non-congeniality related biases in the results by contributing to the explanation of ability (von Davier et al., 2019).

Process data is methodologically distinct from traditional contextual variables, as process data may not be measured independently of the item responses. Consequently, the assumption of conditional independence between item responses and contextual variables given ability, made in the IRT-latent regression (Mislevy, 1984, 1985), may be less reasonable for process data. For example, use of the correct strategy to solve an item may be directly related to the accuracy of the response, even after controlling for ability.

Response time is a type of process data that has been extensively studied (e.g., Molenaar et al., 2015; van der Linden, 2007; van Rijn and Ali, 2018). The issue of potential conditional dependences between item responses and response time has been considered, and several models to jointly model item responses and response time while accounting for conditional dependencies have been proposed (e.g., Bolsinova et al., 2017; Bolsinova and Molenaar, 2018; Molenaar et al., 2016).

While models that account for item response and process data conditional dependencies are available (De Boeck and Jeon, 2019), the applicability of these models to the methodology in LSES is less clear. While IRT-latent regression models that include parameters to account for conditional dependencies have been proposed, accounting for the conditional dependencies in the analysis of plausible values may not be straightforward (Jewsbury et al., 2021b). In the context of response time, the conditional dependencies are small relative to the improvement in the measurement of ability, potentially justifying inclusion of response time into the IRT-latent regression model even if the conditional dependencies are not modeled (Shin et al., under review; Shin et al., in press). However, this result may not generalize to other types of process data, and the accumulation of conditional dependencies from inclusion of many process data variables may result in non-negligible estimation biases.

Questionnaire matrix sampling

Many LSES collect additional contextual information about the test taker through a questionnaire, often administered after the cognitive assessment. While the cognitive assessment is generally matrix sampled, the questionnaire typically involves only one form administered to every test taker. This questionnaire form is necessarily constrained in length, given time constraints on assessment time and findings that lengthy questionnaires reduce response quality (Galesic and Bosnjak, 2009). Consequently, administering one of multiple questionnaire forms to each test taker has been proposed to increase the total number of questionnaire items administered. These forms may involve a matrix sampling design, akin to the cognitive assessment, to ensure all pairs of items are observed (Bertling et al., 2016a).

Early work proposed and evaluated modeling approaches for questionnaire matrix sampling designs (Adams et al., 2013), and several concerns were raised about the validity of these approaches (von Davier, 2013). In particular, concerns were raised that questionnaire matrix sampling necessarily introduces missing data into the questionnaire responses and therefore the predictors in the IRT-latent regression. The standard approach to missing data in the predictors in LSES, the missing indicator method (Cohen et al., 2014), can introduce bias into the regression coefficient estimates (Jones, 1996; Rutkowski, 2011; Rutkowski and Zhou, 2015). However, Grund et al. (2021) recently showed that the use of the missing indicator method in the IRT-latent regression can facilitate unbiased estimation with plausible values, provided that the missing data is MCAR and that the plausible value analysis uses pairwise estimation.

Several alternative approaches have been proposed to the missing data indicator method. Kaplan and Su (2016, 2018) evaluated imputing the missing data on the predictors prior to the IRT-latent regression and showed that bias in reproducing correlations depended more on the matrix sampling design than the imputation method. Weirich et al. (2014) showed that a two-stage multiple imputation involving multiple imputations for the questionnaire missingness prior to the imputation of ability was superior to the use of a single imputation. Finally, when test takers receive items for all questionnaire constructs (or scales), but only a subset of items for any given construct, construct scores may substitute for the original questionnaire response variables in the IRT-latent regression (Shin et al., in preparation).

Alternatives to plausible values

Plausible values generated from IRT-latent regression models are an excellent solution to the design issues common in LSES. Plausible values provide analysts with numbers for latent ability that enable analysis with standard statistical methods commonly available in general purpose statistical software. Complexities associated with the latency of ability, as well as missing data due to both matrix sampling and potentially adaptive testing, are resolved by the institution prior to the analysis. Furthermore, plausible values are particularly suited for group-level inference over common alternative scores. For example, the distribution of common IRT-based scores such as maximum likelihood estimates (MLEs) and expected a posteriori scores (EAPs) can substantially diverge from the distribution of ability (Mislevy et al., 1992), making these scores suboptimal for group-level inference.

Plausible values do, however, have two notable limitations. First, while plausible values were motivated to simplify analysis, plausible value-based analysis is still relatively complex. Recommended use of plausible values, as described above, requires repeating the analysis multiple times for each set of plausible values and each replicate weight. Accounting for bridge study linking designs is a further complication in standard error calculation on certain trend comparisons (Jewsbury, 2019). Plausible values are also theoretically complex, with issues such as the meaning of plausible values, the use of contextual information to generate plausible values, and that different estimates of ability may be suited for different levels of inference, which has necessitated several tutorials on plausible values (e.g., Braun and von Davier, 2017; Mislevy et al., 1992; Wu, 2005). With the rise in popularity of structural equation modeling and associated software (e.g., Mplus, Muthén and Muthén, 2017), many contemporary analysts may be more familiar with IRT-latent regression, a type of structural equation model, than plausible values.

Second, the requirement that all inferences are conducted with a single dataset of plausible values places technical constraints on the institutional analysis of LSES. While institutions intend to ensure as many analyses as possible are approximately congenial with the IRT-latent regression imputation model (Mislevy, 1991), risk of overfitting the IRT-latent regression model may limit how many analyses can be accommodated. Not all contextual variables may be accommodated, and not all uses of those contextual variables may be supported (e.g., Schofield et al., 2015). Finally, certain extensions such as some methods to better handle missing data in the contextual variables (von Davier, 2013) and model process data and item response dependencies (Bolsinova et al., 2017) may be difficult or impossible to implement with plausible value methodology.

Several methods as alternatives to plausible values could be implemented via user-friendly tools to conduct analysis directly with the item response data. For example, the National Center of Education Statistics (NCES) and the Organization for Economic Co-operation and Development (OECD) provide Data Explorers of various surveys (see List of Relevant Websites, below). The Data Explorers can be used to conduct numerous common analyses online. While such tools involve the analysis of plausible values, the tools could be modified to conduct an IRT-based analysis using a model customized for the desired inference. This approach would provide institutions greater flexibility to incorporate methodological advances and innovations, while avoiding any constraints required to support plausible value-based analysis.

Concluding remarks

This chapter describes the core methodology common across many LSES. The core methodology, analysis based on plausible values generated from IRT-latent regression models, was described. Select extensions to the methodology and current areas of research were summarized. Note that summarizing several ongoing research areas was beyond the scope of this introductory chapter, such as analyzing locally dependent item responses and multidimensional constructs (e.g., von Davier et al., in press), research on hierarchical latent regression (e.g., Li et al., 2009), fully Bayesian estimation (e.g., Johnson and Jenkins, 2004), model selection (e.g., Culpepper and Park, 2017), and alternate methods that relax technical assumptions in the estimation (e.g., Andersson and Xin, 2021; von Davier and Sinharay, 2010).

Extensions to the LSES methodology may be categorized as either evolutionary or revolutionary. Evolutionary extensions advance the method in terms of either power or flexibility, but maintain the constraint that analysis is conducted with plausible values. Revolutionary extensions consider approaches that are not premised on the use of plausible values. Underpinning any discussion of LSES methodology is LSES assessment design: plausible values were developed to support group-level inference when individual-level reliability of the assessment is low (Mislevy et al., 1992). While alternatives to the assessment design of LSES is not considered in this chapter, assessment designs that ensure an adequate minimum level of individual-level reliability may facilitate other methodological approaches. For example, the replacement of plausible values with another score that is approximately distributed as ability under these conditions (e.g., weighted likelihood estimates, WLEs; Warm, 1989).

Regardless of the approach taken, LSES methodology remains an active and exciting area of research. Indeed, since the original introduction of the LSES core methodology (Mislevy, 1984, 1985, 1991), LSES remain one of the most prominent examples of advanced psychometrics (Oranje and Kolstad, 2019).

References

- Adams, R.J., Lietz, P., Berezner, A., 2013. On the use of rotated context questionnaires in conjunction with multilevel item response models. *Large Scale Assess. Educ.* 1 (1), 1–27.
- Andersson, B., Xin, T., 2021. Estimation of latent regression item response theory models using a second-order Laplace approximation. *J. Educ. Behav. Stat.* 46 (2), 244–265.

- Asparouhov, T., Muthén, B.O., 2010. Plausible Values for Latent Variables Using Mplus. Mplus Technical Appendix. Muthén and Muthén, Los Angeles, CA.
- Beaton, A.E., 1990a. Introduction. In: Johnson, E.G., Zwick, R. (Eds.), *Focusing the New Design: The NAEP 1988 Technical Report*. Educational Testing Service, Princeton, NJ, pp. 3–9. Report No. 19-TR-20.
- Beaton, A.E., 1990b. Introduction. In: Beaton, A.E., Zwick, R. (Eds.), *The Effect of Changes in the National Assessment: Disentangling the NAEP 1985–86 Reading Anomaly*. Educational Testing Service, Princeton, NJ, pp. 1–13. Report No. 17-TR-21.
- Bergner, Y., von Davier, A.A., 2019. Process data in NAEP: past, present, and future. *J. Educ. Behav. Stat.* 44 (6), 706–732.
- Bertling, J.P., Borgonovi, F., Almonte, D.E., 2016a. Psychosocial skills in large-scale assessments: trends, challenges, and policy implications. In: Lipnevich, A.A., Preckel, F., Roberts, R.D. (Eds.), *Psychosocial Skills and School Systems in the 21st Century: Theory, Research, and Practice*. Springer, New York, NY, pp. 347–372.
- Bertling, J.P., Marksteiner, T., Kyllonen, P.C., 2016b. General noncognitive outcomes. In: Kuger, S., Klieme, E., Jude, N., Kaplan, D. (Eds.), *Assessing Contexts of Learning*. Springer, New York, NY, pp. 255–281.
- Birnbaum, A., 1968. On the Estimation of Mental Ability. Series Report No. 15). USAF School of Aviation Medicine, Randolph Air Force Base, TX.
- Bock, R.D., Aitkin, M., 1981. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika* 46 (4), 443–459.
- Boeschoten, L., Oberski, D.L., De Waal, T., Vermunt, J.K., 2018. Updating latent class imputations with external auxiliary variables. *Struct. Equ. Model.* 25, 750–761.
- Bolsinova, M., Molenaar, D., 2018. Modeling nonlinear conditional dependence between response time and accuracy. *Front. Psychol.* 9 (1525), 1–12.
- Bolsinova, M., de Boeck, P., Tijmstra, J., 2017. Modelling conditional dependence between response time and accuracy. *Psychometrika* 82 (4), 1126–1148.
- Braun, H., von Davier, M., 2017. The use of test scores from large-scale assessment surveys: psychometric and statistical considerations. *Large Scale Assess. Educ.* 5 (1), 1–16.
- Bray, B.C., Lanza, S.T., Tan, X., 2015. Eliminating bias in classify-analyze approaches for latent class analysis. *Struct. Equ. Model.* 22, 1–11.
- Carlin, J.B., 1992. Meta-analysis for 2×2 tables: a Bayesian approach. *Stat. Med.* 11, 141–158.
- Cohen, P., West, S.G., Aiken, L.S., 2014. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Psychology press, New York, NY.
- Collins, L.M., Schafer, J.L., Kam, C.M., 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol. Methods* 6, 330–351.
- Culpepper, S.A., Park, T., 2017. Bayesian estimation of multivariate latent regression models: gauss versus Laplace. *J. Educ. Behav. Stat.* 42 (5), 591–616.
- De Boeck, P., Jeon, M., 2019. An overview of models for response times and processes in cognitive tests. *Front. Psychol.* 10, 102.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B* 39 (1), 1–22.
- Dicke, T., Parker, P.D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., Leutner, D., 2015. Beginning teachers' efficacy and emotional exhaustion: latent changes, reciprocity, and the influence of professional knowledge. *Contemp. Educ. Psychol.* 41, 62–72.
- Eggen, T.J.H.M., Verhelst, N.D., 2011. Item calibration in incomplete testing designs. *Psicologica* 32, 107–132.
- Embretson, S.E., Reise, S., 2000. *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates Inc., Mahwah, NJ.
- Erican, K., Pellegrino, J.W. (Eds.), 2017. *Validation of Score Meaning for the Next Generation of Assessments: The Use of Response Processes*. Routledge, New York, NY.
- Galesic, M., Bosnjak, M., 2009. Effects of questionnaire length on participation and indicators of response quality in a web survey. *Publ. Opin. Q.* 73 (2), 349–360.
- Gortler, R., Fox, J.P., Twisk, J.W., 2015. Why item response theory should be used for longitudinal questionnaire data analysis in medical research. *BMC Med. Res. Methodol.* 15 (55), 1–12.
- Graham, J.W., 2003. Adding missing-data-relevant variables to FIML-based structural equation models. *Struct. Equ. Model.* 10, 80–100.
- Jewsbury, P.A., van Rijn, P.W., 2020. IRT and MIRT models for item parameter estimation with multidimensional multistage tests. *J. Educ. Behav. Stat.* 45 (4), 383–402.
- Jewsbury, P.A., van Rijn, P.W., 2021. Item calibration in multistage tests. In: Yan, D. (Ed.), *Research for Practical Issues and Solutions in Computerized Multistage Testing*. Chapman and Hall/CRC, Boca Raton, FL.
- Jewsbury, P.A., Finnegan, R., Xi, N., Jia, Y., Rust, K., Burg, S., 2020. 2017 NAEP Transition to Digitally Based Assessments in Mathematics and Reading at Grades 4 and 8: Mode Evaluation Study. National Center of Education Statistics. Retrieved from: https://nces.ed.gov/nationsreportcard/subject/publications/main2020/pdf/transitional_whitepaper.pdf.
- Jewsbury, P.A., Jia, Y., Xi, N., 2021a. Effects of mode transition on instruments and subpopulation performance in NAEP. In: Khorramdel, L., von Davier, M., Yamamoto, K. (Eds.), *Innovative Computer-Based International Large-Scale Assessments—Foundations, Methodologies and Quality Assurance Procedures*. Springer, New York, NY.
- Jewsbury, P.A., Shin, H.J., van Rijn, P.W., 2021b. Time-Accuracy Conditional Dependencies and Latent Regression Models. Presented at the National Council on Measurement in Education, Virtual Conference, 9 June 2021.
- Grund, S., Lüdtke, O., Robitzsch, A., 2021. On the treatment of missing data in background questionnaires in educational large-scale assessments: an evaluation of different procedures. *J. Educ. Behav. Stat.* 46 (4), 430–465.
- Jewsbury, P.A., 2019. Error Variance in Common Population Linking Bridge Studies. Research Report No. RR-19-42. Educational Testing Service, Princeton, NJ.
- Johnson, M.S., Jenkins, F., 2004. A Bayesian Hierarchical Model for Large-Scale Educational Surveys: An Application to the National Assessment of Educational Progress. Report No. RR-04-38. Educational Testing Service, Princeton, NJ, pp. 1–28.
- Johnson, E.G., Zwick, R., 1990. *Focusing the New Design: The NAEP 1988 Technical Report*. Report No. 19-TR-20. Educational Testing Service, Princeton, NJ, pp. 3–9.
- Johnson, E.G., 1992. The design of the national assessment of educational progress. *J. Educ. Meas.* 29 (2), 95–110.
- Joinson, C., Heron, J., Araya, R., Paus, T., Croudace, T., Rubin, C., et al., 2012. Association between pubertal development and depressive symptoms in girls from a UK cohort. *Psychol. Med.* 42, 2579–2589.
- Jones, M.P., 1996. Indicator and stratification methods for missing explanatory variables in multiple linear regression. *J. Am. Stat. Assoc.* 91 (433), 222–230.
- Kaplan, D., Su, D., 2016. On matrix sampling and imputation of context questionnaires with implications for the generation of plausible values in large-scale assessments. *J. Educ. Behav. Stat.* 41 (1), 57–80.
- Kaplan, D., Su, D., 2018. On imputation for planned missing data in context questionnaires using plausible values: a comparison of three designs. *Large Scale Assess. Educ.* 6 (1), 1–31.
- Li, D., Oranje, A., Jiang, Y., 2009. On the estimation of hierarchical latent regression models for large-scale assessments. *J. Educ. Behav. Stat.* 34 (4), 433–463.
- Lord, F.M., 1980. *Applications of Item Response Theory to Practical Testing Problems*. Lawrence Erlbaum, Hillsdale, NJ.
- Marsman, M., Maris, G., Bechger, T., Glas, C., 2016. What can we learn from plausible values? *Psychometrika* 81 (2), 274–289.
- Mazzeo, J., von Davier, M., 2013. Linking scales in international large-scale assessments. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. Chapman and Hall/CRC, Boca Raton, FL.
- Meng, X.L., 1994. Multiple-imputation inferences with uncongenial sources of input. *Stat. Sci.* 9 (4), 538–558.
- Messick, S.J., Beaton, A.E., Lord, F.M., 1983. NAEP Reconsidered: A New Design for a New Era. NAEP Report No. 83-1. Educational Testing Service, Princeton, NJ.
- Michaelides, M.P., Ivanova, M., Nicolaou, C., 2020. The relationship between response-time effort and accuracy in PISA science multiple choice items. *Int. J. Test.* 20 (3), 187–205.
- Mislevy, R.J., Beaton, A.E., Kaplan, B., Sheehan, K.M., 1992. Estimating population characteristics from sparse matrix samples of item responses. *J. Educ. Meas.* 29 (2), 133–161.
- Mislevy, R.J., 1984. Estimating latent distributions. *Psychometrika* 49, 359–381.
- Mislevy, R.J., 1985. Estimation of latent group effects. *J. Am. Stat. Assoc.* 80, 993–997.
- Mislevy, R.J., 1991. Randomization-based inference about latent variables from complex samples. *Psychometrika* 56, 177–1.
- Mislevy, R.J., 2016. Missing responses in item response theory. In: van der Linden, W.J. (Ed.), *Handbook of Item Response Theory*, vol. 2. Chapman and Hall/CRC, Boca Raton, FL.
- Molenaar, D., Tuerlinckx, F., van der Maas, H.L., 2015. A generalized linear factor model approach to the hierarchical framework for responses and response times. *Br. J. Math. Stat. Psychol.* 68 (2), 197–219.
- Molenaar, D., Oberski, D., Vermunt, J., De Boeck, P., 2016. Hidden Markov item response theory models for responses and response times. *Multivariate Behav. Res.* 51 (5), 606–626.

- Muraki, E., 1992. A generalized partial credit model: application of an EM algorithm. *Appl. Psychol. Meas.* 16, 159–176.
- Muthén, B., Muthén, L., 2017. *Mplus User's Guide*, eighth ed. Muthén & Muthén, Los Angeles, CA.
- National Center for Education Statistics, 2018. NAEP Technical Documentation on the Web. Retrieved from: <https://nces.ed.gov/nationsreportcard/tdw>.
- Oranje, A., Kolstad, A., 2019. Research on psychometric modeling, analysis, and reporting of the National Assessment of Educational Progress. *J. Educ. Behav. Stat.* 44 (6), 648–670.
- Oranje, A., Mazzeo, J., Xu, X., Kulick, E., 2014. A multistage testing approach to group-score assessments. In: Yan, D., von Davier, A.A., Lewis, C. (Eds.), *Computerized Multistage Testing: Theory and Applications*. Chapman and Hall/CRC, Boca Raton, FL, pp. 371–390.
- Organization for Economic Co-operation and Development, 2016. Technical Report of the Survey of Adult Skills (PIAAC), second ed. Organization for Economic Co-operation and Development, Paris, France. Retrieved from: https://www.oecd.org/skills/piaac/PIAAC_Technical_Report_2nd_Edition_Full_Report.pdf.
- Organization for Economic Co-operation and Development, 2018. PISA 2018 Technical Report. Organization for Economic Co-operation and Development, Paris, France. Retrieved from: <https://www.oecd.org/pisa/data/pisa2018technicalreport>.
- Perie, M., Moran, R., Lutkus, A.D., 2005. NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics. NCES 2005-464. Government Printing Office, Washington, DC.
- Rässler, S., Rubin, D.B., Zell, E.R., 2013. Imputation. *Wiley Interdiscip. Rev.* 5 (1), 20–29.
- Rhee, S.H., Friedman, N.P., Boeldt, D.L., Corley, R.P., Hewitt, J.K., Knafo, A., Young, S.E., 2013. Early concern and disregard for others as predictors of antisocial behavior. *J. Child Psychol. Psychiatr.* 54, 157–166.
- Rubin, D.B., 1976. Inference and missing data. *Biometrika* 63 (3), 581–592.
- Rubin, D.B., 1987. *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York, NY.
- Rubin, D.B., 1996. Multiple imputation after 18+ years. *J. Am. Stat. Assoc.* 91 (434), 473–489.
- Rust, K.F., Johnson, E.G., 1992. Sampling and weighting in the national assessment. *J. Educ. Stat.* 17, 111–129.
- Rust, K., 2013. Sampling, weighting, and variance estimation in international large-scale assessments. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. Chapman and Hall/CRC, Boca Raton, FL, pp. 117–153.
- Rutkowski, L., Zhou, Y., 2015. The impact of missing and error-prone auxiliary information on sparse-matrix sub-population parameter estimates. *Methodology* 11, 89–99.
- Rutkowski, L., 2011. The impact of missing background data on subpopulation estimation. *J. Educ. Meas.* 48 (3), 293–312.
- Schafer, J.L., Graham, J.W., 2002. Missing data: our view of the state of the art. *Psychol. Methods* 7 (2), 147–177.
- Schofield, L.S., Junker, B., Taylor, L.J., Black, D.A., 2015. Predictive inference using latent variables with covariates. *Psychometrika* 80 (3), 727–747.
- Shin, H.J., Yamamoto, K., Khorramdel, L., Robin, F., 2021. Increasing measurement precision of PISA through multistage adaptive testing. In: Wiberg, M., Molenaar, D., González, J., Böckenholt, U., Kim, J.S. (Eds.), *Quantitative Psychology, Springer Proceedings in Mathematics & Statistics*, vol. 353. Springer, Cham, Switzerland, pp. 325–334.
- Shin, H.J., Jewsbury, P.A., Van Rijn, P.W., Robin, F., Lee, S.S., (in preparation). Latent Regression Item Response Models for Missing Data in Background Questionnaires in Educational Large-Scale Assessments.
- Shin, H.J., von Davier, M., Yamamoto, K., (in press). Incorporating timing data in to the PISA population modeling. In: Khorramdel, L., von Davier, M., Kentaro, Y. (Eds.), *Innovative Computer-Based International Large-Scale Assessments Foundations, Methodologies and Quality Assurance Procedures*. New York, NY: Springer.
- Shin, H.J., Jewsbury, P.A., van Rijn, P.W., (under review). Conditional Dependencies Between Cognitive Item Responses and Response Times in Large-Scale Educational Assessments.
- Sinharay, S., Stern, H.S., Russell, D., 2001. The use of multiple imputation for the analysis of missing data. *Psychol. Methods* 6 (4), 317–329.
- Thomas, N., Gan, N., 1997. Generating multiple imputations for matrix sampling data analyzed with item response models. *J. Educ. Behav. Stat.* 22 (4), 425–445.
- Thomas, N., 1993. Asymptotic corrections for multivariate posterior moments with factored likelihood functions. *J. Comput. Graph Stat.* 2 (3), 309–322.
- Ulitzsch, E., von Davier, M., Pohl, S., 2020a. A hierarchical latent response model for inferences about examinee engagement in terms of guessing and item-level non-response. *Br. J. Math. Stat. Psychol.* 73 (S1), 83–112.
- Ulitzsch, E., von Davier, M., Pohl, S., 2020b. Using response times for joint modeling of response and omission behavior. *Multivariate Behav. Res.* 55 (3), 425–453.
- van der Linden, W.J., 2007. A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika* 72, 287–308.
- van der Linden, W., 2016. *Handbook of Item Response Theory 1*. Chapman and Hall/CRC, Boca Raton, FL.
- van Rijn, P.W., Ali, U.S., 2018. A generalized speed–accuracy response model for dichotomous items. *Psychometrika* 83 (1), 109–131.
- von Davier, M., Sinharay, S., 2010. Stochastic approximation methods for latent regression item response models. *J. Educ. Behav. Stat.* 35 (2), 174–193.
- von Davier, M., Sinharay, S., Oranje, A., Beaton, A., 2006. The statistical procedures used in National Assessment of Educational Progress: recent developments and future directions. In: Rao, C.R., Sinharay, S. (Eds.), *Handbook of Statistics, Psychometrics*, vol. 26. Elsevier, Amsterdam, Netherlands, pp. 1039–1055.
- von Davier, M., Gonzalez, E., Mislevy, R., 2009. What are plausible values and why are they useful? In: von Davier, M., Hastedt, D. (Eds.), *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments*, vol. 2, pp. 9–36.
- von Davier, M., Khorramdel, L., He, Q., Shin, H.J., Chen, H., 2019. Developments in psychometric population models for technology-based large-scale assessments: an overview of challenges and opportunities. *J. Educ. Behav. Stat.* 44 (6), 671–705.
- von Davier, M., Khorramdel, L., Shin, H.J., (in press). Scoring and scaling items from innovative domains. In: Khorramdel, L., von Davier, M., Yamamoto, K. (Eds.), *Innovative Computer-Based International Large-Scale Assessments—Foundations, Methodologies and Quality Assurance Procedures*. New York, NY: Springer.
- von Davier, M., 2013. Imputing proficiency data under planned missingness in population models. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. Chapman and Hall/CRC, Boca Raton, FL, pp. 175–201.
- Wainer, H., Dorans, N.J., Flaugher, R., Mislevy, R.J., 1990. *Computerized adaptive testing: A primer*. Lawrence Erlbaum, Mahwah, NJ.
- Warm, T.A., 1989. Weighted likelihood estimation of ability in item response theory. *Psychometrika* 54 (3), 427–450.
- Weirich, S., Haag, N., Hecht, M., Böhme, K., Siegle, T., Lüdtkke, O., 2014. Nested multiple imputation in large-scale assessments. *Large Scale Assess. Educ.* 2 (1), 1–18.
- Wu, M., 2005. The role of plausible values in large-scale surveys. *Stud. Educ. Eval.* 31, 114–128.
- Yamamoto, K., Lennon, M.L., 2018. Understanding and detecting data fabrication in large-scale assessments. *Qual. Assur. Educ.* 26 (2), 196–212.
- Yamamoto, K., Mazzeo, J., 1992. Item response theory scale linking in NAEP. *J. Educ. Behav. Stat.* 17 (2), 155–173.
- Yamamoto, K., Shin, H.J., Khorramdel, L., 2018. Multistage Adaptive Testing Design in International Large-Scale Assessments. *Educational Measurement: Issues and Practice* 37 (4), 16–27.
- Zwick, R., 1992. Chapter 7: statistical and psychometric issues in the measurement of educational achievement trends: examples from the National Assessment of Educational Progress. *J. Educ. Stat.* 17, 205–218.

List of relevant websites

International Data Explorer, <https://nces.ed.gov/surveys/international/ide/>.
 NAEP Data Explorer, <https://www.nationsreportcard.gov/ndecore/landing>.
 PIAAC Data Explorer, <https://piaacdataexplorer.oecd.org/ide/idepiaac/>.
 PISA Data Explorer, <https://pisadataexplorer.oecd.org/ide/idepisa/>.

A brief introduction to contemporary standard setting methods

Sonya Powers, Angela D. Bowzer, Mariann Lemke, Matthew A. Brunetti, and Matthew N. Gaertner, WestEd, San Francisco, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	369
Pre-standard setting meeting activities	370
Establishing the purposes and goals of standard setting	370
Stakeholder engagement	371
Developing preliminary performance level descriptors	371
Selecting a standard setting method	372
Data analysis	372
Preparing materials	373
Typical standard setting meeting activities	373
Common standard setting methods	374
Angoff and related methods	374
Bookmark method	375
Item-descriptor matching method	376
Embedded standard setting	376
Contrasting groups method	377
Body of work method	377
Training	377
Range-finding	378
Pinpointing	378
Post-standard setting meeting activities	378
Vertical articulation	378
Cut score approval and final PLD revisions	379
Technical report and ongoing validation	379
Conclusion	379
Acknowledgment	379
References	379

Glossary

Criterion-referenced A comparison of performance to a fixed standard, such as the knowledge and skills an assessment is designed to measure, without regard to the performance of others

Cut score The point that separates one level of performance from another (e.g. “proficient” vs. “not proficient”) on a score scale

Feedback data Data provided to panelists during the standard setting meeting to help inform their cut score judgments

Impact data The percentage of examinees classified into each performance level, based on a set of cut scores

Norm-referenced A comparison of performance to the performance of others, such as comparing one examinee’s score to the scores of a reference group of examinees

Performance level A scale score range, bounded by the minimum scale score, the maximum scale score, and/or the cut scores, which reduces the granularity of scores to facilitate score interpretation

Performance level descriptors (PLDs) A qualitative description of the knowledge and skills represented by a given performance level. Another common name for these is “achievement level descriptors.”

Standard setting A process that typically uses stakeholder judgments and data to determine cut scores that signify different levels of performance on an assessment

Introduction

At one point or another, most people in modern society have had to take an assessment. Often the score received is only important in relation to a threshold one aims to surpass. For example, a score of 1250 is better than a score of 1200 on the SAT, but if the score needed to be admitted to one’s first-choice university is 1300, the difference between 1250 and 1200 is of little relevance because

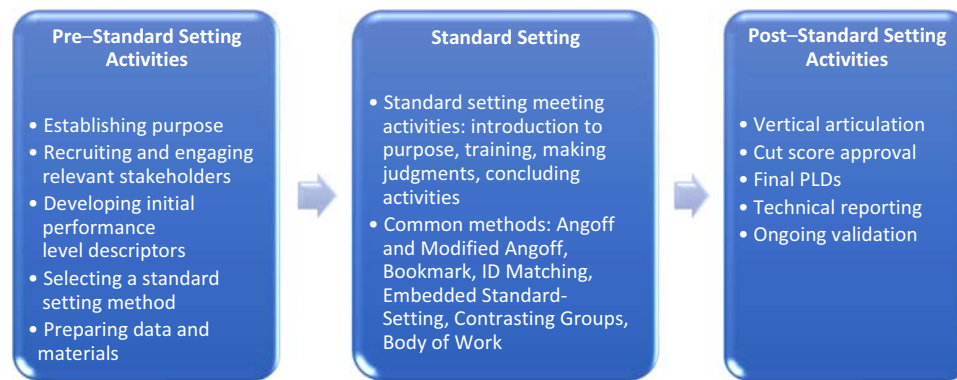


Fig. 1 The standard setting process.

neither score meets the threshold. Standard setting, the process of deriving those thresholds or cut scores, has been an area of concerted research focus in the psychometric community for decades, though it may remain opaque to the general public. New standard setting methods and refinements to existing methods continue to be developed in an ongoing effort to increase the reliability and validity of the resulting cut scores and performance levels.

This article provides a brief summary of the art and science of standard setting (Cizek, 2001). Fig. 1 provides an overview of the standard setting process and of the structure of this article. In most cases, standard setting is conducted through meetings with subject-matter experts. This article begins with a discussion of the activities that must be completed prior to a standard setting meeting. Next, the typical flow of activities during a standard setting meeting is described along with a few common methods used for standard setting. Finally, post-standard setting meeting activities are detailed, including standards articulation, stakeholder approval, and finalizing performance level descriptors (PLDs). For readers who are interested in learning more about particular standard setting methods or activities, Cizek and Bunch (2007) and Cizek (2012b) provide book-length treatments of these topics.

Pre-standard setting meeting activities

This section describes several of the most common activities that occur prior to a standard setting meeting, including:

- establishing the purpose and goals of standard setting,
- engaging stakeholders,
- developing preliminary performance level descriptors (PLDs),
- selecting a standard setting method,
- analyzing data, and
- preparing materials.

Establishing the purposes and goals of standard setting

Many decisions need to be made during the standard setting design process, including selecting a standard setting method and planning for a standard setting meeting. However, making these decisions without first understanding the reasons why performance levels are desired is difficult. For example, users of certification assessment scores may want a single cut score that differentiates passing performance from failing performance. This cut score is used to classify examinees as qualified or unqualified for a particular occupation. The cut score will have to be justified, based on job analysis information and predictive validity evidence, to provide evidence that examinees who fail are truly unqualified for a particular job and that examinees who pass are qualified for the job and will likely be successful at it. If a previous version of the assessment exists, there may also be an implicit assumption that the pass rate on the new assessment will be similar to the previous pass rate.

Setting standards on an assessment that is required for statewide summative accountability purposes, on the other hand, requires having at least two cut scores, separating students into three performance levels, per federal requirements (Every Student Succeeds Act, 20 U.S.C. § 6301, 2015), and states that have had four or five performance levels may want to have the same number of categories on a new or revised version of their assessment. These assessments are federally required to be “criterion-referenced”—tied to performance relative to the state’s academic content standards. Nevertheless, policymakers often have “norm-referenced” assumptions about the percentages of students expected to score within each category, and substantial deviations from the expected percentages may be politically untenable.

Those who are responsible for facilitating a standard setting meeting must understand the myriad of purposes for which assessment results, and especially performance levels, are likely to be used. Working with policymakers and test-score users is a critical first

step in understanding the purposes and goals of the standard setting endeavor and designing a standard setting process that will achieve those goals.

Stakeholder engagement

There are two very important groups of people to engage with prior to a standard setting meeting. The first group is the group of stakeholders who have ultimate responsibility for setting the cut scores. Understanding their priorities and what they view as acceptable, versus unacceptable, outcomes is important. Those in charge of facilitating the standard setting meetings should meet with decision-makers early and often and find ways to embed their priorities into the standard setting process so that everyone is working toward a common goal.

The second group of people to engage with is the group of panelists who will be involved in the standard setting process. There are three important considerations when determining the composition of the standard setting panel:

1. What are important stakeholder groups to include?
2. What knowledge or skills do individuals need in order to contribute to the standard setting process?
3. How many panelists should be included on the panel?

Although examinees are rarely included in a standard-setting panel, subject-matter experts and other individuals who will be impacted by the cut score placement should be included in the process, so that their perspectives are included, they buy into the process, and they endorse the results. Groups to be represented on standard setting panels will differ, depending on the type of assessment. For example, for a statewide summative assessment, teachers are an essential group to include on the standard setting panel because they should be familiar with the state standards to which the assessments align, the students who will take the assessments, and the ways in which scores are used. Other relevant panelists might include curriculum specialists, school board members, and community members.

Decision-makers will want recommended cut scores to be based on the informed judgments of subject-matter experts, and, ideally, they will want to be able to accept those recommendations. If decision-makers reject the recommendations, they risk alienating standard setting panelists and the groups they represent. Therefore, careful panelist selection and facilitation that keeps all stakeholders aligned toward the same goals are needed in order to ensure the success of the standard setting process.

There are no strict criteria for determining the number of panelists needed for a standard setting meeting. However, the number of panelists must be sufficient to ensure diverse stakeholder representation across geographic or other background characteristics, and inclusion of necessary expertise. Often, past precedent (e.g. 10 panelists were included during a previous standard setting, so approximately 10 panelists will be included again) and budget constraints also factor into panel size.

Standard setting meetings are often multi-day or even weeklong events; getting this degree of time commitment requires advance planning, and sufficient time must be allotted for panelist recruitment. Depending on how large and diverse the panelist group should be, in terms of both demographics and subject-matter expertise, recruiting enough willing panelists with desired characteristics may take several months. Once panelists are recruited, it is important to carefully coordinate communications so that they remain engaged. The most important communications are those that establish the legitimacy and importance of the process and provide all of the logistical details that panelists need, such as information about meeting location, travel arrangements, lodging arrangements, and per diems. Loomis (2012) provides thorough descriptions of pre-meeting recruitment and communications with panelists.

Developing preliminary performance level descriptors

Cut score recommendations are typically an output of a standard setting meeting, but the number of expected levels, and descriptions of what those levels should mean, must be established ahead of time. Policymakers, sometimes in conjunction with various stakeholder groups, determine the number of performance levels, based on the goals of classifying examinee performance and the number of performance levels into which examinees can be meaningfully differentiated. For some tests, two performance levels are sufficient. In many large-scale educational assessments, four or five performance levels are common.

Each performance level typically has an associated label, such as “proficient,” that provides a shorthand description of the level (Perie, 2008; Zieky and Perie, 2006). Policymakers should select labels that are clear, concise, and accurate; that avoid ambiguity; and that they are comfortable putting in front of the public. Table 1 lists performance level labels for a variety of assessments (international, national, and state).

Performance level descriptors (PLDs), also known as achievement level descriptors, communicate the knowledge and skills that examinees performing at each performance level possess and how this knowledge deepens at higher performance levels (Perie, 2008; Schneider and Egan, 2014). PLDs are often viewed as the heart of the standard setting process (Egan et al., 2009) because panelists must understand what performance levels mean in order to classify examinees (and/or test items) into each level.

Egan et al. (2012) define four types of PLDs, each with a different goal and purpose in the standard setting process:

1. Policy PLDs: General statements about the knowledge and skills demonstrated by examinees performing at a given level and that communicate policymakers’ vision for the performance levels.
2. Range PLDs: Detailed descriptions of the knowledge and skills embodied within a performance level.

Table 1 Performance level labels for international, national, and state assessments.

Assessment	Lowest level	←————→			Highest level
Trends in International Mathematics and Science Study (TIMSS)	Low	Intermediate	High	Advanced	
National Assessment of Educational Progress (NAEP)	Below basic	Basic	Proficient	Advanced	
Smarter Balanced Assessment Consortium (SBAC)	Standard not met	Standard nearly met	Standard met	Standard exceeded	
Florida Standards Assessments	Inadequate	Below satisfactory	Satisfactory	Proficient	Mastery
Georgia Milestones	Beginning learners	Developing learners	Proficient learners	Distinguished learners	
Ohio's State Tests	Limited	Basic	Proficient	Accelerated	Advanced
State of Texas Assessments of Academic Readiness (STAAR)	Did not meet grade level performance	Approaches grade level performance	Meets grade level performance	Masters grade level performance	
Virginia Standards of Learning (SOL)	Fail/below basic	Fail/basic	Pass/proficient	Pass/advanced	

3. Target PLDs: Descriptions of the minimum level of performance for a given performance level.
4. Reporting PLDs: Concise versions of the range PLDs, used to communicate the meaning of the performance levels to examinees and the public.

Not all types of PLDs are developed for all assessment programs. The goals of the assessment, the standard setting method, and how performance is reported determine the types of PLDs needed. For example, some standard setting methods require the identification of the knowledge and skills of an examinee performing at the minimum level or the bottom of each performance level; these examinees are sometimes referred to as “borderline” examinees (Downing, 2006; Zieky and Perie, 2006). For these methods, target PLDs are needed (Schneider and Egan, 2014); for other methods, target PLDs are not needed.

Preliminary PLDs developed to support the standard setting meeting are typically refined using the final set of cut scores (Perie, 2008). For K–12 assessments where content standards are grade-specific, PLDs will necessarily be grade-specific, and attention should be paid to the knowledge and skills that students need at a performance level at a given grade, to ensure that the knowledge and skills represent a progression across grade levels (Perie, 2008; Zieky and Perie, 2006).

Selecting a standard setting method

Many approaches to setting performance standards exist. Some are better suited than others for a given assessment, based on identified goals or item types used, but there is no one right method (Egan et al., 2009; Ferrara et al., 2014). However, the selection of the method should occur early in the standard setting planning process, so that necessary components of the method can be planned for and carried out at the appropriate times (Downing, 2006). The method identified for use should allow for defensible performance standards, aligned with the purpose and content of the test and with the number of performance levels. A description of several common standard setting methods is provided later in this article.

Data analysis

During the standard setting meeting, data are often presented to panelists to inform their judgments. These data are called “feedback data” and are carefully curated to help facilitate panelists’ judgmental process. Depending on the type of standard setting method employed, various graphical plots and/or descriptive statistics may be provided to panelists during the standard setting meeting, to help them understand how their judgments would be translated into cut scores and to illustrate the consistency of judgments within the panel. These feedback data are prepared during the standard setting meeting, using panelist judgments as input. Analytic code should be developed and validated prior to the standard setting meeting, to automate analysis of panelist judgments so that results can be turned around quickly. This code should be well tested, using a variety of hypothetical human judgments, different numbers of panelists, different agreement levels among panelists, and other potential scenarios, to verify that it will run smoothly during the meeting. There should also be quality control checks built into the process. For example, a lower performance level requires a lower cut score than a higher performance level. If a panelist’s set of judgments are not internally consistent, their judgments should be flagged immediately, so that the facilitator can follow up with the panelist and receive corrected judgments. If a panelist does not provide one or more needed values, this should also be flagged for follow-up.

Some feedback data requires analysis of examinee scores prior to the standard setting meeting. Two common examples are item difficulty, quantified as the percent of examinees who answered each item correctly (i.e. item p -values), and impact data, which is the percentage of examinees classified into each performance level based on a given set of cut scores. For some methods, item maps must be generated, in such a way that each item is located on the score scale. The standard setting method, the availability of

data, and the standard setting philosophy play important roles in determining what types of feedback data to include in the standard setting meeting. For example, a philosophical approach that views standard setting as a primarily content-oriented endeavor implies measuring what examinees know and can do only in reference to the relevant standards. Therefore, providing impact data may be viewed as an incongruent norm-referenced method. Additionally, normative data on examinee performance may not be available at the time of a standard setting meeting, which will eliminate the ability to provide impact data and item difficulty information as feedback data to panelists.

It is increasingly common for standard setting to involve several studies prior to the standard setting meeting, to provide additional context for making judgments. These studies require careful planning and much longer timelines. Methods such as the Briefing Book method (Haertel et al., 2012) and the Benchmark method (Phillips, 2012) elevate the importance of external data in the standard setting process. Although study reports should be written and provided to interested audiences, and studies are often described in detail for the Briefing Book method, study results must often be summarized in concise visuals, so that panelists can see, at a glance, how their recommendations compare to study information. Elements of these methods are often blended with the Bookmark method (Lewis et al., 2012), because it is especially easy to illustrate particular benchmarks and external sources of data by associating benchmarks with specific items in an ordered item booklet (OIB).

Examples of potential analyses that could be used to inform cut scores include:

- comparisons to performance on a previous assessment system;
- comparisons to performance on other relevant assessments or measures (e.g. course grades);
- grade-to-grade comparisons using vertical scale scores and/or impact data; and
- contrasting groups (e.g. have college students take an end-of-course high school assessment and identify the score that best differentiates students who receive a B or better from students who do not).

Even with years to prepare, most testing programs will approach standard setting with incomplete or inadequate data. Item statistics from stand-alone field tests where examinees are not motivated to do their best, legacy items that only partially align with new PLDs, and idiosyncratic field-test populations are common issues that can sow confusion in standard setting panel meetings, and bias in the cut scores, if left unaddressed. Gaertner and Powers (2021) view these issues as missing data problems and offer suggestions, based on results from simulation studies, on potential approaches to address them.

Preparing materials

Some materials needed for a standard setting meeting are specific to the standard setting method selected. For example, an OIB is used for some standard setting methods (e.g. the Bookmark method), but not others. Likewise, examinee work samples are needed for some standard setting methods (e.g. the Body of Work method), but not others. Therefore, standard setting materials should be prepared after a standard setting method has been selected. Materials that are typically prepared for a standard setting meeting include meeting agendas, panelist and facilitator trainings, assessment items and/or OIBs, examinee performance data, preliminary PLDs, panelist judgment forms, and panelist evaluation surveys.

Typical standard setting meeting activities

In most cases, standard setting is conducted through meetings with subject-matter experts. Standard setting meetings for most methods include four key phases, each of which is discussed in this section:

- introduction to the purpose of standard setting,
- training on the standard setting method,
- content experts making judgments, and
- concluding activities (e.g. collection of secure materials, and evaluation).

At the start of the standard setting meeting, panelists should be oriented to the purposes and goals of the assessment and of the meeting. Providing panelists with appropriate background information and training is critical for ensuring that they understand the importance of their task, they are aware of the uses and consequences of the cut scores informed by their recommendations, and that they buy into the procedures and results of the meeting (Schafer, 2005).

The introduction often includes panelists “experiencing the assessment.” During this activity, panelists review each item, or a subset of items, from the assessment and respond to it as if they were an examinee. For constructed-response items, panelists are not required to draft a complete response, but are encouraged to think about the knowledge and strategies that examinees would use to engage with the item. After responding to the items, panelists are given an answer key so that they can self-score the assessment. This activity is meant to give panelists a better feel for the examinee experience and the relative difficulty of the items contained on a test form. For some standard setting methods, the assessment items are provided in an OIB, in which items are ordered from least difficult to most difficult. Although the OIB does not reflect the examinee experience, it is a helpful tool for gaining further insight into the relative difficulty of test items.

Another important part of the introductory training is to familiarize panelists with preliminary PLDs. Understanding the intended meaning of the performance levels is critical for making informed judgments about the level that best describes particular items or examinees.

Following the introduction to the purpose of the standard setting meeting, panelists should be trained on the method being used (Raymond and Reid, 2001). This training typically involves practicing the judgment activity at the center of the standard setting method, with opportunities for questions and clarifications. During the practice activity, explicit connections should be made between the PLDs and the judgments to be made (Ferrara and Lewis, 2012). The judgments that panelists make determine the numerical boundaries that separate performance levels. The mechanics for how the judgments are translated into cut scores should be explained and illustrated. Understanding of the process should be confirmed by all panelists, typically using a survey, prior to moving forward to the first round of judgments.

Once panelists are trained to use the standard setting method, they typically apply the method in 2–4 rounds of judgments, with discussion and feedback data provided between each round. Feedback data may include information about the consistency of panelist judgments, item performance, and/or examinee performance. Judgments may be inconsistent at the beginning of the process, resulting in a wide range of potential cut scores. Across rounds, panelists may revise their judgments based on discussion and feedback data. These revisions tend to improve the consistency of judgments across panelists. Exact consensus across panelists is almost never required, but a fairly narrow range of cut score recommendations is desirable because it suggests that panelists have developed a common understanding of the standard setting method and the meaning of the performance levels. It also means that the final recommended cut scores will not be far from any panelist's judgment, which promotes panelist acceptance of the final results.

Following the last round of judgments, panelists should be given a survey to evaluate the standard setting activities, including the degree to which they agree with where the performance standards were set (Cizek and Bunch, 2007). Results from this survey, in conjunction with results from the survey administered after the initial training and any other surveys administered over the course of the meeting, can be analyzed to provide data in support of the validity of the standard setting process (Hambleton and Pitoniak, 2006).

For many assessments, test security is an important consideration. In such cases, panelists should be briefed on test security requirements during the standard setting training. Panelists should be informed about their roles in ensuring test security and the consequences of violating security protocols. At the end of the meeting, all secure materials should be collected from panelists, ensuring that all items from all panelists are accounted for. Panelists should wait to be dismissed from the standard setting meeting by meeting facilitators after all materials are collected.

Common standard setting methods

The following sections provide brief descriptions of six common standard setting methods: the Angoff method and related methods, the Bookmark method, the ID Matching method, Embedded Standard Setting, the Contrasting Groups method, and the Body of Work method. Very detailed descriptions of these methods, and of several others, can be found in Cizek and Bunch (2007). Embedded Standard Setting is a relatively new method described in Lewis and Cook (2020).

Angoff and related methods

The Angoff method is a criterion-referenced process in which trained panelists make judgments about the number of points an examinee would need to earn on a test in order to be placed into a given performance level. For each performance level, each panelist independently reviews each item on the test and estimates the probability that a borderline or, equivalently, a minimally competent examinee would respond correctly to each item.

There are several variations on the Angoff method (Angoff, 1971), such as the modified yes/no procedure (Impara and Plake, 1997) and the extended Angoff procedure (Hambleton and Plake, 1995). The original method described by Angoff (1971) requires panelists to determine whether a borderline examinee would correctly answer each item, as indicated by a yes/no response. The cut score would be determined by converting "yes" responses to a value of 1 and "no" responses to a value of 0, summing the scores for each panelist, and averaging the total scores across the group of panelists (Impara and Plake, 1997). Table 2 provides an example of the yes/no version of the Angoff method.

Angoff (1971) provided a variation on the yes/no procedure that requires panelists to estimate the proportion of borderline examinees who would answer the item correctly. Despite being a footnote in the Angoff (1971) paper, this variation on the yes/no method is traditionally referred to as the Angoff method. The cut score is determined by summing the probabilities for the individual items. A hypothetical example of the Angoff method is provided in Table 3. This example shows five panelists' judgments for the probabilities of a borderline examinee responding correctly to each item on a 10-item test. The sum of the probabilities across all the items becomes each panelist's suggested cut score for the test.

The extended Angoff procedure applies to tests that include polytomous items (i.e. items worth more than one point). In this case, panelists answer the following question: "How many score points would the borderline examinee earn on this item?"

Once each panelist has provided a cut score, a single cut score based on the group of panelists must be determined. A common approach is to use the median of the panelists' cut scores; however, the mean (the average of all panelists' cut scores) or the trimmed

Table 2 Yes/no Angoff method example.

<i>Item</i>	<i>Panelist 1</i>	<i>Panelist 2</i>	<i>Panelist 3</i>	<i>Panelist 4</i>	<i>Panelist 5</i>
1	No (0)	Yes (1)	No (0)	No (0)	No (0)
2	No (0)	No (0)	No (0)	No (0)	No (0)
3	Yes (1)	Yes (1)	No (0)	Yes (1)	Yes (1)
4	No (0)	Yes (1)	No (0)	Yes (1)	Yes (1)
5	No (0)	Yes (1)	No (0)	No (0)	Yes (1)
6	No (0)	Yes (1)	No (0)	No (0)	No (0)
7	Yes (1)	Yes (1)	Yes (1)	Yes (1)	Yes (1)
8	Yes (1)	Yes (1)	Yes (1)	Yes (1)	Yes (1)
9	No (0)	No (0)	No (0)	No (0)	No (0)
10	Yes (1)	Yes (1)	Yes (1)	Yes (1)	Yes (1)
Sum/cut score	4	8	3	5	6

Table 3 Angoff method example.

<i>Item</i>	<i>Panelist 1</i>	<i>Panelist 2</i>	<i>Panelist 3</i>	<i>Panelist 4</i>	<i>Panelist 5</i>
1	0.35	0.60	0.25	0.45	0.40
2	0.25	0.45	0.30	0.35	0.30
3	0.65	0.80	0.50	0.80	0.85
4	0.50	0.70	0.50	0.65	0.55
5	0.40	0.65	0.40	0.45	0.55
6	0.35	0.55	0.25	0.40	0.50
7	0.65	0.70	0.60	0.65	0.60
8	0.80	0.90	0.75	0.85	0.85
9	0.30	0.50	0.30	0.35	0.40
10	0.75	0.90	0.65	0.80	0.80
Sum/cut score	5.00	6.75	4.50	5.75	5.80

mean (the average of the panelists' cut scores, excluding highest and lowest cut scores) may also be used. The benefit of using the median or the trimmed mean is that any extreme outlier scores will not impact the final cut score. For the example in Table 2, the final cut score could be 5.56 (mean), 5.75 (median), or 5.52 (trimmed mean).

Bookmark method

The Bookmark method (Lewis et al., 1996) is a common alternative to the Angoff method. The defining feature of the Bookmark method is the OIB. Panelists select the point in the OIB at which they judge the items to be too hard for a borderline examinee more often than not. To quantify "more often than not," various response probability values have been used, including, most commonly, 0.5 and 0.67 (Lewis et al., 2012; Huynh, 2006). Panelists have generally reported understanding the Bookmark procedure (Lewis et al., 2012). The Bookmark method also allows for relevant feedback data to be easily included within the OIB. Such feedback data might include external benchmark information, so that panelists understand how their bookmark judgments, made in reference to target PLDs and conceptualizations of borderline examinees, relate to scores and to impact data on other relevant assessments.

For assessments with more than two performance levels, panelists must repeat the process of conceptualizing a borderline examinee and flagging an item in the OIB. For example, an assessment with four performance levels requires three cut scores, three definitions of a borderline examinee, and three bookmarks. For the Angoff method, panelists would have to assign three different probabilities to each item. The Bookmark method greatly simplifies this process by eliminating many of the item-level judgments needed. For example, after a panelist places a bookmark differentiating Level 1 from Level 2, there is no need for the panelist to consider harder items for the first cut score. Likewise, when determining where to set the bookmark that differentiates Level 2 from Level 3, there is no need to consider items that fall before the first bookmark in the OIB—these items have already been judged to be Level 1 items, so they are not relevant to determining the bottom of Level 3.

Although the streamlined bookmarking process saves substantial time for panelists, the success of the process depends heavily on a well-constructed OIB. Challenges with OIBs include difficulty imbalance and item-to-PLD mismatch. Difficulty imbalance occurs when there are gaps in item difficulty in the OIB and/or clusters of items with very similar difficulty within the OIB. Item-to-PLD mismatch occurs when content experts judge items to be ordered differently than their empirically derived order. Difficulty imbalance should be avoided when items are selected for the OIB. Item-to-PLD mismatch is more difficult to control for ahead of time, although having both stable item parameters and the right difficulty balance are essential. Even with well-developed OIBs, item ordering based on empirical data does not always align with panelists' notions of item difficulty.

Item-descriptor matching method

In contrast to the Angoff and Bookmark standard setting methods, the ID Matching method does not require panelists to envision a hypothetical borderline examinee and use this examinee to assist them in determining judgments. Instead, this method is focused solely on test content. At the heart of the ID Matching task is consideration of the PLD that most closely matches the knowledge and skills that examinees need in order to correctly respond to an item. For K–12 academic assessments, invited panelists are often educators, because they tend to be most familiar with academic content standards, knowledge which is critical for making informed ID Matching judgments (Cizek, 2006; Cizek and Bunch, 2007; Ferrara and Lewis, 2012; Ferrara et al., 2014). In addition, developing preliminary PLDs prior to the standard setting meeting is critical to the success of this method. In most implementations of ID Matching, an OIB is provided to panelists, so that they can consider the relative difficulty of items, as well as the specific item content, when determining the performance level to which an item best aligns (Ferrara and Lewis, 2012).

According to Ferrara and Lewis (2012, p. 263), panelists should consider two questions as they make their item judgments: “What does an examinee need to know and be able to do in order to respond successfully to this item or at this score level?” and “What makes this item more difficult than previous items?” A panelist considers these questions for each item and records a judgment as a performance level. These judgments are often recorded in a table like Table 4, which provides an ordering of the items being reviewed, along with metadata about the items being considered, such as a content standard alignment and item difficulty (Cizek and Bunch, 2007; Ferrara et al., 2014). The final column is initially blank; each panelist uses it to record judgments, which are then used to identify cut scores. In the example in Table 4, the judgments do not move strictly from low to high performance levels (in this case, Basic to Proficient). The ID Matching method allows for reversals in judgments in cases where a panelist thinks that a more difficult item aligns best to a lower performance level. In this example, once the performance level cut score is determined, items below the Proficient performance level cut score will be classified as Basic, while all items above that cut score will be classified as Proficient.

In many implementations of ID Matching, patterns in the panelists’ judgments are considered after the first round. Items on the border of two performance levels—referred to as the threshold region—become the focus of subsequent rounds of review. In the example in Table 4, the threshold region between Basic and Proficient includes items 5 through 10. The Proficient cut score is presumed to be a scale score somewhere between the scale scores of item 5 and item 10.

For the ID Matching method, impact data shared with panelists focuses on the locations at which a cut could be placed in the threshold region, and panelist discussions focus on the match of item content to PLDs within the threshold region. In some implementations of ID Matching, the goal by the end of the standard setting meeting is to determine the item in each threshold region that most closely matches the content of the next performance level. Cut scores are typically determined using logistic regression (Ferrara and Lewis, 2012; Ferrara et al., 2014). However, as described in the following section, a new method that improves on logistic regression has been proposed for locating the cut score.

Embedded standard setting

Embedded Standard Setting (ESS; Lewis and Cook, 2020) coordinates and extends existing assessment development practices such that standard setting becomes a logical extension of assessment design and item development, rather than as a post hoc activity. In the standard implementation of ESS, PLDs are developed prior to both standard setting and item writing, and items are written explicitly to elicit evidence for the PLDs. Judgments of performance level alignment, then, are incorporated into the item-writing process, and cut scores are established empirically, based on the agreement of item-writer judgments and items’ empirical

Table 4 Example ID matching judgment form.

<i>OIB item number</i>	<i>Content standard</i>	<i>Empirical item difficulty (p-value)</i>	<i>Item-descriptor judgment</i>
1	C.2a	0.92	Basic
2	C.1b	0.90	Basic
3	C.1b	0.89	Basic
4	M.2	0.88	Basic
5	C.2c	0.86	Proficient
6	M.1	0.84	Proficient
7	F.6	0.80	Basic
8	B.1a	0.79	Proficient
9	M.1	0.78	Proficient
10	B.2	0.77	Basic
11	F.5a	0.75	Proficient
12	F.3	0.72	Proficient
13	C.1a	0.71	Proficient
14	F.6	0.70	Proficient
15	M.1	0.70	Proficient

difficulties. A key advantage of ESS is that subjective evaluation of item alignment to PLDs by a second set of subject-matter experts is not necessary when there is alignment between item-writer judgments and item difficulty parameters.

Because the first published description of the ESS method came out in 2020, many testing programs implementing ESS at the time of writing this article are doing so “in flight”—that is, after at least some items have been written and field tested—rather than following the “full” ESS protocol (i.e. item-to-PLD alignment judgments are being made as the items are being written, prior to field testing). Furthermore, for some programs, a set of cut scores in a given grade is one part of a larger system of cut scores, where the system is articulated vertically across multiple grades and typically placed on a single underlying vertical scale. Therefore, articulation of cut scores across grade levels must be considered, ideally early in the process.

ESS is particularly helpful for assessment programs that have the time, resources, and discretion to iterate. Every cycle of ESS (writing items aligned to PLDs and then reconciling those alignment judgments with empirical data) will surface misaligned items—for example, items that are judged to align with the lowest placement level but that are among the most difficult in the bank. As long as the field test was robust and the empirical data show no obvious signs of bias, these discrepancies can be corrected through revisions to the PLDs. Within reason, misalignments are therefore useful because they provide an opportunity for the assessment developer to make data-driven adjustments to PLDs.

An additional innovation of ESS is the mechanism for translating ID Matching results into cut scores. Rather than logistic regression, a weighted penalty that minimizes the number of items that fall on the “wrong” side of the cut score, based on expert judgment, is calculated by identifying the cut scores. [Lewis and Cook \(2020\)](#) proposed a method called ESS-Count to optimize cut score placement, and [Lewis et al. \(2021\)](#) introduced an additional cut score placement method, which they called ESS-Weight. Simulations conducted by [Gaertner and Powers \(2021\)](#) indicated that ESS-Weight has better properties for setting cut scores, compared to ESS-Count, and produces superior results compared to logistic regression.

Contrasting groups method

For the Contrasting Groups method, instead of having content experts align items with performance levels, panelists must be familiar enough with examinee ability to classify examinees into performance levels. For example, when examinees are students, the panelists should be teachers who are familiar with students’ knowledge and skills in the subject of interest. The panelists review the PLDs, think about which description best describes a particular student, and provide a judgment for the student. This method requires a sample of students that spans the performance continuum, as well as a sufficient number of panelists who are familiar with the students to provide an informed judgment.

As with ID Matching, Contrasting Groups cut scores are often calculated using logistic regression ([Hambleton and Pitoniak, 2006](#)). However, the newer ESS-Weight method ([Lewis et al., 2021](#)), described in the previous section, can also be used to translate judgments into cut scores.

Body of work method

Many standard setting methods have been developed for use with multiple-choice items. Although some such methods can be modified for use with performance assessments or mixed-format assessments, these methods are often not well suited to such assessments. To address this need, the Body of Work method ([Kingston et al., 2001](#)) was designed especially for performance assessments. As the name suggests, Body of Work is a method in which examinees’ responses to all items on an assessment are provided to panelists. Panelists review all responses systematically to identify the set of scores that constitutes the minimum required knowledge and skill for a particular performance level. Judgments are not made on an item-by-item basis, but are made holistically, taking into consideration examinee performance on all items within the set. This method is often applied to writing assessments, portfolio-based assessments, video-recorded performances, and observation-based assessments.

Many of the steps in the Body of Work method are similar to steps in other standard setting methods (e.g. recruiting panelists, conducting a meeting with multiple rounds). One salient difference is the inclusion of sets of examinee work, which are then classified by panelists as part of the standard setting judgment process. These examinee work samples must be strategically selected prior to the standard setting activity. Prior to selecting examinee work samples, all items must be scored. These scores must be reliable in order to provide useful input to the standard setting method. Thus, examinee responses would ideally be double-scored by trained human raters. Once items for all examinees are scored, a subset of examinees is systematically selected, and these examinees’ responses to all items on the assessment are included in the standard setting meeting. These sets of examinee responses are used for three different purposes: training, range-finding, and pinpointing.

Training

During the training portion of the standard setting meeting, panelists are often asked to perform the standard setting task, using a small “training” set, to get them comfortable with completing the task and entering their judgments into a paper or electronic form. For this task, a small set of examinee assessment responses (“bodies of work”) is needed, such that panelists can independently evaluate the responses and determine which performance level is best aligned with the quality of the responses. This training set is typically not part of the range-finding or pinpointing sets, though it could be a very small subset.

Range-finding

The first round of judgments for the Body of Work method is called range-finding. During this round, panelists review examinee work that spans the full performance continuum. The range-finding activity involves at least one round of independent judgments, but often also includes time for group discussion and for panelists to modify judgments, based on that discussion. The outcome of the range-finding activity is the identification of a narrow region of the score scale associated with each cut score. This region includes examinee work for which there is substantial disagreement among panelists as to the performance level it represents. This type of disagreement is most common at the threshold of a performance level.

Pinpointing

The second round of judgments for the Body of Work method is called pinpointing, because this round focuses on a narrower region of the scale, to home in on the cut scores. Typically, new bodies of work, representing score points close to where the cut score is likely to be, are added in this round. The new examinee work samples are selected prior to the standard setting meeting, so that they are available at the meeting. Because the cut scores are not known ahead of time, facilitators must be prepared with many alternative pinpointing sets. During this round, panelists repeat the process of classifying bodies of work into performance levels. After independent judgments, there is often a group discussion, after which panelists are able to change their judgments if they wish.

Logistic regression is used to calculate the raw score at which point the probability is 0.5 of being in a particular performance level or higher (e.g. “proficient” or higher) versus being in the next lower level (e.g. the category below “proficient”). Using logistic regression, the panelist judgments are converted into a cut score (see [Cizek and Bunch \(2007\)](#) for an in-depth example). After the cut scores are computed, feedback data can be provided. In some applications of the Body of Work method, a third round of judgments is conducted, during which panelists are asked to provide the minimum score that examinees in each performance level should obtain, based on feedback data and final group discussions. This provides a quick and intuitive mechanism for panelists to adjust their final recommendations.

The Body of Work method can be very time consuming, depending on the number of achievement levels and the complexity of examinee performances (e.g. number of item types, number of items, range of rubrics). Ideally, a pilot standard setting study would be conducted before the meeting—even if only with internal staff—to gauge how long the range-finding and pinpointing rounds are expected to take. If necessary to save time, tasks can be strategically split among the standard setting panel. For example, different panelists could focus on different cut scores. Splitting tasks among the panel will require that the panel is of sufficient size so there is a balance between how many days the standard setting meeting will take and how many panelists are needed for the meeting. The Body of Work method is also dependent upon a large pool of examinee responses to select from. A wide range of examinee performance is needed in order to select a good sample for range-finding. Additionally, several examples of examinee work are needed at each score point along the raw score scale in order to be able to successfully pinpoint the cut scores. Furthermore, for panelists to be able to properly identify cut scores, the assessment must not have obvious floor effects, ceiling effects, and/or scoring artifacts that result in gaps in the score scale.

Post-standard setting meeting activities

Vertical articulation

When standards are set across multiple grades and/or subjects—as is often the case for statewide summative assessment programs—an additional check is needed in order to verify that the standards, as a whole, make sense. Two common ways for evaluating the reasonableness of standards are via vertical scales, when they exist, and via impact data. In general, the cut scores should follow a smoothly increasing trajectory across grade levels. The cut scores need not follow a linear trajectory, since growth curves are typically nonlinear, but they should not be erratic or decrease as grade level increases. Likewise, impact data can differ from one grade to the next, but it should not change dramatically without explanation. Although examinee cohorts are known to differ in performance, dramatic differences are not expected when cohorts are large. Similarly, large differences in percent proficient across subject areas are often rightfully viewed with suspicion by stakeholders. The process of smoothing cut scores to achieve coherence is called vertical articulation, or vertical moderation. Vertical articulation can be included as a culminating activity during the standard setting meeting, in which the panelists (or a subset of panelists from each panel) come together to consider the cut scores holistically. In this case, the panelists can recommend changes to result in more consistent patterns across grades, keeping in mind the impact on the level of knowledge and skills that a change would imply and whether such changes have implications for the PLDs. Alternatively, vertical articulation can be completed after the standard setting meeting with the policy group responsible for approving the cut scores.

Whether vertical articulation occurs within the standard setting meeting or not, there will need to be a review of the full set of standards, with consideration to their coherence as a system. This review will occur when the organization responsible for carrying out the standard setting meetings presents the full set of recommended cut scores to those individuals responsible for approving the standards. Although impact data are typically not the focus of a standards-based assessment program, these data are still an important consideration for those who approve the standards. There are often expectations around how many examinees will pass the test and around the consistency of pass rates across grade levels. These expectations may be based on pass rates from a prior assessment program or on messaging about how the new assessment program is supposed to differ from prior assessment programs in terms of rigor.

If other data, beyond impact data, are used as feedback to panelists, this information should be reviewed with the policy group that is responsible for approving the standards. In doing so, consideration of how to present information concisely, limiting the presented information to only those pieces that are most relevant for decision-makers, is important. When the policy group has been engaged throughout the process, and when its priorities have been embedded into the standard setting process, it is not uncommon for the recommended cut scores that come out of a standard setting meeting to be accepted, or accepted with minor modifications. When this group has not been consulted, or when its priorities have not been emphasized in the design and implementation of the panel meetings, the recommendations are often heavily modified before final cut scores are approved.

Cut score approval and final PLD revisions

Once the cut scores are approved, those leading the standard setting process should complete a final review of the PLDs, especially if the final approved cut scores differed from the recommended cut scores. The final PLDs should reflect what examinees in a performance level actually know and can do, as opposed to what they should know and be able to do. These final PLDs are used to develop reporting PLDs to help test-score users understand the meaning of the performance levels. Therefore, reporting PLDs are empirical in nature, and they are often validated using item mapping methods where IRT is used to determine each item's location on the underlying scale. Item mapping clarifies which skills are represented within each performance level.

Technical report and ongoing validation

After the cut scores are approved, the standard setting process, as well as the final approved cut scores and PLDs, should be documented in a technical report. The technical report should summarize the standard setting method, describe the panelists, provide information about judgments and feedback data from each round, and provide results from panelist surveys (see [Cizek \(2012a\)](#) for examples of panelist surveys). Evidence of internal consistency, external consistency, and procedural validity should be provided to support the standard setting results ([Kane, 2001](#)). Finalized sets of reporting PLDs and cut scores are provided to those responsible for reporting results and included in the technical report. The percentage of examinees classified into each performance level, based on the cut scores, should be monitored over time, and additional external validity evidence should be collected periodically.

Conclusion

Although policymakers, educators, and the public may not be familiar with the details of how standard setting is carried out, the results of standard setting matter for individuals and for institutions. The percentage of examinees scoring at or above a "proficient" level, for instance, is an important part of most state accountability policies for K–12 schools. Similarly, for students, scoring below "proficient" may mean summer school, and for examinees looking to begin a career, it may mean failure to receive a license. Ensuring that assessment cut scores are determined in an accurate, transparent, and fair manner is critical to high-stakes decision-making. Ongoing research extending the "art and science" ([Cizek, 2001](#)) of standard setting is leading to improvements in the process, a better understanding of the strengths and shortcomings of various methods, and, ultimately, assessment information that improves the validity of score interpretations for examinees and other test-score users.

Acknowledgment

Our coauthor, colleague, and dear friend Matthew Gaertner inspired us and many others with his brilliance, wit, and passion. Our time with Matt was cut far too short, but we are honored to have known him, proud to have been part of his team at WestEd, and privileged to carry on his legacy and vision through this book article.

References

- Angoff, W.H., 1971. Scales, norms, and equivalent scores. In: Thorndike, R.L. (Ed.), *Educational Measurement*, second ed. American Council on Education.
- Cizek, G.J., Bunch, M.B., 2007. *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Sage.
- Cizek, G.J., 2001. Conjectures on the rise and call of standard setting: an introduction to context and practice. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Erlbaum, Mahwah, NJ, pp. 3–17.
- Cizek, G.J., 2006. Standard setting. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum, pp. 225–258.
- Cizek, G.J., 2012a. An introduction to contemporary standard setting: concepts, characteristics, and contexts. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 3–14.
- Cizek, G.J., 2012b. *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge.
- Downing, S.M., 2006. Twelve steps for effective test development. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum, pp. 3–25.
- Egan, K.L., Ferrara, S., Schneider, M.C., Barton, K.E., 2009. Writing performance level descriptors and setting performance standards for assessments of modified achievement standards: the role of innovation and importance of following conventional practice. *Peabody J. Educ.* 84 (4), 552–577.
- Egan, K.L., Schneider, M.C., Ferrara, S., 2012. Performance level descriptors: history, practice, and a proposed framework. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 79–106.

- Every Student Succeeds Act, 20 U.S.C. § 6301, (2015). <https://www.congress.gov/bill/114th-congress/senate-bill/1177>.
- Ferrara, S., Lewis, D., 2012. The Item-Descriptor (ID) matching method. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 255–282.
- Ferrara, S., Perie, M., Johnson, E., 2014. Matching the judgmental task with standard setting panelist expertise: the Item-Descriptor (ID) matching method. *J. Appl. Test. Technol.* 9 (1), 1–20.
- Gaertner, M.N., Powers, S., 2021. How to Win at Standard Setting With Imperfect Data. Paper Presented at the Annual Meeting of the National Council on Measurement in Education.
- Haertel, E.H., Biemers, J., Miles, J., 2012. The briefing book method. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 3–14.
- Hambleton, R.K., Pitoniak, M.J., 2006. Setting performance standards. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. Praeger Publishers, pp. 433–470.
- Hambleton, R.K., Plake, B.S., 1995. Using an extended Angoff procedure to set standards on complex performance assessments. *Appl. Meas. Educ.* 8, 41–55.
- Huynh, H., 2006. A clarification on the response probability criterion RP67 for standard setting based on bookmark and item mapping. *Educ. Meas.* 25 (2), 19–20.
- Impara, J.C., Plake, B.S., 1997. Standard setting: an alternative approach. *J. Educ. Meas.* 34 (4), 353–366.
- Kane, M., 2001. So much remains the same: conception and status of validation in setting standards. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Erlbaum, Mahwah, NJ, pp. 53–88.
- Kingston, N.M., Kahl, S.R., Sweeney, K.P., Bay, L., 2001. Setting performance standards using the body of work method. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Erlbaum, Mahwah, NJ, pp. 53–88.
- Lewis, D., Cook, R., 2020. Embedded Standard Setting: aligning standard-setting methodology with contemporary assessment design principles. *Educ. Meas.* 1, 8–21.
- Lewis, D.M., Mitzel, H.C., Green, D.R., 1996. *Standard Setting: A Bookmark Approach*. Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.
- Lewis, D.M., Mitzel, H.C., Mercado, R.L., Schulz, M., 2012. The Bookmark standard setting procedure. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 225–254.
- Lewis, D., Lee, S., Choi, H.-J., 2021. A Comparison of Two ESS Cut Score Estimation Algorithms. Paper Presented at the Annual National Council on Measurement in Education.
- Loomis, S.C., 2012. Selecting and training standard setting participants: state of the art policies and procedures. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 225–254.
- Perie, M., 2008. A guide to understanding and developing performance level descriptors. *Educ. Meas.* 27, 15–29.
- Phillips, S.E., 2012. Legal issues for standard setting in K–12 educational contexts. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Foundations, Methods, and Innovations*, second ed. Routledge, pp. 225–254.
- Raymond, M.R., Reid, J.B., 2001. Who made thee a judge? Selecting and training participants for standard setting. In: Cizek, G.J. (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Lawrence Erlbaum, pp. 119–158.
- Schafer, W.D., 2005. Criteria for standard setting from the sponsor's perspective. *Appl. Meas. Educ.* 18 (1), 61–81.
- Schneider, M.C., Egan, K.L., 2014. *A Handbook for Creating Range and Target Performance Level Descriptors*. National Center for the Improvement of Educational Assessment.
- Zieky, M.J., Perie, M., 2006. *A Primer on Setting Cut Scores on Tests of Educational Achievement*. Educational Testing Service.

Achievement growth modeling

Scott Monroe and Mina Lee, University of Massachusetts, Amherst, MA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	381
Historical and policy background	382
Achievement growth models	382
The gain score model	383
Model overview	383
Modeling requirements	383
Determining growth adequacy	383
Advantages and disadvantages	384
The categorical model	384
Model overview	384
Modeling requirements	385
Determining growth adequacy	385
Advantages and disadvantages	385
The student growth percentile model	385
Model overview	385
Modeling requirements	386
Determining growth adequacy	386
Advantages and disadvantages	386
The value-added model	386
Model overview	386
Modeling requirements	387
Determining growth adequacy	387
Advantages and disadvantages	387
Statistical properties of achievement growth measures	387
References	388

Introduction

Achievement growth modeling is the process of summarizing student achievement for students or groups of students, typically corresponding to teachers, schools, or districts, across multiple timepoints. Achievement growth is often contrasted with achievement status, that is, the achievement of a student or group at a single timepoint. Achievement growth is also distinct from within-grade change over time, which focuses on a specified grade-level or age, and uses multiple cohorts to study change in achievement across multiple timepoints. This entry focuses on achievement growth modeling in state accountability systems in the United States, which have been widely adopted to comply with various state and federal policies.

There are numerous approaches to achievement growth modeling, which can differ in important respects. For example, models can differ in whether growth is characterized in absolute or relative terms, and whether the target of inference is the individual or group. Growth models can also differ in other ways, including the types and features of data that are required, underlying technical properties and statistical assumptions, complexity of the modeling approach, and interpretability of the results. Finally, most growth models can be used to characterize growth in multiple ways, for example, by describing observed growth, or by determining the growth needed to meet a future target level of performance.

At the same time, states vary in their practical and technical considerations with respect to achievement growth modeling. Practical considerations include the intended use of the growth model (e.g., reporting individual student growth, supporting school accountability), the ease with which the model and results can be communicated to stakeholders, and whether the growth model satisfies any policy requirements. Technical considerations include the capacity of the state's longitudinal data system (e.g., what data are collected and stored each year), the degree to which these data support technical properties of growth models, and the testing program's analytical capabilities.

Each state is motivated to select a modeling approach that yields results that are reliable and valid for their intended uses. There are several practical and statistical considerations that factor into selecting a growth model (D'Brot, 2017). However, because each state has unique practical and technical considerations, no single growth model can be considered a gold-standard for all settings. Instead, each model will have a mix of more and less desirable features depending on the context. Consequently, some states use different models for different purposes, for instance, to report individual student growth, and to support school accountability. For these reasons, there is great variety in achievement growth modeling across the states, especially in the details of the methodologies.

The remainder of this entry is organized as follows. First, a brief historical and policy background of growth modeling is provided. Next, several models and their major features are presented. Finally, some relevant statistical properties of achievement growth measures, notably measurement error, are discussed.

Historical and policy background

In the field of educational measurement, achievement growth modeling has been studied since the 1950s, focusing on measuring growth for individual students. Through the 1970s, much of the research concerned the simple gain score (a current score minus a prior score) or residual gain score (a current score minus a predicted current score), measures designed to quantify individual growth or learning (e.g., Cronbach and Furby, 1970; Lord, 1956; McNemar, 1958; Rogosa et al., 1982; Traub, 1967). Measurement properties of these scores, in particular the reliability of gain scores, received a great deal of attention.

In the 1980s, more sophisticated approaches to modeling growth, such as the use of hierarchical linear modeling, began to appear in the literature (e.g., Bryk and Raudenbush, 1987; Rogosa and Willett, 1985). Notably, some of these approaches integrated background variables to directly model correlates of change (i.e., the relationship between background variables and individual growth). Such approaches could be used to study questions such as “What kinds of persons grow (learn) fastest?” (Cronbach and Furby, 1970, p. 77). From a historical perspective, these approaches can also be understood as precursors to value added modeling (Sanders and Horn, 1998; McCaffrey et al., 2004), which aims to isolate the contribution of individual teachers and schools to student growth.

Around this period, using test scores for accountability purposes became more common. Though testing for accountability has a long history in the United States (Shepard, 2008), the 1980s and 1990s saw state and federal policies, such as the Improving America’s School Act of 1994 and the No Child Left Behind Act (NCLB) of 2001, encouraging more comprehensive accountability systems. Initially, most accountability efforts focused on achievement status, in contrast to growth, using the current year’s student achievement to make judgments about the quality of education provided by teachers, schools, districts, and states (Shepard, 2008). For example, under NCLB, states were required to publicly report proficiency rates for all students, as well as select subgroups of students, within a school. States then used these rates to compare schools and districts, and to evaluate effectiveness.

Since the 2000s, researchers and policymakers have increasingly recognized that such status-based accountability systems may not capture the contributions of teachers and schools to students’ academic progress or growth (Linn, 2003). For example, such systems may fail to recognize effective teachers and schools that tend to serve students with low initial achievement levels, as well as ineffective teachers and schools that tend to serve students with high initial achievement levels. In contrast, growth measures are more closely tied to a student’s learning or individual progress, and are typically viewed as being more equitable than status measures (Linn, 2008). Thus, coupled with status measures, growth measures provide stakeholders with more complete pictures of student learning.

Recent federal and state policies and programs reflect this recognition of the importance of growth measures. Examples include the Growth Model Pilot Program of 2005, Race to the Top competitive grant program of 2009, and Elementary and Secondary Education Act (ESEA) flexibility waivers (2012), all of which incentivized states to adopt student growth measures. Also, the Every Student Succeeds Act of 2015, the current reauthorization of the ESEA, encourages state adoption of growth measures, mandating that state assessment systems must include multiple indicators of school quality and student success. Under ESSA, for elementary and secondary schools, states must include a student growth measure, or another valid and reliable indicator, that supports meaningful differentiation in student performance.

Achievement growth models

The purpose of this section is to provide an overview of common achievement growth models. We classify the models into one of four general approaches: gain score, categorical, student growth percentile, and value-added model. Though other classifications of the models are possible (Castellano and Ho, 2013; Goldschmidt et al., 2012; Martineau, 2016), we believe the proposed groupings, adapted from those proposed by Domaleski and Hall (2016), are conducive to covering the approaches used by most states.

The presentation for each model is organized around four topics. The first topic is a model overview. This includes the primary question the model is designed to answer, with respect to characterizing observed achievement growth. It also covers other fundamental features, such as whether the model characterizes growth in absolute or relative terms, and whether it yields individual or group-level growth estimates. When applicable, the overview also includes how the model may be used to determine the growth needed to meet a future target performance level, which we refer to as “growth to standard” (Domaleski and Hall, 2016). Also, where feasible, the overview includes a brief numerical example. Finally, examples of recent operational uses are provided.

The second topic is modeling requirements. In general, achievement growth modeling results must satisfy typical validity and reliability standards to justify their use (American Educational Research Association et al., 2014). Note that the results may be used for a variety of purposes (e.g., describing student growth, aiding in teacher or school evaluation), and the validity and reliability of each of these uses could be studied separately (e.g., Castellano and McCaffrey, 2020a,b; Hill, 2009). In general, however, the validity and reliability of growth measures relies, in part, on the extent that modeling requirements are met. To this end, we present

modeling requirements specific to the growth models, such as a vertical scale and cut scores articulated across grades. General issues of reliability will be discussed in a separate section, “Statistical Properties of Achievement Growth Measures.”

The third topic is determining growth adequacy. For reporting and accountability, stakeholders are often interested in determining whether growth is “adequate,” at both the individual and group levels. The possible ways in which growth standards may be set, however, depends on the specific growth model. To clarify these issues, for each model, we discuss how individual growth measures can be aggregated, and how growth adequacy determinations can be made.

The fourth and final topic is model advantages and disadvantages. Here, we present some of the more notable benefits and drawbacks for each model.

The gain score model

Model overview

Gain scores are arguably the most intuitive growth measure and characterize observed student growth in absolute terms. They are designed to address the question, “What is the magnitude of student growth?” A straightforward answer to this question is the gain score, defined as $G = X_2 - X_1$, where X_1 and X_2 are the student’s test scores at times 1 and 2, respectively (Lord and Novick, 1968). As an example, suppose a student has mathematics test scores of 400 in grade 3 and 450 in grade 4. Then, the corresponding gain score is simply $50 = 450 - 400$, indicating a 50-point increase or improvement. Gain scores have a long history in educational measurement (e.g., Lord, 1956), and have seen recent operational use in states such as Alabama (Alabama Department of Education, 2019) and Texas (Texas Education Agency, 2018).

Gain scores may also serve as the basis of a growth-to-standard measure. Two preliminary steps in calculating growth to standard are specifying a future target grade and a target performance level. Returning to our example student, assume these targets are grade 6 and the Proficient performance level, respectively. Further, assume the corresponding cut score is 490. Then, using the gain score model, the annual target growth needed by our student to progress from the initial grade 3 score of 400 to a grade 6 target score of 490 is $30 = (490 - 400)/3$. This growth-to-standard measure based on gain scores is sometimes referred to as the “trajectory model,” and has seen recent use in Connecticut (Connecticut State Department of Education, 2017) and Idaho (Idaho State Board of Education, 2018).

Modeling requirements

A requirement for gain scores to be meaningful is that the two test scores must be on a common interval scale. In a conventional pretest-posttest design, this requirement is satisfied if the same test or parallel forms of the same test are used at the two timepoints. When the two tests are different (as with the example mathematics tests in grades 3 and 4), this requirement can be satisfied through vertical scaling, a process used to place scores onto a common scale when tests measure similar domains, but at different grade levels (Kolen and Brennan, 2004). Vertical scaling is a complex endeavor that incorporates several psychometric tasks, including selecting measurement models and scaling procedures, linking, and standard setting (Dorans and Moses, this volume; Patz, 2007).

For the growth-to-standard approach of the trajectory model, the vertical scale requirement is more stringent, because the studied grade span is greater. That is, the simple gain score model assumes that properties of the vertical scale hold for 2 years, whereas the extended trajectory model assumes the properties hold for the entire studied grade span (e.g., grades 3 through 6 for the example student).

In addition, as used in practice, the trajectory model requires that cut scores across grades be articulated (Cizek, 2005). This means that performance levels across grades should have the same meaning and rigor, and that performance level classification rates for adjacent grades should be similar. These cut score characteristics facilitate cross-grade comparisons, and without them, predicting future proficiency is less sensible. For example, if cut scores substantially vary in stringency across years, then the percentage of students who are “on track” will be overly dependent on the choice of target year.

Determining growth adequacy

Determining adequate growth for students based on gain scores is usually done in one of two ways. First, standards can be set on the gain score scale, which corresponds to the underlying vertical scale. For example, a state may use the difference in performance level cut scores in consecutive grades to define adequate growth. If, for instance, the cut scores for the Basic performance level in grades 3 and 4 are 400 and 430, respectively, then under this approach, $G \geq 30$ may represent adequate growth. Second, standards can be norm-referenced, based on percentile ranks of gain scores in a reference group, such as all students in the same grade and subject in a given district. For example, a percentile rank of at least 50 may represent adequate growth.

Similarly, determining adequate growth for students using the growth-to-standard trajectory model is usually carried out in one of two ways. The first approach simply compares the student’s observed growth to the target growth, yielding a yes/no determination of whether the student is “on track.” Returning to the example, the student is deemed “on track” because the student’s observed growth (50) exceeds the growth target (30). This “on track” determination can be made another way with equivalent results, by comparing the target score of interest to the student’s projected score found by extrapolating their current gain. Assuming constant gains, the student’s projected grade 6 score is $550 = 450 + 2 \times 50$. Because the student’s projected grade 6 score (550) is greater than the target cut score (490), the student is deemed “on track.” The second approach also compares the observed growth and target growth, but uses an achieved percentage of the target growth. For the example student with observed growth of 50 and target growth of 30, this achieved percentage is $167\% = 50 \div 30$. Note that any percentage greater than 100 corresponds to “on track.”

To facilitate group-level reporting based on gain scores, the gain scores for students within the group are often averaged. As with individual student gain scores, standards for adequate average group gains can be set on the gain score scale or norm-referenced. Alternatively, the student-level classifications of adequate growth can be aggregated to produce a group-level statistic, the proportion of students exhibiting adequate growth.

Group-level growth adequacy under the growth-to-standard extension of the gain score model (i.e., the trajectory model) can also be determined in several ways. First, projected scores within a group can be averaged, and this average can be compared to a target score (e.g., the cut score for Proficiency). Second, if achieved percentages are used, then these can be averaged and compared to some standard, such as 100%. Finally, as with the gain score model, the proportion of students exhibiting adequate growth may be used.

Advantages and disadvantages

Gain scores and the related growth-to-standard measures have numerous strengths. Due to their simplicity and clear connection to achievement scores, they are straightforward to understand, explain, and calculate. Arguably, they are easy to interpret, because they are on the same scale as the underlying test scores. Also, much is known about the measurement properties of gain scores because educational researchers have been studying them for decades (e.g., Lord, 1956). A strength specific to the growth-to-standard measure is that there is no need to set scale-based standards, as it relies on existing cut scores.

At the same time, these gain score measures have several notable weaknesses. First, they require a vertical scale, and therefore the model is not a viable choice for some states. Second, the accuracy and meaningfulness of these measures depends on the soundness of the underlying vertical scale and, in the case of the growth-to-standard measure, articulated cut scores. To the extent that the vertical scale or cut scores are poorly developed or maintained, the justification for typical uses and interpretations of these growth measures may be threatened (Tong and Kolen, 2007). For example, the vertical scale could exhibit floor or ceiling effects, or initial status could be related to the gain scores. Third, direct comparisons of gain scores can be challenging. Vertical scales are designed to facilitate direct comparisons of test scores across grades. However, depending on properties of the scale, typical gain scores may differ by grade. A fourth weakness, touched on above, is that raw gain scores are hard to interpret without a corresponding growth category (e.g., “high”).

The categorical model

Model overview

The categorical model (also referred to as the value table or transition model) is similar to the gain score model and characterizes observed student growth as transitions across performance levels over time. The categorical model aims to address the question, “How has the student grown in terms of performance level classifications?” To illustrate the categorical model, assume that the following ordered performance level categories are used for each grade: Below Basic, Basic, Proficient, and Advanced. These levels are mapped to the test scale using cut scores. Then, to characterize observed growth, all of the possible combinations of levels for the prior and current grades can be represented in a transition matrix, as shown in Table 1.

Table 1 shows how the possible combinations correspond to positive growth, no growth, or negative growth, with respect to the performance levels. Returning to the example student with test scores of 400 and 450 in grades 3 and 4, assume these scores correspond to performance levels of Basic for the prior grade and Proficient for the current grade, respectively. Then, the categorical model would describe this student’s performance as positive growth of one level. The categorical model has recently seen operational use in states such as Florida (Florida Department of Education, 2018) and Virginia (Virginia Board of Education, 2018).

The categorical model may also be used as the basis for a growth-to-standard measure, much like the trajectory model. Whereas the categorical model characterizes observed growth by considering all combinations of prior and current performance levels (as in Table 1), the growth-to-standard approach considers all possible combinations of current and projected performance levels. The result is a transition matrix identical in structure to Table 1, but based on current and projected performance levels. A student’s

Table 1 Example transition matrix.

Prior grade performance levels	Current grade performance levels			
	Below basic	Basic	Proficient	Advanced
Below basic	0	1	2	3
Basic	−1	0	1	2
Proficient	−2	−1	0	1
Advanced	−3	−2	−1	0

Note. Entries indicate the direction (positive or negative) and number of performance levels of growth in the transition. For a growth-to-standard approach, “Prior Grade” and “Current Grade” are replaced by “Current Grade” and “Projected Grade,” respectively.

projected performance level is found by classifying the student's projected scale score in the target year, which can be found using various approaches (e.g., assuming constant gains). A recent proposal by Kentucky ([Kentucky Department of Education, 2018](#)) provides an example of this growth-to-standard measure based on the categorical model.

Modeling requirements

The categorical model requires a vertical scale and articulated performance levels. In some cases, the original set of performance levels, used to describe achievement, is further divided into sublevels, creating additional combinations of levels to characterize growth. For example, the Basic level may be split into Basic-1 and Basic-2 levels. An advantage of this practice is that it has the potential to provide more detailed descriptions of growth. Increasing the number of performance levels, however, introduces additional demands. First, with this practice, the task of creating articulated performance levels across grades becomes more challenging. And second, the underlying achievement scores must be sufficiently precise to support finer-grained descriptions of growth (e.g., [Lee, 2010](#)).

Determining growth adequacy

The categorical model provides a flexible framework for defining growth adequacy at the student and group levels. In general, these approaches depend on a Value Table, which assigns numerical values to all possible transitions. A Value Table can be based on the number of performance levels of growth, as in [Table 1](#), but can also be constructed to reflect more nuanced valuations of growth. For example, in [Table 2](#), students who transition zero performance levels are assigned different values depending on the student's initial performance level. For instance, students maintaining Proficiency each year are assigned more points than those maintaining the Basic level each year.

At the student level, standards can be set for which transitions constitute adequate growth. For example, using [Table 2](#), perhaps all values of 50 or more are deemed adequate growth. The group level calculations can be made in one of two ways. One possibility is to aggregate the student level determinations, resulting in a percentage of students with adequate growth that can be compared to some benchmark (e.g., 50%). A second possibility is to find the average transition value. In this case, a standard is set for the average transition value (e.g., 100). Note that all of the approaches can be used with the growth-to-standard version of the categorical model as well.

Advantages and disadvantages

A notable strength of the categorical model is its clear connection to performance levels, which facilitates communication with various stakeholders. Student and group-level results are also easy to calculate. Another strength is the flexibility afforded by using value tables; these tables can clarify how states value different types of transitions.

The categorical model also has a few weaknesses. Like the trajectory model, the categorical model has stringent modeling requirements, particularly when it is used to project future performance levels. Another weakness is that the coarse nature of the transition categories can obscure variability in underlying gain scores, sometimes with counter-intuitive results. For example, consider two students with the same raw gain score G , but different initial scores. Depending on the locations of the cut scores, the students may have unequal growth transitions. This weakness may be alleviated by subdividing the original performance levels. However, such splitting of levels may not involve a typical systematic standard setting process, which can complicate the interpretation of the sublevels. Finally, the categorical model is not well-suited for students scoring at extreme performance levels. For example, it is unclear what should count as adequate growth when the initial score is classified as Advanced.

The student growth percentile model

Model overview

In its most basic form, the Student Growth Percentile (SGP, [Betebenner, 2009](#)) model is used to characterize observed student growth in normative terms. Though there are different implementations, in general, the model is designed to answer the question, "How does the student's current performance compare to those students with similar score histories?" This is the same question addressed by the residual gain score model, which was of historical interest (e.g., [Cronbach and Furby, 1970](#)) as well as in use by some states (e.g., [California Department of Education, 2021](#)). A detailed description of how SGPs are estimated is beyond

Table 2 Example value table.

Grade 3 performance levels	Grade 4 performance levels			
	Below basic	Basic	Proficient	Advanced
Below basic	0	100	150	150
Basic	0	50	125	125
Proficient	0	0	100	100
Advanced	0	0	100	100

the scope of this entry. However, assume that the scores of 400 in grade 3 and 450 in grade 4 for our example student result in an SGP of 75. The SGP model supports the interpretation that the student's growth was greater than approximately 75% of students with similar starting points (i.e., grade 3 scores near 400). The SGP model has recently seen operational use in states such as New York ([New York State Education Department, 2018](#)) and Oregon ([Oregon Department of Education, 2017](#)).

As with the gain score and categorical models, the SGP model can be extended to create a growth-to-standard measure. Once the target year and performance level have been identified, the SGP model can be used to find the target growth (i.e., target SGP) needed to progress from the initial score to the target score (e.g., the cut score for the target performance level). Returning to our example, recall that the student's grade 3 score is 400. Again, assume the target year and performance level are grade 6 and Proficient, respectively. Then, the SGP model can be used to find the corresponding target SGP, which, for illustrative purposes, we assume is 40. The interpretation is that for the student to progress from the initial score to the target score the student needs a minimum SGP of 40 over the next 3 years. This target SGP is sometimes referred to as an adequate growth percentile (AGP). The AGP growth-to-standard measure has seen operational use in [Colorado Department of Education \(2018\)](#) and [Michigan Department of Education \(2017\)](#).

Modeling requirements

In contrast to the gain score-based methods discussed thus far, SGP models are relatively complex statistical models that can require specialized software, such as the [R Core Team \(2019\)](#) package SGP ([Betebenner et al., 2021](#)). Unlike the gain score-based methods, SGP models do not require a vertical scale. The original implementation ([Betebenner, 2009](#)), which is also the most widely used, depends on a sophisticated and flexible type of regression modeling, known as quantile regression. This approach to SGPs does not require linearity between past and current scores, and does not assume homoskedastic errors. These features increase the flexibility of the approach to fit a wider range of data, but also increase the need for a large sample size. Readers interested in greater technical detail of this approach are referred to [Betebenner \(2009\)](#) and [Betebenner et al. \(2021\)](#). Alternative approaches to SGPs rely on ordinary linear regression ([Castellano and Ho, 2013](#)), which has stricter modeling requirements, but is likely to perform better when sample sizes are smaller. SGP model implementations can also differ in how measurement error in the achievement scores is handled ([Shang et al., 2015](#)).

Determining growth adequacy

At the student level, determining adequate growth can proceed in two ways. The first approach depends solely on observed growth, and sets a standard for the observed SGP. For example, SGPs of at least 50 could be deemed to reflect adequate growth. The second approach simply compares the observed SGP with the student's AGP. If the SGP exceeds the AGP, the student's growth is deemed sufficient—the student is “on track.” As with the trajectory model, alternatively, but equivalently, the student's projected score assuming constant growth (SGP) over time can be found and compared to the target score; if the projected score exceeds the target score, the growth is deemed sufficient or the student “on track.”

To summarize student level growth, SGPs and AGPs are typically aggregated using the mean or median. Alternatively, the percentage of students demonstrating adequate growth can be used. Standards can then be set for these aggregate measures, or they can be compared normatively.

Advantages and disadvantages

A major advantage of SGPs is that they are relatively easy to interpret. An SGP of 75 has the same interpretation, regardless of the underlying score scale or performance levels. Also, an intuitive reference of typical growth, an SGP of 50, is built into the measure. A final advantage is that SGPs do not require a vertical scale.

SGPs, however, also have several weaknesses. First, they require large sample sizes, and even then, SGPs for students with outlier scores can be somewhat unstable. Second, compared to the simpler gain score-based models, the SGP model is more complex to implement although they can be computed with the SGP ([Betebenner et al., 2021](#)) package in [R Core Team \(2019\)](#). Finally, the SGP model, in its most basic form, does not communicate any information about performance levels.

The value-added model

Model overview

Value-added models are used to characterize the contributions of individual teachers, schools, or districts to student achievement, in normative terms. These models are designed to answer the question: “How does the performance for a group of students compare to expectations?” For the simplest approach, expectations depend on students' test scores for the prior year in the same subject. For more complex approaches, such as the popular Educational Value-Added Assessment System (EVAAS) ([Sanders and Horn, 1998](#)) model, expectations may also depend on students' scores in other grades and subjects, students' teachers in other grades and subjects, and average district scores for each grade and subject. Expectations may also depend on student background covariates and classroom aggregates of such covariates. Conceptually, the difference between the expectations and the observed student scores are then attributed to the teacher (or school, depending on the level of interest). These attributions are referred to as effects (e.g., teacher effects), and are relative to those of the average group for the population (e.g., the average teacher within a district). As with SGPs, a more detailed description of how teacher effects are estimated is beyond the scope of this entry (see [Broatch et al., this volume](#); [Koedel et al., 2015](#)). In practice, value-added models are most useful for identifying teachers with especially high or low effect estimates, as opposed to producing precise rankings of the estimates ([Lockwood et al., 2002](#)).

Because value added models have been used for high-stakes decisions within accountability systems, the appropriate use of value-added models and interpretations of results has proven controversial (e.g., [Baker et al., 2013](#)). There is sometimes a tendency to interpret these results as authoritative indicators of teacher effectiveness, perhaps because value-added model results are obtained using sophisticated statistical models and standardized test scores. However, more careful, qualified interpretations are appropriate for several reasons ([Morganstein and Wasserstein, 2014](#)). For example, because value-added models operationalize teacher effectiveness in terms of standardized test scores, they do not directly reflect other dimensions of teacher effectiveness. Another example is that value-added model estimates may be biased (e.g., undervalue some teachers) when background variables related to classroom-level differences are not included in the model ([McCaffrey et al., 2004](#)). Value-added models have seen recent use in numerous states including Ohio ([Ohio Department of Education, 2018](#)) and Tennessee ([Tennessee Department of Education, 2018](#)).

Modeling requirements

Like the SGP model, value-added models are relatively complex statistical models. Most value-added models, including the EVAAS model, are linear mixed models, designed for analyzing hierarchical data ([SAS, 2021](#)). This modeling framework is flexible, and value-added models can differ in several respects, including how many years of test scores are analyzed, and which background variables are included. Unlike the SGP model, value added models explicitly incorporate group membership variables, which correspond to the estimated teacher effects. For most forms of value-added modeling, a vertical scale is not required. Among the growth models covered in this entry, value-added models place the greatest demands on the underlying longitudinal data system. For example, for the EVAAS model, the data system must track students, and their teachers, schools, and districts for each year included in the analysis. Finally, value-added models require large numbers of students per teacher for more precise teacher effect estimates.

Determining growth adequacy

Unlike the other growth models covered in this entry, value-added models are not typically used to obtain student-level growth results. Instead, the teacher (or school or principal) is the focus of the analysis. Standards can be set by classifying a top or bottom percentage of teachers, or by classifying teachers with effect estimates different than a cut-score, using the standard error of measurement.

Advantages and disadvantages

One advantage of value-added models is that many factors believed to influence achievement and growth can be included in the model, and used to adjust expectations. Two other advantages, owing to the modeling framework, are that value added models can accommodate missing data, and that no vertical scale is required. A final advantage of value-added models is that they have received a great deal of attention from researchers (e.g., [Ballou et al., 2004](#); [Chetty et al., 2014](#); [Papay, 2011](#)).

Value added models also have weaknesses. They are complex to implement, as they require large samples, sophisticated longitudinal data systems, specialized software, and considerable statistical expertise. Also, like SGPs, though teacher effects can be explained at a conceptual level fairly easily, it is often challenging to provide a detailed, transparent explanation of the methodology. A final weakness of value-added models is that they are often not designed to produce student level growth results.

Statistical properties of achievement growth measures

This section covers some of the statistical properties and concerns of achievement growth measures. Though these topics are more technical in nature, their inclusion in this entry are important because the utility of growth measures can depend on, or be limited by, statistical properties. Before discussing growth measures, we discuss some aspects of achievement tests (i.e., status measures) and their development, as they provide useful points of comparison. Also, this section focuses on properties of student growth measures, though aggregated growth measures are also briefly discussed.

Design and development of achievement tests for state accountability systems is a complex endeavor, and includes the development of psychometric and scoring specifications. An important psychometric consideration is test precision, which refers to the consistency of the scores across testing occasions. In general, measurement frameworks (e.g., classical test theory) assume that examinee's scores are inconsistent from one testing occasion to another due to measurement error. And regardless of the measurement framework, the observed score, say X , is considered just an estimate of the true or latent score, say T . Notably, test precision can be improved, for example, by adjusting the test length. As such, state tests are designed to have relatively high precision, with reliabilities typically around 0.9 ([Castellano and McCaffrey, 2020a](#)), which facilitates meaningful score reporting.

An important scoring consideration for achievement tests is determining how scores X will be calculated, for example, using sum scores or IRT-based maximum likelihood scoring. Due to decades of study, the properties of popular scoring methods for estimating the latent score T are well-known. A related question is how to communicate the precision of X in score reporting, such as with confidence intervals or conditional standard errors of measurement.

Reviewing these topics as they relate to growth measures reveals some limitations of the methodology and concerns with how growth modeling is used in practice. One limitation of the methodology is that there is general consensus that in the context of state summative testing, individual-level growth measures are substantially less precise than the achievement scores on which they depend ([Castellano and McCaffrey, 2020a](#); [McCaffrey et al., 2015](#); [Monroe and Cai, 2015](#)). This phenomenon is primarily due

to the high correlation of achievement scores across years (McCaffrey et al., 2015, p. 20) and affects all popular growth measures. Unfortunately, this means that it is impractical to improve the precision of individual-level growth measures (e.g., by substantially increasing the length of summative assessments), so long as they are based on conventional achievement tests.

A second limitation is that compared to established methods for scoring achievement tests, less is known about properties of methods for calculating growth, for several reasons. One issue is that the methods themselves (e.g., SGPs) are relatively new, and thus have been studied less. Another issue is that in contrast to conventional test scoring, there are a greater number of approaches to producing growth scores. Even with one approach (e.g., value added modeling), implementations can differ in important respects that affect statistical properties. A final issue is that with growth measures, there are arguably more statistical properties that should be evaluated. One important example is whether the growth estimates exhibit systematic bias, such as whether overestimation or underestimation of true growth depends on prior achievement (McCaffrey et al., 2015; Shang et al., 2015).

Turning to concerns with how growth modeling is used in practice, one shortcoming is that it is sometimes unclear how latent growth is being defined. However, such definitions are critical for evaluating psychometric properties of growth measures, such as precision and bias (Lord, 1958; McCaffrey et al., 2015; Shang et al., 2015).

A related concern is that in score reporting, the precision of growth score estimates or classifications is typically ignored. Excluding this information, however, makes it more difficult for students, parents, and teachers to fairly evaluate student learning. As noted above, in contrast, achievement scores are typically reported with accompanying measures of uncertainty, in accordance with Standard 2.14 of the *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014).

Some, but not all, of these limitations and concerns extend to aggregate growth measures, including concerns with systematic bias (McCaffrey et al., 2015; Shang et al., 2015). One important difference is that aggregate growth measures, such as means, tend to have greater precision than individual growth measures. A notable complication for aggregate growth modeling, however, is student sorting, that is, the systematic difference between classes in terms of student characteristics (Castellano and McCaffrey, 2017, 2020b; Shang et al., 2015).

References

- Alabama Department of Education, 2019. Every Student Succeeds Act Alabama State Plan. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/alconsolidatedstateplanfinal.pdf>.
- American Educational Research Association, American Psychological Association, National Council of Measurement in Education, 2014. *Standards for Educational and Psychological Testing*. American Educational Research Association, American Psychological Association, & National Council of Measurement in Education, Washington, DC.
- Baker, B.D., Oluwole, J., Green, P., 2013. The legal consequences of mandating high stakes decisions based on low quality information: teacher evaluation in the race-to-the-top era. *Educ. Eval. Pol. Anal. Arch.* 21, 1–71.
- Ballou, D., Sanders, W., Wright, P., 2004. Controlling for student background in value-added assessment of teachers. *J. Educ. Behav. Stat.* 29, 37–65.
- Betebenner, D., 2009. Norm- and criterion-referenced student growth. *Educ. Meas.* 28, 42–51.
- Betebenner, D.W., VanIwaarden, A., Domingue, B., Shang, Y., 2021. SGP: Student Growth Percentiles & Percentile Growth Trajectories. R Package Version 1.9-5.0.
- Broatch, J.E., Manski, S.E., Green, J.L., *Statistics: Value-Added Models*. (This Volume).
- Bryk, A.S., Raudenbush, S.W., 1987. Application of hierarchical linear models to assessing change. *Psychol. Bull.* 101, 147–158.
- California Department of Education, 2021. Student Growth Score File Release Webinar. Retrieved from. <https://www.cde.ca.gov/ta/ac/documents/growthwebjune2021.pdf>.
- Castellano, K.E., Ho, A.D., 2013. A Practitioner's Guide to Growth Models. Council of Chief State School Officers.
- Castellano, K.E., McCaffrey, D.F., 2017. The accuracy of aggregate student growth percentiles as indicators of educator performance. *Educ. Meas.* 36, 14–27.
- Castellano, K.E., McCaffrey, D.F., 2020a. Comparing the accuracy of student growth measures. *J. Educ. Meas.* 57, 71–91.
- Castellano, K.E., McCaffrey, D.F., 2020b. Estimating the accuracy of relative growth measures using empirical data. *J. Educ. Meas.* 57, 92–123.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014. Measuring the impacts of teachers I: evaluating bias in teacher value-added estimates. *Am. Econ. Rev.* 104, 2593–2632.
- Cizek, G.J., 2005. Adapting testing technology to serve accountability aims: the case of vertically moderated standard setting. *Appl. Meas. Educ.* 18, 1–9.
- Colorado Department of Education, 2018. Colorado Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/coconsolidatedstateplanfinal.pdf>.
- Connecticut State Department of Education, 2017. Connecticut Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/ctconsolidatedstateplanfinal.pdf>.
- Cronbach, L.J., Furby, L., 1970. How we should measure "change": or should we? *Psychol. Bull.* 74, 68–80.
- D'Brot, J., 2017. Considerations for Including Growth in ESSA State Accountability Systems. Council of Chief State School Officers. Retrieved from. <https://ccsso.org/sites/default/files/2017-10/CCSSOGrowthInESSAAccountabilitySystems1242017.pdf>.
- Domaleski, C., Hall, E., 2016. Guidance for Estimating and Evaluating Academic Growth. Center for Assessment. Retrieved from. <http://www.ncscpartners.org/Media/Default/PDFs/Resources/EstimatingandEvaluatingStudentGrowth.pdf>.
- Dorans, N., Moses, T., *Equating, Linking, and Scaling*. (This Volume).
- Florida Department of Education, 2018. Florida Department of Education Every Student Succeeds Act (ESSA) State Plan. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/flconsolidatedstateplanfinal.pdf>.
- Goldschmidt, P., Choi, K., Beaudoin, J.P., 2012. Growth Model Comparison Study: Practical Implications of Alternative Models for Evaluating School Performance. Council of Chief State School Officers.
- Hill, H.C., 2009. Evaluating value-added models: a validity argument approach. *J. Pol. Anal. Manag.* 28, 700–709.
- Idaho State Board of Education, 2018. Idaho's Consolidated State Plan: The Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act. Retrieved from. <https://www.sde.idaho.gov/topics/consolidated-plan/files/Idaho-Consolidated-State-Plan-Final-March-28-2018.pdf>.
- Kentucky Department of Education, 2018. Commonwealth of Kentucky Revised Consolidated Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/kyconsolidatedstateplan.pdf>.
- Koedel, C., Mihaly, K., Rockoff, J.E., 2015. Value-added modeling: a review. *Econ. Educ. Rev.* 47, 180–195.
- Kolen, M.J., Brennan, R.L., 2004. *Test Equating, Scaling, and Linking: Methods and Practices*, second ed. Springer, New York.
- Lee, W.C., 2010. Classification consistency and accuracy for complex assessments using item response theory. *J. Educ. Meas.* 47, 1–17.

- Linn, R.L., 2003. Accountability: responsibility and reasonable expectations. *Educ. Res.* 32, 3–13.
- Linn, R.L., 2008. Methodological issues in achieving school accountability. *J. Curric. Stud.* 40, 699–711.
- Lockwood, J.R., Louis, T.A., McCaffrey, D.F., 2002. Uncertainty in rank estimation: implications for value-added modeling accountability systems. *J. Educ. Behav. Stat.* 27, 255–270.
- Lord, F.M., 1956. The measurement of growth. *Educ. Psychol. Meas.* 16, 421–437.
- Lord, F.M., 1958. Further problems in the measurement of growth. *Educ. Psychol. Meas.* 18, 437–451.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.
- Martineau, G.J., 2016. A Guide to Understanding and Selecting Measures of Growth for Smarter Balanced Members. Retrieved from. https://www.nciea.org/sites/default/files/publications/Understanding-selectingimplementing-growth-measures_5-27-16.pdf.
- McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T.A., Hamilton, L., 2004. Models for value-added modeling of teacher effects. *J. Educ. Behav. Stat.* 29, 67–101.
- McCaffrey, D.F., Castellano, K.E., Lockwood, J.R., 2015. The impact of measurement error on the accuracy of individual and aggregate SGP. *Educ. Meas.* 34, 15–21.
- McNemar, Q., 1958. On growth measurement. *Educ. Psychol. Meas.* 18, 47–55.
- Michigan Department of Education, 2017. Michigan's Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/miconsolidatedstateplan.pdf>.
- Monroe, S., Cai, L., 2015. Examining the reliability of student growth percentiles using multidimensional IRT. *Educ. Meas.* 34, 21–30.
- Morganstein, D., Wasserstein, R., 2014. ASA statement on value-added models. *Stat. Publ. Pol.* 1, 108–110.
- New York State Education Department, 2018. New York Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/nyconsolidatedstateplanfinal.pdf>.
- Ohio Department of Education, 2018. Ohio's Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/ohconsolidatedstateplanfinal.pdf>.
- Oregon Department of Education, 2017. Oregon's Consolidated State Plan Under the Every Student Succeeds Act. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/orconsolidatedstateplan.pdf>.
- Papay, J.P., 2011. Different tests, different answers: the stability of teacher value-added estimates across outcome measures. *Am. Educ. Res. J.* 48, 163–193.
- Patz, R.J., 2007. Vertical Scaling in Standards-Based Educational Assessment and Accountability Systems. Council of Chief State School Officers, Washington, DC.
- R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. <https://www.R-project.org/>.
- Rogosa, D.R., Willett, J.B., 1985. Understanding correlates of change by modeling individual differences in growth. *Psychometrika* 50, 203–228.
- Rogosa, D.R., Brandt, D., Zimowski, M., 1982. A growth curve approach to the measurement of change. *Psychol. Bull.* 92, 726.
- Sanders, W.L., Horn, S.P., 1998. Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: implications for educational evaluation and research. *J. Person. Eval. Educ.* 12, 247–256.
- SAS, 2021. SAS EVAAS for K-12 Statistical Models. Retrieved from. https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/sas-evaas-k12-statistical-models-107411.pdf.
- Shang, Y., VanIwaarden, A., Betebenner, D.W., 2015. Covariate measurement error correction for student growth percentiles using the SIMEX method. *Educ. Meas.* 34, 4–14.
- Shepard, L.A., 2008. A brief history of accountability testing, 1965–2007. In: Ryan, K.E., Shepard, L.A. (Eds.), *The Future of Test-Based Educational Accountability*, pp. 25–46.
- Tennessee Department of Education, 2018. Every Student Succeeds Act: Building on Success in Tennessee. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/tnessastateplan122018.pdf>.
- Texas Education Agency, 2018. Every Student Succeeds Act Texas State Plan Appendices. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/txtateplanappendices.pdf>.
- Tong, Y., Kolen, M.J., 2007. Comparisons of methodologies and results in vertical scaling for educational achievement tests. *Appl. Meas. Educ.* 20, 227–253.
- Traub, R.E., 1967. A note on the reliability of residual change scores. *J. Educ. Meas.* 4, 253–256.
- Virginia Board of Education, 2018. Virginia Board of Education Every Student Succeeds Act State Plan. Retrieved from. <https://www2.ed.gov/admins/lead/account/stateplan17/vaconsolidatedstateplanfinal.pdf>.

Statistics: Value-Added Models☆

Jennifer E. Broatch^a, Sarah E. Manski^b, and Jennifer L. Green^{b, a} ^a Arizona State University, Glendale, AZ, United States; and ^b Michigan State University, East Lansing, MI, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of D.F. McCaffrey, Value-Added Models, Editor(s): P. Peterson, E. Baker, B. McGaw, International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 489–494, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01374-9>.

Features of the value-added approach	391
Leading value-added modeling approaches	391
Econometric models	391
Statistical models	392
Two-step VAMs	392
Alternative modeling approaches	393
Limitations and concerns with application	393
Bias	393
Non-random assignment	393
Persistence of inputs to achievement	394
Scaling and limitations of achievement measures	394
Concerns about precision and stability of estimates	394
Summary	395
References	395
Further reading	396

Glossary

Bias The expected value of the difference between a statistical estimator and the quantity that is being estimated

Fixed effects Terms included as explicit explanatory variables in linear models in which model parameters are estimated via least squares or maximum likelihood methods. Typically used to refer to models that include indicator variables for individual units such as schools, teachers, or students

Longitudinal data Data from multiple units with repeated measurement on one or more units over time.

Precision The reciprocal of the sampling variance of a statistical estimator

Random effects Term included in statistical models that are latent and assumed to be random variables from a common distribution

Value-Added Models (VAMs) are methods used in education to estimate contributions of individual teachers, or other educational inputs, toward student achievement. Broadly speaking, VAMs attempt to estimate the “value added” on student achievement by a teacher or other educational input, while controlling for other factors, such as baseline student and classroom performance and characteristics (prior achievement, student/family background, school demographics).

VAMs, which were originally introduced in economics literature on the contribution of current inputs or processes to an output after accounting for previous inputs, have become a popular educational measurement tool. In response to the Race to the Top Act of 2009 (U.S. Department of Education, 2016), almost all states adopted teacher accountability systems that utilized VAMs; VAMs’ claimed ability to isolate a teacher’s effect while controlling for a student’s background made them attractive for measuring teacher effectiveness. Although the Every Student Succeeds Act (2015) has recently lessened the required dependence on VAMs for high-stakes accountability policies, VAMs continue to be used across the country as a part of teacher evaluation. VAMs are also used to evaluate school performance and assess the value of educational interventions or programs (Raudenbush and Wilms, 1995).

Studies have shown that high student achievement and highly effective teachers, as measured by VAMs, have been linked to improved life-long outcomes, such as income and educational attainment (Chetty et al., 2014b). Yet, much controversy and discussion surround the use of VAMs in high-stakes evaluation decisions. Although the goal of value-added modeling is to understand how a teacher or another educational input changes student achievement relative to other teachers, most educational settings are not conducive to being able to make causal interpretations of estimated teacher effects (Rothstein, 2009).

☆ *Change History:* Broatch, Manski and Green updated the text and further reading for the entire article.

This article provides an overview of VAMs, along with their limitations and the concerns about their use for high-stakes evaluation. In the following sections, we describe features of the value-added approach and present different types of VAMs that are commonly used in education. Then we discuss several limitations and concerns regarding the use and application of VAMs within education, including bias, stability, and precision of estimates. Although this article focuses on estimating value-added teacher effects, the models, limitations, and discussions presented similarly apply to estimating the effect of other educational factors, such as schools and programs.

Features of the value-added approach

The primary advantage of VAMs compared to other measures used for accountability is their explicit accounting for competing inputs when estimating teacher or school effects. Other accountability metrics—such as comparing schools on average student scores or the percentage of students who surpass an achievement threshold (e.g., the percentage proficient)—typically make no distinctions between appropriated effects of schools and teachers and differences in other inputs that might exist between students in different schools or classrooms. In a VAM, the difference between the student's potential outcome with their assigned teacher and their potential outcome with a different teacher is the effect of the assigned teacher for that student relative to the other teacher. Given that the goal of a VAM is to compare the performance of many teachers, a relevant teacher-level effect is the difference between a student's potential outcome for their assigned teacher and the average of their potential outcomes for all teachers. The average of these teacher-level effects for individual students equals the teacher's overall effect; this average can be over all students in the population or only the students in or likely to be in a teacher's class. In practice, the estimated teacher-level effects are often considered causal effects. However, the claim that the teacher effects are causal effects depends on the random assignment of students to teachers, which is not feasible in a typical educational setting.

Many of the leading value-added modeling approaches, presented in this article, control for the potential differences in schools or teachers by including prior achievement scores rather than simply using demographic or other broad descriptors of students. By incorporating prior achievement, VAMs place greater emphasis on measuring the inputs of schools and teachers and do so more effectively, but not without their own limitations as discussed in the remainder of this article.

Leading value-added modeling approaches

There are many variations of VAMs. This section describes representative examples of the econometric, statistical, and alternative approaches to value-added modeling and their essential features.

Econometric models

The econometric modeling approach to value-added modeling emerged from the cumulative achievement model. The model starts with a general education production function for student achievement:

$$A_{it} = A_t[F_i(t), E_i(t), \mu_{i0}, \varepsilon_{it}] \quad (1)$$

where A_{it} , individual i 's achievement level at the end of the t th year of schooling, depends on the cumulative achievement function A_t and the entire input histories of family (including neighborhood), $F_i(t)$, and school-based educational inputs, $E_i(t)$, a composite variable representing time-invariant characteristics an individual is endowed with at birth (such as innate ability), μ_{i0} , and a normally distributed, mean-zero error, ε_{it} .

Economists make a series of assumptions to specify model (1) in a form that supports estimation of teacher effects using established estimators. In particular, they assume: the cumulative achievement function A_t does not depend on age and is additively separable (i.e., cumulative achievement equals the sum of the effects of the component inputs, and the effect of any one input does not depend on the levels of the others); family inputs are time-invariant and can be combined into a single term with student inputs; and educational inputs can be separated into additive annual contributions of schools, teachers, peers and other classroom inputs (Todd and Wolpin, 2003). Without further assumptions, estimation of the value added of the current teacher would require data on the school, teacher, peer, and classroom inputs for each student's current and entire prior schooling experience. Because these data do not exist, economists typically assume that all educational inputs decay geometrically at a rate λ and that the family inputs change at that same rate, yielding the following model:

$$A_{it} = \gamma_s S_{it} + \gamma_T T_{it} + \gamma_P P_{it} + \gamma_C Z_{it} + \phi_i + \lambda A_{it-1} + \xi_{it} \quad (2)$$

where T_{it} denotes teacher inputs and can be a single factor for a teacher effect, S_{it} , P_{it} , Z_{it} denote school, peer, and other classroom inputs, respectively, ϕ_i is the fixed contribution of student and family, and the error term is $\xi_{it} = \varepsilon_{it} - \lambda \varepsilon_{it-1}$. The parameters of the model are estimated using least squares with models that include indicators for students and teachers (i.e., student and teacher fixed effects) along with the other variables.

Numerous applications use variants of model (2) to study teachers. Some applications assume $\lambda = 1$ and model the first differences in achievement or gain scores; however, others have used $\lambda = 0$ which results in fixed effects analysis on current-level scores

rather than gain scores. In other applications, λ is estimated by including the lagged prior scores as a predictor in a linear regression model. Some variants of model (2) treat ϕ_i as zero and some attempt to estimate both λ and the student coefficients while accounting for the complexities this creates for estimation.

If the students' potential outcomes for a teacher and the associated context are described by a model analogous to (2), then model (2) is a structural model and estimated effects of teachers are defined by the contrasts of the coefficients in γ_T . Teacher-level estimates can be obtained using linear regression techniques under the assumption that students assigned to any teacher's class differ from students assigned to other teachers' classes in terms of their general levels of family inputs (ϕ_i) and their prior achievement scores (A_{it-1}), but not in terms of unobserved factors related directly to their current level of achievement or to their achievement with their assigned teacher. That is, for students assigned to any given teacher, the distribution of the ξ_{it} must not differ from the distribution for other students and must have mean zero.

Statistical models

Statistical modeling approaches to value-added modeling evolved from the hierarchical linear model (HLM) framework which specifies the sources of variance at each level of a hierarchical data structure linking repeated scores to students and students to classrooms, teachers, and schools. One example of this class of models is the "layered model" originally used by the Tennessee Value-Added Assessment System (Sanders and Rivers, 1996; Sanders et al., 1997). HLMs are all variants of the following model, sometimes referred to as the variable persistence model, presented here for 3 years of scores on a single test:

$$\begin{aligned} A_{i1} &= \eta_1 + \gamma_{s1}S_{i1} + \gamma_{T1}T_{i1} + \gamma_{X1}X_{i1} + \zeta_{i1} \\ A_{i2} &= \eta_2 + \gamma_{s2}S_{i2} + \gamma_{T2}T_{i2} + \gamma_{X2}X_{i2} + \omega_{21}\gamma_{s1}S_{i1} + \alpha_{21}\gamma_{T1}T_{i1} + \zeta_{i2} \\ A_{i3} &= \eta_3 + \gamma_{s3}S_{i3} + \gamma_{T3}T_{i3} + \gamma_{X3}X_{i3} + \omega_{32}\gamma_{s2}S_{i2} + \alpha_{32}\gamma_{T2}T_{i2} + \omega_{31}\gamma_{s1}S_{i1} + \alpha_{31}\gamma_{T1}T_{i1} + \zeta_{i3} \end{aligned} \quad (3)$$

where η_t denotes overall means for the population each year, S_{it} is a vector of indicators identifying each school's share of the instruction for student i in school year t and T_{it} is an analogous vector for teachers. The elements of γ_{st} and γ_{Tt} are normally distributed random variables representing school and teacher effects that are assumed to be independent across schools, teachers, and years. X_{it} denotes vectors of student variables. Finally, $\zeta_i = (\zeta_{i1}, \zeta_{i2}, \zeta_{i3})'$ are multivariate normal variables with mean zero and an unspecified covariance matrix Σ , and are independent of teacher, school, and other student effects. The α and ω parameters measure the fadeout of teacher and school effects and can be estimated from the data. Model (3) may be extended to include more years of testing or multiple subject tests. A demonstration of how to fit model (3) can found in Doran and Lockwood (2006).

Restrictions to this model recover most of the other commonly used statistical VAMs. For example, the layered model is a special case of model (3) with no school effects (i.e., all the elements of all the γ_{st} are set to zero), no explicit controls for students' characteristics (i.e., all the elements of all the γ_{Xt} are set to zero) and all the α are set to one so that there is no fadeout of teacher effects. Another special case of model (3), referred to as cross-classified hierarchical linear models, also sets all the α (and ω) to one but specifies $\zeta_{it} = a_i + b_it$ with (a_i, b_i) assumed to be jointly normally distributed variables, to account for the situation where students are grouped differently in classes over time. Traditional HLMs assume students progress with the same classmates and are special cases with α (and ω) set to zero and correlation among ζ_{it} treated as zero.

The model parameters, including the α (and ω) parameters, can be obtained via maximum likelihood or Bayesian methods (Boedeker, 2017; Mai and Zhang, 2018). Bayesian methods have computational advantages and provide a principled method for making many desired inferences about teachers that account for uncertainty in all the model parameters.

Two-step VAMs

To reduce the increasing computational burden of calculating the estimates in model (2) and model (3), a two-step approach was developed. This approach is commonly referred to as an average residual model. The VAM presented in the Measures of Effective Teaching (MET) Research Report (Kane et al., 2013) and another presented in the work of Chetty et al. (2014b), among others, are examples of a two-step model. For instance, the MET fixed effects model, similar to model (2), estimates the expected achievement for each student based on their own and their peers' baseline scores and demographic characteristics:

$$A_{it} = X_{it}\beta + \bar{X}_{jkt}\gamma + \theta A_{it-1} + \lambda \bar{A}_{jkt-1} + \varepsilon_{it} \quad (4)$$

where the subscripts i, j, k, t represent the student, teacher, section/school, and year, respectively. The model statistically controls for the student's baseline score (lag score), A_{it-1} , the student's demographic characteristics, X_{it} , classroom-averaged baseline achievement score, $\bar{A}_{jkt-1}$, and classroom-averaged demographics, $\bar{X}_{jkt}$. The teacher-level value-added estimates are then calculated in step two. The second step estimates a teacher effect for achievement measure A_v as the average of the differences between each student's expected achievement based on model (4) and their observed achievement for year t . Technically, the value-added estimate for achievement measure A , teacher j , section k , and year t , is the average of the student model (4) residuals or the unexplained heterogeneity. Model (4) is limited to estimating a separate teacher value-added estimate for each univariate measure of student achievement.

Alternative modeling approaches

In other applications, analysts develop approaches for estimating value-added teacher effects that have face-validity to teachers and other stakeholders. For example, applications sometimes estimate teacher effects with simple averages of their students' gain scores or their residuals from a linear regression model predicting current achievement scores with a single prior score. Although these modeling approaches might be special cases of the models listed above, they are applied in practice with the simple notion of measuring growth as a means of estimating value added in a way that is more easily understood by stakeholders.

One prevalent example of a growth modeling approach that was developed to be more accessible to teachers and stakeholders is the student growth percentile (SGP) model (Betebenner, 2009), which can be readily estimated in the SGP package (Betebenner et al., 2020) for R. Similar to models (3) and (4), the SGP model estimates teacher effects by relating each student's current year score to their prior year score, but these estimated effects are instead presented as changes in percentiles, which are more easily understood by stakeholders. Technically, to calculate a student's growth percentile, a SGP model uses quantile regression to estimate the conditional density of a student's current achievement score conditioned on their prior achievement scores. Individual SGPs are then aggregated at a teacher or school level to provide evaluation estimates (e.g., the mean or median of the individual student SGPs in a class are used for a teacher estimate).

Each SGP represents the probability of observing a student's current achievement score conditioned solely on their prior achievement. Prior achievement can be determined by any subject and from any measure of student achievement, but in practice, this is often limited to one or two prior achievement scores, resulting in SGPs that are well approximated by percentile ranks of residuals from a linear regression model of current on prior scores (Castellano and Ho, 2013). The model can also be conditioned on other covariates such as student background, but SGPs in practice have excluded other covariates and relied solely on prior achievement measures.

Although these alternative modeling approaches might be appropriate for developing an accountability system that is acceptable to educators, the statistical properties of the resulting estimators still suffer from the same limitations as and potentially more than those resulting from models (3) and (4) (e.g., Castellano and McCaffrey, 2017). Therefore, methods that evaluate the properties of value-added modeling procedures are needed, as are tactics that emphasize the value of both face-validity and good statistical properties to educators and policymakers.

Limitations and concerns with application

As with any model, applications of VAMs are subject to limitations due to the assumptions of the model. Much of the concern around value-added modeling has been related to its widespread application in education to measure individual teacher effects, especially after the passing of federal and state regulations that require teacher evaluation systems primarily rely on measures of student achievement. Though research has shown teachers vary dramatically in effectiveness as measured by VAMs, the use of VAMs in high-stakes decisions about individual teachers has resulted in controversy and led to the critical evaluation of common VAM estimates. In response to the increasing debate on the use of VAMs in education, the American Statistical Association (ASA) issued a statement (2014) on the use of VAMs in educational assessment. This statement raised numerous cautions about the use of VAMs, including concerns of bias, precision, and stability of value-added modeling estimates. The following sections discuss each of these limitations and concerns.

Bias

VAMs are expected to produce unbiased estimates of a teacher's contribution to student learning. The three sources of bias that have consistently been raised as most problematic are (1) nonrandom assignment of students to schools and teachers within schools; (2) improper assumptions about persistence/fadeout of different factors influencing student achievement; and (3) improper assumptions about test score scales.

Non-random assignment

The fundamental problem of inferring teacher effects from observational data is that students are not randomly assigned to schools or teachers. The students taught by different teachers have systematically different characteristics, and unless properly controlled through statistical adjustment, these differences can introduce bias into estimated teacher effects. Bias of teacher-level effect estimates is an issue when students are disproportionately assigned to schools or classrooms based on background characteristics related to student outcomes and true teacher effects. The non-random assignment of students to schools and teachers based on such characteristics introduces systematic differences in student achievement across schools and classrooms that may be confounded with teacher effects (Ballou et al., 2004). This type of confounding makes it difficult to disentangle systematic differences in achievement due to non-random assignment from a teacher's true effect on student learning and may result in biased value-added estimates that over- or underestimate a teacher's true effect.

To combat the systematic differences in student achievement, student-level background characteristics can be introduced as covariates in the model. However, when student background characteristics are correlated with teacher effects, including student-level covariates can mask the effects of teachers and overcorrects for background characteristics, resulting in the underestimation of true,

latent teacher-level effects. For example, if highly effective teachers are assigned to affluent students, socioeconomic status becomes a proxy for teacher quality and true teacher effects are underestimated. [Lockwood and McCaffrey \(2007\)](#) suggest that under the right conditions and sufficiently many prior tests, covariates are unnecessary to control for heterogeneity among classrooms, but in practice this decision to include certain demographic covariates is often not statistically driven, but rather dictated by state, federal and local regulations. There may also be practical limitations on the number of prior test scores available to include in the model. Ultimately, there is a trade-off between potentially confounding teacher effects with student-level variables and potentially biasing teacher effects when deciding on the inclusion of demographic covariates in the model.

Several recent studies, including quasi-experimental and experimental studies, have examined the extent of bias in VAM estimates resulting from non-random assignment ([Kane and Staiger, 2008](#); [Chetty et al., 2014a](#)). The studies do not provide evidence to suggest that the VAM estimates are biased. As in all hypothesis tests, this does not prove the absence of bias, rather that studies have only indicated that the estimates are not significantly biased.

Persistence of inputs to achievement

It is generally accepted that achievement is a cumulative process and that achievement measured by standardized tests reflects both past and current inputs of educational and other factors as shown in model (1). As noted, models make numerous assumptions about the persistence of past inputs to allow for estimation of the effect of current teachers. For instance, many applications either implicitly or explicitly assume that teacher effects persist undiminished to all future years of student achievement.

However, there is evidence that teacher effects fade on standardized assessments and this can result in biased estimates. Studies that estimate α s in model (3) find past teacher inputs into future test scores decay rapidly ([McCaffrey et al., 2004](#); [Lockwood et al., 2007](#)). Studies using alternative methods also find that effects fade rapidly and never find persistence close to 100%. Limited empirical results demonstrate that errors in the assumptions about persistence can lead to biased estimation of the variability of teachers and systematic errors in teacher effect estimates.

Scaling and limitations of achievement measures

The construction and scaling of instruments used to measure student achievement also present issues when estimating teacher effects. Typically, measures of student achievement are assumed to be on a comparable interval scale, where any difference in scores has the same meaning across the entire scale and scores are directly comparable across years ([Schafer, 2006](#)). For example, a student with test scores of 60 and 70 in consecutive years increases their score by 10 and is assumed to have made the same amount of growth as a student with scores of 30 and 40. However, linking scores from different tests to a single comparable scale may require nonlinear transformations, where rates of growth differ in meaning across all ability levels. That is, students with different levels of ability may have the same rates of achievement on the original test scales, as demonstrated in the previous example, but after transforming the scores nonlinearly to a single developmental scale, the students could potentially have different rates of change given an increase of 10 which may subsequently impact teacher effect estimates ([Martineau, 2006](#)).

Teacher effect estimates may also be sensitive to the alignment of achievement test content with curricula, and the weighting of test content to obtain one overall achievement score. VAMs can only be used to estimate the student aptitude that is measured by the achievement test. Varying curricula and test content across grades amplify the complexity of measuring growth in a multidimensional domain. The potential impact of such test construction and scaling issues on teacher effect estimates should be considered. Most VAMs rely on only modeling quantitative measures of student achievement, such as standardized test scores, but the RealVAMs ([Broatch and Lohr, 2012](#)) model is a multivariate VAM that allows the estimation of teacher effects on other types of measures of achievement too (see [Broatch et al., 2018](#) for details on using the R package “RealVAMS” to estimate this model). Although sensitive to the same scaling issues and limitations of typical achievement measures, the RealVAMs model allows the simultaneous modeling of achievement test scores with discrete measures of student success, such as whether a student graduated from high school, to compare teacher effect estimates across multiple potential measures of success.

Different modeling approaches place more or less stringent requirements on the scaling of measures of student achievement. Model (3) makes less stringent assumptions about scaling than models that assume complete persistence, and this model or variants of it might be better suited for some test score data. However, limitations in the scale and test score transformations might invalidate model (3)’s assumptions of normality and implicit linearity in the relationships between scores from pairs of years. Another approach to combat the issue with test scaling is to use SGPs. By comparing student “growth” rather than achievement scores, SGPs do not rely on scaling for estimation, yet aggregated SGPs are known to suffer from the same bias issues as aggregated residuals. Teachers with lower achieving students have underestimated effects and vice versa for teachers with higher achieving students ([Castellano and McCaffrey, 2017](#); [McCaffrey et al., 2015](#)).

Concerns about precision and stability of estimates

Bias in estimated effects is of particular concern because the goal of VAMs is to evaluate teacher performance on student achievement without conflating teachers’ effects with their students’ backgrounds. However, bias is not the only threat to the utility of value-added modeling estimates. Value-added modeling estimates typically lack sufficient precision to support desired inference, especially those related to high-stakes decisions. The lack of precision and the instability, or variability, of teacher effect estimates across years leads educational practitioners to question whether the value-added modeling estimates are capturing teacher effects, or noise. Studies that report precision of individual effects generally show that annual teacher-level estimates have very large errors. For

example, standard errors are typically sufficiently large that about two-thirds of estimated teacher effects are not significantly different from average. Confidence intervals of teacher-level effects are wide and relative rankings contain substantial errors, so that only about 30–35% of teachers ranked in either the top or bottom quintile in 1 year remain there in the next year (Schochet and Chiang, 2013). Similarly, studies consistently find that the correlation of estimated effects from pairs of adjacent years is low, ranging from less than 0.2 to typically no higher than 0.5 (McCaffrey et al., 2009).

This lack of precision in estimates may be a direct result of model specification and the measurement error inherent in measures of student achievement. It has been reported that two different tests on the same subject can result in vastly different value-added modeling estimates of teacher effects (Lockwood et al., 2007; Papay, 2011). When a VAM includes student- or teacher-level covariates, the estimates of teacher effects tend to increase in variability, thus decreasing the precision of the estimates. Although their shown benefits are limited, shrinkage estimators where the teacher value-added estimates are shrunk toward a Bayesian prior (typically an average teacher value-added) have been used to reduce the standard error of these estimates. These shrinkage estimates have been used in the District of Columbia, New York City, and Los Angeles (Herrmann et al., 2016).

Due to the lack of precision, value-added modeling estimates have been compared with other measures of effectiveness. Evidence from large scale projects including the MET Project (Kane et al., 2013), report moderate positive correlations between value-added measures of teacher effectiveness and other measures of teacher effectiveness, such as principal evaluations, parent evaluations, and classroom observations. These positive correlations with other metrics have supported the use of value-added modeling teacher estimates, but researchers advocate for these estimates to comprise only a portion of a more comprehensive teacher evaluation system.

Summary

Within education, VAMs are a popular tool for measuring the effects of individual teachers on student achievement. Although there is no consensus about the best VAM, there is a general consensus that estimating effects with complex models like models (2), (3), and (4) with multiple years of test scores provides more accurate estimates of teacher effects than simple approaches or ad-hoc models. Concerns of bias, precision, and stability of value-added modeling estimates have been researched extensively in recent years and researchers encourage careful use of VAMs to appropriately make inferences about the effects of educational inputs. Regardless of the shortcomings of the current models, value-added assessments are clearly an improvement over other metrics used in accountability and provide a valuable attempt to use sophisticated methods to draw the most from increasingly available longitudinal test score data.

References

- American Statistical Association (ASA), 2014. ASA Statement on Using Value-Added Models for Educational Assessment. <https://www.amstat.org/asa/files/pdfs/POL-ASAVAM-Statement.pdf>.
- Ballou, D., Sanders, W., Wright, P., 2004. Controlling for student background in value-added assessment of teachers. *J. Educ. Behav. Stat.* 29 (1), 37–65. <https://doi.org/10.3102/10769986029001037>.
- Betebenner, D.W., Van Iwaarden, A.R., Domingue, B., Shang, Y., 2020. SGP: Student Growth Percentiles & Percentile Growth Trajectories. R package version 1.9-5.0.
- Betebenner, D.W., 2009. Norm- and criterion-referenced student growth. *Educ. Meas.* 28 (4), 42–51. <https://doi.org/10.1111/j.1745-3992.2009.00161.x>.
- Boedeker, P., 2017. Hierarchical linear modeling with maximum likelihood, restricted maximum likelihood, and fully Bayesian estimation. *Pract. Assess. Res. Eval.* 22. <https://doi.org/10.7275/5wy-8613>.
- Broatch, J., Lohr, S., 2012. Multidimensional assessment of value added by teachers to real world outcomes. *J. Educ. Behav. Stat.* 37 (2), 256–277. <https://doi.org/10.3102/1076998610396900>.
- Broatch, J., Green, J., Karl, A., 2018. RealVAMS: an R package for fitting a multivariate Value-Added Model (VAM). *R J.* 10 (1), 22–30. <https://doi.org/10.32614/RJ-2018-033>.
- Castellano, K.E., Ho, A.D., 2013. Contrasting OLS and quantile regression approaches to student “growth” percentiles. *J. Educ. Behav. Stat.* 38 (2), 190–215. <https://doi.org/10.3102/1076998611435413>.
- Castellano, K.E., McCaffrey, D.F., 2017. The accuracy of aggregate student growth percentiles as indicators of educator performance. *Educ. Meas.* 36 (1), 14–27. <https://doi.org/10.1111/emip.12144>.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014a. Measuring the impacts of teachers I: evaluating bias in teacher value-added estimates. *Am. Econ. Rev.* 104 (9), 2593–2632. <https://doi.org/10.1257/aer.104.9.2593>.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014b. Measuring the impacts of teachers II: teacher value-added and student outcomes in adulthood. *Am. Econ. Rev.* 104 (9), 2633–2679. <https://doi.org/10.1257/aer.104.9.2633>.
- Doran, H.C., Lockwood, J.R., 2006. Fitting value-added models in R. *J. Educ. Behav. Stat.* 31 (2), 205–230. <http://www.jstor.org/stable/3701365>.
- Every Student Succeeds Act, 20 U.S.C. § 6301, 2015. <https://www.congress.gov/bills/114th-congress/senate-bill/1177>.
- Herrmann, M., Walsh, E., Isenberg, E., 2016. Shrinkage of value-added estimates and characteristics of students with hard-to-predict achievement levels. *Stat. Publ. Pol.* 3 (1), 1–10. <https://doi.org/10.1080/2330443X.2016.1182878>.
- Kane, T.J., Staiger, D.O., 2008. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. National Bureau of Economic Research. <https://doi.org/10.3386/w14607>.
- Kane, T.J., McCaffrey, D.F., Miller, T., Staiger, D.O., 2013. Have we identified effective teachers?. In: Validating Measures of Effective Teaching Using Random Assignment. Bill and Melinda Gates Foundation. <https://usprogram.gatesfoundation.org/-/media/dataimport/resources/pdf/2016/12/met-validating-using-random-assignment-research-paper.pdf>.
- Lockwood, J.R., McCaffrey, D.F., 2007. Controlling for individual heterogeneity in longitudinal models, with applications to student achievement. *Electron. J. Stat.* 1, 223–252. <https://doi.org/10.1214/07-EJS057>.
- Lockwood, J.R., McCaffrey, D.F., Hamilton, L.S., Stecher, B., Le, V., Martinez, J.F., 2007. The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *J. Educ. Meas.* 44 (1), 47–67. <https://doi.org/10.1111/j.1745-3984.2007.00026.x>.

- Mai, Y., Zhang, Z., 2018. Software packages for Bayesian multilevel modeling. *Struct. Equ. Model.* 25 (4), 650–658. <https://doi.org/10.1080/10705511.2018.1431545>.
- Martineau, J., 2006. Distorting value-added: the use of longitudinal, vertically scaled student achievement data for growth-based value-added accountability. *J. Educ. Behav. Stat.* 31 (1), 35–62. <https://doi.org/10.3102/10769986031001035>.
- McCaffrey, D.F., Lockwood, J.R., Koretz, D., Louis, T.A., Hamilton, L., 2004. Models for value-added modeling of teacher effects. *J. Educ. Behav. Stat.* 29 (1), 67–101. <https://doi.org/10.3102/10769986029001067>.
- McCaffrey, D., Sass, T., Lockwood, J., Mihaly, K., 2009. The intertemporal variability of teacher effect estimates. *Educ. Finance Pol.* 4 (4), 572–606. <https://doi.org/10.1162/edfp.2009.4.4.572>.
- McCaffrey, D.F., Castellano, K.E., Lockwood, J.R., 2015. The impact of measurement error on the accuracy of individual and aggregate SGP. *Educ. Meas.* 34 (1), 15–21. <https://doi.org/10.1111/emip.12062>.
- Papay, J.P., 2011. Different tests, different answers: the stability of teacher value-added estimates across outcome measures. *Am. Educ. Res. J.* 48 (1), 163–193. <https://doi.org/10.3102/0002831210362589>.
- Raudenbush, S.W., Wilms, J.D., 1995. The estimation of school effects. *J. Educ. Behav. Stat.* 20 (4), 307–335. <https://doi.org/10.3102/10769986020004307>.
- Rothstein, J., 2009. Student sorting and bias in value-added estimation: selection on observables and unobservables. *Educ. Finance Pol.* 4 (4), 537–571. <https://doi.org/10.1162/edfp.2009.4.4.537>.
- Sanders, W.L., Rivers, J.C., 1996. Cumulative and Residual Effects of Teachers on Future Academic Achievement. University of Tennessee Value-Added Research and Assessment Center.
- Sanders, W., Saxton, A., Horn, B., 1997. The Tennessee Value-Added Assessment System: a quantitative outcomes-based approach to educational assessment. In: Millman, J. (Ed.), *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* Corwin Press, Inc, pp. 137–162.
- Schafer, W.D., 2006. Growth scales as an alternative to vertical scales. *Pract. Assess. Res. Eval.* 11. <https://doi.org/10.7275/xjkz-7n67>.
- Schochet, P., Chiang, H., 2013. What are error rates for classifying teacher and school performance using value-added models? *J. Educ. Behav. Stat.* 38 (2), 142–171. <https://doi.org/10.3102/1076998611432174>.
- Todd, P.E., Wolpin, K.I., 2003. On the specification and estimation of the production function for cognitive achievement. *Econ. J.* 113 (485), F3–F33. <https://doi.org/10.1111/1468-0297.00097>.
- U.S. Department of Education, 2016. Race to the Top: Implementation and Relationship to Student Outcomes Executive Summary. Institute of Education Sciences. <https://ies.ed.gov/ncee/pubs/20174001/pdf/20174000.pdf>.

Further reading

- AERA statement on use of Value-Added Models (VAM) for the evaluation of educators and educator preparation programs. *Educ. Res.* 44 (8), 2015, 448–452. <https://doi.org/10.3102/0013189X15618385>.
- Braun, H., 2005. Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models. Educational Testing Center. <https://www.ets.org/Media/Research/pdf/PICVAM.pdf>.
- Everson, K.C., 2017. Value-added modeling and educational accountability: are we answering the real questions? *Rev. Educ. Res.* 87 (1), 35–70. <https://doi.org/10.3102/0034654316637199>.
- Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., Williams, A., 2005. Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ? The Council of Chief State School Officers. https://www.isbe.net/Documents/growth_models_policymaker_guide.pdf.
- Koedela, C., Mihaly, K., Rockoff, J.E., 2015. Value-added modeling: a review. *Econ. Educ. Rev.* 47, 180–195. <https://doi.org/10.1016/j.econedurev.2015.01.006>.
- Lohr, S., 2015. Red beads and profound knowledge: deming and quality of education. *Educ. Pol. Anal. Arch.* 23 (80). <https://doi.org/10.14507/epaa.v23.1972>.
- McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., Hamilton, L.S., 2003. Evaluating Value-Added Models for Teacher Accountability (MG-158-EDU). RAND Corporation. <https://www.rand.org/pubs/monographs/MG158.html>.
- National Center for Analysis of Longitudinal Data in Education Research-Official Website <http://www.caldercenter.org/>.
- Raudenbush, S.W., Bryk, A.S., 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*, second ed. Sage Publications.
- Wainer, H., (Ed.), Value-added assessment [Special issue]. *J. Educ. Behav. Stat.* 29(1).

Automated scoring of constructed-response items in educational assessment

Mo Zhang and Randy E. Bennett, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Validation of automated scoring	398
Operational uses of automated scoring	399
Automated scoring of essays	399
Automated scoring of spoken responses	399
Automated scoring of short text responses	400
Automated scoring of STEM responses	400
Automated scoring of response processes	400
Emerging trends	401
References	401

Automated scoring usually refers to a collection of batch-process or real-time machine-grading approaches that can differ dramatically depending upon the nature of the constructed-response (CR) task being posed and the expected answer. Two classes of CR task might be distinguished (Bennett, 2011). One class involves items that can be scored by exact matching, where the correct answer(s) are known in advance and a response only needs to be compared to the correct answer(s). A second class consists of questions that call for more complex answers that cannot be graded using exact-matching techniques. The automated scoring of these “complex CR” items is often accomplished through a “scoring model,” where features extracted from the responses are aggregated mathematically to produce a score. The resulting automated scores may be binary (e.g., right/wrong) or polytomous, depending on the grading rubric. This entry focuses on automated scoring of this complex CR class and on approaches that are used operationally in assessment programs and learning products.

Automated scoring has been deployed operationally in consequential assessments since the 1990s, when it was introduced for essay grading, medical problem solving, and architectural design (Clyman et al., 1995; Williamson et al., 1999).¹ Since that time, systems have been developed to grade low-entropy speech (e.g., reading a presented sentence aloud), high-entropy speech (e.g., describing one’s favorite sport), short text responses, mathematical expressions, graphical responses, and problem-solving processes in math and science, among other things.

Williamson et al. (1999) outlined the benefits that can be gained from automated scoring (see Table 1). Their list still applies.

We next discuss validation considerations that are common across content domains, uses, and systems. Then, we provide brief overviews of automated scoring in different domains. We close by offering some comments on several important trends in the field.

Table 1 Potential benefits of automated scoring.

<i>Benefit</i>	<i>Explanation</i>
Reproducibility	A response receives the same score if processed again.
Consistency	The same scoring criteria are applied to all responses.
Tractability	Depending on the method, the reasoning behind the automated scores can be traced, controlled, and explained.
Item specification	Automated scoring can encourage test developers to specify, construct, and modify CR items more finely.
Granularity	Responses can be evaluated at finer levels of detail.
Objectivity	Subjectivity is reduced through strict implementation of criteria.
Reliability	Objectivity, consistency, and reproducibility enhance test reliability.
Efficiency	The time and cost of grading open-ended CR responses is reduced.

Adapted from Williamson et al. (1999).

¹Research on automated scoring dates to Page’s (1966) seminal work on essay grading.

Validation of automated scoring

Kane (2013) argued that validation is iterative and adaptive, involving a network of claims and inferences, each of which depends on assumptions, with some assumptions relying on the plausibility of other assumptions (Kane, 2006, 2013). These claims, inferences, and assumptions are woven together in an interpretive argument that underlies score meaning and use. A single study is insufficient to validate the interpretive argument; instead, a more programmatic approach is typically required.

For the case of automated scoring, on what should that research focus? An important efficiency motivation for many testing programs is to determine whether human rating can be reduced or replaced without any practically important validity loss. In this context, the level of validity provided by human rating is generally taken at face value based on rater selection criteria, training, daily qualification, and operational monitoring. As a result, most validation research has focused on agreement between automated and human scores. However, many authors have argued that concordance with human ratings is not nearly enough (Attali, 2007; Bennett and Zhang, 2016; Williamson et al., 2012). Bejar (2012) argues convincingly that any attempt to validate automated scores through agreement with human ones must first marshal evidence to support the validity of that criterion. But more fundamentally, justifying the uses and interpretations of automated scores calls for a broad collection of evidence consistent with the general argumentative approach (Bennett and Zhang, 2016; Drasgow et al., 2006).

What should constitute this collection of evidence has been discussed extensively in the literature. Even though most of the proposed frameworks are for the scoring of text and spoken responses, the key elements are applicable across domains. For example, Williamson et al. (2012) emphasized five areas: construct representation and relevance, relationship to human scores, association with external measures, generalization across tasks and test forms, and consequences. Xi (2010) suggested five similar areas but with special emphasis to impact on students' test preparation behavior and on teaching and learning practice. Bejar et al. (2016) begin with design, distinguishing evidence extraction (i.e., detection and judgment of construct-relevant features) and evidence synthesis (i.e., quantitative aggregation of judged features). Following Bejar et al., Evanini et al. (2017) proposed designing spoken-language assessment with automated scoring in mind, validity being the most important scoring-design and implementation factor.

Whereas assembling a broad collection of evidence is important, that evidence should not be assumed to speak to validation of the automated scoring system itself; that is, results do not characterize the *system's* validity. In fact, validation is highly conditional (Bennett and Zhang, 2016; AERA, APA, NCME, 2014), characterizing *scores* that are dependent on the purpose of the assessment, its stakes, task types, and the examinee and user populations (Zhang et al., 2019).

The conditional nature of validation is recognized in the various frameworks through their concerns for population-group fairness, both with respect to the meaning of scores and their use. That is, automated scoring should *function* in the same way across groups, especially socially protected ones. *Functioning* can be taken broadly to mean all the areas noted by Williamson et al. (2012): construct representation and relevance, relationship to human scores, association with external measures, generalization across tasks and test forms, and consequences.

Fairness issues have recently received more general attention in the artificial-intelligence (AI) community due to highly publicized differential functioning in a wide array of AI applications (e.g., facial recognition, recidivism risk) (Angwin et al., 2016; Grother et al., 2019). Those issues arose in part from failing to appropriately represent population groups in the development of AI models and neglecting to evaluate system functioning with sufficient samples from those groups. In automated scoring—a subfield of AI—very few such studies have been published. Results to date suggest the presence of potentially important disparities in how some systems function across at least some demographic groups (Bridgeman et al., 2012; Ramineni and Williamson, 2018), implying a need to investigate differential functioning much more comprehensively.

Perhaps also under-emphasized in discussions of the validity of automated scoring is the generalizability component, specifically that related to reliability. Haberman et al. (2015) distinguish two reliability types: assessment accuracy and scoring accuracy. *Assessment accuracy* is the reliability of the total test score to which the CR scores contribute in whole or in part. When graded by humans, the *scoring accuracy* for a CR task will be imperfect because of subjective judgment. Automated scoring has the advantage of perfect accuracy, primarily because the same response will receive an equivalent score if the model remains unchanged (Yao et al., 2019; Zhang, 2013), or the system is not swapped out for another.

Researchers have investigated reliability where the prediction target was the human true score (i.e., the expected score taken over an infinite number of independent human judges). Yao et al. (2019) employed the summed human true score taken across essay responses as a target. They reported improved scoring and assessment accuracy when using automated feature scores (e.g., grammar, usage) plus one human rating, over using only one human rating. Using double human scores led to even higher scoring accuracy in two of three testing programs investigated and generated the highest assessment accuracy. In the context of automated scoring of spoken responses, Zhang et al. (2019) showed that using human scores alone can produce fairly high scoring accuracy and assessment accuracy, with only marginal improvement from adding machine-generated feature scores.

For additional discussion of automated scoring validation, see Bejar (2011), (2017), Bennett (2006), Bennett and Bejar (1998), Bennett and Zhang (2016), Williamson et al. (2006), Clauser et al. (2002), Xi et al. (2012), Yang et al. (2002), Yan and Bridgeman (2020), Zhang et al. (2019).

Operational uses of automated scoring

Automated scoring of essays

The most common operational uses of automated scoring in educational assessment are for grading essay responses. The essays are typically written in response to a short prompt, though in some cases the prompt may reference source material (e.g., several short documents) to be consulted as part of the composition task. Prompts may direct examinees toward different writing purposes, including informational, persuasive, and narrative.

Essay scoring systems are often used alone or in conjunction with human raters to provide scores and/or qualitative feedback (Stevenson and Phakiti, 2013). The evaluated characteristics may include linguistic ones (e.g., grammar, mechanics, text coherence, sentence structure and variety, vocabulary sophistication and choice, aspects of content) and/or *n*-grams (i.e., contiguous sequences of *n* items composed of letters, syllables, or words). Features that quantify those characteristics are aggregated to generate a score intended to predict the one a human rater would assign. The prediction models may range from multiple linear regression and cumulative logit, which are simpler to explain, to more inscrutable machine-learning models, such as support-vector machines and K-nearest neighbor, trained on human-rated responses. Combinations of models may also be used with decision rules to accumulate over their results.

The most well-known of the operationally deployed English-language systems are PEG (MI, 2021), IntelliMetric (IntelliMetric, 2021), Intelligent Essay Assessor (Pearson, 2021a), Turnitin Scoring Engine (Turnitin, 2021), and e-rater (ETS, 2021). Some of these systems have also been employed to score text in other languages (e.g., Intellimetric). In addition, scoring systems are found as components of instructional writing products, including Criterion (ETS Global, 2021), WriteToLearn (Pearson, 2021b), Turnitin Scoring Engine (Bolkan, 2015), and WritingPal (Adaptive Literacy Technology, n.d.). These products generally offer instant structured feedback, as well as scores, intended to help users improve English composition skills.

With respect to validation, of critical importance is that, regardless of some vendors' claims, automated essay scoring systems do not understand what they process in the way that a human does (Bennett and Zhang, 2016). Consequently, they are in principle considerably less capable with respect to content-related factors than are human graders. This fact makes the systems susceptible to construct underrepresentation in that they do not appear to accurately judge such characteristics as factual correctness, unusual organizational structures, quality of argument, appropriateness of evidence supporting a position, or quality of rebuttal (Zhang et al., 2016). They may also be subject to construct-irrelevant variance emanating from such gaming strategies as irrelevant text or employing rare vocabulary (Bejar et al., 2014; Zhang et al., 2016). For these reasons, automated scoring may often be combined with human rating in consequential testing programs (e.g., TOEFL iBT, GRE Analytical Writing, GMAT Analytical Writing Assessment), but used as the sole scoring method in lower-stakes applications (e.g., Smarter Balanced interim assessments).

Given the conditional nature of validation, it is impossible in the context of this entry to summarize the relevant literature for the various uses to which the most-commonly employed systems have been put. Bennett and Zhang (2016, Tables 7.5 and 7.7) give an illustration for e-rater and GRE Analytical Writing, in which evidence is assembled relevant to the meaning of human scores as a criterion, as well as the meaning of automated scores in relation to that and other criteria. Their example shows how claims might (or might not) be supportable given the available evidence. See Shermis and Burstein (2013) for more information about particular systems and research relevant to them.

Automated scoring of spoken responses

Speaking tasks are often divided into low complexity (or "entropy") and high complexity (high entropy). Responses to low-entropy tasks are predictable, as when one is asked to read aloud from written text or restate an orally presented sentence. High entropy tasks, in contrast, produce more spontaneous speech.

Among the notable differences between the automated scoring of essays and speech is the need in the latter to recognize what was said, thereby adding a processing layer. That layer initially introduced significant error because a portion of what was said would inevitably be misinterpreted, more so for spontaneous than for low-entropy speech. Over the last decade, the accuracy of speech recognition has improved considerably so that the error has diminished considerably. However, for nonnative speakers of English, recognition may still represent a source of nonnegligible error (Chen et al., 2018). A second important difference is that during recognition, construct-relevant process features can be extracted (e.g., prosody, fluency, pronunciation) for inclusion in scoring. The subsequent steps are not unlike text scoring, involving the use of techniques from natural language processing to compute linguistic features. Then, the full set of extracted features is aggregated to create automated scores intended to predict what a human rater would assign (see Zhang et al., 2019, for a list of sample features). Of note in feature aggregation is that the same variety of modeling techniques as found in essay scoring may be applied, ranging from more explainable hierarchical regression models to impenetrable machine-learning ones, as well as combinations involving multiple models.

The most well-known operationally used English speech-scoring systems are probably Versant (Pearson, 2021a) and Speechrater (Chen et al., 2018; Zechner and Evanini, 2020; Zechner and Loukina, 2020). Those systems, or ones similar to them, are found in both operational assessments and learning products. Among the testing programs using such systems are PTE Academic, Duolingo English Test, TOEFL iBT, and TOEFL Essentials. Commercial products include TOEFL Practice Online (TPO), EduSpeak, Carnegie Speech, Duolingo, and Liulishuo. Some of these products, like Duolingo, evaluate speech in multiple languages, though typically in low-rather than high-entropy form.

With respect to validity, like essay scoring, issues of construct underrepresentation and irrelevant variance are pertinent. For example, automated speech scoring may inaccurately process responses with unusual characteristics. Yoon and Zechner (2017) listed several categories of such response, including no response/total silence, noise constant in the background, distorted recording, partially missing samples in the response audio, not the target expected language, and off-topic. Proper treatment of such responses is essential to fairness, as well as to establishing and maintaining the credibility of automated scoring.

For validation research related to the use of Speechrater for scoring TPO, see Chen et al. (2018) and Wang et al. (2018) for an example of human-machine agreement in (what appears to be) TOEFL iBT. A more dated discussion of Versant research can be found in Bernstein et al. (2010).

Automated scoring of short text responses

Short text responses run from a phrase to a few sentences intended to give facts, analysis, explanation, or conceptual understanding. Traditional approaches to automated scoring presume that the associated tasks have a small set of semantically correct answers which can be expressed as one or more concepts (Sukkarieh and Blackmore, 2009). Each of these concepts can then be represented in a canonical form as an answer key. The automated scoring system then attempts to recognize when a student's response paraphrases, or is implied by, that key. More contemporary methods, however, utilize ML approaches based on n-grams which, given enough training data, can distinguish among score categories and assign new responses to them.

Automated scoring of short text responses may be done in a variety of domains, including English language arts (reading comprehension, writing), history, science, and mathematics (to explain an answer). Among the operationally used systems are the Turnitin Scoring Engine (Bolkan, 2015), Intellimetric (Intellimetric, 2021), and the Intelligent Essay Assessor (Pearson, 2021a). The Smarter Balanced interim assessments use such scoring operationally (Cambium Assessment, undated).

Automated scoring of STEM responses

Automated scoring has been used operationally with a range of response types in disciplines related to science, technology, engineering, and mathematics (STEM). In science-related disciplines, the longest-standing use for consequential assessment has been in medical problem solving, both for licensure as well as for instruction.

The United States Medical Licensing Examination (USMLE) includes a section of computer-based case simulations (USMLE, 2018). A unique aspect of the measure is that the presented problem changes as the examinee interacts with it (e.g., the patient's condition may change, test results may arrive). USMLE automated scoring consequently must consider not only such outcomes as the final diagnosis and prescribed treatment but also the *process* used to manage the patient. Examinee responses related to that process (e.g., tests ordered or not ordered, initial treatments prescribed or not prescribed) are combined with outcome information by using rules derived from domain experts as to what constitutes a professionally responsible performance. Rather than choosing features and aggregation rules empirically, in this approach, humans define those parameters, as well as perform quality control when the scores are produced (Clauser et al., 2016). Margolis and Clauser (2020) describe the USMLE simulations in detail, as well as evidence related to some validity aspects. Ward et al. (2019) review validity evidence related to the use of computer-based case simulations in health-care licensure and education more generally.

A similar approach was used up until recently in the design section of the Architect Registration Examination (ARE; Bejar, 1991; Braun et al., 2006), also a licensure assessment employing methodology relevant to educational measurement. Unlike USMLE, the ARE design problems were not dynamic—that is, the examinee created a drawing in response to a given specification. Scoring was based on domain-expert propositions and aggregated to produce a pass-fail decision.

Among instructional applications, Inq-ITS is an intelligent tutor, or personalized learning system, intended to develop middle-school students' skills in science inquiry (Gobert et al., 2018). Tasks take the form of scaffolded investigations. As in the USMLE computer-based case simulations, both process and end products are evaluated but, in this case, to judge proficiency on specific inquiry practices and their respective sub-skills. Both propositional and ML techniques are employed, along with Bayesian Knowledge Tracing to amalgamate information (Corbett and Anderson, 1995). See Mislevy et al. (2020) for a detailed treatment of the method, as well as a more general discussion of automated scoring in learning systems.

Far simpler methods are generally used in the scoring of mathematical problems. Aside from short text responses, several additional tasks are amenable to automated scoring, including equations or expressions (where there is a single mathematically correct answer that can take many different surface forms) and the symbolic work leading to the final answer. With respect to the former, in combination with its Equation Editor, Pearson's Math Reasoning Engine is operationally employed to score tasks in consequential assessments (Pearson, 2021a). (See Bennett et al., 1997, for an early example of the scoring technique.) Scoring of symbolic work leading to a final answer is more challenging unless enough constraint is added to limit the paths that can be taken (e.g., breaking the problem into smaller steps, as in the case of such learning systems as Carnegie Learning's MATHia).

Automated scoring of response processes

In the above sections, process data figured into operational scoring in medical problem solving, science inquiry, and speech, among other areas. In addition, process data have been collected in operational international and national assessments, where they have been used to enhance interpretation and inform decisions about how to treat specific cases. For example, the 2019 TIMSS Problem

Solving and Inquiry Tasks were administered in over two dozen countries (Mullis et al., 2021). Event timing, navigation from screen to screen, scrolling, and the use of calculators and rulers were evaluated to examine student progress through the tasks so that non-response data could be more carefully analyzed, distinguishing between students who stopped responding early and those who ran out of time. Process data also permitted examination of how interactive task features were employed and whether that use might have contributed to performance that was lower than expected. (Similar data were collected in the 2017 NAEP eighth grade mathematics assessment and are available for analysis in restricted use data sets, NCES, 2020.)

Of note is that, although process data have been collected in these assessment programs for some time, there is as yet no indication that the underlying analyses are anything more than custom point-in-time applications. That such data are being reported, however, suggests that formalized automated systems will eventually result.

Emerging trends

As has been suggested throughout this entry, the rise of ML approaches is perhaps the most important development in the past decade. Tippins et al. (2021) raise concerns with respect to the use of these approaches in personnel-selection contexts, some of which pertain to high-stakes educational assessment. Chief among those concerns are the lack of a theoretical basis for predictors (e.g., *n*-grams in writing or speaking assessment); transparency (due to the black-box nature of the methods and their oftentimes-designation as trade secret); and limited validity evidence (compounded by the lack of a theoretical basis and limited transparency).

Concerns for fairness are also an issue. Although fairness may be investigated empirically, as with validity, the absence of a theoretical basis and lack of transparency make meaningful evaluation challenging. To be sure, these concerns are arguably more serious with respect to high-stakes uses of ML than with formative assessment in learning products. However, the potential for inequity in learning should not be ignored.

A second emerging trend is legal. Because machine algorithms are being used to make consequential decisions in contexts ranging from lending to employment, governments have begun to issue regulations. The European Union's General Data Protection Regulation (GDPR), for example, granted the right to an explanation of an algorithmic decision that "significantly affects" an individual (Meyer, 2018). Similarly, the OECD Recommendation of the Council on Artificial Intelligence (OECD, 2019) states that "AI Actors should commit to transparency and responsible disclosure regarding AI systems ... to enable those adversely affected to challenge its outcome based on plain and easy-to-understand information on the factors and the logic that served as the basis for the prediction, recommendation or decision" (Section 1. IV.1.1). The G20 Ministerial Statement on Trade and Digital Economy (METI, 2019) contains the same language as the OCED Recommendation.

The growth of regulations and recommendations has encouraged the development of explainable artificial intelligence (XAI) (Adadi and Berrada, 2018; Gunning et al., 2019). This new field seeks to make AI decisions transparent to, and traceable by, humans. To the degree it is successful, when applied to automated scoring, XAI should increase validity by, for example, making construct relevance easier to evaluate. This evaluation, in turn, should make it easier to locate and eliminate predictors that function well empirically but have no substantive import.

References

- Adadi, A., Berrada, M., 2018. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Adaptive Literacy Technology, (n.d.). WritingPal. Available <http://www.adaptiveliteracy.com/writing-pal>.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA, APA, NCME), 2014. Standards for Educational and Psychological Testing. AERA, Washington, DC.
- Angwin, J., Larson, J., Mattu, S., Kirchner, L., 2016. Machine bias. Propublica. Available. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Attali, Y., 2007. Construct Validity of E-Rater[®] in Scoring TOEFL[®] Essays (Research Report 07-21). Educational Testing Service, Princeton, NJ.
- Bejar, I.I., 1991. A methodology for scoring open-ended architectural design problems. J. Appl. Psychol. 76, 522–532.
- Bejar, I.I., Flor, M., Futagi, Y., Ramineni, C., 2014. On the vulnerability of automated scoring to construct-irrelevant response strategies (CIRS): an illustration. Assess. Writ. 22, 48–59.
- Bejar, I.I., Mislevy, R., Zhang, M., Rupp, A.A., 2016. Automated scoring with validity in mind. In: Rupp, A.A., Leighton, J. (Eds.), The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications. Wiley-Blackwell.
- Bejar, I., 2011. A validity-based approach to quality control and assurance of automated scoring. Assess Educ. Princ. Pol. Pract. 18 (3), 319–341.
- Bejar, I.I., 2012. Rater cognition: implications for validity. Educ. Meas. 31 (3), 2–9.
- Bejar, I.I., 2017. Threats to score meaning in automated scoring. In: Erclikan, K., Pellegrino, J.W. (Eds.), Validation of Score Meaning for the Next Generation of Assessments: The Use of Response Processes. Taylor & Francis.
- Bennett, R.E., Bejar, I.I., 1998. Validity and automated scoring: it's not only the scoring. Educ. Meas. 17, 9–17.
- Bennett, R.E., Zhang, M., 2016. Validity and automated scoring. In: Drasgow, F. (Ed.), Technology in Testing: Measurement Issues. NCME Applications of Educational Measurement and Assessment Series. Taylor & Francis, New York.
- Bennett, R.E., Steffen, M., Singley, M.K., Morley, M., Jacquemin, D., 1997. Evaluating an automatically scorable, open-ended response type for measuring mathematical reasoning in computer-adaptive tests. J. Educ. Meas. 34, 163–177.
- Bennett, R.E., 2006. Moving the field forward: some thoughts on validity and automated scoring. In: Williamson, D.M., Bejar, I.I., Mislevy, R.J. (Eds.), Automated Scoring of Complex Tasks in Computer-Based Testing. Laurence Erlbaum Associates, Mahwah, NJ, pp. 403–412.
- Bennett, R.E., 2011. Automated Scoring of Constructed-Response Literacy and Mathematics Items. Bill and Melinda Gates Foundation.
- Bernstein, J., Van Moere, A., Cheng, J., 2010. Validating automated speaking tests. Lang. Test. 27 (3), 355–377.

- Bolkan, J., 2015. Turnitin releases scoring engine for essays, short answers. Journal. Available. <https://thejournal.com/articles/2015/04/28/turnitin-releases-scoring-engine-for-essays-short-answers.aspx>.
- Braun, H., Bejar, I.I., Williamson, D.M., 2006. Rule-based methods for automated scoring: application in a licensing context. In: Williamson, D.M., Mislevy, R.J., Bejar, I.I. (Eds.), Automated scoring of complex tasks in computer-based testing. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 83–122.
- Bridgeman, B., Trapani, C., Attali, Y., 2012. Comparison of human and machine scoring of essays: differences by gender, ethnicity, and country. *Appl. Meas. Educ.* 25 (1), 27–40.
- Cambium Assessment, (undated). Frequently Asked Questions: Smarter Balanced Interim Assessment Automated Scoring (2020–2021 Academic Year). Available. <https://ct.portal.cambiumast.com/-/media/project/client-portals/connecticut/pdf/2019/ct-sbac-interim-automated-scoring-faq.pdf>.
- Chen, L., Zechner, K., et al., 2018. Automated Scoring of Nonnative Speech Using the SpeechRaterSM v.5.0. Engine (RR-18-10). Educational Testing Service, Princeton, NJ.
- Clauser, B.E., Kane, M.T., Swanson, D.B., 2002. Validity issues for performance-based tests scored with computer-automated scoring systems. *Appl. Meas. Educ.* 15, 413–432.
- Clauser, B.E., Margolis, M.J., Clauser, J.C., 2016. Issues in simulation-based assessment. In: Drasgow, F. (Ed.), Technology and Testing: Improving Educational and Psychological Measurement. Routledge, New York, pp. 49–76.
- Clyman, S.G., Melnick, D.E., Clauser, B.E., 1995. Computer-based simulations. In: Mancall, E.L., Bashook, E.G. (Eds.), Assessing Clinical Reasoning: The Oral Examination and Alternative Methods. American Board of Medical Specialties, Evanston, IL, pp. 139–149.
- Corbett, A.T., Anderson, J.R., 1995. Knowledge tracing: modeling the acquisition of procedural knowledge. *User Model. User Adapted Interact.* 4 (4), 253–278.
- Drasgow, F., Luecht, R.M., Bennett, R.E., 2006. Technology and testing. In: Brennan, R.L. (Ed.), Educational Measurement, fourth ed. Praeger, Connecticut, pp. 471–515.
- ETS Global, 2021. Criterion. Available. <https://www.etsglobal.org/gp/en/test-type-family/criterion>.
- ETS, 2021. About the E-Rater Scoring Engine. Available. <https://www.ets.org/erater/about>.
- Evanini, K., Hauck, M.C., Hakuta, K., 2017. Approaches to Automated Scoring of Speaking for K-12 English Language Proficiency Assessment (RR-17-18). Educational Testing Service, Princeton, NJ.
- Gobert, J., Moussavi, R., Li, H., Sao Pedro, M., Dickler, R., 2018. Real-time scaffolding of students' online data interpretation during inquiry with Inq-ITS using educational data mining. In: Azad, A.K.M., Auer, M., Edwards, A., de Jong, T. (Eds.), Cyber-Physical Laboratories in Engineering and Science Education. Springer, Cham, pp. 191–217. https://doi.org/10.1007/978-3-319-76935-6_8.
- Grother, P., Ngan, M., Hanaoka, K., 2019. Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects (NISTIR 8280). U.S. Department of Commerce, Washington, DC. Available. <https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf>.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., Yang, G.-Z., 2019. XAI-Explainable artificial intelligence. *Sci. Robot.* 4 (37). <https://doi.org/10.1126/scirobotics.aay7120>.
- Haberman, S.J., Yao, L., Sinharay, S., 2015. Prediction of true test scores from observed item scores and ancillary data. *Br. J. Math. Stat. Psychol.* 68, 363–385.
- IntelliMetric, 2021. IntelliMetric. Available. <https://www.intellimetric.com/direct/>.
- Kane, M., 2006. Validation. In: Brennan, R.L. (Ed.), Educational Measurement, fourth ed. Praeger, Connecticut, pp. 17–64.
- Kane, M., 2013. Validating the interpretations and uses of test scores. *J. Educ. Meas.* 50, 1–73.
- Margolis, M.J., Clauser, B.E., 2020. Automated scoring in medical licensing. In: Yan, D., Rupp, A.A., Foltz, P.W. (Eds.), Handbook of Automated Scoring: Theory Into Practice. Routledge, New York, pp. 445–468.
- Meyer, D., 2018. AI has a big privacy problem and Europe's new data protection law is about to expose it. *Fortune*. Available. <http://fortune.com/2018/05/25/ai-machine-learning-privacy-gdpr/>.
- MI, 2021. Automated Essay Scoring. Available. <https://www.measurementinc.com/products-services/automated-essay-scoring>.
- Ministry of Economy, Trade, and Industry (METI), 2019. G20 Ministerial Statement on Trade and Digital Economy. Available. https://trade.ec.europa.eu/doclib/docs/2019/june/tradoc_157920.pdf.
- Mislevy, R.J., Yan, D., Gobert, J., San Pedro, M., 2020. Automated scoring in intelligent tutoring systems. In: Yan, D., Rupp, A.A., Foltz, P.W. (Eds.), Handbook of Automated Scoring: Theory Into Practice. Routledge, New York, pp. 445–468.
- Mullis, I.V.S., Martin, M.O., Fishbein, B., Foy, P., Moncaleano, S., 2021. Findings From the TIMSS 2019 Problem Solving and Inquiry Tasks. IEA, Boston, MA. Available. <https://timss2019.org/psi/>.
- National Center for Education Statistics (NCES), 2020. National Process Data From the 2017 NAEP Grade 8 Mathematics Assessment (NCES 2020068). Washington, D.C. Available. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020068>.
- Organization for Economic Cooperation and Development (OECD), 2019. Recommendation of the Council on Artificial Intelligence. OECD, Paris, France. Available. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#_ga=2.100756866.2085259285.1559316172-1236900936.1559134188.
- Page, E.B., 1966. The imminence of grading essays by computer. *Phi Delta Kappan* 48, 238–243.
- Pearson, 2021a. Automated Scoring. Available. <https://www.pearsonassessments.com/large-scale-assessments/k-12-large-scale-assessments/automated-scoring.html>.
- Pearson, 2021b. WriteToLearn. Available. <https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/WriteToLearn/p/100000030.html>.
- Ramineni, C., Williamson, D., 2018. Understanding Mean Score Differences Between the E-rater[®] Automated Scoring Engine and Humans for Demographically Based Groups in the GRE[®] General Test (RR-18-12). Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12192>.
- Shermis, M.D., Burstein, J., 2013. Handbook of Automated Essay Evaluation: Current Applications and New Directions. Routledge, New York.
- Stevenson, M., Phakiti, A., 2013. The effects of computer-generated feedback on the quality of writing. *Assess. Writ.* 19, 51–65.
- Sukkarieh, J.Z., Blackmore, J., 2009. C-rater: automatic content scoring for short constructed responses. In: Florida Artificial Intelligence Research Society (FLAIRS) (Ed.), Proceedings of the Twenty-Second International FLAIRS Conference. Sanibel Island, Florida.
- Tippins, N.T., Oswald, F.L., McPhail, S.M., 2021. Scientific, legal, and ethical concerns about ai-based personnel selection tools: a call to action. *Pers. Assess. Decis.* 7 (2). Available. <https://scholarworks.bgsu.edu/pad/vol7/iss2/1>.
- Turnitin, 2021. Turnitin. Available. <https://www.turnitin.com/>.
- United States Medical Licensing Examination (USMLE), 2018. Sample Test Questions: Step 3. Available. https://www.usmle.org/pdfs/step-3/Step3_Sample_Items.pdf.
- Wang, Z., Zechner, K., Sun, Y., 2018. Monitoring the performance of human and automated scores for spoken responses. *Lang. Test.* 35 (1), 101–120. <https://doi.org/10.1177/0265532216679451>.
- Ward, R.C., Muckle, T.J., Kremer, M.J., Krough, M.A., 2019. Computer-based case simulations for assessment in health care: a literature review of validity evidence. *Eval. Health Prof.* 42 (1), 82–102.
- Williamson, D.M., Bejar, I.I., Hone, A.S., 1999. "Mental model" comparison of automated and human scoring. *J. Educ. Meas.* 36, 158–184.
- Williamson, D.M., Bejar, I.I., Mislevy, R.J., 2006. Automated scoring of complex tasks in computer-based testing: an introduction. In: Williamson, D.M., Bejar, I.I., Mislevy, R.J. (Eds.), Automated Scoring of Complex Tasks in Computer-Based Testing. Lawrence Erlbaum Associates, New Jersey, pp. 1–13.
- Williamson, D.M., Xi, X., Breyer, F.J., 2012. A framework for evaluation and use of automated scoring. *Educ. Meas.: Issues Pract.* 31, 2–13. <https://doi.org/10.1111/j.1745-3992.2011.00223.x>.
- Xi, X., Higgins, D., Zechner, K., Williamson, D.M., 2012. A comparison of two scoring methods for an automated speech scoring system. *Lang. Test.* 29, 371–394.
- Xi, X., 2010. Automated scoring and feedback systems: where are we and where are we heading? *Lang. Assess.* 27, 291–300.
- Yan, D., Bridgeman, B., 2020. Validation of automated scoring systems. In: Yan, D., Rupp, A.A., Foltz, P.W. (Eds.), Handbook of Automated Scoring: Theory Into Practice. Routledge, New York, pp. 445–468.
- Yang, Y., Buckendahl, C., Juszkievicz, P., Bhola, D., 2002. A review of strategies for validating computer-automated scoring. *Appl. Meas. Educ.* 15, 391–412.
- Yao, L., Haberman, S.J., Zhang, M., 2019. Penalized best linear prediction of true test scores. *Psychometrika* 84, 186–211.

- Yoon, S., Zechner, K., 2017. Combining human and automated scores for the improved assessment of non-native speech. *Speech Commun.* 93, 43–52.
- Zechner, K., Evanini, K. (Eds.), 2020. *Automated Speech Technologies for Speaking Assessment*. Routledge, New York, pp. 21–31.
- Zechner, K., Loukina, A., 2020. Automated scoring of extended spontaneous speech. In: Yan, D., Rupp, A.A., Foltz, P.W. (Eds.), *Handbook of Automated Scoring: Theory Into Practice*. Routledge, New York, pp. 445–468.
- Zhang, M., Chen, J., Ruan, C., 2016. Evaluating the Advisory Flags and Machine Scoring Difficulty in the E-Rater® Automated Scoring Engine (RR-16-30). Educational Testing Service, Princeton, NJ. Available. <https://onlinelibrary.wiley.com/doi/full/10.1002/ets2.12116>.
- Zhang, M., Bridgeman, B., Davis, L., 2019. Validity considerations for using automated scoring in speaking assessment. In: Zechner, K., Evanini, K. (Eds.), *Automated Speech Technologies for Speaking Assessment*. Routledge, pp. 21–31.
- Zhang, M., 2013. Contrasting automated and human scoring of essays. *R&D Connections*. No. 23, Princeton, NJ: Educational Testing Service.

The current trends and opportunities for machine learning in learning analytics

Carol M. Forsyth, Caitlin Tenison, and Burcu Arslan, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	404
Learning Analytics	405
Student modeling	405
Proficiency modeling	406
Behavior modeling	406
Adaptation and recommendation	407
Feedback and score reporting	407
Machine learning in learning analytics	408
Supervised learning	409
Unsupervised learning	410
Reinforcement learning	410
Trends and opportunities for the future of learning analytics	411
Theoretically-grounded educational data mining	411
Computational cognitive models	411
Computational linguistic analysis and artificial agent-based interaction	412
Future opportunities	412
Conclusions	412
References	413

Introduction

Imagine a world where every student is provided with a dedicated expert who could instantly review their work and provide them with feedback tailored to their interests and experiences. This expert could inspect the records of numerous other students to predict what educational intervention would be most helpful to their individually assigned student. These experts could inform teachers by providing an accurate real-time picture of their class, so they can focus their effort where it is most needed. Furthermore, each expert would move with their student across classes as their student progressed from one grade to the next, and provide personalized, up-to-date guidance to help their student navigate the educational system and eventually start a career.

Currently, private one-on-one tutoring provides some of the support necessary to fulfill this dream, but the cost of these tutors limits their use. For example, a high-quality tutor can teach based on the student knowledge, adapt to this student's misunderstandings, create opportunities to apply knowledge, provide motivation, adjust based on the student's emotions and provide tailored instructions on exactly which aspect of the learning experience a student can improve on next. The gold standard result of one-to-one tutoring is an increase by two standard deviations compared to students who learn via classroom-based instruction (Bloom, 1984), an opportunity that would be transformative to provide to all students but is simply not affordable to most parents and students.

With the advances in computation, modeling of learners, and Machine Learning (ML) methods, the field of Learning Analytics (LA) emerged to not only fulfill the role of an individual tutor to all students (regardless of socioeconomic status) by personalizing instruction to the student often within digital environments but also to report this information to the teacher. In sum, the goal of LA is to measure, collect, analyze, and to report data about learners and their contexts, for purposes of understanding the learner and identify ways for providing the above-mentioned type of personalized, adaptive instruction in digital learning environments (LAK 2011: 1st International Conference Learning Analytics and Knowledge, 2011).

A well-designed digital learning or interactive assessment environment provides opportunities for human experts (e.g., teachers, assessment developers) to adapt the environment to their needs and make valid inferences about the layers of knowledge, cognitive processing, and affect for each student with the help of LA. These inferences can inform how personalized feedback and one-on-one tutoring is provided to students at a low cost as well as alert teachers when students are struggling (for a review of computerized *Intelligent Tutoring Systems* see Koedinger et al., 1997; Mousavinasab et al., 2021; VanLehn, 2011). One of the main goals of LA research is to optimize each student's learning in and out of the classroom providing teachers with detailed information about the student and class. For example, with LA operating in a digital learning environment, an ML algorithm may identify the misconception at an early stage in students' learning and provide feedback directly to the students and/or to the teacher to quickly correct the misconception. This opportunity may not occur without a detailed level of personalization brought by LA and ML resulting in some students being disadvantaged. Thus, bringing LA and ML approaches together offers each student a closer chance of an equal opportunity in learning as well as assessment of learning.

In this chapter, we provide a general introduction to the primary concepts and areas of research within LA, and highlight the different analytic approaches used to understand and support both students and teachers. As research within LA increasingly applies and extends ML methods, we highlight research along the three branches of ML (i.e., supervised, unsupervised, and reinforcement learning) which has made significant advancement toward the vision of high quality, adaptive or individualized instruction. These advancements make it possible to adapt and recommend instruction, provide immediate feedback, and present score reports as students' progress through the learning experience. We conclude by pointing out the trends and opportunities for the future of LA.

Learning Analytics

Learning Analytics (LA) and Educational Data Mining (EDM) are two fields that aim to measure, collect, analyze, and to report data about learners. Despite some differences, the co-founders of both fields admit that researchers in the field often cross-publish between these two fields because of the common goal of aiding education through analytics and theory (Baker and Siemens, 2014). In this article, we use the term LA in a broader sense, and we include research from both EDM and LA to convey to the readers how analyzing educational data can provide actionable information to students, teachers, and other educational stakeholders. This is a timely message as remote and distance learning becomes increasingly common and solutions such as LA are needed.

LA describes a broad array of analytic approaches that provide scientific and data-driven insights to improve education. While the field of LA brings together diverse topics of research that touch many different types of educational environments, much of the work takes advantage of technology to collect data and drive change. Common educational environments studied within LA include learning management systems, Intelligent Tutoring Systems, and Massive Open Online Courses (MOOCs). These computerized educational environments produce complex data about the learning process such as log files resulting from student interaction. Through the analytic methods central to LA, this data feeds the knowledge discovery process and inspires the development of theories and methods to further improve learning. Fig. 1 illustrates the cycle of research in LA (based on similar figures by Khalil and Ebner, 2015; Romero and Ventura, 2020).

LA has dramatically increased our understanding of student learning by adapting student learning environments and providing teacher reports (for a review, see Baker and Siemens, 2014), via detecting affective states and modify learning experiences accordingly (e.g., D'Mello and Graesser, 2012; D'Mello et al., 2008), representing student cognition to improve math learning (e.g., Koe-dinger et al., 1997), creating and implementing optimal spaced practicing methods (e.g., Pavlik and Anderson, 2008; Walsh et al., 2018) and providing profiles of students characteristics in complex domains such as collaborative problem-solving (e.g., Andrews-Todd and Forsyth, 2020; Forsyth et al., 2020). Below, we briefly review the core concepts in LA research and applications of LA.

Student modeling

For many years, an important aspect of LA has been student modeling (Chrysafiadi and Virvou, 2013; Webb et al., 2001) which is a method of tailoring to the individual learner's needs based on detection (or assessment) of the immediate state of the learner. The focus of student modeling is to capture the affective, behavioral, and cognitive components which are important for student learning and complex assessments. Perhaps the simplest instantiation of a student model can be seen in educational assessments where the goal is identifying what a student needs to know to master a topic. For example, perhaps additional evidence is required to determine a student's mastery of a given topic in a computerized adaptive testing environment. Therefore, the environment adapts to present the student with additional questions to gather the needed evidence before moving on to a new topic (e.g., see Evidence-Centered Design; Mislevy et al., 2003). LA researchers have extended this concept of adaptation by considering a range of different learner models at various levels of granularity (see Zapata-Rivera and Arslan, 2021 for a brief review of different learner modeling approaches).

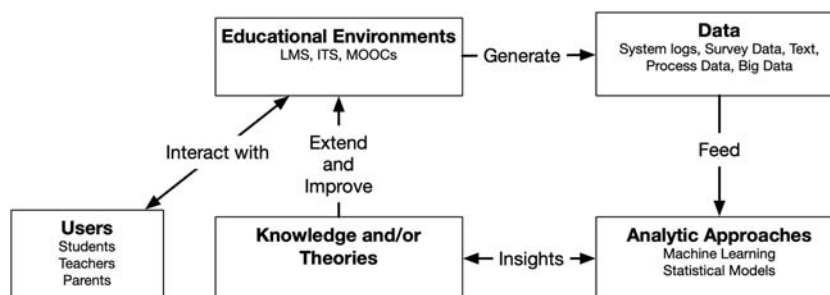


Fig. 1 The traditional research cycle within learning analytics.

Proficiency modeling

The focus of proficiency modeling is to identify what a student knows at a given point in time to either adapt content or provide instructional feedback to help the student reach their optimal performance (for review see [Pelánek, 2017](#)). One of the earliest proficiency modeling approaches, Bayesian Knowledge Tracing (BKT), was developed based on the psychological theory of skill acquisition ([Corbett and Anderson, 1994](#)) and has been adapted and used within numerous digital learning environments (for more details, see [Pelánek, 2017](#); [Ritter et al., 2015](#)). BKT considers not only what the student *knows* but whether or not the student is able to *do* the given skill (roughly corresponding to the concepts of *declarative* vs. *procedural* knowledge). The model requires that instructional designers identify knowledge components encompassing both types of knowledge that underlie each activity a student interacts with. For each Knowledge Component, BKT estimates a set number of states with specific semantics about the what the learner knows based on students item responses (e.g., “know”: a correct answer that accurately reflects student knowledge and/or skills; “guess”: an unlikely correct answer given the student’s previous answers; “slip”: an unintentional error). By observing the sequence of types of item responses generated by a learner, BKT estimates the probabilities that students are in one of those latent states. For example, a student may answer a problem correctly because they *know* the material or because they *guessed*. As the model is provided examples of a student performing a specific skill, it learns the probability that their actions reflect those latent states. These probabilities have valuable psychometric properties and use cases.

Since the development of BKT, there has been considerable work to refine, extend, and develop alternative approaches for estimating student proficiency. BKT has undergone many developments to include additional aspects such as general item difficulty (e.g., [Pardos and Heffernan, 2011](#)), adaptation (e.g., [Yudelson et al., 2013](#)), and skill dependencies ([Meng et al., 2019](#)). Logistic based models, such as Additive Factors Model (AFM; for review see [Durand et al., 2017](#)), can be implemented to derive the student’s current level of proficiency, difficulty of a skill, and learning rate of the skill (for more details, see [Cen et al., 2007, 2008](#); [Thai-Nghe et al., 2010](#)). AFM is a type of multi-dimensional item response theory that has been adapted to capture several assumptions about how people learn within an adaptive computerized learning environment focusing on tutoring, often referred to as Intelligent Tutoring System (ITS) (for review see [Graesser et al., 2012a](#)). Given the interpretability of the parameters and the numerous implementations available open source (e.g., [MacLellan et al., 2015](#); [Pardos and Dadu, 2018](#)), AFM has become a popular tool for both estimating student proficiency and identifying issues within tutoring systems (e.g., [Cen et al., 2006](#); [Goutte and Durand, 2020](#); [Long et al., 2018](#)). More recently, deep learning approaches which utilize ML algorithms (described in Section [Machine learning in learning analytics](#)) have started to be implemented to estimate student proficiency. Approaches such as Deep Knowledge Tracing has demonstrated how a ML method (i.e., a long short-term memory model) can be used to successfully predict student performance over sequences of items (for more details, see [Piech et al., 2015](#); [Wilson et al., 2016](#); [Xiong et al., 2016](#)). Finally, theories of how humans forget information have informed the creation of several proficiency models that consider both the impact of practice and time between practice opportunities when estimating student mastery of skills (e.g., [Eglington and Pavlik, 2020](#); [Pavlik and Anderson, 2008](#); [Walsh et al., 2018a,b](#)). These proficiency models have proved to be especially important in accurately capturing student knowledge across many, spaced sessions. Research comparing different proficiency modeling approaches, illustrates that there is no single approach that works best in all contexts but rather taking into account the domain and environment can reveal the optimal algorithm (e.g., see [Doroudi and Brunskill, 2019](#) for on the equitability of Knowledge Tracing methods). The most appropriate proficiency modeling approach varies depending on the type of data available, the structure of the learning task, the assumptions about skill dependencies and learning processes, and the desired use of these algorithm’s results.

Behavior modeling

Behavior modeling focuses on estimating features of the learner’s state beyond what they know, that may affect learner’s activity within an educational environment. Research on behavioral modeling targets a wide range of different states from affective states, motivational states, physical environment, learning relevant constructs (e.g., self-regulated learning) to states that capture more complex interactions between learners, their peers, and their instructors (e.g., collaborative learning). One type of behavioral modeling that is a flourishing area of research within the LA field focuses on emotion detection (for a complete review, see [D’Mello and Kory, 2015](#); [D’Mello and Graesser, 2012](#)). Specifically, previous research suggests there are six common cognitive-affective states experienced during learning: engaged concentration, confusion, boredom, frustration, delight, surprise ([Baker et al., 2010](#)). Interestingly, confusion is the cognitive-affective state most likely to lead to high levels of understanding ([D’Mello and Graesser, 2014](#)). [D’Mello \(2017\)](#) increased learning gains in an existing ITS by detecting and adapting to human emotion. He was able to detect emotions via multiple sources of evidence (i.e., facial expression, body movements, and linguistic cues). Although some pattern matching algorithms can automatically code some of these cues (e.g., FACETS which can match 14 facial movements to an input of a student face based on the research of [Ekman and Friesen, 1978](#)), these labels are still needed as input to an overall learning prediction model. This is certainly an exciting prospect, but the intuitively obvious issue was the difficulty of getting students in regular classrooms to wear sensors. Therefore, [Ocumpaugh et al. \(2012, 2015\)](#) discovered a method of detection without sensors for identifying student’s affective states in classrooms in an unobtrusive method referred to as the BROMP protocol, which identifies non-verbal body movements to detect engaged versus disengaged students. This method provides feedback to the teachers which allows them to intervene to maximize student learning and provides an additional source of evidence for validating and inspiring other detectors. Specifically, LA researchers introduced new detectors and methods for detecting cognitive-affective states in adaptive learning systems, such as ASSISTments ([Botelho et al., 2017](#); [Heffernan and Heffernan, 2014](#)) and Physics Playground ([Shute and Ventura, 2013](#)). Thus, behavior modeling continues to grow as an area where students’ current knowledge, affective states, and other factors can be included to personalize the learning experience.

Adaptation and recommendation

Proficiency and behavior modeling serve as a foundation for supporting adaptive instruction. However, how to change the system in response to information about the learner can vary depending on what actions are available within the system, the time scale of focus, and the learning goals the system is meant to support. The area of LA research focused on adaptation and recommendation targets the algorithms for implementing how the instructional system changes in response to the information provided by the student models described in the previous section. Adaptation and recommendation can exist at multiple levels of granularity within an educational system: within a practice item, across practice items, between lessons, or even between courses.

Within a given lesson researchers often make the distinction between *inner loop* and *outer loop* adaptation (VanLehn, 2006). Inner loop adaptation takes advantage of model-tracing, a means of tracking a student's problem-solving process and providing support based on students' actions. These methods often involve the use of rules or scripts to identify when the system should intervene (Brown and VanLehn, 1982). For example, Aleven et al. (2006) found that by developing 'bug' rules, a type of production rule (or *if-then* statement) that captured common misuse of hints within a tutoring system, could be used to intervene and provide feedback when these misuses occur. In a pilot study, Aleven et al. found that by providing this adaptation, they were able to correct maladaptive help seeking strategies. This approach of authoring bug rules has been extended to work in a wide range of learning environments from Intelligent Tutoring Systems (e.g., Tacoma et al., 2021) to MOOCs (e.g., Aleven et al., 2015). Inner loop adaptation can adapt to offer instructional support in less structured learning environments. For example, conversational agents have been used across a wide variety of domains in order to structure problem solving, probe student's knowledge and provide input similar to human tutors in the form of feedback which is discussed in detail in the next section (e.g., Graesser et al., 2001, 2014; Hobert and Meyer von Wolff, 2019; Rus et al., 2013). These agents offer input to students based on production rules that are structured to operate like real expert humans. Although agents have been used in numerous capacities across a variety of domains, conversational agents allow for natural language conversations between the student and artificial agent (e.g., Biswas et al., 2010; Gholson et al., 2009; McNamara et al., 2006; Olney et al., 2012; Ward et al., 2013). During natural language conversations, the human input is analyzed with natural language processing algorithms (e.g., Jurafsky and Martin, 2008; Landauer et al., 2013) to determine the intent of the human input and appropriately respond. This type of adaptive dialogue simulates the types of conversations human tutors might engage in when working with students and have been shown to produce similar learning gains to human tutoring (e.g., Graesser et al., 2004; Graesser et al., 2012a,b; Hu and Graesser, 2004; McNamara et al., 2006; Olney et al., 2012; VanLehn et al., 2007; VanLehn, 2011).

Adaptation that involves the selection of what activity students should perform next is referred to as *outer loop* adaptation. The area of research in outer loop adaptation referred to as a *Stopping Policies* which are set rules to determine whether a student should continue practicing a given skill or stop practicing a given skill and progress to new content (e.g., Pelánek and Řihák, 2018; Ritter et al., 2016; Sales and Pane, 2019). The stopping policy of mastery learning is based on student proficiency models such as BKT to estimate when a student knows or has "mastered" a skill. This type of adaptation is a means for an ITS to alter the pacing of material to match students' needs which lead to claims that ITS can produce double the learning gains in half the time of a standard brick-and-mortar class (Koedinger and Corbett, 2006). Distinct from *Stopping Policies*, research on *Instructional Policies* focuses on determining the order in which new content should be practiced (or order of instruction). This area of research encompasses a diverse array of analytic methods and can focus on a range of different levels of decision-making (e.g., Bassen et al., 2020; Iglesias et al., 2009; Khalid et al., 2020; Koedinger et al., 2013; Tarus et al., 2018). In subsection **Reinforcement learning**, we highlight in more detail efforts for ML to identify the optimal sequencing of instruction and practice; however, researchers have identified numerous promising approaches for generating recommendations often in the form of feedback that is sensitive to the data available and needs of the educational system.

Feedback and score reporting

Adapting to a specific human's needs by providing immediate feedback has a long documentation of being incredibly helpful to student learning (for a review, see Shute, 2008). Feedback is provided to students based on recommendations from LA which can be challenging as the feedback must be given at the perfect critical moment and be the correct type of feedback. LA helps identify critical moments where an intervention would be most helpful to a student (e.g., Pardo et al., 2019). However, there are various types of feedback that can be provided to help these students that have been outlined in a meta-review by Shute (2008). Although covering all types would be beyond the scope of this article, we would like to provide a couple of examples. One example is *verification* or simply telling a student whether an answer is correct or incorrect. This can help the student know when to either move on or try again but does not explain to the student how the answer is correct or incorrect. Another example is *elaboration* where the feedback explains multiple aspects about why the given answer is correct or incorrect. Immediate feedback during natural language conversations for tutoring may be more defined to mimic the types of feedback given by real human tutors in dialogic frameworks which are beyond the scope of this article to define (see chapter in this Encyclopedia by Graesser et al.). Moreover, it's important to note that there are still many more possible types of feedback which is why the research on the effect of different types of feedback on student learning is still continuously being explored (for a more recent review, see Mertens et al., 2021) and other types of information are given to stakeholders in the form of score reports.

Score reporting often refers to reports provided to stakeholders after a longer period of time (rather than immediate feedback) requiring information about the audience which could include students, teachers, parents and other stakeholders (Shute, 2008; Zapata-Rivera and Katz, 2014). The research on score reporting can be invaluable to understand what should be provided to each of the various stakeholders (e.g., Kannan et al., 2018a; Kannan et al., 2018b; Zapata-Rivera et al., 2018). Specifically, the needs, knowledge and attitudes of the intended stakeholder should be taken into account when providing a score report or creating a dashboard (Zapata-Rivera and Katz, 2014). However, providing a comprehensible summary information about student performance, strategies to solve a problem, knowledge and skills based on process data (i.e., log files that include user and system events together with their timestamps) can be challenging (Forsyth et al., 2017).

Teachers may have difficulty interpreting reported detailed information collected via LA such as scores relating to cognitive psychology constructs that are correlated with learning (e.g., number of words generated; Zapata-Rivera et al., 2018). Score reporting with information from process data must be understandable by all stakeholders (e.g., the student, teacher, and parent). For example, after realizing that teachers were not understanding the reported material, Forsyth et al. (2017) used an artificial agent to explain each construct to teachers whenever they needed a clarification. This example highlights the gap between what LA researchers may think is important based on literature and what teachers perceive as needed in the classroom. For more details on reporting this type of process data to teachers, see the chapter on Conversation-Based Assessment (CBA; e.g., Jackson and Zapata-Rivera, 2015; Lopez et al., 2017, 2021) in this Encyclopedia. This example highlights the gap between what LA researchers may think is important based on literature and what teachers perceive as needed in the classroom.

Incorporating this information into dashboards requires a good understanding of Human-Computer Interaction, a field focused on how people interact with computer operated machines, to ensure the human knows what to do during interaction. With the help of LA, dashboards have provided information necessary for many students to see their progress in open learner models (an important area in the field; see Bodily et al., 2018). Perhaps the adaptive diagnostic formative assessment system ASSISTments (Heffernan and Heffernan, 2014) has the most notable dashboard as teacher promotion has been the driving force of the mass level adaptation of student use of this free system (over half a million since the onset of COVID-19) and cited in external evaluations as an extremely effective system that met the challenge of aiding students in need (Barshay, 2021). Teachers can see on a sub-skill level what their students understand on an individual and classroom level (Zapata-Rivera and Katz, 2014) and have additional control to assign assignments, researcher created interventions, and display of reports in classes to guide class instruction without embarrassing one student. The developers of ASSISTments accredit their ability to listen to teachers input, thus keeping teachers in the driver's seat of this both adaptive (controlled by the computer) and adaptable (controlled by the teacher) digital environment as the reason for their great success (Heffernan and Heffernan, 2021; Murphy et al., 2020; Sahni et al., 2021).

These are just a few applications of LA. To be clear, LA and ML are not always implemented together. However, if one is new to LA, it may be useful to understand the various types of ML algorithms often used. Therefore, we provide a primer below.

Machine learning in learning analytics

The discipline of Machine Learning (ML) focuses on constructing computer systems and algorithms that automatically improve in accuracy through experience. Although ML has been an active field of research since the 1950s (see Samuel, 1959 for the implementation of ML in the game of checkers), the increased use of technology and resulting data has created opportunities for these algorithms to address numerous practical problems. Education has seen a similar growth in digital learning and assessment technologies over the past 30 years. ML is a primary set of analytic tools used within LA to analyze data resulting from these complex digital learning and assessment environments (Baker and Yacef, 2009; Hernández-Blanco et al., 2019; Romero and Ventura, 2020; Salloum et al., 2020). We review the main types of ML algorithms by providing examples from the field of education to explain how these methods work.

In traditional software engineering, programmers distill knowledge into computer code which instructs the computer to perform specific processes based on input to generate the desired output. In ML, rather than have a human author the computer program, the human provides the computer machine learning algorithms together with training data (i.e., data examples) to identify an optimal computer program that learns from the training data and produces the desired behavior according to a specific metric (e.g., performance). Take, for example, the task of learning to detect whether a student is engaged (or focused on the task) or distracted in a digital reading exercise (such as described in the Section [Behavior modeling](#)). It may be hard to build a software program to detect engagement. However, a human expert (such as a trained researcher) may be able to provide examples of students who are engaged and distracted. These examples can be decomposed into a set of inputs from the corresponding log files (i.e., features) for the ML algorithm. These features might include information like how quickly students were scrolling in their reading task, how often they spoke to their classmates, how many times they looked away from the screen, or even capture more low-level data like computer logs generated by the reading task. With the help of each of these examples, the ML algorithm produces an output (i.e., label), such as "engaged" or "distracted". The number and types of examples the teacher provides the ML algorithm are called *training data*. Through the training data, the system may learn a specific set of features that predicts engagement. Researchers have developed a large array of ML algorithms to handle different constraints and challenges and these methods have been leveraged and adapted within the field of LA to support student modeling, adaptation and recommendation, feedback, and assessments. These algorithms can vary in terms of: (a) the type of program or knowledge the system can learn, (b) how the algorithm represents that information, and (c) the requirements surrounding the training experience. Distinguishing these algorithms in terms of their training experience

Table 1 Examples of research that use similar classes of ML algorithms across different LA applications.

LA challenges		ML approach	Examples
Student modeling	Supervised	Logistic regression models, markov process models, recurrent neural networks	Gervet et al. (2020)
	Unsupervised	K-means, hidden markov models	Sharma et al. (2020)
	Reinforcement learning	Inverse reinforcement learning	Yang et al. (2020)
Adaptation and recommendation	Supervised	Recurrent neural networks	Pardos & Jiang (2020)
	Unsupervised	Pairwise constrained K-means	ALSaad & Alawini (2020)
	Reinforcement learning	Multi-armed bandit algorithms	Li et al. (2020)
Feedback and score reporting	Supervised	Random forest classifiers	Cavalcanti et al. (2020)
	Unsupervised	K-means	Li and Paquette (2020)
	Reinforcement learning	Deep reinforcement learning, deep Q Networks	Ju et al. (2020)

Note: We highlight recent studies from the 2020 Proceedings of the International Conference on EDM and the Learning Analytics and Knowledge (LAK) Conference.

requirements (i.e., supervised, unsupervised, and Reinforcement Learning), we highlight studies that demonstrate how ML is being applied to the core research areas in the field of LA. Across these examples, we highlight the independence between the type of ML algorithm and the challenge targeted within LA, noting that often different types of ML approaches have been used successfully to tackle the same challenge in LA (see Table 1).

Supervised learning

Supervised learning approaches capture a diverse set of algorithms that take feature-label pairs as training examples (i.e., correct answer is given as feedback) and submit these to a program for generating a label for a set of features that are not in the training examples. Supervised classifiers can learn to predict discrete labels (e.g., classification) or continuous outputs (e.g., regression). We previously described a binary classification problem of detecting whether a student was “engaged” or “disengaged”. Supervised learning approaches provide an optimal solution to this problem as these algorithms “learn” different types of and numbers of label predictions based on a dataset where labels are identified by human experts. During this process, it’s possible a human expert identifies someone as not only engaged but also confused. This challenge is met with algorithms that support multiclass classification (e.g., more than two labels) and multilabel classification (e.g., each input can be associated with multiple labels). Regression based classifiers then can learn the function that approximates the mapping between input variables (e.g., engagement and confusion) and some continuous output variable (e.g., learning).

Supervised approaches can vary in terms of the complexity of programs. Deep learning approaches, such as artificial neural networks use complex multilayer networks of nodes that each learn a simple function when trained on feature-label pairs. When unlabeled data is provided to a trained model, the combination of these functions produce a label prediction. These approaches can be especially useful when working with large input features sets (such as pictures and videos). In educational context, Jiang et al. (2018) demonstrated how a neural network-based approach could be used to successfully classify the affect and behaviors of students, comparing a set of neural network-based approaches to regression-based supervised learning approaches. Both types of supervised approaches were trained with labels generated by humans observing a student’s affect during the class session using the BROMP assessment paradigm described earlier (Ocumpaugh et al., 2012, 2015). However, the regression-based approach required human-labeled features whereas the neural network based approach only processed raw computer logs as input. In discussing their results, the authors highlight the trade-off between these two types of supervised classification approaches. Specifically, they point out that although neural networks can derive meaningful features from the raw interaction data, the representation of the relationship between these features and specific student states is too complex to interpret creating a lack of transparency and interpretability. On the other hand, the regression-based approaches required additional steps from humans which can be time consuming and expensive to generate the features; however, the representation that these models learn lends itself to generating psychological or educationally plausible explanations for why certain features are associated with certain student behaviors.

Note that not all supervised algorithms that require a large amount of human effort represent knowledge in a way that is easy to introspect. This directly impacts the interpretability of its outcomes. For example, Decision Trees have a tree structure that organizes labeled data according to a series of sequential decisions; therefore, it is easier to interpret the output of a Decision Tree. However, Decision Trees have some drawbacks, such as efficiency of its algorithms, overfitting, and its predictive power compared to other ML algorithms (Ali et al., 2012). Random Forests (Breiman, 2001), which are an extension of Decision Trees, overcome above-mentioned drawbacks with a tradeoff with the interpretability of their outcome. However, it is possible to increase the interpretability of this type of so-called “blackbox” ML models (i.e., models where it’s difficult to interpret what exactly is occurring because the inputs are unknown). For example, Hutt et al. (2018) used different classifiers (e.g., Random Forest, Gradient Boosted Decision Tree, Logistic Regression) to predict whether students would graduate college based on features from their college applications. The Random Forest classifier outperformed other types of more interpretable classifiers (Gradient Boosted Decision Tree, Logistic Regression) and achieved a predictive accuracy of 71.4%. The authors combined several approaches to

determine a minimal feature set and feature selection (i.e., stochastic hill-climbing approach and logistic regression) to gain insight into what features of a student's application contributed the most to the classifier's performance. Although this type of approach can help to understand the correlation between the features and the outcome variable, the authors concluded that future work combining top-down feature encoding approach together with bottom-up approach could help to illuminate causal relationships beyond correlations.

Unsupervised learning

Unsupervised learning includes a suite of methods for identifying patterns within the input features without the guidance of output labels. These approaches can be used to identify important trends in the input data and generate the output labels for unlabeled data (identification of important features in the log data). These types of approaches are especially useful for problems in which the outcome is difficult to label or represents a probabilistic state (e.g., where there is uncertainty surrounding output labels). Unsupervised clustering approaches learn "latent" (e.g., missing) labels through identifying categories within the set of training objects. These categories reflect structures or patterns present within the data.

For example, [Mostafavi and Barnes \(2017\)](#) used hierarchical clustering, an unsupervised learning method to discover different strategies when students were solving particular problems presented in an intelligent deductive logic tutor. The authors created a data-driven problem profiler, which used the hierarchical clustering of exemplar scores of students who interacted with the tutor in the previous semester to identify different student profiles (i.e., clustering students based on their application of specific Knowledge Components throughout the tutor). This allowed for their system to identify students that use distinct strategies for solving problems within the tutor and adapt the system to provide appropriate support. In addition to providing valuable descriptive information about the student, [Mostafavi and Barnes \(2017\)](#) describe how data-driven clustering can improve how accurately the system could determine student proficiency. This purely data-driven approach was not able to predict student proficiency more accurately than proficiency based on an expert-based approach. However, unlike an expert approach, the data-driven approach could be easily scaled to support greater numbers of students without the needing significant human input.

While this example highlights one application of unsupervised learning to identify groups of students, these approaches can be used across a wide variety of learning contexts to identify patterns across student behavior, student problem solving, and different types of learning content (see [Table 1](#)). The meaningfulness of the output labels these algorithms identify is largely shaped by the data submitted as input labels. For example, when profiling types of collaborative problem solvers there are numerous dimensions that need to be captured within the input data about the students working together and the way in which they interact ([Forsyth et al., 2020](#)). Without considering the corresponding social or cognitive dimensions and constructs when constructing the dataset these approaches are trained on, the unsupervised algorithm may not have the evidence needed to identify meaningful clusters of features for identifying types of collaborative problem solvers. For this reason, unsupervised clustering within LA should be an iterative process where after identification, the identified clusters should be analyzed and compared to hypotheses about the constructs or external assessments that are well-validated to determine if the output of the approach is meaningful. Without "ground-truth" student profile labels to compare the final output to, different resulting clusters can be interpreted using theory and compare resulting clusters to external surveys providing different data from the input to the ML algorithm. As with supervised learning, there is a wide range in the type of knowledge these algorithms can learn including, exclusive cluster membership (e.g., an input is associated with a single label), overlapping clustering (e.g., an input is associated with multiple labels), hierarchical clustering (e.g., labels indicate a position in a hierarchy of labels), probabilistic clustering (e.g., input features are probabilistic and there is uncertainty in how they relate to labels).

Reinforcement learning

Reinforcement learning (RL) is when the algorithm is optimized based on "feedback" from other algorithms or human interaction in the digital environment. In RL, the learning system, often referred to as an "agent", takes actions within the environment and is provided "feedback" from other algorithms or humans. Unlike supervised methods where complete examples are provided with feedback, RL agents can learn *what to do* by exploring the environment, taking actions, and receiving feedback. Take for example learning to play Tic-Tac-Toe. An untrained RL agent will take random actions within the environment. These actions will result in changes to the environment and moves by the opponent, eventually the agent will experience a win or loss. This feedback, referred to as a *reward*, helps the agent build a predictive model of the environment, its opponent, and what moves it should make in different situations. While in this learning example, reward is provided at the end of a series of actions, RL agents can also learn from throughout the learning process. As with other approaches there is diversity in how agents learn, what they can learn, and how they represent that information. Passive learning systems have fixed policies (e.g., strategies) to learn a utility function on the environment states, whereas active learning agents explore an environment and adjust the actions or rules of the algorithm in response. When developing an RL model, the researcher must consider the action space (e.g., what actions the agent can take), the state space (e.g., the representation of the environmental states), the reward function (e.g., what feedback to provide the agent), and the policy optimization (e.g., the algorithmic method for improving the agent).

RL is especially useful when learning programs/knowledge in complex domains where there are a large number of choices available making it difficult for a human to accurately evaluate all of these choices. There has been a lot of recent interest in the LA community in RL algorithms to develop guidelines for adaptation and recommendation especially within Instructional Policy optimization. In this work, RL methods explore a variety of different decisions including the scheduling of educational activities, introduction of new concepts, deciding what feedback to give, and even what class to take next (Bassen et al., 2020; Doroudi et al., 2017; Rafferty et al., 2015). In a recent example of this, Bassen et al. (2020) designed a linear algebra online continuing education course that tested the educational utility of using an RL model to schedule student's activities. The RL agents navigated the action space by choosing between educational activities. Performing these activities would change the state space, what skills the agent was estimated to know (e.g., from test scores and activity performance). Bassen et al. chose a reward function that rewarded test performance improvements and penalized the inclusion of activities covering already mastered skills which sets the ground rules for how the RL algorithm should change based on input. The researchers used data from simulated learners to train the RL agents but also updated their model with actual student data as students began to take the class. Students that were recommended activities using the schedule provided by the RL agents were less likely to drop out of the course than self-directed students and achieved comparable learning gains. While RL is well suited for the complexity of scheduling a diverse set of activities, these methods do require a considerable amount of training data to build an accurate model of the state space and suggest reasonable actions. In their paper, Bassen et al. relied on simulated learners to generate data, even so they acknowledge that it is still unclear if the approach converged. This highlights the importance of having an accurate simulation environment when training RL agents.

Trends and opportunities for the future of learning analytics

Learning Analytics and Machine Learning have been rapidly progressing in development and implementation over the last decades. Given the need for distance learning since the advent of COVID-19, we expect scientists and developers in the field to continue to push the boundaries to better aid student learning and assessment. Specifically, we highlight three areas of exciting growth within the field that use human-in-the-loop approaches: (a) theoretically-grounded educational data mining, (b) computational cognitive modeling, and (c) computational linguistic analysis and artificial agent-based interaction. For a more detailed review on the importance of human-in-the-loop approaches, see also Rosé et al. (2019) and Zapata-Rivera and Arslan (2021).

Theoretically-grounded educational data mining

Theoretically-grounded data mining (Forsyth et al., 2014) is a methodology that combines the principles of Evidence-Centered Design (ECD; Mislevy et al., 2003) with LA/EDM (Baker and Yacef, 2009). ECD is a very careful linking of pieces of evidence gathered from human interaction to higher level skills and subskills corresponding to over-arching cognitive constructs. This principled methodology has been a staple of the assessment community for decades and now more than ever is necessary for establishing the features that provide evidence of student behaviors which can inform researchers of higher-level skills as well as more abstract constructs such as collaborative problem solving. When ECD is implemented in complex digital environments, psychometrics created for traditional items such as multiple-choice items cannot necessarily be implemented appropriately, and more complex predictive modeling embodied in LA is needed. One may ask the question, why is it important to have complex environments with this level of validation required of ECD. Our answer is that we want to provide teachers and students with information that can be captured in environments designed to elicit evidence of cognition and then validated by multiple sources and even provide additional information for personalization that is grounded in theory. Without this level of scrutiny and rigor, we will not be able to substantiate claims for assessment and adaptation in learning environments and therefore are in danger of hitting a commonly scrutinized point of analysis of large log files. This criticism of EDM without ECD is that with so many variables available, spurious correlations are bound to exist (i.e., correlations that are meaningless and simply artifacts of a particular dataset) which can be addressed by combining the two methodologies (see also Goldhammer et al., 2021). After all, every level of adaptation is based on inferences from these pieces of evidence. Multiple methods for implementing theoretically grounded data mining have been proposed and can be explored (Forsyth et al., 2012, 2014, 2020).

Computational cognitive models

Computational cognitive models (CCM) are algorithmic implementations of micro theories of cognition (e.g., theories of learning, or theories of human memory) that represent fixed mechanisms of cognition together with possible free parameters to represent individual differences. Much of the research developing CCMs is based on unified theories of cognition that have been operationalized within cognitive architectures such as ACT-R (Anderson, 2007; Anderson et al., 2004), PRIMs (Taatgen, 2013), Soar (Laird, 2019). Unlike LA approaches that analyze the student log data, CCMs can simulate data (i.e., generate data) to make predictions about learner behavior without collecting data from learners (e.g., Arslan et al., 2017a; MacLellan and Koedinger, 2020). These predictions about learner behavior can then be confirmed using student data as a means of validating and improving the approach. If the model's predictions are not in line with student data, the model can be revised considering the information gathered from the data. Fig. 2 illustrates this research cycle.

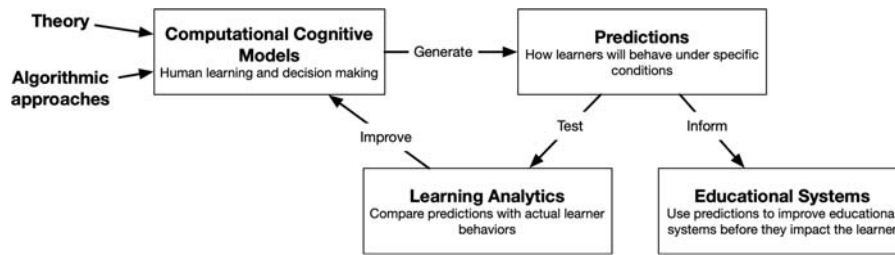


Fig. 2 Proposed research cycle for using computational cognitive models within learning analytics.

While the predictive capabilities of these models have clear educational utility in testing and improving educational systems, they also offer insight into the cognitive mechanisms that underlie the observed behaviors of learners (Arslan et al., 2017b; MacLellan et al., 2016). There is potential for CCMs to contribute to research across the core areas of LA (e.g., student modeling, adaptation and recommendation, and assessment and feedback). Work within the field of cognitive science has illustrated that these CCMs can be fitted to the data to estimate the free parameters at an individual learner level and demonstrated how these free parameters can update models of student proficiency and adapt training methods to improve learning (e.g., Cranford et al., 2020; Walsh et al., 2018a). Furthermore, efforts to build models that simulate human like learning (e.g., *simulated students*) has been shown to have numerous applications within the field of LA (VanLehn et al., 1994) including authoring Intelligent Tutoring Systems at scale (MacLellan and Koedinger, 2020), providing students with opportunities to learn through teaching (Matsuda et al., 2013, 2020), and generating simulation environments to test out different instructional policies (MacLellan, 2017).

Computational linguistic analysis and artificial agent-based interaction

Computational linguistic analysis is a set of tools and procedures that support modeling the students based on the student's words. Computational linguistic analysis can help identify student understanding of material (for more details, see McNamara et al., 2017), affective states (D'Mello and Kory, 2015; Dowell and Graesser, 2014), and identify features that align to social constructs (e.g., Pennebaker, 2011) important for 21st century skills such as collaborative problem solving (e.g., Dowell et al., 2020). This information can provide rich evidence to better understand the students' thinking, motivational and affective states guiding the personalization of adaptive and adaptable components of complex digital learning and assessment environments (for a review, see McNamara et al., 2017).

An important application of computational linguistic analysis as well as theoretically-grounded data mining is the implementation of artificial agent guided learning and assessment environments. The artificial agents can be in the form of talking heads that can tutor students and dynamically adapt to these conversations. This type of work has resulted in effect sizes 0.8–2.0 compared to comparable controls in dozens of studies (for a review, see Graesser, 2016). Furthermore, these types of systems can assess students during Conversation-based Assessment and there are many other Intelligent Tutoring Systems with natural language processing including those that provide recommendations on what to do next in complex environments, which are addressed in more detail in other chapters in this Encyclopedia.

Future opportunities

While above-mentioned areas of research (theoretically-grounded educational data mining, computational cognitive modeling, computational linguistics and artificial agent-based interaction) reflect the current research efforts of the authors, we believe that they have great potential for advancing the field of Learning Analytics. The prominent theme across these areas is the emphasis of theory as the driving force behind LA and the potential refinement of these theories from the insights we draw from educational systems. Theoretically-grounded data mining and computational cognitive modeling attempt to integrate theory with analytics including ML in an iterative fashion to increase the validity of findings and ensure transparency and interpretability via an interdisciplinary approach. Computational linguistics and artificial agent-based interactions focus on understanding human words to interpret the underlying knowledge and state of the human for adaptation based on the human's needs. Agents can add to this endeavor by interacting with students in a manner that mimics a personalized tutor or peer for the purposes of optimal learning. We acknowledge that many more applications of ML and LA may also make great contributions to the field and encourage researchers to explore these areas while considering how these topics can provide a lens through which we can build a better understanding of human learning.

Conclusions

We began this chapter with a vision of a future in which personalized learning affords universal access to the highest quality education. While this future is still distant, research within the field of Learning Analytics has demonstrated a path to this future. In recent years, Machine Learning has featured prominently within the community as a powerful tool through which we can translate data into insights, approaches, and systems for supporting learning and assessment.

We provide a general introduction to three major areas of research within LA and discuss where the application of ML has advanced these research areas. The primer on applications of LA discusses a variety of student modeling or attempts to understand the student's current state or level of knowledge to inform algorithms for personalizing instruction to a student's needs. We discussed modeling based on student knowledge where the algorithms (e.g., Bayesian Knowledge Tracing; for more details, see Pelánek, 2017) adapt based on the student's performance on previous items to inform the computer program to present material that is best suited to the students current understanding of the material. Another type of student modeling discussed is based on the student's behavior and current state. This type of modeling allows for adaptation based on detected affective states or other corresponding states within the learning experience which can increase student learning (D'Mello, 2017). The applications of LA for student modeling include adapting the computerized system for personalized support as well as providing computerized feedback by a computer or arming a teacher with needed information for individual and classroom level interventions.

The primer on ML discusses three common types of modeling including unsupervised learning, supervised learning, and Reinforcement Learning. Unsupervised learning requires minimal human input and allows for the algorithm to determine the human state whereas supervised learning requires human tagging of information for input to the algorithm. Reinforcement Learning algorithms have been used more recently to aid the computer in adjusting algorithms for optimal strategies. These three models are staples within the LA community but it is important to note that ML is not synonymous with LA but rather tools to interpret student state or knowledge and then adapt accordingly for optimal learning.

Our main purpose of this article was to outline how LA and ML can contribute to student learning now more than ever. Given the pandemic and associated equity issues, adaptation comparable to human tutoring is necessary for "closing the gap" in education (Heffernan and Heffernan, 2021). This gap refers to a long list of inequities that have been exacerbated by the pandemic such as differences in socio-economic status that may result in some students having access to internet whereas others do not, access to books, and parents with advanced degrees with time to help their children versus parents that must work outside the home and may or may not have the education to help their children. There are many potential facets, many likely yet to be discovered that have created a myriad of educational experiences for students during this unique time in history including geographical location and school funding situations. This brings us to one final potential use of LA which is to potentially identify students based on groups of features affecting post-pandemic education and determine the best intervention(s) to help affected students. Furthermore, LA often with carefully applied and interpretable ML techniques has laid the foundation for students to receive personalized learning experiences that adapt based on student's knowledge and states. Thus, appropriate applications of learning analytics may be more important now than ever before in history.

References

- Aleven, V., McLaren, B., Roll, I., Koedinger, K., 2006. Toward meta-cognitive tutoring: a model of help seeking with a Cognitive Tutor. *Int. J. Artif. Intell. Educ.* 16 (2), 101–128.
- Aleven, V., Sewall, J., Popescu, O., Xhakaj, F., Chand, D., Baker, R., Wang, Y., Siemens, G., Rosé, C., Gasevic, D., 2015. The beginning of a beautiful friendship? Intelligent tutoring systems and MOOCs. *Int. Conf. Arti. Intell. Edu.* 525–528.
- Ali, J., Khan, R., Ahmad, N., Maqsood, I., 2012. Random forests and decision trees. *Int. J. Comput. Sci. Iss.* 9 (5), 272.
- AlSaad, F., Alawini, A., 2020. Unsupervised approach for modeling content structures of MOOCs. In: *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)*, pp. 18–28.
- Anderson, J.R., 2007. *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.
- Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y., 2004. An integrated theory of mind. *Psychol. Rev.* 111, 1036–1060.
- Andrews-Todd, J., Forsyth, C.M., 2020. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Comput. Hum. Behav.* 104, 105759.
- Arslan, B., Taatgen, N.A., Verbrugge, R., 2017a. Five-year-olds' systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: a computational modeling study. *Front. Psychol.* 8 (275). <https://doi.org/10.3389/fpsyg.2017.00275>.
- Arslan, B., Verbrugge, R., Taatgen, N., 2017b. Cognitive control explains the mutual transfer between dimensional change card sorting and first-order false belief understanding: a computational modeling study on transfer of skills. *Biol. Inspired Cogn. Archit.* 20, 10–20.
- Baker, R.S., Yacef, K., 2009. The state of educational data mining in 2009: a review and future visions. *J. Educ. Data Min.* 1 (1), 3–17.
- Baker, R., Siemens, G., 2014. Educational data mining and learning analytics. In: Sawyer, R. (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology). Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.016>.
- Baker, R.S., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C., 2010. Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum. Comput. Stud.* 68 (4), 223–241.
- Barshay, J., 2021. PROOF POINTS: Free, No Frills Programs Lead the Class in New Federal Study of Remote Learning. *The Hechinger Report*.
- Bassen, J., Balaji, B., Schaarschmidt, M., Thille, C., Painter, J., Zimmario, D., Games, A., Fast, E., Mitchell, J.C., 2020. Reinforcement learning for the adaptive scheduling of educational activities. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–12.
- Biswas, G., Jeong, H., Kinnebrew, J.S., Sulcer, B., Roscoe, R.O.D., 2010. Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Res. Pract. Technol. Enhanc. Learn.* 5 (02), 123–152.
- Bloom, B.S., 1984. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educ. Res.* 13 (6), 4–16.
- Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., Verbert, K., 2018. Open learner models and learning analytics dashboards: a systematic review. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 41–50.
- Botelho, A.F., Baker, R.S., Heffernan, N.T., 2017. Improving sensor-free affect detection using deep learning. *Int. Con. Arti. Intell. Educ.* 40–51.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
- Brown, J.S., VanLehn, K., 1982. Towards a generative theory of "bugs". *Addit Subtract Cognit Perspect* 117–135.
- Cavalcanti, A.P., Diego, A., Mello, R.F., Mangaroska, K., Nascimento, A., Freitas, F., Gašević, D., 2020. How good is my feedback? A content analysis of written feedback. *Proceed. Tenth Int. Conf. Learn. Anal. Knowl.* 428–437.
- Cen, H., Koedinger, K., Junker, B., 2008. Comparing two IRT models for conjunctive skills. *Int. Conf. Intell. Tutoring Syst.* 796–798.
- Cen, H., Koedinger, K., Junker, B., 2006. Learning factors analysis—a general method for cognitive model evaluation and improvement. *Int. Conf. Intell. Tutoring Syst.* 164–175.

- Gen, H., Koedinger, K.R., Junker, B., 2007. Is over practice necessary?-improving learning efficiency with the cognitive tutor through educational data mining. *Front. Artif. Intell. Appl.* 158, 511.
- Chrysafiadi, K., Virvou, M., 2013. Student modeling approaches: a literature review for the last decade. *Expert Syst. Appl.* 40 (11), 4715–4729.
- Corbett, A.T., Anderson, J.R., 1994. Knowledge tracing: modeling the acquisition of procedural knowledge. *User Model. User-Adapted Interact.* 4 (4), 253–278.
- Cranford, E.A., Gonzalez, C., Aggarwal, P., Cooney, S., Tambe, M., Lebiere, C., 2020. Toward personalized deceptive signaling for cyber defense using cognitive models. *Top. Cognit. Sci.* 12 (3), 992–1011.
- D'Mello, S., Graesser, A., 2014. Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychol.* 151, 106–116.
- D'Mello, S., Kory, J., 2015. A review and meta-analysis of multimodal affect detection systems. *ACM Comput. Surv.* 47 (3), 1–36.
- D'Mello, S., 2017. Emotional learning analytics. *Handbook Learn. Anal.* 115.
- D'Mello, S., Graesser, A., 2012. Emotions during learning with AutoTutor. In: Durlach, P., Lesgold, A. (Eds.), *Adaptive Technologies for Training and Education*. Cambridge Univ. Press, pp. 117–139.
- D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., El Kaliouby, R., Picard, R., 2008. AutoTutor detects and responds to learners affective and cognitive states. In: *Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems*, pp. 306–308.
- Doroudi, S., Aleven, V., Brunskill, E., 2017. Robust evaluation matrix: towards a more principled offline exploration of instructional policies. In: *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*, pp. 3–12.
- Doroudi, S., Brunskill, E., 2019. Fairer but not fair enough on the equitability of knowledge tracing. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, pp. 335–339.
- Dowell, N., Lin, Y., Godfrey, A., Brooks, C., 2020. Exploring the relationship between emergent sociocognitive roles, collaborative problem-solving skills, and outcomes: a group communication analysis. *J. Learn. Anal.* 7 (1), 38–57.
- Dowell, N., Graesser, A.C., 2014. Modeling learners' cognitive, affective, and social processes through language and discourse. *J. Learn. Anal.* 1 (3), 183–186.
- Durand, G., Goutte, C., Belacel, N., Bouslimani, Y., Leger, S., 2017. Review, computation and application of the additive factor model (AFM). *Nat. Res. Council Canada, Winnipeg, MB, Canada, Tech. Rep.* 23002483.
- Eglington, L.G., Pavlik, P.I., 2020. Optimizing practice scheduling requires quantitative tracking of individual item performance. *Npj Sci. Learn.* 5 (1), 15. <https://doi.org/10.1038/s41539-020-00074-4>.
- Ekman, P., Friesen, W., 1978. *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press.
- Forsyth, C., Andrews-Todd, J., Steinberg, J., 2020. Are You Really a Team Player? Profiling of Collaborative Problem Solvers in an Online Environment. *International Educational Data Mining Society*.
- Forsyth, C., Graesser, A., Pavlik Jr., P.I., Millis, K., Samei, B., 2014. Discovering theoretically grounded predictors of shallow vs. deep-level learning. In: Stamper, J., Pardos, Z., Mavrikis, M., McLaren, B.M. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*. International Educational Data Mining Society, pp. 229–232.
- Forsyth, C.M., Peters, S., Zapata-Rivera, J.D., Lentini, J., Graesser, A., Cai, Z., 2017. Interactive score reporting: an AutoTutor-based system for teachers. *Int. Conf. Arti. Intell. Educ.* 506–509.
- Forsyth, C., Pavlik Jr., P., Graesser, A.C., Cai, Z., Germany, M., Millis, K., Dolan, R.P., Butler, H., Halpern, D., 2012. Learning gains for core concepts in a serious game on scientific reasoning. In: *Proceedings of the 5th International Conference on Educational Data Mining*, pp. 172–175.
- Gervet, T., Koedinger, K., Schneider, J., Mitchell, T., 2020. When is deep learning the best approach to knowledge tracing? *J. Educ. Data Min.* 12 (3), 31–54.
- Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J.K., Coles, R., Graesser, A.C., Sullins, J., Craig, S.D., 2009. Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instr. Sci.* 37 (5), 487–493.
- Goutte, C., Durand, G., 2020. Confident Learning Curves in Additive Factors Modeling. *International Educational Data Mining Society*.
- Graesser, A.C., 2016. Conversations with AutoTutor help students learn. *Int. J. Artif. Intell. Educ.* 26, 124–132.
- Graesser, A.C., Conley, M.W., Olney, A., 2012a. Intelligent tutoring systems. In: Graham, S., Harris, K. (Eds.), *APA Educational Psychology Handbook, Vol 3: Application to Learning and Teaching*. American Psychological Association, pp. 451–473.
- Graesser, A.C., D'Mello, S., Hu, X., Cai, Z., Olney, A., Morgan, B., 2012b. AutoTutor. In: *Applied Natural Language Processing: Identification, Investigation and Resolution*. IGI Global, pp. 169–187.
- Graesser, A.C., Li, H., Forsyth, C., 2014. Learning by communicating in natural language with conversational agents. *Curr. Dir. Psychol. Sci.* 23 (5), 374–380.
- Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H.H., Ventura, M., Olney, A., Louwerse, M.M., 2004. AutoTutor: a tutor with dialogue in natural language. *Behav. Res. Methods Instrum. Comput.* 36 (2), 180–192.
- Graesser, A.C., VanLehn, K., Rosé, C.P., Jordan, P.W., Harter, D., 2001. Intelligent tutoring systems with conversational dialogue. *AI Mag.* 22 (4), 39.
- Goldhammer, F., Hahnel, C., Kroehne, U., Zehner, F., 2021. From byproduct to design factor: on validating the interpretation of process indicators based on log data. *Large-scale Assess. Educ.* 9 (1), 1–25.
- Heffernan, N., Heffernan, C., 2021. Commentary: Short Tutoring Sessions on Demand, Chatbots to Help with Homework Problems — Innovations that Could Get Students Back on Track.
- Heffernan, N.T., Heffernan, C.L., 2014. The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *Int. J. Artif. Intell. Educ.* 24 (4), 470–497.
- Hernández-Blanco, A., Herrera-Flores, B., Tomás, D., Navarro-Colorado, B., 2019. A systematic review of deep learning approaches to educational data mining. *Complexity* 2019.
- Hobert, S., Meyer von Wolff, R., 2019. Say hello to your new automated tutor – a structured literature review on pedagogical conversational agents. In: *14th International Conference on Wirtschaftsinformatik*, pp. 301–314.
- Hu, X., Graesser, A.C., 2004. Human use regulatory affairs advisor (HURAA): learning about research ethics with intelligent learning modules. *Behav. Res. Methods Instrum. Comput.* 36 (2), 241–249.
- Hutt, S., Gardener, M., Kamentz, D., Duckworth, A.L., D'Mello, S.K., 2018. Prospectively predicting 4-year college graduation from student applications. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 280–289.
- Iglesias, A., Martínez, P., Aler, R., Fernández, F., 2009. Reinforcement learning of pedagogical policies in adaptive and intelligent educational systems. *Knowl. Base Syst.* 22 (4), 266–270.
- Jackson, T., Zapata-Rivera, D., 2015. Conversation-based assessment. *R&D Connec.* 25, 1–8.
- Jiang, Y., Bosch, N., Baker, R.S., Paquette, L., Ocumpaugh, J., Andres, J.M.A.L., Moore, A.L., Biswas, G., 2018. Expert feature-engineering vs. deep neural networks: which is better for sensor-free affect detection? *Int. Conf. Arti. Intell. Educ.* 198–211.
- Ju, S., Zhou, G., Barnes, T., Chi, M., 2020. Pick the Moment: Identifying Critical Pedagogical Decisions Using Long-Short Term Rewards. *International Educational Data Mining Society*.
- Jurafsky, D., Martin, J., 2008. *Speech and Language Processing*. Prentice Hall.
- Kannan, P., Zapata-Rivera, J.D., Leibowitz, E.A., 2018a. Interpretation of score reports by diverse subgroups of parents. *Educ. Assess.* 23 (3), 173–194.
- Kannan, P., Zapata-Rivera, J.D., Mikeska, J., Bryant, A., Long, R., Howell, H., 2018b. Providing formative feedback to pre-service teachers as they practice facilitation of high-quality discussions in simulated mathematics and science methods classrooms. In: *Society for Information Technology & Teacher Education International Conference*, pp. 1570–1575.
- Khalid, A., Lundqvist, K., Yates, A., 2020. Recommender systems for moocs: a systematic literature survey (january 1, 2012–july 12, 2019). *Int. Rev. Res. Open Dist. Learn.* 21 (4), 255–291.

- Khalil, M., Ebner, M., 2015. Learning Analytics: Principles and Constraints. EdMedia+ Innovate Learning, pp. 1789–1799.
- Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.A., 1997. Intelligent tutoring goes to school in the big city. *Int. J. Artif. Intell. Educ.* 8 (1), 30–43.
- Koedinger, K.R., Brunskill, E., Baker, R.S., McLaughlin, E.A., Stamper, J., 2013. New potentials for data-driven intelligent tutoring system development and optimization. *AI Mag.* 34 (3), 27–41.
- Koedinger, K.R., Corbett, A., 2006. Cognitive tutors: technology bringing learning sciences to the classroom. In: Sawyer, R.K. (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, pp. 61–78.
- Laird, J.E., 2019. *The Soar Cognitive Architecture*. MIT press.
- LAK 2011 : 1st International Conference Learning Analytics and Knowledge, 2011.
- Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W., 2013. *Handbook of Latent Semantic Analysis*. Psychology Press.
- Li, T.W., Paquette, L., 2020. Erroneous Answers Categorization for Sketching Questions in Spatial Visualization Training. *International Educational Data Mining Society*.
- Li, Z., Yee, L., Sauerberg, N., Sakson, I., Williams, J.J., Rafferty, A.N., 2020. Getting Too Personal (ized): The Importance of Feature Choice in Online Adaptive Algorithms. *International Educational Data Mining Society*.
- Long, Y., Holstein, K., Aleven, V., 2018. What exactly do students learn when they practice equation solving? Refining knowledge components with the Additive Factors Model. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 399–408.
- Lopez, A.A., Guzman-Orth, D., Zapata-Rivera, D., Forsyth, C.M., Luce, C., 2021. Examining the Accuracy of a Conversation-Based Assessment in Interpreting English Learners' Written Responses. *ETS Research Report Series*.
- Lopez, A.A., Luce, C., Zapata-Rivera, D., Forsyth, C., 2017. Using formative conversation-based assessments to support students' English language development. *IEEE Tech. Committ. Learn. Technol. Bull.* 19 (1), 6–9.
- MacLellan, C.J., Harpstead, E., Patel, R., Koedinger, K.R., 2016. The Apprentice Learner Architecture: Closing the Loop between Learning Theory and Educational Data. *International Educational Data Mining Society*.
- MacLellan, C.J., Koedinger, K.R., 2020. Domain-general tutor authoring with apprentice learner models. *Int. J. Artif. Intell. Educ.* 1–42.
- MacLellan, C.J., Liu, R., Koedinger, K.R., 2015. Accounting for Slipping and Other False Negatives in Logistic Models of Student Learning. *International Educational Data Mining Society*.
- MacLellan, C.J., 2017. *Computational Models of Human Learning: Applications for Tutor Development, Behavior Prediction, and Theory Testing*. Carnegie Mellon University.
- Matsuda, N., Weng, W., Wall, N., 2020. The effect of metacognitive scaffolding for learning by teaching a teachable agent. *Int. J. Artif. Intell. Educ.* 1–37.
- Matsuda, N., Yarbuzinski, E., Keiser, V., Raizada, R., Cohen, W.W., Stylianides, G.J., Koedinger, K.R., 2013. Cognitive anatomy of tutor learning: lessons learned with SimStudent. *J. Educ. Psychol.* 105 (4), 1152–1163. <https://doi.org/10.1037/a0031955>. American Psychological Association.
- McNamara, D.S., Allen, L., Crossley, S., Dascalu, M., Perret, C.A., 2017. Natural language processing and learning analytics. In: Siemens, G., Lang, C. (Eds.), *Handbook of Learning Analytics*. Society for Learning Analytics Research, pp. 93–104.
- McNamara, D.S., O'Reilly, T.P., Best, R.M., Ozuru, Y., 2006. Improving adolescent students' reading comprehension with iSTART. *J. Educ. Comput. Res.* 34 (2), 147–171.
- Meng, L., Zhang, M., Zhang, W., Chu, Y., 2019. CS-BKT: introducing item relationship to the Bayesian knowledge tracing model. *Interact. Learn. Environ.* 1–11.
- Mertens, U., Finn, B., Linder, M., 2021. Effects of computer-based feedback on learning and motivation: a network meta-analysis. *Manuscr. Under Rev.*
- Mislevy, R.J., Almond, R.G., Lukas, J.F., 2003. A brief introduction to evidence-centered design. *ETS Res. Rep. Series* 2003 (1), 1–29.
- Mostafavi, B., Barnes, T., 2017. Evolution of an intelligent deductive logic tutor using data-driven elements. *Int. J. Artif. Intell. Educ.* 27 (1), 5–36.
- Mousavinasab, E., Zarifsanaiy, N., Niakan Kalhori, S.R., Rakhshan, M., Keikha, L., Ghazi Saedi, M., 2021. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interact. Learn. Environ.* 29 (1), 142–163.
- Murphy, R., Roschelle, J., Feng, M., Mason, C.A., 2020. Investigating efficacy, moderators and mediators for an online mathematics homework intervention. *J. Res. Educ. Effectiv.* 13 (2), 235–270.
- Ocuppaugh, J., Baker, R.S., Rodrigo, M.M.T., 2012. Baker-Rodrigo Observation Method Protocol (BROMP) Training Manual Version 1.0.
- Ocuppaugh, J., Baker, R.S., Rodrigo, M.M.T., 2015. Baker Rodrigo Ocuppaugh Monitoring Protocol (BROMP) 2.0 Technical and Training Manual. Technical Report. Teachers College, Columbia University, New York, NY.
- Olney, A.M., Graesser, A.C., Person, N.K., 2012. Question generation from concept maps. *Dial. Discour.* 3 (2), 75–99.
- Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., Mirriahi, N., 2019. Using learning analytics to scale the provision of personalised feedback. *Br. J. Educ. Technol.* 50 (1), 128–138.
- Pardos, Z.A., Jiang, W., 2020. Designing for serendipity in a university course recommendation system. In: *Proceedings of the 10th International Conference on Learning Analytics & Knowledge*, pp. 350–359.
- Pardos, Z.A., Dadu, A., 2018. dAFM: fusing psychometric and connectionist modeling for Q-matrix refinement. *J. Educ. Data Min.* 10 (2), 1–27.
- Pardos, Z.A., Heffernan, N.T., 2011. KT-IDEM: introducing item difficulty to the knowledge tracing model. *Int. Conf. User Model. Adap. Personaliz.* 243–254.
- Pavlik, P.I., Anderson, J.R., 2008. Using a model to compute the optimal schedule of practice. *J. Exp. Psychol. Appl.* 14 (2), 101.
- Pelánek, R., 2017. Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. *User Model. User-Adapted Interact.* 27 (3), 313–350.
- Pelánek, R., Rihák, J., 2018. Analysis and design of mastery learning criteria. *New Rev. Hypermedia Multimedia* 24 (3), 133–159.
- Pennebaker, J.W., 2011. The secret life of pronouns. *N. Sci.* 211 (2828), 42–45.
- Piech, C., Spencer, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., Sohl-Dickstein, J., 2015. Deep knowledge tracing. *ArXiv Preprint ArXiv:1506.05908*.
- Rafferty, A.N., LaMar, M.M., Griffiths, T.L., 2015. Inferring learners' knowledge from their actions. *Cognit. Sci.* 39 (3), 584–618.
- Ritter, S., Yudelson, M., Fancsali, S.E., Berman, S.R., 2016. How mastery learning works at scale. In: *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*, pp. 71–79.
- Ritter, S., Carlson, R., Sandbothe, M., Fancsali, S.E., 2015. *Carnegie Learning's Adaptive Learning Products*. Educational Data Mining.
- Romero, C., Ventura, S., 2020. Educational data mining and learning analytics: an updated survey. *Wiley Interdisciplin. Rev. Data Min. Knowl. Discov.* 10 (3), e1355.
- Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R., 2019. Explanatory learner models: why machine learning (alone) is not the answer. *Br. J. Educ. Technol.* 50 (6), 2943–2958.
- Rus, V., D'Mello, S., Hu, X., Graesser, A., 2013. Recent advances in conversational intelligent tutoring systems. *AI Mag.* 34 (3), 42–54.
- Sahni, S.D., Polarin, J.R., Zhang, Q., Michaelson, L.E., Caverly, S., Polese, M.L., Yang, J., 2021. A what works clearinghouse rapid evidence review of distance learning programs. WWC 2021-005REV. In: *What Works Clearinghouse*. ERIC.
- Sales, A.C., Pane, J.F., 2019. The role of mastery learning in an intelligent tutoring system: principal stratification on a latent variable. *Ann. Appl. Stat.* 13 (1), 420–443.
- Salloum, S.A., Alshurideh, M., Elnagar, A., Shalan, K., 2020. Mining in educational data: review and future directions. *Joint Europ.-US Workshop Appl. Invarian. Comput. Vision* 92–102.
- Samuel, A.L., 1959. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* 3 (3), 210–229.
- Sharma, K., Papamitsiou, Z., Olsen, J.K., Giannakos, M., 2020. Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 480–489.
- Shute, V.J., 2008. Focus on formative feedback. *Rev. Educ. Res.* 78 (1), 153–189.
- Shute, V., Ventura, M., 2013. *Stealth Assessment: Measuring and Supporting Learning in Video Games*. The MIT press.
- Taatgen, N.A., 2013. The nature and transfer of cognitive skills. *Psychol. Rev.* 120 (3), 439–471.
- Tacoma, S., Drijvers, P., Jeurig, J., 2021. Combined inner and outer loop feedback in an intelligent tutoring system for statistics in higher education. *J. Comput. Assist. Learn.* 37 (2), 319–332.

- Tarus, J.K., Niu, Z., Mustafa, G., 2018. Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. *Artif. Intell. Rev.* 50 (1), 21–48.
- Thai-Nghe, N., Horváth, T., Schmidt-Thieme, L., 2010. Factorization Models for Forecasting Student Performance. *Educational Data Mining* 2011.
- VanLehn, K., 2006. The behavior of tutoring systems. *Int. J. Artif. Intell. Educ.* 16 (3), 227–265.
- VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* 46 (4), 197–221.
- VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., Rosé, C.P., 2007. When are tutorial dialogues more effective than reading? *Cognit. Sci.* 31 (1), 3–62.
- VanLehn, K., Ohlsson, S., Nason, R., 1994. Applications of simulated students: an exploration. *J. Artif. Intell. Educ.* 5, 135.
- Walsh, M.M., Gluck, K.A., Gunzelmann, G., Jastrzembski, T., Krusmark, M., 2018a. Evaluating the theoretic adequacy and applied potential of computational models of the spacing effect. *Cognit. Sci.* 42, 644–691.
- Walsh, M.M., Gluck, K.A., Gunzelmann, G., Jastrzembski, T., Krusmark, M., Myung, J.I., Pitt, M.A., Zhou, R., 2018b. Mechanisms underlying the spacing effect in learning: a comparison of three computational models. *J. Exp. Psychol. Gen.* 147 (9), 1325.
- Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., Weston, T., 2013. My science tutor: a conversational multimedia virtual tutor. *J. Educ. Psychol.* 105 (4), 1115.
- Webb, G.I., Pazzani, M.J., Billsus, D., 2001. Machine learning for user modeling. *User Model. User-Adapted Interact.* 11 (1), 19–29.
- Wilson, K.H., Xiong, X., Khajah, M., Lindsey, R.V., Zhao, S., Karklin, Y., Van Inwegen, E.G., Han, B., Ekanadham, C., Beck, J.E., 2016. Estimating student proficiency: deep learning is not the panacea. In: *Neural Information Processing Systems. Workshop on Machine Learning for Education*, p. 3.
- Xiong, X., Zhao, S., Van Inwegen, E.G., Beck, J.E., 2016. Going Deeper with Deep Knowledge Tracing. *International Educational Data Mining Society*.
- Yang, X., Zhou, G., Taub, M., Azevedo, R., Chi, M., 2020. Student Subtyping via EM-Inverse Reinforcement Learning. *International Educational Data Mining Society*.
- Yudelson, M.V., Koedinger, K.R., Gordon, G.J., 2013. Individualized Bayesian knowledge tracing models. In: *International Conference on Artificial Intelligence in Education*, pp. 171–180.
- Zapata-Rivera, J.D., Arslan, B., July 2021. Enhancing personalization by integrating top-down and bottom-up approaches to learner modeling. In: *International Conference on Human-Computer Interaction*. Springer, Cham, pp. 234–246.
- Zapata-Rivera, J.D., Kannan, P., Forsyth, C., Peters, S., Bryant, A.D., Guo, E., Long, R., 2018. Designing and evaluating reporting systems in the context of new assessments. *Int. Conf. Augmen. Cogni.* 143–153.
- Zapata-Rivera, J.D., Katz, I.R., 2014. Keeping your audience in mind: applying audience analysis to the design of interactive score reports. *Assess Educ. Princ. Pol. Pract.* 21 (4), 442–463.

Interdisciplinary frontiers: computer-based process data analysis in educational measurement

Björn Nicolay^a, Florian Krieger^b, and Samuel Greiff^a, ^a University of Luxembourg, School of Humanities, Education and Social Sciences, Department of Behavioral and Cognitive Sciences, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, Esch-sur-Alzette, Luxembourg; and ^b Technical University Dortmund, Department of Rehabilitation Sciences, Dortmund, Germany

© 2023 Elsevier Ltd. All rights reserved.

Facets of computer-based process data analysis in educational measurement featured in this article	417
The role of computer-based process data in educational measurement	418
History of computer-based process data analysis in educational measurement	419
Public computer-based process data repositories	419
The current state of computer-based process data analysis	420
Psychometric models	420
Formative assessment	421
Summative assessment	422
Applied computer-based process data analysis in educational measurement: the example of complex problem solving	423
Summary, challenges, and future directions of computer-based process data analysis in educational measurement	424
Acknowledgments	425
References	426

In the overarching realm of educational measurement, the collection and analysis of data in order to gather knowledge about students' skill sets and their learning progress has always been of focal interest for educational stakeholders (OECD, 2013, 2017). This is particularly the case since such data have several implications that can shape students' future educational trajectories. For instance, a student's grade point average (GPA) to a large extent determines whether they are allowed to advance in their educational journey or are required to repeat a class or even an entire grade (e.g., Huebener and Marcus, 2017).

However, teachers, parents, and students themselves are not only interested in data indicating a passing or failing score in a particular subject or the GPA. Instead, data on factors that contribute to one's chances of meeting—or falling short of meeting—the standards required to achieve educational progress carry comparatively more fine-grained and thus crucially valuable information. Such data may include the time spent on individual learning exercises or the number of interactions made when working on a computerized training program focusing on the acquisition or refinement a specific skill (i.e., computer-based process data).

Such computer-based process data include all information collected during a student's interaction with a given computer program, ranging from time spent and number of clicks made on a particular page to the number and content of questions asked to a virtual tutor in a computerized learning simulation (Goldhammer et al., 2017). We will now provide an overview of the contents of this article, which, in brief, centers on the characteristics, history, contemporary use of process data analysis in educational measurement, a concrete example, and future challenges and directions in order to advance educational policy and practice.

Facets of computer-based process data analysis in educational measurement featured in this article

In this article, we first address the question of what computer-based process data are and illustrate their relevance for the educational measurement context (Section **The role of computer-based process data in educational measurement**). Thereafter, in **History of computer-based process data analysis in educational measurement** section, we present the historical evolution of computer-based process data analysis in educational measurement from its inception to its widespread use today. Subsequently, the availability of public computer-based process data repositories and the importance of computer-based process data in large-scale educational assessments will be described (Section **Public computer-based process data repositories**). In **The current state of computer-based process data analysis** section, we discuss the current state of computer-based process data analysis in educational measurement in greater depth with regard to three central facets: (4a) psychometric models; (4b) formative assessment; and (4c) summative assessment. Psychometric models are included as they serve as the statistical backbone for the thorough and valuable analysis and interpretation of computer-based process data in the educational context (particularly with regard to valid and reliable student assessment; DiCerbo et al., 2017; Mislevy et al., 2012). Meanwhile, formative and summative assessment represent today's two most widely used assessment techniques for monitoring students' learning progress, and have at the same time become a hallmark

of evaluating and advancing educational practice (Gallardo, 2021; Harlen, 2011). Thereafter, in **Applied computer-based process data analysis in educational measurement: the example of complex problem solving** section, we provide an in-depth showcase of applied computer-based process data analysis in educational measurement by referring to the skill complex problem solving, which has been shown to be a key competence for students' educational success and beyond (Mainert et al., 2015; Schweizer et al., 2013). The article concludes with a summary of the relevant aspects of educational computer-based process data analysis and some final thoughts on current challenges and potential future directions for computer-based process data analysis in the educational measurement context (Section **Summary, challenges, and future directions of computer-based process data analysis in educational measurement**).

Notably, this article aims to present a thorough overview and theoretical background of process data analysis in educational measurement, and to provide indications for interested researchers and students on exemplary deeply vetted resources facilitating individual process data analysis endeavors. However, due to the vast amount of theory, practical aspects, and considerations involved in computer-based process data analysis, it is unrealistic to assume that this article provides information on each and every nuance of process data analysis in educational measurement. Therefore, we will at times point the interested reader to further in-depth scholarly sources, such as journal articles or books, about specific phenomena in educational process data analysis. Where applicable, these suggestions are given at the end of each respective section.

The role of computer-based process data in educational measurement

Computer-based process data are automatically stored in computerized log files (Zoanetti and Griffin, 2017) that contain raw information on each action a student performs with a software (e.g., a task or an exercise; see Fig. 1). Compared to other well-established types of process data in educational measurement research, such as think-aloud protocols or eye tracking (Branch, 2001; Horstmann et al., 2009), computer-generated process data are generally collected in a less "invasive" way while a student performs different actions on a computer; that means, the student is typically not explicitly aware of being monitored (Adams et al., 2015). In this regard, the generation and collection of such computer-based process data in digital gamified learning environments (i.e., educational game software) has also been termed "stealth assessment" (Shute, 2011; Shute et al., 2016b). The ability to automatically log computer-based process data in such a non-interfering way is a particularly valuable asset, since students' subjective reports of their perceived skill levels can be susceptible to bias and inaccuracy, as they put additional demands on the student and could interfere with actual processing of the task (e.g., Brown et al., 2015).

The thorough scrutiny of the vast amount of computer-based process data inherent in such log files (i.e., computer-based process data analysis) is typically referred to as data mining, which comprises various steps in order to arrive at data that is ready to be interpreted (Aldowah et al., 2019; Ali, 2013). As this process is generally complex and requires advanced methods, time and computer programming skills, it is usually performed by (semi-) automated programs provided by experts in the field (Gobert et al., 2013; Romero and Ventura, 2007, 2020; Slater et al., 2017). Importantly, these programs are tailored to produce the pieces of information desired for a given purpose, such as universities looking into reasons why students drop out of a given study program or teachers trying to uncover the reasons for their students' poor test results (Hsia et al., 2008; Mostow et al., 2005).

```

<logEntry xsi:type="cbaloggingmodel:LogEntryTimeStamp" timeStamp="2016-03-21T08:58:56.302+0100">
  <logEntry xsi:type="cbaloggingmodel:MicroDynButtonPressLogEntry" button="Execute" phase="control">
    <variable name="Pelikortit" userDefinedId="EndoA" value="22"/>
    <variable name="Pisteet" userDefinedId="EndoC" value="22"/>
    <variable name="Pelinappulat" userDefinedId="EndoB" value="22"/>
    <variable name="Vihreät pelimerkit" userDefinedId="ExoB" value="1"/>
    <variable name="Siniset pelimerkit" userDefinedId="ExoA" value="1"/>
    <variable name="Punaiset pelimerkit" userDefinedId="ExoC" value="1"/>
  </logEntry>
</logEntry>
<logEntry xsi:type="cbaloggingmodel:LogEntryTimeStamp" timeStamp="2016-03-21T08:58:57.128+0100">
  <logEntry xsi:type="cbaloggingmodel:MicroDynButtonPressLogEntry" button="Execute" phase="control">
    <variable name="Pelikortit" userDefinedId="EndoA" value="24"/>
    <variable name="Pisteet" userDefinedId="EndoC" value="24"/>
    <variable name="Pelinappulat" userDefinedId="EndoB" value="24"/>
    <variable name="Vihreät pelimerkit" userDefinedId="ExoB" value="1"/>
    <variable name="Siniset pelimerkit" userDefinedId="ExoA" value="1"/>
    <variable name="Punaiset pelimerkit" userDefinedId="ExoC" value="1"/>
  </logEntry>
</logEntry>

```

Fig. 1 Part of a raw log file in .xml format containing process data depicting the actions applied by a student during a complex problem solving task (see also **Summary, challenges, and future directions of computer-based process data analysis in educational measurement** section and Fig. 2 of this article) in a computer-based large-scale assessment in Finland. Based on Vainikainen (2014).

History of computer-based process data analysis in educational measurement

The computerized administration of not only tests but also classroom learning exercises has become increasingly common since the late stages of the 20th century (Bunderson et al., 1989). This trend continues to be observed today (Danniels et al., 2020). Accordingly, the mid-1990s mark the inception of the first research articles referring to the use of computer-based process data to evaluate and improve policies in educational institutions (for an overview on data mining in the educational context between 1995 and 2005 see Romero and Ventura, 2007). One of the first articles explicitly using the term “computer-based process data,” albeit in an organizational context, was published by Langley (1999). Soon thereafter, further researchers began scrutinizing computer-based process data from different types of log files for educational purposes (Srivastava et al., 2000), leading to rapid growth of the field. One of the earliest published computer-based process data analyses presents algorithms used to thoroughly map a variety of student characteristics (e.g., courses taken and credits earned in different disciplines) in order to more efficiently allocate staff and resources in educational institutions (Luan, 2002). Other pioneering studies address the leveraging of computer-based process data from digital learning and communication platforms for uploading presentation slides and exercises, submitting reports, and directly asking questions to peers and teachers (Mazza and Milani, 2005; Mostow, 2004; Paulsen, 2003).

The aforementioned selection of different ways to investigate and use computer-based process data for educational purposes further expanded over the first decade of the 21st century. During this period, studies began to focus on particular benefits for both students and educators, such as improving communication and adapting curricula by clustering and classifying students into different groups, for example (Romero and Ventura, 2007). A comprehensive overview of the most influential papers in the early years of computer-based process data research can be found in Baker and Yacef (2009).

At around the same time, the educational data mining sub-field known as process mining emerged, which focuses on the analysis computer-based process data (i.e., log files) with the aim of improving existing processes, or skills, in a variety of domains, from workers’ deadline management in organizations to how people book vacation lodging online (Van der Aalst et al., 2010). While initially primarily targeted at business process models (Van der Aalst, 2012), process mining was rapidly extended to the educational domain. Existing research in this vein draws upon computer-based process data collected from students watching different educational video sequences, or from their interactions with digital learning platforms, in order to adapt said videos and platforms accordingly (Bogarín et al., 2014; Van der Aalst et al., 2015). Similar research has been conducted to increase the appeal of computer-based training programs based on inferences about people’s motivation to take part in such programs drawn from computer-based process data (Cairns et al., 2015). Moreover, contemporary research investigating computer-based process data has targeted the evaluation and improvement of a variety of additional educational measurement facets. As such, recent research articles examine interactive learning environments, different assessment approaches, and individual students’ skills that have been shown to be particularly related to educational success, and not seldom make use of publicly available computer-based process data repositories for their analyses (Cerezo et al., 2020; Greiff et al., 2015; Nicolay et al., 2021).

In sum, it can be seen that computer-based process data analysis for educational purposes has become more “fine-grained” in the last 15–20 years. In fact, an increased specialization of individual process data analyses has primarily been made possible due to technical advancements allowing researchers to capture the desired processes accurately using increasingly powerful computers and software (Adjerdj and Kelley, 2018). Additionally, the availability of public process data repositories has put increased amounts of more easily accessible data at the disposal of both researchers and educational stakeholders.

Public computer-based process data repositories

Nearly simultaneously with the widespread recognition of the scholarly potential of computer-based process data analysis, researchers became particularly interested in using such data to improve educational practice (e.g., Branch, 2001). Initially, scientists engaging in computer-based process data analysis largely had to work with their own datasets. However, around 15 years ago, the first public computer-based process data repository was established, creating a new possibility for researchers to carry out computer-based process data analysis without being forced to collect their own data, and easing the comparison and verification of published results (Koedinger et al., 2008). At the same time, research began to increasingly rely on data from digital learning environments used for communication among teachers and students (Baker and Yacef, 2009).

Currently, several public computer-based process data repositories from different sources exist. We want to showcase two contexts of data sources, which are frequently employed by researchers.

First, numerous public repositories for computer-based process data exist from *online tutoring systems* (e.g., “ASSISTments,” “DataShop,” “vMS-Share,” “KEEL”). These repositories are continuously growing, are freely accessible to researchers all over the world, and provide data from people’s interactions with computer-based systems across multiple domains, either in their raw log file format, or in item-level responses per individual participant. Given the amount and variety of available data, these public repositories represent a crucial resource for the advancement of educational policy and practice (Alcalá-Fdez et al., 2011; Blonder et al., 2019; Feng et al., 2009; Koedinger et al., 2010).

Second, data from international large scales assessments (ILSAs) represent another valuable source for data mining. The two presumably most widely known ILSAs “Program for International Student Assessment” (PISA) and “Program for the International Assessment of Adult Competencies” (PIAAC) offer great insights into (educationally) relevant competencies, including reading ability, numeracy, or innovative competencies, like problem solving, among both school students (PISA) and adults (PIAAC) across the globe. Some of these competencies are assessed with computer-based assessments and thus, are stored in computer-based process data (Jude, 2016; Kirsch and Lennon, 2017). As the logged process data of both assessments are publicly available on the official website of the Organization for Economic Co-operation and Development (OECD; see <https://www.oecd.org/pisa/data/> for PISA, and <https://www.oecd.org/skills/piaac/data/piaaclogfiles/> for PIAAC), it is not surprising that multiple studies have used these data for fine-grained analyses of students’ and adults’ educational achievement (Goldhammer et al., 2020; Greiff et al., 2015; He et al., 2017; Liao et al., 2019; Teig et al., 2020). For both assessments, the process data are both available in raw log file format, and in item-based responses per country. Thus, it can be seen that the computer-based process data stored in the publicly available data repositories for PISA and PIAAC represent a valuable resource for scholars to learn more about how both students and adults interact with digital environments as well as their skills in key educational competencies.

The current state of computer-based process data analysis

Having provided a thorough overview of the purpose, history, and application of computer-based process data analysis in educational measurement, and introducing some of the most relevant publicly available computer-based process data repositories, we will now discuss the current state of computer-based process data analysis. In this regard, we will focus on three major facets of computer-based process data analysis due to their importance for successfully conducting and drawing valid inferences from computer-based process data analysis: (a) psychometric models (see 4.1), (b) formative assessment (see 4.2), and (c) summative assessment (see 4.3). In brief, psychometric models serve as statistical foundations for computer-based process data analysis. Formative assessment refers to the evaluation of students’ learning progress during a particular learning exercise or course, while summative assessment measures students’ learning progress after the learning exercise or course has been completed.

Psychometric models

Psychometrics can be defined as “the science of psychological assessment” (Rust and Golombok, 2014, p. 4) and is concerned with how psychology research measures designated phenomena to arrive at meaningful conclusions (Jones and Thissen, 2006). Computer-based process data pose several challenges for psychometric modeling that should be highlighted before moving on to psychometric models that have been specifically developed to deal with these challenges (Mislevy et al., 2012).

Firstly, computer-based process data differ markedly from “traditional” data (i.e., data that were not collected automatically via a digital entity during an assessment process). Such “traditional” data are usually limited to an absolute preference or performance score obtained from each individual participant over a given number of items. For instance, when analyzing such “traditional” data, it is nearly impossible to detect a learning effect for a participant over the course of multiple assessment items apart from a potentially higher overall success rate. Computer-based process data, on the other hand, might detect such a learning effect not only in the form of higher overall performance scores, but also using other indicators like consistently lower response times or fewer intermediate errors committed over several items.

Secondly, logged process data are not limited to a particular type of data such as responses on a Likert scale or number of items correctly recalled from memory. Instead, a single log file contains numerous data points in various formats, such as written information the participant typed, time stamps indicating when the participant started and completed a task, and the exact times and instances the participant clicked on each button shown on the computer screen. Therefore, compared to traditional data, computer-based process data pose unique challenges for psychometric models due to their fine-grained nature.

Today, however, a variety of different psychometric models (i.e., statistical frameworks aiming to explain how data are generated and can subsequently be analyzed in a valid and reliable manner; American Psychological Association, 2020) suitable for computer-based process data analysis are available that seek to address these challenges. One example is the evidence-centered design (ECD) framework, which has received increasing attention for computer-based process data analysis in the educational context (Mislevy et al., 2003). The ECD framework encompasses several different parts. The *student model* consists of the variables to be measured, while the *task model* refers to the situations in which these variables can be observed. Notably, the psychometric model, together with evidence identification rules, serves as a link between the two former models, creating a final evidence model. Based on this ECD framework, Levy (2020) identifies the four psychometric categories of preparing data, analyzing data, presenting data, and situating the work in each of these three categories. In order to account for the varying nature of computer-based process data, Levy (2020) argues that the interdependency between such process data (indicative of a person’s behavior during the task completion process) and product data (an end result or performance score after task completion) needs to be considered. This can be accomplished, for instance, by adding variables from the student model in order to distinguish factors related to the process from those constituting the product, and additional related techniques (p. 223).

Next to the ECD, the Extended Learning and Assessment System (XLAS) framework (Von Davier et al., 2019) represents a second psychometric model for computer-based process data. The developers of this framework argue that the three dimensions of *assessment*, *learning*, and *navigation* must be considered when working with multimodal data such as computer-based process data. They additionally emphasize that, when trying to recover information from computer-based process data, how to extract and create meaningful results from initial raw and unorganized bunches of data should be prioritized.

In addition, Kroehne and Goldhammer (2018) propose using finite state machines to extract the desired data points from initially cumbersome-to-analyze log files. Such finite state machines act as a measurement tool of how the participant interacts with the computer during a specific part of the overall assessment; thus, they act as encoders of the information stored in a given log file. This specific part (e.g., one item within a large questionnaire) represents the finite state, which can be defined theoretically a priori, but also changed later on—for example, to achieve comparability with other assessment approaches incorporating similar action sequences. Importantly, participants' behavior as they progress through this collection of subsequent finite items can then be investigated not only on the within-person level, but also across the different items. Kroehne and Goldhammer (2018) illustrate this procedure using computer-based process data from PISA, discovering that their finite state machine approach produces valid results.

Taken together, although the complexity of computer-based process data poses a unique challenge for researchers, several psychometric models are currently available to facilitate scholarly work with such data. Indeed, psychometric model approaches such as XLAS, ECD, and finite state machines represent valuable tools for both educational stakeholders and scholars. Importantly, as psychometric models are generally based on a theoretical foundation, they are only as useful as their underlying theory. Thus, a strong and well-established theory serves as the best precursor for the development of a meaningful and useable psychometric model (Rupp et al., 2012; Zehner et al., 2020). Consequently, building one's psychometric model on a sound theoretical foundation crucially facilitates formative and summative assessment of students based on computer-based process data (Drachler and Goldhammer, 2020; Zhu et al., 2016).

Formative assessment

Evaluating students' learning progress by means of different assessment techniques has always been an integral part of educational measurement. Nevertheless, formative assessment has received increasing attention in education since the beginning of the 21st century (e.g., Bell and Cowie, 2001). Formative assessment entails the evaluation of students' learning in order to adapt learning policies (e.g., curricula) both on an individual level as well as on a broader class or school level (Van der Kleij et al., 2015). Formative assessment practices are widely known as "assessment for learning" and typically performed during the course of a learning activity or semester rather than afterward, as is the case for summative assessment (Saeed et al., 2018; see 4.3). In addition, a core purpose of formative assessment is to support students' learning by giving and receiving feedback. This feedback does not inherently center on monitoring a student's progress with respect to the learning content, but rather seeks to provide suggestions for improvement with regard to a student's learning strategies (Gikandi et al., 2011). Furthermore, formative assessment can also be computer-based (Bergeson, 2019; Wijesooriya et al., 2015). As such, it can be put into practice by asking students to carry out a classroom task in several smaller groups, for example, and subsequently having the students provide feedback to their peers in the same group. This not only allows students to point out aspects of tasks their peers are approaching differently, but also directly engage in a discussion about how these different learning strategies might be more or less likely to (in)efficiently lead to (in)correct results. Moreover, such discussions can increase students' knowledge about learning strategies and their flexibility in applying and adapting them. Overall, formative assessment represents an important complement to summative assessment, as it can significantly help students achieve higher performance scores in summative assessment tests (Harlen, 2011; Zheng et al., 2019).

The analysis of computer-based process data from formative assessment contexts is spurred by its increasing relevance in educational contexts. Several studies have developed digital learning environments that provide feedback to students throughout their interaction with the environment, while simultaneously yielding a rich set of computer-based process data. For instance, analyzing computer-based process data from the personalized e-learning system PELS, which specifically tailors the content and frequency of feedback to a student's individual needs, revealed a significant increase in students' mathematical abilities after completing learning sessions in the digital learning environment involving more immediate feedback (Chen and Chen, 2009). Similar results were obtained for a related program in the field of mathematics called MATHia, which students used over the course of an entire school year (Zheng et al., 2019). In their study, Zheng et al. (2019) analyzed the computer-based process data employing different statistical models, which produced different results. Thus, the authors emphasize the importance of specifying the statistical methodology of choice in a given context prior to diving into a fine-grained investigation of computer-based process data (e.g., cluster analysis versus hierarchical modeling; Miller et al., 2015). Likewise, researchers should carefully consider how to use and interpret student scores obtained from computer-based process data in order to ensure the validity of formative assessments and their results (Hopster-den Otter et al., 2019).

Alongside established school subjects such as mathematics, domain-general 21st century skills (i.e., skills such as complex problem solving that have been identified as especially relevant for today's learners; Trilling and Fadel, 2009) have been located at the core of digital learning environments producing computer-based process data. In this regard, an analytical framework has been established addressing how to immediately provide feedback to students while they are still working on a given task (Choi and Cho, 2020). This allows students' learning paths to be modeled in more depth, tracing not only the final result of a particular

task, but also the individual steps the student took to arrive at this final result. Moreover, [Choi and Cho \(2020\)](#) analyzed the computer-based process data from their formative assessment methodology by contrasting the observed learning trajectory of individual students to a predefined expected learning trajectory at several time points during a task. This allowed the specific time points at which feedback should be provided to students to be identified, leading to improved feedback timing and content ([Choi and Cho, 2020](#)).

Summative assessment

At present, the most widely applied assessment technique in Western classrooms is summative assessment ([Saeed et al., 2018](#)). Summative assessment refers to assessing students' learning outcomes once at the end of a course in order to make decisions about the students' ability to transition to other courses or the next grade level ([Anthony and Susan, 2005](#)). **Table 1** can help readers understand the differences between formative and summative assessment at a glance.

One example of summative assessment in practice is to require university students to complete a final exam in an introductory module at the end of the semester, and subsequently using each student's score on this final exam to determine whether they are permitted to enroll in the advanced module next semester. In general, summative assessments are deployed for purposes at the individual student level, such as admissions, ability grouping, grading, norming, or mastery of a particular subject ([McGarr and Clifford, 2013](#); [Van Groen and Eggen, 2020](#)).

Multiple research studies have already used computer-based process data analysis to shed light on factors influencing students' summative assessment results. For instance, [Aojula et al. \(2006\)](#) compared the results of a computer-based summative assessment of undergraduate students to teachers' marking by hand and to the results obtained from multiple-choice marking machines that read students' answer sheets. They demonstrate that using computer-based process data analysis represents a viable and beneficial alternative to traditional summative assessment modalities, given its higher practicality and accuracy (e.g., the results are provided directly and automatically and do not need to be calculated by hand). Furthermore, [Gocheva-Ilieva et al. \(2020\)](#) employed common data mining techniques such as cluster analysis to evaluate the impact of different elements of the summative assessment approach on students' final grades in mathematics, discovering a significant influence of students' sex and selected elective subjects. Likewise, computer-based process data from a large-scale summative assessment in France were used to examine the potential benefit of students' employment of strategies such as trial and error versus a planned, structured approach when solving a computer-based mathematics problem ([Salles et al., 2020](#)). The authors concluded that although neither of the two strategic approaches resulted in better or worse overall performance, all students could validly be classified into one of the two strategy groups using the employed cluster analysis. Another recently published study presents the assessment program Winsight[®], which targets potential improvements in individual summative assessment components (e.g., by providing information about students' proficiency and ensuring full accessibility for students with disabilities or limited knowledge of the language of instruction) in order to foster student learning ([Stone and Wylie, 2019](#)). The program primarily seeks to funnel such information to educational stakeholders such as school headmasters and curricular decision-makers. The authors apply their new approach using computer-based process data analysis in order to assess test accessibility (i.e., how different populations of students navigate through the test, select and refine their answers) and thus provide good validity indicators for their assessment program.

In sum, we have seen that caution is advised when analyzing computer-based process data to ensure that valid, meaningful and interpretable results are obtained that carry useful implications for education. However, when suitable statistical methodology and psychometric models capable of accurately capturing and presenting the desired information are applied, computer-based process data analysis represents a valuable tool for the advancement of both formative and summative assessment practices. This includes but is not limited to providing feedback about and evaluating students' learning trajectories.

Table 1 Overview of the purpose, administration, usability, and examples of formative and summative assessment in education.

	<i>Formative assessment</i>	<i>Summative assessment</i>
Goal	Providing feedback and facilitating learning and instruction	Assessing students' knowledge of a particular subject or domain
Time of application	During a learning exercise or course	After a learning exercise or course
Use for teachers	Evaluating students' learning methods and the quality of the learning material, adaptation of exercises and curricula	Understanding individual and class performance, comparing scores to other schools, grading
Use for students	Assessing one's own learning methods in relation to peers, self-evaluation	Evaluating one's own progress in relation to peers and curricular standards
Examples	Self-tests, practice exam	Final exam, individual report

Applied computer-based process data analysis in educational measurement: the example of complex problem solving

At this point, having presented a general overview of the current state of computer-based process data analysis in educational measurement, we would like to illustrate how computer-based process data analysis can be used in practice to fill knowledge gaps regarding how students apply a skill that has been shown to be particularly relevant for educational success and beyond: the domain-general 21st century skill complex problem solving (CPS).

CPS skills can be defined as the ability to solve problems that possess unique features, such as hidden relationships between variables, multiple goals to be reached simultaneously, and the potential for variable values to change dynamically at any given time during the solution process (Greiff et al., 2013a,b). One example of a complex problem could be trying to use a recently updated self-service ticket machine to purchase a train ticket, an example that has already been used in PISA (OECD, 2014; here CPS was labeled as creative problem solving). In order to ultimately make the correct purchase, it is crucial to apply a strategic approach (i.e., deliberately clicking on different buttons on the machine) to build up one's knowledge instead of performing random clicks (i.e., applying the vary-one-thing-at-a-time or VOTAT strategy, also referred to as the control of variables or CVS strategy; Schwichow et al., 2016).

Computer-based microworlds have become the current standard for assessing students' CPS skills (e.g., Stadler et al., 2015). These microworlds involve different hypothetical scenarios as complex problems that must be solved, such as creating a lemonade using different ingredients with arbitrary labels or evaluating the impact of changing the dosage of several unknown types of medication on different body functions (Greiff et al., 2013a,b). In CPS assessment, students can explore the initially hidden relationships among the variables within a given item by moving sliders and clicking different buttons. These actions, as well as others, can be used for process data analysis. See Fig. 2 below for an exemplary CPS item.

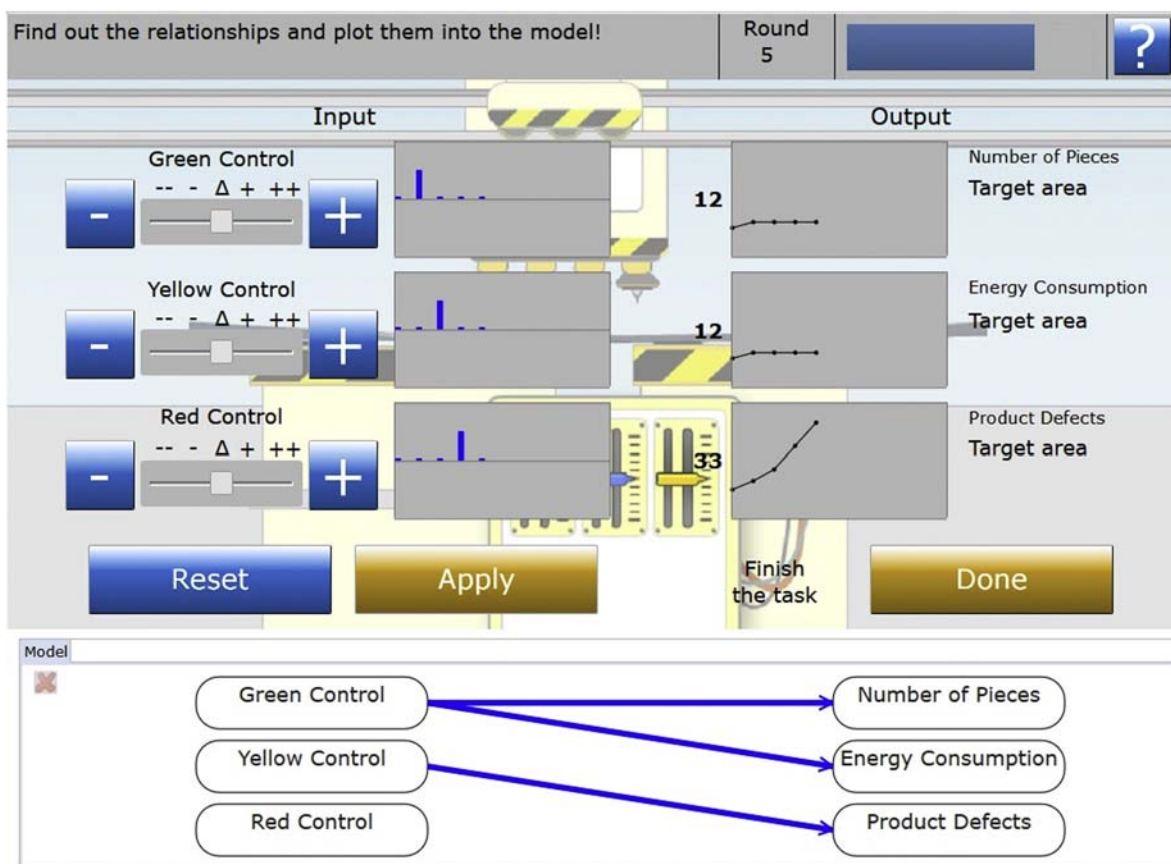


Fig. 2 A complex problem microworld involving multiple variables with hidden relationships (see also Fig. 1 in The role of computer-based process data in educational measurement section for a corresponding log file containing process data). After the variable relationships have been discovered in the problem space (top part) by applying strategies such as VOTAT, the student is asked to indicate each present relationship in the model space (bottom part).

Importantly, computer-based process data are collected and stored in log files while students work on the complex problems within such a microworld. These log files capture all relevant components of a student's interaction with a given problem, including time spent on an item, all variable manipulations and their corresponding time stamps, and whether the student achieved or at least approximated the predefined goals (Xu et al., 2018). A common approach to extracting valuable information from this computer-based process data after CPS assessment is to write an automated program that iterates through all log files and parses the relevant data points for further subsequent statistical analyses. For instance, if data are stored in XML-format, this can be achieved using so-called XML parsers (i.e., programs that automatically extract the desired information from raw log files; e.g., Appen and McDaniel, 2009).

The importance of CPS for both educational and career success as the ability to manage novel and dynamic situations in a systematic way has been discussed in multiple studies (Schweizer et al., 2013; Wüstenberg et al., 2012). Thus, it does not come as a surprise that a large body of research on CPS using computer-based process data from large-scale assessments, including PISA, has accumulated in recent years (Greiff et al., 2015; Han et al., 2019; Liu et al., 2020). For instance, research has established the particular benefit of applying strategies such as the aforementioned VOTAT strategy alone or in concert with other related strategies for successful CPS performance (Molnár and Csapó, 2018). Additionally, the time students spend on individual complex problems has been shown to influence their probability of solving the problem successfully or unsuccessfully (Scherer et al., 2015). A study by Ren et al. (2019) uncovered how students can successfully balance multiple goals to be achieved simultaneously in CPS. Additional research has sought to identify different levels of CPS proficiency in students, for example based on how they approach a CPS assessment test (Greiff et al., 2015; Stadler et al., 2020). Eichmann et al. (2020) used PISA computer-based process data to uncover differences in students' CPS performance based on their sex and ethnic background: the higher CPS success rates found among boys compared to girls were attributable to differences in exploration behavior between the sexes, whereas the behavioral differences investigated were unable to account for the performance differences between students with versus without migration backgrounds.

In summary, as demonstrated by the selected research endeavors described above, computer-based process data analysis has greatly supported the investigation of students' competence levels in crucial skills such as CPS (for an overview of computer-based process data use in CPS, see also Herde et al., 2016). At the same time, computer-based process data analysis has helped to uncover opportunities to improve students' skills in these areas by means of training programs in digital learning environments (e.g., Azevedo, 2007). Moreover, as computers and digital learning and assessment tools become increasingly prevalent in the educational context, we can expect further advancements based on computer-based process data analysis in the near future. Consequently, we will now discuss potential challenges and future focus areas for computer-based process data analysis, after providing a broad summary of its current state of application in the educational measurement domain.

Summary, challenges, and future directions of computer-based process data analysis in educational measurement

We have seen that computer-based process data analysis in educational measurement carries a variety of benefits and challenges simultaneously for both educational stakeholders and researchers. Despite the fact that this promising research tool emerged not long ago, it has already produced several noteworthy advancements, as described throughout this article. Computer-based process data analysis has evolved into a versatile method for evaluating students' learning progress in order to create specifically targeted assessment and training programs for educationally relevant competencies. As such, computer-based process data analysis represents an important means of drawing inferences about students' performance in both formative and summative assessment contexts (Greiff, 2020). On the other hand, the increasingly fine-grained, rich nature of computer-based process data necessitates the development and continuous refinement of psychometric models that are able to extract meaningful information from increasingly complex raw log files. Today, computer-based process data analysis has become an established and crucial tool for improving educational measurement across the globe. Moreover, as technology continues to advance, we can expect that use of computer-based process data analysis in and beyond the educational context will expand rather than contract in upcoming years (Bergner and von Davier, 2019; Dosta et al., 2020).

Turning to potential future directions for computer-based process data analysis in educational measurement, three aspects must be considered: (a) the ongoing changing nature of assessment; (b) the evolution of digital learning environments; and (c) the increasing richness and diversity of computer-based process data, which calls for powerful and flexible statistical analysis tools. The opportunities and challenges associated with each of these aspects will be considered individually below, in order to provide some suggestions about where the use of computer-based process data analysis for educational purposes may be headed in the near future.

Firstly, we have already discussed the increasing use of computer-based process data analysis in both formative and summative assessment contexts. Both approaches represent particularly relevant educational measurement practices, and, today, they are frequently used in conjunction (Owen et al., 2014; Yang and Li, 2018). In addition, diagnostic and portfolio assessment are also becoming increasingly relevant (Chen et al., 2007; Saeed et al., 2018). Diagnostic assessment is strongly related to formative

assessment, and primarily seeks to unravel how well students are prepared for the specific difficulties they experience when working on a particular task (Csapó and Molnár, 2019). In portfolio assessment, students hand in different assignments to their teacher and are then given the possibility to revise these assignments based on their teacher's feedback. After these revisions have been made, each assignment is then handed in again and added to the student's portfolio, which serves as the primary method of assessing their overall learning progress over the course of a full semester or school year. As such, portfolio assessment features elements of both formative and summative assessment (Steen-Utheim and Hopfenbeck, 2019). Thus, as educational practice moves away from exclusive reliance on summative assessment, additional assessment approaches have emerged in recent years and are likely to continue to emerge and be refined in the future. These developments will involve the collection of a multitude of computer-based process data, and new types of computer-based process data will become available to both educational stakeholders and researchers (Shute et al., 2016a).

Secondly, throughout this article, we have presented several digital learning environments that have been developed to guide students through the acquisition and/or refinement of a specific skill. Over the last decade, these digital learning environments have largely evolved into digital game-based learning environments (Loh, 2012). The game-based nature of these environments is beneficial due to students' high levels of satisfaction when working with such programs, thereby keeping students interested, keen to learn, and entertained at the same time (Sawyer et al., 2018). Such digital game-based learning environments inherently produce a considerable amount of computer-based process data. For instance, in a meta-analysis of game-based learning, Qian and Clark (2016) discovered that asking students to work with digital game-based learning environments led to improved levels of critical thinking, creativity, communication, and collaboration. We can therefore see these types of digital learning environments being used increasingly often in educational contexts in upcoming years.

All in all, as with the observed changes in assessment approaches described above, the fact that digital learning environments are continuously evolving will impose new challenges but also new opportunities for computer-based process data analysis in educational measurement. These noteworthy developments in assessment type and methodology will lead to the availability of an increased amount of even more fine-grained computer-based process data in the future. This brings us to the final, and most crucial, aspect to be considered: the trade-off between the richness of computer-based process data and the power and suitability of statistical tools to analyze these data in a meaningful way. Indeed, as a result of the increasingly complex and comprehensive assessment approaches employed in educational contexts, the log files representing the basis for computer-based process data analysis are becoming increasingly fine-grained and difficult to scrutinize (Kumar and Chadha, 2011). Moreover, how to address the potential issue of missing and/or inaccurate data has been discussed in multiple studies, highlighting the importance of quality assurance in computer-based process data analysis (Weeks et al., 2016; Yamamoto and Lennon, 2018). Furthermore, in addition to computer-based process data generated by students completing tasks on laptops and tablets, which can at this point be seen as rather "traditional," other avenues providing complementary computer-based process data are likely to evolve in the future. For instance, technology-based process data from eye tracking (Bergner and von Davier, 2019; Emerson et al., 2020) and physiological sources such as blood pressure, heart rate, and electrodermal activity may be collected as a student works on a digital learning environment task (Sharma et al., 2020; Vrzakova et al., 2020). Moreover, computer-based process data from outside school, including students' interactions with smartphone learning apps they engage with in their leisure time, present an additional source of information relevant for their educational progress (Markowitz et al., 2014; Sin and Muthu, 2015). In other words, the current and potential future sources of computer-based process data for educational purposes seem almost unlimited. Consequently, it will become paramount for educational stakeholders to prioritize the analysis of certain kinds of computer-based process data over others due to the increasing variety of available information. In turn, the creation and refinement of statistical tools for successful, valid, and interpretable computer-based process data analysis will be necessary. As discussed in this article, several such tools already exist today, and we are confident that their development and refinement will continue in the future.

Overall, in its current form, computer-based process data analysis represents a valuable tool for both educational stakeholders and researchers to evaluate and advance educational measurement policies and practices. Since its inception in the final stages of the 20th century, computer-based process data analysis has become a standard method for investigating students' learning progress and learning success. Given the promising possibilities of this evolving technology, we expect computer-based process data analysis to become increasingly present in the future in the everyday lives of educational decision-makers, teachers, and researchers, who share the common goal of providing today's students with a state-of-the-art education in order to support and prepare them for the challenges they will face throughout and after their educational journey.

Acknowledgments

This research was supported by a grant from the Fonds National de la Recherche Luxembourg (Luxembourgish National Research Found) [CORE "TRIOPS"].

References

- Adams, R., Vista, A., Scoular, C., Awwal, N., Griffin, P., Care, E., 2015. Automatic coding procedures for collaborative problem solving. In: *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht, pp. 115–132.
- Adjerid, I., Kelley, K., 2018. Big data in psychology: a framework for research advancement. *Am. Psychol.* 73 (7), 899–917. <https://doi.org/10.1037/amp0000190>.
- Alcalá-Fdez, J., Fernández, A., Luengo, J., Derrac, J., García, S., Sánchez, L., Herrera, F., 2011. Keel data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. *J. Mult. Valued Log. Soft Comput.* 17.
- Aldowah, H., Al-Samarraie, H., Fauzy, W.M., 2019. Educational data mining and learning analytics for 21st century higher education: a review and synthesis. *Telematics Inf.* 37, 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>.
- Ali, M.M., 2013. Role of data mining in education sector. *Int. J. Comput. Sci. Mobile Comput.* 2 (4), 374–383.
- American Psychological Association, 2020. Psychometric Model. <https://dictionary.apa.org/psychometric-model>.
- Anthony, J.N., Susan, M.B., 2005. *Education Assessment of Students*. Pearson Education Ltd.
- Aojula, H., Barber, J., Cullen, R., Andrews, J., 2006. Computer-based, online summative assessment in undergraduate pharmacy teaching: the Manchester experience. *Pharm. Educ.* 6 (4), 229–236. <https://doi.org/10.1080/15602210600886209>.
- Applen, J.D., McDaniel, R., 2009. *The Rhetorical Nature of XML: Constructing Knowledge in Networked Environments*. Routledge.
- Azevedo, R., 2007. Understanding the complex nature of self-regulatory processes in learning with computer-based learning environments: an introduction. *Metacognition Learn.* 2 (2–3), 57–65. <https://doi.org/10.1007/s11409-007-9018-5>.
- Baker, R.S., Yacef, K., 2009. The state of educational data mining in 2009: a review and future visions. *J. Educ. Data Min.* 1 (1), 3–17. <https://doi.org/10.5281/zenodo.3554657>.
- Bell, B., Cowie, B., 2001. The characteristics of formative assessment in science education. *Sci. Educ.* 85 (5), 536–553. <https://doi.org/10.1002/Osce.1022>.
- Bergeson, K.T., 2019. Reading specialists use verbal protocols as a formative assessment tool. *Read. Teach.* 73 (2), 185–193. <https://doi.org/10.1002/trtr.1815>.
- Berger, Y., von Davier, A.A., 2019. Process data in NAEP: past, present, and future. *J. Educ. Behav. Stat.* 44 (6), 706–732. <https://doi.org/10.3102/1076998618784700>.
- Blonder, N., Orsburn, B.C., Blonder, J., Gonzalez, C.A., 2019. Visual mass-spec share (VMS-Share): a new public web-based mass spectrometry visualization and data mining repository. *J. Proteonomics Bioinf.* 12. <https://doi.org/10.4172/0974-276X.1000495>.
- Bogarín, A., Romero, C., Cerezo, R., Sánchez-Santillán, M., 2014. Clustering for improving educational process mining. In: *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, pp. 11–15. <https://doi.org/10.1145/2567574.2567604>.
- Branch, J.L., 2001. Junior high students and think alouds: generating information-seeking process data using concurrent verbal protocols. *Libr. Inf. Sci. Res.* 23 (2), 107–122. [https://doi.org/10.1016/S0740-8188\(01\)00065-2](https://doi.org/10.1016/S0740-8188(01)00065-2).
- Brown, G.T., Andrade, H.L., Chen, F., 2015. Accuracy in student self-assessment: directions and cautions for research. *Assess Educ. Princ. Pol. Pract.* 22 (4), 444–457. <https://doi.org/10.1080/0969594X.2014.996523>.
- Bunderson, C.V., Inouye, D.K., Olsen, J.B., 1989. The four generations of computerized educational measurement. In: Linn, R.L. (Ed.), *Educational Measurement*, third ed. American Council on Education, pp. 367–407.
- Cairns, A.H., Gueni, B., Fhima, M., Cairns, A., David, S., Khelifa, N., 2015. Process mining in the education domain. *Int. J. Adv. Intell. Syst.* 8 (1), 219–232.
- Cerezo, R., Bogarín, A., Esteban, M., Romero, C., 2020. Process mining for self-regulated learning assessment in e-learning. *J. Comput. High Educ.* 32 (1), 74–88. <https://doi.org/10.1007/s12528-019-09225-y>.
- Chen, C.M., Chen, M.C., 2009. Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Comput. Educ.* 52 (1), 256–273. <https://doi.org/10.1016/j.compedu.2008.08.005>.
- Chen, C.M., Chen, M.C., Li, Y., 2007. Mining key formative assessment rules based on learner profiles for web-based learning systems. In: *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies*, Niigata, 2007, pp. 584–588. <https://doi.org/10.1109/ICALT.2007.189>.
- Choi, Y., Cho, Y.I., 2020. Learning analytics using social network analysis and Bayesian network analysis in sustainable computer-based formative assessment system. *Sustainability* 12 (19), 7950. <https://doi.org/10.3390/su12197950>.
- Csapó, B., Molnár, G., 2019. Online diagnostic assessment in support of personalized teaching and learning: the eDia system. *Front. Psychol.* 10, 1522. <https://doi.org/10.3389/fpsyg.2019.01522>.
- Danniels, E., Pyle, A., DeLuca, C., 2020. The role of technology in supporting classroom assessment in play-based kindergarten. *Teach. Teach. Educ.* 88, 102966. <https://doi.org/10.1016/j.tate.2019.102966>.
- DiCerbo, K., Shute, V., Kim, Y.J., 2017. The future of assessment in technology rich environments: psychometric considerations. In: *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*. Springer International, pp. 1–21.
- Dosta, M., Litster, J.D., Heinrich, S., 2020. Flowsheet simulation of solids processes: current status and future trends. *Adv. Powder Technol.* 31 (3), 947–953. <https://doi.org/10.1016/j.appt.2019.12.015>.
- Drachler, H., Goldhammer, F., 2020. Learning analytics and eAssessment—towards computational psychometrics by combining psychometrics with learning analytics. In: Burgos, D. (Ed.), *Radical Solutions and Learning Analytics*. Lecture Notes in Educational Technology. Springer, Singapore, pp. 67–80. https://doi.org/10.1007/978-981-15-4526-9_5.
- Eichmann, B., Goldhammer, F., Greiff, S., Brandhuber, L., Naumann, J., 2020. Using process data to explain group differences in complex problem solving. *J. Educ. Psychol.* 112 (8), 1546–1562. <https://doi.org/10.1037/edu0000446>.
- Emerson, A., Cloude, E.B., Azevedo, R., Lester, J., 2020. Multimodal learning analytics for game-based learning. *Br. J. Educ. Technol.* 51 (5), 1505–1526. <https://doi.org/10.1111/bjjet.12992>.
- Feng, M., Heffernan, N.T., Koedinger, K.R., 2009. Addressing the assessment challenge in an intelligent tutoring system that tutors as it assesses. *J. User Model. User Adapt. Interact.* 19, 243–266. <https://doi.org/10.1007/s11257-009-9063-7>.
- Gallardo, K., 2021. The importance of assessment literacy: formative and summative assessment instruments and techniques. In: Babo, R., Dey, N., Ashour, A.S. (Eds.), *Workgroups eAssessment: Planning, Implementing and Analysing Frameworks*. Intelligent Systems Reference Library, vol. 199. Springer, pp. 3–25. https://doi.org/10.1007/978-981-15-9908-8_1.
- Gikandi, J.W., Morrow, D., Davis, N.E., 2011. Online formative assessment in higher education: a review of the literature. *Comput. Educ.* 57 (4), 2333–2351. <https://doi.org/10.1016/j.compedu.2011.06.004>.
- Gobert, J.D., Sao Pedro, M., Raziuddin, J., Baker, R.S., 2013. From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *J. Learn. Sci.* 22 (4), 521–563. <https://doi.org/10.1080/10508406.2013.837391>.
- Gocheva-Ilieva, S., et al., 2020. Data mining for statistical evaluation of summative and competency-based assessments in mathematics. In: Martínez Álvarez, F., Troncoso Lora, A., Sáez Muñoz, J., Quintián, H., Corchado, E. (Eds.), *International Joint Conference: 12th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2019) and 10th International Conference on European Transnational Education (ICEUTE 2019)*, Advances in Intelligent Systems and Computing, vol. 951. Springer, pp. 207–216. https://doi.org/10.1007/978-3-030-20005-3_21.
- Goldhammer, F., Naumann, J., Rölke, H., Stelter, A., Tóth, K., 2017. Relating product data to process data from computer-based competency assessment. In: *Competence Assessment in Education*. Springer, Cham, pp. 407–425.
- Goldhammer, F., Hahnel, C., Kroehne, U., 2020. Analyzing log file data from PIAAC. In: *Large-Scale Cognitive Assessment*. Springer, Cham, pp. 239–269.

- Greiff, S., Holt, D.V., Funke, J., 2013a. Perspectives on problem solving in cognitive research and educational assessment: analytical, interactive, and collaborative problem solving. *J. Probl. Solving* 5, 71–91. <https://doi.org/10.7771/1932-6246.1153>.
- Greiff, S., Wüstenberg, S., Holt, D.V., Goldhammer, F., Funke, J., 2013b. Computer-based assessment of complex problem solving: concept, implementation, and application. *Educ. Technol. Res. Dev.* 61 (3), 407–421. <https://doi.org/10.1007/s11423-013-9301-x>.
- Greiff, S., Wüstenberg, S., Avisati, F., 2015. Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Comput. Educ.* 91, 92–105. <https://doi.org/10.1016/j.compedu.2015.10.018>.
- Greiff, S., 2020. Technology-based assessment in 21st century education. Hrsg. In: Zender, R., Ifenthaler, D., Leonhardt, T., Schumacher, C. (Eds.), *DELF 2020—Die 18. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.* Bonn: Gesellschaft für Informatik e.V., pp. 21–22.
- Han, Z., He, Q., von Davier, M., 2019. Predictive feature generation and selection using process data from PISA interactive problem-solving items: an application of random forests. *Front. Psychol.* 10, 2461. <https://doi.org/10.3389/fpsyg.2019.02461>.
- Harlen, W., 2011. On the relationship between assessment for formative and summative purposes. In: Gardner, J. (Ed.), *Assessment and Learning*. Sage Publishing, pp. 95–110. <https://doi.org/10.4135/9781446250808.n6>.
- He, Q., von Davier, M., Greiff, S., Steinhauer, E.W., Borysewicz, P.B., 2017. Collaborative problem solving measures in the Programme for International Student Assessment (PISA). In: *Innovative Assessment of Collaboration*. Springer, Cham, pp. 95–111. https://doi.org/10.1007/978-3-319-33261-1_7.
- Herde, C.N., Wüstenberg, S., Greiff, S., 2016. Assessment of complex problem solving: what we know and what we don't know. *Appl. Meas. Educ.* 29 (4), 265–277. <https://doi.org/10.1080/08957347.2016.1209208>.
- Hopster-den Otter, D., Wools, S., Eggen, T.J., Veldkamp, B.P., 2019. A general framework for the validation of embedded formative assessment. *J. Educ. Meas.* 56 (4), 715–732. <https://doi.org/10.1111/jedm.12234>.
- Horstmann, N., Ahlgrimm, A., Glöckner, A., 2009. How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgm. Decis. Mak.* 4, 335–354.
- Hsia, T.C., Shie, A.J., Chen, L.C., 2008. Course planning of extension education to meet market demand by using data mining techniques—an example of Chinkuo Technology University in Taiwan. *Expert Syst. Appl.* 34 (1), 596–602. <https://doi.org/10.1016/j.eswa.2006.09.025>.
- Huebener, M., Marcus, J., 2017. Compressing instruction time into fewer years of schooling and the impact on student performance. *Econ. Educ. Rev.* 58, 1–14. <https://doi.org/10.1016/j.econedurev.2017.03.003>.
- Jones, L.V., Thissen, D., 2006. 1 A history and overview of psychometrics. *Handb. Stat.* 26, 1–27. [https://doi.org/10.1016/S0169-7161\(06\)26001-2](https://doi.org/10.1016/S0169-7161(06)26001-2).
- Jude, N., 2016. The assessment of learning contexts in PISA. In: Kuger, S., Klieme, E., Jude, N., Kaplan, D. (Eds.), *Assessing Contexts of Learning. Methodology of Educational Measurement and Assessment*. Springer, Cham. https://doi.org/10.1007/978-3-319-45357-6_2.
- Kirsch, I., Lennon, M.L., 2017. PIAAC: a new design for a new era. *Large Scale Assess. Educ.* 5 (1), 11. <https://doi.org/10.1186/s40536-017-0046-6>.
- Koedinger, K.R., Cunningham, K., Leber, B., 2008. An open repository and analysis tools for fine-grained, longitudinal learner data. In: *Proceedings of the 1st International Conference on Educational Data Mining*, pp. 157–166.
- Koedinger, K.R., Baker, R.S., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J., 2010. A data repository for the EDM community: the PSLC DataShop. In: *Handbook of Educational Data Mining*, vol. 43. CRC Press, pp. 43–56.
- Kroehne, U., Goldhammer, F., 2018. How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika* 45, 527–563. <https://doi.org/10.1007/s41237-018-0063-y>.
- Kumar, V., Chadha, A., 2011. An empirical study of the applications of data mining techniques in higher education. *Int. J. Adv. Comput. Sci. Appl.* 2 (3), 80–84.
- Langley, A., 1999. Strategies for theorizing from process data. *Acad. Manag. Rev.* 24 (4), 691–710.
- Levy, R., 2020. Implications of considering response process data for greater and lesser psychometrics. *Educ. Assess.* 25 (3), 218–235. <https://doi.org/10.1080/10627197.2020.1804352>.
- Liao, D., He, Q., Jiao, H., 2019. Mapping background variables with sequential patterns in problem-solving environments: an investigation of United States adults' employment status in PIAAC. *Front. Psychol.* 10, 1–32. <https://doi.org/10.3389/fpsyg.2019.00646>.
- Liu, Y., Yang, B., Wu, L., Li, B., Yu, S., 2020. Middle-school students' behavior pattern and strategy selection in problem solving: a study based on data from PISA 2012. In: So, H.J., et al. (Eds.), *Proceedings of the 28th International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education.
- Loh, C.S., 2012. Information trails: in-process assessment of game-based learning. In: Ifenthaler, D., Eseryel, D., Ge, X. (Eds.), *Assessment in Game-Based Learning*. Springer, Cham, pp. 123–144. https://doi.org/10.1007/978-1-4614-3546-4_8.
- Luan, J., 2002. Data mining, knowledge management in higher education, potential applications. In: *Workshop Associate of Institutional Research International Conference*, Toronto, pp. 1–18.
- Mainert, J., Kretschmar, A., Neubert, J.C., Greiff, S., 2015. Linking complex problem solving and general mental ability to career advancement: does a transversal skill reveal incremental predictive validity? *Int. J. Lifelong Educ.* 34 (4), 393–411. <https://doi.org/10.1080/02601370.2015.1060024>.
- Markowetz, A., Błaskiewicz, K., Montag, C., Switala, C., Schlaepfer, T.E., 2014. Psychoinformatics: big data shaping modern psychometrics. *Med. Hypotheses* 82 (4), 405–411. <https://doi.org/10.1016/j.mehy.2013.11.030>.
- Mazza, R., Milani, C., 2005. Exploring usage analysis in learning systems: gaining insights from visualisations. In: *Workshop on Usage Analysis in Learning Systems at the 12th International Conference on Artificial Intelligence in Education*, pp. 65–72.
- McGarr, O., Clifford, A.M., 2013. "Just enough to make you take it seriously": exploring students' attitudes towards peer assessment. *High Educ.* 65 (6), 677–693. <https://doi.org/10.1007/s10734-012-9570-z>.
- Miller, L.D., Soh, L.K., Samal, A., Kupzyk, K., Nugent, G., 2015. A comparison of educational statistics and data mining approaches to identify characteristics that impact online learning. *J. Educ. Data Min.* 7 (3), 117–150. <https://doi.org/10.5281/zenodo.3554731>.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., 2003. On the structure of educational assessments. *Measurement* 1 (1), 3–62. https://doi.org/10.1207/S15366359MEA0101_02.
- Mislevy, R.J., Behrens, J.T., Dicerbo, K.E., Levy, R., 2012. Design and discovery in educational assessment: evidence-centered design, psychometrics, and educational data mining. *J. Educ. Data Min.* 4 (1), 11–48. <https://doi.org/10.5281/zenodo.3554641>.
- Molnár, G., Csapó, B., 2018. The efficacy and development of students' problem-solving strategies during compulsory schooling: logfile analyses. *Front. Psychol.* 9, 302. <https://doi.org/10.3389/fpsyg.2018.00302>.
- Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., Heiner, C., 2005. An educational data mining tool to browse tutor-student interactions: time will tell. In: *Proceedings of the Workshop on Educational Data Mining, National Conference on Artificial Intelligence*. AAAI Press, pp. 15–22.
- Mostow, J., 2004. Some useful design tactics for mining its data. In: *Proceedings of the ITS2004 Workshop on Analyzing Student-Tutor Interaction Logs to Improve Educational Outcomes*, pp. 20–28.
- Nicolay, B., Krieger, F., Stadler, M., Gobert, J., Greiff, S., 2021. Lost in transition—Learning analytics on the transfer from knowledge acquisition to knowledge application in complex problem solving. *Comput. Hum. Behav.* 115, 106594. <https://doi.org/10.1016/j.chb.2020.106594>.
- OECD, 2013. PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy. OECD Publishing, Paris, France. Retrieved from http://www.oecd-ilibrary.org/education/pisa-2012-assessment-and-analyticalframework_9789264190511-en.
- OECD, 2014. PISA 2012 Results: Creative Problem Solving: Students' Skills in Tackling Real-Life Problems (Volume V). OECD Publishing.
- OECD, 2017. PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science. OECD Publishing.

- Owen, V.E., Ramirez, D., Salmon, A., Halverson, R., 2014. Capturing learner trajectories in educational games through ADAGE (assessment data aggregator for game environments): a click-stream data framework for assessment of learning in play. In: American Educational Research Association Annual Meeting, pp. 1–7.
- Paulsen, M.F., 2003. Online Education: Learning Management Systems: Global E-Learning in a Scandinavian Perspective. NKI Forlaget.
- Qian, M., Clark, K.R., 2016. Game-based learning and 21st century skills: a review of recent research. *Comput. Hum. Behav.* 63, 50–58. <https://doi.org/10.1016/j.chb.2016.05.023>.
- Ren, Y., Luo, F., Ren, P., Bai, D., Li, X., Liu, H., 2019. Exploring multiple goals balancing in complex problem solving based on log data. *Front. Psychol.* 10, 1975. <https://doi.org/10.3389/fpsyg.2019.01975>.
- Romero, C., Ventura, S., 2007. Educational data mining: a survey from 1995 to 2005. *Expert Syst. Appl.* 33 (1), 135–146. <https://doi.org/10.1016/j.eswa.2006.04.005>.
- Romero, C., Ventura, S., 2020. Educational data mining and learning analytics: an updated survey. *Wiley Interdiscip. Rev.* 10 (3), e1355. <https://doi.org/10.1002/widm.1355>.
- Rupp, A.A., Levy, R., Dicerbo, K.E., Sweet, S.J., Crawford, A.V., Calico, T., Benson, M., Fay, D., Kunze, K.L., Mislevy, R.J., Behrens, J.T., 2012. Putting ECD into practice: the interplay of theory and data in evidence models within a digital learning environment. *J. Educ. Data Min.* 4 (1), 49–110. <https://doi.org/10.5281/zenodo.3554643>.
- Rust, J., Golombok, S., 2014. *Modern Psychometrics: The Science of Psychological Assessment*. Routledge.
- Saeed, M., Tahir, H., Latif, I., 2018. Teachers' perceptions about the use of classroom assessment techniques in elementary and secondary schools. *Bull. Educ. Res.* 40 (1), 115–130.
- Salles, F., Dos Santos, R., Keskpaik, S., 2020. When didactics meet data science: process data analysis in large-scale mathematics assessment in France. *Large Scale Assess. Educ.* 8, 1–20. <https://doi.org/10.1186/s40536-020-00085-y>.
- Sawyer, R., Rowe, J., Azevedo, R., Lester, J., 2018. Filtered Time Series Analyses of Student Problem-Solving Behaviors in Game-Based Learning. Paper Presented at the 11th International Conference on Educational Data Mining (Raleigh, NC, July 16–20).
- Scherer, R., Greiff, S., Hautamäki, J., 2015. Exploring the relation between time on task and ability in complex problem solving. *Intelligence* 48, 37–50. <https://doi.org/10.1016/j.intell.2014.10.003>.
- Schweizer, F., Wüstenberg, S., Greiff, S., 2013. Validity of the MicroDYN approach: complex problem solving predicts school grades beyond working memory capacity. *Learn. Individ. Differ.* 24, 42–52. <https://doi.org/10.1016/j.lindif.2012.12.011>.
- Schwichow, M., Croker, S., Zimmerman, C., Höfler, T., Härtig, H., 2016. Teaching the control-of-variables strategy: a meta-analysis. *Dev. Rev.* 39, 37–63. <https://doi.org/10.1016/j.dr.2015.12.001>.
- Sharma, K., Papamitsiou, Z., Olsen, J.K., Giannakos, M., 2020. Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In: *Proceedings of the 10th International Conference on Learning Analytics & Knowledge*, pp. 480–489.
- Shute, V.J., Leighton, J.P., Jang, E.E., Chu, M.W., 2016a. Advances in the science of assessment. *Educ. Assess.* 21 (1), 34–59. <https://doi.org/10.1080/10627197.2015.1127752>.
- Shute, V.J., Wang, L., Greiff, S., Zhao, W., Moore, G., 2016b. Measuring problem solving skills via stealth assessment in an engaging video game. *Comput. Hum. Behav.* 63, 106–117. <https://doi.org/10.1016/j.chb.2016.05.047>.
- Shute, V.J., 2011. Stealth assessment in computer-based games to support learning. *Comput. Games Instr.* 55 (2), 503–524.
- Sin, K., Muthu, L., 2015. Application of big data in educational data mining and learning analytics—a literature review. *ICTAC J. Soft Comput.* 5 (4), 1035–1049. <https://doi.org/10.21917/ijsc.2015.0145>.
- Slater, S., Joksimović, S., Kovanovic, V., Baker, R.S., Gasevic, D., 2017. Tools for educational data mining: a review. *J. Educ. Behav. Stat.* 42 (1), 85–106. <https://doi.org/10.3102/1076998616666808>.
- Srivastava, J., Cooley, R., Deshpande, M., Tan, P.N., 2000. Web usage mining: discovery and applications of usage patterns from web data. *ACM SIGKDD Explor. Newsl.* 1 (2), 12–23.
- Stadler, M., Becker, N., Gödker, M., Leutner, D., Greiff, S., 2015. Complex problem solving and intelligence: a meta-analysis. *Intelligence* 53, 92–101.
- Stadler, M., Hofer, S., Greiff, S., 2020. First among equals: log data indicates ability differences despite equal scores. *Comput. Hum. Behav.* 106442. <https://doi.org/10.1016/j.chb.2020.106442>.
- Steen-Utetheim, A., Hopfenbeck, T.N., 2019. To do or not to do with feedback. A study of undergraduate students' engagement and use of feedback within a portfolio assessment design. *Assess. Eval. High Educ.* 44 (1), 80–96. <https://doi.org/10.1080/02602938.2018.1476669>.
- Stone, E., Wylie, E.C., 2019. Building a validity argument while developing and using an assessment: a concurrent approach for the Winsight® summative assessment. *ETS Res. Rep. Ser.* 2019 (1), 1–20. <https://doi.org/10.1002/ets2.12261>.
- Teig, N., Scherer, R., Kjærnsli, M., 2020. Identifying patterns of students' performance on simulated inquiry tasks using PISA 2015 log-file data. *J. Res. Sci. Teach.* 57 (9), 1400–1429. <https://doi.org/10.1002/tea.21657>.
- Trilling, B., Fadel, C., 2009. *21st Century Skills, Learning for Life in Our Times*. John Wiley & Sons.
- Vainikainen, M.P., 2014. Finnish Primary School Pupils' Performance in Learning to Learn Assessments: A Longitudinal Perspective on Educational Equity. Report 360. Unigrafia, Helsinki, Finland.
- Van der Aalst, W.M.P., Pesic, M., Song, M., 2010. Beyond process mining: from the past to present and future. In: *International Conference on Advanced Information Systems Engineering*. Springer, Cham, pp. 38–52.
- Van der Aalst, W.M.P., Guo, S., Gorissen, P., 2015. Comparative process mining in education: an approach based on process cubes. In: Ceravolo, P., Accorsi, R., Cudre-Mauroux, P. (Eds.), *Data-Driven Process Discovery and Analysis. SIMPDA 2013, Lecture Notes in Business Information Processing*, vol. 203. Springer. https://doi.org/10.1007/978-3-662-46436-6_6.
- Van der Aalst, W.M.P., 2012. Process mining: overview and opportunities. *ACM Trans. Manag. Inf. Syst.* 3 (2), 1–17. <https://doi.acm.org/10.1145/0000000.0000000>.
- Van der Kleij, F.M., Vermeulen, J.A., Schildkamp, K., Eggen, T.J.H.M., 2015. Integrating data-based decision making, assessment for learning and diagnostic testing in formative assessment. *Assess. Educ. Princ. Pol. Pract.* 22 (3), 324–343. <https://doi.org/10.1080/0969594X.2014.999024>.
- Van Groen, M.M., Eggen, T.J., 2020. Educational test approaches: the suitability of computer-based test types for assessment and evaluation in formative and summative contexts. *J. Appl. Test. Technol.* 21 (1), 12–24.
- Van Davier, A.A., Deonovic, B.E., Yudelson, M., Polyak, S., Woo, A., 2019. Computational psychometrics approach to holistic learning and assessment systems. *Front. Educ.* (4), 69. <https://doi.org/10.3389/educ.2019.00069>.
- Vrzakova, H., Amon, M.J., Stewart, A., Duran, N.D., D'Mello, S.K., 2020. Focused or stuck together: multimodal patterns reveal triads' performance in collaborative problem solving. In: *Proceedings of the 10th International Conference on Learning Analytics & Knowledge*, pp. 295–304.
- Weeks, J.P., von Davier, M., Yamamoto, K., 2016. Using response time data to inform the coding of omitted responses. *Psychol. Test. Assess. Model.* 58, 671–701.
- Wijesooriya, C., Heales, J., Clutterbuck, P., 2015. Forms of formative assessment in virtual learning environments. In: *21st Americas Conference on Information Systems*, pp. 1–16.
- Wüstenberg, S., Greiff, S., Funke, J., 2012. Complex problem solving. More than reasoning? *Intelligence* 40, 1–14. <https://doi.org/10.1016/j.intell.2011.11.003>.
- Xu, H., Fang, G., Chen, Y., Liu, J., Ying, Z., 2018. Latent class analysis of recurrent events in problem-solving items. *Appl. Psychol. Meas.* 42 (6), 478–498. <https://doi.org/10.1177/0146621617748325>.
- Yamamoto, K., Lennon, M.L., 2018. Understanding and detecting data fabrication in large-scale assessments. *Qual. Assur. Educ.* 26, 196–212. <https://doi.org/10.1108/QAE-07-2017-0038>.

- Yang, F., Li, F.W., 2018. Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Comput. Educ.* 123, 97–108. <https://doi.org/10.1016/j.compedu.2018.04.006>.
- Zehner, F., Harrison, S., Eichmann, B., Deribo, T., Bengs, D., Anderson, N., Hahnel, C., 2020. The NAEP EDM competition: on the value of theory-driven psychometrics and machine learning for predictions based on log data. In: *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)*, pp. 302–312.
- Zheng, G., Fancsal, S.E., Ritter, S., Berman, S., 2019. Using instruction-embedded formative assessment to predict state summative test scores and achievement levels in mathematics. *J. Learn. Anal.* 6 (2), 153–174. <https://doi.org/10.18608/jla.2019.62.11>.
- Zhu, M., Shu, Z., von Davier, A.A., 2016. Using networks to visualize and analyze process data for educational assessment. *J. Educ. Meas.* 53 (2), 190–211. <https://doi.org/10.1111/jedm.12107>.
- Zoanetti, N., Griffin, P., 2017. Log-file data as indicators for problem-solving processes. In: Csapó, B., Funke, J. (Eds.), *The Nature of Problem Solving: Using Research to Inspire 21st Century Learning*. OECD Publishing. <https://doi.org/10.1787/9789264273955-13-en>.

Modeling of network structures

Tracy M. Sweet^a and Samrachana Adhikari^b, ^a Measurement, Statistics and Evaluation Program, Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, United States; and ^b Division of Biostatistics, Department of Population Health, New York University Grossman School of Medicine, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	430
Network notation and terminology	430
Social selection models	432
An example	433
Multilevel extensions to social selection models	434
Longitudinal extensions to social selection models	435
Social influence models	436
Multilevel and longitudinal extensions of social influence models	437
Modeling selection and influence	438
Discussion	438
References	439

Introduction

A network is defined as a set of ties or connections among a group of nodes or entities. In education, relational data or data with network structure can arise from a variety of situations. For example, the friendship relationships among a group of students can be represented by a network where the nodes are the students and the ties are the friendships among them. The nodes need not be students, they could be teachers, schools, or even actions on an assessment. Similarly, the ties or relationships could range from collaboration to romantic interest to even temporal sequences.

Data structure often informs the methods or models that should be used for analysis. For example, student observations nested within teachers or schools often require multilevel models to accommodate the hierarchical sampling design of selecting schools and then students for data collection. Similarly, repeated measures data generally require models that account for temporal dependence. Alternative models are required because model assumptions, such as independent and identical distribution of observations, of other models are not met.

Network data have quite complicated and often ill-defined dependence assumptions. Consider a simple network of children in an elementary classroom where the tie is defined as playing together at recess. Whether two children play together at recess is not necessarily independent from whether two other children play at recess. Consider three children—A, B, and C—where A and B play together. Whether A and C play together is dependent on whether A and B play together and whether B and C play together. In fact, whether A and C play also depends on how many other children are playing with A and how many other children are playing with C and if any of their mutual friends are playing together. As a result, general/generalized linear models that assume independent observations are not appropriate for these data.

Due to the complex interdependency among network ties, specialized models for network data are recommended. The purpose of this paper is to discuss some of the models that have been and could be used in education research. In the next section, we present network notation and terminology and briefly discuss exploratory analysis. Our presentation of network models is divided across the following sections, and we conclude with a discussion of future work.

Network notation and terminology

Consider a peer network of students in fifth grade in an elementary school. The network consists of the students, and the tie is defined to be friendship. The way in which this information is collected is often through a survey in which students nominate or list their friends. Other methods for network data collection include observations in which a third party observes which students interact or in which students list the relationships among all other students, including themselves.

Regardless of the data collection method, the resultant data generally include an ordered pair of students and whether they are friends for all possible pairs of students. One way to organize this information is through an adjacency matrix, an $n \times n$ matrix A such that A_{ij} represents the presence of friendship from node i to node j . Thus $A_{12} = 1$ indicates that node 1 nominated node 2 as a friend whereas $A_{34} = 0$ suggests that node 3 did not consider node 4 a friend.

Depending on the type of network tie, adjacency matrices may be symmetric or asymmetric. The relationship, being on the same soccer team, results in a symmetric adjacency matrix whereas something like romantic interest or even friendship could result in asymmetric matrices. The corresponding ties are considered undirected (soccer team) or directed (romantic interest). Furthermore, the values of the entries of A may also represent the nature of the relationships. A friendship adjacency matrix with binary entries

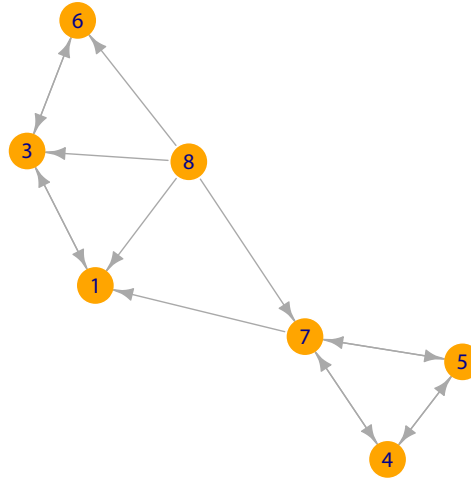


Fig. 1 The sociogram or network plot shows the ties (arrows) among a set of nodes (vertices).

represent the presence or absence of friendship whereas a continuous matrix could represent the strength of the relationship or frequency of interaction. For simplicity, the statistics and models that appear throughout this article assume binary asymmetric networks.

Network data are commonly visualized with a sociogram or network plot with nodes as vertices and ties indicated either by edges (symmetric ties) or arrows (asymmetric ties). As an example, consider the 8-node network given in [Fig. 1](#). The corresponding adjacency matrix is given as

$$A = \begin{pmatrix} \text{NA} & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{NA} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & \text{NA} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \text{NA} & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & \text{NA} & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & \text{NA} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & \text{NA} & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & \text{NA} \end{pmatrix}$$

where $A_{ij} = 1$ indicates a tie from i to j . Note that we write NA (or zero) along the diagonal to signal that nodes do not have ties to themselves.

Consider [Fig. 1](#) as one of many possible networks we might observe among these eight children. There are certain statistics used to characterize and summarize networks, which are calculated directly from the adjacency matrix. We review the most common statistics but recommend texts such as [Wasserman and Faust \(1994\)](#) or [Carolan \(2013\)](#) for a comprehensive list of descriptive statistics.

Let us assume the network in [Fig. 1](#) is a network of fifth grade friendships; there are 8 nodes in this network. Density is the proportion of observed ties given all possible ties defined as $\frac{\sum_{i,j} A_{ij}}{n(n-1)}$ where n is the number of nodes. Note that we do not allow nodes to have ties to themselves so that total number of all possible ties among n nodes is $n(n-1)$. The density of this network is $\frac{15}{56} = 0.27$.

Another statistic is reciprocity, the extent to which directed ties are reciprocated. For every tie from a to b , what proportion of those ties also has a tie from b to a ? Reciprocity can be calculated as the count of the number of bidirectional ties divided by the total number of observed ties; in our example network the reciprocity is 0.5. Transitivity is similar except that it enumerates the number of transitive ties instead of bidirectional ties. A transitive tie is a tie in either direction between a and c given ties in either direction among a and b and b and c .¹ The transitivity of our network is approximately 0.57.

¹There are additional formulas for reciprocity and transitivity. See [Wasserman and Faust \(1994\)](#).

Other common network statistics are at the node or edge level and focus on centrality. Degree is one measure of node centrality and is defined as the number of ties coming in and out of a specified node. For directed ties, in-degree is the number of ties pointing toward a node and out-degree is the number of ties pointing out from a node. In- and out-degrees for a node x are calculated as $\sum_i A_{ix}$ and $\sum_j A_{xj}$ respectively. In [Fig. 1](#), node 8 has an in-degree of 0 and an out-degree of 4 whereas node 5 has an in-degree of 2 and out-degree of 2. Many network statistics can be calculated easily using the `igraph` package ([Csardi and Nepusz, 2006](#)) which is available for R and Python.

Network statistics help researchers systematically characterize the networks of interest. A network that is highly reciprocal might indicate something either about the type of network tie, the individuals in the network, or some external effect on the network. Similarly, differences in density or transitivity or distributions of node degree help researchers better understand their networks and may inspire research questions about the networks they have observed.

Although network statistics are an important part of network analysis, they do not necessarily address why or how a network has certain features. Social scientists are often interested in why network ties form or dissolve and how these network ties impact individual outcomes such as behavior or opinions. In the next two sections, we present two types of social network models: social selection models and social influence models. Social selection models can help researchers quantify how networks form whereas social influence models quantify the relation between an existing network and node outcomes.

Social selection models

Consider our network of fifth grade students in a classroom. The network summaries (graphical and numerical) may reveal that nodes with similar attributes are more likely to be friends or that certain types of students are more likely to be nominated. A social selection model can formalize these patterns and quantify the extent to which node-level attributes predict network ties. That is, social selection models allow us to estimate the association of certain attributes such as gender and race with friendship.

To appropriately accommodate the inter-dependence among network ties, broadly two classes of social selection models have emerged in the literature: exponential random graph models (ERGMs; [Anderson et al., 1999](#)) and conditionally independent dyad (CID) models ([Dabbs et al., 2020](#)). Although these classes are not mutually exclusive, standard ERGMs and CID models generally differ in how they accommodate tie dependence.

ERGMs explicitly define the dependence structure by including network statistics in the model ([Frank and Strauss, 1986](#); [Anderson et al., 1999](#)), and which statistics are included aligns with the assumed dependence structure among ties. The term ERGM is based on the fact that the probability distribution of a graph can be written as an exponential family; the network statistics are the sufficient statistics for the distribution ([Goldenberg et al., 2009](#); [Holland and Leinhardt, 1981](#); [Robins et al., 2007](#)). Common examples of network statistics are reciprocity, density and transitivity, although more complicated network statistics are recommended for model estimation ([Snijders et al., 2006](#)).

Alternatively, CID models incorporate latent variables to account for tie dependence; ties are modeled as independent conditional on latent variables that explain the network structure. Common examples of the CID models include the latent space model ([Hoff et al., 2002](#)), the stochastic block model ([Holland et al., 1983](#)) and the mixed membership stochastic block model ([Airoldi et al., 2008](#)). Note that some of the earliest network models, such as the p_1 model ([Holland and Leinhardt, 1981](#)) and the p_2 model ([van Duijn et al., 2004](#)), could be classified as simple versions of either ERGMs or CID models. The p_1 and p_2 models account for dyad-level dependencies using sender effects, receiver effects, and reciprocity terms.

While ERGMs allow users to estimate the effects of network structures in their models, complex model specification can result in computational complications and convergence issues, thus making them slightly less desirable in real world applications ([Handcock et al., 2003](#)). While CID models do not have these estimation issues, many of these models have either some type of identifiability issue or label switching issue. Further, some researchers may be interested in precise estimations of latent variables, which requires some additional modeling constraints for comparable latent variable estimates ([Sweet and Adhikari, 2020](#); [Sweet and Zheng, 2017](#)).

Many R packages are available to fit social selection network models. For example, to fit ERGMs, there is an R package `ergm` ([Handcock et al., 2021](#)), which is part of a suite of packages in the `statnet` R package. Parameters are generally estimated using Markov chain Monte Carlo (MCMC; [Gelman et al., 2013](#)) although maximum likelihood estimation algorithms are also used, and alternatively pseudo-maximum likelihood estimation can be specified by the user. Within the `statnet` suite, there is also a package called `latentnet` ([Krivitsky and Handcock, 2020](#)) for fitting latent space models for some types of network data using MCMC. Additional specifications of latent space models for multiplex and bipartite networks can be fit using the `lvm4net` package ([Gollini, 2019](#)), which relies on variational inference for parameter estimation. Further, there are many R packages available for additional types of CID models and specific types of network data. For example, the `hlsm` ([Adhikari et al., 2014](#)) package is available for multi-level latent space models, the `NetMix` ([Olivella et al., 2021](#)) for mixed membership stochastic blockmodels, and `blockmodels` ([Leger et al., 2020](#)) for stochastic blockmodels.

Despite the number of available R packages, some network models are not accommodated in an existing R package. Users may choose to write their own MCMC algorithm to meet their needs. However, MCMC sampling software such as [Stan Development Team \(2016\)](#) and JAGS Plummer (2009) allows users without extensive coding experience to specify and quickly estimate a unique model.

An example

Although ERGMs and CID models are philosophically different, they are often used in similar ways in education research, namely to estimate the effect of node- and edge-level attributes on predicting network ties. The interpretation of the parameter estimates for these attributes is also similar. Therefore, we present only one type of model as an example.

A latent space model is a CID model where the latent variables are latent positions for each node in a low-dimensional latent social space. The probability of a tie is a function of the distance from one node to another node in the latent space. Pairs of nodes that are relatively far from each other are less likely to have a tie than nodes that are relatively close to each other.

Conditional on the latent space positions, network ties are independent of one another. Thus, one assumption is that the latent space explains network structure. For example, three nodes with mutual ties are expected to all be close to one another in the latent space, and a node without any ties to other nodes has a latent position far in the latent space.

A latent space model is given as

$$\begin{aligned} \text{logit}[P(A_{ij} = 1)] &= X_{ij}\beta - d(Z_i, Z_j) \\ Z_i &\sim \text{Normal}(0, \Sigma) \end{aligned} \quad (1)$$

where the metric $d(Z_i, Z_j)$ can be any distance metric such as Euclidean distance. A standard prior for Z_i is a d -dimensional vector from a normal distribution with mean 0, and Σ is a diagonal variance-covariance matrix. Thus, the probability of a tie from i to j is a function of covariates X and their parameters β as well as the distance between Z_i and Z_j in the latent space. In the presence of covariates, the latent space positions represent residual network structure not explained by the covariates.

Social science researchers are generally most interested in the regression coefficients for covariates $X=(X_1, X_2, \dots, X_p)$, which are typically attributes about the nodes or the edges among them. In a directed network, a covariate could be information about tie sender (nominator) or the tie receiver (nominated); otherwise a node-covariate is used for an undirected network. Examples of node-level covariates include age, gender, and race. An edge-covariate relates to the pair of nodes or dyad itself. Analogous examples include difference in age, being the same gender or race.

Consider the following toy example in Fig. 3 of an undirected friendship network of 15 fifth grade students for whom we have their choice of favorite recess activity. These categories are arbitrary so let us assume that nodes 1–4 enjoy sports, 5–8 enjoy playground equipment, 9–10 enjoy art, and 11–15 enjoy pretend play. The nodes in the network plot are colored based on these categories.

The ties in Fig. 2 are undirected so we do not have sender and receiver specific variables, but we can include node-level covariates and edge-level covariates when modeling undirected networks. For example, suppose X_{1ij} is indicator for whether the node i 's favorite recess activity is sports. Then we are interested in whether students who love sports are more or less likely to have friends. An edge-level covariate might instead focus on whether i and j enjoy the same favorite activity; that is, $X_{2ij} = 1$ indicates i and j have the same favorite recess activity. Other edge-level variables could be difference in age or having a tie at a previous time.

If the ties in Fig. 2 were directed, we could have sender and receiver specific covariates; X_{1ij} might indicate that the sender's favorite activity is sports, and we might be interested in whether sports players are more likely to nominate others as friends than other types of students. Alternatively, X_{1ij} could indicate that the receiver's favorite activity is sports so we are now interested if sports players are more popular (more likely to be nominated as a friend).

For our example, Fig. 2 suggests that students with the same favorite recess activity are more likely to be friends because nodes of the same color tend to share ties, and nodes of different colors do not share many ties. We could then fit the LSM given in Eq. (1) where X is a single covariate, such that $X_{ij} = 1$ if students i and j selected the same favorite recess activity.

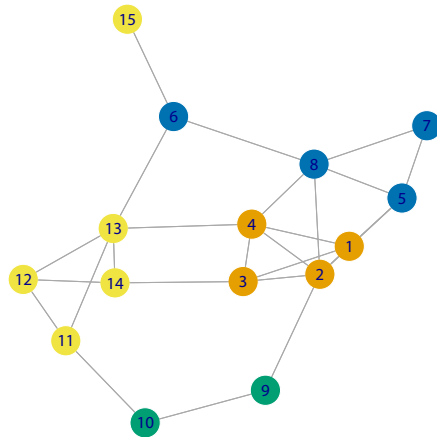


Fig. 2 Example of a 15-node friendship network based on favorite recess activity; Nodes 1–4 enjoy sports, 5–8 enjoy playground, 9–10 enjoy art, and 11–15 enjoy pretend play.

of fifth grade networks within each school as distinct and independent. Similar to how multilevel regression models model the variability both within and between schools, multilevel network models account for variability in tie probability within and between networks.

Multilevel social selection models have been used to examine the associations of teacher attributes on advice-seeking network ties. Research has found positive effects of being a teacher leader (Spillane et al., 2012) and being within relatively close proximity to each other (Spillane et al., 2017) on advice-seeking ties. There were negative effects of teaching English Language Learner students (Hopkins et al., 2015) and being of a racial minority (Bristol and Shirrell, 2019) on advice-seeking ties. These models have also been used to explore relationships among students; Frank et al. (2013) explored the associations with several covariates and found being in the same class predicted new friendships among adolescents.

There are multilevel extensions for both ERGMs and CID models although multilevel CID models have been used more often in the education literature. For that reason and for brevity, we will limit our discussion to those models. Snijders and Kenny (1999) first introduced a multilevel social selection model for family relationship data where the network was defined by the family; this model included sender, receiver, and dyad random effects. Zijlstra et al. (2006) introduced a similar model, the multilevel p_2 model, and Sweet et al. (2013) introduced a multilevel extension to the latent space model.

A multilevel network model assumes networks are independent and isolated; that is ties across networks do not exist. A multilevel latent space model is given as

$$\begin{aligned} \text{logit}[P(A_{ijk} = 1)] &= X_{ijk}\beta_k - d(Z_{ik}, Z_{jk}) \\ Z_{ik} &\sim \text{Normal}(0, \Sigma) \end{aligned} \quad (2)$$

where A_{ijk} now denotes the presence of a tie from node i to node j in network k . As in other random effects models, regression coefficients β_k may vary across the networks. Note that the latent space positions are different for each network, and the distribution for each network's latent position could also vary. Finally, X_{ijk} may represent a sender-, receiver-, edge- or network-level covariate.

Multilevel social selection models both accommodate the nested data structure common in education, and also allow researchers to explore differences among networks. For example, the effect of choosing the same favorite recess activity may vary across schools, and understanding the extent to which this effect varies might be of interest. In addition, the models also accommodate network-level variables, so examining which school-level variables explain some of the variability observed across networks is also relevant. Furthermore, network-level variables include covariates such as a experimental condition, and these models can be used to estimate treatment effects in cluster randomized trials (Sweet et al., 2013).

Longitudinal extensions to social selection models

Longitudinal data is an example of a nested data structure in which repeated observations are collected over time for each unit. Network data can also have a longitudinal structure. For example, a friendship network of students at a school can be collected over multiple years, and these network ties are correlated over time. If two students were friends in the past they are also likely to be friends in the future. Thus, these correlations in longitudinal network data need to be accounted for in social selection models. The longitudinal network models account for such correlation over time and are similar in structure to the multilevel network models.

Both ERGMs and CID models have been extended to model longitudinal network data. Using models called temporal ERGMs, Hanneke et al. (2010) have studied the networks observed over discrete time points. Temporal ERGMs generally condition on the network at time $t - 1$ to model the network at time t , but similar to the ERGMs for cross-sectional network data temporal ERGMs also use network statistics in the model to define the within-network tie dependence. Examples of network statistics can include statistics representing stability, density, and overall reciprocity, observed in the networks at time t and $t - 1$. Other examples of longitudinal ERGMs include Robins and Pattison (2001) and Almqvist and Butts (2013).

CID models have modeled networks observed over multiple time points by introducing longitudinal dependence on the latent variables (e.g., Sewell and Chen (2015), Westveld and Hoff (2011), Xing et al. (2010)). Longitudinal CID models also condition on the network at time $t - 1$ to model the network at time t . And analogous to cross-sectional CID models, the network ties at each time point are assumed to be independent given the latent variables at that time point and in the previous time points.

For simplicity and consistency, we present the longitudinal version of the LSM. Sarkar and Moore (2005) introduced a predictive latent space model for temporal networks, and Sewell and Chen (2015) have similar work extending the latent space model to account for longitudinal networks. A longitudinal latent space model for networks $A_1, \dots, A_T$ observed over T time points, can be written as

$$\begin{aligned} \text{logit}[P(A_{ijt} = 1)] &= X_{ijt}\beta_t - d(Z_{it}, Z_{jt}) \\ Z_{it} &\sim \text{Normal}(Z_{i(t-1)}, \Sigma), \text{ for } t = 2, \dots, T \\ Z_{i1} &\sim \text{Normal}(0, \Sigma) \end{aligned} \quad (3)$$

where A_{ijt} now denotes the presence of a tie from node i to node j at time t . The longitudinal correlation in the networks is accounted for by a random walk plus noise dependence in the latent positions Z_{it} over time. We could further extend this modeling framework to introduce dependence on the coefficients ($\beta_1, \dots, \beta_T$) for each covariate, analogous to time-varying growth models. The idea behind longitudinal LSMs is that nodes that are close in latent positions in previous time points are more likely to also be close in future, and thus, nodes that have ties in the previous time points have higher likelihood to have ties also in the future.

We also note that longitudinal network data has an additional complication in that the nodes in the network may change over time. Immigration of individuals into a network and emigration out of a network is common in educational data. For example, networks of students are likely to vary over years as students move or graduate; similarly, teacher turnover is quite common in schools between teachers leaving the profession, transferring to new schools, and new teachers graduating. Some longitudinal social selection models accommodate node immigration and emigration (Adhikari et al., 2021).

Social influence models

Social influence refers to the phenomenon where an individual's behaviors, opinions, or beliefs change as a result of their network ties, often becoming more similar to those with whom they are connected. Social influence in education is quite common, from students learning dances from each other (Marshall, 2019) and adolescents spending money on items that their peers would see (Bachmann et al., 1993) to more rebellious behaviors such as smoking or drug use (Ennett et al., 2008). Students are not the only individuals susceptible to influence; teachers and other staff members can be professionally influenced by one another (Supovitz et al., 2010), and school leaders may be influenced by the policies and practices implemented in other schools (Muncey and McQuillan, 1996).

Thus, social influence models aim to estimate the effect of these network ties on some outcome of interest, these models incorporate network ties as explanatory variables rather than outcomes. In our context of fifth grade students, we might be interested in whether interest in coding is influenced by friendship ties. We then hypothesize that students will become more interested in coding if they are friends with someone with a strong interest in coding; likewise, students with friends who are very uninterested in coding may be less inclined to participate in coding.

In the social network literature, there are a variety of models for social influence. One prominent model is called the network autocorrelation model (Doreian, 1989; Leenders, 2002), which is similar to the autologistic actor attribute model (ALAAM; Darganova and Robins, 2013; Robins et al., 2001). The autocorrelation results from the outcome of interest being modeled as a function of that outcome and the network. Consequently, autocorrelation models make the assumption that the network has reached a state of equilibrium, which is a strong assumption and may not hold in many educational situations.

A temporal version of these models relaxes this assumption, and these versions are called exposure models (Valente, 2005), temporal versions of network autocorrelation models (Leenders, 2002) or social influence models. These models use a temporal relationship such that an outcome variable Y at time t is predicted by the network ties and the nodes' values of Y at time $t - 1$. This model is given as;

$$Y^t = \alpha WY^{t-1} + X\beta + \varepsilon; \quad (4)$$

where the continuous outcome Y is measured at two time points, W is a function of the network, X is a set of additional covariates, α is the estimate of network influence, and ε is random noise. In many applications W is a weight based on the network ties. For example, if nodes are influenced by those with whom they are connected, then W_i could be defined as a vector indicating with which nodes i is connected and then scaled by the number of ties i has, $W_{ij} = \frac{A_{ij}}{A_i}$. Note that in directed networks, W could be defined in terms of in-ties or out-ties. Other specifications of W certainly exist; Sweet and Adhikari (2020) proposed an influence model in which node influence is based on their relative latent space position.

Consider the following example of the toy network given in Fig. 4 (top left) in which the nodes are colored based on their interest in coding. The colors shown, going from light yellow to peach to red to magenta indicate the interval from no interest to strong interest at time 1. The other plots in Fig. 4 illustrate how interest in coding may change based on our assumption about social influence. If we assume nodes are only influenced by those nodes they nominate (out ties), then Fig. 4 (top right) shows the resulting node attributes at time 2. Under this assumption, nodes 1 and 8 become less interested in coding, for example.

Another assumption is that nodes are influenced by all of their ties, regardless of direction, and Fig. 4 (bottom left) shows that node 1 is now influenced by both node 3 and node 8, and their interest in coding no longer decreases. Finally, in the plot given in Fig. 4 (bottom right), we assume nodes are influenced by all of their ties but also by themselves; that is, now $X = Y^{t-1}$ in Eq. (4) as opposed to $X = 0$ in the other examples. Under this assumption, node 8 has the smallest decrease in coding interest compared to our other assumptions.

Which mechanism of social influence one employs in the model depends on the context, including both the type of tie and the outcome of interest. Including a time 1 node attribute in the model is standard, but selecting whether nodes are influenced by in-ties, out-ties, or both depends on the context. Nodes are likely most influenced by the nodes they nominate (out-ties) in relationships such as advice-seeking or admiration whereas relationships such as friendship or information-sharing tend to be more reciprocal in terms of influence.

In the education research literature, there have been several studies investigating social influence among students. For example, Broda et al. (2018) fit social influence models similar to Eq. (4) to estimate the impact of networks on supportive practices for student college admission; DiGuseppi et al. (2018) fit network autocorrelation models to study resistance to binge drinking; and Huang et al. (2014) found that student exposure to risk behaviors through online friendship networks was associated with their own risk behaviors.

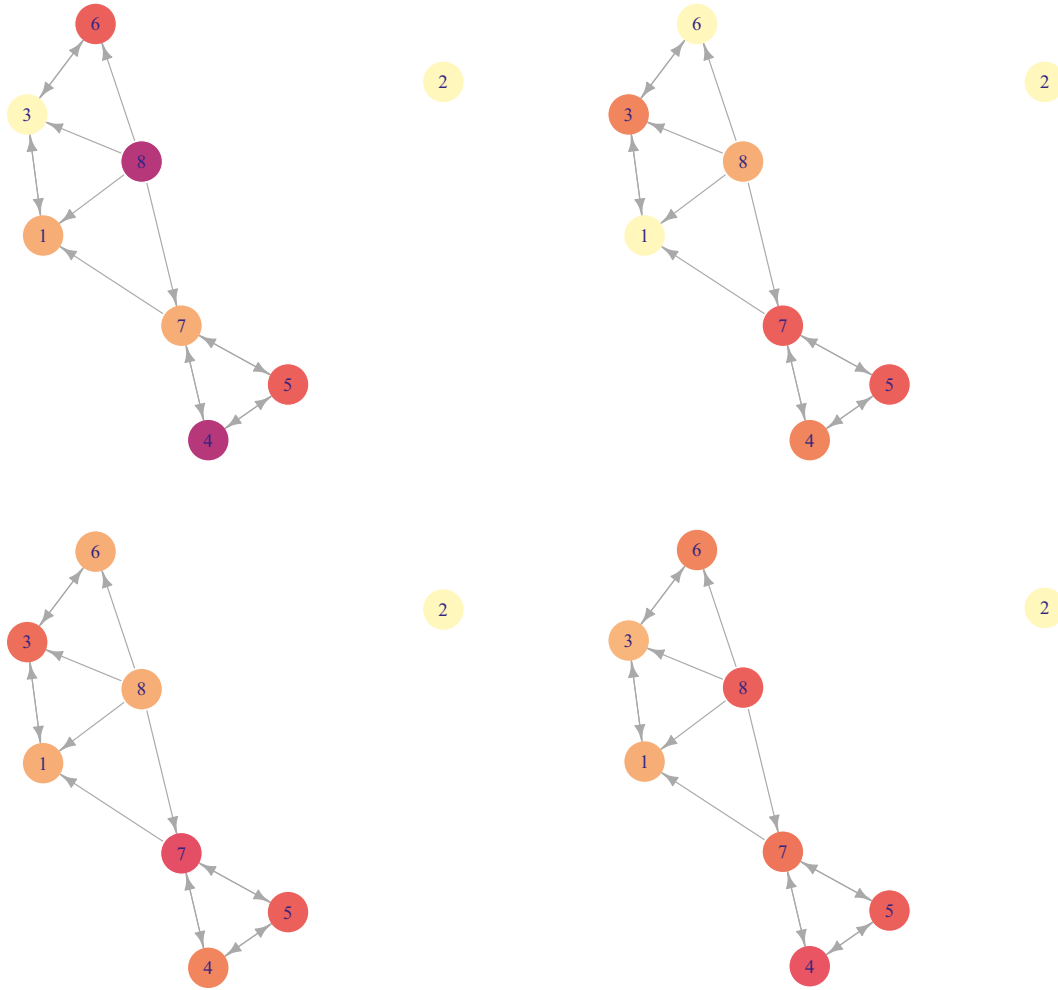


Fig. 4 Changes in node color illustrate changes in node attributes as a result of social influence where assumptions of influence may vary. Ties may be influenced by out ties, all ties, or all ties as well as themselves.

Multilevel and longitudinal extensions of social influence models

Social influence studies in education often require a multilevel version of Eq. (4) because these studies focus on networks of students nested in classrooms/schools or networks of teachers nested in schools. A multilevel social influence model is given as

$$Y_k^t = \alpha_k W_k Y_k^{t-1} + X_k \beta_k + \varepsilon, \quad (5)$$

where k indexes the network. The outcome of interest is still measured at two time points and the ties in network k are used to determine W_k . Eq. (5) may include fixed or random regression coefficients.

Multilevel social influence models are quite prevalent among analyses of teacher network data in education. Frank et al. (2004) and Frank et al. (2014) employed multilevel social influence models to explore teacher social influence on the use of technology and instructional practices respectively. Penuel et al. (2012) studied network effects on professional development for writing instruction. There are few examples of multilevel influence models of students, but Cooc and Kim (2017) found that reading with someone with a higher reading ability score was associated with improved reading ability.

Longitudinal social influence models also exist. O'Malley et al. (2020) extended the social influence model for longitudinal networks and longitudinal outcome data. The core model is given as

$$Y_{it} = \beta_1 Y_{i(t-1)} + X_{i(t-1)} \beta_2 + \alpha W_{(t-1)} Y_{i(t-1)} + \varepsilon_{it}, \quad (6)$$

where the weight matrix W_t at time t is derived from the weights based on the network at time t so that W_t also changes over time. Note that the model assumes the influence effect α is constant over time. O'Malley et al. (2020) also assumes Eq. (6) to be a Markov transition model such that the time between observations at time $t - 1$ and t is assumed to be sufficiently small such that any change

in the outcome of node i has yet to impact those of nodes $j \neq i$. A similar model in which W_t was defined as the weight from out-ties in the network at time t was used by [Spillane et al. \(2018\)](#) to estimate the effect of longitudinal network ties on changes in beliefs about teaching mathematics.

Another version of the longitudinal influence model given in [Eq. \(6\)](#) is to include an interaction term to estimate moderation by network characteristics; [O'Malley et al. \(2020\)](#) included a term in their model to estimate how peer effects are moderated by how important/focal a node is in the network. For example, in an advice-seeking network, leadership status could moderate the influence of collaboration between teachers on their students' performance.

Modeling selection and influence

One criticism of modeling cross-sectional data is that we are often unable to disentangle social selection from social influence to make a causal inference on the network. That is, if two children are friends and they both like playing sports at recess, are they friends because they both like playing sports or do they both like playing sports because they are friends? One solution is to use longitudinal data, because it provides a temporal constraint on how ties can impact nodal attributes and vice versa ([Frank and Xu, 2020](#)). Friendship at time 1 can be used to predict changes in recess activities at time 2, and favorite recess activities at time 2 can be modeled as predictors of new friendships at time 3.

Alternatively, co-evolution models involving stochastic actor oriented models ([Snijders, 1996](#)) model the co-evolution processes of selection and influence over time ([Steglich et al., 2010](#)). These models view network evolution and influence as a co-evolving system, and the selection and influence processes are separated by modeling these two phenomena jointly with sequence of longitudinal observations. These models are fit using the software Simulation Investigation for Empirical Network Analysis (SIENA; [Snijders et al., 2008](#)), and they require longitudinal network data for at least two consecutive waves. One assumption is that ties evolve between waves as a result of nodal attributes. SIENA is a flexible software that allows for small amounts of missing data as well as emigration and immigration of nodes (e.g., students leaving the school or new students entering the school).

Research applying SIENA have mainly been used in the psychological and public health literature that involve networks of students in schools. For example, both [Wölfer and Scheithauer \(2014\)](#) and [Rambaran et al. \(2020\)](#) studied the social influence of bullies; [Osgood et al. \(2013\)](#) found selection and social influence effects of alcohol use in friendship networks; and [Rivas-Drake et al. \(2017\)](#) found selection and social influence effects of ethnic-racial identity. Co-evolution models have great potential for education research when longitudinal data is available.

Discussion

The organization of students and teachers in schools has created opportunities for networks since group learning requires some kind of interaction. Similarly, schools provide more than simply academic instruction; student interactions are also important for the development of social and emotional competencies. Therefore, understanding why certain individuals connect and how those connections impact other outcomes is important for education research. By highlighting classes of network models, our aim was to showcase models commonly used in education research. Social selection models focus on associations between nodal attributes and network ties, and social influence models focus on the effects of the network on nodal outcomes.

We have focused on networks of students or teachers because these networks are the most prevalent in education research, but network nodes need not be individuals. For example, nodes could be items on an assessment (e.g., [Jin and Jeon, 2019](#)), and we might be interested in modeling the co-occurrence of correct items. Alternatively, an item may require several steps or actions for a correct answer; one could think of the possible actions as nodes and the temporal sequence of steps from one action to the next as the network ([Zhu et al., 2016](#)). Measurement models for network data are an emerging area, and we look forward to future developments. A related field is that of network psychometrics ([Epskamp et al., 2017](#); [Marsman et al., 2018](#)), in which the relations among these items or the constructs represented by these items are modeled using a graphical modeling approach. These models are fairly different from the network models described in this manuscript. Rather than using an observed network as a predictor or outcome for inference; the goal is rather to infer the network, defined as the connections among the nodes/constructs.

Furthermore, we have focused on network models common in education research, but this article is not a comprehensive survey of network models more generally. Methodological research on networks exists throughout a variety of disciplines, including research with applications in computer science, statistics, medicine, and public health. Some recent methodological work includes methods for missing data ([Krause et al., 2020](#)) and measurement error ([Lee, 2019](#)) along with new models for co-evolution ([Minhas et al., 2016](#)) and mediation ([Liu et al., 2021](#); [Sweet, 2019](#)). Social networks have even appeared in the causal inference literature ([Lee and Ogburn, 2020](#)).

With respect to education research, there are several areas of future research in network modeling that should be particularly impactful. Future work on missing data will be useful for modeling teacher networks which tend to gain/lose a large percentage of its nodes over time, making longitudinal modeling difficult. Similarly, additional research on mediation models for social networks will help researchers employ these models to analyze experimental intervention data. Finally, there is an area of research

around spillover effects or network interference, that is, the idea that someone in a different experimental condition could be impacted due to their relationships with individuals in another condition. Network models have potential to be useful in this area of research as well.

References

- Adhikari, S., Junker, B.W., Sweet, T.M., Thomas, A.C., 2014. HLSM: Hierarchical Latent Space Network Model. Carnegie Mellon University r package version 3.0.0.
- Adhikari, S., Sweet, T., Junker, B., 2021. Analysis of longitudinal advice-seeking networks following implementation of high stakes testing. *J. Roy. Stat. Soc.* 184 (4).
- Airoldi, E.M., Blei, D.M., Fienberg, S.E., Xing, E.P., 2008. Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* 9, 1981–2014.
- Almqvist, Z.W., Butts, C.T., 2013. Dynamic network logistic regression: a logistic choice analysis of inter-and intra-group blog citation dynamics in the 2004 us presidential election. *Polit. Anal.* 21, 430–448.
- Anderson, C.J., Wasserman, S., Crouch, B., 1999. A p* primer: logit models for social networks. *Soc. Network.* 21, 37–66.
- Bachmann, G.R., John, D.R., Rao, A.R., 1993. Children's Susceptibility to Peer Group Purchase Influence: An Exploratory Investigation. *ACR North American Advances*.
- Bristol, T.J., Shirrell, M., 2019. Who is here to help me? The work-related social networks of staff of color in two mid-sized districts. *Am. Educ. Res. J.* 56, 868–898.
- Broda, M., Ekholm, E., Schneider, B., Hutton, A.C., 2018. Teachers' social networks, college-going practices, and the diffusion of a school-based reform initiative. *SAGE Open* 8, 2158244018817397.
- Carolan, B.V., 2013. *Social Network Analysis and Education: Theory, Methods & Applications*. SAGE Publications.
- Cocc, N., Kim, J.S., 2017. Peer influence on children's reading skills: a social network analysis of elementary school classrooms. *J. Educ. Psychol.* 109, 727.
- Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. *Int. J. Complex Syst.* 1695.
- Dabbs, B., Adhikari, S., Sweet, T., 2020. Conditionally Independent Dyads (CID) network models: a latent variable approach to statistical social network analysis. *Soc. Network.* 63, 122–133.
- Daragonova, G., Robins, G., 2013. Autologistic actor attribute models. In: Lusher, D., Koskinen, J., Robins, G.L. (Eds.), *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*. Cambridge University Press, Cambridge, MA, pp. 102–114.
- DiGuseppi, G.T., Meisel, M.K., Balestrieri, S.G., Ott, M.Q., Cox, M.J., Clark, M.A., Barnett, N.P., 2018. Resistance to peer influence moderates the relationship between perceived (but not actual) peer norms and binge drinking in a college student social network. *Addict. Behav.* 80, 47–52.
- Doreian, P., 1989. Network autocorrelation models: problems and prospects. In: Griffith, D. (Ed.), *Spatial Statistics: Past, Present, Future*. Michigan Document Services, Ann Arbor, MI, pp. 369–393.
- Ennett, S.T., Faris, R., Hipp, J., Foshee, V.A., Bauman, K.E., Hussong, A., Cai, L., 2008. Peer smoking, other peer attributes, and adolescent cigarette smoking: a social network analysis. *Prev. Sci.* 9, 88–98.
- Epskamp, S., Rhemtulla, M., Borsboom, D., 2017. Generalized network psychometrics: combining network and latent variable models. *Psychometrika* 82, 904–927.
- Frank, K.A., Lo, Y.-J., Sun, M., 2014. Social network analysis of the influences of educational reforms on teachers practices and interactions. *Zeitschrift für Erziehungswissenschaft* 17, 117–134.
- Frank, K.A., Muller, C., Mueller, A.S., 2013. The embeddedness of adolescent friendship nominations: the formation of social capital in emergent network structures. *Am. J. Sociol.* 119, 216–253.
- Frank, K.A., Xu, R., 2020. Causal inference for social network analysis. In: Moody, J., Light, R. (Eds.), *Oxford Handbook of Social Network Analysis*. Oxford University Press, pp. 288–310.
- Frank, K.A., Zhao, Y., Borman, K., 2004. Social capital and the diffusion of innovations within organizations: the case of computer technology in schools. *Sociol. Educ.* 77, 148–171.
- Frank, O., Strauss, D., 1986. Markov graphs. *J. Am. Stat. Assoc.* 81, 832–842.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2013. *Bayesian Data Analysis*, third ed. Taylor & Francis, Boca Raton, FL.
- Goldenberg, A., Zheng, A., Fienberg, S., Airoldi, E., 2009. A survey of statistical network models. *Found. Trends Mach. Learn.* 2, 129–133.
- Gollini, I., 2019. *lvm4net: Latent Variable Models for Networks*. r package version 0.3.
- Handcock, M.S., Hunter, D.R., Butts, C.T., Goodreau, S.M., Krivitsky, P.N., Morris, M., 2021. *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks*. The Statnet Project, r package version 4.1.2.
- Handcock, M.S., Robins, G., Snijders, T., Moody, J., Besag, J., 2003. Assessing degeneracy in statistical models of social networks. *Tech. Rep. Citeseer*. 76, 33–50.
- Hanneke, S., Fu, W., Xing, E.P., 2010. Discrete temporal models of social networks. *Electron. J. Stat.* 4, 585–605.
- Hoff, P.D., Raftery, A.E., Handcock, M.S., 2002. Latent space approaches to social network analysis. *J. Am. Stat. Assoc.* 97, 1090–1098.
- Holland, P.W., Leinhardt, S., 1981. An exponential family of probability distributions for directed graphs. *J. Am. Stat. Assoc.* 76, 33–65.
- Holland, P.W., Laskey, K.B., Leinhardt, S., 1983. Stochastic blockmodels: first steps. *Soc. Network.* 5, 109–137.
- Hopkins, M., Lowenhaupt, R., Sweet, T.M., 2015. Organizing instruction in new immigrant destinations: district infrastructure and subject-specific school practice. *Am. Educ. Res. J.* 52, 408–439.
- Huang, G.C., Unger, J.B., Soto, D., Fujimoto, K., Pentz, M.A., Jordan-Marsh, M., Valente, T.W., 2014. Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. *J. Adolesc. Health* 54, 508–514.
- Hunter, D., Goodreau, S., Handcock, M., 2008. Goodness of fit of social network models. *J. Am. Stat. Assoc.* 103, 248–258.
- Jin, I.H., Jeon, M., 2019. A doubly latent space joint model for local item and person dependence in the analysis of item response data. *Psychometrika* 84, 236–260.
- Krause, R.W., Huisman, M., Steglich, C., Snijders, T., 2020. Missing data in cross-sectional networks—an extensive comparison of missing data treatment methods. *Soc. Network.* 62, 99–112.
- Krivitsky, P.N., Handcock, M.S., 2020. *latentnet: Latent Position and Cluster Models for Statistical Networks*. The Statnet Project, r package version 2.10.5.
- Lee, W., 2019. *Latent Variable Models for Indirectly or Imprecisely Measured Networks*. Ph.D. thesis. University of Washington.
- Lee, Y., Ogburn, E.L., 2020. Network dependence can lead to spurious associations and invalid inference. *J. Am. Stat. Assoc.* 1–15.
- Leenders, R.T.A., 2002. Modeling social influence through network autocorrelation: constructing the weight matrix. *Soc. Network.* 24, 21–47.
- Leger, J.-B., Barbillon, P., Chiquet, J., 2020. Blockmodels: Latent and Stochastic Block Model Estimation by a "V-EM" Algorithm. r package version 1.1.4.
- Liu, H., Jin, I.H., Zhang, Z., Yuan, Y., 2021. Social network mediation analysis: a latent space approach. *Psychometrika* 86, 272–298.
- Marshall, W., 2019. Social dance in the age of (anti-) social media: fortnite, online video, and the jock at a virtual crossroads. *J. Popular Music Stud.* 31, 3–15.
- Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., Van Bork, R., Waldorp, L.J., van der Maas, H.L.J., Maris, G., 2018. An introduction to network psychometrics: relating Ising network models to item response theory models. *Multivariate Behav. Res.* 53, 15–35.
- Minhas, S., Hoff, P.D., Ward, M.D., 2016. A new approach to analyzing coevolving longitudinal networks in international relations. *J. Peace Res.* 53, 491–505.
- Muncey, D.E., McQuillan, P.J., 1996. Reform and Resistance in Schools and Classrooms: An Ethnographic View of the Coalition of Essential Schools. Yale University Press.
- Olivella, S., Lo, A., Pratt, T., Imai, K., 2021. *NetMix: Dynamic Mixed-Membership Network Regression Model*. r package version 0.2.0.
- O'Malley, A.J., Moen, E.L., Bynum, J.P., Austin, A.M., Skinner, J.S., 2020. Modeling peer effect modification by network strength: the diffusion of implantable cardioverter defibrillators in the US hospital network. *Stat. Med.* 39, 1125–1144.

- Osgood, D.W., Ragan, D.T., Wallace, L., Gest, S.D., Feinberg, M.E., Moody, J., 2013. Peers and the emergence of alcohol use: influence and selection processes in adolescent friendship networks. *J. Res. Adolesc.* 23, 500–512.
- Penuel, W.R., Sun, M., Frank, K.A., Gallagher, H.A., 2012. Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. *Am. J. Educ.* 119, 103–136.
- Plummer, M., 2009. JAGS: Just Another Gibbs Sampler. Version 1.0. 3, Manual.
- Rambaran, J.A., Dijkstra, J.K., Veenstra, R., 2020. Bullying as a group process in childhood: a longitudinal social network analysis. *Child Dev.* 91, 1336–1352.
- Rivas-Drake, D., Umaña-Taylor, A.J., Schaefer, D.R., Medina, M., 2017. Ethnic-racial identity and friendships in early adolescence. *Child Dev.* 88, 710–724.
- Robins, G., Pattison, P., 2001. Random graph models for temporal processes in social networks. *J. Math. Sociol.* 25, 5–41.
- Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007. An introduction to exponential random graph (p^*) models for social networks. *Soc. Network.* 29, 173–191 special Section: Advances in Exponential Random Graph (p^*) Models.
- Robins, G.L., Pattison, P.E., Elliott, P., 2001. Network models for social influence processes. *Psychometrika* 66, 161–189.
- Sarkar, P., Moore, A.W., 2005. Dynamic social network analysis using latent space models. *ACM SIGKDD Explor. Newsl.* 7, 31–40.
- Sewell, D.K., Chen, Y., 2015. Latent space models for dynamic networks. *J. Am. Stat. Assoc.* 110, 1646–1657.
- Snijders, T.A.B., 1996. Stochastic actor-oriented models for network change. *J. Math. Sociol.* 21, 149–172.
- Snijders, T.A.B., Kenny, D., 1999. The social relations model for family data: a multilevel approach. *Pers. Relat.* 6, 471–486.
- Snijders, T.A.B., Steglich, C.E., Schweinberger, M., Huisman, M., 2008. Manual for SIENA Version 3.2. University of Groningen: ICS, Department of Sociology, Groningen, The Netherlands.
- Snijders, T.A.B., Pattison, P.E., Robins, G.L., Handcock, M.S., 2006. New specification for exponential random graph models. *Socio. Methodol.* 36, 99–153.
- Spillane, J.P., Hopkins, M., Sweet, T.M., 2018. School district educational infrastructure and change at scale: teacher peer interactions and their beliefs about mathematics instruction. *Am. Educ. Res. J.* 55, 532–571. <https://doi.org/10.3102/0002831217743928>.
- Spillane, J.P., Kim, C.M., Frank, K.A., 2012. Instructional advice and information providing and receiving behavior in elementary schools exploring tie formation as a building block in social capital development. *Am. Educ. Res. J.* 119, 72–102.
- Spillane, J.P., Shirrell, M., Sweet, T.M., 2017. The elephant in the schoolhouse: the role of propinquity in school staff interactions about teaching. *Sociol. Educ.* <https://doi.org/10.1177/0038040717696151>.
- Stan Development Team, 2016. Stan Modeling Language User's Guide and Reference Manual. version 2.9.0 ed.
- Steglich, C., Snijders, T.A., Pearson, M., 2010. Dynamic networks and behavior: separating selection from influence. *Socio. Methodol.* 40, 329–393.
- Supovitz, J., Sirinides, P., May, H., 2010. How principals and peers influence teaching and learning. *Educ. Adm. Q.* 46, 31–56.
- Sweet, T., Adhikari, S., 2020. A latent space network model for social influence. *Psychometrika* 1–24.
- Sweet, T.M., 2019. Modeling social networks as mediators: a mixed membership stochastic blockmodel for mediation. *J. Educ. Behav. Stat.* 44, 210–240.
- Sweet, T.M., Thomas, A.C., Junker, B.W., 2013. Hierarchical network models for education research: hierarchical latent space models. *J. Educ. Behav. Stat.* 38, 295–318.
- Sweet, T.M., Zheng, Q., 2017. A mixed membership model-based measure for subgroup integration in social networks. *Soc. Network.* 48, 169–180.
- Valente, T.W., 2005. Network models and methods for studying the diffusion of innovations. In: Carrington, P., Scott, J., Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. Cambridge University Press, New York, NY, pp. 98–116.
- van Duijn, M., Snijders, T., Zijlstra, B., 2004. p2: a random effects model with covariates for directed graphs. *Statistica Neerlandica* 58, 234–254.
- Wasserman, S., Faust, K., 1994. *Social Network Analysis: Methods and Applications*, vol. 8. Cambridge university press.
- Westveld, A.H., Hoff, P.D., 2011. A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict. *Ann. Appl. Stat.* 843–872.
- Wölfer, R., Scheithauer, H., 2014. Social influence and bullying behavior: intervention-based network dynamics of the fairplayer.manual bullying prevention program. *Aggress. Behav.* 40, 309–319.
- Xing, E.P., Fu, W., Song, L., et al., 2010. A state-space mixed membership blockmodel for dynamic network tomography. *Ann. Appl. Stat.* 4, 535–566.
- Zhu, M., Shu, Z., von Davier, A.A., 2016. Using networks to visualize and analyze process data for educational assessment. *J. Educ. Meas.* 53, 190–211.
- Zijlstra, B., van Duijn, M., Snijders, T., 2006. The multilevel p 2 model. *Methodology* 2, 42–47.

Certification and licensure testing

Richard M. Luecht, University of North Carolina at Greensboro, Greensboro, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Practice analysis: determining what to assess	441
Designing the test items: how to assess the KSAJs	443
Designing an operational testing system for processing the examinations	444
Scores and score scales: construction, maintenance, and interpretation	445
Determining the credentialing cut score	446
Efficient test design and form assembly	447
Credentialing classification accuracy and validation	448
Conclusions	449
References	449

The *Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 2014) combine certification and licensure under the heading of *credentialing* examinations. They are certainly similar in some ways, but they are not one and the same (Clauser et al., 2006). According to the United State Bureau of Labor Statistics (USBLs), forty-seven percent of all civilian workers require some type of credential, with 6% requiring a specific certification and almost 20% holding a professional license. Professionals in health care and technical occupations account for almost 95% of all required credentials (USBL, 2021).

Certifying organizations tend to be non-governmental agencies that grow out of professional societies or professional interest groups. Most certifications result in titles or low-to-moderate-stakes credentials issued as part of a voluntary process. That is, the certification candidates choose to undergo the evaluative process to acquire a particular badge, title or credential. Each certificate provides official verification that an individual has met a specified set of standards of demonstrated knowledge, skills or/other professional competencies. The credentialing process can range from some informal verification that the candidate completed specialized or continuing education to more formal and extensive evaluation process that includes meeting eligibility qualifications and undergoing one or more competency assessments. Certifications vary insofar as what is being credentialled—from entry-level of knowledge or proficiency skills to highly specialized and in-depth demonstrations of excellence.

Licensure is a regulatory process implemented via legislation or executive order to protect the public against potential harm or injury from unsafe or otherwise inappropriate professional practices. In addition to meeting specific eligibility requirements including academic credentials or degrees, amount-of-experience, or other training qualifications, licensure applicants typically have to demonstrate that they have attained a specified degree of competency to practice—often in an independent unsupervised manner. For example, in the United States, most health care professionals—physicians, nurses, physician assistants, and other allied health professionals—are licensed by state jurisdictional authorities such as state medical boards. State medical practices acts empower those state boards to enforce the law. In some cases, certification may be required as part of a licensure process. For example, teaching licenses in the US are typically regulated by an agency within the state department of education and a national certification test may be used to verify pedagogical and grade-appropriate content knowledge. Candidates applying for a law license must usually pass a comprehensive state bar examination and be admitted to the state legal bar in order to hold a license to practice law in the state or territory. Other types of licenses can cross local, state or territorial lines and either be regulated by a federal licensing authority or by reciprocal state licensing agreements. In some cases, a national licensing authority acts on behalf of its constituent state licensing boards for purposes of testing and establishing licensing eligibility. Examples include the National Association of State Boards of Accountancy and the Federation of State Medical Boards.

Recognition of credentialing as specialized area of assessment is relatively recent (Shimberg, 1981, 1982; Clauser et al., 2006; Raymond and Luecht, 2012; AERA, APA, and NCME, 2014; Buckendahl, 2017). In most ways, credentialing examinations are built on the same principles and standards of other educational, psychological and employment assessments. However, there are also some issues that remain unique and challenging for credentialing assessments. This chapter highlights seven issues relevant to designing and implementing an effective credentialing examination program: (1) carrying out a practice analysis to understand what to assess; (2) selecting or designing items and tasks to assess sometimes complex knowledge and skills; (3) designing and implementing an operational testing system for examination processing; (4) developing and maintaining a score scale; (5) engaging in standard setting activities to determine the credentialing cut scores; (6) designing and assembling efficient test forms to optimize credentialing decisions; and (7) evaluating the accuracy and validity of credentialing tests.

Practice analysis: determining what to assess

Credentialing examinations are similar in some respects to employment tests in that they should be constructed to reflect the tangible *knowledge, skills* (cognitive or physical), *abilities* or psychological traits, and professional *judgments* (critical decision-

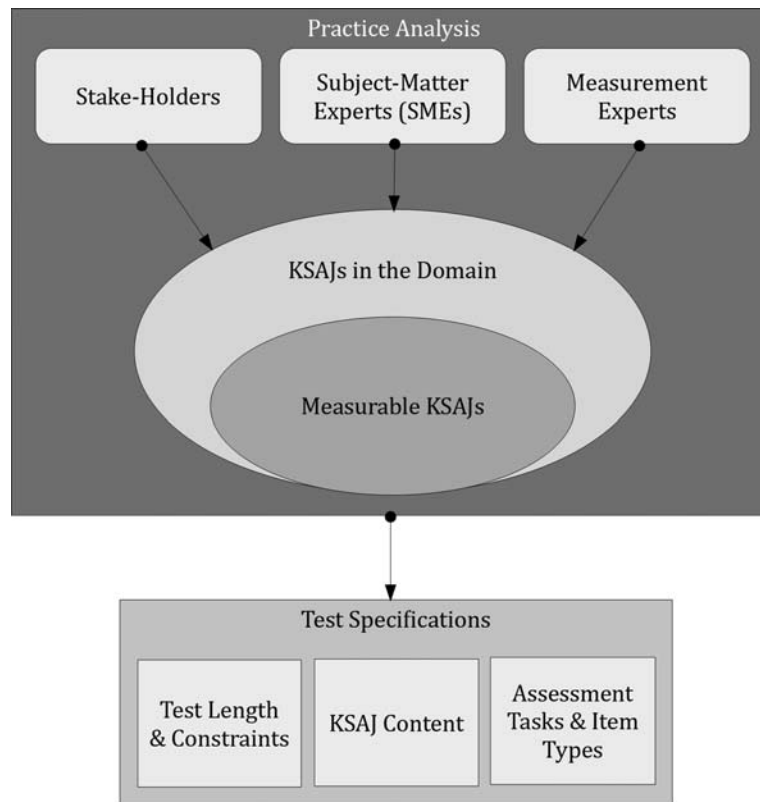


Fig. 1 Determining credentialing test specifications.

making) for entry into a profession, to be licensed, or to otherwise maintain or renew a credential (Note that the acronym “KSAJs” is used in this chapter to reflect the constellation of attributes that could be considered in designing a credentialing test). However, unlike many employment tests, credentialing examinations do not merely rank order the candidates and select as qualified those with the highest scores (Clauser et al., 2006). Credentialing examinations instead undergo a process called *practice analysis* to determine the minimum level of KSAJ-related attributes required to be credentialed or licensed, independent of how many eligible candidates are ultimately successful. The challenge resides with deciding which KSAJs to assess and how to best assess them.

There is no single best way of carrying out a practice analysis (Clauser and Raymond, 2017; Raymond, 2001; Raymond and Neustel, 2006). In fact, Clauser and Raymond outlined eight different methods for carrying out a practice analysis. In general, most practice analysis methods fall into one of two classes: *candidate-focused* or *task-focused* methods. Candidate-focused methods tend to describe the domain in terms of the KSAJs demonstrated by qualified individuals who hold the credential. For example, a candidate-focused method might delineate and prioritize the functional roles that qualified candidates should possess. Task-focused methods emphasize the types of job-related activities (tasks) in which qualified individuals apply the KSAJs. Most task-focused practice-analysis methods incorporate statements that are then rated in terms of the amount of time typical qualified candidates spend performing each task and the perceived importance of the task.

Practice analyses can also differ with respect to the groups included in the process. Professional organization leaders, qualified subject-matter experts (SMEs) and individuals with psychometric (measurement) expertise are all likely to have input into the process of defining and refining the KSAJs to be included on a credentialing examination (Davis-Becker and Muckle, 2017). The purpose and value (or stakes) of the credential, the diversity of the eligible candidate population, and the feasibility of actually measuring the potential wide array of KSAJs in a professional domain all influence the decision as to what is ultimately assessed (Raymond and Luecht, 2012). For example, a credentialing test targeted for entry-level candidates may be substantially different than a test design to credential highly experienced professionals, or those renewing a credential after some designated period of time.

Ultimately, the goal of any practice analysis is to identify the critical KSAJs in a profession (Kane, 1982, 1997). As depicted in Fig. 1, those target KSAJs must further be refined to represent the *measurable* characteristics of competent professional practice that can be converted into a set of test specifications that guide item writing and test assembly.¹ For example, the test content

¹The term “test blueprint” is commonly used to represent the content and (crudely) the cognitive specifications for assembling tests. However, test specifications can and should also incorporate statistical/psychometric requirements such as the distribution of item difficulty and the amount of precision allocated along a score scale.

specifications for the United States Medical Licensing Examination (USMLE) Step 1 covers ten basic science disciplines (e.g., pathology, physiology, pharmacology, biochemistry & nutrition) as well as being cross-referenced to multiple physician competencies involving general principles of medicine, organ systems, and allied areas related to patient care (FSMB and NBME, 2022).

Designing the test items: how to assess the KSAs

Following the practice analysis phase, a credentialing organization must design and generate the test items to meet the test specifications and to assemble test forms. The item design and development process usually involves five steps: (1) selecting or designing appropriate item types; (2) developing item writing guidelines; (3) engaging item writers and test editors to craft and refine the individual items²; (4) pilot-testing the items to obtain performance statistics used in evaluating the items; and (5) engaging in psychometric analysis to “bank” the items for use on current or future test forms.

There are four rather broad classes of items used on credentialing assessments: (i) knowledge-based items; (ii) simulated performance items; (iii) authentic work samples; and (iv) personality and other non-cognitive assessment items (Luecht, 2017a). Some credentialing assessments include a combination of items from these four classes.

Knowledge-based items are very popular for credentialing examination because they can be developed with minimal item-writing training and editing. They tend to take less time than other item types to administer relatively large numbers of items covering a broad range of topics. And they can furthermore be easily and inexpensively scored.

Most knowledge-based items employ some type of stimulus question or scenario and a selected-response (SR) mechanism. Examples of SR include the ubiquitous multiple-choice question (MCQ) that can be machine-scored using a scanner (for paper-and-pencil tests) or using response pattern-matching on a computer-based test (CBT)—see Haladyna (2004) and Luecht (2017a). Other SR item types typically considered for CBTs include multiple-select items, select-and-reorder (prioritize) items; drag-and-drop items, items incorporating digital media (e.g., images or video segments) or exhibits, hot-spot items and problem-based sets that include two or more items associated with a reading passage or scenario (Luecht, 2017a, 2017b; Parshall and Harnes, 2008; Parshall and Brunner, 2017). Short-answer item formats are sometimes used but can be difficult to score given the large number of possible matches of candidate responses to the correct answer(s).

As implied above, these item types are typically used to assess candidates’ declarative knowledge of the professional content, including prevailing theories, principles, professional standards, standardized protocols, and common scenarios or applications. Oral interviews, essays and other types of constructed-response items (CRs) are sometimes used to assess content knowledge, but typically require human scoring to evaluate the quality and substance of the written content. There is, however, a growing interest in using natural language processing (NLP), text analytics, and artificial intelligence for *automated scoring* (Shermis and Burnstein, 2003; Shermis et al., 2015; Drasgow et al., 2006; Williamson et al., 2006).

Simulated performance item types ideally provide a realistic problem scenario that requires the examinee to engage in problem-solving, analysis and/or critical decision making (Luecht and Clauser, 2001). These types of items often attempt to measure the application of specific procedural or decision-making skills using realistic manipulatives such as reference materials or software application tools (e.g. spreadsheets or analytical software). The simulated environment may present a type of “mission” or task challenge that each candidate must attempt using the tools and resources provided by the simulated environment. One example are the computerized case simulations included on the United States Medical Licensing Examination (USMLE™) developed by the National Board of Medical Examiners (Dillon, Clyman, Clauser, & Margolis, 2002; Clauser et al., 2002). The candidates are presented with a simulated patient exhibiting particular symptoms. They may need to order diagnostic tests or therapies and manage the patient over an extended period of time. Another example is the Uniform Certified Public Accountants Examination jointly developed and administered by the American Institute for Certified Public Accountants and the National Association of State Boards of Accountancy. That examination includes simulated performance assessments that require hands-on use of spreadsheets, preparation of written memos, queries of the authoritative accounting literature, and analysis of financial documents (Devore, 2002; Devore et al., 2011).

Other types of simulations may require examinees to construct job-related, written correspondence, or to use a set of software tools to analyze a data set or solve a computer networking problem. Although there is a great deal of intrinsic appeal in the “face validity” of simulated performance item types (see, for example, Clauser et al., 2002; Ward et al., 2019), there can be substantial costs involved in developing and piloting the items, potential security risks if the number of tasks in the item bank is relatively small, and possibly limited domain sampling and contribution to the overall measurement precision of the scores or classification decisions given the amount of time required for candidates to complete each simulation (e.g., Williamson et al., 2006).

Authentic work samples have a relatively long history in both employment and credentialing settings. These types of assessments require an examinee to actually perform a work-related activity to demonstrate an integrated set of knowledge and skills in an applied setting. An example would be requiring a candidate to successfully operate heavy construction equipment such as a crane, excavator or some other material handler to carry out various designated tasks in different settings (see nccco.org). The ubiquitous

²Recently, researchers have recommended less focus on crafting individual items, emphasizing the need to focus on families of *isomorphic* items that all measure the same constellation of KSAs and behave in a statistically similar manner (Luecht and Burke, 2020). Interest in mass-production of items using a combination of templates and computer algorithms—called automatic item generation (AIG)—is also growing given the high costs of producing large numbers test items for professional credentialing examinations (see Gierl and Haladyna, 2012).

on-the-road driving test used to obtain a driver's licensure would also qualify as an authentic performance-based work sample. Given the potential sources of noise that can come into play with work samples (e.g., uncontrollable variations in the assessment setting or inconsistent application of scoring protocols in different contexts by the examiners), it can be difficult to create measurement conditions and scoring protocols that are sufficiently standardized to ensure a comparable experience for all examinees (Ness et al., 2004).

Non-cognitive assessment such as personality questionnaires or situational judgment tasks (SJTs) are often used to measure personality affects such as temperament or attributes like self-reflection, honesty, professional ethics and accountability. Measuring leadership characteristics is also an area of growing interest in some professional settings. Although there is certainly interest by certification or licensing agencies in assessing non-cognitive affects to develop blended definitions of competency that include knowledge, skills, attitudes and professional judgment, few if any certification or licensing agencies have actually attempted to implement non-cognitive affective measures as part of the initial certification/licensure steps, or even for re-certification or license renewal. Instead, continuing education and documentation of designated professional development activities are used.

Designing an operational testing system for processing the examinations

Operational testing is far more complicated than merely writing test items, building test forms, administering the forms to candidates and reporting scores. There are many test development and psychometric support procedures involved in any testing enterprise (Dragow et al., 2006). Fig. 2 provides a high-level overview of the behind-the-scenes functional processes for a typical credentialing examination program. The statistical and content test specifications guide item writing. New items or tasks are ideally pretested³ on operational test forms to gather performance statistics and evaluate their quality and suitability for use on future test forms. Items that survive the quality control vetting process can be statistically calibrated to an underlying *item bank scale*—discussed below—and made eligible in the *item bank* (i.e., the item and task repository). Items in the bank can include calibrated pretest items and existing operational items. The test assembly process brings together the items from the bank (the supply) and the test specifications (the demands).

Scores on the operational test forms may undergo additional statistical adjustments to ensure that, ultimately, all scores and interpretation—including the final pass-fail credentialing decision—fairly reflect performance on a common metric.

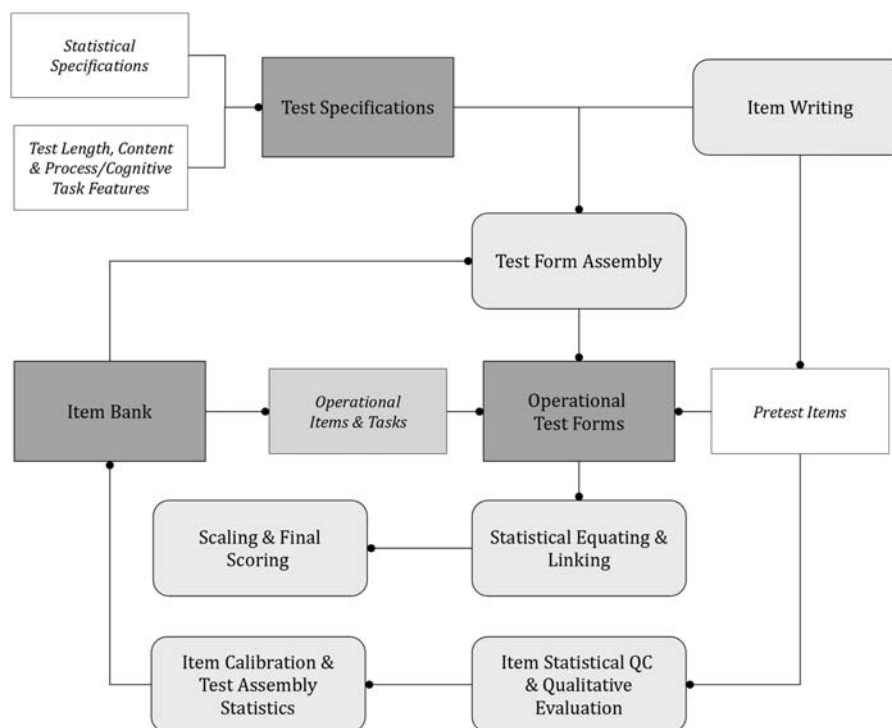


Fig. 2 Operational test development, quality control and item banking.

³Pretest items are typically not included in operational scoring. For this reason, the number of pretest items administered to any single candidate is relatively small (e.g., less than 20% of the total test length).

Scores and score scales: construction, maintenance, and interpretation

Most credentialing examinations are fundamentally *mastery tests*. That is, these types of examinations result in a discrete, pass-fail classification decision. There may indeed be an underlying score scale that is psychometrically maintained over time, but the score(s) are usually converted to a binary score scale: 0 = *fail* or 1 = *pass*. Although some may naively believe that an absolute cut score is a percent-correct score like 75%, the reality is that most credentialing examinations need to: (a) build the underlying statistical score scale; (b) engage in some reasonable standard-setting process to establish the cut score(s) for “mastery” on the test; (c) establish reasonable “data connectivity” and carry out appropriate scale linking analyses to maintain the same score scale and cut score(s) over time⁴; and (d) carry out various statistical and psychometric quality procedures to ensure that the scale properties and interpretations associated with the classification decision remain consistent across test forms.

The terms *scores* and *scoring* are often used in different ways. Item- or task-level scores are the dichotomous (binary) or polytomous (integer, partial-credit) points assigned to each item (e.g., 0 if wrong, 1 if correct). These scores are generated by scoring evaluators that include simple pattern matching algorithms, human scorers using a rubric, or automated scoring algorithms that employ artificial intelligence (AI) scoring models that attempt to replicate the scoring outcomes of human scorers.

Raw test scores are the weighted or unweighted number-correct or total-points earned on a test form obtained by summing the item scores. However, raw scores or percent-correct scores are not necessarily comparable across test forms that differ in average difficulty. For example, consider two forms of a 150-item credentialing examination, Form X and Y, each administered to equally proficient groups of examinees. Suppose Form X has an average score of 100 and Form Y has an average score of 105. The raw scores on these two forms are not comparable unless we make a statistical adjustment to one or both sets of scores to put them on a *common* scale.

Psychometricians use statistical *equating* or *linking* methods to establish comparable (exchangeable) scores from different test forms. A common score scale may be established early during the tenure of a credentialing examination program using one or more test forms. Scores on future test forms are statistically adjusted to make them comparable on that same common scale. For example, given our two 150-item test forms above, we might add five points to all Form X scores to make them statistically comparable⁵ to the Form Y scores. This would be called “*mean equating*”. Classical Test Theory (CTT) methods of establishing comparability between scores on different test forms include mean equating, linear equating, equipercentile equating and an array of non-equivalent groups equating methods (Kolen and Brennan, 2014).

Many credentialing-examination programs employ a different psychometric methodology for establishing a common metric across test forms called Item Response Theory (IRT) that offers many benefits that are simply not feasible with Classical Test Theory. There are numerous mathematical models that can be employed under the IRT umbrella to accommodate different types of items and response data. The technical details of implementing IRT are far beyond the scope of this chapter (see, for example, Hambleton, Swaminathan, & Rodgers, 1991; Lord, 1980). Suffice to say, IRT *calibrates* the statistical item difficulties (and, under some IRT models, other psychometric characteristics) of the individual items to a common metric that is referred to by the Greek letter “ θ ” (theta). The examinees’ scores are likewise represented on the same scale allowing psychometricians to view the location and density of items relative to the concentration of examinees and/or relative to key decision points like a credentialing cut score.

A key advantage of IRT is that all items in the bank are *calibrated* (statistically linked) to the common scale and can then be used to score examinees on that same θ scale regardless of which test form or collection of items they happen to take. IRT calibration further makes computer-adaptive testing (CAT) and multistage testing (MST) possible, where the difficulty of individual items, or larger modules comprised of multiple items or tasks, are algorithmically matched to the apparent proficiency of each examinee (Luecht and Sireci, 2011; Luecht, 2012, 2014).

Fig. 3 displays an IRT-calibrated item bank (at the left). Each circle or dot represents a single item with the shading denoting differential test content. The location of each item relative to the proficiency scale at the left signifies its difficulty. The test-assembly process (center oval) uses the test specifications (statistical target at the top and content requirements at the bottom) to build test forms. The calibrated item statistics associated with each test form are used to score the examinees on the common scale, θ . Differences in test form difficulty are accounted for by the IRT scoring process. As discussed below (see *Efficient Test Design and Form Assembly*), the statistical target at the top (middle) of Fig. 3 is called a *test information function* (TIF) target. Using an IRT-calibrated item bank and target TIF can help ensure that every test form constructed not only meets the sometimes-complicated content requirements but also provides optimal measurement precision to ensure highly accurate pass-fail classification decisions (van der Linden, 1998, 2005; Luecht, 2006, 2017a).

Reported scale scores are the point values or outcomes distributed to candidates and credentialing agencies. These scores may reflect one or more mathematical transformations of either the raw scores, equated raw scores, or IRT proficiency scores. For example, some credentialing programs use a simple transformation that specifies two points such as the mean score of a defined *reference group*⁶ and the cut score (e.g., 200 as the cut score and 250 as the reference group average scale on the reported scale). Those

⁴Smaller credentialing examination programs may elect to apply the established mastery standard to each test form. That amounts to “judgmental equating” and is generally less desirable than establishing the score scale and setting the cut once relative that established scale.

⁵In practice, we would have to employ random assignment of Form X and Y to examinees for this type of statistical equating to be feasible (i.e., we would be establishing *randomly equivalent groups* so that any apparent differences in the score distributions could be attributed to differences in test-form difficulty, see Kolen and Brennan, 2014).

⁶*Reference groups* are usually defined by characteristics such as only including first-time test takers undergoing specified training or a degree program from accredited professional schools or university programs. The theory is that the reference group remains fairly stable in its composition over time.

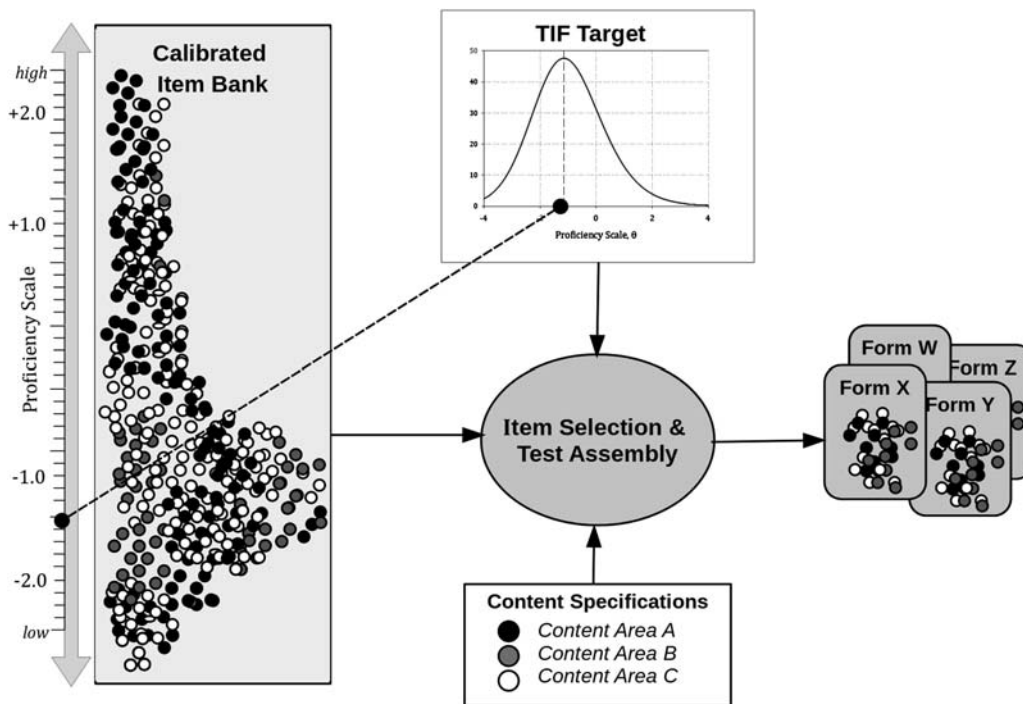


Fig. 3 Test assembly using an IRT-calibrated item bank.

two points can then be used to create a linear transformation equation to compute the reported scale scores. Finally, the mastery (pass-fail) classification scores reflect simple binary transformations of raw scores, equated raw scores, IRT proficiency scores, or reported scale scores. Values at or above an established cut score result in a “pass”, otherwise, the candidate receives a “fail” classification. The cut score is typically set as a specific value on an underlying IRT bank scale or CTT-equated raw-score scale and then transformed to the reported scale. It can be maintained over time by equating or linking future test forms to that same scale.

Determining the credentialing cut score

Passing a test can imply minimally acceptable knowledge and/or skills-based competency for an entry-level credentialing examination. It can alternatively signify a higher level of demonstrated KSAs (e.g. a credential for an advanced certification). The process of establishing a “cut score” is not arbitrary. Rather, a defensible *standard-setting* process is essentially mandated by the *Standards for Educational and Psychological Testing*:

Defining the minimum level of knowledge and skill required for licensure or certification is one of the most important and difficult tasks facing those responsible for credentialing. The validity of the interpretation of the test scores depends on whether the standard for passing makes an appropriate distinction between adequate and inadequate performance. Often, panels of experts are used to specify the level of performance that should be required. Standards must be high enough to ensure that the public, employers, and government agencies are well served, but not so high as to be unreasonably limiting. Verifying the appropriateness of the cut score or scores on a test used for licensure or certification is a critical element of the validation process. (AERA/APA/NCME, 2014, p. 176)

There are multiple methods for setting standards but virtually all of them start with a process for selecting the panelists from an eligible pool of policy makers, subject-matter experts, or other stake-holders. The panelists typically undergo training in a particular standard-setting method and then engage in one or more rounds of rating or classifying candidate performances. The resulting cut score is usually a weighted or unweighted average of the panelists’ individual values. A credentialing policy board or other authorized group may approve the final cut score.

Three common standard-setting methods are the modified Angoff method and associated variations (Angoff, 1971), the Bookmark method (Lewis et al., 1996), and variations on the Contrasting Groups method (Berk, 1976). Cizek and Bunch (2007) provide a very thorough discussion of the array of standard setting methods, their similarities and differences, and issues related to implementation and use.

The modified Angoff method employs probability ratings of individual test items on one or more reference test forms for a theoretical group of *just-qualified candidates* (e.g., 100 credentialing candidates who are just at the threshold in terms of their performance on the reference form[s]). The cut score is a function of the summed probabilities on the reference form(s) with the results usually averaged across standard-setting panels.

The Bookmark procedure orders the items on a prescribed test form from easiest to hardest and instructs the panelists to place the conceptual “bookmark” between items that separate qualified from unqualified candidates answering the items at a designated probability of correctly answering the items. When used in an IRT context, the bookmark for each panelist corresponds to their recommended cut score.

Finally, Contrasting Groups and related holistic methods separate individual candidates who took the test or their performance profiles into two groups, qualified and unqualified candidates, based on the panelists’ judgments about what it takes to deserve the credential. The cut score is determined as the point that optimally separates the two distributions of scores or profiles on the test. Individual panelists’ cut scores are normally averaged or otherwise aggregated to produce a single, recommended cut score. Most standard-setting events further employ two or more rounds of ratings to allow the panelists to feel comfortable with and fine tune their classification decisions (e.g., after reviewing impact data based on the selected cut score).

Once the final credentialing cut score is established and approved for use by the appropriate authority (e.g., a governing board), most credentialing organizations maintain the cut score by equating or linking new test forms to the underlying score scale. When IRT is employed, the cut score is on the item-bank scale, θ . This makes it relatively straightforward to design test forms to optimize current as well as future pass-fail classification decisions. The amount of test information targeted in the region of the cut score determines the statistical accuracy of classification decisions (see the target TIF in Fig. 3). The interplay between decision accuracy associated with cut scores and measurement precision is also discussed in more detail below.

Efficient test design and form assembly

Test-form assembly is a process of building unique or semi-unique test forms that meet designated test specifications—both statistical targets and content-related targets. For example, the test specifications may include a detailed list of relevant topics or content-related categories gleaned from the practice analysis, as well as process-oriented categories or ordered cognitive-complexity dimensions like Bloom’s taxonomy (Bloom et al., 1956). The test specifications can also include professional settings or scenarios to be covered. In contrast, the statistical test specifications reflect either the targeted average difficulty of the test and/or the desired level of precision of the scores.

An important function for IRT-based test assembly is called the *test information function* or TIF. A TIF mathematically expresses the precision of the IRT proficiency scores across the underlying metric, θ . The TIF for a given test form is comprised of item information functions (IIFs). That implies that we can select items to specifically meet a target TIF, thus ensuring that the test form provides exactly the amount of information where it is most needed—in the region of the cut score for a credentialing test (Luecht, 2006; 2017a).

Fig. 4 provides two plots that might be associated with a credentialing exam to certify minimum competence in a profession. The proficiency scale is shown on the horizontal axis. The solid-line plot represents the target TIF with the left-hand vertical scale representing the amount of information (higher is better and implies more precision). The target TIF peaks exactly at the cut score of $\theta_{cut} = -1.28$, which corresponds to the 10th percentile of a normally distributed population of candidates with average proficiency of $\mu(\theta) = 0$ and a variance of $\sigma^2(\theta) = 1$. This distribution of candidate proficiency represents a reference sample of candidates such

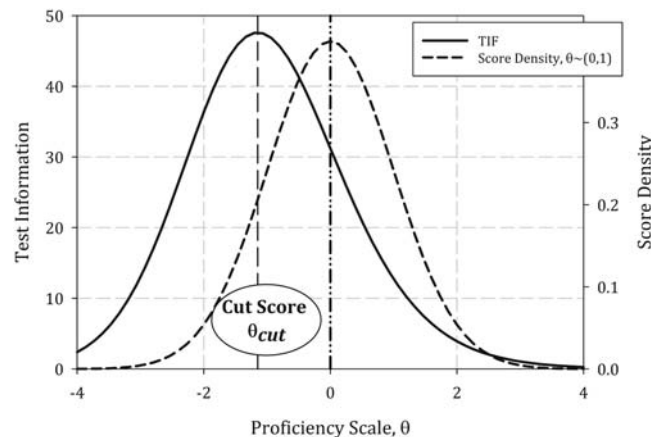


Fig. 4 A target TIF for a credentialing examination with a cut score at $\theta_{cut} = -1.28$.

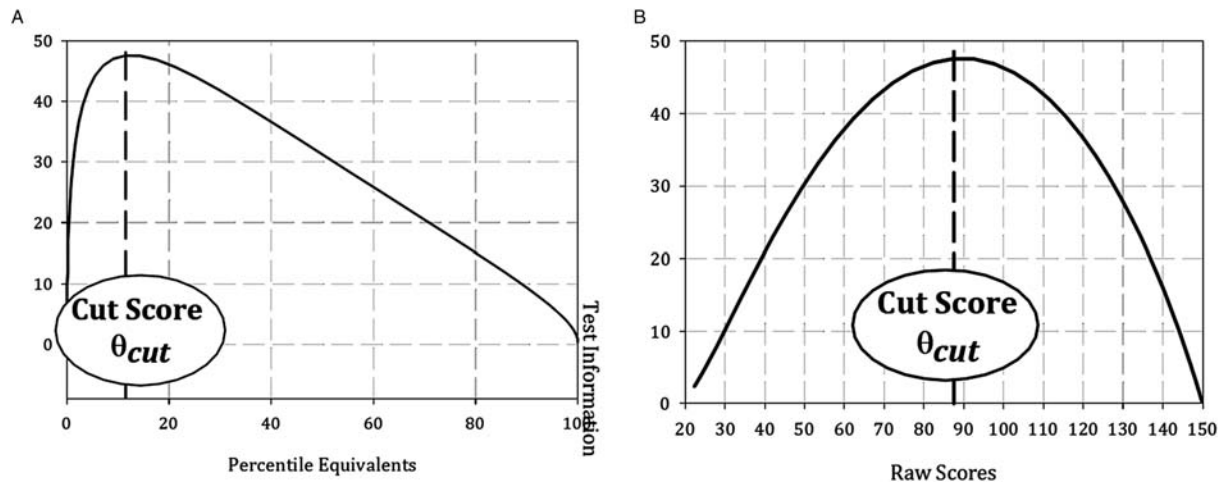


Fig. 5 Alternative ways to display TIF functions.

as first-time test takers from a designated time frame and is indicated by the dashed line in Fig. 4—corresponding to the right-hand vertical scale. This credentialing examination would therefore pass about 90% of the eligible candidates.

The test assembly process can be performed manually by test editors or a test committee. More commonly, testing agencies are beginning to employ sophisticated mathematical optimization models that can very quickly build numerous test forms that meet all of the content and cognitive/process constraints and hit the target TIF. This latter capability is known as *automated test assembly* or ATA (van der Linden, 1998, 2005; Luecht, 1998, 2017a).

Test information can also be plotted relative to other performance-related metrics. For example, Fig. 5A shows the measurement precision as a function of normal-equivalent percentiles. Here, it is clear that the information peaks near a cut score that normatively implies a 10% failure rate. That is, this type of plot provides a normative interpretation of the test information by showing how the measurement precision of a test form is allocated relative to the target distribution of the candidates. It does not represent an absolute percentage of failing candidates. The right-hand plot shown in Fig. 5B presents the amount of test information relative to the [expected] number-correct scores on a reference test form. The horizontal axis can also be easily converted to a proportion- or percent correct-score metric.

Credentialing tests assembled without carefully designed statistical targets may exhibit score or classification inconsistencies or comparability issues across test forms. The rationale for using an unambiguous target statistical specification is three-fold. First, TIFs or other comparable psychometric targets can reflect optimal and feasible test-design specification for assembling every test form. Second, the specifications can help guide item writing/production over time to ensure that the item bank (supply) can meet those optimal measurement information demands. Third, when combined with automated test assembly, TIFs can ensure a high degree of consistency (statistical parallelism) for every test form.

Credentialing classification accuracy and validation

Test scores are rarely the sole determinant for credentialing decisions. For example, candidates may also have to meet stringent eligibility requirements that include verified training or completion of an academic degree. That said, a large part of the justification for using test scores in credentialing stems from the practice analysis and choice of test content (Clauser et al., 2006; Raymond, 2006; Clauser and Raymond, 2017). That is, as noted earlier, the test content is designed to specifically measure the knowledge, skills, abilities and judgments (KSAs) demanded of qualified [credentialed] candidates in situations typically encountered in practice.

Standard setting, in turn, provides a type of empirical validation as to the amount KSAs required for each credential. That empirical validation process justifies the choice of cut score adopted by each credentialing authority. In addition, using appropriate statistical targets for test assembly—like a target test information function (TIF) that consistently amasses most of the measurement precision for each test form in the region of that cut score—ensuring maximal classification (decision) accuracy. For example, a target TIF like the one shown in Fig. 4 provides the desired amount of precision associated with classifying candidates as “passers” or

“failers”. More information in the region of the cut score implies greater precision and fewer false positive or false negative decision errors.

Although external or predictive validation might be considered—that is, statistically correlating the credentialing examination test scores with external measures of “qualified” or “successful” candidates to empirically verify the “accuracy” of decisions,⁷ it is rarely done in practice for an important reason. Simply put, it is extremely difficult to obtain a universally acceptable criterion measure that reflects the same KSAs as the credentialing test. Rather, most credentialing examinations depend on the integrity and generalizability of the entire test design, standard setting, test-form assembly practices, and underlying psychometric processes as providing sufficient and compelling evidence for the interpretation and use of the test scores and classification decisions (Kane et al., 2017).

Conclusions

Credentialing examinations—both certification and licensure tests—present some unique challenges that distinguish them from other assessments like educational achievement tests. A primary difference is the scope and nature of the domain(s) covered. Most credentialing examinations reflect the measurable professional knowledge, skills, abilities and judgments (KSAs) required of qualified individuals in the candidate population across a wide array of typical jobs or credentialing situations. The credential normally reflects minimum competency relative to those KSAs. A practice analysis therefore serves the essential purpose of providing empirical validity evidence about the KSAs that should be included as part of the test specifications.

The choice(s) of item or task types to be included on a credentialing test (multiple-choice or other selected-response items, essays or constructed-response items, simulations, etc.) are sometimes made based on cost and other pragmatic factors. However, it seems advisable to start with the practice analysis and KSAs to determine the nature and scope of the performance evidence required to support the proficiency claims inherent in acquiring the credential. Only then should credentialing organizations select item types that will efficiently and accurately provide that evidence.

Implementing a credentialing examination program can require a fairly complex system of test development, item writing and editing, pilot testing of the test content, and psychometric analyses. The purpose of that system is to process scores and classification decisions for candidates currently taking the test. However, the system should also address future item production given test development demands, item banking, and maintenance of the score scale and cut score over time. Item response theory (IRT) was noted as a solid choice to support these system-level requirements.

Standard setting is an integral part of any credentialing examination program. Simplistically, it establishes the cut score for pass-fail classification decisions. More comprehensively, standard setting provides a formal way for mapping reasoned interpretation of the level of the KSAs expected of qualified candidates to the score scale. Once the cut score is determined—by any of various standard-setting methods—test design and assembly procedures should attempt to optimize the accuracy of pass-fail decisions (i.e., to minimize classification errors). A method of using target IRT test information functions was suggested along with automated test assembly, especially if substantial numbers of test forms are required for operational use.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014. *Standards for Educational and Psychological Testing*. American Educational Research Association, Washington, DC.
- Angoff, W.H., 1971. Scales, norms, and equivalent scores. In: Thorndike, R.L. (Ed.), *Educational Measurement*. American Council on Education, Washington, DC, pp. 508–600.
- Berk, R.A., 1976. Determination of optimal cutting scores in criterion-referenced measurement. *J. Exp. Educ.* 45, 4–9.
- Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R., 1956. *Taxonomy of Educational Objectives; the Classification of Educational Goals; Handbook I: Cognitive Domain*. Longmans, Green, New York, NY.
- Buckendahl, C.W., 2017. Credentialing: a continuum of measurement theories, policies and practices. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, pp. 1–20.
- Clauser, B.E., Margolis, M.J., Case, S.M., 2006. Testing for licensure and certification in the professions. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education (ACE)/Praeger, pp. 701–731.
- Clauser, A.L., Raymond, M., 2017. Specifying the content of credentialing examinations. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, pp. 64–84.
- Clauser, B.E., Kane, M.T., Swanson, D.B., 2002. Validity issues for performance-based tests scored with computer-automated scoring systems. *Appl. Meas. Educ.* 15, 413–432.
- Davis-Becker, S., Muckle, T.J., 2017. Test design: laying out the road map. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, pp. 41–63.
- Devore, R., 2002. *Considerations in the Development of Accounting Simulations* [Technical Report]. American Institute of Certified Public Accountants, Ewing, NJ.
- Devore, Brittingham, P., Maslott, J., 2011, March. Taking MCQ items to the next level—using simulation technology to enhance measurement and face validity. In: Paper Presented at the Association of Test Publishers Annual Conference.
- Drasgow, F., Luecht, R.M., Bennett, R.E., 2006. Technology and testing. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. American Council on Education (ACE)/Praeger, pp. 471–515.

⁷Employment tests are sometimes validated using an external indicator of success such as supervisor ratings. High correlations between the employment test scores and the external indicator are taken as evidence of validity. Likewise, educational assessments are sometimes validated by demonstrating an appropriate level of correlation with other, similar assessments.

- Federation of State Medical Boards and National Board of Medical Examiners, 2022. USMLE Content Outline. Retrieved from: https://www.usmle.org/sites/default/files/2022-01/USMLE_Content_Outline_0.pdf. (Accessed 27 January 2022).
- Haladyna, T.M., 2004. Developing and Validating Multiple-Choice Test Items, third ed. Erlbaum, Mahway, NJ.
- Kane, M.T., 1982. The validity of licensure examinations. *Am. Psychol.* 37, 911–918.
- Kane, M.T., 1997. Model-based practice analysis and test specifications. *Appl. Meas. Educ.* 10, 5–18.
- Kane, M.T., Clauser, B.E., Kane, J., 2017. A validation framework for credentialing tests. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, pp. 21–40.
- Kolen, M.J., Brennan, R.L., 2014. *Test Equating. Methods and Practices*, third ed. Springer-Verlag New York, Inc, New York, NY.
- Lewis, D.M., Mitzel, H.C., Grene, D.R., 1996, June. Standard setting: a bookmark approach. In: Green (Chair), D.R. (Ed.), *IRT - Based Standard Setting Procedures Using Behavioral Anchoring*. Symposium Conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.
- Luecht, R.M., 1998. Computer-assisted test assembly using optimization heuristics. *Appl. Psychol. Meas.* 22, 224–336.
- Luecht, R.M., 2006. Designing tests for pass-fail decisions using item response theory. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 575–596.
- Luecht, R.M., 2012. Computer-based and computer-adaptive testing. In: Ercikan, K., Simon, M., Rousseau, M. (Eds.), *Improving Large Scale Assessment in Education: Theory, Issues, and Practice*. Taylor-Francis/Routledge, New York, pp. 91–114.
- Luecht, R.M., 2014. Computerized adaptive multistage design considerations and operational issues. In: Yan, D., von Davier, A.A., Lewis, C. (Eds.), *Computerized Multistage Testing: Theory and Applications*. CRC Press/Taylor & Francis Group, London, UK, pp. 69–83.
- Luecht, R.M., 2017a. Data and scale analysis for credentialing examinations. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions*. Routledge, New York, pp. 123–152.
- Luecht, R.M., 2017b. Calibrating technology-enhanced items. In: van der Linden, W.J. (Ed.), *Handbook of Item Response Theory, Volume Three: Applications*. CRC Press, New York, pp. 86–193.
- Luecht, R.M., Sireci, S., 2011. A Review of Models for Computer-Based Testing [Research Report 2011-12]. The College Board, New York.
- Luecht, R., Burke, M., 2020. Reconceptualizing items: from clones and automatic item generation to task model families. In: Lissitz, R., Burke, M. (Eds.), *Applications of Artificial Intelligence to Assessment*. Information Age Publishers, Baltimore, MD.
- Ness, W.J., Tepe, V., Ritzer, D., 2004. *The Science and Simulation of Human Performance*. Elsevier, New York.
- Parshall, C.G., Brunner, B., 2017. Content development and review. In: Davis-Becker, S., Buckendahl, C.W. (Eds.), *Testing in the Professions: Credentialing Policies and Practices*. Routledge, New York, pp. 85–104.
- Parshall, C.G., Harmes, J.C., 2008. The design of innovative items: targeting constructs, selecting innovations, and refining prototypes. *CLEAR Exam Rev.* 19 (2), 18–25.
- Raymond, M.R., 2001. Job analysis and the specification of content for licensure and certification examinations. *Appl. Meas. Educ.* 14, 369–415.
- Raymond, M.R., Luecht, R.M., 2012. Licensure and certification testing. In: Geisinger, K.F. (Ed.), *APA Handbook of Testing and Assessment in Psychology*, vol. 3. APA Publications, Washington, DC.
- Raymond, M.R., Neustel, S., 2006. Determining the content of credentialing examinations. In: Downing, S.M., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 181–223.
- Shermis, M.D., Burstein, J., Brew, C., Higgins, D., Zechner, K., 2015. Recent innovations in machine scoring. In: Lane, S., Haladyna, T., Raymond, M. (Eds.), *Handbook of Test Development*, second ed. Routledge, New York, pp. 335–354.
- Shimberg, B., 1981. Testing for licensure and certification. *Am. Psychol.* 36, 1138–1146.
- Shimberg, B., 1982. *Occupational Licensing: A Public Perspective*. Educational Testing Service, Princeton, NJ.
- U.S. Bureau of Labor Statistics, 2021. www.bls.gov/ors/factsheet/credentials.htm. (Accessed 31 October 2021).
- van der Linden, W.J., 1998. Optimal assembly of psychological and educational tests. *Appl. Psychol. Meas.* 22, 195–211.
- van der Linden, W.J., 2005. *Linear Models for Optimal Test Design*. Springer, New York.
- Ward, R.C., Muckle, T.J., Kremer, M.J., Krogh, M.A., 2019. Computer-based case simulations for assessment in health care: a literature review of validity evidence. *Eval. Health Prof.* 42, 82–102.
- Williamson, D.M., Mislevy, R.J., Bejar, I.I., 2006. *Automated Scoring for Complex Tasks in Computer Based Testing*. Lawrence Erlbaum, Hillsdale, NJ.

Admissions testing in higher education

Rebecca Zwick, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of R. Zwick, Admissions Testing*, In International Encyclopedia of Education (Third Edition), edited by Penelope Peterson, Eva Baker and Barry McGaw, Elsevier Ltd., <https://doi.org/10.1016/B978-0-08-044894-7.00230-X>.

Background	451
Higher education admissions testing in the US	452
Main admissions tests used in the US today	452
Today's SAT	452
The ACT	452
The GRE	452
The MCAT	453
The LSAT	453
The GMAT	453
The use of higher education admissions tests in the US	454
Use of admissions tests at four-year undergraduate institutions	454
Use of admissions tests at graduate and professional schools	454
Predictive validity of admissions tests	454
Use of personal characteristics to predict success in higher education	456
Fairness of admissions tests to people of color, women, and other special populations	456
A sampling of higher education admissions testing around the world	456
References	457
Testing program websites	459

Glossary

Correlation coefficient (multiple) A measure of association, ranging from 0 to 1, between a set of (at least two) variables and a single variable. For example, a multiple correlation coefficient could be used to assess the degree to which test scores and high school grades, considered jointly, are associated with college grade-point average

Correlation coefficient (simple) A measure of linear association, ranging from -1 to 1, between two variables (e.g., an admissions test score and a college grade-point average)

Differential item functioning (DIF) A phenomenon in which two (or more) groups of test takers (e.g., men and women) tend to perform differently on a test question despite having equal skill in the area that is being tested. One possible cause of DIF is the inclusion in the test question of material that is irrelevant to the proficiency that is intended to be assessed

Linear regression analysis A statistical technique used to assess the relationship between a set of predictors (consisting of one or more variables) and a criterion variable. In the admissions context, for example, test scores and high school grades are often used to predict college grade-point average

Predictive validity The degree to which a test or other measuring instrument can predict a criterion of interest, typically academic or job performance. In the case of admissions tests, the criterion is typically a grade-point average in a subsequent academic program

Range restriction A reduction in the range (distance between maximum and minimum values) of a variable that results from a selection process. For example, the SAT scores of students admitted to college typically have a smaller range than the SAT scores of applicants

Background

This article provides a brief historical perspective and then describes several aspects of higher education admissions testing in the US: the main tests used, the role of tests in the admissions process, and the evidence on predictive validity. (URLs for each of the tests are given at the end of the article). Following this is a short discussion of test fairness issues. Admissions testing in several other countries is then described.

The administration of tests for regulating admissions to universities most likely began in Europe during the eighteenth century (Webber, 1989: 37), although some reports give an earlier date (e.g., Office of Technology Assessment, 1992; Stewart, 1998). Most

accounts agree that admissions testing had been introduced in Germany and England by the mid-1800s. Standardized admissions testing took root in the US in 1900, when the leaders of 12 northeastern universities formed the College Entrance Examination Board. The College Board developed a common set of examinations that were administered by member institutions and then shipped back to the Board for scoring. Initially, the Board created essay tests in several subject areas; it later developed a new exam that contained mostly multiple-choice questions—the SAT. The first SAT contained questions similar to those included in the Army Alpha tests, which had been used for selecting and assigning military recruits during World War I. These Army tests, in turn, were directly descended from intelligence tests, which had made their US debut in the early 1900s. The SAT was first administered in 1926 to about 8000 candidates (Zwick, 2017b).

Higher education admissions testing in the US

Main admissions tests used in the US today

Admissions processes in the US are largely under the control of individual institutions and instructional programs. In most cases, these entities decide whether and how to use tests in making admissions decisions. The main admissions tests used in the US today are the SAT and ACT, used for college admission; the Graduate Record Examinations (GRE), used for graduate school admission; the Medical College Admission Test (MCAT); the Law School Admission Test (LSAT); and the Graduate Management Admission Test (GMAT), used for admission to business school. As described below, these tests are developed and produced by private organizations. Tests that are not discussed in this article include tests of English that are required of foreign students by some US institutions and admissions tests for certain professional schools, such as schools of dentistry and pharmacy.

A world event that had a substantial impact on test administration policies and practices for the tests described in this article was the outbreak of the coronavirus pandemic in 2020. The administration of tests in traditional testing centers was limited. Applicants could take the GRE, MCAT, LSAT, and GMAT, at home, with remote proctoring. In some cases, modified versions of these tests were used. The future of these alternative procedures, which are described on the websites of the testing programs, is not known as of this writing.

Today's SAT

The SAT testing program is owned by the College Board; the tests, which are primarily paper-based, are administered by Educational Testing Service (ETS) under a contract with College Board. (A computer-delivered version is used in some states and school districts). According to the College Board website, the SAT is “[d]eveloped with input from high school teachers, college faculty, and enrollment professionals [and] covers core content areas deemed essential for success in college—reading, writing and language, and mathematics.” (“SAT” is no longer considered to be an acronym, but the actual name of the test. Originally, it stood for Scholastic Aptitude Test, which was later changed to Scholastic Assessment Test. More recently, the test was known as the SAT I: Reasoning Test).

The SAT has undergone many changes in content in its nearly 100 years of existence; a history of some of these changes is provided by Lawrence et al. (2004). The most recent version of the test, which was introduced in 2016, includes a section on Reading, Writing, and Language, which consists entirely of multiple-choice questions, and a Mathematics section, which contains both multiple-choice items and questions to which the test-takers must supply an answer. The main scores are two section scores—one in “Evidence-Based Reading and Writing” and one in Math—and a total score. (An optional essay test was discontinued in 2021 except for its use in high school accountability programs. The SAT Subject Tests, which assessed the candidates’ knowledge in particular areas, have also been discontinued).

The ACT

In 1959, a competitor to the SAT program emerged—the American College Testing Program (Today, the company is ACT, Inc. and the test is simply the ACT. Like “SAT”, “ACT” is no longer considered an acronym). In its early years, the ACT was administered primarily in Midwestern states, but it is now considered interchangeable with the SAT by most institutions. Like the SAT, the ACT is primarily paper-based but is delivered via computer in some states and school districts.

From the beginning, the ACT was somewhat different from the SAT in terms of underlying philosophy: While the SAT consisted only of verbal and mathematical sections, the ACT was more closely tied to instructional objectives. Today, the content of the test is informed by regular surveys of secondary school teachers and curriculum experts. The ACT consists of four multiple-choice sections: English, Mathematics, Reading, and Science. Test takers receive a score in each subject area, as well as a composite score and a “STEM score,” which is based on performance on the Math and Science sections. Those who choose to take an optional essay test in Writing receive a Writing score and an English Language Arts score based on the English, Reading, and Writing results.

The GRE

The Graduate Record Examinations program of the Carnegie Foundation for the Advancement of Teaching administered its first exam in 1937. In 1948, the GRE program was transferred to ETS, which had been founded the previous year. Today, the GRE is developed and administered by ETS under the direction of the GRE Board, an independent organization that is affiliated with the Association of Graduate Schools and the Council of Graduate Schools. Although it is primarily used to evaluate candidates for admission to graduate school, the GRE General Test is accepted by some business schools and law schools as well. According

to the GRE Website, the test “measures skills that graduate and professional programs have identified as prerequisite for success at the graduate level. One such essential skill is critical thinking, and the GRE General Test measures aspects of critical thinking in each of the test’s three sections,” which are Verbal Reasoning, Quantitative Reasoning, and Analytical Writing. The General Test is computer-administered and is available in test centers around the world.

Since 2011, the Verbal Reasoning and Quantitative Reasoning sections of the GRE General Test have been administered as multi-stage tests (MSTs). In each of these areas, the first of two subsections is of medium difficulty, for purposes of routing. The test taker’s score on the routing subsection determines whether the second section will be easy, medium, or difficult (Robin et al., 2014). (A paper-delivered non-MST version of the GRE is offered in regions of the world where computer-based testing is unavailable).

At the time of the transition to MST format, the content of the Verbal, Quantitative, and Analytical Writing measures was revised, along with the scale on which Verbal and Quantitative scores are reported. The Verbal Reasoning section includes multiple-choice items and items that involve identifying a sentence in a passage that meets a certain description. The Quantitative Reasoning section includes multiple-choice items and those that require numerical entry, as well as quantitative comparison items, in which test takers must compare two quantities and determine whether one quantity is greater, the two are equal, or there is insufficient information to decide. Some multiple-choice items in both the Verbal and Quantitative sections require the test taker to select all correct answers from a set of possible responses, as more than one response may be correct. The Analytical Writing section consists of two essay tasks, which are intended to assess candidates’ ability to “to articulate and support complex ideas with relevant reasons and examples, and examine claims and accompanying evidence,” according to the GRE Website. Test takers receive separate scores in Verbal Reasoning, Quantitative Reasoning, and Analytical Writing.

In addition to the General Test, GRE Subject Tests in Biology, Chemistry, Literature in English, Mathematics, Physics, and Psychology are available, delivered in paper form.

The MCAT

The Medical College Admission Test, first given in 1946, is sponsored by the Association of American Medical Colleges (AAMC), which currently represents 172 accredited US and Canadian medical schools, more than 400 teaching hospitals, and various academic societies and medical professionals. (An earlier test, called the Scholastic Aptitude Test for Medical Students or the Moss Test, was administered between 1928 and 1946.)

The current version of the MCAT, launched in 2015, is a computer-administered, multiple-choice examination that, according to the AAMC Website, is “designed to assess ... problem solving, critical thinking, and knowledge of natural, behavioral, and social science concepts and principles prerequisite to the study of medicine.” The test consists of four sections: Biological and Biochemical Foundations of Living Systems, Chemical and Physical Foundations of Biological Systems, Psychological, Social, and Biological Foundations of Behavior, and Critical Analysis and Reasoning Skills. Test takers receive a score on each section, along with a total score.

The LSAT

The LSAT was conceived at a 1947 meeting between College Board staff and representatives of an association of nine law schools, the precursor of the Law School Admission Council (LSAC). In 1948, the first LSAT was administered by ETS. Since 1979, the LSAT has been administered by the LSAC itself, which today has 222 member institutions in the US, Canada, and Australia.

According to the LSAC Website, the test is “designed to assess critical reading, analytical reasoning, logical reasoning, and persuasive writing skills—key skills needed for success in law school.” The LSAT is administered in two parts. The first part of the test is a multiple-choice exam administered at test centers around the world. In North America, this portion is administered by computer. This part of the exam includes one Reading Comprehension section, one Analytical Reasoning section, and two Logical Reasoning sections. Performance is summarized in a single score.

The second part of the test, called LSAT Writing, is an unscored writing sample that is submitted online using secure proctoring software that can be installed on the candidate’s own computer. Test takers are given a decision problem and are asked to write an essay in which they choose between two positions and defend their choice. Copies of the writing sample are sent to all schools to which candidates apply.

The GMAT

In the late 1940s and early 1950s, the GRE was used to screen business school applicants. In 1953, however, representatives of nine graduate schools of business agreed that they needed an admissions test of their own and, just a year later, the precursor to the GMAT was administered (Schmoller, 1993). For five decades, the test was developed and administered by ETS for the Graduate Management Admission Council (GMAC). Since 2006, the GMAC has contracted with ACT, Inc. to develop the GMAT and with Pearson VUE, a company specializing in computer-based testing, to administer it.

The GMAT is a computer-administered test of “skills required to be successful in graduate business school,” according to the test handbook (Graduate Management Admission Council, 2020). It contains three multiple-choice sections—Verbal Reasoning, Quantitative Reasoning, and Integrated Reasoning, as well as an essay test, the Analytical Writing Assessment.

The Verbal and Quantitative Reasoning sections are adaptive, meaning that they are tailored to the test taker’s skill level. Each of these sections begins with a medium-difficulty item; the difficulty of subsequent items depends on the responses to previous items. The Integrated Reasoning section is the newest part of the test. Added in 2012, it measures the “ability to analyze data and evaluate information presented in multiple formats,” according to the GMAT Website. Many of the Integrated Reasoning questions require

more than one response, and an online calculator with basic functions is available for test takers to use. The Analytical Writing Assessment is intended to measure the ability to think critically and communicate ideas. GMAT test takers receive a score for each of the four sections of the test as well as a total score based on the Verbal and Quantitative sections.

The use of higher education admissions tests in the US

Use of admissions tests at four-year undergraduate institutions

In the US, use of admissions tests varies widely across institutions. In academic year 2018–2019, there were approximately 2300 four-year institutions offering degrees at the bachelor's level or higher. (Only schools that enrolled first-year undergraduates and participated in Title IV federal financial aid programs were included in this count). About 25% of these four-year colleges and universities were open-door, meaning that they accepted all applicants, 30% accepted at least three-quarters of their applicants, and 30% accepted between half and three-quarters. Only 14% had acceptance rates lower than 50% ([National Center for Education Statistics, 2020](#)). The US also has about 1300 two-year institutions, most of which are open-door and therefore do not use admissions tests.

According to the National Association for College Admission Counseling (NACAC), which conducts annual surveys of postsecondary institutions, “[t]he factors that admission officers use to evaluate applications from first-time freshmen have remained largely consistent over the past 20 years. Students’ academic achievements—which include grades, strength of curriculum, and admission test scores—constitute the most important factors in the admission decision” ([Clinedinst, 2019](#), p. 4). However, it is worth noting that the percentage of responding institutions attributing “considerable importance” to test scores declined from 52% to 46% between 2017 and 2018. (Because of the nature of the sample, some caution is warranted in interpreting the NACAC results. In 2019, NACAC surveyed its member institutions, which represented 78 % of four-year institutions in the US that met certain criteria. The response rate was 35% ([Clinedinst, 2019](#), p. 23)).

A recent trend that is not discussed in the NACAC report is the increase in test-optional admissions policies in the US. Under these policies, which began in a few small liberal arts colleges in the 1960s, college applicants themselves decide whether to submit admissions test scores. Submitters’ scores are considered in the admissions process, but nonsubmitters are not at a disadvantage. The National Center for Fair & Open Testing, better known as FairTest, keeps track of schools that do not require admissions test scores. As of late 2020, the organization’s website includes a list titled, “1600+ Accredited, 4-Year Colleges & Universities with ACT/SAT-Optional Testing Policies for Fall, 2021 Admissions” (see <https://www.fairtest.org/university/optional>).

The test-optional movement gained steam in the early 21st century. In most cases, institutions adopting test-optional policies did so in the hope that a more diverse student body would result ([Lopez et al., 2020](#)). The movement grew stronger in the wake of the coronavirus pandemic that struck in 2020. Because most test centers closed, many institutions waived test score requirements—in some cases, for multiple admissions cycles ([Jaschik, 2020](#)).

Use of admissions tests at graduate and professional schools

Graduate school programs tend to be smaller, and admissions policies much more flexible than at undergraduate institutions. Admissions policies and rates vary widely over the hundreds of fields of graduate study available in the US. Decisions are typically in the hands of program faculty, rather than admissions officers and admission rates tend to be much lower.

Surveys have shown that standardized admissions tests are widely used in graduate admissions, in combination with undergraduate grades and other factors ([Kuncel et al., 2001](#)). According to the GRE website, the GRE General Test “is accepted at thousands of graduate schools, including business and law.” In recent years, however, there has been a growing concern that an overemphasis on GRE scores could constitute a barrier to campus diversity (e.g., see [Posselt, 2016](#)). Some graduate departments, therefore, have elected to eliminate the requirement to submit admissions test scores and have moved toward more holistic reviews of applications.

Use of admissions tests is nearly universal at professional schools. The MCAT is required by almost all US medical schools and many Canadian medical schools, the GMAT is used by more than 2300 schools of business and management, according to its website, and the LSAT is the “only test accepted for admission purposes by all ABA [American Bar Association]-accredited law schools and Canadian common-law law schools,” according to its website. (An ABA standards document states that “[a] law school shall require each applicant ... to take a valid and reliable admission test ... A law school that uses an admission test other than the [LSAT] shall demonstrate that such other test is a valid and reliable test ...” ([American Bar Association, 2020](#)). The document provides for some exceptions to this rule). Undergraduate grades are also considered very important in professional school admissions, and medical schools place a high value on interviews as well ([Monroe et al., 2013](#)).

Predictive validity of admissions tests

College, graduate school, and professional school admissions tests are intended to be useful for predicting the grades students will receive in the educational programs they enter. Although a comprehensive evaluation of a test’s validity must involve a consideration of the test’s design, development, content, administration, and use, the validity of admissions tests as a selection tool for higher education institutions is, in practice, judged largely by the degree to which test scores can predict students’ grade-point averages in college, graduate school, or professional school (GPAs). Most often, first-year grades are used.

Conducting a predictive validity study requires that the GPAs for the cohort of interest be available so that the predicted GPAs can be compared to the GPAs actually earned by the admitted students. Predictive validity studies are usually conducted within a single institution, although results may later be averaged across institutions.

Linear regression analysis is typically applied to estimate an equation for predicting college, graduate school, or professional school GPA. The usual predictors are admissions test scores and high school grades (in the case of college admissions) or undergraduate grades (in the case of graduate or professional school admissions). The resulting multiple correlation, which can range from 0 to 1, provides an index of the effectiveness of the prediction equation. The regression analysis can then be repeated using prior grades alone as a predictor. Comparing the predictive effectiveness of the two equations yields an estimate of the value added by using admissions test scores. The simple correlations between test scores and subsequent GPAs are often examined as well.

A factor that complicates the interpretation of these correlations, or validity coefficients, is restriction of range: students whose test scores are too low to allow admission to a higher education institution will not have GPAs. Because of this restriction of range (of test scores, and, as a result, of other predictors and of GPAs as well), validity coefficients tend to be smaller for the admitted students than they would be for the entire population of applicants. Statistical corrections are often applied in an attempt to estimate how big the association would have been if the range had not been restricted (see Gulliksen, 1987: 165–166). In addition, corrections are sometimes applied to adjust for *criterion unreliability*, which is the imprecision of the measure that is being predicted. Except where noted, the correlations reported in this article have not been corrected in any way. Typically, corrected correlations are larger by 0.15–0.20.

A recent review of the predictive validity of admissions tests in college, graduate school, and professional school (Camara et al., Manuscript in preparation) reveals some consistent patterns in the findings. For both the ACT and the SAT, the correlation with college GPA was recently found to be 0.51 in large-scale studies after corrections were applied (Marini et al., 2019a; Westrick et al., 2021). In the case of the ACT, these results are based on analyses of the ACT composite score. In the case of the SAT, results are based on the total of the Math and Evidence-Based Reading and Writing scores. The correlations between admissions test scores and college GPA were somewhat lower than the correlations between high school grades and GPA, which were found to be 0.61 for ACT takers and 0.53 for SAT takers. Analyses that included admissions test scores as predictors along with high school grades yielded multiple correlations with GPA of 0.61 for the SAT (Marini et al., 2019a) and 0.64 for the ACT (Westrick et al., 2021).

In separate studies, the optional SAT Essay (no longer offered) was found to add useful information above and beyond the SAT Evidence-Based Reading and Writing section score in predicting college GPA (Marini et al., 2019b) and the optional ACT Writing test was found useful for predicting grades in entry-level college English composition courses (Radunzel, 2019).

Only a few studies have investigated the validity of the current version of the GRE, introduced in 2011, for predicting graduate school performance. In one such study, Schwager et al. (2015) evaluated the utility of GRE scores and undergraduate GPA as predictors of academic performance in a Dutch university. The research was based on a sample of 282 international student volunteers who took the GRE after they had already been admitted to master's programs. GRE scores were found to improve the prediction of final graduate school GPA beyond the level of accuracy attained with undergraduate GPA alone. Corrected correlations were 0.21 for Verbal Reasoning, 0.17 for Quantitative Reasoning, 0.31 for Analytical Writing, and 0.21 for undergraduate GPA.

Liu et al. (2016) studied the predictive validity of the GRE General Test based on data from more than 3000 students who enrolled in graduate study at the National University of Singapore between 2011 and 2014. Correlations between GRE scores and first-year graduate school GPA were 0.20 for Verbal Reasoning, 0.32 for Quantitative Reasoning, and 0.14 for Analytical Writing. These values were obtained using the combined data for master's and doctoral programs, with correction for range restriction. These correlations exceeded the correlation between undergraduate GPA and graduate school GPA.

Zimmerman et al. (2017) studied the validity of the GRE General Test for admission to master's programs at a technical university in Western Europe. Data was collected between 2006 and 2013; therefore, only a portion of the sample had taken the most recent version of the GRE. Concordance tables were used to place all GRE scores on a common scale. Based on a sample of 369 students, correlations with graduate school GPA were found to be 0.33 for Verbal Reasoning, 0.17 for Quantitative Reasoning, 0.30 for Analytical Writing, and 0.16 for undergraduate GPA.

Additional studies have led to favorable conclusions about the utility of the GRE as a predictor of performance in law school (Klieger et al., 2018) and in business school (e.g., Young et al., 2014).

In their review, Camara et al. (Manuscript in preparation) cite recent validity findings for the GMAT, LSAT, and MCAT. The GMAT total score (Verbal plus Quantitative) was found to have a correlation with graduate school GPA of 0.38, compared to 0.32 for undergraduate GPA. For the Integrated Reasoning and Analytical Writing sections, median correlations of 0.27 and 0.13, respectively, were reported, and a correlation of 0.51 was obtained when all four GMAT sections and undergraduate GPA were used as predictors (Talento-Miller, 2017). For the LSAT, the correlations with graduate GPA were 0.38 for LSAT score, 0.26 for undergraduate GPA, and 0.47–0.48 for both predictors combined (Anthony et al., 2016), and for the MCAT, a recent study found that the median corrected correlation over 16 schools was 0.57 for MCAT Total, 0.52 for undergraduate GPA, and 0.65 for both predictors combined (Busche et al., 2020).

In summary, college admissions test scores tend to be slightly weaker than high school grades as predictors of college GPA, but including test scores in addition to past grades in the prediction equation increases the correlation coefficient. (An exception to this pattern is a large study at the University of California that found that “[a]t UC, test scores are currently better predictors of first-year GPA than high school grade point average” (University of California Academic Senate, 2020: 20)). In recent research, test scores have been found to be more effective than undergraduate grades in predicting graduate school and professional school grades.

Although first-year GPA is the most common criterion variable in admission test validity studies, other criteria have been studied as well. For example, there is considerable evidence that admissions test scores are helpful in predicting grades beyond the first year (Kuncel and Hezlett, 2007; Zwick, 2006, 2007). The evidence on prediction of degree attainment is mixed (Zwick, 2006, 2007). The ACT and SAT appear to be useful for this purpose (e.g., Radunzel and Noble, 2012; Mattern et al., 2013), while graduate and professional school admission tests are less so (Kuncel and Hezlett, 2007). Some studies have examined prediction of other criteria, such as performance on licensing exams and measures of career success, with inconsistent results (see Zwick (2006) for a summary).

Use of personal characteristics to predict success in higher education

An area of ongoing debate is whether including formal measures of personal characteristics (sometimes called noncognitive factors) could improve the prediction of success in higher education while simultaneously making the admissions process more equitable. In a landmark 1985 study, Warren Willingham identified a characteristic he called “productive follow-through,” defined as “persistent and successful extracurricular accomplishment” in high school. Productive follow-through proved to be a poor substitute for high school grades and admissions test scores in predicting which students would be identified as most successful by their colleges, but taking this quality into account along with traditional academic qualifications substantially improved the prediction of success (Willingham, 1985: 184).

The question of whether measures of personal characteristics should be included in college admission decisions, however, is far from straightforward. The use of such characteristics can, intentionally or unintentionally, lead to the exclusion of certain demographic groups (Zwick, 2017a). In addition, personal characteristics are hard to measure in a fair and accurate way. One limitation in using self-reports about these attributes in the admissions context is that the desired answers are often obvious to the respondent. Another drawback is that some measures of these characteristics have shown patterns of group differences similar to those typically seen on standardized tests, on which white students tend to receive higher scores than other students (Zwick, 2017b). Other measures have been shown to have little or no power to predict college performance (e.g., see Thomas et al., 2007: 648).

Although there is widespread agreement that factors such as perseverance, grit, and ability to delay gratification are relevant to success in higher education, measuring these characteristics accurately and without bias is difficult. Thus, relying on observable evidence of the desired characteristics (such as the completion of a long-term research project) or on information from respondents other than the candidates themselves (possibly in the context of a standardized letter of recommendation) may yield more useful data.

Fairness of admissions tests to people of color, women, and other special populations

Standardized admissions test results often reveal substantial average score differences across ethnic, gender, and socioeconomic groups. Among the general public, these differences are often regarded as sufficient evidence that these tests are biased. From a psychometric perspective, a test’s fairness is inextricably tied to its validity. According to Cole and Moss (1989), test bias occurs (i.e., fairness is violated) “when a test score has meanings or implications for a relevant, definable subgroup of test takers that are different from the meanings or implications” for other test-takers. “... [B]ias is differential validity of a given interpretation of a test score ...” (p. 205).

Psychometric assessment of the fairness of admissions tests typically comprises two broad types of investigations. One type consists of analyses of differential prediction of a criterion, usually a GPA. Two distinct questions are typically addressed: first, are the test scores equally predictive of later grades for all groups? This question is typically investigated by obtaining separate prediction equations for each group and then comparing correlation or regression coefficients across groups. Second, if we obtain a single prediction equation for students as a whole, does this equation produce predicted GPAs that are systematically too high or too low for some groups? (See Zwick, 2017b).

Another common component of fairness assessment is an examination of item content. Before a test question is approved for inclusion in an admissions test, it undergoes a “sensitivity review” to eliminate any content that may be disturbing or offensive to test takers. Later, after the test is administered, a differential item functioning (DIF) screening is performed to determine whether equally skilled members of different groups (e.g., men and women) have statistically different rates of correct response on some items. The purpose of DIF analysis is to identify test items with content that may be problematic for some student group. For example, the item may include content that is irrelevant to the construct being assessed, and is more familiar to some student groups than others, such as sports content in a mathematics test. Items found to have DIF are either discarded immediately or flagged for further study.

Zwick (2012) summarizes findings on differential prediction of GPAs for key student groups, including people of color, women, people with disabilities, and students with limited English skills, provides a discussion of DIF findings, and addresses fairness issues associated with test coaching.

A sampling of higher education admissions testing around the world

Higher education admissions processes vary internationally in terms of the degree of government oversight, the degree of centralization, and the role, if any, of admissions tests. Some examples follow.

University admissions in China is based on a highly competitive government-controlled system that relies almost entirely on the National College Entrance Examination, or gaokao (Reshetar and Pitts, 2020). The 9-hour test, which is administered annually, is

typically taken by students in their last year of high school. Each provincial administrative region is assigned an admissions quota; provincial administrators then determine the minimum gaokao score required for admission to each tier of Chinese universities for candidates from that province. Prospective students may only apply to universities in tiers for which they exceed the minimum score. Admissions decisions are made based on additional score requirements associated with each university and program (Zhang and Wang, 2018).

There are two versions of the gaokao—one for students who want to study scientific fields and one for those who wish to pursue studies in the social sciences. All students must respond to sections in Mandarin Chinese, Math, and a foreign language. Those taking the natural sciences exam complete additional sections that may include Physics, Chemistry, or Biology, while those taking the social sciences exam complete sections that may include Political Sciences, History, or Geography (Reshetar and Pitts, 2020). The composition of the gaokao has been in flux in recent years and the Chinese Ministry of Education has reported that some students now have the opportunity to choose “new subject combinations instead of following the former mandatory path of either liberal arts or science” (Gaokao Reform, 2018).

Two major controversies in Chinese university admissions in the last decade have involved the exclusive focus on test scores in admissions and the under-representation of residents of particular provinces (Zhang, 2010). According to the Chinese Ministry of Education, some experimental admissions programs have taken into account students’ past academic achievements along with their gaokao scores. The Ministry also reports an increase in the admission of “rural and underprivileged students” (Gaokao Reform, 2018).

In Sweden, the admissions process for all public institutions (which constitute 30 of the 47 postsecondary institutions) is regulated by the government (Lyrén and Wikström, 2020). A distinction is made between eligibility requirements and selection criteria. In order to be admitted, all applicants must meet eligibility requirements; selection criteria are imposed only when a program has more eligible applicants than there are spaces. Eligibility is typically achieved by successfully completing secondary school. Selection may be based on upper-secondary-school grades, scores on the Swedish Scholastic Aptitude Test (SweSAT), or institution-specific requirements. While there are some exceptions, government regulations require that each institution select at least one-third of its students on the basis of grades and at least one-third on the basis of SweSAT scores. In practice, institution-specific requirements are rarely used (Lyrén and Wikström, 2020). The SweSAT consists of a Verbal section, which includes vocabulary; Swedish reading comprehension; sentence completion; and English reading comprehension, and a Quantitative section, which includes mathematical problem-solving; quantitative comparisons; data sufficiency; and diagrams, tables, and maps. Various proposals to reform Sweden’s eligibility and selection policies are currently under consideration.

In Israel, admission to higher education is generally based on two criteria—the Bagrut Certificate, or BC, administered by the Ministry of Education, and the Psychometric Entrance Test (PET), developed and administered by the National Institute for Testing and Evaluation, an independent body governed by the leaders of Israel’s research universities (Allalouf et al., 2020; Beller, 2001). The BC is the average score on tests administered by the Ministry and assessments conducted within the applicant’s high school. The PET, which is available in Hebrew, Arabic, English, and five other languages, contains sections in Verbal Reasoning (including an essay), Quantitative Reasoning, and English as a foreign language. National concerns about the fairness of the PET to various groups and the possible advantages of test coaching parallel those in the US. In response to social concerns about the admissions process, some alternate routes have been developed for use in certain cases, such as the admission of the top five percent of students in a high school or the inclusion of noncognitive tools or interviews (Allalouf et al., 2020).

Chile has a centralized national admissions system that is used to select studies for nearly all public and private universities. All applicants must take the national admissions test, the Prueba de Selección Universitaria (PSU) and submit a rank-ordered list of up to ten programs to which they wish to apply. A matching algorithm is used to assign students to programs, based on their PSU scores, high school grades, and rank in high school. The PSU consists of two mandatory tests—Language and Math—and two optional tests in Social Science and Science (which, in turn, involves a choice among Biology, Physics, and Chemistry). Program requirements determine which of these optional tests an applicant must take. Recently, the PSU has been the subject of major protests, mainly because of its perceived unfairness to students who participate in the technical-professional curriculum in high school rather than the scientific-humanistic curriculum (Koljatic and Silva, 2020). Various test reforms have been proposed, but their future is unknown.

References

- Allalouf, A., Cohen, Y., Gafni, N., 2020. Higher education admissions practices in Israel. In: Oliveri, M.E., Wendler, C. (Eds.), *Higher Education Admissions Practices: An International Perspective*. Cambridge University Press, Cambridge, UK, pp. 174–189.
- American Bar Association, 2020. 2020–2021 Standards and Rules of Procedure for Approval of Law Schools. Retrieved from: https://www.americanbar.org/groups/legal_education/resources/standards/.
- Anthony, L.C., Dalessandro, S.P., Trierweiler, T.J., 2016. Predictive Validity of the LSAT: A National Summary of the 2013 and 2014 LSAT Correlation Studies. Law School Admission Council Technical Report 16-01. LSAC, Newtown, PA.
- Beller, M., 2001. Admission to higher education in Israel and the role of the Psychometric Entrance Test: educational and political dilemmas. *Assess. Educ.* 8 (3), 316–337.
- Busche, K., Elks, M.L., Hanson, J.T., Jackson-Williams, L., Manuel, R.S., Parsons, W.L., Wofsy, D., Yuan, K., 2020. The validity of scores from the new MCAT exam in predicting student performance: results from a multisite study. *Acad. Med.* 95 (3), 387–395.

- Camara, W., Liu, L., Mattern, K., (Manuscript in preparation). Assessment for admission, placement, and outcomes in higher education. In: L. Cook, M. J. Pitoniak, (Eds.), Educational Measurement, fifth ed. Oxford University Press.
- Clinedinst, M., 2019. 2019 State of College Admission. National Association for College Admission Counseling. Retrieved from: https://www.nacacnet.org/globalassets/documents/publications/research/2018_soca/soca2019_all.pdf.
- Cole, N.S., Moss, P.A., 1989. Bias in test use. In: Linn, R.L. (Ed.), Educational Measurement, third ed. American Council on Education/Macmillan, New York, pp. 201–219.
- Gaokao Reform: China's College Admission Process to Be Fairer and More Merit-Based, 2018. Ministry of Education of the People's Republic of China. Retrieved from: http://en.moe.gov.cn/Specials/Review/Highlights/201806/t20180606_338546.html.
- Graduate Management Admission Council, 2020. GMAT Handbook. Retrieved from: <https://www.mba.com/-/media/files/mba2/assessments/2018/gmat/gmat-handbook-2020-07-10.pdf>.
- Gulliksen, H., 1987. Theory of Mental Tests. Erlbaum, Hillsdale, NJ.
- Jaschik, S., 2020. Coronavirus Drives Colleges to Test Optional. Inside Higher Education. Retrieved from: <https://www.insidehighered.com/admissions/article/2020/03/30/coronavirus-leads-many-colleges-including-some-are-competitive-go-test>.
- Klieger, D.M., Bridgeman, B., Tannenbaum, R.J., Cline, F., Olivera-Aguilar, M., 2018. The Validity of GRE General Test Scores for Predicting Academic Performance at U.S. Law Schools (ETS Research Report 18-26). Retrieved from: <https://onlinelibrary.wiley.com/doi/ful/10.1002/ets2.12213>.
- Koljatic, M., Silva, M., 2020. Chile's admissions tests: pending changes and revisions. In: Oliveri, M.E., Wendler, C. (Eds.), Higher Education Admissions Practices: An International Perspective. Cambridge University Press, Cambridge, UK, pp. 145–160.
- Kuncel, N.R., Hezlett, S.A., 2007. Standardized tests predict graduate students' success. *Science* 315, 1080–1081.
- Kuncel, N.R., Hezlett, S.A., Ones, D.S., 2001. A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: implications for graduate student selection and performance. *Psychol. Bull.* 127, 162–181.
- Lawrence, I., Rigol, G., Van Essen, T., Jackson, C., 2004. A historical perspective on the content of the SAT. In: Zwick, R. (Ed.), Rethinking the SAT: The Future of Standardized Testing in University Admissions. RoutledgeFalmer, New York, pp. 57–74.
- Liu, O.L., Klieger, D.M., Bochenek, J.L., Holtzman, S.L., Xu, J., 2016. An Investigation of the Use and Predictive Validity of Scores from the GRE Revised General Test in a Singaporean University (GRE Board Research Report No. GRE-16-01). Educational Testing Service, Princeton, NJ.
- Lopez, S., Schwartz, E., Ward, J.D., Pisacreta, E.D., 2020. Going Test-Optional with Equity in Mind. Ithaka S+R. Retrieved from: <https://sr.ithaka.org/blog/going-test-optional-with-equity-in-mind/>.
- Lyrén, P.-E., Wikström, C., 2020. Admissions practices in Sweden. In: Oliveri, M.E., Wendler, C. (Eds.), Higher Education Admissions Practices: An International Perspective. Cambridge University Press, Cambridge, UK, pp. 203–216.
- Marini, J.P., Westrick, P.A., Young, L., Ng, H., Shmueli, D., Shaw, E.J., 2019a. Differential Validity and Prediction of the SAT: Examining First-Year Grades and Retention to the Second Year (College Board Report). Retrieved from: <https://files.eric.ed.gov/fulltext/ED597325.pdf>.
- Marini, J.P., Westrick, P.A., Young, L., Shmueli, D., Shaw, E.J., Ng, H., 2019b. Validity of SAT Essay Scores for Predicting First-Year Grades (College Board Report). Retrieved from: <https://files.eric.ed.gov/fulltext/ED599389.pdf>.
- Mattern, K.D., Shaw, E.J., Marini, J., 2013. Does College Readiness Translate to College Completion? (College Board Research Note No. 2013-9). The College Board, New York.
- Monroe, A., Quinn, E., Samuelson, W., Dunleavy, D.M., Dowd, K.W., 2013. An overview of the medical school admission process and use of applicant data in decision making: what has changed since the 1980s? *Acad. Med.* 88 (5), 672–681. <https://doi.org/10.1097/ACM.0b013e31828bf252>.
- National Center for Education Statistics, 2020. Characteristics of degree-granting postsecondary institutions. In: The Condition of Education 2020. Retrieved from: https://nces.ed.gov/programs/coe/indicator_csa.asp.
- Office of Technology Assessment, 1992. Testing in American Schools: Asking the Right Questions. U.S. Government Printing Office, Washington, DC.
- Posselt, J.R., 2016. Inside Graduate Admissions: Merit, Diversity, and Faculty Gatekeeping. Harvard University Press, Cambridge, MA.
- Radunzel, J., Noble, J., 2012. Predicting Long-Term College Success Through Degree Completion Using ACT Composite Score, ACT Benchmarks, and High School Grade Point Average (ACT Research Report 2012(5)). ACT, Inc, Iowa City.
- Radunzel, J., 2019. Relating ACT Writing Scores to First-Year English Composition Grades (ACT Research & Policy Technical Brief). Retrieved from: <https://www.act.org/content/dam/act/unsecured/documents/R1789-writing-to-comp-grades-2019-11.pdf>.
- Reshetar, R., Pitts, M.F., 2020. General academic and subject-based examinations used in undergraduate higher education admissions. In: Oliveri, M.E., Wendler, C. (Eds.), Higher Education Admissions Practices: An International Perspective. Cambridge University Press, Cambridge, UK, pp. 237–255.
- Robin, F., Steffen, M., Liang, L., 2014. The multistage test implementation of the GRE revised general test. In: Yan, D., von Davier, A., Lewis, C. (Eds.), Computerized Multistage Testing. CRC Press, Boca Raton, Florida, pp. 325–341.
- Schmötter, J.W., 1993. The graduate management admission council: a brief history 1953–1992. *Selections* 9 (2), 1–11.
- Schwager, I.T.L., Hülshager, U.R., Lang, J.W.B., Bridgeman, B., 2015. Graduate student selection: GRE, socioeconomic status and undergraduate grade point average as predictors of study success in a Western European University. *Int. J. Sel. Assess.* 23, 71–79.
- Stewart, D.M., 1998. Why Hispanic Students Need to Take the SAT. The College Board. Retrieved from: <http://www.collegeboard.org>.
- Talento-Miller, E., 2017. Differential Validity and Differential Prediction of the GMAT Exam. GMAC. Retrieved from: <https://www.gmac.com/-/media/files/gmac/research/validity-and-testing/tr-17-01-differential-validity-talento-miller-web-release-2.pdf>.
- Thomas, L.L., Kuncel, N.R., Credé, M., 2007. Noncognitive variables in college admissions: the case of the Non-Cognitive Questionnaire. *Educ. Psychol. Meas.* 67 (4), 635–657.
- University of California Academic Senate, 2020. Report of the UC Academic Council Standardized Testing Task Force. Retrieved from: https://senate.universityofcalifornia.edu/_files/committees/sttf/sttf-report.pdf.
- Webber, C., 1989. The Mandarin mentality: university admissions testing in Europe and Asia. In: Gifford, B.R. (Ed.), Test Policy and the Politics of Opportunity Allocation: The Workplace and the Law. Kluwer, Boston, MA, pp. 33–57.
- Westrick, P.A., Schmidt, F.L., Le, H., Robbins, S.B., Radunzel, J., 2021. The road to retention passes through first-year academic performance: a meta-analytic path analysis of academic performance and persistence. *Educ. Assess.* 26 (1), 35–51.
- Willingham, W.W., 1985. Success in College: The Role of Personal Qualities and Academic Ability. College Entrance Examination Board, New York.
- Young, J.W., Klieger, D., Bochenek, J., Li, C., Cline, F., 2014. The Validity of Scores from the GRE® Revised General Test for Forecasting Performance in Business Schools: Phase One (GRE Board Research Report No. 14-01; ETS Research Report No. RR-14-17). Educational Testing Service, Princeton, NJ.
- Zhang, H., Wang, J., 2018. Effective predictions of Gaokao admission scores for college applications in mainland China. In: Annual International Conference on Computational Mathematics. Computational Geometry & Statistics, pp. 163–170.
- Zhang, R., 2010. Media, litigation, and regional discrimination in college admission in China. *Chin. Educ. Soc.* 43 (4), 60–74. <https://doi.org/10.2753/CED1061-1932430406>.
- Zimmermann, J., von Davier, A.A., Buhmann, J.M., Heinemann, H.R., 2017. Validity of GRE General Test scores and TOEFL scores for graduate admission to a technical university in Western Europe. *Eur. J. Eng. Educ.* <https://doi.org/10.1080/03043797.2017.1343277>.
- Zwick, R., 2006. Higher education admissions testing. In: Brennan, R.L. (Ed.), Educational Measurement, fourth ed. American Council on Education/Praeger, Westport, CT, pp. 647–679.
- Zwick, R., 2007. College Admissions Testing. Paper Commissioned by the National Association for College Admission Counseling. <http://www.nacacnet.org>.
- Zwick, R., 2012. Admissions testing in college and graduate education. In: Secolsky, C., Denison, B. (Eds.), Handbook on Measurement, Assessment, and Evaluation in Higher Education. Routledge, New York, pp. 382–404.

- Zwick, R., 2017a. The Risks of Focusing on Character in Admissions. Chronicle of Higher Education. Retrieved from: <http://www.chronicle.com/article/The-Risks-of-Focusing-on-240787?cid=wsinglestory>.
- Zwick, R., 2017b. Who Gets In? Strategies for Fair and Effective College Admissions. Harvard University Press, Cambridge, MA.

Testing program websites

- ACT, <https://www.act.org>.
- GMAT, <https://www.mba.com/exams/gmat>.
- GRE, <https://www.ets.org/gre/>.
- LSAT, <https://www.lsac.org/lSAT>.
- MCAT, <https://students-residents.aamc.org/applying-medical-school/taking-mcat-exam/>.
- SAT, <https://collegereadiness.collegeboard.org/sat>.

Intelligent tutoring systems: a history and an example of an ITS for science

Janice D. Gobert^{a,b}, Michael A. Sao Pedro^b, Haiying Li^c, and Christine Lott^a, ^a Rutgers Graduate School of Education, New Brunswick, NJ, United States; ^b Apprendis, LLC, Massachusetts, MA, United States; and ^c Iowa College Aid, IA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to ITS	460
Core components of ITS	461
History of ITS	461
Computer-assisted instruction	461
From CAI to ITS	462
ITS developments from the 1980s to present	462
Buggy-based approach	462
Model tracing	462
Case-based reasoning	462
Natural language processing	462
Discovery worlds and simulations	463
Mental models	463
Virtual reality	463
Learner control	463
Collaborative learning	463
Tutoring for affective states and engagement	464
Game-based ITSs	464
Student modeling	464
Machine learning/data mining	464
The move towards real time performance assessment & real time scaffolding of inquiry competencies	465
Scaffolds in Inq-ITS	467
Conclusions	467
Acknowledgments	468
References	468

Introduction to ITS

The origins of ITSs can be traced back to the 1950s and what is known as Computer Assisted Instruction, or CAI, which refers to the use of computers in the field of education. It was the introduction of “intelligence” to CAI in the 1970s that paved the way for a new genre known as ITS (Ahuja and Sille, 2013; Corbett et al., 1997). An ITS is a computer program that uses computational techniques to provide personalized instruction to students. This modern day intelligent tutor “knows” what to teach and, in theory, how best to do so. Perhaps most importantly, the tutor knows who is being taught, including their knowledge level, the misconceptions they may hold, their skills and abilities, and in some cases, their motivation and affective state. Research and development of ITSs weave together ideas from three different disciplines, including, computer science, education, and psychology (Alkhatlan and Kalita, 2018; Nwana, 1990). The interconnection between the three disciplines has advantages in terms of the breadth of research questions that can be conducted, which in turn, can contribute to advances in ITS technology. However, there also have been some disadvantages noted, namely that researchers need a firm grasp on all the relevant disciplines, and that those of different backgrounds may have competing goals, varying foci, and disagreement on frameworks and terminology (Nwana, 1990). Nonetheless, this area has generated technological innovations that have begun to successfully bridge the gap between human and computer when it comes to tutoring in education.

An important milestone in ITSs was when the target in terms of efficacy for student learning was set by a classic study by Bloom (1984), when he demonstrated that students who were guided by a human tutor performed two standard deviations better than those who studied in a traditional classroom or group learning environment. This quickly became the goal of ITSs in terms of efficacy at improving students’ learning.

A number of meta-analyses have been conducted on the effectiveness of ITSs, since this is critical to their improvement (Kulik and Fletcher, 2016; Ma et al., 2014; VanLehn, 2011). For example, VanLehn (2011) compared the effectiveness of human tutoring, computer tutoring, and no tutoring, and found that ITSs can be nearly as effective as human tutors.

Ma et al. (2014) also conducted a meta-analysis to compare the learning outcomes of students who used ITSs and those who did not; here other factors such as type of ITS, instructional setting (individual, small group, etc.), and subject domain (math, physics, etc.) were included. In this meta-analysis, it was found that a learning environment that included an ITS was more effective than teacher-led or large group instruction, non-ITS computer-based instruction, and textbooks or workbooks. Further, positive effect

sizes were found at differing levels of education, across different domains, and for different types of ITSs. Notably, however, no significant difference was found between ITS use and individual human tutoring or small group instruction. In another meta-analytic study by Kulik and Fletcher (2016), for 92% of the 50 studies examined, students who used ITSs outperformed students who did not. Further, the use of ITSs raised test scores, whether they were locally developed or standardized, over conventional classroom teaching environments. Thus, meta-analyses such as these provide evidence that suggests that ITSs can be effective at helping students learn.

Core components of ITS

It is widely accepted that typical intelligent tutoring systems (ITS) have four basic components (Corbett et al., 1997; Nwana, 1990). The *domain model*, also known as the expert model or cognitive model, which represents the knowledge, facts, concepts, and rules that belong to the domain being taught (e.g., physics or mathematics). The *student model*, which represents the students' knowledge, including the cognitive and, in some cases, the affective states of the students while they interact with the tutoring system. As the student responds to questions and solves problems, this information can be used to construct appropriate feedback. The *tutoring model*, also known as the pedagogical or teaching model, devises instructional strategies that can be used to provide customized feedback to students by integrating information from the domain/expert model and the student model. For example, the tutoring model can initiate actions, such as giving a hint to the student. Finally, the *user-interface model* is the interactive environment that the student uses to engage the system. It includes the graphics, text, and multimedia sources available to the student and presents the available actions a student can take during the tutoring process. Later in the entry we will provide linkages between each of these (domain model, student model, etc.) and Inq-ITS (Inquiry Intelligent Tutoring System, inqits.com; Gobert et al., 2013, 2016, 2019; Mislevy et al., 2020), where relevant, for illustrative purposes.

History of ITS

Computer-assisted instruction

In its earliest stage, CAI was based on the work of B.F. Skinner and his principles of operant conditioning (Skinner, 1948). In a Skinnerian system, a student was presented with simple questions and then given the correct answer. The goal was to lead the student toward a desired end behavior. The series of "frames" were presented regardless of correctness because wrong answers did not help students learn the correct behavior, according to Skinner. In a frame-oriented system like this, designed to proceed regardless of how the student responds, there is a lack in the capacity to provide any personalized feedback, which later systems compensated for (Nwana, 1990). Further, in this linear style, all of the questions and answers had to be pre-determined, and as such, may not provide appropriate or helpful responses for students (Rickel, 1989).

In the 1960s, Crowder (1959) introduced branching, or intrinsic programs that no longer ignored student responses. Instead, they were used to determine the next frame to be presented to a student (Edwards, 1968; Nwana, 1990). In intrinsic programs, each response to a multiple-choice question is linked to another page in the program. When the student responded with the correct answer, they were presented with subsequent information. If the student selected the wrong choice, they were taken to a page that explained why that choice was incorrect, and then were prompted as needed to retry with a similar question (Edwards, 1968). These branching types of ITSs were also limited as students were confined to multiple choice questions and given no robust feedback, critical to improving learning.

Emerging in the late 1960s and early 1970s, generative systems elucidated and implemented the idea that tutoring material could be generated by the system itself; this soon followed as an improvement to the limited, step-wise nature of frame-oriented systems. Generative systems work by using textual templates, problem-generation grammars, and random numbers so the system can supply as many problems or examples as a student might need to improve learning (Nwana, 1990). An early example is Wexler's system, in which information is encoded in a network and then used to generate new questions and responses (Nwana, 1990; Rickel, 1989). Wexler grouped objects into classes that could then be linked to other classes, such as "family," which could be linked to "mother, father, and children" (Rickel, 1989). These systems were still insufficient however, in that they failed to interact with or adjust to the student's needs and responses. Moreover, they were restricted to drill-type exercises in well-defined domains such as mathematics that have structured content with objective answers (Nwana, 1990; Rickel, 1989). Further, generative systems have an inadequate representation of the students' domain knowledge, making it difficult to tailor feedback in a personalized or adaptive manner (Nwana, 1990; Rickel, 1989).

In early CAI, the student model component used in typical ITS architecture often lacked fine-grained user information and the functionality to provide an adaptive and personalized tutoring experience, which is ultimately the goal of ITSs. In one example, namely, the Leeds Adaptive Arithmetic System (Nwana, 1990), a rudimentary attempt at fulfilling the student model involved using a single integer to represent a student's competence level. In another example of an ITS with a simple student model, the KEYSTROKE model was able to determine the productivity of a user by counting their keystrokes in a basic task such as searching for information (Sleeman, 1985). Another issue, according to Sleeman, is that many early systems did not distinguish an expert from a novice in a given subject area and thus, did not lead to effective tutoring. Sleeman acknowledged that student modeling is an open-ended task and thus requires a system with the capability to interpret and respond to the diverse

knowledge, skills, and abilities of students; however, in well-defined domains, the system can make more simplified assumptions about a user than in ill-defined domains. Later in the chapter we provide a more detailed section on student modeling, as it is a key factor in ITS development and efficacy in supporting students' learning.

From CAI to ITS

With a lack of fine-grained user knowledge (i.e., student model) in early CAI systems, experts soon realized the potential gains of incorporating "intelligence" into these systems. Self (1974) argued that to be most effective, CAI should include a representation of what is being taught, who is being taught, and how to teach it. These categories, respectively, correspond to the current ITS architecture of the domain model, student model, and tutoring model. These improved systems have a more sophisticated representation of the student model, one in which the current knowledge state is used to determine appropriate and personalized instruction (Rickel, 1989).

The first intelligent tutoring program, known as SCHOLAR, marked the introduction of AI to CAI, and ushered in a new generation of systems, which then became labeled as Intelligent Tutoring Systems (Nwana, 1990). SCHOLAR engaged students in dialog, allowing them to ask and answer questions (Carbonell, 1970). This system used a semantic network of units, which form a complex network of facts, concepts, and in some cases, procedures (Ahuja and Sille, 2013). Nodes in the network were tagged if the student recognized a concept. Further, this system exemplified a mixed-initiative approach, whereby the student as well as the tutor could ask (initiate) questions and dialog (Shute and Psotka, 1994). Since this system used a natural language dialog, if the system did not understand the student's question or response, the learning process was halted.

ITS developments from the 1980s to present

Buggy-based approach

By the 1980s, ITS development grew and various systems were developed with different foci in mind. One example is the buggy approach (Brown and Burton, 1978), which not only identifies a student's misconceptions, or bugs, but helps to explain why a student is making a mistake. One demonstration of a system based on the buggy approach is PROUST (Johnson and Soloway, 1985), which targets errors in students' computing of descriptive statistics, such as calculations of range and averages (Shute and Psotka, 1994). A report is submitted to PROUST after the student completes the program, and then the system locates errors and sends a diagnostic report (Shute and Psotka, 1994).

Model tracing

Another approach involves model tracing in which the system tracks a student step-by-step during a problem solving task and then compares this to the expert domain to provide personalized feedback as the student progresses (Merrill et al., 1992). The objective is to sequence the student's problem-solving actions within a complex problem space that may contain thousands of solutions (Corbett et al., 1997). One criticism of model-tracing, however, is the frequent interference from the tutor, which can limit a student's opportunity to learn from or reflect on his own mistakes (Merrill et al., 1992).

Case-based reasoning

Case-based reasoning (CBR) also appeared in the 1980s, which involves applying prior experiences and memories, or cases, to new information in order to interpret problems and find solutions (Kolodner, 1992). This can be applied to the human reasoning process as well as in the approach and design of ITS systems. Kolodner (1983) introduced the first CBR system, known as CYRUS, which is an intelligent fact retrieval system that stores and retrieves events of former Secretaries of State, Cyrus Vance and Edmund Muskie. If given a new fact, it will organize it within the existing database and can also call up a fact when asked, such as "When was the last time Vance was in Egypt?". Another example is CHEF, which created new recipes from recipes it already knew. By accessing a recipe with chicken and peas, it was able to substitute beef and broccoli and adjust the steps in the directions accordingly (Kolodner, 1992). Case-based reasoning has the potential to provide a more adaptive learning environment because it suggests approximate answers rather than providing an absolute response. This also makes it more suitable for ill-defined domains (i.e., philosophy), or environments in which there are too many rules or a rule that can be applied in many ways, such as game playing or programming (Ahuja and Sille, 2013; Shute and Psotka, 1994). Science inquiry is another ill-defined domain, and will be addressed later in the entry (Gobert et al., 2013).

Natural language processing

Another focus of development in ITSs in the 1980s was Natural Language Processing (NLP; Burstein, 2009; Burton, 1976; Chowdhury, 2003). NLP research has helped create a more natural flow in communication between a human and machine, or student and tutor by gathering information about how humans use and understand language and then using it to interact with the user (Chowdhury, 2003). An early example of a system built on NLP is SOPHIE (Burton, 1976), which applied semantic grammar techniques in a well-defined domain. SOPHIE's expert domain was electronic troubleshooting, which simulated faults on which the student could apply their theoretical knowledge to form hypotheses and attempt experiments to troubleshoot the problem (Nwana, 1990).

Discovery worlds and simulations

Another product of ITS development throughout the 1980s was discovery worlds (Abuja and Sille, 2013) or simulations. A simulation is a model of scientific (or social) phenomena that represents the domain-specific properties and conceptual representations of that phenomena. A simulation may emulate some features of real phenomena but not others. Originally referred to as micro-worlds by Papert as a "... subset[s] of reality or a constructed reality so ... as to allow a human learner to *exercise* particular powerful ideas or intellectual skills" (Papert, 1980, p. 204). For example, digital simulations enable learners to change aspects of phenomena, i.e., speed up, slow down, change size, or simplify aspects of real-world phenomena. Learners can explore phenomena, which are not possible to explore in real life due to their size and or time scale (Gobert and Clement, 1999), or engage in due to safety issues.

The goal with simulations is for students to use a computer simulation environment to learn content and skills on their own (Shute and Psotka, 1994). The notion of student-led discovery, however, is at odds with an ITS since the goal of ITS is to provide feedback to guide students' learning. For example, White (1984) created a system that allowed students to explore Newton's laws of motion in a discovery world (a simulation) by controlling a spaceship and navigating it toward a target or through a maze. While these types of systems allow for more freedom on the part of the student to guide their own learning processes, it does demand some knowledge and skill to operate within the microworlds. Students must know what to manipulate and how to do so, i.e., design an experiment or support a hypothesis, etc. (Shute and Psotka, 1994).

Mental models

Related to discovery worlds and simulations is the psychological construct invoked to describe the form of the cognitive representations that are constructed when learners engage in deep learning (Johnson-Laird, 1983). More specifically, model-based learning assumes that understanding requires the construction of mental models of the phenomena under study, and that all subsequent problem-solving, inferencing, or reasoning are done by means of manipulating or "running" these mental models (Gobert and Buckley, 2000). Using the notion of mental models, the 1980s saw systems that leveraged progressively more complex causal models presented in ITSs to, in turn, support the development of students' mental models (Abuja and Sille, 2013). These systems, like QUEST (Qualitative Understanding of Electrical System Troubleshooting; White and Frederiksen, 1990), point out errors in students' thinking and guide students through increasingly complex microworlds. They then allow students to test their ideas by applying the set of laws to see if it predicts real-world events (Shute and Psotka, 1994).

Virtual reality

Going beyond simulations are more deeply immersive-type environments called virtual reality, another innovation in ITSs during the 1990s. These were influenced by situated learning, which emphasized the importance of the physical (and social) context on students' learning processes (Brown et al., 1989). With virtual reality, users can experience a simulated 3-dimensional environment in which they can move about and manipulate aspects of the environment to learn about phenomena as well as new skills (Shute and Psotka, 1994). Virtual reality provides learning opportunities and experiences in a similar manner to simulations, i.e., experiences that would otherwise not be possible or feasible. For example, Perez-Valle and Sagasti (2012) developed a system that allowed students to visit 16th century Spain and tour buildings and temples that no longer exist.

Learner control

Moving into the 1990s, one major issue among researchers in ITS development centered on the topic of learner control. On one side of the issue are those who believe the learners should be in control of their learning processes, while others believe that structure and guidance are necessary components to a learner's success. When applied to the field of ITS, more learner control equates to a learning environment in which students can explore without many restraints imposed by the system. In a program called Smithtown (Shute and Psotka, 1994), which teaches basic microeconomics, students can explore the environment freely and then design their own experiments by choosing variables, testing a hypothesis, and analyzing their data. The system provides a notebook and table to collect and organize data, and graphing tools to allow students to plot their data. For example, students can select a market to explore, such as gasoline, and then adjust the price, and analyze the effects on the market by making a graph reflecting the supply and demand (Shute et al., 1988). In science, the degree of learner versus system "control" used to guide students' activities in learning environments, particularly those involving simulations and/or microworlds, is a topic that has been hotly debated (Hmelo-Silver, 2004; Kirschner et al., 2006). Some have encouraged open-ended exploration by students, and others have offered guidance or structure within the microworld to promote optimized learning. Later, we describe how Inq-ITS scaffolds students by both structuring inquiry into phases and using hints provided by a virtual tutor to help the student when the system detects that the student is struggling with a particular inquiry competency.

Collaborative learning

Further research in education and technology suggested collaborative learning can also enhance learning outcomes in the context of ITSs, which became another focus of ITS development in the 1990s (Alkhatlan and Kalita, 2018). There are certain characteristics that must be in place to achieve success in a collaborative learning group, such as participation and performance analysis (Soller, 2001). For example, in one such system for learning fractions, students could use audio chat to talk through the problem; students were also assigned roles such as helper or problem solver to guide them through the interaction and promote accountability (Olsen et al., 2014).

Tutoring for affective states and engagement

By the end of the 1990s, an emphasis on the emotional state of the learner became a new focus of ITS development. This was in part due to the hypothesis that engagement can influence learning (Gobert et al., 2015). Further, by being able to infer the emotional state of the learner, the system can then respond more appropriately to the needs of the student both cognitively and psychologically (Picard, 1997). For example, student engagement can be tracked via logfiles, which are fine-grained traces of students' interactions in a learning environment for science (Gobert et al., 2015). There also exist various methods to capture the affective state of the student including a computer's camera and microphones, specialized mice, and neuro-headsets, which collectively provide feedback regarding a student's facial expressions, voice, and physiological measures such as heart-rate and skin-conductivity (Alkhatlan and Kalita, 2018; D'Mello and Graesser, 2009). For example, in Auto-Tutor, a natural language ITS that has been used in domains such as computer literacy and physics, three sensors monitored students' affective state using posture patterns, videos of their face, and audio of their speech. This data was used to detect and differentiate emotions such as boredom, delight, and frustration (D'Mello and Graesser, 2009), and in turn was used to devise a targeted response, such as presenting an engaging task if the system detected a student's boredom. For more information on engagement and affective tutoring, see Andres et al. (2021), Baker et al. (2021), Henderson et al. (2021), Lester et al. (2019), and Mott et al. (2021).

Game-based ITSs

One of the more recent developments regarding ITSs is that of game-based learning, which also presupposes that students will learn more effectively if they are more deeply engaged in learning (Alkhatlan and Kalita, 2018). Though, a more apt statement in our opinion is that engagement is a *necessary but not sufficient condition* for learning, since deep learning is influenced by a variety of factors (Gobert, 2022). Rai and Beck (2012) created a math tutoring system with game-like elements, such as naming the monkey character students interact with or receiving badges after completing a level of learning. Despite many benefits to game-based systems, such as maximization of student participation, there also exist potential drawbacks associated with these types of systems (Alkhatlan and Kalita, 2018). Games often demand high cognitive load of users, which can lead students to feel overwhelmed and, in turn, diminish long-term learning of the material (Rai and Beck, 2012); this focus may be better utilized on content rather than the dynamics of the game.

Student modeling

Earlier we highlighted the core components of ITSs, including the student model. It is important to note that the adaptability and personalization of an ITS ultimately relies on the construction of an effective student model. By this, we mean that the student model leads to an understanding of the user by drawing on relevant information including a student's performance, misconceptions, prior knowledge, and in some cases, their personality characteristics. In brief, there are a number of approaches to building a student model including the overlay model, perturbation model, machine learning, cognitive theories, fuzzy logic, and Bayesian networks, with most ITSs using some combination of these (Chrysafiadi and Virvou, 2013), though it is beyond the scope of this entry to articulate the differences between all of these development approaches.

Three approaches that are noteworthy due to their popularity in the development of ITSs are the overlay model, stereotype model, and perturbation model (Chrysafiadi and Virvou, 2013). The overlay model was developed by Stansfield et al. (1976) and views the student's knowledge as a subset, or overlay, of the expert's knowledge, as represented by the domain model. The focus of the ITS then, is to minimize the gap in knowledge between the expert's and the novice's knowledge, skills, etc. (Alkhatlan and Kalita, 2018). The perturbation model is often viewed as an extension of the overlay model because it in part is a subset of the expert's knowledge, but unlike the traditional overlay model, it incorporates misconceptions or incorrect beliefs into this knowledge gap (Alkhatlan and Kalita, 2018). The stereotype model categorizes students based on shared characteristics, or stereotypes in order to guide them through an appropriate task or level based on their current knowledge state. An ITS may classify a new user into a category based on their level of knowledge about the content and can then infer information about the user based on other shared characteristics of that particular stereotype (Chrysafiadi and Virvou, 2013).

There are certain drawbacks, however, to each of these approaches. The overlay model is limited in that it does not take into account the learner's behavior and personality characteristics or the way users integrate new information. The perturbation model has demanding computational requirements and is limited in that the predetermined "bugs" in the system may not apply to some students. The stereotype model is inflexible in that the stereotype classes are pre-set by the system designer, which is time-consuming, and the users must be "of a type" that allow for discrete categorization in order to apply the stereotypes (Chrysafiadi and Virvou, 2013).

Machine learning/data mining

Through the early 2000s, the overlay, stereotype, and perturbation models were the most popular student modeling approaches. However, due to the limitations described above, a newer more adaptive and automated approach, known as machine learning or data mining, gained momentum (Chrysafiadi and Virvou, 2013). Though the term dates back to the 1980s, by the early 90s it was viewed as a sub-process of Knowledge Discovery in Databases or KDD (Coenen, 2011). Machine learning fulfills the need to automate the processes by which a user's actions and behaviors can be observed (Chrysafiadi and Virvou, 2013; Luan et al., 2020). Machine learning analyzes large amounts of data in order to extract meaningful information (Baker and Yacef, 2009; Coenen, 2011), and its techniques allow for better construction and greater adaptability of the student model (Ahuja and Sille, 2013). In the 1990s, machine learning saw a decline and then a resurgence because of the need for large data sets, labeled data, and

computational complexity (Webb et al., 2001). Presently, machine learning is increasingly used for the explanation, prediction, and description of student behavior (Yang et al., 2021), due in part, to an increase in internet usage and information retrieval technologies (Webb et al., 2001) of student data.

Machine learning techniques have mainly been used in two areas of student modeling research. First, to induce a student model from many student behaviors, and second to create a bug library of student models. For instance, Baker (2007) constructed a machine learning model that was able to detect when a student was off-task. Baker et al. (2012) detected emotions of students based on log data of actions with an Algebra ITS system, rather than with physical sensors that other affect detection systems may use. Finally, Baker et al. (2006) used data mining techniques to detect students who were trying to game the system by getting to a bottom-out hint rather than engaging in deep learning; students were then given supplementary exercises based on the ones they attempted to bypass. Gobert et al. (2015) used machine learning on students' logfiles within a science inquiry assessment and learning environment to detect students' disengagement from learning; Gobert et al. (2012, 2013) also used machine-learning for assessment of students' inquiry competencies.

The move towards real time performance assessment & real time scaffolding of inquiry competencies

As addressed earlier in the entry, many ITSs have been designed for well-defined domains. By this, we mean, domains that have one correct answer, such as ITSs for math problems (Corbett and Anderson, 1995; Feng et al., 2009; Gertner and van Lehn, 2000) like the addition of $2 + 3$, or double-digit addition, etc. For domains in which there is one correct answer, the assessment of students' knowledge and skills is fairly straight-forward, and computational techniques including machine-learning, described above, are not typically necessary. In these cases, rather, knowledge-engineering (cf. Koedinger and MacLaren, 2002) is used for both assessing students and scaffolding them. Briefly, in knowledge engineering, a domain expert constructs by hand a cognitive model to capture what it means to demonstrate a sub-component based on the constructs of the domain and/or cognitive task analyses (Ericsson and Simon, 1980), then techniques such as model tracing are used to compare students' knowledge and skills to that of an expert, and intervene when the system detects a difference between the two. Additionally, there are science domains, such as Physics and Chemistry, which include mathematics in their underlying formulas, equations, etc., and they too can use knowledge-engineering for these mathematical aspects of the domain. Some examples here are VanLehn's et al. (2005) Physics Tutoring system for college level Physics problems, which helps students by giving them assistance as they set up free-body diagrams, write equations, and solve them, and Walsh et al.'s (2002) intelligent tutoring system for balancing chemical equations.

Science inquiry, which is the means by which a learner engages in a series of tasks related to forming a testable hypothesis, conducting an experiment to test a hypothesis, analyzing a resulting data, and then communicating findings, is an ill-defined domain (Kuhn, 2005). It is ill-defined because there are many ways in which students can complete a task both productively as well as unproductively. Thus, both scaffolding and the assessment of students' competencies that is required in order to scaffold students (as is the goal of ITSs), is much more complex. Described next in this entry is Inq-ITS Inquiry-Intelligent Tutoring System (inqits.com; Gobert et al., 2013) in order to exemplify an ITS for science inquiry, and how machine learning is used in an ITS of this type. We also use Inq-ITS to describe the core components described above (where applicable for Inq-ITS). Lastly, we provide an overview of Inq-ITS scaffolds and an overview of some results of our scaffolding research.

Inq-ITS is a computer-based system, i.e., an ITS, that provides both real time assessment and scaffolding of science inquiry practices as described in national and international science frameworks, including the Next Generation Science Standards (NGSS, 2013). In brief, using simulations and a series of inquiry tools, widgets, etc., students formulate hypotheses and questions, test them, interpret their data with graphs and other mathematical means, warrant their claims with evidence, and communicate their findings. As students conduct their investigations, the work products they create and processes they use by conducting their investigations are automatically assessed using integrated machine-learning as well as knowledge-engineering, where relevant (Gobert et al., 2013, 2016, 2017, 2019; Mislevy et al., 2020). In addition, Inq-ITS supports teachers' real time assessment and instruction of inquiry practices (NGSS, 2013) using real time reports and actionable alerts via a dashboard called Inq-Blotter (Dickler et al., 2021; Gobert, 2019; Gobert et al., 2019; Gobert and Sao Pedro, 2016; Sao Pedro, 2015, 2016), though these are not the focus of the present chapter (see Dickler et al., 2021 for details on Inq-Blotter).

Inq-ITS was explicitly designed to prioritize the assessment and scaffolding of inquiry rather than the learning of science, which many other curriculum-focused inquiry systems emphasize (cf., Linn and Hsi, 2000). Inq-ITS was informed by the assessment triangle, which includes cognition, observations, and interpretation (Pellegrino et al., 2001), and by Evidence-Centered Design (Mislevy et al., 2020) in which the system elicits and collects evidence of students' proficiencies at inquiry practices (Gobert and Sao Pedro, 2017), i.e., this is performance assessment. In brief Evidence-Centered Design (ECD; Mislevy et al., 2009, 2020) is a design framework that helps specify all the key elements of an assessment system and an intelligent tutoring system, including the domain model, the interface model, the student model, and the tutoring model. ECD also specifies the fine-grained details and the computational processes (critical to interpretation of student data) that are to be carried out in the assessment system in order to generate fine-grained, rigorous assessment data on students' competencies, and in turn, scaffolding. This approach is different from the assessment of rote knowledge of facts or formulas, which can be assessed via multiple choice items. Rather, performance assessment of science inquiry is done as students engage in rich, authentic inquiry tasks with simulations so that the system can act *intelligently* to support students' learning.

Inq-ITS builds on the inquiry and assessment research of others (some of which was reviewed earlier) and forges new ground for inquiry assessment and scaffolding of science with its application of computational analytics including educational data mining, knowledge-engineering, and natural language processing (Gobert et al., 2013, 2016, 2017, 2019). The machine-learned algorithms allow the system to better capture what kinds of difficulties students have with inquiry, and then based on students' difficulties, Rex, a cartoon dinosaur that acts as a virtual pedagogical agent, jumps in to help the student in real time (See Fig. 1 below). These algorithms were built on diverse samples and tested for generalizability, which ensures equity with respect to the new students who later use these systems (Sao Pedro et al., 2013b).

To illustrate how the system works, as well as how the architectural components of an ITS are represented, a small vignette is presented. An eighth grade teacher, Mrs. Donahue, is having her class conduct an inquiry with Inq-ITS' States of Matter Physical Science virtual lab (See Fig. 1, left). In this lab, students determine the effects of the independent variables (e.g., level of heat, amount of substance) on the dependent variables (e.g., time to melt, temperature when melted) by forming a hypothesis, collecting data to test their hypothesis using the simulation, analyzing their data, and communicating their findings.

At a high level, students engage in virtual inquiry with our widgets and a domain-based simulation of a science phenomenon; this includes, forming a question/hypothesis, collecting data with a simulation, interpreting data, warranting claims with data, using mathematics to develop deeper understandings of the phenomena under study and communicating findings about the investigation (Gobert et al., 2012, 2013, 2016, 2019). The simulation environment represents the *domain model* of the science phenomena under investigation, while the simulations, widgets and clickable options presented to students in Inq-ITS demonstrate the *user-interface model*. The Inq-ITS activities leverage the logging infrastructure to store all students' interactions (logfiles) at a fine-grained level, including students' interactions with widgets, representational tools, and simulations, and timestamps for those interactions (Mislevy et al., 2020; Gobert et al., 2012; Sao Pedro et al., 2013a); these are used to construct a student model of the students' competencies across the full range of inquiry practices, which informs the tutoring model.

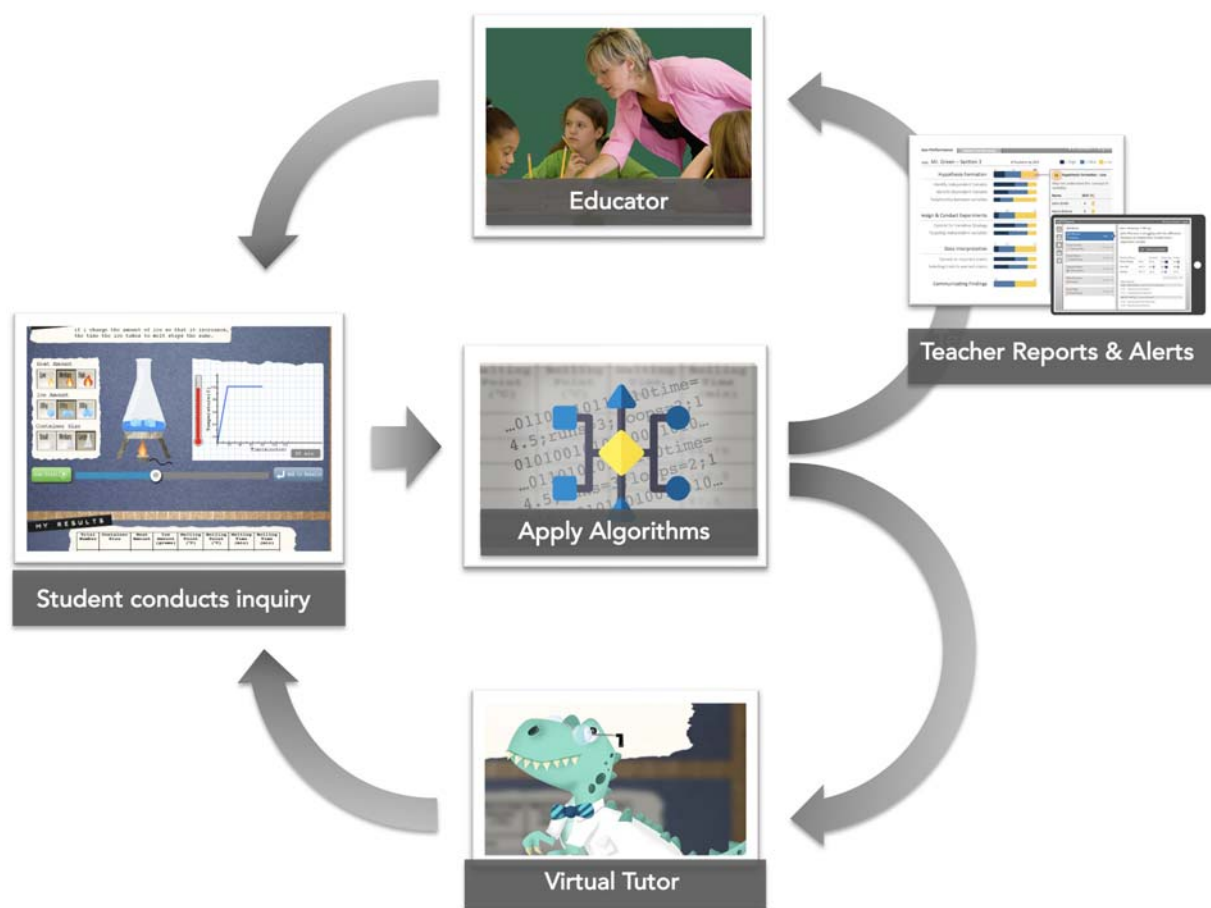


Fig. 1 As students conduct inquiry with Inq-ITS, their inquiry is automatically assessed by underlying algorithms. Rex reacts in real time to support students' learning in the moment. Real time reports and alerts are also generated in real time to support teachers' assessment and instructional needs.

Having a logging infrastructure that collects all students' interactions provides many benefits for intelligent tutoring systems. In terms of *research*, the system can effortlessly and accurately monitor and record every student action; this is important for testing the efficacy of scaffolds in an ITS. Regarding *assessment*, by seamlessly doing assessment in real time and at a fine-grained level, it can assess and remediate students' specific inquiry difficulties in the rich contexts in which they are learned (Mislevy et al., 2003). In terms of *instruction*, it can seamlessly integrate assessment with instruction so additional class time is not needed; that is, the tutor Rex can respond to individual students in real-time as they conduct inquiry, when it is most beneficial for learning; and it can provide formative data and actionable alerts to teachers about his/her students easily and automatically via our dashboard Inq-Blotter (Dickler et al., 2021; Gobert, 2019; Gobert et al., 2019; Gobert and Sao Pedro, 2016; Sao Pedro, 2015, 2016). With regard to *scalability*, because these materials are web-based, virtually any teacher with access to the web can use the ITS.

Furthermore, there are great advantages for analyzing rich logfiles of students' learning using machine-learning, as is done in Inq-ITS. These are often referred to as the 4Vs (Laney, 2001), namely: the large volume of log data, the velocity at which it is generated, the veracity in fine-grained logfiles, and the variability in students' logfiles since there are many ways in which inquiry can be done both productively and unproductively. However, theory is needed to both distill and aggregate data (Quellmalz et al., 2012), and to design categories a priori so that results are pedagogically meaningful (Gobert et al., 2013). In Inq-ITS, algorithms are used to derive rules based on cognitive task analysis (Ericsson and Simon, 1980) of student interaction data and/or human-classified labels of those data that were developed from the national inquiry frameworks (NRC, 1996; NGSS, 2013) as well as data (Gobert and Sao Pedro, 2017) and literature on students' difficulties with conducting inquiry (cf. Kuhn, 2005). Then, machine learning is leveraged to derive algorithms for process-oriented practices that can be exhibited in several different ways, e.g., collecting data (Sao Pedro et al., 2013a,b). The results are aggregated into performance indicators that provide evidence of students' proficiencies for each inquiry practice and their respective sub-skills, and the system is instrumented to jump in to provide real time feedback to students when the system detects that students are struggling on the inquiry competencies of interest.

Inq-ITS delivers scaffolds to students in text format via the pedagogical agent, Rex (see Fig. 1). This proactive scaffolding approach helps to support students in their inquiry processes (Schauble, 1990; deJong, 2006) by preventing students from engaging in unproductive behaviors (Buckley et al., 2006; Sao Pedro, 2013). This proactive approach is also important because students may not be aware that they need help (Aleven and Koedinger, 2000; Aleven et al., 2004).

Scaffolds in Inq-ITS

In Inq-ITS, the primary scaffolding strategies used by the pedagogical agent are explanations provided by Rex (see Fig. 1). Such scaffolding is based on principles from cognitive psychology (McKendree, 1990), and this approach to scaffolding has been shown to benefit learning in other domains (Beck and Mostow, 2008). The scaffolds involve multiple steps or levels (cf., Wood and Wood, 1999), similar to Cognitive Tutors (Anderson et al., 1995). In Inq-ITS, scaffolds are triggered by our assessment algorithms (cf., Gobert et al., 2013, 2016; Sao Pedro, 2013) when the systems detects students' difficulties with inquiry. An example set of the progression of scaffolds that is triggered when a student is not designing controlled experiments, a key skill in science inquiry (Kuhn, 2005), is presented below.

1. Orienting scaffolds are designed to orient the student to the current task within the inquiry cycle since students have difficulties monitoring their progress (de Jong, 2006). For this, Rex may state that in the experiment phase, data should be collected to help support or refute a stated hypothesis.
2. Conceptual scaffolds are designed to provide a high-level hint to the student, e.g., if a student has collected confounded trials (Sao Pedro, 2013), Rex may ask, "How will you know about the effects of x on y, given your trials?"
3. Procedural scaffolds are designed to provide information about the procedure to use on the current task, e.g., Rex may instruct the student to "construct a controlled trial, *relative* to your last trial run." Explicitly teaching this strategy has been shown to be effective (Klahr and Nigam, 2004; Sao Pedro et al., 2009, 2010).
4. Instrumental scaffolds are designed to tell the student exactly what to do on the current task, i.e., a "bottom-out hint" (cf. Aleven and Koedinger, 2000). For example, Rex may instruct the student to "conduct a set of trials in which you change only one variable while keeping the others the same."

Using scaffolds like these, Gobert and her team have demonstrated the effectiveness of our assessment and scaffolding approach at the middle school level for several NGSS practices, including forming questions, planning and carrying out investigations, analyzing and interpreting data, and warranting claims in which scaffolding helped students both acquire and transfer these practices to new Inq-ITS topics over long periods of time, tested up to 170 days (Li et al., 2018, 2019a,b; Gobert et al., 2018; Moussavi et al., 2016; Moussavi, 2018; Sao Pedro et al., 2013a,c).

Conclusions

In summary, the last half century has seen dramatic changes in ITS developments and innovations, enabling ITSs with new capacities, which greatly impacts learning. Important among them is fine-grained student modeling. This, in turn, allows for fine-grained assessment, diagnosis, and personalized tutoring.

This entry also highlights how another innovation, namely machine learning, can be used in an ITS, specifically Inq-ITS. Machine learning in Inq-ITS is used to elicit students' inquiry performances while they engage in inquiry, assess their competencies in real time, and trigger scaffolding to them in real time in order to help them learn the science competencies expected by national frameworks (NGSS, 2013). Furthermore, well-designed, targeted scaffolds that are based on fine-grained, rigorous assessment of students' competencies *can* lead to robust student learning even for ill-defined domains such as inquiry. In sum, this approach to assessment and scaffolding can and should inform the design of assessment and intelligent tutoring systems so that systems are instrumented to assess and support *all* students, regardless of zip code, ELL status, etc. Furthermore, it is imperative that technological resources be leveraged so that ITSs are scalable to provide learning benefits far beyond what an individual teacher can provide.

Acknowledgments

This research was funded by the National Science Foundation (NSF-IIS-1629045, NSF-DRL-1252477, NSF-IIS-1902647) and the US Department of Education (R305A210432, R305A120778, ED-IES-15-C-0018, ED-IES-16-C-0014). Any opinions, findings, etc., expressed are those of the authors and do not necessarily reflect those of the funding agencies.

References

- Ahuja, N.J., Sille, R., 2013. A critical review of development of intelligent tutoring systems: retrospect, present and prospect. *Int. J. Comput. Sci. Issues* 10 (4), 39.
- Aleven, V., Koedinger, K., 2000. Limitations of student control: do students know when they need help? In: Gauthier, G., Frasson, C., VanLehn, K. (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems, ITS 2000*. Springer-Verlag, Berlin, pp. 292–303.
- Aleven, V., McLaren, B., Roll, I., Koedinger, K., 2004. Toward tutoring help seeking. In: Lester, J.C., Vicari, R.M., Paraguaçu, F. (Eds.), *Proceedings of the Seventh International Conference on Intelligent Tutoring Systems, ITS 2004*. Springer-Verlag, Berlin, pp. 227–239.
- Alkhatlan, A., Kalita, J., 2018. Intelligent Tutoring Systems: a Comprehensive Historical Survey With Recent Developments arXiv preprint arXiv:1812.09628.
- Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R., 1995. Cognitive tutors: lessons learned. *J. Learn. Sci.* 4 (2), 167–207.
- Andres, J.M., Hutt, S., Ocupaugh, J., Baker, R.S., Nasir, N., Porter, C., 2021. How anxiety affects affect: a quantitative ethnographic investigation using affect detectors and data-targeted interviews. In: *International Conference on Quantitative Ethnography*. Springer, Cham, pp. 268–283.
- Baker, R.S., Yacef, K., 2009. The state of educational data mining in 2009: a review and future visions. *J. Educ. Data Min.* 1 (1), 3–17.
- Baker, R.S., Corbett, A.T., Wagner, A.Z., 2006. Human classification of low-fidelity replays of student actions. In: *Proceedings of the Educational Data Mining Workshop at the 8th International Conference on Intelligent Tutoring Systems*, vol. 2002, pp. 29–36.
- Baker, R.S.J., Gowda, S.M., Wixon, M., Kalka, J., Wagner, A.Z., Aleven, V., et al., 2012. Towards sensor-free affect detection in cognitive tutor algebra. In: Yacef, K., Zaiane, O., Hershkowitz, H., Yudelson, M., Stamper, J. (Eds.), *Proceedings of the 5th International Conference on Educational Data Mining*, p. 126e133. Chania, Greece.
- Baker, R.S., Nasir, N., Ocupaugh, J.L., Hutt, S., Andres, J.M., Slater, S., Schofield, M., Moore, A., Paquette, L., Munshi, A., Biswas, G., 2021. Affect-targeted interviews for understanding student frustration. In: *International Conference on Artificial Intelligence in Education*. Springer, Cham, pp. 52–63.
- Baker, R.S., 2007. Modeling and understanding students' off-task behavior in intelligent tutoring systems. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1059–1068.
- Beck, J.E., Mostow, J., 2008. How who should practice: using learning decomposition to evaluate the efficacy of different types of practice for different types of students. In: *Intelligent Tutoring Systems*. Springer Berlin Heidelberg, pp. 353–362.
- Bloom, B.S., 1984. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educ. Res.* 13 (6), 4–16.
- Brown, J.S., Burton, R.R., 1978. Diagnostic models for procedural bugs in basic mathematical skills. *Cognit. Sci.* 2 (2), 155–192.
- Brown, J.S., Collins, A., Duguid, P., 1989. Situated cognition and the culture of learning. *Educ. Res.* 18 (1), 32–42.
- Buckley, B., Gobert, J., Horwitz, P., 2006. Using log files to track students' model-based inquiry. In: *The Proceedings of the Seventh International Conference of the Learning Sciences (ICLS)*. Erlbaum, Mahwah, NJ, pp. 57–63.
- Burstein, J., 2009. Opportunities for natural language processing research in education. In: *International Conference on Intelligent Text Processing and Computational Linguistics*. Springer, Berlin, Heidelberg, pp. 6–27.
- Burton, R.R., 1976. *Semantic Grammar: An Engineering Technique for Constructing Natural Language Understanding Systems*. BBN Report No. 3453. Bolt Beranek & Newman, Cambridge, MA.
- Carbonell, J.R., 1970. AI in CAI: an artificial-intelligence approach to computer-assisted instruction. *IEEE Trans. Man Mach. Syst.* 11 (4), 190–202.
- Chowdhury, G.G., 2003. Natural language processing. *Annu. Rev. Inf. Sci. Technol.* 37 (1), 51–89.
- Chrysafiadi, K., Virvou, M., 2013. Student modeling approaches: a literature review for the last decade. *Expert Syst. Appl.* 40 (11), 4715–4729.
- Coenen, F., 2011. Data mining: past, present and future. *Knowl. Eng. Rev.* 26 (1), 25–29.
- Corbett, A., Anderson, J., 1995. Knowledge-tracing: modeling the acquisition of procedural knowledge. *User Model. User-Adapted Interact.* 4, 253–278.
- Corbett, A.T., Koedinger, K.R., Anderson, J.R., 1997. Intelligent tutoring systems. In: *Handbook of Human-Computer Interaction*, pp. 849–874. North-Holland.
- Crowder, N.A., 1959. Automatic tutoring by means of intrinsic programming. In: *Automatic Teaching: The State of the Art*. Wiley, New York, pp. 109–116.
- de Jong, T., 2006. Computer simulations—technological advances in inquiry learning. *Science* 312, 532–533.
- Dickler, R., Gobert, J., Sao Pedro, M., 2021. Using innovative methods to explore the potential of an alerting dashboard for science inquiry. *J. Learn. Anal.* 8 (2), 105–122. <https://doi.org/10.18608/jla.2021.7153>.
- D'Mello, S., Graesser, A., 2009. Automatic detection of learner's affect from gross body language. *Appl. Artif. Intell.* 23 (2), 123–150.
- Edwards, S.J., 1968. Development of a Self-Instruction Program on the Basic Principles of Fitting Clothing. Doctoral Dissertation. Texas Tech University.
- Ericsson, K.A., Simon, H., 1980. Verbal reports as data. *Psychol. Rev.* 87, 215–251.
- Feng, M., Beck, J., Heffernan, N., 2009. Using learning decomposition and bootstrapping with randomization to compare the impact of difference educational interventions on learning. In: Barnes, T., Desmarais, M., Romero, C., Ventura, S. (Eds.), *Proceedings of the 2nd International Conference on Educational Data Mining, EDM 2009*, pp. 51–60. Cordoba, Spain.
- Gertner, A.S., VanLehn, K., 2000. Andes: a coached problem solving environment for physics. In: *International Conference on Intelligent Tutoring Systems*. Springer, Berlin, Heidelberg, pp. 133–142.
- Gobert, J., Buckley, B., 2000. Special issue editorial: introduction to model-based teaching and learning. *Int. J. Sci. Educ.* 22 (9), 891–894.

- Gobert, J., Clement, J.J., 1999. Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *J. Res. Sci. Teach.* 36 (1), 39–53.
- Gobert, J., Sao Pedro, M., 2016. A Real Time Alerting Tool to Transform Teachers' Assessment of Science Inquiry Practices, (NSF-IIS-1629045). Awarded by the. National Science Foundation.
- Gobert, J.D., Sao Pedro, M.A., 2017. Digital assessment environments for scientific inquiry practices. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*, pp. 508–534. West Sussex, UK.
- Gobert, J., Sao Pedro, M., Baker, R., Toto, E., Montalvo, O., 2012. Leveraging educational data mining for real time performance assessment of scientific inquiry skills within microworlds. *J. Educ. Data Min.* 4 (1), 111–143.
- Gobert, J.D., Sao Pedro, M., Raziuddin, J., Baker, R.S., 2013. From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *J. Learn. Sci.* 22 (4), 521–563.
- Gobert, J.D., Baker, R.S., Wixon, M.B., 2015. Operationalizing and detecting disengagement within online science microworlds. *Educ. Psychol.* 50 (1), 43–57.
- Gobert, J.D., Baker, R.S., Sao Pedro, M.A., 2016. Inquiry Skills Tutoring System. US Patent no. 9,373,082 (issued).
- Gobert, J., Sao Pedro, M., Betts, C., Baker, R.S., 2017. Inquiry Skills Tutoring System (Child Patent for Alerting System). US Patent no. 9,564,057 (issued).
- Gobert, J.D., Moussavi, R., Li, H., Sao Pedro, M., Dickler, R., 2018. Real-time scaffolding of students' online data interpretation during inquiry with Inq-ITS using educational data mining. In: *Cyber-physical Laboratories in Engineering and Science Education*. Springer, Cham, pp. 191–217.
- Gobert, J., Sao Pedro, M., Betts, C., Baker, R.S., 2019. Inquiry Skills Tutoring System (Child Patent for Additional Claims to Inq-ITS). US Patent no. 10,186,168 (issued).
- Gobert, J., 2019. Inq-Blotter: Designing Supports for Teachers' Real Time Instruction (NSF-IIS- 1902647). Awarded by the. National Science Foundation.
- Gobert, J., 2022. Engagement Alone Does Not Lead to Deep Science Learning. Inq-ITS. <https://www.inqits.com/post/engagement-alone-does-not-lead-to-deep-science-learning>.
- Henderson, N., Min, W., Rowe, J., Lester, J., 2021. Enhancing multimodal affect recognition with multi-task affective dynamics modeling. In: 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, pp. 1–8.
- Hmelo-Silver, C.E., 2004. Problem-based learning: what and how do students learn? *Educ. Psychol. Rev.* 16, 235–266.
- Johnson, W.L., Soloway, E., 1985. PROUST: knowledge-based program understanding. *IEEE Trans. Software Eng.* (3), 267–275.
- Johnson-Laird, P., 1983. *Mental Models: Toward a Cognitive Science of Language, Inference and Consciousness*. Harvard University Press.
- Kirschner, P.A., Sweller, J., Clark, R.E., 2006. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41 (2), 75–86.
- Klahr, D., Nigam, M., 2004. The equivalence of learning paths in early science instruction effects of direct instruction and discovery learning. *Psychol. Sci.* 15 (10), 661–667.
- Koedinger, K., MacLaren, B., 2002. Developing a Pedagogical Domain Theory of Early Algebra Problem Solving. CMU-HCIL Tech Report, Pittsburgh, PA, pp. 02–100.
- Kolodner, J.L., 1983. Reconstructive memory: a computer model. *Cognit. Sci.* 7 (4), 281–328.
- Kolodner, J.L., 1992. An introduction to case-based reasoning. *Artif. Intell. Rev.* 6 (1), 3–34.
- Kuhn, D., 2005. *Education for Thinking*. Harvard University Press, Cambridge, MA.
- Kulik, J.A., Fletcher, J.D., 2016. Effectiveness of intelligent tutoring systems: a meta-analytic review. *Rev. Educ. Res.* 86 (1), 42–78.
- Laney, D., 2001. 3D data management: controlling data volume, velocity and variety. *META Group Res. Note* 6, 70–73.
- Lester, J.C., Boyer, K.E., Wiebe, E.N., Mott, B., Smith, A., 2019. Board 144: PRIME: engaging STEM undergraduates in computer science with intelligent tutoring systems. In: 2019 ASEE Annual Conference & Exposition.
- Li, H., Gobert, J., Dickler, R., Moussavi, R., 2018. The impact of multiple real-time scaffolding experiences on science inquiry practices. In: *Lecture Notes in Computer Science*. Springer, pp. 99–109.
- Li, H., Gobert, J., Dickler, R., 2019a. Scaffolding during science inquiry. In: Mitchell, J.C., Porayska-Pomsta, K., Joyner, D. (Eds.), *Proceedings of the Sixth Annual ACM Conference on Learning at Scale*. ACM, New York, NY, pp. 1–10.
- Li, H., Gobert, J., Dickler, R., 2019b. Testing the robustness of inquiry practices once scaffolding is removed. In: Coy, A., Hayashi, Y., Chang, M. (Eds.), *Proceedings of the Fifteenth International Conference on Intelligent Tutoring Systems*. Springer, Cham, Switzerland, pp. 204–213.
- Linn, M.C., Hsi, S., 2000. *Computers, Teachers, Peers: Science Learning Partners*. Erlbaum Associates, Mahwah, NJ.
- Luan, H., Peter, G., Lai, H., Gobert, J., Yang Stephen, J.H., Ogata, H., Baltés, J., Rodrigo, G., Li, P., Chin-Chung, T., 2020. Challenges and future directions of big data and artificial intelligence in education. *Front. Psychol.* 11, 1–11.
- Ma, W., Adesope, O.O., Nesbit, J.C., Liu, Q., 2014. Intelligent tutoring systems and learning outcomes: a meta-analysis. *J. Educ. Psychol.* 106 (4), 901.
- McKendree, J., 1990. Effective feedback content for tutoring complex skills. *Hum. Comput. Interact.* 5, 381–413.
- Merrill, D.C., Reiser, B.J., Ranney, M., Trafton, J.G., 1992. Effective tutoring techniques: a comparison of human tutors and intelligent tutoring systems. *J. Learn. Sci.* 2 (3), 277–305.
- Mislevy, R., Chudowsky, N., Draney, K., Fried, R., Gaffney, T., Haertel, G., 2003. *Design Patterns for Assessing Science Inquiry*. SRI International, Menlo Park, CA.
- Mislevy, R.J., Riconscente, M.M., Rutstein, D.W., 2009. *Design Patterns for Assessing Model-Based Reasoning (Large-Scale Assessment Technical Report 6)*. SRI International, Menlo Park, CA. Downloaded May 31, 2011, from: http://ecd.sri.com/downloads/ECD_TR6_Model-Based_Reasoning.pdf.
- Mislevy, R., Yan, D., Gobert, J., Sao Pedro, M., 2020. Automated scoring with intelligent tutoring systems. In: Yan, D., Rupp, A., Foltz, P. (Eds.), *Handbook of Automated Scoring: Theory Into Practice*. Chapman and Hall, London, UK.
- Mott, M., Mott, B., Rowe, J., Ozer, E., Giovanelli, A., Berna, M., et al., 2021. "What's important to you, max?": the influence of goals on engagement in an interactive narrative for adolescent health behavior change. In: *International Conference on Interactive Digital Storytelling*. Springer, Cham, pp. 379–392.
- Moussavi, R., Gobert, J., Sao Pedro, M., 2016. The effect of scaffolding on the immediate transfer of students' data interpretation skills within science topics. In: *Proceedings of the 12th International Conference of the Learning Sciences*. Singapore, pp. 1002–1006.
- Moussavi, R., 2018. *Design, Development, and Evaluation of Scaffolds for Data Interpretation Practices During Inquiry*. Doctoral Dissertation. Worcester Polytechnic Institute, Worcester, MA.
- National Research Council, 1996. *National Science Education Standards*. National Academies Press.
- NGSS Lead States, 2013. *Next Generation Science Standards: For States, by States*. The National Academies Press, D.C.: National Academies Press, Washington, DC.
- Nwana, H.S., 1990. Intelligent tutoring systems: an overview. *Artif. Intell. Rev.* 4 (4), 251–277.
- Olsen, J.K., Belenky, D.M., Aleven, V., Rummel, N., 2014. Using an intelligent tutoring system to support collaborative as well as individual learning. In: *International Conference on Intelligent Tutoring Systems*. Springer, Cham, pp. 134–143.
- Papert, S., 1980. Computer-based microworlds as incubators for powerful ideas. In: Taylor, R. (Ed.), *The Computer in the School: Tutor, Tool, Tutee*. Teacher's College Press, New York, NY, pp. 203–201.
- Pellegrino, J., Chudowsky, N., Glaser, R., 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academy Press, Washington, DC.
- Perez-Valle, A., Sagasti, D., 2012. A novel approach for tourism and education through virtual Vitoria-Gasteiz in the 16th century. In: 2012 18th International Conference on Virtual Systems and Multimedia. IEEE, pp. 615–618.
- Picard, R.W., 1997. *Affective Computing*. MIT Press, Boston, MA.
- Quellmalz, E.S., Timins, M.J., Silbergliitt, M.D., Buckley, B.C., 2012. Science assessments for all: integrating science simulations into balanced state science assessment systems. *J. Res. Sci. Teach.* 49, 363–393.
- Rai, D., Beck, J.E., 2012. Math learning environment with game-like elements: an incremental approach for enhancing student engagement and learning effectiveness. In: *International Conference on Intelligent Tutoring Systems*. Springer, Berlin, Heidelberg, pp. 90–100.

- Rickel, J.W., 1989. Intelligent computer-aided instruction: a survey organized around system components. *IEEE Trans. Syst. Man Cybern.* 19 (1), 40–57.
- Sao Pedro, M., Gobert, J., Heffernan, N., Beck, J., 2009. Comparing pedagogical approaches for teaching the control of variables strategy. In: Taatgen, N.A., vanRijn, H. (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*. Cognitive Science Society, Austin, TX, pp. 1294–1299.
- Sao Pedro, M.A., Gobert, J.D., Raziuddin, J., 2010. Comparing pedagogical approaches for the acquisition and long-term robustness of the control of variables strategy. In: Gomez, K., Lyons, L., Radinsky, J. (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (CLS 2010)—Volume 1, Full Papers*. International Society of the Learning Sciences, Chicago, IL, pp. 1024–1031.
- Sao Pedro, M.A., 2015. *Inq-Blotter: Revolutionizing How Teachers Identify and Support Students Needing Help During Inquiry*. Awarded by US DoE Institute of Education Sciences (ED-IES-15-C-0018).
- Sao Pedro, M.A., 2016. *Inq-Blotter: Recognizing How Teachers Identify and Support Students Needing Help During Inquiry*. Awarded by US DoE Institute of Education Sciences (ED-IES-16-C-0014).
- Sao Pedro, M., Baker, R., Gobert, J., 2013a. Incorporating scaffolding and tutor context into bayesian knowledge tracing to predict inquiry skill acquisition. In: *Proceedings of the 6th International Conference on Educational Data Mining*. Memphis, TN, pp. 185–192.
- Sao Pedro, M., Baker, R., Gobert, J., 2013b. What different kinds of stratification can reveal about the generalizability of data-mined skill assessment models. In: *Proceedings of the 3rd Conference on Learning Analytics and Knowledge*, pp. 190–194. Leuven, Belgium.
- Sao Pedro, M., Baker, R., Gobert, J., Montalvo, O., Nakama, A., 2013c. Leveraging machine-learned detectors of systematic inquiry behavior to estimate and predict transfer of inquiry skill. *User Model. User-Adapted Interact.* 23, 1–39.
- Sao Pedro, M.A., 2013. *Real-Time Assessment, Prediction, and Scaffolding of Middle School Students' Data Collection Skills Within Physical Science Simulations*. In: *Social Science and Policy Studies: Learning Sciences and Technologies Program Ph.D. Dissertation*. Worcester Polytechnic Institute. Technical Report etd-042513-062949.
- Schauble, L., 1990. Belief revision in children: the role of prior knowledge and strategies for generating evidence. *J. Exp. Child Psychol.* 49, 31–57.
- Self, J.A., 1974. Student models in computer-aided instruction. *Int. J. Man Mach. Stud.* 6 (2), 261–276.
- Shute, V.J., Psotka, J., 1994. *Intelligent Tutoring Systems: Past, Present, and Future*. Report No. AL-TP-94-0005. Air Force Materiel Command, Brooks Air Force Base, TX.
- Shute, V., Glaser, R., Raghavan, K., 1988. *Inference and Discovery in an Exploratory Laboratory*. University of Pittsburgh Learning Research and Development Center.
- Skinner, B.F., 1948. "Superstition" in the pigeon. *J. Exp. Psychol.* 38 (2), 168–172. <https://doi.org/10.1037/h0055873>.
- Sleeman, D., 1985. UMFE: a user modelling front-end subsystem. *Int. J. Man Mach. Stud.* 23 (1), 71–88.
- Soller, A., 2001. Supporting social interaction in an intelligent collaborative learning system. *Int. J. Artif. Intell. Educ.* 12 (1), 40–62.
- Stansfield, J.L., Carr, B.P., Goldstein, I.P., 1976. *Wumpus Advisor 1: A First Implementation Program That Tutors Logical and Probabilistic Reasoning Skills*. AI Lab Memo 381. MIT, Cambridge, Massachusetts.
- VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M., 2005. The Andes physics tutoring system: lessons learned. *Int. J. Artif. Intell. Educ.* 15 (3), 147–204.
- VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* 46 (4), 197–221.
- Walsh, M.B., Moss, C.M., Johnson, B.G., 2002. Quantitative impact of a cognitive modeling intelligent tutoring system on student performance in balancing chemical equations. *Chem. Educ.* 7, 379–383. <https://doi.org/10.1007/s00897020620a>.
- Webb, G.I., Pazzani, M.J., Billsus, D., 2001. Machine learning for user modeling. *User Model. User-Adapted Interact.* 11 (1), 19–29.
- White, B.Y., Frederiksen, J.R., 1990. Causal model progressions as a foundation for intelligent learning environments. *Artif. Intell.* 42 (1), 99–157.
- White, B.Y., 1984. Designing computer games to help physics students understand Newton's laws of motion. *Cognit. Instruct.* 1 (1), 69–108.
- Wood, H., Wood, D., 1999. Help seeking, learning and contingent tutoring. *Comput. Educ.* 33, 153–169.
- Yang, C., Chiang, F.-K., Cheng, Q., Ji, J., 2021. Machine learning-based student modeling methodology for intelligent tutoring systems. *J. Educ. Comput. Res.* 59 (6), 1015–1035. <https://doi.org/10.1177/0735633120986256>.

Simulation- and game-based assessment

Kristen DiCerbo, Khan Academy, Avondale, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	471
Assessment in the digital ocean	471
Process data	472
Simulation-based assessment examples	473
Game-based assessment examples	473
Validity, reliability, and fairness	474
Challenges to the field	475
References	475

Introduction

The primary goal of educational assessment is to infer attributes about a learner from observation of their performance on a given activity. Historically, this has been accomplished by presenting students with a defined set of items, usually framed as questions, to which they respond. The items are then scored based on the correctness of the responses and aggregated into a score meant to tell us about what the learner knows and can do. However, we now live in a digital world where students engage in a variety of rich activities from which we can gather detailed data. This ability to capture and analyze information from students' daily learning activity should fundamentally change how we think about assessment (DiCerbo and Behrens, 2014). It opens up the kinds of activities that can be considered as assessments, the kinds of evidence that might emerge from those activities, the ways we score that evidence, and the way we aggregate it. Ultimately, it could change the kinds of attributes about which we are able to make valid, reliable, and fair inferences.

Simulations and games are two types of digital activity that have the potential to be used for assessment purposes. Educational simulations are representations of systems or elements of systems that students interact with for learning purposes. Even prior to digital education, simulations were used, particularly in medical and business education, to allow learners to practice skills and explore concepts with reduced risk, in this case to actual patients or businesses, respectively. With the advent of the digital world, simulations have expanded both in their potential capabilities and educational usage. Cell simulations in biology and simulations of data packets moving through computer networks, for example, allow learners to see otherwise unobservable processes. A variety of science simulations allow learners to ask "what if" questions and try things that could be dangerous or impossible in the real world. A review of 59 studies examining the impact of simulations on learning indicated that, overall, simulations have a beneficial effect on achievement over instructional practices in which there were no simulations (D'Angelo et al., 2014).

Educational games can be hard to differentiate from simulations. Generally they have a better defined goal state and rewards system (D'Angelo et al., 2014) and more defined rules (Salen and Zimmerman, 2003) than simulations. Games are seen as attractive potential learning tools because they are so engaging. Teachers and parents wish their learners were as engaged with learning material as they are with entertainment games. As with simulations, many teachers employed games prior to the digital era, but computing technology has brought a new batch of games from the early days of *The Oregon Trail*, which is a series of educational computer games originally produced in 1971, to games that provide reinforcement for practice of basic skills to games that encourage deep investigation of concepts. Reviews of games as learning tools consistently show a positive impact of games on learning but not on motivation (e.g., Clark et al., 2016; Wouters et al., 2013).

However, in addition to being good learning tools, simulations and games lend themselves well to assessment. As Gee (2010), a pioneer in games for learning, pointed out "if players finish the game Halo, there is no question that they are experts in playing Halo, we don't need to give them a separate test to find out if they are good." In the early 2000s, this statement prompted many in the assessment community to wonder, whether a game that required the use of target knowledge and skills more aligned with standards than Halo could become an educational assessment of academic skill. In addition, the increase in our ability to store, capture, and extract digital data caused the community to begin to explore using process data, or the record of the interactions of learners with the system, as evidence of the process of problem solving, providing information beyond a final answer's correctness to increase our understanding of what learners know and can do.

Assessment in the digital ocean

Historically, in order to observe and evaluate learner behavior at scale, for example, in order to make judgments for certification of proficiency for professionals, it was necessary to employ fixed-time, forced-response tasks that were distal proxies of day-to-day activities. Specific data needed to be collected in a specific amount of time, from a large group of people, and so efficiently implemented, forced tasks were required.

In contrast, in the educational practice of teachers, the assessment model has been much more fluid: interactions occur repeatedly over time, the context of activity is consistent with current goals, observations and performance are made and characterized formally and informally, and the subsequent activities of instruction are introduced in light of previous performance. In such contexts, assessment becomes a natural ubiquitous and unobtrusive (Behrens et al., 2007) side effect in the natural environment or what Shute calls *stealth assessment* (Shute et al., 2009).

The first paradigm benefits from formalization of rules of evidence, standardization of experience (allowing for comparability), collection of data, and data analytics while suffering from decontextualized tasks. It also tends to emphasize one correct answer and one right path to get there. The second paradigm benefits from contextualized and “authentic” tasks while suffering from informal data collection, rules of evidence and analytics, and difficulty with comparability and communication outside a specific classroom. The emerging universality of digital tasks and contexts provides the opportunity to unify the strengths of each of these approaches. As the digital instrumentation needed for educational assessment increasingly becomes part of our natural educational activity, the need for intrusive assessment practices that conflict with learning activities diminishes.

This vision, described in Shute et al. (2016), does not involve administering assessments more frequently (e.g., each week, each day) but, rather, continually collecting data as students interact with digital environments. It relies on data that are unobtrusively gathered while students are playing/learning in carefully designed environments that then allow inferences to be made about relevant competencies. As the various data streams coalesce, we obtain more evidence about what students know and can do across multiple contexts. The vision of assessment in technology-rich environments involves high-quality, ongoing, unobtrusive assessments that can be aggregated to inform a student’s evolving competency levels (at various grain sizes) and also aggregated across students to inform higher-level decisions (e.g., from student to class to school to district to state to country).

However, even without arriving at this comprehensive vision of ongoing assessment, there are other levels of digital assessment implementation that take advantage of the possibilities the digital world has to offer. DiCerbo and Behrens (2012) layout a categorization scheme for thinking about these possibilities. Level one merely replicates traditional assessment activities in a digital environment. Level two involves new assessment activities that allow new types of performances and new types of information. This includes, for example, complex simulation tasks that require integration of multiple skills and collect data about the process learners use to solve the problem. Level three involves embedding assessment in the learning environment, for example, as a game that is threaded through semester-long courseware that serves as both a learning tool and a means to understand what learners know and can do. Level four is the vision described above of a system that integrates inputs from multiple digital activities over time to build an understanding of what learners know and can do. Even at level two, there is potential progress being made toward not just technology-enhanced assessment, but assessment-enhanced technology.

Process data

One objection to traditional, standardized measures of assessment is that they produce a single, binary data point from each question: correctness or incorrectness of the final answer. This does not help us understand how a learner arrived at a particular response, which can be key in remediation (and also key in identifying whether there are potentially other correct answers). Digital environments allow us to capture many features of a learner’s performance on an activity, for example time, sequence of activity, and counts of actions. For example, the game *SimCityEDU* is designed to assess systems thinking, particularly the ability to understand how actions in a system have multiple, interrelated effects. Players are dropped into a city that is already built and is experiencing problems, where they are asked to diagnose and fix the issues. It turns out that there is significant air pollution in the city, and that coal plants are causing a lot of it, however, coal plants also provide power to the city. By observing a sequence of actions, whether the player installs alternative forms of energy prior to bulldozing the coal plants or not, provides us evidence about whether the learner understands that the coal plants have multiple effects on the city, as opposed to just the single effect of air pollution. Gathering multiple observations of similar choices, allows us to build models of players’ levels of systems thinking (DiCerbo et al., 2015).

This ability to capture log file data opens the possibilities of providing insight into a learner’s process of problem solving, not just their final answer. This insight can then help point to particular areas where a learner’s solution might differ from the “prescribed” solution. Over time, and thousands or tens of thousands of users, we will be able to identify even less common solution paths in digital experiences, often with the help of exploratory data analysis (Behrens et al., 2012; DiCerbo et al., 2015). In the *SimCityEDU* game, over time it became apparent that a small percentage of players “solved” the air pollution problem by bulldozing the entire city. Rather than simply tagging these as “wrong,” it was instructive instead to explore why learners follow a particular path. Interviews with learners revealed that bulldozing things is fun, that’s why players chose it as a solution, and we should not infer anything about the systems thinking ability, positively or negatively, for those who bulldozed the whole city.

In the classroom, conversation around why a learner took a particular path of action can open up a rich dialog about how a learner has framed a problem and selected a solution option. Clearly, this kind of discussion is beyond the means of our digital capabilities; it is the domain of a capable teacher. Ideally, the digital experience provides a gateway or launching point for these conversations that go far beyond discussions of correctness. Teachers have been exhorting students to “show your work” for decades and the digital world can help make that a default for all students, and help uncover patterns in the process. However, the collection and manipulation (both technically and statistically) of process data requires new infrastructure and new ways of thinking (Behrens and DiCerbo, 2014). As we will see, most projects attempting to make use of this new data have not been implemented at scale.

Simulation-based assessment examples

Given that a goal of education is to help learners apply knowledge and skills in the real world, simulations are attractive as assessment tools because they can place students in scenarios that require this application, even if they don't make use of all of the potential of process data. There are a number of simulation-based assessment that ask students to engage with a simulation but then answer traditional question types, which provide the evidence for scoring. For example, the state of Minnesota has released a science assessment item based on water bottle rockets from their end of year summative state science assessment (http://minnesota.pearsonaccessnext.com/resources/item-samplers/science/item-samplers/grade8/Grade8_Science_MCA_ItemSamplerScript.pdf).

Students select what size bottle to use, how much water to put in, and the pressure of the air to be pumped in. They click Run button on the screen and can observe the launch of the rocket and its return to earth. Students then answer a series of questions about the simulation in traditional item types. For example, they are asked to run multiple trials with different amounts of water and then asked a multiple choice question about what happens to the height of the rocket as the volume of water increases (decreases, stays the same, increases then decreases, or decreases then increases). There are different item types available, including one that asks students to drag and drop arrows to indicate the direction of forces, but there is no data captured about students' interactions with the bottle rocket simulator. The items therefore do allow students to interact with more real world examples and carry out experiments, but the evidence available is still limited to the correctness of their response.

In the formative space, in which assessment is to inform instruction rather than evaluate it, *SimScientists* (simscientists.org) has complex simulations for students to interact with, and has students answer questions in more traditional item types, which are then aggregated using Bayesian Networks to estimate probabilities of mastery of science-based standards (Quellmalz et al., 2020). *Inq-ITS* (inqits.com), a commercially available simulation and assessment tool, has created virtual labs that walk students through the process of developing a hypothesis, changing variables, analyzing the results, making a claim, and choosing evidence for that claim, all through multiple choice and other selection based items. They have broken down the scientific inquiry process into sufficiently granular steps that they can provide feedback about the steps students take when designing an experiment without providing a wide open environment that makes it difficult to manage data and score responses. *Inq-ITS* also lets students explain their findings in more open-ended writing and scores the open-ended responses with Natural Language Processing-based Artificial Intelligence algorithms. These scores are then aggregated and reported as measures of proficiency on science inquiry skills like collecting data and explaining findings to teachers and students.

Attempts to use process data as evidence in more open simulations have remained in the realm of research projects. The *PhET Interactive Simulations* are a well-known set of science simulations used widely for learning. Since creating data hooks in their HTML5, there have been a number of research projects attempting to model the process data from them. For example, LaMar et al. (2017) attempted to create detectors of students' strategy in a PhET chemistry simulation on solutes using process data. They were able to use the behaviors of increasing or decreasing solute across trials to infer students' goals and strategy.

There have also been attempts to combine process data with more traditional data. For example, Choi et al. (2019), a research team in partnership with the state of Minnesota, developed a simulation to assess engineering design processes around developing and optimizing solutions. Students were asked to build a submarine to meet certain specifications (e.g., depth it could go, speed, and capabilities), test it, and revise based on results. Information about student actions was gathered, including the number and kinds of parts used, pattern of testing, and responses to failed tests. In addition students were asked forced-choice questions about relevance of parts to solving the missions and causes of failure. This information was combined via measurement models to produce assessments of engineering design skill.

However, using process data as evidence is currently difficult to move into production for a variety of reasons, including: (a) it is labor intensive to identify the potential evidence pieces that will be aggregated, and to write code to extract those evidence pieces from large activity log files, (b) the technical requirements to capture the amount of data required over unstable school internet connections are substantial, and (c) there are significant infrastructure changes needed to store vast amounts of unstructured log file data when existing systems are designed to capture highly structured response data.

Game-based assessment examples

The above-mentioned challenges exist for game-based assessments too, along with the challenge of adding creating an engaging game along with an assessment. There are three potential approaches to game-based assessment: (a) creating a game-based assessment from scratch, (b) modifying an existing game, and (c) using an existing game as-is. However, in each case, those who have built them generally agree it is essential to start the process with a model of the skills and knowledge to be assessed, usually in the form of some kind of taxonomy or learning progression.

Mars Generation One: Argubot Academy is an example of a made-from-scratch game-based assessment. Designed by GlassLab in the mid-2010s, this game place players on the first colony on Mars, where there are many decisions to be made about how to run the colony, and decisions are made through robot battles. The game is meant to assess argumentation, and design began by identifying a learning progression describing how learners move from novice to expert levels in different areas of argumentation, such as constructing arguments and evaluating others' arguments. Players establish claims they want to make, for example, whether the primary source of food should be insect, plant, or mammal. They then move through the environment collecting evidence to support their claims. They build robots by matching claims to evidence. Robot battles consist of having built a robot with strong claim-evidence

pairs and finding weaknesses in the opposing robots' claim and supporting evidence. Pieces of evidence for inferences included: which pieces of evidence were collected, how claims and evidence were linked, and ability to identify weaknesses in opposing robots.

SimCityEDU was an effort by the same group who created *Mars Generation One*. In this case, EA, the game company, allowed use of *SimCity*, an open-ended city building game for the creation of the game-based assessment, so this was an example of modding an existing game. The player controls where to place infrastructure like roads and power plants, landmarks, and public services such as schools, parks, hospitals and fire stations. The player also determines the tax rate, the budget, and social policy. The city is populated by "Sims", simulated persons, who live in the city created by the player. The *SimCityEDU* team sought to use this structure to assess students' systems thinking, or understanding how parts of a system are interrelated and how changes in one part will impact changes in other parts. In order to constrain the game sufficiently to ensure it was possible to observe evidence related to systems thinking, the developers created scenarios where cities were already built and players had to troubleshoot problems in the city. The scenario of the excess air pollution was described above. Creating these scenarios involved significant development work to modify the existing game but it did allow access to a popular game made by a large team of designers, already known to be engaging. In the end, however, there were very few evidence pieces that actually provided information about learners' systems thinking skills.

A final option is to simply add in data hooks and an assessment model to an existing game, but not introduce any new game play. This was done with the popular game *Plants vs. Zombies 2* (Shute et al., 2015). In *Plants vs. Zombies 2*, players must plant a variety of special plants on their lawn to prevent zombies from reaching their house. Each of these plants has different attributes. For example, some plants (offensive ones) attack zombies directly, while other plants (defensive ones) slow down zombies to give the player more time to attack the zombies. A few plants generate "sun," an in-game resource needed to purchase more plants. The challenge of the game comes from determining which plants to use and where to place them in order to defeat all zombies in each level of the game. The authors identified four facets of problem solving: (a) analyzing givens and constraints, (b) planning a solution pathway, (c) using tools and resources effectively, and (d) monitoring and evaluating progress and then identified actions in the game that aligned to these facets. These actions were aggregated to create problem solving scores for players.

One of the most researched game-based assessments is *Physics Playground* (originally called Newton's playground; Shute et al., 2013). Based on the puzzle video game *Crayon Physics*, the game requires the player to guide a green ball to a red balloon by drawing/creating simple machines (which are called agents of force and motion in the game) with the mouse. The objects come to life once the object is drawn. Everything obeys the basic rules of physics relating to gravity and Newton's three laws of motion. There are more than 70 levels to the game and players win trophies for completing them. Initial studies used these trophies as evidence of proficiency while later studies developed models with multiple in-game actions providing evidence about players understanding of qualitative physics.

Through these descriptions, it should be clear that there is a tension between the openness of game play and the need to observe specific behaviors in assessment. Some of the attractiveness of games comes from player choice and agency, and that can be diminished to the point of losing the fun when the design forces players into paths to observe certain behaviors. However, without this forcing function, the assessment data are often lacking and there is not enough evidence to make reliable estimates. These challenges face game-based assessment on top of all of the technical challenges described for simulations.

Validity, reliability, and fairness

To quote Messick (1994), "such basic assessment issues as validity, reliability, comparability, and fairness need to be uniformly addressed for all assessments because they are not just measurement principles, they are social values that have meaning and force outside of measurement wherever evaluative judgments and decisions are made" (p. 13). These concerns do not go away in the new digital assessment world. The validity and reliability of the assessment data affect the accuracy of the student diagnosis, and the diagnosis informs instructional support. If the first part of the chain is weak, the rest (i.e., diagnostic accuracy and effective instructional support) would consequently be compromised (Shute and Zapata-Rivera, 2010). However, many of the techniques used for measuring the psychometric properties of assessment were developed in the context of standardized assessment, consisting of discrete items specifically designed to assess a single construct, scored as correct or incorrect. Much of the evidence gathered from simulation- and game-based assessments (e.g., time spent and sequence of events) is not scored as correct/incorrect and often relates to multiple constructs. In addition, there is often variability even in what activity is presented to learners. In game-based assessment, for example, players' previous choices and actions, and the immediate feedback received, will result in differing activity sequences from player to player and from play occasion to play occasion for the same player (Kim and Shute, 2015). As a result of these differences, in some cases the models and techniques for estimating particular types of reliability and validity evidence will need revision to be applicable for new types of activity and data.

A review of validity and reliability evidence for ongoing and stealth assessments can be found in DiCerbo et al. (2017). Regarding reliability, both inter-rater and internal consistency measures can be employed with traditional methods (with some modification allowing for correlation among constructs and errors). However, our ideas of test-retest validity often fall apart because they rely on the construct remaining stable over time. In the new digital environments from which we collect technology, people are often learning as they engage. Similarly, common methods employed in generalizability theory studies cannot be applied because not all students complete all activities and many pieces of evidence are related to multiple constructs. In these cases, more work is

required of the field to go back to the foundations of what Generalizability Theory (Brennan, 1992) addresses in order to develop methods that apply to new conceptualizations of assessment.

Similarly, examining the five types of validity evidence in relation to simulation- and game-based assessment presents a mix of results and demand for new methods. Evidence based on content can be evaluated with known tools and efforts to do so suggest assessments in new digital environments are at particular risk for introducing construct-irrelevant variance, such as game play experience, into estimates of knowledge and skill. Evidence based on response processes and internal structure can be gathered through usual methods. Evidence based on relations to other variables presents a demand for adjustment to traditional rules of thumb. Scores from simulations and games can clearly be correlated to other measures of the same construct. However, it is not clear what constitutes acceptable levels of correlation. If complex performances are measuring constructs at a deeper level than traditional measures, through the inclusion of process data, for example, very high correlations would not be expected between the digital and traditional measures. It is not clear that the existing measure should be the “gold standard” if the more complex performance is closer to the real-world performance we want to make inferences about. However, traditional psychometricians would argue that if correlations are not that high, then the measures are assessing different constructs. In some cases, Cohen’s (1992) definition of a large correlation as 0.50 or above is used as a cutoff for whether a test demonstrates convergent validity. In other cases, a significant, non-zero correlation is viewed as acceptable evidence. The field has not had these discussions as a whole to determine what counts as compelling evidence. Finally, as with traditional assessments, evidence based on consequences of use of simulation- and game-based assessments is sparse and lacking.

Challenges to the field

To be accomplished appropriately, development of simulation- and game-based assessment involves time and money. Here are just a few key tasks required to make a learning game with assessment baked in:

- Understanding how learners progress in the content area. Is there existing research about stages a student goes through when learning a concept? Are there common misconceptions? Research on learning progressions often aligns synergistically to game levels. However, this research must be synthesized and translated into game developers’ language.
- Aligning a core loop of game play to the learning loop. The core loop of game play is the repeating series of activities that drive game play forward. In an educational game, that game play loop should be aligned to the skills to be taught and assessed. For example, in *Mars Generation One: Argubot Academy* a key loop of activity is: identifying evidence, linking evidence to claims, and then identifying problems in an opposing person’s argument. Game play was designed to follow that same loop as players moved through the environment on Mars, interacting with characters to gather evidence, built robots for battles by making claim and evidence pairs, and then engaging in robot battles by putting their claims and evidence against others’.
- Creating evidence models. It is necessary to determine what choices in the game tell us about a players’ knowledge and skill. These choices must be designed into the game so the options available differentiate between people who are proficient in a skill of interest and those who are not. In other words, someone who is proficient is likely to make one choice and someone who is not is more likely to make the other. These choices become pieces of evidence. However, the complexity of game play often means that there may be non-hypothesized activity patterns that also help differentiate skill levels. Use of exploratory data analysis (e.g., Behrens et al., 2012; DiCerbo et al., 2015a,b) and machine learning (e.g., Gobert et al., 2013) or combination of both (e.g., see Rosé et al., 2019) can help uncover these. However, all of this work takes time, expertise, and skill.

It takes many iterations of building and testing to get both the game play and the evidence models right. All of the work requires experts in content, assessment, and game design to work in close collaboration. The amount of time and budget it currently takes to make a quality game does not match up to the budgets and timelines of curriculum development at scale. There have been countless interesting, quality games developed that cover a small portion of curriculum content. These individual games often find homes with passionate teachers who adopt them and work them into their classroom practice. However, this is not deployment at scale. Teams have tried to create minigames (a very short game), built game templates for reuse, or rely on a single game mechanic. However, in order to get to scale, we likely need both new business models for partnerships and new processes and procedures for development. Perhaps in the future artificial technology, for example, can do more of the evidence identification work. In current reality, there is a large gap between what we can build in an Research & Development space and what we can build at scale.

References

- Behrens, J.T., DiCerbo, K.E., 2014. Technological implications for assessment ecosystems: opportunities for digital technology to advance assessment. *Teach. Coll. Rec.* 116, p1–22.
- Behrens, J.T., Frezzo, D.C., Mislevy, R.J., Kroopnick, M., Wise, D., 2007. Structural, functional, and semiotic symmetries in simulation-based games and assessments. In: Baker, E., Dickieson, J., Wulfek, W., O’Neil, H.F. (Eds.), *Assessment of Problem Solving Using Simulations*. Routledge, New York, pp. 59–80.
- Behrens, J.T., DiCerbo, K.E., Yel, N., Levy, R., 2012. Exploratory data analysis. In: Weiner, I.B., Schinka, J.A., Velicer, W.F. (Eds.), *Handbook of Psychology: Research Methods in Psychology*, second ed. Wiley, New York, pp. 34–70.
- Brennan, R.L., 1992. Generalizability theory. *Educ. Meas.: Issues Pract.* 11 (4), 27–34.

- Choi, J., DiCerbo, K., Ventura, M., Lai, E., Wood, J., Iverson, J., 2019. Measuring Proficiency Using Interactive Simulation Data: Empirical Comparison of Evidence Aggregation Methods. National Council on Measurement in Education Annual Meeting, Toronto, Ontario, Canada.
- Clark, D.B., Tanner-Smith, E.E., Killingsworth, S.S., 2016. Digital games, design, and learning: a systematic review and meta-analysis. *Rev. Educ. Res.* 86 (1), p79–122.
- Cohen, J., 1992. A power primer. *Psychol. Bull.* 112 (1), 155.
- D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., Haertel, G., 2014. Simulations for STEM Learning: Systematic Review and Meta-analysis. SRI International, Menlo Park.
- DiCerbo, K.E., Behrens, J.T., 2012. Implications of the digital ocean on current and future assessment. In: Lissitz, R., Jiao, H. (Eds.), *Computers and Their Impact on State Assessment: Recent History and Predictions for the Future*. Information Age Publishing, Charlotte, North Carolina, pp. 273–306.
- DiCerbo, K.E., Behrens, J.T., 2014. *The Impact of the Digital Ocean on Education*. Pearson, London.
- DiCerbo, K.E., Bertling, M., Stephenson, S., Jie, Y., Mislevy, R.J., Bauer, M., Jackson, T., 2015a. The role of exploratory data analysis in the development of game-based assessments. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (Eds.), *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*. Springer, New York, pp. 319–342.
- DiCerbo, K.E., Castellano, K., Jia, Y., Mislevy, R.J., Lin, J., 2015b. Analysis Approaches for Modified and Original Game-Based Assessment. Annual Meeting of the National Council on Measurement in Education, Chicago, IL.
- DiCerbo, K.E., Shute, V., Kim, Y.J., 2017. The future of assessment in technology rich environments: psychometric considerations of ongoing assessment. In: Spector, J.M., Lockee, B., Childress, M. (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*. Springer, New York, NY, pp. 1–21.
- Gee, J.P., 2010. James Paul Gee. Retrieved from: <http://vimeo.com/15732568>.
- Gobert, J.D., Sao Pedro, M., Raziuddin, J., Baker, R.S., 2013. From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *J. Learn. Sci.* 22 (4), 521–563.
- Kim, Y.J., Shute, V.J., 2015. Opportunities and challenges in assessing and supporting creativity in video games. In: Green, G., Kaufman, J. (Eds.), *Research Frontiers in Creativity*. Academic Press, San Diego, CA, pp. 100–121.
- LaMar, M., Baker, R.S., Greiff, S., 2017. Methods for assessing inquiry: machine-learned and theoretical. In: *Design Recommendations for Intelligent Tutoring System*, vol. 5, p. 137. *Assessment Methods*.
- Messick, S., 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educ. Res.* 32 (2), 13–23.
- Quellmalz, E.S., Silbergliitt, M.D., Buckley, B.C., Loveland, M.T., Brenner, D.G., 2020. Simulations for supporting and assessing science literacy. In: *Learning and Performance Assessment: Concepts, Methodologies, Tools, and Applications*. IGI Global, Hershey, PA, pp. 760–799.
- Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R., 2019. Explanatory learner models: why machine learning (alone) is not the answer. *Br. J. Educ. Technol.* 50 (6), 2943–2958.
- Salen, K., Zimmerman, E., 2003. *Rules of Play: Game Design Fundamentals*. MIT Press, Cambridge.
- Shute, V.J., Zapata-Rivera, D., 2010. Intelligent systems. In: Baker, E., Peterson, P., McGaw, B. (Eds.), *Third Edition of the International Encyclopedia of Education*. Elsevier Publishers, Oxford, UK, pp. 75–80.
- Shute, V.J., Ventura, M., Bauer, M.I., Zapata-Rivera, D., 2009. Melding the power of serious games and embedded assessment to monitor and foster learning: flow and grow. In: Ritterfeld, U., Cody, M., Vorderer, P. (Eds.), *Serious Games: Mechanisms and Effects*. Routledge, Taylor and Francis, Mahwah, NJ, pp. 295–321.
- Shute, V.J., Ventura, M., Kim, Y.J., 2013. Assessment and learning of qualitative physics in Newton's playground. *J. Educ. Res.* 106, 423–430.
- Shute, V.J., Moore, G.R., Wang, L., 2015. Measuring Problem Solving Skills in Plants vs. Zombies 2. *Proceedings of the 8th International Conference on Educational Data Mining (EDM 2015)*. International Educational Data Mining Society, Madrid, Spain.
- Shute, V.J., Leighton, J.P., Jang, E.E., Chu, M.W., 2016. Advances in the science of assessment. *Educ. Assess.* 21 (1), 34–59.
- Wouters, P., Van Nimwegen, C., Van Oostendorp, H., Van Der Spek, E.D., 2013. A meta-analysis of the cognitive and motivational effects of serious games. *J. Educ. Psychol.* 105 (2), 249.

Assessment of interpersonal and intrapersonal skills

Patrick C. Kyllonen, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	478
Terminology issues	478
What are the interpersonal and intrapersonal skills?	479
Findings related to interpersonal and intrapersonal skills	480
Interpersonal and intrapersonal skills predict academic outcomes	480
Interpersonal and intrapersonal skills predict non-academic outcomes	481
Unlike the case with cognitive abilities, there are multiple, distinct interpersonal and intrapersonal skills	482
Interpersonal and intrapersonal skills appear to change over time but not in exactly the same way cognitive skills do	482
Interpersonal and intrapersonal skills can be improved	483
There are some subgroup differences in interpersonal and intrapersonal skills, but they are smaller than and different from cognitive skill differences	483
Others'-assessments of soft skills are more predictive of outcomes than self-assessments, but the two provide complementary, not duplicative information	484
Noncognitive and cognitive skills are distinct but they overlap and are difficult to disentangle	485
Uses for interpersonal and intrapersonal skills assessments	485
Approaches for assessing interpersonal and intrapersonal skills	486
Ratings	486
Multidimensional forced-choice (ranking)	487
Situational judgment test (SJT)	487
Performance tasks	488
Conclusions	488
Acknowledgment	488
References	488
Further reading	493

Glossary

d A statistic indicating a standardized difference between groups, that is, a mean difference between groups expressed in a standard deviation metric, that is as a proportion of a standard deviation

r A correlation coefficient indicating level of association between two variables on a scale from -1.0 (perfect negative correlation) through 0 (no association at all) to $+1.0$ (perfect positive correlation)

Soft skills Social, emotional, and self-management skills (including traits, attitudes, values, beliefs). Mostly synonymous with interpersonal and intrapersonal skills. Soft indicates a distinction with hard skills, which typically refer to mathematics, language, and technical skills

Interpersonal and Intrapersonal Skills See *Soft Skills*

Multidimensional forced-choice A testing item type in which a respondent is asked to select (or rank) one of two (or more) statements with respect to which statement they agree with more, or prefer, or believe better characterizes them. The statements reflect different dimensions, such as diligence, cooperativeness, curiosity, and sociability

Situational Judgment Test (SJT) An item type in which a situation is presented (as text, audio, or video) and respondents are asked to indicate their response to the situation. The response typically involves either ranking or rating options with respect to how one might best or how one might typically respond to the situation

Likert scale An item type used to measure personality, attitudes, values, beliefs, feelings, and behaviors. A statement is presented followed by a rating scale that asks the respondent to indicate a level of agreement, endorsement, or frequency, on a 5-point scale (i.e., 5 response categories, such as "strongly disagree," "disagree," "neither agree nor disagree," "agree," and "strongly agree"). Variations on the Likert scale have come to include fewer or more response categories. Named after its originator, Rensis Likert, the developer of the method (Likert, 1932)

Subgroup Typically, in the context of testing, a group identified primarily through self-report, based on gender, race, ethnicity, age, country of origin, socio-economic status, or other demographic characteristics

Outcomes, academic Variables such as grades, graduation, disciplinary infractions, time to degree, teacher evaluations, and participation in school activities, that serve as measures of success in school

Performance tests Measures based on individual (or group) performance on a task (e.g., problem solving, decision-making), in contrast to self-ratings or others'- (teacher, guardian) ratings

Big Five Personality Model (or Five Factor Model) A model to characterize the dimensions of personality based on factor analyses of item responses. The model was identified following analyses of personality item responses in many, diverse settings, across diverse sets of items. The model has served as a standardization in terminology to enable meta-analyses and to evaluate proposals for new personality constructs

Meta-analysis A quantitative synthesis of primary studies that include common variables or variables that can be classified that way. The technique is typically used to estimate correlations or treatment effects by aggregating results from primary study samples weighted by size

Factor analysis A statistical technique for analyzing a matrix (e.g., a correlation matrix of test scores or of item responses), to reduce the number of dimensions needed to summarize or characterize the matrix. It is used in individual differences research to identify the main factors needed to account for the pattern of test score or item response correlations

Introduction

Schools teach skills and assess them to monitor student progress and to award scholarships and favorable admissions decisions. But what is a skill? In the context of education and human capital theory, skills, broadly speaking, refer to the knowledge and capabilities students acquire that prepare them for future learning and successful performance in the workplace. Skills have typically been thought of as a kind of knowledge of how to do something—drive a car, solve a puzzle, program a computer. But there is a growing appreciation for a more expansive view of skill that encompasses interpersonal and intrapersonal aspects. Working hard, being curious, and working well with others can also be thought of as skills under this broader, contemporary perspective.

Two important distinctions are often made about skills. One is between general skills, such as literacy and numeracy, applicable to a broad variety of domains, and domain-specific skills, such as horse-race handicapping or network programming, applicable to a narrower set of tasks or circumstances. The other distinction is between *hard* and *soft* skills. Hard skills refer to those recognized and assessed as part of the formal academic curriculum, for example, reading, writing, mathematics, and science. Soft skills refer to those not taught as part of the formal curriculum, but that are acquired as a part of school anyway. These include being prompt and reliable in attendance and homework, listening to the teacher, being curious and acquiring a love of learning, being in control of one's emotions, and being able to communicate effectively with teachers and classmates. These skills have not always been considered skills at all and have historically been downplayed if not overlooked entirely in assessment, particularly large-scale school assessments. For example, state accountability assessments, licensure examinations, and international student comparative surveys such as the Program for International Student Assessment (PISA) (OECD, 2014) focus almost exclusively on cognitive skills, hard skills, with soft skills measured only for the contextual information they provide to facilitate interpretation of hard skills results. This chapter addresses assessment of general, soft skills, specifically, interpersonal and intrapersonal skills.

Terminology issues

Terminology has been a longstanding problem and a source of confusion in this field. On one side—the noun side of the description—*skills*, *abilities*, *dispositions*, *competencies*, *attributes*, *traits*, *factors*, *constructs*, and *dimensions* are used, sometimes interchangeably. Factors and dimensions are substantively neutral terms derived from the mathematical modeling of item responses and scale scores, particularly factor analysis, which has been a dominant approach used in interpersonal and intrapersonal assessment. But as technical terms, they are probably less frequently used than others. Traits is a widely used term in personality psychology, and in psychometrics it reflects mathematical modeling (e.g., item-response theory), but traits is also used in genetics, and therefore is sometimes incorrectly assumed to suggest a permanent, inherited disposition. Disposition also tends to be associated with permanence. Abilities tend to be associated with cognitive factors, such as general, verbal, spatial, reading, and mathematical abilities, and less so with the attributes discussed in this chapter, but references to noncognitive abilities can still be found (Lindqvist and Vestman, 2011). *Attitudes*, *values*, *beliefs* (or *mindsets*), and *behaviors*, are narrower terms that define smaller classes of psychological constructs. *Constructs* itself is a neutral, and broader term used in psychology to label behavioral domains, including but not limited to interpersonal and intrapersonal behavior. Attributes are neutral, but overly broad, for example, attributes can include features such as hair color. Competencies is a popular descriptor. Skill is a term used in psychology (Anderson, 1983) and philosophy (Ryle, 1946) to represent acquired procedural knowledge (“knowledge how”) as distinct from declarative knowledge (“knowledge that”), and as distinct from abilities, which are assumed less malleable. Thus, skill reflects the notion of a malleable psychological attribute.

On the other (adjective) side of the description, terms such as *soft*, *interpersonal* and *intrapersonal*, *21st century*, *personality*, *noncognitive*, *character*, *behavioral*, *transversal*, and other descriptors (including the term *other*) are also used, sometimes interchangeably, as modifiers of skills, traits, abilities, competencies, factors, and the like. *Soft skills* is a term widely used in the workforce, social-emotional and character skills are terms used in K-12 education, transversal is used by UNESCO, psychosocial and interpersonal and intrapersonal are used in higher education, 21st century is a term describing a change in the kinds of tasks people perform due to the change from a manufacturing to an information economy, noncognitive has been used widely in economics (although

with origins in psychology (Messick, 1979)), and personality by a branch of psychology. Although there are nuances in their distinctions, in general those nuances are not universally understood, and thus these terms can mostly be understood as interchangeable even though some constituencies may argue about the nuances.

What are the interpersonal and intrapersonal skills?

There are many, different, and varied interpersonal and intrapersonal skills. Peterson and Seligman (2004) list 24 interpersonal and intrapersonal skills, which they refer to as “character strengths” (e.g., bravery, love, teamwork, forgiveness) organized as 6 “virtues” (wisdom, courage, humanity, justice, temperance, transcendence). The Collaborative for Academic, Social, and Emotional Learning’s (CASEL) Assessment Guide (2018) presents 28 assessments (from ACT TESSERA to ZOO U) measuring 157 different “SEL Competencies,” from “academic motivation,” “behavioral self-control,” “classroom effort,” “decision-making,” and “future orientation” to “vision” and “zest.” The RAND Education Assessment Finder (Schweig et al., 2018) provides an even more exhaustive list of 98 interpersonal and intrapersonal assessments, each of which provides “subscores” (typically 6–10) measuring various “constructs.” The National Research Council (2012) compiled many sources of lists of such constructs. Kyllonen (2016) compiled lists of various noncognitive skills, attitudes, values, and beliefs from various sources including personality psychology, economics, social attitudes, values, cross-cultural psychology, and others. Trilling and Fadel (2009) compiled a list of 184 “21st century skills” from various sources.

This is a confusing and overwhelming set of constructs. This is particularly so because some constructs go by different names and some names stand for different constructs, a long-recognized phenomenon in psychology referred to as jingle-jangle fallacies (Kelly, 1927, p. 64). There have been several approaches to organizing or categorizing such constructs and placing them into a more manageable framework (Blythe and Borowski, 2018; Berg et al., 2017). CASEL (Durlak et al., 2011, 2015) has proposed an organizing heuristic comprising five categories of constructs—self-awareness, social awareness, self-management, relationship skills, and responsible decision making. The basis for the framework is to enable easy categorization of skills themes in intervention programs (CASEL, 2005). Because of CASEL’s influence in the social emotional learning field, this framework has been quite influential. Another influential categorization comes from the Consortium of Chicago Schools research (Farrington et al., 2012), which proposes 5 categories of “noncognitive factors”—academic behavior, academic perseverance, academic mindsets, learning strategies, and social skills, each of which comprises several narrower factors (e.g., study skills, metacognitive strategies, self-regulated learning, time management, and goal setting as narrower learning strategies factors)—on the basis of evidence that such factors are malleable and relate to academic outcomes. Jones et al. (2019) provided a crosswalk between frameworks, based on expert judgments of similarity between factors from various frameworks.

An alternative to the rational frameworks is the Big 5 model from personality psychology. The model proposes five orthogonal factors—Conscientiousness, Extraversion, Agreeableness, Neuroticism, and Openness—retrievable with the use of the mnemonics OCEAN or CANOE. The Big 5 model is based on the lexical hypothesis, that language has evolved to encode the most salient differentiating characteristics of people in the form of words—such as assertive, generous, organized, insecure, and creative (Goldberg, 1993; John et al., 1988; Saucier and Ostendorf, 1999). A factor analysis of self-ratings based on such words results in five dimensions. This result has been replicated across a variety of languages (Saucier et al., 2000). Saucier and Iurino (2020) recently pointed out that the five-factors are not necessarily the best characterization of the structure of language, at least in English, and that as many as 21 dimensions may be justified. Although the more elaborated structures are probably not as robust, particularly cross-culturally (Lui et al., 2020), there are advantages in predictive capacity for a more high-dimensional model (Paunonen and Ashton, 2001).

Although the lexical hypothesis gave rise to the Big 5 model, the same five dimensions have been found to capture the structure of statements as well as words (e.g., “I am a hard worker,” “I enjoy thinking about the origins of the universe,” “I ask a lot of questions in class”) (John and Srivastava, 1999). Cross cultural generalizability also is present for statement-based measures (Heine and Buchtel, 2009). There have been attempts to discover additional dimensions beyond the Big 5. Drasgow et al. (2012) proposed a “lower order trait taxonomy” empirically identified, comprising several facets of the Big 5 factors, for example, achievement, order, self-control, responsibility, non-delinquency, and virtue as facets of conscientiousness. In addition, the taxonomy proposed 6 lower order Openness facets (intellectual efficiency, ingenuity, curiosity, esthetics, tolerance, depth), and three lower order facets for each of agreeableness (consideration, selflessness, cooperation), extraversion (dominance, sociability, attention seeking), and emotional stability (adjustment, even-tempered, optimism); a total of 21 factors (or facets). Similar efforts have arrived at somewhat similar set of factors (Condon, 2018, found 27 factors; Soto and John, 2017 identified 15 factors; Lee and Ashton, 2004, identified 24 factors). Notably, Primi et al. (2021b) identified 18 factors, and provided a crosswalk to the CASEL and Chicago consortium frameworks, the Soto and John (2017) framework, and a framework used by the Office of Economic Cooperation and Development (OECD; Kanakas and Suarez-Alvarez, 2019, Table 1, p. 3). Soto et al. (2022, (see also Soto et al., 2021) suggest a framework (including a 192-item assessment derived from the framework), that unites many of these frameworks.

The value of the Big 5 and to a smaller extent, its extensions into lower order factors (sometimes referred to as *facets*), is that it serves as a standard framework. It has become a standard because it is based not on experts’ rational judgments, but on empirical relationships (correlations) between self- (and other-) evaluations based on words and statements. That is, people who describe themselves as hard-working tend to also describe themselves as able to meet deadlines. But they are equally likely to be curious or incurious. The five factors and associated facets emerge from the analysis of this pattern of predictive relationships.

There are many advantages to standards, such as the Big 5. One is that meta-analyses can be done based on the framework. The Big 5 makes it possible to classify findings from studies, for example prediction studies of Big 5 factors with academic outcomes. Some studies might not use Big 5 terminology (e.g., Robbins et al., 2004; Richardson et al., 2012), but can still be classified as Big 5 studies in meta-analyses because of the known relationships between the factors used in such studies and Big 5 factors. For example, Robbins et al.'s (2004) "achievement motivation," and Richardson et al.'s (2012) "motivational factors" can be treated in a meta-analysis, such as Poropat's (2009), as measures of conscientiousness for the purpose of estimating the strength of relationship between conscientiousness and academic outcomes. The proliferation of factors identified in the beginning of this section, without a standard such as the Big 5, works against such compiling of studies and the resultant accumulation of knowledge. Further research might begin standardizing on the more specific facet dimensions of the Big 5.

A second advantage is that the standard, the Big 5, becomes a way to understand how concepts fit into a larger picture of the ways people differ generally and to address the issue of the jingle-jangle fallacy. Credé et al.'s (2017) analysis of Duckworth's (2016) grit factor as a measure of two facets of conscientiousness (perseverance of effort, which is responsible for most of the academic outcome predictions, and consistency of interest) is a case in point. Another nice illustration of this organizational capability of the Big 5 is a study by John and Mauskopf (2015) (John and De Fruyt, 2015) in which students were asked to rate themselves on the 157 21st century skills from Trilling and Fadel (2009). A factor analysis of those ratings resulted in a structure remarkably close to the Big 5. In such a way, the Big 5 is a useful organizing framework.

One limitation of the Big 5 model is that although it is based on data, it is based exclusively on correlational data, with correlations limited to those from a specific form of assessment, primarily self-ratings of descriptors and statements, with some research from others' ratings. Based strictly on correlational data, a question is whether the factors identified in the Big 5 are important for human (and student) development. We know that the Big 5 factors are useful metrics for describing an individual (self-description or other description), and useful for predicting outcomes, such as academic, workforce, and life outcomes (see next section, **Interpersonal and intrapersonal skills predict academic outcomes**). But are the Big 5 factors useful for diagnosing individuals and organizing and adapting treatments (interventions), for example? Would training targeting Big 5 factors result in individual changes in those Big 5 factors? Recall that the CASEL framework emerged from attempts to categorize treatments. Is there comparable value in using the Big 5 to categorize treatments? Or is there an alternative that similarly to the Big 5 has an empirical basis that might be used to categorize treatments and target interventions?

A study that addressed that issue was conducted by the National Academies of Science, Engineering, and Medicine (NASEM, 2017), which was given a charge "to more clearly define interpersonal and intrapersonal competencies, to examine whether and to what extent a range of these competencies may be related to each other and to persistence and success in undergraduate education (especially in STEM) and to examine the extent to which these competencies can be enhanced through intervention" (p. 2). The study identified 8 competencies—*behaviors related to conscientiousness, sense of belonging, academic self-efficacy, growth mindset, utility goals and values, intrinsic goals and interest, prosocial goals and values, and positive future self*. Several of these could undoubtedly be aligned to Big 5 factors (the first to Conscientiousness; prosocial goals and values to Agreeableness; sense of belonging, growth mindset, and academic self-efficacy to the Neuroticism factor in the Big 5 according to Judge et al., 2003). Unfortunately, although the proposed 8-factor taxonomy may be useful for categorizing research that has been shown to identify "promising competencies that appear to be related to college success" (p. 4), the taxonomy is limited as a guiding framework for organizing soft skills for a couple of reasons. First, all 8 competencies are intrapersonal competencies—the committee identified no interpersonal competencies (e.g., social skills) associated with success in higher education, and therefore leaves out an important aspect of human skill. Second, as the committee pointed out, the criteria for inclusion were strict, leading to the conclusion that "Only limited research has been conducted to date on the potential relationships between various intra- and interpersonal competencies and students' college success. There are major gaps in the research evidence" (p. 4). Thus, although suggestive, and a useful moment-in-time summary, the NASEM (2017) taxonomy is of limited usefulness for characterizing the nature of interpersonal and intrapersonal competencies.

Findings related to interpersonal and intrapersonal skills

Interpersonal and intrapersonal skills predict academic outcomes

Interpersonal and intrapersonal skills are correlated with academic outcomes such as grades in school. This has been demonstrated at all levels of education, primary (elementary), secondary (high school), and tertiary (college). The correlation, based on a meta-analysis (Poropat, 2009), is at least modest, in the $r = 0.20$ s range. The correlation is only slightly lower than the correlation between cognitive ability tests and grades, which Poropat reported as $r = 0.23$. The correlation estimate, particularly for Conscientiousness seems not to depend on the measure used (McAbee and Oswald, 2013). A more recent and more comprehensive meta-analysis (Mammadov, 2022; see also Zell and Lesick, 2021) supports the key findings of Poropat's analysis: Interpersonal and intrapersonal skills, specifically, all the Big 5 personality dimensions, predict school performance across all levels of school. The two highest correlates were conscientiousness (being hard-working, self-disciplined, and responsible), with a correlation of $r = 0.27$ and openness (being curious, creative, and open-minded) with a correlation of $r = 0.16$ (both correlations corrected for measurement error, but not for range restriction). The strong correlation with conscientiousness did not decline with education level, but the correlation of school performance with openness did decline with education level, from elementary to secondary to post-secondary, where it was smallest. Cognitive ability also predicted school performance ($r = 0.42$), but controlling for cognitive ability, there was still a strong relationship between school performance and conscientiousness. The correlation with openness,

controlling for cognitive ability mostly disappears, suggesting that the relationship between openness and school performance is mostly explained by cognitive ability, which is not surprising given that some openness items are merely self-assessments of one's academic abilities. The importance of conscientiousness (and openness), particularly in higher education, has been found in other meta-analyses (O'Conner and Paunonen, 2007), along with the finding that extraversion (being social, outgoing, and socially dominant) was sometimes negatively related to academic success.

Meta-analyses, such as those just reviewed, are useful for estimating the strength of association between inter-intrapersonal skills and school performance, but a weakness is that they combine many different measures in diverse settings, which might distort the relationship between the constructs. Large-scale studies conducted with common measures in relatively homogenous contexts provide useful supplements to the meta-analyses. Such studies support the findings from the meta-analyses. For example, California's Office to Reform Education (CORE) Districts administered assessments of four noncognitive skills—self-management, growth mindset, self-efficacy, and social awareness—to over 400,000 students in grades 3–12 in over 1200 schools in 8 California school districts in a couple of successive cohorts (2014, 2015). All skills measures correlated positively with measures of academic performance (GPA and state test scores) and negatively with suspensions and absences (West et al., 2018a,b).

Similar results were obtained in a study of District of Columbia Public Schools (DCPS) by Kautz et al. (2021). In this study, over 43,000 students were administered a set of interpersonal and intrapersonal (social-emotional learning, or SEL) measures (perseverance, self-management, self-efficacy, social awareness) and school experiences (i.e., climate) (rigorous expectations, sense of belonging, student satisfaction) which were correlated to a set of outcomes (achievement, attendance, suspensions). The highest correlate was self-management, which correlated $r = 0.23$ with Math and ELA achievement, and with number of AP credits ($r = 0.18$) and on-time graduation ($r = 0.16$). SEL measures did not add to the prediction of these outcomes when demographics and prior year academic measures were included in the prediction of current year academic measures.

This relationship between soft skills and academic outcomes is not limited to the United States. Brazilian researchers (Primi et al., 2021a,b) developed a battery from existing measures that factor analyses showed aligned with the Big 5 (while also including an additional locus of control factor; Santos and Primi, 2014, Table 4). The assessment has been administered to over 600,000 Brazilian students (Ricardo Primi, personal communication, December 2021). Prediction studies showed that the Big 5 factors correlated with both math and language achievement in the moderate ranges close to what Poropat's (2009) and Mammadov's (2022) meta-analyses suggested (Primi et al., 2019) (Interestingly, the correlations were substantially higher for reverse-keyed items, that is items, such as "I miss deadlines," in which a negative response, such as "strongly disagree," is treated as a positive indicator of the trait).

Another major study on the relationship between inter-intrapersonal skills and academic outcomes was the Organization for Economic Co-operation and Development's Study on Social and Emotional Skills (Chernyshenko et al., 2018; Kankaras and Suarez-Alvarez, 2019; OECD, 2021), which was conducted with 10 and 15-year-olds in 10 cities around the world. The OECD assessment measured 15 factors, aligned to the Big 5 (3 facets per each of the Big 5 factors). The factors with the strongest prediction of reading, mathematics, and arts grades were persistence and curiosity (for both 10- and 15-year-olds), and negative correlations with stress resistance and sociability (for 15-year-olds), again, findings aligning with the results of the meta-analyses.

Interpersonal and intrapersonal skills predict non-academic outcomes

There have been many studies that have examined the relationship between interpersonal and intrapersonal skills measures and outcomes going beyond academic outcomes. Roberts et al. (2007) showed that personality predicted mortality (/longevity), divorce, and occupational attainment to a degree comparable to the prediction given by cognitive ability and socioeconomic status. The strongest correlate of longevity was conscientiousness (effect size of 0.09), but other Big 5 factors also were associated with longevity, perhaps partly as the authors suggest because conscientiousness is related to engaging in more healthy behaviors. The strongest correlates of avoiding divorce were emotional stability and agreeableness (in addition to a smaller relationship with conscientiousness). The strongest correlates of occupational attainment (i.e., occupational status, typically indicated by average earnings and amount of education required) were four of the Big 5 with an average beta weight of 0.23, comparable to cognitive abilities (beta = 0.27) and socioeconomic status (beta = 0.09) and parental income (beta = 0.14).

Almlund et al. (2011) and Heckman and Kautz (2014) document a wide variety of relationships between personality and labor market outcomes (e.g., job performance, wages, job search effort), along with health and crime and school outcomes (e.g., attainment, attendance, course difficulty, grades, high-school graduation, test scores). Summaries of this line of work are found in OECD (2015, 2021) and Chernyshenko et al. (2018), and the fact of these documented relationships is what motivated OECD's (2021) Survey of Social Emotional Skills. Wilmot and Ones (2019) reviewed 92 meta-analyses (2500 studies involving over a million participants) concerning the relationship between personality and a variety of occupational outcomes, focusing on conscientiousness (see Zell and Lesick, 2021, for a similar quantitative synthesis of over 50 meta-analyses). Conscientiousness is the most predictive interpersonal-intrapersonal skills variable, and Wilmot and Ones showed that this predictiveness extended over 10 clusters (or themes) of occupational variables, themes related to job performance, counterproductive behavior, leadership, organizational commitment, and others. They also found that conscientiousness was most important in certain sectors, notably Health Care, the Military, Sales, and Customer Service occupations.

Unlike the case with cognitive abilities, there are multiple, distinct interpersonal and intrapersonal skills

An important distinction between soft (social-emotional/interpersonal-intrapersonal) and hard (cognitive) skills is in the importance of a general factor. Cognitive skills, such as quantitative and verbal abilities, despite their apparent differences, tend to be highly correlated, a phenomenon noted since Spearman (1904) and Kelley (1927). For example, the program for international student assessment (PISA) finds that latent correlations between tests of mathematics, science, reading, problem solving, and financial literacy range from $r = 0.76$ to $r = 0.91$, with a median of $r = 0.86$ (OECD, 2014, pp. 230–231, Table 12.6, 12.7, 12.8). In fact, two mathematical subscale latent correlations are lower than this median (OECD, 2014, p. 231, Table 12.9). This suggests a strong general factor governing diverse measures of cognitive ability.

In contrast, although some have argued for a general factor in social emotional skills (Dunkel et al., 2021; Musek, 2007; Oltmans et al., 2018), a better characterization is that interpersonal and intrapersonal skills are strongly multidimensional: Being hard-working and intellectually curious are independent (i.e., orthogonal) characteristics of people, for example. The discussion of soft skills dimensionality in the previous sections (e.g., the Big 5 and its extensions) documents this evidence. There are several implications of this finding of multidimensionality for assessment. One is that social-emotional diagnostic modeling is likely to be more fruitful substantively than cognitive diagnostic modeling (CDM; Rupp and Templin, 2008) has proven to be so far. The original premise of cognitive diagnosis was that targeted training could be based on an individual's profile of knowledge and skill at any given moment. However, there is a large literature attesting to the lack of evidence for differentiation available from typical cognitive tests (Haberman, 2008)—total scores are typically more predictive of subscores than like subscores are (Sinharay et al., 2007). This is a cognitive ability phenomenon related to the strength of the general cognitive ability factor on cognitive test performance (combined with the effects of measurement error related to brief measures on prediction). There is not a comparable powerful general factor in soft skills assessment. Therefore, cognitive diagnosis modeling based on independent (not highly correlated) social emotional factors is likely to be a more productive pursuit.

Interpersonal and intrapersonal skills appear to change over time but not in exactly the same way cognitive skills do

On average, test scores, such as mathematics and language achievement scores, or intelligence test scores, rise steadily from early childhood to adulthood where they flatten out until a decline in old age, with some abilities (fluid intelligence) peaking earlier, and others (crystallized knowledge) later (McArdle et al., 2000). This reflects the accumulation of knowledge, skill, and experience over the life span. This can be seen in K-12 education as average achievement scores rise steadily and monotonically from grade to grade (e.g., Ugo and Hill, 2017, Figure 4, p. 9).

As with cognitive abilities, there does seem to be growth in many soft skills dimensions, particularly social dominance, conscientiousness, agreeableness, and emotional stability (Roberts and Mroczek, 2008; Roberts et al., 2006; Srivastava et al., 2003).¹ In that sense there is some similarity with cognitive changes that occur over the lifespan. Life events may also affect these trajectories both for cognitive (Deary, 2014; Nisbett et al., 2012) and noncognitive skills (De Vries et al., 2021; Jackson et al., 2012; Lüdtke et al., 2011).

The picture of development in childhood and adolescence looks quite different. Agreeableness, conscientiousness, and openness tend to decline in late childhood into adolescence, and incline from late adolescence into adulthood (Soto et al., 2011), although there may be some variations by gender in these patterns (Borghuis et al., 2017). The dip is not an artifact of self-reporting: The same decline is found in parent reports (Soto, 2016). Nevertheless, large-scale studies of social-emotional skills find much the same pattern of non-monotonic changes from childhood to late adolescence. Consider, for example, OECD's (2021) study on social emotional skills, which found that on 15 of the 17 scales evaluated, 15-year-olds scored lower than 10-year-olds by up to 0.6 of a standard deviation. The drop in scores was particularly strong in persistence, optimism, trust, curiosity, sociability, energy, and achievement motivation. In only two cases, tolerance and assertiveness, did scores rise, slightly, from age 10 to 15. In the California CORE study (Meyer et al., 2018; West et al., 2018b, 2020), growth mindset rose from grades 4 to 12, with a slight dip in grade 7; social awareness dropped substantially from grade 4 to a nadir in grade 8, rising slightly after that; self-management rose, then fell, then rose again during the grade 4 to grade 12 span; and self-efficacy fell from grade 6 through grade 10, with a slight upturn in grades 11 and 12. The DC public schools study (Kautz et al., 2021) similarly found that scores on all of the social emotional (perseverance, self-management, self-sufficiency, social awareness) and school experience (rigorous expectations, sense of belonging, student satisfaction) scales declined from age 8 (grade 3) to 16 (grade 9–10), gently rising to age 18 (grade 12) (Figure C1; see also Figure 1, p. 8). Interestingly, scales tended to fall and rise in unison rather than some going in one direction and others going in a different direction, which might be expected with multidimensionality. Considering all these studies, whether the findings reflect adolescence "Sturm und Drang" (storm and stress) (Hall, 1916), or are simply an artifact of measurement technique is unknown because all these studies have been conducted exclusively with rating scales. For example, some of the rise and fall of assessment scores based on rating scales during adolescence might be due to response style (e.g., the tendency to agree or disagree regardless of the question).

¹The studies have similar findings, but Srivastava et al. (2003) found that growth in emotional stability occurred only for women; Roberts et al. (2006) found that the social dominance but not sociability facet of extraversion grew over adulthood; Roberts and Mroczek (2008) showed increases in self-confidence, warmth, self-control, and emotional stability.

Interpersonal and intrapersonal skills can be improved

The previous section made the case for the malleability of interpersonal and intrapersonal (soft) skills, primarily due to the passage of time over the lifespan, but a question is whether soft skills improvement can be accelerated, for example through an intervention. In K-12 there is evidence for efficacy of social-emotional learning programs, both in-school (Durlak et al., 2011) and after school (Durlak et al., 2010). These effects appear to persist over time (Taylor et al., 2017). Such programs have been shown not only to affect the social-emotional skills that are the target of the programs, but also to affect behavior and achievement. Fricke et al. (2021) showed that schools themselves (school effects) might contribute to student social-emotional learning gains, particularly the schools in the top or bottom of the effect distribution, meaning that certain schools are good (or poor) at contributing to social-emotional skills development.

For adults, a meta-analysis by Roberts et al. (2017) found evidence for substantial improvement (effect size of $d = 0.37$) in personality measures following an average of a 24-week intervention, which persisted in follow-up data collections. Such intervention effects were particularly strong for anxiety, although other factors were also examined and showed differences. A limitation of this study is that because pre- and post-intervention measures were self-reports, demand characteristics (i.e., situational cues that indicate to the respondent what the experimenter expects or wishes to happen) might have influenced responding.

Interpersonal and intrapersonal outcomes in the studies cited above were typically measured by self- and other- (typically teacher) rating scales. However, the conclusion of the malleability of interpersonal and intrapersonal (noncognitive) skills is not dependent on studies that rely on rating scales. Kautz et al. (2014) review evidence from a variety of types of measures from a broad array of programs ranging from infant programs, early childhood programs, targeted non-cognitive skill programs, adolescent mentorship, and others on a wide range of academic and life outcomes. Jackson (2018) took the approach of computing two kinds of teacher effects for over 10,000 9th grade teachers in North Carolina: a cognitive teacher effect based on test scores (the variable used in typical teacher value added models) and a noncognitive teacher effect, which he referred to as the *behavioral index*, based on the average of a set of proxy variables—absences, suspensions, course grades, and grade repetition (these two sets of variables were correlated $r = 0.56$, with the behavioral index having a much stronger effect on high-school graduation). Jackson showed that the two kinds of teacher effects influenced subsequent outcomes differently: cognitive teacher effects influenced subsequent SAT scores; noncognitive teacher effects (conditional on teacher effects on test scores) influenced high school graduation and other longer run outcomes including subsequent grade-point average. Jackson et al. (2020) also found school effects on not only interpersonal-intrapersonal self-reports but also on school-based arrests, high school completion, and college enrollment and completion. Kautz and Zanoni (2019) similarly showed that a noncognitive composite of such measures was responsible for increased college enrollment due to an intervention program called *OneGoal*. These two studies represent some of the rare instances in which interpersonal-intrapersonal skills are assessed with behavior other than responses to rating scale items.

A question is whether an intervention designed to enhance interpersonal and intrapersonal (soft) skills can affect outcomes other than self-reports in adults. Within education, a major study showed smaller effects from a mindset intervention on college grades (standardized effect of $d = 0.11$, or $b = 0.10$ [i.e., one tenth of a grade point] in the grade point metric) (Yaeger et al., 2019). The National Academy of Sciences (2017) reviewed evidence for the effects of interventions targeting interpersonal and intrapersonal skills on academic persistence and success and identified numerous programs that did so (see their Table B-1). Solanki et al.'s (2020) meta-analysis of social-psychological interventions in college found moderate effect sizes, and that interventions tended to be particularly effective for at-risk student groups. Much of the evidence in these reviews comes from the United States. Leight (2021) reviewed evidence for efficacy of soft skills training on a variety of workforce and life outcomes outside the United States and identified several positive findings.

Organizations routinely conduct soft skills training for employees, often following Kirkpatrick and Kirkpatrick (2006) evaluation model, which successively examines attitudes toward the training, knowledge of the trained material, transfer of training to the workplace, and return on investment. Arthur et al. (2003) showed evidence for organizational soft-skills training efficacy for each of Kirkpatrick's stages. Martin-Raugh et al. (2020) reviewed 39 meta-analyses concerning the development of soft skills primarily in organizations. Intervention approaches varied and included modeling, goal-setting, character education, sports participation, adventure therapy, role-playing, and reward-based programs. They found at least some evidence for improvement of interpersonal skills (social skills, communication skills, leadership, teamwork) and intrapersonal skills (personality, job-related attitudes, self-concept and self-efficacy, emotions, motivation) from training. Effect sizes tended to be in the 0.2–0.5 standard deviation range. Thus, although there is a widespread belief that noncognitive skills or personality attributes are permanent and fixed, in fact there is considerable evidence for the malleability of such skills through interventions, and that such improvements transfer to outcomes other than self-assessments. Nevertheless, as was pointed out in the National Academies report (2017), the evidence for efficacy for many of these programs likely has been weakened by inadequate assessment of intervention effects.

There are some subgroup differences in interpersonal and intrapersonal skills, but they are smaller than and different from cognitive skill differences

Numerous studies have examined mean gender, race/ethnicity, and cultural subgroup differences in noncognitive skills. Smaller subgroup differences are observed on noncognitive skills measures (typically with Cohen's (1988) $d < 0.3$) compared to cognitive skills measures (often $d > 1.0$). For example, in K-12, gaps between the highest and lowest fifths of socio-economic status

distributions are typically over 1 standard deviation in reading and math, but only between one-third and one-half a standard deviation for self-control and approaches to learning (García and Weiss, 2017).

Also, the pattern of differences is more complex than it is with cognitive skills, with some groups scoring higher on some factors, and other groups higher on other factors. In adults including college students, women tend to be higher in neuroticism, extraversion, agreeableness, and conscientiousness, whereas men tend to be higher in assertiveness and openness to ideas (Costa et al., 2001; Schmitt et al., 2008). This pattern tends to hold up across cultures, although there may be slightly larger gender differences in gender equal countries (Mac Giolla and Kajonius, 2019). This pattern of gender differences, cross-cultural similarity, and age-related changes holds up with others' reports as well as self-reports (McCrae et al., 2005). Differences between sexes tend to emerge in adolescence (Brandes et al., 2021), with Neuroticism differences emerging by age 14, and little evidence for differences between cultures in either self-reports or parents' reports (De Bolle et al., 2015; Soto, 2016).

Regarding racial/ethnic subgroup differences, Foides et al. (2008) conducted a meta-analysis of personality scale differences comparing 5 U.S. subgroups, focusing on Big 5 and facet measures. They concluded that in general, race/ethnic group differences were small, with a few exceptions. The few at least moderate differences ($d > 0.2$) in which the lower and upper 90% CV did not contain zero, were anxiety (Black [B] higher than [>] White [W]), self-esteem (Asian [A] & Hispanic [H] > W), order (A > W), agreeableness (W > American Indian [AI]), emotional stability (B > A), extraversion (B > A), and sociability (H > B).

Also important in comparing subgroups is establishing measurement invariance, a statistical procedure to evaluate subgroup effects on item factors, item means, and item covariances, to ensure that scores from the measure can be interpreted in the same way regardless of the subgroup from which the score is obtained (Millsap, 2011). Only after ensuring measurement invariance can mean and correlation differences between groups be accurately interpreted. Although measurement invariance analysis has not been a universal practice for the past few decades, studies are beginning to include measurement invariance analyses as part of reporting of group and cultural differences (e.g., Cunha et al., 2021), and this practice ought to become a routine part of such reporting.

Others'-assessments of soft skills are more predictive of outcomes than self-assessments, but the two provide complementary, not duplicative information

In the immediately preceding section, a finding was that self-reports and reports by others resulted in similar correlational structures (correlational patterns to support the Big 5 model) and similar conclusions about gender differences, cultural differences, or lack of cultural differences, and age-related changes (McCrae et al., 2005). However, there is also evidence that self-reports and others' reports are not the same. In fact, correlations between self-reports and others' reports on the same construct are not necessarily that high. Corrected for measurement error, self-other correlations range from 0.46 for agreeableness and 0.62 for extraversion (Connolly et al., 2007), but that presents a rosy picture due to the correction. Without correcting for measurement error, the correlation tends to be much lower. For example, for ETS's Personal Potential Index (Kyllonen, 2008), a rating form in which multiple faculty members rated graduate school applicants on six scales (Knowledge & Skill, Creativity, Communication Skills, Teamwork, Motivation, Self-Organization and Professionalism & Maturity), McCaffrey et al. (2018) estimated that as many as 9–12 raters (depending on the scale) would be needed to obtain a generalizability coefficient (a measure of reliability) of 0.7. In the CORE study, Hough et al. (2017, Table 5, Table A3) found low correlations between students' perceptions of school culture and climate and parents' and staffs' perceptions: Correlations between culture and climate scales within a group of respondents were substantially higher than correlations of the same scale across respondent groups, although respondent groups tended to increase their similarity as they moved from elementary to middle to high school students. Kautz et al. (2021) found that student and teacher reports were only modestly correlated, for example $r = 0.24$ for perseverance and $r = 0.45$ for rigorous expectations (Figure 5, p. 13). Also, interpersonal and intrapersonal measures had different trajectories from elementary to high school, with student reports showing a dip from 5th to 9th grade and a shallow rise thereafter and teacher reports showing a steady decline from elementary through high school. ETS's Mission Skills Assessment (MSA), administered to over 13,000 middle school students and 1200 teachers, similarly found student-teacher correlations ranging from $r = 0.24$ (Resilience scale) to $r = 0.43$ (Time Management scale) for 6 MSA scales (Petway et al., 2016).

The relatively modest size of self-other correlations and the difference in trends suggests different information carried in student and teacher reports. A question is which one is more accurate, or at least, which one is more predictive of future outcomes? In K-12, Poropat (2014) showed that others' (teachers) ratings were more predictive of student academic outcomes than self-ratings were. Feng et al. (2022) similarly found that teacher reports were most predictive of children's achievement and behavioral outcomes in school and neither children self-reports nor guardian reports added much to the prediction. They suggest this advantage of teacher reports may be due either to teachers' ability to calibrate due to their exposure to many children, to halo effects that incorporate students' achievement, or to teachers' judgments being limited to the school context, which is the context in which the outcomes occur. Lindqvist and Vestman (2011) attributed their relatively high predictions of 20-year-later labor market outcomes to the use of ratings by others (in this case, experts) versus the more typical self-ratings. Segal (2013) showed that teacher ratings of 8th grade students' noncognitive behavior predicted long-term outcomes, including earnings above and beyond the prediction given by test scores, and that noncognitive behavior predicted earnings at all education levels unlike test scores, which predicted earnings only for postsecondary degree holders.

In the adult literature, Oh et al. (2011) found that others' reports of the Big 5 were stronger predictors of job performance than self-reports were, and that others' reports added substantially to the prediction given by self-reports of job performance for all 5 of

the Big 5 scales (the reverse was not true). Prediction benefit increased incrementally with the number of raters (from 1 to 3). Connelly and Ones (2010) similarly found others' reports to be substantially better than self-reports for predicting outcomes, and that there was a benefit of familiarity with the target of rating, but even relatively modest familiarity was useful for predicting others' behavior. Like McCaffrey et al. (2018) they also suggested that numerous raters, perhaps 5–10, were required to reach typical interrater reliability standards. Nevertheless, Klieger et al. (2022) found in examining doctoral STEM programs using standardized ratings from ETS's Personal Potential Index (PPI) that two dimensions—Planning and Organization (a conscientiousness dimension) and Communication Skills positively predicted school grade point average and lower risk of academic probation controlling for undergraduate grade-point average and Graduate Record Examination (GRE) scores. The PPI also showed considerably smaller mean race/ethnicity and gender differences than undergraduate grade-point average and GRE scores.

Letters of recommendation also have been proposed as a means for obtaining information about noncognitive skills of students and applicants by others. Some research has suggested that there is useful information in letters of recommendation, albeit weak (Kuncel et al., 2014). A recent large study, which analyzed 37,000 letters of recommendation (Dalal et al., 2022) found that letters do not predict admissions offers, after GRE scores and undergraduate GPA is considered, and that there was evidence for effects on the content of the letters attributable to gender, race/ethnicity, and program level applied to (masters vs. doctoral). This suggests that the standardized letter approach in Klieger et al. (2022), which found stronger relationships than those found by Kuncel et al. (2014) and Dalal et al. (2022) might be advantageous over the conventional, and much more common unstandardized letter approach.

Noncognitive and cognitive skills are distinct but they overlap and are difficult to disentangle

Many of the discussions in this field concern the relative importance of cognitive and noncognitive skills to academic and workforce success, as well as to other outcomes. This is an important topic because from an educational assessment perspective, it is important to monitor development and growth of skills, both cognitive and noncognitive, to ensure students are on track. Most attention in such monitoring has been on cognitive skills (e.g., state accountability tests, which began in the 2000s, national and international monitoring assessments such as the National Assessment of Educational Progress, which began in the 1980s, and the Program for International Student Assessment, which began in 2000). Only recently has there been comparable attention given to noncognitive skills development from a large-scale assessment standpoint (OECD, 2021). The importance of cognitive and noncognitive skills also is an important topic from a policy perspective as it may inform where educational investments might best be placed.

There have been attempts to tease apart the effects of cognitive and noncognitive skill on education, labor market, and life outcomes (Bowles et al., 2001; Duckworth et al., 2012), but a problem is that test scores are not pure, stable measures of constructs. A test score can reflect cognitive ability if all respondents are equally and highly attentive and motivated. But a test score may reflect something else, such as the degree to which one is typically motivated to take tests, if all respondents are not equally and highly attentive and motivated (Borghans et al., 2008). Soland (2018) evaluated the degree to which test effort differences could account for achievement gaps. Segal (2012) showed that motivation as indicated by applying effort on low-stakes test is a driving factor for long-term economic success. Kyllonen and Kell (2018) discuss numerous such potential confounding influences on test scores, and similarly confounding cognitive ability influences on noncognitive assessment responses and scores.

Uses for interpersonal and intrapersonal skills assessments

There are many ways in which interpersonal and intrapersonal skills assessments might be used (Kyllonen et al., 2005). Perhaps the three most significant use categories are selection (e.g., college admissions, scholarship reward, job recruitment and selection), student development (e.g., counseling, study skills coursework, leadership development program), and educational system monitoring (e.g., program evaluation, large-scale domestic or international assessment). These use cases vary on several dimensions, which has implications for assessment. The selection use is high stakes, which means that respondents' scores may affect consequential decisions made regarding them, such as whether they are admitted to a particular program, receive a scholarship, or get hired for a job. In these cases, respondents may be motivated to fake on the assessment or receive coaching on how to do so, and so the assessment should be resistant to such attempts. Also, for high-stakes assessments a premium is put on the psychometric properties of the test, that it is reliable, valid for the selection use, and fair, in the sense that the scores can be interpreted in the same way for all users and user subgroups. The AERA/APA/NCME (2014) *Standards for Educational and Psychological Testing* defines high stakes as "a test used to provide results that have important direct consequences for individuals, programs, or institutions ..." (p. 219) and documents the stringent evidence standards needed for this high-stakes use. Naemi et al. (2014) describe a multidimensional forced-choice assessment of noncognitive skills designed for graduate and professional school admissions and workforce personnel selection that falls into this category. Tannenbaum and Kane (2019) argue that there are less-than-high-stakes uses in which some of these stringent standards might be relaxed somewhat. An example might be the case of a placement tests for remedial versus regular instruction.

The development use is generally considered to be low-stakes because one's test scores will not affect consequential decisions about the individual but are instead designed to help the individual develop. AERA et al. (2014) defines low stakes as "a test used to provide results that have only minor or indirect consequences for individuals, programs, or institutions" Here, faking and coaching are not concerns and instead feedback, such as normative feedback, or formative feedback is useful, and suggestions

tied to scores and score profiles are valuable. GMAC's (2013) Reflect™ tool, which was a Likert-scale survey measuring 10 key competencies (innovation, operational thinking, decision-making, strategic vision, strategic self-awareness, resilience, drive, interpersonal intuition, valuing others, and collaboration) is an example of such an assessment. It provided a personalized action plan, library, and benchmarking data from 14 jobs, information designed to prepare business school students for the workplace. ETS's SuccessNavigator™ (Burrus et al., 2013; Markle et al., 2013), was similarly a low-stakes guidance tool, but here the primary application was to serve as an early alert for students who were at risk of leaving college.

The system monitoring use is also low-stakes, at least for the individual test taker. However, in monitoring uses, the data are typically used in policy decisions and so data quality at least at the aggregate level is important. For this application individual scores do not even have to be produced, and a group scores approach can be taken (Mazzeo et al., 2006). For example, different respondents can receive different item subsets, as was done in the DCPS study (Kautz et al., 2021) and is routinely done for measuring cognitive abilities in large-scale assessments such as NAEP and PISA.

Approaches for assessing interpersonal and intrapersonal skills

There are many approaches to assessing interpersonal and intrapersonal skills. However, the predominant approach is ratings, either self-ratings or ratings by others. In this section the ratings approach is reviewed along with several other approaches, which are particularly popular, useful, or promising: multidimensional forced-choice self-assessments, situational judgment tests, and performance tests.

Ratings

Almost all the evidence for the relationship between student social-emotional skills and academic performance is based on ratings, and for students in middle school and above, the vast majority of these are self-ratings (for young students and early childhood targets, teacher or care-giver ratings are the norm) (See McKown et al., 2021, for an exception.). The predominance of ratings and self-ratings can easily be seen in a search through the RAND Assessment Finder (Hamilton et al., 2018), which allows filtering on assessment method. Table 1 provides examples of rating scale items.

Ratings have several problems (Wetzel et al., 2016). Self-ratings are susceptible to reference-group, response style, and social desirability bias. Reference group bias refers to the bias induced by one's reference group—the group to which one compares oneself—when doing a self-evaluation (although this can also apply to others' ratings, it is most often cited in discussions of self-ratings). For example, consider the CORE Self-Management item "I came to class prepared" (the first item listed in Table 1). The

Table 1 Examples of interpersonal-intrapersonal items measuring different constructs.

Assessment	Construct	Instructions/framing statement	Item statement	Response categories
CORE	Self-management	Please answer how often you did the following during the past 30 days. During the past 30 days ...	I came to class prepared.	Almost never/once in a while/ sometimes/often/almost all the time
CORE	Social awareness	Please answer how often you did the following during the past 30 days. During the past 30 days ...	How often did you compliment others' accomplishments?	Not carefully at all/slightly carefully/ somewhat carefully/quite carefully/ extremely carefully
NAEP	Motivation/performance goal	How much does each of the following statements describe a person like you? Select one answer choice on each row	I want to look smart in comparison to the other students in my math class.	Not at all like me/a little bit like me/ somewhat like me/quite a bit like me/very much like me
ETS SN	Social support/connectedness	Please indicate how often each of the statements below applies to you. Please answer each statement separately and honestly.	I feel connected to my peers.	Strongly disagree/disagree/slightly disagree/slightly agree/agree/ strongly agree
PISA 2012	Sense of belonging	Thinking about your school: to what extent do you agree with the following statements? (Please darken only one circle in each row)	I feel happy at school	Strongly agree/agree/disagree/ strongly disagree
BFI-2	Open-mindedness/esthetic sensitivity	I am someone who ...	___Has few artistic interests.	Disagree strongly/disagree a little/ neutral: no opinion/agree a little/ agree strongly

Notes. CORE = California Office to Reform Education (West et al., 2018a); NAEP = National Assessment of Educational Progress (2018) Pilot Mathematics Student Grade 12 (Item 24; NCES, 2018); ETS SN = Educational Testing Service Success Navigator (Markle et al., 2013); PISA 2012 = Student Questionnaire from the Program for International Student Assessment (OECD, 2013; NCES, 2012 [Q40g, Form B]); BFI-2 = Big Five Inventory Version 2 (Soto and John, 2017).

reference group argument is this: There is leeway in the definition of “prepared” and so to answer this, respondents at least implicitly compare themselves to others in class with respect to their level of preparedness. If the respondent is relatively more prepared than classmates “sometimes” or “often,” then they will respond accordingly. But note that this judgment depends on the classmates. Being surrounded by a set of always highly prepared classmates might cause a respondent to judge themselves as relatively unprepared (or to select, for example, the “once in a while” category). That identical level of preparedness, in the context of being surrounded by classmates who themselves are usually unprepared, might warrant an “often” or “almost all the time” response. Different respondents have different reference groups and thus a response given by two respondents is not strictly comparable. Methods such as anchoring vignettes (King and Wand, 2007; King et al., 2004), behaviorally anchored rating scales (Klieger et al., 2018), and Rasch/Guttman Scenario Scales (Ludlow et al., 2020a,b) are attempts to create a common metric to enable comparisons across individuals through anchor items. Kyllonen and Bertling (2014) and He et al. (2017) demonstrated the efficacy of anchoring vignettes in PISA 2012.

Response style bias refers to influences on responding that are independent of the construct being measured, including acquiescence (tendency to agree to statements), extreme response style (tendency to choose the endpoints of the response scales, such as “strongly agree,” “strongly disagree”), and modesty (tendency to choose the middle point[s] in a rating scale). There are useful statistical corrections for some of these (Primi et al., 2019).

Others’ ratings—in which a peer, instructor, advisor, or supervisor is asked to rate an individual student or applicant (e.g., “Please rate the applicant relative to others ... can overcome challenges and setbacks”; from Kyllonen, 2008)—also at least somewhat mitigate the biases in self ratings. But others’ ratings have their own problems, particularly halo/horn effects, which are revealed in high correlations among ratings on disparate attributes, ones that might be expected to have lower correlations (Klieger et al., 2022; McCaffrey et al., 2018).

Multidimensional forced-choice (ranking)

Multidimensional forced-choice (MFC), or ranking methods are alternatives to ratings, designed to mitigate some ratings biases. Response style biases are eliminated because of the choice format. In an MFC item, two (or more) statements are presented together and the respondent is asked to choose which of the two statements better characterizes them, is more true of them, or they agree with more. For example, considering items from Table 1, a respondent might be asked which of two statements better characterizes them: “I came to class prepared” or “I feel connected to my peers.” It is multidimensional because the two statements reflect different dimensions—in this example, self-management and connectedness, respectively. In a ranking version of the task, the respondent is asked to rank the statements (e.g., three, four, or even more statements) with respect to level of agreement with them. Social desirability is addressed by having respondents choose between different statements, where statements are balanced for social desirability (typically, social desirability is assessed in a separate study in which respondents are asked to evaluate the social desirability of the statements, or to select or rate statement that would maximize their chances of getting admitted, getting a scholarship, or getting the job).

Simple sum scoring of forced-choice and ranking methods—in which a dimension score is the sum of the number of times a statement from that dimension is selected—produces what is known as *ipsative* data. With ipsative data respondents may have different dimension scores, but all respondents have the same total score (sum of dimension scores). Ipsative data results in several undesirable measurement properties: it is impossible to have all high or all low dimension scores, correlations between dimension scores must sum to zero, correlations with external measures must sum to zero, and internal consistency reliability estimates are distorted (Brown and Maydeu-Olivares, 2013). Ipsative data with a low number of dimensions results in dimension scores that are negatively correlated, and this can render a factor analysis of item responses uninterpretable. All these properties lie in contrast to the properties of *normative* data, which is the type of data produced by rating scales. Normative data do not constrain scores: it is possible to be high or low on all dimensions, correlations with other dimensions or with external variables do not sum to be zero, and in general, psychometric analyses such as factor analysis and internal consistency reliability produce results that are interpretable.

However, it is not necessary to use simple sum scoring for MFC tasks. Item-response theory (IRT) methods, alternatives to classical sum scoring methods, have been developed that produce normative scores from forced-choice or ranking item responses. It is possible to use IRT scoring approaches for MFC items that produce scores with normative properties, overcoming the problems of ipsative data (Brown and Maydeu-Olivares, 2013; Stark et al., 2005).

There is evidence that forced-choice methods successfully mitigate socially desirable responding (Cao and Drasgow, 2019). There is also evidence that treatments of forced-choice responses that produce normative data, such as the IRT scoring approach or other quasi-ipsative methods (Salgado et al., 2014), result in stronger predictions with outcomes compared to purely ipsative methods or rating scale methods (Salgado and Tauriz, 2014; Salgado et al., 2014). This is an active area of psychometrics development (Brown and Böckenholt, 2022; Bürkner, 2022).

Situational judgment test (SJT)

One alternative to ratings and rankings of statements is the situational judgment test (SJT), which presents a situation and asks the respondent to indicate how they would respond to the situation, or what the best response to the situation would be (can be multiple choice, option ranking, option rating, or short answer). This form of assessment has been used in high stakes admissions

assessments, for business school (Hedlund et al., 2006), medical school (AAMC, 2021), and others. There is an active literature on the kinds of constructs that can be measured with SJTs (Christian et al., 2010) and what the best scoring methods are (De Leng et al., 2017; Zu and Kyllonen, 2020).

Performance tasks

Another alternative to ratings and rankings is a category that might be called performance tasks, in which actual performance is observed and treated as an indicator of various soft skill dimensions. This includes collaborative problem-solving tasks (von Davier et al., 2017), *real effort* tasks (tasks such as repetitively moving objects on the screen with a mouse as an indicator of effort) (Charness et al., 2018), and economic and other games (Falk et al., 2016), and includes both response and process analysis of performance (Hao et al., 2017). These tasks are attractive because they measure actual performance rather than self (or other) assessments of performance, assessments which, as noted earlier, are subject to various biases. Performance measures are often found to be weakly correlated with self-reports (Dang et al., 2020). This may be because performance measures do not achieve the reliability and other psychometric properties traditional questionnaire measures do (Boon-Falleur et al., 2022), but these are early days and we can expect this to be a growing area of continued research.

Conclusions

Considerable research conducted over the past two decades has demonstrated the importance of interpersonal and intrapersonal skills—noncognitive skills, social emotional competencies, personality, or soft skills—on education, workforce, and life outcomes. Consequently, there has been a push to assess these skills in students and workers at all levels, for a variety of purposes, including selection, personal development, and system monitoring. For example, there has been steady growth in the use of interpersonal and intrapersonal skills assessment in hiring (Meinert, 2015), and in offerings in soft skills training for development (Kyllonen et al., 2021), and the OECD (2021) just launched a new comparative (monitoring) study to assess social and emotional skills in cities around the world. This is a dramatic change from the situation two decades ago, when the focus of assessment attention had been almost exclusively on cognitive skills for the purposes of selection, student development, and educational system monitoring. Part of this change reflects growing evidence for the importance of interpersonal and intrapersonal skills; part of the change reflects a more sophisticated understanding of and more sophisticated approaches for assessing such skills.

Many challenges remain. Terminology is a source of confusion and likely will remain that way because different research and application sectors have different terminology preferences. Assessment can serve an important role in standardization through empirical findings: The Big 5 emerged to provide standardization in terminology because of recurring empirical findings of the relationships among certain kinds of self-evaluations despite a wide variety of words and statements used to describe behavior, attitudes, and beliefs. Another challenge is that the dominant assessment methodology is the rating scale, by self or by other, and rating scales have well documented limitations. Future research on the use of actual performance measures, direct or indirect, will likely prove fruitful, and should elevate the status of noncognitive skills assessment to the level of cognitive skills assessment. Performance tasks, including games, real effort tasks, and collaborative tasks and process analysis of performance on such tasks have the potential to overcome the limitations of self-reports, which make comparability over time and across contexts problematic, and prevent unambiguous statements about student growth and development. There is even now skepticism about the interpretability of noncognitive assessment scores, due to reference bias, response style effects, and social desirability and “faking good” bias, phenomena related to the use of rating scale measures (Ziegler et al., 2011). The development of performance measures of interpersonal and intrapersonal skills is rare relative to rating scale measures, and there are many challenges in developing such performance assessments. Still, it is likely that if evaluations of math, reading, and science achievement depended on self-assessments the field of educational assessment would not have the influence on policy and practice that it currently enjoys. Analogously, it is likely that only by moving interpersonal and intrapersonal assessment from ratings to performance will score users and stakeholders become satisfied with the information from such assessments.

Acknowledgment

Preparation for this article was supported by Educational Testing Service, R&D Division, Research & Measurement Sciences, Center for Education and Career Development allocation funding.

References

- Almlund, M., Duckworth, A., Heckman, J., Kautz, T., 2011. Personality psychology and economics. In: Hanushek, E.A., Machin, S., Woessmann, L. (Eds.), *Handbook of the Economics of Education*. Elsevier, Amsterdam, Netherlands, pp. 1–181.
- American Association American Medical Colleges (AAMC), 2021. PREview™ Professional Readiness Exam. <https://www.aamc.org/services/admissions-lifecycle/aamc-preview-professional-readiness-exam>.

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. American Educational Research Association.
- Anderson, J.R., 1983. *The Architecture of Cognition*. Harvard University Press, Cambridge, MA.
- Arthur Jr., W., Bennett Jr., W., Edens, P.S., Bell, S.T., 2003. Effectiveness of training in organizations: a meta-analysis of design and evaluation features. *J. Appl. Psychol.* 88 (2), 234–245. <https://doi.org/10.1037/0021-9010.88.2.234>.
- Berg, J., Osher, D., Same, M.R., Nolan, E., Benson, D., Jacobs, N., 2017. Identifying, Defining, and Measuring Social and Emotional Competencies—Final Report. American Institutes for Research, Washington, DC. Retrieved from: <https://www.air.org/resource/identifying-defining-and-measuring-social-and-emotional-competencies>.
- Blythe, D., Borowski, T., 2018. Ten Criteria for Describing and Selecting SEL Frameworks. CASEL Framework Briefs. <https://measuringSEL.casel.org/wp-content/uploads/2018/09/Framework-A.3.pdf>.
- Boon-Falleur, M., Bouguen, A., Charpentier, A., et al., 2022. Simple questionnaires outperform behavioral tasks to measure socio-emotional skills in students. *Sci. Rep.* 12 (442). <https://doi.org/10.1038/s41598-021-04046-5>.
- Borghans, L., Duckworth, A.L., Heckman, J.J., Weel, B.T., 2008. The economics and psychology of personality traits. *J. Hum. Resour.* 43, 972–1059.
- Borghuis, J., Denissen, J.J.A., Oberski, D., Sijsma, K., Meeus, W.H.J., Branje, S., Koot, H.M., Bleidorn, W., 2017. Big Five personality stability, change, and codevelopment across adolescence and early adulthood. *J. Pers. Soc. Psychol.* 113 (4), 641–657. <https://doi.org/10.1037/pspp0000138>.
- Bowles, S., Gintis, H., Osborne, M., 2001. The determinants of earnings: a behavioral approach. *J. Econ. Lit.* 39 (4), 1137–1176.
- Brandes, C.M., Kushner, S.C., Herzoff, K., Tackett, J.L., 2021. Facet-level personality development in the transition to adolescence: maturity, disruption, and gender differences. *J. Pers. Soc. Psychol.* 121 (5), 1095–1111. <https://doi.org/10.1037/pspp0000367>.
- Brown, A., Böckenholt, U., 2022. Intermittent faking of personality profiles in high-stakes assessments: a grade of membership analysis. *Psychol. Methods*. <https://doi.org/10.1037/met0000295>. Advance online publication.
- Brown, A., Maydeu-Olivares, A., 2013. How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychol. Methods* 18, 36–52. <https://doi.org/10.1037/a0030641>.
- Bürkner, P.C., 2022. On the information obtainable from comparative judgments. *Psychometrika*. <https://doi.org/10.1007/s11336-022-09843-z>.
- Burrus, J., Elliott, D., Brennan, M., Markle, R., Carney, L., Moore, G., Betancourt, A., Jackson, T., Robbins, S., Kyllonen, P., Roberts, R.D., 2013. Putting and Keeping Students on Track: Toward a Comprehensive Model of College Persistence and Goal Attainment. Research Report ETS RR. Educational Testing Service, Princeton, NJ, pp. 13–14. <https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.2013.tb02321.x>.
- Cao, M., Drasgow, F., 2019. Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *J. Appl. Psychol.* 104 (11), 1347–1368. <https://doi.org/10.1037/apl0000414>.
- Collaborative for Academic, Social, and Emotional Learning (CASEL), 2018. SEL Assessment Guide. <https://measuringSEL.casel.org/assessment-guide/>.
- Charness, G., Gneezy, U., Henderson, A., 2018. Experimental methods: measuring effort in economics experiments. *J. Econ. Behav. Organ.* 149, 74–87.
- Chernyshenko, O., Kankaraš, M., Drasgow, F., 2018. Social and Emotional Skills for Student Success and Well-Being: Conceptual Framework for the OECD Study on Social and Emotional Skills. OECD Publishing, Paris. <https://doi.org/10.1787/db1d8e59-en>. OECD Education Working Paper No. 173.
- Christian, M.S., Edwards, B.D., Bradley, J.C., 2010. Situational judgment tests: constructs assessed and a meta-analysis of their criterion-related validities. *Person. Psychol.* 63, 83–117.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*. Routledge. ISBN 978-1-134-74270-7.
- Collaborative for Academic, Social, and Emotional Learning, 2005. *Safe and Sound: An Educational Leader's Guide to Evidence-Based Social and Emotional Learning Programs—Illinois Edition*. Author, Chicago.
- Condon, D.M., 2018. The SAPA Personality Inventory: an empirically—derived, hierarchically—organized self—report personality assessment model. *PsyArXiv* 1–444.
- Connelly, B.S., Ones, D.S., 2010. An other perspective on personality: meta-analytic integration of observers' accuracy and predictive validity. *Psychol. Bull.* 136 (6), 1092–1122. <https://doi.org/10.1037/a0021212>. PMID: 21038940.
- Connolly, J.J., Kavanagh, E.J., Viswesvaran, C., 2007. The convergent validity between self and observer ratings of personality: a meta-analytic review. *Int. J. Sel. Assess.* 15 (1), 110–117. <https://doi.org/10.1111/j.1468-2389.2007.00371.x>.
- Costa, P.T., Terracciano, A., McCrae, R.R., 2001. Gender differences in personality traits across cultures: robust and surprising findings. *J. Pers. Soc. Psychol.* 81 (2), 322–331.
- Crédé, M., Tynan, M.C., Harms, P.D., 2017. Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Pers. Soc. Psychol.* 113 (3), 492–511. <https://doi.org/10.1037/pspp0000102>.
- Cunha, N.M., Martinez, A., Kyllonen, P., Gates, S., 2021. Cross-country comparability of a social-emotional skills assessment designed for youth in low-resource environments. *Int. J. Test.* 21 (3–4), 182–219. <https://doi.org/10.1080/15305058.2021.1995867>.
- Dalal, D.K., Randal, J.G., Cheung, H.K., Gorman, B.C., Roch, S.G., Williams, K.J., 2022. Is there bias in alternatives to standardized tests? An investigation into letters of recommendation. *Int. J. Test.* 22 (1), 21–42. <https://doi.org/10.1080/15305058.2021.2019751>.
- Dang, J., King, K.M., Inzlicht, M., 2020. Why are self-report and behavioral measures weakly correlated? *Trends Cognit. Sci.* 24 (4), 267–269. <https://doi.org/10.1016/j.tics.2020.01.007>.
- De Bolle, M., De Fruyt, F., McCrae, R.R., et al., 2015. The emergence of sex differences in personality traits in early adolescence: a cross-sectional, cross-cultural study. *J. Pers. Soc. Psychol.* 108 (1), 171–185. <https://doi.org/10.1037/a0038497>. PMID: 25603371; PMCID: PMC4327943.
- De Leng, W.E., Stegers-Jager, K.M., Husbands, A., Dowell, J.S., Born, M.P., Themmen, A.P.N., 2017. Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality? *Adv. Health Sci. Educ.* 22 (2), 243–265.
- De Vries, J.H., Spengler, M., Frinrup, A., Mussel, P., 2021. Personality development in emerging adulthood-how the perception of life events and mindset affect personality trait change. *Front. Psychol.* 12, 671421. <https://doi.org/10.3389/fpsyg.2021.671421>. PMID: 34234715; PMCID: PMC8256263.
- Deary, I., 2014. The stability of intelligence from childhood to old age. *Curr. Dir. Psychol. Sci.* 23 (4), 239–245.
- Drasgow, F., Stark, S., Chernyshenko, O.S., Nye, C.D., Hulin, C.L., White, L., 2012. Development of the Tailored, Adaptive Personality Assessment System (TAPAS) to Support Army Selection and Classification Decisions. United States Army Research Institute for the Behavioral and Social Sciences, Fort Belvoir, VA. Technical Report 1311.
- Duckworth, K., Duncan, G.J., Kokko, K., Lyra, A.-L., Metzger, M., Simonton, S., 2012. The Relative Importance of Adolescent Skills and Behaviors for Adult Earnings: A Cross-National Study. Quantitative Social Science—UCL Social Research Institute, University College London. DoQSS Working Papers 12-03.
- Duckworth, A., 2016. *Grit: The Power of Passion and Perseverance*. Scribner/Simon & Schuster.
- Dunkel, C.S., van der Linden, D., Kawamoto, T., Oshio, A., 2021. The general factor of personality as ego-resiliency. *Front. Psychol.* 12. <https://www.frontiersin.org/article/10.3389/fpsyg.2021.741462>.
- Durlak, J.A., Weissberg, R.P., Pachan, M., 2010. A meta-analysis of afterschool programs that seek to promote personal and social skills in children and adolescents. *Am. J. Community Psychol.* 45, 294–309.
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., 2011. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev.* 82 (1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (Eds.), 2015. *Handbook of Social and Emotional Learning: Research and Practice*. The Guilford Press, New York. ISBN 9781462520152.
- Falk, A., Becker, A., Dohmen, T.J., Sunde, U., 2016. The preference survey module: a validated instrument for measuring risk, time, and social preferences. *SSRN Electron. J.* 1–66. <https://doi.org/10.2139/ssrn.2725874>.

- Farrington, C.A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T.S., Johnson, D.W., Beechum, N.O., 2012. Teaching Adolescents to Become Learners. The Role of Noncognitive Factors in Shaping School Performance: A Critical Literature Review. University of Chicago Consortium on Chicago School Research, Chicago.
- Feng, S., Han, Y., Heckman, J.J., Kautz, T., 2022. Comparing the reliability and predictive power of child, teacher, and guardian reports of noncognitive skills. *Proc. Natl. Acad. Sci. U. S. A.* 119 (6). <https://doi.org/10.1073/pnas.2113992119>.
- Foldes, H.J., Duehr, E.E., Ones, D.S., 2008. Group differences in personality: meta-analyses comparing five U. S. racial groups. *Person. Psychol.* 61, 579–616. <https://doi.org/10.1111/j.1744-6570.2008.00123.x>.
- Fricke, H., Loeb, S., Meyer, R.H., Rice, A.B., Pier, L., Hough, H., 2021. Stability of school contributions to student social-emotional learning gains. *Am. J. Educ.* 128 (1). <https://doi.org/10.1086/716550>.
- García, E., Weiss, E., 2017. Education Inequalities at the School Starting Gate: Gaps, Trends, and Strategies to Address Them. Economic Policy Institute. <https://files.eric.ed.gov/fulltext/ED588751.pdf>.
- GMAC, 2013. Reflect™ by GMAC: A Soft-Skills Solution Brought to You by GMAC. <https://www.gmac.com7/media/Files/gmac/Marketing-and-Recruiting/reflect-overview.pdf>.
- Goldberg, L.R., 1993. The structure of phenotypic personality traits. *Am. Psychol.* 48 (1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>.
- Haberman, S.J., 2008. When can subscores have value? *J. Educ. Behav. Stat.* 33 (2), 204–229. <https://doi.org/10.3102/1076998607302636>.
- Hall, G.S., 1916. Adolescence, 2 vols. Appleton, New York.
- Hamilton, L.S., Stecher, B.M., Schweig, J., Baker, G., 2018. RAND Education Assessment Finder. RAND Corporation. <https://www.rand.org/pubs/tools/TL308.html>.
- Hao, J., Chen, L., Flor, M., Liu, L., von Davier, A.A., 2017. CPS-Rater: Automated Sequential Annotation for Conversations in Collaborative Problem-Solving Activities. ETS Research Report Series, ETS RR–17–58. Educational Testing Service, Princeton NJ.
- He, J., Buchholz, J., Klieme, E., 2017. Effects of anchoring vignettes on comparability and predictive validity of student self-reports in 64 cultures. *J. Cross Cult. Psychol.* 48, 319–334.
- Heckman, J.J., Kautz, T., 2014. Achievement tests and the role of character in American life. In: Heckman, J.J., Humphries, J.E., Kautz, T. (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life*. University of Chicago Press, Chicago, IL, pp. 3–56.
- Hedlund, J., Wilt, J.M., Nebel, K., Ashford, S.E., Sternberg, R.J., 2006. Assessing practical intelligence in business school admissions: a supplement to the Graduate Management Admissions Test. *Learn. Individ. Differ.* 16, 101–127.
- Heine, S.J., Buchtel, E.E., 2009. Personality: the universal and the culturally specific. *Annu. Rev. Psychol.* 60, 369–394. <https://doi.org/10.1146/annurev.psych.60.110707.163655>.
- Hough, H., Kalogrides, D., Loeb, S., 2017. Using Surveys of Students' Social-Emotional Learning and School Climate for Accountability and Continuous Improvement. Policy Analysis for California Education. https://www.edpolicyinca.org/sites/default/files/SEL-CC_report.pdf.
- Jackson, J.J., Thoemmes, F., Jonkmann, K., Lüdtke, O., Trautwein, U., 2012. Military training and personality trait development: does the military make the man, or does the man make the military? *Psychol. Sci.* 23 (3), 270–277. <https://doi.org/10.1177/0956797611423545>.
- Jackson, C.K., Porter, S., Easton, J.Q., Blanchard, A., Kiguel, S., 2020. School Effects on Social Emotional Development, School-Based Arrests, and Educational Attainment. NBER Working Paper No. 26759. JEL No. 120.J0. https://www.nber.org/system/files/working_papers/w26759/w26759.pdf.
- Jackson, C.K., 2018. What do test scores miss? The importance of teacher effects on non-test score outcomes. *J. Polit. Econ.* 126 (5). <https://doi.org/10.1086/699018>.
- John, O.P., De Fruyt, F., 2015. Education and Social Progress: Framework for the Longitudinal Study of Social and Emotional Skills in Cities. OECD Publishing, Paris. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/CERI/CD\(2015\)13&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/CERI/CD(2015)13&docLanguage=En).
- John, O.P., Mauskopf, S., 2015. Self-Reported Socio-emotional Qualities: Five Factors for 21st Century Skills? Poster Presented at the Bi-annual Meetings of the Association for Personality Research, Saint Louis, Missouri, USA.
- John, O.P., Srivastava, S., 1999. The Big Five Trait taxonomy: history, measurement, and theoretical perspectives. In: Pervin, L.A., John, O.P. (Eds.), *Handbook of Personality: Theory and Research*. Guilford Press., pp. 102–138.
- John, O.P., Angleitner, A., Ostendorf, F., 1988. The lexical approach to personality: a historical review of trait taxonomic research. *Eur. J. Pers.* 2 (3), 171–203. <https://doi.org/10.1002/per.2410020302>.
- Jones, S., Bailey, R., Brush, K., Nelson, B., 2019. Introduction to the Taxonomy Project: Tools for Selecting & Aligning Frameworks. <https://measuringSEL.caseli.org/wp-content/uploads/2019/02/Frameworks-C.1.pdf>.
- Judge, T.A., Erez, A., Bono, J.E., Thoreson, C.J., 2003. The core self-evaluations scale: development of a measure. *Person. Psychol.* 56, 303–331.
- Kankaraš, M., Suarez-Alvarez, J., 2019. Assessment Framework of the OECD Study on Social and Emotional Skills. OECD Publishing, Paris. OECD Education Working Papers, No. 207 (EDU/WKP(2019)15).
- Kautz, T., Zanoni, W., 2019. Measuring and Fostering Non-cognitive Skills in Adolescence: Evidence From Chicago Public Schools and the OneGoal Program. Mathematica, Princeton, NJ. Unpublished manuscript.
- Kautz, T., Heckman, J.J., Diris, R., Ter Weel, B., Borghans, L., 2014. Fostering and Measuring Skills: Improving Cognitive and Noncognitive Skills to Promote Lifetime Success. OECD, Paris. OECD Education Working Papers.
- Kautz, T., Feeney, K., Chiang, H., Lauffer, S., Bartlett, M., Tilley, C., 2021. Using a Survey of Social and Emotional Learning and School Climate to Inform Decision-Making (REL 2021–114). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <http://ies.ed.gov/ncee/edlabs>.
- Kelley, T.L., 1927. Interpretation of Educational Measurements. World Book Co.
- King, G., Wand, J., 2007. Comparing incomparable survey responses: evaluating and selecting anchoring vignettes. *Polit. Anal.* 15, 46–66.
- King, G., Murray, C.J.L., Salomon, J.A., Tandon, A., 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am. Polit. Sci. Rev.* 44, 341–355.
- Kirkpatrick, D.L., Kirkpatrick, J.D., 2006. *Evaluating Training Programs: The Four Levels*, third ed. Berrett-Koehler, San Francisco. ISBN: 978-1-57675-348-4.
- Klieger, D.M., Kell, H.J., Rikoon, S., Burkander, K.N., Bochenek, J., Shore, J.R., 2018. Development of the Behaviorally Anchored Rating Scales for the Skills Demonstration and Progression Guide. Research Report. ETS RR-18–24. Educational Testing Service, Princeton, NJ.
- Klieger, D., Bochenek, J., Ezzo, C., Holtzman, S., Cline, F., Olivera-Aguilar, M., 2022. Using third-party evaluations to assess socioemotional skills in graduate and professional school admissions. *Int. J. Test.* 22, 72–99. <https://doi.org/10.1080/15305058.2021.2019748>.
- Kuncel, N.R., Kochevar, R.J., Ones, D.S., 2014. A meta-analysis of letters of recommendation in college and graduate admissions: reasons for hope. *Int. J. Sel. Assess.* 22 (1), 101–107. <https://doi.org/10.1111/jisa.12060>.
- Kyllonen, P.C., Bertling, J., 2014. Innovative questionnaire assessment methods to increase cross-country comparability. In: Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*. CRC Press, Boca Raton, pp. 247–285.
- Kyllonen, P.C., Kell, H., 2018. Ability tests measure personality, personality tests measure ability: disentangling construct and method in evaluating the relationship between personality and ability. *J. Intell.* 6 (3), 32. <https://doi.org/10.3390/jintelligence6030032>.
- Kyllonen, P., Walters, A.M., Kaufman, J.C., 2005. Noncognitive constructs and their assessment in graduate education: a review. *Educ. Assess.* 10, 153–184.
- Kyllonen, P.C., Graesser, A.C., Haviland, S.B., Robbins, S.B., Williams, K., 2021. Implementing soft skills training in GIFT, 2022. In: Sinatra, A.M., Graesser, A.C., Hu, X., Goldberg, B., Hampton, A.J., Johnston, J.H. (Eds.), *Design Recommendations for Intelligent Tutoring Systems, Competency-Based Scenario Design*, vol. 9. US Army Combat Capabilities Development Command—Soldier Center, Orlando, FL. ISBN 978-0-9977258-1-0. Available at. <https://www.gifttutoring.org/documents/158>.
- Kyllonen, P.C., 2008. The Research behind the ETS® Personal Potential Index (PPI). Educational Testing Service, Princeton, NJ.

- Kyllonen, P.C., 2016. Socio-emotional and self-management variables in learning and assessment. In: Rupp, A., Leighton, J. (Eds.), *Handbook of Cognition and Assessment*. Wiley/Blackwell, New York.
- Lee, K., Ashton, M.C., 2004. Psychometric properties of the HEXACO personality inventory. *Multivariate Behav. Res.* 39 (2), 329–358. https://doi.org/10.1207/s15327906mbr3902_8.
- Leight, J., 2021. Targeting the Development of Soft Skills in Developing Countries: Evidence From a Growing Literature. *Education in Focus*. International Food Policy Research Institute. <https://blogs.iadb.org/educacion/en/targeting-the-development-of-soft-skills-in-developing-countries-evidence-from-a-growing-literature/>.
- Likert, R., 1932. A technique for the measurement of attitudes. *Arch. Psychol.* 22 (140), 55.
- Lindqvist, E., Vestman, R., 2011. The labor market returns to cognitive and noncognitive ability: evidence from the Swedish enlistment. *Am. Econ. J. Appl. Econ.* 3 (1), 101–128.
- Ludlow, L.H., Anghel, E., Szendey, O., O'Keefe, T., Howell, B., Matz-Costa, C., Braun, H., 2020a. The Boston college living a life of meaning and purpose (BC-LAMP) portfolio: an application of Rasch/Guttman scenario methodology. *J. Appl. Meas.* 21 (2), 134–153.
- Ludlow, L.H., Baez-Cruz, M., Chang, W.-C.C., Reynolds, K., 2020b. Rasch/Guttman Scenario scales: a methodological framework. *J. Appl. Meas.* 21 (4), 361–378.
- Lüttke, O., Roberts, B.W., Trautwein, U., Nagy, G., 2011. A random walk down university avenue: life paths, life events, and personality trait change at the transition to university life. *J. Pers. Soc. Psychol.* 101 (3), 620–637. <https://doi.org/10.1037/a0023743>. PMID: 21744977; PMCID: PMC3596260.
- Lui, P.P., Samuel, D.B., Rollock, D., Leong, F.T.L., Chang, E.C., 2020. Measurement invariance of the five factor model of personality: facet-level analyses among euro and Asian Americans. *Assessment* 27 (5), 887–902. <https://doi.org/10.1177/1073191119873978>.
- Mac Giolla, E., Kajonius, P.J., 2019. Sex differences in personality are larger in gender equal countries: replicating and extending a surprising finding. *Int. J. Psychol.* 54 (6), 705–711. <https://doi.org/10.1002/ijop.12529>.
- Mammadov, S., 2022. Big Five personality traits and academic performance: a meta-analysis. *J. Pers.* 90, 222–255. <https://doi.org/10.1111/jopy.12663>.
- Markle, R., Olivera-Aguilar, M., Jackson, T., Noeth, R., Robbins, S., 2013. Examining Evidence of Reliability, Validity, and Fairness for the SuccessNavigator™ Assessment. Research Report. ETS RR-13–12. ETS Research Report Series.
- Martin-Raugh, M.P., Williams, K.M., Lentini, J., 2020. The Malleability of Workplace-Relevant Noncognitive Constructs: Empirical Evidence From 39 Meta-analyses and Reviews. ETS RR-20–23. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12306>.
- Mazzeo, J., Lazer, S., Zieky, M., 2006. Monitoring educational progress with group-score assessments. In: Brennan, R.L. (Ed.), *Educational Measurement*, fourth ed. Praeger, Westport, CT, pp. 681–699.
- McAbee, S.T., Oswald, F.L., 2013. The criterion-related validity of personality measures for predicting GPA: a meta-analytic validity competition. *Psychol. Assess.* 25 (2), 532–544. <https://doi.org/10.1037/a0031748>.
- McArdle, J., Hamagami, F., Meredith, W., Bradway, K.P., 2000. Modeling the dynamic hypotheses of Gf–Gc theory using longitudinal life-span data. *Learn. Individ. Differ.* 12, 53–79. [https://doi.org/10.1016/S1041-6080\(00\)00036-4](https://doi.org/10.1016/S1041-6080(00)00036-4).
- McCaffrey, D.F., Oliveri, M.E., Holtzman, S., 2018. A Generalizability Theory Study to Examine Sources of Score Variance in Third-Party Evaluations Used in Decision-Making for Graduate School Admissions (ETS GRE-18–03. ETS RR-18-37). Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12225>.
- McCrae, R.R., Terracciano, A., 78 members of the Personality Profiles of Culture Project, 2005. Universal features of personality traits from the observer's perspective: data from 50 cultures. *J. Pers. Soc. Psychol.* 88 (3), 547–561.
- McKown, C., Kharitonova, M., Russo-Ponsaran, N.M., Aksu-Dunya, B., 2021. Development and validation of a shortened form of SELweb EE, a web-based assessment of children's social and emotional competence. *Assessment*. <https://doi.org/10.1177/10731911211046044>.
- Meinert, D., 2015. What do personality tests really reveal? Personality assessments are being used more often in the hiring process. In: *Are They Accurate? News & Features: HR Magazine*. Society for Human Resource Management (SHRM), p. 4384120. <https://www.shrm.org/hr-today/news/hr-magazine/pages/0615-personality-tests.aspx>.
- Messick, S., 1979. Potential uses of noncognitive measurement in education. *J. Educ. Psychol.* 71 (3), 281–292. <https://doi.org/10.1037/0022-0663.71.3.281>.
- Meyer, R.H., Wang, C., Rice, A., 2018. Measuring Students' Social-Emotional Learning Among California's Core Districts: An IRT Modeling Approach [Working Paper]. Policy Analysis for California Education. <https://edpolicyinca.org/publications/measuring-students-social-emotional-learning-among-californias-core-districts-irt>.
- Millsap, R.E., 2011. *Statistical Approaches to Measurement Invariance*. Routledge, New York.
- Musek, J., 2007. A general factor of personality: evidence for the Big One in the five-factor model. *J. Res. Pers.* 41, 1213–1233.
- Naemi, B., Seybert, J., Robbins, S., Kyllonen, P., 2014. Examining the WorkFORCE™ assessment for job fit and core capabilities of FACETS™. *ETS RR* 2014 (2), 1–43. <https://doi.org/10.1002/ets2.12040>.
- National Academies of Sciences, Engineering, and Medicine, 2017. *Supporting Students' College Success: The Role of Assessment of Intrapersonal and Interpersonal Competencies*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24697>.
- National Center for Education Statistics (NCES), 2012. *OECD Program for International Student Assessment 2012 Student Questionnaire Form B*. https://nces.ed.gov/surveys/pisa/pdf/MS12_StQ_FormB_ENG_USA_final.pdf.
- National Center for Education Statistics (NCES), 2018. *NAEP 2018 Student Questionnaire Math Grade 12*. https://nces.ed.gov/nationsreportcard/subject/about/pdf/bgg/student/2018_sq_student_math_g12_pilot.pdf.
- National Research Council, 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/13398>.
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F., Turkheimer, E., 2012. Intelligence: new findings and theoretical developments. *Am. Psychol.* 67 (2), 130–159. <https://doi.org/10.1037/a0026699>. Erratum in: *Am Psychol.* 2012 Feb;67(2):129. PMID: 22233090.
- O'Connor, M.C., Paunonen, S.V., 2007. Big Five personality predictors of post-secondary academic performance. *Pers. Individ. Differ.* 43 (5), 971–990. <https://doi.org/10.1016/j.paid.2007.03.017>.
- OECD (Organization for Economic Co-operation and Development), 2013. *Context questionnaires framework*. In: *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264190511-8-en>.
- OECD (Organization for Economic Co-operation and Development), 2014. *PISA 2012 Technical Report*. OECD Publishing, Paris. <https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf>.
- OECD (Organization for Economic Co-operation and Development), 2015. *Skills for Social Progress: The Power of Social and Emotional Skills*, OECD Skills Studies. OECD Publishing, Paris.
- OECD, 2021. *Beyond Academic Learning: First Results From the Survey of Social and Emotional Skills*. OECD Publishing, Paris. <https://doi.org/10.1787/92a11084-en>.
- Oh, I.-S., Wang, G., Mount, M.K., 2011. Validity of observer ratings of the five-factor model of personality traits: a meta-analysis. *J. Appl. Psychol.* 96 (4), 762–773. <https://doi.org/10.1037/a0021832>.
- Oltmanns, J.R., Smith, G.T., Oltmanns, T.F., Widiger, T.A., 2018. General factors of psychopathology, personality, and personality disorder: across domain comparisons. *Clin. Psychol. Sci.* 6 (4), 581–589. <https://doi.org/10.1177/2167702617750150>.
- Paunonen, S.V., Ashton, M.C., 2001. Big five factors and facets and the prediction of behavior. *J. Pers. Soc. Psychol.* 81 (3), 524–539. <https://doi.org/10.1037/0022-3514.81.3.524>.
- Peterson, C., Seligman, M.E.P., 2004. *Character Strengths and Virtues: A Handbook and Classification*. American Psychological Association; Oxford University Press.
- Petway, K., Rikoon, S., Brenneman, M.W., Burrus, J., Roberts, R.D., 2016. *Development of the Mission Skills Assessment and Evidence of Its Reliability and Internal Structure*. ETS Research Report ETS RR-16–19. Educational Testing Service, Princeton, NJ.
- Poropat, A.E., 2009. A meta-analysis of the five-factor model of personality and academic performance. *Psychol. Bull.* 135 (2), 322–338. <https://doi.org/10.1037/a0014996>.
- Poropat, A.E., 2014. Other-rated personality and academic performance: evidence and implications. *Learn. Individ. Differ.* 34, 24–32. <https://doi.org/10.1016/j.lindif.2014.05.013>.

- Primi, R., De Fruyt, F.D., Santos, D., Antonoplis, S., John, O.P., 2019. True or false? keying direction and acquiescence influence the validity of socioemotional skills items in predicting high school achievement. *Int. J. Test.* 20, 1–25. <https://doi.org/10.1080/15305058.2019.1673398>.
- Primi, R., Santos, D., John, O.P., De Fruyt, F., 2021a. SENNA Inventory for the Assessment of Social and Emotional Skills: Technical Manual. Instituto Ayrton Senna, São Paulo, Brazil. <https://doi.org/10.31234/osf.io/byvpr>.
- Primi, R., Santos, D., John, O.P., De Fruyt, F., 2021b. SENNA inventory for the assessment of social and emotional skills in public school students in Brazil: measuring both identity and self-efficacy. *Front. Psychol.* 12, 716639. <https://doi.org/10.3389/fpsyg.2021.716639>.
- Richardson, M., Abraham, C., Bond, R., 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138 (2), 353–387. <https://doi.org/10.1037/a0026838>.
- Robbins, S.B., Lauver, K., Le, H., Davis, D., Langley, R., Carlstrom, A., 2004. Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol. Bull.* 130 (2), 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>.
- Roberts, B.W., Mroczek, D., 2008. Personality trait change in adulthood. *Curr. Dir. Psychol. Sci.* 17 (1), 31–35. <https://doi.org/10.1111/j.1467-8721.2008.00543.x>.
- Roberts, B.W., Walton, K.E., Viechtbauer, W., 2006. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychol. Bull.* 132 (1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>.
- Roberts, B.W., Kuncel, N.R., Shiner, R., Caspi, A., Goldberg, L.R., 2007. The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspect. Psychol. Sci.* 2, 313–345.
- Roberts, B.W., Luo, J., Briley, D.A., Chow, P.I., Su, R., Hill, P.L., 2017. A systematic review of personality trait change through intervention. *Psychol. Bull.* 143 (2), 117–141. <https://doi.org/10.1037/bul0000088>.
- Rupp, A.A., Templin, J.L., 2008. Unique characteristics of diagnostic classification models: a comprehensive review of the current state-of-the-art. *Measurement* 6 (4), 219–262. <https://doi.org/10.1080/15366360802490866>.
- Ryle, G., 1946. Knowing how and knowing that: the presidential address. *Proc. Aristot. Soc.* 46 (1), 1–16. <https://doi.org/10.1093/aristotelian/46.1.1>.
- Salgado, J.F., Tauriz, G., 2014. The five-factor model, forced-choice personality inventories and performance: a comprehensive meta-analysis of academic and occupational validity studies. *Eur. J. Work. Organ. Psychol.* 23 (1), 3–30. <https://doi.org/10.1080/1359432X.2012.716198>.
- Salgado, J.F., Anderson, N., Tauriz, G., 2014. The validity of ipsative and quasi-ipsative forced-choice personality inventories for different occupational groups: a comprehensive meta-analysis. *J. Occup. Organ. Psychol.* <https://doi.org/10.1111/joop.12098>.
- Santos, D., Primi, R., 2014. Social and Emotional Development and School Learning: A Measurement Proposal in Support of Public Policy. EduLab21, Instituto Ayrton Senna, São Paulo. Available online at: <https://global.ed.gse.harvard.edu/files/geii/files/social-and-emotional-developmente-and-school-learning1.pdf>.
- Saucier, G., Iurino, K., 2020. High-dimensionality personality structure in the natural language: further analyses of classic sets of English-language trait-adjectives. *J. Pers. Soc. Psychol.* 119 (5), 1188–1219. <https://doi.org/10.1037/pspp0000273>.
- Saucier, G., Ostendorf, F., 1999. Hierarchical subcomponents of the Big Five personality factors: a cross-language replication. *J. Pers. Soc. Psychol.* 76 (4), 613–627. <https://doi.org/10.1037/0022-3514.76.4.613>.
- Saucier, G., Hampson, S.E., Goldberg, L.R., 2000. Cross-language studies of lexical personality factors. In: Hampson, S.E. (Ed.), *Advances in Personality Psychology*, vol. 1. Psychology Press, New York, pp. 1–36.
- Schmitt, D.P., Realo, A., Voracek, M., Allik, J., 2008. Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *J. Pers. Soc. Psychol.* 94 (1), 168–182.
- Schweig, J., Baker, G., Hamilton, L.S., Stecher, B.M., 2018. Building a Repository of Assessments of Interpersonal, Intrapersonal, and Higher-Order Cognitive Competencies. RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR2508.html.
- Segal, C., 2012. Working when no one is watching: motivation, test scores, and economic success. *Manag. Sci.* 58 (8), 1438–1457.
- Segal, C., 2013. Misbehavior, education, and labor market outcomes. *J. Eur. Econ. Assoc.* 11, 743–779.
- Sinharay, S., Haberman, S., Puhani, G., 2007. Subscores based on classical test theory: to report or not to report. *Educ. Meas.* 26 (4), 21–28. <https://doi.org/10.1111/j.1745-3992.2007.00105.x>.
- Soland, J., 2018. Are achievement gap estimates biased by differential student test effort? Putting an important policy metric to the test. *Teach. Coll. Rec.* 120 (12), 1–26.
- Solanki, S., Fitzpatrick, D., Jones, M.R., Lee, H., 2020. Social-psychological interventions in college: a meta-analysis of effects on academic outcomes and heterogeneity by study context and treated population. *Educ. Res. Rev.* 31, 100359. <https://doi.org/10.1016/j.edurev.2020.100359>.
- Soto, C.J., John, O.P., 2017. The next Big Five Inventory (BFI-2): developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *J. Pers. Soc. Psychol.* 113 (1), 117–143. <https://doi.org/10.1037/pspp0000096>.
- Soto, C.J., John, O.P., Gosling, S.D., Potter, J., 2011. Age differences in personality traits from 10 to 65: Big-Five domains and facets in a large cross-sectional sample. *J. Pers. Soc. Psychol.* 100, 330–348.
- Soto, C.J., Napolitano, C.M., Roberts, B.W., 2021. Taking skills seriously: toward an integrative model and agenda for social, emotional, and behavioral skills. *Curr. Dir. Psychol. Sci.* 30 (1), 26–33. <https://doi.org/10.1177/0963721420978613>.
- Soto, C.J., Napolitano, C.M., Sewell, M.N., Yoon, H.J., Roberts, B.W., 2022. An integrative framework for conceptualizing and assessing social, emotional, and behavioral skills: the BESSI. *J. Pers. Soc. Psychol.* <https://doi.org/10.1037/pspp0000401>. Advance online publication.
- Soto, C.J., 2016. The little six personality dimensions from early childhood to early adulthood: mean-level age and gender differences in parents' reports. *J. Pers.* 84 (4), 409–422. <https://doi.org/10.1111/jopy.12168>.
- Spearman, C., 1904. "General intelligence," objectively determined and measured. *Am. J. Psychol.* 15, 201–293.
- Srivastava, S., John, O.P., Gosling, S., Potter, J., 2003. Development of personality in early and middle adulthood: set like plaster or persistent change? *J. Pers. Soc. Psychol.* 84 (5), 1041–1053.
- Stark, S., Chernyshenko, O.S., Drasgow, F., 2005. An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: the multi-unidimensional pairwise-preference model. *Appl. Psychol. Meas.* 29 (3), 184–203. <https://doi.org/10.1177/0146621604273988>.
- Tannenbaum, R.J., Kane, M.T., 2019. Stakes in testing: not a simple dichotomy but a profile of consequences that guides needed evidence of measurement quality. ETS RR 2019 (1), 1–16. <https://doi.org/10.1002/ets2.12255>. Princeton, NJ: Educational Testing Service.
- Taylor, R.D., Oberle, E., Durlak, J.A., Weissberg, R.P., 2017. Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. *Child Dev.* 88 (4), 1156–1171.
- Trilling, B., Fadel, C., 2009. *21st Century Skills: Learning for Life in Our Times*. Jossey-Bass, San Francisco, CA.
- Ugo, I., Hill, L., 2017. Student Achievement and Growth on California's K-12 Assessments. Public Policy Institute of California. https://www.ppic.org/wp-content/uploads/content/pubs/report/R_0417IUR.pdf.
- Von Davier, A.A., Zhu, M., Kyllonen, P.C. (Eds.), 2017. *Innovative Assessment of Collaboration (Methodology of Educational Measurement and Assessment)*. Springer, Cham, Switzerland.
- West, M.R., Buckley, K., Krachman, S.B., Bookman, N., 2018a. Development and implementation of student social-emotional surveys in the CORE districts. *J. Appl. Dev. Psychol.* 55, 119–129. <https://doi.org/10.1016/j.appdev.2017.06.001>.
- West, M.R., Pier, L., Fricke, H., Loeb, S., Meyer, R.H., Rice, A.B., 2018b. Trends in Student Social-Emotional Learning: Evidence From the CORE Districts. Policy Analysis for California Education.
- West, M.R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R.H., Rice, A.B., 2020. Trends in student social-emotional learning: evidence from the first large-scale panel student survey. *Educ. Eval. Pol. Anal.* 42 (2), 279–303.

- Wetzel, E., Böhnke, J.R., Brown, A., 2016. Response biases. In: Leong, F.T.L., Bartram, D., Cheung, F.M., Geisinger, K.F., Iliescu, D. (Eds.), *The ITC International Handbook of Testing and Assessment*. Oxford University Press, pp. 349–363. <https://doi.org/10.1093/med:psych/9780199356942.003.0024>.
- Wilmot, M.P., Ones, D.S., 2019. A century of research on conscientiousness at work. *Proc. Natl. Acad. Sci. U. S. A.* 116 (46), 23004–23010. <https://doi.org/10.1073/pnas.1908430116>.
- Yeager, D.S., Hanselman, P., Walton, G.M., et al., 2019. A national experiment reveals where a growth mindset improves achievement. *Nature* 573, 364–369. <https://doi.org/10.1038/s41586-019-1466-y>.
- Zell, E., Lesick, T.L., 2021. Big five personality traits and performance: a quantitative synthesis of 50+ meta-analyses. *J. Pers.* 1–15. <https://doi.org/10.1111/jopy.12683>.
- Ziegler, M., MacCann, C., Roberts, R.D. (Eds.), 2011. *New Perspectives on Faking in Personality Assessment*. Oxford University Press, Oxford.
- Zu, J., Kyllonen, P.C., 2020. Nominal response model is useful for scoring multiple-choice situational judgment tests. *Organ. Res. Methods* 23 (2), 342–366. <https://doi.org/10.1177/1094428118812669>.

Further reading

- Abrahams, L., Pancorbo, G., Primi, R., Santos, D., Kyllonen, P., John, O.P., De Fruyt, F., 2019. Social-emotional skill assessment in children and adolescents: advances and challenges in personality, clinical, and educational contexts. *Psychol. Assess.* 31 (4), 460–473. <https://doi.org/10.1037/pas0000591>.
- McKown, C., 2019. *Assessing Students' Social and Emotional Learning: A Guide to Meaningful Measurement (SEL Solutions Series)*. Norton Professional Books, New York, 978-0-393-71335-0.

Assessment of collaborative problem solving skills

Jessica Andrews-Todd^a and Carol M. Forsyth^b, ^a Educational Testing Service, Princeton, NJ, United States; and ^b Educational Testing Service, San Francisco, CA, United States

© 2023 Elsevier Ltd. All rights reserved.

Importance of collaborative problem solving	494
What is collaborative problem solving?	494
Challenges with CPS assessment	495
In-task assessment framework	495
Context for example applications	496
Example application for operationalizing CPS	497
Example applications for evidence identification and accumulation for CPS skills	499
Descriptions of learner behaviors	499
Profiles	500
Discussion and ongoing work	500
Acknowledgment	501
References	501

Importance of collaborative problem solving

Changes in our world (e.g., demographic changes, digitalization, globalization) have led to shifts in the kinds of skills necessary for success in the 21st century. One such skill includes collaborative problem solving (CPS), as the complexity of problems faced in the 21st century are difficult for a single individual to solve alone. In fact, among the top five skills considered “very important” for successful job performance are teamwork/collaboration and critical thinking/problem solving (Casner-Lotto and Barrington, 2006). Other sources have further identified interpersonal and communication skills as important for current and future success (National Research Council, 2008; Burrus et al., 2013; Whorton et al., 2017; McGunagle and Zizka, 2020). As students go through the education system and into the workforce, they will be expected to work with others, in face-to-face and virtual contexts, to solve complex problems, make decisions, and generate innovative ideas. These kinds of activities require individuals to have competency in skills associated with CPS.

Furthermore, decades of research have shown benefits of engaging in collaborative activities. In particular, there are positive effects of collaboration for a number of outcomes, including learning, performance, social, emotional, and psychological well-being (Lou et al., 2001; Uribe et al., 2003; Gillies, 2004; Andrews and Rapp, 2015; Jeong et al., 2019). Additionally, work has shown that problem solutions generated in groups are often better than those generated by individuals working alone (Schwartz, 1995; Dillenbourg, 1999; Barron, 2000). Such results can be attributed to division of labor afforded by groups and the opportunity to incorporate solutions from other group members with different perspectives, knowledge, and experiences (Graesser et al., 2018b). These benefits further highlight the utility of CPS in many everyday contexts such as school, the workplace, the military, and home.

What is collaborative problem solving?

Collaborative problem solving can be defined as, “the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills, and efforts to reach that solution” (OECD, 2013, p. 6). As can be seen from the definition, there are some features that differentiate CPS from other forms of collaboration. One factor is that for CPS, the team has the goal of solving a novel problem. Team members come together to develop a plan to reach a solution to a given problem. A second factor is that the quality of the solution is transparent for the team. As they develop and implement a plan for getting from their current state to the goal state, there is opportunity to determine the extent to which the problem is solved during the course of problem solving. This is different from collaborative learning which has the goal of using collaboration to support learning of some particular content for team members without the requirement that a problem be solved. A third factor is that there is often a differentiation of roles in which team members may take on different roles (e.g., facilitator or recorder) to support solving a problem. This can differentiate CPS from collaborative decision making which involves team members having the same questions to answer or decisions to make, and often access to the same information for all team members. A fourth factor is there is typically interdependency among the team members wherein team members bring to the situation different perspectives, knowledge, and experiences that support solving a problem (Graesser et al., 2018a). This is important for many of the problems we now face in the 21st century which are often unable to be solved by a single person and require input from diverse experiences and expertise of others who bring relevant knowledge to the table.

The growing need for individuals to pool their knowledge, skills, and expertise together to solve the challenging problems of today and the general acknowledgment of the importance of CPS has motivated interest in national assessments, international assessments and curricular reform geared toward the measurement and development of CPS skills (National Research Council, 2011; OECD, 2013; Hesse et al., 2015; Care et al., 2016; Fiore et al., 2017). This interest has brought to light particular challenges with assessing an individual's competency in CPS. In the next section, we outline key challenges that make CPS assessment a relatively difficult undertaking.

Challenges with CPS assessment

The definition of CPS alone implies the need for assessment of numerous complex skills such as sharing important information and monitoring problem solving to discover an optimal solution to the task at hand. Thus, the definition in and of itself creates some obstacles for measuring how one performs on CPS tasks, as it requires students to display evidence of skills associated with both a cognitive dimension corresponding to problem solving processes and a social dimension corresponding to collaboration or teamwork processes. This makes for a diverse set of skills across a range of disciplines such as individual problem solving, linguistics, computer-supported collaborative learning, and critical thinking (Care et al., 2016). The process-oriented nature of the construct undermines the use of traditional types of assessment (e.g., multiple choice questions) (Davey et al., 2015). For example, with other domains like mathematics, one can determine a student's proficiency according to whether or not a student has correctly answered a question associated with the specific isolated concept under investigation (e.g., the concept of function). CPS is more complicated, as determining proficiency requires individuals to go through the process of interacting with others to solve a problem. The behaviors that the individuals engage in during this interaction need to allow for gathering appropriate evidence to signify an individual's CPS skills.

To accomplish the goal of assessing CPS skills, students must interact in a space that captures evidence of individual skills. Creating and administering CPS activities for individuals in face-to-face settings is an option, but can be quite expensive, time consuming, and difficult to carry out at a large scale. Difficulty with measuring interactive constructs with traditional types of tests and challenges with face-to-face administrations has led a turn to the use of digital environments which have seen a great deal of technological advancements over the years. Such environments offer opportunities for individuals to demonstrate performance around complex problems, and further afford the recording of all individuals' actions and discourse during task engagement, providing sources of evidence for CPS proficiency. However, the use of digital environments can introduce additional challenges. One challenge is associated with operationalizing CPS at the level of granularity suitable for data output from the digital environment (Kerr et al., 2016). A secondary challenge is identifying targeted skills from fine-grained log data generated during individuals' task engagement. Digital environments (e.g., games, simulations, scenario-based tasks) can log every action that each user takes thus making it difficult for analysts to determine the signal from the noise. Each action is typically at a small grain size (e.g., mouse clicks, utterances) which requires multiple levels of parsing, aggregating information to create features, and extracting the features, and there are few guiding principles for how to carry out such work (Gobert et al., 2012). Additionally, log file data are often sparse in nature, meaning that one variable or logged action may have records for many individuals whereas another action may only have few if any instances of a student behaving in a certain way. This makes it difficult to understand the generalizability of discovered student actions (Romero et al., 2009; Honey and Hilton, 2011).

These challenges have led to a number of design approaches for CPS assessment, many of which incorporate constraints on the computerized environment and thereby reduce the degrees of freedom for individuals' interaction with the task to aid in establishing traditional measurement properties such as reliability and validity (all of which require a certain amount of control over the problem space). For example, the 2015 Program for International Student Assessment's (PISA) large-scale international assessment of CPS utilized a computer simulated human-agent approach such that each human participant interacted with one, two, or three computer artificial agents as their team members rather than humans. Communication between the human participants and computer agents were carried out using a predetermined set of message options that human participants could select for each problem. This design was utilized to provide consistency and control over the interactions (OECD, 2013, 2017). Other CPS assessment work has utilized highly constrained environments in similar ways in an effort to narrow the problem space (Chung et al., 1999; Hsieh and O'Neil, 2002; Herborn et al., 2017). Some work has incorporated open chat dialogue among human collaborators, but often analyses have treated the chat messages as simply another action type without considering the content of the chat messages (Care and Griffin, 2014; Adams et al., 2015) or the problem has been constrained so that little variation can occur for the kinds of actions and communication in which individuals engage (Hao et al., 2015; Liu et al., 2015). All of these design decisions, while chosen to deal with various constraints (e.g., limited capabilities for natural language processing), do not allow for detailed measurement of both the actions and discourse of individuals at a fine grain size (Andrews-Todd and Forsyth, 2020). In the next section, we describe one approach developed at Educational Testing Service that was designed to address many of the aforementioned challenges for CPS assessment and to support measurement of CPS skills when the previously mentioned constraints are undesirable.

In-task assessment framework

The in-task assessment framework (I-TAF) was developed to provide guiding principles to support the development of assessments for complex constructs in digital environments (Kerr et al., 2016). I-TAF is informed by evidence-centered design (ECD) (Mislevy et al., 2003), and specifically serves as an approach to instantiate the conceptual assessment framework layer of ECD (i.e., the

student model, task model, evidence model) when evaluating complex constructs in digital environments where the grain size of the data is small. The student model represents what is being measured and includes defining the variables that make up the targeted knowledge, skills, and abilities. The task model addresses where we are measuring the targeted skills and includes the features or affordances available in a task that would be needed to attain evidence of targeted skills. The evidence model concerns how we will measure the targeted skills and includes how we will identify features corresponding to what individuals do and how we will use that information to make inferences about what individuals know and can do. I-TAF provides principles for carrying out each of these models, particularly how one might operationalize a construct, extract evidence of the construct, and link the evidence back to the construct of interest when dealing with low-level log data rather than traditional item response data (Andrews-Todd and Forsyth, 2020).

With standard item response data, variables for the student model are associated with whether a person answers a question correctly or incorrectly; however, with digital environments and process-oriented constructs like CPS, more complex processes are needed to define variables associated with the targeted constructs. In particular, the construct needs to be operationalized in terms of behaviors. For this purpose, I-TAF advocates for the use of an ontology (similar to a concept map) which lays out the construct in a principled way and does so at the level of granularity of data output from a targeted task environment. Specifically, an ontology provides a graphical representation in which high- and low-level concepts are represented by nodes and relationships between those concepts are represented by links (an example of a portion of an ontology is given in Fig. 2 and discussed in the [Example application for operationalizing CPS](#) section). This representation provides a unified definition and shared vocabulary for a domain that can make explicit the presumed dependencies among concepts and skills and also provide a format that can be communicated among humans and computational systems. Because data on complex constructs in digital environments are often process-oriented rather than the selection of a correct or incorrect answer choice, nodes in the ontology should not only include concepts and skills associated with a targeted construct or domain, but also an additional layer that provides behaviors or actions one can take to demonstrate evidence of targeted concepts and skills (e.g., “take measurement” or “ask expert”). A further layer of the ontology introduces tactics that are available in a targeted task environment for carrying out strategies in the ontology. This is important since tactics for carrying out a given strategy can differ based on affordances available in a task. For example, one task may have the option to visit the library to carry out the “ask expert” strategy while another task may afford speaking to a non-playable character. With this, the ontology provides key evidence elements to be identified in the data to make inferences about individuals (Kerr et al., 2016).

I-TAF also provides principles for identifying features using combined top-down and bottom-up approaches. For example, human (e.g., qualitative coding) or machine (e.g., natural language processing) driven approaches can be used to identify features in the data that were outlined in the ontology. Any unattributed actions found in the data can utilize data mining or machine learning approaches to aggregate individual actions into meaningful units of behavior. Extracted behaviors can then be aggregated into variables with processes such as counting the number of instances the feature is observed or calculating the proportion of instances of a correct action. Such variables can then serve as inputs into different measurement or statistical models. The development of the ontology also informs task design, as it can make salient affordances that are (or are not) available in a task to facilitate its measurement properties. For example, if a strategy is deemed important and included in the ontology, but there are no affordances available in the task for carrying out the strategy, then changes to the task design are needed to accommodate the desired strategy. For additional insights related to CPS task design, see Hao et al. (2019) and Coiro et al. (2018). For more detailed information about I-TAF, see Andrews-Todd and Kerr (2019) and Kerr et al. (2016). In the sections to follow, we highlight some examples of how to apply some principles of I-TAF to deal with aforementioned challenges associated with CPS assessment work.

Context for example applications

To describe how to address CPS assessment challenges with I-TAF, we will set examples in the context of a task environment called the Three-Resistor Activity (Horwitz et al., 2017). The Three-Resistor Activity is a collaborative simulation-based task on electronics concepts associated with current, resistance, and voltage. In the task, students work in groups of three, each on a separate computer that runs a simulation of a portion of an electronics circuit. The three simulations are linked to form a series circuit. On each computer screen, students have access to a digital multimeter (DMM) to take measurements, probes from the DMM that can attach to a resistor, a resistor where students can view and change the circuit’s resistance, a zoom button to view other teammates’ boards, a calculator to perform calculations, a text-based chat window to communicate with teammates, and submit buttons to submit work (see Fig. 1 for a screenshot of the task interface).

In the Three-Resistor Activity, students are tasked with setting their resistors so that the voltage across them reaches a specified goal voltage. With the circuits connected in series, any change that one person makes will affect readings on other students’ boards, thus making it important for students to work together to coordinate how to reach their goal voltage values. The task includes four levels of increasing difficulty, with higher levels requiring students to not only reach their goal voltage values, but also figure out the unknown resistance and supply voltage of a fourth circuit in the series. As students interact in the task, all of their relevant actions (e.g., measurements, calculations, resistor changes, chat messages, submissions) are captured and logged to a database.

Teaching Teamwork: Level D

Circuit 3 (User: Hawk, Group: Birds)

Online

We got it! View All Circuits

E is unknown, R0 is unknown, your goal is to make your voltage 4 volts.

2.45

Welcome! You can find your goal at the top of the screen. Use this area to chat with your teammates.

System: Eagle has joined on Circuit 1.

System: Seagull has joined on Circuit 2.

System: Hawk has joined on Circuit 3.

Hawk: hello

Enter chat message here...

Send Chat Message

Enter Unknown Values

E:

R0:

Submit Unknown Values

Fig. 1 Screenshot of the three-resistor activity (level 4, circuit 3).

Example application for operationalizing CPS

In an NSF-funded grant (DUE-1535224), researchers utilized I-TAF to support the operationalization of CPS. Specifically, they developed a CPS ontology for use in assessing CPS skills for undergraduate students in engineering, electronics, and physics courses who completed the Three-Resistor Activity. To develop the CPS ontology, they first conducted an extensive literature review to explore existing frameworks on relevant constructs and work in related areas such as individual problem solving, computer-supported collaborative learning, linguistics, and organizational psychology. This review yielded a list of candidate concepts and skills for the ontology. Subject matter experts (SME's) then reviewed the list and determined relationships between agreed-upon components and definitions. Any disagreements among the SME's were resolved through discussion. This process yielded a construct definition for CPS with concepts or skills associated with CPS and the relationship between them. As noted in the previous section on the "In-Task Assessment Framework," additional steps in developing an I-TAF ontology includes incorporating additional layers to specify the construct at multiple layers of granularity to appropriately reason up from low level behaviors in the digital environments (e.g., button clicks) to higher level proficiencies.

The additional layers in the ontology operationalize CPS in terms of behaviors or strategies and then the tactics or affordances available in a task to carry out the strategies. For these additional layers, team members came together to determine what kinds of behaviors a student would need to show to demonstrate evidence of the skills in the construct definition portion of the ontology. For example, for the "executing" skill, one sub-skill is "enact strategies" and an example behavior for this skill was "take action to carry out plan" with the tactics or affordances available in the Three-Resistor activity to carry out these behaviors being "calculate," "modify," and "measure." As another example, for the "negotiating" skill, the sub-skills are "agree," "disagree," and "resolve conflict" and corresponding example behaviors or strategies are "make disagreement statement," "make agree statement," and "propose alternate" with tactics or affordances available in the Three-Resistor Activity to carry out these behaviors being the "chat" function. A screenshot of a portion of the CPS ontology can be found in Fig. 2. For additional information on how the CPS ontology was created, see Andrews-Todd and Kerr (2019).

Altogether, the CPS ontology includes nine high level skills divided into a social (teamwork, collaboration) and a cognitive (problem solving, taskwork) dimension. In the social dimension, the four skills include maintaining communication, sharing information, establishing shared understanding and negotiating. Maintaining communication includes the subskills rapport

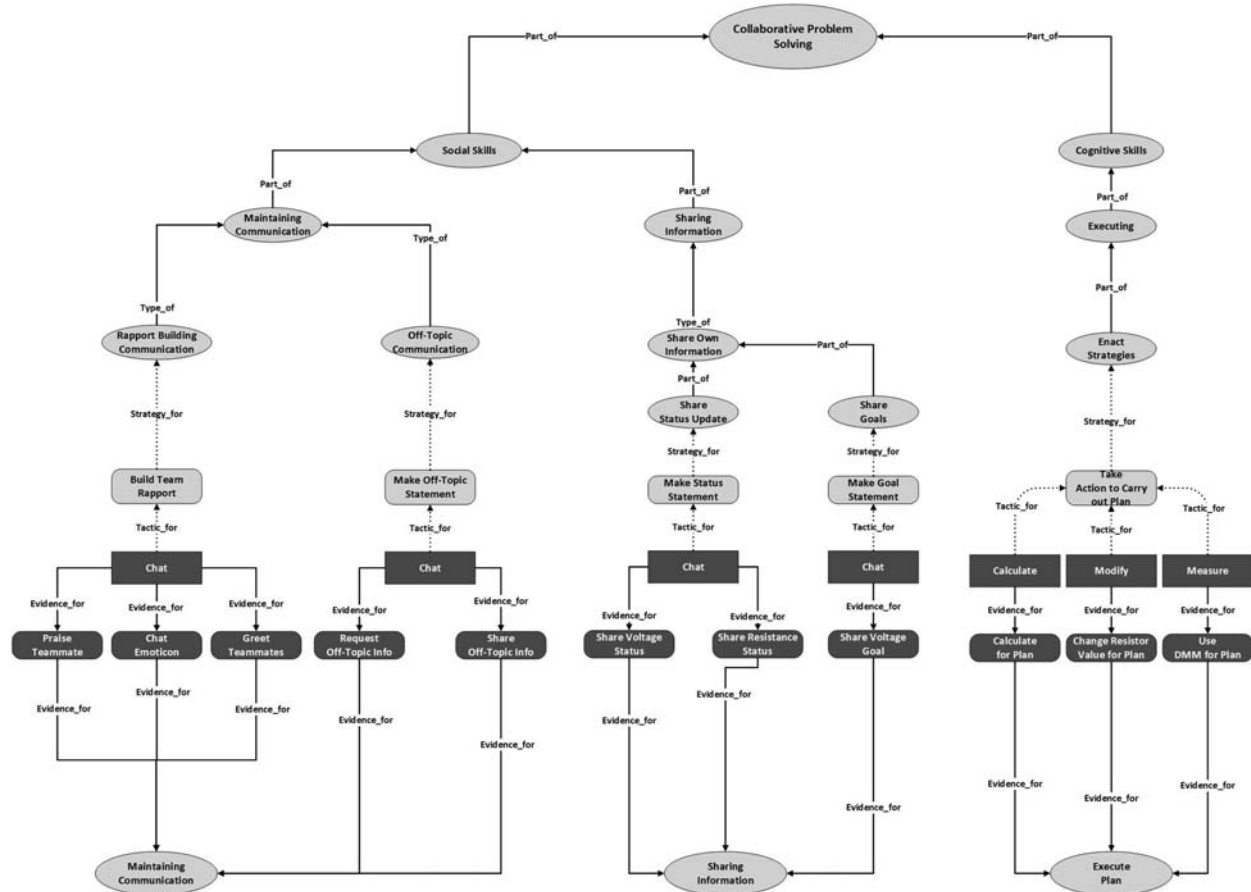


Fig. 2 Portion of the CPS ontology.

building communication (e.g., greeting teammates, praising teammates), off-topic communication (e.g., discussing what was had for breakfast), and inappropriate communication (e.g., degrading teammates). Sharing information includes the sub-skills sharing one's own information (e.g., sharing one's goal voltage value), sharing task or resource information (e.g., sharing where the zoom button is located), and sharing the state of one's own understanding (e.g., metacognitive statements such as "I don't know"). Establishing shared understanding includes subskills associated with requesting information from teammates (e.g., "what is your resistance value?"), providing statements to indicate comprehension of teammates' communication (e.g., "ok," "I got you," or requests for clarification), and communication to make repairs when misunderstandings arise. Negotiating includes the sub-skills expressing agreement (e.g., "yes," or "you're right"), expressing disagreement (e.g., "that's not right"), and resolving conflicts.

In the cognitive dimension, the five skills include exploring and understanding, representing and formulating, planning, monitoring and executing. Exploring and understanding includes sub-skills associated with exploring the problem space and trying to understand the problem. Representing and formulating includes sub-skills for representing the problem (e.g., "this is a series circuit") and formulating hypotheses. Planning includes subskills for setting goals, managing resources (e.g., establishing roles), and developing strategies for solving the problem. Executing includes sub-skills for enacting strategies (e.g., changing to a resistance value suggested by a teammate), suggesting actions for teammates to take to carry out a plan (e.g., "adjust your resistor to 300 ohms"), and reporting own actions to others for carrying out a plan (e.g., "I am going to set mine higher"). Monitoring includes sub-skills for monitoring success toward solving the problem (e.g., "I've got my goal voltage") and monitoring the organization of the team (e.g., "Let's get a move on Hawk"). A list of the high level CPS skills and their definitions can be found in [Table 1](#).

To test the generalizability of the CPS ontology, researchers have applied it in multiple other projects with various contexts and those applications showed the CPS ontology was suitable across different age groups, domains, group compositions, and communication modalities. One example is an IES-funded grant (#R305A170432) which included CPS assessment for middle and high school aged students working in pairs to complete two tasks. One task was an online mathematics task on linear function concepts and argumentation called the T-Shirt Math Task ([Andrews-Todd et al., 2019](#)) in which students worked to determine which of three t-shirt companies was the best choice to purchase t-shirts for their class. The second task was an open-ended physics educational game called Physics Playground ([Shute et al., 2013](#)) in which students, across multiple task levels, had the goal of drawing various objects (e.g., ramp, springboard) to make a ball hit a balloon target. In both tasks, students had multiple modalities available for

Table 1 CPS skills and definitions.

<i>Dimension/skills</i>	<i>Definitions</i>
Social dimension	
Maintaining communication	Content irrelevant social-oriented communication (off topic, rapport building).
Sharing information	Sharing content relevant information (status, role, task information).
Establishing shared understanding	Communication to get on same page as teammates and ensure understanding/comprehension of teammates' contributions.
Negotiating	Communication to express agreement, disagreement, and resolve conflicts that arise.
Cognitive dimension	
Exploring and understanding	Actions to build a mental representation of components of the problem
Representing and formulating	Communication to represent the problem or develop hypotheses.
Planning	Communication to set goals, manage resources, or develop strategies for solving the problem.
Executing	Actions to carry out a strategy or communication to suggest actions for teammates or report own actions.
Monitoring	Actions to check the status of teammates or communication to state task progress and determine whether teammates are doing as they should.

communication, including video and audio channels. For this work, the CPS ontology was modified at only the tactic level to align the ontology with the affordances available in the project tasks. For example, for the executing skill which has the behavior of “take action to carry out plan,” the tactics in the T-Shirt Math Task for carrying out this behavior would be to “type” a response or “submit” an answer choice rather than the affordances of “measure,” “modify,” and “calculate” from the Three-Resistor Activity. In situations in which targeted tasks have similar affordances, no modifications are needed at the tactic layer of the ontology.

Example applications for evidence identification and accumulation for CPS skills

As previously mentioned, discovering targeted behaviors within fine-grained log data aligning to the skills of complex constructs such as CPS can be challenging. With the Three-Resistor Activity, features were extracted from the log data that were identified a priori from the CPS ontology. In the given context, the majority of the data were chat data so qualitative coding was used to support feature extraction. The CPS ontology was used to create a rubric for the human raters. To do this, the nodes and links corresponding to the CPS skills in the ontology were transformed into written protocols that included the high level CPS skills, sub-skills within each skill, definitions associated with each skill, example behaviors from log data that provide evidence of each skill, and action types corresponding to each skill (e.g., chat, measurement, calculation). The raters were trained on the rubric using a subset of log data from the Three-Resistor Activity where they would identify each row corresponding to a student action or discourse as evidence of one CPS skill. Two raters then coded 20% of the data to establish inter-rater reliability. Any disagreements that arose for raters were discussed for the purpose of coming to a common understanding of the rubric interpretation. The remaining data were divided among the raters to code independently. Raters have demonstrated good agreement across two data collections with this CPS rubric (Kappa = 0.84 and 0.93, respectively) (Andrews-Todd and Forsyth, 2020; Forsyth et al., 2020).

Note that machine-driven approaches for feature extraction are also applicable, and can potentially save time and money. For example, in the previously mentioned IES-funded CPS assessment work, natural language processing (NLP) algorithms were used to automatically analyze the CPS discourse. Specifically, Random Forest Classifiers, a deep transfer learning approach called the Bidirectional Encoder Representations from Transformers (BERT) model, and a version of the BERT model that incorporates contextual utterances (i.e., previous and subsequent utterances) have been explored, with the latter performing best in predicting CPS skills within two tasks focused on math and physics (Pugh et al., 2021). Recent work has also applied machine learning methods, specifically a k-nearest neighbors classifier, to automate the classification of CPS skills from the CPS ontology for data from the Three-Resistor Activity. This method showed strong agreement with human annotations (Flor and Andrews-Todd, in press). Additional work has also explored various NLP approaches for automating the identification of CPS skills with promising results (e.g., Flor et al., 2016; Hao et al., 2017; Pugh et al., 2022; Stewart et al., 2019). Once features are extracted, analyses can be carried out with the features to accumulate evidence and report high level proficiencies associated with CPS for individuals. In the following subsections, we provide some examples for how this has been carried out in prior work on the Three-Resistor Activity to show descriptions of learner behaviors and profiles to characterize individuals as types of collaborative problem solvers.

Descriptions of learner behaviors

One strategy for aggregating evidence about individuals' CPS skills is to provide simple descriptions of behaviors. Methods for describing learner behaviors can be approaches such as generating frequencies or counts associated with the number of times an individual displayed targeted CPS skills as evidenced by the extracted features or proportions that take into account all CPS behaviors displayed or totals from a population of interest (e.g., other students from the class) (Zapata-Rivera et al., 2021). With this strategy, we can examine the number of displayed behaviors for each skill not only to understand what the students did but to

also ensure that the behaviors from the CPS ontology were indeed exhibited. This is important because if individuals are not displaying behaviors associated with important CPS skills, we may either need to modify the environment to better provide opportunities for individuals to display targeted CPS skills, consider the possibility that a certain skill cannot be measured with the current task, or infer that individuals may not possess the necessary level of CPS proficiency to display targeted CPS behaviors when given the opportunity. Across two different datasets for the Three-Resistor Activity, students did indeed display the targeted CPS skills from the CPS ontology, but variance did exist. Specifically, students displayed a high level of executing actions whereas representing and formulating and negotiation were displayed less frequently (Andrews-Todd and Forsyth, 2020). Other work has similarly applied such strategies for characterizing CPS behaviors through counts of descriptive behaviors (e.g., Scoular et al., 2017). These types of insights can inform inputs for creating profiles of types of collaborative problem solvers as shown in the next subsection.

Profiles

Another strategy for aggregating evidence about individuals' CPS skills is a profile approach which can provide a way to characterize learner behaviors at a high level. Andrews-Todd and Forsyth have used two different techniques to create learner profiles, including comparing proportional values of displayed skills on social and cognitive CPS dimensions via median splits (Andrews-Todd and Forsyth, 2020) and hierarchical clustering techniques (Andrews-Todd et al., 2018; Forsyth et al., 2020), both of which build on the frequency strategy described previously. In the first technique, students were categorized based on aggregate frequencies of CPS skills corresponding to the high level CPS skills. The frequencies of CPS skills for each dimension (i.e., social vs. cognitive) were ranked and split into roughly 2 equal groups of high and low display of CPS skills for each dimension, respectively. After calculating these for each dimension independently, four groups were then created, corresponding to a 2×2 social and cognitive dimension matrix (i.e., high cognitive with high social, high cognitive with low social, low cognitive with high social, and low cognitive with low social). Aligning with expectations, students in the low cognitive/low social profile tended to underperform relative to students in other profiles on the Three-Resistor Activity (Andrews-Todd and Forsyth, 2020).

In the second technique, an exploratory clustering method was conducted with the CPS skill frequencies that were generated from qualitative coding of the Three-Resistor Activity log data. A four cluster solution was utilized, as these clusters aligned with interpretations from relevant theory and literature associated with collaboration, cognitive, and communication sciences. The four clusters were labeled as *Chatty Doers*, *Social Loafers*, *Active Collaborators*, and *Group Organizers*. Chatty Doers included individuals who had a high level of executing actions (e.g., resistor changes, calculations) and content irrelevant social communication (maintaining communication). Social Loafers were students that displayed below average frequencies on all CPS skills whereas Active Collaborators displayed above average frequencies on the social and cognitive CPS skills relative to the other profiles. Group Organizers tended to display behaviors associated with organizing the problem and the group (e.g., monitoring actions, representing and formulating) (Andrews-Todd et al., 2018). Similar cluster analyses were run on another dataset from the Three-Resistor Activity and yielded four clusters, two of which were different from the first cluster analysis. The two new clusters were *Super Socials* who displayed high levels of social CPS skills relative to other clusters, but did not demonstrate many important cognitive CPS behaviors (other than representing and formulating) and *Low Collaborators* who communicated little with teammates, but did carry out a high level of executing actions (e.g., resistor changes, measurements) (Forsyth et al., 2020). It is important to note that only profiles that aligned to theory or the ontology were accepted as a theoretically-grounded data mining technique was employed to ensure the meaningfulness of the findings (Forsyth et al., 2014). Each of the profiles were validated with external measures and results were in expected directions (e.g., Active Collaborators tended to perform better than Social Loafers). Other work has similarly utilized profile methods with CPS data and has also yielded promising results (e.g., Herborn et al., 2017; Coiro et al., 2019; Dowell et al., 2020).

Discussion and ongoing work

In recent years, more attention has been given to the assessment of CPS skills through interactive tasks, particularly with international assessments such as PISA (OECD, 2013) and the Assessment and Teaching of 21st Century Skills (ATC21S) (Griffin et al., 2012) and national assessments such as the National Assessment of Educational Progress (NAEP) (Fiore et al., 2017). Such attention, motivated by new challenges in today's world that require competency in CPS, has led to advancements that seek to deal with a number of challenges associated with assessment of complex constructs like CPS. In this article, we describe one such advancement, the I-TAF framework, that has utility in providing guiding principles for assessing complex constructs in digital environments. Specifically, we describe and provide examples for how I-TAF can be used to support the operationalization of CPS at a level of grain size appropriate for data from digital environments. We further provide examples for how I-TAF can support evidence identification or extraction of features related to targeted CPS skills, and evidence accumulation, including how to aggregate information about individuals' CPS behaviors to make inferences about higher level CPS proficiencies. The use of such an approach that incorporates a competency model a priori (considered top-down) along with data-driven (or bottom-up) approaches is important when considering a complex construct such as CPS and the thousands of rows of log data produced by student interaction in digital environments. Using purely bottom-up approaches can present challenges for linking features or behaviors to constructs in a targeted domain. Specifically, features without theoretical grounding can produce spurious correlations to skills that are not well aligned with the targeted construct. Purely top-down approaches can also present challenges such as not allowing for the emergence or identification of new features or behaviors that may be important to capture.

There still remain some areas in the space of CPS work that were not addressed in the current article. One area includes exploration of the generalizability of individuals' CPS skills from one context to the next. This is an area that still requires further investigation, though preliminary work suggests mixed results. For example, [Martin-Raugh et al. \(2020\)](#) found evidence for the generalizability of negotiation skills across tasks. Alternatively, [Andrews-Todd et al. \(2021\)](#) found that 37% of participants demonstrated behaviors associated with the same CPS profile across a math and physics task; however, only 10% of participants showed change from high levels of CPS skills to low levels of CPS skills (or vice versa) from one task to the next. The generalizability of CPS skills across contexts likely depends on a number of factors such as the type of task being completed, the composition of the group, and the communication modality.

A second area includes development of quantitative scoring strategies for CPS assessment. In the current article, we discuss strategies for aggregating information about individuals to characterize their CPS behaviors using more descriptive means. Such strategies can support providing meaningful information to learners, teachers, and other stakeholders that can be used to support learning interventions to target identified gaps. These characterizations of learner behaviors require replication across numerous domains and environments to ensure validity, a process that is a part of ongoing work. However, some work has also sought to instead create ways to provide quantitative scores for individuals regarding their CPS skill (e.g., the PISA 2015 CPS assessment). Some of this work has done so as to accommodate various constraints related to issues such as budget, time, particular requirements associated with psychometric quality, or technological limitations. One challenge with such strategies is that CPS is often mapped to a unidimensional scale with approaches such as item response theory (IRT) models; however, CPS is a multidimensional construct. Using methods that assume unidimensionality oversimplifies the richness of CPS and information gathered from corresponding tasks and can also compromise the validity of the assessment ([Hao et al., 2019](#)). There is ongoing work that seeks to develop and employ statistical and psychometric models that take into account the multidimensional nature of CPS. Example models include network analyses, Andersen/Rasch multivariate IRT models, and Bayesian Knowledge Tracing (e.g., [Andrews et al., 2017](#); [He et al., 2017](#); [Polyak et al., 2017](#); [Zhu et al., 2020](#)).

A third area that builds on the assessment work being done for CPS is learning interventions and training materials that target development of CPS skills. With acknowledgment of the importance of CPS for success in today's world, it is important to not only be able to inform individuals about their CPS proficiency, but also to facilitate improvement in CPS skills that are needed. Unfortunately, little work has been put forth to tackle training to help support development of CPS skills for individuals in colleges, industry, and the military, particularly work that is theoretically driven ([Fiore et al., 2018](#); [Graesser et al., 2018a](#)), though work in the computer-supported collaborative learning literature has explored supports for collaborative learning such as adaptive feedback (e.g., [Soller et al., 2005](#); [Walker et al., 2009](#); [Hassan et al., 2019](#)). Andrews-Todd and her colleagues have begun such work through different avenues such as developing and piloting online training modules for CPS skills for higher education populations and conducting research to better understand how to provide real-time feedback to individuals in groups solving problems in computerized environments focusing on types of feedback and timing of delivery. Additional future work includes providing real-time reports in the form of dashboards to teachers and students for intervention during collaborative problem solving. Presenting dashboards that are intuitive to users requires its own line of research, including exploring the needs, knowledge, and attitudes of users, the usability of the dashboard itself, how to best present data in a dashboard, and how to support interpretation of information for decision making ([Zapata-Rivera and Katz, 2014](#); [Kannan et al., 2018](#); [Zapata-Rivera et al., 2018](#)).

Although challenges still remain, significant progress in defining and measuring CPS skills in digital assessments has been made. This is important as CPS is becoming increasingly more valued in the age of technology. Students will need to be able to demonstrate proficiency in working in multidisciplinary teams to solve problems. Developing and implementing CPS assessments and supports for improving CPS skills will remain of strong interest for researchers for years to come.

Acknowledgment

This material is based upon work supported by the Institute of Education Sciences under Grant R305A170432 awarded to the first author and in collaboration with CRESST and the University of Colorado, Boulder. The opinions expressed are those of the authors and do not necessarily represent the views of the Institute of Education Sciences.

References

- Adams, R., et al., 2015. Automatic coding procedures for collaborative problem solving. In: Griffin, P., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer, pp. 115–132.
- Andrews, J.J., Rapp, D.N., 2015. Benefits, costs, and challenges of collaboration for learning and memory. *Transl. Issues Psychol. Sci.* 1, 182–191.
- Andrews, J.J., et al., 2017. Modeling collaborative interaction patterns in a simulation-based task. *J. Educ. Meas.* 54 (1), 54–69.
- Andrews-Todd, J., Forsyth, C.M., 2020. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Comput. Hum. Behav.* 104, 105759. <https://doi.org/10.1016/j.chb.2018.10.025>.
- Andrews-Todd, J., Kerr, D., 2019. Application of ontologies for assessing collaborative problem solving skills. *Int. J. Test.* 19 (2), 172–187.
- Andrews-Todd, J., et al., 2018. Identifying profiles of collaborative problem solvers in an online electronics environment. In: Boyer, K.E., Yudelson, M. (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining*. International Educational Data Mining Society, Buffalo, NY, pp. 239–245.
- Andrews-Todd, J., Jackson, G.T., Kurzum, C., 2019. Collaborative Problem Solving Assessment in an Online Mathematics Task. Research Report RR-19-24. Educational Testing Service, Princeton, NJ.

- Andrews-Todd, J., Steinberg, J., Jiang, Y., 2021. Stability of Collaborative Problem Solving Profile Membership Across Tasks. Educational Testing Service.
- Barron, B., 2000. Achieving coordination in collaborative problem-solving groups. *J. Learn. Sci.* 9 (4), 403–436.
- Burrus, J., et al., 2013. Identifying the Most Important 21st Century Workforce Competencies: An Analysis of the Occupational Information Network (O* NET). ETS RR-13-21. Educational Testing Service, Princeton, NJ.
- Care, E., Griffin, P., 2014. An approach to assessment of collaborative problem solving. *Res. Pract. Technol. Enhanc. Learn.* 9 (3), 367–388.
- Care, E., Scoular, C., Griffin, P., 2016. Assessment of collaborative problem solving in education environments. *Appl. Meas. Educ.* 29 (4), 250–264.
- Casner-Lotto, J., Barrington, L., 2006. Are They Really Ready for Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century U.S. Workforce. The Conference Board, Partnership for 21st Century Skills, Corporate Voices for Working Families, and Society for Human Resource Management, Washington, DC. Available at: <http://eric.ed.gov/?id=ED519465>. (Accessed 8 July 2015).
- Chung, G.K.W.K., O'Neil, H.F., Herl, H.E., 1999. The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Comput. Hum. Behav.* 15 (3), 463–493.
- Coiro, J., Sparks, J.R., Kulikowich, J.M., 2018. Assessing online collaborative inquiry and social deliberation skills as learners navigate multiple sources and perspectives. In: Braasch, J.L., Braten, I., McCrudden, M.T. (Eds.), *Handbook of Multiple Source Use*. Routledge, New York, pp. 485–501.
- Coiro, J., et al., 2019. Students engaging in multiple-source inquiry tasks: capturing dimensions of collaborative online inquiry and social deliberation. *Liter. Res.* 68 (1), 271–292.
- Davey, T., et al., 2015. Psychometric Considerations for the Next Generation of Performance Assessment. Educational Testing Service, Princeton, NJ.
- Dillenbourg, P., 1999. Collaborative Learning: Cognitive and Computational Approaches. Pergamon, Elsevier Science, Amsterdam.
- Dowell, N.M., et al., 2020. Exploring the relationship between emergent sociocognitive roles, collaborative problem-solving skills, and outcomes: a group communication analysis. *J. Learn. Anal.* 7 (1), 38–57.
- Fiore, S.M., et al., 2017. Collaborative Problem Solving: Considerations for the National Assessment of Educational Progress. National Center for Education Statistics, Alexandria, VA.
- Fiore, S.M., Graesser, A., Greiff, S., 2018. Collaborative problem-solving education for the twenty-first-century workforce. *Nat. Hum. Behav.* 2 (6), 367–369.
- Flor, M., Andrews-Todd, J., in press. Towards automatic annotation of collaborative problem solving skills in technology-enhanced environments. *J. Comput. Assist. Learn.*
- Flor, M., et al., 2016. Automated classification of collaborative problem solving interactions in simulated science tasks. In: Tetreault, J., et al. (Eds.), *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications*. Association for Computational Linguistics, Stroudsburg, PA, pp. 31–41.
- Forsyth, C.M., et al., 2014. Discovering theoretically grounded predictors of shallow vs. deep-level learning. In: Stamper, J., et al. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*. International Educational Data Mining Society, London, UK, pp. 229–232.
- Forsyth, C., Andrews-Todd, J., Steinberg, J., 2020. Are you really a team player? Profiling of collaborative problem solvers in an online environment. In: Rafferty, A.N., et al. (Eds.), *Proceedings of the 13th International Conference on Educational Data Mining*. International Educational Data Mining Society, Irfane, Morocco, pp. 403–408.
- Gillies, R.M., 2004. The effects of cooperative learning on junior high school students during small group learning. *Learn. Instruct.* 14 (2), 197–213.
- Gobert, J.D., et al., 2012. Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. *J. Educ. Data Mining* 4, 111–143.
- Graesser, A.C., Fiore, S.M., et al., 2018a. Advancing the science of collaborative problem solving. *Psychol. Sci. Publ. Interest* 19 (2), 59–92.
- Graesser, A.C., Foltz, P.W., et al., 2018b. Challenges of assessing collaborative problem solving. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, Cham, Switzerland, pp. 75–91.
- Griffin, P., McGaw, B., Care, E. (Eds.), 2012. *Assessment and Teaching of 21st Century Skills*. Springer, New York.
- Hao, J., et al., 2015. Assessing collaborative problem solving with simulation based tasks. In: *Proceedings of the 11th International Conference on Computer-Supported Collaborative Learning*. International Society for the Learning Sciences, Gothenburg, Sweden.
- Hao, J., et al., 2017. CPS-Rater: Automated Sequential Annotation for Conversations in Collaborative Problem-Solving Activities. RR-17-58. Educational Testing Service, Princeton, NJ, pp. 1–9. <https://doi.org/10.1002/ets2.12184>. Available at.
- Hao, J., et al., 2019. Psychometric Considerations and a General Scoring Strategy for Assessments of Collaborative Problem Solving. ETS RR-19-41. Educational Testing Service, Princeton, NJ, pp. 1–17.
- Hassan, M.A., et al., 2019. An adaptive feedback system to improve student performance based on collaborative behavior. *IEEE Access* 7, 107171–107178.
- He, Q., et al., 2017. Collaborative problem solving measures in the Program for International Student Assessment (PISA). In: von Davier, A.A., Zhu, M., Kyllonen, P.C. (Eds.), *Innovative Assessment of Collaboration*. Springer, New York, NY, pp. 95–111.
- Herborn, K., Mustafić, M., Greiff, S., 2017. Mapping an experiment-based assessment of collaborative behavior onto collaborative problem solving in PISA 2015: a cluster analysis approach for collaborator profiles. *J. Educ. Meas.* 54 (1), 103–122.
- Hesse, F., et al., 2015. A framework for teachable collaborative problem solving skills. In: Griffin, P., Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer, New York, pp. 37–56.
- Honey, M.A., Hilton, M.L. (Eds.), 2011. *Learning Science Through Computer Games and Simulations*. National Academies Press, Washington, DC.
- Horwitz, P., et al., 2017. Teaching Teamwork: electronics instruction in a collaborative environment. *Community Coll. J.* 41 (6), 341–343.
- Hsieh, I.L.G., O'Neil, H.F., 2002. Types of feedback in a computer-based collaborative problem-solving group task. *Comput. Hum. Behav.* 18 (6), 699–715.
- Jeong, H., Hmelo-Silver, C.E., Jo, K., 2019. Ten years of computer-supported collaborative learning: a meta-analysis of CSCL in STEM education during 2005–2014. *Educ. Res. Rev.* 28, 100284.
- Kannan, P., Zapata-Rivera, D., Leibowitz, E.A., 2018. Interpretation of score reports by diverse subgroups of parents. *Educ. Assess.* 23 (3), 173–194.
- Kerr, D., Andrews, J.J., Mislavy, R.J., 2016. The in-task assessment framework for behavioral data. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*. Wiley-Blackwell, Hoboken, NJ, pp. 472–507.
- Liu, L., et al., 2015. A tough nut to crack: measuring collaborative problem solving. In: Rosen, Y., Ferrara, S., Mosharraf, M. (Eds.), *Handbook of Research on Computational Tools for Real-World Skill Development*. IGI-Global, Hershey, PA, pp. 344–359.
- Lou, Y., Abrami, P.C., d'Apollonia, S., 2001. Small group and individual learning with technology: a meta-analysis. *Rev. Educ. Res.* 71 (3), 449–521.
- Martin-Raugh, M.P., et al., 2020. Negotiation as an interpersonal skill: generalizability of negotiation outcomes and tactics across contexts at the individual and collective levels. *Comput. Hum. Behav.* 104, 105966.
- McGonagle, D., Zizka, L., 2020. Employability skills for 21st-century STEM students: the employers' perspective. *High Educ. Skills Work. Base Learn.* 10 (3), 591–606.
- Mislavy, R.J., Almond, R.G., Lukas, J.F., 2003. A Brief Introduction to Evidence-Centered Design. ETS Research Report RR-03-16. Educational Testing Service, Princeton, NJ.
- National Research Council, 2008. *Research on Future Skill Demands: A Workshop Summary*. National Academies Press, Washington, DC.
- National Research Council, 2011. *Assessing 21st Century Skills*. National Academies Press, Washington, DC.
- OECD, 2013. *PISA 2015 Collaborative Problem Solving Framework*. OECD Publishing, Paris.
- OECD, 2017. *PISA 2015 Results (Volume V): Collaborative Problem Solving*. OECD Publishing, Paris.
- Polyak, S.T., von Davier, A.A., Peterschmidt, K., 2017. Computational psychometrics for the measurement of collaborative problem solving skills. *Front. Psychol.* 8, 2029. <https://doi.org/10.3389/fpsyg.2017.02029>.
- Pugh, S.L., et al., 2021. Say what? Automatic modeling of collaborative problem solving skills from student speech in the wild. In: Hsiao, I.-H., et al. (Eds.), *Proceedings of the 14th International Conference on Educational Data Mining*. International Educational Data Mining Society, Paris, France, pp. 55–67.
- Pugh, S.L., Rao, A., Stewart, A.E., D'Mello, S.K., 2022. Do speech-based collaboration analytics generalize across task contexts? LAK22: 12th International Learning Analytics and Knowledge Conference, 208–218.
- Romero, C., et al., 2009. Evolutionary algorithms for subgroup discovery in e-learning: a practical application using Moodle data. *Expert Syst. Appl.* 36, 1632–1644.
- Schwartz, D.L., 1995. The emergence of abstract representations in dyad problem solving. *J. Learn. Sci.* 4 (3), 321–354.

- Scoular, C., Care, E., Awwal, N., 2017. An approach to scoring collaboration in online game environments. *Electron. J. eLearn.* 15 (4), 335–342.
- Shute, V.J., Ventura, M., Kim, Y.J., 2013. Assessment and learning of qualitative physics in Newton's Playground. *J. Educ. Res.* 106, 423–430.
- Soller, A., et al., 2005. From mirroring to guiding: a review of state of the art technology for supporting collaborative learning. *Int. J. Artif. Intell. Educ.* 15 (4), 261–290.
- Stewart, A.E., Vrzkova, H., Sun, C., Yonehiro, J., Stone, C.A., Duran, N. D., Shute, V., D'Mello, S.K., 2019. I say, you say, we say: using spoken language to model socio-cognitive processes during computer-supported collaborative problem solving. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–19.
- Uribe, D., Klein, J.D., Sullivan, H., 2003. The effect of computer-mediated collaborative learning on solving ill-defined problems. *Educ. Technol. Res. Dev.* 51 (1), 5–19.
- Walker, E., Rummel, N., Koedinger, K., 2009. Beyond explicit feedback: new directions in adaptive collaborative learning support. In: O'Malley, C., et al. (Eds.), *Computer Supported Collaborative Learning Practices- CSDL 2009 Conference Proceedings*. International Society of the Learning Sciences, pp. 552–556.
- Whorton, R., et al., 2017. Critical skills for the 21st century workforce. In: Burrus, J., et al. (Eds.), *Building Better Students: Preparation for the Workforce*. Oxford University Press, New York, pp. 47–72.
- Zapata-Rivera, D., Katz, I.R., 2014. Keeping your audience in mind: applying audience analysis to the design of interactive score reports. *Assess Educ. Princ. Pol. Pract.* 21 (4), 442–463.
- Zapata-Rivera, D., et al., 2018. Designing and evaluating reporting systems in the context of new assessments. In: Schmorow, D., Fidopiastis, C. (Eds.), *Augmented Cognition: Users and Contexts*. AC 2018. Lecture Notes in Computer Science. Springer, Cham, pp. 143–153.
- Zapata-Rivera, D., Andrews-Todd, J., Oliveri, M.E., 2021. Communicating assessment information in the context of a workplace formative task. *J. Writing Anal.* 5, 324–341. <https://doi.org/10.37514/JWA-J.2021.5.1.10>.
- Zhu, M., Andrews-Todd, J., Zhang, M., 2020. Application of network analysis in understanding collaborative problem solving processes and skills. In: Hong, J., Lissitz, R.W. (Eds.), *Innovative Psychometric Modeling and Methods*. Information Age, Charlotte, NC.

Conversation-based assessment: current findings and future work

Diego Zapata-Rivera, Jesse R. Sparks, Carol M. Forsyth, and Blair Lehman, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	504
Dialog systems	504
Conversation-based assessments	505
CBA development and scoring	506
CBA examples	507
CBAs and caring assessments	508
CBAs and student emotions	509
CBAs and individual differences	510
Moving forward	511
Summary	513
Appendix A. Table of resources	514
References	515

Introduction

Conversation-based assessments (CBAs) offer a way to engage students in computer-based dialogs aimed at gathering evidence of what they know or can do. Advances in natural language understanding, dialog systems, and interaction with artificial agents make it possible to design and implement guided natural language conversations that can be used to collect evidence of students' knowledge, skills, and other attributes (KSAs). These CBAs can be used in conjunction with other types of tasks and can be embedded in scenario-based tasks, simulations, and other technology rich tasks.

In this article, we: (a) review relevant work in the area of dialog systems; (b) describe how CBAs were designed to assess complex 21st century constructs such as science inquiry and English language communication skills, elaborate on development and scoring issues, and provide examples of research on the development and evaluation of CBA tasks; (c) connect this work to work on designing and evaluating personalized, “caring” assessments, which extend the scope of the student model to include cognitive and non-cognitive aspects of the student as well as aspects of the learning context to create assessment experiences that provide valued evidence of student performance while considering students' learning needs (Zapata-Rivera, 2017), and elaborate on the connection of this work with Universal Design for Learning (UDL; CAST, 2018) as a framework to support the varying needs of diverse learners; (d) discuss work on exploring student emotions and individual differences in CBAs; (e) report on current challenges and areas for future work elicited from experts in dialog systems; and (f) discuss our view of challenges and future work in the area of CBAs. In addition, we provide an annotated list of resources in [Appendix A](#) for new students and researchers compiled from experts in the field.

Dialog systems

Conversation-based systems for learning and assessment provide opportunities for students to discuss various topics with artificial agents—embodied characters or “talking heads”—that can respond to students in real time. These types of educational environments began emerging in the 21st century (Biswas et al., 2010; Gholson et al., 2009; McNamara et al., 2006; Olney et al., 2012; Ward et al., 2013). These systems cover academic topics ranging from physics, biology, scientific reasoning, and computer literacy to constructs such as self-regulation and meta-cognition (see Graesser, 2016 for review).

Dialog systems often make use of artificial virtual pedagogical agents (Johnson et al., 2000; Johnson and Lester, 2016)—virtual characters or avatars designed to engage the learner and provide a simulated social context in which the conversational interactions occur with the virtual agent, with the aim of supporting students' learning. Virtual agents promote deep learning and motivation through adaptive dialogic interactions (Atkinson, 2002; Baylor and Kim, 2005; Moreno et al., 2001). Conversational agents can convey information, pose questions, provide feedback, guide attention, and serve as a virtual teammate for the human learner (Johnson et al., 2000; Johnson and Lester, 2016). Agents' features may vary in terms of their epistemic roles or personas (e.g., knowledgeable expert or supportive motivator; Baylor and Kim, 2005; Baylor and Ryu, 2003), the quality of the language used such as formality or politeness (e.g., Wang and Johnson, 2008; Li and Graesser, 2017, 2021), whether they communicate via text, voiced audio, or some combination (Johnson and Lester, 2016; Mayer, 2005), among others (see Schroeder et al., 2013).

Perhaps one of the better-known systems that has repeatedly shown high learning gains across two decades is AutoTutor, deployed to middle-schoolers and college students. AutoTutor teaches multiple topics in STEM areas and has shown effect sizes

ranging from 0.6 to 2.0 which is comparable to human tutors (for a review, see [Graesser, 2016](#)). This is not surprising as AutoTutor was created based on analysis of expert tutors' and students' conversations ([Graesser and Person, 1994](#)). The dialogic framework uses scaffolding moves which are positively correlated with learning in AutoTutor ([Jackson and Graesser, 2006](#)). The positive relationship to learning may be explained by Chi's ICAP (Interactive, Constructive, Active, Passive) model of learning which predicts that active interaction aids learning more so than passive interaction (for more information see [Chi, 2009](#)). (Note: Additional constructs may also aid in learning with this framework, for more information [Forsyth et al., 2020b](#)). The primary scaffolding moves correspond to a framework called Expectation-Misconception tailored dialog. The main idea of this framework is that a student is asked a question with an expected correct answer and known misconceptions that are identified and implemented within the system a priori. Scaffolding techniques are then used to adaptively help the student articulate these expected answers. Such scaffolding moves include hints (broad clues), prompts (requiring single word or phrase answers), misconceptions correction, and assertions. A mixture of two different algorithms (Regular Expressions and Latent Semantic Analysis) are used to analyze the human input. Specifically, regular expressions are implemented to capture a specific word or phrase ([Jurafsky and Martin, 2008](#)). Latent Semantic Analysis (LSA) is a geometric pattern matching algorithm that allows the human input to be interpreted in context (see [Landauer et al., 2007](#)). Together, the evaluation of human responses made by these two algorithms are highly correlated with expert human raters; that is, the automated raters and expert human raters have moderate/high inter-rater reliability ([Cai et al., 2011](#)), making possible the illusion of intelligence when artificial agents provide appropriate responses. After the human input is analyzed and categorized, the dialog manager or dialog engine is used to determine the appropriate response (e.g., what should be said? And who should say it?).

There have been several AutoTutor derivatives based on this conversational framework, including OperationARA ([Millis et al., 2011](#); [Forsyth et al., 2014](#)). OperationARA is a serious game designed to teach research methodology to college students and could potentially also be well-suited for high-school students. Students learn about twelve topics of research methodology while engaging with two or more virtual agents across three different teaching modules. The students begin by learning the basic shallow information (such as vocabulary) and then progress to applying information and finally generating questions (three levels of student interaction that research suggests will require students to have a progressively deeper understanding of the content; [Forsyth et al., 2013, 2020b](#)).

The conversational structure of OperationARA is based on the AutoTutor framework but with the extension of one or more additional artificial agents. The resulting multi-party conversations are referred to as "trialogues" ([Graesser et al., 2017](#)). Within ARA, trialogues increase the adaptivity of the conversation so that students receive personalized instruction based on their current level of knowledge. Specifically, students with low prior knowledge may benefit the most from simply observing two artificial agents, an expert and a student agent, interact (vicarious learning mode; [Driscoll et al., 2003](#)). Conversely, students with high prior knowledge may benefit more from teaching another agent, referred to as the teachable agent mode. The intermediate students receive the regular AutoTutor framework. In addition to providing this level of personalization, trialogues create more assessment possibilities as novel dialogic conversations are possible.

Another application of dialog-based approaches involves the assessment of complex constructs such as collaborative problem solving (CPS; [Andrews-Todd and Forsyth, 2020, 2022](#); [Graesser et al., 2017a,b, 2018](#); [Herborn et al., 2020](#); [Rosé et al., 2017](#)). In fact, many 21st century skills such as CPS lend themselves to be assessed using interactive tasks, as assessing CPS in a simulated environment is necessary to capture evidence of both cognitive and social skills ([Andrews-Todd and Forsyth, 2020](#); [Coiro et al., 2019](#)). For example, PISA has implemented dialog-based tasks to assess CPS on a world-wide scale using a simulated environment ([Csapo et al., 2012](#)).

The simulated environment implemented in PISA for this purpose used artificial agents to guide students in highly constrained, text-based, menu-driven conversations. This environment includes some of the characteristics of CBAs, but it is designed for high-stakes assessment. Specifically, rather than the mixed-initiative free flowing conversations implemented in other tutorial systems (e.g., AutoTutor; for review, see [Graesser, 2016](#); [GuruTutor](#); [Olney et al., 2012](#)) as well as other adaptive modeling for student interactions that provide students the needed level of support based on their current knowledge level to aid them in solving physics problems (e.g., [Katz et al., 2021](#)), students were given forced choice options to respond to the agents. Although, this design choice can be used to elicit clear evidence of particular aspects of CPS skills ([OECD, 2013](#)), it limits the types of conversations that could be observed in real-world contexts. Researchers have continued to explore methods of analysis to measure CPS in less constrained environments to improve assessment of this construct in lower-stakes situations ([Forsyth et al., 2020a](#)). In these formative contexts, it is possible to integrate more naturalistic conversations to elicit fine-grained evidence of student KSAs. Specifically, we can capture information about students with natural language conversations in the context of conversation-based assessments.

Conversation-based assessments

Conversation-based assessments (CBAs) result from the need to gather additional evidence of the decisions students make on performance-based assessments such as computer-based scenarios, simulations, or other forms of technology rich tasks ([Zapata-Rivera et al., 2013](#)). CBAs involve the use of artificial agents that are designed to engage students in conversations about the nature of their decisions in the assessment environment ([Zapata-Rivera et al., 2015a,b](#)). CBAs build on research results in areas such as dialog systems ([Adamson et al., 2014](#); [Graesser et al., 2001](#); [Millis et al., 2011](#)), artificial virtual agents ([Johnson and Lester, 2016](#)), and technology enhanced assessments ([Bennett et al., 2007](#); [Clarke-Midura et al., 2011](#); [Quellmalz et al., 2011](#)).

CBAs have been used to gather evidence of middle-school students' KSAs that may be difficult to obtain using traditional assessment approaches. CBAs are useful for eliciting additional explanatory evidence for a variety of KSAs. For example, guided conversations or dialogs with artificial agents can be used to assess 21st century skills such as argumentation, science inquiry, critical thinking, and collaborative problem solving. CBAs can also be used to assess communication skills (e.g., English language skills such as reading, writing, listening, and speaking). CBAs can be used in conjunction with other item types in a variety of assessment situations, including simulation-based, scenario-based, and game-based assessments (So et al., 2015; Song et al., 2014; Zapata-Rivera et al., 2014).

CBAs can be used to gather evidence of students' explanations about decisions that they make in technology rich environments (e.g., Why were particular pieces of data used to support a prediction? Why was a particular combination of parameters selected to run the simulation?). Because they leverage an adaptive dialogic framework similar to the one previously described for the Auto-Tutor system` (see Zapata-Rivera et al., 2015a), CBAs can provide students with multiple opportunities to elaborate on their responses depending on the quality and content of those responses (e.g., rephrasing the question and asking the student to provide additional information after receiving an initial response that is incomplete or irrelevant), thus presenting students with a more dynamic, personalized assessment experience.

Conversations with artificial agents can be considered as an interactive constructed response item type that allows for dynamic exchanges of information about target constructs. CBAs can be used to create interactive, yet open assessment situations for students to demonstrate what they know or can do. Compared to other item types, CBAs are characterized by high levels of openness and interactivity, including the potential for students to take different conversational paths, or to introduce branching at key points based on students' interactions within the environment. By attaching CBA to other item types (e.g., responses to a MC question or interactions within a simulation), it is possible to create enhanced or hybrid items that provide opportunities for exploring the cognitive processes and strategies the student is demonstrating (Jackson and Zapata-Rivera, 2017). Taken together, these features make CBAs a suitable environment for exploring questions around personalized assessment and learning.

In the next section, we describe aspects of the development of CBA prototypes as well as some results of evaluation efforts.

CBA development and scoring

The development process of CBAs is iterative and follows the principles of evidence-centered design (Mislevy et al., 2003; Yin and Mislevy, 2022), a principled approach to assessment design that focuses on the evidence-based argument structure connecting tasks to the target constructs.

The CBA development process takes into account the context(s) where intended conversations can take place, the construct(s) that are intended to be measured, and conversation structure and adaptivity issues such as conversation openings, main questions, conversation moves/patterns used to follow-up on student responses, role of artificial agents (e.g., teacher and/or peer, who asks questions and provides feedback, type of feedback provided), initial scoring rules (e.g., specifying responses or response sequences that earn full or partial credit), task sequencing, and conversation closings. These components are represented in a conversation space diagram. This structure is converted into (a) a conversation script used by a dialog engine to manage the conversation (Zapata-Rivera et al., 2015a) and (b) scoring rules associated with each path in the conversation space diagram.

A key aspect of the CBA development process is to carefully design the interactions (e.g., questions and feedback) to address the constructs and reduce the complexity of the conversation space. Response categories are used to deal with a variety of student responses and to respond following a limited number of pre-defined paths corresponding to those categories. These branching conversations are refined based on expert input and student data. Automated testing tools and crowdsourcing have been used to speed up the development process of conversation elements such as regular expressions and language models (Zapata-Rivera et al., 2015).

Scoring approaches used in the development of CBAs include: (a) assigning initial path-based scores based on expert judgment; (b) revising initial scoring rules based on qualitative analyses of student responses; and (c) scoring produced automatically by the dialog engine based on natural language processing techniques such as LSA and regular expressions, which match student responses to expected correct responses throughout the conversation.

Performance on CBA prototypes is determined by aggregating evidence from different types of items including multiple-choice questions, constructed responses and conversations with artificial agents into a comprehensive total score. In situations where multiple constructs are measured, we compute a single measure of performance per each construct.

CBAs are evaluated by conducting small-scale studies such as usability studies, cognitive labs, and path validation studies (used to explore how the system responds to student utterances), and large-scale studies examining issues like the effects of different types of questions or characters on CBA performance, the types of emotions participants experience in these environments, how the type of evidence gathered through CBAs compares to external measures, and how subgroups of students (e.g., English language learners) interact with CBAs. CBAs provide a rich amount of data and offer researchers with an opportunity to explore issues surrounding the design and evaluation of advanced measurement modeling and scoring approaches that are appropriate for these kinds of tasks.

CBA examples

Several CBA prototypes have been designed to assess skills such as science inquiry (Zapata-Rivera et al., 2014; Liu et al., 2016), mathematical reasoning (Cayton-Hodges et al., 2015), formulating and justifying arguments (Song et al., 2014), collaborative problem solving (Hao et al., 2015); and English language skills for English learners (So et al., 2015; Evanini et al., 2014). An additional prototype assessed both English language and mathematics skills in an integrated environment (ELLA-Math; Lopez et al., 2021). The development of these prototypes involved an interdisciplinary team with expertise in areas such as dialog systems, assessment development, cognitive science, human-computer interaction, game design, and psychometrics.

These CBA prototypes make use of a variety of question types (e.g., multiple-choice, inline choice, constructed response) in addition to conversations with artificial agents. The type of evidence gathered using CBAs compared to other types of tasks has been analyzed (Forsyth et al., 2018; Liu et al., 2016; Zapata-Rivera et al., 2016). Results of these studies have shown that English language learners interacting with CBAs produce relevant responses comparable to those that are generated in human-to-human communication (Forsyth et al., 2018) and that CBAs can be used to collect additional evidence of targeted constructs (e.g., additional dimensions of the constructs), compared to the evidence collected using other types of tasks (Liu et al., 2016; Zapata-Rivera et al., 2016). Thus, it is possible to develop isomorphic CBA prototypes (i.e., tasks with similar underlying structure, but with different surface features in terms of topic and content) to collect similar types of evidence in different scenario contexts with the same underlying skills, which has implications improving scalability by facilitating repurpose of conversation elements such as dialog scripts and other components of the environment (Liu et al., 2016; Zapata-Rivera et al., 2015a).

CBA prototypes have been implemented using various levels of fidelity, ranging from text-based graphical interfaces with static images representing the agents to interactive immersive environments with animated, embodied agents. Fig. 1 shows a screenshot of the ELLA-Math prototype, which was designed to assess English language skills and math content knowledge of middle school English learners. In this prototype, the student works with two artificial fellow students on solving small-group activities. Students take turns reading the questions and engaging in conversations aimed at supporting understanding of the questions and scaffolding the math concepts required to solve each task. The prototype makes use of open response and multiple-choice tasks. Conversations are triggered based on what the student types (open response task) or selects (multiple-choice tasks). Multiple-choice tasks include choices such as “I do not understand ...” which triggers conversations aimed at clarifying the question. Response categories handled by CBA prototypes usually include correct, partially correct, incorrect, irrelevant, meta-communicative (e.g., “What did you say?” “Please repeat”), meta-cognitive (e.g., “I have no idea,” “I forgot”), and blank (i.e., no response provided) responses (Zapata-Rivera et al., 2015a).



Fig. 1 A screenshot of ELLA-Math.

Lopez et al. (2021) report on a study evaluating the accuracy of the ELLA-Math prototype in correctly identifying and classifying student responses generated by middle school students. The system was found highly accurate in interpreting many of the sentences provided by students. Also, several areas for improvement were identified (e.g., handling additional types of responses and improving the robustness of the system to accommodate various types of misspellings and typographical errors). Information collected can be used to refine the internal components of the system (e.g., regular expressions and dialog scripts) following an iterative development and evaluation approach.

Zapata-Rivera et al. (2018) report on the results of a study exploring the different kinds of unexpected responses found in data collected using CBA prototypes. Unexpected responses were identified and classified into several categories by the researchers. These categories included: (a) confused (response made little sense or lacked an explanation); (b) frustrated (the student seemed annoyed by his or her interaction with the characters); (c) repetitive (the student repeated the same answers to different questions); (d) unmotivated (the student did not care to answer or provided short answers); (e) irrelevant (the student provided off topic responses); (f) asking for help (the student asked for additional information or materials); (g) gaming the system (the student tried to test the capabilities of the system); (h) attempting to communicate (the student did not answer the question and referred to previous interactions); (i) different languages (the student responded in a language other than English); and (j) giving up (the student mentioned that they want to quit). Identification of these categories may reflect students with different levels of motivation and shows the potential for the assessment system to identify these situations and react to them in a way that addresses students' frustrations and provides them the necessary supports to persist through the task. In the next section, we describe an approach for designing and implementing personalized, caring assessments that offer potential strategies for dealing with some of these situations.

CBAs and caring assessments

Personalized, "caring" assessment systems are designed to consider additional information about the student (e.g., cognitive, meta-cognitive and non-cognitive aspects) and the learning context to create assessment situations that students find engaging, and which can be used to collect valid and reliable evidence (Zapata-Rivera, 2017). Information about the learning context including socio-cultural aspects can be useful in implementing personalization approaches that positively influence student learning and can be adopted within educational ecosystems (see Fig. 2).

Different levels of "caring" (i.e., student support) can be implemented depending on the purposes of the assessment. Two common use cases include dealing with (a) student motivation issues and (b) prior knowledge or opportunity to learn. When dealing with student motivation aspects, assessments could implement student-centered approaches to deal with instances of low motivation, such as providing information to the student about why the test is administered and how the data will be used (e.g., as a learning opportunity that will help the student identify strengths and areas for improvement), incorporating motivation-oriented feedback, offering other assessment options for students to choose, and taking into account the level of motivation when interpreting the evidence collected during the assessment. Similarly, in the case where a student does not know or did not have the opportunity to learn a particular concept, the assessment could offer relevant learning materials or scaffolded activities before asking additional questions about the concept. The assessment can then take this information about the students' prior knowledge and experiences into account when interpreting the data collected (Zapata-Rivera et al., 2020).

This personalized, caring assessment approach is consistent with a broader educational emphasis on the use of Universal Design for Learning (UDL; CAST, 2018) as a framework for improving and optimizing learning for all students. UDL is a framework for the design of learning materials and environments, grounded in the idea that supporting diverse learners where they are should include careful consideration and design of learning opportunities such that they reduce barriers to access and engagement, providing learners with multiple means of (1) engagement, (2) representation, and (3) action and expression. Our framework for caring assessments introduces the possibility to incorporate supports consistent with the UDL guidelines into an adaptive assessment experience (e.g., providing multiple ways to make a response), in order to optimize the assessment experience for students, while simultaneously permitting collection of more precise evidence of students' KSAs.

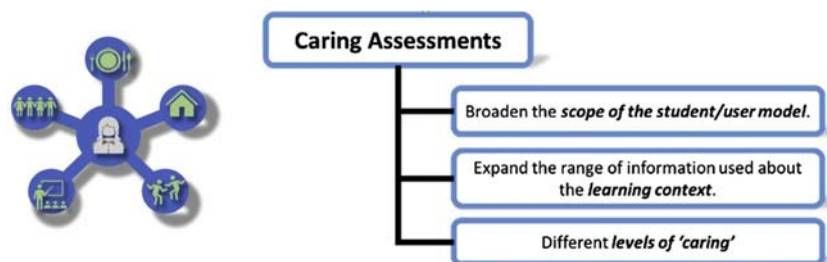


Fig. 2 Personalized, "caring" assessments.

One approach to meeting students where they are and providing support are cognitive bandwidth recovery strategies (Verschelden, 2017). Cognitive bandwidth recovery strategies are designed to help students minimize the negative impacts of cognitive resources being dedicated to dealing with the effects of poverty, racism, and social marginalization, which allows students to “recover the cognitive bandwidth” required to successfully perform learning and assessment tasks. These strategies promote a growth mindset (Dweck, 2006) and self-efficacy and have potential for improving the chances of minority students to demonstrate what they know or can do. These strategies include teachers asking students who already took a particular assessment to share their experiences with other students who will take that same assessment; asking students to write about particular values or topics that decrease identity threat, increase belongingness, or promote growth mindset; co-designing assessment tasks and feedback with teachers and students taking into account aspects such as inclusion and growth mindset; taking breaks during assessments to engage students in activities that lower anxiety; providing additional time for students who start assessments late; and taking the assessment in a familiar environment where students feel safe. Such strategies could be implemented within a caring assessment framework to provide adaptive supports for a variety of social-emotional characteristics.

CBAs and student emotions

The importance of emotions during the learning process has been well documented (e.g., Baker et al., 2010; D’Mello et al., 2006; Graesser et al., 2007; Pekrun et al., 2011). In particular, research on tutorial conversations (human-human, computer-human) has documented that students experience a variety of emotions and that the experience of those emotions impact learning outcomes (e.g., Baker et al., 2010; Forbes-Riley and Litman, 2011; Graesser et al., 2007). The learning-centered emotions (anxiety, boredom, confusion, curiosity, delight, engagement/flow, frustration, happiness, surprise) were first identified during interactions with AutoTutor and were later found to be prominent in short learning sessions (30–90 min) across a variety of domains and interaction styles (D’Mello, 2013). Conversation-based learning environments in the last decade have begun detecting student emotions in real time and deploying emotion-sensitive support to students (D’Mello et al., 2011; Forbes-Riley and Litman, 2011). These emotion-sensitive interventions have shown to be particularly beneficial for students with lower prior knowledge.

Emotions during testing have received relatively less attention (see Pekrun et al., 2011, 2004, and Spangler et al., 2002, for exceptions). Early research focused primarily on test anxiety (Zeidner, 1998) and while more recent research has expanded the emotions investigated (anger, anxiety, enjoyment, hope, hopelessness, pride; Pekrun, 2006), the context has remained limited to the context of tests with traditional item formats (e.g., multiple-choice items). It is unclear how the testing context will impact the emotions experienced and how those emotions relate to test outcomes, but it is important to understand the relationship between test context and emotions to provide both support to students and accurate test outcome reporting. Thus, the creation of CBAs (and other non-traditional assessment formats) has necessitated a new round of investigations into student emotions during assessment.

A series of studies using retrospective affect judgment tasks were conducted to investigate the emotional experiences of students during CBAs (Lehman and Zapata-Rivera, 2018b). The first goal of these investigations was to identify those emotions that are most relevant to this specific testing context. Initially, a large set of emotions was explored because CBAs are a combination of conversation-based learning environments (evoking learning-centered emotions; D’Mello, 2013) and assessments (evoking achievement emotions; Pekrun et al., 2011). Ekman’s (1973) basic emotions (anger, contempt, disgust, fear, joy, sadness, surprise) were also included in this initial investigation. A consistent set of prominent emotions were found for interactions with CBAs of science inquiry and math argumentation skills: boredom, confusion, curiosity, delight, engagement, enjoyment, frustration, hope, pride. However, the findings revealed no relationship between these prominent emotions and performance on the CBAs.

A follow-up study was conducted to directly address the lack of relationship between emotion and performance (Lehman and Zapata-Rivera, 2018a). The prior studies only allowed students to report whether an emotion did or did not occur at a given point in the CBA. However, research on emotions during learning has shown that the relationship between emotions and learning outcomes may not be as simple as the presence or absence of an emotion during learning (D’Mello and Graesser, 2012; D’Mello et al., 2007). For example, research on confusion during learning has found that confusion resolution and duration impact the confusion-learning relationship (D’Mello and Graesser, 2012). In this vein, the third study allowed students to report the intensity of their experienced emotion (low, medium, high). The findings revealed that curiosity, engagement, happiness, and pride were all positively correlated with performance, but the relationship between negative emotions and performance varied based on emotion intensity (see Fig. 3). For boredom, confusion, and frustration the results showed that they were positively correlated with performance at low intensity, not correlated at medium intensity, and negatively correlated at high intensity. These opposing relationships based on emotion intensity add an interesting new twist to understanding students’ emotional experience during testing. Performance on conversations was determined by applying the Expectation-Misconception framework described above (Lehman and Zapata-Rivera, 2018a).

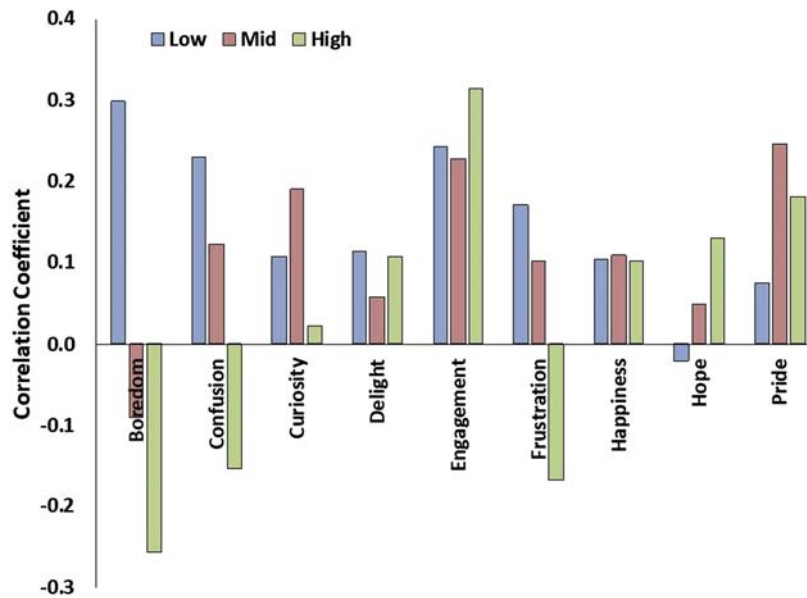


Fig. 3 Relationship between emotion intensity and conversation-based assessment performance.

Documenting the emotional experience is an important first step, but it is critical to determine how these findings can be applied to improve both the students' experience and the validity of test outcomes. In the case of CBAs, the structure of conversations can be evaluated through the lens of student emotions. One motivation to develop CBAs was to increase student engagement. Thus, the different paths in the conversation structure can be evaluated as to the extent that students maintain the productive state of engagement or resolve less productive states and return to an engaged state. A follow-up question, for example, that frequently elicits frustration may need to be re-designed to ensure that it provides students with the best opportunity to show what they know and can do.

Real-time, emotion-sensitive adaptations and supports are another way that information about student emotions can be leveraged to improve student experience and the validity of test outcomes. As mentioned previously, this type of support is already provided in some conversation-based learning environments (D'Mello et al., 2011; Forbes-Riley and Litman, 2011). Similar efforts targeting test-taking motivation during testing have also shown positive results in terms of both decreasing disengaged responding and increasing test reliability (Wise et al., 2006). Thus, it is not out of the question to provide this type of support in the assessment context. However, there are still potential issues with providing emotion-sensitive adaptations during assessment that must be explored. One issue is that the ability to reliably detect student emotions is not perfect, and in high stakes testing contexts, there could be serious consequences to incorrectly delivering (or not delivering) emotion-sensitive support to a student. Another important issue is fairness. Any kind of personalized, adaptive support will inherently mean that students will not have a standardized experience with the test. While this lack of standardization may have a positive impact for most students, it is vital to ensure that the differential emotion-sensitive support that students receive does not disadvantage or advantage any particular group of students.

Beyond emotions, other student-level characteristics like self-efficacy could be used to provide adaptive support to students within a caring assessment framework; in the following section, we review research examining the role of various individual differences in students' interactions with CBAs.

CBAs and individual differences

There has been long-standing interest in understanding the role of individual differences in learning and instruction (e.g., Cronbach and Snow, 1977; Jonassen and Grabowski, 1993; Snow, 1986). In general, such work has been concerned with investigating how individual-level variation in characteristics such as cognitive abilities, personality, prior knowledge, or other attributes may interact with features of the instructional tasks learners are asked to complete, resulting in different learning outcomes (Jonassen and Grabowski, 1993). Individual differences have been conceptualized in terms of both academic skills (e.g., reading or math ability) and non-academic personal qualities (Duckworth and Yeager, 2015), also referred to as non-cognitive or social-emotional competencies (e.g., persistence, self-efficacy, growth mindset, personality, etc.). Recent research has examined the utility of different approaches to assessing or measuring such individual differences via self-report surveys, situational judgment tasks, or other more direct measures

(e.g., Abrahams et al., 2019; Duckworth and Yeager, 2015; Kyllonen, 2016; Lipnevich et al., 2013; Soland et al., 2019; West et al., 2018), and has documented the individual differences that most strongly correlate with or predict academic achievement in various domains (e.g., Hattie, 2009; Kyllonen, 2012; Nofle and Robins, 2007; Richardson et al., 2012; Schneider and Preckel, 2017; West et al., 2016). Individual differences in self-efficacy and self-regulation skills appear to be particularly important, showing moderate correlations with college GPA (Richardson et al., 2012).

As with any learning or assessment experience, when students engage with CBA, they bring with them different constellations of resources and capabilities and will differ in their potential to experience barriers. For example, a meta-analysis (Schroeder et al., 2013) reported that artificial agent features such as communication modality (text-only or voice and text vs. narration-only), agent voice (no voice vs. human or computer-generated voice), appearance (non-humanoid or humanoid vs. human), animation (animated vs. static), the content domain (STEM vs. humanities), the student's grade level (K-12 vs. postsecondary), and degree of prior knowledge (moderate vs. low or high) all have a significant impact on learning from virtual agents; however, virtual agents' potential impact on assessment performance is less clear. It is thus important to understand how students' individual resources and capabilities may relate to and potentially impact their perceptions of and performance within CBA tasks.

A series of studies was conducted to investigate the role of individual differences in the context of a CBA measuring science inquiry skills (see Liu et al., 2016). The first study examined whether students' perceptions of virtual agents in the CBA (a scientist and a peer) related to CBA performance (Sparks et al., 2018). Specifically, this work evaluated whether experimental manipulations of features of the peer agent (i.e., high vs. low topic knowledge) and the type of question posed by the peer agent (i.e., asking the student to compare their observations to the peer's and choose the stronger observation, vs. asking whether the student agrees or disagrees with the peer's choice of the stronger observation) had an apparent impact on students' perceptions of both agents (measured by self-report ratings) or on task performance (measured by scores aggregated from student responses to selected response, constructed response, and conversation-based items included in the task, as described previously). Results indicated that question type had greater impact on students' perceptions than the virtual peer's topic knowledge level, such that students rated the virtual scientist as less knowledgeable and authoritative in the agree/disagree condition than in the compare condition. Agent perceptions also correlated with students' task performance, such that students who performed better on the CBA were more likely to give higher ratings to the virtual scientist, but lower ratings to the virtual peer. Students who earned "A" grades in science, or had higher performance on a standardized science assessment, were also less likely to rate the virtual peer as knowledgeable.

A second study more directly examined the individual differences that correlated with and predicted CBA performance by asking students to complete a variety of self-report measures in addition to completing the science inquiry CBA (Sparks et al., 2019). Analyses indicated that individual differences in social-emotional variables like persistence (i.e., grit; Duckworth et al., 2007), growth mindset (Dweck, 2006), cognitive flexibility (Martin and Rubin, 1995), achievement goal orientation (Elliot and Murayama, 2008), self-efficacy (Pintrich and De Groot, 1990), and self-reported interest in science (Maltese and Tai, 2010) were positively correlated with CBA performance, and test anxiety (Pintrich and De Groot, 1990) was negatively correlated. Regression analyses revealed that after controlling for demographics (urban vs. rural school, grade level) and prior achievement (standardized test scores), individual differences in cognitive flexibility, growth mindset, science self-efficacy, and test anxiety were also significant predictors of performance on the science CBA (test anxiety had a negative relationship to scores).

In a third study, those four individual difference measures found to significantly predict CBA scores were used as inputs for a hierarchical cluster analysis, which yielded four distinct profiles of students with different patterns of background variables and performance (Sparks et al., 2020). While the study found one high-performing cluster with high prior achievement, very low-test anxiety, and high levels of growth mindset, self-efficacy, and cognitive flexibility, three low-performing clusters were also observed, with potentially different contributing factors. For example, one low-performing cluster had low prior achievement as well as lower levels of growth mindset (i.e., highly fixed mindset), while another low-performing cluster had average prior achievement but relatively low cognitive flexibility, self-regulation, persistence, intrinsic value, and science interest. Pending further validation, these profiles could be leveraged to streamline the design and delivery of adaptations to CBA tasks to address and ameliorate different reasons for poor performance (e.g., encouraging a growth mindset vs. strategies to manage lack of self-regulation, perseverance, or interest), in support of the larger goals of the caring assessment framework previously described (Zapata-Rivera, 2020). Additional research is needed to understand precisely how CBAs can be effectively tailored to provide an engaging and valid experience for students within particular profiles by using information about the range of characteristics students bring to the assessment.

As illustrated above, CBAs provide fertile ground for conducting innovative research in support of designing and developing personalized learning and assessment systems, suggesting many fruitful possible future directions for further investigations.

Moving forward

Several experts were invited to provide their insights on issues regarding the use of dialog systems in assessment and learning. The following experts accepted our invitation to provide comment: Arthur C. Graesser, The University of Memphis; Sandra Katz, University of Pittsburgh; Lewis Johnson, Alelo; Andrew Olney, The University of Memphis; and Carolyn Rosé, Carnegie Mellon University.

Experts identified the main challenges for implementing assessments that make use of dialog systems. These challenges include:

- Developing authoring tools that are easy to use, linking discourse to external representations (pictures, diagrams, simulations), supporting the implementation of regular expressions that usually need to be hand-crafted, generating good hints, and designing approaches for providing feedback to conversation participants while considering assessment goals (Graesser, 2021).
- Improving natural language understanding to assess student knowledge and conceptual understanding at various levels. For example, understanding responses to deep-reasoning questions (e.g., why and how questions) and supporting student engagement with dialog-based interactions with tutoring systems (Katz, 2021). Also, assessing the forgiving tendency of dialog can be challenging (e.g., the listener usually fills in missing details). Dialog-based assessment needs to be adaptive in diagnosing student understanding (Olney, 2021).
- Validating these assessments with human judgments, which have a subjective element, and finding appropriate validation methodologies that work in the lab and translate well into educational practice (Johnson, 2021).
- Considering differences in how students treat computer agents compared to humans (Rosé and Torrey, 2005). However, computer-based dialog could be compared to using a structured problem-solving interface and thus it should be as effective in assessing knowledge and cognitive skills (Rosé, 2021).

Experts were asked to elaborate on areas for future work on CBAs. In addition to addressing the challenges mentioned above, the experts mentioned the following directions for future research:

- Exploring the use of Deep learning models known as Transformers in natural-language understanding, and investigating trade-offs typically considered in reinforcement learning to precisely measure ability while taking into account time constraints (Olney, 2021).
- Continuing to investigate the generalizability of trained models in automated assessment of conversational processes (Rosé, 2021).
- Integrating agent systems with smart phones and developing conversational guides at a meta-level (i.e., outside of the content being assessed) (Graesser, 2021).
- Measuring a variety of communicative competencies in dialog including persuasiveness and empathy. For example, communicative competences such as those relevant to the community of health workers (e.g., communicate using language at the level of the individual; communicate with individuals in a non-judgmental and appropriate manner; practice active, reflective listening and remain solution-oriented in all interactions; recognize and respond to non-verbal behavior; ask open-ended questions to solicit information from the individual; utilize affirming statements to provide positive reinforcement) (Johnson, 2021).
- Exploring the potential benefits of integrating Open Learner Modeling (OLM) technology with dialog-based tutoring systems (Katz et al., 2021). Opening the student model in a dialog-based tutoring system raises some interesting challenges such as maintaining the illusion that students are interacting with a conversational partner while exposing the system's adaptive processes, and finding approaches to share dialog-based system's evidence with students (e.g., when evidence includes lengthy explanations) (Katz, 2021).

Finally, we put together a list of resources that these experts recommended for graduate students and other researchers who are new to this area of research, in order to inspire scholars who are just being introduced to these topics to pursue related lines of work. These resources can be found in [Appendix A](#).

Building on the ideas mentioned above, we elaborate on additional challenges and future work in this area:

- As mentioned above, additional tools to support the development of CBAs that support naturalistic communication is one of the areas for future work. As more technology rich tasks are integrated into assessment and learning systems, the need for rapid development of conversations that can be used to gather evidence of constructs (e.g., using particular conversation patterns) should be considered. These authoring tools can benefit from domain, assessment and dialog systems experts as well as information collected through crowdsourcing approaches (e.g., Ramanarayanan et al., 2016).
- Continue improving NLP capabilities used in the implementation of CBAs (e.g., improving the accuracy of speech and dialog engines), developing and evaluating measurement models to score CBAs, and developing procedures for validating CBAs under different conditions (e.g., formative and summative contexts), for different subgroups of learners (e.g., English language learners, and underserved groups), and to measure additional constructs (e.g., communicative competencies).
- Additional work on implementing and evaluating cognitive bandwidth recovery strategies (Verschelden, 2017) may result in insights for improving assessment and learning systems. Additional challenges in the area of caring assessments include data privacy and security issues, the development of new communication channels and feedback loops with teachers and students, and supporting appropriate use, trust, and adoption of these assessment systems and the information they provide. Results from research in areas such as OLM and interactive score reporting can inform work on communicating assessment results and support trust in dialog systems (Zapata-Rivera, 2020).

- Future work on emotions and CBAs should focus on how to (1) utilize student emotions to evaluate and refine the design of CBAs and (2) provide on-demand support for students during the completion of CBAs. The current findings can be replicated during the development process of future CBAs to identify components of the CBA that is or is not supporting student engagement. Providing on-demand support, on the other hand, is a more complex matter as it requires both the reliable detection of emotions and the deployment of appropriate supports. The support provided to students may need to be further personalized based on student individual differences, as these characteristics are likely to impact the effectiveness of interventions. There is also the issue of fairness, as mentioned previously, so it is important to understand the impact of any support provided on the validity of test outcomes.
- With respect to individual differences in CBA, future research and development could take several different directions. One direction would be to focus on integrating CBAs within instructional programs designed to measure and develop students' social-emotional learning (SEL) competencies to further explore linkages between SEL and CBA performance (e.g., adapting CBAs based on students pre-task levels of variables like self-efficacy and growth mindset and evaluating whether those adaptations have any impact on the growth of SEL competencies when they are measured again later in the school year; developing and evaluating CBAs for learning and assessing SEL competencies, etc.). Another direction would be to expand efforts to align the structure of CBA within a caring assessment framework to the principles of the UDL framework, developing affordances of CBAs to better support the range of universal designs that could be implemented and evaluating their effects on learning and performance (e.g., as compared to the highly adaptive, dynamic context of human tutoring, or to more traditional, static assessments), thus positioning CBAs as a platform for delivering assessments that align to UDL requirements. Both strands of work would yield foundational and applied insights to support the design of adaptive, personalized assessment and learning systems.

Summary

In this article, we described foundational and applied work in the area of CBA. Through the analysis of unexpected responses found in data collected with CBA prototypes, we introduced the concept of personalized, “caring” assessments and elaborated on the implications for CBAs that consider a broader conception of the student and the context in which assessment and learning take place.

Interactive simulated environments with agents can help students learn as well as assess their learning. We clearly show the learning gains resulting from this approach to creating these types of environments for learning (see Graesser, 2016); this article extends that work to the realm of assessment with conversation-based assessment environments that elicit evidence beyond what might be possible with traditional test items.

We reviewed preliminary findings from two current areas of research on factors that may influence student performance in CBAs: student emotions and individual differences. The research on emotions has a long history of showing importance to learning and this work has been extended to the area of emotions in assessments. Specifically, emotions may affect student performance on CBAs adding construct-irrelevant variance that could obstruct measurement of content knowledge. Furthermore, emotions may dictate whether a student can display their knowledge and thus should be considered during design and personalization. The work on individual differences further highlights the need to take into account how different non-cognitive or social-emotional competencies, or other background characteristics not traditionally considered in assessment, may influence the types of evidence of students' KSAs gathered using CBAs. Thus, it is imperative that we continue to explore these variables and how they interact to affect engagement and performance in interactive simulated environments.

Furthermore, we elicited information from several experts in the area of dialog systems who elaborated on gaps in the current literature including creating authoring tools that can easily be used by teachers, integrating new deep learning models as well as exploring open-learning models, finding more sources of validation (including human ratings), discovering ways to scale up such environments by hosting them on multiple platforms, including measurement of additional competencies, and more. This feedback helps solidify the importance of taking into account a broader view of the student that includes non-cognitive factors and contextual information as part of the student model, and using this information to tune the learning and assessment experience so it is more personalized to where students are in terms of their current level of KSAs.

In sum, the goal of this article is to showcase the possibilities for assessment environments with agents that engage students in personalized conversations about target KSAs. The opportunities afforded by such environments could provide valuable feedback to stakeholders, improve measurement of target constructs, and improve students' experience with these systems. We expect this article will provide researchers with valuable information to begin (or continue) exploring the potential of using dynamic conversations within learning and assessment systems.

Appendix A. Table of resources

Resource	Description
Adamson, D., Bharadwaj, A., Singh, A., Ashe, C., Yaron, D., Rosé, C.P., 2014. Predicting Student Learning From Conversational Cues, Proceedings of Intelligent Tutoring Systems Volume 8474 of the Series Lecture Notes in Computer Science, pp. 220–229.	Describes how low-level textual features, models based on sequential discourse act labels and evidence from initial stages of collaborations can predict student success.
Chi, M., 2009. Active–constructive–interactive: a conceptual framework for differentiating learning activities. <i>Topics Cogn. Sci.</i> 1, 73–115.	Provides a framework to differentiate active, constructive, and interactive in terms of observable overt activities and underlying learning processes. This framework is used to analyze the effectiveness of interactive and constructive tasks and suggest coding schemes for activities.
Forsyth, C.M., Graesser, A.C., Millis, K., 2020. Predicting learning in a multi-component serious game. <i>Technol. Knowl. Learn.</i> 25, 251–277.	Describes a methodology for interpreting data on a fine-grained level resulting from student interaction with two or more artificial agents in a learning environment.
Graesser, A.C., Hu, X., Rus, V., Cai, Z., 2020. AutoTutor and other conversation-based learning and assessment environments. In: Rupp, A., Yan, D., Foltz, P. (Eds.), <i>Handbook of Automated Scoring: Theory Into Practice</i> . New York: CRC Press/Taylor and Francis.	Provides an overview of AutoTutor and CBAs. The significance includes the original dialogic structure of AutoTutor focused on learning and how this structure is augmented and applied for assessment in CBAs.
Graesser, A.C., 2016. Conversations with AutoTutor help students learn. <i>Int. J. Artif. Intell. Educ.</i> 26, 124–132.	Provides an overview of evidence gleaned from dozens of studies showing learning gains resulting from AutoTutor over a decade and multiple contexts.
Graesser, A.C., 2020. Learning science principles and technologies with agents that promote deep learning. In: Feldman, R.S. (Ed.), <i>Learning Science: Theory, Research, and Practice</i> , New York: McGraw-Hill, pp. 2–33.	Describes the theoretically-grounded principles that are implemented in a system for learning that helps students form a complete mental model (vs. Shallow memorization) of content with a focus on conversation-based environments.
Howley, I., Rosé, C.P., 2016. Toward careful practices for automated linguistic analysis of group learning. <i>J. Learn. Anal.</i> 3(3), pp. 239–262, http://doi.org/10.18608/jla.2016.33.12	Discusses bridging how linguistics views discourse and the actual operationalization to create frameworks of human discourse in collaborative environments based on theories from the learning sciences.
Jackson, G.T., Graesser, A.C., 2006. Applications of human tutorial dialog in AutoTutor: an intelligent tutoring system. <i>Revista Signos</i> , 39, 31–48.	Discusses the role of interactivity and the interactivity hypothesis in AutoTutor. The interactivity hypothesis suggests that any interaction aids learning above passive learning.
Johnson, W.L., 2019. Data-driven development and evaluation of Enskill English. <i>Int. J. Artif. Intell. Educ.</i> 29, 425–457, https://doi.org/10.1007/s40593-019-00182-2	Discusses a bottom-up, machine learning approach to creating dialogic systems.
Katz, S., Albacete, P., Chounta, I.A., Jordan, P., McLaren, B.M., Zapata-Rivera, D., 2021. Linking dialog with student modeling to create an adaptive tutoring system for conceptual physics. <i>Int. J. Artif. Intell. Educ.</i> 1–49.	Describes work on implementing and evaluating student modeling in a tutorial dialog system. The authors discuss future work to improve the performance of the system and to integrate tutorial dialog systems in the classroom.
McTear, M.F., Callejas, Z., Griol, D., 2016. <i>The Conversational Interface</i> , Cham: Springer, vol. 6, No. 94, p. 102.	Provides an overview of types of interfaces on various devices where humans have natural language conversations.
Naik, A., Rosé, C.P., 2020. Toward Open Domain Event Trigger Identification Using Adversarial Domain Adaptation, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.	Explains a method for creating non-domain specific natural language conversations with a completely unsupervised approach (data enter black box models).
Naik, A., Lehman, J.F., Rosé, C.P., 2021. Adapting Event Extractors to Medical Data: Bridging the Covariate Shift. Proceedings of the 2021 Conference of the European Chapter of the Association for Computational Linguistics.	Explains an approach to analyzing conversations by comparing several different unsupervised models.
Olney, A.M., 2018. The unreasonable effectiveness of AutoTutor. In: Millis, K.K., Long, D., Magliano, J., Wiemer, K. (Eds.), <i>Deep Comprehension: Multi-disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension</i> . New York: Routledge.	Discusses the success of AutoTutor producing learning gains comparable to expert raters and resulting implications. There is a particular focus on the cost-effectiveness and relative simplicity of this system.

(Continued)

Resource	Description
Olney, A.M., Graesser, A.C., Person, N.K., 2010. Tutorial dialog in natural language. In: Nkambou, R., Bourdeau, J., Mizoguchi, R. (Eds.), <i>Advances in Intelligent Tutoring Systems, Studies in Computational Intelligence</i> , Berlin: Springer-Verlag, vol. 308, pp. 181–206.	Discusses the importance of basing natural language tutorial dialog conversations on expert human tutors rather than novices as investigated in an ITS referred to as GURU.
Paladines, J., Ramírez, J., 2020. A systematic literature review of intelligent tutoring systems with dialog in natural language. <i>IEEE Access</i> , 8, 164246–164267, https://doi.org/10.1109/ACCESS.2020.3021383 .	Provides an overview of systems across domains that teach students with natural language conversations.
Rosé, C.P., Torrey, C., 2005. Interactivity Versus Expectation: Eliciting Learning Oriented Behavior With Tutorial Dialog Systems, <i>Proceedings of 10th IFIP TC13 International Conference on Human-Computer Interaction—Interact 2005, Lecture Notes in Computer Science</i> , vol. 3585, pp. 323–336.	Explains an empirical investigation along with recommendations to attempt to help students provide similar elaborated conversations to conversational agents as with humans.
Rosé, C.P., Howley, I., Wen, M., Yang, D., Ferschke, O., 2017. Assessment of discussion in learning contexts, invited chapter. In: von Davier, A., Zhu, M., Kyllonen, P. (Eds.), <i>Innovative Assessment of Collaboration</i> , New York, NY: Springer Verlag.	Discusses the difficulties and methodologies of assessing conversational data resulting from computerized environments
VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. <i>Educ. Psychol.</i> 46(4), 197–221.	Provides a meta-analysis of research evaluating the effectiveness of intelligent tutoring systems compared to other methods of instruction.
Wen, M., Yang, D., Rosé, D., 2014. Linguistic Reflections of Student Engagement in Massive Open Online Courses, <i>Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media</i> , pp. 525–534.	Discussion of discoveries based on detectors of student affect displayed in conversation in MOOC forums.

References

- Abrahams, L., Pancorbo, G., Santos, D., John, O.P., Primi, R., Kyllonen, P., De Fruyt, F., 2019. Social-emotional skill assessment in children and adolescents: advances and challenges in personality, clinical, and education contexts. *Psychol. Assess.* 31, 460–473.
- Adamson, D., Dyke, G., Jang, H.J., Rosé, C.P., 2014. Towards an agile approach to adapting dynamic collaboration support to student needs. *Int. J. AI Educ.* 24 (1), 91–121.
- Andrews-Todd, J., Forsyth, C.M., 2020. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Comput. Hum. Behav.* <https://doi.org/10.1016/j.chb.2018.10.025>.
- Andrews-Todd, J., Forsyth, C.M., 2022. Assessment of Collaborative Problem Solving Skills, *International Encyclopedia of Education*, 4th ed. Elsevier (in press).
- Atkinson, R.K., 2002. Optimizing learning from examples using animated pedagogical agents. *J. Educ. Psychol.* 94 (2), 412–427.
- Baker, R., D'Mello, S., Rodrigo, M., Graesser, A., 2010. Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum. Comput. Stud.* 68, 223–241.
- Baylor, A.L., Kim, Y., 2005. Simulating instructional roles through pedagogical agents. *Int. J. Artif. Intell. Educ.* 15, 95–115.
- Baylor, A.L., Ryu, J., 2003. The effects of image and animation in enhancing pedagogical agent persona. *J. Educ. Comput. Res.* 38 (4), 373–394.
- Bennett, R.E., Persky, H., Weiss, A., Jenkins, F., 2007. Problem-Solving in Technology Rich Environments: A Report From the NAEP Technology-Based Assessment Project. NCES 2007-466. U.S. Department of Education, National Center for Educational Statistics, U.S. Government Printing Office, Washington, DC.
- Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., Roscoe, R., 2010. Measuring self-regulate learning skills through social interactions in a teachable agent environment. *Res. Pract. Technol. Enhanc. Learn.* 5, 123–152.
- Cai, Z., Graesser, A.C., Forsyth, C., Burkett, C., Millis, K., Wallace, P., Halpern, D., Butler, H., 2011. Trialog in ARIES: user input assessment in an intelligent tutoring system. In: Chen, W., Li, S. (Eds.), *Proceedings of the 3rd IEEE International Conference on Intelligent Computing and Intelligent Systems*. IEEE Press, Guangzhou, pp. 429–433.
- CAST, 2018. Universal Design for Learning Guidelines Version 2.2. Retrieved from. <http://udlguidelines.cast.org>.
- Cayton-Hodges, G., Bauer, M., Bertling, M., Katz, I., Lehman, B., 2015. Assessing Mathematical Argumentation Through Automated Trialogues. Presented at the 45th Annual Meeting of the Jean Piaget Society. Toronto, Canada.
- Chi, M., 2009. Active–constructive–interactive: a conceptual framework for differentiating learning activities. *Topics Cogn. Sci.* 1, 73–115.
- Clarke-Midura, J., Code, J., Dede, C., Mayrath, M., Zap, N., 2011. Thinking outside the bubble: virtual performance assessments for measuring complex learning. In: Mayrath, M.C., Clarke-Midura, J., Robinson, D. (Eds.), *Technology-Based Assessments for 21st Century Skills: Theoretical and Practical Implications From Modern Research*. Information Age, Charlotte, NC, pp. 125–147.
- Coiro, J., Sparks, J.R., Kiili, C., Castek, J., Lee, C.H., Holland, B.R., 2019. Students engaging in multiple-source inquiry tasks: capturing dimensions of collaborative online inquiry and social deliberation. *Liter. Res. Theor. Method Pract.* <https://doi.org/10.1177/2381336919870285>.
- Cronbach, L.J., Snow, R.E., 1977. *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. Irvington, New York, NY.
- Csapo, B., Funke, J., Schleicher, A. (Eds.), 2012. *On the Nature of Problem Solving: A Look Behind PISA 2012 Problem Solving Assessment*. OECD Series, Heidelberg, Germany.
- D'Mello, S., Graesser, A., 2012. Dynamics of affective states during complex learning. *Learn. InStruct.* 22, 145–157.
- D'Mello, S., Craig, S.D., Sullins, J., Graesser, A., 2006. Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialog. *Int. J. Artif. Intell. Educ.* 16, 3–28.

- D'Mello, S., Taylor, R., Graesser, A., 2007. Monitoring affective trajectories during complex learning. In: McNamara, D.S., Trafton, J.G. (Eds.), *Proceedings of the 29th Annual Cognitive Science Society*. Cognitive Science Society, Austin, TX, pp. 203–208.
- D'Mello, S., Lehman, B., Graesser, A., 2011. A motivationally supportive affect-sensitive AutoTutor. In: Calvo, R.A., D'Mello, S.K. (Eds.), *New Perspectives on Affect and Learning Technologies*. Springer, New York, pp. 113–126.
- D'Mello, S.K., 2013. A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *J. Educ. Psychol.* 105, 1082–1099.
- Driscoll, D.M., Craig, S.D., Gholson, B., Ventura, M., Hu, X., 2003. Vicarious learning: effects of overhearing dialog and monologue-like virtual tutoring sessions. *J. Educ. Comput. Res.* 12, 431–450.
- Duckworth, A.L., Yeager, D.S., 2015. Measurement matters: assessing personal qualities other than cognitive ability for educational purposes. *Educ. Res.* 44, 237–251.
- Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: perseverance and passion for long-term goals. *J. Pers. Soc. Psychol.* 92, 1087–1101.
- Dweck, C.S., 2006. *Mindset: The New Psychology of Success*. Ballantine Books, New York, NY.
- Ekman, P., 1973. Cross-cultural studies of facial expression. In: Ekman, P. (Ed.), *Darwin and Facial Expressions: A Century of Research in Review*. Academic Press, New York, NY, pp. 169–222.
- Elliot, A.J., Murayama, K., 2008. On the measurement of achievement goals: critique, illustration, and application. *J. Educ. Psychol.* 100, 613–628.
- Evanini, K., So, Y., Tao, J., Zapata-Rivera, D., Luce, C., Battistini, L., Wang, X., 2014. Performance of a triologue-based prototype system for English language assessment for young learners. In: *WOCCI-2014*, pp. 79–84.
- Forbes-Riley, K., Litman, D., 2011. Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialog computer tutor. *Speech Commun.* 53, 1115–1136.
- Forsyth, C.M., Graesser, A.C., Walker, B., Millis, K., Pavlik, P., Halpern, D., 2013. Didactic galactic: types of knowledge learned in a serious game. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (Eds.), *Proceedings of Artificial Intelligence in Education: 16th International Conference (AIED 2013)*. Springer Verlag, Berlin Heidelberg, pp. 832–835.
- Forsyth, C.M., Graesser, A.C., Pavlik, P., Millis, K., Samei, B., 2014. Discovering theoretically grounded predictors of shallow vs. deep-level learning. In: Stamper, J., Pardos, Z., Mavrikis, M., McLaren, B.M. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*, pp. 229–232.
- Forsyth, C.M., Luce, C., Zapata-Rivera, D., Jackson, J., Evanini, K., So, Y., 2018. Evaluating English language learners' conversations: man vs. machine. *Int. J. Comput. Assist. Lang. Learn.* <https://doi.org/10.1080/09588221.2018.1517126>.
- Forsyth, C.M., Andrews-Todd, J., Steinberg, J., 2020a. Are you really a team player? Profiles of collaborative problem solvers in an online environment. In: Rafferty, A.N., Whitehill, J., Cavalli-Sforza, V., Romero, C. (Eds.), *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)*, pp. 403–408.
- Forsyth, C.M., Graesser, A.C., Millis, K., 2020b. Predicting learning in a multi-component serious game. *J. Technol. Knowl. Learn.* 25, 251–277.
- Gholson, B., Witherspoon, A., Morgan, B., Brittingham, J.K., Coles, R., Graesser, A.C., Sullins, J., Craig, S.D., 2009. Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instr. Sci.* 37, 487–493.
- Graesser, A.C., Person, N.K., 1994. Question asking during tutoring. *Am. Educ. Res. J.* 31, 104–137.
- Graesser, A.C., Person, N.K., Harter, D., 2001. The tutoring research group: teaching tactics and dialog in AutoTutor. *Int. J. Artif. Intell. Educ.* 12, 257–279.
- Graesser, A., Chipman, P., King, B., McDaniel, B., D'Mello, S., 2007. Emotions and learning with AutoTutor. In: Luckin, R., Koedinger, K., Greer, J. (Eds.), *Artificial Intelligence in Education: Building Technology Rich Learning Contexts that Work*. IOS Press, Amsterdam, pp. 569–571.
- Graesser, A.C., Forsyth, C., Lehman, B., 2017a. Two heads are better than one: learning from agents in conversational triologues. *Teach. Coll. Rec.* 119, 1–20.
- Graesser, A.C., Forsyth, C.M., Foltz, P., 2017b. Assessing conversation quality, reasoning, and problem solving performance with computer agents. In: Csapo, B., Funke, J., Schleicher, A. (Eds.), *On the Nature of Problem Solving: A Look Behind PISA 2012 Problem Solving Assessment*. OECD Series, Heidelberg, Germany, pp. 245–261.
- Graesser, A.C., Foltz, P.W., Rosen, Y., Shaffer, D.W., Forsyth, C., Germany, M., 2018. Challenges of assessing collaborative problem solving. In: Care, E., Griffin, P., Wilson, M. (Eds.), *Assessment and Teaching of 21st Century Skills*. Springer Publishers, Heidelberg, Germany, pp. 75–91.
- Graesser, A.C., 2016. Conversations with AutoTutor help students learn. *Int. J. Artif. Intell. Educ.* 26, 124–132.
- Graesser, A., 2021. Questions for Chapter. Personal Communication, pp. 1–2.
- Hao, J., Liu, L., von Davier, A.A., Kyllonen, P., 2015. Assessing collaborative problem solving with simulation based tasks. In: *Proceedings of the 11th International Conference on Computer-Supported Collaborative Learning*. International Society for the Learning Sciences, Gothenburg, Sweden.
- Hattie, J.A.C., 2009. *Visible Learning: A Synthesis of Meta-analyses Relating to Achievement*. Routledge, New York, NY.
- Herborn, K., Stadler, M., Mustafić, M., Greiff, S., 2020. The assessment of collaborative problem solving in PISA 2015: can computer agents replace humans? *Comput. Hum. Behav.* 104. <https://doi.org/10.1016/j.chb.2018.07.035>.
- Jackson, G.T., Graesser, A.C., 2006. Applications of human tutorial dialog in AutoTutor: an intelligent tutoring system. *Rev. Signos* 39, 31–48.
- Jackson, T., Zapata-Rivera, D., 2017. Conversation-based assessment. In: Rupp, A.A., Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment*. John Wiley & Sons, Inc, pp. 263–279.
- Johnson, W.L., Lester, J.C., 2016. Face-to-face interaction with pedagogical agents, twenty years later. *Int. J. Artif. Intell. Educ.* 26 (1), 25–36.
- Johnson, W.L., Rickel, J.W., Lester, J.C., 2000. Animated pedagogical agents: face-to-face interaction in interactive learning environments. *Int. J. Artif. Intell. Educ.* 11, 47–78.
- Johnson, L., 2021. Questions for Chapter. Personal Communication, pp. 1–2.
- Jonassen, D.H., Grabowski, B.L., 1993. *Handbook of Individual Differences, Learning, and Instruction*. Routledge, New York, NY.
- Jurafsky, D., Martin, J., 2008. *Speech and Language Processing*. Prentice Hall, Englewood, NJ.
- Katz, S., Albacete, P., Chounta, I.A., Jordan, P., McLaren, B.M., Zapata-Rivera, D., 2021. Linking dialog with student modeling to create an adaptive tutoring system for conceptual physics. *Int. J. Artif. Intell. Educ.* 1–49.
- Katz, S., 2021. Questions for Chapter. Personal Communication, pp. 1–2.
- Kyllonen, P.C., 2012. The importance of higher education and the role of noncognitive attributes in college success. *Pensamiento Educativo* 49 (2), 84–100.
- Kyllonen, P.C., 2016. Designing tests to measure personal attributes and noncognitive skills. In: Lane, S., Raymond, M.R., Haladyna, T.M. (Eds.), *Handbook of Test Development*. Routledge/Taylor & Francis Group, New York, NY, pp. 190–211.
- Landauer, T., McNamara, D.S., Dennis, S., Kintsch, W., 2007. *Handbook of Latent Semantic Analysis*. Erlbaum, Mahwah, NJ.
- Lehman, B., Zapata-Rivera, D., 2018a. Frequency, intensity, and Deep Emotions, Oh My! Investigating Control-Value Theory in Conversation-Based Assessments. Presented at the Annual Meeting of the American Educational Research Association. New York, New York.
- Lehman, B., Zapata-Rivera, D., 2018b. Student emotions in conversation-based assessments. *IEEE Trans. Learn. Technol.* 11, 1–13.
- Li, H., Graesser, A., 2017. Impact of pedagogical agents' conversational formality on learning and engagement. In: Andre, E., et al. (Eds.), *AIED 2017 Proceedings*, pp. 188–200.
- Li, H., Graesser, A.C., 2021. The impact of conversational agents' language on summary writing. *J. Res. Technol. Educ.* 53 (1). <https://doi.org/10.1080/15391523.2020.1826022>.
- Lipnevich, A.A., MacCann, C., Roberts, R.D., 2013. Assessing non-cognitive constructs in education: a review of traditional and innovative approaches. In: *The Oxford Handbook of Child Psychological Assessment*. <https://doi.org/10.1093/oxfordhb/9780199796304.013.0033>.
- Liu, L., Steinberg, J., Qureshi, F., Bejar, I., Yan, F., 2016. Conversation-based assessments: an innovative approach to measure scientific reasoning. *Bull. IEEE Tech. Comm. Learn. Technol.* 18 (1), 10–13.
- Lopez, A.A., Guzman-Orth, D., Zapata-Rivera, D., Forsyth, C.M., Luce, C., 2021. Examining the Accuracy of a Conversation-Based Assessment in Interpreting English Learners' Written Responses (Research Report No. RR-21-03). Educational Testing Service. <https://doi.org/10.1002/ets2.12315>.
- Maltese, A.V., Tai, R.H., 2010. Eyeballs in the fridge: sources of early interest in science. *Int. J. Sci. Educ.* 32, 669–685.

- Martin, M.M., Rubin, R.B., 1995. A new measure of cognitive flexibility. *Psychol. Rep.* 76, 623–626.
- Mayer, R.E., 2005. Principles based on social cues: personalization, voice, and presence principles. In: Mayer, R.E. (Ed.), *Cambridge Handbook of Multimedia Learning*. Cambridge University Press, New York, pp. 201–212.
- McNamara, D.S., O'Reilly, T., Best, R., Ozuru, Y., 2006. Improving adolescent students' reading comprehension with iSTART. *J. Educ. Comput. Res.* 34, 147–171.
- Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A.C., Halpern, D., 2011. Operation ARIES! A serious game for teaching scientific inquiry. In: Ma, M., Oikonomou, A., Lakshmi, J. (Eds.), *Serious Games and Edutainment Applications*. Springer-Verlag, London, pp. 169–196.
- Mislevy, R.J., Steinberg, L.S., Almond, R.G., 2003. On the structure of educational assessments. *Meas. Interdiscip. Res. Perspect.* 1, 3–62.
- Moreno, R., Mayer, R.E., Spires, H.A., Lester, J., 2001. The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognit. InStruct.* 19 (2), 117–213.
- Nottle, E.E., Robins, R.W., 2007. Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. *J. Pers. Soc. Psychol.* 93 (1), 116–130.
- OECD, 2013. *PISA 2015 collaborative problem solving framework*. OECD Publishing, Paris.
- Olney, A., D'Mello, S.K., Person, N., Cade, W., Hays, P., Williams, C., Lehman, B., Graesser, A.C., 2012. Guru: a computer tutor that models expert human tutors. In: Cerri, S., Clancey, W., Papadourakis, G., Panourgia, K. (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems*. Springer-Verlag, Berlin Heidelberg, pp. 256–261.
- Olney, A., 2021. Questions for Chapter. Personal Communication, pp. 1–2.
- Pekrun, R., Goetz, T., Perry, R., Kramer, K., Hochstadt, M., Molfenter, S., 2004. Beyond test anxiety: development and validation of the Test Emotions Questionnaire (TEQ). *Hist. Philos. Logic* 17, 287–316.
- Pekrun, R., Goetz, T., Frenzel, A., Barchfield, P., Perry, R., 2011. Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemp. Educ. Psychol.* 36, 36–48.
- Pekrun, R., 2006. The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341.
- Pintrich, P.R., De Groot, E.V., 1990. Motivational and self-regulated learning components of classroom academic performance. *J. Educ. Psychol.* 82, 33–40.
- Quellmalz, E.S., Timms, M.J., Buckley, B.C., Davenport, J., Loveland, M., Silbergitt, M.D., 2011. 21st century dynamic assessment. In: Mayrath, M.C., Clarke-Midura, J., Robinson, D. (Eds.), *Technology-Based Assessments for 21st Century Skills: Theoretical and Practical Implications From Modern Research*. Information Age, Charlotte, NC, pp. 55–90.
- Ramanarayanan, V., Suendermann-Oeft, D., Lange, P., Ivanov, A.V., Evanini, K., Yu, Z., Truprun, E., Qian, Y., 2016. Bootstrapping Development of a Cloud-Based Spoken Dialog System in the Educational Domain From Scratch Using Crowdsourced Data. Research Report No. RR-16-16. Educational Testing Service, Princeton, NJ. <https://doi.org/10.1002/ets2.12105>.
- Richardson, M., Abraham, C., Bond, R., 2012. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138 (2), 353–387. <https://doi.org/10.1037/a0026838>.
- Rosé, C.P., Howley, I., Wen, M., Yang, D., Ferschke, O., 2017. Assessment of discussion in learning contexts, invited chapter. In: von Davier, A., Zhu, M., Kyllonen, P. (Eds.), *Innovative Assessment of Collaboration*. Springer Verlag, New York, NY.
- Rosé, C.P., Torrey, C., 2005. Interactivity versus Expectation: Eliciting Learning Oriented Behavior with Tutorial Dialogue Systems. In: *Proceedings of 10th IFIP TC13 International Conference on Human-Computer Interaction - Interact 2005, Lecture Notes in Computer Science, Volume 3585*, pp. 323–336.
- Rosé, C., 2021. Questions for Chapter. Personal Communication, pp. 1–2.
- Schneider, M., Preckel, F., 2017. Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol. Bull.* 143, 565–600.
- Schroeder, N., Adesope, O., Gilbert, R., 2013. How effective are pedagogical agents for learning? A meta-analytic review. *J. Educ. Comput. Res.* 49 (1), 1–39.
- Snow, R.E., 1986. Individual differences and the design of educational programs. *Am. Psychol.* 41 (10), 1029–1039. <https://doi.org/10.1037/0003-066X.41.10.1029>.
- So, Y., Zapata-Rivera, D., Cho, Y., Luce, C., Battistini, L., 2015. Using dialogues to measure English language skills. In: García Laborda, J., Sampson, D.G., Hambleton, R.K., Guzman, E. (Eds.), *Journal of Educational Technology and Society. Special Issue: Technology Supported Assessment in Formal and Informal Learning*, pp. 21–32.
- Soland, J., Zammaro, G., Cheng, A., Hitt, C., 2019. Identifying naturally occurring direct assessments of social-emotional competencies: the promise and limitations of survey and assessment disengagement metadata. *Educ. Res.* 48 (7), 466–478.
- Song, Y., Sparks, J., Brantley, W., Oliveri, M., Zapata-Rivera, D., 2014. Designing Game Activities to Assess Students' Argumentation Skills. Paper Presented at the Annual Meeting of the American Educational Research Association (AERA), Philadelphia.
- Spangler, G., Pekrun, R., Kramer, K., Hofmann, H., 2002. Students' emotions, physiological reactions, and coping in academic exams. *Hist. Philos. Logic* 15, 413–432.
- Sparks, J.R., Zapata-Rivera, D., Lehman, B., James, K., Steinberg, J., 2018. Simulated dialogues with virtual agents: effects of agent features in conversation-based assessments. In: Penstein Rosé, C., et al. (Eds.), *Proceedings of AIED 2018: Artificial Intelligence in Education*. Springer Cham, Basel, Switzerland, pp. 469–474.
- Sparks, J.R., Peters, S., Steinberg, J., James, K., Lehman, B.A., Zapata-Rivera, D., 2019. Individual Difference Measures that Predict Performance on Conversation-Based Assessments of Science Inquiry Skills. Paper Presented at the Annual Meeting of the American Educational Research Association (AERA), Toronto, Canada.
- Sparks, J.R., Steinberg, J., Castellano, K., Yan, F., Lehman, B., Zapata-Rivera, D., 2020. Generating Individual Difference Profiles Via Cluster Analysis: Toward "Caring Assessments" for Science. Poster Accepted at the Annual Meeting of the National Council for Measurement in Education (NCME), San Francisco, CA (Conference canceled).
- Verschelden, C., 2017. Bandwidth Recovery: Helping Students Reclaim Cognitive Resources Lost to Poverty, Racism, and Social Marginalization. Stylus, Sterling, VA.
- Wang, N., Johnson, W.L., 2008. The politeness effect in an intelligent foreign language tutoring system. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (Eds.), *Intelligent Tutoring Systems. ITS 2008, Lecture Notes in Computer Science, vol. 5091*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69132-7_31.
- Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., Weston, T., 2013. My science tutor: a conversational multimedia virtual tutor. *J. Educ. Psychol.* 105, 1115–1125.
- West, M.R., Kraft, M.A., Finn, A.S., Martin, R.E., Duckworth, A.L., Gabrieli, C.F.O., Gabrieli, J.D.E., 2016. Promise and paradox: measuring students' non-cognitive skills and the impact of schooling. *Educ. Eval. Pol. Anal.* 38, 148–170.
- West, M.R., Buckley, K., Krachman, S.B., Bookman, N., 2018. Development and implementation of student social-emotional surveys in the CORE Districts. *J. Appl. Dev. Psychol.* 55, 119–129.
- Wise, S.L., Bhola, D., Yang, S., 2006. Taking the time to improve the validity of low-stakes tests: the effort-monitoring CBT. *Educ. Meas.* 25, 21–30.
- Yin, C., Mislevy, R.J., 2022. Evidence-centered Design, *International Encyclopedia of Education*, 4th ed. Elsevier (in press).
- Zapata-Rivera, D., Liu, L., Katz, I.R., Vezzu, M., 2013. Exploring the use of game elements in the development of innovative assessment tasks for science. *Cogn. Technol.* 18 (1), 43–50.
- Zapata-Rivera, D., Jackson, T., Liu, L., Bertling, M., Vezzu, M., Katz, I.R., 2014. Science inquiry skills using dialogues. In: *12th International Conference on Intelligence Tutoring Systems*, pp. 625–626.
- Zapata-Rivera, D., Jackson, T., Katz, I.R., 2015a. Authoring conversation-based assessment scenarios. In: Sottolare, R.A., Graesser, A.C., Hu, X., Brawner, K. (Eds.), *Design Recommendations for Intelligent Tutoring Systems Volume 3: Authoring Tools and Expert Modeling Techniques*. U.S. Army Research Laboratory, pp. 169–178.
- Zapata-Rivera, D., Lehman, B., Jackson, T., 2015b. Conversation-based assessments at educational testing service. Science inquiry skills using dialogues. In: Conati, C., Heffernan, N., Mitrovic, A., Verdejo (Eds.), *Proceedings of Artificial Intelligence in Education. Book Series Lecture Notes in Artificial Intelligence*. Springer, New York, NY, pp. 906–910.
- Zapata-Rivera, D., Liu, L., Chen, L., Hao, J., von Davier, A., 2016. Assessing science inquiry skills in immersive, conversation-based systems. In: Daniel, B.K. (Ed.), *Big Data and Learning Analytics in Higher Education*. Springer International Publishing, pp. 237–252. https://doi.org/10.1007/978-3-319-06520-5_14.

- Zapata-Rivera, D., Lehman, B., Sparks, J.R., Por, H.-H., James, K., 2018. Identifying and Addressing Unexpected Responses in Conversation-Based Assessments. Research Memorandum No. RM-18-13. Educational Testing Service, Princeton, NJ.
- Zapata-Rivera, D., Lehman, B., Sparks, J.R., 2020. Learner Modeling in the Context of Caring Assessments. In: Sottolare, R., Schwarz, J. (Eds.), Adaptive Instructional Systems. HCI 2020. Lecture Notes in Computer Science, 12214, pp. 422–431.
- Zapata-Rivera, D., 2017. Toward caring assessment systems. In: Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP'17). ACM, New York, NY, USA, pp. 97–100. <https://doi.org/10.1145/3099023.3099106>.
- Zapata-Rivera, D., 2020. Open student modeling research and its connections to educational assessment. Int. J. Artif. Intell. Educ. <https://doi.org/10.1007/s40593-020-00206-2>.
- Zeidner, M. (Ed.), 1998. Test Anxiety: The State of the Art. Plenum Press, New York, NY.

Classroom assessment

James H. McMillan, Virginia Commonwealth University, Richmond, VA, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction to classroom assessment	519
Two main purposes of classroom assessment: summative and formative	520
Summative classroom assessment	520
Types of summative assessment and learning targets	521
Grading	521
Feedback	521
Formative classroom assessment	522
Characteristics	522
Embedded formative assessment	522
Formative use of summative assessment	523
Classroom assessment literacy	523
Efforts to identify assessment literacy	523
Components of classroom assessment literacy	524
Foundational concepts	525
Foundational influences	526
Action	526
Consequences	526
References	527

Introduction to classroom assessment

In the field of education there is a long tradition of using measures like tests to establish evidence of student proficiency. Measurement is the assignment of numbers to characterize a student performance event that indicates knowledge, understanding or skill. It is a quantitative approach to documenting what students know and can do. Educators also use qualitative judgments to discern student proficiency. Whether quantitative or qualitative, the process of assessment in education includes interpretation and use of the information gathered to document and/or improve student learning.

Educational assessment includes a range of different ways of gathering information, including large scale, standardized tests, locally designed quizzes and homework assignments, observations, and school district measures. In the classroom, the goal is to use assessment for improving student proficiency in relatively fine-grained learning segments or units. It is used to support teaching and learning. As such, effective classroom assessment is closely based upon and integrated with contemporary research on learning and motivation (Shepard, 2019).

Classroom assessment, then, “is a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes, including diagnosing student strengths and weaknesses, monitoring student progress toward meeting desired levels of proficiency, assigning grades, providing feedback to students and parents, and enhancing student learning and motivation.” (Brookhart and McMillan, 2020, p. 4). In this chapter the primary emphasis is on what teachers do in gathering, evaluating, and using evidence of student proficiency.

The collection of evidence in classroom assessment occurs in the context of individual classrooms (DeLuca et al., 2019; Xu and He, 2019), and is locally controlled by the teacher, whether designed by the teacher or an external source. It is different from large-scale assessment, such as the use of standardized tests and year-end accountability tests, in both design and purpose. As pointed out by Shepard (2019), tests designed for monitoring, accountability, and selection are unable to be used effectively to improve student learning, and often undermine what is taught and student motivation. In contrast, effective classroom assessment that is based on recent sociocultural learning theories supports a culture that can enhance instruction and deep student learning. Shepard states (p. 184):

Because classroom assessment is intended to aid directly in the learning process (not merely to measure learning outcomes), it necessarily must be closely tied to instructional practices and to relevant research: on learning (in subjects such as mathematics, science, and literacy; on motivation, feedback, and self-regulation; on cognitive and sociocultural aspects of learning and identify; on curriculum; on adaptive teaching and teacher-learning; and on theories of formative assessment and grading.

Classroom assessment, as a process, includes a number of steps (see Fig. 1). Initially, fine-grained learning outcomes or targets and success criteria need to be clearly defined in terms that students can understand. This may involve deconstructing required

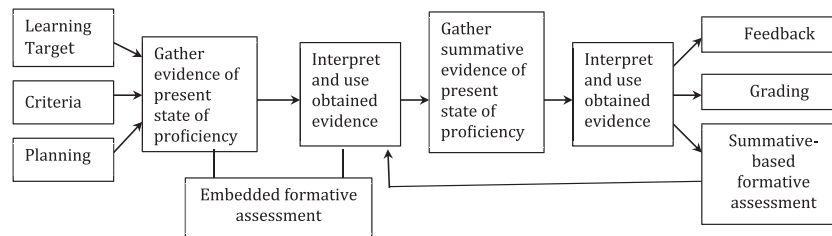


Fig. 1 The process of classroom assessment.

learning goals standards to clarify the nature of knowledge and skills that lead to proficiency. The purpose of the assessment is determined, as primarily formative, to enhance learning, or summative, to document current levels of proficiency. Evidence is then gathered to provide a beginning point for instruction. Embedded formative assessment (William and Leahy, 2015) is conducted concurrently with instructional activities. Following instruction and embedded formative assessment, teachers implement appropriate assessments that are primarily summative (design and selection of implemented assessments occurs both prior to and following instruction). In the final stage, use is determined, and this action needs to be based on an understanding of how grading, feedback and reporting impact student learning and motivation. This includes the use of summative assessment to improve instruction and student learning.

Two main purposes of classroom assessment: summative and formative

While classroom assessment is used for many reasons, such as serving as evidence for grading, tracking student progress, reporting to parents, improving instruction, remediation, diagnosing student misconceptions, motivating students, and teacher and school accountability, there are two core purposes – summative and formative. This section describes each of these purposes.

Summative classroom assessment

The primary purpose of summative classroom assessment is to document what students know, their skills, and other cognitive outcomes. It can be described as assessment of student performance. Summative assessments, most notably tests, papers, projects, and quizzes, are used to monitor and record student proficiency in achieving well-defined learning outcomes. Key characteristics of summative assessment include the following:

- Administered after instruction is completed.
- Focused on knowledge, comprehension, and application learning outcomes.
- Limited opportunity for students to shape, personalize or modify the assessment.
- Pre-planned, structured format.
- Used primarily for grading.
- Encourages performance-oriented student motivation.
- Whether selected-response or constructed-response items are used, tends to be objectively scored, often automatically by digital programs.
- Interpreted with reference to levels of proficiency.

Since the principle use of summative assessment is to document student learning, it has historically been tied to traditional measurement principles as applied to large-scale testing, including approaches to establish technical qualities, with an emphasis on objective scoring. This has often led to implementation of assessment principles and techniques that are not useful in promoting effective instruction and student learning. If teachers use traditional measurement concepts as a basis for classroom assessment that use selected-response items, then, classroom assessment tends to be focused more on what was learned than improving learning. As shown by recent research, in contrast, most teachers are guided in their assessment practices by a core value toward using evidence of performance to enhance student learning. This translates to a situated and differentiated perspective on classroom assessment practices, stressing the importance of individual teachers and context (DeLuca et al., 2019).

For summative assessment it is important that the nature of the assessment represents what was targeted in instruction and learning materials. The match between the content of the assessment and the learning goals and objectives is determined by a review and categorization of the items used in the assessment. In a selected-response test, in which students select an answer from provided alternatives (e.g., multiple-choice), the review of the items provides content-related evidence of validity (*validity* is the extent to which the inferences from the test, such as acceptable or unacceptable performance, are supported as reasonable). That is, with appropriate evidence of the match, it is reasonable to infer that when students perform well, they know all aspects of the learning that was planned. The match is sometimes assured by use of a table of specifications. The table is a two-dimensional chart that lists

key learning areas and types of knowledge or understanding that is emphasized. Items are then categorized according to both dimensions. The results indicate the percentage of items that measure each learning area and type of knowledge, understanding or application.

Types of summative assessment and learning targets

It is important to match the type or method of the summative assessment with the nature of the learning that is targeted. Selected-response items (e.g., multiple-choice, binary-choice, and matching) are best for assessing knowledge and simple understanding, such as recognition, recalling facts, definitions, and spelling. Selected-response items are effective when the goal is to sample a larger domain of knowledge. Constructed-response items, including essay questions, which require construction of an answer without the presence of possible answers, are well-suited to assessing more in-depth understanding, processes, relationships, and application. Performance assessments, such as making presentations, complete projects, and making products, are used to demonstrate student skills and application of knowledge in what they construct and present proficiency in the form of a report or demonstration. While performance assessments are more time consuming, they have the advantage of requiring more complex thinking to what are often authentic tasks closely related to living outside of school. The increased relevance of the tasks and resultant student engagement enhances motivation and learning but requires time to carefully develop rubrics and score answers subjectivity for grading and feedback. When used by multiple teachers, common interpretation of rubric levels is needed.

In the last two decades summative classroom assessments have been increasingly tied to year-end accountability tests. This is done to familiarize students with the format of accountability tests and assure coverage of the content on these “high-stakes” tests. Often, classroom tests include items that have been used and publicly released from previous years. However, it is clear that there are many deleterious impacts of this trend on student learning, instruction and teacher morale (Amrein and Berliner, 2003; Heritage, 2018; Shepard et al., 2018). If classroom assessments mimic high-stakes tests the items tend to focus on lower-level cognitive skills, narrow instruction and curriculum, and employ multiple-choice formats. Very recently, there is increased pressure to use items obtained from online test banks, with little systematic review of item quality or match to instruction (McMillan, 2018). School districts may also develop and employ common classroom assessments, the same test given by all teachers in the same courses and grade levels. Unless utilized for department level discussions about instruction within a unit, curriculum materials, and other instructional adjustments, these practices tend to undermine the primary purpose of classroom assessment tied to improving student learning and enhancing student motivation. Significant developments in large-scale testing have recently increased the usefulness of such assessment practices to inform teaching and learning (Brookhart and DePascale, *in press*). Large-scale testing companies have devoted extensive resources to provide assessments that can be used for assessing student growth and instruction, though there is disagreement about the utility and value of these tests as substitutes for classroom assessment.

Grading

An important function of summative assessment is to assign grades to student work. Grades are typically calculated on the basis of the percentage of selected-response items answered correctly, points given for short answer answers, or letters (e.g., A, B) for papers, presentations and projects. Percentages and points are usually converted to letter grades. Grades are also assigned to performance in relation to established standards of learning (standards-based grading).

The practice of grading is both longstanding and controversial (Guskey and Brookhart, 2019; Brookhart et al., 2016). Brookhart and Guskey (2019), in a recent review of 100 years of grading practices, found that teachers continue to use multiple factors in the determination of grades, and that even when grading tests and papers, there are often significant differences in the criteria for judging work, teacher severity or leniency, differences in tasks and item difficulty, teacher bias, differences in weighting various aspects of papers, and different grading scales. This is not surprising since teachers’ and school districts’ standards differ, especially when focused on classroom assessments. However, individual teachers also vary their grading practices. There are specific strategies that teachers can employ to support making grading more reliable, including the use of clear criteria, teacher training in assessment and grading to enhance the use of high quality test items, the use of scoring guides and rubrics to focus on dimensions of criteria and levels of performance, grading each short-answer item for all students then grading the next item, grading anonymously, and collaborating with others to review grading practices.

Feedback

Feedback from summative assessments has typically been straightforward and simple – an indication of what is answered correctly or incorrectly, total score, and grade (verification feedback). When verification is accompanied by explanations of why answers are incorrect and suggestions for further learning, students are encouraged to take further action based on that information that will enhance their learning. Time and supportive resources are needed to enable students to review and respond to feedback, especially to specify what and how they will address the feedback and revise their work. In other words, students need to act on the feedback for it to be effective. Because each student is unique, this often requires differentiated feedback to best address individual strengths and needs (Brookhart, 2017). Such differentiation is accomplished with the use of targeted, specific comments that essentially coach students to engage in further learning toward desired outcomes (Brookhart, 2020; Ruiz-Primo and Brookhart, 2018). In the end, if students do not act on the feedback, it is not effective.

The results of summative assessments for groups of students can be used to provide similar feedback to each group, and inform subsequent instruction with the group to address inadequate understanding and to modify future instruction. The teacher and/or students first identify what content and process learning outcomes from the unit need attention, what specific further steps are needed, and how eventual learning is assessed. The goal of these steps is to improve student learning once it has been documented.

Formative classroom assessment

Characteristics

The primary purpose of formative classroom assessment is to improve student learning. Since the seminal publication by Black and Wiliam (1998), formative assessment has become a major focus for classroom assessment. Heritage (2018) contends that the purpose of all classroom assessment “is to provide information that can be used to support student learning” (p. 39). Whether or not all classroom assessment can be used primarily to improve learning, as well as to document and report learning, it is clear that formative assessment, or what some have called assessment *for* learning (AfL), dominates research and writing about classroom assessment.

Formative assessment has been defined in many ways, with most a variation or extension of Black and Wiliam (2009, p. 9):

[Assessment] practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learning, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited.

This definition emphasizes that assessment *forms* instruction and learning experiences (Wiliam and Leahy, 2015). That is, the “evidence” of current student proficiency that is assessed by the teacher is used to feed back to the teacher and student to determine next steps in learning. Formative assessment is purposeful, a process of continuous gathering of evidence of student learning in relation to learning goals and then using the evidence for teacher decision making about instructional adjustments, and to give feedback to students. It is important to emphasize that formative assessment is not a specific tool or test (i.e., a formative assessment). Rather, it is a more comprehensive process that includes five essential elements that expand upon three questions – Where are students trying to go? Where are they now? How can they get there?

1. The identification of clear learning goals and criteria for evaluating student work.
2. Gathering evidence of student proficiency.
3. Evaluating gathered evidence.
4. Providing feedback to students.
5. Making instructional adjustments.

Other definitions of formative assessment include involving students and teachers collaboratively, student self-evaluation, learning progressions, a planned system that integrates metacognition and self-regulation, and domain specific and domain general differentiation across disciplines (Andrade and Cizek, 2010; Cizek et al., 2019; Brookhart and DePascale, *in press*). More expansive definitions of formative assessment have been promulgated to integrate research on instruction and learning that align with how formative assessment is implemented. In particular, as pointed out by Shepard (2019), formative assessment that is coherently linked to motivation and cognitive development is able to impact learning in an equitable manner.

Formative assessment is implemented to various extents in different ways. While teachers may only use some of the essential components, most effective formative assessment occurs when all five elements are evident. For example, a quiz may be employed to gauge student learning, but if there is ineffective feedback to the student and little or no instructional adjustment simply giving the quiz and noting results leaves out these two elements. Two main types of formative assessment have emerged – embedded and summative-based (McMillan, 2018).

Embedded formative assessment

The most powerful and effective type of formative assessment is embedded in the course of instruction. It is implemented as teaching and learning are taking place, essentially “embedded” in instruction. One type of embedded formative assessment can be described as “on-the-fly,” “at-the-moment,” or informal. It occurs spontaneously, depending how a lesson is being received, how engaged students are in learning, and how students respond to questions and prompts. This involves continuous monitoring of students while teaching, gathering evidence about how learning is developing. The evidence is then interpreted immediately to help the teacher know whether instructional correctives are needed (e.g., slowing down, giving more examples, using small groups to discuss, calling on students). Based on student responses and behavior, feedback is immediate.

Effective “at-the-moment” embedded formative assessment involves keen observation and appropriate questioning of students. Teachers need to observe both nonverbal and verbal cues to ascertain confidence, engagement, and understanding. This includes body language and facial expressions, as well as vocal loudness, pitch and rate of speaking. Questioning needs to allow sufficient wait time for student responses; questions need to be clear, succinct, and appropriately sequenced, with probing, as well as differentiated based on student ability, grade level, cultural backgrounds, and the subject matter being taught.

A second type of embedded formative assessment uses planned prompts, problems, assignments and questions that a teacher develops as an integral part of lesson planning. Less formal than a quiz or homework, use of the techniques provides the teacher with evidence of student learning to which feedback and instructional correctives can be made. Student responses are ungraded.

Formative use of summative assessment

Summative-based formative assessment, also referred to as formal formative assessment, is conducted through the use of more prescribed, structured, and scheduled sets of questions or problems that require student responses. This includes homework, quizzes, unit tests, and other tasks for which there would be a formal score to indicate student proficiency. Most of these assessments follow instruction, rather than occur during instruction. The results are gathered to inform teachers about progress of an entire class, subgroups of students, and individual students in order to contribute to a grade and inform whether any instructional course corrections are needed.

Preassessments of what students know, and their skills, are used for several purposes. While these ungraded assessments do not occur after instruction, like summative assessments, they help determine an appropriate point to begin instruction, one that matches with student knowledge to avoid either redundancy or expectations that are too high. This allows teachers to plan instruction that is at the right level of difficulty – challenging students to motivate them. Preassessments can also be used to make instruction more interesting, relevant, and engaging. Both structured exercises and quizzes are employed for assessing students' level of understanding to be able to plan instruction to build on these levels. Structured activities could include a writing assignment, group work, or targeted questioning. By combining the information obtained with other data, such as test scores, student records, and other teacher observations to provide confirmation of student understanding, student strengths and weaknesses are identified tied directly to the next instructional topic. Short, ungraded pretests of targeted learning outcomes can also identify current levels of understanding, acquaint students with what will be learned, and provide a baseline to gauge improvement in proficiency.

Homework is used to provide additional practice in applying what students are taught. It acts as a check on student learning to help determine whether students have reached proficiency. Homework performance can be used diagnostically to structure further instruction and individualized assistance. Homework is most effective when students provide unassisted answers to constructed-response questions, and when the teacher is able to provide meaningful feedback.

Summative quizzes and unit tests can be used formatively by providing teachers with an indication of student strengths and weaknesses that lead to instructional correctives, and for providing meaningful feedback to students. Beyond grades, feedback can specify areas needing remediation or further clarification or elaboration. This can enhance the impact of future instruction and improve student motivation that leads to further learning. Quizzes are effective because they are short, targeted to specific learning outcomes, and lead to immediate feedback to students about what they do and do not understand. Quiz-based, targeted feedback makes it relatively easy for students to relate their performance to how they learned the material. Unit tests, which may cover a week or more of learning, are more difficult to use formatively. This is due to the coverage of more material, the delayed nature of feedback, and the more extended gap between the result and learning process.

Classroom assessment literacy

It is generally recognized that classroom assessment is a process of inquiry and action in which teachers collect, interpret, and use fine-grained evidence of targeted student proficiency (e.g., specific outcomes related to short classroom lessons) to document and improve student learning. This process is not unlike what is used in conducting research more broadly. Questions are initiated for which data are collected, obtained evidence is interpreted, and conclusions are reached about the meaning of the results.

Defined in this way, there are three characteristics that are unique to how the assessment process is used by teachers in the classroom: 1) the emphasis on relatively fine-grained student proficiency rather than broader competencies and abilities; 2) the central role of teachers in determining and enacting the process and interpreting evidence (though both student self-assessment and peer-assessment strategies can be utilized); and 3) the dual purpose of both documenting and enhancing student learning.

These characteristics differentiate classroom assessment from a larger set of skills needed to interpret and use results from large-scale testing that focuses on larger-grained student proficiency. In both cases, testing is linked to student learning and instruction, but because the theoretical basis for how information is gathered and used is fundamentally different, the nature of teacher competencies to effectively use each type of assessment also differs. Fine-grained assessment for learning is informed most by learning theory and how teachers interact with students, while the use of larger-grained (long-cycle) assessment, from interim and year end tests, is implemented with more of a psychometric emphasis. For example, Brookhart (2011) points out that formative assessment as employed by teachers, based on information they obtain from classroom interactions with students and other fine-grained assessment, is quite different from what is advocated by testing companies in promoting the use large-scale test results for formative purposes. The dual purpose of classroom assessment also sets it apart from most large-scale testing and reflects the recognition that assessment is most productive when it can be used to improve student learning.

Efforts to identify assessment literacy

Researcher and expert efforts to suggest what teachers should know about assessment span decades, even if the term “assessment” is more contemporary than earlier language that stressed different but related concepts of measurement, testing and evaluation.

Changes from initial efforts, most of which were based on what prospective teachers need in their preparation programs, to more recent work, shows the development of recommended knowledge and skills from an application of psychometric principles to what is needed to improve as well as document learning.

The initial formal initiative to describe what teachers should know about assessment, the *Standards for Teacher Competence in Educational Assessment of Students* (1990), emphasized the application of measurement concepts for both classroom and standardized assessment. These standards have been important in framing what has been presented in textbooks, training programs, and measures to document the level of teachers' assessment competencies. Notably, Stiggins (1991), who probably first coined the term "assessment literacy," conceptualized assessment-literate individuals as having "a basic understanding of the meaning of high- and low-quality assessment and are able to apply that knowledge to various measures of student achievement" (p. 535).

Two significant efforts, involving multiple professional groups, have promulgated classroom assessment standards. *The Student Evaluation Standards: How to Improve Evaluations of Students* (Joint Committee on Standards for Educational Evaluation, 2003) was the first set of competencies dedicated to assessment use at the classroom level, with minimum attention to standardized testing. The 2003 *Standards* represents a massive, systematic effort to identify classroom assessment competencies and demonstrates movement toward practices that focus on improving student learning. Interestingly, though, there is little explicit language that relates assessment to instruction (the term formative assessment is not used). The emphasis is on the gathering of summative evidence that results in accurate and fair reporting of student proficiency. The overall message is action-oriented – what do teachers, administrators, and policy makers need to understand to enhance teachers' "ability to design and implement the most useful, proper, feasible, and accurate classroom evaluation practices?" (p. 13).

The Classroom Assessment Standards for PreK-12 Teachers (Joint Committee on Standards for Educational Evaluation, 2015) is a revision of the 2003 *Standards*. The 2015 *Standards* shows further evolution toward the central role of student learning in assessment, reflecting an effort to move both the organization and nature of the competencies toward having a positive impact on student learning. Over several years, sources of input, including a substantial review of literature, feedback of drafts from groups of practitioners, researchers, members of the Joint Committee, field trials, national hearings, and a validation panel, were used to establish and verify the standards. The 2015 *Standards* are organized into three domains: Foundations, Use, and Quality. The Foundations domain contains six standards that address essential assessment principles that will enhance instruction and student learning (purpose, learning expectations. For example, the first standard in the Foundations domain, Purpose, states clearly that results from assessment should be used to improve learning. The second standard, Learning Expectations, points to the importance of appropriate alignment between assessment and learning opportunities. The third, Assessment Design stresses student opportunity to show learning, and the fourth standard addresses student engagement in assessment that will improve learning. The second domain, Use, moves from purpose and design to showing how results should be used to support learning, including standards on how student performance is analyzed, effective feedback, instructional follow up, and grading. The Quality domain contains standards such as the importance of cultural and linguistic diversity, exceptionality, fairness, reliability, and validity. Overall, 8 of 11 standards in the first two domains directly reference student learning and/or instruction.

While the emphasis in the 2003 *Standards* was on describing teacher knowledge and skills, the 2015 *Standards* extend this focus to using assessment evidence to inform and improve instruction and student learning. The newer standards are much more concise and user friendly, with a comprehensive emphasis throughout on how implementation of the right kind of quality assessment practices will improve instruction and student learning.

The move toward greater emphasis on the impact of assessment on students is demonstrated in several research efforts to identify essential teacher assessment competencies. DeLuca et al., (2016) examined 15 documents from six countries published since 1990 that explicitly addressed standards or competencies for literacy related to student assessment. Their analysis showed a swing in conceptualizations of teacher assessment competencies to focus less on summative and standardized tests, and more on fairness, assessment for learning, communication of assessment results, and support for teachers. In another comprehensive review, Xu and Brown (2016) identified 100 peer-reviewed publications from 1985 to 2015 that focused on teachers' assessment literacy, competence, and professional development. The purpose was to synthesize findings from empirical studies and expert recommendations to suggest a framework that integrates knowledge and skills with measures of assessment literacy and needs for pre- and in-service education. The three major themes presented (knowledge and skills within assessment literacy, assessment education and its relationship with various mediating factors, and contextual considerations of assessment literacy) and ten subthemes are used to suggest a conceptual framework in which knowledge about assessment is integrated with classroom contextual factors, individual proclivities and values of teachers, and socio-cultural and institutional factors.

Pastore and Andrade (2019), employing a Delphi inquiry with 35 international experts on educational assessment and teacher education, propose a model of teacher assessment literacy organized by three components: (1) a conceptual knowledge dimension (what, why and how to assess, data analysis, and effective reporting); (2) a praxeological dimension (assessment in practice, navigating multiple assessment demands and use of assessment data); and (3) a socio-emotional dimension (sense-making of assessment practices, fairness, and awareness of power dynamics).

Components of classroom assessment literacy

Components (i.e., competencies) of classroom assessment literacy build upon and update existing lists of assessment standards using theories of learning and motivation, research on formative assessment, and the growing importance of student perceptions

about assessment as a critical link between what is shared with students about their performance and subsequent impacts on learning and motivation. The components are based on the following:

1. Practicality is critical to use. Teachers will best understand and implement classroom assessment practices that are meaningfully related to the process of assessment and instruction. Hence, suggested competencies are organized by how teachers plan, implement and use assessment.
2. Classroom assessment for students is a sociocultural and emotional, as well as cognitive, experience (Brown and Harris, 2016). As a process that involves interactions between students and teachers, it is fundamentally relational in which teacher values and perspectives are integrated with students' cultural backgrounds (Looney et al. (2017)). As a result, teacher-student relationships shape how students react to and gain meaningfulness from assessment results.
3. Classroom assessment is best conceptualized as a set of practices that are implemented *with* and *for* students, not something done *to* students. These suggested competencies emphasize how assessment practices are implemented to have positive impacts on not only standards-based knowledge and skills, but other important outcomes, such as emotions, motivation, attributions, self-regulation and self-efficacy.
4. To be effective, classroom assessment competencies must be framed within what we know about how students learn and how they are motivated. In particular, recent constructivist and sociocultural models of learning need to be employed to provide meaningful connections between assessment and instruction. Implementing these models promotes the use of challenging, authentic assessment tasks that increase intrinsic motivation and feelings of competence.
5. Like instruction, the process of classroom assessment is contextually bound and influenced. Application of assessment methods, interpretation, and use is influenced by the nature of students in a particular class (e.g., skill level, responses to different types of assessments and feedback), subject matter, level of schooling, teaching style, teacher beliefs about what is important to emphasize, the relative influence of high-stakes tests, and other contextual factors. As related to a process of inquiry, then, these factors need to be understood and considered in decision-making, especially with respect to grading. This suggests that while it is helpful to articulate a single set of competencies, the reality is that the process of inquiry and action is individualized to particular settings.

The classroom assessment components presented here are organized into four categories. The first and second categories, Foundational Concepts and Foundational Influences, list areas of knowledge and understanding that are needed for planning and using classroom assessment. The third category, Action, shows skills needed for how assessment is designed, implemented, scored. Finally, based on foundational knowledge and evidence gathered, teachers need to consider the consequences of classroom assessment – how the process influences students cognitively, emotionally, and behaviorally beyond the immediate assessment task and receipt of scores.

Foundational concepts

- *Classroom Assessment as a Process of Inquiry.* Teachers need a thorough understanding of the nature of classroom assessment as a sequenced progression of steps that involve inquiry and action, and how this process is different from “testing” and large-scale assessment. There is a need to pay particular attention to the role of professional judgment in all phases of the process (Looney et al., 2017).
- *Purpose.* Teachers need to understand the formative and summative purpose(s) of a particular classroom assessment event. Formative assessment focuses on improving student learning; summative on documenting proficiency. A complete consideration of purpose should encompass the extent to which other reasons for assessment are important, such as whether they need to help prepare students for standardized testing, sort students for selection, use for providing feedback, identify misconceptions or learning errors, to craft next instructional steps that will support students in their learning, to support students' engagement in assessment of their own work (self-assessment), and develop dispositions and traits. Being clear about all aspects of purpose will enable teachers to select the most appropriate types of assessment to impact students as aligned with learning goals, standards and targets, and will provide a comprehensive set of considerations in designing and implementing assessment.
- *Types of Assessment Items and Tasks.* Like purpose, knowledge about different types of assessment items (e.g., selected-response, constructed-response, oral questioning) and tasks (e.g., performance-based, portfolio, student self-assessment, peer-assessment), is fundamental to good assessment. Teachers need an understanding of the strengths and limitations of each (which types are best for gathering evidence on different learning targets), how students are affected by different types of assessment (e.g., in test preparation), and what is needed to prepare students for the assessment.
- *Quality.* Validity, reliability, and fairness form the foundation of high-quality assessment. In particular, teachers need a thorough understanding of sources of error that contribute to the results obtained from summative assessment, especially factors intrinsic to the student, bias in scoring, and limitations from sampling content to be tested. Evidence of validity and reliability of classroom assessment (whether qualitative insights into student understanding or scores on an assessment) is more a matter of alignment between learning goals and elicited evidence and repeated convergent evidence and collaborative evaluation and judgment than calculating correlations among different measures.
- *Content and Pedagogical Knowledge.* A key element of effective classroom assessment is being able to apply appropriate models of learning in a field (Pellegrino et al., 2001; Shepard, 2006). This requires in depth content and pedagogical knowledge in the domain to enable teachers to deconstruct standards, know what is most important to assess, help students understand learning

targets, and anticipate and plan for student misconceptions (Schoenfeld, 2017). This leads to improved alignment between what is most important for students to learn with assessment, and enables teachers to provide clear explanations of targets and criteria.

Foundational influences

- *Assessment Environment.* Similar to other aspects of classroom climate or environment, teachers establish norms and expectations for summative and formative classroom assessment. The resulting classroom assessment environment influences how students prepare for, take, and respond to assessment. A positive environment, one that values the formative purpose of assessment, uses authentic, meaningful and relevant tasks, engages students in self-assessment, promotes the use of feedback, and values an expectation to learn from being wrong, will enhance learning and motivation.
- *Student Perceptions.* Student perceptions of assessment include their beliefs and feelings about the assessment experience. These perceptions, occurring prior to, during, and after a summative or formative assessment event, are key elements in determining the meaning of the performance and teacher feedback, as well as subsequent student behavior. While a relatively new topic of research in the assessment field, extensive work in other fields has established the central role of perceptions in mediating meaningfulness and behavior. For example, students who exert little effort in preparing for an assessment (perhaps because they know it will not be challenging), may respond indifferently to good performance compared to students who exert much effort; novel forms of assessment are perceived as more interesting and engaging; and emotional responses to assessment-based achievement can vary from hope, pride, and contentment, to anxiety, anger, and hopelessness, each of which impacts the receptivity of feedback, attributions, and further behavior (Vogel and Pekrun, 2016). Student perceptions also play a significant role in sociocultural dynamics within the classroom assessment environment.
- *Student Engagement.* Student involvement in assessment is an essential aspect of high-quality classroom assessment. Research shows how engagement in formative assessment is related to the development of self-regulation and self-determinism, and engagement in summative classroom assessment is related to goal orientation and self-efficacy (Andrade and Heritage, 2018). Student engagement and involvement is important so that assessment is a shared experience, rather than something that is done to them. Students can be included in many aspects of assessment that enhance their understanding of what needs to be learned (targets), how criteria are used to evaluate performance (e.g., using criteria to distinguish between surface and deep understanding), and how feedback can be used to move from current to targeted levels of proficiency. Engagement is enhanced with both self-assessment and peer-assessment. Each involves the student in thinking that promotes better understanding of current levels of proficiency, what further learning is needed, and how feedback from formative assessment can be used in self-regulation.

Action

- *Designing and Selecting Appropriate Assessments.* Teachers need to know the strengths and limitations of different types of summative and formative assessments (e.g., oral, traditional tests, performance-based, portfolio, student self-reports of dispositions, student self-assessment), as well as principles for designing high quality items and tasks and formatting (e.g., spacing, use of diagrams, computer-based formats). Overall, students respond positively to appropriate, high quality assessment, and tend to discount poorly constructed items or what they perceive as hastily constructed tasks. Furthermore, when students believe an assessment is fair and challenging, preparation effort is greater and more appropriate attributions are formed from performance. It is not simply a matter of applying rules for what makes items good – it is just as important to be aware of item difficulty and student perceptions of different types of assessments.
- *Administration.* A widely recognized, key element to successful summative assessment is knowing how to appropriately administer the test, assignment or task. Setting the right climate, knowing how to prepare students, determining time frame, responding to questions, preparing directions, and other administrative details are important. Students also need to understand the importance of responding honestly to informal questions about their understanding and to self-assessments for effective formative assessment. Honesty is in part developed as students see that teachers respond constructively to those admissions and try alternative strategies to support student learning.
- *Scoring.* Virtually all standards document and lists of competencies recognize the need for knowing how to score student performance. While scoring selected-response items is straightforward, the use of criteria and rubrics for constructed-response short answer and essay items, performance assessments and portfolios requires skill in developing clearly worded rubrics that show how specific criteria develop across increasing levels of sophistication and selecting exemplars of work to illustrate each level. For summative assessment, teachers need to know how to calculate composite scores. For formative assessment scores are accompanied by teacher feedback that identifies areas for improvement.

Consequences

- *Instruction.* Assessment informs instruction in a number of ways. Teacher questioning as part of formative assessment, based on an interpretation of student performance, can include use of a specific sequence of questions, questions that scaffold learning progressions, and appropriately challenging questions that can be directed to individual students or to groups. What do students understand and/or misunderstand? What new materials or approaches to learning should students engage in? Are further

examples needed? As a result of summative assessments, is there a need for individual tutoring, review of content, or different assignments? From evidence provided, what further instruction is needed to deepen understanding? Is additional modeling of a skill needed? What common misconceptions were revealed? What are implications for instruction for future groups of students? Teachers need to know how to make conscious, well-planned connections between assessment data and subsequent teaching and learning experiences that move students toward greater proficiency.

- *Feedback.* Appropriate feedback can have powerful positive impacts on student performance, particularly from formative assessment, as well as on motivation (Ruiz-Primo and Brookhart, 2018). Characteristics of effective feedback are well documented; integrating this knowledge as part of classroom assessment literacy has great potential to benefit students. Beyond knowing how to provide effective feedback, it is also essential for teachers to understand how students respond to teachers' responses to their work (William and Leahy, 2015).
- *Learning.* The consequences of assessment on learning comprise a key focus of recent efforts to define classroom assessment literacy. This includes how assessment practices influence student learning throughout the assessment process. Learning is affected by how students prepare for different types of assessments (e.g., selected-response compared to constructed-response) on different types of learning targets (deep understanding or surface knowledge), level of difficulty, and the amount learned. Learning is impacted as students do the assessment, as well as what they learn from their results and feedback. Learning is not restricted to curriculum-based targets. There is good evidence that through the process of assessment students can learn important self-regulation skills (Andrade and Heritage, 2018), positive self-efficacy, responsibility, and persistence in face of difficulties and barriers (McMillan, 2018).
- *Motivation.* While there has been much emphasis on assessment for learning, it is also important to understand how assessment impacts student motivation in the classroom (e.g., preparing for and taking tests), as well as how the development of motivational traits is influenced by the assessment experience. Extensive research in achievement motivation, self-regulation, attributions and goal orientation, for example, show how assessment performance is related to the development of self-efficacy, intrinsic motivation, responsibility, internal attributions, and risk-taking. Recent research in the relationship between self-regulation and assessment is especially persuasive (Andrade and Brookhart, 2016; Andrade and Heritage, 2018). Through self-assessment and opportunities to plan their learning based on proficiency in relation to learning targets, monitor progress, and evaluate their performance (especially when unsuccessful), students enhance their self-regulation skills.
- *Grading and Reporting.* Summative classroom assessment is used for grading and reporting. Given the long-standing variations in teachers' use of different factors in how they grade (Brookhart et al., 2016) teachers need a full understanding of how grading is done fairly and accurately, and how it affects students. For example, including a score of zero has an overly negative impact on grades due to unequal scale intervals, and averages can mask improvement and give undue influence to isolated, extreme scores.

These components have been selected to help educators frame what is most important about what teachers need to know and do when implementing classroom assessment. Compared to what would be appropriate for assessment literacy or data literacy more broadly, the conceptualization does not include the need to know about, interpret and use standardized tests. The contention is that classroom assessment has emerged as much more than assessment in the classroom, or the application of measurement principles to student assessment. Classroom assessment is now conceptualized as a thoughtful and complex process, focused as much on improving student performance as documenting it.

References

- Amrein, A., Berliner, D., 2003. The testing divide: new research on the intended and unintended impact of high-stakes testing. *Peer Rev.* 5 (2), 31–32.
- Andrade, H.L., Brookhart, S.M., 2016. The role of classroom assessment in supporting self-regulated learning. In: Laveault, D., Allal, L. (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation*. Springer, Heidelberg, pp. 293–309.
- Andrade, H.L., Cizek, G.J. (Eds.), 2010. *Handbook of Formative Assessment*. Routledge, New York.
- Andrade, H.L., Heritage, M., 2018. Using Formative Assessment to Enhance Learning, Achievement, and Academic Self-Regulation. Routledge, New York.
- Black, P., William, D., 1998. Assessment and classroom learning. *Assessment in Education* 5 (1), 103–110.
- Brookhart, S.M., 2011. Educational assessment knowledge and skills for teachers. *Educ. Meas.* 30 (1), 3–12.
- Brookhart, S.M., 2017. *How to Give Effective Feedback to Your Students*. ASCD, Arlington.
- Brookhart, S.M., 2020. Feedback and measurement. In: Brookhart, S.M., McMillan, J.H. (Eds.), *Classroom Assessment and Educational Measurement*. Routledge, New York, pp. 63–78.
- Brookhart, S.M., DePascale, C.A., in press. Assessment to inform teaching and learning. *Educational Measurement* (fifth ed.).
- Brookhart, S.M., Guskey, T.R., 2019. Reliability in grading and grading scales. In: Guskey, T.R., Brookhart, S.M. (Eds.), *What We Know About Grading: What Works, what Doesn't, and What's Next*. ASCD, Alexandria, VA, pp. 13–31.
- Brookhart, S.M., McMillan, J.H., 2020. *Classroom Assessment and Educational Measurement*. Routledge, New York.
- Brookhart, S.M., Guskey, T.R., Bowers, A.J., McMillan, J.H., Smith, J.K., Smith, L.F., Stevens, M.T., Welsh, M.E., 2016. A century of grading research: meaning and value in the most common educational measure. *Rev. Educ. Res.* 68, 803–848.
- Brown, G.T.L., Harris, L.R. (Eds.), 2016. *Handbook of Human and Social Conditions in Assessment*. Routledge, New York.
- Cizek, G.J., Andrade, H.L., Bennett, R.E., 2019. Formative assessment: history, definition, and progress. In: Andrade, H.L., Bennett, R.E., Cizek, G.J. (Eds.), *Handbook of Formative Assessment in the Disciplines*. Routledge Taylor and Francis, New York, pp. 3–19.
- DeLuca, C., LaPointe-McEwan, D., Luhanga, U., 2016. Teacher assessment literacy: a review of international standards and measures. *Educ. Assess. Eval. Account.* 28 (3), 251–272.
- DeLuca, C., Coombs, A., MacGregor, S., Rasooli, A., 2019. Toward a differential and situated view of assessment literacy: studying teachers' responses to classroom assessment scenarios. *Front. Educ.* 4, 94. <https://doi.org/10.3389/educ.2019.00094>.

- Guskey, T.R., Brookhart, S.M., 2019. What We Know About Grading: What Works, What Doesn't, and What's Next. ASCD, Alexandria.
- Heritage, M., 2018. Making assessment work for teachers. *Educ. Meas.* 37 (1), 39–41.
- Joint Committee on Standards for Educational Evaluation, 2003. The Student Evaluation Standards. Corwin Press, Thousand Oaks.
- Joint Committee on Standards for Educational Evaluation, 2015. The classroom assessment standards for preK-12 teachers. Retrieved from the Joint Committee website. www.jcsee.org.
- Looney, A., Cumming, J., van Der Kleij, F., Harris, K., 2017. Reconceptualizing the role of teachers as assessors: teacher assessment identity. *Assess. Educ. Princ. Pol. Pract.* 25 (4), 1–26. <https://doi.org/10.1080/0969594X.2016.1268090>.
- McMillan, J.H., 2018. Classroom Assessment: Principles and Practice that Enhance Student Learning and Motivation. Routledge, New York.
- Pastore, S., Andrade, H.L., 2019. Teacher assessment literacy: a three-dimensional model. *Teach. Teach. Educ.* 84, 128–138.
- Pellegrino, J.W., Chudowsky, N., Glaser, R., 2001. Knowing what Students Know: The Science and Design of Educational Assessment. National Academies Press, Washington.
- Ruiz-Primo, M.A., Brookhart, S.M., 2018. Using Feedback to Improve Learning. Routledge, New York.
- Schoenfeld, A.H., 2017. On learning and assessment. *Assess Educ. Princ. Pol. Pract.* 24 (3), 369–378. <https://doi.org/10.1080/0969594X.2017.1336986>.
- Shepard, L.A., 2006. Classroom assessment. In: Brennan, R.I. (Ed.), *Educational Measurement*, fourth ed. Praeger, Westport, pp. 623–646.
- Shepard, L.A., 2019. Classroom assessment to support teaching and learning. *Annals Am. Acad.* 683, 183–200.
- Shepard, L.A., Penuel, W.R., Pellegrino, J.W., 2018. Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educ. Meas.* 37 (1), 21–34.
- Standards for teacher competence in educational assessment of students, 1990. National Council on Measurement in Education, Washington, DC. American Federation of Teachers, Washington, DC; National Education Association, Washington, DC.
- Stiggins, R.J., 1991. Assessment literacy. *Phi Delta Kappan* 72 (7), 534–539.
- Vogel, E., Pekrun, R., 2016. Emotions that matter to achievement: student feels about assessment. In: Brown, G.T.L., Harris, L.R. (Eds.), *Handbook of Human and Social Conditions in Assessment*. Routledge, New York, pp. 111–128.
- William, D., Leahy, S., 2015. Embedding Formative Assessment: Practical Techniques for K 12 Educators. Learning Sciences International, West Palm Beach.
- Xu, Y., Brown, G.T.L., 2016. Teacher assessment literacy in practice: a reconceptualization. *Teach. Teach. Educ.* 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>.
- Xu, Y., He, L., 2019. How pre-service teachers' conceptions of assessment change over practicum: implications for teacher assessment literacy. *Front. Educ.* 4, 145. <https://doi.org/10.3389/feduc.2019.00145>.

Analysis of variance: univariate and multivariate approaches

Patrick J. Rosopa and Bruce M. King, Clemson University, Clemson, SC, United States

© 2023 Elsevier Ltd. All rights reserved.

This is an update of B.M. King, Analysis of Variance, In International Encyclopedia of Education (Third Edition), edited by Penelope Peterson, Eva Baker and Barry McGaw, Elsevier Ltd., 2010, <https://doi.org/10.1016/B978-0-08-044894-7.01306-3>.

Introduction	529
The logic of ANOVA	530
Computation of F : an example	531
Interpreting a significant F value	532
The repeated-measures design	532
Two-way ANOVA	533
Planned comparisons	534
References	535
Further reading	535

Glossary

Alpha The probability of a Type I error (rejection of a true null hypothesis)

Analysis of variance (ANOVA) A statistical test used to compare the means of three or more populations

Factor An Independent (treatment) variable

F ratio $S^2_{\text{bet}}/S^2_{\text{W}}$; between-groups variance estimate divided by the within-groups variance estimate

Grand mean The mean of all scores in all the treatment conditions

Interaction effect In analysis of variance, the joint effect of two or more independent variables that cannot be predicted from the main effects

Level(s) Different values of an Independent variable (factor)

Main effect In analysis of variance, the mean differences among the levels of one factor averaged across the levels of the other factor

Multivariate analysis of variance (MANOVA) A multivariate version of analysis of variance used to compare means among groups on multiple dependent variables

Omega-squared ω^2_{bet} ; a measure of effect size for analysis of variance

Orthogonal comparisons Comparisons that are unique and do not overlap

Planned comparisons Statistical tests used in place of ANOVA when the researcher knows in advance which comparisons are of interest

Post hoc comparisons Statistical tests performed after obtaining a significant value of F that indicates which means are significantly different

Power The probability of rejecting a false null hypothesis

Sum of squares (SS) Sum of the squared deviations from the mean

Type I error Rejection of a true null hypothesis

Introduction

Imagine that you are superintendent of a school district and you wish to compare the mathematics competency of students enrolled in seven high schools. You randomly select 30 seniors from each school and give them a mathematics achievement test. Your null hypothesis is that the mean mathematics achievement scores in the seven populations (the seven high schools) are equal, that is, there is no difference in scores among the schools. How should you analyze the results? If you had decided to test for differences between all pairs of means with use of a t -test, you would have to conduct 21 tests. Suppose that you had decided in advance to set the probability of a Type I error (rejection of a true null hypothesis) at 0.05. This means that when testing for a difference between two population means, in 20 replications of the same study (drawing independent random samples of the same size each time) you would expect the results of one of the t tests to be statistically significant even if there is no difference between the population means. Although for any one test, alpha (probability of a Type I error) is 0.05, with 21 independent tests, the probability of making at least one Type I error is $p = 1 - (0.95)^{21} = 1 - 0.34 = 0.66$. Thus, if you found a statistically significant difference for one or two of your 21 tests, there is a high probability that the result(s) was due to a Type I error rather than to a true difference between groups.

You need a test that allows you to make multiple comparisons among means while holding alpha steady at a preselected level. To this end, Ronald Fisher (1890–1962) developed analysis of variance (ANOVA) in the 1920s. An interesting feature of ANOVA is that although it is a test for differences among population means, it can be conducted without calculating the value of any of the sample means. It may seem odd that we are interested in testing for differences in population means and yet we use a statistical procedure referred to as analysis of “variance.” We will learn below that the test simply involves the ratio of two variances.

The logic of ANOVA

ANOVA involves a categorical independent variable, often referred to as a factor or treatment, and a continuous dependent variable (or outcome). When testing for differences among population means on the dependent variable across the levels of the independent variable, variation in scores can be due to two things: chance and treatment effect. Here, a treatment effect refers to true differences among population means. When there is no treatment effect, all population means are the same. As the scores in any one sample are drawn randomly from the same population, the scores can differ from one another, and from the sample mean, only due to chance. However, scores in different samples, as well as sample means, can differ from one another due to chance, but also because the experimental treatment has an effect. If there is no treatment effect, scores in different groups will differ only because of chance. Thus, when the null hypothesis (there is no treatment effect) is true, the variation of scores around the sample means within each sample ($X - \bar{X}$) and the variation of sample means from the grand mean ($\bar{X} - \bar{\bar{X}}$; the grand mean, $\bar{\bar{X}}$, is the mean of all scores) will both be a reflection of variation in scores due to chance factors.

For any study of two or more samples, the amount by which a score differs from the grand mean ($X - \bar{\bar{X}}$) can be partitioned into two parts: the amount by which the score differs from its own sample mean ($X - \bar{X}$) and the amount by which the sample mean differs from the grand mean ($\bar{X} - \bar{\bar{X}}$). The partition of $X - \bar{\bar{X}}$ is represented by the following equation:

$$X - \bar{\bar{X}} = (X - \bar{X}) + (\bar{X} - \bar{\bar{X}}) \quad (1)$$

As the name suggests, in ANOVA, we are going to work with variance estimates. Variance estimates are calculated by dividing the sum of the squared deviations from the mean (often called sum of squares, and abbreviated SS) by the degrees of freedom (df). We can calculate the sum of squares total (SS_{total}) for two or more samples by calculating the deviation from the grand mean for every score, squaring each deviation score, and then summing (Σ) all the squared deviations from the grand mean:

$$SS_{\text{total}} = \sum^{\text{all}} (X - \bar{\bar{X}})^2$$

Look again at Eq. (1). You can see that we could also calculate SS_{total} by squaring $(X - \bar{X}) + (\bar{X} - \bar{\bar{X}})$:

$$\sum^{\text{all}} (X - \bar{\bar{X}})^2 = \sum^{\text{all}} (X - \bar{X})^2 + \sum^k n (\bar{X} - \bar{\bar{X}})^2$$

where:

n = the number of scores in a particular group

k = the number of samples

When we square $(a + b)$ there is a middle term, here $2(\bar{X} - \bar{\bar{X}}) \sum (X - \bar{X})$, but in this case, it is always zero. What has been accomplished is the partitioning of the total sum of squares into two parts: (1) $\sum^{\text{all}} (X - \bar{X})^2$, called SS_{within} (or SS_{error}), a reflection only of chance factors when drawing samples, and (2) $\sum^k n (\bar{X} - \bar{\bar{X}})^2$, called SS_{between} , a reflection of both chance factors and treatment effect (when the population means are different). Thus,

$$SS_{\text{total}} = SS_{\text{within}} + SS_{\text{between}}$$

We can now divide the separate sums of squares by their respective degrees of freedom to obtain variance estimates. The degrees of freedom for SS_{total} is the total number of scores minus one: $n_{\text{total}} - 1$. For SS_{within} (abbreviated SS_{w}), the degrees of freedom is calculated by adding $n - 1$ for each sample: $n_{\text{total}} - k$. For SS_{between} (abbreviated SS_{bet}), degrees of freedom is the number of samples minus one: $k - 1$.

The end result is two variance estimates. One is the within-groups variance estimate, $s_{\text{w}}^2 (= SS_{\text{w}} / df_{\text{w}})$, which is an estimate of the variance due to chance. It is sometimes called mean square within or mean square error. The other is the between-groups variance estimate, $s_{\text{bet}}^2 (= SS_{\text{bet}} / df_{\text{bet}})$, which is an estimate of the variance due to chance plus any treatment effect. It is sometimes called mean square between. When there is no treatment effect, both are estimates of the same thing. Thus, when the null hypothesis

is true (i.e., there is no treatment effect), the ratio s_{bet}^2/s_w^2 , called the F ratio, should be about 1.0. As noted above, because s_{bet}^2 is an estimate of the variance due to chance plus any treatment effect, when the null hypothesis is false (i.e., there is a treatment effect), s_{bet}^2 will exceed s_w^2 , and the greater the treatment effect, the more is the amount by which s_{bet}^2 will exceed s_w^2 . Due to sampling variation, we expect the value of the F ratio to be different with every replication of a study (employing new randomly drawn samples of the same size). However, if we set $\alpha = 0.05$ and obtain a value of F that would have occurred by chance less than 5% of the time (if the null hypothesis were true), we reject the null hypothesis and conclude that there is a treatment effect (i.e., the treatment effect reliably exceeds effects due to random chance). Finding the critical value for F will be discussed in the example below.

Computation of F : an example

Let us look at an example. Suppose that we wish to compare two new methods for teaching mathematics with the standard method. We randomly draw a sample of ten students for each method and use each method for 1 year, at the end of which we administer a mathematics achievement test to each student. The hypothetical results are as follows:

Standard method (X_1)	Method 2 (X_2)	Method 3 (X_3)
68	85	62
70	90	76
75	80	74
65	78	70
55	82	58
80	86	78
72	92	74
78	76	80
60	94	65
74	84	75
$\sum X_1 = 697$	$\sum X_2 = 847$	$\sum X_3 = 712$
$\bar{X}_1 = 69.7$	$\bar{X}_2 = 84.7$	$\bar{X}_3 = 71.2$

$$\bar{\bar{X}} = 2256 / 30 = 75.2$$

We now calculate SS_{total} , SS_w , and SS_{bet} :

$$\begin{aligned} SS_{\text{total}} &= (68 - 75.2)^2 + (70 - 75.2)^2 + \dots (65 - 75.2)^2 \dots \\ &\quad + (75 - 75.2)^2 = 2722.8 \\ SS_w &= (68 - 69.7)^2 + (85 - 84.7)^2 + (62 - 71.2)^2 \dots \\ &\quad + (70 - 69.7)^2 + (90 - 84.7)^2 + (76 - 71.2)^2 \\ &\quad \cdot \quad \cdot \quad \cdot \\ &\quad \cdot \quad \cdot \quad \cdot \\ &\quad \cdot \quad \cdot \quad \cdot \\ &\quad + (74 - 69.7)^2 + (84 - 84.7)^2 + (75 - 71.2)^2 = 1357.8 \end{aligned}$$

$$SS_{\text{bet}} = 10(69.7 - 75.2)^2 + 10(84.7 - 75.2)^2 + 10(71.2 - 75.2)^2 = 1365$$

There are formulas to calculate SS_{total} , SS_w , and SS_{bet} that are easier to use for hand calculations than using deviations from means, but they do not provide you with a basic understanding of ANOVA (see King et al., 2018). (It is these formulas that allow you to compute F without ever having calculated a single mean.) It is customary to place the results in a summary table, often called an ANOVA table:

Source	SS	df	s^2	F
Between groups	1365	2	682.5	13.571
Within groups	1357.8	27	50.29	
Total	2722.8	29		

Be sure to place between groups first, above within groups, to set up the proper ratio. As a check on your work, SS_{bet} plus SS_w should equal SS_{total} , and df_{bet} plus df_w should equal df_{total} (s_{bet}^2 plus s_w^2 does not equal s_{total}^2). Note that textbooks also refer to the variance estimates (s^2) as mean squares (MS) because they represent a type of average of a set of squared numbers.

The calculated F ratio is a random variable that follows a particular probability distribution known as the F distribution. In order to determine if our obtained value of F is statistically significant, we must use a table for the F distribution, found in any introductory statistics textbook. The table provides critical values of F . When our obtained value of F is equal to or greater than the critical value, it indicates that our result is statistically significant (i.e., has a low probability of occurring by chance, assuming the null hypothesis is true). The F distribution varies as a function of two parameters: df_{bet} and df_{w} . Critical values of F are found at the intersection of df for the numerator (2 in our example) and df for the denominator (27 in our example), and in this case, the critical values are 3.35 for $\alpha = 0.05$ and 5.49 for $\alpha = 0.01$. Thus, in our example, the probability of a Type I error was less than 0.01, that is, the result is statistically significant. In addition to your F statistic and p -value, you should also provide a measure of effect size when reporting results. One commonly used measure for independent-groups ANOVA is omega-squared (ω_{bet}^2) (see King et al., 2018). It gives us a population-based, rather than a sample-based, estimate of how much of the proportion of the variance in the dependent variable (in this case, mathematics achievement) was attributed to the different levels of treatment (in this case, teaching method). The formula for estimating omega-squared ($\hat{\omega}_{\text{bet}}^2$) can be written as:

$$\hat{\omega}_{\text{bet}}^2 = \frac{SS_{\text{bet}} + (k - 1)s_{\text{w}}^2}{SS_{\text{total}} + s_{\text{w}}^2}$$

where, in our example, $k = 3$. The other quantities can be obtained from the ANOVA table. Here, $\hat{\omega}_{\text{bet}}^2 = 0.53$.

Interpreting a significant F value

Independent-groups ANOVA can be used with two samples, in which case F is the square of the t -statistic that compares the two sample means. A statistically significant result indicates that one population mean is either less than or greater than the other. What does it mean when we obtain a statistically significant value of F for three or more samples? In this case, it tells us only that there is a difference among the population means. It does not tell us the manner in which they differ. For three groups, all three population means could be different from one another, or one could be greater than the other two, etc. To determine which means are significantly different from others, we normally use *post hoc* (*a posteriori*) comparisons. Some of the most commonly used tests are Duncan's multiple-range test, the Newman-Keuls test, Tukey's Honestly Significant Difference (HSD) test, and the Scheffé test. Duncan's test is the least conservative with regard to Type I error and the Scheffé test is the most conservative. An explanation of these tests is beyond the scope of this article, but most textbooks will provide a full explanation of one or more of them. However, before you can use any of them you must first have obtained a significant value of F . In our example, all four *post hoc* tests would reveal that teaching method 2 is superior to the other two methods, which did not significantly differ from one another.

Although the variance decomposition described above always remains true in ANOVA, some underlying assumptions are required when testing hypotheses. The first is that the populations from which the samples are drawn are normally distributed. Moderate departure from the normal bell-shaped curve does not greatly affect the outcome, especially with large-sized samples (Glass et al., 1972). However, results are much less accurate when populations of scores are very skewed or multimodal (Tomarken and Serlin, 1986), which is frequently the case in the behavioral sciences (Micceri, 1989). In this case, you should consider using the Kruskal-Wallis test, an assumption-free (nonparametric) test for the independent-groups design (see King et al., 2018). This is especially true when using small samples. A second assumption is that of homogeneity of variance; that is, the variances in the populations from which samples are drawn are the same (Rosopa et al., 2013). A variety of statistical tests exist to determine whether this assumption is violated (Rosopa et al., 2013, 2016). This can be a major problem when variances differ considerably across samples and is less of a problem if you use samples that are of the same size (Milligan et al., 1987; Tomarken and Serlin, 1986). Fortunately, existing research has found that alternative procedures (e.g., weighted least squares regression and the use of heteroscedasticity-consistent covariance matrices) are relatively unaffected by the violated assumption regardless of differences in sample sizes and are readily available in modern statistical software (Rosopa et al., 2018, 2013).

Although ANOVA is useful when studying a single dependent variable (e.g., mathematics achievement), it can be extended to designs where data is collected on more than one dependent variable. For example, an educational researcher may be interested in the effects of three modes of teaching (face-to-face, online, or hybrid) on three dependent variables—mathematics achievement, reading comprehension, and school self-efficacy. With students randomly assigned to treatments and measured on the three dependent variables at the end of the school year, the researcher could analyze the data using a multivariate version of ANOVA known as multivariate analysis of variance (MANOVA) (Rencher and Christensen, 2012). While ANOVA relies on the F ratio to compare the population means of three or more groups, in MANOVA, there are four statistical tests available to compare population means on multiple dependent variables across three or more groups. Rencher and Christensen (2012) describe the general situations in which one multivariate test tends to outperform the others.

The repeated-measures design

ANOVA can also be used with the repeated-measures design. For example, assume that 20 students had achievement scores under three conditions (pretest, posttest, 6-month follow-up). This simple design is also referred to as a one-way within-subjects design.

The major difference from the independent-groups design is that SS_w is partitioned into two parts: (1) SS_{subjects} , which is a reflection of the variability in scores resulting from individual differences, and (2) $SS_{\text{residuals}}$, a reflection of the variability in scores due to chance. Although three variance estimates can be calculated (s^2_{subjects} , s^2_{bet} , and s^2_{resid}), normally we are interested only in the F ratio $s^2_{\text{bet}}/s^2_{\text{resid}}$ to determine if there is a statistically significant difference among groups. (We are usually not interested in whether or not individuals differ.)

An additional assumption with this design, when there are more than three conditions, is sphericity (King et al., 2018). It requires that the population variances of the differences between the levels of the within-subjects factor are the same. When there are only two conditions, there is only one set of differences, and therefore, only one variance. Thus, the assumption cannot be violated.

When you have three or more conditions and the sphericity assumption is violated, one alternative data-analytic approach is the use of an F ratio with an adjusted- df that accounts for violating the assumption. Of the adjusted- df approaches that are available for the univariate F ratio, the method by Greenhouse-Geisser strikes a good balance between controlling Type I error and having adequate statistical power. In other words, when the null hypothesis is actually false, it has a high likelihood of detecting a treatment effect compared to other approaches (Maxwell et al., 2018). Another alternative is the use of MANOVA which does not require the sphericity assumption. The MANOVA approach for analyzing repeated measures data actually focuses on the differences between the levels of the within-subjects factor. For example, with three levels, the difference between level 1 vs. level 2 and level 2 vs. level 3 is analyzed in MANOVA. Although the MANOVA approach for analyzing data from a repeated measures design does not require the sphericity assumption, if the assumption is not violated, the usual F ratio tends to have greater statistical power than MANOVA (Maxwell et al., 2018).

Two-way ANOVA

In the previous example, we examined the effects of three different teaching methods on mathematics achievement scores using an independent-groups design. There was only one treatment variable—teaching method. However, ANOVA allows us to study two or more treatment variables (called factors) simultaneously. Suppose, for example, that we were also interested in whether the amount of time devoted to teaching mathematics affected achievement scores. Rather than conducting two separate studies, we can study both factors (teaching method and time spent teaching) simultaneously using a two-factor ANOVA design. If we study all three teaching methods (referred to as levels of the factor teaching method) and two different durations of time spent teaching (two levels; e.g., 45 min per day vs. 90 min per day), we have a 3×2 design. Let us suppose that we select eight students randomly for each of the six independent conditions (standard teaching method for 45 min per day, standard teaching method for 90 min per day, etc.) and after 1 year we obtain the following mathematics achievement scores:

		Teaching method			
Teaching time	Standard method	Method 2	Method 3		
45 min	60	73	62		
	80	87	74		
	68	76	63		
	72	84	73		
	64	83	64		
	76	77	72		
	69	81	67		
	71	79	69		
	$\bar{X} = 70$	$\bar{X} = 80$	$\bar{X} = 68$		
90 min	65	92	63		
	81	106	77		
	68	103	70		
	78	95	68		
	70	97	70		
	76	101	72		
	72	99	67		
	74	99	73		
	$\bar{X} = 73$	$\bar{X} = 99$	$\bar{X} = 70$		
	$\bar{X}_{\text{std}} = 71.5$	$\bar{X}_2 = 89.5$	$\bar{X}_3 = 69$		$\bar{\bar{X}} = 76.67$

The six combinations of rows and columns are called cells. The six cell means, two row means, three column means, and the grand mean are provided. In two-way ANOVA, we want to know if there is a main effect for any of the factors, or, in other words, if there are differences in the means of the levels of one factor averaged across the levels of the other factor. In our example, there are

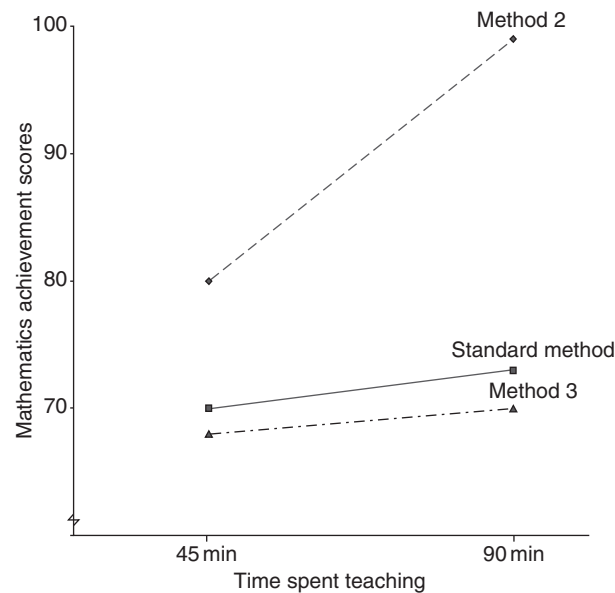


Fig. 1 Hypothetical scores in a 3×2 factorial study of the effects of teaching method and time spent teaching on mathematics achievement scores.

two possible main effects and the essential questions can be phrased as: (1) Is there an overall influence of teaching method (for both the 45 min and 90 min conditions), and (2) is there an overall influence of time devoted to teaching (across all three teaching methods)?

There is an equally, if not more, important question that we can examine with a factorial design: Is there an interaction effect—are the differences among the levels of one factor the same for all levels of the other factor? To better understand this, let us first look at the results of our study in graph form (Fig. 1).

In our hypothetical study, if an increase in time spent teaching mathematics had an equal effect for all three teaching methods, the three lines would be parallel. A sizable gap between the lines indicates a possible main effect for teaching method. In our graph, we see evidence for a main effect, but the difference between method 2 and the other two teaching methods depends on the amount of time used in the classroom. Method 2 results in moderately better mathematics achievement scores when used for 45 min per day, but dramatically better scores when students are taught mathematics by this method for 90 min per day. If there had been a main effect for teaching method (method 2 better than the other two) without an interaction effect, the differences between method 2 and the others would have been the same for both levels of the other factor (teaching time per day). In other words, the lines in Fig. 1, would be approximately parallel to one another. A significant interaction effect tells us that before we can interpret the effects of a factor, we must examine that factor at each level of the other factor. This will, of course, involve examining cell means. Three-factor experimental designs (e.g., teaching method $\times$ time spent teaching $\times$ gender) and even higher-order designs are possible, but interpretation of interaction effects becomes more difficult.

When calculating two-way ANOVA, SS_{total} is partitioned into four parts. SS_{wc} (within cell) is the equivalent of SS_{w} (within groups) in the one-factor design. SS_{bet} is split into three parts: SS_{rows} , SS_{columns} , and $SS_{\text{rows} \times \text{columns}}$. To see how to calculate two-way ANOVA, see King et al. (2018). In our example, we obtain a significant effect for rows [$F(1,42) = 30.78$, $p < 0.001$], columns [$F(2,42) = 80.21$, $p < 0.001$], and rows $\times$ columns interaction [$F(2,42) = 14.59$, $p < 0.001$]. As at least one of our main effects was statistically significant, we may now conduct a *post hoc* test to see which levels differed from the others (e.g., method 2 vs. the other two).

As we learned above, MANOVA can be used for one-way ANOVA—both independent groups and repeated-measures designs. Not surprisingly, it can also be extended to two-way designs. Using the current example, we could have a 3×2 MANOVA where we measure not only mathematics achievement, but also math self-efficacy. This would allow us to test for both main effects (teaching method and time spent teaching) and a potential interaction effect but on two dependent variables simultaneously.

Planned comparisons

When conducting any statistical test one wants the test to have considerable power, that is, a high probability of rejecting a false null hypothesis. *Post hoc* tests allow us to make all possible pairwise comparisons, but to protect us from making Type I errors, the differences between means must be large enough to be declared statistically significant. If one's study is exploratory in nature, then ANOVA is a very good statistical test. However, if one knows in advance of conducting a study which comparisons are important

to him or her, one should use planned (*a priori*) comparisons in place of ANOVA and *post hoc* comparisons (Rosnow and Rosenthal, 1989; Wilkinson and The Task Force on Statistical Inference, 1999; Winer et al., 1991). To do so, one should conduct independent-groups *t* tests using s^2_w (calculated from all groups), but to protect against Type I errors, all the comparisons must be orthogonal (unique, with no overlap). When comparing k groups, there are only $k-1$ orthogonal comparisons. In our example, we might be interested in the two comparisons standard method versus (method 2 + method 3) and method 2 versus method 3. Unlike *post hoc* comparisons, one does not need a significant overall *F* in order to use planned comparisons. Even Ronald Fisher recognized the value of planned comparisons:

When the [*F*] test does not demonstrate significant differentiation, much caution should be used before claiming significance for special comparisons. Comparisons, which the experiment was designed to make, may, of course, be made without hesitation (Fisher, 1949: 57).

References

- Fisher, R.A., 1949. The Design of Experiments, fifth ed. Oliver and Boyd, London.
- Glass, G.V., Peckham, P.D., Sanders, J.R., 1972. Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. *Rev. Res. Educ.* 42, 237–288.
- King, B.M., Rosopa, P.J., Minium, E.W., 2018. Statistical Reasoning in the Behavioral Sciences, seventh ed. Wiley, Hoboken, NJ.
- Maxwell, S.E., Delaney, H.D., Kelley, K., 2018. Designing Experiments and Analyzing Data: A Model Comparison Perspective, third ed. Routledge, New York.
- Micceri, T., 1989. The unicorn, the normal curve, and other improbable creatures. *Psychol. Bull.* 105, 156–166.
- Milligan, G.W., Wong, D.S., Thompson, P.A., 1987. Robustness properties of nonorthogonal analysis of variance. *Psychol. Bull.* 101, 464–470.
- Rencher, A.C., Christensen, W.F., 2012. Methods of Multivariate Analysis, third ed. Wiley, Hoboken, NJ.
- Rosnow, R.L., Rosenthal, R., 1989. Statistical procedures and the justification of knowledge in psychological science. *Am. Psychol.* 44, 1276–1284.
- Rosopa, P.J., Schaffer, M.L., Schroeder, A.N., 2013. Managing heteroscedasticity in general linear models. *Psychol. Methods* 18, 335–351.
- Rosopa, P.J., Schroeder, A.N., Doll, J., 2016. Detecting between-groups heteroscedasticity in moderated multiple regression with a continuous predictor and a categorical moderator: a Monte Carlo Study. *Sage Open* 6, 1–16.
- Rosopa, P.J., Brawley, A.M., Atkinson, T.P., Robertson, S.A., 2018. On the conditional and unconditional Type I error rates and power of tests in linear models with heteroscedastic errors. *J. Mod. Appl. Stat. Methods* 17, eP2647.
- Tomarken, A.J., Serlin, R.C., 1986. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychol. Bull.* 99, 90–99.
- Wilkinson, L., The Task Force on Statistical Inference, 1999. Statistical methods psychology journals. Guidelines and explanations. *Am. Psychol.* 54, 594–604.
- Winer, B.J., Brown, D.R., Michels, K.M., 1991. Statistical Principles in Experimental Design, third ed. McGraw Hill, New York.

Further reading

- Agresti, A., 2018. Statistical Methods for the Social Sciences, fifth ed. Pearson, Boston, MA.
- Gravetter, F.J., Wallnau, L.B., 2008. Essentials of Statistics for the Behavioral Sciences, sixth ed. Thomson Wadsworth, Belmont, CA.
- Meyers, L.S., Gamst, G., Guarino, A.J., 2017. Applied Multivariate Research: Design and Interpretation, third ed. Sage, Thousand Oaks, CA.

Analysis of covariance: univariate and multivariate approaches

Bryan Keller and Dobrin Marchev, Teachers College, Columbia University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	536
The general linear <i>F</i> test	536
Univariate ANOVA	537
Data demonstration	538
Comparing group means	539
Univariate ANCOVA	539
Data demonstration	540
Adjusted group means	541
Covariate balance, overlap and ANCOVA adjustment	541
Assumptions and limitations	542
Multivariate ANOVA	543
Multivariate ANCOVA	545
Assumptions and limitations	545
Data demonstration	545
Discussion	546
References	547

Introduction

The analysis of covariance (ANCOVA) is an inferential statistical method for analyzing experimental data that allows for the comparison of two or more group means while controlling for one or more covariates. As Fisher, who invented the procedure, explained, the analysis of covariance “combines the advantages and reconciles the requirements of the two very widely applicable procedures known as regression and analysis of variance” (1934, p. 269). ANCOVA improves upon the analysis of variance (ANOVA) in two key ways. First, by controlling for covariates that are correlated with the outcome, the ANCOVA test of treatment efficacy (i.e., difference in group means) is typically more powerful than the analogous ANOVA-based test. Second, in situations where random assignment fails to provide adequate balance on a confounding variable, or in cases where random assignment was not implemented, ANCOVA may be used to “control for” covariates by making linear adjustments to estimated group means that can reduce or remove bias due to confounding.

The use of ANCOVA to identify average treatment effects—either in the context of a well-implemented randomized experiment to reduce error variance or in a non-randomized study to control for confounding—requires assumptions, some of which are more likely to be tenable in practice than others. In modern usage, the meaning of ANCOVA is sometimes broadened to include nonlinear transformations of covariates or treatment by covariate interactions (cf. Huitema, 2011). Our focus herein, however, will be on the use of ANCOVA in the classical sense, as a method used to test for constant treatment effects assuming linear functional forms.

Multivariate analysis of variance (MANOVA) extends ANOVA to allow for simultaneous testing for treatment effects in the presence of multiple outcome variables. For correlated outcome variables, MANOVA may provide substantially better power to detect effects than the associated univariate ANOVAs. This same rationale holds when also controlling for one or more covariates via multivariate ANCOVA (MANCOVA); thus, we find reason to recommend MANCOVA for scenarios involving multiple correlated outcome variables.

The remainder of this chapter is organized as follows. In the next section we briefly define some key terms and describe the general linear *F* test, which is used to construct hypothesis tests in ANOVA and ANCOVA. In the third section, because central concepts in ANCOVA are most meaningful when compared against analogous concepts in ANOVA, we describe the univariate one-way ANOVA model, give a data example, and discuss pairwise comparisons of group means through contrast testing. We present the univariate ANCOVA model in **Univariate ANCOVA** section, where we examine the omnibus *F* statistic in relation to ANOVA, give a data example, describe adjusted group means, and discuss the role of covariate balance in ANCOVA adjustment. We then follow a similar structure for the multivariate procedures, introducing MANOVA first and following with MANCOVA and a data demonstration. We conclude with some discussion.

The general linear *F* test

Consider a simple linear regression model for outcome *Y* and predictor *X*,

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

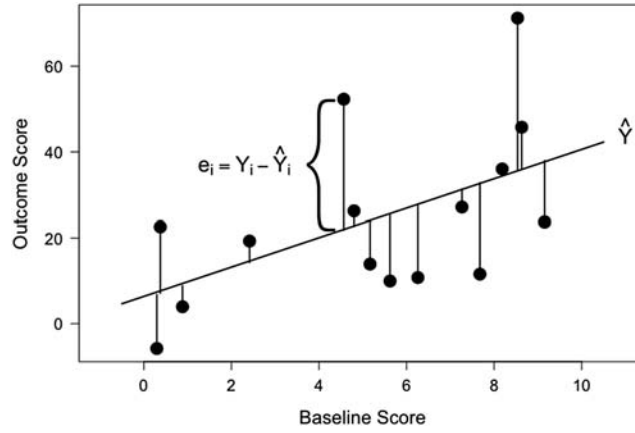


Fig. 1 Scatterplot of baseline and outcome data with fitted regression line and residuals.

where errors are independent and identically (iid) normally distributed with mean zero and constant variance; i.e., $\varepsilon_i \sim N(0, \sigma^2)$. When data are used to estimate the parameters of the model, the resulting function may be used to predict the mean of the outcome, Y , conditional on the values of the predictor, X . The outputs from this function are called *predictions* or *fitted values*, and the fitted value for the i th unit is denoted as $\hat{Y}_i$ as follows,

$$E[Y_i|x_i] = \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (2)$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are, respectively, estimates for β_0 and β_1 . The *residual* for participant i , e_i , is defined as the difference in observed and predicted values; that is, $e_i = Y_i - \hat{Y}_i$; see Fig. 1.

Sums of squared residuals may be used to create measures of model fit. In particular, sums of squared residuals give a quantitative summary of variation leftover (i.e., not explained) by the model. The *general linear F test* provides a framework for constructing hypothesis tests by comparing regression models using sums of squares (SS). As an example, to test $H_0: \beta_1 = 0$ for the simple regression model in Eq. (1), a *restricted* model is specified in which the parameter to be tested (or parameters, if a joint test is desired) is fixed to the hypothesized value. Here the restricted model would be $Y_i = \beta_0 + \varepsilon_i$. The general linear F statistic is defined as

$$F = \frac{(SS_R - SS_F)/(df_R - df_F)}{SS_F/df_F} \quad (3)$$

where SS_R and SS_F are the sums of squared residuals, and df_R and df_F are degrees of freedom, for the restricted and full models, respectively. If errors $\varepsilon_i \sim N(0, \sigma^2)$, then the statistic F will follow an F distribution under the null hypothesis with $df_R - df_F$ numerator and df_F denominator degrees of freedom.¹

Univariate ANOVA

It is helpful to use the analysis of variance (ANOVA) as a starting point from which to describe the ANCOVA model. For a single-factor experiment, the one-way ANOVA model may be described as follows,

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (4)$$

where,

- Y_{ij} is the outcome variable for participant i in group j ,
- μ is the unweighted average of group means,
- α_j is the treatment effect (i.e., $\mu_j - \mu$) for the j th group, and
- ε_{ij} is idiosyncratic (random) error for participant i in level j .

The ε_{ij} are assumed to be independent $N(0, \sigma^2)$ random variables. To uniquely define model parameters, a constraint is typically imposed such that the treatment effects add to zero, $\sum_{j=1}^a \alpha_j = 0$.² The one-way ANOVA model assumes that any systematic variability in the outcome is solely due to the treatment factor groupings, and each α_j represents the j th group's deviation from the overall

¹See Maxwell et al. (2018) and Neter et al. (1990) for texts that build up test statistics for ANOVA and ANCOVA via model comparisons using the general linear F test.

²Another possibility is to require a *weighted* sum of the to be zero. In that case, μ is a weighted average of group means. For a useful discussion of weighting schemes and implications, see Neter et al. (1990, pp. 554–556).

population mean, μ . Suppose there are a groups in total with n_j participants in the j th group. The omnibus null hypothesis for the one-way ANOVA is that all treatment effects are identically equal to zero; that is,

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$

Let $\bar{y}$ and $\bar{y}_j$ be estimates of the overall mean and group j mean, respectively. Full and restricted models and least squares estimators are shown in Table 1. The sums of squares (SS) decomposition for ANOVA is motivated by the model in Eq. (4). After subtracting the mean μ from both sides, Eq. (4) may be expressed as follows,

$$Y_{ij} - \mu = \alpha_j + \varepsilon_{ij} \quad (5)$$

and with sample analogs as

$$(y_{ij} - \bar{y}) = (\bar{y}_j - \bar{y}) + (y_{ij} - \bar{y}_j) \quad (6)$$

The three terms in parentheses in Eq. (6) are referred to as *deviations*. Squaring both sides, canceling out the middle term on the right hand side because $\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j) = 0$, and summing over all i and j leads to the sums of squares decomposition for one-way ANOVA; see Eq. (7). Note the sum of squares “total” does not represent the total squared variation in the outcome variable because the mean has been subtracted off the left hand side of the equation. Thus, SST is typically referred to as “corrected for the mean”.

$$\begin{aligned} SS_{T; \text{total (corrected)}} &= SS_{B; \text{between}} + SS_{W; \text{within}} \\ \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2 &= \sum_{j=1}^a n_j (\bar{y}_j - \bar{y})^2 + \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2 \end{aligned} \quad (7)$$

After substitution and some algebra, the F statistic for the one-way ANOVA omnibus null hypothesis may be written in terms of SS_B and SS_W as follows,

$$F = \frac{(SS_R - SS_F)/(df_R - df_F)}{SS_F/df_F} = \frac{SS_B/(a - 1)}{SS_W/(N - a)} \quad (8)$$

where N is the total number of participants. If $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$, the F statistic will follow an F distribution with $a - 1$ numerator and $N - a$ denominator degrees of freedom under the null hypothesis.

Data demonstration

Consider a hypothetical experiment in which a total of $N = 15$ participants were randomly assigned to three groups (two treatments and a control) with equal frequency (i.e., $n_1 = n_2 = n_3 = 5$). All participants were measured at baseline on a variable known to be positively correlated with the outcome. The one-way ANOVA restricted model assigns predicted values to participants based on the overall mean of the outcome variable, regardless of group membership; whereas, the full model assigns a predicted value to each participant based on their group mean. The left and right panels of Fig. 2 depict the data by group along with restricted and full models for the one-way ANOVA omnibus null hypothesis along with dotted lines tracing residuals that contribute to SS_R and SS_F .

ANOVA results are often displayed in tabular form as in Table 2, which reveals that $SS_B = 1890.1$ and $SS_W = 3710.0$. The mean squares of 945.1 and 309.8 represent, respectively, the numerator and denominator of the F statistic. Their quotient is 3.06 which, when compared with an F distribution with $3 - 1 = 2$ numerator and $15 - 3 = 12$ denominator df yields a p value of 0.084.

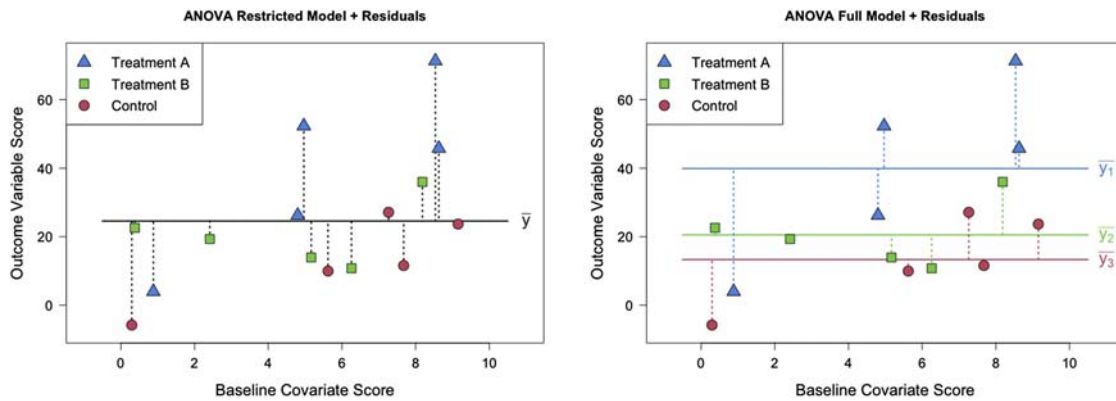


Fig. 2 ANOVA restricted (left panel) and full (right panel) model fits and residuals.

Table 1 Full and restricted models, parameters, and least-squares estimators for the one-way ANOVA omnibus test.

ANOVA	Model	Parameters	Least-squares estimators
Full	$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$	μ $\alpha_j, \alpha_2, \dots, \alpha_J$	$\hat{\mu} = \bar{y}$ $\hat{\alpha}_j = \bar{y}_j - \bar{y}$
Restricted	$Y_{ij} = \mu + \varepsilon_{ij}$	μ	$\hat{\mu} = \bar{y}$

Table 2 ANOVA output table; SS = sum of squares; MS = mean squares; df = degrees of freedom.

Source	SS	df	MS	F	P
Between	1890.1	2	945.1	3.06	0.084
Within	3710.0	12	309.8		

Comparing group means

A rejection of the omnibus test null hypothesis signals that at least one of the treatments had an effect, or, equivalently, that at least one pair of group means differ. Questions about which group means differ and by how much may be addressed by estimating and testing contrasts. A *contrast* is a linear combination of the group means such that the coefficients add up to 0. That is, ψ is a contrast if it may be written as follows,

$$\psi = c_1\mu_1 + c_2\mu_2 + \dots + c_a\mu_a = \sum_{j=1}^a c_j\mu_j \quad (9)$$

where $\sum_j c_j = 0$. Sample group means may be substituted to obtain an estimate,

$$\hat{\psi} = c_1\bar{y}_1 + c_2\bar{y}_2 + \dots + c_a\bar{y}_a = \sum_{j=1}^a c_j\bar{y}_j \quad (10)$$

and null hypotheses of the form $H_0 : \psi = 0$ may be tested by constructing an F statistic,

$$F_\psi = \frac{\hat{\psi}^2 / \sum_{j=1}^a \left(\frac{c_j}{n_j} \right)}{MS_W} \quad (11)$$

The statistic F_ψ will follow an F distribution under the null hypothesis with 1 numerator and $df_F = N - a$ denominator degrees of freedom. Contrasts for pairwise comparisons of means for treatment groups A and B and the control group, along with estimates, F statistics, and p -values, are given in [Table 3](#). The estimated group mean differences of 19.4, 26.2, and 7.2 may be directly observed as vertical distances between prediction lines in the full ANOVA model depicted in the right panel of [Fig. 2](#).

Univariate ANCOVA

The one-way ANCOVA model extends ANOVA by adding one or more numeric predictors (i.e., covariates). With a single covariate, X_{ij} , the ANCOVA model may be written as follows,

$$Y_{ij} = \mu + \alpha_j + \beta_1 X_{ij} + \varepsilon_{ij} \quad (12)$$

where β_1 is the slope that quantifies the strength of the linear relationship between the covariate, X , and the outcome, Y , and other parameters are defined as in [Eq. \(4\)](#).

Table 3 Table of contrasts, estimates, F statistics, and adjusted p -values for pairwise comparisons of group means.

Comparison	Ψ	$\hat{\psi}$	F	p
A vs. B	$1\mu_A - 1\mu_B + 0\mu_C$	19.4	3.03	0.23
A vs. control	$1\mu_A + 0\mu_B - 1\mu_C$	26.6	5.71	0.08
B vs. control	$0\mu_A + 1\mu_B - 1\mu_C$	7.2	0.42	0.80

Note. p -values are adjusted via Tukey's HSD for a family of three pairwise comparisons.

Table 4 Full and restricted models, parameters, and least-squares estimators for the one-way ANCOVA omnibus test for a model with one covariate.

ANCOVA	Model	Parameters	Least-squares estimator
Full	$Y_{ij} = \mu + \alpha_j + \beta_1 X_{ij} + \varepsilon_{ij}$	μ $\alpha_1, \alpha_2, \dots, \alpha_a$ β_1	$\hat{\mu} + \hat{\alpha}_j = \bar{y}_j - b_W \bar{x}_j$ $\hat{\beta}_1 = b_W$
Restricted	$Y_{ij} = \mu + \beta_1 X_{ij} + \varepsilon_{ij}$	μ β_1	$\hat{\mu} = \bar{y} - b_T \bar{x}$ $\hat{\beta}_1 = b_T$

The least squares estimator for coefficient β_1 under the restricted ANCOVA model, b_T , is referred to as the total or overall regression slope because it ignores group membership:

$$b_T = \frac{\sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x})(y_{ij} - \bar{y})}{\sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2} \quad (13)$$

The least squares estimator for β_1 under the full ANCOVA model, b_W , is referred to as the pooled within-group regression slope because it estimates a single slope by pooling the numerator and denominator components that would have contributed to the within-group slopes had they been estimated separately³:

$$b_W = \frac{\sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_j)}{\sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2} \quad (14)$$

As with the ANOVA model, the F statistic for testing the omnibus null hypothesis, $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_j$, under the ANCOVA model may be built up by making model comparisons under the general linear F framework. In particular,

$$F = \frac{\left(SS_B - \left[b_T^2 \sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 - b_W^2 \sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 \right] \right) / (a - 1)}{\left(SS_W - b_W^2 \sum_{j=1}^a \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 \right) / (N - a - 1)} \quad (15)$$

$$= \frac{SS_{B(\text{adj})} / (a - 1)}{SS_{W(\text{adj})} / (N - a - 1)} \quad (16)$$

where $SS_{B(\text{adj})}$ and $SS_{W(\text{adj})}$ are sums of squared residuals that have been *adjusted* for the linear regression of the outcome on the baseline covariate. The value of $SS_{B(\text{adj})}$ may be larger or smaller than the unadjusted (ANOVA) value of SS_B ; however, when participants are randomly assigned to groups, the value of $SS_{B(\text{adj})}$ will, on average, be identical to the unadjusted value of SS_B .

In contrast, $SS_{W(\text{adj})}$ will always be less than or equal to SS_W due to the subtraction of a non-negative term, which removes within-group variability due to the linear relationship between the outcome, Y , and the covariate, X , from the residuals. The stronger the linear relationship between X and Y within groups, the greater the reduction in $SS_{W(\text{adj})}$ relative to SS_W . If the covariate and outcome have no linear relationship in the sample, the value of b_W^2 will be zero, and the inclusion of the covariate will have no effect on SS_W . Note that the denominator df , $N - a - 1$, is decreased by one relative to ANOVA because one additional parameter, β_1 , is included in the ANCOVA model. Full and restricted models and least squares estimators for the one-way ANCOVA model are given in [Table 4](#).

Data demonstration

The left and right panels of [Fig. 3](#) depict the restricted and full models for the one-way ANCOVA omnibus null hypothesis along with dotted lines tracing residuals. Though it may be difficult to discern from the plot, the slope of the regression line in the left panel ($b_T = 3.51$) is slightly less steep than the three identical regression slopes in the right panel ($b_W = 3.71$).

³See [Wildt and Ahtola \(1978\)](#) for derivations of b_T and b_W .

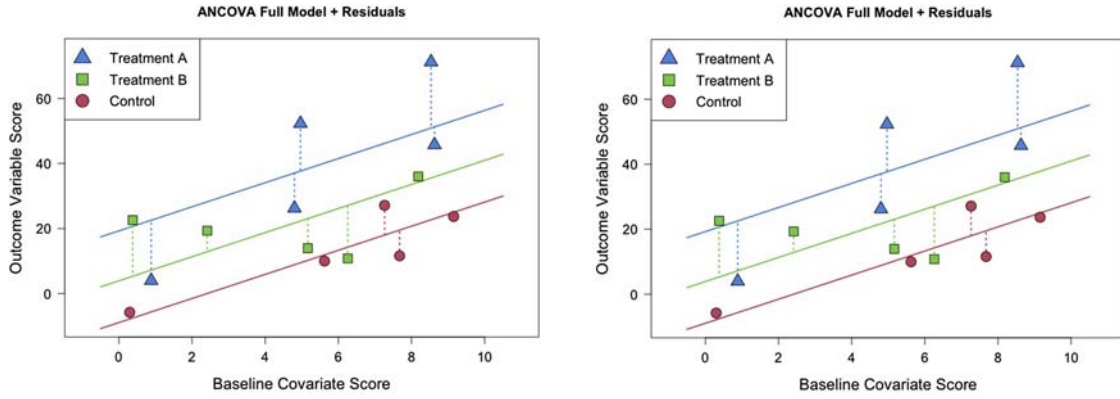


Fig. 3 ANCOVA restricted (left panel) and full (right panel) model fits and residuals.

Table 5 ANOVA output table for the ANCOVA model; SS = sum of squares; MS = mean squares; df = degrees of freedom.

Source	SS	df	MS	F	p
Between (adjusted)	1990.3	2	995.2	5.56	0.021
Covariate	1740.6	1	1740.6	9.72	0.010
Within (adjusted)	1969.4	11	179.0		

ANCOVA results are displayed in [Table 5](#). It is of interest to compare results with those of the ANOVA, shown in [Table 2](#). Here we see that $SS_{B(\text{adj})} = 1990.3$ as compared with $SS_B = 1890.1$. That this value changed by a small amount, only about 5% here, is expected because participants were randomly assigned to treatment groups and, therefore, were well-balanced on the baseline covariate. The change in within-subject variation is more drastic: $SS_{W(\text{adj})} = 1969.4$ as compared with $SS_W = 3710.0$. This large reduction of approximately 50% is also expected due to the strong positive linear relationship between the covariate and the outcome, and can be seen graphically by noting that the residuals in the right panel of [Fig. 3](#) (ANCOVA full model) are substantially smaller, on average, than those in the right panel of [Fig. 2](#) (ANOVA full model).

The ANCOVA-based test of the omnibus null hypothesis ($F(2, 11) = 5.56$; $p = 0.021$) produced a larger F statistic and smaller p -value than the ANOVA-based test ($F(2, 12) = 3.06$; $p = 0.084$), largely due to the reduction in $SS_{W(\text{adj})}$ relative to SS_W . That is, accounting for the linear relationship between the covariate and the outcome carved away a large portion of the variation that had previously contributed to the error variance.

Adjusted group means

Estimated group means under the ANCOVA full model are no longer equal to observed sample group means; instead, they are conditional on the value of the covariate (or covariates, if more than one are included).

$$\begin{aligned} E[Y_{ij}|x_{ij}] &= \hat{\mu} + \hat{\alpha}_j + b_W x_{ij} \\ &= \bar{y}_j + b_W(\bar{x}_j - \bar{x}_{ij}) \end{aligned} \quad (17)$$

By conditioning on a single value of the covariate, ANCOVA allows us to assess what the impact of each treatment *would have been* had all the groups been centered at the same value on the covariate. For each group, $j = 1, \dots, a$, the *adjusted group mean*, $\bar{y}'_j$, is defined as the predicted mean of the outcome score given that the covariate is set to its overall mean, $\bar{x}$.

$$\bar{y}'_j = E[Y_{ij}|x_{ij} = \bar{x}] = \bar{y}_j + b_W(\bar{x}_j - \bar{x}) \quad (18)$$

Pairwise differences in adjusted group means, shown in [Table 6](#), may be visualized as vertical distances between regression lines in the right panel of [Fig. 3](#).

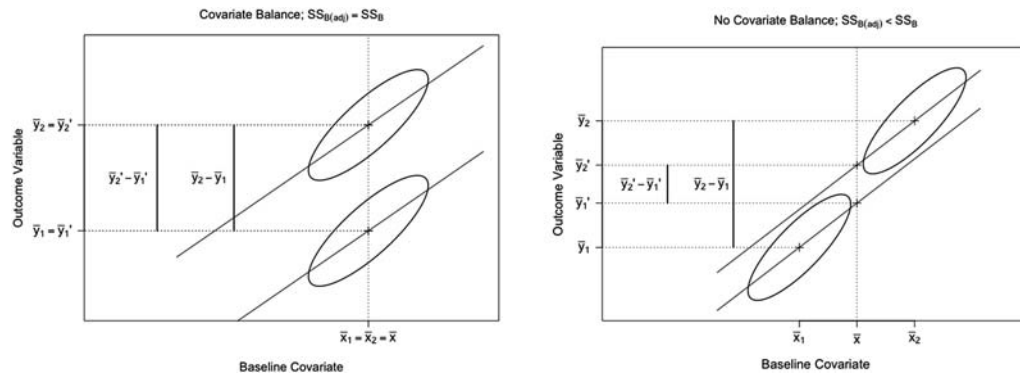
Covariate balance, overlap and ANCOVA adjustment

A covariate is said to be *balanced* with respect to the treatment groups if it is identically distributed across those groups. Covariate balance is expected in well-implemented randomized experiments because random assignment renders all covariates, X , statistically

Table 6 Table of contrasts for pairwise comparisons of adjusted group means.

Comparison	ψ	$\hat{\psi}$	F	p-value
A vs. B	$1\mu_A - 1\mu_B + 0\mu_C$	15.4	3.22	0.216
A vs. control	$1\mu_A + 0\mu_B - 1\mu_C$	28.2	11.08	0.017
B vs. control	$0\mu_A + 1\mu_B - 1\mu_C$	12.9	2.21	0.335

Note. p-values are adjusted via Tukey's HSD for a family of three pairwise comparisons.

**Fig. 4** A two-group experiment with baseline covariate balance (left panel) and imbalance (right panel).

independent of the treatment assignment, T ; that is, $X \perp T$. In non-randomized studies or when randomization becomes complicated by post-treatment factors such as differential attrition, covariates may be systematically imbalanced.⁴ When an important baseline covariate is imbalanced, adjustment via ANCOVA may cause unadjusted and adjusted group means and, hence, $SS_{B(adj)}$ and SS_B , to differ systematically. The direction of the difference depends on the severity of the imbalance and the magnitudes of b_T and b_W ; see the numerator of Eq. (15).

Consider the stylized diagrams in Fig. 4. The data in the left panel display good covariate balance at baseline, as evidenced by overlapping ranges of covariate values across groups and identical covariate group means (i.e., $\bar{x}_1 = \bar{x}_2 = \bar{x}$). Despite the strong positive linear relationship between the baseline covariate and the outcome, covariate adjustment will make no systematic difference to the estimates of treatment effect in cases of good balance; note that $\bar{y}'_2 - \bar{y}'_1 = \bar{y}_2 - \bar{y}_1$ in the left panel of Fig. 4.

Data in the right panel display covariate imbalance at baseline. If X and T were truly independent, we would not expect such drastic lack of overlap in covariate values across groups and differing covariate group means (i.e., $\bar{x}_1 \neq \bar{x}_2$). In this case, the covariate X appears to be a confounding variable because it is associated both with the group assignment and with the outcome variable. If the assumed linear functional forms, as specified in Eq. (12), are correct, estimates based on the adjusted means will remove confounding bias due to the covariate from effect estimates; note that $\bar{y}'_2 - \bar{y}'_1 < \bar{y}_2 - \bar{y}_1$. The trouble is that when covariate imbalance leads to lack of covariate overlap, assumptions about correct model specification must be made based on extrapolation (i.e., in the absence of data), which is not advised.

Although lack of overlap is easy to spot when controlling for only one covariate, it is more difficult to visualize as more covariates are included. One solution is to estimate the propensity score, which serves as a unidimensional balance proxy for the full set of observed covariates (Rosenbaum and Rubin, 1983). There are a number of resources that provide guidance on best practices for the use of propensity scores for checking overlap (see, e.g., Keller and Tipton, 2016; Schafer and Kang, 2008; and Steiner and Cook, 2013).

Assumptions and limitations

The assumptions required for valid statistical inference in ANCOVA may be summarized succinctly by the statement, $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$.

1. *Errors are statistically independent.* That is, after accounting for the variation explained by the covariates and by group membership, residual variations are assumed to be independent. This assumption could be violated if, for example, participants were repeatedly measured. In that case, each participant's set of measurements would likely be more highly correlated with one another than with those of other participants, even after controlling for fixed effects in the model. It is typically not possible to

⁴While it is possible that a well-implemented randomized experiment may result in poor baseline covariate balance due to an unlucky randomization, this becomes less and less likely as sample sizes increase. For small to moderate samples, where chance imbalances are more likely, rerandomization (Morgan and Rubin, 2014; Zhou et al., 2018) may be used to ensure a minimum level of balance is achieved.

test the independence assumption with data; instead, logical arguments based on how the data were collected must be made to defend it.

2. *Errors are normally distributed with mean 0 and constant variance.* For any combination of covariate values and/or group indicators, outcome data are assumed to have been generated from identical normal distributions, each with mean 0 and constant variance, σ^2 .
 - (a) *Linearity.* The assumption of zero conditional mean implies that the functional form given in Eq. (12) is correctly specified. This can go wrong if the true relationship between the covariate and the outcome is nonlinear or if there is effect modification (i.e., the magnitudes and/or directions of treatment effects vary with the values of a covariate).
 - (b) *Constant variance.* When group sizes are equal, or nearly so, ANCOVA-based tests are generally robust to violations of the constant variance assumption. For unbalanced group size pairings, however, heterogeneous variances may cause problematic departures from expected type I error rates. The Welch-Satterthwaite correction is recommended for such cases because it maintains nominal type I error rates in the face of group variance heterogeneity and still is quite powerful relative to alternatives (Tomarken and Serlin, 1986).
 - (c) *Normal distribution.* The ANCOVA omnibus null hypothesis test is somewhat robust to non-normally distributed residuals, in particular as sample sizes grow larger. Violation of the distributional assumption may invalidate the nominal type I error rate for smaller samples, especially when coupled with unbalanced group sizes.

When verifying the tenability of the latter three assumptions with data, we recommend using a multifaceted approach that includes graphical exploration, estimation of relevant effect sizes, and the use of hypothesis tests. Other issues such as multicollinearity and influential observations can have an impact on the stability of results. For more detailed discussions on checking assumptions see, e.g., Keppel and Wickens (2004), Huitema (2011), Maxwell et al. (2018), and Fox (2008).

Multivariate ANOVA

Multivariate analysis of variance (MANOVA) is a generalization of ANOVA to multiple dependent variables (DVs). Paralleling Eq. (4), the MANOVA model is:

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}, \quad i = 1, \dots, n_j, j = 1, \dots, a \quad (19)$$

where each term is now a p -dimensional vector, and the errors ε_{ij} are assumed to be independent multivariate normal $N_p(0, \Sigma)$ variables. As in the univariate case, the treatment effects may be parameterized to satisfy $\sum_{j=1}^a \alpha_j = 0$. Note that each component of the observed vector Y_{ij} satisfies the univariate model 3 and that the error components are correlated, but the overall covariance matrix Σ is the same for all groups.

The overall omnibus null hypothesis to be tested is that the vector of treatment effects for each outcome variable is identical across all groups:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0$$

Furthermore, the decomposition in Eq. (6) still holds in its multivariate version:

$$(y_{ij} - \bar{y}) = (\bar{y}_j - \bar{y}) + (y_{ij} - \bar{y}_j) \quad (20)$$

Since each term in Eq. (20) is a vector, it can't simply be squared; instead outer products are used, and the resulting equation consists of *matrices* of sums of squares and cross products (SSCP):

$$\text{SSCP}_{T; \text{ total (corrected)}} = \text{SSCP}_{B; \text{ between}} + \text{SSCP}_{W; \text{ within}} \\ \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y})(y_{ij} - \bar{y})' = \sum_{j=1}^a n_j (\bar{y}_j - \bar{y})(\bar{y}_j - \bar{y})' + \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)(y_{ij} - \bar{y}_j)' \quad (21)$$

The following matrices are multivariate analogs to SS_B and SS_W :

$$B = \sum_{j=1}^a n_j (\bar{y}_j - \bar{y})(\bar{y}_j - \bar{y})' \quad (22)$$

$$W = \sum_{j=1}^a \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)(y_{ij} - \bar{y}_j)' \quad (23)$$

In the univariate case the sum of squares SS_B and SS_W are independently distributed with χ^2 distributions, and in the multivariate case B and W have independent Wishart distributions. To test the overall omnibus null hypothesis, a measure of how *large* the matrices B and W are is needed. There is *no unique* way to extend the F -test from Eq. (8) to the multivariate case. Different researchers have proposed different multivariate analogies to the F -statistic. One of the most commonly used is Wilks' (1932) Lambda statistic:

$$\Lambda = \frac{\det(W)}{\det(B + W)} \quad (24)$$

As Wilks' Λ is a ratio of the determinants of two Wishart distributed matrices, exact results for its distribution are not available in general, but there exist relationships to the F distribution for some special cases of p and a , or an asymptotic χ^2 approximation due to Bartlett (1938). Essentially, Wilks' Λ arises from the likelihood ratio principle and so does its asymptotic behavior. Note that more separation among the groups will drive the W matrix closer to 0; therefore, H_0 is rejected for *small* values of Λ . Wilks' Λ can also be written as $\Lambda = \frac{1}{\det(I + W^{-1}B)}$, and consequently, using the eigenvalues $\lambda_1, \dots, \lambda_r$ of the matrix $W^{-1}B$, where $r = \text{rank}(B)$ as:

$$\Lambda = \prod_{i=1}^r \frac{1}{1 + \lambda_i} \quad (25)$$

As groups become more separated, W becomes closer to being a singular matrix, so the eigenvalues grow larger. Other approaches for testing H_0 due to Pillai, Lawley-Hotelling and Roy are all based on the same eigenvalues but use them in a different way. For example, Pillai's trace statistic is defined as

$$V = \text{tr}((W + B)^{-1}B) = \sum_{i=1}^r \frac{\lambda_i}{1 + \lambda_i} \quad (26)$$

and H_0 is rejected for *large* values of V . For very large samples, all four criteria are indistinguishable. For small to moderate sample sizes, some simulation-based evidence suggests that Pillai's trace is more robust to non-normality (Johnson and Wichern, 2007, p. 336).

If the multivariate omnibus null hypothesis is rejected, it is typically of interest to know which variables contributed to the rejection of the null hypothesis and which did not. Rencher and Christensen (2012, p. 195) argue that follow-up univariate ANOVA F -tests do not change the overall nominal significance level, provided that they are performed *only after rejection* of the MANOVA null hypothesis. However, this does not mean that these univariate F -tests might not suffer from inflated Type II error rates (i.e., lack of power).

This is illustrated in Fig. 5, where the true population distribution is a bivariate normal with mean $(0, 0)$, variances equal to 2 and 1, and a covariance equal to 1, which results in a high correlation of $\sqrt{2}/2 \approx 0.71$. This can be thought of as the distribution of the difference between two groups, which under the null hypothesis should have mean $(0, 0)$. The outermost red ellipse corresponds to the 95% contour of the distribution, which demarcates the acceptance and rejection regions for the multivariate null hypothesis. The shaded rectangular area is the Cartesian product of the acceptance regions of the two univariate tests.

The dot represents a potential sample with mean $(2, -1)$, which is not an unusual data point if looking at each variable separately, as shown by the fact that the point is inside the rectangle corresponding to the Bonferroni adjusted simultaneous 95% CIs for the univariate tests. Such data would result in rejection of the *multivariate* hypothesis because under the null hypothesis it is very unlikely to observe data with a large positive Y_1 mean together with a negative Y_2 mean. However, neither of the follow-up univariate hypotheses would be rejected.

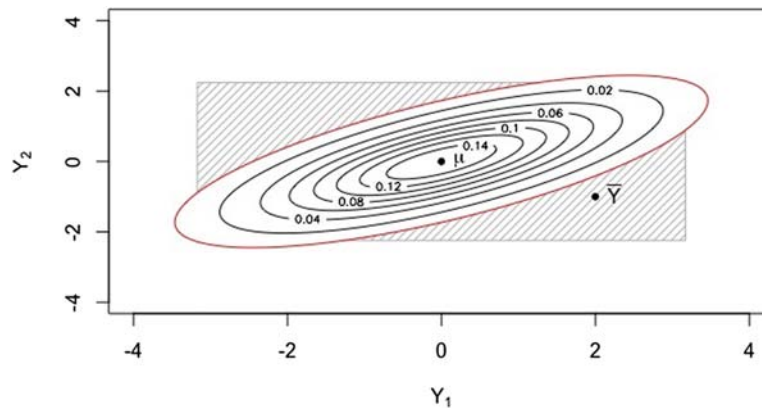


Fig. 5 A bivariate normal distribution with correlation 0.71. The area outside the red ellipse represents the rejection region for the multivariate omnibus null hypothesis test. The area outside the shaded rectangle represents the Bonferroni-corrected rejection region for (either of the two) univariate omnibus null hypothesis tests. For the sample mean vector $\bar{y}$, the multivariate test rejects while the univariate tests do not. Neither $y_1 = 2$ nor $y_2 = -1$ is an unlikely observation when *taken alone*, however, given the correlation between Y_1 and Y_2 , they are quite unlikely to have been observed *together*.

Multivariate ANCOVA

Multivariate analysis of covariance (MANCOVA) is a combination of the MANOVA model with a multivariate regression model with one or more predictors. The inclusion of these covariates leads to *noise reduction* in the sense that the variance associated with the covariates is removed from the error variance, thus providing a more powerful test for the difference between the groups. The model with a single covariate may be written as follows,

$$Y_{ij} = \mu + \alpha_j + \beta_1 x_i + \varepsilon_{ij}, \quad i = 1, \dots, n_j, j = 1, \dots, a \quad (27)$$

where each term is a $p \times 1$ vector. The general case is easier to write in matrix notation,

$$\underset{n \times p}{Y} = \underset{n \times a \times p}{Z} \underset{n \times a}{\alpha} + \underset{n \times b \times p}{X} \underset{n \times b}{\beta} + \underset{n \times p}{E} \quad (28)$$

where Z is the MANOVA design matrix, and X is the multivariate regression matrix containing b covariates. The estimates of the coefficients are produced via standard multivariate regression formulas (see, e.g., [Timm, 2002](#), p. 226). For example, using the generalized inverse matrix $(Z'Z)^-$ for the case when the MANOVA is overparametrized,⁵ we have:

$$\hat{\alpha} = (Z'Z)^- Z' (Y - X\hat{\beta}) \quad (29)$$

Observe that this estimate of the group effects is adjusted by the residual from a regression on the covariates.

To test hypotheses about the treatment effects, α , the standard approach is to obtain the “within” matrix (now called the *residual matrix*) W_{full} from the full model in [Eq. \(28\)](#) and the residual matrix $W_{reduced}$ from the model under the null hypothesis, and then proceed the same way as in the MANOVA hypothesis, using the eigenvalues of $W_{full}^{-1}(W_{full} - W_{reduced})$. To summarize, the MANCOVA test enjoys the benefits of both MANOVA (i.e., increased power for rejecting the omnibus test with correlated outcomes) and ANCOVA (i.e., reduction of within-groups error terms via regression).

Assumptions and limitations

Significance tests for MANOVA, MANCOVA, and many other multivariate statistical analyses are based on the assumption that the data are multivariate normal. More specifically, the error matrix E in [Eq. \(28\)](#) is assumed to contain on each row independent vectors which are distributed $N_p(0, \Sigma)$.

1. *Row vectors are independent.* This assumption is analogous to the univariate case with exception that the observations associated with the same participant are assumed to be correlated as specified in Σ .
2. *Error vectors are multivariate normal with zero mean vector and constant covariance matrix.*
 - (a) *Linearity.* In line with the univariate case, the relationships between covariates and outcome variables are assumed to be correctly specified.
 - (b) *Constant covariance matrix.* The assumption that each error vector is distributed $N_p(0, \Sigma)$ implies identical covariance matrices. This assumption is the multivariate analog to the univariate assumption of constant variance. The assumption may be tested with Box’s M test, though the test is sensitive to violations of other assumptions such as multivariate normality.
 - (c) *Multivariate normal distribution.* Multivariate normality implies that each outcome variable must be normally distributed, which can be tested with a number of tests such as Shapiro-Wilks, Anderson-Darling or Kolmogorov-Smirnov, as well as graphs such as a Q-Q plot. However, multivariate normality is a stronger assumption than individual univariate normality, and there are multivariate tests available in many software packages such as Mardia’s test and the Cox-Small test. [Mardia \(1971\)](#) argues that reasonably large sample sizes in each group ensure robustness and mitigates the need for extensive testing, though [Rencher and Christensen \(2012, p. 210\)](#) suggest checking for gross violations.

MANOVA and MANCOVA procedures are sensitive to outliers. Thus, it is recommended that tests for outliers are performed with any MANOVA and MANCOVA analysis. For more discussion of assumptions and their implications see, e.g., [Rencher and Christensen \(2012\)](#), [Johnson and Wichern \(2007\)](#) and [Tabachnick and Fidell \(2013\)](#).

Data demonstration

[Woodworth et al. \(2017\)](#) replicated a randomized efficacy study of three web-based positive psychology interventions and a control condition, and [Woodworth et al. \(2018\)](#) made the data publicly available. Participants assigned to the *gratitude* condition were asked to send a letter of gratitude to someone deserving in their life. Participants assigned to the *three good things* condition were asked to write down three good things that happened each day for a week. Participants assigned to the *signature strengths* condition completed a survey to assess character strengths and were asked to use one of their top five strengths each day for a week. Participants in the comparison group were asked to write about early memories each day for a week.

⁵Following [Timm \(2002\)](#), we use the superscript $-$ to denote the generalized inverse.

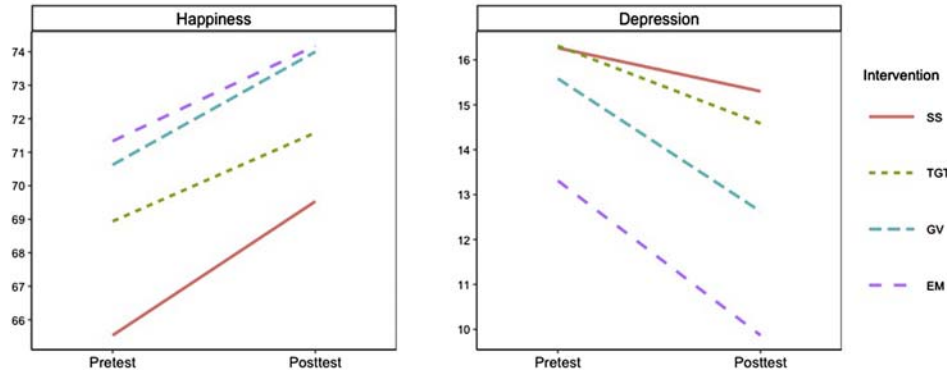


Fig. 6 Pretest and posttest scores on two dependent variables for the data demonstration; note the lack of balance across intervention group means at baseline.

Of interest is the effect of the intervention type ($a = 4$ groups) on the depression and happiness index scores ($p = 2$ DVs, correlated at -0.76). For this analysis we focus on the 147 participants who completed baseline and posttest measures; see Fig. 6. Neither Wilks nor Pillai MANOVA omnibus tests are significant ($p = 0.28$). A follow-up analysis with univariate F -tests on each of the DVs separately agrees with this result. Running a MANCOVA, controlling for baseline scores as covariates, also results in a nonsignificant omnibus test result ($p = 0.45$) via both Wilks and Pillai. The error matrices for the MANOVA and MANCOVA are:

$$W_{\text{MANOVA}} = \begin{pmatrix} 17736.79 & -15240.07 \\ -15240.07 & 22637.71 \end{pmatrix}, W_{\text{MANCOVA}} = \begin{pmatrix} 8458.66 & -5233.99 \\ -5233.99 & 7714.14 \end{pmatrix}$$

As expected, adjusting for baseline scores reduced important components of the error variance matrix. It is instructive to note that $\det(W_{\text{MANCOVA}}) = 3.8 \times 10^7$ is 4.5 times smaller than $\det(W_{\text{MANOVA}}) = 1.7 \times 10^8$. A similar comparison of the “Between” matrices (for formulas see Timm, 2002, p. 227), yields

$$B_{\text{MANOVA}} = \begin{pmatrix} 724.04 & -525.16 \\ -525.16 & 481.89 \end{pmatrix}, B_{\text{MANCOVA}} = \begin{pmatrix} 216.39 & -23.68 \\ -23.68 & 18.7 \end{pmatrix}$$

In general, for randomized experiments, we do not expect to see large differences in the MANOVA and MANCOVA covariate-adjusted between matrices. In this example, however, the determinant of the between matrix decreased by a factor of almost 21. The explanation for this large reduction in the magnitudes of the components of the between matrix is the lack of baseline balance in the analytic sample, perhaps due to post-randomization attrition. MANOVA fails to account for the baseline differences and attributes all outcome variation to the treatment. Whereas, with MANCOVA, the baseline variation in outcomes across treatments is accounted for and, thus, $\det(B)$ is substantially smaller. Note that we use these data to demonstrate the mechanics of the multivariate procedures. We have skipped over important details such as the verification of the tenability of assumptions and the reporting of and interpretation of results in context that should not be neglected in practice.

Discussion

Random assignment leads to covariate balance, which renders both ANOVA and ANCOVA-based estimators of effects unbiased (see the left panel of Fig. 4). The purpose of making ANCOVA adjustments with randomized experimental data is to reduce error variance, which leads to more precise estimates and more powerful tests. With observational data, ANCOVA may still be useful in reducing error variance, but it also is used for another purpose: to reduce bias due to confounding covariates.

The ANCOVA and MANCOVA models, as described in Eqs. (11) and (26), require (a) that relationships between covariates and outcomes are linear and (b) that treatment effects are constant across all covariate values. If either of these assumptions is violated, estimates based on adjusted means are no longer valid. Furthermore, because extrapolation is required when covariate imbalance leads to lack of overlap, the validity of the estimates may be quite sensitive to functional form misspecifications with observational data. That said, some empirical evidence suggests that in some cases the basic ANCOVA model can perform quite well with respect to bias reduction with observational data (Shadish et al., 2008); however, simulation evidence suggests that when functional form assumptions are violated, performance may be substantially improved by allowing for more flexible modeling, for example, by including covariate by treatment interaction terms to handle non-constant treatment effects (Schafer and Kang, 2008).

There is a rich and expanding literature on alternatives to ANCOVA for effect estimation in observational study settings in the causal inference literature; some of these approaches are reviewed in Schafer and Kang (2008). In recent years, a number of machine learning alternatives to ANCOVA for causal effect estimation with observational data have been proposed. These newer methods lack the simplicity and interpretability of ANCOVA but benefit from algorithmic handling of nonlinearities and interactions in their

naive implementations. Bayesian additive regression trees (Hill et al., 2011) and targeted maximum likelihood estimation (TMLE; van der Laan and Gruber, 2010) are two examples that have performed well in data competitions (Dorie et al., 2019). Furthermore, a number of new methods have been proposed for explicitly detecting and estimating causal effects in the presence of treatment effect heterogeneity. A special issue in the journal *Observational Studies* (Volume 5, Issue 2, 2019) is devoted to the topic.

Our focus here has been on one-way (i.e., single-factor) models for ANCOVA and MANCOVA. Of course, it may be of interest to adjust for covariates in experimental contexts involving two or more factors and their interactions. It poses no problem to extend to multifactor studies, and a number of texts provide relevant test statistics and formulas for adjusted cell means in the univariate case (e.g., Neter et al., 1990; Kirk, 1982). In the multivariate case, one would need to alter the MANOVA design matrix, i.e., Z in Eq. (27), to include the appropriate deviation-coded variables; see Timm (2002, p. 246) for details.

Finally, after a rejection of the MANCOVA omnibus null hypothesis, as in the multivariate data demonstration given above, univariate ANCOVA F tests are not the only options for following up to better understand what contributed to the rejection. Instead, for example, one might choose to search for the linear combination of outcome variables that lead to the best separation among treatment groups. This approach, called *discriminant analysis*, is beyond the scope of this chapter but is covered in the multivariate texts referenced herein.

References

- Bartlett, M.S., 1938. Further aspects of the theory of multiple regression. *Math. Proc. Camb. Phil. Soc.* 34, 33–40.
- Dorie, V., Hill, J., Shalit, U., Scott, M., Cervone, D., 2019. Automated versus do-it-yourself methods for causal inference: lessons learned from a data analysis competition. *Stat. Sci.* 34, 43–68.
- Fisher, R.A., 1934. *Statistical Methods for Research Workers*, fifth ed. Oliver and Boyd, London, England.
- Fox, J., 2008. *Applied Regression Analysis and Generalized Linear Models*, second ed. Sage, Thousand Oaks, CA.
- Hill, J., Weiss, C., Zhai, F., 2011. Challenges with propensity score strategies in a high-dimensional setting and a potential alternative. *Multivariate Behav. Res.* 46, 477–513.
- Huitema, B., 2011. *Analysis of Covariance and Alternatives*, second ed. Wiley, New York.
- Johnson, R.A., Wichern, D.W., 2007. *Applied Multivariate Statistical Analysis*, sixth ed. Prentice Hall, Upper Saddle River, NJ.
- Keller, B., Tipton, E., 2016. Propensity score analysis in R: a software review. *J. Educ. Behav. Stat.* 41, 326–348.
- Keppel, G., Wickens, T.D., 2004. *Design and Analysis: A Researcher's Handbook*, fourth ed. Pearson Education, Inc, Upper Saddle River, New Jersey.
- Kirk, R.E., 1982. *Experimental Design: Procedures for the Behavioral Sciences*, second ed. Brooks/Cole Publishing Company, Pacific Grove, CA.
- Mardia, K.V., 1971. The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. *Biometrika* 58, 105–121.
- Maxwell, S.E., Delaney, H.D., Kelley, K., 2018. *Designing Experiments and Analyzing Data: A Model Comparison Perspective*, third ed. Routledge, New York.
- Morgan, K.L., Rubin, D.B., 2014. Rerandomization to improve covariate balance in experiments. *Ann. Stat.* 40, 1263–1282.
- Neter, J., Wasserman, W., Kutner, M.H., 1990. *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*, third ed. Irwin, Burr Ridge, IL.
- Rencher, A.C., Christensen, W.F., 2012. *Methods of Multivariate Analysis*, third ed. John Wiley & Sons, Inc, Hoboken, NJ.
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55.
- Schafer, J., Kang, J., 2008. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol. Methods* 13, 279–313.
- Shadish, W.R., Clark, M.H., Steiner, P., 2008. Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *J. Am. Stat. Assoc.* 103, 1334–1343.
- Steiner, P.M., Cook, T.D., 2013. Matching and propensity scores. In: Little, T. (Ed.), *Oxford Handbook of Quantitative Methods*. Oxford University Press, Oxford.
- Tabachnick, B.G., Fidell, L.S., 2013. *Using Multivariate Statistics*, sixth ed. Pearson, Boston.
- Timm, N.H., 2002. *Applied Multivariate Analysis*. Springer, New York.
- Tomarken, A.J., Serlin, R.C., 1986. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychol. Bull.* 99, 90–99.
- van der Laan, M.J., Gruber, S., 2010. Collaborative double robust targeted maximum likelihood estimation. *Int. J. Biostat.* 6, Article 17.
- Wildt, A.R., Ahtola, O.T., 1978. *Analysis of Covariance*. Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 12. Sage Publications, Beverly Hills and London.
- Wilks, S.S., 1932. Certain generalizations in the analysis of variance. *Biometrika* 24, 471–491.
- Woodworth, R.J., O'Brien-Malone, A., Diamond, M.R., Schüz, 2017. Web-based positive psychology interventions: a reexamination of effectiveness. *J. Clin. Psychol.* 73, 218–232.
- Woodworth, R.J., O'Brien-Malone, A., Diamond, M.R., Schüz, 2018. Data from, "web-based positive psychology interventions: a reexamination of effectiveness". *J. Open Psychol. Data* 6, 1.
- Zhou, Q., Ernst, P.A., Morgan, K.L., Rubin, D.B., Zhang, A., 2018. Sequential rerandomization. *Biometrika* 105, 745–752.

Linear regression analysis

Sijia Huang, Indiana University Bloomington, Bloomington, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	548
Formulation of linear regression model	548
Simple linear regression model	549
Multiple linear regression model	550
Estimation of linear regression models	551
Ordinary least squares (OLS)	552
Geometric interpretation of OLS	553
Standard errors, confidence intervals and hypotheses testing of β	554
Maximum likelihood estimation (MLE)	554
Assumptions of linear regression model	555
Linearity	555
No multicollinearity among independent variables	555
Distribution of disturbance term	555
Homoscedasticity	555
Extensions of linear regression model	556
Generalized linear model	556
Hierarchical linear model	556
Multivariate linear model	557
Conclusion	557
References	557

Introduction

When two or more than two attributes are measured, it is often of interest that if there exist any associations or dependence among these attributes. Linear regression is one of the most broadly applied statistical techniques for the modeling of relationships among variables. The term *regression* dates back to [Galton's \(1886\)](#) study on the relationship between the heights of parents and their adult children. [Galton \(1888\)](#) also invented the *index of correlation* and used it to measure the closeness of two variables. Pearson, who was an English mathematician and a notable student of Galton, later expanded Galton's ideas and presented a mathematical treatment of regression ([Pearson, 1896](#)). [Pearson \(1895\)](#) also developed the formula of the still commonly-used Pearson product-moment correlation coefficient (PPMCC). [Slutsky \(1913\)](#) contributed to the development of regression analysis through introducing a goodness of fit test that adopts the assumption of normality. [Fisher \(1922\)](#) synthesized the work of Pearson and the theory of errors of [Gauss's \(1809\)](#), and introduced what is considered the modern regression model. For more information regarding the history of regression analysis, we refer interested readers to sources such as [Stanton \(2001\)](#) and [Aldrich \(2005\)](#).

Since its invention, regression analysis gradually became an active research area. The standard linear regression model, which is the focus of this entry, has been extended to generalized linear models (GLMs; [Nelder and Wedderburn, 1972](#)), hierarchical linear models (HLMs; [Raudenbush and Bryk, 2002](#)), and multivariate regression model. The linear regression model and its extensions have also been applied extensively in a broad spectrum of scientific fields, such as biology and education, for the purposes of making predictions and inferring causal relationships. In addition, the model is considered the foundation of the statistical learning theory ([James et al., 2013](#)), which deals mainly the issue of prediction.

Formulation of linear regression model

Unlike a correlation coefficient such as the PPMCC, which is one single value that quantifies the linear relationship between two variables, a linear regression model describes the relationship between one dependent variable and one or more than one independent variable. The dependent variable, also known as the response or outcome variable, is assumed to be a linear combination of the regression coefficients and independent variable(s), which is also known as the explanatory variable(s) or predictor(s). When only one independent variable is considered, the linear regression model is called *simple linear regression model*. When more than one independent variable is included in the model, it is named *multiple regression model*.

Simple linear regression model

In a simple linear regression model, both the independent variable and the dependent variable are scalars. **Table 1** shows ten simulated observations, each of which consists of measurements of two variables, x and y . Assuming that x is the independent variable and y is the dependent variable, we generate a scatter plot (**Fig. 1**) whose horizontal axis indicates values of x of the observations and the vertical axis shows values of y . The dots in the figure represent the observations (x_i, y_i) , for $i = 1, \dots, 10$. A simple linear regression model that describes the relationship between x and y takes the following form:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \text{for } i = 1, \dots, n \quad (1)$$

where n is the total number of observations, and $n = 10$ in the simulated example. The assumed linear relationship between x and y is captured by two unknown but fixed regression coefficients: an intercept β_0 and a slope β_1 . As shown in **Fig. 1**, β_0 indicates the value of y when $x = 0$, and β_1 represents the averaged change in y that is associated with a unit increase in x (e.g., the change in y as x changes from 4 to 5, and from 10 to 11). Once estimated, the regression coefficients can be used to make predictions of y given values of x . ε_i is a disturbance term which represents the part of y that cannot be explained by the independent variable x . In practice, ε_i is often assumed to follow a normal distribution with mean zero and variance σ^2 , $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. However, this normality assumption is not required for the estimation of the regression coefficients or the variance of the disturbance term.

Note that in this example, the independent variable x is continuous, but it can be categorical. When the independent variable is categorical, the process of *dummy coding* is usually performed to convert the categorical variable into a series of dummy variables, which can only take the values of 0 and 1. For example, if the variable *region*, which has five categories (*Northeast*, *Southwest*, *West*, *Southeast*, and *Midwest*), is included in a regression model as the independent variable, we can let the category *Northeast* be the

Table 1 Simulated data.

i	y	x
1	27	9
2	5	0
3	23	7
4	21	11
5	20	6
6	26	5
7	34	9
8	21	3
9	16	3
10	37	12

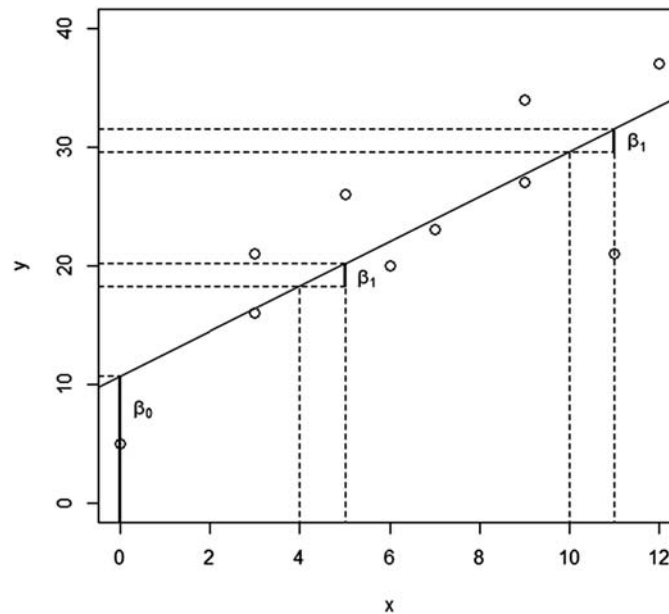


Fig. 1 Simple linear regression Values of the independent variable x are plotted on the horizontal axis, while values of the dependent variable y are plotted on the vertical axis. Each dot in the figure represents an observation. The fitted line is defined by an intercept β_0 and a slope β_1 .

Table 2 An example of dummy coding.

Region	Dummy variable 1	Dummy variable 2	Dummy variable 3	Dummy variable 4
Northeast	0	0	0	0
Southwest	1	0	0	0
West	0	1	0	0
Southeast	0	0	1	0
Northeast	0	0	0	1

reference group and create four dummy variables. Each of the four dummy variables indicates if the realization of the variable *region* is *Southwest*, *West*, *Southeast*, or *Midwest* or not, respectively. Table 2 shows the dummy coding process.

The created dummy variables instead of the original variable will be included in the linear regression model. In this case, the intercept coefficient β_0 indicates the average value of y in the reference group (*Northeast*), and the slope coefficients associated with the dummy variables represent the differences in y between the other four groups and the reference group.

In the linear regression analysis, a quantity that is often of interest is the so-called *coefficient of determination*, which is usually denoted by R^2 . R^2 represents the proportion of variation in the dependent variable that can be explained by the independent variable(s), and it ranges from 0 to 1. A R^2 that is close to 1 indicates that the regression model fits the data (i.e., the observations) very well, and it is reliable in terms of making predictions. On the other hand, a R^2 that is close to 0 means that the relationship between the independent and dependent variables is very weak.

In the case of the simple linear regression model, R^2 simply equals the square of the correlation coefficient between the independent and dependent variables (x and y in the simulated example), and can be derived from the slope coefficient β_1 and the data,

$$R^2 = \frac{\beta_1^2 \times \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2)$$

where $\bar{x}$ and $\bar{y}$ represent the means of x_i and y_i ($i = 1, \dots, n$), $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ and $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$, respectively.

Multiple linear regression model

When it comes to the case in which more than one independent variable is assumed to have an impact on the dependent variable, the linear regression process is termed *multiple linear regression*. The dependent variable is expressed as a linear combination of all independent variables. A multiple regression model with p independent variables takes the form:

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p} + \varepsilon_i. \quad (3)$$

In Eq. (3), y_i is the realization of the scalar dependent variable y in the i th observation, and $x_{i,k}$ ($k = 1, \dots, p$) represents the value of the k th independent variables in the i th observation. The multiple regression model contains $p + 1$ parameters: one intercept β_0 and p slope coefficients, each of which corresponds to an independent variable. ε_i is the disturbance term and represents the unexplained part of y_i . ε_i is often assumed to be normally distributed, but this is not necessary for the estimation of the model parameters. A compact form of Eq. (3) is:

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{i,k} + \varepsilon_i. \quad (4)$$

Table 3 shows 10 simulated observations. Each of the ten observations includes values of a dependent variable (y) and two independent variables (x_1 and x_2). If a linear relationship can be assumed between the dependent variable and the two independent variables, a multiple regression model can be written down,

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i. \quad (5)$$

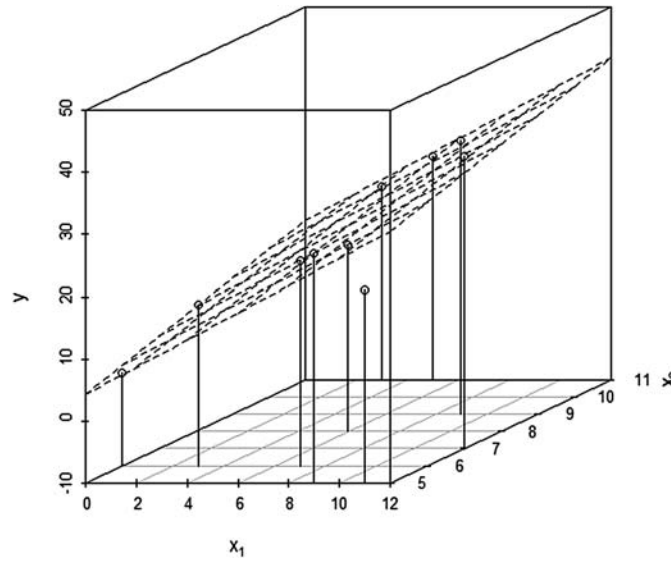
The intercept β_0 indicates the value of y when both independent variables are zeros, $x_1 = x_2 = 0$. β_1 is the slope coefficient associated with the first independent variable x_1 , and shows the change in y as x_1 increases by one unit, holding x_2 constant. β_2 is the slope coefficient that corresponds to x_2 , and represents the change in y associated with every one unit increase in x_2 , controlling for x_1 . ε_i is the disturbance term and has a variance of σ^2 .

The linear relationship among x_1 , x_2 , and y can be depicted by a hyperplane in a three-dimension space. The three axes in Fig. 2 indicate values of x_1 , x_2 and y of the simulated observations, respectively, so that each point in the box represents a simulated observation. The hyperplane indicates how y changes as the two independent variables change. To facilitate the understanding of regression coefficients, two more plots are generated. Fig. 3A shows the amount that y would increase if x_1 increases by one unit while fixing x_2 to be 5 and 10, respectively. The two lines in this figure can be viewed as two “slices” of the box shown in Fig. 2. The two parallel lines in Fig. 3B indicate how y changes as x_2 increases while fixing x_1 to be 4 and 8.

The multiple regression model can be expressed in the following matrix form,

Table 3 Simulated data (two predictors).

i	Y	x_1	x_2
1	27	9	5
2	5	0	6
3	23	7	6
4	21	11	5
5	20	6	8
6	26	5	11
7	34	9	9
8	21	3	11
9	16	3	6
10	37	12	7


Fig. 2 Multiple linear regression The three axes show values of x_1 , x_2 and y of the observations. The linear relationship among them is described by the hyperplane, which is define by β_0 , β_1 and β_2 .

$$y = X\beta + \epsilon, \quad (6)$$

where

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{1,1} & \dots & x_{1,p} \\ \vdots & \ddots & & \vdots \\ 1 & x_{n,1} & \dots & x_{n,p} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \text{and} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}. \quad (7)$$

In Eq. (6), y is a vector that consists of realizations of the dependent variable in all n observations. X is the so-called *design matrix* and is of size $n \times (p + 1)$. All elements of the first column of X are ones, and correspond to the intercept coefficient β_0 . β is the regression coefficient vector, and consists of one intercept and p slope coefficients, each of which is associated with one independent variable. ϵ is a vector of size n and consists of the disturbance terms. ϵ is often assumed to follow a multivariate normal distribution with mean zeros and a diagonal covariance matrix, whose diagonal elements are all σ^2 and off-diagonal elements are all zeros, $\epsilon \sim \mathcal{N}_n(0, I_n \sigma^2)$. Again, this normality assumption is not required for the estimation of model parameters.

Estimation of linear regression models

The estimation of linear regression models refers to obtaining the model parameters (i.e., the intercept and slope coefficients) that define the “best-fitting” line or hyperplane of the data. The discussion in this section focuses on the ordinary least squares (OLS)

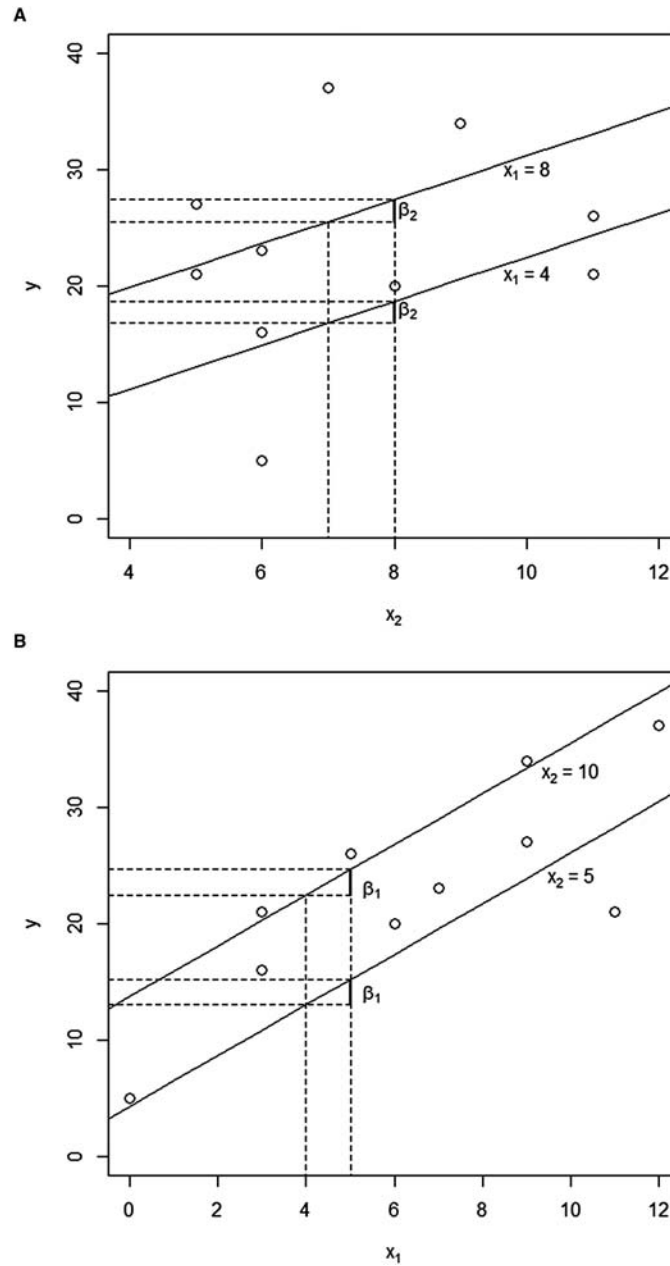


Fig. 3 Illustration of regression coefficients in multiple regression models (A) Change in y as x_1 changes, holding x_2 constant (B) Change in y as x_2 changes, holding x_1 constant.

method, which is one of the most frequently adopted approaches to estimate linear regression models and obtains the “best-fitting” line by minimizing the sum of squared residuals. We demonstrate the OLS approach using a simple regression model. By the end of this section, we introduce briefly the maximum likelihood estimation (MLE) approach.

Ordinary least squares (OLS)

Recall the simple regression model shown in Eq. (1),

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i. \quad (8)$$

Let b_0 and b_1 denote some random guesses of the unknown intercept coefficient β_0 and the slope coefficient β_1 . For any set of b_0 and b_1 , a model implied value of y_i can be computed by plugging values of the two guesses into Eq. (1). The differences between the model-implied and observed values of y_i are the so-called *residuals* (illustrated in Fig. 4).

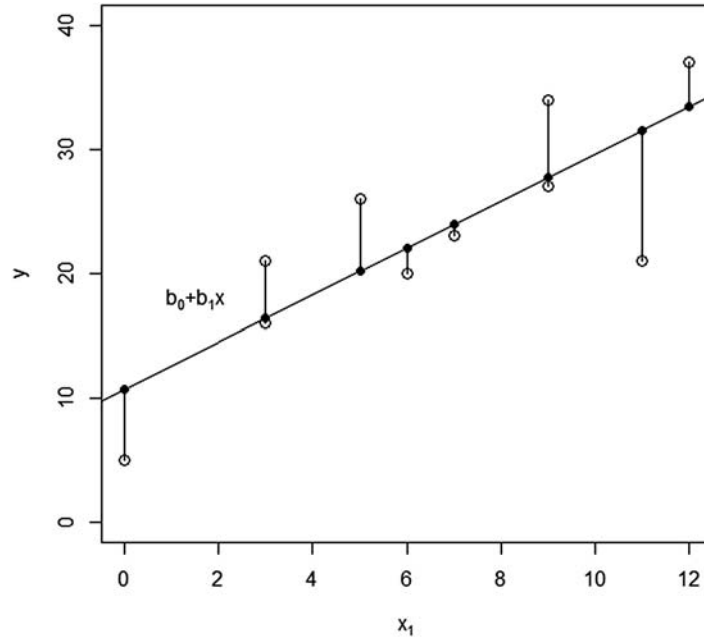


Fig. 4 Illustration of residuals The white dots are the observed data, while the black dots on the line defined by b_0 and b_1 (guesses of the regression coefficients) are the model-implied values. Differences between the observed and model-implied values of y are the residuals.

$$e_i = y_i - (b_0 + b_1 x_i). \quad (9)$$

The least squares estimator of the intercept and slope coefficients, denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$, are the values that minimize the sum of the squared residuals (SSR),

$$SSR = \sum_{i=1}^n e_i^2. \quad (10)$$

The values of $\hat{\beta}_0$ and $\hat{\beta}_1$ can be obtained analytically as,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}. \quad (12)$$

The $\bar{x}$ and $\bar{y}$ are, as defined in the previous section, the means of x_i and y_i ($i = 1, \dots, n$), respectively. In matrix form, the solution is,

$$\hat{\beta} = (X'X)^{-1}X'y, \quad (13)$$

where $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)$ is a vector that consists of estimates of all regression coefficients, and $(\cdot)^{-1}$ represents the inverse of the matrix in the parentheses.¹

Geometric interpretation of OLS

Let $\hat{y}_i$ ($i = 1, \dots, n$) denote the model-implied values of the dependent variable associated with the OLS estimator $\hat{\beta}_0$ and $\hat{\beta}_1$, and $\hat{e}_i$ be the corresponding residual,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (14)$$

$$\hat{e}_i = y_i - \hat{y}_i. \quad (15)$$

In other words, $\hat{y}_i$ is a dot on the “best-fitting” line of the data. All model predicted values can be stacked in a vector, denoted by $\hat{\mathbf{y}}$, and

$$\hat{\mathbf{y}} = X\hat{\beta}. \quad (16)$$

¹Note that this assumes that X is full column rank so that $X'X$ is invertible. When X is not full rank, alternative solutions such as generalized inverses (Searle, 1971) can be applied.

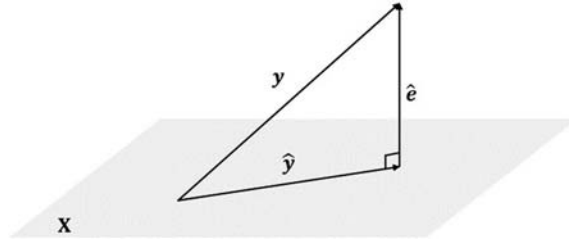


Fig. 5 Geometric interpretation of ordinary least squares estimation.

Similarly, let $\hat{e}$ denote the vector of $\hat{e}_i$ for $i = 1, \dots, n$.

The ordinary least squares estimation has a geometric interpretation. As shown in Fig. 5, the predicted value vector $\hat{y}$ is the projection of the observed values y onto the space spanned by columns of the design matrix X (a vector of ones and a vector of all x_i in the simple linear regression case). The vector y is the sum of $\hat{y}$ and $\hat{e}$, which is orthogonal to the space spanned by the columns of X . Since $\hat{e}$ is orthogonal to the space, its length is the smallest among residual vectors associated with all possible values of b_0 and b_1 .

Standard errors, confidence intervals and hypotheses testing of β

Once $\hat{\beta}$ are obtained, we can quantify the uncertainty around these point estimates of regression coefficients via the calculation of the associated standard errors. The OLS approach does not require that the disturbance term ε_i follows a normal distribution to produce unbiased point estimates of the regression coefficients and the variance of ε_i . The covariance matrix of $\hat{\beta}$, represented by the letter Σ , can be calculated as:

$$\Sigma = (X'X)^{-1}\hat{\sigma}^2. \quad (17)$$

In Eq. (17), $\hat{\sigma}^2$ denotes an unbiased estimator of the unknown variance of the disturbance terms σ^2 . $\hat{\sigma}^2$ is computed as:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \mathbf{x}_i' \hat{\beta})^2}{n - p - 1}, \quad (18)$$

in which $\mathbf{x}_i = (1, x_{i,1}, \dots, x_{i,p})$ consists of a one and values of p independent variables in the i th observation. The standard errors of $\hat{\beta}$ can be obtained by taking the square root of the diagonal elements of Σ .

If the normality assumption of ε_i holds, we can construct a confidence interval of a specific regression coefficient for any significance levels. For example, a 95% confidence interval of β_k (the k th element of β) is:

$$\left(\hat{\beta}_k - t^{0.975} \sqrt{\Sigma_{k,k}}, \quad \hat{\beta}_k + t^{0.975} \sqrt{\Sigma_{k,k}} \right), \quad (19)$$

where $t^{0.975}$ is the 97.5% percentile of a t distribution with $n - p - 1$ degrees of freedom. $\Sigma_{k,k}$ represents the k th diagonal element of the covariance matrix Σ . We are 95% confident that the true value of β_k lies within this confidence interval.

We can also conduct hypothesis testing of the regression coefficients if ε_i is assumed to follow a normal distribution. For example, to examine if β_k equals a constant c , we can form the below test statistic:

$$t_k = \frac{\hat{\beta}_k - c}{\sqrt{\Sigma_{k,k}}}. \quad (20)$$

Under the null hypothesis that $\beta_k = c$, this test statistic follows a t -distribution with $n - p - 1$ degrees of freedom. A t_k that is too large or too small would lead to the rejection of the null hypothesis and the conclusion that β_k is unlikely to be equal to c .

Maximum likelihood estimation (MLE)

In addition to the OLS approach, another widely-used approach to estimate linear regression models is the maximum likelihood estimation (MLE). Recall that the disturbance term ε_i is independent of the design matrix X and also independent across observations. If the normality assumption of ε_i holds, conditional on x_i , y_i would also follow a normal distribution with the mean $\beta_0 + \beta_1 x_i$ and variance σ^2 , $y_i \sim \mathcal{N}(\beta_0 + \beta_1 x_i, \sigma^2)$. The conditional probability density of y_i given x_i and the true model parameters is,

$$p(y_i | x_i; \beta_0, \beta_1, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y_i - (\beta_0 + \beta_1 x_i)}{\sigma} \right]^2}. \quad (21)$$

The conditional probability density of the data can be obtained by multiplying together the probability densities of all n observations,

$$\prod_{i=1}^n p(y_i|x_i; \beta_0, \beta_1, \sigma^2) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{y_i - (\beta_0 + \beta_1 x_i)}{\sigma} \right]^2}. \quad (22)$$

Let $L(b_0, b_1, s^2)$ denote the likelihood associated with some random guesses of the model parameters β_0 , β_1 , and σ^2 . With the MLE approach, we would like to find the values of b_0 , b_1 , and s^2 that maximizes $L(b_0, b_1, s^2)$. For convenience, we usually work with the log-likelihood, which is obtained by simply taking the logarithm of the likelihood. Parameter values that maximize the log-likelihood also maximize the likelihood,

$$l(b_0, b_1, s^2) = \log L(b_0, b_1, s^2) = \log \prod_{i=1}^n p(y_i|x_i; b_0, b_1, \sigma^2) = -\frac{n}{2} \log 2\pi - n \log s - \frac{1}{2s^2} \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2. \quad (23)$$

The MLE approach yields the same point estimates of regression coefficients as the OLS approach.

Assumptions of linear regression model

As alluded in the previous sections, the linear regression model along with the associated estimation approaches (such as the OLS) require certain assumptions about the independent variables, the disturbance term and their relationships to be satisfied. In this section, we briefly review a few major assumptions made by the linear regression model.

Linearity

The linear regression model assumes that the expectation of the scalar dependent variable is a linear combination of the independent variables and the regression coefficients. Note that the assumption of linearity is about the regression coefficients in the sense that the independent variables can be transformed arbitrarily. Thus, some basic expansions of the independent variables, such as $x_2 = x_1^2$, can also be included in the linear regression models. In other words, the following model:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2^2 + \varepsilon_i \quad (24)$$

is recognized as a linear regression model.

No multicollinearity among independent variables

A major assumption made by the linear regression model on the independent variables is that there exists no multicollinearity among them. Multicollinearity refers to the situation where values of a specific independent variable can be obtained by linearly combining other independent variables, such as $x_2 = 2x_1$. If there exists strong linear relationships among independent variables, columns of the design matrix $\mathbf{X}$ would not be linearly independent. In other words, $\mathbf{X}$ would not be full rank. This will lead to the consequence that estimates of the regression coefficients are not uniquely defined with the OLS approach. Multicollinearity is also used to describe situations where a combination of independent variables is highly correlated with another independent variable. Note that the independent variables remain linearly independent but the regression coefficients will be poorly estimated.

The variance inflation factor (VIF) quantifies the severity of multicollinearity and can be used to examine the assumption in the linear regression model. A large VIF indicates that there exists a strong correlation between the independent variables. If there is complete collinearity, then the VIF will be infinity.

Distribution of disturbance term

Assumptions are also made about the distribution of the disturbance term and the relationship between the disturbance term and the independent variable(s). For example, with the OLS approach, the disturbance term has to follow a normal distribution to create confidence intervals and conduct hypothesis testing of the regression coefficients.

In addition, the disturbance term is assumed to be independent with the independent variable(s) and also with each other. This independence assumption can be easily violated in time series data where an individual is measured at multiple time points. Approaches such as the generalized least squares (GLS) can be used to handle the correlated disturbances.

Homoscedasticity

The assumption of homoscedasticity, also known as the homogeneity of variance assumption, states that the variances of the disturbance terms are the same across observations. Stated in another way, all elements on the diagonal of the covariance matrix of $\boldsymbol{\varepsilon}$ have the same value.

The homoscedasticity assumption can be examined by plotting the residuals against the dependent variable. If this assumption is violated, the weighted least squares (WLS) approach can be applied to deal with the problem.

Extensions of linear regression model

In a standard linear regression model, the dependent variable is a continuous variable and would follow a normal distribution if the disturbance term is normally distributed. In this section, we introduce three extensions of the standard linear regression model: the generalized linear model (GLM), the hierarchical linear model (HLM), and the multivariate linear model. The GLM allows the dependent variable to follow an arbitrary distribution. The HLM accommodates data with more than one level. The multivariate linear model allows the modeling of the relationship among independent variables and multiple dependent variables.

Generalized linear model

The generalized linear model (GLM; [Nelder and Wedderburn, 1972](#)) extends the linear regression model and can be viewed as a general framework. With a GLM, data whose dependent variable follows distributions other than the normal distribution can be modeled appropriately. A GLM consists of three elements: a dependent variable that can follow an arbitrary distribution, a linear predictor, which is a linear combination of the independent variables and regression coefficients, and a link function,

$$\mathbb{E}[y_i | \mathbf{x}_i] = \mu_i \quad (25)$$

$$g(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta}. \quad (26)$$

In [Eq. \(25\)](#), μ_i is the conditional expectation of the dependent variable y_i given the independent variables $\mathbf{x}_i$, $\mathbf{x}_i' \boldsymbol{\beta}$ is the linear predictor, and $g(\cdot)$ the so-called *link function*, which is a known function that maps the linear predictor to the desired distribution. For example, when the dependent variable is assumed to follow a Bernoulli distribution (i.e., y_i can only take the values of 0 and 1), the logit link function can be applied so that the linear predictor (which ranges from $-\infty$ to ∞) is mapped to the 0–1 range,

$$\mu_i = \frac{1}{1 + \exp(-\mathbf{x}_i' \boldsymbol{\beta})}. \quad (27)$$

Another example is that if the dependent variable follows a Poisson distribution and takes the values of nonnegative integers, a log link function can be applied,

$$\log(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta}. \quad (28)$$

The standard linear regression model presented in previous sections can be viewed as a special case of the generalized linear model. The link function in a standard linear regression model is the identity link function,

$$\mu_i = \mathbf{x}_i' \boldsymbol{\beta}. \quad (29)$$

Hierarchical linear model

The hierarchical linear model (HLM; [Raudenbush and Bryk, 2002](#)), which is also known as the multilevel model ([Goldstein, 2011](#)), is another extension of the standard linear regression model. The HLM can be applied to data with more than one level, which are prevalent in social and behavioral sciences. A canonical example in the field of education of two-level data is that students are nested within classrooms. A two-level HLM can be formulated as follows: at level 1 (the student level), the outcome of the i th student in classroom j is expressed as the linear combination,

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij} \quad (30)$$

in which β_{0j} and β_{1j} are the group-specific intercept and slope coefficients, x_{ij} is a level-1 independent variable (i.e., a student-related variable), and ε_{ij} is the associated disturbance term, $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. At level 2 (the classroom level), the group-specific intercept and slope coefficients are modeled,

$$\beta_{0j} = \gamma_{00} + \gamma_{01}w_j + u_{0j} \quad (31)$$

$$\beta_{1j} = \gamma_{10} + u_{1j}, \quad (32)$$

where γ_{00} and γ_{10} are the average intercept and slope across groups, w_j is a level-2 independent variable (i.e., a classroom-related variable), γ_{01} is the associated regression coefficients and represents how the group-specific intercept changes as w_j changes. u_{0j} and u_{1j} are deviations of group-specific coefficients to the averages, and are assumed to follow a bivariate normal distribution,

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix} \right].$$

Note that the off-diagonal elements σ_{01} and σ_{10} can be nonzero, indicating the deviations of the intercept and slope can be correlated.

In addition to the example discussed above, more complex models for multiple level-1 and level-2 independent variables and more complex models for level-1 coefficients are possible and widely used, such as the cross-classified multilevel model (Goldstein, 1994).

Multivariate linear model

If there are more than one dependent variable in the observations, and a linear relationship is assumed between the independent and dependent variables, the multivariate regression model can be used to model the relationship. The multivariate regression model takes the following form:

$$\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \boldsymbol{\varepsilon}_i. \quad (33)$$

In Eq. (33), $\mathbf{y}_i$ is a vector of size m and includes values of m dependent variables in the observation i . $\mathbf{x}_i$ is a vector of independent variables and is of size p , and $\mathbf{B}$ is the corresponding regression coefficient matrix, the size of which is $m \times p$. $\boldsymbol{\varepsilon}_i$ is a vector of the disturbance terms and is often assumed to follow a multivariate normal distribution with mean zeros and covariance matrix Ψ , $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_m(0, \Psi)$. The disturbance terms in the same observation may have different variances and can be correlated. The generalized least squares (GLS) approach can be used to obtain estimates of the model parameters in the multivariate regression model.

Conclusion

In this article, we present the formulation of the standard linear regression model, which is one of the most broadly applied statistical techniques to investigate the relationship between one dependent variable and one or more than one independent variable. We review how to obtain estimates of the regression coefficients and the variance of the disturbance term via the approaches of ordinary least squares (OLS) and maximum likelihood estimation (MLE). We also discuss a few assumptions that underlie the linear regression model, including assumptions about independent variables, distribution of the disturbance term, and the relationship among variables. In addition, we show three major extensions of the linear regression model, the generalized linear model (GLM), hierarchical linear model (HLM), and multivariate regression model.

Last but not least, although not discussed in this entry, we would like to highlight that linear regression model also plays a fundamental role in the field of statistical learning, which is also known as machine learning. We conclude that the regression analysis will remain an active research area, and the linear regression model and its extensions will be applied successfully in various scientific fields.

References

- Aldrich, J., 2005. Fisher and regression. *Stat. Sci.* 401–417.
- Fisher, R.A., 1922. The goodness of fit of regression formulae, and the distribution of regression coefficients. *J. Roy. Stat. Soc.* 85, 597–612.
- Galton, F., 1886. Regression towards mediocrity in hereditary stature. *J. Anthropol. Inst. G. B. Ireland* 15, 246–263.
- Galton, F., 1888. Co-relations and their measurement, chiefly from anthropometric data. *Proc. Roy. Soc. Lond.* 45, 135–145.
- Gauss, C.F., 1809/1963. *Theoria Motus Corporum Coelestium*. Dover, New York. Reprinted 1963.
- Goldstein, H., 1994. Multilevel cross-classified models. *Socio. Methods Res.* 22, 364–375.
- Goldstein, H., 2011. *Multilevel Statistical Models*, fourth ed. Wiley, London.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning*, vol. 112. Springer.
- Nelder, J.A., Wedderburn, R.W., 1972. Generalized linear models. *J. Roy. Stat. Soc.* 135 (3), 370–384.
- Pearson, K., 1895. Note on regression and inheritance in the case of two parents. *Proc. Roy. Soc. Lond.* 58, 240–242.
- Pearson, K., 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. Lond.* 187, 253–318.
- Raudenbush, S.W., Bryk, A.S., 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*, second ed. SAGE.
- Searle, S.R., 1971. *Linear Models*. John Wiley & Sons.
- Slutsky, E., 1913. On the criterion of goodness of fit of the regression lines and on the best method of fitting them to the data. *J. Roy. Stat. Soc.* 77, 78–84.
- Stanton, J.M., 2001. Galton, Pearson, and the Peas: a brief history of linear regression for statistics instructors. *J. Stat. Educ.* 9.

Nonlinear regression analysis

Hsin-Hsiung Huang and Qing He, Department of Statistics and Data Science, University of Central Florida, Orlando, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	558
An overview of the nonlinear regression models	558
Estimation methods	558
Gauss-Newton method for nonlinear least-squares problems	559
Maximum likelihood estimation	560
Asymptotic statistical inference	560
Parameterization and curvature	561
Newton-Raphson method	562
Variance heterogeneity	564
Model building	564
Simulation data analysis	566
Further reading	566
Acknowledgment	567
References	567

Introduction

An overview of the nonlinear regression models

Nonlinear regression models have been widely used in various fields including statistics, chemistry, physics, psychology, health science, and biology. Some of them have a linear relationship in the parameters (i.e., linear in the θ). For example, a polynomial regression fits a curved relationship between the response variable and predictors using higher-ordered values of the predictors, but it is linear in terms of the parameters. Statistical inference can be derived from nonlinear regression models (Eq. 1). The main advantages of nonlinear regression models include interpretability, parsimony, and prediction (Bates and Watts, 1988). In general, nonlinear models are capable of accommodating various mean functions, so nonlinear models are appropriate for applications with parsimonious parameters and easily interpretable due to the fact that the parameters can be associated with meaningful factors.

Assume that there are n observations of responses $y_1, \dots, y_n$ and predictors $\mathbf{x}_{ik}$ associated with y_i on $k = 1, \dots, p$ independent variables. In nonlinear regression, the y 's and $\mathbf{x}$'s satisfy the nonlinear regression model

$$y_i = f(\mathbf{x}_i, \theta) + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where f is a function that is nonlinear in the p elements of parameter θ with the unknown true value θ^* , and the ε_i 's are independent and identically distributed (i.i.d) random errors with mean zero and constant variance σ^2 . The conditional mean response $E(y_i | \mathbf{x}_i) = f(\mathbf{x}_i, \theta)$, so we will refer to f as the mean function. For example, the Beverton-Holt model (Beverton and Holt, 2012) which is similar to the Michaelis-Menten model (Sheiner and Beal, 1980) and used for modeling discrete-time population gives the expected population $y = n_{t+1}$ as a function of the previous population $\mathbf{x} = n_t$ as

$$n_{t+1} = f(\mathbf{x}, (\alpha, \beta)) = \frac{\alpha \mathbf{x}}{1 + \mathbf{x}/\beta},$$

where $\theta = (\alpha, \beta)$ and the parameter α is the slope at 0 and β is the concentration between 0 and the upper limit $\alpha\beta$. It is a nonlinear regression since the two variables are related in a nonlinear (curved) relationship.

Estimation methods

In order to estimate the parameters in a nonlinear regression model, we minimize the sum of squared residuals, which measures how far the y observations vary from the nonlinear function $f(\mathbf{x}_i, \theta)$ used to predict y . Such estimators are called the least squares estimators $\hat{\theta}$ of θ , that is,

$$S(\theta) = \sum_{i=1}^n (y_i - f(\mathbf{x}_i, \theta))^2$$

is used to find

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} S(\boldsymbol{\theta}).$$

For example, to estimate α and β in the Beverton-Holt model, we minimize $S(\alpha, \beta) = \sum \left(y_i - \frac{\alpha x_i}{1 + x_i/\beta} \right)^2$. We first derive the gradients g_α, g_β and Hessian matrix with elements H_{ij} , $i, j \in \{\alpha, \beta\}$ as follows:

$$\begin{aligned} g_\alpha &= \frac{\partial S(\boldsymbol{\theta})}{\partial \alpha} = -2 \sum \mathbf{x}_i \left(y_i - \frac{\alpha x_i}{1 + x_i/\beta} \right), \\ g_\beta &= \frac{\partial S(\boldsymbol{\theta})}{\partial \beta} = -2 \sum \left(y_i - \frac{\alpha x_i}{1 + x_i/\beta} \right) \frac{\alpha x_i^2}{\beta^2 (1 + x_i/\beta)^2}, \\ H_{\alpha, \alpha} &= \frac{\partial g_\alpha}{\partial \alpha} = 2 \sum \frac{x_i^2}{1 + x_i/\beta}, \\ H_{\alpha, \beta} &= \frac{\partial g_\alpha}{\partial \beta} = 2 \sum \frac{\alpha x_i^3}{\beta^2 (1 + x_i/\beta)^2}, \\ H_{\beta, \alpha} &= -2 \sum \frac{x_i^2 y_i}{\beta^2 (1 + x_i/\beta)^2} - \frac{x_i^3}{\beta^2 (1 + x_i/\beta)^3}, \text{ and} \\ H_{\beta, \beta} &= -2 \sum \frac{-2\alpha x_i^2 y_i}{(\beta + x_i)^3} + \frac{2\alpha^2 x_i^3}{\beta^3 (1 + x_i/\beta)^3} - \frac{3\alpha^2 x_i^4}{\beta^4 (1 + x_i/\beta)^4}. \end{aligned}$$

Then we can use the Gauss-Newton or Newton-Raphson methods to iteratively estimate α and β introduced in the **Gauss-Newton method for least squares estimation** and **Newton-Raphson method** sections, respectively. The Gauss-Newton method is essentially the Newton-Raphson method with the modification that the Gauss-Newton method uses the approximation $2J^T J$, where J is the Jacobian matrix, for the Hessian matrix. Note that for a scalar nonlinear function, its Jacobian matrix is the same as the gradient.

Gauss-Newton method for nonlinear least-squares problems

When f is twice differentiable with respect to $\boldsymbol{\theta}$, we have the gradient and Hessian which can be written as:

$$\begin{aligned} g_j &= \frac{\partial S(\boldsymbol{\theta})}{\partial \theta_j} = -2 \sum_{i=1}^n (y_i - f_i) \frac{\partial f_i}{\partial \theta_j}, \\ H_{jk} &= \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k} = -2 \sum_{i=1}^n \left((y_i - f_i) \frac{\partial^2 f_i}{\partial \theta_j \partial \theta_k} - \frac{\partial f_i}{\partial \theta_j} \frac{\partial f_i}{\partial \theta_k} \right) \end{aligned}$$

where $f_i = f(\mathbf{x}_i, \boldsymbol{\theta})$.

Further, we solve for the least-squares solution $\hat{\boldsymbol{\theta}}$ in the following system of equations

$$\left. \frac{\partial S(\boldsymbol{\theta})}{\partial \theta_j} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0; \quad j = 1, \dots, p, \quad (2)$$

which are called the normal equations with

$$\sum_{i=1}^n (y_i - f_i) \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_j} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0.$$

Its matrix form is

$$\mathbf{J}(\hat{\boldsymbol{\theta}})^T \mathbf{r} = 0$$

with (i, j) -th element $J_{ij}(\hat{\boldsymbol{\theta}}) = \frac{\partial f_i}{\partial \theta_j}$ and $\mathbf{r} = \mathbf{y} - f(\hat{\boldsymbol{\theta}})$ with the i -th element $r_i = y_i - f_i$. The matrix $\mathbf{J}(\hat{\boldsymbol{\theta}})$ of size $n \times p$ is called the Jacobian matrix. Define the $p \times p$ matrices $G_1, \dots, G_n$ so that the (j, k) -th element of G_i is $\frac{\partial^2 f_i}{\partial \theta_j \partial \theta_k}$, and then the gradient and Hessian are given by

$$\mathbf{g} = -2\mathbf{J}^T \mathbf{r}, \quad \mathbf{H} = 2\mathbf{J}^T \mathbf{J} - 2 \sum_{i=1}^n r_i G_i. \quad (3)$$

In a linear regression model, the Jacobian is simply the data design matrix $\mathbf{X}$ and does not depend on the parameter values $\boldsymbol{\theta}$.

The normal equations do not have an analytic solution for $\hat{\theta}$ in most cases, so that numerical iterative procedures are needed. The Gauss-Newton method estimates the parameters in nonlinear regression using the first-order Taylor's expansion

$$f(\mathbf{x}, \theta) \approx f(\mathbf{x}, \theta^*) + \nabla f(\mathbf{x}, \theta)^T (\theta - \theta^*)$$

to approximate a nonlinear regression function (Nadaraya, 1964; Ruppert and Wand, 1994) considering a small neighborhood of θ^* . The linear approximation of $f(\mathbf{x}, \theta)$ in the neighborhood of θ^* results in an approximate residual sum of squares

$$S(\theta) \approx \sum_{i=1}^n \left(y_i - f(\mathbf{x}_i, \theta^*) - \sum_j^p J_{ij}^T (\theta_j - \theta_j^*) \right)^2,$$

and

$$S(\theta) - S(\theta^*) \approx (\hat{\theta} - \theta^*)^T \mathbf{J}^T \mathbf{J} (\hat{\theta} - \theta^*). \quad (4)$$

The corresponding normal equations at the t -th iteration are given by

$$\mathbf{J}(\theta^{(t)})^T \mathbf{J}(\theta^{(t)}) (\theta - \theta^{(t)}) = \mathbf{J}(\theta^{(t)})^T (y - f(\theta^{(t)})).$$

The updating increment

$$\delta^{(t)} = \left(\mathbf{J}(\theta^{(t)})^T \mathbf{J}(\theta^{(t)}) \right)^{-1} \mathbf{J}(\theta^{(t)})^T (y - f(\theta^{(t)}))$$

in the t -th iteration, and then $\theta^{(t+1)} = \theta^{(t)} + \lambda_t \delta^{(t)}$ where λ_t is the step size. The update direction $\delta^{(t)}$ is derived from the tangent plane approximation to the solution locus.

The estimate of the asymptotic covariance of the least squares estimate is given by

$$\text{cov}(\hat{\theta}) = \sigma^2 (\mathbf{J}^T \mathbf{J})^{-1}.$$

The statistical inference about θ is based on the property that $\hat{\theta}$ follows $N(\theta, \sigma^2 (\mathbf{J}^T \mathbf{J})^{-1})$ asymptotically. σ^2 is estimated by $s^2 = S(\hat{\theta})/(n-p)$, where p is the number of parameters and $\mathbf{J}$ is estimated by $\mathbf{J}(\hat{\theta})$.

Maximum likelihood estimation

Considering a normal distributed error term

$$y_i - f(\mathbf{x}_i, \theta) \stackrel{i.i.d.}{\sim} N(0, \sigma^2),$$

the log-likelihood function of θ and σ^2 is

$$l(\theta, \sigma^2) = -\frac{S(\theta)}{2\sigma^2} - \frac{n}{2} \log(2\pi\sigma^2).$$

Maximum likelihood estimation (MLE) is a method of estimating the parameters by maximizing $l(\theta, \sigma^2)$ with respect to θ and σ^2 for which we need the following derivatives

$$\frac{\partial l}{\partial \sigma^2} = \frac{1}{2\sigma^4} S(\theta) - \frac{n}{2\sigma^2}, \quad (5)$$

$$\frac{\partial l}{\partial \theta} = -\frac{1}{2\sigma^2} \frac{\partial S(\theta)}{\partial \theta}. \quad (6)$$

Setting Eqs. (5) and (6) equal to 0, we have $\hat{\sigma}_{\text{MLE}}^2 = \frac{S(\hat{\theta})}{n}$ and the normal Eq. (2). As a result, $\hat{\theta}_{\text{MLE}} = \hat{\theta}$. In summary, when the noise follows a normal distribution, the Gauss-Newton iterations to minimize the least squares $S(\theta)$ are exactly the same as the Newton-Raphson iterations to maximize the log-likelihood function with respect to θ , and hence the maximum likelihood estimator is equivalent to the least squares estimator (Bates and Watts, 1988).

Asymptotic statistical inference

Let us consider the nonlinear regression model (Eq. 1) where the $\varepsilon_i \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. We note that as $n \rightarrow \infty$, the least squares estimate $\hat{\theta}$ is asymptotically $N(\theta^*, \sigma^2 (\mathbf{J}^T \mathbf{J})^{-1})$, where θ^* is the true value of θ , and the matrix $\mathbf{J} = [\partial f(\mathbf{x}_i; \theta) / \partial \theta_j]_{\theta^*}$ plays the same role as X in the linear regression (Saber and Wild, 1989). In particular, by analogy with the linear confidence region

$$\left\{ \theta \mid (\theta - \hat{\theta})^T \hat{J} (\theta - \hat{\theta}) \leq p s^2 F_{p, n-p}^\alpha \right\}, \quad (7)$$

is an approximate $100(1 - \alpha)\%$ confidence region for θ . Here $s^2 = S(\hat{\theta})/(n - p)$, $\hat{J} = J(\hat{\theta})$, and $F_{p, n-p}^\alpha$ is the upper critical value of the $F_{p, n-p}$ distribution. As the linear approximation is valid asymptotically, (Eq. 7) will have the correct confidence level of $100(1 - \alpha)\%$ asymptotically. As α varies, the regions (Eq. 7) are ellipsoid contours of the approximate multivariate normal density function of $\hat{\theta}$ (with J replaced by $\hat{J}$). Since $S(\theta)$ measures how close the observations are to the fitted equation for any θ , it would seem appropriate to also base confidence regions for θ on the contours of $S(\theta)$. Such a region could take the form

$$\left\{ \theta \mid S(\theta) \leq c S(\hat{\theta}) \right\}$$

for some c , $c > 1$. Regions of this type are often called the exact confidence regions, as they are not based on any approximations. However, the confidence levels, or coverage probabilities, of such regions are generally unknown, though approximate levels can be obtained from asymptotic theory.

For large enough n , the consistent estimator $\hat{\theta}$ will be sufficiently close to θ^* for the approximation (Eq. 4) to hold with $F_{p, n-p}^\alpha$ in Eq. (7). Therefore, substituting

$$S(\theta^*) - S(\theta) \approx (\hat{\theta} - \theta^*)^T \hat{J}^T \hat{J} (\hat{\theta} - \theta^*) \quad (8)$$

from Eq. (4), we obtain the confidence region

$$\left\{ \theta \mid S(\theta) \leq S(\hat{\theta}) \left(1 + \frac{p}{n - p} F_{p, n-p}^\alpha \right) \right\} \quad (9)$$

that has the required asymptotic confidence level of $100(1 - \alpha)\%$. The regions Eqs. (7) and (9) are asymptotically the same, and they are identical for linear models. However, for finite n , these regions may be very different, and it indicates inadequacy of the linear approximation (Eq. 4). Under the assumption of normal errors we see that the above clash between the two confidence regions embodies a wider principle. We have a choice between confidence regions based on the asymptotic normality of the maximum likelihood estimator $\hat{\theta}$, which is also the least squares estimator, and those based on the contours of the likelihood function via the likelihood ratio test.

Parameterization and curvature

There are some nonlinear models which are intrinsically linear because they can be made linear in the parameters by a simple transformation. For example, the reciprocal of the Michaelis-Menten regression mean function

$$f(x, \mathbf{b}) = \frac{b_1 x}{b_2 + x} \quad (10)$$

results in the model

$$\frac{1}{y} = \frac{1}{b_1} + \frac{b_2}{b_1} \frac{1}{x} + \varepsilon = \beta_1 + \beta_2 \frac{1}{x} + \varepsilon,$$

which is linear in the transformed parameters β_1 and β_2 . In such cases, transforming a model to its linear form often provides better inference procedures and confidence intervals. However, parameterization may affect the asymptotic variance and convergence of the estimation algorithm (Bates and Watts, 1988; Huang et al., 2010). The re-parameterization is an important issue. Although the shape of the solution locus is fixed, the performance of the Gauss-Newton estimator varies with respect to different parameterizations. This depends on the nonlinearity in the model. There are two kinds of nonlinearity—*intrinsic nonlinearity* and *parameter-effects nonlinearity* (Hamilton and Watts, 1985; Bates and Watts, 1988). The *intrinsic nonlinearity* is associated with the modeling and is invariant under parameterization. The *parameter-effects nonlinearity*, however, can be lessened through a proper parameterization. If either component of the nonlinearity is large, the least squares estimate may not converge. Furthermore, the estimated covariance for $\hat{\mathbf{b}}$ given by $\hat{V}(\hat{\mathbf{b}}) = \hat{\sigma}^2 (J(\hat{\mathbf{b}})^T J(\hat{\mathbf{b}}))^{-1}$, would change greatly in each step of the iteration, and the statistical inference based on the asymptotic normality becomes unreliable. In other words, the least squares estimator in nonlinear regression depends on the curvedness of the underlying model as well as the parameterization adopted.

The quality of the linear approximation can be summarized by *intrinsic curvature* $\kappa_d^N = \|d^T C_N d\|$ and *parameter-effects curvature* $\kappa_d^T = \|d^T C_T d\|$ (Bates and Watts, 1988), where d has unit length, C_N is the normal direction curvature matrix, and C_T is the tangent direction curvature matrix. For details on curvatures, please read the Appendix in Huang et al. (2010).

The two curvatures correspond to two assumptions of linear approximation of a nonlinear mean function. First, the planar assumption ensures that the nonlinear mean function is approximated by its tangent plane at a given point. Second, the uniform coordinate assumption means that an equi-spaced, straight and parallel linear coordinate system is placed on the approximation tangent plane. *Intrinsic curvature* is related to the planar assumption, and it depends on the dataset considered and the mean function but not on the parameterization used in the mean function. *Parameter-effects curvature* is related to the uniform coordinate

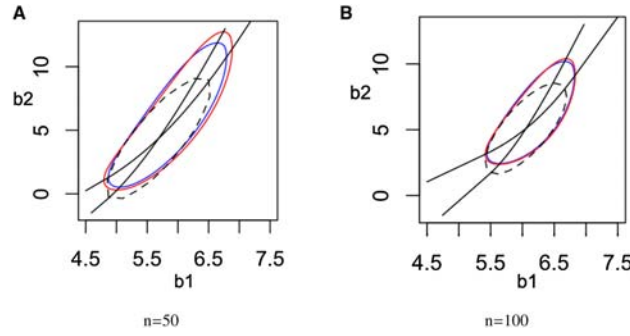


Fig. 1 The profile plots of the simulation data analysis. The 95% confidence region of b_1 and b_2 using the Wald statistic (dashed), the likelihood ratio (blue) and Skovgaard's approximation (red) with simulated data from the Michaelis-Menten model. The left panel is with sample size $n = 50$, and the right panel is with sample size $n = 100$.

assumption, and it depends on all aspects of the model, including the parameterization. Large values of these two curvature measures indicate a poor linear approximation. The function `rms.curv()` in the R package MASS can be used to calculate the two measures for a given `nls()` model fit (Venables and Ripley, 2002). The intrinsic curvature is generally relatively negligible compared with the parameter-effects curvature (Bates and Watts, 1988), and in such cases the profile likelihood approach is useful.

The Wald, likelihood ratio (LR) and score tests are asymptotically equivalent. They all have the same sampling distribution which is a chi-square distribution with the same degrees of freedom. However, the Wald test statistics $(\hat{\mathbf{b}} - \mathbf{b}_0)^T [\hat{\mathbf{V}}(\hat{\mathbf{b}})]^{-1} (\hat{\mathbf{b}} - \mathbf{b}_0)$, which results in an ellipsoidal confidence region, uses $\hat{\mathbf{b}}$ and depends on curvature of the likelihood at $\hat{\mathbf{b}}$. The score test depends on the slope and curvature at $\hat{\mathbf{b}}_0$, while the LR test uses information from both $\hat{\mathbf{b}}$ and $\hat{\mathbf{b}}_0$. The differences among them vanish in large samples if the null hypothesis is true. If the null hypothesis is false, they may take very different values. The Wald procedures are influenced by both intrinsic curvature and parameter-effects curvature. A model with a lower parameter-effects curvature is preferable.

The curvature can affect the quality of statistical inferences based on the asymptotic normality. When the mapped parameter curves onto the tangent plane are not uniform grid lines, the resulting confidence region may not be reliable. We use the Michaelis-Menten model (Eq. 10) as an example by simulating data with $\varepsilon_i \stackrel{i.i.d.}{\sim} N(0, 1^2)$, $b_1 = 6$, $b_2 = 5$, and $\mathbf{x}$ generated from the uniform distribution (1, 100). We demonstrate how confidence regions of b_1 and b_2 are affected by curvatures with two different sample sizes $n_1 = 50$ and $n_2 = 100$ displayed in Fig. 1 using the contour method of the nlreg package (Brazzale, 2005). The dashed, blue and red lines represent the confidence regions of (b_1, b_2) obtained with the Wald test, the likelihood ratio test, and Skovgaard's approximation, respectively (Skovgaard, 1996). The confidence regions of the latter two methods are closer to the exact one than the Wald (Bates and Watts, 1988; Skovgaard, 1996). When the sample size is small, the shapes and coverage between the Wald confidence region and the two others are very different. As the sample size grows, they become less different.

Newton-Raphson method

The Newton-Raphson method (also called Newton's method) which has been widely used for estimating the parameters of nonlinear regression models due to its advantage of fast convergence using the second order Taylor's expansion to approximate the nonlinear function $l(\boldsymbol{\theta})$ as a quadratic polynomial as follows:

$$l(\boldsymbol{\theta}) \approx l(\boldsymbol{\theta}^{(t)}) + (\boldsymbol{\theta} - \boldsymbol{\theta}^{(t)})^T \nabla l(\boldsymbol{\theta}^{(t)}) + \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}^{(t)})^T \nabla^2 l(\boldsymbol{\theta}^{(t)}) (\boldsymbol{\theta} - \boldsymbol{\theta}^{(t)}), \quad (11)$$

where $\boldsymbol{\theta}^{(t)}$ is in a neighborhood of the true $\boldsymbol{\theta}$, ∇ is the first derivative and ∇^2 is the second derivative with respect to $\boldsymbol{\theta}$. The Newton-Raphson method can be computationally challenging and heavily dependent on good starting values (Gill et al., 1986). To maximize the quadratic approximation, we set the first derivative of the approximate quadratic function (Eq. 11) equal to 0 at a new point $\boldsymbol{\theta}^{(t+1)}$.

$$\begin{aligned} \nabla l(\boldsymbol{\theta}^{(t)}) + [\nabla^2 l(\boldsymbol{\theta}^{(t)})] (\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^{(t)}) &= \mathbf{0}_p, \\ \boldsymbol{\theta}^{(t+1)} &= \boldsymbol{\theta}^{(t)} - [\nabla^2 l(\boldsymbol{\theta}^{(t)})]^{-1} \nabla l(\boldsymbol{\theta}^{(t)}) \\ &= \boldsymbol{\theta}^{(t)} + [-\nabla^2 f(\boldsymbol{\theta}^{(t)})]^{-1} \nabla f(\boldsymbol{\theta}^{(t)}) \end{aligned}$$

which approximates $\boldsymbol{\theta}$ in each iteration t . The value $\boldsymbol{\theta}^{(t+1)}$ should be a better guess than the initial point $\boldsymbol{\theta}^{(t)}$ is.

Here, we focus on deriving the Newton-Raphson method for the exponential family (Nelder and Wedderburn, 1972) including the Gaussian (normal) distributions for continuous variables, binomial distributions for binary response variables, Poisson

distributions for count variables, and the gamma and inverse-Gaussian families of distributions for modeling positive continuous data, where the conditional variance of y_i increases with its expectation. The log-likelihood for n independent observations from an exponential family is given by

$$L_n(\boldsymbol{\theta}, \phi; \boldsymbol{y}) = \sum_{i=1}^n \left\{ \frac{y_i \theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi) \right\}. \quad (12)$$

Assume that the mean μ_i depends linearly on the covariates through a link function g as follows:

$$g(\mu_i) = \eta_i = \mathbf{X}_i^T \boldsymbol{\beta}.$$

Because the link function is invertible, we can also write

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(\mathbf{X}_i^T \boldsymbol{\beta}).$$

Note that

$$\mu_i = b'(\theta_i), \theta_i = (b')^{-1} \circ g^{-1}(\mathbf{X}_i^T \boldsymbol{\beta}) := h(\mathbf{X}_i^T \boldsymbol{\beta})$$

where

$$b'(\theta_i) = \frac{\partial b(\theta_i)}{\partial \theta_i}, g'(\mu_i) = \frac{\partial g(\mu_i)}{\partial \mu_i}, \text{ and } h'(\boldsymbol{\beta}) = \frac{\partial h(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}.$$

According to the chain rule, we have

$$\begin{aligned} \frac{\partial L_n}{\partial \beta_j} &= \sum_{i=1}^n \frac{\partial L_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \beta_j} \\ &= \sum_{i=1}^n \frac{y_i - \mu_i}{\phi} h'(\mathbf{X}_i^T \boldsymbol{\beta}) X_{i,j} := \sum_{i=1}^n (y_i - \mu_i) W_i X_{i,j}, \end{aligned}$$

where

$$W_i = h'(\mathbf{X}_i^T \boldsymbol{\beta}) / g'(\mu_i) \phi.$$

Define $\mathbf{W} = \text{diag}(W_1, \dots, W_n)$, then the gradient is

$$\nabla L_n(\boldsymbol{\beta}) = \sum_{i=1}^n \frac{(y_i - \mu_i) \mu_i'(\eta_i)}{\sigma_i^2} \mathbf{X}_i = \mathbf{X}^T \mathbf{W}(\mathbf{y} - \boldsymbol{\mu}).$$

For the Hessian, we have

$$\frac{\partial^2 L_n}{\partial \beta_j \partial \beta_k} = \sum_{i=1}^n \frac{y_i - \mu_i}{\phi} h''(\mathbf{X}_i^T \boldsymbol{\beta}) X_{i,j} X_{i,k} - \frac{1}{\phi} \sum_{i=1}^n \left(\frac{\partial \mu_i}{\partial \beta_k} \right) h'(\mathbf{X}_i^T \boldsymbol{\beta}) X_{i,j}.$$

Note that

$$\frac{\partial \mu_i}{\partial \beta_k} = \frac{\partial b'(\theta_i)}{\partial \beta_k} = \frac{\partial b'(h(\mathbf{X}_i^T \boldsymbol{\beta}))}{\partial \beta_k} = b''(\theta_i) h'(\mathbf{X}_i^T \boldsymbol{\beta}) X_{i,k}.$$

So,

$$-E[\mathbf{H}_{L_n(\boldsymbol{\beta})}] = \mathbf{X}^T \mathbf{W} \mathbf{X},$$

where $\mathbf{W} = \text{diag}\left(\frac{h'(\mathbf{X}_i^T \boldsymbol{\beta})}{g'(\mu_i)}\right)$. It leads to the Hessian matrix $\mathbf{H}(\boldsymbol{\theta}) = -\mathbf{X}^T \mathbf{W} \mathbf{X}$. Fisher scoring is equivalent to the Newton-Raphson method when we use the canonical link in logistic regression. Since $y_i \in (0, 1)$, $-\mathbf{X}^T \mathbf{W} \mathbf{X}$ is strictly negative definite. When y_i is too close to 0 or 1, weights are close to 0 and $\mathbf{H}$ may have singular values close to 0 and therefore may be computationally singular. The Newton-Raphson algorithm (Green, 1984; Cleveland and Devlin, 1988; Simonoff and Tsai, 1991) iterates according to

$$\begin{aligned} \boldsymbol{\beta}^{(t+1)} &= \boldsymbol{\beta}^{(t)} + (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{y} - \boldsymbol{\mu}) \\ &= (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{y} - \boldsymbol{\mu} + \mathbf{X} \boldsymbol{\beta}^{(t)}), \end{aligned} \quad (13)$$

which is the iteratively reweighted least squares (IRLS) algorithm (Ruppert and Wand, 1994).

Remarks: IRLS solves a generalized linear model in each iteration

$$(\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{z}^{(t)},$$

where $z^{(t)} = y - \mu + X\beta^{(t)}$ is the residual. When W is symmetric positive definite this is equivalent to the solution of the weighted least squares problem:

$$\min_{\beta} \sum_{i=1}^n \left(W^{1/2} (X_i^T \beta - y_i) \right)^2.$$

Note that in the logistic regression case $g'(\mu) = g(\mu)(1 - g(\mu))$ is never zero, so that the diagonal entries of W are never zero. However, other link functions might lead to degenerate cases with indefinite weighting matrices. When the link function $g(\mu)$ is the identity function and the error follows Gaussian with mean zero and variance one ($N(0, 1)$), then the IRLS algorithm becomes

$$\begin{aligned} \beta^{(1)} &= 0, \\ \mu &= X\beta^{(1)} = 0, \\ z &= \eta + \frac{y - g(\mu)}{g'(\mu)} = 0 + \frac{y - 0}{1} = y, \\ W &= \text{diag} \left(\frac{g'(\mu)^2}{(g'(\mu))^2} \right) = \text{diag}(1^2/1) = I_n, \\ \beta^{(2)} &= 0 + (X^T X)^{-1} X^T y, \end{aligned}$$

which is the ordinary least squares estimator.

Variance heterogeneity

When the residual errors show a trend (e.g., increasing variability as the explanatory variable increases), this can be addressed by modeling the variance as a function of the independent variable or the fitted values. Different methods such as the likelihood ratio test, score test, nonparametric test, and certain modeling approaches can be used for the diagnosis of heteroscedasticity in normal nonlinear regression models (Tsai, 1986; Simonoff and Tsai, 1991; Diblasi and Bowman, 1997; Lin and Wei, 2003). For non-normal models, the test of departure from the nominal dispersion, including over-dispersion and under-dispersion were also discussed (Dean, 1992; Wei et al., 1998).

If variance heterogeneity is ignored, the parameter estimates may not be influenced much, but it may result in severely misleading confidence and prediction intervals (Carroll and Ruppert, 2017). One way of taking into account variance heterogeneity is by explicitly modeling it by formulating a regression model for the variance. For example, Chung and Park (2007) investigated regression models in which the conditional variance of the error is an unscaled function of an integrated time series.

The generalized least squares method (Nelder and Wedderburn, 1972) is typically used for correlated residuals with heterogeneous variances using the IRLS (Eq. 13). Formally, we replace the assumption of the noise distribution $\varepsilon \sim N(0, \sigma^2 I_n)$ with the assumption of the noise distribution as $\varepsilon \sim N(0, \Sigma)$ where $\Sigma = V^{1/2} R V^{1/2}$. Here V is a diagonal matrix of potentially different (heterogeneous) variance terms and R is a correlation matrix. V has a constant diagonal and R can be modeled by a variance function (e.g., $a_i(\phi)$ in an exponential family in Eq. (12)) or an autoregressive-moving-average temporal structure. Additionally, Gaussian processes (Rasmussen, 2003) with the Bayesian nonparametric inference have been widely used for heteroscedastic nonlinear regression in which we have $f \sim N(\mu, K)$. The covariance matrix K has elements $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$, and function κ measures similarity between two observations $\mathbf{x}_i$ and $\mathbf{x}_j$. For example, the radial basis function kernel has $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\{-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \gamma\}$ with a hyperparameter γ .

Model building

Model building aims at finding more realistic ways to describe the stochastic behavior observed in data. We listed a few widely used models, which are classified by their shape and applied to real data analysis in various fields (Farebrother, 1983).

- Polynomial models. Linear and quadratic polynomial regression have been widely applied to solve real-world data problems for scientists and engineers. Quadratic or higher-order polynomials are curved with respect to the predictors, but they are linear in the parameters so that they can be fitted by using linear regression. However, they tend to overfit the responses with a large number of predictors. As the availability of nonlinear regression algorithms increased, the use of polynomials has sensibly decreased. Linear or quadratic polynomials are mainly used to approximate the observed response within a small range of each predictor for understanding the relationship with the response.
- Concave/Convex curves models. Concave/convex curves describe nonlinear relationships with asymptotes and without inflection points.
 1. Exponential model. The exponential equation describes an increasing or decreasing trend with constant relative rate. The most common parameterization is

$$f(x, \theta) = \theta_1 \exp(\theta_2 x),$$

where the slope of the tangent line through x is θ_2 . Therefore, the response increases by an amount that is proportional to x positively when $\theta_2 > 0$ (exponential growth), and negatively when $\theta_2 < 0$ (exponential decay). The exponential curve is used to describe the growth of a population in unconstrained environmental conditions. The exponential function is nonlinear and can be fitted by using `nls()` or `drm()` functions in R (Team et al., 2018).

2. Asymptotic model. The asymptotic regression model describes a limited growth, where y approaches a horizontal asymptote as $x \rightarrow \infty$. There are several different parameterizations known as the monomolecular growth model (Fekedulegn et al., 1999). The most widely used parameterization is

$$y = \theta_1 - (\theta_1 - \theta_2)\exp(-\theta_3 x) + \varepsilon,$$

where θ_1 is the maximum attainable y , θ_2 is y at $x = 0$, and θ_3 is proportional to the relative rate of y increase while x increases.

3. Negative exponential model. This model can be given by

$$y = \theta_1(1 - \exp(-\theta_2 x)),$$

which has a similar shape to the asymptotic regression, but $y = 0$ when $x = 0$, and is often used to model absorbed photo-synthetically active radiation (Kiniry et al., 1989).

4. Power curve model. The power curve is also known as the Freundlich equation or allometric equation (Thommes et al., 2015) and the most common parameterization is

$$y = \theta_1 x^{\theta_2} + \varepsilon.$$

This curve is equivalent to an exponential curve on the logarithm of x since $\theta_1 x^{\theta_2} = \theta_1 \exp(\log(x^{\theta_2})) = \theta_1 \exp(\theta_2 \log(x))$, and it does not have an asymptote for $x \rightarrow \infty$.

5. Logarithmic model. After x is log-transformed,

$$y = \theta_1 + \theta_2 \log(x)$$

is a linear model with $x > 0$. The parameter θ_2 dictates the shape, as in the exponential equation, Indeed, if $\theta_2 > 0$, the curve is convex up and y increases as x increases. If $\theta_2 < 0$, the curve is concave up and y decreases as x increases.

6. Rectangular hyperbola model. The Michaelis-Menten equation is a rectangular hyperbola, often parameterized as

$$y = \frac{\theta_1 x}{\theta_2 + x} + \varepsilon,$$

which is a convex-up curve and y increases as x increases up to a plateau level. The parameter θ_1 represents the asymptote as $x \rightarrow \infty$, while θ_2 is the x value giving a response equal to $\theta_1/2$. This is because $\theta_2 = x_{50}$ when $\theta_1/2 = \frac{\theta_1 x_{50}}{\theta_2 + x_{50}}$ where x_{50} is the median.

- Sigmoidal curve model. Sigmoidal curves are S-shaped and may be increasing, decreasing, or symmetric around the inflection point. We show a common parameterization.

1. Logistic model. The logistic curve is derived from the cumulative logistic distribution function; the curve is symmetric around the inflection point and it may be parameterized as

$$y = \theta_1 + \frac{\theta_2 - \theta_1}{1 + \exp(\theta_3(x - \theta_4))} + \varepsilon,$$

where θ_2 is the upper asymptote, θ_1 is the lower asymptote, and θ_4 is the x value producing a response half-way between θ_2 and θ_1 . θ_3 is the slope around the inflection point, and it can be positive or negative. Hence, y may increase or decrease as x increases.

2. Gompertz model. The Gompertz curve has the following parameterization

$$y = \theta_1 + (\theta_2 - \theta_1)\exp\{-\exp[\theta_3(x - \theta_4)]\} + \varepsilon,$$

where the parameters have the same meaning as those in the logistic model, but the difference is that this curve is not symmetric around the inflection point.

- Log-logistic models. The sigmoidal response curve is symmetric on the logarithm of x , which requires a log-logistic curve. For example, in biologic assays, the log-logistic curve is defined as follows:

$$y = \theta_1 + \frac{\theta_2 - \theta_1}{1 + \exp\{\theta_3[\log(x) - \log(\theta_4)]\}} + \varepsilon,$$

where the parameters have the same meaning as the logistic model.

1. Weibull-type 1 model. The Weibull distribution is a widely used lifetime distribution in reliability and life data analysis. The type-1 Weibull curve is an alternative for the Gompertz curve like the log-logistic curve is for the logistic curve. The model function is as follows:

$$f(x, \theta) = \theta_1 + (\theta_4 - \theta_1)\{1 - \exp\{-\exp[\theta_2(\log(x) - \log(\theta_3))]\}\}.$$

The parameters have the same meaning as in the sigmoidal curves.

2. Weibull-type 2 model. The type-2 Weibull curve is for the Gompertz curve what the log-logistic curve is for the logistic curve. The equation is as follows:

$$f(x, \theta) = \theta_1 + (\theta_4 - \theta_1) \exp\{-\exp[\theta_2(\log(x) - \log(\theta_3))]\},$$

again with the same parameter interpretation as in the sygmoidal curves.

Simulation data analysis

We compare the nonlinear least squares estimation with different initial values of the parameters in the nonlinear regression using a simulated dataset where the error term ε is not normally distributed. We show that the Bayesian approach is more appropriate to find reliable estimates as well as the confidence regions. The simulation setting is similar to the one shown in Fig. 1. We use the Michaelis-Menten model (Eq. 10) to simulate 100 data points with $\varepsilon \stackrel{i.i.d.}{\sim} \text{Gamma}(10, 0.25)$, $\theta_1 = 100$, and $\theta_2 = 0.05$. The covariates $\mathbf{X}$ are generated from the absolute values of $N(0, 20^2)$.

We implement the least squares method through the `nls()` function in R (Baty et al., 2015). Two sets of starting values for the parameters θ_1 and θ_2 are tried, (10, 3) and (50, 0.1), to validate if least squares estimates can converge to the true values given bad starting values.

Both the estimates and 95% confidence intervals for θ_1 and θ_2 using the least squares and Bayesian approaches are shown in Table 1. With good initial values, both the least squares and Bayesian method give good estimates and confidence regions for the parameters. However, when the starting values are bad, the estimates are far from the true values for both methods. The 95% confidence regions still contain the true values for Bayesian modeling, but the bands are too wide since the posterior samples tend to converge to local optimums.

Further reading

- Levenberg-Marquardt algorithm. The use of the Gauss-Newton algorithm with Levenberg-Marquardt modifications (Levenberg, 1944; Marquardt, 1963) is to speed up the convergence as well as to stabilize the computation for near-singular $\mathbf{J}^T \mathbf{J}$ in the least squares normal equations. The Levenberg-Marquardt method compromises the Gauss-Newton method and the steepest descent method.
- Multicollinearity. This occurs when some columns in the Jacobian matrix $\mathbf{J}$ are highly correlated and leads to an ill-conditioned matrix of normal equations. This implies that the model may be over-parameterized, so that a simpler model or a transformation of the predictors or parameters may be considered.
- Nonparametric and semiparametric regression. If the target of data analysis is to fit the response curve on the explanatory variables, a nonparametric regression could be a better alternative than a nonlinear regression model. However, the fitted curve may be less interpretable. Semiparametric regression which consists of both parametric and nonparametric components is an alternative to nonlinear regression modeling.
- Kernel estimator. A one-dimensional smoothing kernel is any smooth, symmetric function $K: \mathbb{R} \rightarrow \mathbb{R}$ such that $K(x) \geq 0$ and $\int K(x)dx = 1$, $\int xK(x)dx = 0$, $0 < \int x^2 K(x)dx < \infty$. Given a bandwidth $h > 0$, the Nadaraya-Watson kernel regression estimator (Nadaraya, 1964; Watson, 1964) is defined as

$$\hat{f}_h(x) = \frac{\sum_{i=1}^n K\left(\frac{\|x-x_i\|}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{\|x-x_i\|}{h}\right)} = \sum_{i=1}^n w_i(x) y_i,$$

Table 1 The estimates of θ_1 and θ_2 with the 95% confidence intervals using least squares and Bayesian approach with two different settings of initial values.

Model	Initials	$\theta_1 = 100$	$\theta_2 = 0.05$
NLS	(10, 3)	81.78 (73.94, 89.61)	-0.22 (-0.23, -0.21)
	(50, 0.1)	102.53 (102.38, 102.69)	0.05 (0.05, 0.05)
Bayesian modeling	(10, 3)	84.74 (25.48, 102.67)	-0.25 (-1.18, 0.05)
	(50, 0.1)	102.53 (102.37, 102.69)	0.05 (0.05, 0.05)

where $w_i(x) = \frac{K(\|x - X_i\|/h)}{\sum_{j=1}^n K(\|x - X_j\|/h)}$. Therefore, $\hat{f}_h(x)$ is a local weighted average of the y_i 's.

- Initial values for parameter estimation. For a linear model the Gauss-Newton method finds the minimum in one iteration from any initial parameter estimates. If the nonlinear mean function is nearly linear, the convergence of the Gauss-Newton method is fast and does not depend heavily on the initial parameter estimates. However, as the magnitude of model nonlinearity is large, convergence can be slow or even may not occur, and the resulting parameter estimates may not be reliable. In that case, good initial values are important.
- Differential geometric view. The measures of curvature described above are related to the parameterization. When a normal error assumption is assumed, a Euclidean metric is imposed. Changing the error distribution results in a different metric and geometric structures of the distributions. The study of probability and information using differential geometry is called information geometry. We refer the reader to Amari (2016) for theory and applications of information geometry.

Acknowledgment

Hsin-Hsiung Huang is supported by the National Science Foundation grants (DMS-1924792).

References

- Amari, S.-i., 2016. Information Geometry and Its Applications, vol. 194. Springer.
- Bates, D., Watts, D., 1988. Nonlinear Regression Analysis and Its Applications, Wiley Series in Probability and Mathematical Statistics. Wiley, New York [u.a.].
- Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.-P., Delignette-Muller, M.-L., 2015. A toolbox for nonlinear regression in R: the package nlstools. *J. Stat. Software* 66, 1–21.
- Beverton, R.J., Holt, S.J., 2012. On the Dynamics of Exploited Fish Populations, vol. 11. Springer Science & Business Media.
- Brazzale, A.R., 2005. Hoag: An R Package Bundle for Higher Order Likelihood Inference. *Rnews* 5/1 May 2005, pp. 20–27. ISSN 609-3631.
- Carroll, R.J., Ruppert, D., 2017. Transformation and Weighting in Regression. Chapman and Hall/CRC.
- Chung, H., Park, J.Y., 2007. Nonstationary nonlinear heteroskedasticity in regression. *J. Econom.* 137, 230–259.
- Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* 83, 596–610.
- Dean, C.B., 1992. Testing for overdispersion in Poisson and binomial regression models. *J. Am. Stat. Assoc.* 87, 451–457.
- Dibiasi, A., Bowman, A., 1997. Testing for constant variance in a linear model. *Stat. Probab. Lett.* 33, 95–103.
- Farebrother, R.W., 1986. Nonlinear regression modelling: a unified practical approach: David A. Ratkowsky, (Dekker, New York, 1983). *Int. J. Forecast.* 2 (1), 125.
- Fekedulegn, D., Siurta, M.M., Colbert, J., 1999. Parameter estimation of nonlinear growth models in forestry. *Silva Fenn.* 33.
- Gill, P.E., Murray, W., Saunders, M.A., Wright, M.H., 1986. User's Guide for NPSOL (Version 4.0): A Fortran Package for Nonlinear Programming. Technical Report. Stanford Univ CA Systems Optimization Lab.
- Green, P.J., 1984. Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives (with discussion). *J. R. Stat. Soc. Series B Methodol.* 46, 149–192.
- Hamilton, D.C., Watts, D.G., 1985. A quadratic design criterion for precise estimation in nonlinear regression models. *Technometrics* 27 (3), 241–250. <https://doi.org/10.1080/00401706.1985.10488048>.
- Huang, H., Hsiao, C., Huang, S., 2010. Nonlinear Regression Analysis. *International Encyclopedia of Education*, 3rd. Elsevier, pp. 339–346. <https://doi.org/10.1016/B978-0-08-044894-7.01352-X>.
- Kiniry, J., Jones, C., O'toole, J., Blanchet, R., Cabelguenne, M., Spanel, D., 1989. Radiation-use efficiency in biomass accumulation prior to grain-filling for five grain-crop species. *Field Crop. Res.* 20 (1), 51–64.
- Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* 2 (2), 164–168.
- Lin, J.-G., Wei, B.-C., 2003. Testing for heteroscedasticity in nonlinear regression models. *Commun. Stat. Theor. Methods* 32 (1), 171–192.
- Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.* 11 (2), 431–441. <https://doi.org/10.1137/0111030>.
- Nadaraya, E.A., 1964. On estimating regression. *Theor. Probab. Appl.* 9, 141–142.
- Nelder, J.A., Wedderburn, R.W.M., 1972. Generalized linear models. *J. R. Stat. Soc. Series A General* 135, 370–384.
- Rasmussen, C.E., 2003. Gaussian processes in machine learning. In: *Summer School on Machine Learning*. Springer, pp. 63–71.
- Ruppert, D., Wand, M.P., 1994. Multivariate locally weighted least squares regression. *Ann. Stat.* 22, 1346–1370.
- Saber, G.A.F., Wild, C.J., 1989. Nonlinear Regression. Technical Report.
- Sheiner, L.B., Beal, S.L., 1980. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. *J. Pharmacokinet. Biopharm.* 8 (6), 553–571.
- Simonoff, J.S., Tsai, C.-L., 1991. Improved Tests for Nonconstant Variance in Regression Based on the Modified Profile Likelihood. Leonard N. Stern School of Business, New York University.
- Skovgaard, I.M., 1996. An explicit large-deviation approximation to one-parameter tests. *Bernoulli* 2 (2), 145–165. <https://doi.org/10.3150/bj/1193839221>.
- Team, R.C., et al., 2018. R: A Language and Environment for Statistical Computing; 2018.
- Thommes, M., Kaneko, K., Neimark, A.V., Olivier, J.P., Rodriguez-Reinoso, F., Rouquerol, J., Sing, K.S., 2015. Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report). *Pure Appl. Chem.* 87 (9–10), 1051–1069.
- Tsai, C.-L., 1986. Score test for the first-order autoregressive model with heteroscedasticity. *Biometrika* 73 (2), 455–460.
- Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics With S, fourth ed. Springer, New York, ISBN 0-387-95457-0.
- Watson, G.S., 1964. Smooth regression analysis. *Sankhyā Ser. 26*, 359–372.
- Wei, B.-C., Shi, J.-Q., Fung, W.-K., Hu, Y.-Q., 1998. Testing for varying dispersion in exponential family nonlinear models. *Ann. Inst. Stat. Math.* 50 (2), 277–294.

Hierarchical linear models

D. Rindskopf, CUNY Graduate Center, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Example 1: two levels, one predictor at each level, linear model	568
Example 2: repeated measures and longitudinal data	570
Generalized linear and nonlinear models	571
Shrinkage estimates	571
Meta-analysis	572
Bayesian estimation	572
Three underutilized applications of hierarchical models	572
Complex repeated measures designs	572
Modeling panel results (repeated surveys)	573
Parameter shrinkage (and reducing effects of multicollinearity) by modeling the structure of regression parameters	573
Recommended reading	573
Summary and conclusion	573
References	574
Further reading	574
Relevant websites	574

In many educational contexts, data occur in a nested structure. For example, we may measure the achievement level of a child in a particular subject matter. That child is located (nested) within a particular class of students, and we might predict that the child's performance would have been different had she been in a different class of students. Further, the child's class may be one of many within a particular school, and that school may be one of many within a school system, which is located within a larger educational or political unit (e.g., county, state, and region). Each level of the structure may have its own characteristics that influence the achievement of the child. Hierarchical linear models (and more general models, to be discussed later) are designed to properly model influences on outcome variables at all levels of such a hierarchy, including the proper handling of data dependencies created by the nested structure. (Note that not all nesting produces a multilevel model; only when both levels are random effects is this true. For example, people nested within treatments, as in a typical ANOVA, do not result in the type of model discussed here because levels of treatment are typically fixed in number and are not a random sample from a population of treatments.)

In this article, we describe a variety of models that are designed to accommodate a nested data structure. Some models are linear, but others are nonlinear; some involve nesting of people within higher-level units (such as classes), while other models are applicable to repeated measures that are nested within individuals (including longitudinal studies). We begin with a very simple example, and then show how the principles involved can be used to expand the scope and type of models that may be used. These principles are all expansions or generalizations of those used in ordinary least squares (OLS) regression; often, the models can be viewed as a series of regressions for different levels of the data structure. Those unfamiliar with regression should read the article on univariate linear regression in this volume before reading the rest of this one.

Although we will not discuss it in detail, these methods also resolve the problem known as aggregation bias. It has been well known that correlations among variables at a higher level (e.g., measures on states) do not equal (and may be opposite in sign to) correlations among variables at the individual level. (A related problem that arises in categorical data analysis is known as Simpson's paradox.) Multilevel models are constructed so that inferences can be made at the right level, whichever level that happens to be for the problem at hand.

We discuss two examples of multilevel models. The first example is for a situation in which students are nested within colleges, and we wish to predict academic performance from one student-level and one college-level variable. The second example considers longitudinal data, where a series of observations is made on each person over a period of time. Here, the nesting is of a different sort: Observations (over time) are nested within individuals.

Example 1: two levels, one predictor at each level, linear model

For a first example, assume that the administration in 10 colleges in a particular country wants to see how well scores on an academic-achievement test given to all students in the last year of secondary school predict performance of students in their first year of college. If we consider each college separately, we may imagine that within each college, researchers fit a regression model to the data; the equation for school j would then be:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$$

In this equation, Y represents the grade of a student after the first year of college, X represents the score of the student on the secondary school-achievement test, i is used to refer to a specific student, and j to refer to a specific school. The coefficients are as in a linear regression model, except that each school has its own intercept (β_{0j}) and slope (β_{1j}).

As we shall be examining the intercept and slope of each school, and modeling them, it is important to be sure that these are quantities of interest. For the slope, there is ordinarily no issue (except perhaps multiplication or division by a constant in order to make a one-unit change more easily interpreted). In order to make the intercept meaningful, we will often have to transform X . One common strategy is centering, which means to create a new variable by subtracting the mean of X from each observed value of X . (For this article, we assume that the mean is computed for the whole data set, not for each college individually; this is called grand mean centering.) In any regression, the intercept represents the predicted value of Y when all predictors have a value of 0; therefore centering X will cause the intercept in each college's regression equation to equal the predicted value of Y for a student who is average on X (based on students from all included colleges). Using the exact sample mean is not necessary; sometimes, it is more important to use a round number for ease of interpretation. For example, if the mean is 497, it might be easier to subtract 500. Or, if a score of 600 is typically required for admissions, one might subtract 600 from all X values instead.

In addition to subtracting some value from X , one might also want to control the scaling further. In the United States, for example, where Scholastic Aptitude Test (SAT) scores have a mean near 500 and a standard deviation near 100, one might divide by 100 so that a one-unit change in the transformed variable is more meaningful than a one-unit change on the original scale. This makes the slope more easily interpreted (and sometimes improves numerical stability.) From this point on, we will assume that X has been suitably scaled.

The term r_{ij} is a residual, indicating that the students' scores on the secondary school test will typically not predict their grades in college perfectly. The variance of the residuals in each college can differ among schools, but often modeling is begun by assuming that it is the same in each; we will denote the within-school residual variance by σ^2 , in which we do not use a subscript for college (j).

The intercepts and slopes will vary from college to college; some of this variation is just sampling variability due to using a sample of students within each college, but the other variability is due to true differences among colleges. The true variability in intercepts will be denoted by τ_{00} , and the true variability in slopes will be denoted by τ_{11} . (There is also a covariance between the intercepts and slopes, denoted τ_{01} , but this is seldom of much interest.)

If τ_{00} and τ_{11} are small, then the intercepts and slopes do not vary much among colleges, and one prediction equation will work equally well in each. The colleges will gain by knowing this, because now their prediction equation is known with greater precision due to the contribution of information from other colleges. Of course, unless the colleges are very similar, it is unlikely that both τ_{00} and τ_{11} will be small, which leads to a consideration of two issues: First, can we explain, based on characteristics of the colleges, why some have larger and others smaller intercepts and/or slopes? Second, even though the colleges differ, can we still use information from other colleges to help improve prediction in each college?

We may believe that characteristics of the colleges will have some relationship to either the intercept or the slope of their prediction equations. For example, it may be that schools differ in the difficulty of their courses, or in their grading standards. These may affect the intercepts or slopes. One might use the college's average value on X as a proxy for selectivity or difficulty; this could be a predictor for intercepts and/or slopes. Another possible measure would be the selection ratio, that is, the proportion of applicants who are admitted to the college.

We will now write the model for variability among colleges out more formally. The equations, which will first be presented without predictors, are:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

The first equation says that the intercept in a particular college (j) is equal to some overall average plus a deviation of that college's intercept from the overall average. The second equation expresses the same idea for the slopes from each college. The variances of the residuals (u_{0j} in the first equation and u_{1j} in the second) are the values τ_{00} and τ_{11} mentioned previously. In this instance, they are called unconditional residual variances, because they are not conditioned (dependent) on any college-level predictors. When we analyze the data we will obtain estimates of both the mean intercept and slope (called fixed effects in most software) and the variances of the residuals (called random effects in most software, though it is the u_{ij} s that are random variables).

Now suppose that the residual variances are large enough that we believe it worthwhile to try to account for some of their variation by using a college-level variable such as proportion of applicants admitted. In realistic applications, there are often many potential variables, and variables that help predict differences in intercepts may not be the same variables that help predict differences in slopes. In this simple introduction, however, we will use only one such variable, which we will call W ; individual values of W are denoted W_j , with the subscript j indicating that the value of W will vary across colleges.

The extended model at the college level will now consist of the following equations:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}$$

The variances τ_{00} and τ_{11} of the residual terms u_{0j} and u_{1j} now represent conditional variances; that is, the variances remaining after accounting for variation in intercepts and slopes using information provided by the predictor W . These variances should now be smaller than in the unconditional model if W proves to be a good predictor, and measures analogous to R -square in the usual regression context may be calculated.

Multilevel models are very much like ordinary regression models in many ways: the outcome variable may be continuous or categorical. If continuous, it may be normally distributed or have some other distribution (e.g., exponential for survival models). If the dependent variable is categorical, models similar to logistic regression (for binary outcomes) or its extensions to polytomous or ordinal outcomes are possible. The right-hand side of any of the equations may require polynomial coding (e.g., the square and cube) for continuous predictors, dummy variables (or other coding methods) for categorical predictors, and product terms for interactions. The model may be intrinsically nonlinear, so that more complex models must be fit. We deal briefly with some of these issues later, and other articles have more details.

Example 2: repeated measures and longitudinal data

In many educational applications, people are measured more than one time. In some cases, these measurements extend over a long period of time; measurements may be daily, weekly, monthly, or yearly, for example. In other cases, the time sequence is not of great importance; for example, most models for test scores do not take into account that some items are administered before others. In either case, we can use the concept of nesting to reconceptualize the model for this type of data: we view observations as being nested within an individual. (In some instances, it might seem more natural to view time of observation as crossed with individuals, but this viewpoint is not usually considered in the multilevel literature, because doing so would reduce the utility of the model in many common situations, as we hope will become apparent shortly.)

As an example, suppose that a researcher wants to examine the growth of children's ability to sight-read words. Each month during the school year, she has every child in a class attempt to read a list of 20 words, and counts the number of words that the child can correctly read. We will treat the number of words read as if it were a continuous outcome variable to simplify the discussion in this section, and will later deal with the complexities that this type of outcome necessitates. The researcher wants to model the number of words read by each child as a function of the number of months into the school year. The time variable may be coded in a number of ways; one might be to start the variable Months at 0 for the first measurement period, and increase by one for each month. Another approach would be to use 0 for the last period of measurement, and number preceding periods as -1 , -2 , etc. Each possible choice determines the meaning of the intercept in the model. Again, for simplicity we will use a linear model, but realize that a straight line is unlikely to be a perfect description of such growth; we write the equation for each student as follows:

$$Y_{ij} = \pi_{0j} + \pi_{1j}M_{ij} + e_{ij}$$

where the subscript j is used to represent the individual student, i is used to represent a measurement occasion, Y_{ij} is the number of words read correctly by student j at occasion i , M_{ij} is the number of months (on whatever scale is chosen; here we will assume 0 is used for the first time of measurement) for measurement of student j at occasion i , and e_{ij} is the residual. The intercept π_{0j} and slope π_{1j} are different for each student, and again we will allow these to be predicted by variables at the higher level, which in this case is the individual student. For example, boys and girls may start at different levels, and they may progress at different rates. We may represent this by the following level-2 equations, in which F is equal to one for females and zero for males:

$$\pi_{0j} = \beta_{00} + \beta_{01}F_j + r_{0j}$$

$$\pi_{1j} = \beta_{10} + \beta_{11}F_j + r_{1j}$$

Of course the structure of all of these equations is exactly the same as the case in the first example where students were nested within colleges; only the context has changed. In these equations, the intercepts β_{00} and β_{10} represent the beginning number of words read (π_0) and the rate of change in words per month (π_1) for an average male (i.e., $F = 0$) in the study. The slopes in these equations, β_{01} and β_{11} , represent the difference in beginning number of words read and rate of change for females as compared to males. Again, the choice of coding makes a difference; gender could be coded in other ways, which would change the interpretation of each of the coefficients in these equations. The residual terms represent the amount by which the true values of π_{0j} and π_{1j} differ from that predicted for the average male or female.

This method of representing the model has several advantages over other methods. Most importantly, from a conceptual point of view, this is a model for each individual's growth pattern and for deviations of individuals from the average pattern. The focus is on the individual, rather than a group. As a consequence of this focus, in principle (and at present, in practice also) missing data are not a problem, as long as the cause for missingness is not related to the outcome variable. Each child could have a different pattern of months during which he or she was measured, and in fact the times of measurement do not even have to follow a common pattern. A child could be measured at months 1.25, 3.75, and so on, rather than exactly at the beginning of each month—the model remains the same. (Of course, in practice one must measure each individual a reasonable number of times, and with those times spaced in a reasonable way, in order to have accurate estimates of that child's growth pattern, but this is another matter.)

Generalized linear and nonlinear models

Two general categories of complications frequently arise with multilevel models. As mentioned previously, not all dependent variables are continuous, and not all continuous variables have a normal distribution. In addition, not all relationships are linear, so nonlinear models must sometimes be considered. This section briefly describes some approaches to each issue.

In the second extended example described, the outcome was the number of words read correctly from a list of 20 words presented to the student. Equivalently, one might calculate the percentage or proportion of the list that was read correctly. If proportions were not extreme (e.g., if most were between 0.2 and 0.8), then using these in a model with a normal distribution might be a good approximation to the correct analysis. A more accurate approach would be to consider each word as a trial, and the number correct as coming from a binomial distribution with an underlying probability of success for each trial that is constant across words. (An even more accurate model would allow words to have varying difficulties, and possibly to differ in other ways also, as in item-response theory. Often the simplifying assumption of equal difficulty is a good enough approximation to work well in practice.)

The binomial model is closely related to a more widely known model, logistic regression. For these models, the outcome that is modeled is usually the logarithm of the odds of a correct response, commonly called the logit of the probability of a correct response. The right-hand side of the level-1 equation for individual responding still has a linear form in these models; the left-hand side is the transformed value $\ln(\omega/(1-\omega))$, also denoted $\text{logit}(\omega)$ in some sources, where ω is the probability of a student responding correctly. The model for each individual is written in the following form:

$$\ln\left(\frac{\omega_{ij}}{1 - \omega_{ij}}\right) = \pi_{0j} + \pi_{1j}M_{ij}$$

where ω_{ij} is the expected (not observed) probability of person j answering correctly during month i . The residual is not explicitly listed in this equation; instead, we state that $Y_{ij} \sim \text{Binomial}(20, \omega_{ij})$; that is, the observed number of words correctly read has a binomial distribution based on 20 trials, with probability ω_{ij} of being correct on each trial.

This is an example of a generalized linear model; these are extensions of linear models that include a linear model, a transform of the dependent variable, and a distributional form that is more general than a normal distribution. In this case, the logit is the transformation and the binomial is the distributional form.

Another example requiring the use of generalized linear models is when the outcome is a count, but the count is not the number of successes in a fixed number of trials, or it is a generally small count out of a very large number of fixed trials (in which case we are using an approximation to the binomial distribution). A researcher might count the number of times a student shows an aggressive behavior toward classmates, or the number of times the student raises his or her hand for the attention of the teacher rather than responding without being called on. If the period of observation is the same for each student and time of observation, then generally the Poisson distribution provides the simplest probability model for counts. A slight extension is required when the time of observation varies, so that the model is for the rate of behavior per period of time (number of hand raises per hour, e.g., when the time of observation may vary between 5 and 20 minutes). If F represents the expected number of acts and Z represents the amount of time for an observational period, so that F/Z is a rate per unit time, the level-1 equation can be written in the following form:

$$\ln(F_{ij} / Z_{ij}) = \pi_{0j} + \pi_{1j}M_{ij}$$

As with the binomial model, there is no explicit error term in the equation because the count variable in the equation is an expected rather than observed frequency, and in this case it has a Poisson distribution.

A related model with the same form is the discrete-time survival model, in which we model the influences on the occurrence of an event. In education, we might wish to determine factors affecting whether, and if so when, a student in college will drop out. Modeling whether or not the student drops out involves an ordinary logistic regression (or possibly with students nested within colleges if we observe many colleges). But to model the length of time to dropout is slightly more complicated. We observe each student over a series of semesters, and note whether he or she has dropped out or is still enrolled for that semester. The main aspects of this situation that are different from others with a binary outcome in each time period are that (1) drop out can only occur during one period, and (2) the sequence of observations can be stopped (censored) due to graduation so that the event we are looking for (dropout) may not be observed for that individual. It may not be obvious, but information from students who never drop out is nonetheless useful. Survival models developed in somewhat parallel ways in three areas: medicine (for obvious reasons), engineering (where they are generally known as failure time models), and sociology (where they are known as event history models).

Shrinkage estimates

In the original example, predicting grades of students in college based on their grades in secondary school, each of the 10 researchers could do a study and estimate a slope and intercept for their own college. But the 10 colleges, while not identical, should be similar enough so that each researcher can use the information from all 10 colleges to estimate the slope and intercept in their college. Using a phrase common in this type of situation, they can each borrow strength from related data sets. This generally will shrink their

estimate of the intercept and slope toward the average intercept and slope. The amount of shrinkage depends on how well their own parameters are estimated; if they have little data, shrinkage will be large, but if they have a great deal of data on their own school, shrinkage will be small.

One can also adjust shrinkage for characteristics of the college. Perhaps large colleges have different slopes or intercepts than small colleges; if so the size of the college (or a function of it, such as the logarithm) can be used as a predictor. (Centering or other scaling of these predictors is also recommended for more useful interpretation of parameters.)

Note that the more similar the colleges are, the more is gained from shrinkage estimates. The extreme instance, with identical colleges, would make the random effect variances zero, and the situation would revert to one large sample, with more accurate estimates due to the larger sample size applying to each. Even if the colleges are not identical, if the differences among them are explained by their characteristics, then the residual variances are small and result in more accurate estimates of the individual colleges' intercepts and slopes.

Meta-analysis

A situation in which it is not obvious that multilevel models can be applied is meta-analysis. In meta-analysis, a large number of studies about a topic are summarized by a measure of effect size, and by other characteristics of the study (age of subjects, whether a randomized design was used, etc.). The objective is to determine what the average effect size is, whether it varies (beyond sampling variability) from study to study, and if so whether study characteristics can be used to predict that variability. If we view subjects as nested within studies, it is easy to see that this falls into the realm of multilevel models, but with the individual level data missing (i.e., we only see summary statistics for the study as a whole). Most programs for multilevel models (and some additional specialized programs) can perform such analyses.

Bayesian estimation

Multilevel models have two types of parameters: Those generally called fixed effects, represented in this article by γ_{ij} , and the variances of random effects, represented by τ_{ij} and σ^2 . Broadly speaking there are two kinds of estimation methods for these parameters in hierarchical models: Frequentist and Bayesian. Frequentist methods (such as maximum likelihood and least squares) treat parameters as fixed unknowns; Bayesians treat parameters as random variables that have distributions (or at least our beliefs about them have distributions). Frequentist results are consequently estimates of the parameters and standard errors, which tell how accurate the estimates are. Bayesians summarize their beliefs about the parameters in terms of probability distributions.

Bayesian methods have several advantages. They take into account all uncertainty; in particular the estimation of regression parameters takes into account uncertainty about the variance parameters. This is important when there are a small number of higher-level units, because their variance is poorly estimated. Bayesian methods also lead naturally to shrinkage estimators, and handle missing data in a more natural way than frequentist methods. The interpretation of interval estimates is more natural for Bayesians: There is a 95% chance that the parameter value is in the 95% credible interval (analogous to a frequentist confidence interval). Other probabilities can be calculated that cannot be using frequentist methods; for example the probability that an effect is greater than zero (or any other number). Software for general Bayesian modeling is somewhat involved to use; popular packages include BUGS/WinBUGS/OpenBUGS, jags, and Stan. Common models can be dealt with more simply, for example with the brms or MCMCglmm packages in R. For further information on Bayesian inference, see the entry "Bayesian Statistical Analysis" and [Kruschke and Vanpaemel \(2015\)](#).

Three underutilized applications of hierarchical models

The most common applications of hierarchical models are (i) people nested in groups, (ii) longitudinal data (observations over time nested in people), and (iii) meta-analysis (people nested in studies). There are other applications that are less-often utilized, but quite useful when they are appropriate.

Complex repeated measures designs

Suppose you are examining papers written by a group of people, and rating each on how many errors of a particular type were made out of the number of opportunities to make them. Different people have different number of opportunities, so the structure of the responses is different for each person. These subjects, in turn, are measured on a variable thought to be predictive of making that type of errors (e.g., age or gender). This would be difficult or impossible to analyze from a typical repeated measures perspective, but is straightforward using hierarchical models. Note this is different from the usual application that emphasizes growth modeling (that is, in this case we are not modeling behavior as a function of time/age.) For more information on such designs, see [Schonfeld and Rindskopf \(2007\)](#).

More generally, consider any situation in which there is a data set for each individual, and a model for each person's behavior. Then one can consider what the average of each parameter is across individuals, how much variation there is across individuals in

the parameters, and whether that variation can be explained as a function of person-level characteristics. Single case designs common in psychology can be analyzed in this manner; in those studies time is a nuisance variable rather than a variable of substantive interest (see e.g., [Rindskopf and Ferron, 2014](#)).

Modeling panel results (repeated surveys)

If a survey or a testing program is repeated at regular intervals, a model may be desired for each administration. These could be fit separately to each set of data for a time point, in which case the estimates of similar quantities (parameters) are not strengthened by what should be similar estimates in the other years. On the other hand, complete pooling of data, making the coefficients the same in each time point, is probably not correct either. A hierarchical model allows each coefficient to be estimated separately, but with support from data at other time points. This could be as simple as making each parameter in a model a random effect, or modeling a linear or quadratic time change for parameters.

For example, suppose a survey were administered every year, and we wish to fit the same regression model every year. Our model for person i in year j is:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + r_{ij}$$

At one extreme, if we believed that the coefficients were the same in every year, we could pool data from all years and fit one equation. At the other extreme, if we believed coefficients could change wildly from year to year, we would fit each year's data separately. But if we believed that the coefficients were similar, though not identical, we would fit level 2 equations, here represented as having a linear trend over time:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}t_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}t_j + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}t_j + u_{2j}$$

where t represents time, perhaps scaled so that 0 is the current time. The time effect may be modeled as quadratic if it is believed that the parameters do not change in a linear fashion over time. Additional characteristics of years (such as whether a pandemic was taking place) could also be entered. This use of HLM is similar to having people nested within groups, except people are nested within time-structured groups, similar to some repeated measures data.

Parameter shrinkage (and reducing effects of multicollinearity) by modeling the structure of regression parameters

In regression a frequent problem is collinearity, the high correlation among predictors. Many possible solutions to this problem have been suggested, some of which include shrinking the (standardized) coefficients toward zero. One such technique is known as ridge regression, and another is the LASSO. Such methods can be put in the multilevel framework. This can be more general than similar methods such as ridge regression, because all coefficients are not necessarily shrunk toward the same value (such as zero or the mean).

The meaning of "hierarchical" is somewhat different in these models. People (or observations) are not nested; the hierarchy is one of parameters. The parameters in which we are interested (here the regression coefficients) are modeled at the next level of the hierarchy of parameters, which is of smaller dimension than the original set of parameters. [Lindley and Smith \(1972\)](#) were the first to suggest a Bayesian version of ridge regression and related models (in addition to the more common version of multilevel models we have discussed).

Recommended reading

Several excellent textbooks are available to learn more about multilevel models. [Raudenbush and Bryk \(2002\)](#), whose notation is similar to that used here, and [Goldstein \(2011\)](#) cover a wide range of material on these models. Each of these also is associated with software to fit these models, HLM ([Raudenbush et al., 2019](#)) for the former and MLWin ([Rasbash et al., 2020](#)) for the latter. Other introductory books include [Hox et al. \(2018\)](#) and [Snijders and Bosker \(2012\)](#). [Gelman and Hill \(2007\)](#) have a unique viewpoint, and many interesting approaches not found elsewhere. Other references in the reading list include more advanced works, as well as specialized references on various aspects of multilevel models.

Summary and conclusion

When units are nested within higher level units, multilevel modeling should be used for three major reasons: (1) it is the correct analysis (which ordinarily would be a sufficient reason), (2) viewing the model in this way leads to insights that are otherwise

elusive about relationships, and (3) we avoid certain perplexing problems, such as aggregation bias. Multilevel models can be used not only in obvious situations such as when we have children in classes (and or schools), but also where subjects are nested within studies (meta-analysis), or where observations are nested within people (repeated measures and longitudinal studies). Using these models enables us to estimate effects within each level of a study, as well as between levels. Specialized software is readily available, and several major statistical packages now include the capability to fit multilevel models.

References

- Gelman, A., Hill, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.
- Goldstein, H., 2011. *Multilevel Statistical Models*, fourth ed. Wiley, Chichester.
- Hox, J.J.C.M., Moerbeek, M., van de Schoot, R., 2018. *Multilevel Analysis: Techniques and Applications*. Routledge.
- Kruschke, J.K., Vanpaemel, W., 2015. Bayesian estimation in hierarchical models. In: Busemeyer, J.R., Wang, Z., Townsend, J.T., Eidels, A. (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology*. Oxford University Press, Oxford, UK, pp. 279–299.
- Lindley, D., Smith, A., 1972. Bayes estimates for the linear model. *J. Roy. Stat. Soc. B* 34 (1), 1–41. Retrieved March 28, 2021, from <http://www.jstor.org/stable/2985048>.
- Rasbash, J., Steele, F., Browne, W.J., Goldstein, H., 2020. *A User's Guide to MLwiN*, v3.05. Centre for Multilevel Modelling, University of Bristol, Bristol.
- Raudenbush, S.W., Bryk, A.S., 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*, second ed. Sage, Thousand Oaks, CA.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R., 2019. *HLM 8 for Windows* [Computer Software]. Scientific Software International, Inc, Skokie, IL.
- Rindskopf, D.M., Ferron, J.M., 2014. Using multilevel models to analyze single-case design data. In: Kratochwill, T.R., Levin, J.R. (Eds.), *Single-Case Intervention Research: Methodological and Statistical Advances*. APA, pp. 221–246. <https://doi.org/10.1037/14376-008>.
- Schonfeld, I.S., Rindskopf, D., 2007. Hierarchical linear modeling in organizational research: longitudinal data outside the context of growth modeling. *Organ. Res. Methods* 10 (3), 417–429. <https://doi.org/10.1177/1094428107300229>.
- Snijders, T.A.B., Bosker, R.J., 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, second ed. Sage, London.

Further reading

- Braun, H., 1989. Empirical Bayes methods: a tool for exploratory analysis. In: Bock, R.D. (Ed.), *Multilevel Analysis of Educational Data*. Academic Press, New York, NY, pp. 19–55.
- Davidian, M., Giltinan, D.M., 1995. *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall, Boca Raton, FL.
- Efron, B., Morris, C., 1977. Stein's paradox in statistics. *Sci. Am.* 236, 119–127.
- Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2006. *SAS for Mixed Models*, second ed. SAS Institute, Cary, NC.
- Longford, N.T., 1993. *Random Coefficient Models*. Clarendon Press, Oxford.
- Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS—a Bayesian modeling framework: concepts, structure, and extensibility. *Stat. Comput.* 10, 325–337.
- McCulloch, C.E., Searle, S.R., 2001. *Generalized, Linear, and Mixed Models*. Wiley-Interscience, New York.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed-Effects Models in S and S-PLUS*. Springer, New York.
- Singer, J.D., Willett, J.B., 2003. *Applied Longitudinal Data Analysis*. Oxford University Press, Oxford, UK.
- Singer, J.D., 1998. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *J. Educ. Behav. Stat.* 24, 323–355.
- Skrondal, A., Rabe-Hesketh, S., 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. Chapman and Hall/CRC Press, Boca Raton, FL.
- Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D., 2003. *WinBUGS User Manual*. MRC Biostatistics Unit, Cambridge (Version 1.4). <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>.
- Spiegelhalter, D.J., Abrams, K.R., Myles, J.P., 2004. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, Chichester.

Relevant websites

- <http://www.cmm.bristol.ac.uk>, Center for Multilevel Modeling.
- <http://www.ssicentral.com>, Scientific Software International, (Contains many good references to books, articles, and websites).
- <http://stats.idre.ucla.edu>, UCLA Statistical Computing, (Contains useful resources for multilevel models and other advanced statistical procedures).
- <http://www.stats.ox.ac.uk/~snijders/multilevel.htm>, (Tom Snijders Website; contains much useful information, including data sets).

Categorical data analysis

Dandan Chen and Carolyn J. Anderson, Department of Educational Psychology, University of Illinois, Urbana-Champaign, Champaign, IL, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	575
Notations	575
Sampling processes	575
Sampling distributions	576
Types of probability	577
Estimation of the unknown probability	578
Estimation method: maximum likelihood estimation	578
Hypothesis testing	578
Measures of association	580
Odds ratio	580
Pearson correlation coefficient	580
Tests of no association	581
Chi-square test statistics	581
Independence	581
Homogeneous distributions	581
Test of linear association	581
Concluding remarks	582
References	582
Relevant websites	582

Introduction

Valid inference requires proper methods to handle discrete data. Categorical data consist of counts or discretely measured variables. Examples of categorical variables include race or ethnicity, responses to a multiple-choice question, status of an online purchase, students' letter grades, gender, responses to survey items, graduation status, and whether an individual receives a vaccination. The methods described in this chapter are for analyses where a categorical variable is the response, outcome, or dependent variable; predictor or explanatory variables are either categorical or continuous. The levels of categorical variables can be either nominal (e.g., type of high school attended) or ordinal (e.g., letter grade).

Whether a variable is nominal or ordinal depends on context. For example, whether you received a vaccine seems to be inherently nominal. Either you have or have not received a vaccine. However, if the question "Have you received a COVID-19 vaccine?" is an item on a scale developed to measure a person's perceived susceptibility or the consequences of getting sick, the "yes" and "no" categories are ordered concerning what the scale is measuring.

In this chapter, we use examples of two categorical variables. Fundamentals here include contingency tables, sampling processes, sampling distributions, estimation of probabilities with hypothesis testing, measures of association, and testing for relationships using statistics based on different distributions.

Notations

Throughout this chapter, notations for categorical variables are in Roman letters, and parameters of models and distributions are represented by lower-case Greek letters. For instance, the data can be presented as a two-way contingency table with two categorical variables. In this situation, X and Y denote categorical variables; a cross-classification of these two variables is shown in a contingency table, where n_{ij} is a count or frequency within the cell (i, j) (i.e., row i and column j), n_{i+} and n_{+j} are the row and marginal column totals, and n is the total sample size. When there are more than two rows or columns, the numbers of rows and columns are denoted by I and J , respectively. μ represents the mean of a variable (X or Y), π denotes a probability, p is a sample proportion. For example, μ_{ij} , π_{ij} and $p_{ij} = n_{ij}/n$ represent the mean, probability, and sample proportion for cell (i, j) , respectively. df is the short form for degrees of freedom.

Sampling processes

The data about fatal shootings from on-duty police officers in the United States from January 2015 to December 2020, compiled by the Washington Post, were cross-classified in **Table 1** for two variables: gender (female, male) and race (non-white, white). We

Table 1 Counts of fatal police shootings cross-classified in a two-way table by race and gender.

	Woman	Man	Row Totals
Non-white	98	2892	2990
White	155	2552	2707
Column Totals	253	5444	5697

dropped three incidents that involved young children. The values within each cell are the frequencies of cases for each combination of gender and race, and the numbers in the margins are totals by row and by column.

Incidents in **Table 1** are assumed to be independent of each other, and each incident falls into only one contingency table cell. In the theory of population sampling, the data in **Table 1** reflect the *Poisson sampling*, in which each population unit has a certain probability of being included in the sample data (Särndal et al., 1992). In this case, the marginal frequencies and the total number of observations are not bounded or fixed by design; therefore, cell proportions and the row and column proportions reflect actual probability distributions in the population, which can differ.

Table 2 shows an example of independent *Bernoulli sampling*, a special Poisson sampling case. The probability of being sampled into a category for each subject is the same in this sampling process. Specifically, this probability remains 0.5 because there are only two categories for the corresponding variable. Data in **Table 2** offer an example. They come from a phase-3 clinical trial of the Moderna vaccine for COVID-19. In this trial, 15,210 subjects were randomly assigned to the placebo and the vaccine groups (Baden et al., 2021), and they were later classified as being “symptomatic” or “asymptomatic.” In **Table 2**, the row margins are fixed by design.

Sampling distributions

When binomial sampling is used, as shown in **Table 2**, the row distributions are independent, and row margins follow a *binomial distribution*. Specifically, the *binomial distribution* is helpful for modeling counts or proportions of a variable (or event) that has two categories (or outcomes), such as students’ success or failure in a college course and getting head or tail when flipping a coin. The binomial distribution assumes that (1) there is a fixed total of independent trials for this event, in which each trial has two outcomes (known as a *Bernoulli trial*), and that (2) the probability of the targeted outcome (e.g., “success”) is the same for each trial. A Bernoulli trial with the quality of (2) is also known as an *identical trial*.

The mean and variance of the probability of the outcome Y in a fixed number of Bernoulli trials are

$$E(Y) = n\pi, \text{VAR}(Y) = n\pi(1 - \pi), \quad (1)$$

where n denotes the total number of trials and π denotes the probability of the occurrence of the targeted in one trial. The probability of having the targeted outcome occurring in y trials follows the binomial probability mass function:

$$P(y) = \frac{n!}{y!(n-y)!} \pi^y (1 - \pi)^{n-y}, y = 0, 1, 2, \dots, n \quad (2)$$

For example, if the probability of selecting a student enrolled in the National School Lunch Program in a school district is 0.2, then the probability of having seven students who are enrolled in this program, out of a random sample of 10 students in this school district, is $P(7) = \frac{10!}{7!(10-7)!} 2^7 (1 - .2)^{10-7} = 0.0008$, which is 0.08%. It is worth noting that the sample size matters here to ensure accuracy in making statistical inferences: The fixed total n should be at least 10 when π equals 0.5 to make the binomial distribution symmetric. This requirement is necessary to meet normality, a fundamental assumption undergirding frequentist statistical modeling. The required minimum gets larger when π gets closer to 0 or 1 to achieve this symmetric, bell-shaped distribution. In this case, we say the sample size is “large enough” and see the binomial distribution as a normal one.

In a more generalized case, if the “response” or outcome variable (i.e., symptomatic) has three or more categories (e.g., none, mild, severe), we would get contingency tables with three or more columns, and each row would follow a *multinomial distribution*. The multinomial distribution assumes that (1) there is a fixed total of independent trials for this event, in which each trial has more

Table 2 A 2×2 table of counts of treatment group by COVID-19 infection.

	Asymptomatic	Symptomatic	Row Totals
Placebo	15,025	185	15,210
Vaccine	15,199	11	15,210
Column Totals	30,224	196	30,420

than two outcomes, and that (2) the probability of each outcome (e.g., none, mild, severe) is the same for each trial. When the probability of each outcome is denoted by π_j (such that we have $\pi_1, \pi_2, \dots, \pi_j$), we must have $\sum_j \pi_j = 1$. Given the fixed total of identical trials, n , we must have the numbers of occurrences for each outcome, denoted by y_j (such that we have $y_1, y_2, \dots, y_j$), meet $\sum_j y_j = n$. Hence, the probability of having y_1 trials with the first outcome, y_2 trials with the second outcome, ..., and y_j trials with the last outcome follow the multinomial probability mass function below:

$$P(y_1, y_2, \dots, y_j) = \left(\frac{n!}{y_1! y_2! \dots y_j!} \right) \pi_1^{y_1} \pi_2^{y_2} \dots \pi_j^{y_j}, y_j = 0, 1, 2, \dots, n \quad (3)$$

Evidently, the binomial distribution described above is a special case of the multinomial distribution in which $j = 2$.

The *Poisson distribution* is typically used for counts or proportions of an event occurring randomly over time or space. It helps analyze the categorical data, such as the number of students that completed an activity online in an hour of a day, and the number of people who arrive at a university gym within 5 min during Day 1 of a new semester, as shown in Table 3. This distribution assumes the following:

(1) The numbers of people arriving at the university gym in non-overlapping time intervals are independent, (2) the probability of precisely one person arriving at a university gym, π , is proportional to the length of the time interval, and (3) the probability of two or more people arriving at the university gym in a sufficiently short interval is essentially 0. This Poisson probability distribution is unimodal and skews to the right. Its mean and the variance (of the probability of occurring) for the outcome Y are equal to π :

$$E(Y) = \text{VAR}(Y) = \pi. \quad (4)$$

For the outcome Y occurring in y trials, the Poisson probability mass function is

$$P(y) = \frac{e^{-\mu} \mu^y}{y!}, y = 0, 1, 2, \dots, n. \quad (5)$$

As π increases, the Poisson distribution becomes more symmetrical and closer to a bell shape.

Types of probability

For two-way tables, we distinguish between three types of probability: joint probability, marginal probability, and conditional probability. The *joint probability* of X and Y is the probability that an observation falls into cell (i, j) , a combination of categories from two variables. For example, given the fatal police shooting in Table 1, we can calculate maximum likelihood estimates for the joint probabilities, as shown in Table 4. For example, the estimated joint probability that the victim was a non-white woman equals $\hat{\pi}_{11} = 0.017$.

In some situations, our focus may be on the row or column margin, investigating the *marginal probability*. The marginal probabilities equal $\pi_{i+} = \pi_{i1} + \pi_{i2} + \dots + \pi_{ij}$ for row margins, and $\pi_{+j} = \pi_{1j} + \pi_{2j} + \dots + \pi_{ij}$ for column margins. For instance, using the police shooting data in Table 1, we can obtain maximum likelihood estimates of marginal probabilities of victims being non-white or white, or of marginal probabilities of victims being woman or man. They are shown in the row and column margins of Table 3, respectively. This marginal probability of victims being non-white is $\hat{\pi}_{1+} = 0.525$, and this marginal probability of victims being white is $\hat{\pi}_{2+} = 0.475$.

The *conditional probability* refers to the probability of an outcome conditioning on the probability of another outcome. For the two-way contingency tables, the conditional probability of the category j of one variable given the category i of another variable equals

$$\pi_{j|i} = \pi_{ij} / \pi_{i+}. \quad (6)$$

Given the vaccine data in Table 2, we can calculate conditional probabilities for each row and column, shown in Table 5. When the conditional probabilities for the two rows are approximately the same for a column (i.e., either the asymptomatic or symptomatic group), we may conclude that the vaccine is not effective for the asymptomatic/symptomatic group.

Table 3 Example: number of people who arrive at a university gym during day 1.

Time	Observed frequency
8:00 a.m.–8:05 a.m.	2
8:00 a.m.–8:05 a.m.	11
8:00 a.m.–8:05 a.m.	32
...	...

Table 4 Maximum likelihood estimates of joint and marginal probabilities of fatal police shootings.

	Woman	Man	Row Totals
Non-white	98/5697 = 0.017	2892/5697 = 0.508	2990/5697 = 0.525
White	155/5697 = 0.027	2552/5697 = 0.448	2707/5697 = 0.475
Column Totals	253/5697 = 0.044	5444/5697 = 0.956	1.000

Table 5 Maximum likelihood estimates of conditional probabilities from COVID-19 vaccine data.

	Asymptomatic	Symptomatic	Row Totals
Placebo	15025/15210 = 0.9878	185/15210 = 0.0122	1.00
Vaccine	15199/15210 = 0.9993	11/15210 = 0.0007	1.00
Column Totals	30224/30420 = 0.9936	196/30420 = 0.0064	1.00

Estimation of the unknown probability

Estimation method: maximum likelihood estimation

As hinted, our probability estimates presented in [Tables 4–5](#) are *maximum likelihood estimates*. In the practice of categorical data analysis, the probability parameter π in the binomial, multinomial, or Poisson distribution is often unknown. In this case, we need to *infer* the value of the unknown based on the knowns like n and y , which researchers often have control of in a research design. In an inferential method, a (point) *estimate* is a value derived from some observed data. It is an inference of the *estimator*, which can be seen as a random variable that follows a specific sampling distribution. In the most popular inferential method, *maximum likelihood estimation* method, the *estimate* is a value of the *estimator* following a *likelihood function*. The estimate maximizing this function is known as the *maximum likelihood estimate*, and it can be interpreted as the *most likely* value of the unknown parameter given the values of known parameters.

The maximum likelihood estimation method is popular because of an important property: Its estimator of probability (e.g., π) can be approximated by a sample proportion (e.g., $\hat{p}$), which follows a normal distribution when the sample size is large (i.e., the *Law of Large Numbers*, which claims that a proportion in a sample, $\hat{p}$, gets closer to the proportion in the population, p , as the sample size increases). This normality minimizes standard errors in parameter estimation. In other words, the maximum likelihood estimators have good “large-sample behaviors.” Other estimators tend to have relatively large standard errors because they usually do not follow a normal distribution. The mean and the variance of the sampling distribution of the maximum likelihood estimator π are the same as that of the binomial distribution for p :

$$E(\hat{\pi}) = E(\hat{p}) = \pi, \text{VAR}(\hat{\pi}) = \text{VAR}(\hat{p}) = \frac{\pi(1 - \pi)}{n}. \quad (7)$$

The likelihood function is mathematically equivalent to the joint probability mass function evaluated at the observed values ([Hogg et al., 2015](#)). For n trials that follow the binomial distribution, when the targeted outcome shows up $y_1, y_2, \dots, y_n$ times, this joint probability mass function is

$$P(y_1, y_2, \dots, y_n) = \prod_{i=1}^n \pi^{y_i} (1 - \pi)^{n - y_i} = \pi^{\sum y_i} (1 - \pi)^{n - \sum y_i}. \quad (8)$$

When treated as a function of π , the above function is known as the likelihood function $L(\pi)$:

$$L(\pi) = f(y_1; \pi) f(y_2; \pi) \dots f(y_n; \pi) = \pi^{\sum y_i} (1 - \pi)^{n - \sum y_i}. \quad (9)$$

When solving this function, we can obtain the maximum likelihood estimate of the unknown parameter π , denoted by $\hat{\pi}$. This $\hat{\pi}$ equals the observed proportion $\hat{p}$ of an event in a sample, expressed as

$$\hat{\pi} = \hat{p} = \frac{y}{n}. \quad (10)$$

Hypothesis testing

The maximum likelihood estimate of the unknown parameter in a population, as described above, is based on a sample. One sample is generally not a perfectly accurate reflection of the whole population. For example, while we know the probability of

obtaining a head when flipping a coin is $\pi = 0.5$ (i.e., the population parameter), we may get three heads when flipping a coin ten times (i.e., a sample) instead of five heads. This sample statistic about the probability of getting a head is $\hat{\pi} = 3/10 = 0.3$, which is not 0.5. The difference between a sample statistic and the population parameter is called *sampling error* or *margin of error*. In this example about flipping a coin, the sampling error is 0.2 (i.e., from 0.5 to 0.3).

Considering the existence of the sampling error, *hypothesis testing* is commonly used as a decision-making process to draw inferences about the population based on the sample data. It is about rejecting one of two *hypotheses* related to a research question: the *null hypothesis* and the *alternative hypothesis* (also known as the scientific/research hypothesis). In the prior example about flipping a coin, we can set two competing hypotheses about whether the sample statistic is close enough to the population parameter, corresponding to a two-sided hypothesis test. The null hypothesis is

$$H_0 : \pi = .5. \quad (11)$$

The alternative hypothesis is

$$H_1 : \pi \neq .5. \quad (12)$$

In hypothesis testing, *p-value* is the probability of observing the data at least as extreme as the observed sample, given that the H_0 is true. The *alpha level*, also known as *level of significance* or *Type-I error rate*, is the probability of rejecting the null hypothesis when the null is true; it can be seen as the acceptable error level.

To examine whether the sample statistic is close enough to the population parameter, we check whether the difference between the two is *statistically significant*. If yes, we shall reject the null hypothesis and go with the alternative hypothesis, concluding that the sample statistic is not close to the population parameter. By “statistically significant,” we mean the corresponding *p-value* is *lower* than the alpha level. How can we obtain the *p-value* then? Researchers use the *standard normal distribution* for converting the difference between the sample statistic and the population parameter into a *z score* and then into a *p-value*. In the previous example about flipping a coin, the standard error of the population parameter $\pi = 0.5$ is calculated as follows:

$$SE = \sqrt{\text{VAR}(\hat{\pi})} = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{.5(1-.5)}{10}}. \quad (13)$$

With this standard error, the difference between the sample statistic and the population parameter can be converted into a *z score*:

$$z = \frac{\hat{\pi} - \pi}{SE} = \frac{.3 - .5}{\sqrt{\frac{.5(1-.5)}{10}}} = -1.265, \quad (14)$$

which positions the sample statistic in a standard normal distribution, with which we can find the *p-value* is .898 for $z = -1.265$. When setting the alpha level at 0.05, which is common in educational research, we see the *p-value* 0.898 is not statistically significant, and so we fail to reject the null hypothesis and conclude that the sample statistic 0.3 is close enough to the population parameter 0.5.

As an alternative to the *p-value*, a *confidence interval* is a range of values centered around the sample statistic for hypothesis testing. *Confidence level*, $(1-\alpha) \times 100\%$, defines the width of this range. It is the probability that this range contains the actual value of the population parameter, associated with the alpha level. For example, when the alpha level is 0.05, the confidence level is 95%. The following formula is used to construct this confidence interval for the population parameter π :

$$CI = \hat{\pi} \pm z_{\alpha/2}(\widehat{SE}), \quad \widehat{SE} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \quad (15)$$

where $\hat{\pi}$ is the sample statistic. If the confidence interval contains 0, we fail to reject the null hypothesis and conclude that the statistic $\hat{\pi}$ is not statistically significant; if not, we reject the null hypothesis. Using the prior example, we get $\widehat{SE} = \sqrt{0.3(1-0.3)/10} = 0.145$, $z_{\alpha/2} = z_{0.025} = 1.960$, and so the confidence interval is $CI = 0.3 \pm 1.960 \times 0.145 = (0.02, 0.58)$. Because this interval does not contain 0, we reject the null hypothesis and conclude that the estimate 0.3 is statistically significant at the 0.05 level.

Broadly speaking, there are three pillars in the literature about hypothesis testing, and all of them can use a likelihood function to infer the population parameters. They are namely the likelihood-ratio test, the Wald test, and Rao's score test. All three tests are asymptotically equivalent under the null hypothesis. In other words, if the sample size is very large, we would not expect to see a noticeable difference among them. However, if the sample size is small to moderate, the values of the three statistics will diverge. The Wald test is the least reliable when sample size is small, and the likelihood ratio test and the score test are preferred in this case.

What is specified above is *Rao's score test*. In the *likelihood-ratio test*, a ratio of two maximum likelihood estimates is used as the test statistic, called the *likelihood-ratio test statistic*. It is

$$\chi^2 = -2\ln\left(\frac{l_0}{l_1}\right) \quad (16)$$

In this test statistic, l_0 is the maximum likelihood estimate of the parameter when H_0 is assumed to be true, and l_1 is the maximum likelihood estimate over all possible parameter values. In this case, l_1 is at least as large as l_0 . This test statistic takes

the natural log transform with the coefficient 2, as this conversion makes the statistic approximate a *chi-square* statistic, easier for computation. The corresponding *p*-value of this test statistic is the right-tailed probability of the chi-square distribution with $df = 1$.

In the *Wald test*, we evaluate the unrestricted standard error of the estimate of the parameter $\hat{\pi}$, *SE*, by substituting the maximum likelihood estimate for the unknown parameter in the formula of the true standard error (e.g., $SE = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$, because π follows the binomial distribution). The confidence interval in the Wald test using this *SE* is known as the *Wald confidence interval*. The test statistic in the Wald test is z^2 , which takes the form in Eq. (14), with the standard error being evaluated via the maximum likelihood estimation. It approximately follows a chi-square distribution with $df = 1$. The *p*-value in this Wald test is the right-tailed probability for z^2 in this chi-square distribution. In this test, the distributions for hypothesis testing are often *normal distributions* or *chi-square distributions* when the sample size is large enough.

Measures of association

The two most common measures of association for categorical variables are the odds ratios and Pearson's correlation coefficient. In this chapter, we primarily use the odds ratio as a measure and definition of association.

Odds ratio

Odds are the ratio of two conditional probabilities. For example, the odds of category *i* given column category *j* equal

$$odds = \pi_{ij} / (1 - \pi_{ij}). \quad (17)$$

Using the data in Table 1, the observed odds that a male victim was non-white (vs. white) equals

$$\frac{\hat{\pi}_{non-white|man}}{\hat{\pi}_{white|man}} = \frac{p_{12}/p_{+2}}{p_{22}/p_{+2}} = \frac{n_{12}}{n_{22}} = 2892/2552 = 1.133. \quad (18)$$

The probability of the victim being a non-white man is 1.133 times the probability of the victim being a white man. That is, non-white males were more likely to be shot than white males. For women, the odds equal $98/155 = 0.632$.

The ratio of odds, known as *odds ratio*, ranges from 0 to positive infinity. When the odds ratio is 1, it indicates the odds are equal, and so no relationship exists between the two variables (e.g., race and gender, as the row and column variables). Let θ_{i^*,jj^*} represent the odds ratio of two row categories *i* and *i** and two column categories *j* and *j**. Then the odds ratio equals $\hat{\theta}_{12,12} = 1.133/0.632 = 1.792$ in our example. It says that the *odds* that a victim was a non-white for men is 1.792 times the odds for women. It is worth noting that the odds ratio refers to ratios of odds and not to probabilities, as it would be wrong to state that "the probability of the victim being non-white for men is 1.792 times the probability for women."

Odds ratios can be computed when we calculate a "cross-product ratio" with joint probabilities:

$$\theta_{ii^*} = (\pi_{ij}\pi_{i^*j^*}) / (\pi_{i^*j}\pi_{ij^*}). \quad (19)$$

If we switch the odds in the numerator and denominator, we get $0.632/1.133 = 1/1.792 = 0.558$, and the interpretation would "flip": The odds that a victim was a woman for non-white victims is 0.558 times that for white victims.

Pearson correlation coefficient

The *Pearson correlation coefficient* measures the linear relationship between pairs of variables, which requires numerical codes for categories of each variable. For example, when two survey question items contain the unique responses "strongly disagree," "disagree," "agree," and "strongly agree," we use integers 1, 2, 3, and 4 to code these responses and denote these responses with u_s and v_s , respectively for two question items, where *u* and *v* denote the items and $s = 1, 2, 3, \dots, n$, indexing individual respondents. Thus, if Respondent 1 chooses "disagree" on one item, the response is $u_1 = 2$; if this respondent chooses "agree" on the other item, the response is $v_1 = 3$. Given this coded data, we can set a contingency table with *I* rows and *J* columns corresponding to unique responses to the two items for calculation. Alternatively, we can organize the data in the respondent by item format, displaying all individual respondents' responses in two columns corresponding to the two items. In the latter case, the formula for computing correlation is

$$r_{XY} = s_{XY} / (s_X s_Y) = \frac{\sum_{s=1}^n (u_s - \bar{u})(v_s - \bar{v})}{\sqrt{\sum_{s=1}^n (u_s - \bar{u})^2} \sqrt{\sum_{s=1}^n (v_s - \bar{v})^2}}. \quad (20)$$

where s_{XY} denotes the covariance between the two items, s_X and s_Y denote the standard deviation of the two items, respectively, $\bar{u}$ and $\bar{v}$ denote the means for the two items. For dichotomous variables, (20) is referred to as the *phi coefficient*. The choice of codes makes a difference for variables with more than two categories; we may choose the codes leading to the largest possible correlation,

which can be found from correspondence analysis. For the police shooting data in [Table 1](#), the correlation between gender and race is $r = -0.059$.

Tests of no association

We discuss tests for three possible null hypotheses: independence, homogeneous distributions, and linear relationship. The tests differ regarding the hypothesis being tested and thus the conclusions. After presenting the test statistics, we give examples of tests for nominal and ordinal variables.

Chi-square test statistics

Two common tests statistics have sampling distributions that are approximately chi-square for large sample sizes, requiring five or more counts per cell to ensure the shape of the distribution is chi-square. The two statistics are *Pearson's chi-square statistic*,

$$X^2 = \sum_i \sum_j \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}, \quad (21)$$

and the *likelihood ratio chi-square statistic* or *deviance*,

$$G^2 = 2 \sum_i \sum_j n_{ij} \log \left(\frac{n_{ij}}{\hat{\mu}_{ij}} \right), \quad (22)$$

where $\hat{\mu}_{ij}$ is the estimated expected frequency in cell (i, j) . Note that Pearson's chi-square statistic X^2 is an integral part of Pearson's correlation: $|r_{XY}| = \sqrt{X^2/n}$.

Both X^2 and G^2 equal 0 when $n_{ij} = \hat{\mu}_{ij}$ for all the cells (i.e., the observed frequencies equal the estimated expected frequencies). For testing independence and homogeneous association, specified below, X^2 and G^2 have the same value for the estimated expected frequency $\hat{\mu}_{ij} = n_{+j} n_{i+}/n$. That said, because the logic for arriving at $\hat{\mu}_{ij}$ differs in the sampling design, these two test statistics can lead to with different conclusions.

Independence

When the margins of a table are not fixed, we may run a test of independence to check if there is a relationship between variables. The null hypothesis of independence is $H_0 : \pi_{ij} = \pi_{i+} \pi_{+j}$ for all i and j , and the alternative hypothesis is $H_a : \pi_{ij} \neq \pi_{i+} \pi_{+j}$ for at least some combination of i and j . Estimated expected frequencies equal $\hat{\mu}_{ij} = n_{i+} n_{+j}/n$. If the null hypothesis is true (and cell counts > 5), then (21) and (22) have approximately chi-square distributed with $df = (I-1)(J-1)$.

For [Table 1](#), testing whether victims' gender and race are independent yields $X^2 = 20.068$ and $G^2 = 20.137$ with $df = (2-1)(2-1) = 1$ and the p -values < 0.01 . The data leads us to the conclusion that gender and race are related. The strength and nature of the dependence, as reflected by the odds ratio, are given in (18).

Homogeneous distributions

When the data follow independent binomial or multinomial distributions, we may be interested in whether the conditional distributions of the response variable are the same (i.e., homogeneous). The null hypothesis is $H_0 : \pi_{j|i} = \pi_{+j}$ or in terms of frequencies, $H_0 : \mu_{j|i} = \pi_{+j} n_{i+}$.

Using the clinical trial data in [Table 2](#), we test whether the distribution of participants receiving the placebo is the same as that of participants receiving the vaccine and get $X^2 = 155.47$, $G^2 = 187.88$ with $df = 1$ and the p -values < 0.01 . These values align with what we find with the odds ratio, which shows a discrepancy between the placebo group and the vaccine group: The odds of people receiving the vaccine are 17.01 times the odds of people receiving the placebo for the asymptomatic group.

Test of linear association

For ordinal variables, there are more powerful association tests that use the order of categories. The *Mantel-Haenszel statistic* is

$$M^2 = (n_{++} - 1)r^2, \quad (23)$$

where r is Pearson's correlation coefficient defined in (3). The hypotheses are $H_0 : \rho = 0$ and $H_a : \rho \neq 0$. M^2 follows a chi-square distribution with one degree of freedom. [Table 6](#) is a cross-classification of responses to two items of the *General Social Survey 2018*. The two items are "In general, would you say your quality of life is ..." and "In general, what would you rate your physical health?". The possible responses are "Excellent," "Very good," "Good," "Fair," and "Poor." To test whether physical health and

Table 6 Ordinal data for two survey questions in *General Social Survey 2018*.

Quality of life	Physical health					Row Totals
	Excellent	Very good	Good	Fair	Poor	
Excellent	221	160	66	29	2	478
Very good	120	410	328	81	11	950
Good	29	71	341	172	27	640
Fair	7	5	40	138	34	224
Poor	1	1	2	11	22	37
Column Totals	378	647	777	431	96	2329

quality of life are related, we use equally spaced integers from 1 to 5 for coding response categories, and obtain $r = 0.599$, $M^2 = 834.937$, and $p\text{-value} < 0.01$. The data support the conclusion of a linear relationship.

Generally speaking, to test a relationship, we needed $(I - 1)(J - 1)$ non-redundant odds ratios to characterize the association between two variables completely. Given our example, this number equals $(5 - 1)(5 - 1) = 16$. Making use of the ordering categories, we can summarize the association using one single statistic, r or M^2 , which is more powerful than X^2 and G^2 in (21) and (22).

Concluding remarks

This chapter scratches the surface of categorical data analysis, part of which overlaps with another chapter in this book about generalized linear models. For comprehensive coverage, a good starting place is Agresti (2019). Most statistical software programs have the functionality to analyze categorical data, including RStudio with R packages named “binom,” “Epi,” and “MASS.”

References

- Agresti, A., 2019. *An Introduction to Categorical Data Analysis*, third ed. Wiley, New York.
- Baden, L., El Sahly, H., Essink, B., Kottloff, K., Frey, S., Novak, R., Diemert, D., Spector, S., Rouphael, N., Creech, C., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., et al., 2021. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* 384, 403–416.
- Hogg, R., Tanis, E., Zimmerman, D., 2015. *Probability and Statistical Inference*, ninth ed. Pearson Education, Inc, Upper Saddle River, NJ.
- Särndal, C., Swensson, B., Wretman, J., 1992. *Model Assisted Survey Sampling*. Springer, New York.

Relevant websites

- Fatal Force, <https://www.washingtonpost.com/graphics/investigations/police-shootings-database/>.
- General Social Survey (GSS), <https://www.norc.ox.ac.uk/Research/Projects/Pages/general-social-survey.aspx>.

Generalized linear models[☆]

Peter K. Dunn, School of Science, Technology and Engineering, University of the Sunshine Coast, Sippy Downs, QLD, Australia

© 2023 Elsevier Ltd. All rights reserved.

This is an update of J.J. Faraway, Generalized Linear Models, in International Encyclopedia of Education (Third Edition), Edited by: Penelope Peterson, Eva Baker, Barry McGaw, Elsevier, 2010, <https://doi.org/10.1016/B978-0-08-044894-7.01331-2>.

Introduction	583
Definition	583
Fitting: estimating the regression parameters	584
Evaluating models	585
Inference	585
Estimating ϕ	585
Evaluating model fit	586
Example 1: binomial GLM	586
Example 2: rate model	587
Example 3: categorical data	588
Summary	589
References	589

Introduction

Univariate linear regression models, or just linear regression models, describe how the values of a response variable vary when values of the predictors vary. Linear regression models require that the response variable be continuous (and normally distributed), and the variation of the response observations around the regression line must be approximately constant. However, many situations do not meet these criteria. For example, many response variables are counts or proportions, and for many continuous variables the variation increases as the mean of the response variable increases (that is, the variance is a function of the mean). In these cases, another approach is to use generalized linear models (GLMs). GLMs apply for a wide variety of response variables, and unite many other commonly-used statistical models (such as logistic regression) under a common structure, which enables a common discussion of estimation, inference and residual analysis.

Definition

GLMs have two components: The probability function describing the variation in the values of the response variable (the random component), and a function connecting the mean of the response variable with the values of the predictors (the structural component).

The probability functions that describe how the response variable y is distributed (the random component) have a common structure; collectively, these distributions are called exponential dispersion models (EDMs). EDMs include normal distributions, Poisson distributions (for modeling *counts*), binomial distributions (for modeling *proportions* of a total) and gamma distributions (for modeling right-skewed positive continuous data). Some other distributions, such as the negative binomial and the Weibull, are not EDMs, but are sufficiently closely related to be used in a GLM-like manner.

The probability function for all distributions in the EDM family can be written in the form

$$f(y; \theta, \phi) = a(y, \phi) \exp \left[\frac{y\theta - \kappa(\theta)}{\phi} \right].$$

In this expression, θ (the *canonical parameter*) is a known, invertible function of the mean μ (specifically $\mu = \kappa'(\theta)$). The known function $\kappa(\theta)$ is called the *cumulant function*, and ϕ is a positive *dispersion parameter*. The variance of the EDM is $\text{var}(Y) = \phi \kappa''(\theta) = \phi V(\mu)$, where $V(\mu)$ is called the *variance function* (Table 1). The variance function uniquely identifies the distribution within the EDM family. This common structure allows general statements to be made about all GLMs for fitting models, inference and diagnostic analysis.

[☆] *Change History:* January 2021. P. K. Dunn updated and expanded the text initial sections of the text, changed Example 1, and slightly edited Examples 2 and 3 and the remaining text.

Table 1 Common EDMs and their variance function, dispersion parameter, commonly-used link function, and unit deviance (from which the deviance can be computed). For the binomial distribution, y represents the proportion of “successes” out of m independent trials, where m is known.

EDM	Variance function, $V(\mu)$	ϕ	Commonly-used link function	Unit deviance $d(y, \mu)$
Normal	1	σ^2	$\eta = \mu$ (identity)	$(y - \mu)^2$
Poisson	μ	1	$\eta = \log \mu$ (logarithmic)	$2\{y \log(y/\mu) - (y - \mu)\}$
Binomial	$\mu(1 - \mu)$	1	$\eta = \log \frac{\mu}{1-\mu}$ (logistic)	$2\left\{y \log\left(\frac{y}{\mu}\right) - (m - y) \log\left(\frac{m-y}{m-\mu}\right)\right\}$
Gamma:	μ^2	ϕ	$\eta = \log \mu$ (logarithmic)	$2\{-\log(y/\mu) + (y - \mu)/\mu\}$

Since θ is a known, invertible function of μ , the probability function can also be written directly in terms of the mean μ as

$$f(y; \mu, \phi) = b(y, \phi) \exp \left[-\frac{d(y, \mu)}{2\phi} \right]$$

where $d(y, \mu)$ is called the **unit deviance**. The notation $y \sim EDM(\mu, \phi)$ denotes that the response variable y has a distribution in the EDM family with mean μ and dispersion parameter ϕ .

GLMs are *linear* models, meaning that the predictors combine linearly in the structural component. For each of the $i = 1, 2, \dots, n$ observations, the p predictor variables combine in the form

$$\eta_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_p x_{p,i} = \mathbf{x}_i^T \boldsymbol{\beta}$$

where $x_{j,i}$ is the i th observation for the j th predictor variable, $\mathbf{x}_i$ is the vector of predictor variables for observation i , and $\boldsymbol{\beta}$ is the vector of length $p + 1$ of unknown regression parameters (and β_0 is called the *intercept*). η is called the *linear predictor*. The predictor values may involve transformed variables (for example, a predictor may take the form $\log x$), categorical variables (using indicator variables), or interactions between variables. The relationship between μ and η is defined by

$$\eta_i = g(\mu_i)$$

for all $i = 1, 2, \dots, n$ observations, where the function $g(\cdot)$ is called the *link function*. The link function must be a differentiable, monotonic function. A suitable choice of link function for many GLMs is the logarithmic function (Table 1) because this ensures $\mu > 0$ for any value of η . For example, the logarithmic link function is almost always used with a Poisson distribution since then $\eta = \log(\mu)$ and hence $\mu = \exp(\eta)$, and so $\mu > 0$ without needing to place restrictions on the values of η (and hence β). Using the logarithmic link function also means that additive effects in η have a multiplicative effect on y . Linear regression models use the identity link function $\mu = \eta$. Binomial EDMs require special link functions to ensure $0 < \mu < 1$, since then μ represents a proportion.

The complete GLM can be written as

$$\begin{cases} y_i \sim EDM(\mu_i, \phi) & \text{(The random component)} \\ g(\mu_i) = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_p x_{p,i} & \text{(The systematic component)} \end{cases}$$

Fitting a GLM is not the same as fitting a linear regression model with a transformed response variable. For example, a GLM fitted using a log link has the form $\mu_i = E[y_i] = \exp \eta_i$ (where $E[\cdot]$ means “expected value”). Transforming the response variable and fitting a linear regression model leads to a model with systematic component $\mu_i = E[\log y_i] = \eta_i$; that is, the GLM models the means directly, whereas the regression model models the logarithm of the response variable. Note that there are $p + 1$ unknown regression parameters in the systematic component.

Fitting: estimating the regression parameters

The estimates of the regression parameters $\boldsymbol{\beta}$ are almost always maximum likelihood estimates (MLEs), written $\hat{\boldsymbol{\beta}}$. For a GLM based on a normal distribution, the MLEs are the same estimates obtained using least-squares. For other EDMs, the MLEs are found numerically using the iteratively reweighted least squares algorithm (IRLS). Since EDMs all have a similar structure, a general algorithm for fitting GLMs can be developed that applies for all GLMs. The algorithm generally converges rapidly.

The IRLS algorithm for estimating $\boldsymbol{\beta}$ only depends on the mean and variance of the EDM, and not the full EDM distribution. In addition, the value of ϕ is not needed for computing $\hat{\boldsymbol{\beta}}$, so (if necessary) the value of ϕ can be estimated separately, hence estimating ϕ is discussed later. This estimate $\hat{\boldsymbol{\beta}}$ can then be used to compute the fitted values $\hat{\mu} = g(\hat{\eta}) = X\hat{\boldsymbol{\beta}}$ where X is the $n \times (p + 1)$ matrix of predictor values, and $\hat{\mu}$ and $\hat{\eta}$ are vectors of length n .

The estimate of the variance of $\hat{\beta}$, written $\widehat{\text{var}}(\hat{\beta})$, is easily computed from the results of the IRLS fitting algorithm. The standard error of $\hat{\beta}_j$, written $\text{s.e.}(\hat{\beta}_j)$, is the j th diagonal element of the square root of $\widehat{\text{var}}(\hat{\beta})$.

Fitting a GLM using software requires the user to nominate the EDM and the link function to use, the response variable, and the list of predictors. All major statistical software packages can fit GLMs.

Evaluating models

Competing GLMs can be compared using a measure of discrepancy between the observations y_i and the fitted means $\hat{\mu}_i$. One discrepancy measure is the **residual deviance**

$$D(y, \hat{\mu}) = \sum_{i=1}^n d(y_i, \hat{\mu}_i)$$

where $d(y, \mu)$ is called the **unit deviance** (Table 1). The **scaled residual deviance** is defined as $D^*(y, \hat{\mu}) = D(y, \hat{\mu})/\phi$. The residual deviance is often just written as D and the scaled deviance as D^* . The distribution of the scaled deviance approximately follows a chi-squared distribution (with $n - (p + 1)$ degrees of freedom), and is related to the log-likelihood of the EDM (Dunn and Smyth, 2018, p. 236).

Another discrepancy measure is the **Pearson statistic** X^2 :

$$X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\text{var}(\hat{\mu}_i)} = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\phi V(\hat{\mu}_i)}.$$

Both the deviance and Pearson statistics are equivalent to the residual sum-of-squares for normal distributions: $\sum_{i=1}^n (y_i - \hat{\mu}_i)^2$. If the value of ϕ is unknown, an estimate is used. Software generally reports the value of the deviance (or scaled deviance) after fitting a model.

Inference

Since GLMs all have a similar structure, inference can be developed in general terms. Competing models can be compared by observing the *changes* in the discrepancy for nested models. If the *change* in residual deviance is used, this is equivalent to a likelihood-ratio test (LRT).

Consider two models: a larger model L with residual deviance D_L with $p_L + 1$ unknown parameters in the linear predictor, and a second (smaller) model S with $(p_S + 1) < (p_L + 1)$ unknown regression parameters and residual deviance D_S . Model S is *nested* in Model L if the regression parameters in Model S are a subset of those in model L ; then $D_L \leq D_S$.

If the value of ϕ is known, then $(D_S - D_L)/\phi$ has an approximate χ^2 distribution, with $p_L - p_S$ degrees of freedom. When the value of ϕ is unknown and is estimated, then

$$\frac{(D_S - D_L)/(p_L - p_S)}{\tilde{\phi}}$$

has an approximate F -distribution, with $(p_L - p_S, n - (p_L + 1))$ degrees of freedom (where $\tilde{\phi}$ is some estimate of ϕ).

Regression coefficients can be tested individually for significance using *Wald tests*. "The test statistic $(\hat{\beta}_j - \beta_j)/\text{s.e.}(\hat{\beta}_j)$ for testing if the regression coefficient has the value β_j has an approximate normal distribution when the value of ϕ is known (or an approximate t -distribution when ϕ is estimated)." The standard errors of the regression parameters are computed as a by-product of fitting the GLM, as noted above, so values of the Wald test statistic (and the corresponding p -value) are readily computed and reported by software.

For linear regression models only, the LRT and Wald tests are equivalent with exact distributions. For all other EDMs, the tests are approximate only, but the distributional assumptions are better, in general, for the LRTs. All major software packages include facilities for conducting LRTs and Wald tests.

Estimating ϕ

Although an estimate of ϕ is not needed to estimate β , a value of ϕ is needed to compute D^* and $\text{s.e.}(\hat{\beta}_j)$. The MLE of ϕ is not usually used since it is difficult to compute and has poor properties anyway (Dunn and Smyth, 2018); usually, the mean deviance or the Pearson estimate is used.

The mean deviance estimator of ϕ for a model with $p + 1$ estimated regression parameters in the linear predictor η is $D(y, \hat{\mu}) / (n - (p + 1))$, where D is the value of the residual deviance for the model. The Pearson estimate of the dispersion parameter is $X^2 / (n - (p + 1))$, where X^2 is the value of the Pearson statistic for the model. Different software may, by default, report either the Pearson or deviance estimate of ϕ .

Evaluating model fit

Since GLMs all have a similar structure, common methods of evaluating the model (diagnostic analysis) can be developed. However, the standard residuals used in linear regression (that is, $r_i = y_i - \hat{\mu}_i$ for each observation) do not make sense in the GLM context, since the variation of $\hat{\mu}_i$ is a function of the value of $\hat{\mu}_i$. Instead, Pearson and deviance residuals are commonly used. *Pearson residuals* are defined as

$$r_{P,i} = \frac{y_i - \hat{\mu}_i}{\sqrt{\text{var}(\hat{\mu}_i)}}$$

where the denominator rescales the residuals so that the variance is approximately constant for all $\hat{\mu}_i$. Notice that $X^2 = \sum r_{P,i}^2$.

By analogy to the Pearson residuals, write $D = \sum r_{D,i}^2$. Then, the *deviance residuals* are $r_{D,i} = \sqrt{d(y_i, \hat{\mu}_i)}$ and where $r_{D,i}$ takes the sign of $y_i - \hat{\mu}_i$. Both the Pearson and deviance residuals have approximate normal distributions (and so behave similar to the familiar residuals from linear regression models), but normality is not guaranteed. Hence, if the deviance or Pearson residuals do not show approximate normality, either the model is inadequate or the model is adequate but the residuals are simply non-normally distributed. This non-normality can be very evident when using EDMs with discrete values of y (i.e., binomial and Poisson EDMs). A better choice of residual, suitable even for discrete EDMs, is to use quantile residuals (Dunn and Smyth, 1996) which are always approximately normally distributed (apart from sampling variation) whenever the correct EDM is used.

Other statistics used for diagnostic analysis—such as hat values, and influence measures such as Cook's distance—also have approximate analogies with linear regression models in the context of GLMs.

All major software packages compute deviance and Pearson residuals, other diagnostic statistics, and can plots these statistics. R allows the computing of quantile residuals (using the `statmod` package; Smyth et al., 2020).

Example 1: binomial GLM

The most common type of GLM is the binomial GLM. The data summarized in Table 2, obtained from the Australian Bureau of Statistics web page (<https://www.abs.gov.au>), contains the number of four-year-old students enrolled at school in Queensland, Australia, in 2019, classified by Sex, School Type and indigenous status. Consider modeling the *proportion* of First Nation students, say μ_i , using a *binomial* distribution, where m_i is the total number of students in each category. This assumes that each student is independent of other students, which may not be true.

A simplistic systematic component is $\log[\mu_i / (1 - \mu_i)] = \beta_0$, which (unrealistically) states that the proportion of First Nation students is not related to the Sex of the student or the School Type. The link function used here is the *logit link*, and the model is commonly called a *logistic regression model*. Other link functions can be used (such as the *probit link*), but the logistic link function is commonly-used and allows interpretation in terms of odds and odds ratios: $\mu / (1 - \mu)$ represents the odds that a randomly-chosen student is First Nation, so the systematic component can be written as $\log\text{-odds} = \beta_0$, or $\text{odds} = \exp(\beta_0)$. This shows that the regression coefficients are related to the *odds* of the event of interest (here, the odds that a student is First Nation). All popular statistical software can fit this GLM. Using R (R Core Team, 2020), the systematic component is $\log[\mu_i / (1 - \mu_i)] = -2.351$. This suggests that the *odds* that a student is First Nation is $\exp(-2.351) = 0.0953$. Using the data in Table 2 directly, the odds are $6027/63252 = 0.0953$ also.

Table 2 The number of students in Queensland schools in 2019, classified by the sex of the student, the type of school and whether the students are First Nation or not. Note: The cohort sizes are $m = m_1 + m_2$.

Sex of student: x_1	School type: x_2	Number of First Nation students: m_1	Number of non-First Nation students: m_2	Proportion First Nation students: y_i
Females	Government	2540	21,219	0.1069
Males	Government	2734	22,574	0.1080
Females	Non-government	391	9496	0.0395
Males	Non-government	362	9963	0.0351
	Totals	6027	63,252	0.0870

More realistically, the predictors are incorporated into the model. Categorical variables are included through numerical coding: the predictor x_1 can be coded such that 0 represents females and 1 represents males; the predictor x_2 can be coded using 0 for government schools, and 1 for non-government schools. (This coding is arbitrary.) Then a possible systematic component is

$$\log[\mu_i / (1 - \mu_i)] = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i}$$

where x_1 and x_2 take the values 0 or 1 as appropriate, as described above. Using the R function `glm()` with `family = "binomial"`, the following estimates are found (output slightly edited):

Coefficient	Estimate	Std. error	z-value	p-value
Intercept $\hat{\beta}_0$	-2.113	0.0202	-104.7	<0.0001
Sex (female: 0; male: 1): $\hat{\beta}_1$	-0.006564	0.027147	-0.242	0.809
School type (gov: 0; non-gov: 1): $\hat{\beta}_2$	-1.135	0.0399	-28.456	<0.0001
Null deviance: 1033.1798 on 3 degrees of freedom				
Residual deviance: 2.9359 on 1 degree of freedom				

The output shows that the systematic component of the model is

$$\log[(\mu_i / (1 - \mu_i))] = -2.133 - 0.006564x_{1,i} - 1.135x_{2,i}$$

where the values of x_1 and x_2 are defined above. The random component is $y_i \sim \text{Bin}(\mu_i)$.

The p -values (final column), from a Wald-test, show that the School Type is significantly associated with the proportion of First Nation students, but the Sex of the student is not. Alternatively, a LRT can be conducted by fitting a binomial GLM *without* School Type as a predictor. This yields a binomial GLM with a residual deviance of 1033.17 (with 2 df). The *change* in the residual deviance between these two models is substantial ($1033.17 - 2.9359 = 1030.2$), by adding just one extra parameter (that is, on $3 - 2 = 1$ degree of freedom). This produces a very small p -value after referring to a χ^2_1 distribution, which means the model *with* School Type as a predictor is a far better model than the model *without* this predictor.

The deviance residuals (one for each row in Table 2) are -0.4497, 0.4358, 1.1280, and -1.1276. While deviance residuals do not follow a normal distribution exactly, none of these residuals are large in value (say, greater than three in absolute value), so they raise no concerns with the fitted model.

Example 2: rate model

Consider the (hypothetical) data in Table 3. The number of incidents in eight schools are shown along with the numbers of students and an average school test score. We would like to know if the number of incidents is related to the test score. However, larger schools naturally will tend to generate more incidents.

The response variable is a count, best modeled using a Poisson distribution. The number of incidents out of the school size μ/N could be modeled using a binomial distribution. Instead, we will model the counts using a Poisson distribution with the log link (to ensure a positive mean μ), since the counts may be of greater interest (for example, it may inform the number of support staff employed). The systematic component of the model is:

$$\log(\mu / N) = \beta_0 + \beta_1 x$$

where we have omitted the subscripts i for convenience. Rearranging:

$$\log \mu = \log N + \beta_0 + \beta_1 x$$

Table 3 The number of incidents occurring in schools of given sizes and average test scores.

Incidents, y	7	5	1	5	2	1	5	2
Number of students, N	985	850	241	495	838	616	759	842
Score, x	32	74	44	20	50	73	48	36

This is called a *rate model*. The term $\log N$ will result in multiplying the rate at which the incidents occur by the number of students as we might wish here. This term has no regression coefficient (or, rather, it is fixed at one). Such a term is called an *offset*. The remaining part of the predictor is linear in the score, x . All popular statistical software can fit this GLM. Using R, the following estimates are found (output slightly edited):

Coefficient	Estimate	Std. error	z-value	p-value
Intercept $\hat{\beta}_0$	-4.6703	0.5319	-8.78	<0.0001
Score	-0.0139	0.0114	-1.22	0.22
Null deviance: 8.0380 on 7 degree of freedom				
Residual deviance: 6.4944 on 6 degree of freedom				

The negative coefficient for Score of -0.0139 suggests that the rate of incidents decreases with increasing test Scores. The z -statistic of -1.22 can be used to test the significance of this effect (p -value: 0.22) but, as before, using the change in the residual deviance is more reliable. Here the null model (fitted with only the intercept) is the appropriate smaller model, with a *null deviance* of 8.0380. The test statistic is $8.0380 - 6.4944 = 1.5436$. Comparing to a χ^2_1 distribution, the p -value is 0.21, as there is one degree of freedom difference between the two models. Either test suggests there is insufficient evidence of a relationship between test scores and the rate of incidents.

Example 3: categorical data

A GLM approach is also useful for modeling categorical data. Consider the data in [Table 4](#) from [Faraway \(2016\)](#). Subjects from the 1996 US National Election Study (NES) study were cross-classified by political identification (PID) and level of education.

Suppose we consider the counts as Poisson distributed with mean μ_{ij} and use a GLM with log link

$$\log \mu_{ij} = \mu + \text{PID}_i + \text{Education}_j$$

where the both predictors are considered as nominal (unordered) factors, each with seven levels. The rate at which observations fall in each of the cells will be some overall mean rate multiplied by a parameter reflecting the level of party identification, times another parameter representing the level of education. This is equivalent to claiming that PID and level of education are independent. The estimated parameters for the factor levels represent the relative proportions of observations falling in each row or column of the data. These are usually not of interest, and say nothing about the *relationship* between the two factors. Instead, the interaction terms, not the main effects, are usually of interest as they model the dependence between the factors.

Fitting the main-effects model gives a residual deviance of 40.7 on 36 degrees of freedom. A model that allows dependence between the two factors would require a separate parameter for each cell. Then, the number of parameters would equal the number of data points, so the model would fit perfectly with a deviance of zero on zero degrees of freedom. If we compare this latter saturated model to the independence model above, the difference in deviance is 40.7 on 36 degrees of freedom resulting in a p -value of 0.27. Hence, we prefer the smaller model which represents independence of the two factors.

Traditionally, the test for independence uses a Pearson's χ^2 which we might compute here as the sum of the squared Pearson residuals, giving a value of 38.8 again on 36 degrees of freedom. Hence, the outcome of the test is similar to the deviance-based test.

The analysis so far has not used the ordinality of both factors (they have been treated as nominal). One way to accommodate these in the model is by using scores. We assign scores u_i and v_j to the rows and columns such that $u_1 \leq u_2 \leq \dots \leq u_7$ and $v_1 \leq v_2 \leq \dots \leq v_7$.

Table 4 A subset of subjects drawn from the 1996 US NES study, cross-classified by political identification and level of education.

	MS	HS dropout	HS	Some college	CC degree	BA degree	MA degree
Strong democratic	5	19	59	38	17	40	22
Weak democratic	4	10	49	36	17	41	23
Independent democratic	1	4	28	15	13	27	20
Independent	0	3	12	9	3	6	4
Independent republican	2	7	23	16	8	22	16
Weak republican	0	5	35	40	15	38	17
Strong republican	1	4	42	33	17	53	25

Notes: MS: 8 years of education or less; HS Dropout: High-school dropout; HS: High school diploma (or equivalent); Some college; CC degree Community or junior college degree; BA degree: BA degree; MA degree: postgraduate degree.

$\dots \leq v_7$. The assignment of scores requires some judgment. If one has no particular preference, even spacing allows for the simplest interpretation. The model then becomes

$$\log \mu_{ij} = \mu + \text{PID}_i + \text{Education}_j + \gamma u_i v_j$$

where γ represents the amount of association and is like an (unscaled) correlation coefficient: $\gamma = 0$ means independence. We assigned the scores 1–7 respectively to each of the factors and fit this model. We obtained $\gamma = 0.0287$; the positive value indicates a positive association, although the size of the effect cannot be directly interpreted. The residual deviance is 30.6 which represents a difference of 10.1 for one extra parameter (i.e., on one degree of freedom) from the independence model. This is a strongly significant result, indicating an association here. Notice that it was essential to use the information about the ordinality of the factors to obtain this result.

Summary

Generalized linear models provide a common approach to a broad range of response modeling problems. Normal, Poisson, and binomial responses are the most commonly used, but other distributions can be used as well. GLMs also need a link function to be specified which allows further flexibility in the modeling. The GLM can be fitted using a common procedure and a mechanism for hypothesis testing is available. Diagnostics using residuals provide a way to check that chosen models are adequate.

Even though GLMs have commonalities, GLMs have aspects particular to the response type. For example, a logistic regression is a GLM but the user still needs to understand the particular interpretation of odds in this type of model. With categorical data, the imposition of a GLM sometimes requires parametrizations that are difficult to interpret and other approaches may be more revealing. Thus while GLMs might provide some convenient structure across seemingly different models, we must recognize that they do not provide a comprehensive solution.

The basic GLM has been extended in many directions. [Wedderburn \(1974\)](#) introduced the idea of quasi-likelihood, which avoids explicitly specifying the distribution of the response, requiring only that we specify the link and variance functions. This idea was extended to the notion of generalized estimating equations with particular application to longitudinal data by [Liang and Zeger \(1986\)](#). [Fahrmeir and Tutz \(2001\)](#) describe GLMs for multivariate responses, and random effects can be added to the model using generalized linear *mixed* models ([Stroup, 2012](#)).

Further reading about GLMs can be found in several books. [McCullagh and Nelder \(1989\)](#) is the classical reference although this can be difficult for readers without a background in statistical theory. More accessible texts include [Fox \(2008\)](#) and [Dobson and Barnett \(2008\)](#). [Dunn and Smyth \(2018\)](#) discuss theory, and then explain the application of specific GLMs using R software. [Faraway \(2016\)](#) has an excellent, short introduction to GLMs and extensions based on using R, while [Hardin and Hilbe \(2007\)](#) discuss fitting GLMs in Stata.

References

- Dobson, A., Barnett, A., 2008. An Introduction to Generalized Linear Models, third ed. Chapman and Hall, Boca Raton, FL.
- Dunn, P.K., Smyth, G.K., 1996. Randomized quantile residuals. *J. Comput. Graph Stat.* 5 (3), 236–244.
- Dunn, P.K., Smyth, G.K., 2018. Generalized Linear Models With Examples in R. Springer.
- Fahrmeir, L., Tutz, G., 2001. Multivariate Statistical Modelling Based on Generalized Linear Models. Springer, New York.
- Faraway, J.J., 2016. Extending the Linear Model With R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. CRC press.
- Fox, J., 2008. Applied Regression Analysis and Generalized Linear Models, second ed. Sage, Thousand Oaks, CA.
- Hardin, J., Hilbe, J., 2007. Generalized Linear Models and Extensions, second ed. Stata Press, College Station, TX.
- Liang, K., Zeger, S., 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73 (1), 13–22.
- McCullagh, P., Nelder, J., 1989. Generalized Linear Models, second ed. Chapman and Hall, London.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Smyth, G.K., Hu, Y., Dunn, P.K., Phipson, B., 2020. statmod: statistical modeling. R Package.
- Stroup, W.W., 2012. Generalized Linear Mixed Models: Modern Concepts, Methods and Applications. CRC press.
- Wedderburn, R., 1974. Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika* 61 (3), 439–447.

PCA and other dimensionality-reduction techniques

Nikolay Trendafilov and Michele Gallo, Department of Human and Social Sciences, University of Naples "L'Orientale", Naples, Italy

© 2023 Elsevier Ltd. All rights reserved.

Introduction	590
Matrices and matrix notations	590
Data matrices	590
Data matrix pre-processing	591
Matrix definitions and operations	592
Matrix rank and matrix decompositions	592
PCA	592
PCA of correlation matrix (EVD)	592
PCA of data matrix (SVD)	594
EFA	595
LDA	597
Concluding remarks	598
References	599

Introduction

This article is about techniques for reduction of the data dimensionality. Arguably, the most popular of them is the principal component analysis (PCA) (Jolliffe, 2002). It is used practically in any research field, sometime under different names. Other techniques to be considered here are exploratory factor analysis (EFA) and linear discriminant analysis (LDA). We will see later that both, EFA and LDA, elaborate PCA in some way. For this reason, PCA will be given more attention.

To really appreciate the dimensionality-reduction techniques, it is essential to speak in terms of matrices. So we start with some matrix refreshment.

Matrices and matrix notations

Data matrices

Great number of applications require collecting data arranged in a table, called data matrix. What happens is that for a given set of p variables, measurements are available/made for every involved subject on each variable. The resulting p values are ordered in a p -dimensional vector usually called (multivariate) observation. Further, the already available/made observations are arranged one under another in a table X called a (data) matrix. Suppose that $n(>p)$ observations are available/made altogether. Then the matrix X contains exactly np values: n observations consisting of p values each. We say that the matrix X has size $n \times p$, and assuming that the data are real numbers we write $X \in \mathbb{R}^{n \times p}$. This shows that the matrix X can be considered as either a collection of n row-vectors $(x_i, i = 1, \dots, n)$, called observations, or by p column-vectors $(x_j, j = 1, \dots, p)$, called variables, i.e., $X = [x_1, \dots, x_p]$. We also say that the (data) matrix X has n rows and p columns. Occasionally, the n observations are written as column-vectors of size $p \times 1$ and arranged one to another, resulting in a matrix with n columns and p rows. This is in fact, the transposed matrix to the constructed above X , and is denoted by X^T , i.e. $X^T = [x_1^T, \dots, x_n^T] \in \mathbb{R}^{p \times n}$. For example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \in \mathbb{R}^{3 \times 2} \text{ and } A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \in \mathbb{R}^{2 \times 3}.$$

Consider the 9×7 data matrix in Table 1, which is a subset of the 24 Psychological Tests given to 145 students in the Grant high school in Chicago (Harman, 1976, p. 123) and (Holzinger and Swineford, 1939). The full data matrix itself is available from the R package MBESS and is obtained by typing data (HS). For simplicity, we consider only six tests, namely Word Recognition (V14), Number Recognition (V15), Figure Recognition (V16), Object-Number (V17), Number-Figure (V18) and Figure-Word (V19) given to the first 9 students, with "M" for male and "F" for female:

Even for this very small data matrix, it is difficult to say much about the involved variables and/or observations (students). Indeed, they can be considered as points in $\mathbb{R}^9$ and $\mathbb{R}^7$ respectively, and thus, difficult to visualize. Unfortunately, this is always the case in real applications, because as a rule the dimension p is much greater than 2 or 3.

Table 1 A subset of 24 psychological tests (Harman, 1976, p. 123).

#	Sex	V14	V15	V16	V17	V18	V19
1	M	173	91	96	8	2	10
2	F	174	81	106	9	15	17
3	M	168	84	101	1	7	16
4	M	130	84	101	10	15	14
5	M	184	98	99	9	9	15
6	F	188	86	116	10	10	16
7	F	192	95	118	16	19	20
8	F	168	97	113	5	6	16
9	F	168	84	107	6	4	7
Mean		171.67	88.89	106.33	8.22	9.67	14.56
Var		327.0	41.6	61.5	16.9	32.0	15.0

That is why, one of the most important tasks in data analysis is the reduction of the data dimensionality p . The rationale behind doing this is twofold: first, not all of the involved variables are equally important, and second, some weighted sum (linear combination) of them can be more informative than each of the individual variables. The purpose of the dimension reduction techniques is to find solutions involving as fewer of the original variables as possible while preserving and explaining as much of the information carried out by the data as possible.

Data matrix pre-processing

For now, let us forget about the categorical variable “Sex” in [Table 1](#), and concentrate on the remaining continuous variables V14 to V19. We note that V14 has much bigger mean and variance than the rest of them. Usually this is caused by measuring the variables in different units, suggesting that they are not readily comparable. The remedy for this is to standardize them, so that they all have zero mean and unit variance. This is achieved by subtracting the mean of each variable/column from its observed values, and then dividing them to the square root of the variance. For example, the value 173 from [Table 1](#) for the first student on V14 is standardized as:

$$\frac{173 - 171.67}{\sqrt{327.0}} = 0.07354917.$$

The easiest way to standardize all variables V14 to V19 from [Table 1](#), is to insert them into an 9×6 matrix, say, X , and use the following R function:

```
scale(X, center = TRUE, scale = TRUE).
```

The resulting standardized variables are given in [Table 2](#):

If the standardized variables from [Table 2](#) are collected in a matrix Z , then the sample correlation matrix of the data is given by

$$\frac{Z^T Z}{n-1}.$$

Nowadays, it is widely accepted first to center the data matrix X , e.g., using R this can be done with:

```
X0 <- scale(X, center = TRUE, scale = FALSE),
```

Table 2 Standardized variables V14 to V19 from [Table 1](#).

#	V14	V15	V16	V17	V18	V19
1	0.07	0.33	-1.32	-0.05	-1.36	-1.18
2	0.13	-1.22	-0.04	0.19	0.94	0.63
3	-0.20	-0.76	-0.68	-1.75	-0.47	0.37
4	-2.30	-0.76	-0.68	0.43	0.94	-0.14
5	0.68	1.41	-0.94	0.19	-0.12	0.11
6	0.90	-0.45	1.23	0.43	0.06	0.37
7	1.12	0.95	1.49	1.89	1.65	1.40
8	-0.20	1.26	0.85	-0.78	-0.65	0.37
9	-0.20	-0.76	0.09	-0.54	-1.00	-1.95
Mean	0	0	0	0	0	0
Var	1	1	1	1	1	1

and then, make its columns/variables have unit lengths (sum of squares):

$$X \leftarrow X0 \% \% \text{diag}(1/\text{sqrt}(\text{colSums}(X0 * X0))).$$

Such pre-processing is known as the data matrix sphering. For sphered X , the product $X^T X$ gives the sample correlation matrix of the data. In other words, the sphered data are obtained from the standardized ones after division by $\sqrt{n-1}$, i.e., $X = Z/\sqrt{n-1}$.

Matrix definitions and operations

The following facts about matrices will be useful in the sequel:

- $A \in \mathbb{R}^{n \times n}$ is called symmetric matrix, if $A = A^T$;
- $D \in \mathbb{R}^{n \times n}$ is diagonal, if its off-diagonal entries are 0 and is denoted by $D = \text{diag}(d_1, \dots, d_n)$; the identity matrix is $I_n = \text{diag}(1, \dots, 1) \in \mathbb{R}^{n \times n}$;
- for $A, B \in \mathbb{R}^{n \times p}$ and $\alpha, \beta \in \mathbb{R}$, we define $\alpha A \pm \beta B = C \in \mathbb{R}^{n \times p}$, and the elements of C are $c_{ij} = \alpha a_{ij} \pm \beta b_{ij}$, where $i = 1, \dots, n, j = 1, \dots, p$;
- for $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, $AB = C \in \mathbb{R}^{m \times p}$, and the elements of C are $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$, where $i = 1, \dots, m, j = 1, \dots, p$; note, that the matrix product is not commutative ($AB \neq BA$)! Indeed:

$$A^T A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

but

$$AA^T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 11 & 17 \\ 11 & 25 & 39 \\ 17 & 39 & 61 \end{bmatrix} \in \mathbb{R}^{3 \times 3}.$$

- $A \in \mathbb{R}^{p \times p}$ is invertible, if there exists $W \in \mathbb{R}^{p \times p}$, such that $AW = WA = I_p$; then W is called the inverse of A and is denoted by A^{-1} ;
- $A \in \mathbb{R}^{p \times p}$ is singular, if it is not invertible; $A \in \mathbb{R}^{p \times p}$ is singular, if and only if its determinant is 0, and we write $\det(A) = 0$.

Matrix rank and matrix decompositions

The rank of a matrix $A \in \mathbb{R}^{n \times p}$ is a kind of its intrinsic dimension. It is defined as the size of its largest invertible (square) sub-matrix. Let $n > p$, then $A \in \mathbb{R}^{n \times p}$ can have rank at most p . If A does have rank p , it is called full rank matrix. If its rank is $r < p$, then A is called rank deficient. In particular, a square $A \in \mathbb{R}^{p \times p}$ is full rank if and only if it is invertible.

By definition, A and A^T have same ranks. Later we will see how to find the rank of a matrix, as well as that $\text{rank}(A) = \text{rank}(A^T A) = \text{rank}(AA^T)$. These identities imply that if $A \in \mathbb{R}^{n \times p}$ is full rank, so is $A^T A$, and AA^T is rank deficient.

The following useful inequalities hold for any $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{p \times q}$:

$$\text{rank}(A) + \text{rank}(B) - p \leq \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}. \quad (1)$$

The concept of matrix rank plays a central role in the dimension reduction. To see this, look at the product AA^T from the example above: it demonstrates that a symmetric matrix $S (= AA^T)$ can be represented by a matrix A of a smaller size. In real applications, the dimension reduction task requires solution of the opposite problem: we are given S , and need to find suitable A of a smaller size and/or rank, such that $S = AA^T$ or at least $S \approx AA^T$. The former problem gives an idea of what is meant by matrix decomposition, while the latter one suggests the purpose of the low-rank matrix approximation which aims precisely at dimensionality reduction.

PCA

PCA of correlation matrix (EVD)

PCA is one of the first dimension reduction techniques, introduced by Hotelling (1933) and published (note!) in Journal of Educational Psychology. In fact, Hotelling re-invented the eigenvalue decomposition (EVD) as a tool to find successive and mutually orthogonal principal components (PCs). Later it was realized that the Hotelling's first PC, coincides with the Pearson's (straight) line passing through a cloud of points and minimizing the sum of squared distances between the data points and the line (Pearson, 1901).

EVD states that every symmetric matrix $S \in \mathbb{R}^{p \times p}$ can be written as $S = Q\Lambda Q^T$, where $Q \in \mathbb{R}^{p \times p}$ is orthogonal matrix, i.e., $Q^T Q = QQ^T = I_p$ (identity matrix), and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^{p \times p}$ is diagonal matrix, with $\lambda_1 \geq \dots \geq \lambda_p$. If all $\lambda_i > (\geq) 0$, then S is called

positive (semi-)definite. The columns of $Q (= [q_1, \dots, q_p])$ are called eigenvectors of S and λ_i is the eigenvalue corresponding to q_i . If all λ_i are different, then Q is unique. Note that $\det(S) = \lambda_1 \dots \lambda_p > 0$, i.e., a positive definite S is always invertible.

Suppose $X \in \mathbb{R}^{n \times p}$ is a full rank sphered data matrix. Then, its sample correlation matrix $X^\top X$ is a positive definite matrix with unit main diagonal. This assumption is not too restrictive, because most real data need standardization of the original variables. Then, applying EVD to $X^\top X$ gives:

$$X^\top X = Q\Lambda Q^\top = \lambda_1 q_1 q_1^\top + \dots + \lambda_p q_p q_p^\top, \quad (2)$$

where $\lambda_1 \geq \dots \geq \lambda_p > 0$. Very small eigenvalues of $X^\top X$ usually indicate data problems, e.g., some variable (column of X) is a multiple of another one, known as collinearity.

Let us see now, how the neat mathematical representation (2) of $X^\top X$ translates into statistical terms. First, the elements of the eigenvectors q_i of $X^\top X$ are called component loadings. Multiplying (2) on the left hand side by $q_1^\top$ and on the right hand side by q_1 gives:

$$q_1^\top X^\top X q_1 = (X q_1)^\top X q_1 = y_1^\top y_1 = \lambda_1. \quad (3)$$

The linear combination $y_1 = X q_1$ is called the first PC. As X is centered (as sphered), y_1 is centered too. This shows that the variance of y_1 is simply given by the sum of its squared elements and thus, equal to the corresponding eigenvalue λ_1 . More generally, the variances of the PCs are given by the eigenvalues of $X^\top X$, which explains, why they cannot be negative.

Consider the transformed data matrix $Y = XQ \in \mathbb{R}^{n \times p}$, known as the matrix of component scores. As Q is orthogonal matrix, this implies that Y is obtained from X by a high-dimensional orthogonal rotation, which by definition preserves angles and lengths. In this connection it is worth reminding that the oldest approach to EVD proposed by [Jacobi \(1846\)](#) relies on successive planar rotations. As X is centered, so is Y . It follows from (3), that the covariance matrix $Y^\top Y$ of the PCs, is diagonal Λ .

The sum of the diagonal elements of a square matrix $A \in \mathbb{R}^{p \times p}$ is called trace and it is denoted by $\text{trace}(A)$. Direct calculations show that $\text{trace}(AB) = \text{trace}(BA)$ for any two matrices $A, B \in \mathbb{R}^{n \times p}$. The total variance contained in the data X is given by the sum of the variances of all variables, i.e., by the trace of the covariance matrix. In our situation, the covariance matrix reduces to the sample correlation matrix $X^\top X$, which trace is p : the sum of its unit diagonal entries. Then, it follows from (2) that:

$$\text{trace}(X^\top X) = \text{trace}(Q\Lambda Q^\top) = \text{trace}(Q^\top Q\Lambda) = \text{trace}(\Lambda) = \sum_{i=1}^p \lambda_i = p,$$

This shows that the PCs in Y have the same total variance as the original variables, $\text{trace}(Y^\top Y) = \text{trace}(\Lambda)$, and that the average PC variance is 1.

The expression (2) immediately appeals for dimension reduction: we can choose any number $r < p$ to approximate $X^\top X$ as follows:

$$X^\top X \approx \lambda_1 q_1 q_1^\top + \dots + \lambda_r q_r q_r^\top, \quad (4)$$

where the right hand side of (4) is a sum of matrices of rank 1 each, as products of vectors. Moreover, $\text{trace}(X^\top X) = p \approx \lambda_1 + \dots + \lambda_r$, which helps to measure how good our approximation is. It is common practice to take into account only the eigenvalues with values greater than average, i.e., $\lambda_i > 1$, and use their eigenvectors in (4) for low-rank approximation of $X^\top X$.

It is convenient to have (4) expressed in matrix form as follows:

$$X^\top X \approx Q_r \Lambda_r Q_r^\top = (Q_r \Lambda_r^{1/2}) (Q_r \Lambda_r^{1/2})^\top, \quad (5)$$

where $\Lambda_r = \text{diag}(\lambda_1, \dots, \lambda_r) \in \mathbb{R}^{r \times r}$ is diagonal matrix containing the first largest r eigenvalues of $X^\top X$, and $Q_r = [q_1, \dots, q_r] \in \mathbb{R}^{n \times r}$ is composed by the first r columns of Q (corresponding eigenvectors). A matrix that contains a subset of the columns of an orthogonal matrix is called orthonormal, i.e., Q_r is orthonormal and $Q_r^\top Q_r = I_r$ only.

Let us apply what we know so far on the small data set in [Table 1](#), or more precisely to the sphered variables V14 to V19. Their sample correlation matrix is given by:

$$X^\top X = \begin{bmatrix} 1.00 & 0.41 & 0.45 & 0.29 & 0.01 & 0.32 \\ 0.41 & 1.00 & 0.15 & 0.26 & -0.08 & 0.24 \\ 0.45 & 0.15 & 1.00 & 0.43 & 0.43 & 0.49 \\ 0.29 & 0.26 & 0.43 & 1.00 & 0.71 & 0.38 \\ 0.01 & -0.08 & 0.43 & 0.71 & 1.00 & 0.74 \\ 0.32 & 0.24 & 0.49 & 0.38 & 0.74 & 1.00 \end{bmatrix}.$$

If we insert this correlation matrix into an R data frame and apply `eigen()`, we find that its eigenvalues are: 2.84, 1.37, 0.74, 0.61, 0.41 and 0.03. Only the first two are greater than the average PC variance 1, which suggests to use the first two PCs, and produce

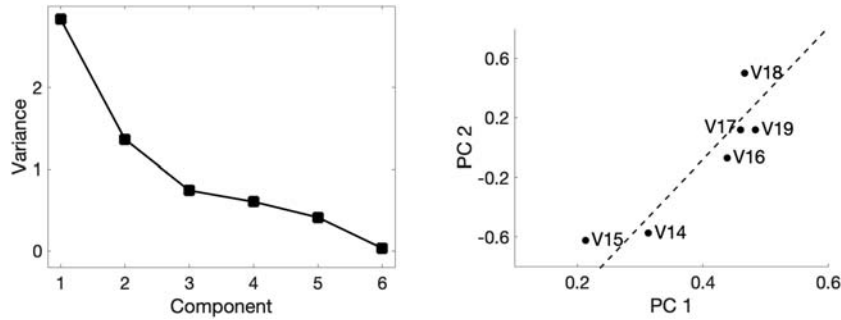


Fig. 1 Scree plot (left panel) and component loadings plot.

a rank-2 approximation to $X^T X$. They explain 70.08% $\left(= \frac{2.8395+1.3654}{6}\right)$ of the total variance, and as a result, we are able to analyze the data using only two PCs, rather than the initial six variables, i.e. we achieve two-thirds dimension reduction. With this PCA solution we achieve a rank-2 approximation of $X^T X$, namely $X^T X \approx \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T = Q_2 \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q_2^T$.

In this case, the decision for the number of useful PCs is easy, because their variances are well away from 1. In practice, it is common to draw a scree plot of the variances as in **Fig. 1**(left), and consider only the PCs before the graph flattens. Following this rule, one can conclude that, in fact, three PCs are the best choice. Indeed, they explain 82.46% $\left(= \frac{2.8395+1.3654+0.7427}{6}\right)$, which brings considerable improvement in the fit, but the level of dimension reduction is compromised: we need three PCs, which is half the number of the original variables. This situation gives a flavor of the inevitable bargaining of fit for simplicity and vice versa accompanying the dimension reduction.

Suppose, we make our choice in favor of the simplicity and work only with the first two PCs. Their component loadings are given in the left two columns of **Table 3**. Clearly, the first PC represents the relative weights of the involved six variables, while the second PC separates the “recognition” variables (V14, V15 and V16) from the “comparative” ones (V17, V18 and V19). Their graphical representation in **Fig. 1**(right) is quite interesting. It reveals that V16 is, in fact, closer to the “comparative” tests. Moreover, if the coordinate system is rotated such that the first axis follows the direction of the dashed line, then the separation of the two groups of tests can be achieved in a single dimension. Note, that many statistical packages report the PC loadings as they are given in the right couple of columns of **Table 3**.

PCA of data matrix (SVD)

Nearly at the same time when Hotelling proposed his PCA with EVD of covariance/correlation matrix, another matrix decomposition was proposed by [Eckart and Young \(1936\)](#). This is the singular value decomposition (SVD) and its great importance is that PCA can be performed directly on the data matrix X . Any potential noise and/or errors in the data multiply when forming the covariance/correlation matrix. Thus, working with SVD is beneficial in two ways: saves calculations and offers more precise analysis.

To simplify the notations, we consider the so-called economy SVD, which states that for every matrix $X \in \mathbb{R}^{n \times p}$, there exists orthonormal matrix $U \in \mathbb{R}^{n \times p}$ ($U^T U = I_p$) and orthogonal matrix $V \in \mathbb{R}^{p \times p}$, such that:

$$X = U \Sigma V^T = \sigma_1 u_1 v_1^T + \dots + \sigma_p u_p v_p^T, \quad (6)$$

Table 3 Two ways to report component loadings.

	q_1	q_2	$q_1 \sqrt{\lambda_1}$	$q_2 \sqrt{\lambda_2}$
V14	0.31	−0.57	0.53	−0.67
V15	0.21	−0.62	0.36	−0.73
V16	0.44	−0.07	0.74	−0.08
V17	0.46	0.12	0.78	0.14
V18	0.47	0.50	0.79	0.58
V19	0.48	0.12	0.82	0.14
SS	1	1	2.84	1.37

where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{p \times p}$ with $\sigma_1 \geq \dots \geq \sigma_p \geq 0$, called singular values of X . The number of positive singular values of X gives its rank.

A left multiplication of (6) by $X^\top$ gives $X^\top X = V\Sigma^2 V^\top$, which connects the SVD of X to the EVD of $X^\top X$ by taking $Q = V$ and $\lambda_i = \sigma_i^2, i = 1, \dots, p$ in (2). Similarly, a right multiplication of (6) by $X^\top$ shows that the columns of U are eigenvectors of $XX^\top$, and $\sigma_1^2, \dots, \sigma_p^2$ are also their eigenvalues. Thus, $X^\top X$ and $XX^\top$ have the same (non-zero) eigenvalues, and same ranks.

As with (2) for the EVD case, we can achieve dimension reduction with (6), by choosing any number $r < p$ and approximate X as follows:

$$X \approx \sigma_1 u_1 v_1^\top + \dots + \sigma_r u_r v_r^\top, \quad (7)$$

where the right hand side of (7) is again a sum of matrices of rank 1 each. Here, we measure the goodness of our approximation by $\text{trace}(X^\top X) = p \approx \sigma_1^2 + \dots + \sigma_r^2$.

It is convenient to have (7) expressed in matrix form as follows:

$$X \approx U_r \Sigma_r V_r^\top, \quad (8)$$

where $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$ contains the first largest r singular values of X , with the corresponding columns (singular vectors) in U_r and V_r .

The matrix of component scores $Y = XQ = XV$ from Section **PCA of correlation matrix (EVD)** is also given by $Y = U\Sigma$ from (6). Going back to the data set in **Table 2**, we realize that PCA can be performed without forming $X^\top X$. By applying the R function `svd()` to X , we find the component loadings matrix V , which coincides with Q from (2). The right number of components to use is found in the same way as in **PCA of correlation matrix (EVD)** section.

Finally, note that some people use $V\Sigma$ and U for loadings and scores instead. For example, this is the case with SPSS. Of course, this is just a matter of scaling and does not affect the interpretation. One can easily recognize which scaling is used by adding the squared loadings. If they add up to 1, then V is the loadings matrix. If they add up to the variance of the corresponding component, $V\Sigma$ present the loadings.

Our personal opinion is that loadings with unit lengths are easier to interpret within a component, and to compare between components. Thus, if the loadings are of interest, we would take V and the scores become $U\Sigma$. However, if the scores are of interest, we would work with (orthonormal) U and the loadings become $V\Sigma$.

EFA

EFA is another dimension reduction technique, which model is much more complicated than PCA. As a result EFA needs whole books to cover different aspects and approaches ([Bartholomew et al., 2011](#)). Applying EFA is also more complicated. The simplest option seems to be the R function `factanal`; for more complex EFA needs there are R libraries like `psych` or `fungible`.

Not surprisingly, all these features make EFA far less popular than PCA. To simplify our task we concentrate on the dimension reduction aspect of EFA and view it as another (data) matrix decomposition ([Harman, 1976](#)).

Let $X \in \mathbb{R}^{n \times p}$ be a standardized data matrix. The EFA model is in fact a specific matrix decomposition/factorization of X ([Harman, 1976, 2.9](#)):

$$X \approx FL^\top + P\Psi, \quad (9)$$

where the matrix parameters $F \in \mathbb{R}^{n \times r}$, $L \in \mathbb{R}^{p \times r}$, $P \in \mathbb{R}^{n \times p}$ and $\Psi \in \mathbb{R}^{p \times p}$ (diagonal) are unknown. Note that, as X is centered, so are F and P .

It is very helpful to rewrite (9) using block matrices as follows:

$$X \approx FL^\top + P\Psi = [F \ P][L \ \Psi]^\top = WA^\top, \quad (10)$$

where $W = [F \ P] \in \mathbb{R}^{n \times (r+p)}$ means that F and P are put next to each other and form a “bigger” matrix W with the same number of rows, but with $r + p$ columns. As seen from (10), the blocks are treated as single entries in matrix manipulations, i.e., $W = [F \ P]$ and $A = [L \ \Psi]$ are treated as 1×2 vectors.

The product $FL^\top$ in (9) provides the (low) rank- r approximation of X , while each column of $P\Psi$ only adjusts the corresponding column in $FL^\top$. That is why F and P are called respectively common and specific/unique factors. Clearly, their roles are very different and they need to be well separated. For this reason, the EFA model (9) requires F and P to be uncorrelated/orthogonal to each other: $F^\top F = I_r, P^\top P = I_p, F^\top P = O_{r \times p}$. Equivalently, the block matrices in (10) fulfill $W^\top W = I_{(r+p)}$.

The classical EFA makes use of (10) to derive:

$$X^\top X \approx (WA^\top)^\top (WA^\top) = AA^\top = LL^\top + \Psi^2, \quad (11)$$

which, in turn is an EFA model for the sample correlation matrix $X^\top X$. The model (11) has only two unknowns L and Ψ . The term $LL^\top$ is a (low) rank- r approximation to the covariances (off-diagonal elements) of $X^\top X$, while Ψ^2 simply adjusts its diagonal entries

to fit the main (unit) diagonal of $X^\top X$. As Ψ^2 is sparse (contains only $1/p$ non-zero entries), this makes the EFA model (11) the oldest matrix completion problem.

Comparing (5) and (11) clearly shows the difference between PCA and EFA: PCA finds the best low rank r approximation to $X^\top X$, while EFA finds the best low rank r approximation to the covariances of $X^\top X$. Note, that in general the approximation $LL^\top + \Psi^2$ to $X^\top X$ is not a low rank matrix.

The considerable difference between the EFA model (10) and SVD is in the specific structure of $A (= [L \ \Psi])$. As a result, W and A cannot be found such that the model fits X exactly. Instead, F, L, P and Ψ are sought such that the difference/error $E = X - WA^\top$ is made as small as possible. Here we want to make the sum of squared errors $\sum_i \sum_j e_{ij}^2 = \text{trace}(E^\top E)$ as small as possible, which is known as the least-squares EFA (LS-EFA).

First, making use of $W^\top W = I_{(r+p)}$, we find that the sum of squares is:

$$\text{trace}(E^\top E) = \text{trace}(X^\top X) - 2\text{trace}(W^\top XA) + \text{trace}(A^\top A), \quad (12)$$

and note that the first term in (12) is constant. Second, we observe that the last term in (12) depends on A only. This suggests a simple alternating procedure to find W and A . First, fix $A = [L \ \Psi]$, and solve:

$$\max_W \text{trace}(W^\top XA), \quad (13)$$

for $W = [F \ P]$. To do this, consider the SVD of $XA = U\Sigma V^\top$, which suggests that the solution of (13) is $W = UV^\top$. Second, update $A = [L \ \Psi] = [X^\top F \text{diag}(X^\top P)]$. Plug the new A back in (13) and repeat the process until, say, $\text{trace}(E^\top E) < 10^{-6}$. This can be implemented in a small R code.

Unfortunately, this solution is not unique, because $XA \in \mathbb{R}^{n \times (r+p)}$ is always rank deficient. This conclusion follows from (1), by noting that X and A have ranks at most p . This lack of uniqueness is the notorious factor indeterminacy in EFA.

We apply the proposed alternating EFA procedure to the small data from **Table 1**. Again, we work with the sphered data matrix X already subject to PCA in **PCA of correlation matrix (EVD)** section. We look for two-factor ($r = 2$) solution, to keep the results comparable to the PCA ones. We also apply another LS-EFA procedure fals from the R library fungible, which finds L_R and Ψ_R by LS minimization of the difference/error $E = X^\top X - LL^\top - \Psi^2$ in (11).

The results are summarized in **Table 4**. The unique variances in Ψ_R^2 and Ψ^2 are reasonably similar, except the entry for V14. At first glance, the factor loadings L_R and L seem quite different. However, their plots in **Fig. 2** show that they form very similar configurations of points in $\mathbb{R}^2$, possibly after some reflection.

The last couple of columns of **Table 4** show that L matches L_R very well after multiplication by the orthogonal $Q_R = \begin{bmatrix} 0.5427 & -0.8400 \\ -0.8400 & -0.5427 \end{bmatrix}$. One notes that $\det(Q_R) = -1$, i.e., indeed, Q_R is a reflection, which makes an angle -28.57° with the first axis $e_1 = (1, 0)$. Both factor loadings L and L_R testify for distinction of V14, V15 from V16, and for their separation from V17, V18 and V19. This separation is also clearly visible in **Fig. 2**, as well as that it can be realized in a single dimension. Variables like V18 with zero unique variance are considered singular and are known as Heywood cases. Note, that many EFA numerical procedures, e.g., fals, keep such variances at 0.005.

The component loadings from the last two columns of **Table 3** (given by $V\Sigma$ from (6)) can be roughly compared to the factor loadings in **Table 4**. Their values are reasonably similar to the factor loadings L_R . Their plots depicted respectively in **Fig. 1** and **Fig. 2**, also support such a conclusion.

It is interesting to visualize the students as presented by the component and factor scores. As $F^\top F = I_r$, we take U from (6) as component scores for this comparison. Indeed, we also have for them $U^\top U = I_r$, and the scaling with Σ is already applied above to the loadings $V\Sigma$. **Fig. 3** shows that the male and female students are not separated.

Table 4 Factor loadings from two LS-EFA methods.

Vars	<i>Fals/fungible</i>			<i>LS-EFA (13)</i>			<i>Fitted</i>	
	L_R		Ψ_R^2	L		Ψ^2	LQ_R	
	1	2		1	2		1	2
V14	0.45	-0.58	0.46	0.93	0	0.12	0.51	-0.79
V15	0.28	-0.53	0.64	0.48	0.01	0.76	0.25	-0.41
V16	0.62	-0.14	0.59	0.46	-0.44	0.59	0.62	-0.15
V17	0.68	0.06	0.54	0.29	-0.65	0.46	0.71	0.11
V18	0.83	0.54	0.01	0.01	-0.98	0.00	0.83	0.52
V19	0.74	0.06	0.44	0.34	-0.73	0.32	0.80	0.11
SS	2.38	0.94	2.68	1.52	2.11	2.25	2.53	1.10

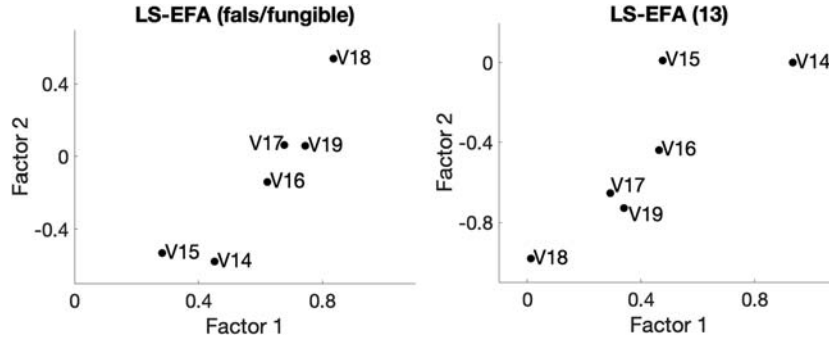


Fig. 2 Plots of factor loadings from two LS-EFA methods.

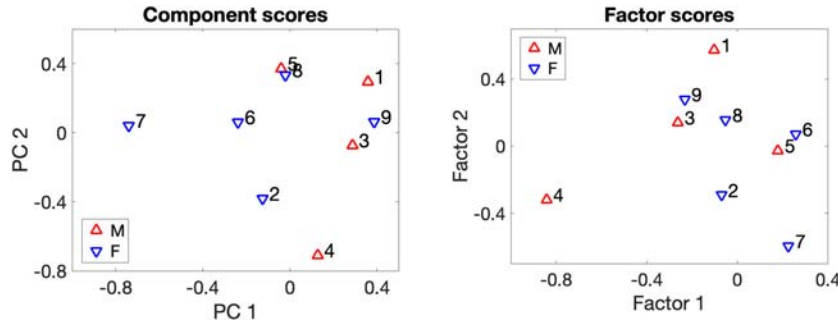


Fig. 3 Component scores U and factor scores F .

LDA

LDA is designed to analyze data with observations divided into, say, g groups, i.e., the rows of X are grouped. We will see that LDA achieves dimensionality reduction, which is additionally facilitated by the presence of groups.

As with PCA, LDA considers linear combinations $Y = XA$ of the original variables, usually called canonical variates (CVs), or discriminant functions. The LDA goal is to find a full-rank matrix $A \in \mathbb{R}^{p \times r}$, such that the groups are better separated/discriminated in the transformed data $Y (=XA)$, than with respect to any of the original variables. The number r of “new” variables (CVs) is usually much smaller than p . In fact, it should be $r < \min\{p, g\}$. In this sense, LDA is another dimension reduction technique.

As before, n observations are made on p variables which are collected in a $n \times p$ ($n > p$) data matrix X . The difference is that these n observations are a priori divided into g groups/populations. Thus, if n_i is the number of individuals in the i th group, then we have $n_1 + n_2 + \dots + n_g = n$. For example, as indicated by the variable Sex in Table 1, we have male and female students, i.e., the observations of these data are divided into two groups, $g = 2$.

The observations group membership is expressed by an indicator matrix $G \in \mathbb{R}^{n \times g}$ of zeros and ones. Each column of G corresponds to one group and has 1 at its i th position if the i th observation of X belongs to that particular group and zero otherwise. Clearly, we have $G^T G = \text{diag}(n_1, \dots, n_g) = N$. The matrix containing the means of the groups is given by $\bar{X} = N^{-1} G^T X$.

For example, here is the group indicator matrix G for the data in Table 1:

$$\text{Sex} = (\text{M F M M M F F F})^T \Rightarrow G = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

and respectively $G^T G = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = N$ and $\bar{X} = N^{-1} G^T X$ gives:

$$\bar{X} = \begin{bmatrix} 0.12 & -0.02 & 0.26 & 0.08 & 0.07 & 0.06 \\ -0.15 & 0.02 & -0.32 & -0.11 & -0.09 & -0.07 \end{bmatrix}.$$

Let X be centered/sphered data matrix. In LDA, the sample covariance/correlation matrix $X^T X$ is usually denoted by S_T (for Total). The following representation of S_T plays a central role in LDA:

$$\begin{aligned} S_T &= X^T X + \bar{X}^T N \bar{X} - \bar{X}^T N \bar{X} = (G\bar{X})^T (G\bar{X}) + (X - G\bar{X})^T (X - G\bar{X}) \\ &= X_B^T X_B + X_W^T X_W = S_B + S_W, \end{aligned} \quad (14)$$

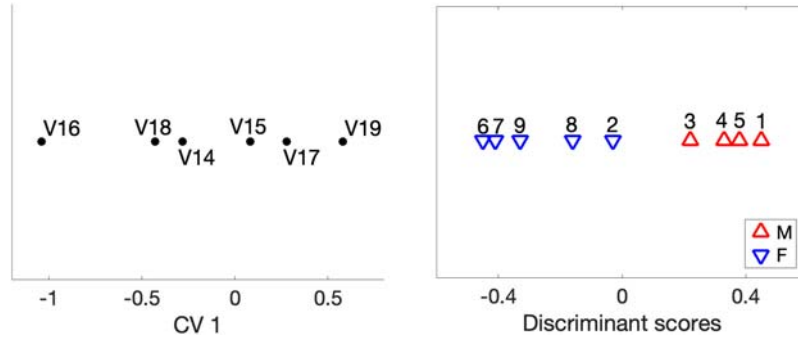


Fig. 4 Discriminant function (CV) and scores.

where S_B and S_W express the between- and within-groups scatter/variability of the data. There are two important implications following from (14):

1. S_B is a product of G and $\bar{X}$ with ranks $\min\{n, g\}$ and $\min\{p, g-1\}$ respectively; then, it follows from (1) that its rank can be at most $\min\{p, g-1\}$;
2. S_T is given/fixed, so the separation among the groups can be improved, either by “increasing” S_B , or by “reducing” S_W (Fisher, 1936).

The latter observation helps to define LDA as the following problem:

$$\max_a \frac{a^\top S_B a}{a^\top S_T a}, \quad (15)$$

which is solved in turn for every column of the transformation $A = [a_1, \dots, a_r]$.

The LDA problem (15) is equivalent to the following generalized EVD:

$$(S_B - \mu S_T)a = 0_p, \quad (16)$$

and can be solved by the function `geigen` from the R package `geigen`.

For our example with $g = 2$, LDA can produce only one CV. Thus, the dimension is reduced to $r = 1$, and $A \equiv a \in \mathbb{R}^p$. Indeed, the solution of (16) gives a single non-zero eigenvalue $\mu = 0.8487$ with corresponding eigenvector:

$$a = (-0.2820, 0.0822, -1.0398, 0.2788, -0.4286, 0.5809) \in \mathbb{R}^6,$$

satisfying $a^\top S_T a (= a^\top X^\top X a) = 1$. The coefficients of (the CV) a order the variables on the line as seen in Fig. 4: V16, V18, V14, V15, V17, V19. This is quite a different interpretation from the one of PCA, which orders them as V15, V14, V16, V17, V19, V18 on the hypothetical dashed line in Fig. 1.

Finally, the one-dimensional LDA projection $y = Xa$ ($y^\top y = 1$) of the students is given by their discriminant scores:

$$y = (0.45, -0.03, 0.22, 0.33, 0.38, -0.45, -0.41, -0.16, -0.33) \in \mathbb{R}^9,$$

where, since students 2, 6, 7, 8, and 9 are female, the male and female students are clearly separated with their positive and negative scores. This outperforms the two-dimensional plots of the component and factor scores depicted in Fig. 3. This is not surprising, because the LDA primary goal is the separation of the groups of observations. On the other hand, the LDA variables representation in Fig. 4 looks less satisfying than the outcomes from PCA and EFA, depicted in Fig. 2.

As with PCA and EFA, we can do LDA using X and X_B only, without forming S_T and S_B needed for (15) and (16). This is achieved by their generalized SVD and the function `gsvd` from the same package `geigen`. The only non-zero generalized singular values of X_B and X are 0.6776 and 0.7355. Their squared quotient gives the same eigenvalue μ as above. However, the corresponding generalized singular vector b is only proportional to a , obtained from (16), i.e., $b = \alpha a$ for some $\alpha > 0$. The correct CV a is reconstructed by taking $\alpha = \sqrt{b^\top S_T b}$.

Concluding remarks

Nowadays, the methods for dimension reduction routinely analyze huge data sets with thousands or even millions of variables. This requires modifications to drastically reduce their number by dropping off less important variables and produce sparse solutions, involving only fewer important ones. Thus, the goodness-of-fit of the dimension reduction is further compromised and traded for a clearer and/or easier interpretation. This is a very active research area and brings together artificial intelligence, machine learning and statistical communities. Some modern aspects of PCA, EFA and LDA are considered in (Trendafilov and Gallo, 2021).

References

- Bartholomew, D., Knott, M., Moustaki, I., 2011. *Latent Variable Models and Factor Analysis: A Unified Approach*, third ed. John Wiley & Sons Ltd., Chichester, UK.
- Eckart, C., Young, G., 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1, 211–218.
- Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugenics* 7, 179–184.
- Harman, H.H., 1976. *Modern Factor Analysis*, third ed. University of Chicago Press, Chicago, IL.
- Holzinger, K., Swineford, F., 1939. A study in factor analysis: the stability of a bifactor solution. In: *Supplementary Educational Monograph*, vol. 48. University of Chicago Press, Chicago, IL.
- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441, 498–520.
- Jacobi, C., 1846. Über ein leichtes verfahren die in der theorie der sacularstroungen vorkommendern gleichungen numerisch aufzulösen. *Crelle's J.* 90, 51–94.
- Jolliffe, I.T., 2002. *Principal Component Analysis*, second ed. Springer-Verlag, New York, NY.
- Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. *Phil. Mag.* 2, 559–572.
- Trendafilov, N., Gallo, M., 2021. *Multivariate Data Analysis on Matrix Manifolds (With Manopt)*. Springer, New York, NY.

Exploratory factor analysis

Nickolay Trendafilov^a and Kei Hirose^b, ^a Department of Human and Social Sciences, University of Naples “L’Orientale”, Naples, Italy; and ^b Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan

© 2023 Elsevier Ltd. All rights reserved.

Introduction	600
EFA model	600
Number of factors	602
Two ways to find Λ and Ψ	603
EFA interpretation	603
Simple structure concept	604
Simple structure criteria and rotation methods	604
Rotation to simple structure target	605
Concluding remarks	606
References	606

Introduction

Factor analysis (FA) originated in psychometrics as a tool to model and measure the human abilities and associated individual differences. In 1904, Charles Spearman published his seminal paper (Spearman, 1904), which is considered the beginning of FA. The purpose was to measure the human intelligence, thought at that time to be a single unobservable entity, called general intelligence g . He used regression-like equations but with the difference that g is not available (observable) as in the regression problems. Soon, the theory of intelligence evolved and the Spearman’s (one factor) model with a single latent construct became less popular and replaced by the so called multiple FA (Thurstone, 1947).

There are two major approaches to FA: exploratory (EFA) or confirmatory (CFA) (Mulaik, 2010, Ch 15). Here we concentrate on EFA. Its difference from CFA will be briefly outlined in the next section.

EFA model

Suppose, we are given p centered variables $x_i, i = 1, \dots, p$. EFA models every variable x_i as the following linear combination:

$$x_i \approx \lambda_{i1}f_1 + \dots + \lambda_{ir}f_r + \psi_i u_i = \sum_{j=1}^r \lambda_{ij}f_j + \psi_i u_i, \quad (1)$$

where $f_1, \dots, f_r$ and $u_1, \dots, u_p$ are called respectively common and unique or specific factors. Thus, there are $r(< p)$ common factors for all x_i , and p unique factors, one for each of the input variables, which explains their name.

A graphical expression of the EFA model can be given by a path diagram as depicted in Fig. 1. The variable x_i is affected by the r common factors $f_1, \dots, f_r$ and by the unique factor u_i . We will see later, that the factor loading λ_{ij} gives the correlation between x_i and f_j .

Usually we work with a $n \times p$ ($n > p$) data matrix $X = [x_1, \dots, x_p]$ of n observations made on p standardized variables divided by $\sqrt{n-1}$, i.e. $X^T X$ is their sample correlation matrix. Then the classical EFA model in (1) is more convenient to be rewritten in matrix form as follows (Mulaik, 2010, p. 136):

$$X \approx F\Lambda^T + U\Psi, \quad (2)$$

where the common and unique factors from (1) are collected as columns of the $n \times r$ matrix F and the $n \times p$ matrix U respectively. As X is centered, F and U are centered, i.e. $1_n^T F = 0_r$ and $1_n^T U = 0_p$ (zero vectors). Λ is known as the matrix of factor loadings, while Ψ is diagonal and Ψ^2 gives the variances of the unique factors U . Thus, ψ_i^2 is the part of the variance of the i th variable x_i that the common factors cannot explain.

Eq. (2) is called the r -factor EFA model. The number r of factors is usually much smaller than the number p of input variables and will be discussed in Section Number of factors.

By definition, the unique factors U are uncorrelated and with unit lengths, that is, $U^T U = I_p$, which is to say that U is an orthogonal matrix. Moreover, the common and unique factors are assumed uncorrelated, i.e. $F^T U = 0_{r \times p}$ (zero matrix), meaning they are orthogonal to each other. Then, the EFA model (2) implies the following model for the correlation matrix of the data:

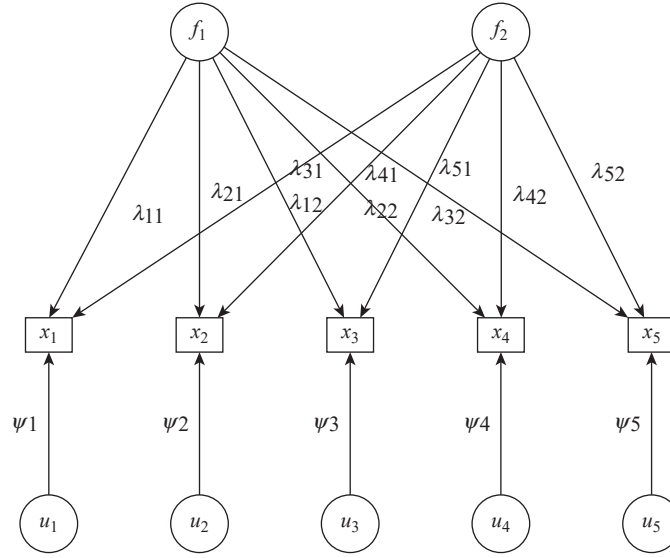


Fig. 1 Path diagram of the EFA model (1) for $p = 5$ and $r = 2$.

$$\begin{aligned} X^T X \approx R_{XX} &= (F\Lambda^T + U\Psi)^T (F\Lambda^T + U\Psi) = \\ &= \Lambda F^T F \Lambda^T + \Lambda F^T U \Psi^T + \Psi U^T F \Lambda^T + \Psi U^T U \Psi = \\ &= \Lambda F^T F \Lambda^T + \Psi^2. \end{aligned} \quad (3)$$

Clearly, $F^T F$ can be seen as the covariance matrix of the common factors. EFA assumes that the common factors F also have unit variances, i.e. have unit lengths, which makes of $F^T F$ the correlation matrix of the common factors. Most frequently, they are considered uncorrelated, i.e. $F^T F = I_r$, and F is an orthonormal matrix. Then, (3) simply reduces to:

$$X^T X \approx R_{XX} = \Lambda \Lambda^T + \Psi^2, \quad (4)$$

well known as the EFA model for the sample correlation matrix. The EFA problem is to find Λ and Ψ such that R_{XX} fits $X^T X$ as close as possible.

The EFA model (4) of the correlation structure of the data clearly shows the EFA idea: $\Lambda \Lambda^T$ fits the off-diagonal elements of $X^T X$, i.e. the sample correlations, while the unique variances Ψ^2 simply adjust the diagonal elements of $\Lambda \Lambda^T$ to fit the unit main diagonal of $X^T X$. This exhibits the different roles of the common and unique factors in EFA, expressed by $F^T U = O_{r \times p}$. We stress that EFA fits the correlations among the p input variables by means of only r factors, which makes it a dimension reduction technique.

Left multiplication of (2) by F^T gives $\Lambda = X^T F$, suggesting that λ_{ij} can be treated as correlation between the i th variable and the j th factor. Then, λ_{ij}^2 is the contribution of the j th factor to the variance of the i th variable.

Now, we can mention briefly the difference between EFA and CFA. In EFA, there are no constraints on the factor loadings Λ . In contrast, CFA may assume/speculate that some λ_{ij} take zero values, and thus impose a priori lack of relationships among certain factors and variables. Thus, EFA and CFA represent different approaches of using the same common-factor model (4). The purpose of CFA is to utilize preliminary knowledge or hypothesis.

It follows from (4), that the diagonal elements of $\Lambda \Lambda^T$ and Ψ^2 should add up to 1, which can be written as:

$$\text{diag}(X^T X) = I_p = \text{diag}(\Lambda \Lambda^T) + \text{diag}(\Psi^2) = \text{diag}(\Lambda \Lambda^T) + \Psi^2,$$

where $\text{diag}()$ makes a diagonal matrix from its argument. The entries of $\text{diag}(\Lambda \Lambda^T) = I_p - \Psi^2$ are called communalities of the original variables. They give the part of the variance of the observed variables due to just common factors. Note, that the entries of $\text{diag}(\Lambda^T \Lambda)$ give the contributions of each common factor to the total variance.

It is worth reminding that the trace of a square $p \times p$ matrix A is the sum of its diagonal elements. It is denoted by $\text{trace}(A)$ and the definition implies that $\text{trace}(A^T) = \text{trace}(A)$ and $\text{trace}(AB) = \text{trace}(BA)$. In particular, $\text{trace}(\Lambda \Lambda^T) = \text{trace}(\Lambda^T \Lambda)$. By taking trace from both sides of (4), we find another useful identity of the EFA model:

$$\text{trace}(X^T X) = p = \text{trace}(\Lambda^T \Lambda) + \text{trace}(\Psi^2),$$

indicating that the total variance of the input variables is composed by the total contribution of the common factors and the variances of the unique factors.

Number of factors

Sometimes, we want a solution with a specific or prescribed number r of factors. Then, we simply look for a r -factor EFA solution. Particularly, the choice of $r = 2$ is very popular because the resulting $p \times 2$ matrix of factor loadings Λ can be plotted on the plane.

However, in general, this is not the case: r is unknown in advance and needs to be found. There exist two very simple and popular heuristic ways to find, or at least to suggest, a suitable number of factors (Mulaik, 2010, 8.2.11). They both use the eigenvalues of the sample correlation matrix $X^T X$. The Kaiser's rule makes r equal to the number of eigenvalues greater than one. The second way is graphical and uses the so-called "elbow" or "scree" plot. The ordered in decreasing order eigenvalues are plotted (against their order) and r is made equal to the number before which the graph flattens.

In reality, their application is not so straightforward as it may look by their simplicity. For example, eigenvalues close to 1 make the Kaiser's rule inconclusive. Similarly, the "elbow" plot may flatten at two or more eigenvalues, or not flatten at all. The standard practice is to try several plausible values of r and compromise between simpler model (smaller r), or better fit of the model to the data.

We can illustrate these potential complications with the small well-known data set of Five Socio-Economic Variables (Harman, 1976, Table 2.1, p. 14). It contains $n = 12$ observations (census tracts of small areal subdivisions of Los Angeles, CA) and $p = 5$ variables, Total Population (POP), Median School Years (SCH), Total Employment (EMP), Misc. Professional Services (SER) and Median Value House (HOU). They are measured in different units and we need to standardize them and analyze their sample correlation matrix:

$$X^T X = \begin{bmatrix} 1.000 & .0098 & .9724 & .4389 & .0224 \\ .0098 & 1.000 & .1543 & .6914 & .8631 \\ .9724 & .1543 & 1.000 & .5147 & .1219 \\ .4389 & .6914 & .5147 & 1.000 & .7777 \\ .0224 & .8631 & .1219 & .7777 & 1.000 \end{bmatrix}. \quad (5)$$

The eigenvalues of $X^T X$ are 2.87, 1.80, 0.21, 0.10 and 0.02. According to the Kaiser's rule we conclude without hesitation that two factors are appropriate for these data. Moreover, they carry 93.4% $\left(= \frac{2.8733+1.7967}{5} \times 100\% \right)$ of the total variance of the data.

Now, let us look at the scree plot of these eigenvalues in Fig. 2. We see with surprise that it clearly suggests using three factors. Apparently, this is not a very good idea: the three eigenvalues explain 97.7% of the total variance, which is only 4% more than the first two, but result in a more complicated EFA model (involving three factors for five original variables).

A likely source for problems with these heuristic methods is that they are based on the eigenvalues of $X^T X$, while the common factors F are responsible for part of it only. Nowadays, the Horn's parallel method is becoming a popular alternative of the elbow/Kaiser's rule. The idea is to compare the available eigenvalues of $X^T X$ to the average eigenvalues of simulated zero-mean uncorrelated multivariate normal data of the same size. The rule is to count only the actual eigenvalues greater than the simulated ones. For our example, we find average simulated eigenvalues to be 1.91, 1.32, 0.92, 0.58 and 0.28, that is, only the first two eigenvalues of $X^T X$ are greater than the simulated ones and thus, $r = 2$. More details on how to evaluate the number r of common factors can be found in (Mulaik, 2010, 8.2.11).

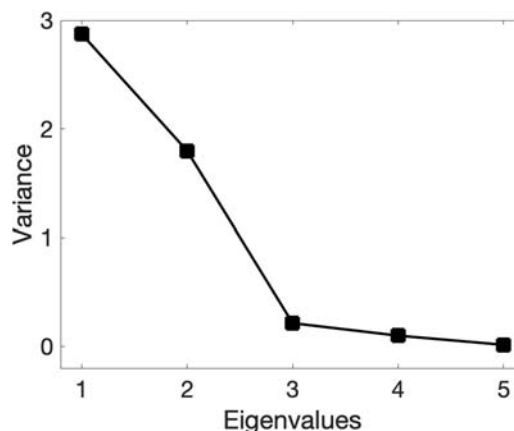


Fig. 2 Scree plot of the sample correlation matrix (5).

Two ways to find Λ and Ψ

In order to evaluate the closeness of R_{XX} to $X^\top X$, we need a discrepancy measure. Probably, the most natural way to approach the problem is to consider their difference $E = X^\top X - R_{XX}$. Then we can find Λ and Ψ such that the sum of squared errors $\sum_{ij}^p e_{ij}^2 = \text{trace}(E^\top E) = \text{trace}(E^2)$ is made as little as possible. This problem is known as the least-squares EFA (LS-EFA) and can be summarized as:

$$\min_{\Lambda, \Psi} \text{trace} \left[\left(X^\top X - \Lambda \Lambda^\top - \Psi^2 \right)^2 \right]. \quad (6)$$

If the standardized data are assumed normally distributed with zero mean and correlation matrix given by R_{XX} as in (4), then the likelihood function for $X^\top X$ can be used as a discrepancy measure. Its minimization with respect to Λ and Ψ gives the following maximum likelihood EFA problem (ML-EFA):

$$\min_{\Lambda, \Psi} \log \left(\det \left(\Lambda \Lambda^\top + \Psi^2 \right) \right) + \text{trace} \left(\left(\Lambda \Lambda^\top + \Psi^2 \right)^{-1} X^\top X \right). \quad (7)$$

It is important to stress that ML-EFA (7) can be used safely with non-normally distributed data (Mulaik, 2010, 8.3.9).

Let us apply LS-EFA and ML-EFA to the sample correlation matrix (5) of the Five Socio-Economic Variables. We solve both EFA problems (6) and (7) by using `fals` and `faml` respectively from the R package `fungible`. The resulting factor loadings Λ and unique variances Ψ^2 are given in Table 1.

The values of Ψ^2 show how well the corresponding variables are fit by the common factors. However, zero unique variances are traditionally considered unrealistic and are referred to as “Heywood cases” (Harman, 1976, 7.3). For this reason, many EFA programs (including `fungible`) do not permit unique variances less than 0.005. This phenomenon is illustrated in the ML-EFA solution in Table 1, where a Heywood case is obtained for POP.

In general, the Heywood cases are considered statistical phenomenon (Bartholomew et al., 2011), which may result from a range of factors: from simple sampling errors to wrong number r of factors. Nowadays many researches relax this restriction and admit zero unique variances.

EFA interpretation

After the EFA solution $\{\Lambda, \Psi\}$ is found by solving LS-EFA (6) or ML-EFA (7), we need to interpret it. This second stage of EFA is very important by itself and is the subject of this Section.

The EFA interpretation is mainly based on the factor loadings Λ . We already know that λ_{ij} is the correlation between the i th input variable and the j th factor. In other words, λ_{ij} expresses the importance of the j th factor in the linear approximation of the i th input variable. In this sense, the factor loadings interpretation will be simplified if their magnitudes are either near zero or near one.

Let us interpret the LS-EFA and ML-EFA factor loadings obtained in Table 1. At first glance, the two solutions are very different. Clearly, the ML loadings look more understandable. There is only one loading of medium magnitude. All the rest are either very small (< 0.1), or large, with three near unity. In contrast, almost all LS loadings are of rather medium magnitude, and thus difficult to interpret. A possible way out, at least for small data sets, is to plot and inspect the factor loadings for some interesting patterns. We plot these LS and ML solutions respectively on the left and the right panels of Fig. 3. It is easy to see that both solutions form very similar configurations of points on the plane. This indicates that, in fact, the LS- and ML-EFA solutions are very close. We will see in the next sections that this can be explained in formal terms.

Table 1 Two EFA solutions for Five Socio-Economic Variables.

VARS	LS-EFA			ML-EFA		
	Λ		Ψ^2	Λ		Ψ^2
POP	0.622	−0.776	0.010	−0.026	0.997	0.005
SCH	0.715	0.549	0.187	0.897	0.038	0.194
EMP	0.710	−0.686	0.026	0.091	0.978	0.036
SER	0.882	0.149	0.200	0.777	0.460	0.185
HOU	0.763	0.584	0.077	0.961	0.046	0.074

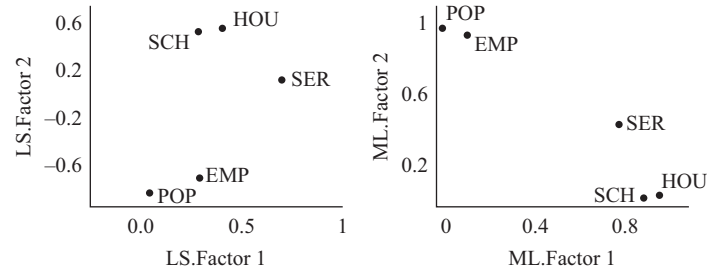


Fig. 3 Plots of factor loadings from two EFA methods.

Simple structure concept

The EFA model (2) is not unique. Indeed, it can be rewritten as:

$$X \approx F\Lambda^T + U\Psi = FQQ^{-1}\Lambda^T + U\Psi = (FQ)(\Lambda Q^{-T})^T + U\Psi,$$

where Q is any $r \times r$ non-singular matrix, i.e. $\det(Q) \neq 0$ and $Q^{-T} \equiv (Q^{-1})^T$. Now, we can recalculate the EFA correlation model (3) in the new notations as follows:

$$\begin{aligned} X^T X \approx R_{XX} &= (FQQ^{-1}\Lambda^T + U\Psi)^T (FQQ^{-1}\Lambda^T + U\Psi) = \\ &= \Lambda Q^{-T} Q^T Q Q^{-1} \Lambda^T + \Psi^2 = \\ &= (\Lambda Q^{-T})(Q^T Q)(\Lambda Q^{-T})^T + \Psi^2, \end{aligned}$$

where ΛQ^{-T} is the new transformed factor loading matrix and $Q^T Q$ is the covariance matrix of the transformed common factors FQ .

The most frequent EFA choice for Q is an $r \times r$ orthogonal matrix, that is, $Q^T Q = Q Q^T = I_r$ and $Q^{-1} = Q^T$. Such matrices express either rotation or reflection, but in both cases they preserve lengths and angles. Note that for orthogonal Q , the transformed common factors FQ remain uncorrelated.

Another popular EFA choice are the $r \times r$ non-singular matrices Q for which $Q^T Q$ is a correlation matrix, i.e. $\text{diag}(Q^T Q) = I_r$. Such matrices are called obliques, because the transformed common factors FQ are not mutually orthogonal. In EFA, such Q s are called oblique rotations.

It is a long tradition in EFA to benefit from this lack of uniqueness of its model. It opens the opportunity to transform the initial loadings Λ into new ones, which are possibly easier to interpret, by making use of orthogonal and/or oblique rotations. However, to do this in an objective manner, we need to define the meaning of “easier to interpret”.

For this reason, the concept of simple structure was introduced in EFA (Thurstone, 1947, p. 335). In more contemporary language they look as follows (Harman, 1976, p. 98):

1. Each row of the factor matrix should have at least one zero,
2. If there are r common factors each column of the factor matrix should have at least r zeros,
3. For every pair of columns of the factor matrix there should be several variables whose entries vanish in one column but not in the other,
4. For every pair of columns of the factor matrix, a large proportion of the variables should have vanishing entries in both columns when there are four or more factors,
5. For every pair of columns of the factor matrix there should be only a small number of variables with non-vanishing entries in both columns.

It makes sense to remind here that the mathematical meaning of vanish is “to become zero” (Webster’s II, The Riverside Publishing Company, 1984). In modern jargon, the simple structure concept, in fact, requires sparse loadings. Broadly speaking the vector $a \in \mathbb{R}^p$ is called sparse if the number of its non-zero entries is very small compared to p . Unfortunately, such a clear-cut sparseness can not be achieved by the rotation methods. The best they can do is to make the loadings’ magnitudes more distinguishable.

Simple structure criteria and rotation methods

Let us take a closer look at how the simple structure concept is implemented in rotation methods. To make the interpretation process objective, analytic rotation methods were introduced (Thurstone, 1947, XVII). At the heart of these methods is the simplicity criterion, which is supposed to express the Thurstone’s five rules in a single mathematical formula. Unfortunately, it turns out that this is a very difficult task. As a result there exists a huge number of criteria, because none of them is capable to produce satisfying solution for every problem.

Table 2 Two VARIMAX rotated loadings and the orthogonal matrices used to rotate the corresponding initial factor loadings from Table 1.

VARS	LS-EFA		ML-EFA	
	I	II	I	II
POP	0.020	−0.995	0.021	0.997
SCH	0.902	0.001	0.898	0.004
EMP	0.144	−0.976	0.137	0.972
SER	0.790	−0.420	0.798	0.423
HOU	0.961	−0.002	0.962	0.001
Rotation	0.793	−0.610	0.999	−0.047
Matrix <i>Q</i>	0.610	0.793	0.047	0.999

Arguably the most popular rotation criterion is VARIMAX (Mulaik, 2010, p. 310). It produces satisfactory results in many applications and is based on a simple to explain concept. Let $\Xi = \Lambda Q$ be the $p \times r$ matrix of orthogonally rotated factor loadings. The variance of the squared loadings of the j th rotated column ξ_j is:

$$V_j = V(\xi_j^2) = \sum_{i=1}^p \xi_{ij}^4 - \frac{1}{p} \left(\sum_{i=1}^p \xi_{ij}^2 \right)^2, \quad j = 1, \dots, r. \quad (8)$$

The variance V_j will be large if there are few large squared loadings and all the rest are near zero. The variance V_j will be small when all squared loadings have similar magnitudes. The VARIMAX rotation problem is to find an $r \times r$ orthogonal matrix Q such that the total variance $V = \sum_{j=1}^r V_j$ is maximized.

We apply varimax from fungible to rotate the LS-EFA and the ML-EFA factor loadings depicted in Table 1. The resulting VARIMAX loadings are given in Table 2. The rotated loading are nearly identical, which formally confirms our conclusion after observing the plots of the initial loadings in Fig. 3. We note that the orthogonal matrix Q that rotates the ML initial loadings is nearly the identity I_2 , that is, the initial ML loadings nearly maximize the VARIMAX criterion and need very “little” rotation.

Rotation to simple structure target

There are a number of alternative rotation methods which do not rely on any particular simplicity criterion. Instead, they “refine” some initial or rotated solution Λ . They construct a target matrix $\mathcal{Y}$ reflecting the simple structure of the problem, based on the already available loadings Λ . Then, transformed factor loadings are obtained by orthogonal or oblique rotation of Λ such that to match the target $\mathcal{Y}$ as close as possible.

PROMAX is probably the most popular method of this kind and constructs a target $\mathcal{Y}$ with elements $v_{ij} = \lambda_{ij}^5 / |\lambda_{ij}|$ (Mulaik, 2010, p. 342–347). As $|\lambda_{ij}| \leq 1$, the magnitudes of the small loading become smaller, while the larger ones remain relatively large.

The VARIMAX loadings found in Table 2 show that both EFA solutions agree in their interpretation, which is also quite unambiguous. In this sense, looking for their “refinement” is not really necessary, but we are interested to see PROMAX in action. We apply promaxQ from fungible with the suggested “fourth power” target. What happens is that first $\mathcal{Y}$ is constructed from the VARIMAX loadings from Table 2. Then, an oblique rotation is found to transform and make them as much like the target $\mathcal{Y}$ as possible. The resulting “refined” factor loadings are given in Table 3.

The PROMAX factor loadings from the two EFA methods are practically identical. The transformed common factors are correlated, and the correlations are also very similar.

Table 3 Two PROMAX factor loadings refining Table 2 and the correlations between the LS and ML oblique common factors.

VARS	LS-EFA		ML-EFA	
	I	II	I	II
POP	−0.110	1.02	−0.104	1.02
SCH	0.924	−0.117	0.920	−0.121
EMP	0.020	0.982	0.018	0.977
SER	0.754	0.327	0.764	0.329
HOU	0.984	−0.123	0.985	−0.124
Correlations between common factors	0.251		0.246	

Concluding remarks

The article introduces the EFA model and its main aspects. The EFA of a small data set gives a flavor and illustrates its application to a real problem. In reality, before doing EFA one is likely to face a preprocessing stage and deal with missing data and/or outliers. They affect considerably the sample correlation matrix, and thus, invalidate the EFA results.

It is understandable, that there is a great number of theoretical and applied EFA details that are left untouched in this article. The provided references can be starting point for a deeper study of EFA. In addition, many specialized issues can only be found in research papers. For example, we already know that EFA is derived and works with continuous data. There are ways to adapt EFA for data from questionnaires with dichotomous or discrete variables, or even multi-level data. Of course, such issues cannot be accommodated in this introductory article.

References

- Bartholomew, D., Knott, M., Moustaki, I., 2011. *Latent Variable Models and Factor Analysis: A Unified Approach*, third ed. John Wiley & Sons Ltd, Chichester, UK.
- Harman, H.H., 1976. *Modern Factor Analysis*, third ed. University of Chicago Press, Chicago, IL.
- Mulaik, S.A., 2010. *The Foundations of Factor Analysis*, second ed. Chapman and Hall/CRC, Boca Raton, FL.
- Spearman, C., 1904. "General intelligence", objectively determined and measured. *Am. J. Psychol.* 15, 201–292.
- Thurstone, L.L., 1947. *Multiple Factor Analysis*. University of Chicago Press, Chicago, IL.

Confirmatory factor analysis: foundations and extensions

Larry R. Price, Department of Methodology, Measurement and Statistical Analysis, Psychometrics and Statistics, Texas State University, San Marcos, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	607
Data considerations	607
Sample size and power	607
The common factor model	608
Multilevel confirmatory factor analysis	609
Second-order factor analysis	611
Confirmatory bifactor analysis	613
Bayesian confirmatory factor analysis	615
Summary and conclusions	617
References	617

Introduction

Confirmatory factor analysis (CFA) is a technique used to analyze the efficacy of measurement models where the number of factors and their direct relationship is specified. Conducting CFA requires an *a priori* strategy where a hypothesized factor structure is assessed by evaluation of model fit to empirical (or simulated) data. Application of CFA involves confirming that variables (e.g., items or total test scores) fit a particular pattern or cluster to form a certain dimension according to theory (McDonald, 1999; Fabrigar and Wegner, 2012). Also, researchers synthesize information about the factors and their contribution reflected by examinee performance on the observed variables (e.g., scores on tests or other measurement acquisition formats). Researchers using CFA attempt to understand *why* the variables are correlated and the *degree* or level of accuracy variables and factors provide relative to a theory (Price, 2016).

Application of confirmatory factor analysis (CFA) involves defining a latent variable of interest-based on theory or acquired knowledge. Researchers then construct observed variables (e.g., test items or other variables) to measure continuous latent variables (a.k.a. factors). The selection of a scaling model and variable construction follows sound principles based on rules of correspondence in consideration of properties and levels of measurement (Torgerson, 1958; Krantz et al., 1971). Confirmatory factor analysis is a special case of the common factor model. This model postulates that total scores or test items correlate because they are determined in part from common but unobserved influences. These common influences are due to common factors meaning that variables are correlated to some degree. The unobserved influences manifest as a latent factor (or simply a *factor*) in CFA.

Traditionally, conducting CFA involved developing a measurement model or models within the tradition of structural equation modeling (SEM) to provide validity evidence of scores on tests or other measurement instruments or protocols (Floyd and Widaman, 1995). During the past 30 years, the confirmatory analytic framework has evolved substantially. It includes (a) multilevel CFA (Heck and Thomas, 2015), (b) second and higher-order CFA, (c) bifactor CFA, (c) Bayesian CFA (Kaplan, 2014), (d) longitudinal CFA (Little, 2013; McCardle, 2007), and (g) dynamic CFA of intra-individual change Browne and Zhang (2007) and Molenaar (2004). This article provides an overview of some of these approaches and recent developments in the CFA framework.

Data considerations

Initially, factor analysis was suited for variables scaled on an interval level of measurement and that the variable distributions were assumed to adhere to linear model assumptions (e.g., normality and linearity). However, item-level factor models that include items measured at the binary or ordinal level are now commonly estimated (Flora and Curran, 2004; Muthén, 1984; Muthén and Kaplan, 1985), as are models with nonnormally distributed data.

Sample size and power

As previously noted, CFA is conducted using structural equation modeling (SEM). Planning for an adequate level of statistical power in any study using SEM is critical. Frequently, the use of SEM involves testing or confirming hypotheses about relationships in measurement models (e.g., in test or instrument development, Price, 2016) and among structural relations in a full SEM (Wang and Wang, 2012). Planning for adequate statistical power in a CFA is crucial because it is a hypothesis-driven approach targeting

confirmation of an apriori hypothesis or hypotheses. Specifically, sufficient statistical power in CFA is necessary to ensure that the precision and stability of standard errors and parameter estimates are accurate and do not contribute to inflated Type I and II error rates of hypothesis tests (Barker et al., 2015). Like other inferential statistical methods, estimating statistical power apriori involves determining the minimum required sample size to achieve the maximum level of power for a hypothesized effect size under a specified level of statistical significance (Wang et al., 2013). Estimation of statistical power in CFA (and SEM generally) depends on (a) if the study design is cross-sectional or longitudinal, (b) the number of relationships among test items or variables, (c) reliability of scores based on test items or variables, (d) level of measurement of indicator variables—continuous, ordered categorical or dichotomous and the statistical estimator used (e.g., ML, WLSMV, Robust ML, etc.), (e) magnitude of the factor loadings, (f) pattern and amount of missing data, (g) whether indicator variables are individual items or composites or aggregates of items (e.g., sum scores, average scores, parcels or testlets), and (h) model complexity. Although SEM (and CFA conducted with SEM) is a large sample multivariate method, excessively large samples are costly to acquire and often overstate negligible effects. The challenge in planning for adequate power is to strike a balance between obtaining a sample that is too small to detect essential or hypothesized effects and one that is too large, resulting in wasted time, energy, and resources. However, when a large sample is available, researchers can conduct cross-validation of the CFA thereby providing greater evidence of the tenability of the model (Price et al., 2002; Tulskey and Price, 2003). For guidelines on sample size requirements in CFA see Brown (2015), Wolf et al. (2013), MacCallum et al. (1999), Marsh and Hau (1999), Boomsma (1985). Muthén and Muthén (2002) provide an example of how to conduct power analysis using Markov chain Monte Carlo (MCMC) approach. The MCMC simulation approach to power estimation for model parameters is particularly useful in multilevel and complex factor and structural equation models.

The common factor model

Factor analysis is an essential technique for investigating relationships between constructs (i.e., unobservable underlying attributes or processes) and observed indicators (e.g., test items or other observed measures). The common factor model provides a mathematical framework linking observations of an underlying process to the theory that provides understanding and synthesis. A common factor is a latent variable or factor that influences more than one observed variable. For example, Y1–Y3 are influenced by the η_1 latent variable or factor, and Y4–Y6 are influenced by the η_2 latent variable or factor. The common factor model in Fig. 1 is a two-factor oblique model because the latent variables or factors are correlated (e.g., a double-headed arrow connecting the two factors exists).

Using factor analytic techniques such as CFA, researchers can evaluate how latent or unobservable constructs such as intelligence, achievement, aggression, depression, leadership style, personality traits, etc. affect the daily lives of individuals in various settings. In structural equation modeling, a measurement model is a confirmatory factor analysis model that specifies a set of theoretically informed relationships between a latent variable or factor and observable indicators reflecting intricate, unique information about that factor. Fig. 1 depicts an oblique two-factor CFA.

The general form for the measurement model in a CFA is expressed in Eq.(1) (Heck and Thomas, 2015) as

$$y_i = v + \Lambda \eta_i + e_i \quad (1)$$

In Eq.(1), persons or subjects are indexed by the subscript i , with y_i representing a p -dimensional vector of observed dependent variables, v is a p -dimensional vector of measurement intercepts, η (eta) is an m -dimensional vector of latent variables, Λ (lambda) is $p \times m$ matrix of factor loadings, and e_i (epsilon) is a $p \times 1$ vector of measurement errors (Raykov and Marcoulides, 2006; Heck and Thomas, 2015). Measurement errors are uncorrelated with factors but can correlate with other errors of measurement based on research design or measurement protocols. However, models with extreme flexibility in measurement error may not be identified statistically. Assuming $E(e_i) = 0$, the mean structure is defined as

$$\mu = v + \Lambda \alpha \quad (2)$$

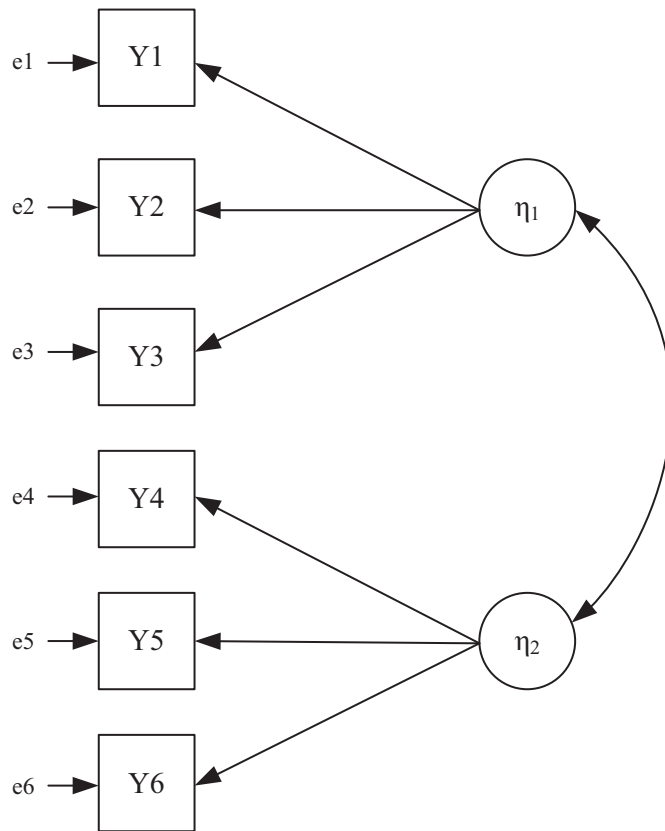
In Eq.(2), note the addition of α , which is a vector of factor means of dimension m .

The covariance structure is

$$COV(y) = \Lambda \Psi \Lambda' + \Theta \quad (3)$$

In Eq.(2), Ψ (psi) is a matrix of factor variance and covariances, and Θ is a $p \times p$ covariance matrix of normally distributed residuals with mean zero and variance of some magnitude (Heck and Thomas, 2015).

The goal of CFA is to reproduce the relationships among indicators with as few parameters required for the estimation process as possible. Confirmatory factor analysis works by imposing theory-driven constraints into the common factor model. Specifically, in CFA, develops and tests a model by prespecifying the number of factors, the pattern of loadings and appropriate error structure (e.g., allowing correlated indicator errors due to a measurement-driven method effect). For example, in Fig. 1, indicators Y1–Y3 load only on factor 1 (η_1) and Y4–Y6 load only on factor 2 (η_2), no cross-loadings exist (i.e., the restriction in the CFA is that cross-loadings are fixed at a value of zero). The simple structure in Fig. 1 is achieved by the specification of each factor loading on a unique set of indicators. In some parameterizations (depending on the specific program used) factor correlations are estimated by default in models with two or more factors. Parameter estimation proceeds by evaluating the sample covariance matrix (S) relative to the model implied population covariance $\Sigma(\theta)$ matrix using $F[S, \Sigma(\theta)]$ yielding a minimum fit function (F_{min}) that indexes the location



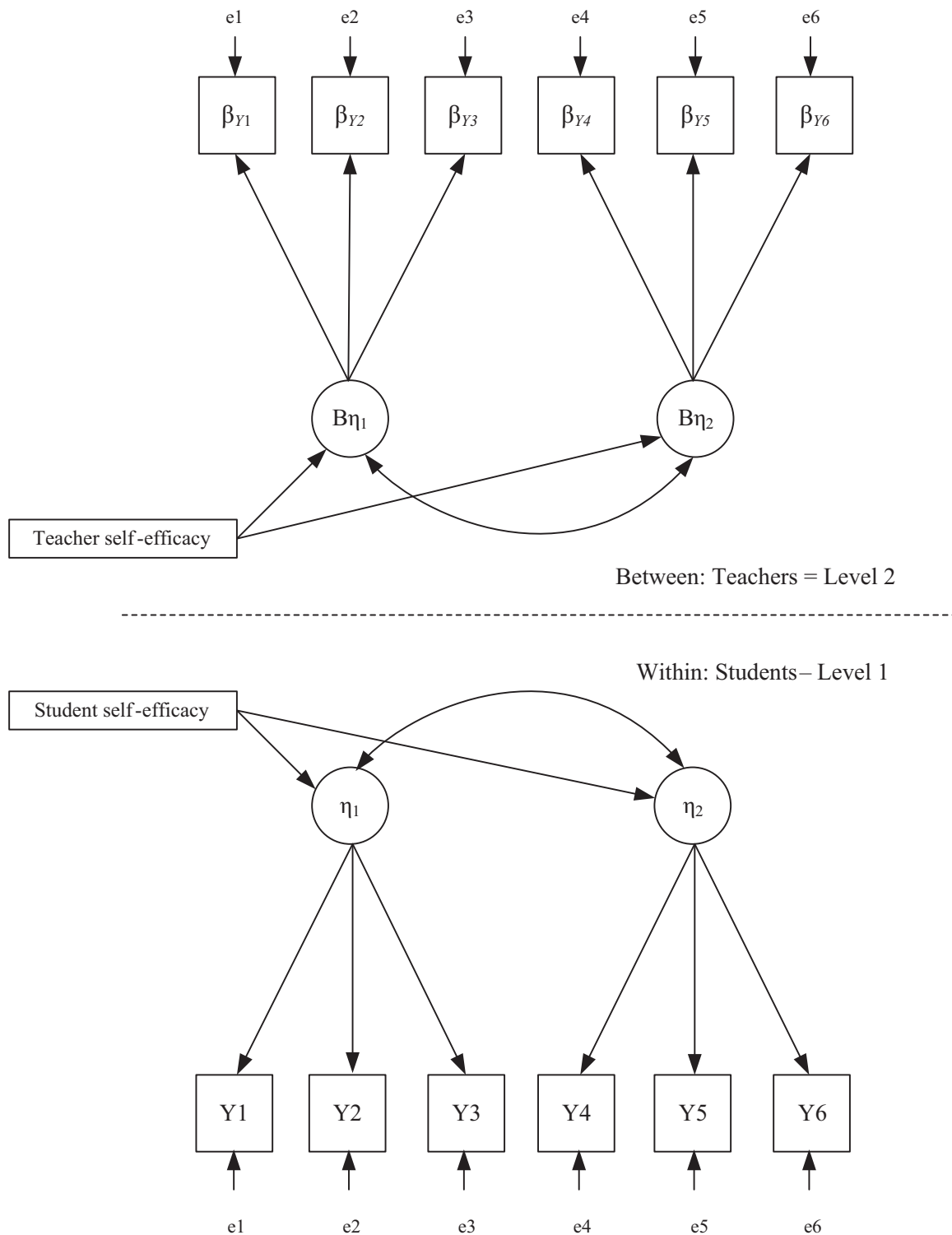
Note. A common factor is a factor that influences more than one observed variable. For example, Y1 – Y3 are influenced by the η_1 latent variable or factor and Y4 – Y6 are influenced by the η_2 latent variable or factor. The common factor analysis model above is oblique because the latent variables or factors are correlated (e.g. a double-headed arrow connecting the two factors is present).

Fig. 1 Two-factor model.

in the estimation process where the difference between S and $\Sigma(\theta)$ is the smallest. However, most often estimation proceeds using the method of maximum likelihood (ML), where the likelihood is induced by parametric assumptions of the latent variables and residuals. In the case where indicator variables are nonnormally distributed, alternative estimators are available (e.g., see [Muthén and Muthén, 2007–2010](#)) as are estimators that produce rescaled chi-square fit statistics. After estimation of model parameters, the tenability of results assessed by interpretation of adequacy of model-data fit using model fit statistics such as overall chi-square test, comparative fit index (CFI), and the Root Mean Square Error of Approximation (RMSEA) and associated levels of confidence. Additional fit statistics are available and produced by SEM programs and may be useful depending on the research goal and model specification. For guidance on use of fit statistics post estimation see [Brown \(2015\)](#).

Multilevel confirmatory factor analysis

In standard CFA, observations or measurements are structured at a single level with the assumption that they are statistically independent and identically distributed. However, in many instances, research calls for conducting CFA in settings where individuals are clustered or nested within organizations (e.g., schools, businesses, healthcare settings, or other organizations). Also, measurements on individuals are often nested within individuals either cross-sectionally or longitudinally. Multilevel CFA (ML-CFA) provides a multilevel measurement modeling framework that accounts for clustering or nested structures in CFA resulting in, for example, analyses of construct validity and measurement invariance at different levels of measurement (i.e., at level-1 vs. level-2). An essential use of multilevel modeling is to determine whether the effects at level-1 (micro-level) are the same as those at level-2 (macro-level; [Sniders and Bosker, 1999](#)). For example, is it correct or reasonable to assume the factor structure holds equally well (i.e., is invariant) across different schools, businesses, healthcare settings, etc.? If the answer is no, to what degree do the level-2 units vary? When parameter estimates of the CFA model are restricted to be the same (or have little variation) at both levels-1 and 2, the result is a fixed-effects model with no need to estimate and interpret variation across level 2 units or clusters. If the parameter estimates level-1 are allowed to vary across level-2, the result is a random-effects or random coefficients model. The ML-CFA model also allows for the inclusion of covariates at level-1 and level-2 to investigate hypothesized effects that may influence the parameter estimates at either level (e.g., see [Fig. 2](#)). Extending the initial CFA in [Fig. 1](#), consider a two-factor CFA model that measures the level of writing



Note. The symbols $B\eta_1$ and $B\eta_2$ in the upper half of the figure signify the group-level factors. The bottom half of the figure represents the within-subjects factor model. The Y variables remain consistent across levels. In the within model, solid circles on arrows are random slopes and on Y variables are random intercepts. In the between model, random slopes are depicted as circles or latent variables varying across teachers or clusters. $\beta_{Y1} \dots \beta_{Y6}$ at the level-2 represent intercepts.

Fig. 2 Multilevel confirmatory factor analysis model with covariates.

fluency (Factor 1: Y1–Y3 are measures of fluency) and essay composition (Factor 2: Y4–Y6 of writing composition) in middle school for student with disabilities as the outcomes at level-1. Evidence from the literature on writing performance reveals that students' level of self-efficacy influences their writing performance. Therefore, student self-efficacy is included at level-1 as a covariate to test the within-subjects effect of self-efficacy as a predictor of writing fluency and composition. At level-2 (between-clusters or teachers), teachers' self-efficacy related to teaching writing serves as a covariate. The research questions are: Does student writing performance in fluency and essay composition vary across teacher classrooms? Is teacher level of self-efficacy (level-2 covariate) a significant predictor the student writing performance in fluency and essay composition? Is student level of self-efficacy (level-1 covariate) a significant predictor of student writing performance in fluency and essay composition?

Muthén (1991, 1994) provided a general framework for the analysis of multilevel covariance structures for SEM. Following Muthén's framework, analysis of the ML-CFA in Fig. 2 involves analysis of a total variance-covariance matrix representing a single population that can be disaggregated into within-person and between-groups matrices reflecting levels-1 and 2 of the hierarchy (see Fig. 2). Total scores for persons at level-1 capture the variation within-persons (i.e., using the deviation of each person's score from the group mean) and a between-group matrix that captures group-level mean composition. In the level-1 or within-part of the model in Fig. 2, the regression of the two factors on the student-level covariate self-efficacy captures variation in the random slopes for students (solid dots marked as S1 and S2). The level-1 random intercepts (displayed as solid dots on the Y variables) capture how means on the variables randomly vary across students. In the level-2 (between part of the model), the regression of the two factors on the teacher-level covariate self-efficacy captures the random variation across teachers (clusters) for each factor. The within and between matrices are orthogonal and additive as in other linear modeling frameworks (e.g., linear mixed-models). Eqs.(4)–(9) follow (Heck and Thomas, 2015; Muthén, 1991, 1994). The ML-CFA overall population equation for person i in cluster j is

$$\text{COV}(y_{ij}) = \Sigma_T = \Sigma_B + \Sigma_W \quad (4)$$

The subscripts T , B , and W represent the total, between-cluster, and within-group covariance matrices. Decomposition of the total matrix is expressed as the between-cluster covariance matrix is

$$\Sigma_B = A_B \Psi_B A_B' + \Theta_B \quad (5)$$

the within-group covariance matrix is

$$\Sigma_W = A_W \Psi_W A_W' + \Theta_W \quad (6)$$

Finally, factor means are modeled at the cluster level as

$$\mu_j = A_j \alpha \quad (7)$$

Returning to Fig. 2, I provide an overview of how the ML-CFA analysis proceeds. The multilevel measurement model in Fig. 2 is

$$y_{ij} = v_j + A \eta_{ij} + \varepsilon_{ij} \quad (8)$$

In Eq.(8), y_{ij} represents the vector of variables Y1–Y6, v_j $p \times 1$ vector of intercepts, v is a p -dimensional vector of measurement intercepts, A (lambda) is $p \times m$ matrix of factor loadings, η_{ij} (eta) is an m -dimensional vector of latent factors, and ε_i (epsilon) is a $p \times 1$ vector of residuals errors (Raykov and Marcoulides, 2006; Heck and Thomas, 2015). Independence in observations is assumed at the cluster or teacher level—not the level of individual persons.

Latent factor means randomly vary across teachers as

$$\eta_{ij} = \alpha + \eta_{Bj} + \eta_{Wij} \quad (9)$$

where α is the grand mean for all η_{ij} , η_{Bj} is a randomly varying factor component capturing teacher effects, and η_{Wij} is a random factor for students that varies within teachers. The structure of this framework provides a partition of the variability of students' scores into (a) between teachers (capturing the teacher effect) on scores or outcomes. Students' scores are group (teacher-centered), rendering them uncorrelated. Finally, the between-teachers and within students' covariance components are independent like the random effects analysis of variance model providing a way to estimate the random effect of teachers' conditional on student i residing in a specific teacher classroom j . For details on ML-CFA model specification (including program coding in Mplus), estimation and interpretation of results, see Heck and Thomas (2015). The ML-CFA can also be extended to longitudinal models. For details on model specification, estimation, and interpretation for the longitudinal ML-CFA see Little (2013).

Second-order factor analysis

In contrast to the two-factor oblique model in Fig. 1, a second-order CFA provides an alternative way to evaluate the factor structure where multiple factors exist. A second-order factor model is appropriate when first-order factors are explained by a higher-order structure (Schumacker and Lomax, 2010, p. 271). In keeping with CFA tradition, the application of a second-order structure is informed by theory to justify its use. Recall that in a single level CFA, including only first-order factors, the focus is on the measurement model that maps the items or indicator variables onto a hypothesized latent or factor structure. Provided a first-order structure

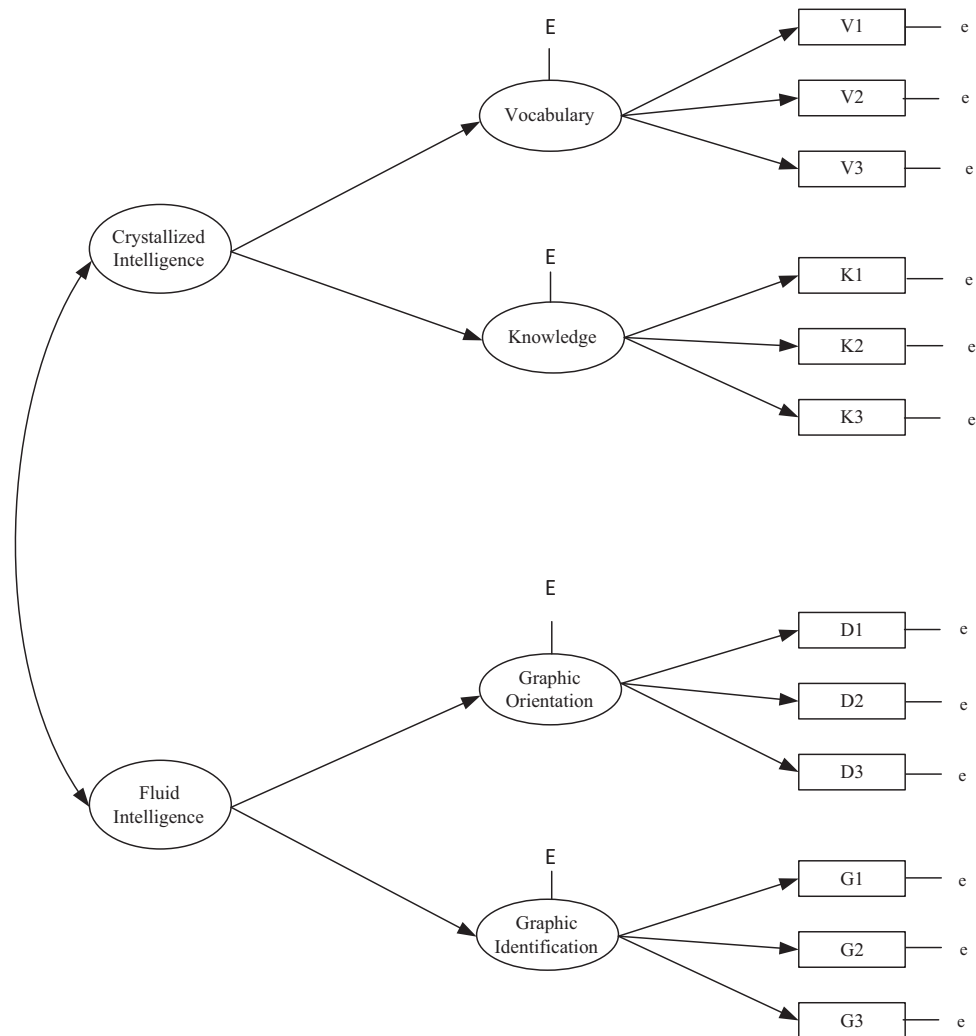


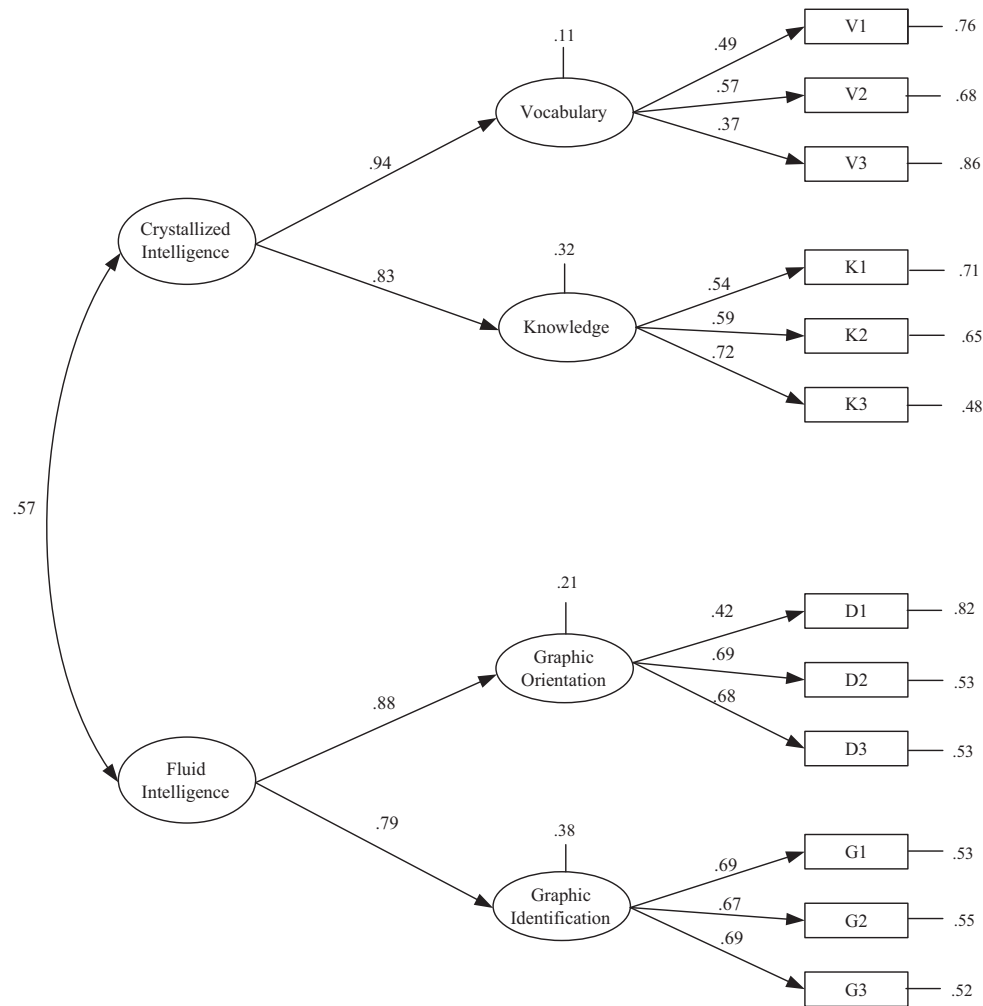
Fig. 3 Second-order CFA.

is tenable, in a second-order CFA, the focus is on the intercorrelations and/or linear relations among the factors rather than the item or variable relationships with the factors.

An example of a second-order CFA is provided in Figs. 3 and 4. Factors and indicator variables in the model are based on a small subset of data based on the general theory of intelligence (Gf/Gc) presented in Price (2016). In Fig. 3, the four first-order factors are Vocabulary, Knowledge, Graphic Orientation, and Graphic Identification. These factors are latent indicators of second-order factors of Crystallized Intelligence and Fluid Intelligence. The two second-order factors are predicted to account for the correlations among the first-order factors. According to the general theory of intelligence, Crystallized and Fluid Intelligence are posited to be distinct yet positively correlated parts of intelligence. The tenability of the second-order CFA should be substantiated by the pattern of correlations among the first-order factors. For example, relatively high, positive correlations are expected between the first-order factors comprising Crystallized Intelligence (Vocabulary and Knowledge) and Fluid Intelligence (Graphic Orientation, and Graphic Identification). However, only moderate, positive associations are expected between the first-order factors Vocabulary and Knowledge when correlated with Graphic Orientation and Graphic Identification.

The second-order CFA in Fig. 3 includes the following assumptions. First, first-order factors, Vocabulary, Knowledge, Graphic Orientation, and Graphic Identification are continuous latent variables or factors. Intercepts and residual variances (e) of the indicator variables comprising the first-order factors are freely estimated as are the residual variances (a.k.a. disturbances) of the first-order factors (E). Error variances of indicator variables and first-order factors are uncorrelated. Finally, second-order factors freely correlate based on empirical and theoretical considerations.

The general sequence for conducting a second-order (or higher-order) CFA includes (1) developing a conceptually valid, well-fitting first-order model, (2) examining the pattern of correlations among first-order factors, (3) fit the second-order model based on empirical and theoretical grounds (Brown, 2015, p. 290). Review of results of the analysis with the example data for a four-factor oblique CFA provide evidence that the model fit is acceptable (e.g., $\chi^2(66) = 108.59, p < 0.001$, CFI = 0.97; RMSEA = 0.04; 90%



Chi-square = 109.50, df = 49, $p = 0.00$; Comparative Fit Index = .97; Root Mean Square Error of Approximation = .04 (90% C.I.: .03 - .04).

Fig. 4 Second-order CFA standardized solution.

C.I.: 0.03–0.04). Furthermore, all factor loadings on the indicators display reasonable magnitudes. The pattern of the correlations among the four first-order factors (not shown in Fig. 4) supports the tenability of proceeding with fitting a second-order CFA. For example, the correlation between Knowledge and Vocabulary is 0.78. However, the correlations between Vocabulary and Graphic Orientation and Graphic Identification are moderate at 0.49 and 0.41, respectively.

Similarly, the correlations between Knowledge and Graphic Orientation and Graphic Identification are moderate at 0.41 and 0.39 respectively. Importantly, the observed patterns of correlations among the first-order factors generally align with the *GfGc* theory of intelligence (i.e., if the first-order correlations among factors were all equal or approximately so, there would be no need for a second-order CFA). Finally, the Schmid-Leiman (1957) transformation can be applied to the standardized solution in a second-order CFA if a researcher is interested in the relationship on the second-order factor(s) to observed indicator variables. For example, the effect of Crystallized Intelligence (second-order factor) on observed variable V1 is obtained as $0.94(0.49) = 0.46$. The value of 0.46 is interpreted as a total effect or as a mediated effect (e.g., Vocabulary mediates the relationship between Crystallized Intelligence and V1). An excellent treatment on the application and interpretation of the Schmid-Leiman transformation in higher-order CFA is provided in Brown (2015).

Confirmatory bifactor analysis

The bifactor model (Reise, 2012; Chen et al., 2006; Harman, 1976) provides another way to examine CFA models with second- or higher-order structures. In a bifactor model, a general factor exists that accounts for significant covariance in observed indicator

variables. Also, specific domains exist that account for unique variance in the indicators of a particular domain over and above the general factor. [Reise \(2012, p. 668\)](#) notes that bifactor models are well-suited for modeling construct-relevant multidimensionality that arises in responses to measures of broad constructs where multiple distinct domains of item content are included to increase evidence of content-related validity. To this end, bifactor models provide a cogent framework measure construction and conceptualization and for the psychometric evaluation of measures existing in multidimensional models. Importantly, there is a connection between the bifactor and second-order CFA. The second-order model in [Figs. 3 and 4](#) is nested within the bifactor model, a characteristic of the model that provides a way to compare a bifactor baseline model to a second-order CFA. To illustrate the bifactor model, I reconfigure the second-order CFA in [Figs. 3 and 4](#) as depicted in [Fig. 5](#).

[Fig. 6](#) provides the standardized solution to the bifactor model in [Fig. 5](#). The fit of the bifactor model is worse than that of the second-order model as evidenced by a smaller CFI and larger RMSEA. Additionally, the difference between the chi-square model fit statistics for the two models is 67.76 (177.26, bifactor, and 109.50, second-order) with a difference in degrees of freedom of 7 (statistically significant at $p < 0.01$). Also, in the bifactor model, the variance of the Graphic Orientation factor is insignificant and factor loadings are small in magnitude, indicating that when adjusting for general intelligence, Graphic Orientation does not add significant unique variance to the solution. As a next step, a researcher can eliminate the Graphic Orientation domain factor and rerun the analysis with a respecified bifactor model. This decision to take the step of specifying the model without Graphic Orientation versus using the second-order CFA results depends on the goal(s) of the researcher. However, it is not always easy to determine whether to opt for a second-order model or a bifactor model. For information and guidance on this topic, see [Mansolf and Reise \(2017\)](#). Factor-model based reliability is estimated for bifactor models using OmegaH and OmegaS coefficients ([Reise et al., 2012; McDonald, 1999](#)). In summary, using a nested modeling approach (where the bifactor model is nested within the second-

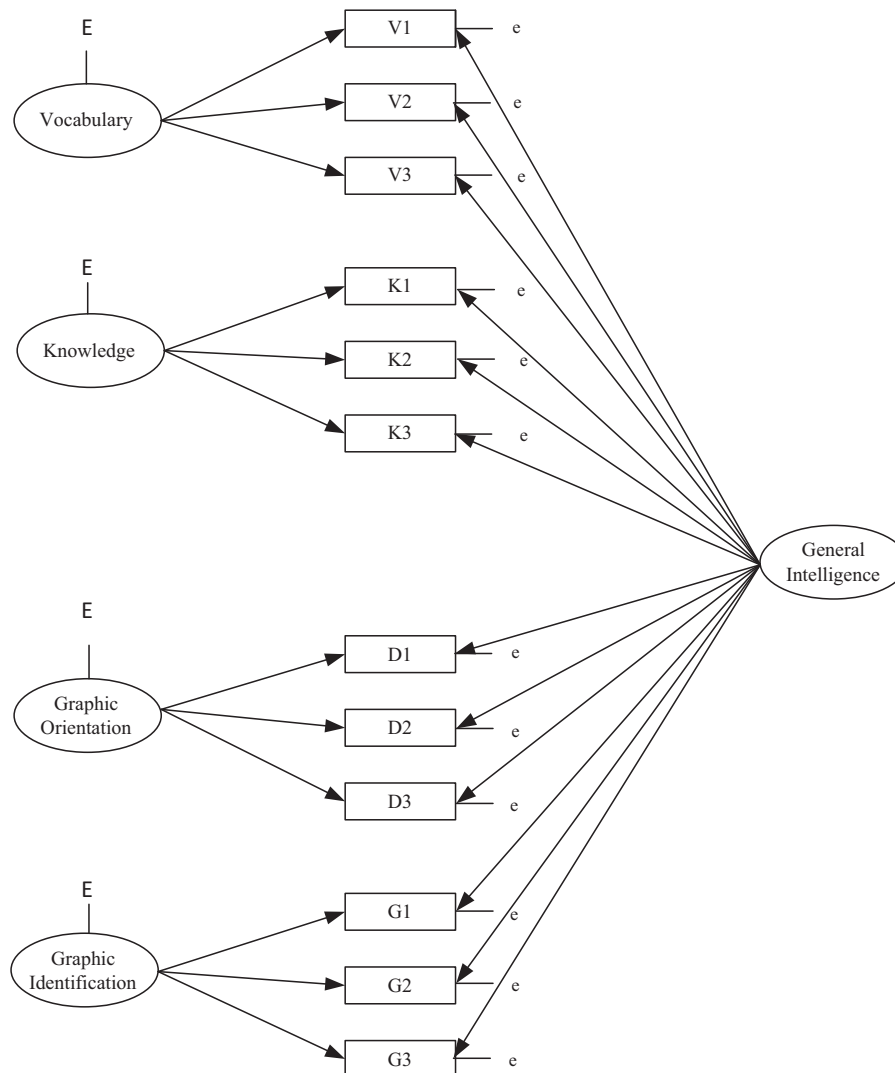
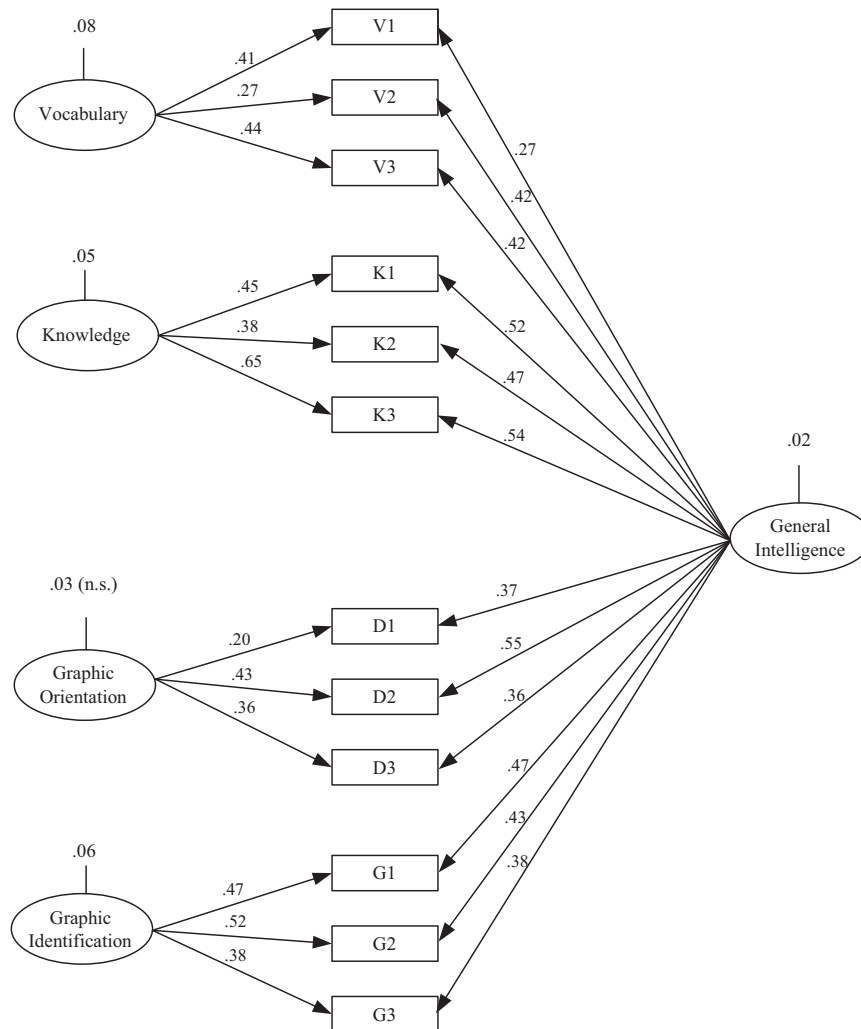


Fig. 5 Bifactor model.



Chi-square = 177.26, $df=42$, $p<.001$; Comparative Fit Index = .94; Root Mean Square Error of Approximation = .06 (90% C.I.: .05-.07)

Fig. 6 Bifactor model—standardized solution.

order model) to investigate the structure of multidimensional scales provides a comprehensive way to investigate the factor structure of multifactor scales.

Bayesian confirmatory factor analysis

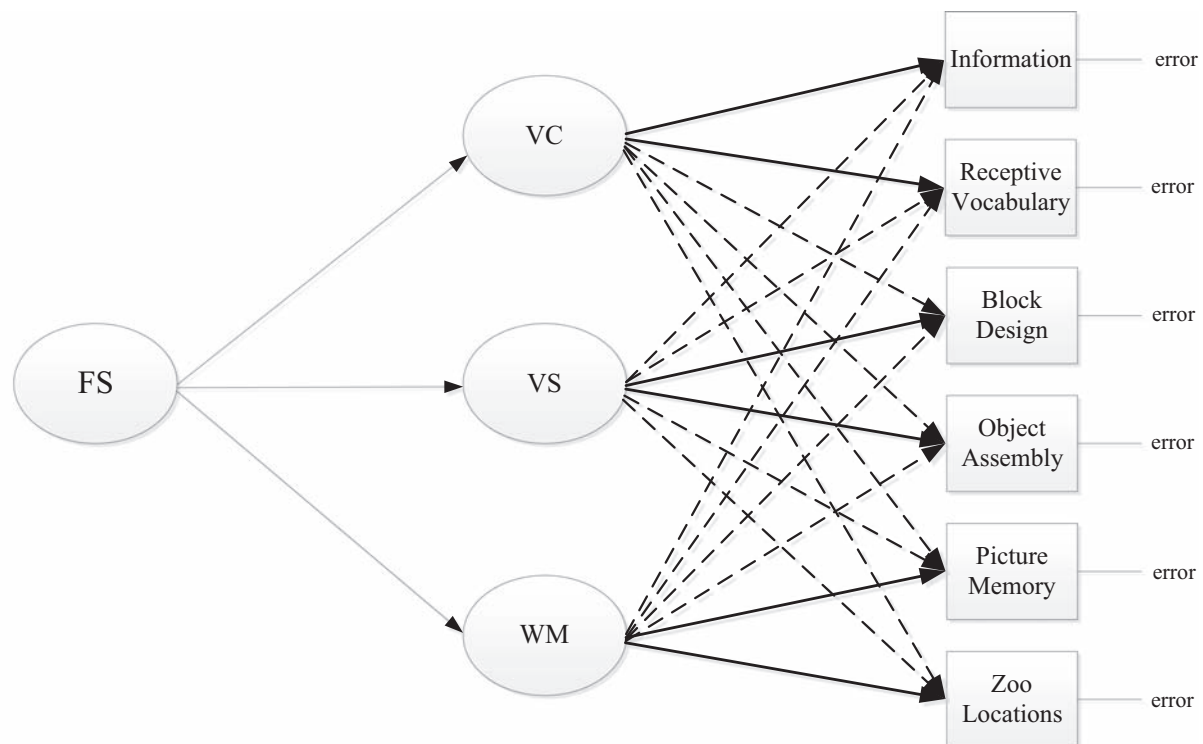
Bayesian CFA is an approach that has recently emerged and provides increased flexibility in CFA. Bayesian probability theory and statistical methods have existed since the 18th century. However, until recently the use of Bayesian statistical methods has been limited. Lack of use of Bayesian methods is due to (a) a different paradigm than that of the well-established long-run (a.k.a. frequentist) school of probability and (b) access to software to enable the use of the methods. The software *Mplus* and R are available to conduct Bayesian analyses. Also, many published articles and books are available to allow researchers to learn and become proficient applied users of Bayesian statistical methods.

Bayesian statistical thinking can be viewed as an extension of the traditional or frequentist approach, in that it formalizes aspects of the statistical analysis that are left to uninformed judgment by researchers in classical statistical analyses (Price, 2016; Lee, 2007). A critical difference between Bayesian and Classical probability is that in the Bayesian framework, the data are fixed and the parameters are random. Conversely, in classical probability, the data are random and the estimated parameters are fixed. Because the parameters estimated in a Bayesian analysis are random, researchers are able to make direct probabilistic statements about (a) point estimates such as the mean, median, mode, or regression weights and (b) random parameters lying within specific upper and lower limits (i.e., credibility intervals). Also, Bayesian credible intervals often contain frequentist coverage probabilities, often being close to Bayesian coverage levels (Lee, 2007).

Attractive features of using Bayesian methods include the following. First, Bayesian estimation does not assume that posterior distributions of parameters are normal (i.e., no reliance on large-sample, traditional statistical theory), making for better performance in small sample size conditions. Second, the estimation of complex models with multiple latent variables (factors) with categorical indicators or mixtures of continuous and categorical is possible where traditional maximum likelihood estimation fails. Third, in scenarios where missing data is present, full information maximum likelihood (FIML) or multiple imputation (MI) are automatically implemented based on the analytic model based on user preference. Fourth, higher statistical power under conditions where the distributions of indicator variables are not normal. Finally, multilevel and mixture models are more easily and accurately estimated. As Marsh et al. (2009, 2012) notes, poor model fit occurs using traditional ML methods in large samples and in multi-dimensional measurement models with a large set of indicators. Therefore, Bayesian CFA is ideal for fitting longitudinal, multilevel, and multiple group analyses where models are often complex.

An example of a second-order Bayesian CFA is provided in Fig. 7. The impetus for this study arose from the observation of perfect model fit in the second-order CFA conducted on the Wechsler Preschool and Primary Test of Intelligence-Fourth Edition (WPPSI-IV; Wechsler, 2012) for ages 2 years, months to 3 years, 11 months ($N = 1100$) during the test development phase (Price, 2017). In the original second-order CFA conducted using traditional CFA with maximum likelihood estimation perfect model fit was observed ($\chi^2 = 2.0(6)$, $p = 0.91$; CFI = 1.0; RMSEA = 0.0; SRMR = 0.0). However, a non-positive definite factor covariance matrix was observed yielding the solution inadmissible. To examine this solution more thoroughly, Bayesian CFA with informative priors was employed to resolve the non-positive definite matrix solution between the measurement and latent variable parts of the second-order model in a way that provided an optimal tradeoff between simple structure and model-data fit.

Recall that in CFA, cross-loadings are constrained to be exactly zero (a requirement often resulting in poor model fit). In Bayesian CFA, parameters such as factor loadings and indicator error variances can be set to near or approximately zero but not exactly at zero without compromising a properly specified measurement model. Model estimation using Bayesian CFA proceeded in the current example by specifying informative priors on cross-loadings (dashed paths in Fig. 7) by using mean zero (0) and normal distributions (N) with small variance (0.01), or formally as $\sim N(0, 0.01)$. Factor variances were set to a value of 1.0 and the metrics of the priors align with standardized loading solutions. The variance of the prior distribution captures the uncertainty about parameter statistics with value of 0.01 yielding a highly informative prior on the magnitude of cross-loadings (e.g., ± 0.03 in a standardized loading metric; Muthén and Asparouhov, 2012). Assignment of the ± 0.03 yields a 95% standardized cross-loading limit of ± 0.28 . Bayesian analysis proceeds by estimating the posterior distributions of parameters using Markov chain Monte Carlo (MCMC)



Note. Heavy bold paths = anchor test; Dotted paths are set to an approximate zero value; Dashed paths are required to allow the assignment of non-zero priors. FS=full scale IQ, VC=vocabulary, VS=visual spatial, WM=working memory.

Fig. 7 Bayesian second-order confirmatory factor analysis.

sampling (e.g., see Gelman et al., 2004 for details). Given the use of MCMC methods in Bayesian analysis, additional evaluation of MCMC convergence to the posterior distribution is required. The convergence of the MCMC algorithm is evaluated by the potential scale reduction statistics (PSR). Gelman et al. (2004) recommends a PSR of 1.10 for all model parameters as evidence of operational convergence. Additionally, a review of convergence plots produced by the MCMC chains is recommended and typically reported in Bayesian structural equation modeling (BSEM), including Bayesian CFA. Finally, Bayesian posterior predictive checking (PPP) is conducted and reflects the discrepancy between the actual data and the posterior distribution replicated by the model (Gelman et al., 2004; Muthén and Asparouhov, 2012). Results of the Bayesian analysis revealed that (a) latent (factor) regressions were inflated in the original second-order CFA, (b) assigning informative priors provided a way to capture the impact of the cross-loadings and account for the skewness in the parameter distributions, and (c) inadmissible solutions due to a non-positive latent variable matrix were eliminated. In summary, Bayesian CFA provided higher sensitivity to model specification informed by theory.

Summary and conclusions

Confirmatory factor analysis is a technique used to analyze the efficacy of measurement models where the number of factors and their direct relationship are specified. Conducting CFA requires an *a priori* strategy where a hypothesized factor structure is assessed by evaluation of model fit to empirical (or simulated) data. Specifically, application of CFA involves (a) confirming that variables (e.g., items or test total scores) fit a particular pattern or cluster to form a certain dimension according to a theory (McDonald, 1999; Fabrigar and Wegner, 2012), and a (b) synthesis of information about the factors and their contribution reflected by examinee performance on the observed variables (e.g., scores on tests or other measurement acquisition formats). Importantly, researchers using CFA attempt to understand *why* the variables are correlated and the *degree* or level of accuracy the variables and factors provide relative to a theory (Price, 2016).

Traditionally, conducting CFA involved developing a restricted measurement model or models within the tradition of structural equation modeling (SEM) to provide validity evidence of scores on tests or other measurement instruments or protocols (Floyd and Widaman, 1995). This article reveals the change and evolution of CFA during the past 30 years. The confirmatory analytic framework has evolved substantially and now includes (a) longitudinal factor analysis (Little, 2013; McCardle, 2007), (b) second and higher-order CFA, (c) bifactor CFA, (c) Bayesian CFA Kaplan (2014), (d) multilevel CFA and (g) dynamic CFA of intra-individual change Browne and Zhang (2007) and Molennar (2004). This article provides an overview of these developments in the CFA framework. Finally, although not included in this article, CFA models are often incorporated into multiple indicator multiple cause (MIMC; Gunzler and Morris, 2015) models and Multitrait-Multimethod (MTMM; Marsh and Grayson, 1995) models in a variety of disciplines.

References

- Barker, C., Pistrang, N., Elliot, R., 2015. *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners*, third ed. John Wiley & Sons, Ltd, Oxford, UK.
- Boomsma, A., 1985. Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika* 50, 229–242.
- Brown, T.A., 2015. *Confirmatory Factor Analysis for Applied Research*, second ed. Guilford Press, New York, NY.
- Browne, M.W., Zhang, G., 2007. Developments in the factor analysis of individual time series. In: Cudeck, R., MacCallum, R.C. (Eds.), *Factor Analysis at 100: Historical Developments and Future Directions*. Lawrence Erlbaum, Mahwah, NJ, pp. 265–292.
- Chen, F.F., West, S.G., Sousa, K.H., 2006. A comparison of bifactor and second-order models of quality of life. *Multivariate Behav. Res.* 41, 189–225.
- Fabrigar, L.R., Wegner, D.T., 2012. *Exploratory Factor Analysis*. Oxford University Press, New York, NY.
- Flora, D.B., Curran, P.J., 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol. Methods* 9, 466–491.
- Floyd, F.J., Widaman, K.F., 1995. Factor analysis in the development and refinement of clinical assessment instruments. *Psychol. Assess.* 7, 286–299.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. *Bayesian Data Analysis*, second ed. Chapman-Hall, Boca Raton, FL.
- Gunzler, D.G., Morris, N., 2015. A tutorial on structural equation modeling for analysis of overlapping symptoms in co-occurring conditions using Mplus. *Stat. Med.* 34, 3246–3280. <https://doi.org/10.1002/sim.6541>.
- Harman, H.H., 1976. *Modern Factor Analysis*, third ed. University of Chicago Press, Chicago.
- Heck, R.H., Thomas, S.L., 2015. *An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches Using Mplus*, third ed. Routledge, New York, NY.
- Kaplan, D., 2014. *Bayesian Statistics for the Social Sciences*. Guilford Publications, New York, NY.
- Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A., 1971. *Foundations of Measurement*, vol. 1. Academic Press, San Diego: CA.
- Lee, S.Y., 2007. *Structural Equation Modeling: A Bayesian Approach*. Wiley, New York.
- Little, T.D., 2013. *Longitudinal Structural Equation Modeling*. Guilford Publications, New York.
- MacCallum, R.C., Widaman, K.F., Zhang, S., Hong, S., 1999. Sample size in factor analysis. *Psychol. Methods* 4, 84–99. <https://doi.org/10.1037/1082-989X.4.1.84>.
- Mansolf, M., Reise, S.P., 2017. When and why the second-order and bifactor models are distinguishable. *Intelligence* 61, 120–129.
- Marsh, H.W., Grayson, D., 1995. Latent variable models of multi-trait-multimethod data. In: Hoyle, R.H. (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*. Sage Publications, Thousand Oaks, CA, pp. 177–198.
- Marsh, H.W., Hau, K.T., 1999. Confirmatory factor analysis: strategies for small sizes. In: Hoyle, R. (Ed.), *Statistical Strategies for Small Sample Research*, pp. 251–284. Chapter 10.
- McArdle, J.J., 2007. Five steps in the structural factor analysis of longitudinal data. In: Robins, R., MacCallum, R.C. (Eds.), *Factor Analysis at 100: Historical Developments and Future Directions*. Lawrence Erlbaum, Mahwah, NJ.
- McDonald, R.P., 1999. *Test Theory: A Unified Treatment*. Lawrence Erlbaum, Hillsdale, NJ.
- Molennar, P.C.M., 2004. In: Cudeck, R., MacCallum, R.C. (Eds.), *Factor Analysis at 100: Historical Developments and Future Directions*. Lawrence Erlbaum, Mahwah, NJ, p. 100.
- Muthén, B.O., Asparouhov, T., 2012. Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychol. Methods* 17, 313–335.
- Muthén, B.O., Kaplan, D., 1985. A comparison of methodologies for the factor analysis of non-normal Likert variables. *Br. J. Math. Stat. Psychol.* 38, 171–189.

- Muthén, L.K., Muthén, B.O., 2002. How to use Monte Carlo study to decide on sample size and determine power. *Struct. Equ. Model.* 4, 599–620.
- Muthén, L.K., Muthén, B.O., 2007–2010. *Mplus User's Guide*, Version 6. Muthén & Muthén, Los Angeles.
- Muthén, B.O., 1984. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 49, 115–132.
- Muthén, B.O., 1991. Multilevel factor analysis of class and student achievement components. *J. Educ. Meas.* 28, 338–354.
- Muthén, B.O., 1994. Multilevel covariance structure analysis. *Socio. Methods Res.* 22, 376–398.
- Price, L.R., Tulskey, D., Millis, S., Weiss, L., 2002. Redefining the factor structure of the Wechsler Memory Scale-III: confirmatory factor analysis with cross-validation. *J. Clin. Exp. Neuropsychol.* 24 (5), 574–585.
- Price, L.R., 2016. *Psychometric Methods: Theory Into Practice*. Guilford Publications, New York, NY.
- Price, L.R., 2017. A didactic investigation of perfect fit in second-order confirmatory factor analysis: exploratory structural equation modeling and Bayesian approaches. *J. Biomet. Biostat.* 2 (2), 1011. ISSN: 2573-5470.
- Raykov, T., Marcoulides, G.A., 2006. *A First Course in Structural Equation Modeling*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Reise, S.P., Bonifay, W.E., Haviland, M.G., 2012. Scoring and modeling psychological measures in the presence of multidimensionality. *J. Pers. Assess.* 95 (2), 129–140.
- Reise, S.P., 2012. The rediscovery of bifactor measurement models. *Multivariate Behav. Res.* 47, 667–696.
- Schmid, J., Lieman, J.M., 1957. The development of hierarchical factor solutions. *Psychometrika* 22, 53–61.
- Schumacker, R.E., Lomax, R.G., 2010. *A Beginner's Guide to Structural Equation Modeling*, third ed. Routledge, New York, NY.
- Sniders, T.A.B., Bosker, R.J., 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage Publications, Thousand Oaks, CA.
- Torgerson, W., 1958. *Theory and methods of scaling*. Wiley, New York.
- Tulskey, D., Price, L.R., 2003. The joint WAIS-III and WMS-III factor structure: development and cross-validation of a six-factor model of cognitive functioning. *Psychol. Assess.* 15 (2), 149–169.
- Wang, J., Wang, X., 2012. *Structural Equation Modeling: Applications Using Mplus*. Wiley Higher Education Press, Hoboken, NJ. <https://doi.org/10.1002/9781118356258>.
- Wang, J., Watts, A.S., Anderson, R.A., Little, T.D., 2013. Common fallacies in quantitative research methodology. In: Little, T.D. (Ed.), *The Oxford Handbook of Quantitative Methods*. Oxford University Press, New York, pp. 718–758. <https://doi.org/10.1093/oxfordhb/9780199934898.013.0031>.
- Wechsler, D., 2012. *Wechsler Preschool and Primary Scale of Intelligence*, fourth ed. NCS Pearson, San Antonio, TX.
- Wolf, E.J., Harrington, K.M., Clark, S.L., Miller, M.W., 2013. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educ. Psychol. Meas.* 73, 913–934. <https://doi.org/10.1177/013164413495237>.

A tutorial on structural equation modeling

Pui-Wa Lei and Hui Zhao, Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA, United States

© 2023 Elsevier Ltd. All rights reserved.

Common SEM terminologies and notations	619
Model conceptualization and specification	620
Model estimation	621
Choice of estimation methods	622
Model evaluation	623
Conclusion	626
References	626

Structural equation modeling (SEM) is a flexible general modeling framework that encompasses conventional general linear models such as multiple regression and factor analytic models and is capable of many plausible extensions. This article focuses on modeling procedures for the basic covariance structure models that include directional or structural relationships among observed or latent variables. Structural equation modeling generally starts with model specification, followed by estimation, evaluation, and if necessary, modification (e.g., Lei and Wu, 2007; Kline, 2016, pp. 117–118). Important considerations at each step and commonly encountered modeling issues along with recommended guidelines, where applicable, are discussed and illustrated with a concrete example below.

Common SEM terminologies and notations

A general SEM may include a *measurement part*, in which relationships between observed indicator variables and the unobserved latent constructs they purport to measure are stipulated, and a *structural part* that specifies directional relationships among latent and/or observed variables. Fig. 1 illustrates a relatively simple full SEM model. It contains a measurement part in which the item scores (Item 1–Item 5) serve as indicators for the latent construct (reading efficacy) as well as a structural part in which the hypothesized regression relations among the other observed variables (gender, grade, signaling baseline, intervention, signaling posttest) and the latent construct (reading efficacy) are specified. Models that contain only the measurement part are considered a special case and known as *confirmatory factor analysis models* (CFA). CFA is often used to examine internal structure validity evidence

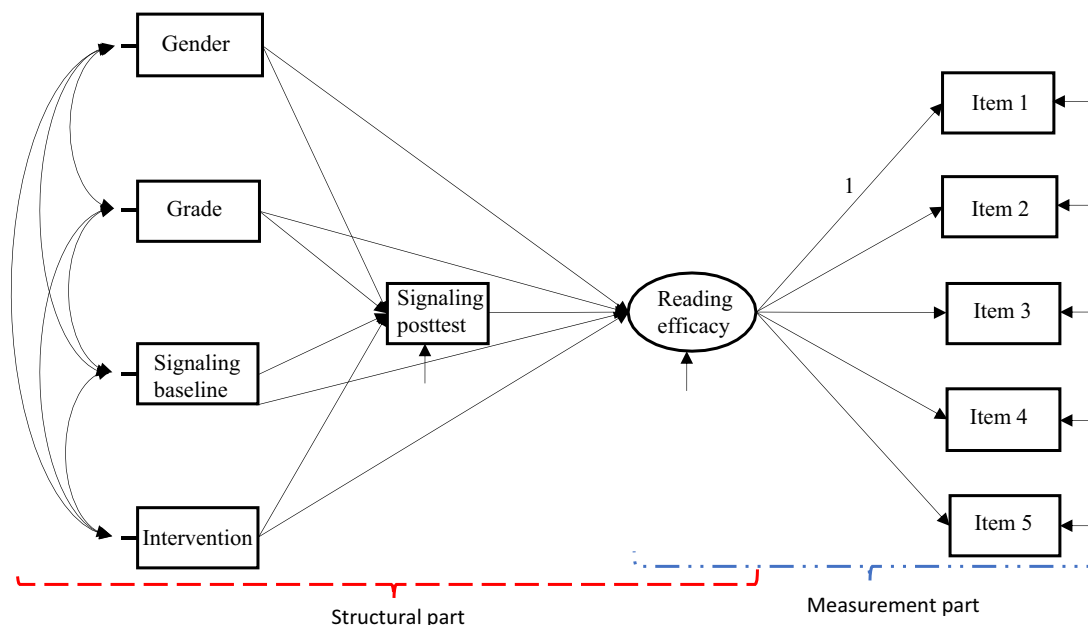


Fig. 1 Conceptual diagram of indirect intervention effect on self-efficacy via learning signaling words.

for measurement scores of latent constructs. *Path models* refer to the application of the structural part only on observed variables and can be considered a multivariate extension of regression to allow simultaneous estimation of multiple outcome variables, mediation effects, and correlated errors.

SEM models in education research are often complex. As pictures speak louder than words, *path diagrams* become an effective medium of communication for model specification. By SEM convention, square or rectangular shapes are used to enclose observed variables in a path diagram and round or oval shapes are used to enclose latent variables. Directional relationships between variables, latent or observed, are indicated by directional arrows with the hypothetical cause pointing to the presumed outcome (i.e., cause → outcome). In SEM, outcome or dependent variables that have other variables pointing directly to them are considered *endogenous* because their variances are being explained in the model (e.g., signaling posttest, reading efficacy, and Item 1–Item 5 in Fig. 1). Explanatory or independent variables that point to other variables only are considered *exogenous* as their variances are not being explained (e.g., gender, grade, signaling baseline, and intervention in Fig. 1). Inter-variable associations unexplained by other variables in the model are connected by bidirectional curvy arrows (↔; e.g., those among the observed exogenous variables). Equation residuals (referred to as *disturbances* in the structural part or *measurement errors* in the measurement part) can be represented by small circles (as they are latent) or simply by small arrows pointing to the endogenous variables in question without source variables.

Model conceptualization and specification

SEM models can be specified to test specific research questions or hypotheses that are developed from literature review of theories, prior research findings, or both. Important considerations for model specification include directions of structural relations, whether observed or latent variables are analyzed, and whether the model is theoretically identified (see definition below). Existing literature and previous findings can often shed light on the direction of structural relations among variables (e.g., whether self-efficacy predicts achievement, achievement predicts self-efficacy, or both in a reciprocal fashion). Timing or order of experimental events and measurements in research designs also helps to inform the direction as events that occur earlier should naturally be possible causes rather than outcomes of later events. Moreover, many constructs in Education and Psychology are not directly observable (e.g., self-efficacy, achievement) but are inferred from participant or informant responses to survey or test items that are designed to tap the latent constructs. As such, indicators (item scores or subscale scores) are reflective of the latent constructs (i.e., latent factors → indicator variables) in CFA or measurement part of SEM models and are called *effect indicators* (Kline, 2016, p. 196).

The choice between modeling observed or latent variables could depend on whether psychometric qualities, particularly internal structures, of the measures are well established in the relevant literature and whether sample size is sufficiently large to incorporate a measurement part which will increase the number of model parameters to be estimated. In general, larger sample sizes are required to provide stable estimates for larger models with more parameters. A common rule of thumb is that sample size should be no less than 10 or 20 times the number of parameters to be estimated in the model (e.g., Kline, 2016, p. 16). When sample size is inevitably small, one way to reduce model size without sacrificing the structural part is to approximate latent constructs by observed scores from well-established measurement scales, which essentially reduces the measurement part to a single-indicator-per-construct CFA. Reliabilities of the measure scores, if available, can be incorporated in the SEM model to account for measurement errors.

Once a model of interest is conceived, it is prudent to check whether it is *identified*. A model is said to be identified if all model parameters can be uniquely estimated. There are some basic necessary conditions for model identification, such as the number of unknowns (model parameters to be estimated) cannot be more than the number of knowns (observed variances and covariances in SEM), every latent variable must be assigned a scale, and a variable cannot be a predictor of itself, but they are not sufficient. Necessary and/or sufficient identification conditions have been established for certain model types (see Kline, 2016, chapters 7, 9, and 10 for more information). In general, reciprocal or feedback relations, correlated errors, complex indicators that cross-load on multiple factors, and small number of indicators for latent factors (e.g., two per factor) are potential causes of under-identification and require special attention.

To illustrate some of the concepts, suppose a researcher would like to test the logic model for a reading intervention program. The intervention program is designed to teach children conjunctive words that signal the type of text structure to improve reading comprehension and reading self-efficacy. In the research design, children were randomly assigned to the intervention or control conditions. Students' knowledge about signaling words were measured at baseline and after the intervention. Besides taking signaling word knowledge tests after the intervention, students were also asked to rate several items about reading self-efficacy on a 4-point Likert type scale. The hypothetical conceptual model is presented in Fig. 1. As the logic model hypothesizes that the intervention program will increase signaling word knowledge which will in turn increase reading self-efficacy, Fig. 1 shows intervention → signaling posttest → reading efficacy with the demographic variables (gender and grade) and the signaling baseline measure controlled for. Intervention also points directly to reading efficacy to allow for partial mediation through signaling posttest. Reading efficacy was latent and indicated by five items in the model. All other variables were treated as observed with the implicit assumption that they were measured without errors, which would likely hold for demographic and intervention condition indicators. Signaling knowledge was not likely measured without errors and could also be modeled as latent with a large sample. However, it was treated as observed (due to reasonably good internal structure evidence) to keep the model simple for illustration purpose. The directional arrows pointing to signaling posttest could be justified by the order of measurement (i.e., the source

predictors were measured prior to the posttest). The path from signaling posttest to reading efficacy was based on the logic model. Because signaling posttest and reading efficacy were measured at approximately the same time, the opposite direction or direct feedback between them could not be ruled out based on time precedence and the results should be interpreted with caution. Nevertheless, the model is theoretically identified because there are more knowns than unknowns, more than three indicators for the reading efficacy factor, and there are no correlated errors or feedback structural relations.

Model estimation

The general SEM model can be expressed as a system of equations. Using Bollen's (1989, p. 320) notation, the measurement part includes: $y = \Lambda_y \eta + \varepsilon$ and $x = \Lambda_x \xi + \delta$,

where y and x are vectors of observed variables; Λ_y and Λ_x are matrices of factor loadings (λ 's) for latent endogenous variables η and latent exogenous variables ξ , respectively; and ε and δ are vectors of measurement errors for y and x , respectively. The structural part given by Bollen (1989, p. 319) is: $\eta = B\eta + \Gamma\xi + \zeta$, where B is a matrix of coefficients (β 's) among the latent endogenous variables with 0 on the main diagonal (i.e., each endogenous variable is allowed to regress on other endogenous variables but not on itself), Γ is a matrix of coefficients (γ 's) between latent endogenous variables and latent exogenous variables, and ζ is a vector of disturbances (or equation residuals or errors). Expected values of errors (ε , δ , and ζ) are zero. Assuming that the latent and observed variables are centered at their means (i.e., not modeling mean structures), and errors are uncorrelated with the latent variables and with each other (see Bollen, 1989, p. 320 for more details), the covariance matrix among observed variables can be written as a function of the model parameters (Bollen, 1989, p. 325):

$$\Sigma = \begin{bmatrix} \sum_{yy} & \sum_{yx} \\ \sum_{xy} & \sum_{xx} \end{bmatrix} = \begin{bmatrix} \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}(\Gamma\Phi\Gamma' + \Psi)[(\mathbf{I} - \mathbf{B})^{-1}]'\Lambda_y' + \Theta_\varepsilon & \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}\Gamma\Phi\Lambda_x' \\ \Lambda_x\Phi\Gamma'[(\mathbf{I} - \mathbf{B})^{-1}]'\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta \end{bmatrix} \quad (1)$$

where $\mathbf{I}$ is an identity matrix, $(\mathbf{I} - \mathbf{B})$ is nonsingular, Φ is the covariance matrix for ξ , Ψ is the covariance matrix for ζ , Θ_ε and Θ_δ are covariance matrices for ε and δ , respectively. Therefore, factor loadings in Λ_y and Λ_x , path coefficients in $\mathbf{B}$ and Γ , variance-covariance elements among errors in Θ_ε , Θ_δ , and Ψ , as well as variance-covariance elements among latent exogenous variables in Φ are model parameters to be estimated.

Most applications do not need all matrices mentioned above or all elements in them. Non-applicable elements would become zeros, and some values are fixed to known constants for identification purpose (e.g., one λ for each latent variable is commonly fixed to 1 or the factor variance is fixed to 1 to provide a scale for the latent variable; the diagonal elements of $\mathbf{B}$ are fixed to 0 as an endogenous variable cannot be a cause for itself). For the model depicted in Fig. 1, because the exogenous variables are not latent, $\Lambda_x = \mathbf{I}$, $\Theta_\delta = \mathbf{0}$, $x = [\text{gender, grade, signaling baseline, intervention}] = \xi$, and Φ will take the values of the observed variances and covariances among the x -variables. The measurement model for y is,

$$y = \begin{bmatrix} \text{Item 1} \\ \text{Item 2} \\ \text{Item 3} \\ \text{Item 4} \\ \text{Item 5} \\ \text{SgnlPo} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ \lambda_{41} & 0 \\ \lambda_{51} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \text{efficacy} \\ \text{ISgnlPo} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ 0 \end{bmatrix}$$

where SgnlPo is signaling posttest, ISgnlPo is a phantom latent variable set to be the same as the observed variable by fixing the loading of SgnlPo on ISgnlPo (i.e., λ_{62} in Λ_y) to 1 and error variance of SgnlPo to 0. The structural part of the model is (SgnlBl is signaling baseline),

$$\eta = \begin{bmatrix} \text{efficacy} \\ \text{ISgnlPo} \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \text{efficacy} \\ \text{ISgnlPo} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \end{bmatrix} \begin{bmatrix} \text{Gender} \\ \text{Grade} \\ \text{SgnlBl} \\ \text{Intervention} \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}.$$

For readers unfamiliar with matrix notations, the model can be simplified to a set of simple linear equations, one for each of the endogenous variables as a function of the variables (including errors) pointing directly to them. For example, the model in Fig. 1 can be expressed as:

$$\begin{aligned} \text{Item 1} &= 1 * \text{efficacy} + \varepsilon_1 \\ \text{Item 2} &= \lambda_{21} * \text{efficacy} + \varepsilon_2 \end{aligned}$$

$$\text{Item 3} = \lambda_{31} * \text{efficacy} + \varepsilon_3$$

$$\text{Item 4} = \lambda_{41} * \text{efficacy} + \varepsilon_4$$

$$\text{Item 5} = \lambda_{51} * \text{efficacy} + \varepsilon_5$$

$$\text{efficacy} = \beta_{12} * \text{signaling posttest} + \gamma_{11} * \text{Gender} + \gamma_{12} * \text{Grade} + \gamma_{13} * \text{signaling baseline} + \gamma_{14} * \text{Intervention} + \zeta_1$$

$$\text{Signaling posttest} = \gamma_{21} * \text{Gender} + \gamma_{22} * \text{Grade} + \gamma_{23} * \text{signaling baseline} + \gamma_{24} * \text{Intervention} + \zeta_2$$

Model parameters (e.g., the λ 's, β 's, γ 's, variances of ε 's in Θ_ε , and variances of ζ in Ψ for the model in Fig. 1) are estimated by minimizing a fit function which summarizes the discrepancy between the observed covariance matrix and the model implied covariance matrix (the rightmost side of Eq. 1). Different estimation methods differ by the way fit functions are defined and may require different distributional assumptions. Because there is no close-form solution for most SEM models, iterative procedures are generally used to find the set of parameter values that minimize the fit function. Iterative procedures require a set of starting values to begin, and parameter values are adjusted iteratively until some convergence criteria (usually in terms of negligibly small differences in parameter values or fit function values between adjacent iterations) are met. The model chi-square goodness-of-fit test statistic, where applicable, is also a function of the converged fit function value. Starting values and convergence criteria are usually supplied by SEM software programs by default and can be overwritten by users. Because iterative procedures may stop when a local minimum (or maximum for maximum likelihood estimation) is reached, SEM modelers are recommended to try different starting values to ensure that the model converges with the same or largely similar set of parameter estimates.

Choice of estimation methods

An important decision for applied SEM researchers is the choice among estimation methods available in SEM software programs. Practical considerations for estimator choice include, but are not limited to, satisfaction of distributional assumption, measurement scale of the observed (particularly endogenous) variables, sample size, and missing data amount and mechanism. The normal-theory-based maximum likelihood (ML) estimation method is probably the most widely used and the default in many SEM software programs. However, it assumes multivariate normal distribution or no excessive kurtosis of observed variables (e.g., Lei and Wu, 2012), an assumption that is hardly satisfied in practice (Micceri, 1989). Robust versions of ML have been developed to deal with violations of normality. A version of robust ML provides regular ML parameter estimates but incorporates Satorra and Bentler's (1988, 1994) scaled statistics to provide robust standard errors and adjusted model chi-squares. Model chi-squares can be mean-adjusted (e.g., MLM in Mplus) or mean- and variance-adjusted (e.g., MLMV in Mplus), but the mean-adjusted version has been more extensively examined and more commonly used. Although this version of robust ML has been found to perform well for non-normally distributed continuous variables with complete data, it should be noted that it does not address the nonlinearity issue associated with categorical endogenous variables (e.g., McDonald and Ahlwat, 1974; Mislevy, 1986; Lei and Shiverdecker, 2020).

Moreover, full-information maximum likelihood (FIML) estimation method, that is considered one of the current state-of-the-art methods for handling missing data (e.g., Graham, 2009), utilizes a modified likelihood function based on individual data (e.g., Enders, 2001). That is, the default ML in SEM programs is not necessarily FIML and not all ML is FIML. Users should consult manuals of SEM software programs to figure out how FIML is specified (see Lei and Wu, 2012, Appendix 10.1 for some examples). If one encounters both missing data and nonnormality issues, another version of robust ML called MLR can be considered because MLR uses FIML and yields robust statistics (e.g., Lei and Shiverdecker, 2020).

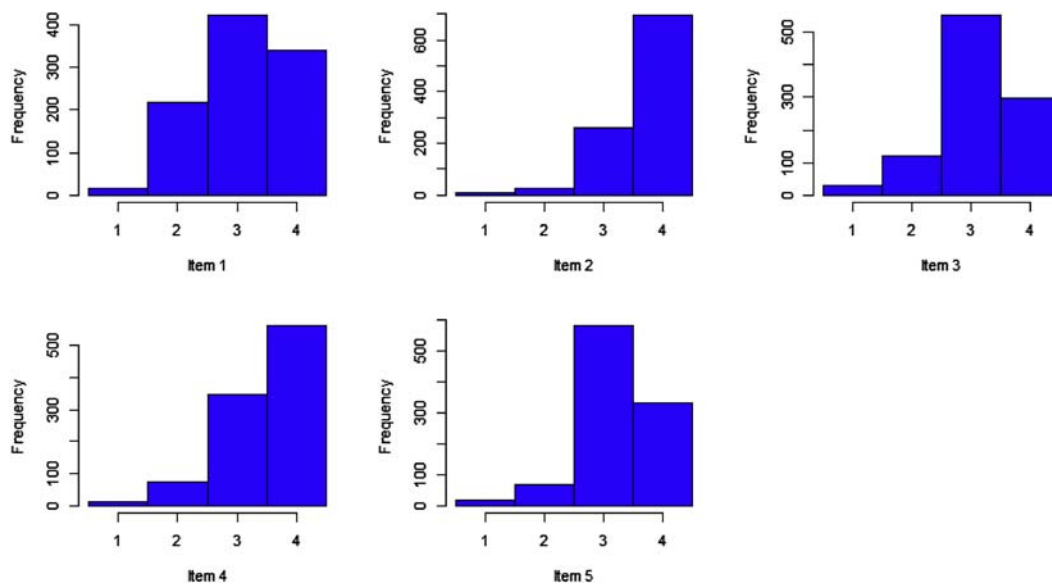
Ordinal variables (Likert-type scales) are commonly used in measurement models where item scores serve as indicators for the latent constructs. If normal-theory-based ML, with or without robust statistics, is used for ordinal endogenous variables, parameter estimates may be biased except perhaps when the number of ordinal categories is large (5 or more) and the scores are symmetrically distributed (e.g., Rhemtulla et al., 2012). In general, categorical estimation methods should be used for ordinal endogenous variables (see Lei and Shiverdecker, 2020 for a brief discussion of different approaches to dealing with categorical data in SEM). In short, categorical endogenous variables need to be explicitly declared as such in SEM software programs to overwrite the default assumption of modeling continuous variables. Both ML and weighted least squares (WLS) estimators can be used with categorical variables but their robust versions are more generally recommended (Lei and Shiverdecker, 2020). Specifically, previous empirical studies have showed that the full WLS requires very large sample sizes to perform well, while robust (or diagonal) WLS tends to outperform full WLS at realistic sample sizes (e.g., Flora and Curran, 2004). WLSMV is a robust WLS that provides mean- and variance-adjusted chi-squares as well as robust standard errors. It is the default estimator in Mplus when categorical endogenous variables are declared and estimates polychoric/polyserial correlations as sample statistics for model estimation. However, WLSMV currently handles missing data by listwise deletion or pairwise deletion and might not be a good choice when missing data amount is not small or missingness is related to observed variables (i.e., MAR; Asparouhov and Muthén, 2010). Full-information ML estimators such as MLR might be better in dealing with missing data.

For illustration purpose, the model presented in Fig. 1 was estimated using a random sample of 1000 observations drawn from a larger study. Sample descriptive statistics for the observed variables are given in Table 1. Frequency distributions of the categorical indicators for the latent variable are shown in Fig. 2. It was clear from Fig. 2 that the 4-point indicator variables were quite skewed and more so for Items 2 and 4. As such, the item scores should be treated as categorical. Three estimation methods available in Mplus and examined by Lei and Shiverdecker (2020) for categorical variables were compared to demonstrate their similarities and differences. The first was MLR without declaring item scores as categorical, a common practice in applied research. The second

Table 1 Descriptive statistics for observed variables.

	<i>Item 1</i>	<i>Item 2</i>	<i>Item 3</i>	<i>Item 4</i>	<i>Item 5</i>	<i>SgnlPo</i>	<i>SgnlBl</i>	<i>Gender</i>	<i>Grade</i>	<i>Intrvntn</i>
Item 1	1	<i>0.506</i>	<i>0.416</i>	<i>0.688</i>	<i>0.469</i>	<i>0.303</i>				
Item 2	0.394	1	<i>0.392</i>	<i>0.518</i>	<i>0.465</i>	<i>0.187</i>				
Item 3	0.352	0.293	1	<i>0.444</i>	<i>0.426</i>	<i>0.205</i>				
Item 4	0.576	0.419	0.357	1	<i>0.449</i>	<i>0.276</i>				
Item 5	0.382	0.353	0.348	0.362	1	<i>0.148</i>				
SgnlPo	0.320	0.195	0.218	0.285	0.163	1				
SgnlBl	0.190	0.192	0.109	0.205	0.115	0.360	1			
Gender	0.028	0.012	−0.030	−0.008	−0.063	0.069	0.040	1		
Grade	0.006	0.015	0.020	−0.014	−0.029	0.144	0.138	0.023	1	
Intrvntn	0.029	−0.022	0.045	0.009	0.053	0.217	−0.005	0.002	0.000	1
Mean	3.09	3.65	3.12	3.47	3.23	16.42	13.48	0.49	0.54	0.50
SD	0.78	0.59	0.72	0.69	0.64	7.16	5.82	0.50	0.50	0.50

Note. Table values on the lower triangle are Pearson correlations. Italicized values on the upper triangle, where applicable, are polychoric (among items) or polyserial (with SgnlPo) correlations. SgnlPo = signaling posttest score; SgnlBl = baseline signaling score; Intrvntn = intervention condition. $N = 1000$.

**Fig. 2** Frequency distributions of item scores for reading efficacy.

was MLR with item scores declared as categorical (MLR-CAT) and used probit link to put estimates on the same standard scale as WLSMV. The third was the default WLSMV with probit link for categorical endogenous variables. The sample size should be sufficient for the model size and these estimators (Lei and Shiverdecker, 2020).

Both unstandardized and standardized (with respect to endogenous and latent variables, STDY) parameter estimates provided by the three estimation methods are shown in Table 2. STDY was chosen because all but one exogenous variable was dichotomous. As expected, the path coefficients between the continuous endogenous variable and the exogenous variables were identical across methods. The largest difference between treating item scores as categorical versus not was found in the factor loadings. Standardized factor loadings provided by MLR without declaring items as categorical were consistently lower than those by MLR-CAT and WLSMV (differences ranged from 0.06 to 0.12). Standardized MLR-CAT and WLSMV estimates were largely similar. Interestingly, differences among methods in other standardized parameter estimates (≤ 0.02) were essentially negligible. Since there was no missing data in this example, the computationally intensive MLR-CAT was not necessary and WLSMV would be preferred.

Model evaluation

Before interpreting results from any model, goodness-of-fit of the model to data should be assessed. A chi-square goodness-of-fit test is available for most estimators. However, the model chi-square statistic (simply the product between the converged fit function value and sample size or sample size minus 1) is very sensitive to sample size. Therefore, alternative fit statistics that are less affected

Table 2 Parameter estimates (standard errors) provided by different estimation methods.

Model parameters	MLR		MLR-CAT		WLSMV	
	Unstd. Est. (SE)	StdY. Est. (SE)	Unstd. Est. (SE)	StdY. Est. (SE)	Unstd. Est. (SE)	StdY. Est. (SE)
Direct effects on SgnlPo						
Gender	0.76 (0.41)	0.11 (0.06)	0.76 (0.41)	0.11 (0.06)	0.76 (0.41)	0.11 (0.06)
Grade	1.37 (0.42)	0.19 (0.06)	1.37 (0.42)	0.19 (0.06)	1.37 (0.42)	0.19 (0.06)
SgnlBI	0.43 (0.03)	0.06 (0.00)	0.43 (0.03)	0.06 (0.00)	0.43 (0.04)	0.06 (0.01)
Intrvntn	3.13 (0.41)	0.44 (0.06)	3.13 (0.41)	0.44 (0.06)	3.13 (0.43)	0.44 (0.06)
Factor loadings/direct effects on reading efficacy						
Item 1	1.00 (0.00)	0.74 (0.02)	1.00 (0.00)	0.82 (0.02)	1.00 (0.00)	0.81 (0.02)
Item 2	0.57 (0.05)	0.56 (0.03)	0.66 (0.08)	0.68 (0.03)	0.82 (0.04)	0.68 (0.03)
Item 3	0.62 (0.05)	0.51 (0.03)	0.50 (0.06)	0.57 (0.03)	0.72 (0.04)	0.59 (0.03)
Item 4	0.87 (0.05)	0.75 (0.03)	1.09 (0.14)	0.84 (0.03)	1.02 (0.04)	0.83 (0.02)
Item 5	0.58 (0.05)	0.53 (0.03)	0.55 (0.07)	0.61 (0.03)	0.77 (0.04)	0.63 (0.03)
Gender	<i>−0.04 (0.04)</i>	<i>−0.07 (0.07)</i>	<i>−0.11 (0.10)</i>	<i>−0.08 (0.07)</i>	<i>−0.07 (0.06)</i>	<i>−0.09 (0.07)</i>
Grade	<i>−0.09 (0.04)</i>	<i>−0.15 (0.07)</i>	<i>−0.19 (0.10)</i>	<i>−0.14 (0.07)</i>	<i>−0.12 (0.06)</i>	<i>−0.14 (0.07)</i>
SgnlPo	0.03 (0.00)	0.36 (0.04)	0.07 (0.01)	0.36 (0.04)	0.04 (0.00)	0.35 (0.04)
SgnlBI	0.02 (0.00)	0.03 (0.01)	0.03 (0.01)	0.02 (0.01)	0.02 (0.01)	0.03 (0.01)
Intrvntn	<i>−0.07 (0.04)</i>	<i>−0.09 (0.07)</i>	<i>−0.12 (0.10)</i>	<i>−0.09 (0.07)</i>	<i>−0.06 (0.06)</i>	<i>−0.08 (0.07)</i>
Indirect effects on reading efficacy via SgnlPo						
Intrvntn	0.09 (0.02)	0.16 (0.03)	0.22 (0.04)	0.16 (0.03)	0.13 (0.02)	0.15 (0.03)
Residual variances						
SgnlPo	41.46 (1.62)	0.81 (0.02)	41.46 (1.62)	0.81 (0.02)	41.46 (2.30)	0.81 (0.02)
Reading efficacy	0.29 (0.02)	0.82 (0.03)	1.64 (0.28)	0.83 (0.03)	0.58 (0.03)	0.83 (0.03)

Note. Variances of measurement errors are not shown. Italicized parameter estimates are not statistically significant at the 0.05 level. Unstd. Est. = unstandardized estimates; StdY. Est. = standardized estimates (standardizing latent and endogenous variables but not exogenous variables); SE = standard error; SgnlPo = Signaling posttest score; SgnlBI = Signaling baseline score; Intrvntn = intervention condition.

by sample size have been developed and mainly used for overall model fit evaluation in practice. There are two general categories of alternative fit indices for standalone overall model fit evaluation depending on whether comparison to a default, often independent model was used (incremental fit) or not (absolute fit). Because different fit indices might be sensitive to misspecifications in different types of parameters, the consensus is to examine multiple indices from different categories. Kline (2016) recommended reporting a minimum set of four fit statistics: model chi-square, Comparative Fit Index (CFI; Bentler, 1990), Root Mean Square Error of Approximation (RMSEA; Steiger, 1990), and Standardized Root Mean Square Residual (SRMR). CFI is an incremental fit index, and higher values indicate better fit (≥ 0.9 or 0.95 is desirable). An RMSEA value of ≤ 0.05 is considered a “close fit” (Steiger, 1990), ≤ 0.06 is desirable (Hu and Bentler, 1999), and > 0.10 is considered bad (Steiger, 1990). An SRMR value of ≤ 0.08 is desirable (e.g., Hu and Bentler, 1999) and > 0.10 is generally not good. Even though these rough rules-of-thumb might be helpful for understanding the metrics of these fit statistics, they were recommended based on studies under a very limited set of conditions and should not be used rigidly (see Kline, 2016, pp. 268–269 for a list of systematic steps for model fit evaluation).

Regardless of overall (or global) model fit, it is important to evaluate local fit and locate any sources of misfit. Local fit loosely refers to the appropriateness of model estimation results other than global fit statistics. Specifically, it is useful to check parameter estimates (e.g., whether there are improper solutions such as negative variances and greater than 1 correlations, whether estimates are significant and/or in expected directions), standard error estimates (whether they are reasonably small because large values that are improper for the parameters in question often indicate estimation problems or identification issues), R-squares for endogenous variables (to see if their variances are well explained), and residuals (standardized value $> |2|$ or correlation residual > 0.10 suggests non-negligible unexplained association). Model results are interpreted when both global fit and local fit are deemed adequate. Otherwise, the model could be modified to fit the data better.

Post hoc model modification is commonly aided by modification index, which is computed for each fixed parameter in the current model and approximates the amount of model chi-square drop should the fixed parameter be freely estimated (assuming other parameters remain unchanged). Modification indices are often reported along with the corresponding expected parameter change (EPC) values. As modification indices are sensitive to sample sizes, it would be helpful to consider the magnitude of EPC and whether the path in question makes theoretical sense if modification is necessary.

Joreskog (1993) discussed three general modeling approaches to SEM. One is strictly confirmatory such that modeling is considered complete regardless of goodness-of-fit of the hypothesized model to data; well-fitting models are adopted and ill-fitting ones are discarded. This all-or-nothing approach might be risky as the initial conceptualization might overlook paths that have not been the focus of theories or explored/reported in previous research. The second approach utilizes model comparison in which goodness-of-fit between competing models predicted by different theories or hypotheses is compared to determine which one is better supported by data. Comparison of fit between models can be based on information criteria such as AIC or BIC (lower is generally

better) for non-nested models or likelihood ratio chi-square difference test for nested models (see Kline, 2016, pp. 286–287 and p. 281, respectively, for more information). Model choice is usually based on a balance between goodness-of-fit and parsimony because a good fitting parsimonious model is more likely to generalize. The third approach is model generating which refers to the practice of post hoc model modification based on statistical indices. Restricted empirical searches for modification with reasonable justifications are generally acceptable. However, relying on a large number of post hoc modifications to derive a well-fitting model is NOT recommended because it is unlikely to lead to a true model (MacCallum, 1986; MacCallum et al., 1992). Regardless of the modeling approach, it is important to point out that goodness-of-fit does not necessarily mean that the model is true or correct, much like the maxim in basic statistical inference that failing to reject a null hypothesis does not mean it is true. Potentially many alternative models might fit the data just as well or even better.

To demonstrate model evaluation, results for the hypothesized model in Fig. 1 from the three estimation methods provided by Mplus were examined. Table 3 shows the overall model fit statistics available from Mplus and R-squares for the endogenous variables. The model estimated by MLR did not treat item scores as categorical, while the model estimated by MLR-CAT and WLSMV did. The models were not nested, and both AIC and BIC produced for ML estimators favored the MLR-CAT model. This suggested that the item scores should better be treated as categorical. Since the recommended minimum set of global fit statistics were not available for MLR-CAT and both categorical methods yielded largely similar parameter estimates, we focused on results from WLSMV.

Model chi-square was statistically significant indicating that the model did not fit well. However, the other global fit statistics suggested good fit ($CFI = 0.98 > 0.95$, $RMSEA = 0.04 < 0.05$, $SRMR = 0.03 < 0.08$). All model parameter and standard error estimates appeared reasonable (i.e., no out-of-bound or questionable values). Standardized factor loadings were acceptably large (≥ 0.59), and all but three path coefficients were statistically significant ($ps < 0.05$). R-square for reading efficacy and signaling posttest was 0.17 and 0.19, respectively, and considered moderately large using Cohen's thresholds for regression. All residual correlations were small ($\leq |0.076|$ and less than the 0.10 threshold). Modification indices suggested that some of the measurement errors might be correlated. For example, modification index for the error covariance between Item 1 and Item 4 was 17.32 (> 3.84) and standardized EPC was 0.45. Since the correlation residual for the pair was small (0.029) and there was no conceivable explanation for the correlation besides the latent factor, the model was not modified. Local fit was judged to be satisfactory. The large model chi-square was likely due to the relatively large sample size.

As for the hypothesized indirect effect of intervention on reading self-efficacy via signaling word knowledge, the model results appeared to support the hypothesis. The estimated indirect effect was statistically significant ($p < 0.001$). The standardized coefficient of 0.15 suggested that students in the intervention condition were 0.15 standard deviation higher than students in the control condition on reading self-efficacy via signaling knowledge, holding gender, grade level, and baseline knowledge constant. Even though the data did not seem to contradict the hypothesized mediation effect, it should be interpreted with caution because reading efficacy and signaling posttest were assessed at approximately the same time. A longitudinal design that allows model specification based on time precedence would be more suitable for testing mediation effects and require fewer assumptions.

Table 3 Model fit statistics and R-squares provided by different estimation methods.

	MLR	MLR-CAT	WLSMV
Fit statistics			
Loglikelihood	−7824.06	−7461.43	–
AIC	15,700.13	14,984.86	–
BIC	15,827.73	15,137.00	–
N-adjusted BIC	15,745.15	15,038.54	–
Model chi-square	52.41	–	67.99
df	25	–	25
RMSEA	0.03	–	0.04
CFI	0.98	–	0.98
TLI	0.97	–	0.98
SRMR	0.03	–	0.03
R-squares			
Reading efficacy	0.18	0.18	0.17
Signaling posttest	0.19	0.19	0.19
Item 1	0.55	0.67	0.66
Item 2	0.32	0.47	0.46
Item 3	0.26	0.33	0.35
Item 4	0.56	0.70	0.69
Item 5	0.28	0.38	0.40

Conclusion

In conclusion, SEM is a popular technique in many disciplines, including Education and Psychology. This article briefly introduces some fundamental concepts of SEM with a focus on covariance structure analysis and provides a modeling tutorial on procedures and related practical considerations. It is the authors' hope that it will inspire more judicious applications of SEM in applied research. It should also be noted that although many research questions can be addressed using the basic covariance structure models introduced here, the SEM field is fast growing and has seen prolific developments of models and estimation techniques. Advanced developments include multilevel SEM (e.g., Asparouhov and Muthén, 2016; Preacher et al., 2010), Bayesian SEM (e.g., Dombrowski et al., 2018; Muthén and Asparouhov, 2012), and exploratory SEM (e.g., Asparouhov and Muthén, 2009; Marsh et al., 2020), to name a few.

References

- Asparouhov, T., Muthén, B.O., 2009. Exploratory structural equation modeling. *Struct. Equ. Model.* 16, 397–438.
- Asparouhov, T., Muthén, B.O., 2010. Weighted Least Squares Estimation With Missing Data [Unpublished technical report]. Available at: <https://www.statmodel.com/download/GstrucMissingRevision.pdf>.
- Asparouhov, T., Muthén, B.O., 2016. General random effect latent variable modeling: random subjects, items, contexts, and parameters. In: Harring, J.R., Stapleton, L.M., Beretvas, S.N. (Eds.), *Advances in Multilevel Modeling for Educational Research: Addressing Practical Issues Found in Real-World Applications*. Information Age Publishing, Inc, Charlotte, NC, pp. 163–192.
- Bentler, P.M., 1990. Comparative fit indexes in structural models. *Psychol. Bull.* 107, 238–246.
- Bollen, K.A., 1989. *Structural Equations With Latent Variables*. John Wiley & Sons, Inc, New York.
- Dombrowski, S.C., Golay, P., McGill, R.J., Canivez, G.L., 2018. Investigating the theoretical structure of the DAS-II core battery at school age using Bayesian structural equation modeling. *Psychol. Sch.* 55 (2), 190–207.
- Enders, C.K., 2001. The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychol. Methods* 6, 352–370.
- Flora, D.B., Curran, P.J., 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol. Methods* 9, 466–491.
- Graham, J.W., 2009. Missing data analysis: making it work in the real world. *Annu. Rev. Psychol.* 60, 549–576.
- Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55.
- Joreskog, K.G., 1993. Testing structural equation models. In: Bollen, K.A., Lang, J.S. (Eds.), *Testing Structural Equation Models*. Sage, Newbury Park, CA, pp. 294–316.
- Kline, R.B., 2016. *Principles and Practice of Structural Equation Modeling*, fourth ed. The Guilford Press, New York.
- Lei, P.-W., Shiverdecker, L.K., 2020. Performance of estimators for confirmatory factor analysis of ordinal variables with missing data. *Struct. Equ. Model.* 27 (4), 584–601.
- Lei, P.-W., Wu, Q., 2007. Introduction to structural equation modeling: issues and practical considerations. *Educ. Meas.* 26 (3), 33–43.
- Lei, P.-W., Wu, Q., 2012. Estimation in structural equation modeling. In: Hoyle, R.H. (Ed.), *Handbook of Structural Equation Modeling*. Guilford Press, New York, pp. 164–180.
- MacCallum, R.C., Roznowski, M., Necowitz, L.B., 1992. Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychol. Bull.* 111, 490–504.
- MacCallum, R.C., 1986. Specification searches in covariance structure modeling. *Psychol. Bull.* 100, 107–120.
- Marsh, H.W., Guo, J., Dicke, T., Parker, P.D., Craven, R.G., 2020. Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and Set-ESEM: optimal balance between goodness of fit and parsimony. *Multivariate Behav. Res.* 55 (1), 102–119.
- McDonald, R.P., Ahlward, K.S., 1974. Difficulty factors in binary data. *Br. J. Math. Stat. Psychol.* 27 (1), 82–99.
- Micceri, T., 1989. The unicorn, the normal curve, and other improbable creatures. *Psychol. Bull.* 105, 156–166.
- Mislevy, R.J., 1986. Recent developments in the factor analysis of categorical variables. *J. Educ. Stat.* 11, 3–31.
- Muthén, B.O., Asparouhov, T., 2012. Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychol. Methods* 17, 313–335.
- Preacher, K.J., Zyphur, M.J., Zhang, Z., 2010. A general multilevel SEM framework for assessing multilevel mediation. *Psychol. Methods* 15 (3), 209–233.
- Rhemtulla, M., Brosseau-Liard, P., Savalei, V., 2012. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychol. Methods* 17, 354–373.
- Satorra, A., Bentler, P.M., 1988. Scaling corrections for chi-square statistics in covariance structure analysis. In: *Proceedings of the Business and Economic Section of the American Statistical Association*. American Statistical Association, Alexandria, VA, pp. 308–313.
- Satorra, A., Bentler, P.M., 1994. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye, A., Clogg, C.C. (Eds.), *Latent Variables Analysis: Applications for Developmental Research*. Sage, Thousand Oaks, CA, pp. 399–419.
- Steiger, J.H., 1990. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behav. Res.* 25, 173–180.

Clustering approaches

James E. Corter, Department of Human Development, Teachers College, Columbia University, New York, NY, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	627
An overview of clustering	627
Types of cluster models	628
Partitioning	628
K-means	628
Latent class cluster analysis	629
Hierarchical clustering	629
Agglomerative methods for hierarchical clustering	630
Additive trees	630
Overlapping clusters	630
Data preparation for a cluster analysis	631
Assessing fit and validity of clustering solutions	632
Illustrative applications of clustering in education	632
Clustering of countries	632
Clustering of educational institutions, classrooms, and teachers	633
Clustering of students	633
Clustering of educational test items	634
Clustering of learner knowledge	634
Conclusions	635
References	635
Further reading	636
Relevant websites	636

Introduction

Cluster analysis refers to a set of exploratory data analysis (EDA) methods for multivariate data, that sort instances into groups of similar items. Cluster analysis (CA) has been used widely in educational research, and applications are increasing due to the surge of interest in educational data mining. Application areas within education have included clustering of schools or other educational institutions, curricula, test items, and students. In particular, students may be grouped according to their similarity on demographic, social, or cognitive attributes, but often are classified based on their test scores, academic attainment or other aspects of their learning history. A highly selective review of these applications is given toward the end of this article, organized by the type of educational problem to which the method has been applied.

In order to better understand both these prior contributions and the potential of clustering methods in educational research, some background material may be helpful. First, the goals and uses of cluster analysis should be considered. Cluster analysis (CA) can be used merely as a technique for exploratory data analysis (EDA) or a data reduction tool, a way of grouping cases described by their profile on a set of variables so that groups of similar cases (clusters) can be analyzed separately, or themselves used as the cases for further analysis. CA is sometimes used for theory confirmation or theory generation, wherein the clusters are taken to indicate important subcategories of the modeled instances. This type of uses is common in education and psychology. Finally, CA is sometimes used in machine learning or learning analytics research as a form of “unsupervised learning,” a way of modeling or mimicking intelligent behavior, specifically the categorization of instances.

Clustering methods can themselves be grouped according to the type of underlying model that is fit: partitions, hierarchical clusterings, or sets of overlapping clusters. Within each of these general types of clustering models, a wide variety of criteria for defining optimal solutions and a wide variety of heuristic optimization methods have been proposed. For example, multivariate mixture models and latent class or latent profile cluster techniques (McCutcheon, 1987; Bauer, 2021) can be viewed as alternative methods to fit partitions. The next section below provides a brief primer on these methodological definitions and distinctions.

An overview of clustering

The general clustering problem can be described as follows. Given a set of conceptual objects, instances or data cases, each described as a vector of values on a set of predictor variables (usually numeric), allocate these instances to a set of clusters so that the instances within each cluster are relatively similar or homogeneous, and any two instances not in the same cluster are relatively dissimilar or

heterogeneous. In educational contexts, the instances may correspond to students, test items, classrooms, or other entities. Thus, a cluster is merely a set of instances, and a clustering (or cluster solution) is a set of clusters. Different types of clustering models can be defined formally in terms of the set relations among the clusters (Carroll and Corter, 1995).

Some clustering methods (e.g., k-means) take as input data a rectangular matrix of multivariate data $\mathbf{X}$, with N rows representing cases or instances (in educational contexts, the cases are often students) and P columns representing variables. Cluster methods are most often used to cluster the *rows* of such a data matrix; i.e., to discover groups of cases or instances that are highly similar. However, cluster methods could also be used to find clusters of *variables* (columns) that have similar patterns across cases or instances (rows). More often, principal components analysis (PCA) or factor analysis (FA) is employed to analyze relationships among the variables. However, factor-analytic rotation methods are often employed to find a more interpretable “simple structure” of factor loadings, wherein each variable heavily loads on only a single factor, and every variable loads on some factor. This type of ideal simple structure is equivalent to describing each factor as a list of associated high-loading variables, thus the PCA or FA solution, interpreted in this fashion, can be viewed as a way to identify a clustering (ideally a partitioning) of the variables. In fact, this is often the goal of a PCA or FA analysis; in applications where the variables are test items, each subset or “cluster” of variables may be taken as an interpretable dimension or subscale of the test. However, cluster methods offer a more direct way to achieve such a grouping of highly associated variables.

Some widely used clustering methods such as the k-means partitioning method and Ward’s method for hierarchical clustering (Ward, 1963) take as input the multivariate data matrix described above. These particular methods attempt to group (multivariate) cases so as to minimize the within-cluster variation and/or maximize the between-cluster variation on each numeric variable. However, other clustering methods, including most methods for hierarchical clustering and for fitting sets of overlapping clusters, take as input a matrix of *proximities*, that is an $N \times N$ matrix giving the similarities, distances, associations, or correlations between each pair of the cases or instances to be clustered. For these methods, if the user has begun with a rectangular $N \times P$ matrix $\mathbf{X}$ of multivariate data, it is necessary to first compute or select some measure of proximity between each pair of cases. Proximity is a general term that includes similarity and distance measures, as well as measures of association. Most statistics packages have functions or procedures to aid in this step of generating a proximity matrix, and some clustering software will perform this step automatically if given a rectangular multivariate data matrix as input. A commonly used measure of distance is the Euclidean distance between cases. For cases i and j , the Euclidean distance in the P -dimensional data space is

$$d(i, j) = \sqrt{\sum_{p=1}^P (x_{ip} - x_{jp})^2}$$

Correlations can also be considered to be a measure of similarity among cases (or variables), although the meaning of negative correlations must be carefully considered.

The choice of a proximity coefficient can have implications for the type of structure that is discovered. For example, measures of correlation are sensitive to the multivariate profile shape, but not to the overall level of the clustering variables. Measures of distance, in contrast, are sensitive to both shape and overall level. Most clustering software packages offer a wide variety of distance, similarity, and association coefficients that can be used to transform multivariate observations into proximities.

Types of cluster models

As already mentioned, cluster methods may be broadly grouped by the type of clustering model or solution they employ, into partitioning methods, hierarchical clustering methods, and methods for fitting sets of (arbitrarily) overlapping clusters. Thus, the term “non-hierarchical clustering” should be avoided, because it is ambiguous.

Partitioning

The most widely used clustering methods are methods for fitting partitions. A *partition* is a set of clusters that are mutually exclusive and exhaustive, meaning that no instance belongs to more than one cluster, and all instances belong to at least one cluster. As a consequence, in a partition each pair of clusters is disjoint (i.e., does not overlap). A partitioning is usually represented as a list of clusters, with each cluster represented as a list of the instances that belong to it. For example, the following list represents a hypothetical partitioning of five animals into three clusters C1, C2, and C3:

C1 = (dog, cat)
C2 = (bat)
C3 = (robin, sparrow)

K-means

The k-means algorithm (Sebestyn, 1962; MacQueen, 1967; Steinley, 2006) is probably the most widely used method to find partitions. It works as follows. The researcher decides on the number (k) of clusters to be fit to the data, and the algorithm begins by selecting some set of instances or cases as initial cluster seeds. These initial seeds could be randomly chosen cases, or well-

separated cases selected by some systematic method. Or, the researcher might select the initial cluster seeds for theoretical reasons or by some other external criterion. The algorithm commences by doing a “sorting pass” to assign all the remaining cases to the nearest cluster (represented by the seed on this first pass). “Nearest” means the cluster with the smallest Euclidean distance from the case in question. In subsequent iterations, the case is assigned to the cluster with the nearest *centroid* (defined as the profile of means on the P variables for that cluster). For cluster k , the centroid vector is

$$\bar{X}_k = [(\bar{X}_1)_k, (\bar{X}_2)_k, \dots, (\bar{X}_P)_k]$$

Following assignment of all instances to the nearest cluster, the cluster centroids are recomputed and another sorting pass commences. This iterative process of sorting and recomputing cluster centroids continues until it stabilizes, meaning that the cluster membership does not change in a sorting pass, at which point the cluster solution is determined.

The k-means algorithm can be thought of as an effective heuristic method for finding the partition into K clusters that minimizes the trace of the $P \times P$ within-cluster sum-of-squares and cross-products matrix W . i.e., if SSW_p denotes the sum of squares within clusters for variable p , and n_k denotes the number of cases in cluster k ,

$$SSW_p = \sum_{k=1}^K \sum_{i=1}^{n_k} (x_{ip} - \bar{X}_k)^2 = \sum_{i=1}^N (x_{ip} - \bar{X}_k)^2$$

$$tr(W) = \sum_{p=1}^P SSW_p = \sum_{p=1}^P \sum_{i=1}^N (x_{ip} - \bar{X}_k)^2$$

This minimization effectively minimizes the weighted average of the variances of the variables within the clusters. And because for each variable $SST = SSW + SSB$, it also maximizes the weighted average sum of squares between clusters (SSB), a measure of cluster distinctiveness.

However, the k-means algorithm is not guaranteed to find the optimal allocation of cases to the K clusters. Thus, it is recommended best practice to run k-means a large number of times for a given specified number of clusters K , using random starting seeds. The solution with the best fit measure (minimal SSW or $tr(W)$) is then taken to be the optimal solution. If the algorithm is run only once, it is best to let the program automatically choose well-separated cases for the initial cluster seeds.

Algorithms for k-means clustering are implemented in all of the most popular statistical packages, for example the QUICK CLUSTER procedure in SPSS, PROC FASTCLUS in SAS, and the *kmeans* function in the R package *stats*.

Latent class cluster analysis

Latent Class Cluster Analysis refers to partitioning clustering performed using Latent Class Analysis (LCA) or Latent Profile Analysis (LPA) methods. LCA and LPA are examples of the general class of techniques known as latent structure analysis (McCutcheon, 1987). Within the clustering field, these types of models are sometimes referred to as mixture models, or “model-based clustering.” The distinction between LCA and LPA is that the former assumes categorical observed variables, and the latter assumed continuous predictor variables. But note that LCA is sometimes used as a term referring to either case. The object of Latent Class Cluster Analysis is to classify each observation or instance into one of a set of mutually exclusive and exhaustive latent classes, i.e., clusters. The inferred class membership is expressed in probabilistic terms, thus these methods are sometimes referred to in the data mining literature as examples of “fuzzy partitioning” algorithms. However, a given observation or instance is generally assumed to be assigned to its modal class or cluster, so the results can be treated deterministically.

Latent Class Cluster models are fit by Maximum Likelihood Estimation, often using the EM algorithm (Dempster et al., 1977). Software to fit the Latent Class Cluster model is available from commercial packages such as LatentGOLD, and also from R packages including *poLCA* and *mclust*. The package *mclust* assumes continuous predictors (the LPA case), while *poLCA* assumes categorical predictors (as in LCA).

Hierarchical clustering

A hierarchical clustering is a set of clusters for which each pair of clusters is either disjoint, or one cluster includes the other (Carroll and Corter, 1995). Put another way, the clusters resulting from a hierarchical clustering form a nested set. A hierarchical clusterings is often displayed as a tree graph or dendrogram (see Fig. 1).

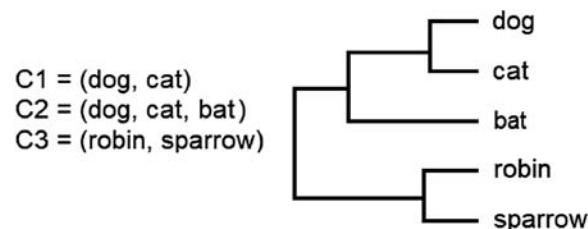


Fig. 1 A hierarchical clustering of five animals displayed as an ultrametric tree graph or dendrogram.

Agglomerative methods for hierarchical clustering

Hierarchical clustering (HC) is usually accomplished by *agglomerative* algorithms, whereby all n instances are initially treated as separate clusters. Then, at each step the most similar pair of instances or clusters are chosen and combined into a new cluster. This process continues until all clusters have been merged into a single cluster. Examples of such agglomerative algorithms include Ward's method (Ward, 1963), single-link, complete-link, and average-link clustering (Sokal and Sneath, 1963; Johnson, 1967). Evaluation studies of hierarchical clustering methods, some conducted with simulated data and some with real data, have tended to show Ward's method and average-link clustering (more specifically, the "unweighted pair-groups method using arithmetic averaging") as yielding the most satisfactory solutions.

In the tree diagram, the "height" of the node joining two clusters is an index of how dissimilar objects in the two clusters are. This node height for clusters is often reported by hierarchical clustering software, and is sometimes referred to as the "fusion coefficient." Interestingly, hierarchical clustering is often applied by researchers as a flexible means of finding partitions. Basically, any perpendicular (vertical) slice through the (horizontally oriented) dendrogram determines a partition (Cortier, 1996). For example, in Fig. 1, one such vertical line determines the partition $C1 = (\text{dog}, \text{cat})$, $C2 = (\text{bat})$, $C3 = (\text{robin}, \text{sparrow})$. Thus, the hierarchical clustering simultaneously determines partitionings with $k = 1$ to N clusters. In fact, Murtagh and Legendre (2014) recommend using hierarchical clustering with Ward's method to identify a good partition level, then seeking to refine that partition solution using k-means.

Algorithms for hierarchical clustering are implemented in most statistical packages. In SPSS, the CLUSTER procedure offers both average-link clustering and Ward's method (single-link and complete-link are also available). Note that the SPSS code for Ward's method is written in such a way as to require inputting a matrix of *squared* Euclidean distances rather than distances (Murtagh and Legendre, 2014). In SAS, PROC CLUSTER offers these methods. In R, several packages provide hierarchical clustering algorithms, with the *hclust* function of package *stats* and the *agnes* function of package *cluster* providing the basic hierarchical clustering methods mentioned above.

Additive trees

The distances in a tree diagram representing a hierarchical clustering satisfy the ultrametric inequality, which states that for any two instances x and y that are in the same branch or subtree, the distance between them ($d(x,y)$) and their distances to any other instance z not in that subtree satisfy the ultrametric inequality $d(x,y) \leq d(x,z) = d(y,z)$. A more general type of tree is the additive tree, which satisfies the more general condition called the additive inequality (Sattath and Tversky, 1977). The additive inequality states that for two objects x and y in one branch of the tree, and object u and v in another, $d(x,y) + d(u,v) \leq d(x,u) + d(y,v) = d(x,v) + d(y,u)$. Graphically, an additive tree resembles an ultrametric tree or dendrogram, but the leaves of the tree may vary in their distance to the root of the tree (see Fig. 2). This allows distances in the tree to reflect unique features of the instances that may make it more or less similar to all other instances in the tree (e.g., "bat" in Fig. 2).

The additive tree is a very general model that can accommodate even unidimensional structures (Cortier, 1996). An additive tree displayed in rooted form can be interpreted as a hierarchical clustering, with additional parameters or object constants (the leaf arcs) that can model an instance being relatively near or distant to the other objects in the tree (see Fig. 2).

Effective algorithms to fit additive trees to proximity matrices include ADDTREE/P (Sattath and Tversky, 1977; Cortier, 1982), GTREE (Cortier, 1998), de Soete's (1983) sequential unconstrained minimization technique, the iterative reduction method (Roux, 1988), and the iterative projection method (Hubert and Arabie, 1995). The last three methods above are implemented in the *ls_fit_addtree* function of the R package *clue*. With these methods, it is recommended to run the function a large number of times with a given data set and to take the best-fitting solution, just as for k-means. The ADDTREE/P and GTREE programs are available in the NETLIB repository of statistical software source code.

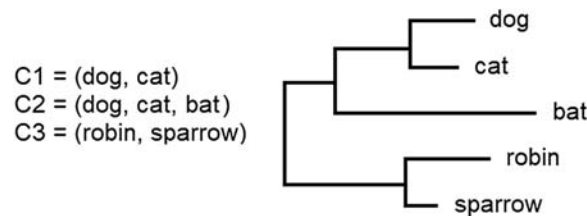


Fig. 2 A rooted additive tree representing the proximities among 5 objects, with branches interpreted as clusters.

Overlapping clusters

In an overlapping cluster solution, clusters may overlap in arbitrary patterns. This means that a given case or instance can be a member of 0, 1, or more clusters. An overlapping clustering can of course be represented as a list of clusters. But graphically representing sets of overlapping clusters is not straightforward. One representation involves first plotting points representing the cases or instances in the plane, then representing each cluster by a closed loop. An example is shown below in Fig. 3. Another possible representation is the extended tree (Cortier and Tversky, 1986), in which some clusters are represented as a tree, but others are

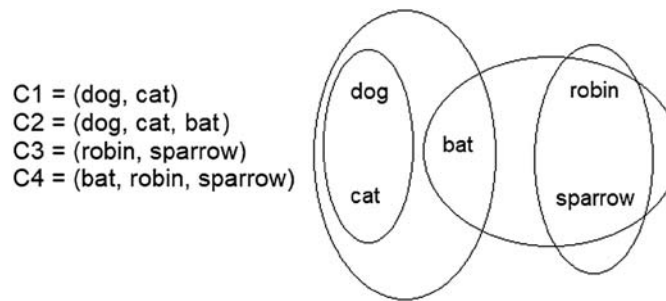


Fig. 3 An overlapping clustering of five animal terms, represented by the “loops” convention.

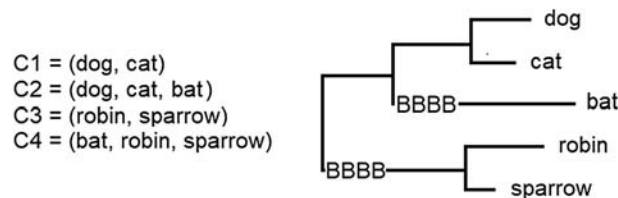


Fig. 4 An overlapping clustering of five animal terms, represented as an extended tree.

indicated by marked segments on the tree branches, as in Fig. 4. For example, “bat” shares a marked segment with “robin” and “sparrow”, presumably because they share the attribute of flight.

Algorithms to fit overlapping cluster solutions include ADCLUS (Shepard and Arabie, 1979), MAPCLUS (Arabie and Carroll, 1980), EXTREE (Cortier and Tversky, 1986), the additive profiles method (Mirkín, 1987), additive biclustering (Wilderjans et al., 2013; see also Van Mechelen et al., 2004) and overlapping k-means (Cleuziou, 2008). Overlapping cluster methods have been little used in education, due to their computational complexity and the fact that few mass-market statistical software packages offer such algorithms. However, the ADPROCLUS package to fit the additive profiles clustering model is available as a MATLAB program (Wilderjans et al., 2011), and the EXTREE program is available as PASCAL source code in the NETLIB repository of statistical software.

In spite of the limited availability of software for fitting these models, overlapping clustering methods seem potentially valuable for applications in education. One reason is that this model is a very general one, subsuming both partitions and hierarchical clustering as special cases. Thus, overlapping clustering methods can be argued to best fulfill the promise of a general “unsupervised learning” technique. Two examples that illustrate this potential are described below, one an application to the problem of inferring a set of cognitive or educational attributes in a diagnostic measurement context, and another involving the discovery of structure in student knowledge of arithmetic.

Data preparation for a cluster analysis

The above descriptions of clustering algorithms and software assume that the input data are either numeric predictors in the form of a rectangular multivariate data matrix, or a matrix of the pairwise proximities among the N cases to be clustered. If the data are in raw multivariate form, and include categorical predictors, dummy coding of the category levels may be necessary before computing distances or correlations.

One of the thorny problems in cluster analysis of multivariate data concerns standardization of variables. To illustrate, if we have students in a high-school classroom described by their age, gender (dummy coded as 1 = Female or 0 = Other), and SAT mathematics test score, it usually will not make sense to compute the Euclidean distances among students using these three raw-score variables, because the scale of the SAT score variable will cause it to dominate the other two measures, yielding essentially unidimensional structure in the distances matrix. For this reason, many researchers standardize the variables before inputting the data to k-means or Ward’s method, or before computing a proximity measure. However, this practice can sometimes cause noise to obscure signal. For example, if 48 of 50 students are age 16, but two are 17, standardizing the age variable will cause it to have equal weight with gender and SAT score, even though a one-year difference in age may be trivial or unimportant, and distributed in an arbitrary way. As another example, if the clustering variables are all Likert-scale rating items defined with a 5-point scale, it may not make sense to standardize them, even though some items show a larger variance than others. Simulation studies (Milligan and Cooper, 1988; Steinley, 2006) suggest that standardizing variables by their range is a good generic solution.

A data preparation technique that is fairly widespread in practice, but not always advisable, is to prepare for a cluster analysis by first conducting a principal components analysis (PCA) of the data, then using the scores on the important components (usually defined as those components with eigenvalues ≥ 1) as the input variables to the clustering, rather than the original variables. While

this practice seems to have advantages, such as the convenience of dealing with a smaller number of input variables, and ensuring that those input variables are uncorrelated, it has disadvantages as well. First, the continuous nature of the component scores may obscure cluster-like structure in the data. Second, using factor or component scores may hide, rather than solve, issues surrounding variable selection, weighting and standardization.

Assessing fit and validity of clustering solutions

The quality of a clustering solution can be assessed in any of several ways, and these may vary among methods. One criterion is fit of the cluster model to the data: does the derived solution do a good job of predicting variance in the input variables? For clustering methods such as Ward's method and k-means that take multivariate input, we can assess the RSQ due to the model for each clustering variable as $RSQ = SSB/SST$. The quality of the clustering is often assessed by a MANOVA checking if the clusters differ on the variables used in the clustering, but it should be noted that the p -values associated with such a MANOVA should be taken as merely descriptive, due to the circular logic in performing such a test (we find clusters differing on a set of variables, then test if they differ on these variables). For methods that take a proximity matrix as input, we can calculate model distances for each pair of cases from the derived solution, then correlate these model distances with the input proximities. This is referred to as the "cophenetic correlation" in some clustering references. This correlation squared also gives an estimate of the proportion of variance in the data accounted for by the model. In contrast, Latent Class Clustering models are fit by Maximum Likelihood Estimation, thus likelihood-based model fit statistics such as AIC or BIC can be computed and used.

Stability of the clustering solution is also of interest, because stability indicates a strong cluster structure that emerges regardless of minor changes in the analysis settings or parameters. For a method such as k-means, a given data set can (and should) be fit a large number of times with randomly chosen initial cluster seeds. The percentage of these runs that result in the maximal value of RSQ gives a good indication of the stability of the solution. Cluster stability might also be assessed using a resampling technique such as the jackknife or the bootstrap (e.g., Hennig, 2007).

The validity of a clustering solution can be assessed in a number of ways. First, of course, is the criterion of face validity: the clusters should be interpretable in terms of what is known about the application domain and the data. Also, the clusters should seem internally homogeneous and distinct from one another. Ideally, the clusters should also exhibit external validity, that is, the clusters ought to differ on relevant external criterion variables. For example, a clustering of students based on their patterns of interaction with an online course might reasonably be expected to predict learning outcomes. Thus, it is often informative to present profiles of the derived clusters (say, in a partitioning) not only on the variables used for the clustering, but also for other relevant demographic or outcome variables. MANOVAs or ANOVAs testing for differences in these related variables can often provide useful evidence of external validity.

Illustrative applications of clustering in education

Clustering methods have been applied to a wide variety of data pertaining to education, featuring analyses of the relationships among schools or other educational institutions, classrooms and teachers, curricula, test items, and students, among other entities. Many of these applications have been described in education-focused publications, but relevant work can be found in other literatures, including statistics and various subdisciplines of psychology, such as organizational behavior, cognitive, and clinical/counseling. The highly selective sample of educational applications given below are organized by their area of application within education, moving from the most general types of case or instance (clustering of countries, educational institutions) to the highly specific (clustering of individual students or test items).

Clustering of countries

Large international comparative studies of educational achievement such as Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) yield comprehensive databases that offer rich opportunities for secondary data analysis using multivariate methods, including clustering. For example, Olsen and Lie (2011) analyzed 52 science interest items embedded in PISA 2006 to explore 15-year old students' interest in science in 57 countries, in order to ascertain if science interest profiles differ across groups of culturally related countries. They analyzed the correlations among these interest items, first "double-centering" or residualizing the score on each item by subtracting the overall means for that country and that item. They then used average-link hierarchical clustering on these correlations to obtain a dendrogram of countries. The structure of the tree diagram showed a number of culturally related groups of countries. Based on this tree, using measures of cluster distinctiveness, the authors proposed a partition with clusters labeled *Chinese*, *Central/South Asian*, *Ibero-American*, *English*, *Nordic*, *German*, *French*, *Baltic*, and *Slavic*. They conclude that linguistic and other cultural factors seem to influence students' interest in science, with the distinction between scientific issues involving living systems and those addressing physical/technological systems playing a role in producing different profiles of item responses across countries and regions.

In another use of cluster analysis to explore educational differences among countries, Forbes et al. (2020) investigated the relationship between use of inquiry-based instruction and science achievement. Using data from PISA 2015, Forbes et al. first used

Latent Profile Analysis to find clusters of students with different profiles of inquiry-based science teaching within each of $n = 13$ countries selected to represent a range of science teaching practices and student science achievement. They found fundamental differences in the frequency and nature of inquiry-based instruction reported by students associated with the highest and lowest levels of science achievement. They then investigated differences among these countries in the profiles of inquiry-based teaching methods associated with the highest-performing cluster in each country using complete-link hierarchical clustering on the Euclidean distances among profiles. The hierarchical clustering revealed four main clusters of countries, including one cluster that showed low levels of inquiry-based teaching (Bulgaria, Hungary, South Korea) and one that showed overall high levels (Denmark, France). The two intermediate-level clusters were differentiated largely by the reported amount of classroom debate and student testing of their own ideas.

Clustering of educational institutions, classrooms, and teachers

Cluster analysis methods have been used successfully to study diversity in educational institutions, classrooms, and teachers. In a classic study by Moos (1978), student perceptions of classroom climate were used to analyze 200 junior high and high school classrooms. The classroom climate instrument used yielded nine subscales. The intraclass correlations for each classroom were used in a hierarchical cluster analysis, which yielded five groups of classrooms (with four classrooms left unclassified): control oriented, innovation oriented, affiliation oriented, task oriented and competition oriented. External validation of the clusters showed differences in student satisfaction and mood, and in teacher satisfaction.

Dorman et al. (2006) studied classroom environments, including incorporation of technology, in a sample of 286 secondary-school classrooms in Australia, using student responses to an instrument assessing 10 classroom environment dimensions: student cohesiveness, teacher support, involvement, task orientation, investigation, cooperation, equity, differentiation, computer usage and young adult ethos. A hierarchical cluster analysis (not further specified) of the student responses was used to find a partition into five groups of classrooms, which were labeled *exemplary*, *safe and conservative*, *non-technological teacher-centered*, *contested technological*, and *contested non-technological*.

A study by de Kruif et al. (2000) investigated patterns of teachers' interaction behaviors in early childhood classrooms. Sixty-three pre-K teachers were rated on their use of eight interaction behaviors, including such behaviors as redirecting, acknowledging, praising, etc. The authors used hierarchical clustering (not further specified) followed by k-means clustering to obtain four clusters of teacher interaction styles, labeled average, elaborating, controlling, and nonelaborative. Validation of the clusters by external variables indicated that the clusters were differentiated by teachers' education, scores on a caregiver sensitivity measure, classroom quality, group child engagement, and licensing level of the pre-K center.

Clustering of students

Clustering methods seem ideally suited to studying student attitudes, beliefs, and achievement, because these methods seek to describe and find common patterns in diversity. In particular, students may be grouped according to their similarity on demographic, social, or cognitive attributes, but often are classified based on their academic attainment or other aspects of their learning history.

As an example of clustering for the purposes of theory development or confirmation, Speece et al. (1985) used hierarchical clustering to identify seven distinct behavioral subtypes of children identified by their schools as learning-disabled. The clustering was based on subscale scores on a teacher rating of classroom behaviors. The obtained subtypes were stable and reliable, and the subtype profiles on the clustering variables differed from a sample of children without identified learning disabilities. The subtypes differed significantly on independent observational measures and on ratings by special education teachers, but not on achievement or intelligence. Interpretation of the subtypes confirmed a multiple-syndrome view of learning disabilities.

A study by Rowley (2000) used hierarchical clustering to explore African American students' beliefs about the value of getting a good education. Using Ward's method, five clusters of African American college students ($N = 126$) were identified with differing profiles on two measures of idealistic and context-specific educational utility and on grades. These clusters were then related to profiles on a racial ideology/Black identity scale. The results indicated that African American students hold a variety of perspectives about the value of education, and that certain types of beliefs naturally occur in students with certain academic performance records.

In a study of student motivation and self-regulated learning motivated by self-determination theory, Ng et al. (2016) collected data from 782 students in science courses from eight secondary schools in Singapore. Based on a questionnaire measuring motivational beliefs and learning strategies, they used Ward's method for hierarchical clustering to derive four clusters of students, then confirmed these clusters using k-means. The clusters differed significantly (by a MANOVA test) not only on the motivational beliefs and strategies variables used in the clustering, but also on the scales of a questionnaire based on self-determination theory, and on term grades in science.

Clustering has been frequently employed within the educational data mining (EDM) community as a means of finding structure in large data sets, where it is sometimes referred to as a technique for "unsupervised learning" (Romera and Ventura, 2007; Dutt et al., 2017). For example, Rodrigo et al. (2008) analyzed the logs of student interaction with Aplusix, an intelligent tutoring system for algebra. Based on a k-means analysis, they determined that students fell into two primary groups: one with a more collaborative interaction style and one with a more solitary style of work. In another study, Trivedi et al. (2011) clustered middle school students based on their work with the ASSISTments tutoring system, using both scores on problems and certain "dynamic" measures of

interaction with the system, including number of problems done, percent correct on the help questions, average time spent per item, average number of attempts per item, and average numbers of hints per item. Trivedi et al. used k-means clustering, with 200 random-seed starts, for the cluster analyses, and evaluated the results for $k = 2-7$ cluster solutions. They then developed separate prediction models for each cluster, using the above variables (plus the “static” measure of average score on the main problems) to predict the criterion of test score on the MCAS arithmetic subtest, a statewide Massachusetts test. These cluster-specific prediction models, and ensemble combination of their predictions, improve prediction of the criterion achievement scores.

Clustering of educational test items

When the achievement of students is measured, and item-level data is retained, the resulting persons $\times$ items matrix of scores may be fruitfully analyzed by cluster analysis. While cluster analysis of individuals based on their test performance has a long history (e.g., Thorndike, 1953), cluster analyses of the correlations among test items is relatively rare. This may be due in part to the traditional use of dimensional models such as PCA or factor analysis to analyze the correlations among items. But as mentioned above, FA or PCA is often employed to find homogeneous subsets of items that might be considered as subscales, a goal that can be accomplished more directly with clustering methods. A few such analyses (e.g., Revelle, 1979) are found in the literature, with some studies finding better model fit for hierarchical clustering models than for the traditional dimensional models (Beller, 1990).

In an application to diagnostic testing, Corter (1995) analyzed correlations from a diagnostic test of mixed-fraction subtraction skills, applying a hierarchical clustering method (additive trees) and the EXTREE method for finding overlapping clusters, to discover subsets of items that correspond to necessary subskills or “attributes” of the Q-matrix used in many diagnostic measurement models. Because the attributes of a Q matrix are not necessarily hierarchically structured (though they may be for some tests), the use of overlapping cluster methods seems most appropriate for this purpose. Recent work using clustering in the context of diagnostic classification models (DCMs) has returned to the problem of classifying examinees rather than items, because a DCM can be thought of as a method for sorting examinees into latent test profiles, essentially an LCA problem. Chiu et al. (2009) explored the use of similarity-based partitioning and hierarchical clustering methods as an alternative robust technique for accomplishing this diagnostic classification of examinees. They proved that for educational test data meeting certain assumptions, the probability that hierarchical cluster analysis assigns examinees to their true proficiency classes approaches 1 as the number of test items increases.

Clustering of learner knowledge

A study that illustrates the use of clustering as an unsupervised learning method employed overlapping clustering to discover structure in the arithmetic knowledge that mathematics learners possess about the numbers 0 through 9 (Shepard et al., 1975). Shepard et al. asked college students (expert-level learners) to rate the similarity of these numbers, presented as written numerals but judged as abstract concepts. The six-cluster additive clustering solution for these similarity judgments (obtained using the ADCLUS program, Shepard and Arabie, 1979) is shown in Fig. 5. Here, the clusters are depicted as loops enclosing the set of numbers associated with that cluster. The positions of the points representing the numbers 0–9 are arbitrary for the clustering, but are positioned here in a way that is convenient for the loops representation, as determined by a prior multidimensional scaling analysis.

The solution features highly interpretable clusters. One cluster picks out 0 and 1, the additive and multiplicative identities, another is comprised of the odd primes, and the powers and multiples of 2 and of 3 define the other four clusters. This overlapping cluster analysis demonstrates that student knowledge of the mathematical properties of the numbers can be inferred from a cluster analysis of the students’ similarity judgments.

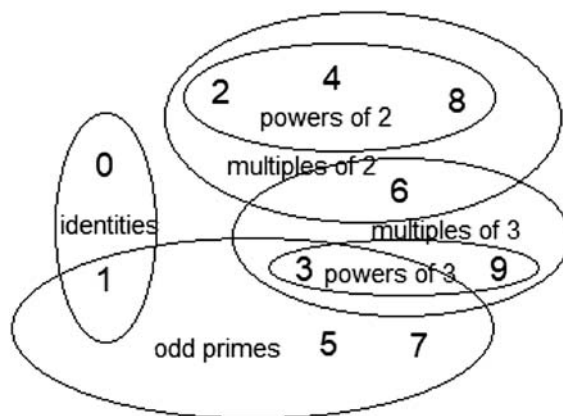


Fig. 5 Six-cluster additive clustering (ADCLUS) solution reported by Shepard et al. (1975) for the similarity ratings among the numbers 0–9. Cluster interpretations are given within each loop.

Conclusions

Techniques for cluster analysis are useful exploratory techniques for data reduction and data exploration. As the applications in this article suggest, they can also be useful for theory discovery and theory confirmation in educational research. A broad and deep understanding of cluster analysis models and techniques is helpful in ensuring that these methods are used appropriately, and that the results of such analyses are correctly interpreted.

One fundamental issue is the choice of an appropriate clustering model/method, a choice that is seldom explicitly discussed. In the educational literature, techniques for partitioning and for hierarchical clustering are widely used. However, as noted, sometimes hierarchical clustering techniques are used merely as a quick and convenient way of identifying partitions. Techniques for fitting overlapping sets of clusters have been less widely available, and arguably underused in educational research. However, as these techniques become better known and the software for fitting them become more widely available, they will perhaps become more popular.

Other issues that researchers should be aware of when using cluster analysis include proper preparation of the data, choice of software, and implementation parameters for the analysis. While best practices have been briefly outlined here, these issues have been more deeply explored elsewhere in the methodological literature. Cluster analysis techniques are recommended as useful technique in educational research, useful (and promising) enough to justify these continuing development efforts.

References

- Arabie, P., Carroll, J.D., 1980. MAPCLUS: a mathematical programming approach to fitting the ADCLUS model. *Psychometrika* 45, 211–235.
- Bauer, J., 2021. A Primer to Latent Profile and Latent Class Analysis. Downloaded 12/29/2021. [Psyarxiv.com](https://psyarxiv.com).
- Beller, M., 1990. Tree versus geometric representation of tests and items. *Appl. Psychol. Meas.* 14 (1), 13–28. <https://doi.org/10.1177/014662169001400102>.
- Carroll, J.D., Corter, J.E., 1995. A graph-theoretic method for organizing overlapping clusters into trees, multiple trees, or extended trees. *J. Classif.* 12, 283–314.
- Chiu, C.Y., Douglas, J.A., Li, X., 2009. Cluster analysis for cognitive diagnosis: theory and applications. *Psychometrika* 74 (4), 633–665.
- Cleuziou, G., 2008. An extended version of the k-means method for overlapping clustering. In: 2008 19th International Conference on Pattern Recognition. IEEE, pp. 1–4.
- Corter, J.E., Tversky, A., 1986. Extended similarity trees. *Psychometrika* 51, 429–451.
- Corter, J.E., 1982. ADDTREE/P: a PASCAL program for fitting additive trees based on Sattath & Tversky's ADDTREE algorithm. *Behav. Res. Methods. Instrum.* 14, 353–354.
- Corter, J.E., 1995. Using clustering methods to explore the structure of diagnostic tests. In: Nichols, P., Chipman, S., Brennan, R. (Eds.), *Cognitively Diagnostic Assessment*. Lawrence Erlbaum Associates, Hillsdale NJ, pp. 305–326.
- Corter, J.E., 1996. Tree models of similarity and association. Sage University. In: *Papers Series: Quantitative Applications in the Social Sciences*, Series No. 07-112. Sage, Thousand Oaks, CA.
- Corter, J.E., 1998. An efficient metric combinatorial algorithm for fitting additive trees. *Multivariate Behav. Res.* 33 (2), 249–271.
- de Kruif, R.E., McWilliam, R.A., Ridley, S.M., Wakely, M.B., 2000. Classification of teachers' interaction behaviors in early childhood classrooms. *Early Child. Res. Q.* 15 (2), 247–268.
- De Soete, G., 1983. A least squares algorithm for fitting additive trees to proximity data. *Psychometrika* 48, 621–626. <https://doi.org/10.1007/BF02293884>.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.* 39 (1), 1–22.
- Dorman, J.P., Aldridge, J.M., Fraser, B.J., 2006. Using students' assessment of classroom environment to develop a typology of secondary school classrooms. *Int. Educ. J.* 7 (7), 906–915.
- Dutt, A., Ismail, M.A.B., Herawan, T., 2017. A systematic review on educational data mining. In: IEEE Access. Digital Object Identifier. <https://doi.org/10.1109/ACCESS.2017.2654247>.
- Forbes, C.T., Neumann, K., Schiepe-Tiska, A., 2020. Patterns of inquiry-based science instruction and student science achievement in PISA 2015. *Int. J. Sci. Educ.* 42 (5), 783–806.
- Hennig, C., 2007. Cluster-wise assessment of cluster stability. *Comput. Stat. Data Anal.* 52 (1), 258–271. <https://doi.org/10.1016/j.csda.2006.11.025>. ISSN 0167-9473.
- Hubert, L., Arabie, P., 1995. Iterative projection strategies for the least squares fitting of tree structures to proximity data. *Br. J. Math. Stat. Psychol.* 48, 281–317. <https://doi.org/10.1111/j.20448317.1995.tb01065.x>.
- Johnson, S.C., 1967. Hierarchical clustering schemes. *Psychometrika* 32 (3), 241–254.
- MacQueen, J., 1967. Some methods of classification and analysis of multivariate observations. In: Le Cam, L.M., Neyman, J. (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. University of California Press, Berkeley, CA, pp. 281–297.
- McCutcheon, A.L., 1987. Latent class analysis. Sage University. In: *Papers Series: Quantitative Applications in the Social Sciences*, Series No. 07-064. Sage, Thousand Oaks CA.
- Milligan, G.W., Cooper, M.C., 1988. A study of standardization of variables in cluster analysis. *J. Classif.* 5, 181–204.
- Mirkin, B.G., 1987. The method of principal clusters. *Autom. Rem. Control* 10, 131–143.
- Moos, R.H., 1978. A typology of junior high and high school classrooms. *Am. Educ. Res. J.* 15 (1), 53–66. <https://doi.org/10.3102/00028312015001053>.
- Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J. Classif.* 31, 274–295.
- Ng, B.L., Liu, W.C., Wang, J.C., 2016. Student motivation and learning in mathematics and science: a cluster analysis. *Int. J. Sci. Math. Educ.* 14 (7), 1359–1376.
- Olsen, R.V., Lie, S., 2011. Profiles of students' interest in science issues around the world: analysis of data from PISA 2006. *Int. J. Sci. Educ.* 33 (1), 97–120.
- Revelle, W., 1979. Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behav. Res.* 14 (1), 57–74.
- Rodrigo, M.M.T., Anglo, E.A., Sugay, J.O., Baker, R., 2008. Use of unsupervised clustering to characterize learner behaviors and affective states while using an intelligent tutoring system. *Int. Conf. Comput. Educ.* 57–64.
- Romera, C., Ventura, S., 2007. Educational data mining: a survey from 1995 to 2005. *Expert Syst. Appl.* 33, 135–146.
- Roux, M., 1988. Techniques of approximation for building two tree structures. In: Hayashi, C., Diday, E., Jambu, M., Ohsumi, N. (Eds.), *Recent Developments in Clustering and Data Analysis*. Academic Press, New York, pp. 151–170.
- Rowley, S.J., 2000. Profiles of African American college students' educational utility and performance: a cluster analysis. *J. Black Psychol.* 26 (1), 3–26.
- Sattath, S., Tversky, A., 1977. Additive similarity trees. *Psychometrika* 42, 319–345.
- Sebestyen, G.S., 1962. *Decision Making Processes in Pattern Recognition*. Macmillan, New York.
- Shepard, R.N., Arabie, P., 1979. Additive clustering: representation of similarities as combinations of discrete overlapping properties. *Psychol. Rev.* 86, 87–123.
- Shepard, R.N., Kilpatrick, D.W., Cunningham, J.P., 1975. The internal representation of numbers. *Cognit. Psychol.* 7, 82–138.
- Sokal, R.R., Sneath, P.H.A., 1963. *Principles of Numerical Taxonomy*. W. H. Freeman, San Francisco.

- Speece, D.L., McKinney, J.D., Appelbaum, M.I., 1985. Classification and validation of behavioral subtypes of learning-disabled children. *J. Educ. Psychol.* 77 (1), 67–77.
- Steinley, D., 2006. K-means clustering: a half-century synthesis. *Br. J. Math. Stat. Psychol.* 59, 1–34.
- Thorndike, R.L., 1953. Who belongs in the family? *Psychometrika* 18 (4), 267–276.
- Trivedi, S., Pardos, Z.A., Heffernan, N.T., 2011. June. Clustering students to generate an ensemble to improve standard test score predictions. In: *International Conference on Artificial Intelligence in Education*. Springer, Berlin, Heidelberg, pp. 377–384.
- Van Mechelen, I., Bock, H.H., De Boeck, P., 2004. Two-mode clustering methods: a structured overview. *Stat. Methods Med. Res.* 13 (5), 363–394.
- Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244.
- Wilderjans, T.F., Ceulemans, E., Van Mechelen, I., Depril, D., 2011. ADPROCLUS: a graphical user interface for fitting additive profile clustering models to object by variable data matrices. *Behav. Res. Methods*, 43 (1), 56–65.
- Wilderjans, T.F., Depril, D., Van Mechelen, I., 2013. Additive biclustering: a comparison of one new and two existing ALS algorithms. *J. Classif.* 30, 56–74. <https://doi.org/10.1007/s00357-013-9120-0>.

Further reading

- Aldenderfer, M.S., Blashfield, R.K., 1984. Cluster analysis. Sage University. In: *Papers Series: Quantitative Applications in the Social Sciences*, Series No. 07-44. Sage, Thousand Oaks, CA. <https://doi.org/10.4135/9781412983648>.
- Jain, A.K., Dubes, R.C., 1988. *Algorithms for Clustering Data*. Prentice-Hall, Inc.
- Kaufman, L., Rousseeuw, P.J., 2009. *Finding Groups in Data: An Introduction to Cluster Analysis*, vol. 344. John Wiley & Sons.
- Sneath, P.H.A., Sokal, R.R., 1973. *Numerical Taxonomy*. Freeman, San Francisco.

Relevant websites

- NETLIB Repository of Statistical Software, <https://www.netlib.org>.
- Source and Executable Code for ADDTREE/P, GTREE, EXTREE, www.columbia.edu/~jec34.
- The AdProClus Program for Additive Clustering, <https://ppw.kuleuven.be/okp/software/adproclusstandalone/>.
- The R Project for Statistical Computing, <https://www.r-project.org/>.

Latent class analysis

Jeroen K. Vermunt, Department of Methodology and Statistics, Tilburg University, Tilburg, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

This is an update of J.K. Vermunt, *Latent Class Models*, Editor(s): Penelope Peterson, Eva Baker, Barry McGaw, *International Encyclopedia of Education* (Third Edition), Elsevier, 2010, Pages 238–244, ISBN 9780080448947, <https://doi.org/10.1016/B978-0-08-044894-7.01340-3>.

Introduction	637
Simple LC and mixture models	637
Generalized linear models for item probabilities/means	639
Some restricted models for categorical items	640
Models with explanatory variables	640
Three-step latent class analysis	641
Extensions	641
Maximum likelihood estimation	642
Model selection issues	643
Software	644
References	644

Introduction

A statistical model can be called a latent class (LC) or mixture model if it assumes that some of its parameters differ across unobserved subgroups, latent classes, or mixture components. This rather general idea has several seemingly unrelated applications, the most important of which are clustering, scaling, density estimation, and random-effects modeling. It should be noted that in applied fields the terms LC model and mixture model are often used interchangeably, which is also what I will do here. In the more technical statistical literature on mixture modeling, the term LC analysis is reserved for a specific type of mixture model, that is, a mixture model for a set of categorical items (for the classical LC model).

LC analysis was introduced in 1950 by Lazarsfeld as a tool for building typologies (or clustering) based on dichotomous observed variables. More than 20 years later, Goodman (1974) made this model applicable in practice by developing an algorithm for obtaining maximum likelihood estimates of the model parameters, as well as proposed extensions for polytomous manifest variables and did important work on the issue of model identification. Many important extensions of this classical LC model have been proposed since then, such as models containing explanatory variables (Dayton and Macready, 1988), models that relax the local independence assumption (Hagenaars, 1988), constrained models similar to IRT models (Lindsay et al., 1991; Heinen, 1996), models with multiple latent variables (Magidson and Vermunt, 2001), models for longitudinal data (Van de Pol and Langeheine, 1990; Collins and Lanza, 2010), models for multilevel data (Vermunt, 2003, 2010b), three-step LC analysis (Bolck et al., 2004; Vermunt, 2010a), and LC tree models (Van den Bergh et al., 2018).

Whereas the classical LC model and its extensions are conceived primarily as a clustering and scaling tool for categorical data analysis, LC and finite mixture models can be useful in several other areas as well. One of these is as a probabilistic cluster analysis tool for continuous observed variables, an approach that offers many advantages over traditional cluster techniques such as K-means clustering (Wolfe, 1970; McLachlan and Peel, 2000; Vermunt and Magidson, 2002). Another application area is dealing with unobserved heterogeneity, as happens in mixture regression analysis of multilevel or repeated measurement data (Wedel and De Sarbo, 1994; Vermunt and Van Dijk, 2001).

The remainder of this article is organized as follows. After introducing the simplest type of LC models, I discuss various restricted LC models as well as models with explanatory variables. Next, I give an overview of other types of LC and mixture models, which includes various recently proposed extensions. The remaining sections focus on parameter estimation, model selection, and software. An annotated list of further reading is provided at the end of the chapter.

Simple LC and mixture models

LC analysis is typically used as a tool for analyzing multivariate response data; that is, data consisting of several dependent variables, response variables, or items. We will denote the response of subject i on dependent variable j by y_{ij} , and the number of dependent variables by J . The full response vector of a subject is denoted by y_i . To make things more concrete, Table 1 presents a small illustrative data set consisting of three dichotomous responses, y_{i1} , y_{i2} , and y_{i3} ($0 = \text{incorrect}$; $1 = \text{correct}$). This is a subset of items from a mathematics test administered to 2156 children; that is, the first three items (out of a total of 18) from the Latent GOLD demo data file “cito.dat.” The column “Frequency” contains the observed frequency count for each of the eight possible answer patterns.

Table 1 Small data set with three dichotomous responses.

y_{i1}	y_{i2}	y_{i3}	Frequency	$f(\mathbf{y}_i)$	$f(\mathbf{y}_i v_i = 1)$	$f(\mathbf{y}_i v_i = 2)$	$f(v_i = 1 \mathbf{y}_i)$	$f(v_i = 2 \mathbf{y}_i)$	Modal
0	0	0	239	0.111	0.004	0.272	0.020	0.980	2
0	0	1	101	0.047	0.010	0.102	0.128	0.872	2
0	1	0	283	0.131	0.038	0.271	0.175	0.825	2
0	1	1	222	0.103	0.104	0.102	0.605	0.395	1
1	0	0	105	0.049	0.020	0.092	0.248	0.753	2
1	0	1	100	0.046	0.054	0.035	0.703	0.297	1
1	1	0	348	0.161	0.208	0.091	0.774	0.226	1
1	1	1	758	0.352	0.562	0.034	0.961	0.039	1

Note: These are the first three items (out of a total of 18) from the Latent GOLD demo data file called "cito.dat."

In addition to the J observed dependent variables, a LC model contains a discrete latent variable. We will denote a subject's unobserved score on this latent variable by v_i , the number of LCs by C , and a particular class by c , where $c = 1, 2, \dots, C$. The aim of a LC analysis of the data set in **Table 1** could be to classify pupils into two groups, masters and non-masters, which differ with respect to the probability of answering the test items correctly. The results obtained with a 2-class model will be used to illustrate the various components of a LC model.

LC analysis defines a model for $f(\mathbf{y}_i)$, the probability density of the multivariate response vector $\mathbf{y}_i$. In the above example this is the probability of answering the items according to one of the eight possible response patterns, for example, of answering the first two items correctly and the last one incorrectly ($y_{i1} = 1, y_{i2} = 1, y_{i3} = 0$), which as can be seen in **Table 1** equals 0.161 for the estimated 2-class model. The assumption underlying any type of LC or mixture model is that the density $f(\mathbf{y}_i)$ is a weighted average (or mixture) of the C class-specific densities $f(\mathbf{y}_i|v_i = c)$. This is expressed mathematically as follows:

$$f(\mathbf{y}_i) = \sum_{c=1}^C P(v_i = c) f(\mathbf{y}_i|v_i = c) \quad (1)$$

Here, $P(v_i = c)$ denotes the probability that a subject belongs to LC c . For our small empirical example, the estimates of these (prior) class membership probabilities are 0.601 and 0.399 for class 1 and 2, respectively (see **Table 2**). The assumed mechanism by **Eq. (1)** is that each individual belongs to one of C exhaustive and mutually exclusive classes with probability $P(v_i = c)$ and that given membership of LC c an individual provides responses according to the probability density associated to this class. **Table 1** shows the estimated values of $f(\mathbf{y}_i|v_i = 1)$ and $f(\mathbf{y}_i|v_i = 2)$ for our data sets. As can be seen, LC 1 has higher probabilities for the response patterns with 2 or 3 correctly answered items, whereas LC 2 has higher probabilities for the response patterns with 0 or 1 item correct.

The classical LC model combines the assumption of **Eq. (1)** shared by all mixture models with the assumption of local independence. Local independence means that the J responses are mutually independent given a subject's class membership. It can be expressed as follows:

$$f(\mathbf{y}_i|v_i = c) = \prod_{j=1}^J f(y_{ij}|v_i = c) \quad (2)$$

Independence implies that the joint density $f(\mathbf{y}_i|v_i = c)$ is obtained as a product of the J item-specific densities $f(y_{ij}|v_i = c)$. In our example, $f(y_{ij} = 1|v_i = c)$, is the class-specific probability of giving a correct answer to item j . As reported in **Table 2**, for a subject belonging to the first LC, these equal 0.844, 0.912, and 0.730 for items 1, 2 and 3, respectively. The local independence assumption implies, for example, that the probability of answering the first two items correctly and the last one incorrectly for someone in LC 1 equals $0.844 \times 0.912 \times (1 - 0.730) = 0.208$. Note that the local independence assumption is also used in other types of latent variables models, such as in factor analysis and IRT modeling, and is thus not specific for LC analysis.

Combining the two basic **Eqs. (1) and (2)** yields the following model for $f(\mathbf{y}_i)$:

Table 2 Parameters (class proportions and probability of a correct answer) obtained with a 2-class model for data in **Table 1**.

	$c = 1$	$c = 2$
$P(v_i = c)$	0.601	0.399
$\pi_{11c} = P(y_{i1} = 1 v_i = c)$	0.844	0.252
$\pi_{21c} = P(y_{i2} = 1 v_i = c)$	0.912	0.499
$\pi_{31c} = P(y_{i3} = 1 v_i = c)$	0.730	0.273

$$f(\mathbf{y}_i) = \sum_{c=1}^C P(v_i = c) \prod_{j=1}^J f(y_{ij} | v_i = c) \quad (3)$$

To complete the model specification, we need to define the form of the conditional densities $f(y_{ij} | v_i = c)$. In the classical LC model for categorical items these are multinomial probability densities; that is,

$$f(y_{ij} | v_i = c) = \prod_{r=0}^{R_j-1} \pi_{jrc}^{y_{ijr}^*}$$

where R_j is the number of categories of item j , $0 \leq y_{ij} \leq R_j - 1$, and $y_{ijr}^* = 1$ if $y_{ij} = r$ and 0 otherwise. Note this is a slightly complicated, but mathematically elegant, way to express that someone in LC c has a probability equal to $\pi_{jrc} = P(y_{ij} = r | v_i = c)$ of giving response r to item j . In the special case of a dichotomous response, the multinomial distribution reduces to the Bernoulli distribution with success probability $\pi_{jc} = \pi_{j1c} = P(y_{ij} = 1 | v_i = c)$. **Table 2** presents these probabilities for our small example.

It is important to note that LC models can not only be used with categorical responses, but also with continuous responses and counts. The density $f(y_{ij} | v_i = c)$ could be a binomial, Poisson, or negative binomial distribution for counts, and a normal or gamma distribution for continuous responses. The mixture model for continuous response variables is sometimes referred to as the latent profile model. The parameters of this model are the class proportions and class-specific item means and variances (μ_{jc} and σ_{jc}^2).

By comparing the J sets of item parameters across classes, one can name the classes. The parameter estimates presented in **Table 2** show that the first class can be named the masters because pupils belonging to that class have much higher probabilities of answering the test items correctly than pupils belonging to the second non-masters class.

Similar to cluster analysis, one of the purposes of LC analysis might be to assign individuals to LCs. The probability of belonging to LC c given responses $\mathbf{y}_i$ —often referred to as posterior membership probability—can be obtained by the Bayes rule:

$$P(v_i = c | \mathbf{y}_i) = \frac{P(v_i = c) f(\mathbf{y}_i | v_i = c)}{f(\mathbf{y}_i)} \quad (4)$$

Table 1 reports $P(v_i = c | \mathbf{y}_i)$ for each answer pattern. For example, $P(v_i = 1 | \mathbf{y}_i)$ equals 0.774 for the (1,1,0) pattern, which is obtained as $0.601 \times 0.208 / 0.161$. The most common classification rule is modal assignment, which amounts to assigning each individual to the LC with the highest $f(v_i = c | \mathbf{y}_i)$. The last column of **Table 1** reporting the modal assignments shows that pupils with at least 2 correct answers are assigned to class 1 and the others to class 2.

In the introduction we stated that mixture models are statistical models in which parameters are assumed to differ across LCs. But what is the statistical model used in the simple LC models discussed so far? It is the independence model: we assume responses to be independent, with different parameter values for each class. Depending on the scale type of the response variables, these parameters are Bernoulli probabilities, multinomial probabilities, normal means and variances, Poisson rates, etc.

Generalized linear models for item probabilities/means

Haberman (1979) showed that the LC model for categorical response variables can also be specified as a log-linear model for an expanded table, including the latent variable v_i as an additional dimension. Using such a log-linear specification is equivalent to parameterizing the response probability for item j as follows:

$$\log\left(\frac{\pi_{jrc}}{\pi_{j0c}}\right) = \log\left(\frac{P(y_{ij} = r | v_i = c)}{P(y_{ij} = 0 | v_i = c)}\right) = \alpha_{jr} + \beta_{jcr} \quad (5)$$

for $1 \leq r \leq R_j - 1$; that is, as a multinomial logistic regression model with intercepts α_{jr} and slopes β_{jcr} (note that we use the first item category, $r = 0$, as baseline). One identification constraint needs to be imposed, for example, $\beta_{j1r} = 0$ (the parameters for class 1 are fixed to 0) or $\alpha_{jr} = 0$ (intercepts are fixed to 0).

For dichotomous responses and binomial counts, the regression model could be a binary logit or probit model, for Poisson counts a log-linear model, and for continuous responses a standard linear model. These are generalized linear models (GLMs) of the form

$$g[E(y_{ij} | v_i = c)] = \alpha_j + \beta_{jc} \quad (6)$$

where $g[\cdot]$ is the link function transforming the expected value of y_{ij} to the linear term. For ordinal polytomous variables, one may use an ordinal regression model, such as an adjacent-category or cumulative logit model. These are models that restrict the item response probabilities π_{jrc} .

Some restricted models for categorical items

Many interesting types of restricted LC models for categorical items have been proposed which involve imposing (linear) constraints on either the conditional probabilities π_{jrc} or the logit coefficients of Eq. (5). One of these are probabilistic Guttman scaling models for dichotomous responses, which are LC models with $C = J + 1$ classes, one for each possible total score. The idea is that apart from measurement error, class c should provide a positive (correct) answer to the $c-1$ easiest items and a negative (incorrect) answer to the remaining $J-(c-1)$ items. The various types of probabilistic Guttman models differ in constraints they impose on the measurement error. The simplest and most restricted model is the Proctor (1970) model. Table 3 presents the parameter estimates obtained when fitting the Proctor model to the data set in Table 1. As can be seen, the probability of a correct response is either 0.833 or $0.167 = 1 - 0.833$. The measurement error—or the probability of giving a response which is not in agreement with the class—is estimated to be equal to 0.167. Whereas the Proctor model assumes that the measurement error is constant across items and classes, less restricted models can be defined which allow the error probabilities to differ across items, classes, or both (see, e.g., Dayton, 1999). Note that these equality constraints on the error probabilities can also be defined using linear constraints on the logit parameters: $\alpha_{j1} = 0$ and $\beta_{j1c} = -\beta^*$ for $c \leq j$ and $\beta_{jc} = \beta^*$ otherwise.

Croon (1990) proposed a restricted LC model that similar to non-parametric IRT (Sijtsma and Molenaar, 2002) assumes monotonic item response functions; that is, $\pi_{j1c} \leq \pi_{j1,c+1}$, or, equivalently, $\beta_{j1c} \leq \beta_{j1,c+1}$. A more restricted version, in which not only classes but also items are ordered, is obtained by imposing the additional set of restriction $\pi_{j+11c} \leq \pi_{j1c}$; that is, by assuming double monotony. Vermunt (2001) discussed various generalizations of these models.

Various authors described the connection between restricted LC analysis and parametric IRT modeling (see, for example, Heinen, 1996; Lindsay et al., 1991); that is, IRT models with a discrete specification of the distribution of the underlying trait or ability can be defined as LC models with restrictions on the logistic parameters. The key restriction is $\beta_{jrc} = \beta_{jr}^* \cdot \theta_c$ for nominal items and $\beta_{jrc} = \beta_j^* \cdot r \cdot \theta_c$ for ordinal items, where the θ_c are LC locations representing the C possible values of the discretized latent trait. These locations may be fixed a priori, for example, at $-2, -1, 0, 1$, and 2 in the case of $C = 5$, but may also be treated as free parameters to be estimated. Depending on whether the items are dichotomous, ordinal, or nominal, this yields a 2-parameter logistic, generalized partial credit, or nominal response model. Further restrictions involve equating β_j^* across items, yielding Rasch and partial credit models, and imposing across category and across item restrictions on α_{jr} parameters as in rating scale models for ordinal items.

Models with explanatory variables

The most important extension of the LC models discussed so far is the possibility to include explanatory variables (covariates) affecting the responses (Wedel and DeSarbo, 1994) or the class memberships (Dayton and Macready, 1988). Denoting the vector with explanatory variables for subject i by $\mathbf{x}_i$, the LC model of interest can be formulated as follows:

$$f(\mathbf{y}_i | \mathbf{x}_i) = \sum_{c=1}^C P(v_i = c | \mathbf{x}_i) \prod_{j=1}^{J_i} f(y_{ij} | v_i = c, \mathbf{x}_{ij}) \quad (7)$$

The main difference compared to the model defined in Eq. (3) is that now we have a model for $f(\mathbf{y}_i | \mathbf{x}_i)$ —the conditional density of $\mathbf{y}_i$ given $\mathbf{x}_i$.

Similar to the regression models for the response variables introduced in Eqs. (5) and (6), one can define a mixture regression model with explanatory variables; i.e.,

$$g\left[E(y_{ij} | v_i = c, \mathbf{x}_{ij})\right] = \alpha_c + \sum_{p=1}^P \beta_{pc} \mathbf{x}_{ijp} \quad (8)$$

As before, y_{ij} may refer to the response on item j by pupil i , in which case the explanatory variables in a long (two-level) format data file will consist of a design matrix defining the item parameters. For our small example with 3 items, the data file will contain 3 records per person—one per item. Each record has 4 columns with the first column containing an identifier variable linking the records, the second column indicating the response to a single item, and 2 additional columns for dummy-coded predictors to

Table 3 Parameters (class proportions and probability of a correct answer) obtained with Proctor model for data in Table 1.

	$c = 1$	$c = 2$	$c = 3$	$c = 4$
$P(v_i = c)$	0.160	0.155	0.126	0.559
$\pi_{21c} = P(y_{12} = 1 v_i = c)$	0.167	0.833	0.833	0.833
$\pi_{11c} = P(y_{11} = 1 v_i = c)$	0.167	0.167	0.833	0.833
$\pi_{31c} = P(y_{13} = 1 v_i = c)$	0.167	0.167	0.167	0.833

indicate which item the response is for. Taking the third item as the reference item, the first predictor takes on the value 1 for a record containing the first item response and 0 for the other records, and the second predictor equals 1 for the second response and 0 for the other two responses.

However, the model in Eq. (8) can also be used for many other purposes. In fact, it is a model for analyzing two-level data sets, where regression parameters are allowed to differ across LCs (of higher-level units). For example, γ_{ij} could be the test score of pupil j belonging to school i , and $\mathbf{x}_{ij}$ a set of pupil characteristics (e.g., IQ). A mixture regression model would identify LCs of schools with different intercepts and different effects of child characteristics on the test scores. Another possible application is in the analysis of longitudinal data, where j is a time point for subject i , and where vector $\mathbf{x}_{ij}$ contains time variables. This yields a LC growth model in which subjects are grouped based on their developmental trajectories (Vermunt, 2007). A fourth possible application is in experiments in which subjects are observed in multiple conditions, such as in conjoint studies. The mixture regression model can be used to group subjects based on their reactions on the experimental conditions. In fact, in each of these application types, the LC model is used as a random-coefficient model without parametric assumptions about the distribution of the random effects (Aitkin, 1999; Vermunt and Van Dijk, 2001).

As shown in Eq. (7), an individual's class membership may also be predicted using covariates. This is achieved by defining a multinomial logistic regression model for $P(v_i = c | \mathbf{x}_i)$:

$$\log \frac{P(v_i = c | \mathbf{x}_i)}{P(v_i = 1 | \mathbf{x}_i)} = \gamma_{0c} + \sum_{q=1}^Q \gamma_{pc} x_{iq}$$

Strongly related are multiple-group LC models (Hagenaars and McCutcheon, 2002; Magidson and Vermunt, 2004; Kankaraš et al., 2010). These can be defined by using the grouping variable as a nominal explanatory in the model.

Three-step latent class analysis

Rather than including covariates directly within the estimated LC model, one may also use the following type of three-step approach:

1. Perform model selection and estimation in the usual way, thus without the inclusion of covariates;
2. Obtain class assignments w using the selected model from step 1, as explained in Eq. (4);
3. Perform subsequent analyses with covariates or other types of external variables using the class assignments w of step 2.

Although this stepwise approach of separating the LC analysis from the analyses one would like to do after the latent classes are constructed is very practical and intuitive, it is also problematic. More specifically, as a result of the classification errors introduced in step 2, it yields underestimated associations between external variables and latent classes. The larger the classification errors, the larger the bias in the estimates of these associations.

However, building upon the work by Bolck et al. (2004), Vermunt (2010a) proposed a solution to this problem. He showed how to perform a valid step-3 analysis by adjusting for the classification errors introduced in step 2. Basically, a new LC model is estimated in which the class assignments w are used as the single indicator with known conditional response probabilities $P(w | v_i = c)$. This adjusted step-3 analysis can not only be used with covariates predicting class membership (via a logistic model), but also for investigating how classes differ with respect to a distal outcome variable (Bakk et al., 2013). Bakk and Vermunt (2016) recommended using the more robust BCH adjustment method for continuous distal outcomes (dependent variables), while for covariates and categorical distal outcomes the maximum likelihood (ML) approach is preferred.

Extensions

The most common model-fitting strategy in LC analysis is to increase the number of classes until the local independence assumption holds. This may, however, yield solutions which are difficult to interpret. One alternative approach is to relax the local independence assumption by allowing for associations between particular item pairs. Hagenaars (1988) showed how to define LC models with local dependencies for categorical responses. With continuous responses this is easily achieved by using multivariate instead of univariate normal distributions for locally dependent items (see, e.g., McLachlan and Peel, 2000; Vermunt and Magidson, 2002).

Another alternative strategy involves increasing the number of discrete latent variables instead of the number of LCs, which is especially useful if the items measure several dimensions, such as different math subskills. This so-called discrete factor modeling approach (Magidson and Vermunt, 2001) is a special case of the path modeling approach for discrete latent variables developed by Hagenaars (1990), Vermunt (1997). Many other interesting models can be defined within this framework, such as latent Markov (or transition) models for the analysis of longitudinal data (Collins and Lanza, 2010; Van de Pol and Langeheine, 1990; Vermunt et al., 2008) and LC models for cognitive diagnosis (De la Torre and Douglas, 2004).

Another interesting extension is the LC tree approach (Van den Bergh et al., 2018), in which after starting with a small number of classes at the root of the tree, a hierarchical structure of mutually linked classes is obtained by sequentially splitting classes into two

subclasses as long as the model fit improves. This approach, which is similar to (divisive) hierarchical clustering, can also be combined with mixture growth modeling and three-step LC analysis.

Various types of models have been developed that contain both discrete and continuous latent variables, examples of which include mixture factor models (Yung, 1997; McLachlan and Peel, 2000), mixture structural equation models (Dolan and Van der Maas, 1997), and mixture IRT models (Rost, 1990). Recently, extensions of mixture factor modeling have been proposed for detecting sources of measurement non-invariance in intensive longitudinal data, multiple-group data for many groups, and multi-level data (De Roover et al., 2017, 2022; Vogelsmeier et al., 2019).

Another important extension is the multilevel LC model (Vermunt, 2003, 2010b). One of its variants is a model with discrete latent variables at multiple levels of a hierarchical structure: e.g., children belong to LCs with different performances on a set of test items, and schools belong to LCs with different distributions of children across the child-level performance classes. Multilevel LC models can be used for the analysis of two-level multivariate and three-level univariate response data.

Lanza et al. (2013), Clouth et al. (2022) showed how to combine LC analysis with tools for causal inference in observational (non-experimental) studies. More specifically, Lanza et al. (2013) illustrated the use of inverse propensity weighting and matching based on propensity scores to determine the effect of a treatment on class membership. As an alternative, Clouth et al. (2022) proposed using a three-step LC approach with inverse propensity weighting.

Maximum likelihood estimation

The parameters of LC models are typically estimated by means of maximum likelihood (ML). The log-likelihood function that is maximized is based on the probability densities defined in Eqs. (1), (2), and (3); that is,

$$\ln L = \sum_{i=1}^N \ln f(\mathbf{y}_i)$$

with categorical responses one will typically group the data and construct a frequency table as we did in Table 1. The log-likelihood function for grouped data equals

$$\ln L = \sum_{k=1}^K n_k \ln f(\mathbf{y}_k)$$

where k is a data pattern, K the number of different data patterns and n_k the cell count corresponding to data pattern k . Notice that only nonzero observed cell entries contribute to the log-likelihood function, a feature that is exploited by several more efficient LC software packages that have been developed within the past few years (see the Software section for more discussion of these packages).

One of the problems in the estimation of LC models for discrete y_{ij} is that model parameters may be nonidentified, even if the number of degrees of freedom—the number of independent cells in the J -way cross-tabulation minus the number of free parameters—is larger or equal to zero. Nonidentification means that different sets of parameter values yield the same maximum of the log-likelihood function or, worded differently, that there is no unique set of parameter estimates. The formal identification check is via the Jacobian matrix (matrix of first derivatives of $f(\mathbf{y}_i)$), which should be column full rank. Another option is to estimate the model of interest with different sets of starting values. Except for local solutions (see below), an identified model gives the same final estimates for each set of the starting values.

Although there are no general rules with respect to the identification of LC models, it is possible to provide certain minimal requirements and point to possible pitfalls. For an unrestricted LC analysis, one needs at least three responses (y_{ij} 's) per individual, but if these are dichotomous, no more than two LCs can be identified. One has to be careful with four dichotomous response variables, in which case the unrestricted three-class model is not identified, even though it has a positive number of degrees of freedom. With five dichotomous items, however, even a five-class model is identified. Usually, it is possible to achieve identification by constraining certain model parameters.

A second problem associated with the estimation of LC models is the presence of local maxima. The log-likelihood function of a LC model is not always concave, which means that hill-climbing algorithms may converge to a different maximum depending on the starting values. Usually, we are looking for the global maximum. The best way to proceed is, therefore, to estimate the model with different sets of random starting values. Typically, several sets converge to the same highest log-likelihood value, which can then be assumed to be the ML solution. Current LC analysis software packages have automated the use of multiple sets of random starting values to reduce the probability of getting a local solution.

Another problem in LC modeling is the occurrence of boundary solutions, which are probabilities equal to 0 (or 1) or logit parameters equal to minus (or plus) infinity. These may cause numerical problems in the estimation algorithms, occurrence of local solutions, and complications in the computation of standard errors and number of degrees of freedom of the goodness-of-fit tests. Boundary solutions can be prevented by imposing constraints or by taking into account other kinds of prior information on the model parameters.

The most popular methods for solving the ML estimation problem are the expectation-maximization (EM) and Newton-Raphson (NR) algorithms. EM is a very stable iterative method for ML estimation with incomplete data. NR is a faster procedure that, however, needs good starting values to converge. The latter method makes use of the matrix of second-order derivatives of the log-likelihood function, which is also needed for obtaining standard errors of the model parameters.

Model selection issues

The goodness-of-fit of LC models for categorical responses can be tested using Pearson and likelihood-ratio chi-squared tests. The latter is defined as

$$L^2 = 2 \sum_{k=1}^K n_k \ln \frac{n_k}{N \cdot f(y_k)}$$

As in log-linear analysis, the number of degrees of freedom (df) equals the number of cells in the frequency table minus 1, minus the number of independent parameters. In an unrestricted LC model,

$$df = \prod_{j=1}^J R_j - C \cdot \left[1 + \sum_{j=1}^J (R_j - 1) \right]$$

Although it is no problem to estimate LC models with 10, 20, or 50 indicators, in such cases, the frequency table may become very sparse and, as a result, asymptotic p -values can no longer be trusted. An elegant but somewhat time-consuming solution to this problem is to estimate the p -values by parametric bootstrapping. Another option is to assess model fit in lower-order marginal tables (e.g., in the two-way marginal tables). Magidson and Vermunt (2004) refer to Pearson chi-squared statistics in two-way tables as bivariate residuals (or BVRs).

Even though models with C and $C + 1$ are nested, one cannot test them against each other using a standard likelihood-ratio ($-2 \ln L$ difference) test because it does not have an asymptotic chi-squared distribution. A solution to this problem is to approximate its sampling distribution using bootstrapping. But since this method is computationally demanding, usually alternative methods are required for comparing models with different numbers of classes. One popular method is the use of information criteria, such as the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC), and the AIC3 (a variant of the AIC which uses a penalty of 3 instead of 2 per parameter), where smaller values indicate better model fit.

Usually, we are not only interested in goodness-of-fit but also in the performance of the modal classification rule (see Eq. 4). The estimated proportion of classification errors under modal classification equals

$$E = \sum_{i=1}^N \frac{1}{N} \{1 - \max[P(v_i = c | y_i)]\}$$

This number can be compared to the proportion of classification errors based on the unconditional probabilities $P(v_i = c)$, yielding a reduction of errors measure

$$\lambda = 1 - \frac{E}{1 - \max[P(v_i = c)]}$$

The closer this nominal R^2 -type measure is to 1, the better the classification performance of a model. Other types of classification error reduction measures have been proposed based on entropy or qualitative variance (Collins and Lanza, 2010; Vermunt and Magidson, 2016).

To illustrate the use of model selection statistics for deciding about the number of classes, Table 4 presents the results for 1- to 6-class models estimated with the "cito.dat" data set using all 18 items, where p -values for the goodness-of-test and the likelihood-

Table 4 Fit measures for the LC models estimated using all 18 items of the "cito.dat" data set.

Model	BIC	AIC	AIC3	L^2	Bootstrap p -value L^2	Largest BVR	$-2 \ln L$ difference	Bootstrap p -value $-2 \ln L$ difference	Entropy R^2
1 Class	44,667.15	44,564.98	44,582.98	15,115.33	0.00	213.86			
2 Classes	41,055.03	40,845.02	40,882.02	11,357.36	0.01	43.34	3757.96	0.00	0.80
3 Classes	40,779.08	40,461.22	40,517.22	10,935.57	0.38	11.89	421.79	0.00	0.69
4 Classes	40,830.76	40,405.06	40,480.06	10,841.40	0.41	6.35	94.17	0.00	0.61
5 Classes	40,929.23	40,395.69	40,489.69	10,794.04	0.37	5.74	47.37	0.07	0.60
6 Classes	41,029.75	40,388.36	40,501.36	10,748.70	0.34	5.92	45.33	0.12	0.59

Note: The selected model is printed in bold. BIC, AIC, and AIC3 select the model with the lowest value, L^2 the simplest model with a non-significant value, largest BVR the simplest model showing a large drop compared to the previous model, and $-2 \ln L$ difference the most complex model with a significant value. Entropy R^2 should not be used for model selection.

ratio test were obtained using bootstrapping (with 500 replications). The simplest model with a non-significant L^2 goodness-of-fit value and moreover with the lowest BIC is the 3-class model, and should thus be selected according to these two statistics. The 4-class model would be selected based on the AIC3, the bootstrap likelihood-ratio ($-2 \ln L$ difference) test, and the BVR: the AIC3 is lowest, the likelihood-ratio test shows it is significantly better than the 3-class model and non-significantly worse than the 5-class model, and the largest BVR drops quite a bit when going from 3 to 4 classes, but not anymore afterward. The AIC indicates that at least 6 classes are needed, but simulation studies have shown that it tends to overestimate the number of classes.

As also happens in this example application, different criteria typically point at different “best” models, meaning that either preference for parsimony and/or content knowledge should determine the final decision, in this example, on whether to retain the 3- or 4-class model. The last column of **Table 4** reports the entropy-based R^2 value indicating how well one can predict the individuals’ class memberships based on their observed responses. As in our example, this measure typically decreases with the number of classes. It should be noted that it should not be used for model selection, but only as a measure indicating how well the selected model performs in terms of classification (similar to a reliability coefficient of a scale).

Software

One of the first LC analysis programs, MLLSA, made available by Clifford Clogg in 1977, was limited to a relatively small number of nominal variables. Today’s programs can handle many more variables, as well as other scale types. For example, the LEM program (Vermunt, 1997) provides a command language that can be used to specify a large variety of models for categorical data, including LC models. Mplus is a command language-based structural equation modeling package that implements many types of LC and mixture models. In addition, routines and packages for the estimation of specific types of LC models are available for SAS, R, and Stata (see, for example, Lanza et al., 2007; Linzer and Lewis, 2011; Skrondal and Rabe-Hesketh, 2004). Haughton et al. (2009) reviewed the Latent GOLD program and the R packages poLCA and MCLUST.

Latent GOLD (Vermunt and Magidson, 2016, 2021) is a stand-alone program that was especially developed for LC analysis, and which contains both an SPSS-like point and click user interface and a syntax language. It implements all important types of LC models, such as models for response variables of different scale types, restricted LC models, models with predictors, models with local dependencies, models with multiple discrete latent variables, path models with discrete latent variables, latent Markov/transition models, mixture regression and mixture growth models, mixture factor analysis and IRT, multilevel LC models, three-step LC modeling, and LC tree models, as well as features for dealing with partially missing data, performing bootstrapping, generating multiple imputed data sets, performing simulation studies, and providing correct inference with complex sampling designs.

References

- Aitkin, M., 1999. A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics* 55, 218–234.
- Bakk, Z., Vermunt, J.K., 2016. Robustness of stepwise latent class modeling with continuous distal outcomes. *Struct. Equ. Model.* 23, 20–31.
- Bakk, Z., Tekle, F.B., Vermunt, J.K., 2013. Estimating the association between latent class membership and external variables using bias adjusted three-step approaches. *Socio. Methodol.* 43, 272–311.
- Bolck, A., Croon, M., Hagenaars, J., 2004. Estimating latent structure models with categorical variables: one-step versus three-step estimators. *Polit. Anal.* 12, 3–27.
- Clouth, F.J., Pauws, S., Mols, F., Vermunt, J.K., 2022. A new three-step method for using inverse propensity weighting with latent class analysis. *Adv. Classif. Data Anal.* in press.
- Collins, L.M., Lanza, S.T., 2010. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Wiley, New York.
- Croon, M.A., 1990. Latent class analysis with ordered latent classes. *Br. J. Math. Stat. Psychol.* 43, 171–192.
- Dayton, C.M., Macready, G.B., 1988. Concomitant-variable latent class models. *J. Am. Stat. Assoc.* 83, 173–178.
- Dayton, C.M., 1999. *Latent Class Scaling Analysis*. Sage Publications, Thousand Oaks.
- De la Torre, J., Douglas, J., 2004. Higher-order latent trait models for cognitive diagnosis. *Psychometrika* 69, 333–353.
- De Roover, K., Vermunt, J.K., Timmerman, M., Ceulemans, E., 2017. Mixture simultaneous factor analysis for capturing differences in latent variables between higher-level units of multilevel data. *Struct. Equ. Model.* 24, 506–523.
- De Roover, K., Vermunt, J.K., Ceulemans, E., 2022. Mixture multigroup factor analysis for unraveling factor loading non-invariance across many groups. *Psychol. Methods* in press.
- Dolan, C.V., Van der Maas, H.L.J., 1997. Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika* 63, 227–253.
- Goodman, L.A., 1974. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* 61, 215–231.
- Haberman, S.J., 1979. *Analysis of Qualitative Data: Vol. 2. New Developments*. Academic Press, New York.
- Hagenaars, J.A., McCutcheon, A.L., 2002. *Applied Latent Class Analysis*. Cambridge University Press, Cambridge, UK.
- Hagenaars, J.A., 1988. Latent structure models with direct effects between indicators: local dependence models. *Socio. Methods Res.* 16, 379–405.
- Hagenaars, J.A., 1990. *Categorical Longitudinal Data: Loglinear Analysis of Panel, Trend and Cohort Data*. Sage, Newbury Park, CA.
- Haughton, D., Legrand, P., Woolford, S., 2009. Review of three latent class cluster analysis packages: latent Gold, poLCA, and MCLUST. *Am. Statistician* 63, 81–91.
- Heinen, T., 1996. *Latent Class and Discrete Latent Trait Models: Similarities and Differences*. Sage, Thousand Oaks, CA.
- Kankaraš, M., Moors, G., Vermunt, J.K., 2010. Testing for measurement invariance with latent class analysis. In: Davidov, E., Schmidt, P., Billiet, J. (Eds.), *Cross-Cultural Analysis: Methods and Applications*. Routledge, New York, pp. 359–384.
- Lanza, S.T., Collins, L.M., Lemmon, D.R., Schafer, J.L., 2007. PROC LCA: a SAS procedure for latent class analysis. *Struct. Equ. Model.* 14, 671–694.
- Lanza, S.T., Coffman, D.L., Xu, S., 2013. Causal inference in latent class analysis. *Struct. Equ. Model.* 20, 361–383.
- Lazarsfeld, P.F., 1950. The logical and mathematical foundation of latent structure analysis & the interpretation and mathematical foundation of latent structure analysis. In: Stouffer, S.A., et al. (Eds.), *Measurement and Prediction*. Princeton University Press, Princeton, NJ, pp. 362–472.
- Lindsay, B., Clogg, C.C., Grego, J., 1991. Semiparametric estimation in the Rasch model and related models, including a simple latent class model for item analysis. *J. Am. Stat. Assoc.* 86, 96–107.

- Linzer, D.A., Lewis, J.B., 2011. polCA: an R package for polytomous variable latent class analysis. *J. Stat. Software* 42 (1), 1–29.
- Magidson, J., Vermunt, J.K., 2001. Latent class factor and cluster models, bi-plots and related graphical displays. *Socio. Methodol.* 31, 223–264.
- Magidson, J., Vermunt, J.K., 2004. Latent class models. In: Kaplan, D. (Ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences*. Sage Publications, Thousand Oaks, pp. 175–198.
- McLachlan, G.J., Peel, D., 2000. *Finite Mixture Models*. John Wiley & Sons, New York.
- Proctor, C.H., 1970. A probabilistic formulation and statistical analysis of Guttman scaling. *Psychometrika* 35, 73–78.
- Rost, J., 1990. Rasch models in latent classes. An integration of two approaches to item analysis. *Appl. Psychol. Meas.* 14, 271–282.
- Sijtsma, K., Molenaar, I.W., 2002. *Introduction to Nonparametric Item Response Theory*. Sage, Thousand Oaks.
- Skrondal, A., Rabe-Hesketh, S., 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall/CRC, London.
- Van de Pol, F., Langeheine, R., 1990. Mixed Markov latent class models. *Socio. Methodol.* 20, 213–247.
- Van den Bergh, M., van Kollenburg, G.H., Vermunt, J.K., 2018. Deciding on the starting number of classes of a latent class tree. *Socio. Methodol.* 48, 303–336.
- Vermunt, J.K., Magidson, J., 2002. Latent class cluster analysis. In: Hagenaars, J., McCutcheon, A. (Eds.), *Applied Latent Class Analysis*. Cambridge University Press, pp. 89–106.
- Vermunt, J.K., Magidson, J., 2016. *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. Statistical Innovations Inc, Belmont, MA.
- Vermunt, J.K., Magidson, J., 2021. *Upgrade Manual for Latent GOLD Basic, Advanced/Syntax and Choice Version 6.0*. Statistical Innovations Inc, Arlington, MA.
- Vermunt, J.K., Van Dijk, L., 2001. A nonparametric random-coefficients approach: the latent class regression model. *Multilevel Model. Newsl.* 13, 6–13.
- Vermunt, J.K., Tran, B., Magidson, J., 2008. Latent class models in longitudinal research. In: Menard, S. (Ed.), *Handbook of Longitudinal Research: Design, Measurement, and Analysis*. Elsevier, Burlington, MA, pp. 373–385.
- Vermunt, J.K., 1997. *Log-Linear Models for Event Histories*. Sage, Thousand Oaks, CA.
- Vermunt, J.K., 2001. The use restricted latent class models for defining and testing nonparametric and parametric IRT models. *Appl. Psychol. Meas.* 25, 283–294.
- Vermunt, J.K., 2003. Multilevel latent class models. *Socio. Methodol.* 33, 213–239.
- Vermunt, J.K., 2007. Growth models for categorical response variables: standard, latent-class, and hybrid approaches. In: van Montfort, K., Oud, H., Satorra, A. (Eds.), *Longitudinal Models in the Behavioral and Related Sciences*. Erlbaum, Mahwah, NJ, pp. 139–158.
- Vermunt, J.K., 2010a. Latent class modeling with covariates: two improved three-step approaches. *Polit. Anal.* 18, 450–469.
- Vermunt, J.K., 2010b. Mixture models for multilevel data sets. In: Hox, J., Roberts, J.K. (Eds.), *The Handbook of Advanced Multilevel Analysis*. Routledge, New York, pp. 59–81.
- Vogelsmeier, L.V.D.E., Vermunt, J.K., Böing-Messing, F., De Roover, K., 2019. Continuous-time latent Markov factor analysis for exploring measurement model changes across time. *Methodology* 15 (Suppl. 1), 29–42.
- Wedel, M., DeSarbo, W., 1994. A review of recent developments in latent class regression models. In: Bagozzi, R.P. (Ed.), *Advanced Methods of Marketing Research*. Blackwell Publishers, Cambridge, MA, pp. 352–388.
- Wolfe, J.H., 1970. Pattern clustering by multivariate cluster analysis. *Multivariate Behav. Res.* 5, 329–350.
- Yung, Y.F., 1997. Finite mixtures in confirmatory factor-analysis models. *Psychometrika* 62, 297–330.

An introduction to growth mixture models (GMM)

Tae Kyoung Lee^a, Kandauda A.S. Wickrama^b, and Catherine Walker O'Neal^b, ^a Department of Social Innovation Convergence, Child Psychology and Education, Sungkyunkwan University, Seoul, South Korea; and ^b Department of Human Development and Family Science, University of Georgia, Athens, GA, United States

© 2023 Elsevier Ltd. All rights reserved.

A latent growth model	646
Potential heterogeneity in individual trajectories	648
Growth mixture modeling	649
Specifying a growth mixture model	650
Building a growth mixture model using <i>Mplus</i>	650
Specify a latent growth model (step one)	650
Estimating a growth mixture model (step two)	651
Addressing estimation problems (step three)	652
Selecting the optimal class model (step four)	653
Illustrative example: identifying the optimal model	654
Conclusion	655
References	655

Education researchers are often interested in findings heterogeneous patterns in longitudinal data. Growth mixture modeling (GMM) is a statistical approach that models such heterogeneity by classifying individuals with similar longitudinal patterns into groups (termed latent classes) (Muthén and Shedden, 1999; Wickrama et al., 2016). The purpose of this article is to illustrate the use of GMM within the field of education research via an examination of students' mental health over time (i.e., repeated measures of depressive symptoms). To accomplish this objective, we provide an overview of GMM, highlight the strengths of this analytical approach, and review specific GMM analytical strategies, in addition to providing specific step-by-step examples following the three steps of GMM model specification recommended by Wickrama et al. (2016): (a) model estimation, (b) addressing estimation problems, and (c) model selection and interpretation. We begin by briefly reviewing latent growth modeling (LGM), which serves as the measurement (or baseline) model for GMM.

A latent growth model

A latent growth model (LGM) is a type of longitudinal structural equation model that allows researchers to evaluate an individual's change (trajectory) over time (Meredith and Tisak, 1990; Wickrama et al., 2016). A LGM can be considered as two different processes capturing intra- and inter-individual change. First, a latent initial level and a latent slope (i.e., the growth factors) are estimated based on each individual's trajectory (known as intra-individual changes; Wickrama et al., 2016). Second, growth factors estimate the *average* trajectory and *variation* across individuals' trajectories (known as inter-individual variation). Therefore, a LGM allows researchers to examine individual trajectories over time and individual differences in these trajectories simultaneously. The LGM begins with the premise that a set of repeated measures are functionally related to the passage of time. For example, with a set of repeated measure of students' depressive symptom scores, a LGM can be specified as follows.

$$\text{Depressive Symptoms}_{it} = \pi_{0i} + \pi_{1i} \times t + \varepsilon_{it}, \quad \varepsilon_{it} \sim \text{NID}(0, \sigma_t^2) \quad (1)$$

where there are $i = 1, 2, \dots, n$ respondents, $t = 0, 1, 2, \dots, t-1$ timepoints (assuming linear changes), and ε_{it} = the error term for the i th individual at time t (assuming that error, ε , is normally and independently distributed with a mean of 0 and a variance of σ^2). Because each person has their own intercept (π_0) and slope (π_1), we have as many intercepts and slopes as we have study participants (n). The univariate (i.e., unconditional) version of this model can be written as:

$$\pi_{0i} = \mu_{00} + \zeta_{0i}, \quad \text{where } \zeta_{0i} \sim \text{NID}(0, \psi_{00}) \quad (2)$$

$$\pi_{1i} = \mu_{11} + \zeta_{1i}, \quad \text{where } \zeta_{1i} \sim \text{NID}(0, \psi_{11}) \quad (3)$$

$$\psi = \begin{bmatrix} \psi_{00} & \\ \psi_{10} & \psi_{11} \end{bmatrix} \quad (4)$$

where there are $i = 1, 2, \dots, n$ respondents in the study, μ_{00} and μ_{11} represent the average intercept and slope, and ζ_{0i} and ζ_{1i} indicate the errors of the estimated intercepts and slopes around their means (assuming that the errors are normally and independently

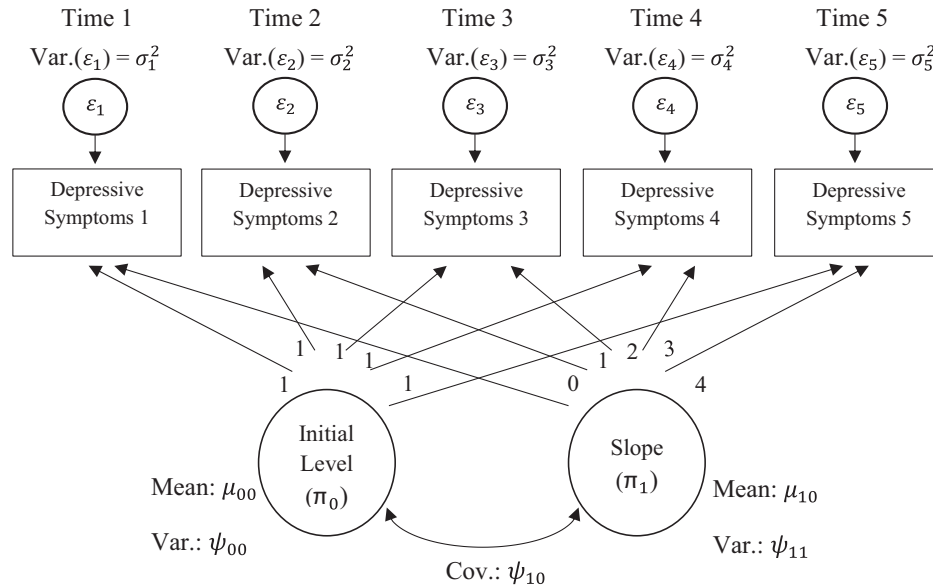


Fig. 1 A specification of linear latent growth model with five repeated measures of depressive symptoms. Note. Var. = variance. Cov. = covariance.

distributed). ψ_{00} and ψ_{11} represent the error variance of the intercept and slope, respectively. The parameter μ_{11} describes the average overall rate of change in individuals' depressive symptoms, or the mean trajectory of depressive symptoms across the sample, while the population variance for the change parameter (ψ_{11}) reflects individual differences in the rate of change. ψ_{10} denotes the covariance between π_{0i} and π_{1i} .

In a LGM, these two processes are represented through the specification of latent (i.e., unobserved) variables referred to as (a) *initial level* and (b) *slope* (rate of change). Using five repeated measures of students' depressive symptoms, **Fig. 1** shows a LGM model specification. The *initial level* variable reflects the level of depressive symptoms at the first measurement occasion (i.e., depressive symptom 1). The initial level estimate is the expected initial value (i.e., students' scores of latent initial level factor) of depressive symptoms for each individual in the sample. Because the initial level factor is constant over time, the factor loadings are fixed to 1 for the repeated measures. The latent slope factor estimates the constant (absolute) change in depressive symptoms over time (e.g., the rate of change for the repeated variable, depressive symptoms, from timepoint 1 to 5). This constant change of an attribute is sometimes termed a maturation effect to represent the effect of aging or development over time. The *slope* variable (i.e., students' scores for the latent slope factor) reflects the rate of change in depressive symptoms across the five timepoints. Time (e.g., age, day, or wave of measurement) is built into the factor loading matrix to form the hypothesized trajectory. The most common form, or shape, for a trajectory models a linear trend over time. For instance, with data from five time points equally spaced one year apart, the factor loadings for the slope are 0, 1, 2, 3, and 4 for timepoints 1 through 5, respectively (see a latent slope factor in **Fig. 1**). Note that non-linear trajectories, such as quadratic trajectories, can also be specified by adjusting the time loadings (e.g., $t = 0, 1, 4, 9$, and 16). See Wickrama et al. (2016) for more details about non-linear trajectory specification for a LGM. Residuals (ϵ) for the repeated measures reflect the unexplained variance (σ^2) of repeated outcomes (e.g., depressive symptoms) at each timepoint after accounting for effects of trajectories.

Means (μ_{00} and μ_{10} ; known as fixed effects) and variances (ψ_{00} and ψ_{11} ; known as random effects reflecting the degree of inter-individual variability, or differences, in trajectories) are estimated for the latent growth factors (initial levels and slopes). The covariance between initial level and slope, ψ_{10} , representing the (rank-order) association between the initial level and the growth (i.e., slope) is also estimated for the sample. A negative covariance between the initial level and slope factors would suggest that students with greater values at timepoint 1 (relative to other students in the study) tended to have a steeper decline (or less positive growth) in depressive symptoms over time.

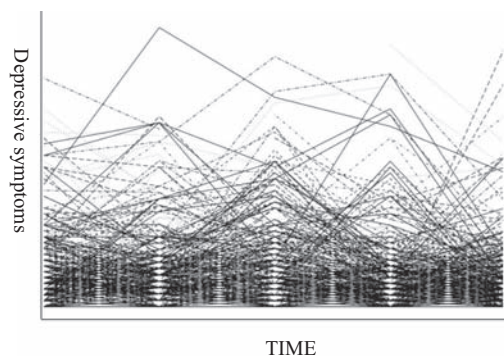
To evaluate model fit, the chi-square (χ^2) statistic is typically examined. A χ^2 with a significant p -value indicates that the hypothesized model is a poor fit to the data. However, the χ^2 statistic is sensitive to sample size and is commonly significant when the sample size is large (Little, 2013; Wickrama et al., 2016). Alternatively, other fit indices can be examined, such as the Comparative Fit Index (CFI, ≥ 0.90 acceptable fit; Little, 2013) and Root Mean Square Error of Approximation (RMSEA, ≤ 0.08 acceptable fit; Little, 2013).

Potential heterogeneity in individual trajectories

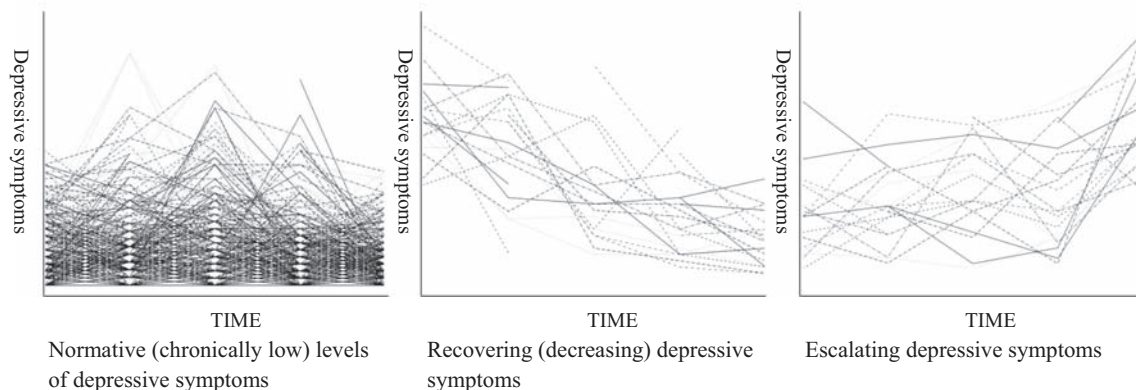
In some instances, latent growth modeling is not well-suited for studying change because of the inherent heterogeneity in the trajectories of the specific attributes examined. For example, students' health attributes are typically not uniformly distributed throughout the population. Instead, qualitatively different trajectories may exist within the population. As a result, individual trajectories of various health attributes may be heterogeneous, or clustered. For instance, classes, or groups of students, may exist with variations in depressive symptoms, including one group with consistently low and stable symptoms and another group with moderate and increasing symptoms. There may also be instances where study attributes do not have a common growth trend (e.g., linear or non-linear trajectory).

In the longitudinal context, heterogeneity, or clustering, of individual trajectories may be characterized into distinct sub-populations for two reasons. First, the heterogeneity, or clustering, of individual trajectories can be based merely on the rates of change (slopes) across individuals. That is, even though all individuals share a common growth trend/function (e.g., all respondents share a linear growth pattern), there could be three [or more] clusters based on the rate of change. For example, sub-populations with varying rates of change could include linear and increasing, linear and decreasing, and linear but constant. The heterogeneous rate of change across individuals produces different patterns resulting in unobserved variation. In considering student mental health trajectories, sub-populations may include those who are chronically mentally ill, those with deteriorating mental health, and those who are generally healthy or have improving mental health. Although all of these sub-populations share a common linear growth function, they have meaningfully different average rates of change (and/or different initial levels). These differences result in different growth patterns.

For example, research suggests several distinct sub-populations exist for youths' depressive symptoms trajectories (e.g., [Lee et al., 2017](#)). These trajectories represent normative (chronically low) levels of depressive symptoms, escalating depressive symptoms, and recovering (decreasing) depressive symptoms. Panel A of [Fig. 2](#) illustrates the observed individual trajectories of depressive symptoms for all individuals in the present sample of 436 students between the ages of 13 and 16. This figure does not show an overall trend in depressive symptoms over time. Panel B of [Fig. 2](#) shows randomly selected trajectories of depressive symptoms belonging



Panel A. Observed Individual Trajectories (Growth Model Approach)



Panel B. Unobserved Heterogeneity in Depressive Symptom Trajectories for the Consistently Low, Recovering, and Escalating Classes.

Fig. 2 Individual depressive symptom trajectories for members of each trajectory class.

to the three identified trajectory classes. The individuals in these classes exhibited different patterns of change. These distinct patterns of change suggest the assumption of a single population and/or a single growth function is an oversimplification. Instead, the results support the existence of heterogeneous growth with different growth patterns across unique sub-populations. Although it is not possible to extract distinct patterns of trajectories in a LGM, one way to examine this further is through the use of growth mixture modeling, which we discuss in detail in the next section.

Growth mixture modeling

A growth mixture model (GMM) is a specific type of LGM. A major analytic goal of growth mixture modeling is to examine distinct patterns (or Heterogeneity) in growth parameters (i.e., means and variances of initial levels and slope) over time (Wickrama et al., 2016). This approach assumes that the population consists of distinct sub-groups of trajectories (that is, a “mixture” of distributions), thereby accommodating for the existence of heterogeneous trajectories or latent classes of trajectories. Thus, using a GMM is more appropriate than a simple LGM for the investigation of individual trajectories when there is heterogeneity in the trajectories of interest (Muthén and Muthén, 2000).

As shown in Fig. 3, a GMM allows for heterogeneity by extending a LGM to incorporate latent categorical variables (C) in addition to continuous latent growth factors (initial level and slope) within the same modeling framework. (Note that C represents latent classes of individual trajectories). Thus, a GMM fully captures inter-individual variation (i.e., within-class variation) in individual growth trajectories like a LGM, while also taking into account between-class variation (i.e., unobserved heterogeneity). In

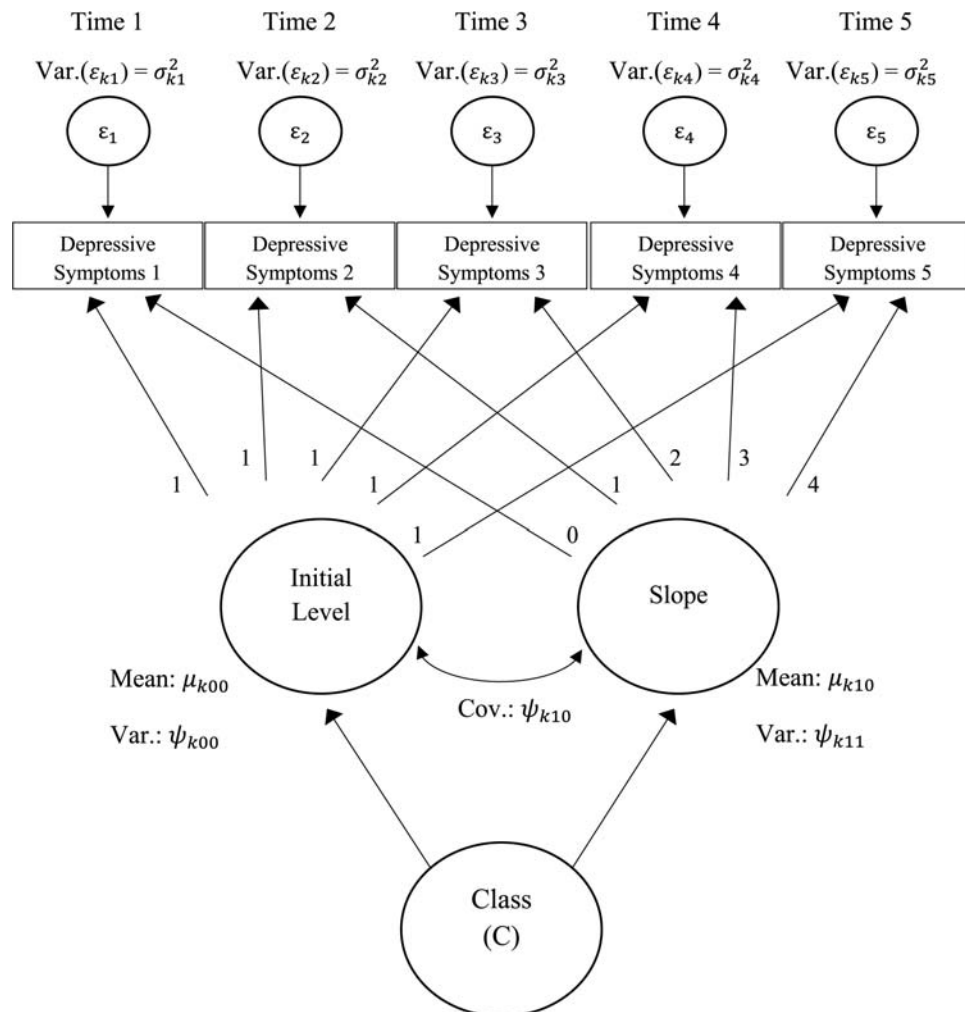


Fig. 3 A linear growth mixture model (GMM). Note. Var. = Variance. Cov. = Covariance. Class $k = 1, 2, 3, \dots K$.

other words, a GMM assumes that the mixture of within-class normal distributions of latent growth factors (initial level and slope) corresponds to several classes that have between-class variation (Bauer and Curran, 2003; Feldman et al., 2009).

Specifying a growth mixture model

As previously discussed, a GMM includes a latent categorical variable, C , that specifies each individual's membership in a certain class. In a GMM, individuals are assigned to one of the categories (or latent classes) with distinct growth patterns (intercept and slopes) according to the individual's growth pattern. Unlike other strategies, such as multi-group growth modeling which requires *a priori* knowledge of each individual's group membership (e.g., a gender which is a categorical grouping variable with two categories), the group assignment in a GMM is latent and "determined" as part of the estimation routine. That is, a GMM identifies *unobserved (latent) subpopulations* and describes the growth patterns of each unobserved sub-population along with examining differences in growth patterns for these unobserved (latent) sub-populations.

Fig. 3 shows the growth parameters in class k ($= 1, 2, \dots, K$). For example, μ_{k00} and μ_{k10} represent the average initial levels and slope in latent trajectory class k . ψ_{k00} and ψ_{k11} are the error variances of the estimated intercept and slope factors in class k , and ψ_{k10} is the covariance between the intercept and slope factors within the latent trajectory class k . The k subscript indicates that most parameters are allowed to vary between the estimated latent trajectory classes. Thus, each latent trajectory class could be defined by its own LGM based on class-specific parameters consisting of (1) means (i.e., μ_{k00} and μ_{k10}) of growth factors, (2) variance and covariance structures (ψ_{k00} , ψ_{k11} , and ψ_{k10}) of the growth factors, and (3) residual variances of manifest indicators (i.e., σ_{kt}^2).

Building a growth mixture model using Mplus¹

A GMM with unbiased model parameters can be estimated in Mplus in four steps.

- Step One Specify a latent growth model (LGM)
- Step Two Specify a growth mixture model (GMM)
- Step Three Address estimation problems
- Step Four Select the optimal unconditional model

For the illustrative purpose of these four steps, we use data from adolescents from rural families in the Midwest who participated in a panel study; Family Transition Project, (FTP; Principal Investigator [PI]: R. D. Conger). Additional information regarding the study procedures is available from Wickrama et al. (2016). Nine items from the SCL-90-R (Derogatis and Melisaratos, 1983) were used to assess adolescents' self-reported depressive symptoms in 1991, 1992, 1994, 1995, and 1997 on a 5-point Likert scale from 1 = *not at all* to 5 = *extremely*. For each timepoint, a mean score was computed with higher values indicating more depressive symptoms. The measure's internal consistency (α) was greater than .90 across timepoints.

Specify a latent growth model (step one)

Before fitting a GMM to the data, a LGM can be estimated to determine if the LGM assumptions of uniform variation with a common growth function are valid. That is, fitting a LGM to the data allows the researcher to examine the overall fit of the data to a trajectory with a single shape or form before considering trajectory heterogeneity. We recommend examining two important model results from this stage: model fit indices and variances of the growth factors. Poor fit indices and statistically significant variances of growth factors suggest the existence of heterogeneous trajectories, or sub-populations, with distinct trajectories.

Illustrative example: Mplus syntax for a latent growth model

Mplus syntax (version 8.4) for the latent growth model is as follow:

```
VARIABLE: NAMES ARE DEP1-DEP5; USEVAR = DEP1-DEP5; MISSING = all (-999);
MODEL:
  I S | DEP1@0 DEP 2@1 DEP 3@3 DEP 4@4 DEP 5@6;
OUTPUT: SAMPSTAT;
PLOT: SERIES = DEP1-DEP5(S); TYPE = PLOT3;
```

Note that DEP1–DEP5 indicate depressive symptoms at timepoints 1, 2, 3, 4, and 5. This syntax specifies factor loadings for a linear change proportionate to the unequal time intervals of the current example data ($t = 0, 1, 3, 4$, and 6). For the purpose of illustrating the model, we specified a linear growth function model; however, as mentioned previously, a non-linear growth function, such as a quadratic function, could also be estimated. Regarding the model results, the model fit indices suggested that the LGM fit the data poorly ($\chi^2(df) = 70.741(10)$, $p < 0.001$, RMSEA (90%) = 0.117 (0.093, 0.145). CFI/TLI = 0.892/0.892, SRMR = 0.067). Furthermore, the variances for the growth factors were statistically significant (0.165, $p < 0.001$ and 0.004, $p < 0.001$, for the intercept and slope factors, respectively), indicating significant variation in the growth function between individuals. Together the poor model fit and the significant variances of growth factors provide some evidence of potential heterogeneity in longitudinal changes of depressive symptoms.

¹While we focus on utilizing this approach in Mplus, the analytical approach can be implemented with other software packages (e.g., R, SAS, and STATA).

Estimating a growth mixture model (step two)

Building on the LGM results that supported possible heterogeneity in the trajectories of depressive symptoms, a GMM can be estimated as a next step to investigate potential classes of adolescents with distinct symptom trajectories.

Illustrative example: *Mplus* syntax for a growth mixture model

In *Mplus*, a GMM model can be specified using the syntax shown in Figs. 4 and 5:

The bold portions of the syntax are the parts that add the GMM to the existing univariate LGM.

The TYPE = MIXTURE option specifies the mixture model algorithm. By specifying CLASSES = C(2), a two-class model is examined. If the model converges successfully and the results are interpretable, the number of classes can be increased by specifying CLASSES = C(3) followed by CLASSES = C(4), CLASSES = C(5), and so on until the optimal class model is selected. The STARTS and STITERATIONS syntax commands are not required for estimating a GMM because *Mplus* automatically sets these parameters at default values. Unless other values are specified, *Mplus* assumes 20 random sets of starting values and 4 optimizations in the final stage. However, adjusting these values may aid in obtaining successful convergence. These convergence issues are described in detail in Step 3 (Addressing estimation problems).

In the OUTPUT command, TECH7 can be specified to request sample statistics for each class. TECH8 provides the optimization history (i.e., iteration process), which is used in conjunction with STARTS and STITERATIONS. Specifying TECH11 requests the Lo-Mendell-Rubin likelihood ratio test (LMR-LRT) of model fit (Lo et al., 2001). If TECH13 is specified, univariate, bivariate, and multivariate skewness and kurtosis values are provided in the output (i.e., Mardia's measure of multivariate normality test; Muthén, 2003). TECH14 provides the bootstrapped likelihood ratio test (BLRT). We return to these tests when describing how to select the optimal class model (Step 4). Finally, the PLOT command can be used to request graphical displays of the model results. The observed and estimated means and trajectories for each class can be viewed in the graphs provided in the *Mplus* output. Using the dropdown menus in *Mplus* allows the user to view the graphs specified by the PLOT syntax (Plot → View Plots → Estimated Means and Observed Individual Values).

Mplus automatically constraints all variances and covariances of the growth factors to be equal across classes by default (i.e., variance-covariance equality). This variance-covariance equality of growth parameters can lead to problems. Simulation studies have shown that similar restrictions (i.e., equal variance-covariances) could result in the over-extraction of latent classes and biased parameter estimates (Bauer and Curran, 2004; Enders and Tofghi, 2008; Magidson and Vermunt, 2004). Thus, it may be necessary

Title: Growth Mixture Model (GMM)

DATA: File is EXAM.dat;

VARIABLE: names are DEP1-DEP5;

usevar = DEP1-DEP5;

CLASSES=C(2);

missing = all (-999);

ANALYSIS: **TYPE=MIXTURE;**

STARTS = 500 10;

STITERATIONS = 10;

LRTBOOTSTRAP=50;

MODEL: **%OVERALL%**

1 S | DEP1@0 DEP 2@1 DEP 3@3 DEP 4@4 DEP 5@6;

The **CLASSES** option defines "C" and specifies the number of classes in the mixture model.

TYPE=MIXTURE is required for mixture models.

STARTS is used to specify the number of randomly selected starting values.

STITERATIONS is used to specify the maximum number of iterations allowed in the initial stage.

LRTBOOTSTRAP is used in conjunction with the TECH14 option to specify the number of bootstrap draws to be completed when estimating the *p*-value of the true 2 time log-likelihood distribution.

%OVERALL% specifies an overall mixture model.

OUTPUT: SAMPSTAT STANDARDIZED MOD (5.00) **TECH7 TECH8 TECH11 TECH13 TECH14;**

PLOT: SERIES = DEP1-DEP5(S); TYPE=PLOT3;

TECH7 requests the sample statistics for each class.

TECH8 requests the optimization history (i.e., iteration process).

TECH11 requests the Lo-Mendell-Rubin likelihood ratio test that compares *k*-1 class model to the *k*-class model.

TECH13 requests the multivariate skewness and kurtosis test.

TECH14 requests a parametric bootstrapped likelihood ratio test that compares the *k*-1 class model to *k*-class model.

Fig. 4 *Mplus* syntax for a growth mixture model (GMM). Note. I = initial level. S = slope. Dep = depressive symptoms.

to free some (or all) of the within-class variances. Adding the syntax below to the model command of the GMM syntax frees the variances and covariances:

```
MODEL: %OVERALL%
I S | DEP1@0 DEP2@1 DEP3@3 DEP4@4 DEP5@6;
%C#1%
[I-S](M1-M2);
I-S (V1-V2);
I WITH S (COV1);
%C#2%
[I-S](M3-M4);
I-S (V3-V4);
I WITH S (COV2);
MODEL TEST:
M1=M3;
```

The bold commands allow the unique variances and covariances of the intercepts and slopes to be freely estimated for each class. Also, these syntax options make it possible to test hypotheses about all growth parameters (i.e., mean and variance-covariance parameter equality across classes; invariance test of growth parameters) using a Wald chi-square test (i.e., likelihood ratio test, LRT; [Morin et al., 2011](#)). This test is often utilized when selecting the optimal modeling approach (e.g., a model with equal variance-covariances in all classes or a model with freely estimated variances-covariances). We assigned names in parentheses to all growth parameters (e.g., M1, M2, and V1, etc.) to enable use of the MODEL TEST command to test a series of parameter constraints. The lines following the MODEL TEST command provide the specific constraints to be tested. For descriptive purposes, in the current example, we specify the null hypothesis that $M1 = M3$. This corresponds to the null hypothesis that the mean trajectories of depressive symptoms are equal across two classes.

Researchers can also fix the variance and covariance of specific growth parameters to zero (a partially constrained GMM) by adding the appropriate syntax under each class specification. For example, to fix the growth parameters of Class 2 to 0, the following syntax is added:

```
%C#2%
I-S@0; I WITH S@0
```

By assigning labels to the variance-covariance parameters, these constrained variance-covariances can be tested using a Wald chi-square test (i.e., specifying MODEL TEST in the model command).

Regardless of the GMM specification, this approach is exploratory and driven by the data utilized. As such, the results may not align with theoretical expectations or hypotheses. Consequentially, when employing this type of exploratory approach to model specification, results should be treated circumspectly. The decisions about class identification should be based on theory, interpretability, prior research, and practical considerations (e.g., model convergence and computational speed time).

Addressing estimation problems (step three)

Certain estimation difficulties are relatively common with mixture models and can lead to a non-optimal or non-estimated class solution. Previous literature has reported several reasons for these estimation problems. First, the probability distribution of the indicators used for constructing the classes may not be normal ([Bauer and Curran, 2003](#)). Second, there can be convergence problems related to a “local solution” ([Hipp and Bauer, 2006](#)). Third, the model may not converge due to inappropriateness of the data and/or non-identification of the model ([Collins and Lanza, 2010](#); [Wurpts and Geiser, 2014](#)). We explain each of these common problems in the sections that follow.

Estimation problems related to the non-normal probability distribution

When estimating a GMM, the number of latent classes identified may be incorrect due to the nature of the distribution of the observed data (e.g., multivariate skewness or non-normality; [Bauer and Curran, 2003](#)). In latent growth modeling, it is assumed that the repeated measures are normally distributed (i.e., multivariate normality). This implies that the random coefficients (intercept and slope) and time-specific residuals are also normally distributed ([Verbeke and Molenberghs, 2000](#)). However, based on the statistical theory of finite normal mixture models, in a GMM, the normality assumption is relaxed to accommodate the mixture of several normal distributions corresponding to the different classes. These classes may have different mean structures and variance-covariance structures. That is, in a GMM, normality is only assumed within the classes (i.e., within-class normality), and heterogeneity is a characteristic of the residuals. However, if the repeated measures data are not normally distributed, finite normal mixture models, such as a GMM, can erroneously estimate latent class trajectories in the absence of population heterogeneity. Particularly, there is the possibility of over-extracting latent trajectory classes in a homogeneous, non-normal distribution ([Bauer and Curran, 2003](#)). This over-extraction makes generalizing the finding of heterogeneous classes to the larger population doubtful.

To increase the accuracy of latent class trajectory identification, we recommend assessing the normality of the data before selecting the optimal class model. To assess multivariate normality, [Hayduk \(1987\)](#) suggested that each univariate distribution should have a normal distribution. Thus, the univariate skewness and kurtosis statistics should be investigated. Values less than ± 2.00 are generally accepted for normal distributions ([George and Mallery, 2010](#)).

Estimation problems related to local maxima

In the estimation process, the highest log-likelihood value should be replicated when the optimal class solution is achieved. Thus, “local maxima” problems can result in an incorrect class solution. With maximum likelihood estimation (MLE), the model solution is provided when the likelihood ratio reaches a replicable maximum value through an iterative optimization process. This iterative optimization process consists of two stages: (1) the initial stage, which is generated from the given starting values, and (2) the final optimization stage (Myung, 2003). The iterative optimization process could stop prematurely and return a sub-optimal set of parameter values depending on the choice of the initial starting values (chosen either at random or by the researcher). When the highest log-likelihood value is not replicated, a “local maxima” or “local solution” was obtained.

It is good practice to run the model with multiple random sets of starting values to guarantee that the global solution is obtained and avoid “local maxima” (Hipp and Bauer, 2006). The global solution has the highest log-likelihood value (i.e., the best solution). Jung and Wickrama (2008) suggested that between 100 and 500 random sets of starting values are necessary to locate the global solution (i.e., the highest log-likelihood value as opposed to a local solution) (see the “STARTS = 500 10” statement in Fig. 4).

Estimation problems due to model non-identification and inappropriate data

When the model is not identified (i.e., an under-identified model), the model does not converge and no standard errors (and related p -values) or other meaningful estimates are estimated. Non-identification is often due to model complexity; that is, models often fail to converge when too many parameters are estimated simultaneously. In these instances, an error message is printed in the output, and the user needs to re-specify the model, most commonly, by constraining some parameters (Muthén and Muthén, 2012). Non-convergence may also occur when data are used inappropriately (Muthén and Muthén, 2012). For example, some researchers may be interested in examining trajectories of major depressive disorder (0 = No or 1 = Disorder). In these situations, it would be important to model the data as having a logistic or probit distribution rather than as having a continuous normal distribution (Lee et al., 2018).

Selecting the optimal class model (step four)

After successful convergence and support for a global solution, the final step is to select the optimal class model. With other model types, chi-square values are the primary criteria for selecting the optimal model. However, in a mixture model, the chi-square values are not very useful for determining the optimal model because the likelihood ratios between the k -class and $k-1$ class model do not follow a chi-square distribution (Feldman et al., 2009; Muthén, 2003). Thus, it is necessary to investigate multiple model fit indices to select the optimal model. Various indices have been suggested; they fall into three categories (Feldman et al., 2009; Jung and Wickrama, 2008; Nylund et al., 2007):

- (1) Information criteria (IC) statistics, such as the Bayesian information criterion (BIC) and the sample size-adjusted BIC (SSABIC),
- (2) Entropy values, and
- (3) likelihood ratio tests (LRT), such as the Lo-Mendell-Rubin likelihood ratio-test (LMR-LRT) and the bootstrapped LRT (BLRT) (Tein et al., 2013).

Information criteria (IC) statistics

IC statistics are generally computed using the deviance statistics (-2 log-likelihood values), the number of model parameters (p), and sample size (n). IC statistics incorporate a parameterization penalty for models with a large number of parameters (as an indicator of high model complexity) and/or small sample size. In general, lower values of the Bayesian information criterion (BIC) and sample size adjusted BIC (SSABIC) values indicate a better model fit (Feldman et al., 2009).

Entropy and average posterior probabilities

Entropy is a standardized index (i.e., range = 0 to 1) of model-based classification accuracy. Higher values indicate improved enumeration accuracy, which signals clear class separation (Feldman et al., 2009; Nagin, 2005). Entropy is estimated based on the average posterior probabilities (Muthén, 2004). Thus, an entropy value of zero indicates that the posterior probability of class membership does not vary across classes (no class separation). In contrast, an entropy value of one indicates that, for any given individual, there is perfect posterior probability (value of 1) for membership in one class with no posterior probability (value of 0) for membership in other classes, signaling perfect class separation). According to Clark and Muthén (2009), entropy values of 0.40, 0.60, and 0.80 represent low, medium, and high class separation. The entropy statistic is provided in the *Mplus* output along with the average posterior probability for each class. Like the entropy statistic, in general, high probabilities indicate that the classes are distinct from each other. For example, a probability of 0.91 suggests that 91% of subjects in the assigned class fit that category, while 9% of the subjects in that given class do not accurately fit that category (Fanti and Henrich, 2010).

Likelihood ratio test (LRT): LMR-LRT and bootstrapped LRT (BLRT)

Two LRTs, the Lo-Mendell-Rubin likelihood ratio test (LMR-LRT; Lo et al., 2001) and the bootstrapped likelihood ratio test (BLRT; McLachlan and Peel, 2000) are often used for model comparison when determining the optimal number of classes. The LMR-LRT compares two adjacent class models ($k-1$ and k -class model) (Lo et al., 2001). Like the SSABIC, an adjusted LMR-LRT (Adj. LMR-LRT) statistic accounting for sample size is provided in *Mplus* (Muthén, 2004).

McLachlan and Peel (2000) also developed another LRT known as the bootstrapped LRT (BLRT). Using this approach, the likelihood ratio test (LRT) between the $k-1$ and k -class models is conducted through a bootstrapped procedure (Asparouhov and Muthén, 2012). The bootstrap draw can be manually specified using the LRTBOOTSTRAP option in the ANALYSIS command (The default number varies from 2 to 100). For example, the following syntax specifies 50 bootstrap draws (see the corresponding syntax in Fig. 4).

LRTBOOTSTRAP = 50; 100 bootstrap draws has been suggested by past research for the estimation of a more precise p -value (McLachlan and Peel, 2000). In both tests (LMR-LRT and BLRT), the difference in log-likelihood values between the $k-1$ and k -class models is calculated and inference statistics are provided (p -values). Significant p -values indicate that the k -class model (i.e., the current model) provides a better fit than the $k-1$ class model (i.e., a model with one less class).

In the *Mplus* output, both LRTs (i.e., LMR-LRT and BLRT) provide the difference of log-likelihood values between the $k-1$ class model and the k -class model. A low p -value for both statistics indicates that the current k -class model is a significantly better fit to the data than a model with one less class. In most cases, the BLRT outperforms the LMR-LRT and other IC statistics (Nylund et al., 2007; Tein et al., 2013). However, the BLRT is more sensitive to non-normal distributions and model complexity, which can lead to convergence problems (Nylund et al., 2007). If a local solution is detected in either the $k-1$ class or the k -class model (or both models), the difference between the log-likelihood values will be biased. In turn, this leads to biased inference statistics (p -values). As such, the more robust LMR-LRT is generally preferred (Nylund et al., 2007). In the "Model fit information" section of the *Mplus* output, IC statistics and the entropy statistic are provided. However, the LMR-LRT and BLRT are only shown when TECH11 (for LMR-LRT) and TECH14 (for BLRT) are specified in the syntax (see the corresponding syntax in Fig. 4).

Other considerations

In addition to these three classes of model fit indices used for selecting the optimal number of classes, there are other factors that should be considered (Feldman et al., 2009). Enumerating the number of latent classes in a GMM is not a simple identification procedure based solely on model fit indices. For example, researchers can consider conventional model comparison approaches. If similar information (e.g., fit indices) is provided for both $k-1$ and k -class models, the more parsimonious model (i.e., the model with fewer classes) is preferred (Feldman et al., 2009). An additional consideration is the sample size of the smallest class. If the smallest class contains less than 5% of the sample and/or the sample size for the smallest class is less than 25, it is recommended that the model only be retained as the optimal model if the researcher can accurately defend what is gained from this small class given the possibility of low power and a lack of statistical precision (Berlin et al., 2014; Lubke and Neale, 2006). Finally, the interpretability of each class trajectory is essential. Plots often provide useful information when selecting the optimal class model. When viewing the plots, each class trajectory should be clearly distinct and separate from other classes. The classes should also be interpretable and consistent with theory.

Illustrative example: identifying the optimal model

Using our example model (i.e., repeated measures of depressive symptoms from Time 1 to Time 5; DEP1 to DEP5), we illustrate how to select the optimal class model. We first examined the model fit indices for the two-class model and compared those indices with indices from the model with three classes. This process continued incrementally by fitting models with a larger number of classes until model fit ceased to improve. Furthermore, to avoid the possibility of a local maxima, 500 random sets at the initial stage and 10 final optimizations were specified for all models (STARTS = 500 10). Note that, for model simplicity, only a linear trajectory was assessed. Table 1 provides the model fit indices for these incremental GMMs with two to five classes.

As shown in Table 1, no convergences problems (i.e., local maxima) were detected across the class models. The 5-class model had the lowest BIC and SSABIC values, supporting the 5-class model as the preferred model. However, the 4- and 5-class models had classes with small group sizes. The smallest class sizes were 5 (1.1%) and 6 (1.4%) for the 4- and 5-class models, respectively. Also, the output suggests that estimation of the BLRT in the 5-class model is biased. Taking these results into consideration, the 3-class model was selected as the optimal model. The model fit indices suggest that the 3-class model was a better fit to the data than the 2-class model (lower BIC and SSABIC values, higher entropy value [with average posterior probabilities ranging from 0.898 to 0.990], and a statistically significant p -value for the BLRT). While the p -value for the LMR-LRT was not statistically significant (indicating that the 3-class model does not improve the model fit), Nylund et al. (2007) reported that the BLRT outperforms the LMR-LRT when a local solution is not detected.

Growth parameters for each optimal model are shown in Table 2. Most growth parameters (i.e., means and variances) in this 3-class model were statistically significant. The distinct patterns of (mean) trajectories for each class suggest that longitudinal changes in students' depressive symptoms differ by classes. Also, within a class, there was significant inter-individual variation in growth factors (initial level and slope), indicating that individual trajectories varied within a class.

Consistent with these previous findings noting heterogeneity in depressive symptoms trajectories (Wickrama et al., 2008), the results in Table 2 suggest the existence of three general trajectory patterns. These three patterns include: (1) a consistently low symptom trajectory, (2) an initially low or moderate and increasing symptom trajectory, and (3) an initially high and decreasing symptom trajectory. Combining the information gathered from the model fit statistics, plots, and relevant theory, we selected the three-class GMM model as the optimal class model.

Table 1 Fit statistics for incremental GMMs (total N = 436).

Fit statistics	2 classes	3 classes	4 classes	5 classes
LL (no. of parameters)	−963.282 (13)	−918.234 (16)	−887.040 (19)	−853.867 (22)
BIC	2005.564	1933.711	1889.555	1841.443
SSABIC	1964.318	1882.935	1829.258	1771.627
Entropy	0.974	0.953	0.957	0.939
Adj. LMR-LRT (<i>p</i>)	<i>216.795 (0.00)</i>	85.411 (0.20)	58.615 (0.12)	62.895 (0.16)
BLRT (<i>p</i>)	<i>228.685 (0.00)</i>	<i>90.096 (0.00)</i>	61.830 (0.00)	66.344 (0.00) ^a
Group size (%)				
C1	27 (6.2%)	20 (4.6%)	5 (1.1%)	6 (1.4%)
C2	409 (93.8%)	38 (8.7%)	21 (4.8%)	11 (2.5%)
C3		378 (86.7%)	33 (7.6%)	32 (7.4%)
C4			377 (86.5%)	34 (7.7%)
C5				353(81.0%)

Note. LL = Log-Likelihood value. No. of Parameters = Number of estimated (freed) parameters. BIC = Bayesian Information Criteria. SSABIC = Sample Size Adjusted BIC. LMR-LRT = Lo-Mendell-Rubin Likelihood Ratio Test. Adj. LMR-LRT = Adjusted LMR. BLRT = Bootstrap Likelihood Ratio Test. *p* = *p*-value. Bold and italic values indicate the best values for the model fit statistics in the set of models listed in the corresponding row.

^aNo replicated log-likelihood value (H0 log-likelihood value) of the *k*-1 class model.

Table 2 Growth parameters across the three classes in the optimal GMM.

	Intercept		Linear slope		Factor covariance
	Mean	Variance	Mean	Variance	
C1 (high and decreasing)	2.651***	0.050***	−0.187***	0.001*	−0.003*
C2 (low and increasing)	1.432***	0.050***	0.181***	0.001*	−0.003*
C3 (consistently decreasing)	1.374***	0.050***	−0.023***	0.001*	−0.003*

Note. C = Class. Unstandardized coefficients are shown. **p* < 0.05. ***p* < 0.01. ****p* < 0.001.

Conclusion

This chapter provides an overview of growth mixture modeling (GMM), characterized as an extension of latent growth modeling (LGM) and demonstrates the strengths of this approach via a practical illustration. Whereas LGM applications typically assume that the sample is comprised of a single population with a single set of growth parameters (e.g., means, variances, covariances). With GMM, this assumption is relaxed to allow for unobserved subpopulations drawn from latent classes with different growth parameters for each subpopulation. Thus, using GMM, researchers can identify and better understand various groups and patterns of change that may exist within a dataset. In illustrating the application of GMM with an example exploring adolescents' depressive symptoms, we found three distinct patterns of longitudinal change in their depressive symptoms. We hope our illustrations serve as a guide for other researchers and prompt increased interest in the use of GMM as we believe this approach is a useful tool for examining research questions centered on the students' health developments.

References

- Asparouhov, T., Muthén, B., 2012. Using Mplus Tech11 and Tech14 to Test the Number of Latent Classes. Retrieved from. <https://www.statmodel.com/examples/webnotes/webnote14.pdf>.
- Bauer, D.J., Curran, P.J., 2003. Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychol. Methods* 8, 338–363.
- Bauer, D.J., Curran, P.J., 2004. The integration of continuous and discrete latent variable models: potential problems and promising opportunities. *Psychol. Methods* 9, 3–29.
- Berlin, K.S., Williams, N.A., Parra, G.R., 2014. An introduction to latent variable mixture modeling (Part 1): cross-sectional latent class and latent profile analyses. *J. Pediatr. Psychol.* 39, 188–203.
- Clark, S., Muthén, B., 2009. Relating Latent Class Analysis Results to Variables Not Included in the Analysis. Retrieved from. <https://www.statmodel.com/download/relatinglca.pdf>.
- Collins, L.M., Lanza, S.T., 2010. *Latent Class and Latent Transition Analysis: With Applications in the Social Behavioral and Health Sciences*. John Wiley & Sons Inc, Hoboken, NJ.
- Derogatis, L.R., Melisaratos, N., 1983. The brief symptom inventory: an introductory report. *Psychol. Med.* 13, 595–605.
- Enders, C.K., Tofighi, D., 2008. The impact of misspecifying class-specific residual variances in growth mixture models. *Struct. Equ. Model.* 15, 75–95.
- Fanti, K.A., Henrich, C.C., 2010. Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: findings from the national institute of child health and human development study of early child care. *Dev. Psychol.* 46, 1159–1175.
- Feldman, B., Masyn, K.E., Conger, R., 2009. New approaches to studying problem behaviors: a comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Dev. Psychol.* 45, 652–676.

- George, D., Mallery, M., 2010. SPSS for Windows Step by Step: A Simple Guide and Reference, 16.0 Update, tenth ed. Pearson, Boston.
- Hayduk, L., 1987. Structural Equation Modeling With LISREL. John Hopkins University Press, Baltimore.
- Hipp, J.R., Bauer, D.J., 2006. Local solutions in the estimation of growth mixture models. *Psychol. Methods* 11, 36–53.
- Jung, T., Wickrama, K.A.S., 2008. An introduction to latent class growth analysis and growth mixture modeling. *Soc. Personal. Psychol. Compass* 2, 302–317.
- Lee, T.K., Wickrama, K.A.S., O'Neal, C.W., Lorenz, F.O., 2017. Social stratification of general psychopathology trajectories and young adult social outcomes: a second-order growth mixture analysis over the early life course. *J. Affect. Disord.* 15, 373–383.
- Lee, T.K., Wickrama, K.A.S., O'Neal, C.W., 2018. Application of latent growth curve analysis with categorical responses in social behavioral research. *Struct. Equ. Model.* 25, 294–306.
- Little, T.D., 2013. Longitudinal Structural Equation Modeling. Guilford Press.
- Lo, Y., Mendell, N., Rubin, D., 2001. Testing the number of components in a normal mixture. *Biometrika* 88, 767–778.
- Lubke, G.H., Neale, M.C., 2006. Distinguishing between latent classes and continuous factors: resolution by maximum likelihood. *Multivariate Behav. Res.* 4, 499–532.
- Magidson, J., Vermunt, J.K., 2004. Latent class models. In: Kaplan, D. (Ed.), *Handbook of Quantitative Methodology for the Social Sciences*. Sage, Newbury Park, CA, pp. 175–198.
- McLachlan, G., Peel, D., 2000. *Finite Mixture Models*. Wiley, New York.
- Meredith, W., Tisak, J., 1990. Latent curve analysis. *Psychometrika* 55, 107–122.
- Morin, A.J.S., Maïano, C., Negengast, B., Marsh, H.W., Morizot, J., Janosz, M., 2011. General growth mixture analysis of adolescents' developmental trajectories of anxiety: the impact of untested invariance assumptions on substantive interpretations. *Struct. Equ. Model.* 18, 613–648.
- Muthén, B., Muthén, L.K., 2000. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin. Exp. Res.* 24, 882–891.
- Muthén, L.K., Muthén, B.O., 2012. *Mplus User's Guide*, seventh ed. Authors, Los Angeles.
- Muthén, B.O., Shedden, K., 1999. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 55, 463–469.
- Muthén, B., 2003. Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). *Psychol. Methods* 8, 369–377.
- Muthén, B., 2004. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan, D. (Ed.), *Handbook of Quantitative Methodology for the Social Sciences*. Sage, Newbury Park, CA, pp. 345–368.
- Myung, I.J., 2003. Tutorial on maximum likelihood estimation. *J. Math. Psychol.* 47, 90–100.
- Nagin, D.S., 2005. *Group-Based Modeling of Development*. Harvard University Press, Cambridge, MA.
- Nylund, K.L., Asparouhov, T., Muthén, B., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equ. Model.* 14, 535–569.
- Tein, J.-Y., Cox, S., Cham, H., 2013. Statistical power to detect the correct number of classes in latent profile analysis. *Struct. Equ. Model.* 20, 640–657.
- Verbeke, G., Molenberghs, G., 2000. *Linear Mixed Models for Longitudinal Data*. Springer, New York.
- Wickrama, K.A.S., Conger, R.D., Abraham, W.T., 2008. Early family adversity, youth depressive symptoms trajectories, and young adult socioeconomic attainment: a latent trajectory class analysis. *Adv. Life Course Res.* 13, 161–192.
- Wickrama, K.A.S., Lee, T.K., O'Neal, C.W., Lorenz, F.O., 2016. *Higher-Order Growth Curves and Mixture Modeling With Mplus: A Practical Guide*. Routledge/Taylor & Francis Group.
- Wurpts, I.C., Geiser, C., 2014. Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Front. Psychol.* 5, 920–934.

Nonparametric statistical methods

Wilco Emons^a and Klaas Sijtsma^b, ^a Cito, National Institute for Educational Measurement, Arnhem, the Netherlands; and ^b Tilburg University, Tilburg, the Netherlands

© 2023 Elsevier Ltd. All rights reserved.

Introduction	657
Nonparametric statistical tests	658
Testing hypotheses about group differences	658
Measures of association and tests of independence for ordinal variables	660
Miscellaneous methods	662
Nonparametric data exploration for single variables	662
Nonparametric regression models	663
Preliminary: parametric regression analyses	663
Nonparametric regression analysis	664
Advanced smoothers	665
Order-restricted nonparametric regression	666
Gauging uncertainties: nonparametric confidence intervals	667
Illustration using PISA data	667
Concluding remarks	667
References	669

Introduction

The goal of a statistical analysis is to estimate properties of a population from sampled data. Examples are the distribution of a test score for arithmetic ability and the difference between boys and girls with respect to the mean test score for arithmetic ability. The researcher may test hypotheses, for example, whether the distribution is normal and the difference is significant. Nonparametric methods are appropriate when certain assumptions about distributions that common parametric methods make are questionable. For example, when the sample size is small (say, less than 20) and the population distribution is skewed, the sample means do not follow a parametric Student's *t*-distribution. Consequently, the one-sample *t*-test results in a probability of exceedance (*p*-value) that is biased and therefore cannot be trusted, and valid conclusions about the population based on sample data cannot be drawn (e.g., Wilcox, 2001, pp. 1–2). A nonparametric alternative would be the Wilcoxon signed-rank test (e.g., Hollander and Wolfe, 1999, p. 36), which only assumes that the observations come from a distribution that is symmetrical about the mean, but without assuming normality or another parametric shape.

For introductory textbooks on nonparametric statistics see, for example, Siegel and Castellan (1988), Hollander and Wolfe (1999), Sheskin (2007), Kolassa (2020), and Kloeke and McKean (2020). In this entry, we consider two particular classes of widely used nonparametric methods. One class concerns methods for testing hypotheses about population properties, but without distributional assumptions that may be unreasonable in certain circumstances. Nonparametric statistical methods derive the sampling distribution without assuming a particular, restrictive parametric shape for the population quantity of interest, and for that reason, they are often referred to as *distribution-free methods*. The signed-rank test example referred to belongs to this class of nonparametric methods.

The other class concerns regression methods that can be used to study the functional relation between variables without imposing a particular parametric function on this relationship. For example, linear regression assumes a linear relationship between two interval-level variables. Many examples are conceivable where the assumption of linearity is untenable. Examples include educational progress as a function of age (e.g., Kiefer et al., 2018) and test performance (e.g., number of correct answers) as a function of ability (Sijtsma and van der Ark, 2021). One possible solution is to use a non-linear parametric function, such as the logistic function (e.g., Fox, 2015). Alternatively, one can estimate the functional relationship directly from the data (e.g., Fan and Gijbels, 1996; Fox, 2000). Nonparametric regression methods may be particularly useful to study the functional relation between an interval-level independent variable and a binary or ordinal-level dependent variable. Furthermore, estimating the shape of the relationship from the data may lead to different and interesting results that a particular parametric function, such as the linear or the logistic function, would obscure when estimated.

Nonparametric methods have several interesting properties. First, they are suited for variables at all measurement levels, including categorical data, which are common in educational research. Categorical variables have either a nominal or an ordinal measurement level, but may also represent counts of particular events. Examples of nominal variables include the kinds of arithmetic errors made by students and the assignment of each student to one of three remedial teaching methods for arithmetic. Examples of ordinal variables include the teacher's rank ordering of students with respect to perceived need of remedial teaching and the

student's attitude to learning arithmetic. An example of a count variable is the number of students in each remedial teaching program who receive help.

Nonparametric methods are also useful for interval-level variables, such as students' motivation, when samples are small and assuming a parametric population distribution is not justified. The scores are then treated as ordinal rather than interval-level variables. The price paid is that information about the distances between scores is lost and the power of statistical tests reduced. Researchers often consider this unacceptable, and incorrectly analyze their data as if the sampled scores on a variable stem from a known distribution. Research has shown that in such cases the loss of information by using nonparametric methods is not as bad as one might suspect (e.g., [Hollander and Wolfe, 1999](#), p. 1).

Second, nonparametric methods are based on weaker assumptions than parametric methods. For example, instead of *assuming* a parametric distribution for a particular statistic such as the *t*-distribution, nonparametric methods use all possible rank-orderings of the data to derive the exact distribution for the statistics of interest. Statistical models based on weak assumptions are important because their assumptions are realistic about the level of knowledge expressed in many theories in educational science. However, although distribution-free, nonparametric methods do make assumptions about population distributions. Distribution-free only means that the sampling distribution is not parametrically restricted, such as the normal distribution or the Poisson distribution would do.

Third, nonparametric statistical models are convenient because they are easy to use, but some restraint may be in order. Nonparametric methods often involve complex combinatorial mathematics that are more difficult to grasp than the assumptions underlying parametric methods. Combinatorics is notorious for running into computations that soon become unmanageable even with today's computers as sample size increases or the numbers of variables increases, and all possible patterns of scores must be computed.

This entry has the following structure. First, we review a few well-known nonparametric statistical tests from the myriad of test procedures (e.g., [Siegel and Castellan, 1988](#); [Wasserman, 2006](#)) based on known, exact sampling distributions or simulated sampling distributions when exact sampling distributions are unknown or numerically too complex to evaluate. Second, we discuss methods for nonparametric data exploration and nonparametric regression. Third, we explain how to use computer-based methods for estimating standard errors and confidence intervals, both for univariate sample statistics (e.g., a distribution's median) and regression functions. The nonparametric principles and techniques discussed are a selection of widely used techniques, but do not cover the entire domain of nonparametric methods.

To illustrate the methods, we use small numerical examples and a real-data example from the Program for International Student Assessment (PISA; [OECD, 2019](#)). PISA is a triennial international survey addressing 15-year old's performance in reading, mathematics, and science. Each wave also includes additional questionnaires on various socio-economic and education-related attitudes, such as school well-being, school environment, and attitude toward science. We used data from the Dutch population collected in 2018 ([OECD, 2019](#)). The 2018-version of the student questionnaire included several scales related to student's attitude toward reading and the accessibility of reading materials for these students such as the number of books available in their home.

Nonparametric statistical tests

Testing hypotheses about group differences

To illustrate some of the foundations of nonparametric statistics, we first consider the situation where we compare two groups. Say, the researcher studies whether students receiving game-based instruction engage in deeper learning than students receiving traditional instruction (e.g., [Bado, 2019](#)). An experiment is done in which two groups of five students are assigned to one of these two conditions. One group is instructed using the traditional approach (e.g., classroom teaching) and the other group receives game-based instructional activities. After completing the program, the students are assessed on learning outcomes, motivation and learning experience. [Table 1](#) shows hypothetical data from such an experiment. The numbers represent number-correct scores

Table 1 Fictional data for academic performance in two groups trained with different instructional methods.

Instructional method			
Traditional		Game-based	
Score	Rank among all students	Score	Rank among all students
3	1	20	10
4	2	14	6
5	3	15	7
7	5	6	4
18	9	17	8
Rank sum:	20	Rank sum:	35

on a test of 20 items reflecting learning results. The statistical question is whether the mean difference in number-correct scores between the two instructional conditions is significant. Because the samples are small, a nonparametric approach is in order.

Adopting a nonparametric statistical test, we take the students from the two groups together and order them by their test score from smallest to largest and assign rank numbers (see ranks among all students in [Table 1](#)). If the two instructional methods have the same effect, implying that the test scores in each group come from the same underlying distribution, we expect the rank numbers are evenly distributed across the two conditions. However, if one instructional method is more effective than the other method, meaning that the students in one group perform generally better, we expect most low rank numbers in one group and most high rank numbers in the other group. Hence, if we sum the ranks in each condition separately, then a higher the rank sum in, say, the game-based condition provides evidence of a positive effect of game-based teaching. The test statistic of interest is the sum of the ranks in the game-based group, which in [Table 1](#) equals 35.

Let W_s denote the rank-sum variable and w_s the observed sample value in group s (here, $s = 1$ for the traditional group and $s = 2$ for the game-based group). To test whether the treatment had a significant effect, we need the distribution of W_2 under the null hypothesis (H_0) of “no effect”; this is the null distribution. Nonparametric methods use all possible rank-orderings in the data to derive the exact distribution of the statistic of interest. Assuming random sampling, all possible arrangements of the students are equally likely under H_0 . This means that under H_0 the observed rank numbers in the game-based condition can be treated

as a random sample without replacement of five numbers out of the rank numbers 1 through 10. In total, there are $\binom{10}{5} = 252$ possible combinations, each having a probability of $1/252$. For each combination, we compute the associated value of the rank sum W_2 . The distribution of the rank sums across these combinations is the exact null distribution ([Fig. 1](#)). The probability that W_2 exceeds observed sample value = 35 is the one-tailed p -value, which equals 0.076 and corresponds to a two-tailed p -value of 0.151. These p -values are called exact p -values because they are based on the exact sampling distribution. Adopting a significance level $\alpha = 0.05$, we conclude that the sample does *not* provide evidence that the game-based instructional method is more effective than the traditional method.

The test we described is the Wilcoxon rank-sum test, better known as Mann-Whitney’s U test ([Mann and Witney, 1947](#); [Wilcoxon, 1945](#)), with the difference that the latter defines the test statistic using a rescaled version of the rank sum. The rescaled value has a fixed lower bound of 0, which is obtained when in the common ranking the persons from one group occupy the lowest positions and the persons from the other group the highest positions. Then, the groups are maximally separated, that is, the maximum effect is found. Hence, statistic U thus has a more straightforward interpretation. For the Mann-Whitney U test, let n_1 be the sample size and R_{+1} the rank sum in group 1, and let n_2 and R_{+2} be the sample size and rank sum in group 2. Mann-Whitney U is defined as U_1 or U_2 , whichever is smaller, where

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_{+1},$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_{+2}.$$

In [Table 1](#), $U_1 = 20$ and $U_2 = 5$; hence, Mann-Whitney’s U equals 5. Notice that if group 1 contains the five lowest rank numbers (i.e., 1 through 5), we have $U_1 = 0$ and $U_2 = 25$, and thus $U = 0$.

The example shows some important features of nonparametric methods. First, nonparametric methods use combinatorics to derive exact null distributions. This works well for small samples (say, less than 20 observations), but as sample size increases, the number of combinations grows exponentially and the exact evaluation soon is impossible. Fortunately, for many nonparametric statistics the sampling distribution converges to the normal distribution as sample size increases. Then, one uses large-sample

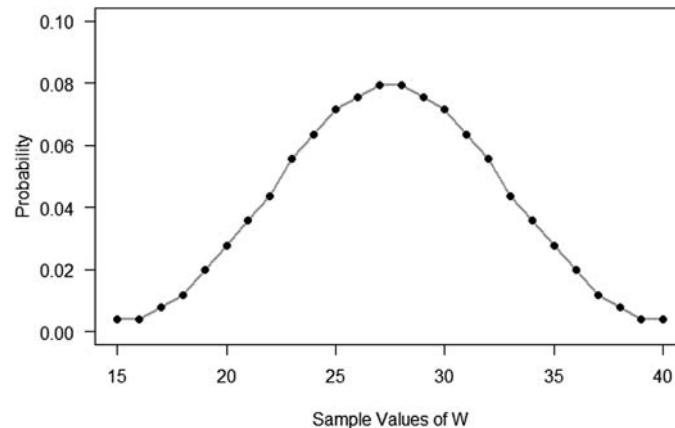


Fig. 1 Exact sampling distribution for Wilcoxon’s rank-sum statistic in the traditional group (W_2).

approximations to test the hypotheses of interest. For example, for the Wilcoxon rank-sum statistic W_2 , the large-sample approximation is the normal distribution with mean

$$E_{H_0}(W_2) = \frac{n_2(N+1)}{2} = \frac{5(10+1)}{2} = 27.5$$

and variance

$$\sigma_{H_0}^2(W) = \frac{n_1 n_2 (N+1)}{12} = \frac{5 \cdot 5 \cdot 11}{12} = 22.92.$$

The normalized value of W_2 then equals

$$Z_{W_2} = \frac{w_2 - E_{H_0}(W_2)}{\sqrt{\sigma_{H_0}^2(W)}} = \frac{35 - 27.5}{\sqrt{22.92}} = 1.57$$

and the corresponding two-tailed p -value, which in this case equals 0.117, is obtained from tables for the normal distribution or using standard software. The normal distribution is continuous, whereas statistic W_s is discrete. Continuity corrections exist to account for the discreteness of the rank sums. In our example, the continuity corrected two-tailed p -value equals 0.144. Finally, the independent samples t -test would yield a two-tailed p -value of 0.088. If we were to test at the 10% significance level, we would reject H_0 using the independent samples t -test, but not if we would use the Mann-Whitney test. In this case, it matters which statistical test one chooses for the conclusions of the study.

Second, the example shows another feature of nonparametric statistics, which is the use of permutation tests. Permutation tests are used to evaluate hypotheses defined by rank-order information. If the number of permutations is too large to evaluate hypotheses, then one replaces the ambition to evaluate all by a large albeit limited number of random permutations and computes the Monte Carlo p -value (Davison and Hinkley, 1997, pp. 141–142). This approach circumvents either making assumptions about the distribution of the variables, as the independent samples t -test does, or relying on asymptotic large-sample approximations, but the number of simulated replications must be very large.

Third, a feature with which nonparametric methods do *not* distinguish themselves from parametric methods is that they make assumptions about sampling and distributions underlying the variable of interest. A critical assessment of the plausibility of assumptions is important for both nonparametric methods and parametric methods. An assumption both kinds of methods share is that samples are drawn at random. This assumption also underlies the Wilcoxon rank sum test. Another assumption of the Wilcoxon rank-sum test is that under the null hypothesis the observations in both study groups come from the same population distribution. The alternative hypothesis assumes that observations come from populations that are only different by a shift in location, but otherwise have the same shape and spread. If the latter assumption is unreasonable, the Wilcoxon rank-sum test produces a p -value that cannot be trusted.

The Wilcoxon rank-sum test for testing mean differences between three or more populations as in analysis of variance is available. A widely used test is the Kruskal-Wallis test (Kruskal and Wallis, 1952). Its technical discussion is beyond the scope of this article, but the basic idea is the same as for the Wilcoxon rank-sum test. One obtains for all respondents their rank in the combined samples, and then determines the rank sums per condition. Under the null hypothesis, we expect the same rank sums across all conditions. Variation in the rank sums across the groups suggests an effect. If the observed variance in the rank sums has a probability of exceedance smaller than 0.05 under the null hypothesis (i.e., $p < 0.05$), the effect is significant at the 5% level. The permutation test yields an exact p -value. Normal approximations exist as well; see Hollander and Wolfe (1999). Generalizations to multivariate analyses are possible (e.g., Katz and McSweeney, 1980; Finch and Davenport, 2009).

Measures of association and tests of independence for ordinal variables

The degree of association between variables, say, X and Y , provides information about either the underlying mechanisms of the phenomenon under study, or may be used to predict future events as in educational screening and for selection tests. Educational researchers may study whether motivation, home support, and access to learning materials are important variables for predicting study success. They may study the effect of a new teaching method on arithmetic ability or an operating educational policy on social inequalities. The most commonly used measure of association is Pearson's product-moment correlation (Pearson, 1895; also, see Agresti et al., 2017, Chap. 3), denoted ρ in the population, and r in the sample. Correlation ρ describes the degree of linear association between two variables. A correlation of 0 means absence of a *linear* association. Assuming bivariate normality for X and Y and a sample of size N , the null hypothesis that $\rho = 0$ can be tested using a t -test.

We assume that variables X and Y are ordinal, their relationship is non-linear, or that one is reluctant to assume bivariate normality, in each case rendering the product-moment correlation unsuitable. An alternative to the product-moment correlation is Spearman's rank correlation ρ_s (Spearman, 1904; also, see Liebetrau, 1983). The scores on each variable are replaced by their ranks, which we also denote X and Y . For example, for a respondent indexed i , the pair of ranks ($X_i = 7$, $Y_i = 15$) shows that (s) he was ranked 7 on variable X and 15 on variable Y . The difference between these ranks equals $d_i = y_i - x_i = 8$. We calculate the difference of the ranks for all other persons. When several people have the same score on a variable, they are put in random order, and their rank is the mean of the ranks assigned had no ties been observed. Thus, for each person i , we have converted their

scale-level measurements into scale-free and distribution-free rank numbers. Spearman's ρ_s is defined as the sample product-moment correlation between the rank numbers X and Y , and for sample data is denoted r_s . Ignoring ties for simplicity, and using differences d between paired ranks, r_s equals

$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}.$$

For small samples, a permutation test may be used for testing the null hypothesis that $\rho_s = 0$. Spearman's ρ_s is especially useful for studying associations between continuous variables that fail to meet the assumptions of bivariate normality. However, when X and Y are already discrete and ordinal, each with a limited number of categories, there will be many ties and Spearman's ρ_s may not be optimal. Kruskal's gamma (γ) and different versions of Kendall's τ are more appropriate (see [Hollander and Wolfe, 1999](#); [Liebetrau, 1983](#)).

Suppose an educational researcher investigates the association between students' satisfaction with school support and engagement in extracurricular activities. Perceived school support is measured using a three-point Likert item having ordered response categories *low*, *medium*, and *high*. Extra-curricular engagement is measured on a three-point scale having categories *never*, *occasionally*, and *often*. **Table 2** gives fictional data for this research question. Kruskal's γ and Kendall's τ coefficients use the concepts of concordant and discordant observation pairs to quantify ordinal association. A student who never engages in extracurricular activities and has low satisfaction and a student who often engages in such activities and has high school-support satisfaction form a concordant pair, because they are equally ordered on both variables. Hence, concordant pairs suggest positive association. Students ordered oppositely on the two variables form a discordant pair, suggesting a negative association. A tie occurs if a person pair has the same score on one or both variables. Ties allow different interpretations giving rise to different association measures.

Let C be the number of concordant pairs in the sample, D the number of discordant pairs, and T the number of ties. Kruskal's γ uses only C and D , and is defined as

$$\gamma = \frac{C - D}{C + D}.$$

Under the null hypotheses of independence we expect $C = D$, implying $\gamma = 0$. In the example in **Table 2** we have $C = 164$, $D = 26$, and $\gamma = 0.726$. Kruskal's γ treats ties as uninformative or undecisive about the association. [Kendall \(1938\)](#) took a different approach. Kendall's τ_a includes T and equals

$$\tau_a = \frac{C - D}{\frac{1}{2}[N(N - 1)]} = \frac{C - D}{C + D + T} \quad (1)$$

Kendall's τ_a is similar to gamma except that in the denominator it has all possible pairs that can be formed, including all possible ties. Conceptually, this means that ties provide evidence of a lack of association. Kendall's τ_b refines τ_a by making additional adjustments for the number of ties on each single variable. These ties are considered uninformative of the association and they are therefore subtracted from T in **Eq. (1)**. The resulting denominator contains all the relevant ordinal information available in the sample. Because T is reduced, we have $\tau_b \geq \tau_a$. In the example, we have $\tau_b = 0.512$. Kendall's τ_b is the preferred method when the ordinal variables have the same number of categories (i.e., the contingency table is square), otherwise, Stuart-Kendall's τ_c is preferred.

The γ and τ -coefficients can also be written as a function of the rank numbers, but the formulas are rather complex and therefore beyond the scope of this entry. What matters here is that the statistics use rank information and thus the permutation test may be used to test the association for significance. If the sample size permits, one can evaluate all possible permutations to derive the exact null distribution, but otherwise one could use the bootstrap method ([Davison and Hinkley, 1997](#)). Large sample approximations to the coefficients exist as well; see, for example, [Hollander and Wolfe \(1999, pp. 365–367\)](#).

Kruskal's γ and Kendall's τ_b are two of the many coefficients that are available; see [Liebetrau \(1983\)](#) for a review. These coefficients are so-called symmetric measures of association because their value does not depend on which variable is considered the dependent variable and which one is considered the independent variable. Sometimes, the researcher has a hypothesis about the causal mechanism underlying the association. Asymmetric measures of association, such as Somer's d ([Somers, 1962](#)), take on different values depending on which variable is the dependent variable and which one is the independent variable, and thus require articulated hypotheses of which of the two variables causally precedes the other.

Table 2 Fictional frequencies for student satisfaction with school support and engagement in extra-curricular activities.

Engagement in extra-curricular activities	Satisfaction with school support		
	Low	Medium	High
Never	4	2	0
Occasionally	2	8	3
Often	1	3	6

Miscellaneous methods

The methods discussed can be generalized in different ways. First, instead of comparing groups on a single outcome variable, one can compare them on a *combination* of several related outcome variables (e.g., [Pituch and Stevens, 2016](#)). For example, a researcher may study how different teaching methods influence student motivation using [Keller's \(2010\)](#) motivational design framework. Keller's model comprises four factors—attention, relevance, confidence, and satisfaction—that in combination invoke and maintain student motivation. Using a multivariate approach, one can test whether groups that received different teaching methods show significant differences in the combination (i.e., the profile) of these four aspects, and thus whether the teaching methods motivate students differently. If samples are small or modeling assumptions are violated, one may use a nonparametric approach to multivariate analysis of variance (MANOVA; e.g., [Puri and Sen, 1971](#)) to test between-group mean differences on multiple outcomes (i.e., H_0 : no mean differences; H_1 : mean difference on one or more outcomes). [Finch \(2005\)](#) showed that the nonparametric M-ANOVA has good statistical properties in situations in which the normality assumptions underlying the parametric MANOVA are violated.

Nonparametric data exploration for single variables

The first step in any statistical analysis should be inspecting the characteristics of the individual variables or pairs of variables. This enables researchers to make articulated decisions about subsequent statistical analyses, such as choosing between a nonparametric and parametric method. Furthermore, inspection may reveal anomalies such as outlying observations, out-of-range values, and missing data. Data screening involves descriptive statistics for central tendency, such as the mean, the mode, and the median, and spread, such as the variance, the range, and the interquartile-range. An important question is whether the data follow the normal distribution. The Shapiro–Wilk test is a well-known nonparametric test for evaluating whether the observations deviate from the normal curve. The Kolmogorov–Smirnov test is a more general, often-used nonparametric method for testing whether the data come from a hypothesized distribution, such as the normal. Often, it has less power than the Shapiro–Wilk test to detect violations of normality. Neither test makes assumptions about the population distribution of the data.

Data screening also involves graphs and other data visualization techniques. For example, the histogram is a simple nonparametric graphical method for inspecting the underlying distribution of continuous variables. A histogram is obtained by dividing the scale of the continuous variable into a finite number of consecutive non-overlapping intervals, also known as bins, and then counting the number of observations within each interval. [Fig. 2A](#) shows the histogram of reading proficiency for a sample of Dutch 15-year old children. The data were collected from the 2018 PISA study ([OECD, 2019](#)). The number and size of the bins govern

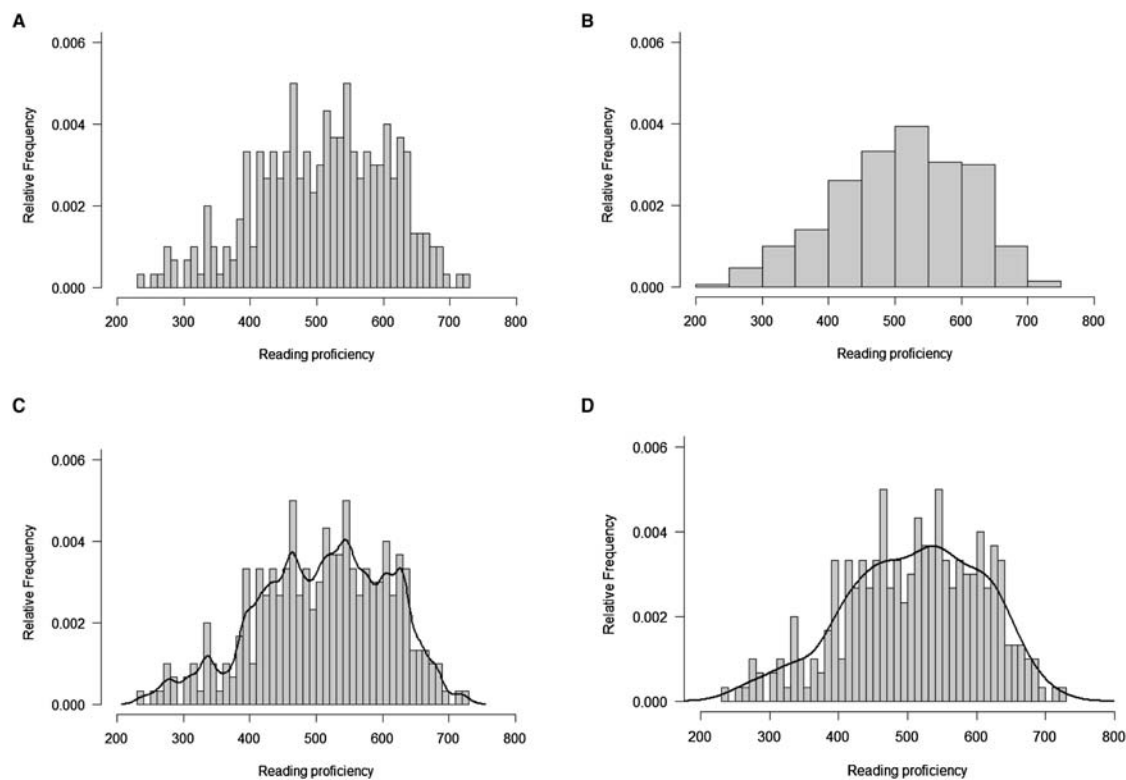


Fig. 2 Distribution of proficiency scores in a Dutch sample of 300 fifteen-year-old students using (A) the histogram and a small bin (B) the histogram and a wide bin; (C) a kernel smoother based on a small bandwidth; and (D) a kernel smoother based on a sufficiently large bandwidth.

the shape of the histogram. Small samples only allow a limited number of bins of which the height varies greatly, while some bins may be empty. Consequently, the histogram looks irregular, and gives an inaccurate impression of the underlying distribution. Reducing the number of bins yields a histogram that shows a more gradual pattern (see Fig. 2B), but has the risk that important local fluctuations will be obscured introducing bias. In practice, it may be difficult to find the bin size that optimally balances smoothness (i.e., absence of rapid local fluctuations) with accuracy.

Taking the study of a variable's distribution a step further, one may start with a fine-grained histogram and use a smoother to estimate the distribution (technically: the density) from a number of targeted focal points x_k ($k = 1, \dots, K$); see, for example, Silverman (1986) and Sheather and Jones (1991). The probability of each focal point is estimated using a weighted average of the bin to which the observation belongs and the neighboring bins, where the weights are defined by a kernel function. Kernel smoothing has the effect of bringing out the salient features of the distribution while filtering out irregularities that probably are due to sampling fluctuations. The degree of smoothness is controlled by the bandwidth-parameter, which determines the span of the neighboring observations, which are taken into account to estimate the probability of the focal point. A small bandwidth only gives weight to the observations close to the focal point, but a larger bandwidth also assigns weight to observations further away from the focal point, which results in a smoother curve. Fig. 2C and 2D show the smooth estimates of the distribution of proficiency scores obtained with kernel smoothing for small and large bandwidth, respectively. We discuss kernel smoothing, including the role of the bandwidth, in the next section.

Nonparametric regression models

Preliminary: parametric regression analyses

This section presents nonparametric regression analyses as an illustration of a more advanced nonparametric statistical method that allows researchers to study associations. Regression analysis (e.g., Fox, 2015) plays a prominent role in the social sciences. The basic model relates a single dependent variable, denoted Y , to one or more independent variables, denoted X_j ($j = 1, \dots, J$). The additive linear model is defined as

$$Y = \beta_0 + \sum_{j=1}^J \beta_j X_j + \varepsilon.$$

Coefficient β_0 is the intercept, coefficient β_j weights variable X_j , and ε is the residual representing the contribution of all neglected factors and random influences such as relevant independent variables, sampling errors and measurement errors. The model assumes that random influences cancel each other out, which implies that $E(\varepsilon) = 0$. As a result, the regression model is a model for the mean of Y conditional on the independent variables. Let $E(Y|X)$ be the expected mean dependent variable conditional on the independent variables collected in vector X , then the regression function $f(x)$ is defined as

$$f(x) \equiv E(Y|X) = \beta_0 + \sum_{j=1}^J \beta_j X_j. \quad (2)$$

The regression function in Eq. (2) also accommodates nominal or ordinal independent variables (Fox, 2015). Parameters β_j ($j = 1, \dots, J$) are of primary substantive interest. They express a partial effect, and their interpretation depends on the study design. Most often, cross-sectional data are used, and the effects have to be interpreted as follows: If two groups differ by one unit on independent variable X_j and are equal on the other independent variables, the expected mean difference in Y equals β_j . Frequently encountered interpretations like, "if X_j increases by one unit, Y is expect to increase by β_j units" are unwarranted if the regression model is fitted to cross-sectional data. For example, we may find in cross-sectional data a negative association between teacher's age and the use of new digital teaching tools. However, that does not mean that as teachers grow older, their use of digital tools will decrease. In fact, as the digitalization in education is growing fast, it is more likely that the use of digital tools will increase among all teachers. Cause-effect interpretations require at least an articulated causal diagram that describes the hypothesized causal processes, and the calculus of causation can then be used to decide whether the observed data allow consistent causal effect estimates (Pearl, 1996; Pearl and Mackenzie, 2019).

Hypotheses often tested are whether single variables contribute significantly to explaining variance in Y (i.e., $H_0: \beta_j = 0$ against $H_1: \beta_j \neq 0$), whether subsets of independent variables explain significant parts of the variance in Y (i.e., $H_0: \beta_j = \beta_k = 0$, against H_1 : "not H_0 "), and whether the amount of variance explained by the whole model, denoted R^2 , is significant (i.e., $H_0: R^2 = 0$ against $H_1: R^2 > 0$). For testing hypotheses about the regression coefficients or the explanatory power of the whole model, it is assumed (among others) that at all levels of X , the conditional distribution of Y is normal with mean 0, and has a constant variance; that is $\sigma_{\varepsilon|x}^2 = \sigma_{\varepsilon}^2$. This assumption is known as homoscedasticity. By implication we assume $\varepsilon \sim N(0, \sigma^2)$, where σ^2 is the marginal variance of the residuals.

Let e_ν be the observed residual for person ν under the estimated model. The estimated residual variance ($\hat{\sigma}^2$) equals $\frac{\sum_{\nu} e_\nu^2}{(N-K-1)}$, where K is the number of estimated regression weights β_j ($j \geq 1$). Variance $\hat{\sigma}^2$ is the pooled estimate of the conditional variances $\sigma_{\varepsilon|x}^2$. If the assumption of homoscedasticity is violated, the use of the pooled estimate $\hat{\sigma}^2$ for hypothesis testing is no longer valid,

and alternatives, such as weighted-least-squares regression should be considered (Fox, 2015). Variance $\hat{\sigma}^2$ is also invalid if there is systematic misfit due to model misspecification, for example, when fitting a linear model to data when a curvilinear relationship would be in order.

Nonparametric regression analysis

Suppose, the researcher has insufficient evidence supporting linearity, normality, and homoscedasticity, or that his/her substantive theory does not imply this degree of structure. Then, nonparametric regression methods provide an alternative for studying the relationships between variables. The terminology of nonparametric pertains to two different aspects. First, one refrains from postulating a parametric regression function $f(x)$, as in Eq. (2), but instead estimates the functional form directly from the data. Data-driven estimates of the functional form provide interesting information about relationships that may otherwise remain concealed by a priori assuming a parametric form. Second, when the assumptions about the distribution of the residuals ε are incorrect, one may use nonparametric methods to test hypotheses about the regression parameters.

For simplicity, we consider models with one explanatory variable, X . For didactic reasons we use an artificially created data set to illustrate concepts and methods. To give some context to the results, we pretend that the data concern children's age (X) and intelligence (Y). Fig. 3A shows the scatterplot. The straight solid line is the least-squares estimate of the linear regression function. In this

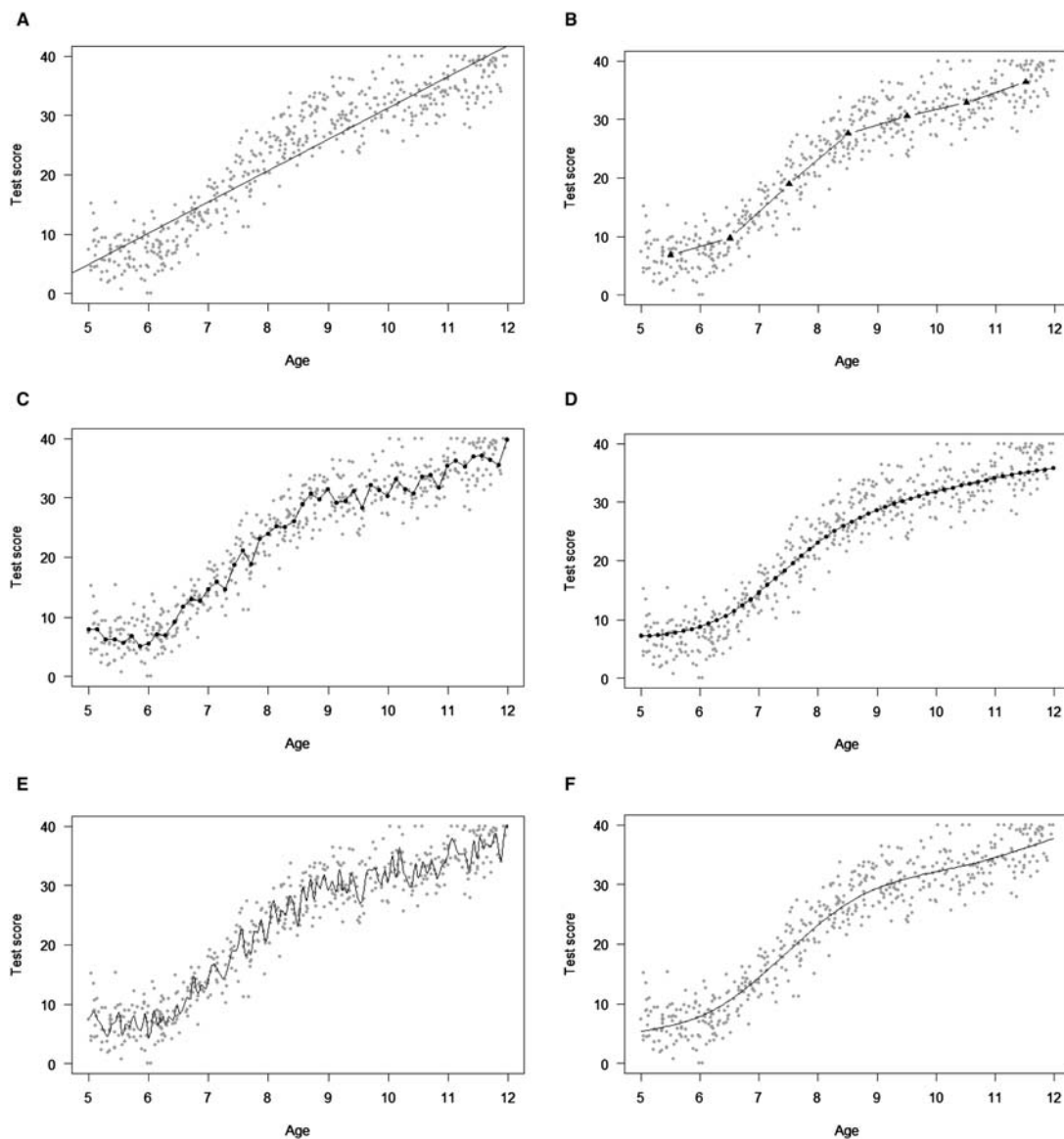


Fig. 3 Examples of different smoothers using fictional educational data: (A) the linear model; (B) binning using a large bandwidth; (C) kernel smoothing using a small bandwidth; (D) kernel smoothing using a larger bandwidth; (E) spline smoother with no roughness penalty; (F) spline smoother with a roughness penalty of 1.

example, a linear model does not describe the relationship between age and test scores well. To improve the fit of the model to the data, one could estimate a polynomial function that includes higher-order terms such as X^2 and even X^3 , but that did not yield an adequate fit either. This lack of fit for polynomial models is a common finding (e.g., Fan and Gijbels, 1996). Fan and Gijbels (1996; see also Fox (2015, p. 511) pointed out that to achieve adequate fit using polynomial functions one often needs a high-degree polynomial, which results in large variability of the parameters and unstable results, as well as unpredictable behavior of $f(x)$ for X values remote from the bulk of the predictor values.

If one cannot or is not prepared to assume a specific parametric form for the regression function, the simplest alternative is to compute the mean Y conditional on separate values of X , which we will denote by $\bar{Y}|x$, and then draw the graph that connects adjacent conditional estimates $\bar{Y}|x$. Problems similar to computing histograms may arise when only few values of Y are tied to separate values of X . Bins may solve this problem, similar to the computation of histograms. Let the bins be indexed by $k = 1, \dots, K$ and let $\bar{Y}|k$ denote the mean of all observations within the k th bin. The triangles in Fig. 3B give the local estimates for seven bins, each bin spanning a one-year age-range. The connected solid line is the estimated regression function. The nonparametric estimate better describes the trend and thus has better fit than the linear model in Fig. 3A.

Binning is simple, but the categorization of the independent variable into non-overlapping categories makes it inefficient. An improvement is to use a moving local average. Let x_0 be a point on the X -scale and let $E(Y|x_0)$ be the unknown conditional mean for which we seek the best possible estimate based on the data; that is, $E(Y|x_0)$ is the value of the intended regression function f at point x_0 . The moving average, also known as nearest-neighbor smoother, takes the m nearest values to x_0 and takes the average of these neighbors as the estimate $\hat{f}(x_0)$; that is, $\hat{f}(x_0) = \frac{\sum_{i=1}^N d_i y_i}{\sum_{i=1}^N d_i}$, where $d_i = 1$ if x_i belongs to the m nearest neighbors, and 0 otherwise.

Kernel smoothing takes local averaging a step further by assigning weights to observations that are greater the closer the observation is to the focal point x_0 . Denoting weights by w_i , the locally weighted average is obtained as

$$\hat{f}(x_0) = \frac{\sum_{i=1}^N w_i y_i}{\sum_{i=1}^N w_i}, \quad (3)$$

where the weights w_i are defined by the kernel function that uses the distance of x_i to the focal point x_0 . Eq. (3) is the Nadaraya-Watson estimator (Fan and Gijbels, 1996). Among the most frequently used kernel functions are the Epanechnikov and the Gaussian, because of their efficiency and their sound theoretical basis. The Epanechnikov kernel defines the weights w_i using a parabolic function centered around x_0 . The Gaussian kernel is defined as the bell-shaped curve

$$K(x_i, x_0) = \exp \left[\frac{-(x_i - x_0)^2}{\lambda} \right]. \quad (4)$$

The critical factor in kernel smoothing is the choice of the bandwidth parameter λ that determines the width and with that the steepness of the bell-shape. As the curve is narrower, it is also steeper and only data points close to the focal point are effective in estimating $f(x_0)$. Fig. 3C shows the estimated regression function using Gaussian-kernel smoothing for a small bandwidth. The resulting curve shows an erratic trend, which is due to the use of few data points for each estimate of the regression model. Fig. 3D shows the curve for a (sufficiently) large bandwidth, showing that as λ increases, the estimated regression curve gets smoother, but also flatter and less informative about the general trend. This is due to the use of both data points that are close and data points much further away from the focal point, so that estimates for different focal point are rather similar.

The examples show that there is a trade-off between (local) accuracy and precision, also known as the bias-variance trade-off. Choosing the bandwidth is somewhat arbitrary and difficult, because one wants to avoid capitalizing on random fluctuation (small bandwidth), while preventing potentially interesting trends from being obscured (large bandwidth). Automated procedures for selecting the optimal bandwidth exist, but we advise not to rely solely on automated procedures because the optimal *statistical* solution may not always be the optimal *substantive* solution for the application envisaged (e.g., Fox, 2015, p. 534; also Sheather and Jones, 1991). Instead, we recommend trying different bandwidths and visually decide which bandwidth value constitutes the tipping point where the regression function changes from a sufficiently smooth function to a function too erratic. The automated bandwidth may be a good starting point. For more in depth discussions on bandwidth selection, see Fox (2015, Chap. 18) and Loader (1999, Chap. 10).

Advanced smoothers

Kernel smoothers are local-constant fit methods. They are flexible, easy to implement, but also have some drawbacks (Fan and Gijbels, 1996; Fox, 2015). First, they tend to suffer from end-point bias. This means that the estimated regression curve may be biased for x -values at the boundaries of the observed X domain, because only neighboring data points on one side of the focal point are available. Second, local constant fit methods ignore the local shape of the regression function around x_0 , rendering the method less efficient and less accurate. To overcome these drawbacks, local-linear or local-polynomial regression smoothers may be used (e.g., Fan, 1993; Fox, 2015). The method assumes that a k -order polynomial regression model accurately describes the local trend

in the data. In particular, for each focal point x_0 , we estimate a polynomial regression function using weighted least-squares; that is, we find parameters β_0 and β_m ($m = 1, \dots, k$) in

$$\hat{Y} = \beta_0 + \sum_{m=1}^k \beta_m (X - x_0)^m,$$

that minimize $\sum_{i=1}^I w_i (Y - \hat{Y})^2$, where w_i is the kernel function of interest such as the Gaussian kernel (Eq. 4). Parameter estimate $\hat{\beta}_0$ is the estimated focal function value; that is, $\hat{f}(x_0) = \hat{\beta}_0$. The estimates for a carefully selected series of focal points provide the estimated regression curve. Research has shown that the local linear smoother (i.e., $k = 1$) already greatly improves local-constant kernel smoothing (Fan, 1993). We see this, for example, when the regression function around x_0 is linear, but the observations x are not symmetrically distributed around the x_0 . The local-constant is biased, but not the local-linear smoother (Fox, 2015, pp. 536–537).

Kernel smoothing and local-polynomial regression are two of many smoothing techniques that are available. They construct the regression function using locally weighted averages for a series of focal points x_0 . The resulting regression function is obtained using linear interpolation between these discrete points. The bandwidth parameter, the choice of the focal points, and the kernel function together govern the degree of smoothness. Although effective in studying trends in the data, the methods do not guarantee that the regression function is smooth and continuous everywhere along the X -domain. Piece-wise spline regression and spline smoothing (De Boor, 2001; Green and Silverman, 1994; Marsh and Cormier, 2002; Ramsay and Silverman, 1997) produce smooth and continuous regression functions.

First we consider *piece-wise spline regression*. Suppose we have N distinct observations (x_i, y_i) without ties on either x or y , which are grouped into K bins. One can estimate a regression function in each bin such that the curves in adjacent bins connect smoothly. Not only should curves from adjacent bins connect the points on the boundaries of the bins, also called knots, but in passing a knot, the slope of the curve should not change abruptly. This desired property requires at least a cubic polynomial function in each bin, so-called cubic splines. The result of all K splines is a continuous regression function, which can be uniquely defined by the function value and the second derivative at each scale point x . The number of bins and the location of the knots determine the level of smoothness.

Spline smoothing does not rely on binning, but solves the smoothness problem at the knots by putting a penalty on the “roughness” of the function. In particular, let $f''(x)$ be the second derivative of (x) . The penalized spline smoother determine the function f by minimizing

$$\sum_i [y_i - f(x_i)]^2 - \alpha \int_{x_{\min}}^{x_{\max}} f''(x)^2 dx$$

where $\alpha > 0$.

The integral $\int_x f''(x)^2 dx$ describes the roughness and parameter α expresses the desired balance between fit and degree of roughness. If $\alpha = 0$, the function is a piece-wise cubic polynomial function between adjacent x -values. It has maximum fit as it is as close to the data as possible, but shows many rapid local fluctuations as in Fig. 3E. As α increases, the function f converges to a straight line. Fig. 3F shows the curve based on the same data for an appropriate roughness penalty. Spline smoothing is a particular case of a general class of penalized optimization methods having advantageous properties. It does not require binning uses a parameter directly linked to the roughness of the estimated function, and tends to be less sensitive to end-point bias.

Order-restricted nonparametric regression

The methods discussed do not impose restrictions on the form of the regression function. They are useful descriptive tools to inform the researcher about possible relationships in the data. However, a researcher may have a priori hypotheses about the shape of the relationship. For example, when studying educational progress, it may be reasonable to assume a monotone progression in the development of children’s cognitive abilities. To test this hypothesis, one may fit a regression model that is monotonically increasing; that is, for every pair of predictor values $x_1 < x_2$, the researcher expects $\hat{f}(x_1) \leq \hat{f}(x_2)$. Using the unconstrained model as the baseline model (i.e., the full model), one can test whether the restricted model describes the data equally well as an unconstrained model. In other words, whether imposing restrictions does not lead to a model having significantly worse fit. Order-restricted regression methods have a long history (e.g., Barlow et al., 1972). Developments that are more recent include Dette et al. (2006), whose approach results in a strictly increasing function. Other approaches to constrained regression methods include monotone spline smoothing (e.g., Green and Silverman, 1994; Ramsay, 1988).

Exploring the relationship between variables is often just the first step in educational research. The ultimate goal is to answer questions about the interplay between several variables, which leads to a better understanding of teaching methods, learning strategies, and the effectiveness of educational policies. Nonparametric regression is available for multivariate data (i.e., multiple regression) and for different types of outcome variables (e.g., ordinal indicators, binary outcomes; Fox, 2015). For example, if we have two independent variables, instead of estimating a line, multivariate models estimate a smooth plane through the data points (x_1, x_2, y) . Furthermore, we can evaluate how much each single independent variable contributes to the explained variance, and use

hierarchical regression analysis to test the contribution for significance. Spline smoothing accommodates hybrid approaches, in which, for example, one independent variable is modeled using a spline smoother, and the others using a linear model. The contribution of each single independent predictor can be evaluated using their contribution to the penalized sums of squared residuals.

Gauging uncertainties: nonparametric confidence intervals

Sample statistics provide uncertain information about parameters. This uncertainty has to be taken into account to fully appreciate the results of empirical studies. Inferential statistics addresses the question to what extent similar results would be found if the study was replicated using a new sample. Uncertainty is quantified using standard errors. They express the degree of variability in the sample statistic because of sampling errors. Parametric models derive a statistic's standard error assuming that the sampling distribution follows a parametric distribution, such as the normal. For the mean, the assumption is justified by the central limit theorem, which states that if the sample size is sufficiently large (say, >25) the distribution of sample means based on random samples of the same size approximates the normal distribution. Let s denote the sample standard deviation and $\bar{X}$ the sample mean, then the standard error equals $s_{\bar{X}} \approx s_X / \sqrt{n}$. If samples are small or the normal approximation does not apply, asymptotically derived standard errors are biased and nonparametric methods may be in order. An example is the nonparametric bootstrap (e.g., [Efron and Tibshirani, 1993](#)).

The starting point for bootstrapping is the assumption that the sample is representative of the target population. Using computer technology, we create a new data set by resampling—that is, repeatedly drawing a sample—of N observations, with replacement, from the data set. Because the observations are sampled with replacement, the newly simulated data set likely contains several original observations more than once and will be different from the original data set. Thus, sample statistics will also be somewhat different and vary across samples. By repeating this process, say 1000 times, we obtain 1000 simulated data sets. The standard deviation of the test statistic across the 1000 data sets is our standard error. This idea can easily be extended to more complex situations, including nonparametric regression methods. The data set consists of N pairs of scores (x_i, y_i) . Using resampling with replacement, we construct simulated samples of pairs (x_i, y_i) and fit the nonparametric regression curve to each simulated data set. It is imperative that one uses the same focal points and the same bandwidth parameters. At each focal data point x_0 , the fifth and ninety-fifth quantile can be determined, producing a confidence envelope (to be illustrated in the next section).

Nonparametric bootstrapping is easy to implement, but [Davison and Hinkley \(1997\)](#) explain that bootstrap methods should be used carefully, and not as an escape for inadequate sampling designs, failed data collections, or inappropriate study designs. Moreover, as is true for all standard errors, nonparametric bootstrapping results are sensitive to sampling errors themselves; that is, anything we compute from a sample is liable to uncertainty. For example, if we take a small sample from the population, and by coincidence, the sample is homogeneous, then the nonparametric bootstrap standard errors may be too small, suggesting more certainty than is justified. In addition to the nonparametric bootstrap, parametric bootstrap also exists ([Davison and Hinkley, 1997](#); [Efron and Tibshirani, 1993](#)). The parametric bootstrap generates simulated data sets using a statistical model.

Illustration using PISA data

As an example of data analysis using nonparametric statistics, we consider measurements of reading proficiency from the PISA project and additional data on students' attitude toward reading. Students were asked to express whether they think reading is a waste of time on a four-points Likert scale ranging from strongly disagree ($=1$) to strongly agree ($=4$). For convenience of interpretation, we have reverse-scored this variable so that a higher score indicates a more positive attitude toward reading. [Fig. 4A](#) shows the raw data for a random sample of 500 students. Clearly, the linear model is inappropriate here. Alternatively, an ordinal logistic model (e.g., [Fox, 2015](#), Chap. 14) might be considered, but logistic models assume linearity at the logit scale and are still restrictive. Nonparametric methods are good candidates in this case for studying the relation between interval-level independent variables and an ordinal-polytomous response variable.

We used spline smoothing to model mean attitude toward reading as a function of reading proficiency. Spline smoothing captures nonlinear trends, and does not force the lower and upper asymptotes equal to the minimum item score and maximum item score, respectively. [Fig. 4B](#) shows the estimated regression function along with the bootstrapped 90% confidence bands for another subsample of 700 respondents. Notice that the regression function gives the expected item score as a function of reading proficiency. For low levels of proficiency, there is only a weak relation with attitude toward reading. However, for reading proficiency of 350 or higher, the relationship is positive. The confidence bands suggest that the trend is real and not caused entirely by sampling fluctuation.

Concluding remarks

Nonparametric methods are appropriate when certain assumptions about variable distributions are questionable and sample sizes are small. These advantages render their use especially appealing in studies where large samples are not feasible, such as studies that involve expensive experimental equipment (e.g., virtual reality glasses) or studies that require intensive data gathering procedures

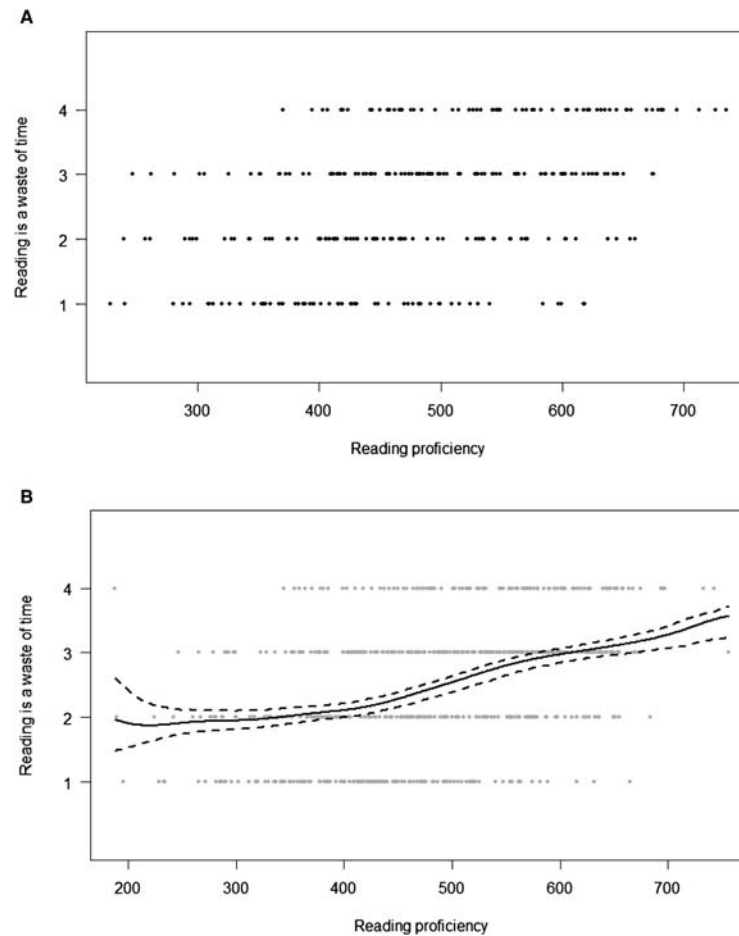


Fig. 4 Empirical example of the nonparametric regression approach to study the relation between a continuous independent variable and a categorical dependent variable having four ordered response categories.

(e.g., using think-aloud methods). Nonparametric statistical methods are also useful in preliminary research on the possible effectiveness of new educational methods at a small scale. Using non-parametric methods, and perhaps a liberal alpha level (e.g., $\alpha = 0.10$, or $\alpha = 0.20$) to reduce the chance of committing type II error (i.e., falsely rejecting an effective method), often gives a good first impression of the possible effectiveness of the new method envisaged but making as few assumptions as possible. Researchers can use this information to refine the subsequent large-scale study further.

This entry also discussed several exploratory data-driven techniques (smoothers) to study relations between variables. With that, we are also connecting to the growing field of artificial intelligence (AI) and related machine learning methods. AI has become popular in many social and scientific areas. For example, companies use data-driven exploratory techniques to uncover patterns in consumer behavior and use the knowledge they acquire to recommend products that are attractive to individual consumers. These companies usually have large amounts of data at their disposal. To take full advantage of the information in the data, they prefer starting their analyses with as few assumptions as possible rendering non-parametric modeling techniques as a perfect basis. Education shows similar trends. Due to the increasing use of computers and smart devices (e.g., tablets, smart phones, virtual reality) in the classroom, increasingly richer data become available for educational purposes. This trend led to a fast growing field of learning analytics and personalized learning, in which “smart” learning algorithms are used to tailor the instructions to the specific needs of the student. Flexible self-learning models are needed to support students optimally and nonparametric methods are excellent candidates to realize this goal.

Machine learning methods put a strong emphasis on data exploration and prediction. Cross-validation techniques are used to validate the model and to evaluate the precision of predictions in the future. Prediction-based modeling certainly also has its benefit in an educational context, but caution has to be exercised when adopting such techniques in high-stakes educational settings. Uncritical use of prediction models without a solid theoretical basis may, for example, perpetuate educational disadvantages. When it comes to important high-stakes decisions, such as assigning students to special education or certification, it is imperative to also *understand* why and how certain trends in the data emerge. This goal can only be achieved from carefully designed studies, driven by theoretically sound and articulated hypotheses and using carefully selected samples, and using appropriate statistical tools that do justice to the characteristics of the underlying data. The nonparametric methods presented in this entry serve this goal.

References

- Agresti, A., Franklin, C.A., Klingenberg, B., 2017. *Statistics: The Art and Science of Learning From Data*, fourth ed. Pearson Education Limited, Harlow, Essex England.
- Bado, N., 2019. Game-based learning pedagogy: a review of the literature. *Interact. Learn. Environ.* <https://doi.org/10.1080/10494820.2019.1683587>.
- Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D., 1972. *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression*. Wiley, New York, NY.
- Davison, A.C., Hinkley, D.V., 1997. *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge, NY.
- De Boor, C., 2001. *A practical guide to splines*. Springer-Verlag, New York, NY.
- Detle, H., Neumeyer, N., Pilz, K.F., 2006. A simple nonparametric estimator of regression function. *Bernoulli* 12. <https://doi.org/10.3150/bj/1151525131>.
- Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York, NY.
- Fan, J., Gijbels, I., 1996. *Local Polynomial Modeling and Its Applications*. Chapman and Hall, London, UK.
- Fan, J., 1993. Local linear regression smoothers and their minimax efficiencies. *Ann. Stat.* 21, 196–216.
- Finch, W.H., Davenport, T., 2009. Performance of Monte Carlo permutation and approximate tests for multivariate means comparisons with small sample sizes when parametric assumptions are violated. *Methodol.* 5, 60–70.
- Finch, H., 2005. Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology* 1, 27–38.
- Fox, J., 2000. *Nonparametric Simple Regression: Smoothing Scatterplots*. Sage, Thousand Oaks, CA.
- Fox, J., 2015. *Applied Regression Analysis and Generalized Linear Models*, third ed. Sage, Thousand Oaks, CA.
- Green, P.J., Silverman, B.W., 1994. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman & Hall, London, UK.
- Hollander, M., Wolfe, D.A., 1999. *Nonparametric Statistical Methods*, second ed. Wiley, New York.
- Katz, B.M., McSweeney, M., 1980. A multivariate Kruskal-Wallis test with post hoc procedures. *Multivariate Behav. Res.* 15, 281–297.
- Keller, J.M., 2010. *Motivational Design for Learning and Performance*. Springer, New York, NY.
- Kendall, M., 1938. A new measure of rank correlation. *Biometrika* 30, 81–89.
- Kiefer, C., Rosseel, Y., Wiese, B.S., Mayer, A., 2018. Modeling and predicting non-linear changes in educational trajectories: the multilevel latent growth components approach. *Psychol. Test Assess Model.* 60, 189–221.
- Kloke, J., McKean, J.W., 2020. *Nonparametric Statistical Methods Using R*. Chapman & Hall/CRC, Boca Raton, FL.
- Kolassa, J.E., 2020. *An Introduction to Nonparametric Statistics*. Chapman & Hall/CRC, Boca Raton, FL.
- Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621.
- Liebetrau, A.M., 1983. *Measures of Association*. Sage, Beverly Hills, CA.
- Loader, C., 1999. *Local Regression and Likelihood*. Springer, New York, NY.
- Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18, 50–60.
- Marsh, L.C., Cormier, D.R., 2002. *Spline regression models*. Sage, Los Angeles, CA. <https://doi.org/10.4135/9781412985901>.
- OECD, 2019. *PISA 2018 Assessment and Analytical Framework*. OECD Publishing, Paris, France. <https://doi.org/10.1787/b25efab8-en>.
- Pearl, J., Mackenzie, D., 2019. *Book of Why: The New Science of Cause and Effect*. Penguin Books.
- Pearl, J., 1996. *A Causal Calculus for Statistical Research*. Springer-Verlag, New York, NY.
- Pearson, K., 1895. Notes on regression and inheritance in the case of two parents. *Proc. Roy. Soc. Lond.* 58, 240–242.
- Pituch, K.A., Stevens, J.P., 2016. *Applied Multivariate Statistics for the Social Sciences*. Routledge, New York, NY.
- Puri, M.L., Sen, P.K., 1971. *Nonparametric Methods in Multivariate Analysis*. John Wiley & Sons, New York, NY.
- Ramsay, J.O., Silverman, B.W., 1997. *Functional Data Analysis*. Springer-Verlag, New York.
- Ramsay, J.O., 1988. Monotone regression splines in action. *Stat. Sci.* 425–461. <https://doi.org/10.1214/ss/1177012761>.
- Sheather, S.J., Jones, M.C., 1991. A reliable data-based bandwidth selection method for kernel density estimation. *J. Roy. Stat. Soc. B* 53, 683–690.
- Sheskin, D.J., 2007. *Parametric and Nonparametric Statistical Procedures*, fourth ed. Chapman and Hall/CRC Press, Boca Raton, FL.
- Siegel, S., Castellan Jr., N.J., 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.
- Sijtsma, K., van der Ark, L.A., 2021. *Measurement Models for Psychological Attributes*. CRC Press, Boca Raton, FL.
- Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis*. CRC Press, Boca Raton, FL.
- Somers, R.H., 1962. A new asymmetric measure of association for ordinal variables. *Am. Socio. Rev.* 27 (6), 799–811.
- Spearman, C., 1904. The proof and measurement of association between two things. *J. Psychol.* 15, 72–101.
- Wasserman, L.A., 2006. *All of Nonparametric Statistics*. Springer, New York.
- Wilcox, R.R., 2001. *Fundamentals of Modern Statistical Methods*. Springer-Verlag, New York, NY.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bull.* 1, 80–83.

Causal models

Wen Wei Loh^a and Jee-Seon Kim^b, ^a Department of Quantitative Theory and Methods, Emory University, Atlanta, GA, United States; and ^b Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	670
Articulating causal effects using potential outcomes	671
Determining which covariates generate confounding of treatment and outcome	672
Estimation of causal effects when unconfoundedness can be established	674
Regression	674
Matching	675
Estimators that combine propensity scores and outcome regression models	677
Sensitivity analysis to unmeasured confounding	678
Estimating causal effects when unconfoundedness is precluded	678
Summary and recent advances	679
Acknowledgments	680
References	680

Introduction

Identifying causal effects is critical to educational research and practice, because many questions in education concern causal relationships between interventions and (un)intended outcomes. Does repeating a grade help struggling learners in early grades, or is retention harmful to students in the long run? Do investments in private charter schools pay off in terms of improved outcomes in students and communities? Do gifted learning programs in public schools help students achieve future successes, or do they widen socioeconomic and racial disparities? Despite their importance and prevalence, the causal impacts of interventions, or treatments, can be difficult to identify in practice. The difficulty stems from each unit being observable in only either the treatment or the control condition, but never in both conditions simultaneously. For example, a student can either repeat a grade or advance to the next grade, but they cannot experience both. This so-called “fundamental problem of causal inference” (Holland, 1986) precludes simply calculating each unit’s causal effect as the difference between their treated and control outcomes, thereby necessitating assumptions to estimate causal impacts.

Randomized experiments are considered the ideal study design for causal inference but are often unethical or impractical to implement in many educational settings. Therefore, answering causal questions requires subject-matter knowledge surrounding a particular context. Assumptions are inevitably made and incorporated into research designs and analytic techniques for estimating the impacts of treatment while accounting for threats such as selection bias, omitted variables, and measurement error. This article aims to provide the fundamental concepts, issues, and analytic tools that can be used to draw causal inferences in education and more broadly in the social sciences. An added complication is that clustered or multilevel data are common in educational research, as students are generally nested within classes and schools, and thus units within clusters are often dependent. Challenges and strategies with multilevel data are described throughout the article in the context of causal models.

The rest of the article is structured as follows: **Articulating causal effects using potential outcomes** defines the causal effect of a treatment on an outcome using the potential outcomes framework, and explains core assumptions for identifying and estimating average treatment effects, such as the *stable unit treatment value assumption*, also known by its acronym SUTVA (Rubin, 1986), and *unconfoundedness*. **Determining which covariates generate confounding of treatment and outcome** presents *causal diagrams*, also referred to as *causal graphs* or *directed acyclic graphs (DAGs)*, that facilitate intuitive graphical representations of causal relations between variables grounded in theory. Causal diagrams grounded in established theory and empirical evidence can enhance understanding of underlying causal mechanisms, and improve subsequent causal inferences. In particular, they can be an essential research tool for determining which covariates should (not) be adjusted for when seeking an unbiased estimate of the causal effect. We provide examples of simple causal diagrams involving various variables, such as *confounders*, *mediators*, and *colliders*, and explain why it is important to distinguish between these different types of variables. **Estimation of causal effects when unconfoundedness can be established** reviews widely used analytic techniques for estimating treatment effects, such as covariate adjustment, *propensity scores*, and other *matching* techniques, as well as combinations of treatment selection and outcome models, and their modifications and extensions to multilevel data. Because the aforementioned estimation methods rely on the unconfoundedness assumption, we describe sensitivity analyses for assessing how different causal inferences can be due to unmeasured confounding. **Estimating causal effects when unconfoundedness is precluded** describes an example of natural experiments called *instrumental variables* for estimating causal effects when unconfoundedness cannot be reasonably assumed to hold. The final section summarizes the major concepts discussed in the article and concludes with recent advances in causal modeling. Relevant references and suggested further readings are provided throughout the article.

Articulating causal effects using potential outcomes

Suppose that research interest is in whether education in a private high school leads to higher academic achievement scores than education in a public high school. Such causal questions can be readily and precisely formalized using the *potential outcomes* framework, commonly referred to as the Neyman-Rubin Causal Model (Splawa-Neyman et al., 1990; Rubin, 1990). The framework encourages conceptualizing and explicating the cause-effect relation in terms of well-defined treatment and control conditions, e.g., education in either a private or a public high school; and the outcome of interest, e.g., math, reading, or vocabulary scores. The potential outcome of each student can then be defined as the outcome that would have been observed had that student been, possibly contrary to actuality, enrolled in either treatment or control (Rubin et al., 2004). Precise articulation of treatment and control conditions is thus essential to ensure that these potential outcomes exist and are well-defined for all members of the population of interest. For example, let Z denote a student's enrollment in either a private ($Z = 0$) or public ($Z = 1$) high school, and let Y denote an academic outcome such as their reading comprehension score. Let $Y(0)$ and $Y(1)$ denote their potential outcomes had they been enrolled in a private and public high school, respectively. Hence, it must be conceptually conceivable that each student in the population of interest could be enrolled in either a private or a public high school. We will be agnostic to the philosophical debate over, and metaphysical nature of, causation and instead refer interested readers to Morgan and Winship (2015) and Pearl and MacKenzie (2018).

Hypothetically, if the values of both potential outcomes for each student were available for analysis, each student's individual treatment effect can be determined simply as a contrast between the potential outcomes. For example, the individual treatment effect of education in a public versus private high school for each student (on the additive scale) is the difference $Y(1) - Y(0)$. The corresponding average treatment effect (ATE) across all students in the population would then be the population expectation of the individual treatment effects, i.e., $E[Y(1) - Y(0)]$. However, because exactly one of these potential outcomes can be observed for each student, depending on their observed treatment, the unobservable potential outcome exists as a counterfactual (Pearl, 2010). In other words, the observed outcome can be determined from the potential outcomes and the observed treatment as $Y = Y(0)(1 - Z) + Y(1)Z$. Thus, a cornerstone of the potential outcomes framework is the maintenance of an assumption known as the stable unit treatment value assumption (SUTVA) (Rubin, 1986). SUTVA is a basic assumption of causal effect stability that requires (i) the treatment and control conditions to be well-defined so that there are "no hidden treatment variations" among individuals experiencing the same condition, and (ii) each individual's potential outcomes to be unaffected by treatments experienced by other individuals so that there is "no (treatment) interference" (Imbens and Rubin, 2015). As discussed in Hong and Raudenbush (2006), SUTVA becomes more challenging in multilevel data because of the intrinsic dependency within clusters and potential peer or spillover effects. We review recent advances under such settings in a later section.

When treatment is randomized, routine statistical methods comparing the outcomes between treated and control individuals can be used to estimate the average treatment effect. The estimate is unbiased for the causal impact of treatment in the population of interest, regardless of the result from a null hypothesis test of no treatment effect. Random assignment is desirable for causal inference because it renders the treatment and control groups equal in expectation on all observed and unobserved dimensions except for the treatment status. A thoughtfully designed and properly conducted randomized experiment ensures that all competing explanations for any differences in the observed outcomes between the treatment and control conditions are disavowed. A treatment assignment mechanism is therefore termed as being *unconfounded* when the observed condition to which a participant is assigned possesses no information about their potential outcome under the unobserved condition.

When treatment is not randomly assigned, a more involved approach that carefully adjusts for common causes of both treatment and outcome is required to estimate the causal effect without bias. Continuing the example above, any differences in academic outcomes when comparing students in private and public schools are likely not due to the choice by itself causing the outcomes, but rather due to the (unmeasured) resources that are available to students who are more likely to enroll in a private school, and thus obtain better outcomes. For example, private school students may be economically advantaged to pay the tuition required in private schools, and experience more engaging environments for learning outside of school. Hence in observational studies where individuals can select their experienced treatment, non-causal (or "spurious") correlations between the treatment and outcome can arise, due to their mutual associations with such confounders. When these confounders are not adjusted for, because they were either unmeasured, or measured but excluded from the analysis, they induce confounding bias in the treatment effect estimate. Let L denote a determined subset of covariates that generate confounding of the treatment-outcome relation. The unconfoundedness assumption states that the probability of selecting either treatment or control is independent of the joint potential outcomes (and any function of them), within strata defined by all combinations of values of L . When all confounders L are measured and this assumption is plausibly satisfied, individuals with the same values of L are equally likely to have selected treatment (also termed "comparable" or "exchangeable") regardless of their observed treatment status, so that treatment is effectively randomized within strata defined by L . Confounding is therefore a causal concept that does not apply when simply assessing statistical correlations (Hernán and Robins, 2020). The *positivity* assumption further states that conditional on L , the probabilities of selecting treatment, and selecting control, are both non-zero; i.e., $0 < \Pr(Z = 1|L = l) < 1$ for all unique values of $L = l$. We describe in the next section how researchers may use theoretical knowledge and prior empirical evidence to determine which covariates associated with treatment, outcome, or both treatment and outcome, are appropriate for adjustment to eliminate confounding bias.

Determining which covariates generate confounding of treatment and outcome

When treatment is non-randomized, covariates that generate confounding of the treatment-outcome relation must be adjusted for to eliminate as much spurious correlation as possible from the causal effect estimate. But routinely including all measured variables that covary with either treatment, or outcome, or both, in the analysis can yield biased estimates. Therefore, determining whether to adjust for a particular covariate when assessing the causal effect requires carefully considering the causal relations between the covariate and treatment, and between the covariate and outcome. The causal diagram approach to causality (Pearl, 2009, 2012) is invaluable for visualizing and explicitly evaluating such causal relations. Causal diagrams that accurately represent data-generating processes can provide clear justifications for systematically determining which covariates must be adjusted for to eliminate non-causal associations when estimating the causal effect (Elwert, 2013). Causal diagrams are particularly appealing because they resemble traditional path diagrams in social science research (Pearl, 1998; Bollen and Pearl, 2013). However, causal diagrams encode causal relationships that are completely nonparametric. Unlike path diagrams, it is unnecessary to specify the nature of the functional dependence of each variable on its causally antecedent variables. Furthermore, causal diagrams are drawn based on extant subject matter theory and prior empirical evidence, whereas path diagrams are (often) drawn based on the observed data.

We briefly introduce basic terminology and features of causal diagrams using simple examples. In each example we will focus on a single covariate and its causal relations with treatment and outcome. Variables are represented in a causal diagram using nodes. For example, in the causal diagrams shown in Fig. 1, we will use the node Z to represent treatment, and the node Y to represent outcome. The directed *edge* (i.e., single-headed arrow) from one node to another symbolizes a causal effect that the variable at the origin exerts on the variable at the terminus. For example, in the causal diagrams shown in Fig. 1, the arrow from Z to Y depicts the causal effect exerted by treatment Z on outcome Y . The presence of an arrow merely represents a causal effect of an unknown magnitude, whereas the absence of an arrow precludes any possible causal effect. A *path* is any sequence of edges pointing in any direction that connects one variable to another, with no recurring variable on that path. For example, in the causal diagram of Fig. 1a, the path $Z \leftarrow L \rightarrow Y$ links treatment Z and outcome Y via a shared common cause L . In general, any path with one arrow pointing to treatment, and another pointing to outcome, is termed a “back-door” path from treatment to outcome.

When the confounder L is unadjusted for in the analysis, its absence will generate a non-causal (or spurious) empirical association between Z and Y that induces bias in the causal effect estimate. In other words, adjusting for L “blocks” the path between Z and Y so that the causal effect of Z on Y can be estimated without bias; see e.g., Hayduk et al. (2003). However, confounders are not necessarily direct causes of both treatment and outcome. Consider the causal diagrams in Figs. 1b and 1c, where L is associated with either Z due to an unmeasured common cause U_{Lz} , or Y due to an unmeasured common cause U_{Ly} , respectively. Even though L causally influences only either Z or Y , failing to adjust for L will lead to a spurious association between Z and Y . Therefore, the back-door paths between Z and Y can be blocked by adjusting for L , even when U_{Lz} or U_{Ly} are unmeasured and cannot be adjusted for. In all three scenarios presented in Figs. 1a to 1c, it is appropriate (and recommended) to adjust for the confounders L .

We now turn to other covariates that may be associated with both treatment and outcome. A *directed path* is a path in which all edges point in the same direction. For example, in Fig. 1d, $Z \rightarrow M \rightarrow Y$ is a directed path from treatment to outcome that intersects the intermediate variable M , commonly referred to as a *mediator*. Notwithstanding M being associated with both Z and Y , M does not lie on a back-door path from Z to Y , and should not be adjusted for. It is only appropriate to adjust for M when interest is in the so-called “direct effect” of Z on Y transmitted via causal pathways that do not intersect M . Evaluating the treatment effect transmitted along directed paths intersecting possible mediator(s) is the focus of (causal) *mediation analyses* that seek to partition the treatment effect into direct and indirect components. Otherwise, M should be unadjusted for when assessing the (“total”) effect of Z on Y via all possible causal pathways leading from Z to Y .

A variable is a *collider* along a particular path if it has two arrows pointing directly at it. For example, in the causal diagram of Fig. 1e, C is a collider along the path $Z \rightarrow C \leftarrow Y$. An unadjusted collider blocks a path so that there are no non-causal associations between the origin and terminus along that path. However, adjusting for a collider, such as the common effect C of both Z and Y , will create a spurious correlation between Z and Y . Hence C should not be adjusted for when estimating the causal effect of Z on Y . It is important to note that C need not be directly caused by both Z and Y , as shown in Figs. 1f and 1g, for C to be a collider along a path between Z and Y . For example, even when C is merely associated with Z due to an unmeasured common cause U_{Cz} , C is nonetheless a collider along the path $Z \leftarrow U_{Cz} \rightarrow C \leftarrow Y$. Additionally, when C is merely associated with Y due to an unmeasured common cause U_{Cy} , C is a collider along the path $Z \leftarrow U_{Cz} \rightarrow C \leftarrow U_{Cy} \rightarrow Y$. Adjusting for C induces “collider (stratification) bias,” and colliders should thus remain unadjusted for (Greenland, 2003b; Cole et al., 2009; Ding and Miratrix, 2015; Thoemmes, 2015). Note that the causal diagrams in Figs. 1b and 1f differ only in whether Y is an effect of L , or a cause of C , respectively. The direction of the causal effect thus determines whether to adjust for the covariate, or not. A variable is a *descendant* of another variable if it can be reached by a directed path. For example, in the causal diagram of Fig. 1h, D is a descendant of C . No collider (nor any descendant of a collider) along a causal path from Z to Y should be adjusted for, otherwise it can create spurious correlations between treatment and outcome that induce bias in the effect estimate.

It may therefore be tempting to simply rule out adjusting for all variables—besides outcome—that are measured temporally after treatment, but it is important to note that such variables need not be causally influenced by treatment. For example, in Fig. 1b, if L shares a common cause U_{Lz} with Z , such as when L and Z are different measures of the same underlying latent

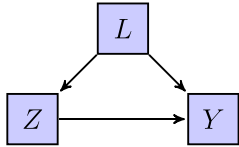
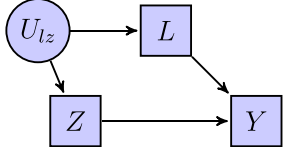
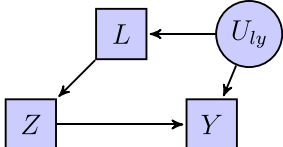
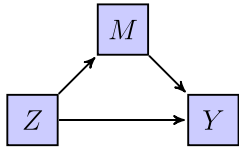
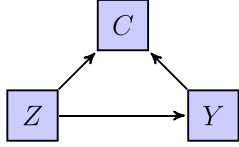
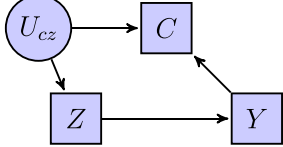
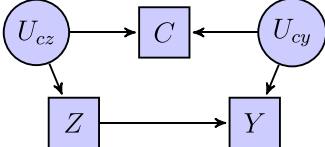
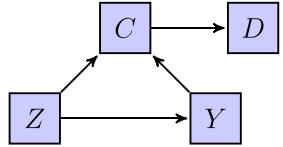
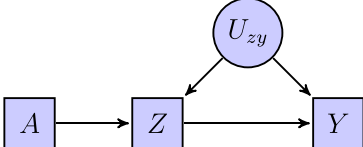
Causal diagram	Whether a covariate should or should not be adjusted for
a 	Confounders L that are common causes of Z and Y must be adjusted for.
b 	Confounders L that causally influence Y , but are only associated with Z due to an unmeasured common cause U_{lz} , must be adjusted for.
c 	Confounders L that causally influence Z , but are only associated with Y due to an unmeasured common cause U_{ly} , must be adjusted for.
d 	Intermediate variables M on a causal path from Z to Y should not be adjusted for when estimating the (total) effect of Z on Y via all possible causal paths.
e 	C is a collider on a path from Z to Y , and should not be adjusted for.
f 	C is a collider on a path from Z to Y (with U_{cz} being unmeasured), and should not be adjusted for. Another possibility for C to be a collider on a path from Z to Y is if U_{cy} is an unmeasured common cause of C and Y , and Z causally influences C instead.
g 	C is a collider on a path from Z to Y (with U_{cz} and U_{cy} both being unmeasured), even though C is not a causal descendant of either Z or Y , and thus should not be adjusted for. Bias resulting from adjusting for C is variously referred to as “collider (stratification) bias,” “included variable bias,” or “M-bias.”
h 	D is a descendant of a collider C along a path from Z to Y , and should not be adjusted for, regardless of whether C is adjusted for or not.
i 	A affects only Z , and is unassociated with Y , and thus should not be adjusted for. When identifying the causal effect of Z on Y is prohibited by the unmeasured confounder U_{zy} , A can act as an instrumental variable to facilitate estimating the effect of Z on Y under certain assumptions.

Fig. 1 Causal diagrams describing examples of causal relations between a third observed variable with treatment Z and outcome Y . Confounders are denoted by L , mediators are denoted by M , colliders are denoted by C , and instruments are denoted by A . Rectangular nodes denote observed variables, while round nodes denote unmeasured variables.

construct U_{lz} , then L should be adjusted for even if it is measured after treatment Z is recorded (and a causal effect of Z on L can be precluded). Therefore, the causal influence of treatment on such “post-treatment” or “treatment-induced” variables, such as M , C or D , must be grounded in established theoretical knowledge or prior thoughtful experimentation. Finally, we consider the setting depicted in the causal diagram of **Fig. 1i** where the variable A causes Z but does not causally influence Y directly.

Instrumental variables such as A should not be adjusted for when estimating the causal effect of Z on Y because they produce inefficient estimators without reducing confounding bias (Brookhart et al., 2006; Wooldridge, 2016). When an unmeasured confounder U_{zy} precludes blocking the back-door path $Z \leftarrow U_{zy} \rightarrow Y$, the instrumental variable A permits causal inference under certain scenarios. We discuss this in a later section.

Mediators M , colliders C , descendants of colliders D , and instrumental variables A are examples of covariates that must be handled cautiously when carrying out confounding adjustment; see Pearl (2013) for further details on their impacts on confounding bias. A causal diagram containing all the (un)observed variables in a study can be constructed by starting with the treatment-outcome causal relation, then adding covariate nodes and arrows based on relevant existing literature and procured expert knowledge. A causal diagram that accurately represents structural relations grounded in established theory can improve understanding and reasoning of underlying causal mechanisms (Elwert, 2013; Morgan and Winship, 2015; Rohrer, 2018; Grosz et al., 2020).

When more complex causal relations among the covariates are known, all possible back-door paths between treatment and outcome must be carefully considered. A more general “back-door criterion” provides sufficient conditions to determine whether all non-causal associations can be eliminated by adjusting for a (minimal) set of confounders L (Pearl, 2009, Chapter 3). The criterion consists of two parts. First, the variables in L block every back-door path between the treatment and the outcome (i.e., with an arrow pointing to treatment). Second, no variable in L is a causal descendant of treatment. Suppose L satisfies the back-door criterion with respect to a treatment Z and outcome Y . In that case, unconfoundedness of the causal effect of Z on Y given L holds regardless of the specific functional form or distribution of the variables (L, Z, Y) . Assessing whether the back-door criterion holds for a candidate adjustment set can be challenging, because a variable that is a non-collider on one path may be a (descendant of a) collider on another path. The R (R Core Team, 2021) package DAGitty (Textor et al., 2017) can aid visualizing causal diagrams, and automate determining whether a minimally sufficient subset of covariates exists to establish unconfoundedness. A minimal adjustment set is desirable because it reduces the number of variables for which functional dependencies with the treatment, outcome, or both treatment and outcome, must be specified when estimating the causal effect in the subsequent data analysis.

The back-door criterion provides a simple graphical criterion for determining whether unconfoundedness is possible when the structure of the causal diagram is known. But establishing a correct and realistic causal structure for all variables based solely on theoretical knowledge or prior empirical evidence can be difficult, if not impossible, in large observational studies. Under such settings, one possibility is to select covariates for confounding adjustment based on the observed data, provided that there exists a subset of (measured) covariates sufficient to establish unconfoundedness. However, valid inference following variable selection using the same data is no longer guaranteed when using routine variance estimators, with confidence intervals that are too narrow because the uncertainty introduced by the variable selection process is unaccounted for (Huang, 2020). Such data-adaptive covariate (or confounder) selection techniques for causal inference are beyond the scope of this article; see VanderWeele (2019) and Witte and Didelez (2019) for recent reviews. The plausibility of the unconfoundedness assumption that relies on a set of measured covariates should thus be scrutinized in each unique application toward a more resilient justification that the causal effect of interest can be unbiasedly estimated. Well-grounded justifications of the covariates either included in or omitted from adjustment should thus form part of the analysis (Foster, 2010; Austin and Stuart, 2015; Kainz et al., 2017).

Estimation of causal effects when unconfoundedness can be established

Now suppose that a (minimal) subset of covariates L has been determined, so that there are no variables - measured or unmeasured - that would undermine the unconfoundedness assumption. We may then proceed to estimate the ATE while adjusting for L . We describe a few approaches that build on prevalent regression models in the education research literature, but differ in terms of how the models are utilized to obtain an effect estimator.

Regression

A straightforward and common estimation approach is to simply regress the outcome Y on the treatment Z and the covariates L jointly in a multiple regression model. Traditional regression models are already widely used to describe the associations between independent and dependent variables, without necessarily alluding to causal relations. However, these techniques can also be employed as a parametric tool to adjust for confounding when assessing the causal effect of the treatment of interest on the outcome. When the causal effect is assumed to be homogenous and additive, the regression coefficient for the treatment is endowed with a causal interpretation, by essentially comparing the predicted outcomes between individuals with different treatment levels but having the same values of L . However, such an approach requires correctly specifying the associations between the covariates and the outcome in the assumed regression model.

Two possible issues arise when relying solely on routine outcome regression models for causal inference. First, the covariates may be moderators of the treatment-outcome relation, so that the treatment effect is modified by (possibly non-linear) combinations of the covariates. When the assumed model is overly parsimonious and includes only a single main effect for treatment, misspecification biases can arise because the heterogeneity of the treatment effect is implicitly averaged over the distribution of the covariates. A sufficiently flexible regression model with non-linear and higher order covariate-treatment interaction terms may be used to explicitly parameterize possibly heterogeneous treatment effects. However, conceptualizing such a model that

gives unbiased predictions - without impractically high variance of prediction - over the whole covariate space can be difficult or unfeasible in practice when there are many covariates, due to the curse of dimensionality. Second, biases due to extrapolation can arise when there is little overlap in the distributions of L between the treatment and control groups, particularly when L are strong predictors of Z , resulting in multicollinearity in the outcome model. The lack of empirical information about the treatment-outcome relation within disparate values of L that differ between the treated and untreated individuals requires extrapolating the covariate-specific outcome mean for the treated to the observed covariate levels for the untreated (and vice versa). Such extrapolation biases cannot be detected using routine checks of model assumptions, and can thus be difficult to diagnose and resolve (Vansteelandt and Daniel, 2014).

When the outcome of interest is collected in a nested or clustered setting, we can extend standard regression or single-level linear modeling approaches to multilevel models to account for the hierarchical structure of the data and dependency among units within clusters (Raudenbush and Bryk, 2002; Snijders and Bosker, 2011; Goldstein, 2011; Hox et al., 2017). Multilevel models, also referred to as hierarchical linear models or mixed-effects models, can incorporate different functional forms of covariates at different levels of the hierarchy, can attend to cross-level interactions, and allow for random (varying) coefficients of the intercepts and predictor effects across clusters. A treatment variable Z can be at a lower or upper level (e.g., treatment at the student level or school level), and the treatment effects at a lower level can be assumed to be constant across all units (i.e., fixed) or to vary across the clusters (i.e., random) in multilevel models. Strategies and techniques for causal inference specifically designed for such data structures will be described later in this section.

Matching

The aforementioned extrapolation biases can be avoided by eschewing a model-based approach for the associations between the outcome and covariates. Matching is a (non-)parametric method of adjustment for non-randomized treatment assignment that forms strata with exchangeable treated and untreated individuals (Rosenbaum, 2002). For example, suppose that the covariates L are low-dimensional and discrete, so that there are only a few unique values of L . Then individuals within non-overlapping strata defined by values of L are entirely indistinguishable from each other in all ways with respect to L , except for the observed treatment Z and (differences in) their potential outcomes. A straightforward comparison of observed outcomes between treated and untreated individuals within each stratum, followed by a weighted average of the (mean) differences with weights being the proportions of individuals in each stratum, would yield an unbiased estimator of the ATE. Matching therefore seeks to mimic a block-randomized design where each stratum corresponds to a block in which the distribution of L is identical among the treated and untreated individuals, and treatment is randomly assigned within each stratum.

However, full stratification on L to achieve unconfoundedness unavoidably results in fine-grained stratum-specific effects, even if interest is merely in the ATE. Furthermore, in most realistic settings, there are typically many variables in L , (some of) which may be continuous (i.e., taking on many possible values). Hence, unique strata defined using all possible combinations of values of L may not contain both treated and untreated individuals, even in very large sample sizes, due to the curse of dimensionality. This violates the aforementioned positivity assumption because individuals within a particular stratum have zero probabilities of being assigned to either treatment or control. While stratum-specific effects can be computed only for strata within which positivity and unconfoundedness hold, such strata may not be representative of the original population.

A commonly used alternative to a full stratification on the covariates is to model instead the associations between the treatment and covariates. The *propensity score* is the conditional probability that an individual would have been assigned treatment given their covariates (Rosenbaum and Rubin, 1983b), and is widely used in educational and psychological research (McCaffrey et al., 2004; Cook et al., 2009; Harder et al., 2010; Shadish and Steiner, 2010; Austin, 2011; Thoemmes and Kim, 2011; Lanza et al., 2013; West et al., 2014; Keller and Tipton, 2016; Dong et al., 2020; Powell et al., 2020). A key property of the propensity score is that the covariate distributions are rendered (approximately) similar, or “balanced,” across the treated and untreated groups by conditioning on the propensity score, provided unconfoundedness is met, and the model for the treatment given the covariates is correctly specified.

When the covariates in L are sufficient to eliminate all confounding of the treatment-outcome relation, then adjustment for the propensity score is also sufficient to achieve unconfoundedness. In other words, the potential outcomes and treatment are conditionally independent among individuals with the same propensity score, so that treated and untreated individuals with the same propensity score are exchangeable. The propensity score can then be viewed as a single stratifying variable to construct matched sets of treated and untreated individuals with identical propensity score values. The outcomes between treated and untreated individuals within each stratum - defined by a matched set - can then be contrasted. The average treatment effect is the weighted average of the average stratified treatment effect estimates. Because the propensity score is scalar, it permits using simple and transparent methods for confounding adjustment that would otherwise be impossible in the presence of high-dimensional covariates. Note that estimated propensity scores do a better job of balancing the observed covariates than the true propensity scores would, because the estimated propensity scores correct for the chance imbalances that characterize any finite sample (Rosenbaum, 1987a).

However, in finite samples, it may be unfeasible for different individuals (with different treatment levels) to have exactly the same estimated propensity score. Individuals with similar but unequal propensity scores must therefore be matched to the same stratum using a pre-determined similarity criterion. For example, subclassification (or interval matching) partitions individuals into equally-sized subgroups based on the sample quintiles of the propensity score distribution (Cochran, 1968). But there is

no guarantee that each subgroup will contain both treated and untreated individuals so that the positivity assumption is upheld. A data-adaptive alternative is full matching (Hansen, 2004), which creates partitions that minimize the within-subgroup average difference in the estimated propensity score (and thus the distribution of the covariates), while ensuring that at least one treated and one untreated individual is in each subgroup. Other matching algorithms seek to optimize the within-sample balance of the (observed joint distribution of the) covariates for the treated and untreated individuals within each subgroup (Diamond and Sekhon, 2013). Various R packages implementing different matching algorithms include *cem* (Iacus et al., 2009), *Matching* (Sekhon, 2011), and *MatchIt* (Ho et al., 2011). Stuart (2010) provides a review of matching algorithms.

An alternative to constructing matched sets using the estimated propensity scores is to directly employ the propensity scores as inverse weights when comparing the outcomes between treated and control individuals (Lunceford and Davidian, 2004). Let $p(L_i)$ be the propensity score for individual i whose covariate values are L_i . Define the inverse probability (of treatment) weights as:

$$W_i = \frac{Z_i}{p(L_i)} + \frac{1 - Z_i}{1 - p(L_i)} \quad (1)$$

Each individual is weighted by the inverse of their probability to receive their observed treatment given the observed covariates, thus creating a pseudo-population where Z is no longer associated with L and there is no confounding bias when comparing treated and untreated individuals in such a pseudo-population. Both the treatment and control groups then have the same weighted distribution on (observed) common causes of treatment and outcome as the pseudo-population. Estimation predominantly proceeds by first fitting a logistic regression model for the treatment given the covariates, commonly referred to as the *propensity score model*, to the observed data, such that covariate balance is achieved. The fitted propensity score for each individual is obtained by plugging in the maximum likelihood estimates of the model parameters for their unknown values. Let $\widehat{W}_i$ denote the estimated weight from plugging in $\widehat{p}(L_i)$ for $p(L_i)$ in (1). An estimator of the ATE is then the treatment coefficient in a weighted linear regression of the outcome versus treatment (only) using $\widehat{W}_i$ as the individual weights. An asymptotic Wald $100(1 - \alpha)\%$ confidence interval (CI) can be calculated using the robust (sandwich) variance estimator, and the critical values from a standard Normal distribution (Austin and Stuart, 2017). Note that no model for the outcome, conditional on treatment and the covariates, is required to be (correctly) specified for unbiased estimation of the ATE. A benefit of weighting individuals by their inverse propensity score is that known survey weights - constructed to account for complex observational study design features - can be readily accommodated by simply multiplying each individual's survey weight by W_i .

A shortcoming of inverse propensity score weighting is its reliance on the estimated propensity scores. Estimated values close to zero or one can lead to disproportionately large weights that give undue influence to a small subset of individuals relative to the remainder of the sample. In particular, the resulting skewness and sampling variability in the tails of the propensity score distribution occasionally result in erratic estimates relative to traditional outcome regression adjustment. Nonetheless, the larger standard errors more honestly reflect the uncertainty about the causal effect because no empirical information (or assumptions) about the joint distributions of covariates and potential outcomes are explicitly used (Tan, 2006). One solution in practice is to trim back the weights (Petersen et al., 2012), by progressively truncating extreme values to interior percentiles of the initial estimated distribution (i.e., first back to the 1st and 99th percentiles, then the 2.5th and 97.5th percentiles, and so on.) However, this may render bias in the effect estimate. A recommended solution is to consider alternatives to standard logistic regression for the propensity score. For example, *covariate balancing propensity scores* (Chattopadhyay et al., 2020; Imai and Ratkovic, 2014) directly maximize balance of the (observed) covariate distributions across the treated and untreated groups when estimating the optimal model parameter values that determine the weights. Alternatively, propensity scores may be estimated using computationally intensive but flexible machine learning techniques (Lee et al., 2010), such as generalized boosted models (McCaffrey et al., 2004; Setodji et al., 2017), with an R implementation in the *twang* package (Cefalu et al., 2021); classification trees (Luellen et al., 2005); or neural networks (King and Zeng, 2001; Setoguchi et al., 2008). Webster-Clark et al. (2021) provide a recent review of propensity score methods, while Kainz et al. (2017) offer practical recommendations for using propensity scores.

In the presence of multilevel data, propensity scores that account for the nested or clustered structure while directly optimizing covariate balance between the treatment groups can be employed (Kim et al., 2017; Yang, 2018; Fuentes et al., 2021). Because clusters constitute natural strata, matching can be conducted within clusters if sample sizes are not too small and the numbers of both treated and untreated cases are sufficient in all clusters (Kim and Seltzer, 2007). In many empirical studies, however, sample sizes may be too small or the positivity assumption will be violated in some or many clusters, and thus within-cluster matching is not feasible. In such circumstances, across-cluster matching techniques can effectively "borrow units" from other clusters to improve overlap between treated and untreated cases (Thoemmes and West, 2011; Steiner et al., 2012). Combinations of within- and across-cluster matching have also been proposed (Arpino and Cannas, 2016; Zubizarreta and Keele, 2017). As a matching strategy for multilevel data, Kim and Steiner (2015) applied latent class selection models for evaluating heterogeneity in the selection processes and classifying clusters into a small number of homogeneous groups (i.e., latent classes), and then pooling cases within latent class. This "across-cluster within-classes" matching method can be used when the class memberships are known (manifest classes) or unknown (latent classes). Similarly, Kim et al. (2016) used finite mixture models for estimating heterogeneous treatment effects where private tutoring was effective for a small class of students but not for the rest. The across-cluster within-class matching also weakens an important assumption for causal models in that the unconfoundedness assumption is required within each class but not across the classes. There has also been increasing use of machine learning methods for causal inference with observational multilevel data (Suk et al., 2020, 2021). In summary, a variety of procedures are possible for matching in the presence of multilevel data structures.

Estimators that combine propensity scores and outcome regression models

Without a perfect specification of the propensity score model, one cannot be assured that consistent estimates can be obtained. Because propensity scores are intended to eliminate the treatment-confounder associations by weighting the observed sample so that it is representative of a pseudo-population, systematically incorrect estimated propensity scores can generate systematically incorrect weighting schemes that yield inconsistent estimates of treatment effects.

We now discuss three possible estimation methods that combine the use of propensity scores and outcome regression models. Such “doubly-robust” estimators are appealing because they offer protection against model misspecification biases when either model is correctly specified (Robins, 2000; Bang and Robins, 2005). The estimators differ in terms of how the estimated propensity scores are utilized in the outcome regression model. A straightforward approach is to include the propensity score as an additional predictor in an outcome model, referred to as “regression adjustment for the propensity score” (D’Agostino, 1998). Suppose that the outcome is continuous, and regressed against treatment, covariates, and the propensity score in a linear model. Provided that the relation between the outcome and the true (but unknown) propensity scores is correctly specified, the ordinary least squares estimator of the treatment coefficient, with the estimated propensity scores used in place of their true values, is an unbiased estimator of the ATE in large samples when the propensity score model is correctly specified and the exposure effect is homogenous. Under these conditions, the ordinary least squares estimator of the ATE does not necessitate correct specification of the outcome regression model and is thus not vulnerable to extrapolation biases under that model. Because the propensity scores are a summary of the covariates that suffices to ensure ignorability, the propensity scores completely remove the bias in estimating the ATE and any additional information provided by them merely serves to make the estimate more precise (Vansteelandt and Daniel, 2014).

Conversely, suppose that the outcome model is correctly specified before the propensity scores are added, in the sense that the outcome model given treatment and covariates (i.e., without the propensity scores) are unbiased for the true data-generating mechanism of the potential outcomes. Adding (any function of) the propensity scores merely results in an overfitted outcome model, and adds mean-zero noise to the predicted outcomes, so that the choice of propensity score model does not affect the validity of the estimator. Hence the ATE estimator is unbiased even if the propensity score model is misspecified (Vansteelandt and Daniel, 2014). In other words, an unbiased estimator of the ATE can be obtained so long as either the outcome model or the propensity score model is correctly specified. In general, functions of the estimated propensity scores can be incorporated into the outcome model as additional predictors. For example, smoothing splines with knots at the sample quintiles of the logit of the propensity score (i.e., linear predictor in the propensity score model) (Little and An, 2004), or indicators of strata membership following subclassification on the propensity score using the sample quintiles (Kang and Schafer, 2007), can be used.

However, when a nonlinear model for the outcome is employed, such as with a canonical link function in a generalized linear model, the benefits of adjusting for the (estimated) propensity score in the outcome model are more subtle. The protection offered against model misspecification is then limited to only tests of the null hypothesis of no treatment effect that are based on a robust variance estimator. Under such settings, the hypothesis tests have the correct type I error rate (in large samples) when either the propensity score model, or the outcome regression model, is correctly specified (Vansteelandt and Daniel, 2014).

An alternative estimator can be obtained by weighting the observations in the outcome regression model (given treatment and covariates) using the weights W_i as defined in (1). Estimation proceeds by first estimating W_i , then fitting the outcome regression model with each observation weighted using the estimates of W_i . The estimator of the ATE is then the weighted least squares estimator of the treatment coefficient, and is consistent so long as either the outcome model or the propensity score model is correctly specified (Robins et al., 2000).

The approaches so far encode the ATE as the treatment coefficient in a parametric model for the outcome. A third approach avoids such an explicit parametrization by augmenting the inverse probability (of treatment) weights estimator with an outcome model to exploit information about the treatment effects. The potential outcomes can then be imputed, and an estimator of the ATE calculated without necessarily being encoded as a parameter in the outcome model. An example of a so-called *augmented inverse propensity weighted* (AIPW) estimator (Robins et al., 1994) can be defined as follows. Let $\widehat{W}_i$ denote the estimated weights, as defined in (1), for individual i . Let $\widehat{E}(Y_i|Z_i, L_i)$ denote the outcome regression model fitted to the observed data that we will use to predict the (potential) outcomes for each individual under different (hypothetical) treatments. The estimator of the ATE given a sample of size n is then:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n Z_i \widehat{W}_i \{Y_i - \widehat{E}(Y_i|Z_i = 1, L_i)\} - (1 - Z_i) \widehat{W}_i \{Y_i - \widehat{E}(Y_i|Z_i = 0, L_i)\} \\ + \{\widehat{E}(Y_i|Z_i = 1, L_i) - \widehat{E}(Y_i|Z_i = 0, L_i)\}. \end{aligned} \quad (2)$$

Such AIPW estimators have a number of appealing properties (Glynn and Quinn, 2010). In particular, they are consistent if either the propensity score model or outcome model is correctly specified, and asymptotically unbiased when both are correctly specified. Hence, they permit reducing the reliance on traditional parametric regression models that demand (correctly) specifying the exact relationships between the covariates and treatment, or outcome, or both. These relationships may be complex and inadequately understood using existing theory, resulting in biases due to inaccurate predictions of the propensity scores, or outcomes, or both. To mitigate the risks of biases due to misspecified models, richly parameterized models with multiple higher-order interaction terms and flexible functional forms (e.g., cubic or smoothing splines) for the predictors in L can be fitted instead. However, it may be difficult or unfeasible to fit such models in practice, because the number of parameters will be very large relative to the sample size.

Such constraints may be overcome by considering (*supervised*) machine, or statistical, learning algorithms designed to accurately predict dependent variables given a set of independent variables (or features). For example, nonparametric regression methods (such as generalized additive models; [Hastie et al., 2009](#)), or tree-based algorithms (such as random forests; [Breiman, 2001](#)) and Bayesian additive regression trees (BARTs; [Chipman et al., 2010](#)), can be used to estimate the conditional expectations of the potential outcomes, while permitting flexible and arbitrarily complex functional forms of L . Because these algorithms can be trained to optimize predictive accuracy while incorporating a large number of parameters, they can outperform traditional parametric models for the accurate prediction of conditional expectations. For example, AIPW estimators using machine learning techniques are implemented in the R package AIPW ([Zhong et al., 2021](#)). But caution needs to be exercised to avoid overfitting that can render invalid inferences when using predictive machine learning algorithms ([Vansteelandt, 2021](#)). In the concluding section, we provide recent examples of machine learning-based causal inference methods for estimating heterogeneous causal effects.

Sensitivity analysis to unmeasured confounding

Applied researchers intending to conduct causal inference in an observational study should seek to measure and adjust for all relevant confounders toward eliminating biases due to unmeasured confounding. But in most realistic settings, confounders likely remain unadjusted for, raising the prospect of invalid inferences. A common setting is when there are confounders for which it is unfeasible or impractical to observe or collect data. Therefore, it is judicious to conduct a sensitivity analysis addressing unmeasured confounding by judging the (minimum) extent to which substantive conclusions would differ after adjusting for an additional hypothetical unmeasured confounder.

A widely-adopted approach focuses on characterizing the (strength and direction of) association between the hypothetical confounder and the treatment as a sensitivity parameter ([Rosenbaum, 1987b](#)). However, a large confounding bias requires a strong confounder-treatment association and a strong confounder-outcome association (conditional on the treatment). Hence another approach further incorporates the (strength and direction of) association between the hypothetical confounder and the outcome as an additional sensitivity parameter ([Rosenbaum and Rubin, 1983a](#); [Lin et al., 1998](#); [Imbens, 2003](#)). The sensitivity analysis proceeds by (i) specifying different values for the sensitivity parameters, (ii) repeating estimation of the causal effect while adjusting for the hypothetical confounder under each fixed value of the sensitivity parameters in turn, and then (iii) assessing the relative sensitivity of the results across the range of values. Graphical tools, such as bivariate sensitivity contour plots that visualize different combinations of how strongly the hypothetical confounder must be (simultaneously) associated with the exposure and outcome for the conclusions to be affected, can be especially useful ([Carnegie et al., 2016](#); [Ding and VanderWeele, 2016](#); [Dorie et al., 2016](#); [VanderWeele and Ding, 2017](#)). Such methods require users to calibrate the range and granularity of the values of the sensitivity parameters. Alternatively, Bayesian sensitivity analyses for unmeasured confounding ([Greenland, 2003a](#); [McCandless and Gustafson, 2017](#)) avoid such user calibration through the application of Bayes Theorem to obtain a posterior distribution for the sensitivity parameters in the outcome model. We refer readers to [Liu et al. \(2013\)](#) for an overview of other approaches.

Estimating causal effects when unconfoundedness is precluded

It may be unfeasible to record all relevant confounders in practice, thus precluding unconfoundedness from being reasonably assumed to hold. Under such settings, confounding cannot be removed, even with an infinite sample size, and the ATE is non-parametrically (point) unidentifiable. Quasi-experimental designs and natural experiments, such as *regression discontinuity* ([Thistlethwaite and Campbell, 1960](#)), *difference-in-differences* ([Card and Krueger, 1994](#)), and instrumental variable analysis ([Angrist et al., 1996](#)), permit alternative causal inference approaches that can reduce the reliance on the unconfoundedness assumption ([Cook et al., 2002](#); [Angrist and Pischke, 2008](#); [Kim and Steiner, 2016](#)).

However, the validity of these approaches are contingent on their own set of causal assumptions. We briefly describe the assumptions for the instrumental variable approach. Consider the setting depicted in the causal diagram of [Fig. 1i](#) where the causal effect cannot be identified due to unmeasured confounding by U_{ZY} . An instrumental variable such as A can permit causal inference about the effect of Z on Y . A specific setting in the context of randomized experiments is when partial (or non-) compliance arises, so that certain individuals are less likely than others to adhere to the (randomized) treatment. The randomized treatment then plays the role of an instrumental variable for the treatment experienced by (or provided to) each individual; see e.g., [Swanson et al. \(2018\)](#). However, the validity of the instrumental variables approach demands three core assumptions about the causal structure ([Glymour et al., 2012](#)): (i) there is no unmeasured confounding of the instrument and the outcome of interest, (ii) the instrument must be associated with the treatment of interest, and (iii) every directed path from the instrument to the outcome must pass through the treatment (commonly termed “exclusion restriction”). Assumption (i) is depicted by the absence of back-door paths between A and Y ; assumption (ii) is depicted by the arrow from A to Z ; and assumption (iii) is depicted by the absence of a directed path from A to Y . Identification of the causal effect demands a further so-called “monotonicity” assumption that the conditional expectation of (counterfactual) treatment given the instrument satisfies certain distributional constraints. Efforts have thus focused on evaluating these assumptions; see e.g., [Swanson and Hernán \(2013\)](#), [Yang et al. \(2014\)](#), [Pizer \(2016\)](#), [Labrecque and Swanson \(2018\)](#), and [Lousdal \(2018\)](#). We refer readers to [Goetghebeur et al. \(2020\)](#) for an accessible introduction to instrumental variable based methods.

Summary and recent advances

Studying cause-effect relationships to elevate understanding of the behavioral, social, or other mechanisms leading to an outcome is central in education research and practice. While well-designed randomized experiments may be the gold standard for causal inference, they are often unfeasible in practice due to ethical concerns or logistic challenges. In observational studies where participants can select their treatment, spurious correlations between the treatment and the outcome - arising from their mutual association with a confounder that is unadjusted for - can lead to misleading causal conclusions.

We presented the potential outcomes framework for formalizing the causal impacts of interventions, or treatments, on outcomes. We have stated the key assumptions for identifying and estimating average treatment effects, namely the stable unit treatment value assumption (SUTVA), unconfoundedness, and positivity. We have introduced causal diagrams for representing causal relations between variables that are grounded in theory. In particular, causal diagrams are an essential tool to determine which covariates should and should not be adjusted for when seeking to identify the causal effect. We provided examples of simple causal diagrams explaining why certain covariates should or should not be adjusted for. We have reviewed widely used analytic techniques for estimating the treatment effects, such as covariate adjustment, propensity scores, and other matching techniques, as well as combinations of treatment selection and outcome models. However, these estimation methods are contingent on assuming that unconfoundedness can be met, which is not always plausible. Therefore, we have additionally described sensitivity analyses for unmeasured confounding. Natural experiments such as instrumental variable based methods are attractive alternatives for causal inference when unconfoundedness cannot be reasonably assumed to hold. However, they require their own set of causal assumptions for valid inference. Due to the prevalence of multilevel data structures in education, we also presented strategies and statistical tools for conducting causal inference with observational multilevel data.

We state a few recent advances in causal models that may drive education research in exciting directions. Scientific interest often lies in disentangling causal effects transmitted through intermediate variables on the causal path between treatment and outcome, such as M in the causal diagram of Fig. 1d. Using potential outcomes to conceptualize (in)direct effects can be especially fruitful under the *causal mediation analysis* framework. Model-free definitions of *natural direct and indirect effects*, first developed by Robins and Greenland (1992) and Pearl (2001), do not rely on any assumed statistical models, and therefore enable extensions to non-additive and non-linear models (MacKinnon et al., 2020). Furthermore, the framework facilitates formalizing the unconfoundedness (or “sequential ignorability”) assumptions necessary for decomposing the total effect into a direct and an indirect effect (Imai et al., 2010; Pearl, 2014). When there are multiple simultaneous mediators, valid causal inferences about the indirect effects via each mediator can be obtained, even without assuming a causal structure among the mediators, or when the mediators share unmeasured confounders (Loh et al., 2021). We refer readers to Nguyen et al. (2021) for an introduction, and VanderWeele (2015) for a book-length exposition.

Multilevel data structures are common in educational settings where repeated measures, individuals, and educational clusters correspond to different levels (Raudenbush and Bryk, 2002; Snijders and Bosker, 2011; Goldstein, 2011; Hox et al., 2017). Under such nested or clustered data structures, omitted variable bias in estimation is likely to arise due to unmeasured confounding at upper (such as classroom or school) levels (Kim and Frees, 2006; Kim, 2009; Arpino and Mealli, 2011; Li et al., 2013; Fuentes et al., 2021). Because random effects estimation approaches are inadequate to adjust for such unmeasured upper-level confounding, fixed effects estimation approaches are recommended to eliminate such bias (Hazlett and Wainstein, 2020; Imai and Kim, 2020; Kim and Steiner, 2021). However, they may be unreliable, e.g., when cluster sizes are small or some clusters contain few or no (un)treated observations; or undesirable, e.g., when the effects of school, teacher, or student characteristics on the outcome are of interest. Kim and Frees (2007) introduced the generalized method of moments (GMM) estimator as an alternative estimator when the random effects estimator is biased and the fixed effects estimator is impractical. Omitted variable tests for multilevel data and robust GMM estimation are implemented in the R package REndo (Gui et al., 2021) that fits linear models with endogenous regressors.

A component of SUTVA restricts an individual’s potential outcomes to depend only on their experienced treatment. This may be overly restrictive in settings where individuals are causally connected (Raudenbush and Schwartz, 2020). Continuing the example in the introduction, if the impact of private schooling on the academic outcomes of Alex depended on whether their next-door neighbor and closest friend Cameron went to the same school, then this would violate SUTVA. More general definitions of the potential outcomes under (treatment) interference (Cox, 1958) permit assessing peer (influence), or spillover, effects when evaluating the impacts of educational interventions. Analytic strategies and techniques that build on the potential outcomes framework include Sobel (2006), Hong and Raudenbush (2006), Aronow et al. (2021), and Loh and Ren (2020) among many others; see Halloran and Hudgens (2016) for a comprehensive review.

We have focused on the average treatment effect in an entire population of interest. But variation of the effect across different groups of individuals is often expected because individuals vary in their susceptibilities to treatment. Hence interest is often in discerning effect heterogeneity to inform future interventions targeted at specific subgroups (Mayer et al., 2019). Effect heterogeneity across strata of a given pre-treatment modifier V can be assessed by simply computing the average treatment effect of Z on Y within each level of V . Researchers should thus collect data on pre-treatment covariates, even in randomized experiments, to investigate effect heterogeneity. In an observational study, V may be a subset of the confounders L needed for unconfoundedness to hold. The remaining variables in L must then be adjusted for (within each stratum of V) following the above procedures. When either the variables V that induce treatment effect heterogeneity, or the functional forms reflecting how V modify the treatment

effect, are unknown, machine learning algorithms can be exploited to adaptively uncover V from among a large set of variables in L , while permitting flexible functional forms of L (including V) to estimate the conditional ATE (CATE). Recent developments such as *causal forests* (Athey and Wager, 2019), *Bayesian causal forests* (Hahn et al., 2020), *targeted minimum loss estimation* (TMLE) (Van der Laan and Rose, 2011), and *metalearners* (Künzel et al., 2019), employ different machine learning algorithms to estimate the propensity score model, or the outcome model, or both. Comparisons of these cutting edge methods have been made in causal inference data analysis competitions (Carvalho et al., 2019; Dorie et al., 2019).

Acknowledgments

The authors thank Peter Steiner, Dan Bolt, Louise Poppe, and Caroline Keck for their detailed and insightful comments on an earlier version of this manuscript.

References

- Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. *J. Am. Stat. Assoc.* 91 (434), 444–455.
- Angrist, J.D., Pischke, J.-S., 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press.
- Aronow, P.M., Eckles, D., Samii, C., Zonszein, S., 2021. Spillover effects in experimental data. *Advances in Experimental Political Science*, p. 289.
- Arpino, B., Cannas, M., 2016. Propensity score matching with clustered data. an application to the estimation of the impact of caesarean section on the apgar score. *Stat. Med.* 35 (12), 2074–2091.
- Arpino, B., Mealli, F., 2011. The specification of the propensity score in multilevel observational studies. *Comput. Stat. Data Anal.* 55 (4), 1770–1780.
- Athey, S., Wager, S., 2019. Estimating Treatment Effects with Causal Forests: An Application. *Obs. Stud.* 5 (2), 37–51.
- Austin, P.C., 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar. Behav. Res.* 46 (3), 399–424.
- Austin, P.C., Stuart, E.A., 2015. Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* 34 (28), 3661–3679.
- Austin, P.C., Stuart, E.A., 2017. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Stat. Methods Med. Res.* 26 (4), 1654–1670.
- Bang, H., Robins, J.M., 2005. Doubly robust estimation in missing data and causal inference models. *Biometrics* 61 (4), 962–973.
- Bollen, K.A., Pearl, J., 2013. Eight myths about causality and structural equation models. In: *Handbook of Causal Analysis for Social Research*. Springer, pp. 301–328.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
- Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J., Stürmer, T., 2006. Variable selection for propensity score models. *Am. J. Epidemiol.* 163 (12), 1149–1156.
- Card, D., Krueger, A.B., 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *Am. Econ. Rev.* 84 (4), 772.
- Carnegie, N.B., Harada, M., Hill, J.L., 2016. Assessing sensitivity to unmeasured confounding using a simulated potential confounder. *J. Res. Educ. Eff.* 9 (3), 395–420.
- Carvalho, C., Feller, A., Murray, J., Woody, S., Yeager, D., 2019. Assessing treatment effect variation in observational studies: results from a data challenge. *Obs. Stud.* 5 (1), 21–35.
- Cefalu, M., Ridgeway, G., McCaffrey, D., Morral, A., Griffin, B.A., Burgette, L., 2021. *Twang: toolkit for weighting and analysis of nonequivalent groups*. R Package Version 2.0.
- Chattopadhyay, A., Hase, C.H., Zubizarreta, J.R., 2020. Balancing vs modeling approaches to weighting in practice. *Stat. Med.* 39, 3227–3254.
- Chipman, H.A., George, E.I., McCulloch, R.E., et al., 2010. Bart: bayesian additive regression trees. *Ann. Appl. Stat.* 4 (1), 266–298.
- Cochran, W.G., 1968. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 295–313.
- Cole, S.R., Platt, R.W., Schisterman, E.F., Chu, H., Westreich, D., Richardson, D., Poole, C., 2009. Illustrating bias due to conditioning on a collider. *Int. J. Epidemiol.* 39 (2), 417–420.
- Cook, T.D., Campbell, D.T., Shadish, W., 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin Boston, MA.
- Cook, T.D., Steiner, P.M., Pohl, S., 2009. How bias reduction is affected by covariate choice, unreliability, and mode of data analysis: results from two types of within-study comparisons. *Multivariate Behav. Res.* 44 (6), 828–847. PMID: 26801798.
- Cox, D.R., 1958. *Planning of Experiments*. John Wiley & Sons, New York.
- D'Agostino Jr., R.B., 1998. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat. Med.* 17 (19), 2265–2281.
- Diamond, A., Sekhon, J.S., 2013. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Rev. Econ. Stat.* 95 (3), 932–945.
- Ding, P., Miratrix, L.W., 2015. To adjust or not to adjust? sensitivity analysis of m-bias and butterfly-bias. *J. Causal Inference* 3 (1), 41–57.
- Ding, P., VanderWeele, T.J., 2016. Sensitivity analysis without assumptions. *Epidemiology* 27 (3), 368.
- Dong, N., Stuart, E.A., Lenis, D., Nguyen, T.Q., 2020. Using propensity score analysis of survey data to estimate population average treatment effects: a case study comparing different methods. *Eval. Rev.* 44 (1), 84–108.
- Dorie, V., Harada, M., Carnegie, N.B., Hill, J., 2016. A flexible, interpretable framework for assessing sensitivity to unmeasured confounding. *Stat. Med.* 35 (20), 3453–3470.
- Dorie, V., Hill, J., Shalit, U., Scott, M., Cervone, D., 2019. Automated versus do-it-yourself methods for causal inference: lessons learned from a data analysis competition. *Stat. Sci.* 34 (1), 43–68.
- Elwert, F., 2013. *Graphical Causal Models*. Springer Netherlands, Dordrecht, pp. 245–273.
- Foster, E.M., 2010. Causal inference and developmental psychology. *Dev. Psychol.* 46 (6), 1454–1480.
- Fuentes, A., Lütke, O., Robitzsch, A., 2021. Causal inference with multilevel data: a comparison of different propensity score weighting approaches. *Multivariate Behav. Res.* 0 (0), 1–24. PMID: 34128730.
- Glymour, M.M., Tchetgen Tchetgen, E.J., Robins, J.M., 2012. Credible mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am. J. Epidemiol.* 175 (4), 332–339.
- Glynn, A.N., Quinn, K.M., 2010. An introduction to the augmented inverse propensity weighted estimator. *Polit. Anal.* 18 (1), 36–56.
- Goetghebuer, E., le Cessie, S., De Stavola, B., Moodie, E.E., Waernbaum, I., the topic group Causal Inference (TG7) of the STRATOS initiative, 2020. Formulating causal questions and principled statistical answers. *Stat. Med.* 39 (30), 4922–4948.
- Goldstein, H., 2011. *Multilevel Statistical Models*, vol. 922. John Wiley & Sons.

- Greenland, S., 2003a. The impact of prior distributions for uncontrolled confounding and response bias: a case study of the relation of wire codes and magnetic fields to childhood leukemia. *J. Am. Stat. Assoc.* 98 (461), 47–54.
- Greenland, S., 2003b. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 14 (3), 300–306.
- Grosz, M.P., Rohrer, J.M., Thoemmes, F., 2020. The taboo against explicit causal inference in nonexperimental psychology. *Perspect. Psychol. Sci.* 15 (5), 1243–1255. PMID: 32727292.
- Gui, R., Meierer, M., Algesheimer, R., Schilter, P., 2021. REndo: fitting linear models with endogenous regressors using latent instrumental variables. *R Package Version 2.4.2*.
- Hahn, P.R., Murray, J.S., Carvalho, C.M., 2020. Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Anal.* 15 (3), 965–1056.
- Halloran, M.E., Hudgens, M.G., 2016. Dependent happenings: a recent methodological review. *Curr. Epidemiol. Rep.* 3 (4), 297–305.
- Hansen, B.B., 2004. Full matching in an observational study of coaching for the sat. *J. Am. Stat. Assoc.* 99 (467), 609–618.
- Harder, V.S., Stuart, E.A., Anthony, J.C., 2010. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol. Methods* 15 (3), 234–249.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media.
- Hayduk, L., Cummings, G., Stratkotter, R., Nimmo, M., Grygoryev, K., Dosman, D., Gillespie, M., Pazderka-Robinson, H., Boadu, K., 2003. Pearl's D-separation: one more step into causal thinking. *Struct. Equ. Model.* 10 (2), 289–311.
- Hazlett, C., Wainstein, L., 2020. Understanding, choosing, and unifying multilevel and fixed effect approaches. *Polit. Anal.* 1–20.
- Hernán, M.A., Robins, J.M., 2020. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC.
- Ho, D., Imai, K., King, G., Stuart, E.A., 2011. Matchit: nonparametric preprocessing for parametric causal inference. *J. Stat. Software* 42 (8), 1–28.
- Holland, P.W., 1986. Statistics and causal inference. *J. Am. Stat. Assoc.* 81 (396), 945–960.
- Hong, G., Raudenbush, S.W., 2006. Evaluating kindergarten retention policy: a case study of causal inference for multilevel observational data. *J. Am. Stat. Assoc.* 101 (475), 901–910.
- Hox, J.J., Moerbeek, M., Van de Schoot, R., 2017. *Multilevel Analysis: Techniques and Applications*. Routledge.
- Huang, P.-H., 2020. Postselection inference in structural equation modeling. *Multivariate Behav. Res.* 55 (3), 344–360. PMID: 31304796.
- Iacus, S., King, G., Porro, G., 2009. CEM: software for coarsened exact matching. *J. Stat. Software* 30 (9), 1–27.
- Imai, K., Keele, L., Tingley, D., 2010. A general approach to causal mediation analysis. *Psychol. Methods* 15 (4), 309.
- Imai, K., Kim, I.S., 2020. On the use of two-way fixed effects regression models for causal inference with panel data. *Polit. Anal.* 1–11.
- Imai, K., Ratkovic, M., 2014. Covariate balancing propensity score. *J. Roy. Stat. Soc. B* 76 (1), 243–263.
- Imbens, G.W., 2003. Sensitivity to exogeneity assumptions in program evaluation. *Am. Econ. Rev.* 93 (2), 126–132.
- Imbens, G.W., Rubin, D.B., 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.
- Kainz, K., Greifer, N., Givens, A., Swietek, K., Lombardi, B.M., Zietz, S., Kohn, J.L., 2017. Improving causal inference: recommendations for covariate selection and balance in propensity score methods. *J. Soc. Soc. Work. Res.* 8 (2), 279–303.
- Kang, J.D., Schafer, J.L., 2007. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. *Stat. Sci.* 22 (4), 523–539.
- Keller, B., Tipton, E., 2016. Propensity score analysis in R: a software review. *J. Educ. Behav. Stat.* 41 (3), 326–348.
- Kim, G.-S., Paik, M.C., Kim, H., 2017. Causal inference with observational data under cluster-specific non-ignorable assignment mechanism. *Comput. Stat. Data Anal.* 113, 88–99.
- Kim, J., Seltzer, M., 2007. *Causal Inference in Multilevel Settings in Which Selection Processes Vary across Schools*. cse technical report 708. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Kim, J.-S., 2009. Multilevel analysis: an overview and some contemporary issues. In: *The SAGE Handbook of Quantitative Methods in Psychology*. SAGE, London, pp. 337–361.
- Kim, J.-S., Frees, E.W., 2006. Omitted variables in multilevel models. *Psychometrika* 71 (4), 659.
- Kim, J.-S., Frees, E.W., 2007. Multilevel modeling with correlated effects. *Psychometrika* 72 (4), 505–533.
- Kim, J.-S., Steiner, P.M., 2015. Multilevel propensity score methods for estimating causal effects: a latent class modeling strategy. In: *Quantitative Psychology Research*. Springer, pp. 293–306.
- Kim, J.-S., Steiner, P.M., Lim, W.-C., 2016. Mixture modeling methods for causal inference with multilevel data. *Advances in Multilevel Modeling for Educational Research: Addressing Practical Issues Found in Real-World Applications*, pp. 335–359.
- Kim, Y., Steiner, P.M., 2016. Quasi-experimental designs for causal inference. *Educ. Psychol.* 51 (3–4), 395–405. PMID: 30100637.
- Kim, Y., Steiner, P.M., 2021. Causal graphical views of fixed effects and random effects models. *Br. J. Math. Stat. Psychol.* 74, 165–183.
- King, G., Zeng, L., 2001. Improving forecasts of state failure. *World Polit.* 53 (4), 623–658.
- Künzel, S.R., Sekhon, J.S., Bickel, P.J., Yu, B., 2019. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proc. Natl. Acad. Sci. Unit. States Am.* 116 (10), 4156–4165.
- Labrecque, J., Swanson, S.A., 2018. Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools. *Curr. Epidemiol. Rep.* 5 (3), 214–220.
- Lanza, S.T., Moore, J.E., Butera, N.M., 2013. Drawing causal inferences using propensity scores: a practical guide for community psychologists. *Am. J. Community Psychol.* 52 (3–4), 380–392.
- Lee, B.K., Lessler, J., Stuart, E.A., 2010. Improving propensity score weighting using machine learning. *Stat. Med.* 29 (3), 337–346.
- Li, F., Zaslavsky, A.M., Landrum, M.B., 2013. Propensity score weighting with multilevel data. *Stat. Med.* 32 (19), 3373–3387.
- Lin, D.Y., Psaty, B.M., Kronmal, R.A., 1998. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 54 (3), 948–963.
- Little, R., An, H., 2004. Robust likelihood-based analysis of multivariate data with missing values. *Stat. Sin.* 949–968.
- Liu, W., Kuramoto, S.J., Stuart, E.A., 2013. An introduction to sensitivity analysis for unobserved confounding in nonexperimental prevention research. *Prev. Sci.* 14 (6), 570–580.
- Loh, W.W., Moerkerke, B., Loeys, T., Vansteelandt, S., 2021. Disentangling indirect effects through multiple mediators without assuming any causal structure among the mediators. *Psychol. Methods*. <https://doi.org/10.1037/met0000314>. Advance online publication.
- Loh, W.W., Ren, D., 2020. Estimating social influence in a social network using potential outcomes. *Psychol. Methods*. <https://doi.org/10.1037/met0000356>. Advance online publication.
- Lousdal, M.L., 2018. An introduction to instrumental variable assumptions, validation and estimation. *Emerg. Themes Epidemiol.* 15 (1), 1.
- Luellen, J.K., Shadish, W.R., Clark, M., 2005. Propensity scores: an introduction and experimental test. *Eval. Rev.* 29 (6), 530–558.
- Lunceford, J.K., Davidian, M., 2004. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Stat. Med.* 23 (19), 2937–2960.
- MacKinnon, D.P., Valente, M.J., Gonzalez, O., 2020. The correspondence between causal and traditional mediation analysis: the link is the mediator by treatment interaction. *Prev. Sci.* 21 (2), 147–157.

- Mayer, A., Zimmermann, J., Hoyer, J., Salzer, S., Wiltink, J., Leibing, E., Leichsenring, F., 2019. Interindividual differences in treatment effects based on structural equation models with latent variables: an EffectLiteR tutorial. *Struct. Equ. Model.* 1–19.
- McCaffrey, D.F., Ridgeway, G., Morral, A.R., 2004. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol. Methods* 9 (4), 403.
- McCandless, L.C., Gustafson, P., 2017. A comparison of bayesian and Monte Carlo sensitivity analysis for unmeasured confounding. *Stat. Med.* 36 (18), 2887–2901.
- Morgan, S.L., Winship, C., 2015. *Counterfactuals and Causal Inference*. Cambridge University Press.
- Nguyen, T.Q., Schmid, I., Stuart, E.A., 2021. Clarifying causal mediation analysis for the applied researcher: defining effects based on what we want to learn. *Psychol. Methods* 26 (2), 255–271.
- Pearl, J., 1998. Graphs, causality, and structural equation models. *Socio. Methods Res.* 27 (2), 226–284.
- Pearl, J., 2001. Direct and indirect effects. In: *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, pp. 411–420.
- Pearl, J., 2009. *Causality: Models, Reasoning and Inference*, second ed. Cambridge University Press, New York, NY, USA.
- Pearl, J., 2010. On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology* 21 (6), 872.
- Pearl, J., 2012. *The Causal Foundations of Structural Equation Modeling*. Guilford Press, New York, NY, USA, pp. 68–91.
- Pearl, J., 2013. Linear models: a useful “microscope” for causal analysis. *J. Causal Inference* 1 (1), 155–170.
- Pearl, J., 2014. Interpretation and identification of causal mediation. *Psychol. Methods* 19 (4), 459.
- Pearl, J., MacKenzie, D., 2018. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, NY, USA.
- Petersen, M.L., Porter, K.E., Gruber, S., Wang, Y., van Der Laan, M.J., 2012. Diagnosing and responding to violations in the positivity assumption. *Stat. Methods Med. Res.* 21 (1), 31–54.
- Pizer, S.D., 2016. Falsification testing of instrumental variables methods for comparative effectiveness research. *Health Serv. Res.* 51 (2), 790–811.
- Powell, M.G., Hull, D.M., Beaujean, A.A., 2020. Propensity score matching for education data: worked examples. *J. Exp. Educ.* 88 (1), 145–164.
- R Core Team, 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raudenbush, S.W., Bryk, A.S., 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*, vol. 1. sage.
- Raudenbush, S.W., Schwartz, D., 2020. Randomized experiments in education, with implications for multilevel causal inference. *Annu. Rev. Stat. Appl.* 7, 177–208.
- Robins, J.M., 2000. Robust estimation in sequentially ignorable missing data and causal inference models. In: *Proceedings of the American Statistical Association*, vol. 1999, pp. 6–10. Indianapolis, IN.
- Robins, J.M., Greenland, S., 1992. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 3 (2), 143–155.
- Robins, J.M., Hernan, M.A., Brumback, B., 2000. Marginal Structural Models and Causal Inference in Epidemiology.
- Robins, J.M., Rotnitzky, A., Zhao, L.P., 1994. Estimation of regression coefficients when some regressors are not always observed. *J. Am. Stat. Assoc.* 89 (427), 846–866.
- Rohrer, J.M., 2018. Thinking clearly about correlations and causation: graphical causal models for observational data. *Adv. Meth. Pract. Psychol. Sci.* 1 (1), 27–42.
- Rosenbaum, P.R., 1987a. Model-based direct adjustment. *J. Am. Stat. Assoc.* 82 (398), 387–394.
- Rosenbaum, P.R., 1987b. Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika* 74, 13–26.
- Rosenbaum, P.R., 2002. *Observational Studies*. Springer, New York, New York.
- Rosenbaum, P.R., Rubin, D.B., 1983a. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *J. Roy. Stat. Soc. B* 45 (2), 212–218.
- Rosenbaum, P.R., Rubin, D.B., 1983b. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70 (1), 41–55.
- Rubin, D.B., 1986. Comment: which ifs have causal answers. *J. Am. Stat. Assoc.* 81 (396), 961–962.
- Rubin, D.B., 1990. [On the application of probability theory to agricultural experiments. Essay on principles. Section 9.] comment: neyman (1923) and causal inference in experiments and observational studies. *Stat. Sci.* 5 (4), 472–480.
- Rubin, D.B., Stuart, E.A., Zanotto, E.L., 2004. A potential outcomes view of value-added assessment in education. *J. Educ. Behav. Stat.* 29 (1), 103–116.
- Sekhon, J.S., 2011. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *J. Stat. Software* 42 (7), 1–52.
- Setodji, C.M., McCaffrey, D.F., Burgette, L.F., Almirall, D., Griffin, B.A., 2017. The right tool for the job: choosing between covariate-balancing and generalized boosted model propensity scores. *Epidemiology* 28 (6), 802–811.
- Setoguchi, S., Schneeweiss, S., Brookhart, M.A., Glynn, R.J., Cook, E.F., 2008. Evaluating uses of data mining techniques in propensity score estimation: a simulation study. *Pharmacoepidemiol. Drug Saf.* 17 (6), 546–555.
- Shadish, W.R., Steiner, P.M., 2010. A primer on propensity score analysis. *N.born Infant Nurs. Rev.* 10 (1), 19–26 (Quantitative Research Methodology).
- Snijders, T.A., Bosker, R.J., 2011. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage.
- Sobel, M.E., 2006. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *J. Am. Stat. Assoc.* 101 (476), 1398–1407.
- Splawa-Neyman, J., Dabrowska, D.M., Speed, T.P., 1990. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Stat. Sci.* 5 (4), 465–472.
- Steiner, P.M., Kim, J.-S., Thoemmes, F., 2012. Matching strategies for observational multilevel data. In: *Joint Statistical Meeting Proceedings, Social Statistics Section*. American Statistical Association, pp. 5020–5032.
- Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. *Stat. Sci.* 25 (1), 1–21.
- Suk, Y., Kang, H., Kim, J.-S., 2020. Random forests approach for causal inference with clustered observational data. *Multivariate Behav. Res.* 56 (6), 829–852.
- Suk, Y., Kim, J.-S., Kang, H., 2021. Hybridizing machine learning methods and finite mixture models for estimating heterogeneous treatment effects in latent classes. *J. Educ. Behav. Stat.* 46 (3), 323–347.
- Swanson, S.A., Hernán, M.A., 2013. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology* 24 (3), 370–374.
- Swanson, S.A., Hernán, M.A., Miller, M., Robins, J.M., Richardson, T.S., 2018. Partial identification of the average treatment effect using instrumental variables: review of methods for binary instruments, treatments, and outcomes. *J. Am. Stat. Assoc.* 113 (522), 933–947. PMID: 31537952.
- Tan, Z., 2006. A distributional approach for causal inference using propensity scores. *J. Am. Stat. Assoc.* 101 (476), 1619–1637.
- Textor, J., van der Zander, B., Gilthorpe, M.S., Liśkiewicz, M., Ellison, G.T., 2017. Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *Int. J. Epidemiol.* 45 (6), 1887–1894.
- Thistlethwaite, D.L., Campbell, D.T., 1960. Regression-discontinuity analysis: an alternative to the ex post facto experiment. *J. Educ. Psychol.* 51 (6), 309–317.
- Thoemmes, F.J., 2015. M-bias, butterfly bias, and butterfly bias with correlated causes – a comment on Ding and Miratrix (2015). *J. Causal Inference* 3 (2), 253–258.
- Thoemmes, F.J., Kim, E.S., 2011. A systematic review of propensity score methods in the social sciences. *Multivariate Behav. Res.* 46 (1), 90–118. PMID: 26771582.
- Thoemmes, F.J., West, S.G., 2011. The use of propensity scores for nonrandomized designs with clustered data. *Multivariate Behav. Res.* 46 (3), 514–543.
- Van der Laan, M.J., Rose, S., 2011. *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer Science & Business Media.
- VanderWeele, T., 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*, first ed. Oxford University Press.
- VanderWeele, T.J., 2019. Principles of confounder selection. *Eur. J. Epidemiol.* 34 (3), 211–219.
- VanderWeele, T.J., Ding, P., 2017. Sensitivity analysis in observational research: introducing the E-value. *Ann. Intern. Med.* 167, 268–274.
- Vansteelandt, S., 2021. Statistical modelling in the age of data science. *Obs. Stud.* 7 (1), 217–228.
- Vansteelandt, S., Daniel, R.M., 2014. On regression adjustment for the propensity score. *Stat. Med.* 33 (23), 4053–4072.

- Webster-Clark, M., Stürmer, T., Wang, T., Man, K., Marinac-Dabic, D., Rothman, K.J., Ellis, A.R., Gokhale, M., Lunt, M., Girman, C., Glynn, R.J., 2021. Using propensity scores to estimate effects of treatment initiation decisions: state of the science. *Stat. Med.* 40 (7), 1718–1735.
- West, S.G., Cham, H., Thoemmes, F., Renneberg, B., Schulze, J., Weiler, M., 2014. Propensity scores as a basis for equating groups: basic principles and application in clinical treatment outcome research. *J. Consult. Clin. Psychol.* 82 (5), 906–919.
- Witte, J., Didelez, V., 2019. Covariate selection strategies for causal inference: classification and comparison. *Biom. J.* 61 (5), 1270–1289.
- Wooldridge, J.M., 2016. Should instrumental variables be used as matching variables? *Res. Econ.* 70 (2), 232–237.
- Yang, F., Zubizarreta, J.R., Small, D.S., Lorch, S., Rosenbaum, P.R., 2014. Dissonant conclusions when testing the validity of an instrumental variable. *Am. Statistician* 68 (4), 253–263.
- Yang, S., 2018. Propensity score weighting for causal inference with clustered data. *J. Causal Inference* 6 (2), 20170027.
- Zhong, Y., Kennedy, E.H., Bodnar, L.M., Naimi, A.I., 2021. AIPW: an R package for augmented inverse probability weighted estimation of average causal effects. *Am. J. Epidemiol.* 190 (12), 2690–2699.
- Zubizarreta, J.R., Keele, L., 2017. Optimal multilevel matching in clustered observational studies: a case study of the effectiveness of private schools under a large-scale voucher system. *J. Am. Stat. Assoc.* 112 (518), 547–560.

Advanced machine learning

Oren E. Livne, Educational Testing Service, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	684
Types of machine learning	685
Supervised learning	685
Unsupervised learning	685
Other types of learning	686
Deep learning	686
The artificial neuron	686
Basic fully-connected architecture	687
Model training	688
Convolutional Neural Networks	689
Software frameworks	689
Applications to educational measurement	690
Automatic feature extraction: off-topic spoken response detection	690
Time-series modeling: automated essay scoring	691
Recurrent neural networks	691
Automated essay scoring	692
Transfer learning	692
Generative adversarial networks and content generation	692
Summary	693
Future directions	693
References	693

Glossary

Artificial Intelligence (AI) the simulation of human intelligence processes by machines, especially computer systems. Specific applications include expert systems, natural language processing, speech recognition and machine vision

Artificial Neural Network a collection of connected units called artificial neurons, which loosely models a biological brain and used to learn data representations

Convolutional Neural Network a class of artificial neural network, most commonly applied to analyze visual imagery and contains convolutional layers

Deep Learning a family of machine learning methods based on artificial neural networks that contain many hidden layers

Machine Learning the use and development of computer systems that learn from examples and automatically draw inferences and detect patterns in data

Recurrent Neural Network a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence, used for language translation, natural language processing and speech recognition

Stochastic Gradient Descent (SGD) an iterative algorithm for optimizing an objective function, used for training neural networks

Supervised Learning the machine learning task of learning to predict the label of an input example

Transfer Learning a machine learning method where a model developed for a task is reused as the starting point for a model on a second task

Reinforcement Learning the machine learning task of operating in an environment using feedback, without a fixed training data set.

Unsupervised Learning the machine learning task of learning patterns and relationships in data from unlabeled examples

Introduction

Machine learning (ML) is the use and development of computer systems that can *learn through experience*, i.e., gradually improve their performance on a task, by using algorithms and statistical models that draw inferences and detect patterns in data (Wikipedia, 2021a,b,c). This is accomplished by automatically adapting data structures and parameters to optimize an objective function, as

more data are shown to the system. ML is considered a subset of Artificial Intelligence (AI), and has become the dominant data-driven approach in a great number of fields where an enormous amount of data are available, but it is difficult or unfeasible to develop conventional algorithms to perform the task. These include image recognition for medical diagnosis (Varela-Santos and Melin, 2021), speech recognition (Hussain et al., 2022), automatic language translation (Lakew et al., 2018), stock market trading (Sonare et al., 2021), and online fraud detection and e-mail spam filtering (Fu et al., 2016). From “Google Search by Voice” to self-driving cars, ML is all around us nowadays.

Educational measurement is undergoing a similar transformation. Digitally Based Assessments (DBAs) enable the capture of a broad range of test takers’ response process information at finer time granularity than traditional psychometrics, in areas of problem solving (Hao et al., 2017; Polyak et al., 2017), communication and collaboration (von Davier et al., 2017), via chats and games (Wu et al., 2016; Kim and Shute, 2015). These rich process data can provide new opportunities for validating assessments, improving measurement precision, examining response patterns, and identifying new constructs (Neiriz et al., 2017; Mislevy et al., 2014). But these potential benefits do not come for free. The significantly increased volume, velocity, and variety of data pose new challenges to psychometricians for handling, analyzing, and interpreting the data to materialize their value. ML is a critical component of the analysis, combined with natural language processing (Clark et al., 2010), educational data mining, and computer vision (Huang and Khan, 2021).

Types of machine learning

Supervised learning

Supervised learning describes a class of problem that involves using a model to learn a mapping between input examples and a target variable (Brownlee, 2019). Supervised learning problems are either classification problems (predicting a categorical class label) or regression problems (predicting a numerical label). For instance, the input example could be a student’s essay (say, encoded into a vector of words), and the label would be a human-graded score (either categorical or continuous).

In a typical pipeline (cf. Fig. 1), input examples are annotated and divided into a training, validation, and test sets (also called “folds”; a typical splitting is 70%-15%-15%). The machine learning model of choice, e.g., decision trees (Topirceanu and Grossecck, 2017) or Support Vector Machines (SVM) (Cristianini and Ricci, 2008) is first trained on the training set, and its prediction accuracy is evaluated on the validation set. This is repeated for multiple values of global model hyper-parameters (e.g., the soft margin value γ in SVM, which controls the margin between different classes and the number of allowed classification mistakes), and the optimal hyper-parameters values that approximately maximize validation accuracy are chosen. Finally, the test set accuracy is measured and reported. The test set must not be used during the training process, otherwise the prediction accuracy may be inflated due to overfitting.

Unsupervised learning

Unsupervised learning refers to a class of problems that involves using a model to describe or extract relationships in data. As opposed to supervised learning, the input data is unlabeled. This does not require data annotation, which is typically a tedious and costly task (e.g., scoring thousands of essays by hand); on the other hand, unsupervised learning does not have a teacher correcting the model (Goodfellow et al., 2016).

The most common unsupervised learning task is **clustering**: finding groups in data. For instance, an automated essay evaluation system might gradually develop a concept of “good essays” and “bad essays” without ever being given scored essays. Examples of clustering algorithms are k-Means, mean-shift clustering, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), expectation-maximization clustering using Gaussian mixture models, and agglomerative hierarchical clustering (Seif, 2018). Clustering quality is measured by the *silhouette score*, a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation) (Rousseeuw, 1987). If the data is high-dimensional, it is typically first *projected* into a lower-dimensional space representation before clustering (Fig. 2).

Another unsupervised learning task is *density estimation*: summarizing the distribution of data. A popular solution method is kernel density estimation (Gramacki, 2017).

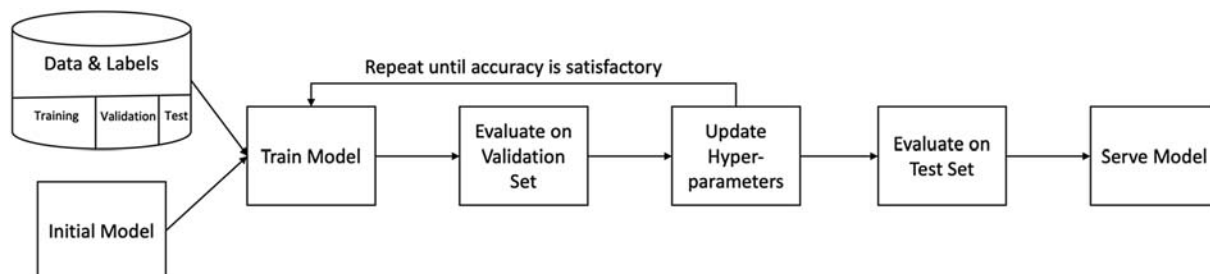


Fig. 1 Supervised machine learning model life cycle. Adapted from Miao et al. (2017).

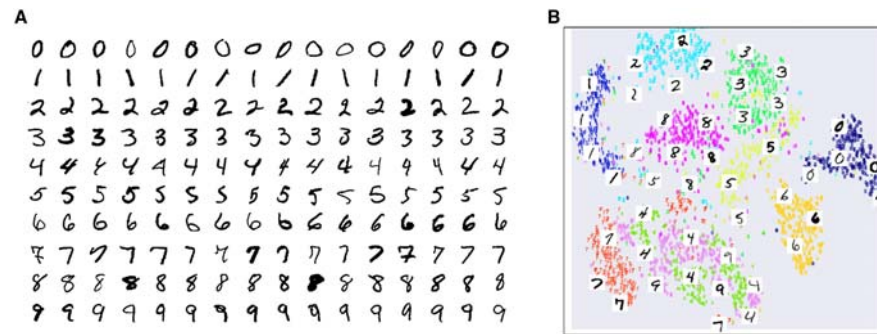


Fig. 2 The MNIST database (Wikipedia, 2021a,b,c) is a large database of handwritten digits used as a common machine learning benchmark. (A) Supervised learning. Each input example is a 28×28 pixel black-and-white image labeled as 0, 1, ..., or 9. Both the images and labels are used during training. New unlabeled images can then be fed into the trained model for prediction. (B) Unsupervised learning example of projecting MNIST images into a two-dimensional plane and clustering them into 10 groups. (A) Reproduced from Wikipedia (2021a,b,c); (B) Reproduced from Silbernagel (2019).

Other types of learning

Reinforcement learning describes a class of problems where an agent operates in an environment and must learn to operate using feedback (Sutton and Barto, 2018). For instance, a collaborative learning system may include a “bot” that can interact with learners (offering hints or assisting in problem solving) based on their inputs. Here there is no fixed training data set, rather “a goal or set of goals that an agent is required to achieve, actions they may perform, and feedback about performance toward the goal” (Brownlee, 2019).

Other types of learning include *semi-supervised learning* (supervised learning where the training data contains very few labeled examples and a large number of unlabeled examples), *multi-instance learning* (individual examples are unlabeled, but groups of examples are), and *self-supervised learning* (unsupervised learning problem framed as a supervised learning problem for which supervised learning algorithms can be applied) (Brownlee, 2019).

Deep learning

Deep learning is a family of machine learning methods for extracting patterns from data using *Artificial Neural Networks (ANNs)*. Deep learning can be supervised, semi-supervised or unsupervised. DNN frameworks can model complex non-linear relationships. Early ANNs date to the 1960s (Rosenblatt, 1962), but their application was limited by data availability and computing power. Renewed interest started in the late 1990s when DNNs were applied to reading over 10% of all the checks in the United States (LeCun et al., 2010). Following breakthrough work (Krizhevsky et al., 2012) in 2012, where DNNs trained on massively labeled data using Graphical Processing Units (GPUs) halved the error rate of state-of-the-art computer vision systems on the ImageNet object recognition challenge, deep learning has emerged as a hot research field, with thousands of publications every year, and prestigious conferences including the Conference on Computer Vision and Pattern Recognition and International Conference on Computer Vision.

DNNs have been successful in many fields, outperforming traditional ML methodologies and sometimes surpassing human expert performance. Application areas include non-linear system control (robotics, self-driving car, process control) (Grigorescu et al., 2020), game-playing and decision making (backgammon, chess, racing, and Go (Silver et al., 2016); cf. Fig. 3), pattern recognition (radar systems, face identification, object, speech and hand-writing recognition) (LeCun et al., 2010), natural language processing (language translation, language generation, and time-series analysis) (Devlin et al., 2019), medical diagnosis (Amato et al., 2013), financial applications (check recognition, fraud detection) (Fu et al., 2016), visualization, Email spam filtering (Dada et al., 2019).

DNNs are attractive because they perform well without external human-designed structures or time-intensive feature engineering. Despite their advantage, their use in educational applications has been limited (Huang and Khan, 2021). Section **Applications to educational measurement** showcases illustrative examples of many recent papers, and there is much room for the development of deep neural networks for educational data classification with complex standards and taxonomies, time-series modeling, recommendation systems that identify learners’ individual styles, and new content generation.

The artificial neuron

A biological neurons receives real-valued input $\mathbf{x} = (x_1, \dots, x_m)$ from other neurons (hundreds in small organisms to 10^4 – 10^5 in the human brain (Herculano-Houzel, 2009)) via synapses (Fig. 4A). Inputs are approximately summed, and if the value exceeds



Fig. 3 AlphaGo versus Lee Sedol, also known as the Google DeepMind Challenge Match, was a five-game Go match between 18-time world champion Lee Sedol and AlphaGo, a deep-learning Go program developed by Google DeepMind, played in Seoul, South Korea between 9 and March 15, 2016. AlphaGo won 4-1. The match has been compared with the historic chess match between Deep Blue and Garry Kasparov in 1997, where computers surpassed the best humans in the world in an advanced cognitive skill. Reproduced from [Armasu \(2016\)](#).

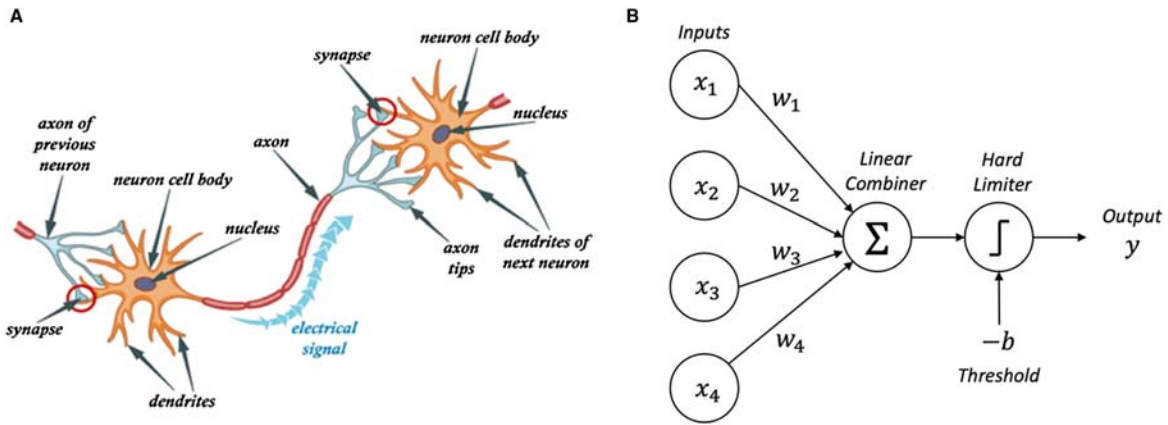


Fig. 4 (A) A biological neuron connected to input neurons. (B) The artificial neuron (perceptron), a nonlinear gate mapping $m = 4$ inputs to an output. (A) Reproduced from [Shen \(2013\)](#).

a threshold the neuron sends an electrical spike to the next neurons. All neurons send signals simultaneously, enabling massively parallel processing.

The basic building block of an artificial neural network is a *perceptron* ([Rosenblatt, 1958](#)). It is a binary classifier, i.e., a function f that maps a real-valued input vector $\mathbf{x} = (x_1, \dots, x_m)$ to a scalar binary output value:

$$f(\mathbf{x}; \boldsymbol{\theta}) := \begin{cases} 1, & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where $\mathbf{w} \cdot \mathbf{x} = \sum_{i=1}^m w_i x_i$, and $\boldsymbol{\theta} = (\mathbf{w}, b)$ are *learned parameters* ([Fig. 4B](#)).

Basic fully-connected architecture

The goal of a neural network is to approximate a function f^* , for instance, $y = f^*(x)$ mapping the input x to the class y . The network defines a mapping $f(\mathbf{x}; \boldsymbol{\theta})$ and learns the value of the parameter vector $\boldsymbol{\theta}$ that results in the best function approximation. In a *feed-forward* network data flows from the input layer to the output layer without looping back. A network with feedback connection is called a *recurrent neural network* (cf. [Sec. Recurrent neural networks](#)) [19, Chap. 6].

An l -layer neural network (also known as a “multi-layer perceptron network” ([Rosenblatt, 1958](#))) is a mapping f of an input vector $\mathbf{x}$ to an output $y = f(\mathbf{x})$ (typically a scalar), comprising of layers $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(l)}$, where the first layer is the input $\mathbf{x}^{(1)} = \mathbf{x}$, followed by $l - 2$ hidden layers, and the output layer $\mathbf{x}^{(l)} = y$. Each layer is a vector of *neurons*, also called *units*. A neuron is a perceptron whose hard-thresholding is replaced by a differentiable *non-linear activation function* σ (such as tanh, sigmoid or rectified linear).

In a *feed-forward* network data flows from the input layer to the output layer without looping back. Each neuron is connected to some of the previous layer’s neurons. In this section we are discussing a *fully-connected* layer, in which it is connected to *all* of them

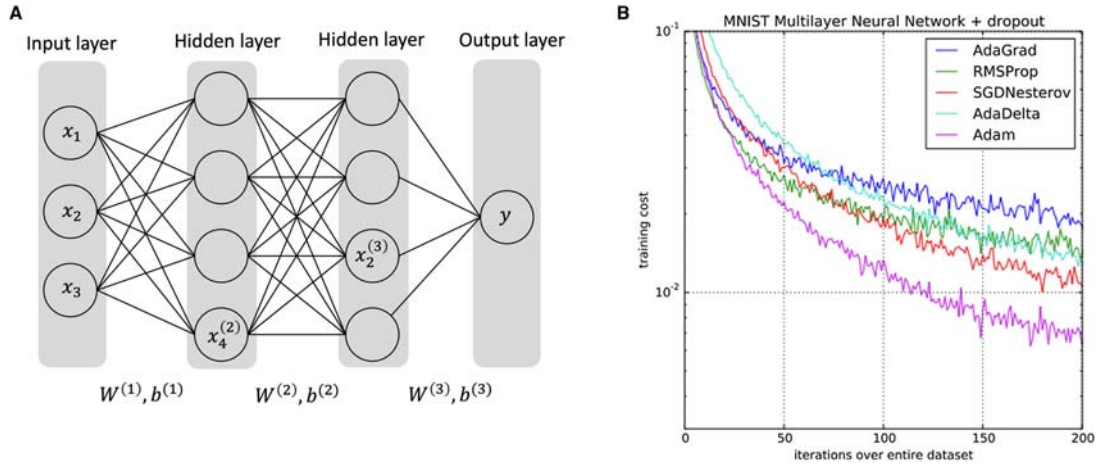


Fig. 5 (A) A four-layer fully-connected neural network. Each link contains a learned parameter, e.g., $w_{4,3}^{(2)}$ connects $x_3^{(2)}$ to $x_4^{(3)}$. Each neuron (except the input layer) has a learned bias parameter. (B) Comparison of training algorithms (optimizers) for the MNIST hand-written digit recognition task. Reproduced from Diederik (2017).

(as opposed to convolutional layers, for instance, which are sparse; cf. Sec. **Convolutional Neural Networks**). Layer $k + 1$ is calculated from layer k by

$$x^{(k+1)} = \sigma(W^{(k)}x^{(k)} + b^{(k)}), \quad (2)$$

where $W^{(k)}$ is a weight matrix, $b^{(k)}$ is a bias vector, and σ is applied element-wise. The last layer may employ a different function g to combine $x^{(L-1)}$ into the final output (e.g., softmax (Wikipedia, 2021a,b,c)). For instance, a four-layer network (Fig. 5A) defines the mapping

$$y = f(x; \theta) = g(W^{(3)}\sigma(W^{(2)}\sigma(W^{(1)}x + b^{(1)}) + b^{(2)}) + b^{(3)}), \quad (3)$$

where θ denotes the vector of learned parameters, consisting of all the elements of $W^{(1)}, b^{(1)}, \dots, W^{(L-1)}, b^{(L-1)}$.

Model training

During training, the parameters θ are *learned* by minimizing a *loss function* $J(\theta)$, which is a performance measure (e.g., prediction accuracy) evaluated over the entire training set. For instance, an L_2 loss is often used for a regression problem:

$$J(\theta) = \sum_i \|y_i - f(x_i; \theta)\|_2^2 \quad (4)$$

where x_i is the i th training instance, y_i the corresponding label, and f is the network's feed-forward mapping.

Starting from an initial guess θ , Stochastic Gradient Descent (SGD) iterations are applied to gradually reduce J using the update

$$\theta \leftarrow \theta - \eta \nabla J(\theta), \quad (5)$$

where the step size η is the *learning rate*. The gradient ∇J is calculated using the *back-propagation algorithm* (Jing et al., 2012). All standard software frameworks provide an auto-differentiation module, as long as J is differentiable almost everywhere. Since the training set is large, SGD reduces gradient computation time by calculating ∇J only approximately on a *mini-batch* of training examples. Thus loss reduction is not monotonic and contains stochastic noise (Fig. 5B) [19, Chap. 8]. SGD converges to a local minimum. Convergence may be slow and require up to hundreds of *epochs* (passes over the training set) [19, Chap. 8]. Batch computation increases efficiency via GPU parallelism; the mini-batch size is chosen based on a trade-off between model performance and GPU memory capacity (Keskar et al., 2017).

Training and optimizing a neural network on a data set is a somewhat of an art. The practitioner can experiment with, vary, and substitute different elements to improve training time and the final model accuracy:

- Varying the number of layers, neurons (units) per layer, and the neuron connections, all of which comprise the network's *architecture*. Increasing the model capacity (number of learned parameters) is a double-edge sword: it can boost accuracy, but is also more prone to over-fitting.
- In particular, introducing *residual connections* or *skip connections*, connecting two layers that are farther apart in the network (i.e., non-consecutive). These are used to allow gradients to flow through a network directly without passing through non-linear activation functions, to mitigate training instability or slowness due to exploding or vanishing gradients. ResNet (He et al., 2015) is a popular architecture that demonstrates this training speed-up in pattern recognition applications.

Table 1 Some of the top neural network MNIST hand-written digit classification results, which surpass human performance.

Network	Prediction accuracy	Learned parameters
Branching/merging CNN with capsules (Byerly et al., 2021)	99.87	1,514,187
Efficient-CapsNet (Mazzia et al., 2021)	99.84	161,824
SOPCNN single model (Assiri, 2020)	99.83	1,400,000
Human (LeCun et al., 1995)	≈ 99.8	—

More results are available at [Anonymous \(2021\)](#).

- Training using different optimization algorithms (SGD; momentum-based methods, such as Adam ([Diederik, 2017](#))) and varying the learning rate reduction schedule to speed-up convergence.
- Adding data augmentation, drop-out layers, weight decay and loss regularization terms to avoid over-fitting in case of very large networks ([Goodfellow et al., 2016](#)).

As training on large benchmarks (e.g., ImageNet) may require weeks and/or many GPUs, many pre-trained models are available for download and use (e.g., [Koh, 2021](#)) ([Table 1](#)).

Convolutional Neural Networks

A Convolutional Neural Networks (CNN) is a specialized network for processing data with a known grid topology, for instance, time-series data (1D grid of samples) or image data (2D grid of pixels $\times$ #channels) [19, Chap. 9]. With the new reality shaped by COVID-19, remote testing and proctoring are taking center stage. Image processing tasks such as tracking a test subject's eye movements, or even merely classify when they appear on camera or not, are becoming crucial, and CNNs are a powerful tool to perform such tasks.

CNNs are frequently used in image and pattern recognition especially because they exploit the geometry to reduce the computational complexity of fully-connected DNNs. For instance, a fully-connected layer for a small image of size 100×100 has 10,000 weights *per neuron* in the second layer. In contrast, CNN layers consist of a set of learned *convolution kernels*, which have a small receptive field but extend through the full depth of the input volume [19, Chap. 9]. As each neuron is connected to a small image region (e.g., 3×3) of input neurons, layers represent features of increasingly larger scales ([Fig. 6](#)).

Software frameworks

Python has emerged as the dominant programming language for machine learning and data science. Notwithstanding, there are plenty of alternatives, from the traditional R, C/C++, Java, to the newer Scala and Julia ([Gossett, 2021](#)).

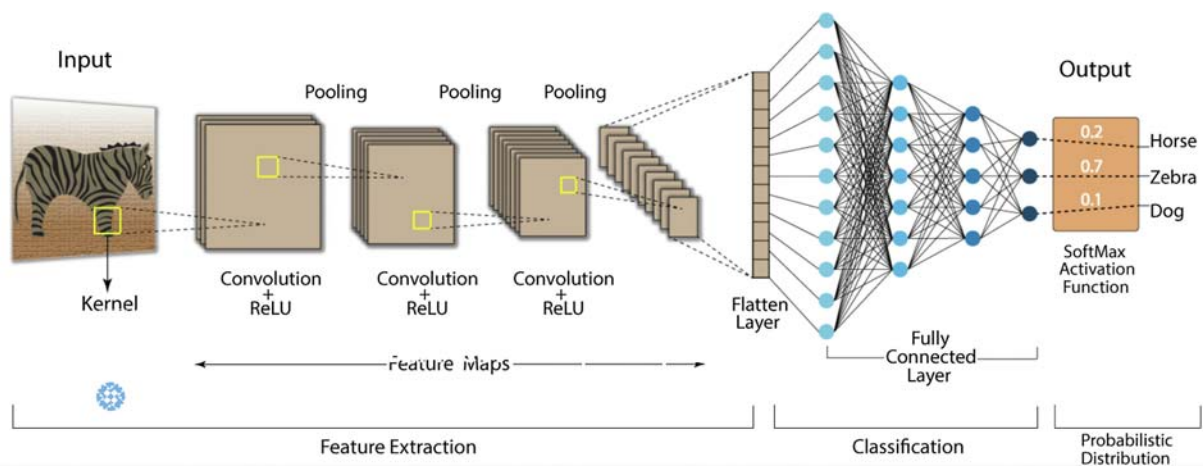


Fig. 6 A typical Convolutional Neural Network (CNN) architecture. A multi-scale feature representation is first extracted in convolutional layers (from edges to corners to object parts to entire objects). Fully-connected layers then perform the abstract object classification. Reproduced from [Developers Breach \(2021\)](#).


```

# Import needed packages.
import torch
import torchvision
from torchvision.datasets import CIFAR10
from torch.utils.data import DataLoader

# Check if GPU support is available.
cuda_avail = torch.cuda.is_available()

# Load VGG model with pre-trained weights.
net = torchvision.models.vgg19(pretrained=True)

# Load the test set (10,000 images).
batch_size = 32
test_set_size = 10000
# Define pre-processing transformations (normalize mean and std).
test_transformations = torchvision.transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize((0.5,0.5,0.5), (0.5,0.5,0.5))
])
test_set = CIFAR10(root="./data", train=False, transform=test_transformations, download=True)
test_loader = DataLoader(test_set, batch_size=batch_size, shuffle=False, num_workers=4)

# Compute the average accuracy and loss over all test images.
model.eval()
test_acc = 0.0
for i, (images, labels) in enumerate(test_loader):
    if cuda_avail:
        images = Variable(images.cuda())
        labels = Variable(labels.cuda())
    # Predict classes using images from the test set.
    outputs = model(images)
    _, prediction = torch.max(outputs.data, 1)
    prediction = prediction.cpu().numpy()
    test_acc += torch.sum(prediction == labels.data)
test_acc = test_acc / test_set_size
print("Test Accuracy: {:.3f}".format(test_acc))

```

Fig. 7 Python code example for measuring the prediction accuracy of the popular VGG CNN architecture (Simonyan and Zisserman, 2015) with pre-trained parameters on the CIFAR10 data set (Krizhevsky et al., 2010), using the PyTorch framework.

Python has become widely used for several reasons. It is easy to learn and use, prioritizes readability, and supports a rich ecosystem of libraries for data pre-processing and analysis (NumPy, SciPy, Pandas, scikit-learn), visualization (Matplotlib, Seaborn, Bokeh), and neural networks (the Facebook-developed PyTorch and the Google-developed TensorFlow, and Keras). Fig. 7 shows a basic NN implementation using PyTorch.

Applications to educational measurement

Automatic feature extraction: off-topic spoken response detection

In a speaking language proficiency test that elicits unconstrained spontaneous speech, especially if it is a low-stakes practice test, test takers may not provide good-faith responses if they are fatigued, unmotivated, distracted, etc. For instance, students may recite their response to a previous question (which would be off-topic), and this may help them reduce disfluencies and generate fluent speech. As a result, it may make it more difficult for an automated scoring system to provide a valid assessment of the proficiency of the speakers.

In (Lee et al., 2017), Min Lee et al. developed an off-topic response detection model in the context of automated scoring of non-native English speakers' spontaneous speech. The model consists of the three components shown in Fig. 8: a sentence modeling layer using *Siamese CNNs* (one for modeling questions, the other for modeling responses), a similarity distance calculation layer, and a fully-connected neural network layer. The model outputs an off-topic confidence score.

A data set of 100,000 spoken responses from a large-scale, standardized test of English speaking proficiency was collected. Each question asked to provide information about or opinions on familiar topics, and covered diverse topics: classroom activities, education and learning, and jobs, and work. The responses were converted into word hypotheses using an Automated Speech Recognition System (ASR). The NN was trained on 60% of the data and evaluated on 40%. The NN-based model had a 97% accuracy in detecting off-topic responses, compared with 94% for a baseline model trained using traditional ASR techniques. It also had a 3% higher recall (the fraction of responses flagged as off-topic by the model that were true off-topic responses). This illustrates the power of NNs in automatically extracting meaningful features from data. More generally, neuron activation values from the last few layers of a NN can be used as low-dimensional feature embedding for exploratory analysis (e.g., clustering).

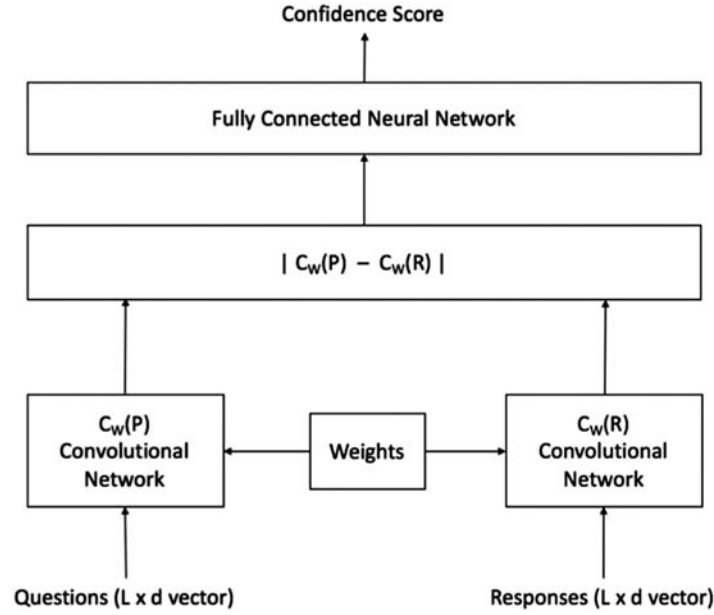


Fig. 8 Siamese Convolutional Neural Network architecture for detecting off-topic spoken language responses. Reproduced from Lee et al. (2017).

Time-series modeling: automated essay scoring

Recurrent neural networks

In a traditional NN, all inputs (and outputs) are assumed to be independent of each other; but in some tasks that's not true. Predicting the next word in a sentence depends on which words came before it (Huang and Khan, 2021).

Recurrent neural networks (RNNs) (Miljanovic, 2012) are designed to recognize patterns in sequential data (text, genomes, stock prices, time-series sensor data). Two prominent RNN variations are Long Short Term Memory unit (LSTM) (Hochreiter and Schmidhuber, 1997) and Gated Recurrent Unit (GRU) (Chung et al., 2014). Unlike feed-forward neural networks, they have *feedback connections*, and can process not only single data points, but also entire sequences of data. They can be thought of having a “memory” which captures information about what has been calculated so far.

Fig. 9 depicts a basic RNN (cf. (Bao et al., 2017)).

1. x_t the input at time t , e.g., a one-hot vector of the t th word in a sentence for $t = 1, 2, \dots, T$. (In natural language processing, *one-hot vector encoding* (One, 2021) is a mapping of a word in a sentence into a vector of binary bits, such that each word corresponds to a unique vector representation.)
2. $s_t = \sigma(Ux_t + Ws_{t-1})$ is the hidden state at time t , which contains the network's “memory” and depends on the current input and previous hidden state via the activation function σ (e.g., a ReLU). As in a hidden Markov model, the memory depends mostly on recent values and gradually “forgets” older states.
3. $o_t = \sigma(Vs_t)$ is the output at time t . The final output o_T is fed to a loss function as in feed-forward networks. The difference is that in feed-forward networks, all the inputs and outputs are independent of each other, whereas in an RNN we generate different intermediate outputs o_t that depend on all inputs seen till that point $(x_1, \dots, x_t)$.

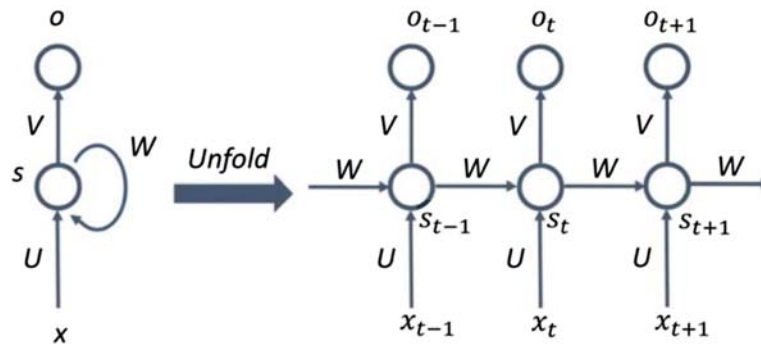


Fig. 9 A simple RNN architecture and its unfolded version. U , V and W are the weights of the hidden layer, the output layer and the hidden state, respectively. x_t and o_t are the input vector and output result at time t , respectively. Reproduced from Bao et al. (2017).

Table 2 The Quadratic weight Kappa (QWK) values of three automated essay scoring models on the ASAP data set.

Model	Prompts								Average
	1	2	3	4	5	6	7	8	
Score range	2 – 12	1 – 6	0 – 3	0 – 3	0 – 4	0 – 4	0 – 30	0 – 60	–
LSTM-CNN-att	0.822	0.682	0.672	0.814	0.803	0.811	0.801	0.705	0.764
SKIPFLOW	0.832	0.684	0.695	0.788	0.815	0.810	0.800	0.697	0.764
SBLSTMA	0.861	0.731	0.780	0.818	0.842	0.820	0.810	0.746	0.801

Reproduced from Liang et al. (2018).

Automated essay scoring

In (Liang et al., 2018), Liang et al. applied RNNs to Automated Essay Scoring (AES). AES helps reduce manual workload and speed up learning feedback. In that work, both the sample essays and rating criteria provided by domain experts were used as an input pair to a Siamese Bidirectional Long Short-Term Memory Architecture (SBLSTMA) model.

The model was evaluated on the Kaggle Automated Student Assessment Prize (ASAP) benchmark data set against two state-of-the-art baselines: SKIPFLOW and LSTM-CNN-att, a recurrent convolutional neural network (Dong et al., 2017). ASAP contains eight prompts, each of which is a different genre.

The SBLSTMA outperforms both baseline models (LSTM-CNN-att and SKIPFLOW) by approximately 5% on average Quadratic weight Kappa (QWK) (Sim and Wright, 2005) (cf. Table 2).

Other applications of RNNs in education include speech recognition, grammar learning, human action recognition, and in particular, natural language processing tasks such as semantic parsing and question answering, document summarization, and machine translation (Huang and Khan, 2021).

Transfer learning

Transfer learning (TL) is a machine learning approach aiming to exploit the knowledge retrieved from Problem A for improving the predictive performance of a learning model for a different but related Problem B (Tsiakmaki et al., 2020). TL is particularly useful when the available data for B at hand is set is small, but there is plenty of data about A. For instance, knowledge gained while learning to recognize cars (A) could apply when trying to recognize trucks (B); similarly, understanding language in general (A) can inform understanding the sentiment in student essays (B). DNNs provide an attractive approach to TL: a pre-trained model on the generic task (say, language) serves as the initial guess. It is fine-tuned by only training the last few layers, typically the fully-connected layers, on the data set of the new task. This reduces the number of learned parameters to avoid over-fitting, reduce training time, and exploit the universal representations of lower-level constructs already learned in the earlier layers.

TL has become an emerging research area in ML for education. In particular, they allow building predictive models of student behavior. Two examples: (a) (Tsiakmaki et al., 2020) demonstrates that predicting student performance (or at least identifying students at risk of failure) in a course can be achieved with satisfactory accuracy, provided that data sets of students who have attended other related courses are available; (b) in a Massive Open Online Course (MOOC), predictive models of student behavior can support multiple aspects of learning, including instructor feedback and timely intervention. Notwithstanding, when student outcomes are yet unknown, models trained on historical data have poor prediction performance. In (Ding et al., 2018), an automated transductive transfer learning approach using a temporal organization of the MOOC click stream data. An LSTM neural network was used to generate embedding of the data.

Generative adversarial networks and content generation

A *generative model* is a method to learn the true data distribution of real world from a training set, and then generate new training instances with some variations.

Generative adversarial networks (GANs) (first introduced by Goodfellow et al. (2014)) have dramatically advanced AI content generation. A GAN is a pair of neural networks, a *generator* and a *discriminator*, contesting with each other in a zero-sum game. The generator $G(z)$ learns to map random inputs z from a latent space to a particular data distribution of interest, while the discriminator $D(x)$ learns to discriminate between instances from the true data distribution and candidates produced by the generator.

The generator's training objective is to increase the error rate of the discriminator (i.e., to "fool" the discriminator network by producing fake synthesized instances that appear to have come from the true data distribution). The discriminator's training objective is to accurately predict whether the input x is real or generated (a binary classification) (Huang and Khan, 2021).

After training, the generative network can then be used to create new plausible samples on demand. GANs have been extremely successful in generating photographs of human faces; realistic cartoon characters; image-to-image translation; text-to-image translation; 3D object generation; and more (Brownlee, 2021). The next generation of learning and assessment systems will greatly

benefit by these advances. For instance, a game-based adaptive testing environment may need generate items and objects on the fly based; a collaborative learning environment may feature a computer-generated agent whose appearance and content are generated by GANs.

Summary

- Machine learning is a sub-domain of artificial intelligence, which focuses on the use and development of computer systems that can learn from examples. Parameters and data structures are automatically adjusted to optimize an objective function (a performance on a task), based on training examples. These structures learn patterns in the data that generalize to examples not seen before by the system.
- Machine learning includes supervised learning (which uses data labels), unsupervised learning (no labels/exploratory), semi-supervised learning (only part of the data is labeled), and reinforcement learning (operating against an adversary).
- Deep learning is a particularly popular family of machine learning methods that uses artificial neural networks, which have outperformed competing algorithms on many pattern recognition, natural language processing, and decision-making tasks.
- A neural network comprises of a sequence of layers, each containing many neurons. The strength of connection between neurons is updated during the learning process. Layers can either be convolutional or fully-connected. Training is performed using an optimizer such as stochastic gradient descent; open source Python implementations such as PyTorch and Tensorflow are typically used to train and test.
- Neural networks have been applied with great success to educational measurement tasks, including off-topic spoken response detection, time-series modeling, automated scoring of structured and essay responses, and content generation.

Future directions

ML and NN will likely continue to influence educational measurement, perhaps more than ever. And will influence educational measurement. Detecting sentiment, argumentation and cheating in essays might utilize features generated by powerful language models such as GPT-2 and GPT-3 (Brown et al., 2020). Novel Item Response theory (IRT) that use deep knowledge tracing models relying on time-dependent neural networks have been suggested (Converse, 2021).

Under the hood, field-agnostic fundamental research on better NN training algorithms, outlier identification, and increasing accuracy, has been booming, and will continue to rapidly evolve (Perez et al., 2020).

References

- Amato, F., López, A., Peña-Méndez, E.M., Vanihara, P., Hampl, A., Havel, J., 2013. Artificial neural networks in medical diagnosis. *J. Appl. Biomed.* 11 (2), 47–58.
- Armasu, L., 2016. AlphaGo AI Defeats Sedol Again, With "Near Perfect Game". <https://www.tomshardware.com/news/alphago-defeats-sedol-second-time,31377.html>. (Accessed 24 August 2021).
- Assiri, Y., 2020. Stochastic Optimization of Plain Convolutional Neural Networks With Simple Methods.
- Bao, W., Yue, J., Rao, Y., 2017. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS One* 12.
- Brown, T.B., Mann, B., Ryder, N., et al., 2020. Language Models Are Few-Shot Learners.
- Brownlee, J., 2019. 14 Different Types of Learning in Machine Learning. <https://machinelearningmastery.com/types-of-learning-in-machine-learning/>. (Accessed 24 August 2021).
- Brownlee, J., 2021. 18 Impressive Applications of Generative Adversarial Networks (GANs). <https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/>. (Accessed 24 August 2021).
- Byerly, A., Kalganova, T., Dear, I., 2021. No Routing Needed Between Capsules.
- Chung, J., Gulcehre, C., Cho, K.H., Bengio, Y., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling.
- Clark, A., Fox, C., Lappin, S. (Eds.), 2010. *The Handbook of Computational Linguistics and Natural Language Processing*. Wiley-Blackwell.
- Converse, G., 2021. *Neural Network Methods for Application in Educational Measurement*. PhD Thesis. University of Iowa.
- Cristianini, N., Ricci, E., 2008. *Support Vector Machines*. Springer, US, pp. 928–932.
- Dada, E.G., Stephen Bassi, J., Chiroma, H., Abdulhamid, S.M., Adetunmbi, A.O., Ajibuwa, O.E., 2019. Machine learning for email spam filtering: review, approaches and open research problems. *Heliyon* 5 (6), e01802.
- Developers Breach, 2021. Convolutional Neural Network. <https://developersbreach.com/convolution-neural-network-deep-learning/>. (Accessed 24 August 2021).
- Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding.
- Diederik, P., 2017. Kingma and Jimmy Ba. Adam: A Method for stochastic Optimization.
- Ding, M., Wang, Y., Hemberg, E., O'Reilly, U.-M., 2018. Transfer Learning Using Representation Learning in Massive Open Online Courses.
- Dong, F., Zhang, Y., Yang, J., 2017. Attention-based recurrent convolutional neural network for automatic essay scoring. In: *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*. Association for Computational Linguistics, pp. 153–162.
- Fu, K., Cheng, D., Tu, Y., Zhang, L., 2016. Credit card fraud detection using convolutional neural networks. In: Hirose, A., Ozawa, S., Doya, K., Ikeda, K., Lee, M., Liu, D. (Eds.), *Neural Information Processing*. Springer International Publishing, pp. 483–490.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (Eds.), *Advances in Neural Information Processing Systems*, vol. 27. Curran Associates, Inc.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Gossett, S., 2021. 9 Data Science Programming Languages to Know. <https://builtin.com/data-science/data-science-programming-languages>. (Accessed 24 August 2021).
- Gramacki, A., 2017. *Nonparametric Kernel Density Estimation and Its Computational Aspects*, first ed. Springer Publishing Company, Incorporated.
- Grigorescu, S., Trasnea, B., Cocias, T., Macesanu, G., 2020. A survey of deep learning techniques for autonomous driving. *J. Field Robot.* 37 (3), 362–386.
- Hao, J., Chen, L., Flor, M., Liu, L., von Davier, A.A., 2017. CPS-Rater: Automated Sequential Annotation for Conversations in Collaborative Problem-Solving Activities. Techreport. ETS Research Report.

- He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition.
- Herculano-Houzel, S., 2009. The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* 3, 31.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Huang, Y., Khan, S.M., 2021. Advances in AI and machine learning for education research. In: Von Davier, A.A., Mislevy, R., Hao, J. (Eds.), *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment*, Chapter 11. Springer International Publishing.
- Hussain, S., Nazir, R., Javed, U., Khan, S., Sofi, R., 2022. Speech recognition using artificial neural network. In: Raj, J.S., Ram, P., Perikos, I., Shi, Y. (Eds.), *Intelligent Sustainable Systems*. Springer, Singapore, pp. 83–92.
- Jing, L., Cheng, J.H., Shi, J.Y., Huang, F., 2012. Brief introduction of back propagation (bp) neural network algorithm and its improvement. In: Jin, D., Lin, S. (Eds.), *Advances in Computer Science and Information Engineering*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 553–558.
- Keskar, N.S., Mudigere, D., Nocedal, J., Smelyanskiy, M., Tang, P.T.P., 2017. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima.
- Kim, Y.J., Shute, V.J., 2015. The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Comput. Educ.* 87 (C), 340–356.
- Koh, J.Y., 2021. Model Zoo. <https://modelzoo.co/>. (Accessed 24 August 2021).
- Krizhevsky, A., Nair, V., Hinton, G., 2010. Cifar-10 (Canadian Institute for Advanced Research). <http://www.cs.toronto.edu/kriz/cifar.html>. (Accessed 24 August 2021).
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems. NIPS'12*, vol. 1. Curran Associates Inc., pp. 1097–1105.
- Lakew, S.M., Cettolo, M., Federico, M., 2018. A Comparison of Transformer and Recurrent Neural Networks on Multilingual Neural Machine Translation.
- LeCun, Y., Jackel, L.D., Bottou, L., Brunot, A., Cortes, C., Denker, J.S., Drucker, H., Guyon, I., Muller, U.A., Sackinger, E., Simard, P., Vapnik, V., 1995. Comparison of learning algorithms for handwritten digit recognition. In: Fogelman, F., Gallinari, P. (Eds.), *International Conference on Artificial Neural Networks. EC2 & Cie*, pp. 53–60.
- LeCun, Y., Kavukcuoglu, K., Farabet, C., 2010. Convolutional networks and applications in vision. In: *ISCAS 2010—2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems*, pp. 253–256. Copyright: Copyright 2010 Elsevier B.V., All rights reserved; 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, ISCAS 2010; Conference date: 30-05-2010 Through 02-06-2010.
- Lee, C.M., Yoon, S.-Y., Wang, X., David Mulholland, M., Choi, I., Evanini, K., 2017. Off-Topic Spoken Response Detection Using Siamese Convolutional Neural Networks. *INTERSPEECH*.
- Liang, G., On, B.-W., Jeong, D., Kim, H.-C., Choi, G.S., 2018. Automated essay scoring: a siamese bidirectional LSTM neural network architecture. *Symmetry* 10 (12).
- Mazzia, V., Salvetti, F., Chiaberge, M., 2021. Efficient-Capsnet: Capsule Network With Self-Attention Routing.
- Miao, H., Li, A., Davis, L.S., Deshpande, A., 2017. Towards unified data and lifecycle management for deep learning. In: *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pp. 571–582.
- Miljanovic, M., 2012. Comparative analysis of recurrent and finite impulse response neural networks in time series prediction. *Indian J. Comput. Sci. Eng.* 3.
- Mislevy, R.J., Oranje, A., Bauer, M.I., Hao, J., Von Davier, A., 2014. *Psychometric Considerations in Game-Based Assessment*. Createspace Independent Pub.
- Neiriz, R., Ercikan, K., Pellegrino, J.W., 2017. Validation of score meaning for the next generation of assessments: the use of response processes. New York, NY: Routledge. *J. Educ. Meas.* 57 (1), 152–155.
- One, J., 2021. Hot Encoding of Text Data in Natural Language Processing. In: <https://medium.com/analytics-vidhya/one-hot-encoding-of-text-data-in-natural-language-processing-2242feb2148>. (Accessed 24 August 2021).
- PapersWithCode.com. Image Classification on MNIST, 2021. <https://paperswithcode.com/sota/image-classification-on-mnist>. (Accessed 24 August 2021).
- Perez, H., Joseph, H., Tah, M., 2020. Improving the accuracy of convolutional neural networks by identifying and removing outlier images in datasets using t-SNE. *Mathematics* 8 (5).
- Polyak, S.T., von Davier, A.A., Peterschmidt, K., 2017. Computational psychometrics for the measurement of collaborative problem solving skills. *Front. Psychol.* 8, 2029.
- Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65 (6), 386–408.
- Rosenblatt, F., 1962. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Cornell Aeronautical Laboratory. Report No. VG-1196-G-8. Spartan Books.
- Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65.
- Seif, G., 2018. The 5 Clustering Algorithms Data Scientists Need to Know. <https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68>. (Accessed 24 August 2021).
- Shen, J., 2013. The Science of Practice: What Happens When You Learn a New Skill. <https://lifehacker.com/the-science-of-practice-what-happens-when-you-learn-a-510255025>. (Accessed 24 August 2021).
- Silbernagel, C., 2019. Investigation of the Design, Manufacture and Testing of Additively Manufactured Coils for Electric Motor Applications. PhD Thesis. University of Nottingham, p. 10.
- Silver, D., Huang, A., Maddison, C.J., et al., 2016. Mastering the game of go with deep neural networks and tree search. *Nature* 529, 484–503.
- Sim, J., Wright, C.C., 2005. The Kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys. Ther.* 85 (3), 257–268.
- Simonyan, K., Zisserman, A., 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition.
- Sonare, B., Padgal, A., Gaikwad, Y., Patil, A., 2021. Video-based sign language translation system using machine learning. In: *2021 2nd International Conference for Emerging Technology (INCEIT)*, pp. 1–4.
- Sutton, R.S., Barto, A.G., 2018. *Reinforcement Learning: An Introduction*, second ed. The MIT Press.
- Topircceanu, A., Grossecck, G., 2017. Decision tree learning used for the classification of student archetypes in online courses. *Proc. Comput. Sci.* 112, 51–60.
- Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., Ragos, O., 2020. Transfer learning from deep neural networks for predicting student performance. *Appl. Sci.* 10 (6).
- Varela-Santos, S., Melin, P., 2021. A new approach for classifying coronavirus Covid-19 based on its manifestation on chest x-rays using texture features and neural networks. *Inf. Sci.* 545, 403–414.
- von Davier, A.A., Hao, J., Liu, L., Kyllonen, P., 2017. Interdisciplinary research agenda in support of assessment of collaborative problem solving: lessons learned from developing a collaborative science assessment prototype. *Comput. Hum. Behav.* 76, 631–640.
- Wikipedia, 2021a. Machine Learning. https://en.wikipedia.org/wiki/Machine_learning. (Accessed 24 August 2021).
- Wikipedia, 2021b. MNIST Database. https://en.wikipedia.org/wiki/MNIST_database. (Accessed 24 August 2021).
- Wikipedia, 2021c. Softmax Function. https://en.wikipedia.org/wiki/Softmax_function. (Accessed 24 August 2021).
- Wu, W., Chen, L., Yang, Q., 2016. Students' Personality and Chat Room Behavior in Synchronous Online Learning. *UMAP*.

Research study design

Garret J. Hall^a and Craig A. Albers^b, ^a Florida State University, Tallahassee, FL, United States; and ^b University of Wisconsin, Madison, WI, United States

© 2023 Elsevier Ltd. All rights reserved.

Experimental design and statistical procedure considerations	696
Research designs	696
Internal validity and causal inference	697
Descriptive and correlational studies	697
Causal studies	698
Randomized designs	698
Statistical analyses of causal designs	701
Summary	701
References	701

Glossary

Causal studies A type of research study designed to determine whether one or more variables actually cause changes in one or more outcome variables

Control group The group that does not receive the intervention within a research study

Correlational studies A type of research study that involves relating two or more variables without inferring causal relations

Descriptive studies A type of research study that involves the collection of data to simply describe a phenomenon

Internal validity The extent to which a researcher can accurately state that the independent variable caused the changes in the outcome variable(s)

Intervention group The group that receives the intervention within a research study

Exposure A variable (e.g., an intervention, a policy, another phenomenon) to which a unit (e.g., a person, a school, a community) is subjected that is theorized to cause change in one or more outcomes. Exposure may occur randomly or non-randomly

Exogenous Describes something that is unrelated to characteristics or qualities of a unit (i.e., random and/or external)

Endogenous Describes something that is related or attributable to internal qualities or characteristics of a unit (i.e., systematic)

Observational study A research design in which assignment to intervention or control groups is not done by random assignment. These studies may also be referred to as quasi- or non-experimental designs. Observational studies are more inclusive of all non-randomized studies, including studies designed to detect causal effects or correlate two or more variables

Random assignment A process or procedure in which units are assigned to the intervention or control group by a random process

Randomized experimental design A research study in which assignment to intervention or control groups is done by random assignment

Single-case design Studies involving one or a small number of individuals that are exposed to different treatments over time. Repeated measures are taken on each person and relations of repeated measures to changes in treatment are analyzed

“What works” has always been the primary question underlying educational research, and no more prevalent has that question been than since the rise of “education science” in the last 20–30 years. In the United States, much of this field of education science gained traction because of policy initiatives in the early 2000s through the No Child Left Behind (NCLB) legislation to systematically determine effective practices and use assessment as an accountability mechanism to promote more effective instruction. Many research and accountability efforts had been in place in the US and elsewhere prior to this; however, NCLB was one of the largest scale education policy initiatives at the time and promoted international attention to educational research as a quantitative science.

Like other social sciences concerned with the influence of policies and exposures (e.g., economics, epidemiology, sociology) on social outcomes, identifying what specific practices, curricula, and policies in education promote high-quality education and student outcomes is a complex causal question. In addition to having socially and organizationally complex theoretical relations in an education system, the specific statistical analyses used to examine these questions are directly linked to the theories and practices of determining educational effectiveness. Thus, in addition to articulating clear theories and hypotheses about effective educational practices and policies, education researchers must adhere to a set standards of sound research design and analyses. The goal of this

entry is to (a) discuss the importance of research in education, (b) outline the legal and professional guidelines for research in education, and (c) provide an overview of popular designs and statistical techniques used in education research, particularly those used for causal inference.

The United States in the early 2000s saw a massive push to establish an evidence base for effective educational practices after years of evolving and growing calls for evidence in education (see Eisenhart and Towne, 2003, for historical perspective of the evolution of the definition of scientifically-based education research in NCLB). The tension between the policy initiatives calling for evidence-based practices and the perceived lack of rigor in education research promoted a more scientific perspective in the assessment, testing, and promotion of evidence. This was further legitimized through federal funding and organizations to promote education science as an interdisciplinary field (i.e., U.S. Department of Education's Institute of Education Science).

Experimental design and statistical procedure considerations

Randomization to conditions (i.e., experimental designs) have typically been viewed as the “gold standard” of causal inference. That is, when an intervention is properly designed to sample and then randomize participating individuals or units to treatment(s) and control conditions (in addition to using appropriate statistical analyses), it is clear to infer that the intervention(s) caused observed changes relative to the control group. The sampling properties of that design and the statistical analyses used allow inference as to whether that observed effect is meaningfully different from zero.

Standards for experimental designs in education research have become increasingly rigorous as new methods are developed and tested. Perhaps the most well-known standards of research in education have been developed at the U.S. Department of Education's Institute of Education Sciences (IES), particularly IES's What Works Clearinghouse (WWC; <https://ies.ed.gov/ncee/wwc/>). WWC provides not only practitioner and researcher resources on evidence-based practices but also procedural and design standards for randomized, observational, and single-case design research. WWC's Standards Handbook (2020a) and Procedures Handbook (2020b), now in their fourth editions, provide clear and detailed information regarding the design and analysis of educational experiments and observational studies. Likewise, Kratochwill et al. (2010) developed WWC's Single-Case Design Technical Documentation (2010) that provides information regarding the design and analysis of single-case research.

Research designs

Detailed and technical information on all relevant research designs in education is outside the scope of this entry. We provide a broad overview of relevant (a) observational and (b) experimental designs in education. However, unlike much of the other work in education, we do not draw distinctions between causal and non-causal studies. Rather, we focus on distinguishing between experimental versus observational, regardless of whether causal inferences are warranted. We take this approach for two primary reasons.

First, causal inference is based on a set of assumptions, not a particular design. Randomization is not immune to bias or failures to infer causation. Likewise, observational studies can meet assumptions needed to infer causation with crafty identification of temporal exposures or mechanisms by which exposures are induced and/or advanced statistical methods that help meet assumptions for causal inference. We also believe that few, if any, questions in education are truly non-causal. Instead, purely descriptive or correlational studies can provide a foundation for understanding key relations that aid in theory development and understanding context. Although descriptive and correlational questions might not, on their own, tell us whether A causes B, they might tell us why A and B matter and how causal questions can be addressed in future studies.

Second, the distinction between causal and non-causal studies further obfuscates the assumptions of causal inference in all studies and reinforces the idea that large-scale, randomized studies are the only way to meaningfully assess causal effects. Of course, experimentation offers the most straight-forward way to say whether A causes B. However, that does not always require randomizing individuals to groups, nor does it always require group designs at all. Moreover, randomization may not be feasible or ethical in some cases. For example, epidemiological work has recently shown that an ideal randomized experiment (a “target trial”) can be emulated under certain conditions with observational data, potentially alleviating the need to design an experiment and collect new data (Hernán and Robins, 2016). Breaking down distinctions between causal and non-causal studies, we believe, will help researchers critically assess assumptions of causal inference, even for studies that are randomized and those that do not involve large group comparisons.

Below, we first introduce some key assumptions for causal inference (internal validity) before describing descriptive, correlational, and causal research in observational studies. Then, we describe group-randomized experiments and their common designs. Finally, we introduce single-case designs.

Internal validity and causal inference

Internal validity is the extent to which a design meets assumptions necessary to infer causal relations between one or more independent variables and one or more dependent variables. The primary rationale for conducting randomized experiments is that the nature of randomization alleviates many (but not all) internal validity problems. Randomization to conditions ensures that differences between conditions are attributable only to actual differences between conditions. This is related to an assumption called *exchangeability*: Person A in the control group and Person B in the treatment group could be exchanged with no consequences except for a difference in their outcomes due to the treatment(s). Randomization ensures this can be easily met. Many observational studies make it more difficult to assume exchangeability. Thus, some observational studies have a similar assumption called *conditional independence* or *ignorability*: conditional on certain factors (e.g., control variables), treatment assignment is essentially random. The methods by which conditional independence can be met are statistically technical and beyond the scope of this entry. Regardless of the presence of randomization, other key assumptions for causal inference are positivity and consistency. *Positivity* means that everyone has a non-zero probability of treatment exposure. *Consistency* means that the observed treatment is the same for everyone and independent of everyone; one person's treatment is the same as the next and they do not affect each other.

A simple example of exchangeability/conditional independence, positivity, and consistency can be demonstrated using covariates in an observational study. Students are nonrandomly assigned to either a math intervention group or a control; everyone has a non-zero chance of receiving either condition (i.e., positivity), but some process determines some students are more likely than others to get one of the two conditions. The intervention is administered individually in the exact same way to all participants. It is assumed that there is no cross-over between conditions (i.e., students in the control group would not affect the intervention group and vice versa; consistency), and students in the intervention or control do not affect the outcomes of other students in the same group (i.e., no interference). A researcher collects a large set of demographic and prior achievement data as well as interventions status and the outcome variable. The researcher might assess the difference in outcomes between groups while controlling for the background variables. Under strong assumptions, it might be possible to assume that the nonrandom treatment is essentially random once it is conditioned on the covariates, thus making it easier to assume that observed differences between conditions are due to the effect of the intervention. In reality, conditional independence, positivity, and consistency cannot be measured directly, and the researcher must try to determine how much these concepts influence their design.

Exchangeability and conditional independence are inherently tied to the selection mechanism into treatment. [Shadish et al. \(2002\)](#) offered clear definitions for nine factors of internal validity, with selection being a primary factor. In a randomized experiment where all individuals adhere to their assigned treatment condition, the only mechanism of selection into treatment is the randomization; randomization is an *instrument* determining treatment exposure. In observational studies, there may be other non-random instruments that determine exposure to a treatment (hence *instrumental variables* analysis). In other cases, the selection mechanism might be unknown and/or one or more groups can be observed longitudinally prior to and following exposure to a treatment. Observing variables over time helps address temporal precedence, which can also help understand selection processes into treatments or mechanisms of exposures.

History and maturation are also important. Processes acting simultaneously with the treatment may make inferences about a treatment ambiguous (i.e., history), and processes that naturally change over time may be confounded with a treatment (i.e., maturation). Maturation may occur naturally, but this idea is also related to testing effects, whereby being measured on an outcome systematically changes subsequent measurements. Similarly, the measures themselves may take on different characteristics (or be changed by researchers in a study) across time or groups, which is called instrumentation. Regression to the mean can also conflate causal effects, as extreme scores are more likely to converge toward the average. Missing data (e.g., attrition), if it occurs for systematic reasons related to the measurements or the treatment itself, may introduce bias, as those remaining in the study or those who provided data might represent certain participant characteristics that conflate treatment assignment with outcomes. The ninth internal validity threat [Shadish et al. \(2002\)](#) discussed is the potential for these factors to add to or compound one another. These threats are not mutually exclusive.

Importantly, randomized experiments are not immune from these threats, and these threats can just as easily add to or compound each other in randomized or observational settings.

Thus, although successful randomization helps mitigate issues of selection or temporal precedence, the threats of history, maturation, regression, missing data, testing, and instrumentation can all work independently or together in threatening the validity of a randomized design as well as observational studies. Thus, although successful randomization helps mitigate issues of selection or temporal precedence, history, maturation, regression, missing data, testing, and instrumentation can all work independently or together in threatening the validity of a randomized design.

Descriptive and correlational studies

Descriptive and correlational studies are primarily characterized by their inability to meet the necessary assumptions of causal inference in addition to the goal of the study. Descriptive work may not (and likely cannot) meet assumptions of causal effects, but that is not the goal of descriptive studies. These studies seek to highlight levels, trends, and contexts of certain variables. For example, developmental psychologists might seek to describe the level, trend, shape, and variability of a variable over time, in addition to the context in which the variable was measured. This is important work describing how a variable changes over time, but it likely makes no inferences regarding the causal process of growth or the relations of growth to other variables.

Correlational studies might extend or supplement descriptive and causal studies by articulating hypotheses about the extent to which two or more variables are related and why. A specific statistical analysis is neither correlational nor causal (these differences depend on assumptions and design); however, researchers might use a regression analysis to determine the magnitude of relations between students' social-emotional-behavioral well-being and academic achievement. Researchers might be interested in these well-being and achievement relations when controlling for a variety of variables, yet that study still would not likely warrant clear causal inferences just by controlling for potential confounders. Nonetheless, it may inform theory as to why and to what degree social-emotional-learning relates to achievement, which then may help other researchers and practitioners more clearly articulate underlying causal relations.

Causal studies

Causal studies may or may not differ in design and data collection procedures from descriptive and correlational studies. Randomized experiments, of course, would differ drastically from a purely correlational study. In reality, descriptive, correlational, and causal questions can all be addressed with similar data. For example, extant data sources such as administrative records or non-intervention secondary datasets from governmental or independent sources may lend themselves to purely descriptive, correlational, or causal questions. The tenability of causal inferences depends solely on the assumptions that can be made about the design and analysis, which applies to randomized and observational settings. We divide our discussion of causal questions into (a) randomized designs and (b) observational designs. In both sections, we attend to both large-*N* and single-case designs.

Randomized designs

Randomized group designs consist of randomizing many units (e.g., individuals, schools, communities) to different conditions. Many different design configurations can be used depending on the unit of randomization. Below, we review a few popular designs.

Factorial designs

Factorial designs can take place with any level of analysis (e.g., people, groups, organizations). Factorial designs consist of two or more categorical variables; only one categorical variable would simplify a factorial design into a one-way design. With multiple conditions, their interactions, and potentially repeated measures, factorial designs can easily become complex, necessitating caution when making many comparisons to avoid Type 1 errors and planning adequate sampling to ensure statistical power. More complex analysis techniques to account for the many different sources of variance in large factorial designs in education are advancing, however, such as Bayesian methods (Gelman et al., 2012; Kassler et al., 2020).

As an example, 200 students might be randomized to four conditions that are combinations of Treatments A and B. Individuals could receive Treatment A (treatment or control) with or without Treatment B (treatment or control). This is a 2×2 factorial design where Treatments A and B are crossed factors (i.e., both treatments appear in all levels of the other treatment). We simulated data for this design, the values of which are presented in Table 1. In this design, there are a number of different estimates that might be of interest to a researcher. For example, a researcher might be interested in computing the effects of Treatment B at each level of Treatment A (called simple effects), or a researcher might wish to examine whether the treatment effect of Treatment B is larger at one level of Treatment A as compared to another. Main effects would include comparing marginal means of the treatment and control groups (average values) across all other treatment values.

Factorial designs can be crossed, partially crossed, or nested. In a fully crossed design, all values of each factor occur in all levels of other factors. A partially crossed design is one in which only some factors are crossed. Finally, a nested design is one in which levels of one factor only occur in certain levels of another factor. For example, students might be randomly assigned to receive one of two interventions, but within each intervention, there are two different kinds of intervention settings: within one intervention condition, students are randomly assigned to one of two intervention settings. In the other intervention condition, students are randomly assigned to one of two different settings. The nested factorial design is closely related to what is commonly observed in applied research where classrooms are nested in schools: the same classroom cannot appear in different schools.

Blocked design

A blocked design occurs when individuals are randomized to conditions within blocks. These blocks are typically meaningful units of analysis, and this within-block randomization helps control for between-block variability that would not be of substantive interest to the study. However, blocking can leverage existing groupings to ease implementation of treatment allocation (using

Table 1 Means by condition for 2×2 factorial analysis of variance example.

		Condition B <i>M</i> (<i>SD</i>)		Marginal means: condition A
		Treatment B	Control B	
Condition A <i>M</i> (<i>SD</i>)	Treatment A	49.6 (10.6)	54.6 (11.2)	52.1
	Control A	51.2 (13.3)	49.7 (9.1)	47.95
	Marginal means: condition B	50.4	49.7	

Note. Data generated using the *faux* package (DeBruine, 2021) in R (R Core Team, 2021).

schools as blocks being a prime example). Typically, the blocking variable is included as a control variable when analyzing a randomized blocked design. As an example, students might be randomized to treatments within classrooms or schools. Then, the classroom or the school is accounted for in some way when analyzing the intervention effects. Blocking can be used with any grouping variable, however, especially if that grouping variable would be meaningful to account for when implementing a treatment. For example, a treatment might be randomized within grade levels or age ranges to account for variability across age levels. This would be a natural choice when it is assumed the blocking variable might relate to the outcome but not in a way that is of interest to the study. A blocked design could easily be incorporated within a nested, partially, or fully factorial design. Typically, the blocking factor in a blocked design is not of substantive interest because the goal is to remove variation in the outcome that could be attributed to the blocking variable. As a result, blocking can increase statistical power to detect a treatment effect relative to the same experiment conducted without blocking.

Cluster-randomized design

A cluster-randomized design (CRD) involves the randomization of existing groups to treatment conditions. In education, the most common example of the CRD is the randomization of classrooms or schools to treatment conditions. This differs from the blocked design in that the higher-level unit (e.g., teachers and/or schools) is of substantive interest because it is the unit assigned to treatment condition, whereas blocking variables in blocked designs are frequently not of direct interest to the study. However, CRDs can also be blocked (e.g., randomizing teachers within schools to treatment conditions).

CRDs inherently involve nesting of units within one or more higher-level groupings (e.g., students nested in classrooms and schools). This nesting creates a correlation among units within a higher-level grouping (e.g., students' achievement within a school possesses similarity for a variety of reasons). Nesting requires special care when analyzing data, and although nesting could apply to many different factorial and blocked designs, the nesting is an explicit component of CRDs. Multiple methods can be used to handle nesting in data, including multilevel modeling, cluster-robust standard error estimation, fixed effects estimation (e.g., including dummy codes for each higher-level grouping), generalized estimating equations (GEE), and design effects (Huang, 2018; McNeish et al., 2017), although multilevel modeling and cluster-robust standard error estimation tend to be the most prevalent in education when analyzing cluster-randomized trials.

Nesting is important to take into account because the correlation among units within groups can bias downward the estimation of variability if that correlation within units is ignored. The methods mentioned above all take into account nesting but do so in very different ways. If one is interested in estimating the variance in the outcome between higher-level groupings in a CRD, then multilevel modeling should be used. Cluster-robust estimation, GEE, and design effects all seek to correct the variance in the presence of clustering but do not model the variability across higher-level groups (see McNeish et al., 2017). We refer to the reader to McNeish et al. (2017) and Huang (2018) for further discussion of these modeling issues when analyzing nested data.

Another important implication of the CRD is that the correlation among units within groups reduces the statistical power to detect a treatment effect. The sample size in a CRD needs to be increased to obtain the same level of statistical power that one might have with an individually-randomized trial. This occurs because the correlation among units (e.g., students) implies less independence of outcome values within groupings, so outcome data within units are less unique than they would be if treatment conditions were allocated at the individual level. This dependence makes it harder to detect an effect because each data point is not contributing as uniquely in a CRD as compared to treatments randomized at the individual level. Recent studies detail issues of power in CRDs in education, and we refer readers to these articles for more discussion of power issues in CRDs (Spybrook et al., 2016, 2020).

Single-case research designs

Over the last decade or so, significant developments have been made with respect to single-case research. SCDs are conducted on one or more (usually a small number) units. These units, frequently individuals, are followed over time, and the time period over which they are followed includes changing exposures to interventions or stimuli. Although there are many different kinds of single-case designs, we focus on three widely used types: ABAB/withdrawal designs, multiple-baseline designs, and alternating treatments designs. ABAB designs consist of two conditions that are sequentially implemented for an individual; the individual is exposed to A, then B, then A again, then B again. This could extend on for as many time periods as feasible. With multiple individuals, the sequencing of A and B phases can be randomized. This combination of randomization and replication helps rule out internal validity threats including maturation and history. Multiple baseline designs involve a baseline and a treatment phase (as well as a follow-up phase in some cases), and the baseline length is staggered and randomized across individuals. With alternating treatments designs, individuals typically receive a baseline phase with one treatment, a treatment phase in which treatments A and B are alternated between immediate sessions, and a post-treatment phase based on the relative effectiveness of the interventions in the treatment phase.

The key factor across all these designs is the emphasis on replication of treatments and phases across individuals, although some designs also allow within-person replication (e.g., ABAB designs). Randomization has become a core component of single-case designs (SCDs; Kratochwill and Levin, 2010), whereas for many years SCDs were considered quasi- or non-experimental. These advancements have led to significant improvements in the rigor of SCDs in a variety of ways. Just like with group designs, randomization helps establish that the treatment is not systematically related to specific participant or unit characteristics. In addition, replication helps establish whether other factors are systematically related to the treatment (e.g., maturation, history).

Observational studies

As aforementioned, observational studies can be anything that does not involve researcher manipulation of treatment assignment. When seeking to identify causal effects in an observational setting, researchers are frequently interested in an intervention, policy, or phenomenon that is theorized to impact one or more outcomes. Often, mechanisms of exposure to those factors are nonrandom. Individuals may self-select into a treatment (or others may select them for treatment), or they may be exposed to a treatment in their setting due to treatment occurring at a context level (e.g., schools implementing curricula). Random exposures that are not created by a researcher are also possible in observational studies, such as lottery studies. For example, lottery enrollment is frequently used for dual-language education programs, which can be used to estimate the effects of dual-language education on outcomes (Steele et al., 2017). Importantly, the method of analyzing the observational study should align to the context of the treatment. Below, we briefly review some of these methods that are used to assess causal effects in observational settings. All of these methods can be combined, and sometimes that is a necessity for answering a causal question with observational data.

Regression discontinuity

Regression discontinuity designs (Lee and Lemieux, 2010) involve the use of a variable that determines assignment to conditions as a means of obtaining causal inferences. For example, a cut score might determine the placement of individuals into educational programs (e.g., Umansky, 2016). The variable creating the discontinuity (usually called the running variable) can be controlled for, which helps reduce any additional confounding between the treatment and the outcome that is due to the running variable. In education research, discontinuities frequently occur based on test scores. Even with a nonrandom running variable, individuals just above and just below the cutoff are likely to be very similar, so at the cutoff, a discontinuity design emulates a randomized experiment. Inferences about the causal effect that the discontinuity has are less tenable as one moves further away from the cutoff. As a recent example, Bonilla et al. (2021) examined how a high school ethnic studies class impacted long-term academic outcomes by using the grade-point average (GPA) cutoff that determined students' exposure to the ethnic studies class ($GPA = 2.0$). In their analysis, they controlled for a variety of demographic characteristics in addition to GPA, which determined treatment exposure.

Instrumental variable model

A frequent use of instrumental variables analysis is when there is an endogenous treatment or exposure (e.g., individuals self-select into treatment), but there may be another variable (or multiple) that predicts exposure or selection (the instrument[s]). Provided that the instrument does not also impact the outcome of interest (referred to as the exclusion criterion), specific statistical methods (often two-stage least squares) can be used to identify the extent to which the treatment or exposure impacted the outcome by using the instrumental variable to predict treatment exposure. Methods used in instrumental variable analyses help identify the part of the originally endogenous treatment that can be assumed to be *exogenous*, or external to specific characteristics of the units being investigated. Baiochi et al. (2014) provided a tutorial on using instrumental variables for causal inference in health settings, but these same concepts can be applied in education settings. Instrumental variable analysis can also be used when there is noncompliance to treatment assignment in randomized controlled trials and discontinuity designs. See Huang (2018) for an example of using instrumental variable analysis for noncompliance in experiments.

Weighting/matching

Often an instrument is not available, but treatment condition or exposure status is in addition to many other pre-treatment variables. In these cases, weighting or matching methods may be appropriate. These methods seek to create equivalent treatment groups by modeling the probability of treatment based on a large set of covariates. These variables would be used to predict treatment status, and the probabilities of treatment status (the propensity score) would then be used to match individuals across conditions. This helps establish the conditional independence assumption. Weighting methods simply transform the propensity score into a weight. Regression adjustment accomplishes a similar goal as weighting and propensity scores under certain conditions. We refer readers to Powell et al. (2020) for more on propensity score analyses in education.

Time-series designs

Time series designs involve following units (e.g., people, groups, organizations) over time. Interrupted time-series (ITS) designs occur when an intervention or exposure occurs during the time-series. In a typical ITS, the unit (or all units if more than one) in the time-series are exposed to the same intervention at the same time. Comparative interrupted time-series (CITS) designs involve the comparison of units that receive the intervention at different times or some units may not receive it at all. It should be noted that the CITS design also includes the difference-in-difference (DID) design, which compares the change in an outcome from pre-to post-exposure among an exposed group to a group that was not exposed (see Hall et al., 2021, for a recent example involving math achievement). Typically, the CITS design includes multiple pre- and post-exposure data points, whereas DID might include only one pre- and one post-exposure measurement. We refer readers to Somers et al. (2013) for more on CITS and DID designs and estimation for group-based studies. Importantly, single-case designs are, in fact, ITS and CITS designs. However, (C)ITS designs used in education and policy analysis frequently involve analysis at the group level and nonrandom treatment exposure, yet the principles of these designs are parallel to those of single-case designs.

Statistical analyses of causal designs

The statistical procedures available to researchers are innumerable and can be difficult to choose. More importantly, it is critical to align the statistical analyses with the data-generating process under investigation. This requires careful consideration of the design of the study and the properties of the independent and dependent variables under analysis. An incorrect application of a statistical procedure could ruin an analysis of even the best randomized design. For example, a study of the effect of a social-emotional curriculum on suspensions and office-discipline referrals might involve not only school-level assignment of the curriculum but a zero-inflated count outcome distribution (i.e., a large number of zeros relative to other values, as would likely be the case with suspension data). There are multiple ways to account for not only the nested nature of students in schools and school-level intervention assignment as well as the count outcome, and the choice of the appropriate analysis depends primarily on the inferences one would wish to draw from this study. Failure to account for the clustered design of the study as well as the type of outcome distribution could invalidate causal inferences about the intervention, despite a well-designed and implemented study.

Matching the design of one's study—including nesting, longitudinal data, and measurement scales—to the appropriate statistical analysis can be a challenging task in some cases. For example, some data might require nonparametric techniques or unique model specifications to account for difference sources of variance and dependence structures. In many cases, however, the analysis choices are a bit clearer. Using the 2×2 design described above, one would use analysis of variance (ANOVA)/linear regression.

In this design, one might be interested in the effects of either Treatment A (averaged across Treatment B) or Treatment B (averaged across Treatment A). One might also be interested in whether Treatment B works better in the presence of Treatment A or not, which is called an interaction effect or a moderation effect (i.e., Treatment A moderates or changes the effect of Treatment B). Using a two-way ANOVA that models the main effect of Treatments A and B as well as their interaction, one could conclude that Treatment A works (averaging across Treatment B), but Treatment B is not effective on average. However, Treatment B does work when it is combined with Treatment A, whereas it does not work without Treatment A.

The choice of this analysis is very clear as this design is the basic factorial ANOVA model. However, the decision for a particular analysis depends on a theoretical and practical understanding of the data-generating process being studied. Many designs have clear analysis decisions, but others do not, and these analysis issues should be understood as much before conducting a study as possible. When this is not possible, researchers should take care to select a method that comports with their theory and the inferences they wish to draw. For example, if Treatment A were a time variable (e.g., pre-test, post-test), the analytic model would need to change to accommodate the repeated observations of individuals in the study. With this addition, there might be many analytic choices, all of which offer strengths and drawbacks, although some may better align with specific theories than others.

Summary

The past two decades have been marked by continued growth in the amount of research conducted in educational settings and by growth in advanced research designs and analytical approaches. Within education, the expectations for what constitutes acceptable education practices have undergone a significant and meaningful transformation during this time with an increased emphasis in the application of doing what scientifically-sound research has clearly demonstrated to be effective practices.

Although randomized experimental designs are often held as the preferred method for determining evidence-based practice (e.g., by the U.S. Department of Education), it would be erroneous to automatically presume that such designs are the best design in all settings. Using such a design does not guarantee that various threats to the study's validity are not present. Instead, each of the research designs should be recognized to have relative strengths and weaknesses, with the results being interpreted with these in mind. In recognition of this, we described alternate research designs, including descriptions of observational versus experimental designs.

References

- Baiocchi, M., Cheng, J., Small, D.S., 2014. Tutorial in biostatistics: instrumental variable methods for causal inference. *Stat. Med.* 33 (13), 2297–2340.
- Bonilla, S., Dee, T.S., Penner, E.K., 2021. Ethnic studies increases longer-run academic engagement and attainment. *Proc. Natl. Acad. Sci. U. S. A.* 188 (37) e2026386118.
- DeBruine, L., 2021. Faux: Simulation for Factorial Designs.
- Eisenhart, M., Towne, L., 2003. Contestation and change in national policy on “scientifically based” education research. *Educ. Res.* 32 (7), 31–38.
- Gelman, A., Hill, J., Yajima, M., 2012. Why we (usually) don't have to worry about multiple comparisons. *J. Res. Educ. Eff.* 5, 189–211.
- Hall, G.J., Schaefer, P., Hedges, T., Grodsky, E., 2021. Examining Bridges in Mathematics and differential effects among English language learners. *Sch. Psychol. Rev.* <https://doi.org/10.1080/2372966X.2020.1871304>.
- Hernán, M.A., Robins, J.M., 2016. Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidemiol.* 183 (8), 758–764.
- Huang, F., 2018. Using instrumental variable estimation to evaluate randomized experiments with imperfect compliance. *Practical Assess. Res. Eval.* 23 (2).
- Kassler, D., Nichols-Barrer, I., Finucane, M., 2020. Beyond “treatment versus control”: how Bayesian analysis makes factorial experiments feasible in education research. *Eval. Rev.* 44 (4), 238–261.
- Kratochwill, T.R., Levin, J.R., 2010. Enhancing the scientific credibility of single-case intervention research: randomization to the rescue. *Psychol. Methods* 15 (2), 124–144.
- Kratochwill, T.R., Hitchcock, J., Horner, R.H., Levin, J.R., Odom, S.L., Rindskopf, D.M., Shadish, W.R., 2010. Single-case Designs Technical Documentation. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. *J. Econ. Lit.* 48 (2), 281–355.

- McNeish, D., Stapleton, L.M., Silverman, R.D., 2017. On the unnecessary ubiquity of hierarchical linear modeling. *Psychol. Methods* 22 (1), 114–140. <https://doi.org/10.1037/met0000078>.
- Powell, M.G., Hull, D.M., Beaujean, A.A., 2020. Propensity score matching for education data: worked examples. *J. Exp. Educ.* 88 (1), 145–164.
- R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
- Shadish, W.R., Cook, T.D., Cambell, D.T., 2002. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton-Mifflin Publishers.
- Somers, M.A., Zhu, P., Jacob, R., Bloom, H., 2013. The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation. MDRC Working Paper on Research Methodology. MDRC.
- Spybrook, J., Shi, R., Kelcey, B., 2016. Progress in the past decade: an examination of the precision of cluster randomized trials funded by the U.S. Institute of Education Sciences. *Int. J. Res. Method Educ.* 39 (3), 255–267.
- Spybrook, J., Zhang, Q., Kelcey, B., Dong, N., 2020. Learning from cluster randomized trials in education: an assessment of the capacity of studies to determine *what works, for whom, and under what circumstances*. *Educ. Eval. Pol. Anal.* 42 (3), 354–374. <https://doi.org/10.3102/0162373720929018>.
- Steele, J.L., Slater, R.O., Zamarro, G., Li, T.M.J., Burkhauser, S., Bacon, M., 2017. Effects of dual-language immersion programs on student achievement: evidence from lottery data. *Am. Educ. Res. J.* 54 (1_Suppl.), 282S–306S.
- Umansky, I.L., 2016. To be or not to be EL: an examination of the impact of classifying students as English learners. *Educ. Eval. Pol. Anal.* 38 (4), 714–737.
- What Works Clearinghouse, 2020a. What Works Clearinghouse Standards Handbook. Version 4.1. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- What Works Clearinghouse, 2020b. What Works Clearinghouse Procedures Handbook. Version 4.1. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Sampling designs for educational research

Sharon L. Lohr, Arizona State University, Tempe, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Data collection from surveys	703
Probability and nonprobability samples	704
Probability sampling designs	705
Simple random sampling	705
Stratified sampling	705
Balanced sampling	705
Cluster sampling	706
Probability-proportional-to-size sampling	706
Systematic sampling	707
Stratified multistage sampling	707
Standard errors of estimates from probability samples	707
Survey design and nonresponse	708
Nonprobability samples	708
Decisions for survey designs	709
References	709

Glossary

Balanced sample A sample with the property that totals estimated from the sample for a set of variables equal the population totals for those variables

Cluster sample A sample in which each unit in the population belongs to a group (cluster), and the clusters are sampled according to the survey design

Nonprobability sample A sample that is drawn by means other than random selection. Samples that consist of units conveniently at hand are nonprobability samples

Nonresponse Failure of units selected for the sample to provide data on some or all survey items

Oversampling Sampling a population subgroup so that it comprises a higher percentage of the sample than of the population

Population The set of units for which estimates are desired

Probability-proportional-to-size sampling A sampling design in which the probability of selecting a sampling unit is proportional to the number of subunits it contains

Probability sampling A method of sample selection in which units are drawn using a random selection method and each possible subset of the population has a predetermined probability of being the sample that is drawn

Sampling weight The reciprocal of the probability with which a unit is selected to be in the sample. The sampling weight tells how many population units are represented by a unit in the sample

Simple random sample A sampling design in which every subset of n population units has the same probability of being selected as the sample

Stratified sample A sampling design in which the population units are partitioned into strata and a probability sample is selected from each stratum

Data collection from surveys

What percentage of fourth-grade children in Texas read at the level judged to be proficient by the state? What is the distribution of number of hours worked per week for public school teachers in a particular city? What percentage of elementary schools in a province employ a music teacher? How many students who take calculus in high school end up completing a college degree in mathematics, science, or engineering within six years of high school graduation? How satisfied are parents with their children's schools? What is the difference in literacy proficiency between adult men and adult women in a country?

These are all questions that may be addressed through taking a survey. While the research questions cover many different topics, their common feature is that they all are concerned with characteristics of a well-defined population at a specific period of time. The question asked is something that, given the operationalizations of the concepts, could be known exactly if every member of the population were accurately measured.

If the number of hours worked during the week of April 15 were ascertained accurately for every teacher in the city, then any statistic about number of hours worked per week would be known exactly for that population—the mean or median number of hours worked, the total number of hours worked for all teachers in the city, the percentage who work more than 40 hours per week, the difference in average working hours by gender and grade taught, and more.

The goal of the survey design is to obtain information from a sample that can be used to estimate the statistics that would be obtained if everyone in the population were measured, and to be able to assess the accuracy of the estimates.

The population is assumed to have N units, where N may be known or unknown. For the teacher workload survey, N would be the number of teachers in the city. For an assessment of reading proficiency among Texas fourth-graders, N would be the number of fourth-grade students. A sample of n units is to be taken from the N units in the population. The survey designer needs to decide what n should be, how units should be selected for the sample, and how information from the sample should be used to estimate characteristics of the population.

Probability and nonprobability samples

If the goal of taking a survey is to be able to describe features of the population of interest, the sample must be chosen in a way that allows the survey-taker to use the information from units in the sample to make inferences about units that are not in the sample.

Some samples consist of units that are easy to obtain or that are conveniently at hand. A sample that consists of persons who volunteer to participate in the survey (perhaps by clicking on a website), or a sample that consists of teachers who are personal acquaintances of the investigator, is a convenience sample. In a judgment sample, the investigator might deliberately select schools to be in the sample that are thought to be “typical” for the city.

Results from convenience or judgment samples are difficult to generalize to the population because no one knows whether statistics calculated from units in the sample reflect the characteristics of units not in the sample. Convenience samples often consist of the most accessible members of the population—for example, university students who pass by the library, or computer users who volunteer to take surveys—and it is unknown whether the sampled units are similar to less accessible units in the population. Judgment samples may reflect conscious or subconscious biases of the investigators. When samples consist of volunteers, a small group of persons with strong views can determine the survey results. Statistics calculated from convenience or judgment samples apply to the sample members; one must assume that the sampled units are like the nonsampled units in order to apply them to the population.

Probability sampling uses random selection to choose units to be in the sample, and thus eliminates many of the sources of bias that are associated with nonprobability samples, which do not employ random selection. In probability sampling, each subset of the population has a predetermined probability of being included in the sample. A randomizing mechanism such as a random number generator is used to select the sample according to those probabilities. Because the sample is selected randomly, the investigator need not worry that the results will be skewed from bias in selecting the sample.

Statistics from probability samples can be generalized to apply to the population without having to make assumptions about the unsampled units. In a probability sampling framework for inference, the probability distributions for statistics calculated from the sample are based on the distributions of the random variables that indicate which units are included in the sample. Suppose the population contains N units. Let π_i denote the probability that unit i is included in the sample, and let π_{ij} denote the probability that units i and j are both included in the sample. These probabilities give all the information that is needed to estimate population quantities and calculate the standard errors of the estimates. In a probability sample, it is assumed that every unit in the population can be selected for the sample (has $\pi_i > 0$).

The sampling weight of unit i in the sample S is $w_i = 1/\pi_i$, the reciprocal of the probability with which the unit is selected for the sample, and can be interpreted as the number of population units represented by a particular unit in the sample. Let y_i be a characteristic associated with unit i in the population. If a census were taken, y_i would be known for every unit in the population; when a sample S is taken, y_i is measured only for units in S .

The population total, $t_y = \sum_{i=1}^N y_i$, or population mean, $\bar{y}_p = t_y/N$, are often of special interest in a survey sample and are commonly estimated by

$$\hat{t}_y = \sum_{i \in S} w_i y_i \quad (1)$$

and

$$\hat{\bar{y}} = \frac{\sum_{i \in S} w_i y_i}{\sum_{i \in S} w_i} \quad (2)$$

Consider a survey in which 200 schools are randomly selected from a population of 2000 schools and y_i equals the number of music teachers in school i . With this sampling design, each school in the population has probability $\pi_i = 200/2000$ of being included in the sample. The weight of every school in the sample is $w_i = 1/\pi_i = 10$; each school in the sample represents itself

plus 9 other schools that were not selected for the sample. The weights for units in this sample sum to the population size, 2000. Here, calculating the estimated population total in [Eq. \(1\)](#) is equivalent to creating a new data set of 2000 pseudo-observations in which each sampled school's value of y_i is repeated w_i times, and then summing the number of music teachers in the new, replicated, data set. The sampling weights can be thought of as a way of using the sample to construct a pseudo-population whose characteristics are thought to resemble those of the original population.

Probability sampling designs

All probability sampling designs require that population units be selected randomly with the specified probabilities. Within that framework, though, many designs can be used. The design features described below can be used singly or in combination. For a probability sampling design with a sufficiently large sample, [Eqs. \(1\) and \(2\)](#) give approximately unbiased estimators of the population total and mean of a variable y . The variances of those estimators depend on the particular sampling design used. [Lohr \(2022\)](#) describes the theory of probability sampling and shows how to design and analyze surveys using the features described in this section. [Kalton \(2020\)](#) provides a concise, nontechnical summary of survey design and weight calculation.

Simple random sampling

In simple random sampling, each possible subset of the population of size n has an equal probability of being selected as the sample. Each unit in the population has sampling weight $w_i = N/n$. The example in the [Probability and nonprobability samples](#) section is a simple random sample of size 200.

Stratified sampling

For many education surveys, information is known about the population that can be used to make the survey design and estimation more efficient. In stratified sampling that information is used to divide the population into mutually exclusive groups called strata. The strata are often subpopulations of interest. A population of schools might be stratified by public or private status, urbanicity, and size; a population of students in a university might be stratified by gender, race, ethnicity, age, and field of study.

A stratified random sample is conducted by taking an independently selected simple random sample from each stratum. Suppose the population of 2000 schools contains 1500 public and 500 private schools. A proportionally allocated stratified random sample of size 200 would have 150 public and 50 private schools; each sampled school has weight 10, just as in simple random sampling. The difference is that the number of private schools in a simple random sample varies from sample to sample, while the stratified random sample is forced to have exactly 50 private schools. This forces the proportionally allocated sample to mirror the population with respect to the stratification variable of public/private status. If variable y is associated with the stratification variables, its distribution is likely to be more similar to the population distribution as well, and the standard error of $\hat{y}$ from a proportionally allocated stratified random sample of size n will usually be smaller than the standard error of the sample mean from a simple random sample of the same size.

Alternatively, a disproportional allocation may be desired in order to obtain a prespecified precision for certain strata. For example, one could sample 100 public and 100 private schools, so that the weight for each sampled public school is 15 and the weight for each sampled private school is 5. In this disproportional design, private schools are oversampled—that is, the proportion of private schools in the sample is greater than their proportion in the population of schools—and the weights must be used to obtain estimates that are unbiased for population quantities. The sampling weights are used in estimation to compensate for the oversampling. In this example, the sum of the weights for the sampled private schools is 500 and the sum of the weights for the sampled public schools is 1500. Even though the sample has equal numbers of schools of each type, the weighting brings them back to their proportions in the population.

In general, the more strata that can be formed, the better. A stratified sample with proportional allocation almost always provides more information per unit sampled than a simple random sample of the same size. Many surveys are stratified to the degree that 2 units are selected per stratum.

Balanced sampling

Sometimes more information is known about the population than can be accommodated in a stratified sampling design, or it is desired to have a sample that mirrors the population with respect to numerous variables. In the United States, for example, the College Scorecard system ([US Department of Education, 2021](#)) contains information on more than 1500 variables for each post-secondary education institution in the database, including information about number and demographic composition of undergraduate and graduate students, types of degrees offered, admission rates, retention and graduation rates, median student debt, tuition fees, early career earnings of graduates, and a wealth of other topics.

This information could be used to form strata for sampling. But stratification requires that each population unit belongs to exactly one stratum. Thus, a stratification by 5 categories of institution size, 3 categories of highest degree offered, public or private status, and high or low tuition results in 60 strata that are formed by the cross-classification of the stratification variables.

Balanced sampling (Tillé, 2011) allows a probability sample to be selected that reflects the distribution of each balancing variable individually (sometimes at the expense of the cross-classifications), and may be preferred to simple stratification when detailed information is available about the population before sampling. A balanced sample is selected so that

$$\hat{t}_x = \sum_{i \in S} w_i x_i = t_x \quad (3)$$

for each balancing variable x . Thus, a sample that is balanced by institution size, highest degree, public/private status, and high/low tuition will contain the same proportion in each category of those variables as the population. Balanced sampling has the additional advantage that the balancing variables need not be categorical—for example, if the admission rate is a balancing variable, then the sample will have the property that the estimated mean admission rate from the colleges in the sample will equal the mean admission rate of all colleges in the population. If an exact balanced sample cannot be achieved, Eq. (3) may be modified to require that $\hat{t}_x \approx t_x$.

A stratified sample is a special case of a balanced sample when the x variables are indicators for the cross-classified categories defined by the stratification variables. Balanced sampling is thus a generalization of stratified sampling that guarantees that the sample will mirror the population for each individual x variable considered for balancing. By omitting the cross-classifications from the set of balancing variables, a balanced sample can match the population on more characteristics than a stratified sample.

Cluster sampling

In simple and stratified random sampling, the individual is the unit that is sampled. Thus, to draw a stratified random sample of students, one would first construct a list of all students, divide the students on the list into strata, and then draw a simple random sample from each stratum. This may be feasible for sampling students within a single school or university, but may be impractical for drawing a sample from the set of all students in a province or country. Typically, in these situations, a cluster sampling design is used. A stratified sample of schools (clusters) is selected, and then all or a subsample of students are sampled from each selected school. Here, a school is the primary sampling unit (psu) that is selected at the first stage of sampling, and the student is the secondary sampling unit (ssu) that is selected at the second stage of sampling.

Students in the same school often share common environmental factors. They may live in the same area, learn from the same teachers, or have similar socioeconomic backgrounds. It would thus be expected that students in the same school would have more similar values of y than students selected randomly from the population of all schools. Thus, selecting a sample of n students through a cluster sample typically does not provide as much information as a simple random or stratified random sample of size n .

In a one-stage cluster sample, all ssus within a sampled psu are sampled. In a survey where the survey instrument is administered to every student within a sampled school (for example, if each student is given a paper or computer-administered assessment), a one-stage sample makes sense because there is little additional cost to sample all, rather than some, students in the school. For a more labor-intensive assessment (for example, if sampled students are given a medical examination or interviewed in person), it may be more cost-effective to use a two-stage cluster sample, in which a subsample of ssus is taken from each psu.

The sampling weights are determined by the probabilities of inclusion at each stage. In a two-stage cluster sample with schools as psus and students as ssus, the weight for student i is $w_i = 1/[P(\text{student } i\text{'s school is included in the sample}) \times P(\text{student } i \text{ is included in the sample of students from that school})]$.

While a cluster sample typically has less precision relative to a simple random sample of the same number of students, it is commonly used for education surveys because schools or classrooms form a natural sampling unit. Much of the expense for the survey comes from securing cooperation of sampled schools and ensuring that the instrument is administered at the school according to the survey protocols; in some cases, field investigators may need to travel to the school to work with school personnel or administer the instrument. Thus, a cluster sample is often less expensive or easier to conduct than a design that involves sampling students directly (which may not even be possible in the absence of a list of all students). A cluster sample of schools has the additional advantage that school-level or classroom-level effects and similarities can be studied.

The precision of estimates from a cluster sample depends primarily on the number of psus that are sampled. To obtain greater precision for estimates from a cluster sample in which students are subsampled from schools, it is generally necessary to sample additional schools as opposed to subsampling more students within a fixed number of sampled schools.

Probability-proportional-to-size sampling

The psus used in cluster sampling are often groupings that occur naturally such as geographic regions, communities, school districts, schools, or classrooms. The psus may vary greatly by size—some schools may have 3000 students while other schools may have 100 students. An equal-probability sample of schools gives a large school the same probability of selection as a tiny school, even though the large school contributes much more to the population total.

Instead of selecting psus with equal probabilities, many survey designs call for selecting them with probabilities proportional to their sizes. This selection is often done after the psus are grouped into strata to improve precision and ensure that large psus have high probability of being included in the sample; some very large psus may be selected with probability one so that they are guaranteed to be in the sample.

If schools are sampled with probability proportional to size and then the same number of students is selected from each sampled school, then each student in the population has an equal probability of being selected for the sample. This design has many advantages in practice. The large schools have higher probability of being represented in the sample, yet the workload of conducting interviews or administering assessments is approximately the same for each sampled school. If population totals of interest are roughly proportional to the size of the psu (for example, larger high schools might be expected to have a larger number of students who enroll in college than smaller high schools), then the probability-proportional-to-size sampling design will also produce estimates that have greater precision. If the number of psus in population is unknown before sampling, an estimate or a related measure of size may be used to assign the probabilities of selection.

Systematic sampling

In systematic sampling, every k th unit is selected from a list of population units. The starting point is chosen randomly. If the list is in random order, a systematic sample will behave much the same way as a simple random sample. Systematic sampling may produce biased estimates if the population units are arranged in periodic order: if, for example, men and women alternate in the population list and $k = 10$, the systematic sample will consist of either all men or all women.

Systematic sampling, however, can produce estimates with higher precision than simple random sampling if the population is sorted before sampling according to characteristics thought to be related to quantities of interest. It is often used for selecting students or teachers in the second or third stage of a cluster sample. Suppose that in a cluster sample of teachers, schools are selected with probability proportional to the number of students in the first stage of sampling. Within each sampled school, the teachers are listed in decreasing order of years of experience. Then a systematic sample of teachers within the school is guaranteed to contain teachers with a range of values for years of experience.

Stratified multistage sampling

Many large education surveys combine stratification with several stages of clustering. For a national survey, the first stage of sampling might select geographic areas of the country, the second stage might select school districts in the sampled areas, the third stage might select schools within districts, the fourth stage might select classrooms within schools, and the fifth stage might select students within the selected classrooms.

As an example, consider the sampling design structure for the 2019 Trends in International Mathematics and Science Study (TIMSS), a survey intended to provide information on student knowledge and skills in mathematics and science in countries around the world, with emphasis on measuring trends over time. Each participating country devises a design for sampling fourth-grade and/or eighth-grade students in the country, following the basic two-stage stratified sampling design described by LaRoche et al. (2020).

The first step for the basic TIMSS design is to construct or obtain a list of all schools containing fourth- or eighth-grade students. The schools may be placed into strata according to country-specific variables such as state or province, degree of urbanization, language of instruction, socioeconomic variables, or school scores on national examinations. Then schools are sampled within strata with probability proportional to size.

The stratification improves the precision of estimators and also allows specific types of schools to be oversampled if desired. In some countries, the list of schools is sorted by the stratification variables and a systematic sample of schools is selected from the sorted list, which forces an implicit stratification.

In the second stage of sampling, a fixed number of classes (usually one or two) is randomly selected from the set of classes for the target grade at each sampled school. All students in the sampled classes then participate in the assessment. Selecting a fixed number of classes in the sampled schools largely offsets the unequal probabilities of selection at the first stage and results in a sample of students in which each student has a similar probability of being selected for the sample and hence a similar sampling weight.

Standard errors of estimates from probability samples

The design features discussed above affect the precision of estimates from surveys. Estimates from a stratified sample with proportional allocation usually have smaller standard errors than estimates from a simple random sample of the same size; estimates from a cluster sample usually have larger standard errors than estimates from a simple random sample of the same size.

Standard errors can be calculated in multiple ways. Formulas for calculating standard errors of statistics from stratified and cluster samples, which depend on the inclusion probabilities π_i and π_{ij} , are given in Lohr (2022) and other sampling books. Standard errors for statistics such as regression coefficients can be calculated using linearization (Taylor series) methods.

Many surveys that employ a stratified multistage design use replicate weight methods to calculate standard errors. With these methods, which include jackknife, bootstrap, and balanced repeated replication, the variability among estimates calculated with

the replicate weight variables is used to compute standard errors of estimates. Valliant et al. (2018) give step-by-step instructions for constructing and using replicate weights.

Survey design and nonresponse

The sampling designs and estimators in the **Probability and nonprobability samples** and **Probability sampling designs** sections are constructed assuming that (1) every unit in the population has a positive probability of being selected for the sample (for example, all schools in the country are included in the list of schools from which the sample is drawn), (2) every unit selected for the sample provides data (there is no nonresponse), and (3) every unit provides accurate answers to the survey instrument (there is no measurement error and no missing data for items on the instrument).

Few surveys achieve this ideal. Survey samplers have a number of tools that can attempt to compensate for nonresponse and potential bias arising from nonresponse. These include adjusting the sampling weights of respondents so that they account for similar nonrespondents' share of the population as well as their own (weighting and calibration; see Särndal and Lundström, 2005) or substituting predicted values for the missing data (imputation; see van Buuren, 2018).

The best way to deal with nonresponse and measurement error is to try to anticipate and prevent them while designing the survey. Measurement error can be reduced through designing survey instruments that accurately and reliably measure the concepts of interest (Saris and Gallhofer, 2014). The choice of how to administer the survey can also affect measurement error: for sensitive topics such as cheating or drug use, a self-administered survey may elicit more accurate responses, while in other situations responses may be more accurate when the survey is administered by an interviewer who can explain the survey items (Dillman et al., 2014).

Potential nonresponse, too, should be considered during the survey design. A pilot study or previous surveys can indicate the anticipated amount of nonresponse. Designed experiments can be conducted to explore methods for improving response rates or for reducing nonresponse bias or measurement error. It may be desired to oversample strata that are anticipated to have low response rates.

Nonprobability samples

Under ideal circumstances, where every unit in the population has a positive probability of being selected for the sample and there is no nonresponse or measurement error, a probability sampling design is guaranteed to give estimates that are approximately unbiased and that have known precision. In some situations, though, it may be impossible or overly expensive to conduct a high-quality probability sample, or a nonprobability sample may better meet the research needs.

Convenience or judgment samples may provide useful information in some situations. Many psychology experiments, for example, are conducted on convenience samples of students enrolled in Psychology 101 classes. But to generalize beyond the students participating in the experiment, the investigator must assume that the experimental results observed for the college students will apply more generally. This assumption can often be tested only by repeating the experiment on other types of participants.

Sometimes probability samples are infeasible when it is desired to take a survey of a subpopulation that is relatively uncommon in the larger population, such as students who are recent immigrants or teachers who are military veterans. A probability sample of schools might yield relatively few students or teachers in the population of interest. But a respondent-driven sample, a nonprobability sample in which an initially recruited set of persons in the desired subpopulation then recruits other persons in the subpopulation to be in the sample, may yield a large sample of persons in the subpopulation of interest, even though it is not guaranteed to be representative or produce unbiased estimates. Gile et al. (2018) discuss how to calculate estimates from these types of nonprobability samples and assess their validity.

The main concern with nonprobability samples is that some parts of the population may be systematically excluded from the sample and that estimates calculated from the nonprobability samples may thus be biased. This may occur even if the nonprobability sample is extremely large. The mean squared error (MSE) for the sample mean $\bar{y} = \sum_{i \in S} y_i / n$ can be expressed as the product of three factors (Meng, 2018):

$$MSE(\bar{y}) = E[(\bar{y} - \bar{y}_p)^2] = E[\{Corr(R, y)\}^2] \times \frac{N-1}{N} \left(\frac{N}{n} - 1 \right) \times \sigma^2 \quad (4)$$

where $Corr(R, y)$ is the correlation between the variable of interest y_i and the variable R_i that indicates participation in the sample ($R_i = 1$ if unit i is in the sample and 0 otherwise), and σ^2 is the population variance of y . From Eq. (4), the MSE is low if there is little association between the value of y and being in the sample, if the sample accounts for a large part of the population (the second factor is 0 if $n = N$), or if the values of y in the population are all similar to each other (σ^2 is small).

For a simple random sample, the expected squared correlation between R and y is $1/(N-1)$, and it is typically of similar magnitude for other types of probability samples. Thus, in a simple random sample, $MSE(\bar{y}) \leq \sigma^2/n$; the precision of the estimate increases when a larger sample size n is taken—regardless of the population size.

In a nonprobability sample, however, even small correlations between y and participating in the sample can result in substantial bias that cannot be mitigated by taking a larger sample. A correlation of 0.05 is often considered to be small in the social sciences. But if the expected correlation between y and R is of that magnitude, so that $E[\{\text{Corr}(R, y)\}^2] \geq 0.0025$, then a sample of half the population has larger MSE for estimating the population mean than a simple random sample of size 400. Thus, with a modest correlation of 0.05 between y and the sample selection, a convenience sample of 500 million people from a population of one billion people provides less accuracy than a simple random sample of size 400 from the same population.

The accuracy of estimates from a nonprobability sample thus depends on the degree of correlation between the variables of interest and the sample selection mechanism. Tourangeau et al. (2013), summarizing results of studies about volunteer web surveys, found that even after weighting adjustments there were often substantial biases that could shift the estimates by 20% or more. On the other hand, if inclusion in the nonprobability sample is not related to the variables of interest, the nonprobability sample may produce valuable information at much less cost than a probability sample. The challenge for nonprobability samples is that, unlike in a probability sample with full response, the correlations between inclusion in a nonprobability sample and variables of interest—and hence the biases—are unknown.

Decisions for survey designs

The best survey design is one that fits the needs of the research. Sometimes it is desired to have highly accurate statistics that describe the entirety of the population. In that case, it may be desired to take a high-quality probability sample with adequate resources to follow up with nonrespondents. In other situations, particularly in preliminary stages of research, a nonprobability sample may provide the type of information needed to inform further investigations.

A number of decisions need to be made when designing a survey. The first involves whether a survey should be taken at all. It may be possible to answer the research questions (or modifications of them) with existing survey data or by combining data from multiple sources (National Academies of Sciences, Engineering, and Medicine, 2017). Flawed statistics from a convenience sample with high bias, or from a poorly conducted, high-nonresponse probability survey (which may have the same types of biases as the nonprobability samples discussed in the Nonprobability samples section) may result in more harm than having no statistics.

If it is determined that a survey is the best method for answering the research questions, the sample should be designed to give estimates with desired precision for the quantities to be studied, for the population as a whole and for subpopulations of interest. Information that is already known about the population can be used for stratifying or balancing the sample, or for oversampling certain subpopulations. If clustering is to be employed, simulation studies can provide information for evaluating candidate designs. Lohr (2022) describes methods for optimizing and evaluating prospective survey designs.

If the survey is to be repeated at different points in time, a repeated cross-sectional (independent samples are taken at each time period), longitudinal (the persons in the initial sample are followed over time), or panel (cohort samples are taken at different start times and followed for multiple interviews) design may be used. All of these designs allow the investigator to study changes in the mean levels of variables over time (for example, the average score in year 2 can be compared with the average score in year 1). Longitudinal and panel studies may be more complicated to conduct, but they also allow the researcher to study intra-personal changes over time (for example, the trajectory of tobacco product use in individual high school students).

Finally, the survey design should anticipate possible problems that may arise during implementation. Pre-survey experimentation and testing can help identify some of these problems. Sufficient resources should be set aside for following up with nonrespondents. The survey-taker should also, if possible, collect information in the survey that can be used later to adjust for nonresponse and evaluate nonresponse bias.

References

- Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, fourth ed. John Wiley & Sons, Hoboken, NJ.
- Gile, K.J., Beaudry, I.S., Handcock, M.S., Ott, M.Q., 2018. Methods for inference from respondent-driven sampling data. *Annu. Rev. Stat. Appl.* 5, 65–93.
- Kalton, G., 2020. *Introduction to Survey Sampling*, second ed. Sage, Thousand Oaks, CA.
- LaRoche, S., Joncas, M., Foy, P., 2020. Sample design in TIMSS 2019. In: Martin, M.O., von Davier, M., Mullis, I.V.S. (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report*. International Association for the Evaluation of Educational Achievement, Boston, pp. 3.1–3.33.
- Lohr, S.L., 2022. *Sampling: Design and Analysis*, third ed. CRC Press, Boca Raton, FL.
- Meng, X.L., 2018. Statistical paradises and paradoxes in big data (I): law of large populations, big data paradox, and the 2016 US Presidential Election. *Ann. Appl. Stat.* 12, 685–726.
- National Academies of Sciences, Engineering, and Medicine, 2017. *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*. National Academies Press, Washington, DC.
- Saris, W.E., Gallhofer, I.N., 2014. *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. John Wiley & Sons, Hoboken, NJ.
- Särndal, C.E., Lundström, S., 2005. *Estimation in Surveys with Nonresponse*. John Wiley & Sons, Hoboken, NJ.
- Tillé, Y., 2011. Ten years of balanced sampling with the cube method: an appraisal. *Surv. Methodol.* 37, 215–226.
- Tourangeau, R., Conrad, F.G., Couper, M.P., 2013. *The Science of Web Surveys*. Oxford University Press, New York.
- US Department of Education, 2021. *Technical Documentation: College Scorecard Institution-Level Data*. U.S. Department of Education, Washington, DC.
- Valliant, R., Dever, J.A., Kreuter, F., 2018. *Practical Tools for Designing and Weighting Survey Samples*, second ed. Springer, Cham, Switzerland.
- Van Buuren, S., 2018. *Flexible Imputation of Missing Data*, second ed. CRC Press, Boca Raton, FL.

Continuous probability distributions

Sandip Sinharay, ETS, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Uniform distribution	710
Exponential distribution	711
Normal distribution	711
Lognormal distribution	711
Student's <i>t</i> distribution	712
Beta distribution	712
Gamma distribution	713
Cauchy distribution	714
χ^2 distribution	714
<i>F</i> distribution	715
Dirichlet distribution	715
Wishart distribution	715
Conclusions	716
References	717

Continuous probability distributions are distributions that are used to describe continuous random variables. The values of variables observed in statistical applications are mostly discrete, often due to the use of rounded units. For example, the weight of a person is a continuous variable, but is typically measured in the nearest pounds or kilograms. Therefore, strictly speaking, continuous distributions may not be appropriate tools in statistical applications. That is one reason [Holland \(1979\)](#) was surprised by the tyranny of continuous models in a world full of discrete data. However, continuous distributions are convenient approximations of the reality, facilitating mathematical and statistical analysis ([Johnson et al., 1994](#)). For example, think of a random variable X that follows the binomial distribution (see the article on discrete probability distributions by Sinharay in this encyclopedia) with number of trials = 30 and success probability = 0.4. To compute the probability that X is less than or equal to 10 using binomial probability values, one has to sum 11 terms of the form $\binom{30}{x} 0.4^x 0.6^{30-x}$. However, if the investigator approximates the binomial distribution (a discrete distribution) by the normal distribution (a continuous distribution—described later in this article), the computation is much simpler as shown below

$$P(X \leq 10) = P(X \leq 10.5) = P\left(\frac{X - 30 \times 0.4}{\sqrt{30 \times 0.4 \times 0.6}} \leq \frac{10.5 - 30 \times 0.4}{\sqrt{30 \times 0.4 \times 0.6}}\right) = \Phi(-0.56)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function (cdf) of the standard normal distribution and the addition of 0.5 to 10 denotes a “continuity correction”. Such approximations are satisfactory in several statistical applications. For example, while the probability $P(X \leq 10)$, computed from the binomial distribution, is 0.2917, the aforementioned normal approximation provides a value of 0.2881—so the approximation is quite accurate.

A key issue with a continuous random variable is that the probability is zero that it will take any specific value. However, there is usually a positive probability that such a variable will fall within an interval. While a discrete random variable has a probability mass function, a continuous random variable has a probability density function (pdf). The integral of the pdf over an interval is the probability that the continuous random variable will fall in that interval; The probability is known as the cdf of the variable when the interval includes all values smaller than or equal to a given value x .

In the following sections, several continuous distributions are discussed. For each distribution, its genesis, pdf, cdf, moments, estimation of the moments, methods to generate values of a random variable following that distribution, and relationships to other distributions (if any) are discussed.

Uniform distribution

The uniform or rectangular distribution with range (a, b) , often denoted as $U(a, b)$, has the pdf $\frac{1}{b-a}$ if $a \leq x \leq b$ and 0 otherwise. The cdf is 0 if $x \leq a$, $\frac{x-a}{b-a}$ if $a \leq x \leq b$ and 1 if $x \geq b$ otherwise. When $a = 0$ and $b = 1$, the uniform distribution becomes the standard or unit uniform distribution.

The mean of the distribution is $\frac{a+b}{2}$ and the variance is $\frac{(b-a)^2}{12}$.

For any continuous random variable Y , its cdf $F(y)$ follows a $U(0,1)$ distribution and $F^{-1}(R)$ has the cdf F for R following the $U(0,1)$ distribution. This property is used to generate random numbers from other continuous distributions starting from standard uniform random numbers (e.g., [Evans et al., 2000](#)).

The method of moments estimates of a and b are $\bar{x} - 3^{1/2}s$ and $\bar{x} + 3^{1/2}s$ respectively where $\bar{x}$ is the sample mean and s is the sample standard deviation. The maximum likelihood estimates of a and b are the minimum and maximum values, respectively, over the sample. To generate a random number x following a $U(a,b)$ distribution, one has to generate a $U(0,1)$ random number R , and then use the transformation $x = a + (b - a)R$.

The uniform distribution has found several applications as the prior distribution of the parameters of IRT models in Bayesian analyses. For example, in Bayesian IRT analyses, [Karadavut \(2019\)](#) and [Gao and Chen \(2005\)](#) assumed the prior distribution of both the examinee ability parameter and item difficulty parameter to be the uniform distribution between -3 and 3 .

Exponential distribution

The exponential distribution has been employed to model variables such as the time to decay of radioactive atoms and time to failure of components with constant failure rates.

The pdf of the exponential distribution is given by

$$f(x; b) = \frac{1}{b} e^{-x/b}, 0 \leq x \leq \infty$$

The mean of the distribution is b and the variance is b^2 . The maximum likelihood estimate of b is the sample mean.

A random number from the exponential distribution can be generated by generating a $U(0,1)$ random number R and then computing $-b \log(R)$.

[Scheiblechner \(1979\)](#) suggested the *latent exponential model* for modeling the response times of examinees on test items, where the mean parameter of the distribution is the sum of an examinee speed parameter and an item speed parameter.

Normal distribution

The uniform or rectangular distribution is arguably the most frequently used continuous distribution in statistics and in educational measurement. The normal distribution with mean μ and variance σ^2 , often denoted as $\mathcal{N}(\mu, \sigma^2)$, has the pdf

$$\phi(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ if } -\infty \leq x \leq \infty$$

The cdf is given by

$$\Phi\left(\frac{x-\mu}{\sigma}\right) = \int_{t=-\infty}^x \phi(t; \mu, \sigma^2) dt$$

The standard normal distribution is a special case of the normal distribution with $\mu = 0$ and $\sigma = 1$. The cdf of the standard normal distribution is given by

$$\Phi(x) = \int_{t=-\infty}^x \phi(t; 0, 1) dt = \int_{t=-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Given $x_1, x_2, \dots, x_n$, which is a random sample of observations of size n from a normal distribution with mean μ and variance σ^2 , the method of moments estimates as well as the maximum likelihood estimates of μ and σ^2 are the sample mean $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ and sample variance $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, respectively. There exist numerous approaches for generating random draws from the normal distribution. A popular approach, that suggested by [Box and Muller \(1958\)](#), involves first simulating two random numbers X_1 and X_2 from a uniform distribution between 0 and 1 and then applying the transformations $Z_1 = \sqrt{-2\log(X_1)}\cos(2\pi x_2)$ and $Z_2 = \sqrt{-2\log(X_1)}\sin(2\pi x_2)$; Z_1 and Z_2 would be independent and follow the normal distribution with mean 0 and variance 1.

The normal distribution has found several applications in education. For example, in applications of item response theory (IRT) (see, for example, articles by Ackerman and Cai in this encyclopedia and [Hambleton and Swaminathan, 1985](#)), the examinee ability distribution is typically assumed to follow the normal distribution.

Lognormal distribution

If a random variable V has a normal distribution with mean μ and variance σ^2 , then e^V has a lognormal distribution with parameters μ and σ^2 . In other words, if a variable has a lognormal distribution, then its logarithm has a normal distribution. The pdf of the distribution is given by

$$f(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log(x) - \mu}{\sigma}\right)^2}$$

where x and σ are both positive. If X follows a lognormal distribution with parameters μ and σ^2 , then $Y = e^a X^b$ follows a lognormal distribution with parameters $a + b\mu$ and $b^2\sigma^2$.

The expectation and variance of the lognormal distribution are given by $e^{\mu + \frac{\sigma^2}{2}}$ and $e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$, respectively.

Using its relationships to the normal distribution, the parameters of the lognormal distribution from a sample $x_1, x_2, \dots, x_n$ of draws from the distribution can be estimated as

$$\hat{\mu} = \frac{1}{n} \sum \log(x_i), \hat{\sigma}^2 = \frac{1}{n-1} \sum (\log x_i - \hat{\mu})^2$$

The easiest way to generate random numbers from a log-normal distribution with parameters μ and σ^2 is to generate random numbers from a normal distribution with mean μ and variance σ^2 and then exponentiate them.

In some applications of Bayesian methods to IRT models, the prior distribution on the slope parameters is sometimes assumed to be a lognormal distribution. For example, the PARSCALE software program (e.g., see [du Toit, 2003](#)), which is used to fit IRT models by several operational testing programs, assumes a lognormal distribution as the prior distribution for the slope parameters. In addition, a popular model for response times (e.g., [van der Linden, 2006](#)) is based on the lognormal distribution.

Student's t distribution

The t distribution is often applied in testing hypotheses for normally distributed data. For example, for testing the hypothesis that the mean of a normal distribution is 0, the standard test statistic follows a central t distribution under the null hypothesis.

The pdf of the distribution is given by

$$\frac{\nu^{1/2} e^{-\delta^2/2}}{\Gamma(\nu/2) \pi^{0.5} (\nu + x^2)^{(\nu+1)/2}} \sum_{i=0}^{\infty} \Gamma\left(\frac{\nu+i+1}{2}\right) \frac{(x\delta)^i}{i!} \left(\frac{2}{\nu+x^2}\right)^{i/2}$$

where ν is the degree of freedom (df) and δ is the non-centrality parameter.

The mean of the distribution is $\frac{\delta(\nu/2)^{1/2} \Gamma((\nu-1)/2)}{\Gamma(\nu/2)}$ and the variance is given by

$$\frac{\nu}{\nu-2} (1 + \delta^2) - \frac{\nu}{2} \delta^2 \left(\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \right)^2$$

If the non-centrality parameter is 0, then the distribution is called a central t distribution that has mean 0 and variance $\frac{\nu}{\nu-2}$.

The t distribution has a much heavier tail than the normal distribution for small df. [Fig. 1](#) shows a comparison of the pdf of the Student's t distribution with 4 df against that of the standard normal distribution. The tail of the t distribution is heavier than that of the normal distribution.

For large df's, the central t distribution becomes close to the standard normal distribution. If X and Y are independent random variables and they follow the standard normal distribution and the χ^2 distribution with ν df, respectively, then $\frac{X}{\sqrt{Y/\nu}}$ follows the t distribution with ν df. The distribution of the square of a random variable that follows a central t distribution with ν df multiplied by ν is an F distribution with df's 1 and ν .

To generate a random number following the central t distribution with ν df, one can compute the ratio of a standard normal random number and the square root of Y/ν , where Y is a random number following the χ^2 distribution with ν df and both of these random numbers are generated independently of each other.

The t statistic, whose distribution under null hypothesis is the t distribution, has been widely used in education (e.g., [Zimmerman, 1997](#)). Other types of application of the distribution to education include [Fox et al. \(2014\)](#) where, in a Bayesian IRT approach, the prior distribution of the item parameters was assumed to be a t distribution.

Beta distribution

The beta distribution is used to model continuous random variables whose range is between 0 and 1. For example, in Bayesian analyses, the beta distribution is often used as a prior distribution of the parameter p (which is bounded between 0 and 1) of the binomial distribution. See, for example, [Novick and Jackson \(1974\)](#).

The pdf of the beta distribution is given by

$$f(x; a, b) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}$$

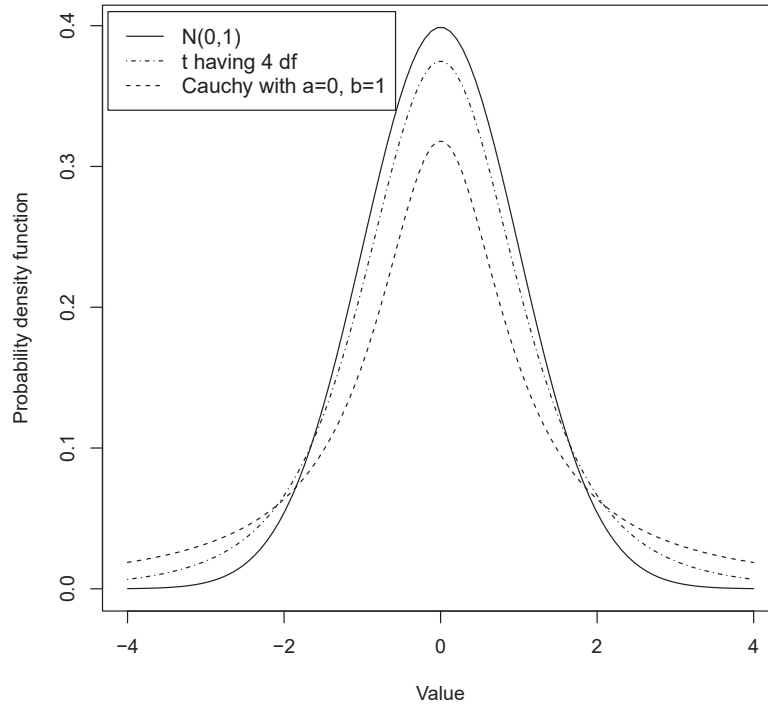


Fig. 1 Plot of the pdfs of the standard normal, Cauchy and the Student's t distribution with 4 df.

where $0 \leq x \leq 1$, $a > 0$ and $b > 0$ are called shape parameters, and $B(a, b)$ is the beta function defined as

$$B(a, b) = \int_{x=0}^1 x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

where $\Gamma(a)$ is the gamma function defined as $\Gamma(a) = \int_{t=0}^{\infty} t^{a-1} e^{-t} dt$. Note that when $a = b = 1$, the Beta distribution becomes identical to the uniform distribution between 0 and 1.

The expectation of the beta distribution is $\frac{a}{a+b}$ and the variance is $\frac{ab}{(a+b)^2(a+b+1)}$. The method-of-moments estimates of the parameters a and b can be obtained by setting the above-mentioned mean and variance equal to the sample mean $\bar{x}$ and sample variance s^2 and solving them for a and b . The solutions in this case are defined only when $\bar{x}(1-\bar{x}) > s^2$ and are given by

$$\hat{a} = \bar{x} \left[\bar{x} (1 - \bar{x}) / s^2 - 1 \right], \hat{b} = (1 - \bar{x}) \left[\bar{x} (1 - \bar{x}) / s^2 - 1 \right]$$

It is also possible to estimate a and b using maximum likelihood, but the method requires an iterative procedure such as the Newton-Raphson algorithm.

Random numbers following the beta distribution can be generated in several ways. If V_1 and V_2 are independent and follow the χ^2 distribution (the section on the χ^2 distribution that follows will discuss generating random numbers following that distribution) with parameters c_1 and c_2 respectively, then $\frac{V_1}{V_1 + V_2}$ follows a beta distribution with parameters $c_1/2$ and $c_2/2$. Alternatively, if $Y_1, Y_2, \dots, Y_n$ are independent random variables following the $U(0, 1)$ (that is, the continuous uniform distribution between 0 and 1) distribution, then the s -th order statistic Y'_s follows a beta distribution with parameters s and $n - s + 1$.

The beta distribution, as described above, is often referred to as the two-parameter beta distribution. An extension of the two-parameter beta distribution is the four-parameter beta distribution whose pdf is proportional to $(x - \alpha)^{a-1} (\beta - x)^{b-1}$, where $\alpha < x < \beta$; the variable $\frac{x - \alpha}{\beta - \alpha}$ has a two-parameter beta distribution in this case.

A well-known application of the beta distribution in education can be found in Lord (1965), where the true test score was modeled using the four-parameter beta distribution. In addition, as in Patz and Junker (1999), in applications of the Markov chain Monte Carlo (MCMC) algorithm models, the prior distribution for the guessing parameter is often assumed to be the beta distribution.

Gamma distribution

The gamma distribution is used to model continuous random variables whose range is between 0 and ∞ . The pdf of the gamma distribution is given by

$$f(x; b, c) = \frac{1}{b\Gamma(c)} \left(\frac{x}{b}\right)^{c-1} e^{-\frac{x}{b}}, b \geq 0, c > 0$$

where b is called the scale parameter and c is the shape parameter. In many applications of the distribution, b is assumed to be equal to 1. The mean of the gamma distribution is bc and the variance is b^2c .

If $2c$ is an integer, then the gamma random variable is half of a χ^2 random variable with $2c$ df. If $c = 1$, then the gamma distribution becomes the exponential distribution. If V_1 and V_2 are gamma random variables, with $b = 1$ and $c = c_1$ and c_2 respectively, then $\frac{V_1}{V_1 + V_2}$ follows a beta distribution with parameters c_1 and c_2 .

By the method of moments, the estimate of b is $s^2/\bar{x}$ and that of c is $(\bar{x}/s)^2$. It is possible to estimate the parameters using the maximum likelihood method, but that requires the use of an iterative algorithm.

If c is an integer, then $-b \prod_{i=1}^c R_i$ follows a gamma distribution with parameters b and c , where the R_i 's are independent $U(0,1)$ random variables—this result can be used to generate random numbers following the gamma distribution.

The reciprocal of a gamma random variable follows a distribution called the inverse gamma distribution. In several applications of Bayesian methods to IRT models, the prior distribution for the variance parameters is often assumed to be the inverse gamma distribution. See, for example, Bradlow et al. (1999).

Cauchy distribution

The Cauchy distribution is often of mathematical interest due to the absence of any moments (Evans et al., 2000). The pdf of the distribution is given by

$$f(x; a, b) = \left\{ \pi b \left[1 + \left(\frac{x-a}{b} \right)^2 \right] \right\}^{-1}$$

and the cumulative distribution function (cdf) is given by

$$\frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{x-a}{b} \right)$$

The mode and the median of the distribution is a , but no moments exist. The Cauchy distribution is unimodal and symmetric but has tails that are much heavier than those of the normal distribution. Fig. 1 compares the pdf of the Cauchy distribution with $a = 0$ and $b = 1$ (that is a special case of the Student's t distribution with 1 df) against that of a t distribution with 4 df and the standard normal distribution. The Cauchy distribution has the heaviest tail among the three distributions.

The ratio of two independent standard normal random variables follows a Cauchy distribution with $a = 0$ and $b = 1$. A Cauchy distribution with $a = 0$ and $b = 1$ is a special case of the Student's t distribution with one degree of freedom (df).

One set of estimates of a and b are the sample median and half of the sample interquartile range, respectively. Alternatively, one can compute maximum likelihood estimates of a and b using an iterative approach.

A random number following the Cauchy distribution can be generated by generating a random number R from the $U(0,1)$ distribution and then setting $F(x) = R$, where $F(x)$ is the cdf of the distribution. The solution is given by $a + b \tan[\pi(R - 0.5)]$.

Verhagen et al. (2016) assumed, in a Bayesian analysis using item response theory (IRT) models, that the prior distribution on the item difficulty parameters is a Cauchy distribution; the reason of this choice was that they wanted a prior distribution that has a wider tail than a normal distribution.

χ^2 distribution

The χ^2 distribution is the distribution of the sum of squares of several independent normal random variables with unit variance. A χ^2 random variable with d df and non-centrality parameter $\delta \geq 0$ has the pdf

$$f(x; d, \delta) = \frac{e^{-(x+\delta)/2}}{2^{d/2}} \sum_{j=0}^{\infty} \frac{x^{d/2+j-1} \delta^j}{\Gamma(d/2+j) 2^{2j} j!}$$

If the aforementioned normal random variables have means $\mu_1, \mu_2, \dots$, then $\delta = \sum_i \mu_i^2$.

The expectation and variance of the χ^2 distribution are $d + \delta$ and $2(d + 2\delta)$ respectively. In most common applications of the χ^2 distribution, the noncentrality parameter δ is 0. The distribution is then called a central χ^2 distribution (that is the distribution of the sum of squares of several independent standard normal variables), the summation in the pdf reduces to a single term (that for $j = 0$), the mean is d , and the variance is $2d$.

If d is an integer, a central χ^2 random variable with df d can be generated by generating d standard normal random numbers and adding their squares.

Several goodness-of-fit statistics (see the article by Tendeiro on this topic in this encyclopedia) used in educational applications follow the χ^2 distribution when the model fits the data. For example, the M_2 and χ_{pf} statistics that were suggested by

Maydeu-Olivares and Joe (2005) and Sinharay (2018), respectively, for assessing the goodness-of-fit of IRT models and person fit for response-time models follow the χ^2 distribution.

F distribution

The pdf of the F distribution is given by

$$f(x; \nu, w, \delta) = k \frac{e^{-\delta/2} \nu^{\nu/2} w^{w/2} x^{\nu/2}}{B(\nu/2, w/2)(w + \nu x)^{(w+\nu)/2}}$$

where ν and w are df's and positive integers, $\delta \geq 0$ is the noncentrality parameter, and $k = 1 + \sum_{j=1}^{\infty} \left(\frac{\nu w x / 2}{w + \nu x} \right)^j \frac{(\nu + w)(\nu + w + 2)(\nu + w + 2j - 2)}{j! \nu (\nu + 2) \dots (\nu + 2j - 2)}$. Often, $\delta = 0$, in which case $k = 1$ and the F distribution is called the central F distribution.

The mean and variance of the F distribution are $\frac{w(\nu + \delta)}{\nu(w - 2)}$ and $2 \left(\frac{w}{\nu} \right)^2 \frac{(\nu + \delta)^2 + (\nu + 2\delta)(w - 2)}{(w - 2)^2(w - 4)}$, respectively.

If two independent χ^2 random variables χ_1^2 and χ_2^2 have df's ν and w , then $\frac{\chi_1^2/\nu}{\chi_2^2/w}$ has an F distribution with ν and w df. This result can be used to generate random numbers following the F -distribution.

The F distribution has often been applied in education in analysis of variance (ANOVA; see the article by King), where the ratio of the mean square of an effect and the mean square of error has a central F distribution under the null hypothesis that the effect is zero. For example, Sinharay and Lu (2008) used ANOVA to examine if the mean of a fit statistic differed over three conditions and concluded that no difference over the conditions because of values of F statistics that were not statistically significant.

Dirichlet distribution

The Dirichlet distribution is the multivariate generalization of the beta distribution. In Bayesian analyses, the Dirichlet distribution is often used as a prior distribution of the parameters of the multinomial distribution. See, for example, Novick and Jackson (1974, chapter 10-7). Consider a k -dimensional random variable $\mathbf{X} = (X_1, X_2, \dots, X_k)$ so that the X_i 's are all positive and the sum of them is 1. If $\mathbf{X}$ follows the Dirichlet distribution, its pdf is given by

$$f(x_1, x_2, \dots, x_k; c_1, \dots, c_k) = \frac{\Gamma\left(\sum_{i=1}^k c_i\right)}{\prod_{i=1}^k \Gamma(c_i)} \prod_{i=1}^k x_i^{c_i-1}$$

where c_i , $i = 1, 2, \dots, k$ are the parameters of the distribution. The marginal distribution of X_i is the beta distribution with parameters c_i and $\sum_{j=1}^k c_j - c_i$. This property can be used to estimate the parameters c_i using the method of moments as discussed earlier for the beta distribution. For $k = 2$, the Dirichlet distribution becomes the beta distribution with parameters c_1 and c_2 .

To generate a random number $\mathbf{x} = (x_1, x_2, \dots, x_k)$ from the Dirichlet distribution with c_i s as parameters, one can generate a sequence of independent draws y_i , $i = 1, 2, \dots, k$, where y_i is a draw from the gamma distribution (to be discussed later) with parameters $b = 1$ and $c = c_i$, and compute $x_i = \frac{y_i}{\sum_{j=1}^k y_j}$, $i = 1, 2, \dots, k$

van Duijn and Jansen (1995) modeled spelling errors made by 721 school students on 4 dictation tests each of which includes 20 Dutch words. They assumed that the number of errors made by a student on a test follows a Poisson distribution (see the article on discrete distributions by Sinharay in this encyclopedia) where the mean parameter of the distribution is a product of an examinee ability parameter and a difficulty parameter specific to that examinee-test combination. The difficulty parameters are assumed to sum to 1 for each examinee over the four tests and are assumed to follow a Dirichlet distribution whose parameters are to be estimated. In an analysis of mathematics assessment data from early childhood longitudinal study using latent classes, Depaoli et al. (2016) assumed a Dirichlet prior distribution on the proportion of examinees belonging to each latent class.

Wishart distribution

The Wishart distribution is the multivariate generalization of the χ^2 distribution and is the distribution of a scalar multiple of the maximum-likelihood estimator of the covariance matrix of a multivariate normal distribution. A k -dimensional random variable $\mathbf{X}$ following the Wishart distribution has a pdf equal to

$$|\mathbf{X}|^{\frac{n-k-1}{2}} e^{-\frac{1}{2}\text{tr}(\Sigma^{-1}\mathbf{X})}$$

$$\frac{2^{nk/2} \pi^{k(k-1)/4} |\Sigma|^{n/2}}{\prod_{i=1}^k \Gamma\left(\frac{n+1-i}{2}\right)}$$

where Σ is a $k \times k$ matrix, n is the df, $\text{tr}(A)$ is the trace of matrix A and $|A|$ is the determinant of the matrix A . The distribution has a mean of $n\Sigma$.

If $Y_1, Y_2, \dots, Y_n$ are k -dimensional random numbers independently drawn from a multivariate normal distribution with mean $\mathbf{0}$, where $\mathbf{0}$ denotes a vector of zeroes, and variance matrix Σ , then the distribution of $\sum_{i=1}^n Y_i Y_i'$ is a k -dimensional Wishart distribution with k df and parameter Σ . This result can be used to generate random matrices following the Wishart distribution.

The inverse of a random variable following a Wishart distribution follows the *inverse Wishart distribution*. The inverse Wishart distribution is often assumed to be the prior distribution on the variance matrix in a Bayesian approach to IRT modeling (e.g., [Johnson and Sinharay, 2005](#)).

Conclusions

Among the numerous continuous statistical distributions, this article provides a brief overview of a few distributions that have found applications in the field of education. [Table 1](#) shows the pdf, mean, and variance of all the distributions discussed in this article (for simplicity, the table includes the expression of the variance of only one element/component for the Dirichlet and Wishart distributions). See, for example, [Evans et al. \(2000\)](#), ([Johnson et al. 1994, 1995](#)), [Kotz et al. \(2000\)](#) for more detailed descriptions of these and many more continuous distributions.

Several continuous distributions (such as the beta, gamma, exponential, and normal distributions) belong to the *exponential family of distributions* (e.g., [Haberman, 2016](#)) that have pdf of the form

$$\exp \left[\sum_i A_i(\theta) B_i(x) + C(x) + D(\theta) \right]$$

Table 1 Some features of the continuous distributions discussed in this article.

Distribution	Pdf	Mean	Variance
Uniform	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$\frac{1}{b} e^{-x/b}$	b	b^2
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Lognormal	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2}$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
Student's t	$\frac{\frac{1}{\Gamma(\nu/2)} \frac{v^{\nu/2} e^{-v/2}}{\pi^{0.5} (v+x^2)^{(\nu+1)/2}} \sum_{i=0}^{\infty} \Gamma\left(\frac{\nu+i+1}{2}\right) \frac{(x\delta)^i}{i!} \left(\frac{2}{v+x^2}\right)^{i/2}}$	$\frac{\delta(v/2)^{1/2} \Gamma((\nu-1)/2)}{\Gamma(\nu/2)}$	$\frac{v}{v-2} (1 + \delta^2) - \frac{v\delta^2}{2} \left(\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \right)^2$
Beta	$\frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
Gamma	$\frac{1}{b\Gamma(c)} \left(\frac{x}{b}\right)^{c-1} e^{-\frac{x}{b}}$	bc	b^2c
Cauchy	$\left\{ \pi b \left[1 + \left(\frac{x-a}{b} \right)^2 \right] \right\}^{-1}$	Does not exist	Does not exist
χ^2	$\frac{e^{-(x+\delta)/2}}{2^{d/2}} \sum_{j=0}^{\infty} \frac{x^{d/2+j-1} \delta^j}{\Gamma(d/2+j) 2^{2j} j!}$	$d + \delta$	$2(d + 2\delta)$
F	$k \frac{e^{-\delta/2} v^{v/2} w^{w/2} x^{(v-2)/2}}{B(v/2, w/2) (w+vx)^{(v+w)/2}}$	$\frac{w(v+\delta)}{v(w-2)}$	$2\left(\frac{w}{v}\right)^2 \times \frac{(v+\delta)^2 + (v+2\delta)(w-2)}{(w-2)^2(w-4)}$
Dirichlet	$\frac{\Gamma\left(\sum_{i=1}^k c_i\right)}{\prod_{i=1}^k \Gamma(c_i)} \prod_{i=1}^k x_i^{c_i-1}$	$E(x_i) = \frac{c_i}{\sum_{i=1}^k c_i}$	$V(x_i) = \frac{c_i \left(\sum_{i=1}^k c_i - c_i \right)}{\left(\sum_{i=1}^k c_i \right)^2 \left(\sum_{i=1}^k c_i + 1 \right)}$
Wishart	$ x ^{\frac{n-k-1}{2}} e^{-\frac{1}{2} \text{tr}(\Sigma^{-1}x)}$ $2^{nk/2} \pi^{k(k-1)/4} \Sigma ^{1/2} \prod_{i=1}^k \Gamma\left(\frac{n+1-i}{2}\right)$	$n\Sigma$	$V(x_{ij}) = 2n\sigma_{ij}^2$

where $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ is the p -dimensional parameter vector of the distribution. Uniform and Cauchy distributions are examples of distributions that do not belong to the exponential family. The exponential family of distributions have several attractive properties, such as the set of $B_i(x)$'s being jointly sufficient for the parameters θ . Hence, the family has been applied to several applications in education. See, for example, [Haberman \(2008\)](#), for an application of the family of distributions to equating.

The choice of the specific continuous distribution to be used in a specific application depends on several issues. If a specific distribution is used in practice for data of the type an investigator has that distribution should be the first choice. For example, a Dirichlet distribution is often applied in practice to model the probabilities from a multinomial distribution. If such knowledge is lacking, a sensible approach is to graphically plot the data and try to find a distribution that is expected to describe the data adequately. After fitting a distribution to a data set, methods of goodness-of-fit can be used to examine whether the distribution provides an adequate description of the data set.

References

- Box, G.E.P., Muller, M.E., 1958. A note on the generation of random normal deviates. *Ann. Math. Stat.* 29, 610–611.
- Bradlow, E.T., Wainer, H., Wang, X., 1999. A Bayesian random effects model for testlets. *Psychometrika* 64, 153–168.
- Depaoli, S., Yang, Y., Felt, J., 2016. Using Bayesian statistics to model uncertainty in mixture models: a sensitivity analysis of priors. *Struct. Equ. Model.* 24, 198–215.
- du Toit, M., 2003. IRT from SSL. Scientific Software International, Lincoln, IL.
- Evans, M., Hastings, N., Peacock, B., 2000. *Statistical Distributions*. Wiley, New York.
- Fox, J.-P., Marsman, M., Mulder, J., Verhagen, J., 2014. Complex latent variable modeling in educational assessment. *Commun. Stat. Simulat. Comput.* 45, 1499–1510.
- Gao, F., Chen, L., 2005. Bayesian or non-Bayesian: a comparison study of item parameter estimation in the three-parameter logistic model. *Appl. Meas. Educ.* 18, 351–380.
- Haberman, S.J., 2008. Continuous Exponential Families: An Equating Tool (ETS Research Rep. No. RR-08-05). ETS, Princeton, NJ.
- Haberman, S.J., 2016. Exponential family distributions relevant to IRT. In: van der Linden, W. (Ed.), *Handbook of Item Response Theory, Statistical Tools*, vol. 2. Chapman and Hall/CRC, Boca Raton, FL, pp. 47–69.
- Hambleton, R.K., Swaminathan, H., 1985. *Item Response Theory: Principles and Applications*. Kluwer Academic Publishers, Boston.
- Holland, P.W., 1979. The tyranny of continuous models in a world of discrete data. *IHS J.* 3, 29–42.
- Johnson, M.S., Sinharay, S., 2005. Calibration of polytomous item families using Bayesian hierarchical modeling. *Appl. Psychol. Meas.* 29, 369–400.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. *Continuous Univariate Distributions*, second ed., vol. 1. Wiley, New York.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. *Continuous Univariate Distributions*, second ed., vol. 2. Wiley, New York.
- Karadavut, T., 2019. The uniform prior for Bayesian estimation of ability in item response theory models. *Int. J. Assess. Tools Educ.* 6, 568–579.
- Kotz, S., Balakrishnan, N., Johnson, N.L., 2000. *Continuous Multivariate Distributions*, vol. 1. Wiley, New York.
- Lord, F.M., 1965. A strong true score theory, with applications. *Psychometrika* 30, 239–270.
- Maydeu-Olivares, A., Joe, H., 2005. Limited and full information estimation and goodness-of-fit testing in 2^n contingency tables: a unified framework. *J. Am. Stat. Assoc.* 100, 1009–1020.
- Novick, M.R., Jackson, P.H., 1974. *Statistical Methods for Educational and Psychological Research*. McGraw-Hill, New York.
- Patz, R., Junker, B., 1999. Applications and extensions of MCMC in IRT: multiple item types, missing data, and rated responses. *J. Educ. Behav. Stat.* 24, 342–366.
- Scheiblechner, H., 1979. Specifically objective stochastic latency mechanisms. *J. Math. Psychol.* 19, 18–38.
- Sinharay, S., Lu, Y., 2008. A further look at the correlation between item parameters and item fit statistics. *J. Educ. Meas.* 45, 1–45.
- Sinharay, S., 2018. A new person-fit statistic for the lognormal model for response times. *J. Educ. Meas.* 55, 457–476.
- van der Linden, W.J., 2006. A lognormal model for response times on test items. *J. Educ. Behav. Stat.* 31, 181–204.
- van Duijn, M.A.J., Jansen, M.G.H., 1995. Modeling repeated count data: some extensions of the Rasch Poisson counts model. *J. Educ. Behav. Stat.* 20, 241–258.
- Verhagen, J., Levy, R., Millsap, R.E., Fox, J.-P., 2016. Evaluating evidence for invariant items: a bayes factor applied to testing measurement invariance in IRT models. *J. Math. Psychol.* 72, 171–182.
- Zimmerman, D.W., 1997. A note on interpretation of the paired-samples t test. *J. Educ. Behav. Stat.* 22, 349.

Discrete probability distributions

Sandip Sinharay, ETS, Princeton, NJ, United States

© 2023 Elsevier Ltd. All rights reserved.

Bernoulli distribution	718
Binomial distribution	719
Geometric distribution	719
Hypergeometric distribution	719
Negative binomial distribution	720
Poisson distribution	720
The multinomial distribution	721
Conclusions	721
References	722

A discrete random variable (denoted X) is one that can take one of an enumerable or countably infinite set of values x_i , $i = 1, 2, \dots$, each with a specific probability p_i so that the probabilities over all the values add up to 1 (that is, $\sum_i p_i = 1$). The function that specifies the probabilities for the possible values is defined as the probability mass function (pmf). The cumulative distribution function (cdf) of a discrete distribution, which provides the probability that the corresponding random variable is smaller than or equal to a specific value, is a step function. Discrete random variables occur naturally in statistics and several other fields including education; for example, the number of items correctly answered on a test, the number of examinees passing a test etc. constitute discrete random variables. [Holland \(1979\)](#) commented that the world is full of discrete data. Hence several types of discrete random variables have been suggested in the literature. There also exist several statistical methods specifically designed to handle discrete data—see, for example, [Bishop et al. \(1975\)](#) and the article by Anderson on categorical data analysis in this encyclopedia.

In the following sections, several discrete distributions are discussed. For each distribution, its genesis, pmf, moments, estimation of the moments, methods to generate values of a random variable following that distribution, and relationships to other distributions (if any) are discussed. Applications of these distribution to educational problems are also mentioned, when appropriate.

Bernoulli distribution

The Bernoulli distribution is the most basic discrete distribution. A variable that follows the distribution can take one of two possible values, 1 (usually referred to as a success) or 0 (a failure), where the probability of success is p , $0 \leq p \leq 1$. An example of a Bernoulli random variable (that is, a variable that follows the Bernoulli distribution) is the outcome of a coin toss, where the outcome is either a head (success) or a tail (failure) and the probability of a head is a number between 0 and 1 (the probability is 0.5 if the coin is balanced).

The pmf of a Bernoulli distribution is given by

$$P(X=x) = p^x(1-p)^{1-x}$$

where x can be either 0 or 1. The cdf $F(x)$ of the distribution is 0 if $x < 0$, $1-p$ if $0 \leq x < 1$, and 1 if $x \geq 1$. The mean and the variance of the distribution are p and $p(1-p)$.

The maximum likelihood estimate (MLE) of p from a sample $x_1, x_2, \dots, x_s$ from the Bernoulli random variable is the sample mean $\bar{x} = \frac{1}{s} \sum_{i=1}^s x_i$, which is the proportion of successes in the sample. For example, six heads in 10 coin tosses results in the sample mean and hence the MLE of p being equal to 0.6.

To generate a random number X from the Bernoulli distribution, one has to first generate a random number R from the $U(0, 1)$ distribution, that is, the uniform distribution between 0 and 1 (A review of the uniform distribution can be found in the article by Sinharay on continuous distributions in this encyclopedia). If R is less than p , then X is set to 0; otherwise, X is set to 1.

The Bernoulli distribution is employed in the basic building block of item response theory (IRT) models (see, for example, articles by Ackerman et al. and Cai in this encyclopedia, and [Hambleton and Swaminathan, 1985](#)), which are heavily used to analyze data from educational tests. Several educational tests include dichotomously scored (0/1) items. In an application of an IRT model, these item scores are assumed to follow the Bernoulli distribution with p equal to a suitably chosen nonlinear function of the examinee and item characteristics.

Binomial distribution

If $Y_1, Y_2, \dots, Y_n$ are independent Bernoulli random variables, all with success probability $p, 0 \leq p \leq 1$, then $\sum_{i=1}^n Y_i$, which denotes the number of successes among the Y_i s, is a binomial random variable with number of trials = n and success probability p .

The pmf of a Binomial distribution is given by

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, 2, \dots, n$$

The mean and variance of the distribution are np and $np(1-p)$.

The MLE of p from a sample $x_1, x_2, \dots, x_n$ from the Binomial distribution with n trials and success probability p is the ratio of the sample mean $\bar{x}$ and n . To generate a random number from a binomial distribution, one can generate n Bernoulli random variables, all with probability p , and add them up. Alternatively, one can compute the cdf of the binomial distribution, generate $R \sim U(0, 1)$, and then take x as the random number from the binomial distribution if the cdf for $x-1$ is less than R and the cdf for x is greater than or equal to R (this is often called the *cdf-inversion* method).

Several applications of the binomial distribution to educational data can be found in [Novick and Jackson \(1974\)](#). In one such application, the number of words spelled correctly (out of a total of n words) by a student, where each word is equally difficult to spell, is assumed to follow the binomial distribution. The binomial distribution has also played a major role in the modeling of raw scores on educational tests. For example, [Keats and Lord \(1962\)](#) denoted the true proportion-correct score (or relative true score) and observed raw score on a test with I dichotomously scored items of a randomly chosen examinee as p and X , respectively. The distribution of X given p was assumed to be the binomial distribution with I trials and success probability p . If $f(p)$ denotes the population distribution of the true proportion-correct score, then the marginal distribution of X is given by

$$P(X=x) = \int_p \binom{I}{x} p^x (1-p)^{I-x} f(p) dp$$

The discrete distribution provided by the above equation is referred to as the *beta-binomial distribution*, which is another, albeit less popular, discrete distribution.

Geometric distribution

If a sequence of independent Bernoulli random variables are generated with the same success probability $p, 0 < p \leq 1$, the number of trials or failures before the first success follows a geometric distribution.

The pmf of a variable X following the geometric distribution is given by

$$P(X=x) = p(1-p)^x, x = 0, 1, 2, \dots, \infty$$

Thus, the number of possible values of a geometric random variable is countably infinite.

The mean and variance of a geometric distribution are $\frac{1-p}{p}$ and $\frac{1-p}{p^2}$. The maximum likelihood estimate of p from a sample from the geometric distribution is $\frac{1}{1+\bar{x}}$.

The cdf of the geometric distribution for an integer r is given by $P(X \leq r) = 1 - (1-p)^{r+1}$. Using this expression, a random number from a geometric distribution can be generated by first generating $R \sim U(0, 1)$ and then computing the integer part of $\frac{\log(R)}{\log(1-p)}$.

In an approach to control item exposure in computerized adaptive tests (CATs) suggested by [Lewis \(2007\)](#), one assigns each item in the item pool a probability r of actually being administered given that it has been selected by the CAT algorithm used to deliver the test. Suppose that at each stage of the CAT, the algorithm gives an ordered list of the remaining items in the item pool in terms of attractiveness to be administered. The first item in the list would have a probability r of being administered at this stage. The second item has a probability $r(1-r)$ of being administered, the third a probability of $r(1-r)^2$ and so on. In general, if G denotes the random position in the list of the item that is actually administered at this stage, then G follows the geometric distribution.

Hypergeometric distribution

Suppose that a sample of n items was drawn without replacement from a population of N elements of which M elements have a certain attribute. The number of elements in the sample with the attribute is a hypergeometric random variable. This distribution could arise when, for example, a group of N examinees includes M examinees who passed a test and $(N-M)$ who failed; if one draws a sample of n examinees from the group, the variable denoting the number of examinees who passed in the sample, X , is a hypergeometric random variable. The parameters for the distribution are N and M .

The pmf of the distribution is given by

$$P(X=x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

The mean and variance of the distribution are $\frac{nM}{N}$ and $\frac{(nM/N)(1-M/N)(N-n)}{N-1}$, respectively.

When $\frac{n}{N}$ is small and N is large, the hypergeometric distribution can be approximated by a binomial random variable with n trials and $p = \frac{M}{N}$.

The parameters of the distribution can be estimated from a sample of size n using the MLE. The estimate of N is the largest integer less than or equal to $\frac{nM}{x}$ and that of M is the largest integer less than or equal to $\frac{(N+1)x}{n}$.

To draw a random number from the distribution, one can construct a routine that simulates an experiment, by drawing elements at random, that will generate a hypergeometric random variable. For example, for given values of N (the number of examinees), M (the number of examinees who passed), and n (the number of draws), one can draw n elements from a sample of N examinees among whom M have been marked as “passed”; the number of examinees marked as “passed” in the sample of n would be a random draw from the hypergeometric distribution. Alternatively, one can compute the cdf of the distribution and then generate a hypergeometric random number with the help of a single $U(0, 1)$ random number and the cdf-inversion method.

Meijer (2004) suggested a nonparametric approach to assess person fit. In the approach, one considers a pair of subtests, one easier and the other more difficult, and then examines whether the number correct score on the more difficult subtest (x_D) is exceptionally low given the number correct score on the two subtests combined (x_T). Suppose I_T denotes the number of items in the two subtests combined and I_D denotes the number of items on the more difficult subtest. To perform the examination, one computes the probability of an exceptionally low subtest score under the assumption of the hypergeometric distribution, with I_T , I_D , and x_T playing the roles of N , M and n , respectively.

Negative binomial distribution

The number of failures before the n -th success in a sequence of draws of Bernoulli random variables, where the success probability is p in each draw, is a negative binomial random variable. The pmf of the distribution is given by

$$P(X=x) = \binom{n+x-1}{n-1} p^n (1-p)^x, 0 < p \leq 1$$

If X follows the negative binomial distribution with n successes and success probability of p , then X is also the distribution of the sum of n independent geometric random variables with success probability of p .

The mean and variance of a negative binomial distribution are $n \frac{1-p}{p}$ and $n \frac{1-p}{p^2}$.

The MLE of p from a sample from the negative binomial distribution is $\frac{n}{n+\bar{x}}$, where $\bar{x}$ is the sample mean.

If p is small, one can generate a negative binomial random number by adding up n geometric random numbers. Another way is to generate a sequence of $U(0, 1)$ random variable values. When the number of those values less than p first reaches n , the number of values that are greater than p is a draw from the negative binomial distribution. Oliveira et al. (2018) used the negative binomial to model data on the number of allocated students for several combinations of institution and program under the Portuguese higher education system.

Poisson distribution

The Poisson distribution describes the probability of finding exactly x events in a given length of time if the events occur independently at a constant rate. In addition, the Poisson distribution can be obtained as an approximation of a binomial distribution when the number of trials n of the latter distribution is large, success probability p is small, and np is a finite number.

The pmf of the Poisson distribution is given by

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, \dots, \infty$$

where λ is a positive number. Both the mean and variance of the Poisson distribution are equal to λ .

The MLE of λ from a sample from the Poisson distribution is the sample mean.

It is possible to generate a Poisson random variable by generating a random number following the $U(0, 1)$ distribution and then using the cdf-inversion method.

The Poisson distribution has been applied in education, for example, by Novick (1969), for modeling the number of errors committed by an examinee while reading a text—the λ parameter was assumed to be a product of a text difficulty parameter for the corresponding text and an error propensity parameter for the corresponding examinee.

An extension of the Poisson distribution is the zero-inflated Poisson (ZIP) distribution that is used for data that include an excessive number of individuals with a zero value of the variable of interest. The pmf of the ZIP distribution is given by

$$P(X=x) = \begin{cases} \rho + (1-\rho)e^{-\lambda} & \text{if } x = 0 \\ (1-\rho) \frac{e^{-\lambda} \lambda^x}{x!} & \text{if } x > 0 \end{cases}$$

where the parameter ρ is included to account for the extra zero values. A closely related model is the hurdle Poisson distribution whose pmf is given by

$$P(X=x) = \begin{cases} \rho & \text{if } x = 0 \\ \frac{1-\rho}{1-e^{-\lambda}} \frac{e^{-\lambda} \lambda^x}{x!} & \text{if } x > 0 \end{cases}$$

Salehi and Roudbari (2015) used the ZIP distribution to model the number of failed courses and semesters for the students of a university in Iran and Desjardins (2016) used the ZIP distribution to model data on the number of days suspended for students from a large urban school district.

The multinomial distribution

The multinomial distribution is a multivariate generalization of the binomial distribution. Consider a trial that results in exactly one of some fixed finite number k of possible outcomes, with probabilities $p_1, p_2, \dots, p_k$ (so that $p_i \geq 0$ for $i = 1, \dots, k$ and $\sum_{i=1}^k p_i = 1$), and there are n independent trials. Then let the random variables X_i indicate the number of times outcome number i was observed over the n trials. Then $X = (X_1, X_2, \dots, X_k)$ follows a multinomial distribution with parameters n and $\mathbf{p}$, where $\mathbf{p} = (p_1, p_2, \dots, p_k)$. Note that if X_i has the value x_i , $i = 1, \dots, k$, for a sample, then $\sum_{i=1}^k x_i = n$. An example where a multinomial random variable could occur is during the throw of a die. Let $X_i, i = 1, 2, \dots, 6$, denote the number of times i is observed in n throws of a die. Then $X = (X_1, X_2, \dots, X_6)$ has a multinomial distribution and $\sum_{i=1}^6 x_i = n$. If the die is fair, then $p_i = \frac{1}{6}$ for all i .

The pmf of the multinomial distribution is given by

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$

where $\sum_{i=1}^k x_i = n$. If $k = 2$, the multinomial distribution becomes a binomial distribution with n trials and success probability p_1 . When $X = (X_1, X_2, \dots, X_k)$ follows a multinomial distribution with the pmf given above, X_i follows a binomial distribution with n trials and success probability p_i . Hence, the mean and variance of X_i are np_i and $np_i(1 - p_i)$, respectively. The covariance between X_i and X_j is $-np_i p_j$.

The MLE of p_i for a multinomial distribution is the ratio of the sample mean of x_i 's and n .

The straightforward way to generate a multinomial random variable is to simulate an experiment (by drawing n uniform random numbers that are assigned to specific bins according to the cumulative value of the $\mathbf{p}$ vector) that will generate a multinomial random variable.

The multinomial distribution has been applied to build statistical models for the number of examinees who obtained various raw scores on a test (e.g., Holland and Thayer, 2000) and for the number of examinees who obtained specific patterns of scores such as $\{0, 1, 1, 0, 1, 0, 0, \dots, 1\}$ on the items of a test (e.g., Holland, 1990).

Conclusions

Among the numerous discrete probability distributions, this article provides a brief overview of arguably the most popular seven such distributions, all of which have found applications in education. Table 1 shows the pmf, mean, and variance of all the distributions discussed in this article. See, for example, Evans et al. (2000), Johnson et al. (2005, 1997) for more details on these and many more discrete distributions. The choice of the specific distribution to be used in a specific application depends on several issues. If a specific distribution is used in practice for data of the type an investigator has, that distribution should be the first choice. If such knowledge is lacking, a sensible approach is to graphically plot the data and try to find a distribution that is expected to describe the data adequately. After fitting a distribution to a data set, methods of goodness-of-fit (see the article by Ames Boykin et al. in this encyclopedia) can be used to examine whether the distribution provides an adequate description of the data set.

Table 1 Some features of the discrete distributions discussed in this article.

Distribution	pmf	Mean	Variance
Bernoulli	$p^x(1-p)^{1-x}$	p	$p(1-p)$
Binomial	$\binom{n}{x} p^x(1-p)^{n-x}$	np	$np(1-p)$
Geometric	$p(1-p)^x$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
Hypergeometric	$\frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$	$\frac{nM}{N}$	$\frac{(nM/N)(1-M/N)(N-n)}{N-1}$
Negative binomial	$\binom{n+x-1}{n-1} p^n(1-p)^x$	$n \frac{1-p}{p}$	$n \frac{1-p}{p^2}$
Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}$	λ	λ
Multinomial	$\frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$	np_i	$np_i(1-p_i)$

References

- Bishop, Y.M.M., Fienberg, S.E., Holland, P.W., 1975. Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, MA.
- Desjardins, C.D., 2016. Modeling zero-inflated and overdispersed count data: an empirical study of school suspensions. J. Educ. 84, 449–472.
- Evans, M., Hastings, N., Peacock, B., 2000. Statistical Distributions. Wiley, New York.
- Hambleton, R.K., Swaminathan, H., 1985. Item Response Theory: Principles and Applications. Kluwer Academic Publishers, Boston.
- Holland, P.W., Thayer, D.T., 2000. Univariate and bivariate loglinear models for discrete test score distributions. J. Educ. Behav. Stat. 25, 133–183.
- Holland, P.W., 1979. The tyranny of continuous models in a world of discrete data. IHS J. 3, 29–42.
- Holland, P.W., 1990. The Dutch identity: a new tool for the study of item response models. Psychometrika 55, 5–18.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1997. Discrete Multivariate Distributions. Wiley, New York.
- Johnson, N.L., Kemp, A.W., Kotz, S., 2005. Univariate Discrete Distributions. Wiley, New York.
- Keats, J.A., Lord, F.M., 1962. Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. Psychometrika 27, 59–72.
- Lewis, C., 2007. Some Thoughts on Controlling Item Exposure in Adaptive Testing. Paper Presented at the 2007 GMAC® conference on Computerized Adaptive Testing.
- Meijer, R.R., 2004. Using patterns of summed scores in paper-and-pencil tests and computer-adaptive tests to detect misfitting item score patterns. J. Educ. Meas. 41, 119–136.
- Novick, M.R., Jackson, P.H., 1974. Statistical Methods for Educational and Psychological Research. McGraw-Hill, New York.
- Novick, M.R., 1969. Bayesian Methods in Psychological Testing (ETS Research Bulletin No. RB-69-31). ETS, Princeton, NJ.
- Oliveira, R., Gonçalves, A.M., Vasconcelos, R.M., 2018. A Negative Binomial Model for Student Allocation to Higher Education in Portugal During the Post-Bologna Process. Paper Presented at the International Conference of Numerical Analysis and Applied Mathematics (ICNAAM).
- Salehi, M., Roudbari, M., 2015. Zero inflated poisson and negative binomial regression models: application in education. Med. J. Islam. Repub. Iran 29, 1–6.

Descriptive statistics

Jennifer L. Green^a, Sarah E. Manski^a, Timothy A. Hansen^a, and Jennifer E. Broatch^{b, a} Michigan State University, East Lansing, MI, United States; and ^b Arizona State University, Glendale, AZ, United States

© 2023 Elsevier Ltd. All rights reserved.

Introduction	723
Describing univariate categorical data	723
Describing univariate numerical data	724
Measures of center	724
Measures of variability	725
Standard deviation	725
Interquartile range	726
Reporting measures of variability	726
Measures of shape	727
Skewness	727
Kurtosis	729
Measures of location	729
Describing multivariate data	729
Describing bivariate numerical data	730
Example	730
Conclusion	733
References	733

Introduction

The aim of descriptive statistics is to summarize categorical and numerical data in an informative way, both numerically and graphically. Descriptive statistics are used to help describe key features or characteristics of data, such as the shape of their distribution, where the center lies, and how the data vary about that center. Ideally, a collection of descriptive statistics is chosen to convey the data's salient features in a more concise, but just as effective, manner as reporting each individual value.

Descriptive statistics form a basis for all quantitative analysis and are a precursor for inferential statistics, which uses properties of a data set to make inference and predictions beyond the data. Depending on the scope of a project, descriptive statistics alone may be sufficient, or they can be used to inform further statistical analyses. In this entry, we provide an overview of commonly used descriptive statistics and how they are interpreted in summaries and analyses.

Descriptive statistics begin with a data set consisting of a set of variables with values or observations for each observational unit. We call the set of all observational units of interest the *population*, and any numerical descriptor of the population is called a *parameter*. In most analyses it is difficult to collect data for the entire population and thus the data set will only contain information about a subset of the population of interest. We call this subset of the population a *sample*, and any numerical descriptor of a sample is called a *statistic*. For example, suppose our population of interest is all high school students at a particular school, and consider a set of data on high school students that records the name, gender identity, year in school (9th, 10th, 11th, 12th), grade point average (GPA), and Scholastic Aptitude Test (SAT) score of each student. If our data set contains observations for every student in the school, then the set contains the entire population. If our data set instead contains only a subset of students in the school, we would call this set a sample. In this example, each student is an observational unit or case, and name, gender identity, year, GPA, and SAT score are variables.

In the following sections, we focus on statistics that describe *univariate data*, that is, data on a single variable, and we divide variables into two types: categorical and numerical. The variable type determines which types of descriptive statistics we can use and the types of analyses we can perform. We first discuss numerical and graphical methods to describe categorical variables, such as frequency tables and bar charts. For describing numerical data, we discuss commonly used measures of center, such as mean and median; commonly used measures of variability, such as standard deviation and interquartile range; commonly used measures of shape, such as skew and kurtosis; and commonly used measures of location, such as percentiles and standardized scores. We conclude by briefly discussing methods, such as a scatterplot and correlation, used to summarize bivariate numerical data and then present a practical example of how many of these descriptive statistics may be used and interpreted in an educational context.

Describing univariate categorical data

Many data sets contain *categorical variables* whose possible values are categories or levels of an attribute, allowing us to compare the frequency or percentage of observations in each category. For instance, in our previous example, gender identity and year are

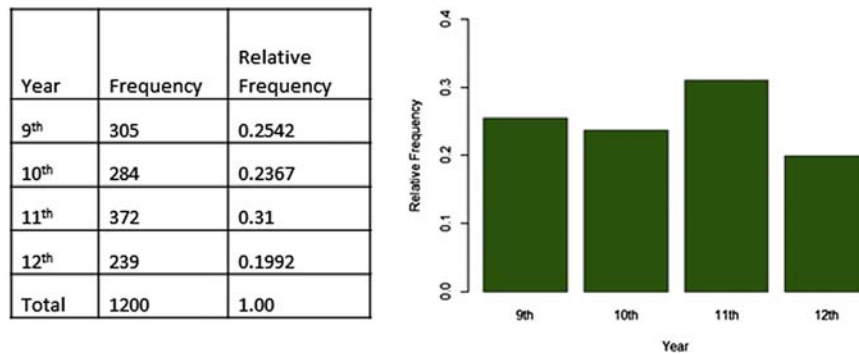


Fig. 1 Frequency table (left) and bar chart (right) of students' year in high school.

categorical variables that each have distinct categories. When the categories have a natural ordering, such as 9th, 10th, 11th, and 12th grade for year, we call the variable an ordinal categorical variable and report the categories in their natural order. When the categories have no such ordering, such as male, female, non-binary, etc. for gender identity, we call the variable a nominal categorical variable and can report the categories in any order.

To summarize categorical variables, counts or frequencies of observations in each category, as well as relative frequencies, can be displayed numerically and visually. The relative frequency or proportion of observations in a particular category is calculated by dividing the number of observations in the category by the total number of observations in the sample. Numerically, these descriptive statistics can be displayed using a frequency table that presents the name, frequency, and relative frequency of each category. These statistics can also be displayed and examined visually using bar charts, where vertical bars are drawn for each category according to their frequency or relative frequency. To demonstrate each of these methods, [Fig. 1](#) provides a frequency table and a bar chart for the nominal categorical variable year for a fictional high school. We can interpret these summaries using statements such as, "In the data set, the year with the most students is 11th grade with 372 students" or "Approximately 25.42% of the students in the data set are in 9th grade." In this way, we can summarize the data for a categorical variable more succinctly than listing all the observations.

Describing univariate numerical data

Numerical variables are variables whose values are represented on a numeric or quantitative scale, like GPA and SAT score in our initial example. When summarizing numerical variables, we seek to describe properties of the data distribution, that is, to describe information about the values observed and the overall pattern of how often the possible values occur. The simplest way to examine the distribution for a given variable is through graphical methods, such as histograms or box plots. Here, we focus on numerical summaries used to describe key features of a single variable's distribution—its center, variability, and shape—as well as measures of location, such as percentiles and standardized scores, that describe how an individual observation relates to the variable's distribution.

Measures of center

One salient feature of data is the center of its distribution, or the typical value of a numeric variable. Most commonly, the mean and the median are calculated, but other methods are sometimes useful. Choosing which measure(s) of center to report requires an understanding of the context of the question to be answered and the meaning behind the data, as well as consideration of the presence of extreme values and the overall shape of the distribution. Each measure of center tells us something different about the data, and each is impacted differently by the presence of extreme values. In this entry, we focus on the mean and median.

The *mean* is the arithmetic average, calculated by adding all data values together and dividing by the count or the total number of values. It tells us the fair-share value, the value at which the total would be unchanged if each observation were equal. The mean is also the balancing point where the total distance each observation lies from the mean is balanced or equal on each side. The *median* represents the middle value, separating the top and bottom 50% of the ordered observations. To find the median, the data are ordered from least to greatest, and a value that splits the observations into two subsets of equal size is identified. If there are an odd number of observations, the median is the value of the single observation that lies in the middle of the ordered data, but if there are an even number of observations, the median is typically the mean of the two observations that lie in the middle. For example, the median of the values 1, 3, and 5 would be 3, whereas the median of the values 1, 3, 5, and 10 would be $(3 + 5)/2$.

$2 = 4$. In this way, the median tells us the middle of the data, with approximately half above and half below and with little regard to the magnitude of each observation.

We often consider the presence of extreme values when deciding which measure(s) of center to report. The mean gives equal weight to all data values, so values that are extremely small or large relative to the other values have large impact on the value of the mean. In contrast, the median relies on the relative ordering rather than the magnitude of the data values, so extreme values have little or no impact. This resistance to extreme values makes the median useful when describing data with relatively extreme observations or outliers.

The context of the data is also important to consider when reporting a measure of center. The mean and the median offer different information, and we must often decide whether to report the mean or the median, or both. For example, when examining the values of properties in a school's area of service, the median value would let us know that approximately half the properties are worth less and approximately half are worth more. This could tell us something about the typical financial situation of families in the area, and would not change much if a new family moved in and bought a relatively expensive property. On the other hand, the mean, which gives equal weight to each observation, would be influenced by such a relatively expensive property value, resulting in a value that may not represent what is typical for the area. The mean tells us the value of a property if they were all worth the same and can better represent the typical value in the absence of extreme values. Because each measure of center conveys different information, it is important to understand how the statistic will be used when deciding which to report.

Measures of variability

A measure of center or typical value of a numeric variable only describes a portion of the overall data distribution. For a more complete description of the distribution, we also need to summarize the amount of variability in the data. For example, Fig. 2 illustrates midterm exam scores for two different classes that have the same mean midterm exam score of 70 points, but noticeably different distributions; the way the scores vary from one another differs from class to class. This variation is illustrated by the horizontal spread of the data.

Variability can be measured in multiple ways. The most simple measure of variability is *range*, calculated by the difference in the maximum and minimum values of a variable:

$$\text{Range} = \text{Maximum value} - \text{Minimum value}.$$

However, this measure does not utilize all the information contained in the data, so other measures of variability, such as standard deviation and interquartile range, are typically presented.

Standard deviation

A common measure of variability or spread, *standard deviation*, provides a typical distance of how far the observations are from the mean. This measure helps us understand whether most observations are close to the mean or far from it. The sample standard deviation, s , is calculated using the formula:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

where $\bar{x}$ is the sample mean. Note that we divide by the degrees of freedom, $n - 1$, versus the complete sample size, n , so that the *sample variance*, s^2 , is an unbiased estimate of the population variance, σ^2 . If the entire set of N population values and their mean μ are known, then the population variance, σ^2 , is calculated based on the formula for the population standard deviation, σ :

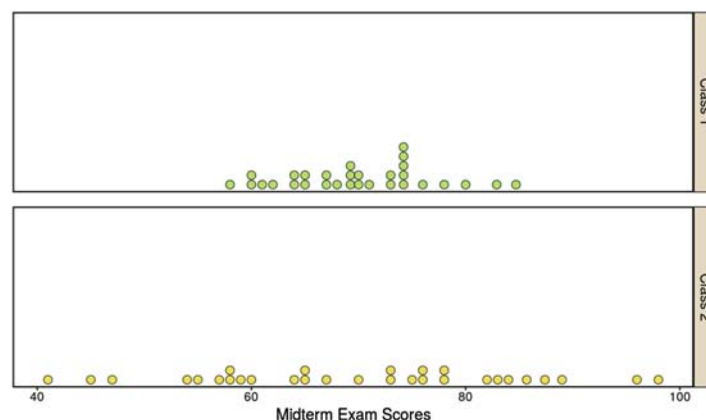


Fig. 2 Dot plots of students' midterm exam scores for two different classes. Plots were made using "Explore Quantitative Data" from "The Art and Science of Learning From Data Web Apps" <https://artofstat.com/web-apps>.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{N}}$$

Standard deviation conveniently measures the typical distance of an observation from the mean in the data's original units. However, this measure of variability, like the mean, is affected by extreme observations and is most appropriate for a roughly symmetric distribution (see [Skewness](#) section). Sample variance, s^2 , which is the square of standard deviation, is similarly affected by extreme observations but not as easily interpreted because it is measured in squared units.

Interquartile range

The *interquartile range* (IQR) can be used in lieu of, or in addition to, standard deviation to measure variability. Similar to the median, the IQR is resistant to extreme observations or outliers, because it depends on the order rather than the numeric values of the observations. The IQR represents the distance between the top 25% of the observations and the bottom 25% of the observations. That is, the IQR is the range of the middle 50% of the observations. Formally, IQR is calculated as the difference between the third quartile value, Q_3 , (75th percentile) and the first quartile value, Q_1 , (25th percentile):

$$IQR = Q_3 - Q_1$$

where the first quartile is calculated as the median of the sorted observations below the median and the third quartile is calculated as the median of the sorted observations above the median. Thus, the values of Q_1 , the median, and Q_3 split the data into approximately fourths. These three measures coupled with the minimum and maximum values form the *5 number summary*.

The IQR is often displayed in a box plot, where the size of the box is proportional to the IQR, visually displaying a measure of the variability in the data. [Fig. 3](#) displays box plots of the same two classes' exam scores presented in [Fig. 2](#), where Class 1 has less variability than Class 2. Here, the length of the box for Class 1 is visibly smaller than that of Class 2.

Reporting measures of variability

Within reports and statistical software output, tables are often used to summarize measures of center and variability, as well as other descriptive statistics. An example is given for the midterm exams in [Table 1](#). We see from these summary statistics that in Class 1, the middle 50% of students scored between 65 and 74.4 points, whereas the Class 2 scores were much more spread out, with the middle 50% between 58 and 82 points. This observation is supported by Class 2 having a standard deviation more than twice as large as Class 1, meaning that Class 2 scores are, in general, more than twice as far from their mean as scores from Class 1. By examining the descriptive statistics presented in the table, we can succinctly characterize and understand the center and variability of the data distribution.

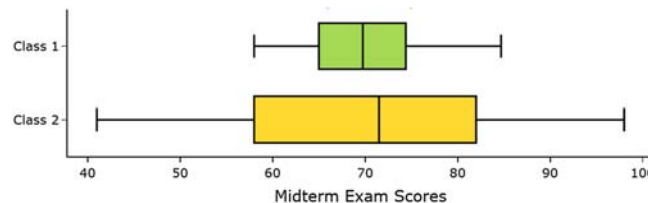


Fig. 3 Box plots of students' midterm exam scores for two different classes. Plots were made using "Explore Quantitative Data" from "The Art and Science of Learning From Data Web Apps" <https://artofstat.com/web-apps>.

Table 1 Descriptive statistics of the students' midterm exam scores displayed in [Figs. 2 and 3](#) for each class.

Class	n	Mean	Standard deviation	Minimum	Q_1	Median	Q_3	Maximum	IQR
Class 1	30	70.0	6.77	58.0	65.0	69.8	74.4	84.7	9.4
Class 2	30	70.0	14.83	41.0	58.0	71.5	82.0	98.0	24.0

Measures of shape

When summarizing data, we want to consider multiple characteristics of a distribution, not just its center and variability. Considering additional characteristics, such as a data distribution's shape, offers useful information that measures of center and variability may not provide, such as how the data are distributed across the range of values. Graphical displays and plots help depict these characteristics by illustrating not only the center and variability of a data distribution, but also its shape or overall pattern. For example, the plots in Fig. 4 display the midterm exam scores for two different classes of 30 students that each have the same mean of 75.2 points and the same standard deviation of 10 points, as reported in Table 2. However, the two distributions of exam scores are not identical. Even though Class 1 and Class 2 have the same mean score and the same standard deviation, the plots illustrate how each set of exam scores is distributed differently over the range of values from 49 to 89 points. For instance, the dot plots show that Class 1's midterm exam scores are primarily concentrated around two different values or modes, 70 points and 84 points, creating a bimodal distribution, whereas Class 2's midterm exam scores are primarily concentrated around 80 points, creating a unimodal distribution.

By displaying data visually, we are able to notice and compare aspects of each distribution's shape, such as its modality, that the mean and standard deviation do not convey. Other aspects we may notice about a data distribution's shape include skewness and kurtosis. Although numerical measures of skewness and kurtosis are not typically reported in educational research, these various aspects of a data distribution's shape, as well as its center and variability, allow us to more holistically describe and understand key features of the data. In the following sections, we discuss each of these distributional characteristics, summarizing what each represents and how to interpret statistical measures of skewness and kurtosis for unimodal distributions.

Skewness

Skewness is a distributional characteristic we use to help describe the lack of symmetry in a data distribution. Graphically, we can visualize a data distribution's symmetry, or lack thereof. When a distribution of values is approximately symmetric, such as those

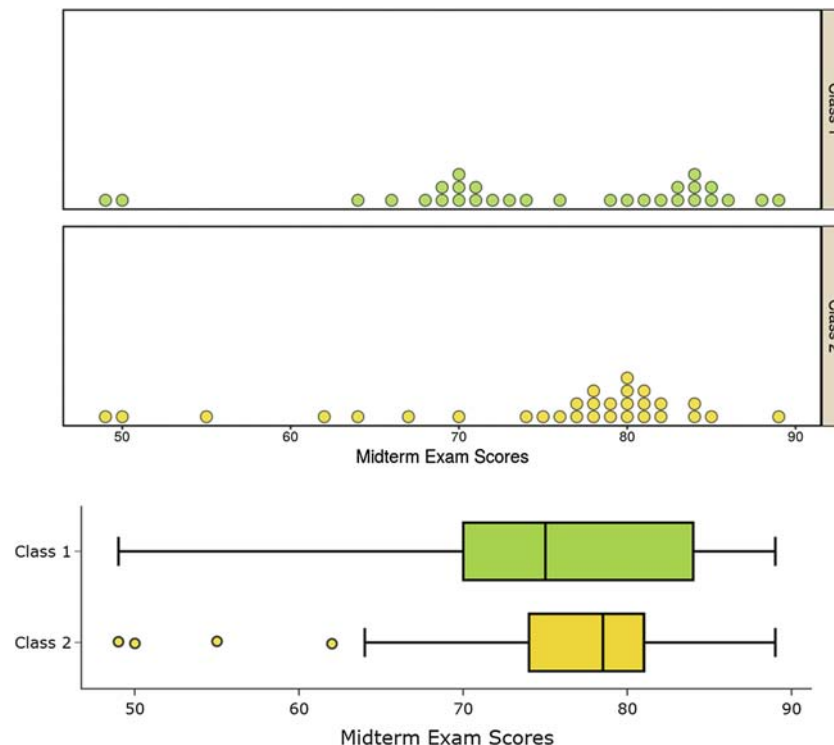


Fig. 4 Dot plots (top panel) and box plots (bottom panel) of students' midterm exam scores for two different classes. Plots were made using "Explore Quantitative Data" from "The Art and Science of Learning From Data Web Apps" <https://artofstat.com/web-apps>.

Table 2 Descriptive statistics of the students' midterm exam scores displayed in Fig. 4 for each class.

Class	<i>n</i>	Mean	Standard deviation	Minimum	Q1	Median	Q3	Maximum	IQR	Skew
Class 1	30	75.2	10.00	49.0	70.0	75.0	84.0	89.0	14.0	−0.85
Class 2	30	75.2	10.00	49.0	74.0	78.5	81.0	89.0	7.0	−1.33

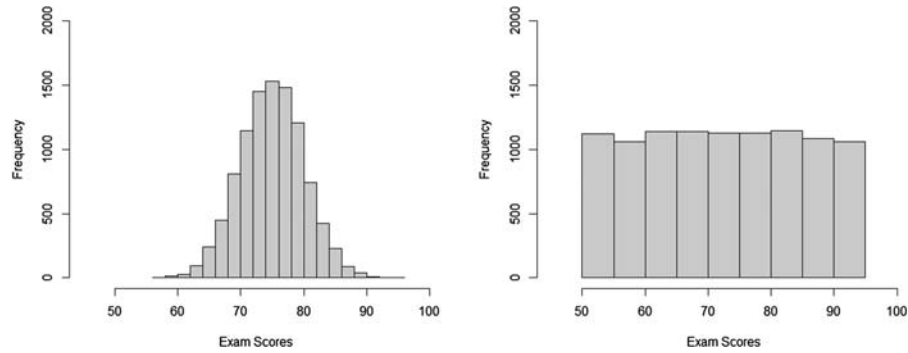


Fig. 5 Histograms of approximately symmetric distributions, one that is unimodal and bell-shaped like a Normal distribution (left) and one whose frequencies are nearly uniform across the range of scores (right).

shown in **Fig. 5**, the mean is close to the median and each side of the distribution along the mean or median is roughly a mirror image of itself. However, when one side, or tail, of a distribution stretches further from the center than the other tail, this asymmetry is called skew. For example, the two distributions of midterm exam scores displayed by the dot plots in **Fig. 4** are asymmetric or skewed, because they both have observations that extend or stretch their lower tails further to the left than their upper tails extend to the right. The boxplots also visually depict this negative or left skew by showing how the distance between the minimum score and the median is greater than the distance between the median and the maximum score for each class's distribution of midterm exam scores. Based on these characteristics, Class 1's distribution of midterm exam scores can be described as a left-skewed bimodal distribution, and Class 2's as a left-skewed unimodal distribution.

In general, a unimodal distribution's skew or lack of symmetry is described by the direction in which one of its tails stretches further from the mean or median than its other tail. A left- or negative-skewed unimodal distribution has a lower tail that creates an asymmetry in the distribution by extending far to the left, as illustrated by the plots of Class 2's midterm exam scores in **Fig. 4**, whereas a right- or positively-skewed unimodal distribution's upper tail extends far to the right, as illustrated in **Fig. 6**. Numerically, the skew of a distribution of quantitative observations can be calculated using the following formula:

$$\text{Estimated Skew} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3$$

where $\bar{x}$ and s are the sample mean and standard deviation, respectively. Distributions that are fairly symmetric, such as those in **Fig. 5**, with sample mean approximately equal to the sample median, have values of skew equal or close to zero. Positively- or right-skewed unimodal distributions, such as **Fig. 6**, have values of skew greater than zero, where the sample mean is typically larger than the sample median (e.g., see **Table 3**). In contrast, negatively- or left-skewed unimodal distributions, such as Class 2's distribution of scores in **Fig. 4**, have values of skew less than zero, and the sample mean is typically smaller than the sample median (e.g., see **Table 2**). The magnitude of skew gives a measure of the amount of asymmetry in a data distribution, with larger values indicating more asymmetric distributions. When considering multimodal distributions that are skewed, such as Class 1's distribution of scores in **Fig. 4**, categorizations of skewness values, and the corresponding relationships between the sample mean and median, may differ from those that apply to unimodal distributions, so it is important to also visually explore these characteristics with graphs.

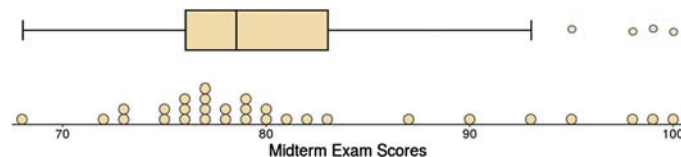


Fig. 6 Dot plot and corresponding box plot depicting a right- or positively-skewed distribution of students' midterm exam scores. Plots were made using "Explore Quantitative Data" from "The Art and Science of Learning From Data Web Apps" <https://artofstat.com/web-apps>.

Table 3 Descriptive statistics of students' midterm exam scores displayed in **Fig. 6**.

n	Mean	Standard deviation	Minimum	Q1	Median	Q3	Maximum	IQR	Skew
30	81.1	8.40	68.0	76.0	78.5	83.0	100.0	7.0	0.97

Kurtosis

When describing a data distribution's shape, we can also characterize how extreme its tails are to help identify potential issues with outlying or extreme observations. *Kurtosis* is a descriptive statistic used to help measure how data disperse between a distribution's center and tails, with larger values indicating a data distribution may have "heavy" tails that are thickly concentrated with observations or that are long with extreme observations. A distribution's kurtosis can be calculated using different formulas, including the following:

$$\text{Estimated Kurtosis} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4$$

Because the standardized distance each value falls from the mean is raised to the fourth power, kurtosis increases for every observation that differs from the mean but is most impacted by observations that are more than one standard deviation above or below the mean and lie in a data distribution's tails. For example, the normal distribution, which is a unimodal, bell-shaped distribution that looks similar to the data distribution on the left in Fig. 5, has kurtosis equal to three, where most of the distribution is concentrated around the mean and its tails quickly approach zero in height. Because the normal distribution often serves as a reference distribution for comparing other distributions, the value of a distribution's kurtosis is typically reported as the difference between its kurtosis and the kurtosis of the normal distribution, or as its "excess kurtosis." Data distributions with positive values of excess kurtosis, such as Class 2's distribution of midterm exam scores (Fig. 4), tend to have heavier tails than the normal distribution, whereas those with negative values of excess kurtosis, such as the approximately uniform data distribution shown in Fig. 5, tend to be broader distributions with few to no extreme observations. In general, higher positive values of excess kurtosis suggest that there may be issues with outliers in the data, offering additional information about the data's distribution that should be explored further by checking the data's accuracy and summarizing it graphically.

Measures of location

In addition to collectively summarizing a variable's data distribution, we often want to describe a single value or data point in relation to the full set of values. This may be because we want to quantify how far it is from the center of the data, or because we wish to communicate something about that data point, such as how a single test score compares to the whole group of test takers. There are two common approaches to describing location. The location of a data point can be reported as a *percentile*, stating the proportion of the data at or below the selected value. We could also report a *standardized score* to describe how the observation lies in relation to the mean, above or below, using the standard deviation as our measuring stick.

Percentiles are commonly used to report standardized test scores to students and parents. Typically, in this context, percentiles are defined to be the proportion of the test scores that are less than but not equal to the selected score, expressed as a percentage. This definition excludes the possibility of a 100th percentile, since each value is included in the set and cannot be less than itself. However, percentiles can be defined in other ways, so it is important for writers to include the definition they are applying. In all uses, percentiles divide the data into 100 slices and use these cuts as measures of location. Other -iles can be used, such as quartiles, quintiles, and deciles, which divide the data into four, five, or ten slices, respectively. These slices each contain approximately the same number of data points but are often unequal in length. For example, test scores falling in the first decile may range from 0 to 46 and those falling in the second decile may range from 47 to 52, but each decile still contains the same number (10%) of test scores. The median is defined as the 50th percentile.

Standardized scores are most often used with a symmetric unimodal data distribution and are found by subtracting the mean from the value of interest and dividing the difference by the standard deviation:

$$\text{Standardized Score} = \frac{x - \bar{x}}{s}$$

This standardized value tells us how many standard deviations above or below the mean a data point falls. Standardized scores are typically used to gauge a point's extremeness, or to compare values from multiple variables with different scales. The process of calculating a standardized score is an example of standardization, and similar calculations are used in finding scale scores on large-scale high-stakes testing. When data are approximately normally distributed, common rules of thumb can be developed and used. For example, values that are greater than three standard deviations away from the mean ($|\text{standardized score}| \geq 3$) would be unusual in a normal distribution. Although such rules only apply to normally distributed data, standardized scores can be calculated for any data.

Describing multivariate data

Although this entry primarily focuses on summarizing univariate data, data sets often contain data on multiple variables, not just one, that may be related. Statistics that describe *multivariate data*, that is, data collected on two or more variables for each observational unit, help us summarize and explore potential relationships between variables that may not be evident when summarizing each variable individually. For instance, the previously described data set on high school students contains multivariate data because data were collected on multiple variables for each student, including their gender identity, year in school, GPA, and SAT score. With these multivariate data, we can describe each variable not only individually, but also in relation to the other variables

to help uncover and understand different relationships, such as the relationship between a student's GPA and SAT score and the relationship between a student's year in school and SAT score.

When summarizing multivariate data, we first describe each variable individually, using descriptive statistics such as those discussed in previous sections. After exploring each variable of interest individually, we then numerically and graphically summarize them together to explore their potential relationship(s). Similar to univariate data, the number and types of variables we want to simultaneously summarize determines which types of descriptive statistics we can use and the types of analyses we can perform. "In this entry, we focus on graphical and numerical summaries used to describe the relationship between two numerical variables, that is, *bivariate numerical data*, but descriptive statistics and analyses also exist for other types of multivariate data, including the combination of numerical and categorical variables."

Describing bivariate numerical data

We use scatterplots to visualize and identify patterns and trends in bivariate numerical data. In a scatterplot, each observation is represented by a point or dot whose position on the horizontal and vertical axes indicates its value for each respective variable. Plotting each observation in a scatterplot and identifying patterns among the points can help us determine if a relationship exists between two numerical variables.

Often, we use scatterplots to assess the linear relationship between two numerical variables. Scatterplots might show positive linear trends, negative linear trends, or no linear trend at all. A scatterplot shows a positive linear trend if the variable on the vertical or y-axis tends to increase at a constant rate as the variable on the horizontal or x-axis increases. Conversely, if the variable on the y-axis tends to decrease at a constant rate as the variable on the x-axis increases, then the scatterplot shows a negative linear trend.

The *correlation coefficient*, r , is a statistic for measuring the strength and direction of a bivariate linear relationship and is calculated using the following formula:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

The correlation coefficient can assume any value between -1 and 1 ($-1 \leq r \leq 1$), where its sign indicates the direction of the linear trend and its magnitude indicates the strength of the linear relationship between the two numerical variables. For example, the correlation coefficient for a perfect positive linear relationship where all observations lie in a straight, positive line is $r = 1$, whereas the correlation coefficient for a perfect negative linear relationship is $r = -1$. The weaker the linear relationship tends to be, the closer in magnitude the correlation coefficient is to 0. Because the correlation coefficient is easily affected by influential observations and may not accurately reflect the true pattern in the data, it is important to also look at the corresponding scatterplot when describing the relationship between two numerical variables. For example, if the correlation coefficient is close to or approximately 0, looking at the scatterplot will determine whether there is no relationship between the two variables or if the relationship is simply not linear.

Example

Descriptive statistics are reported and used in educational reports and research. For example, [Table 4](#) presents several different statistics that summarize 2016–2017 United States high school seniors' most recent SAT scores for each state in the South Census Region ([United States Census Bureau, 2013](#)) and the entire United States ([Snyder et al., 2019](#)). To obtain each state's reported means ("Mean score") and standard deviations ("SD"), along with its "percentage distribution of seniors who had taken the SAT, by total SAT score range," the original data set would have contained multiple numerical variables for each student, including their total, evidence-based reading and writing, and math SAT scores, as well as one categorical variable, their state. To obtain the percent of graduating high school seniors in 2017 who took the SAT for each state, another data set would have contained each state's total number of graduating high school seniors in 2017, as well as the number of those seniors who took the SAT.

The descriptive statistics reported in [Table 4](#) summarize the center, variability, and shape of each state's distribution of total SAT scores. For example, the United States' mean and standard deviation for total SAT scores indicate that, on average, 2016–2017 United States high school seniors who took the SAT earned a total SAT score of 1060 points, and scores tended to vary from that mean by 195 points on average. In South Carolina, 2016–2017 high school seniors who took the SAT had a mean total SAT score similar to the national mean, at 1064 points, but their total SAT scores varied less from their mean on average ($SD = 174$) than those in the United States did ($SD = 195$). In contrast, the mean total SAT score for those in the District of Columbia, at 950 points, was lower than the national average, but students' scores were more variable, varying an average of 240 points from the mean. Across the states in the South Census Region, the mean total SAT scores form a bimodal distribution, with eight of the states having mean total SAT scores between 1000 and 1100 points, and four of the states having mean total SAT scores between 1200 and 1250 points (see [Fig. 7](#)). The variability among 2016–2017 seniors' SAT total scores was also different across the South Census Region states, with scores in West Virginia ($SD = 171$), South Carolina ($SD = 174$), North Carolina ($SD = 176$), and Georgia ($SD = 177$) varying the least from each of their respective means on average and those in the District of Columbia ($SD = 240$) and Oklahoma ($SD = 257$) varying the most.

Each state's percentage distribution of total SAT scores for 2016–2017 high school seniors who took the SAT offers information about how students' scores were distributed across the range of possible scores from 400 to 1600 points. Visually summarizing these

Table 4 Table of Scholastic Aptitude Test (SAT) mean scores, Standard Deviations (SD), and score ranges for high school seniors who took the SAT in 2017, as well as the percentage of the 2017 graduating class of high school seniors who took the SAT by state in the South Census Region.

State (abbreviation)	Evidence-based reading and writing						Percentage distribution of seniors who had taken the SAT, by total SAT score range							Percent taking SAT
	Total SAT		writing		Math		SAT score range							
	Mean score	SD	Mean score	SD	Mean score	SD	1400–1600	1200–1390	1000–1190	800–990	600–790	400–590		
United States (US)	1060	195	533	100	527	107	5	20	36	30	9	#	48	
Alabama (AL)	1165	199	593	100	572	111	15	29	33	20	#	0	5	
Arkansas (AR)	1208	198	614	103	594	107	19	36	28	14	#	0	3	
Delaware (DE)	996	197	503	104	492	103	#	14	31	36	16	#	100	
District of Columbia (DC)	950	240	482	126	468	123	6	13	18	32	29	#	90	
Florida (FL)	1017	190	520	97	497	104	#	16	33	35	12	#	83	
Georgia (GA)	1050	177	535	92	515	95	3	18	39	34	7	#	61	
Kentucky (KY)	1247	188	631	95	616	105	26	36	27	10	#	0	4	
Louisiana (LA)	1198	204	611	102	586	112	20	33	27	17	#	0	4	
Maryland (MD)	1060	199	536	102	524	108	5	21	35	30	9	#	69	
Mississippi (MS)	1242	187	634	91	607	107	25	36	28	10	#	0	2	
North Carolina (NC)	1081	176	546	92	535	95	4	22	41	28	5	#	49	
Oklahoma (OK)	1047	257	530	133	517	132	12	20	20	27	21	#	7	
South Carolina (SC)	1064	174	543	91	521	94	3	20	40	31	5	#	50	
Tennessee (TN)	1228	190	623	95	604	107	22	36	30	9	#	#	5	
Texas (TX)	1020	193	513	102	507	101	#	15	34	35	12	#	62	
Virginia (VA)	1102	184	561	95	541	100	6	25	40	24	4	#	65	
West Virginia (WV)	1086	171	558	88	528	95	5	20	44	28	#	0	14	

#Rounds to zero.

Adapted from Digest of Education Statistics 2017 (NCES 2018-070, p. 258), by Snyder, T.D., de Brey, C., Dillow, S.A., 2019. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. In the public domain. Adapted with permission.

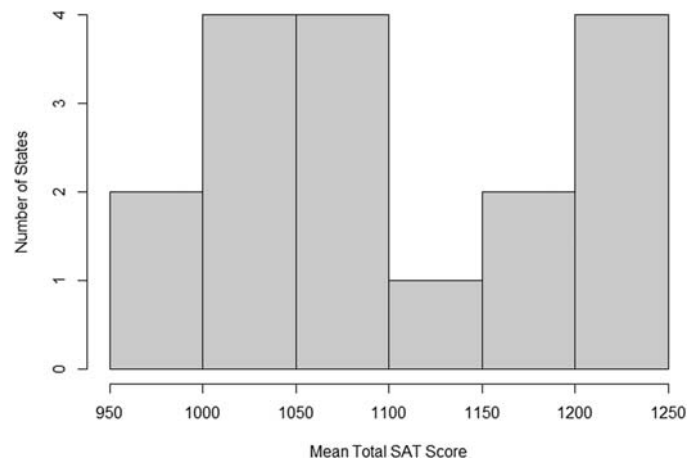


Fig. 7 Histogram of states' mean total SAT scores for 2016–2017 high school seniors in the South Census Region who took the SAT.

distributions with histograms (see Fig. 8) illustrates similarities and differences across the states' distributions, and the grayscale shading conveys the percentage of the state's 2017 graduating class of high school seniors who took the SAT. States with darker shadings, such as Delaware, Florida, and District of Columbia, had over 80% of their 2017 graduating classes take the SAT. However, when comparing these states' distributions of total SAT scores, we notice that Delaware's and Florida's percentage distributions are unimodal and fairly symmetric, whereas the District of Columbia's distribution of scores, while also unimodal, is right-skewed. In general, Delaware's and Florida's distributions of total SAT scores are similarly centered (Mean = 996 and 1017, respectively), variable (SD = 197 and 190, respectively), and shaped, with most of the students scoring between 800 and 1190 points. In contrast, the total SAT scores of 2016–2017 high school seniors in District of Columbia who took the SAT tended

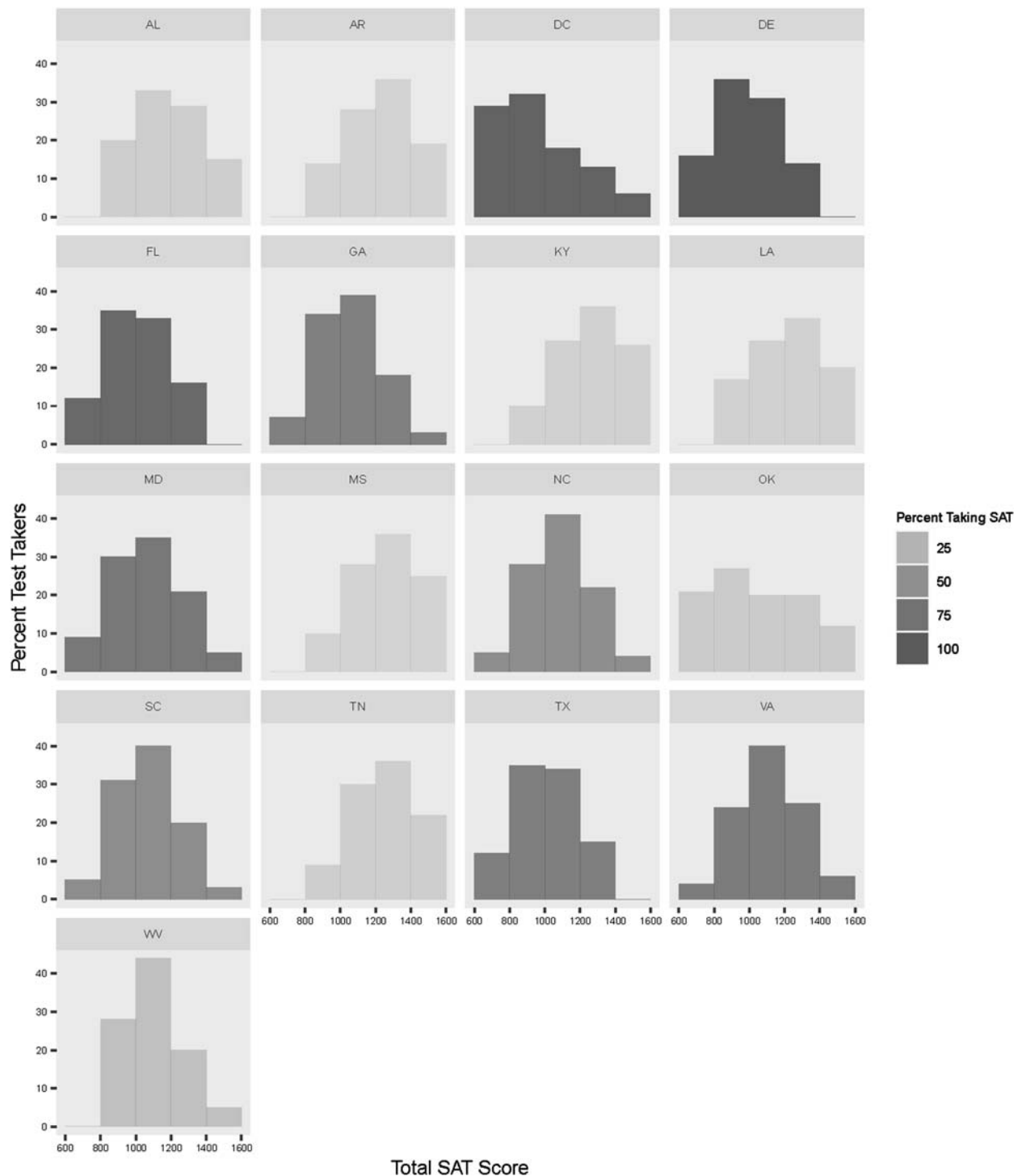


Fig. 8 Histograms of total SAT scores for 2016–2017 United States high school seniors by state, where the shading represents the percentage of the state's graduating class of high school seniors in 2017 who took the SAT.

to be relatively lower on average (Mean = 950), with most lying between 600 and 990 points, and a few higher than that, creating a right-skewed distribution with greater variability ($SD = 240$). Across all the distributions, there are noticeable similarities and differences in how total SAT scores were distributed, not just in mean and variability, but also in shape. Most of the South Census Region states' distributions of total SAT scores are unimodal and fairly symmetric (e.g., DE, FL, MD, NC, SC, TX, and VA), but other distributions are unimodal and right skewed (e.g., DC and WV) or unimodal and left skewed (e.g., KY, MS, and TN). Further, the differences in shading illustrate how approximately half the states had less than 15% of their 2017 graduating classes take the SAT, whereas the other states had at least 49% take the exam.

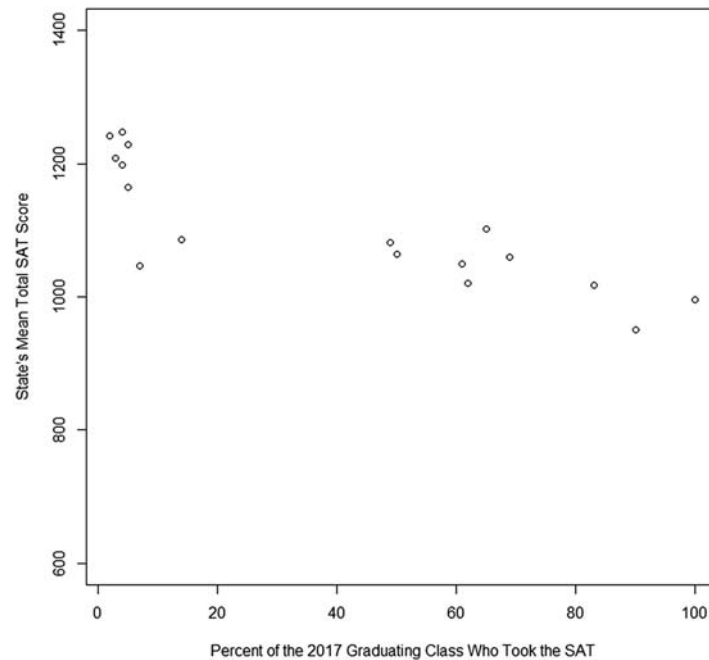


Fig. 9 Scatterplot of each South Census Region state's mean total SAT score for 2016–2017 high school seniors by its percent of graduating class of high school seniors in 2017 who took the SAT.

The notable differences in the percent of the 2017 graduating class who took the SAT in each state and the bimodal distribution of the states' mean total SAT scores (see Fig. 7) suggest that there might be a relationship between these two numerical variables. Visualizing these variables in a scatterplot (see Fig. 9), we notice that there are two distinct groups of points—states in which less than 15% of their 2017 graduating classes took the SAT and states in which more than 45% took it. The scatterplot and correlation ($r = -0.84$) both suggest that there is a fairly strong, negative linear relationship between these two numerical variables for these South Census Region states; as the percent of the 2017 graduating class who took the exam increases, the mean total SAT score tends to decrease for these states. When exploring this relationship separately for each group, the correlations still suggest fairly strong, negative linear relationships ($r = -0.76$ and $r = -0.74$, respectively, for states with less than 15% and more than 45% of their 2017 graduating classes taking the SAT). These bivariate relationships suggest that the bimodal distribution of mean total SAT scores we discovered in Fig. 7 is at least partially explained by the percent of the 2017 graduating seniors in a South Census Region state who took the SAT, allowing us to uncover additional information beyond what was offered by the univariate summaries alone.

Conclusion

When dealing with data in any context, an important first step is to explore properties of the data using descriptive statistics. We have discussed how variables can be divided into two types, categorical and numerical, and what tools are appropriate for each type. For categorical variables, counts and relative frequencies are key and are often displayed using frequency tables and bar charts. For numerical variables, we seek to describe features of the variable's distribution including its center, variability, and shape, and how the location of individual observations relate to that distribution. By developing meaningful summary statistics about variables in a data set, descriptive statistics not only allow for their own interpretations, but also lay the groundwork for further analysis such as exploratory data analysis, statistical modeling, and statistical inference; for further reading, see Çetinkaya-Rundel and Hardin (2021). Using and interpreting descriptive statistics helps us move beyond complex lists of variables and values to make meaningful observations about the population of interest and the variables measured.

References

- Çetinkaya-Rundel, M., Hardin, J., 2021. Introduction to Modern Statistics, first ed. OpenIntro [online]. <https://www.openintro.org/book/ims/>.
- Snyder, T.D., de Brey, C., Dillow, S.A., 2019. Digest of Education Statistics 2017 (NCES 2018-070). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.
- The Art and Science of Learning From Data Web Apps, <https://artofstat.com/web-apps>.
- United States Census Bureau, 2013. Census Regions and Divisions of the United States. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

Foundational methods: descriptive statistics: bivariate and multivariate data (correlations, associations)

Debbie L. Hahs-Vaughn, University of Central Florida, Orlando, FL, United States

© 2023 Elsevier Ltd. All rights reserved.

Correlations and associations	734
Introduction	734
History of correlations	734
Types of bivariate correlations	735
Parametric correlations	735
Nonparametric correlations	741
Considerations with correlations	747
Correlation and causation	747
Spurious correlations	747
Restriction of range	747
Acknowledgment	748
References	748

Nomenclature

- V Cramer's phi
- τ Kendall's tau
- ρ_ϕ Phi coefficient
- ρ_{XY} or ρ Population Pearson correlation coefficient
- σ_{XY} Population covariance
- r_{XY} or r Sample Pearson correlation coefficient
- s_{XY} Sample covariance
- ρ_s Spearman's rho
- r_t Tetrachoric correlation

Correlations and associations

Introduction

It can be argued that correlations are the first statistical concept that most individuals learn, albeit perhaps unbeknownst at the time. From a very young age, children learn about, and through, relationships, and they continue to learn through relationships as they develop (e.g., consequences that may result from good or poor behaviors). In the statistical world, correlations are arguably the foundation of basic as well as advanced statistics given that relationships are either implicit or explicit in these statistics (Chen and Popovich, 2011).

Correlation coefficients are the statistics that allow us to describe relationships in terms of strength and direction. Correlation coefficients generally range in absolute value from 0 to 1, with statistics closer to an absolute value of 1 reflecting stronger relationships. Directionality can be assessed by the sign of the correlation. Positive correlations suggest that values that increase on one variable also increase on the second variable. Negative correlations suggest that values that increase on one variable decrease on the second variable. Correlations at or near zero suggest little or no relationship; in other words, there is no discernible pattern to the relationship between variables.

Correlations have many applications and can be used as descriptives statistics or inferential statistics. This article will focus on bivariate (i.e., two variables) parametric and non-parametric correlation indexes as well as characteristics of correlations, and will begin with a concise history of correlations.

History of correlations

Karl Pearson, the scholar for whom the Pearson product moment correlation is named, is considered by many to be the founder of scientific statistics (Eves, 2002) and the first modern statistician (Yasmin, 2007). Pearson (1887–1936) was a scientist in England and directed the University of London's Francis Galton Laboratory for National Eugenics (Eves, 2002). Pearson's work was

influenced by Sir Francis Galton (1822–1911) (Yasmin, 2007; Blyth, 1994), the scientist and cousin of Charles Darwin who created new statistical methods for studying heredity (Eves, 2002), including correlation and regression (Blyth, 1994).

Within the history of statistics, the discovery of correlation and regression is regarded as one of the most important developments (Fienberg, 1992; Stigler, 1986). Pearson credited Galton for the creation of correlation, noting no preceding work by Gauss or Bravais. Pearson also recognized that Galton himself realized the unprecedented discovery with which he had made (Pearson, 1920). As noted by Pearson (Pearson, 1920):

If you think Galton did not appreciate the width of his new methods, you must turn to the last paragraph of his *Introduction to the Natural Inheritance*.

The conclusions cannot, however, be intelligibly presented in an introductory article. They depend on ideas that must first be well comprehended, and which are now novel to the large majority of readers and unfamiliar to all. But those who care to brace themselves to a sustained effort, need not feel much regret that the road to be travelled over is indirect, and does not admit of being mapped beforehand in a way they can clearly understand. It is full of interest of its own. It familiarizes us with the measurement of variability, and with curious laws of chance that apply to a vast diversity of social subjects. This part of the inquiry may be said to run along a road on a high level that affords wide views in unexpected directions, and from which easy descents may be made to totally different goals to those we have now to reach. I have a great subject to write upon

Galton (1889).

According to Pearson (Pearson, 1920), it was Galton's book (Galton, 1889) that enticed Pearson, along with two other scholars (Walter F. R. Weldon and Francis Y. Edgeworth) to study correlation. Weldon applied correlation to study organs in shrimps and other species, while Edgeworth applied the methods to moral or social sciences (Stephen, 1978). Interestingly, Galton termed the statistic the "index of co-relation" (Galton, 1889), Weldon termed it "Galton's function", and Edgeworth replaced both these terms with "coefficient of correlation" (Pearson, 1920), the latter term of which has seemed to stand the test of time.

Although Galton is credited for introducing correlations in 1888, Pearson is credited for refining many of the ideas introduced by Galton (Blyth, 1994). Pearson's student, G. U. Yule, found that if the regression was linear, the regression slope was still found to be $\frac{\rho\sigma_Y}{\sigma_X}$ (Blyth, 1994), where the correlation coefficient ρ is the regression slope when both variables were scaled by their standard deviations, σ_Y is the standard deviation of Y , and σ_X is the standard deviation of X . Thus, by the early 1900s, the application of regression and correlation to the linear model had been made (Blyth, 1994). Pearson realized the need, and worked on albeit unsuccessfully, generalizing the correlation to nonlinear models (Pearson, 1920). Much of the continued development of the correlation coefficient after this time is the result of Sir Ronald Aylmer Fisher, including calculating the sampling distribution of an estimate of the correlation coefficient (Fisher, 1915) and demonstrating the value of the correlation in clarifying aspects of heredity and genetics (Fisher, 1918).

The history of correlations is quite interesting and not without controversy (e.g., Norton and Pearson, 1976). Readers interested in learning more are directed to the references cited within as what is presented here is but a small snippet of the historical underpinnings.

Types of bivariate correlations

There are a number of different types of bivariate correlation indexes (i.e., correlation coefficients computed for two variables). As we will see, most were developed to overcome limitations with the Pearson correlation, the correlation with which we will begin the discussion. Likely the most common correlation coefficient, the Pearson correlation coefficient is a parametric correlation coefficient.

Parametric correlations

Fisher formulated the framework for parametric statistical inference (Fisher, 1922) (see Hald, 2007, for a detailed account of the history of parametric statistics during this time). However, it was not until Wolfowitz (1942) that the term *parametric* was coined, as he distinguished parametric from nonparametric functional forms. Parametric statistics are those for which the variables used within the statistical procedure have a functional form that is known. In other words, parametric statistics are those statistical procedures that assume the data, which is sampled from the population, can be sufficiently modeled by a probability distribution with a parameter set that is fixed (Geisser, 2006). The normal distribution is the most common distributional assumption for parametric procedures. Within the umbrella of correlations, there exist both parametric and non-parametric (i.e., distribution-free) correlations.

Pearson product moment correlation

The Pearson product moment correlation (hereinafter referred to as Pearson), as noted previously, is credited in large part to Karl Pearson's work proceeding Galton. The Pearson correlation, also referred to as the Galton-Pearson correction (Blyth, 1994), is arguably the most utilized correlation coefficient, and there is some research to support this. For example, a review of more than 18,000 psychology articles published in 2014 found that an effect size measure of some type was reported in about 25% of the publications

(de Winter et al., 2016). Of correlations reported, Pearson was reported about twice as often as Spearman's rho (discussed in a later section) (de Winter et al., 2016).

The Pearson correlation is most appropriate to compute when both variables are at least interval in measurement scale and when there is a linear (as compared to non-linear) relationship between those variables (specifically, a linear relationship between X and Y , in other words, a bivariate normal distribution, is assumed for the Pearson), although there are special cases of the Pearson when this is not the case. (A variable that is interval in measurement scale is one where the units can be ordered and equal differences between the values imply equal distance in terms of the characteristic being measured (Hahs-Vaughn and Lomax, 2020a)) An initial understanding of covariance is required, prior to the introduction of correlation.

Covariance

The variance and correlation share commonalities in that the correlation is simply the standardized covariance. The population covariance is denoted by σ_{XY} and calculated as follows:

$$\sigma_{XY} = \frac{\sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)}{N}$$

where X_i and Y_i are the scores for case i on variables X and Y , respectively, and μ_X and μ_Y are the population means for variables X and Y , respectively. N represents the population size.

The sample covariance, s_{XY} , is computed as follows:

$$s_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}$$

where $\bar{X}$ and $\bar{Y}$ are the sample means for variables X and Y , respectively, and n is the sample size. The denominator of $n-1$ is the correction necessary to obtain an unbiased estimate of the population parameter.

The formula for the covariance is essentially an average of the product of the paired deviation scores. When variables X and Y are positively related, then the deviation scores will tend to be of the same sign, their products will tend to be positive, and the covariance will be a positive value (i.e., $\sigma_{XY} > 0$). When variables X and Y are negatively related, then the deviation scores will tend to be of opposite signs, their products will tend to be negative, and the covariance will be a negative value (i.e., $\sigma_{XY} < 0$). Finally, in the case of unrelated variables X and Y , then the deviation scores will consist of both the same and opposite signs, their products will be both positive and negative and sum to zero, and the covariance will be zero (i.e., $\sigma_{XY} = 0$).

A computational formula for the population covariance, which is less wieldy than the previously provided formula, is as follows:

$$\sigma_{XY} = \frac{N \left(\sum_{i=1}^N X_i Y_i \right) - \left(\sum_{i=1}^N X_i \right) \left(\sum_{i=1}^N Y_i \right)}{N^2}$$

where the first summation involves the cross-product of X multiplied by Y for each case summed across all N cases, and the other terms should be familiar.

The corresponding computational formula for the sample covariance is as follows:

$$s_{XY} = \frac{n \left(\sum_{i=1}^n X_i Y_i \right) - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n(n - 1)}$$

where the denominator is $n(n-1)$ so as to yield an unbiased sample estimate of the population covariance.

The covariance serves a useful purpose in many contexts, but the interpretation can be challenging as the magnitude of a covariance can take on any value. This results in further difficulty in making comparisons between covariance values. A large covariance may reflect a strong relationship, but it may also simply be an artifact of the metric of the variables. Correlation coefficients, which are standardized covariance values (as noted earlier), measure the relationship between variables and are easier to interpret than covariances.

Computation of the Pearson correlation coefficient

The Pearson correlation coefficient can be computed in various ways. Using standardized scores, the population Pearson correlation can be computed as follows:

$$\rho_{XY} = \frac{\sum_{i=1}^N (z_X z_Y)}{N}$$

where z_X and z_Y are the z scores (i.e., standardized scores) for variables X and Y respectively, whose product is taken for each case, and then summed across all N cases.

Derived from the standardized equation, the Pearson can also be computed as follows, where the numerator is the covariance and the denominator is the product of the standard deviations. The denominator provides the standardization of the covariance, and in this sense, the Pearson correlation is essentially a standardized covariance.

$$\rho_{XY} = \frac{\frac{\sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)}{N}}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

The computational formula for the Pearson is written as follows:

$$\rho_{XY} = \frac{N \left(\sum_{i=1}^N X_i Y_i \right) - \left(\sum_{i=1}^N X_i \right) \left(\sum_{i=1}^N Y_i \right)}{\sqrt{\left[N \left(\sum_{i=1}^N X_i^2 \right) - \left(\sum_{i=1}^N X_i \right)^2 \right] \left[N \left(\sum_{i=1}^N Y_i^2 \right) - \left(\sum_{i=1}^N Y_i \right)^2 \right]}}$$

The sample correlation is denoted by r_{XY} (or r , where the correlated variables are implied). The formulas for the sample correlation r_{XY} and the population correlation ρ_{XY} are essentially the same. The only difference is notational, where n is substituted for N . For example, the computational formula for the sample correlation is as follows:

$$r_{XY} = \frac{n \left(\sum_{i=1}^n X_i Y_i \right) - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{\sqrt{\left[n \left(\sum_{i=1}^n X_i^2 \right) - \left(\sum_{i=1}^n X_i \right)^2 \right] \left[n \left(\sum_{i=1}^n Y_i^2 \right) - \left(\sum_{i=1}^n Y_i \right)^2 \right]}}$$

The sample correlation can also be computed as follows using the covariance and standard deviations, where the covariance of X and Y is the numerator and the denominator reflects the standard deviations of X and Y .

$$r_{XY} = \frac{s_{XY}}{(s_X)(s_Y)}$$

Unlike the sample variance and covariance, the sample correlation has no correction for bias.

As an example, in [Table 1](#) we have data from a hypothetical sample of 20 students who have been assessed on academic achievement (as measured by GPA, column 1, $M = 3.127$, $SD = 0.255$) and emotional intelligence (column 2, $M = 45.742$, $SD = 11.240$). Using the formula for the correlation coefficient which includes the covariance ($s_{XY} = 0.872$) and standard deviation terms, we find a correlation coefficient of approximately 0.304.

$$r_{XY} = \frac{s_{XY}}{(s_X)(s_Y)} = \frac{0.872}{(11.240)(0.255)} = 0.304$$

This coefficient reflects a slightly stronger relationship between emotional intelligence and academic achievement than found by [MaccAnn et al. \(2020\)](#) who found an overall effect of $\rho = 0.20$ in their meta-analysis of 158 citations.

Interpretations of the Pearson correlation

There are many interpretations that can be made from the correlation coefficient. The *sign* of the Pearson correlation denotes the direction of the relationship (e.g., positive or negative), and the *value* of the Pearson denotes the strength of the relationship. The Pearson ranges within a scale from -1.00 to $+1.00$, where -1.00 indicates a perfect negative relationship, 0 indicates no relationship, and $+1.00$ indicates a perfect positive relationship. The larger the correlation coefficient (in absolute value terms), the more interdependence between the variables is suggested.

The squared Pearson is the *coefficient of determination* and is interpreted in many ways, all of which are stating the same concept: the proportion of shared variance; the proportion of variance in X accounted for by Y (and vice versa); the proportion of variance explained; and the proportion of variance of X predicted by Y (and vice versa). While the coefficient of determination is interpreted as a proportion (and thus an $r^2 = 0.10$ is one-half of the shared variance of $r^2 = 0.20$), r cannot be interpreted similarly (i.e., $r = 0.10$ cannot be interpreted as one-half the relationship of $r = 0.20$).

Hypothesis testing

The Pearson correlation can be used as a descriptives statistic or as an inferential statistic. Inferences can be made to determine the extent to which a single sample correlation differs from zero (i.e., no relationship). An example would be to test if the correlation between age and income differs from zero. In this example, a correlation of zero would indicate there is no relationship between age and income. Inferences can also be made to examine the extent to which two (or more than two) independent sample correlations are different. An example of this test of inference would be the examination of the extent to which the correlation between age and income differs for individuals who have (as compared to do not have) a college degree. Tests of inference between two dependent correlations can also be computed. An example would be when multiple measures are taken on a sample, and the researcher wants to determine if the relationship between two of the measures differs from the relationship between two different measures. Readers

Table 1 Academic achievement and emotional intelligence scores.

Academic achievement	Emotional intelligence	Academic achievement (rank) ^a	Emotional intelligence (rank) ^b	Academic achievement (ordinal) ^c	Emotional intelligence (ordinal) ^d	Academic achievement (binary) ^e	Emotional intelligence (binary) ^f
2.938	51.173	6.00	13.00	1.00	4.00	0.00	1.00
3.352	34.816	14.00	4.00	3.00	2.00	1.00	0.00
3.542	61.304	20.00	19.00	4.00	5.00	1.00	1.00
3.424	51.538	17.00	14.00	3.00	4.00	1.00	1.00
3.041	48.014	10.00	10.00	2.00	3.00	1.00	0.00
3.504	54.895	19.00	16.00	4.00	4.00	1.00	1.00
3.371	63.016	16.00	20.00	3.00	5.00	1.00	1.00
3.008	26.567	8.00	2.00	2.00	1.00	1.00	0.00
2.839	55.193	2.00	17.00	1.00	4.00	0.00	1.00
3.022	43.616	9.00	9.00	2.00	3.00	1.00	0.00
2.922	51.992	5.00	15.00	1.00	4.00	0.00	1.00
3.116	38.732	11.00	6.00	2.00	2.00	1.00	0.00
3.211	48.778	13.00	11.00	2.00	3.00	1.00	0.00
2.870	25.346	4.00	1.00	1.00	1.00	0.00	0.00
2.840	40.631	3.00	7.00	1.00	3.00	0.00	0.00
3.119	30.030	12.00	3.00	2.00	2.00	1.00	0.00
3.425	35.933	18.00	5.00	3.00	2.00	1.00	0.00
3.364	60.075	15.00	18.00	3.00	5.00	1.00	1.00
2.676	50.928	1.00	12.00	1.00	4.00	0.00	1.00
2.957	42.259	7.00	8.00	1.00	3.00	0.00	0.00

^aAcademic achievement (rank): rank ordered values.^bEmotional intelligence (rank): rank ordered values.^cAcademic achievement (ordinal): four ordered categorized values, 1 = less than 3.000; 2 = 3.000–3.249; 3 = 3.250–3.499; 4 = 3.50 or greater.^dEmotional intelligence (ordinal): five ordered categorized values, 1 = 20.000–29.999; 2 = 30.000–39.999; 3 = 40.000–49.999; 4 = 50.000–59.999; 5 = 60.000–69.999.^eAcademic achievement (binary): 0 = less than 3.00; 1 = 3.00 or greater.^fEmotional intelligence (binary): 0 = less than 50.00; 1 = 50.00 or greater.

interested in more in-depth coverage of hypothesis testing with correlations are referred to any number of excellent resources (e.g., [Hahs-Vaughn and Lomax, 2020a,b](#); [Wilcox and Muska, 2001](#); [Wilcox, 2017](#)), in addition to encyclopedia articles “Statistical Inference: Frequentist Approaches” and “Statistical Inference: Bayesian Approaches” for coverage of hypothesis testing/statistical inference more generally.

Assumptions of the Pearson correlation coefficient

The Pearson correlation coefficient is a parametric statistic. As such, there are distributional assumptions associated with it. Specifically, a linear relationship between X and Y , in other words, a bivariate normal distribution, is assumed for the Pearson. When linearity is violated, the Pearson correlation is not robust ([Wilcox, 2017](#)), and more specifically, “highly not robust... even a single aberrant point can alter r , the usual estimate of ρ , by a large amount. One practical consequence is that an association can be masked even by a single outlier” ([Wilcox, 2004](#)). In addition to outliers, additional data features that can affect the magnitude of the Pearson correlation include non-linearity, restriction of range, the magnitude of residuals, and rotating points (i.e., the magnitude of the slope around which points cluster) ([Wilcox, 2016](#)). Pearson correlation coefficient of zero may occur even if there is a strong non-linear relationship. As a result of these limitations, a number of alternatives to the Pearson have been developed, some of which are more robust than others. Many of these alternative coefficients are presented in [Nonparametric Correlations](#) section.

Visualizing relationships

When X and Y are measured on at least an interval scale, scatterplots can be valuable in understanding the relationship between variables, both the strength and the direction. Using Cohen’s conventions for interpreting correlations as an effect size ([Cohen, 1988](#)), [Figs. 1–4](#) show examples of weak ($r = 0.10$), moderate ($r = 0.30$), and strong ($r = 0.50$) correlations. [Figs. 5 and 6](#) provide illustrations of very strong positive relationships ($r = 0.95$), and [Fig. 7](#) provides an illustration of a strong negative relationship ($r = -0.70$).

When $r = 0.10$ (see [Fig. 1](#)), there is a lack of definition to the points in the scatterplot. In other words, relatively no concentration to the points (seemingly random) and no identifiable trend to suggest either a positive or negative relationship.

In the case of a moderate relationship, such as $r = 0.30$ (see [Fig. 2](#)), the points in the scatterplot adhere a bit tighter together such that there begins to be a discernible relationship evident. In this case, with a positive correlation, as the values of X increase, the values of Y decrease. This suggests a positive relationship. Additionally, there is more definition in the figure such that the points are adhering more closely together in a pattern as compared to those in [Fig. 1](#), although there is still somewhat of a loose scatter to the points.

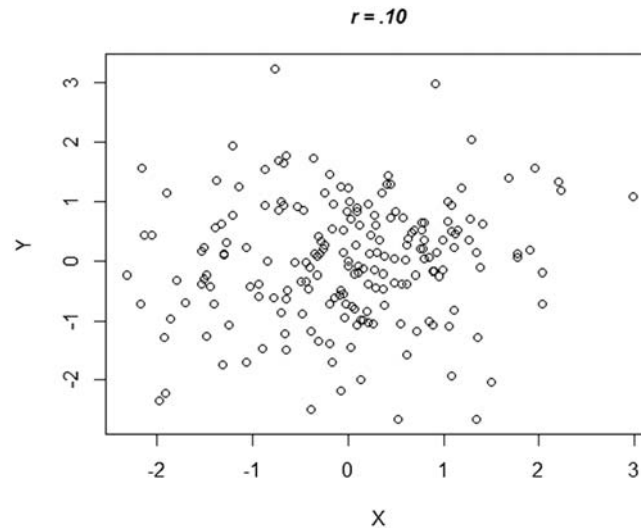


Fig. 1 Scatterplot ($r = 0.10$).

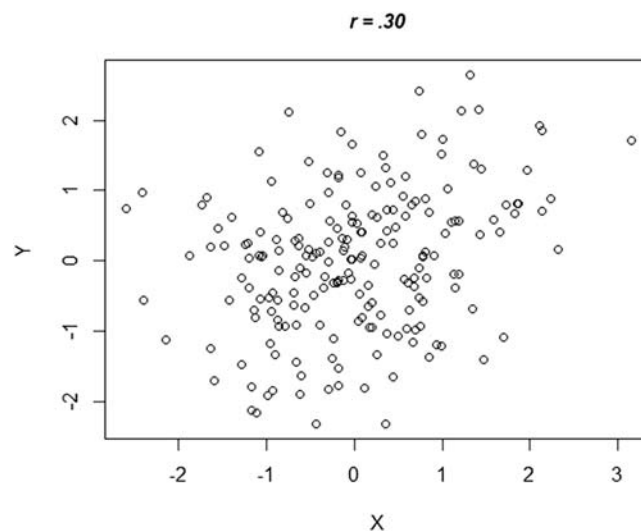


Fig. 2 Scatterplot ($r = 0.30$).

With a smaller sample size ($n = 20$ in Fig. 3, see data in Table 1; vs. $n = 200$ in Fig. 2), the scatterplot of the relationship between emotional intelligence and academic achievement ($r = 0.30$ in both Figs. 2 and 3) becomes somewhat more difficult to discern in terms of the direction and strength given the relatively looseness of points coupled with fewer points.

When $r = 0.50$ (Fig. 4), the concentration of points in the scatterplot is tighter, providing a better visual to understand the strength and direction (i.e., positive) of the relationship.

When $r = 0.95$ (Fig. 5), the concentration of points in the scatterplot helps the reader to understand the strength of the relationship (i.e., the tight concentration suggests a very strong association) as well as the direction of the relation (i.e., positive; as values on X increase, values on Y also increase).

Fig. 6 presents a scatterplot with variables again correlated strongly ($r = 0.95$). However, the scales for X and Y differ. X has a mean of 10 and standard deviation of 1, and Y has a mean of 30 and standard deviation of 5. Thus, even when the correlation coefficient is very strong, the scales of the variables may be very dissimilar. In other words, the metrics of the variables do not need to be similar to manifest a relationship.

A strong negative relationship is presented in Fig. 7. The negative relationship is reflected in the direction of the points; as the values for X increase, the values for Y decrease (and vice versa). The concentration of points in the scatterplot, while not as tight as those in Figs. 5 and 6, suggest a strong relationship between the variables.

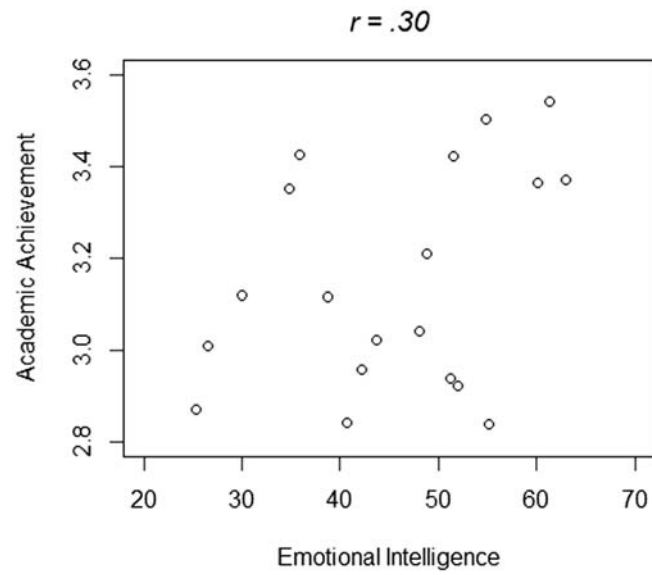


Fig. 3 Scatterplot of academic achievement and emotional intelligence ($r = 0.30$).

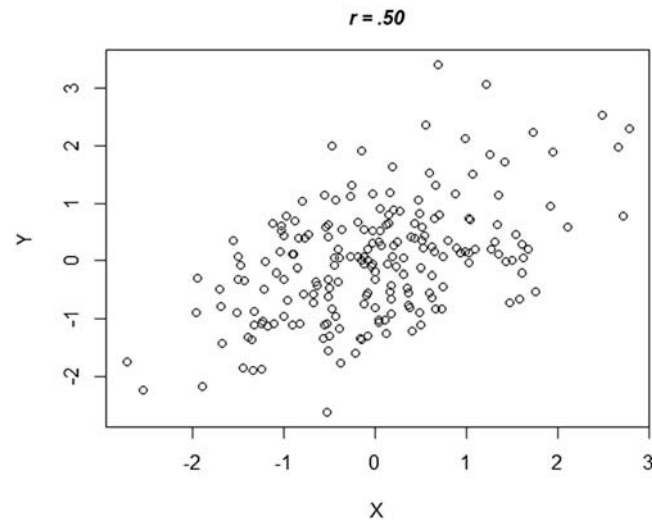


Fig. 4 Scatterplot ($r = 0.50$).

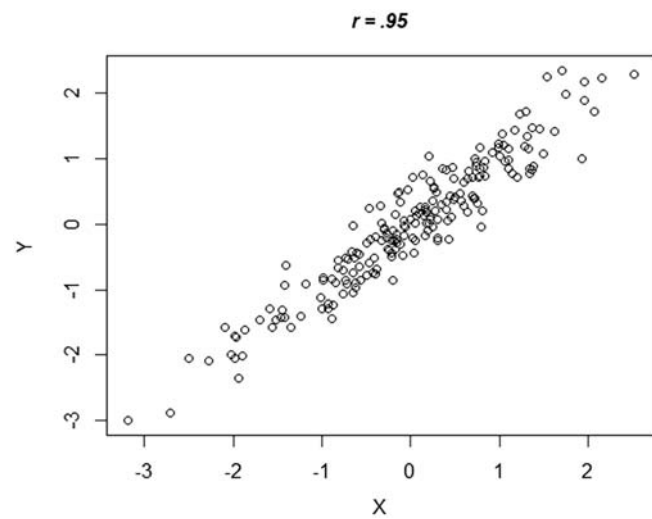


Fig. 5 Scatterplot ($r = 0.95$).

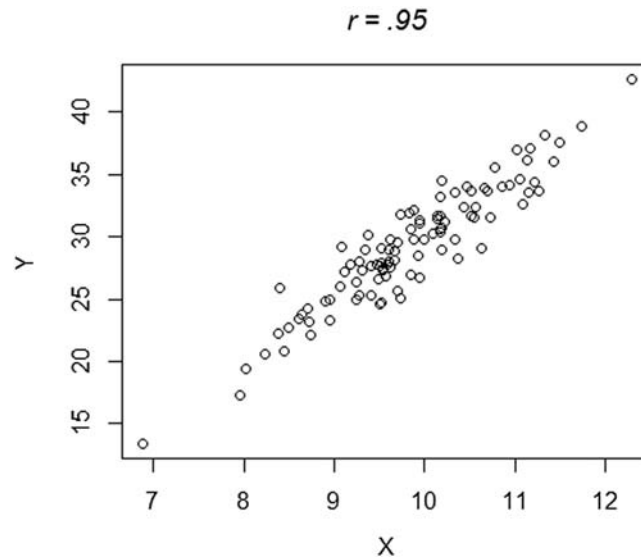


Fig. 6 Scatterplot with variables of differing scales ($r = 0.95$).

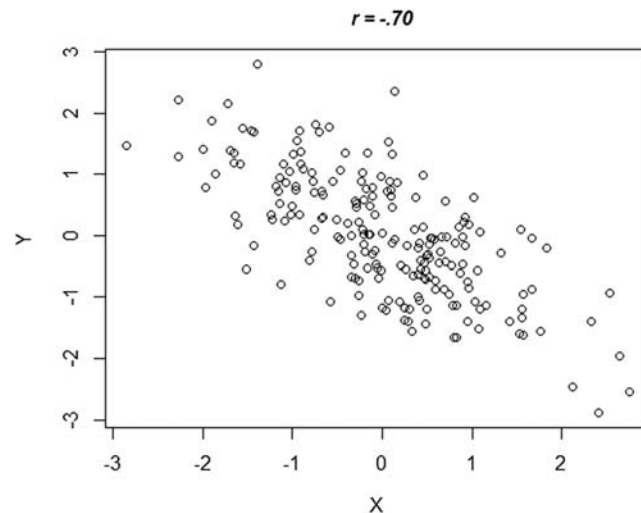


Fig. 7 Scatterplot ($r = -0.70$).

Nonparametric correlations

There are a number of correlation coefficients that fall under the umbrella of nonparametric correlations. Nonparametric correlations have fewer associated assumptions as compared to parametric correlations. Some of the most common that we will discuss include: phi; Cramer's phi; Spearman; and Kendall's tau. Many additional alternatives to the Pearson are then also presented.

Phi coefficient

The phi coefficient, ρ_ϕ is an appropriate correlation to use when both variables being correlated are dichotomous in nature (i.e., binary categories coded as 0 and 1). The phi represents the geometric mean of the percentage differences across rows and columns in variables that are nominal in scale (Hernes, 1970).

When correlating two binary variables, consider a 2×2 contingency table as shown in Table 1. The columns correspond to the two levels of the first variable (coded as 0 and 1) and the rows correspond to the two levels of the second variable (again coded as 0 and 1). The cells within the contingency table indicate the frequencies for the particular combinations of the levels of the two variables. The phi is computed as follows (Pearson, 1904), where the letters represent the frequencies of the respective cells within the contingency table (see Table 2):

$$\rho_\phi = \frac{(bc - ad)}{\sqrt{(a + c)(b + d)(a + b)(c + d)}}$$

Table 2 Example contingency table.

X	Y		
	0	1	
1	a	b	a + b
0	c	d	c + d
	a + c	b + d	a + b + c + d

A perfect relationship for the phi occurs with either of the following conditions: (a) empty cells are present in either of the two diagonally opposite cells; or (b) the margins of each row and column are identical (e.g., $a + b = a + c$ or $b + d = c + d$).

The phi is statistically equivalent to the Pearson, and is thus considered a special case of the Pearson correlation. In comparison to Pearson which can range from -1.0 to $+1.0$, due to the binary nature of the data, the range for phi often does not reach -1.0 or $+1.0$. A challenge in using the phi is that the value of the phi coefficient is commensurate with the table marginals, with less precise results when the marginals are more unequal (Liu, 1980). Standardization of the table marginals has been suggested so that the phi coefficient can be applied to comparisons (Deming and Frederick, 1940). An iterative technique, for 2×2 and larger $R \times C$ tables, to standardize the table marginals and allow for cross-table comparison has been offered (Bonacich, 1972; Liu, 1980).

In addition to early work by Pearson (1904), readers interested in early discussions of the phi coefficient are also referred to Yule (1912).

As an example, correlating the binary measures of academic achievement (0 = below 3.000; 1 = above 3.000) and emotional intelligence (0 = below 50.000; 1 = 50.000 or above), we find a phi coefficient of -0.179 (see Table 3 for contingency table).

$$\rho_{\phi} = \frac{(bc - ad)}{\sqrt{(a + c)(b + d)(a + b)(c + d)}} = \frac{(5)(3) - (4)(8)}{\sqrt{(4 + 3)(5 + 8)(4 + 5)(3 + 8)}} = \frac{-17.00}{94.916} = -0.179$$

While binary (e.g., pass/fail, yes/no) and ordered categories (e.g., Likert scales) are common to find in practice, it is not recommended practice to force order continuous data into binary or ordered categories as information on the responses is lost in the process. Categorizing the values into dichotomous or polytomous ordered categories has been done within this article for illustrative purposes only. As seen here, this loss of information results in a different estimated correlation coefficient relative to what was computed using the Pearson correlation coefficient.

Cramer's V

Cramer's V is an appropriate correlation to use when the variables being correlated are multichotomous in nature (Cramér, 1946). Cramer's V can be computed when both variables are nominal (with at least one variable having more than two categories) or when one variable is nominal and the other variable is ordinal (and at least one variable has more than two categories). Although the chi-squared statistic and contingency coefficient can be computed in the same situations (in terms of measurement scales of variables) as Cramer's V, Cramer's V is the preferred statistic to compute (Chen and Popovich, 2011) (albeit challenges that exist as will be noted subsequently). Cramer's V can range from 0 when X and Y are unrelated, to $+1.0$ when at most one non-zero probability appears in every row or column (assuming non-zero marginal probabilities), and is computed as follows, where χ^2 refers to chi squared and $\min(\text{row}, \text{column})$ refers to the smallest number of rows or columns in the contingency table:

$$V = \sqrt{\frac{\chi^2}{n[\min(\text{row}, \text{column}) - 1]}}$$

There are challenges with Cramer's V. Likely most problematic is that Cramer's V has no meaningful interpretation; a large coefficient does not necessarily indicate a strong association, and likewise, a small coefficient does not necessarily indicate a weak association. As a result, several alternatives to Cramer's V have been proposed that offer meaningful interpretations (Goodman and Kruskal, 1979; Kvalseth, 2018). Cramer's V also has challenges in notation. While $\hat{V}^2$ was originally proposed by Cramér

Table 3 Emotional intelligence and academic achievement contingency table.

Emotional intelligence	Academic achievement		
	0 (below 3.0)	1 (3.0 or above)	
1 (50 or above)	a = 4	b = 5	a + b = 9
0 (below 50)	c = 3	d = 8	c + d = 11
	a + c = 7	b + d = 13	a + b + c + d = 20

(1946), this notation has not been universally adopted. Many software packages use V , but not all. Adding to the notation confusion is the squared canonical correlation coefficient, also noted as $\hat{V}^2$ (Kendall and Stuart, 1958).

Computationally, because χ^2 is an element in the calculation for Cramer's, small sample size and small cell frequencies that plague χ^2 are also problematic with Cramer's V . When small sample size and small cell frequencies are present for chi-squared, the approximation of chi-squared will be inaccurate and Cramer's V will be over or underestimated, leading to Type I errors (Kvalseth, 2018). Additionally, if either the number of rows or columns changes, small changes in proportions can create a substantial change in the Cramer's V coefficient (Kvalseth, 2018).

It was noted previously that Cramer's V can reach +1.0 given certain marginal probability values. When one or more marginal probabilities are zero, Cramer's V can be impacted. For example, Cramer's V may not reach +1.0 even if there is a perfect relationship and the number and rows and columns are the same (Kvalseth, 2018). Additionally, Cramer's V coefficients are only comparable in situations where the contingency tables which are being compared have the same values of $\min(\text{row}, \text{column})$.

As an example, correlating the binary measure of academic achievement (0 = below 3.00; 1 = 3.00 or above) and ordered category measure of emotional intelligence (5 categories), we find a Cramer's V coefficient of nearly 0.60 (specifically, $V = 0.577$):

$$V = \sqrt{\frac{\chi^2}{n[\min(\text{row}, \text{column}) - 1]}} = \sqrt{\frac{6.667}{20[2 - 1]}} = 0.577$$

In comparison, correlating the binary measure of emotional intelligence (0 = below 50; 1 = 50 or greater) and ordered category measure of academic achievement (4 categories), we find a slightly larger Cramer's V coefficient (specifically, $V = \sqrt{\frac{8.225}{20[2-1]}} = 0.641$). While binary (e.g., pass/fail, yes/no) and ordered categories (e.g., Likert scales) are common to find in practice, it is not recommended practice to force order continuous data into binary or ordered categories as information on the responses is lost in the process.

Spearman's rho

Spearman's rho rank correlation coefficient was proposed by Spearman as a correction for attenuation in Pearson's coefficient (Spearman, 1904), and is appropriate to compute when both variables are ordinal in scale. In his first discussion of this correlation, Spearman not only presented the first approximation of the coefficient but also decried the benefits of ranks, which included, among others, that this method is not influenced by outliers as happens in the case of Pearson's coefficient (Spearman, 1904).

The history of Spearman's rho is quite lively and scholars have noted that "no one really 'discovered' SRC [Spearman's rho coefficient], or perhaps at best that SRC somehow mysteriously emerged from the various bouts of fisticuffs between Spearman and Pearson" (Lovie, 1995). Lovie (1995) concludes that Spearman brought rank data to the forefront and provided the first discussion of Spearman's rho coefficient, while Pearson named Spearman's rho as ρ and explored the properties of the coefficient. Kendall, however, revived the coefficient in the late 1930s and early 1940s, repositioned it within nonparametrics, and named it Spearman's rho coefficient (Lovie, 1995).

Spearman's correlation is a measure of relationship of ranked data and can be computed as follows (assuming no tied ranks) (Hotelling and Pabst, 1936):

$$\rho_s = 1 - \frac{6 \left[\sum_{i=1}^N (X_i - Y_i)^2 \right]}{N(N^2 - 1)}$$

An example of a tied rank would be if two cases had the same value on either variable being corrected. When one or more ties in rank exist, this formula is only approximate, depending on the number of ties. Thus, when there are ties, especially when there are more than a few ties, many researchers recommend using Kendall's tau as an alternative correlation (e.g., Wilcox, 1995). Spearman's coefficient can also be computed via Pearson on the ranked X and Y .

Like the Pearson correlation, Spearman's rho can range in value from -1.0 to $+1.0$ indicating the degree of monotonic relationship between the variables. When the correlation is 0, there is no relationship in how the variables have their data rank ordered. A value of $+1.0$ indicates that the ordered points in the two variables are exactly the same. A value of -1.0 indicates that the ordered points in the two variables are exactly opposite (i.e., one data point ranking first in one variable and one data point ranking last in the other variable). The sign of the coefficient can be interpreted as with the Pearson. Negative correlation coefficients indicate that as the values for one variable increase, the values for the second variable decrease. Positive correlation coefficients indicate that as one variable increases in value, the value of the second variable as well.

Spearman's rho and Kendall's tau, both functions of rank, have been shown to correlate at least 0.98, approaching 1 as the sample size approaches infinity, when X and Y are independent (Kendall and Stuart, 1958).

Had the data in columns 1 and 2 of Table 1 (the continuous values) been treated as rank ordered (columns 3 and 4) instead of continuous, a similar relationship would be present with Spearman ($\rho_s = 0.263$) as what as computed using continuous data with Pearson ($r = 0.30$). If the data were ordered categories (columns 5 and 6 of Table 1), the Spearman coefficient weakens slightly ($\rho_s = 0.220$).

Kendall's tau

Kendall's tau (Kendall, 1938), τ , is also appropriate when both variables are ordinal in scale and is also computed from ranked data. With Kendall's tau, the values for one variable are rank ordered and then the order of the second variable is examined to see if/how many pairs of values are out of order. Kendall's tau can be computed as follows when there are no ties in rank (denoted as τ_a), where P represents the number of concordant pairs (i.e., agreeing) and Q represents the number of discordant (i.e., disagreeing) pairs. More specifically, with any pair of observations, (X_A, Y_A) and (X_B, Y_B) , the pairs are concordant if the sort order of the observations agree (e.g., $X_A > X_B$ and $Y_A > Y_B$ or $X_A < X_B$ and $Y_A < Y_B$). If the sort order does not agree, in other words, if the observations are in reverse order or opposite directions (e.g., $X_A > X_B$ and $Y_A < Y_B$ or $X_A < X_B$ and $Y_A > Y_B$), the pairs are discordant.

$$\tau_a = \frac{2(P - Q)}{n(n - 1)}$$

When there are ties, Kendall's tau- b can be computed. In this formula, P represents the number of concordant pairs and Q represents the number of discordant pairs. X_0 represents the number of paired observations that are tied only on X , and Y_0 represents the number of paired observations tied only on Y .

$$\tau_b = \frac{P - Q}{\sqrt{(P + Q + X_0)(P + Q + Y_0)}}$$

Kendall's is interpreted in the same fashion as Spearman's. When all scores of the two variables are ordered the same (i.e., no scores of the two variables are out of order), a perfect positive correlation (+1.0) is achieved. When scores of the two variables are in exact opposite order, a perfect negative correlation (−1.0) is obtained. When there is no relationship in how the variables have their data rank ordered, the correlation is 0. Negative correlation coefficients indicate that $P - Q$ is negative. Positive correlation coefficients indicate that $P - Q$ is positive. When tied ranks exist, Kendall's τ_b (Kendall, 1945) and Stuart's τ_c (Stuart, 1953), which are modifications of Kendall's τ_a , can be computed. When there are many tied ranks, Goodman and Kruskal's gamma correlation (Goodman and Kruskal, 1954) can be computed.

While Spearman's rho and Kendall's tau are similar in that they are both functions of ranked data, they are calculated differently. As such, it is not uncommon to find different results when calculating Spearman's and Kendall's tau. Kendall claimed preference for tau, relative to rho, for practical and theoretical reasons including a simpler interpretation for tau when estimating the population parameter (Kendall and Gibbons, 1990; Kendall, 1970). While both are appropriate when ordinal data are being correlated, Kendall's tau (as compared to Spearman's rho) better estimates the population correlation coefficient given sample data (Howell, 2010), especially when the sample sizes are smaller (e.g., $n < 10$). Kendall's tau has been shown to be more powerful than Spearman's rho for testing the null hypothesis that the correlation coefficient equals zero (van der Waerden, 1969).

Instances where the ordinal variables being correlated are affected by a third variable (e.g., blocking variable) have been discussed (e.g., Korn, 1984). In the presence of a blocking variable, the weighted sum of Kendall's tau or Spearman's rho can be computed. Simulation research suggests they are equally powered when using the optimally weighted sum of taus or rhos, but Spearman's rho is preferred when ties are present as its variance has a simpler form (Taylor, 1987).

Generalizations to Kendall's tau which allow the examination of dichotomous, ordinal, and quantitative measures have been developed (Zhang et al., 2010). Kendall's tau has also been applied to single-subject research designs as tau- U , correlating time and the observed outcome (e.g., Brossart et al., 2018), within Bayesian methods (e.g., van Doorn et al., 2019), and within multilevel models (e.g., Romdhani et al., 2014).

Table 4 presents the continuous and rank ordered data previously presented in Table 1, with the exception that academic achievement (continuous variable) has been rank ordered from the smallest to largest values. Had the data in columns 1 and 2 of Table 4 (the continuous values) been treated as rank ordered (columns 3 and 4) instead of continuous, a weaker correlation is evidenced with Kendall's tau ($\tau_b = 0.179$, where P represents concordant pairs and sum to 112, and Q represents discordant pairs which sum to 78) relative to what was computed using continuous data with Pearson ($r = 0.30$).

If the data were ordered categories (columns 5 and 6 of Table 1), which is common to find in practice, Kendall's tau weakens further ($\tau_b = 0.141$).

Additional alternative correlation coefficients

The correlations presented thus far are arguably the most common correlation coefficients. However, there are several additional correlations that can be computed which are more robust to violations of assumptions associated with the Pearson correlation. A concise overview for some of these lesser known correlations is presented. Readers interested in formula and technical notes are referred to the resources presented.

Percentage bend correlation

The percentage bend correlation is a correlation coefficient that can be computed when linearity is not a reasonable assumption (e.g., skewed distributions) (Wilcox, 1994). As one robust analog to the Pearson coefficient, the percentage bend correlation protects against outliers in the marginal distributions without taking the overall data structure into account (Wilcox, 2017). The percentage bend correlation is intended to serve as a correlation estimate that is not overly sensitive to minor distributional changes, and it is not intended to replace or serve as an estimate of ρ (Wilcox, 2017). An example of the use of the percentage bend correlation is in neuropsychological studies that have small samples due to rare disorders and often extreme observations (Kulesz et al., 2015).

Table 4 Concordant and discordant pairs for academic achievement and emotional intelligence scores.

Academic achievement	Emotional intelligence	Academic achievement (rank)	Emotional intelligence (rank)	Discordant pairs (Q)	Concordant pairs (P)
2.676	50.928	1.00	12.00	11.00	8.00
2.839	55.193	2.00	17.00	15.00	3.00
2.840	40.631	3.00	7.00	6.00	11.00
2.870	25.346	4.00	1.00	0.00	16.00
2.922	51.992	5.00	15.00	11.00	4.00
2.938	51.173	6.00	13.00	9.00	5.00
2.957	42.259	7.00	8.00	5.00	8.00
3.008	26.567	8.00	2.00	0.00	12.00
3.022	43.616	9.00	9.00	4.00	7.00
3.041	48.014	10.00	10.00	4.00	6.00
3.116	38.732	11.00	6.00	3.00	6.00
3.119	30.030	12.00	3.00	0.00	8.00
3.211	48.778	13.00	11.00	2.00	5.00
3.352	34.816	14.00	4.00	0.00	6.00
3.364	60.075	15.00	18.00	3.00	2.00
3.371	63.016	16.00	20.00	4.00	0.00
3.424	51.538	17.00	14.00	1.00	2.00
3.425	35.933	18.00	5.00	0.00	2.00
3.504	54.895	19.00	16.00	0.00	1.00
3.542	61.304	20.00	19.00	0.00	0.00

Point-biserial and biserial correlation

When the Pearson correlation coefficient formula is used with a dichotomous (0, 1) variable and a continuous variable, the correlation is known as a point-biserial correlation. The formula for the point-biserial correlation, as presented by [Sheskin \(2011\)](#), is as follows:

$$r_{pb} = \left(\frac{\bar{Y}_0 - \bar{Y}_1}{s_Y} \right) \sqrt{\frac{Np_0(1-p_0)}{N-1}}$$

Where $\bar{Y}_1$ represents the mean of the continuous variable for the cases coded as 1 on the binary variable, $\bar{Y}_0$ is the mean of the continuous variable for cases coded as 0 on the binary variable, and s_Y represents the sample standard deviation for the continuous variable. The overall sample size is N , with p_0 representing the proportion of cases coded 0 in the binary variable.

When the binary variable is generated by dichotomizing a continuous variable which originates from a distribution that is normally distributed, the Pearson correlation coefficient computed between these variables is termed the biserial correlation ([Cheng and Liu, 2016](#)). The biserial correlation is computed as follows:

$$r_b = \left(\frac{\bar{Y}_0 - \bar{Y}_1}{\sigma_Y} \right) \left(\frac{n_1 n_0}{\gamma N^2} \right)$$

The notation follows that from the point-biserial correlation. The new term, γ , represents a probability density, specifically the ordinate or height of the standard normal distribution curve at the point on the curve which separates p_0 (proportion of cases coded as 0) from $1-p_0$ (proportion of cases coded as 1). The sample sizes of the cases coded as 1 and 0 are reflected in n_1 and n_0 , respectively.

The biserial correlation coefficient can be computed from the point biserial as follows:

$$r_b = \frac{r_{pb} \sqrt{p_0(1-p_0)}}{\gamma}$$

In both cases, the biserial and point-biserial correlation coefficients are appropriate to compute when correlating a dichotomous (0, 1) variable with a continuous variable ([Tate, 1954](#)). Using the binary data for academic achievement and the continuous data for emotional intelligence from [Table 1](#) and applying the Pearson correlation coefficient, we compute a sample point-biserial correlation coefficient as follows:

$$r_{pb} = \left(\frac{\bar{Y}_0 - \bar{Y}_1}{s_Y} \right) \sqrt{\frac{Np_0(1-p_0)}{N-1}} = \left(\frac{45.947 - 45.360}{11.240} \right) \sqrt{\frac{20(0.35)(1-0.35)}{20-1}} = 0.026$$

Had the point-biserial correlation been computed using the binary data for emotional intelligence and continuous data for academic achievement (see Table 1), the point-biserial correlation coefficient would be quite a bit stronger, specifically 0.176. Thus, depending on the variable which is dichotomized, the point-biserial may produce similar results to other correlation coefficients.

Computing the biserial using the point-biserial for the relationship between the binary data for academic achievement and the continuous data for emotional intelligence, we find the following. In this case, using the standard normal curve table, the value of z that demarcates the upper 35% of the distribution (p_0) from the remaining 65% ($1-p_0$) is approximately 0.39. When z is 0.39, y is 0.3697 (where y is obtained from a table for ordinates of the standard normal curve such as Bernstein et al. (1988)). Accordingly, the biserial correlation can be computed from the point-biserial as follows:

$$r_b = \frac{r_{pb} \sqrt{p_0(1-p_0)}}{y} = \frac{0.026 \sqrt{0.35(1-0.35)}}{0.3697} = 0.034$$

Early scholars noted that the range of the point-biserial correlation never reaches the same value as the biserial correlation (Lord and Novick, 1968), specifically reaching a maximum of 0.798 when the continuous variable is normally distributed (Cheng and Liu, 2016). Thus interpreting the size of the point-biserial correlation based on the usual range of the Pearson correlation is discouraged (Allen and Yen, 1979). More recent research suggests that the point-biserial coefficient becomes more restricted when the probability that the binary variable is 1 (vs. 0) departs from 0.50. However, the maximum value of the point-biserial correlation may actually range from -1.0 to $+1.0$ (Cheng and Liu, 2016). Thus, researchers should not assume the point-biserial has a restricted range (Cheng and Liu, 2016). The point-biserial correlation is often used to calculate item-total score coefficients (i.e., how strongly each item correlates with the total score) as a way of estimating item discrimination.

Tetrachoric correlation

A tetrachoric correlation, denoted as r_t , is a special case of the Pearson coefficient. The tetrachoric coefficient assumes bivariate normality for two continuous latent variables measured as dichotomous (i.e., a 2×2 contingency table). When there is bivariate normality of the underlying distribution prior to dichotomization, the phi coefficient can be computed from the tetrachoric correlation (and vice versa) (Demirtas, 2016). The tetrachoric coefficient is most often used to correlate two independent binary variables but can also be used to examine rater reliability of a single rater when two raters independently rate using a dichotomous scale (Bonett and Price, 2005).

Pearson presented this coefficient in an 1899 memoir, providing illustrations that included a horse's coat color and a person's eye color (i.e., correlating maternal grandmother eye color with granddaughter eye color with eyes of two shades, gray or lighter or dark gray or darker) (Pearson, 1900). Later researchers offered alternate approximations of the tetrachoric (e.g., Castellan, 1966; Digby, 1983).

The tetrachoric correlation has been described as "quite possibly, the single most difficult correlation coefficient to compute" (Long et al., 2009). Brown (1977) presented an algorithm for the tetrachoric correlation, and the coefficient is becoming more mainstreamed in many disciplines given today's high speed computing (Greer et al., 2003). Its use has also "been surrounded by a certain amount of mystique and a separation of researchers into two camps, believers and non-believers in underlying continuous normal variables" (Muthén and Hofacker, 1988), in other words, the extent to which a "true" dichotomy actually exists. This has led to the creation of an empirical test to determine the appropriateness of using tetrachoric correlations for binary variables (Muthén and Hofacker, 1988).

The tetrachoric correlation has also been found to be a biased estimated of the population correlation, with low expected cell frequency contributing to this bias. (Brown and Benedetti, 1977). Brown and Benedetti found that expected frequencies of 10 or more are needed to obtain reasonable standard error estimates.

Polychoric correlation

The polychoric correlation generalizes the tetrachoric correlation coefficient to more than two categories and, like the tetrachoric, has an assumption that the underlying bivariate distribution is normal. The use of polychoric correlations is common when fitting models in structural equation modeling to ordinal data (Monroe, 2018). Maximum likelihood methods have traditionally been used to estimate polychoric correlations, and when estimated in this fashion, have been found to be an unbiased estimate of the correlation between the bivariate normal variables (Olsson, 1979). More recent research using Bayesian suggests that maximum a posteriori (MAP) estimation is a better estimator of the polychoric (relative to maximum likelihood) and overcomes the shrinkage effect evident in expected a posteriori (EAP).

Research on the polychoric suggests that, for ordinal data, it outperforms the Pearson, Spearman's rhos, and Kendall's tau-b coefficients (Babakus and Ferguson, 1988) and is robust even where there is modest violation of normality (Quiroga, 1992). Additionally, MAP estimators with uniform prior (MAPU) has been found to be less sensitive to small sample size than maximum likelihood and other estimators, and for all sample sizes examined ($n = 25$ to 400) consistently outperformed maximum likelihood (Choi et al., 2011).

Additional correlations

Although an attempt has been made to address the more commonly known correlations, it is important to recognize that there are many additional correlation coefficients that have been created. Among others, these include the biweight midcorrelation and the distance correlation which are mentioned very briefly here.

The biweight midcorrelation is another robust correlation alternative to the Pearson, and is less sensitive to outliers as it is computed using the median (Langfelder and Horvath, 2012; Wilcox, 2017). In genetics, the biweight correlation has been found to outperform mutual information (MI), a generalized correlation index (Song et al., 2012).

The distance coefficient, based on Euclidean distances between sample elements (not sample moments), is analogous to the Pearson coefficient. Distance correlations are measures of dependence between random vectors but equate to zero only when there is independence in the random vectors. The distance correlation can be applied to examine dependence and is more general than the Pearson coefficient in that normality is not assumed (Székely et al., 2007).

Considerations with correlations

There are a number of statistical, as well as more theoretical, considerations that should be made when working with correlations. These include the consideration of causation in the examination of relationships, spurious correlations, restriction of range, and data considerations.

Correlation and causation

Variables that are related indicate that there is an identifiable pattern to the data. This is not meant to imply a causal relationship. Correlation coefficients allow the examination of relationships or associations but do not allow the inference of causality—albeit quite tempting to make that inference. Any attempt to draw causal inference from data which are correlated, by merely examining the correlation, is impossible and “fraught with pitfalls” (Haig, 2011). Causal inference between two variables, such as X and Y , can occur only under the following conditions: (a) X occurs prior to Y ; (b) X and Y covary; and (c) there are no confounders that occur in the relationship between X and Y . Readers interested in learning more about causation and causal inference are referred to any number of excellent choices (e.g., Morgan and Winship, 2014; Pearl et al., 2016).

Spurious correlations

A spurious correlation occurs when the correlation between two variables occurs not because one is caused by the other but rather because of a third variable. Examples of spurious correlations that have been found using actual data, with common sense indicating they are spurious, include the relationship between sour cream consumption and fatal motor cycle accidents ($r = 0.916$), the relationship between consumption of margarine and the divorce rate in Maine ($r = 0.993$), and science and technology expenditures and suicide by hanging, strangulation, and suffocation ($r = 0.86$) (Moosa, 2017). Spurious correlations are the underlying reason that cointegration analysis techniques, developed within the field of economics, were developed (Moosa, 2017). Cointegration proposes that if two integrated variables are strongly correlated, the correlation is spurious unless there is cointegration of the variables (Moosa, 2017). Moosa (2017) demonstrated that spurious correlations cannot be rejected with cointegration testing when there is a strong correlation coefficient between the underlying series, suggesting that “spurious relations cannot be revealed by testing alone and that theory, common sense, and intuition should come first”. This thought is echoed by many researchers who see the phenomena of spurious correlations particularly problematic with big data (Calude and Longo, 2017).

Restriction of range

Restriction of range occurs in correlated variables when data observed from a sample are not available across the entire range of scores for one of the correlated variables. In other words, one of the variables has very little variation. This is termed “restriction of range” as to produce a relationship between variable, there must be a certain range of values in the variables being correlated so that the correlation manifests. Recall that the computation for a correlation is based on how values deviate from the mean. If there is very little or no variation in a variable, the correlation produced using that variable will suggest little to no correlation, even if in reality the correlation is very strong. An example of restriction of range may be correlating college entrance exam scores and first semester college cumulative grade point average for gifted students. It is reasonable to suspect that the variation of college entrance exam scores for gifted students will be relatively limited. There is not enough variation in this variable to manifest into a strong relationship and the resulting correlation will be attenuated (although some research suggests range restriction may increase the correlation in the presence of outlier prone distributions such as exponential and lognormal, among others (Zimmerman and Williams, 2000)). This is an example of indirect range restriction. There is not direct truncation but rather it is produced when selection of the sample is based on another correlated variable (in this example, identification as gifted) (Sackett and Wade, 1983; Meade, 2012). Direct range restriction can also occur. In direct range restriction, the score distribution on the variable is directly truncated (Meade, 2012). For example, direct range restriction would occur in a situation where individuals complete instrument X and Y , and only those in the top scoring 10% of X are selected for the sample.

Researchers who encounter restriction of range must address it statistically. Additionally, it is important to first understand if the range restriction is direct or indirect as the steps and sequence are dependent on the type or range restriction. Hunter and colleagues have shown that using direct range restriction practices when the range restriction is indirect produces substantial negative biases in

corrected values. Steps and formulas for correcting for range restriction are available (e.g., Hunter et al., 2006), and though illustrated in the context of meta-analyses, the methods are applicable to correlations in individual studies. Thorndike's case II correction formula is commonly used in the event of direct range restriction when restriction occurs on only one variable (Thorndike, 1949) and was built upon earlier work of Pearson (Pearson, 1903). Corrections for multivariate range restriction are also available (e.g., Lawley, 1943; Johnson and Ree, 1994). More recently, researchers have proposed viewing selection as a missing data mechanism and have studied the application of missing data techniques (e.g., multiple imputation or full information maximum likelihood) (e.g., Pfaffel and Spiel, 2016; Pfaffel et al., 2016a,b; Wiberg and Sundstrom, 2009).

Data considerations

A number of data considerations have been presented previously. More general considerations, including sample size and distributional shape, will be presented here.

Sample size

As with many statistics, small sample sizes can detrimentally impact the correlation coefficient. Specifically, due to capitalizing on chance, the expected value of the squared Pearson correlation will be larger than zero when samples are small (Wishart, 1931). In the presence of other conditions sometimes experienced with correlations, sample size can be a further issue with which researchers must handle (e.g., Pfaffel and Spiel, 2016). Both the Pearson and Spearman have been found to underestimate the population correlation value for small sample sizes (de Winter et al., 2016). The underestimation in the presence of small samples (e.g., $N = 5$ and $N = 20$) is relatively minimal if the population Pearson coefficient is small or moderate (e.g., $\rho = 0.20$) but becomes more severe between 0.30 and 0.90.

The sample size needed to obtain a Fisher confidence interval for desired width with Pearson, Kendall's tau, or Spearman can be computed (Bonett and Wright, 2000), including the use of bootstrapped replicates to estimate the Pearson correlation confidence interval (Olivoto et al., 2018).

Distributional shape

When both variables being correlated are symmetrical and have the same distributional shape, the Pearson correlation between those variables can range from -1.0 to $+1.0$. However, to the extent that asymmetry and dissimilar shapes occur, the coefficient will not reach $|1|$. Researchers have also found that the Spearman and Pearson behave similarly when the correlated variables are normally distributed. However, Spearman is robust to outliers while the Pearson is quite sensitive, even in samples as large as 200 (de Winter et al., 2016). Researchers working with distributions that have kurtosis present, particularly in the presence of leptokurtic distributions, are encouraged to compute the Spearman rather than Pearson as the Pearson coefficient has been found to be more variable in the presence of leptokurtic distributions (de Winter et al., 2016).

Acknowledgment

No support or financial assistance was provided for this project.

References

- Allen, M.J., Yen, W.M., 1979. *Introduction to Measurement Theory*. Brooks/Cole, Monterey, CA.
- Babakus, E., Ferguson, C.E., 1988. On choosing the appropriate measure of association when analyzing rating scale data. *J. Acad. Market. Sci.* 16, 95–102.
- Bernstein, I.H., Garbin, C.P., Teng, G.K., 1988. Appendix A: Tables of the Normal Curve. *Applied Multivariate Analysis*. Springer-Verlag, New York, NY.
- Blyth, S., 1994. Karl Pearson and the correlation curve. *Int. Stat. Rev.* 62, 393–403.
- Bonacich, P., 1972. Technique for analyzing overlapping memberships. *Socio. Methodol.* 4, 176–185.
- Bonett, D.G., Price, R.M., 2005. Inferential methods for the tetrachoric correlation coefficient. *J. Educ. Behav. Stat.* 30, 213–225.
- Bonett, D.G., Wright, T.A., 2000. Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika* 65, 23–28.
- Brossart, D.F., Laird, V.C., Armstrong, T.W., 2018. Interpreting Kendall's tau and tau-U for single-case experimental designs. *Cogn. Psychol.* 5.
- Brown, M.B., Benedetti, J.K., 1977. On the mean and variance of the tetrachoric correlation coefficient. *Psychometrika* 42, 347–355.
- Brown, M.B., 1977. The tetrachoric correlation and its asymptotic standard error. *J. Roy. Stat. Soc.* 26, 343–361.
- Calude, C.S., Longo, G., 2017. The deluge of spurious correlations in big data. *Found. Sci.* 595–612.
- Castellan Jr., N.J., 1966. On the estimation of the tetrachoric correlation coefficient. *Psychometrika* 31, 67–73.
- Chen, P.Y., Popovich, P.M., 2011. *Correlation*. Sage, Thousand Oaks, Ca.
- Cheng, Y., Liu, H., 2016. A short note on the maximal point-biserial correlation under non-normality. *Br. J. Math. Stat. Psychol.* 69, 344–351.
- Choi, J., Kim, S., Chen, J., Dannels, S., 2011. A comparison of maximum likelihood and Bayesian estimation for polychoric correlation using Monte Carlo Simulation. *J. Educ. Behav. Stat.* 36, 523–549.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum, Hillsdale, NJ.
- Cramér, H., 1946. *Mathematical Methods of Statistics*. Princeton University Press, Princeton.
- de Winter, J.C.F., Gosling, S.D., Potter, J., 2016. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: a tutorial using simulations and empirical data. *Psychol. Methods* 21, 273–290.
- Deming, W.E., Frederick, F.S., 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* 11, 427–444.
- Demirtas, H., 2016. A note on the relationship between the phi coefficient and the tetrachoric correlation under nonnormal underlying distributions. *Am. Statistician* 70, 143–148.

- Digby, P.G.N., 1983. Approximating the tetrachoric correlation coefficient. *Biometrics* 39, 753.
- Eves, H.W., 2002. A very brief history of statistics. *Coll. Math. J.* 33, 306–308.
- Fienberg, S.E., 1992. A brief history of statistics in three and one-half chapters: a review essay. *Stat. Sci.* 7, 208–225.
- Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* 10, 507–521.
- Fisher, R.A., 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52, 399–433.
- Fisher, R.A., 1922. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond.* 222, 309–368.
- Galton, F., 1889. *Natural Inheritance*.
- Geisser, S., 2006. *Modes of Parametric Statistical Inference*. Wiley-Interscience, Hoboken, NJ.
- Goodman, L.A., Kruskal, W.H., 1954. Measures of association for cross classifications. *J. Am. Stat. Assoc.* 49, 732–764.
- Goodman, L.A., Kruskal, W.H., 1979. *Measures of Association for Cross Classification*. Springer Verlag, New York, NY.
- Greer, T., Dunlap, W.P., Beatty, G.O., 2003. A Monte Carlo evaluation of the tetrachoric correlation coefficient. *Educ. Psychol. Meas.* 63, 931–950.
- Hahs-Vaughn, D.L., Lomax, R.G., 2020a. *An Introduction to Statistical Concepts*. Routledge/Taylor & Francis, New York, NY.
- Hahs-Vaughn, D.L., Lomax, R.G., 2020b. *Statistical Concepts: A First Course*. Routledge/Taylor & Francis, New York, NY.
- Haig, B., 2011. Spurious correlation. In: Salkind, N.J. (Ed.), *Encyclopedia of Measurement and Statistics*. Sage.
- Hald, A., 2007. *A History of Parametric Statistical Inference From Bernoulli to Fisher, 1713–1935*. Springer, New York, NY.
- Hermes, G., 1970. A Markovian approach to measures of association. *Am. J. Sociol.* 75, 992–1011.
- Hotelling, H., Pabst, M.R., 1936. Rank correlation and tests of significance involving no assumption of normality. *Ann. Math. Stat.* 7, 29.
- Howell, D.C., 2010. *Statistical Methods for Psychology*. Thomson Wadsworth, Belmont, CA.
- Hunter, J.E., Schmidt, F.L., Le, H., 2006. Implications of direct and indirect range restriction for meta-analysis methods and findings. *J. Appl. Psychol.* 594.
- Johnson, J.T., Ree, M.J., 1994. Range: A Pascal Program to Compute the Multivariate Correction for Range Restriction.
- Kendall, M.G., Gibbons, J.D., 1990. *Rank Correlation Methods*. Oxford University Press, New York, NY.
- Kendall, M.G., Stuart, A., 1958. *The Advanced Theory of Statistics*. Hafner Publishing Co, New York, NY.
- Kendall, M.G., 1938. A new measure of rank correlation. *Biometrika* 30, 81–93.
- Kendall, M.G., 1945. The treatment of ties in ranking problems. *Biometrika* 33, 239–251.
- Kendall, M.G., 1970. *Rank Correlation Methods*. Griffin, London, England.
- Korn, E.L., 1984. Kendall's tau with a blocking variable. *Biometrics* 40, 209–214.
- Kulesz, P.A., Tian, S., Juranek, J., Fletcher, J.M., Francis, D.J., 2015. Relations between volumetric measures of brain structure and attentional function in spina bifida: utilization of robust statistical approaches. *Neuropsychology* 29, 212–225.
- Kvalseth, T.O., 2018. An alternative to Cramer's coefficient of association. *Commun. Stat. Theor. Methods* 47, 5662–5674.
- Langfelder, P., Horvath, S., 2012. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Software* 46.
- Lawley, D.N., 1943. A note on Karl Pearson's selection formulae. *Proc. R. Soc. Edinb. Sect. A* 62, 28–30.
- Liu, R., 1980. Note on phi-coefficient comparison. *Res. High. Educ.* 13, 3–8.
- Long, M.A., Berry, K.J., Mielke Jr., P.W., 2009. Tetrachoric correlation: a permutation alternative. *Educ. Psychol. Meas.* 69, 429–437.
- Lord, F.M., Novick, M.R., 1968. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Reading, MA.
- Lovie, A.D., 1995. Who discovered Spearman's rank correlation? *Br. J. Math. Stat. Psychol.* 48, 255–269.
- Maccann, C., Jiang, Y., Brown, L.E.R., Double, K.S., Bucich, M., Minbashian, A., 2020. Emotional intelligence predicts academic performance: a meta-analysis. *Psychol. Bull.* 146, 150–186.
- Meade, A.W., 2012. Restriction of range. In: Salkind, N.J. (Ed.), *Encyclopedia of Research Design*. Sage, Los Angeles, CA.
- Monroe, S., 2018. Contributions to estimation of polychoric correlations. *Multivariate Behav. Res.* 53, 247–266.
- Moosa, I.A., 2017. Blaming suicide on NASA and divorce on margarine: the hazard of using cointegration to derive inference on spurious correlation. *Appl. Econ.* 49, 1483–1490.
- Morgan, S.L., Winship, C., 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, Cambridge, England.
- Muthén, B., Hofacker, C., 1988. Testing the assumptions underlying tetrachoric correlations. *Psychometrika* 53, 563–577.
- Norton, B., Pearson, E.S., 1976. A note on the background to, and refereeing of, R. A. Fisher's 1918 paper "On the correlation between relatives on the supposition of Mendelian inheritance". *Notes Rec. R. Soc. Lond.* 31, 151–162.
- Olivoto, T., Lúcio, A.D.C., DE Souza, V.Q., Nardino, M., Diel, M.I., Sari, B.G., Kryszczun, D.K., Meira, D., Meier, C., 2018. Confidence interval width for Pearson's correlation coefficient: a Gaussian-independent estimator based on sample size and strength of association. *Agron. J.* 110, 503–510.
- Olsson, U., 1979. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 44, 443–460.
- Pearl, J., Glymour, M., Jewell, N.P., 2016. *Causal Inference in Statistics: A Primer*. Wiley, West Sussex, England.
- Pearson, K., 1900. Mathematical contributions to the theory of evolution. VII. On the correlation of characters not quantitatively measurable. *Philos. Trans. Roy. Soc. Lond.* 195A, 1–47.
- Pearson, K., 1903. Mathematical contributions to the theory of evolution XI on the influence of natural selection on the variability and correlation of organs. *Philos. Trans. Roy. Soc. Lond.* 200, 1–66.
- Pearson, K., 1904. *On the Theory of Contingency and Its Relation to Association and Normal Correlation*. Drapers Company Research Memoires, London, England.
- Pearson, K., 1920. Notes on the history of correlation. *Biometrika* 13, 25–45.
- Pfaffel, A., Spiel, C., 2016. Accuracy of range restriction correction with multiple imputation in small and moderate samples: a simulation study. *Pract. Assess. Res. Eval.* 21, 1–20.
- Pfaffel, A., Kollmayer, M., Schober, B., Spiel, C., 2016a. A missing data approach to correct for direct and indirect range restrictions with a dichotomous criterion: a simulation study. *PLoS One* 11 (3), e0152330.
- Pfaffel, A., Schober, B., Spiel, C., 2016b. A comparison of three approaches to correct for direct and indirect range restrictions: a simulation study. *Pract. Assess. Res. Eval.* 21, 1–15.
- Quiroga, A.M., 1992. Studies of the Polychoric Correlation and Other Correlation Measures for Ordinal Variables. *Acta Universitatis Upsaliensis*, Sweden.
- Romdhani, H., Lakhali-Chaieb, L., Rivest, L.P., 2014. Kendall's tau for hierarchical data. *J. Multivariate Anal.* 128, 210–225.
- Sackett, P.R., Wade, B.E., 1983. On the feasibility of criterion-related validity: the effects of range restriction assumptions on needed sample size. *J. Appl. Psychol.* 68, 374–381.
- Sheskin, D., 2011. *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press, Boca Raton.
- Song, L., Langfelder, P., Horvath, S., 2012. Comparison of co-expression measures: mutual information, correlation, and model based indices. *BMC Bioinf.* 13, 1–21.
- Spearman, C., 1904. The proof and measurement of association between two things. *Am. J. Psychol.* 15, 72–101.
- Stephen, M.S., 1978. Francis Ysidro Edgeworth, statistician. *J. Roy. Stat. Soc.* 141, 287–313.
- Stigler, S.M., 1986. *The History of Statistics: The Measurement of Uncertainty Before 1900*. Belknap Press of Harvard University Press, Cambridge, MA.
- Stuart, A., 1953. The estimation and comparison of strengths of association in contingency tables. *Biometrika* 40, 105–110.
- Székely, G.J., Rizzo, M.L., Bakirov, N.K., 2007. Measuring and testing dependence by correlation of distances. *Ann. Stat.* 35, 2769–2794.
- Tate, R.F., 1954. Correlation between a discrete and a continuous variable: point-biserial correlation. *Ann. Math. Stat.* 25, 603–607.
- Taylor, J.M.G., 1987. Kendall's and Spearman's correlation coefficients in the presence of a blocking variable. *Biometrics* 43, 409–416.
- Thorndike, R.L., 1949. *Personnel Selection: Test and Measurement Techniques*. Wiley.
- van der Waerden, B.L., 1969. *Mathematical Statistics*. Springer-Verlag, Berlin, Germany.

- van Doorn, J., Ly, A., Marsman, M., Wagenmakers, E.J., 2019. Bayesian Estimation of Kendall's Tau Using a Latent Normal Approach.
- Wiberg, M., Sundstrom, A., 2009. A comparison of two approaches to correction of restriction of range in correlation analysis. *Pract. Assess. Res. Eval.* 14.
- Wilcox, R.R., Muska, J., 2001. Inferences about correlations when there is heteroscedasticity. *Br. J. Math. Stat. Psychol.* 54, 39–47.
- Wilcox, R.R., 1994. The percentage bend correlation coefficient. *Psychometrika* 59, 601–616.
- Wilcox, R.R., 1995. *Statistics for the Social Sciences*. Academic Press, San Diego, CA.
- Wilcox, R.R., 2004. Inferences based on a skipped correlation coefficient. *J. Appl. Stat.* 31, 131–143.
- Wilcox, R.R., 2016. *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press.
- Wilcox, R.R., 2017. *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier, Burlington, MA.
- Wishart, J., 1931. The mean and second moment coefficient of the multiple correlations coefficient in samples from a normal population. *Biometrika* 22, 353–361.
- Wolfowitz, J., 1942. Additive partition functions and a class of statistical hypotheses. *Ann. Math. Stat.* 13, 247–279.
- Yasmin, H.S., 2007. On the eras in the history of statistics and data analysis. *J. Wash. Acad. Sci.* 93, 17–34.
- Yule, G.U., 1912. On the methods of measuring association between two attributes. *J. Roy. Stat. Soc.* 75, 579–652.
- Zhang, H., Liu, C.-T., Wang, X., 2010. An association test for multiple traits based on the generalized Kendall's tau. *J. Am. Stat. Assoc.* 105, 473–481.
- Zimmerman, D.W., Williams, R.H., 2000. Restriction of range and correlation in outlier-prone distributions. *Appl. Psychol. Meas.* 24, 267.

Statistical inference: frequentist approaches

Jeffrey R. Harring, University of Maryland, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Foundational concepts	751
Population vs. sample	751
A brief introduction to probability	752
Probability distributions	752
Hypothesis testing	754
Translating hypotheses	754
Sampling distribution	755
Central limit theorem	755
Standard error	756
Using the sampling distribution	757
Test statistics and confidence intervals	758
Hypothetical example	759
Confidence intervals	759
Discussion	760
References	761

Statistical inference is primarily concerned with making claims about unknown features of a population based on data that are randomly sampled from that population. From a frequentist perspective, these inferences or conclusions are made using probability models that are thought to have generated the data. In contrast to activities of data exploration which focus on distilling information and searching for interesting patterns using the sample data, statistical inference is focused on using the sample data to answer specific questions about the population—posed before the data are collected. Conclusions based on exploration are typically informal—inferred based on characteristics (i.e., individuals, settings, and circumstances) of the data whereas statistical inferential claims are more formal, backed by a statement of our confidence in them, and applied to a wider universe of individuals and a broader set of circumstances and settings (Garfield et al., 2008). According to Moore (2004), there are two central themes in statistical inference. The first is parameter estimation—obtaining a point estimate of a quantity of interest from the data. An example might be estimating parameters of a multiple regression model or even something as simple as estimating the mean of a continuous variable. Estimation approaches are discussed in detail in “Estimation Approaches (Least-squares, ML variants, Bayesian)”, another entry in this volume. The second theme and more germane to this entry is hypothesis testing, which consists of making probabilistic statements about the values of estimated quantities in the population. The Pearson product moment correlation (r) calculated from the sample might be used, for instance, to test the bivariate linear relation between two continuous variables in the population. Testing hypotheses often include computing confidence intervals and reporting test results using p -values. These approaches are intended to quantify uncertainty inherent when using sample data to make claims about the population as well as to evaluate the plausibility of chance as an alternative explanation to patterns in the observed data. This article discusses the components of statistical inference from a frequentist perspective including a brief introduction to probability, probability distributions, sampling distributions, sampling error, point and interval estimates and hypothesis testing.

Foundational concepts

A number of foundational concepts related to statistical inference was previously mentioned but not discussed in any detail. A brief description of each is now provided.

Population vs. sample

In general, a *population* is an entire collection of elements whose characteristics one may want to study (e.g., fifth-grade students' scores on a standardized test, length of stay in hospitals in Massachusetts, infant mortality rates in large urban cities). As suggested by these examples, a population can be of any size depending on how it is defined. It could range from a relatively small set of elements, which might be easily collected, to a large but finite number that would be impractical to collect in its entirety. Seldom are data collected from an entire population. Instead, a *random sample* from the population is drawn to infer something about the characteristics of the population. Often referred to as *parameters* denoted by Greek letters, these characteristics are numerical quantities [e.g., mean (μ), variance (σ^2), correlation (ρ)] describing the distributions of and relations among the variables of interest. These corresponding characteristics

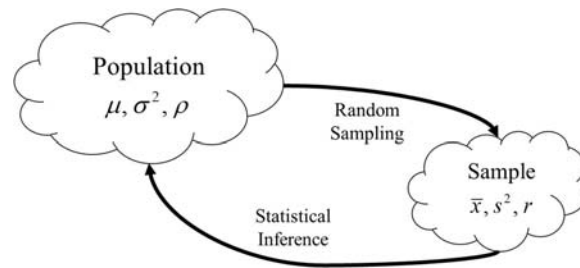


Fig. 1 A schematic of the foundational concepts of statistical inference.

have counterparts in the sample data. These are referred to as *statistics* [e.g., mean ($\bar{x}$), variance (s^2), correlation (r)] and are estimates of the corresponding population parameters. Presuming that the sample that is gathered is truly random, and representative of the population, not only can population parameters be estimated by statistics computed from the sample data, but the accuracy of these estimates can also be assessed. These foundational concepts and their relations are summarized in **Fig. 1**.

A brief introduction to probability

The foundation of the frequentist approach to statistical inference is largely probabilistic in nature and so a brief discussion of basic probability tenets is needed to make sense of statistical results presented later. The concept of *probability* can be viewed in different ways, but most common definitions take either an analytic or frequentist perspective. The analytic perspective defines probability formally. Take for example flipping a coin and computing the probability of obtaining the outcome of a tail on a single flip. If an event (flipping of a coin and getting a tail) can occur in A ways and can fail to occur in B ways, and if all possible ways are equally likely (e.g., a tail or a head has an equal chance of happening), then the probability of its occurrence is $A/(A + B)$, and the probability of its failing to occur is $B/(A + B)$. The coin has two possible outcomes, heads or tails, and flipping a tail can only be obtained one way. Thus, the probability of obtaining a tail on a single flip of a coin using the analytic method is $Pr(tails) = 1/(1 + 1) = 0.5$.

In the previous example, it was trivial to list all of the possible outcomes of the *event*. Many real-life scenarios involving computing probabilities are not so straightforward because listing all of the possible outcomes is impractical. For instance, what is the probability of winning a game of solitaire? Here, a frequentist approach to computing probability may be more prudent. Determining the probability of flipping a tail through a frequentist approach necessitates conducting a series of trials (i.e., flips of the coin) and calculating the proportion of flips that come up tails. To keep it simple, suppose a coin was flipped 20 times and the following pattern of heads and tails was observed:

T H T H H T H T H T T H H T T T T H T T

To compute the probability of flipping a tail one simply takes the number of tails (n_T) and divides that number by the total number of independent trials ($n = 20$). The estimated probability of flipping a tail comes out to $(12/20)$ or 0.60. Of course, this value is much different from the probability computed from the analytic perspective. Our frequentist probability would be more accurate if the number of flips increased. Thus, the probability might be defined as the limit of the relative frequency of occurrence of the desired event that is approached as the number of trials increases.

$$Pr(\text{tail}) = \lim_{n \rightarrow \infty} \frac{n_T}{n}$$

That is, if the coin were flipped 100, 1000, 10,000 times or more, the probability of a tail occurring based on relative frequency would approach 0.50. Statistical inferences, as we will see, are largely based on decisions made with respect to probabilities computed from frequentist or analytical approaches.

Probability distributions

A probability distribution underlying a random variable in the population governs the probability of obtaining a specific value or the probability of a value that falls within a particular interval. This distinction is necessary because the first applies to distributions of *discrete* variables while the latter applies to distributions of *continuous* variables.

A discrete variable typically takes on one of a relatively small number of possible values (e.g., a five-point scale of socioeconomic status). A variable that can take on one of infinitely many possible values is generally treated as a continuous variable if the values represent at least an interval or ratio scale.¹ For example, consider a binomial distribution—a discrete probability distribution—to

¹The level of measurement or scale of measure refers to a classification that describes the nature of information within the numbers assigned to variables. According to [Stevens \(1946\)](#) there are four different scales of measurement: nominal, ordinal, interval, and ratio.

represent the number of heads obtained when flipping three coins. The binomial distribution deals with situations in which each of a number of independent trials results in one of two mutually exclusive outcomes. The binomial distribution is defined formally as

$$Pr(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where $Pr(X=x)$ is the probability of random variable X taking on a specific value x , p is the probability of a success, and n represents the number of trials. **Table 1** and **Fig. 2** show the possible outcomes of flipping three coins and observing the number of heads. For example, the probability of obtaining exactly $x=2$ heads when flipping $n=3$ coins is 0.375. For discrete distributions like the binomial distribution modeling the number of heads observed when flipping three coins, the probabilities for each discrete event can be computed and are the sum of the areas of the rectangles in **Fig. 2**. These probabilities, and thus the area of the rectangles, sum to 1. This is a characteristic of all probability distributions.

As another example, consider a continuous random variable such as weight. Weights of newborn babies in the US vary according to the normal distribution (the normal distribution is a continuous probability distribution) with approximate mean of $\mu = 7.50$ pounds and standard deviation of $\sigma = 1.25$. **Fig. 3** shows this hypothetical distribution of weights of US newborns in the population.

Here, it does not make sense to seek the probability of a particular weight, say 7.321476 lbs., which would be very small—essentially zero. However, it does make sense to determine the probability over an interval of weights, say from 6.5 lbs. to 8.5 lbs. That is, what is the probability that a randomly selected newborn from the population will weigh between 6.5 and 8.5 pounds, inclusively? This probability is denoted by $Pr(6.5 \leq x \leq 8.5)$. **Fig. 4** shows this probability as the shaded area under the normal curve and the computed probability is 0.567. While it may not be obvious, the total area under this normal distribution is 1—again a characteristic of all probability distributions. This cursory introduction to both probability and probability distributions is meant to acquaint readers to concepts that will be used later when describing frequentist procedures, such as hypothesis testing, for making statistical inferences.

Table 1 Probabilities of obtaining heads from tossing three coins using the binomial distribution.

Number of heads	Probability
$x=0$	$\binom{3}{0} 0.5^0 0.5^3 = 0.125$
$x=1$	$\binom{3}{1} 0.5^1 0.5^2 = 0.375$
$x=2$	$\binom{3}{2} 0.5^2 0.5^1 = 0.375$
$x=3$	$\binom{3}{3} 0.5^3 0.5^0 = 0.125$

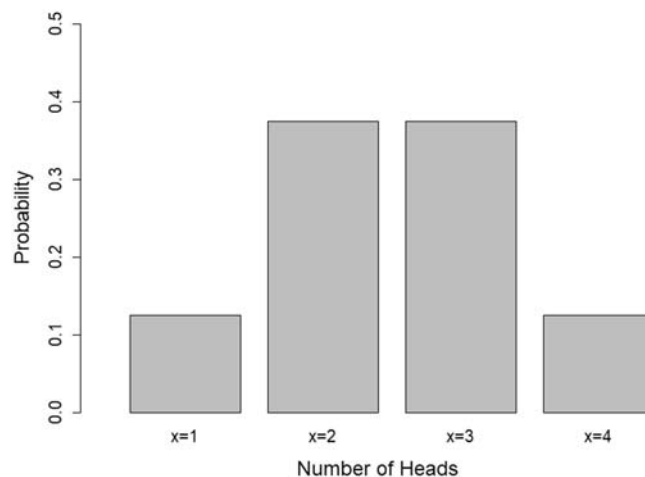


Fig. 2 Barplot of probabilities computed using the binomial distribution.

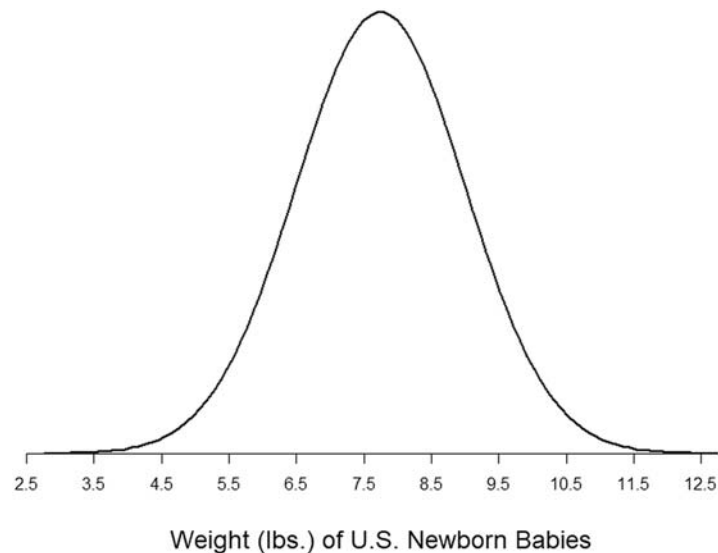


Fig. 3 Normal probability distribution of US newborn babies' weights in pounds.

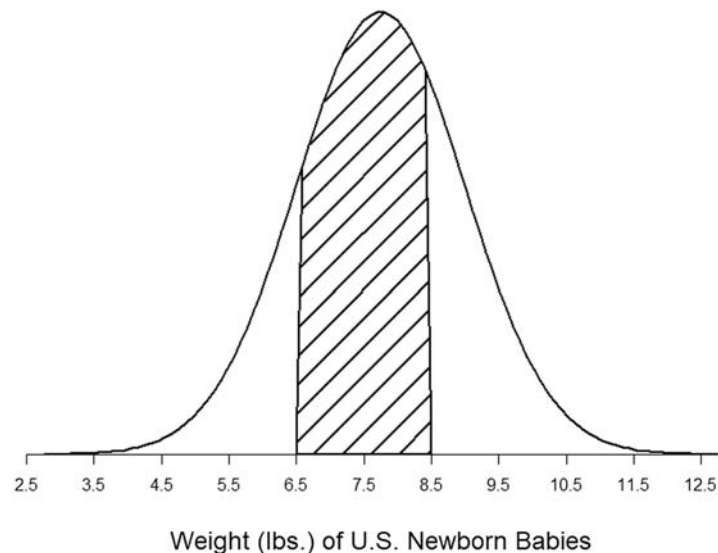


Fig. 4 Probability between an interval denoted by the shaded area under the curve with endpoints at 6.5 and 8.5 pounds.

Hypothesis testing

An essential, initial component of testing any hypothesis is the researcher's articulation of a research question. Most research questions in education are formulated based on an investigator's insightful observations or intuition founded in an explicit understanding of the underlying phenomenon. For example, a school district's reading specialist may conjecture that a new reading curriculum targeting bilingual children will produce higher reading comprehension scores as measured on a state-wide assessment of learning than the current curriculum (the research question). Stemming from his/her pedagogical content knowledge or expertise in bilingual education, the reading specialist may generate a research hypothesis that this new approach will impact student achievement. A research question (or research hypothesis) generally reflects a researcher's expectation of reality and should directly correspond to the statistical hypotheses to be empirically evaluated.

Translating hypotheses

A prerequisite to testing a research hypothesis is translating the question into a statistical hypothesis. A statistical hypothesis is a numerical statement regarding a hypothesized population parameter that captures the essence of the quantity articulated in

the research question. In the prior example involving a comparison between a new reading curriculum and current reading curriculum, a researcher might posit that average test scores on a standardized reading assessment for students randomly assigned to the new curriculum will be different than average test scores for students randomly assigned to the current curriculum. The corresponding statistical hypothesis operationalizes the description and formalizes the relation using appropriate parameters

$$H_0 : \mu_n - \mu_c = 0$$

$$H_1 : \mu_n - \mu_c \neq 0$$

where μ_n is mean of test scores in the population of students taking the new curriculum while μ_c is the mean of test scores in the population of students taking the current curriculum. Hypothesis testing involves comparing two hypotheses: (1) the null hypothesis and (2) the alternative hypothesis. The null hypothesis states that there is no relation (no effect) between the phenomena whose relation is under investigation. This corresponds to $H_0 : \mu_n - \mu_c = 0$ where average student test scores under the two reading curricula are the same (i.e., their difference is zero). The alternative hypothesis, as the name suggests, is the alternative to the null hypothesis stating that some kind of relation exists, $H_1 : \mu_n - \mu_c \neq 0$. Although introductory statistics courses highlight that the alternative hypothesis in this relation (e.g., one group's average test scores will be greater or less than the other group) may be directional (i.e., one-sided), the vast majority of research studies state a non-directional, two-sided, alternative hypothesis (e.g., the average test scores in the two groups are not the same). In general, for any population parameter, θ , the statistical hypothesis takes on the form

$$H_0 : \theta = K$$

$$H_1 : \theta \neq K$$

where K is any constant value—often, but not always—defined to be zero. Recall, that a parameter (a numeric quantity in the population) is unknown and a statistic computed from sample data is used as our best guess of this parameter. In the running example, $\bar{y}_n - \bar{y}_c$ would be used as a *point estimate* of $\mu_n - \mu_c$. This quantity is almost always non-zero—conforming to the alternative hypothesis. The obvious question is, “Is the observed average group difference in test scores small enough to attribute to chance variability or does the observed average group difference represent a real effect?” In other words, is the observed difference simply a result of the particular sample data we happened to obtain. If the group means in the population are actually the same, then any observed difference is attributable to *sampling error*. Answering this question will require the use of what is referred to as the statistic's sampling distribution.

Sampling distribution

A key concept of statistical inference from a frequentist perspective is the *sampling distribution* of a statistic, $\hat{\theta}$, and conveys specifically what degree of sample-to-sample variability (i.e., sampling error) can be expected by chance. In other words, the sampling distribution of a statistic (e.g., group mean difference) can be thought of as a distribution of values obtained for that statistic over repeated sampling (i.e., running the experiment, or drawing samples, an unlimited number of times). In hypothesis testing, sampling variability of a statistic, $\hat{y}_n - \hat{y}_c$ is the foundation for estimating the probability of observing a statistic equal to or more extreme than what is obtained under the true null hypothesis (i.e., no effect). To make this idea more concrete, assume we were interested in the difference in average performance on a standardized reading assessment for all second grade students assigned to receive either a new curriculum or the current curriculum (population). Our best guess of that population group mean difference ($\mu_n - \mu_c$) would be the sample group mean difference ($\bar{y}_n - \bar{y}_c$) based on a random sample of n_n and n_c observations from these two populations. If this process of drawing random samples of size n_n and n_c from the same populations were repeated many times and for each sample drawn, the mean difference between the groups was computed, the group mean differences from all the samples would vary and would form a distribution—the sampling distribution of the group mean differences. As the number of repeated samples increases, the sampling distribution is more well-defined. In the case of $\bar{y}_n - \bar{y}_c$ theoretically, the average of all possible sample group mean differences, or its expected value, $E(\bar{y}_n - \bar{y}_c)$, is equal to the population group mean difference, $\mu_n - \mu_c$. Statistics with this property, that $E(\hat{\theta}) = \theta$, are referred to as *unbiased estimators*.

Central limit theorem

The central limit theorem (CLT) provides a factual statement about the sampling distribution of the mean. Specifically, the CLT states that given a population with mean μ and variance σ^2 , the sampling distribution of the mean (i.e., the distribution of sample means) will have a mean equal to the mean of the population, μ (i.e., $\mu_{\bar{y}} = \mu$) and variance, $\sigma_{\bar{y}}^2$, equal to σ^2/n . This implies that as the sample size of the samples drawn from a population increases, the standard error of the means decreases, leading to less variability of sample means and smaller sampling error. The distribution of sample means will approach the normal distribution as n , the sample size, increases no matter what the shape of the distribution in the population is.

In the hypothetical example involving the group mean difference, testing the null hypothesis will require examining the sampling distribution of the difference between means. To be concrete, suppose $y_1 \sim N(\mu_1, \sigma_1^2)$ and $y_2 \sim N(\mu_2, \sigma_2^2)$ coming from sample sizes of n_1 and n_2 , respectively. Using the fact that the variance of the sum or difference between two independent random variables is the sum of their variances, and from the CLT that the variance of the distribution of $\bar{y}_1$ is σ^2/n_1 and the variance of the distribution of $\bar{y}_2$ is σ^2/n_2 , the variance of the difference of these two variables is the sum of their variances. Thus, the variance of the mean difference is

$$\sigma_{\bar{y}_1 - \bar{y}_2}^2 = \sigma_{\bar{y}_1}^2 + \sigma_{\bar{y}_2}^2 = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right) \quad (1)$$

Coupled with the information regarding the variability of the sample mean difference in Eq. (1), the implications of the CLT are shown in Fig. 5. As the sizes of the samples drawn from the two populations increase, the variance of the sample mean difference decreases. Fig. 5 presents two sampling distributions of the mean differences, one based on $n_1 = n_2 = 10$, and the other based on $n_1 = n_2 = 50$.

The black line represents the group mean difference in the population given that the means of the two groups are equal as they would be under the null hypothesis. They are at 0 in both plots. In fact, it can be shown that $E(\bar{y}_1 - \bar{y}_2) = \mu_1 - \mu_2$. Examining the scale on the x-axis, it is clear that the group mean differences comprising the sampling distribution in the left panel are more variable around the mean value of 0. Whereas, the group mean differences in the right panel are less variable—more tightly clustered around 0. The final point to be made about this distribution concerns its shape. An important theorem in statistics states that the sum or difference of two independent normally distributed variables is itself normally distributed. Because Fig. 5 represents the difference between two sampling distributions of the mean, and because the sampling distribution of means is at least approximately normal for reasonable sample sizes, the distributions in Fig. 5 must be at least approximately normal.

Standard error

Sample-to-sample variability, or sampling error, of a sample statistic is captured by the *standard error*. In the case of the group mean difference as described previously, the standard error is the standard deviation of its sampling distribution. When samples are randomly selected from corresponding populations, the resulting distribution will be approximately normally distributed for reasonable sample sizes, n_1 and n_2 . When the standard deviations in these two populations are known, the standard error of the sampling distribution of the group mean difference is the square root of the variance in Eq. (1)

$$\sigma_{\bar{y}_1 - \bar{y}_2} = \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)}$$

Clearly, the standard error ($\sigma_{\bar{y}_1 - \bar{y}_2}$) depends on the size of the variability of scores in each population (i.e., σ_1^2 and σ_2^2) as well as the sample sizes n_1 and n_2 . While not a requirement, assume that the sample sizes are equal (i.e., $n_1 = n_2$). As group sample size increases, $\sigma_{\bar{y}_1 - \bar{y}_2}$ decreases to its lower limit of 0. That is, $n_1 = n_2 \rightarrow \infty$, $\sigma_{\bar{y}_1 - \bar{y}_2} \rightarrow 0$. In a similar fashion, as group sample size decreases to its lower limit of 1, $\sigma_{\bar{y}_1 - \bar{y}_2}$ increases to its upper limit of the population standard deviation, $\sqrt{(\sigma_1^2 + \sigma_2^2)}$. Generally

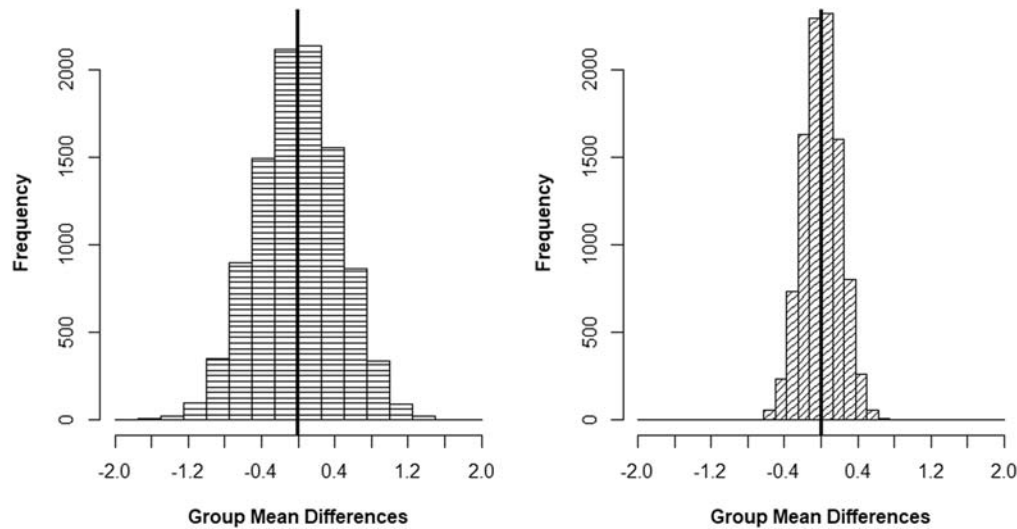


Fig. 5 Distributions of group mean differences from repeated sampling from populations in which the means are equal. Sample size in the left panel is $n_1 = n_2 = 10$. Sample size in the right panel is $n_1 = n_2 = 50$.

speaking, over repeated sampling, a sample statistic $\hat{\theta}$ based on a larger sample size will have less variability (i.e., less sampling error) than those based on a smaller sample size. In practice, the population standard deviations are almost never known. As we will see, a reasonable tact is to use sample standard deviations as our best guess at these unknown quantities.

Why is understanding these basic tenets of the sampling distribution of a statistic so crucial? Imagine in our running example that a sample mean difference of 0.75 was computed from the sample data (i.e., $\bar{y}_1 - \bar{y}_2 = 0.75$). Is this 0.75-point mean difference sufficiently large to justify the conclusion that the two samples were drawn from different populations as suggested by the alternative hypothesis. Or as Jones and Tukey (2000) proposed, is the difference sufficiently large for us to identify the direction of the difference in population means? Under the null hypothesis of no group mean difference, a mean difference of 0.75 does not seem so unusual based on the sampling distribution in the left panel of Fig. 5 in which the group sample size is 10, but seems quite improbable from examining the right panel in which the group sample size is 50. In addition, Fig. 5 intuitively demonstrates a number of properties of sampling distributions that will have significant implications for hypothesis testing.

Using the sampling distribution

The sampling distribution of a statistic under the presumption that the null hypothesis (H_0) is true is used to test statistical hypotheses. Because the sampling distribution is a probability distribution, this involves estimating the probability of observing a sample statistic (e.g., $\bar{y}_1 - \bar{y}_2$) equal to or more extreme than what is observed from the sample. For this probability distribution, the farther away the observed sample statistic (e.g., $\bar{y}_1 - \bar{y}_2$) is from the population parameter under the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$ (i.e., out on the tails of the distribution), the more likely that the difference between $\bar{y}_1 - \bar{y}_2$ and $\mu_1 - \mu_2$ represents a real mean difference in the population and is not due to sampling variation. A sample statistic value near the population value of the sampling distribution is more indicative of sample-to-sample variation. Sample statistic values in the tails of the distribution occur less frequently in repeated sampling, and consequently, are less probable if in fact H_0 was true. In fact, a probability quantifying how likely a sample statistic is can be calculated.

Under H_0 , if this probability is judged to be smaller than a predetermined threshold probability level (denoted α), the conclusion would be that H_0 is likely not true and the alternative hypothesis is likely operating in the population. This is the basic idea behind frequentist hypothesis testing—making conclusions about population parameters using probabilistic statements regarding the likelihood of observing a sample statistic if indeed the null hypothesis in the population was true. A small probability suggests that observing such a statistic is quite unlikely. The threshold value of $\alpha = 0.05$ is conventionally used in research (although this is arbitrary), leading to the following decision rules with regard to H_0 . In general, if the probability of observing the sample statistic is smaller than the threshold α given the null hypothesis is true (i.e., $p(\hat{\theta}|H_0 \text{ is true}) \leq \alpha$), then the null hypothesis should be rejected in favor of the alternative hypothesis. The null hypothesis would be retained if $p(\hat{\theta}|H_0 \text{ is true}) > \alpha$. As Moore (2004) points out, this decision rule does not suggest that such values of the sample statistic cannot occur under the null hypothesis, H_0 . It merely states that values so extreme (i.e., those out on the tails of the sampling distribution) arise with much less frequency under H_0 and thus seem implausible if in fact the null hypothesis was true.

It is important to keep in mind that hypothesis testing is unable to address the question of whether H_0 is true given the observed statistic computed from the sample (Howell, 2011) and a chosen threshold, α . The process only involves the determination of the probability of obtaining sample statistic values equal to or more extreme than the sample result under the assumption that the null hypothesis is true. Other threshold values like $\alpha = 0.01$ or $\alpha = 0.10$ can be specified and used to make decisions when evaluating H_0 . A particular value of α is chosen by the researcher before examining the data and is based on his or her consideration of committing Type I and Type II errors. Table 2 provides a schematic of the decision making process in frequentist hypothesis testing.

The process of distinguishing between the null hypothesis and the alternative hypothesis is aided by considering two conceptual types of errors. Type I error occurs when wrongly rejecting the null hypothesis. In the running example using group mean differences one might conclude that the two groups have different means in the population when in actuality they do not. The second type of error, Type II error, occurs when the null hypothesis is wrongly retained (or not rejected). That is, there was a population mean difference between the two groups, but we wrongly said that there wasn't. The other two scenarios correspond to correct decisions regarding the state of reality in the population.

Table 2 Correct decisions and errors in hypothesis testing.

		Decisions based on sample data	
		Reject H_0	Fail to reject H_0
True state in the population	H_0 is true	Type I error (significance level) α	Correct decision (confidence level) $1 - \alpha$
	H_0 is false	Correct decision (power) $1 - \beta$	Type II error β

Test statistics and confidence intervals

Sampling distributions are critical to conducting hypothesis tests because they place the one value of the statistic computed from sample data into the broader context of all other possible values that could have been obtained over repeated sampling from the population. In practice, this allows assessing the probability of observing values more extreme than a given sample statistic under H_0 . The sampling distribution can either be derived theoretically using key concepts like the CLT or the sampling distribution can be approximated using *bootstrapping* techniques (see, *Statistical inference: Resampling approaches* in this volume). For the former, a sample statistic is typically re-expressed as a test statistic for which characteristics of the sampling distribution are known.

For example, the test statistic for evaluating a sample mean difference between two groups, $\bar{y}_1 - \bar{y}_2$ when the standard deviations, σ_1^2 and σ_2^2 in the populations are known is:

$$z_{obs} = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{y}_1 - \bar{y}_2}}$$

This test statistic, z_{obs} , represents the distance between the sample mean difference and hypothesized population mean difference in standard deviation units, where the standard deviation is the standard error of the sampling distribution of the mean difference. A larger absolute value of z_{obs} indicates that $\bar{y}_1 - \bar{y}_2$ is further away from the hypothesized population mean under H_0 —in our running example, zero. The exact probability of observing a value of z_{obs} or larger can be obtained from tabled values found in many introductory statistics books or can be easily computed using numerous mainstream statistical software packages, even Microsoft Excel.

Applying the z -test as described above requires that the population standard deviations, σ_1^2 and σ_2^2 , be known. When these values are unknown, the standard error of the mean difference must be computed from the sample data, $s_{\bar{y}_1 - \bar{y}_2}$, where $s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}$.

The test statistic in this case is a t -statistic

$$t_{obs} = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{s_{\bar{y}_1 - \bar{y}_2}}$$

which follows a t -distribution. The t -distribution is actually a family of distributions which is characterized by degrees of freedom (df). Three members of the t -distribution family are displayed in Fig. 6 with an increasing number of df ($df = 3, 10$, and ∞). Note, that as the total sample size, $N = n_1 + n_2$, increases, the t -distribution converges on the z -distribution. Beyond $df = 120$, the difference between the z - and t -distributions is negligible—the t -test becomes the z -test for all practical purposes. Test statistics other than z and t may be needed as required by a statistical analysis. For example, the χ^2 distribution is used to test whether a variance deviates from a population value (e.g., 0); or, the F distribution is used for testing two variance components. Regardless of the specific test statistic used in a particular research situation, the logic for hypothesis testing described here, remains the same.

Adjudicating whether to reject the null hypothesis can be accomplished in three different, but congruent, ways. The first method directly compares the observed sample test statistic (e.g., t_{obs} , z_{obs} , or χ_{obs}^2) against a critical value of the statistic (e.g., t_{cv} , z_{cv} , or χ_{cv}^2) as determined by the predetermined significance level, α . Similarly to computing probabilities, the critical value of a t -test statistic (i.e., t_{cv}) corresponding with a given value of α and a given df is readily available from statistical software. The second method is one discussed earlier that focuses on carrying out the hypothesis test by comparing the probability of obtaining sample statistic values more extreme than a given sample statistic, assuming the H_0 is true, against a predetermined threshold probability level (i.e., $\alpha = 0.05$). The third approach uses a $100(1 - \alpha)\%$ confidence interval of the population parameter (see upcoming section on confidence intervals). The following decision rules for testing H_0 can be established for the first two of the three approaches.

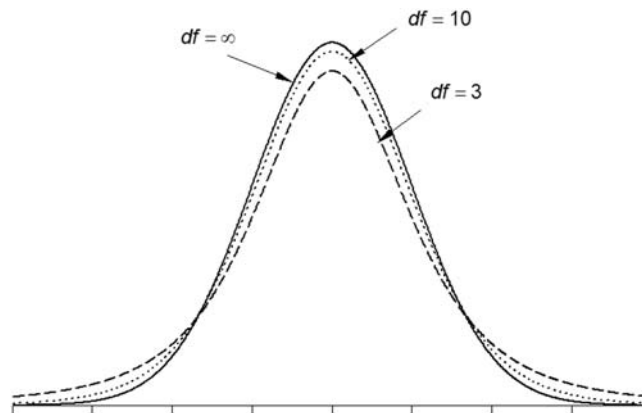


Fig. 6 T-distributions with $df = 3$ (— — —), $df = 10$ (·····), and $df = \infty$ (—).

The first decision rule compares t_{obs} against t_{cv} leading to

If $|t_{obs}| \geq |t_{cv}| \rightarrow \text{Reject } H_0$
 If $|t_{obs}| < |t_{cv}| \rightarrow \text{Fail to Reject } H_0$

This decision rule is equivalent to the previous, probability-based decision rule for testing H_0 :

If $p(t_{obs}|H_0) \leq \alpha \rightarrow \text{Reject } H_0$
 If $p(t_{obs}|H_0) > \alpha \rightarrow \text{Fail to Reject } H_0$

Hypothetical example

Returning to the running hypothetical example involving testing the mean difference in reading scores between students assigned to receive a new curriculum and those students assigned to receive the current reading curriculum. A researcher obtains test scores for $n_n = n_c = 100$ students and computes the following sample statistics: $\bar{y}_n = 75$, $\bar{y}_c = 66$, $s_n = 25$, and $s_c = 23$ of students in the two groups and is interested in testing the non-directional null hypothesis, $H_0 : \mu_n - \mu_c = 0$ at an $\alpha = 0.05$ significance level. The test statistic is

$$t_{obs} = \frac{(\bar{y}_n - \bar{y}_c) - (\mu_n - \mu_c)}{s_{\bar{y}_n - \bar{y}_c}} = \frac{(75 - 66) - 0}{3.40} = 2.65$$

where $s_{\bar{y}_n - \bar{y}_c} = \sqrt{\left(\frac{s_n^2}{n_n} + \frac{s_c^2}{n_c}\right)}$. Under the assumption of unequal population variances, $|t_{obs}| = 2.65$ exceeds the $t_{cv} [t_{(0.975, df=196.639)} = 1.972, \alpha = 0.05]$.² Thus, the null hypothesis, $H_0 : \mu_n - \mu_c = 0$, is rejected as it is very improbable ($p = 0.009$) to observe sample mean difference values equal to or greater than 9 if in fact H_0 was true.

It is highly recommended that researchers go beyond simply reporting of the results obtained from testing hypotheses to also include some measure of the size of an effect, often referred to as the *effect size*. The importance of using effect sizes in educational and psychological research was emphasized in a report by the American Psychological Association (APA) task force on statistical inference [Wilkinson et al. \(1999\)](#), which stated unequivocally that effect sizes of primary outcomes should always be presented with additional comments that place these effect sizes in a practical and theoretical context. In fact, professional journals that regularly publish research studies that use quantitative methods, currently insist on them.

In addition, some methodologists have strongly argued that statistical significance is the least interesting thing about the results and should also be described in terms of its magnitude. An appropriate effect size for our running example of the mean difference in reading scores in two experimental conditions would be a standardized mean difference as measured by Cohen's d or Hedge's g . The type of effect size would convert the actual mean difference of 9 points onto a standard deviation scale. Colloquially, a statistically significant result suggests that the new reading curriculum affects students, but an effect size provides information on how much does it affect them. For our running example, Cohen's d would be calculated as

$$d = \frac{(\bar{y}_n - \bar{y}_c)}{s_{pooled}} = \frac{(75 - 66)}{24.021} = 0.375$$

The value 0.375 can be interpreted as the standardized mean difference in reading scores between the two groups. This might be categorized as a small-to medium-size effect ([Cohen, 1988, 1994; Kelley, 2007](#)).

Confidence intervals

The method of testing a hypothesis described above is a point-estimate approach that results in a dichotomous decision (i.e., reject or fail to reject H_0). An interval estimation approach can also be used to test a statistical hypothesis and provides the additional benefit of producing a range of values within which a population parameter is likely to reside. That is, interval estimates are attempts to set limits that have a high probability of encompassing the population parameter. From a frequentist perspective, these intervals are called *confidence intervals* (CI). In general, a CI for an unknown parameter is a recipe that, in repeated samples, yields intervals that contain the true parameter with a prespecified probability, the confidence level, $1 - \alpha$. For instance, a 95% CI would be interpreted to mean that 95% of similarly constructed intervals will contain the population parameter while 5% of them will not. This concept is illustrated in [Fig. 7](#) for a hypothetical situation where 20 random samples were drawn from two populations and the sample mean difference was subsequently computed and a 95% confidence interval was constructed from each sample. Out of the 20 CIs for $\mu_1 - \mu_2$, 19 (or 95%) contained the population parameter (i.e., true mean difference), and 1 did not.

²If equal variances are assumed in the two groups, then the formula for the standard error reduces to a pooled estimate, $s_{\bar{y}_n - \bar{y}_c} = s_{pooled} \sqrt{(1/n_n) + (1/n_c)}$,

where $s_{pooled} = \sqrt{\frac{[(n_n - 1)s_n^2 + (n_c - 1)s_c^2]}{(n_n + n_c - 2)}}$.

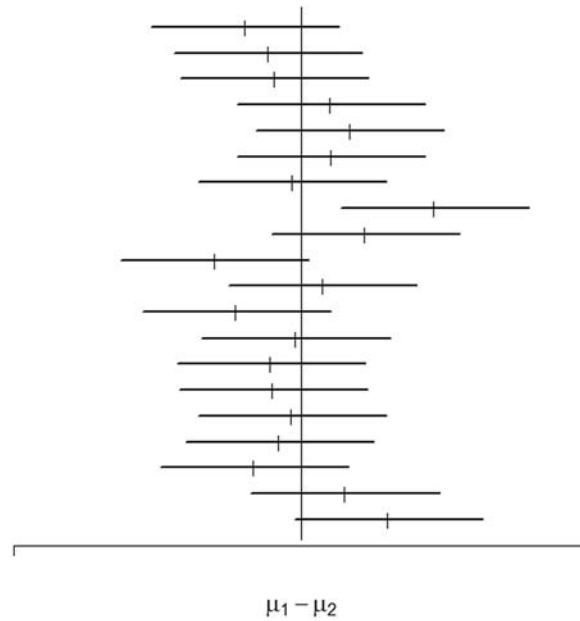


Fig. 7 Twenty 95% confidence intervals for the population mean difference, $\mu_1 - \mu_2$.

It is very tempting to say that the probability is 0.95 that the interval includes the population mean difference. Confidence intervals are random variables and will vary from one study to the next while the population parameter is fixed. A specific interval—once calculated from the sample data—is no longer a random variable and therefore does not have a probability associated with it.

A CI can be constructed for a population parameter, θ , using $\hat{\theta}$ estimated from the sample. For a given significance level, α ,

$$100 \cdot (1 - \alpha)\% \text{ CI} = \hat{\theta} \pm t_{\left(1-\frac{\alpha}{2}\right)} \cdot SE_{\hat{\theta}}$$

This is the conventional definition of a symmetric CI under a true H_0 , where $t_{\left(1-\frac{\alpha}{2}\right)}$ is the appropriate test statistic (e.g., z , t) value at the $100 \times \left(1 - \frac{\alpha}{2}\right)$ percentile point of its distribution, and $SE_{\hat{\theta}}$ is the estimated standard error of the sampling distribution. Upper and lower limits of the CI can be obtained by adding and subtracting the quantity $t_{\left(1-\frac{\alpha}{2}\right)} \cdot SE_{\hat{\theta}}$ resulting in the CI:

$\left(\hat{\theta} - T_{\left(1-\frac{\alpha}{2}\right)} \cdot SE_{\hat{\theta}}, \hat{\theta} + T_{\left(1-\frac{\alpha}{2}\right)} \cdot SE_{\hat{\theta}} \right)$. This has a $(1 - \alpha)$ probability of containing the unknown population parameter, θ . As was mentioned previously, using a CI is the third method of adjudicating a hypothesis test. Given H_0 is true, the hypothesized population parameter value falling outside a CI constructed in this manner leads to the rejection of H_0 . Conversely, a CI containing the parameter value would lead to the decision of failing to reject H_0 .

Returning once again to our hypothetical running example about evaluating a new reading curriculum, where $\bar{y}_n - \bar{y}_c = 9$, $s_{\bar{y}_n - \bar{y}_c} = 3.40$, and t critical value is $t_{(0.975, df=196.639)} = 1.972$ at the $\alpha = 0.05$ level. The 95% CI for $\mu_n - \mu_c$ would be (2.30, 15.71) with the lower limit computed as $9 - (1.972 \times 3.4)$ and the upper limit computed as $9 + (1.972 \times 3.4)$. As this CI does not contain 0, the value under the null hypothesis, then H_0 would be rejected (just as it was using the other two decision rules).

Discussion

Indisputably, modern scientific thinking has been greatly enhanced by thoughtful statistical reasoning. Part of that reasoning is directed at making inferences about a population based on sample data. From a frequentist viewpoint, statistical inference involves many interrelated activities. This non-exhaustive set includes, (1) choosing a statistic that best represents the phenomena to be investigated, (2) obtaining a random sample from a target population, (3) estimating the statistic using sample data, and (4) using the null hypothesis significance testing (NHST) paradigm and confidence intervals to make probabilistic statements about the true values of corresponding parameters in the population. This paradigm has dominated the educational, social and behavioral science research landscape over the past 90+ years since the seminal works by Fisher (1925) and Neyman and Pearson (1928). Perezgonzalez (2015) provides an informative account of the chronology of hypothesis testing up through the beginning of the 21st century

and Lehmann (1993) provides a readable commentary regarding the similarities and differences between Fisher and Neyman and Pearson's perspectives on hypothesis testing. Yet, the NHST paradigm is not without its critics, which have led to alternative statistical inferential paradigms.

In 2019, the American Statistical Association (ASA) published a special issue related to statistical inference in the 21st century. In the leading article, Wasserstein et al. (2019) wrote on behalf of the ASA stating.

The ASA statement on p -values and statistical significance stopped just short of recommending that declarations of “statistical significance” be abandoned. We take that step here. We conclude, based on our review of the articles in this special issue and the broader literature, that it is time to stop using the term “statistically significant” entirely. Nor should variants such as “significantly different”, $p < 0.05$, and “nonsignificant” survive, whether expressed in words, by asterisks in a table, or in some other way. (p. 2).

The sentiments (or mandates) embodied in this quote and throughout this special issue in the American Statistician should give each applied scientist and quantitative methodologist pause. While the NHST paradigm is unlikely to disappear anytime soon, there are a few take-away messages that are worth highlighting with these criticisms in mind. Be open, thoughtful, and modest when interpreting statistical results and drawing inferences. An inferential statement provides only an approximation to the truth and therefore, should be accompanied and contextualized by some measure of uncertainty associated with it. As Wasserstein later asserted, “... small p -values are like a right-swipe in Tinder. It means you have an interest. It doesn't mean you're ready to book the wedding venue.”

Two alternatives to the NHST paradigm already exist and worth mentioning here. The first is statistical inference from a Bayesian point-of-view. Bayesian inference (see *Statistical inference: Bayesian approaches*, an entry in this volume) is an approach to statistical inference in which Bayes' theorem is used to update the *probability for a hypothesis* as more evidence or information about model parameters and additional data becomes available. The second inferential paradigm is model-based (see, e.g., Anderson, 2008; Burnham and Anderson, 2002) using information criteria such as Akaike's information criterion (AIC; Akaike, 1974) or the Bayesian information criterion (BIC; Schwarz, 1978). This paradigm focuses on the principle of multiple working hypotheses that encourages researchers to think hard to identify several plausible scientifically-justifiable hypotheses—where inferences are conditional on this set of alternatives. Each scientific hypothesis corresponds to a statistical model thought to have generated the data whose parameters capture some interesting facet of the phenomenon. The researcher provides several pieces of *evidence* arguing why one should prefer a particular hypothesis over another—with the data as the final arbiter. Both Bayesian and informatic-theoretic approaches are relatively new to the social and behavioral sciences and therefore are not yet well established in the applied literature. It is unclear whether one paradigm for statistical inference is better, or more effective, than another. However, one thing is certain. The landscape for making statistical inferences in a frequentist framework is changing and will continue to do so.

References

- Akaike, H., 1974. A new look at the statistical model identification. In: IEEE Transactions on Automatic Control AC-19, pp. 716–723.
- Anderson, D.R., 2008. Model Based Inference in the Life Sciences: A Primer on Evidence. Springer, New York, NY.
- Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference, a Practical Information-Theoretic Approach. Springer, New York, NY.
- Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed.
- Cohen, J., 1994. The earth is round ($p < 0.05$). Am. Psychol. 49, 997–1003.
- Fisher, R.A., 1925. Statistical methods and scientific induction. J. Roy. Stat. Soc. B 17, 69–78.
- Garfield, J.B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., Zieffler, A.S., 2008. Learning to reason about statistical inference. In: Garfield, J.B., Ben-Zvi, D. (Eds.), Developing Students' Statistical Reasoning. Springer, Dordrecht, Germany, pp. 261–288.
- Howell, D.C., 2011. Statistical Methods for Psychology, eighth ed. Wadsworth Cengage Learning, Belmont, CA.
- Jones, L.V., Tukey, J.W., 2000. A sensible formulation of the significance test. Psychol. Methods 5, 411–414.
- Kelley, K., 2007. Confidence intervals for standardized effect sizes: theory, application, and implementation. J. Stat. Software 20, 1–24.
- Lehmann, E.L., 1993. The Fisher, Neyman–Pearson theories of testing hypotheses: one theory or two? J. Am. Stat. Assoc. 88, 1242–1249.
- Moore, D.S., 2004. The Basic Practice of Statistics, third ed. W. H. Freeman, New York, NY.
- Neyman, J., Pearson, E.S., 1928. On the use and interpretation of certain test criteria for purposes of statistical inference: part I. Biometrika 20A, 175–240.
- Perezgonzalez, J.D., 2015. Fisher, Neyman–Pearson or NHST? A tutorial for teaching data testing. Front. Psychol. 6, 223.
- Schwarz, G., 1978. Estimating the dimensions of a model. Ann. Stat. 6, 461–464.
- Stevens, S.S., 1946. On the theory of scales of measurement. Science 103, 677–680.
- Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a world beyond $p < 0.05$. Am. Stat. 73, 1–19.
- Wilkinson, L., Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999. Statistical methods in psychology journals: guidelines and explanations. Am. Psychol. 54, 594–604.

Resampling methods

Dennis D. Boos, Department of Statistics, North Carolina State University, Raleigh, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	762
Fundamentals	763
Statistical populations	763
Random samples	763
Monte Carlo simulation	763
The bootstrap	764
Bootstrap confidence intervals	765
Bootstrap hypothesis testing	766
Parametric bootstrap	766
When does the nonparametric bootstrap “work”?	767
Bootstrap resources	767
The jackknife	767
Examples	768
Further comments	768
Permutation tests	769
One sample (symmetry) or two paired samples	769
Two independent samples	770
General notation and approach	770
Further comments on permutation tests	770
Summary	770
References	771

Overview

Resampling in statistics generally refers to creating new data sets from a random sample $Y = (Y_1, \dots, Y_n)^T$ to help with inference based on estimators or test statistics, or any quantity that is simply a function of Y like the sample mean or sample median or the t statistic. Actually, we might be interested in $t = \sqrt{n}(\bar{Y} - \mu)/s_{n-1}$ as a test statistic with $\mu = \mu_0$ a known constant or as a *pivotal quantity* with unknown μ used to get confidence intervals. Denote any of these random quantities generally by $T(Y)$. The focus of this entry is on the bootstrap, the jackknife, and permutation tests. All three methods are used to estimate certain aspects of the sampling distribution of $T(Y)$ such as the variance or percentiles of $T(Y)$. Other resampling methods (not discussed here) include cross-validation, random subsampling, and balanced repeated replications—see [Efron \(1982\)](#) and [Efron and Tibshirani \(1993\)](#) for discussion of these.

It is very easy to be confused by the two levels of distributions in play here. The population that Y is sampled from may have a mean μ or a 95th percentile $\eta_{0.95}$, but those are not the direct focus of the resampling methods. To be clear, let us consider the simplest situation where $Y_1, \dots, Y_n$ are independent and identically distributed from a population with mean μ and variance σ^2 . In this case, we know that the sample mean $\bar{Y}$ estimates μ unbiasedly, $E(\bar{Y}) = \mu$, and its variance is $\text{Var}(\bar{Y}) = \sigma^2/n$. These latter calculations are based on the sampling distribution of $\bar{Y}$, which is inherited from the distribution of the sample. In fact, we know from the Central Limit Theorem that the sampling distribution of $\sqrt{n}(\bar{Y} - \mu)/\sigma$ converges to a standard normal distribution in large samples. In this simple case, the nonparametric bootstrap estimate of the variance of $\bar{Y}$ is s_n^2/n , where $s_n^2 = n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$ is

the “ n ” version of the sample variance. The jackknife estimator is s_{n-1}^2/n , where $s_{n-1}^2 = (n-1)^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$ is the “ $n-1$ ” version of the sample variance.

Now, I give a brief introduction to our three resampling methods.

1. The bootstrap introduced by [Efron \(1979\)](#) is the simplest of the resampling methods and is based on drawing random samples from the original sample, mimicking the original sampling plan as closely as possible. Its scope of application is very wide and can be used for bias and variance estimation, confidence interval construction, hypothesis testing, and for other less well-known inferential goals.
2. The jackknife was introduced by [Quenouille \(1949, 1956\)](#) as a general method to remove bias from estimators, but [Tukey \(1958\)](#) noticed it also led to a method for estimating variances, and that quickly became its main use. Resampling with the

jackknife is simply creating new samples by dropping out each data point one by one, thereby creating n data sets of size $n-1$. The $T(Y)$ of interest is calculated on each of the “delete-1” data sets, and the jackknife variance estimate is based on the sample variance of these n T s.

3. In the early 1930s, R. A. Fisher discovered a very general method of testing hypotheses based on permuting the data in ways that do not change its joint distribution under the null hypothesis. These permutations (data modifications) are the resamples. A test statistic calculated for each resample allows construction of a null distribution for the test statistic. This *permutation* method does not require standard parametric assumptions such as normality of the data. It does require, however, certain invariance properties under the null hypothesis that restricts application to fairly simple designs. But in such situations, the method results in exact tests with level α under very weak distributional assumptions. Moreover, the method is *statistic-inclusive* in the sense that any test statistic can be used and inherits the level- α property, although some test statistics are much more powerful than others. Tests based on this method are called permutation tests (or randomization tests).

Fundamentals

Here we discuss very briefly the components required to understand the bootstrap: random sampling from populations and Monte Carlo simulation.

Statistical populations

To understand our three resampling methods, we need to go back to the basics underlying statistical inference about a population. A statistical population is simply a set of numbers such as

$$\{x_1, x_2, \dots, x_N\}, \quad \text{for example } \{10, 10, 20, 20, 20, 30, 40, 50\}. \quad (1)$$

If there are replications like the two 10s and three 20s above with $N = 8$, we would just give the unique values $\{10, 20, 30, 40, 50\}$ and the relative frequencies $p_1 = 2/8$, $p_2 = 3/8$, $p_3 = 1/8$, $p_4 = 1/8$, and $p_5 = 1/8$. These relative frequencies may be thought of as probabilities of randomly drawing one member from the population. Taken together, the possible values and the relative frequencies define a population and are also called the *distribution* of the possible values. Here we might be interested in estimating the population mean

$$\mu = (2/8) * 10 + (3/8) * 20 + (1/8) * 30 + (1/8) * 40 + (1/8) * 50 = 25.$$

Often, we use statistical models, also called parametric models, to describe populations. The two best-known models for discrete populations are the *binomial distribution* and the *Poisson distribution*. The best-known continuous data models are the *normal distribution* and the *exponential distribution*.

The goal of statistical inference is to estimate population parameters like the mean μ or variance σ^2 with some notion of estimator variability, or form a confidence interval that gives an interval of plausible values for the population parameter, or carry out a hypothesis test of a specified null value θ_0 . In order to carry out statistical inference, we must obtain a random sample from the population of interest.

Random samples

Suppose that we get a randomly selected member of a population, Y_1 . The distribution of that Y_1 is the same as the population distribution, and it has the same mean and variance and other characteristics of the population. However, to carry out statistical inference, we need a bunch of similarly randomly selected members of the population denoted by $Y_1, \dots, Y_n$ and called a random sample. Characteristics of the sample like the sample mean $\bar{Y}$ and sample variance s_{n-1}^2 are used to estimate the analog population characteristics. (There are many other types of random samples, like stratified or cluster survey sampling, but for simplicity we stick with the simplest form.)

Monte Carlo simulation

This last “fundamental” is crucial for understanding the standard way to compute the bootstrap. The field of statistics has a long history of studying parameter estimators and associated confidence intervals and hypothesis tests. However, statistical theory has trouble dealing with the sampling distributions of estimators that are more complex than $\bar{Y}$ and s_{n-1}^2 . So, how can we evaluate whether new procedures in complicated situations are useful? The answer is that we run computer experiments to evaluate the new procedures, and we do something very similar with the bootstrap. On a computer, it is very easy to generate N independent

random samples, each of size n , from a population like the Poisson or normal. Here is what these N samples look like (Y is a column vector, hence the use of the superscript T for transpose).

$$\begin{aligned} Y_1 &= (Y_{1,1}, \dots, Y_{1,n})^T \\ Y_2 &= (Y_{2,1}, \dots, Y_{2,n})^T \\ &\dots \\ &\dots \\ &\dots \\ Y_N &= (Y_{N,1}, \dots, Y_{N,n})^T \end{aligned} \quad (2)$$

Let $T(Y)$ be any sample statistic of interest, for example, a parameter estimate or a test statistic. Compute $T()$ for each of the above N samples resulting in $T(Y_1), \dots, T(Y_N)$. This is not a random sample from the original population, but rather a random sample from the sampling distribution of $T()$ when samples of size n are generated from the original population. We can use this sample to estimate anything regarding the sampling distribution of $T()$. For example, we can estimate the mean using $\bar{T} = N^{-1} \sum_{i=1}^N T(Y_i)$

and the variance using $s_T^2 = (N-1)^{-1} \sum_{i=1}^N (T(Y_i) - \bar{T})^2$. The mean would tell us about the bias of an estimator. The variance could be compared with the variance of competing estimators. We could evaluate the rejection rate of test statistics under the null (Type I error rates) and under alternatives (power = 1 - Type II error rates). Keep in mind, we can do all this because we are generating the data from a known distribution. When using the bootstrap, we do not know the true population that the sample comes from.

The bootstrap

Suppose that we have a sample $Y = (Y_1, \dots, Y_n)^T$ from an unknown population and an estimator $\hat{\theta}(Y)$ for some population parameter θ . How can we estimate the variance of $\hat{\theta}(Y)$ using the nonparametric bootstrap method? We simply mimic the Monte Carlo method of the previous section, but instead of generating the samples from the true population, we treat the sample itself as a population and draw B samples from it, resulting in

$$\begin{aligned} Y_1^* &= (Y_{1,1}, \dots, Y_{1,n})^T \\ Y_2^* &= (Y_{2,1}, \dots, Y_{2,n})^T \\ &\dots \\ &\dots \\ &\dots \\ Y_B^* &= (Y_{B,1}, \dots, Y_{B,n})^T \end{aligned} \quad (3)$$

We then calculate $\hat{\theta}$ for each Y_i^* resulting in a random sample $\hat{\theta}_i^* = \hat{\theta}(Y_i^*), \dots, i = 1, \dots, B$. Note that these are conditionally independent, holding the original sample $Y_1, \dots, Y_n$ fixed. The first thing of interest is to estimate the variance of $\hat{\theta}(Y)$ by the sample variance of the $\hat{\theta}_i^*$,

$$\hat{V}_{\text{Boot}} = s_{B-1}^2 = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i^* - \bar{\hat{\theta}}^*)^2.$$

Note that we have just described the *nonparametric bootstrap*; later we discuss the *parametric bootstrap* approach that draws samples exactly like the Monte Carlo method of the previous section (but using estimated parameters).

For illustration, consider the following $n = 13$ college course evaluations ("How effective was the instructor?") for the same class and instructor over multiple semesters, ordered from smallest to largest:

3.60 3.85 4.02 4.03 4.09 4.14 4.22 4.44 4.52 4.52 4.55 4.62 4.67

The sample median is the middle value, 4.22. Let us use $B = 10$ resample data sets to estimate the variance of the sample median—without the bootstrap, this is a non-trivial task. To draw a bootstrap resample, just randomly select $n = 13$ values with replacement from the above data set. Then repeat $B - 1$ more times. Here are the first three resamples, each sorted from smallest to largest:

3.60 3.85 4.02 4.02 4.09 4.14 4.22 4.44 4.52 4.52 4.52 4.55 4.67
3.60 3.60 4.02 4.02 4.14 4.14 4.22 4.44 4.52 4.62 4.62 4.67 4.67
3.60 3.85 3.85 3.85 4.03 4.03 4.44 4.44 4.52 4.52 4.55 4.62 4.67

Note that in the first resample, 4.02 shows up twice, and 4.52 shows up three times. These repetitions are the result of using sampling with replacement (Note that using sampling *without replacement* would give the original sample every time). The medians of these three resamples are 4.22, 4.22, and 4.44, respectively. The full set of $B = 10$ medians is

4.22 4.22 4.44 4.22 4.09 4.52 4.44 4.44 4.03 4.09.

The sample variance of these 10 medians is 0.0390. If repeat this process with $B = 10$ three more times, we get bootstrap estimates of variance $\hat{V}_{\text{boot}}$ of the median = 0.0390, 0.0354, 0.0347, and 0.0370. Why are these four bootstrap estimates different? It is because each estimate is based only on samples of size $B = 10$. We call this the variability due to Monte Carlo sampling. Raising B to $B = 1000$, we get bootstrap estimates 0.0304, 0.0294, 0.0301, 0.0301, with only differences in the 3rd decimal place. Letting B get very large, essentially gives the “official” bootstrap variance estimate 0.030 with no Monte Carlo variability.

However, let us consider a hypothetical second sample of $n = 13$ course evaluations from the same population. The sample median would of course be different from 4.22 of the original sample, but the “official” bootstrap variance estimate of the median, with no Monte Carlo variability, is likely to be much different than the 0.030 of the first sample. That difference comes from the sample-to-sample variability of the “official” bootstrap variance estimate. Thus, when using $B = 10$ or even $B = 1000$ to estimate the “official” bootstrap variance estimate, there are two sources of variation: (i) the Monte Carlo variation that can be driven to zero by choosing B larger and larger, and (ii) the sample-to-sample variation that cannot be made smaller unless we can increase the sample size n of the original sample.

For variance estimation, the typical recommendation is to use between $B = 100$ and $B = 1000$ resamples because the sample-to-sample variance of the “official” variance estimate (based on conceptually repeating the original collection of $n = 13$ evaluations) is much higher than the Monte Carlo variability in the variance estimates due to using $B < \infty$.

Bootstrap confidence intervals

Since the sample median has an approximate normal distribution in large samples, we might use $4.22 \pm 1.96 * \sqrt{0.030}$ yielding the approximate 95% interval (3.88, 4.56). Efron and Tibshirani (1993, p. 154) call this the *standard confidence interval*. Efron’s (1981) first bootstrap interval method was merely to take the 2.5th and 97.5th percentiles of the bootstrap sample $\hat{\theta}_1^*, \dots, \hat{\theta}_B^*$, and he called it the *bootstrap percentile interval*. Efron generally recommends B to be larger for confidence intervals than for variance estimation, perhaps 2000 or more. Using $B = 2000$ here, we obtain (4.02, 4.52). Efron (1981, 1987) also proposed more advanced bootstrap intervals call the bias-corrected (BC) and BC_a (bias-corrected accelerated), but for this example they are essentially the same as the percentile method. These more complex intervals have theoretically better convergence properties but often do not improve much over the percentile method in small samples.

The proportion of A’s received in the 13 classes above was also available. Here are both variables in the original order. A plot of the data is in Fig. 1

class	1	2	3	4	5	6	7	8	9	10	11	12	13
Eval	4.02	4.14	4.09	3.85	4.03	3.60	4.22	4.52	4.44	4.67	4.62	4.52	4.55
A’s	0.22	0.34	0.42	0.29	0.26	0.30	0.38	0.52	0.44	0.45	0.56	0.74	0.32

The Pearson correlation coefficient estimate is $r_p = 0.66$, and the Spearman rank correlation estimate is $r_s = 0.76$. Using $B = 10,000$, the bootstrap estimate of variance for r_p is 0.013 and for r_s it is 0.022. For observing Monte Carlo variation, the

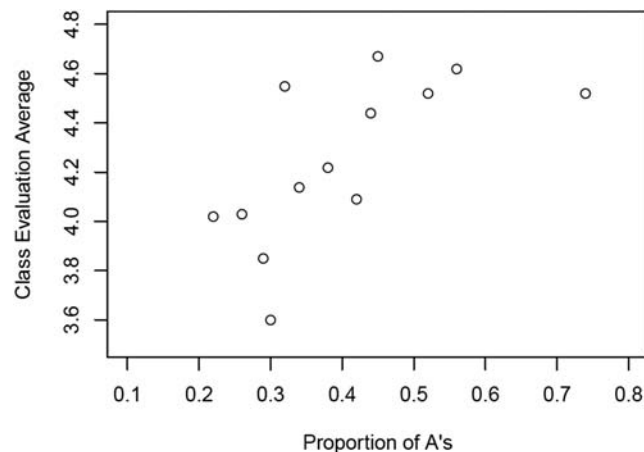


Fig. 1 Class average evaluations versus proportion of A’s in each course.

bootstrap estimate of variance for r_p using five independent runs of $B = 100$ was 0.012, 0.017, 0.013, 0.014, and 0.010. To completely avoid that variation, we used $B = 10,000$.

The associated standard 95% confidence interval for ρ_p is $0.66 \pm 1.96 * \sqrt{0.013}$, or (0.44, 0.88). For the Spearman correlation ρ_s , it is $0.76 \pm 1.96 * \sqrt{0.0215}$, or (0.47, 1.00), where we truncated at 1.00. The bootstrap percentile intervals are (0.44, 0.88) and (0.36, 0.93), respectively. Perhaps the percentile interval for ρ_s is shifted to the left of (0.47, 1.00) because the estimate 0.76 is somewhat close to its upper bound of 1.00.

Bootstrap hypothesis testing

Suppose that we have two independent samples $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ and would like to test the null hypothesis of equal means $H_0 : \mu_X = \mu_Y$. For simplicity, let us use the pooled t statistic, $t_p = (\bar{X} - \bar{Y}) / \sqrt{s_p^2(1/m + 1/n)}$, where s_p^2 is the usual pooled variance estimate. Following the above methods, we might draw B sets of independent bootstrap samples from the X and Y samples, and calculate t_p for each one. Unfortunately, this would not be appropriate for hypothesis testing because we need to estimate the sampling distribution of t_p under the null hypothesis, that is with equal means. But here, the bootstrap sampling would be from populations with means $\bar{X}$ and $\bar{Y}$, respectively. One way to resample and force equal means is to get draw bootstrap samples with replacement from $(X_1, \dots, X_m, Y_1, \dots, Y_n)$. This approach actually forces both resamples to be drawn from the exact same population (and is similar to the permutation approach of the next section). An alternative that allows for different “population” distributions (and different variances) but equal means is to sample X^* with replacement from $(X_1 - \bar{X}, \dots, X_m - \bar{X})$ and Y^* from $(Y_1 - \bar{Y}, \dots, Y_n - \bar{Y})$. The key idea when bootstrap sampling for hypothesis testing is that the null hypothesis must be preserved (or forced) in the bootstrap pseudo-populations.

Parametric bootstrap

In some situations, one might be comfortable assuming that the sample $Y = (Y_1, \dots, Y_n)^T$ are independent and identically distributed from a parametric family like the exponential family having density $f(y; \sigma) = (1/\sigma)\exp(-y/\sigma)$, $y > 0, \sigma > 0$. The natural estimator of σ is the sample mean $\bar{Y}$, and the estimated density is then $f(y; \bar{Y})$. The parametric bootstrap draws resamples from this estimated density instead of treating the sample as a population and drawing resamples from it. This method is exactly like the Monte Carlo method of the [Fundamentals](#) section except that here σ is replaced by $\bar{Y}$. The main downside of the parametric bootstrap is that it depends on the parametric assumption being correct, which is the reason that the nonparametric bootstrap is dominant in practice. The following example, however, is a perfect place for the parametric bootstrap leading to an exact hypothesis test (up to Monte Carlo approximation).

Table 11.1 of [Boos and Stefanski \(2013, p. 414\)](#) reports 35 yearly maximum flow rates (gallons per second) at one river gauging station in North Carolina. For subsequent analysis it was important to confirm that the maxima actually follow an extreme value (Gumbel) distribution with cumulative distribution function (CDF) $F(x; \mu, \sigma) = F_0((x - \mu)/\sigma)$, where $F_0(y) = \exp\{-\exp(-y)\}$. The left panel of [Fig. 2](#) gives a histogram of the 35 sample values overlaid with the estimated extreme value density $(1/\hat{\sigma})f_0((x - \hat{\mu})/\hat{\sigma})$, where $(\hat{\mu}, \hat{\sigma})$ are the maximum likelihood estimates. A formal assessment is given by the Anderson-Darling test statistic

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[\log \left\{ F \left(Y_{(i)}; \hat{\theta} \right) \right\} + \log \left\{ 1 - F \left(Y_{(n+1-i)}; \hat{\theta} \right) \right\} \right],$$

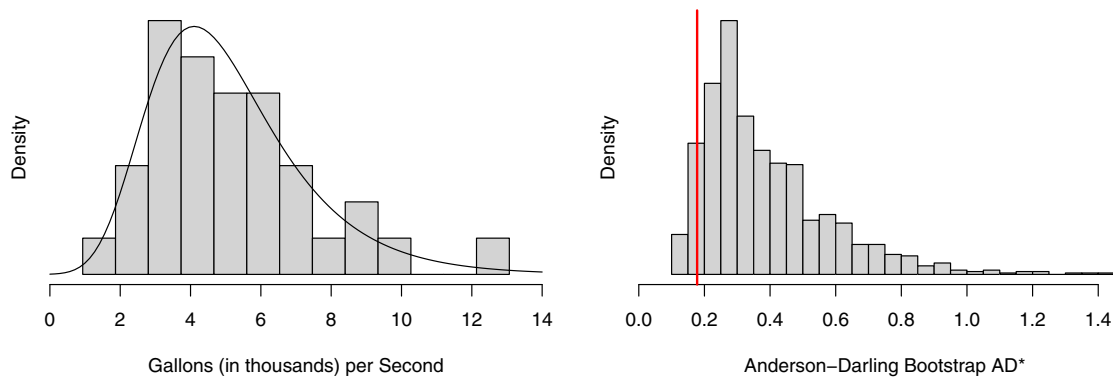


Fig. 2 Left panel: histogram of 35 maxima of river flow with estimated extreme value density. Right panel: histogram of $B = 1000$ Anderson-Darling bootstrap test statistics for the extreme value distribution. $T_0 = 0.178$ for original data is marked with vertical line.

where $Y_{(1)} \leq \dots \leq Y_{(n)}$ are the sample ordered values. Here the parametric bootstrap can be used to estimate the sampling distribution of AD when the sample is drawn from an extreme value distribution (the null hypothesis). It turns out that this sampling distribution is the same for any set of parameter values; so we just.

1. Generate B samples of size n from the extreme value distribution with $\mu = 0$ and $\sigma = 1$.
2. Find the maximum likelihood estimators $\hat{\mu}^*$ and $\hat{\sigma}^*$ for each bootstrap sample.
3. Compute AD above for each bootstrap sample and using $\hat{\mu}^*$ and $\hat{\sigma}^*$. Call these $AD_1^*, \dots, AD_B^*$.
4. The bootstrap estimated p -value is $\hat{p}_{\text{Boot}} = [\text{Number of } AD_i^* \geq AD_0]/B$, where AD_0 is the test statistic computed on the original sample.

Carrying this out for the max flow rates where $n = 35$, $AD_0 = 0.178$, and using $B = 1000$ leads to $\hat{p}_{\text{Boot}} = 0.93$ with standard error 0.01. The right panel of [Fig. 2](#) gives a histogram of the AD_i^* along with a vertical line at the observed value $AD_0 = 0.178$. We could also carry this out this approach to test goodness-of-fit for any location-scale family by merely replacing the base CDF F_0 by any other F_0 of interest. For example, letting F_0 be the standard normal leads to an estimated p -value against normality $\hat{p}_{\text{Boot}} = 0.20$. That seems a little unusual because the data are clearly skewed, but the AD statistic is an omnibus test and may not fully catch the skewness. A more focused test is to use the sample third moment skewness measure $m_3/m_2^{1.5}$, where $m_k = n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^k$. The skewness estimate here is 0.96, but the normal distribution population skewness is 0. Sampling again from the standard normal, the two-sided parametric bootstrap p -value is $\hat{p}_{\text{Boot}} = 0.017$ with standard error 0.001 using $B = 10,000$, certainly evidence against a normal distribution.

When does the nonparametric bootstrap “work”?

First, for the bootstrap to “work,” we mean that estimated bootstrap distribution of a sample statistic T has the same distribution in large samples as the large sample limit of the true sampling distribution of T . Basically, this happens when T has approximately a normal (or χ^2) distribution in large samples. Some typical theorems are given in [Shao and Tu \(1995, Ch. 3\)](#). But it is much more enlightening to look at examples where the nonparametric bootstrap does not work. A counterexample from [Bickel and Freedman \(1981\)](#) is for a sample drawn from a uniform $(0, \theta)$ distribution with density $f(y; \theta) = 1/\theta$ if $y \in (0, \theta)$, and 0 otherwise. The maximum likelihood estimator of θ is the sample maximum, say $Y_{(n)}$ where $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}$ are the ordered sample values. The pivotal quantity $T = n(\theta - Y_{(n)})/\theta$ has a limiting standard exponential distribution in large samples. However, the analog bootstrap quantity is $T^* = n(Y_{(n)} - Y_{(n)}^*)/Y_{(n)}$, where $Y_1^* \leq Y_{(2)}^* \leq \dots \leq Y_{(n)}^*$ is an ordered nonparametric resample (Note that $Y_{(n)}$ plays the role of θ in the bootstrap pseudo-population). It turns out that T^* does not converge in large samples, so the nonparametric bootstrap fails. However, the parametric bootstrap based on sampling from $f(y; \bar{Y}) = (1/\bar{Y})\exp(-y/\bar{Y})$ has the correct limiting distribution. An alternative nonparametric bootstrap that resamples with sample size $m < n$ and having special constraints on the growth rate of m with n leads to the correct convergence of T^* (see, e.g., [Swanepoel, 1986](#)). This “ m out of n ” nonparametric bootstrap can be used in other *nonregular* cases, but it is not as appealing because of the choice of m .

Bootstrap resources

Some of the early important papers in the statistics literature are [Efron \(1979, 1981\)](#), [Bickel and Freedman \(1981\)](#), [Efron \(1982, 1987\)](#), [Hall \(1988\)](#). Important theoretical books are [Hall \(1992\)](#), [Shao and Tu \(1995\)](#), and [Davison and Hinkley \(1997\)](#). For general audiences, the best single source is the book by [Efron and Tibshirani \(1993\)](#). Expository articles include [Hinkley \(1988\)](#), [Efron \(2003\)](#), [Boos \(2003\)](#), [Boos and Stefanski \(2010, 2013, Ch. 11\)](#), and [Hesterberg \(2015\)](#). The two main R packages for bootstrap calculations are `bootstrap` (v2019.6, [Tibshirani et al., 2019](#)) and `boot` (v1.3–26, [Canty and Ripley, 2021](#)); `MonteCarlo.se` ([Boos et al., 2019](#)) has simple R functions for getting bootstrap standard errors in Monte Carlo studies.

The jackknife

The jackknife method preceded the bootstrap by 30 years ([Quenouille, 1949, 1956](#); [Tukey, 1958](#)), but has a much narrower focus, basically bias estimation/correction and variance estimation. To start, consider a random sample $Y = (Y_1, \dots, Y_n)^T$ and an estimator $\hat{\theta}(Y)$ of a population parameter θ . Then let $\hat{\theta}_{[i]}$ denote the “leave-1-out” estimator obtaining by computing $\hat{\theta}$ with Y_i deleted from the sample. Denote the average of these “leave-1-out” estimators by $\bar{\theta}_1 = n^{-1} \sum_{i=1}^n \hat{\theta}_{[i]}$ and define the pseudo-values by

$$\hat{\theta}_{ps,i} = n\hat{\theta} - (n-1)\hat{\theta}_{[i]}. \quad (4)$$

The average of these pseudo-values is the bias-adjusted jackknife estimator

$$\hat{\theta}_J = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{ps,i}.$$

The jackknife variance estimator for $\hat{\theta}$ can be defined in terms of the sample variances of the leave-1-out $\hat{\theta}_{[i]}$ or of the pseudo-values Eq. (4) as follows,

$$\hat{V}_J = \frac{(n-1)^2}{n} \frac{1}{n-1} \sum_{i=1}^n (\hat{\theta}_{[i]} - \bar{\theta}_1)^2 = \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^n (\hat{\theta}_{ps,i} - \hat{\theta}_J)^2. \quad (5)$$

The latter form shows that we treat the pseudo-observations somewhat like a Monte Carlo random sample of independent $\hat{\theta}'$ s.

Examples

The simplest example is the sample mean $\hat{\theta} = \bar{Y}$. In this case, the leave-1-out estimators are $\hat{\theta}_{[i]} = (n\bar{Y} - Y_i)/(n-1)$, and the pseudo-values are $n\bar{Y} - (n\bar{Y} - Y_i) = Y_i$, the observations themselves. Thus $\hat{V}_J = s_{n-1}^2/n$. Recall that the parallel bootstrap variance estimator is $\hat{V}_{\text{Boot}} = s_n^2/n$, which uses the biased “n” version of the sample variance.

The simplest nonlinear estimator is the sample variance $\hat{\theta} = s_n^2 = n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$, which estimates the population variance $\sigma^2 = E(Y_1 - \mu)^2$. The leave-1-out estimators are a bit ugly, but the pseudo-values are

$$\hat{\theta}_{ps,i} = \left(\frac{n}{n-1}\right) (Y_i - \bar{Y})^2, \quad i = 1, \dots, n.$$

Taking the average of the pseudo-values, we obtain

$$\hat{\theta}_J = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{ps,i} = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2,$$

the usual unbiased version of the sample variance. This illustrates how $\hat{\theta}_J$ is less biased than $\hat{\theta}$ where in this case $E_F\{\hat{\theta}\} = \{(n-1)/n\}\sigma^2$. Note also that

$$\hat{\theta}_{ps,i} - \hat{\theta}_J = \left(\frac{n}{n-1}\right) \left\{ (Y_i - \bar{Y})^2 - s_n^2 \right\}.$$

Finally, using the $(n-1)$ version sample variance of $\hat{\theta}_{ps,i}$ divided by n , we get

$$\hat{V}_J = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\frac{n}{n-1}\right)^2 \left\{ (Y_i - \bar{Y})^2 - s_n^2 \right\}^2 = \left(\frac{n}{n-1}\right)^3 \frac{(m_4 - s_n^4)}{n},$$

where $m_4 = n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^4$ is the sample central fourth moment, which estimates the population fourth moment $\mu_4 = E(Y_1 - \mu)^4$. Thus, we see that in large samples $\hat{V}_J$ is estimating the asymptotic variance of s_n^2 given by $(\mu_4 - \sigma^4)/n$.

Recall the correlation estimates of [The bootstrap](#) section for the 13 class evaluations. The bootstrap gave $\hat{V}_{\text{Boot}} = 0.013$ for the Pearson correlation and $\hat{V}_{\text{Boot}} = 0.022$ for the Spearman correlation. The jackknife gives $\hat{V}_J = 0.013$ and $\hat{V}_J = 0.014$, respectively. It is curious that $\hat{V}_J = 0.014$ is so much smaller than $\hat{V}_{\text{Boot}} = 0.022$ for the Spearman correlation. [Efron and Stein \(1981\)](#) gave the interesting result that often $\hat{V}_J$ is biased upwards, that is, $E(\hat{V}_J) \geq \text{Var}(\hat{\theta})$, but there is no general result suggesting which of $\hat{V}_{\text{Boot}}$ or $\hat{V}_J$ will be larger.

Further comments

The jackknife can be extended to multiple independent samples, regression settings, and dependent samples. Generally it can be used for variance estimation in most places where the bootstrap can be used, but not for estimating the variance of the sample median. It is not as intuitive as the bootstrap to extend to more complicated situations, but it is one of the popular methods available for variance estimation in complex surveys ([Rust and Rao, 1996](#)). [Wu \(1986\)](#) gives extensions for regression. [Miller \(1974\)](#) provides an early review of the jackknife, but accessible introductions are in [Efron and Tibshirani \(1993\)](#) and [Boos and Stefanski \(2013, Ch. 10\)](#). Asymptotic theory is given in [Shao and Tu \(1995\)](#). R code may be found in the packages `bootstrap` (v2019.6, [Tibshirani et al., 2019](#)) and `MonteCarlo.se` ([Boos et al., 2019](#)).

Permutation tests

Permutation tests were invented by R. A. Fisher sometime in the early 1930s with examples found in his books (e.g., Fisher 1934, 1935). The exact first mention seems a mystery, but the main scope of application was quickly discovered by others, for example, Pitman (1937a,b, 1938) and Welch (1937), and the general theorem on validity given by Hoeffding (1952). See Berry et al. (2011) for a detailed history. Then the permutation method languished for many years except for use with rank methods where it provided one way to compute null distributions of rank-based test statistics like the Wilcoxon Signed Rank and Wilcoxon Rank Sum statistics. With increased computing power, the exact permutation distributions became available in certain situations, but more importantly, it was discovered that one only need to randomly sample from the full set of permutations. So computations are no longer an issue.

Before getting into details and examples, the key points are these:

1. The permutation method can only be used in 5 basic designs (plus stratified versions and other extensions). Contingency tables belong to these designs but they are presented in a different form, so we add them as a separate situation below.
 - (a) One sample, where the test statistic has a symmetric distribution around the null hypothesis value. Or two paired samples, where differences within pairs are used. Key examples: paired t -test, Wilcoxon Signed Rank Test.
 - (b) Two independent samples, where each sample comes from the same population under the null hypothesis. Key examples: two-sample pooled variance t -test, Wilcoxon Rank Sum Test.
 - (c) k independent samples, where each sample comes from the same population under the null hypothesis. Key examples: one-way analysis of variance F test for no mean differences, Kruskal-Wallis Rank Test.
 - (d) k treatments in a randomized complete block design. Key examples: two-way analysis of variance F test for no treatment mean differences, Friedman Rank Test.
 - (e) Random (X, Y) pairs. Key examples: Pearson correlation, simple linear regression t -test for slope = 0, Spearman and Kendall Rank Tests of Independence.
 - (f) m by k contingency tables. Key examples: McNemar's Test (2 by 2 tables, paired binary samples), Fisher's Exact Test (2 by 2 tables, independent or paired binary samples), Extension of Fisher's Exact Test (independence of two nominal-valued variables), Kruskal-Wallis Test (k independent samples of one ordinal-valued variable), Spearman and Rank Correlation Tests of Independence (two paired ordinal-valued variables).
2. There are no distributional *type* restrictions like normal or exponential. So the method will apply to any type of distribution as long as the basic random sampling restrictions of the above 5 designs hold.
3. There are no restrictions on the test statistic used. That is, one can dream up any test statistic for the above 5 designs, and the permutation method will make its use valid, that is, it will have correct Type I error properties. However, the basic permutation theory says nothing about power properties (Type II errors), that is, which test statistics will have good rejection rates under alternatives.

One sample (symmetry) or two paired samples

Fisher introduced the permutation approach for the matched-pairs problem in a discussion of Darwin's data on self-fertilized and cross-fertilized plants. There were 15 pairs of plants, and the differences

$$49, -67, 8, 16, 6, 23, 28, 41, 14, 29, 56, 24, 75, 60, -48$$

have mean $\bar{D} = 20.933$, $s = 37.744$, and $t = 2.148$ for testing $H_0 : \mu_D = 0$ versus $H_a : \mu_D \neq 0$, where μ_D is the population mean difference. The two-sided p -value is 0.0497 from the t table with 14 degrees of freedom. Alternatively, consider Fisher's permutation argument. There were 2^{15} possible random assignments of types of seeds to the 15 blocks of size 2. Thus, Fisher considered all 2^{15}

sums $\sum_{i=1}^{15} D_i$, where D_i is the i th difference and D_i is replaced by $-D_i$ in all ways possible (because switching X_i and Y_i changes $D_i = X_i - Y_i$ to $-D_i = Y_i - X_i$). For example, here are three such modifications:

$$\begin{aligned} & -D_1, D_2, \dots, D_{15}, \\ & -D_1, -D_2, \dots, D_{15}, \\ & -D_1, -D_2, \dots, D_{15}. \end{aligned}$$

Fisher found only $835 + 28 = 863$ which are greater than or equal to the observed sum = 314. The two-sided p -value is $(2) (863)/32,768 = 0.0527$ (by symmetry there are 863 sums ≤ -314). Note that $t = \sqrt{n}\bar{D}/s$ is permutationally equivalent to $\sum_{i=1}^{15} D_i$ because t is a monotonic function of $\sum_{i=1}^{15} D_i$ that depends on $\sum_{i=1}^{15} D_i^2$, which is constant over all 2^{15} permutations.

Let us consider the theory behind Fisher's approach. The population null model is that the differences $D_1, \dots, D_n$ are independent, each with a symmetric distribution about some θ_0 ; often $\theta_0 = 0$. The distributions do not need to be the same, merely symmetric about θ_0 . Thus $D_i - \theta_0$ and $-(D_i - \theta_0)$ have the same distribution under the null hypothesis that θ_0 is the center of the D_i distribution. So there are 2^{15} ways of manipulating the original sample by multiplying by pluses and minuses, all resulting

in samples with exactly the same joint distribution under the null hypothesis. We call these modifications of the sample *permutations* even though in this case they are not permutations.

Two independent samples

Let the samples be labeled $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$. For permutation tests, the null hypothesis is that the samples come from identical populations. This is easily justified when randomly assigning a batch of experimental units to two treatments, but it can be used with observational data if that “equal-under-the-null” assumption is justified. Then, the data modifications are to partition the full data $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ into all possible X and Y samples of size m and n ; for example, exchanging X_1 and Y_1 gives $(Y_1, X_2, \dots, X_m)$ and $(X_1, Y_2, \dots, Y_n)$. There are $(m+n)/(m!n!)$ such sets of two samples.

General notation and approach

In each situation, we use N for the total number of data points and M_N for the number of unique modifications that preserve the same joint distribution of the data under the null hypothesis. For the first example above (paired samples), $N = n$ and $M_N = 2^n$. For two independent samples, $N = m + n$ and $M_N = (m+n)/(m!n!)$. For k independent samples of size $n_1, n_2, \dots, n_k$, $N = n_1 + n_2 + \dots + n_k$, and $M_N = N!/(n_1!n_2!\dots n_k!)$. For the randomized complete block design with n blocks and k treatments, $N = n * k$, and $M_N = (k!)^n$. Finally, for n pairs (X, Y) and the test of independence, $N = 2n$ and $M_N = n!$.

For each situation and any statistic T , we calculate T for all the permutations, resulting in $T_1, \dots, T_{M_N}$. Let T_0 be the statistic calculated on the original data, and suppose that we reject the null hypothesis for large values of T . Then the permutation p -value is

$$p_{\text{perm}} = \frac{[\text{Number of } T_i \geq T_0]}{M_N}.$$

The permutation theory guarantees that if you do formal testing as “reject the null if $p_{\text{perm}} \leq \alpha$,” then $P(\text{Type I error}) \leq \alpha$. Note that the nonparametric bootstrap method cannot give this kind of guarantee. If we reject the null hypothesis for small values of T , then we merely change $\geq$ to $\leq$ in p_{perm} above. For two-sided tests, we often work with absolute values or squared versions like $|t|$ or t^2 .

Because the number of permutations M_N grows rapidly with N , we typically randomly sample (with replacement) B of the permutations, compute T for each sampled permutation resulting in $T_1^*, \dots, T_B^*$, and estimate p_{perm} with

$$\hat{p}_{\text{perm}} = \frac{[\text{Number of } T_i^* \geq T_0]}{B}.$$

The numerator of $\hat{p}_{\text{perm}}$ is binomial (B, p_{perm}) distributed so that $E(\hat{p}_{\text{perm}}) = p_{\text{perm}}$ and the estimated standard deviation of $\hat{p}_{\text{perm}}$ is $\sqrt{\hat{p}_{\text{perm}}(1 - \hat{p}_{\text{perm}})/B}$. I suggest choosing B large enough so that the estimated standard deviation is less than the first non-zero digit of $\hat{p}_{\text{perm}}$.

Further comments on permutation tests

Why are permutation tests not used more? The core of statistical applications use linear model theory where t and F approximations are usually quite close to permutation results. How close is close? The overriding point is that the sample-to-sample variability of any kind of p -value is quite large compared to the difference between good approximate p -values and the true permutation p -value p_{perm} or its estimate $\hat{p}_{\text{perm}}$ (see, e.g., [Boos and Stefanski, 2011](#)). So approximations are often good enough for inferential purposes. However, with categorical data, contingency tables, and rank tests, the chi-squared approximations are not as good, and permutation methods are very useful. One area where permutation methods are especially useful but not well-known is in meta-analysis (see [Follmann and Proschan, 1999](#)). Some books on permutation tests include [Edgington et al. \(2007\)](#), [Good \(1994\)](#), and subsequent editions), [Berry et al. \(2014\)](#), and [Boos and Stefanski \(2013, Ch. 12\)](#). The history of permutation tests with many references may be found in [Berry et al. \(2011\)](#), and [Berry et al. \(2018\)](#).

Summary

Resampling is a way to estimate the sampling distribution of random quantities computed from random samples. The main applications are for estimating the bias and variance of estimators, give confidence intervals for population parameters, and estimate the null distributions of hypothesis testing statistics.

1. Chronologically, permutation tests appeared first, in the 1930s. The scope of the permutation approach is very limited, applying to basically only five sampling designs and their extensions. However, for the those designs, the approach is *distribution-free* (does not depend on parametric family assumptions like normality) and is *statistic-inclusive* (any test statistic can be used). Moreover, the level- α control of Type I error holds for all samples sizes, small and large. In practical applications, it is simplest to estimate

the p -value by sampling from the set of permutations, and choosing B large enough to make the Monte Carlo standard error small.

2. The jackknife appeared next historically in development. For samples of size n , the i th resample of size $n - 1$ is created by dropping Y_i from the sample. The statistic of interest, say T , is then computed for each of the n resamples of size $n - 1$, and the jackknife variance estimate $\hat{V}_1$ is formed from the sample variance of those T^* . $\hat{V}_1$ is only theoretically justified in large samples and for regular estimators, but typically it performs well in small samples.
3. The bootstrap appeared last of the three main resampling methods, introduced by Efron in his famous 1979 paper. The nonparametric bootstrap is based on sampling from the sample as if it were a population. It can estimate virtually any aspect of regular estimators and test statistics. For variance estimation, its scope is wider than the jackknife because it can handle non-smooth estimators like the sample median. Otherwise, choosing between the nonparametric bootstrap and the jackknife is mostly personal preference—in large samples, such as found in Monte Carlo studies, they give almost identical results. However, the bootstrap can help with a myriad of other statistical inference needs such as confidence intervals and hypothesis tests. When compared with permutation tests on their home turf, the nonparametric bootstrap is not as good because all results are justified only in large samples and are limited to regular test statistics. However, there are many situations outside the home base of permutation tests where the bootstrap can be used. Typical recommendations for bootstrap resample sizes B are 100 to 1000 for variance estimation and 2000 or more for confidence intervals and percentile estimation. The “official” bootstrap estimate is the result of letting B approach infinity. If computing time is not a problem, choosing a large B can help others match results published in papers.

The parametric bootstrap is based on assuming a family of distributions like the normal (μ, σ^2) distribution, and drawing bootstrap resamples from the family with estimated parameters. It can handle nonregular statistics and also give exact test results as we saw with goodness-of-fit testing. However, the parametric bootstrap is less used in applications because we often want to avoid parametric assumptions. For example, we might choose a parametric family to get a maximum likelihood estimator but then use the nonparametric bootstrap to assess variability.

References

- Berry, K.J., Johnson, J.E., Mielke Jr., P.W., 2011. Permutation methods. *Comput. Stat.* 3, 527–542.
- Berry, K.J., Johnson, J.E., Mielke Jr., P.W., 2014. *A Chronicle of Permutation Statistical Methods*. Springer.
- Berry, K.J., Johnson, J.E., Mielke Jr., P.W., Johnson, L.A., 2018. Permutation methods. Part II. *Comput. Stat.* 10, 1–31.
- Bickel, P.J., Freedman, D.A., 1981. Some asymptotic theory for the bootstrap. *Ann. Stat.* 9, 1196–1217.
- Boos, D.D., Stefanski, L.A., 2010. Efron's bootstrap. *Significance* 7, 186–188.
- Boos, D.D., Stefanski, L.A., 2011. P-value precision and reproducibility. *Am. Stat.* 65, 213–221.
- Boos, D.D., Stefanski, L.A., 2013. *Essential Statistical Inference*. Springer.
- Boos, D., Matthew, K., Osborne, J., 2019. MonteCarlo.se: Monte Carlo Standard Errors. R Package Version 0.1.0. <https://CRAN.R-project.org/package=MonteCarlo.se>.
- Boos, D.D., 2003. Introduction to the bootstrap world. *Stat. Sci.* 18, 168–174.
- Canty, A., Ripley, B., 2021. Boot: Bootstrap R (S-Plus) Functions. R Package Version 1, pp. 3–26. <https://CRAN.R-project.org/package=boot>.
- Davison, A.C., Hinkley, D.V., 1997. *Bootstrap Methods and Their Application*. Cambridge University Press.
- Edgington, E.S., Edgington, E., Onghena, P., 2007. *Randomization Tests*, fourth ed. Taylor and Francis.
- Efron, B., Stein, C., 1981. The jackknife estimate of variance. *Ann. Stat.* 9, 586–596.
- Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman & Hall.
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26.
- Efron, B., 1981. Nonparametric standard errors and confidence intervals (with discussion). *Can. J. Stat.* 9, 139–172.
- Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. *SIAM (Society for Industrial and Applied Mathematics)*.
- Efron, B., 1987. Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–185.
- Efron, B., 2003. Second thoughts on the bootstrap. *Stat. Sci.* 18, 135–140.
- Fisher, R.A., 1934. *Statistical Methods for Research Workers*, fifth ed. Oliver & Boyd.
- Fisher, R.A., 1935. *The Design of Experiments*, eighth ed., 1966. Hafner Press.
- Follmann, D.A., Proschan, M.A., 1999. Valid inference in random effects meta-analysis. *Biometrics* 55, 732–737.
- Good, P.I., 1994. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer-Verlag.
- Hall, P., 1988. Theoretical comparison of bootstrap confidence intervals. *Ann. Stat.* 16, 927–953.
- Hall, P., 1992. *The Bootstrap and Edgeworth Expansion*. Springer-Verlag Inc.
- Hesterberg, T.C., 2015. What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. *Am. Stat.* 69, 371–386.
- Hinkley, D.V., 1988. Bootstrap methods. *J. Roy. Stat. Soc. B* 50, 321–337.
- Hoeffding, W., 1952. The large-sample power of tests based on permutations of observations. *Ann. Math. Stat.* 23, 169–192.
- Miller, R.G., 1974. The jackknife—a review. *Biometrika* 61, 1–15.
- Pitman, E.J.G., 1937a. Significance tests which may be applied to samples from any populations. *J. Roy. Stat. Soc. Suppl.* 4, 119–130.
- Pitman, E.J.G., 1937b. Significance tests which may be applied to samples from any populations. II. The correlation coefficient test. *J. Roy. Stat. Soc. Suppl.* 4, 225–232.
- Pitman, E.J.G., 1938. Significance tests which may be applied to samples from any populations. III. The analysis of variance test. *Biometrika* 29, 322–335.
- Quenouille, M.H., 1949. Approximate use of correlation in time series. *J. Roy. Stat. Soc. B* 11, 18–84.
- Quenouille, M.H., 1956. Notes on bias in estimation. *Biometrika* 43, 353–360.
- Rust, K.F., Rao, J.N.K., 1996. Variance estimation for complex surveys using replication techniques. *Stat. Methods Med. Res.* 5, 283–310.
- Shao, J., Tu, D., 1995. *The Jackknife and Bootstrap*. Springer-Verlag Inc.
- Swanepoel, J.W.H., 1986. A Note on proving that the (modified) bootstrap works. *Commun. Stat. Theor. Methods* 15, 3193–3203.
- Tibshirani, R., Leisch, F., Kostyshak, S., 2019. Bootstrap: Functions for the Book “An Introduction to the Bootstrap”. R package version 2019.6. <https://CRAN.R-project.org/package=bootstrap>.

Tukey, J., 1958. Bias and confidence in not quite large samples (abstract). *Ann. Math. Stat.* 29, 614.

Welch, B.L., 1937. On the Z-test in randomized blocks and Latin squares. *Biometrika* 29, 21–52.

Wu, C.F.J., 1986. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Stat.* 14, 1261–1295. This page intentionally left blank

Statistical inference: Bayesian approaches

Yang Liu and Tracy Sweet, Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	773
Bayesian inference	774
Posterior distribution	774
A running example	774
A decision-theoretic treatment	775
Decision rule, loss, and risk	775
Point estimation	775
Hypothesis testing	776
Set estimation	776
Prediction	777
Connection with frequentist inference	777
Prior specification	778
Uncertainty quantification	778
Objective Bayes	778
Computation	778
Analytical solution by conjugacy	778
Monte Carlo sampling from the posterior	779
Importance sampling	779
Markov chain Monte Carlo	780
Model fit assessment and model selection	780
Posterior predictive checking	780
Model comparison	781
Out-of-sample prediction	781
Marginal likelihood	781
Concluding remarks	782
References	782

Overview

In Bayesian statistics, probability is viewed as a universal gauge for uncertainty (O'Hagan, 2004, 2008). Not only is it able to describe randomness due to sampling—the source of variation dominating the frequentist definition of probability, it also broadly quantifies the lack of precise knowledge in any aspect of the data generating mechanism. For example, it is valid in Bayesian statistics to formalize one's personal belief on the values of model parameters as a probability distribution, whereas in frequentist statistics model parameters are treated as non-random quantities attached to data generating models and thus are not personal. As a consequence, Bayesian inference focuses on how the prior belief about parameters is updated after acquiring data and hence is “post-data” in nature. In contrast, frequentist inference concerns the stochastic behavior of statistical procedures over repeated sampling; such inference is “pre-data” as probability statements are made for the generative process that produces the observed data (Hinkley, 1983). Although Bayesian and frequentist statistics are based on fundamentally different philosophies, the discrepancy may be hardly noticeable in large samples (van der Vaart, 1998).

Prior specification is crucial for Bayesian inference in both subjective and objective modes. In more subjective uses of Bayesian statistics, analysts rely on probability distributions to make coherent decisions under uncertainty (DeGroot, 1970; Savage, 1954). Accordingly, a prior distribution should properly encode noisy *a priori* knowledge about the statistical model, which is then updated following the Bayes' rule when new data are observed. It is also possible to form priors when no *a priori* information is available (Berger, 2006; Kass and Wasserman, 1996). Nevertheless, there is no ubiquitous criterion of objectivity and hence no unique definition of objective priors.

In spite of the long-standing origin that can be traced back to the 18th century (Bayes, 1763), Bayesian methodology has not been popularized in education research and practices until recent decades due to computational intensity. Only in very few lucky cases, e.g., by conjugacy of well-known families of distributions, can one have closed-form expressions for the posterior density and the summary quantities (e.g., mean and variance) thereof. More frequently, Monte Carlo sampling techniques, such as Markov chain Monte Carlo (MCMC; Metropolis et al., 1953) and sequential Monte Carlo (SMC; Del Moral et al., 2006) must be invoked to generate random samples from the posterior distribution, from which approximations to summary quantities are constructed.

Statistical inference made from models that are incompatible with the observed data can be misleading. In the Bayesian framework, model-data fit can be conveniently assessed by comparing the observed data with the posterior distribution of the predictive data (Rubin, 1984). When there are competing hypotheses on the joint model for data and parameters, analysts often need to identify the optimal candidate in a principled manner. Model comparison criteria in Bayesian statistics can be defined based on the accuracy of out-of-sample prediction, the marginal likelihood of the observed data implied by joint model, and so forth (Kass and Raftery, 1995; Spiegelhalter et al., 2002).

Bayesian inference

Posterior distribution

The primary task of Bayesian inference is to specify the joint distribution of data $\mathbf{y} \in \mathcal{Y}$ and parameters $\boldsymbol{\theta} \in \Theta$, where $\mathcal{Y}$ and Θ denote data and parameter spaces, respectively. Let $f(\boldsymbol{\theta}, \mathbf{y})$ be the density¹ of the joint distribution, which can be factorized by the definition of a conditional density:

$$f(\boldsymbol{\theta}, \mathbf{y}) = f(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta}) \quad (1)$$

In Eq. (1), $f(\boldsymbol{\theta})$ is the marginal density of $\boldsymbol{\theta}$ and is commonly referred to as the *prior* density, which reflects one's uncertainty about model parameters prior to observing any data. Subjective beliefs, expert opinions, and/or existing empirical results may all serve as sources for prior specification. The second component on the right-hand side of Eq. (1), $f(\mathbf{y}|\boldsymbol{\theta})$, is termed the *likelihood*, which indicates how the data values $\mathbf{y}$ are likely to be observed given the parameter values $\boldsymbol{\theta}$. The likelihood uniquely characterizes the data generating mechanism and plays the central role in both frequentist and Bayesian inference.

For making inference about parameters $\boldsymbol{\theta}$, one is often more interested in the density of the "inverse problem," i.e., $f(\boldsymbol{\theta}|\mathbf{y})$, which expresses one's updated or adjusted belief of parameters $\boldsymbol{\theta}$ after seeing the observed data $\mathbf{y}$. In modern terminology, $f(\boldsymbol{\theta}|\mathbf{y})$ is referred to as the *posterior* density. Using the relationship of joint and conditional densities in the other direction, we arrive at the celebrated Bayes' (1763) formula:

$$f(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\boldsymbol{\theta}, \mathbf{y})}{f(\mathbf{y})} = \frac{f(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta})}{f(\mathbf{y})} \quad (2)$$

In Eq. (2), the denominator $f(\mathbf{y})$ is often called the *marginal likelihood* of $\mathbf{y}$, which by definition can be further expressed as

$$f(\mathbf{y}) = \int_{\mathcal{Y}} f(\boldsymbol{\theta}, \mathbf{y}) d\mathbf{y} = \int_{\mathcal{Y}} f(\boldsymbol{\theta})f(\mathbf{y}|\boldsymbol{\theta}) d\mathbf{y} \quad (3)$$

It is remarked that $f(\mathbf{y}) \neq 0$ is implicitly assumed in the Bayes' formula Eq. (2): This is not a restrictive assumption because any reasonable specification of $f(\boldsymbol{\theta}, \mathbf{y})$ should guarantee that data values arbitrarily close to the observed instance $\mathbf{y}$ can be generated with a positive probability.

A running example

As an illustration, we consider a one-parameter probit regression example (Agresti, 2013, Section 7.1). Let $\mathbf{y} = (y_1, \dots, y_n)'$ be a collection of n independent response variables. Each $y_i \in \{0, 1\}$ is a Bernoulli trial with success probability

$$\pi_i(\boldsymbol{\theta}) = P\{y_i = 1|\boldsymbol{\theta}\} = \Phi(\boldsymbol{\theta} + x_i) \quad (4)$$

in which Φ denotes the standard normal distribution function, x_i represents the known covariate value for observation i , and $\boldsymbol{\theta}$ is the intercept parameter.² By independence, the likelihood can be expressed as

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^n \pi_i(\boldsymbol{\theta})^{y_i} [1 - \pi_i(\boldsymbol{\theta})]^{1-y_i} \quad (5)$$

Consider $n = 2$ and a standard normal prior $f(\boldsymbol{\theta})$.³ The data space $\mathcal{Y}$ comprises four response patterns: $(0, 0)'$, $(1, 0)'$, $(0, 1)'$, and $(1, 1)'$. As Fig. 1 depicts, the joint density $f(\boldsymbol{\theta}, \mathbf{y})$ is the product of the prior and likelihood. The area under each $f(\boldsymbol{\theta}, \mathbf{y})$ equals to the marginal likelihood $f(\mathbf{y})$: The four areas add up to one, which is the total area under the standard normal prior. The posterior is obtained by renormalizing the corresponding joint density so that the area under the curve is one.

¹The term "density" is slightly abused to refer to either the probability density function for a continuous random variable or the probability mass function for a discrete random variable. The letter f is reserved as a generic notation for density functions.

²For ease of illustration, we fix the slope for the covariate to 1. In most applications of probit analysis, however, both the intercept and slope are estimated.

³The covariate values are fixed at $x_1 = 0.5$ and $x_2 = -1$.

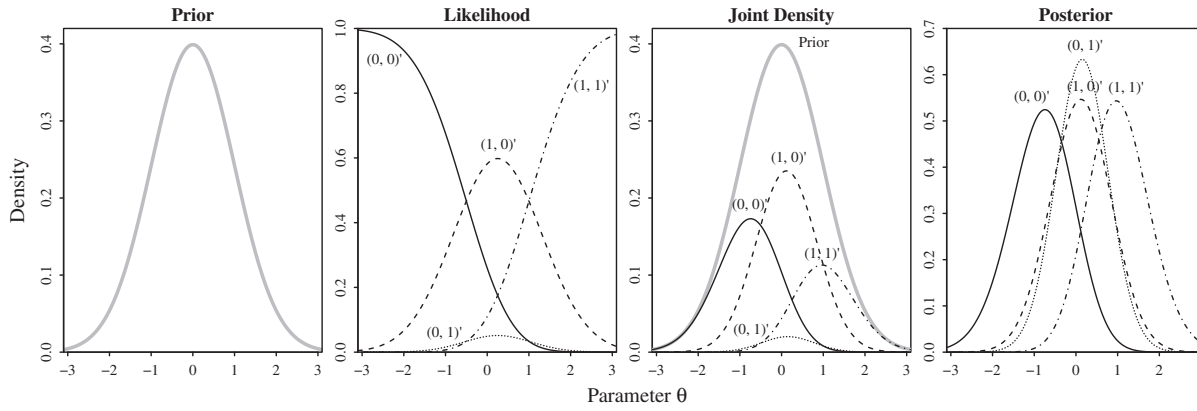


Fig. 1 Prior $f(\theta)$, likelihood $f(\mathbf{y}|\theta)$, joint density $f(\theta, \mathbf{y})$, and posterior $f(\theta|\mathbf{y})$ for the probit regression example ($n = 2$). There are in total four possible response patterns: $\mathbf{y} = (0, 0)'$, $(0, 1)'$, $(1, 0)'$, and $(1, 1)'$. The standard normal prior is highlighted by thick solid lines in gray color. Likelihoods, joint densities, and posteriors for different response patterns are shown in different line types.

A decision-theoretic treatment

Given the posterior of θ given $\mathbf{y}$, various kinds of inference can be made for the parameters θ , transformations of θ , or future data whose generative process depends on θ . Inferential procedures can be conveniently presented and evaluated under the decision theoretic framework (Berger, 1985), which we now elaborate.

Decision rule, loss, and risk

A statistical decision can be formally defined as a map d from the data space $\mathcal{Y}$ to the action space $\mathcal{A}$. The action space is determined by the type of statistical inference of interest: For example, $\mathcal{A} = \Theta$ if one aims to find a point estimator of θ . Given a pair of data $\mathbf{y} \in \mathcal{Y}$ and parameters $\theta \in \Theta$, every statistical decision $d(\mathbf{y}) \in \mathcal{A}$ carries a “cost” measured by a non-negative *loss function* $L(d(\mathbf{y}), \theta)$ —as the name suggests, a better decision comes with a smaller loss. Because data $\mathbf{y}$ and parameters θ are both stochastic in the Bayesian configuration, one may evaluate the overall quality of a decision rule d by the *Bayes risk*:

$$R(d) = \iint_{\Theta \times \mathcal{Y}} L(d(\mathbf{y}), \theta) f(\theta, \mathbf{y}) d\theta d\mathbf{y} \quad (6)$$

which is the expectation of the loss function under the joint density. We may further rearrange Eq. (6) as

$$R(d) = \int_{\mathcal{Y}} L(d(\mathbf{y}), \theta) f(\theta|\mathbf{y}) f(\mathbf{y}) d\theta d\mathbf{y} = \int_{\mathcal{Y}} \underbrace{\left\{ \int_{\Theta} L(d(\mathbf{y}), \theta) f(\theta|\mathbf{y}) d\theta \right\}}_{\rho(d, \mathbf{y})} f(\mathbf{y}) d\mathbf{y} \quad (7)$$

In the second line of Eq. (7), the term $\rho(d, \mathbf{y})$ is labeled the *posterior risk*, which amounts to the posterior expectation of the loss function. As a corollary of Eq. (7), a decision rule d that minimizes the posterior risk for every $\mathbf{y}$ also minimizes the Bayes risk.

Point estimation

Let $d(\mathbf{y})$ be a point estimator of θ , and $\mathcal{A} = \Theta$. We may define the loss function by $L(d(\mathbf{y}), \theta) = \|d(\mathbf{y}) - \theta\|$ with any norm $\|\cdot\|$ on Θ . For instance, the (squared) L^2 -norm corresponds to the squared error loss, and the L^1 -norm corresponds to the absolute error loss. A point estimator that minimizes the Bayes risk Eq. (6) is considered optimal. It can be shown that the posterior mean, also known as the expected *a posteriori* (EAP) estimator, minimizes the posterior risk Eq. (7) under the squared error loss; similarly, the posterior median is optimal under the absolute error loss. More generally, we might be interested in estimating transformed parameters defined by a map $\eta: \Theta \rightarrow H$, $\theta \mapsto \eta(\theta)$. Now the action space $\mathcal{A}$ becomes the transformed parameter space H ; the loss and risk functions can be defined accordingly with a suitable norm on H .

In our running example, suppose that we observed the response pattern $\mathbf{y} = (1, 0)'$. The posterior density of θ and the EAP estimator $\hat{\theta}(\mathbf{y}) = E(\theta|\mathbf{y}) = \int_{\Theta} \theta f(\theta|\mathbf{y}) d\theta = 0.12$ are graphed in the first panel of Fig. 2. We also consider estimating the expected total of y_1 and y_2 :

$$\eta(\theta) = E(y_1 + y_2|\theta) = \pi_1(\theta) + \pi_2(\theta) \quad (8)$$

which is a strict monotonic transform of θ . The EAP estimator of $\eta(\theta)$ is $\hat{\eta}(\mathbf{y}) = E(\eta(\theta)|\mathbf{y}) = \int_{\Theta} \eta(\theta) f(\theta|\mathbf{y}) d\theta = 0.93$. All the integrals involved in the above calculations do not have closed-form solutions and thus are approximated using numerical quadrature.

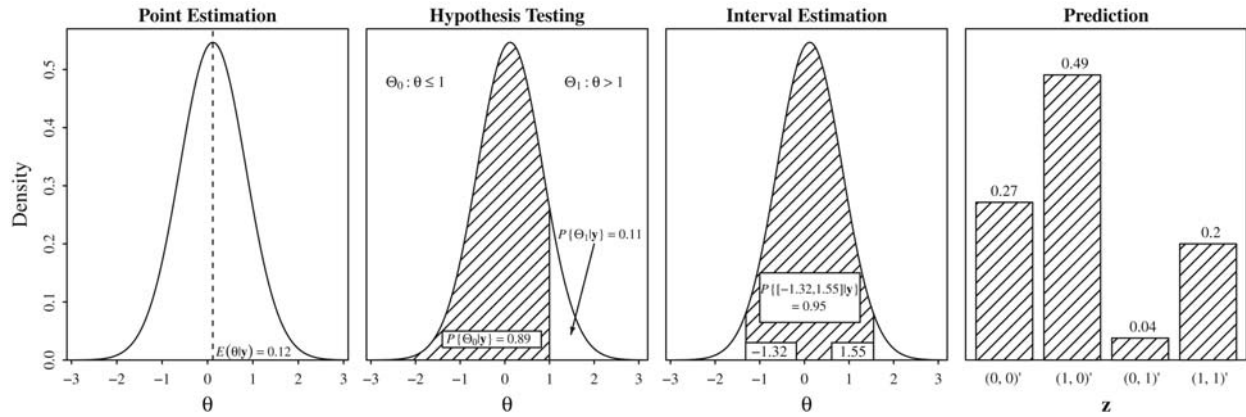


Fig. 2 Inference made based on the observed response pattern $\mathbf{y} = (1, 0)'$. First panel: Expected *a posteriori* estimator of θ . Second panel: Hypothesis testing of $\Theta_0 : \theta \leq 1$ versus $\Theta_1 : \theta > 1$. Third panel: 95% highest posterior density region. Fourth panel: Predictive probabilities for the four future response patterns.

Hypothesis testing

Let the null hypothesis Θ_0 and the alternative hypothesis Θ_1 be disjoint subsets of the parameter space Θ . A hypothesis test is conducted to decide whether Θ_0 or Θ_1 is favored given one's prior knowledge as well as the observed data. The binary action space for the testing problem is denoted $\mathcal{A} = \{\Theta_0, \Theta_1\}$. Accordingly, a decision rule d maps data $\mathbf{y}$ to one of the two possible actions, and the loss function for evaluating $d(\mathbf{y})$ may take the following form:

$$L(d(\mathbf{y}), \theta) = \begin{cases} c_1, & \theta \in \Theta_0 \text{ and } d(\mathbf{y}) = \Theta_1 \\ c_2, & \theta \in \Theta_1 \text{ and } d(\mathbf{y}) = \Theta_0 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

In Eq. (9), c_1 and c_2 are positive constants assigned to penalize Type I and Type II error, respectively. If we set $c_1 = c_2$, which equally weighs the two types of error, then the optimal decision that minimizes the posterior risk boils down to selecting the subspace with a larger posterior probability.

In the probit regression example, suppose that we are interested in testing $\Theta_0 = \{\theta : \theta \leq 1\}$ against $\Theta_1 = \{\theta : \theta > 1\}$ given the observed response pattern $\mathbf{y} = (1, 0)'$; see the second panel of Fig. 2. The posterior probabilities for the two subspaces are $P\{\Theta_0|\mathbf{y}\} = 0.89$ and $P\{\Theta_1|\mathbf{y}\} = 0.11$, which clearly favors Θ_0 .

Set estimation

A set estimate of θ , denoted $d(\mathbf{y})$ using the decision-theoretic notation, is a data-dependent subset of the parameter space Θ . The corresponding action space contains all subsets of Θ , i.e., $\mathcal{A} = 2^\Theta$. Averaged over the joint distribution of data and parameters, a good set estimate is expected to strike a balance between a high chance to include θ and a small size (Winkler, 1972). It then motivates the following loss function:

$$L(d(\mathbf{y}), \theta) = 1\{\theta \notin d(\mathbf{y})\} + w|d(\mathbf{y})| \quad (10)$$

in which $1\{B\} = 1$ if event B holds and 0 otherwise, $|A|$ denotes the size of set $A \subset \Theta$, and $w > 0$ is the weight controlling the relative importance of coverage and size. Eq. (10) gives rise to the following posterior risk:

$$\rho(d, \mathbf{y}) = P\{\theta \notin d(\mathbf{y})|\mathbf{y}\} + w|d(\mathbf{y})| \quad (11)$$

One may optimize Eq. (11), which implies a compensatory relation between the coverage and size of the resulting set estimate, upon observing data $\mathbf{y}$. As an alternative, one may fix a required level of posterior probability covered by $d(\mathbf{y})$ while optimize only for its size. This latter approach is more commonly seen in practice and leads to the so-called *high posterior density* (HPD) regions (Box and Tiao, 1973, p. 122).

In the probit regression example, the 95% HPD interval estimate⁴ of θ is defined by

$$d(\mathbf{y}) = \{\theta : f(\theta|\mathbf{y}) \geq \lambda\} \quad (12)$$

in which λ satisfies $P\{f(\theta|\mathbf{y}) \geq \lambda|\mathbf{y}\} = 0.95$; both λ and the end points of the HPD interval need to be determined numerically. For the response pattern $\mathbf{y} = (1, 0)'$, as shown in the third panel of Fig. 2, the 95% HPD interval is $[-1.32, 1.55]$, and the associated $\lambda = 0.08$.

⁴Here, the HPD region is an interval because the posterior density is unimodal.

Prediction

Sometimes, it is desired to make predictive inference about data that have not yet been observed. Let $\mathbf{z} \in \mathcal{Z}$ be the future data: $\mathbf{z}$ and $\mathbf{y}$ need not be independent or identically distributed in general, but we assume that the conditional likelihood of $\mathbf{z}$ given $\mathbf{y}$, denoted $f(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})$, is governed by the same set of parameters $\boldsymbol{\theta}$. The *posterior predictive density* for future data given the observed data is

$$f(\mathbf{z}|\mathbf{y}) = \int_{\Theta} f(\boldsymbol{\theta}|\mathbf{y})f(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})d\boldsymbol{\theta} \quad (13)$$

To recast Bayesian prediction in the decision-theoretic framework, we just need to treat $(\boldsymbol{\theta}', \mathbf{z}')'$ as “augmented parameters” defined on $\Theta \times \mathcal{Z}$ and apply suitable loss functions on the $\mathbf{z}$ -marginal.

In the probit regression example, suppose that the future data consist of two additional responses with the same covariate values x_1 and x_2 —that is, $\mathbf{z}$ is an independent and identically distributed (i.i.d.) copy of $\mathbf{y}$. In this case, the predictive probability Eq. (13) reduces to $f(\mathbf{z}|\mathbf{y}) = \int_{\Theta} f(\boldsymbol{\theta}|\mathbf{y})f(\mathbf{z}|\boldsymbol{\theta})d\boldsymbol{\theta}$. Given the responses $\mathbf{y} = (1, 0)'$ to the two original items, the predictive probabilities for the four possible future patterns were computed and plotted in the last panel of Fig. 2.

Connection with frequentist inference

Notwithstanding the disparity in their underlying philosophies, Bayesian and frequentist approaches often yield comparable results in many applications, especially when the likelihood is formed based on a large number of independent samples. Let the data $\mathbf{y}$ be composed of n independent observations $(y_1, \dots, y_n)'$ so that the likelihood is subject to the factorization $f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^n f_i(y_i|\boldsymbol{\theta})$, where

$f_i(y_i|\boldsymbol{\theta})$ denotes the per-observation likelihood. Suppose for the moment that the sample data $\mathbf{y}$ are generated under $f(\mathbf{y}|\boldsymbol{\theta}_0)$ with true parameters $\boldsymbol{\theta}_0$, which is a standard frequentist setup. It can be established under virtually no additional assumption that the posterior distribution $f(\boldsymbol{\theta}|\mathbf{y})$ converges weakly to a point mass concentrated at $\boldsymbol{\theta}_0$ as the sample size n increases—a result known as the Doob’s Consistency Theorem (Doob, 1949). Under stronger regularity conditions, the Bernstein-von Mises Theorem (Le Cam and Yang, 1990; van der Vaart, 1998) states that the posterior distribution of $\boldsymbol{\theta}$ approaches the (random) normal measure $\mathcal{N}(\hat{\boldsymbol{\theta}}_n(\mathbf{y}), \mathcal{I}_n(\boldsymbol{\theta}_0)^{-1})$ as $n \rightarrow \infty$, in which $\hat{\boldsymbol{\theta}}_n(\mathbf{y})$ denotes the maximum likelihood estimator (MLE) and $\mathcal{I}_n(\boldsymbol{\theta}_0)$ denotes the sample Fisher information matrix. As such, Bayesian point estimators, such as the posterior mean, behave essentially identically to the MLE in large samples, and Bayesian credible sets have asymptotically the correct frequentist coverage.

To illustrate the large-sample concordance of frequentist and Bayesian inference in our probit regression example, we fixed $\theta_0 = -1$ and simulated 40 independent responses variables.⁵ In Fig. 3, we contrast the normalized likelihood and posterior, as well as the MLE and the posterior mean (i.e., EAP), computed using the first two, first 20, and all 40 response variables. As n increases, the normalized likelihood and posterior becomes closer to one another, and both the MLE and EAP tend to $\theta_0 = -1$.

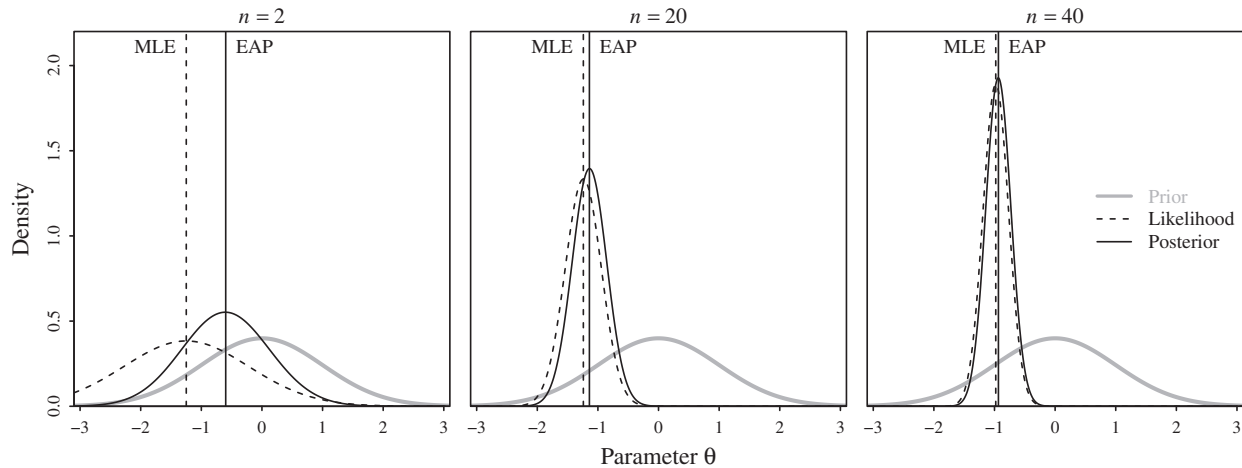


Fig. 3 Prior (gray solid), likelihood (black dashed), and posterior (black solid) for $n = 2, 20$, and 40 . Because the likelihood is smaller than the prior and posterior by many orders of magnitude when n is large, we normalize the likelihood so that the area under the curve is one. The maximum likelihood estimator (MLE) and the expected *a posteriori* (EAP) estimator are marked as the solid and dashed vertical lines, respectively.

⁵The covariates were randomly generated: $x_i \sim \text{Uniform}(0.5, 2)$, $i = 1, \dots, 40$. Once sampled, those values were treated as fixed when making inference about $\boldsymbol{\theta}$.

Prior specification

Uncertainty quantification

Howson and Urbach (2006) argued that Bayesian inference serves as a universal foundation for scientific inductions: In particular, it dictates how knowledge is accumulated or adapted in the emergence of new evidence. In line with the perspective of inductive reasoning, a prior distribution of model parameters should be specified to reflect the analyst's preknowledge and the associated degree of uncertainty. Though it is not invalid to select priors based solely on subjective beliefs, scientific inquiries often call for more principled prior elicitation protocols to reduce the impact of personal bias. For example, prior information may be elicited from content experts by the techniques of "elaboration" (O'Hagan et al., 2006). It may also be possible to find meta-analytical evidence or even historical data sets pertaining to the statistical model under investigation. Methods that turn historical information into priors, such as meta-analytic-predictive prior and power prior, have been extensively studied in the literature (Ibrahim and Chen, 2000; Schmidli et al., 2014).

Objective Bayes

When no *a priori* knowledge is present, one may still be able to conduct Bayesian inference using objective priors, which are derived based on various underpinnings of "objectivity" or "noninformativeness"; see Kass and Wasserman (1996) for a comprehensive review. One classic rule is the principle of insufficient reason, also known as the Bayes-Laplace postulate, which assigns "equal likelihood" to all parameter values to indicate absence of prior knowledge. In case the parameter space is unbounded so that a uniform prior is improper, a popular workaround amounts to the use of diffuse proper priors, such as normal distributions with large variances (Box and Tiao, 1973, p. 23). An obvious problem for the rule of insufficient reason is that the uniform distribution fails to be invariant under reparameterization: In the probit regression example, a uniform prior on expected total $\eta(\theta)$, a nonlinear transform of the intercept θ given by Eq. (8), induces a non-uniform prior on θ . To resolve this issue, one may alternatively treat reparameterization invariance as the defining property of non-informativeness. One prominent example of this kind is Jeffreys' (1946) prior: For one-parameter problems, Jeffreys' prior is proportional to the square-root of the Fisher information. A third possible definition of objective priors is to "maximize the discrepancy" between the prior and posterior distributions. For instance, Berger and Bernardo (1989) measured the discrepancy by the expected Kullback-Leibler divergence and defined a "reference prior" via an asymptotic argument. Yet another interesting rule to define objective priors is by matching frequentist probabilities. Simply put, we would like to find prior distributions such that the resulting credible intervals attain almost exact frequentist coverage (Datta and Mukerjee, 2004; Welch and Peers, 1963).

Computation

Analytical solution by conjugacy

As discussed earlier, Bayesian decision making entails computing or approximating descriptive quantities (e.g., moments, quantiles, etc.) of the posterior distribution. For certain combinations of likelihood functions $f(\mathbf{y}|\theta)$ and prior densities $f(\theta)$, the posterior and prior distributions belong to the same family, leading to the phenomenon of *conjugacy* (Raiffa and Schlaifer, 1961). Such priors are referred to as *conjugate priors* with respect to the parameters θ for the likelihood $f(\mathbf{y}|\theta)$.

A standard textbook example of conjugacy concerns estimating the mean of a normal distribution with a known variance (Gelman et al., 2013, Section 2.5). Let $y_1^*, \dots, y_n^*$ be i.i.d. $\mathcal{N}(\theta, \sigma^2)$ data, where $\sigma^2 > 0$ is known, and $\mathcal{N}(\theta_0, \sigma_0^2)$ be the prior for the mean parameter θ . It can be shown that the resulting posterior is again normal with mean $(n\bar{y}^* \sigma_0^2 + \theta_0 \sigma^2) / (n\sigma_0^2 + \sigma^2)$ and variance $\sigma_0^2 \sigma^2 / (n\sigma_0^2 + \sigma^2)$, in which $\bar{y}^* = \sum_{i=1}^n y_i^* / n$. Inference for the mean parameter can therefore be made with ease since descriptive quantities of the normal family are often readily computable.

In certain scenarios, conjugacy, albeit not directly applicable to the likelihood of the observed data, may simplify posterior-based computations (e.g., Monte Carlo sampling) via data augmentation (Tanner and Wong, 1987). In our probit model example, we introduce i.i.d. auxiliary variables $y_i^* \sim \mathcal{N}(\theta, 1)$, $i = 1, \dots, n$, and re-express the success probability Eq. (4) as a function of both θ and y_i^* (Albert and Chib, 1993):

$$\pi_i(\theta, y_i^*) = P\{y_i^* \geq -x_i | \theta\} \quad (14)$$

The augmented Bernoulli model Eq. (14) has parameters $(\theta, \mathbf{y}^*)'$ where $\mathbf{y}^* = (y_1^*, \dots, y_n^*)'$. While the joint posterior distribution $p(\theta, \mathbf{y}^* | \mathbf{y})$ is still difficult to characterize, the conditionals $p(\theta | \mathbf{y}^*, \mathbf{y}) = p(\theta | \mathbf{y}^*)$ and $p(\mathbf{y}^* | \theta, \mathbf{y}) = \prod_{i=1}^n p(y_i^* | \theta, y_i)$ turn out to be tractable: $\theta | \mathbf{y}^* \sim \mathcal{N}(n\bar{y}^* / (n+1), 1 / (n+1))$ by the normal conjugacy result; meanwhile, Eq. (15) suggests that y_i^* follows $\mathcal{N}(\theta, 1)$ truncated to $[-x_i, \infty)$ if $y_i = 1$ and to $(-\infty, -x_i]$ if $y_i = 0$. As both conditional distributions are easy to sample from, we can generate random samples from the augmented posterior (approximately) via Gibbs sampling (Geman and Geman, 1984).

Despite the convenience that conjugate priors may bring, they can only be found for few combinations of likelihood functions and prior distributions. Forcing a specific likelihood that contains unrealistic assumptions about the data generating mechanism

ultimately results in poor model-data fit. Meanwhile, restricting to certain prior distributional families also adversely affects the flexibility in describing prior beliefs.

Monte Carlo sampling from the posterior

When it is impossible to derive closed-form expressions for posterior descriptive quantities, analysts may still count on Monte Carlo methods to make Bayesian inference. A Monte Carlo procedure generates samples (often approximately) from the target posterior distribution and estimates the desired quantities by appropriate sample statistics. For ease of exposition, suppose that the quantity of interest is the posterior expectation of the transformed parameters $\eta(\theta)$ for the rest of the current section.⁶ Usually, we are only able to analytically express the posterior density up to a normalizing constant: That is, only the joint density Eq. (1) is tractable, but the marginal likelihood Eq. (3) is not. A plethora of algorithms have been devised to tackle sampling problems of this sort. Due to limited space, we only briefly discuss two popular methods and apply them to the probit regression example.

Importance sampling

Importance sampling (IS) circumvents the difficult task of directly sampling from the target posterior. Instead, it only requires sampling from an *importance distribution* with a density $g(\theta)$ such that $f(\theta, \mathbf{y}) > 0$ implies $g(\theta) > 0$: Typically, we choose $g(\theta)$ that is easy to sample from (e.g., a normal or t distribution). It is not difficult to see that

$$E(\eta(\theta)|\mathbf{y}) = \frac{\int \eta(\theta) f(\theta, \mathbf{y}) d\theta}{\int f(\theta, \mathbf{y}) d\theta} = \frac{\int \eta(\theta) [f(\theta, \mathbf{y})/g(\theta)] g(\theta) d\theta}{\int [f(\theta, \mathbf{y})/g(\theta)] g(\theta) d\theta} = \frac{E_g[\eta(\theta) w(\theta)]}{E_g w(\theta)} \quad (15)$$

in which $w(\theta) = f(\theta, \mathbf{y})/g(\theta)$ is the *importance weight*, and E_g denotes the expectation with respect to the importance distribution. Eq. (15) indicates that it suffices to find Monte Carlo estimates of the two expectations with respect to g on the right-hand side, which can be much simpler than directly estimating the left-hand side expectation as sampling from g is assumed to be a trivial task. Note that IS remains valid even when g is not normalized, because the unknown normalizing constant appears in both the numerator and denominator and thus is canceled out on the right-hand side of Eq. (15).

We now apply IS to the probit regression example with the observed responses $\mathbf{y} = (1, 1)'$ and the item parameters listed in Footnote 1. We set the importance distribution to Cauchy with a location parameter 0.97, which is the mode of $f(\theta, \mathbf{y})$. With 10,000 Monte Carlo samples from the importance distribution, we obtained the IS estimate of $E(\theta|\mathbf{y})$ as 1.03—which coincides with the estimate based on numerical quadrature up to the second decimal place. The associated IS estimate of the target posterior density is graphed in the left panel of Fig. 4, along with the target and importance densities.

The efficiency of importance sampling hinges upon finding an importance distribution that closely approximates the target density, which becomes increasingly difficult as the dimensionality increases. One successful attempt to scale up the IS scheme to high-dimensional sampling problems is SMC (Chopin and Papaspiliopoulos, 2020; Del Moral et al., 2006). As the name

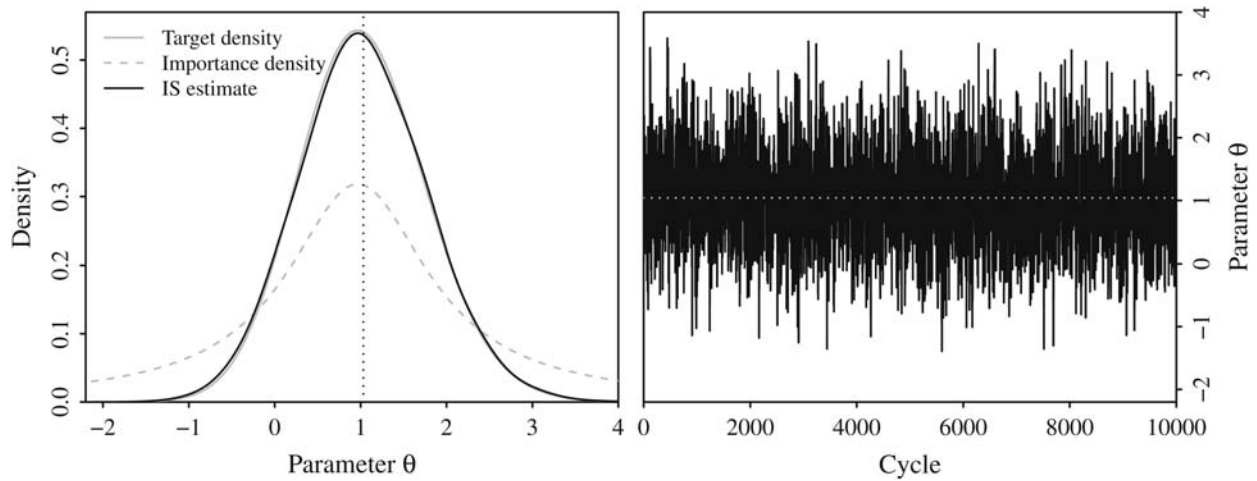


Fig. 4 Sampling from the posterior of the probit regression intercept θ given the responses $\mathbf{y} = (1, 1)'$. Left: The target, importance, and empirically estimated densities are shown in different combinations of line types and grayscales. The vertical dotted line marks the importance sampling (IS) estimate of the posterior mean. Right: A sample path produced by the random-walk Metropolis algorithm. The estimated posterior mean is highlighted by the horizontal dotted line.

⁶It is implicitly assumed that the expectation exists.

suggests, SMC constructs an importance density in a sequential manner—starting with a density that is easy to sample from and adjusting iteratively toward closer approximations to the target density. SMC has been implemented in generic software packages such as SMCTC (Johansen, 2009) and LibBi (Murray, 2013).

Markov chain Monte Carlo

MCMC sampling relies on the simulation of an ergodic Markov chain that has the target posterior as the stationary distribution (Robert and Casella, 2005, Definition 7.1). A Markov chain is characterized by a transition density $p(\theta_{t+1}|\theta_t)$, $t = 0, 1, \dots$, which designates the conditional distribution of the state at time $t + 1$ given the state at time t . We say that the target posterior $f(\theta|y)$ is the *stationary distribution* of the Markov chain if

$$f(\theta_{t+1}|y) = \int p(\theta_{t+1}|\theta_t)f(\theta_t|y)d\theta_t \quad (16)$$

Intuitively, Eq. (16) implies that if at a certain time t the marginal distribution of θ_t is exactly $f(\theta_t|y)$ then it remains unchanged forever afterward. If such a Markov chain further visits every event that has a positive posterior probability infinitely often regardless of the starting state (i.e., Harris recurrence; Meyn and Tweedie, 2009, p. 199), then it can be shown that the target expectation under the posterior is the limit of the running average along the sample path of the Markov chain: As $T \rightarrow \infty$,

$$\frac{1}{T} \sum_{t=0}^T \eta(\theta_t) \rightarrow E(\eta(\theta)|y) \quad (17)$$

almost surely regardless of the initial distribution of θ_0 . The above result is known as the Ergodic Theorem (Meyn and Tweedie, 2009, Theorem 17.1.7) and lays the groundwork for MCMC-based estimation of posterior expectations. So long as we can devise a Markov chain that satisfies the required technical conditions and let it run for sufficiently long time, the running average along the sample path is a valid estimate of the desired expectation.

One machinery to build ergodic Markov chains with desired stationary distributions is the Metropolis-Hastings algorithm (MH; Hastings, 1970; Metropolis et al., 1953). At time t , let $q(\theta_{t+1}|\theta_t)$ be the density of a *proposal distribution* such that $q(\theta_{t+1}|\theta_t) > 0$ for every θ_t and θ_{t+1} in the parameter space. An MH update is composed of the following steps: 1) generate a new state θ_{t+1} from the proposal, 2) compute the *acceptance ratio*

$$\alpha(\theta_t, \theta_{t+1}) = \min \left\{ \frac{f(\theta_{t+1}, y)q(\theta_t|\theta_{t+1})}{f(\theta_t, y)q(\theta_{t+1}|\theta_t)}, 1 \right\} \quad (18)$$

and 3) generate $u \sim \text{Uniform}[0, 1]$; retain θ_{t+1} if $u \leq \alpha(\theta_t, \theta_{t+1})$ and set $\theta_{t+1} = \theta_t$ otherwise.

To sample from the posterior given $y = (1, 1)'$ in the probit regression example, we considered a symmetric random-walk proposal $\theta_{t+1}|\theta_t \sim \mathcal{N}(\theta_t, 2^2)$. Using symmetric proposals further simplifies the calculation of Eq. (18) as the proposal densities are canceled out from the ratio. The trace plot of θ_t for $t = 1, \dots, 10000$ can be found in the right panel of Fig. 4. We estimated $E(\theta|y)$ by the average of the 10,000 generated θ_t 's, i.e., 1.04, which is very close to the previous IS estimate.

Varieties of efficient MCMC samplers have been developed based on the MH rule. For example, a Gibbs sampler (Geman and Geman, 1984) updates each coordinate of θ by sampling from its conditional distribution given the most recent values of the others. A Hamilton/hybrid Monte Carlo algorithm (HMC; Neal, 2011), as another example, generates proposals by simulating Hamiltonian dynamics on an augmented parameter space. Popular software implementations of the Gibbs and HMC samplers include BUGS (Spiegelhalter et al., 1996), JAGS (Plummer, 2015), and Stan (Stan Development Team, 2019), to name a few.

Model fit assessment and model selection

Posterior predictive checking

Posterior predictive checking (PPC) exploits the simple rationale that the observed data should be “representative” of the predictive data were the latter generated from a well-fitting model. Let $z \in \mathcal{Z}$ denote the predictive data that is an i.i.d. copy of the observed data y given model parameters θ . Recall that the joint posterior density of z and θ can be written as

$$f(z, \theta|y) = f(\theta|y)f(z|\theta) \quad (19)$$

which is a special case for the integrand of Eq. (13) under the i.i.d. assumption. Because the dimensionality of data can be high, it is often more handy to conduct PPC using a summary statistic $T(y, \theta)$, which is allowed to be dependent on both the data and parameters. Define the posterior predictive p -value (PPP-value) as

$$p_T(y) = \int \int_{T(z, \theta) \geq T(y, \theta)} f(z, \theta|y) dz d\theta \quad (20)$$

which amounts to the upper tail probability under the joint posterior of z and θ given y . An excessively low or high PPP-value indicates that the observed data is unlikely to be predicted based on the fitted posterior and hence bears witness to poor

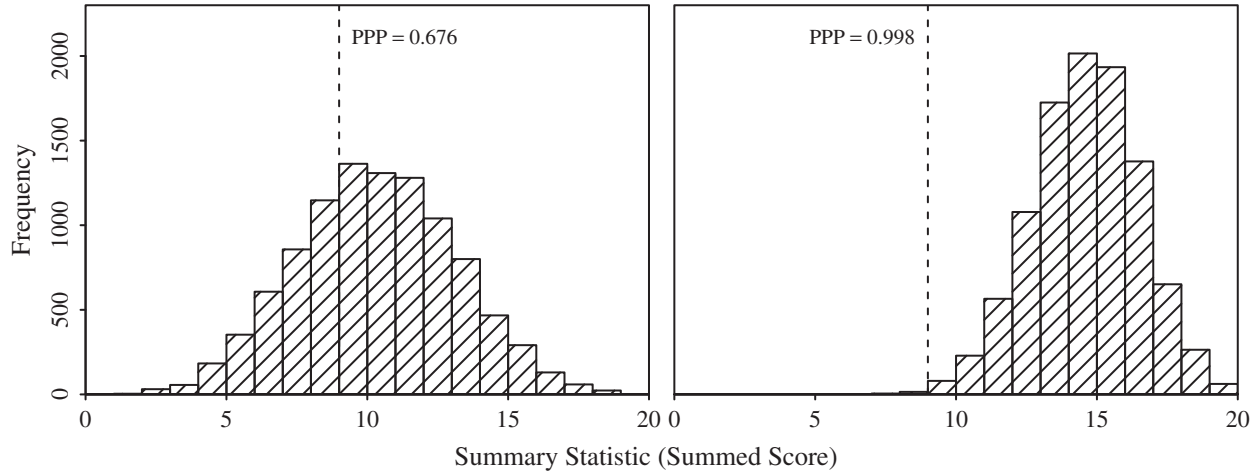


Fig. 5 Posterior predictive distribution of test statistics and the posterior predictive p -value (PPP). The summary statistic is the total score of 20 responses; the observed total score, i.e., 9, is marked by the vertical dashed lines. The histograms were produced based on 10,000 sets of simulated predictive data. Left: Results based on the probit model with a $\mathcal{N}(0, 1)$ prior. Right: Results based on the assumption that $\theta = -0.5$.

model-data fit. A PPP-value should not be confused with a frequentist p -value because its sampling distribution is in general not uniform (Bayarri and Berger, 2000; Robins et al., 2000).

Consider the probit regression example⁷ with $n = 20$ and the observed response vector $(1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0)'$. We used the total score $T(\mathbf{y}, \theta) = \sum_{i=1}^n y_i$ as the summary statistic, which equals to 9 when evaluated at the observed responses. Under the 2 PL model with a $\mathcal{N}(0, 1)$ prior, we estimated the PPP-value by simulating predictive responses conditional on 10,000 MCMC samples. A histogram of the predicted total scores is displayed in the left panel of Fig. 5; the observed total score is marked by the vertical dashed line, associated with a PPP-value of 0.676. To demonstrate how PPC is responsive to lack of fit, we also consider a poorly behaving model in which θ is fixed at -0.5 (or equivalently, specifying a point-mass prior concentrated at -0.5). The resulting PPP-value is 0.998: It can be seen from the right panel of Fig. 5 that the observed total score lies on the far lower tail of the posterior predictive distribution.

Model comparison

Out-of-sample prediction

A common strategy for model selection is to favor models that better predict future data. The accuracy of out-of-sample prediction can be measured in numerous ways. For instance, one may prefer models with high expected log predictive density (ELPD; Gelman et al., 2013), which is defined as follows:

$$\text{ELPD}(\mathbf{y}) = E_{\mathbf{z}} \log f(\mathbf{z}|\mathbf{y}) \quad (21)$$

in which $f(\mathbf{z}|\mathbf{y})$ is the posterior predictive density of future data $\mathbf{z}$ given by Eq. (13), and $E_{\mathbf{z}}$ denotes the generating distribution of $\mathbf{z}$. While the posterior predictive density can be approximated using Monte Carlo methods, the generating distribution of $\mathbf{z}$ is often not known to the analyst. If an observed $\mathbf{z}$ is available as a hold-out sample, one may estimate Eq. (21) by suitable sample statistics. There are also same-sample estimators of Eq. (21), such as the (minus) Deviance Information Criterion (DIC; Spiegelhalter et al., 2002), which do not need a hold-out sample to approximate $E_{\mathbf{z}}$.

With the same response vector we have used to demonstrate PPC, we obtained the posterior of θ conditional on the first ten responses and evaluated the corresponding log posterior predictive density of the remaining ten responses as an estimate of the ELPD. The estimated ELPD is -5.29 when a $\mathcal{N}(0, 1)$ prior is used, and -6.38 when a point-mass prior at -0.5 is used. It is then concluded that the first model has a better performance in out-of-sample prediction.

Marginal likelihood

Another useful criterion for model comparison is the marginal likelihood (Eq. 3), which reflects the propensity to reproduce the observed data under the hypothesized joint statistical model (Eq. 1). A model with a higher marginal likelihood is deemed more compatible with the observed data. It is then natural to gauge the relative fit of two competing models, whose marginal likelihood functions are respectively denoted $f(\mathbf{y}|M_1)$ and $f(\mathbf{y}|M_2)$, by the Bayes factor

⁷Item parameters were fixed at the values used to produce Fig. 3.

$$\text{BF}_{12}(\mathbf{y}) = \frac{f(\mathbf{y}|M_1)}{f(\mathbf{y}|M_2)} \quad (22)$$

$\text{BF}_{12}(\mathbf{y}) > 1$ implies that M_1 fares better than M_2 and *vice versa*. In practice, the marginal likelihood is often intractable and needs to be approximated numerically (e.g., via bridge sampling; Meng and Wong, 1996). When the sample size is large and the prior's influence is negligible, the marginal log-likelihood can be approximated by the Bayesian information criterion (BIC; Schwarz, 1978).

In the 20-item probit regression example we have been using throughout the current section, let M_1 be the joint model with a $\mathcal{N}(0, 1)$ prior and M_2 be that with a point-mass prior at -0.5 . The marginal log-likelihood values for M_1 and M_2 are -12.92 and -14.06 , respectively, leading to a Bayes factor $\text{BF}_{12}(\mathbf{y}) = \exp(14.06 - 12.92) = 3.12$. So far, all the model-fit diagnostics suggest that the model with a standard normal prior fits the observed data better.

Concluding remarks

Active research on Bayesian inferential methods includes but is not limited to model development and estimation. Convenience of Bayesian inferential framework and advances in computer software allow analysts to specify flexible models with less restrictive parametric assumptions (e.g., using process priors; Ghosh and Ramamoorthi, 2003). Recent development of Bayesian estimation techniques entails not only more general and/or efficient Monte Carlo sampling algorithms but also optimization algorithms that find good approximations to the posterior distribution and can scale up to large-size problems (e.g., variational inference; Blei et al., 2017).

References

- Agresti, A., 2013. *Categorical Data Analysis*, Wiley Series in Probability and Statistics. Wiley.
- Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. *J. Am. Stat. Assoc.* 88 (422), 669–679.
- Bayarri, M.J., Berger, J.O., 2000. *p*-values for composite null models. *J. Am. Stat. Assoc.* 95 (452), 1127–1142.
- Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances. *Philos. Trans. R. Soc.* 45, 293–315. London V61.53, 370–418. Reprinted with biographical note by G. A. Barnard in *Biometrika*.
- Berger, J.O., Bernardo, J.M., 1989. Estimating a product of means: Bayesian analysis with reference priors. *J. Am. Stat. Assoc.* 84 (405), 200–207.
- Berger, J.O., 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer.
- Berger, J.O., 2006. The case for objective Bayesian analysis. *Bayesian Anal.* 1 (3), 385–402.
- Blei, D.M., Kucukelbir, A., McAuliffe, J.D., 2017. Variational inference: a review for statisticians. *J. Am. Stat. Assoc.* 112 (518), 859–877.
- Box, G.E.P., Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*. Addison-Wesley.
- Chopin, N., Papaspiliopoulos, O., 2020. *An Introduction to Sequential Monte Carlo*, Springer Series in Statistics. Springer International Publishing.
- Datta, G.S., Mukerjee, R., 2004. *Probability Matching Priors: Higher Order Asymptotics*, Lecture Notes in Statistics. Springer, New York.
- DeGroot, M.H., 1970. *Optimal Statistical Decisions*. McGraw-Hill.
- Del Moral, P., Doucet, A., Jasra, A., 2006. Sequential Monte Carlo samplers. *J. R. Stat. Soc. B* 68 (3), 411–436.
- Doob, J.L., 1949. Application of the theory of martingales. In: *Le calcul des probabilités et ses applications*. Colloque International du CNRS, Paris, pp. 23–27.
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D., 2013. *Bayesian Data Analysis*, third ed. Taylor & Francis.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* (6), 721–741.
- Ghosh, J.K., Ramamoorthi, R.V., 2003. *Bayesian Nonparametrics*. Springer-Verlag.
- Hastings, W.K., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications.
- Hinkley, D.V., 1983. Can frequentist inferences be very wrong? A conditional “yes”. In: Box, G.E.P., Leonard, T., Wu, C.-F. (Eds.), *Scientific Inference, Data Analysis, and Robustness*. Academic Press, pp. 85–103.
- Howson, C., Urbach, P., 2006. *Scientific Reasoning: The Bayesian Approach*, third ed. Open Court Publishing Company.
- Ibrahim, J.G., Chen, M.-H., 2000. Power prior distributions for regression models. *Stat. Sci.* 15 (1), 46–60.
- Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 186. The Royal Society, pp. 453–461.
- Johansen, A.M., 2009. SMCTC: sequential Monte Carlo in C++. *J. Stat. Software* 30 (6), 1–41.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90 (430), 773–795.
- Kass, R.E., Wasserman, L., 1996. The selection of prior distributions by formal rules. *J. Am. Stat. Assoc.* 91 (435), 1343–1370.
- Le Cam, L., Yang, G., 1990. *Asymptotics in Statistics*. Springer-Verlag, New York.
- Meng, X.-L., Wong, W.H., 1996. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Stat. Sin.* 831–860.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21 (6), 1087–1092.
- Meyn, S., Tweedie, R., 2009. *Markov Chains and Stochastic Stability*, second ed. Cambridge University Press.
- Murray, L.M., 2013. Bayesian state-space modelling on high-performance hardware using LibBi. *J. Stat. Software* 67 (10). <https://arxiv.org/pdf/1306.3277.pdf>.
- Neal, R.M., 2011. MCMC using Hamiltonian dynamics. In: Brooks, S., Gelman, A., Jones, G., Meng, X. (Eds.), *Handbook of Markov Chain Monte Carlo*. CRC Press.
- O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley and Sons, Chichester.
- O'Hagan, A., 2004. Dicing with the unknown. *Significance* 1 (3), 132–133.
- O'Hagan, A., 2008. The Bayesian approach to statistics. In: Rudas, T. (Ed.), *Handbook of Probability: Theory and Applications*. Sage Publications, Inc., Thousand Oaks, CA, pp. 85–100.
- Plummer, M., 2015. JAGS Version 4.0. 0 User Manual.
- Raiffa, H., Schlaifer, R., 1961. *Applied Statistical Decision Theory*. Harvard Business School Publications, Division of Research, Graduate School of Business Administration, Harvard University.
- Robert, C., Casella, G., 2005. *Monte Carlo Statistical Methods*, Springer Texts in Statistics, second ed. Springer, New York.

- Robins, J.M., van der Vaart, A., Ventura, V., 2000. Asymptotic distribution of p -values in composite null models. *J. Am. Stat. Assoc.* 95 (452), 1143–1156.
- Rubin, D.B., 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Stat.* 1151–1172.
- Savage, L.J., 1954. *The Foundations of Statistics*. Wiley, New York.
- Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D., Neuenschwander, B., 2014. Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 70 (4), 1023–1032.
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Stat.* 6 (2), 461–464.
- Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., 1996. Bugs 0.5: Bayesian Inference Using Gibbs Sampling Manual (Version ii), pp. 1–59.
- Spiegelhalter, D., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* 64 (4), 583–639.
- Stan Development Team, 2019. *Stan Modeling Language User's Guide and Reference Manual*, version 2.26 ed.
- Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. *J. Am. Stat. Assoc.* 82 (398), 528–540.
- van der Vaart, A.W., 1998. *Asymptotic Statistics*. Cambridge University Press.
- Welch, B.L., Peers, H.W., 1963. On formulae for confidence points based on integrals of weighted likelihoods. *J. R. Stat. Soc. B* 25 (2), 318–329.
- Winkler, R.L., 1972. A decision-theoretic approach to interval estimation. *J. Am. Stat. Assoc.* 67 (337), 187–191.

Foundational methods: power analysis

Qi Zhang^a, Jessaca Spybrook^b, Benjamin Kelcey^c, and Nianbo Dong^d, ^a American Institutes for Research, Washington, VA, United States; ^b Western Michigan University, Kalamazoo, MI, United States; ^c University of Cincinnati, Cincinnati, OH, United States; and ^d University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

© 2023 Elsevier Ltd. All rights reserved.

Power for single level RTs	785
Model without covariates	785
Model with a covariate	786
Power calculation for individual level RCTs	786
Power for two-level CRTs	786
Model without covariates	786
Model with a cluster-level covariate	788
Power calculation for two-level CRTs	789
Future directions	789
References	791

Randomized trials (RTs) are the gold standard for education impact studies as a well-designed and implemented RT yields an unbiased estimate of the treatment effect (Hedges, 2017). In an RT, individuals or groups of individuals are randomly assigned to either receive the treatment that is being tested or to a comparison condition, which often utilizes the business as usual (BAU) model or program. The intervention being tested might be a reading improvement program, an online math curriculum, or a professional development program for teachers.

A RT is highly regarded in evaluating program impact because random assignment creates two or more equivalent groups based on observed or unobserved characteristics (Shadish et al., 2002). For example, the process of randomly assigning students to treatment and comparison groups would theoretically create two groups that have a similar composition of gender, race or ethnicity, and socioeconomic characteristics, or observed characteristics. In addition, random assignment also ensures that the two groups also have similar composition of unobserved characteristics, or characteristics that cannot be directly measured, such as attitude and self-efficacy (U.S. Department of Education, 2003). Creating equivalent treatment and comparison groups is an essential element in designing RTs because it partitions out the preexisting differences between groups to ensure that the observed treatment effect is the result of the intervention.

There are many components that contribute to the validity of the inferences in an RT, including the primary outcome measures and analytic plan (Shadish et al., 2002). Although all of these are important, in this article we focus specifically on the statistical power to detect an effect of a given magnitude. Statistical power is the probability of detecting an effect of a given magnitude when the true effect is non-zero. The probability of making a Type II error increases as the statistical power of the study decreases. Low statistical power may lead to invalid inferences about the presence or absence of a treatment effect.

A statistical power analysis is typically carried out to determine the probability of an RT to detect the main treatment effect (Hedges and Rhoads, 2010). The main treatment effect is the average effect of the intervention. It is often quantified in terms of an effect size, defined as the standardized difference in the mean outcome between the treatment and comparison conditions. The sensitivity of an RT to detect the main treatment effect often focuses on three overlapping topics: the standard error associated with the treatment effect, the power to detect the main effect, and the smallest effect size a study is designed to detect (Hedges et al., 2012). The latter two topics are typically the focus of *a-priori* power analysis, or a power analysis conducted during the planning stage of an RT.

Demonstrating that a study is sufficiently powered to detect an effect with a certain magnitude of interest is an imperative component of planning an RT. For example, suppose a team of researchers is developing a new online mathematics tutoring program for students and plans to test the impact of the program on mathematics achievement. They design an RT in which students are assigned to either the new online tutoring program or the current program. At the conclusion of the study, the team finds that there is no difference in mean outcomes for the two groups. A key question one might ask is whether there is no difference because the new program is not better than the current program or if the study simply is not sufficiently powered to detect a meaningful treatment effect? That is, was the sample size too small so that the estimated treatment effect was imprecise leading to a failure to reject the null hypothesis? This is not a question that one should ask at the end of a study. An *a-priori* power analysis is critical to ensure that the study is designed to enable strong evidence of whether an intervention is effective.

The purpose of this article is to present an overview of statistical power for RTs seeking to test the impact of an educational intervention. We lay the foundation by beginning with the statistical power to detect the treatment effect for a single-level RT, or a single level study in which students are the unit of random assignment. Then we discuss the statistical power to detect the treatment effect for a cluster randomized trial (CRT) or a study in which clusters, such as schools, are assigned to conditions and outcomes are

assessed at the student level. CRTs are common in impact studies in education given the nested structure of schools, students within schools, and the fact that interventions are often implemented at the school level. We end with a discussion of other types of statistical power analyses that are relevant for impact studies of educational interventions.

Power for single level RTs

Model without covariates

Recall the online mathematics tutoring program example. Students are randomly assigned to either the new program or the current program. Student mathematics achievement is the primary outcome. The level 1, or student-level model is:

$$Y_i = \beta_0 + \beta_1 T_i + e_i \quad e_i \sim N(0, \sigma^2) \quad (1)$$

Where Y_i is the mathematics score for individual $i = \{1, \dots, n\}$; β_0 is the overall mean mathematics score; β_1 is the mean difference between the treatment and comparison group or the main treatment effect; T_i is a treatment indicator with $-\frac{1}{2}$ for control and $\frac{1}{2}$ for treatment, and e_i is the residual error associated with students within a treatment condition with variance σ^2 .

The treatment effect is estimated by $\hat{\beta}_1 = \bar{Y}_E - \bar{Y}_C$, where $\bar{Y}_E$ is the mean for the treatment group and $\bar{Y}_C$ is the mean for the control group. The variance of the estimated treatment effect, which describes uncertainty of the estimate effect, is:

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{P(1-P)n} \quad (2)$$

Note the variance of the treatment effect is a function of the total sample size (n), the proportion of students assigned to the treatment condition (P), and the between-student variance within a treatment condition (σ^2). The associated hypothesis test is:

$$\begin{aligned} H_0 : \beta_1 &= 0 \\ H_1 : \beta_1 &\neq 0 \end{aligned} \quad (3)$$

This test is based on the F statistic, which is defined as:

$$F_{\text{statistic}} = \frac{(MS_{\text{treatment}})}{(MS_{\text{error}})} \quad (4)$$

which denotes the relationship between treatment variance and error variance.

The power for F-test is (Kirk, 1982):

$$\begin{aligned} \text{Power} &= \text{Prob}(\text{Reject } H_0 | H_0 \text{ is false}) \\ &= \text{Prob}(F > F_{\alpha;1, J-2}) \\ &= 1 - \text{Prob}(F < F_{\alpha;1, J-2}) \end{aligned} \quad (5)$$

where $F_{\alpha;1, J-2}$ is the critical value under the null hypothesis with 1 numerator degree of freedom and $n - 2$ denominator degrees of freedom. Under the null hypothesis, the F -statistic follow the central F -distribution. Under the alternative hypothesis, the F -statistic follows the noncentral F -distribution, represented by a non-centrality parameter λ . As Eq. (6) shows, the non-centrality parameter is a function of the true treatment effect and variance of the estimated treatment effect.

$$\lambda = \frac{\beta_1^2}{\text{Var}(\hat{\beta}_1)} = \frac{\beta_1^2}{\sigma^2/[P(1-P)n]} \quad (6)$$

It is common to standardize the parameters. The standardized treatment effect can be expressed as $\delta = \beta_1/\sigma$. Set $\sigma^2 = 1$ and substituting δ into Eq. (6), we get the standardized version as:

$$\lambda = \frac{\delta^2}{1/[P(1-P)n]} \quad (7)$$

The non-centrality parameter is related to the power of the test. That is, the power increases as the λ increases. Eq. (7) suggests λ is a function of the sample size (n), the standardized treatment effect (δ), and the proportion of individuals in the treatment condition (P). As n increases, the non-centrality parameter increases, thus the power also increases. Applying this concept to the example of designing a study to test an online mathematics tutoring program, the researcher could improve the power to detect the treatment effect by increasing the number of students in the sample. This approach is often adopted by researchers since sample size typically under the control of the researcher. The power also increases as δ increases. However, the magnitude of the treatment effect is a function of the intervention and not something that is typically under the control of the researcher.

Model with a covariate

Researchers often include covariates to increase the precision of the estimated treatment effect. A pre-treatment measure, such as the students' mathematics achievement prior to receiving the intervention, is often incorporated in the model. In this case, the level-1 or the student level model is an extension of Eq. (1):

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + e_i \quad e_i \sim N(0, \sigma_{|x}^2) \quad (8)$$

where X_i is the pre-treatment covariate and β_2 is the effect associated with the covariate. The proportion of variance within a treatment condition explained by the covariate is $R_{|x}^2 = 1 - \frac{\sigma_{|x}^2}{\sigma^2}$ where σ^2 is the between-person variance within a treatment condition and $\sigma_{|x}^2$ is the variance conditional on the covariate.

Under the alternative hypothesis, the F -statistic follows a non-central F -distribution with 1 numerator and $n - 3$ denominator degrees of freedom. Note that the inclusion of a covariate reduces the degrees of freedom by 1. The standardized version of the non-centrality parameter is:

$$\lambda = \frac{n\delta^2}{1/[P(1-P)](1-R_{|x}^2)} \quad (9)$$

As Eq. (9) suggests, λ is a function of the sample size (n), the percent of variance within a treatment condition explained by the covariate ($R_{|x}^2$), the treatment effect which was standardized the same way as in the case without the covariate (δ), and the proportion of individuals assigned to the treatment condition (P). As $R_{|x}^2$ increases, the non-centrality parameter increases, thus the power also increases.

Continuing with the example of designing a study to test an online mathematics tutoring program, the researcher could include a pre-intervention measure that is highly correlated with the mathematics achievement outcome, such as the student's mathematics achievement outcome from the previous year, to increase the power of test for a given sample size and treatment effect.

Power calculation for individual level RCTs

There are two main approaches to conducting a power analysis: the sample size approach and the effect size approach. The sample size approach calculates the number of individual level units necessary for a study to detect a standardized main treatment effect of a given magnitude with a specified level of power. Note that power = 0.80 is generally the acceptable level of power for designing educational studies. Software programs, like Optimal Design Plus (Raudenbush et al., 2011), CRT Power (Borenstein and Hedges, n.d.), and PowerUp (Dong and Maynard, 2013)! are freely available for conducting power analysis for individual-level and multi-level RTs. These software programs can also be used to conduct power analysis for studies with non-randomized designs.

Suppose the researchers designing the online mathematics tutoring program are interested in determining the number of students they need to detect a treatment effect of 0.20 standard deviations (SD). Fig. 1 shows the study needs at least 790 students to detect a treatment effect of 0.20 at power = 0.80 if using a model without covariates (solid line). However, as the dashed line in Fig. 1 demonstrates, the study only needs to recruit 288 students if using a model with a pretest covariate that explains 64% of the between-student variance within a treatment condition ($R_{|x}^2 = 0.64$). This is an example where choosing appropriate covariate(s) in the model could significantly reduce the cost associated with the study. Note that we set the target treatment effect size as 0.20 SD in this example. This is based on the effect sizes of 0.20–0.30 SD commonly found in education interventions (Hill et al., 2008).

The effect size approach determines the minimum detectable effect size (MDES) a study can detect with a given number of individual-level units at a given level of power, typically power = 0.80 (Bloom, 1995). Continuing with the example of an online mathematics tutoring program, suppose the research team is limited to a sample of 500 students. They want to know the MDES assuming power = 0.80. Fig. 2 demonstrates that the MDES is 0.25 based on the model without covariates (solid line). However, the dashed line in Fig. 2 shows that the MDES is 0.15 when considering a pretest covariate that explains 64% of the between-student variance within a treatment condition. In other words, the study could detect a smaller treatment effect with power = 0.80 when the model includes a covariate, while other parameters remain constant.

Power for two-level CRTs

In a CRT, groups of individuals, or clusters are randomly assigned to treatment and comparison conditions. CRTs are common in educational settings as interventions are often implemented at the school level, e.g., whole school reform models, new curricula, and hence entire schools are assigned to condition. However, outcomes are measured at the student level creating a two-level design.

Model without covariates

Suppose a research team is interested in assessing the effectiveness of a new reading program and plans to assign schools to either implement the new program or carry on with the program that is already in place. The team seeks to determine the impact of the

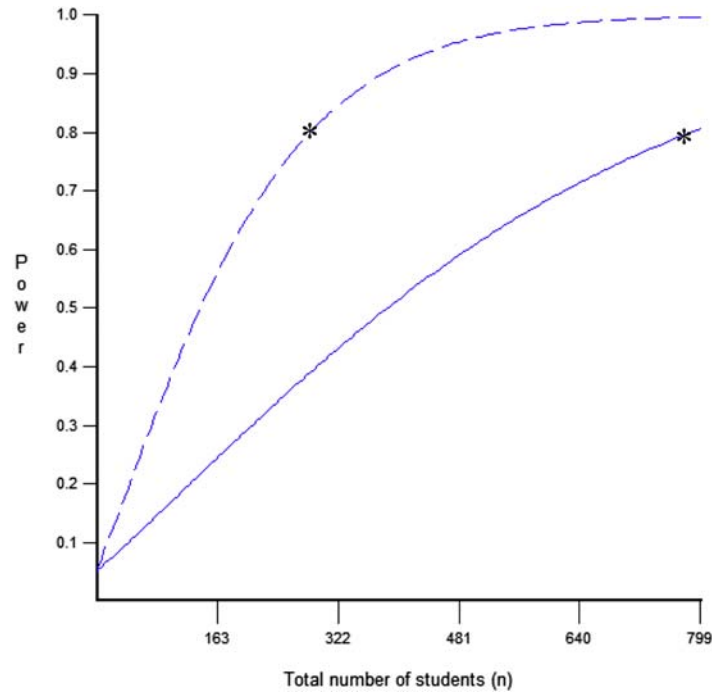


Fig. 1 Power versus total number of individuals for individual-level RTs based on the model without covariates (*solid line*) and the model with a covariate that explains 64% of the between-person variance within treatment condition (*dashed line*) at $\alpha = 0.05$. The asterisks represent points where power = 0.80. Note: This figure was retrieved from Optimal Design Plus (Raudenbush et al., 2011).

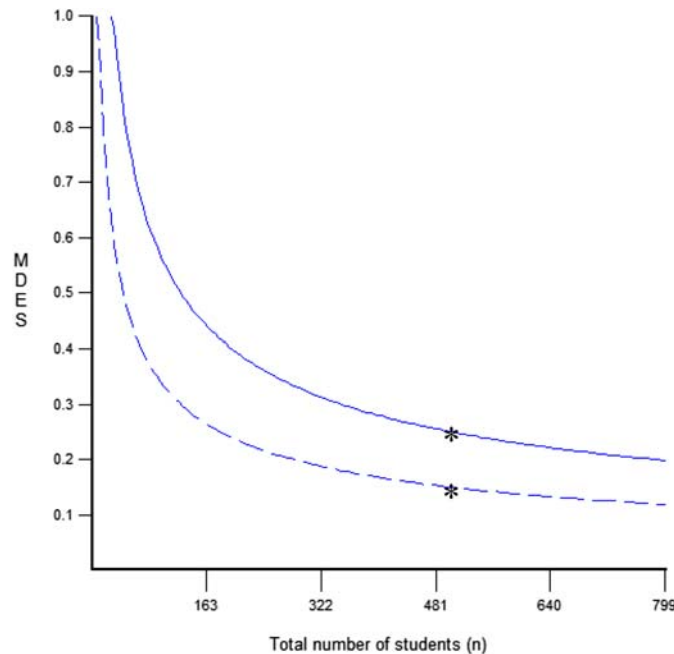


Fig. 2 MDES versus total number of individuals for individual-level RTs based on the model without covariates (*solid line*) and the model with a covariate that explains 64% of the between-person variance within a treatment condition (*dashed line*) at $\alpha = 0.05$. The asterisks represent points where $n = 500$. Note: This figure was retrieved from Optimal Design Plus (Raudenbush et al., 2011).

program on student reading achievement. This is a typical two-level CRT with students nested within schools, schools randomly assigned to treatment and comparison conditions and outcome measured at the student level.

The level-1, or the student level model is:

$$Y_{ij} = \beta_{0j} + e_{ij} \quad e_{ij} \sim N(0, \sigma^2) \quad (10)$$

where Y_{ij} is the reading score for individual $i = \{1, \dots, n\}$; β_{0j} is the mean reading score for school j ; e_{ij} is the residual error associated with students with variance σ^2 .

The level 2, or the school-level model is:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T_j + r_{0j} \quad r_{0j} \sim N(0, \tau_{00}) \quad (11)$$

where γ_{00} is the grand mean reading achievement; γ_{01} is the mean difference between the treatment and control groups or the main treatment effect; T_j is a treatment indicator with $-1/2$ for control and $1/2$ for treatment, and r_{0j} is the residual error associated with schools within a treatment condition variance τ_{00} .

The treatment effect is estimated by $\hat{\gamma}_{01} = \bar{Y}_E - \bar{Y}_C$ where $\bar{Y}_E$ is the mean for the treatment group and $\bar{Y}_C$ is the mean for the control group. The variance of the estimated treatment effect is:

$$\text{Var}(\hat{\gamma}_{01}) = \frac{(\tau_{00} + \sigma^2/n)}{JP(1-P)} \quad (12)$$

Eq. (12) shows that the variance of the estimated treatment effect is a function of the between-cluster variance within a treatment condition (τ_{00}), within-cluster or between-student variance (σ^2), the sample size within a cluster (n), and the total number of clusters (J). The associated hypothesis is:

$$\begin{aligned} H_0 : \gamma_{01} &= 0 \\ H_1 : \gamma_{01} &\neq 0 \end{aligned} \quad (13)$$

and the power for the test is:

$$\begin{aligned} \text{Power} &= \text{Prob}(\text{Reject } H_0 | H_0 \text{ is false}) \\ &= \text{Prob}(F > F_{\alpha; 1, J-2}) \\ &= 1 - \text{Prob}(F < F_{\alpha; 1, J-2}) \end{aligned} \quad (14)$$

In the case of 2-Level CRT, the F -statistic = $\frac{(MS_{\text{treatment}})}{(MS_{\text{cluster}})}$. Under the null hypothesis, the F statistic follows a central F distribution with 1 numerator degree of freedom and $J - 2$ denominator degrees of freedom. Under the alternative hypothesis, the F statistic follows a non-central F distribution represented by the non-centrality parameter λ .

$$\lambda = \frac{\gamma_{01}^2}{\text{Var}(\hat{\gamma}_{01})} = \frac{JP(1-P)\gamma_{01}^2}{(\tau_{00} + \sigma^2/n)} \quad (15)$$

As in the case of the single-level RT, it is common to standardize the parameters, Hence we can re-express λ as:

$$\lambda = \frac{JP(1-P)\delta^2}{[\rho + (1-\rho)/n]} \quad (16)$$

where δ is the standardized treatment effect, $\delta = \frac{\gamma_{01}}{\sqrt{\tau_{00} + \sigma^2}}$ and ρ is the intraclass correlation (ICC), $\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$, or the proportion of between-cluster variance within a treatment condition relative to the total variance within a treatment condition.

Recall that the power increases as the noncentrality parameter increases. Eq. (16) demonstrates that power increases as the treatment effect (δ) increases and the sample size increases, and the ICC (ρ) decreases. As noted earlier, the magnitude of the treatment effect is a function of the intervention and hence not typically under the control of the researcher. The ICC is a function of the clusters. That is, the more homogenous the clusters, schools for example, the smaller the ICC. Therefore, the researcher may have some control in reducing the ICC during the sampling stage of the study. However, the sample size is still the parameter that the researcher can influence the most. It is important to note that in a two-level CRT, increasing the total number of clusters (J) has a much stronger effect on the power to detect the main treatment effect than increasing the number of individuals per cluster (n).

Model with a cluster-level covariate

Similar to the case of an individual-level RT, it is common to include covariates to increase the precision of the estimated treatment effect (Bloom et al., 2007). In a two-level CRT, cluster-level covariates, such as school mean achievement scores are frequently incorporated in impact studies because they are more readily available and less costly to collect compared to the data for individual-level covariates. In addition, they explain the between-cluster variance and thus drive down the ICC, which increases the power to detect the treatment effect, holding everything else constant. Hence, we limit our discussion to the case of cluster-level covariates.

The level-1 model is the same as the level-1 model without covariates. The level-2, or the cluster-level model is

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T_j + \gamma_{02}W_j + r_{0j} \quad r_{0j} \sim N(0, \tau_{00|W}) \quad (17)$$

where W_j is the cluster-level covariate and γ_{02} is the effect associated with the covariate. The proportion of between-cluster variance within a treatment condition explained by the covariate is $R^2_{W|T} = 1 - \frac{\tau_{00|W}}{\tau_{00}}$.

Under the alternative hypothesis, the F -statistic follows a non-central F -distribution with 1 numerator and $J - 3$ denominator degrees of freedom. Note that the inclusion of a covariate reduces the degrees of freedom by 1. The standardized version of the non-centrality parameter is:

$$\lambda = \frac{JP(1-P)\delta^2}{\left[(1 - R^2_{W|T})\rho + (1 - \rho)/n \right]} \quad (18)$$

Note that the difference between Eq. (18) and the non-centrality parameter for the model without covariates (Eq. 16) is the factor $(1 - R^2_{W|T})$, which reduces the between-cluster variance within a treatment condition. However, since a cluster-level covariate is at level-2, it does not reduce the within-cluster variance. The more variance explained by the cluster-level covariate, the greater the power, holding everything else constant.

The proportion of variance between-clusters within a treatment condition (ICC) and the proportion of variance explained by a cluster-level covariate ($R^2_{W|T}$) are referred to as the design parameters for conducting a power analysis. These design parameters are context-specific and depend on the specific outcome, sample, nesting structure, etc. Empirical estimates of these design parameters for some outcomes, such as student achievement outcomes, are widely available in the literature (Hedges and Hedberg, 2007; Hedges et al., 2012; Kelcey and Shen, 2016; Spybrook et al., 2016a,b). The literature generally suggests the school-level ICCs, such that involved in the power calculations for two-level CRTs where schools are at level-2, are in the range of 0.14–0.26 for student mathematics and reading achievement outcomes (Hedges and Hedberg, 2007) and 0.17–0.31 for student science achievement outcomes (Spybrook et al., 2016a,b). Similarly, studies of teacher development that draw on two-level CRTs where teachers are nested within schools 2 suggest that ICCs are in the range of 0.16–0.35 for mathematics knowledge for teaching and 0.08–0.24 for reading knowledge for teaching (Kelcey and Phelps, 2013a,b). The school-level R^2 values associated with achievement pretest covariates are 0.71–0.80 for student mathematics, reading, and science achievement outcomes (Hedges and Hedberg, 2007; Spybrook et al., 2016a,b). These design parameters are pertinent in the accuracy of power calculation for CRTs. In the event the empirical estimates are not available for a particular outcome or context, a small pilot study in similar schools with a similar outcome may be useful for generating estimates of the design parameters.

Power calculation for two-level CRTs

The two approaches for conducting power calculation for individual-level RTs also apply to CRTs. Suppose a research team assigns the new reading program to schools, such that half of the schools receive the intervention and the other half continues with the current program. The team is interested in examining the program's impact on students' reading achievement. Using the sample size approach, the team wants to determine the number of schools the study needs to detect a treatment effect of 0.20 with 50 students per school at power = 0.80. The team assumes that 15% of the variance is between the clusters (i.e., ICC = 0.15). Fig. 3 shows the study needs at least 132 schools based on the model without any covariates (solid line). Suppose the team plans to include the school average reading score from a year prior to implementing the intervention as a level-2 covariate. Further, this covariate explains 80% of the between-cluster variance within a treatment condition. Fig. 3 shows the study only needs approximately 40 schools to achieve a power of 0.80 (dashed line), which is approximately 1/3 of those needed for the model without any covariate.

Similarly, the research team could use the effect size approach for power analysis. Suppose the same the research team has a limited budget and they could only recruit 50 schools and 50 students per school. Fig. 4 demonstrates that the MDES is 0.33 based on the model without any covariates (solid line), which is larger than the 0.20–0.30 effect size that studies in education typically are designed to detect. However, when the school average pretest covariate is included as a cluster-level covariate that explains 80% of the between-cluster variance within a treatment condition, the MDES is 0.18 (dashed line in Fig. 4). This new MDES is below the range of 0.20–0.30, which suggests that the study is sufficiently powered to detect a meaningful treatment effect. This example highlights the importance of cluster-level covariates in reducing the sample sizes needed to sufficiently power a study.

Future directions

As the field is accumulating more knowledge on the main treatment effects of educational interventions, which answers the "what works?" question, researchers are also interested in the capacity of RTs to answer more questions. For example, "for whom?", "under what conditions?", and "how?" questions are quintessential for understanding the applicability of interventions across educational contexts and the underlying mechanisms the interventions work to improve student achievement. Researchers often rely on moderator and mediator analyses for answering these questions. As such, more researchers are planning studies and conducting power analyses for main, moderator, and mediator effects (Spybrook et al., 2020).

Moderator effects represent differential treatment effects. For example, one may ask, does the treatment have a differential effect based on gender or pretest score? From a modeling perspective, they are represented by an interaction term between the moderator

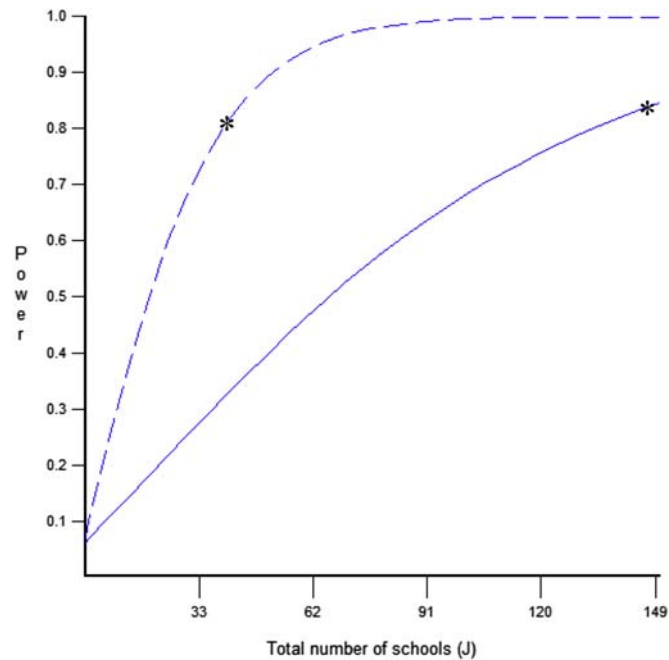


Fig. 3 Power versus total number of individuals for CRTs based on the model without covariates (*solid line*) and the model with a covariate that explains 80% of the between-cluster variance within a treatment condition (*dashed line*) at $\alpha = 0.05$. The asterisks represent points where power = 0.80. Note: This figure was retrieved from Optimal Design Plus (Raudenbush et al., 2011).

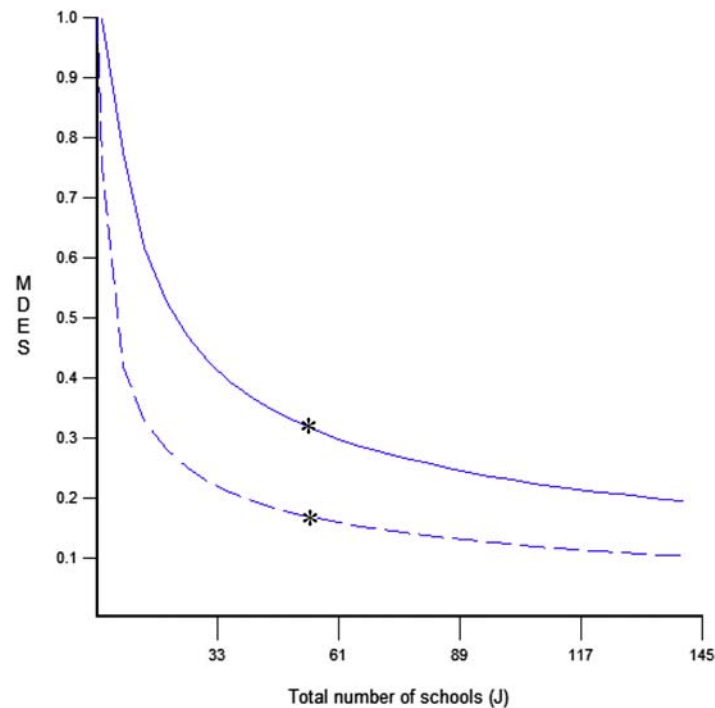


Fig. 4 MDES versus total number of individuals for CRTs based on the model without covariates (*solid line*) and the model with a covariate that explains 80% of the between-cluster variance within a treatment condition (*dashed line*) at $\alpha = 0.05$. The asterisks represent points where $J = 50$. Note: This figure was retrieved from Optimal Design Plus (Raudenbush et al., 2011).

and the treatment. In an individual-level RT, moderators are the individual-level. In a two-level CRT, moderators may be at the individual-level or cluster-level. The literature on calculating power for moderator effects in CRTs is growing (e.g., Bloom, 2005; Jaciw et al., 2016; Spybrook et al., 2016a,b; Dong et al., 2021; Dong et al., 2020) and software is also available for conducting these analyses (PowerUp!-Moderator, (Dong et al., 2017a)).

Mediator effects are key to understanding the mechanisms through which an intervention comes to be effective. For example, the researcher may want to test the extent to which a new science curriculum operates by first increasing students' interest in science such that the increased interest leads to greater science achievement. Interest in science represents the mediator in this example. Similar to moderators, mediators may be at either the individual-level or the cluster-level. Literature for power calculations for mediator effects is also growing, particularly for CRTs (e.g., Kelcey et al., 2020; Kelcey et al., 2019; Kelcey et al., 2017) and software is also available for conducting these analyses (PowerUp!-Mediator, (Dong et al., 2017b)).

However, there are many challenges ahead. One challenge being the moderator and mediator effect sizes studies should aim to detect. Education researchers typically plan CRTs to detect the main treatment effect of 0.20–0.30 SDs (Hill et al., 2008). The same reference effect sizes for planning studies to detect a moderator or mediator effects are in need of further exploration. As more CRTs are incorporating these types of analyses in their designs, we hope to see more evidence of moderation and mediation effects accumulate in the field.

References

- Bloom, H.S., Richburg-Hayes, L., Black, A.R., 2007. Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educ. Eval. Pol. Anal.* 29 (1), 30–59.
- Bloom, H.S., 1995. Minimum detectable effect: a simple way to report the statistical power of experimental designs. *Eval. Rev.* 19, 547–556.
- Bloom, H.S., 2005. Randomizing groups to evaluate place-based programs. In: Bloom, H.S. (Ed.), *Learning More from Social Experiments: Evolving Analytic Approaches*. Russell Sage, New York, NY, pp. 115–172.
- Borenstein, M., Hedges, L.V., (n.d.). *CRT-power*. Biostat, Englewood, NJ.
- Dong, N., Maynard, R.A., 2013. PowerUp!: a tool for calculating minimum detectable effect sizes and sample size requirements for experimental and quasi-experimental designs. *J. Res. Educ. Effect.* 6 (1), 24–67.
- Dong, N., Kelcey, B., Spybrook, J., Maynard, R.A., 2017a. PowerUp!-Moderator: A Tool for Calculating Statistical Power and Minimum Detectable Effect Size of the Moderator Effects in Cluster Randomized Trials (Version 1.08) [Software]. Available from: <http://www.causalevaluation.org/>.
- Dong, N., Kelcey, B., Spybrook, J., Maynard, R.A., 2017b. PowerUp!-Mediator: A Tool for Calculating Statistical Power for Causally-Defined Mediation in Cluster Randomized Trials (Beta Version 1.0) [Software]. Available from: <http://www.causalevaluation.org/>.
- Dong, N., Kelcey, B., Spybrook, J., 2020. Design considerations in multisite randomized trials to probe moderated treatment effects. *J. Educ. Behav. Stat.* 1–33.
- Dong, N., Kelcey, B., Spybrook, J., Bulus, M., 2021. Power analyses for moderator effects with (non)randomly varying slopes in cluster randomized trials. *Methodology*. In press.
- Hedges, L.V., Hedberg, E.C., 2007. Intraclass correlation values for planning group-randomized trials in education. *Educ. Eval. Pol. Anal.* 29 (1), 60–87.
- Hedges, L.V., Rhodes, C., 2010. *Statistical Power Analysis in Education Research*. U.S. Department of Education. National Center for Special Education Research, Institute of Education Sciences, Washington, DC.
- Hedges, L.V., Hedberg, E.C., Kuyper, A.M., 2012. The variance of intraclass correlation in three- and four-level models. *Educ. Psychol. Meas.* 72 (6), 893–909.
- Hedges, L.V., 2017. Challenges in building useable knowledge in education. *J. Res. Educ. Effect.* 11 (1), 1–21.
- Hill, C.J., Bloom, H.S., Black, A.R., Lipsey, M., 2008. Empirical benchmark for interpreting effect sizes in research. *Child Dev. Perspect.* 2 (3), 172–177.
- Jaciw, A.P., Lin, L., Ma, B., 2016. An empirical study of design parameters for assessing differential impacts for students in group randomized trials. *Eval. Rev.* 40 (5), 410–443.
- Kelcey, B., Phelps, G., 2013a. Strategies for improving power in school randomized studies of professional development. *Eval. Rev.* 37, 520–554.
- Kelcey, B., Phelps, G., 2013b. Considerations for designing group randomized trials of professional development with teacher knowledge outcomes. *Educ. Eval. Pol. Anal.* 35, 370–390.
- Kelcey, B., Shen, Z., 2016. Multilevel design of school effectiveness studies in sub-saharan Africa. *Sch. Effect. Sch. Improv.* 27, 492–510.
- Kelcey, B., Dong, N., Spybrook, J., Shen, Z., 2017. Experimental power for Indirect effects in group-randomized studies with group-level mediators. *Multivariate Behav. Res.* 52 (6), 699–719.
- Kelcey, B., Spybrook, J., Dong, N., 2019. Sample size planning for cluster-randomized interventions probing multilevel mediation. *Prev. Sci.* 20 (3), 407–418.
- Kelcey, B., Xie, Y., Spybrook, J., Dong, N., 2020. Power and sample size determination for multilevel mediation in three-level cluster randomized trials. *Multivariate Behav. Res.* 1–18.
- Kirk, R.E., 1982. *Experimental Design*. John Wiley, New York, NY.
- Raudenbush, S.W., Spybrook, J., Congdon, R., Liu, X.F., Martinez, A., Bloom, H., 2011. *Optimal Design Software for Multi-Level and Longitudinal Research (Version 3.01)* [Software].
- Shadish, W.R., Cook, T.D., Campbell, D.T., 2002. *Experimental and Quasi-Experimental Designs*. Wadsworth Cengage Learning, Belmont, CA.
- Spybrook, J., Kelcey, B., Dong, N., 2016a. Power for detecting treatment by moderator effect in two and three-level cluster randomized trials. *J. Educ. Behav. Stat.* 41 (6), 605–627.
- Spybrook, J., Westine, C.D., Taylor, J.A., 2016b. Design parameters for impact research in science education: a multistate analysis. *AERA Open* 2 (1), 1–15.
- Spybrook, J., Zhang, Q., Kelcey, B., Dong, N., 2020. Learning from Cluster Randomized Trials in Education: an assessment of the capacity of studies to determine what works, for whom, and under what condition. *Educ. Eval. Pol. Anal.* 42 (3), 354–374.
- U.S. Department of Education, 2003. *Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide*. Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC.

Estimation approaches: least-squares, maximum likelihood and Bayesian

Jeffrey R. Harring^a and Emma K. Harring^b, ^a University of Maryland, College Park, MD, United States; and ^b University of Maryland, College Park, MD, United States

© 2023 Elsevier Ltd. All rights reserved.

Least squares estimation	792
Generalized least squares	795
Nonlinear least squares	795
Maximum likelihood estimation	795
Maximizing the loglikelihood function	798
Hypothesis testing and confidence intervals for MLEs	799
Model comparisons	799
Properties of frequentist estimators	800
Bayesian estimation	800
Prior distributions	802
Parameter estimation via Gibbs sampling	802
Conclusion	803
References	804

To achieve understanding, researchers use statistical models to posit theoretical relations between variables in an effort to approximate basic mechanisms underlying processes being studied. Statistical models are comprised of two main elements: (1) a deterministic component and (2) a stochastic, or random, component. The former often contains parameters, which are unknown constants that convey information about the effects or associations among the variables. These parameters are estimated using the sample data. The random component of the model are also variables that are assumed to be sampled from a population distribution with known parameters. A multiple linear regression model is an example of a statistical model in which the relation between a single continuous dependent variable, y and the p independent predictors or explanatory variables, $\mathbf{x} = (x_1, \dots, x_p)$ is defined as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + e \quad (1)$$

In Eq. (1), $\beta_1, \dots, \beta_p$ are parameters to be estimated and capture the effect of each corresponding x on y . The unobserved random variable, e , in Eq. (1) is the residual or error that captures all other variables that might influence y other than the predictors, $\mathbf{x}$. In other words, it represents all the variation in y that is not explained by $\mathbf{x}$. And because e is a random variable, it carries with it a distributional assumption in the population—often assumed to be a normal distribution.

Using the sample data, the process of estimating the parameters of a statistical model is called *fitting the model*. From a frequentist perspective, many methods exist to estimate parameters of a model such as least squares, method of moments, and maximum likelihood (see, e.g., Lehmann and Casella, 1998, for a comprehensive overview regarding the derivation of several estimators and discussion highlighting their statistical properties). Another popular approach is Bayesian estimation. This approach differs both philosophically and mechanically from frequentist estimation methods. This entry provides an overview of the basic tenets of these approaches for point estimation and how variability of these estimates is defined.

Least squares estimation

A popular method of estimating parameters of a statistical model with a single dependent variable is least squares (LS) estimation (see, e.g., Weisberg, 2014). This approach is based on deriving parameter estimates that minimizes the sum of squared residuals or errors. A general model can be specified for sample size n as

$$y_i = f(\mathbf{x}_i, \boldsymbol{\theta}) + e_i, \quad i = 1, \dots, n \quad (2)$$

where y_i is the dependent variable, $\boldsymbol{\theta}$ is a vector comprised of parameters to be estimated, $f(\mathbf{x}_i, \boldsymbol{\theta})$ is some function of $\boldsymbol{\theta}$ and known values on independent variables $\mathbf{x}_i$, and e_i is the residual. By rearranging the terms in Eq. (2), the residuals can be expressed as $e_i = y_i - f(\mathbf{x}_i, \boldsymbol{\theta})$. That is, residuals are defined as the difference between the outcome, y_i , and the function value, $f(\mathbf{x}_i, \boldsymbol{\theta})$. The LSs solution, $\hat{\boldsymbol{\theta}}$, are estimates of the parameters that minimize the residual sum of squares (RSS)

$$\text{RSS} = \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \hat{\boldsymbol{\theta}})]^2 \quad (3)$$

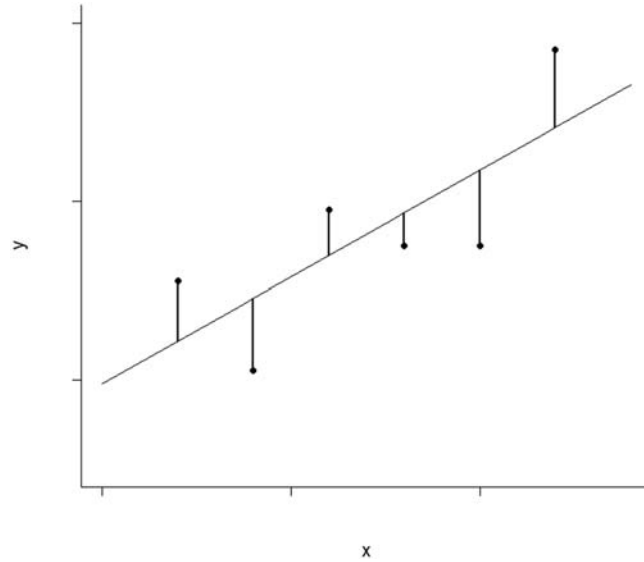


Fig. 1 Residuals, $\hat{e} = y - \hat{y}$, shown as solid vertical lines for 6 hypothetical bivariate data.

Under the principle of least squares, there is only one unique solution, $\hat{\theta}$ for which RSS is a minimum. **Fig. 1** shows the residuals (solid lines), the discrepancies between y_i and the fitted value, $\hat{y}_i = f(x_i, \hat{\theta})$ at the corresponding value of x_i across all i for six hypothetical data points.

Fig. 2 shows the least squares principle where the residuals are literally “squared” and the total area of those squares when summed, is a minimum. **Fig. 3** shows the same calculation but using $\hat{y} = \bar{y}$ as the best fitting line.

For the ubiquitous linear model (Searle and Gruber, 2016) of Eq. (1), the expression in Eq. (2) can be written in matrix notation familiarly as

$$y = X\theta + e \quad (4)$$

where X is a design matrix whose columns are the independent variables (predictors) in the model augmented with a column of ones while the rows of X represent the n cases in the sample. The predictors can be continuous (e.g., age, height) and can take on an infinite number of possible values or the predictors can be categorical (e.g., gender, ethnicity) taking on a finite, discrete number of values.

The *ordinary least squares* (OLS) solution for $\hat{\theta}$ is solved by determining the parameter estimates that minimize the sum of squared residuals. In matrix notation, this sum can be computed as follows:

$$RSS(\theta) = (y - X\theta)'(y - X\theta) \quad (5)$$

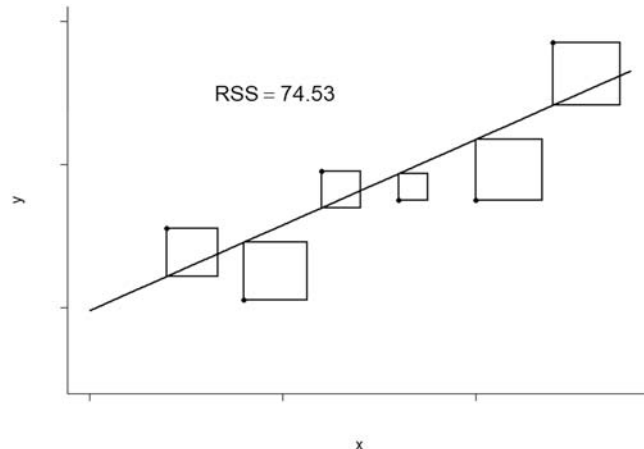


Fig. 2 Schematic of the least squares principle. The residual sum of squares for the OLS best fitting regression line is a minimum.

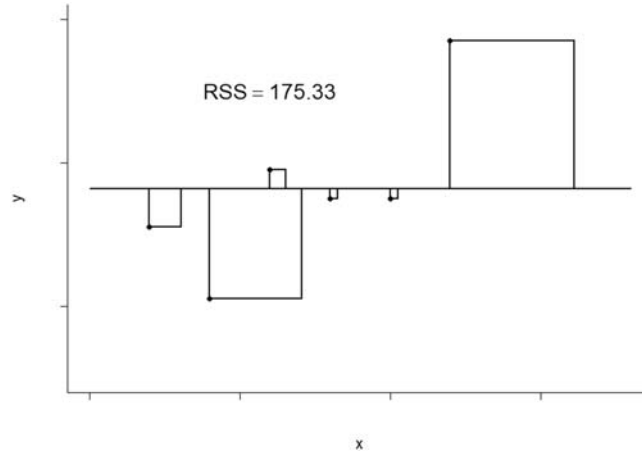


Fig. 3 Residual sum of squares where $\hat{y} = \hat{\beta}_0$, the mean of y , is the best fitting regression line.

where the prime symbol (') superscript denotes the transpose operator. The minimum of RSS can be computed by multiplying out the right-hand side of Eq. (5)

$$\text{RSS}(\theta) = \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\theta + \theta'\mathbf{X}'\mathbf{X}\theta$$

and taking the differential of RSS with respect to θ , setting this expression equal to zero, and solving for θ . The OLS estimates from the normal equations are

$$\frac{\partial \text{RSS}(\theta)}{\partial \theta} = \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\theta + \theta'\mathbf{X}'\mathbf{X}\theta$$

$$0 = -2\mathbf{y}'\mathbf{X} + 2\mathbf{X}'\mathbf{X}\theta$$

$$\hat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (6)$$

provided that the inverse, $(\mathbf{X}'\mathbf{X})^{-1}$, in Eq. (6) exists. The mean square error (MSE), $\hat{\sigma}^2$, can also be expressed in matrix notation as

$$\text{MSE} = \hat{\sigma}^2 = \frac{\mathbf{y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}}{n - p - 1} \quad (7)$$

where n is the sample size, and p is the number of predictors in the model. Other statistical components of the linear model such as the sum of squares decomposition in analysis of variance (ANOVA) for comparing group mean differences and numeric quantities such as R^2 can also be expressed in matrix notation (see, e.g., Kutner et al., 2005).

Standard errors. The vast majority of linear models are estimated via OLS. A point estimate of a parameter computed from fitting a linear model to the data presents our best guess of the unknown quantity in the population. A point estimate or confidence interval is at the heart of testing hypotheses about these coefficients and drawing statistical inferences. The other main ingredient to facilitate the testing of hypotheses of model parameters is the estimate's *standard error*. Variances of and covariances among the $p+1$ parameter estimates (including an intercept) capture the precision of these estimates and are computed using the residual variance defined in Eq. (7)

$$\text{var}(\hat{\theta}) = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$$

Standard errors are then obtained by taking the square root of the variance components which are the values on the diagonal of this $(p+1) \times (p+1)$ matrix.

$$\text{SE}(\hat{\theta})_{kk} = \sqrt{\text{var}(\hat{\theta})_{kk}} \quad \text{for } k = 1, 2, \dots, p+1$$

These standard errors can then be used to test hypotheses regarding the estimates, often in the form of a t -test, as well as to derive confidence intervals about corresponding population parameters. For a given significance level, α , (i.e., $\alpha = 0.05$)

$$t = \frac{\hat{\theta}}{\text{SE}(\hat{\theta})} \quad \hat{\theta} - t_{(1-\alpha/2)} \text{SE}(\hat{\theta}) \leq \theta \leq \hat{\theta} + t_{(1-\alpha/2)} \text{SE}(\hat{\theta})$$

where $t(1-\alpha/2)$ is the critical value from the appropriate t -distribution characterized by degrees of freedom (df) at the $100 \times (1-\alpha/2)$ percentile point—the boundary value separating the rejection and acceptance regions.

Generalized least squares

Assumptions of the linear model described in Eq. (4) include independence of observations and homogeneity of variance. That is the conditional variance of the residuals is the same regardless of the values of the independent variables associated with it. If the residuals are correlated or do not have equal conditional variance or both, then the linear model can be fitted to the data using *generalized least squares* (GLS) estimator. Parameters estimated using GLS can be calculated analytically

$$\hat{\theta} = (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{y} \quad (8)$$

where $\mathbf{\Sigma}$ is the variance-covariance matrix among the residuals with known elements given $\mathbf{X}$ with rows and columns equal to the number of observations, n . The diagonal elements of $\mathbf{\Sigma}$ are the variances of the residuals across cases and the off-diagonal elements are the covariances between the corresponding pair of residuals. When the covariance matrix $\mathbf{\Sigma}$ is diagonal (i.e., the residuals are uncorrelated), the GLS estimator is referred to as the *weighted least squares* (WLS) estimator.

The GLS structure allows for heteroscedastic patterns in the residuals. A common type of heteroscedasticity is an increasing spread (i.e., fanning) of residuals around the expected mean function. A linear model weights each observation equally when estimating model parameters with OLS. This strategy is fine as long as data are equally clustered around the fitted function across the values of the independent variables. Observations that deviate from this data cluster can have an undue influence on the parameter estimates themselves, and thus, where the fitted function is located. GLS estimation allows for differential weighting of observations through the inclusion of $\mathbf{\Sigma}$ with known elements. These weights are proportional to the reciprocal of the variance of the responses for each case. In practice, this weight matrix is almost never known a priori and must be estimated (i.e., $\hat{\mathbf{\Sigma}}$, which can be particularly challenging to compute for large datasets. The estimator that is obtained

$$\hat{\theta} = (\mathbf{X}'\hat{\mathbf{\Sigma}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{\Sigma}}^{-1}\mathbf{y} \quad (9)$$

is called the *feasible generalized least squares* estimator. There is no universal method for estimating the elements in $\hat{\mathbf{\Sigma}}$, although *iteratively reweighted least squares* has been suggested (see, e.g., Ruppert and Carroll, 1998). Most procedures typically use the residuals from fitting the model to data using OLS as a first-step. How the problem is approached depends on the specific application and on additional assumptions that may be made about the process generating the errors of the model (i.e., moving average process). Note, that if $\mathbf{\Sigma} = \mathbf{I}_n$, where $\mathbf{I}_n$ is an $n \times n$ identity matrix, then the GLS estimator reverts back to the OLS estimator.

Nonlinear least squares

Nonlinear regression models (Bates and Watts, 1988; Seber and Wild, 2003) represent a broad class in which the behavior between the dependent and independent variable is best captured and summarized by a nonlinear function—a function that has at least one parameter that enters the function in a nonlinear fashion (i.e., exponential, logistic, or Michaelis-Menten). Nonlinear functions tend to be used either when they are grounded in theory or to build known curvilinear behavior into a model. Some phenomena like growth, for example, exhibit asymptotic or leveling off behavior. Learning a task has this same type of gradual behavior both at the beginning of the learning process and toward mastery. For the former, an exponential function or a variant of Michaelis-Menten or hyperbolic function (see, e.g., Thurstone, 1919) may be useful to model data with this asymptotic behavior while a logistic function conforms to relational characteristics of learning. Even when a linear approximation works well, a nonlinear model may still be used to retain a clear interpretation of the parameters.

If y is not a linear function of θ in Eq. (2), then the least squares solution (see Eqs. (6) and (8)) can generally not be derived analytically and iterative methods, like Gauss-Newton or Newton-Raphson (see, e.g., Seber and Wild, 2003, for a complete description), must be employed to estimate model parameters. This leads to an estimation method called *nonlinear least squares*. The idea is exactly the same as for linear models—find parameter estimates of the nonlinear function that minimize the sum of the squared residuals.

Maximum likelihood estimation

Estimating model parameters based on the least squares criterion—minimizing the sum of squared residuals—remains a popular technique for statistical methods falling under the *general linear model*. These include multiple linear regression, ANOVA, and their many variants. However, least squares is limited as it can only be used with certain models and sampling designs (i.e., simple random sampling from the population, independence of observations). In practice, a more flexible estimation approach such as maximum likelihood (ML) is used. ML estimation involves defining a likelihood function indicating how likely the observed sample data are as a function of possible parameter values of a statistical model. Thus, maximizing the likelihood function

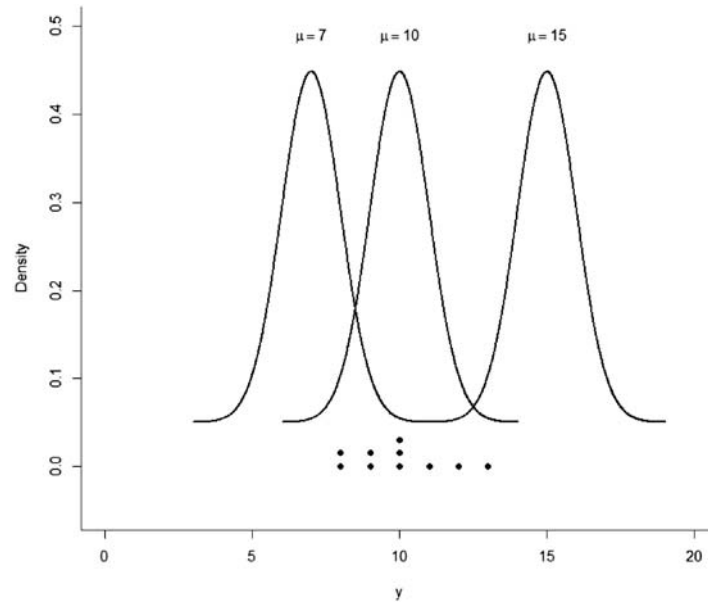


Fig. 4 Three candidate normal distributions with means at $\mu = 7$, $\mu = 10$, and $\mu = 15$ all with variance equal to 1. A dotplot of the hypothetical data is superimposed below the normal distributions.

determines the parameters that are most likely to have produced the observed data given that the data came from a particular distribution. There are generally three steps in ML estimation:

1. define the distributional assumption on which the statistical model is based,
2. construct the likelihood function—the joint density of the observations conditional on the parameters, and
3. maximize the likelihood function with respect to the parameters.

The basic methodology for ML estimation for a single parameter will be illustrated using a hypothetical data from a normal distribution. Assume that the following ten scores were sampled from a normal distribution:

$$8, 8, 9, 9, 10, 10, 10, 11, 12, 13$$

where it is known that the variance is $\sigma^2 = 1$. The mean of this distribution is unknown, however, and could take on any real value. What value of μ is most reasonable in light of the fact that (1) the population from which the sample was drawn is $y \sim N(\mu, 1)$, and (2) these particular ten observations have been obtained?

To illustrate the idea behind ML estimation, consider the three different normal distributions in Fig. 4—each is a potential candidate to have produced these data. On the bottom of Fig. 4, the ten data values are graphed as a dot plot. Above the data are three representative normal distributions which differ in terms of μ , but whose variance is $\sigma^2 = 1$. Of course, there are many other possible normal distributions, but these three illustrate the kind of match of distribution to data we want.

The likelihood function is essentially a measure of closeness that describes numerically how well the data conform to the distribution for various values of the parameter. The way this is done operationally is to compute the product of function values evaluated at each data point given particular values of the parameters. The parameter values that give the best performance in terms of similarity of the distribution to the data—maximizing the likelihood function—are called the maximum likelihood estimates (MLEs) of the parameters.

To carry on with our example, the likelihood function is constructed using the normal distribution with mean, μ and variance, $\sigma^2 = 1$

$$f(y_i | \mu, \sigma^2 = 1) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2}(y_i - \mu)^2 \right\} \quad (10)$$

Then the likelihood functions for our three candidate values of μ evaluated at the 10 data points are

$$\begin{aligned} L(\mu = 7 | y, \sigma^2 = 1) &= \prod_{i=1}^{10} f(y_i | \mu = 7, \sigma^2 = 1) \\ &= f(y_1 = 8 | \mu = 7, \sigma^2 = 1) \times \cdots \times f(y_{10} = 13 | \mu = 7, \sigma^2 = 1) \\ &= 0.24197 \times \cdots \times 0.0000061 \approx 1.796037E - 29 \end{aligned}$$

$$\begin{aligned}
 L(\mu = 10 | \mathbf{y}, \sigma^2 = 1) &= \prod_{i=1}^{10} f(y_i | \mu = 10, \sigma^2 = 1) \\
 &= f(y_1 = 8 | \mu = 10, \sigma^2 = 1) \times \cdots \times f(y_{10} = 13 | \mu = 10, \sigma^2 = 1) \\
 &= 0.05399 \times \cdots \times 0.00443 \approx 6.274323\text{E} - 10 \\
 \\
 L(\mu = 15 | \mathbf{y}, \sigma^2 = 1) &= \prod_{i=1}^{10} f(y_i | \mu = 15, \sigma^2 = 1) \\
 &= f(y_1 = 8 | \mu = 15, \sigma^2 = 1) \times \cdots \times f(y_{10} = 13 | \mu = 15, \sigma^2 = 1) \\
 &= 9e - 12 \times \cdots \times 0.0539 \approx 3.241579\text{E} - 64
 \end{aligned}$$

Because $L(\mu = 10 | \mathbf{y}, \sigma^2 = 1)$ is greater than $L(\mu = 7 | \mathbf{y}, \sigma^2 = 1)$ and $L(\mu = 15 | \mathbf{y}, \sigma^2 = 1)$, $\mu = 10$ explains the data better than the other two possibilities. For this reason it may be said that with these alternatives, the MLE is $\hat{\mu} = 10$. For the data $\mathbf{y} = (8, 8, 9, 9, 10, 10, 10, 11, 12, 13)$ the likelihood function across values of μ is shown in Fig. 5. The maximum occurs at $\hat{\mu} = 10$.

Note that in the previous hypothetical example the value of likelihood function is a very, very small number. For a variety of reasons, but especially true for computing MLEs using statistical software, the loglikelihood function is used instead of the likelihood function. The logarithm function is monotonically increasing and order-preserving, and thus, maximization of the logarithm of a function is equivalent to maximization of the function itself. With the logarithmic transformation, the very small numbers computed using the likelihood function will be converted to large negative numbers—ameliorating the analytic computations as well as underflow of the computer's numerical precision.

In general, there are more than three alternative values for a parameter—as was the case in the hypothetical example above. For instance, μ could take on any real numbered value (i.e., $\mu \in \mathbb{R}$). To determine the MLE, one evaluates the loglikelihood,

$$l(\mu | \mathbf{y}, \sigma^2 = 1) = \log L(\mu | \mathbf{y}, \sigma^2 = 1) = -\frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2 \quad (11)$$

taking $\hat{\mu}$ to be the value for which the loglikelihood is largest. This is done algebraically or numerically (i.e., using the computer), not by examining values of the likelihood empirically. The algebraic solution for this single parameter ML problem is based on the fact that the slope of the tangent line (i.e. the first partial derivative of the loglikelihood function) at the maximum is 0 (see Fig. 6).

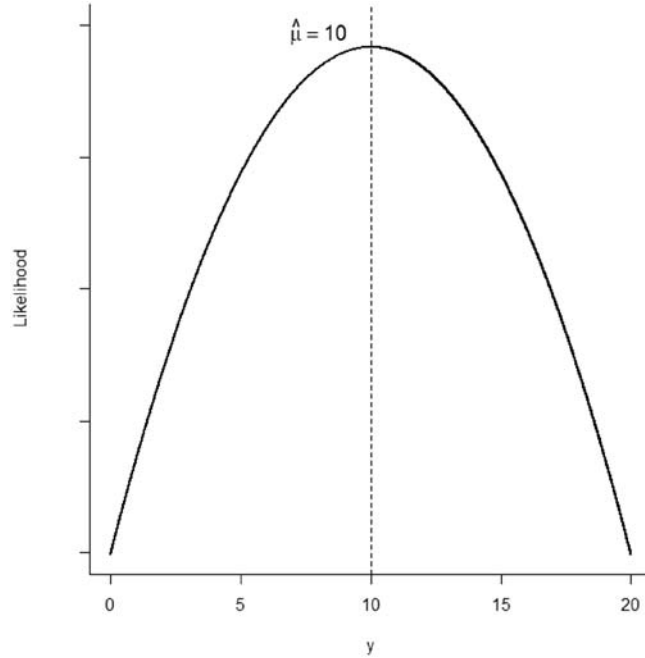


Fig. 5 Maximum likelihood estimate, $\hat{\mu} = 10$, for the mean of a normal distribution computed using the hypothetical data.

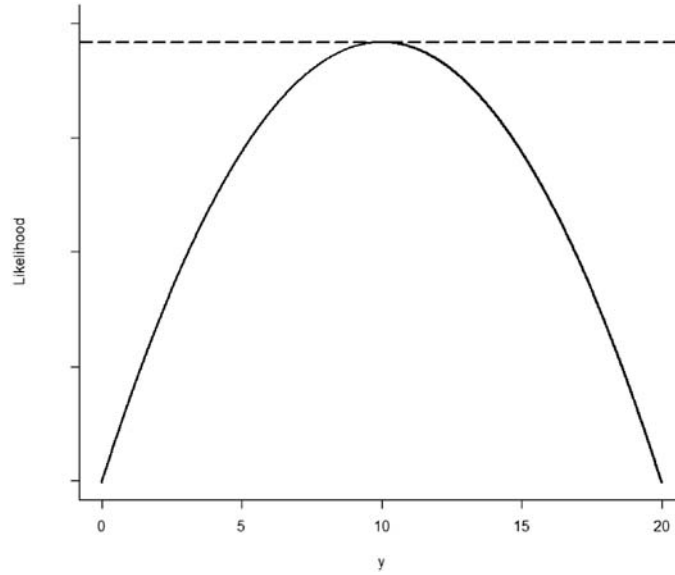


Fig. 6 The slope of the tangent line (— —) evaluated at the parameter estimate that produces a maximum value of the likelihood function, is zero.

Thus, finding the MLE requires taking the first derivative of the loglikelihood in [Eq. \(11\)](#),

$$\frac{\partial}{\partial \mu} l(\mu | y, \sigma^2 = 1) = \sum_{i=1}^n (y_i - \mu)$$

setting the expression equal to zero, and solving for the unknown parameter, μ .

$$0 = \sum_{i=1}^n (y_i - \mu)$$

$$\hat{\mu} = n^{-1} \sum_{i=1}^n y_i = \bar{y}.$$

It may come as no surprise that the ML estimate of μ is the sample mean, $\bar{y}$. Maximum likelihood can also be used to estimate several parameters, $\theta = (\theta_1, \dots, \theta_q)$ simultaneously, with $q > 1$. For example, to estimate the coefficients and residual variance in a multiple linear regression model assuming normality of the residuals. In that case it is necessary to maximize the likelihood with respect to all parameters. This can be done by computing the partial derivatives of the loglikelihood with respect to each parameter and setting each partial derivative equal to zero. It is then necessary to solve a system of equations equal to the number of parameters being estimated.

$$\frac{\partial l(\theta)}{\partial \theta_k} = 0, \quad k = 1, \dots, q$$

Generally speaking, although it is theoretically possible to write out, differentiate the loglikelihood function with respect to the various model parameters and solve the resultant system of equations, this becomes increasingly tedious even for the relatively simple models with few parameters. With the aid of computers, iterative numerical methods are used instead to solve the system of equations.

Maximizing the loglikelihood function

Several iterative methods have been developed to maximize loglikelihood functions with multiple parameters. These methods are generally categorized into derivative-free methods and methods based on using first- and/or second-derivatives of the loglikelihood function with respect to the parameters. Derivative-free methods include the Nelder-Mead algorithm ([Nelder and Mead, 1965](#)) which uses function values of the loglikelihood and a series of pre-determined steps to explore the loglikelihood surface in search for a maximum. Another such derivative-free method is simulated annealing ([Kirkpatrick et al., 1983](#)). Simulated annealing is a stochastic local search technique that approximates the global maximum of the loglikelihood function through an iterative progression in which an initial solution is repeatedly improved by making small, local alterations until no such modification yields a better solution ([Fouskakis and Draper, 2002](#)).

Newton-Raphson (N-R) or one of its variants is perhaps the most common method for facilitating ML estimation. This method uses first- and second-derivatives of the loglikelihood function and initial values (educated guesses) of the parameters must be provided by the analyst to get the algorithm started. At the $w + 1$ iteration of the algorithm, N-R updates parameters, θ as

$$\theta^{w+1} = \theta^w - \mathbf{H}(\theta^w)^{-1} \mathbf{g}(\theta^w)$$

where θ^w are the current values of the parameters, $\mathbf{g}(\theta^w)$ is the gradient (or score) vector—that is the first derivative of $l(\theta)$ with respect to each parameter in θ , whereas $\mathbf{H}(\theta^w)$ is the *Hessian* matrix—that is the matrix of second partial derivatives of $l(\theta)$. The observed information matrix from which standard errors for the parameters are obtained is the negative of the Hessian matrix and is often denoted as $\mathbf{J}(\theta^w) = -\mathbf{H}(\theta^w)$. A common variant of the N-R algorithm is the Fisher-scoring algorithm, which at each iteration, updates parameters using expected, not observed, information, $\mathbf{I}(\theta^w) = -E[\mathbf{H}(\theta^w)]$. These two approaches of approximating the information matrix from which standard errors are extracted are asymptotically (i.e., in large samples) equivalent and in practice, either could be used.

One of the reasons for its popularity is that it is said to be quadratically (rapidly) convergent in the neighborhood of the solution. This is an advantage of the N-R method but also one of its potential liabilities. If in fact, the initial parameter estimates are far away from the maximum, there is no guarantee that the N-R algorithm will converge to a maximum at all. Typically, the N-R algorithm stops and claims that a maximum has been reached when some convergence criterion such as when the set of partial derivatives of the loglikelihood function are close to zero, is met. Though it is generally possible for the N-R algorithm to find a local maximum of the loglikelihood function, it is not guaranteed that this also corresponds to the global maximum of the likelihood surface. When there are multiple local maxima (i.e., for increasingly complex statistical models) a critical analytic decision point is that of the choice of the starting values for the algorithm. Approaches like the method of moments are frequently used to obtain reasonable starting values when user-defined values or the default settings in statistical software are inadequate.

Hypothesis testing and confidence intervals for MLEs

In addition to deriving parameter estimates, it is also important to determine the accuracy of the estimates. Like least squares estimation, the standard errors of the MLEs are used for this purpose to facilitate hypothesis testing and create confidence intervals of population parameters. Standard errors are obtained from using either the expected or observed information matrix previously discussed. Again, adding to its popularity as a general optimization scheme for ML estimation, the N-R algorithm uses the observed information matrix. This is computed as the negative of the Hessian matrix, second partial derivatives of the loglikelihood, and is obtained as a byproduct of the N-R algorithm. As a result of likelihood asymptotic theory (see, e.g., [Lehmann and Casella, 1998](#)), the variance-covariance matrix of the MLEs of the parameter estimates at convergence of the algorithm, Σ , is simply the inverse of the observed information matrix. That is, $\Sigma(\hat{\theta}) = \mathbf{J}(\hat{\theta})^{-1}$. The diagonal elements will be the variances of the MLEs, and the off-diagonal elements will be the covariances between them. These are needed to test hypotheses based on linear functions of the parameters.

Model comparisons

Maximum likelihood estimation also facilitates the testing of two competing, nested models via the likelihood ratio test (LRT). The LRT assesses the goodness-of-fit of two statistical models based on the ratio of their likelihood function values (or difference of their loglikelihood function values), specifically one calculated by maximizing over the entire parameter space and the other computed after imposing some constraint(s). In a likelihood ratio test the loglikelihood value obtained under two alternative hypotheses are compared. In the model under the null hypothesis termed the *reduced* model, one or more of the parameters that are maximized in the alternative hypothesis are assumed fixed to a value (i.e., zero) or set equal to other parameters in the competing model. For example, residual variances in two linear regression models are set equal to one another. The model under the alternative hypothesis is termed the *full* model because maximum likelihood estimates are derived for all parameters. Often, the likelihood ratio test statistic is expressed as a difference between the loglikelihood values

$$\lambda_{LR} = -2 \left[l(\hat{\theta}_{\text{reduced}}) - l(\hat{\theta}_{\text{full}}) \right]$$

Under the assumption that the null hypothesis is correct, this difference will be asymptotically distributed as χ^2 on degrees of freedom equal to the difference in the number of parameters that are estimated in the two models. Large values of λ_{LR} lead to rejecting the reduced model (i.e., the null hypothesis) in favor of the full model while small values of λ_{LR} lead to a failure to reject the null hypothesis. In other words, small λ_{LR} values lead to p -values greater than the nominal 0.05 suggesting that the reduced model with fewer parameters does not fit appreciably worse than the full model. The LRT requires that the models be nested (i.e., the full model can be transformed into the reduced model by imposing constraints on the parameters of the full model). Comparing two non-nested models can be accomplished using popular information criteria such as AIC [Akaike \(1974\)](#) or BIC [Schwarz \(1978\)](#), which are formulated using the loglikelihood.

Properties of frequentist estimators

Point estimators and interval estimators are two primary types of estimators in statistics. A point estimator is a single value of a statistic from the sample data representing an estimate of its corresponding unknown parameter in the population. It is a random variable and therefore varies from sample to sample. A *confidence interval* is an interval estimator from a frequentist perspective which uses the sample data to compute a range of values that is likely to include population parameters with a certain degree of confidence. This is often expressed as a percentage (i.e., 95% confidence interval). The endpoints of the intervals are referred to as the upper and lower confidence limits.

For many classes of statistical models there are a number of estimation procedures that can be called upon when fitting models—usually specified as an option within the particular statistical software. Frequently, different procedures solve the same estimation problem. That is, the parameter estimates (and corresponding standard errors) produced by these seemingly disparate approaches, are identical. OLS and ML estimation approaches will arrive at the same parameter estimates for a multiple linear regression model, for example, if the errors are assumed to come from a normal distribution. However, this is an exception and not the rule. Parameter estimates obtained by different estimation methods (i.e., estimators) typically have different statistical properties (i.e., distributional assumptions, variance, bias). The choice between competing estimation principles is often made on the basis of these properties. There are three desirable properties every good estimator should possess: (1) unbiasedness, (2) consistency, and (3) efficiency.

A point estimator, $\hat{\theta}$, is an unbiased estimator of the true population parameter, θ , if its expected value is equal to the population parameter (i.e., $E[\hat{\theta}] - \theta = 0$; for finite sample size, n). Put another way, the unbiasedness property of an estimator implies that the average of an estimated value computed from many samples approaches the true, but unknown, population parameter. A non-zero difference indicates bias.

An estimator, $\hat{\theta}$, of the population parameter, θ , is said to be *consistent* if its sampling distribution becomes more concentrated around the true value of the population parameter as the sample size increases (i.e., $n \rightarrow \infty$). Intuitively, a consistent estimator tends to get closer to the parameter it is estimating as the sample size increases.

The most *efficient* point estimator is the one with the smallest variance of all the unbiased and consistent estimators considered with finite sample sizes. The variance measures the level of dispersion of the estimate from sample-to-sample. In short, if two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are considered, the estimator with a smaller variance—more precision—is preferred.

Choosing an estimator (i.e., maximum likelihood) for a particular application may seem rather challenging, but the reality is much less intimidating. Many classes of models typically have a default estimator that capitalizes on the statistical properties previously mentioned. For example, the default estimation approach for the broad class of generalized linear models (e.g., logistic regression) in many statistical software programs is maximum likelihood whereas ordinary least squares is the default estimation scheme for linear regression models. The choice is not as clear for models accommodating multiple categorical, dependent variables (e.g., structural equation models, generalized linear mixed effects models) and a final decision is often based on ease of implementation, computational effort, in addition to sound statistical properties. Even if the default estimator is ultimately chosen, it is important to know what trade-offs exist in the context of your data and statistical model.

Bayesian estimation

For frequentist estimation approaches like least squares and maximum likelihood, data are viewed as random while population parameters such as linear regression coefficients are treated as unknown, but fixed quantities. The implication is that a single population value can be uncovered for each unknown parameter. The primary goal in estimating such quantities is to facilitate statistical inference—that is to construct confidence intervals around the sample point estimates and to carry out hypothesis tests based on assuming an appropriate statistical null hypothesis in hopes that the data will provide sufficient evidence against it. It may be interest, for example, to determine which independent variables are the best predictors of high school dropout using a logistic regression model. In such a model, testing whether individual regression parameters are zero in the population is of interest as is constructing confidence intervals to convey ranges of values that these parameters are likely to be with a given probability.

In the Bayesian paradigm, unknown population parameters are treated as random variables that follow a certain distribution and the data, once obtained, are viewed as fixed. Prior knowledge or beliefs about the possible shape, location, and variability of this distribution are modeled by specifying a prior distribution on the parameters that will be updated by the data through the likelihood function to form a posterior distribution for the parameters. This relation between prior distribution of a parameter θ , the distribution of the likelihood and the posterior distribution is displayed in Fig. 7.

The rationale behind Bayesian estimation is grounded in Bayes's theorem, which states

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (12)$$

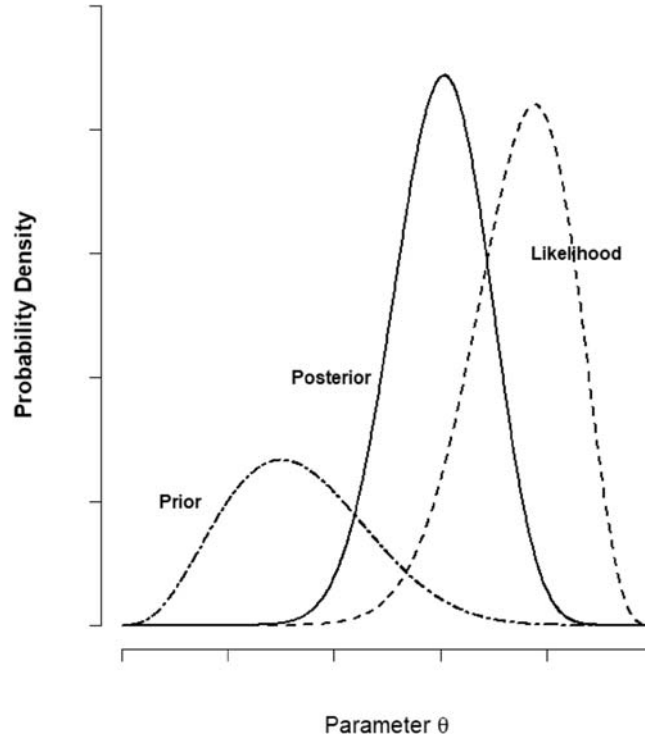


Fig. 7 Prior, likelihood, and posterior distributions for parameter θ .

The numerator of the expression on the right-hand side of Eq. (12) represents the conditional probability distribution, $p(y|\theta)$ of the data y given model parameters, θ . For fixed, known data this distribution is the likelihood function and is multiplied by the marginal distribution, $p(\theta)$ which is referred to as *prior probability distribution* of model parameters. The denominator in Eq. (12) is $p(y) = \int p(y|\theta)p(\theta)d\theta$ and represents the marginal probability distribution of the observed data in the case of continuous parameters or $\sum_{\theta} p(y|\theta)p(\theta)$ for discrete parameters with the sum being taken over all possible values of θ . In general, direct computation of the integral (or sum) in the denominator of Eq. (12) is difficult to carry out analytically—even for simple models. Yet, this barrier can be bypassed by noticing that $p(y)$ does not vary as a function of θ and consequently, is often dropped from the expression. Bayes's theorem is frequently represented as a proportional relation (Carlin and Thomas, 2009),

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (13)$$

The expression on the left-hand side of Eq. (13), $p(\theta|y)$, is referred to as the posterior distribution, which captures what is believed about the parameters after having integrated the information contained in the data. According to Levy (2016), the posterior distribution is the “solution” or “answer” from a Bayesian analysis. In other words, inference from a Bayesian analysis is focused on constructing and sampling from the posterior distribution, and as a distribution, it is often summarized and communicated with typical graphics and summary statistics. For example, histograms or density plots may be used to visualize patterns of symmetry or asymmetry. The mean, median or mode of the posterior distribution may be used as a measure of central tendency while the standard deviation of the posterior could be used as a measure of dispersion or variability. Other quantities of interest that summarize the posterior include Bayesian posterior p values and Bayes Factors (BFs) for hypothesis testing as well as constructing a *credibility interval* (i.e., Bayesian “confidence interval”) which indicates a region of the distribution with a given probability. The 90% highest posterior density (HPD) credibility interval of θ , for instance, is the interval for which the probability that θ is inside that interval in the posterior distribution is 0.90. The HPD credibility interval is constructed in such a way that no value outside the interval has a higher posterior density than any value inside the interval.

The Bayesian approach to estimation is an attractive alternative to frequentist methods like maximum likelihood because Bayesian statistics do not rely on large sample theory (i.e., central limit theorem) to make valid inferences and therefore, may produce reasonable results even with small to moderate sample sizes. This is particularly true when background knowledge about the model parameters is available. Generally, the more information a researcher can specify before observing the data, the smaller the sample size required to obtain the same certainty compared to an analysis that incorporates no prior knowledge. The Bayesian framework allows the researcher to interact in estimating model parameters in a way that he or she is unable to do in the frequentist estimation paradigm. This type of interaction is done through the specification of prior distributions of model parameters.

Prior distributions

Mechanically, incorporating prior distributions of model parameters into the estimation process distinguishes the Bayesian approach to estimation from frequentist methods. Prior distributions are probabilistic in nature that reflect an investigator's previous beliefs and knowledge of certainty about these parameters, such as their location or scale. These beliefs are specific to the investigator (i.e., two investigators may assert dissimilar beliefs about characteristics of a distribution regarding the same parameter). Great uncertainty in model parameter values is quantified by specifying a diffuse or non-informative prior distribution that is characterized by large dispersion. Examples of diffuse priors might include a uniform distribution over the real number line or a normal distribution with an enormous variance. Greater certainty about model parameter values translates into specifying so-called informative or subjective prior distributions in which the strength of one's belief is inversely related to the variability of the distribution. In fact, there is a continuum of informativeness that can be built into estimating parameters with a Bayesian framework. Fig. 8 shows three different normal distributions labeled as non-informative, weakly informative and informative to provide some visual context to what these descriptors mean. Note that the location of each of the normal distributions is centered at 2 (i.e., unstandardized regression coefficient) but have varying degrees of variability.

The combination of large sample size and diffuse prior distributions yield results that appear similar to those from a frequentist analysis. The posterior mean tends to be close to the MLE, the posterior standard deviation appears to be close to the standard error, and the limits of a posterior credibility interval tends to be close to a confidence interval with the same coverage. In large samples, the likelihood, not the prior distribution, seems to dominate the distributional characteristics of the posterior. When conducting a Bayesian analysis with small sample sizes, the prior distribution can play a more important role. In practice, prior distributions are often not carefully chosen, and many applied studies using Bayesian analysis rely on default software settings or generic heuristic values. These naive priors (McNeish, 2016) are in contrast to non-informative prior distributions, both of which can actually be quite informative if poorly chosen and when the sample size is small. A sensible justification should be provided for why a particular prior is suitable for a specific parameter. Other characteristics of prior distribution specification such as choosing the functional form of the distribution, conjugate and improper priors, and hyperpriors in a fully Bayesian analysis are more fully discussed by (Levy and Mislevy, 2016; Rupp et al., 2004).

Parameter estimation via Gibbs sampling

In the vast majority of modeling situations in which Bayesian estimation is utilized, a closed-form solution for the posterior does not exist or is analytically intractable. In these cases, inference proceeds by implementing a computational-intensive approach that

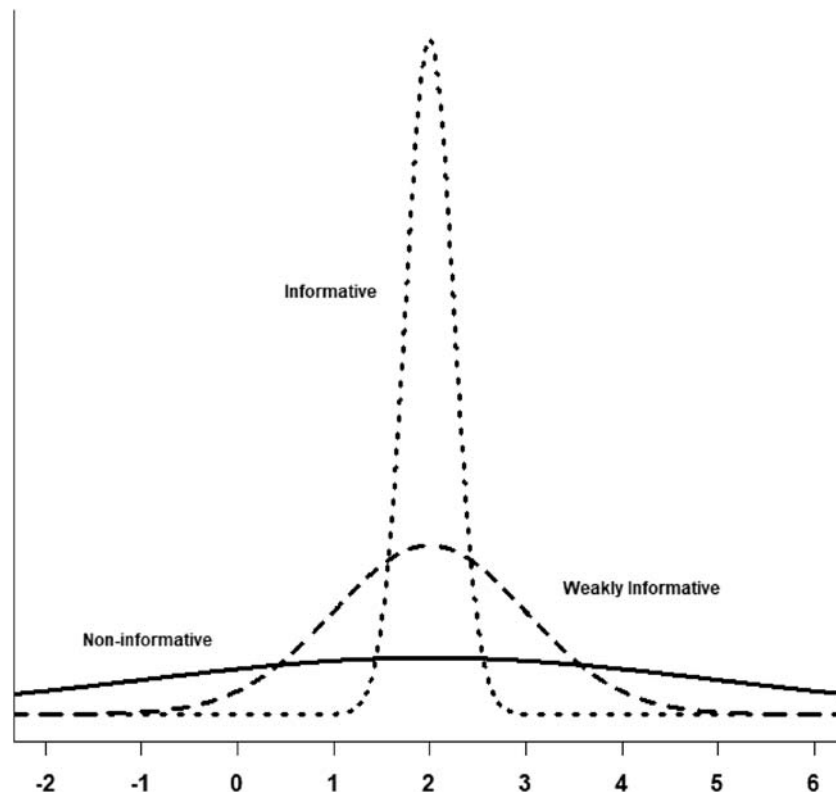


Fig. 8 Non-informative (—), weakly informative (— — —), informative (·····) normal prior distributions for an unstandardized regression coefficient located at 2.

approximates the posterior distribution instead. Sampling from the posterior distribution can be obtained via Markov chain Monte Carlo (MCMC) simulation methods. MCMC methods are useful for sampling from high-dimensional multivariate densities by separating these densities into more tractable univariate or smaller dimensional multivariate densities [Lynch \(2007\)](#). That is, MCMC methods, avoids sampling directly from the joint posterior distribution by drawing samples from the full conditional posterior distribution of each parameter conditioned on the data and the most recent value of all other parameters (see, e.g., [Levy and Mislevy, 2016](#)). Popular MCMC algorithms that facilitate estimation of model parameters include the random-walk Metropolis-Hastings algorithm, the “No-U-turn” sampler, an extension of the Hamiltonian Monte Carlo algorithm, and the ubiquitous Gibbs sampler.

The idea behind the Gibbs sampler is to generate a sequence of samples from the joint probability distribution (i.e., the posterior distribution) of two or more random variables by iterating through each variable (or block of variables), drawing a sample from its conditional distribution with the remaining variables fixed to their current values evaluated with the sample data. The sequence of draws is linked, forming a chain, in that the current value in the chain depends on the previous value in the chain. In a Bayesian analysis, the chain is constructed such that it converges to the posterior distribution. In other words, a large enough number of draws in the chain may be considered drawing samples from the posterior distribution. Gibbs sampling is applicable when the joint distribution is not known explicitly or is difficult to sample from directly, but the conditional distribution of each variable (or small set of variables) is known and is easier to sample. To be concrete, suppose a Bayesian analysis was to be performed for a multiple linear regression model with three predictors $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e$. The parameters of the model to be estimated are $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \sigma_e^2)$. The joint distribution of the parameters given the data, $p(\theta|y, x)$ makes up the posterior distribution. The basic idea of Gibbs sampling is to iteratively sample from the conditional distribution $p(\theta_d|y, x, \theta_{-d})$ where θ_{-d} denotes the parameter vector, θ without the parameter d . Then the algorithm proceeds in the following manner over $B(1, \dots, B)$ draws:

Step 1: Provide initial starting values for the parameters, $\theta^{(0)} = (\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}, \beta_3^{(0)}, \sigma_e^{2(0)})$

Step 2: Generate draws from the conditional distributions at iteration $b + 1$ as follows:

$$\begin{aligned}\beta_0^{(b+1)} &:= \beta_0^* \sim p(\beta_0^{(b)} | y, x, \beta_1^{(b)}, \beta_2^{(b)}, \beta_3^{(b)}, \sigma_e^{2(b)}) \\ \beta_1^{(b+1)} &:= \beta_1^* \sim p(\beta_1^{(b)} | y, x, \beta_0^{(b+1)}, \beta_2^{(b)}, \beta_3^{(b)}, \sigma_e^{2(b)}) \\ \beta_2^{(b+1)} &:= \beta_2^* \sim p(\beta_2^{(b)} | y, x, \beta_0^{(b+1)}, \beta_1^{(b+1)}, \beta_3^{(b)}, \sigma_e^{2(b)}) \\ \beta_3^{(b+1)} &:= \beta_3^* \sim p(\beta_3^{(b)} | y, x, \beta_0^{(b+1)}, \beta_1^{(b+1)}, \beta_2^{(b+1)}, \sigma_e^{2(b)}) \\ \sigma_e^{2(b+1)} &:= \sigma_e^{2*} \sim p(\sigma_e^{2(b)} | y, x, \beta_0^{(b+1)}, \beta_1^{(b+1)}, \beta_2^{(b+1)}, \beta_3^{(b+1)})\end{aligned}$$

Step 3: Repeat until convergence.

A primary advantage of Gibbs sampling is its simplicity. In contrast to the Metropolis-Hastings algorithm, proposed links (the next draw from the conditional distribution) in the chain are always accepted. The primary disadvantage is that the conditional distributions, like those in the previous example, must be derived for successful implementation. This is a tractable problem when the prior probability distribution $p(\theta_d)$ is conjugate to the posterior distribution. This distinction is more theoretical than practical as most statistical software that conducts Gibbs sampling, as well as other MCMC algorithms, to some extent automates the computational “heavy-lifting,” ameliorating concerns that data analysts or researchers might have.

Lastly, this short introduction to Bayesian estimation did not include any discussion of important computational issues, challenges, and analytic decision points such as assessing convergence, strategies of assigning multiple chains, choosing a burn-in period when sampling, evaluating auto-correlation, assessing data-model fit, and model selection. The interested reader is directed to texts by [Levy and Mislevy, 2016](#) and [Kaplan, 2014](#) for detailed explanation of these, and other, important concepts when taking a Bayesian approach to estimating parameters of a model.

Conclusion

George [Box et al. \(2005\)](#) is often credited with the insightful maxim “While all models are wrong, some are useful” (p. 440). Although many statistical models do not perfectly map onto the complex underlying processes found in educational research studies, they are often useful approximations. These sentiments were nicely captured in the title of [Micceri’s \(1989\)](#) paper, “The unicorn, the normal curve, and other improbable creatures.” The normal curve does not actually occur in nature, but it is a very useful model that adequately summarizes many continuously-scaled variables. Whether these models are good representations of the underlying theoretical mechanism or empirically driven through data exploration, parameters corresponding to interesting facets of the phenomenon summarized by these models will need to be estimated. This model fitting activity is at the heart of *statistical inference*. From both frequentist and Bayesian perspectives, many estimation approaches exist that are regularly used in this process. Although there are many estimators that were not discussed in this entry (i.e., method of moments, iteratively reweighted least squares), omnipresent approaches, ordinary least squares, maximum likelihood, and Bayesian methods via MCMC algorithms, were. Their popularity notwithstanding, OLS, ML, and Bayesian estimators have desirable statistical properties and can be effectively used in a variety of data analytic situations (i.e., nested data structures, longitudinal and cross-sectional designs, increasingly complex models) and across nuanced data conditions (i.e., missing data).

References

- Akaike, H., 1974. A new look at the statistical model identification. In: IEEE Transactions on Automatic Control AC, 19, pp. 716–723.
- Bates, D.M., Watts, D.G., 1988. Nonlinear Regression Analysis and its Applications. Wiley, New York, NY.
- Box, G.E.P., Hunter, J.S., Hunter, W.G., 2005. Statistics for Experimenters, second ed. Wiley, New York, NY.
- Carlin, B.P., Thomas, A.L., 2009. Bayesian Methods for Data Analysis, third ed. Chapman & Hall/CRC, Boca Raton, FL.
- Fouskakis, D., Draper, D., 2002. Stochastic optimization: a review. Int. Stat. Rev. 70, 315–349.
- Kaplan, D., 2014. Bayesian Statistics for the Social Sciences. The Guilford Press, New York, NY.
- Kirkpatrick, S., Gelatt, C., Vecchi, M., 1983. Optimization by simulated annealing. Science 220, 671–680.
- Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., 2005. Applied Linear Statistical Models, fifth ed. McGraw-Hill Irwin, New York, NY.
- Lehmann, E.L., Casella, G., 1998. Theory of point estimation, second ed. McGraw-Hill Irwin, New York, NY.
- Levy, R., 2016. Advances in bayesian modeling in educational research. Educ. Psychol. 51, 368–380.
- Levy, R., Mislevy, R.J., 2016. Bayesian Psychometric Modeling. CRC Press, New York, NY.
- Lynch, S.M., 2007. Introduction to Applied Bayesian Statistics and Estimation for Social Scientists. Springer, New York, NY.
- McNeish, D.M., 2016. On using bayesian methods to address small sample problems. Struct. Equ. Model.: A Multidisciplinary Journal 23, 750–773.
- Micceri, T., 1989. The unicorn, the normal curve, and other improbable creatures. Psychol. Bull. 105, 156–166.
- Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7 (4), 308–313.
- Rupp, A.A., Dey, D.K., Zumbo, B.D., 2004. To bayes or not to bayes, from whether to when: applications of bayesian methodology to modeling. Struct. Equ. Model. 11, 424–451.
- Ruppert, D., Carroll, R., 1998. Transformation and Weighting in Regression. Chapman Hall, New York, NY.
- Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
- Searle, S.R., Gruber, M.H.J., 2016. Linear Models, second ed. Wiley, Hoboken, NJ.
- Seber, G.A.F., Wild, C.J., 2003. Nonlinear Regression, second ed. Wiley, Hoboken, NJ.
- Thurstone, L.L., 1919. The learning curve equation. Psychol. Monogr. 26, i–51.
- Weisberg, S., 2014. Applied Linear Regression, fourth ed. Wiley, Hoboken, NJ.

Missing data analysis

Brenna Gomer and Ke-Hai Yuan, University of Notre Dame, Notre Dame, IN, United States

© 2023 Elsevier Ltd. All rights reserved.

Overview	805
Missing-data mechanisms	805
Missing completely at random	806
Missing at random	806
Missing not at random	806
Remarks on the definitions of missing-data mechanisms	806
Methods to handle missing values	807
Ad hoc methods	807
Deletion-based methods	807
Single-imputation	807
Weighting approaches	808
Likelihood-based methods	808
Maximum likelihood	808
Multiple imputation	809
Models for nonignorable missingness	810
Bayesian methods	811
Statistical learning methods	812
CART	812
Random forest analysis	812
Other methods	812
Threats to proper handling of missing values and possible remedies	812
Violated distribution conditions	812
Tests of normality	813
Robust methods	813
The use of auxiliary variables	813
Missing-data mechanism uncertainty	813
Tests of the missingness mechanism	814
Sensitivity analysis	814
Conclusion and recommendations	815
References	815

Overview

Missing data analysis has been an active area of research for several decades and is an interdisciplinary subject. Scientific contributions to this area have been published in journals from statistics, biostatistics, survey statistics, and quantitative methods in psychology and education, among others. Missing data analysis is a broad term that is typically used to refer to both statistical aspects of missing data as well as statistical methods that are used when there are missing values in a dataset. Such methods may seek to identify characteristics of missing values, facilitate valid statistical inferences in the context of missing data, or assess the extent to which statistical results are influenced by uncertainty regarding characteristics of missing values. Missing data, missingness, and nonresponse are all terms that describe such missing values.

Statistical approaches used in missing data analysis operate by making assumptions about the underlying reason for missingness, most popularly categorized by the so-called missing-data mechanisms ([Rubin, 1976](#)).

Missing-data mechanisms

Missing-data mechanisms (or, alternatively, missingness mechanisms) describe characteristics of missing values with respect to “the relationship between missingness and the values of variables in the data matrix” ([Little and Rubin, 2002](#)). Missing values can be classified as belonging exclusively to one of the three missing-data mechanisms according to their dependence on the observed values in the dataset, missing values in the dataset, or neither. Formally, this is described through the conditional distribution of missingness, which is given by

$$f_{M|Y}(m_i|y_i, \psi) \quad (1)$$

where y_i is the set of data values for subject i in the complete dataset $Y = (y_{ij})$, ψ are parameters associated with the model of the missingness process, and $M = (m_{ij})$ is a matrix of missing-data indicators with $m_{ij} = 1$ if y_{ij} is missing and $m_{ij} = 0$ if y_{ij} is observed. The missing-data mechanisms describe how this conditional distribution of missingness can or cannot be simplified.

Missing completely at random

Under the Missing Completely At Random (MCAR) mechanism, missingness does not depend on the values of variables in the dataset—neither missing nor observed. This implies that the conditional distribution of missingness can be simplified to (Little and Rubin, 2002)

$$f_{M|Y}(m_i|y_i, \psi) = f_{M|Y}(m_i|\psi) \forall y_i, \psi \quad (2)$$

For example, suppose respondents are given a random subset of items on a survey from a larger item bank such that items are assigned to respondents with equal probability. In such planned missingness designs, the missing responses on those unassigned items for a given participant fall under the MCAR mechanism.

Missing at random

Under the Missing At Random (MAR) mechanism, missingness does not depend on the missing values themselves after the observed values have been conditioned on. That is, the conditional distribution of missingness can be simplified to (Little and Rubin, 2002)

$$f_{M|Y}(m_i|y_i, \psi) = f_{M|Y}(m_i|y_{i(0)}, \psi) \forall y_{i(1)}, \psi \quad (3)$$

where $y_{i(0)}$ denotes the set of observed values for subject i and $y_{i(1)}$ denotes the set of missing values for subject i . For example, suppose respondents are given an educational assessment that consists of two sections. Suppose also that the second section of the assessment is only administered to respondents who score above a certain threshold on the first section. This planned missingness design will create missing responses on the second section for respondents who scored below the chosen threshold. Such missing responses fall under the MAR mechanism.

Missing not at random

Under the Missing Not At Random (MNAR) mechanism, missingness depends on the missing values themselves, even after the observed values are conditioned on. In other words, there is no way to remove the dependency on such missing values statistically with the observed data. The MNAR mechanism can be further subdivided into two subtypes, called focused and diffuse MNAR (Gomer and Yuan, 2021). Missingness under the focused MNAR subtype depends only on the missing values themselves, while missingness under the diffuse MNAR subtype depends on both the missing values and the observed values in the dataset. Historically, the MNAR mechanism has typically been described in words rather than re-expressions of the conditional distribution of missingness, but recent work has provided specific formulations for the MNAR subtypes. For focused MNAR, the conditional distribution of missingness can be simplified to (Gomer and Yuan, 2021)

$$f_{M|Y}(m_i|y_i, \psi) = f_{M|Y}(m_i|y_{i(1)}, \psi) \forall y_{i(0)}, \psi \quad (4)$$

For diffuse MNAR, the conditional distribution of missingness cannot be simplified at all, and instead is given by Gomer and Yuan (2021)

$$f_{M|Y}(m_i|y_{i(0)}, y_{i(1)}, \psi) \forall y_{i(1)}, \psi \quad (5)$$

which is repeated from Eq. (1), in which y_i is replaced with $(y_{i(0)}, y_{i(1)})$.

Remarks on the definitions of missing-data mechanisms

Note that in the definitions of MAR and diffuse MNAR, $y_{i(0)}$ represents the variables that are observed and explicitly used in the formulation of the likelihood function. A survey may contain many items; those that are observed but not used in the formulation of the likelihood function are equivalent to missing values. Also note that the classification of the different missing-data mechanisms is motivated by the need to simplify the likelihood function for statistical modeling, and not just mathematical brevity.

Methods to handle missing values

Methods to handle missing values, or missing-data methods, are statistical approaches that provide either (1) an artificially complete dataset that no longer contains missing values, or (2) facilitate statistical inference even in the presence of missing data. In both cases, the end-goal of missing-data methods is typically to obtain statistical inferences for the analysis model of interest from a dataset that originally contained missing values.

Missing-data methods can also be classified according to the strategies they use to handle missing values, and typically fall under one or more of the following categories: (1) deletion-based procedures, (2) weighting procedures, (3) imputation, and (4) model/distribution-based methods.

Another important distinction is between ad hoc methods and principled methods: ad hoc methods are convenience-based approaches that generally provide a single artificially complete dataset and can only provide valid statistical inferences for certain population parameters in very restrictive conditions. Principled approaches, on the other hand, have better statistical justification and provide valid statistical inferences for parameters of the population distribution or the analysis model of interest under less restrictive settings.

Ad hoc methods

The most commonly used ad hoc approaches fall into one of two categories: (1) deletion-based methods or (2) single imputation.

Deletion-based methods

Listwise deletion, also called complete case analysis, is a deletion-based missing-data method that removes all data values associated with cases or respondents that have missing values. Under this approach, if a participant responds to 20 out of 21 survey items, the entire row of the dataset that corresponds to that participant is deleted—and even the information from the 20 observed items for that participant is lost. This also means that if a dataset contains 20 such nonrespondents out of 100 respondents in total, the sample size after listwise deletion is reduced to 80. The operating assumption of listwise deletion is that missing values are MCAR. Under this assumption, the data from the complete cases can be regarded as a random sample from the population distribution of interest and thus will still yield statistical inferences that are unbiased. However, a potent drawback to this approach is that even when the assumption of MCAR is satisfied, the reduced sample size leads to decreased statistical power. When missingness is instead MAR or MNAR, listwise deletion generally leads to invalid statistical inferences. In many statistical software programs, listwise deletion is used by default when data contain missing values and no other missing-data method is specified.

Pairwise deletion, sometimes referred to as available case analysis, is another deletion-based missing-data method. This approach uses all the observed (i.e., available) values from incomplete cases in the calculation of statistical estimates. In the case of correlations, all pairwise complete observations are used from the variables of interest. Thus, even if a participant has a missing response on item 21 of a survey, her values on the other 20 items are included when calculating those item correlations. The operating assumption of pairwise deletion is that missing values are MCAR. Under this assumption, pairwise deletion provides unbiased estimates for several types of parameters. Pairwise deletion will result in greater statistical power compared to listwise deletion. However, pairwise deletion suffers from problems of comparability: statistical estimates obtained via pairwise deletion may not be based on the same members of the sample and thus are no longer strictly comparable. The overall sample size may also be unclear when covariances are based on different numbers of observations, which can lead to biased test statistics in contexts such as structural equation modeling (Marsh, 1998). Another drawback of pairwise deletion is that it can sometimes result in correlation and covariance matrices that are not positive definite, which is problematic for multivariate statistical analysis (Wothke, 1993).

Single-imputation

Mean substitution, also called unconditional mean imputation or simply mean imputation, is a single-imputation missing-data method that replaces missing values on a variable with the mean calculated from its observed values. This will result in an artificially complete dataset that has the same sample size as the original data. Stated generously, the operating assumption of mean substitution is MCAR—but even under this assumption, mean substitution can only provide unbiased estimates with respect to the mean. This unbiasedness is possible because there is no systematic difference between the missing and observed values under MCAR, so replacing missing values with the mean will not bias the results. However, statistical estimates of other parameters are invalid even under the assumption of MCAR because the variation of the values themselves is reduced while the sample size is artificially inflated (Peugh and Enders, 2004).

Regression imputation, also sometimes called conditional mean imputation, is a single-imputation missing-data method that replaces missing values with their predicted values from a regression model. This model regresses the missing variable on the variables which are completely observed for the given case or respondent, and is often based on the subset of the sample whose values are observed for that particular set of variables. Thus, stated generously, the operating assumption of regression imputation is that missingness is MAR or MCAR. However, as with mean substitution, the variability of values on the missing variable is underrepresented by regression imputation which produces biased variance and covariance estimates even when the missing-data mechanism is MAR or MCAR (Little and Rubin, 2020; Peugh and Enders, 2004). A modification of regression imputation is stochastic regression imputation, which aims to solve the issue of biased variance-covariance estimates by adding random residual terms to the imputed

values. Some types of parameter estimates are statistically consistent under MAR and MCAR using this method (Peugh and Enders, 2004). However, regardless of which form of regression imputation is used, the implementation becomes more complicated when missingness occurs on more than one variable and respondents differ on which variables are missing.

Weighting approaches

Weighting adjustments are missing-data methods that approach the phenomenon of missingness from a sample selection perspective, in which each member of the sample has some positive probability of being selected for the sample. It is supposed that the population of interest is composed of several subpopulations, and that one of these subpopulations is prone to nonresponse and is thus undersampled compared to the others. However, some participants with observed data may belong to the subpopulation that is prone to nonresponse. The rationale of weighting approaches is to differentially weight such participants to correct the bias that comes from undersampling the subpopulation that is prone to missingness. Such weighting approaches include propensity weighting (Czajka et al., 1992; Little, 1986; Rubin, 1985), and inverse probability weights (Kish, 1995; Potthoff et al., 1992; Stapleton, 2002). The values of observed variables may be used to help determine such sampling weights. However, the bias and sampling variance of mean estimates are sensitive to the particular choice of such variables: in some cases, the bias and sampling variance are only reduced when such variables are highly associated with both nonresponse and with the outcome variable (Little and Rubin, 2020, Table 3.1).

Likelihood-based methods

Likelihood-based missing-data methods fall under the umbrella of “model/distribution-based” methods and are also principled approaches. These missing-data methods are based on the likelihood from a model/distribution for the variables of interest (denoted here as Y), missingness indicator (denoted M), and, optionally, a set of fully-observed covariates (X). The “full likelihood” involves modeling of the missingness process and is given by

$$L_{full}(\theta, \psi | Y_{(0)}, X, M) = \int f(Y_{(0)}, Y_{(1)}, M | X, \theta, \psi) dY_{(1)} \propto f(Y_{(0)}, M | X, \theta, \psi) \quad (6)$$

where $f(\cdot)$ is the probability function of the involved variables, θ are the population parameters of interest, $Y_{(0)}$ are the set of observed values in the dataset, and ψ , X , and M are as previously defined. Starting from the joint density $f(Y, M | X, \theta, \psi)$ of the data Y and missingness indicator M , $Y_{(1)}$ can be integrated out to obtain $f(Y_{(0)}, M | X, \theta, \psi)$ used in the full likelihood of Eq. (6). Specifically, $\int f(Y_{(0)}, Y_{(1)}, M | X, \theta, \psi) dY_{(1)}$ can be simplified by factoring the joint distribution according to the selection model or pattern-mixture model factorizations described later in this article.

As the missingness process is typically unknown, it may be difficult to find its correct formulation. Thus, there is a risk that the full likelihood also may not be specified correctly even in cases when the data distribution is known. This is a threat to the validity of statistical inferences obtained from missing-data methods that are based on the full likelihood. In practice, the data distribution is often not known: this means there are two places in which the full likelihood may be incorrectly specified and thus two opportunities for reaching invalid statistical inferences.

In contrast, the “ignorable likelihood” does not involve modeling of the missingness process, which is given by

$$L_{ign}(\theta | Y_{(0)}, X) = \int f(Y_{(0)}, Y_{(1)}, X | \theta) dY_{(1)} \propto f(Y_{(0)} | X, \theta) \quad (7)$$

Thus, the ignorable likelihood is a simplified version of the full likelihood. Missing-data methods based on the ignorable likelihood may be used in cases when missingness is from the MCAR or MAR mechanisms. For this reason, the MCAR and MAR mechanisms are sometimes referred to as “ignorable” missingness. Missing-data methods that are based on the full likelihood should be used when missingness is from the MNAR mechanism. For this reason, the MNAR mechanism is sometimes referred to as “nonignorable” missingness.

Maximum likelihood

In general, maximum-likelihood yields estimates of population parameters that maximize the likelihood of the data. In the context of missing data and under the assumption of ignorable missingness, maximum-likelihood accomplishes this purpose via an algorithm and maximizes the ignorable likelihood given in Eq. (7). Maximum-likelihood can also be used when missingness is non-ignorable, in which case the method maximizes the full likelihood given in Eq. (6). Maximum-likelihood estimation under the assumption of ignorable missingness is sometimes referred to as full information maximum likelihood (FIML) or, in cases when the population distribution is assumed to be normal, normal-distribution-based maximum likelihood (NML). The term FIML is particularly common in the context of structural equation modeling (SEM).

When data contain missing values, maximum likelihood is most commonly implemented via the expectation maximization (EM) algorithm or variants thereof. Working with the full or ignorable likelihood, this algorithm alternates between an “E” step and an “M” step, in which the “E” step calculates various conditional expectations and the “M” step maximizes the expected likelihood of the complete data given the current estimates of those conditional expectations. Extensions or modifications of the EM algorithm are sometimes used in order to speed convergence of the algorithm or when the “M” step requires iteration. Such extensions of the EM algorithm include the ECM, ECME, AEEM, and PX-EM algorithms. Hybrid approaches such as gradient EM (Lange,

1995) or the accelerated EM algorithm (Jamshidian and Jennrich, 1993) may modify the EM algorithm by combining it with Newton-Raphson or Fisher scoring methods (Louis, 1982; Meilijson, 1989).

As a missing-data method, maximum-likelihood can provide estimates of the model parameters of interest directly when the analysis model (i.e., regression, SEM, growth curve model, etc.) is included in the formulation of the likelihood. Otherwise, maximum-likelihood can provide estimates of the mean and variance-covariance matrix of the population distribution that can be used for further statistical analyses. In such cases, the formulation of the likelihood involves the assumed population distribution (such as the normal distribution) and omits its relation to the analysis model.

The performance of the maximum-likelihood method has been extensively studied in the literature under a variety of circumstances and compared to other sophisticated missing-data methods. When the likelihood function is correctly specified, maximum-likelihood yields consistent and asymptotically normally distributed parameter estimates. The standard errors of maximum-likelihood estimates can be consistently estimated via the inverse of the corresponding observed information matrix. A disadvantage of the method is that maximum-likelihood does not supply predictions for the missing values themselves. This may be incompatible with certain statistical calculations such as residual diagnostics in regression.

Maximum-likelihood under the assumption of ignorable missingness is available as a missing-data method in several statistical software packages, such as SPSS, Mplus, SAS, and the “norm” package in R (Ported to R by Alvaro A. Novo. Original by J.L. Schafer, 2013).

Multiple imputation

Multiple imputation is a missing-data method that can be applied under both ignorable and nonignorable missingness, depending on the particular specifications involved. This missing-data method uses an algorithm to impute multiple plausible values for missing responses. Imputing several plausible values for a missing response is a strategy that reflects the inherent uncertainty associated with the missing response. The algorithm that is used in multiple imputation consists of four steps: (1) Impute missing values a total of D times according to the specified imputation model, yielding D complete datasets, (2) Conduct the statistical analyses of interest on each of the D complete datasets as if no missing values occurred, (3) Combine the set of D statistical estimates into a single set of “pooled” estimates, and (4) Obtain final inference(s) on the pooled estimates according to a set of pooling rules that are designed to correctly account for the variability associated with the pooled estimates.

The statistical assumptions that are made by multiple imputation come from the specified imputation model. The imputation model may assume ignorable missingness or allow for nonignorable missingness through modeling of the missingness mechanism. The imputation model also may make distributional assumptions: for example, continuous variables may be assumed to follow a normal distribution, nominal variables may be assumed to follow a multinomial distribution, and dichotomous variables may be assumed to follow an underlying latent normal distribution and modeled through logistic regression. The imputation procedure traditionally falls under the Bayesian framework, drawing from the predictive posterior distribution of missing responses. However, there are several other choices for creating imputations. A popular choice is predictive mean matching, in which missing values are replaced with observed responses from a pool of “donors”—subjects who are deemed to be similar to the subjects with missing responses on the completely observed variables. This method of creating imputations will always produce “permissible” values for the variable type; for example, missing responses on a discrete numeric variable will always be imputed with discrete values. Thus, predictive mean matching can be used for any variable type. However, this method has limitations: the quality of imputations is only as good as the pool of donors, which may suffer when the pool of donors is too small compared to the rate of missingness or when missingness is nonignorable. Imputations may also be created from random forest analysis or from CART, which are nonparametric methods borrowed from the statistical learning literature. Imputation models for nonignorable missingness include Heckman’s selection model (Galimard et al., 2016, 2018).

A popular multiple imputation approach is Chained-Equation Multiple Imputation, also called Multiple imputation via Chained Equations (MICE). This approach breaks from traditional MI methods in that it does not use the joint distribution of the variables to create imputations. Instead, missing values are imputed via a sequence of conditional distributions that describe how a given variable is related to the others in the dataset. Thus, this approach to multiple imputation requires the user to specify the number of imputations (D) as well as the number of iterations (t). The set of conditional distributions in Multiple imputation via Chained Equations may not necessarily correspond to a coherent joint distribution. Nevertheless, multiple imputation via chained equations can have good performance when the set of conditional distributions fit the data well, even though these may lack strict theoretical justification (Li et al., 2014; Little and Rubin, 2020). In such cases, it is recommended that the analyst checks that the imputed values are “reasonable” via diagnostic checks, perhaps checking that the distribution of imputed values is similar to that of observed values conditional on missingness propensities. There are several approaches to perform diagnostic checks, some of which are supported by statistical software (Abayomi et al., 2008; Bondarenko and Raghunathan, 2016; van Buuren and Groothuis-Oudshoorn, 2011). However, such checks can offer no guarantee regarding the true accuracy of the imputations—imputations may be far from their true values but still appear reasonable.

Special care must be taken when obtaining pooled estimates of certain statistical quantities. For example, standard errors will be underestimated if they are pooled by averaging across the D completed datasets due to the artificially increased sample size. Thus, pooling rules that correctly account for the variability of such statistical estimates have been developed. Such rules exist for standard errors, test statistics, and adjustments to the reference distribution for such test statistics and confidence intervals. In the context of regression, pooling rules have also been developed for R^2 and R^2_{adj} (Harel, 2009).

Choosing an appropriate number of imputations (and iterations, if using Multiple Imputation via Chained Equations) is an important issue for conducting multiple imputation. Standard recommendations have stated that as few as 5 or 10 imputations are sufficient to obtain reliable statistical estimates (Little and Rubin, 2020). However, statistical power and Monte Carlo error are important considerations in practical data analysis which may suggest that a greater number of imputations is optimal (Bodner, 2008; Graham et al., 2007; Horton and Lipsitz, 2001; Royston, 2004; Schafer, 1997; Von Hippel, 2009; White et al., 2011).

In such cases, the optimal number of imputations is related to the “fraction of missing information,” although this quantity may be difficult or impractical to calculate in some applications (White et al., 2011). One recommendation is to choose the number of imputations so that it is proportional to the actual rate of missingness. For example, if there is 40% missingness overall, then 40 imputations are sufficient (Bodner, 2008; Von Hippel, 2009).

Advancements in multiple imputation methods have been made for several data contexts, including for multilevel models (Enders et al., 2020; Grund et al., 2018). Multiple imputation is also available in many statistical software packages, such as SPSS, SAS, Mplus, and the mice package in R (van Buuren and Groothuis-Oudshoorn, 2011; RStudio Team, 2018). The miceadds package in R supplements the information provided by the mice package, such as pooled estimates of R^2 and R^2_{adj} according to the pooling rules in Harel (2009), Robitzsch and Grund (2021), van Buuren and Groothuis-Oudshoorn (2011).

Multiple imputation has several advantages. First, the imputation model can be specified very flexibly. This allows the missing-data method to be used for ignorable or nonignorable missingness. Under the assumption of ignorable missingness, the adoption of a nonparametric imputation model such as random forest analysis, CART, or predictive mean matching can sidestep issues arising from a misspecified imputation model. Second, multiple imputation provides a set of complete datasets which are easier to analyze. Disadvantages of multiple imputation are that proper pooling rules may still need to be developed for some multicomponent test statistics, a misspecified imputation model can lead to overestimated variances, and distribution violations of the imputation model can lead to estimated standard errors that are statistically inconsistent (Little and Rubin, 2020; Raghunathan and Rubin, 1998; Yuan et al., 2012). Furthermore, choosing an appropriate number of imputations may not always be straightforward, and a large number of imputations increases computational burden.

The performance of multiple imputation has been compared to that of maximum likelihood in many settings, including item response theory and analyses with mediation or moderation (Collins et al., 2001; Finch, 2008; McNeish, 2017; Schafer and Graham, 2002; Shin et al., 2017; Yuan et al., 2012; Zhang and Wang, 2013, 2017). However, the two methods are asymptotically equivalent under the assumption of normality and when Bayesian linear regression is used as the imputation model.

Models for nonignorable missingness

Models for nonignorable missingness are based on the full likelihood in Eq. (6) and operate by factoring the joint distribution of the data and missingness, denoted as $f(Y, M|X, \theta, \psi)$. Thus, such missing-data methods assume a distribution for the data and a distribution for missingness. A further assumption is that the statistical relationships in the missingness process itself are specified correctly—all relevant variables must be included together with any interaction or nonlinear terms that are involved in the true (but unknown) missingness process. Models for nonignorable missingness generally fall under one of two approaches, whose implementations differ across types of analysis model.

Selection models

Selection models aim to model the MNAR mechanism by factoring the joint distribution of Y and M as $f(Y|X, \theta)f(M|X, Y, \psi)$. That is, the joint distribution is factored into a distribution of the data given optional covariates and the distribution of missingness given the data and optional covariates.

A popular approach in the selection model framework for cross-sectional data is the Heckman selection model (Heckman, 1976). The Heckman selection model assumes the data are normally distributed and that $f(Y|X, \theta)$ can be modeled by regression, in which the outcome variable contains missing values. An equation for $f(M|X, Y, \theta)$ models the missingness process via regression, which is linked to the regression equation for $f(Y|X, \theta)$ through the error terms. Specifically, the residuals of these two regression equations are assumed to be correlated when missingness is nonignorable. Maximum likelihood estimation is typically used to obtain parameter estimates of the regression models. Heckman's selection model is sensitive to the assumption of normality, model misspecification, and requires that missingness is isolated to the outcome variable.

Selection models have also been developed for repeated measures. The approaches of Wu and Carroll (1988) and Diggle and Kenward (1994) assume that the analysis model is a latent growth curve model. The formulation of Wu and Carroll (1988) regresses the probability of missingness at a given time point on the intercepts and slopes of the individual growth trajectories through logistic regression equations. In this way, missingness is assumed to depend on all variables, including those after the timepoint missingness occurred. This model operates by making distributional assumptions regarding the latent intercepts and slopes for the growth trajectories. The formulation of Diggle and Kenward (1994) differs from that of Wu and Carroll (1988) in that the probability of missingness is regressed on the measured variables themselves. Missingness is allowed to depend on the current and previous measurement occasions and distributional assumptions are made regarding the repeated measures variables.

Software support for selection models under the assumption of normality and cross-sectional data are available in the “sampleSelection” R package and Mplus (Toomet and Henningsen, 2008). Software support for selection models with repeated measures are available in Mplus. This software uses the model suggested by Diggle and Kenward (1994).

Pattern-mixture models

As implied by its name, the method of pattern-mixture models assumes that each observed data pattern follows a specific model/distribution, and that the whole sample is a mixture of these patterns. This method factors the joint distribution $f(Y, M|X, \phi, \pi)$ as $f(Y|X, M, \phi)f(M|X, \pi)$, where π are parameters related to the probability of missingness patterns and ϕ are population parameters with components $\phi^{(k)}$ that correspond to the set of parameters associated with the k th missingness pattern. That is, the joint distribution is factored into: (1) a distribution of the data given optional covariates which is stratified by the missingness patterns observed, and (2) the distribution of missingness given optional covariates and probabilities of such missingness patterns. As a result of this stratification, parameters in ϕ associated with incompletely observed variables may not be identified for the corresponding groups of nonrespondents without specifying constraints.

Such an issue can be addressed by imposing certain restrictions on the values of ϕ . For example, for nonignorable missingness that depends on certain covariates (i.e., from the diffuse MNAR mechanism), a restriction may be that a subset of the parameters from the conditional distribution of the completely observed variables (say, y_1) given the incompletely observed variable (say, y_2) $\phi_{1.2}^{(k)}$ are the same across missingness patterns, i.e., $\phi_{1.2}^{(0)} = \phi_{1.2}^{(1)} = \phi_{1.2}$, where $\phi_{1.2}$ designates the set of parameters for the distribution of y_1 conditional on y_2 (Little and Wang, 1996). With a single variable containing nonignorable missing values, a common assumption is that the distributions of observed and missing responses differ by an intercept: that is, the mean of the distribution for nonrespondents is shifted from that of the distribution for respondents (as in Example 15.8 of Little and Rubin, 2020).

When two or more variables contain missing values, the number of restrictions on ϕ required for model identification depends on the number of completely observed variables p_1 relative to the number of variables with missing responses p_2 . Little and Wang (1996) discuss cases in which a pattern-mixture model may be just-identified, underidentified, or overidentified in the context of regression analysis. For overidentified pattern-mixture models, iterative methods are needed to obtain parameter estimates. However, the performance of the pattern-mixture-models suggested by Little and Wang (1996) suffers when there are more than two patterns of missingness (Little and Rubin, 2020; Tang et al., 2003).

In the context of repeated measures, pattern-mixture models may fit a growth model for each missingness pattern and place restrictions on the resulting growth parameter estimates. The approach of Hedeker and Gibbons (1997, 2006) uses a drop-out indicator as a predictor for the latent slopes and intercepts, and assumes equivalence of growth trajectories for two or more patterns of missingness—for example, between early and late drop-outs. The model also assumes that the covariance structure is the same across missingness patterns and that missingness arises from subjects who “drop out” of the study altogether. Other choices of constraints for the set of inestimable parameters in ϕ in the context of repeated measures are also possible (Enders, 2010, 2011; Demirtas and Schafer, 2003; Molenberghs et al., 1998; Thijs et al., 2002; Verbeke, 2000). Software support for pattern-mixture models is available in Mplus and SAS.

Bayesian methods

Bayesian methods for statistical inference can also be applied in the context of missing data for both ignorable and nonignorable forms of missingness. As with complete data, the method operates by drawing from the posterior distribution of the parameters of interest. Under ignorable missingness and with a chosen prior distribution $p(\theta)$ for the model parameters, the posterior distribution is $p(\theta|Y_0) \propto p(\theta) \times f(Y_0|\theta)$. In cases with a monotone missingness pattern and when a one-to-one function of θ can be found in which the corresponding parameters are *a priori* independent, the parameters of interest can be simulated from the posterior distribution directly based on the factored likelihood (Little and Rubin, 2020, Chapter 7). When these conditions do not hold, iterative methods should be used. Such iterative methods include data augmentation or Gibbs' sampler.

Data augmentation is an iterative Bayesian method that alternates between an “I” step and a “P” step to eventually produce a draw of the joint posterior distribution of missing values and the parameters of interest given the observed data (Tanner and Wong, 1987). Specifically, for the t^{th} iteration, the “I” step imputes missing values $Y_1^{(t)}$ based on a draw from the conditional density $p(Y_1|Y_0, \theta^{(t-1)})$, and the “P” step draws values of $\theta^{(t)}$ from the conditional density $p(\theta|Y_0, Y_1^{(t)})$. Once the algorithm has proceeded long enough to converge to a draw from the target joint posterior distribution, the values of θ and Y_1 are kept from the final iteration T . This process is repeated independently a total of D times to obtain D draws from the posterior distribution of θ and the posterior predictive distribution of Y_1 . Inference about θ is based on the D draws from its posterior distribution $p(\theta|Y_0)$ which may be summarized by the posterior mean and variance-covariance.

In cases when it is difficult to obtain draws from the joint posterior distribution directly—such as scenarios in which θ is multivariate and its components are not independent—Gibbs' sampler can be used. For the t^{th} iteration, the “P” step of data augmentation now involves a series of conditional distributions, in which the values of a parameter $\theta_j^{(t)}$ are drawn from conditional distributions given the current draws of other parameters $\theta_{j-1}^{(t)}, \theta_{j-2}^{(t)}, \dots$. Like data augmentation, the algorithm eventually converges to a draw from the target joint posterior distribution and should be independently run a total of D times to obtain D draws from the target distribution $p(\theta|Y_0)$.

Both data augmentation and Gibbs' sampler may not work well if the number of iterations is too small. However, several methods have been developed to assess or improve the convergence of the sequence of draws (Gelman et al., 2013; Gelman and Rubin, 1992; Geyer, 1992; Liu and Rubin, 1996, 2002). The usual limitations of Bayesian methodology are also applicable to their use as missing-data methods—such as when $f(Y_0|\theta)$ is misspecified as normal for convenience (see e.g., Yuan and Gomer,

2021) or the potential for a prior to overwhelm the information provided from the data, particularly with small sample sizes (Gelman, 2006; McNeish, 2016; Smid and Winter, 2020; Smid et al., 2020). Software support for Bayesian missing-data methods includes Mplus, SAS, and the “brms” R package (Bürkner, 2017).

Statistical learning methods

CART

A relatively new missing-data method adapted from the statistical learning framework is Classification and Regression Trees (CART) (Breiman et al., 1984). This nonparametric approach relies on inverse probability weights to operate (McArdle, 2013). The implementation of CART typically assumes missing values are MCAR, although a few authors claim that the particular relationship between missingness and the outcome variable has a greater impact on CART’s performance than the missing-data mechanism (Ding and Simonoff, 2006; Saar-Tsechansky and Provost, 2007).

As a missing-data method, CART aims to find the best cut-off point on values of a variable that will separate or “split” the group of respondents from the group of nonrespondents. However, this process may need to be repeated several times in order to achieve good separation of the groups. This yields a tree-like structure in which each “node” corresponds to a different variable or different cut-point on a variable that was considered in a previous step. The end result classifies subjects into several subgroups; each subgroup is composed entirely of respondents or nonrespondents (to the fullest extent possible). The predicted probabilities associated with being classified into the subgroups are inverted and then used as sampling weights. CART may be used to create a single artificially complete dataset or paired with multiple imputation, although the results obtained from a single dataset may be problematic (van Buuren, 2018).

A potential drawback of CART is that it is prone to overfitting—in other words, it produces a tree structure that optimizes separation in the current dataset at the cost of generalizing well to others (Crawford, 1989). This shortcoming can be addressed by preventing the tree from “growing” too large through the use of stopping criteria, in which the final subgroups are required to have a minimum sample size. A different strategy is to “prune” a large tree structure via cost-complexity pruning, in which the misclassification rate of respondents vs. nonrespondents is balanced against tree complexity. At the time of this writing, there is still no consensus regarding which of the two strategies is best, although it is a subject of ongoing debate (Berk, 2009; Hayes et al., 2015; Louppe, 2014).

Random forest analysis

Random forest analysis is another statistical learning approach that has been adapted for use as a missing-data method (Breiman, 2001). Random forest analysis operates in a similar way to CART and under the same assumptions of missingness. However, while CART creates a single tree structure, random forest analysis is an ensemble method that creates a “forest” of many such regressions trees and each is based on a bootstrapped sample of the data. Missing values can be imputed based on the values of complete cases and the resulting sampling weights to yield a single complete dataset (Breiman, 2003). However, both random forest analysis and CART can be used as an imputation model in multiple imputation and are supported in the “mice” R package (van Buuren and Groothuis-Oudshoorn, 2011).

Other methods

Several other statistical learning approaches have also been adapted for use as missing-data methods. These include classification methods such as k-Nearest-Neighbors (kNN) (Batista and Monard, 2001; Huang and Lee, 2004; Wasito and Mirkin, 2006; Zhang, 2011, 2012), k-means or clustering methods (Fujikawa and Ho, 2002; Li et al., 2004; Patil et al., 2010; Zhang et al., 2008), artificial neural networks (Nelwamondo et al., 2007; Setiawan et al., 2008), and support vector machines (SVM) (Mallinson and Gamerman, 2003; Sivapriya et al., 2012; Suykens et al., 2002). These missing-data methods are often used as single-imputation strategies, although it is possible to use them in a multiple imputation context (Sivapriya et al., 2012). Note that even these advanced methods will lead to underestimation of variances and covariances if they are used as single-imputation strategies, and that articles often evaluate the performance of these methods in terms of prediction accuracy of the missing values themselves rather than on the quality of parameter estimates.

Threats to proper handling of missing values and possible remedies

Violated distribution conditions

Note that likelihood-based missing-data methods such as ML and MI only produce valid statistical inferences when the likelihood function is correctly specified. When these inferences are based on the ignorable likelihood of Eq. (7), the missingness mechanism must be MAR or MCAR and the distribution of the data must match the distribution assumed by the likelihood. It is a well-known fact that real data are seldom normally distributed, as has been repeatedly pointed out in meta analyses (Blanca et al., 2013; Cain et al., 2017; Micceri, 1989). Nevertheless, most researchers assume normality when implementing ML and MI (NML, NMI) in data analysis due to a lack of alternative distributions with multivariate data. Practical distribution violations cause NML to become “pseudo ML,” and NMI to become “pseudo MI.” FIML (full information maximum likelihood) is just a nickname for a widely used methodology unless the population distribution is truly multivariate normal. With skewed and kurtotic data, the NML

missing-data method can still yield consistent parameter estimates when the population distribution satisfies certain conditions (Yuan and Bentler, 2010). Consistent standard errors of the NML estimates can also be obtained via sandwich-type covariance matrices (Yuan, 2009b). However, such conditions are hard to verify in practice. It is likely that NMI also yields consistent parameter estimates under the same conditions specified in Yuan and Bentler (2010); see Yuan et al. (2012) for a discussion of the rationale. However, NMI cannot generate consistent standard errors for the NMI estimates because nonnormally distributed data are replaced with normally distributed ones (Yuan et al., 2012).

Tests of normality

Methods exist to test multivariate normality with complete data (e.g., Liang and Ng, 2009). However, with incomplete data, and especially under the MAR and MNAR mechanisms, the distribution of the observed data can be skewed and kurtotic even when the underlying population distribution is multivariate normal (Jamshidian and Yuan, 2014). This makes it hard to justify the NML and NMI methods in operation, because they depend on conditions that are not verifiable. Nevertheless, it is still desirable to acknowledge possible distribution violations in practice for these methods to be of a scientific nature. In this direction, the measure of multivariate kurtosis (Mardia, 1970) has been extended to incomplete data (Yuan et al., 2004) which is also available in statistical software (Bentler, 2006).

Robust methods

There are developments pertaining to NML methods that yield consistent standard errors and rescaled test statistics for mean and covariance structural models with incomplete data (Yuan, 2009b; Yuan and Bentler, 2000). With real data whose distributions possess heavier tails, robust methods have been shown to yield much more reliable results in statistical modeling (Wilcox, 2017; Yuan and Gomer, 2021). Robust methods via the *t*-distribution-based maximum likelihood (TML) have been developed by Little (1988a). Yuan et al. (2016) developed more general robust methods via estimating equations, which not only yield robust estimates but also provide an approach to obtain consistent standard errors for the robust parameter estimates as well as more reliable test statistics in mean and covariance structure analysis (Tong et al., 2014).

Comparing distribution violations against non-ignorable missingness, results show that data from a heavier-tailed distribution can cause much more damage to the NML missing-data method than MNAR missingness does (Yuan and Savalei, 2014), especially when data are contaminated. Monte Carlo results also show that robust methods yield close to optimal results even when the population distribution is normally distributed (Yuan et al., 2015).

The use of auxiliary variables

Datasets in the social sciences often contain a large number of variables. Some of these variables may not be of particular research interest and are thus excluded from statistical analysis. However, as has been presented earlier, the MAR mechanism is defined via conditional distributions; and omitting such observed variables from missing data procedures is equivalent to failing to observe these variables in the first place. This presents a scenario for the introduction of auxiliary variables: variables that are related to missingness but are not involved in substantive analysis. Proper use of auxiliary variables can lead to improved analysis and more reliable results, and it is best carried out using a two-stage procedure. In the first stage, auxiliary variables are included in the estimation of saturated models (means and variances-covariances). The second-stage analysis is conducted using the estimated means and variances-covariances corresponding to the substantive variables from the first stage. This two-stage procedure is easy to implement both under NML and under robust methods (Savalei and Bentler, 2009; Yuan and Lu, 2008; Yuan and Zhang, 2012). Software support for this procedure is also available in the context of structural equation modeling in the R package “rsem” (Yuan and Zhang, 2020).

There exist recommendations to include a large number of auxiliary variables, or, in other words, “the more the better” (e.g., the inclusive strategy in Collins et al., 2001). However, these recommendations are based on Monte Carlo studies with normally distributed populations. In practice, including auxiliary variables whose distributions contain heavier tails can make the resulting NML estimates less accurate, even though the missing data mechanism becomes MAR (Yuan and Savalei, 2014). There are methods for selecting auxiliary variables based on their effects on the existing estimates (Yuan, 2009a). However, regardless of how auxiliary variables are chosen, there is no existing procedure that allows one to conclude that the resulting mechanism actually becomes MAR. This is because the MNAR mechanism is on a continuum that is characterized by the extent to which missingness is influenced by the missing values themselves. Including auxiliary variables may nevertheless bring the resulting estimates closer to the population values, but it is hard to claim that the missingness in any given analysis of incomplete data is ignorable.

Missing-data mechanism uncertainty

In practice, the underlying missing-data mechanism is often unknown. This can be problematic if the chosen missing-data method assumes ignorable missingness while the true missingness process is actually nonignorable. Missing-data methods designed for nonignorable missingness do so by making untestable assumptions, so their performance may also be uncertain in practice. For this reason, it is better not to use these methods when missingness is ignorable. An incorrect diagnosis of the missing-data mechanism has consequences for the validity of statistical inferences (Collins et al., 2001; Gomer and Yuan, 2021; Little and Rubin, 2020; Smuk et al., 2017). Thus, uncertainty regarding the true missing-data mechanism is an important issue in missing data analysis.

Tests of the missingness mechanism

One way to approach the issue of missing-data mechanism uncertainty is through statistical methods that seek to identify or test the missing-data mechanism directly. Although this is an appealing idea, such methods fall short of achieving this goal. A further complication is that such tests of the missingness mechanism are often misused in practice.

Tests of MCAR

Tests of the MCAR mechanism operate by testing the homogeneity of means and variances-covariances across missingness patterns. Under MCAR, the estimated means and variances-covariances from each of these patterns are unbiased for the overall population mean and variance-covariances, so equality of means and variance-covariances across missingness patterns is a necessary condition for the MCAR mechanism (Rhoads, 2012; Yuan et al., 2018). Thus, rejecting the null hypothesis suggests that the true missing-data mechanism is not MCAR. Perhaps the most famous test of the MCAR mechanism is Little's test of MCAR (Little, 1988b), which tests equality of means and variances-covariances through a likelihood ratio test statistic. A modification of this test using generalized least squares was developed by Kim and Bentler (2002). Other developments for tests of the MCAR mechanism have been made for scenarios with small samples sizes within missingness patterns (Jamshidian and Jalal, 2010), categorical data (Park and Davis, 1993), repeated measures (Park et al., 1993), and generalized estimating equations (Chen and Little, 1999; Qu and Song, 2002).

Unfortunately, homogeneity of means and variances-covariances is not a sufficient condition for MCAR (Rhoads, 2012); and indeed, there are many conditions in which homogeneity of means and variances-covariances hold even under the MAR and MNAR mechanisms and under multivariate normality (Yuan et al., 2018). For this reason, tests of MCAR can only provide credible evidence against the MCAR mechanism, rather than in favor. Nevertheless, misuse of these tests is common.

Tests of MAR

Tests of the MAR mechanism may test the MAR assumption against a specific MNAR alternative. Such tests may take the form of a model-based likelihood ratio test (Zhou et al., 1999), which assume that the full likelihood for nonignorable missingness has been correctly specified; or, equivalently, that the model of the missingness process is correct (Rhoads, 2012). Unfortunately, it has been shown that for every MNAR model, an MAR model can be found that has equivalent fit to the data (Molenberghs et al., 2008). This means that such likelihood ratio tests can only provide evidence against the MAR mechanism in cases when there is strong *a priori* belief in the MNAR model (Molenberghs et al., 2008).

Several methods claim to test the MAR mechanism or ignorable missingness against a more general alternative. However, such tests actually evaluate a weaker condition, such as that data are Observed At Random (OAR), $P(Y_{(0)}|R) = P(Y_{(0)})$ (Rubin, 1976), which may not even be a necessary condition for ignorable missingness in some analysis model contexts (Rhoads, 2012). Tests of the MAR mechanism that actually evaluate OAR include those of Katz (2006) and Sherman (2000). Tests may instead evaluate OAR conditional on a fully observed vector of covariates (Park and Davis, 1993) or OAR conditional on baseline responses in the context of repeated measures (Donaldson and Moinpour, 2005). Thus, the stated goal of such tests differs from what is actually evaluated. This problem is likely to exist for any test of the MAR mechanism that claims to evaluate MAR against a general alternative (Rhoads, 2012).

Sensitivity analysis

Sensitivity analysis is a statistical approach to assess the sensitivity of results to the missing-data mechanism or to the departure from the MAR assumption (Little and Rubin, 2020). Typically, sensitivity analysis involves explicit modeling of the missingness process, in which the user specifies a series of values for a "sensitivity parameter" that controls the degree of departure from the MAR mechanism. By examining the fluctuations in final statistical inferences across distinct values of the sensitivity parameter, researchers can assess the stability of their statistical inferences with respect to missing-data mechanism uncertainty. Sensitivity analysis is a widely studied topic, although many developments have come from biostatistics or related fields.

In general, sensitivity analysis consists of three steps: (1) choosing a range of values for the sensitivity parameter; (2) carrying out statistical analyses for each of the chosen values; and (3) incorporating the results to obtain final statistical inferences. With respect to the last step, there are three popular approaches used for decision-making regarding such final statistical inferences (The Prevention and Treatment of Missing Data in Clinical Trials, 2010). One approach is to report the upper and lower bounds of parameter estimates obtained across a plausible range of values for the sensitivity parameter or the upper and lower bounds of such estimates (Leacy et al., 2017; Molenberghs and Kenward, 2007). Another approach is to conduct statistical analyses under the assumption of MAR and determine which values of the sensitivity parameter would overturn the current set of statistical conclusions (Campbell et al., 2011; Liublinka and Rubin, 2012; Yan et al., 2009). The third approach is to obtain a summarized version of statistical inference by averaging over values of the sensitivity parameter according to a set of principles (Carpenter et al., 2007; Siddique et al., 2012).

Several methods for conducting sensitivity analysis are multiple-imputation-based approaches. One of the most popular methods is "multiple imputation with delta adjustment," which adds a fixed offset to the linear predictor of the imputation model to shift the mean of the distribution of nonrespondents compared to that of respondents (Carpenter and Kenward, 2008; Carpenter et al., 2013; Leacy et al., 2017; Rubin, 1977; van Buuren et al., 1999). An alternative is to switch the order of operations such that the

fixed offset is added to the imputations themselves instead of the linear predictor (Giusti and Little, 2011; Little and Rubin, 2020; van Buuren et al., 1999). Many other multiple-imputation-based approaches exist (Carpenter et al., 2007; Hsu et al., 2020).

Reference-based approaches to sensitivity analysis are sometimes used in the context of clinical trials, in which statistical analyses are conducted under various assumptions regarding missing values or protocol deviations. One such assumption might be that the mean drug effectiveness for drop-outs from the treatment group is equal to the observed mean of the placebo group at unobserved time points. Another assumption might be that drug effectiveness for drop-outs at unobserved time points is equal to the marginal mean at the last observed time-point. Several other assumptions may also be considered (Ayele et al., 2014; Carpenter et al., 2013; Cro et al., 2016; Smuk et al., 2017).

Tipping-point analysis is an approach to sensitivity analysis that finds the values of missing responses required to overturn the current statistical conclusion (Yan et al., 2009). The plausibility of these values can then be evaluated by the researcher or agency tasked with reviewing the findings. Statistical results obtained under this strategy have an intuitive interpretation. Graphic displays have also been developed to visualize such tipping-point analyses (Campbell et al., 2011; Liublinska and Rubin, 2012).

Conclusion and recommendations

There are many statistical procedures available to the data analyst when handling missing values. In our opinion, the best practice is to use substantive judgment to diagnose the missing-data mechanism and use a likelihood-based missing-data method that is appropriate for such a missingness mechanism or a robust method when the distribution of the observed variables contains heavy tails. The particular circumstances of the data (such as the modeling context) may naturally suggest one missing-data method over another. If there is uncertainty regarding the underlying missing-data mechanism, we recommend conducting sensitivity analysis to determine the extent to which statistical conclusions are influenced by the assumption of missingness mechanism. Tipping-point analysis or approaches that provide a single summarized version of statistical inferences may be the most straightforward to interpret, and thus may be a suitable choice for work that will be published in a journal where such analyses are uncommon.

We caution against the use of single imputation missing-data methods in practice. Note that several sophisticated missing-data methods borrowed from the statistical learning literature fall under this category, unless they are instead used as the imputation model in multiple imputation. We also caution against tests of the missingness mechanism, but such tests can be conducted responsibly if their results are interpreted carefully in the context of their limitations. In practice, it might be the case that the underlying population distribution is unknown and all three types of missing-data mechanisms (MCAR, MAR, MNAR) co-exist within the same pattern of missingness. In such cases, methods based on multivariate normality or that assume a specific missing-data mechanism may not perform well (Gomer, 2019). Caution should be used in the application of existing missing-data methods whenever such conditions are suspected.

References

- Abayomi, K., Gelman, A., Levy, M., 2008. Diagnostics for multivariate imputations. *J. Roy. Stat. Soc. C* 57 (3), 273–291.
- Ayele, B.T., Lipkovich, I., Molenberghs, G., Mallinckrodt, C.H., 2014. A multiple-imputation-based approach to sensitivity analyses and effectiveness assessments in longitudinal clinical trials. *J. Biopharm. Stat.* 24 (2), 211–228.
- Batista, G., Monard, M.C., 2001. Proceedings of the Argentine Symposium on Artificial Intelligence. A Study of K-Nearest Neighbor as a Model-Based Method to Treat Missing Data, vol. 30, pp. 1–9.
- Bentler, P., 2006. EQS6 Structural Equations Program Manual. Multivariate Software, Encino, CA.
- Berk, R., 2009. Statistical Learning From a Regression Perspective. Springer, New York.
- Blanca, M., Arnau, J., López-Montiel, D., Bono, R., Bendayan, R., 2013. Skewness and kurtosis in real data samples. *Methodology* 9 (2), 78–84.
- Bodner, T., 2008. What improves with increased missing data imputations? *Struct. Equ. Model.* 15, 651–675.
- Bondarenko, I., Raghunathan, T., 2016. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. *Stat. Med.* 35 (17), 3007–3020.
- Breiman, L., Friedman, J., Stone, C., Olshen, R., 1984. Classification and Regression Trees. In: The Wadsworth Statistics Probability Series, vol. 19. Wadsworth, Belmont, CA.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
- Breiman, L., 2003. Manual for Setting Up, Using, and Understanding Random Forests V4. https://www.stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf.
- Bürkner, P.-C., 2017. BRMS: an R package for Bayesian multilevel models using Stan. *J. Stat. Software* 80 (1), 1–28.
- Cain, M., Zhang, Z., Yuan, K.H., 2017. Univariate and multivariate skewness and kurtosis for measuring nonnormality: prevalence, influence and estimation. *Behav. Res. Methods* 49 (5), 1716–1735.
- Campbell, G., Pennelo, G., Yue, L., 2011. Missing data in the regulation of medical devices. *J. Biopharm. Stat.* 21 (2), 180–195.
- Carpenter, J.R., Kenward, M.G., 2008. Missing Data in Clinical Trials—A Practical Guide. NIFH, Birmingham.
- Carpenter, J.R., Kenward, M.G., White, I.R., 2007. Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Stat. Methods Med. Res.* 16 (3), 259–275.
- Carpenter, J.R., Roger, J.H., Kenward, M.G., 2013. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *J. Biopharm. Stat.* 23 (6), 1352–1371.
- Chen, H., Little, R., 1999. A test of missing completely at random for generalized estimating equations with missing data. *Biometrika* 86 (1), 1–13.
- Collins, L., Schafer, J., Kam, C., 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol. Methods* 6 (4), 330–351.
- Crawford, S., 1989. Extensions to the CART algorithm. *Int. J. Man Mach. Stud.* 31 (2), 197–217.
- Cro, S., Morris, T.P., Kenward, M.G., Carpenter, J.R., 2016. Reference-based sensitivity analysis via multiple imputation for longitudinal trials with protocol deviation. *Stata J.* 16 (2), 443–463.

- Czajka, J., Hirabayashi, S., Little, R., Rubin, D., 1992. Projecting from advance data using propensity modeling: an application to income and tax statistics. *J. Bus. Econ. Stat.* 10, 117–132.
- Demirtas, H., Schafer, J.L., 2003. On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Stat. Med.* 22 (16), 2553–2575.
- Diggle, P., Kenward, M.G., 1994. Informative drop-out in longitudinal data analysis. *J. Roy. Stat. Soc. C* 43 (1), 49–73.
- Ding, Y., Simonoff, J.S., 2006. An Investigation of Missing Data Methods for Classification Trees. NYU Working Paper No. SOR-2006-3.
- Donaldson, G., Moinspour, C., 2005. Learning to live with missing quality-of-life data in advanced-stage disease trials. *J. Clin. Oncol.* 23 (30), 7380–7384.
- Enders, C., Du, H., Keller, B., 2020. A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and nonlinear terms. *Psychol. Methods* 25 (1), 88–112.
- Enders, C., 2010. *Applied Missing Data Analysis*. Guilford press.
- Enders, C.K., 2011. Missing not at random models for latent growth curve analyses. *Psychol. Methods* 16 (1), 1–16.
- Finch, H., 2008. Estimation of item response theory parameters in the presence of missing data. *J. Educ. Meas.* 45 (3), 225–245.
- Fujikawa, Y., Ho, T., 2002. Cluster-Based Algorithms for Dealing With Missing Values, Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, pp. 549–554.
- Galimard, J., Chevret, S., Protopopescu, C., Resche-Rigon, M., 2016. A multiple imputation approach for MNAR mechanisms compatible with Heckman's model. *Stat. Med.* 35 (17), 2907–2920.
- Galimard, J.-E., Chevret, S., Curis, E., Resche-Rigon, M., 2018. Heckman imputation models for binary or continuous MNAR outcomes and MAR predictors. *BMC Med. Res. Methodol.* 18 (1), 1–13.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7 (4), 457–472.
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, B., 2013. *Bayesian Data Analysis*, third ed. CRC Press.
- Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* 1 (3), 515–534.
- Geyer, C.J., 1992. Practical Markov Chain Monte Carlo. *Statistical Science*, pp. 473–483.
- Giusti, C., Little, R.J.A., 2011. An analysis of nonignorable nonresponse to income in a survey with a rotating panel design. *J. Off. Stat.* 27 (2), 211–229.
- Gomer, B., Yuan, K.-H., 2021. Subtypes of the missing not at random missing data mechanism. *Psychol. Methods* 26 (5), 559–598.
- Gomer, B., 2019. MCAR, MAR, and MNAR in the same dataset: a realistic evaluation of methods for handling missing data. *Multivariate Behav. Res.* 54 (1), 153.
- Graham, J.W., Olchowski, A.E., Gilreath, T., 2007. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev. Sci.* 8 (3), 206–213.
- Grund, S., Lüdtke, O., Robitzsch, A., 2018. Multiple imputation of missing data for multilevel models: simulations and recommendations. *Organ. Res. Methods* 21 (1), 111–149.
- Harel, O., 2009. The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. *J. Appl. Stat.* 36 (10), 1109–1118.
- Hayes, T., Usami, S., Jacobucci, R., McArdle, J.J., 2015. Using classification and regression trees (CART) and random forests to analyze attrition: results from two simulations. *Psychol. Aging* 30 (4), 911–929.
- Heckman, J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Ann. Econ. Soc. Meas.* 5 (NBER), 475–492.
- Hedeker, D., Gibbons, R.D., 1997. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychol. Methods* 2 (1), 64–78.
- Hedeker, D., Gibbons, R.D., 2006. *Longitudinal Data Analysis*. John Wiley & Sons.
- Horton, N.J., Lipsitz, S.R., 2001. Multiple imputation in practice: comparison of software packages for regression models with missing variables. *Am. Statistician* 55, 244–254.
- Hsu, C.-H., He, Y., Hu, C., Zhou, W., 2020. A multiple imputation-based sensitivity analysis approach for data subject to missing not at random. *Stat. Med.* 39 (26), 3756–3771.
- Huang, C.-C., Lee, H.-M., 2004. A grey-based nearest neighbor approach for missing attribute value prediction. *Appl. Intell.* 20 (3), 239–252.
- Jamshidian, M., Jalal, S., 2010. Tests of homoscedasticity, normality, and missing completely at random for incomplete multivariate data. *Psychometrika* 75 (4), 649–674.
- Jamshidian, M., Jennrich, R.I., 1993. Conjugate gradient acceleration of the EM algorithm. *J. Am. Stat. Assoc.* 88 (421), 221–228.
- Jamshidian, M., Yuan, K.-H., 2014. Examining missing data mechanisms via homogeneity of parameters, homogeneity of distributions, and multivariate normality. *Wiley Interdiscip. Rev.* 6, 56–73.
- Katz, M., 2006. *Study Design and Statistical Analysis: A Practical Guide for Clinicians*. Cambridge University Press, Cambridge.
- Kim, K., Bentler, P., 2002. Tests of homogeneity of means and covariance matrices for multivariate incomplete data. *Psychometrika* 67 (4), 609–623.
- Kish, L., 1995. Methods for design effects. *J. Off. Stat.* 11 (1), 55–77.
- Lange, K., 1995. A gradient algorithm locally equivalent to the EM algorithm. *J. Roy. Stat. Soc. B* 57 (2), 425–437.
- Leacy, F.P., Floyd, S., Yates, T.A., White, I.R., 2017. Analyses of sensitivity to the missing-at-random assumption using multiple imputation with delta adjustment: application to a tuberculosis/HIV prevalence survey with incomplete HIV-status data. *Am. J. Epidemiol.* 185 (4), 304–315.
- Li, D., Deogun, J., Spaulding, W., Shuart, B., 2004. Towards Missing Data Imputation: A Study of Fuzzy K-Means Clustering Method, International Conference on Rough Sets and Current Trends in Computing. Springer, pp. 573–579.
- Li, F., Baccini, M., Mealli, F., Zell, E.R., Frangakis, C.E., Rubin, D.B., 2014. Multiple imputation by ordered monotone blocks with application to the anthrax vaccine research program. *J. Comput. Graph Stat.* 23 (3), 877–892.
- Liang, J., Ng, K., 2009. A multivariate normal plot to detect nonnormality. *J. Comput. Graph Stat.* 18 (1), 52–72.
- Little, R., Rubin, D., 2002. *Statistical Analysis With Missing Data*, second ed. John Wiley and Sons.
- Little, R., Rubin, D., 2020. *Statistical Analysis With Missing Data*, third ed. John Wiley and Sons.
- Little, R.J.A., Wang, Y., 1996. Pattern-mixture models for multivariate incomplete data with covariates. *Biometrics* 52 (1), 98–111.
- Little, R., 1986. Survey nonresponse adjustments. *Int. Stat. Rev.* 54, 139–157.
- Little, R., 1988a. Robust estimation of the mean and covariance matrix from data with missing values. *Appl. Stat.* 37, 23–38.
- Little, R.J., 1988b. A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* 83 (404), 1198–1202.
- Liu, C., Rubin, D.B., 1996. Markov-normal analysis of iterative simulations before their convergence. *J. Econom.* 75 (1), 69–78.
- Liu, C., Rubin, D.B., 2002. Model-based analysis to improve the performance of iterative simulations. *Stat. Sin.* 12, 751–767.
- Liubinska, V., Rubin, D.B., 2012. Enhanced Tipping-Point Displays. JSM Proceedings, Section on Survey Research Methods, San Diego, CA, 3861–3868.
- Louis, T.A., 1982. Finding the observed information matrix when using the EM algorithm. *J. Roy. Stat. Soc. B* 44 (2), 226–233.
- Louppe, G., 2014. Understanding Random Forests: From Theory to Practice arXiv preprint arXiv:1407.7502.
- Mallinson, H., Gammerman, A., 2003. *Imputation Using Support Vector Machines*. Department of Computer Science. Royal Holloway, University of London, Egham, UK.
- Mardia, K., 1970. Measure of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530.
- Marsh, H.W., 1998. Pairwise deletion for missing data in structural equation models: nonpositive definite matrices, parameter estimates, goodness of fit, and adjusted sample sizes. *Struct. Equ. Model.* 5 (1), 22–36.
- McArdle, J., 2013. Dealing with longitudinal attrition using logistic regression and decision tree analyses. In: McArdle, J., Ritschard, J. (Eds.), *Contemporary Issues in Exploratory Data Mining in the Behavioral Sciences*. Routledge, New York, NY, pp. 282–311.
- McNeish, D., 2016. On using Bayesian methods to address small sample problems. *Struct. Equ. Model.* 23 (5), 750–773.
- McNeish, D., 2017. Missing data methods for arbitrary missingness with small samples. *J. Appl. Stat.* 44 (1), 24–39.
- Meilijson, I., 1989. A fast improvement to the EM algorithm on its own terms. *J. Roy. Stat. Soc. B* 51 (1), 127–138.
- Micceri, T., 1989. The unicorn, the normal curve, and other improbable creatures. *Psychol. Bull.* 105, 156–166.
- Molenberghs, G., Kenward, M., 2007. *Missing Data in Clinical Studies*, vol. 61. John Wiley and Sons.
- Molenberghs, G., Michiels, B., Kenward, M.G., Diggle, P.J., 1998. Monotone missing data and pattern-mixture models. *Stat. Neerl.* 52 (2), 153–161.

- Molenberghs, G., Beunckens, C., Sotito, C., Kenward, M.G., 2008. Every missingness not at random model has a missingness at random counterpart with equal fit. *J. Roy. Stat. Soc. B* 70 (2), 371–388.
- Nelwamondo, F.V., Mohamed, S., Marwala, T., 2007. Missing data: a comparison of neural network and expectation maximization techniques. *Curr. Sci.* 1514–1521.
- Park, T., Davis, C.S., 1993. A test of the missing data mechanism for repeated categorical data. *Biometrics* 49 (2), 631–638.
- Park, T., Lee, S., Woolson, R.F., 1993. A test of the missing data mechanism for repeated measures data. *Commun. Stat. Theor. Methods* 22 (10), 2813–2829.
- Patil, B.M., Joshi, R.C., Toshniwal, D., 2010. Missing Value Imputation Based on K-Mean Clustering With Weighted Distance. *International Conference on Contemporary Computing*. Springer, pp. 600–609.
- Peugh, J.L., Enders, C.K., 2004. Missing data in educational research: a review of reporting practices and suggestions for improvement. *Rev. Educ. Res.* 74 (4), 525–556.
- Ported to R by Alvaro A. Novo Original by J.L. Schafer, 2013. Norm: Analysis of Multivariate Normal Datasets With Missing Values. R package version 1.0-9.5. <https://CRAN.R-project.org/package=norm>.
- Potthoff, R., Woodbury, M., Manton, K., 1992. "Equivalent sample size" and "equivalent degrees of freedom" refinements for inference using survey weights under superpopulation models. *J. Am. Stat. Assoc.* 87, 383–396.
- Qu, A., Song, P., 2002. Testing ignorable missingness in estimating equation approaches for longitudinal data. *Biometrika* 89 (4), 841–850.
- Raghunathan, T.E., Rubin, D.B., 1998. Proceedings of the Silver Jubilee Meeting of the Statistical Society of Canada. Roles for Bayesian Techniques in Survey Sampling, vol. 51, p. 55.
- Rhoads, C.H., 2012. Problems with tests of the missingness mechanism in quantitative policy studies. *Stat. Polit. Pol.* 3 (1).
- Robitzsch, A., Grund, S., 2021. Miceadds: Some Additional Multiple Imputation Functions, Especially for "Mice". R Package Version 3.11-6. <https://CRAN.R-project.org/package=miceadds>.
- Royston, P., 2004. Multiple imputation of missing values. *Stata J.* 4, 227–241.
- RStudio Team, 2018. RStudio. Integrated Development Environment for R, RStudio, Inc., Boston, MA. <http://www.rstudio.com/>.
- Rubin, D., 1976. Inference and missing data. *Biometrika* 63, 581–592.
- Rubin, D., 1977. Formalizing subjective notions about the effect of nonrespondents in sample surveys. *J. Am. Stat. Assoc.* 72, 538–543.
- Rubin, D., 1985. The use of propensity scores in applied Bayesian inference. In: Bernardo, J., De Groot, M., Lindley, D., Smith, A. (Eds.), *Bayesian Statistics*, vol. 2, pp. 463–472. North-Holland, Amsterdam.
- Saar-Tschansk, M., Provost, F., 2007. Handling missing values when applying classification models. *J. Mach. Learn. Res.* 8, 1625–1657.
- Savalei, V., Bentler, P.M., 2009. A two-stage approach to missing data: theory and application to auxiliary variables. *Struct. Equ. Model.* 16 (3), 477–497.
- Schafer, J., Graham, J., 2002. Missing data: our view of the state of the art. *Psychol. Methods* 7 (2), 147–177.
- Schafer, J., 1997. *Analysis of Incomplete Multivariate Data*. CRC Press.
- Setiawan, N.A., Venkatachalam, P.A., Hani, A.F.M., 2008. Missing attribute value prediction based on artificial neural network and rough set theory. In: 2008 International Conference on BioMedical Engineering and Informatics, vol. 1. IEEE, pp. 306–310.
- Sherman, R.P., 2000. Tests of certain types of ignorable nonresponse in surveys subject to item nonresponse or attrition. *Am. J. Polit. Sci.* 44 (2), 362–374.
- Shin, T., Davison, M.L., Long, J.D., 2017. Maximum likelihood versus multiple imputation for missing data in small longitudinal samples with nonnormality. *Psychol. Methods* 22 (3), 426–449.
- Siddique, J., Harel, O., Crespi, C., 2012. Addressing missing data mechanism uncertainty using multiple-model multiple imputation: application to a longitudinal clinical trial. *Ann. Appl. Stat.* 6 (4), 1814–1837.
- Sivapriya, T., Kamal, A., Thavavel, V., 2012. Imputation and classification of missing data using least square support vector machines—a new approach in dementia diagnosis. *Int. J. Adv. Res. Artif. Intell.* 1 (4), 29–33.
- Smid, S., Winter, S., 2020. Dangers of the defaults: a tutorial on the impact of default priors when using Bayesian SEM with small samples. *Front. Psychol.* 11.
- Smid, S., Depaoli, S., Van De Schoot, R., 2020. Predicting a distal outcome variable from a latent growth model: ML versus Bayesian estimation. *Struct. Equ. Model.* 27 (2), 169–191.
- Smuk, M., Carpenter, J.R., Morris, T.P., 2017. What impact do assumptions about missing data have on conclusions? A practical sensitivity analysis for a cancer survival registry. *BMC Med. Res. Methodol.* 17 (1), 1–7.
- Stapleton, L., 2002. The incorporation of sample weights into multilevel structural equation models. *Struct. Equ. Model.* 9, 475–502.
- Suykens, J.A., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J.P., 2002. *Least Squares Support Vector Machines*. World Scientific.
- Tang, G., Little, R., Raghunathan, T., 2003. Analysis of multivariate missing data with nonignorable nonresponse. *Biometrika* 90 (4), 747–764.
- Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. *J. Am. Stat. Assoc.* 82 (398), 528–540.
- The Prevention and Treatment of Missing Data in Clinical Trials, 2010. *Principles and Methods of Sensitivity Analyses*. National Academies Press (US).
- Thijs, H., Molenberghs, G., Michiels, B., Verbeke, G., Curran, D., 2002. Strategies to fit pattern-mixture models. *Biostatistics* 3 (2), 245–265.
- Tong, X., Zhang, Z., Yuan, K.-H., 2014. Evaluation of test statistics for robust structural equation modeling with nonnormal missing data. *Struct. Equ. Model.* 21, 553–565.
- Toomet, O., Henningsen, A., 2008. Sample selection models in R: package sample selection. *J. Stat. Software* 27 (7). <https://www.jstatsoft.org/v27/i07/>.
- van Buuren, S., Groothuis-Oudshoorn, K., 2011. Mice: multivariate imputation by chained equations in R. *J. Stat. Software* 1–68.
- van Buuren, S., Boshuizen, H., Knook, D., 1999. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat. Med.* 18 (6), 681–694.
- van Buuren, S., 2018. *Flexible Imputation of Missing Data*. CRC Press.
- Verbeke, G., 2000. *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics, New York.
- Von Hippel, P., 2009. How to impute squares, interactions, and other transformed variables. *Socio. Methodol.* 39, 265–291.
- Wasito, I., Mirkin, B., 2006. Nearest neighbors in least-squares data imputation algorithms with different missing patterns. *Comput. Stat. Data Anal.* 50 (4), 926–949.
- White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: issues and guidance for practice. *Stat. Med.* 30 (4), 377–399.
- Wilcox, R., 2017. *Introduction to Robust Estimation and Hypothesis Testing*, fourth ed. Academic Press.
- Wothke, W., 1993. Nonpositive Definite Matrices in Structural Modeling, vol. 154. Sage Focus Editions, p. 256.
- Wu, M.C., Carroll, R.J., 1988. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* 44 (1), 175–188.
- Yan, X., Lee, S., Li, N., 2009. Missing data handling methods in medical device clinical trials. *J. Biopharm. Stat.* 19 (6), 1085–1098.
- Yuan, K.-H., Bentler, P., 2000. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Socio. Methodol.* 30, 167–202.
- Yuan, K.-H., Bentler, P., 2010. Consistency of normal distribution based pseudo maximum likelihood estimates when data are missing at random. *Am. Statistician* 64, 263–267.
- Yuan, K.-H., Gomer, B., 2021. An overview of applied robust methods. *Br. J. Math. Stat. Psychol.* 74, 199–246.
- Yuan, K.-H., Lu, L., 2008. SEM with missing data and unknown population using two-stage ML: theory and its application. *Multivariate Behav. Res.* 62, 621–652.
- Yuan, K.-H., Savalei, V., 2014. Consistency, bias and efficiency of the normal-distribution-based MLE: the role of auxiliary variables. *J. Multivariate Anal.* 124, 353–370.
- Yuan, K.-H., Zhang, Z., 2012. Robust structural equation modeling with missing data and auxiliary variables. *Psychometrika* 77, 803–826.
- Yuan, K.-H., Zhang, Z., 2020. RSEM: Robust Structural Equation Modeling With Missing Data and Auxiliary Variables. R package version 0.5.0. <https://CRAN.R-project.org/package=rsem>.
- Yuan, K.-H., Lambert, P., Fouladi, R., 2004. Mardia's multivariate kurtosis with missing data. *Multivariate Behav. Res.* 39, 413–437.
- Yuan, K.-H., Yang-Wallentin, F., Bentler, P.M., 2012. ML versus MI for missing data with violation of distribution conditions. *Socio. Methods Res.* 41 (4), 598–629.

- Yuan, K.-H., Tong, X., Zhang, Z., 2015. Bias and efficiency for SEM with missing data and auxiliary variables: two-stage robust method versus two-stage ML. *Struct. Equ. Model.* 22, 178–192.
- Yuan, K.-H., Chan, W., Tian, Y., 2016. Expectation-robust algorithm and estimating equations for means and dispersion matrix with missing data. *Ann. Inst. Stat. Math.* 68, 329–351.
- Yuan, K.-H., Jamshidian, M., Kano, Y., 2018. Missing data mechanisms and homogeneity of means and variances–covariances. *Psychometrika* 83 (2), 425–442.
- Yuan, K.-H., 2009a. Identifying variables responsible for data not missing at random. *Psychometrika* 74, 233–256.
- Yuan, K.-H., 2009b. Normal distribution based pseudo ML for missing data: with applications to mean and covariance structure analysis. *J. Multivariate Anal.* 100, 1900–1918.
- Zhang, Z., Wang, L., 2013. Methods for mediation analysis with missing data. *Psychometrika* 78 (1), 154–184.
- Zhang, Q., Wang, L., 2017. Moderation analysis with missing data in the predictors. *Psychol. Methods* 22 (4), 649–666.
- Zhang, S., Zhang, J., Zhu, X., Qin, Y., Zhang, C., 2008. Missing Value Imputation Based on Data Clustering, *Transactions on Computational Science I*. Springer, pp. 128–138.
- Zhang, S., 2011. Shell-neighbor method and its application in missing data imputation. *Appl. Intell.* 35 (1), 123–133.
- Zhang, S., 2012. Nearest neighbor selection for iteratively kNN imputation. *J. Syst. Software* 85 (11), 2541–2552.
- Zhou, X.-H., Castelluccio, P., Hui, S.L., Rodenberg, C.A., 1999. Comparing two prevalence rates in a two-phase design study. *Stat. Med.* 18 (10), 1171–1182.

Displaying information: from woolly mammoths to the Great Migration

Howard Wainer^a and Michael Friendly^b, ^aPennington, NJ, United States; and ^bYork University, Toronto, ON, Canada

© 2023 Elsevier Ltd. All rights reserved.

Evolution of recording numbers	819
Using numbers	820
How to build a ship from a forest	822
What does a woolly mammoth look like?	822
A practical type I graphic— train schedule without showing locomotives	823
The first soaring national debt plot—William Playfair's iconic 1801 graphic	824
Hybrid displays: combining type 1 and type 2 displays	826
More than just facts: graphs as poetry	828
One plain graphical poem	830
The Kovno Ghetto	830
The graphic poetry of Charles Joseph Minard	831
Using graphs in a narrative argument	832
W.E.B. Du Bois	833
The Great Migration	835
A <i>gedanken</i> collaboration	837
Conclusion	839
References	840

Graphs are the intermediary between information and understanding.

with our apologies to Hans Hofmann (Friedel and Dickey, 1998).

Successfully negotiating the modern world often requires full and facile comprehension of quantitative phenomena. Experience gathered over more than three centuries has taught us this is best done by transforming the evidence into various kinds of pictures, so that we can enlist the remarkable power of the human visual system to rapidly decode and store complex information. Tracing the evolution of data visualization reveals many mysteries whose solution requires that we begin at the beginning.

Evolution of recording numbers

The idea of a number is very old indeed, and ancient ways of writing down numbers can be traced back to Paleolithic tally sticks (dating from the Aurignacian, approximately 30,000 years ago), which had notches cut into a bone, ostensibly to represent counts of something of interest. For example, such counting sticks were subsequently used to keep track of domestic animals. A notch was made when an animal was released to pasture and then later, when they returned, the shepherd's thumb would move down the stick, notch by notch. If no notches remained after the last animal returned, the shepherd would be assured that all was well. Such a system, though vastly better than relying on memory, had room for improvement. Adding an animal was easy, just another notch, but subtracting (perhaps because of a predator in the neighborhood or mutton for lunch) was more difficult and might require carving a new stick. Keeping separate track of several kinds of animals (e.g., goats and sheep or males and females) could be done with separate sticks for each type, but it would quickly become cumbersome to carry them around and to remember which was which.

Over time, the system of counting sticks evolved. The pre-Columbian Incans transformed their counting sticks into Quipus, ropes into which they tied knots that served the same role as notches. But knots could be untied, allowing subtraction. Also, two or more such ropes, each representing a different kind of animal, could be tied together and easily wrapped around the waist or thrown over the shoulder. Over time the Quipu evolved to include a summary: knots on the bottom portion of the rope each represented a single animal, but when ten knots accumulated, they were untied and a single knot representing ten was tied higher up on the rope. Some Quipus had three levels of knots, with a knot on the topmost level representing one hundred. These typically were above the joining of several strands (say goats, sheep and pigs), to indicate a grand total. In the evening, Quipus were sometimes hung on the wall of its owner's dwelling as a display of the family's wealth. We assume that the use of a base-ten for the construction of Quipus was strongly related to having ten fingers, which almost surely was the basis for all subsequent counting devices. Had the Incans had six fingers on each hand, there is little doubt that their mathematics would have been base twelve (Fig. 1).

The same forces that propelled the evolution of counting to yield the Quipu were at work in the Roman counting system, where a single knot became the written symbol "I" and a count of three was simply "III". But instead of untying the knots and replacing them with another knot in a different place, Romans substituted a new symbol "V" to mean five and still another symbol (X) for ten, and so on. The position of the symbol had little meaning except when indicating subtraction (e.g., IV) to avoid possibly confusing repetition (IIII). The Roman system was an obvious formalism of the humble tally (/, //, ///) still in use today.

Roman numerals were systematic, but their use for anything beyond record keeping and decoration was cumbersome. Try to imagine how Roman engineers could have built anything whose construction required computation of a precise measurement (e.g., how would you multiply 6342 by 18,327 with Roman numerals?). Despite this grave limitation, the evolution of written numbers, at least in Europe, was in stasis until the great Italian mathematician Leonardo Fibonacci (c.1175–c.1250) visited the Middle East during the time of the third Crusade.

In his travels in North Africa Fibonacci met with Middle-Eastern merchants and learned about the Hindu-Arabic numeral system, whose key component was zero. Upon his return Fibonacci told an engineer friend about this marvelous system, which used the digits 0–9 and place value to represent numbers but met with no enthusiasm. His friend rebuffed him, explaining that although this new method of recording numbers sounded interesting, he had a pressing problem in long division that had to be completed by next November and so he had no time to spare. This experience motivated Fibonacci to publish his justly famous 1202 treatise *Liber Abaci*. His book laid out in detail how the new system could be used in bookkeeping; interest calculations, money changing and other practical matters, and soon Roman numerals were relegated to being primarily decorative.

Using numbers

The search for truth has been a theme throughout human history. This search comprises two hierarchical categories—Do we use evidence? And, if so, what sorts of evidence have validity for the situation at hand?

The exploration of these great questions is called epistemology.

Epistemology is almost a foreign word and concept in common discourses. It typically makes its initial appearance, adjoined to metaphysics and phenomenology, as part of a trinity in the requisite undergraduate introductory philosophy course. But it rarely makes an impact on everyday life. Its definition usually contains a phrase like a “method for gaining knowledge,” which might translate into “Scientific method” which might then be specialized to “procedures used by actual practicing scientists.”

The iconic American physicist Richard Feynman provided, as only he could, a clear description of modern epistemology. In his famous Messenger Lectures, given at Cornell in 1964 (subsequently anthologized in his 1965 book *The Character of Physical Law*) he said,

In general, we look for a new law by the following process. First, we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science.

It does not make any difference how beautiful your guess is. It does not make any difference how smart you are, who made the guess, or what his name is—if it disagrees with experiment it is wrong. That is all there is to it.

Feynman answered the epistemological question of how and when evidence should be used by placing it in an exalted position. Evidence vetoed all else. Yet strangely, at least to us, this point of view has never been universal.

We hear the term “evidence-based decision making” in many fields: medicine, education, economics and political policy, to pick four. The implication is commonly that this is a new and modern way to try to solve contemporary problems. If what we are doing now is evidence-based, what were we doing previously?¹

How can we consider the use of evidence in science new? Hasn’t evidence been at the very core of science for millennia? The short answer is no. Making decisions evidence-based has always been a tough row to hoe, for once you commit to it, no idea, no matter

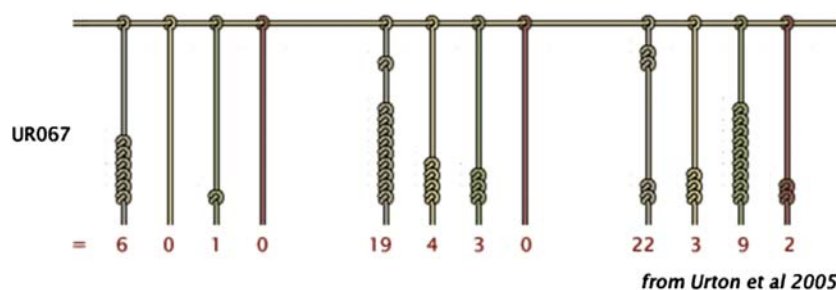


Fig. 1 A descriptive Quipu. Urton, Gary, 1998. From knots to narratives: reconstructing the art of historical record keeping in the Andes from Spanish transcriptions of Inka Khipus. *Ethnohistory* 45(5): 409–438, doi:10.2307/483319. JSTOR 483319.

¹Googling the term “Evidence-Based Medicine” reveals that the earliest reference is 1992. When I suggested that prior to that time medicine must’ve been faith-based my boss, a physician, said that he thought that it was intelligently designed.

how beautiful, no matter how desirable, can withstand an established contrary fact, regardless of how ugly that fact might be. The conflict between evidence and faith in the modern world is all around us, even in scientific issues for which faith is not required. So it is not surprising that using evidence to make decisions is taking a long time to catch on. The formal idea of using evidence as a method for gaining knowledge is often dated, as are so many things, with Aristotle (384 BC–322 BC). Aristotle diverged from his teacher Plato's purely rational approach by advocating the crucial inclusion of experience.² Aristotle famously studied the prenatal development of chickens by collecting a number of freshly laid eggs and then sequentially opening them, one each day, and noting the changes in the development of the chick embryos that he observed within them.

But the path from Aristotle to Feynman was not smooth. Once one commits to using evidence to make decisions, facts take precedence over opinion. And not all supporters of an empirical approach had the muscular support of Alexander the Great. Hence it took almost 1500 years before it briefly reappeared with Roger Bacon [1214–1292], who observed that,

Reasoning draws a conclusion, but does not make the conclusion certain, unless the mind discovers it by the path of experience.

But once again it slipped away, only to gain a firmer foothold with the work of Francis Bacon [1561–1626] who re-popularized the formal use of evidence, which was subsequently expanded and amplified by the British empiricists John Locke [1632–1704], George Berkeley [1685–1753], and especially the influential Scot David Hume [1711–1776], whose 1738 *Treatise on Human Nature*, as well as his 1741 *Essays, Moral and Political*, had a profound influence on Adam Smith, Jeremy Bentham and Immanuel Kant.

Although it took a very long time, by the late 18th century the stage was set. There was the beginning of a reasonably broad agreement on what constituted evidence, coupled with the by now well-established Hindu-Arabic number system, which itself provided a workable format for describing quantitative phenomena. And there were important and vexing problems requiring solutions, as well as a growing impetus to gather informative data. But what was lacking was a way to look at numbers and see meanings.

It is natural to ask how best to extract meaning from numbers. The answer is to transform numbers, which must be read, into pictures that could be seen. In doing so we enlist the extraordinarily powerful visual system, which allows us to see structure and, most importantly, to find what we were not expecting. This unanticipated benefit, obtained by visualizing quantitative evidence, profoundly changed the way science operated. It made the theoretical plotting of points with the goal of finding suggestive patterns possible and, because of an evolving string of successes, acceptable. Those wildly successful applications helped the evolution of what has become one of the most powerful tools for scientific discovery and for the communication of phenomena to those who could not be present when they occurred.

The story of the origin and development of the various kinds of data visualization forms the basis of this article. Although the notion of looking at quantitative measures of observations (numbers) to uncover meaning is surprisingly modern, the idea of visualizations is ancient.

There are two types of graphics that in a sense are antithetical to one another, yet both are critically useful tools.

The marvelous discovery of graphical methods to visualize data is usually dated to 1786, specifically the publication of William Playfair's *Atlas*, though there were earlier stirrings of the idea. Playfair's work, however, represented a dramatic development: spatial representation of data—e.g., time, value of imports, and value of exports—that were not, in their essence, spatial. There was no *a priori* logical basis for believing that such a representation would bear fruit, but it did. In 1822, Playfair happily reported that in 1787 France's King Louis XVI, after reading his *Atlas*, said, "they (the graphs) spoke all languages and were very clear and easily understood." Plotting of points in the search for unexpected patterns is now in full force. Mining big data intelligently would be insuperably more difficult without data graphics.

The sorts of graphics just described are defined by what is usually a synthetic spatial representation of data that are not, in their essence, spatial. They extend the visual idea of a two-dimensional map or town plan to data maps. There is a very long list of contemporary accounts explaining, exploring, and expanding the range of graphic possibilities. A modern list must include Bertin (1973) keystone explication, Tukey (1977) classic that called for a radically new graphical approach to statistics (the birth of "data science"), and Tufte (1983) masterful exposition showing that effective displays could provide "beautiful evidence." There has been a torrent of valuable additions since.³ Such applications, in which the word "graphic" is a noun, we will refer to as Type I graphics.

But there is a second definition of the word graphic, when it is used as an adjective with the meaning of *literal lifelikeness*—as in "graphic violence" or "graphic novel." This definition is quite the opposite of the synthetic displays that now dominate modern science. Ironically, this second definition is far older than the one that Merriam-Webster lists first, as are many of the data displays that represent it. These graphics not only satisfy the first goal of visualization, which is to answer basic questions, such as how do I get there, how do I do it, what happened, or even the causal, why did it happen? But they also serve the deeper purpose of acting as an intermediary between information and understanding.

²Aristotle was certainly not the first to advocate the critical advantages provided by empirical evidence. Sun Tzu [544–496 BC] predated Aristotle by two centuries and was the author of the widely cited *The Art of War*. In it he was very clear about the value of direct experiential knowledge. He wrote that, "Foreknowledge cannot be gotten from ghosts and spirits, cannot be had by analogy, cannot be found out by calculation. It must be obtained from people; people who know the conditions of the enemy."

³To learn more, the interested reader could do no better than the rest of Edward Tufte's trio of graphical wonders (1990 and 1996) and, being unafflicted with modesty, we also recommend some of our earlier descriptions; more specifically Wainer (1984, 1993, 1997, 2000, 2005, 2009, 2012, 2014) and, most recent, Friendly and Wainer (2021).

To illustrate the power of these “graphics of the second kind” and the questions they answer, we have chosen to look at two especially compelling examples drawn from the 18th and 19th century, specifically at examples of shipbuilding and the physical character of a woolly mammoth; of course, there are many more. The common characteristic of these graphics is that they provided a uniquely effective way to think about, reason about, and understand scientific phenomena; they illustrate how powerful and how broadly useful visualizations can be in the service of scientific and practical explanation.

How to build a ship from a forest

Looking at history from a modern perspective often reveals mysteries. One such mystery emerges when trying to imagine the combination of science and craft required for shipwrights to efficiently transform the contents of a local forest into a 19th century, 24-gun frigate for the British Royal navy. This task must have involved a shipwright looking carefully at each tree to see what parts could be used for different purposes and somehow communicating those arboreal insights to a sawyer for execution.

Fig. 2 from [Panckoucke \(1783\)](#) *marine Encyclopedia*, shows that communication was graphical. Shipwrights drew a picture of a portion of the forest and on each tree were inscribed the parts of the ship they needed that were to be the contribution of that tree. Long straight tree trunks would become the ship’s masts; shorter straight sections would be its spars or yardarms, and forked sections would naturally fill and buttress the bow or various overhangs. This graphic is an early example of visual thinking and visual explanation that compactly answers the questions: How do I do it? How do I transform the trees like those found in a local forest into the necessary parts to build a ship? Such instructions could be expressed verbally, similar to asking for turn-by-turn directions when you are lost. But, as with asking for directions and being drawn a map, this graphic gives immediate and memorable instructions that can be used and the lessons implicit in it are easily generalized for future use.⁴

What does a woolly mammoth look like?

In 1801, Charles Willson Peale, a curator of one of the first museums of natural history in the US, put a mammoth skeleton on display in Philadelphia’s Philosophical Hall. The skeleton was a mixture of anatomy and guesswork, with missing bones replaced by wood or papier-mâché. The exhibit proved so successful that Peale took it on the road, displaying it in both London and Bristol. A contemporaneous drawing based on Peale’s skeleton, ([Fig. 3](#)), however, shows that, his display had numerous anatomical mistakes—he even had the tusks upside down.⁵

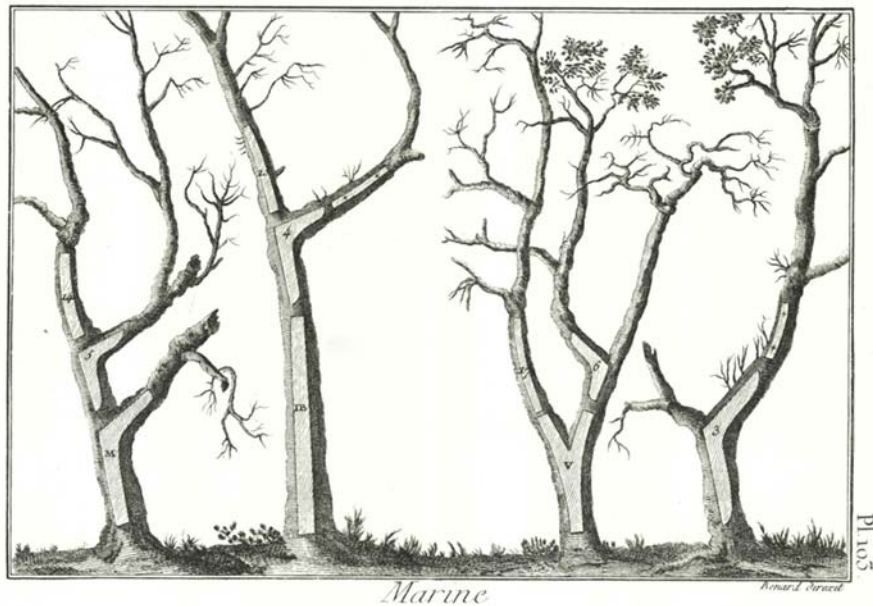


Fig. 2 Plate 103 in [Panckoucke \(1783\)](#) “*Encyclopédie Méthodique Marine*” showing how to harvest specific ship parts from trees (Our thanks to Kaicong Wu, who brought this example to our attention) http://www.epaves.corsaires.culture.fr/mediatheque/communs/images/grand_formats/1/449_4.jpg.

⁴Panckoucke was, famously, the publisher of the *Encyclopédie méthodique*, the successor to the *Encyclopédie* of Diderot & d’Alembert whose 1751 publication marked the editors valiant attempt to represent the thought of the Enlightenment and to “change the way that people think.”

⁵Peale’s construction was based on available fossilized bones, contemporary anatomical knowledge, analogical reasoning based on his knowledge of elephants. But there had to be other, even less accurate, sources based on his confusion about the mammoth’s character—for example he described it as carnivorous.

But modern reconstructions have it right—upward curling tusks, humped shoulders, downward sloping spine. We should note that our modern conception was not solely obtained by studying bones but by looking closely at the art of those who observed the animals first-hand. More than half of the hundreds of drawings left by Paleolithic artists on the walls of the Rouffignac Cave in the south of France were of woolly mammoths. The drawings answer the question: what did the animal look like? See Fig. 4. What led Peale astray was that he didn't have the benefit of the visualizations of Paleolithic artists to guide him. Having an accurate answer to "what did they look like?" means that investigators can at least begin on the right path. The value of such visual communications has often proved to be profound. Through such depictions, the original artists, their very bones long dust, have reached out across the centuries and aided modern understanding.

A practical type I graphic—Marey (1878) train schedule without showing locomotives

The popularity of visualizations owes much to the almost religious fervor of scientists and epistemologists of the 19th century who sought to banish subjectivity from science. They "drew on a number of techniques including inferential statistics, double-blind clinical trials, and self-registering instruments to hold subjectivity at bay".⁶ But, the oldest and most important, of these was visualization.

By 1878 the French physiologist Etienne Marey, whose graphic schedule of all the trains between Paris and Lyons (reproduced in Fig. 5) provides a powerful illustration of the breadth of value of this approach, expressed the feelings of most natural scientists of the value of graphical representation,

There is no doubt that graphical expression will soon replace all others whenever one has at hand a movement or change of state—in a word, any phenomenon. Born before science, language is often inappropriate to express exact measures or definite relations.⁷

Marey (1878, p. iii).

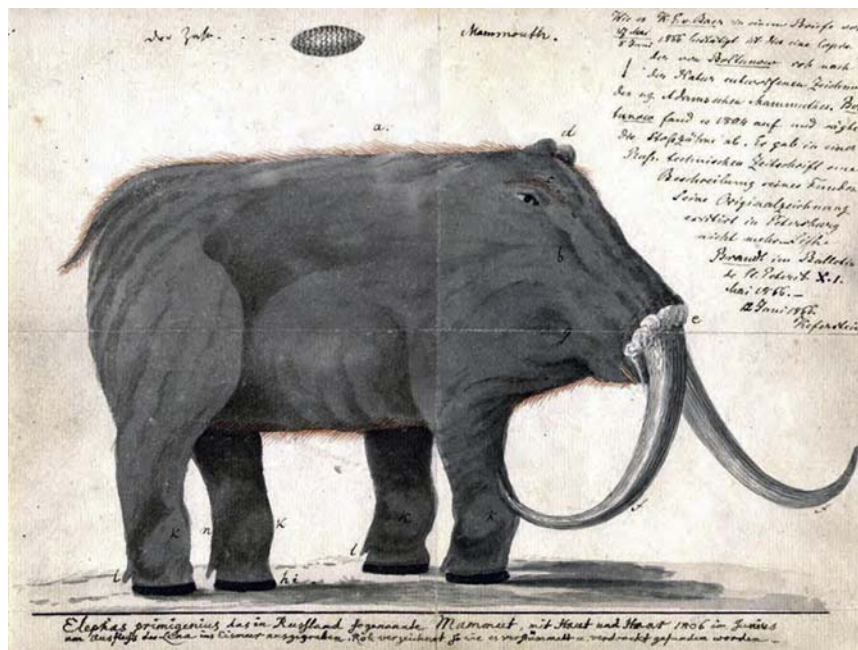


Fig. 3 An 1802 drawing of the mammoth, based on Peale's skeletal reconstruction which they characterized as a "carnivorous animal of immense size" <https://goo.gl/images/8xCpmd>- our thanks to Stephen Stigler for this figure.

⁶From p. 17 in Lorraine, J., Daston and Peter Galison, *Marvelous* 2007 Book, Objectivity.

⁷Marey (1878, p. vi).



Fig. 4 One selected panel of a woolly mammoth, from the Rouffignac cave. Courtesy of the Bradshaw Foundation.

Marey was also giving voice to the movement away from the sorts of subjectivity that had characterized prior science in support of the more modern drive toward objectivity. Although some cried out for the “insights of dialectic,” “the power of arguments,” and the “flowers of language,” their protestations were lost on Marey, who dreamed of a wordless science that spoke instead in high speed photographs and mechanically generated curves; in images that were, as he put it, in the “language of the phenomena themselves”.

The first soaring national debt plot—William Playfair’s iconic 1801 graphic

Fig. 6 is one of the most remarkable displays ever produced, containing a mixture of data and design innovations to provide a convincing visual argument that connected wars and debt. The design begins with the recognition that the grid of a plot serves the same purpose as the scaffolding that workers erect to construct a building; and just as the scaffolding is removed after the building is completed, so too should the graphical gridlines. Playfair recognized that no one is likely to be interested in what happened in 1700 or 1710 or 1720, but instead interest on specific dates would arise in two different kinds of situations:

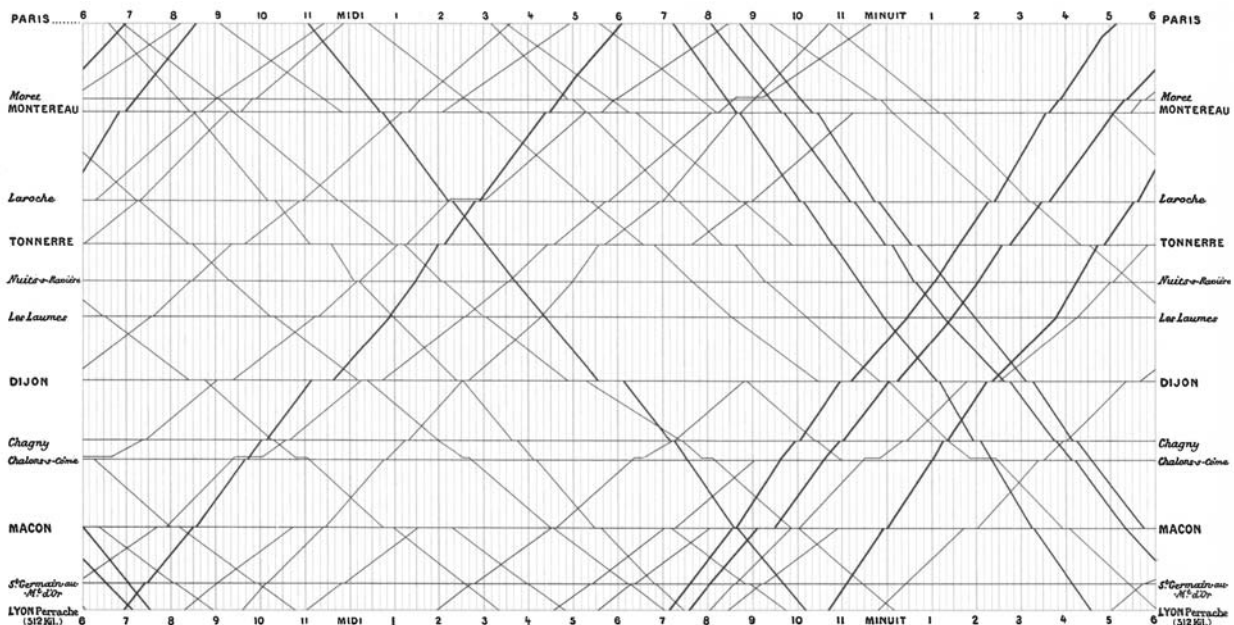


Fig. 5 Marey (1878) graphical train schedule, showing all trains between Paris and Lyons each day.

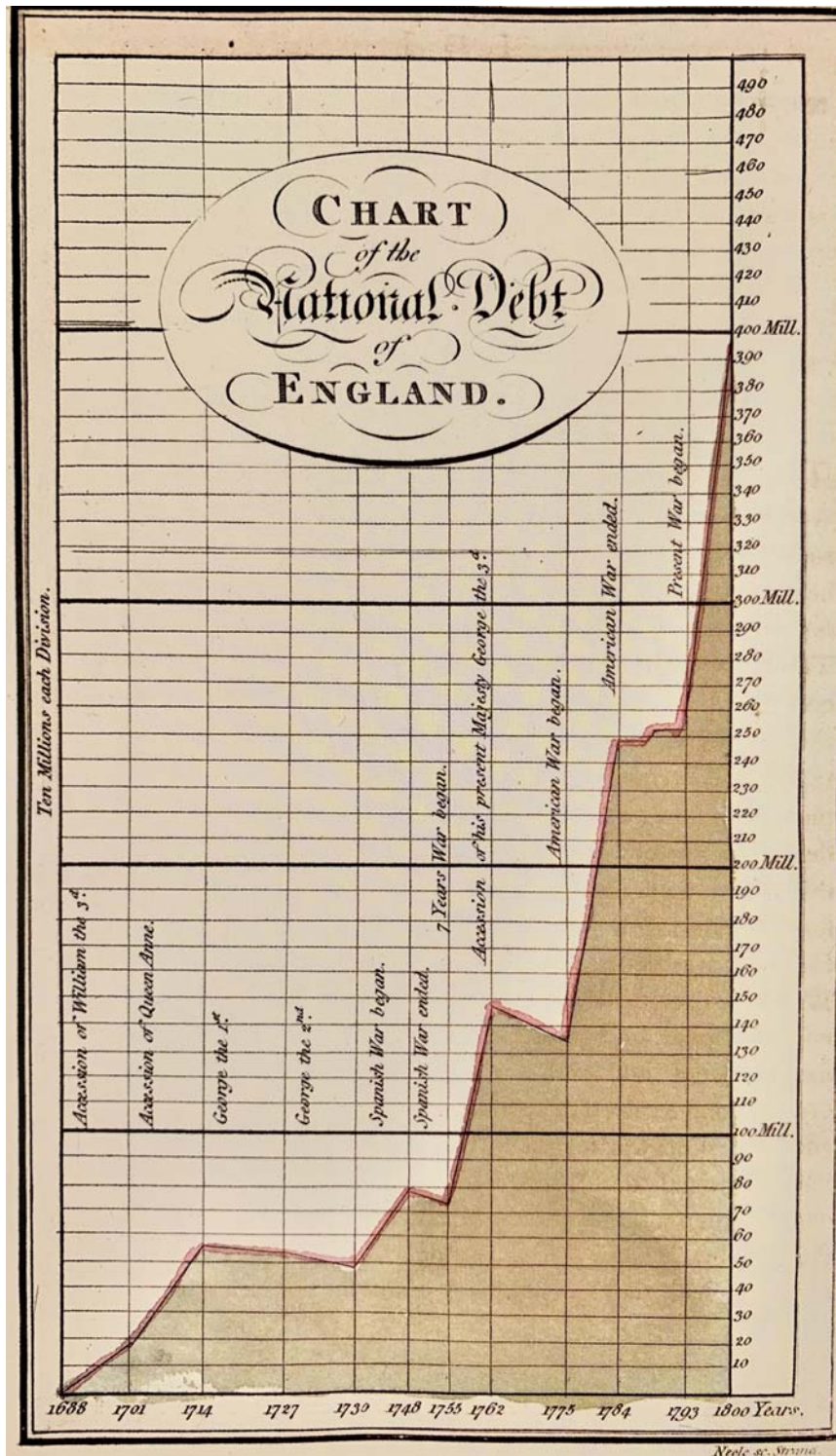


Fig. 6 The first “Skyrocketing National Debt” plot containing a lot of explanatory collateral information brilliantly integrated (Playfair’s Atlas, Plate 20).

- When the outcome one sees in the data suggests a question, e.g., what happened in 1730 or in 1775 to make the debt suddenly rise? Or in 1748 or 1784 when the debt fell or stabilized?
- When a plausibly impactful event occurs and we want to know what were its consequences on national debt, e.g., did the accession of Queen Anne in 1701 or that of George the second in 1727 have any effect?

With this idea in mind Playfair built the graph shown in Fig. 6. He then replaced the grid that he used to construct it with gridlines at interesting times. The vertical grid was no longer spaced equally across the century, and each vertical line is annotated with a plausible causal event.

We see that national debt increased in 1730 when the Spanish War began and dropped in 1748 when it ended; it jumped markedly in 1755 at the beginning of the Seven Years War and dropped in 1762 at its end (and the accession of George the third), it jumped in 1775 when the American War began and abated in 1784 when the war ended. The message is clear; wars are bad for national debt.

Last, Playfair obviously understood that perception of the data is easily manipulated by the ratio of the width of the graph to its height—the aspect ratio. The determination of this ratio is entirely subjective. If the horizontal axis is much larger than the vertical scale a soaring mountain can be transformed to a gradual plain. To make his point about the soaring national debt more emphatically, Playfair chose to make this plot higher-than-wide.

Modern research on effective design⁸ reveals that the arbitrary choice of aspect ratio is most often best resolved by selecting a ratio that makes the plot as close to a diagonal line as possible. Playfair's prescient design adhered to this advice even though the experimental evidence that revealed this result lay almost 200 years in the future.

Hybrid displays: combining type 1 and type 2 displays

Maps, which date back at least 4000 years, are among the oldest and most widely used of Type 2 displays. Their development isn't surprising, for it is a natural metaphor to represent space as space. But adding a third variable onto a geographic background is a more recent innovation. Of course, there were partially thematic maps (in China) as early as 164 BC, but the additional variables plotted were geographic variables such as altitude, which were ancillary to the main purpose of the map. A thematic map whose focus was primarily on the other variable did not appear until the 17th century.

The initial variables plotted on a geographic background were physical variables such as magnetic phenomena, currents, and geology. Such maps have obvious and immediate uses. Edmond Halley published what is believed to be the first meteorological chart in 1686. And, in 1815 William Smith (1769–1839) prepared a map depicting the geology of England that showed specific fossils were always found in corresponding layers of rock. This severely conflicted with the biblical interpretation, in which the rocks were laid down during the Great Flood. Smith's production has been called the "map that changed the world," instigating changes in thinking so that by the end of the 19th century, geology was—at least in terms of chronology and identification—advancing as a modern, evidence-based science.

The gathering of what might be called social statistics is often dated from John Graunt's 1662 analysis of the London Bills of Mortality. Graunt analyzed and published birth and death data collected by London parishes, but there were no graphs or maps included in his collection. John Arbuthnot (1710) famously used Graunt's data to illustrate his invention of null hypothesis testing, but he too did not include any graphs. In fact, the clerical errors in Arbuthnot's paper provide strong evidence that he never graphed the data. If he had prepared such a plot, the errors would have stuck out (literally) like sore thumbs and been corrected. That they weren't corrected provides support for the thesis that Arbuthnot never looked at Graunt's data in this way.

There were bits and pieces of graphic display of social data in the late 17th and early 18th centuries. Most notably, Christiaan Huygens drew a survival graph in a letter he sent to his brother in 1669. But graphs of social (economic) data only entered the scientific mainstream in 1786 with the publication of William Playfair's *Atlas*.

Playfair prepared wonderful, elegant graphs of economic data, but no maps. So far as is known, few maps of such social subjects as population, religion, or industrial production appeared before 1820. The preparers of the earliest such maps merely wrote the statistic of interest on the map.

An example of this practice is Wyld (1815) map, "Chart of the World Shewing the Religion, Population, and Civilization of Each Country." But such maps only barely "qualify as thematic since they do not graphically portray the character of the distribution effectively" (Robinson, 1982).

While it is surely true that putting numbers on a map to indicate the location where something is happening conveys more information than would merely listing them in a table, such a practice does not use the graphic medium at anywhere near its capability. This changed in 1830, when Frère de Montizon produced a map showing the population of France in which he represented the population by dots, each dot representing 10,000 people. Although no one remarked on it for the better part of a century, it remains perhaps the most important conceptual breakthrough in thematic mapping (Fig. 7).

After Montizon's map, there were a number of others that used shading to convey population between the years of 1830 and 1845. Often, they were roughly drawn and clearly meant to be of only secondary importance to the document in which they appeared. But a few were clever, remarkably polished displays. Exemplary among these were the population maps designed and prepared by Henry Drury Harness, a young lieutenant in the Royal Engineers that appeared in an 1837 *Atlas* that accompanied a report prepared by railway commissioners.

In 1847, a 34-year-old barrister named Joseph Fletcher published the first of several long papers. This paper was filled with tables and a single, rudimentary map. In fact, he rejected the use of maps explicitly, suggesting you could get all you needed by simply

⁸Cleveland (1994).

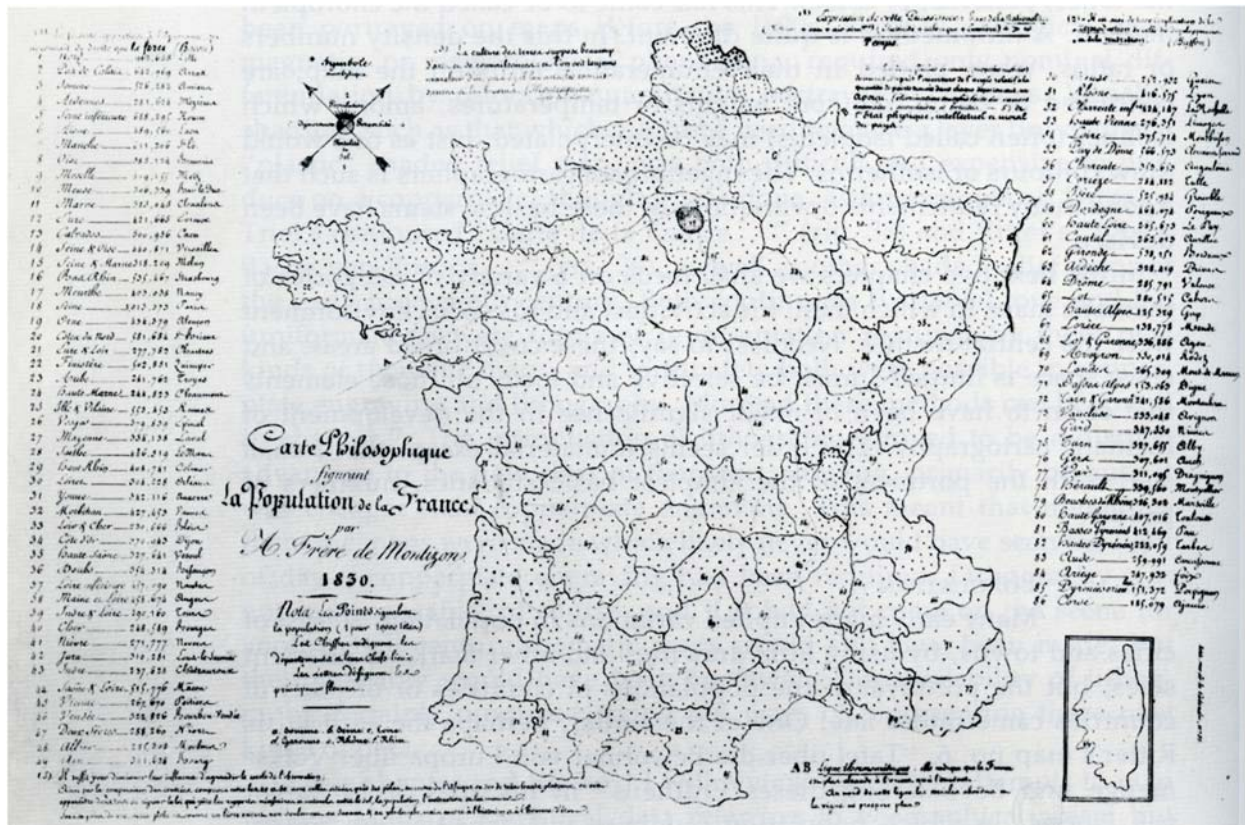


Fig. 7 Frère de Montizon's 1830 map of the population of France using irregularly spaced dots to convey population density. Original at the Bibliothèque Nationale. This one taken from Robinson (1982, p. 112).

perusing the columns of numbers and thus avoid the bother and expense of drafting and printing maps. He reversed himself completely⁹ two years later in two articles with the same title (but much greater length) as his 1847 paper. These included many shaded maps that were innovative in both content and format. Format first.

Fletcher, having a background in statistics, or at least what was thought of as statistics in the mid-19th century, did not plot the raw numbers, but instead their deviations from the mean. And so, the scale of tints used varied from the middle. This approach is common now. For example, Andrew Gelman—in *Red State, Blue State, Rich State, Blue State*—uses increasing saturations of red to show the proportion of voters who voted Republican in states where they were the majority and increasing saturations of blue for those states that voted Democratic. The closer the vote came splitting 50:50, the closer the shading was to white, regardless of the final outcome.

But, in 1849, this format was new. The advantage of centering all variables at zero was that it allowed Fletcher to prepare a number of maps, on different topics and scales, and place them side-by-side for comparison without worrying about the location of the scale. He also oriented the shading, as much as possible, in the same way. In his words, "In all the maps, it will be observed that the darker tints and the lower numbers are appropriated to the unfavorable end of the scale."

Of course, with variables such as population density, it is hard to know which end is favorable. He chose lesser population as more favorable because, we speculate, it seemed to accompany favorable outcomes on the other variables.

Although many of Fletcher's innovations in format were noteworthy, it is the content he chose that makes him special. He did not make maps of obvious physical variables such as wind direction or altitude, or even the distribution of the population (although he did make one map of that for comparative purposes). No, he had more profound purposes in mind. Joseph Fletcher made maps of moral statistics and opposed these maps with others to suggest causal interpretations to his viewers.

For example, juxtaposed next to his plot of "Ignorance in England and Wales" (Fig. 8), he also had maps of the incidence of crime (Fig. 9). He opted for this juxtaposition on the basis of detailed analysis of subsets of the data. For example, he wrote, "We thus find that the decline in total ignorance to be slowest in the most criminal and the most ignorant districts ..."

And then, looking into the phenomenon at a more microscopic level, he observed:

⁹Ordinarily, it would be hard to know exactly what it was that got Fletcher to change his approach so completely, but an evocative hint was provided in his 1849 paper when he states, "A set of shaded maps accompanies these tables, to illustrate the most important branches of the investigation, and I have endeavored to supply the deficiency which H.R.H. Prince Albert was pleased to point out, of the want of more illustrations of this kind."

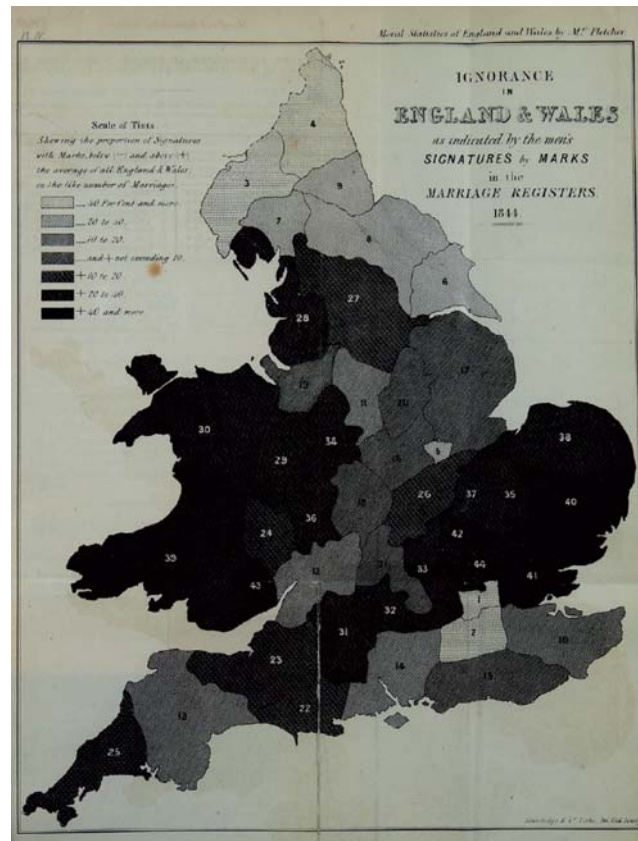


Fig. 8 Ignorance in England and Wales taken from [Fletcher \(1849b\)](#).

The two least criminal regions are at the opposite extremes in this respect (the Celtic and the Scandinavian), with this important difference, that in the region where there is the greatest decline of absolute ignorance among the criminals (the Scandinavian), there is not one-half of the amount of it in the population at large which exists in the other

In addition to these maps, he also produced parallel plots of bastardy in England and Wales, improvident marriages, persons of independent means, pauperism, and many other variables that could plausibly be thought of as either causes or effects of other variables.

His goal in doing this was to generate and test hypotheses, which might then be used to guide subsequent social action and policy. His epistemological viewpoint was both wise and surprisingly modern. It is fitting that we close with three sentences of his on the use of the data and his maps (note the similarity to Richard Feynman's views quoted earlier):

The man who studies society, however, labors under a great difficulty in being entirely denied the use of experiment, and limited most rigidly to observation; the observation of elements most subtly combined, in a state of unceasing change, and wholly beyond his control. Analysis, therefore, in the sense of the chemist, is absolutely impossible; but by exhaustive enumerations of facts, which are strongly indicative of the existence of many others, or are their invariable concomitants, we get a means of detecting the excess or deficiency of certain social elements in definite classes or localities; and by multiplying these lines of observation, and the combinations in which they are arranged for purposes of comparison, we gradually arrive at higher and safe inductions, which will sometimes corroborate principles which we have reached by deductive reasoning from the moral elements of individual character, and by observations on society in the limited field of our person experience, and at others will present irreconcilable results, which the bigot of theory will despise, but which the man of science knows how to prize as gems. He who seeks facts merely to illustrate a hypothesis in which he believes with a blind faith, will throw them away in disgust; but he who uses a hypothesis merely to discover truth, will, on the contrary, abandon its use in the moment that he arrives at acts which resist all efforts to reduce them into accordance with them.

More than just facts: graphs as poetry

So far, we have portrayed visual displays of empirical information as compact summaries that, at their best, can clarify a muddled situation. This is true, as far as it goes, but omits the magic. In this section we will show how the visualization of important data can be an alchemist that can make good scientists great and transform great scientists into giants.

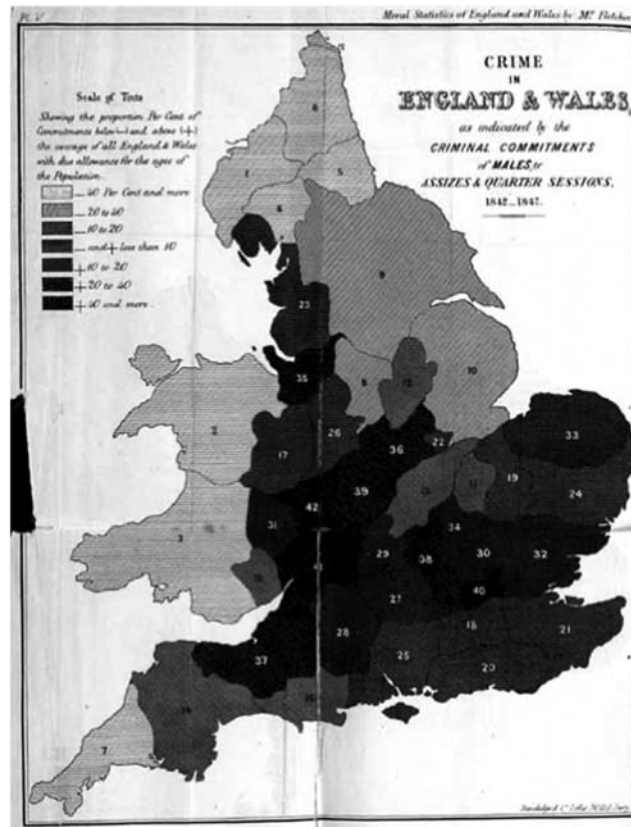


Fig. 9 Crime in England and Wales taken from Fletcher (1849b).

In this coda we argue that sometimes, albeit too rarely, the combination of critical questions addressed by important data and illuminated by evocative displays can achieve a transcendent, and often wholly unexpected, result. At their best, visualizations can communicate emotions and feelings in addition to cold, hard facts.

The American poet Robert Frost [1874–1963] famously pointed out that:

Poetry is when an emotion has found its thought and the thought has found its words.

We can reverse the idea, inferring that one goal of the poet is to present the words that will engender the emotion.

The communication of an emotion is a difficult task that is a goal of many other means of communication. Surely music is the most obvious; in his *Ode to Joy* Beethoven transformed his emotions through his ideas to music. The listener starts with the music and ends with exaltation. Visual artists have long followed this same path; Picasso transmitted his grief for the hundreds killed in the April 1937 bombing of the Basque town of Guernica during the Spanish Civil War in his mural *Guernica*, completed just two months after the bombing. For the receptive viewer “the emotion is so lacerating that the next step beyond would be either insanity or suicide.”¹⁰

Not all music is as transformational as *Ode to Joy*, nor are all paintings as evocative as *Guernica*; and most sequences of words are not poetic. But, happily, by Frost’s definition, there have been poets in many media, and we can rejoice in their accomplishments. Goethe’s observation that “Architecture is frozen music” suggests we might find the soul of a poet in many architects. One undeniable example is Maya Lin, whose design for the simple chronological carving of 57,661 fallen soldiers’ names into polished blocks of black granite yielded the heart-rending Vietnam Memorial in Washington DC.

We can include some databased graphs in this elite company—graphs whose impressions on the viewer are so evocative to fully deserve the adjective “poetic.” To reach this level requires a combination of data that have impact and a design of compact clarity. Reaching back to Frost’s definition, it happens when an emotion has found its data and the data have found their design.

¹⁰Mumford (2000, p. 12).

As should be obvious, a data display can only achieve this transcendent property when the data are worthy. Consider the formally competent plot shown in Fig. 10, that shows the fluctuating preferences for three flavors of ice cream in NJ over the past 30 years. It is both informative and instantly forgettable.

Now, consider an identical plot in Fig. 11 with the labels changed to reflect what the data actually were, augmented with one plausible causal covariate (the president during the time periods indicated).

The responses to this plot that we have received so far, have led us to conclude that it is difficult, if not impossible, to avoid making some causal conjectures and to immediately begin looking for other sources of data to confirm or reject those conjectures. Those conjectures arose with considerable affect. Obviously, it wasn't the graph, it was the data; the graph was the *intermediary between information and understanding*.

Let us begin slowly, by first introducing one graphic poem that might otherwise be overlooked because of its plain appearance. Then we introduce two giants in their respective fields: Charles Joseph Minard, whose use of flow maps to communicate the horrors of war defied the limitations ordinarily associated with pen and ink, and W.E.B. Du Bois, who dedicated his long career to the difficult task of improving the lives and circumstances of African-Americans. He did this by broadly communicating more than a century of African-Americans' accomplishments despite their near-universal suffering under the yokes of slavery and racism. After this introduction to these men and their work we describe a plausible result of a *gedanken* collaboration between them.

One plain graphical poem

The conductor Ignat Solzhenitsyn, in a prelude to an all-Mozart concert, made the surprising observation, "*Mozart's music is often underestimated because it is so beautiful.*" Similarly, the poetry in some graphic designs is often unappreciated because those designs are so mundane. Few graphic artists have Picasso's or Maya Lin's eye, but we ought not be blind to poetry from others with less refined skills.

The Kovno Ghetto

During the Nazi occupation of Lithuania the Nazis initiated a series of actions that resulted in the deaths of over 136,000 Jews. Initially, the Germans, in elaborately illustrated reports, meticulously documented these murders, but their production of records was reduced as the war continued. Indeed, in October of 1943, SS chief Heinrich Himmler, in a speech to his subordinates, rationalized the minimized record keeping, saying, "This is an unwritten and never-to-be-written page of glory in our history." Nonetheless, the Nazis did create some documents that tabulated what they considered their triumphs, and others also kept records of their atrocities.

In many ghettos, Jewish leaders organized committees to keep official chronicles of daily life. Indeed, from the outset of the establishment of the Kovno Ghetto, Elkhanan Elkes, chairman of the Jewish council (the *Ältestenrat*), asked Kovno's Jews to write their own history as a legacy for future generations. Artists drew pictures, writers wrote stories, musicians composed music, poets wrote verse. They all used the skills and talent that they had to record both facts and emotions for a posterity and an audience that they feared they would not be able to meet directly. Ghetto residents without such artistic talent used whatever skills they had. Those with scientific training recorded data and presented them in a variety of formats, tabular and graphical.

Fig. 12 is one of those graphs, a traditional population pyramid, whose familiar simplicity and apparent banality belie its horrifying content. The bars represent the number of Jews in the ghetto by age, youngest at the bottom. The left side represents males, the

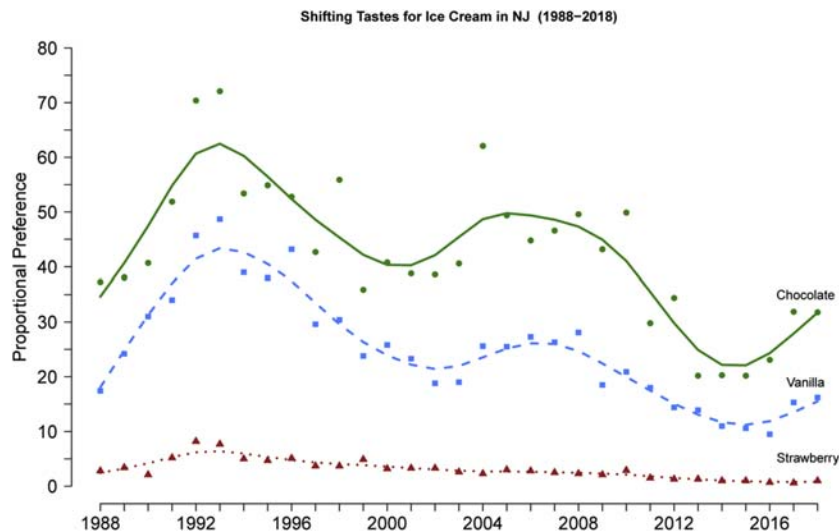


Fig. 10 Preferences for three flavors of ice cream in NJ from 1988 to 2018.

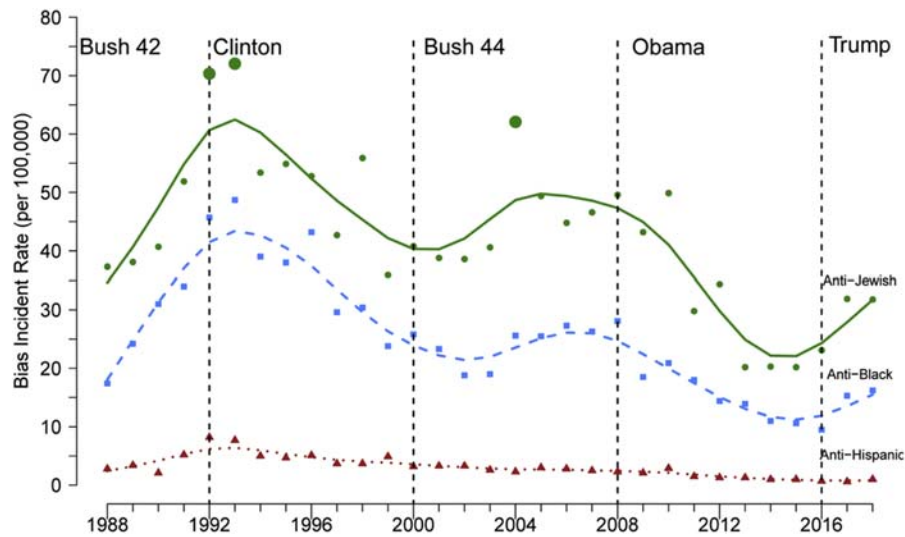


Fig. 11 The rates per 100,000 of bias crimes in NJ from 1988 to 2018.

right females. The total length of each bar is the number of inhabitants of the ghetto of that age and sex in the beginning of October 1941. The shaded portion represents those still alive in November, only a month later.

The graphic poetry of Charles Joseph Minard

Charles Joseph Minard [1781–1870] had a long and productive life. But official regulations forced him to retire from his long-held position at the National School of Bridges and Roads on March 27, 1851, his 70th birthday. Nevertheless, he continued working and teaching for the remainder of his life. The liberty he enjoyed in retirement allowed him to devote himself to projects he had

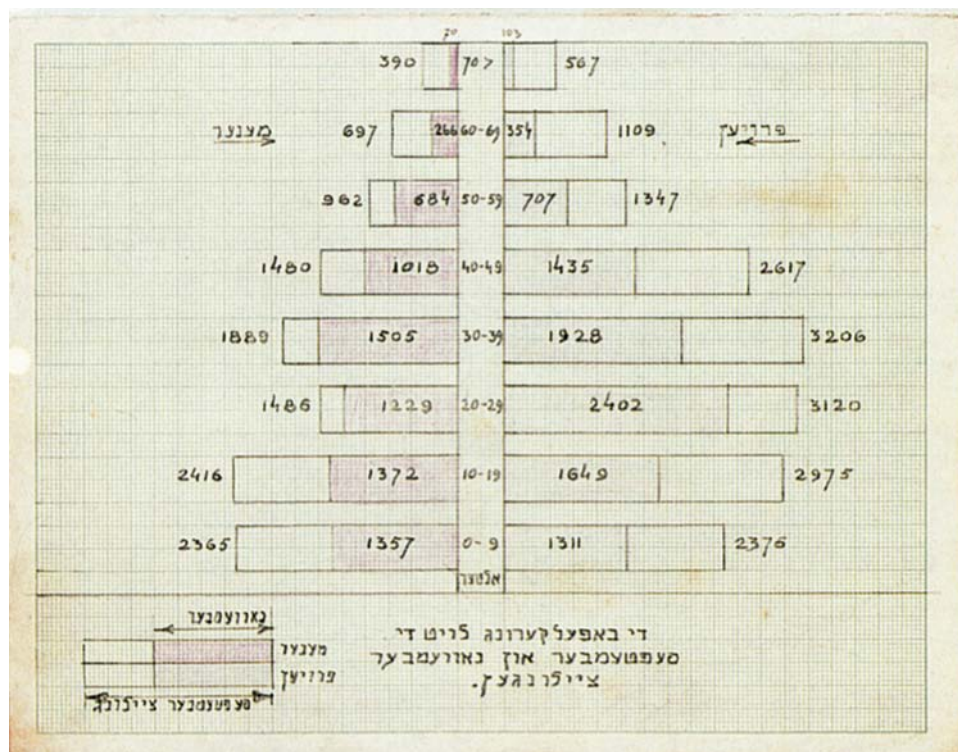


Fig. 12 Kovno Ghetto: the population losses in the Kovno Ghetto due primarily to the “Great Action” of October 28, 1941. Males are represented on the left, females on the right. The shaded portion represents those still surviving in November. The central column indicates the age groups from 0 to 9 year olds at the bottom to 70 and older at the top. Source: [Braun and Wainer \(2004\)](#).

begun earlier but that had been interrupted by the obligations of his job. Freed of other responsibilities, his development of new graphic forms and themes nearly doubled in rate for 20 years, and continued up to his death at age 89.¹¹

Fig. 13 is Minard's justly famous map of Napoleon's Russian campaign, in which the initially mighty metaphorical river of the 422 thousand men of the French Grand Army begins on the left by crossing the Nieman River into Russia, it gradually diminishes and is juxtaposed with the trickle of the surviving 10 thousand returning at the end. When the French army reached Moscow in October of 1812 with only 100,000 men, the moment captured at the far right of Minard's map, they found it sacked and largely deserted. They turned back and marched across the steppes into the teeth of the Russian winter—Minard coupled the returning march with the descending temperature, shown in an accompanying panel below the map. He used the falling temperature as one plausible causal variable to connect the decline of the army's size with something in addition to the sniping of Russian soldiers. When this graphic was first published, the French physiologist and chronophotographer E.J. Marey¹² [1830–1904] was in awe, saying that it “defied the pen of the historian by its brutal eloquence.”

What was the inspiration for Minard's graphical poetry? Minard's personal history makes the source clear. As a young engineer in Anvers in 1813, he had witnessed the horrors of war in a siege by the Prussian army. In his 89th year, 1869, he was deeply troubled by what he foresaw as an inevitable new war with Prussia and the havoc that it would wreak on the Second French Empire. His “figurative map” of Napoleon was published on November 20, 1869.

Minard's fears proved correct. The Franco-Prussian War began on July 19, 1870. France was ultimately defeated, and, following a long siege, Paris fell on January 28, 1871. Minard foresaw this too and, though now infirm and requiring crutches, he left Paris for Bordeaux on September 11, 1870, leaving behind nearly all his books and papers. He carried with him only a few works-in-progress, but these apparently have been lost. Sadly, six weeks later he contracted a fever, and on October 24, 1870, he died.

Very little of Minard's personal life is known; the main historical source is the obituary written by his son-in-law, Victor Chevallier in 1871.¹³ He said, “Finally, ... as if he could sense the terrible disaster that was about to disrupt the country, he illustrated the loss of lives that had been caused by ... Napoleon The graphical representation is gripping; ... it inspires bitter reflections on the human cost of the thirst for military glory.” This may well be the reason that this graphic story has become the iconic example of Minard's graphic poetry.

Using graphs in a narrative argument

Using a series of graphical displays to construct an empirical narrative was not common in Minard's time. More often, a single display was used to convey a single idea. Napoleon's March (Fig. 13), as breathtakingly marvelous as it is, is still a short story,

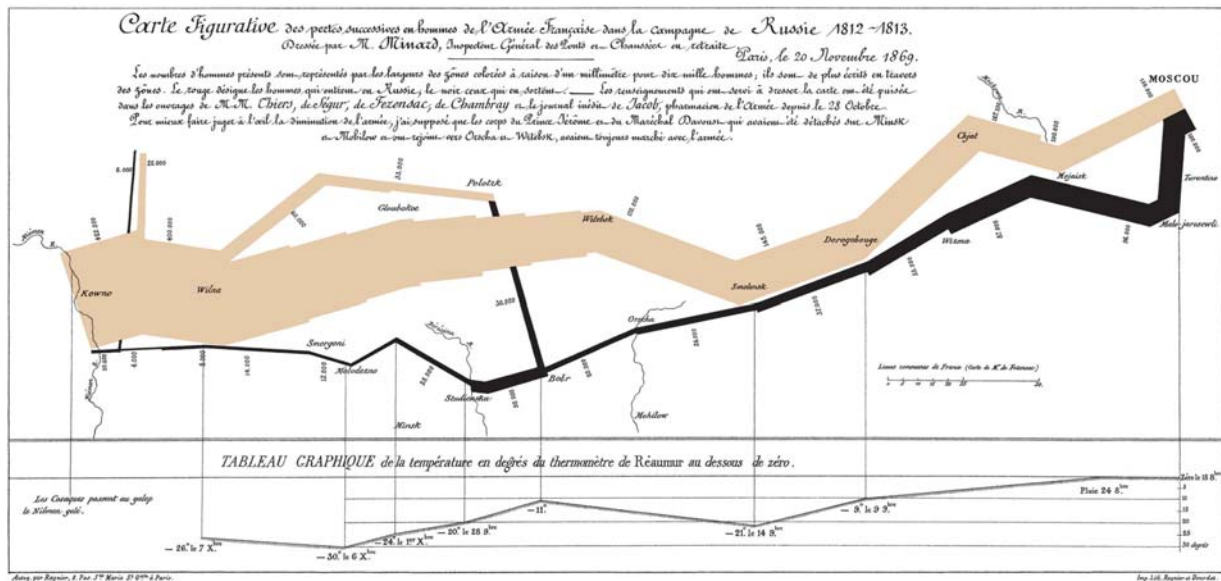


Fig. 13 Napoleon march graphic: Charles Joseph Minard's narrative map of Napoleon's disastrous 1812 Russian campaign. The width of the orange “river” is proportional to the size of Napoleon's invading army; its black continuation shows the size of the returning army. Source: Archives, École Nationale des Ponts et Chaussées, by permission.

¹¹See Friendly (2002) for a timeline of Minard's works.

¹²This is the same Étienne-Jules Marey we met previously when we discussed his well-known and easy-to-use graphical schedule for all trains between Paris and Lyon shown in Fig. 5.

¹³V. Chevallier, Notice nécrologique sur M. Minard, inspecteur général des ponts et chaussées, en retraite, Annales des ponts et chaussées, 2 (1871), 1–22. An English translation was prepared by Dawn Finley, <https://www.edwardtufte.com/tufte/minard-obit>. For further detail on Minard's life, see Friendly (2002) and a biography, <http://datavis.ca/gallery/minard/biography.pdf>.

with the limitations commonly associated with that form. Embedding a sequence of displays in an empirical narrative has been too rarely used. But when used effectively, this approach forms a memorable image that has remarkable impact.

Minard himself sometimes used comparative graphs to tell a story. As one example, he prepared a sequence of plots to show how the Union blockade of Confederate ports during the American Civil War affected the supply of cotton to English mills. In Fig. 14 we reproduce his three-part plot (for 1858, 1864, and 1865) which made it clear that before the US Civil War the southern US states were the principal supplier for British mills. But when that supply was effectively cut off by the blockade, the mill owners replaced it with cotton from India and Egypt, which continued, at a slightly diminished level, even after the war had ended.

Arguably, the most famous (and perhaps the first) use of a coordinated sequence of displays to complement an integrated verbal explanation was Leonardo's magnificent tale (Fig. 15) of how a human fetus is carried in a womb. It challenged Galen's 1400-year-old description of the human womb as bicameral (having two sinuses) to carry multiple births.

W.E.B. Du Bois

William Edward Burghardt (W.E.B.) Du Bois was born in Great Barrington, Massachusetts on February 23, 1868, shortly after the end of slavery in the United States. He died in Accra, Ghana on August 27, 1963. During the course of his 95 years he accumulated an extraordinary list of accomplishments. He earned a Ph.D. from Harvard—the first African-American to do so. He was by turns a sociologist, historian, civil rights activist, and author. His many books included: *The Souls of Black Folk*, *The Suppression of the African Slave Trade*, *Dusk of Dawn*, *Darkwater*, and *The Talented Tenth*. He was also a skilled data scientist devoted to the power of unadorned facts.

It is natural to ask how DuBois, more than a century ago, was able to look at numbers and extract meaning. Certainly the consensus then, and now, is that the extraction of meaning from the huge tables of numbers that might carry that meaning is wearisome to the eye and difficult for the brain; as the American economists Farquahr & Farquahr so memorably put it in 1898, it is a task akin to extracting sunbeams from cucumbers. We can plausibly conclude that DuBois was best able to read meaning from numbers after they had been transformed to visual displays. His remarkable success at this means that as a data scientist, he contributed to the development of the modern approach to constructing a narrative argument based on evidence through his use of a series of graphic displays. In this way he is a direct, lineal, descendant of Leonardo and Minard.

In 1900 Du Bois collaborated with Booker T. Washington on "The Exhibit of American Negroes" at the Paris World Fair (the *Exposition Universelle Internationale*). It included 400 patents by African-Americans as well as 200 books with African-American

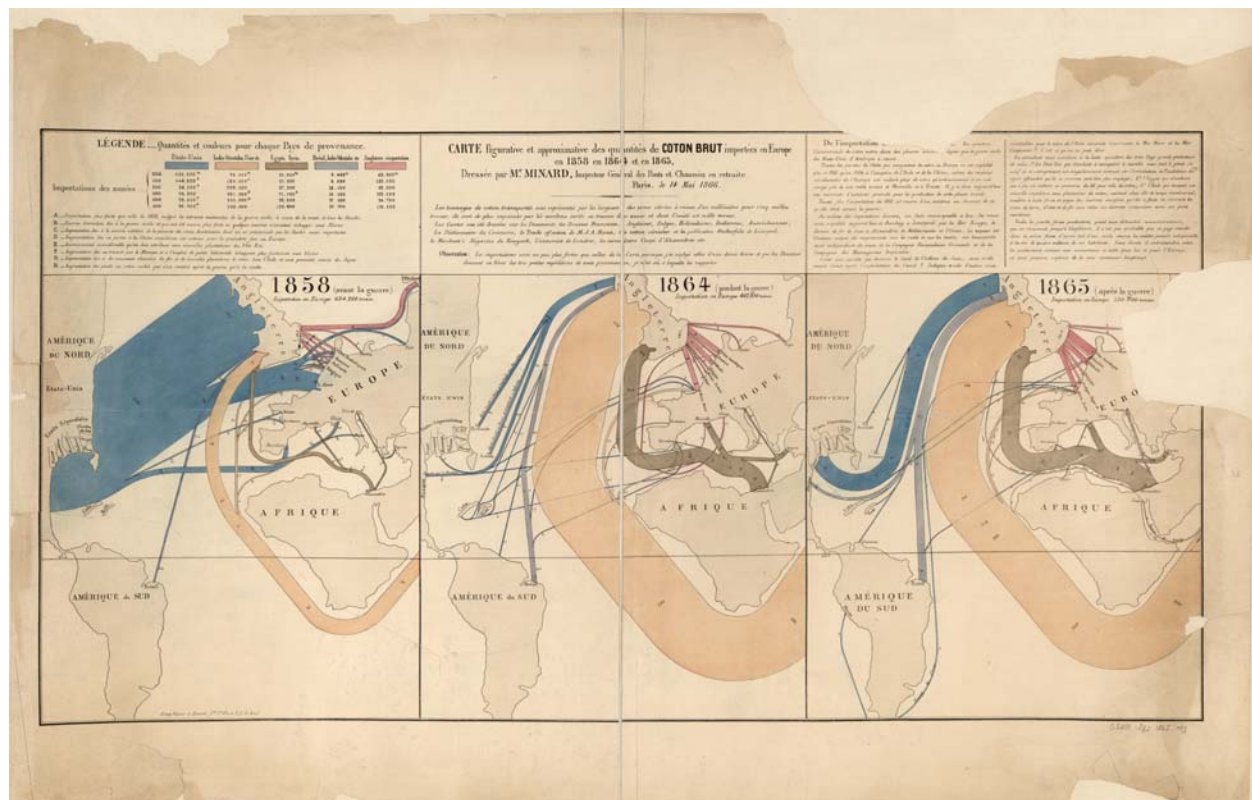


Fig. 14 Minard: imports of cotton: a sequence of three graphs showing the sources of cotton imports to Europe for 1858, 1864, and 1865. The metaphorical blue river represents imports from the US, the orange from India. Source: National Archives <https://www.loc.gov/item/99463789/>.



Fig. 15 Da Vinci's notebook: Leonardo's graphic story of a fetus told with words and incomparable pictures that disproved Galen of Pergamum's claim about women's bicameral uterus that had persisted unchallenged for 1400 years. Source: Wikipedia, public domain.

authors. In addition, a large number of facts about African-Americans were woven into a memorable narrative by transforming them into data graphics.¹⁴

The nearly 60 graphs and thematic maps spanned a wide range of characteristics of African-Americans and their lives. This section of the exhibit was titled "A series of statistical charts illustrating the condition of the descendants of former African slaves now in residence in the United States of America."

Du Bois began his story with a graph dramatically showing the profound effect of Lincoln's Emancipation Proclamation (Fig. 16). It shows the balance of enslaved versus free African-Americans from 1790 until 1870. From this we learn that the proportion of free African-Americans during slavery hovered around 12% of the total African-American population until January of 1863, when Lincoln's Emancipation Proclamation¹⁵ was the key to the dramatic change shown.

Having established the status of African-Americans in the United States, he then used a simple bar chart (Fig. 17) to show that there were seven-and-a-half million African-Americans in 1890 and also the exponential growth of that population over the prior

¹⁴Du Bois's account of this exhibit appears in Du Bois (1900). Nearly the entire collection, comprising over 200 photographs and numerous charts and maps, has been digitized by the Library of Congress. See: <https://blogs.loc.gov/picturethis/2014/02/du-bois-american-negro-exhibit-for-the-1900-paris-exposition/>.

¹⁵Although all slaves in the Confederate States were freed by Lincoln's proclamation in 1863, it was not until the ratification of the 13th amendment on December 6, 1865 that all slavery was truly abolished in the United States.

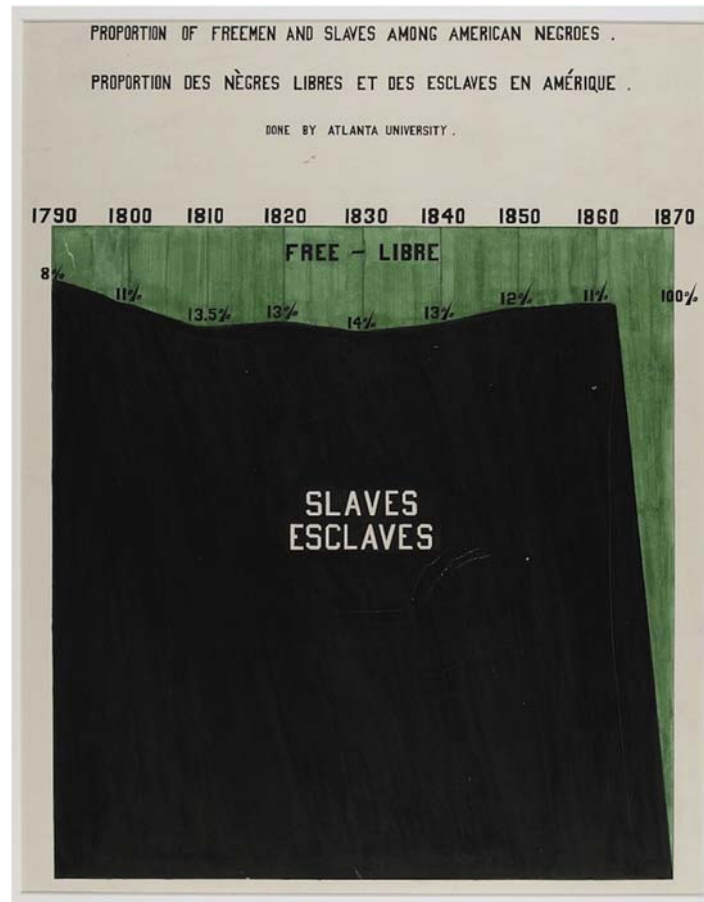


Fig. 16 Effect of the emancipation proclamation: proportion of slaves and free Negroes from 1790 until 1870. Source: Library of Congress, <http://hdl.loc.gov/loc.pnp/ppmsca.33913>.

150 years. Although a simple time series line graph in the style of Playfair might have been used, Du Bois chose the format of a horizontal bar chart in the style of Minard to also show the exact numbers.

Du Bois used these displays to answer fundamental questions of existence: How many of us are there? Are we increasing or decreasing in number? How fast? The next obvious questions are “Where do we live?” and “Where did we come from?” He addressed the first of these in Fig. 18, which showed dramatically the heavy concentration of African-Americans in what had been the states of the Confederacy, but there were signs of a movement out of the Deep South into the Northeast. He addressed the second question (Fig. 19) with the very first display of his exhibit, showing the sources and destinations of the slave trade.

He continued with 53 other plots, expanding and enriching the narrative. This fact-based description was made possible by the 1870 expansion of the US census, which, for the first time, included the African-American citizens in the national accounting.

Absent, however was any sort of coherent representation of the gigantic migration of African-Americans out of the South that took place after Reconstruction and continued for the first half of the 20th century. This phenomenon could have been beautifully communicated with a collaborative effort between Du Bois and Minard.

The Great Migration

As we have seen in Figs. 17 and 18, there were 7.5 million African-Americans in 1890, the great majority of whom lived in the rural south of the United States. Between 1916 and 1970 more than 6 million had migrated to the industrial North and West. Historians have dubbed this the Great Migration, commemorating it as the largest peacetime movement of people within a country in history.¹⁶

Such a movement cries out for description and explanation. The biggest initial question is, of course, “Why?” But before one can tackle such an interpretative question, one faces the descriptive questions: “From where to where?” “When?” and “How Many?” These questions are fully answered with data, indeed, the sorts of data that are routinely and rigorously gathered by the U.S.

¹⁶Lemann (1991).

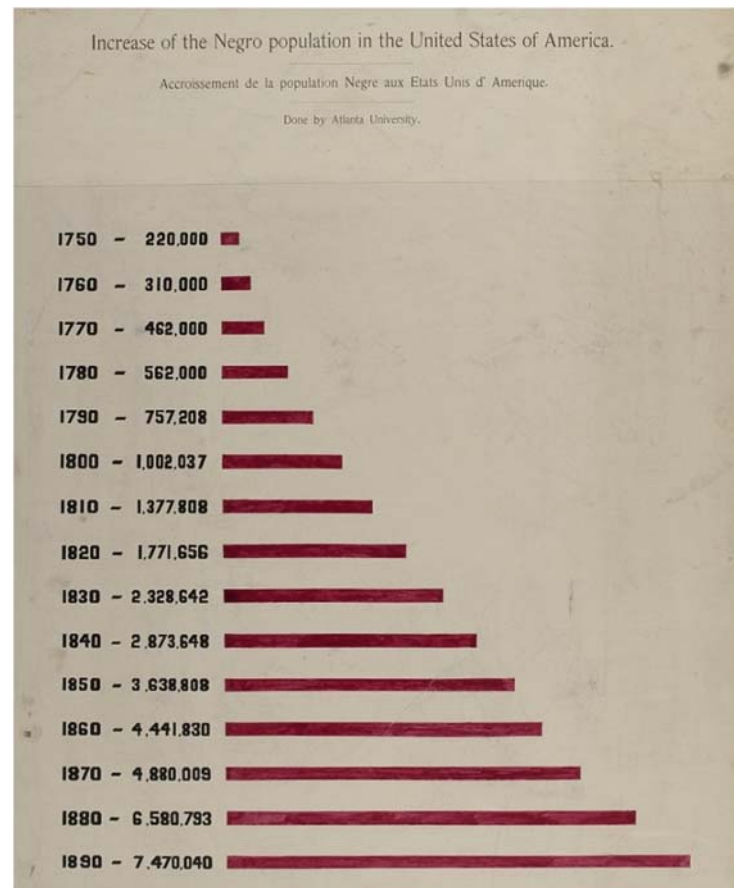


Fig. 17 Du Bois's bar chart: increase of the Negro population of the United States from 1750 until 1890. Source: Library of Congress, <http://hdl.loc.gov/loc.pnp/ppmsca.33901>.

decennial census. But extracting geographic structure from the entangled tables of numbers describing it brings to mind the cucumbers of the brothers Farquahr. And communicating the messages contained in these census tables beyond the eyes of experts is far better done graphically. But how are we to transform the tables to graphs?

To answer this question we adjoin the life work of the two giants whose work forms the center of this article: Du Bois's passion for improving the lives of African-Americans and widely communicating their accomplishments, and Minard's genius for transforming complex data into compelling pictures that impacts both the intellect and the emotions.

Although their long lives overlapped by but two years (and even during that brief period they were separated by language, culture, and an ocean) the work and goals of Minard and Du Bois were complementary. In the balance of this article we demonstrate of how a blending of their work yields an evocative narrative of such awakening power that it can rightfully be described as poetic.

The census has always been tasked with counting all residents of the United States; in 1863 its' scope was expanded to include those African-Americans who had been emancipated from slavery. Thus, the 1870 census, just after the Civil War, included counts of African-Americans for the first time.¹⁷ The census data provided counts, by race, place of birth and place of residence, making it straightforward to study migration patterns of people.

The newly freed African-Americans faced impossibly difficult times. They were desperately poor, largely illiterate, and living in a region whose plantation economy offered little chance for improvement. Most worked as sharecroppers, tenant farmers, or farm laborers. In addition to economic woes, the post Reconstruction South was racist and violent: Jim Crow laws instituted *de jure* racial segregation, and racist organizations like the Ku Klux Klan, the White League, and the Red Shirts made terror a part of daily life for African-Americans living in the post-Civil War Confederate South. Thus, it isn't surprising that large numbers of African-Americans opted to migrate to the North for greater opportunities. World War I created a huge demand for workers in northern factories; northern railroads needed workers so badly they provided free train passes for thousands of southern blacks.

¹⁷Previously the census included slaves as merely counts within various categories (e.g., age and sex) as part of the household goods of their owners.

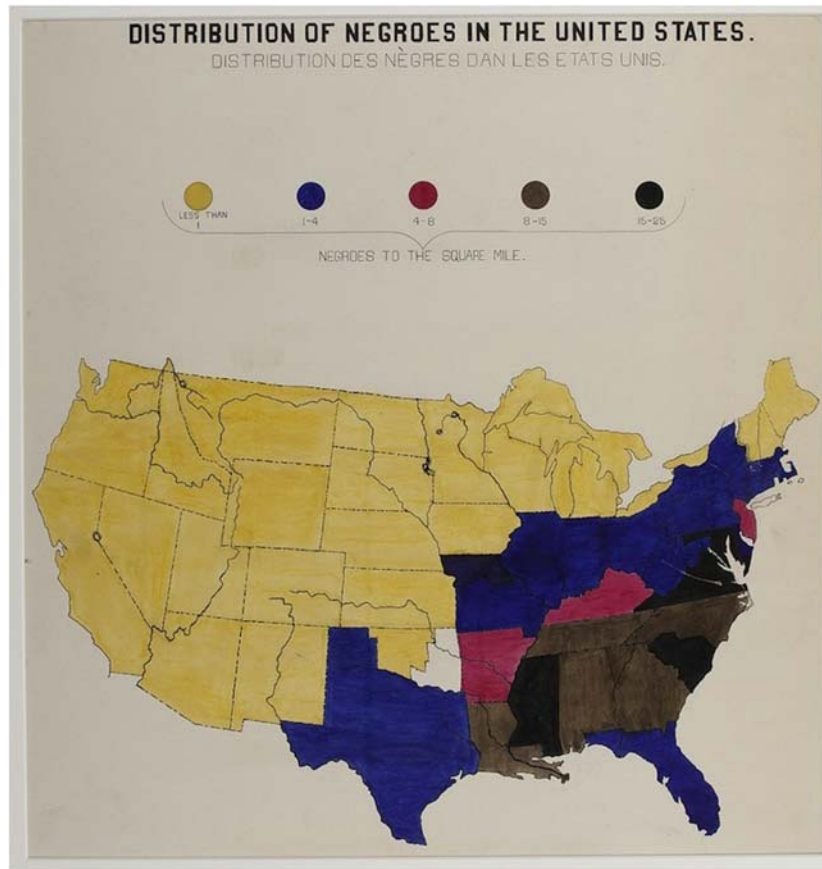


Fig. 18 Population density map: the distribution of the population of Negroes in the United States. Shading color shows “Negroes per square mile.” Source: Library of Congress, <http://hdl.loc.gov/loc.pnp/ppmsca.33900>.

A gedanken collaboration

Let us now imagine a collaboration between Du Bois, who chose the topic on which we focus our attention, the Census, which supplied the data, and Minard, who invented and perfected the graphical method that would allow us to see the character of the Great Migration, which, over the 55 year span from 1915 until 1970 saw more than six million African-Americans leave the South.¹⁸

Shown in **Fig. 20** is a compound plot of four flow maps drawn in the style of C.J. Minard. The four panels are spaced 20 years apart, showing the movement of non-whites (the census designation, which consisted primarily of African-Americans). The arrows show the net migration (outflow MINUS inflow), and the widths of the arrows are proportional to the number of migrants. The panel at the bottom depicts the annual number of lynchings of African-Americans; 94% of these were in the former Confederacy. We are using this variable in the same way that Minard used temperature in his map of Napoleon’s march—as one possible causal variable.

The left-most panel of **Fig. 20** (for 1880) shows the beginning of the post-reconstruction exodus. Most of the migrants left the South for the industrial cities of the North and Northeast (Chicago, Detroit, Pittsburgh, D.C.). The number of lynchings increased, peaking in 1890, and in 1900 the migration more than doubled. The need for labor expanded during World War I and is manifested by the further increases shown in the 1920 panel, when 454,000 left the South in the decade 1910–1920, and the 1940 panel, when another 1,500,000 left during the Depression and the build-up to World War II. Another principal cause of this part of the migration was the great Mississippi Flood of 1927, in which more than 200,000 African-Americans were displaced from their homes on the lower Mississippi River.

The Great Depression affected huge numbers of people, both White and Black, many of whom lost their homes (particularly in the Dust Bowl states) and left for a better life elsewhere. This is shown for African-Americans in the top panel of **Fig. 21** (an enlargement of the 1940 panel of **Fig. 20**). The bottom panel (1940 data for Whites) provides a dramatic contrast. Although African-Americans fled the South in the decades of the Great Depression in much the same patterns as they did in the decades before and after, the

¹⁸This section borrows from Andrews and Wainer (2017).

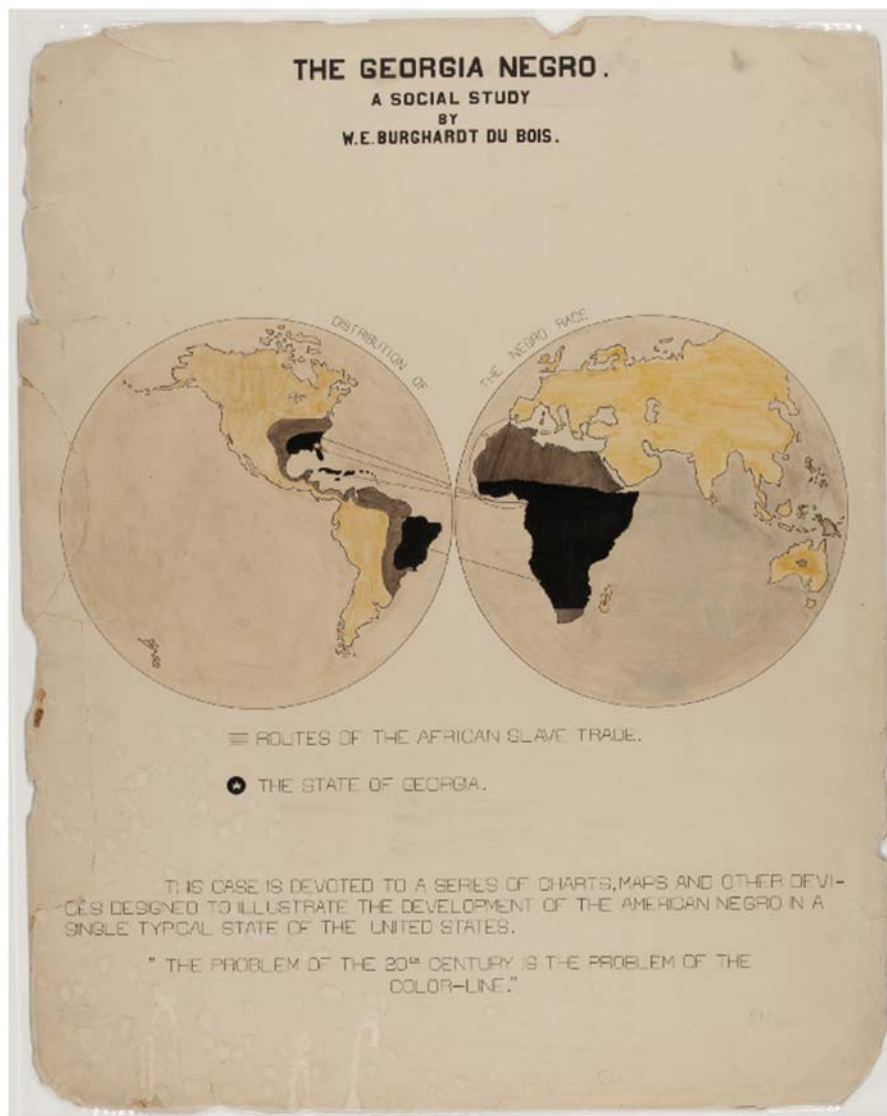


Fig. 19 The routes of the African slave trade. Source: Library of Congress, <http://hdl.loc.gov/loc.pnp/ppmsca.33866>.

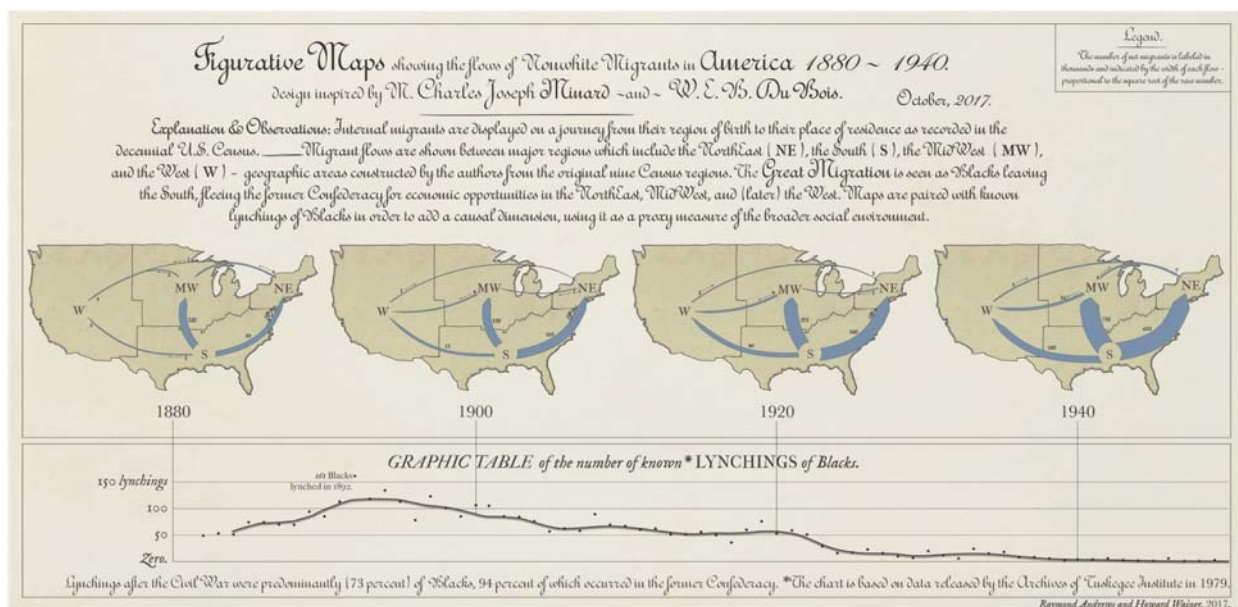


Fig. 20 Flow maps of migration: figurative maps showing the flows of non-white migrants in America, 1880–1940, using a design inspired by Minard and Du Bois. Source: Graphic designed and executed by Raymond J. Andrews & Howard Wainer reprinted with their permission.

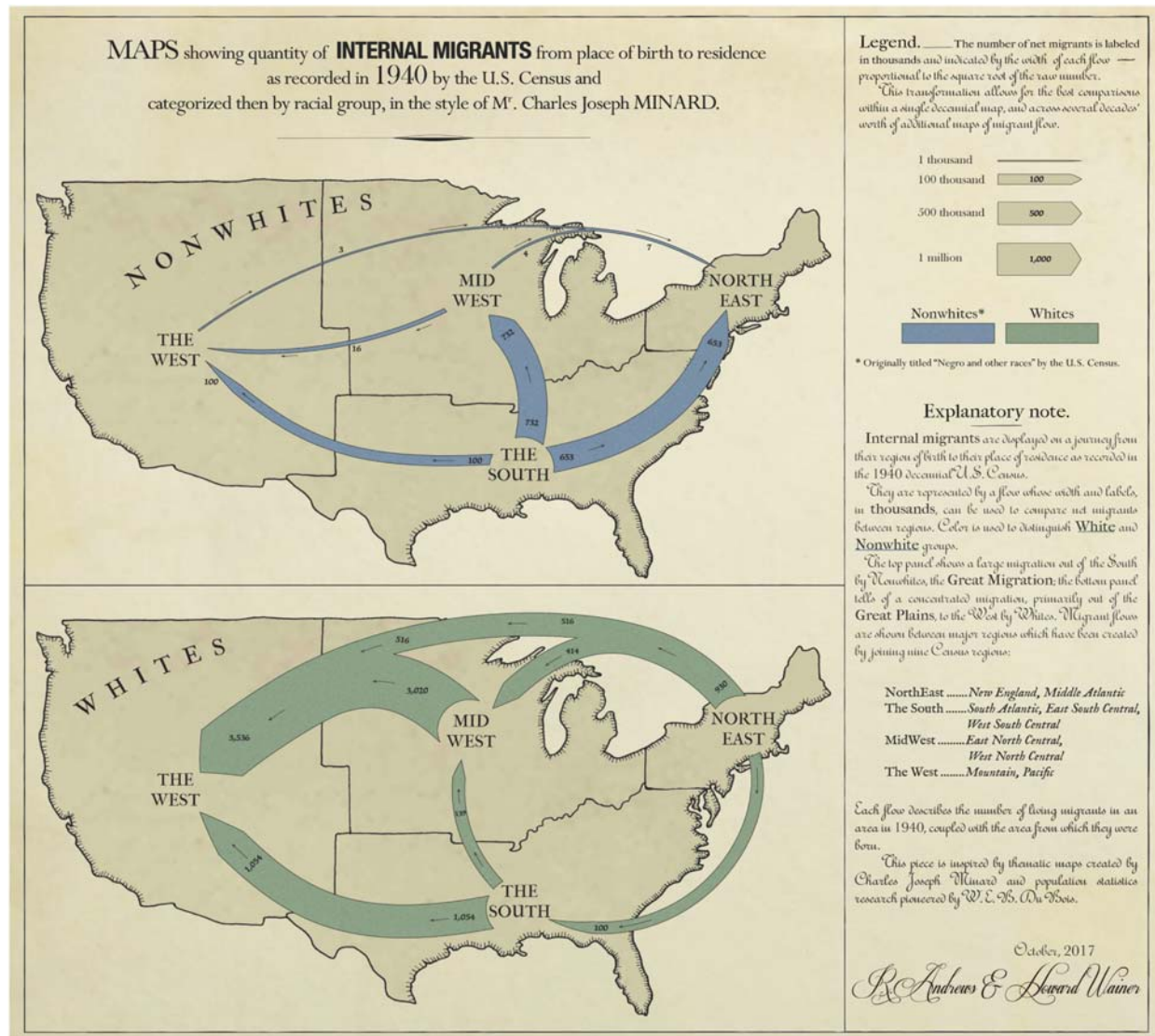


Fig. 21 Flow maps for 1940: maps showing the numbers of internal migrants by birthplace and place of residence, as recorded in the 1940 US Census, categorized by racial group, using a design inspired by Minard. Source: Graphic designed and executed by Raymond J. Andrews & Howard Wainer and reprinted with their permission.

Whites were often farmers whose economic lives were ruined by the Dust Bowl catastrophe (think of the Joads in Steinbeck's *Grapes of Wrath*, who piled all their belongings onto the family truck and headed West to California).

Conclusion

So we beat on, boats against the current, borne back ceaselessly into the past.

Nick Carraway's haunting last line from F. Scott Fitzgerald's *The Great Gatsby*.

This article, in a very real sense, provides a history and framework for modern science. We have found that, whenever we face a new problem it is always wise to look backward first; this provides the benefit of helping us to understand the thinking of those who have preceded us. The hubris that drives some to believe that we are wiser than our forebears rarely yields a happy result. This article is driven by the belief that we are best prepared to move into the unknown future by better understanding the past. In addition to illuminating the past, we wished to draw attention to the extraordinary power of visual depiction of quantitative phenomena to communicate efficiently both facts and emotion—how a graph can be a poem.

To illustrate these points, we have provided a portrait of the Great Migration of African-Americans from the South of the United States in the century following their emancipation from slavery. Sensible space limitations forced us to make some simplifications. In doing so we found that Mies van der Rohe's dictum "less is more" was spectacularly true and faced some of the same choices for level of detail vs. fidelity that Du Bois and Minard encountered in their work.

Looking at decade-by-decade plots added little that was not more clearly shown by collapsing to 20-year periods; showing migration destinations to the nine census regions, rather than the four super-regions we constructed added more noise than structure; including a 1960 plot that contains the final stage of the exodus, when almost 3.5 million left the South for the North and the West adds quantity to what we have already shown, but does not change the overall message. Thus, we believe that this summary is both accurate and complete, and provides a suitable continuation of the task that W.E.B. Du Bois began more than a century ago.

Du Bois's graphical narrative showed facts about African-Americans; how they started out and what they became. We added some details of how they got there and where they came from. The tool we chose to do this we snatched from the past, mining the ideas of the great C.J. Minard. We shamelessly borrowed his metaphor of the flow of goods or people across a geographic background. Minard's adjoining the delicate curve of declining temperature with the rapid shrinkage of Napoleon's army in their death march across the Russian steppes in the winter of 1812 forces the viewer to share the great sadness that enveloped all of France. Our use of the ancillary variable of lynchings was meant to generate a similar emotion, as one empathizes with the terror that African-American southerners must have felt as they made the decision to leave their homes and seek to rebuild their lives elsewhere.

It doesn't take an excessive amount of imagination to connect this same methodology to the Jewish diaspora from Nazi Germany, or the Hebrews from Pharaoh's Egypt, or peasants during the eastern European pogroms, or the Cherokees' Trail of Tears, or the Bataan Death March. We will leave it to others to prepare graphic accounts so that we can better understand and remember the facts and the emotions of these tragic events. For poetic data visualizations, combining both the communication of facts that can be learned, with emotions that can live in your heart forever can help provide a bulwark against future horrors, for "*there is a magic in graphs. The profile of a curve reveals in a flash a whole situation—the life history of an epidemic, a panic, or an era of prosperity. The curve informs the mind, awakens the imagination, convinces.*"^{19,20}

References

- Andrews, R.J., Wainer, H., 2017. The Great Migration: a graphics novel featuring the contributions of W.E.B. Du Bois and C.J. Minard. *Significance* 14 (3), 14–19.
- Arbutnot, J., 1710. An argument for divine providence taken from the constant regularity in the births of both sexes. *Philos. Trans. Roy. Soc.* 27, 186–190.
- Bertin, J., 1973. *Sémiologie Graphique*. The Hague: Mouton-Gautier. In: Berg, W., Wainer, H. (Eds.), *Semiology of Graphics*, second ed. University of Wisconsin Press, Madison, Wisconsin. 1983.
- Braun, H., Wainer, H., 2004. Numbers and the remembrance of things past. *Chance* 17 (1), 44–48.
- Brinton, W.C., 1939. *Graphic Presentation*. Brinton Associates, New York.
- Chevallier, V., 1871. In: Finley, D. (Ed.), *Notice nécrologique sur M. Minard, inspecteur général des ponts et chaussées, en retraite*. *Annales des ponts et chaussées*, 5e série; 2e semestre, pp. 1–22. <https://www.edwardtufte.com/tufte/minard-obit>.
- Cleveland, W.S., 1994. *The Elements of Graphing Data*. Hobart Press, Summit, New Jersey.
- Daston, L., Gallison, P., 2007. *Objectivity*. Zone Books, New York.
- Du Bois, W.B., 1900. The American Negro at Paris. *Am. Monthly Rev. Rev.* 22 (5), 576.
- Feynman, R.P., 1965. The Character of Physical Law (Intro quote from Chap. 7, "Seeking New Laws." p. 156). MIT Press, Cambridge, MA.
- Fletcher, J., 1847. Moral and educational statistics of England and Wales. *J. Stat. Soc. Lond.* 10 (3), 193–233.
- Fletcher, J., 1849b. Summary of the Moral Statistics of England and Wales. Privately printed and distributed.
- Friedel, H., Dickey, T., 1998. Hans Hoffman. Hudson Hills Press, New York.
- Friendly, M., Wainer, H., 2021. A History of Data Visualization and Graphic Communication. Harvard University Press, Cambridge.
- Friendly, M., 2002. Visions and re-visions of Charles Joseph Minard. *J. Educ. Behav. Stat.* 27 (1), 31–51.
- Graunt, J., 1662. *Natural and Political Observations on the Bills of Mortality*. John Martyn and James Allestry, London.
- Harness, H.D., 1837. Atlas to Accompany the Second Report of the Railway Commissioners. Irish Railway Commission, Ireland, Dublin.
- Huygens, C., 1669. *Oeuvres Complètes, Tome Sixième Correspondance*, 515–518, 526–539. Martinus Nijhoff, Den Hague.
- Lemann, N., 1991. *The Promised Land: The Great Black Migration and How It Changed America*. Knopf, New York.
- Marey, E.J., 1878. *La Méthode graphique dans les sciences expérimentales et particulièrement en physiologie et en médecine*. Paris.
- Mumford, L., 2000. *Art and Technics (The Bampton Lectures in America)*. Columbia University Press, NY.
- Panckoucke, C.-J., 1783. *Encyclopédie Méthodique Marine*. Chez Panckoucke, Paris.
- Robinson, A.H., 1982. *Early Thematic Mapping in the History of Cartography*. University of Chicago Press, Chicago.
- Tufte, E.R., 1983. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.
- Tufte, E.R., 1990. *Envisioning Information*. Graphics Press, Cheshire, CT.
- Tufte, E.R., 1996. *Visual Explanations*. Graphics Press, Cheshire, CT.
- Tukey, J., 1977. *Exploratory Data Analysis*. Addison-Wesley, Boston.
- Wainer, H., 1984. How to display data badly. *Am. Statistician* 38, 137–147.
- Wainer, H., 1993. Graphical answers to scientific questions. *Chance* 6 (4), 48–50.
- Wainer, H., 1997. *Visual Revelations: Graphical Tales of Fate and Deception From Napoleon Bonaparte to Ross Perot*. Copernicus Books, New York (Reprinted in 2000, Hillsdale, NJ: Lawrence Erlbaum Associates).

¹⁹Nick Carraway's haunting last line from F. Scott Fitzgerald's *The Great Gatsby*.

²⁰Henry D. Hubbard in the preface to Brinton (1939).

- Wainer, H., 2000. *Visual Revelations: Graphical Tales of Fate and Deception From Napoleon Bonaparte to Ross Perot*, second ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Wainer, H., 2005. *Graphic Discovery: A Trout in the Milk and Other Visual Adventures*. Princeton University Press, Princeton, NJ.
- Wainer, H., 2009. *Picturing the Uncertain World: How to Understand, Communicate and Control Uncertainty Through Graphical Display*. Princeton University Press, Princeton, NJ.
- Wainer, H., 2012. Moral statistics and the thematic maps of Joseph Fletcher. *Chance* 25, 43–46.
- Wainer, H., 2014. *Medical Illuminations: Using Evidence, Visualization & Statistical Thinking to Improve Healthcare*. Oxford University Press, London.
- Wylid, J., 1815. Some early demographic maps. In: Jarcho, S., 1973 *Bull. N. Y. Acad. Med.* 49, 837–844.

Meta-analysis

Betsy Jane Becker^a and Christopher G. Thompson^b, ^a Department of Educational Psychology and Learning Systems, College of Education, Florida State University, Tallahassee, FL, United States; and ^b Department of Educational Psychology, College of Education, Texas A & M University, College Station, TX, United States

© 2023 Elsevier Ltd. All rights reserved.

What is meta-analysis?	843
Problem formulation	843
Theory and frameworks in problem development	843
Inclusion and exclusion rules	844
Examples of problem formulation	844
Literature search	844
Data evaluation	845
Computing effect sizes	846
Mean differences: the d family of effect sizes	846
Correlational effects: r and associated measures	847
Count based effect sizes	847
Other effect sizes	847
Measurement error	847
Coding theoretically important variables	847
Coding study quality and methodological variables	848
Reliability of coding	848
Data analysis	848
Notation and initial graphics	848
Notation	848
Initial graphics	849
Example	849
Fixed-effect or common-effect model	849
Model	849
Example of fixed-effect-mean estimation	850
Investigating between-studies variation	850
Estimating τ^2	850
The index I^2	850
Testing for homogeneity	851
Example of between-studies variance assessment	851
Random-effects model	851
Model	851
Investigating the mean effect	851
Example of random-effects-mean estimation	852
Explaining variation in effects	852
Meta-regression	852
Example of meta-regression	852
ANOVA-like models	853
Example of ANOVA	853
Publication bias	853
Detection of publication bias	853
Adjusting for publication bias	854
Other summaries	854
Advanced meta-analysis methods	856
Dependence in meta-analysis	856
Bayesian approaches	856
Interpretation and reporting of results	856
References	857

We begin by introducing the idea of meta-analysis. We then present the steps in the process of conducting a meta-analysis of summary data, with special attention to statistical analyses. We illustrate each step using recent meta-analyses of studies of

game-based instruction for second-language learning. We also briefly mention advanced meta-analysis techniques including multi-variate meta-analysis and Bayesian meta-analysis.

What is meta-analysis?

Meta-analysis, or the “analysis of analyses” is a term created by [Gene Glass \(1976\)](#). Meta-analysis is the systematic accumulation and analysis of results from series of studies addressing a common question. Glass was motivated to coin this term and create procedures for summarizing results by the vast numbers of disparate empirical studies on the use of psychotherapy. He wanted to statistically analyze the large set of results on psychotherapy, and integrate them in a more rigorous fashion than he could achieve with a traditional narrative review. Meta-analysts assess typical study findings and also explore why findings vary across studies. Many scholars use the word “meta-analysis” to mean the entire process of integrating results of related studies, whereas others use the terms research synthesis, integrative review, or systematic review for this purpose. We use the term in the broader sense in this chapter.

Glass revolutionized the synthesis of treatment studies by using effect sizes to represent study findings. An effect size is an index that is comparable across studies and captures information about the research question of interest. Effect sizes may not be needed if complete data sets are available, as in individualized-participant-data (IPD) meta-analysis ([Cooper and Patall, 2009](#)). The term effect size often refers to the standardized mean difference, but many other indices (correlations, proportions, etc.) can serve as effect sizes in meta-analyses. (Effect sizes are discussed further below.)

Glass’s first meta-analyses concerned the effectiveness of psychotherapy and the impact of class size on student achievement, and he used a standardized mean difference as the effect size, specifically

$$g_i^G = \frac{(\bar{Y}_i^T - \bar{Y}_i^C)}{S_i^C}, \quad (1)$$

where $\bar{Y}_i^T$ is the mean of the treatment group in study i on some outcome Y , $\bar{Y}_i^C$ is the analogous mean from the control group, and S_i^C is the standard deviation (SD) of Y for the control group in study i . This index resembles a z score, except that its numerator is a difference between means, not the difference between a score and a mean. Thus g_i^G can be interpreted as the number of standard deviations of difference between the two group means, in the control-group’s score scale. The population analogue to g_i^G is the parameter $\delta_i^G = (\mu_i^T - \mu_i^C)/\sigma_i^C$ where μ_i^T and μ_i^C are the treatment and control population means, respectively, and σ_i^C is the SD of the control population.

Meta-analyses have the same components as primary research ([Jackson, 1980](#); [Cooper, 1982, 1989, 2017](#)). A simple version of Cooper’s steps in the meta-analysis process includes:

1. Problem formulation, or the crafting of a clear question about a relationship, group difference, or other effect that typically specifies which study characteristics may explain patterns of results across studies;
2. Literature search, where a systematic search of the literature uses explicit rules to identify relevant studies;
3. Data evaluation, including calculation of effect sizes and extraction of study features as well as ongoing evaluation of study relevance;
4. Data analysis, where statistical techniques are used to relate effect sizes to features such as sample characteristics, treatment intensity, methods of study, etc.;
5. Interpretation of synthesis results, to translate the results; and
6. Public presentation, via published manuscript or some other form of report.

Other descriptions of the process use different labels for these steps or add steps.

Problem formulation

A successful meta-analysis is based on specific questions that are not so broad as to be unanswerable, or so narrow that few studies address them. Consider the question “Does the use of educational technology improve learning?”. Many kinds of technology exist, and it is likely they do not all work for all students or all kinds of learning. We may want to specify a content area (e.g., math learning or second-language acquisition), and perhaps an age level or range of grades for the learners. Should the focus be narrowed, to mobile learning or game-based interventions? Should we study learning inside or out of school? Each consideration allows the meta-analyst to narrow or broaden the scope of the review and the degree to which its conclusions can be generalized.

Theory and frameworks in problem development

Relevant theory, a practical framework, or a scoping review can support good question development. Prior reviews (if any) and scoping reviews ([Munn et al., 2018](#)) can identify how many studies exist, clarify constructs examined therein, and possibly identify gaps in knowledge in a domain.

A general framework such as [Cronbach's \(1982\)](#) UTOS framework for generalizability may also be useful. UTOS stands for Units, Treatments, Outcomes (originally "observing operations"), and Settings. Specifying each of these limits the domains to which meta-analysis inferences may apply. In medicine the PICO framework (for Patient/Problem, Intervention, Comparison, and Outcomes; [Richardson et al., 1995](#)) has become a popular way to develop "well-built" questions ([1995](#), p. A12). [Becker \(1996\)](#) and [Matt and Cook \(2009\)](#) have applied the UTOS framework to the meta-analysis context. [Becker \(1996\)](#) pointed out that the factors that lead to differences in study results often reflect the units, treatments, etc., that were studied. These may be dictated by a theory of an intervention, a model of interrelationships, or practical considerations.

Attributing differences in study results to characteristics of treatments, participants, or settings is a key part of a meta-analysis. Considering these components up front focuses the literature search, study screening, data extraction, and analysis. Selecting these features is both a conceptual and practical task. At the theoretical level one must first define the constructs of interest – such as language learning in our example. Then the meta-analyst must operationalize those constructs as part of defining inclusion rules. Clear operationalization is important not only for outcome constructs (e.g., tests of language skill), but also for treatments (e.g., What kinds of learning games should be examined? What components must they include?) and other predictors. This task can be challenging when reviewing studies of complex interventions ([Higgins et al., 2019](#)), or when many factors influence the outcomes of interest (e.g., [Brown et al., 2016](#)).

The degree to which the meta-analyst wants to generalize beyond their collected studies also determines what statistical models should be used. If a meta-analyst only wants to make inferences about the studies at hand (i.e., does not wish to generalize) the fixed-effect model can be used. Fixed-effect models are also appropriate when all studies share a common parameter value, or when a set of predictors explains all of the variation in the study outcomes ([Hedges and Vevea, 1998](#)). In contrast random-effects models assume that the true effects vary across studies and are sampled from a theoretical population of parameter values. In spite of the fact that the studies in meta-analyses are rarely randomly sampled, random-effects models are very common in current practice. Other details of these two models are discussed below.

Inclusion and exclusion rules

Inclusion and exclusion rules developed in the problem-formulation stage help the reviewer define the limits of the search (useful at step 2), and of later generalizations. Setting criteria for participant characteristics, intervention features, methods used, and publication-date limits can contract the size of the relevant literature and ensure that collected studies will be useful. Using too many exclusion rules may lead to very small sets of studies; here scoping reviews can help in planning the details of the meta-analysis.

The nature of the question asked determines which effect sizes should be used. Questions typically imply a target effect size that requires certain data (e.g., means and SDs, correlations); these can be listed in the inclusion rules. Whether the target effect size can be calculated in all studies depends on what is reported. Differences in methods used in primary studies, such as the presence or absence of a control group, or the type of sample used, affect what can be computed as an effect magnitude and impact study validity. Our question on the effectiveness of educational technology did not specify whether the technology would be compared to another means of instruction (e.g., to non-technology-based instruction). However, two-group studies provide stronger causal inferences and better internal validity than one-group designs, thus bolstering meta-analytic inferences. Our question asked about improvement in learning, implying that we should collect studies of change in performance. Methods considerations like these should be specified as part of problem formulation.

Examples of problem formulation

Two examples illustrate problem formulation in meta-analysis. [Thompson and von Gillern \(2020\)](#) reviewed studies of game-based instruction for vocabulary learning. They examined only English language learners, but the native language of the learners was not restricted. Participants ranged in school level from kindergarten through college. The authors required that studies use digital video games (the T or treatment) on several platforms, versus some alternative form of instruction. Their outcome (the O) was vocabulary acquisition, and the settings were largely school based. Both experiments and quasi-experiments were included.

In contrast [Cho et al. \(2018\)](#) studied technology for second-language learning, but only on mobile devices – a different "T". Cho et al. did not require that the mobile-learning instruction involve games, and they examined a broader set of Os: vocabulary plus pronunciation skills and general reading and language-arts outcomes. They also required that all studies have a control (non-mobile) condition such as textbook- or paper-based vocabulary lessons. Last, Cho et al. included only experiments. Only one study ([Wu and Huang, 2017](#)) appeared in both reviews.

Literature search

Once a suitable problem is in hand, a thorough, unbiased, and replicable search is conducted for both published and unpublished documents. Ideally the meta-analyst will find all available sources, but that can be hampered by publication bias, discussed below. Databases should always be used in the search. Inspecting the thesaurus of search terms for each database helps identify keywords for the most effective search logic. In our view it is critical to search for dissertations and theses. They are rich with data, and while

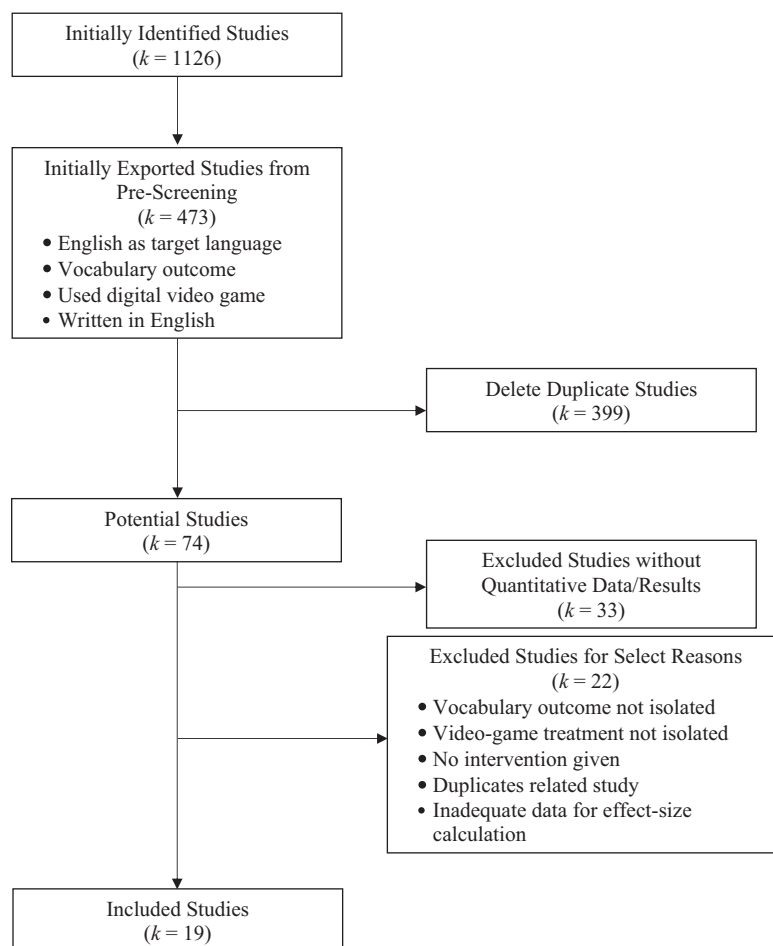


Fig. 1 PRISMA flow chart. From Thompson, C.G., von Gillern, S., 2020. Video-game based instruction for vocabulary acquisition with english language learners: a Bayesian meta-analysis. *Educ. Res. Rev.* 30, 100332.

technically unpublished have usually been thoroughly vetted. Also, because they are not formally published, the threat of publication and reporting bias is less.

Existing reviews may also be a source of studies, depending on the overlap in the review questions and any date restrictions. Online searching (e.g., via google, google scholar, and other web sites), writing to relevant listservs, and contacting scholars active in a field can also help identify unpublished sources. For some questions that may have been of secondary importance to the original research (e.g., gender differences) it may be useful to scan issues of relevant journals for articles not tagged with appropriate key words.

Once the search is done, the titles, abstracts, or full texts of identified sources must be screened for appropriateness. This avoids wasting time extracting information from studies without data or that do not meet inclusion and exclusion rules. During this process one must track deleted studies and the reasons for their deletion so that a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart (Page et al., 2021) can be created. Fig. 1 shows the flow chart from Thompson and von Gillern's review of game-based language-learning studies (2020, p. 5).

Meta-analysts routinely screen thousands of sources to cull out unsuitable studies. Recent advances in searching and screening techniques automate these processes. Software products such as RobotSearch and Abstrackr (Wallace et al., 2012, 2014), and SWIFT-Review (Howard et al., 2016), among others, speed the work of screening. Marshall and Wallace (2019) describe these and other tools for automating research syntheses.

Data evaluation

Data evaluation includes both data extraction and evaluation. The meta-analyst, optimally with a team of assistants, computes effect sizes to represent study results and extracts or "codes" features of the studies. These features represent variables that may relate to study results, or that one hopes do not relate to results.

Computing effect sizes

Kelley and Preacher (2012, p. 140) defined effect size as “a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest”; they extensively discuss effect-size properties. The effect sizes for any meta-analysis should flow naturally from the question driving the synthesis. However, variations in study design lead to a diversity of possible effect sizes, and the meta-analyst must choose among them based on the goals of the meta-analysis and the data available. Consider our question about improvements in learning due to educational technology. Our effect size must incorporate measures of language skill taken before and after the intervention. Studies that measure language only post intervention would be useless. Preliminary knowledge of the study designs used in the domain under review will help the meta-analyst identify appropriate effect sizes.

To cover all potential effect sizes would require several book chapters, thus we mention only the most common effect sizes. Each has a standard error that can be computed with only one observed instance of the effect size. Thus, meta-analysts can conduct analyses that (1) account for each study’s uncertainty, and (2) provide measures of model fit, unlike most analyses learned in introductory statistics classes.

We use the symbols T_i and θ_i to refer to the effect size of interest in study i and the parameter it is estimating and v_i to denote the variance of each T_i . We use this notation as we introduce each of the effects in this section, and below as we discuss analyses.

Mean differences: the d family of effect sizes

In Eq. (1) we introduced the effect size used by Glass (1976) to represent average differences between two groups. Hedges (1981; Hedges and Olkin, 1985) argued that if one accepts the typical assumption that the population variances of the two groups are equal to σ_i^2 (i.e., the two-sample t -test homoskedasticity assumption), a better estimator is

$$T_i = g_i = \frac{(\bar{Y}_i^T - \bar{Y}_i^C)}{S_i^P}. \quad (2)$$

Here the denominator is the pooled within groups standard deviation

$$S_i^P = \sqrt{\frac{(n_i^T - 1)[S_i^T]^2 + (n_i^C - 1)[S_i^C]^2}{(n_i^T + n_i^C - 2)}},$$

where n_i^T and n_i^C are the sample sizes for the treatment and control groups, respectively, and $\bar{Y}_i^T$, $\bar{Y}_i^C$, S_i^T , and S_i^C are defined for (1) above. Hedges noted that in small samples g_i is a biased estimator of the population effect size $\delta_i = (\mu_i^T - \mu_i^C)/\sigma_i$. He derived the approximate unbiased estimator

$$d_i = J(m_i)g_i, \quad (3)$$

where $J(m_i) = 1 - 3/(4m_i - 1)$ and $m_i = n_i^T + n_i^C - 2$, the degrees of freedom of the two sample t test. Hedges also provided a large-sample asymptotic variance for d_i , which is

$$v_i = \text{Var}(d_i) = [J(m_i)]^2 \left[\frac{n_i^T + n_i^C}{n_i^T n_i^C} + \frac{\delta_i^2}{2(n_i^T + n_i^C)} \right], \quad (4)$$

with d_i being substituted for δ_i in practice.

Other study designs necessitate the use of different effect sizes for mean differences; Lin and Aloe (2020) provide a good overview. For example, Becker (1988) proposed an effect size for pre-post designs, $T_i = g_i^{\text{Change}} = (\bar{Y}_i - \bar{X}_i)/S_{X_i}$, where X and Y are pretest and posttest measures, respectively, and S_{X_i} is the pretest standard deviation in study i . Hedges (2011) and Lai and Kwok (2014, 2016) suggested ways to compute effect sizes for multi-level analyses. Keef and Roberts (2004) developed a partial standardized-mean-difference effect size which divides a mean difference adjusted for a covariate by the square root of the residual standard error. This latter estimator will be larger than an effect computed using Eq. (2) if the covariate reduces residual error.

The key issue for meta-analysts as they choose among the available estimators is whether effects will be comparable across studies. Sometimes effects from different designs can be combined, especially if adjustments are made to enhance comparability. For example, d values from t tests may be combined with effects from cluster randomized trials if the latter have been standardized using a measure of within-cluster variation (Hedges, 2007). Similarly, a partial standardized mean difference that controls for socioeconomic status (SES) in a broad national sample may be comparable to an effect size with no covariate control from a sample with participant SES levels in a narrow range (e.g., a low SES sample). Sometimes, though, disparate designs inform us about very different phenomena, and in such cases their effects should not be combined with the two-sample d_i above. For one, indices measuring mean change for a single group (like g_i^{Change}) are best synthesized separately from the two-sample d given in Eq. (3) and other two-group effects.

Correlational effects: r and associated measures

Another common effect size is Pearson's correlation, $T_i = r_i$, which estimates the population correlation ρ_i . Bivariate correlations are widely reported and have a simple variance,

$$v_i = \text{Var}(r_i) = \frac{(1 - \rho_i^2)^2}{n_i - 1}.$$

This variance decreases as ρ_i approaches either -1 or 1 , and extreme ρ_i produce skewed sampling distributions of r . Many meta-analysts therefore analyze the variance-stabilized [Fisher \(1921\)](#) transform of r ,

$$T_i = z_i = 0.5 \ln \left(\frac{1 + r_i}{1 - r_i} \right),$$

which has variance $v_i = 1/(n_i - 3)$. Many sources cover the synthesis of correlations, including [Field \(2001\)](#) and the writings of [Hunter and Schmidt \(2004\)](#).

Meta-analysts have also summarized partial and semi-partial correlations (e.g., [Polanin et al., 2021](#)) and correlation matrices. Methods for estimating and analyzing partial indices are covered by [Aloe \(2014\)](#), [Aloe and Becker \(2012\)](#), [Aloe and Thompson \(2013\)](#), and [Becker and Wu \(2007\)](#). Summaries of correlation matrices are now also common ([Becker, 1992, 1995](#); [Becker and Aloe, 2019](#)) as is the synthesis of data on complex models (e.g., [Brown et al., 2015, 2016](#)).

Count based effect sizes

Studies in medicine often deal with counts or percentages of patients who are ill versus healthy. [Borenstein and Hedges \(2019\)](#) give an extensive overview of available effects for such data. Count data are often analyzed as proportions or "risks" $T_i = p_i = n_{i1}/n_i$ or logits $\ln [p_i/(1 - p_i)]$, where the count of events of interest is denoted as n_{i1} and the sample size is n_i . The variance of p_i is $v_i = \pi_i (1 - \pi_i)/n_i$, which has its maximum at $\pi_i = 0.5$ but can be much smaller near 0 and 1. The logit has a distribution that is asymptotically more normal than that of p_i .

When count data arise from two groups, the meta-analyst has many options for handling the counts in a synthesis. The risk difference ($p_i^T - p_i^C$), risk ratio (p_i^T / p_i^C) (or its log, $\ln [p_i^T / p_i^C]$), or the odds ratio $[p_i^T / (1 - p_i^T)] / [p_i^C / (1 - p_i^C)]$ can be used. The log-odds ratio (LOR) is most common due to its superior statistical properties ([Borenstein and Hedges, 2019](#)). It is computed as

$$T_i = \text{LOR}_i = \ln \left(\frac{p_i^T / (1 - p_i^T)}{p_i^C / (1 - p_i^C)} \right).$$

Denoting the number of cases with the event of interest as n_{i1} and the number of non-event cases as $n_{i0} = n_i - n_{i1}$, the variance of the LOR can be computed as

$$v_i = \text{Var}(\text{LOR}_i) = \frac{1}{n_{i1}} + \frac{1}{n_{i0}^T} + \frac{1}{n_{i1}^C} + \frac{1}{n_{i0}^C}.$$

Other effect sizes

Other effect indices may be relevant for special situations. For example, one can summarize variances or their ratios, data from diagnostic tests such as sensitivities and specificities ([Chu et al., 2009](#)), or the area under the curve (AUC) index for the receiver operating characteristic curve. Survival-curve analyses (see [Combes et al., 2014](#)), models that predict risk of disease ([Debray et al., 2019](#)), and studies in genetics ([Kavvoura and Ioannidis, 2008](#)) each have their own effect indices.

We are not fans of the use of p values to represent effects, as they confound effect size with sample size. Only when commensurate indices cannot be found does it make sense to use p values.

Measurement error

A problem for many effect measures is measurement error or unreliability. Differential measurement error is a central concern in the literature on validity generalization – the synthesis of test validities (see, e.g., [Schmidt and Hunter, 1977](#)). If measures of reliability can be obtained for each study's measures, corrections can be applied. If not, other aspects of the measures used, such as test length, may be proxies for the degree of measurement error.

If $r_{YY'}$ is the reliability of outcome measure Y , a standardized mean difference d can be corrected using the formula $d^C = d / \sqrt{r_{YY'}}$. Dividing the correlation by the square root of the product of the reliabilities of both measures disattenuates the observed r . This correction always increases effect sizes, as unreliability lowers observed effect sizes and measures of association. Techniques to adjust for differential reliability (mentioned above) are sensible to use when reliabilities vary widely across studies. For further details see [Hunter and Schmidt \(2004\)](#).

Coding theoretically important variables

Features of substantive or theoretical interest should always be coded as potential explanatory variables for differences among studies. Variables capturing the type of treatment, treatment duration (e.g., [Thompson and von Gillern, 2020](#)) or intensity, or

treatment fidelity also often have theoretical significance. Meta-analyses of complex treatments may examine a set of key program components. Such study features may be correlated. Thus, the meta-analyst should check for collinearity before including multiple features as moderators. Participant characteristics such as age or grade, gender, and level of experience with content or degree of expertise (e.g., in a sport) are often coded.

Educational meta-analyses often examine tests of subject-matter content taught in an intervention, or other academic measures (such as grades), as well as psychological constructs such as motivation, attention, and the like. If the outcomes measured vary across studies, information should be coded so the meta-analyst can analyze whether mean effects differ across constructs or measures. [Cho et al. \(2018\)](#) coded four categories of language learning, whereas [Thompson and von Gillern \(2020\)](#) examined only vocabulary acquisition.

Coding study quality and methodological variables

Variables reflecting study methods and quality must be coded. Characteristics of how a study is conducted may be considered nuisance variables by some – but failing to check on their relationships to the effect size would be an error. Some methodological features will bring a critic's eye to the synthesis if left uncoded, with variables capturing risk of bias being especially critical for meta-analyses of interventions. Ignoring whether studies used random assignment overlooks a key aspect of internal validity. Similarly, failing to note the presence of attrition, confounded treatments, or evidence of group incomparability raises doubts about the quality of the synthesis. Variables like date of publication (to allow examination of time trends or changes in methods used), or author gender ([Eagly and Carli, 1981](#)) and affiliation allow for tracking of potential biases.

[Valentine and Cooper \(2008\)](#) developed the Study Design and Implementation Assessment Device which captures many aspects of quality in education studies, and very detailed risk-of-bias tools have been developed for medical syntheses (RoB2: [Sterne et al., 2019](#); ROBINS-I: [Sterne et al., 2016](#)). These tools do not fit all meta-analyses but provide key aspects of bias for meta-analysts to consider.

In our examples, Cho et al. differentiated standardized tests from researcher-made tests (which may be more targeted to the skills taught in the interventions); studies using researcher-made tests showed higher effects. Thompson and von Gillern were unable to fully code type of assessment due to poor reporting in their primary studies, but noted that about 25% of their studies used standardized instruments. These authors also coded whether each study had used random allocation to assign participants to groups. Neither meta-analysis corrected for attenuation.

Reliability of coding

The extraction of information for meta-analysis, like any other coding process, involves human judgment. Meta-analysts must enhance reliability by training their coders and providing them with clear instructions. In state-of-the-art syntheses two or more trained persons code all studies. The meta-analyst develops a coding sheet and instructions, then selects a few studies for coding practice. Coders discuss discrepancies, unusual cases, or potential additional variables before further coding. A check for coder drift, with a reliability check of each coded variable, may be done after an initial set of studies has been coded. After coding is complete, the reliability of each coded feature should be assessed and discrepancies resolved. Each feature's reliability should be reported in the meta-analysis write-up. For example, Cho and colleagues reported initial inter-coder reliabilities between .78 and 1; coder reliability was not reported by Thompson and von Gillern.

Data analysis

The data-analysis stage of the review often begins with assessments of the average size of the effect of interest, and of the degree to which effect sizes vary across studies. Variation can be estimated or tested, though initial homogeneity testing has become less common as more meta-analysts begin with the assumption that true study effects vary (i.e., they adopt random-effects models for their analyses). Statistical and graphical procedures help the meta-analyst to explore and explain the nature of their accumulated results.

Notation and initial graphics

Notation

We denote the effect size of interest as T_i , for $i = 1$ to k independent studies; each T_i estimates a parameter θ_i . The simplest fixed-effect model for the effect size within study i is

$$T_i = \theta + e_i, \quad (5)$$

where e_i reflects sampling variance resulting from the specific sample of cases examined within study i . For most common effect-magnitude indices the errors e_i are treated as approximately normal in large samples, that is, $e_i \sim \text{Normal}(0, v_i)$, for $i = 1$ to k , where v_i is the variance of the residual e_i , or of T_i conditional on θ . Under this model θ is a common effect size shared by all

populations under study. Formulas for v_i were given above. The v_i are typically treated as known, which is plausible because all v_i are asymptotically efficient, but this efficiency is not guaranteed for small samples.

Our examples use the studies of game-based instruction for second-language learning from Thompson and von Gillern (henceforth TvG; 2020). Treatment-control differences were represented with the standardized-mean-difference index $T_i = d_i$ shown in Eq. (3). Positive T_i occur when the game-based instructional group outscores the comparison group on a post-instruction language test.

Initial graphics

Because each effect size has its own standard error, a typical first step in a meta-analysis is to create a forest plot showing $100(1 - \alpha)\%$ confidence intervals (CIs; typically 95%) for the effect sizes under fixed effects using $T_i \pm z_{\alpha/2} \sqrt{v_i}$ with $\alpha = 0.05$. Either fixed- or random-effects CIs can be made, but we prefer fixed-effect CIs which give a better picture of the degree of between-studies variation because only within-study variation is captured in the CI bars.

Forest plots provide information on central tendency and a sense of how variable the effect sizes are relative to sampling error. Each study effect is marked with a box inversely proportional to study size, plus a wider interval for the CI that reflects sampling error. The mean effect (formulas given below) is shown as the center of a small diamond below the study CIs. The width of the diamond shows the 95% CI for the mean.

A “quick and dirty” test of agreement among the results asks whether any vertical line can be drawn across all the fixed-effect CIs in the forest plot. Often such a line falls near the mean effect. A vertical reference line may be added to show the point at which no difference exists for many T_i (for odds measures, a reference line may appear at 1). It is fairly unusual for meta-analyses to include histograms of results.

Example

Fig. 2 shows 95% fixed-effect confidence intervals in effect-size order for the standardized mean differences from TvG, specifically $T_i \pm 1.96\sqrt{v_i}$. Narrower intervals are from large studies, such as study 4. The TvG studies show positive effects overall, but they do not all show the same advantages for game-based instruction. Indeed, one effect size is -0.01 , and 8 intervals include zero, consistent with a null effect. In contrast the largest value shows almost a two standard-deviation difference in mean vocabulary scores. Also, no line can be drawn through all CIs for these studies, suggesting notable heterogeneity.

Fixed-effect or common-effect model

Model

The simplest model for a series of effect sizes is the fixed-effect (FE) model in Eq. (5), $T_i = \theta + e_i$, for $i = 1$ to k . All studies are assumed to have sampled from a single population and estimated one common θ . Given the diverse nature of studies in education and the social sciences this may be a rare event, but the fixed-effect model can also be adopted when the meta-analyst does not wish to generalize beyond the collected studies. Under this model all $\theta_i = \theta$, and variation among the T_i from different studies is presumed to be from sampling variation, which is captured by the v_i estimates. We can thus write

$$T_i \sim \text{Normal}(\theta, v_i), \text{ for } i = 1 \text{ to } k, \quad (6)$$

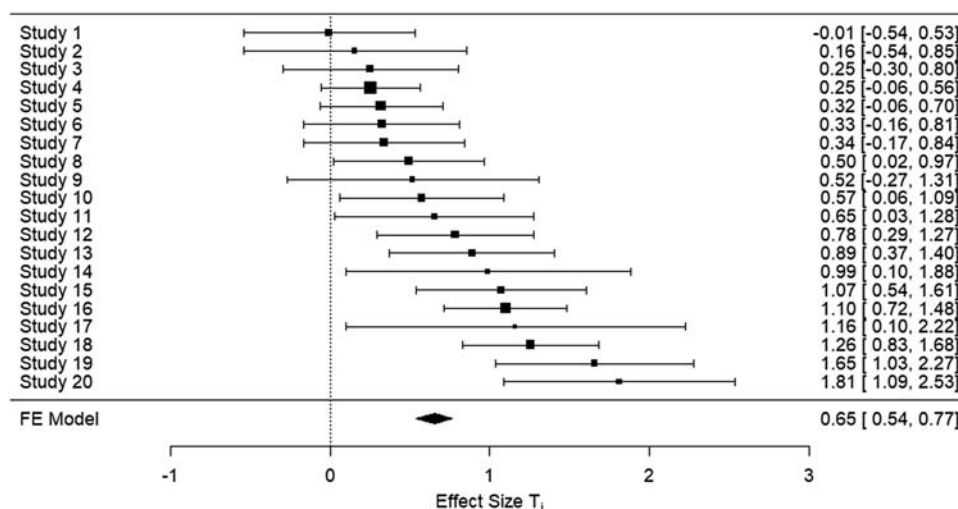


Fig. 2 The 95% confidence intervals for results of the TvG game-based learning studies.

and we can use either weighted-least-squares or maximum-likelihood methods to estimate θ as the minimum-variance unbiased inverse-variance weighted mean (Hedges, 1981) using

$$\bar{T}^{FE} = \frac{\sum_{i=1}^k w_i T_i}{\sum_{i=1}^k w_i}, \quad (7)$$

where $w_i = 1/v_i$ serves as the weight for study i . Because nearly all forms of v_i depend on the sample size n_i , the FE mean is very close to a sample-size weighted mean, and the influence of each effect size is greater when it is from a large sample.

The variance of $\bar{T}^{FE}$ is

$$\text{Var}(\bar{T}^{FE}) = \frac{1}{\sum_{i=1}^k w_i}. \quad (8)$$

The distribution of $\bar{T}^{FE}$ is also asymptotically normal, thus we can use the estimate and its variance to obtain a $100(1 - \alpha)\%$ confidence interval for the common θ as

$$\bar{T}^{FE} \pm z_{\alpha/2} \text{SE}(\bar{T}^{FE}). \quad (9)$$

Here the standard error of the mean is the square root of the variance in Eq. (8). A test of $H_0: \theta = 0$ is also available as $\bar{T}^{FE} \sqrt{\sum_{i=1}^k w_i}$ which is compared to critical values of the standard normal distribution. The **metafor** program for R (Viechtbauer, 2010) command `rma(T, v, method = "FE")` generates the values in Eqs. (8) and (9) and a set of associated statistics from the T_i and v_i .

Example of fixed-effect-mean estimation

For the 20 studies from TvG the fixed-effect estimate of the common population effect size is $\bar{T}^{FE} = 0.65$ standard-deviation units ($SE = 0.058$), leading to a 95% confidence interval for the common θ covering from 0.54 to 0.77. This interval does not include zero implying that game-based instruction is significantly superior to lecture-based instruction under the FE model.

Investigating between-studies variation

Because the forest plot of the TvG effect sizes shows heterogeneity, the effect sizes are probably not consistent across studies. In such cases if the meta-analyst is not devoted to using the fixed-effect model, they should model the between-studies differences reflected in the data.

Estimating τ^2

If the meta-analyst believes that the θ_i are not equal, the between-studies variance of the θ_i , denoted τ^2 , must be estimated. Many scholars have provided estimators of τ^2 ; Viechtbauer (2005) provides an excellent review of available methods. The earliest estimator (Hedges, 1983) is a simple method-of-moments estimator: Using the sample variance of the T_i (S_T^2) to reflect all observed variation, and the mean of the v_i ($\bar{v}$) to capture within-study variation, the between-studies variance is the difference

$$\hat{\tau}_{HE}^2 = S_T^2 - \bar{v}. \quad (10)$$

This estimator can take on negative values and thus it has been replaced by better-behaving estimators such as the restricted maximum likelihood estimator (REML, Thompson and Sharp, 1999), which is not shown here as it is iterative and cannot be written in closed form. We use the REML estimate below, obtained via `rma(T, v, method = "REML")` in **metafor**.

The index I^2

One other index of variation typically reported in meta-analyses is I^2 , a measure of the percent of observed effect-size variance that is true between-studies variation (Higgins and Thompson, 2002, Eq. 9). The program **metafor** calculates $I^2 = 100 \left[\frac{\hat{\tau}^2}{\hat{\tau}^2 + \bar{v}} \right] \%$, where $\bar{v}$ is a function of the v_i that captures sampling variation. Cutoffs of 25%, 50%, and 75% have been suggested for low, moderate, and high levels of true variation, but these are arbitrary.

Testing for homogeneity

Last, it is not necessary to test the hypothesis of no between-studies variance, that is $H_0: \theta_1 = \theta_2 = \dots = \theta_k$ or equivalently $H_0: \tau^2 = 0$, but it is easily accomplished using the weighted variance of the effect sizes (Hedges, 1983). The Q test is computed as

$$Q = \sum_{i=1}^k w_i (T_i - \bar{T}^{FE})^2 \quad (11)$$

where $w_i = 1/v_i$ and $\bar{T}^{FE}$ is the mean under the fixed-effect model from Eq. (7). The statistic Q follows an approximate chi-square distribution with $df = k - 1$ for k studies (though Hoaglin, 2016, points out problems with treating Q as a chi-square).

Example of between-studies variance assessment

We compute two variance estimates for the game-based learning studies. The method-of-moments estimator in Eq. (10) above (known as the HE estimator in *metafor*) is $\hat{\tau}_{HE}^2 = 0.157$ ($SE = 0.084$). The restricted maximum likelihood estimate is very similar: $\hat{\tau}_{REML}^2 = 0.155$ ($SE = 0.076$), with square root $\hat{\tau}_{REML} = 0.394$. If the θ_i values are normally distributed, most θ_i will fall in a range of about 1.54 standard-deviation units ($\pm 1.96 \times 0.394$) around the mean.

The value of I^2 for the TvG studies is 69%, near the benchmark for high variability. Over two-thirds of the observed variance in the effects reflects true differences among the populations, according to I^2 . Last, using Eq. (11), $Q = 60.6$, $df = 19$, and $p < .0001$ for the TvG data, also suggesting considerable between-studies heterogeneity.

Random-effects model

Model

Researchers typically design their studies to differ from others to increase their chances of successfully publishing something novel. This and other features of the process of science make it reasonable to expect that each study might have a unique effect θ_i . This is represented by the random-effects model,

$$T_i = \theta_i + e_i = \theta. + u_i + e_i \quad \text{for all } i. \quad (12)$$

This model is a special case of a multilevel model where the level-one (L1) model is $T_i = \theta_i + e_i$ and level two is given by $\theta_i = \theta. + u_i$ as seen in Eq. (12). Unlike in a typical multilevel model, the L1 variance in meta-analysis is not estimated using grouped data, because the within-study variance v_i is treated as known. The level-2 variance is the between-studies variance τ^2 introduced above that reflects variation among the true study parameters, the θ_i . It is also typical to treat the θ_i (or u_i) as normal, thus $\theta_i \sim N(\theta., \tau^2)$, where $\theta.$ is the mean of the θ_i .

Even when variation in true effects is anticipated, the meta-analyst often wants to estimate or test hypotheses about the mean $\theta.$ of the population of effects. Under random-effects assumptions all estimators and tests must address variation in the θ_i values. Assuming a normal distribution for the θ_i with mean $\theta.$ and variance τ^2 leads to the large-sample approximate result that $T_i \sim \text{Normal}(\theta., v_i + \tau^2)$, for $i = 1$ to k . We are no longer conditioning the variance of T_i on a common θ value. The mean effect size is typically estimated via either least-squares or maximum-likelihood estimation, and both require an estimate of the between-studies variance of the θ_i , τ^2 .

Investigating the mean effect

If the meta-analyst has adopted the random-effects (RE) model, the estimate of the overall mean effect size is typically obtained via the inverse-variance weighted estimator

$$\bar{T}^{RE} = \frac{\sum_{i=1}^k w_i^* T_i}{\sum_{i=1}^k w_i^*}, \quad (13)$$

where $w_i^* = 1/(v_i + \hat{\tau}^2)$ is the random-effects weight for study i using the estimate of τ^2 (e.g., $\hat{\tau}_{REML}^2$). The presence of both v_i and $\hat{\tau}^2$ in the weights reflects both within-study and between-studies variation. If $\hat{\tau}^2$ is much larger than the individual v_i , the weights for the T_i will be very similar across studies. In such cases $\bar{T}^{RE}$ is close to the unweighted mean.

The variance of $\bar{T}^{RE}$ reflects both sources of uncertainty – within-study uncertainty captured by v_i and between-studies variation represented by $\hat{\tau}^2$:

$$\text{Var}(\bar{T}^{RE}) = \frac{1}{\sum_{i=1}^k w_i^*} = \frac{1}{\sum_{i=1}^k (v_i + \hat{\tau}^2)^{-1}}.$$

The $100(1 - \alpha)\%$ random-effects confidence interval for the mean effect size θ , is

$$\bar{T}^{\text{RE}} \pm z_{\alpha/2} \sqrt{\text{Var}(\bar{T}^{\text{RE}})}. \quad (14)$$

The RE mean and the between-studies variance also can be used to estimate a “credible interval” that tells us where $100(1 - \alpha)\%$ of the population θ_i values are expected to lie. The interval $\bar{T}^{\text{RE}} \pm z_{\alpha/2} \hat{\tau}_{\text{REML}}$ is usually wider than the CIs given in Eq. (9) and Eq. (14) above.

Example of random-effects-mean estimation

The REML estimator of τ^2 for the TvG effects of 0.155 produces a mean effect of $\bar{T}^{\text{RE}} = 0.70$ ($SE = 0.109$). The random-effects SE of the mean is about twice the size of the fixed-effects SE of 0.058, producing a 95% confidence interval for the mean of 0.49–0.91. This interval’s location is shifted from where the FE interval was centered (at 0.65) and is also wider than the FE interval which ran from 0.54 to 0.77. As this RE interval also does not include zero and the mean is large, we still conclude that game-based instruction is superior to lecture-based instruction. However, the variation in the θ_i has increased our uncertainty – making us more unsure of where the mean effect is located.

We also use the REML variance estimate to create the credible interval including 95% of the population θ_i values, assuming normality of the θ_i . Using the RE mean, we compute $\bar{T}^{\text{RE}} \pm 1.96 \times \hat{\tau}_{\text{REML}}$, where $\hat{\tau}_{\text{REML}} = 0.393$. The credible interval covers 0.70 ± 0.77 or from -0.07 to 1.47 . Game-based instruction may have a small negative effect for some populations, but at the other extreme, effects of nearly 1.5 standard-deviation units are plausible. We can also use the normal distribution to estimate what portion of effects will be positive – for our data almost 95% of populations are predicted to have positive standardized mean differences under the RE model.

Explaining variation in effects

If considerable variation is found in the study results, or if the meta-analyst has prior hypotheses about the roles of coded study features, explanatory models are needed. Building on the asymptotic normal distribution for the effect sizes T_i , we can examine whether individual predictors or sets of predictors explain variation in effect sizes. Models like analysis of variance (for categorical predictors) and meta-regressions are available.

Meta-regression

Suppose that the meta-analyst wants to know whether study results have changed over time, or in relation to some other study feature. The model

$$\theta_i = \beta_0 + \beta_1 X_i \quad (15)$$

is a linear model where X (e.g., publication year) may explain part or all of the variance in the θ_i . Substituting this model into $T_i = \theta_i + e_i$, we obtain $T_i = \beta_0 + \beta_1 X_i + e_i$, which is a fixed-effect model with only sampling error. Alternately we may estimate the model under a mixed-model framework, adding a second error term u_i' to the model in Eq. (15), which accounts for any residual between-studies error not explained by all predictors. The analysis of linear models with predictors in meta-analysis has come to be called meta-regression.

Assuming independence of the T_i , we can use weighted least squares or other methods to estimate the key coefficient β_1 (or a set of β_j in the case of a multiple regression). We compute

$$\hat{\beta} = (X' \hat{\Sigma}^{-1} X)^{-1} X' \hat{\Sigma}^{-1} T$$

with variance $\text{Var}(\hat{\beta}) = (X' \hat{\Sigma}^{-1} X)^{-1}$ where X is a vector form of the X variable plus a column of ones to allow for an intercept (Hedges and Olkin, 1985). Also $\hat{\Sigma}^{-1} = \text{diag}(v_1, v_2, \dots, v_k)$ or in the case of a mixed model, $\hat{\Sigma}^{-1} = \text{diag}(v_1 + \hat{\tau}_X^2, v_2 + \hat{\tau}_X^2, \dots, v_k + \hat{\tau}_X^2)$, where $\hat{\tau}_X^2$ is the estimated variance of the between-studies residuals u_i' .

Together $\hat{\beta}$ and $\text{Var}(\hat{\beta})$ allow testing of the slope coefficient(s), and intercept if needed. A chi-square test of the predictor or set of predictors is also available, with $df = p$, the number of predictors (here $p = 1$). The weighted sum of squared residuals from this regression provides a chi-square test of whether X has explained all between-studies variance beyond sampling error.

Example of meta-regression

For the TvG data the model can be obtained from `metafor` using the fixed-effects command `rma(T, v, mods = ~ year, method = "FE")` for the predictor publication year. The estimated model is $\hat{T}_i = 36.7 - 0.018 \text{ Year}_i$, suggesting that the standardized mean difference drops by 0.18 every ten years. The earliest publication appeared in 2006 and the last in 2017, so the predicted decrease during the time frame of the studies is about 0.2 standard-deviation units. The standard error of the slope is $SE(b_1) = 0.018$, and the slope’s confidence interval covers from -0.054 to 0.018 , not a notable change.

The test for the slope is not significant ($Q(1) = 0.97$, $p = .33$) and residual variation from the model is considerable, with I^2 remaining at 69% and a significant residual test ($Q(18) = 59.6$, $p < .0001$). A mixed model (using method = “REML”) would

be a better choice so that we account for unexplained variation given the predictors. Year does not show significant predictive power when random effects are added, but for reporting we would use the mixed-model slope estimate of -0.014 ($SE = 0.037$).

ANOVA-like models

Another useful approach for examining multi-category predictors is the analogue to analysis of variance (ANOVA) described by Hedges (1982). As ANOVA is a linear model, estimates can be obtained using the meta-regression methods described above. However the ANOVA approach's separate within-group tests of homogeneity Q_g for each of the G groups help illuminate differences among study subgroups. Likelihood ratio and chi-square tests are available, and analyses can be done under fixed- or random-effects assumptions (though to test fit one must adopt the FE model).

Example of ANOVA

Table 1 shows results for the TvG studies grouped according to their gender composition. All groups showed significant treatment effects for game-based instruction, with the $k_1 = 2$ all-male samples showing the strongest effect with $\bar{T}_{.1} = 1.32$ standard-deviation units. All-female samples had a mean of 0.89, followed by mixed-gender samples and samples with unreported gender composition. All but the last mean differ from zero.

Three of the four subsets of studies showed little variation, and two of those had large effects (the all-male and all-female samples). This effect homogeneity allows us to say that the means for the single-gender samples and the studies with no gender information are common to all studies in those study subgroups. Mixed-gender samples showed a lower mean, and the greatest amount of spread. Unexplained variation remains within this study subset, and we interpret its mean of 0.63 as the average across the population of studies with mixed-gender samples. The between-groups test of differences due to gender composition was significant, with $Q(3) = 16.9$, $p < .001$.

Publication bias

A key threat to the validity of meta-analysis is publication bias (e.g., Dickersin, 1990; Smith, 1980), a tendency for collections of published studies to show a preponderance of significant results. Significant results can of course arise because researchers have examined effective treatments or strong relationships. But the documented tendency for journals to publish significant results, and for researchers to withhold non-significant ones, suggests that collections of studies may be filled with Type I errors. The published literature contains more significant results than would be reasonable if all results were published without regard to their significance, even when controlling for post hoc levels of power of the analyses.

A variant of this issue is reporting bias, where some research results within a published study are reported but others are not. Early studies of these biases (e.g., Coursol and Wagner, 1986) and explicit statements by journal editors confirm that both editors and referees are influenced by the strength of results. Selective reporting and related questionable research practices or QRPs (e.g., HARKing: Kerr, 1998; p hacking: Simmons et al., 2011; QRPs: John, Loewenstein and Prelec, 2012) make it critical for meta-analysts (1) to assess whether publication or reporting bias may be affecting their results (i.e., to "detect" publication bias) and if that seems likely, (2) to mitigate or adjust for it. A full treatment of publication bias can be found in Rothstein et al. (2005).

Detection of publication bias

Funnel Plot and Effect Asymmetry. The funnel plot (Egger et al., 1997) is a visual approach to detecting potential publication bias. The meta-analyst looks for asymmetry in a scatterplot of the effect sizes (typically displayed on the horizontal axis) versus a measure of their precision such as the sample size, the effect-size standard error (SE) or variance, or the inverse of the SE or variance. Sterne and Egger (2001) favor the standard error, plotted so that smaller values are near the top of the vertical axis.

If all effects arise from a single population, effects will be arrayed on either side of the mean, leading to symmetry in the plot. More dispersion is expected in less precise studies (near the bottom of the scatterplot). However, heterogeneity can result from many reasonable sources of variation (e.g., different durations or intensities of treatments, diversity in the studied samples), chance, true heterogeneity in effects, poor study design or analysis, or bias in reporting (Sterne et al., 2011). The funnel plot is a popular approach to bias detection in spite of the need to decide on asymmetry based only on one's judgment. Tests of asymmetry (e.g., by Egger et al., 1997) reduce this concern.

Table 1 Mean effect size and group-level analyses by gender composition of samples.

Sample gender composition	k_g	$\bar{T}_{.g}$	$SE(\bar{T}_{.g})$	z_g	$p(z_g)$	95% confidence limits		$\hat{\tau}_{REMLg}$	Q_g	$p(Q_g)$
						Lower	Upper			
All male	2	1.32	0.272	4.86	<.0001	0.79	1.85	0.29	2.22	0.14
All female	2	0.89	0.207	4.32	<.0001	0.49	1.30	0	0.99	0.32
Mixed gender	13	0.63	0.133	4.74	<.0001	0.37	0.89	0.38	34.9	0.0005
Unspecified	3	0.43	0.223	1.95	.051	-0.00	0.87	0.29	4.59	0.10

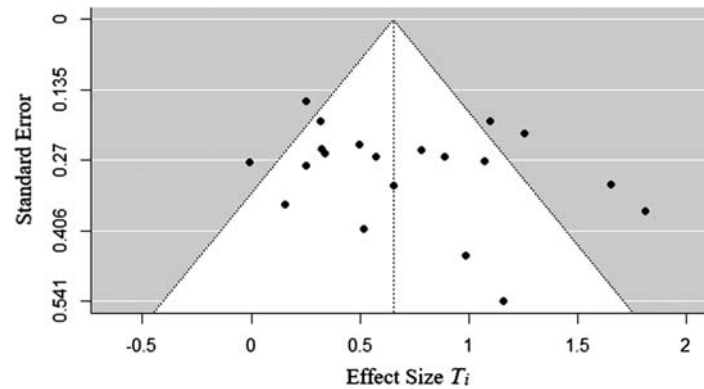


Fig. 3 Funnel plot with fixed-effects contours.

Comparisons of Published and Unpublished Studies. A very common approach (e.g., [Smith, 1980](#)) is to compare the results of published and unpublished studies. An indicator of publication type can be added to a meta-regression. This approach requires that unpublished reports be collected, which is considered state-of-the-art for data collection.

Example. [Fig. 3](#) shows a scatterplot of our example effects made under the fixed-effects assumption. Following Thompson and von Gillern (p. 20) we use $\sqrt{v_i}$ as the precision measure. Contour curves allow the meta-analyst to identify effects that are not consistent with an underlying model of homogeneity ([Peters et al., 2008](#)). Two extreme studies are seen to the left and four to the right of the contours but the effects appear to be largely symmetrical. However, Egger et al.'s test reveals a significant degree of asymmetry ($z = 2.21$, $p = .03$).

We next compare results from published and unpublished sources. A mixed-effects meta-regression estimated a nonsignificant difference between means from the 16 published and 4 unpublished sources of 0.14 standard deviations ($SE = 0.29$). Published studies show a larger mean of 0.72 versus 0.58 for effects from other sources, consistent with the idea that smaller effects are less likely to be published.

TvG's original meta-analysis used Bayesian methods to examine this difference. [Fig. 4](#) shows the posterior densities for the means and standard deviations of the published and unpublished studies. Bayesian posterior distributions show where the bulk of the population effects might lie if we treat all model parameters as random quantities ([Sutton and Abrams, 2001](#)). The posterior density of the mean for the unpublished studies on the left in the upper panel of [Fig. 4](#) has lower mode and mean values than that for the published studies, and is considerably more dispersed (with $SD = 0.22$ vs. 0.13 for published sources). In the lower panel the posteriors of the SDs τ differ dramatically and show that unpublished studies (with a mode of zero) have minimal between-studies variance compared to the more-dispersed published studies.

Adjusting for publication bias

If potential publication bias is seen, it is common to adjust for its presence. Adjustments usually shrink estimates toward zero, because publication bias favors larger effects as seen in TvG. A popular adjustment method from [Duval and Tweedie \(2000\)](#) – the trim-and-fill technique – keys off of symmetry in the funnel plot. The mathematical details are too complex for this chapter, but in short, the plot is populated with the original effects plus hypothetical points that might have been unreported, making the funnel more balanced. Then a new mean effect size and confidence interval are computed. [Fig. 5](#) shows the trim-and-fill based on the fixed-effect model. Four negative effects are added, and the mean is adjusted downwards to 0.52 ($SE = 0.054$) from the original mean of 0.65. Under the random-effects model the trim-and-fill analysis does not add any data points, and the mean is not adjusted.

Other adjustment techniques use selection models that adjust based on hypothesized chance of publication, such as “all results with $p < .05$ will be published, but only half of studies with $p > .05$ will appear.” [Hedges \(1984\)](#), [Iyengar and Greenhouse \(1988\)](#), and [Vevea \(Vevea and Hedges, 1995; and Vevea and Woods, 2005\)](#), among others, have modeled selection of effect sizes.

Other summaries

A variety of other overall summaries are available for meta-analysis. We advise avoiding summaries based on p values, including the popular failsafe N ([Becker, 2005](#); [Rosenthal, 1979](#)). These address an overly general hypothesis ([Becker, 1987](#)) and inherently confound information about sample size and effect size.

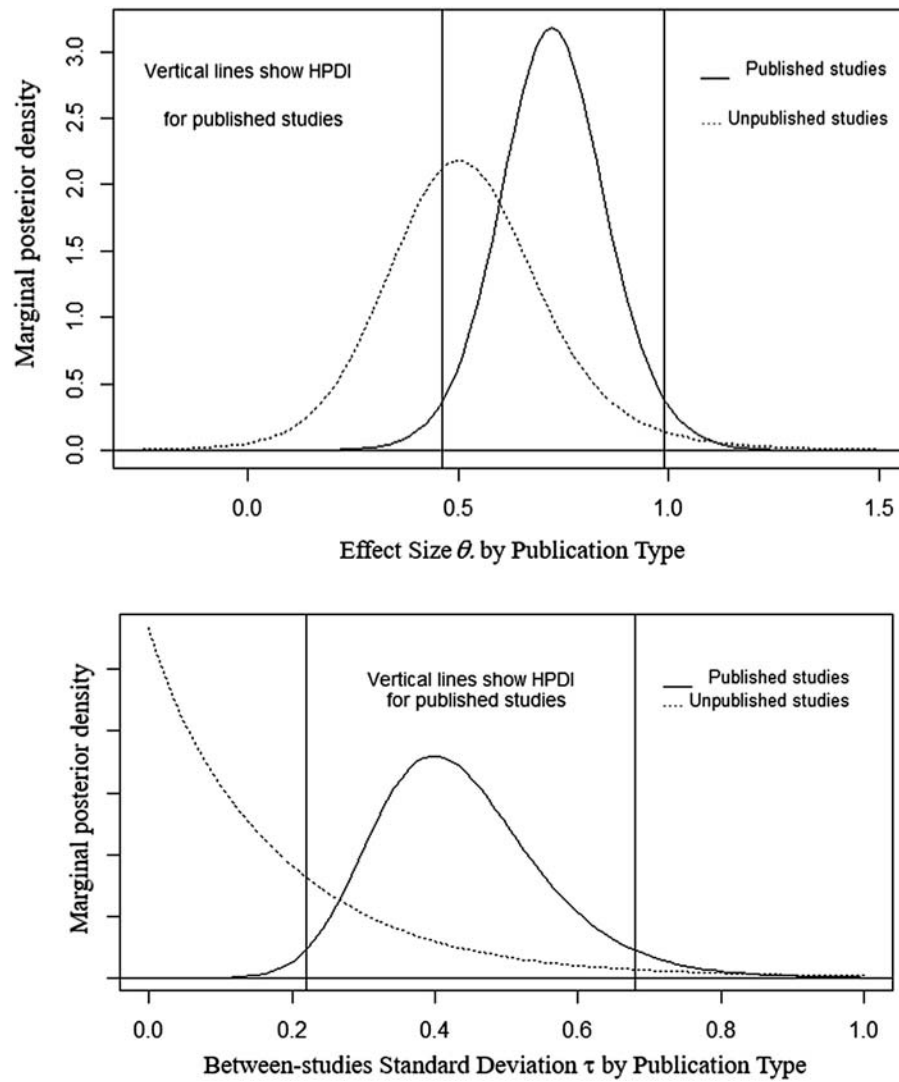


Fig. 4 Posterior distributions of the means θ . and standard deviations τ for studies grouped by publication status.

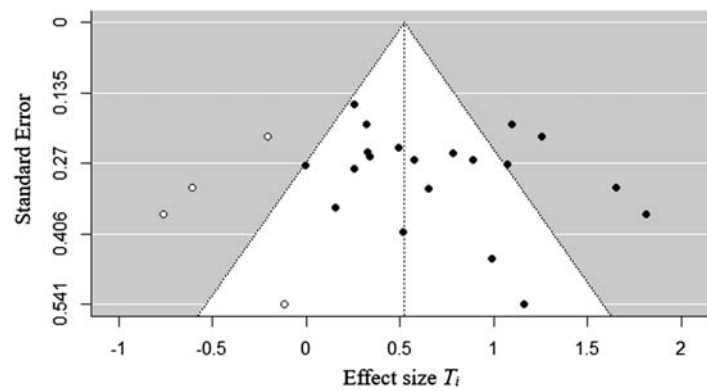


Fig. 5 Trim-and-fill plot with fixed-effect contours.

Advanced meta-analysis methods

Advanced meta-analytic techniques provide sounder or more informative analyses for some common meta-analytic situations. We mention just a few.

Dependence in meta-analysis

One prevalent complexity of primary research is dependent outcomes (Becker, 2000). Ignoring dependence can lead to higher levels of Type I error (Riley, 2009), and studies with many correlated outcomes can unduly influence a synthesis if those outcomes are treated as independent.

Approaches to dependence include ignoring it (if only a few studies in a large synthesis report multiple outcomes), creating subsets of independent results, and data reduction, which can lead to loss of information on within-study differences. Raudenbush et al. (1988) were the earliest to propose multivariate effect-size analyses, and robust-variance based methods can also be used (Hedges et al., 2010). Multiple comparison groups within and across studies can be addressed via network meta-analysis (Lu and Ades, 2004) which links sets of studies of multiple treatments to show the evidence on all conditions that have been compared.

Bayesian approaches

Bayesian analyses incorporate the meta-analyst's prior beliefs with the accumulated study effects and provide extensive information about the potential distributions of study parameters. Bayesian posterior distributions show where each population of effects – from individual study effect sizes to overall means and between-studies variances – might lie. These densities provide richer detail on location, dispersion, and shape than simpler credible intervals or CIs. While full explication of Bayesian analyses is beyond the scope of this chapter, we note that computational advances have made these analyses more accessible. The program **bayesmeta** (Roever, 2020), used by TvG and to obtain the plots shown above, is one example.

Interpretation and reporting of results

Detailed guidance for reporting meta-analyses is provided in numerous sources, including PRISMA, APA's Meta-analysis Reporting Standards (Appelbaum et al., 2018), the online Cochrane Handbook (<https://training.cochrane.org/handbook/current/chapter-iii>), and the Stroup et al. (2000) guidelines for observational studies. We mention only points that we view as vital.

A quality report is complete, transparent, and helps the reader understand what was done and the relevance of the results. The goal is to craft a report that would allow for replication of the review. Meta-analysts must give details of the process (inclusion and search rules, data-evaluation quality, etc.) as well as of the results. Reporting the PRISMA flow chart is a must. Reports should include tables of coded features and their reliabilities, and a table of study results and coded features. We advocate for listing all effect sizes, either in a table or forest plot. With the advent of open science this task has become somewhat easier.

The reviewer should begin by describing the problem that was formulated and noting whether the collected studies matched the target population. Matt and Cook (2009) argued that considering the universe of interest to the meta-analyst (the populations of U's, T's, O's, and S's) is key to understanding how far meta-analysis results can be generalized. If the studies match the target population, the meta-analyst can report what the effect-size analyses have revealed without conditions. If the studies do not map well onto the target, the meta-analyst should describe any gaps – in sample composition, outcomes or relationships examined, and treatment instantiations; these are candidate areas for future research. Of late gap maps have become popular for this purpose. Meta-analyses should report what is not known, as well as what is known.

A primary part of interpretation is to evaluate the magnitude of the typical effect size. Meta-analysts should contextualize effect magnitudes in the domain of interest rather than focusing on significance. Reports often default to using Cohen's (1988) rules of thumb, but even Cohen argued that his benchmarks were not universal. An effect of 0.16 may be small according to Cohen but seem large in a field where treatments rarely lead to strong effects.

A second goal is to convey an understanding about dispersion in results. Measures of uncertainty should be reported for all point effect-size estimates, and reports of random-effects meta-analysis should include not only confidence intervals for means, but also statements about the potential range of the distribution of population effects using credible intervals. Graphical methods are extremely useful in this regard. Minimal dispersion between studies suggests that the observed mean effects apply widely in the population, that is, effects do not depend on moderators. Here Cook's (1993) concept of heterogeneous irrelevancies is useful. Cook argues that nonsignificant moderator analyses support broader generalizations than significant ones, by identifying features that do not matter to the size of an effect or to its variability. If findings depend on participant, outcome, or intervention features, their applicability is constrained. The meta-analyst should report on all moderator analyses that were done regardless of significance. Omitting nonsignificant results adds reporting bias to a literature that may already suffer from it.

A complete report also describes the quality of evidence in the studies such as biases or method-related differences. Thompson and von Gillern reported a slightly larger mean of 0.71 for randomized studies, versus 0.61 for other studies. While not significant the 0.10 difference is not far from Lipsey and Wilson's (1993) finding of a 0.05 SD advantage for randomized studies in 74

syntheses. Visualizations of the degree of biases can be produced by software like **robvis**, an R package (McGuinness and Higgins, 2020). The meta-analysis report should also cover limitations in the review itself (Higgins et al., 2013) such as issues with reliability of coding or publication bias.

Finally, review results should be placed into the broader context of research and practice by connecting to past reviews in the field, if any, and to dominant theories relevant to the review. Meta-analytic findings that conflict with theoretical predictions raise questions about elements of relevant theories. Gaps (as mentioned above) and limitations can lead to recommendations for future research.

In many fields meta-analyses will have implications for practice and policymaking. Meta-analyses now form the foundation of evidence-based medicine. Government agencies such as the Agency for Healthcare Research and Quality use meta-analyses to influence physician decision making. Meta-analyses are critical to the U. S. Food and Drug Administration's (FDA) drug-approval process, as evidenced by the FDA's development of a draft Guidance document for the use of meta-analyses (Food and Drug Administration, U.S. Department of Health and Human Services, 2018). Likewise, the U. S. Department of Education's What Works Clearinghouse's primary focus is to scrutinize educational-intervention studies, and where justified, to synthesize them. While meta-analyses conducted by individuals may not have the immediate impact of government funded works, by conducting high quality meta-analyses and setting out the lessons learned for practice, all meta-analysts can reach toward that goal.

References

- Aloe, A.M., 2014. An empirical investigation of partial effect sizes in meta-analysis of correlational data. *J. Gen. Psychol.* 141 (1), 47–64.
- Aloe, A.M., Becker, B.J., 2012. An effect size for regression predictors in meta-analysis. *J. Educ. Behav. Stat.* 37 (2), 278–297. <https://doi.org/10.3102/1076998610396901>.
- Aloe, A.M., Thompson, C.G., 2013. The synthesis of partial effect sizes. *J. Soc. Soc. Work. Res.* 4 (4), 390–405. <https://doi.org/10.5243/jsswr.2013.24>.
- Appelbaum, M., Cooper, H., Kline, R.B., et al., 2018. Journal article reporting standards for quantitative research in psychology: APA publications and communications board task force report. *Am. Psychol.* 73 (1), 3–25. <https://doi.org/10.1037/amp0000191>.
- Becker, B.J., 1987. Applying tests of combined significance in meta-analysis. *Psychol. Bull.* 102, 164–171.
- Becker, B.J., 1988. Synthesizing standardized mean-change measures. *Br. J. Math. Stat. Psychol.* 41, 257–278.
- Becker, B.J., 1992. Using results from replicated studies to estimate linear models. *J. Educ. Stat.* 17, 341–362.
- Becker, B.J., 1995. Corrections to "Using results from replicated studies to estimate linear models". *J. Educ. Stat.* 20, 100–102.
- Becker, B.J., 1996. The generalizability of empirical research results. In: Benbow, C.P., Lubinski, D. (Eds.), *From Psychometrics to Giftedness: Papers in Honor of Julian C. Stanley*. Johns Hopkins Press, Baltimore.
- Becker, B.J., 2000. Multivariate meta-analysis. In: Tinsley, H.E.A., Brown, S. (Eds.), *Handbook of Applied and Multivariate Statistics and Mathematical Modeling*. Academic Press, San Diego, CA.
- Becker, B.J., 2005. Failsafe *N* or file-drawer number. In: Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Wiley, Chichester, West Sussex, England, pp. 111–125.
- Becker, B.J., Aloe, A.M., 2019. Model based meta-analysis and related approaches. In: Valentine, J.C., Cooper, H.M., Hedges, L.V. (Eds.), *Handbook of Research Synthesis and Meta-Analysis*, third ed. Russell Sage Foundation, New York, pp. 339–363.
- Becker, B.J., Wu, M.-J., 2007. The synthesis of regression slopes in meta-analysis. *Stat. Sci.* 22 (3), 414–429.
- Borenstein, M., Hedges, L.V., 2019. Effect sizes for meta-analysis. In: *Model based meta-analysis and related approaches*. In: Valentine, J.C., Cooper, H.M., Hedges, L.V. (Eds.), *Handbook of Research Synthesis and Meta-Analysis*, third ed. Russell Sage Foundation, New York, pp. 207–243.
- Brown, S.A., García, A.A., Brown, A., Becker, B.J., Conn, V.S., Ramirez, G., Winter, M.A., Sumlin, L.L., García, T.J., Cuevas, H.E., 2016. Biobehavioral determinants of glycemic control in type 2 diabetes: a systematic review and meta-analysis. *Patient Educ. Counsel.* 99, 1558–1567.
- Brown, S.A., Becker, B., García, A.A., Brown, A., Ramirez, G., 2015. Model-driven meta-analyses for informing health care: a diabetes meta-analysis as an exemplar. *West. J. Nurs. Res.* 37, 517–535. <https://doi.org/10.1177/0193945914548229>.
- Cho, K., Lee, S., Joo, M., Becker, B.J., 2018. The effects of using mobile devices on student achievement in language. *Educ. Sci.* 8 (3), 105. <https://doi.org/10.3390/educsci8030105>.
- Chu, H., Nie, L., Cole, S.R., Poole, C., 2009. Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: alternative parameterizations and model selection. *Statistics in Medicine* 28 (18), 2384–2399. <https://doi.org/10.1002/sim.3627>. PMID: 19499551.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, second ed. Erlbaum, Hillsdale, NJ.
- Combescurre, C., Foucher, Y., Jackson, D., 2014. Meta-analysis of single-arm survival studies: a distribution-free approach for estimating summary survival curves with random effects. *Stat. Med.* 33 (15), 2521–2537.
- Cook, T.D., 1993. A quasi-sampling theory of the generalization of causal relations. In: Sechrest, L.B., Scott, A.G. (Eds.), *Understanding Causes and Generalizing about Them, New Directions for Program Evaluation*, vol. 57. Jossey-Bass, San Francisco.
- Cooper, H.M., 1982. Scientific guidelines for conducting integrative research reviews. *Rev. Educ. Res.* 52 (2), 291–302. <https://doi.org/10.3102/00346543052002291>.
- Cooper, H.M., 1989. *Integrating Research: A Guide for Literature Reviews*. Sage, Beverly Hills.
- Cooper, H.M., 2017. *Research Synthesis and Meta-Analysis: A Step-by-step Approach*, fifth ed. Sage, Beverly Hills.
- Cooper, H., Patall, E.A., 2009. The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychol. Methods* 14 (2), 165–176. <https://doi.org/10.1037/a0015565>.
- Coursol, A., Wagner, E.E., 1986. Effect of positive findings on submission and acceptance rates: a note on meta-analysis bias. *Prof. Psychol.* 17, 136–137.
- Cronbach, L.J., 1982. *Designing Evaluations of Educational and Social Programs*. Jossey-Bass, San Francisco.
- Debray, T.P., Damen, J.A., Riley, R.D., et al., 2019. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat. Methods Med. Res.* 28 (9), 2768–2786. <https://doi.org/10.1177/0962280218785504>.
- Dickersin, K., 1990. The existence of publication bias and risk factors for its occurrence. *J. Am. Med. Assoc.* 263 (10), 1385–1389.
- Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56, 455–463.
- Eagly, A.H., Carli, L.L., 1981. Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: a meta-analysis of social influence studies. *Psychol. Bull.* 90, 1–20.
- Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.* 315, 629–634.

- Field, A.P., 2001. Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed- and random-effects models. *Psychol. Methods* 6, 161–180. <https://doi.org/10.1037/1082-989X.6.2.161>.
- Fisher, R.A., 1921. On the “probable error” of a coefficient of correlation deduced from a small sample. *Metron* 1, 3–32.
- Food and Drug Administration, U.S. Department of Health and Human Services, 2018. Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products Guidance for Industry (Draft Guidance). Docket number FDA-2018-D-3710. U.S. Department of Health and Human Services, Washington, DC. downloaded from: <https://www.fda.gov/media/117976/download>. on September 19, 2021.
- Glass, G.V., 1976. Primary, secondary, and meta-analysis of research. *Educ. Res.* 5, 3–8.
- Hedges, L.V., 1981. Distribution theory for Glass's estimator of effect size and related estimators. *J. Educ. Stat.* 6, 107–128.
- Hedges, L., 1982. Fitting categorical models to effect sizes from a series of experiments. *J. Educ. Stat.* 7 (2), 119–137. <https://doi.org/10.2307/1164961>.
- Hedges, L.V., 1983. A random effects model for effect size. *Psychol. Bull.* 93 (2), 388–395.
- Hedges, L.V., 1984. Estimation of effect size under nonrandom sampling: the effects of censoring studies yielding statistically insignificant mean differences. *J. Educ. Behav. Stat.* 9, 61–85.
- Hedges, L.V., 2007. Effect sizes in cluster-randomized designs. *J. Educ. Behav. Stat.* 32 (4), 341–370.
- Hedges, L.V., 2011. Effect sizes in three-level cluster-randomized experiments. *J. Educ. Behav. Stat.* 36 (3), 346–380.
- Hedges, L.V., Olkin, I., 1985. *Statistical Methods for Meta-Analysis*. Academic Press, Orlando, FL.
- Hedges, L.V., Tipton, E., Johnson, M.C., 2010. Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* 1, 39–65.
- Hedges, L.V., Vevea, J.L., 1998. Fixed- and random-effects models in meta-analysis. *Psychol. Methods* 3 (4), 486–504.
- Higgins, J.P.T., Lane, P.W., Anagnostis, B., Anzures-Cabrera, J., Baker, N.F., Cappelleri, J.C., Haughie, S., Hollis, S., Lewis, S.C., Moneuse, P., Whitehead, A., 2013. A tool to assess the quality of a meta-analysis. *Res. Synth. Methods* 4 (4), 351–366. <https://doi.org/10.1002/jrsm.1092>.
- Higgins, J.P.T., Lopez-Lopez, J.A., Becker, B.J., Davies, S.R., Dawson, S., Grimshaw, J.M., McGuinness, L.A., Moore, T.H.M., Rehfuss, E.A., Thomas, J., Caldwell, D.M., 2019. Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Glob. Health* 4, e000858. <https://doi.org/10.1136/bmjgh-2018-000858>.
- Higgins, J.P.T., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* 21 (11), 1539–1558. <https://doi.org/10.1002/sim.1186>.
- Hoaglin, D.C., 2016. Misunderstandings about *Q* and ‘Cochran’s *Q* test’ in meta-analysis. *Stat. Med.* 35 (4), 485–495.
- Howard, B.E., Phillips, J., Miller, K., et al., 2016. SWIFT-Review: a text-mining workbench for systematic review. *Syst. Rev.* 5, 87.
- Hunter, J.E., Schmidt, F.L., 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, second ed. Sage, Thousand Oaks, CA.
- Iyengar, S., Greenhouse, J.B., 1988. Selection models and the file drawer problem. *Stat. Sci.* 3, 109–117.
- Jackson, G.B., 1980. Methods for integrative reviews. *Rev. Educ. Res.* 50 (3), 438–460. <https://doi.org/10.3102/00346543050003438>.
- John, L.K., Loewenstein, G., Prelec, D., 2012. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychol. Sci.* 23, 524–532. <https://doi.org/10.1177/0956797611430953>.
- Kavvoura, F.K., Ioannidis, J.P.A., 2008. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Hum. Genet.* 123, 1–14. <https://doi.org/10.1007/s00439-007-0445-9>.
- Keef, S.P., Roberts, L., 2004. The meta-analysis of partial effect sizes. *Br. J. Math. Stat. Psychol.* 57 (1), 97–129.
- Kelley, K., Preacher, K.J., 2012. On effect size. *Psychol. Methods* 17 (2), 137–152.
- Kerr, N.L., 1998. HARKing: hypothesizing after the results are known. *Pers. Soc. Psychol. Rev.* 2 (3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4.
- Lai, M.H.C., Kwok, O., 2014. Standardized mean differences in two-level cross-classified random effect models. *J. Educ. Behav. Stat.* 39, 282–302. <https://doi.org/10.3102/1076998614532950>.
- Lai, M.H.C., Kwok, O., 2016. Estimating standardized effect sizes for two-and three-level partially nested data. *Multivariate Behav. Res.* 51, 740–756. <https://doi.org/10.1080/00273171.2016.1231606>.
- Lin, L., Aloe, A.M., 2020. Evaluation of various estimators for standardized mean difference in meta-analysis. *Stat. Med.* 40 (2), 403–426. <https://doi.org/10.1002/sim.8781>.
- Lipsey, M., Wilson, D., 1993. The efficacy of psychological, educational, and behavioral treatment: confirmation from meta-analysis. *Am. Psychol.* 48 (12), 1181–1209.
- Lu, G., Ades, A.E., 2004. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat. Med.* 23, 3105–3124.
- Marshall, I.J., Wallace, B.C., 2019. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst. Rev.* 8, 163. <https://doi.org/10.1186/s13643-019-1074-9>.
- Matt, G.E., Cook, T.D., 2009. Threats to the validity of generalized inferences. In: Cooper, H., Hedges, L.V., Valentine, J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation, New York, pp. 537–560.
- McGuinness, L.A., Higgins, J.P.T., 2020. Risk-of-bias visualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. *Res. Synth. Methods*. <https://doi.org/10.1002/jrsm.1411>.
- Munn, Z., Peters, M.D.J., Stern, C., et al., 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* 18, 143.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br. Med. J.* 372, n71. <https://doi.org/10.1136/bmj.n71>.
- Peters, J., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J. Clin. Epidemiol.* 61, 991–996.
- Polanin, J.R., Espelage, D.L., Grotper, J.K., et al., 2021. A meta-analysis of longitudinal partial correlations between school violence and mental health, school performance, and criminal or delinquent acts. *Psychol. Bull.* 147 (2), 115–133.
- Raudenbush, S.W., Becker, B.J., Kalaian, S., 1988. Modeling multivariate effect sizes. *Psychol. Bull.* 103 (1), 111–120.
- Richardson, W.S., Wilson, M.C., Nishikawa, J., Hayward, R.S.A., 1995. The well-built clinical question: a key to evidence-based decisions. *ACP J. Club* 123, A–12.
- Riley, R.D., 2009. Multivariate meta-analysis: the effect of ignoring within-study correlation. *J. Roy. Stat. Soc.* 172 (4), 789–811.
- Roever, C., 2020. Bayesian random-effects meta-analysis using the bayesmeta R package. *J. Stat. Software* 93 (6). <https://doi.org/10.18637/jss.v093.i06>.
- Rosenthal, R., 1979. The file drawer problem and tolerance for null results. *Psychol. Bull.* 86, 638–641.
- Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), 2005. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Wiley, Chichester, West Sussex, England.
- Schmidt, F.L., Hunter, J.E., 1977. Development of a general solution to the problem of validity generalization. *J. Appl. Psychol.* 62, 529–540.
- Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22 (11), 1359–1366.
- Smith, M.L., 1980. Publication bias and meta-analysis. *Eval. Educ.* 4, 22–24.
- Sterne, J.A., Egger, M., 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J. Clin. Epidemiol.* 54 (10), 1046–1055. [https://doi.org/10.1016/S0895-4356\(01\)00377-8](https://doi.org/10.1016/S0895-4356(01)00377-8). PMID: 11576817.
- Sterne, J.A.C., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., et al., 2016. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *Br. Med. J.* 355, i4919. <https://doi.org/10.1136/bmj.i4919>.
- Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., et al., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Br. Med. J.* 366, i4898.
- Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., et al., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Br. Med. J.* 342, d4002. <https://doi.org/10.1136/bmj.d4002>.
- Stroup, D.F., Berlin, J.A., Morton, S.C., et al., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *J. Am. Med. Assoc.* 283 (15), 2008–2012.

- Sutton, A.J., Abrams, K.R., 2001. Bayesian methods in meta-analysis and evidence synthesis. *Stat. Methods Med. Res.* 10, 277–303.
- Thompson, C.G., von Gillern, S., 2020. Video-game based instruction for vocabulary acquisition with English language learners: a Bayesian meta-analysis. *Educ. Res. Rev.* 30, 100332.
- Thompson, S., Sharp, S., 1999. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat. Med.* 18, 2693–2708.
- Valentine, J.C., Cooper, H., 2008. A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: the Study Design and Implementation Assessment Device (Study DIAD). *Psychol. Methods* 13 (2), 130–149.
- Vevea, J.L., Hedges, L.V., 1995. A general linear model for estimating effect size in the presence of publication bias. *Psychometrika* 60, 419–435.
- Vevea, J.L., Woods, C.M., 2005. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. *Psychol. Methods* 10, 428–443.
- Viechtbauer, W., 2005. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J. Educ. Behav. Stat.* 30 (3), 261–293. <https://doi.org/10.3102/10769986030003261>.
- Viechtbauer, W., 2010. Conducting meta-analysis in R with the metafor package. *J. Stat. Software* 36 (3), 1–48. <http://www.jstatsoft.org/v36/i03/>.
- Wallace, B.C., Dahabreh, I.J., Schmid, C.H., Lau, J., Trikalinos, T.A., 2014. Modernizing evidence synthesis for evidence-based medicine. In: Greenes, R.A. (Ed.), *Clinical Decision Support*, second ed. Academic Press, Amsterdam, pp. 339–361.
- Wallace, B.C., Small, K., Brodley, C.E., Lau, J., Trikalinos, T.A., 2012. Deploying an interactive machine learning system in an evidence-based practice center: Abstract. In: *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*. ACM, New York, pp. 819–824.
- Wu, T.T., Huang, Y.M., 2017. Mobile game-based English vocabulary practice system based on portfolio analysis. *Educ. Technol. Soc.* 20, 265–277.

This page intentionally left blank

INDEX

Notes

Cross-reference terms in *italics> are general cross-references, or refer to subentry terms within the main entry (the main entry is not repeated to save space). Readers are also advised to refer to each article for additional cross-references – not all of these cross-references have been included in the index cross-references.*

The index is arranged in set-out style with a maximum of three levels of subheading. Major discussion of a subject is indicated by bold page numbers. Page numbers suffixed by t, f, and b refer to Tables, Figures, and Boxes respectively. *vs.* indicates a comparison.

The index entries are presented in word-by-word alphabetical sequence in which a group of letters followed by a space is filed before the same group of letters followed by a letter. For example, entries beginning 'air density' are alphabetized before 'aircraft.' Prefixes and terms in parentheses are excluded from the initial alphabetization.

A

ABER. *See* arts-based educational research (ABER)
 ability, man2 and 9:212–213
 ABMMR. *See* arts-based mixed methods research (ABMMR)
 aboriginal education
 culturally grounded education warriors 3:19
 dedication 3:9–10
 ethnic fraud 3:10–11
 Eurocentric education as tool of
 assimilation and oppression 3:13
 instruments of possible change and
 transformation 3:12–13
 positive signs of promise and
 transformation 3:18–19
 self-determination to education as
 sovereign right 3:12
 as social reproduction 3:13–14
 Aboriginal incarceration dots 3:17
 Aboriginal rights-based education model
 3:12
 abstract modeling 6:612
 abstract principles to research practice 12:36
 ACAD. *See* activity-centered analysis and
 design (ACAD)
 academic and professional organizations
 14:16
 academic capitalism 8:16
 and market thinking in higher education
 counterproductive side of academic
 meritocracy 8:44
 macro level 8:37–40
 meso level 8:40–41
 micro level 8:41–44

academic freedom protections
 challenges and opportunities for
 8:179–181
 contemporary roots of 8:175
 as human right 8:178–179
 promising start to legal morass 8:176–177
 US universities and 8:175–178
 academic habitus of PhD students
 career opportunities 8:434–435
 employment prospects 8:434–435
 getting acquainted with processes and
 procedures 8:431
 graduate attributes 8:434
 managing own learning 8:432–433
 methods 8:430
 peer relationships 8:433–434
 PhD supervision 8:431–432
 socialization 8:433
 support from peers 8:434
 teaching and learning 8:431–433
 academic literacy 10:468
 academic meritocracy, counterproductive
 side of 8:44
 academic mobility
 barriers in policy and practice 1:233
 brain drain and brain circulation
 1:233–234
 climate change 1:235
 current issues and future challenges
 1:233–235
 faculty mobility 1:232
 futures of mobility 1:235–236
 geopolitical currents 1:234–235
 international student mobility 1:230–232
 academic performance
 self-efficacy as predictor of 6:290

Academic Ranking of World Universities
 (ARWU) 2:261–263
 academic remuneration
 augmenting salary by non-conventional
 tactics 8:197
 full spectrum of 8:196–197
 moonlighting 8:197
 negotiation 8:196–197
 academic reward system, sociological studies
 of 8:159–160
 academic self-concept
 and academic achievement 6:244–246
 across achievement tracks 6:245
 across age groups 6:245–246
 controlling for other variables 6:246
 correlational evidence 6:244
 across cultures 6:245
 generalizability of evidence 6:245–246
 learning outcomes 6:247
 link between academic self-concept and
 educational outcomes 6:244–247
 longitudinal evidence 6:244–245
 meta-analytic evidence for moderators
 6:246
 impact of multidimensional structure on
 self-concept research 6:243–244
 structure of 6:243–244
 across subject areas 6:246
 academic self-efficacy 11:390–391, 6:291
 and academic self-regulation 11:390–391
 self-efficacy as predictor of 11:390
 academic self-regulation 11:390–391, 6:291
 academic tenacity, grit and resilience to
 actually matter 6:194–195
 grit and resilience as independent factors
 6:193

- academic tenacity, grit and resilience to (*continued*)
 international validation of Bolton Uni-stride scale 6:194
 investigating construct of grit and relationship to factors 6:192–193
 overview of resilience 6:190
 relevance 6:193–194
 of grit for education 6:191–192
 of resilience for education 6:190–191
 short review of grit 6:191
- academy, case of 3:85
- accessibility of ideas 6:332
- accommodation approach to access 11:327–328
- accountability 13:194, 13:298, 4:293–294
 accountability and professional practice 4:294
 accountability-driven data 4:170–172
 building professional ownership of 4:297
 consequences for assessment 13:299–300
 constraints on accountability assessment 13:299–301
 developing system of professional accountability 4:296–301
 through large-scale assessments 13:197–199
 learning 13:301–302
 New Zealand research 13:302–303
 impact of public accountability systems 4:294–296
- accountable talk 6:604–605
- accuracy
 challenge of 10:637–639
 illusion of 13:227
 of knowledge 6:212
- achievement
 growth models 14:382–387
 gain score model 14:383–384
 historical and policy background 14:382
 statistical properties of 14:387–388
 student growth percentile model 14:385–386
 value-added model 14:386–387
 peer relationships and academic achievement 6:290
 teacher–student relationships and 6:289
 achievement goal theory (AGT) 6:492–493
- ACL. *See* African Children's literature (ACL)
- acquisition 6:15, 6:213–215
- ACT. *See* American College Testing Program (ACT)
- action research (AR) 12:568–569, 5:236, 7:359
 advantages of intersecting MMR with AR and CBPR approaches in education 12:570–571
 challenges of intersecting MMR with AR and CBPR approaches in education 12:571
 common features of 12:569–570
 methodological considerations of intersecting MMR with AR and CBPR 12:571–576
 mixed methods 12:431
 model 5:470
 active knowledge making 6:263
 active mode
 of cognitive engagement 6:691
 strategies 6:694
 active unothering 5:679
 activism among young women, modes of 2:368–370
 activist lenses 10:611–614
 activist practitioner inquiry into life in schools/families/communities in US South 7:397
 activity-centered analysis and design (ACAD) 6:394, 6:417–418
 and innovative or learning environments 6:395–396
 lens to micro, meso and macro levels 6:395–396
 and maori marae 6:395
 actors in curriculum design 7:202–204
 curriculum policy and actors 7:202–203
 role of networks 7:203–204
 ad hoc methods 14:807–808
- ADAPT21. *See* adaptability in complex problem solving situations (ADAPT21)
- adaptability in complex problem solving situations (ADAPT21) 13:252
- adaptive expertise 6:181–182
- adaptive testing
 designs 14:364–365
 effects of 14:213–215
 measurement efficiency 14:214
 test-taking motivation 14:215
 validity 14:214
- additional alternative correlation coefficients 14:744–747
- additional language (L2) 6:364
 translating theories of learning to L2 context 6:365–366
- additive trees 14:630
- adequate anti-racist policies, lack of 5:416–417
- administrators as catalysts for school improvement 4:199–200
- admissions tests
 at four-year undergraduate institutions 14:454
 at graduate and professional schools 14:454
 to people of color, women, and special populations 14:456
 predictive validity of 14:454–456
- adolescence
 academic motivation development 6:492–493
 AGT 6:492–493
 cognitive development and adolescent learning 6:487–488
 expanding social networks 6:490
 expectancies and values 6:492
 exploring and changing identities 6:490–491
 gender identity 6:491
 identity development and adolescent learning 6:490–492
 key developments in 6:486–487
 motivational beliefs change during adolescence 6:493
 possible selves and self-concept 6:491
 racial and ethnic identity 6:491–492
 roles of adolescent social relationships and experiences 6:489–490
 social/moral development 6:489–490
 theoretical perspectives on social/moral development 6:489
 adult education as response 1:548–550
 language acquisition 1:548–549
 recognizing prior learning 1:549–550
- adult literacy
 functional literacy 10:458
 and identity formation 10:458–459
 international surveys of adults' literacy skills 10:457–458
 life skills 10:457
 non-valorized practices 10:459
 as paradigm 10:458–459
 as political praxis 10:459–461
 and political struggle 10:459–460
 as problem 10:456–458
 problematizing technology 10:460–461
- adult-learning 10:302
- adulthood
 bridge period of positives and negatives 6:499
 cognitive developments in midlife 6:500
 emerging adulthood 6:497
 generativity 6:500
 identity and work commitment 6:498
 middle adulthood 6:498–501
 myths of middle age 6:501
 personality stability *vs.* development 6:498
 sustained intimacy 6:498
 young adulthood 6:497–498
- advertising literacy 2:305
- AE. *See* assessment engineering (AE)
- AES. *See* automated essay scoring (AES)
- affect theory 10:912
 assemblage perspectives on affect 10:911–914
 expanding conceptions of 10:912–914
 openings 10:916
 as orienting and disorienting 10:913
 recognizing affect and textual variations 10:914–916
 as transformative and normative 10:913–914
- affective engagement 6:356
- affirming methodologies 12:179–180
 in education, oral traditions and 12:180
- affordances 10:505
- AfL. *See* assessment for learning (AfL)
- AfrEA. *See* African Evaluation Association (AfrEA)
- Africa
 COVID-19 impact on higher education in 8:91–92
 inclusive education in 9:405
 literacy as regimes of African decolonial burden 10:731–734

- African precolonial writing practices 10:724–729
- connecting colonial and contemporary regimes of truth with regimes of literacy 10:729–731
- transfer credit 8:347–348
- African Children's literature (ACL) 10:688
- accessibility issues 10:693–694
- critical evaluation of studies and essays/articles surveyed 10:689–690
- gaps in scholarship 10:694
- publishing of books with black and brown faces 10:692–693
- representations of 10:691–692
- research and literary criticism on 10:690
- scholarship of annotations 10:690
- African Evaluation Association (AfrEA) 13:267
- African evaluation guidelines 13:267
- African higher education
- addressing issues inequality, equity, and access in educational systems 8:88
- advancing meaningful partnerships 8:87
- advancing technological infrastructure 8:87
- cases of selected African countries 8:86–87
- Egypt 8:84
- implications pandemic for future planning 8:87
- innovative modes of education delivery 8:88
- Kenya 8:86
- Nigeria 8:85–86
- South Africa 8:85
- African legacy 10:656–657
- alternative terms to name region, and consequences for children's literature 10:656–657
- denominations of area, and consequences for children's literature 10:658–659
- African universities
- establishment of national universities in Africa and quality of academic freedom in practice 8:115–118
- intensity and quality of academic freedom in 8:120–122
- post-2000 developments 8:120–122
- African-American language from classroom 10:312–313
- Afrocentric education
- See also* global and Indigenous education
- authors' self locations 2:162
- contemporary implications and offerings of 2:167–169
- major tenets & philosophies of 2:165–167
- origins and historical overview of 2:162–165
- struggling for 2:164–165
- age roles in self-efficacy formation 6:252
- agency enacts motivation 13:55
- Agenda 2030 and UN's sustainable development goals 7:406–407
- AGT. *See* achievement goal theory (AGT)
- AI. *See* appreciative inquiry (AI); artificial intelligence (AI)
- AIC. *See* Akaike information criterion (AIC)
- AIG. *See* automatic item generation (AIG)
- AIMMR. *See* arts-informed mixed methods research (AIMMR)
- 'aina-based education
- addressing settler traditions of place 3:230
- articulating what transformation looks like 3:236–237
- Centering Kanaka 'Oiwī perspectives of place and decolonizing place-based education 3:229
- claim and advance indigenous peoples' processes of learning and teaching 3:230–231
- findings 3:233–234
- focus on transformation 3:231
- He Leo Mahalo 3:237
- Ho'omana 3:236
- Huli Ka Lima I lalo 3:234–235
- knowledge sources 3:232
- kuleana 3:235
- Mo'olelo as methodology 3:232
- naming of framework 3:233
- opening with story 3:228–229
- projects of 3:234
- revisiting 3:230
- structural issues of access 3:235
- weaving collaborative story 3:232
- 'āina Momona and 3:235–236
- Akaike information criterion (AIC) 14:286
- Alaska Native Claims Settlement Act on education 3:25–26
- ANSCA education consortium 3:27
- BBNCEF 3:28–29
- challenges and opportunities 3:29
- coalition of funders 3:29
- colonial periods 3:23
- eligibility 3:27
- formation of native association 3:25
- geography and diversity of Alaska's Indigenous peoples 3:22–23
- impact 3:28
- modern American period 3:24–25
- native sovereignty and Western education 3:26–27
- traditional education 3:24
- westernized education 3:24
- Alaska native values 3:147–148
- Alexander's dialogic teaching framework 6:604
- aligning PBL with NGSS
- approach to design 11:120–122
- context 11:119–120
- of implementation 11:120
- data collection 11:122
- design cycles 11:123–127
- discussion 11:127–129
- employing design-based research methods 11:121
- logic model 11:121–122
- methodology 11:119–123
- related research and theoretical foundation 11:118–119
- studying implementation of materials 11:122
- Aloha 'āina Theory and Guerilla Praxis applications 3:74–78
- development of 3:73–74
- establishing native theory & praxis 3:73–74
- higher education, neoliberalism, and indigenous knowledge 3:70–73
- principle of centering genealogy and nationality 3:77
- principle of kuleana and moral obligations 3:77
- principle of sovereignty and de-occupation 3:76
- principle of truth telling, radicalism and protest 3:77–78
- toward an 3:74–78
- alphabetic literacy, tyranny of writing and hegemonic ideology of 10:729–730
- alterable family variables 4:263–266
- elements of parental involvement 4:265
- parental involvement including parental expectations 4:264–265
- parental style and importance of love and structure in home 4:265
- social capital 4:265–266
- alternative assessment 13:96–97
- alternative formats 13:97
- alternative forms of classroom assessment 13:98–100
- assessment by teachers for external purposes 13:98–99
- diagnostic assessment 13:101
- dynamic assessment 13:101
- forms 13:97–100
- theoretical underpinnings 13:99–100
- uses 13:100–101
- ambiguity 4:18–20
- America through literature and popular culture 10:229–230
- American College Testing Program (ACT) 14:452
- American study on phenomenology 12:353–356
- amnesty international 2:61
- amount of knowledge 6:212
- Amparo Clavijo 2:199
- analogical reasoning process 6:614
- analysis 12:748–749
- analysis of covariance (ANCOVA) 14:536
- general linear *F* test 14:536–537
- multivariate 14:545–546
- multivariate ANOVA 14:543–544
- univariate 14:539–543
- univariate ANOVA 14:537–539
- analysis of variance (ANOVA) 14:530
- ANOVA-like models 14:853
- computation of *F* 14:531–532
- interpreting significant *F* value 14:532
- logic of 14:530–532
- planned comparisons 14:534–535
- repeated-measures design 14:532–533
- two-way ANOVA 14:533–534
- analysis-centric MMR designs 12:430
- anchor test designs 14:239
- best practices for 14:243–244
- anchors 12:36–37

- ANCOVA. *See* analysis of covariance (ANCOVA)
- Andre's writing sample reinterpreted 10:859
- Andres Manuel López Obrador proposals 2:80
- anganwadis 1:535
- Angoff and related methods 14:374–375
- ANN. *See* artificial neural networks (ANN)
- ANOVA. *See* analysis of variance (ANOVA)
- anthropocene
- alternatives 7:304–305
 - curricular challenge and reciprocal device 7:297–299
 - curricular strategies 7:301–302
 - education in 1:695
 - long duration 7:299–300
 - master narratives' rule and role 7:302–304
 - new historical culture 7:300
 - short story 7:300
 - storylines for 7:299–300
- anthropocentrism 2:206–207
- anti-black linguistic racism 10:144–145
- in preservice teacher education 10:145–146
- anti-black racism 10:270–272
- anti-colonial indigenous social media 10:803–806
- dogface208 vibes 10:804–806
 - something else 10:803–804
- anti-racist approaches in teacher education 5:420–422
- creating safe spaces injustice can be critically examined from multiple perspectives 5:420–421
 - critical self-reflection, examination, and autobiography 5:420
 - discomfort and processing emotions 5:420–421
- anti-racist education 5:414–416
- theoretical underpinnings of 5:415
- anti-racist teacher education
- experiences of Black, Indigenous, and teacher candidates of color 5:416
 - necessity of 5:416–420
 - programmatic challenges 5:416–417
 - recommendations 5:423–424
 - whiteness of teaching and teacher education 5:415–416
- anti-racist teacher educators, hire and train 5:423
- anti-testing movement 14:6–7
- anxiety and mismatch between presentation and experience of remote school 10:932–933
- ANZEA. *See* Aotearoa New Zealand Evaluation Association (ANZEA)
- Aotearoa 3:16–17
- Aotearoa New Zealand, pacific people in 8:400
- Aotearoa New Zealand Evaluation Association (ANZEA) 13:267
- APP. *See* aspiring principals program (APP)
- Appalachia regional comprehensive center's evaluation of capacity-building program, responsive metaevaluation of 13:271–272
- appreciative inquiry (AI) 5:237
- apprenticeship learning
- in academic fields 6:572–573
 - characteristics of 6:568–569
 - in communities of practice 6:570–571
 - developmental perspective on 6:571
 - historical roots and development 6:569–570
 - from learning perspective 6:570–571
 - roots and developments of 6:569–570
 - educational research 6:570
 - from teaching perspective 6:571–573
 - in vocational training 6:571–572
- April revolutionary black arts 2:361
- AQF. *See* Australian Qualifications Framework (AQF)
- AR. *See* action research (AR)
- Arab children's publishing scene 10:678–680
- Arab countries, scholastic landscape in 10:680–681
- Arabic picture books
- ecosystem and opportunities 10:678–683
 - trends and opportunities 10:681–683
 - illustration and type 10:685–686
 - language 10:683–684
 - poetry and rhyme 10:685
- Archer's morphogenetic social realism 7:89–90
- horizontal and vertical discourses 7:91–92
 - social justice and social and epistemic access to knowledge 7:91
- argument-based approach to validity 14:36–40
- enacted interpretations and uses 14:37–38
 - inferences in IUA 14:37
 - IUA and validity argument use case 14:38–40
 - principled approaches in test design to support IUA 14:38
 - sources of validity evidence 14:38
- argument-based approach to validity and validation 13:214
- interpretation-and-use argument 13:214
- argument-based pedagogical methods
- evidence from designed to foster students' critical thinking skill and dispositions 6:582–583
 - from skills to dispositions 6:582–583
 - transfer from social to individual 6:582
 - transfer to other topics/contexts 6:582
- argumentation as alternative construct to critical thinking 6:576–578
- acceptability 6:578
 - reasonableness 6:577–578
 - relevance 6:578
 - sufficiency 6:578
- argumentative discourse, reasoning abilities, skills and dispositions implied in 6:579–581
- appraisal abilities 6:580
- epistemic and epistemological awareness 6:581
- evaluation of arguments 6:580
- evaluation of context 6:580
- evaluation of information 6:580
- generative abilities 6:579
- inferential abilities 6:580
- metastrategic skills 6:580–581
- reasoning/arguing disposition 6:581
- Army Alpha test 14:6
- arrival of GERM as alternative typology 1:168–171
- art and science of reading fluency and fluency instruction 10:388–397
- art installations 2:369–370
- art of governing 7:106–107
- artificial agent-based interaction 14:412
- artificial intelligence (AI) 10:530–531, 6:83
- automated tutoring 10:534–536
 - challenges 10:539–540
 - components of reading skill 10:529–530
 - current AI tools 10:532–533
 - limitations in embodying us 10:539–540
 - and literacy 10:529
 - mentoring by social robots 10:537–538
 - mining databases to optimize reading complexity 10:537
 - ORF assessment 10:533–534
 - predicting future 10:538–539
 - research foundations for 10:536–538
 - retrospective 10:540
 - spoken and written language *vs.* formal/informal 10:530
 - technology aided assessment 10:533
 - traditional, current, future approaches to reading instruction 10:531–536
- artificial neural networks (ANN) 6:86
- artificial neuron 14:686–687
- artistic approach in curriculum design 7:201
- artistry of teaching, reclaiming 5:648–654
- arts 5:305
- arts-based practices for literacy & language learning
 - oral language development 10:359
 - phonics 10:360–361
 - phonological awareness 10:360
 - phonological awareness and phonics through arts 10:359–361
 - reading fluency and comprehension 10:361
 - vocabulary learning 10:361–362
 - writing 10:362–363
- arts-based research
- using art as different way of knowing 12:164–165
 - art in arts-based research 12:166
 - art teachers researching academically 12:162–163
 - using art to collect data 12:163–164
 - using art to share findings of research 12:163
 - arts-based qualitative research 12:161
 - expanding existing forms of art to facilitate research 12:161–162
 - knowledge-making through 12:166
 - photograph 12:159–160
 - positionality 12:160
 - research in 12:165

- researching through and in art 12:160–161
- arts-related research
- challenges from tensions between 12:708–709
 - challenges to using and integrating 12:707–709
- challenges faced in arts-related approaches 12:707–708
- clear discussion of methods/practices 12:708
 - using clear terminology 12:708
 - level of training 12:707–708
 - support for art in mixed methods study 12:708
- telling stories through 7:394
- arts-based educational research (ABER) 2:313–314
- and critical race theory 2:317
 - historical and theoretical foundations and characteristics of 2:314–315
 - implications of ABER for diversity, democracy, and social justice 2:318
 - and New Materialism 2:317–318
 - paradigm shifts in 2:316–318
 - qualities and rigors of 2:315–316
- arts-based mixed methods research (ABMMR) 12:704–705
- arts-informed mixed methods research (AIMMR) 12:704–707
- ascriptive metaevaluation 13:265
- Asia, growing interest in 8:25–26
- Asia-Pacific-realist
- alternative constructions 1:412–414
 - constructions of education within multiple geographies of 1:414–417
 - geographies as pathways for student mobility 1:416–417
 - multiple ways of understanding 1:412–414
 - realist perspective 1:412
- Asia/Pacific, transfer credit 8:346–347
- Asian American immigrant students in schools/families/communities, research on education of 7:396
- ASPIRES 11:377
- aspiring principals program (APP) 4:362–363
- assemblage perspectives on affect 10:911–914
- assessment engineering (AE) 14:155–157
- assessment for learning (Afl.) 13:21, 13:23–28, 13:65–66, 13:147–148
- and cultural values 13:30–32
 - definitions and key characteristics 13:24–26
 - distinguishing assessment for learning from formative assessment 13:28–30
 - edge of implementation 13:148–152
 - evidence, dialog, feedback, reflection 13:25
 - and formative assessment 13:29–30
 - implementing 13:148
 - increasing teachers' awareness 13:32
 - leadership support 13:32
 - making learning explicit, behaving virtuously 13:30–31
 - in moment, simultaneously with teaching and learning 13:25–26
 - within normative context of practice, teachers' understanding Afl 13:66–67
 - principles and classroom learning culture 13:30–32
 - principles and strategies 13:26–28
 - promoting learner autonomy, respecting hierarchy 13:31
 - reasserting distinctiveness of assessment for learning 13:30
 - reconciling Afl principles and East Asian classroom assessment practices 13:32
 - resistance to reform due to continued good performance 13:31–32
 - strategies 13:27–28
 - student agency 13:24–25
 - ten principles 13:26
 - three principles 13:26–27
- assessment of 21st century skills (ATC21S) 13:250–251
- assessments 10:370–371, 13:22, 13:192–193, 7:283–285, 9:398–399
- competence and standards 7:283–284
 - delivery layer 14:167
 - development of 10:370–371, 9:400–401
 - fairness 13:196
 - frameworks and formats for measuring student learning in higher education 8:313–314
 - implementation layer 14:167
 - key characteristics of 13:195–196
 - learning to learn assessing and learning 13:10–20
 - literacy 14:523–524
 - policy reform 13:120–122
 - process 13:226
 - coursework 13:228–229
 - desirable properties of 13:225–226
 - importance of tasks 13:226
 - as part of instruction 13:227–228
 - routes to better assessment 13:232–233
 - short item multiple choice tests 13:230–232
 - spectrum of assessment models 13:227–232
 - timed examinations 13:229–230
 - tools and processes 13:225–227
 - reform 13:119–128
 - reliability 13:196
 - standardized assessment 5:24
 - standards as instrument in accountability and monitoring performance 7:284–285
 - task format 8:317
 - validity 13:195–196, 13:300–301
- asset 6:210
- asset-based *vs.* deficit-based approaches 10:79–80
- assimilation 9:72
- associative learning in infancy 6:451–452
- asynchronous discussion tool/forum 11:210
- ATA. *See* automated test assembly (ATA)
- ATC21S. *See* assessment of 21st century skills (ATC21S)
- attention 6:612
- aspects of 6:122–123
 - behavioral phases 6:119
 - brain networks 6:120–122
 - compensatory systems 6:125
 - and developmental disorders 6:121
 - digital media and 6:124–125
 - directed-attention model 6:123–124
 - dyadic attention 6:123
 - to faces 6:118–119
 - genes and 6:122
 - in infancy 6:117–119
 - intention requires attention 6:124
 - joint attention 6:124
 - and learning 6:123–125
 - memory and 6:122
 - networks 6:119–122
 - novelty attracts attention 6:122
 - pupil dilation and 6:122–123
 - selective attention 6:118
- attitudes
- about disability 9:328–329
 - as life flows 9:330–331
 - psychology of 9:329–330
 - toward inclusion 9:328–333
- Australia 11:286, 8:443–448
- Australia-India research fund 8:208
 - careers advice and quality in 8:444
 - competition and choice 5:31–32
 - disabled children and schooling in 9:511
 - EAL/D policy and context 10:68
 - GERM and education reform in 5:31–34
 - increased focus on reading, mathematics, and science 5:33
 - pathways into and within Australia's tertiary education 8:445–446
 - principal preparation 4:360
 - reviews of education for students with disability in
 - Australian context 9:376–378
 - Australian reviews and inquiries 9:377–378
 - integration 9:382
 - review and inquiry findings 9:378–382
 - student exclusion 9:380–382
 - student segregation 9:380
 - school education in 5:30
 - school leadership 4:67
 - standardization of teaching and learning 5:32–33
 - system leader and system leadership 4:131–133
 - teacher accountability in 5:64–65
 - transfer credit 8:347
- Australian, changing context of children's development and learning 4:341
- findings 4:339–342

- Australian, changing context of (*continued*)
 governments, systems, schools, families and communities creates ways of leading and learning 4:340–341
 integrated service delivery approach 4:341–342
 leadership for learning 4:342–344
 methodology 4:338
 strengths-based approach to indigenous education potential to reframe deficit assumptions 4:339–340
 Australian case study 13:149–150, 5:148
 Australian Government Productivity Commission 3:11
 Australian labor market 8:443–444
 Australian Qualifications Framework (AQF) 8:447
 and effects on pathways 8:447–448
 Australian schools
 creative writing and spoken word poetry in 10:70–72
 real talk in 10:71
 Australian tertiary education system, issues in developing more coherent 8:446
 Australian tertiary system 8:444–445
 Australian universities. See *oneself in*
 designing frameworks that embrace indigenous visibility 3:114–115
 relationality and creating sense of belonging 3:112–114
 universities, whiteness and racism 3:110–111
 Australian vocational education and training system 8:491–492
 Austria's educational system, overview of 5:217
 authentic leadership 4:503, 4:505
 authentic learning 5:702
 authenticity
 challenge of 10:637–639
 in engineering curriculum 11:37–38
 authorship 10:730–731
 autism
 identities
 analysis of topics and themes 9:490–493
 discussion 9:493–494
 and girls 9:486
 methodology 9:487
 literacies, and social communication 10:436–437
 characteristics of home literacy and autism 10:436–437
 literacy and social communication success and post-school outcomes 10:437
 student literacy profiles and autism 10:436
 and literacy challenges 10:437–438
 medical model perspective 10:438
 social model of disability 10:438
 supporting inclusion of students on autism spectrum through new literacies 10:438
 and multimodal approaches to literacies and social communication 10:439–440
 research 10:439
 spectrum 10:436
 embracing new literacy studies 10:436
 autistic voice 9:487
 autoethnography 12:102–103, 12:105
 case for using 9:139–140
 as critical reflection 9:138
 in teacher preparation 9:140–141
 critique of 9:138–139
 evocative autoethnography 12:104
 futures 12:105
 historical view 12:103
 outlets for autoethnographic scholarship 12:105
 tales of school experience 9:488–490
 Bella's tale 9:489–490
 Hannah's tale 9:489
 Jodie's tale 9:488–489
 Rachel's tale 9:488
 and voice 12:103–104
 automated analysis of student writing, discourse analytics in 6:77–79
 automated essay scoring (AES) 14:692
 automated scoring
 emerging trends 14:401
 of essays 14:399
 operational uses of 14:399–401
 of response processes 14:400–401
 of short text responses 14:400
 of spoken responses 14:399–400
 of STEM responses 14:400
 validation of 14:398
 automated test assembly (ATA) 14:201
 adaptive test assembly 14:206–207
 as adjunct to manual assembly 14:207
 constraints 14:202
 limitations 14:207
 objective functions 14:202–205
 simultaneous assembly of multiple forms or item blocks 14:206
 solution algorithms 14:205–206
 test assembly
 as combinatorial optimization problem 14:202–205
 use cases 14:206–207
 test blueprints 14:201–202
 automatic item generation (AIG) 14:194
 non-template-based AIG and example items 14:197–198
 template-based AIG 14:194–195
 for test development 14:194–199
 automating oral reading fluency assessment (ORF assessment) 10:533–534
 automation
 Locke on nature 9:211–212
 man1 9:210–211
 man2 and ability 9:212–213
 man3 and 9:215–217
 autonomous and ideological models of literacy in classroom 10:310–311
 autonomy, arranging for 10:403–404
 AutoTutor 6:637, 6:641–643
 auxiliary variables, use of 14:813
 aversive control 6:17
 award bearing model 5:470
 awareness-raising 10:830
 axiology 12:382
- ## B
- Bakhtin's dialogism 6:602
 balwadis 1:535
 BAS. See BEAR assessment system (BAS)
 basic fully-connected architecture 14:687–688
 Bate-papo and pláticas 2:376–377
 Bayesian approaches 14:856
 Bayesian inference 14:774–777
 computation 14:778–780
 model fit assessment and model selection 14:780–782
 prior specification 14:777
 Bayesian CFA 14:615–617
 Bayesian data analysis 14:130–131
 Bayesian estimation 14:572, 14:800–801
 Bayesian inference 14:774–777
 connection with frequentist inference 14:777
 decision-theoretic treatment 14:775–777
 posterior distribution 14:774
 running example 14:774
 Bayesian information criterion (BIC) 14:285–286
 Bayesian methods 14:811–812
 Bayesian networks 14:120–123
 additional topics in 14:126–127
 to causal networks 14:126
 d-separation algorithm 14:121
 discrete item response theory model 14:123–126
 graphical models 14:119–120
 maps 14:120–121
 message-passing, belief propagation, and probability updating 14:122–123
 parameter learning 14:126–127
 probability factorization 14:121–122
 structure learning 14:126
 Bayesian psychometrics 14:130
 model specification 14:131–133
 posterior inference 14:132–133
 posterior predictive model checking 14:134–136
 real data example 14:134
 BBNCEF. See Bristol Bay Native Corporation Education Foundation (BBNCEF)
 BEAR assessment system (BAS) 14:147–153
 construct map 14:148–149
 items design 14:149
 measurement model 14:151–153
 outcome space 14:149–151
 beginning teachers' professional identity
 work and support
 aim of identity work 5:164–165
 characteristics of identity work 5:165
 conceptualizing beginning teachers' professional identity 5:161–162

- constraining and cultivating professional identity experiences 5:162–164
 experiences confirm or reinforce beginning teachers' identity 5:163–164
 identity issues beginning teachers face 5:163
 implications 5:167–168
 moments of contact with students 5:163–164
 positioning toward students, colleagues and school leaders 5:164
 program aimed at identity work 5:166–167
 recognizing growth 5:165
 role of reflection 5:165
 safe environment 5:165
 stories to live by 5:164
 supporting beginning teacher's professional identity development 5:164–167
 tailor-made support aimed at strength 5:165
 behavioral dispositions 6:356
 behavioral psychology and education 6:19
 behavioral strategies 6:356
 behaviorism 6:37, 6:365, 6:368
 behaviorist perspective, learning from
 beginning 6:12
 behavioral psychology and education 6:19
 classical conditioning 6:14–17
 distinctive features 6:13
 habituation 6:13–14
 history 6:12–13
 legacy of behavioristic approach 6:20
 operant conditioning 6:17–19
 beliefs 6:356–357
 Bernoulli distribution 14:718
 Bernstein's pedagogic models 7:10–11
 beta distribution 14:712–713
 BETTER project in Mozambique 10:477
 betweenness centrality 4:530
 bibliometrics
 analyses 4:549–553
 review method
 database for 4:547–549
 rationale for 4:546–547
 role in research assessment 8:160–165
 aggregate indicators 8:161–162
 impact 8:161
 metrics as product 8:163
 productivity 8:160
 publication venues 8:162
 recommendations for good practice 8:163–165
 BIC. *See* Bayesian information criterion (BIC)
 bifactor/testlet models 14:91–92
 Bildung 7:252, 7:335–336
 bilingual education, types of 2:189–190
 biliteracy practices
 creativity in 10:49–50
 creativity in and across languages 10:48–49
 drama and play pedagogies as performative spaces in 10:50–51
 identity development & translanguaging 10:49
 language brokering 10:49–50
 personal, linguistic, cultural and social resources 10:51
 translanguaging as human capital building 10:50
 Binet, Alfred 14:5–6
 contribution 6:199–200
 binomial distribution 14:719
 binomial GLM 14:586–587
 biologically primary skills 6:128
 biologically secondary skills 6:128
 biserial correlation correlation 14:745–746
 bivariate numerical data, describing 14:730
 black colleges and environments 8:425
 black educators 6:674–675
 George Edmond Haynes 6:675
 W. E. B. DuBois 6:674–675
 black faces, books with 10:692–693
 black feminist concept 2:116–118
 black language in teacher education 10:145–146
 #BlackLivesMatter 2:352
 grassroots organizing for racial justice in schools 2:356–359
 at school national days of action 2:360–361
 streets to school house 2:353–356
 struggling for union support and BLMAS in higher education 2:358–359
 week of action to year of purpose and lifetime of practice 2:359
 black queer literacies 10:262–263
 Black Radical Educator Day 2:361
 black social settlement movement 6:674
 Black to School 2:360
 blacks in higher education 8:353–354
 blended course designs
 challenge 8:74–75
 combining self-report and observational data in investigations into 8:79–80
 emergence 8:79
 entanglement 8:78–79
 in 3P model 8:79
 materiality data used to complement self-report data as evidence of learning 8:78–79
 quantitative approaches to identifying qualitative variation in learning 8:76–78
 research into student approaches to learning 8:75–78
 blended learning 6:426
 research and effectiveness of online and 6:427–428
 blocked design 14:698–699
 BLRT. *See* bootstrapped LRT (BLRT)
 Bock-Aitkin EM algorithm 14:95
 Bologna 5:606–607
 Bolton Uni-stride scale, international validation of 6:194
 bookmakers
 challenge of availability and access 10:636–637
 social responsibility of 10:636–639
 Bookmark method 14:375
 books with black and brown faces 10:692–693
 bootstrap method 14:764–767
 bootstrap confidence intervals 14:765–766
 bootstrap hypothesis testing 14:766
 bootstrap resources 14:767
 nonparametric bootstrap work 14:767
 parametric bootstrap 14:766–767
 bootstrapped LRT (BLRT) 14:653–654
 border-crossing across multitopic spaces 10:702
 border-hispanic adolescents 10:291
 borderlands 10:288–290
 Bedouin border life in Negev of Israel 10:292–293
 border affect and emotional encounters 10:290
 characterizing traversing 10:288–289
 crossing of borders 10:289
 hispanic population and educational research 10:290
 language and literacy educational research on Texas-side of U. S. Mexico border 10:290–292
 rampant stereotypes of border lands and youth populations 10:291
 research to uncover realities of United States–Mexico border adolescent populations 10:291–292
 unique teacher education for border adolescents 10:293–294
 youth 10:290–292
 Botswana case study 1:104–117
 boundary crossing model 5:385
 Braille 10:423
 adoption of Braille in United States 10:425
 contractions and governing rules 10:430–431
 current practices for Braille instructions 10:427–429
 history 10:423–427
 reading 10:431–432
 research implications 10:429–432
 system and contractions 10:423–425
 writing 10:428–429
 brain
 aging 6:514–515
 circulation 1:233–234
 drain 1:233–234
 bridge study designs 14:364
 Bristol Bay Native Corporation Education Foundation (BBNCEF) 3:28–29
 broaden-and-built theory 6:722–723
 Brown, leadership frameworks and taxonomies 4:406
 brown faces, books with 10:692–693
 bubbling knowledge economy 1:282–283
 Buber's philosophy of dialogue 6:602
 buggy-based approach 14:462
 building blocks 14:210–213
 item pool 14:210–211
 item selection 14:211–212
 person parameter estimation 14:211
 start of test 14:211
 bureaucracy 4:103–104
 business sector, IEBs' practices impact on GBs and 4:425–426
- ## C
- C-I model. *See* construction-integration model (C-I model)
 CA. *See* cluster analysis (CA)

- Cabaret academy 10:524–526
- CAF. *See* conceptual assessment framework (CAF)
- CAMFED. *See* Campaign for Female Education (CAMFED)
- Campaign for Female Education (CAMFED) 9:472
- intervention 9:472–474
- CAMPUS ASIA program 1:452
- campus visit 8:338
- Canada 3:14–15
- principal preparation 4:359–360
- system leader and system leadership 4:132–133
- Canadian case 13:150–151
- Canadian State 2:39
- Canadian teachers, testing person-environment fit model with 5:332–333
- Canadianness, perfect stranger and contradictions of 2:276–277
- capability approach 9:455–456
- CARE
- building practices for disruption and critical conversations to build equitable classrooms 10:477
- coaching with 10:476–477
- developing trusting relationships through caring for/with preservice teachers 10:477
- existing literacy teaching practices in placement classroom critically examined as starting point 10:477
- reflecting within community 10:477
- using video and student work to focus on students as source of feedback 10:477
- care-related technologies, UK research on 12:356–358
- career pathways
- federal investment in career pathways in US 8:472–473
- in international context 8:469–470
- models and approaches 8:471–472
- research needed on career pathways 8:474
- in United States 8:470–471
- WWC meta-analysis on 8:473–474
- Caribbean form of communication 12:183
- caring assessments, CBAs and 14:508–509
- caring teachers 6:101
- Carol Bacchi's WPR approach 8:358
- CART. *See* classification and regression trees (CART)
- CAS. *See* complex adaptive systems (CAS)
- cascade model 5:470
- case methods in teacher education 5:405–413
- case study
- design 12:86–87
- activities 12:87
- choosing cases 12:86
- example of case study 12:88
- humanistic study 12:87
- issues 12:87
- narrative study 12:87
- size 12:86
- triangulation 12:87
- types 12:86–87
- research 12:539–540
- applying three frameworks 12:542
- case study design and case selection 12:543–544
- examining cases within mixed methods study 12:540–541
- frameworks for combining 12:540–542
- intersecting two approaches to form mixed methods case study 12:542
- key decisions when combining 12:543–546
- mixing methods within case study 12:541–542
- case-based knowledge system for teaching 5:410
- case-based learning (CBL) 6:707–709
- design of learning task in 6:707–708
- level of student/teacher centeredness in 6:708
- sequence of information acquisition in 6:708–709
- case-based reasoning (CBR) 14:462
- CASEL. *See* collaborative for academic, social, and emotional learning (CASEL)
- CAT. *See* computerized adaptive testing (CAT)
- catalysts for school improvement 4:199–200
- categorical data analysis 14:575
- measures of association 14:580–581
- notations 14:575
- probability 14:577
- sampling distributions 14:576–577
- sampling processes 14:575–576
- tests of no association 14:581–582
- unknown probability estimation 14:578–580
- categorical model
- achievement growth 14:384–385
- advantages and disadvantages 14:385
- determining growth adequacy 14:385
- model overview 14:384–385
- modeling requirements 14:385
- categorical variables 14:723–724
- Cattell model 6:203
- Cauchy distribution 14:714
- causal models
- articulating causal effects using potential outcomes 14:671–672
- causal effects unconfoundedness
- established 14:674–678
- covariates generate confounding of treatment and outcome 14:672–674
- estimating causal effects unconfoundedness 14:678–679
- estimators that combine propensity scores and outcome regression models 14:677–678
- matching 14:675–677
- regression 14:674–675
- sensitivity analysis to unmeasured confounding 14:678
- causal studies 14:698
- caution indices family 14:275
- caveat emptor 9:219–220
- CB design. *See* counterbalanced design (CB design)
- CBAs. *See* conversation-based assessments (CBAs)
- CBES. *See* community-based educational spaces (CBES)
- CBL. *See* case-based learning (CBL)
- CBPR. *See* community-based participatory research (CBPR)
- CBR. *See* case-based reasoning (CBR)
- CBT. *See* competency-based training (CBT)
- CBT systems. *See* computer-based training systems (CBT systems)
- CCAF. *See* critical curriculum analysis framework (CCAF)
- CCCS. *See* Center for Contemporary Cultural Studies (CCCS)
- CCCs. *See* crosscutting concepts (CCCs)
- CCM. *See* computational cognitive models (CCM)
- CCT. *See* Critical Cosmopolitan Theory (CCT)
- CDR. *See* curricular design research (CDR)
- CDS. *See* cognitive design system (CDS)
- cedar basket 3:190
- CELÄNEN I SNEPENEK ŁTE 3:90–92
- respecting honoring and living indigenous ways of knowing and being 3:90–92
- Center for Contemporary Cultural Studies (CCCS) 2:25
- centering women of color feminisms 12:57–58
- Central America, transfer credit 8:348
- central limit theorem (CLT) 14:755–756
- centrality of race 12:70
- centralization mode 7:76–77
- centralized curriculum
- centralization work 7:22–23
- historical approach to 7:20–21
- making sense of centralization 7:23–25
- certification and licensure testing
- credentialing classification accuracy and validation 14:448–449
- designing operational testing system for processing examinations 14:444
- designing test items 14:443–444
- determining credentialing cut score 14:446–447
- efficient test design and form assembly 14:447–448
- practice analysis 14:441–443
- score scales 14:445–446
- certification for principalship (CFP) 4:362
- CFA. *See* confirmatory factor analysis (CFA)
- CFG. *See* conceptual Flow Graphics (CFG)
- CFP. *See* certification for principalship (CFP)
- chained equating methods 14:242
- chalkboard, use of 13:67–69
- challenge-based learning (ChBL) 6:709

- Charles Joseph Minard, graphic poetry of 14:831–832
- CHAT. *See* cultural historical activity theory (CHAT)
- ChBL. *See* challenge-based learning (ChBL)
- chi-square test statistics 14:581
- Chicago Consortium leadership studies, results of 4:245–246
- Chicago Public Schools (CPS) 4:244
- school leadership studies in 4:244–245
- childhood/children 10:606–607
- development and learning 4:341
- digitally mediated literature for 10:623–624
- earliest insights
- children's mobilization of action schemas 10:40–43
 - early writing patterns 10:42–43
 - foundational schemas 10:40–41
 - meaning attribution and written language 10:43–45
 - pre-phonological roots of literacy 10:39–40
 - sensitivity to orthographic patterns 10:41–42
- equity in early childhood education 5:93–98
- focus 10:184
- Greek-Cypriot children's literature 10:673–674
- historical development of Greek children's literature 10:672
- interactivity 10:626–627
- interests
- as inquiry and identity development 7:176–177
 - theoretical framings of 7:174–176
- literation, explorations in visualization in 10:666–669
- literature 10:609
- literature between 1950s and 1970 10:665–666
- literature in Greece and Cyprus
- Greek-Cypriot children's literature 10:673–674
 - historical development of Greek children's literature 10:672
 - springtime of children's literature in Greece 10:673
 - talent and influence of Eugene Trivizas 10:674–675
- literature in Latin America 10:644–661
- materiality 10:625–626
- metafictive literature for 10:622–623
- multimodality 10:624–625
- narrative structures and metalepsis 10:628–629
- paratextual features 10:627–628
- parental involvement for children of different ages 4:268
- parental unemployment and children's educational outcomes 1:118–127
- postmodern and digital turns in 10:620–622
- rights 7:429–431
- school garden and nature-based participatory research with 2:216–217
- springtime of children's literature in Greece 10:673
- talent and influence of Eugene Trivizas 10:674–675
- traits and trends of Greek children's literature 10:675–676
- transmedial narratives and storytelling 10:629–630
- US teachers learned about indigenous children in Guatemala 10:139–140
- Chile, education in developing countries 1:208
- Chimamanda Ngozi Adichie 8:357–358
- China
- challenge 1:658–664
 - enduring legacy 1:660–661
 - genesis and development 1:658–659
 - reform and opening era 1:661
 - rural-urban differences 1:661
 - science and innovation 1:659
 - comparisons with 5:712–716
 - complexity of educational leadership in 4:478–479
 - engagement with Southeast Asia in higher education 1:666–672
 - implantation and development of 8:23–24
 - learning from 4:480–481
 - re-turning to cultural-historical-philosophical foundations of China's curriculum knowledge 7:69–70
 - shapes educational leadership in 4:479–480
 - uneven rise of 4:476–478
 - university in East Asia 8:26
- Chinese children's literature 10:665–666
- back to childhood and recreation 10:665–666
 - at beginning of 20th century 10:662–664
 - multiple perspectives on understanding Chinese children's literature 10:666–669
 - politicalization of children's images 10:664–665
- Chinese culture 6:607
- Chinese literacy acquisition
- challenges embedded in Chinese orthography 10:750–752
 - Chinese writing system 10:748–750
 - language and writing system in 10:747–753
 - literacy acquisition in Chinese 10:753–755
 - oral language in 10:747–748
- Chinese Minzu diversity and education
- Chinese education to contextualize 10:760
 - as complement to reflecting on languages and literacies in education 10:760–761
 - as complex notion 10:761–762
 - education as mirror for education in other contexts 10:763–764
 - education taking place 10:764–767
 - as political-economic and/or cultural-linguistic notion 10:762–763
- Chinese Minzu education
- identify Minzu knowledge 3:221–222
 - Minzu as companion to indigeneity 3:219
 - Minzu knowledge in Chinese education 3:220–221
 - Mongolian STEM teachers reflecting on use of Minzu knowledges in teaching 3:222–226
 - Problematicizing indigeneity for 3:217–218
- Christian missionaries, impact of 10:775
- church literacies 10:201–202
- citizen science in STEM education
- awareness to science and environment 11:227
 - case studies 11:228–230
 - citizen science in schools 11:226–228
 - civic action 11:227–228
 - classroom as platform for community engagement 11:229–230
 - enhancing learning and environmental stewardship through 11:228–229
 - integration 11:230–231
 - learning about complexity of science through 11:229
 - learning through citizen science 11:225–226
 - societal aspects of science 11:226–227
- citizenship 2:38–39
- Canadian State and multiculturalism 2:39
- immigration to security to pandemic politics 2:41–43
- implications for education for citizenship 2:43
- racism and dispossession 2:39–43
- research centered on 10:494
- CIVED. *See* civic education study (CIVED)
- civic education 1:415–416, 6:436
- civic education study (CIVED) 14:349
- civic learning
- re-envisioning 1:85–87
 - clarity 1:86
 - communion 1:86
 - compassion 1:87
 - courage 1:87
 - creativity 1:86–87
 - with social media 6:435–437
 - affordances and constraints of social media for civic engagement 6:436–437
 - civic engagement on social media 6:435–436
 - emerging issues in education 6:437
 - understanding civic learning and engagement on social media 6:436
- civic life, re-envisioning 1:84–85
- civic media literacies 2:307
- civic pedagogy
- of empowerment 1:89–90
 - of enlightenment 1:88–89
 - insight 1:88
 - inspiration 1:88–89
 - re-envisioning 1:87–90
 - social advocacy 1:90

- civic pedagogy (*continued*)
 social inquiry 1:89–90
- civic pluralism 6:259
- civil rights, 1960s and 8:425
- civil war amendments and equal rights
 10:272–273
- class-economic status 9:251
- classic cognitive theories 6:677
- classical conditioning 6:14–17
- classical test theory (CTT) 14:21–26, 14:51
 congeneric 14:54
 generalizability 14:55–56
 interpretational contexts 14:54–56
 London school and development of
 14:2–5
 measurement precision 14:25
 methods 14:23–24
 models 14:22–23, 14:52–54
 parallel 14:53
 real-data example 14:24–25
 reliability 14:55
 as special case 14:54
 test design 14:22
 test development 14:56–57
 true score 14:51–52
 validity 14:55
- classification and regression trees (CART)
 14:812
- classroom
 affordances for mathematics learning
 6:331–332
 African-American language from
 10:312–313
 assessment 14:519–520
 action 14:526
 classroom assessment literacy
 14:523–527
 components of 14:524–527
 consequences 14:526–527
 formative 14:522–523
 foundational concepts 14:525–526
 foundational influences 14:526
 identify assessment literacy 14:523–524
 literacy 14:523–527
 main purposes of 14:520–523
 summative 14:520–522
 autonomous and ideological models of
 literacy in 10:310–311
 based research 10:557
 communication about knowledge and
 rationality 10:314–315
 ecology of learning in 10:559
 engage religious literacies pluralistically,
 practices in 10:213–214
 exemplar from classroom-based research
 project 5:521–524
 fostering responding to and creating
 multimodal texts in 10:507–512
 optimal learning in 6:369–370
 practice 13:89–95
 extending socio-psycholinguistic
 viewpoint for 10:841
- climate change 1:235, 3:156–157, 5:745,
 8:208
 in education policy 1:598–599
 micro, meso, and macro sociopolitics of
 climate change in action 1:597–598
- climate justice & land sovereignty 2:368
- clinical outcome assessment (COA) 14:96
- CLT. *See* central limit theorem (CLT);
 communicative language teaching
 (CLT)
- cluster analysis (CA) 14:627
 data preparation for 14:631–632
- cluster models types 14:628–631
 k-means 14:628–629
 partitioning 14:628–629
- cluster randomized trials (CRTs)
 14:784–785
 power calculation for two-level CRTs
 14:789
 power for two-level CRTs 14:786–789
- cluster-level covariate, model with 14:785
- cluster-randomized design (CRD) 14:699
- clustering approaches 14:627–628
 assessing fit and validity of clustering
 solutions 14:632
 cluster models types 14:628–631
 of countries 14:632–633
 data preparation for cluster analysis
 14:631–632
 of educational institutions, classrooms, and
 teachers 14:633
 of educational test items 14:634
 illustrative applications of clustering in
 education 14:632–634
 of learner knowledge 14:634
 of students 14:633–634
- CNN. *See* convolutional neural networks
 (CNN)
- co-authorship, transnational academic
 collaboration in 1:312–314
- co-citation analysis 4:551–553
- co-created/co-creating/co-creation 7:236
 business course, University of St Andrews,
 Scotland, UK 7:240
 chemistry course at Uppsala University,
 Sweden 7:239
 courses within Masters in Education,
 Universiti Utara Malaysia, Sintok,
 Malaysia 7:240
 of curriculum 7:239–240
 curriculum in higher education 7:235–237
 essay titles, University of Vienna, Austria
 7:240
 interdisciplinary program, University of
 Technology, Sydney, Australia
 7:239–240
 science skills curriculum at University of
 Glasgow, Scotland, UK 7:239
- co-occurrence analysis 4:553
- co-regulation of learning 13:38–39
- COA. *See* clinical outcome assessment
 (COA)
- coaching 4:384
 in action 5:149–150
 early career teachers 5:148
 investment to professional capital in school
 5:149
 model 5:470
- purpose of 5:148–149
 in teachers' professional development
 5:145–147
 Coda 5:737–738
 codifying racism 10:274
- coding
 method 12:113
 practices 12:303–304
 reliability of 14:848
 study quality and methodological variables
 14:848
 theoretically important variables
 14:847–848
- cognition-centered design principles
 affordances 14:175
 designing tasks for process data modeling,
 analysis, and reporting 14:179–181
 digital assessment tasks and technology-
 enhanced items 14:172–173
 gestalt principles of visual perception
 14:174–175
 goals and scope 14:171–172
 interactivity 14:175
 multimedia 14:176
 principles for 14:173–179
 visual media 14:176–179
 working memory and constraints of human
 information processing
 14:173–174
- cognitive apprenticeship framework
 6:572–573
 learning content 6:572
- cognitive demand 6:332
- cognitive design system (CDS) 14:155
- cognitive development and adolescent
 learning 6:487–488
 concrete operations stage 6:487
 formal operations stage 6:487
 limitations of Piagetian perspective 6:488
 metacognitive and information processing
 perspectives 6:488
 Piagetian perspective of adolescent learning
 and development 6:487
 preoperational stage 6:487
 sensorimotor stage 6:487
 Vygotskian perspective of adolescent
 learning and development 6:488
- cognitive elaboration perspective on
 communication process 6:593
- cognitive interviews 11:451
- cognitive load 14:174
- cognitive load theory
 biologically primary, generic-cognitive skills
 and biologically secondary,
 domain-specific skills 6:128
 borrowing and reorganizing principle 6:129
 collective working memory effect 6:132
 element interactivity effect 6:131
 environmental organizing and linking
 principle 6:129
 evolutionary educational psychology
 6:127–128
 expertize reversal effect 6:131
 generic-cognitive knowledge and domain-
 specific knowledge 6:128

- goal-free effect 6:130
guidance fading effect 6:131
human cognitive architecture 6:127–129
human movement effect 6:132
imagination effect 6:132
information store principle 6:129
instructional consequences of human
cognitive architecture 6:130–132
isolated elements effect 6:132
modality effect 6:131
narrow limits of change principle 6:129
problem completion effect 6:130
processing biologically secondary
information 6:129
randomness as genesis principle 6:129
redundancy effect 6:130–131
self-management effect 6:131
some cognitive load effects 6:130–132
spacing effect 6:132
split-attention effect 6:131
transient information effect 6:132
variability effect 6:132
worked example effect 6:130
cognitive neuroscience 6:226
cognitive revolution 6:324, 6:365
cognitive skills 14:485
cognitive strategies 6:356
cognitive theories 6:210
of learning 6:5–6
cognitive-behavior modification 6:13
cognitive-behavior therapy 6:13
coherence 6:212–213
creating context for 4:212–213
coherent educational systems 6:549
cohorts 4:375
COLA. *See* cost of living adjustments
(COLA)
cold war and public diplomacy 1:641–642
collaboration
learning 5:702
as strategy for promoting equity in
education 4:281
anomalies and organizational learning
4:287–289
approach 4:282
family of approaches 4:283–284
framework for analysis 4:284–286
global agenda 4:281–282
inquiry-based developments 4:283
nature of practice 4:289
sense of process 4:286–287
social learning 4:289
collaborative for academic, social, and
emotional learning (CASEL) 6:95
SEL framework 6:95–96
collaborative functions on student
achievements 11:221
collaborative intelligence 6:263
collaborative learning 5:702
in STEM
assess and analyze TECL in STEM
11:211–212
benefits of 11:208
facilitating group written discourse
11:209–210
generate, deposit, and visualize collective
knowledge 11:211
guiding collaborative process 11:210
parameters and forces of 11:207–208
remote collaborative learning
environments 11:208–209
across space 11:208–209
collaborative model 5:384
collaborative problem solving (ColPS)
13:234–235
classification framework for collaborative
problem-solving skill assessments
13:235–237
current landscape 13:239–241
future directions and agendas 13:241–242
past landscape 13:237–239
collaborative problem solving skills (CPS
skills) 14:494–495
applications for evidence identification and
accumulation for 14:499–500
challenges with CPS assessment
14:495–496
classification framework for collaborative
problem-solving skill assessments
13:235–237
context for example applications
14:496–497
example application for operationalizing
CPS 14:497–499
importance of 14:494
learner behaviours 14:500
profiles 14:500
collaborative reasoning (CR) 10:330–335
academic vocabulary 10:332
analogical reasoning 10:333
argumentative reasoning 10:332–333
creates opportunities for exploratory talk
10:334
engages students and teachers in social
exchange of ideas 10:334–335
fosters zone of proximal development
10:334
language coherence 10:332
listening and reading comprehension
10:331
multilink causal reasoning 10:333
pragmatic awareness 10:332
promotes collaborative learning 10:334
promotes development of language
10:331–332
promotes development of thought
10:332–333
social mechanism of 10:334–335
speaking and writing proficiency
10:331–332
transfer of thinking and reasoning
10:333
collected data development 12:367–368
quantitizing 12:367–368
collective leadership development, model of
4:198–199
collective learning 5:423–424
college reading
assessment 10:466
contextualization 10:469
curricular structures, past and present
10:465–466
entextualization 10:470
expectations and purposes of 10:469–470
instructional foci 10:466
perspectives focused on learner attributes
10:468
positioning theory 10:470
professional associations 10:466–467
theoretical assumptions 10:467–470
college student activism, abbreviated
historical overview of student
activism 2:399–402
Colombia, indigenous bilingual school in
10:140
colonialism
and ambivalent legacies 1:389–390
as culture, unlearning and ethical
relationality 2:277–278
and impact on globalization 1:110
color-evasive racism 5:417–418
ColPS. *See* collaborative problem solving
(ColPS)
combinatorial speech, onset of 6:469–470
combined MMR designs 12:431
combining methods
for causal explanation 12:444–445
for generalization 12:445–446
combining words 6:470
commodification
of education-choice 1:65
process 8:234–235
of self 1:66–67
of students-accountability 1:65–66
common core in France 7:13–14
common item equating, best practices for
14:243–244
common standard setting methods
14:374–378
common-effect model 14:849–850
communicating/communication
and conflict management in Thailand
4:496
information for students 10:557
mixed methods study design procedure
12:764–768
strategies 10:831–832
communicative approach
in curriculum design 7:201
into practice 6:370–371
communicative competence 6:366
communicative language teaching (CLT)
6:364
learner-centered model of 6:367–368
teacher-led model of 6:366–367
communicative purposes, expanding 6:470
communicators 4:487
community 9:183–189
collaboration 2:369–370
community-based participatory approaches
in qualitative research
building community capacity
12:171–173
culture, curriculum and pedagogy
12:174–175

- community (*continued*)
- ethics in participatory and action research 12:176
 - information technology and social media for community based action research 12:173–174
 - localized educational practices 12:170
 - organizational and systems change 12:175–176
 - situating community-based, participatory action research 12:169–170
 - community-based participatory MMR designs 12:430
 - community-campus relationships and partnerships 8:482–483
 - community-centered indigenous language recovery, restoration, revitalization, and renewal
 - collaboration, networking & resources 10:793–794
 - lighting way home 10:787–788
 - community-engaged teacher education 5:376
 - funds of knowledge and cultural intelligence to dismantle deficit thinking 5:378–379
 - power of dialogic interactions between teachers, students, and communities 5:377
 - raising critical consciousness by discussing evidence with social impact 5:377–378
 - development
 - HRE in 2:62–63
 - learning in 2:325–326
 - engagement 3:214–215
 - developing STEM communities of practice 11:240
 - establishing STEM professional-school collaboration 11:240
 - highlighting STEM professional contribution to society 11:241
 - practices with 11:238–240
 - principles to support STEM education across various settings 11:237–238
 - recommendations 11:240–241
 - in STEM education 11:235
 - students STEM aspirations 11:237
 - importance of generational influences and connectivity to 10:181
 - lead 3:92
 - literacies 10:200
 - mapping 2:215
 - and networks 6:411
 - of practice framework 6:438
 - of practice model 5:470
 - resistance from 4:162–163
 - schools as promising strategy 4:278
 - service-learning as community-campus engagement 8:477–478
 - community-based educational spaces (CBES) 2:297
 - activism and organizing 2:298
 - challenges in field 2:299
 - considering sociopolitical context 2:299
 - features of 2:298
 - historical context of 2:297–298
 - pedagogical and philosophical diversity 2:298
 - possibility and future 2:299–300
 - relationships, social capital, and advocacy 2:298
 - community-based participatory research (CBPR) 12:430–431, 12:568–569, 12:593
 - advantages of intersecting MMR with AR and CBPR approaches in education 12:570–571
 - challenges of intersecting MMR with AR and CBPR approaches in education 12:571
 - common features of 12:569–570
 - methodological considerations of intersecting MMR with AR and CBPR 12:571–576
 - comparative education
 - counter-balancing inequity and injustice in comparative education research 1:295
 - end of “history” and fragmentation 1:14–15
 - globalization and governance 1:15
 - invention of traditions 1:11–12
 - traditions and tensions 1:12–13
 - comparative judgment 13:73–74
 - educational applications of 13:75–76
 - reliability and validity of 13:76–77
 - comparative model fitting 6:204
 - John Carroll and key outcome of 6:204
 - comparative teacher education research in turbulent times
 - climate change 5:745
 - global instability 5:745–746
 - market models and performativity in teacher education 5:741–742
 - national populism 5:742–745
 - resources of hope 5:746–748
 - competence 4:258–259
 - arranging for sense of 10:404
 - competence-based curricular reforms in two contrasting education systems 7:11–14
 - mediate conflict 4:258–259
 - nurture competence of lead 4:259
 - competency
 - assessment in workplace 13:97–98
 - competency-based approach
 - attracting prospective teachers using 5:333
 - selecting prospective teachers using 5:334–336
 - for selecting teachers, implementing 5:336
 - for teacher recruitment 5:332
 - examples of competency-based recruitment outside of education 5:331–332
 - person-vocation fit and competency-based recruitment 5:331–333
 - competency-based training (CBT) 8:491
 - background on Australian vocational education and training system 8:491–492
 - current nature of curriculum 8:496–497
 - extracts from two units of competency and assessment requirements 8:499
 - first decade of 8:493–494
 - getting out of corner 8:498–499
 - method and evidence base for chapter 8:492–493
 - second decade of 8:494–495
 - system painted itself into corner 8:497–498
 - third decade of 8:495–496
 - complementary model 5:383
 - complex adaptive systems (CAS) 12:616
 - adaptations to emergent CAS conditions
 - inform integration procedural decisions 12:621–622
 - attention to interactions of CAS informs data source and sampling decisions 12:620–621
 - representations of nested structure of CAS informs initial integration decisions 12:617–620
 - transparency of accounts of CAS outcomes informs credible interpretation decisions 12:623–624
 - complex crises, emerging issues and challenges 2:71
 - complex education systems 9:28–29
 - inclusion and diversity as educational value in complex system 9:31–32
 - path forward 9:32–33
 - theoretical models for understanding complexity in 9:29–30
 - complex repeated measures designs 14:572–573
 - complexity
 - complexity-informed mixed methods research in education design practices, implications for 12:624
 - theory 4:68
 - comprehension skills 6:356
 - comprehensive high school, basic issue and 9:80–81
 - comprehensive sexuality education (CSE) 2:171
 - community, culture, ceremonies 2:173–176
 - masculinity, culture and context 2:172–173
 - theorizing masculinity beyond hegemony 2:173
 - comprehensive teacher evaluation system 13:288–290
 - evaluation criteria 13:288–290
 - evaluation purposes 13:288
 - sources for collecting relevant data 13:290
 - computation 14:778–780
 - computational cognitive models (CCM) 14:411–412
 - computational linguistic analysis 14:412
 - computational power, increased use of 14:11
 - computational thinking (CT) 11:23–24, 11:312

- assessment of 11:90
- borderlands 11:92–93
- as core 21st century skills relevant for all teachers 5:265–266
- dual challenge for CT to address 11:312–313
- in educational policy and practice 11:89–90
- framings of 11:93
- heart of computer science 11:91–92
- history of 11:88–89
- learning spaces for teacher candidates 11:316–317
- preparing teachers with CT for equity pedagogy 11:314–315
- preparing teachers with CT pedagogy 11:313–314
- as research construct 11:91–93
- STEM + C3 urban teacher residency program 11:313–317
- teachers connect CT, equity, and data 11:317–318
- computer simulations 6:681
 - implications and recommendations for future research 6:686–687
 - learning with simulations, games, and hybrid forms 6:685–686
 - typology of simulations and games 6:681–685
- computer-assisted instruction 14:461–462
 - CAI to ITS 14:462
- computer-based performance assessments in higher education 8:314
- computer-based process data in educational measurement 14:418
 - challenges, and future directions of 14:424–425
 - complex problem solving 14:423–424
 - current state of 14:420–422
 - featured in article 14:417–418
 - formative assessment 14:421–422
 - history of 14:419
 - psychometric models 14:420–421
 - public computer-based process data repositories 14:419–420
 - role of 14:418
 - summative assessment 14:422
- computer-based training systems (CBT systems) 6:637
- computer-supported collaborative learning (CSCL) 6:49–50, 6:77
 - designing and supporting CSCL 6:590
 - discourse analytics in 6:77
 - future perspectives 6:590–591
 - productive 6:588
 - shared meaning making in social interactions 6:589
- computerized adaptive testing (CAT) 14:209, 14:276
 - building blocks 14:210–213
 - effects of adaptive testing 14:213–215
 - works 14:209–210
- computing effect sizes of data evaluation 14:846–847
 - correlational effects 14:847
 - count based effect sizes 14:847
 - effect sizes 14:847
 - mean differences 14:846
 - measurement error 14:847
- conceptual assessment framework (CAF) 14:163
- conceptual Flow Graphics (CFG) 11:157–159
- conceptual interests 10:29
- conceptual knowledge 6:382
 - assessing conceptual knowledge competencies 13:317
 - cognitive competencies 13:317
 - interpersonal competencies 13:318–319
 - metacognitive competencies 13:317–318
 - personal competencies 13:318
- conceptual memory, embodiment of
 - for abstract concepts 6:70–71
 - for events 6:71
 - for numbers 6:70
 - for objects 6:69
- conceptual models 12:746
- conceptual tool for emergent disciplines 6:311–312
- conceptualization of policy isomorphism 1:271–273
- conceptualizing/conceptualization knowledge 6:210
 - teacher educator development as learning problem 5:477–478
- conduct user-centered needs assessments, focus groups and surveys to 14:223–226
- conducting large-scale MM studies in 12:627–628
- conference submission system, self-reported epistemologies and methodologies in 10:904
- confirmatory factor analysis (CFA) 14:607, 14:613–614
 - Bayesian 14:615–617
 - common factor model 14:608–609
 - data considerations 14:607
 - multilevel 14:609–611
 - sample size and power 14:607–608
 - second-order factor analysis 14:611–613
- conflict
 - detection 6:355
 - and migration 5:746
 - regulation 6:355
 - resolution 6:355
- conjugacy, analytical solution by 14:778–779
- conservationist/resourceist current 2:212–213
- constructed-response items 14:187–188
 - automated scoring of 14:397–403
 - developing and scoring constructed-response item types 14:190–191
- constructed-response task (CR task) 14:397
- construction via literacy 10:210–212
- construction-integration model (C-I model) 6:354
- constructive generating mode 6:695
- constructive mode
 - of cognitive engagement 6:691–692
 - strategies 6:694–695
- constructivism 12:40–41, 12:385, 6:368
 - and mixed methods evaluation 12:533–534
- constructivist 13:42
- constructivist theory of school reform 4:227–228
- contact theory (CT) 6:677
- contemporary contexts for learning
 - addressing changing contexts in formal education 6:264–265
 - on contexts of learning 6:258
 - digital media as major contemporary contextual factor 6:262–264
 - learner diversity as major contemporary contextual factor 6:260–262
 - learning in formal and informal settings 6:258–259
 - social functions of education in context 6:259–260
- contemporary curriculum policies
 - Bernstein's pedagogic models 7:10–11
 - curriculum policies in France and Quebec 7:11–16
- contemporary standard setting methods
 - common standard setting methods 14:374–378
 - post-standard setting meeting activities 14:378–379
 - pre-standard setting meeting activities 14:370–373
 - typical standard setting meeting activities 14:373–374
- content
 - area literacy 10:449–451
 - in practice 10:450
 - gamification 6:685
 - sources of 4:364
 - national and local contexts 4:364
 - professional and academic content 4:364
 - student and community need 4:364
 - teacher interventions 6:595–596
- content-source integration model (CSI model) 6:355
- contexts
 - of learning 6:258
 - variables 14:353–354
- contextually relevant approach 4:189–190
- continental pedagogy 7:246–248
 - as "science" 7:252–253
 - and curriculum studies 7:253–254
- contingent leadership 4:49–50
- continuous probability distributions
 - See also* discrete probability distributions
 - beta distribution 14:712–713
 - Cauchy distribution 14:714
 - Dirichlet distribution 14:715
 - exponential distribution 14:711
 - F distribution 14:715
 - gamma distribution 14:713–714
 - lognormal distribution 14:711–712
 - normal distribution 14:711
 - student's *t* distribution 14:712

- continuous probability distributions
(*continued*)
uniform distribution 14:710–711
 χ^2 distribution 14:714
- contrasting groups method 14:377
- contrasting theories of learning 6:364
- conventional approaches to work and learning 2:237–238
- conventional humanistic qualitative research, sampling processes and strategies in 12:202–203
- convergence 2:321–322
- convergent MMR designs 12:427
- conversation skill development 6:471
- conversation-based assessments (CBAs) 14:505–506
and caring assessments 14:508–509
development and scoring 14:506
dialog systems 14:504–505
examples 14:507–508
and individual differences 14:510–511
and student emotions 14:509–510
- conversational agents 6:637
in informal learning setting 6:644
learning environments with 6:639–644
- convolutional neural networks (CNN) 14:689
- cooperative learning 6:593–599
- Copernican revolution 1:692–693
- CoPO. *See* cycle of Peer Observation (CoPO)
- copying-detection and similarity analyses 14:300
- core analysis methodology 14:360–364
- core domains 14:352
additional and innovative domains 14:353
- core MMR designs 12:427
- core vocabulary 10:369
- corporate models of change 5:33–34
- corporations impact on higher education 8:32–33
skilled and diverse workforce 8:32–33
strategic relationships 8:33
- correlation coefficient 14:730
- correlations
and associations 14:734–748
considerations with 14:747–748
correlation and causation 14:747
history of 14:734–735
nonparametric correlations 14:741–747
parametric correlations 14:735–739
PCA of correlation matrix 14:592–594
restriction of range 14:747–748
spurious correlations 14:747
of subscores 14:292–293
types of 14:735–747
- corruption 1:421, 1:662
- Cosmopolitan literacies
and asymmetric ignorance 10:710–711
as ways of worldmaking 10:712
worldmaking as practice 10:713–714
- cosmopolitanism literacy 10:453–454
- cost of living adjustments (COLA) 8:195
- counter story-telling 12:51–52
- counterbalanced design (CB design) 14:238
- course co-construction 3:137–138
- course design teams at Elon University, North Carolina, USA 7:239
- COVID-19 pandemic 1:636–637, 9:87–88
and challenges to existing policy and practices 5:63
changes in teacher assistant roles 9:120
and child protection 9:88
disconnecting relationships among students 9:121
higher education in Africa 8:91–92
HRE in era of black lives matter and 2:57
impact 8:93–97
on core missions of universities 8:94–95
issues exposed by 9:119–121
and mega-trends 5:103
online education as response to 1:150–151
overreliance on technology 9:119–120
pandemics 8:92
popular discourses of school space and 10:930–936
physical, affective, and hyperreal productions of school space during 10:931–932
physical and affective reproduction of school space at home
quality turn, return of lifelong learning and 1:516–517
responses 8:93
effect on higher education institutions 8:94
and right to education 1:407–409
strategies to 8:97–98
teachers' work in context of 5:636–647
teachers' working conditions during 5:638–640
teacher voice and control over working condition 5:639–640
work intensification 5:638–639
techno-social discourses of student affect and mental health during 10:933–934
- CPS. *See* Chicago Public Schools (CPS)
- CPS skills. *See* collaborative problem solving skills (CPS skills)
- CR. *See* collaborative reasoning (CR)
- CR task. *See* constructed-response task (CR task)
- Cramer's V 14:742–743
- CRD. *See* cluster-randomized design (CRD)
- Creative Britain, translating power of 1:286
- creative interests 10:29
- creativity 11:64–65
research 6:135–136
operant theory and 6:136–137
Piaget's constructionist theory and 6:137–138
relationship between creativity and learning 6:136–138
in STEM education 11:65
cognitive inhibition 11:68–69
constructing innovative and operational evaluation system for STEM education 11:72–73
cultivation of 11:69–71
cultivation of creativity through projects 11:69
developing effective STEM curriculums 11:71
development of adolescents 11:65–66
exploring effective STEM classroom teaching 11:71–72
factors affecting creativity in educational environment 11:66–69
future research prospects 11:71–73
individual 11:70
knowledge 11:68
personality 11:68
scientific and technological creativity 11:65
strategies 11:71
- credit for experience 8:344
- independent study 8:344
- portfolio demonstration 8:344
- work experience 8:344
- credit hours 8:343–344
- credit recognition 8:344–345
Cambridge International Examinations 8:345
completion of learning opportunities 8:345
Excelsior College exams 8:345
non-credit adult classes 8:345
- criminalization, legal 2:181
- critical approaches to work and learning 2:238–240
- Critical Cosmopolitan Theory (CCT) 1:110–111
- critical curriculum analysis framework (CCAF) 7:38–39
applying CCAF to curriculum policy text ensemble 7:39–44
- critical disabilities 10:611–614
- critical disability studies 10:612–613, 9:244
globalization of 9:244–245
- critical ethnography 12:51
- critical friendship should be mutual 5:466
- critical hip-hop 10:230–231
- critical inquiry communities 5:483–484
- critical legal theory 12:69
- critical literacy/literacies 10:225, 10:320–321
See also sensory literacies
biggest problems facing black community 10:225–227
comparative reading of newspaper reports 10:118–119
counterstory of 10:225
critical approaches to teaching Africanist presence in literature 10:228–229
critical hip-hop and literacy studies 10:230–231
critical media pedagogies 10:231–232
critical pedagogy of English language, literature, and literacy 10:227–228
exploring race and justice in America through literature and popular culture 10:229–230
in foreign language education 10:113–114
katakana project 10:115–116
linguistic landscape project 10:116–118

- through Nancy Fraser's theory of justice 10:219–222
- perspectives on children's and young adult literature 10:605–619
- powerful literacy education 10:228
- praxes in Japanese language classrooms 10:114–119
- recognition 10:221–222
- redistribution 10:220
- representation 10:222
- and social justice 10:217–218
- theories 10:607
- universal critical literacy and revolutionary love 10:233
- YPAR and critical literacies for social action 10:232–233
- critical media education
- blending drama and digital animation 10:500
 - blending promenade performance, literature and digital media 10:500–501
 - critical and creative 10:498–499
 - Le Cinéma 10:499–500
 - snapshot of media education in UK 10:498
- critical media literacy 10:77–80, 2:304–305
- asset-based vs. deficit-based approaches 10:79–80
 - crisis of trust 10:77
 - curriculum interventions 10:81–82
 - disparities of access and information 10:76–77
 - language acquisition 10:79
 - photography 10:80–81
 - putting theory into practice 10:80–83
 - research studies 10:82–83
- critical pedagogy as complementary frame for considering teacher educators' learning 5:478
- critical race 10:611–614
- critical race theory (CRT) 10:612, 12:68, 12:119, 2:22, 2:317, 2:378, 5:361
- and applications in disciplines beyond law 12:72–75
 - under attack 12:68
 - branches of 12:72
 - centrality of experiential knowledge 12:70–72
 - centrality of race 12:70
 - critical legal theory 12:69
 - definition and legal foundations 12:68–72
 - in educational research 12:72–75
 - implications of CRT for future educational scholarship 12:74–75
 - narratives, stories, and counterstories 12:72
 - and qualitative methods in educational research 12:73–74
 - reframing dominant educational theory through critical race perspective 12:73
 - researcher positionality 12:71
- critical realism 12:384, 7:85–89
- Archer's morphogenetic social realism 7:89–90
 - depth ontology 7:86–87
 - implications of critical realism for curriculum 7:88–89
 - intransitive domain and transitive domain 7:87–88
 - knowledge of powerful and powerful knowledge 7:88
 - similarities and differences between natural and social worlds 7:87
 - social realist school in sociology of education 7:90–92
 - transactional realism 7:92–93
- critical service-learning 6:678
- critical theory 12:46–50
- brief history of critical theory and Frankfurt School 12:46–50
 - Kant, Hegel, Marx, and modern foundations of critical philosophy 12:47
 - in qualitative educational research 12:50–52
 - applications of 12:51–52
 - in qualitative research in education 12:46
- critical thinking (CT)
- boundedness to principles 6:576
 - clarity and precision 6:576
 - as endpoint 6:583–584
 - intentionality 6:576
 - media and information literacy 6:584
 - reflexivity 6:583–584
 - sensitivity to context 6:576
 - as starting point 6:575–576
 - values 6:583
- criticality, research centered on 10:493–494
- cross-border higher education in Northeast Asia 1:448–450
- cross-cultural narrative imagination in hard times 7:401–402
- cross-cultural narrative inquiry
- composite auto/biographical narrative method to study bilingual/bicultural and multilingual/multicultural experience 7:393
 - metaphors to identify research themes and understand cross-cultural experiences 7:394
 - origins of 7:392–393
 - representing bilingual and bicultural lives and identities 7:393–394
 - tales in words and tales on canvas 7:394
- cross-cultural relevance 6:33
- cross-cultural studies (CXCS) 3:141
- faculty 3:146
 - indigenous/native model of well-being 3:149
 - mission statement 3:146
 - overall work ethics 3:148–149
 - reciprocity 3:147
 - relevance 3:147
 - respect 3:147
 - responsibility 3:147
 - shared Alaska native values 3:147–148
 - student and faculty videos 3:145
 - values 3:146–147
- cross-dialect communications 10:747–748
- crosscutting concepts (CCCs) 11:444–445
- and mathematizing 11:33–34
- CRT. *See* critical race theory (CRT)
- CRTs. *See* cluster randomized trials (CRTs)
- CSCL. *See* computer-supported collaborative learning (CSCL)
- CSE. *See* comprehensive sexuality education (CSE)
- CSI model. *See* content-source integration model (CSI model)
- CT. *See* computational thinking (CT); contact theory (CT); curriculum theory (CT)
- CTT. *See* classical test theory (CTT)
- cultural capital 2:248–249
- cultural historical activity theory (CHAT) 5:385, 5:477
- value in theorizing teacher educators' learning and activity 5:477
- cultural imperialism 6:607
- cultural justice, counterstory of 10:225
- cultural materialisms 10:608
- cultural political economy 1:469
- cultural practices, (re)activating and (re)centering language in/to 10:801–802
- cultural representations, challenge of 10:639
- cultural studies 10:608
- in education 12:417–418
 - illustrative examples of cultural studies in education 12:419–423
 - and mixed methods research 12:418–419
- cultural-discursive arrangements 5:35
- culturally relevant literacies
- conceptions of family literacy 10:182–183
 - of families of color 10:190
 - family literacy implementation in United States 10:183–186
 - importance of generational influences and connectivity to community 10:181
 - origins of family literacy programs 10:180–181
 - reenvisioning family literacy for families of color 10:181–182
- culturally relevant pedagogy 7:358–359
- culturally responsive leadership 4:65
- culturally responsive pedagogy, impact on 1:109
- culturally tailored interventions 12:680–683
- examples of MMR for culturally tailored interventions 12:682–683
 - facilitating participatory action research for capacity building 12:681–682
- culture(s) 2:21
- collapsing great divide 10:297–300
 - cultural studies and education 2:25–26
 - deconstructing predominantly white cultures 8:417
 - developing multicultural outlook 8:421
 - environment of 7:81
 - great divides 10:296–297
 - and indigeneity 2:26–28
 - learning about cultures 8:420
 - matters in SLD identification practices 10:283–284

- culture(s) (*continued*)
- methodologies and research constructs 10:300
 - multicultural education 2:23–24
 - recognizing and deconstructing white culture 8:420–421
 - recognizing legitimacy of multiple cultures 8:421
 - roles in self-efficacy formation 6:252
 - in sociology and anthropology of education 2:21–23
 - understanding culture 8:420–421
 - unsettling local 10:303–304
 - ways with words 10:298–299
- curiosity 6:142
- definitions, characteristics, and case examples 6:141–143
 - development 6:144
 - educational implications 6:143–145
 - interest development 6:144–145
 - sustained and developing individual interest 6:143
 - triggered situational interest that is sustained 6:142–143
- curricula, premises for knowledge divisions in 7:257–258
- curricular design research (CDR) 7:202
- curriculum 10:565–566, 4:343–344, 4:375, 7:333–334, 7:340, 7:381–382, 7:424
- assessment 7:52, 7:283–285
 - capacity building 7:53
 - change 7:261–263
 - changing role of competences in 7:281–282
 - citizenship, migration 2:46–52
 - competences in differentiated curriculum and in mass education 7:282–283
 - competences internationally 7:285–286
 - competing views of knowledge in Swedish curriculum discourses 7:262–263
 - contemporary perspectives on 7:329–330
 - contemporary trends in curriculum research centralization, decentralization, and recentralization 7:76–77
 - curriculum, knowledge, and view of realism 7:79–82
 - environment of culture 7:81
 - experience of knowledge 7:80–81
 - hard and soft governance in era of technology 7:78
 - increasing global interest in policy of national school systems 7:74–75
 - knowledge concepts and 7:81–82
 - modes of curriculum governance 7:75–79
 - multilevel governance and 7:78–79
 - NPM to network governance 7:77–78
 - social realism as response to postmodernism 7:80
- context 7:36–37
- from critical discourse studies perspective 7:37–38
- curriculum standardization 7:260–261
- curriculum theorizing on knowledge question 7:263–264
- defining curriculum 7:327–328
- design and innovations
- English curriculum in Philippines 7:217
 - implementation challenges 7:221–222
 - implementing an evidence-based LC5 7:220–221
 - LC5 7:217–219
- design
- actors in 7:202–204
 - approaches 7:200–202
 - artistic approach 7:201
 - CDR 7:202
 - communicative approach 7:201
 - conceptualizing curriculum design 7:198–200
 - instrumental approach 7:201
 - pragmatic approach 7:201–202
 - as process 7:200
- development
- curriculum revisions today 7:151–153
 - historical overview 7:149–151
 - infrastructure and capacity building 7:153
 - terminology 7:149
- dialogic ideal *vs.* transmission of 6:606–607
- and digital citizenship 7:414–423
- English history teachers' experience of 7:366–369
- frameworks and standards, 21st century skills in 7:51–52
- framing
- issue of curriculum structure 7:256–257
 - question of knowledge in 7:262
- and governance in developing countries
- curriculum reform in developing world 7:1–2
 - emergence of transnational curriculum reform discourse 7:2–3
 - social logic of 7:3–5
 - South Africa's curriculum (2005) 7:5
- in higher education 8:353–354
- historical perspectives of 7:328–329
- history teachers' experience of 7:364–365
- ideology
- and curriculum 7:427–429
 - and different views of knowledge in modern schooling 7:258–260
 - knowledge in social efficiency as 7:259
 - knowledge in social reconstruction as curriculum ideology 7:259–260
 - social efficiency as 7:259
 - social reconstruction as 7:259–260
- implementation in 7:52–53
- implications of critical realism for 7:88–89
- and inclusive education 7:440–446
- institutional curriculum 7:258–260
- instructional leadership in decolonizing curriculum 8:352–353
- interventions 10:81–82
- knowledge 4:487
- and curriculum change 7:261–263
 - and literacy
- education as human capital 7:310–311
- expanding literacy 7:308–310
- global trends in literacy curriculum 7:310–312
- literacy curricula in global times 7:311–312
- multiliteracies 7:309–310
- new literacy studies 7:308–309
- literacy for PSTs 7:160–161
- making 7:189–191
- agency and curriculum literacy for PSTs 7:160–161
 - community in 7:169–170
 - and educational change 7:156–158
 - intellectually honest curriculum making 7:177–179
 - in ITE 7:159–160
 - labor in 7:168–169
 - as object of activity 7:164–166
 - rules in 7:167–168
 - teacher agency and 7:192–195
 - and teacher education 7:158–160
 - tools in 7:166–167
- multiplicities as participatory modes of interaction
- curriculum's multiplicities as prompting humane becoming 7:376–377
 - slippery slopes of reconceptualized curriculum 7:371–372
- pathways of work 6:384–385
- policies in France and Quebec 7:11–16
- competence-based curricular reforms in two contrasting education systems 7:11–14
 - recontextualization of common trends into two national contexts 7:14–16
- policy and actors 7:202–203
- position in 7:50–51
- power relations, curriculum, and recognition 7:435
- premises for knowledge divisions in curricula 7:257–258
- programmatic curriculum 7:260
- reconceptualization and emphasis on identity 7:383–385
- feminist theorizing and identities 7:384
 - identity and curriculum 7:386
 - phenomenological theorizing and identity 7:384–385
 - socio-material approaches to identities and 7:386–389
- reform
- changed scientific meaning at turn of 20th century 7:30–31
 - in Europe 7:28–30
 - revolutionary approaches and reinvention of 7:31–33
- in relation to democracy in early childhood education 7:425–429
- and science
- civic driver for science curriculum design 7:319–320
 - common core to sciences 7:320–321
 - cultural driver for science curriculum design 7:318–319

- curriculum progression 7:323
 demarcating science 7:315–316
 developing curriculum models to represent science 7:323
 developmental driver for science curriculum design 7:317–318
 differentiating curriculum 7:322–323
 disciplinary identify 7:320
 economic driver for science curriculum design 7:317
 four frames for understanding purposes of science curriculum 7:317–320
 learning progressions 7:325
 organization of the science curriculum 7:320–325
 science curriculum and educational purpose 7:316–317
 selecting science to teach 7:316
 STEM question 7:321–322
 three phases of constructing 7:315
 standardization 7:260–261
 standardized curriculum 5:22–23
 stratification within secondary schooling 9:80–81
 studies
 east meets west in 7:398
 emergence of 7:382–383
 and globalization & my entry points 7:65–67
 globalization and transnational curriculum 7:57–59
 internationalization of 7:60–62
 textual 7:263
 theoretical distinctions on curriculum governance 7:21–22
 theorizing on knowledge question 7:263–264
 time lag 7:53
 21st century skill in 7:49
 and United Nations' sustainable development goals
 curriculum and nation–state 7:407
 curriculum and SDG 7:408–409
 deliberation and ethical dimension involved 7:409–410
 democratic impasse 7:407–408
 promises and insufficient action 7:407
 recontextualizations, hegemony, and democratic iterations 7:409
 curriculum theory (CT) 7:126
 analytical concepts to understand curriculum change 7:130
 analytical levels of 7:129–130
 contribution of discursive institutionalism to 7:129–130
 curriculum actors as agents of change 7:130
 curriculum studies 7:118–119
 denominational roots of cultural configuration of 7:119–121
 discursive institutionalism 7:127–128
 as discursive practice 7:127–128
 historical social science 7:98–102
 as increasingly complicated discourse 7:129
 national curricula and national epistemologies 7:121–122
 national discourses, international research and 7:122–124
 new institutionalism 7:126–128
 school curriculum 7:102–103
 structures of knowledge 7:96–98
 CXCS. *See* cross-cultural studies (CXCS)
 CyberActivism 2:369
 cycle of Peer Observation (CoPO)
 8:261–262
 restoring professional agency through 8:262–263
 Cyprus, education policymaking in 9:260
- D**
- D-ISC model. *See* discrepancy-induced source comprehension model (D-ISC model)
 D-separation algorithm. *See* directed separation algorithm (D-separation algorithm)
 DAC. *See* development assistance committee (DAC)
 Dakar EFA 1:515–516
 Dakar framework of action 1:533
 Daniella's testimonio 2:378–381
 beyond ESU 2:381
 hope, fear, and action 2:379–380
 undocumented students belong in college 2:378–379
 data 12:201–202, 13:306–308
 awareness of data practices 10:492–493
 collecting specific types of 12:365
 to confirm assumptions and using data to challenge beliefs 4:173
 considerations 14:748
 distributional shape 14:748
 sample size 14:748
 consolidation 12:496
 data-analysis strategies in mixed-methods research 12:376
 decisions related to collecting 12:364–365
 designing tasks for data modeling, analysis, and reporting process 14:179–181
 develop range of data sets 12:455–456
 drive of data and sampling 12:205–207
 evaluation 14:845–848
 extraction and synthesis 6:650
 governance competences 10:598–599
 matrix 14:590–591
 PCA of 14:594–595
 pre-processing 14:591–592
 pre-processing practices 14:239–241
 checking anchor items 14:240
 need to continue discrete distributions of scores 14:240
 observed score equating methods 14:241–243
 presmoothing score distributions 14:240–241
 sample selection 14:240
 preparation for cluster analysis 14:631–632
 processing and analysis 10:420–421
 reconceptualized 12:200–201
 using software for data-based integration 12:503–507
 sources 4:159
 for tracking and flexible grouping 4:173–174
 transformation 12:494–495, 12:756
 visualization principles 14:221–222
 data analysis 11:122–123, 11:166–167, 14:848–854, 4:506
See also meta-analysis
 fixed-effect or common-effect model 14:849–850
 initial graphics 14:849
 and interpretation 12:674–676
 investigating between-studies variation 14:850–851
 limitations 11:168
 notation 14:848–849
 publication bias 14:853–854
 random-effects model 14:851–852
 research question 11:166–167
 data collection 12:747–748, 4:505
 and analysis, specify procedures for 12:767
 combining different types of 12:365–367
 explore differences between quantitative groups or deviating individuals 12:366
 explore expected or unexpected quantitative results using qualitative data collection and analysis 12:366–367
 integrating outcomes of analysis of separate datasets 12:365–367
 provide quantitative overview as part of qualitative research 12:366
 qualitative data to illustrate quantitative findings 12:367
 quantitative data for development in qualitative study 12:367
 quantitative data for sampling for qualitative data collection 12:367
 designs 14:238–239
 mapping potential ways to mix method 12:695–696
 in mixed methods research studies 12:487–489
 different samples, different time points 12:489
 sample, different time points 12:488–489
 sample, one time point 12:488
 online 12:694–696
 qualitative and/or quantitative data online 12:694–695
 and research designs 12:688
 strategies in mixed-methods research 12:376
 from surveys 14:703–704
 databases 10:370
 for bibliometric review 4:547–549
 development of 10:370–371
 datafication 1:673
 datafication in higher education 1:675–677
 deprofessionalization 1:677–678
 in education 1:674–675

- datafication (*continued*)
 emerging challenges 1:677–679
 inequalities 1:678
 privatization 1:678–679
 regulation 1:679
- DBE. *See* design-based education (DBE)
- DBER research artifacts in STEM
 11:322–324
 cross-disciplinary 11:323–324
 integration of research 11:324
 science 11:322–323
 statistics 11:323
- DCIs. *See* disciplinary core ideas (DCIs)
- DCM. *See* diagnostic classification models (DCM)
- De facto* privatization 1:177
- de-schooling higher education systems
 8:49–50
- deaf learners
 closing deaf schools 9:482
 exceptional 9:480–481
 segregated separate school
 9:481–482
- decentralization 1:401–402
 education in South-East Asia 1:421
 mode 7:76–77
- decision considerations
 for accountability uses of through-year
 assessments 13:217–218
 for instructional uses of through-year
 assessments 13:218
- decision making
 practices in Thailand 4:496
 process
 district–school arrangements that shape
 school level 4:238–239
 district–school informal social networks
 4:239
 district–school organizational routines
 4:239
 formal and informal authority
 4:237–238
 jurisdiction 4:239
 organizational routines 4:238
 research in 4:236–237
 school level factors that shape research
 use in 4:237–238
 in schools 4:236
 status 4:238
 reasoning and 6:235–242
- decision-theoretic treatment 14:775–777
 decision rule, loss, and risk 14:775
 hypothesis testing 14:776
 point estimation 14:775
 prediction 14:777
 set estimation 14:776
- decision/use inferences 13:217–218
- declarative knowledge 6:211
- declarative memory 6:220
- declarative sentence 6:470
- decolonial and education 2:32–36
 mobilizing decolonial 2:34–36
 toward decoloniality 2:34
- decolonial resistance 10:810–811
- decolonial turns of southern theory 1:38–39
- decoloniality and impact on globalization
 1:110
- decolonizing/decolonization
 meeting needs of 8:53–54
 place-based education
 Centering Kanaka 'Oiwi perspectives of
 3:229
 psychology 3:172–175
 colonial tool of western psychology and
 eugenics 3:173–174
 cultural appropriation within western
 psychology 3:174–175
 indigenous psychology 3:177–178
 plastic medicine wheel 3:175–177
- decoupling direct international education
 policy 1:255–256
- deep learning 14:686–689
- defiant imaginaries
 engaging politics of identity 5:40–43
 Foucault's challenge 5:43–47
 friendship as form of negative ethics in
 teaching 5:47
 teachers' relationships to curriculum and
 pedagogy 5:41–42
 teachers' relationships with leadership and
 colleagues 5:42
 teachers' understanding of and relationship
 to themselves 5:43
- deficit 6:210
 asset-based *vs.* deficit-based approaches
 10:79–80
 discourses 5:418
 model 5:470
- degree centrality 4:529
- degree of integration 11:413–414
- deletion-based methods 14:807
- deliberate feedback designs 13:43
- Delors Report to futures report 1:517–521
- Delphi, results of 11:351–354
- Delphi method
 guiding theory and philosophy of
 12:608–609
 historical roots and development of 12:608
 issue identification/prioritization 12:612
 types of 12:609
 value 12:609
 of Delphi as part of mixed methods
 research 12:611–612
- Delphi panel, shortcomings of 11:352
- Delphi variants
 characteristic of 12:610
 number of rounds 12:610
 quantitative, qualitative and mixed
 methods Delphi 12:610–611
 sample size 12:610
- democracy 1:581–584, 7:424–425
 curriculum in relation to democracy in early
 childhood education 7:425–429
 through early childhood practice
 7:424–433
 and education 2:11
 contemporary challenges for 2:9
 as neoliberal commodities 2:10
 implications of ABER for 2:318
- democratic racism 5:418
- democratization of knowledge 8:105
- depth of knowledge 11:449
 as evidence of knowledge in use 11:452
- depth ontology 7:86–87
- descriptive and correlational studies
 14:697–700
- descriptive statistics 14:723
 describing multivariate data 14:729–730
 describing univariate categorical data
 14:723–724
 describing univariate numerical data
 14:724–729
 example 14:730–733
 measures of center 14:724–725
 measures of location 14:729
 measures of shape 14:727–729
 measures of variability 14:725–726
- design coaching 11:38–39
- design considerations, narrative inquiry
 12:127–128
- design cycles 11:123–127
 design 11:123–126
 first cycle teaching experiments
 11:123–124
 iterative cycles of design, testing, and
 revising 11:123
 outcomes and reflections from first design
 cycle 11:124
 outcomes and reflections from second
 design cycle 11:125–126
 outcomes and reflections toward fourth
 cycle 11:127
 second cycle pilot studies 11:124–126
 testing and analysis 11:124–127
 third cycle field experiments 11:126–127
- design decisions 12:364
- design research
 agenda
 collaboration *vs.* scale in design research
 10:478
 considerations and cautions in 10:478
 imagination in design research 10:478
 publishing in academy 10:478
 characteristics and principles of
 10:474–477
 literacy teacher preparation 10:475–477
 research in practice 10:475
 research into practice 10:474–475
- design-based education (DBE) 11:198–199
 with computational thinking and STEM+C
 11:201–202
 designing DBE for STEM learning 11:204
 with digital literacy and new media creation
 11:202–204
 with engineering design and maker
 education 11:200–201
 foundation and rationale of 11:198–200
 higher education settings 11:201–204
 interdisciplinarity of the current
 applications 11:204
 K-12 settings 11:200–201
 out-of-school settings 11:201
 rationale of 11:199–200
 in STEM, current applications of
 11:200–204
 theoretical foundation of 11:198–199
- design-based pedagogy 11:293–296

- epistemological shifts for 11:292–293
 implementing science lessons for ELLs and students with disabilities 11:295–296
 integrating robotics in ways to enhance professional learning 11:296
 lesson planning as design partners 11:294–295
 designing quality criteria, selective aspects of 12:519–520
 designing school
 building curriculum and future models of schooling 7:232–233
 classroom curricula 7:226–228
 material transfers of curricular designs 7:230
 materializing curriculum and educational discourses 7:224–226
 relationship between school architecture and educational practice 7:228–230
 DETECT, dimensionality assessment software as 14:293
 development assistance committee (DAC) 13:267–268
 developmental science approaches
 bilingual children learning to write in Chinese and English 10:860–863
 computational *vs.* observation models of literacy development 10:856–857
 development is complex and nonlinear 10:850
 dynamic, historical modeling of early literacy 10:851–853
 importance of showing variation in modeling literacy behavior 10:856–859
 learning to write in Spanish 10:853–856
 models of multidimensionality 10:857
 principles of development 10:849–850
 prospective, developmental models of learning 10:857–859
 retrospective lens on interpreting developmental discontinuities in early reading and writing 10:850–851
 deviance information criterion (DIC) 14:286–287
 DF. *See* differential functioning (DF)
 DI. *See* differentiated instruction (DI)
 diagnostic classification modeling
 general latent class model 14:111–112
 measurement domain 14:109–110
 measurement model and developing items 14:110–112
 specifying measurement model 14:111–112
 specifying structural model 14:111
 diagnostic classification models (DCM) 14:109
 compensatory models 14:115–116
 constraining LCDM to yield common 14:114–116
 extended DCM measurement models 14:116–117
 hierarchical 14:116
 longitudinal 14:116–117
 measurement model parameterization 14:112–117
 measuring misconceptions 14:117
 noncompensatory models 14:114–115
 polytomous 14:116
 dialectical pluralism 12:386
 dialects communications 10:747–748
 dialog systems of CBAs 14:504–505
 dialogic education 6:607–608
 Alexander's dialogic teaching framework 6:604
 Bakhtin's dialogism 6:602
 Buber's philosophy of dialogue 6:602
 contemporary approaches to 6:603–606
 epistemological dialogic education 6:601
 Freire's pedagogy of oppressed 6:602
 historical origins of 6:601–602
 Mortimer and Scott's communicative approach 6:603–604
 Nystrand's dialogically organized instruction 6:603
 ontological dialogic education 6:601
 questions of definition 6:600–601
 Socrates' contrast between "living" and "dead" meaning 6:602
 tensions, dilemmas and future directions 6:606–609
 dialogic educational system 10:738–739
 dialogic teaching 6:607–608
 diaspora 10:811–812
 curriculum theorizing 7:399–401
 as human capital and state possession 1:130
 as marginalized and social construct 1:130–131
 DIC. *See* deviance information criterion (DIC)
 dichotomies 6:658
 dichotomous response models 14:74–75
 didactics 7:250–252
 Didaktik 7:334–335
 denominational roots of cultural configuration of 7:119–121
 DIF. *See* differential item functioning (DIF)
 difference 2:84–85
 differential bundle functioning 14:251–252
 differential functioning (DF) 14:249
 analyses important part of test development and validation 14:257
 practical considerations 14:257
 uniform and non-uniform DIF 14:250–251
 differential item functioning (DIF) 14:249–251
 analyses used operationally 14:256–257
 latent DIF analysis 14:256
 methods used to detect DIF 14:252–256
 uniform and non-uniform DIF 14:250–251
 differential person functioning (DPF) 14:252
 differential test functioning 14:252
 differentiated instruction (DI) 5:712–713
 differentiated learning 6:263
 digital and multimodal learning 10:487–496
 digital animation, blending drama and 10:500
 digital assessment tasks, characteristics of 14:172–173
 digital biliteracy, research on 10:204–205
 digital citizenship
 as citizenship rights 7:417
 for civic, political and economic participation 7:417–418
 development requires social emotional learning 7:415
 as digital competence 7:415
 evolving conceptualizations of citizenship in context 7:416–417
 future of digital citizenship curricula 7:419–421
 internet activism and place-based social action 7:417–418
 need for continuity between citizenship and digital citizenship education 7:420–421
 rise of DIY citizens and borderless netizens 7:417
 in school curriculum 7:415
 sociotechnical regime challenges in context of educational transitions 7:419–420
 tensions between empowering/empowered digital citizens and 7:418–419
 digital communication 10:439–440
 digital divide
 different types of 1:158–159
 digital cultural capital 1:159–160
 role of education in bridging 1:162
 theories of 1:158–160
 digital equity 10:492
 digital formats, writing in 10:572–582
 digital games 10:585–586
 construction of paratexts 10:586
 digital literacies 10:587
 gaming
 fosters complex thinking and increases motivation 10:586
 increase literacy performance 10:586
 promotes development of "soft-skills" 10:586
 as multimodal systems 10:584–585
 perspectives on gaming 10:585–586
 school-based literacies and 10:585–586
 digital geographies of places 10:507
 digital identity, research centered on 10:494
 digital landscape 10:551–552
 digital literacy 2:306, 5:265–266
 and computational thinking as core century skills relevant for teachers 5:265–266
 digital games 10:587
 multimodal learning 10:572–582
 technologies and approaches 3:66
 digital media 10:500–501, 2:369
 as major contemporary contextual factor 6:262–264
 sensorium in digital media practices 10:482–483

- digital methods in qualitative research in
online spaces 12:293
- digital multimodal literacy 10:321–322
- digital ocean
assessment in 14:471–472
process data 14:472
- digital performance assessments 8:314–315
in higher education 8:313, 8:317–319
advantages and challenges of 8:319–321
- digital storytelling 10:511–512, 12:183
- digital teacher learning space, mentoring in
5:735–736
conceptual directions 5:735
curricular directions 5:735–736
- digital teaching space, spheres of knowledge
for mentoring in 5:734–735
- digital technology 4:388, 5:640–641
adapting curriculum and instruction 5:641
children's uses of digital technology in
homes 10:171–173
context mattered 5:642
systems 6:263
unique challenges of hybrid models
5:641–642
- digital texts
e-book and 10:558–559
skills, strategies, dispositions and social
practices for reading 10:553–555
- digital world making, collective engagement
in 10:341
- digitally mediated literature for children
10:623–624
- dimensionality assessment software as
DETECT and DIMTEST 14:293
- dimensionality-reduction techniques
14:590–599
- DIMTEST, dimensionality assessment
software as 14:293
- direct protest/action 2:368–369
- directed separation algorithm (D-separation
algorithm) 14:121
- Dirichlet distribution 14:715
- disability 9:229, 9:232, 9:436
accountability at national and international
level and inclusion of students with
13:176–178
attitudes about 9:328–329
barriers to higher education for students
with 9:371–372
controversy about nature of 9:272–273
and employment
capability approach 9:455–456
connecting 9:453–455
interpreting disability employment data
9:454–455
responding to disability employment
gaps 9:453–454
rethinking employment 9:456–459
establishment of special education
9:440–443
facilitators to higher education for students
with 9:372–373
facts and objectification of 9:192–193
in higher education, students with 9:369
inclusion and nature of 9:275
inclusion and severity of 9:275–276
- institutions 9:438–440
and intersectionality 9:234
- labeling disabilities 9:273
- life with conflicting frames 9:195–196
- modeling story of 9:198–199
- as ordinary and extraordinary diversity
9:272
- orientation of frames 9:191–192
- overrepresentation and social and
intersectional disablism model of
9:501–503
- and poverty 9:464–465
measuring 9:462–464
politics of poverty 9:465–466
in UK, lived experience of
9:466–467
- as predicament 9:272
- related extra costs in education 9:55
- social model and 9:193–195
- special education as response to 9:273
- studies as framer of frames 9:196–198
- studies in posthuman age 9:201
developing posthuman narrative analysis
9:203–204
entanglements with environments 9:206
entanglements with machines
9:205–206
mothers of disabled children 9:202
project 9:202–203
support services 9:370–371
- youth activism
conceptualizing inclusive education
through 9:321–323
within inclusive education, embedding
9:323–325
- disability studies in education (DSE) 9:234,
9:238
- contemplating belonging, cultivating
9:241–245
- differentiating DSE and critical disability
studies 9:244
- emergence of DisCrit and intersectional
approaches 9:243–244
- framework 9:353
encourages examination of social and
cultural context in which children
learn 9:355
partners in pedagogy 9:356–359
problems of traditional individual and
individualizing approaches to
assessment 9:355–356
rethinking research and teacher
professional learning about
assessment and pedagogy, 1 for
9:354–355
supports focus in research on problems
in teaching as professional
dilemmas 9:354–355
- globalization of critical disability studies
9:244–245
- negation of racial disparities by field of
special education 9:243
- neurodiversity movement 9:242–243
- ongoing crippling of curriculum with
9:241–242
- of policies, placements, and dispositions
9:238–239
- popularity of memoir 9:242
- sense of belonging 9:239–241
- disabled children
adulteration of 9:314–315
barriers to play for 9:310–313
emancipating play for 9:315–317
right to play 9:308–309
- disabled students' experiences of schooling
disability 9:229
essence of students' experiences 9:232
friends matter 9:231
inclusive education 9:229–230
pedagogical practices 9:230
pedagogies and curriculum 9:234–235
practicum matters 9:235
school cultures 9:230
"special" education practices 9:231
support 9:231
teacher education 9:233–235
teachers' values, beliefs and assumptions
9:230
theoretical framing of students' experiences
9:232–233
understanding students' experiences of
school 9:229–232
- disciplinary core ideas (DCIs) 11:444–445
- disciplinary knowledge 6:211
- disciplinary literacy 10:449–451, 10:468,
11:43
- development of conceptual framework for
6:306–307
- ongoing research in 10:451–452
- in practice 10:450
- research on expertise 6:306–307
- disciplinary reductionism 10:884–885
- discipline-specific formative assessment
13:6–7
- discourse 7:67–69
analysis 12:29, 12:141–142
peculiarities of discourse analysis as
discipline of study 12:144–146
as qualitative research approach
12:142–144
in terms of QUAL research 12:146–148
- analytics in automated analysis of student
writing 6:77–79
- competence 6:366
- skills 6:470–471
- discrepancy-induced source comprehension
model (D-ISC model) 6:355
- discrete item response theory model
14:123–126
- discrete probability distributions
See also continuous probability
distributions
Bernoulli distribution 14:718
binomial distribution 14:719
geometric distribution 14:719
hypergeometric distribution
14:719–720
multinomial distribution 14:721
negative binomial distribution 14:720
Poisson distribution 14:720–721
- discrimination 6:18

- discriminative stimulus 6:13, 6:18
- DisCrit and intersectional approaches,
emergence of 9:243–244
- dispersion parameter () 14:584
- estimate of 14:585–586
- displaying information
- evolution of recording numbers
14:819–820
 - first soaring national debt plot 14:824–826
 - gedanken collaboration 14:837
 - graphic poetry of Charles Joseph Minard
14:831–832
 - graphs
 - as poetry 14:828–837
 - in narrative argument 14:832–833
 - great migration 14:835–837
 - hybrid displays 14:826–828
 - Kovno Ghetto 14:830–831
 - using numbers 14:820–826
 - plain graphical poem 14:830–831
 - practical type I graphic 14:823–824
 - ship from forest 14:822
 - W.E.B. Du Bois 14:833–835
 - woolly mammoth look like 14:822–823
- dispositional knowledge 6:382
- dispossession 2:39–43
- disproportionality 9:498–500
- distal self report 13:56
- distance learning 6:262
- distributed leadership 4:27–28, 4:48–49,
4:65, 4:92, 4:204
- allocative distribution 4:49
 - blindness to exclusion 4:106
 - bureaucracy 4:103–104
 - colonization by research theories 4:105
 - commentary 4:98
 - contingent leadership 4:49–50
 - critiques of 4:97, 4:101–103
 - different contexts 4:96
 - disciplines 4:98
 - to empirical test 4:92–93
 - exploring contemporary evidence 4:94–95
 - exploring relationships 4:95–96
 - instrument development 4:96–97
 - insufficient interrogation of research
methodology 4:105
 - leadership theory 4:106
 - poor historical awareness 4:104
 - power and 4:102–103
 - and student outcomes 4:93–94
 - sustaining weakness of field 4:104–106
 - teacher leadership 4:49
 - theoretical evangelism 4:104–105
 - in turnaround contexts 4:165
- District Primary Education Program (DPEP)
1:531
- district–school arrangements 4:238–239
- district–school informal social networks
4:239
- district–school organizational routines
4:239
- divergent thinking 6:237–238
- diverse learners
- recruiting, preparing, and retaining teachers
of color to teach 5:359–360
 - teacher education for 5:356–367
- diversity 2:84–85, 6:261
- difference 2:4
- epistemologically 2:5–6
 - ontologically 2:4–5
- discourse beyond Anglo-Saxon contexts
2:290–291
- discourse in Anglo-Saxon contexts 2:290
- in early childhood education 5:600
- supporting learning of 5:600–601
 - workforce 5:600
- implications of ABER for 2:318
- and inclusivity 2:3–4
- need to transform discourse and curriculum
around 2:295
- and race relations 2:1–2
- as social construction 2:289, 2:294–295
- toward socially just world 2:6
- divestment at University of Cape Town
2:401
- DM model. *See* double monotonicity model
(DM model)
- DMF. *See* document models framework
(DMF)
- document analysis
- approaches to analyzing documents
12:285–286
 - discourse analysis 12:287–288
 - ethnography 12:288–289
 - ethnomethodological ethnography
12:289–290
 - meta-synthesis 12:286–287
- document models framework (DMF) 6:354
- domain analysis of ECD 14:165
- domain modeling layer articulates 14:165
- domain module 6:638
- domain-specific skills 6:128
- double art of teaching 5:650–652
- double monotonicity model (DM model)
14:142–143
- DPEP. *See* District Primary Education
Program (DPEP)
- DPE. *See* differential person functioning
(DPE)
- draft 14:72–85
- dramatic inquiry pedagogy 10:347–350
- different approaches to 10:350
 - dramatic conventions 10:349
 - dramatic dialogic inquiry 10:349
 - dramatic imagination 10:348–349
 - four approaches 10:350
 - framing 10:350
 - transforming classroom reality 10:349
- DSE. *See* disability studies in education
(DSE)
- dual process models 6:238–239
- dual strategy model 6:239–241
- DuBois, W.E.B. 6:674–675
- Duende, languaging with
- inventive anti-colonial indigenous social
media 10:803–806
 - Kisizi 10:799–800
- Duignan, leadership frameworks and
taxonomies 4:406
- duoethnography 12:51
- dyad level 4:527–529
- strength of ties 4:528–529
 - typology of ties 4:528
- dynamic system approaches to language
acquisition 10:14–17
- critical reflection 10:23–24
 - findings from studies on first language
acquisition 10:17–20
 - findings from studies on second language
acquisition 10:21–22
 - implications for research methodology
10:16–17
 - layers of organization and time scales 10:16
 - process causality 10:16
 - self-organization and emergence 10:15–16
 - state spaces and trajectories 10:14–15
 - variability intrinsic feature of complex
system's trajectory 10:16
- dysconscious racism and lack of critical
understandings of race 5:417
- dyslexia
- challenges related to assessment and
categorization 10:410
 - competing theories as to cause of 10:411
 - difficulties with phonological processing
10:411
 - early identification 10:411–412
 - etymological history of term 10:408–409
 - perceptual issues 10:411
 - research in neuroscience and 10:411
 - varying definitions of 10:409–410
- ## E
- e-books
- and digital texts 10:558–559
 - bi-directional nature of reading and
response 10:558–559
 - classroom based research 10:557
 - constructing meaning with and from text in
digital world 10:552
 - differences of moving from print to screen
10:552–553
 - digital landscape 10:551–552
 - internet inquiry and research 10:555–557
 - skills, strategies, dispositions and social
practices for reading both online
and digital texts 10:553–555
- e-learning ecologies 6:262
- e-texts
- bi-directional nature of reading and
response 10:558–559
 - classroom based research 10:557
 - constructing meaning with and from text in
digital world 10:552
 - differences of moving from print to screen
10:552–553
 - digital landscape 10:551–552
 - e-book and digital texts 10:558–559
 - internet inquiry and research 10:555–557
 - skills, strategies, dispositions and social
practices for reading both online
and digital texts 10:553–555
- Early Career Academics (ECAs) 8:513
- Early career researchers (ECRs) 8:514

- early career teachers
 aims of induction programs 5:137
 currently supported 5:137–139
 employment status 5:135
 induction phase 5:135–137
 and work 5:135–136
 work be reimagined 5:139–140
- early career teaching and resilience
 agency 5:156–157
 confidence 5:155
 culture 5:157–158
 purposefulness 5:154–155
 resilience 5:153–154
 social support 5:155–156
- early child development 1:366–368
 key developments in 6:477–485
 reconceptualizing family literacies in
 10:192–193
- early childhood care and education (ECCE)
 1:534–535
- early childhood education (ECE) 7:173,
 9:449
 children's interests as inquiry and identity
 development 7:176–177
 critiquing entrenched coalition 7:173–174
 curriculum in relation to democracy in
 7:425–429
 diverse and globally mobile 5:600–601
 diversity and global mobility in 5:600
 "intellectually honest" curriculum making
 7:177–179
 key concepts are informing 6:479–483
 children's agency and motivation 6:481
 cultural tools and symbol systems
 6:480–481
 implications of socio-cultural theories for
 early childhood education
 6:483–484
 learning leads development 6:479
 play as social practice 6:481–483
 professional development 5:596–597
 research literature 5:597–598
 relationship between professional
 development and program quality
 5:598–600
 supporting learning of diverse and globally
 mobile 5:600–601
 theoretical framings of children's interests
 7:174–176
- early childhood education and care (ECEC)
 5:93
 homogenized ECEC workforce and policy
 enactment 5:95–96
- early education 1:368–369
- early language development
 individual differences in early language
 development 6:471–472
 major milestones in 6:468–471
 social context and language development
 6:472–473
- early life experiences, HE 8:411
- East Asia through globalization, university in
 8:25–27
- East Asian languages, critical literacies in
 10:112–122
- ECAs. *See* Early Career Academics (ECAs)
- ECCE. *See* early childhood care and
 education (ECCE)
- ECD. *See* evidence-centered design (ECD)
- ECE. *See* early childhood education (ECE)
- ECEC. *See* early childhood education and
 care (ECEC)
- eco-education currents, systemic, humanist/
 mesological, bioregional,
 ethnographic, and 2:214–215
- ecocritical curriculum and pedagogies
 2:207–209
- ecocritical framework 2:203–205
- ecocriticism 10:611–614
- ecofeminist of southern theory 1:38–39
- ecological systems theory 6:532–533
 to conceptualize survey project 12:667–668
- econometric modeling approach
 14:391–392
- economic hegemony
 colonialism's contributions to economic
 hegemony 1:682–683
 goal of 1:683–685
- economic household decision in education
 9:52–54
- economic mobility, HBCUs and 8:426
- ecopedagogical literacy 1:604–610
 ecopedagogically reading citizenships
 1:607–608
 ecopedagogically reading d/Development
 and sustainability 1:605–607
 ecopedagogically reading epistemologies
 and ecolinguistics 1:608–609
 ecopedagogically reading globalizations
 1:607
 ecopedagogically reading postdigitalism
 and post-truthism 1:609–610
- ecopedagogy 1:601
 essence 1:602–604
 uniqueness as environmental pedagogy
 1:604
- EcoSchools programs 2:213
- ecosystem, active agents in 10:681
- ECRs. *See* Early career researchers (ECRs)
- EDID. *See* equity, diversity, inclusion and
 decolonization (EDID)
- Educación Propia 10:738–739
- educand-world, changing relation between
 7:250–252
- educated communities 2:246
- education 1:48–50, 1:505–506, 1:584,
 6:258
 added value and challenges of MM design
 12:631–633
 Amadeo Giorgi 12:135–136
 analysis of topics and themes 9:490–493
 Andres Manuel López Obrador proposals
 2:80
 arguments to implement interdisciplinarity
 in school education 7:272–274
 assessment to cope with low quality of
 1:405–406
 background and relevance of TePP project
 12:629
 behavioral psychology and 6:19
- bibliometric analyses 4:549–553
 and brain aging 6:514–515
 case studies 12:24–26, 2:79–80
 central line of MM design in TePP project
 12:629–631
 challenges in conducting large-scale MM
 studies in 12:627–628
 in changing global arena 1:480–487
 co-citation analysis 4:551–553
 and cognitive aging 6:515–518
 as complex system 9:28–29
 context in Thailand 4:491–492
 covid-19 and right to 1:407–409
 datafication in 1:674–675
 democracy and 2:8–12
 in developing countries
 broad overview 1:209–210
 Chile 1:208
 development decades five and six 1:211
 fourth decade 1:210–211
 Ghana 1:208–209
 knowledge society 1:212–213
 Mexico 1:208
 partnerships 1:211–212
 poverty reduction strategy papers 1:212
 results-based lending 1:213–214
 third development decade 1:207
 two development decades 1:206–207
 Zambia 1:209
- development, South Asia 1:427–428
- diplomacy 8:205
- and disability, developments in special
 needs 9:429
- disability and rate of returns on investment
 in 9:51
- disability biases and change in economic
 household decision in 9:52–54
- disability related extra costs in 9:55
- disability studies in
 intersectional case-study 9:252–253
 intersectionality 9:249–252
 leveraging intersectionality 9:254–256
 through disability youth activism
 9:321–323
- disciplined study of cases in 12:85–86
- discourse analysis 12:29
- disenfranchised individuals and groups in
 Hong Kong and Mainland China,
 research on, 0 of 7:398
- diversity, democracy, and social justice in
 2:313–319
- and economic development 1:53
- embedding disability youth activism within
 inclusive education 9:323–325
- emergence of private corporate and
 philanthropic actors in 1:392–394
- emergence of WB as prominent actor in
 education for economic
 development 1:480–482
- English as medium of instruction and
 research in
 2008 White paper on Norwegian
 language policy 1:376–377
 academic publishing 1:376
 example from Tanzania 1:377

- higher rewards for academic publications in English 1:376
- Norwegian researchers defending our language as language of research 1:377
- from one colonial language to other in Rwanda 1:375
- prestige planning 1:377–378
- private primary schools 1:373
- rhetoric of English and development 1:373–374
- situation in India 1:374–375
- situation in Nordic countries 1:375
- strengthening of government primary school 1:373
- ethnography 12:28–29
- ethnonationalism 2:77–78
- external aid for 1:443
- financing of 1:421
- first nations & indigenous identities in 9:151–158
- formulating and implementing workable language of 3:66
- fundamental approach 1:513–514
- Gabriel Boric's proposals 2:80–81
- gender equality and measurement 2:69–71
- globalism 2:78
- globalization and public funding of 1:441–443
- globalization and shifting geopolitics of education role in bridging digital divide 1:162
- glance at facts 1:157–158
- highlights from recent empirical research 1:160–162
- inequality of digital access 1:157–158
- inequality of digital use and online skills 1:158
- theories of digital divide 1:158–160
- globalization as spaces and flows 1:2–4
- under globalization
- IB and shifting geopolitics of 1:347–352
- grounded theory 12:27
- research in 12:111–119
- studies in 12:116
- history
- curriculum studies 7:118–119
- denominational roots of cultural configuration of 7:119–121
- national curricula and national epistemologies 7:121–122
- national discourses, international research and 7:122–124
- as human capital 7:310–311
- and human development 1:54–55
- and human rights 1:56–57
- humanizing language and literacy in 10:346–357
- Husserl, Heidegger & Merleau-Ponty 12:131–133
- identifying document database for bibliometric review 4:547–549
- and identity 9:486
- illustrative applications of clustering in 14:632–634
- implementation of interdisciplinary approach in 7:274–277
- implications of globalization for 1:7–8
- inclusive approach to 4:402
- increasing competitive nature of 7:292
- insistence on disability studies in 9:509–510
- integrating art and artistic practices into mixed methods studies in 12:705–707
- centering artistic practices in a mixed methods study 12:707
- embedding art within other methodological approaches 12:705–707
- interdisciplinary conceptions 7:268–272
- and practices in Quebec education 7:270–272
- and international development 2:65–68
- intervention strategies to overcome barriers 9:54
- for just, utopistic, transition 1:25–26
- keyword co-occurrence analysis 4:553
- knowledge 2:75–77
- language of education and literacy issues, English 3:65
- leadership for 4:30–31
- learn from populism 1:653–655
- for liberation 7:358
- living schools and purpose of 7:352–354
- logics 7:268–270
- machine learning in 6:86–89
- market 1:64–67
- Max van Manen 12:136–138
- methodological challenges in evaluating cost benefits of 9:55–57
- methodology 9:487
- mixed methods
- case study in 12:546–547
- and multimethod research in 12:361–371
- within multiple geographies of Asia-Pacific 1:414–417
- civic and social education 1:415–416
- international assessments and education divide 1:414–415
- myths about special 9:279–281
- narrative inquiry 12:27–28
- NDIS and 9:447–448
- negation of racial disparities by field of special 9:243
- neo-colonialism and 1:686
- neoliberalization of 1:63–64
- new pathways in
- implications on framework conditions 8:61–62
- learning pathways 8:58–61
- for new world order 1:488–489
- in Nigeria
- conceptualizing englishization 3:61–66
- educational policies and English in Nigeria 3:63–64
- PAR 12:29–30
- as part of colonialist imperialist project 1:685–686
- and participatory democracy
- interrupting democratic ideal 2:228–229
- schooling and democratic citizenship 2:230–231
- schooling as affective violence 2:232–233
- whiteness against participatory democracy 2:231–232
- pedagogical phenomenology 12:139
- phenomenology 12:26–27
- and methodology 12:134–138
- philosophy to educational research 12:133–134
- planning, technology and language choice in 10:818–821
- design solution 10:820
- end-users 10:820
- framework for language, technology, and education 10:819–820
- language context 10:819
- purposes 10:820
- technology 10:820
- policy transfer in
- global policy spaces 2:155–157
- post-world war II 2:153–155
- travellers and reformers 2:151–153
- polycymaking in global education system 1:274–275
- poverty and additional determinants of schooling 9:54
- power 2:74–75
- privatization of 1:436–438
- progressive education 7:351–352
- qualitative research in 12:14–17
- quality, inequality and equality in 2:68–69
- radical critiques of education and development 1:57
- rationale for bibliometric reviews 4:546–547
- reflections on an exemplary large-scale MM study 12:628–633
- reform
- in Australia 5:31–34
- in India 5:722–723
- politics of 1:201–202
- relevance of grit for 6:191–192
- relevance of resilience for 6:190–191
- research
- oral traditions in 12:180–183
- significance of oral traditions and as affirming research methodologies in 12:180
- as response to disability 9:273
- rethinking critical theory/study of art of governing 7:106–107
- attachments of desires and potentialities 7:109–113
- critical studies 7:113–114
- fabricating kinds of people 7:107–108
- locating homeless and location-less reason 7:111–113

- education (*continued*)
- political 7:106–109
 - politics and political of schooling 7:108–109
 - temporal ordering and desires of differences 7:110–111
 - role in construction of identity in diverse world 2:88–90
 - role in national development 1:390–391
 - science mapping 4:550
 - SDG 4 1:539–546
 - in small island states
 - ecology of educational policies and priorities in small states 1:461–463
 - historiography of small states research 1:459–461
 - toward definition of small states 1:457–459
 - social justice 2:79, 4:39
 - social realist school in sociology of 7:90–92
 - in South-East Asia 1:421
 - STEAM-inspired vision for 11:14–16
 - as struggle against anti-rationalism 1:652–653
 - as struggle against economic deprivation 1:653
 - struggle against populism 1:652–653
 - supports children's overall development 4:341–342
 - and sustainable development 1:55, 7:352
 - systems
 - competence-based curricular reforms in two contrasting 7:11–14
 - shifting focus of financing within 9:37–40
 - theorizing and concepts 4:29–31
 - implications and conclusions 4:30–31
 - theorizing global spaces and relations 1:4–7
 - in times 2:17–18
 - transnationalism as activator of rights in 1:49–50
 - twenty-first century observations on structural adjustment in 1:214
 - unraveling assessment as phenomenon in 13:103–104
 - visual media in education and assessment 14:176
 - well-being in 6:721–722
- Education for Liberation Day 2:361
- Education International (EI) 1:575–576
- history of 1:575–576
 - and multi-lateral institutions 1:576–577
- education policy 13:259–260
- asking spatial questions of 1:33–34
 - discourses
 - analysis 9:262–266
 - critical policy discourse analysis 9:261
 - discursive absences and policy silences 9:265–266
 - education policymaking in Cyprus 9:260
 - minority rights discourses 9:264–265
 - PDA 9:261–262
 - geospatial research development in 1:32–33
 - Marxist political economy and spatial dimensions of 1:30–31
 - (post)spatial turn 1:31
 - power and topological space 1:31–32
 - reform 13:120
 - spatial theories, methods and 1:29–30
 - SSC and 1:322–323
 - in 21st century
 - global context 5:605–606
 - local context 5:607–608
 - regional context 5:606–607
 - education-choice, commodification of 1:65
 - educational accountability 13:181–191
 - assessment across and within international contexts 13:166–169
 - NCLB 13:166
 - SBR 13:165–166
 - use and misuse of assessment in United States 13:164–166
 - passage of ESEA and emergence of NAEP 13:165
 - program evaluation and ability tracking 13:164–165
 - educational and life transitions theory (ELT theory) 6:530, 6:534
 - educational assessments 13:136–146
 - analyzing purposes 13:137–138
 - classifying purposes 13:139–140
 - differentiating between purposes 13:138–139
 - differentiating within purposes 13:139
 - listing purposes 13:136–137
 - multiple perspectives on purposes 13:143–144
 - purpose 13:136
 - purposes and evaluation 13:141–143
 - purposes and validity 13:140–141
 - purposes of 13:196–197
 - research approaches to 13:104–105
 - scholarly debates in 13:104
 - taking multiple purposes into account 13:144–145
 - of 21st century skills 13:252
 - ADAPT21 13:252
 - assessment in 21st century—digitalization 13:248–253
 - assessment of 21st century skills 13:250–251
 - assessment of computer and information literacy 13:251
 - defining 21st century skills 13:245–248
 - evolving frameworks of 21st century skills 13:246–247
 - international assessment initiatives 13:250–251
 - modeling collaborative problem-solving data 13:252–253
 - national 21st century skills
 - assessment—projects 13:251–253
 - from policy to practice 13:247–248
 - TEACH21st 13:252
 - transformative aspects of assessment in 21st century 13:253–254
 - unraveling assessment as phenomenon in education 13:103–104
 - educational change
 - consequences and tests as levers of 14:41–42
 - curriculum making and 7:156–158
 - evaluating consequences of tests used as levers of 14:42
 - educational decisions, types of 13:196–197
 - educational development, conceptions of 8:275
 - educational digital competence 5:266–268
 - educational digital inclusion in Global South, labyrinths of 1:284–285
 - educational diplomacy, power of 1:643
 - educational environment 11:67
 - factors affecting creativity in 11:66–69
 - educational equity, transformative teacher education for 5:361–364
 - educational fads 9:178–180
 - crumbling orthodoxy 9:179
 - cures, therapy and hope 9:179
 - performance enhancing trends 9:179
 - technical trends 9:179
 - trends of near future 9:180
 - educational failure, populism consequence of 1:651–652
 - educational goals 8:454–455
 - educational implications, social cognitive theory 6:32
 - goals and self-evaluations of progress 6:32
 - self-efficacy for learning 6:32
 - self-regulation 6:32
 - educational inclusion 9:12–13
 - probable success of 9:276–281
 - educational institutions, proper functions of 4:1–3
 - educational interventions to reduce gender/racial gaps 11:339–341
 - encouraging diversity and broadening pool of students taking STEM coursework 11:341
 - spatial training 11:339–340
 - values affirmation 11:340–341
 - educational leaders 4:372
 - cohorts 4:375
 - curriculum 4:375
 - defining educational leadership 4:372–373
 - field-based experiences 4:377
 - leadership development 4:374–378
 - in US 4:370–371
 - partnerships 4:377–378
 - program faculty 4:378
 - researching quality preparation 4:373–374
 - student selection 4:374–375
 - teaching 4:375–376
 - educational leadership
 - in China 4:478–479
 - deriving distinctive work of 4:4–8
 - distinctiveness of the proper functions 4:3–4
 - framed from distributed perspective 4:83
 - framed from multilevel distributed framework 4:83–88
 - implications for
 - classroom teaching 4:5
 - leadership of learning and teaching 4:6–8
 - leading and participating in teacher professional learning 4:7–8

- leading by ensuring quality teaching for deeper learning 4:7
 leading by ensuring safe and orderly learning environment 4:8
 leading by resourcing strategically 4:6–7
 leading by setting goals and expectations for deeper learning 4:6
 proper functions of educational institutions 4:1–3
 pruning theories 4:40–41
 refining definitions of theories 4:39–40
 shapes educational leadership in China 4:479–480
 sociological, psychological, and socio-psychological theories 4:39
 students 4:4–5
 theory 4:34–36
 about leadership styles 4:36–38
 in organizational context 4:38
 about social justice in education 4:39
 educational measurement 14:96
 applications to 14:690–693
 automatic feature extraction 14:690
 development of current practice in 14:7–11
 foundations of 14:1–7
 future of 14:11–12
 London school and development of classical test theory 14:2–5
 educational movement, pedagogical progressivism as 5:683–684
 educational neuroscience 6:63–64
 educational values in perspective 6:62–63
 involving learners in research process 6:63–64
 key research domains in 6:59–61
 brain health fundamentals 6:59
 executive functions 6:60
 exercise 6:59
 individual differences in educational outcomes 6:61
 nutrition 6:59
 protection from stress 6:59
 SES 6:61
 sleep 6:59
 social and emotional learning 6:59–60
 research to education, translating 6:62–63
 training researchers for interdisciplinarity 6:63
 educational outcomes
 educational implications of gender gaps in 6:406
 educational pathways
 from education into labor market 8:441
 between educational sectors 8:441
 through lifespan 8:441–443
 typologies of 8:441–443
 educational performance 1:119–121
 educational policies in Nigeria 3:63–64
 educational policy and practice, issues of 1:579–581
 educational governance 1:580
 future of teaching profession 1:580
 shortage of teachers 1:580
 uses of technology 1:579–580
 educational policymaking, broad principles to inform alternative approach to 5:25–26
 educational practice
 learnings from populism for 1:654–655
 potential for 8:320
 educational processes promoting adaptability 8:458–459
 educational project, southern theory as 1:37–44
 educational quality 5:102
 in-service teacher professional development 5:102
 qualifications and pre-service training 5:102
 teacher deployment 5:102
 educational research 4:511–512, 6:26, 6:29, 6:31
 analyzing and synthesizing data 12:662
 critically appraising primary-level studies 12:661
 CRT and qualitative methods in 12:73–74
 data collection from surveys 14:703–704
 decisions for survey designs 14:709
 extending psycholinguistic perspectives in 10:841
 extracting data 12:662
 future of 10:845–846
 learnings from populism for 1:655
 mixed methods research synthesis in 12:660–662
 MMSNA in 12:687–690
 current debates 12:688–690
 data collection and research designs 12:688–689
 human capital limitations 12:689
 methods integrated in 12:688
 practical impact 12:690
 prevalence of 12:688
 nonparametric modeling frameworks for 14:139–140
 nonprobability samples 14:708–709
 probability and nonprobability samples 14:704–705
 probability sampling designs 14:705–707
 qualitative content analysis for 12:315
 selecting sampling strategy, search strategy, and eligibility criteria 12:661
 survey design and nonresponse 14:708
 world-systems analysis in 1:23–25
 writing protocol and selecting design 12:661
 writing report 12:662
 educational simulations 6:685
 educational standards, literacies social values reflected in 10:491
 educational surveys
 adaptive testing 14:355
 collection of contextual information 14:360
 content, cognition and context information 14:352–354
 domain sampling 14:354
 emphasis on trend reporting 14:360
 establishing comparable scales and plausible values 14:355–356
 extensions and advances in analysis methodology 14:364–366
 focus on group-level inference 14:359
 general design considerations 14:354–355
 group adaptive designs 14:355
 matrix sampling 14:354–355
 matrix sampling design 14:359
 measurement characteristics 14:354–356
 multistage adaptive designs 14:355
 multistage sampling design 14:360
 population sampling 14:354
 reporting of results 14:356–357
 trends and innovations 14:356
 educational systems
 addressing issues inequality, equity, and access in 8:88
 indigenous knowledge important to 11:350
 educational technology 5:102–103
 educational telos 7:252
 educational test items, clustering of 14:634
 educational testing
 academic and professional organizations 14:16
 applications 14:19
 contemporary sources 14:16–17
 early attempts 14:15–16
 foundations 14:18
 government agencies 14:17
 issues and future directions 14:19–20
 operations 14:18–19
 principles 14:18–19
 sources 14:15–17
 testing organizations and sponsors 14:16–17
 educational theory, learnings from populism for 1:653–654
 educational transfer
 best practice as ideal norm 1:292–293
 decolonial option 1:295–296
 imagining educational transfer beyond western transcendence 1:294–297
 implications for future research 1:297–299
 interrupting education transfer 1:294–295
 logic of universality 1:292
 post-colonial option 1:295
 spiritual/ontological option 1:296–297
 theory over context 1:293–294
 in western transcendental worldview 1:291–294
 educational trends 9:177–178
 and exclusion 9:180–181
 educator-educand relation 7:248–249
 establishing relation 7:248–249
 personal character of relation 7:249
 educator-world relation 7:249–250, 7:252
 communication 7:250
 experience 7:251–252
 representation 7:250
 educators 4:266–267
 principals educators 4:471–472
 in schools and libraries 10:641
 EFA, Jomtien to twin global agendas of 1:514–515
 EFA model. *See* exploratory FA model (EFA model)

- effective professional development 5:512
 interrogating evidence for 5:512–513
 interrogating primacy of curriculum content in 5:513–514
 interrogating role of coaching in 5:514–515
 investing in 5:511–512
- effective teachers
 attitudes, beliefs & expectations 5:295–296
 characteristics of 5:288–297
 cognition 5:288–289
 collegiality 5:294
 communication 5:292
 cultural competence 5:292
 motivation 5:292
 personality 5:293
 self-efficacy 5:291
 self-reflection and reflection 5:293
 social and emotional competence 5:289–291
 values, morals and ethics 5:296
- efficacy 4:317–318
- EG design. *See* equivalent groups design (EG design)
- EI. *See* Education International (EI)
- eigenvalue decomposition (EVD) 14:590
 PCA of correlation matrix 14:592–594
- EL. *See* English learner (EL)
- ELA. *See* English Language Arts (ELA)
- ELC. *See* engineering literacies club (ELC)
- Elder praxis 10:789–790
- Elders' fish camp 3:190–191
- electracy
 Cabaret academy 10:524–526
 designing of disaster 10:518–527
 heuristics 10:520–521
 theming mystory 10:526–527
 theopraxis 10:521–522
 visceral turn 10:522–524
- electroencephalography (EEG)
 EEG for STEM education research 11:462
 examining sex differences with 11:461–462
 research approaches using 11:459–462
 spatial reasoning map to EEG data 11:460–461
- electronic databases 10:536–537
- elementary education, universalization of 1:429–431
- ELF. *See* English as lingua franca (ELF)
- elite schools
 critical studies of 2:222
 reading studies of elite schools though intersectional lens 2:220–226
- ELLs. *See* English language learners (ELLs)
- ELT theory. *See* educational and life transitions theory (ELT theory)
- EM. *See* expectation-maximization (EM)
- embedded standard setting (ESS) 14:376–377
- embodied cognition
 from amodal over embodiment to hybrid theories of conceptual representations 6:68–69
 embodiment of conceptual memory for numbers 6:70
- embodiment of conceptual memory for objects 6:69
- embodiment of memory for abstract concepts 6:70–71
- embodiment of memory for events 6:71
- embodiment of reading and writing 6:71–72
- embodied discourses 10:923
- embodied knowledge 10:922, 6:211
- embodied practices 10:922–923
- embodiment 10:922–923
 embodied being 10:923
 embodied discourses 10:923
 embodied knowledge 10:922
 embodied practices 10:922–923
- emergent bilingual students, research on home and family literacies of 10:205
- emergent disciplines, conceptual tool for 6:311–312
- emitted behavior 6:16–17
- EMME. *See* eye movement modeling examples (EMME)
- emotional intelligence 6:532
- emotional support and safety 6:293
- emotions 10:506, 11:67–68
 in education 6:150–151
 conceptualization 6:149–151
 groups of academic emotions 6:151
 measurement 6:149–151
 short history of 6:150–151
 and reconciliation 5:315–317
- employers in French ITE 5:76–79
- employment, disability and
 capability approach 9:455–456
 connecting 9:453–455
 interpreting disability employment data 9:454–455
 responding to disability employment gaps 9:453–454
 rethinking employment 9:456–459
- emulation studies 6:453–454
- engagement 6:357
- engineering education 11:36
 authenticity and inclusiveness in engineering curriculum 11:37–38
 design coaching and responsive teaching 11:38–39
 engineering in pre-college education 11:36–37
 epistemology-based framework 11:37
 focus on engineering design reasoning 11:39–40, 11:40f
- engineering literacies club (ELC) 11:50
 exploration of multiliteracies in 11:55–59
- engineering literacy 11:45
- engineering research
 alignment with teacher professional development program and 5E instructional model 11:301
 classroom implementation support and academic year activities 11:303–304
 collaborative research 11:303
 description of competition 11:304–305
 design of project 11:300–304
- evaluation outcomes and observations 11:305–306
- evaluation procedure 11:305
- guided training 11:302
- illustrative examples of student projects 11:307
- industry-based experiences 11:303
- lesson plan development workshop 11:303
- project design and activities 11:302–304
- project results 11:305–308
- project-based learning 11:300–301
- rationale for incorporating mechatronics and robotics with entrepreneurship experiences in classrooms 11:299–300
- technology business idea competition 11:304–305
- theoretical background 11:300–304
- England
 academy schools, teaching in 5:40–43
 on diverging trajectories 5:607–608
 geography GCSE in 7:39–44
 curriculum context 7:39–40
 curriculum drivers/warrants 7:39–40
 curriculum ensemble 7:39
 curriculum knowledge 7:41–44
 curriculum policy deconstruction 7:40–41
 textual deconstruction 7:40–41
 history national curriculum in 7:363–364
 legislation and policies in 9:61–63
 school networks for ITE in 5:75–76
 system leader and system leadership 4:132
 teacher accountability in 5:64–65
- English
 case of 8:168–171
 critical pedagogy of English language 10:227–228
 curriculum in Philippines 7:217
 English, internationalization and Empire 8:170
 history teachers' experience of curriculum reform 7:366–369
 language of education and literacy issues 3:65
 in language policy provisions 3:64–65
 languages and indigenization of 3:63
 multilayered dominance of English and anglophone contexts 8:169
 in Nigeria 3:63–64
 in non-anglophone contexts 8:169–170
 schools
 filming inclusive practice in 9:302
 retention in 5:129–132
- English as lingua franca (ELF) 8:170–171
- English Language Arts (ELA) 6:214
- English language learners (ELLs) 11:295–296
 Australia's EAL/D policy and context 10:68
 English language learners' writing 10:67
 English language learners' writing practices 10:72
 multilingualism, multiculturalism, and meaning making 10:71–72
 pedagogy and 10:66–67

- science lessons for 11:295–296
 in secondary schools 10:67–70
 spoken word in context 10:71
 teaching and learning cycle 10:69–70
 writing and learning across disciplines 10:66
 writing pedagogy 10:66
- English learner (EL) 10:86
 accessibility of state content assessments 10:88
 accommodations for ELs on state content assessments 10:88–89
 first language translation 10:88
 linguistic modification 10:89
 one homogenous subgroup 10:92–93
 performance on state content assessments 10:86–89
 read-aloud 10:89
 reliability and validity of other standardized assessments 10:87–88
 reliability and validity of state content assessments 10:87
 students assessment
 current issues 10:89–93
 need for evidence-based accommodations 10:91–92
 need for guidelines on assigning accommodations 10:92
 unnecessary linguistic complexity 10:90–91
- Englishization
 Anglicization of names 3:63
 as domestication, hybridization and Anglicization 3:62–63
 of indigenous languages and indigenization of English 3:63
 as restlessness 3:61–62
- entrepreneurial universities 8:15, 8:38
 entrepreneurialism 8:16
 environment 11:70
 environmental citizenship education 2:330–332
 public and private realms of 2:331
 rights and responsibilities of 2:330–331
 spatial and temporal dimensions of 2:331–332
- environmental education
 conceptions and currents of 2:210–211
 conservationist/resourcist, problem-solving, scientific, and sustainability currents 2:212–213
 EcoSchools programs 2:213
 moving forward with and in currents 2:217–218
 naturalist, value-centered, and holistic currents 2:211–212
 nature preschools and forest kindergartens 2:212
 praxis, socially critical, and feminist currents 2:215–217
 systemic, humanist/mesological, bioregional, ethnographic, and eco-education currents 2:214–215
- epilogue 7:72
 epistemic capital 6:264
 epistemic cognition
 emerging trends in epistemic cognition research 6:165–166
 issues in epistemic cognition research 6:164–165
 research
 emerging trends in 6:165–166
 issues in 6:164–165
 strands of research related to 6:162–163
 strands of research related to epistemic cognition 6:162–163
- epistemic injustice 8:17
 epistemic practices
 argumentation as 5:257–260
 nature of argumentation in science vs. other domains 5:257
 science vs. RE teachers' views of argumentation 5:258–260
 in science education 5:255–256
- epistemological and pedagogical concept, implications for partnership in ITE as 5:87–88
 epistemological evolution of grounded theory 12:118–119
 epistemological knowledge 6:211
 epistemologies and methodologies in literacy research
 appendix 10:908–909
 data collection and analysis 10:903–904
 epistemology 10:908
 methods 10:903–904, 10:909
 results 10:904
 setting 10:909
 theoretical perspectives 10:909
- epistemology 12:382
 epistemology-based framework 11:37
- equality
 and education 1:581
 in education 2:68–69
 of opportunity 13:259–260
- equipercenile equating function 14:241
 equitable access to key ideas for all students 6:332–333
- equity 1:581–584
 accountability-driven data 4:170–172
 challenges in leading for 4:403–404
 data for tracking and flexible grouping 4:173–174
 data to confirm assumptions and using data to challenge beliefs 4:173
 doctoral curriculum for preparing equity-minded teacher educators 5:478–484
 in early childhood education
 contextual translation 5:95
 equity work 5:94–96
 homogenized ECEC workforce and policy enactment 5:95–96
 moving equity agenda forward 5:96–97
 policy framing 5:94–95
 equity-oriented doctoral coursework 5:481–483
 equity-oriented teaching 5:479
 issues 4:170
 leadership for 4:402–403
- preparing teachers with CT for equity pedagogy 11:314–315
 in scientific inquiry 3:213–214
 teacher educating for 5:480–481
- equity, diversity, inclusion and decolonization (EDID) 4:307–308
 equity-oriented NTE, dimensions of doctoral education for 5:480–484
 equity-oriented teacher education, pedagogies of 5:482–483
 equivalent groups design (EG design) 14:238
 equivalent model 5:384
- Erasmus (+) student mobility
 academic learning 1:219
 contextualizing international student mobility motivations 1:223
 curriculum and learning outcomes 1:225
 employability 1:219–220
 Erasmus+ impact study survey 1:222–223
 and higher education institutions 1:220
 impact on students 1:220–221
 institutional effects of Erasmus participation 1:224–226
 institutional motivations for international student mobility 1:224
 international student mobility 1:223–226
 internationalization of student population 1:224
 motivations and effects 1:219–221
 organization and governance 1:226
 personal development and life experiences 1:220
- erasures analysis and answer changes 14:299–300
- ESS. *See* embedded standard setting (ESS)
- essays, automated scoring of 14:399
- estimation approaches
 Bayesian estimation 14:800–801
 prior distributions 14:802
 GLS 14:795
 hypothesis testing and confidence intervals for MLEs 14:799
 LS estimation 14:792–795
 maximizing loglikelihood function 14:798–799
 ML estimation 14:795–798
 model comparisons 14:799
 nonlinear least squares 14:795
 parameter estimation via Gibbs sampling 14:802–803
 properties of frequentist estimators 14:800
- Estonia
 comparisons with 5:712–716
 educational system, overview of 5:216
- ESU 2:381
- ethical educational leadership 4:138–140
 issue persisting 4:141–144
 organisational authenticity 4:144
 organisational competence 4:142
 organisational integrity 4:143
 personal authenticity 4:143
 personal competence 4:142
 personal integrity 4:143

- ethical educational leadership (*continued*)
 practice 4:140–141
 social authenticity 4:143–144
 social competence 4:142
 social integrity 4:143
- ethics
 ethical work as basis for praxis 4:414–417
 Foucault's concept of 4:413–414
 Foucault's key aspects of 4:413–414
 moral codes and acts 4:413
- Ethiopia 9:128–129
 cases of 9:128
- ethnic identities 4:431
- ethnographic grounded theory 12:116–117
- ethnographic studies 10:190
- ethnography 12:8, 12:28–29, 12:288–289
 attempts at defining 12:94–95
 counts as ethnography 12:93–94
 ethnographic logic of inquiry 12:95–96
 ethnography's roots and contemporary
 quandaries 12:92–93
 historical foundations and challenges in
 understanding of variations and
 12:92–94
 in or out of ethnographic space 12:96–97
 power and potentials of 12:98–99
 primary goals for 12:97–98
 variety of ethnographies 12:93
- ethnomethodological ethnography
 12:289–290
- ethnomethodology 12:39–40
- ethnonationalism 2:77–78
- ETT framework. *See* exploitative,
 transactional, or transformational
 framework (ETT framework)
- EU. *See* European Union (EU)
- Eugene Trivizas, talent and influence of
 10:674–675
- eugenics movement 14:6
- eugenics psychology, colonial tool of
 3:173–174
- Eurocentric education 3:17
- Europe
 curriculum reform rise in 7:28–30
 fighting cuts to higher education in 2:401
 knowledge production in
 Bologna process 8:13
 critiques of knowledge production in
 8:15–17
 entrepreneurial universities 8:15
 entrepreneurialism and Europeanization
 as governmentality 8:16
 epistemic injustice 8:17
 industry 8:14–15
 Lisbon strategy 8:13–14
 science parks 8:14–15
 states and supranational organizations
 8:13–14
 teacher education in 5:606–607
 transfer credit 8:346
 European Bologna Process 8:13
 European classical philosophers 12:46
 European countries, insights from five
 5:213–231
 European education systems support
 colonialist hegemony 1:686–687
- European Union (EU) 5:606–607
 European governance, policy coordination,
 and education 1:497–500
 expert groups and the network governance
 of education 1:498
 policy coordination under European
 Semester 1:500
 political system and education 1:496–497
 supranational integration, Europeanization,
 and education policy 1:494–497
- Europeanization as governmentality 8:16
- evaluation 13:193
 evaluation/value-added evaluation story,
 story of 7:345–346
 through large-scale assessments
 13:197–199
- EVD. *See* eigenvalue decomposition (EVD)
- Every Student Succeeds Act (2015)
 13:110–111
- evidence of knowledge
 depth of knowledge as 11:452
 flexible thinking as 11:452–453
 evidence statements 11:448–449
- evidence-centered design (ECD) 14:126,
 14:153–154, 14:163
 applications and extensions 14:167–168
 assessment as evidentiary argument
 14:163–164
 assessment engineering 14:164
 domain analysis 14:165
 implementation and delivery layers 14:167
 layers in 14:164–167
 test specifications 14:164
- evidence-informed decision making in
 schools and districts 4:169–176
- examination of mediational studies
 reviewed by Wu 4:249–252
- exclusion 9:72, 9:510–511
 assessment for 9:402–403
 core business of 9:506–507
 educational trends and 9:180–181
- executive function (EF)
 classroom conditions supportive of EF
 6:169–172
 training students' EF skills 6:172–173
- exemplars 6:612
 analogical reasoning 6:614–615
 instructional implications 6:618
 learning from worked examples 6:615–617
 observational learning 6:612–614
 problem solving followed by instruction
 6:617–618
 relations between theoretical perspectives
 6:618
- exercising agency self-regulated learning
 13:55
- exile pedagogy/curriculum and teaching in-
 between 7:398–399
- exogenous knowledge systems for
 environmental decision-making
 in Kumaon Himalayas
 11:349–357
- expansive learning model 5:386
- expectancy-value theory 11:377
- expectation-maximization (EM) 14:95
 Bock-Aitkin EM algorithm 14:95
- experiential education, service-learning as
 8:478
- explanatory measurement
 as alternative 14:99–100
 descriptive and explanatory 14:103
 differences between common approach and
 explanatory approach 14:102–103
 effects of item covariates can be dimensions
 14:102
 example of modeling item difficulties
 14:101
 example of modeling item loadings
 14:101
 examples 14:101
 generality of 14:103
 looking beyond data analysis 14:103–106
 measure and explain at same time
 14:102–103
 random respondent coefficients in
 explanative models 14:101–102
 research issues 14:106
 sets of items as indicators 14:102
 software 14:106–107
- explanatory reductionism 10:885
- explanatory sequential MMR designs
 12:428–429
- explicit knowledge 6:211
- exploitative, transactional, or
 transformational framework (ETT
 framework) 6:678
- exploratory FA model (EFA model) 14:590,
 14:595–596, 14:600–601
 interpretation 14:603–605
 number of factors 14:602
 rotation to simple structure target 14:605
 simple structure concept 14:604
 simple structure criteria and rotation
 methods 14:604–605
 two ways to find Ψ 14:603
- exploring case-based knowledge system for
 teaching 5:409–410
- exponential distribution 14:711
- extended assignments/projects 13:98
- external generalization method 12:446
- external regulation of learning 13:37
- external sources of individual differences
 6:471–472
- externally organised/managed networks
 6:418–419
- extinction
 classical conditioning 6:15–16
 and reinforcement schedules 6:17
- extrapolation considerations
 for accountability uses of through-year
 assessments 13:217
 for instructional uses of through-year
 assessments 13:217
- extrapolation inferences 13:216–217
- eye movement modeling examples (EMME)
 6:614
- eye movements in reading
 nonlinear characteristics of 10:418–419
 patterns of regression 10:419
 due to text difficulty and reader purpose
 10:419

- variables influencing fixation location 10:418–419
- eye-tracking
- approaches 11:467–468
 - in special populations
 - behaviorist perspective on 10:416
 - constructivist lens of 10:414–417
 - data processing and analysis 10:420–421
 - design and initial procedure 10:420
 - interpretations 10:421
 - limitations 10:420–421
 - major contributions of 10:417–420
 - overview of theory 10:416–417
 - presentation of 10:415
 - through social-constructivist lens 10:417
 - strategic visual sampling 10:417–418
 - understanding eye-tracking 10:415
- F**
- F* distribution 14:715
- FA. *See* factor analysis (FA)
- face to face to remote mentoring 5:730–732
- study of 5:730
- FACT. *See* foundational assessment of competencies for teaching (FACT)
- factor analysis (FA) 14:21, 14:28–31, 14:268–269, 14:600
- factorial designs 14:698
- faculty compensation and salary equity
- factors for incremental salary increase 8:195
 - faculty compensation around globe 8:198–199
 - faculty development support 8:196
 - full spectrum of academic remuneration 8:196–197
 - life insurance 8:196
 - medical insurance 8:196
 - merit, performance, and productivity 8:195
 - miscellaneous benefits 8:196
 - non-monetary compensation 8:195–196
 - retirement and savings contributions 8:196
 - salary differentials by gender 8:197–198
 - salary differentials by race 8:198
 - salary parity, discrimination, and equity 8:197–198
 - summer compensation 8:195
 - tuition assistance 8:196
- faculty or staff member
- co-facilitation by 8:273
 - co-facilitation by pair of 8:273
 - co-facilitation by pairs or teams of 8:273
 - facilitation by 8:272
- faculty salary
- bonuses and subsidies 8:194–195
 - classification systems 8:192
 - constituents of 8:192–195
 - disciplinary differences 8:193
 - full/part-time status 8:193
 - geographic region within country 8:193
 - institutional control 8:192
 - levels 8:192
 - rank 8:192
 - tenure 8:193
 - union affiliation 8:194
- fairness 14:12, 4:515
- future of fairness in measurement 14:48–49
 - related to fairness 14:46–48
 - setting context 14:45–46
 - simulation-and game-based assessment 14:472
 - as validity 14:35–36
 - linking values 14:36
 - threats 14:36
- false empathy 5:418
- false narratives of Mexican motherland 10:142–144
- family educational cultures on student success
- alterable family variables 4:263–266
 - schools and leaders help parents further develop features of 4:266–269
- family influence 11:370–381
- family literacy/literacies 10:191
- assessing and evaluating 10:173
 - building on cultural capital and funds of knowledge 10:174–175
 - conceptions of 10:182–183
 - cultural and social diversity 10:169–170
 - digital 10:171
 - for families of color 10:181–182
 - gender and family literacy programs 10:175
 - home digital literacy environments 10:171
 - implementation in United States 10:183–186
 - approaches to 10:186
 - in early childhood development 10:192–193
 - for families of color 10:192
 - insights from 10:186–194
 - leveraging equitable literacy practices in 10:193–194
 - programmatic tensions 10:190–191
 - responses to 10:185
 - specific family literacy programs 10:183
 - undergirding philosophies of 10:183
 - maintaining and promoting first/home languages 10:174
 - mediators of literacy 10:170
 - naturalistic studies of 10:169–173
 - origins of family literacy programs 10:180–181
 - programs 10:173–175
 - theoretical frames 10:168–169
- family-school relationships 10:184
- fatherlessness 4:266–267
- federal assessment policies, positive impacts of 13:111–112
- feedback 13:100–101
- authentic feedback 13:46–47
 - classroom feedback informed by responsive pedagogy 13:90–91
 - conceptualizations 13:41–43
 - emerging areas 13:46–47
 - framework for classroom feedback practice as responsive pedagogy 13:91–93
 - internal comparisons 13:46
 - literacy 13:43–44
 - model of effective classroom feedback informed by responsive pedagogy 13:93
 - for student learning in higher education digitally-enabled feedback 6:626
 - feedback in online learning environments 6:627
 - feedback seeking strategies 6:625
 - partnerships in teacher and student feedback literacy 6:624–625
 - peer feedback 6:625–626
 - supporting inclusion and diversity through 13:46
 - technology and 13:44–45
 - trends in feedback research 13:43–45
 - trust, relationships and 13:45
- females' engagement in literacy, assessments of 10:253–254
- feminisms 10:609
- feminist
- and critical social theories 12:58–59
 - current 2:215–217
 - epistemologies and theories 12:57–59
 - genealogy
 - as policy studies approach 2:105–106
 - reminder of policy studies 2:104–105
 - and indigenous research 12:41–42
- field-based experiences, educational leaders 4:377
- filmic sensing, urban humanities 10:954–955
- financing education in Zimbabwe 9:415–416
- financing inclusive education beyond financing general education 9:38
- financing "non-educational" aspects of provision 9:38
 - financing different sectors within education system 9:38
 - financing inclusive education systems 9:39–40
 - financing provided by range of stakeholders 9:38
 - financing special needs education 9:38–39
 - from financing special needs education to financing inclusive education for all learners 9:38–40
- first language (L1) 10:174
- acquisition 6:364
 - bidirectional interactions and collective variables in 10:17–18
 - findings from studies on 10:17–20
 - modeling L1 acquisition 10:20
 - nonlinearity and variability in 10:18–20
 - self-organization and emergence in 10:18
 - maintaining and promoting first/home languages 10:174
 - theories of learning and 6:364–365
- first word 6:469–470
- first-in-family experience 8:392–397
- fixed-effect model 14:849–850
- flanders, legislation and policies in 9:63–64
- flexible performance tasks 11:448–449
- flexible task performance expectations (FTPE) 11:448–449

- flexible thinking as evidence of knowledge in use 11:452–453
- flourishing 6:722–723
- fluency 10:388–390
- as art 10:392
 - assessing reading fluency 10:393–394
 - automatic processes 10:390–391
 - automaticity 10:389
 - brief history of 10:390
 - to comprehension 10:391–392
 - fluency–beyond english 10:392–393
 - prosodic processes 10:391
 - prosody 10:389–390
 - as science 10:390–392
- fluidity across meaning-making environments 10:490
- fMRI for STEM education research 11:465
- focus on forms (FonFS) 6:366–367
- folk epistemologies 4:18–20
- FonFS. *See* focus on forms (FonFS)
- food science class 11:229–230
- forecasting Delphi method 12:609
- foreign aid
- global connections and new forms of 1:199
 - new forms of 1:199–200
 - to Pacific Islands' development and education, framing 1:195–196
- foreign language contexts 6:364
- foreign language education
- critical literacies in 10:113–114
 - mere literacy to critical literacies in 10:113
- formal competence 6:366
- formal credentials 4:365
- formal education 6:298
- addressing changing contexts in 6:264–265
- formal learning 6:212
- formative assessment 13:28–29, 13:197, 6:333
- assessment for learning and 13:29–30
 - assessment reform and school reform in US context 5:583–584
 - characteristics, processes and facet-based perspectives on 5:587–589
 - and co-regulation of learning 13:38–39
 - commentaries, critiques and counter-arguments 13:29–30
 - distinguishing assessment for learning from 13:28–30
 - and external regulation of learning 13:37
 - formative assessment and assessment for learning principles 13:29
 - implications of theories 13:36–37
 - professional growth over time and developmental trajectories perspectives on 5:589–592
 - purpose 13:100–101
 - regulation as function 13:35–36
 - and self-regulation of learning 13:37–38
 - shift from formative evaluation to 5:584–587
 - and summative assessment 13:39
 - teaching and learning 13:5
 - assessment 13:2
 - critical features of effective formative assessment 13:5
- current concepts, confusions, and controversies 13:3–5
- discipline-specific formative assessment 13:6–7
- foundational concepts 13:1–3
- inference 13:1–2
- origins of 13:3
- primary purpose 13:4
- recent research on 13:5–7
- relationship between formative assessment and self-regulated learning 13:5–6
- student involvement 13:4
- technology-enhanced formative assessment 13:7
- timing and feedback 13:4
- formative classroom assessment 14:522–523
- See also* summative classroom assessment
 - characteristics 14:522
 - embedded formative assessment 14:522–523
 - formative use of summative assessment 14:523
- formative evaluation 12:603–605
- formative metaevaluation 13:264
- formative research 12:602–603
- Foucault's challenge 5:43–47
- reconstituting non-hierarchical relationships with colleagues via distancing 5:45–46
 - reconstituting relation to curriculum, pedagogy and assessment via betrayal 5:44–45
 - reconstituting relation with self via frankness 5:46–47
- foundational assessment of competencies for teaching (FACT) 14:312
- pilot data 14:312
- foundational method 14:734–750
- four 21st century social values 10:490–491
- creativity, innovation, and problem-solving 10:491
 - criticality, responsibility, and social change 10:490–491
 - multiplicity, hybridity, and flexibility 10:491
 - new literacies social values reflected in educational standards around world 10:491
 - participation & collaboration 10:490
- fourth industrial revolution (4IR) 8:54–55, 8:279–280
- enabling conditions 8:283
 - imperatives for change flowing from 8:284–285
 - implications for curriculum 8:283–284
 - potential for higher education, teaching and learning 8:281–283
 - socioeconomic realities confronting SA higher education 8:280
 - supportive infrastructure, access to devices and cost-effective connectivity in higher education 8:281
- FPCs
- involvement in education in terms of geopolitics and equity 1:592–594
 - role in education being interpreted in education research 1:592
- frameworks for conceptualizing mixed methods research 12:390–401
- France 11:286
- common core in 7:13–14
 - curriculum policies in 7:11–16
 - on diverging trajectories 5:607–608
- Frankfurt School 12:47–48
- Frederick Douglass high school disciplinary literacy project 10:566–567
- free play 11:170
- Freire's pedagogy of oppressed 6:602
- Freirean teaching for environmental praxis 1:602–604
- generative themes and dialogue 1:603
 - reinventions 1:604
 - utopias and hope 1:603–604
- French ITE
- local hierarchies and market of employers in 5:76–79
- frequentist approaches
- brief introduction to probability 14:752
 - CLT 14:755–756
 - confidence intervals 14:759–760
 - foundational concepts 14:751
 - hypothesis testing 14:754
 - hypothetical example 14:759
 - population *vs.* sample 14:751–752
 - probability distributions 14:752–753
 - sampling distribution 14:755
 - standard error 14:756–757
 - test statistics and confidence intervals 14:809
 - translating hypotheses 14:754–755
- frequentist estimators, properties of 14:800
- friendship as form of negative ethics in teaching 5:47
- FTPE. *See* flexible task performance expectations (FTPE)
- fully online distance education 8:287–289
- academic offerings 8:292–293
 - features 8:290–291
 - graduation 8:294
 - marketing 8:291
 - method 8:289–290
 - onboarding process 8:291–292
 - support structures 8:293–294
- fully-online mixed methods studies 12:698–699
- funding
- education 9:448–450
 - early childhood education 9:449
 - higher education and VET 9:450
 - primary and secondary school 9:449–450
 - environment for higher education institutions 8:185–186
 - higher education 8:183
 - theory and distribution of funding models in higher education 8:108–109

funds
 of identity approach 10:162–165
 funds of knowledge and funds of identity
 as theoretical lenses 10:164–165
 research on funds of identity in subject
 area learning 10:163–164
 of knowledge 10:468
 of knowledge approach 10:160–162
 funds of knowledge in international
 contexts 10:162
 research literature reviews on
 10:160–161
 research on funds of knowledge in
 subject area learning 10:161–162
 future-ready learners
 evolving pedagogies 5:718
 main program components 5:716–718
 multiple perspectives 5:718
 preparing future-ready teachers for
 5:716–718
 professional practice 5:718
 rigorous programs 5:717–718
 teacher professionalism 5:716–717
 future-ready teaching profession, learning
 capacity for 5:268–269
 future-ready teaching profession, lifelong
 learning capacity for 5:268–269

G

Gabriel Boric's proposals 2:80–81
 gain score model 14:383–384
 advantages and disadvantages
 14:384
 determining growth adequacy 14:383–384
 model overview 14:383
 modeling requirements 14:383
 gall on matter 9:213–215
 Galton, Francis 14:2–3
 game-based learning (GBL) 6:681
 games
 game-based assessment examples
 14:473–474
 assessment in digital ocean 14:471–472
 challenges to field 14:473
 validity, reliability, and fairness 14:472
 game-based ITSs 14:464
 hybrid forms of simulations and games
 designed for learning 6:684–685
 learning with simulations, games, and
 hybrid forms 6:685–686
 play 11:171
 typology of simulations and 6:681–685
 gamification 6:685
 gaming
 fosters complex thinking and increases
 motivation 10:586
 gaming promotes development of “soft-
 skills” 10:586
 perspectives on 10:585–586
 gamma distribution 14:713–714
 GANs. *See* generative adversarial networks
 (GANs)
 GATS. *See* General Agreement on Trade in
 Services (GATS)

Gauss-Newton method for least squares
 estimation 14:559–561
 GB. *See* governing body (GB)
 GBG. *See* Good Behavioral Game (GBG)
 GE. *See* global education (GE)
 GEM report. *See* global education
 monitoring report (GEM report)
 gender 10:252–255, 10:609–610, 1:581
 differences
 educational implications of gender gaps
 in educational outcomes 6:406
 in intelligence 6:400
 in international assessments 10:255–256
 interpreting magnitude of gender gaps
 6:401–402
 on literacy assessments and
 interpretations, critiques of
 10:254–255
 in mathematics and science 6:404–406
 in specific cognitive abilities 6:400–406
 in verbal and language abilities
 6:402–404
 gaps in
 educational outcomes 6:406
 interpreting magnitude of 6:401–402
 gender-based violence 12:404–405
 gender-cultural literacies 10:256
 gender-diverse teachers as persons,
 professionals and pedagogues
 5:208–210
 gendered and classed subjectivities in elite
 higher education
 changes to subjectivities in higher
 education 8:393–395
 classed and gendered subjectivities
 during transitional time 8:392–393
 identities 4:431
 sexuality and education 9:159–168
 and literacy practices 10:257–258
 literacy
 as cognitive process or performance and
 role of 10:255–256
 as social practice and role of 10:256–258
 perspectives on 10:250–252
 poverty and disability 9:470–479
 cost-effectiveness of intervention
 9:477
 effectiveness of intervention
 9:474–478
 evaluation approach 9:474
 learning outcomes 9:475–477
 productive investment 9:477–478
 retention outcomes 9:475
 recommendations for future research on
 10:259
 roles in self-efficacy formation 6:252
 social and cultural constructions of 10:257
 stereotypes about mathematics and science
 11:334–337
 parents and caregivers 11:336
 teachers and educators 11:337
 troubling social models of 10:251–252
 gender and sexually diverse (GSD) 9:159
 curricular and pedagogical implications
 9:161

epistemological implications of queer and
 transgender studies in education
 9:165–166
 knowing and being known 9:163
 policy and practice implications 9:162
 students in schools, desubjugated voices of
 9:162–165
 students in schools, experiences of
 9:159–162
 students of color 9:161
 toward desubjugation 9:164–165
 Gender Research Institute 8:208
 General Agreement on Trade in Services
 (GATS) 1:503–505
 generalizability considerations
 for accountability uses of through-year
 assessments 13:216
 for instructional uses of through-year
 assessments 13:216
 generalizability theory (GT) 14:9, 14:21,
 14:26–28, 14:59–60
 and D studies, infinite universe of
 14:61–63
 D study VCs 14:61–62
 error variances 14:62
 G study VCs 14:60–61
 measurement precision 14:28
 methods 14:27–28
 MGT 14:63–64
 model 14:26–27
 specification 14:64
 practical and conceptual considerations in
 application of 14:64
 random *vs.* fixed facets 14:64
 real data illustration 14:64–69
 real-data example 14:28
 reliability-like coefficients 14:62–63
 test design 14:26
 universe of admissible observations and G
 studies 14:60–61
 universe scores 14:61
 generalization 6:18
 inferences 13:216
 generalized graded unfolding model
 (GGUM) 14:78–79
 Generalized Intelligent Framework for
 Tutoring (GIFT) 6:638
 generalized least squares (GLS) 14:795
 generalized linear models (GLMs) 14:556,
 14:583
 binomial GLM 14:586–587
 categorical data 14:588–589
 definition 14:583–584
 estimating 14:585–586
 evaluating model fit 14:586
 evaluating models 14:585
 fitting 14:584–585
 inference 14:585
 for item probabilities/means 14:639
 rate model 14:587–588
 generalized partial credit model (GPCM)
 14:75–76
 generation effects, memory 6:110
 generative adversarial networks (GANs)
 14:692

- generative adversarial networks (GANs)
(*continued*)
and content generation 14:692–693
generic-cognitive skills 6:128
genetic algorithms 6:86
geographic information systems (GIS)
12:645–647
ensure enough frequency in taking GIS
integrated STEM class 11:84
integration in STEM classes, effectiveness of
11:83
as research tool outside classroom
12:648–649
as teaching tool in classroom 12:648
geopolitical currents 1:234–235
geopolitics 1:193–204, 1:597
climate change in education policy
1:598–599
micro, meso, and macro sociopolitics of
climate change inaction 1:597–598
geopolitics of knowledge (GoK) 2:261–262
Gerald Campano 2:198
GERM. *See* Global Education Reform
Movement (GERM)
German-Jordanian University 8:208
gestalt synthesis of science of learning 4:14
incongruent paradigms 4:14
GGUM. *See* generalized graded unfolding
model (GGUM)
Ghana, education in developing countries
1:208–209
Ghetto, Kovno 14:830–831
GIFT. *See* Generalized Intelligent Framework
for Tutoring (GIFT)
Giorgi, Amadeo 12:135–136
girls
analysis of topics and themes 9:490–493
autism and 9:486
with disabilities in Global South
contested accounts in disability studies
2:125–126
contested discourses over girls with
disabilities 2:124–125
findings 2:127–130
research as decolonial praxis 2:126–127
understanding complex experiences of
2:127–128
discussion 9:493–494
education
gender equity and global connection,
contested ideas of 1:613–621
in sub-Saharan Africa 9:471
methodology 9:487
GIS. *See* geographic information systems
(GIS)
GISA. *See* Global Integrated Scenario-based
Assessment (GISA)
GLMs. *See* generalized linear models (GLMs)
global and Indigenous education
See also afrocentric education
affective well-being 2:258
body text 2:251–252
cognitive well-being 2:258
colonial divide, reclaiming non-western
education 2:253–255
colonial divisions 2:252–253
in context of emerging challenges
2:257–259
different approaches to 2:255–257
ecological well-being 2:258
economic well-being 2:259
relational well-being 2:258
global childhood
construction of global childhood education
space 1:490–491
UNICEF and politics of 1:491
global children's literature
researchers of 10:640
university instructors of 10:640–641
global children's policy agency, construction
of 1:489
global citizenship 1:111
education
curricular practices of 2:48–49
moving forward during and after COVID-
19 2:50
theoretical discourses on 2:47–48
global competition over skills
attraction and retention of skills through
selective migration policies
1:632–634
COVID-19 pandemic 1:636–637
Diaspora strategies and return migration
policies 1:634
emergent research agendas 1:635–637
global developments and debates
1:631–632
migration industry 1:636
production of skills through education and
training 1:634–635
skilled immigration policies 1:633–634
student migrants as future skilled migrants
1:634–635
TVET 1:634–635
whither “skills” and “talent” 1:635–636
global connective media
YouTube
as global connective platform
1:623–624
as knowledge platform 1:624–625
videos for school study 1:625–627
global conversations about teachers and
teacher education 5:67–68
global critical philosophies 12:48–50
global democracy
cultivating social-self 1:85–87
divided in darkness 1:80–82
evolving 1:84–85
inspiring principled action 1:87–90
moving toward light 1:90–91
peril and promise of 1:79–80
returning to wellspring of democracy
1:82–84
global education (GE) 2:141–143
critical faultlines and tensions of
2:145–146
demands on education to make
difference and considering
educational of education
2:146
tension between strategic and dedicated
cosmopolitan sovereignties and
subjectivities 2:145–146
as historical social science 7:100–102
industry 5:4–7
isomorphic mechanisms in education
policymaking in 1:274–275
negotiating politics of global education
governance 1:521–523
philanthropy
digital directions of 1:588–589
shift philanthropy into domain of public
1:590
race 1:165–166
reform movement 5:31
variants of 2:144–145
world bank in global education arena
1:485–486
global education monitoring report (GEM
report) 1:553
critiques and evaluations of 1:561–562
education for all global monitoring report
1:555–557
emergence of 1:557–561
historical roots of 1:553–555
looking ahead at monitoring challenges and
pandemic recovery 1:562–563
Global Education Reform Movement
(GERM) 1:168, 5:29
antidotes to 1:171–172
arrival as alternative typology 1:168–171
in Australia 5:31–34
global education race 1:165–166
increasing competition between schools
over enrollments 1:169
increasing priority to basic literacy and
numeracy 1:169–170
and policy effects 5:34–36
privatization of public education
1:170–171
standardization of teaching and learning in
schools 1:169
tentative typology of 20th century whole
system school reforms 1:166–168
test-based accountability 1:170
global educational market 2:155–157
global English supports 1:688
global human rights 1:71
education 1:71–73
global programs for human rights
dissemination 1:73–74
from global to local 1:76
(r-)evolution of 1:74–75
regimes 1:74
global imperatives 1:199–200
national implementation 1:201–202
global indigenous education sites of
struggles 3:14
Global Integrated Scenario-based
Assessment (GISA) 6:358
global languages 1:686–687
global literacies 10:321
global literature
challenge of 10:633–634
circles 10:558

- in lives of readers 10:634–635
resources on 10:641–642
social responsibility of
 bookmakers 10:636–639
 influencers 10:640–641
 readers 10:635–636
global meaning-making
 advocacy 10:700–702
 analyses 10:700
 being ethical, respectful, and responsive
 10:702–703
 border-crossing across multitopic spaces
 10:702
 considerations for 10:703–705
 contemplative, analytic, and advocacy-
 oriented 10:704–705
 as critical culturalists 10:704
 dance of 10:703
 dimensions of 10:697–703
 sites of contemplation 10:698–699
global mobility
 in early childhood education 5:600
 in early childhood education workforce
 5:600
 and middle class families
 living mobility 1:303–304
 sense of belonging 1:304
 strategies for social reproduction
 1:304–305
 supporting learning of 5:600–601
Global North–Global South 8:211
global outcasts, concern of 10:142–143
global politics of quality in education
 5:49–53
global processes 1:250–259
global programs for human rights
 dissemination 1:73–74
global south
 decolonizing research with girls with
 disabilities in
 contested accounts in disability studies
 2:125–126
 contested discourses over “girls with
 disabilities” 2:124–125
 findings 2:127–130
 research as decolonial praxis
 2:126–127
 strengths and capacities 2:128–130
 understanding complex experiences of
 girls with disabilities from the
 Global South 2:127–128
 labyrinths of educational digital inclusion
 in 1:284–285
 strengths and capacities 2:128–130
 understanding complex experiences of girls
 with disabilities from 2:127–128
global south-north partnership
 basic knowledge acquisition 1:113
 development to support quality study
 abroad programs 1:104
 to enhance and support study abroad
 programs 1:106–108
 impact on participant’s professional
 development 1:115
 participants’ most challenging experiences
 1:114
 participants’ most rewarding experiences
 1:113–114
 professional growth and development of
 educators through study abroad
 programs in nations 1:114–115
 research methodology 1:112–113
 results 1:113
 theoretical perspectives on impact of
 globalization, study abroad sustain
 initiatives 1:109–112
 transformation of study abroad programs
 rewards and challenges 1:105–106
global teacher education reform
 gaze 1:264–265
 hands 1:262–264
 infographics 1:266–267
 permanence/change 1:265–266
 themes in visual analysis 1:261–267
 visual regime 1:267–268
 youth 1:261–262
global teaching insights 13:330
global testing regimes
 international organizations and
 comparative data 1:327–328
 mediatization and standardized testing
 1:332–333
 neoliberalism, racialized students, testing
 and accountability 1:329–332
 OECD’s PISA, high performing and high
 equity 1:328
global university rankings (GURs) 2:261
 big three GURs 2:262–264
 and colonial knowledge/power relations
 2:265–268
 and universality of quality and excellence
 GURs 2:264–265
globalism 2:78
 and collective value 2:360
globalization 1:197, 13:306–308
 curriculum 7:57–59
 implications of globalization for education
 research 1:7–8
 perspectives for understanding national
 education policymaking in context
 of 1:273–275
 as spaces and flows 1:2–4
 and the neo-liberal agenda 9:85–86
 theoretical perspectives on impact of sustain
 initiatives 1:109–112
 colonialism, decoloniality and impact on
 1:110
 theorizing global spaces and relations
 1:4–7
GLS. *See* generalized least squares (GLS)
GMAT. *See* Graduate Management
 Admission Test (GMAT)
GMM. *See* growth mixture models (GMM)
GOF. *See* goodness of fit (GOF)
GoK. *See* geopolitics of knowledge (GoK)
Good Behavioral Game (GBG) 6:19
good path. *See* Kayani:yo (good path)
goodness of fit (GOF) 14:83, 14:260–261
 indices 14:263
 numerical examples 14:268–270
 overall goodness of fit statistics 14:262
governance
 by markets 9:224–225
 new forms of 5:73–75
 governing body (GB) 4:420
 IEBs’ practices impact on GBs and business
 sector 4:425–426
 government, impact of 10:776–777
 government primary school, strengthening
 of 1:373
GPCM. *See* generalized partial credit model
 (GPCM)
graded response model (GRM) 14:76–77,
 14:89
 bifactor extension 14:92
 correlated factors extension 14:90–91
Graduate Management Admission Test
 (GMAT) 14:453–454
graduate record examinations (GRE)
 14:452–453
graduate school
 admissions tests at 14:454
 as professionalization process 3:137
grammar
 usage 6:403
 in writing 10:69
grammatical competence 6:366
Gramscian conceptual framework 1:566
grand theories, historical roots of
 12:5–6
graphic poetry of Charles Joseph Minard
 14:831–832
graphical models, Bayesian networks
 14:119–120
grassroots organizing for racial justice in
 schools 2:356–359
grateful refugee, performing identity of
 9:173
GRE. *See* graduate record examinations
 (GRE)
greatest purchase, approaches with
 2:223–225
 complexity 2:224–225
 praxis 2:223–224
 relationality 2:224
 social contexts 2:224
Greece
 disability as matter of rationalizing
 expenditure 9:419–420
 seeking free character of public education
 9:422–423
 separate schools for different students
 9:420–422
 springtime of children’s literature in 10:673
Greek children’s literature
 historical development of 10:672
 traits and trends of 10:675–676
Greek-Cypriot children’s literature
 10:673–674
grit
 grit and relationship to factors 6:192–193
 beyond grit and resilience 6:193–194
 relevance of grit for education 6:191–192
 and resilience as independent factors
 6:193
 short review of 6:191
GRM. *See* graded response model (GRM)

- grounded theory 12:27, 12:108,
12:308–309
coding method 12:113
critical theoretical applications 12:118–119
CRT and grounded theory methodology
12:119
diverging interpretations 12:109–110
ethnographic grounded theory 12:116–117
evaluating 12:115–116
future directions 12:116
grounded theory's contested family of
approaches 12:110–111
and mixed methods designs 12:117–118
ontological and epistemological evolution
12:118–119
origins and history of 12:109–110
overview of article 12:108–109
in practice 12:114–116
purpose statements and research questions
12:114
research in education 12:111–119
studies in education 12:116
theoretical sampling and theory
development 12:113–114
group meetings 5:276
developing group meeting coding scheme
5:276
knowledge exchange addressing factors
5:282
growth mixture models (GMM) 14:646,
14:649–654
addressing estimation problems
14:652–653
estimating 14:651–652
estimation problems due to model non-
identification and inappropriate
data 14:653
estimation problems related to local
maxima 14:653
estimation problems related to non-normal
probability distribution 14:652
using Mplus 14:650–654
Mplus syntax for 14:651–652
potential heterogeneity in individual
trajectories 14:648–649
selecting optimal class model 14:653–654
specifying 14:650
GSD. *See* gender and sexually diverse (GSD)
GT. *See* generalizability theory (GT)
Guanxi 1:662
Guatemala, US teachers learned about
indigenous children and schools
in 10:139–140
guidance to facilitate inquiry-based learning
6:632–633
guided play 11:171
GURs. *See* global university rankings (GURs)
Guttman errors 14:275
Gutzwa's positionality 12:62–63
- H**
Haberman's classical test theory based
method 14:294
habituation 6:13–14
Halaqah 12:182
handling divergence of QUAL and QUAN,
strategies for 12:472–475
initiation 12:474–475
reconciliation 12:473–474
verification 12:473
hard governance in era of technology 7:78
hard-to-staff Chinese rural schools, teacher
resilience in
building teacher resilience in 5:201
conceptual framework 5:192–194
conceptualizing teacher resilience
5:192–193
data analysis 5:196
data collection 5:194
factors influencing development of teacher
resilience 5:193–194
methodology 5:194–196
participants 5:194
policy context for 5:192
results 5:196–200
rural teacher resilience across professional
life phases 5:196–197
Harré's positioning theory 10:920–921
storylines and positioning triangle
10:920–921
#Hashtag 2:342–343
Hawai'i, indigenizing university of 3:72–73
Hawaiian language
infusion of Hawaiian into 3:43–44
linguistic calque 3:44–48
Haynes, George Edmond 6:675
HC. *See* hierarchical clustering (HC)
HE. *See* higher education (HE)
headship, challenges of 4:394–396
health supports children's overall
development 4:341–342
heart rate and electrodermal approaches
11:465–466
hegemonic power 1:198
HEI. *See* higher education institutions (HEI)
heritage languages 2:386–387
heroic refugee, performing identity of
9:173–174
heterogeneity 6:651
heterogeneous networks 6:419–420
heterogeneous profession, transitional fears
of 5:470–471
heterosexuality 2:173–176
hidden curriculum 6:326
hierarchical clustering (HC) 14:629–630
agglomerative methods for 14:630
hierarchical linear models (HML)
14:556–557
Bayesian estimation 14:572
complex repeated measures designs
14:572–573
generalized linear and nonlinear models
14:571
meta-analysis 14:572
modeling panel results 14:573
recommended reading 14:573
repeated measures and longitudinal data
14:570
shrinkage estimates 14:571–572
two levels, one predictor at each level, linear
model 14:568–570
underutilized applications of 14:572–573
hierarchical mode 11:3–4
high needs schools 4:187–188
characterizing high needs schools 4:187
combining contextual relevance, capacity
building, and continuous
improvement 4:191
effective leadership for 4:189–191
effective school leadership 4:188–189
high needs school leader development
4:191–192
increasing organizational capacity 4:190
leading collaborative and continuous
school-wide improvement efforts
4:190–191
resources for 4:187–188
taking contextually relevant approach
4:189–190
high school disciplinary literacy project
10:566–567
high sustainable enthusiasm 5:197
high-quality learning experience for students
4:405
higher education (HE) 3:70–73, 8:145,
8:356, 8:374, 8:406, 9:450
admissions testing in US 14:452–457
ACT 14:452
admissions testing around world
14:456–457
GRE 14:452–453
LSAT 14:453
main admissions tests used in US today
14:452–454
MAT 14:453–454
MCAT 14:453
personal characteristics to predict success
in 14:456
today's SAT 14:452
use of 14:454
in Africa, COVID-19 impact on 8:91–92
biographical approach 8:410
challenges of social media and 8:136
changes to subjectivities in 8:393–395
Kate 8:393–394
Vuong 8:394–395
class 8:409–410
class size intricacies in
admission office perspective 8:337
campus visit 8:338
changing demographics 8:341
costs 8:340
disparities by country/SES 8:339–340
pressures affecting class size 8:341
recruitment 8:337–338
registrar 8:341
retention 8:338
selectivity 8:337
student success 8:340–341
transfer students 8:339
variations across types of institutions
8:339
community-campus relationships and
partnerships 8:482–483

- completion problem for particular higher education systems 8:376–378
- conceptualizing service-learning 8:477–481
- concerns that stimulated HD 8:146
- constructions of and approaches to diaspora in 1:129–131
- critical approach 1:131–132
- critical pedagogy, and challenge of resistance 8:139–144
- curriculum 7:235–236
- curriculum and blacks in 8:353–354
- curriculum making and 7:163–172
- datafication in 1:675–677
- definitions of employability 8:406–407
- diaspora as human capital and state possession 1:130
- diaspora as marginalized and social construct 1:130–131
- disability support services 9:370–371
- early life experiences 8:411
- in Europe, fighting cuts to 2:401
- evaluation of NPM within HE sector 8:147–148
- experiencing HE 8:410–411
- experiencing university life 8:412
- focus on 8:206
- funding across countries 8:109–110
- gender and class 8:408–409, 8:412–413
- for good start in employment life 8:58–59
- governance reform in Portugal 8:154–155
- growth of 1:432–434
- higher education grant funding 8:111–112
- HRE in 2:59–61
- human rights centers 2:59–60
 - scholars at risk student advocacy seminars 2:60
 - teacher training 2:60
 - training of social workers and psychologists 2:60–61
- inequalities in 1:434–435
- institutional organization to support service-learning 8:483–485
- internationalization
- of higher education 1:129
 - ideologies 8:167–168
 - language and 8:168–171
- as kit 8:59–60
- and labor market 8:406
- legislative obligations 9:370
- linking research and teaching to benefit student learning in
- building collaborative relationships between students and staff 8:265–266
 - developing classroom consciousness between students and staff 8:264–265
 - importance of complexity and context in teaching and learning 8:259–260
 - improving teaching and learning through collaborative observation 8:263–264
 - research context and methodology 8:260–261
- restoring professional agency through CoPO 8:262–263
- in MENA, globalization impact on international cooperation and privatization 1:382–384
- missed links between education and development 1:385–387
- nomadic scholars to colonial dominance 1:380–381
- rethinking university in MENA uncertain times 1:384–385
- university as national development project 1:381–382
- minority learner experiences in 8:400
- pacific people in Aotearoa New Zealand 8:400
- new public management and 8:145–150
- in Northeast Asia
- CAMPUS ASIA program 1:452
 - challenges for regionalization of 1:453–454
 - cross-border higher education in Northeast Asia 1:448–450
 - emergence of 1:450–451
 - regionalization of 1:450–452
- North–South links in 8:212–213
- North–South partnerships 8:210
- operationalization of NPM in HE 8:146–147
- as opportunity for change 8:60–61
- performance assessments in 8:312–315
- potential for 8:281–283
- reasonable accommodations/adjustments typically afforded in 9:370
- reasonable adjustments 9:370
- research 8:408
- research funding in 8:108
- responsibilities to indigenous resurgence
- good relations 3:101–102
 - people and practices in good relations 3:105–106
 - policies as good relations 3:104
 - programs in good relations 3:104–105
 - seeing possible through individual and collective actions 3:103
 - tensions navigating transformation 3:102–103
 - transforming indigenous education through good relations 3:106–107
 - unsettling colonialism 3:100–101
- role and limitations of theories for HE
- policy 8:148–149
- selling mobility for 8:235–236
- as solid foundation for development 8:59
- in South Asia 1:431–435
- spent on 8:184–185
- strategy formations
- as double-edged process 8:1–3
 - essence of modern university and strategy formation 8:7–9
 - trends in 8:1–9
 - across world regions 8:3–7
- student completions in global north
- completion problem for particular higher education systems 8:376–378
 - completion problem for particular students and institutions 8:375–376
 - reimaging problem 8:378–379
 - student experience 8:407
 - on students
 - impacts on corporations 8:32–33
 - impacts on society 8:34–35
 - students impact 8:30–32 - for students with disability, barriers to 9:371–372
 - for students with disability, facilitators to 9:372–373
 - students with disability in higher education 9:369
 - supportive infrastructure, access to devices and cost-effective connectivity in 8:281
 - theory and distribution of funding models in 8:108–109
 - transitions and graduate labor market 8:413–414
 - trends in research funding and current challenges 8:110–111
 - truth before reconciliation within 2:275
 - untold sides of single story representations of dropout challenge in 8:359
- higher education institutions (HEI) 8:185–186
- funding environment for 8:185–186
 - internationalization strategies 1:223
 - public funds made available to 8:186–187
 - responses effect on 8:94
- historical social science 7:98–102
- global education as 7:100–102
 - and school curricula 7:99–100
- historically black colleges and universities
- black colleges today and environments 8:425
 - 1960s and civil rights 8:425
 - diversity in HBCU faculty 8:426
 - and economic mobility 8:426
 - federal recognition of 8:425
 - future of 8:426–427
 - HBCU contributions in STEM 8:426
 - historical background 8:423
 - prominent legal cases and 8:424
 - rise of industrial philanthropy 8:423–424
- Hitt, leadership frameworks and taxonomies 4:405
- building professional capacity 4:405
 - connecting with external partners 4:405
 - creating supportive organization for learning 4:405
 - establishing and conveying vision 4:405
 - facilitating high-quality learning experience for students 4:405
- HIV/AIDS 2:366–367
- HML. *See* hierarchical linear models (HML)
- Ho’omana project 3:236
- holism
- path toward voluntary embrace 1:84
 - politically engaged 1:82–84
 - toward transformative social philosophy 1:83

- holistic current 2:211–212
home literacies 10:200–201
Homo sapiens 6:36
homogeneity, testing for 14:851
homogeneous distributions 14:581
honesty 4:257–258
 foster culture of integrity 4:258
 honest and authentic 4:258
Hong Kong, principal preparation 4:362
hope 1:603–604
horizontal expertise 5:386
horizontal layout for visual displays 12:766
horizontal mobility and objectives of
 mobility 1:242
Horn's Gf-Gc model 6:203
How People Learn II (HPL-II) 6:637
Howard Gardner's multiple intelligences
 6:203
HPL-II. *See* How People Learn II (HPL-II)
HRE. *See* human rights education (HRE)
human capital building, translanguaging as
 10:50
human capital development
 on learning individual 6:524
 on social organization of learning
 6:524–525
human capital in knowledge economy
 2:245–246
human challenge of learning 10:845–846
human development 2:153–155
 South Asia 1:427–428
human information process
 working memory and constraints of
 14:173–174
human rights
 cities 2:62–63
 global programs for human rights
 dissemination 1:73–74
human rights 1:581–584
human rights education (HRE) 2:53
 in action 2:59–63
 background on 2:54–56
 in community development 2:62–63
 critiques of human rights and 2:57–58
 definition of 2:54
 in era of black lives matter and Covid-19
 pandemic 2:57
 as evolving practice 2:56–57
 evolving theories of 2:56
 goals and concepts of 2:53–54
 hegemonic, western-centered and “top
 down” 2:58
 in higher education 2:59–61
 history of 2:54–55
 impacts of 2:56
 in K-12 schooling sector 2:59
 legal and normative dimensions of human
 rights 2:55
 in national plans of action 2:59–61
 in non-formal education sector 2:61–62
 possible solutions for 2:58
 purpose and structure of chapter 2:54
 spread of 2:55–56
 transformative and emancipatory HRE
 methodologies 2:56–57
 universalism 2:57
humanitarian relief initiatives 8:208
humanizing language and literacy in
 education
 dramatic inquiry pedagogy 10:347–350
humanizing classroom communities
 10:351–353
humanizing identities 10:355
humanizing relationships and identities
 10:353–355
humanizing students' connections with
 curriculum 10:352–353
transformative crucible 10:354
transformative dialogue 10:354–355
transforming classroom communities
 10:353
hybrid approaches 12:756
hybrid displays 14:826–828
hybrid identity 8:23–24
hybrid off-line mixed methods studies
 12:699–700
hybrid on-line mixed methods studies
 12:699–700
hybrid theories of conceptual
 representations, from amodal over
 embodiment to 6:68–69
hypergeometric distribution 14:719–720
hyperreal school spaces
 affective shortfalls of 10:932–934
 discourse on cameras and student presence
 in 10:932
 on videoconference platforms 10:931–932
hypothesis testing 14:578–580
- I**
- I-TAF. *See* in-task assessment framework (I-TAF)
IB. *See* international baccalaureate (IB)
IBE. *See* intercultural bilingual education (IBE)
IBL. *See* inquiry-based learning (IBL)
IC statistics. *See* information criteria statistics (IC statistics)
ICAP framework. *See* interactive, constructive, active, and passive framework (ICAP framework)
ICILS. *See* international computer and information literacy study (ICILS)
ICT. *See* information and communication technology (ICT)
ICT4E framework in context of language diversity 10:820–821
ideas
 accessibility 6:332
 sophistication of 11:455
identity 10:149, 2:85–87, 6:330–331
 capital 10:153
 crisis 2:86–87
 ideologies of legitimacy 10:152–153
 imagined identities and imagined futures
 10:153–154
 investment, and 10:151–154
 LS and identities 10:198–199
 options 7:381–382
of pentecostal students in church and
 school 10:204
politics 4:411–412
power 10:152
research on 10:203–206
teachers' impact on identity development
 2:89–90
theoretical framework 10:198–199
ideological domain 10:827
IEA's Rosetta Stone 14:351–352
IEBs. *See* independent, external, business-based governors (IEBs)
IEL. *See* integrated educational leadership (IEL)
IEP. *See* individualized educational plan (IEP)
IF-MT. *See* integrated framework of multiple texts (IF-MT)
IHE. *See* international higher education (IHE)
IHERI. *See* international higher education; research and innovation (IHERI)
IJU. *See* International Joint Universities (IJU)
IK. *See* indigenous knowledge (IK)
IL. *See* instructional leadership (IL)
ILSAs. *See* international large-scale assessments (ILSAs)
imagining STEM with and place
 action-regeneratin 11:137
 courage to respect lessons of more-than-human ancestors 11:138
humble to more-than-human intelligences
 brings humans back down to earth
 11:138–139
introductory words 11:131
knowledge-recognizing 11:135–137
relationships-remembering 11:134–135
returning to river 11:134–135, 11:137–138
STEM as integration 11:131–132
styres' framework for indigenous thought
 11:132–138
truth of interdependence among humans
 and more-than-humans
 11:139–140
turning to river, story, and indigenous
 thought 11:132
vision-recentering 11:134
wisdom of land for sustainability 11:140
imitation studies 6:453–454
immediate memory 6:108–109
immersive video technologies 1:153–154
immigrants 12:405–406
 status 13:257–258
 students
 educational outcomes 13:258–259
 in ILSA 13:256–257
 new challenges 13:260–261
 teachers in NYC 5:676–677
IMPACT project 4:67
imperative sentence 6:470
imperialisms 1:196–198
 in “La Isla del Encanto” 10:811
implications for schools 2:181–182
implicit knowledge 6:211
importance sampling (IS) 14:779–780

- inclusive pedagogy in teacher education
 about course 5:248–249
 data analysis 5:247
 data sources 5:247
 ethical issues 5:247
 inclusion 5:244
 inclusive pedagogy 5:244–245
 methodology and methods 5:246–247
 participants 5:247
 presentation of findings 5:247
 professionalism in teaching diverse groups
 of students 5:250–252
 teacher education pedagogy 5:246
 teachers' professional identity 5:245
 theoretical framework 5:244–246
 understanding inclusion and inclusive
 education 5:249–250
 working in inclusive practices 5:247–252
 in-school literacies 10:202–203
 in-service teacher education 11:276
 appropriate time and resources 11:278
 collaborative opportunities 11:278
 design decisions associated with STEM in-
 service teacher education
 11:279–280
 emerging developments in STEM in-service
 teacher education 11:280–281
 high quality STEM in-service teacher
 education 11:278–279
 in-service program and school alignment
 11:278
 knowledge focus 11:278
 ongoing program evaluation 11:279
 STEM disciplines within in-service teacher
 education 11:277
 STEM in-service teacher education
 11:277–278
 studying STEM in-service teacher education
 11:281
 teacher participation in in-service education
 11:277
 in-service teacher professional development
 5:102
 in-task assessment framework (I-TAF)
 14:496
 inclusion 1:581–584, 9:88–91, 9:222–224
 as activity 9:274
 assessment for 9:402–403
 attitudes toward 9:328–333
 and class size 9:276
 confusions about inclusion 9:276–279
 controversy about meaning of 9:273–274
 dissolving 9:331–332
 in education 9:397–398
 education 9:89–91
 and equity
 global trends 9:3–4
 international policy context 9:1–2
 involving wider community 9:7–8
 as principles 9:5
 promising developments 9:3
 role of administrative departments 9:8–9
 school development 9:7
 use of evidence 9:5–7
 whole-system approach 9:4–5
 as location 9:274
 models of disability 9:89
 and nature of disabilities 9:275
 paying for inclusion and exclusion
 including all learners in appropriate
 educational opportunities 9:44
 developing innovative and flexible
 learning environments 9:45
 ensuring learners effectively included in
 appropriate educational
 opportunities 9:42
 ensuring transparent and accountable
 systems of inclusive education 9:43
 financing inclusive education financing
 general education 9:38
 inclusive financing of whole education
 systems 9:41–43
 introducing transparent systems of
 accountability for inclusive
 education 9:45–46
 moving toward financing takes
 preventative policy approach
 9:40–46
 promoting school-development
 approach for inclusive education
 9:42
 providing innovative and flexible
 learning environments 9:42–43
 reinforcing need for prevention-based
 policy approaches 9:43–44
 shifting focus of financing within
 education systems 9:37–40
 underpinning inclusive education by
 a school-development approach
 9:44–45
 practical issues in 9:274–276
 reconciliation 9:90–91
 relevance of paradigms in promoting
 cultural diversity and 12:422–423
 rethinking schooling for 9:506–514
 and severity of disability 9:275–276
 slippery, contentious concept 9:88–89
 and teacher diversity 9:274–275
 inclusive assessment
 examining social, educational and legal
 considerations of 13:172–173
 frameworks for 13:173–176
 education policy perspective
 13:174–175
 legal perspective 13:176
 social perspective 13:173–174
 inclusive autism 10:440–442
 inclusive communities
 insider voices 9:183–187
 in pursuit of social justice 9:188–189
 inclusive education 5:100–101, 9:229–230,
 9:508–509
 in Africa 9:405
 agenda, crisis, austerity and 9:418–425
 for all 9:13
 as catalyst for reform 9:431–432
 closing deaf schools 9:482
 co-designing lessons 9:147–148
 competing agendas and re-fragmentation of
 9:24–25
 concepts 10:443
 conditions for voice 9:145–146
 contemporary change 5:102–103
 current context 5:100–101
 deaf learners exceptional 9:480–481
 debates 9:362–363
 development of inclusive education policies
 9:399–400
 dialogs lead to changes 9:148
 for disabled children and students
 barriers to inclusive education and ways
 to address them 9:297–300
 bringing paradigm shift into education of
 disabled children and students
 9:294–296
 champions of inclusive education
 9:302–303
 developing inclusive education theory
 into practice 9:301–302
 filming inclusive practice in english
 schools 9:302
 identifying and removing barriers to
 9:303
 index for inclusion has three key
 dimensions 9:300–301
 possible solutions 9:304
 diversity of teachers 5:101
 evolution of 9:20
 fields of tension 9:387–388
 final thoughts 9:149
 gaps in 9:415
 getting started 9:146–147
 inclusive education as policy 9:388
 inclusive inquiry 9:146
 international arena to local practices
 9:389–394
 language and teachers 5:101
 in Malawi 9:405–408
 existing policies and strategies for
 9:405–407
 practices, systems and structures of
 9:407–408
 pre-service and in-service teacher training
 9:408
 Nordic model of education and
 development of 9:98–100
 role of teachers in inclusive access 5:101
 segregated separate school 9:481–482
 strengths 9:415
 strengths and gaps on policy and practice in
 9:408–410, 9:415
 students' voices 9:144–145
 Swedish education system 9:388–389
 thinking about voices and dialogs through
 process 9:149
 utilizing global social justice lens to explore
 indicators of
 classroom practices 9:109
 climate 9:108–109
 collaboration and shared responsibility
 9:109–110
 curriculum 9:108
 facilitating purposive search 9:105–106
 frameworks for examining inclusive
 education 9:105

- inclusive education (*continued*)
- inputs 9:106–108
 - leadership 9:107–108
 - outcomes 9:111–112
 - participation 9:111
 - policy 9:106
 - post-school outcomes 9:111–112
 - process 9:108–110
 - resources and finances 9:107
 - role of special schools 9:110
 - school practices 9:109
 - staff professional development and teacher education 9:106–107
 - student achievement 9:111
 - supports to individuals 9:110
 - voice/voices in research 9:143–144
 - in Zimbabwe 9:410–414
 - existing policies and strategies for 9:410–413
 - practices, systems & structures for inclusive education 9:413–414
 - pre-service and in-service teacher training 9:414
- inclusive learning environments in United States
- background of study 8:418
 - conceptual framework 8:418
 - data analysis 8:420
 - deconstructing predominantly white cultures 8:417
 - limitations 8:420
 - methodology 8:419
 - participants and setting 8:419
 - representative preliminary findings 8:420
 - research questions 8:419
 - understanding culture 8:420–421
- inclusive literacies 10:440–442
- inclusive new literacies 10:442–443
- framework of 10:442–443
 - model of 10:442–443
- inclusive pedagogies, new learning, new times
- amplified learner agency 9:73
 - awareness of learner profiles 9:73
 - categorization of differences for inclusive strategy 9:70–72
 - differentiated curriculum 9:73
 - dimensions of learner differences 9:69–70
 - historical responses to learner differences 9:72
 - measuring comparabilities 9:74
 - productive diversity 9:73–74
 - recursive feedback 9:73
 - toward an inclusive pedagogy 9:72–74
- inclusive postsecondary education
- emanating inclusion through commitment to reflective practice 9:365–366
 - facilitation as strategy for advancing inclusion 9:364–365
 - key features taking place of programmatic design 9:364
 - resisting definitions through focus on practice 9:363
 - sustainable 9:366–367
- inclusive practice in schools
- building plane, while flying 9:122–123
 - COVID-19 9:118
 - inclusion or exclusion 9:115–116
 - method 9:118–119
 - objectives 9:118
 - principles of inclusive practice 9:116–118, 9:122–123
 - results and discussion 9:119–121
- inclusive responsive leadership 4:65
- inclusive STEM 11:397
- inclusive strategy
- affinities 9:71
 - agencies 9:71
 - categorization of differences for 9:70–72
 - change 9:71
 - confounding factors of categorization 9:70–71
 - intergroup relationality 9:70
 - internal group variations 9:70
 - intersectionality 9:71
 - “lifeworld differences” framework 9:71–72
 - narratives 9:71–72
 - personae 9:71
 - unmanageable lists of differences 9:70
- inclusiveness in engineering curriculum 11:37–38
- inclusivity 9:72
- incongruent theories of action 4:20
- increasing organizational capacity 4:190
- independence 14:581
- independent, external, business-based governors (IEBs) 4:420
- acquiring educational knowledge through formal and informal training 4:423–424
 - embodying lay knowledge 4:424
 - practices impact on GBs and business sector 4:425–426
- index I² 14:850
- India 9:129
- economic, social and political context of education reforms in 5:722–723
- Indian education for all 3:136–137
- Indian education scenario 1:530
- Indianapolis Regional Office 12:88–90
- image of service 12:90
 - obstacles to implementation 12:89–90
 - RFW training 12:89
 - station 12:88–89
- indigeneity 2:26–28
- indigenous education systems 2:112
 - indigenous land knowledge and connections 2:110–111
 - indigenous philosophies of reciprocity 2:112–113
 - positionality on 2:109–110
- indigenizing universities
- of Hawai’i 3:72–73
 - process 3:70–71
- indigenous doctoral program
- indigenous doctoral program 3:132–133
 - indigenous doctorate 3:129–130
 - indigenous graduate studies 3:131
 - indigenous institution 3:130–131
 - New Zealand accreditation 3:133–134
- indigenous education systems 2:112
- indigenous knowledge (IK) 11:350, 8:104
- dismantling structural barriers 3:213–214
 - to educational system 11:350
 - focus groups 11:354–355
 - interviews 11:355
 - in Kumaon Himalayas 11:351
 - land management 11:354
 - results 11:354–355
 - in science education
 - diverse frames of reference 3:198
 - native American land management 3:203–205
 - systems 3:198–199
 - for environmental decision-making in Kumaon Himalayas 11:349–357
 - water management 11:354
 - by youth for environmental decision-making 11:353
- indigenous language
- curriculum 10:790–791
 - engaging in indigenous language recovery 10:793–794
 - renewal 10:789–790
 - engaging in 10:793–794
 - restoration 10:789–790
 - engaging in 10:793–794
 - resurgence 10:789–790
 - revitalization
 - challenges and supports to 3:58
 - context for language change 3:51–52
 - engaging in 10:793–794
 - indigenous immersion education 3:54–56
 - and indigenous immersion programs 3:54–58
 - indigenous language change 3:52
 - need for 3:53–54
 - and partnering with community 3:93
 - planning for 3:54
 - structural and contextual differences 3:56–57
- indigenous mixed methods evaluation 12:536
- indigenous models of education before contact 9:151
- indigenous peoples 1:582
- processes of learning and teaching 3:230–231
- indigenous philosophies 12:387
- indigenous psychology 3:177–178
- indigenous scholar-practitioners, co-authors as 10:786–787
- indigenous social-psychological theories 6:279–280
- indigenous storywork for education and research 3:189–197
- indigenous students, stereotype threat and 6:278–279
- indigenous-centered programming 3:89–98
- indigenous/native model of well-being 3:149
- individual differences in early language development 6:471–472

- external sources of individual differences 6:471–472
- internal sources of individual differences 6:471
- nature of individual differences 6:471
- individual learning 6:5–6
 - related approaches 6:6–7
- individualized educational plan (IEP) 9:59
 - current amendments to 9:61
- induction in teachers' professional development 5:145–147
- induction programs, aims of 5:137
 - common strategies and components used in 5:137–138
 - influence of school culture on 5:138
 - mentoring 5:138–139
 - role of school leaders in 5:138
- industrial philanthropy, rise of 8:423–424
- inequality in education 2:68–69
- infancy 6:451
 - attention 6:458–459
 - incentive value 6:460
 - learning in 6:451–457
 - moving beyond separate accounts toward integrated frameworks 6:457–461
 - social partners and social contingency 6:459–460
- infant learning, separate accounts toward integrated frameworks for 6:457–461
- influences and connectivity to community 10:181
- informal curriculum 2:136–137
- informal education concerns and technology, genealogy of 2:339–340
- informal environmental citizenship education
 - environmental citizenship education 2:330–332
 - informal and non-formal environmental education 2:330–332
 - informal learning about environment 2:332–333
 - roots and shoots, china 2:334
 - studies of 2:333–335
 - WaYS program 2:334–335
 - women in agroecology leadership for conservation, Guatemala 2:333–334
- informal environmental education, efficacy of 2:333
- informal learning 2:338–339, 6:212, 6:506–508
 - conversational agents in informal learning setting 6:644
 - about environment 2:332–333
 - recognition pivotal for lifelong learning strategies 6:523–524
- informal STEM learning 11:244–245
 - addressing adult life-long 11:248
 - broadening participation in 11:246–247
 - documenting long-term 11:248
 - in larger STEM ecosystem, role of 11:245–246
- learning in informal STEM environments 11:245
 - supporting equity and inclusion for all learners 11:247
- informal technology 6:506–507
- information acquisition, sequence of 6:703
- information and communication technology (ICT) 10:792, 1:157
 - ICT-related competencies 13:314
 - issues associated with ICT use at home 1:160–161
 - principals' work in public schools 4:307
 - problems associated with ICT use at school 1:161–162
- information criteria statistics (IC statistics) 14:653
- information literacy 2:305–306
- information processing speed 11:69
- information transmission 13:41–42
- initial research approaches 10:40
- initial teacher education (ITE) 5:83, 5:690–691, 7:159–160
 - be practical, not just conceptual 5:696
 - centering new voices, frameworks and knowledge 5:695
 - challenges inherent within partnership working 5:85–87
 - competency base 5:340
 - contributions 5:695–696
 - through preservice teachers 5:693–694
 - through program structures, content and design 5:695
 - to social justice 5:693–695
 - through teacher educators 5:694–695
 - criticality and inquiry 5:695
 - ecological model of teacher agency 5:340
 - environmental conditions 5:341–342
 - learning to teach within ITE partnerships and role of clinical practice 5:84–85
 - noticing and assessing student teachers' agency in 5:343
 - place based, embodied, experiential connected teaching and learning 5:696
 - policy background 5:83–84
 - policy context of TPAs in initial teacher education programs 5:427–428
 - principles to inform ITE partnership working 5:88–90
 - professional purpose 5:341
 - supporting teacher agency in 5:340–342
 - teacher agency 5:339–340
 - understanding nature of ITE partnerships 5:84
 - of universities in 5:613
 - university–school partnerships in 5:381–392
- initial teacher training (ITT) 5:612
 - market review in England 5:614–616
- innatism 6:368
- innatist-informed approach 6:367
- innovative pedagogies in relation to curriculum
 - barriers for individual teachers 7:137–138
 - changing learner 7:136–137
- convergent and divergent pedagogy and curriculum 7:132–134
 - discussion 7:138–139
 - divergent pedagogy 7:134–136
 - philosophy for children 7:135
 - productive pedagogies 7:134
 - rolling stone uphill 7:137–138
 - school level barriers 7:138
 - Singapore's IP 7:135
 - thinking through geography 7:135–136
- Inq-ITS, scaffolds in 14:467
- Inq-ITS. *See* Inquiry Intelligent Tutoring System (Inq-ITS)
- inquiry competencies, real time performance assessment & real time scaffolding of 14:465–467
- Inquiry Intelligent Tutoring System (Inq-ITS) 6:644
- inquiry-based developments 4:283
- inquiry-based learning (IBL) 11:96, 6:681, 6:702–703
 - design of learning question in 6:703
 - level of student/teacher centeredness in 6:703
 - conditions for 6:631
 - criticisms of 6:631–632
 - guidance to facilitate 6:632–633
 - inquiry-based learning example of 6:631
 - inquiry-focused questioning for STEM 11:97–99
- institutional competition for accumulation of economic and symbolic capital, superimposition of 8:37–40
- institutional curriculum 7:258–260
- institutional organization to support development of relationships and partnerships 8:486
- service-learning 8:483–485
- institutional service-learning 6:675–676
- institutionalized violence and forced education systems 9:151
- institutions
 - effects on 2:404–405
 - variations across types of 8:339
- instructed SLA (ISLA) 6:364
- instructional knowledge 4:487
- instructional leadership (IL) 4:11, 4:45–46, 4:74–75
 - for academic, social justice, and equity 4:28–29
 - evolution 4:15–17
 - findings 4:15–20
 - framing instructional leadership and learning through lens of management 4:18–19
 - illustrating integration of 4:75
 - leadership content knowledge 4:12
 - problem posing, thinking with theory, and theory building 4:13–15
 - purpose statement 4:13
 - science of learning lens 4:11–13
 - scientific management and facilitating learning 4:12–13
- instructional programs 6:594
- instrumental aid 6:292

- instrumental approach in curriculum design 7:201
- instrumental variable model 14:700
- integrated claim
 - in mixed methods research, real-life example of development of 12:368–369
 - using multiple methods, development of 12:363–364
- integrated educational leadership (IEL) 4:73–75
 - cognitive resources 4:77–78
 - effects 4:75–77
 - improving conceptions of 4:77–80
 - psychological resources 4:79
 - self-efficacy 4:79
 - social resources 4:78–79
 - values and ethics 4:79–80
- integrated framework of multiple texts (IF-MT) 6:356
- integrated mixed method research 12:411
- integrated programme (IP) 7:135
- integrated STEM
 - approaches to assessment of 11:414
 - assessment of student learning in 11:412–415
 - conceptual model for 11:415–417
 - curriculum and assessment
 - background and defining STEM education 11:410
 - conceptualizing “integrated” STEM learning 11:410–412
 - learning 11:410–412
- integrating art and artistic practices in mixed methods research in education
 - framing mixed methods research 12:704
 - framing various forms of arts-related research 12:703–704
 - methodology, design, or methods 12:702–705
- integration 12:501, 12:513–516, 9:382
 - accountability for acting on recommendations 9:383–384
 - alternative curriculum 9:382
 - clear ownership of recommendations 9:383
 - independent review processes 9:383
 - issues raised in reviews and inquiries 9:382
 - mentoring approaches in remote mediation 5:735–736
 - in mixed methods research 12:611–612
 - recommendations 9:382
 - review design considerations 9:382–384
- integrative analysis of qualitative and quantitative data or results 12:755
- integrative mixed methods analysis 12:492–496
 - data consolidation 12:496
 - data transformation 12:494–495
 - empirical example for 12:496
 - linking data 12:494
 - typologies and classifications 12:493–494
- integrative mixed methods research
 - combining methods for
 - causal explanation 12:444–445
 - generalization 12:445–446
 - complementarity of differences 12:443–446
 - components of design 12:442–443
 - conceptual framework 12:442–443
 - goals 12:442
 - illustrative study 12:446–448
 - interactive concept of design 12:442
 - methods 12:443
 - research questions 12:443
 - validity and generalizability 12:443
- integrative mixed methods research, quality criteria in
 - Collins 12:517
 - Hall and Howard 12:517
 - integration 12:513–516
 - O’Cathain 12:517
 - Onwuegbuzie and Johnson 12:516
 - research design 12:515
 - sampling design 12:515–516
 - selective empirical examples of validity designs 12:518–519
 - selective quality criteria frameworks 12:516–517
 - selective validity design concerns 12:517
 - Tashakkori 12:516–517
 - validity design 12:516
- integrity, foster culture of 4:258
- intellectual disability in high-income countries 9:456–459
- intelligence
 - considerations regarding education 6:206–207
 - consistencies across past years 6:206
 - further developments, debate, and integration 6:203
 - gender differences in 6:400
 - general 6:199–201
 - implications for educators and for education 6:206
 - mathematical models renewed 6:204
 - multiple intelligences 6:205–206
 - role of early mathematical models of 6:201–203
 - testing 14:5–7
 - Alfred Binet 14:5–6
 - Army Alpha test 14:6
 - Stanford-Binet and rise of anti-testing movement 14:6–7
- intelligent accountability education 13:160–161
 - accountability reform movement 13:156–157
 - aim of 13:155
 - Coleman report 13:155–156
 - expected of teachers, given working conditions 13:160
 - expert group on schools’ contribution to students’ learning 13:158–159
 - holding teachers accountable 13:157–158
 - quest for simplicity—human and social capital 13:159–160
- intelligent tutoring systems (ITSs) 14:460–461, 6:637–639
 - CAI to 14:462
 - challenges for future 6:645
 - collaborative learning 14:463
 - core components of 14:461
 - developments from 1980s to present 14:462–465
 - discovery worlds and simulations 14:463
 - example ITS systems without conversational agents 6:639
 - game-based ITSs 14:464
 - history of 14:461–465
 - learner control 14:463
 - learning environments with conversational agents 6:639–644
 - mental models 14:463
 - performance in improving learning 6:639
 - real time performance assessment & real time scaffolding of inquiry competencies 14:465–467
 - student modeling 14:464
 - tutoring for affective states and engagement 14:464
 - virtual reality 14:463
- intending teachers, career motivations of 5:662–663
- intentional communication, beginnings of 6:469
- interactive, constructive, active, and passive framework (ICAP framework) 6:594–595, 6:690
 - active 6:691
 - active mode strategies 6:694
 - compliance 6:695
 - constructive 6:691–692
 - constructive mode strategies 6:694–695
 - criteria for discriminating between modes 6:692
 - evidence in support of ICAP hypothesis 6:696–697
 - hypothesis 6:692
 - ICAP theory to evaluate cognitive strategies 6:692–697
 - interactive 6:692
 - more or less complicated processing 6:695
 - passive 6:690–691
 - passive mode strategies 6:694
 - theory 6:690–692
- interactive approach to integration, philosophical grounding that supports 12:413
- interactive mode
 - of cognitive engagement 6:692
 - strategies 6:695–696
- Interactive Strategy Training for Active Reading and Thinking (iSTART) 6:214
- interconnectedness 6:212–213
- intercultural bilingual education (IBE) 10:740–742
- intercultural education initiatives 10:739–740
- interdisciplinary approach in education 7:274–277
- interdisciplinary frontiers 14:417–429
- interdisciplinary science learning 11:1
 - challenges for creating and sustaining 11:7

- connecting disciplines within natural sciences 11:2–3
- connecting science across STEM fields 11:3–5
- connecting science with non-STEM fields 11:5–6
- flat mode 11:4–5
- hierarchical mode 11:3–4
- knowledge categorization and practices in 11:1–2
- levels of boundary crossing 11:2–6
- nature of 11:1–2
- practices in 11:6–7
- promoting 11:6–7
- strategies for facilitating 11:6
- intergenerational knowledge relations 10:789–790
- internal generalization method 12:445–446
- internal sources of individual differences 6:471
- internal validity and causal inference 14:697
- international baccalaureate (IB) 1:344
 - phenomenon of 1:345–346
 - research on 1:346–347
 - and shifting geopolitics of education under globalization 1:347–352
 - as window on international education 1:347–351
 - as window on school internationalization 1:351–352
- international computer and information literacy study (ICILS) 14:349
- international context, mathematics and science achievement in 6:404–405
- international cooperation 1:520–521
- international curriculum 1:338
 - cultural-historical-philosophical foundations of 7:67–69
 - re-turning to cultural-historical-philosophical foundations of 7:70–71
- International Day of Persons with Disabilities 2:360
- international development, NGO and 1:565–573
- international education 1:337–338, 2:140
 - global trade in
 - commercialization of international education in different countries and impacts 1:252–254
 - decoupling direct international education and migration link and emergence of post-study work rights policy 1:255–256
 - economic processes 1:251–252
 - emergence of international education and migration nexus 1:254–255
 - forces underpinning global trade in 1:250–252
 - historical and political processes 1:251
 - liberalization of international trade 1:251–252
 - market forces 1:251
 - neoliberal ideologies 1:252
 - and skilled migration 1:254–256
 - IB as window on 1:347–351
 - international higher education (IHE) 8:202
 - changing higher education landscape 1:642–643
 - curriculum and blacks in higher education 8:353–354
 - decolonial intersectionality 8:351–352
 - educating elites 1:640–641
 - expanding and changing role of 8:202
 - instructional leadership in decolonizing curriculum 8:352–353
 - power of educational diplomacy 1:643
 - soft power as category of practice 1:644
 - toward critical geopolitics 1:644–645
 - international higher education, research and innovation (IHERI) 8:202–209
 - international interaction of academics 1:244–245
 - international investment 1:361–363
 - International Joint Universities (IJU) 8:208
 - international large scale student assessments 13:181–185
 - international large-scale assessments (ILSAs) 13:129, 13:324–325, 14:348–350, 2:155–157
 - approaches to invariance testing 13:335–336
 - assessment of contextual information 13:328–330
 - bringing together ILSAs and national assessments 13:332
 - for building capabilities in low-and middle income countries 14:351–352
 - comparability of measures 13:335–336
 - context, history, and purposes of 13:325–327
 - definition and characteristics of 13:325
 - domains and populations 13:327–328
 - extreme proficiency distributions 13:132–133
 - future directions of 13:327–336
 - historical perspectives 13:325–326
 - improving estimation of performance scores 13:335
 - indicating test-taking effort 13:334–335
 - innovative item and task formats 13:334
 - invariance of multilevel constructs 13:336
 - item response theory 13:131
 - linking ILSA data to national registers 13:332
 - linking ILSA data to non-ILSA data 13:332–333
 - linking multiple 13:330–331
 - linking multiple ILSAs at country level 13:331
 - linking multiple ILSAs at individual, classroom, or school level 13:331
 - longitudinal designs and extensions of 13:331–332
 - longitudinal extensions of ILSAs at individual level within countries 13:332
 - longitudinal ILSA designs at country level 13:331–332
- main purposes of 13:326–327
- making sense of contextual information
 - about countries 13:333
 - measurement invariance 13:335
 - methodological advancements 13:333–336
 - multi-stage adaptive testing 13:334
 - novel task and test designs 13:334
 - once again, correlation is not causation 13:333
 - striving for good indicators of teaching practice 13:329–330
 - types of data and data sources 13:330–333
 - uncovering problem solving strategies and response behaviors 13:334
- international mobility of academics 1:244–245
 - academic mobility 1:244
 - irritating statistics 1:244
 - beyond mobility 1:245
- international organizations (IOs) 13:308, 2:153–155, 5:67–68
 - cultural political economy and strategic-relational approach 1:469
 - historical and sociological institutionalism 1:470–471
 - institutional entrepreneurship 1:471–472
 - international political economy 1:468
 - international relations and global governance 1:469–470
 - organizational legitimacy 1:472–473
 - policy streams and policy entrepreneurs 1:472
 - principal-agent theory 1:471
 - science and technology studies 1:474
 - theorizing work of 1:466–474
 - world culture theory 1:466–468
- international policy
 - context 9:1–2
 - framework for teachers 5:100
 - inclusion as 9:389–390
- international political economy 1:468
- international schools 1:338–341
 - early days 1:336–337
 - going forward 1:341–342
 - international curriculum 1:338
 - international education 1:337–338
 - nature of education 1:340–341
 - terminology and definitions 1:340
- international student mobility 1:230–232, 1:241–244
 - direction of mobility 1:241–242
 - statistics and indicators 1:243
 - temporary mobility *vs.* mobility for complete study program 1:242
 - value of student mobility 1:243–244
 - varied functions 1:242
 - “vertical” *vs.* “horizontal” mobility and objectives of mobility 1:242
- international test commission (ITC) 14:233
 - guidelines 14:233–234
- international testing, accountability in age of 13:306–312
- international trends
 - in school decision making 4:205–206
 - on teacher evaluation 13:287–297

- internationalization
 - of curriculum
 - for all students 8:222–227
 - best practice principles for 8:229–230
 - characteristics of best practice 8:228–229
 - inclusive, holistic and values driven 8:229
 - purposeful, planned, ongoing 8:229–230
 - rethinking 8:227–228
 - systematic, system-wide and context relevant 8:230
 - of higher education 1:129, 8:168–171
 - definitions and conceptualizations 1:240–241
 - exceptional term and undercurrents 1:239–240
 - future perspectives 1:247
 - international mobility and international interaction of academics 1:244–245
 - international student mobility 1:241–244
 - varied modes of knowledge transfer 1:245–246
 - ideologies of 8:167–168
 - of substance of curricula, teaching and learning 1:246
- internationalization and global competence 1:109
 - as perceived by Global North and South 1:109–110
- internet
 - activism 7:417–418
 - characteristics of internet environment 10:488–490
 - assumption of multiliteracies and online mobility, creation, and networking 10:489
 - constancy of change 10:489
 - easier, cheaper, and faster access to information 10:489–490
 - fluidity across meaning-making environments 10:490
 - as information and learning environment in higher education 8:312
 - inquiry and research 10:555–557
 - critical evaluation of online information 10:557
 - forming goals and asking questions 10:555
 - generating search terms and investigating search results 10:555–556
 - locating information 10:556–557
 - search activity
 - methodology 10:544
 - pedagogical implications 10:548
 - research into internet users' interaction processes 10:544–546
 - research into students' criticality on 10:547–548
 - search completion research 10:546–547
 - three broad research themes 10:544–548
- internships 8:462
 - contextual changes impacting 8:464–465
 - in education 8:462–463
 - expansion of internships to work-integrated learning 8:465
 - recommendations for future internship developments 8:465–467
 - reducing participation barriers 8:466
- interpersonal and intrapersonal skills 14:479–480
 - appear to change over time but not in exactly the same way cognitive skills do 14:482
 - approaches for assessing 14:486–488
 - findings related to 14:480–485
 - improved 14:483
 - multidimensional forced-choice 14:487
 - noncognitive and cognitive skills 14:485
 - others-assessments of soft skills 14:484–485
 - performance tasks 14:488
 - predict academic outcomes 14:480–481
 - predict non-academic outcomes 14:481
 - ratings 14:486–487
 - SJT 14:487–488
 - subgroup differences in 14:483–484
 - terminology issues 14:478–479
 - unlike case with cognitive abilities, there are multiple, distinct 14:482
 - uses for 14:485–486
- interpersonal relationships with peers and teachers 6:289–291
 - peer relationships and academic achievement 6:290
 - peer relationships and motivation 6:290
 - peer relationships and social-behavioral outcomes 6:291
 - relationships with peers 6:289–290
 - relationships with teachers 6:289
 - teacher–student relationships and achievement 6:289
 - teacher–student relationships and motivation 6:289
 - teacher–student relationships and social-behavioral outcomes 6:289
- interpretation and use argument (IUA) 14:36–37
 - action to support IUA 14:40
 - alternate perspectives on meaning of validity 14:40
 - decision/use inference 14:40
 - extrapolation inference 14:39–40
 - generalization inference 14:39
 - inferences in 14:37
 - IUA and validity arguments 14:40–42
 - performance inference 14:39
 - principled approaches in test design to support IUA 14:38
 - scoring inference 14:39
 - and validity argument use case 14:38–40
- interpretation-and-use argument 13:214
- interquartile range (IQR) 14:726
- intersectional analysis 2:220–222
- intersectional disablism model of dis/Ability, overrepresentation and social and 9:501–503
- intersectionality 2:116–118
 - Africa 2:120–121
 - Caribbean 2:118–121
 - disability studies in 9:249–252
 - disability and class/socio-economic status 9:251
 - disability and gender 9:251–252
 - disability and migrant/refugee status 9:252
 - disability and race 9:250–251
 - framework for discerning injustice 9:250–252
 - intersectional case-study 9:252–253
 - leveraging intersectionality 9:254–256
 - across disciplines 2:117
 - in education 2:117–118
- intertext model 6:354
- interviews, number of 12:113
- IOs. *See* international organizations (IOs)
- IP. *See* integrated programme (IP)
- IQ testing, progressivism and professional demarcation 9:442
- IQR. *See* interquartile range (IQR)
- IRT. *See* item response theory (IRT)
- IRTree models 14:333–334, 14:336–338
 - comparison of 14:338
- IS. *See* importance sampling (IS)
- ISE model 5:505–508
 - basic stage of 5:506–507
 - be used 5:509
 - contribution to teacher educators 5:509
 - important to get to know 5:508–509
 - intermediate stage of 5:507
 - third stage of 5:508
- ISLA. *See* instructed SLA (ISLA)
- Islamic missionaries, impact of 10:775
- ISM internationalization strategies 1:223
- isolation
 - leadership 4:319–320
- Israel, peace education in 1:96–99
 - educational settings 1:96–97
 - research findings 1:97–99
- iSTART. *See* Interactive Strategy Training for Active Reading and Thinking (iSTART)
- Italy
 - legal framework 9:59–61
 - legislation and policy in 9:59–61
 - networks of professionals in ITE for primary school teachers in 5:77–78
- ITC. *See* international test commission (ITC)
- ITE. *See* initial teacher education (ITE)
- ITE in England, school networks for 5:75–76
 - local hierarchies and market of employers in French 5:76–79
- item formats and evaluation practices
 - developing and scoring constructed-response item types 14:190–191
 - developing common selected-response item types 14:189–190
 - general principles in selecting item types 14:186–188
 - item types 14:191
 - item types in context of test design 14:185–186

principles for developing items 14:188–189

item pool 14:210–211

item response theory (IRT) 13:131, 14:7–9, 14:21, 14:31–33, 14:72–73, 14:86, 14:255–256, 14:269–270, 14:273, 14:309, 14:331, 14:360–361

extensions and applications 14:8–9

goodness-of-fit 14:83

IRT-latent regression 14:361

JMLE 14:80–81

Markov chain Monte Carlo 14:82

measurement precision 14:33

methods 14:32

MMLE 14:81

model 14:31–32

model-based rater effects indices 14:310–312

multidimensional extensions of ordinal IRT models 14:332–333

multiple imputations 14:361–362

parameter estimation of 14:80

Rasch measurement 14:8

rater model results compared to observed score metrics 14:312–318

real-data example 14:32

review of 14:74–80

summed score to IRT scaled scores 14:96

test design 14:31

from unidimensional IRT models to MIRT models 14:87–89

in United States 14:8

item selection function 14:211–212

constraint management 14:212–213

end of test 14:213

fully adaptive item selection 14:211–212

item selection with limited adaptivity 14:212

item-descriptor matching method 14:376

iterative approach to integration, philosophical grounding that supports 12:413

iterative feedback designs 13:43

ITSs. *See* intelligent tutoring systems (ITSs)

ITT. *See* initial teacher training (ITT)

IUA. *See* interpretation and use argument (IUA)

J

jackknife method 14:767–768

examples 14:768

further comments 14:768

January Queer Organizing 2:360

Japan 11:285

implantation and development of 8:22–23

university in East Asia 8:25–26

Japanese language classrooms, critical literacies praxes in 10:114–119

JDR model. *See* job-demands-resource model (JDR model)

Jeff literacy lab 10:564–566

Jenga model 8:59

Jennifer Alford 2:199–200

JLR. *See* Journal of Literacy Research (JLR)

JMLE. *See* joint maximum likelihood estimation (JMLE)

job choice theory 4:446–447

job readiness as well as occupational preparation 8:454

job satisfaction 4:316–317

job-demands-resource model (JDR model) 5:214–215

joint displays in mixed methods research 12:754

approaches to integration 12:755–756

common intents for 12:757–760

guidance 12:756–757

joint maximum likelihood estimation (JMLE) 14:80–81

Journal of Literacy Research (JLR) 10:903–904

content analysis 10:903–904

joy in learning 5:715–716

justice, teacher educating for 5:480–481

Justice for George Day 2:360

justice-minded teacher educators, doctoral curriculum for preparing 5:478–484

justice-oriented doctoral coursework 5:481–483

justice-oriented NTE, dimensions of doctoral education for 5:480–484

justice-oriented teacher education, pedagogies of 5:482–483

justice-oriented teaching 5:479

justifications for narrative inquiry 12:127–128

personal justification 12:128

practical justifications 12:128

social justifications 12:128

K

K-12 classrooms, implementing integrated STEM curricula in 11:159–160

K-12 global/international education terminology and trajectories 2:140–145

K-12 schooling sector 2:59

K-12 schools in United States

challenges of teaching and learning disciplinary literacies 10:54–55

critiques to SFL pedagogy and responses 10:60–62

SFL 10:55–60

K-12 STEM education settings 11:200–203

K-12 teachers' beliefs and practices in STEM integration 11:251–261

k-means algorithm 14:628–629

Ka Lau Ulu, sticking point for 3:44–48

Aia o Māmāi ke kaula i lole 3:47

continuation of 3:47

He aha ka inoa o ka holoholona e umō mai ne 3:47–48

Oko a ka palaoa, oko a ke koholā 3:44–45

Pakalaki 3:45–47

Pupuka o hope o ka pōkake! 3:45

kastom leadership 3:154–156

katakana project 10:115–116

Kate 8:393–394

Kaupapa Māori theory 3:181

application of 3:73

Kayaní:yo (good path) 3:212–213

Ohneganos Ohnegahde:gyo 3:207–208

reconnecting youth to waterways and tradition 3:214

KBD. *See* knowledge-based development (KBD)

KC. *See* knowledge component (KC)

Kendall's tau 14:744

key informants, key knowledgeable, and reputational sampling 12:157–158

Kisizi language 10:799–800

(re)activating and (re)centering language in/to cultural practices 10:801–802

fueling innovation 10:801–803

language 10:799–800

symbolic performances and semiotic resistance 10:800

knowledge 2:75–77, 6:209, 9:163

building 6:213–215

change 7:261–263

competing views of knowledge in Swedish curriculum discourses 7:262–263

curriculum ideologies and different views of 7:258–260

curriculum theorizing on 7:263–264

curriculum theorizing on knowledge question 7:263–264

deficit vs. asset 6:210

defining 6:209–210

depth of 11:449

dimensions of 6:212–213

diplomacy

based on negotiation, collaboration, cooperation 8:207

broader than cultural diplomacy 8:204

build and strengthen relations between and among countries 8:207

characteristics of 8:206–207

commonality of issues 8:207

from cultural, science, education diplomacy 8:204–205

different from soft power 8:205–206

different levels of collaboration 8:207

diversity of actors and partners 8:206

education diplomacy 8:205

examples of 8:207–208

expanding and changing role of IHE 8:202

focus on higher education, research and innovation 8:206

inclusive of science diplomacy but more comprehensive 8:204

issues and challenges 8:209

misconceptions 8:204

reciprocity-mutual but different benefits 8:207

recognition of different needs and collective use of resource of actors 8:206

soft power approach juxtaposed to diplomatic approach 8:205–206

- knowledge (*continued*)
 study knowledge diplomacy 8:204
 toward working definition of 8:203
 divisions in curricula 7:257–258
 divisions in curricula, premises for 7:257–258
 economy
 bubbling knowledge economy 1:282–283
 human capital in 2:245–246
 labyrinths of educational digital inclusion in global south 1:284–285
 missing links of KBD 1:283–284
 through racial capital lens 2:245
 racially minoritized individuals and 2:247–248
 “rational” knowledge economy 1:281–282
 theories for change 2:248–249
 transforming 2:248
 translating power of “Creative Britain” 1:286
 exchange addressing factors 5:281–283
 addressing factors 5:282–283
 group meetings 5:282
 providing support 5:283
 sharing resources 5:282
 speaking same language 5:283
 experience of 7:80–81
 formulating official view of knowledge in Sweden 7:263
 framing issue of knowledge traditions 7:256–257
 framing question of knowledge in curriculum 7:262
 funds of knowledge approach 10:160–162
 implications 6:215
 issues of production and utilization of academic perspectives 8:102–103
 global south perspectives 8:103–104
 knowledge and curriculum change 7:261–263
 learner-centered curriculum ideology 7:259
 in learner-centered curriculum ideology 7:259
 leveraging existing 6:213–214
 measuring 6:210–213
 mobilization 3:214–215
 oligopolistic structures of knowledge production 8:38–39
 oligopolistic structures of knowledge transfer 8:39–40
 of powerful and powerful knowledge 7:88
 production in Europe 8:15–17
 academic capitalism 8:16
 and relational expertise 5:386
 scientifically rational curriculum ideology 7:259
 in scientifically rational curriculum ideology 7:259
 several types of 6:572
 social efficiency as curriculum ideology 7:259
 in social efficiency as curriculum ideology 7:259
 social justice and social and epistemic access to knowledge 7:91
 social reconstruction as curriculum ideology 7:259–260
 in social reconstruction as curriculum ideology 7:259–260
 structures of 7:96–98
 in Swedish curriculum discourses 7:262–263
 for technology integration
 educational digital competence 5:266–268
 learning capacity for future-ready teaching profession 5:268–269
 trends in learning technology development 5:264–265
 traditions and curriculum structure 7:256–257
 types 6:211–212
 updating and revising 6:214–215
 VR and knowledge generation 6:640–641
 knowledge, skills, abilities (KSA) 6:638
 knowledge, skills, abilities and judgments (KSAs) 14:448
 assess KSAs 14:443–444
 knowledge component (KC) 6:638–639
 knowledge-based development (KBD) 1:283–284
 knowledge-level continuum in mixed methods 12:640–641
 with existing data 12:641–642
 Korea
 implantation and development of 8:24–25
 university in East Asia 8:26–27
 KSA. *See* knowledge; skills. *See* , , abilities (KSA)
 KSAs. *See* knowledge; skills. *See* , , abilities and judgments (KSAs)
 kuleana 3:235
 Kumoan Himalayas, indigenous knowledge in 11:351
 Kumoan Himalayas, “our land our life” and educational context in 11:349–350
 kurtosis 14:729
- L**
 LA. *See* learning analytics (LA)
 La Borinqueña 10:812
 labor markets 1:547–552
 HE and 8:406
 and skill formation systems 8:439–440
 transitions and graduate 8:413–414
 Labu, leadership moment in 3:152–153
 laddering programs 3:93–94
 LaNA. *See* literacy and numeracy assessment (LaNA)
 language curriculum 5. 0 (LC5) 7:216–219
 balancing outcome and process 7:217–218
 forms of 7:217
 implementing an evidence-based LC5 7:220–221
 LC5 components 7:219
 outcomes-based pedagogy 7:219
 self-management and life skills 7:219
 sociocultural aspect of 7:218–219
 technology integration 7:218
 language policy and planning (LPP) 10:123
 positioning of multilingualism in 10:124–128
 scope of 10:123–124
 language socialization (LS) 10:198–199
 language(s) 10:210–213, 7:67–69
 acquisition 1:548–549
 development 10:79
 arts as autonomous to dialectical and dialogic social practices 10:314–316
 classroom communication about knowledge and rationality 10:314–315
 epistemological regimes 10:315–316
 change
 context for 3:51–52
 indigenous language change 3:52
 choice and education 10:817–818
 in contact 10:718–719
 creativity in and 10:48–49
 development 10:2
 adding and storing words in memory 10:2
 coining words 10:6–7
 early utterances 10:2–4
 figurative language 10:9
 language in interaction 10:8–9
 language skills 10:9–10
 multilingual exposure and 6:472–473
 politeness 10:9–10
 pretend play 10:9
 social context and 6:472–473
 socioeconomic status and 6:472
 taking up offers of new words 10:4–6
 telling stories 10:10
 English in language policy provisions 3:64–65
 as fluid and variable in relation to social purpose and context 10:312–314
 higher education 8:168–171
 and indigenization of English 3:63
 language-enabling technologies 10:792
 of Latina/o/x lowriders, research on 10:205–206
 learning 10:209–210, 10:794
 and literacy
 of Latina/o/x lowriders 10:205–206
 lenses for building 10:159–167
 myth 10:142
 beyond orthography 10:752–753
 partners 10:140–142
 origins of 10:141
 pedagogy, promoting linguistic diversity in 10:313–314
 perspectives focused on language/text valuing 10:468–469
 play 10:343
 figurative language 10:343
 of possibility and social justice 5:726–727

- power in neocolonial project 1:687–688
in STEM education
 disciplinary literacy 11:43
 engineering literacy 11:45
 language and literacy 11:42–43
 language-rich STEM education 11:48
 mathematics literacy 11:45–46
 multimodality 11:46–47
 practical example of affordances of
 language in STEM 11:48
 science literacy 11:44
 STEM literacy 11:43–46
 technology literacy 11:44–45
system in China 10:747–753
and teachers 5:101
large-scale assessments
 evaluation and accountability through
 13:197–199
 domains and constructs 13:198
 examples of large-scale assessments
 13:198
 measurement characteristics and design
 considerations 13:199
 purpose and policy relevance 13:198
 reporting of results 13:199
 trends and innovations 13:199
 purpose and policy relevance of 14:348
large-scale international assessments,
 leadership in 4:109–114
large-scale survey studies 9:159–162
latent class (LC) 14:637
 cluster analysis 14:629
 extensions 14:641–642
 generalized linear models for item
 probabilities/means 14:639
 maximum likelihood estimation
 14:642–643
 model selection issues 14:643–644
 models with explanatory variables
 14:640–641
 restricted models for categorical items
 14:640
 simple LC and mixture models
 14:637–639
 software 14:644
 three-step latent class analysis 14:641
latent growth model (LGM) 14:646–647
 Mplus syntax for 14:650
 specify 14:650
latent learning 6:23
latent variable mixture modeling (LVMM)
 4:542–543
Latin America
 colonialism and coloniality in 10:736–737
 contemporary forms of resistance in
 10:742–743
 continuing tensions 10:743
 educación Propia 10:738–739
 education
 assessment to cope with low quality of
 education 1:405–406
 and consolidation of national state
 1:399–400
 contextual record and basic concepts
 1:398–399
 Covid-19 and right to education
 1:407–409
 curricula 1:402–405
 decentralization 1:401–402
 in developing countries 1:210
 ending 1:409
 expiration of 20th-century milieu 1:401
 lifelong learning 1:402
 teacher professionalization, identity, and
 collective representation 1:406–407
 IBE from within 10:740–742
 intercultural education initiatives
 10:739–740
 literacy as form of domination during
 colonialism in 10:736–737
 literacy as form of resistance in 10:737–742
 literacy education 10:139
 mechanisms of coloniality 10:737
 in relation to children's literature
 10:655–659
 school leadership in
 field of study and relevance to region
 4:469–470
 policy drivers and areas for future
 research 4:473–475
 policy trends 4:471–473
 research trends 4:470–471
 social configuration and literary heritage
 10:655–656
 transfer credit 8:348
 visions for future 10:743
Latina/o/x lowriders, research on language
 and literacy of 10:205–206
Latino male teachers and bilingual students,
 literacy and identity of
 10:203–204
Latinx immigrant
 community-based educator's lived
 experiences 5:674–676
 lived histories to prepare Latinx immigrant
 students for college and career
 readiness 5:675
Latinx students, cultivating and affirming
 assets of 5:674–675
Law School Admission Test (LSAT) 14:453
Lawrence Stenhouse on artistry of teaching
 5:652–653
LC. *See* latent class (LC)
LC5. *See* language curriculum 5. 0 (LC5)
LCDM. *See* log-linear cognitive diagnosis
 model (LCDM)
LCK. *See* leadership content knowledge
 (LCK)
LDA. *See* linear discriminant analysis (LDA)
leaders 4:266–269
 development 4:393–394
 developing leaders 4:396
 formation 4:411–413
 character of leadership 4:412–413
 character of leadership 4:414–417
 Foucault's concept of ethics 4:413–414
 school leadership and identity politics
 4:411–412
 leader–member exchanges, quality matters
 in 4:517
 leader–member exchanges
 organizational outcomes of 4:518
 leadership 4:109–111, 4:401
 and attainment 4:111–112
 benevolence 4:256–257
 challenges in leading for equity 4:403–404
 changing leadership as key to turnaround
 4:162
 changing student demographics 4:24
 character of 4:414–417
 competence 4:258–259
 as consistent and contiguous 3:157
 conveying caring, empathy, and respect
 4:257
 crucible of leadership character formation
 4:416–417
 cultivating culture of care 4:257
 in decolonizing curriculum 8:352–353
 development 4:374–378, 4:393–394
 in US 4:370–371
 dispositions 4:384–386
 distributed leadership 4:48–49
 for education and learning 4:30–31
 education theorizing and concepts 4:29–31
 effective schools 4:217
 and family factors
 comprehensive reform efforts 4:278
 family engagement 4:273
 family engagement around world
 4:276–278
 historical look at family engagement
 4:273–274
 impacts of family engagement and school
 leadership 4:272–273
 influence of western leadership standards
 in policy borrowing 4:274–275
 international leadership standards and
 family engagement 4:275
 leadership standards related to family
 and community engagement
 4:274–275
 models of family engagement 4:273
 formation
 character of 4:412–417
 crucible of 4:416–417
 finding “sweet spot” between
 fragmentation and coherence
 4:411–413
 Foucault's concept of ethics 4:413–414
 school leadership and identity politics
 4:411–412
 standing for common good 4:414–415
 values and accountability 4:415–416
 frameworks 4:472–473
 frameworks and taxonomies 4:404–406
 and governance for improvement
 4:209–215
 for high needs schools 4:189–191
 history of leadership research
 4:62–63
 honesty 4:257–258
 as hope 3:156–157
 identity construction 4:431–433
 collective endorsement 4:433
 implications for 4:435
 individual internalization 4:432
 leadership development 4:430

- leadership (*continued*)
- relational recognition 4:433
 - toward inclusive leadership identity theories 4:433–435
 - women and 4:430–431
 - inclusive and culturally responsive 4:65
 - inclusive approach to education 4:402
 - indigenous leadership 3:151
 - as influence 4:45
 - influence and impact 4:66
 - instructional leadership 4:45–46
 - international trends 4:401–402
 - isolation 4:319–320
 - job satisfaction 4:316–317
 - as key to reintegration 4:161–162
 - as knowledge guardianship 3:153–154
 - leadership for equity 4:402–403
 - leadership for learning 4:26–29
 - for learning 4:342–344
 - managerial leadership 4:46
 - moments 3:152–157
 - moral leadership model 4:47–48
 - in non-western countries 4:476–483
 - openness 4:260–261
 - pedagogical strategies for leadership development
 - individual reading 4:387
 - leadership dispositions 4:384–386
 - leadership knowledge and understanding 4:386–387
 - leadership skills 4:382–384
 - lectures/discussions and forms of information presentation 4:387
 - multi-use pedagogical strategies 4:388–389
 - pedagogical strategies 4:387
 - pedagogical strategies for enhancing leaders' skills 4:383–384
 - transfer of learning 4:389–390
 - underlying processes 4:382–383
 - underlying processes 4:387
 - perspectives on school improvement 4:217–218
 - planned change 4:217
 - and policy
 - case studies of four secondary schools in England 4:350
 - connecting leadership with policy enactment 4:348–349
 - findings 4:350
 - research 4:349–350
 - policy implementation 4:217
 - policy trends 4:24–26
 - practices in school turnaround 4:164
 - preparation programs
 - common content areas 4:364
 - commonalities, variations and trajectories in 4:363–366
 - frameworks 4:363
 - operational features 4:364–365
 - sources of content 4:364
 - trajectories 4:365–366
 - principal efficacy 4:317–318
 - for professional accountability 4:297–298
 - rationale for leadership changes 4:162
 - reliability 4:259–260
 - school and kastom leadership 3:154–156
 - school culture 4:217
 - school improvement journeys 4:218
 - school leaders 4:255–256
 - school leaders enacting change 4:406–407
 - search parameters and opening considerations 4:316
 - social networks 4:218
 - social processes in shaping of 4:526
 - Solomon islands 3:151–152
 - standing for common good 4:414–415
 - stress/burnout 4:318–319
 - to student learning, disentangling pathways from 4:246
 - styles 4:204
 - theories about 4:36–38
 - for successful school turnaround 4:165–166
 - teacher leadership through collective 4:197–198
 - theory 4:106
 - transactional leadership 4:46
 - transformational leadership 4:46–47
 - trusted and trusting 4:261
 - typology for 4:45–50
 - and values 4:45
 - values and accountability 4:415–416
 - and vision 4:45
 - as worship 3:152–153
 - leadership, character of 4:412–413
 - leadership content knowledge (LCK) 4:12, 4:16–17
 - learner centered approach 10:451–452
 - learner differences
 - age 9:69
 - communities of commitment 9:70
 - corporeal attributes 9:69
 - dimensions of 9:69–70
 - ethnos 9:70
 - family 9:69
 - gender 9:70
 - language 9:70
 - locale 9:69
 - material conditions 9:69
 - physical and mental abilities 9:69
 - race 9:69
 - sex and sexuality 9:69
 - social class 9:69
 - symbolic differences 9:70
 - learner diversity as major contemporary contextual factor 6:260–262
 - learner-centered curriculum ideology, knowledge in 7:259
 - learner-centered model of CLT 6:367–368
 - learning 10:494, 13:22, 13:66, 13:301–302, 4:17, 4:496, 6:37, 8:94–95, 8:281–283
 - academic habitus of PhD students 8:431–433
 - to act in vocational practice 7:209–210
 - designable elements affording learning to act in vocational practice 7:209–210
 - learning processes and activities 7:209
 - adaptation on ground 6:562
 - additional language 6:364
 - analysis 6:650–651
 - from and through relational exchange 10:793–794
 - Hawaiian 10:793
 - Hawaiian and Blackfoot 10:793–794
 - Mohawk 10:793
 - assessments 13:65–66, 5:713–714, 6:210
 - attention and 6:123–125
 - balancing academic rigor with inspiring joy in 5:715–716
 - from behaviorist perspective 6:12–21
 - biomimicry connect to educational theory 10:844–845
 - capacity for future-ready teaching
 - profession 5:268–269
 - from cases 13:152
 - changing paradigm 10:839–842
 - from China 4:480–481
 - in classrooms 10:559
 - commonalities 6:446
 - communicative approach into practice 6:370–371
 - communities 6:606
 - in community development 2:325–326
 - conceptualizations 6:512
 - in context 6:3–5
 - related approaches 6:4–5
 - contrasting theories of learning 6:364
 - through cooperation and joint reciprocal reflection 5:573–574
 - counts as success 6:563
 - creating supportive organization for 4:405
 - cross cultural lenses on
 - practice important, understanding cultural communities of 6:268–271
 - studies of differences in pathways of 6:271–273
 - understanding culture 6:267–268
 - cross-theoretical principles of 6:515–518
 - about cultures 8:420
 - data extraction and synthesis 6:650
 - and development 6:518
 - and development, sociocultural theories of 10:198–199
 - differentiated optimization 6:564
 - disabilities' unsettling intersections
 - intersectional framing of SLD 10:281–284
 - tensions in conceptualization of 10:280–281
 - disciplinary literacies 10:54–55
 - across disciplines 10:66
 - in disciplines 6:305
 - assumptions underlying conceptual framework 6:305–306
 - core constructs for 6:307–312
 - for designing adaptive instruction in 21st century 6:312
 - development of conceptual framework for disciplinary literacies 6:306–307
 - distinctions 6:446–447

- of diverse and globally mobile early childhood education workforce 5:600–601
- ecology framework 6:438
- education profession finds itself today 10:842
- educational research, future of 10:845–846
- embracing biological-evolutionary perspective 10:844–846
- emergence of constructivism 10:840
- emergence of cultural forces and individual participants 10:838–839
- emergence of linguistic perspective 10:840
- emergence of more balanced CLT models 6:368–369
- environments with conversational agents 6:639–644
 - conversational patterns in human and computer tutors 6:640–641
- evidence that mindfulness training improves learning 6:649
- from examples 12:700
- extending psycholinguistic perspectives in educational research 10:841
- extending socio-psycholinguistic viewpoint for classroom practice 10:841
- and feedback. Seeking evidence of 13:67–69
- fidelity to designs 6:561–562
- final notes on reading theories of 6:6
- focusing on higher order and interdependent skills 6:565
- in formal and informal settings 6:258–259
- framing reading and writing education from psychological viewpoint 10:838
- fundamental processes of
 - active and interactive learning with others 6:113
 - comprehension processes 6:111
 - conceptual change and restructuring processes 6:113
 - conditions improve comprehension 6:111–113
 - conditions improve learning 6:113–114
 - conditions improve memory 6:109–111
 - goal-directed or purposeful reading 6:112
 - importance of feedback and guidance 6:114
 - learning by invention, learning from errors, preparation for future learning 6:113–114
 - learning from multiple representations 6:112
 - learning from multiple sources 6:112
 - learning from questions and explanations 6:112
 - memory processes 6:108–109
 - metacognitive monitoring and self-regulated learning 6:114
 - practice effects in memory and skill acquisition 6:109
 - skill acquisition in language, reading, and math 6:110
- goals and learning outcomes 6:326
- governments, systems, schools, families and communities creates ways of 4:340–341
- heterogeneity 6:651
- hinting at theories 4:17
- human challenge of learning 10:845–846
- hybrid forms of simulations and games designed for 6:684–685
- implications for pedagogy from linguistic perspective 10:841–842
- implications for pedagogy from psychological viewpoint 10:839
- within indigenous schools and communities 4:339–342
- in infancy 6:451–457
 - associative learning 6:451–452
 - imitation and emulation 6:453–454
 - social learning 6:453–457
 - statistical learning 6:452–453
- instructional leadership and role of principal 4:26–27
- language ideology 10:830–831
- leadership for 4:30–31, 4:342–344
- leadership for 4:26–29
- learner-centered model of CLT 6:367–368
- learning, development and schooling 6:40–41
- from learning sciences, perspectives on
 - brief history of 6:44–45
 - methods in 6:50–51
 - outlook on advancing research on learning in 6:51–52
 - (self-)concept of 6:45–46
 - theories of learning and guidance in 6:46–50
- across lifespan 6:443–449
- limits to optimizing 6:561–563
- linear, hierarchical, and cause and effect assumptions 4:19
- literacy and language learning from biological-evolutionary perspective 10:845
- and literacy out of school
 - ethical dimensions of technology 6:377–378
 - foregrounding racial justice and power 6:376
 - learning on move 6:376–377
 - new methodologies 6:375
- managing own learning 8:432–433
- materiality data used to complement self-report data as evidence of 8:78–79
- in maths 6:323–334
- methods in 6:572–573, 6:649–651
- mindfulness training improve learning 6:648–649
- mitigating risks to optimizing 6:564
- moderator analysis 6:652
- more to know in instructional science 6:562–563
- from multiple texts 6:353–363
- national framework for 5:714–715
- need to claim and advance indigenous peoples' processes of 3:230–231
- needed for successful optimization 6:563
- networks 6:410–411
- architectures 6:411
- in schools 5:284–285
- from neuroscience, perspectives on 6:57–66
- observational 6:612–614
- optimal learning in classroom 6:369–370
- outcomes, identification and attainment of 8:463–464
- pathways 8:58–61
- through play-in-interaction
 - gameplay, animation and learning 6:661
 - language play 6:662
 - learning 6:662
 - play, mediation and everyday disputes 6:658–659
 - play as participation 6:660
 - play politics 6:660–661
 - role of play and games for learning 6:659–660
- positive or negative influence on 10:585–586
- power in 2:324–325
- presumption of competence 6:302–303
- process 13:51
 - aspects shared among theories of 13:53
 - assessments of 13:60
 - assessments of learning events 13:59–60
 - basic cognitive operations 13:53–54
 - complexities attending operational definitions for 13:55
 - conditions triggering operations to modeling learning processes 13:55
 - conventional instruments for measuring 13:55–59
 - observation systems 13:56
 - operational definitions 13:52–53
 - operationally defining 13:54–55
 - operationally defining cognitive operations and processes 13:53–55
 - process measurements 13:59
 - psychometric issues 13:60–61
 - states, events and processes 13:51–52
 - tracing learning processes using fine-grained data 13:57–59
- process of learning language 10:841
- progressions
 - applications of 6:547–549
 - assessment 6:550
 - assessment 6:551
 - components 6:546
 - content of learning progression components 6:549–550
 - current and future research 6:551–554
 - developing context-specific 6:552–553
 - diversity of learning contexts 6:552–553
 - emphasis on ordering *vs.* pathway 6:546–547
 - equity critiques of 6:549–551
 - exploring student diversity 6:553–554
 - formative (classroom) assessment 6:548
 - generalizability 6:547
 - ordering of learning progression components 6:551
 - providing information about learning context 6:552

- learning (*continued*)
 representing diversity 6:554
 responding to diversity 6:554
 scope and grain size 6:547
 standards, curricula, and instruction
 6:550–551
 standards, curriculum, and instruction
 6:547–548
 summative (classroom and large-scale)
 assessment 6:548
 variation in 6:545–547
 psychological experimentation model
 10:842–843
 quantitative approaches to identifying
 qualitative variation in 8:76–78
 to read and write positioned and framed in
 past 10:838–842
 to read and write
 implications 6:340–341
 knowledge and skills that influence
 reading development 6:335–336
 knowledge and skills that influence
 writing development 6:336–337
 reading and writing relations 6:337–340
 recognizing significance of identity,
 language and culture 6:566
 relationship between conceptual metaphors
 and discourse 10:843–844
 relationship between creativity and
 6:136–138
 resources 5:567–568
 resourcing for sustainable change 6:562
 risk of bias 6:651–652
 in rural East Africa
 assessment guidelines 13:206
 building an Afl culture through
 questions 13:205
 building assessment competence
 13:205–206
 methodology 13:203–204
 research and learning context 13:203
 results 13:204–206
 scholarship of 8:275
 in school 6:723–727
 future perspective on well-being and
 learning 6:727
 science of optimization 6:564
 in sciences
 abstraction 6:346
 failure of common sense 6:345–346
 languages of science 6:346–348
 lexical density 6:347
 logical connectives 6:347
 multimodality 6:348
 nature of curriculum 6:349–350
 nominalizations 6:348
 polysemy or ambiguity of meaning 6:348
 visualizations 6:348
 vocabulary and academic language 6:347
 search strategy and inclusion criteria 6:649
 self-efficacy for 6:32
 sequence, implementing 5:521–524
 teaching episode 1 5:522
 teaching episode 2 5:522–523
 teaching episode 3 5:523–524
 shift in thinking 10:842–844
 with simulations, games, and hybrid forms
 6:685–686
 social construction of 6:298–300
 social context of 6:573
 social media and 8:132–133
 in social movements 2:326–327
 social/cultural constructivist perspective on
 learning “problems” 6:300–301
 sociocultural and cognitive perspectives on
 6:1–11
 sociocultural perspective on 6:37–42
 spaces and materiality
 framing analysis of complex learning
 environments 6:394–395
 present–shaping spaces for networked
 learning in contemporary societies
 6:393–394
 spaces for teacher candidates 11:316–317
 standardization of 5:32–33
 standardized
 broad principles to inform alternative
 approach to educational
 policymaking 5:25–26
 broader, global policy conditions
 5:21–22
 policy problem 5:21
 problem represented 5:20–21, 5:24–25
 standardized assessment 5:24
 standardized curriculum 5:22–23
 standardized teaching 5:23
 study 5:236–237
 teacher-led model of CLT 6:366–367
 from teaching 5:524–525
 technology development, trends in
 5:264–265
 theoretical justification for benefits for
 6:593
 theories of learning and implications for
 assessment 13:10–16
 as acquiring taught, and implications for
 assessment 13:11–12
 behaviorist views of 13:11
 building knowledge through social
 interaction, and implications for
 assessment 13:14–16
 cognitive constructivist views of learning
 13:12–13
 different approaches to learning and
 assessment be combined
 13:16–17
 folk views of learning 13:11
 implications for assessment 13:11–12
 implications for assessment 13:13
 implications for assessment 13:14–15
 importance of curriculum 13:16–17
 as individual sense-making, and
 implications for assessment
 13:12–13
 issues of progression and generalizability
 in sociocultural approaches
 13:15–16
 to learn and implications for assessment
 13:17–19
 sociocultural/situated views of 13:14
 theories of learning and L1 acquisition
 6:364–365
 trajectory 5:521
 transfer of 4:389–390
 translating theories of learning to L2 context
 6:365–366
 turn 4:20–21
 types of 14:686
 21st century teaching and 7:352–354
 understanding importance of complexity
 and context in 8:259–260
 in virtual communities 2:327–328
 ways of framing reading and writing
 education 10:838–839
 through work 6:381–389
 work and 2:236–243
 through work
 adaptability and vocational education
 8:453–454
 augmenting vocational educational
 experiences to meet 21st-century
 work requirements 8:452–453
 developing skillful and adaptable
 workforce 8:460
 educational processes promoting
 adaptability 8:458–459
 goals for and processes of vocational
 education 8:454–455
 institutional-based activities that incite
 authentic work experiences
 8:457–458
 intentionally and actively integrating
 students’ experiences 8:456–457
 organizing and providing workplace
 experiences 8:456
 promoting learning agency 8:459–460
 responsive vocational education
 curriculum and pedagogic practices
 8:455–460
 securing “hard to learn” knowledge 8:459
 in ZPD 6:41
 learning analytics (LA) 14:404–408,
 8:331–332
 adaptation and recommendation 14:407
 adoption of 8:332–333
 analytics of unstructured natural language
 data 6:77–79
 behavior modeling 14:406
 change management 8:333–334
 discourse analytics in automated analysis of
 student writing 6:77–79
 discourse analytics in computer-supported
 collaborative learning 6:77
 feedback and score reporting 14:407–408
 machine learning in 14:408–411
 proficiency modeling 14:406
 student modelling 14:405–406
 trends and opportunities for future of
 14:411–412
 learning management system (LMS)
 8:132–133
 social media incorporated into 8:133–136
 Learning Set Performances (LSP) 11:446
 least squares estimation (LS estimation)
 14:792–795

- leave-one-out cross validation (LOOCV) 14:287–288
- legacy of behavioristic approach 6:20
- legal criminalization 2:181
- Lego Foundation 1:589–590
- Lego model 8:59–60
- lesbian, gay, bisexual, transgender, queer, intersex, agender/asexual people (LGBTQIA people) 2:179
- expanded legal protections 2:183–184
- implications for schools 2:181–182
- increased visibility and protections for 2:182–185
- legal criminalization 2:181
- looking forward 2:185–186
- medical pathologization 2:181
- people as deviant and diseased 2:180–182
- positive, public visibility 2:183
- queer educators 2:182
- school curricula 2:182
- terms and concepts 2:179–180
- lesson design 5:236
- lesson study (LS) 5:570
- background on history and internationalization of 5:570–571
- key features of LS from epistemology of complexity 5:571–574
- and phases 5:572–573
- practice and benefits 5:571
- results regarding practice of 5:574–578
- leveraging existing knowledge 6:213–214
- Levine observed score equating method 14:243
- lexicon 6:467–468
- LFPs. *See* low fee private schools (LFPs)
- LGBTQ+
- MMT of university students identifying as 6:535–536
- LGBTQ+ inclusive curriculum 7:435–436
- critical democratic education and 7:436–437
- greeting 7:436
- narrative and counternarrative 7:437
- rhetoric and discourse 7:436–437
- LGBTQIA people. *See* lesbian, gay, bisexual, transgender, queer, intersex, agender/asexual people (LGBTQIA people)
- LGM. *See* latent growth model (LGM)
- LGTQ+ 1:582–584
- academic freedom 1:583
- child labor 1:582
- education, trade unions, and democracy 1:583–584
- trade union rights and democracy 1:583
- violence, harassment, and safe schools 1:582
- libraries
- educators in 10:641
- in higher education
- diversity and inclusion 8:130
- economics of information 8:127–128
- library collections 8:126
- library personnel 8:125–126
- library services platforms 8:129
- library usage 8:128
- linked data 8:128–129
- movement toward networked systems 8:128–129
- open source 8:129
- organization of 8:124–125
- public services 8:124
- technical services 8:125
- life based literary narratives in multicultural teacher education 7:395–396
- life course theory 6:532
- lifelong learning
- analytic perspectives and theoretical approaches question dominant discourse 6:524–526
- capacity for future-ready teaching profession 5:268–269
- European contexts for 6:522
- interplay between international and national agendas for change 6:522–523
- learning as social practice, focusing on social organization of learning 6:525–526
- learning as social process that focuses on learning individual 6:525
- in Northern Nigeria 10:778–783
- reframing debate and action beyond “formal, non-formal or informal” learning 6:526–527
- lifespan literacy 10:468
- likelihood ratio test (LRT) 14:653–654
- likelihood-based methods 14:808–812
- Linda’s story 8:359–360
- homeless and new in town 8:359
- thrown into deep end 8:359–360
- linear association test 14:581–582
- linear discriminant analysis (LDA) 14:590, 14:597–598
- linear equating function 14:241–242
- linear model 14:568–570
- generalized linear and nonlinear models 14:571
- linear regression analysis
- assumptions of 14:555
- disturbance term 14:555
- estimation of 14:551–555
- extensions of 14:556–557
- formulation of 14:548–551
- homoscedasticity 14:555
- linearity 14:555
- multicollinearity among independent variables 14:555
- linearity 14:555
- linguistic landscape project 10:116–118
- linking data 12:494
- Lisbon strategy in Europe 8:13–14
- listing purposes 13:136–137
- literacy and numeracy assessment (LaNA) 14:352
- literacy education 10:228
- additional neural factors relatable to reading 10:872–875
- of indigenous multilingual children 10:139–140
- Latin America 10:139
- radical hope 10:139
- US teachers learned about indigenous children and schools in Guatemala 10:139–140
- literacy education naturalized 10:875–877
- narrow-gauge synthesis of neuroanatomy related to reading 10:868–872
- in Nigeria 3:65–66
- realism in literacy research 10:877
- useful knowledge about neurophysiology of reading 10:867–868
- literacy research (LR) 10:904, 10:921–922
- data collection and analysis 10:903–904
- about LRA 10:903
- methods 10:903–904
- results 10:904
- literacy research association (LRA) 10:903
- Literacy Research: Theory, Method, and Practice (LR:TMP) 10:903–904
- content analysis 10:904
- literacy/literacies 10:252–255, 10:439–440, 10:529, 10:584, 10:719–720
- acquisition in Chinese 10:753–755
- assessments, critiques of gender differences on 10:254–255
- assessments of male and females’ engagement in 10:253–254
- as autonomous skills and defining literacy as social processes and ideological 10:309–311
- awareness of data practices 10:492–493
- call to action 10:493–494
- character action, design, context and game rules 10:584
- characteristics of internet environment 10:488–490
- church literacies 10:201–202
- in classroom 10:310–311
- as cognitive process or performance and role of gender 10:255–256
- collapsing great divide 10:297–300
- community literacies 10:200–201
- conceptions of 5:451
- connecting colonial and contemporary regimes of truth with regimes of 10:729–731
- consequences of queering 10:265–266
- construction of gender in 10:250–261
- in contact 10:718–719
- contemporary issues in 10:718–720
- critical pedagogy of 10:227–228
- criticality and justice 10:491–492
- criticality and social change, research centered on 10:493–494
- criticism on ACL 10:690
- definitional contrasts between autonomous and ideological 10:309–310
- digital equity 10:492
- digital games as multimodal systems 10:584–585
- digital identity and citizenship, research centered on 10:494
- digital literacies & digital games 10:587
- discourses, rise of 10:593

- literacy/literacies (*continued*)
 enacted and emergent narrative understandings 10:584
 enacted and emergent narrative understandings 10:584
 enhancement and achievement impact 10:779–783
 formulating and implementing workable language of 3:66
 and gender of Latino male teachers and bilingual students 10:203–204
 great divides 10:296–297
 history and definitions of 10:717
 home and family literacies of emergent bilingual students 10:205
 home literacies 10:200
 and identity
 of Latino male teachers and bilingual students 10:203–204
 literature review on literacy in multiple contexts 10:200–203
 LS and identities 10:198–199
 new literacy studies 10:199
 of Pentecostal students in church and school 10:204
 research on 10:203–206
 research on digital biliteracy 10:204–205
 research on home and family literacies of emergent bilingual students 10:205
 research on language and literacy of Latina/o/x lowriders 10:205–206
 theoretical framework 10:198–199
 in-school literacies 10:202–203
 interconnectedness of literacies online 10:493
 interpretations, critiques of gender differences on 10:254–255
 of Latina/o/x lowriders, research on 10:205–206
 learning 10:209–210
 learning with(in) white and straight places 10:262
 life-span perspectives 10:717–718
 literacies, language & schooling 10:308–318
 literacies social values reflected in educational standards around world 10:491
 literacy normativity, black queer literacies, and restorative literacies 10:262–263
 literacy studies 10:199
 literacy studies of spatial theory 10:929–930
 literacy/multiliteracies
 FL multiliteracies post NLG's manifesto 10:106–107
 literacy in late 20th and early 21st centuries 10:105–107
 new and emergent directions in multiliteracies language education 10:107–109
 literate lives of queer students of color 10:263–265
 literature review on 10:200–203
 local literacies 10:301–302
 mentoring in hybrid spaces 10:476
 methodologies and research constructs 10:300
 models 10:886
 abduction 10:888
 analogy 10:888
 epistemology of 10:891–893
 implications of reductionism for 10:887
 implications of unobservables for 10:888–889
 inference 10:888
 issues from epistemology 10:884–891
 literacy models and reductionism 10:886–887
 modeling unobservables 10:888
 models in science 10:883–884
 and reductionism 10:886–887
 reductionism as epistemological issue 10:884–885
 skepticism for 10:891
 transcendental arguments 10:888
 unobservables as epistemological issue 10:887–888
 moving forward 10:720–721
 and multimodal social communication 10:439
 multiple modalities and digital literacies 10:304
 myth 10:142
 in Nigeria 10:774–778
 in Nigerian education 10:773–774
 and Nigerian language ecology 10:774
 non-mainstream literacies 10:201
 normativity 10:262–263
 of online of reading and research 10:488
 of pentecostal students in church and school 10:204
 as political project 10:218–219
 popular discourses of school space and Covid-19 10:930–936
 posthumanism, materiality and networks of 10:304–305
 practical considerations in our lives and in everyday contexts 10:494
 practices 10:258
 in family literacy for families of color 10:193–194
 identity construction via literacy, language, and religious practices 10:210–212
 language and literacy learning 10:209–210
 power via literacy, language, and religious practices 10:212–213
 practices in classroom that engage religious literacies pluralistically 10:213–214
 sensorium in 10:482–483
 pre-phonological roots of 10:39–40
 procedural literacy 10:585
 producing and accessing adequate literacy materials 3:66
 reading, critical imagination in 10:339
 recommendations for future research on 10:259
 research and practice in literacy teacher education 10:474–479
 research on 10:203–206
 research on digital biliteracy 10:204–205
 research on language and literacy of Latina/o/x lowriders 10:205–206
 school-based literacies and digital games 10:585–586
 second phase socio-cultural studies of 10:301–302
 self-criticality 10:492
 shifting social values and opportunities for social change 10:490–491
 sites of normativity 10:263–264
 sites of promise 10:266–267
 and social communication 10:443–444
 social or cultural facts and phenomena in 10:886–887
 as social practice and role of gender 10:256–258
 within socially organized practices 10:299–300
 South Asia 1:427–428
 and spatial production of school during pandemic
 studies 10:230–231, 10:469
 subverting sites of normativity 10:264–265
 teaching and learning 10:494
 theoretical framework 10:198–199
 in times of crisis 10:717
 “transformative” literacy actions 10:584
 under-realized potentials 10:491–493
 unsettling local 10:303–304
 ways with words 10:298–299
 literature 10:500–501
 asserting presence 10:340
 belonging in possible futures 10:340–341
 collective engagement in digital world making 10:341
 critical approaches to teaching Africanist presence in 10:228–229
 critical imagination, play and figured worlds 10:338–339
 critical imagination and transformative activist stance 10:339
 critical pedagogy of 10:227–228
 defining critical imagination in literary reading 10:339
 figurative language 10:343
 figured worlds 10:339
 freeing social signs 10:343
 language play 10:343
 play and figured worlds 10:339
 positioning and repositioning 10:342
 repositioning through ruptures 10:342
 rewriting spaces 10:342
 uncertainty and imagination 10:338
 world making 10:340–343
 world making for presence, belonging, and collective action 10:340–342
 writing selves into texts 10:342
 lived curriculum making stories 7:342–346
 boundaries 7:340
 curriculum 7:340

- images and metaphors of teaching 7:341–342
- literature review 7:340–342
- narrative inquiry 7:342
- teacher knowledge & knowledge-for-teaching 7:340–341
- trends and issues on boundaries 7:346–347
- living life to fullest 9:202–203
- living schools 7:350
- 21st century teaching and learning 7:352–354
- education for sustainable development 7:352
- progressive education 7:351–352
- and purpose of education 7:352–354
- LMIC. *See* low-and middle-income countries (LMIC)
- LMR-LRT. *See* Lo-Mendell-Rubin likelihood ratio test (LMR-LRT)
- LMS. *See* learning management system (LMS)
- Lo-Mendell-Rubin likelihood ratio test (LMR-LRT) 14:653–654
- local agendas, new forms of 1:199–200
- local curriculum-making, children's participation in
- authorized in national curricula 7:183
- child as social actor 7:182
- children as citizens in curriculum-making 7:184
- children as curriculum informants 7:182–183
- children as project protagonists and collaborators 7:182
- children in rhizomatic curriculum-making 7:184
- children's curriculum-making with place, home and community 7:185
- curriculum in contemporary times 7:181–182
- decentering child in posthuman curriculum-making 7:185
- origins of 7:181
- locality partnerships as support for professional accountability 4:300–301
- log-linear cognitive diagnosis model (LCDM) 14:112–114
- constraining 14:114–116
- logic and assumptions 13:185–188
- logical capacities development 6:235–238
- content and development 6:236–237
- reasoning development as form of divergent thinking 6:237–238
- logistic regression 14:255
- lognormal distribution 14:711–712
- LOITASA project 1:372
- Lok Jumbish 1:530–531
- London school and development of classical test theory 14:2–5
- contributions after Spearman 14:4–5
- Francis Galton 14:2–3
- intelligence testing and eugenics movement 14:5–7
- Pearson and Spearman 14:3–4
- long-term memory 6:109
- longitudinal data, features of 4:539
- longitudinal designs in school leadership research 4:540
- longitudinal dynamic networks 4:532
- longitudinal focus, absence of 10:185
- longitudinal MMR designs 12:427–428
- longitudinal studies 4:540
- LOOCV. *See* leave-one-out cross validation (LOOCV)
- loving engagement and empathy 2:360–361
- low fee private schools (LFPSS) 1:177
- access and affordability 1:360–361
- case of 1:180
- chains and international investment 1:361–363
- definitions and complexity of reality 1:355–356
- development of 1:357–358
- factors leading to 1:358–360
- hidden costs and private tuition costs 1:360
- language of instruction 1:360
- over-supply of teachers 1:358
- overall poor performance of public sector 1:358–360
- low income/working class minoritized individuals and knowledge economy 2:246–247
- low-and middle-income countries (LMIC) 9:126
- low-cost private school chains 1:361–363
- LPP. *See* language policy and planning (LPP)
- LR. *See* literacy research (LR)
- LR:TMP. *See* Literacy Research: Theory, Method, and Practice (LR:TMP)
- LRA. *See* literacy research association (LRA)
- LRT. *See* likelihood ratio test (LRT)
- LS. *See* language socialization (LS); lesson study (LS)
- LS estimation. *See* least squares estimation (LS estimation)
- LSAT. *See* Law School Admission Test (LSAT)
- LSP. *See* Learning Set Performances (LSP)
- Lucia's case 5:674
- Lucia's lived experiences, key learnings of 5:676
- LVMM. *See* latent variable mixture modeling (LVMM)
- M**
- M-Decree
- embeds right to reasonable adjustments in educational legislation 9:64–65
- strengthens right to enroll 9:64
- tightens supervision of referral of students to special education 9:65
- M2PL model. *See* multidimensional compensatory model (M2PL model)
- M4BL. *See* Movement 4 Black Lives (M4BL)
- machine learning (ML) 14:404, 14:684–685, 6:83–85
- applications to educational measurement 14:690–693
- approaches to training 6:84–85
- deep learning 14:686–689
- in education 6:86–89
- instructor aids 6:87
- student aids 6:87–88
- future of 6:88–89
- history of 6:85–86
- in learning analytics 14:408–411
- machine learning/data mining 14:464–465
- software frameworks 14:689–690
- types of 14:685–686
- macro level 4:129–130
- framework for system leadership 4:128–131
- makerspace as/for STEM education 11:107–109
- in formal educational settings 11:108
- implementing makerspace in STEM in K-12 settings 11:108–109
- in informal educational settings 11:107–108
- MAL. *See* meaninglessness, alienation and loss of culture and purpose (MAL)
- Malawi
- inclusive education 9:405–408
- policies and practices in 9:404–417
- Malaysia
- accessible and caring 4:486
- diligent and hardworking 4:487
- encourage teamwork 4:486
- findings 4:486–487
- firm and perfectionist 4:487
- good communicators 4:487
- inclusiveness 4:486
- instructional and curriculum knowledge 4:487
- limitations and future studies 4:488
- methodology 4:485–486
- reflections 4:487–488
- research context 4:485
- women leadership in 4:485
- male engagement in literacy, assessments of 10:253–254
- Malta's educational system, overview of 5:216
- Mana Academy 3:165–166
- mana model 6:280–281
- Mana motuhake 6:282
- Mana tangatarua 6:283
- Mana tū 6:282–283
- Mana ūkaipō 6:281–282
- Mana whānau 6:281
- managerial leadership 4:46
- managerialism 4:46
- managerialism 4:46
- managing external demands
- creating context for coherence 4:212–213
- leveraging networks to sustain attention to internal goals 4:212
- in schools 4:211–212
- Mandela, Nelson
- university student learning outcomes initiative 8:310
- Manen, Max van 12:136–138
- mangle of contemporary geopolitics 1:691–696

- MANOVA. *See* multivariate analysis of variance (MANOVA)
- mantel haenszel (MH) 14:253
- manual assembly, as adjunct to 14:207
- Māori language 3:83–84
- Māori student success 6:280–281
- MAR mechanism. *See* missing at random mechanism (MAR mechanism)
- marginal maximum likelihood estimation (MMLE) 14:81
- Marianna Souto-Manning 2:196–197
- markets 9:219–222
- back to 9:222
 - governance 9:224–225
 - models and performativity in teacher education 5:741–742
 - market model, performativity and implications for education and teacher education 5:741–742
 - marketization and performativity are not universally adopted 5:742
- Markov chain Monte Carlo approach (MCMC approach) 14:82, 14:95, 14:780
- Marxist political economy of education 1:30–31
- masculinity 2:173–176
- culture and context 2:172–173
- mass education, competences in differentiated curriculum and in 7:282–283
- massive open online courses (MOOC) 1:246, 6:262, 6:426–427, 6:645
- mastering grammar 6:470–471
- mastery experiences 6:291–292
- use of 11:392
- matching law 6:19
- material-economic arrangements 5:35
- materiality
- of literacy 10:304–305
 - of marae and long-lasting values 6:391–392
 - of print-based picturebooks 10:625–626
- maternal education 6:473
- mathematical content 6:329
- mathematical literacy (ML) 13:220, 13:222–225
- implications for teaching and learning 13:224–225
 - role of tasks 13:222–223
- mathematical literacy 6:326
- mathematical models of intelligence 6:201–203
- basic models 6:201
 - hierarchical models 6:202–203
 - primary mental abilities 6:201–202
 - understanding factor model of g 6:201
- mathematical problem solving 11:463
- mathematical processes and practices 6:329–330
- mathematical reasoning 11:22–23
- for STEM lessons, developing a habit of mind of 11:25
 - for STEM lessons, developing habit of mind of 11:25
- mathematical thinking and learning 6:324, 6:326–331
- mathematician, ways of working as 11:30–32
- different kinds of models for different sort of purposes 11:32
 - generalizing and specializing to generate and test conjectures 11:31
 - problem posing 11:31–32
 - problem solving 11:30–31
- mathematics
- classroom
- conducting group meetings 5:276
 - creating learning networks in schools 5:284–285
 - curating tablet resources for teachers 5:285
 - developing teachers' use of technology 5:274–275
 - factors 5:277–281
 - factors impact teachers' use of technology 5:273–274
 - group meetings 5:275–276
 - knowledge exchange addressing factors 5:281–283
 - methodology 5:275–276
 - purpose 5:275–276
 - results 5:277–283
- coursework and the supply of STEM graduates 11:366
- as language tool in STEM education 11:21–22
- interpreting and explaining 11:22
 - representing and modeling 11:21–22
- learning 6:323
- classroom affordances for 6:331–332
- literacy 11:45–46
- and STEM 11:100–101
- for STEM 11:29–30
- citizenship STEM or cultivation of professionals 11:29–30
 - notion of learning for sake of learning 11:29
 - STEM and relationship with STEM disciplines 11:30
- in STEM education 11:20–24
- implications 11:24–25
 - relationship between mathematics and STEM 11:19–20
 - theoretical frameworks 11:20
- teacher noticing in 5:325–326
- as thinking tool in STEM education 11:22–24
- matrix
- definitions and operations 14:592
 - rank and matrix decompositions 14:592
- maximizing loglikelihood function 14:798–799
- maximum likelihood (ML) 14:642, 14:808–809
- maximum likelihood estimation (MLE) 14:554–555, 14:578, 14:642–643, 14:795–798
- considerations 14:344
- hypothesis testing and confidence intervals for 14:799
- MCAR. *See* missing completely at random (MCAR)
- MCAT. *See* medical college admission test (MCAT)
- McLewis' positionality 12:62
- MCMC approach. *See* Markov chain Monte Carlo approach (MCMC approach)
- MD-TRACE model. *See* multiple-document task-based relevance assessment and content extraction model (MD-TRACE model)
- MDGs. *See* millennium development goals (MDGs)
- MDK-C framework. *See* multidimensional knowledge in text comprehension framework (MDK-C framework)
- MDL. *See* model of domain learning (MDL)
- MDM. *See* midday meal scheme (MDM)
- meaning attribution
- early attribution of 10:43–45
 - and written language 10:43–45
- meaning making 10:71–72
- meaninglessness, alienation and loss of culture and purpose (MAL) 3:17–18
- measurement 13:193
- design 8:316
 - item response times 14:322–323
- media 2:341–343
- documentation 10:792
 - literacy education frameworks 2:304–307
 - and technology 10:791–793
- media and information literacy (MIL) 10:590
- datafication beyond computing 10:592–593
 - definitions and epistemological diversity in global perspective 10:590–593
 - documentarization displacements 10:591–592
 - emergent themes and research areas 10:597–599
 - information troubles and new participation competences 10:597–598
 - innovative research postures and methodologies 10:599–600
 - institutionalization of 10:601
 - mediatization rebooted by increased emergence of participation 10:591
 - MIL and transliteracy for transformative education 10:600–601
 - platform society and new consumption and protection competences 10:597
 - rise of literacy discourses 10:593
 - sustainability challenges and data governance competences 10:598–599
 - transliteracy as frontier field 10:593–597
- media education
- adults as target population for disinformation media literacy 2:308

- assessing needs of marginalized and immigrant communities 2:307–308
 centering affective dimensions 2:307
 checklists, fact-checking & information evaluation frameworks 2:308
 context and significance 2:301–303
 engaging mindfulness techniques 2:308–309
 gaps & future directions 2:307–309
 history of 2:303–304
 teaching lateral reading and process thinking 2:308
 in UK 10:498
 mediation, learning and 6:37–38
 medical college admission test (MCAT) 14:453
 medical pathologization 2:181
 memo-writing practices 12:304
 memory
 in classroom 6:225–226
 cognitive neuroscience 6:226
 conceptualizing memory and its failures 6:222–223
 dissemination and implementation of the established science 6:227
 in everyday contexts 6:224–225
 factors make memory inaccurate 6:222–223
 false autobiographical memories 6:223
 future of memory science 6:226–227
 long-term autobiographical memory 6:224–225
 long-term retention of knowledge from school 6:225–226
 metaphors 6:222
 neuroscience explanations, seductive allure of 6:226
 for people 6:224
 processes 6:108–109
 science 6:226–227
 testing improves retention 6:225
 training 6:225
 varieties of human memory 6:220–222
 MENA. *See* Middle East and North Africa (MENA)
 mental disorders 9:436
 empowerment, patients' voices and critique 9:440
 establishment of special education 9:440–443
 mentoring 4:384
 in new digital teacher learning space 5:730–732
 student teachers in "time of cholera" in public Arab schools 5:732–734
 in teachers' professional development 5:145–147
 meritocracy racism 5:418
 message content 10:31
 meta-analysis 14:572, 14:843
 See also data analysis
 advanced meta-analysis methods 14:856
 Bayesian approaches 14:856
 data analysis 14:848–854
 data evaluation 14:845–848
 dependence in 14:856
 literature search 14:844–845
 problem formulation 14:843–844
 of studies on teacher and school evaluation 13:291–294
 criteria for including studies 13:292
 descriptive analysis 13:292–294
 methods 13:291–292
 selection of studies 13:291
 meta-inference 12:368–369
 meta-reflection and assessments of multimodal practices, engaging in 10:512–513
 meta-regression 14:852
 metacognition 6:263
 metacognitive behavior 6:330
 metacognitive knowledge 6:211
 metacognitive strategies 6:356
 metaevaluation
 African evaluation guidelines 13:267
 distinguishing metaevaluation from synthesis studies 13:265
 evaluation standards 13:267
 of HIV/AIDS program evaluations using African evaluation guidelines 13:272
 idiosyncratic criteria and connoisseurship in metaevaluation 13:268
 mechanisms to support teacher evaluation systems 13:290–291
 methods 13:268–270
 norms and standards for evaluation in United Nations systems 13:268
 practice 13:270–272
 program evaluation standards 13:266–267
 purposes 13:264–265
 quality standards for development evaluation 13:267–268
 questions and criteria 13:265–268
 recommended by Cook and Gruder 13:269–270
 recommended by Scriven 13:270
 recommended by Stufflebeam 13:269
 of technology literacy program in underserved communities 13:271
 methods to handle missing values 14:807–812
 metropolis-Hastings Robbins-Monro algorithm (MH-RM algorithm) 14:95
 Mexican children and schooling experience in Mexico 10:143
 Mexican motherland, false narratives of 10:142–144
 Mexico 10:812–814
 education in developing countries 1:208
 MFC ranking. *See* multidimensional forced-choice ranking (MFC ranking)
 MGT. *See* multivariate generalizability theory (MGT)
 MH. *See* mantel Haenszel (MH)
 MH model. *See* monotone homogeneity model (MH model)
 MH-RM algorithm. *See* metropolis-Hastings Robbins-Monro algorithm (MH-RM algorithm)
 micro level 4:130–131
 framework for system leadership 4:128–131
 microeconomic labor market theory 4:446
 midday meal scheme (MDM) 1:535–536
 Middle East and North Africa (MENA) 1:380
 rethinking university in MENA uncertain times 1:384–385
 middle leader 4:115
 in schools 4:115–119
 migrant children 1:662
 MIL. *See* media and information literacy (MIL)
 milieux 12:696–697
 milieu 12:697
 priority 12:696
 timing 12:696–697
 millennium development goals (MDGs) 1:533
 balwadis and anganwadis 1:535
 ECCE 1:534–535
 implementation challenges 1:536
 local response to MDGs in India 1:533–534
 MDM 1:535–536
 MIMIC model. *See* multiple indicators multiple causes model (MIMIC model)
 mindfulness
 analysis 6:650–651
 data extraction and synthesis 6:650
 evidence that mindfulness training improves learning 6:649
 heterogeneity 6:651
 method 6:649–651
 moderator analysis 6:652
 risk of bias 6:651–652
 search strategy and inclusion criteria 6:649
 training improve learning 6:648–649
 minoritized communities 2:246
 minority ethnic teachers work 5:130
 minority learners, diversity, and educational outcomes
 experiences in higher education 8:400
 negative and "color-blind" discourses on 8:399–400
 social relationships in higher education 8:400–403
 minority rights discourses 9:264–265
 minority teachers leave, potential explanations of why 5:128–129
 racial climate in schools 5:129
 school demographic characteristics 5:128
 school working conditions 5:129
 teacher characteristics 5:128
 Minzu knowledge
 in Chinese education 3:220–221
 identify Minzu knowledge 3:221–222
 Mongolian STEM teachers reflecting on use of 3:222–226
 MIRT. *See* multidimensional item response theory (MIRT)
 "mis-education" of Aotearoa

- “mis-education” of Aotearoa (*continued*)
 history curriculum reset 2023 3:4–5
 indigenous past never colonizers pre-history 3:1–2
 myths 3:2–5
 teaching Māori and iwi histories 3:5
 missing at random mechanism (MAR mechanism) 14:806
 missing completely at random (MCAR) 14:806
 missing data analysis 14:805–818
 missing not at random mechanism (MNAR mechanism) 14:806
 missing paradigm 4:13
 missing values and possible remedies, threats to proper handling of 14:812–815
 missing-data mechanisms 14:805–806
 remarks on definitions of 14:806
 uncertainty 14:813–815
 missingness mechanism, tests of 14:814
 mission schools 9:152–156
 federally run day schools 9:153–155
 key assimilation reports 9:152–153
 missionaries
 impact of 10:775
 impact of Christian missionaries 10:775
 impact of Islamic missionaries 10:775
 mix methods 12:696–697
 milieu 12:697
 priority 12:696
 timing 12:696–697
 mixed analysis 12:749
 findings 12:749–750
 mixed dataset 12:368
 mixed method grounded theory (MM-GTM) 12:558
 adaptability to complexity 12:560
 analytical logic of abduction 12:559–560
 compatibility with multiple paradigms 12:560
 design considerations 12:565
 directions for future research 12:565
 distinct contributions of each method 12:560
 gauging quality of 12:563
 implications for practice 12:565
 interactive designs with integration through embedding 12:562
 iterative exchange embedded in constant comparative method 12:559
 linking distinct qualitative and quantitative phases to develop and verify or elaborate grounded theory 12:561
 methodological compatibilities between 12:559–560
 methodological transparency 12:563
 purpose 12:558–559
 saturation 12:563
 two approaches to implementing 12:561–562
 utility and originality 12:563
 mixed methods 12:581–587, 12:608–614
 analysis 12:500
 integration 12:501
 using software for data-based integration 12:503–507
 using software for results-based integration 12:501–503
 using software for sequential integration 12:508–509
 using software for transformation-based integration 12:507–508
 strategies for integration and required software features 12:510
 challenges faced in mixed methods approaches 12:707–708
 cultural studies in education, philosophical foundations for 12:418
 in culturally tailored interventions
 conceptual 12:678–679
 foundations 12:678–679
 methodological 12:679
 social-cultural-political 12:679
 data analysis 12:491–492
 for data collection 12:693–701
 Delphi 12:610–611
 design decisions applied to NSLES 12:670–672
 in education RCT
 efficacy and effectiveness trials 12:605–606
 evaluation type and design 12:600
 terminology 12:600–601
 in education survey research, approaches to using 12:668
 and evaluation
 constructivism and 12:533–534
 designing mixed methods studies 12:532–536
 distinction between evaluation and research 12:531
 in educational evaluation 12:531
 indigenous 12:536
 paradigm wars and current landscape of philosophical frameworks in evaluation 12:531–532
 postpositivism and 12:532–533
 pragmatism and 12:534
 transformative 12:534–535
 with existing data 12:636–644
 advantages and disadvantages of using 12:642
 existing data in mixed methods research 12:642–643
 existing qualitative data in mixed methods 12:639–640
 existing quantitative data in mixed methods 12:639
 gauge qualitative “depth” of topic, then proceed quantitatively 12:640
 gauge quantitative “breadth” of topic, proceed qualitatively 12:639
 knowledge-level continuum in mixed methods 12:640–641
 grounded theory and 12:117–118
 mixed methods-grounded theory exemplars 12:118
 insights for mixed methods researcher 12:417–425
 with phenomenology, intersecting
 exploring intersections with phenomenological approaches 12:539–540
 exploring philosophical compatibility with phenomenological approach 12:540–542
 mixing descriptive phenomenology 12:540–541
 mixing interpretive phenomenology 12:542
 phenomenological method 12:539
 strategies for mixing within MMPR studies 12:543–546
 rigor 12:612
 sampling 12:486–487
 survey design options 12:672–674
 in survey research 12:666–677
 survey research 12:666–667
 using NSLES example 12:668
 surveys and understanding individual respondents 12:667
 with translational research 12:588–598
 mixed methods action research (MMAR) 12:568
 aligning information with the phases in MMAR framework 12:576
 applying MMAR framework to design studies that intersect MMR with AR and CBPR 12:576–579
 designing reconnaissance and/or evaluation phases of the MMAR study 12:578
 example of MMAR process with MMR in both reconnaissance and evaluation phases 12:576
 example of MMAR process with MMR in evaluation phase 12:573
 example of MMAR process with MMR in reconnaissance phase 12:573
 identifying next steps in MMAR study process 12:578–579
 MMAR framework 12:572
 MMAR study process model 12:573
 orienting to MMAR study process and building community research capacity 12:578
 understanding problem and community needs 12:576
 mixed methods research (MMR) 12:362, 12:383, 12:418, 12:462, 12:478, 12:491, 12:538–540, 12:568, 12:600–601, 12:679, 12:729
 A-paradigmatic combination 12:464
 advantages and limitations of using QGIS as 12:650–652
 illustrating four mixed methods research design decision transformations 12:617–624
 implications for complexity-informed mixed methods research in education design practices 12:624
 situating complexity science in education and mixed methods research 12:616
 advantages of intersecting MMR with AR and CBPR approaches in education 12:570–571

- applications of mixed methods case study in education 12:546–547
- applying three frameworks 12:542
- categories of 12:427
- challenges and implications for future research 12:497–498
- challenges from tensions between 12:708–709
- challenges introduced by the inherent complexity in ethical reasoning within 12:523–524
- challenges of intersecting MMR with AR and CBPR approaches in education 12:571
- challenges to using and integrating 12:707–709
- combinations without integration of worldviews 12:464–465
- combined MMR designs 12:431
- common features of 12:569–570
- common worldviews and combinations 12:463–465
- communicate integration reports fairly recognize contributions 12:528
- complementary strengths combination 12:465
- comprehensive representation of integrated findings 12:526
- comprehensiveness of typologies 12:426
- conceptual characteristics 12:569
- conceptual frameworks describing 12:391–394
- convergence 12:468–471
- convergence and divergence in 12:466–475
 - in synthesis of QUAL and QUAN evidence 12:467–468
 - of scientific worldviews 12:463–466
- convergent MMR designs 12:427
- core MMR designs 12:427
- cross-cultural study in education using 12:421–422
- cultural studies in education and 12:418–419
- definition and examples 12:472
- definition and utility of conceptual frameworks 12:391
- design integration procedures affirm free and informed consent 12:526–527
- dialectical pluralism combination 12:465
- divergence 12:472–475
- in education
 - axiology 12:382
 - elements of philosophical paradigm 12:381–383
 - epistemology 12:382
 - framing mixed methods research and philosophical paradigms 12:383
 - indigenous philosophies 12:387
 - methodology 12:382–383
 - ontology 12:382
 - philosophical paradigm 12:380–381
 - philosophical paradigms in 12:383–387
 - yinyang philosophy 12:387
- ethical dilemmas arising from integration in evolving designs 12:525–526
- examining cases within mixed methods study 12:540–541
- explanatory sequential MMR designs 12:428–429
- fixed and emergent designs 12:426
- foundations of
 - data-analysis strategies in 12:376
 - data-collection strategies in 12:376
 - design of 12:374–376
 - integrated inferences in 12:376–377
 - purposes and utility of 12:373
 - research questions in 12:374
 - transferability of inferences in 12:377
- frameworks for combining 12:540–542
- frameworks
 - to articulate one's positionality for mixed methods research 12:398
 - describing process of mixed methods research study 12:392–393
 - describing prototypical mixed methods research designs 12:393
 - describing structure of field of mixed methods research 12:394
 - describing validity in mixed methods research 12:393–394
 - for field of mixed methods research 12:395–400
 - to organize literature about mixed methods 12:396
 - to plan teaching about mixed methods research 12:399–400
 - to understand published mixed methods study 12:396–397
- illustrative examples of cultural studies in education and 12:419–423
- implement integration strategies
 - authentically acknowledge multifaceted findings 12:527–528
- and integration 12:419
- integration challenge in 12:491
- integration in 12:611–612
 - case-based analysis 12:413–414
 - characterizing mixed priority research 12:415
 - design features that promote integration 12:414–415
 - distinguishing between mixing and integration 12:412
 - estimates of frequency of meaningful integration in 12:412
 - integrated mixed method research 12:411
 - integrating through joint displays 12:414
 - integrating through methodological or conceptual framework 12:414
 - multi-dimensional way 12:413
 - navigating aftershock of paradigm wars 12:412
 - philosophical grounding that supports interactive and iterative approach to integration 12:413
 - procedures that promote integration 12:413–415
- integration strategies for 12:453–460
 - analyze subgroups 12:459
 - develop range of data sets 12:455–456
- incorporate perspectives of all relevant stakeholders 12:454–455
- integrate claims step by step 12:456–457
- integrate qualitative and quantitative data into single data set 12:457–458
- integration of claims 12:451–453
- integration point to reflect on next step 12:457
- plan and re-plan the study 12:459–460
- revisit and re-analyze the data 12:459
- search for contradictions 12:459
- types of integration 12:457
- integrative mixed methods analysis 12:492–496
- intercultural study in education using 12:420–421
- intersecting two approaches to form mixed methods case study 12:542
- key decisions when combining 12:543–546
- levels of intersection 12:571–572
- literature reviews of qualitative, quantitative and 12:466
- longitudinal MMR designs 12:427–428
- methodological considerations of intersecting MMR with AR and CBPR 12:571–576
- methodological implications 12:480–483
- mixed methods data analysis 12:491–492
- mixed methods design and timing 12:544–545
- mixed methods study design, visual displays and communicating 12:764–768
- mixing methods within case study 12:541–542
- mono methodology mixed methods designs 12:431
- multicultural study in education using 12:419–420
- multiple worldviews combination 12:465
- philosophical characteristics 12:569–570
- prioritizing strands 12:427
- problem 12:463
- procedural characteristics 12:570
- proposed solution 12:465–466
- purposeful integration of quantitative and qualitative approaches in 12:423
- qualitative research in conceptualizing mixed method study 12:479–480
- rationale for focus and description of team approach 12:523
- real-life example of development of integrated claim in 12:368–369
- reasoned expectations of participation burdens 12:525
- reasoned reporting of authorship contributions 12:526
- scaffolded MMR designs 12:429
- sequential MMR designs 12:428
- single worldview combination 12:464–465
- socio-ecological framework for 12:395–396
- strategies for handling divergence 12:472–475

- mixed methods research (MMR)
(*continued*)
- studies
 - data collection in 12:487–489
 - sampling in 12:485–487
 - types of sampling schemes 12:486–487
 - substantive theory combination 12:464
 - synthesis designs in 12:466–467
 - synthesis
 - to conduct systematic literature review 12:656–657
 - in educational research 12:660–662
 - need for systematic literature reviews 12:656
 - pitfall for 12:657–658
 - step-by-step guide for 12:658–659
 - theoretical implications of privileging qualitative methods 12:478–479
 - three promising practices for addressing ethical dilemmas in integration 12:526–528
 - use of integration strategies to enhance case interpretation 12:545–546
 - value of Delphi as part of 12:611–612
 - variations in naming 12:428
 - visual displays in 12:763
- mixed methods research proposal (MMRP)
12:729–731
- abstract 12:735–736
 - background 12:731
 - budget 12:735
 - components of 12:730
 - conceptual framework/inquiry worldview 12:731
 - data analysis strategies 12:733–734
 - data collection procedures 12:733
 - data integration 12:734
 - data management 12:736–737
 - data sources 12:733
 - data validation and reliability 12:734
 - extenuating circumstances 12:737
 - institutional Review Board documents 12:734
 - issues to consider developing 12:735–737
 - literature review 12:731–732
 - methods 12:732–734
 - mixed methods designs 12:732
 - mixed methods procedural diagram 12:732–733
 - participants/context/sampling 12:733
 - questionnaires/interview questions 12:735
 - references and appendices 12:734–735
 - research aims 12:731
 - research questions/hypotheses 12:732
 - technology usage 12:737
 - timeline 12:735
 - title 12:735
- mixed methods social network analysis (MMSNA) 12:685
- critique and integration of qualitative and quantitative methods 12:687
 - defining 12:685–686
 - increasing adoption of quantitative social network analysis 12:687
 - methodological history of 12:687
 - MMSNA in educational research 12:687–690
 - qualitative beginnings 12:687
- mixed online and hybrid studies, analyzing 12:697–700
- ML. *See* machine learning (ML); mathematical literacy (ML); maximum likelihood (ML)
- ML-PBL. *See* Multiple Literacies in Project-Based Learning (ML-PBL)
- MLE. *See* maximum likelihood estimation (MLE)
- MM-GTM. *See* mixed method grounded theory (MM-GTM)
- MMAR. *See* mixed methods action research (MMAR)
- MMLE. *See* marginal maximum likelihood estimation (MMLE)
- MMR. *See* mixed methods research (MMR)
- MMRP. *See* mixed methods research proposal (MMRP)
- MMSNA. *See* mixed methods social network analysis (MMSNA)
- MMT theory. *See* multiple and multi-dimensional transitions theory (MMT theory)
- MMT-GM. *See* Multiple and Multi-dimensional Transitions Growth Mindset (MMT-GM)
- MMT-HOH. *See* Multiple and Multi-dimensional Transitions Habits of Heart (MMT-HOH)
- MMT-HOM. *See* Multiple and Multi-dimensional Transitions Habits of Mind (MMT-HOM)
- MNAR mechanism. *See* missing not at random mechanism (MNAR mechanism)
- Mo'olelo as methodology 3:232
- mobile learning in science education 11:216–217
- accessibility and affordability 11:216–217
 - compatibility 11:217
 - multifunctionality 11:217
 - portability 11:216
- mobility for complete study program 1:242
- mobilizing prefigurative hybrid literacies for decolonial resistance 10:809–816
- modal model 6:108
- mode of inquiry, Niu frontiers 3:163
- critical challenge from past 3:164–167
 - critical guiding questions for exploring 3:163
 - positioning myself 3:163
- model comparison, Bayesian approaches 14:781–782
- marginal likelihood 14:781–782
 - out-of-sample prediction 14:781
- model comparisons 14:799
- model fit in factor analysis and structural equation modeling 14:261–265
- identifying source of misfit and model modification 14:262
- model fit with discrete observed data 14:266–268
- obtaining accurate p-values by pooling cells a priori 14:266
 - overall goodness of fit statistics and goodness of fit indices 14:267
 - piecewise assessment of fit and model modification 14:268
- model of domain learning (MDL) 6:447–449
- and learning across lifespan 6:443–449
 - and traditional models of expertise 6:448–449
- model selection 14:260–261
- indices 14:268
- model tracing 14:462
- model training in deep learning 14:688–689
- model-data fit evaluation 14:690
- absolute fit 14:280–284
 - applications 14:276
 - aspects 14:276
 - Bayes factors 14:284–285
 - educational measurement example 14:280–281
 - information criterion 14:285–287
 - item response theory 14:273
 - person-fit analysis 14:274–276
 - relative fit 14:284–288
- modeling collaborative problem-solving data 13:252–253
- modeling item response times 14:323–327
- estimation 14:326
 - illustration 14:327–329
 - joint models for item responses and response times 14:325–326
 - model fit 14:326–327
 - separate models for response times 14:324
- modern schooling
- curriculum ideologies and different views of knowledge in 7:258–260
- modern test theory 14:7–10
- modern university, implantation and development of 8:22–25
- modifiers 13:194–195
- criterion-referenced/norm-referenced 13:194
 - formative/summative 13:194
 - high stakes/low stakes 13:194
 - large-scale 13:195
 - minimum competency 13:195
 - standards-based 13:195
- Mokken Scale analysis (MSA) 14:138–144
- framework 14:140–143
 - models 14:140–143
 - software for 14:143
 - variations on 14:143
- Mongolian STEM teachers reflecting on use of Minzu knowledges in teaching 3:222–226
- mono methodology mixed methods designs 12:431–432
- monolingualism, normalization of 10:124–125
- monotone homogeneity model (MH model) 14:140–142

- Monte Carlo sampling from posterior 14:779–780
- Monte Carlo simulation method 14:763–764
- MOOC. *See* massive open online courses (MOOC)
- moral leadership model 4:47–48, 4:64–65
- Morgan and colleagues, responses to 9:500–501
- morphology 6:467–468
- Mortimer communicative approach 6:603–604
- motivation 11:68, 6:357, 6:612
peer relationships and 6:290
teacher–student relationships and 6:289
- Movement 4 Black Lives (M4BL) 2:352
- MSA. *See* Mokken Scale analysis (MSA)
- multi-data sources in networked systems leadership 4:532–533
- multi-dimensional profiles 13:233
- multi-dimensional way, framing research problem in 12:413
- multi-mode networks and multilevel networks 4:532
- multimodal texts 10:509–512
concepts constituting responding to 10:504–506
fostering responding to and creating multimodal texts in classroom 10:507–512
remix in 10:510–511
responding to 10:508
digital videos 10:509
graphic novels and comics 10:509
images and illustrations 10:508–509
- multi-use pedagogical strategies 4:388–389
cohort and network structures 4:389
digital technology 4:388
needs analysis 4:388
university-district/school partnerships 4:389
- multi-user virtual environments (MUEs) 11:209
- multicultural education 2:23–24
narrative and experience in 7:394–395
- multiculturalism 10:71–72, 2:39
- multidimensional compensatory model (M2PL model) 14:79–80
- multidimensional forced-choice ranking (MFC ranking) 14:487
- multidimensional item response theory (MIRT) 14:86, 14:89–94
correlated-factors MIRT models 14:89–91
models 14:293
and scoring individuals 14:94–96
from unidimensional IRT models to 14:87–89
- multidimensional knowledge in text comprehension framework (MDK-C framework) 6:212
- multilateralism, negotiating politics of 1:521–523
- multilevel CFA 14:609–611
- multilevel distributed framework, educational leadership framed from 4:83–88
- multilingual children, misconceptions and radical hopes in literacy education of 10:139–140
- multilingual education
new trends in 2:190–191
traditional trends in 2:190
types of 2:189–190
- multilingual exposure and language development 6:472–473
- multilingualism
in language policy 10:124–128
hierarchies of value 10:126–127
promoting multilingualism 10:127–128
valuing and validating multilingualism 10:128
at school 2:188–189
- multilingualism 10:71–72
- multiliteracies
framework 10:573
new and emergent directions in multiliteracies language education 10:107–109
aesthetics and affect 10:108–109
and online mobility, creation, and networking, assumption of 10:489
for STEM learning 11:51–54
clarification of terms often used interchangeably or conflated with 11:52–54
history and principles of 11:51–52
and metalanguage to describe and interpret design elements and modes 11:52
multimodality 11:54
- multimedia representations 14:176
- multimethod research 12:362
- multimodal approaches to literacies and social communication, autism and 10:439–440
- multimodal composing processes
affect-driven 10:577–578
agentive identity expression 10:574
areas of exploration 10:578–579
collaborative multimodal composing 10:576
critical stance toward multimodal composing 10:579
developing as multimodal designers 10:575
integration into schools 10:578–579
mediated by composing tools 10:576
multifaceted view of 10:578
multiliteracies 10:573
multimodal composing-to-learn 10:575–576
multimodal meaning-making 10:578
multimodal methods 10:579
prominent theoretical frameworks in 10:572–573
recursive process and modal preferences 10:577
research on adolescents and 10:573–576
- research on multimodal composing processes 10:576–578
- reshaping classroom spaces 10:574–575
- social semiotics 10:573
- sociocultural theories of 10:572–573
- multimodal literacies and engineering design 11:54–59
Ameena and Emma begin to sketch bridge designs 11:56–57
Ameena and Emma gather materials and organize guides 11:59
context and participants 11:55
Emma and Ameena discuss and plan which materials to use 11:59
Emma and Ameena shift their planning to include the template 11:57–58
exploration of multiliteracies in engineering literacies club 11:55–59
- multimodal meaning 6:263
- multimodal response and production in time and space, engaging in 10:506–507
- multimodal social communication 10:439–440
literacies and 10:439
- multimodal social communication 10:443
- multimodal systems, digital games as 10:584–585
- multimodality 10:469
- multinomial distribution 14:721
- Multiple and Multi-dimensional Transitions Growth Mindset (MMT-GM) 6:540–541
- Multiple and Multi-dimensional Transitions Habits of Heart (MMT-HOH) 6:540
- Multiple and Multi-dimensional Transitions Habits of Mind (MMT-HOM) 6:540
- multiple and multi-dimensional transitions theory (MMT theory) 6:534–537
of higher stage medical trainees 6:536–537
successful 6:539–541
of university students identifying as LGBTQ+ 6:535–536
of young adults and significant others 6:535
- multiple imputation 14:809–810
- multiple indicators multiple causes model (MIMIC model) 14:256
- multiple intelligences 6:205–206
organizing broad intelligences 6:205–206
recognized broad intelligences of today 6:205
relations among broad intelligences 6:205
still newer intelligences 6:205
- multiple linear regression model 14:550–551
- Multiple Literacies in Project-Based Learning (ML-PBL) 11:119–120
- multiple methods, development of an integrated claim using 12:363–364
- multiple mini-interviews for teacher selection 5:335–336
- multiple texts 6:353

multiple texts (*continued*)
 measurement of multiple source
 comprehension 6:357–358
 models of learning from 6:354–356
 reading on internet 6:359–361
 role of individual differences 6:356–357
 teaching integration and sourcing skills in
 multiple text contexts 6:358–359
 multiple-document task-based relevance
 assessment and content extraction
 model (MD-TRACE model)
 6:354–355
 multistage MMR designs 12:430
 multivariate analysis of variance (MANOVA)
 14:543–544, 14:662
 multivariate ANCOVA 14:545–546. *See also*
 univariate ANCOVA
 assumptions and limitations 14:545
 data demonstration 14:536
 multivariate data, describing 14:729–730
 multivariate generalizability theory (MGT)
 14:63–64
 multivariate linear model 14:557
 multivariate $p \times i^2$ design 14:68–69
 museum education
 challenges and responses 2:321
 convergence 2:321–322
 and idea of building capacity 2:320
 museums as learning institutions
 2:320–321
 MUVes. *See* multi-user virtual environments
 (MUVes)
 Myran and Sutherland's synthesis of science
 of learning 4:14
 myths about special education 9:279–281

N

NAEP. *See* national assessment of
 educational progress (NAEP)
 Naïve Bayes classifiers 6:86
 native theory & praxis 3:73–74
 Namibia, early career academics in
 discussion of findings 8:517–518
 early career researchers in context
 8:513–514
 literature on early career researchers:
 experiences, challenges and
 prospects 8:514–515
 methodology 8:515
 Namibian ECRs' experiences, prospects and
 challenges 8:515–517
 Namibian ECRs' experiences, prospects and
 challenges 8:515–517
 Nancy Fraser's theory of justice 10:219–222
 recognition 10:221–222
 redistribution 10:220
 representation 10:222
 NAPALI. *See* national Pacific American
 Leadership institute (NAPALI)
 narrative analyses 12:309–310
 narrative assessment case study 9:356–359
 anna and research 9:357
 anna and Ria 9:358
 anna and Ria's parents 9:359
 anna and school management 9:359
 narrative hermeneutics 5:113–114
 enabling constraints 5:115
 learning to count what counts 5:115–116
 new precariate 5:117–118
 other peoples' children 5:114
 possible under circumstances 5:116–117
 reading archive of feelings 5:114
 for teachers and teaching 5:119
 narrative inquiry 12:27–28, 7:342
 design considerations 12:127–128
 engaging participants in 12:125–127
 justifications for 12:127–128
 negotiating being in field 12:126
 negotiating from field texts to interim
 research texts 12:126
 negotiating from field to field texts 12:126
 negotiating from interim research texts to
 final research texts 12:126–127
 negotiating living, telling, retelling and
 reliving stories of experience 12:125
 negotiating ongoing reflexivity 12:125
 negotiating working with response
 communities 12:127
 ontological and epistemological
 considerations in 12:124
 relational ethical considerations 12:129
 on research landscape 12:123–124
 thinking narratively 12:124–125
 touchstones of 12:128
 narrative skill development 6:471
 narrative spaces in teacher education
 5:411
 narrative structures and metalepsis
 10:628–629
 national 21st century skills
 assessment–projects 13:251–253
 National and Community Service Trust Act
 6:675–676
 national assessment of educational progress
 (NAEP) 14:348
 National Council on Teacher Quality
 (NCTQ) 6:298
 National Disability Insurance Scheme
 (NDIS) 9:446–447
 and education 9:447–448
 funding education 9:448–450
 national education systems 1:205
 national identity, education in South-East
 Asia 1:420
 national large-scale assessments
 14:348–350
 national Pacific American Leadership
 institute (NAPALI) 3:167–170
 national peer educator study 8:309–310
 national populism 5:742–745
 appealing to populist ideas and “evidence”
 to dismantle the profession 5:744
 education and nation-state 5:743
 influence on teacher education 5:743–744
 responses from educationists 5:744
 summary 5:745
 national research council (NRC) 14:146
 National Research Council, spatial thinking
 framework introduced by 11:77
 national school systems, increasing global
 interest in policy of 7:74–75
 national strategy for education research and
 enquiry (NSERE) 5:616–619
 National Students Financial Aid Scheme
 8:361
 native association, formation of 3:25
 native sovereignty 3:26–27
 natural acquisition processes 6:367
 natural language processing 14:462
 naturalist current 2:211–212
 naturalistic observation
 ethics, dramatis personae, locations, and
 scenes of action 12:255
 familiar and strange 12:255–257
 fieldwork as basis for validity,
 generalization, reliability, and
 objectivity 12:257
 joining 12:254
 observation 12:250–252
 orienting 12:252–253
 participant observation 12:253
 significance of naturalistic observation
 12:258–259
 truth out of place 12:258
 nature 12:274
 of models and their representation
 11:112–113
 nature-based participatory research with
 children 2:216–217
 NCLB Act. *See* No Child Left Behind Act
 (NCLB Act)
 NCTQ. *See* National Council on Teacher
 Quality (NCTQ)
 NDIS. *See* National Disability Insurance
 Scheme (NDIS)
 negative binomial distribution 14:720
 neocolonialism 1:683–685
 and education 1:686
 neoliberal dispositif
 choice and accountability–the prioritization
 of instrumental ethics 1:64–65
 discourse of numbers 1:67
 dispositive 1:62–63
 education market 1:64–67
 education-choice, commodification of 1:65
 neoliberalism 1:60–61
 neoliberalization 1:61–62
 neoliberalization of education 1:63–64
 self, commodification of 1:66–67
 students-accountability, commodification
 of 1:65–66
 neoliberal educational restructuring
 8:250–251
 ideological motives 8:252–253
 pursuit of profit 8:251–252
 from schooling in disaster capitalism to
 critical pedagogy 8:255–256
 still capitalizing on disaster 8:253–255
 neoliberal forces, push and pull of
 10:141–142
 neoliberal globalization
 consequences of compliance 1:81–82
 crisis of 1:80–82
 neoliberalism 1:60–61, 1:688, 3:70–73

- as justificatory social vision 1:80–81
- and redress 2:273–274
- neoliberalization 1:61–62
 - of education 1:63–64
- network level, group and whole 4:530–531
- network structures model
 - modeling selection and influence 14:438
 - multilevel extensions to social selection models 14:434–435
- network notation and terminology 14:430–432
- social influence models 14:436–438
- social selection models 14:432–436
- networked expertise 6:183
- networked learning 6:410–413
 - origins 6:411–412
 - research and development in 6:412–413
 - transition to social media and to postdigital world 6:413
- networked professional learning spanning social media 6:437–439
 - affordances and constraints of social media as spaces for networked professional learning 6:438
 - definitions 6:437
 - emerging issues in networked professional learning 6:438–439
 - understanding networked professional learning on social media 6:437–438
- networked systems leadership
 - framing networked systems leadership for change 4:531
 - insights from social network research related to leadership and change 4:527–531
 - moving toward networked systems leadership 4:525–526
 - promising developments of 4:532–533
 - real time, multi-data sources in 4:532–533
 - social network theory and analysis 4:526–527
 - social processes in shaping of leadership 4:526
- networking
 - multiliteracies and online mobility, creation, and networking 10:489
- networks 4:526, 7:203–204
 - and learning
 - analyzing online conversations 6:413
 - approaches to research and development 6:413–418
 - learning networks and networked learning 6:410–411
 - mapping participation 6:416–417
 - networked learning 6:411–413
 - orientation 6:409–411
 - personal and organisational perspectives 6:411
 - practical applications 6:418–419
 - social network analysis 6:414–416
 - taking network perspective 6:409–410
 - understanding participants' experiences 6:413
 - of literacy 10:304–305
 - of professionals in ITE for primary school teachers in Italy 5:77–78
 - neurobiology of reading 10:874–875
 - neurocognitive and physiological
 - measurement of STEM learning processes
 - applications of neurocognitive methods to educational research in STEM 11:458–462
 - broader considerations for application 11:468–469
 - EEG for STEM education research 11:462
 - examining sex differences with EEG methods 11:461–462
 - exploring individual differences and causality 11:464–465
 - exploring individual/group differences in STEM 11:460–462
 - fMRI for STEM education research 11:465
 - functional imaging approaches in STEM educational research 11:462–465
 - investigating cognitive strategies 11:459–460
 - investigating cognitive strategies 11:463
 - observing physiological differences in expertise 11:467–468
 - physiological approaches in STEM educational research 11:468
 - physiological measures of learning processes 11:465–468
 - physiology of emotion 11:465–466
 - primer on necessity for neuro and physiological tools 11:458
 - research approaches using electroencephalography 11:459–462
 - spatial reasoning map to EEG data 11:460–461
 - uncovering causal explanations in STEM 11:462
 - neuropsychology of reading, knowledge about 10:867–868
 - comprehension 10:870–872
 - decoding 10:869–870
 - narrow-gauge synthesis of 10:868–872
 - technical issues 10:867–868
 - neuroscience explanations, seductive allure of 6:226
 - new digital teacher learning space, mentoring in 5:730–732
 - new imperialism 1:193–204
 - new literacies 11:54
 - avenues for research 10:493
 - awareness of data practices 10:492–493
 - call to action 10:493–494
 - characteristics of internet environment 10:488–490
 - criticality and justice 10:491–492
 - digital equity 10:492
 - examining new literacies of online of reading and research 10:488
 - interconnectedness of literacies online 10:493
 - practical considerations in our lives and in everyday contexts 10:494
 - research centered on criticality and social change 10:493–494
 - research centered on digital identity and citizenship 10:494
 - self-criticality 10:492
 - shifting social values and opportunities for social change 10:490–491
 - teaching and learning 10:494
 - under-realized potentials 10:491–493
 - new literacy studies (NLS) 10:434–436, 11:52–53
 - inclusive goals linked to literacy research and education 10:435–436
 - inclusive literacy education 10:435
 - and literacy education 10:435
 - New London Group (NLG) 10:105
 - FL multiliteracies post NLG's manifesto 10:106–107
 - manifesto and literacy turn 10:105–106
 - New Materialism 2:317–318
 - New Public Management (NPM)
 - 13:306–308, 8:145, 8:152–153
 - contextualizing policy and broadening perspective 8:149–150
 - evaluation of NPM within HE sector 8:147–148
 - and higher education 8:146–150
 - maintaining benefits from NPM 8:149–150
 - to network governance 7:77–78
 - operationalization of NPM in HE 8:146–147
 - and rise within higher education 8:145–146
 - new public management (NPM)
 - 13:306–307
 - New York City (NYC) 5:676
 - New York integrated learning system metaevaluation 13:271
 - New Zealand accreditation 3:133–134
 - news literacy 2:306
 - newspaper reports, comparative reading of 10:118–119
 - Newton-Raphson method 14:562–564
 - Next Generation Science Standards (NGSS)
 - 11:117, 11:301, 11:444–445
 - challenge presented by 11:119
 - NextGen Scientist Survey 11:378
 - NGO
 - critical scholarship on 1:566–568
 - implications for NGO sector 1:569–570
 - and new aid architecture 1:568–570
 - NGOization 1:567
 - time of reckoning for NGO sector 1:570–571
 - NGSS. *See* Next Generation Science Standards (NGSS)
 - Nigeria
 - educational policies and English in 3:63–64
 - literacy education in 3:65–66
 - Nigerian academicians, impact of 10:777–778
 - Nigerian education
 - gestating and birthing literacy for life-long learning in 10:778–783

- Nigerian education (*continued*)
 literacy and Nigerian language ecology 10:774
 Nigerian indigenous education 10:771–772
 politics of literacy in Nigeria 10:774–778
 situating literacy in 10:773–774
- Niu frontiers
 conceptual underpinnings 3:161
 decolonizing pacific perspectives of place 3:168
 dissonance to slow, purposeful “discovery” 3:160–161
 mode of inquiry 3:163
 Oceania and taking oceanic perspective 3:161
 Oceania as relational space 3:168
 positioning oceanic identity 3:162–167
 scope of Niu conversation 3:160
 setting scene, providing context 3:162
 talanoa 3:164–165
 Talanoa as simply conversation 3:162
- NLG. *See* New London Group (NLG)
- NLS. *See* new literacy studies (NLS)
- No Child Left Behind Act (NCLB Act) 13:109, 13:166, 6:298
 and special populations 13:109–110
- node level network analysis 4:529–530
- nominal response model (NRM) 14:93–94
 multidimensional 14:93–94
- nominal response model 14:77–78
- non-Anglo-Saxon contexts 2:289–296
- non-formal education sector, HRE in 2:61–62
- non-formal environmental education 2:330–332
- non-hierarchical relationships with colleagues via distancing 5:45–46
- non-intending teachers, career motivations of 5:662–663
- non-mainstream literacies 10:201
- non-necessity of human ability-as-skill under life 3. 0 9:215–217
- non-random assignment of students 14:393–394
- non-school-based ILR programs 3:57–58
- non-state actors on teachers’ lives and work
 global education industry 5:4–7
 philanthropic foundations and teacher education 5:6–7
 private school chains and teachers’ work 5:5–6
 supranational organizations 5:1–4
- non-template-based AIG and example items 14:197–198
See also template-based AIG
 evaluation of 14:198–199
- non-uniform DIF 14:250–251
- non-western countries, leadership in 4:476–483
- noncognitive skills 14:485
- nondeclarative memory 6:221
- nonignorable missingness, models for 14:810–811
- nonlinear least squares 14:795
- nonlinear models, generalized linear and 14:571
- nonlinear regression models 14:558–567
 asymptotic statistical inference 14:560–561
 estimation methods 14:558–559
 further reading 14:566–567
 Gauss-Newton method for least squares estimation 14:559–561
 maximum likelihood estimation 14:560
 model building 14:564–566
 Newton-Raphson method 14:562–564
 overview of 14:558
 parameterization and curvature 14:561–562
 simulation data analysis 14:566
 variance heterogeneity 14:564
- nonlinearity of scan paths 10:417–418
- nonparametric confidence intervals 14:667
- nonparametric correlations 14:741–747
- nonparametric data exploration for single variables 14:662–663
- nonparametric methods 14:143–144
 frameworks for educational research 14:139–140
 for modeling item responses 14:138–144
- nonparametric regression
 analysis 14:664–665
 models 14:663–667
 advanced smoothers 14:665–666
 nonparametric regression analysis 14:664–665
 order-restricted nonparametric regression 14:666–667
 preliminary 14:663–664
- nonparametric statistical tests 14:658–662
 association and tests of independence for ordinal variables 14:660–661
 gauging uncertainties 14:667
 illustration using PISA data 14:667
- nonparametric data exploration for single variables 14:662–663
- nonparametric regression models 14:663–667
 testing hypotheses 14:658–660
- nonprobability sample, educational research 14:704–705
- Nordic countries, representations of inclusive education in 9:100
- Nordic disability 9:435–445
- Nordic model of education and development of inclusive education 9:98–100
 representations of inclusive education in Nordic countries 9:100
- Nordic representations of inclusive education
 Danish example 9:102
 intersection of educational agendas 9:101–102
- Nordic model of education and development of inclusive education 9:98–100
 from welfare states to competition states 9:101
- normal distribution 14:711
- normative discourse and challenging socialization 5:726
- normative eye movement patterns from reading 10:418
- North America, transfer credit 8:348
- Northeast Asia
 challenges for regionalization of higher education in 1:453–454
 economic interdependence in 1:448
 geopolitical tensions in 1:453
 growth & diversity of cross-border higher education in 1:448–450
 intraregional dependence among Northeast Asian countries 1:448–450
 regional leadership 1:453–454
 regionalization of higher education in 1:450–452
 US influence on Northeast Asian higher education 1:454
- North–South partnerships
 continued Northern domination 8:213–216
 decoloniality 8:214
 Global North–Global South 8:211
 imbalances in 8:213
 international higher education partnerships 8:210
 internationalization and transformation in South African higher education 8:211–212
 internationalization as recolonization 8:215–216
 new paradigm of 8:216–217
 North–South links in higher education 8:212–213
 southern critique of internationalization 8:214–215
- Norwegian researchers defending our language as language of research 1:377
- notation system 12:746
- novice teacher educators
 case of formal preparation focused on critical, justice-oriented aims 5:500–502
 course background and focus 5:500
 existing forms of learning and development for 5:498–500
 formal study 5:499
 implications of case 5:501–502
 learning and practice 5:501
 learning conceptual and practical tools 5:501
 personal account 5:497–498
 self-study and collaborative inquiry 5:498–499
 unplanned course learning 5:501
 varying emphases of professional learning for 5:499–500
- NPM. *See* New Public Management (NPM); new public management (NPM)
- NRC. *See* national research council (NRC)
- NRM. *See* nominal response model (NRM)
- NSERE. *See* national strategy for education research and enquiry (NSERE)

nStudy of learning process 13:57–59
 numerical variables 14:724
 NYC. *See* New York City (NYC)
 Nystrand's dialogically organized
 instruction 6:603

O

- observational classroom data 13:329–330
 OARS project. *See* oxford argumentation in
 religion and science project (OARS
 project)
 OBE. *See* outcomes-based education (OBE)
 observational learning 6:26, 6:612–614
 observed score equating methods
 14:241–243
 for equating scores in common population
 14:241–242
 observed score rater effects indices
 14:309–310
 observed sum scores 14:134
 occasional waves 5:197
 occupational enthusiasm *vs.* career burnout
 5:198
 Oceania
 and oceanic perspective 3:161
 as relational space 3:168
 odds ratio 14:136, 14:580
 OECD. *See* Organization for Economic Co-
 operation and Development
 (OECD)
 OER. *See* open educational resources (OER)
 off-topic spoken response detection 14:690
 Old Man Coyote's story 3:192–193
 older adults' development
 brain changes 6:504
 hearing 6:504
 informal learning 6:506–508
 men's sheds 6:507
 nonformal learning settings 6:508–509
 older adult learning and education 6:505
 older adults' informal learning through
 volunteering and play 6:507–508
 physical changes 6:503
 studies on
 cognitive training and improved
 cognition 6:505
 effects of cognitive and physical training
 on brain 6:505
 effect of physical training on cognition
 6:504–505
 vision 6:503–504
 ole talk
 analyzing and reporting ole talk 12:186
 application of 12:186
 Caribbean form of communication 12:183
 as oral research method 12:183–186
 in practice 12:184–186
 recording ole talk 12:184
 in research environment 12:183–184
 transcribing ole talk 12:184–186
 OLLIs. *See* Osher Lifelong Learning Institutes
 (OLLIs)
 OLS. *See* ordinary least squares (OLS)
 Oman
 assessment of structural model and
 mediating effect 4:506–507
 authentic leadership 4:503
 teacher engagement and teacher turnover
 intention 4:505
 data analysis 4:506
 data collection 4:505
 instruments 4:506
 method 4:505–507
 preliminary analyses 4:506–507
 school leadership in 4:502–503
 teacher engagement 4:503–504
 TTI 4:504–505
 online communication 10:439–440
 online education 1:149
 challenges in transition to 5:705
 culturally relevant online education 1:152
 development of 1:149–150
 finance issues associated with 1:154
 frameworks 1:151–152
 as response to COVID-19 pandemic
 1:150–151
 student engagement in 1:152–154
 online information, critical evaluation of
 10:557
 online learning
 blended learning 6:426
 COVID-19 pandemic and online learning
 enrollment growth 6:424–425
 development and growth of 6:423–424
 emergency remote instruction 6:426
 future of online learning research 6:428
 MOOCS 6:426–427
 pedagogy of 6:425
 research and effectiveness of online and
 blended learning 6:427–428
 online mobility
 multiliteracies and online mobility,
 creation, and networking 10:489
 online multiplayer games 10:442–443
 online of reading and research, literacies of
 10:488
 online professional development
 5:552–554
 effectiveness of 5:553–554
 forms of 5:553
 Online Research and Comprehension
 Assessment (ORCA) 6:358
 online skills
 inequality of digital use and 1:158
 online teacher education 5:705–707
 factors influencing impact in 5:704–705
 practices influencing impact in 5:701–703
 online teaching and learning practices in
 teacher education
 accessibility 5:701–702
 characteristics of 5:701–702
 driving impact in teacher education in era
 pre-covid 5:700–701
 and factors influencing impact in teacher
 education in era post-covid
 5:701–705
 first study 5:699
 methods 5:699–700
 second study 5:699–700
 use of appropriate digital tools 5:701
 online text, skills, strategies, dispositions
 and social practices for reading
 10:553–555
 online-merge-offline learning (OMO
 learning)
 case study 1 8:297–298
 case study 2 8:298
 case study 3 8:298–299
 case study 4 8:299
 case study collection 8:297–299
 education in post-Covid-19 era 8:296–297
 functions of technology in 8:300–301
 method 8:299
 research motivations and objectives 8:297
 results 8:299–301
 spaces are designed 8:301
 types of learning strategies have been
 implemented in 8:299
 Ontario, system leader and system
 leadership 4:132–133
 onto-epistemological qualitative language
 and literacy research 10:895–901
 ontological evolution, of grounded theory
 12:118–119
 ontological imperative
 digital methods in qualitative research in
 online spaces 12:293
 empirical demonstration 12:295–299
 researcher in web-based qualitative research
 methods 12:299–300
 in web-based qualitative research methods
 12:294
 ontology 12:382
 open access 8:63–64
 arXiv and after 8:65–66
 faculty responses to 8:66–68
 global gap 8:68–69
 publishing 8:69–71
 transformation by declaration 8:64–65
 open educational resources (OER) 6:262
 openness 4:260–261
 extend discretion and shared decision
 making 4:260–261
 operant conditioning 6:17–19
 aversive control 6:17
 choice, matching and self-control 6:18–19
 conditioned reinforcement, chains, and
 delay-of-reinforcement gradient
 6:18
 extinction and reinforcement schedules
 6:17
 shaping 6:17
 stimulus control 6:18
 operant creativity 6:136–137
 operant theory 6:136–137
 operational testing system for processing
 examinations 14:444
 operational values in methods and
 procedures 12:10–11
 optimal learning in classroom 6:369–370
 optimism 4:79
 optimizing learning through retrieval
 practice and spacing 6:664–672
 oral language development 10:359

oral traditions
 in education research 12:180–183
 significance of oral traditions in education 12:180
 storytelling 12:181–183
 orality 10:730
 ORCA. *See* Online Research and Comprehension Assessment (ORCA)
 order-restricted nonparametric regression 14:666–667
 ordinary least squares (OLS) 14:552–554
 geometric interpretation of 14:553–554
 standard errors, confidence intervals and hypotheses testing of β 14:554
 Organization for Economic Co-operation and Development (OECD) 1:327, 13:267–268, 13:308–309
 OECD's PISA 13:308–311
 high performing and high equity 1:328
 organizational citizenship behavior 4:516
 organizational improvement knowledge 4:78
 organizational learning 4:287–289
 organizational legitimacy 1:472–473
 organizational question, inclusion as 9:392–393
 organizational routines 4:238
 organizing schools for improvement
 conceptual framework 4:204–205
 empirical evidence 4:204–205
 international trends in school decision making 4:205–206
 leadership style and distributed leadership 4:204
 need for international perspective 4:205
 power-sharing and trust in schools 4:204
 recommendations 4:207
 recommendations for organizing decision making for school improvement 4:207
 orientations 6:330–331
 origins of emotions 6:152–153
 appraisal antecedents of achievement emotions 6:152–153
 in context of education 6:152–153
 general approaches to 6:152
 social-cognitive antecedents of emotional appraisals 6:153
 Osher Lifelong Learning Institutes (OLLI)s 6:508–509
 demographics, participation and participants' interests at U3As and 6:509
 outcome expectations, symbolic processes 6:28
 outcomes-based education (OBE) 8:51–52
 critique of 8:51–52
 outlier sampling 12:156–157
 overlapping cluster solution 14:630–631
 oxford argumentation in religion and science project (OARS project) 5:258

P

Pacific education systems, foreign development aid and 1:194–195
 Pacific island countries, shifting priorities of foreign aid to 1:193–204
 Pacific Island Literacy and Numeracy Assessment (PILNA) 14:351
 Pacific Islands' development and education, framing foreign aid to 1:195–196
 Pacific people in Aotearoa New Zealand 8:400
 Pacific step-up 1:200
 PAD. *See* principled assessment design (PAD)
 PAL. *See* principal authentic leadership (PAL)
 palimpsest 10:713–714
 pan African university 8:207–208
 pandemic 1:584, 5:745–746
 challenges to delivering tertiary education during 9:127–128
 uses of technology in navigating 8:83–90
 PAR. *See* participatory action research (PAR)
 paradigm shifts 10:890
 paradigm wars
 navigating aftershock of 12:412
 parameter estimation via Gibbs sampling 14:802–803
 parameter learning, Bayesian networks 14:126–127
 parametric correlations 14:735–739
 parametric regression analyses 14:663–664
 paratextual features 10:627–628
 parent stories 10:191–192
 parental engagement 9:94
 parental family structure 4:266–267
 parental involvement
 for children of different ages 4:268
 elements of 4:265
 including parental expectations 4:264–265
 on students 4:266–267
 parental style
 and importance of love and structure in home 4:265
 parental unemployment and children's educational outcomes
 educational performance 1:119–121
 grade repetition 1:121
 mechanisms and institutional context 1:124–125
 noneconomic mechanisms 1:125
 reduction of family income 1:124–125
 results of reviewed studies 1:119–124
 school dropout at secondary level 1:121–122
 tertiary education enrollment and completion 1:123–124
 upper secondary enrollment and completion 1:122
 partial credit model (PCM) 14:75–76, 14:332
 comparison of 14:338
 multidimensional extension of 14:336
 participant demographics 1:112–113
 participatory action research (PAR) 12:29–30, 12:52
 participatory culture specific intervention model (PCSIM) 12:680
 mixed methods in 12:681
 participatory photography 12:271–272
 carefree youth 12:281
 equality 12:277–278
 freedom and safety 12:279–280
 living in bubble 12:281
 methods and data 12:273
 nature 12:274
 as research method 12:272–281
 students' perceived privileges regarding living in Nordic country 12:273–281
 visual research and place in research in education 12:270–271
 welfare state privileges 12:275–276
 partitioning 14:628–629
 partnership 4:377–378
 approaches to facilitation of 8:272
 models of 5:383–385
 collaborative 5:383
 collaborative model 5:384
 community of practice model 5:384
 complementary model 5:383
 democratic professionalism 5:384–385
 equivalent model 5:384
 HEI-led 5:383
 research and innovation 5:384
 separatist model 5:383
 models of partnership facilitation 8:272–273
 national, institutional, disciplinary, and individual influences on approaches to 8:271
 values 8:270
 PASEC. *See* program for analysis of education systems (PASEC)
 passive mode strategies 6:694
 of cognitive engagement 6:690–691
 pattern-mixture models 14:811
 PBL. *See* problem-based learning (PBL); project based learning (PBL)
 PCA. *See* principal component analysis (PCA)
 PCM. *See* partial credit model (PCM)
 PCSIM. *See* participatory culture specific intervention model (PCSIM)
 PD. *See* professional development (PD)
 PDA. *See* policy discourse analysis (PDA)
 PDE. *See* principled design for efficacy (PDE)
 PDP. *See* Piano Didattico Personalizzato (PDP)
 PE. *See* performance expectations (PE)
 peace education in conflict zones
 adding methodological concerns to critical perspectives 1:100
 critical perspectives and more 1:99–100
 educational concerns to critical perspectives 1:99–100
 historical development 1:94–95
 peace education in Israel 1:96–99
 theoretical premises 1:95–96

- Pearson, Karl 14:3
 London school and development of
 classical test theory 14:3–4
 Pearson correlation coefficient 14:580–581
 Pearson product moment correlation
 method 14:735–739
 assumptions of 14:738
 computation of 14:736–737
 covariance 14:736
 hypothesis testing 14:737–738
 interpretations of 14:737
 visualizing relationships 14:738–739
 Pearson statistic 14:585
 pedagogic renewal in Quebec 7:12–13
 pedagogical module 6:638
 pedagogical phenomenology 12:139
 hermeneutic pedagogical phenomenology
 12:136–138
 pedagogical progressivism as educational
 movement 5:683–684
 pedagogical strategies
 for enhancing leaders' skills 4:383–384
 multi-use pedagogical strategies 4:388–389
 pedagogical translanguaging 2:192
 pedagogies 6:383–386
 and assessment via betrayal 5:44–45
 peer(s)
 assessment 13:79–83
 elementary school 13:81–82
 face-to-face 13:80
 high school 13:82
 higher education 13:82–83
 literature reviews 13:80–81
 online 13:80
 reliability 13:80
 theory of 13:80
 communicating norms and expectations
 6:291–292
 distinguishing one's self from 9:212–213
 instrumental aid 6:292
 interaction 11:67
 relationships and academic achievement
 6:290
 relationships and motivation 6:290
 relationships with 6:289–290
 rewards and reinforcements 6:294
 social and emotional support and safety
 6:293
 pentecostal students in church and school
 10:204
 literacy and identity of 10:204
 people
 of color 12:405
 fabricating kinds of 7:107–108
 percentage bend correlation 14:744
 performance assessment 13:97–98
 conceptual-theoretical and methodological
 foundation of 8:315
 framework 8:316–317
 in higher education 8:312–315
 in higher education, challenges of
 implementing and using
 8:320–321
 of student learning in higher education
 8:312–330
 performance expectations (PE) 11:445
 performance level descriptors (PLD)
 14:371–372
 cut score approval and final PLD revisions
 14:379
 performance tasks
 flexible 11:448–449
 testing and redesign of 11:449
 performance-based funding 8:188–189
 perhaps-cutting, exercise in 12:204–205
 permutation tests 14:769–770
 comments on permutation tests 14:770
 general notation and approach 14:770
 one sample or two paired samples
 14:769–770
 two independent samples 14:770
 person response function (PRF) 14:275
 person-fit analysis 14:274–276, 14:300
 comparative performance of 14:276
 Guttman errors 14:275
 lz statistic family 14:274
 personal epistemological practice
 6:386–388
 personal interests and approaches to
 participating in writing 10:29
 personal learning networks (PLNs) 6:418
 personalized teaching plan (PTP) 9:59
 Peru 9:130
 phenomenological theorizing and identity
 7:384–385
 phenomenology 12:8–9
 method 12:26–27
 qualitative research approaches
 12:310
 study
 Amadeo Giorgi 12:135–136
 Husserl, Heidegger & Merleau-Ponty
 12:131–133
 Max van Manen 12:136–138
 pedagogical phenomenology 12:139
 phenomenology and methodology
 12:134–138
 philosophy to educational research
 12:133–134
 Phi coefficient 14:741–742
 philanthrocapitalism 1:588–590
 philanthropic foundations 5:6–7
 Philippines 10:814–815
 context of imperialism and resistance
 10:814
 Ina Bai 10:815
 philosophical wealth 12:36–37
 philosophical-practical contradictions
 1:663–664
 philosophy-centric MMR designs 12:430
 phonology 6:467–468
 photography 10:80–81
 photovoice process 12:262, 12:267
 children and youth photovoice research
 examples 12:264
 children as co-collaborators in research
 12:266
 and data collection 12:264–265
 establishing research framework
 12:262–263
 history of 12:261
 image selection and theme generation
 12:263
 as participatory action research 12:262
 photography results 12:265–266
 presentation 12:263
 preserve preschool 12:264
 taking pictures 12:263, 12:267
 themes 12:266
 understanding green spaces 12:267–268
 and urban youth 12:267–268
 and young children 12:264–266
 youth as co-researchers 12:268
 youth environmental club 12:267
 PIAAC. *See* program for international
 assessment of adult competencies
 (PIAAC)
 Piaget's constructionist theory and creativity
 6:137–138
 Piano Didattico Personalizzato (PDP) 9:59
 amendments to 9:61
 PILNA. *See* Pacific Island Literacy and
 Numeracy Assessment (PILNA)
 pinpointing 14:378
 PIPM. *See* principal's information-
 processing mechanisms (PIPM)
 PIPM questionnaire 4:183
 analyzing external and internal information
 4:183
 disseminating information to teachers
 4:183
 receiving information from students,
 parents, community, and superiors
 4:183
 storing and retrieving information 4:183
 PIRLS. *See* progress in international reading
 literacy study (PIRLS)
 PISA. *See* program for international student
 assessment (PISA)
 PISA for Development program (PISA-D)
 14:351
 PISA-D. *See* PISA for Development program
 (PISA-D)
 place education
 addressing settler traditions of 3:230
 Centering Kanaka 'Ōiwi perspectives of
 3:229
 place-based education 11:273
 plastic medicine wheel 3:175–177
 platform society and new consumption and
 protection competences 10:597
 plausible values
 alternatives to 14:366
 analysis of 14:363–364
 generation of 14:362
 interpretation of 14:362–363
 PLC. *See* professional learning communities
 (PLC)
 PLD. *See* performance level descriptors
 (PLD)
 PLN. *See* professional learning networks
 (PLN)
 PLNs. *See* personal learning networks
 (PLNs)
 pluriversality 1:294–297
 poetry, telling stories through
 7:394

- point-biserial correlation 14:744
Poisson distribution 14:720–721
policy Delphi method 12:609
policy discourse analysis (PDA)
 9:261–262
policy drivers 4:473–474
 and areas for future research 4:474–475
policy dynamics 1:436–443
 external aid for education 1:443
 globalization and medium of instruction
 1:438–439
 globalization and public funding of
 education 1:441–443
 internationalization and ranking of
 institutions 1:439–441
 privatization of education 1:436–438
policy entrepreneurs 1:472
policy isomorphism
 alternative approaches 1:276–277
 boundary conditions for applying the
 analytical framework of 1:274
 in broad context of social sciences
 1:273–274
 conceptualization of 1:271–273
 confirmative views 1:275
 controversial views concerning 1:275–277
 criticisms 1:275–276
 isomorphic mechanisms in education
 policymaking in global education
 system 1:274–275
 national education policymaking in context
 of globalization 1:273–275
policy reform 13:120–122
policy regimes 1:557–561
policy streams 1:472
policy studies approach, feminist genealogy
 as 2:105–106
policy studies present futures 2:106–107
policy trends 4:471–473
 leadership frameworks 4:472–473
 principals prepared for job 4:472
 principals professionals and principals
 educators 4:471–472
political clashes between district and valued
 constituencies 4:229–230
political environment 4:162
political orientation
 influences 2:135–136
 political party orientation 2:137–138
politicalization of children's images
 10:664–665
politics and ethics 12:38–42
politics of global childhood 1:491
politics of researching race and dis/Ability
 and education 9:497–505
politics of scientometrics and datafication
 process 1:309–310
poly-SIBTEST 14:254–255
polychoric correlation 14:746
 coefficient 14:134–136
polysemy or ambiguity of meaning 6:348
polytomous response models 14:75
poor historical awareness 4:104
populism 1:648–649
 consequence of educational failure
 1:651–652
 democratic threat 1:650–651
 education learn from populism 1:653–655
 education struggle against populism
 1:652–653
 learnings from populism for educational
 practice 1:654–655
 learnings from populism for educational
 research 1:655
 learnings from populism for educational
 theory 1:653–654
 populists 1:649–650
populists 1:649–650
portfolio assessment 13:99
Portuguese education, public policies in
 assessment 9:398–399
 assessment for inclusion or exclusion
 9:402–403
 development of assessment
 9:400–401
 development of inclusive education policies
 9:399–400
 inclusion in education 9:397–398
Portuguese higher education
 changing state regulation 8:153–154
 higher education governance reform in
 8:154–155
 institutional governance and management
 structures 8:155–156
 reframing institutional leadership 8:156
positionality 12:60–63
 power of positionality in praxis 12:61–63
 in qualitative research 12:4–5
positioning theory 10:470
 and affect 10:923–925
 and corporeal dynamics 10:924–925
 and embodiment 10:922–923
 Foucault 10:919
 Harré's positioning theory 10:920–921
 Hollway 10:919
 perspectives on 10:918–919
 posthumanism and 10:921–922
 of self and others 10:924
positive deviance 12:157
positive organizational studies
 developing agency 4:515
 educational research 4:511–512
 exchanges 4:517
 and values 4:514
 importance of peers 4:518
 job crafting and expert instruction 4:516
 organizational citizenship behavior 4:516
 organizational outcomes of
 leader–member exchanges 4:518
 paying attention to what supports
 relationships 4:518
 perceptions affect relationships 4:514–515
 positive behaviours 4:513–519
 quality matters in leader–member
 exchanges 4:517
 relationship-building for school
 improvement 4:515
 remembering reciprocity 4:518
 school leader 4:519
 specifics about positive school leadership
 4:512–513
 time and timing matter 4:517
positive school leadership, specifics about
 4:512–513
positivism 12:38–39, 12:383–384
possibility, experiencing language of
 5:726–727
post-baccalaureate fellow 8:273
post-meeting survey 14:226
post-philosophies
 historic traditions and move toward
 10:940–941
 invitation to think and do inquiry
 10:944–945
 post qualitative inquiry is needed
 10:941–942
 produced for literacy studies 10:942–944
post-secondary disciplinary literacy
 10:452–453
post-standard setting meeting activities
 14:378–379
 cut score approval and final PLD revisions
 14:379
 technical report and ongoing validation
 14:379
 vertical articulation 14:378–379
post-stratification equating methods 14:242
post-unit assessment 11:451
postcolonial theory 10:611
postcolonialism in movies and higher
 education
 hybridity and change 8:246–248
 politics 8:244–246
 representations of other and voice
 8:241–244
posterior inference 14:132–133
Posterior predictive checking (PPC)
 14:780–781
posterior predictive model checking (PPMC)
 14:134–136, 14:280–284
 discrepancy measures for 14:282–283
 posterior predictive *p*-value 14:283–284
posthuman narrative analysis 9:203–204
posthumanism 10:610–611, 1:692–693
 datafication 1:695
 education, new geopolitics and 1:693–694
 education in anthropocene 1:695
 of literacy 10:304–305
 reassembling “context” 1:694–695
 and STS in contemporary education
 1:694–695
 theory
 and critical positioning 10:921–922
 Foucault 10:919
 Hollway 10:919
 perspectives on 10:918–919
postmodernism 10:890
 positions 12:386
postpositivism 12:383–384
 and mixed methods 12:532–533
poverty
 disability and 9:464–465
 measuring 9:462–464
 politics of 9:465–466

- in UK, lived experience of 9:466–467
 reduction strategy papers 1:212
 power 2:74–75
 analysis
 calculation for individual level RCTs 14:786
 power calculation for two-level CRTs 14:789
 power for single level RTs 14:785–786
 power for two-level CRTs 14:786–789
 and distributed leadership 4:102–103
 via literacy 10:212–213
 power-sharing and trust in schools 4:204
 power/knowledge 2:251–260
 PPC. *See* Posterior predictive checking (PPC)
 PPMC. *See* posterior predictive model checking (PPMC)
 PPP. *See* Presentation-Practice-Production (PPP)
 PPPs. *See* public private partnerships (PPPs); public-private partnerships (PPPs)
 PQP. *See* principals qualification program (PQP)
 practical domain 10:828–829
 practice
 curriculum 6:383–386
 of different professions, inclusion in 9:393–394
 pedagogies 6:385–386
 practice-based approaches to work and learning 2:240
 practitioner 4:364–365
 knowledge 8:105–106
 pragmatic approach in curriculum design 7:201–202
 pragmatism 12:385
 and mixed methods evaluation 12:534
 pragmatism-centric mixed methods research 12:431
 praxic current 2:215–217
 pre service teachers (PSTs) 7:155
 curriculum making, agency and curriculum literacy for 7:160–161
 pre unit assessment 11:451
 pre-digital reading instruction 10:531–532
 pre-meeting survey 14:224
 pre-phonological roots of literacy 10:39–40
 historical grounding 10:39–40
 initial research approaches 10:40
 neurobiological underpinning 10:39
 sensory motor roots 10:39
 pre-service STEM teacher education
 attributes of STEM teachers 11:272
 challenges of 11:270–271
 models of STEM pre-service programmes 11:270–271
 professional development models for 11:272
 situated learning in course design 11:272–273
 strategies of 11:272–274
 support from stakeholders 11:271, 11:274
 technology in STEM education 11:271
 video-based education 11:273–274
 way forward 11:272–274
 pre-service teachers (PST) 5:393, 5:613–614
 communities of practice 5:395–396
 data analysis 5:397
 descriptive overview 5:397
 epistemic negotiations 5:397–400
 findings 5:397–401
 identity development 5:401
 interconnections, feedback, and blending opportunities 5:398–399
 methodology 5:396–397
 moving back and forth 5:401
 on negotiations 5:399–400
 partnerships in teacher education 5:394–395
 partnerships in third space 5:394–395
 relational factors 5:400–401
 sample and data collection 5:396
 voices and experiences 5:396–401
 working together or working with 5:400–401
 pre-standard setting meeting activities 14:370–373
 data analysis 14:372–373
 PLDs 14:371–372
 preparing materials 14:373
 purposes and goals of 14:370
 selecting standard setting method 14:372
 stakeholder engagement 14:371
 precarity 10:810–811
 predictors of student well-being 6:724–726
 prefigurative hybrid literacies 10:810
 Premack principle 6:13
 Presentation-Practice-Production (PPP) 6:367
 PRF. *See* person response function (PRF)
 primary and secondary education 2:195–201
 primary school 9:449–450
 teachers in Italy 5:77–78
 transition into and out of 9:79–80
 principal authentic leadership (PAL) 4:506
 principal component analysis (PCA) 14:590, 14:592–595
 of correlation matrix 14:592–594
 of data matrix 14:594–595
 matrices and matrix notations 14:590–592
 principal practices in Thailand 4:496–497
 communication and conflict management 4:496
 decision-making and problem-solving 4:496
 developing school's culture 4:497
 leadership vision 4:496
 leading teaching and learning 4:496
 principal preparation
 analytic strategies 4:359
 commonalities, variations and trajectories in leadership preparation programs 4:363–366
 developing trajectories 4:359
 methodology 4:358–359
 program profiles, contexts and frameworks 4:359–363
 site selection 4:358–359
 sources and data collection 4:359
 principal supervisor differentiation strategies 4:331–332
 background 4:331
 developing partnerships 4:331–332
 summary of findings 4:331
 use of context and principal needs 4:332
 principal supervisor interactions with leadership teams 4:332–333
 background 4:332
 degree of interactions 4:333
 focus of interactions 4:333
 relations of principal supervisor interactions to ILT members' perceptions 4:333
 summary of findings 4:333
 principal supervisors coach principals to lead improvement
 communicate clear expectations 4:334
 data analysis 4:327–328
 data sources 4:326–327
 documents 4:327
 engagement in joint work 4:330–331
 establish thoughtful partnerships 4:334
 findings 4:328–333
 implications for practice and policy 4:333–334
 interviews 4:326–327
 methodology 4:326–328
 observations 4:327
 prioritize selection and preparation 4:334
 qualities of productive principal supervisor–principal partnership 4:328–330
 support learning of school-based instructional leadership teams 4:334
 surveys 4:327
 principal turnover
 economic benefits 4:449–450
 factors influencing principal turnover 4:447–451
 nonpecuniary working conditions 4:450–451
 principal characteristics 4:447
 school context 4:448–449
 theoretical framework 4:446–447
 working conditions 4:449–451
 principal-agent theory 1:471
 principal's information-processing mechanisms (PIPM) 4:181
 control variables 4:181
 measures 4:181
 method 4:180–182
 participants and procedures 4:180–181
 results 4:181–182
 teacher's intent to leave 4:181
 principals prepared for job 4:472
 principals professionals and educators 4:471–472
 principals qualification program (PQP) 4:359–360
 principals' work in public schools 4:305–307
 advances in information communication technology and social media 4:307

- principals' work in public schools
(*continued*)
- building and sustaining partnerships 4:306–307
 - changing nature of 4:307–310
 - covid-19 pandemic and occupational health and safety guidelines 4:308
 - dealing with personnel 4:306
 - engaging in new work demands 4:309
 - increased pace of work 4:309
 - increased volume of similar work 4:309
 - lead curriculum 4:305–306
 - maintaining school site 4:306
 - managing budget 4:306
 - no downtime 4:310
 - perceived limited resources 4:309–310
 - policies, programs and accountability initiatives 4:308
 - processing paperwork and forms 4:307
 - role 4:304
 - student wellness 4:308
 - time on tasks 4:305
 - work 4:305
 - work intensification 4:308–310
 - working long hours 4:308–309
- principled assessment design (PAD) 14:146
- BAS 14:147–153
 - going outside triangle 14:159–160
 - selected PAD approaches 14:153–159
- principled design for efficacy (PDE) 14:157–159
- print text 10:569
- print-centrism Dwarfs multimodal textuality 10:569
- prior distributions 14:802
- prior knowledge 6:209–210, 6:356
- prior specification 14:777
- objective Bayes 14:778
 - uncertainty quantification 14:778
- prioritize black, indigenous, and teacher candidates of color 5:423
- private education 1:662
- private primary schools 1:373
- privatization
- by design 1:177
 - of education
 - conceptual demarcation and relevant distinctions 1:174–175
 - definition and scope 1:174–175
 - effects of privatization 1:178–179
 - main drivers 1:175–177
 - neoliberal creed 1:177
 - privatization by design *vs. de facto* privatization 1:177
 - regulating independent schools 1:180
 - regulation and governance of private provision 1:179–180
 - regulation of PPPs in education 1:179–180
 - world trends 1:175
 - of public education
 - charter schools, vouchers, and 1:183–184
 - develop strong base using social media 1:189
 - parents opting children out of tests 1:188–189
 - privatization of education to
 - privatization in education 1:187–188
 - rise of testing and demise of social good 1:184–187
 - rise of unthinkable 1:190
 - probability 14:577
 - probability-proportional-to-size sampling 14:707
 - sampling designs 14:705–707
 - balanced sampling 14:705–706
 - cluster sampling 14:706
 - probability-proportional-to-size sampling 14:706–707
 - simple random sampling 14:705
 - standard errors of estimates from probability samples 14:707
 - stratified multistage sampling 14:707
 - stratified sampling 14:705
 - systematic sampling 14:707
 - probability distributions 14:752–753
 - probability sample, educational research 14:704–705
 - problem formulation of meta-analysis 14:843–844
 - examples of 14:844
 - inclusion and exclusion rules 14:844
 - theory and frameworks in problem development 14:843–844
 - problem solving 6:229–230
 - cognitive processes and types of knowledge in 6:230–231
 - distinction between productive and reproductive thinking 6:232
 - expertize 4:78
 - followed by instruction 6:617–618
 - nature of insight 6:232–233
 - problem space and search processes 6:233
 - problem-solving practices in Thailand 4:496
 - in realistic situations 6:233
 - rigidity in thinking 6:231
 - teaching of thinking skills 6:233–234
 - transfer 6:231–232
 - types of 6:230
 - problem-based learning (PBL) 4:383–384, 6:703–706
 - building on prior knowledge 6:704
 - design of learning task in 6:704–705
 - eliciting discussion 6:704–705
 - encouraging knowledge integration and transfer 6:705
 - level of student/teacher centeredness in 6:705–706
 - relevance for future profession 6:705
 - sequence of information acquisition in 6:706
 - stimulating self-directed learning 6:705
 - problem-solving current 2:212–213
 - procedural interests 10:29
 - procedural knowledge 6:211, 6:382
 - procedural literacy 10:585
 - process 11:70
 - data 14:11, 14:365
 - process-outcome links 6:594–595
 - teacher interventions 6:595–596
 - production rule 6:639
 - productive learning environments 6:331–333
 - productive learning networks 6:411
 - professional accountability
 - building professional ownership of accountability 4:297
 - collaborative and rigorous self-evaluation 4:299–300
 - inspection as support for 4:301
 - leadership for 4:297–298
 - locality partnerships as support for 4:300–301
 - schools as centers of professional learning 4:298–299
 - shift in culture and practice needed 4:296
 - system of 4:296–301
 - professional competence 5:198–199
 - professional development (PD) 5:469–470, 5:511, 5:550
 - artifacts of practice 5:551–552
 - co-design of instructional resources 5:552
 - content-agnostic professional development through equity lens 5:556–557
 - content-specific professional development through equity lens 5:556
 - facilitation of professional development 5:554–555
 - facilitator's role in supporting learning and community development 5:554
 - with focus on equity 5:555–557
 - frameworks for effective PD facilitation 5:555
 - practice-based 5:551–552
 - preparation of PD facilitators 5:554–555
 - relationship between program quality and 5:598–600
 - tensions and unresolved issues 5:554
 - professional education 5:369–370
 - professional identity of university academics
 - "becoming" good teacher 8:507–508
 - "being" good university teacher 8:505–506
 - constructivism and social practice 8:506
 - discussion 8:509–510
 - methods 8:506–507
 - personal dispositions and intrinsic factors 8:507
 - professional development 8:508–509
 - professional identity in education 8:504–505
 - results 8:507–510
 - professional learning 5:469–470, 5:528–530
 - data collection 5:532
 - distinctness and stringency in commission statement 5:535
 - of early career teachers, coaching for 5:144–152
 - findings 5:532–536
 - methods and methodology 5:530–532
 - middle leading as facilitating 5:533–534
 - middle leading as managing 5:533
 - middle leading at different levels 5:534
 - middle leading practice at two levels 5:536–537

- model enables and constrains 5:534–536
 number of levels/positions 5:535
 organizing for 5:533–534
 positive culture of change 5:535–536
 scaffolding 5:535
 shared spaces for communication 5:535
 site 5:530–531
 sociocultural perspectives on 5:386–387
 theory of practice architectures 5:531, 5:537
 trust 5:536
 understanding professional learning in
 workplace 5:471–472
 professional learning communities (PLC)
 13:275–276
 calculating VA and raw school performance
 13:282
 characteristics and processes of schools as
 13:277–278
 contrasting extent of PLCs in primary and
 secondary schools 13:281
 correlations between school performance
 and PLC factual characteristics
 13:283
 correlations between school performance
 and PLC indicators 13:282
 relationship between pupil outcomes and
 PLC processes and contexts
 13:281–282
 stages of PLC development 13:278–280
 summary indicators of 13:280–281
 three approaches to evaluate schools as
 13:276–277
 professional learning networks (PLN)
 5:275
 for teachers' development 5:275
 professional learning networks (PLN) 6:438
 professional literacies 10:468
 professional school, admissions tests at
 14:454
 professional teaching standards, global
 framework of 1:579
 professional unlearning opportunities
 2:278–279
 professionalism
 foster culture of 4:259–260
 new forms of 5:73–75
 professionals
 principals professionals 4:471–472
 profit corporations and education reform
 involved in field of education 1:591–592
 involvement of FPCs in education in terms
 of geopolitics and equity
 1:592–594
 role of FPCs in education being interpreted
 in education research 1:592
 program for analysis of education systems
 (PASEC) 14:350
 program for international assessment of
 adult competencies (PIAAC)
 14:349
 program for international student
 assessment (PISA) 1:328, 13:182,
 14:349, 14:658, 7:285–286
 age of enlightenment to 5:605–606
 illustration using PISA data 14:667

 OECD's PISA, high performing and high
 equity 1:328
 PISA study 7:285–286
 for schools 14:349–350
 program quality, relationship between
 professional development and
 5:598–600
 programmatic curriculum 7:260
 progress in international reading literacy
 study (PIRLS) 14:348
 project based learning (PBL) 10:451, 11:118,
 11:273, 11:445, 6:706–707
 design of learning task in 6:706
 design process 11:446–450
 level of student/teacher centeredness in
 6:707
 major findings 11:454–455
 methodology and analysis 11:450–454
 problem and objective 11:444–445
 promise of project-based learning for
 supporting student learning 11:118
 sequence of information acquisition in
 6:707
 significance 11:456
 significance of work 11:445
 theoretical framework 11:445–446
 prospective memory 6:221–222
 proxies server of learning process 13:56–57
 proximal self report 13:56
 PST. *See* pre-service teachers (PST)
 PSTs. *See* pre service teachers (PSTs)
 psychiatry 9:436–438
 psychological dispositions, leadership
 development 4:385–386
 emotional arousal 4:386
 mastery experiences 4:386
 verbal persuasion 4:386
 vicarious experiences 4:386
 psychological experimentation model
 10:842–843
 psychological resources
 4:79
 psychological theories
 4:39
 psychology, well-being in 6:721–722
 psychometric issues of learning process
 13:60–61
 psychometric models 14:420–421
 psychometric simulation 14:341–342
 analyzing results based on design and
 research question 14:345
 comparing true and estimated parameters
 14:344–345
 defining and justifying conditions 14:342
 estimating parameters 14:344
 replicating procedure specified number of
 times 14:345
 simulating data under specified conditions
 14:343–344
 specifying experimental design and
 outcomes of interest 14:343
 specifying research question 14:342
 PTP. *See* personalized teaching plan (PTP)
 public accountability systems, impact of
 4:294–296

 public arab schools, mentoring student
 teachers in “time of cholera” in
 5:732–734
 public computer-based process data
 repositories 14:419–420
 public diplomacy
 changing higher education landscape
 1:642–643
 cold war and 1:641–642
 educating elites 1:640–641
 and liberal world order 1:641
 power of educational diplomacy 1:643
 soft power as category of practice 1:644
 toward critical geopolitics 1:644–645
 public dissemination matters 3:215
 public education, privatization of
 1:170–171
 public funds made available to higher
 education institutions 8:186–187
 public good 3:81–83
 public private partnerships (PPPs) 1:568
 problematic promise of 1:568
 public schooling 6:298
 public services of libraries 8:124
 public-private partnerships (PPPs) 1:174
 regulation in education 1:179–180
 publication bias 14:853–854
 adjusting for 14:854
 detection of 14:853–854
 publications role in academic careers
 8:158–159
 publish or perish
 adverse effects of quantified view on 8:165
 role of bibliometrics in research assessment
 8:160–165
 role of publications in academic careers
 8:158–159
 sociological studies of academic reward
 system 8:159–160
 Puerto Rico 10:811–812
 punishment 6:17
 Pūrākau 12:182
- ## Q
- Q-matrix 6:638–639
 QA. *See* quality assurance (QA)
 QCA. *See* qualitative comparative analysis
 (QCA)
 QGIS. *See* qualitative geographic
 information systems (QGIS)
 QS 2:263–264
 QTR. *See* quality teaching rounds (QTR)
 qualitative analysis 12:748–749
 qualitative comparative analysis (QCA)
 12:581
 considerations for use 12:584–586
 development 12:581
 examples from field 12:582–584
 limitations 12:586
 method overview 12:581–582
 within mixed methods evaluation 12:582
 process for using 12:584
 qualitative content analysis 12:314–315
 basic techniques 12:316–319

- qualitative content analysis (*continued*)
 deductive category assignment 12:318–319
 different versions of 12:320–321
 for educational research 12:315
 inductive category formation 12:317–318
 methodological backgrounds 12:316
 short history of 12:315–316
 software support 12:319–320
- qualitative data
 collecting 12:364–365
 converting qualitative data into quantitative data 12:367–368
 into single data set 12:457–458
- qualitative evaluation 12:150–158
 comparison-focused 12:156
 connecting evaluation questions to uses and methods 12:153
 contribution of qualitative methods to evaluation 12:151–152
 different evaluation purposes 12:153
 evaluation through observation 12:152
 examples 12:151–152
 focus groups 12:152
 framing practical qualitative evaluation questions to enhance use 12:154–155
 impact inquiries involving qualitative fieldwork 12:153–154
 individual case examples 12:151
 key informants, key knowledgeable, and reputational sampling 12:157–158
 outlier sampling 12:156–157
 positive deviance 12:157
 purpose of 12:150–151
 purposeful sampling for 12:156–157
 sampling high-impact cases 12:156
 typical case sampling 12:157
 unanticipated consequences 12:154
 units of analysis 12:155
 utilization-focused 12:158
- qualitative findings 12:750
- qualitative geographic information systems (QGIS) 12:645–647
 advantages and limitations of using QGIS as mixed methods research approach 12:650–652
 application of 12:649–650
 as mixed methods approach in education 12:647–649
- qualitative inquiry 12:722–723
 data history in 12:199–200
- qualitative interviews 12:210
 analytic approaches to interviewing 12:216–217
 in conversational world 12:210–211
 different forms of 12:212–216
 history of 12:211–212
 individual and group interviews 12:214–215
 interviewing using different media 12:215–216
 structure 12:212–214
- qualitative literacy and language research 10:895–900
 composite stories 10:898
 counter narratives 10:897–898
 intersectionality 10:898–899
 personal stories 10:897
 stories of others 10:897–898
 translanguaging, transliteracies and transnational 10:899–900
- qualitative method (QUAL method) 12:462
 convergence and divergence in synthesis of 12:467–468
 critique and integration of 12:687
 Delphi 12:610–611
 handling divergence of 12:472–475
 and mixed methods studies 12:466
- qualitative research 12:55–57
 applying theory and philosophy in 12:7–11
 approaches 12:305–310
 analytic practices and strategies 12:302–305
 case studies 12:306–307
 coding practices 12:303–304
 ethnographies 12:307–308
 grounded theories 12:308–309
 inductive thematic analysis 12:304–305
 memo-writing practices 12:304
 narrative analyses 12:309–310
 phenomenologies 12:310
- authenticity 12:338
 axiological foundations of 12:8–9
 centering ethical considerations 12:338
 consistency 12:340
 credibility 12:341
 critical subjectivity 12:341
 determinants of quality in 12:338–348
 disrupting dominant ideologies in 12:56–57
 distinguishing ontology from epistemology 12:3–4
 in education 12:14–17
 generalizability and transferability 12:4
 historical roots of grand theories 12:5–6
 histories and traditions in 12:37
 implications for researchers in education 12:349
 ironic validity 12:346
 logics employed in 12:4
 meaningful coherence 12:341
 member checking 12:341–342
 memoing 12:342
 methodology, methods, and analysis 12:59–60
 paralogical validity 12:347
 philosophical assumptions of 12:3–5
 positionality 12:342
 and reflexivity 12:60–63
 after positivism 12:6–7
 power, positionality, and reflexivity in 12:4–5
 proposals 12:711
 revise and share 12:719
 writing 12:711–719
 qualitative sensibilities into mixed methods research 12:11
 queer theory and 12:78–84
 reciprocity 12:342–343
 reflexivity 12:343
 relational ethics 12:343–344
 resonance 12:344
 rhizomatic validity 12:346–347
 rich rigor 12:344
 sacredness 12:344–345
 salad bar approach to considering quality in 12:337–338
 sharing privilege 12:345
 significant contributions 12:345
 sincerity 12:345–346
 sites as arbiters of quality 12:346
 taxonomies of 12:6
 theory and theories in 12:5–7
 transformative paradigm 12:55–56
 transgressive validities 12:346–347
 triangulation 12:347
 trip to salad bar 12:348–349
 voice 12:347–348
 voluptuous validity 12:347
 worthy topic 12:348
- qualitative researchers, orientation among 12:17–19
- qualitative sampling and data 12:198–209
 alternatives to representational sampling strategies 12:203–204
 data as being with 12:201–202
 data history in qualitative inquiry 12:199–200
 data reconceptualized 12:200–201
 exercise in perhaps-cutting 12:204–205
 exercise in ritornello 12:205
 political of data and sampling 12:205–207
 shaping data and sampling futures 12:208
 and strategies in conventional humanistic qualitative research 12:202–203
- qualitative sampling and qualitative data 12:198–209
- qualitative sub-approaches as values-driven 12:9
- qualitative variation in learning, quantitative approaches to identifying 8:76–78
- quality assurance (QA) 1:423
- quality control
 empirical illustration 14:335–338
 IRTree models 14:333–334
 multidimensional extensions of ordinal IRT models 14:332–333
- quality in education 2:68–69
 peeling evaluation onion 5:50–51
- quality retention
 challenge 5:539–540
 developing teacher quality 5:544–546
 efficacious beliefs 5:544
 first, recruitment second 5:540–542
 purpose and values 5:542–543
 quality retention 5:540
 schools matter in developing and sustaining teacher quality 5:545–546
 social ecological approach to researching teacher quality 5:544–545
 teacher quality revisited 5:542–544
- quality teaching rounds (QTR) 5:236
- QUAN method. *See* quantitative method (QUAN method)

- quantitative analysis 12:749
- quantitative approaches to identifying qualitative variation in learning 8:76–78
- quantitative data
- collecting 12:364–365
 - into single data set 12:457–458
- quantitative findings 12:750
- quantitative method (QUAN method) 12:462
- convergence and divergence in synthesis of 12:467–468
 - critique and integration of 12:687
 - Delphi 12:610–611
 - handling divergence of 12:472–475
 - and mixed methods studies 12:466
- quantitative social network analysis
- increasing adoption of 12:687
- quantitative social network analysis, increasing adoption of 12:687
- quasi-experimental designs 4:542
- Quebec
- curriculum policies in 7:11–16
 - education
 - interdisciplinary conceptions and practices in 7:270–272
 - pedagogic renewal in 7:12–13
- queer affirming 2:360
- queer educators 2:182, 2:184
- queer in school 2:179–187
- queer students 2:184–185
- queer studies in education 9:165–166
- queer theory
- complexities of 12:79–80
 - queering qual 12:80–82
- questionnaire matrix sampling 14:365
- ## R
- race-evasive approaches to white student resistance in teacher education 2:284–285
- race-evasive curriculum in teacher education 2:284
- race-evasiveness in teacher education 2:283–285
- racial justice activism 2:367–368
- racial literacy in teacher education 2:285–286
- teacher education programs must actively shift ideological commitments 2:286
 - teacher education programs must actively support diversification 2:285–286
- racial stereotypes about mathematics and science 11:337–339
- racially minoritized individuals and knowledge economy 2:247–248
- racism 2:39–43, 3:110–111
- and “overwhelming presence of whiteness” in teacher education 5:475–477
 - cruel fallout of 10:144
 - investigate survival strategies of those with lived experiences of 5:421
 - counter-narratives 5:421
 - perspective-taking and building empathy 5:421
 - and members of dominant culture in examining racism is resisted 5:421–422
 - analysis of classrooms and schools 5:422
 - analyzing teaching materials 5:421
 - challenging racism in all forms through informed collective action 5:422
 - conceptual frameworks and access to new language 5:421–422
 - facilitating skill development in order to resist oppressive conditions 5:422
 - media literacy and analysis 5:421
 - participatory action research 5:422
 - peer mentorship 5:422
 - in teacher education 2:283–285
- racist discourses and ideologies 5:417–418
- radon home survey 11:228–229
- Ralámuli Nation’s literacies for resistance 10:813–814
- Ralámuli’s present context, historical roots of 10:812–813
- RAND reading comprehension heuristic 10:892
- random forest analysis 14:812
- random samples 14:763
- random-effects model 14:851–852
- random-effects-mean estimation 14:852
- randomized controlled trials (RCT) 12:599
- advantages and disadvantages of mixed methods with 12:601–602
- randomized designs 14:698–700
- randomized trials (RTs) 14:784
- model with covariate 14:786
 - model without covariates 14:785
 - power calculation for individual level 14:786
 - power for single level RTs 14:785–786
- range-finding of standard setting meeting 14:378
- rapid qualitative research 12:324–325
- breadth vs. depth 12:330
 - challenges of 12:330–331
 - ethics and governance 12:331
 - five key features of 12:325–330
 - funding 12:331
 - future of 12:331–332
 - new approaches to 12:332
 - overcoming assumptions 12:330
 - sampling and rapport 12:330–331
 - team-based fieldwork and analysis 12:331
 - theoretical grounding and reflexivity 12:330
- Rara’aba, leadership moment in 3:153–154
- Rasch measurement 14:8
- rater effects modeling 14:309–312
- common rater metrics and rater effects 14:308–309
 - in constructed response scores 14:308
 - factors in model selection 14:318–319
 - identifying aberrant raters 14:318
 - models with rater parameters 14:310–312
 - models without rater parameters 14:312
 - observed score rater effects indices 14:309–310
 - in teacher licensure performance assessment 14:312–318
- “rational” knowledge economy 1:281–282
- RCT. *See* randomized controlled trials (RCT)
- #Reconciliation 2:275–276
- re-envisioning family literacy for families of color 10:192
- re-examining
- beyond 9:288–289
 - continuation 9:288
 - history 9:285
 - needless 9:289–290
 - re-arranged 9:286–288
 - un-together 9:286
 - words 9:284–285
- re-thinking weaving 3:185
- idea 3:185–186
- re-weaving thinking 3:183
- claiming space to theorize art practice as methodology 3:183–184
 - theorists 3:183
- reader(s)
- as assembler 10:554
 - as builder 10:554
 - as composer 10:555
 - as fixer 10:554
 - global literature in lives of 10:634–635
 - as responder 10:555
 - response theories 10:607–608
 - social responsibility of 10:635–636
- reading 6:402
- additional neural factors relatable to 10:872–875
 - to be meaningful for struggling readers 10:403–405
 - both online and digital texts 10:553–555
 - competence in meaningful, engaged reading experiences 10:402–403
 - comprehending together 10:403
 - reading minds 10:403
 - reading strategically 10:402
 - components of reading skill 10:529–530
 - comprehension in troubled times 10:319–320
 - engaging multiple perspectives and positioning 10:323–324
 - evaluating text 10:326–327
 - historical moment 10:322–326
 - orientation toward action 10:325
 - pedagogy 10:323–327
 - practice 10:322–323, 10:325–326
 - practices needed to engage as text users/analysts and critics 10:322–327
 - unearthing ideologies in text 10:322–323
 - development 10:427
 - in international context 6:403
 - intervention class 10:565–566
 - narrow-gauge synthesis of neuroanatomy related to 10:868–872
 - neurobiology of 10:874–875
 - nonlinear characteristics of eye movements in 10:418–419
 - as transformative human experience 10:399–402

- reading (*continued*)
 useful knowledge about neurophysiology
 of 10:867–868
- reading as problem solving model (RESOLV
 model) 6:355
- real data example 14:301–304
 analysis of answer changes 14:302
 person-fit analysis 14:303
 score differencing to detect item
 preknowledge 14:303
 similarity analysis 14:302
- real talk in Australian schools 10:71
- real time performance assessment & real
 time scaffolding of inquiry
 competencies 14:465–467
- realism and theorizing curriculum 7:85–95
- realistic simulation 13:98
- reasoning
 and decision making 6:235
 development as form of divergent thinking
 6:237–238
 logical capacities development 6:235–238
 understanding individual differences in
 6:238–241
- Rebecca Rogers 2:197–198
- recentralization mode 7:76–77
- reciprocal causation 4:19–20
- reciprocal determinism framework and self-
 efficacy 6:289
- reciprocity
 indigenous philosophies of 2:112–113
- reconceptualized curriculum, slippery slopes
 of 7:371–372
- curricular enactment's embrace of
 multiplicities, lived terms of
 7:372–376
- meaning-making
 as elemental to being human 7:373
 as generative, mindful of place and
 process 7:373–374
 as imaginative, instilling embodied
 understandings 7:375
 as in need of other 7:374
 as reflexively informing self-
 understandings in relation to wider
 contexts 7:375–376
 as temporal/spatial negotiation
 7:374–375
- value of compassion and strategic empathy
 5:320
- reconciliation 3:94–96
- pedagogies 5:318–320
 pedagogical principle of mutual
 vulnerability 5:319–320
 in posttraumatic contexts 5:317–318
 significance of pedagogic discomfort
 5:319
- research
 age of apology 2:273
 apologies 2:275
- recording numbers, evolution of
 14:819–820
- recurrent neural networks (RNNs) 14:691
- recursive feedback 6:263
- REDDFORMA 5:463
- reduction, emergence instead of 10:885
- reductionism
 as epistemological issue 10:884–885
 literacy models and 10:886–887
- reflections 3:192–196
 becoming storywork ready
 for education 3:193–194
 for research 3:194–196
- reflexivity 12:60–63
 in qualitative research 12:4–5
- reflexology 6:37
- refugees 12:405–406
- education
 brain drain, repatriation of refugees and
 post-conflict recovery 1:141–142
 challenges compared to other
 international and transnational
 education 1:139–140
 curricular content and pedagogical
 orientation in refugee education
 1:141
 geographical distribution and focus of
 research on 1:138–139
 language policies and in global
 educational landscape 1:140–141
 particular challenges in 1:138
 pedagogical and psychosocial challenges
 in international education 1:139
 position in globalized world 1:136–137
 refugee learner 1:137
 refugee learners and right to education
 1:137–138
- students 9:169
 assimilation, and identity performance
 9:172–174
 deserving refugee, performing identity of
 9:172–173
 facing sudden influx of 5:672–674
 grateful refugee, performing identity of
 9:173
 heroic refugee, performing identity of
 9:173–174
 identities and “wise” educational spaces
 9:174–175
 native, performing identity of 9:172
 otherization and identities of 9:170–171
 and project of answering back to identity
 projections 9:174
 stigma and 9:170–171
 stories and studies of 5:672–674
 teacher identity and capacity to heal
 9:174–175
- refusal, academic gatekeeping and pedagogy
 of 3:71–72
- regional comparative and explanatory study
 (ERCE) 14:350
- regional large-scale assessments
 14:350–351
 for building capabilities in low-and middle
 income countries 14:351–352
- Regional Trade Agreements 1:506–507
- regionalism 1:422–423
- regression
 discontinuity 14:700
 discontinuity designs 14:700
- model 14:674–675
 estimators that combine propensity
 scores and outcome 14:677–678
 parameters
 parameter shrinkage by modeling
 structure of 14:573
- regulation as function of formative
 assessment 13:35–36
- reinforced learning ecosystems for STEM
 success 11:395–400
- reinforcement 6:17, 6:294
- reinforcement learning (RL) 14:410–411,
 6:85
- reinventing teacher agency 5:723–724
 critical spaces to achieve teacher agency
 5:724
 developing capacity to imagine and find
 voice 5:726
 disrupting normative discourse and
 challenging socialization 5:726
 economic, social and political context
 5:722–723
 experiencing language of possibility and
 social justice 5:726–727
 teacher's personal narrative 5:724–725
- reinventions 1:604
- relatedness, arranging for sense of
 10:404–405
- relational expertise 6:183–184
- relationality of belonging 3:112–114
- relationships 13:45
 relationship-building for school
 improvement 4:515
- relevance of emotions
 designing emotionally healthy learning
 environments 6:155–157
 for learning and performance 6:153–155
 optimizing learners' and teachers 6:156
 optimizing test-taking and feedback 6:156
 promoting enjoyment of learning in
 classroom 6:156
 for students 6:153–154
 students' and teachers' emotional
 experiences during classroom
 instruction 6:156
 supporting and fostering adaptive emotion
 regulation 6:157
 supporting students' feelings 6:156
 for teachers 6:154–155
- reliability 4:259–260
 classical test theory 14:22–26
 factor analysis 14:28–31
 foster culture of professionalism
 4:259–260
 generalizability theory 14:26–28
 item response theory 14:31–33
 reliability-like coefficients 14:62–63
 simulation-and game-based assessment
 14:472
 of subscores 14:292–293
- religious practices 10:210–213
- remote learning, affective shortfalls of
 10:932–934
- remote mentoring in diverse contexts
 5:732–735

- remote school, missing connections and affective depletion in 10:933
- RENKEI. *See* research and education network for knowledge economy initiatives (RENKEI)
- reporting mixed methods research studies
context of mixed methods reporting 12:740
framing 12:741–742
integrated mixed methods findings 12:750
integrating mixing 12:746
interpretation 12:750–751
methods 12:745–749
mixed procedures 12:748
problem statement, research questions, and rationale for mixing 12:743–745
qualitative procedures 12:748
qualitative sampling 12:746
quantitative procedures 12:748
quantitative sampling 12:746–747
research design 12:745–746
responding to reporting limitations for mixed methods 12:751
sampling 12:746–747
step-by-step approach to reporting mixed methods research 12:741–751
titling your paper 12:742
voice 12:742–743
- representational sampling strategies, alternatives to 12:203–204
- reproduction 6:612
- reproductive justice, SOGI, and HIV/AIDS 2:366–367
- resampling method
bootstrap 14:764–767
fundamentals 14:763–764
jackknife method 14:767–768
permutation tests 14:769–770
- “reschooling” higher education
curriculum 8:48–49
politicization of curriculum 8:50–51
strategic levers and curriculum 8:52–53
supercomplexity as driver of curriculum change 8:53
systems 8:49–50
- research and education network for knowledge economy initiatives (RENKEI) 8:208
- research as decolonial praxis 2:126–127
- research design 12:515, 14:696–700
blocked design 14:698–699
causal studies 14:698
cluster-randomized design 14:699
descriptive and correlational studies 14:697–700
experimental design and statistical procedure considerations 14:696
factorial designs 14:698
instrumental variable model 14:700
internal validity and causal inference 14:697
observational studies 14:700
randomized designs 14:698–700
regression discontinuity 14:700
single-case research designs 14:699
statistical analyses of causal designs 14:701
- time-series designs 14:700
weighting/matching 14:700
- research ethics 12:35–36
illegitimate medical research to formal research ethics 12:35
medicine to social sciences 12:35–36
- research funding
in higher education 8:108
trends in research funding and current challenges 8:110–111
- Research lesson (RL) 5:571
and ecology of action 5:573
- research methodology 12:382–383
- research study design 14:696–700
Blocked design 14:698–699
causal studies 14:698
CRD 14:699
descriptive and correlational studies 14:697–700
experimental design and statistical procedure considerations 14:696
factorial designs 14:698
internal validity and causal inference 14:697
observational studies 14:700
randomized designs 14:698–700
SCD 14:699
statistical analyses of causal designs 14:701
- research-policy connections “in time of cholera” 5:734–735
- researching educational dialogues, methodological issues for 6:608–609
- residential schools, impact of 9:153–155
- residual deviance 14:585
- resilience 4:79, 6:190
grit and resilience 6:193–194
as independent factors 6:193
relevance of resilience for education 6:190–191
- resistance
in “La Isla del Encanto” 10:811
by teacher educators 5:417
- RESOLV model. *See* reading as problem solving model (RESOLV model)
- response analysis 11:452–454
- response processes, automated scoring of 14:400–401
- response-time models 14:300–301
- responsive pedagogy
classroom feedback informed by 13:90–91
framework for classroom feedback practice as 13:91–93
emotions 13:92–93
feedback content 13:91–92
feedback dialog 13:92
feedback strategy 13:91
goals 13:91
self-efficacy 13:92
model of effective classroom feedback informed by 13:93
- responsive teaching 11:38–39
- restorative justice 2:360
- restorative literacies 10:262–263
- restorying truths and reconcilia(c)tion as future educational research 2:279–280
- restricted models for categorical items 14:640
- results-based integration, using software for 12:501–503
- resurgence 3:94–96
- retention 6:612
- rethinking multiliteracies
anti-black linguistic racism 10:144–145
casting our radical vision 10:138–139
cruel fallout of racism 10:144
disrupting anti-black linguistic racism in teacher education 10:145–146
false narratives of Mexican motherland 10:142–144
language partners 10:140–142
literacy education of indigenous multilingual children 10:139–140
normalizing black language in teacher education 10:145–146
- restricted purchase, approaches with 2:222–223
diminishment 2:222
formulaic 2:222–223
micro-analytic 2:223
- retrieval practice 6:664–666
potential reasons for students’ study decisions 6:668–669
and spacing combined 6:667
students’ use of 6:668
toward successful implementation of retrieval spacing and 6:669–670
- retrieval spacing 6:666–667
potential reasons for students’ study decisions 6:668–669
and retrieval practice combined 6:667
students’ use of 6:668
toward successful implementation of retrieval practice and 6:669–670
- revisiting teacher evaluation decade after reforms
achievement effects outcomes 5:58–59
downward vertical diffusion 5:56
evaluation system bias 5:58
history of state adoption 5:55–56
implementation of state reforms 5:57–58
intermediary organizations 5:57
internal determinants 5:57
mechanisms 5:55
methods 5:55
political dynamics causing state teacher evaluation reform 5:56–57
research question 5:54–55
teacher labor market outcomes 5:58
upward vertical diffusion 5:56
- rewards 6:294
- rhizomatic mapping 10:958–959
- ritornello, exercise in 12:205
- RL. *See* reinforcement learning (RL); Research lesson (RL)
- RNNs. *See* recurrent neural networks (RNNs)
- Road Scholar 6:508–509
- Rob Simon 2:199

robotics
 as/for STEM education 11:105–107
 in formal educational settings 11:105–106
 implementing robotics in K-12 settings 11:106–107
 in informal science education settings 11:105
 with entrepreneurship experiences in classrooms 11:299–300
 robotics-based teacher professional development 11:301
 in ways to enhance professional learning 11:296
 robust community perspective 11:450
 rubric 11:450t
 robust methods 14:813
 role of policy 9:219–220
 RTs. *See* randomized trials (RTs)
 rubrics 11:448–449
 depth 11:450t
 rubrics for depth, flexibility and robust community perspective 11:449–450, 11:450t
 rural Kenya
 young children invent Swahili pidgin in 10:799–800
 rural teacher resilience, factors influencing 5:197–200
 individual 5:197–199
 rural teachers' resilience development 5:199–200

S

S-STEM 11:378
 S-STEP. *See* self-study of teacher education practices (S-STEP)
 SA. *See* South Africa (SA)
 SA higher education, socioeconomic realities confronting 8:280
 SACMEQ. *See* Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ)
 SACP. *See* social action curriculum projects (SACP)
 safety, GSD 9:163–164
 Salamanca statement and framework 9:86
 sampling design 12:515–516
 sampling distribution 14:576–577, 14:755, 14:757
 sampling processes 14:575–576
 sampling schemes
 convenience sampling 12:486
 mixed methods sampling 12:486–487
 probability sampling 12:486
 purposive sampling 12:486
 types of 12:486–487
 Sarv Shiksha Abhiyan (SSA) 1:531–532
 achievement of 1:532
 role of external aid 1:532–533
 #SayHerName Day 2:361
 SBR. *See* standards-based reform (SBR)
 scaffolded equity courses 5:423–424
 scaffolded MMR designs 12:429
 analysis-centric MMR designs 12:430
 multistage MMR designs 12:430
 philosophy-centric MMR designs 12:430
 social justice-centric MMR 12:430
 scale alignment, special case of 14:237–238
 scaled residual deviance 14:585
 Scandinavian school system 9:442–443
 SCDs. *See* single-case designs (SCDs)
 scenario Delphi 12:609
 scepticism, epistemological response to 10:890–891
 scholarly communications, open access and unfinished transformation of 8:63–73
 scholarly domain 10:829
 scholarship
 of annotations 10:690
 gaps in 10:694
 scholarship of teaching and learning (SoTL) 5:237
 analysis of recent SoTL and DBER research artifacts in STEM 11:322–324
 school PIPMs. *See* school principals' information processing mechanisms (school PIPMs)
 school principals' information processing mechanisms (school PIPMs) 4:179–180
 school principal methods 4:456–458
 school inspections 4:456
 self-review 4:456
 student learning indicators 4:457–458
 surveys with school community 4:457
 written assessments 4:456–457
 schooling 10:252–255, 6:40–41
 continued growth and influences from 6:471
 in disaster capitalism to critical pedagogy 8:255–256
 politics and political of 7:108–109
 poverty and additional determinants of 9:54
 schools 10:301, 4:266–269
 capacity 4:209–210
 building school capacity for improvement 4:221–222
 changes and reciprocal effects 4:221
 for improvement 4:218–221
 journeys 4:222
 leadership 4:218–219
 participative decision making 4:220
 school improvement as system-wide phenomenon 4:222
 teacher collaboration 4:219–220
 teacher learning 4:220–221
 teacher motivation 4:220
 trust 4:220
 as centers of professional learning 4:298–299
 chains and teachers' work 5:5–6
 coaching is investment to professional capital in 5:149
 community 4:516
 access to schooling 9:87
 building community and empowering school community 9:94
 building inclusion for all collaboratively 9:91–94
 children's rights and education for all 9:86
 digital divide 9:87–88
 examining boundaries between inclusion and exclusion 9:91–93
 frameworks for exploring policy and practice 9:92–93
 implications for school leadership 9:93
 importance of sense of belonging to school 9:93–94
 impact of inequitable societies on children's wellbeing and life chances 9:86–87
 learning and digital divide 9:87–88
 impact of pandemic on children's education and mental health and wellbeing 9:87–88
 rationale for change 9:91–92
 reflections 9:93–94
 risks posed to children with disabilities 9:88
 state of precarity 9:87
 surveys with 4:457
 surveys with school community 4:457
 transformational leadership impact on 4:56–57
 transforming schools to build inclusion for all 9:94–95
 cultures 4:497, 9:230
 influence on induction programs 5:138
 curricula
 historical social science and 7:99–100
 implications for 2:184–185
 21st century skills in 7:52
 curriculum 7:102–103
 revisiting 10:600
 decision making and student achievement 4:206–207
 disengaged or disenfranchized in engagement 10:569
 learned over decades of work 10:569
 print text and textbook 10:569
 print-centrism Dwarfs multimodal textuality 10:569
 theoretical frameworks 10:567–568
 work with disenfranchised adolescents 10:564–567
 education
 arguments to implement interdisciplinarity in 7:272–274
 educators in 10:641
 evidence-informed decision making in schools and districts 4:169–176
 exo-system 4:153
 experiences 9:490–491
 expert group on schools' contribution to students' learning 13:158–159
 garden research with children 2:216–217
 GB practices
 business sector involvement in 4:421–422
 E-govs' experiences of barriers 4:425
 IEB governors 4:423–424

- IEB governors embodying lay knowledge 4:424
 policy context 4:421
 project design 4:423
 GSD students in 9:159–162
 improvement
 relationship-building for 4:515
 teacher leadership for 4:200–201
 teachers and administrators as catalysts for 4:199–200
 increasing competition between schools over enrollments 1:169
 inspections 4:456
 method 4:456
 international trends in 4:205–206
 internationalization
 was window on 1:351–352
 leader 4:255–256, 4:267–268, 4:519
 animating our conceptual framework 4:239–240
 challenges of headship 4:394–396
 decision making process in schools 4:236
 developing leaders 4:396
 development 4:191–192
 district–school arrangements that shape school level decision making 4:238–239
 enacting change 4:406–407
 index for inclusion to promote equity 4:407
 leadership development and leader development 4:393–394
 opportunities to thrive, survive and revive 4:396–398
 promoting equity in 4:406–407
 research in decision making 4:236–237
 role in induction programs 5:138
 school level factors that shape research use in decision making 4:237–238
 use of research
 leadership 3:154–156, 4:44–45, 4:188–189, 4:411–412
 Australia 4:67
 complexity theory 4:68
 cultural differences 4:68
 distributed leadership 4:65
 essential supports theoretical framework 4:244
 features of longitudinal data 4:539
 field of study and relevance to region 4:469–470
 IMPACT project 4:67
 implications 4:246–248
 inclusive and culturally responsive leadership 4:65
 influence 5:186–188
 instructional leadership 4:64
 international successful school principalship project 4:67–68
 in Latin America
 leadership frameworks 4:65–69, 4:472–473
 leadership influence and impact 4:66
 longitudinal designs in 4:540
 longitudinal studies 4:540
 LVMM 4:542–543
 models of 4:63–65
 moral leadership 4:64–65
 in Oman 4:502–503
 policy drivers and areas for future research 4:473–475
 policy trends 4:471–473
 principals prepared for job 4:472
 quasi-experimental designs 4:542
 research
 research trends 4:470–471
 significance of 4:44
 studies in CPS 4:244–245
 TSL 4:64
 useful designs 4:541–542
 learning, principals' support of 4:179–180
 learning networks in 5:284–285
 managing external demands in 4:211–212
 middle leader 4:115–119
 networks for ITE in England 5:75–76
 perspectives on organization of 5:368–369
 positive and negative experiences at 9:490–491
 power-sharing and trust in 4:204
 principals
 learning-centered leadership impact on teachers' intent to leave 4:180
 literature review 4:454–458
 methodology 4:454
 methods for 4:456–458
 school principals' professional development 4:458
 standards for school principals 4:455–456
 principalship
 in Thai literature 4:493–496
 in Thailand 4:491
 queer in 2:179–187
 recent literature related to research-engaged schools 4:151–156
 recommendations for organizing decision making for 4:207
 reflection 4:119–121
 reform
 constructivist theory of 4:227–228
 culture, structure, and agency 4:227–228
 teacher buy-in introducing reform 4:228–230
 research on evidence and research-informed practice in schools 4:151–154
 research-engaged school 4:150–151
 resistance from 4:162–163
 school factors affecting research engagement 4:154
 school-based literacies 10:585–586
 school-based teacher educators' professional learning 5:472–473
 school-valued literacy 10:191
 school-wide improvement efforts, leading collaborative and continuous 4:190–191
 school/school stories, stories of 7:343–344
 space
 physical and affective reproduction of school space at home 10:931
 popular discourses of 10:930–936
 student challenges to school space on TikTok 10:935
 techno-social transformation of 10:935–936
 standardization of teaching and learning in 1:169
 strength of evidence 4:154–156
 supporting student and teacher well-being in 6:726–727
 system and practitioner patterns of research and evidence-informed practices in schools 4:151–153
 teacher well-being in 6:725–726
 texts
 size of lexicon that students will encounter in 10:366
 theoretical landscape 4:146–149
 turnaround
 leadership practices in 4:164
 principals' role in 4:163–164
 US teachers learned about indigenous schools in Guatemala 10:139–140
 well-being and learning in 6:723–727, 6:723f
 workplace conditions in 5:565–568
 formalization of work processes 5:566
 learning potential of task 5:565–566
 learning resources 5:567–568
 opportunities for feedback, evaluation, and reflection 5:566
 shared norms 5:567
 teacher participation in handling problems and developing work processes 5:566–567
 YouTube videos for school study 1:625–627
 science 11:322–323
 education 11:444–445
 literacy 11:44
 Science, technology, engineering and mathematics (STEM)
 play to
 inclusion and exclusion criteria 11:165
 methods 11:165–168
 play types 11:171–173
 research question one 11:168–170
 research question two 11:170–172
 results 11:168–172
 search and data collection process 11:165
 Science, technology, engineering and mathematics (STEM) 11:11, 11:50–51, 11:96, 11:312, 11:401, 14:400, 7:314–315
 and arts 11:10–11
 automated scoring of 14:400
 development of STEM integration 11:266
 education 11:36, 11:189, 11:262
 factors related to student motivation in 11:402–405
 formal and informal STEM learning promotes student motivation in 11:404–405
 gender, age, family, culture, and student motivation in 11:403–404

- Science, technology, engineering and mathematics (STEM) (*continued*)
- immersion and interactivity two key factors 11:191
 - implications and future directions 11:266
 - inquiry-focused questioning for 11:97–99
 - integration of 11:266
 - mathematics and 11:100–101
 - measurement of motivational constructs in 11:405
 - methodological perceptions, considerations, and approaches in 11:265–266
 - methods 11:194
 - operationalizing 11:265
 - positioning diverse STEM workforce for 11:396
 - practice development and difficulty appropriateness 11:195
 - question 7:321–322
 - results 11:194–195
 - role of motivation in STEM education 11:401
 - STEM + C3 urban teacher residency program 11:313–317
 - synthesis of students' and teachers' perceptions 11:265
 - teacher and student perceptions of 11:262–265
 - theoretical foundation of motivation in 11:401–402
 - toward conceptualization of motivation in 11:402
 - use case 11:192–194
 - virtual reality in 11:190
- science, technology, engineering and mathematics (STEM) 11:401
- attrition 11:366–367
- science and engineering practices (SEPs) 11:444–445
- scientific current 2:212–213
- scientifically rational curriculum ideology, knowledge in 7:259
- score differencing involves test 14:299
- score equating 14:237–238
- concordances within testing programs 14:245–246
 - data collection designs 14:238–239
 - linkages across tests from different testing programs 14:244–245
 - procedures for equating scores 14:239–244
 - reporting score linkings that not equating 14:246
 - score linkings that misconstrued as 14:244–246
- score linkings that misconstrued as score equatings 14:244–246
- score reporting research 14:217–219
- audience specificity 14:218
 - conduct user-centered needs assessments 14:223–226
 - data visualization principles 14:221–222
 - design preliminary mockups 14:220–223
 - develop prototypes and iterate with systematic feedback 14:226–227
 - development of 14:222–223
 - focus groups and surveys to conduct user-centered needs assessments 14:223–226
 - focus on stakeholder interpretation and use 14:219
 - gather assessment information needs 14:220
 - iterative multistep framework for designing and evaluating 14:219–227
 - large-scale surveys to evaluate intended uses and decisions 14:227
 - operational deployment 14:227
 - score scales 14:445–446
 - scoring 8:317
 - inferences 13:215
 - scoring considerations for accountability uses of through-year assessments 13:215
 - Scott's communicative approach 6:603–604
 - SDG 7:410–411
 - curriculum and 7:408–409
 - SDG4 1:547–552
 - SDG 4. *See* sustainable development goal 4 (SDG 4)
 - SDGs. *See* sustainable development goals (SDGs)
 - SDSN. *See* sustainable development solutions network (SDSN)
 - SDT. *See* self-determination theory (SDT)
 - SEA-PLM. *See* Southeast Asia Primary Learning Metrics (SEA-PLM)
 - seclusion 8:381–391
 - second language acquisition (SLA) 6:364
 - crisis of trust 10:77
 - critical media literacy and 10:76–85
 - curriculum interventions 10:81–82
 - disparities of access and information 10:76–77
 - findings from studies on 10:21–22
 - interacting variables in 10:21–22
 - nonlinearity and variability in 10:21–22
 - photography 10:80–81
 - putting theory into practice 10:80–83
 - research studies 10:82–83
 - second language contexts 6:364
 - second-order factor analysis, CFA 14:611–613
 - secondary education 1:431
 - secondary school 9:449–450
 - epistemic practices of science in 5:254–262
 - students' STEM career aspirations 11:382–387
 - supporting English language learners in 10:67–70
 - secondary schooling, curriculum stratification within 9:80–81
 - SEL. *See* social and emotional learning (SEL)
 - selected-response items 14:187
 - developing common selected-response item types 14:189–190
 - selection models 14:810
 - selective social learning in infancy 6:455–457
 - contexts of 6:455–456
 - foundations of 6:456–457
 - self-assessment 13:79–80, 13:83–85
 - face-to-face 13:80
 - higher education 13:85
 - literature reviews 13:83
 - online 13:80
 - primary school 13:84
 - reliability 13:80
 - secondary school 13:84–85
 - theory of 13:80
 - self-authoring literacy 10:143
 - self-determination theory (SDT) 6:686
 - self-efficacy 11:430, 4:79, 6:27–28, 6:250–251, 6:288
 - academic self-efficacy 6:290
 - and academic self-regulation 6:291
 - construct about self 6:289
 - development of 6:289
 - for learning 6:32
 - measurement issues 6:253–254, 6:289–290
 - motivational, behavioral, and achievement consequences of 6:252–253
 - as predictor of academic performance 6:290
 - promoting self-efficacy of learners 6:254–255
 - reciprocal determinism and self-efficacy 6:289
 - roles of age, gender, and culture in self-efficacy formation 6:252
 - sources of 6:251–252, 6:289
 - in STEM 11:391–392, 6:291–294
 - academic self-efficacy 11:390–391
 - development of 11:389–390
 - educational strategies to develop and strengthen 11:392
 - gender differences in 11:391–392
 - measurement issues 11:390
 - reciprocal determinism and 11:389
 - sources of 11:389
 - self-esteem 6:532
 - self-evaluation
 - developing collaborative and rigorous 4:299–300
 - Self-Explanation Reading Training (SERT) 6:214
 - Self-Reflection Day 2:361
 - self-regulated framework 6:30–31
 - self-regulated learning
 - relationship between formative assessment and 13:5–6
 - self-regulated strategy development (SRSD) 6:613
 - self-regulation
 - educational implications 6:32
 - emotional and behavioral 10:400
 - of learning 13:37–38
 - social cognitive theory 6:29–32
 - phases of 6:30
 - self-reported epistemologies and methodologies 10:904
 - in conference submission system 10:904
 - self-review method 4:456
 - self-study 5:237, 5:461–463, 5:489
 - critical friendship, and complexities of teacher education 5:491–492

- and enacting pedagogy of teacher education 5:488–489
- Rodrigo's personal account of self-study of practice 5:462–463
- role and value of critical friendship in 5:463–466
- impact of self-study on Rodrigo's perspectives 5:463
- Tom's personal account of self-study of practice 5:461–462
- self-study of teacher education practices (S-STEP) 5:458
- context 5:459
- outline of argument 5:459–460
- resources for learning more about 5:466
- self-study as methodology for improvement of teacher education 5:460–461
- teacher or teacher educator study personal teaching practices 5:458
- teachers and teacher educators learn to teach 5:458–459
- self-video, researching teacher learning through
- designs 5:122–123
- tensions, challenges and future directions 5:123–124
- theoretical frameworks 5:121
- SEM. *See* structural equation modeling (SEM)
- semantic networks, building 6:470
- semantic processing effects, memory 6:111
- semiotics
- post-structuralism 12:194–195
- in qualitative research 12:195–196
- signs 12:190–191
- signs and meaning 12:191–192
- structure 12:192–194
- SEND. *See* special educational needs and disability (SEND)
- sense of belonging 3:112–114
- sensitivity analysis 14:814–815
- sensitivity to orthographic patterns 10:41–42
- sensor module 6:638
- sensory literacies
- embodied cognition and sensory studies 10:481–482
- across life course 10:483–484
- origins of 10:481–482
- pedagogical challenges and possibilities 10:484
- privileging of vision in theories of communication 10:481
- senses, body, and literacy practices 10:480–484
- sensorium in literacy and digital media practices 10:482–483
- sensory memory 6:108
- sensory motor roots 10:39
- separatist model 5:383
- sequencing learning activities 6:573
- sequential integration 12:755–756
- using software for 12:508–509
- sequential MMR designs 12:428
- SERT. *See* Self-Explanation Reading Training (SERT)
- service learning 5:422, 6:673, 6:675–676, 8:477–481
- authentic evidence 8:482
- black educators 6:674–675
- black social settlement movement 6:674
- as community-campus engagement 8:477–478
- conceptualizing 6:676
- course characteristics 8:482
- defining service-learning 8:478–481
- empirical support for student learning through 8:481–482
- institutional 6:675–676
- institutional organization to support 8:483–485
- longitudinal studies 8:481
- reconceptualizing service-learning for 21st century 6:677–679
- service-learning as experiential education 8:478
- settlement houses 6:673–674
- theorizing 6:676–677
- SES. *See* socioeconomic status (SES)
- set estimation 14:776
- settlement houses 6:673–674
- sex 10:609–610
- perspectives on 10:250–252
- troubling social models of 10:251–252
- sexuality 10:609–610
- sexuality-diverse teachers as persons, professionals and pedagogues 5:206–208
- sexuality and gender identity (SOGI) 2:365–367
- SFL. *See* systemic functional linguistics (SFL)
- shadow curriculum 7:288–290
- characteristics of 7:290–291
- defining shadow curriculum 7:290–292
- definition and characteristics 7:291
- evolution of concept 7:289–290
- focusing on academic success of students in schooling 7:291
- fulfilling academic needs and abilities of individual students 7:291
- for future curriculum studies 7:292–293
- increasing competitive nature of education 7:292
- providing differentiated content based on amount of tuition 7:291–292
- shadow education 2:346
- definitions and scope 2:346
- diversity in supply 2:348–349
- drivers of demand 2:347–348
- scale and nature of 2:346–347
- social and political implications 2:349
- Shanghai 11:285–286
- comparisons with 5:712–716
- shared leadership 4:27–28
- shared powerpoint slides, University College Northern Denmark (UCN), Aalborg, Denmark 7:240
- Shaw Lane girls academy 4:351–353
- contexts 4:351
- opportunity of ownership 4:352–353
- opportunity of purpose 4:351–352
- Shinto
- decolonial potentials 1:39–41
- politics of 1:41–42
- salvaging Shinto 1:42
- short text responses, automated scoring of 14:400
- short-term memory 6:108, 6:221
- SIBTEST 14:254
- siloed equity courses and practica structure 5:417
- simple linear regression model 14:549–550
- simple random sampling 14:705
- simple view of reading (SVR) 10:891–892
- simulation games 6:685
- simulation-based assessment example 14:473
- assessment in digital ocean 14:471–472
- challenges to field 14:473
- validity, reliability, and fairness 14:472
- Singapore 11:284–285
- context 2:291–295
- curricular and pedagogical modifications around academic differences 2:293
- diversity as academic readiness 2:292–293
- insights from Singapore teachers 2:292
- revelations of dominant societal narratives and concerns 2:293–294
- education phases and initiatives 5:711
- inclusive education in republic of
- contextual legacies and priorities 9:427
- cultural-historical framework 9:426–427
- inclusive education as catalyst for reform 9:431–432
- initiatives and support for preschool children with disabilities 9:429–430
- post-2004 9:429
- practice and support within mainstream schools and institutes of higher learning 9:430–431
- situating inclusive education within 9:428–429
- toward inclusive education 9:428
- principal preparation 4:361
- Singapore's Integrated Programme 7:135
- system leader and system leadership 4:132
- single-case designs (SCDs) 14:238, 14:699
- single-case research designs (SCD) 14:699
- single-imputation 14:807–808
- singular value decomposition (SVD) 14:594
- PCA of data matrix 14:594–595
- Sino-ASEAN higher education partnership 1:670–671
- SIT. *See* social identity theory (SIT)
- situational judgment tests (SJT)
- 14:487–488, 5:334–335
- for teacher selection 5:334–335
- situations model 6:354
- SJ. *See* social justice (SJ)
- SJT. *See* situational judgment tests (SJT)
- skepticism
- as epistemological issue 10:889–890

- skepticism (*continued*)
 underdetermination 10:889
 for literacy models 10:891
 skewness 14:727–728
 skilled immigration policies 1:633–634
 skills 1:635–636
 developing 6:471
 production of skills through education and training 1:634–635
 SLA. *See* second language acquisition (SLA)
 SLD. *See* specific learning disabilities (SLD)
 Slovenia's educational system, overview of 5:217
 small-group discussion 14:225–226
 small-scale qualitative studies 9:162–165
 SN-ABC. *See* social network-affective, behavior and cognition (SN-ABC)
 social action, YPAR and critical literacies for 10:232–233
 social action curriculum projects (SACP) 5:422
 social and emotional learning (SEL)
 6:92–97
 efforts on horizon 6:102–103
 escalating mental health problems in children and youth 6:93
 essential ingredients 6:99–102
 frameworks 6:95–96
 learning context 6:99
 need for 6:93–94
 and neuroscience 6:93
 research evidence 6:97–99
 skills are durable 6:99
 skills in childhood predict academic achievement and adult success 6:97–98
 social and emotional skills are malleable 6:98
 of students 6:100
 systemic 6:97
 of teachers 6:101–102
 transformative 6:96–97
 widespread support for 6:93–94
 social and environmental justice as ethical imperatives 5:684–686
 social attitudes 9:170
 social capital 4:267
 social change
 century social values 10:490–491
 creativity, innovation, and problem-solving 10:491
 criticality, responsibility, and social change 10:490–491
 multiplicity, hybridity, and flexibility 10:491
 participation & collaboration 10:490
 research centered on 10:493–494
 shifting social values and opportunities for 10:490–491
 social cognitive career theory 11:376
 social cognitive theory
 conceptual framework 6:24–25
 cross-cultural relevance 6:33
 early modeling studies 6:23
 educational implications 6:32
 enactive and vicarious learning 6:25
 historical influences 6:23
 of human functioning 6:250
 imitation 6:23
 integration with technology 6:33
 latent learning 6:23
 methodological refinements 6:33
 reciprocal interactions 6:24–25
 self-regulation 6:29–32
 social learning 6:23
 symbolic processes 6:27–29
 vicarious, symbolic, and self-regulatory processes 6:25
 vicarious processes 6:25–26
 social comparison 6:29
 social consensus 4:18–20
 social construction of learning 6:298–300
 social constructivism 6:301
 social context and language development 6:472–473
 social context of learning 6:573
 social contract 1:584
 capacities to unfold into gentleman of 9:210–211
 social development, literacy, culture and 10:296–297
 social education 1:415–416
 social efficiency as curriculum ideology, knowledge in 7:259
 social emotional dispositions 4:384–385
 social emotional learning 7:415
 social functions of education in context 6:259–260
 social identity theory (SIT) 6:276–277
 main influences for indigenous students 6:277–278
 social imagination 10:400–401
 social influence models 14:436–438
 multilevel and longitudinal extensions of 14:437–438
 social interactions, shared meaning making in 6:589
 social interests 10:29
 social justice (SJ) 2:79, 4:438, 5:690–691
 barriers to leadership identity construction for 4:443
 belonging to professional group 4:442
 counterstory of 10:225
 critical literacies and 10:217–218
 through critical literacy 2:196–200
 in education 4:39
 and education
 critiques of 2:16–17
 socially just education in our times 2:17–18
 theorizing social justice and education from above 2:13–15
 theorizing socially just education from below 2:15–16
 education policy 4:442–443
 experiencing language of 5:726–727
 factors of leadership identity formation for SJ 4:441–443
 forms of leadership identity building 4:440–441
 foundations of 6:454–455
 implications of ABER for 2:318
 importance of leadership identity building 4:439–440
 inclusion and teachers 5:68
 ITE's contribution to 5:693–695
 leadership identity formation for 4:439–441
 literacy as political project 10:218–219
 mathematics and stem 11:32–33
 mentors/role models 4:442
 personal characteristics 4:443
 pre-service and in-service training of principals 4:441–442
 in schools 2:195–196
 selective social learning 6:455–457
 and social and epistemic access to knowledge 7:91
 social justice-centric MMR 12:430
 community-based participatory MMR designs 12:430
 transformative mixed methods and community-based participatory research 12:430–431
 transformative MMR designs 12:430
 students and teachers as 10:222–223
 teacher education for 5:691–692
 tension in teacher education for 5:692
 theories 1:110–111
 social learning 4:289, 6:23
 in infancy 6:453–457
 social logic of curriculum models 7:3–5
 social media
 challenges of higher education and 8:136
 in higher education
 access issues 8:136–137
 privacy issues 8:136
 incorporated into LMS 8:133–136
 and learning 6:431–433
 affordances and constraints of social media as spaces for literacies 6:434–435
 challenges and opportunities for future of education 6:439
 civic learning with social media 6:435–437
 features and descriptions of social media literacies 6:434
 literacies for changing world 6:433–435
 networked professional learning spanning social media 6:437–439
 social media composition as literacy 6:434
 themes in research on 6:433
 platforms 10:792
 principals' work in public schools 4:307
 in teaching and learning 8:132–133
 and universities collide 8:134–136
 social model and debility studies 9:193–195
 social movements 2:341–343
 learning in 2:326–327
 social network-affective, behavior and cognition (SN-ABC) 6:533
 social networks 4:526–527

- research related to leadership and change 4:527–531
- theory and analysis 4:526–527
- social persuasion 6:292
- use of 11:392
- social processes in shaping of leadership 4:526
- social progressivism as political movement 5:682–683
- social realism as response to postmodernism 7:80
- social realist school in sociology of education 7:90–92
- social reconstruction as curriculum ideology, knowledge in 7:259–260
- social resources 4:78–79
- acting in emotionally appropriate ways 4:79
- managing emotions 4:78
- perceiving emotions 4:78
- social responsibility
- of bookmakers 10:636–639
- of influencers 10:640–641
- of readers 10:635–636
- social robots, mentoring by 10:537–538
- social sciences
- policy isomorphism in broad context of 1:273–274
- social selection models 14:432–436
- longitudinal extensions to 14:435–436
- multilevel extensions to 14:434–435
- social semiotics 10:504, 10:573
- perspective of grammar 10:55–60
- social studies
- affective nature of learning 6:319
- constraints on research in 6:317
- education
- contemporary perspectives on 7:329–330
- future directions for 7:330
- historical perspectives of 7:328–329
- influence of instruction 6:319–320
- learning in 6:317–320
- learning in and out of school 6:318–319
- purposes of learning 6:315–317
- as school subject 6:315
- social support and safety 6:293
- social-behavioral outcomes 6:289
- peer relationships and 6:291
- teacher–student relationships and 6:289
- social-cognitive model 6:613
- social-political arrangements 5:35–36
- social-to-self progression 6:30
- social/cultural constructivist perspective on learning “problems” 6:300–301
- socially critical current 2:215–217
- socially progressive teacher education programs, research for
- conceptualizing 5:681–686
- toward working definition of 5:686–687
- society
- impact on higher education 8:34–35
- life with conflicting frames 9:195–196
- orientation of frames 9:191–192
- social model and debility studies 9:193–195
- socio-cultural literacies 10:256
- socio-cultural studies of literacy 10:301–302
- socio-cultural theories 6:677
- developed over time 6:478
- of feedback 13:42–43
- of learning 6:3–5
- socio-cultural theorizing, influence of 7:386
- socio-ecological framework for field of mixed methods research 12:395–396
- socio-economic status 9:251
- and language development 6:472
- socio-material approaches to identities and curriculum 7:386–389
- socio-political domain 10:827–828
- socio-psychological theories 4:39
- sociocultural approaches, issues of progression and generalizability in 13:15–16
- sociocultural perspective on learning 6:37–42
- learning, development and schooling 6:40–41
- learning and mediation 6:37–38
- tool mediated action 6:39
- socioeconomic status (SES) 6:61, 8:339
- disparities by country/SES 8:339–340
- sociolinguistic competence 6:366
- sociological theories 4:39
- socioscientific issues (SSI) 11:143
- building new curriculum and adopting new repertoires of instruction 11:146–147
- in context of STEM education 11:143–144
- developing teachers’ pedagogical content knowledge in 11:148–150
- goals of SSI in STEM education 11:144–145
- hospitality of school contexts 11:147–148
- teachers’ conceptions of 11:146
- teachers’ sense of purpose in teaching STEM 11:145–146
- Socrates’ contrast between “living” and “dead” meaning 6:602
- soft governance in era of technology 7:78
- soft power
- approach juxtaposed to diplomatic approach 8:205–206
- knowledge diplomacy different from 8:205–206
- “soft-skills” gaming promotes development of 10:586
- software
- for detection of test fraud 14:301
- frameworks 14:689–690
- logs of learning process 13:57–59
- for MSA 14:143
- support 12:319–320
- SOGI. *See* sexuality and gender identity (SOGI)
- solidarity, mele and hula as embodied calls to action and 10:802–803
- Solomon islands 3:151–152
- SoTL. *See* scholarship of teaching and learning (SoTL)
- South Africa (SA) 8:51–52
- South Africa’s curriculum (2005) 7:5
- South African and international students at University of Pretoria, case study of narratives of 8:241–249
- South African education 9:252–253
- South African higher education
- internationalization and transformation in 8:211–212
- South African universities
- Bacchi’s WPR approach to representations of dropout problem amongst financial aid funded students 8:363–364
- findings 8:359–360
- language constraint 8:360–363
- lost in crowd 8:360–361
- method 8:358–359
- persistence 8:360–361
- untold sides of single story representations of dropout challenge in HE 8:359
- Wadzi’s story 8:361–362
- South America, transfer credit 8:348
- South Asia, education development in
- higher education in 1:431–435
- human development, literacy and education development 1:427–428
- policy dynamics 1:436–443
- secondary education 1:431
- universalization of elementary education 1:429–431
- South-East Asia, education in
- colonial legacy 1:420
- corruption 1:421
- demographic challenge 1:420
- diversity 1:419
- dropout 1:421
- education and decentralization 1:421
- financing of education 1:421
- inequality 1:420
- national identity 1:420
- public private balance 1:422
- QA 1:423
- regionalism 1:422–423
- teachers and university faculty 1:423
- south-south cooperation (SSC) 1:317
- first phase 1:320–321
- historical phases in development of 1:320–322
- historizing origins of 1:319–322
- second phase 1:321
- south-south cooperation and education
- policy 1:322–323
- third phase 1:321–322
- toward definition of 1:318–319
- transfer, and sustainable development goals 1:323–325
- Southeast Asia in higher education
- China’s higher education aid to Southeast Asia 1:669–670
- confucius institutes in Southeast Asia 1:668–669
- as host to Southeast Asian students 1:666–668

- Southeast Asia in higher education
(*continued*)
- Sino-ASEAN higher education partnership 1:670–671
- Southeast Asia Primary Learning Metrics (SEA-PLM) 14:351
- Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) 14:350
- southern theory 1:37–38
- ecofeminist and decolonial turns 1:38–39
 - politics of Shinto 1:41–42
 - salvaging Shinto 1:42
 - self-negation 1:42–43
 - Shinto's decolonial potentials 1:39–41
- space
- scalar understandings of 1:4–5
 - topological understandings of 1:5–7
- spaced practice effects, memory 6:110
- spatial dimensions of education 1:30–31
- spatial thinking 11:76–77
- approaching best practice in facilitating students 11:83–86
 - existing concepts of 11:76–77
 - facilitating growth of 11:81
 - gender gap in 11:81
 - geo-technology and 11:82
 - high demand *vs.* low performance in students' 11:80–81
 - low performance
 - in overall 11:80
 - in spatial components that highly relate to STEM 11:80–81
 - to measure 11:77–78
 - in STEM education 11:78–81
 - to cultivate 11:81–86
- Spearman, Charles 14:3
- contributions after Spearman 14:4–5
 - London school and development of classical test theory 14:3–4
- Spearman's g 6:199
- Spearman's ρ 14:743
- special education 9:222–224
- establishment 9:440–443
 - IQ testing, progressivism and professional demarcation 9:442
 - national cases 9:441–442
 - Scandinavian school system 9:442–443
- special educational needs and disability (SEND) 6:607–608
- special & inclusive education, debates in
- controversy about meaning of inclusion 9:273–274
 - controversy about nature of disability 9:272–273
 - disabilities and education 9:271
 - diverging approaches to disabilities 9:270–272
 - particulars of abilities and disabilities 9:271
 - quantification, measurement and categorization 9:271–272
 - variability among situations 9:271
- specific learning disabilities (SLD) 10:279
- culture matters in SLD identification practices 10:283–284
 - dangers of unitary approach 10:282–283
 - disrupting double bind identities 10:281–282
 - future SLD research 10:284–285
 - intersectional framing of 10:281–284
- specificity 6:212
- speech sounds
- discriminating 6:468
 - producing 6:468
 - word segmentation 6:468–469
- speech-print match 10:31–32
- spelling 10:429, 6:403
- assessment 10:379–380
 - closer look at developmental constraints 10:382–383
 - emergent and beginning spellers 10:376
 - explorations in development of spelling
 - from emergent to proficient spellers 10:374–379
 - explorations in other languages 10:383
 - implications and recommendations for assessment and developmentally-based instruction 10:379–383
 - instruction 10:380–382
 - intermediate spellers 10:378
 - proficient spellers 10:378–379
 - transitional spellers 10:376–377
- spoken language *vs.* formal/informal 10:530
- spoken responses, automated scoring of 14:399–400
- spoken word in context 10:71
- SRSD. *See* self-regulated strategy development (SRSD)
- SSA. *See* Sarv Shiksha Abhiyan (SSA)
- SSC. *See* south-south cooperation (SSC)
- SSI. *See* socioscientific issues (SSI)
- stage-environment fit theory 6:532
- stakeholder comprehension and usability
- cognitive interviews and surveys to 14:226
 - surveys and interviews to 14:226–227
- stakeholder engagement 14:371
- standard deviation 14:725–726
- standard error 14:756–757
- standardized scores 14:729
- standards-based model 5:470
- standards-based reform (SBR) 13:165–166
- Starratt, leadership frameworks and taxonomies 4:405
- statistical analyses of causal designs 14:701
- statistical approaches to detect types, test fraud 14:298–301
- statistical inference 14:751–761
- statistical learning methods 14:812
- in infancy 6:452–453
- statistical modeling approach 14:392
- statistical populations method 14:763
- statistics 11:323
- STEAM education
- goals of 11:11–13
 - history and meaning of 11:11
 - to improve educational experiences within systemic constraints 11:14
 - STEAM-inspired vision for education 11:14–16
 - STEAM's reach and barriers 11:13
- STEM
- additional considerations, and limitations 11:435–438
 - affective learning 11:421–422
 - assessment 11:413–414
 - assessment and teacher education 11:415
 - attitudes 11:424–430
 - career aspirations
 - assessing family influence and 11:377–378
 - early childhood experiences 11:373
 - experiences with tools 11:373
 - factors related to 11:383–384
 - families and STEM hobbies 11:374
 - families from marginalized communities and underrepresented groups in 11:375–376
 - family science habitus 11:372
 - formation of 11:382–383
 - gender and 11:344
 - gender differences 11:374–375
 - individual factors 11:383
 - influences 11:383–384
 - modeling 11:377
 - modeling and predicting family influences on 11:376–377
 - parental support 11:372–373
 - perceived characteristics of STEM careers and lifestyle preferences on career decision 11:345–347
 - recommendations for further research 11:385
 - relationships between abilities, interest, self-efficacy and STEM aspirations 11:344–345
 - role of family in 11:371–375
 - science capital 11:371
 - science identity 11:376
 - science self-concept 11:376
 - theoretical perspectives 11:384–385
 - competence 11:412–413
 - coursework, encouraging diversity and broadening pool of students taking 11:341
 - curricula, engineering design as key integrator in 11:411
 - curriculum, innovative approaches to 11:411–412
 - development of STEM integration 11:266
 - disciplines within in-service teacher education 11:277
 - ecosystem, role of informal STEM learning in larger 11:245–246
 - education
 - accommodation approach to access 11:327–328
 - applications of neurocognitive methods to 11:458–462
 - broad approaches to engineering integration 11:156
 - conceptual and definitional issues 11:361
 - conceptual flow graphics 11:157–159
 - considerations for selecting an engaging context 11:157–159
 - to cultivate spatial thinking in 11:78–81

- demand and supply 11:362
 EEG for 11:462
 fMRI for 11:465
 functional imaging approaches in 11:462–465
 global nexus of 11:358–359
 global STEM supply and domestic labor markets 11:363–365
 goals of SSI in 11:144–145
 immersion and interactivity two key factors 6:639
 implementation of 11:114
 implementing integrated STEM curricula in K-12 classrooms 11:159–160
 implications for STEM education 11:115
 integrated STEM curriculum development 11:154
 integrated STEM curriculum development and implementation 11:154–159
 integrated STEM curriculum framework 11:154–159
 integrated STEM curriculum implementation 11:160–161
 integration of models and modeling in STEM curricula and practice 11:114
 mathematics as language tool in 11:21–22
 mathematics as thinking tool in 11:22–24
 mathematics coursework and the supply of STEM graduates 11:366
 models and modeling in nature of models and their representation 11:112–113
 physiological approaches in 11:468
 political history of 11:359–361
 practice development and difficulty appropriateness 11:195
 problems and problematics in supply 11:365–367
 promises and perils of 11:262–269
 rationale for rise of 11:178
 research
 rise of
 roles of models in relation to specific modeling processes, knowledge and skills 11:113
 spatial thinking in 11:78–81
 SSI in context of 11:143–144
 STEM integration curriculum assessment 11:156–157
 strategies for teaching 11:114
 TEM attrition and STEM attraction 11:366–367
 UDE framework 11:329–330
 UDE guidelines and practices in 11:331–332
 understanding integrated STEM implementation through student discourse 11:160
 universal design approach to access 11:328–329
 universal design in
 virtual reality in 6:638–639
 and workforce development
 educational strategies to develop and strengthen STEM self-efficacy 6:291
 fair/competition 11:238
 family engagement in 11:239–240
 gender differences in self-efficacy and STEM achievement 6:291
 HBCU contributions in 8:426
 identity 11:435
 in-service teacher education 11:277–278
 design decisions associated with 11:279–280
 emerging developments in 11:280–281
 high quality 11:278–279
 studying 11:281
 information website 11:238
 integration 11:252–253
 broader and developmental approach to 11:258
 challenges to implementing broader approach to integrated STEM learning 11:259
 interest 11:430
 learning
 ecosystem 11:235–237
 ecosystem 11:397–399
 gap 11:298–299
 role of family and peers in 11:236
 role of individuals in 11:235
 role of school in 11:236
 role of society in 11:236–237
 literacy 11:43–46
 mastery experiences 6:291–292
 for mathematics 11:29–30
 citizenship STEM or cultivation of professionals 11:29–30
 notion of learning for sake of learning 11:29
 STEM and relationship with STEM disciplines 11:30
 methods 11:422–424
 motivation 11:435
 network website 11:239
 outreach activities 11:239
 promoting explicit mathematical connection-making in STEM lessons 11:25
 relationships between abilities, interest, self-efficacy and, STEM aspirations 11:344–345
 schools
 achievement 11:183
 characteristics of 11:181–182
 equity considerations 11:185–186
 inclusive 11:179–180
 inclusive STEM school outcomes 11:182–184
 racial/ethnic, economic, and gender disparities 11:184
 review of STEM school literature 11:182–184
 rigorous course-taking patterns and experiences 11:182–183
 school models 11:180
 selective 11:179
 selective STEM school outcomes 11:182
 STEM interests and aspirations 11:183
 varieties of 11:178–180
 vision for STEM schools of tomorrow 11:185–186
 science museum visit 11:239
 search results and discussion of constructs 11:424–435
 self-efficacy in 11:430, 6:291–294
 service learning 11:238
 setting up grade-appropriate mathematics learning goals in STEM lessons 11:24–25
 social persuasion 6:292
 and STEM education 11:103–104
 teacher, emphasize role of 11:84
 teacher induction, international perspectives on conceptual framework 11:284
 induction programs in Singapore, Japan, Shanghai, Australia, and France 11:284–286
 mentoring and induction programs and beginning STEM teacher effectiveness 11:288–289
 research on mentoring and induction programs promote culturally relevant teaching 11:286–288
 vicarious learning 6:292
 workforce for STEM careers, positioning diverse 11:396
 STEM. *See* Science, technology, engineering and mathematics (STEM); science, technology, engineering and mathematics (STEM)
 STEM Career and Technical Education schools (STEM CTE schools) 11:180
 STEM careers, perceived characteristics of 11:345–347
 STEM communities of practice (STEM CoPs) 11:240
 STEM CoPs. *See* STEM communities of practice (STEM CoPs)
 STEM CTE schools. *See* STEM Career and Technical Education schools (STEM CTE schools)
 STEM Integration Curriculum Assessment (STEM-ICA) 11:156–157
 STEM-ICA. *See* STEM Integration Curriculum Assessment (STEM-ICA)
 step-by-step guide for mixed methods research synthesis 12:658–659
 analyzing and synthesizing data 12:659
 critically appraising primary-level studies 12:659
 extracting data 12:659
 selecting sampling strategy, search strategy, and eligibility criteria 12:658
 writing protocol and selecting design 12:658
 writing report 12:659
 stereotype threat 6:278
 and indigenous students 6:278–279
 stigma

- stigma (*continued*)
 and refugee student identity 9:170–171
 and social attitudes 9:170
 stimulus control 6:18
 storms
 choosing change 3:125
 collective action for change 3:124–125
 mapping our journeys 3:121
 Masi methodology and talanoa 3:121–122
 mentoring for change 3:125–126
 slow change 3:122
 utilizing university to make change 3:123
 working with allies to make change 3:123–124
 story of David 9:183–187
 story of Louisa 9:188
 story of Marilyn 9:187
 storybook reading, focus on 10:184–185
 storytelling 10:629–630, 12:181–183
 mobilizing 10:815
 storywork ready 3:191–192
 strategic competence 6:366
 strategic performance 6:357
 strategic-relational approach, international organizations 1:469
 strategically managing external demands 4:210–211
 strengths-based approach
 to indigenous education potential to reframe deficit assumptions 4:339–340
 stress contagion 6:101–102
 stress/burnout, leadership 4:318–319
 “strong” CLT 6:367
 structural domain 10:827–828
 structural equation modeling (SEM) 14:619
 choice of estimation methods 14:622–623
 common SEM terminologies and notations 14:619–620
 model conceptualization and specification 14:620–621
 model estimation 14:621–622
 model evaluation 14:623–625
 structural gamification 6:685
 structure learning, Bayesian networks 14:126
 struggling readers, meaningfulness for
 agency over future life narratives 10:401–402
 arranging for reading to be meaningful for 10:403–405
 emotional and behavioral self-regulation 10:400
 happiness 10:402
 increasing reading competence in meaningful, engaged reading experiences 10:402–403
 intellectual engagement 10:401
 meaningful reading for 10:399–400
 reading as transformative human experience 10:399–402
 social imagination 10:400–401
 talk about books, expanding relationships, and shifting identities 10:400
 student
 CBAs and student emotions 14:509–510
 completion problem for particular students and institutions 8:375–376
 criticality on internet 10:547–548
 data analysis 11:452
 with disability
 barriers to higher education for 9:371–372
 facilitators to higher education for 9:372–373
 in higher education 9:369
 diversity 6:553–554
 effects on 2:403–404
 engagement in online education 1:152–154
 immersive video technologies 1:153–154
 opportunities for interaction 1:153
 synchronous live meetings 1:153
 video feedback 1:153
 video introductions 1:153
 weekly overviews 1:153
 equitable access to key ideas for all students 6:332–333
 exclusion 9:380–382
 enrollment practices 9:380
 issues raised in reviews and inquiries 9:380–381
 recommendations 9:381–382
 restrictive practices 9:380
 findings 5:626–631
 first study 5:626–629
 growth percentile model 14:385–386
 advantages and disadvantages 14:386
 determining growth adequacy 14:386
 model overview 14:385–386
 modeling requirements 14:386
 high-quality learning experience for 4:405
 impact on higher education 8:30–32
 benefits 8:30–31
 tuition costs 8:30
 variations in benefits across students 8:31–32
 learning 4:209–210
 benefits of 8:307–308
 creating student learning outcomes 8:306–307
 examples 8:309
 expert group on schools’ contribution to 13:158–159
 indicators 4:457–458
 national peer educator study 8:309–310
 Nelson Mandela university student learning outcomes initiative 8:310
 outcomes assessment 8:305–307
 through service-learning 8:481–482
 shared responsibility for 8:305–306
 steps in 8:308
 maintaining motivation in teaching 5:628–629
 methods 5:626
 mixed feelings 5:628
 mobility
 new class formations 8:237–238
 beyond pandemic 8:238–239
 processes of commodification 8:234–235
 selling mobility for higher education 8:235–236
 understanding student demand 8:236–237
 multidimensional and changing nature of teacher professionalism 5:622–623
 opportunities for students to developing productive belief systems 6:333
 parental involvement on 4:266–267
 perceptions of STEM 11:263–265
 power of dialogic interactions between students and communities 5:377
 ratings 13:329–330
 of teaching practice 13:329
 response on squirrel post-unit assessment focus group 11:452
 segregation 9:380
 issues raised in reviews and inquiries 9:380
 recommendations 9:380
 selection 4:374–375
 sensemaking in project-based learning
 design process 11:446–450
 major findings 11:454–455
 methodology and analysis 11:450–454
 problem and objective 11:444–445
 significance 11:456
 significance of work 11:445
 theoretical framework 11:445–446
 social relationships at school
 interpersonal relationships with peers and teachers 6:289–291
 relationship provisions and affordances 6:291–294
 and staff co-creating curriculum in higher education
 benefits 7:240–241
 challenges 7:241–242
 examples of practice 7:239–240
 theoretical frameworks and models 7:237–239
 student-centered family, acquaintances of 6:709
 student-centered instruction 6:701–711
 student-faculty pedagogical partnership 8:268–269
 bringing together practitioners of 8:272
 center-based forms of facilitation 8:272–273
 challenges and ways of addressing them 8:273–274
 consistent outcomes of 8:269–270
 facilitation by faculty or staff member 8:272
 facilitation by single student 8:273
 models of partnership facilitation 8:272–273
 national, institutional, disciplinary, and individual influences on approaches to partnership 8:271
 partnership values 8:270
 teams of students, faculty, and teaching and learning center staff 8:272

- whole-of-class facilitation 8:272
- student-led functions on student achievements 11:221
- students-accountability, commodification of 1:65–66
- success 8:340–341
- t* distribution 14:712
- transformational leadership impact on 4:57–58
 - direct effects of 4:57–58
 - indirect effects of 4:58
- uses of YouTube 1:625–626
- voices 13:122–123
 - on assessment 13:124–126
 - equality implications at system level 13:125–126
 - key messages form research in assessment reform and 13:123–126
 - learning from listening to young people 13:124
 - role in promoting inclusive education 9:143–150
 - system level issues 13:125
 - young people as assessment policy actors 13:125
- well-being 6:723–725
 - empirical evidence 6:724
 - predictors of 6:724–725
 - supporting student and teacher well-being in school 6:726–727
- word knowledge
 - features of words that influence students' word knowledge 10:366–367
- Student Activist Day 2:360–361
- studies and essays/articles surveyed, critical evaluation of 10:689–690
- studies variance assessment
 - example of between-studies variance assessment 14:851
- study abroad
 - programs
 - challenge of linking collaborative partnerships to quality 1:107–108
 - challenge of over-committed faculty from the Global South 1:107
 - developing two-way learning through sustainable partnerships 1:106
 - financial hurdles 1:108
 - issues related to Global North and South 1:105–106
 - need for intercultural competencies and pedagogy 1:105
 - need for sustainable and equitable to enhance and support 1:106–108
 - notions of decoloniality and “othering” 1:106
 - rewards and challenges transformation of 1:105–106
 - studying abroad while staying at home 1:108
 - theoretical perspectives on impact of sustain initiatives 1:109–112
 - studying teachers' work and lives
 - appreciation of pedagogical and collaborative aspects of teaching 5:626–627
 - being teacher in changed circumstances 5:624
 - data collection and analysis 5:626
 - participants 5:624
 - portraying negative picture 5:627–628
 - studies 5:624–626
 - teachers' work and role of emotions 5:623–624
 - sub-discipline of education research, cultural configuration of 7:118–119
 - Sub-Saharan Africa
 - challenges facing education in 1:388–397
 - girls' education in 9:471
 - subscores
 - beta-binomial model 14:293
 - correlation or reliability of 14:292–293
 - dimensionality assessment software as DETECT and DIMTEST 14:293
 - Haberman's classical test theory based method 14:294
 - methods to evaluate subscores worth reporting 14:292–294
 - MIRT models 14:293
 - principal component analysis and factor analysis 14:293
 - psychometric quality of 14:292
 - real data example 14:294–295
 - subscore reporting 14:290–292
 - subscores adequate psychometric quality 14:294
 - sudowriters 10:539
 - summative assessment 13:39, 13:197
 - summative classroom assessment 14:520–522
 - See also* formative classroom assessment
 - feedback 14:521–522
 - grading 14:521
 - and learning targets types 14:521
 - summative metaevaluation 13:264–265
 - superimposition
 - compliance strategies for academic survival
 - on curious search for progress in knowledge 8:41–44
 - bias against novelty 8:42
 - looking for funding 8:42–43
 - split between elite and rank-and-file habitus 8:43–44
 - of managerial control of research on control by peer review 8:40–41
 - consolidating established lines of research and nimble knowledge production, limited scope for creative deviance 8:40–41
 - supervised learning approach 14:409–410, 14:685, 6:84
 - support-vector machines (SVM) 6:86
 - supporting learning 6:47–50
 - design of learning environments in broad range of educational contexts 6:47
 - pedagogies 6:47–49
 - survey project, context of our 12:668–672
 - sustainability current 2:212–213
 - sustainable development goal 4 (SDG 4) 1:539
 - background to 1:539–540
 - emergence and structure of 1:540–542
 - limits and liabilities of 1:543–545
 - optimism and virtues of 1:542–543
 - sustainable development goals (SDGs) 1:539, 9:86
 - background to 1:539–540
 - for education (SDG 4) 9:19
 - competing agendas and re-fragmentation of inclusive education 9:24–25
 - creation of 9:20–24
 - evolution of inclusive education 9:20
 - operationalization of 9:21–24
 - utilitarian capture of quality 9:24
 - SSC, transfer, and 1:323–325
 - sustainable development solutions network (SDSN) 8:208
 - SVD. *See* singular value decomposition (SVD)
 - SVM. *See* support-vector machines (SVM)
 - SVR. *See* simple view of reading (SVR)
 - Sweden
 - case of 7:261–263
 - educational system 5:217–218
 - formulating official view of knowledge in 7:263
 - Swedish curriculum discourses, knowledge in 7:262–263
 - Swedish education system 9:388–389
 - Swedish policy-context, inclusion in 9:390–391
 - Swedish teacher-training, inclusion in 9:391–392
 - “sweet spot” between fragmentation and coherence 4:411–413
 - symbolic processes 6:27–29
 - goal properties 6:28
 - goals and self-evaluations of progress 6:28
 - self-efficacy 6:27–28
 - synchronous discussion forum 11:209–210
 - syntax 6:467–468
 - system leader and system leadership 4:131–133
 - system leadership
 - defining system leadership 4:124–125
 - emergence of 4:127–128
 - global perspectives on 4:128
 - macro and micro level framework for 4:128–131
 - system leader and system leadership 4:131–133
 - for system transformation 4:125–126
 - toward future and reflection 4:133–134
 - system transformation, system leadership for 4:125–126
 - systematic feedback
 - cognitive interviews and surveys to evaluate stakeholder comprehension and usability 14:226
 - develop prototypes and iterate with 14:226–227
 - systematic literature review 8:317

systemic functional linguistics (SFL)
 10:54–60
 critiques to SFL pedagogy and responses
 10:60–62
 genres 10:58–60
 registers 10:57–58
 SFL scholarship in US 10:62–63
 systemic racism
 in education 10:275
 in politics and education tacit knowledge
 10:273–276, 6:211

T

- T3D model. *See* transition-to-trained-doctor model (T3D model)
 Talanoa 12:182, 3:164–165
 introducing rise up academy, rise up trust
 3:166
 Mana Academy 3:165–166
 Mana Moana 3:167
 NAPALI 3:167–170
 as simply conversation 3:162
 talent 1:635–636
 TALIS. *See* Teaching and Learning International Survey (TALIS)
 talking circle 12:182
 Tamagotchi model 8:58–59
 Tanzania, education in developing countries
 1:210
 target language (TL) 6:364
 tau-equivalent 14:53
 essentially 14:53–54
 TBA. *See* Test-based assessment (TBA)
 TBL. *See* team-based learning (TBL)
 TDKRA. *See* Transforming Disability Knowledge Research and Activism (TDKRA)
 TE. *See* teacher education (TE); teacher engagement (TE)
 teach mandated curriculum 5:452
 teacher 5:15–17
 accountability
 policy mutations 5:64
 in US, England and Australia 5:64–65
 advice for 9:493
 Bella 9:493
 Hannah 9:493
 Jodie 9:493
 Rachel 9:493
 advice for 9:493
 Bella 9:493
 Hannah 9:493
 Jodie 9:493
 Rachel 9:493
 agency 5:339–340, 7:191–192
 competency base 5:340
 critical spaces to achieve teacher agency
 5:724
 and curriculum making 7:192–195
 for curriculum making 7:193–194
 curriculum making for 7:194–195
 developing capacity to imagine and find
 voice 5:726
 ecological model of 5:340
 environmental conditions 5:341–342
 noticing and assessing student 5:343
 professional purpose 5:341
 revisiting teacher agency 5:723–724
 supporting teacher agency 5:340–342
 teacher's personal narrative 5:724–725
 alignment with teacher professional
 development program 11:301
 are responsive and flexible 5:16–17
 beliefs
 disciplinary ways of thinking
 11:253–255
 integrated STEM examples based on
 11:255–257
 interdisciplinary 11:257
 multidisciplinary 11:256–257
 and practices 11:253–257
 single CTE discipline 11:255–256
 single S/M discipline 11:256
 in STEM integration 11:253
 transdisciplinary 11:257
 buy-in introducing reform 4:228–230
 political clashes between district and
 valued constituencies 4:229–230
 reculturing teaching practices 4:229
 candidate resistance 5:418–420
 compliance and disrespect 5:420
 denial and defensiveness 5:419
 distancing strategies 5:419
 "good people", emotionalities, and
 maintenance of innocence
 5:418–419
 saying "right thing" 5:419–420
 strategic silence 5:419
 career policies 5:66
 centrality of teacher assessment 13:233
 communicating norms and expectations
 6:291
 community 7:341–342
 teacher community stories, stories of
 7:344–345
 computational thinking as core 21st century
 skills relevant for 5:265–266
 connect CT, equity, and data 11:317–318
 curating tablet resources for 5:285
 deployment 5:102
 through coaching, lessons for 5:150–151
 coaching for 5:144–145
 digital literacy and computational thinking
 as core century skills relevant for
 5:265–266
 with disabilities
 advocating for change through robust
 evidence 9:349–350
 impact of covid-19 9:348
 discussion 9:348
 engagement with colleagues 9:346–347
 experiences of 9:343
 facilitating inclusion 9:348–349
 findings 9:343–348
 impact on pedagogy 9:346
 insights from current literature of
 9:342–343
 note on anonymity 9:343
 official documents 9:343
 parents and community 9:347–348
 reasons for becoming teacher 9:343–344
 research approach 9:343
 students 9:347
 systemic efforts needed to increase
 participation and retention of 9:349
 teacher identity 9:344–346
 disciplinary ways of thinking 11:257–258
 discourse effects in whole class on small-
 group functioning 6:597
 diversity 9:274–275
 case study of 5:129–132
 diversity of 5:101
 education accountability 5:614
 educators
 career trajectories and contractual
 arrangements of 5:614
 general discussion of 5:445–446
 goals for courses 5:450–452
 identity 5:453–454
 learning, critical pedagogy as
 complementary frame for
 considering 5:478
 lives and journeys 5:449–454
 place of doing research 5:452–453
 professional growth 5:505
 professional growth as researchers
 three-stage model of scholarly growth in
 teacher education 5:505–508
 variability in work of 5:471
 epistemologies and epistemic practices of
 science 5:256–257
 examples from teachers' professional
 learning 6:418–419
 expertise
 adaptive expertise 5:175
 alternative perspective 5:177
 deliberate practice 5:172–173
 deliberate practice and reflection in
 teacher education and teacher
 professional development
 5:174–175
 development of 5:172–175
 experience and deliberate practice
 5:172–173
 expert teachers use of memory when
 dealing with classroom situations
 5:171–172
 experts' use of memory resources 5:171
 nature of expertise from cognitive-
 psychological perspective
 5:170–172
 stages in development of teachers as
 professionals 5:173–174
 teacher expertise for specific tasks 5:174
 framing debate 5:180–181
 identity 5:232–235, 7:342
 affective models of 5:350–351
 color fan of possibilities 5:235–240
 from identity research to a pedagogy of
 identity 5:348
 integrated approach 5:351–352
 make explicit 5:232–233
 make public 5:234–235
 make sense 5:233–234

- professional identity and professional learning 5:347–348
- rationalist models of 5:349–350
- theoretical framework 5:348–349
- implications for 5:644–645
 - for teacher professional development 5:176–177
 - for teacher research 12:408–409
- instrumental aid 6:292
- interventions with collaborative groups 6:594–595
- job satisfaction 5:214, 5:219
 - country settings 5:215–218
 - data analysis 5:221
 - data sources and sample 5:218
 - disciplinary climate 5:219
 - job demands and 5:223–224
 - job resources and 5:224
 - limitations and further research 5:224–225
 - measures 5:218–221
 - method 5:218–221
 - needs for professional development 5:219
 - relationship between job demands, resources and 5:215, 5:222–223
 - results 5:222–223
 - teacher autonomy 5:220
 - teacher professional collaboration 5:220
 - teacher self-efficacy 5:220–221
 - workload stress 5:219
- know and perform “works” 5:15–16
- knowledge & knowledge-for-teaching 7:340–341
- knowledge/tool for professional development 6:548–549
- language and 5:101
- leadership 4:49, 4:196–197
 - and administrators as catalysts for school improvement 4:199–200
 - through collective leadership 4:197–198
 - development 4:197
 - for school improvement 4:200–201
- learning
 - in mathematics classroom 5:518–527
 - needs 5:274
 - from noticing 5:524
 - rise of identity work in 5:348
 - as situated in classroom 5:519–520
 - understanding and designing for 5:482
 - as workplace learning 5:562–569
- lived experience
 - contrasts from countries 7:365
 - English history teachers’ experience of curriculum reform since 1985 7:366–369
 - history national curriculum in England 7:363–364
 - history teachers’ experience of curriculum since 2010 7:364–365
 - methodology 7:366
- lives 5:183–188
 - commitment 5:184–185
 - professional identities 5:185
 - resilience 5:186
- school leadership influence 5:186–188
- living curriculum 7:359–361
 - action research 7:359
 - culturally relevant pedagogy 7:358–359
 - education for liberation 7:358
 - notes to Mr. G. 7:359–360
 - theoretical framework 7:358–359
 - transforming understanding of electrical drawings using practical tasks 7:360–361
- mediate curriculum 7:144–146
- mediating of curriculum making processes
 - compliance, resistance, and creative mediation 7:143–144
 - from curriculum implementation to curriculum making 7:142–143
 - curriculum making revisited 7:143
 - role of teacher education and professional development 7:146
 - teachers as mediators of curriculum 7:146
 - teachers mediate curriculum 7:144–146
- as mediators of curriculum 7:146
- mentors 5:389
 - acknowledgment and recognition of 5:389–390
- meta-evaluation mechanisms to support teacher evaluation systems 13:290–291
- model of teacher effectiveness 5:297–299
 - attitudes, beliefs, expectations 5:299
 - behavioral dimensions 5:299
 - capability dimensions 5:299
 - cognitive characteristics 5:299
 - collegiality 5:299
 - commitment and motivation 5:299
 - communication 5:299
 - competence 5:299
 - critical reflection 5:299
 - cultural competence 5:299
 - psychosocial characteristics related to self 5:298
 - social and emotional competency 5:298
 - state dimensions 5:297–299
 - trait dimensions 5:297
- narrative hermeneutics for 5:119
- noticing 5:520–521
 - cognitive-psychological perspective 5:324
 - conceptualizing 5:323–325
 - development during teacher education 5:326–327
 - different perspectives on teacher noticing research 5:323–324
 - for effective classroom management 5:326
 - expert-novice comparisons 5:326–327
 - expertise-related perspective 5:324
 - facets of 5:324–325
 - intervention studies 5:326
 - in mathematics 5:325–326
 - modeling and measuring 5:325–326
 - noticing as part of teacher professional competence 5:325
- part of teacher professional competence 5:325
 - prevalence of perspectives 5:324
- socio-cultural perspective 5:324
- pedagogy 5:204–206, 5:640–642
- perceptions of STEM 11:263
- personal narrative 5:724–725
- perspectives and practice of assessment for learning in classrooms in Tanzania
 - data and results 13:66–70
 - literature 13:65–66
 - methods 13:64–65
 - participants 13:64
 - study 13:64–65
 - teacher development 13:64
- policy frames
- in 2000s 5:63–66
 - accountability in teacher education 5:65–66
 - arguable matters 5:67–68
 - irruption of COVID-19 and challenges to existing policy and practices 5:63
 - new public management policies 5:63–64
 - teacher accountability in US, England and Australia 5:64–65
 - teacher accountability policy mutations 5:64
 - teacher-related policy 5:62–66
- power of dialogic interactions between teachers and communities 5:377
- professional development
 - induction, mentoring, and coaching in 5:145–147
 - and learning 5:563–564
- professional learning 5:518–520
 - about assessment and pedagogy, DSE framework for rethinking research and 9:354–355
 - leading and participating in 4:7–8
 - multidimensional and changing nature of 5:622–623
- professional needs of 5:672–674
- professionalization 5:204–206
 - identity, and collective representation 1:406–407
- quality 5:544–546
 - schools matter in developing and sustaining 5:545–546
 - social ecological approach to researching 5:544–545
- ratings 13:329–330
 - of teaching practice 13:329
- recommendations for teacher training policy and practice 5:423–424
- recruitment
 - competency-based approach for 5:332
 - four opportunities to build teaching workforce 5:330
 - teacher attraction 5:330–331
 - teacher selection 5:331
 - two sides of 5:329–331
- relational work 5:642–644
 - colleagues 5:643–644
 - families 5:643

- teacher (*continued*)
- students 5:642–643
 - relationships
 - to curriculum and pedagogy 5:41–42
 - with leadership and colleagues 5:42
 - research capacity
 - enabling or constraining research capacity 5:619–620
 - four trends affecting 5:612–614
 - as researchers 2:195–196
 - resilience 5:200
 - retention, principal's role in 4:178
 - rewards and reinforcements 6:294
 - role of teacher education and professional development 7:146
 - through self-video
 - designs 5:446
 - tensions, challenges and future directions 5:446–448
 - theoretical frameworks 5:446–448
 - shifts in teachers' work amidst COVID-19 5:637–638
 - shortages 5:656–660
 - attract and retain teachers 5:665–668
 - career motivations of intending teachers and non-intending teachers 5:662–663
 - characteristics of potential teachers 5:661
 - government policies that increase demand for teachers 5:657–658
 - implications for research, policy and practice 5:668
 - key factors differentiating three groups 5:664
 - lag time between pupil and teacher numbers 5:656
 - limitations of teacher supply model 5:657
 - policies that suppress supply of teachers 5:658–660
 - promising approaches for attracting teachers to schools 5:665–667
 - promising approaches for retaining teachers 5:667–668
 - wants to be teacher 5:660–664
 - skills
 - educational digital competence 5:266–268
 - and knowledge for technology integration 5:263–271
 - learning capacity for future-ready teaching profession 5:268–269
 - for technology integration
 - trends in learning technology development 5:264–265
 - social and emotional support and safety 6:293
 - social justice, inclusion and 5:68
 - South-East Asia 1:423
 - stories of 7:343
 - stress 6:101–102
 - student relationships
 - and achievement 6:289
 - and motivation 6:289
 - teach to best and well 5:181–183
 - teacher-led functions on student achievements 11:221
 - teacher-led model of CLT 6:366–367
 - teacher-related policy 5:62–66
 - teacher-research 2:196–200
 - teachers' lives 5:183–188
 - and teaching in low-income countries
 - challenge of inclusive education 5:100–101
 - educational quality 5:102
 - international policy framework for teachers 5:100
 - teaching influenced by
 - methodology 5:288
 - quality assessment of reviewed studies 5:288
 - studies reviewed 5:288
 - understanding effective teachers and teaching 5:300
 - understanding juxtaposition between effective teachers and teaching 5:299–300
 - transformational leadership impact on 4:57
 - transnational policy 5:67–68
 - turnover intention 4:505
 - understanding of and relationship to themselves 5:43
 - use of technology 5:273–274
 - developing teachers' use of technology 5:274–275
 - PLN for teachers' development 5:275
 - professional development 5:274
 - teachers' learning needs 5:274
 - YouTube 1:626–627
 - well-being 6:725–726
 - predictors and outcomes of 6:726
 - supporting student and teacher well-being in school 6:726–727
 - work amid global education reform
 - Australian context 5:30
 - GERM and education reform in Australia 5:31–34
 - GERM and policy effects 5:34–36
 - global education reform movement 5:31
 - work lives and development
 - critical incident vignette 5:109–110
 - narrative and biography in pedagogical practices for 5:109–111
 - narrative–biographical perspective 5:106–107
 - promise and potential of narratives 5:107–108
 - rationale in narrative pedagogies 5:110–111
 - teacher development in action and thinking 5:108–109
 - teacher education (TE) 5:6–7, 5:12–15, 5:368–369, 5:437, 9:233–235
 - accountability in 5:65–66
 - aim of TE effect research 5:437–438
 - analysis methods 5:442
 - anti-black linguistic racism in 10:145–146
 - anti-racist approaches in 5:420–422
 - and arts
 - brief snapshot of relevant research 5:305–306
 - challenges and tensions for teacher education in arts 5:307–309
 - embedding creative arts as core formal curriculum component 5:306–307
 - Finland 5:307
 - opportunities 5:309–311
 - professional learning during pandemic 5:311
 - Singapore 5:307
 - understanding teaching as artistry 5:309
 - Wales 5:307
 - background and basic elements of critical approaches to teacher education policy analysis 5:11–12
 - be held accountable and wants to hold itself accountable 5:13
 - black language in 10:145–146
 - case methods in 5:405–406
 - narrative spaces in 5:411
 - proposing next steps 5:409–411
 - recent research on construction of 5:407–408
 - recent research on impact of 5:406–407
 - research studies 5:408
 - changing knowledge & cultures in 9:136–142
 - in crisis and in need of improvement 5:12–13
 - critical perspectives on history and context of 5:481
 - curriculum making and 7:158–160
 - data collection 5:442
 - design 5:441
 - development of teacher noticing during 5:326–327
 - dismantling traditional teacher education 5:446–448
 - for diverse learners
 - addressing roots of inequity through transformative action 5:360–361
 - culturally relevant pedagogy 5:358
 - culturally sustaining pedagogy 5:358–359
 - engine for transformation 5:362–364
 - multicultural education 5:357–358
 - preparing future teachers to teach diverse students 5:357–360
 - preparing white Teachers for diverse students 5:357–359
 - recruiting, preparing, and retaining teachers of color to teach diverse learners 5:359–360
 - roadmap for transformation 5:361
 - student diversity 5:356–357
 - transformative teacher education for educational equity 5:361–364
 - vehicle for transformation 5:362
 - effects 5:438–441
 - global conversations about teachers and 5:67–68
 - government intervention in 5:446–448
 - implications for 2:90, 5:176–177

- developing routine and adaptive expertise 5:176–177
- practicing for adaptive expertise 5:176–177
- improvement through self-study of practice methodology 5:490–493
- moving forward through self-study 5:494
- providing students with written feedback on assessment 5:493
- reconceptualizing teacher knowledge as tensions 5:491
- self-study improve teacher education practices 5:490–491
- teacher to teacher educator 5:492–493
- theoretical assumptions 5:489–490
- instrumentation 5:441–442
- is evidence-based 5:13–14
- is practical 5:14–15
- knowledge bases for 5:448–449
- large-scale quantitative studies 5:440
- methodology of exemplary studies 5:440–441
- methods in 5:405–406
 - curating case-based knowledge system for teaching 5:410
 - exploring case-based knowledge system for teaching 5:409–410
 - learned from recent research studies 5:408
 - narrative spaces in 5:411
 - proposing next steps for case methods practice and research 5:409–411
 - recent research on construction of 5:407–408
 - recent research on impact of 5:406–407
- mixed-methods studies 5:440–441
- modeling object of research 5:439
- more symmetric partnerships in 5:372–373
- moving from apprenticeship model to academic/intellectual endeavor 5:446
- origins of 9:136–137
- overwhelming presence of whiteness in 5:476
- in pandemic times
 - assessments for and learning 5:713–714
 - and beyond 5:718
 - comparisons with Shanghai, China and Estonia 5:712–716
 - learn for life 5:711–716
 - national framework for teaching and learning 5:714–715
 - preparing future-ready teachers for future-ready learners 5:716–718
 - Singapore's education phases and initiatives 5:711
- pedagogy 5:387
- phenomenographic studies 5:440
- policy and professionalism 5:10–19
- practices 9:234–235
- preserves status quo of whiteness 5:476–477
- and professionalism
 - context and comparison in 5:72–73
- forms of governance and professionalism 5:73–75
- ITE for primary school teachers in Italy 5:77–78
- local hierarchies and market of employers in French ITE 5:76–79
- methodological note 5:75–76
- school networks for ITE in England 5:75–76
- race-evasive approaches to white student resistance in 2:284–285
- racial literacy in 2:285–286
- racism and “overwhelming presence of whiteness” in 5:475–477
- reflective practice in 9:137–138
- relevance of TE effect research 5:437
- research long-term effects of educating teachers 5:439–441
- research on 5:448–449
- research reviews 5:439
- researcher 5:481–482
- as researchers 5:449
- self-study as methodology for improvement of 5:460–461
- seminal studies 5:438
- for social change 2:283–288
- for social justice, tension in 5:692
- suggestions for conducting TE effect studies 5:441–442
- transition from classroom teacher to 5:448
- validity requirements 5:441
- value of research within 5:612
- teacher engagement (TE) 4:503–506
- teacher supply model (TSM) 5:657
- limitations of 5:657
- teachers turnover intention (TTI) 4:504–506
- teaching 10:494, 13:66, 4:375–376, 8:94–95, 8:281–283
- academic habitus of PhD students 8:431–433
- content integration 6:358–359
- critical approaches to teaching Africanist presence in literature 10:228–229
- curating case-based knowledge system for 5:410
- dialogue 6:604–605
- disciplinary literacies 10:54–55
- educating teaching profession
 - characteristics of professions and semi-professions 5:369
- in England's academy schools 5:40–43
- expertise
 - adaptive expertise 6:181–182
 - collective expertise 6:183–184
 - deliberate practice 6:180–181
 - ecologies of expertise 6:184–185
 - ecologies of individual expertise 6:184
 - exceptional performance 6:180–182
 - knowledge building 6:182–183
 - multilevel conceptualizations of expertise 6:185
 - multiplicity of expertises 6:185
 - networked expertise 6:183
 - relational expertise 6:183–184
 - synthesis and methodological implications 6:185–187
- exploring case-based knowledge system for 5:409–410
- friendship as form of negative ethics in 5:47
- graduate courses
 - accessibility 3:139
 - bicultural accountability 3:137
 - challenges 3:139
 - course co-construction 3:137–138
 - final paper 3:138
 - graduate school as professionalization process 3:137
 - Indian education for all 3:136–137
 - presentations/posters 3:138
- as human enterprise/image of teaching as big data, image of 7:346
- images of 7:341
- implications for 5:644–645
- and learning, democratic approach to 5:702
- metaphors of 7:341–342
- method 11:67
- Mongolian STEM teachers reflecting on use of Minzu knowledges in 3:222–226
- narrative hermeneutics for 5:119
- national framework for 5:714–715
- need to claim and advance indigenous peoples' processes of 3:230–231
- point and purpose of 5:649–650
- as research-based profession 5:613–614
- scholarship of 8:275
- social media in 8:132–133
- source evaluation 6:359
- standardization of 5:32–33
- standardized teaching 5:23
- and transfer effects of 21st century skills (TEACH21st) 13:252
- 21st century teaching and learning 7:352–354
- understanding importance of complexity and context in 8:259–260
- Teaching and Learning International Survey (TALIS) 13:325
- 2018 video study 13:330
- age of enlightenment to 5:605–606
- teaching performance assessments (TPA) 5:427
 - common design components of 5:429–432
 - differences and distinctions 5:431–432
 - focus and nature of TPA credential 5:432
 - general characteristics, processes and practices associated with 5:430–431
 - issues and challenges for TPAs and TPA research 5:432–433
 - lesson plans and key instructional resources and materials 5:430–431
 - moderation and assessment processes are bespoke and contextual 5:432
 - perspectives on 5:428–429
 - policy context in initial teacher education programs 5:427–428
 - range of components and cycles/processes are used 5:431–432
 - videos 5:430
 - whole class assessment 5:430
 - work samples from three students 5:430
- team-based learning (TBL) 6:709

- technical and vocational education training (TVET) 1:631, 1:634–635
- technical services of libraries 8:125
- technology 10:719–720, 13:44–45, 2:341–343, 6:608
- characteristics of technology-enhanced items 14:172–173
 - developing teachers' use of 5:274–275
 - utilizing process data from technology-based assessments 13:334–335
 - as double-edged sword 10:793
 - emergence of technology-based internships 8:464–465
 - expanding role of technology in education 10:821–822
 - adaptability through improved data capture 10:821–822
 - assessment 10:822
 - content development 10:821
 - inclusion 10:821
 - student agency and empowerment 10:821
 - hard and soft governance in era of 7:78
 - literacy 11:44–45
 - preparing participants for technology-based workplaces 8:466
 - support language acquisition 10:792–793
 - teacher skills and knowledge for technology integration 5:263–271
 - teachers' use of 5:273–274
 - technology-enhanced formative assessment 13:7
- technology transfer offices (TTOs) 8:14–15
- Technology-Enhanced Collaborative Learning (TECL) 11:207
- in STEM education 11:208–211
 - in STEM, assess and analyze 11:211–212
- TECL. *See* Technology-Enhanced Collaborative Learning (TECL)
- Tehran's public spaces, women's perceptions and experiences in 12:649–650
- template-based AIG 14:194–195
- See also* non-template-based AIG and example items
 - evaluation of 14:196
 - example of 14:195–196
- temporary mobility for complete study program 1:242
- tensions in access and accountability
- access on peripheries 8:370
 - accounting for access 8:369–370
 - epistemic access and accountability 8:371–372
 - fees and funding 8:368–369
 - governance and evaluation 8:370–371
 - widening access 8:367
- tentative typology of 20th century whole system school reforms 1:166–168
- TePP project
- background and relevance of 12:629
 - central line of MM design in 12:629–631
- tertiary education
- in low-and middle-income countries
 - challenges to delivering tertiary education during pandemic 9:127–128
 - globalization traps 9:132–133
 - lesson from high-income countries 9:131
 - lessons from massification of basic education 9:131
 - massification trap 9:131
 - opportunity trap 9:131–132
 - multiple recent shifts in tertiary education landscape 9:126–127
 - pathways 8:440–441
 - into tertiary education 8:445
 - within tertiary education 8:445–446
 - tertiary funding and effects on pathways, approaches to 8:446–447
- test assembly 14:201–208
- adaptive test assembly 14:206–207
 - as adjunct to manual assembly 14:207
 - as combinatorial optimization problem 14:202–205
 - constraints 14:202
 - limitations 14:207
 - objective functions 14:202–205
 - simultaneous assembly of multiple forms or item blocks 14:206
 - solution algorithms 14:205–206
 - test blueprints 14:201–202
 - use cases 14:206–207
- test blueprints 14:201–202
- test development, science of 14:11
- test fraud
- software for detection of 14:301
 - and statistical approaches to detect types 14:298–301
- Test of Science-Related Attitudes (TOSRA) 11:378
- test-based accountability 5:34
- Test-based assessment (TBA) 13:307
- drivers of 13:306–308
- test-taking motivation 14:215
- testamonal narrative 10:730
- testing 13:193
- effects of memory 6:110
 - for homogeneity 14:851
 - tests of normality 14:813
 - tetrachoric correlation 14:746
 - text based reports 12:723
- text matters, choice of 10:559
- textbook 10:569
- textual curriculum 7:263
- textual variations
- book club instigations 10:914–915
 - Pikachu bonding 10:914
 - recognizing affect and textual variations 10:914–916
 - words 10:915–916
- Thailand
- cultural context 4:492–493
 - education context in 4:491–492
 - findings 4:493–497
 - identification of sources 4:493
 - review method 4:493
 - successful principal practices in 4:496–497
 - successful school principalship in 4:491
 - Thailand context 4:491–493
- THE. *See* Times Higher Education (THE)
- thematic analysis 12:304–305
- Theoharis, leadership frameworks and taxonomies 4:406
- theopraxis 10:521–522
- theoretical frameworks, values as manifest in 12:8
- theoretical traditions 10:320–322
- critical literacy 10:320–321
 - digital multimodal literacy 10:321–322
 - global literacies 10:321
- theoretically-grounded educational data mining 14:411
- theoreticians 3:184–185
- theorizing learning 6:46–47
- theory 13:80
- of justice 10:219–222
 - of peer assessment 13:80
 - of self-assessment 13:80
 - theory-centric mixed methods research designs 12:431
 - theory-ladenness 10:889
 - theory-practice problem 5:371
- theory, methods, and practice (TMP) 10:904
- thick mapping, urban humanities 10:954–955
- thinking together dialogic approach 6:605
- "third space" model 5:385
- and transformation 10:150–151
 - transformation and polyphonic identities 10:151
- Thorndike's social intelligence 6:200–201
- three-dimensional (3D)
- learning 11:444–445
 - narrative inquiry space 12:124–125
 - pedagogical framework 11:217–221
 - augmentation 11:218–219
 - collaborative technology 11:220
 - degree to technology is integrated with learning activities 11:217–219
 - ethical approval 11:222
 - learning outcomes varied at different levels of mobile integration 11:219
 - pedagogical roles students and teachers play 11:219–221
 - remarks 11:221–222
 - student perceptions 11:221
 - student-led and collaborative functions
 - outperform teacher-led functions on student achievements 11:221
 - student-led technology 11:219–220
 - substitution 11:218
 - teacher-led technology 11:220
- 3PL model correlated factors extension 14:90
- 3P model, entanglement in 8:79
- three-step latent class analysis 14:641
- through-year assessments 13:212
- additional goals 13:212–213
 - argument-based approach to validation of through-year assessments 13:215–218
 - argument-based approach to validity and validation 13:214

- call for evidence 13:219
 claims and evidence 13:213–214
 current examples 13:212
 validity argument 13:218–219
 Thurstone's primary mental abilities and
 advent of factor analysis 6:200
 Tiktok
 critiques of production and boundaries
 hyper-school 10:934
 student challenges to school space on
 10:935
 time-series designs 14:700
 time-series modeling 14:691–692
 Times Higher Education (THE) 2:261, 2:263
 TIMSS. *See* trends in international
 mathematics and science study
 (TIMSS)
 TISA. *See* Trade in Services Agreement (TISA)
 TL. *See* target language (TL); transfer learning
 (TL); transformational leadership
 (TL)
 TMP. *See* theory, methods, and practice
 (TMP)
 Toronto Teachers' Federation (TTF) 2:385
 backlash 2:393–395
 bilingual education 2:388–389
 daytime HL classes 2:389–395
 early experiments in Toronto and new
 policies in Ontario 2:387–388
 engagement with work group 2:391–392
 historical and policy context 2:386–388
 theoretical and methodological framing
 2:386
 Work Group on Third Language Instruction
 2:389–391
 TOSRA. *See* Test of Science-Related Attitudes
 (TOSRA)
 Tostan, HRE in community development
 2:62
 TPA. *See* teaching performance assessments
 (TPA)
 trace data of learning process 13:57–59
 tracking developing theory 12:113–114
 Trade in Services Agreement (TISA) 1:507
 trade unions 1:577–578
 and teaching profession 1:577–578
 traditional education 3:24
 traditional reading instruction 10:531–532
 training model 5:470
 portion of standard setting meeting 14:377
 trans affirming 2:360
 trans-culturally relevant pedagogy 5:678
 transactional leadership 4:46
 transactional realism 7:80–81, 7:92–93
 transcendental arguments 10:888
 transculturalism and contemporary
 transnationalism 1:46–47
 transdisciplinary approach
 critical cosmopolitanism literacy and
 10:453–454
 disciplinary literacy and content area
 literacy 10:449–451
 future forecast 10:449
 history and epistemological issues
 10:449–451
 ongoing research in adolescent disciplinary
 literacy 10:451–452
 post-secondary disciplinary literacy
 10:452–453
 transdisciplinary literacy and deep
 comprehension 10:452
 vignette 10:448–449
 transfer credit 8:345–348
 transfer learning (TL) 14:692
 transformation-based integration, using
 software for 12:507–508
 transformational leadership (TL) 4:75
 illustrating integration of 4:75
 transformational school leadership (TSL)
 4:27, 4:46–47, 4:54–55, 4:64
 developing people 4:55–56
 holding high performance expectations
 4:55
 impact
 on school conditions 4:56–57
 on students 4:57–58
 on teachers 4:57
 improving instruction 4:56
 managing and securing accountability 4:56
 origin and major developments of 4:54
 redesigning organization 4:56
 setting directions 4:55
 shared vision and building goal consensus
 4:55
 typical models and practices of 4:55–56
 transformative aspects of assessment in 21st
 century 13:253–254
 transformative mixed methods 12:430–431,
 5:470
 and evaluation 12:534–535
 MMR designs 12:430
 research 12:402–409
 application of 12:404–408
 assumptions associated with
 12:402–404
 disabilities and deafness 12:406–407
 implications for teacher research
 12:408–409
 indigenous populations 12:407–408
 transformative teacher education for
 educational equity 5:361–364
 transformative-emancipation 12:385–386
 transformer model 8:60–61
 Transforming Disability Knowledge
 Research and Activism (TDKRA)
 2:126
 project 2:126–127
 Transgender Day of Remembrance 2:360
 transgender studies in education 9:165–166
 transition-to-trained-doctor model (T3D
 model) 6:537–539, 6:539f
 transitions
 conceptualization of 6:530–531
 defining successful 6:532
 successful multiple and multi-dimensional
 transitions 6:539–541
 theoretical perspectives 6:532–539
 translanguaging 10:149, 2:191–192
 affordances of named languages 10:156
 conceptualizing 10:149
 conceptualizing identity 10:149
 identity, and pedagogical context
 10:154–156
 named languages, translanguaging, and
 productive unsettling of identity
 categories 10:149–150
 translation and adaptation 14:231
 importance of 14:230–231
 inappropriate uses of 14:232–233
 ITC guidelines 14:233–234
 poor translation and adaptation 14:232
 process of 14:231–232
 statistical measures of comparability 14:234
 validity and construct irrelevance 14:231
 translational research process 12:589–590
 action research 12:593
 adaptation 12:591–592
 advantages and challenges of intersecting
 mixed methods with 12:596–597
 conceptual model for intersecting mixed
 methods with 12:594–596
 dissemination and implementation
 12:592–593
 evidence-based practice 12:590–591
 methodological components of
 12:590–594
 translational science 12:589–590
 translingual literacy practices and social
 justice 10:107
 transliteracy 10:130–131
 definition of 10:130–131
 epistemology and core concepts and themes
 10:595–597
 ethics of 10:133–134
 as frontier field 10:593–597
 and stakes 10:593–594
 teacher toolkit 10:131–132
 for transformative education 10:600–601
 as translingual, transcultural, and
 cosmopolitan practice 10:132–133
 transmedial narratives 10:629–630
 transmediation 10:504–505
 transmodal language pedagogies 10:108
 transnational academic collaboration
 in co-authorship, ethical implications for
 1:312–314
 organisational culture and social
 psychology of 1:311–312
 politics of scientometrics and datafication
 process 1:309–310
 transnational curriculum 7:57–59
 transnational examples 10:809–816
 transnational feminist approach 2:364–365
 transnational higher education 1:245–246
 transnational literacies 10:143–144
 transnationalism 10:810–811
 formations of 1:47–48
 problematics of transnationalism as 21st
 century concept 1:45–46
 transculturalism and contemporary
 transnationalism 1:46–47
 transnationalism as activator of rights in
 education 1:49–50
 transnationalism nexus 1:48–49
 transparency 4:514–515

trauma informed classrooms
 closing circle with feathers of hope 3:38–40
 journey of un-learning to re-learn2 3:36–38
 opening circle with our “fallen feathers” 3:32–34
 trauma, development and classrooms 3:34–36
 trends in international mathematics and science study (TIMSS) 14:348
 1995 video studies 13:329–330
 1999 video studies 13:329–330
 trends in learning technology development 5:264–265
 triangulation 12:351–352
 American study on phenomenology 12:353–356
 developments in literature on triangulation 12:352–353
 innovative UK research on care-related technologies 12:356–358
 tripartite collaboration 4:365
 trust 13:45
 truth and reconciliation
 at limits of settler colonialism 2:274
 before reconciliation within higher education 2:275
 with teachers 2:278–279
 TSL. *See* transformational school leadership (TSL)
 TSM. *See* teacher supply model (TSM)
 TTF. *See* Toronto Teachers’ Federation (TTF)
 TTI. *See* teachers turnover intention (TTI)
 TTOs. *See* technology transfer offices (TTOs)
 Tucker
 case of 5:672–674
 leadership frameworks and taxonomies 4:405
 building professional capacity 4:405
 connecting with external partners 4:405
 creating supportive organization for learning 4:405
 establishing and conveying vision 4:405
 facilitating high-quality learning experience for students 4:405
 turnaround leaders 4:159–161
 antecedents and factors affecting turnaround 4:163
 changing leadership as key to turnaround 4:162
 conceptual framework 4:158
 creating collaborative school culture and structure 4:163
 data sources 4:159
 district leadership in turnaround contexts 4:165
 external conditions 4:161
 findings 4:159–166
 internal conditions 4:160–161
 leadership 4:161–162
 practices in school turnaround 4:164
 method 4:159
 principals’ role in school turnaround 4:163–164
 as process 4:161
 rationale for leadership changes 4:162

restraints on 4:162–163
 turnaround leadership competence 4:165
 TVET. *See* technical and vocational education training (TVET)
 21st century, DBE 11:199–200
 21st century competencies
 assessing conceptual knowledge competencies 13:317
 assessing ICT-related competencies 13:316
 cognitive competencies 13:314–315
 conceptual competencies 13:314
 domains of knowledge 13:314
 implications for future assessment practice 13:319–320
 international commitments to 13:315
 interpersonal competencies 13:315
 metacognitive competencies 13:315
 personal competencies 13:315
 progress in assessing 13:315–319
 specifying 13:313–315
 21st century skills
 conceptualizing 21st century skills 7:50
 in curriculum 7:49
 in curriculum frameworks and standards 7:51–52
 in school curricula 7:52
 two-mode networks and multilevel networks 4:532
 2PL model Bifactor extension 14:92
 2PL model correlated factors extension 14:89–90
 two-tier model 14:92–93

U

U-curve 5:196–197
 U3As. *See* universities of Third Age (U3As)
 UB. *See* University of Botswana (UB)
 ubiquitous learning 6:263
 ubuntu storytelling 12:181–182
 UCF. *See* University of Central Florida (UCF)
 UD in Education (UDE) 11:329
 framework 11:329–330
 guidelines and practices in STEM education 11:331–332
 UDE. *See* UD in Education (UDE)
 UDL. *See* Universal Design For Learning (UDL)
 UEB. *See* Unified English Braille (UEB)
 UK
 disability and poverty in 9:466–467
 research on care-related technologies 12:356–358
 UN. *See* United Nations (UN)
 un-learning
 language ideology 10:830–831
 to re-learn2 3:36–38
 Unapologetically Black Day 2:360
 UNCRC. *See* United Nations Convention on Rights of Child (UNCRC)
 uncritical allies 5:418
 undocumented activism
 activism, advocacy, allyship, usurping 2:378
 activism spectrum 2:377–378
 analysis 2:381–383
 architecting DREAM 2:375
 Bate-papo and pláticas 2:376–377
 Daniella’s testimonio 2:378–381
 existence as resistance 2:375
 felons not families 2:375
 guiding frameworks 2:377–378
 interpreting signs of university’s commitment to diversity 2:383
 legacy of progress 2:383
 literature overview 2:375
 methodology 2:376–377
 perceiving “wins” 2:382–383
 positionality 2:377
 sacrifice in frontline activism 2:381–382
 testimonio 2:377
 Trump administration 2:376
 undocumented activism globally 2:376
 usurping moment and movement 2:382
 UNEG. *See* United Nations Evaluation Group (UNEG)
 UNESCO. *See* United Nations Educational, Scientific and Cultural Organization (UNESCO)
 UNICEF. *See* United Nations Children’s Fund (UNICEF)
 Unified English Braille (UEB) 10:425–427
 uniform DIF 14:250–251
 uniform distribution 14:710–711
 unit deviance 14:584–585
 unit learning goals 11:446–447
 unitary approach, dangers of 10:282–283
 United Nations (UN) 7:406
 United Nations Children’s Fund (UNICEF) 1:488
 construction
 of global childhood education space 1:490–491
 of global children’s policy agency 1:489
 education for new world order 1:488–489
 organisational configurations in times of change 1:490
 UNICEF and politics of global childhood 1:491
 United Nations Convention on Rights of Child (UNCRC) 9:86
 United Nations Educational, Scientific and Cultural Organization (UNESCO) 1:510
 assertion of UNESCO’s global education discourse 1:517–521
 contrasting goals of 1:391–392
 education for democratic redevelopment 1:513–514
 global education goal setting 1:514–515
 intellectual trajectories 1:513–517
 knowledge and learning 1:518–519
 negotiating politics of multilateralism and global education governance 1:521–523
 positioning UNESCO’s educational agenda 1:511–513
 realizing education vision 1:519–520
 role 1:520–521
 setting problem 1:517–518

- United Nations Evaluation Group (UNEG) 13:268
- United States 3:15–16
- academic freedom and US constitution 8:176–177
 - adjusted group means 14:541
 - African, indigenous, and Spanish literacies 10:269
 - anti-black racism and white supremacy 10:270–272
 - assessment reform and school reform in US context 5:583–584
 - assumptions and limitations 14:542–543
 - codifying racism 10:274
 - codifying anti-literacy through US courts 10:270
 - college students, tactics, issues, and campaigns of 2:402–403
 - comparing group means 14:539
 - contemporary critical accounts of schools in 12:50
 - covariate balance, overlap and ANCOVA adjustment 14:541–542
 - data demonstration 14:538, 14:540–541
 - debates about family literacy in 10:184
 - family literacy implementation in 10:183–186
 - federal assessment policy reforms
 - assessment landscape in US 13:108–111
 - benefits for special populations 13:112–113
 - broader assessment system, impact of accountability assessments on 13:115–116
 - federal assessment policies, negative impacts of 13:113
 - special populations, negative impacts on 13:113–114
 - teacher evaluations, impact of accountability assessments on 13:114–115
 - federal investment in career pathways in 8:472–473
 - gender differences on US National Assessments 10:255
 - increasingly fragile state of tenure in 8:177–178
 - leadership development in 4:370–371
 - looking forward 10:276–277
 - mathematics and science achievement in 6:404
 - principal preparation 4:362–363
 - roots of literacy in United States 10:268–273
 - state court cases 10:276
 - systemic racism
 - in education 10:275
 - in politics and education 10:273–276
 - teacher
 - accountability in 5:64–65
 - learned about indigenous children and schools in Guatemala 10:139–140
 - universities and academic freedom 8:175–178
- US-China tensions univariate ANCOVA 1:663, 14:537–543
- See also* multivariate ANCOVA
- univariate categorical data, describing 14:723–724
- univariate numerical data, describing 14:724–729
- univariate $p \times i$ design 14:65–68
- universal critical literacy and revolutionary love 10:233
- universal design approach to access 11:328–329
- Universal Design For Learning (UDL) 6:712–713, 7:440–442, 9:334–336
- and access to curriculum 7:441–442
 - design cycle 6:715–716
 - embracing learner variability 6:713–714
 - flipped, blended and online 6:717
 - as instructional design framework 6:715–718
 - interpreting UDL in practice 7:442
 - national case studies 7:443–444
 - as orientation 6:713–714
 - origins of 7:440–441
 - policy
 - borrowing and 7:442–443
 - travel, UDL, and education for all 7:442–443
 - progressing forward with 9:338–339
 - reconsidering beliefs about the capabilities of all learners 6:714
 - reframing of barriers to learning and relocating “problem” fixed 6:714
 - reporting criteria 6:718
 - rising to equity 6:715
 - social model approach 9:336–338
- universal design in STEM education 11:326–333
- universities 10:302, 3:110–111
- core missions of 8:94–95
 - economic impact 8:95–96
 - effect on inequalities within faculty and student demographics 8:95–96
 - outreach 8:95
 - research 8:95
 - teaching and learning 8:94–95
 - in East Asia through globalization 8:25–27
 - faculty in South-East Asia 1:423
 - as national development project 1:381–382
 - transformation of university missions
 - implantation and development of modern university 8:22–25
 - transformation of concept of “university” in East Asia through globalization 8:25–27
- universities of Third Age (U3As) 6:508–509
- demographics, participation and participants’ interests at OLLIs and 6:509
- University of Botswana (UB) 1:107
- University of Central Florida (UCF) 1:107
- university-district/school partnerships 4:389
- university–school partnerships
- benefits of 5:382–383
 - equal partners 5:389
 - key considerations for universities and schools 5:387–390
 - models of partnership 5:383–385
 - oversight and responsibility 5:388
 - professional inquiry 5:390
 - purpose of 5:382
 - shared vision 5:388
 - sustainability 5:389
 - teacher education pedagogy 5:387
 - teacher mentors 5:389
 - time 5:388
 - trust 5:388
- unknown probability estimation 14:578–580
- unraveling assessment as phenomenon in education 13:103–104
- unstructured natural language data, analytics of 6:77–79
- unsupervised learning approach 14:410, 14:685, 6:85
- Urban humanities 10:956
- banking *vs.* dialogic models of education 10:951
 - developing urban humanities in situ 10:949–951
 - humanities studio 10:951
 - humans in city 10:948–949
 - life in “undercommons” 10:950–951
 - practical future 10:953–954
 - practices of 10:954–955
 - spatial justice 10:953
 - urbanism as collective life 10:952
 - values of urban humanities 10:952–954
- Urban youth literacies 10:742–743
- Urbanism as collective life 10:952
- utilization-focused evaluation 12:158
- utopias 1:603–604
- ## V
- validating multilingualism 10:128
- validity
- argument 13:218–219
 - argument-based approach to validity 14:36–40
 - consequences 14:40–42
 - design 12:516
 - O’Cathain 12:519
 - Onwuegbuzie and Johnson 12:518–519
 - selective aspects of designing quality criteria 12:519–520
 - selective empirical examples of 12:518–519
 - Tashakkori 12:519
- fairness as validity 14:35–36
- linking values 14:36
- simulation-and game-based assessment 14:472
- theory 14:9–10
- testing practice 14:10–11
- threats 14:36
- value-added models (VAMs) 14:386–387, 14:390

- value-added models (VAMs) (*continued*)
 advantages and disadvantages 14:387
 alternative modeling approaches 14:393
 bias 14:393–394
 of conceptual framework for learning in disciplines 6:312
 concerns about precision and stability of estimates 14:394–395
 determining growth adequacy 14:387
 econometric models 14:391–392
 features of 14:391
 leading VAMs approaches 14:391–393
 limitations and concerns with application 14:393–395
 modeling requirements 14:387
 non-random assignment 14:393–394
 persistence of inputs to achievement 14:394
 scaling and limitations of achievement measures 14:394
 statistical modeling 14:392
 two-step VAMs 14:392
 value-centered current 2:211–212
 values 6:28–29
 valuing multilingualism 10:128
 VAMs. *See* value-added models (VAMs)
 variability
 measures of 14:725–726
 reporting measures of 14:726
 variance components
 D study VCs 14:61–62
 G study VCs 14:60–61
 Vary-One-Thing-at-a-Time strategy (VOTAT strategy) 6:643
 venture philanthropy 1:587
 vernaculars 10:468–469
 vertical layout for visual displays 12:766
 vertical mobility and objectives of mobility 1:242
 VET systems. *See* vocational education and training systems (VET systems)
 vicarious learning 6:292
 use of 11:392
 vicarious processes 6:25–26
 educational research 6:26
 functions of modeling 6:25
 observational learning 6:26
 video feedback 1:153
 video games 10:440–442
 autism, social communication, and literacies of 10:440–441
 conventional and functional literacy skills 10:442
 multimodal potential of gaming literacies for autism 10:441–442
 multimodal potential through gaming literacies 10:441
 social communication and literacy motives and benefits of 10:441
 social communication through literacies of 10:440
 violated distribution conditions 14:812–813
 violence against young women and women 2:366
 virtual communities, learning in 2:327–328
 virtual reality (VR) 11:189, 6:637
 content appropriateness 11:195, 6:645
 details in 11:195
 immersion and interactivity two key factors 11:191
 and knowledge generation 11:193–194, 6:640–641
 laboratory environment 11:192
 laboratory environment 6:639–644
 life like details in VR simulations 6:644
 methods 11:194, 6:641
 results 11:194–195, 6:641–643
 in STEM education 11:190, 6:638–639
 STEM practice development and difficulty appropriateness 11:195
 technological savvy 11:195, 6:644
 use case 11:192–194, 6:639
 visual displays
 arrows use 12:767
 building 12:769
 and communicating
 mixed methods study design procedure 12:764–768
 relationship between research questions and data collection methods 12:763–764
 concise language 12:768
 connecting 12:768–769
 draw boxes for various quantitative and qualitative stages 12:766
 embedding 12:771
 expected products or outcomes 12:767–768
 generate title for 12:764–765
 horizontal or vertical layout for 12:766
 and integration
 of findings 12:771–772
 of methods 12:768–771
 letter case to designate priority 12:766–767
 merging 12:769–770
 in mixed methods research 12:763
 procedures for data collection and analysis 12:767
 visual display simple 12:768
 visual media representations 14:176–179
 dynamic visual media 14:179
 in education and assessment 14:176
 graphs and charts 14:178–179
 representational and decorative pictures 14:176–178
 visual regime of global teacher education reform 1:267–268
 Vivian Vasquez 2:196
 vocabulary curricula, need for 10:369–370
 vocabulary interventions 10:367–368
 benefit from 10:367–368
 core vocabulary 10:369
 development of databases and assessments 10:370–371
 need for vocabulary curricula 10:369–370
 research might support efficacious 10:368–371
 vocabulary unique to literary and informational texts 10:369–370
 vocabulary learning 10:361–362
 vocabulary spurt 6:469–470
 vocabulary unique to literary and informational texts 10:369–370
 vocational education 1:547–552
 adaptability and 8:453–454
 curriculum and pedagogic practices 8:455–460
 designable elements affording developing vocational knowledge 7:210–211
 developing vocational
 identity 7:211–212
 knowledge 7:210–211
 educational goals 8:454–455
 goals for and processes of 8:454–455
 job readiness as well as occupational preparation 8:454
 learning to act in vocational practice 7:209–210
 results 7:209–212
 review process 7:208
 securing “hard to learn” knowledge 8:455
 vocational education and training systems (VET systems) 8:438, 9:450
 approaches to revitalize Australian 8:447
 perceptions of VET and higher education and effects on pathways 8:447
 vocational identity
 designable elements affording developing 7:212
 developing vocational identity 7:211–212
 epistemic elements 7:212
 instrumental elements 7:212
 learning processes and activities 7:211–212
 social elements 7:212
 spatial elements 7:212
 temporal elements 7:212
 vocational knowledge
 designable elements affording developing 7:210–211
 developing vocational knowledge 7:210–211
 epistemic elements 7:210–211
 instrumental elements 7:211
 learning processes and activities 7:210
 social elements 7:211
 spatial elements 7:211
 temporal elements 7:211
 vocational practice
 designable elements affording learning to act in 7:209–210
 epistemic elements 7:209
 instrumental elements 7:210
 social elements 7:210
 spatial elements 7:209
 temporal elements 7:210
 vocational training 1:547–552
 VOTAT strategy. *See* Vary-One-Thing-at-a-Time strategy (VOTAT strategy)
 VR. *See* virtual reality (VR)
 Vuong 8:394–395

W

W. E. B. *See* William Edward Burghardt (W. E. B.)

- Wabanaki Youth Science program (WaYS program) 2:334–335
- WAIC. *See* Watanabe-Akaike information criterion (WAIC)
- Watanabe-Akaike information criterion (WAIC) 14:287
- water 3:208–209
- WaYS program. *See* Wabanaki Youth Science program (WaYS program)
- weaving thrivance 3:177–178
- web-based qualitative research methods, ontological imperative in 12:294
- empirical demonstration 12:295–299
- researcher in 12:299–300
- web-based technologies for data collection 12:693–701
- weighting approaches 14:808
- weighting/matching designs 14:700
- well-being 6:721–723
- future perspective on well-being and learning 6:727
- and learning in school 6:723–727
- predictors and outcomes of 6:722–723
- in psychology and education 6:721–722
- Western education 3:26–27
- Western incursions 1:659–660
- Western psychology
- colonial tool of 3:173–174
- cultural appropriation within 3:174–175
- Western science knowledge systems 3:199
- Western settler colonialism 3:199–203
- different perspectives for fire management 3:201
- impact of diseases 3:199–200
- exclusion of indigenous land management 3:200
- indigenous fire and environmental management 3:201–203
- protection *vs.* sustainable resource utilization 3:200–201
- settler greed 3:200
- shifting land management 3:200
- Westernized education 3:24
- What Works Clearinghouse (WWC) 8:473
- meta-analysis on career pathways 8:473–474
- what you test is what you get (WYTIWYG) 13:227
- Whatu as teacher 3:186–187
- Whatuora
- as methodology 3:186
- radical decolonizing methodology 3:186
- white saviorism 5:418
- white supremacy 10:270–272
- white woman
- historical archetype meets present day 2:98–101
- historical construction of 2:96–98
- as ideal type 2:94–101
- whiteness 3:110–111
- in teacher education, overwhelming presence of 5:476
- whole-group discussion 14:224–225
- whole-of-class facilitation 8:272
- William Edward Burghardt (W. E. B.) 14:833–835
- within school stratification
- basic issue and comprehensive high school 9:80–81
- curriculum stratification within secondary schooling 9:80–81
- early schooling and ability grouping 9:77–79
- internationally 9:82
- transition into and out of primary schools 9:79–80
- WM. *See* working memory (WM)
- women(s)
- in agroecology leadership for conservation, Guatemala 2:333
- conceptual framework 2:324
- in higher education
- in but still out 8:384–386
- conceptual clarity 8:383–384
- curveball of Covid-19 8:388–389
- daily bread of being black in academia 8:387–388
- looking back 8:384
- mirage of work/life balance 8:386–387
- politics of being feminist researcher 8:382–383
- structural framework 8:381–382
- winning ways 8:389–390
- intersecting gender and ethnic identities 4:431
- and leadership 4:430–431
- leadership in Malaysia 4:485
- research institute 8:208
- for women's human rights 2:62
- word(s)
- benefit from vocabulary interventions 10:367–368
- features of words influence students' word knowledge 10:366–367
- groups of words might benefit from instruction 10:367
- research
- might support efficacious vocabulary instruction 10:368–371
- shown about types of words 10:367–368
- on word features clarify words to teach 10:366–367
- segmentation 6:468–469
- size of lexicon students will encounter in school texts 10:366
- spurt 6:469–470
- work
- body of work method 14:377–378
- intensification 5:638–639
- and learning
- changing world of work 2:236–237
- conventional approaches to 2:237–238
- critical approaches to 2:238–240
- practice-based approaches to 2:240
- studies of 2:237–240
- learning through 6:381
- explaining 6:382–383
- personal epistemological practice 6:386–388
- practice curriculum, pedagogies, and epistemologies 6:383–386
- securing occupational capacities through practice 6:388
- with marginalized populations 6:553–554
- Work Group on Third Language Instruction 2:389–391
- work-integrated learning, expansion of internships to 8:465
- workforce development model 10:456–457
- working conditions 4:450–451
- working memory (WM) 14:173, 6:221
- capacity 6:108–109
- and constraints of human information processing 14:173–174
- workplaces 10:302
- curriculum 6:384–385
- pathways of work 6:384–385
- focusing experiences on workplace-readiness competencies 8:467
- learning 5:564–568
- teachers' workplace 5:564–565
- workplace conditions in schools 5:565–568
- workplace-based learning 8:462–468
- world bank
- after 2008 1:484–485
- contrasting goals of 1:391–392
- emergence of WB as prominent actor in education 1:480–482
- in global education arena 1:485–486
- and post-Washington consensus 1:482–484
- world culture theory 1:466–468
- world Englishes
- ideological domain 10:827
- and neighboring/overlapping paradigms 10:825–826
- and pedagogy 10:826–832
- practical domain 10:828–829
- responding to critique of 10:832–833
- scholarly domain 10:829
- sociopolitical and structural domain 10:827–828
- stepping stones toward changing paradigm 10:829–832
- stumbling blocks 10:827–829
- World Trade Organization (WTO) 1:503
- early history of 1:503–504
- education 1:505–506
- forum shifting 1:506–507
- TISA 1:507
- trade related aspects of intellectual property rights 1:507–508
- world-Earth de-distancing 1:602
- world-systems analysis (WSA) 1:19, 7:96
- and education for just, utopistic, transition 1:25–26
- in educational research 1:23–25
- historical social science 7:98–102
- major concepts 1:20–23
- origins 1:19–20
- school curriculum 7:102–103
- structures of knowledge 7:96–98
- writing 10:27–28, 10:32–36, 10:831, 6:403

writing (*continued*)

- assessments method 4:456–457
- children under age three learning about 10:29–32
- contexts for observing two-year-olds learning to write 10:28
- formal/informal *vs.* 10:530
 - meaning attribution and 10:43–45
- forms and directional patterns 10:30
- foundational understandings 10:29–30
- identity 10:32
- intervention class 10:565–566
- language
- participation 10:28–29
- pedagogy 10:66
 - Australia's EAL/D policy and context 10:68
 - and English language learners 10:66–67
 - English language learners' writing 10:67, 10:72
 - and learning across disciplines 10:66
 - multilingualism, multiculturalism, and meaning making 10:71–72
 - spoken word in context 10:71
 - supporting English language learners in secondary schools 10:67–70
 - teaching and learning cycle 10:69–70
- reimagining research on
 - counternarrative 10:960
 - follows 10:960–961
 - helps 10:961–962
 - no counternarrative 10:961
 - beyond orthography 10:752–753
 - researchers imagined writing 10:959
 - rhizomatic mapping 10:958–959
 - shows 10:960
 - writing does 10:962–963
- system in China 10:747–753
- up qualitative research
 - approaches and methods/techniques 12:724–725
 - finding right journal or publisher 12:725
 - forms, formats, purposes and standards 12:722–723
 - getting started 12:721–722
 - quality checklists and guidelines 12:724–725
 - recently developed methods 12:722–723
 - reviewing 12:726–727

submitting and being reviewed

12:725–726

text based reports 12:723

WSA. *See* world-systems analysis (WSA)WTO. *See* World Trade Organization (WTO)WWC. *See* What Works Clearinghouse (WWC)WYTIWYG. *See* what you test is what you get (WYTIWYG)**Y**

yarning type of conversation 12:182

Yin yang philosophy 12:387

young adult

fiction 10:815

MMT of 6:535

young children invent Swahili pidgin in rural Kenya 10:799–800

young emergent bilingual children, assessment of 10:96

approaches to learning 10:100

call for comprehensive 10:96–102

cognitive development 10:99

direct language assessment 10:98

family and early care experiences 10:100–101

language and early literacy development 10:97–98

literacy assessment 10:99

observational approach 10:98

parent and teacher reports 10:97

physical development 10:100

social and emotional development 10:100

societal and policy contexts 10:101–102

young people

boundaries of literature for 10:606–607

literature for 10:606–607

young women and activism

dissent, dissidence, and (global) citizenship from below 2:370–371

environmental issues—climate justice & land sovereignty 2:368

issues—racial justice activism 2:367–368

issues—reproductive justice, SOGI, and HIV/AIDS 2:366–367

issues—violence against young women and women 2:366

modes of activism among young women 2:368–370

notes on transnational feminist approach 2:364–365

(re)framing young women and activism 2:363–364

snapshot of contemporary young women and activism—issues 2:365–368

youth 10:606–607, 2:341–343

audiences 10:242, 10:245

as educators 2:337–345

electoral study 2:134–135

literacies 10:242–243

literature 10:811–812

out-of-school literacies 10:242

as perpetual educators 2:340–343

texts 10:242, 10:244

Youth Participatory Action Research projects (YPAR projects) 10:232

and critical literacies for social action 10:232–233

YouTube

as global connective platform 1:623–624

as knowledge platform 1:624–625

students' uses of 1:625–626

teachers' uses of 1:626–627

videos for school study 1:625–627

YPAR projects. *See* Youth Participatory

Action Research projects (YPAR projects)

Yukl, leadership frameworks and

taxonomies 4:404

Z

Zambia, education in developing countries 1:209

ZIKA rapid response project 8:208

Zimbabwe

financing education in 9:415–416

inclusive education in 9:410–414

inclusive education policies and practices 9:404–417

zone of proximal development (ZPD) 6:41, 6:369, 6:479–480, 6:677

ZPD. *See* zone of proximal development (ZPD)

AUTHOR INDEX

A

- Abdelhay, Ashraf
Regimes of literacy as regimes of truth about Africa: language ideologies and southern voices
- Abdmolaei, Shirin
Global testing regimes: accountability, standardization, racial inequality and the rise of audit culture
- Abedi, Jamal
Assessment of English learner students
- Aberdein, Rachel
Girls, autism identities, and education
- Abiodun-Ekus, Onyedikachi
The politics of literacy languaging and texts: the politics of literacy in Nigeria.
- About Ghaida, Susanne K.
Educational concerns in Arabic picture books
- Abramowitz, Brian
Family influence and STEM career aspirations
- Ackerman, Terry
Item Response Theory
- Adah Miller, E.C.
Using flexible thinking to assess student sensemaking of phenomena in project-based learning
- Adams, Donnie
Women school leadership in Malaysia: perspectives and enactments
- Adamson, David
Learning analytics
- Adeyemo, Kolawole Samuel
The disruption of Black history and critical pedagogy in international higher education curriculum
- Adhikari, Samrachana
Modeling of network structures
- Agbashe, Olaide
Uses of technology in navigating the pandemic: the case of African higher education: Chika Sehoole, Lydia Ikwemesim, Olaide Agbashe
- Ahaotu, Joseph Onyema
The politics of literacy languaging and texts: the politics of literacy in Nigeria.
- Ahenakew, Cash
Power/knowledge: global and Indigenous education
- Ainscow, Mel
Collaboration as a strategy for promoting equity in education
Inclusion and equity: an agenda for educational reform
- Ajjawi, Rola
Feedback for learning
- Akbulut Yıldırım, Meltem
Higher education strategy formations in shifting landscape of higher education
- Akin-Sabuncu, Sibel
Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms
- Akkari, Abdeljalil
Impact of globalization on higher education in the Middle East and North Africa
- Akshay, Nagarajan
STEAM for all: a vision for STEM and arts integration
- Al-Fadala, Asmaa
Framing educational leadership as a *multilevel distributed* practice: a systemwide perspective

- Al-Nawab, Hadeel F.
Researching curriculum and disability in higher education
- Albers, Craig A.
Research study design
- Alexander, Patricia A.
Learning across the lifespan: an overview
- Allal, Linda
Assessment and the regulation of learning
- Allan, Julie
Legislation and policy in Europe
- Allen, Rosie
From grit and resilience to academic tenacity
- Allen, Walter R.
Issues of production and utilization of knowledge: multiple uses of knowledge
- Allington, Richard L.
The meaning(s) of dyslexia
- Allotey, Eugenia
International perspectives on community-engaged teacher education
- Allsop, Yvonne
Key developments during adolescence: implications for learning and achievement
- Alonzo, Alicia C.
Learning progressions: affordances, critiques, and pathways toward equity
- Alrawashdeh, Ghaida S.
Language choice, technology and international education
- Alvarez-Arellano, Lilian
Children's literature in Latin America today: a historical perspective
- Alvermann, Donna E.
Reimagining research on writing: a timely rhizomatic argument
- Alves, Ines
Public policies in Portuguese education: the path to inclusion for all
- Alvunger, Daniel
Curriculum making and teacher agency
- Ames, Allison J.
Simulation studies for psychometrics
- Amevuvor, Jocelyn
The scholarship of African Children's literature: issues of representation of critical voices in the conversation
- Anastasiou, Dimitris
Recent debates in special & inclusive education
- Anderman, Eric M.
Key developments during adolescence: implications for learning and achievement
- Andersen, Hanne
Publish or perish
- Anderson, Richard C.
Collaborative Reasoning and children's development of language and thought
- Anderson, Jim
Family literacies: differing perspectives
- Anderson, Ann
Family literacies: differing perspectives
- Anderson, Jami L.
Intersecting mixed methods with translational research: implications for educational research and practice
- Anderson, Carolyn J.
Categorical data analysis
- Andreotti, Vanessa
Power/knowledge: global and Indigenous education
- Andrews-Todd, Jessica
Assessment of collaborative problem solving skills
- Annala, Johanna
Curriculum making and higher education: an activity-theoretical perspective
- Anson, Emily
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
- Aoyama, Ryosuke
On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT

- Appelbaum, Peter
The rise of STEM education: mathematics learning for STEM
- Archibald Q'um Q'um Xiem, Jo-ann
Indigenous storywork for education and research
- Ardron, Kevin
Document analysis
- Arena, Peter
Economic impact of higher education on students, corporations and society
- Armstrong, Sonya L.
College reading
- Armstrong, David
Educational trends and exclusion
- Arndt, Petra A.
Embodied cognition
- Arnold, Richard
South-South Cooperation in Education
- Aronsson, K.
Learning through play-in-interaction
- Arrington, Laura E.
A constructivist lens of eye-tracking in special populations
- Arslan, Burcu
Cognition-centered design principles for digital assessment tasks and items
The current trends and opportunities for machine learning in learning analytics
- Artiles, Alfredo J.
Learning disabilities' unsettling intersections: disrupting learning, cultural, socioeconomic, and environmental deficits
- Artopoulos, Alejandro
Knowledge economy meets development imaginaries
- Aryee-Price, Awo Okaikor
From #BlackLivesMatter to Black Lives Matter at school: how the fight for Black Liberation entered the schoolhouse
- Asikuk Topkok, Sean
Developing transforming research—The University of Alaska Fairbanks Center for Cross-Cultural Studies as case study
- Aston, Ruth
Teaching is influenced by the teacher
- Ávalos, Beatrice
Teacher policy frames, enactment forms and issues
- Avery, Helen
Education of refugees
- Ayoo, Philip O.
The impact of COVID-19 on higher education in Africa
- Azzopardi, Andrew
Building inclusive communities
- ## B
- Baartman, Liesbeth
Learning environments in vocational education: how to elicit learning for and in practice
- Babchuk, Wayne A.
Grounded theory
Philosophical and theoretical underpinnings of qualitative research
- Bagheri, Nazgol
Using mixed methods research with Geographic Information Systems (GIS)
- Baijnath, Narend
Implications of the Fourth Industrial Revolution in the vortex of higher education in South Africa
- Baizhanov, Sanzhar
Global university rankings and geopolitics of knowledge
- Bakan, Abigail B.
Citizenship
- Baker, Jim
Education International and its global contribution to educational debates
- Baker, Bernadette
From capacity to ability to automation: “Western” conceptions of the figure of man and ableist subjectivities

- Baker, Betsy A.
Epistemologies and methodologies in literacy research: a case study
- Baldrige, Bianca J.
Community-based education
- Baldwin, Doug
Common item formats and evaluation practices
- Bale, Jeff
Language, race, and legitimacy: the Toronto Teachers' Federation's battle against heritage-language programs in Toronto schools, 1982–1986
- Ball, Stephen J.
The neoliberal dispositif: understanding the transformation of the social and ethical space of education
- Ball, Arnetha F.
Teacher education for diverse learners
- Bamberger, Annette
Diaspora and internationalization of higher education: a critical approach
- Barney, Katelyn
The importance of “seeing oneself” in Australian universities: representations and reflections of Indigeneity in higher education
- Barrot, Jessie S.
Language education
- Barrueco, Sandra
Assessment of young emergent bilingual children
- Bartlett, Lora
Teachers' work in the context of COVID-19
- Barton, Keith C.
Learning in history and social studies
- Basarkod, Geetanjali
Academic self-concept: a key predictor of achievement and learning
- Basit, Tehmina N.
“It was a life enhancing activity”: the academic habitus of PhD students
- Basna, Kristyna
Erasmus(+) student mobility: individual and institutional motivations and effects
- Batkie, Ryan
Socioscientific issues and STEM learning
- Batra, Poonam
Reinventing teacher agency: shifting the narrative of teacher education
- Bauer, Scott C.
Leadership contexts and how they matter: leaders' working conditions
- Baumgartner, Lisa M.
Older adults' development, learning, and education
- Bayerlein, Leopold
The future of internships and workplace-based learning
- Beach, Richard
Responding to and creating multimodal texts
- Beadle, Rosalind
Responding to cultural diversity through action and community-based participatory approaches in qualitative research
- Bean, Thomas W.
Shifting from disciplinary versus content area literacy to a transdisciplinary approach
- Becker, Betsy Jane
Meta-analysis
- Beckett, Angharad E.
“But tell me, where do the children play?” Disabled children and the right to play
- Beech, Jason
Policy transfer in education
Volume 1: Globalization and the Shifting Geopolitics of Education
- Beijaard, Douwe
Beginning teachers' professional identity work and its support
- Bekerman, Zvi
Peace education in conflict zones: general trends and Israeli particularities
- Benavot, Aaron
Global Education Monitoring Report
- Bennett, Randy E.
Automated scoring of constructed-response items in educational assessment

- Benozzo, Angelo
Qualitative sampling and qualitative data
- Bergen, Jenn K.
Contemporary challenges and approaches in anti-racist teacher education
- Bernstein, Jared
Artificial intelligence in literacy
- Best, Marnie
Re-examining the education for all agenda in “apocalyptic” times
- Bhana, Deevia
Comprehensive sexuality education: young masculinity, cultural relevance and context matters
- Bianchi-Bustos, Marcelo Emilio
Children’s literature in Latin America today: a historical perspective
- Biernacki, Paulina
Test translation and adaptation
- Biesta, Gert
Reclaiming the artistry of teaching
- Billett, Stephen
Learning through work: augmenting vocational educational experiences to meet 21st-century work requirements
Learning through work: premises, conceptions and practices
- Billot-Vasquez, Kiana
Issues of production and utilization of knowledge: multiple uses of knowledge
- Biraimah, Karen L.
Enhancing a Global South-North partnership to promote global justice in K-12 classrooms: a Botswana case study
- Birbili, Maria
Teachers mediating of curriculum making processes
- Blackburn, Mollie V.
Literacy at the intersections of race, gender, and sexuality
- Blair, Julie
Weaving together mind, body and spirit: envisioning pathways to trauma-informed classrooms from Indigenous perspectives
- Blaisdell, Nicolas
Systemic functional linguistic and disciplinary literacy development in multilingual K-12 schools in the United States
- Bleiberg, Joshua
Revisiting teacher evaluation a decade after reforms
- Bliuc, Ana-Maria
Innovations in university student learning research into blended course designs: incorporating self-report and observational data
- Bloome, David
Literacies, language & schooling: teaching and learning the language arts
- Bo, Yalong
Understanding turnaround leadership research: continuity and change (2017–2021)
- Bobis, Janette
Situating teacher learning in the mathematics classroom and everyday practice
- Bodine Al-Sharif, Mary Ann
Critical race theory and educational research utilizing qualitative methods
- Boler, Megan
Media education: history, frameworks, debates and challenges
- Bonal, Xavier
The World Bank and education in a changing global arena
- Bong, Mimi
Self-efficacy
- Boos, Dennis D.
Resampling methods
- Bordewieck, Kathleen E.
Family influence and STEM career aspirations
- Borko, Hilda
Current trends, tensions and unresolved issues in research on teacher professional learning
- Borrero, Noah
Teachers living the curriculum: culturally relevant pedagogy and action research

- Boswell, Eileen
Grounded theory
Philosophical and theoretical underpinnings of qualitative research
- Boucher, Nicole S.
The efficacy of learning strategies from the ICAP perspective
- Bovill, Catherine
Students and staff co-creating curriculum in higher education
- Bowzer, Angela D.
A brief introduction to contemporary standard setting methods
- Boykin, Allison Ames
Model-data fit evaluation: posterior checks and Bayesian model selection
- Bradfield, Abraham
The importance of “seeing oneself” in Australian universities: representations and reflections of Indigeneity in higher education
- Bradford, Brent
Principles of inclusive practice in schools: what is COVID-19 teaching us?
- Bradshaw, Victoria
Establishing quality in qualitative research: trustworthiness, validity, and a lack of consensus
- Bradshaw, Laine
Diagnostic classification modeling
- Bragg, Debra D.
Career pathways
- Brain, Sam
Reviews of education for students with disability in Australia
- Bramwell, Daniela
School leadership in Latin America: research and policy trends
- Branson, Christopher M.
The ethical imperative in educational leadership
- Bråten, Ivar
Learning from multiple texts
- Braun, Henry I.
Advantages and challenges of performance assessment of student learning in higher education
- Bray, Mark
Shadow education
- Brennen, Tim
Memory
- Bringle, Robert G.
Higher education: service-learning as pedagogy, partnership, institutional organization, and change strategy
- Brinkmann, Svend
Qualitative interviews
- Brissett, Nigel O.M.
The education Sustainable Development Goal 4: a critical appraisal
- Broatch, Jennifer E.
Descriptive statistics
Statistics: Value-Added Models
- Brock-Utne, Birgit
English as a medium of instruction and research in education: an international and comparative analysis
- Brooks, Clare
Teachers’ research capacity and initial teacher education policy
- Broucker, Bruno
About the rise, the characteristics and future of New Public Management in Higher Education
- Brouwer, C.N.
Determining long term effects of teacher education
- Brown, Bailey A.
Regimes of research and development funding in higher education
- Brown, Gavin T.L.
Leading school-based assessment for educational improvement: rethinking accountability
- Brunetti, Matthew A.
A brief introduction to contemporary standard setting methods

- Bryan, Victoria M.
Intelligences
 - Bryant, Darren A.
International patterns and trajectories in principal preparation
 - Bubb, Sara
Understanding what makes school leaders effective: the importance of personal development and support
 - Bullock, Emma P.
Mixed methods research design practices to address complexity in education
 - Bunda, Tracey
The importance of “seeing oneself” in Australian universities: representations and reflections of Indigeneity in higher education
 - Bunn, Matthew
Tensions in access and accountability
 - Burch, Patricia
For profit corporations and education reform: global dimensions and impacts
 - Burgos, Daniel
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis
 - Burgstahler, Sheryl E.
Universal design in STEM education
 - Burke, Penny Jane
Tensions in access and accountability
 - Burkhardt, Hugh
Assessment and mathematical literacy: a brief introduction
 - Burn, Katharine
Can initial teacher education partnerships maintain their epistemological and pedagogical underpinnings in light of market-driven reforms? The case of England
 - Burnheimer, Eric
Cooperative and collaborative learning
 - Burrell-Kim, Danielle
Digital games and literacies
 - Bush, Tony
Leadership practices and why they matter: evolving leadership models
 - Byun, Kiyong
Regionalization of higher education and geopolitics in Northeast Asia
- C**
- Cai, Yuzhuo
Policy isomorphism: a lens for understanding the influence of globalization on national education policymaking
 - Cai, Jinfa
Mathematics in STEM education
 - Cai, Li
Multidimensional item response theory
 - Cain, Leia K.
Establishing quality in qualitative research: trustworthiness, validity, and a lack of consensus
Promising practices for addressing unique ethical dilemmas arising from integration in mixed methods research
 - Calle-Díaz, Luzkarime
Literacy, colonialism and resistance in Latin America
 - Calma-Birling, Destany
Executive function and education
 - Cambourne, Brian L.
The literacy journey: from a psychological to a biological way of thinking about learning to read and write
 - Camicia, Steven
Curriculum and LGBTQ+ inclusion
Curriculum and social studies education
 - Campano, Gerald
Mobilizing prefigurative hybrid literacies for decolonial resistance: three transnational examples
 - Cannon, Michelle
Digital media arts and critical media education: an alchemy of creative literacy practices
 - Carless, David
Feedback for student learning in higher education

- Carpenter, Shana K.
Optimizing learning through retrieval practice and spacing
- Carrillo, Carmen
Online teaching and learning practices in teacher education: past, present and future
- Carrim, Nazir
Diversity
- Carrington, Suzanne
Girls, autism identities, and education
Leadership for equity in schools
- Carroll, Tim
Economic impact of higher education on students, corporations and society
- Carson, Jerome
From grit and resilience to academic tenacity
- Carvalho, Lucila
Learning spaces and materiality: embracing past and present when designing for the future
- Casabianca, Jodi M.
Rater effects modeling
- Casely-Hayford, Jeffrey
The role of job demands and resources for teacher job satisfaction: insights from five European countries
- Casinader, Niranjan
Transnationalism
- Castaneda, Lillian Vega
Uses and abuses of social media in higher education
- Castaneda, Mario E.
Uses and abuses of social media in higher education
- Castek, Jill
Expansive, dynamic, and critical new literacies: widening views of digital and multimodal learning
- Castillo, Nathan M.
Language choice, technology and international education
- Castrillón, Gloria
“Reschooling” higher education: the dynamics of curriculum in South Africa
- Cenoz, Jasone
Multilingualism at school and multilingual education
- Cervetti, Gina N.
Reading comprehension in troubled times
- Chajewski, Michael
Classical test theory
- Chang, Chih Ching
Dialogic education: tensions and dilemmas
- Chang, TingWen
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis
- Chapman, Amy L.
Social media and learning
- Charalambous, Evi
International trends on teacher evaluation
- Chasi, Samia
North–South partnerships in higher education
- Chataika, Tsitsi
Inclusive education policies and practices in Malawi and Zimbabwe
- Chauhan, Purvee
Promoting engineering research with entrepreneurship and industry experiences: a teacher professional development program
- Chen, Zhe
Making the public private: the privatization of public education
- Chen, I.C.
Using flexible thinking to assess student sensemaking of phenomena in project-based learning
- Chen, Dandan
Categorical data analysis
- Chesworth, Liz
Children’s interests and curriculum making in early childhood education
- Chi, Michelene T.H.
The efficacy of learning strategies from the ICAP perspective

- Ching, Kimiko
Key developments during adolescence: implications for learning and achievement
- Chisnall, Georgia
Rapid qualitative research
- Chitiga, Miriam
The disruption of Black history and critical pedagogy in international higher education curriculum
- Choi, Gi Woong
Design-based education in STEM: for learners of the 21st century
- Choi, Yoonjung
Curriculum, citizenship, migration
- Chu-Fuluifaga, Cherie
Minority learners, diversity, and educational outcomes: the case of Aotearoa, New Zealand
- Chung, Simmee
Narrative inquiry
- Cian, Heidi D.
Family influence and STEM career aspirations
- Cintron, Dakota W.
Bayesian networks
- Cizek, Gregory J.
Formative assessment: an overview of history, theory and application
- Clandinin, D. Jean
Narrative inquiry
- Clark, Eve V.
Language development in the early years
- Clark, Heather F.
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program
- Clarke, Matthew
Defiant imaginaries: from a politics of identity to an ethics of friendship in teaching
- Clarke, Anthony
Teacher inquiry: by any other name
- Clauser, Brian E.
Past, present, and future of educational measurement
- Clayton, Patti H.
Higher education: service-learning as pedagogy, partnership, institutional organization, and change strategy
- Clinton, Janet
Teaching is influenced by the teacher
- Coburn, Cynthia E.
School leaders' use of research: viewing research use in decision making through an organizational lens
- Coburn, Elaine
Social justice and education: addressing inequities in schooling
- Codere, S.K.
Using flexible thinking to assess student sensemaking of phenomena in project-based learning
- Coelho, Carla Ruthes
Curriculum and identity options
- Cohen, Dan
Spatial theories, methods and education policy
- Colditz, Jason
Intersecting mixed methods with phenomenology
- Coles, Alf
Researching teacher learning through self-video
- Collins, Riley
Teachers' work in the context of COVID-19
- Collins, Kathleen M.T.
Role of quality criteria in integrative mixed methods research
- Cologon, Kathy
Considerations for implementing Universal Design for Learning: toward anti-oppressive pedagogies
- Colwell, Jamie
Epistemologies and methodologies in literacy research: a case study
- Comber, Barbara
Critical literacies & social justice: social justice approaches in schools
Primary and secondary education: literacy and social justice

- Conklin, Hilary G.
Preparing novice teacher educators for critical, justice-oriented teacher education practice
- Conner, Jerusha O.
Social movements and college student activism
- Connor, David J.
The concept of belonging and other areas of interest in disability studies in education
- Cook-Sather, Alison
Reciprocal teaching and learning through student-faculty pedagogical partnership
- Cook, Loraine D.
Cultural studies in education: insights for the mixed methods researcher
- Cope, Bill
Contemporary contexts for learning: an overview
Inclusive pedagogies, new learning, new times
- Corcoran, Tim
Attitudes toward inclusion
- Corson, Jordan
Leadership and governance for improvement: how schools manage external demands
- Corter, James E.
Clustering approaches
- Cortez Ochoa, Artemio Arturo
Review of evaluation approaches for school principals
- Cowen, Robert
Comparative education: then, now, and next
- Craig, Cheryl J.
Challenges to lived curriculum making: 1990–2020
- Crawford, Jenifer
Critical theory in qualitative research in education
- Creamer, Elizabeth G.
Integration in mixed method research
Mixed methods and grounded theory
- Creamer, Elizabeth
Promising practices for addressing unique ethical dilemmas arising from integration in mixed methods research
- Crisman, Jonathan Jae-an
Urban humanities as experimental research praxis
- Crocker, Jonathan W.
Education and participatory democracy
- Crouch, Debra K.
The literacy journey: from a psychological to a biological way of thinking about learning to read and write
- Croydon, Abigail E.
Disability and employment: finding a way to work
- Cruz, Cindy
Youth as educators
- Cun, Aijuan
Posthumanism and Positioning Theory: explorations of embodiment and affect from a critical perspective
- Cunningham, James W.
Literacy models and epistemology
- Curt Fulwider, G.
Machine learning
- Curwood, Jen Scott
Literacies, language, and schooling: exploring writing pedagogy for English language learners
- Czerniawski, Gerry
Professional development or professional learning: developing teacher educators' professional expertise
- D**
- Daly, Alan J.
Networked systems leadership: the potential of social network theory and analysis
- Dalyot, Keren E.
Citizen science in STEM education: engaging students with real life science
- Damacio, Bri
Issues of production and utilization of knowledge: multiple uses of knowledge
- Damico, Holly
A constructivist lens of eye-tracking in special populations
- Damico, Jack S.
A constructivist lens of eye-tracking in special populations

- Danesi, Marcel
Semiotics in the context of qualitative research
- Danish, Joshua A.
What is learning, for whom, and to what end? An overview
- Dare, Emily A.
Measurement instruments of STEM affective learning: a systematic review
STEM curriculum development and implementation
- Darwich, Lina
Teachers' work in the context of COVID-19
- Da'as, Rima'a
Principal's information-processing mechanisms (PIPMs) as an effective tool for learning-centered leadership that reduces teacher's intent to leave
- Datnow, Amanda
Evidence-informed decision making in schools and districts: the role of equity
School reform from a constructivist perspective
- Davey, Tim
Automated test assembly
- Davis, Allison Crean
Mixed methods in survey research
- Dawson, Georgia
Teaching is influenced by the teacher
- Dawson, Phillip
Feedback for learning
- Day, Christopher
Research on successful school leadership
Teachers willingness and ability to teach to their best and well: the person in the professional
- Daza, Viviana
Inside looking out: a third space perspective on pre-service teachers' practice experiences
- De Boeck, Paul
Explanatory measurement
- de Bruijn, Elly
Learning environments in vocational education: how to elicit learning for and in practice
- DeCuir-Gunby, Jessica T.
Mixed methods research proposal
- Dei, George J.S.
Afrocentric education: traditions, politics and possibilities
- de Jong, Ton
Simulations and games for learning
- Delahunty, Thomas
Neurocognitive and physiological measurement of STEM learning processes
- Deliman, Amanda
Biliteracy, playfulness and performative practices
- Dell, Laura
Photovoice with children and youth
- Delmas, Peggy M.
Qualitative research approaches and their application in education
- DeLuca, Christopher
Implementation of Assessment for Learning
- Dempsey, Majella
Curriculum and teacher education: the pre-service teacher as curriculum maker
- Dempster, Neil
Changing the context of Australian Indigenous education through leadership for learning "both ways"
- Dervin, Fred
"Indigenous" ways of knowing in Chinese Minzu education: teachers reflecting on their engagement with STEM in inner Mongolia
Chinese Minzu (民族) diversity and education
- Desai, Aditi
UNESCO: promising education multilateralism?
- Deschambault, Ryan
On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT
- De Schauwer, Elisabeth
Legislation and policy in Europe
- Desimoni, Victoria
Education in small island (and micro) states
South-South Cooperation in Education

- Devitt, Adam
Design-based pedagogy: a heuristic for developing math and science teachers' professional practice
- Dey, Joyeeta
The mangle of contemporary geopolitics
- Dharamshi, Pooja
A lifetime of experience: an international study on the lives of teacher educators:
- DiBenedetto, Maria K.
Learning from a social cognitive theory perspective
- DiCerbo, Kristen
Simulation- and game-based assessment
- Dickson, Elizabeth A.
Social, educational and legal perspectives on assessment and accountability measures in inclusive education for students with disability
- Dillon, Deborah R.
Digital literacies, multimodality and the engagement of adolescents who are disengaged or disenfranchised in school
- Ding, Meixia
Mathematics in STEM education
- Dobson, Tom
Diverse landscapes of school governing body practices
- Dohmen, Dieter
New pathways in education
- Dolma, Tenzin
Education for all and MDGs: global education policy translation in India
- Domaleski, Chris
Educational testing: a review of standards, guidelines, and practices
- Dong, Nianbo
Foundational methods: power analysis
- Donnelley-Power, Caitlin
Education in small island (and micro) states
- Donnelly, Verity J.
Paying for inclusion and exclusion
- Donovan, Courtney
Mixed methods sampling and data collection
- Dorans, Neil J.
Score equating: an aspirational form of score linking
- Dou, Remy
Family influence and STEM career aspirations
- Doucette, Janice
Learning to count what counts: narrative hermeneutics for teachers and teaching
- Douglas, Daniel
STEM education and workforce development: the history, politics, and evidence
- Down, Barry
Early career teachers' work
- Drisko, James
Writing up qualitative research
- Duckor, Brent
Examining conceptual frameworks for understanding teacher learning in the domain of formative assessment
- Dudley-Marling, Curt
A social/cultural constructivist perspective on learning "problems"
- Duncan Bonokoski, Arden A.
Inclusive postsecondary education: vision, practice, higher education reimaged
- Dunn, Peter K.
Generalized linear models
- Dunning, Darren L.
Mindfulness and learning
- Dunphy, Julie
Socioscientific issues and STEM learning
- Dupuy, Beatrice C.
Literacy/multiliteracies and foreign language education
- Durksen, Tracy L.
Teacher recruitment: a competency-based approach
- Dussel, Inés
Global connective media: YouTube as an educational infrastructure
- Dutt, Khaleda Gani
The role of job demands and resources for teacher job satisfaction: insights from five European countries

Dvořák, Dominik
Curriculum development

Dwyer, Bernadette
E-texts and e-books: constructing meaning with and from texts in an online and digital world

E

Earley, Peter
Understanding what makes school leaders effective: the importance of personal development and support

East, Martin
Learning an additional language

Eberle, Julia
Apprenticeship learning

Eckert, Jonathan
Leading together: developing teacher leadership

Edmiston, Brian
Humanizing language and literacy in education with dramatic inquiry

Edmondson, Macey L.
Academic freedom protections in national and international law

Edwards-Schuth, Brandon
Ecocritical pedagogies and curriculum

Edwards, D. Brent,
The work of international organizations: ten theoretical approaches and a way forward

Edwards, Miriam
Researching curriculum and disability in higher education

Edwards, Patricia A.
Toward culturally relevant literacies with children and families of color

Egbe, Gabriel B.
Englishization and the language of education in Nigeria: literacy issues, challenges and hopes

Ek, Lucila D.
Literacy and identity across home, school, church, and social life

Ell, Fiona R.
The contribution of initial teacher education to social justice

Elliot, Norbert
Fairness

Ellis, Sue
Teacher identity: crossing the technical-rationalist and affective divide

Ellis, Robert A.
Innovations in university student learning research into blended course designs: incorporating self-report and observational data

Elmore, Jacob
International perspectives on STEM teacher induction

Elwood, Jannette
Assessment reform and students' voices

Emil, Serap
Higher education strategy formations in shifting landscape of higher education

Emons, Wilco
Nonparametric statistical methods

Enciso, Patricia
Literature: critical imagination and young adult readers remaking worlds

English, Leona M.
Labor markets, SDG4 and vocational education and training

Engsig, Thomas T.
Nordic representations and paradoxes of inclusive education

Ennes, Megan E.
Family influence and STEM career aspirations

Ennser-Kananen, Johanna
Internationalization in English? Problematizing the role of language in higher education internationalization

Erduran, Sibel
Learning to teach epistemic practices of science in secondary schools

Erstad, Ola
Educational assessment of 21st century skills—novel initiatives, yet a lack of systemic transformation

- Esboldt, Joy
White woman: or, the abused abuser's role in U.S. educational stratification
- Escobar, Martha
Self-efficacy in STEM
- Espinoza, Óscar
Six decades of development: the impact of structural adjustment programs on education in developing countries
- Eustis, Jennifer M.
The changing functions of libraries in higher education
- Evans, Karen
Lifelong learning beyond initial schooling
- Ewing, Robyn Ann
Teacher education and the arts: challenges, tensions, opportunities
- Exley, Beryl
Sensory literacies: the full sensorium in literacy learning
- Ezike, Nnamdi C.
Model-data fit evaluation: posterior checks and Bayesian model selection
- F**
- Fairchild, Nikki
Qualitative sampling and qualitative data
- Faulconer, Emily K.
The scholarship of teaching and learning: cases from one STEM department
- Favero, Luca
Erasmus(+) student mobility: individual and institutional motivations and effects
- Felder, Marion
Recent debates in special & inclusive education
- Feldman, Jennifer
University–school partnerships in initial teacher education
- Fernandes, Domingos
Public policies in Portuguese education: the path to inclusion for all
- Fernandez, Maria Beatriz
Teacher policy frames, enactment forms and issues
- Fernandez, Frank J.
Academic freedom protections in national and international law
- Ferrante, Patricia
Global connective media: YouTube as an educational infrastructure
- Ferraz, Raul
Model-data fit evaluation: item fit and model selection
- Fetters, Michael D.
Mixed methods research designs
- Fiacco, James
Learning analytics
- Files, Martez D.
Critical race theory and educational research utilizing qualitative methods
- Filippakou, Ourania
Higher education, critical pedagogy, and the challenge of resistance
- Finkelstein, Idit
Principal's information-processing mechanisms (PIPMs) as an effective tool for learning-centered leadership that reduces teacher's intent to leave
- Firoozi, Tahereh
Automatic item generation
- Firouzshahi, Narges
A constructivist lens of eye-tracking in special populations
- Fischer, Frank
Perspectives on learning from the learning sciences
- Fischer, Jennifer
Advantages and challenges of performance assessment of student learning in higher education
- Fisher, Douglas
Arts-based practices for literacy & language learning
- FitzGerald, Anne Marie
Leadership and family factors: critical perspectives on international leadership standards, models, and practices
- Fitzgerald, Sarah R.
The changing functions of libraries in higher education

- Flessa, Joseph
School leadership in Latin America: research and policy trends
- Flores, Maria Assunção
Online teaching and learning practices in teacher education: past, present and future
Studying teachers' work and lives: insights from research before and during the COVID-19 pandemic
- Flückiger, Bev
Changing the context of Australian Indigenous education through leadership for learning "both ways"
- Fontdevila, Clara
The World Bank and education in a changing global arena
The privatization of education. Drivers, social effects and regulatory challenges of private sector participation in schooling
- Foote, Lori A.
Combining mixed methods and case study research
- Forsyth, Carol M.
Assessment of collaborative problem solving skills
Conversation-based assessment: current findings and future work
The current trends and opportunities for machine learning in learning analytics
- Forzani, Elena
Expansive, dynamic, and critical new literacies: widening views of digital and multimodal learning
- Frau-Meigs, Divina
Transliteracy and the digital media: theorizing media and information literacy
- Fray, Leanne
Questioning the consensus on effective professional development
- Fredericks, Bronwyn
The importance of "seeing oneself" in Australian universities: representations and reflections of Indigeneity in higher education
- Frenzel, Anne C.
Emotions in education
- Frey, Nancy
Arts-based practices for literacy & language learning
- Frey, Andreas
Computerized adaptive testing and multistage testing
- Friddle, Karole-Ann
Expanding meaningfulness for struggling readers
- Fridrich, Monika
Connecting play to STEM concepts, practices and processes: review of research on play within STEM learning environments
- Friedrich, Daniel
Teach for all and the visual regime of global teacher education reform
- Friendly, Michael
Displaying information: from woolly mammoths to the Great Migration
- Friesen, Sharon
Learning to count what counts: narrative hermeneutics for teachers and teaching
- Friesen, Norm
Continental pedagogy & curriculum
Phenomenology and education: researching pedagogical experience
- Frieson, Brittany L.
Rethinking multiliteracies: toward a radical vision for multilingualism
- Froehlich, Dominik E.
Mixed methods and social network analysis
- Frost, Nollaig
Intersecting mixed methods with phenomenology
- Fránquiz, María E.
Funds of knowledge and funds of identity: lenses for building language and literacy
- Fuentealba Jara, Rodrigo
Self-study of teacher education practices: the complex challenges of learning from experience
- Fusenig, Jannah
Learning across the lifespan: an overview
- G**
- Gabriel, Rachael
The meaning(s) of dyslexia

- Gaertner, Matthew N.
A brief introduction to contemporary standard setting methods
- Gale, Trevor
The problem with higher education student completions in the global north
- Galla, Candace Kaleimamoowahinekapu
(Just) languaging with *duende*: (re)inventing and innovating communicative practices as resistance, healing, and thrivance in Indigenous contexts
Community-centered Indigenous language recovery, restoration, revitalization, and renewal
- Gallagher, Deborah J.
Looking for something outside of ourselves: the politics of researching race and dis/Ability and (special) education
- Gallo, Michele
PCA and other dimensionality-reduction techniques
- Galvin, Sarah M.
Social media and learning
- Gambino, Andrea
Critical media literacy and second language acquisition
- Gamlem, Siv M.
Feedback and classroom practice
- García-Carrión, Rocío
International perspectives on community-engaged teacher education
- Garg, Rabani
Cosmopolitan literacies and the limits of understanding: possible directions in global literacy research
- Gasman, Marybeth
Historically Black colleges and universities: historical and contemporary issues
- Gebhard, Meg
Systemic functional linguistic and disciplinary literacy development in multilingual K-12 schools in the United States
- Geertsema, Johan
Teacher expertise and how it develops during teachers' professional lives
- Gendron, Danielle
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
- Gerber, Hannah R.
The ontological imperative in conducting qualitative research in online spaces
- Gharib, Hoda
Media education: history, frameworks, debates and challenges
- Gholson, Melissa
U.S. federal assessment policy reforms in the twenty first century and their impacts
- Ghosh, Ratna
Education, identity and diversity
- Giannotti, Maureen
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program
- Gibson, Randall
Photovoice with children and youth
- Gibson, Colin M.
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
- Gidra, Aanchal
Language choice, technology and international education
- Gierl, Mark J.
Automatic item generation
- Gilbert, Christine
Reclaiming accountability
- Giles, Rachel L.
Qualitative research approaches and their application in education
- Gilmore, Perry
(Just) languaging with *duende*: (re)inventing and innovating communicative practices as resistance, healing, and thrivance in Indigenous contexts
- Giroux, Henry A.
Higher education, critical pedagogy, and the challenge of resistance
- Glaesser, Judith
Curriculum and competences

- Glanfield, Florence
Imagining STEM *with* and *as* place
- Gobert, Janice D.
Intelligent tutoring systems: a history and an example of an ITS for science
- Gocer Sahin, Sakine
Differential functioning
- Godfrey, David
Schools as research-informed professional learning organizations
- Goetz, Thomas
Emotions in education
- Goldman, Susan R.
Learning in the disciplines: a conceptual framework
- Golumbic, Yaela N.
Citizen science in STEM education: engaging students with real life science
- Gomer, Brenna
Missing data analysis
- Gong, Byoung-gyu
Redefining educational transfer and borrowing in the pluriverse
- Goodley, Dan
Disability studies in a posthuman age
- Goodwin, A. Lin
Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms
- Goodyear, Peter
Networks and learning
- Gorard, Stephen
Tackling teacher shortages: doing the right thing at the right time and with the right evidence
- Gordon, Anna Lise
Early career teaching and resilience
- Gore, Jennifer
Questioning the consensus on effective professional development
- Gorter, Durk
Multilingualism at school and multilingual education
- Gorur, Radhika
The mangle of contemporary geopolitics
- Gothberg, June E.
Reporting mixed methods research studies
- Govender, Navan N.
Teacher identity: crossing the technical-rationalist and affective divide
- Graesser, Arthur C.
Intelligent tutoring systems and conversational agents
- Gravel, Jenna W.
Universal Design for Learning in its 3rd decade: a focus on equity, inclusion, and design
- Green, Jennifer L.
Descriptive statistics
Statistics: Value-Added Models
- Greene, Andrew
Teach for all and the visual regime of global teacher education reform
- Greenhow, Christine M.
Social media and learning
- Greenwood, Janinka
Arts-based research: inquiry through, in, and as art
- Gregory, Kristen
Shifting from disciplinary versus content area literacy to a transdisciplinary approach
- Greiff, Samuel
Assessment of collaborative problem-solving: past achievements and current challenges
Interdisciplinary frontiers: computer-based process data analysis in educational measurement
- Grey, Sue
Educational accountability, international large scale assessments and PISA: powerful influences based on weak foundations
- Grice, Christine
Research on successful school leadership
- Grieshaber, Susan
Conceptualizing professional development with and for a diverse and globally mobile early childhood education workforce

- Griffith, John
The scholarship of teaching and learning: cases from one STEM department
- Griffiths, Tom G.
World-systems analysis and curriculum theory
World-systems analysis
- Griffiths, Miro
Everybody's got to learn sometime: disability youth activism and the pursuit for inclusive education
- Gu, Qing
Leadership and policy: how principals of successful schools enact education policy for improvement
The challenge to quality retention: why schools matter
Volume 4: Administration, Governance, and Leadership
- Guðjónsdóttir, Hafþís
Enacting inclusive pedagogy in teacher education: creating a learning space for teachers to develop their professional identity
- Guerra, Myriam Jimena
Literacy and identity across home, school, church, and social life
- Guerrero, Tricia A.
Fundamental processes of learning: an overview
- Guetterman, Timothy C.
Joint displays in mixed methods research
- Guevara, Jennifer
UNICEF: self-transformation in a reconfigured global childhood space?
- Guilfoyle, Liam
Learning to teach epistemic practices of science in secondary schools
- Guillotte, Amy E.
Teachers and teaching in low-income countries
- Gulson, Kalervo N.
Spatial theories, methods and education policy
- Guo, Chong-je
Community engagement in STEM education
- Gurr, David
A review of research on middle leaders in schools
- Guthrie, Hugh B.
Pathways, labor markets and tertiary education systems
- Gutzwa, Justin A.
Feminist and critical approaches to gender in qualitative research
- Guy, Breonte S.
Positioning diverse students for science, technology, engineering, and mathematics (STEM) careers: reinforced learning ecosystems for STEM success
- Guzzetti, Barbara J.
The construction of gender in literacy
- ## H
- Ha, Seung Yon
Key developments during adolescence: implications for learning and achievement
- Ha, Jesse
The efficacy of learning strategies from the ICAP perspective
- Hadad, Roxana
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program
- Hagedorn, Linda Serra, PHD
Faculty compensation and salary equity
- Hagger, Hazel
Can initial teacher education partnerships maintain their epistemological and pedagogical underpinnings in light of market-driven reforms? The case of England
- Hahs-Vaughn, Debbie L.
Foundational methods: descriptive statistics: bivariate and multivariate data (correlations, associations)
- Häkkinen, Päivi
Computer supported collaborative learning
- Halai, Anjum
Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania
- Hall, Garret J.
Research study design

- Hallinger, Philip
Applying bibliometric review methods in education: rationale, definitions, analytical techniques, and illustrations
Successful school principalship in Thailand: a literature review
- Hallström, Jonas
Models and modeling in STEM education: nature, roles, and implementation
- Hamilton-McKenna, Caroline
Critical literary perspectives on children's and young adult literature
- Hammersley, Martyn
The history of qualitative research in education
- Hamre, Bjørn F.
Writing the history of Nordic disability and mental disorders
- Han, Feifei
Innovations in university student learning research into blended course designs: incorporating self-report and observational data
- Han, Areum
Assessment of collaborative problem-solving: past achievements and current challenges
- Han, Zhuangzhuang
Bayesian psychometrics
- Hanley, Whitney
Looking for something outside of ourselves: the politics of researching race and dis/Ability and (special) education
- Hanna, Amy
Assessment reform and students' voices
- Hansen, Timothy A.
Descriptive statistics
- Hantke, Sharissa Unger
Contemporary challenges and approaches in anti-racist teacher education
- Happel-Parkins, Alison
Ecocritical pedagogies and curriculum
- Hardy, Ian
Standardized learning: the "problem" of teachers' work and learning under performative policy conditions
- Harnar, Michael A.
Metaevaluation: key concepts explained
- Harper-Hill, Keely
Girls, autism identities, and education
- Harper-Hill, Bella
Girls, autism identities, and education
- Harring, Jeffrey R.
Estimation approaches: least-squares, maximum likelihood and Bayesian
Statistical inference: frequentist approaches
- Harring, Emma K.
Estimation approaches: least-squares, maximum likelihood and Bayesian
- Harris, Roger McL.
Pathways, labor markets and tertiary education systems
- Harris, Alma
Distributed leadership: taking a retrospective and contemporary view of the evidence base
System leadership – emergence and maturity
- Harrison, Christopher W.
School leaders' use of research: viewing research use in decision making through an organizational lens
- Harrison, Colin
Evaluating, supporting and enhancing Internet search activity: a review of research, with implications for pedagogy
- Hartmann, Eva
Public diplomacy and international higher education
- Hascher, Tina
Well-being and learning
- Hatch, Thomas
Leadership and governance for improvement: how schools manage external demands
- Havazelet, Michael
Literacy and the spatial production of school during a pandemic
- Hayden, Mary
International schools: shifting sands and winds of change

- Haynes-Brown, Tashane
Cultural studies in education: insights for the mixed methods researcher
- He, Ming Fang
Cross-cultural narrative inquiry—cross-cultural lives
- He, Qing
Nonlinear regression analysis
- Heck, Ronald H.
Longitudinal designs for school leadership research
- Hedges, Helen
Children's interests and curriculum making in early childhood education
- Hedidar, Walid
Literacy and international development
- Hegarty, Seamus
Gender, poverty and disability—responding to a triple exclusion
- Heimann, Mikael
Attention: a prerequisite for learning
- Helgetun, J.B.
The global and the local: idea flows, contexts, and the influencing of education policy in the 21st century
- Hendawy Al-Mahdy, Yasser F.
School leadership practice in Oman: an investigation of authentic leadership, teacher engagement and turnover intention
- Henderson, Annette M.E.
Learning and development in infancy: toward an integrated approach
- Heng, Tang T.
Diversity beyond its discourse: examining the discourse of diversity in non-Anglo-Saxon contexts
- Henninger, Mirka
Quality control: response style modeling
- Henry, Annette
Intersectionality in the Caribbean and Africa
- Herbey, Ivan I.
Intersecting mixed methods with translational research: implications for educational research and practice
- Hernández-Saca, David
Looking for something outside of ourselves: the politics of researching race and dis/Ability and (special) education
- Herrera-Santillán, Donovan
Children's literature in Latin America today: a historical perspective
- Herrington, Michèle Hinton
Teaching Performance Assessments (TPAs)
- Hervas, Gabriel
Lesson study: a practice for teachers' training and professional development at any educational level
- Heydon, Rachel M.
Curriculum and identity options
- Hidi, Suzanne E.
Curiosity and interest: what should educators know and consider?
- Hiebert, Elfrieda H.
Which words to teach: review and reflection
- Hinchman, Kathleen A.
Youth out-of-school (critical) literacies, texts, and audiences
- Hingangaroa Smith, Graham
Volume 3: Ways of Knowing and Indigenous Education
- Hirose, Kei
Exploratory factor analysis
- Hitchcock, John H.
Mixed methods in survey research
- Ho, Li-Ching
Informal environmental citizenship education
- Hoadley, Ursula
Changing forms of curriculum and governance in developing countries
- Hoare, Carol
Key developments in early and middle adulthood
- Hoff, Erika
Early language development: major milestones and individual differences
- Hoffman, James V.
Research and practice in literacy teacher education: imagining, designing, and engaging in "what could be"

- Hofmann, Carmen
Embodied cognition
- Holm, Gunilla
Participatory photography: high school students' awareness of privileges in Nordic welfare states
- Holmberg, Carrie
Examining conceptual frameworks for understanding teacher learning in the domain of formative assessment
- Holmes, Amanda
Community-centered Indigenous language recovery, restoration, revitalization, and renewal
- Holmes, Kathryn
Secondary school students' STEM career aspirations
- Hong, Bawool
Regionalization of higher education and geopolitics in Northeast Asia
- Hong, Sunggye
Issues and current trends of research in braille for people with visual impairments: questions and implications
- Hong, Quan Nha
Convergence and divergence in mixed methods research
- Honward, Sameer
Learning to balance indigenous and exogenous knowledge systems for environmental decision-making in the Kumaon Himalayas
- Hopfenbeck, Therese N.
Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania
Volume 13: Assessment, Evaluation and Accountability
- Hopkins, David
System leadership – emergence and maturity
- Hörmann, Bernadette
Curriculum, Didaktik, and Bildung
- Hornby, Garry
Recent debates in special & inclusive education
- Horowitz, Rosalind
Borderlands: unique sources of transnational language and literacy for educational research
- Hovardas, Tasos
Simulations and games for learning
- Howard, Patrick
Living schools—reframing education for future possibility
- Howard, Adam
Reading studies of elite schools through an intersectional lens
- Howell, Lisa
Truth and then reconciliation research: an emerging field of educational studies
- Hruby, George G.
A naturalized view of literacy education: what the neuro- and life sciences may offer language and literacy research
- Hu, Weiping
Creativity in STEM education
- Huang, Junzi
Rethinking critical theory/study of education: contributions of the “posts/new materialism”
- Huang, Ronghuai
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis
- Huang, Sijia
Linear regression analysis
Multidimensional item response theory
- Huang, Hsin-Hsiung
Nonlinear regression analysis
- Hubbard, Lea
School reform from a constructivist perspective
- Huckaby, M. Francyne
Education and participatory democracy
- Hughson, E. Anne
Inclusive postsecondary education: vision, practice, higher education reimaged
- Huguet, Alice
School leaders' use of research: viewing research use in decision making through an organizational lens

Hull, Glynda A.

Learning and literacy out of school: learning every day and everywhere

Hulme, Moira

The role and influence of non-state actors on teachers' lives and work

Humburg, Martin

Erasmus(+) student mobility: individual and institutional motivations and effects

Hunt-Anderson, Ingrid E.

Cultural studies in education: insights for the mixed methods researcher

Hunt, Patricia Arias

A constructivist lens of eye-tracking in special populations

Hurem, Aida

Designing equitable STEM education: Guidelines for parents, educators and policy-makers to reduce gender/racial achievement gaps

Hursh, David

Making the public private: the privatization of public education

Hurtado, S.

Feminist and critical approaches to gender in qualitative research

Hutchens, Neal H.

Academic freedom protections in national and international law

Hyatt, David

Critical curriculum analysis framework (CCAF)

I

Ideus, Kimberly L.

Family influence and STEM career aspirations

Ifenthaler, Dirk

Adoption of learning analytics

Ikwuemesim, Lydia

Uses of technology in navigating the pandemic: the case of African higher education: Chika Schoole, Lydia Ikwuemesim, Olaide Agbashe

Imants, Jeroen

Teacher learning as workplace learning

Ingram, Jenni

Researching teacher learning through self-video

Ingrey, Jennifer C.

Affirming trans and queer teachers' lives: rethinking teacher professionalization and pedagogy

Iordanou, Kalypso

Argumentation and critical thinking

Ismail, Salma

Women's informal learning in poor communities, social movements and virtual spaces

Ismail, Nashwa

Distributed leadership: taking a retrospective and contemporary view of the evidence base

Ivankova, Nataliya V.

Frameworks for conceptualizing mixed methods research

Intersecting mixed methods with action and community-based participatory research approaches

Intersecting mixed methods with translational research: implications for educational research and practice

Volume 12: Qualitative, Multimethod, and Mixed Methods Research

Ivey, Gay

Expanding meaningfulness for struggling readers

J

Jackson, David F.

A pedagogical framework for mobile learning in science education

Jacobs, Jennifer

Current trends, tensions and unresolved issues in research on teacher professional learning

Jacobson, Erik

(Mis)understanding adult literacy

Jagger, Susan

Exploring currents and confluences in environmental education

Jahnke, Huia T.

A public good? The commodification of Māori language and culture in higher education

- Jahnke, Huia
Volume 3: Ways of Knowing and Indigenous Education
- Jaipal-Jamani, Kamini
Makerspace and robotics as/for STEM education
- James, Mary E.
Assessing and learning, and learning to learn
- James, Carl E.
Social justice and education: addressing inequities in schooling
- Jang, Bong Gee
Epistemologies and methodologies in literacy research: a case study
- Janks, Hilary
Literacies, language & schooling: teaching and learning the language arts
- Järvelä, Sanna
Computer supported collaborative learning
- Jeffries, Michelle
Critical literacies & social justice: social justice approaches in schools
- Jenkins, K. Neil
Document analysis
- Jennings, Derek
Decolonizing psychology: weaving thriving to disrupt colonial eugenics and shift educational practices within Indigenous psychology
- Jenson, Jennifer
Digital games and literacies
- Jeong, Dong Wook
Organizing schools for improvement: the role of school decision making and authority
- Jewsbury, Paul A.
Educational surveys: methodological foundations
- Jeynes, William H.
Effects of family educational cultures on student success at school: directions for leadership
- Jiang, Jieyu
Redefining educational transfer and borrowing in the pluriverse
- Jiang, Shiyao
Learning analytics
- Jindal-Snape, Divya
Multiple and multi-dimensional educational and life transitions: conceptualization, theorization and XII pillars of transitions
- Jocson, Korina M.
Youth poetry and literacies in remaking the world
- Johansson, Stefan
The role of job demands and resources for teacher job satisfaction: insights from five European countries
- John, Theresa Arevgaq
Developing transforming research—The University of Alaska Fairbanks Center for Cross-Cultural Studies as case study
- Johnson-Jennings, Michelle
Decolonizing psychology: weaving thriving to disrupt colonial eugenics and shift educational practices within Indigenous psychology
- Johnson, Bruce
Early career teachers' work
- Johnson, Latrise P.
Literacy at the intersections of race, gender, and sexuality
- Johnson, Joseph A.
Socioscientific issues and STEM learning
- Johnson, Natasha C.
Advancing education research through mixed methods with existing data
- Johnson, Matthew S.
Bayesian psychometrics
- Johnson, Adeerya
Intersectionality in the Caribbean and Africa
- Johnson, Greer
Changing the context of Australian Indigenous education through leadership for learning “both ways”
- Johnston-Ak, Patricia MG.
Is there a Dr in the house: building an indigenous doctoral program
- Johnstone, Christopher
Curriculum and inclusive education: universal design for learning as a “traveling” phenomenon

- Johnstone, Christopher J.
Quality and inclusion in SDG 4
- Jones, Bronwen M.A.
The neoliberal dispositif: understanding the transformation of the social and ethical space of education
- Jones, Ithel
Service-learning and experiential education
- Jones, M. Gail
Family influence and STEM career aspirations
- Jones, Christopher E.
Critical race theory and educational research utilizing qualitative methods
- Jones, Denisha
From #BlackLivesMatter to Black Lives Matter at school: how the fight for Black Liberation entered the schoolhouse
- Jones, Michelle
Distributed leadership: taking a retrospective and contemporary view of the evidence base
- Jongbloed, Ben
The funding of higher education: international perspectives
- Jorge, Miriam
Literacy, colonialism and resistance in Latin America
- Jotia, Agreement L.
Enhancing a Global South-North partnership to promote global justice in K-12 classrooms: a Botswana case study
- Joubert, Venise
Postcolonialism in the movies and higher education: a case study of the narratives of South African and international students at the University of Pretoria
- Joy Cumming, J.
Social, educational and legal perspectives on assessment and accountability measures in inclusive education for students with disability
- Jude, Nina
Educational assessment and the rise of immigrant students
- jules, tavis d.
Education in small island (and micro) states
South-South Cooperation in Education
- Jung, Jung-Hoon
Shadow curriculum as a new form of curriculum: its evolution, definition, and implications for curriculum studies
- Jung, Yong Ju
Design-based education in STEM: for learners of the 21st century
- K**
- Kahwati, Leila
Mixed methods and qualitative comparative analysis
- Kajigaya, Takeshi
On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT
- Kalantzis, Mary
Contemporary contexts for learning: an overview
Inclusive pedagogies, new learning, new times
- Kamalodeen, Vimala Judy
Cultural studies in education: insights for the mixed methods researcher
- Kamat, Sangeeta
NGOs & international development: ongoing trends and new architectures
- Kamberelis, George
Imagining and reimaging focus group research in education
- Kan, Wei
Discovered and reconstructed: a perspective of Chinese children's literature over the last 100 years
- Kane, Heather
Mixed methods and qualitative comparative analysis
- Kang, Jingoo
Gender and STEM career aspiration
- Kannan, Priya
Score reporting: design and evaluation methods informed by research

- Kannangara, Chathurika
From grit and resilience to academic tenacity
- Kapila, Vikram
Promoting engineering research with entrepreneurship and industry experiences: a teacher professional development program
- Kapā'anaokaola Nākoa Oliveira, Katrina-Ann R.
Ke Ala o ka No'eau: an infusion of Hawaiian into the Hawaiian language
- Karkouti, Ibrahim M.
Deconstructing predominantly White cultures: creating inclusive learning environments in the United States
- Karseth, Berit
Contemporary trends in curriculum research
Volume 7: Approaches to the Curriculum and Its Politics
- Kassen, Jenny E.
Gender identities, sexuality and education
- Kauffman, James M.
Recent debates in special & inclusive education
- Kauko, Jaakko
Global politics of quality in education
- Keehner, Madeleine
Cognition-centered design principles for digital assessment tasks and items
- Kelcey, Benjamin
Foundational methods: power analysis
- Kelchtermans, Geert
Narrative and biography in teacher work lives and development
- Keller, Bryan
Analysis of covariance: univariate and multivariate approaches
- Kempf, Arlo
Afrocentric education: traditions, politics and possibilities
- Kenklies, Karsten
Continental pedagogy & curriculum
- Kennedy, Kerry J.
Asia-Pacific-realist and imaginary constructions
- Kennedy, Aileen
Research for socially progressive teacher education programs
- Kenway, Jane Edith
Reading studies of elite schools through an intersectional lens
- Kersh, Natasha
Lifelong learning beyond initial schooling
- Keune, Anna
Perspectives on learning from the learning sciences
- Khorramdel, Lale
Assessment, evaluation, and accountability: a brief introduction
Educational surveys: conceptual overview
- Kiefer, Markus
Embodied cognition
- Kim-Rich, Erin
Literacies, language, and schooling: exploring writing pedagogy for English language learners
- Kim, Terri
Transnational academic collaboration: a critical analysis of the global and national contexts and constraints of co-authorship
- Kim, Young-Suk Grace
Learning to read and write
- Kim, Young Chun
Shadow curriculum as a new form of curriculum: its evolution, definition, and implications for curriculum studies
- Kim, Young Ae
Integrated STEM curriculum and its assessment
- Kim, Eunjeong
Integrated STEM curriculum and its assessment
- Kim, Soo Hyeon
Design-based education in STEM: for learners of the 21st century
- Kim, Daniel
Intersecting mixed methods with translational research: implications for educational research and practice
- Kim, Stella Y.
Generalizability theory and applications
- Kim, Jee-Seon
Causal models

- King, James R.
College reading
- King, Bruce M.
Analysis of variance: univariate and multivariate approaches
- Kirsch, Irwin
Educational surveys: conceptual overview
- Kitchen, Julian
Improving teacher education through the self-study of practice
- Klar, Hans W.
Successful leadership in high needs schools
- Klassen, Robert M.
Teacher recruitment: a competency-based approach
- Klinger, Don A.
Building teachers' capacity in assessment *for* learning in rural East Africa
- Kluijtmans, Manon
Teacher expertise and how it develops during teachers' professional lives
- Knaus, Jake
Case methods in teacher education
- Knight, Jane
Knowledge diplomacy: the role of international higher education, research and innovation in international relations.
- Knobloch, Neil A.
K-12 teachers' beliefs and practices in STEM integration
- Koellner, Karen
Current trends, tensions and unresolved issues in research on teacher professional learning
- Koh, Sin Yee
Global competition over skills
- Kohli, Rita
Teacher education for social change: shifting from race-evasiveness to racial literacy
- König, Johannes
Teacher education and the development of teacher noticing
- Koopman, Maaïke
Learning environments in vocational education: how to elicit learning for and in practice
- Koro, Mirka
Qualitative sampling and qualitative data
- Koskey, Michael
Developing transforming research—The University of Alaska Fairbanks Center for Cross-Cultural Studies as case study
- Kosnik, Clare
A lifetime of experience: an international study on the lives of teacher educators:
- Kovačević, Jasna
Applying bibliometric review methods in education: rationale, definitions, analytical techniques, and illustrations
- Kragh, Jesper V.
Writing the history of Nordic disability and mental disorders
- Krajčik, Joseph
Aligning PBL with the NGSS: What do the NGSS offer PBL?
- Krause-Alzaidi, Lara-Stephanie
Culture and literacy: situated social literacies
- Krieger, Florian
Assessment of collaborative problem-solving: past achievements and current challenges
Interdisciplinary frontiers: computer-based process data analysis in educational measurement
- Krone, Beth
Literature: critical imagination and young adult readers remaking worlds
- Kubota, Ryuko
On pedagogical applications of World Englishes: stumbling blocks, stepping stones, and usefulness of boundaries in ELT
- Kuby, Candace R.
Post-philosophies and the doing of inquiry in literacies studies
- Kuckartz, Udo
Using software for mixed methods analysis
- Kulophas, Dhirapat
Successful school principalship in Thailand: a literature review

- Kumagai, Yuri
Critical literacies in East Asian languages: examples from Japanese in foreign language education
- Kumpunen, Stephanie
Rapid qualitative research
- Kurth-Schai, Ruthanne
The light in the distance: global democracy and humanity's hope for the future
- Kurtz, Brianna A.
Enhancing a Global South-North partnership to promote global justice in K-12 classrooms: a Botswana case study
- Kvamme, Ole Andreas
Curriculum and the United Nations' sustainable development goals
- Kwak, Lydia
The role of job demands and resources for teacher job satisfaction: insights from five European countries
- Kyllonen, Patrick C.
Assessment of interpersonal and intrapersonal skills
- Kyriakides, Leonidas
International trends on teacher evaluation
- L**
- Ladwig, James G.
Within school stratification
- Lake, Vickie E.
Service-learning and experiential education
- Lake, Lucy
Gender, poverty and disability—responding to a triple exclusion
- Lamb, Richard
Virtual reality and science, technology, engineering, and mathematics education
- Lamichhane, Kamal
Economics of inclusive education
- Lane, Suzanne
Validity
- Lara, Gilberto P.
Literacy and identity across home, school, church, and social life
- Largo, Marissa
Arts-Based Educational Research
- Law, Nancy
Curriculum and digital citizenship
Teacher skills and knowledge for technology integration
- Lazonder, Ard W.
Inquiry-based learning
- Leander, Kevin
Literacy and the spatial production of school during a pandemic
- Leask, Betty R.
Reimagining internationalization of the curriculum
- Leat, David
Innovative pedagogies in relation to curriculum
- Lee, Moosung
International patterns and trajectories in principal preparation
- Lee, Jack T.
Academic mobility: global trends and future prospects
- Lee, Crystal Chen
Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms
- Lee, Hyun Ji
Self-efficacy
- Lee, Carol D.
Cross cultural lenses on learning
- Lee, Wing On
Curriculum and digital citizenship
- Lee, Lamoine Samuel
Cultural studies in education: insights for the mixed methods researcher
- Lee, Won-Chan
Generalizability theory and applications
- Lee, Mina
Achievement growth modeling
- Lee, Tae Kyoung
An introduction to growth mixture models (GMM)

- Lee, Tiffany S.
Indigenous language revitalization in the United States and Canada
- Leech, Nancy L.
Mixed methods sampling and data collection
- Lehman, Blair
Conversation-based assessment: current findings and future work
- Lehti, Hannu
Parental unemployment and children's educational outcomes — a Literature review
- Lei, Pui-Wa
A tutorial on structural equation modeling
- Leija, María G.
Funds of knowledge and funds of identity: lenses for building language and literacy
- Leijen, Äli
Supporting and assessing professional agency of student teachers in initial teacher education
- Leithwood, Kenneth
Integrated educational leadership
Pedagogical strategies for leadership development
Volume 4: Administration, Governance, and Leadership
- Leland, Andrew
Queer in school: the experiences of youth and educators in the United States
- Lemke, Mariann
A brief introduction to contemporary standard setting methods
- Lenoir, Yves
Interdisciplinarity in basic education: interpretations, relevance and principles
- Lenters, Kimberly
Affect theory and textual variations
- Leonardo, Zeus
White woman: or, the abused abuser's role in U.S. educational stratification
- Leung, Promail K.Y.
Pre-service STEM teacher education
- Leventhal, Brian C.
Simulation studies for psychometrics
- Lewis, Steven
Globalization and education
Spatial theories, methods and education policy
- Li, Qiong
Thriving on challenges: understanding teacher resilience in hard-to-staff Chinese rural schools
- Li, Haiying
Intelligent tutoring systems: a history and an example of an ITS for science
Intelligent tutoring systems and conversational agents
- Li, Xinyi
Creativity in STEM education
- Li, T.
Using flexible thinking to assess student sensemaking of phenomena in project-based learning
- Li, Siqi
Spatial thinking and STEM education
- Liang, Leming
Teacher skills and knowledge for technology integration
- Liasidou, Anastasia
Cross-cultural perspectives on the rise of special education policy discourses.
- Liddiard, Kirsty
Disability studies in a posthuman age
- Liddicoat, Anthony J.
Language policy and planning as a discursive accomplishment
- Liebenberg, Linda
Thinking qualitatively in the context of mixed methods research: implications for study design
- Lillejord, Sølvi
A more intelligent accountability—future directions
Educating the teaching profession
- Lim, Sirene
Conceptualizing professional development with and for a diverse and globally mobile early childhood education workforce

- Lim, Levan
Inclusive education in the Republic of Singapore:
a situated perspective
- Lim, Sliki N.
Formative assessment: an overview of history,
theory and application
- Lim, Leonel
Diversity beyond its discourse: examining the
discourse of diversity in non-Anglo-Saxon
contexts
- Lindner, Marlit Annalena
Cognition-centered design principles for digital
assessment tasks and items
- Liou, Yi-Hwa
Networked systems leadership: the potential of
social network theory and analysis
- Lipe, Daniel
Seeing through the smoke: creating space for
Indigenous Knowledge in science education
- Lipnevich, Anastasiya A.
Assessment, evaluation, and accountability: a brief
introduction
Emotions in education
- Little, Judith Warren
Teachers' work in the context of COVID-19
- Littlefair, David
Document analysis
- Liu, Peng
Understanding turnaround leadership research:
continuity and change (2017–2021)
- Liu, Min
The context in teacher education and
professionalism: hierarchies, networks and
markets in four countries
- Liu, Katrina
Teacher education for diverse learners
- Liu, Zhichun
Machine learning
- Liu, Qian
Dialogic education: tensions and dilemmas
- Liu, Jing
Transformation of university missions—from
nation-building to a focus on common global
challenges reflections from East Asia, under the
risk of being “lost in translation”
- Liu, Xiufeng
Volume 11: The Rise of STEM Education
- Liu, Xiang
Bayesian psychometrics
- Liu, Yang
Statistical inference: Bayesian approaches
- Livne, Oren E.
Advanced machine learning
- Loder-Jackson, Tondra L.
Critical race theory and educational research
utilizing qualitative methods
- Loh, Wen Wei
Causal models
- Lohr, Sharon L.
Sampling designs for educational research
- Long, Yun
Dialogic education: tensions and dilemmas
- Longlands, Helen
Education, equality and international development
- Loogma, Krista
Conducting large-scale mixed methods studies in
education
- Looking Horse, Makasa
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
- Lopes, Joao
Recent debates in special & inclusive education
- Loreman, Tim
Principles of inclusive practice in schools: what is
COVID-19 teaching us?
- Lott, Christine
Intelligent tutoring systems: a history and an
example of an ITS for science
- Louis, Karen Seashore
The contributions of positive organizational studies
to educational leadership and school
improvement
- Lövdén, Martin
Education and neurocognitive aging - is there
a relation?

Low, Ee-Ling
Rethinking teacher education in pandemic times:
lessons from and for Singapore

Loxley, Andrew
Following the money? Commodification,
segregation and shopping

Loyens, Sofie M.M.
Student-centered instruction: inquiry-, problem-,
project-, and case-based learning

Löytönen, Teija
Qualitative sampling and qualitative data

Lubienski, Chris
Spatial theories, methods and education
policy

Luecht, Richard M.
Certification and licensure testing

Luft, Julie A.
In-service teacher education for STEM

Lumb, Matt
Tensions in access and accountability

Lumby, Jacky
Distributed leadership: critical perspectives

Luo, Tian
Student motivation in STEM: factors related to and
measurement of STEM motivation

Lupinacci, John
Ecocritical pedagogies and curriculum

Luschei, Thomas F.
Organizing schools for improvement: the role of
school decision making and authority

Lyons-Thomas, Juliette
Test translation and adaptation

Lyons, Suzanne
Economic impact of higher education on students,
corporations and society

M

Ma, Josef Kuo-Hsun
Digital learning divide

Ma, Jasmine Y.
What is learning, for whom, and to what end? An
overview

Ma, Shufeng
Collaborative Reasoning and children's
development of language and thought

Ma, Ye
Item Response Theory

Ma, Mingjia
Item Response Theory

Maaka, Margaret
Volume 3: Ways of Knowing and Indigenous
Education

Macalalag, Augusto
Socioscientific issues and STEM learning

MacArthur, Jude
Using disability studies in education to move from
individualized assessment to partners in pedagogy

Macedo, Elizabeth
Transnational curriculum inquiry

Machado, Armando
Learning from a behaviorist perspective

Macintyre Latta, Margaret
Curriculum's multiplicities as participatory modes
of interaction

Macías, Luis Fernando
Undocumented activism: a bate-papo~testimonio

Madden-Smith, Abby
Indigenous student identity development
and school success: mitigating social identity
and stereotype threats in
schools

Magalhães, António
Governance, institutional autonomy and leadership
in Portuguese higher education

Magnussen, Svein
Memory

Magnússon, Gunnlaugur
From Salamanca to Sweden: inclusive education as
policy in transit

Mahuika, Nēpia
"The mis-education" of Aotearoa: the struggle for
Indigenous history in New Zealand

Majewski, Haylee M.
Self-efficacy in STEM

- Majhanovich, Suzanne
Issues: neocolonialism and global languages
- Makoni, Sinfree
Regimes of literacy as regimes of truth about Africa:
language ideologies and southern voices
- Malatesta, Jaime L.
Generalizability theory and applications
- Maldonado-Maldonado, Alma
Knowledge/power and education:
ethnonationalism, globalism, and social justice
- Mali, Amar
Education for all and MDGs: global education
policy translation in India
- Mallette, Marla H.
Are we there yet? Onto-epistemological qualitative
language and literacy research
- Mangin, Melinda M.
Queer in school: the experiences of youth and
educators in the United States
- Manski, Sarah E.
Descriptive statistics
Statistics: Value-Added Models
- Marchev, Dobrin
Analysis of covariance: univariate and multivariate
approaches
- Margolis, Melissa J.
Past, present, and future of educational
measurement
- Mariani, Rachael
Reviews of education for students with disability in
Australia
- Marion, Scott F.
Validity for through-year assessments
- Markauskaite, Lina
Teaching expertise: an activity system's lens
- Markovits, Henry
Reasoning and decision making
- Marsh, Herbert W.
Academic self-concept: a key predictor of
achievement and learning
- Martin-Hill, Dawn
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
- Martinez-Yaden, Camille
Reintroducing “development” into theories of the
acquisition and growth of early literacy:
developmental science approaches and the
cultural-historical perspective of L. S. Vygotsky
- Martino, Wayne J.
Gender identities, sexuality and education
- Mason, Jillian
Are we there yet? Onto-epistemological qualitative
language and literacy research
- Massaro, Dom
Artificial intelligence in literacy
- Massonnié, J.
Perspectives on learning from neuroscience
- Masutha, Mukovhe
Non-completion amongst black working-class
students in South African universities: the dangers
of a single story
- Mathou, Cécile
Contemporary curriculum policies: globalization,
recontextualization and hybridization of
pedagogic models
- Maunakea, Summer Puanani
‘Āina-based education: Huli Ka Lima I Lalo
Transformative praxis
- Mavrou, Katerina
Cross-cultural perspectives on the rise of special
education policy discourses.
- Maxwell, Claire
Global mobility and middle class
families—parenting and education
- Maxwell, Joseph Alex
Designing integrative mixed methods research
- Maxwell, Graham S.
Assessment and accountability global trends
and future directions in 21st century
competencies
- Maydeu-Olivares, Alberto
Model-data fit evaluation: item fit and model
selection
- Mayer, Diane
Volume 5: Teachers’ Lives, Work, and Professional
Education

- Mayer, John D.
Intelligences
- Mayer, Richard E.
Problem solving
- Mayo, Peter
Labor markets, SDG4 and vocational education and training
- Mayoh, Joanne
Intersecting mixed methods with phenomenology
- Mayring, Philipp A.E.
Qualitative content analysis
- McCaffrey, Daniel F.
Volume 14: Quantitative Research and Educational Measurement
- McCarthy, Kathryn S.
Knowledge: a fundamental asset
- McConn, Matthew
Are we there yet? Onto-epistemological qualitative language and literacy research
- McCoy, Whitney N.
Mixed methods research proposal
- McCrudden, Matthew T.
Learning from multiple texts
Visual displays in mixed methods research: integration of methods
- McDermott, Mark A.
The multimodal affordances of language in STEM education
- McDevitt, Seung Eun
Educating a world on the move: rethinking teacher preparation in the context of mass global migration and increasingly diverse classrooms
- McDonald, Sarah
The first-in-family experience: gendered and classed subjectivities in elite higher education
- McGinn, Noel
Six decades of development: the impact of structural adjustment programs on education in developing countries
- McGrane, Joshua A.
Comparative judgment
- McGrath, Simon
Development and education
- McIlroy, Anne-Marie
Using disability studies in education to move from individualized assessment to partners in pedagogy
- McKenzie, Marcia
The geopolitics of climate change and education policy
- McKenzie, James
Indigenous language revitalization in the United States and Canada
- McLachlan, Claire Jane
Curriculum and democracy: building understandings of democracy through early childhood practice
- McLaughlin, Juliana Mohok
Shifting priorities of foreign aid to Pacific Island Countries: some considerations of new imperialism and geopolitics
- McLean Davies, Larissa
Teaching Performance Assessments (TPAs)
- McLewis, Channel C.
Feminist and critical approaches to gender in qualitative research
- McMillan, James H.
Classroom assessment
- McNair, Jonda C.
Toward culturally relevant literacies with children and families of color
- McNamara, Danielle S.
Knowledge: a fundamental asset
- McNaughton, Stuart
Optimizing learning: an overview
- McPherson, Amy
Within school stratification
- McQueen, Kahontiyoha Cynthia Denise
Kayaní:yo (a good path)—Warrior Science
Ohneganos—Indigenous ecological knowledge
Tehahenteh
- McVee, Mary B.
Posthumanism and Positioning Theory: explorations of embodiment and affect from a critical perspective
Using multiliteracies and multimodality to explore engineering literacies and problem-scoping for

- multilingual learners: A perspective on
multiliteracies for STEM learning
- Medina, José Luis
Lesson study: a practice for teachers' training and
professional development at any educational level
- Medina, Carmen Liliana
Biliteracy, playfulness and performative practices
- Mehan, Hugh
School reform from a constructivist perspective
- Meijer, Cor J.W.
Paying for inclusion and exclusion
- Meiser, Thorsten
Quality control: response style modeling
- Menna, Lydia
A lifetime of experience: an international study on
the lives of teacher educators:
- Menon, Kirti
"Reschooling" higher education: the dynamics of
curriculum in South Africa
Implications of the Fourth Industrial Revolution in
the vortex of higher education in South Africa
- Menter, Ian
Comparative teacher education research in
turbulent times
- Merrill, Barbara
Experiencing class and gender: mature women
students and their journeys from HE to the labor
market
- Mertens, Donna M.
Mixed methods and evaluation
Transformative mixed methods research
- Messiou, Kyriaki
The role of students' voices in promoting inclusive
education
- Meyer, Helen
Photovoice with children and youth
- Migliarini, Valentina
Legislation and policy in Europe
- Mihas, Paul
Qualitative research methods: approaches to
qualitative data analysis
- Mihr, Anja
Global human rights
- Mikser, Rain
Conducting large-scale mixed methods studies in
education
- Milana, Marcella
European Union
- Milford, Todd M.
The multimodal affordances of language in STEM
education
- Miller, Richard
Teacher education for diverse learners
- Miller, Emily Adah
Aligning PBL with the NGSS: What do the NGSS
offer PBL?
- Mills, Martin
Retention of teachers from minority ethnic groups
in disadvantaged schools
- Mills, Kathy A.
Sensory literacies: the full sensorium in literacy
learning
- Milne, Catherine
Design-based pedagogy: a heuristic for developing
math and science teachers' professional
practice
- Minces, Victor Hugo
STEAM for all: a vision for STEM and arts
integration
- Mincu, Monica
The context in teacher education and
professionalism: hierarchies, networks and
markets in four countries
- Mishra Tarc, Aparna
Decolonial and education
- Misiaszek, Greg William
Ecopedagogy and ecopedagogical literacy
- Mislevy, Robert J.
Evidence-centered design
- Mitra, Ayan
A naturalized view of literacy education: what the
neuro- and life sciences may offer language and
literacy research
- Mockler, Nicole
Teachers' work amid global education reform: "the
greatest challenge"?

- Mohamed Emam, Mahmoud
School leadership practice in Oman: an investigation of authentic leadership, teacher engagement and turnover intention
- Molla, Tebeje
Equity in early childhood education: policy framing, translation and enactment
- Monroe, Scott
Achievement growth modeling
- Montelongo, Ricardo
Online education
- Moorosi, Pontso
Understanding leadership identity construction: a gendered analysis
- Morales, Maria C.
Assessment of young emergent bilingual children
- Morel, Richard Paquin
Framing educational leadership as a *multilevel distributed* practice: a systemwide perspective
- Morgan, Bob
Toward a rights-based and culturally responsive model of Aboriginal Education
- Morrell, Ernest
Critical literacies and social justice
- Morrell, Jodene
Critical literacies and social justice
- Morris, Rebecca
Tackling teacher shortages: doing the right thing at the right time and with the right evidence
- Morris, Paul
Educational accountability, international large scale assessments and PISA: powerful influences based on weak foundations
- Morton, Missy
Using disability studies in education to move from individualized assessment to partners in pedagogy
- Moschetti, Mauro C.
The work of international organizations: ten theoretical approaches and a way forward
- Moser, Alecia
Learning and development in infancy: toward an integrated approach
- Moses, Tim
Score equating: an aspirational form of score linking
- Motala, Shireen
Exploring the professional identity of university academics: “becoming” a university teacher
Implications of the Fourth Industrial Revolution in the vortex of higher education in South Africa
- Motawy, Yasmine S.
Educational concerns in Arabic picture books
- Mowat, Joan G.
Working collaboratively with the school community to build inclusion for all
- Muench, Richard
Academic capitalism and market thinking in higher education
- Muijs, Daniel
Leadership in large-scale international assessments—diminishing returns?
- Mukherjee, Mousumi
Education for all and MDGs: global education policy translation in India
- Mulholland, Kirstin
Document analysis
- Muller, Johan
Changing forms of curriculum and governance in developing countries
- Murillo, Luz A.
Rethinking multiliteracies: toward a radical vision for multilingualism
- Murphy, Joseph F.
The contributions of positive organizational studies to educational leadership and school improvement
- Mustafa, Muhammad Yasir
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis
- Muñoz, Joaquin
(Just) languaging with *duende*: (re)inventing and innovating communicative practices as resistance, healing, and thrivance in Indigenous contexts

- Mutton, Trevor
Can initial teacher education partnerships maintain their epistemological and pedagogical underpinnings in light of market-driven reforms? The case of England
- Myers, Aaron J.
Model-data fit evaluation: posterior checks and Bayesian model selection
- Myran, Steve
Re-conceptualizing instructional leadership through a science of learning lens
- N**
- Naepi, Sereana
Capturing storms: Pacific women's approaches to change in New Zealand universities
- Naidoo, Jaqueline
Teachers living the curriculum: culturally relevant pedagogy and action research
- Nakhid, Camille
Affirming methodologies—the Caribbean oral tradition of Ole Talk
- Nastasi, Bonnie Kaul
Mixed methods in culturally tailored interventions
- Nava, Imelda
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program
- Näykki, Piia
Computer supported collaborative learning
- Neeganagwedgin, Erica
First Nations & Indigenous identities in education
Indigeneity
- Neepin, Ryan
Weaving together mind, body and spirit: envisioning pathways to trauma-informed classrooms from Indigenous perspectives
- Nelson, Ryan L.
A constructivist lens of eye-tracking in special populations
- Newman, Isadore
Foundations of mixed methods research
- Newton, Paul E.
Taking purpose into account when designing, using and evaluating educational assessments
- Ng-A-Fook, Nicholas
Truth and then reconciliation research: an emerging field of educational studies
- Nguyen, Diep Thi Bich
Global processes: global trade in international education and the international education-migration nexus
- Nguyen, Xuan Thuy
Decolonizing research with girls with disabilities in the Global South
- Niad, Hayley
Curriculum and inclusive education: universal design for learning as a “traveling” phenomenon
- Nicol, Cynthia
Imagining STEM *with* and *as* place
- Nicolay, Björn
Interdisciplinary frontiers: computer-based process data analysis in educational measurement
- Nieveen, Nienke M.
Curriculum design
- Niglas, Katrin
Conducting large-scale mixed methods studies in education
- Niipare, Alina K.
Early career academics in Namibia: experiences, prospects and challenges
- Nilsen, Trude
Accountability for the future? International testing and future directions
- Nind, Melanie AE.
Disability and employment: finding a way to work
- Nolan, Andrea
Equity in early childhood education: policy framing, translation and enactment
- Nordgren, Kenneth
History curriculum in the Anthropocene; how should we tell the story?
- Nordin, Andreas
Curriculum theory and new institutionalism

- Norman-Tichawangana, Verity
NGOs & international development: ongoing trends and new architectures
- North, Amy
Education, equality and international development
- Norton, Bonny
Translanguaging and identity across global sites
- No'eau Warner, Sam L.
Ke Ala o ka No'eau: an infusion of Hawaiian into the Hawaiian language
- Nteropoulou-Nterou, Evdoxia
Crisis, austerity and the inclusive education agenda: the case of Greece
- Nuttall, Joce
Conceptualizing professional development with and for a diverse and globally mobile early childhood education workforce
- Nyberg, Lars
Education and neurocognitive aging - is there a relation?
- O**
- O'Keeffe, Lisa
Inquiry-based teaching and STEM education
- Oanda, Ibrahim Ogachi
The hollowing out of academic freedom from African universities; from external capture to internal implosion
- Oberheidt, Stephanie
Erasmus(+) student mobility: individual and institutional motivations and effects
- Okerson, Justine R.
The intricacies of class size in higher education
- Oliver, J. Steve
Integrated STEM curriculum and its assessment
- Oliveri, Maria Elena
Fairness
- Olsen, Allison Wynhoff
Literacies, language & schooling: teaching and learning the language arts
- Omenugha, Kate
The politics of literacy languaging and texts: the politics of literacy in Nigeria.
- Omercajic, Kenan
Gender identities, sexuality and education
- Onghena, Patrick
Mixed methods research synthesis for literature reviews
- Onukaogu, Chukwuemeka Eze
The politics of literacy languaging and texts: the politics of literacy in Nigeria.
- Opini, Bathseba
Intersectionality in the Caribbean and Africa
- Oplatka, Izhar
The emergence of social justice leaders: identity formation of new principals
- Orland-Barak, Lily
Mentoring "in the time of cholera": perspectives, promises and challenges in the preparation of mentors for diversity in the new digital teacher learning space
- Ornelas, Carlos
Shifting policies of education in Latin América
- Ortiz Aragón, Alfredo
Responding to cultural diversity through action and community-based participatory approaches in qualitative research
- Osborne, Jonathan
Learning in the sciences
- Óskarsdóttir, Edda
Enacting inclusive pedagogy in teacher education: creating a learning space for teachers to develop their professional identity
- O'Brien, David G.
Digital literacies, multimodality and the engagement of adolescents who are disengaged or disenfranchized in school
- O'Leary, Matt
Linking research and teaching to benefit student learning in higher education: complexity, collaboration and a tailored approach
- O'Neal, Catherine Walker
An introduction to growth mixture models (GMM)
- O'Sullivan, John
Leadership formation: in the smithy of the personal and professional

- Ott, Corinna D.
Teacher education for social change: shifting from race-evasiveness to racial literacy
- Õun, Tiia
Conducting large-scale mixed methods studies in education
- Outhred, Rachel
Delivering on the promises of tertiary education in low- and middle-income countries: risks, sustainability and inclusion
- Ozga, Jenny
Accountability in an age of international testing
- P**
- Pacico, Juliana Cerentini
Item Response Theory
- Paige, David D.
The art and science of reading fluency and fluency instruction
- Panaou, Petros
Children's literature in Greece and Cyprus
- Panchamukhi, Pushkarni
Globalization and political economy of education development in South Asia
- Pantaleo, Sylvia
The evolution of children's literature: metafictional devices and digital mediation
- Papadamou, Tryfonas
International trends on teacher evaluation
- Pareja Roblin, Natalie N.
Curriculum and 21st century skills
- Park, Seijoon
School leadership and organizational factors: evaluating pathways to student learning
- Park, Vicki
Evidence-informed decision making in schools and districts: the role of equity
- Parker, Stephen
The problem with higher education student completions in the global north
- Parr, Judy M.
Volume 6: Learning, Cognition, and Human Development
- Parsons, Seth A.
Epistemologies and methodologies in literacy research: a case study
- Pasque, Penny A.
Knowledge production in Europe: actors, policies and critiques
- Patfield, Sally
Questioning the consensus on effective professional development
- Patterson, Ashley N.
Epistemologies and methodologies in literacy research: a case study
- Patton, Michael Quinn
Qualitative evaluation
- Pavlov, Goran
Model-data fit evaluation: item fit and model selection
- Pearson, P. David
Reading comprehension in troubled times
- Pekrun, Reinhard
Emotions in education
- Penn, Helen
False and distracting promises: early childhood education and care in low income countries
- Penna, Olivia
Researching curriculum and disability in higher education
- Pensoneau-Conway, Sandra L.
Autoethnography: storied scholarship
- Percle, Amanda
A constructivist lens of eye-tracking in special populations
- Perez, Leticia
CT for equity: situating computational thinking, computer science, and social justice in a math and science teacher education program
- Pérez, Alexandra
Validity
- Periasamy, Ravadhi
Women school leadership in Malaysia: perspectives and enactments

- Perry, Kekailoa
The salient truthfulness of Aloha 'Āina Theory and
Guerrilla Praxis
- Petrovai, Kinga
The relationship between knowledge exchange and
teachers' use of mobile technologies in the
mathematics classroom
- Pfadt, Julius M.
Reliability
- Phelan, Anne M.
Defiant imaginaries: from a politics of identity to an
ethics of friendship in teaching
- Philippou, Stavroula
Curriculum making and teacher agency
- Phillips, Louise G.
Children's participation in local
curriculum-making
- Pidgeon, Michelle
Being in good relations: higher education's
responsibilities to Indigenous resurgence
- Pillay, Venitha
Women in higher education: exclusion, inclusion
and seclusion
- Pillow, Wanda S.
Feminist genealogy & policy futures
- Plano Clark, Vicki L.
Combining mixed methods and case study research
Frameworks for conceptualizing mixed methods
research
- Pluye, Pierre
Convergence and divergence in mixed methods
research
- Poe, Mya
Fairness
- Poed, Shiralee
Researching curriculum and disability in higher
education
- Pollock, Katina
Principals' work in public schools
- Pope, Raechele
Minoritized individuals and knowledge-economy
- Popkewitz, Thomas S.
Rethinking critical theory/study of education:
contributions of the "posts/new materialism"
- Porou, Ngāti
"The mis-education" of Aotearoa: the struggle for
Indigenous history in New Zealand
- Porter, Rebecca L.
Disability and poverty
- Poth, Cheryl N.
Mixed methods research design practices to address
complexity in education
Promising practices for addressing unique ethical
dilemmas arising from integration in mixed
methods research
- Powers, Sonya
A brief introduction to contemporary standard
setting methods
- Pradhan, Sahara
NGOs & international development: ongoing trends
and new architectures
- Pratesi, Alessandro
Triangulation: uses, abuses and recent
developments
- Presiado, Vivian E.
Rethinking multiliteracies: toward a radical vision
for multilingualism
- Price, Larry R.
Confirmatory factor analysis: foundations and
extensions
- Priestley, Mark
Curriculum making and teacher agency
- Prinsloo, Mastin
Culture and literacy: situated social literacies
- Print, Murray
Political orientation
- Proctor, Chris
Computational thinking
- Puhan, Gautam
Subscores; what they are, why they are in high
demand, and how to evaluate their quality
- Purzer, Senay
Engineering education

Q

- Qadach, Mowafaq
Principal's information-processing mechanisms (PIPMs) as an effective tool for learning-centered leadership that reduces teacher's intent to leave
- Qazi, Mohammed
Self-efficacy in STEM
- Qian, Haiyan
Leadership in non-western countries: China
- Qiu, Fanxiao Wani
Curiosity and interest: what should educators know and consider?
- Qu, Pengtong
Biliteracy, playfulness and performative practices
- Quaynor, Phoebe L.
The scholarship of African Children's literature: issues of representation of critical voices in the conversation
- Que, Hua
Work and learning: power, politics and approaches
- Quist, Isaac
Challenges facing education in Sub-Saharan Africa
- Quiñones, Sandra
Leadership and family factors: critical perspectives on international leadership standards, models, and practices

R

- Rädiker, Stefan
Using software for mixed methods analysis
- Rahimi, Seyedahmad
Machine learning
- Rao, Kavita
Universal Design for Learning in its 3rd decade: a focus on equity, inclusion, and design
- Rapanta, Chrysi
Argumentation and critical thinking
- Rapley, Tim
Document analysis

- Rasinski, Timothy
The art and science of reading fluency and fluency instruction
- Rauschenberger, Emilee
The role and influence of non-state actors on teachers' lives and work
- Rawajfih, Yasmeeen
Self-efficacy in STEM
- Rawi, Roszalina
Assessment for learning
- Read, Robyn
Global Education Monitoring Report
- Reilly, David
Designing equitable STEM education: Guidelines for parents, educators and policy-makers to reduce gender/racial achievement gaps
Gender differences in educational achievement and learning outcomes
- Reimann, Peter
Teaching expertise: an activity system's lens
- Renkl, Alexander
Exemplars
- Renninger, K. Ann
Curiosity and interest: what should educators know and consider?
- Restoule, Jean-Paul
Can we transform the academy? Indigenous-centered programming at the Department of Indigenous Education at University of Victoria
- Revers, Scott
Experiencing class and gender: mature women students and their journeys from HE to the labor market
- Reyes, Coleen
Leading in urban contexts: how principal supervisors coach principals to lead improvement
- Reynolds, Martyn
Journeying to the common good: exploring Indigenous Solomon Islands leadership through five moments

- Rezai-Rashti, Goli M.
Global testing regimes: accountability,
standardization, racial inequality and the rise of
audit culture
- Rhee, Jeong-eun
Volume 2: Diversity, Democracy, and Social Justice
in Education
- Rice, Bethany M.
Changing knowledge & cultures in teacher
education: autoethnography for critical reflection
- Riddle, Stewart
Democracy and education
- Riedinger, Kelly
Informal STEM learning
- Rieser, Richard
Developing inclusive education for disabled
children and students: advocacy in hard times
- Rikers, Remy M.J. P.
Student-centered instruction: inquiry-, problem-,
project-, and case-based learning
- Rindskopf, D.
Hierarchical linear models
- Ring-Whalen, Elizabeth A.
Measurement instruments of STEM affective
learning: a systematic review
STEM curriculum development and
implementation
- Rivas, Axel
Curriculum governance: the dilemma of the
centralized curriculum
- Rix, Jonathan
Re-examining special needs—what could be better?
- Rizvi, Fazal
International higher education and the
commodification of student mobility
Volume 1: Globalization and the Shifting
Geopolitics of Education
- Roach, Whitney Neumeyer
Education and participatory democracy
- Roberts, Thomas
Promises and perils of STEM education:
synthesizing teacher, student, & research
perceptions
- Roberts, Audrey Conway
Promises and perils of STEM education:
synthesizing teacher, student, & research
perceptions
- Robertson, Susan L.
The World Trade Organization
- Robinson, Maureen
University–school partnerships in initial teacher
education
- Robinson, Chazz
Minoritized individuals and knowledge-economy
- Robinson, Viviane
What is distinctive about educational leadership?
- Roehrig, Gillian H.
Measurement instruments of STEM affective
learning: a systematic review
STEM curriculum development and
implementation
- Rogers, Rebecca
Literacy, colonialism and resistance in Latin
America
- Rogers, Theresa
Critical literary perspectives on children's and young
adult literature
Volume 10: Literacies and Languages Education
- Rönnerman, Karin
Middle leading as a practice for enabling
professional learning
- Rosé, Carolyn P.
Learning analytics
- Rose, David H.
Universal Design for Learning in its 3rd decade:
a focus on equity, inclusion, and design
- Rose, Anthea
Diverse landscapes of school governing body
practices
- Rosén Rasmussen, Lisa
Designing school(s): school architecture and
curriculum
- Rosopa, Patrick J.
Analysis of variance: univariate and multivariate
approaches

- Roussel, Linda
Intersecting mixed methods with translational research: implications for educational research and practice
- Rowe, Emma
Spatial theories, methods and education policy
- Rowe, Deborah Wells
Writing before age three
- Rowluck, William
Community-centered Indigenous language recovery, restoration, revitalization, and renewal
- Rubio-Jimenez, Ana
Dialogic education: tensions and dilemmas
- Rubio-Torres, Carlos
Children's literature in Latin America today: a historical perspective
- Runco, Mark A.
Creativity and learning
- Runswick-Cole, Katherine
Disability studies in a posthuman age
- Rüschepöhler, Lilith
Student motivation in STEM: factors related to and measurement of STEM motivation
- Rushby, Jade V.
Teacher recruitment: a competency-based approach
- Russell, Tom
Self-study of teacher education practices: the complex challenges of learning from experience
- Rutherford, Gill
"First you have to see me as a human being": disabled students' experiences of schooling
- Rutkowski, Leslie
Measuring what we cannot see: assessing extreme populations in international large-scale assessments
- Rutkowski, David
Measuring what we cannot see: assessing extreme populations in international large-scale assessments
- Ryen, Anne
Ethics and politics in qualitative research
- Ryoo, Jean J.
Critical theory in qualitative research in education
- ## S
- Saarinén, Taina
Internationalization in English? Problematizing the role of language in higher education internationalization
- Saggers, Beth
Autism, literacies, and social communication: a framework of inclusive new literacies
Girls, autism identities, and education
- Sahlberg, Pasi
Coaching for professional learning of early career teachers
Global education reform movement revisited
- Saldaña, Christopher M.
Educational accountability: the use and abuse of assessment data
- Säljö, Roger
Learning from a sociocultural perspective
- Salman, Casey
Reimagining research on writing: a timely rhizomatic argument
- Salmerón, Ladislao
Learning from multiple texts
- Salmons, Janet
Mixed methods and web-based technologies for data collection
- Saltman, Kenneth J.
Global education philanthropy
Neoliberal education: from capitalizing on disaster to critical pedagogy
- Salzman, Hal
STEM education and workforce development: the history, politics, and evidence
- Sambolín Morales, Astrid N.
Mobilizing prefigurative hybrid literacies for decolonial resistance: three transnational examples
- Sanchez, Belén
Curriculum governance: the dilemma of the centralized curriculum

- Sánchez, Sonia N.
Literacy and identity across home, school, church,
and social life
- Sanchez, Cecilia
Assessment of English learner students
- Sandoval, William A.
Epistemic cognition
- Sanga, Kabini
Journeying to the common good: exploring
Indigenous Solomon Islands leadership through
five moments
- Sant, Edda
Populism and education
- Santana, Anabel Fernández
Affirming methodologies—the Caribbean oral
tradition of Ole Talk
- Sao Pedro, Michael A.
Intelligent tutoring systems: a history and an
example of an ITS for science
- Sarroub, Loukia K.
Religious influences on the growth of literacy
practice
- Sarungi, Veronica
Teachers' perspectives and practice of assessment for
learning in classrooms in Tanzania
- Sato, Mistilina
Case methods in teacher education
- Sayed, Yusuf
UNESCO: promising education multilateralism?
- Schellings, Gonny
Beginning teachers' professional identity work and
its support
- Schellman, Madeline
Diagnostic classification modeling
- Scherer, Ronny
Accountability for the future? International testing
and future directions
- Scherman, Vanessa
Thinking qualitatively in the context of mixed
methods research: implications for study design
- Schoenfeld, Alan H.
Assessment and mathematical literacy: a brief
introduction
- Learning in math(s)
- Schönborn, Konrad J.
Models and modeling in STEM education: nature,
roles, and implementation
- Schonert-Reichl, Kimberly A.
Social and emotional learning
- Schoonenboom, Judith
Mixed methods and multimethod research in
education: six design decisions for collecting,
combining, and developing datasets
Ten mixed methods integration strategies for
obtaining a detailed understanding
- Schostak, J.
Naturalistic observation
- Schroeder, Cassandra
Religious influences on the growth of literacy
practice
- Schucker, Kelly A.
Posthumanism and Positioning Theory:
explorations of embodiment and affect from
a critical perspective
Using multiliteracies and multimodality to explore
engineering literacies and problem-scoping for
multilingual learners: A perspective on
multiliteracies for STEM learning
- Schuelka, Matthew J.
Reform and purpose for inclusive values in complex
education systems
- Schultz, Katherine
Learning and literacy out of school: learning every
day and everywhere
- Schunk, Dale H.
Learning from a social cognitive theory perspective
- Schwab Cartas, Joshua
Community-centered Indigenous language
recovery, restoration, revitalization, and
renewal
- Seago, Nanette
Current trends, tensions and unresolved issues in
research on teacher professional learning
- Sebastian, James
School leadership and organizational factors:
evaluating pathways to student
learning

- See, Beng Huat
Tackling teacher shortages: doing the right thing at the right time and with the right evidence
- Seggie, Fatma Nevra
Higher education strategy formations in shifting landscape of higher education
- Sehoole, Chika
Uses of technology in navigating the pandemic: the case of African higher education: Chika Sehoole, Lydia Ikwuemesim, Olaide Agbashe
Volume 8: Higher Education Management, Policy, and Practice
- Sellar, Sam
Datafication
- Semenov, Andrei
Executive function and education
- Seow, Tricia
Informal environmental citizenship education
- Serafini, Frank
The evolution of children's literature: metafictional devices and digital mediation
- Severance, Samuel
Aligning PBL with the NGSS: What do the NGSS offer PBL?
- Severo, Cristine
Regimes of literacy as regimes of truth about Africa: language ideologies and southern voices
- Shagrir, Leah
Teacher educators' professional growth as researchers
- Shahjahan, Riyad A.
Global university rankings and geopolitics of knowledge
- Shan, Hongxia
Work and learning: power, politics and approaches
- Shank Lauwo, Monica
Translanguaging and identity across global sites
- Shannon-Baker, Peggy
Integrating art and artistic practices in mixed methods research in education
Philosophical underpinnings of mixed methods research in education
- Share, Jeff
Critical media literacy and second language acquisition
- Sharma, Gunjan
UNESCO: promising education multilateralism?
- Sharma, Umesh
Principles of inclusive practice in schools: what is COVID-19 teaching us?
Utilizing a global social justice lens to explore indicators of inclusive education
- Shavelson, Richard J.
Advantages and challenges of performance assessment of student learning in higher education
- Shea, Peter
Online learning
- Shen, Ji
Interdisciplinary science learning
Technology-enhanced collaborative learning in STEM
- Shepard, Lorrie A.
Educational accountability: the use and abuse of assessment data
- Shi, Shengpeng
Dialogic education: tensions and dilemmas
- Shi, Dexin
Model-data fit evaluation: item fit and model selection
- Shields, Robin
Academic mobility: global trends and future prospects
- Shimauchi, Sae
Transformation of university missions—from nation-building to a focus on common global challenges reflections from East Asia, under the risk of being "lost in translation"
- Shin, Jinnie
Automatic item generation
- Short, Kathy G.
Global literature for young people: critical issues and social responsibilities

- Showunmi, Victoria
Understanding leadership identity construction:
a gendered analysis
- Shute, Valerie J.
Machine learning
- Sichombe, Beatrice S.
Early career academics in Namibia: experiences,
prospects and challenges
- Siddiq, Fazilat
Educational assessment of 21st century
skills—novel initiatives, yet a lack of systemic
transformation
- Sijtsma, Klaas
Nonparametric statistical methods
Reliability
- Silova, Iveta
Redefining educational transfer and borrowing in
the pluriverse
- Silvestri, Katarina N.
Using multiliteracies and multimodality to explore
engineering literacies and problem-scoping for
multilingual learners: A perspective on
multiliteracies for STEM learning
- Simpson, Amber
Connecting play to STEM concepts, practices and
processes: review of research on play within STEM
learning environments
- Singal, Nidhi
Teachers with disabilities in mainstream schools:
embodying and enacting inclusive education
- Sinharay, Sandip
Continuous probability distributions
Discrete probability distributions
Modeling item response times
Statistical methods for detection of test fraud on
educational assessments
Subscores; what they are, why they are in high
demand, and how to evaluate their quality
- Sintos Coloma, Roland
Volume 2: Diversity, Democracy, and Social Justice
in Education
- Sipuka, Precious Bupe
Exploring the professional identity of university
academics: “becoming” a university teacher
- Sirota, Sandra L.
Human rights education
- Sirotkin, Josephine
Disability and poverty
- Sivesind, Kirsten
The rise and relevance of curriculum reform
- Skukauskaitė, Audra
Ethnography: foundations, challenges, and spaces
of possibilities
- Slee, Roger
Rethinking schooling for inclusion
Volume 9: Inclusive Education and Disability
Studies in Education
- Sleegers, Peter
Leadership and school capacity for continuous
improvement: toward a systemic approach
- Smit, Brigitte
Thinking qualitatively in the context of mixed
methods research: implications for study design
- Smith, Erica
Thirty years of competency-based training: how
Australia painted itself into a curriculum corner in
vocational education and training
- Smith, Hannah
Girls, autism identities, and education
- Smith, William C.
Quality and inclusion in SDG 4
- Smith, Katie J.
Gender, poverty and disability—responding to
a triple exclusion
- Smith, Blaine E.
Digital literacies & multimodal learning: writing in
digital formats
Responding to and creating multimodal texts
- Smith, Patriann
Toward culturally relevant literacies with children
and families of color
- Smith, Lyric
Integrating art and artistic practices in mixed
methods research in education
- Smith, Jeffery K.
Assessment, evaluation, and accountability: a brief
introduction

- Smith, Hinekura Lisa
Re-weaving thinking, re-thinking weaving:
theorizing Indigenous arts practice as research
methodology
- Snoddon, Kristin
Inclusive education and deaf learners
- So, Wing-mui Winnie
Community engagement in STEM education
- Soini, Tiina
Curriculum making and teacher agency
- Sommerhoff, Daniel
Perspectives on learning from the learning sciences
- Sosnowski, Jim
Rethinking multiliteracies: toward a radical vision
for multilingualism
- Souto-Otero, Manuel
Erasmus(+) student mobility: individual and
institutional motivations and effects
- Sparks, Jesse R.
Conversation-based assessment: current findings
and future work
- Spillane, James P.
Framing educational leadership as a *multilevel
distributed* practice: a systemwide perspective
School leaders' use of research: viewing research use
in decision making through an organizational
lens
- Spina, Nerida
Critical literacies & social justice: social justice
approaches in schools
Primary and secondary education: literacy and
social justice
- Spybrook, Jessaca
Foundational methods: power analysis
- Srinivasa Rao, S.
For profit corporations and education reform:
global dimensions and impacts
- St. Denis, Verna
Contemporary challenges and approaches in anti-
racist teacher education
- Stacey, Meghan
Teachers' work amid global education reform: "the
greatest challenge"?
- Staddon, John
Learning from a behaviorist perspective
- Stahl, Norman A.
College reading
- Stahl, Garth
The first-in-family experience: gendered and classed
subjectivities in elite higher education
- Stake, Robert
Case study research
- Starnawski, Marcin
World-systems analysis and curriculum theory
World-systems analysis
- Staudt Willet, K. Bret
Social media and learning
- Staufert-Reyes, Ericka
Mobilizing prefigurative hybrid literacies for
decolonial resistance: three transnational
examples
- Staus, Nancy
Informal STEM learning
- Stein, Sharon
Power/knowledge: global and Indigenous
education
- Stempfer, Lisa
Emotions in education
- Sterenber Mahon, Alycia J.
Reporting mixed methods research studies
- Stillman, Jamy
Equity- and justice-minded teacher educators
- Stobart, Gordon
Alternative assessment
- Stockinger, Kristina
Emotions in education
- Stone, Jason E.
Credit for experience, credit recognition and
transfer
- Stone, Bessie G.
Autism, literacies, and social communication:
a framework of inclusive new
literacies
- Storksdieck, Martin
Informal STEM learning

- Stornaiuolo, Amy
Cosmopolitan literacies and the limits of understanding: possible directions in global literacy research
- Stosich, Elizabeth Leisy
Leadership and governance for improvement: how schools manage external demands
- Strid, Karin
Attention: a prerequisite for learning
- Stringer, Andrea
Coaching for professional learning of early career teachers
- Stringer, Ernest
Responding to cultural diversity through action and community-based participatory approaches in qualitative research
- Stylianides, Gabriel J.
The relationship between knowledge exchange and teachers' use of mobile technologies in the mathematics classroom
- Styres, Sandra
Weaving together mind, body and spirit: envisioning pathways to trauma-informed classrooms from Indigenous perspectives
- Subban, Pearl
Utilizing a global social justice lens to explore indicators of inclusive education
- Suciu, Lavinia
Qualitative research approaches and designs: discourse analysis
- Sude
Chinese Minzu (民族) diversity and education
- Sugrue, Ciaran
Leadership formation: in the smithy of the personal and professional
- Sülau, Veronica
Middle leading as a practice for enabling professional learning
- Sullivan, Anna M.
Early career teachers' work
- Sumara, Dennis
Learning to count what counts: narrative hermeneutics for teachers and teaching
- Sun, Jingping
Research on successful school leadership
Transformational school leadership
- Sun, Xiaohong
Thriving on challenges: understanding teacher resilience in hard-to-staff Chinese rural schools
- Sun, Jingjing
Collaborative Reasoning and children's development of language and thought
- Sundberg, Daniel
Curriculum and knowledge content
- Suspitsyna, Tatiana
Knowledge production in Europe: actors, policies and critiques
- Švigelj, Melissa Marini
Youth as educators
- Swaffield, Sue
Assessment for learning
- Sweet, Tracy M.
Modeling of network structures
- Sweet, Tracy
Statistical inference: Bayesian approaches
- Sweller, John
Cognitive load theory
- T**
- Taber, Keith S.
Curriculum and science
- Tadesse, Azeb
Issues of production and utilization of knowledge: multiple uses of knowledge
- Taghipoorreynah, Mehdi
Mixed methods and the Delphi method
- Tai, Joanna
Feedback for learning
- Tajima, Chihiro
Mixed methods research designs
- Takasugi, Catherine
Learning to count what counts: narrative hermeneutics for teachers and teaching
- Takayama, Keita
Southern theory as an educational project: Shinto, self-negation and comparative education

- Takeda, Yuya
On pedagogical applications of World Englishes:
stumbling blocks, stepping stones, and usefulness
of boundaries in ELT
- Tan, Sean
Principled assessment design
- Tao, Liqing
Chinese literacy acquisition: an overview
- Tarc, Paul
International Baccalaureate: meanings, uses and
tensions in a globalizing world
K-12 global/international education:
dancing with “diversity, democracy, and social
justice”
- Tashakkori, Abbas
Foundations of mixed methods research
- Tatto, Maria Teresa
Comparative teacher education research in
turbulent times
- Tefera, Adai A.
Learning disabilities' unsettling intersections:
disrupting learning, cultural, socioeconomic, and
environmental deficits
- Teichert, Laura
Family literacies: differing perspectives
- Teichler, Ulrich
Internationalization of higher
education
- Templeton, Shane
Spelling: theory, assessment, and
pedagogy
- Tendeiro, Jorge N.
Model-data fit evaluation: aberrant response
detection
- Tenison, Caitlin
The current trends and opportunities for machine
learning in learning analytics
- Tereshchenko, Antonina
Retention of teachers from minority ethnic groups
in disadvantaged schools
- Ter Vrugte, Judith
Simulations and games for learning
- Thaver, Thana
Inclusive education in the Republic of Singapore:
a situated perspective
- Thessin, Rebecca A.
Leading in urban contexts: how principal
supervisors coach principals to lead improvement
- Thom, Jennifer S.
Imagining STEM *with* and *as* place
- Thomas, Sally M.
Evaluating schools as professional learning
communities
Review of evaluation approaches for school
principals
- Thomas, M.S.C.
Perspectives on learning from neuroscience
- Thomas, Gary
Following the money? Commodification,
segregation and shopping
- Thompson, Jeff
International schools: shifting sands and winds of
change
- Thompson, Alisun
Teachers' work in the context of COVID-19
- Thompson, Kathryn N.
Simulation studies for psychometrics
- Thompson, Christopher G.
Meta-analysis
- Thorpe, Matthew
Datafication
- Tibbitts, Felisa L.
Human rights education
- Tierney, Robert J.
Global meaning-making: critical analysis and
advocacy for a diverse world
- Tilak, Jandhyala B.G.
Globalization and political economy of education
development in South Asia
- Tilton, Zachary D.
Metaevaluation: key concepts explained
- Timoštsuk, Inge
Conducting large-scale mixed methods studies in
education

- Tippett, Christine D.
The multimodal affordances of language in STEM education
- Titchkosky, Tanya
"Disability and society: changing our frames of reference"
- Titus, Shawn
Socioscientific issues and STEM learning
- Tlili, Ahmed
Design of Online-Merge-Offline (OMO) learning environments in the post-COVID-19 era: case study analysis
- Tolchinsky, Liliana
Children's earliest insights in developing literacy
- Topping, Keith J.
Peer- and self-assessment
- Toro, Sandra Annette
Museum education: building the capacity of society and its members
- Toropova, Anna
The role of job demands and resources for teacher job satisfaction: insights from five European countries
- Towns-Bain (Alutiig), Aleesha
Impact of the 1971 Alaska Native Claims Settlement Act on education: 1971–2021
- Tran, Ly Thi
Global processes: global trade in international education and the international education-migration nexus
- Trendafilov, Nickolay
Exploratory factor analysis
PCA and other dimensionality-reduction techniques
- Trigiani, Amanda
Media education: history, frameworks, debates and challenges
- Trigos-Carrillo, Lina
Literacy, colonialism and resistance in Latin America
- Tripet, Kristen
Situating teacher learning in the mathematics classroom and everyday practice
- Tripp, Jennifer N.
STEM schools: their rise, present state, and future potential
Using multiliteracies and multimodality to explore engineering literacies and problem-scoping for multilingual learners: A perspective on multiliteracies for STEM learning
- Tröhler, Daniel
Curriculum theory and education history
- Truesdale, Lindsay M.
Assessment of young emergent bilingual children
- Tschannen-Moran, Megan
Leadership practices and why they matter: trust in leaders
- Tucker-Smith, T. Nicole
Universal Design for Learning in its 3rd decade: a focus on equity, inclusion, and design
- Tulloch, Michelle K.
Early language development: major milestones and individual differences
- Turner, Jennifer
Learning to count what counts: narrative hermeneutics for teachers and teaching
- Turner, Fergal
Delivering on the promises of tertiary education in low- and middle-income countries: risks, sustainability and inclusion
- Turner, Rita
Ecocritical pedagogies and curriculum
- Tzirides, Anastasia O.
Contemporary contexts for learning: an overview
Inclusive pedagogies, new learning, new times
- U**
- Ulmer, Gregory L.
Electracy
- Ümarik, Meril
Conducting large-scale mixed methods studies in education
- Unterhalter, Elaine
What works or what matters? Contested ideas of girls' education, gender equity and global connection

Urban, Mathias

UNICEF: self-transformation in a reconfigured
global childhood space?

Uugwanga, Jafet S.

Early career academics in Namibia: experiences,
prospects and challenges

V

Vanassche, Eline

Teacher education policy and professionalism:
a personal review of teacher education policy
research

van Aswegen, Rachel

International perspectives on STEM teacher
induction

van den Akker, Jan J.H.

Curriculum design

van der Kleij, Fabienne

Implementation of Assessment for Learning

van der Schaaf, Marieke

Teacher expertise and how it develops during
teachers' professional lives

van Dijk, Esther E.

Teacher expertise and how it develops during
teachers' professional lives

van Dijk, Marijn

Dynamic system approaches to language
acquisition

van Geert, Paul

Dynamic system approaches to language
acquisition

Van Hove, Geert

Legislation and policy in
Europe

van Rijn, Peter W.

Modeling item response times

van Ryneveld, Linda

Toward a framework for the implementation of
fully online distance education at a residential
higher education institution

van Tartwijk, Jan

Teacher expertise and how it develops during
teachers' professional lives

van Veen, Klaas

Teacher learning as workplace learning

Vattøy, Kim-Daniel

Feedback and classroom practice

Veck, Wayne

Refugee student identities and educational
exclusion

Veiga, Amélia

Governance, institutional autonomy and leadership
in Portuguese higher education

Verger, Antoni

The privatization of education. Drivers,
social effects and regulatory challenges
of private sector participation in
schooling

Vermunt, Jeroen K.

Latent class analysis

Vieira, Daniella Santos

Undocumented activism: a bate-papo~testimonio

Vincent, Lily

Defiant imaginaries: from a politics of identity to an
ethics of friendship in teaching

Vindrola-Padros, Cecilia

Rapid qualitative research

Visse, Merel

Case study research

Vogl, Susanne

Strategies to integrative mixed methods
analysis

von Davier, Matthias

Educational surveys: conceptual overview

Voogt, Joke M.

Curriculum and 21st century skills
Curriculum design

W

Wachira, Nicholas

Building teachers' capacity in assessment *for* learning
in rural East Africa

Wade, Carmen

The disruption of Black history and critical
pedagogy in international higher education
curriculum

- Wagner, Daniel A.
Language choice, technology and international education
Literacy and international development
Teachers and teaching in low-income countries
- Wahlström, Ninni
Contemporary trends in curriculum research
Volume 7: Approaches to the Curriculum and Its Politics
- Wainer, Howard
Displaying information: from woolly mammoths to the Great Migration
- Walford, Geoffrey
Low-fee private schools
- Walker, Allan
International patterns and trajectories in principal preparation
Leadership in non-western countries: China
- Walker, Amy
Biliteracy, playfulness and performative practices
- Walker, Cindy M.
Differential functioning
- Walton, Elizabeth
Disability studies in education and intersectionality
- Walton, Janet B.
Combining mixed methods and case study research
- Wang, Yinying
Future research on educational leadership: theoretical groundings of educational leadership research
- Wang, Changzhao
Interdisciplinary science learning
Technology-enhanced collaborative learning in STEM
- Wang, Jian
Student motivation in STEM: factors related to and measurement of STEM motivation
- Wang, Hui-Hui
K-12 teachers' beliefs and practices in STEM integration
- Wang, Lei
Volume 11: The Rise of STEM Education
- Wang, Yibo
Item Response Theory
- Warner, Chantelle N.
Literacy/multiliteracies and foreign language education
- Warren, Te Rina
A public good? The commodification of Māori language and culture in higher education
- Watkins, Amanda
Paying for inclusion and exclusion
- Watkins, Daphne C.
Advancing education research through mixed methods with existing data
- Wawrzynski, Matthew R.
Assessing student learning outcomes
- Weaver-Hightower, Marcus B.
Developing qualitative research proposals
- Weaver, Leemoy
Cultural studies in education: insights for the mixed methods researcher
- Weaver, Steve R.
Cultural studies in education: insights for the mixed methods researcher
- Webb, Noreen M.
Cooperative and collaborative learning
- Webber, Melinda
Indigenous student identity development and school success: mitigating social identity and stereotype threats in schools
- Weber, Everard
Postcolonialism in the movies and higher education: a case study of the narratives of South African and international students at the University of Pretoria
Volume 8: Higher Education Management, Policy, and Practice
- Weems, Lisa D.
Young women and activism and dissident global citizenship
- Wegerif, Rupert
Dialogic education: tensions and dilemmas

- Weiland, Steven
Open access and the unfinished transformation of scholarly communications
- Weinstein, José
School leadership in Latin America: research and policy trends
- Welch, Anthony
Education in South-East Asia. Much achieved, much still to do
The China challenge; China's challenges
- Welker, Alyson
Imagining and reimagining focus group research in education
- Weltsek, Gustave J.
Biliteracy, playfulness and performative practices
- Wendt Samu, Tanya
New frontiers in Pacific education, leadership and decolonized practice within Oceania: an initial exploration
- Weninger, Csilla
Curriculum and literacy
- Wentzel, Kathryn R.
Students' social relationships at school
- Wheelahan, Leesa
Realism and theorizing curriculum—the role of knowledge
- Whelan, Alison
Innovative pedagogies in relation to curriculum
- Whitburn, Ben
The National Disability Insurance Scheme, education and funding to support the inclusion of students with disabilities across educational sectors
- White, Bruce
Inquiry-based teaching and STEM education
- White, Howard
Mixed methods in education RCTs
- Wickrama, Kandauda A.S.
An introduction to growth mixture models (GMM)
- Wieczorek, Oliver
Academic capitalism and market thinking in higher education
- Wieselmann, Jeanna R.
Measurement instruments of STEM affective learning: a systematic review
STEM curriculum development and implementation
- Wiggins, Charletta N.
Critical race theory and educational research utilizing qualitative methods
- Wijnia, Lisette
Student-centered instruction: inquiry-, problem-, project-, and case-based learning
- Wilde, Jodie
Girls, autism identities, and education
- Wiley, Jennifer
Fundamental processes of learning: an overview
- Wilkinson, Ian A. G.
Volume 6: Learning, Cognition, and Human Development
- Williams (Tlingit), Maria Shaa Tláa
Impact of the 1971 Alaska Native Claims Settlement Act on education: 1971–2021
- Williams, Rachel E.
Establishing quality in qualitative research: trustworthiness, validity, and a lack of consensus
- Willis, Arlette Ingram
Literacy access, the law, and racism in the United States: a critique
- Wilson-Lopez, Amy
Shifting from disciplinary versus content area literacy to a transdisciplinary approach
- Wilson, Mark
Explanatory measurement
Principled assessment design
- Wind, Stefanie A.
Nonparametric methods for modeling item responses: Mokken Scale Analysis
- Windchief, Sweeney
Teaching graduate courses; indigenous students, course co-construction and bicultural accountability
- Wingate, Lori A.
Metaevaluation: key concepts explained

Wingo, Nancy P.

Intersecting mixed methods with action and community-based participatory research approaches

Winne, Philip H.

Assessing learning processes

Winter, Christine

Critical curriculum analysis framework (CCAF)

Winterbottom, Christian

Service-learning and experiential education

Winters, Niall

The relationship between knowledge exchange and teachers' use of mobile technologies in the mathematics classroom

Wong, Lok-Sze

School leaders' use of research: viewing research use in decision making through an organizational lens

Wong, K. Laiana

Ke Ala o ka No'eau: an infusion of Hawaiian into the Hawaiian language

Wood, Elizabeth Ann

Key developments in early childhood

Woolley, Mary C.

Teachers' lived experience: voices of history teachers on the English National Curriculum

Wozolek, Boni

Queer theory and qualitative research

Wright, Handel Kashope

Culture

Wylie, E. Caroline

U.S. federal assessment policy reforms in the twenty first century and their impacts

Wyman, Leisy

(Just) languaging with *duende*: (re)inventing and innovating communicative practices as resistance, healing, and thriving in Indigenous contexts

X

Xiao, Yao

Culture

Work and learning: power, politics and approaches

Xu, Guanlan

Item Response Theory

Xu, Jiajun

Diagnostic classification modeling

Y

Yaden, David B., Jr.

Reintroducing "development" into theories of the acquisition and growth of early literacy: developmental science approaches and the cultural-historical perspective of L. S. Vygotsky

Yaden, David

Volume 10: Literacies and Languages
Education

Yamamoto, Kentaro

Educational surveys: conceptual overview

Yan, Rui

An overview of literature on factors that influence principal turnover

Yang, Rui

China's engagement with Southeast Asia in higher education

Yang, Na

Frameworks for conceptualizing mixed methods research

Yang, Ting

Mixed methods in education RCTs

Ydesen, Christian

Educational assessment – trends in research, policy and practice

Ye, Tianpeng

Item Response Theory

Yemini, Miri

Global mobility and middle class families—parenting and education

- Yenika-Agbaw, Vivian S.
The scholarship of African Children's literature:
issues of representation of critical voices in the
conversation
- Yildirim, Kasim
The art and science of reading fluency and fluency
instruction
- Yin, Chengbin
Evidence-centered design
- Yip, Valerie W.Y.
Pre-service STEM teacher education
- Ylimaki, Rose M.
Leadership for education and learning in changing
contexts
- Yokota, Junko
Global literature for young people: critical issues
and social responsibilities
- Yonezawa, Akiyoshi
Transformation of university missions—from
nation-building to a focus on common global
challenges reflections from East Asia, under the
risk of being "lost in translation"
- Yoon, Ee-Seul
Spatial theories, methods and education policy
- You, Xiaoye
Transliteracy for a cosmopolitan future
- Young, Michelle D.
The development of educational leaders: history
and contemporary practice in the
United States
- Young, Chase
The art and science of reading fluency and fluency
instruction
- Youngs, Peter
International perspectives on STEM teacher
induction
- Yuan, Mei
Chinese Minzu (民族) diversity and education
"Indigenous" ways of knowing in Chinese Minzu
education: teachers reflecting on their engagement
with STEM in inner Mongolia
- Yuan, Ke-Hai
Missing data analysis
- Yun, Eunju
Conceptualizing professional development with
and for a diverse and globally mobile early
childhood education workforce
- ## Z
- Zacharia, Zacharias C.
Simulations and games for learning
- Zaleski, Michelle
Transliteracy for a cosmopolitan future
- Zammit, Katina
The construction of gender in literacy
- Zancajo, Adrián
The World Bank and education in a changing global
arena
The privatization of education. Drivers,
social effects and regulatory challenges
of private sector participation in
schooling
- Zapata-Rivera, Diego
Conversation-based assessment: current findings
and future work
- Zeegers, Yvonne
Inquiry-based teaching and STEM
education
- Zelazo, Philip David
Executive function and education
- Zembylas, Michalinos
Emotion and reconciliation pedagogies
- Zhai, Xiaoming
A pedagogical framework for mobile learning in
science education
- Zhang, Liang
Economic impact of higher education on students,
corporations and society
- Zhang, Jihong
Item Response Theory
- Zhang, Mo
Automated scoring of constructed-response items in
educational assessment
- Zhang, Qi
Foundational methods: power
analysis

- Zhao, Weili
Re-turning to the cultural, historical, and
philosophical foundations of international
curriculum
- Zhao, Hui
A tutorial on structural equation modeling
- Zheng, Mingying
Item Response Theory
- Zhou, Jingtian
Thriving on challenges: understanding teacher
resilience in hard-to-staff Chinese rural schools
- Zhu, Kejin
China's engagement with Southeast Asia in higher
education
- Zhuang, Rongxia
Design of Online-Merge-Offline (OMO) learning
environments in the post-COVID-19 era: case
study analysis
- Zitter, Ilya
Learning environments in vocational education:
how to elicit learning for and in practice
- Zlatkin-Troitschanskaia, Olga
Advantages and challenges of performance
assessment of student learning in higher
education
- Zwick, Rebecca
Admissions testing in higher education